博士論文（要約）

論文題目　Studies on isolation, structure elucidation, and biosynthesis of secondary metabolites from the thermophilic bacterium *Thermosporothrix hazakensis*

(好熱性細菌 *Thermosporothrix hazakensis* 由来二次代謝産物の単離、構造解析、生合成に関する研究)

氏　　名　朴　眞 洙

平成 25 年度　博士論文

Studies on isolation, structure elucidation, and biosynthesis of secondary metabolites from the thermophilic bacterium *Thermosporothrix hazakensis*

(好熱性細菌 *Thermosporothrix hazakensis* 由来二次代謝産物の単離、構造解析、生合成に関する研究)

東京大学

大学院農学生命科学研究科

応用生命工学専攻

朴　眞 洙

指導教員　　西山　真

ABBREVIATIONS

λ	Wavelength [nm]
A domain	Adenylation domain
ACP	Acyl carrier protein
AMP (ATP)	Adenosine mono (tri) phosphate
AT	Acyl transferase
BLAST	Basic local alignment search tool
BuOH	Butanol
C domain	Condensation domain
calcd	Calculated
CoA	Coenzyme A
COSY	Correlation spectroscopy
Cy domain	Cyclization domain
Da	Dalton
DH	Dehydratase
DNA	Deoxyribonucleic acid
EDTA	N,N,N´,N´-Ethylendiaminetetraacetate
ER	Enoylreductase
ESI	Electrospray ionization
EtOAc	Ethyl acetate
GFC	Gel-filtration chromatography
HMBC	Heteronuclear multiple bond correlation
HPLC	High performance liquid chromatography
HR	High resolution
HSQC	Heteronuclear single quantum correlation
Hz	Hertz
IPTG	Isopropyl-β-D-thiogalactoside
kb	Kilobase
kDa	Kilodalton
KR	Ketoreductase
KS	Ketosynthase
MeOH	Methanol
MS	Mass spectrometry
mult	Multiplicity
MW	Molecular weight
m/z	Mass-to-charge ratio

NMR	Nuclear magnetic resonance
NOESY	Nuclear overhauser effect spectroscopy
NRPS	Non-ribosomal peptide synthetase
obsd	Observed
OD	Optical density
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PCP	Peptidyl carrier protein
PKS	Polyketide synthase
ppm	Part per million
PTFE	Polytetrafluoroethylene
Rt	Retention time
SDS	Sodium dodecylsulfate
SLPM	Standardized liter per minute
TFA	Trifluoroacetic acid
TPP	Thiamine diphosphate

Table of Contents

ABBREVIATIONS i

INTRODUCTION 1

A brief history of microbial natural products 1

New source of novel microbial and chemical diversity 3

Extremophiles 4

Thermophiles 5

Thermosporothrix hazakensis SK20-1^T 6

1. Bioinformatics analysis of the *T. hazakensis* SK20-1^T genome for sequence-based discovery of secondary metabolites 9

1.1. Introduction 9

1.2. Structure-determining factors in NRPS and PKS 12

1.3. Methods 14

1.4. Results and Discussion 16

1.5. Summary 33

2. Isolation, structure elucidation, and biological activity of the secondary metabolites from *T. hazakensis* SK20-1^T 34

2.1. Chemical screening using small-scale culture conditions 34

2.2. Fermentation, isolation and purification for new molecules from *T. hazakensis* SK20-1^T ... 40

2.3. Structure elucidation of the isolated compounds 45

2.4. Biological activity of the isolated compounds 50

2.5. Summary 51

3. Biosynthesis of new compounds isolated from *T. hazakensis* SK20-1^T 52

3.1. Biosynthesis of (-)-sattabacin (**1**) and (-)-hazakacin (**2**) 52

3.2. インターネット公表出来ないために削除 74

3.3. インターネット公表出来ないために削除 74

3.4. インターネット公表出来ないために削除 74

CONCLUSION 75

EXPERIMENTAL 76

APPENDIX 85

REFERENCE 93

ACKNOWLEDGEMENTS 100

INTRODUCTION

A brief history of microbial natural products

Natural products or secondary metabolites derived from microorganisms have been the source for a significant proportion of the new small-molecule chemical entities developed as drugs (**Figure 1**). Since the discovery of the fungal antibiotic penicillin by Alexander Fleming in 1928 and the period during the 1940s and 1950s, i.e., the so-called ‘Golden Age of Antibiotics’, microbial secondary metabolites with unprecedented chemical diversity and potent biological activity have facilitated immense pharmaceutical discovery and drug development (**Figure 2**). However, over the past three decades, the discovery of novel natural products that exhibit beneficial pharmaceutical properties has been dramatically diminishing, and studies on traditional drug discovery have declined because of several challenges such as a low isolation ratio of novel skeletons and the repetitive isolation of known compounds. Consequently, new drug discovery strategies appeared from the 1990s such as combinatorial chemistry, high-throughput screening, and computer-assisted design of small-molecule ligands.

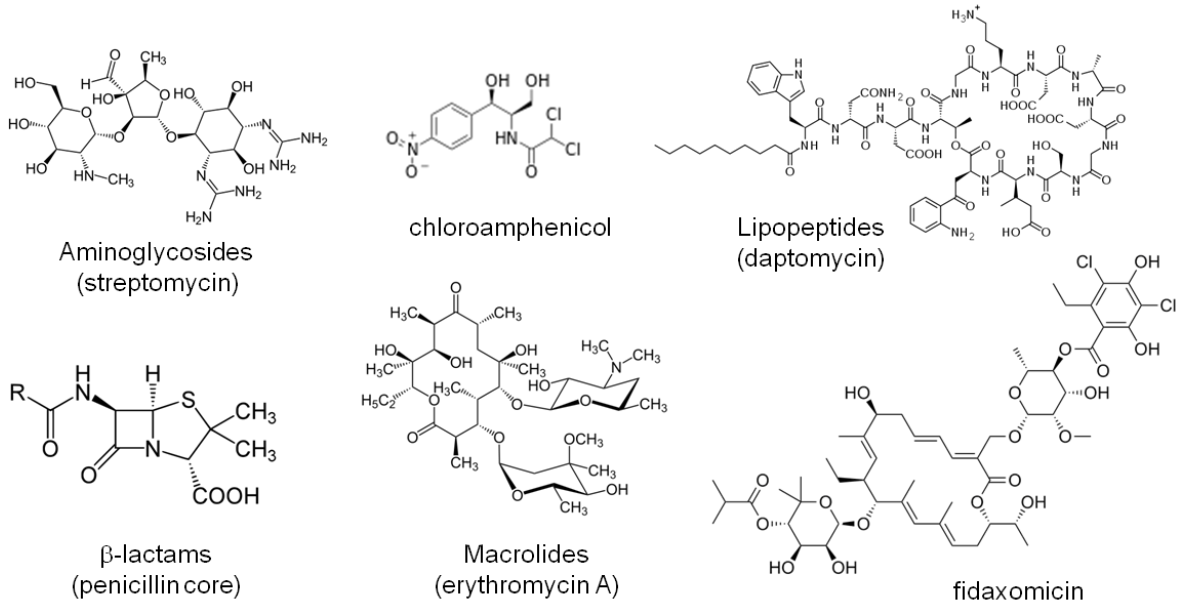


Figure 1 Representative medicinal natural products from microorganisms.

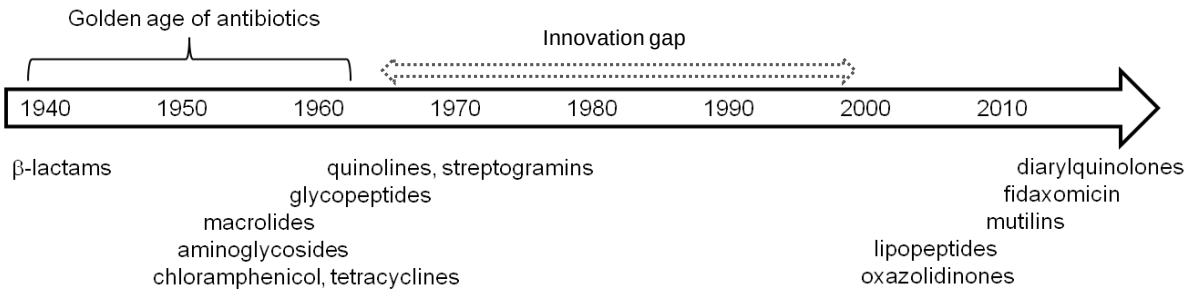


Figure 2 Timeline representing a brief history of drug discovery including the 'Golden Age' of antibiotic discovery (1940–1960). No new structural classes of antibiotics were introduced between 1962 and 2000, representing a serious innovation gap during the genomic era. This figure was adapted from Walsh and Wencewicz (2013).

New source of novel microbial and chemical diversity

Despite this crisis in natural products research, microbial secondary metabolites resulting from biosynthetic reactions that have evolved over several billion years, have tremendous potential to provide novel chemotherapeutic agents. Thus, microbial natural products continue to represent an excellent resource for drug discovery. However, currently, natural products chemistry requires a novel research approach to discover new molecules rather than known compounds. One approach to avoid the re-isolation of known compounds is to study organisms isolated from new ecological niches, i.e., to maximize the chemical diversity available from microorganisms, new sources of microbes are necessary. In most cases, secondary metabolites are only produced under certain environmental conditions, during different stages of growth, or when challenged with another species. These molecules have evolved over time to interact with diverse biological targets. Thus, newly isolated microorganisms are thought to possess substantial potential for the production of novel compounds. Sources of organisms for natural products discovery have moved from land plants in the 19^{th} century to terrestrial microorganisms in the 1940s and 1950s, marine organisms and algae in the 1970s, and marine microorganisms in 2000. In this way, new biological sources have always provided unprecedented chemical scaffolds with new biological activities. In this context, extremophiles have also attracted the attention of natural products chemists as a source of new small molecules.

Extremophiles

An extremophile is an organism that optimally grows in physically or geochemically extreme conditions that are difficult for most life to survive. These extreme conditions have traditionally included high and low temperature, high pressure, high and low pH, and high salt (**Table 1**). Various microorganisms have been previously isolated from uncommon and extreme environments, e.g., *Bacillus infernus* (Boone, Liu et al. 1995), which has been obtained from depths of approximately 2,700 m below the land surface; *Methanococcus jannaschii* (Bult, White et al. 1996), which has been isolated from the surface of the sea floor of a Pacific thermal vent; and *Deinococcus radiodurans*, which is the most radiation-resistant organism and has been isolated from cans of meat exposed to supposedly sterilizing doses of radiation (Cavicchioli & Thomas, 2000).

In recent years, extremophiles have been recognized as valuable sources of novel secondary metabolites with significant biological activities. In particular, numerous secondary metabolites with unprecedented chemical structures have been isolated from marine microorganisms that dwell at low temperature, high pressure, and high salt.

Table 1 Classification of extremophiles.

Extremophile	Environmental factor	Optimal growth conditions
Thermophile	High temperature	50-60 ˚C (moderate thermophile) 61-79 ˚C (thermophile) > 80 ˚C (hyperthermophile)
Psychrophile	Low temperature	< 20 ˚C
Piezophile (barophile)	High pressure	> 35 MPa (1 m depth = 10.5 kPa)
Halophile	High salt	> 3 % NaCl
Alkaliphile	High pH	pH > 9
Acidophile	Low pH	pH < 4

Thermophiles

Among extremophiles, thermophilic bacteria are thought to play important roles for advancement in biotechnology as sources of thermostable proteins and enzymes and as biocatalysts for the conversion of biomass to fuel-related compounds (Wiegel and Ljungdahl 1986). However, there have been few reports that describe bioactive compounds isolated from thermophilic bacteria, fungi, or archaea. Furthermore, the biosynthesis of secondary metabolites isolated from thermophilic bacteria has relatively not been well studied. Since 1964, several molecules have been isolated from thermophiles: thermorubin (actinomycete, 48−53°C, 1964), granaticin and its derivatives (*Streptomyces thermoviolaceus*, 1977), sibyllimycine (*Actinomyces* sp., 60℃, 1996), the siderophore S_{VK21} (*Bacillus licheniformis*, 60℃, 2003), thermozymocidin (*Mycelia sterilia*, 40−45°C, 1972), and the thiazole-containing alkaloid TM-64 (*Thermoactinomyces* sp., 45℃, 1978) (Wilson and Brimble 2009) (**Figure** 3). Thus, thermophilic bacteria may represent an inexhaustible source of undiscovered small molecules.

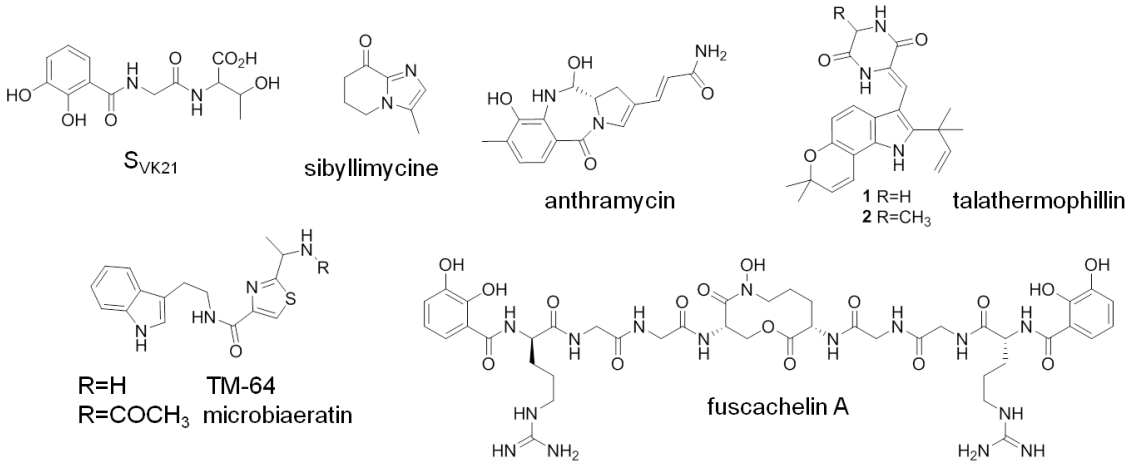


Figure 3 Natural products derived from thermophilic microorganisms.

Thermosporothrix hazakensis SK20-1^T

Thermosporothrix hazakensis SK20-1^T is a novel thermophilic bacterium, which has been isolated from ripe compost produced by a field-scale composter (Yabe, Aiba et al. 2010). The strain SK20-1^T is an aerobic Gram-positive bacterium that grows at 31–58 ℃, with optimum growth at 50 ℃. The SK20-1^T strain forms a branched substrate and aerial mycelia with spores; thus, the colony morphology and life cycle (**Figure 4**) are very similar to aerial mycelium-forming actinomycetes, from which more than 60 % of the known antibiotics have been discovered (Fenical 1993). Nevertheless, *T. hazakensis* SK20-1^T is entirely distinct from actinomycetes in the phylogenetic tree (**Figure 6**). More interestingly, the strain forms multiple exospores per mother cell by budding in branched aerial mycelia (**Figure 5**), which is a unique characteristic of this thermophilic bacterium (Yabe, Aiba et al. 2010). Numerous studies have reported on the relationship between microbial morphological differentiation and the production of secondary metabolites (Ohnishi, Yamazaki et al. 2005). Secondary metabolism is known to be commonly associated with a complex life cycle involving morphological differentiation in microorganisms. These unique characteristics thus led me to consider that the new thermophilic bacterium *T. hazakensis* SK20-1^T possesses the potential to produce unprecedented metabolites.

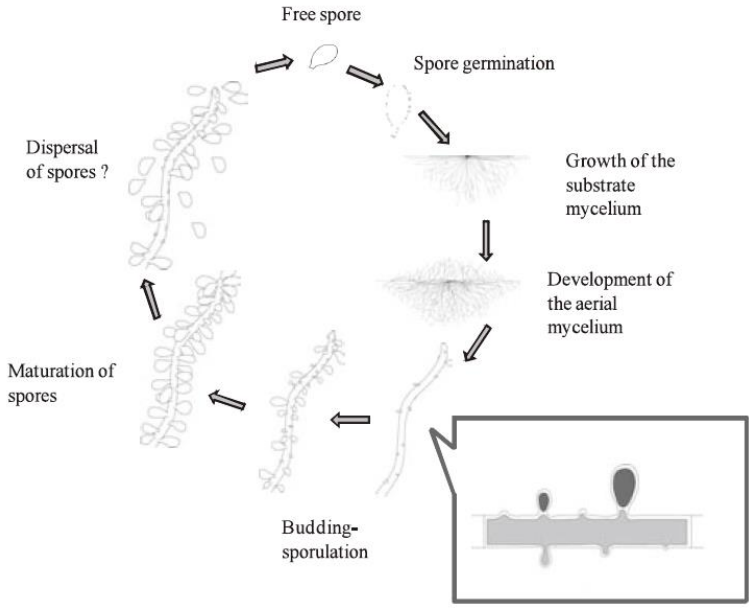


Figure 4 The life cycle of *Thermosporothrix hazakensis* SK20-1^{T}. This diagram was adapted from Yabe, Aiba et al. (2010)[b].

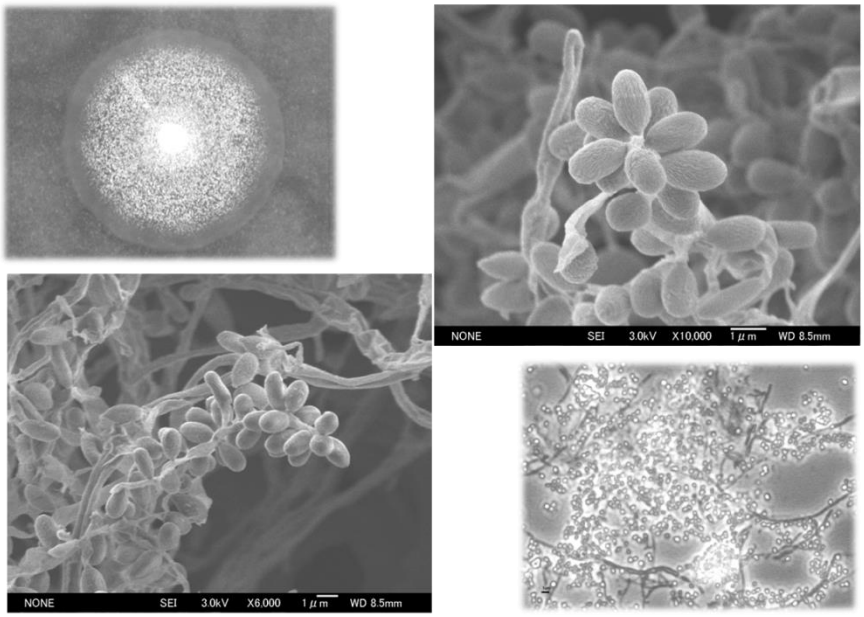


Figure 5 Aerial mycelium forming colony and multiple exospores observed in *T. hazakensis* SK20-1^{T}. These photos are included by courtesy of S. Yabe.

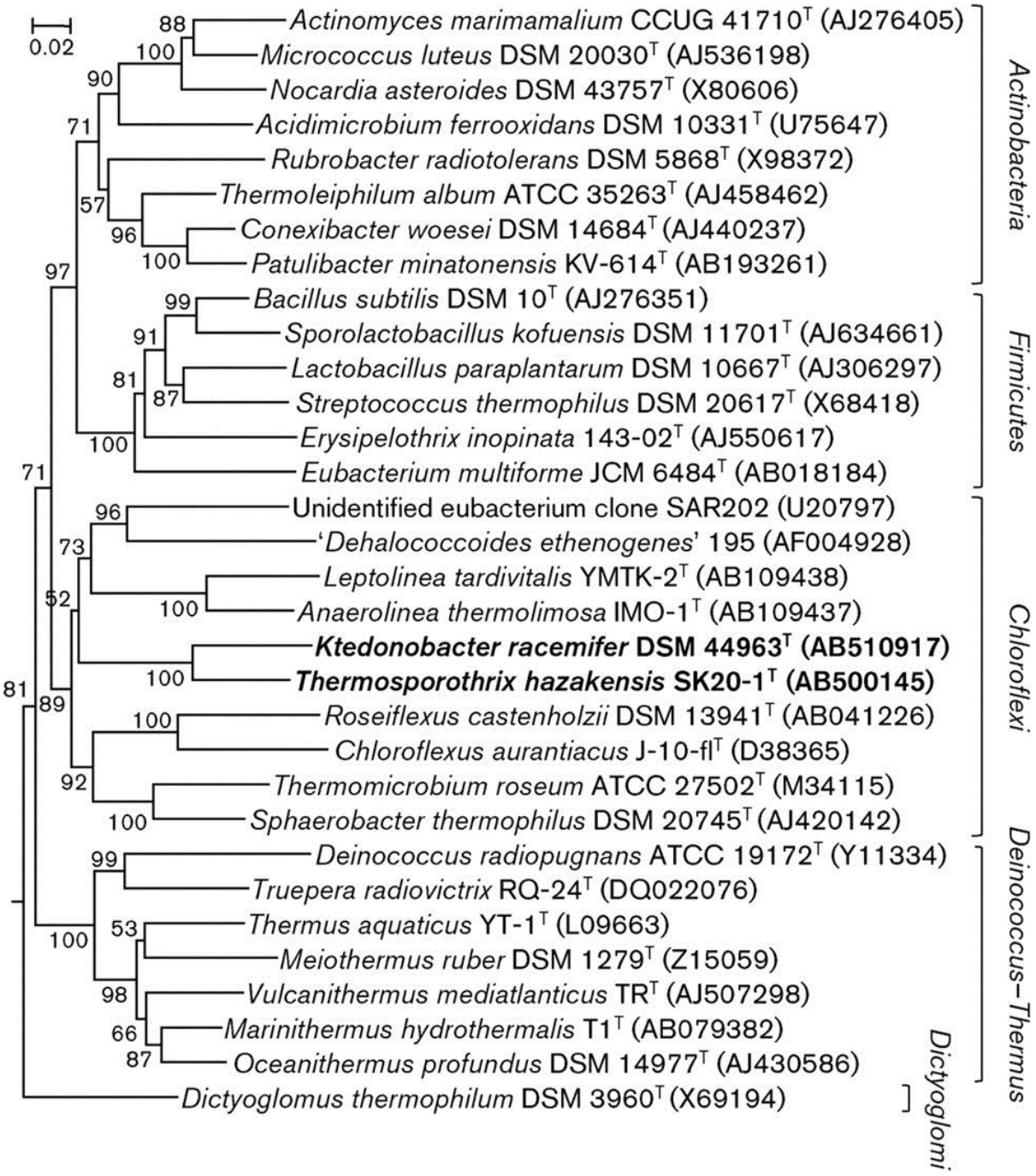


Figure 6 The phylogenetic tree of *T. hazakensis* SK20-1^T, which represents a new member in the phylum *Chloroflexi* and is distinct from actinomycete. This phylogenetic tree was adapted from Yabe, Aida et al. 2010[a].

1. Bioinformatics analysis of the *T. hazakensis* SK20-1^T genome for sequence-based discovery of secondary metabolites

1.1. Introduction

Prior to chemical screening that included culture, isolation, purification and spectral analysis, bioinformatics analysis of the genome sequence of *T. hazakensis* SK20-1^T was conducted to identify the presence of gene clusters, which may be involved in the biosynthesis of secondary metabolites. Furthermore, I expected that additional extensive analyses using bioinformatics tools may enable me to predict the corresponding chemical structures. If it is possible to predict putative structures based on in silico analysis, culture conditions such as the culture medium or isolation method could be optimized to facilitate the determination of the products (Zazopoulos, Huang et al. 2003). Through the rapid advancement of sequencing technologies and the consequently low cost of sequencing, genome sequencing has become accessible to laboratory-scale research groups. Furthermore, excellent software has been developed to aid in handling the large amount of data generated from genome sequencing of microorganisms of interest. Therefore, the innovation of the 1990s enabled access to abundant DNA sequence resources, and many natural products chemists have explored novel compounds using a strategy of genome-guided natural products discovery. In fact, genome sequencing and analysis was utilized in the discovery of previously unprecedented secondary metabolites from well-studied bacteria such as *Streptomyces coelicolor* (Challis and Ravel 2000, Lautru, Deeth et al. 2005). Several structures of natural products derived from genome-utilizing methodologies are illustrated in **Figure 1-1** (Kreutzer and Nett 2012). (Kreutzer and Nett 2012). A report of the initial findings on secondary metabolites from thermophilic bacteria based on genome analysis revealed an orphan NRPS gene cluster that biosynthesizes fuscachelins from the thermophilic

actinomycete *Thermobifida fusca*. The biosynthetic pathway of fuscachelins exhibited atypical features including nonlinear peptide assembly, and the end product was a siderophore molecule that possesses a distinct novel iron-chelating domain (Dimise, Widboom et al. 2008).

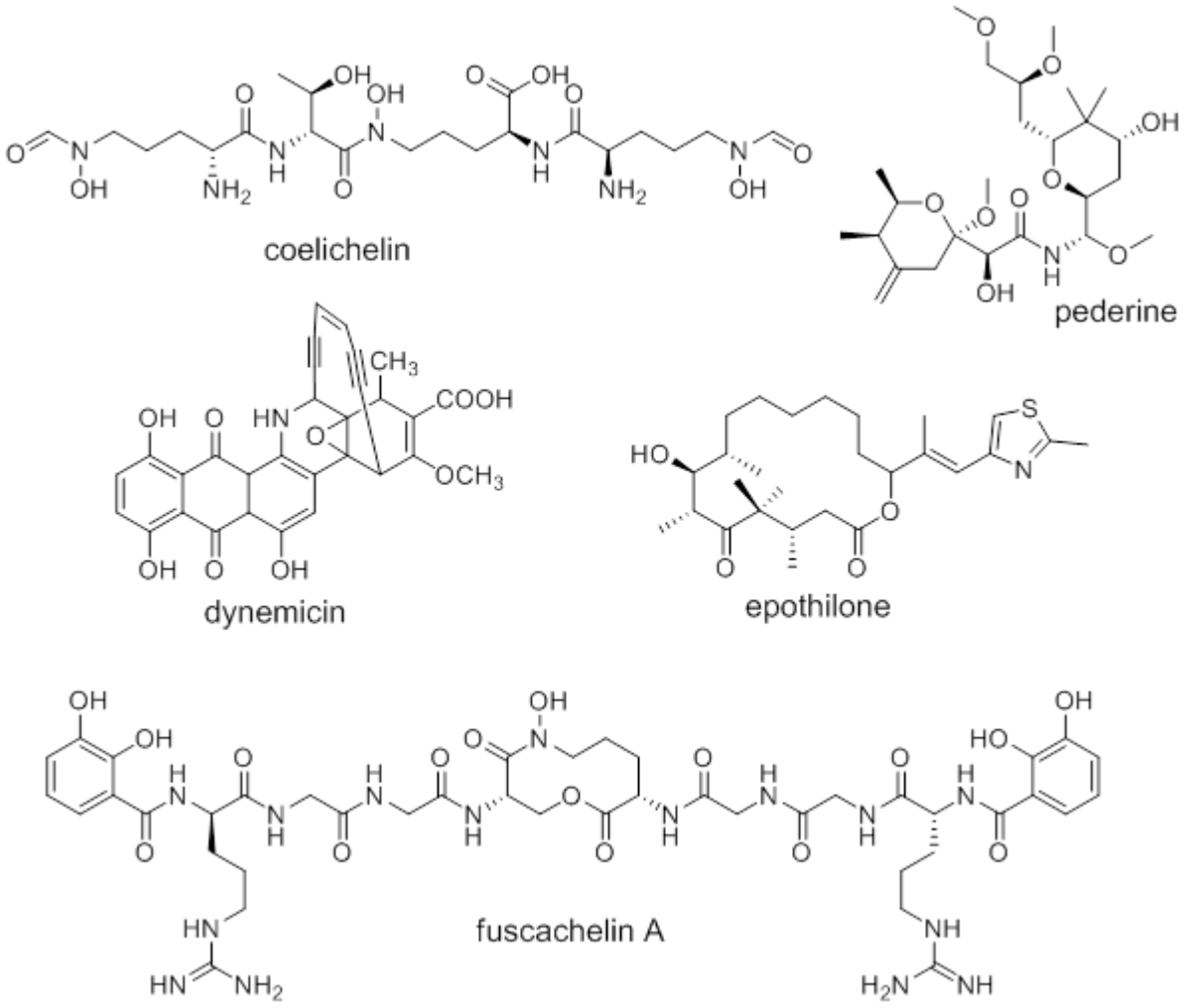


Figure 1-1 Natural products discovered using methods that have benefited from microbial genomics such as genome mining or genome scanning.

Structure prediction from a biosynthetic gene locus is conducted based on the ‘co-linearity’ between the number of modules and the number of building blocks in megasynthases such as modular type I polyketide synthase (PKS) or nonribosomal peptide synthase (NRPS). Generally, the number and order of the modules typically correspond to the sequence of building blocks in the peptide or polyketide. This mechanistic paradigm is often referred to as the co-linearity rule. Therefore, if the organization of the gene cluster is known, the exact structure can be deduced. However, these megasynthases do not always catalyze reactions in a straightforward manner and often produce an end-product that differs from the anticipated structure. This variation in organization and functionality is due to diverse reasons including undiscovered missing domains or stuttering, skipping, and iterating domains.

1.2. Structure-determining factors in NRPS and PKS

1.2.1. Adenylation (A) domain in NRPS

The A domain in NRPS modules recognizes a specific amino acid to activate the amino acid as aminoacyl-adenosine monophosphate (AMP), which is subsequently loaded onto the peptidyl carrier protein (PCP) domain (**Figure 1-2**). Extensive studies on the 3D structure of A domains revealed conserved residues, which are critical for substrate binding and are referred to as the nonribosomal code or the specificity-conferring code (Stachelhaus, Mootz et al. 1999, Challis, Ravel et al. 2000). A prediction of the substrate through the identification of an A domain enables us to propose a primary structure produced by NRPS.

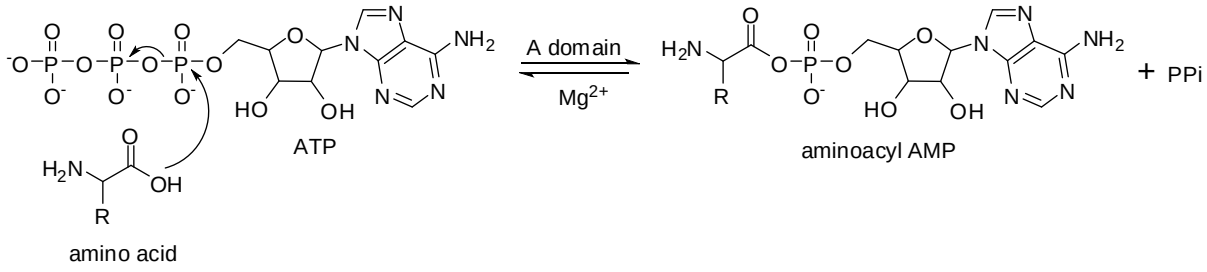


Figure 1-2 The activation (adenylation) of a substrate, amino acid by the A domain.

1.2.2. Acyl transferase (AT) domain in PKS

The A domain activates an amino acid to load onto the PCP domain, whereas the acyl transferase (AT) domain transfers a particular acyl moiety from acyl-coenzyme A (CoA) to the acyl carrier protein (ACP) domains (**Figure 1-3**) (Fischbach and Walsh 2006). As found for the A domain in NRPS, diverse studies on AT domains revealed four conserved motifs that control substrate specificity (Tsai and Ames 2009). The AT is comprised of two domains (a large core domain and a small domain), and a catalytically essential serine (GH**S**XG) in motif2 is located between the two domains. Additionally, motif1 is located 30 residues

upstream of the catalytic serine, and motif3 is located 100 residues downstream of the catalytic serine. These motifs form a substrate-binding pocket to determine the size and shape of the substrate. Furthermore, motif4, which is a hypervariable region at the C-terminus, has been identified to be important in substrate recognition. However, the classification of AT domains has not yet been sufficiently developed, and the accuracy of the prediction of the substrate is relatively low, in contrast to that of the A domain.

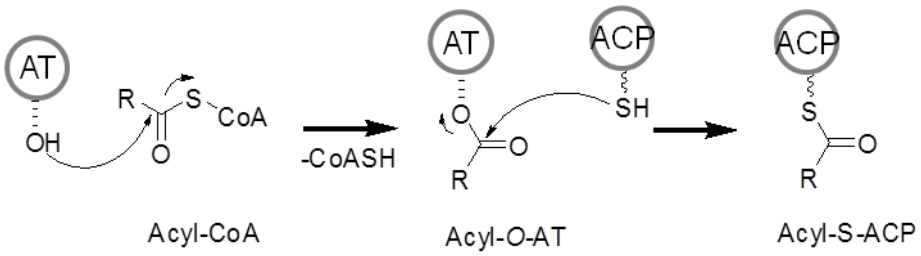


Figure 1-3 Recognition and binding of the substrate by the AT domain and substrate loading on to the ACP domain.

In addition to the A domain in NRPS and the AT domain in PKS, several other determinants affect the chemical structure of the secondary metabolites, such as the epimerization (E) domain, which catalyzes the stereoconversion of L-amino acids to D-amino acids, and the cyclization (Cy) domain, which is involved in heterocycle formation of oxazoline/oxazole or thiazoline/thiazole in NRPS. In PKS, three representative reduction-catalyzing domains (i.e., ketoreductase (KR), enoyl-CoA reductase (ER) and dehydratase (DH)) are involved in the diversification of polyketide compounds. Finally, several enzymatic posttranslational processes such as glycosylation, methylation, halogenation and oxidation sustain the construction of a complete structure. These modifications often influence the bioactivity, selectivity and pharmacokinetic properties of the molecules.

1.3. Methods

1.3.1. Genome sequencing of *T. hazakensis* SK20-1^T

The genomic DNA of *T. hazakensis* SK20-1^T was purified using a generic protocol (Ausubel 1999) and sequenced using the Genome Analyzer IIx system (Illumina) at Takara Bio, Kyoto. The sequence reads were assembled using the ALLPATHS-LG short read genome assembler ver. 39099 (http://www.broadinstitute.org/software/allpaths-lg/blog/). Genome sequences were annotated using the MiGAP (http://www.migap.org/), protein BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) (Altschul, Gish et al. 1990), and FRAMEPLOT 2.3.2 (http://www0.nih.go.jp/~jun/cgi-bin/frameplot.pl) programs (Ishikawa and Hotta 1999).

1.3.2. Genome analysis

The draft genome sequence was subject to automatic annotation using the RAST server (Aziz, Bartels et al. 2008). Next, antiSMASH (Blin, Medema et al. 2013) was used to identify secondary metabolite-related gene clusters in the genome. A graphical representation of the PKS and NRPS domain organization was obtained from amino acid sequences using freely available databases such as PKS/NRPS Analysis (Bachmann and Ravel 2009) and the NRPS/PKS database (Ansari, Yadav et al. 2004). The specificity of the A domains in an NRPS system was identified using a specialized tool called NRPSpredictor (Rausch, Weber et al. 2005) and NRPSpredictor2 (Rottig, Medema et al. 2011). The specificity of the AT domains was analyzed based on Yadav et al. (2003) and Smith and Tsai (Smith and Tsai 2007) by multiple sequence alignment using CLUSTALW (Goujon et al., 2010; Larkin et al., 2007) with similar AT domains obtained from BLAST. All domains including the accessory

domains were analyzed for their activity by identifying the catalytically active residues in their amino acid sequences using BLAST and CLUSTALW. The nucleotide sequences used for cloning were based on the amino acid regions specified by Pfam (http://pfam.sanger.ac.uk/).

Table 1-1 List of tools used in this study to analyze the genome sequences.

Program	Hyperlink	Purpose	Ref
RAST	http://rast.nmpdr.org	Annotation	Aziz, Bartels et al. 2008
antiSMASH	http://antismash.secondarymetabolites.org	NRPS, PKS, Lantipeptide	Blin K et al., 2013
PKS/NRPS analysis	http://nrps.igs.umaryland.edu/nrps/	NRPS, PKS	Bachmann and Ravel, 2009
NRPSpredictor	http://ab.inf.uni-tuebingen.de/toolbox/index.php?view=domainpred	NRPS	Rausch et al., 2005
NRPSPredictor2	http://nrps.informatik.uni-tuebingen.de/Controller?cmd=SubmitJob	NRPS	Rottig et al., 2011

1.4. Results and Discussion

The total contig length, maximum contig size, and average contig length were 7,300,446 bp, 1,061,024 bp, and 114,069 bp, respectively (GC content, 53.1 %). It is known that not all bacteria are capable of producing secondary metabolites and that secondary metabolic clusters are unevenly distributed within the bacteria taxa, with various groups being regarded as highly prolific. It has been reported that bacteria possessing genomes smaller than three megabases tend not to produce secondary metabolites, thereby suggesting that natural product formation is a luxury that only large-genome bacteria can afford (Donadio S et al. 2007). Generally, the genome size of thermophilic bacteria is distributed over a wide range from 1.5–2.5 Mb, and the thermophilic actinobacterium *Thermobifida fusca* possesses a large genome of 3.6 Mb (http://en.wikipedia.org/wiki/List_of_sequenced _bacterial_genomes). In comparison to the genomes of typical thermophilic bacteria, *T. hazakensis* SK20-1^T thus possesses an exceptionally large genome (7.3 Mb), which suggests the existence of additional genes that are involved in a diverse and complicated metabolism. The genomic sequence of *T. hazakensis* SK20-1^T was analyzed using the web-based annotation service, RAST (http://rast.nmpdr.org) (Aziz, Bartels et al. 2008). A total of 6,739 putative protein-coding sequences were predicted, 25 % of which were assigned to a subsystem including carbohydrates, amino acids, fatty acids, protein metabolism, and DNA metabolism. Scanning the genome revealed a pool of PKS and NRPS genes that produce secondary metabolites. According to the annotation of each coding sequence, I identified the presence of gene loci that biosynthesize lysine via α-aminoadipate including *lysW* (*thzk4003*), *lysX* (*thzk4004*), *lysZ* (*thzk4005*), *lysY* (*thzk4006*), *lysJ* (*thzk4007*), and *lysK* (*thzk4008*) that are found in thermophilic bacteria or archaea..

Next, antiSMASH was utilized to identify secondary metabolite-related gene clusters in the genome. Consequently, all 23 gene loci were identified (**Table 1-2**, **Figure 1-4**). A detailed organization of each gene locus is described as follows.

Table 1-2 Secondary metabolite related gene loci identified in *T. hazakensis* SK20-1^{T}.

Cluster 1, 4, and 12: type II PKS
Cluster 2: hybrid NRPS - type I PKS
Cluster 3, 7, 9, 11, 15, 16, 18, 20, 21, 22[a], and 23: lantipeptide (bacteriocin)
Cluster 5: NRPS – type I PKS
Cluster 6: Hserlactone
Cluster 8: type I PKS - NRPS
Cluster 10 [a]: NRPS – type I PKS
Cluster 13 [a]: type I PKS – NRPS
Cluster 14: type I PKS – NRPS
Cluster 17 [a]: NRPS – type I PKS
Cluster19: type I PKS

[a] These clusters are subsequently described.

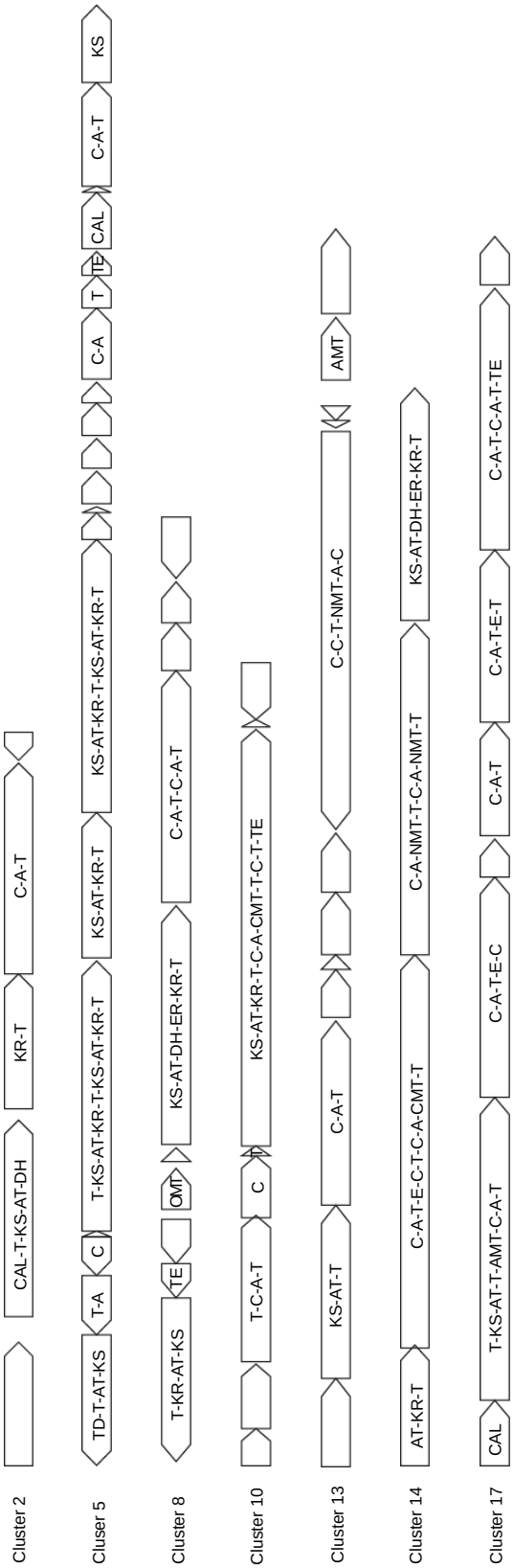


Figure 1-4 Hybrid NRPS-PKS gene clusters identified from analysis of the *T. hazakensis* SK20-1^{T} genome.

In particular, the PKS/NRPS hybrid gene loci are more abundant than the discrete NRPS and PKS gene loci. From the genome sequence, seven PKS/NRPS hybrid gene loci were identified, whereas only four PKS and no NRPS gene loci were recognized. In addition to NRPS/PKS, 11 lantipeptide or bacteriocin gene loci were identified, suggesting that *T. hazakensis* SK20-1^T may be a plausible ribosomal peptide producer. To date, none of these gene clusters can be attributed with certainty to a known natural product.

Most PKS/NRPS gene loci identified in *T. hazakensis* SK20-1^T appear to be insufficient to produce or predict the corresponding products because the organization of the clusters are irregular and specific domains are absent. For example, a TE domain is absent in cluster 2, whereas a nearly similar gene cluster that harbors a TE domain was discovered in *Ktedonobacter racemifer* DSM 44963, which is the most phylogenetically close bacterium to *T. hazakensis* SK20-1^T (**Figure 1-5**). This finding suggests that products synthesized by NRPS/PKS in cluster 2 may be not properly released from the ACP domain in the final step. Additionally, the transcription directions are shuffled, which makes it difficult to predict the biosynthetic pathway for clusters 5 and 8.

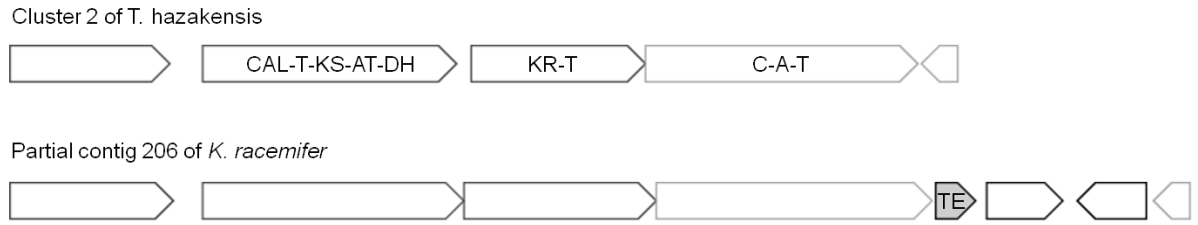


Figure 1-5 Comparison of PKS/NRPS gene loci from *T. hazakensis* SK20-1^T and *K. racemifer.*

Moreover, it was difficult to predict the physiological substrates of the A domains or the AT domains because the residues in the substrate-binding pocket are not fully conserved. This variation in the biosynthetic genetic logic may indicate the potential production of novel natural product scaffolds by *T. hazakensis* SK20-1^T.

[Cluster 10]

Of 23 identified gene regions, cluster 10 displays a relatively well-organized gene cluster for the prediction of a putative chemical skeleton. *In silico* analysis of cluster 10 revealed that the gene loci consist of hybrid NRPS/PKS genes in addition to transcription regulator- and membrane receptor-encoding genes (**Table 1-3**). Through a more detailed investigation of each domain in NRPS, I recognized that only two cysteine-activating A domains are present in the NRPS region, in contrast to the presence of four Cy domains and four PCP domains, suggesting that these A domains presumably iteratively catalyze the activation of cysteine (**Figure 1-6**).

Table 1-3 List of genes in cluster 10 and their putative function.

Gene	Size (bp)	Predicted encoded function
10-1	219	Transcription repressor
10-2	249	Hypothetical
10-3	465	Integrase
10-4	339	Transcription regulator
10-5	276	Heme response regulator
10-6	810	Sensor histidine kinase
10-7	1374	Penicillin-binding protein-4 (PBP-4), 57/77
10-8	912	Thioredoxin reductase, 23/43
10-9	1731	Dipeptide-binding protein (DppE), 25/45
10-10	3915	NRPS (ACP-Cy-A_{cys}-PCP)
10-11	1656	NRPS (Cy)
10-12	282	NRPS (PCP)
10-13	11262	PKS/NRPS (KS-AT-KR-ACP-Cy-A_{cys}-MT-PCP-Cy-PCP-TE)
10-14	189	Hypothetical protein
10-15	1554	Synaptotagmin-1
10-16	321	Hypothetical protein
10-17	429	hypothetical (NAD binding)
10-18	324	hypothetical (NAD binding)
10-19	1044	L-ornithine cyclodeaminase, 30/47
10-20	1062	cysteine synthase
10-21	2632	NRPS (A_{ser}-PCP-C)

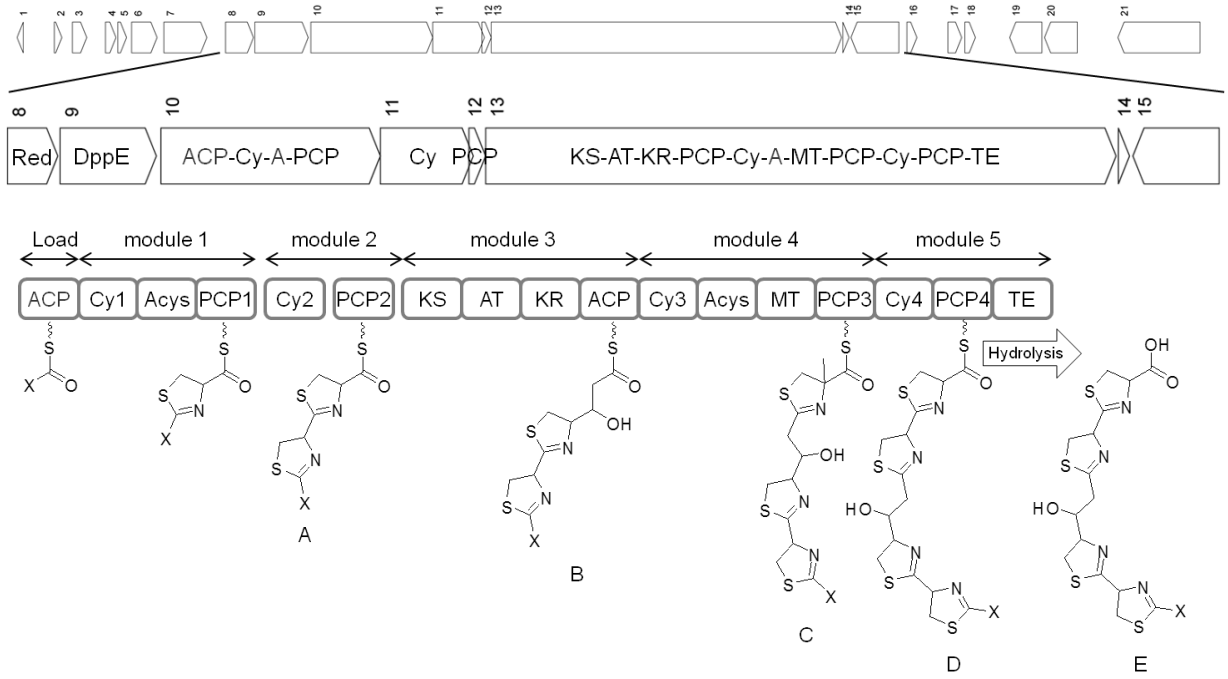


Figure 1-6 Genetic organization of the hybrid PKS/NRPS cluster 10 and the predicted lipopeptide structure following analysis of each domain. A domains in module 1 and module 4 are anticipated to catalyze the activation of cysteine two times.

Functional analysis of the four NRPS modules and a PKS module led me to predict a partial structure that possesses four thiazoline moieties connected by a two-carbon unit (**Figure 1-6E**). In detail, an unidentified starting unit is loaded onto the ACP domain. The A domain is predicted to be specific for Cys activation and will subsequently transfer Cys from a Cys–AMP to both PCP1 and PCP2 domains. Thus, an unidentified acyl group and the two cysteinyl groups are covalently transferred to ACP and PCPs, respectively, of NRPS (10-10–12). Subsequently, Cy1 and Cy2 catalyze the successive condensation and cyclization of each Cys to form two thiazoline rings (A). The mixed PKS/NRPS multifunctional enzyme (10-13) completes the elongation. This protein consists of 11 domains that form three elongation modules and a TE domain. The KS domain loads the malonyl group onto the two-thiazoline intermediate produced by NRPS (10-10–12) that is linked to the ACP domain (B). Module 4 in 10-13 then catalyzes the condensation and cyclization of methyl-Cys, which is presumably produced by methyltransferase (MT) in module 4, to form intermediate C. Final elongation occurs via module 5 to yield D, and D is subsequently hydrolyzed by the TE domain to yield

the end product (E). Subsequent analysis using the SciFinder® program (scifinder.cas.org) revealed that this putative partial structure (E) is a typical structure found in several iron-chelating agents such as yersiniabactin and piscibactin (**Figure 1-7**) that have been isolated from *Yersinia enterocolitica* and *Photobacterium damselae* subsp. *piscicida*, respectively. These two microbes are known to cause infectious disease in humans and fish, and the produced siderophores contribute to the virulence of these pathogens through their involvement in iron uptake in iron-deficient environments. Interestingly, the mixed NRPS/PKS gene organization of cluster 10 is also very similar to that of yersiniabactin and piscibactin biosynthetic clusters (Gehring, DeMoll et al. 1998, Osorio, Juiz-Rio et al. 2006) (**Figure 1-8**).

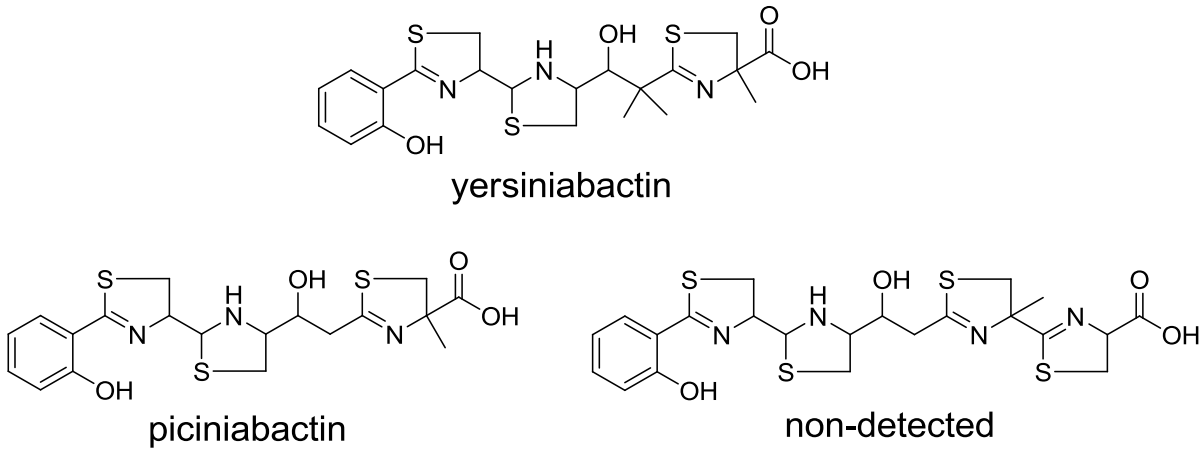


Figure 1-7 Structures of the iron-chelating siderophores yersiniabactin, piscibactin, and the non-detected structure predicted by the gene organization of the piscibactin biosynthetic cluster.

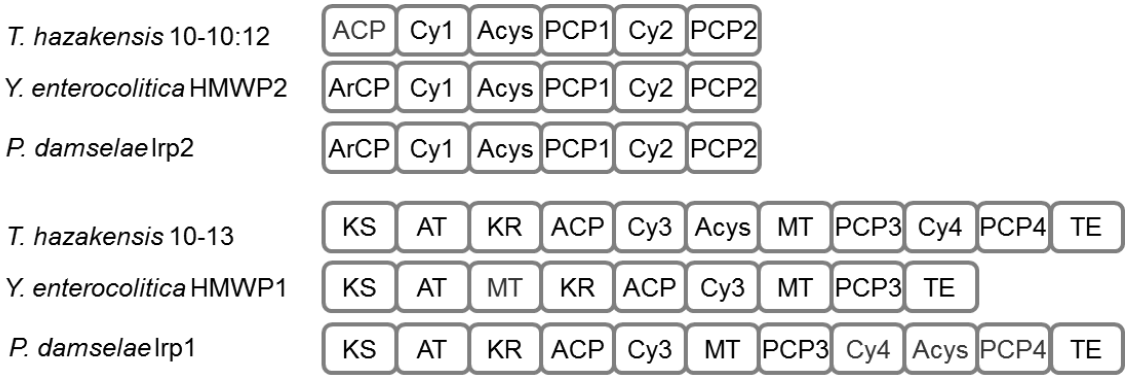


Figure 1-8 Comparative analysis of the NRPS/PKS functional domains in cluster 10 of *T. hazakensis*, the yersiniabactin biosynthetic cluster of *Y. enterocolitica*, and the piscibactin biosynthetic cluster of *P. damselae.*

Although these three gene organizations are very similar, cluster 10 lacks the salicylate adenylation domain that activates an initial precursor such as YbtE and Irp5 in the yersiniabactin and piscibactin biosynthetic pathways, respectively. Thus, I speculated that the product of cluster 10 may differ from the typical salicylate-thiazoline siderophore structure. Additional analysis of cluster 10 identified additional coding sequences that are likely involved in secondary metabolite biosynthesis downstream of 10-13.

In detail, 10-19 and 10-20 encode ornithine cyclodeaminase and cysteine synthase respectively. And 10-21 encodes NRPS that harbors C, A (serine specific), and PCP domains. Bioinformatics analysis suggests that the conversion of L-serine to L-2,3-diaminopropionate (DAP) is catalyzed by the concerted actions of 10-19 and 10-20. 10-20 is a homolog of cysteine synthase enzymes that catalyze the pyridoxal phosphate (PLP)-dependent replacement and elimination of the β substituent of its substrate (L-serine) (**Figure 1-9**). During catalysis, this enzyme forms a PLP-bound α-aminoacrylate intermediate. Therefore, 10-20 is expected to use a similar mechanism to form a Schiff base linkage between PLP and an α-aminoacrylate intermediate (**Figure 1-10**). 10-20 subsequently catalyzes a β-substituent replacement reaction analogous to that observed in cysteine synthases. However, although cysteine synthases use sulfur derived from sulfide as the nucleophile, I predicted that 10-20 uses the nitrogen of ammonia as the nucleophile. The source of this nucleophile is the ammonia released from L-ornithine by 10-19. 10-19 is a homolog of ornithine

cyclodeaminases that convert L-ornithine to L-proline with the release of ammonia (Zhao, Luo et al. 2007). Based on the presence of DAP biosynthetic gene loci (10-19, 10-20) and an additional adenylation domain (10-21), I expected that DAP rather than salicylate may be the precursor for cluster 10 in yersiniabactin and piscibactin biosynthesis. This analysis resulted in a proposed biosynthetic product of cluster 10, as shown in **Figure 1-11**.

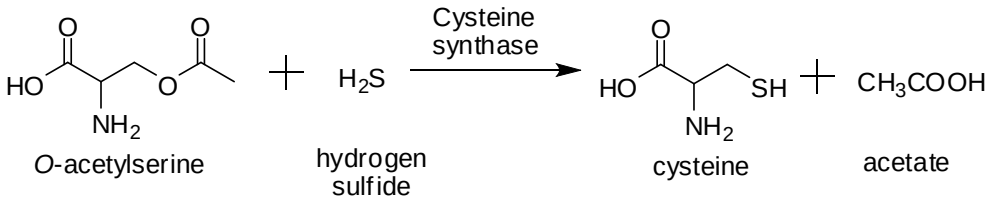


Figure 1-9 The reaction catalyzed by cysteine synthase. Cysteine synthase catalyzes the formation of cysteine and acetate from *O*-acetylserine and hydrogen sulfide.

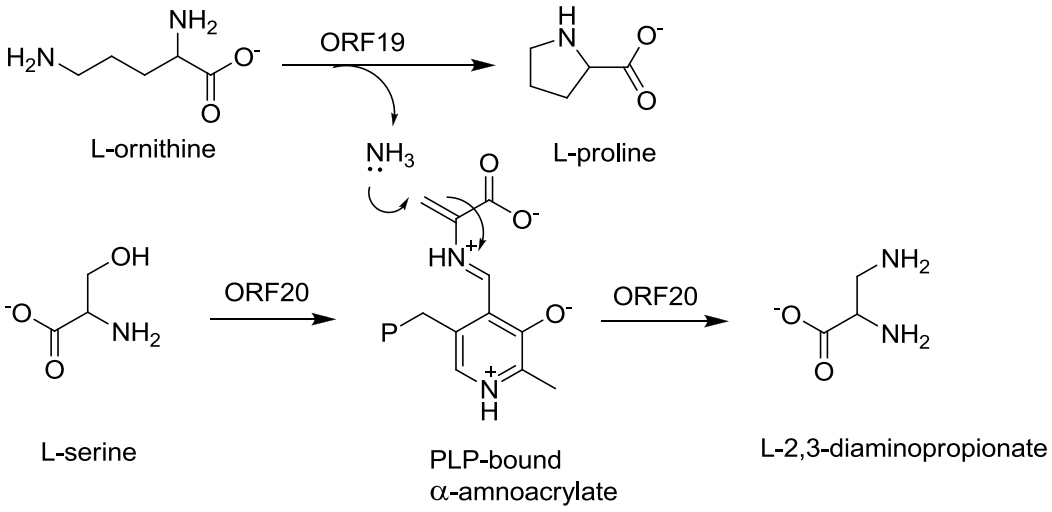


Figure 1-10 Schematic representation of the proposed L-2,3-diaminopropionate biosynthesis through the ORF19- and ORF20-coupled reaction.

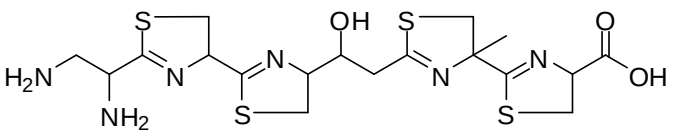


Figure 1-11 The proposed structure of the product produced by the biosynthetic cluster 10.

[Cluster 13]

Additionally, certain features (*thzk3324-3327*) led me to consider cluster 13 (PKS/NRPS) as the gene cluster encoding an unusual hydroxymalonyl-acyl carrier protein (ACP) or aminomalonyl-ACP as extender units (**Figure 1-12**), which are involved in the biosynthesis of zwittermicin (Emmert, Klimowicz et al. 2004, Chan, Boyne et al. 2006). In bacterial polyketide biosynthesis, the utilization of extender units other than malonyl-CoA or methylmalonyl-CoA is rarely observed. Thus, the incorporation of an unusual extender unit may indicate the production of new compounds.

Table 1-4 Annotation of genes involved in the production of hydroxymalonyl-CoA.

Gene	Size (bp)	Proposed function	% Id/Si
thzk3324	861	3-hydroxybutyryl-CoA dehydrogenase (*Herpetosiphon aurantiacus* DSM 785)	66/81
thzk3325	249	phosphopantetheine-binding protein (*Clostridium papyrosolvens*)	61/77
thzk3326	1134	acyl-CoA dehydrogenase (*Nostoc punctiforme* PCC 73102)	63/78
thzk3327	1056	FkbH-like protein (*H. aurantiacus* DSM 785)	59/80

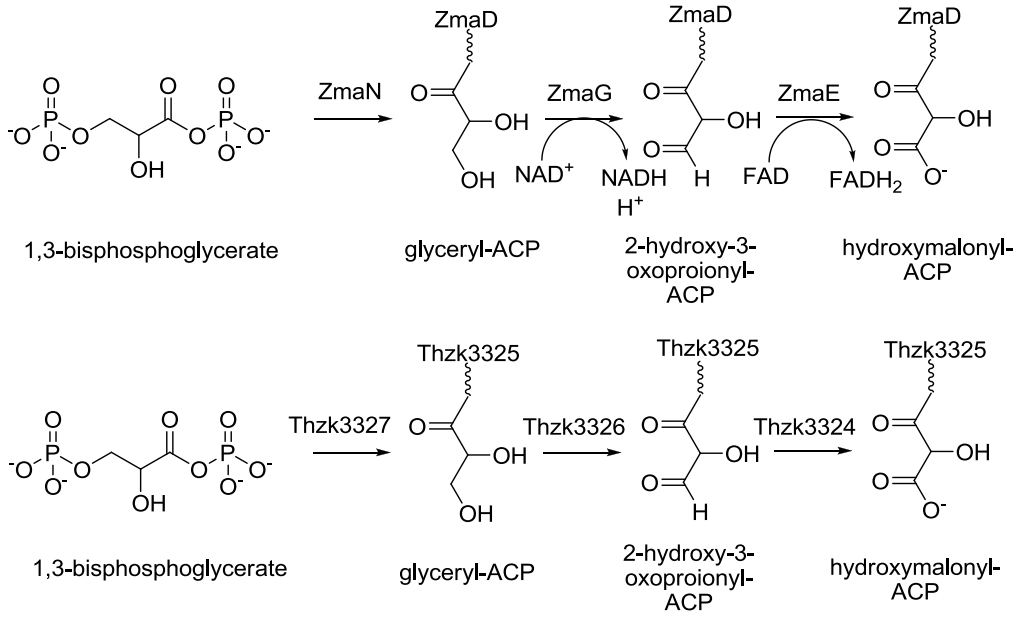


Figure 1-12 The biosynthetic pathway of hydroxymalonyl-CoA in zwittermicin biosynthesis (top) and cluster 13 (bottom) of *T. hazakensis* SK20-1^{T}. Thzk3324, 3-hydroxybutyryl-CoA dehydrogenase; Thzk3325, acyl carrier protein (ACP); Thzk3326, acyl-CoA dehydrogenase; and Thzk3327, FkbH-like protein.

[Cluster 17]

Domain analysis of cluster 17 using bioinformatics tools in **Table 1-1** enabled me to determine the genetic organization of this cluster as shown in **Table 1-5** and **Figure 1-4**.

Table 1-5 Annotation of genes in cluster 17

Gene	Size (bp)	Proposed function	% Id/Si [a]
17-1	1791	Long-chain-fatty-acid:ACP ligase (*Calothrix* sp. PCC 7507)	55/69
17-2	8274	amino acid adenylation domain-containing protein (*Nostoc punctiforme* PCC 731020]	42/59
17-3	6012	NRPS (*Anabaena variabilis* ATCC 29413)	41/58
17-4	1056	taurine catabolism dioxygenase TauD/TfdA (*N. punctiforme* PCC 73102)	52/73
17-5	3123	amino acid adenylation protein (*Crinalium epipsammum* PCC 9333)	37/55
17-6	4659	NRPS (*Microcoleus* sp. PCC 7113)	42/59
17-7	7248	NcpB (*Nostoc* sp. ATCC 53789)	42/60
17-8	1371	Major Facilitator Superfamily transporter (*Cylindrospermum stagnale* PCC 7417)	42/64

[a] Identity (Id); Similarity (Si)

A substrate search of the AT domain in 17-1 indicates the activation of malonate to malonyl-CoA, which is a common extender unit in modular PKS. For NRPS loci, the apparent activation of isoleucine by the A domain in 17-2, histidine by the A domain in 17-3, and serine by the second A domain in 17-7 yields the corresponding aminoacyl-AMPs. In contrast, a fatty acyl-AMP ligase (FAAL) domain in 17-1 that is similar to a fatty acid-adenylating enzyme may play a role as the initiation/loading module. If the starter unit is provided as the free acid, it may be activated and loaded by an NRPS-like adenylation and thiolation-loading di-domain (A-ACP) (Hertweck 2009). Additionally, a histidine residue that is typical for long-chain acyl-CoA synthetase is present 38-residue downstream of the ATP-binding motif. In typical A domains, this residue is phenylalanine, and in small-chain

acyl-CoA synthetase, it is tryptophan (Gulick 2009) (**Figure 1-13**). Thus, I concluded based on the bioinformatics results that an AMP-dependent synthetase and ligase (17-1) from cluster 17 may catalyze the adenylation of a long aliphatic carboxylic acid. Furthermore, three A domains in 17-5, 17-6, and 17-7 did not exhibit substrate specificity, suggesting the possibility of a novel A domain signature (**Table 1-6**).

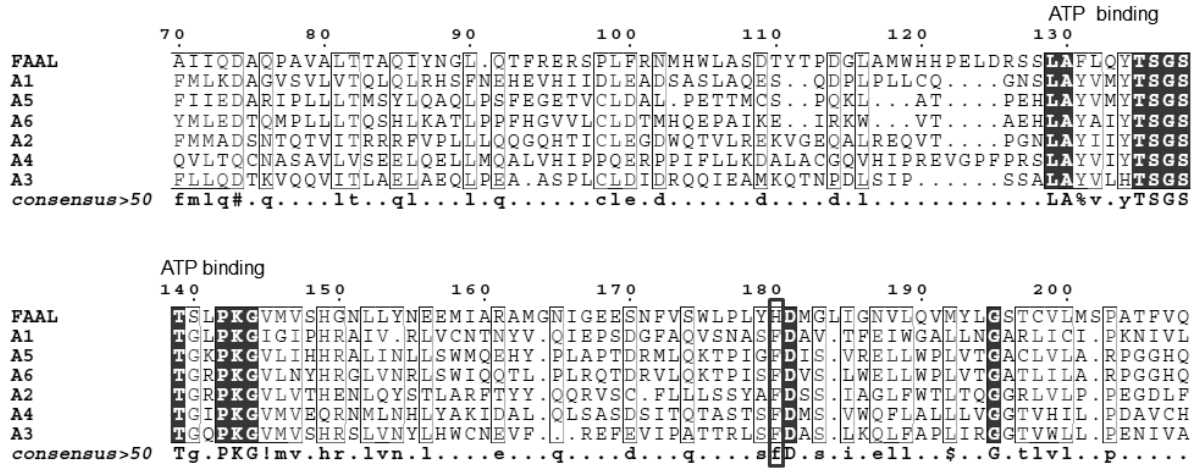


Figure 1-13 Amino acid sequence alignment of FAAL and A domains in cluster 17. His181 of FAAL is the typical residue of the acyl-CoA ligase gene, whereas an adenylation domain is conserved as phenylalanine (F) at the corresponding position.

Table 1-6 The NRPS code extracted from the A domains in cluster 17 and their putative amino acid substrates.

Position	Residues in the binding pocket	Prediction
A-domain 1 (17-2)	D A F F L G G T F -	Ile
A-domain 2 (17-3)	D S A L I A E V W -	His
A-domain 3 (17-5)	D A K N L G L A N -	No hit (orn*)
A-domain 4 (17-6)	D M W L G A D S -	No hit (haorn*)
A-domain 5 (17-7)	D I R H I G L I E -	No hit (hpg*)
A-domain 6 (17-8)	D V W H I S L I E -	Ser

*nearest neighbor amino acid by NRPSpredictor2

orn, ornithine; haorn, N-acetyl-delta-N-hydroxyornithine; hpg, 4-hydoxy-phenyl-glycine

17-3 also contains a predicted epimerization (E) domain, suggesting that the stereochemistry of the corresponding amino acid in the product has the D-configuration. Although the predicted structure is incomplete because of an unknown starting unit from FAAL and three extending building blocks from the A3-A4-A5 domains, a mixed polyketide-nonribosomal peptide structure was putatively constructed, as shown in **Figure 1-14** from the analysis of cluster 17. Additionally, the 17-4 gene product (351 residues) exhibits 52 % identity to the taurine catabolism dioxygenase TauD in *Nostoc punctiforme* PCC 73102. Several TauD homologs in NRPS are known to catalyze the β-hydroxylation of amino acids (Strieker, Kopp et al. 2007). Therefore, I also expected that the A3 domain may adenylate β-hydroxylated amino acids produced by the 17-4 gene product.

As a final step, according to the manner of release from a thiolation domain (i.e., ACP or PCP) of the thioesterase (TE), which catalyzes the cleavage of thioesters, an acyclic acid or a macrolactone/macrolactam may be synthesized (Du and Lou 2010). The TE domain in cluster 17 is a typical hydrolytic enzyme possessing the conserved motif GXSXG (Rajakumari and Daum 2010). Moreover, the TE domain includes a characteristic proline 17 upstream of GXSXG, which is typical for a TE that acts as a cyclase (Roonsawang et al., 2007) (**Figure 1-15**). Thus, this TE domain of cluster 17 is speculated to cyclize the linear putative lipohexapeptide by mediating a nucleophilic attack between the amino group introduced by aminotransferase (AMT) on the fatty acid and the terminal carboxyl group (**Figure 1-14**).

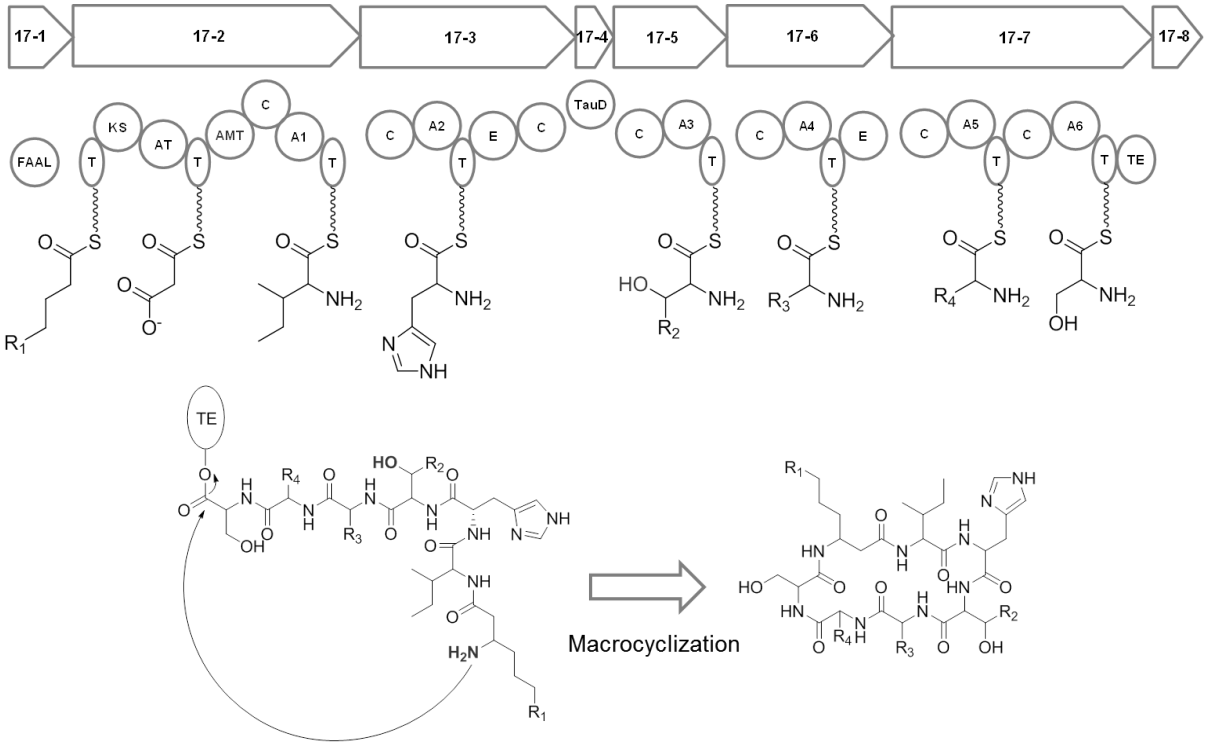


Figure 1-14 Genetic organization of the hybrid PKS/NRPS cluster 17 and the predicted lipopeptide structure following analysis of each domain.

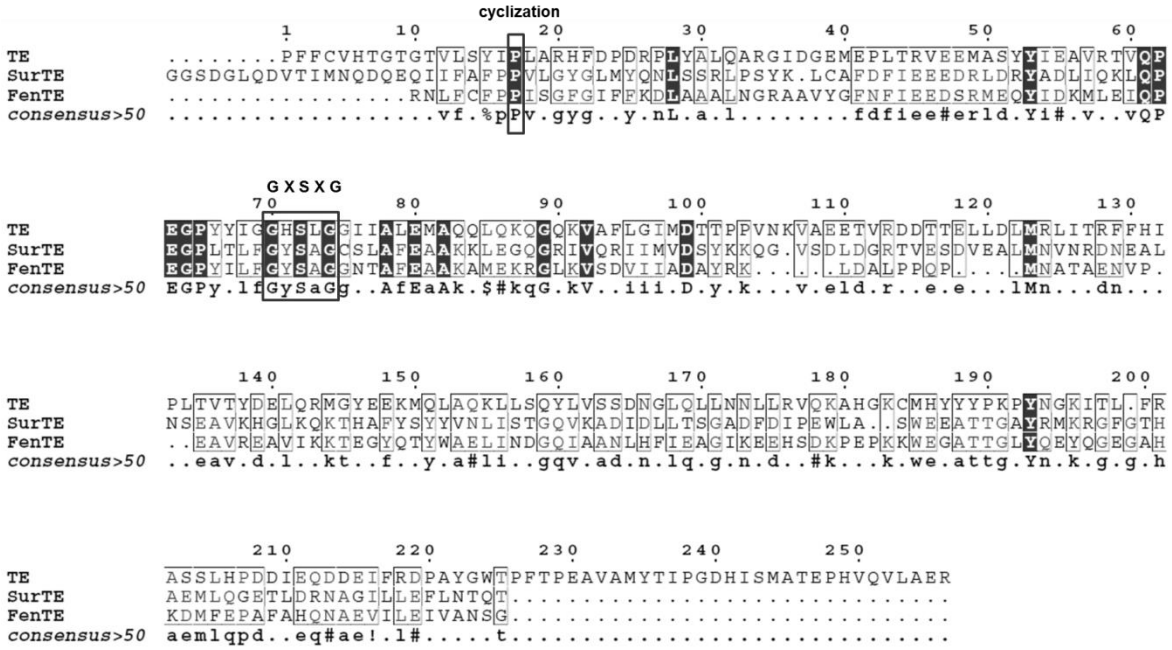


Figure 1-15 Amino acid sequence alignment of the TE domain in 17-7 and macrocyclization-catalyzing TE domains in the surfactin and fengycin biosynthetic clusters (Sur, Fen).

[Cluster 22]

Furthermore, cluster 22 (lantipeptide-bacteriocin) is highly homologous to the cinnamycin biosynthetic gene cluster (**Table 1-7**, **Figure 1-16**) (Widdick, Dodd et al. 2003, Okesli, Cooper et al. 2011). The locus includes the genes required for the synthesis of the precursor peptide (leader-core peptide) and posttranslational modification (CinX: hydroxylation of aspartate 15; CinM: dehydration of serine and threonine residues and cyclization of cysteine; and Cinorf7: the formation of a cross-link between lysine19 and dehydroalanine). Although the sequence of the predicted leader peptide of the putative precursor peptide (Thzk6460) varies, its core region exhibits very high sequence identity to the cinnamycin group (**Figure 1-17**), suggesting that this gene cluster possesses the genetic capacity to generate cinnamycin-like peptides. In detail, the four residues Val2, Ile10, Ala12, and Ile13 of the core peptide in Thzk6460 differ from those of the cinnamycin biosynthetic gene cluster (**Figure 1-17**). These four residues are most likely not involved in posttranslational modification positions for thioester linkage, hydroxylation and cyclization; therefore, the entire lantipeptide structure derived from cluster 22 was expected to be similar to cinnamycin. cluster 22 in *T. hazakensis* SK20-1^T is proposed to encode for the biosynthesis of a cinnamycin derivative, whose putative structure is shown in **Figure 1-17**.

Table 1-7 List of the genes in the lantipeptide biosynthetic gene locus, cluster 22.

# ORF	Length (bp)	Predicted function (% Id/Si)
6455	744	Predicted L-lactate dehydrogenase, YkgE
6456	1398	Predicted L-lactate dehydrogenase, YkgF
6457	657	Predicted L-lactate dehydrogenase, YkgG
6458	**906**	**CinX protein (45/64)**[a]
6459	**3165**	**CinM protein (52/66)**
6460	**261**	**Lantibiotic cinnamycin precursor *cinA* (54/68)**
6461	**393**	**Cinorf7 protein (45/63)**
6462	399	hypothetical protein
6463	372	hypothetical protein

[a] %Id/Si was calculated from corresponding protein of the cinnamycin biosynthetic pathway in *Streptomyces cinnamoneus cinnamoneus* DSM 40005.

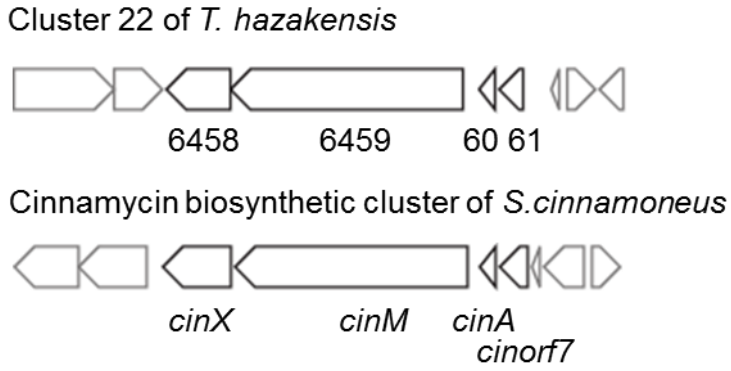


Figure 1-16 Gene organization of cluster 22 and the cinnamycin biosynthetic cluster.

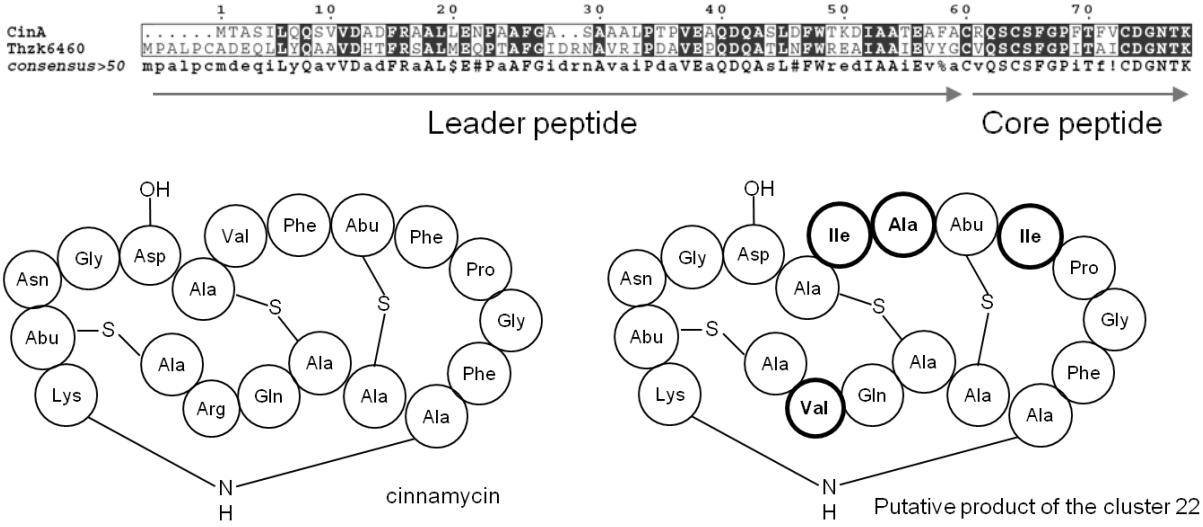


Figure 1-17 Sequence alignment of the precursor peptide between CinA and Thzk6460 (top) and the cinnamycin structure and predicted structure synthesized by cluster 22 (bottom).

1.5. Summary

During the last three decades, numerous biosynthetic pathways of several secondary metabolites have been elucidated primarily from microorganisms. This knowledge of the biosynthesis of secondary metabolites adds value to microbial genome sequences. In the field of natural products chemistry, various novel compounds have been identified using genome-guided discovery and isolation. Therefore, this new methodology to expand chemical diversity would be applicable to as of yet unexplored organisms as well as well-studied microorganisms such as terrestrial actinomycetes.

T. hazakensis SK20-1^{T} appears to exhibit the potential to produce several secondary metabolites based on its morphological characteristic of multiple exospore formation and a complex life cycle. To verify the genetic potential of *T. hazakensis* for the production of secondary metabolites, I subjected the *T. hazakensis* genomic DNA to draft sequencing using an Illumina DNA sequencer and analyzed the genome sequence using various bioinformatics tools. The 7.3-Mb genome is sufficiently large to afford the biosynthesis of diverse secondary metabolites. Annotation of the coding sequences revealed that numerous genes occupying two-third of all genes were hypothetical. Next, I performed an *in silico* investigation of the gene loci that may be involved in the biosynthesis of secondary metabolites. Analysis of the genome sequences indicated the presence of 23 gene clusters including hybrid NRPS/PKS and ribosomal peptide synthase. Moreover, additional analysis of each putative gene function predicted corresponding plausible chemical products.

For example, the hybrid PKS/NRPS in “cluster 17” was proposed to synthesize a lipohexapeptide compound. Functional analysis of “cluster 10” suggested a siderophore compound possessing four thiazoline heterocycles as a putative biosynthetic product. Additionally, “cluster 22” exhibits a gene organization that is very similar to that of the lantipeptide cinnamycin biosynthetic gene cluster. The core peptide sequence suggested “cluster 22” encodes the biosynthetic genes responsible for a cinnamycin-like lantipeptide with a molecular mass of 1,888 Da.

However, I was unable to predict secondary metabolites biosynthesized by other gene loci because of irregular gene organization or insufficient conserved domains to assume biosynthetic precursors. Nevertheless, these irregularities suggested a high probability to produce novel chemical skeletons.

2. Isolation, structure elucidation, and biological activity of the secondary metabolites from *T. hazakensis* SK20-1^T

2.1. Chemical screening using small-scale culture conditions

2.1.1. Methods

T. hazakensis SK20-1^T was cultured to screen for the production of secondary metabolites on a small scale. To explore the metabolic profile of *T. hazakensis* SK20-1^T, different culture media were used, which included K, Q, A1, and ISP1 media. In all conditions, *T. hazakensis* SK20-1^T grown on ISP1 agar at 55℃ for 7 days was used to inoculate 10 mL of liquid preculture. The preculture, which was grown at 55℃ and 140 rpm for 3 days, was used as inoculums for 150-mL cultures in 500-mL Sakaguchi flasks. The primary culture was incubated for 7 days prior to extraction.

After incubation, each culture broth was centrifuged at 5000 rpm for 10 min to separate the mycelium and supernatant. Subsequently, each supernatant was fractionated using two volumes of ethyl acetate, and each mycelium was extracted using 70 % acetone. Each organic fraction was dried in vacuo and dissolved in methanol. Prepared samples were injected into an analytical HPLC equipped with a PDA detector using an aqueous acetonitrile gradient (10–90 %) as the mobile phase. These samples were also analyzed using LC-ESI-MS. LC-MS measurements were conducted using the samples dissolved in methanol (approximately 1 mg/mL) for injection into the LC-MS system. Experiments were recorded in positive mode and analyzed using the PeakView software from AB SCIEX.

Additionally, the genome scanning described in the previous chapter indicated the presence of putative lipohexapeptide and cinnamycin-like lantipeptide biosynthetic gene loci. Thus, peptide-rich extracts were prepared using three methods (**Table 2-1**): solvent extraction (Widdick, Dodd et al. 2003), acetone precipitation (Zendo, Nakayama et al. 2008) and acid precipitation (Romano, Vitullo et al. 2013).

Table 2-1 Extraction methods to prepare peptide-rich extracts.

Solvent extraction
Washing culture broth with EtOAc
Extraction with BuOH
Extraction with 10 % formic acid
Evaporation
Acetone precipitation
Addition iced acetone into culture broth
Precipitation at -20℃ for 3 hours
Harvesting by centrifugation at 5,000 rpm for 20 minutes
Dissolving pellets in 15 % MeCN containing 0.1 % TFA
Acid precipitation
Acidification of supernatant to pH 2.0 with 6N HCl
Incubation at 4℃ for overnight for precipitation
Collecting the precipitant by centrifugation at 10,000 rpm for 20 minutes
Dissolving pellets in $CHCl_3$-MeOH mixuture (v/v, 2:1)
Polytetrafluoroethylene (PTFE) filtration
Removal of organic solvent under vacuum

Peptide-rich extracts prepared likewise were analyzed by LC-ESI-MS using a gradient (15 to 75 %) of aqueous acetonitrile containing 0.1 % formic acid.

To examine the production of the phenolate-type siderophore predicted by analysis of NRPS/PKS cluster 10, I cultured *T. hazakensis* SK20-1^T under an iron-deficient condition by adding 2,2'-dipyridyl (100 μM of final concentration), which is a non-assimilable and nontoxic iron-chelating reagent, to several different culture media including minimal media (CM9, NMMP), TSB, and ISP1. After 5 days of culture, 5 mg of $FeCl_3$ was added to 100 ml of each supernatant, and the cultures were gently stirred for 10 min. After incubation at 4 ℃ overnight, XAD-7HP resin was used to extract the Fe(III) complex.

[Antibacterial susceptibility test]

The bacterial strain was incubated in 5 mL of LB medium in a test tube for 1 day at 37℃, and an MH (Mueller-Hinton, Difco) agar plate was prepared. The next day, 1 % of the bacteria was added to 5 mL of melted soft top agar (0.5 % agar). The soft agar was subsequently mixed using a vortex and then immediately poured onto MH agar. Samples dissolved in MeOH were applied on 6-mm sterile filter discs with MeOH and ampicillin. After the MeOH had evaporated, each disc was transferred onto the agar plate. After overnight incubation on 37℃, antibacterial activity was determined by the sample clear zone in comparison to that of ampicillin.

2.1.2. Results and Discussion

T. hazakensis SK20-1^{T} was cultured on four different media (K, Q, A1, and ISP1) to compare the corresponding metabolic profiles. After culture at 55℃, growth using the Q medium was poor in comparison to the other medium conditions; therefore, the extract from the Q medium was not prepared, and other samples were analyzed using analytical HPLC (**Figure 2-1**). Following HPLC analysis, the chromatogram of the Al-medium extract exhibited simple peaks in comparison to those of the K and ISP1 extracts. According to HPLC analysis, the major product exhibiting an Rt of 18.6 min was detected in all samples, and another distinct peak exhibiting an Rt of 15.7 min was detected in the sample using soybean-based medium (K medium). Although various peaks were observed in the K medium extract, the UV absorbance spectrum of several major peaks was very similar to soybean-derived isoflavones such as genistein and daidzein. These soybean-derived components may hinder the isolation of other compounds; therefore, I chose ISP1 as a fermentation medium for subsequent large-scale culture. Additionally, a comparison of 2-day and 5-day cultures on ISP1 confirmed a substantial enhancement of secondary metabolite production (**Figure 2-2**).

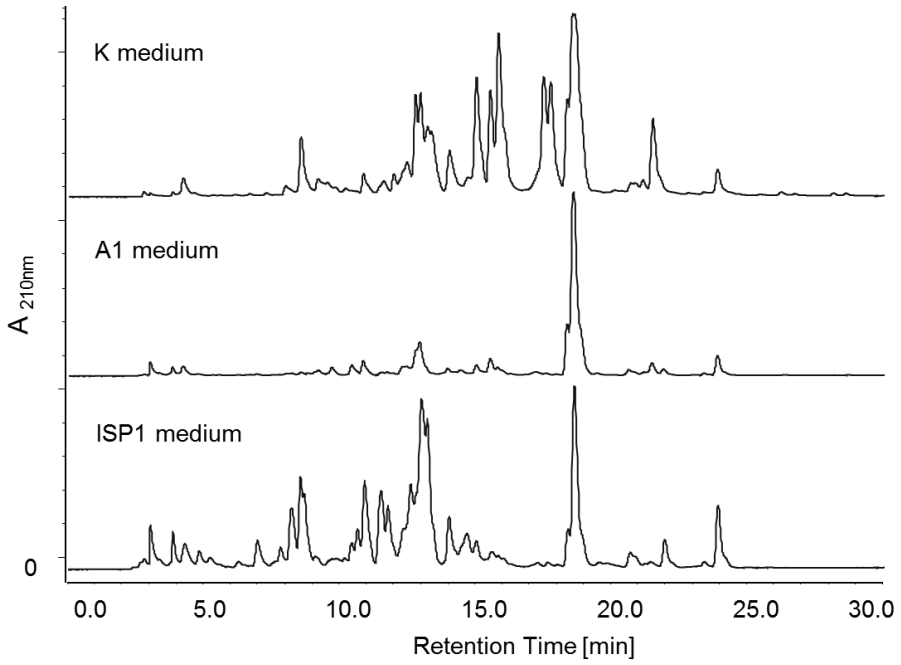


Figure 2-1 Secondary metabolite HPLC profiles according to the culture media.

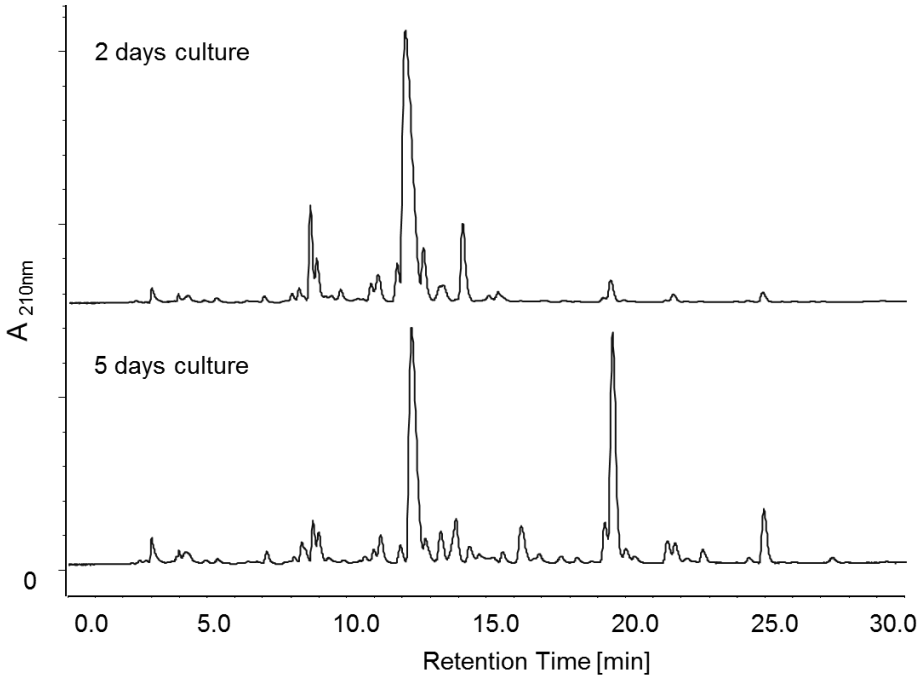


Figure 2-2 Comparison of secondary metabolite profiles according to the culture period using ISP1 medium.

In addition to the aforementioned HPLC analysis, I analyzed each crude extract using LC-MS. LC-MS analysis indicated that most constituents in the crude extract exhibited ion peaks from *m/z* 100 to 300 in positive ion mode, which indicates small molecules that are not related to the NRPS-PKS hybrid compounds predicted using genome scanning. Peptide-rich extracts also did not exhibit specific ion peaks of a cinnamycin-like lantipeptide (the predicted molecular mass, 1888.1 Da) or lipohexapeptide in the LC-MS analysis. Generally, lantipeptides and lipopeptides are known to possess significant antibacterial activities by forming pores in the cytoplasmic membrane of the pathogen. To determine the presence of the target compounds in specific extraction samples, I also performed an antibacterial test against Gram-positive *Micrococcus luteus* and Gram-negative *E. coli* JM109 Δ*tolC*. The *E. coli* JM109 Δ*tolC* strain is antibiotic sensitive via the disruption of the *tolC* gene that encodes a multidrug efflux pump. According to the antibacterial test, each sample did not exhibit any biological activity, indicating that peptide compounds predicted by the genome analysis were not produced under the investigated culture condition (**Figure 2-3**).

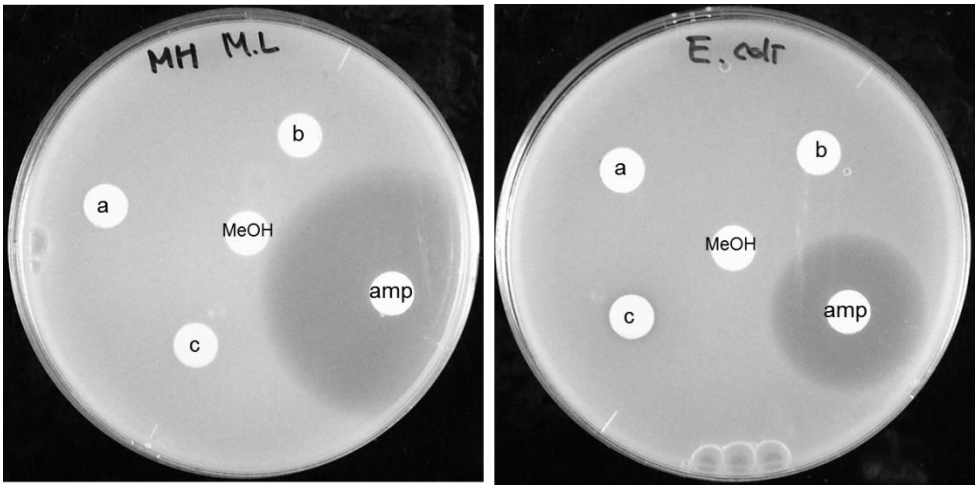


Figure 2-3 Antibacterial test against *M. luteus* (left) and *E. coli* Δ*tolC* (right):
a, solvent extraction; b, acetone precipitation; and c, acid precipitation. MeOH, methanol used to dissolve the samples (a, b, c); amp, ampicillin (100 μg/disc).

The culture grown in iron-limited media did not produce the predicted siderophore compounds. For minimal media, I was unable to obtain a sufficient amount of cells due to

poor growth of *T. hazakensis* SK20-1^T. In contrast, culture extracts of TSB and ISP1 containing 2,2'-dipyridyl exhibited identical HPLC chromatograms to those of culture extracts without 2,2'-dipyridyl. The addition of $FeCl_3$ to the culture media also did not exhibit an Fe(III)-complex in HPLC and LC-MS analyses, suggesting that the targeted siderophore compound was not successfully produced under the iron-limited condition.

The initial chemical screening of *T. hazakensis* SK20-1^T did not reveal the products of the biosynthetic pathways predicted by genome mining, most likely because the amount produced was very low. Therefore, these results suggest that additional studies are required to activate the identified gene clusters to produce sufficient quantities to isolate the compounds. Generally, a bacterial genome harbors diverse biosynthetic gene clusters that are expected to be involved in secondary metabolite production. Nevertheless, most biosynthetic gene clusters are known to not be expressed under normal laboratory conditions. To avoid and overcome the challenge of silent gene clusters, several strategies are available, from traditional methods using various culture conditions to induce the expression of the desired biosynthetic pathway to a synthetic biology approach such as in vitro reconstitution of genes and their expression in a heterologous host system. Activation of the regulatory genes that activate the transcription of the biosynthetic gene cluster has also been effective to activate silent gene clusters.

Moreover, the cause of the lack of production by the analyzed gene cluster was investigated. In particular, I speculate that an absence of a proteolytic site to form the core peptide in the precursor peptide may underlie the lack of production of the expected cinnamycin-like peptide from cluster 22. For cinnamycin biosynthesis, an AXA motif is present between the leader sequence and the core region in CinA. This motif is recognized by type I signal peptidases of the general secretory (*sec*) pathway to yield a cleaved core peptide for the production of the final cinnamycin product. However, the precursor peptide of "cluster 22" lacks this motif as well as the GG or GA protease cleavage motif observed in most class II lantibiotics. Thus, the precursor in cluster 22 would not be appropriately cleaved to yield the core peptide. I assume that the production of the predicted peptide may be achieved by the introduction of an artificial cleavage site using site-directed mutagenesis and heterologous expression using a host that harbors the appropriate endoproteinase.

2.2. Fermentation, isolation and purification for new molecules from *T. hazakensis* SK20-1^T

2.2.1. Fermentation

The SK20-1^T strain was grown on ISP1 medium at 50 ℃ for 7 days and then cultured as a seed culture in ISP1 medium in a 500-ml Sakaguchi flask for 3 days. The seed culture was transferred to a two 5-L bioreactor (Bioneer C500, B.E. Marubishi, Tokyo) containing 3 L of fermentation medium consisting of glucose (0.4%), malt broth (1%), and yeast extract (0.4%) and was cultured for 7 days at 55 ℃ with stirring at 300 rpm. Aeration was maintained at 1.0 SLPM (standardized liter per minute).

2.2.2. Preparation of the crude extract

After 7 days of fermentation, the total culture broth was centrifuged at 5000 rpm for 10 minutes to separate the mycelium and supernatant. The supernatant was extracted twice using an identical volume of ethyl acetate, and the ethyl acetate extract was combined with the acetone-extracted mycelium extract. The remaining water fraction was extracted using an identical volume of saturated butanol to yield the butanol extract.

2.2.3. Fractionation of the EtOAc and BuOH extracts

The EtOAc extract (1.2 g) was fractionated by DIAION® HP-20 (Mitsubishi Chemical, Tokyo) flash column chromatography using mixtures of acetone-water (20, 40, 60, 80, and 100% acetone in water, each 100 mL) as elution solvents (**Figure 2-4**). The butanol extract (0.5 g) was also fractionated by HP-20 flash column chromatography using mixtures of methanol-water (20, 40, 60, 80, and 100 %).

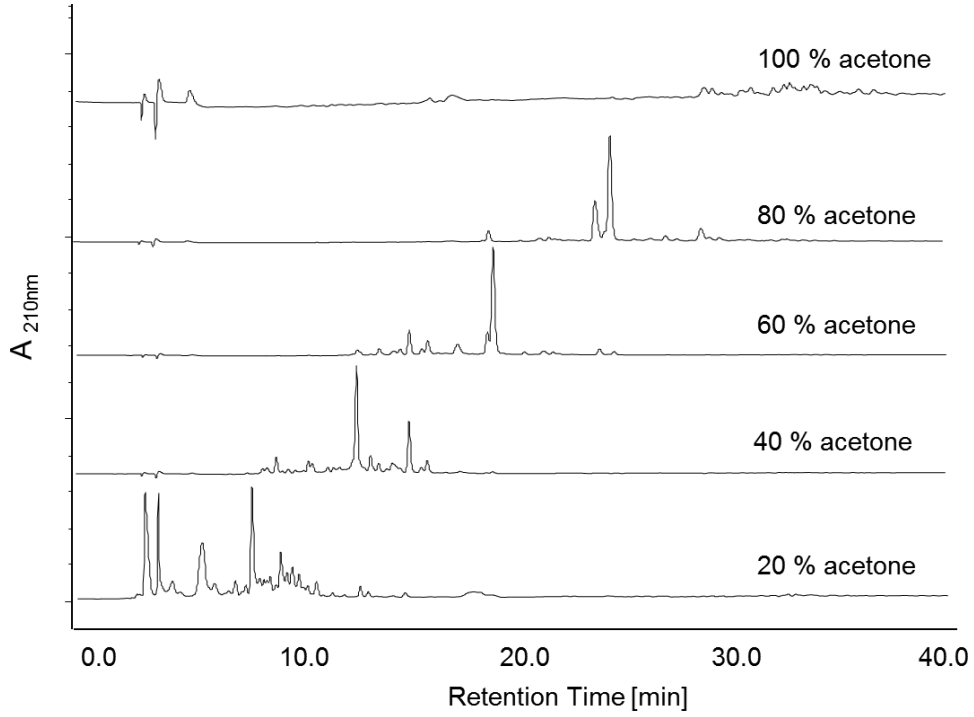


Figure 2-4 HPLC analysis of the fractionation of the EtOAc extract using HP-20 chromatography.

2.2.4. Isolation and purification of secondary metabolites

The overall procedure to isolate and purify secondary metabolites from *T. hazakensis* cultures is illustrated in **Figure 2-5**. Five new compounds (**1**-**5**, **Figure 2-6**) were obtained using preparative HPLC; the 60% (270 mg) and 80% (210 mg) aqueous acetone fractions were separated using isocratic elution of 45% and 50% aqueous acetonitrile, respectively. These purification steps yielded compound **2** (7 mg, **Figure 2-8**) from the 60% acetone fraction and compound **1** (40 mg, **Figure 2-7**), compound **4** (1 mg, **Figure 2-10**), and compound **5** (0.1 mg) from the 80% acetone fraction. Next, a 100% methanol extraction (70 mg) of the BuOH extract was also applied to preparative HPLC to isolate a highly-conjugated resonance structure that exhibited a distinct 350 nm absorption. Preparative HPLC was conducted using 60 % methanol containing 0.1 % TFA, yielding compound **3** (5 mg, **Figure 2-9**).

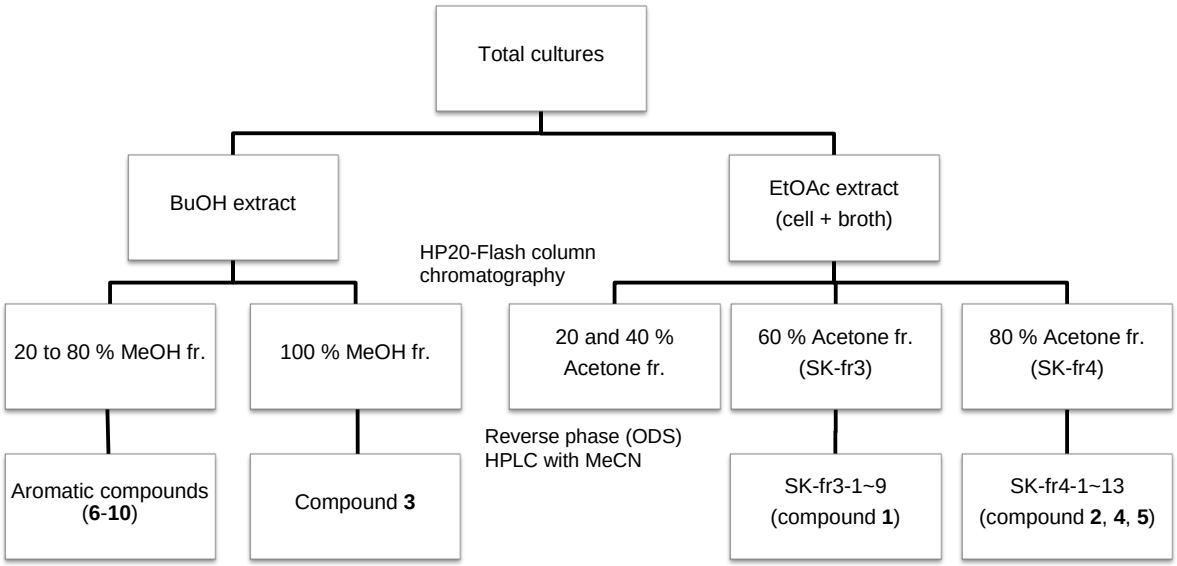


Figure 2-5 Flowchart of the fractionation and purification to yield new compounds **1**-**5**.

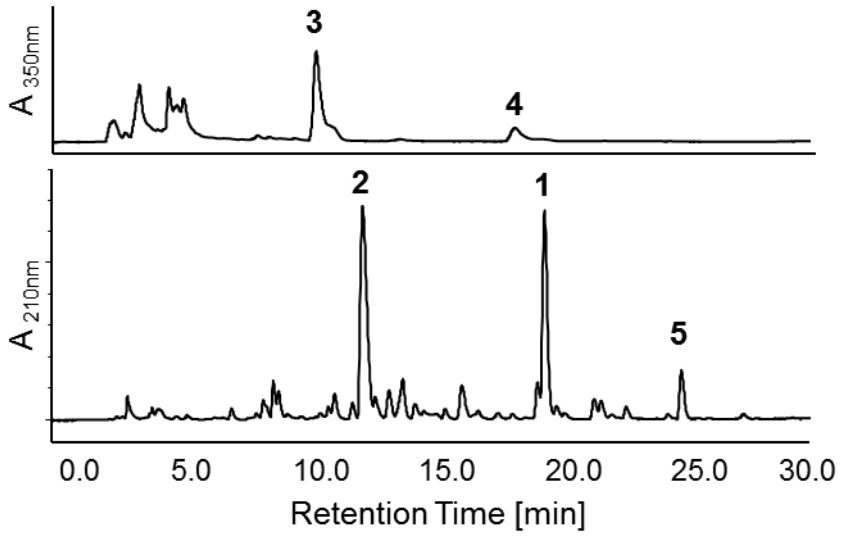


Figure 2-6 Analytical HPLC chromatogram indicating five new compounds (**1**-**5**) in the total culture extract.

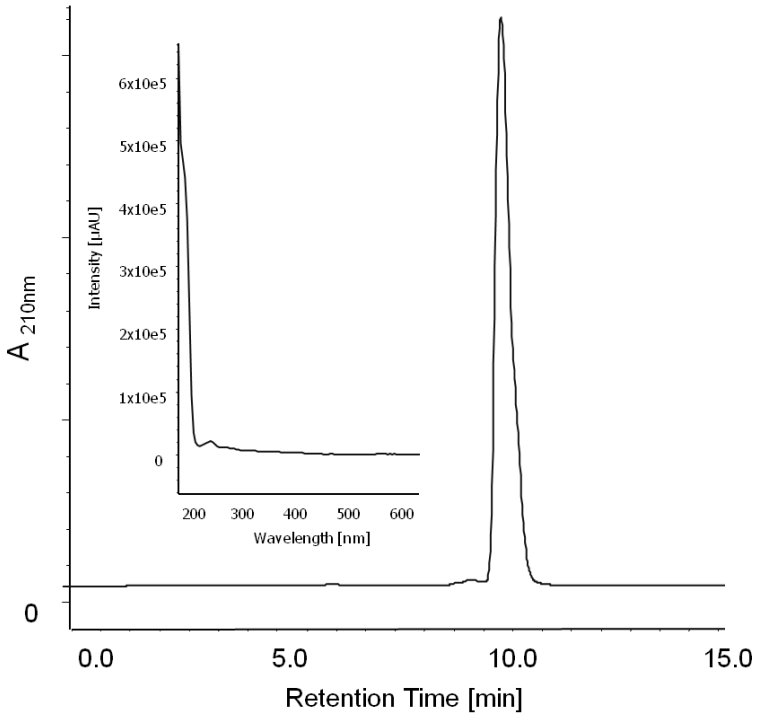


Figure 2-7 HPLC chromatogram of purified compound **1** ((-)-sattabacin) and its UV/Vis absorption spectrum (inset).

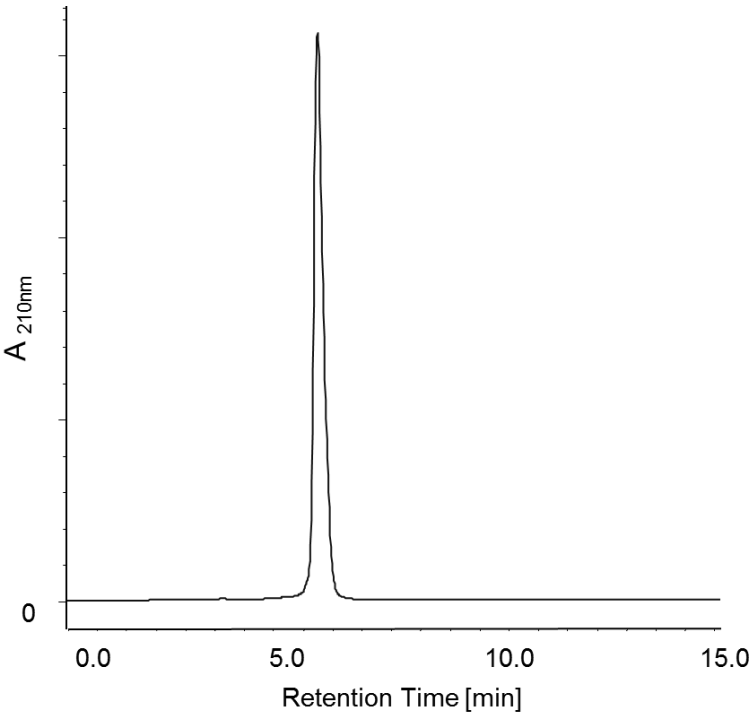


Figure 2-8 HPLC chromatogram of purified compound **2** ((-)-hazakacin).

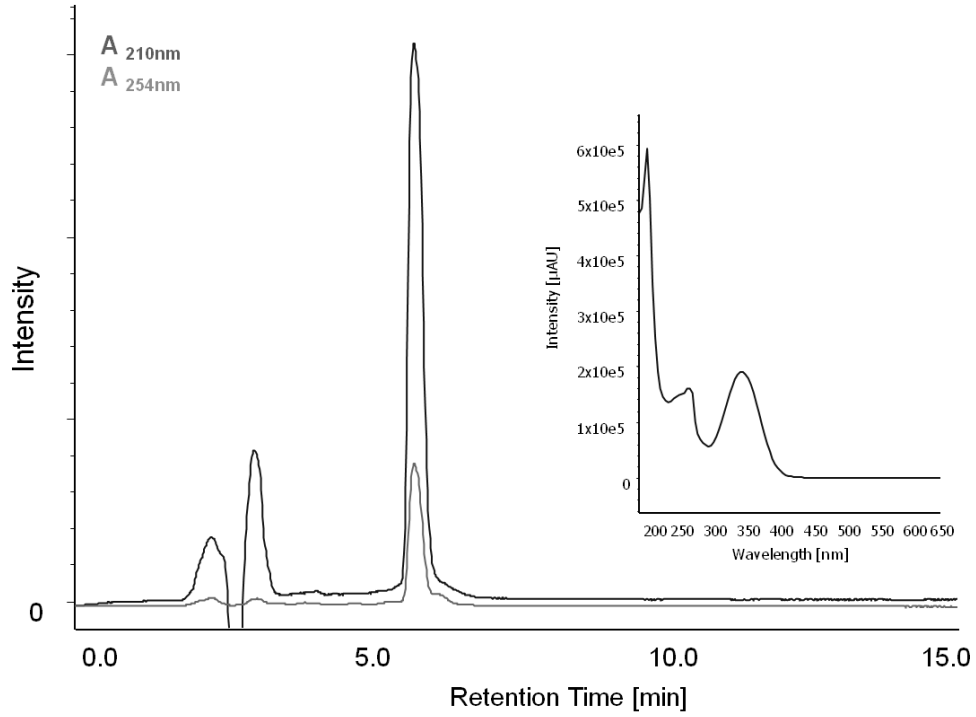


Figure 2-9 HPLC chromatogram of purified compound **3** and its UV/Vis absorption spectrum (inset).

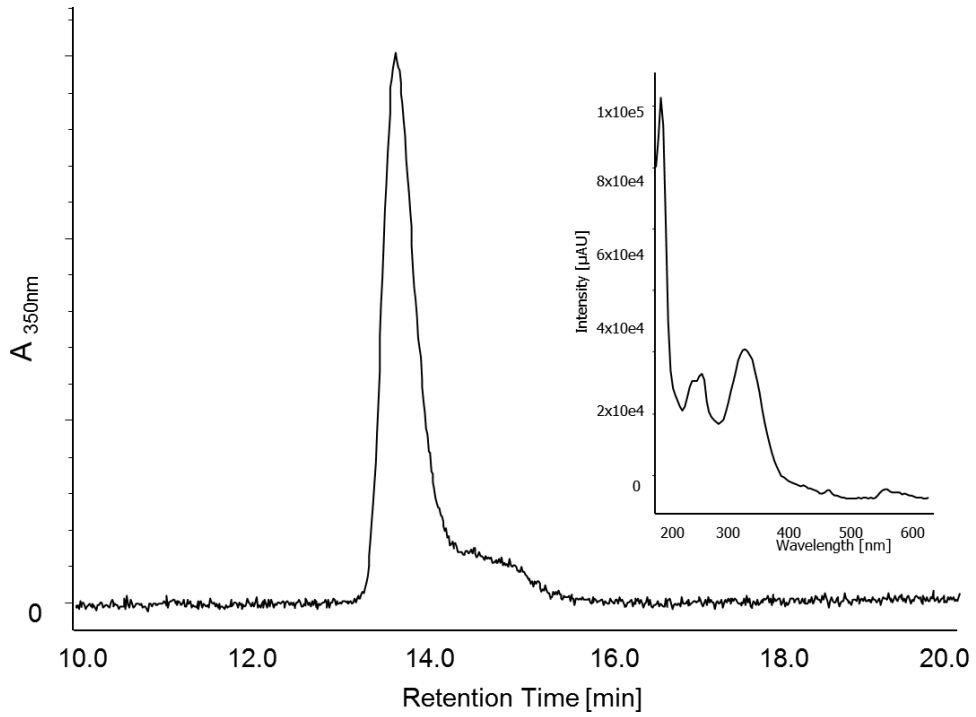


Figure 2-10 HPLC chromatogram of purified compound **4** and its UV/Vis absorption spectrum (inset).

2.3. Structure elucidation of the isolated compounds

2.3.1. Method

2.3.1.1. NMR spectroscopy

All NMR analyses of the extracts were performed using a JEOL ECA-600 spectrometer operating at 600 MHz for ^{1}H and 150 MHz for ^{13}C. The spectra were processed using Delta Version 5.0.3 and MestReNova Version 6.0.2 software. Deuterated solvents, which were purchased from Cambridge Isotope Laboratories (CIL), were used to suspend the fractions for NMR measurement.

2.3.1.2. Mass spectrometry

A high-resolution Triple TOF® 5600 MS (AB SCIEX, Tokyo) was used with a CAPCELL PAK C18 IF column (2.0 × 50 mm; Shiseido, Tokyo). MS analysis was performed using electrospray ionization (ESI) in both positive and negative ion modes.

2.3.1.3. UV measurements

UV spectra were recorded on a Shimadzu UV-2500PC UV-Vis spectrophotometer (Tokyo) using 1.0 cm quartz cells.

2.3.1.4. Measurement of optical rotation

The optical rotations were recorded using a DIP-1000 polarimeter (JASCO, Tokyo).

2.3.2. Result and Discussion

2.3.2.1. (-)-Sattabacin (**1**)

Compound **1** was isolated as a pale yellow oil with the molecular formula $C_{13}H_{18}O_2$, as determined by positive high-resolution mass spectrometry (*m/z* 207.1381 $[M+H]^+$ and *m/z* 189.1275 $[M-H_2O+H]^+$, calculated for $C_{13}H_{19}O_2$ and $C_{13}H_{17}O$: 207.1385 and 189.1279, respectively), indicating five degrees of unsaturation. In the ^1H and ^{13}C NMR spectra (**Table 2-2**), I recognized the presence of a phenyl moiety (5H, δ_H 7.20-7.30), an oxygen-bearing methine (1H, δ_H 4.36), and two magnetically equivalent methyl groups (6H, δ_H 0.91). Two partial structures, from the 1-substituted phenyl to C2' and from C4' to C7', were constructed by interpreting of the chemical shifts, coupling constants, and ^1H-^1H COSY correlations. The two partial structures were then connected via a ketone (C-3') to account for the remaining double bond, which was supported by the HMBC correlations of H-2' (δ_H 4.36) and H-4' (δ_H 2.36) to C-3' (δ_C 211.3). Thus, the planar structure of **1** was established as 2-hydroxy-5-methyl-1-phenylhexan-3-one, or sattabacin, which was previously isolated from *Bacillus* sp.(Lampis, Deidda et al. 1995). However, the optical rotation of 1 ($[\alpha]_D^{25}$ −36 (*c* 0.4, $CHCl_3$)) was opposite that of synthetic (+)-sattabacin ($[\alpha]_D^{25}$ +41 (*c* 0.1 CH_2Cl_2) (Aronoff, Bourjaily et al. 2010). Therefore, **1** was determined to be (−)-sattabacin, and C-2' was assigned the *R* configuration (**Figure 2-11**).

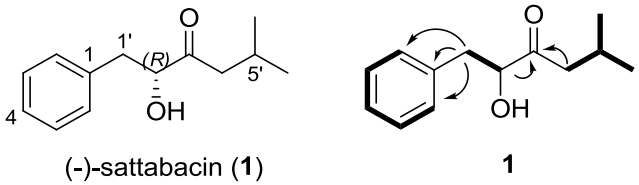
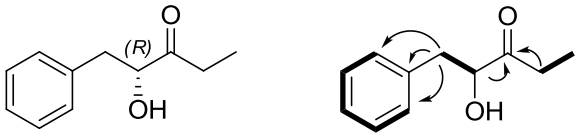


Figure 2-11 Structure of (-)-sattabacin (**1**) and ^1H-^1H COSY correlations (bold lines) and HMBC correlations (arrows) of **1**.

(−)-sattabacin (**1**): pale yellow oil; ^1H NMR and ^{13}C NMR (**Table 2-2**, **Figure A1**, **Figure A2**); HR-ESI-MS (*m/z*): $[M+H]^+$ calcd for $C_{13}H_{19}O_2$, 207.1380; found, 207.1381 (**Figure 2-13**); $[\alpha]_D^{25}$ −36 (*c* 0.4, $CHCl_3$).

2.3.2.2. (-)-Hazakacin (**2**)

Compound **2** was obtained as a pale yellow oil. The molecular formula was assigned to be $C_{11}H_{14}O_2$ as determined by positive high-resolution mass spectrometry (*m/z* 179.1066 $[M+H]^+$; calculated for $C_{11}H_{15}O_2$: 179.1072), indicating that **2** possessed five degrees of unsaturation. The ^1H and ^{13}C NMR spectra (**Table 2-2**) of **2** exhibited a resonance pattern similar to that of **1**, showing signals for a methyl (triplet, 3H, δ_H 0.95), an oxygen-bearing methine (1H, δ_H 4.28), four aliphatic methylene protons (4H, δ_H 3.01, 2.78, 2.53, 2.47), and five aromatic protons (5H, δ_H 7.15-7.25). In particular, the signal for the protons from the phenyl moiety of **2** was nearly identical to that of 1. However, **2** exhibited a signal for only one methyl proton in contrast to the two methyl group signals observed in **1**. One methylene proton signal (δ_H 2.53) did not exhibit any COSY correlation with the methyl proton, suggesting that this methylene was connected to the quaternary carbon. HMBC correlation of the oxygen-bearing methine proton (δ_H 4.28) and an aliphatic methylene proton (δ_H 2.53) to a quaternary ketone carbon (δ_C 213.2) established that **2** was 2-hydroxy-1-phenyl-3-pentanone. Previously, **2** was chemically synthesized by the tandem alkylation and reduction of phenylacetaldehyde, but it has not been found in a natural source (Weiberth and Hall 1985). The absolute configuration of C-2' was determined by optical rotation ($[\alpha]_D^{25}$ −28 (*c* 0.5, $CHCl_3$)), which was identical to the negative optical rotation of (−) sattabacin (**1**). I thus assigned the stereochemistry of the C-2' of **2** as *R* and named **2** "(-)-hazakacin" (**Figure 2-12**).



(-)-hazakacin (**2**) **2**

Figure 2-12 Structure of (-)-hazakacin (**2**) and ^1H-^1H COSY correlations (bold lines) and HMBC correlations (arrows) of **2**

(−)-hazakacin (**2**): pale yellow oil; ^1H NMR and ^{13}C NMR (**Table 2-2**, **Figure A3**, **Figure A4**); HR-ESI-MS (*m/z*): $[M+H]^+$ calcd for $C_{11}H_{14}O_2$, 179.1067 (**Figure 2-13**); found, 179.1066; $[\alpha]_D^{25}$ −28 (*c* 0.5, $CHCl_3$).

Table 2-2 ^{1}H (600 MHz) and ^{13}C (150 MHz) NMR data of **1** and **2**

#	Compound **1**[a]		Compound **2**[b]	
	δ_H, mult (*J* in Hz)	δ_C	δ_H, mult (*J* in Hz)	δ_C
1	-	136.8	-	136.8
2	7.24 m	129.4	7.21 d (6.96)	128.4
3	7.29 dd (7.2, 7.2)	128.6	7.24 dd (7.50, 7.14)	127.2
4	7.24 m	127.0	7.17 dd (6.96, 7.20)	125.4
5	7.29 dd (7.2, 7.2)	128.6	7.24 dd (7.50, 7.14)	127.2
6	7.24 m	129.4	7.21 d (6.96)	128.4
1′	3.12 dd (14.4, 4.2) 2.82 dd (14.4, 7.2)	40.1	3.01 dd (14.0, 4.6) 2.78 dd (14.0, 8.0)	39.0
2′	4.36 dd (7.2, 4.2)	77.6	4.28 dd (4.6, 8.0)	76.7
3′	-	211.3	-	213.2
4′	2.36 d (6.6)	47.5	2.53 qd (7.3, 18.5) 2.47 qd (7.3, 18.5)	30.6
5′	2.17 dq (6.6, 6.6)	24.7	0.95 t (7.3)	5.5
6′	0.91 dd (6.6, 1.2)	22.7		
7′	0.91 dd (6.6, 1.2)	22.7		

The chemical shifts were recorded in $CDCl_3$[a] and CD_3OD[b].

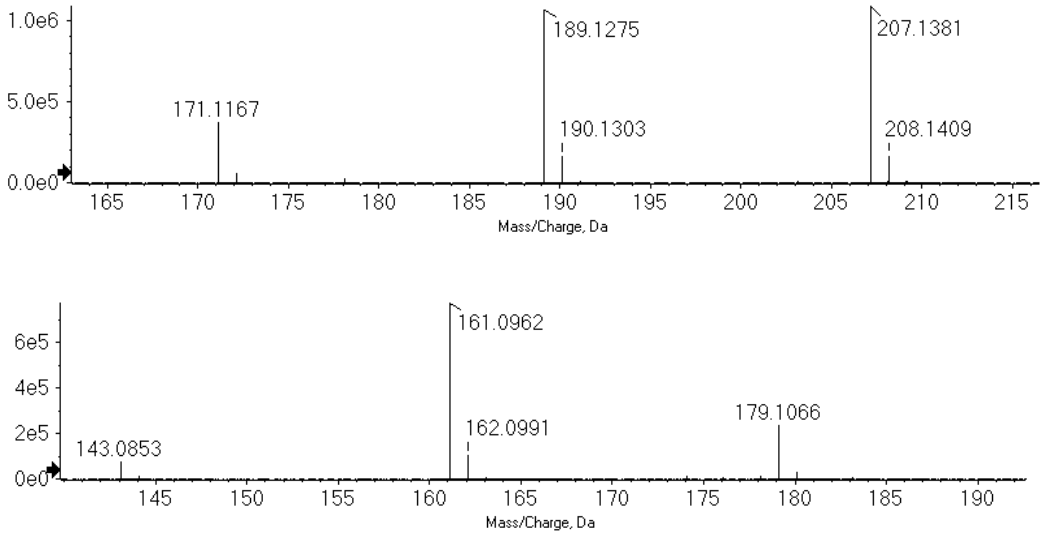


Figure 2-13　HR-ESI(+)-MS spectra of compounds **1** (top) and **2** (bottom).

2.3.2.3. インターネット公表出来ないために削除

2.3.2.4. インターネット公表出来ないために削除

2.3.2.5. インターネット公表出来ないために削除

2.3.2.6. インターネット公表出来ないために削除

2.3.2.7. インターネット公表出来ないために削除

2.4. Biological activity of the isolated compounds

2.4.1. Methods

The bioactivity assay was kindly conducted by Dr. Noritaka Kagaya (Technology Research Association for Next-Generation Natural Products Chemistry) for the assessment of cytotoxicity, by Dr. Junko Hashimoto (Japan Biological Informatics Consortium (JBIC)) for the assessment of antifungal activity, and by Dr. Miho Izumikawa (JBIC) for the assessment of antibacterial activity.

Antimicrobial activity against *C. albicans* NBRC 1594 was assessed by measuring the inhibition of fungal growth. After incubation of *C. albicans* for 1 day, isolated compounds (**1**, **2**, **3**, and **4**) were added to the fungal culture, which had been diluted to $1x10^{-2}$. *C. albicans* was subsequently incubated for 48 hours at 35 ℃. The inhibition of growth was determined by measuring the absorbance at 620 nm, and siccanin was used as a positive control in this assay.

The antimicrobial activity against *M. luteus* ATCC9341 was measured by the paper disk method. A paper disk (diameter, 6 mm) containing 10 μg of a sample was placed on an LB agar plate, and the plate was incubated at 27 °C for 24 h. The cytotoxicity of samples in a human ovarian carcinoma SKOV3 cell line, a mesothelioma Meso-1 cell line, and a T lymphoma Jurkat cell line was determined by colorimetric assay using WST-8 [2-(2-methoxy-4-nitrophenyl)-3-(4-nitrophenyl)-5-(2,4-disulfophenyl)-2H-tetrazolium, monosodium salt] following a detailed procedure that has been previously described (Khan, Komaki et al. 2011, Kawahara, Hwang et al. 2012).

2.4.1.1. Results and Discussion

インターネット公表出来ないために削除

2.5. Summary

インターネット公表出来ないために削除

3. Biosynthesis of new compounds isolated from *T. hazakensis* SK20-1^T

3.1. Biosynthesis of (-)-sattabacin (**1**) and (-)-hazakacin (**2**)

3.1.1. Materials and Methods

3.1.1.1. Cloning and overexpression of ScyA homologs

The ScyA homologs were PCR amplified from genomic DNA using the following primer sets summarized in **Table 3-1**.

Table 3-1 Primer sets used to clone the ScyA homologs

Gene locus	Primer	Sequence (5' to 3')*	Restriction enzyme
thzk0150	0150F	cgtatggatcctcgacatattcagctatag	*Bam*HI
	0150R	gttgttaagcttttaatgggccggttg	*Hin*dIII
thzk1026	1026F	gttacggatccgctcaagtcagcgattttc	*Bam*HI
	1026R	aacctaagcttctacttcttcgttggcagg	*Hin*dIII
thzk3750	3750F	gctgaggatccgcaatcatgcacgggggct	*Bam*HI
	3750R	cgaacgtcgacctatacagcgacttgtgtt	*Hin*dIII
thzk5630	5630F	agagtgggatccactcgtatgtctgcaactg	*Bam*HI
	5630R	gctccaagctttcaggaaggaaccgttgct	*Sal*I

* The sequence sites for restriction enzymes were underlined.

The amplified DNA fragments were digested with the previously mentioned restriction enzymes and ligated with the expression vector pHis-8, which was digested with the corresponding enzymes, using Ligation High (Toyobo, Osaka). Subsequently, the resultant plasmid was transformed into *E. coli* DH5α. The sequence-confirmed plasmid was then transformed into *E. coli* BL21(DE3). The resultant transformant was cultured in 1 L of TB medium containing 50 μg/mL of kanamycin at 37 ℃ until the OD_{600} reached 0.5. The cells were cooled on ice for 10 minutes and incubated at 18 ℃ for 20 hours following

induction with 0.1 mM IPTG. The cells were harvested by centrifugation and resuspended in 100 mL of lysis buffer (50 mM Tris-HCl, pH 8.0, 100 mM NaCl, 20% glycerol, 20 mM imidazole, and 1% Tween 20). The cell suspensions were then sonicated using a Branson Sonifier 250 (Emerson Japan, Tokyo). To separate the cellular debris from the soluble protein, the lysate was centrifuged at 34,700 ×*g* for 20 minutes at 4 ℃. Purification using Ni-NTA chromatography (GE Healthcare Life Sciences) was performed according to the manufacturer's instructions, and the target protein was eluted with 250 mM imidazole in 50 mM Tris-HCl (pH 8.0), 100 mM NaCl, and 20% glycerol. The eluted protein was concentrated by ultrafiltration using Vivaspin 20 centrifugal filters (10,000 molecular weight cut-off; Sartorius Stedim Biotech).

3.1.1.2. Enzymatic assay of ScyA homologs

The ScyA homolog assay was performed in 50 mM HEPES-NaOH (pH 8.0) containing 3 mM $MgCl_2$, 1.5 mM substrate, 1 mM TPP, and 200 μg/ml enzyme. The reaction mixture was incubated at 30 ℃ for 20 hours. After incubation, the reaction mixture was quenched by an identical volume of methanol, and the precipitate was removed by centrifugation. The reaction products were analyzed by LC-ESI-MS using a gradient solvent system consisting of 10-90% aqueous acetonitrile with 0.1% formic acid over 5 minutes. To investigate the optimal temperature for activity, the reaction solution was incubated from 25 ℃ to 55 ℃ for 1 hour.

3.1.1.3. Molecular mass estimation of Thzk0150

The molecular mass of Thzk0150 was estimated by gel-filtration chromatography (GFC) using a Superdex 200 column (GE Healthcare) that was equilibrated with 20 mM Tris-HCl (pH 8.0), 150 mM NaCl, and 10 % glycerol (v/v) at a flow rate of 0.7 mL/min. The molecular mass was estimated by calibration using marker proteins as follows: ferritin (440 kDa), conalbumin (75 kDa), and ovalbumin (43 kDa).

3.1.2. Results and Discussion

3.1.2.1. Literature investigation

I initiated biosynthetic studies on (−)-sattabacin (**1**) and the novel natural acyloin (**2**, (-)-hazakacin). In the report that described the initial steps in the biosynthesis of the UV-blocking compound scytonemin, which is produced by the cyanobacterium *Nostoc punctiforme* ATCC 29133, an earlier acyloin precursor was presumed to be produced through the formation of a C−C bond between two amino acid-derived α-keto acids. During scytonemin biosynthesis, the thiamine diphosphate (TPP)-dependent enzyme (ScyA) condenses two α-keto acids to yield an α-hydroxy-β-keto acid intermediate. (**Figure 3-1**) (Balskus and Walsh 2008, Balskus and Walsh 2009). However, the details of the ScyA reaction mechanism were not clarified. From these findings, (-)-sattabacin (**1**) was considered to be synthesized by the condensation between two α-keto acid substrates followed by decarboxylation, as shown **Figure 3-2**.

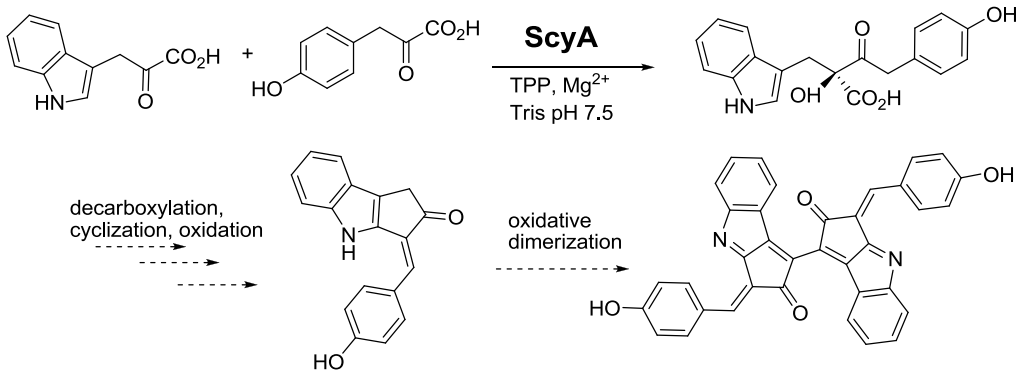


Figure 3-1 Early step in the biosynthetic pathway of scytonemin from the cyanobacterium *N. punctiforme*. This scheme was adapted from Balskus and Walsh 2008.

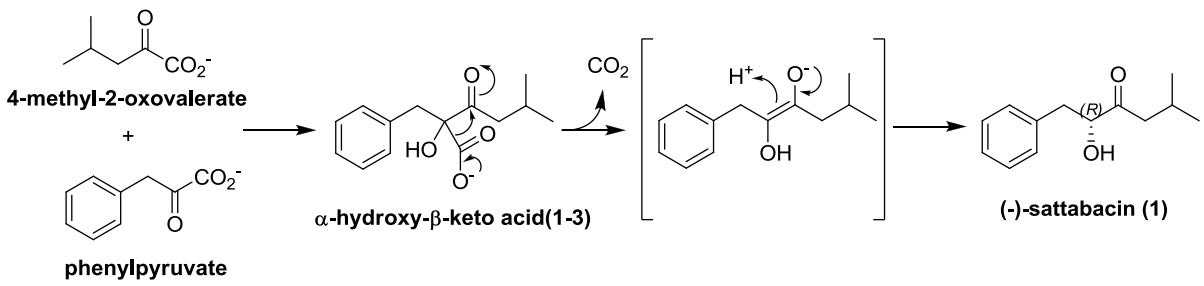


Figure 3-2 Initially presumed biosynthetic route for (−)-sattabacin (**1**).

3.1.2.2. Overexpression of the recombinant ScyA homologs

To evaluate this hypothesis, I identified gene products that were homologous to ScyA from the genomic DNA sequence of *T. hazakensis* SK20-1^T and found four homologous proteins, Thzk0150, Thzk5630, Thzk1026 and Thzk3750 with 29, 26, 26, and 25% sequence identity to ScyA, respectively (**Figure 3-3**, **Table 3-2**). Each gene was then overexpressed with a polyhistidine-tag, and the expressed proteins were purified by nickel-nitrilotriacetic acid (Ni-NTA)-agarose chromatography (**Figure 3-4**).

Table 3-2 List of *scyA* homologous genes identified in the *T. hazakensis* SK20-1^T genome.

Gene	Score (E value)[a]	Proposed function
thzk0150	203 (4e-53)	Acetolactate synthase (ALS) large subunit
thzk5630	157 (2e-39)	Glyoxylate carboligase
thzk2026[b]	139 (8e-34)	ALS large subunit
thzk1026	129 (1e-30)	Pyruvate oxidase
thzk3750	127 (3e-30)	TPP requiring enzyme

[a] Score was calculated using ScyA amino acid sequence as a query.

[b] The gene was excluded from candidates of *scyA* homolog owing to certain function involved in a leucine biosynthesis.

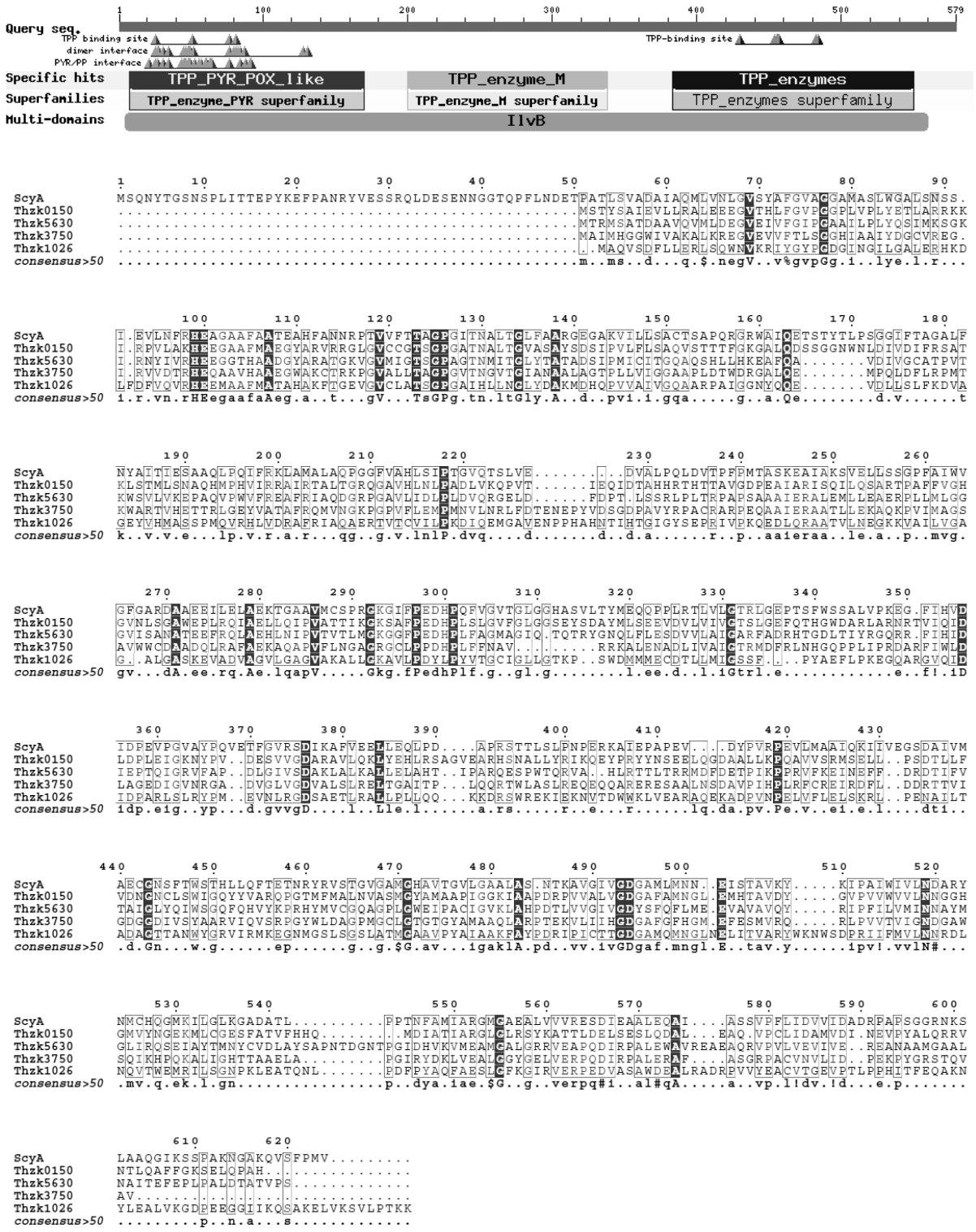


Figure 3-3 Putative conserved domain of Thzk0150 (top) and the sequence alignment of ScyA, Thzk0150, Thzk1026, Thzk3750, and Thzk5630 (bottom).

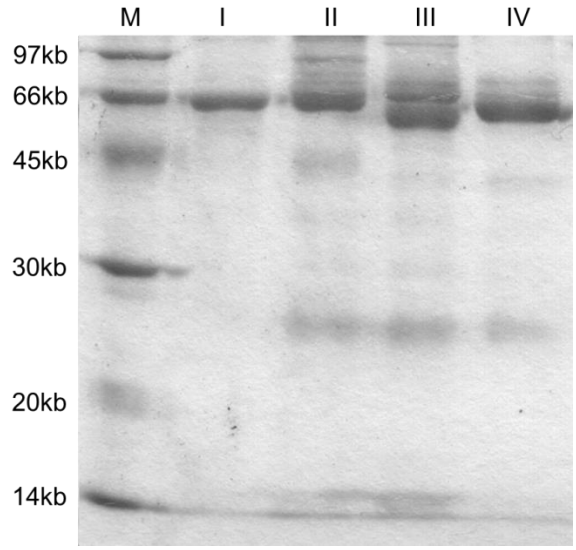


Figure 3-4 SDS-PAGE analysis of purified ScyA homologs of *T. hazakensis* SK20-1^T. M: protein marker; I: Thzk0150 (64.7 kDa); II: Thzk1026 (67.6 kDa); III: Thzk3750 (61.8 kDa); and IV: Thzk5630 (64.6 kDa).

3.1.2.3. Investigation of the enzymatic activity of ScyA homologs

The activity of each purified enzyme was then examined using two possible substrates, phenylpyruvate and 4-methyl-2-oxovalerate, in the presence of TPP and Mg^{2+} as cofactors. This examination revealed that only the Thzk0150-catalyzed reaction yielded new peaks according to LC-MS analysis with *m/z* 249 $[M-H]^-$ and *m/z* 205 $[M-COOH]^-$, which corresponded to a reaction product (**Figure 3-5**). In contrast, no peaks corresponding to sattabacin (*m/z* 207 $[M+H]^+$ and *m/z* 189 $[M-H_2O+H]^+$) were observed from LC-MS analysis. This LC-MS analysis of the reaction product suggested that 2-benzyl-2-hydroxy-5-methyl-3-oxohexanoic acid (**1-3**) was likely the α-hydroxy-β-keto acid intermediate of (-)-sattabacin. Based on the structure of **1**, I hypothesized the mechanism of the TPP-mediated coupling reaction by Thzk0150 as shown in **Figure 3-2** and **Figure 3-6**. I presumed that the decarboxylation of **1-3** by an unidentified decarboxylase produced. (−)-sattabacin (**1**) following the coupling reaction.

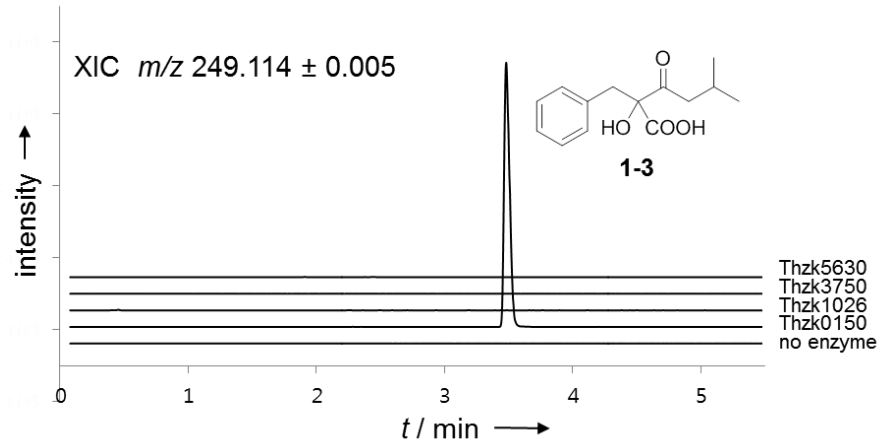


Figure 3-5 LC-MS analysis of the reaction mixtures. Four recombinant enzymes, Thzk0150, Thzk1026, Thzk3750, and Thzk5630, were investigated for the synthesis of the α-hydroxy-β-keto acid intermediate (**1-3**). The extracted ion current (XIC) chromatograms in the LC-MS analysis represent *m/z* 249.114 ± 0.005, which corresponds to **1-3**.

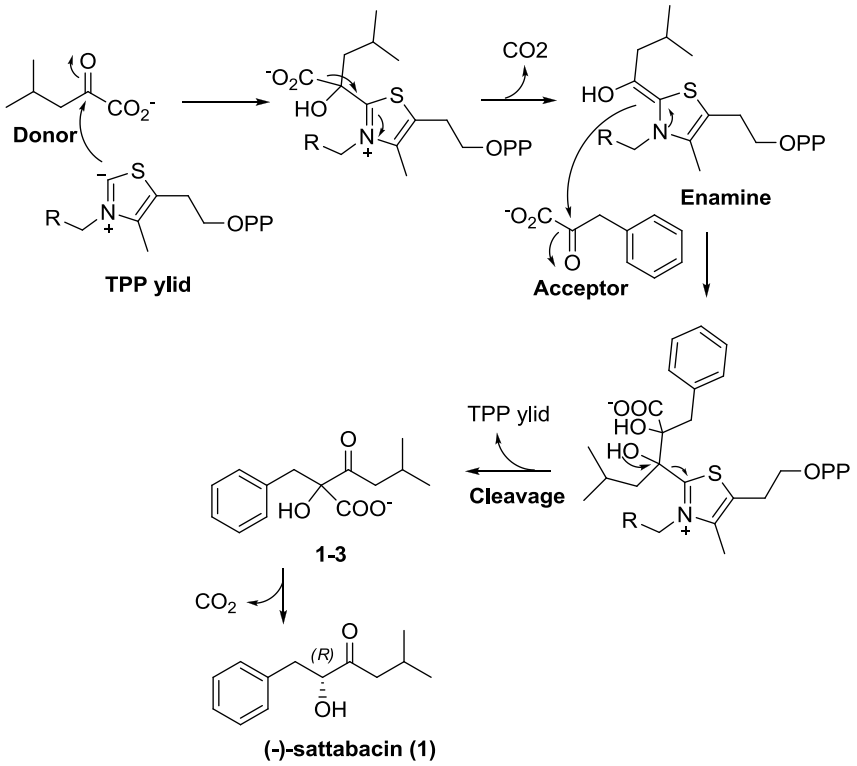


Figure 3-6 Initially presumed reaction mechanism of Thzk0150. In this mechanism, 4-methyl-2-oxovalerate (the donor) first reacts with the TPP ylid to form an alcohol intermediate. The resultant alcohol undergoes decarboxylation, yielding an enamine. Subsequently, the enamine double bond participates in nucleophilic addition to the carbonyl group of phenylpyruvate (the acceptor) to yield a β-hydroxy iminium ion. Finally, the ion decomposes to yield the α-hydroxy-β-keto acid, and the TPP ylid simultaneously regenerates. However, the enamine hypothesized in this mechanism was not detected in the Thzk0150-catalyzed reaction mixture using LC-MS analysis.

3.1.2.4. TPP and Mg^{2+} dependent Thzk0150

According to conserved domain search of Thzk0150, two TPP binding sites were identified in the N- and C-terminal regions of Thzk0150 (**Figure 3-3**). Accordingly, the enzymatic activity of Thzk0150 was abolished in the absence of TPP (**Figure 3-7**)); thus, I could conclude that Thzk0150 is a TPP-dependent enzyme. In addition to TPP dependence, Mg^{2+} was also investigated for its participation in the Thzk0150-catalyzed reaction as a cofactor (**Figure 3-8**), and the participation of Mg^{2+} was subsequently supported by LC-MS analysis (**Figure 3-7**).

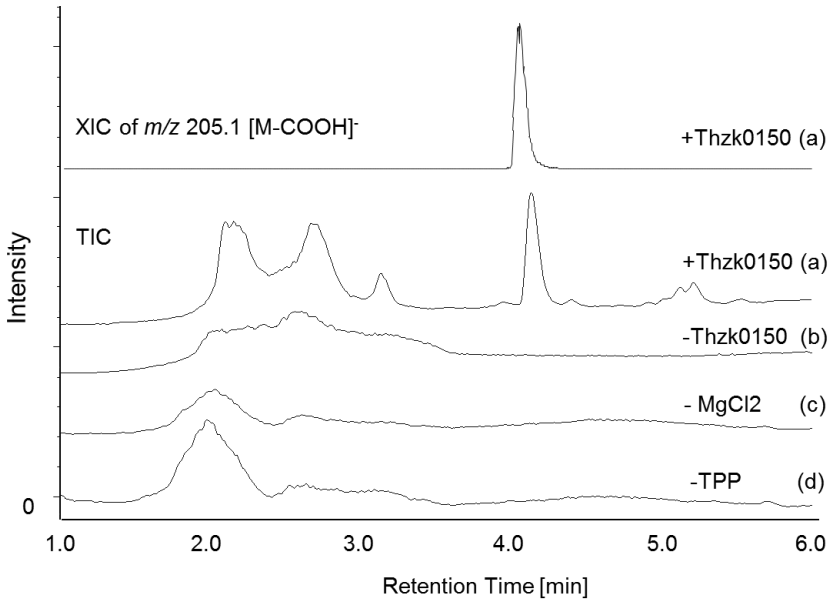


Figure 3-7 LC-ESI(-)-MS analysis to determine the dependency of TPP and Mg^{2+} in the Thzk0150-catalyzed reaction; Thzk0150, $MgCl_2$ and TPP (a), only $MgCl_2$ and TPP (b), Thzk0150 and TPP (c), and Thzk0150 and $MgCl_2$ (d).

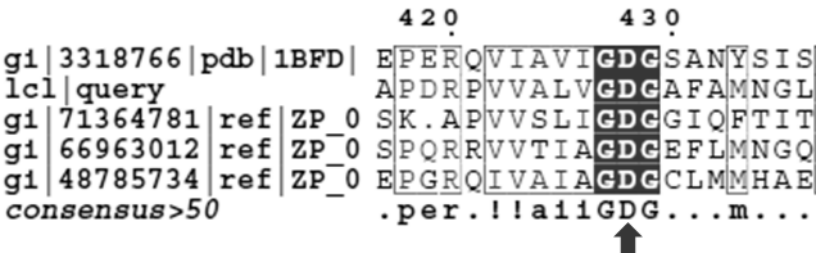


Figure 3-8 Identification of the conserved magnesium binding residue (D428) in Thzk0150.

3.1.2.5. Estimation of the molecular mass and structural organization of Thzk0150

The molecular mass of purified Thzk0150 was estimated using GFC. Two peaks were observed in the elution profile of Thzk0150 (**Figure 3-9**). Based on molecular mass calibration using marker proteins, the second peak was found to have a molecular mass of 230.3 kDa (**Figure 3-10**). Considering that the molecular mass of a subunit of Thzk0150 was found to be 64.7 kDa for an N-terminally $(His)_8$-tagged construct, the second peak is proposed to correspond to a tetramer, and the first peak appears to be an impurity that is present in the injected loading sample. This result was supported by SDS-PAGE analysis of the fractions (Fr. 5–Fr. 14) from GFC. The previously reported acetolactate synthase and acetohydroxyacid synthase have also been identified as tetrameric or dimeric; therefore, the estimated size of Thzk0150 is considered appropriate (Pang, Duggleby et al. 2002, Pang, Duggleby et al. 2004, Wang, Lee et al. 2009).

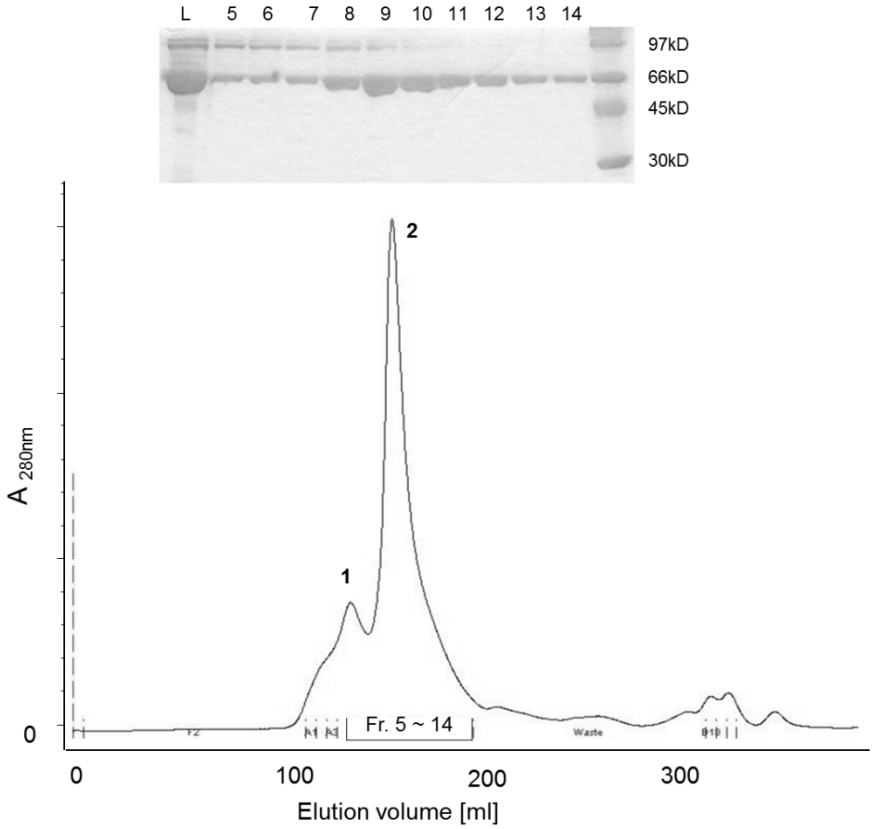


Figure 3-9 Elution profile of Thzk0150 in gel-filtration chromatography (GFC).

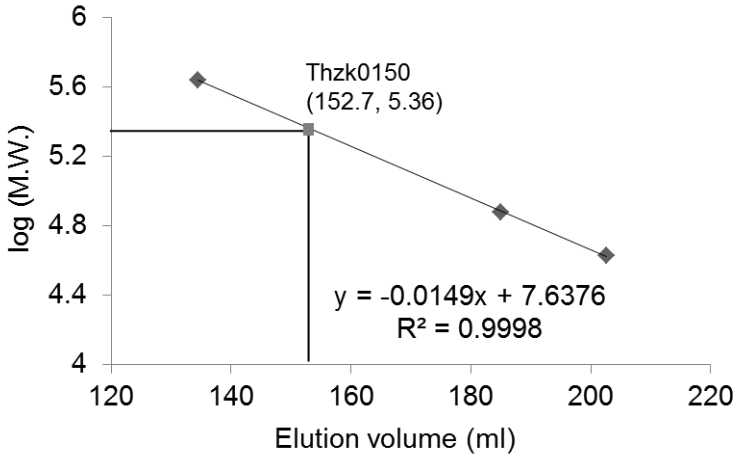


Figure 3-10 Estimation of the molecular mass of Thzk0150. The protein marker used consisted of ferritin (440 kDa), conalbumin (75 kDa), and ovalbumin (43 kDa). The elution volume of peak 2 in **Figure 3-9** is indicated.

3.1.2.6. Enzymatic reaction mechanism of Thzk0150

To clarify the reaction mechanism, I prepared a large-scale Tkzk0150-catalyzed reaction and purified the reaction product. Unexpectedly, the purified product yielded an identical ^{1}H NMR spectrum to that of **1** (**Figure 3-11**). However, the purified product did not possess an identical optical activity ($[\alpha]_D^{25}$ −2.1 (*c* 0.05, $CHCl_3$)). These results indicated that the purified product was a racemic mixture of (±)-sattabacin.

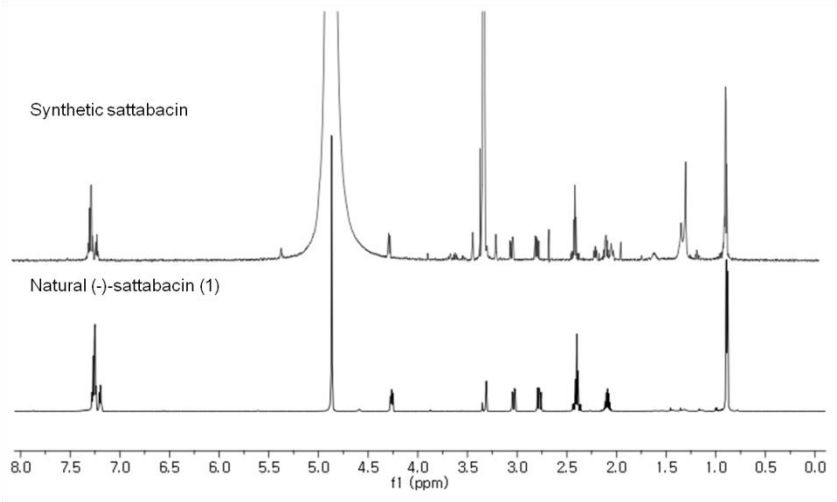


Figure 3-11 ^{1}H NMR spectra (600 MHz, CD_3OD) of enzymatically synthesized sattabacin (top) and (-)-sattabacin (bottom) isolated from *T. hazakensis* SK20-1^{T}.

I thus hypothesized that the nonenzymatic decarboxylation of **1-3** produced the racemic sattabacin during my purification procedure because such nonenzymatic decarboxylation following a TPP-dependent condensation of two α-keto acids has also been observed for ScyA. In addition to **1-3**, I detected a small amount of an additional unknown product (**1-5**) in the reaction mixture (**Figure 3-12**) and determined the structure as 3-hydroxy-1,4-diphenylbutan-2-one (**1-5**, **Figure 3-13**) using NMR and MS analyses .

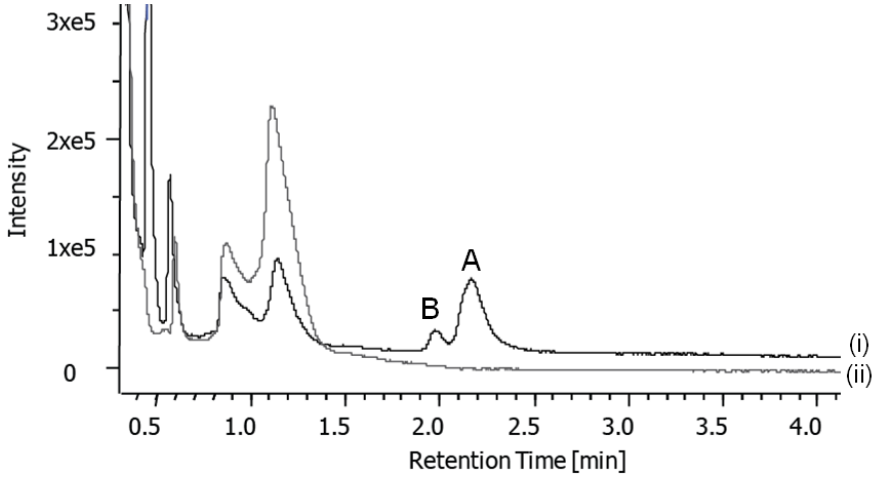


Figure 3-12 HPLC chromatogram of the Thzk0150-catalyzed reaction mixture with enzyme (i) and without enzyme (ii). The two labeled peaks represent the reaction products, α-hydroxy-β-keto acid (A) and compound **1-5** (B).

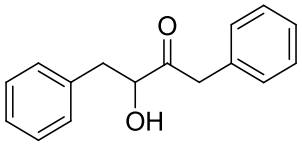


Figure 3-13 Structure of the byproduct (**1-5**) of the Thzk0150-catalyzed reaction.

Compound **1-5**: pale yellow oil; ^{1}H NMR (600 MHz, CD_3OD,

Figure A5) δ 7.16~7.28 (8H, m), 7.09 (2H, d, *J* = 7.2 Hz), 4.38 (1H, *J* = 4.8, 7.8 Hz), 3.82 (1H, *J* = 16.2 Hz), 3.76 (1H, *J* = 16.2 Hz), 3.05 (1H, *J* = 4.2, 13.8 Hz), 2.81 (1H, *J* = 7.8, 13.8 Hz); ^{13}C NMR (150 MHz, CD_3OD) δ 210.8, 137.4, 134.0, 129.4, 129.1, 128.0 (2C), 126.3, 126.2, 77.1, 45.0, 39.4 (δc was assigned by HSQC and HMBC); HR-ESI-MS (*m/z*): $[M-H]^-$ calcd for $C_{16}H_{15}O_2$, 239.1090; found, 239.1090.

However, the identification of **1-5** led me to reject the previously hypothesized reaction mechanism (**Figure 3-6**) because **1-5** could not be produced through this reaction mechanism. Alternatively, I propose a likely mechanism of the TPP-mediated coupling reaction between phenylpyruvate and 4-methyl-2-oxovalerate (**Figure 3-15**). In this alternative mechanism, phenylpyruvate (the donor) first reacts with the TPP ylid to form an alcohol intermediate. The resultant alcohol undergoes decarboxylation, yielding an enamine. Subsequently, the enamine double bond acts as a nucleophile and adds to the carbonyl group of 4-methyl-2-oxovalerate (the acceptor) to produce a β-hydroxy iminium ion. The ion then decomposes to yield 2-hydroxy-4-methyl-2-(2-phenylacetyl)pentanic acid (**1-4**), which is an alternative α-hydroxy-β-keto acid intermediate of sattabacin.

In this alternative mechanism (**Figure 3-15**), the protonation of the enamine (**1-7**) likely yields an additional β-hydroxy iminium ion (**1-8**), followed by decomposition to yield benzyl aldehyde (**1-9**). Subsequently, another molecule of enamine (**1-7**) acts as a nucleophile and adds to the carbonyl group of **1-9** to yield a β-hydroxy iminium ion (**1-10**). Finally, the cleavage of **1-10** by a retro-aldol reaction produces **1-5**. This mechanism was strongly supported by the observation of *m/z* 543.087 $[M-H]^-$ from the LC-MS analysis, which corresponded to the enamine adduct (**1-7**) between TPP and phenylpyruvate (**Figure 3-14**). I thus concluded that Thzk0150 catalyzes the condensation of phenylpyruvate (the donor) and 4-methyl-2-oxovalerate (the acceptor) to form **1-4**, which is the intermediate of sattabacin. The resultant α-hydroxy-β-keto acid intermediate **1-4** undergoes decarboxylation, yielding an enediol intermediate (**1-6**). Finally, deprotonation of a hydroxy group on C3' of **1-6** and enantioselective protonation at C2' yields (−)-sattabacin (**1**). Because the protonation at C2' is enantioselective, one or more enzymes must be involved in the conversion of **1-4** to **1** in SK20-1^T cells. However, I was unable to identify candidate enzymes in the flanking regions of the *thzk0150* gene in the genome sequence of SK20-1^T that could be involved in the decarboxylation and protonation in (−)-sattabacin biosynthesis.

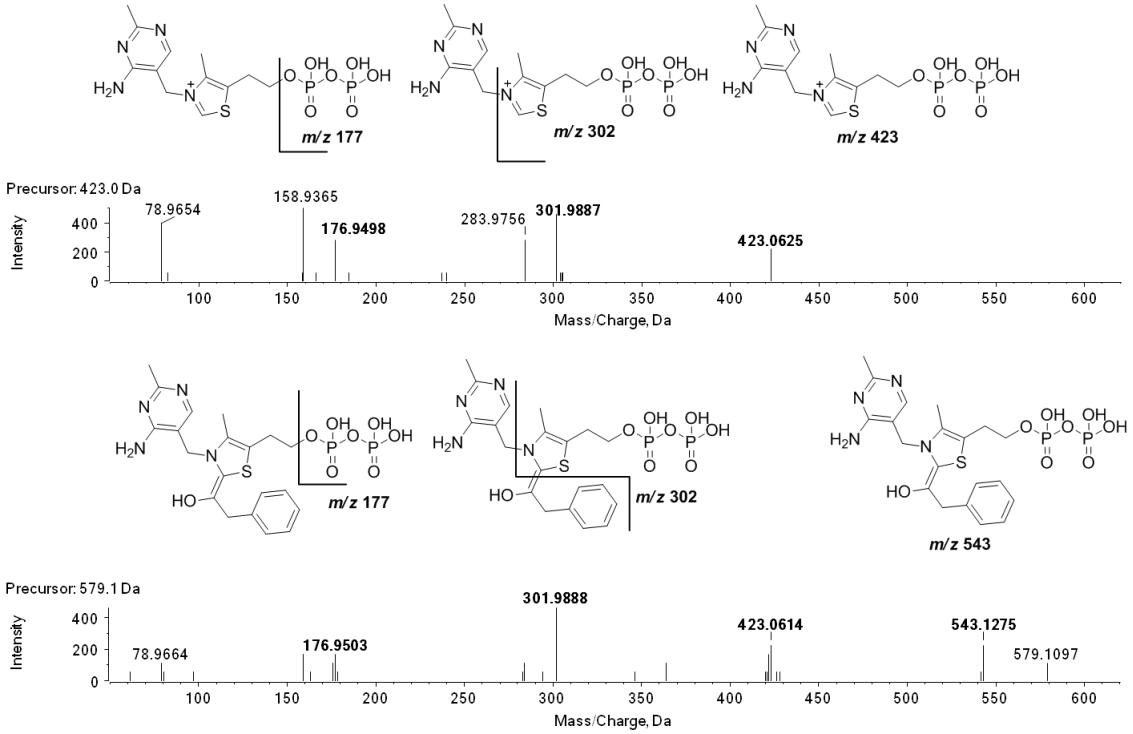


Figure 3-14 MS/MS fragment patterns of thiamine diphosphate (*m/z* 423, top) and an enamine adduct between TPP and phenylpyruvate (*m/z* 543, bottom) observed in the Thzk0150-catalyzed reaction mixture.

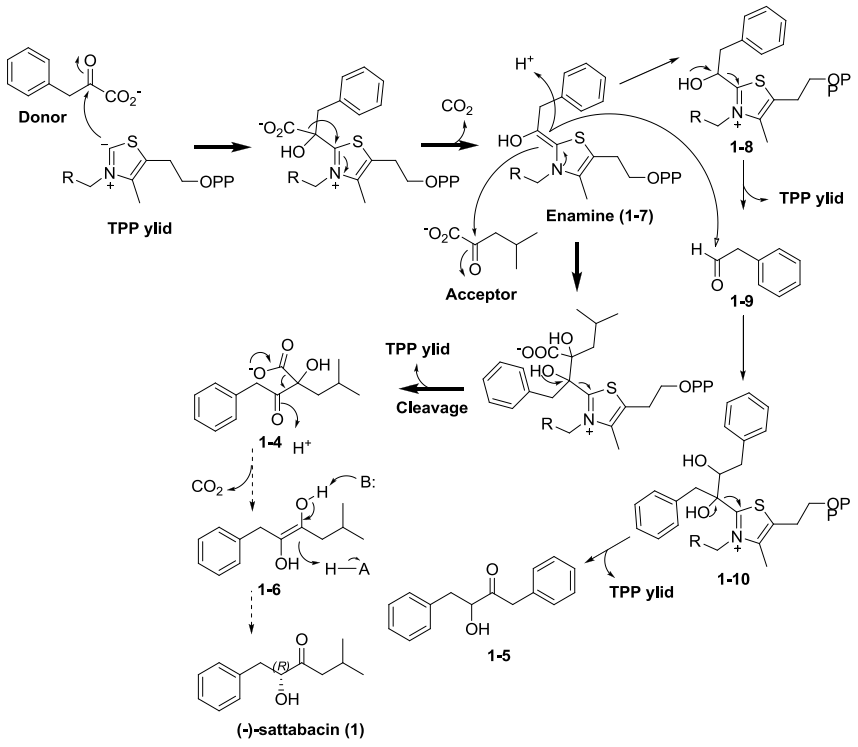
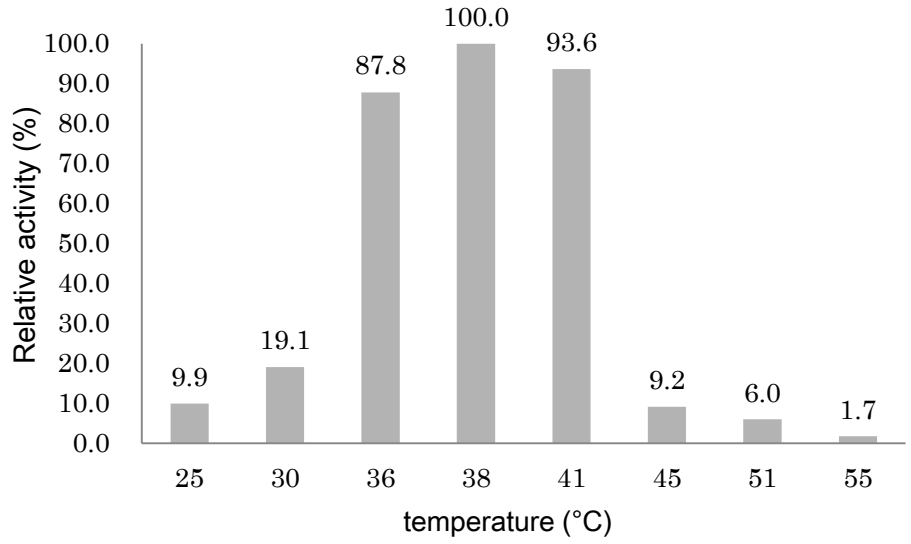


Figure 3-15 Proposed alternative reaction mechanism for the production of acyloin compounds by the TPP-dependent Thzk0150 enzyme.

3.1.2.7. Determination of the optimal temperature for the enzymatic activity of Thzk0150

Next, I examined the optimal temperature of the Thzk0150-catalyzed reaction. The optimal temperature was determined to be 36−41 °C (**Figure 3-16**). Considering that *T. hazakensis* SK20-1^T grows at 31–58 ℃, with optimal growth at 50 ℃, the determined optimal temperature of Thzk0150 was likely too low. Generally, enzymes derived from thermophiles are minimally active at room temperature but exhibit enzymatic activity at their optimal growth temperature (Merz, Yee et al. 2000). Therefore, this result led me to inquire as to the cause for the difference between the optimal Thzk0150 enzymatic temperature and the optimal temperature for cell physiology. To understand this phenomenon, I considered horizontal (lateral) gene transfer, which is the transfer of genes between organisms in a manner other than traditional reproduction (Ochman, Lawrence et al. 2000). Increasing evidence indicates that lateral gene transfer has played an integral role in the evolution of the bacterial genome. Lateral gene transfer can be detected by several methods; a significant difference between the G+C content of a specific gene and the overall G+C content of the entire chromosome can also serve as evidence for lateral gene transfer (Ochman, Lawrence et al. 2000). In fact, the 53.1 % global G+C content of *T. hazakensis* SK20-1^T genome is different from the 59.4 % G+C content of *thzk0150*. Therefore, *thzk0150* presumably has been recently acquired in *T. hazakensis* SK20-1^T and was originally present in a mesophilic bacterium growing at a 37℃ optimal temperature. However, I am not confident that the 6 % difference is significant, which would indicate lateral (horizontal) gene transfer. Furthermore, other genes within the *thzk0150*-flanking regions display a wide range of % GC content, as shown in **Figure 3-17**.



temperature	25	30	36	38	41	45	51	55
Area of XIC	14786	28421	130936	149102	139564	13647	8907	2559
relative activity (%)	9.9	19.1	87.8	100.0	93.6	9.2	6.0	1.7

Figure 3-16 Temperature-dependent Thzk0150 enzymatic activity. The purified enzyme precipitated in the reaction mixture at temperatures greater than 51℃.

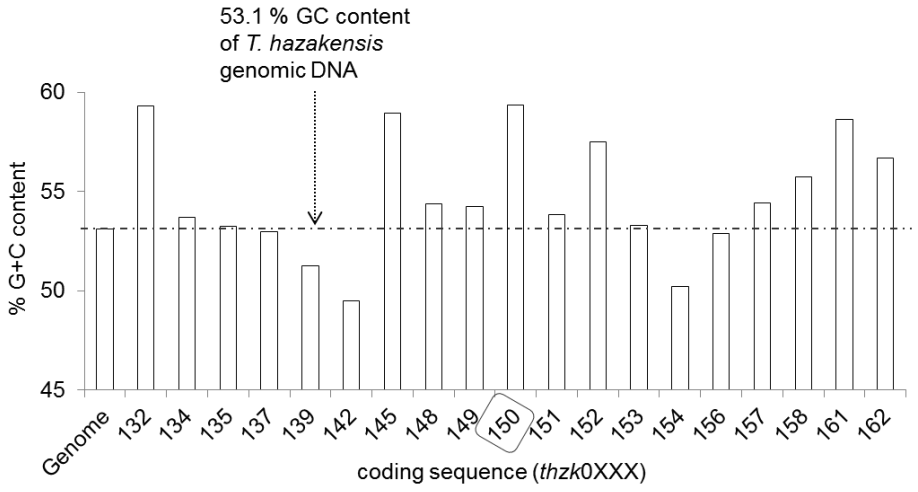


Figure 3-17 The % G+C content of the flanking regions of thzk0150. Coding sequences smaller than 200 bp were ignored.

Although the concept of lateral gene transfer may be appropriate, an unknown factor may have affected the enzymatic properties of Thzk0150, such as thermal stability. Alternatively, although there is insufficient experimental evidence to confirm this theory, we suspect that the N-terminal His-tag in the recombinant Thzk0150 influenced the enzymatic activity and thermal stability; similar influences on protein solubility and catalytic characteristics have frequently been observed (Ledent, Duez et al. 1997).

3.1.2.8. Substrate specificity of Thzk0150

The substrate specificity of Thzk0150 as an acyl acceptor was investigated using four α-keto acids: 2-oxobutanoate, 2-oxoisovalerate, (±)-3-methyl-2-oxovalerate, and pyruvate (rather than 4-methyl-2-oxovalerate) (**Figure 3-18**). With the exception of pyruvate, the remaining three molecules were accepted as acyl acceptor substrates by the recombinant Thzk0150 enzyme to each yield the corresponding α-hydroxy-β-keto acids (**Figure 3-19**). Based on this result, compound **2** ((-)-hazakacin) was also presumably produced by Thzk0150-catalyzed condensation between phenylpyruvate and 2-oxobutanoate (**Figure 3-19d**). Furthermore, two other aromatic α-keto acids, *p*-hydroxyphenylpyruvate and indole-3-pyruvate, were investigated as acyl donors instead of phenylpyruvate. This examination revealed that only indole-3-pyruvate was accepted by Thzk0150 to yield 2-(2-(1*H*-indol-3-yl)acetyl)-2-hydroxy- 4-methylpentanoic acid (**Figure 3-19**), which may represent an α-hydroxy-β-keto acid intermediate for the biosynthesis of sattazolin in the sattabacin-producing *Bacillus* sp. (Lampis, Deidda et al. 1995).

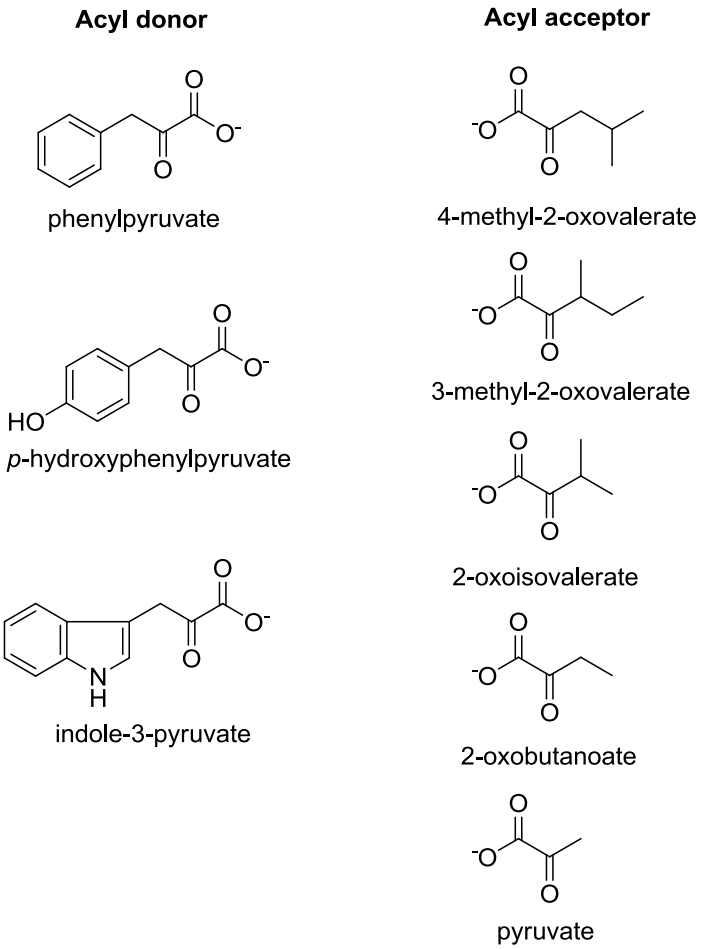


Figure 3-18 Acyl donor and acceptor substrates used to examine the substrate specificity of Thzk0150.

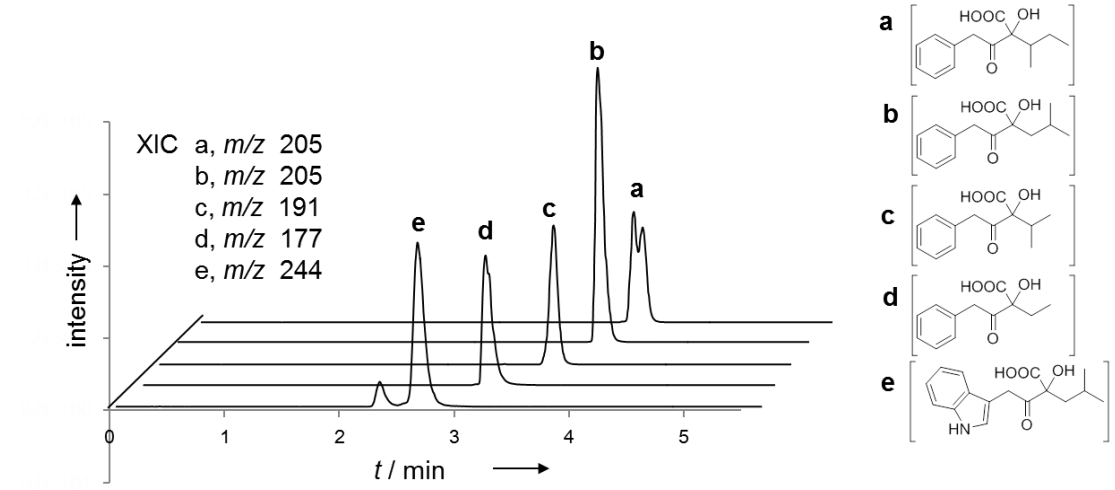


Figure 3-19 LC-MS analysis of the substrate specificity of the Thzk0150 enzyme. When phenylpyruvate was used as a donor substrate, (±)-3-methyl-2-oxovalerate (a), 4-methyl-2-oxovalerate (b), 2-oxoisovalerate (c), and 2-oxobutanoate (d) were accepted as acyl acceptors. When 4-methyl-2-oxovalerate was used as an acyl acceptor, indole-3-pyruvate (e) and phenylpyruvate were accepted as donor substrates. Each chromatogram was represented using the XIC of the $[M-COOH]^-$ derived from α-hydroxy-β-keto acids formed in each reaction mixture. The structures of the α-hydroxy-β-keto acids formed in each reaction mixture are displayed in the chromatograms. The carbonyl carbon and the adjacent β-hydroxy carbon in the displayed structures could be interchangeable. High-resolution MS spectra for each α-hydroxy-β-keto acid are available in **Figure**s **3-20** to **3-24**.

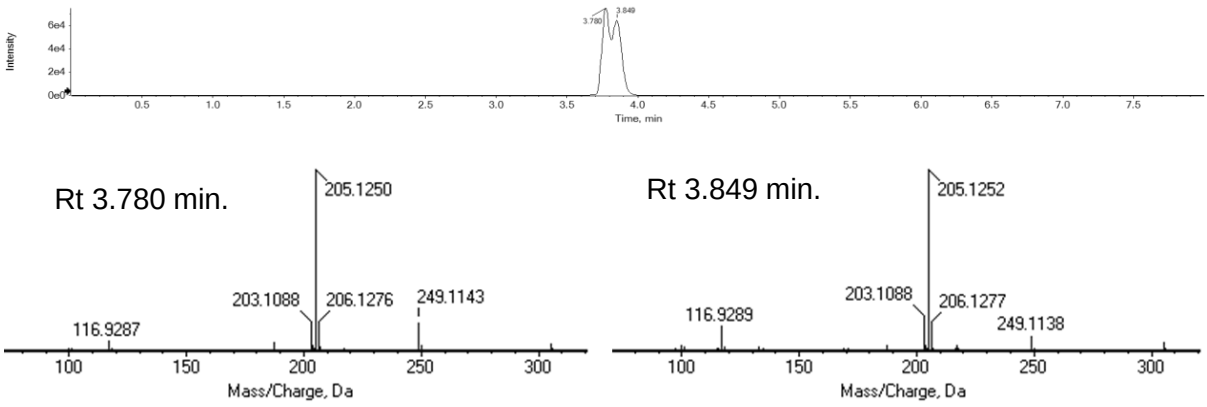


Figure 3-20 LC-MS ion current chromatogram extracted by *m/z* 205 (top) and ESI(−)-MS spectra of α-hydroxy-β-keto acids (*m/z* 205 $[M-COOH]^-$; bottom) synthesized by Thzk0150 in the reaction mixture containing phenylpyruvate and (±)-3-methyl-2-oxovalerate.

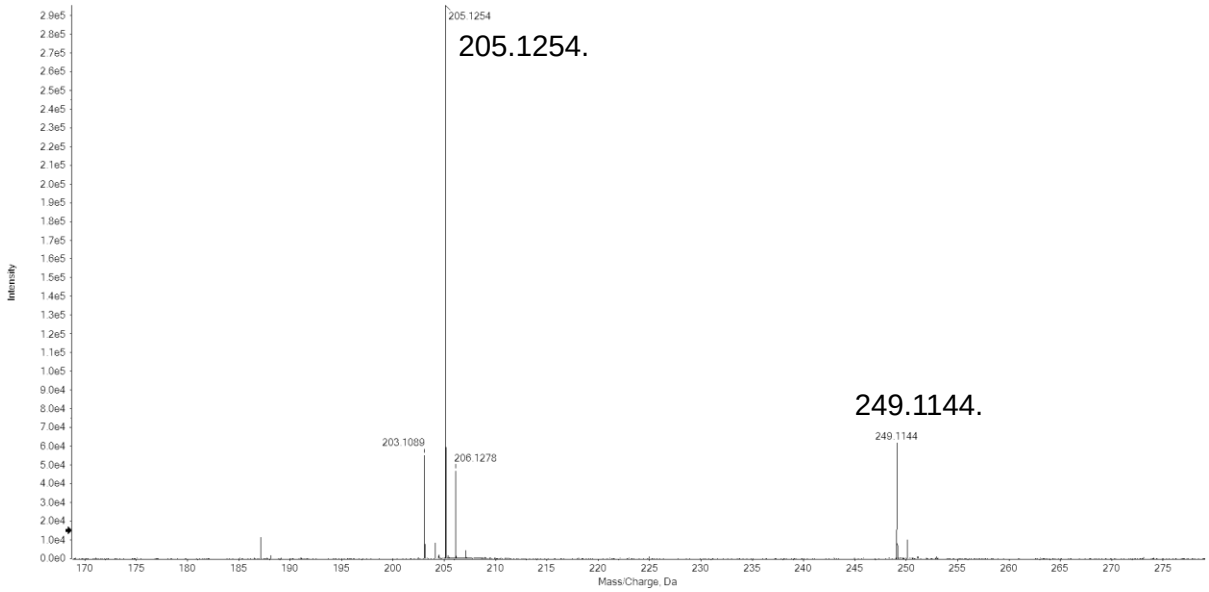


Figure 3-21 ESI(−)-MS spectrum of an α-hydroxy-β-keto acid (*m/z* 249 $[M-H]^-$, 205 $[M-COOH]^-$) synthesized by Thzk0150 in the reaction mixture containing phenylpyruvate and 4-methyl-2-oxovalerate.

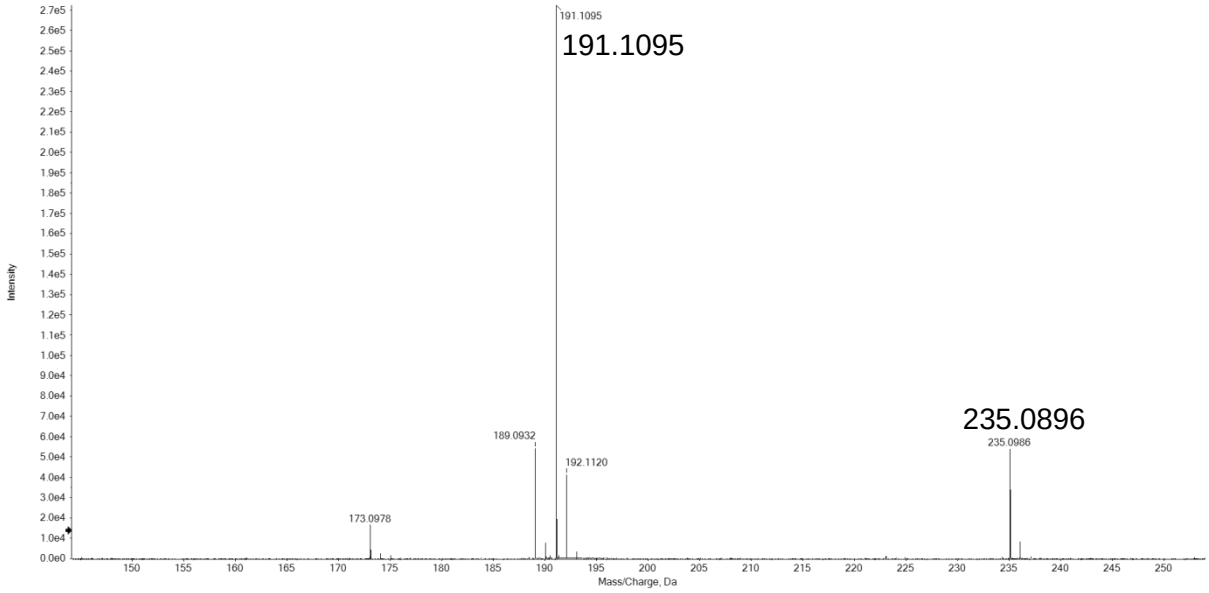


Figure 3-22 ESI(−)-MS spectrum of an α-hydroxy-β-keto acid (*m/z* 235 $[M-H]^-$, 191 $[M-COOH]^-$) synthesized by Thzk0150 in the reaction mixture containing phenylpyruvate and 2-oxoisovalerate.

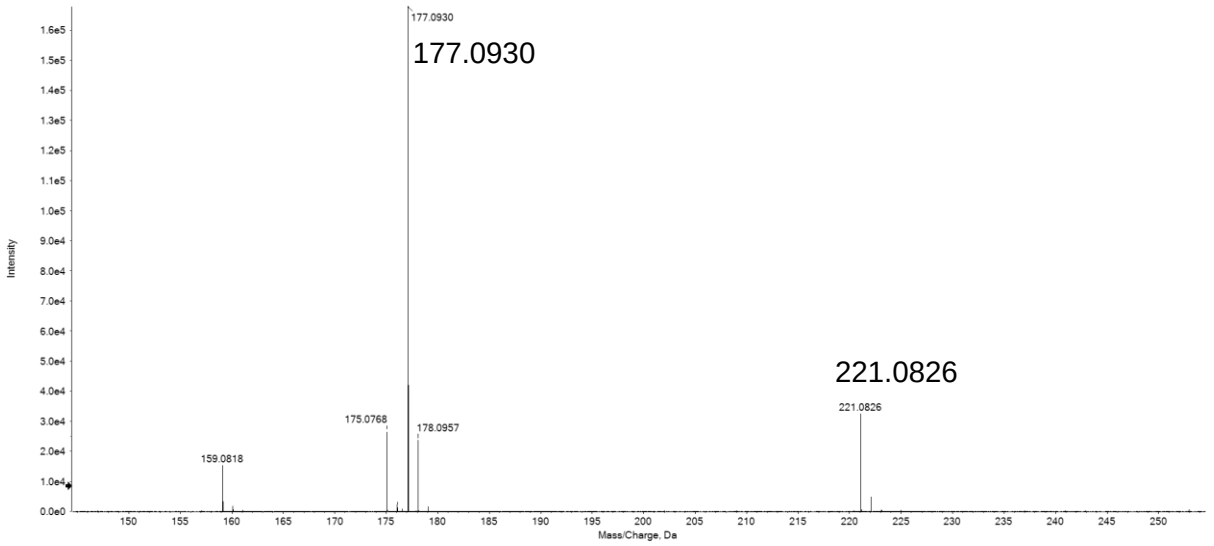


Figure 3-23 ESI(−)-MS spectrum of an α-hydroxy-β-keto acid (*m/z* 221 $[M-H]^-$, 177 $[M-COOH]^-$) synthesized by Thzk0150 in the reaction mixture containing phenylpyruvate and 2-oxobutanoate.

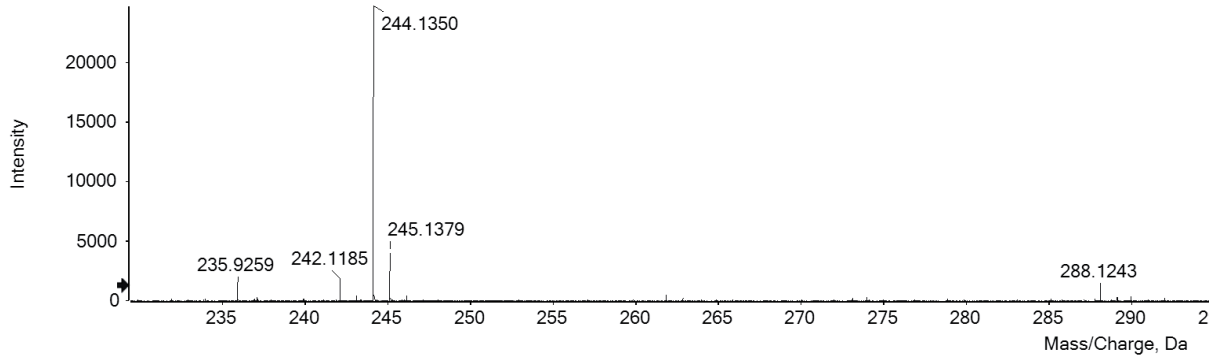


Figure 3-24 ESI(−)-MS spectrum of an α-hydroxy-β-keto acid (*m/z* 288 $[M-H]^-$, 244 $[M-COOH]^-$) synthesized by Thzk0150 in the reaction mixture containing indole-3-pyruvate and 4-methyl-2-oxovalerate.

3.1.2.9. Gene analysis of the regions flanking *thzk0150*

I analyzed the genes flanking *thzk0150* to identify a decarboxylase gene presumably required for the completion of **1** and **2** biosynthesis. However, I was unable to identify candidate enzymes within the flanking regions of the *thzk0150* gene in the genome of *T. hazakensis* SK20-1^{T} that could be involved in the decarboxylation and protonation of α-hydroxy-β-keto acid intermediate (**1-4**) in (−)-sattabacin biosynthesis (**Figure 3-25**, **Table 3-3**); most of the genes near *thzk0150* encode hypothetical proteins and are, therefore, not presumed to be responsible for acyloin biosynthesis. Consequently, I supposed that specific hypothetical proteins in the flanking regions may catalyze the conversion of the α-hydroxy-β-keto acid intermediate (**1-4**) to (-)-sattabacin (**1**) via an unprecedented mechanism or the biosynthetic genes for acyloins are presumed to be scattered throughout the genome in *T. hazakensis.* Additionally, prephenate dehydratase-annotated Thzk0148 was expected to be responsible for the production of phenylpyruvate from prephenate. In contrast, the biosynthetic genes for scytonemin in the cyanobacterium *N. punctiforme* form a gene cluster of the putative genes involved in the production of the precursors indole-3-pyruvate and *p*-hydroxyphenylpyruvate.

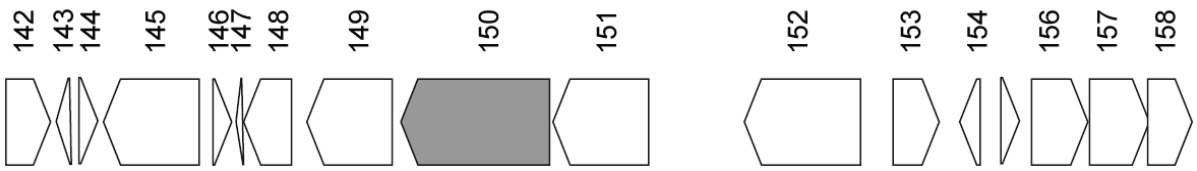


Figure 3-25 Gene organization of the regions flanking *thzk0150.*

Table 3-3 Annotation of *thzk0150* and genes flanking *thzk0150*.

# ORF	Size (bp)	Conserved domain	proposed protein function
142	513	Helix-turn-helix domain	hypothetical protein
143	114	no hit	hypothetical protein
144	165	no hit	hypothetical protein
145	1164	Major Facilitator Superfamily	major facilitator superfamily MFS_1
146	171	no hit	hypothetical protein
147	117	no hit	hypothetical protein
148	**570**	**Prephenate dehydratase**	**Prephenate dehydratase**
149	999	3-Oxoacyl-ACP synthase III	3-oxoacyl-ACP synthase III
150	**1740**	**TPP enzyme**	**Acetolactate synthase large subunit**
151	1113	no hit	hypothetical protein
152	1353	Acetyltransferase (GNAT) domain	hypothetical protein
153	546	Ham1 family	Inosine triphosphate pyrophosphatas
154	216	no hit	hypothetical protein
155	171	no hit	hypothetical protein
156	660	Bacterial regulatory proteins, tetR family	Transcriptional regulator, TetR family
157	702	FAD binding domain	Oxidoreductase
158	531	FAD binding domain	Oxidoreductase

3.2. インターネット公表出来ないために削除

3.3. インターネット公表出来ないために削除

3.4. インターネット公表出来ないために削除

CONCLUSION

インターネット公表出来ないために削除

EXPERIMENTAL

Materials

Media

LB

Bacto tryptone	1.0 %
Bacto yeast extract	0.5 %
NaCl	1.0 %
(Bacto Agar)	1.5 %

TSB

Tryptone Soya Broth	3.0 %
(agar)	1.5 %

K

Soluble starch	2.5 %
Soybean meal	1.5 %
Dry yeast	0.2 %
$CaCO_3$	0.4 %
pH	adjust to 7.2

Q

Glycerin	2.0 %
Molasses	1.0 %
Casein	0.5 %
Polypeptone	0.1 %
$CaCO_3$	0.4 %
pH	adjust to 7.2

A1

Soluble starch	1.0 %
Bacto yeast extract	0.4 %
Peptone	0.2 %
pH	adjust to 7.0

YM

Bacto yeast extract	0.5 %
Malt extract broth	3.0 %
pH	adjust to 7.0

MISP2

Glucose	0.4 %
Malt extract broth	1.0 %
Bacto yeast extract	0.4 %
pH	adjust to 7.0

TB

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

After autoclave, following solution was added.

KH_2PO_4	2.31 g
K_2HP_4	12.54 g
H_2O	fill up to 100 mL

Antibiotics

ampicillin	100 μg/ml
kanamycin	50 μg/ml

Buffers

For plasmid DNA preparation from *E. coli* (alkaline lysis)

P1

Tris base	6.06 g
EDTA	3.72 g
H_2O	fill up to 800 mL
pH	adjust to 8.0 with HCl then added H_2O up to 1 L

P2

NaOH	8.0 g in 900 mL H_2O
20 % SDS	50 mL
H_2O	fill up to 1 L

P3

Potassium acetate	294.5 g in H2O
pH	adjust to 5.5 with glacial acetic acid (appropriate 110 mL)
H_2O	fill up to 1 L

TE

0.5 M Tris-HCl (pH 8.0)	200 µl
1 M EDTA	1 mL
H_2O	98.8 mL

For SDS-PAGE (12%)

Running gel

Acrylamide	12 %
Tris-HCl (pH 8.8)	375 mM
SDS	0.1 %
APS	0.1 %
TEMED	0.1 %

Stacking gel

Acrylamide	4.5 %
Tris-HCl (pH 6.8)	90 mM
SDS	0.05 %
APS	0.1 %
TEMED	0.2 %

Electrophoresis buffer

Tris	3.03 g
Glycine	14.4 g
SDS	1.0 g
H_2O	fill up to 1 L

2x SDS-PAGE sample buffer

Tris-HCl (pH 6.8)	100 mM
DTT	200 mM
SDS	4 % (w/v)
Bromophenol blue	0.2 %
Glycerol	20 % (v/v)

For agarose electrophoresis

Agarose gel

Agarose ME	1.0 g
1 x TAE	fill up to 100 mL

1 x TAE

Tris	4.84 g
Acetate	1.14 mL
EDTA	0.74 g
H_2O	fill up to 1 L

For protein purification

Wash buffer

Tris-HCl (pH 8.0)	50 mM
NaCl	0.5 M
Glycerol	20 % (w/v)
Imidazole (pH 7.5)	20 mM

Lysis buffer

added 1% Tween 20 into wash buffer

Elution buffer

Tris-HCl (pH 8.0)	50 mM
NaCl	0.5 M
Glycerol	20 % (w/v)
Imidazole (pH 7.5)	250 mM

Stock solution

NaCl	0.2 M
Tris-HCl (pH 7.5)	50 mM

Plasmid used in this study

plasmid	Description	reference or source
pT7blue	*E. coli* cloning vector; Amp^r; pUCori	Novagen
pHis8	*E. coli* expression vector; Kan^r; pBR322ori	Jez, Ferrer et al. 2000
pET26b(+)	*E. coli* expression vector;Kan^r; pBR322ori	Novagen

Kits and enzymes used in this study

Article	Manufacturer	Use
GenElute Plamid Miniprep Kit	Sigma	Preparation of plasmid DNA
MinElute Gel Extraction Kit	Qiagen	DNA purification
GoTaq Green Master Mix	Promega	Colony PCR
Ligation high	TOYOBO	DNA ligation
Expand High Fidelity PCR system	Roche	Gene amplification
Restriction enzymes	TaKaRa	Cloning
Lysozyme	Nacalai	Preparation of genomic DNA

Methods

Preservation of bacterial stocks

For long term storage of thee *E. coli* strains and *T. hazakensis*, 500 µl of culture was mixed with 500 µl of 60% glycerol in a cryovial and stored at -80℃.

Genomic DNA preparation using the salting-out method

T. hazakensis was cultured in 10 mL of TSB medium for 2−3 days with shaking at 55℃. Cells from 500 µL of culture were harvested by centrifugation (15,000 rpm, 1 min) and resuspended in an identical volume of solution I. The mixture was then gently mixed with a rotary mixer for 30 min followed by the addition of 500 µL of 1 % SDS and a 1-min incubation. After the addition of 250 µL of phenol/chloroform, 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 the volume of 3 M sodium acetate and an identical volume of isopropanol. After centrifugation (15,000 rpm; 10 min), the supernatant was removed, and the DNA pellet 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 ℃ for 30 min before the quantity and quality of the DNA was determined using agarose gel (0.8 %) electrophoresis.

Plasmid DNA preparation using alkaline lysis method

Plasmid DNA was routinely isolated using the alkaline lysis methods (Birnboim and Doly 1979). Cells from 3 mL of overnight culture were harvested by centrifugation (15,000 rpm; 1 min) and resuspended in 300 µL of P1. The mixture was then lysed with 300 µL of P2 by inverting six times and incubated for 5 min on ice followed by the addition of 300 µL of P3 and a 15-min incubation on ice. The resulting mixture was then centrifuged at 15,000 rpm for 10 min. Next, 700 µL of the supernatant was transferred to a new tube, and an identical 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 volumes 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 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. For sequencing purposes, pure plasmid DNA was prepared using the GenElute Plasmid Miniprep Kit from Sigma.

Transformation of *E. coli*

Frozen competent cells were thawed on ice for 5 min to which 5 µl of the plasmid DNA was added and incubated on ice for 40 min. Subsequently, the cells were heat shocked at 42℃ for 2 min. The cells were then incubated again on ice for 2 min, and 1 mL of LB medium was added to the cold-shocked cells, which were then incubated at 37℃ for 1 h. The cells were centrifuged at 13,000 rpm for 1 min, and the cell pellet was resuspended in 300 µl of fresh LB medium. A volume of 100 µl of the cell suspension was plated on an LB agar plate containing a selection marker, and the plate was incubated overnight at 37℃. Positive transformants were identified by restriction digestion or colony PCR.

Overexpression of recombinant DNA in *E. coli*

E. coli BL21(DE3) was used as the expression host. A transformant was cultured overnight in LB, and a 2 % inoculum from this culture was used to inoculate 200 mL of TB medium containing the selection marker. The culture was incubated at 37℃ for approximately 2 h until the OD_{600} reached 0.5 to 1.0. After starvation on ice for 10 min and induction by IPTG, cells were incubated overnight at 18℃.

Cell lysis

Cells were harvested from the induced cultures by centrifugation at 5000 rpm for 10 min at 4℃. The cell pellet was then resuspended with the lysis buffer and incubated on ice for 30 to 60 min. The cells were lysed on ice by sonication using several short pulses (0.5–1 min) with pauses (0.5–1 min). The cell debris was pelleted by centrifugation at 17,000 rpm for 20 min at 4℃. The supernatant contains the native protein of interest. The soluble (supernatant) and the insoluble (pellet) fractions were analyzed using SDS-PAGE to determine the presence of the protein of interest.

Recombinant protein purification

His-tagged recombinant protein was purified using Ni-NTA chromatography. Two milliliters of Ni-NTA resin was washed with 5 volumes of washing buffer. The supernatant was applied to the column and allowed to flow through. The column was then washed with 15 volumes of washing buffer. Subsequently, the protein was eluted from the column with 3 volumes of elution buffer. Each fraction was then analyzed using SDS-PAGE.

Instruments

UHPLC	
UHPLC Column	Shiseido Capcell Pak C18 2.0mm × 50mm
UHPLC Pump	JASCO X-LC 3185PU
UHPLC PDA	JASCO X-LC 3110MD
UHPLC Auto-sampler	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
Analytical HPLC	
HPLC Column	Senshu Pak PEGASIL ODS 4.6ϕ × 250mm
HPLC Pump	JASCO PU-2089 Plus
HPLC PDA	JASCO MD-2010 Plus
HPLC Auto-sampler	JASCO AS-2057 Plus
Preparative HPLC	
Prep. HPLC Column	Senshu Pak PENGASIL ODS 20ϕ × 250mm
Prep. HPLC UV/Vis Detector	SSC-5410, Senshu Scientific co. ltd
Prep. HPLC Pump	SSC-3462, Senshu Scientific co. ltd
LCMSMS	
LC	Shimadzu UFLC
LC Column	Shiseido Capcell Pak C18 2.0mm × 50mm
MSMS	AB SCIEX Triple TOF 5600
NMR	JEOL Superconducting Magnet 600 MHz (^{1}H NMR), 150 MHz (^{13}C NMR)

APPENDIX

Selected NMR spectra used in this study

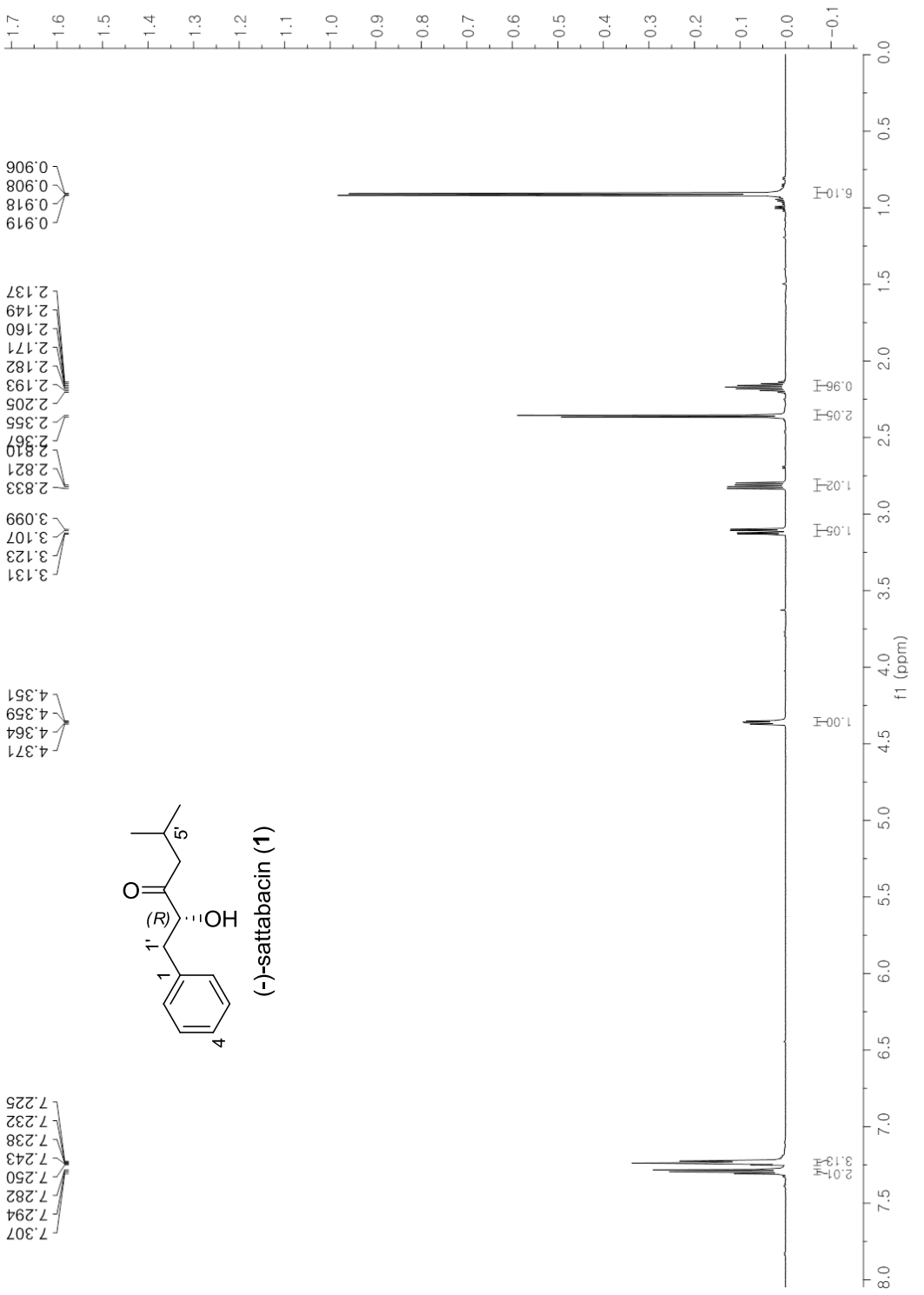


Figure A1 ^{1}H NMR spectrum of compound **1** (600 MHz, $CDCl_3$).

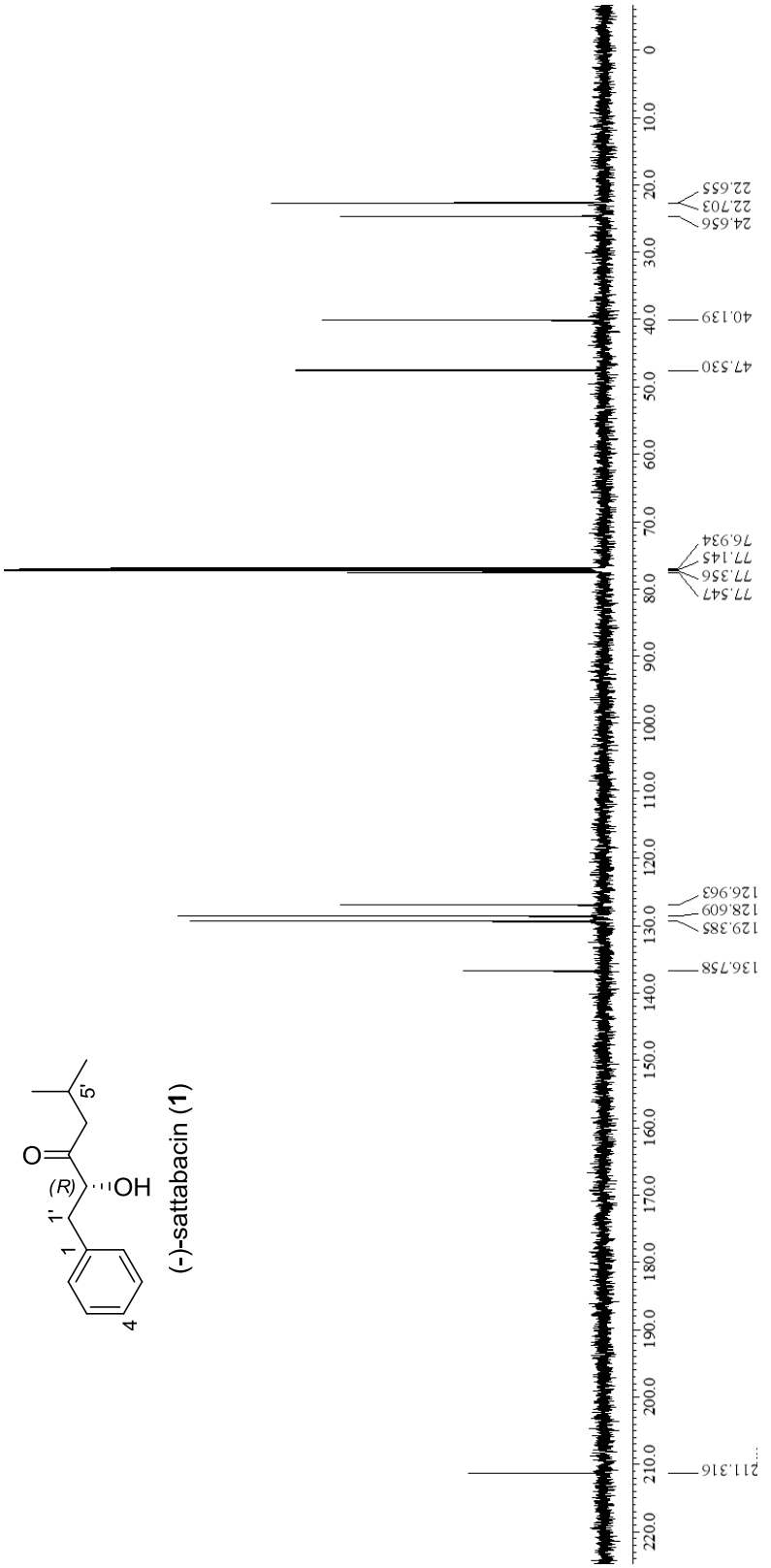


Figure A2 ^{13}C NMR spectrum of compound **1** (150 MHz, $CDCl_3$).

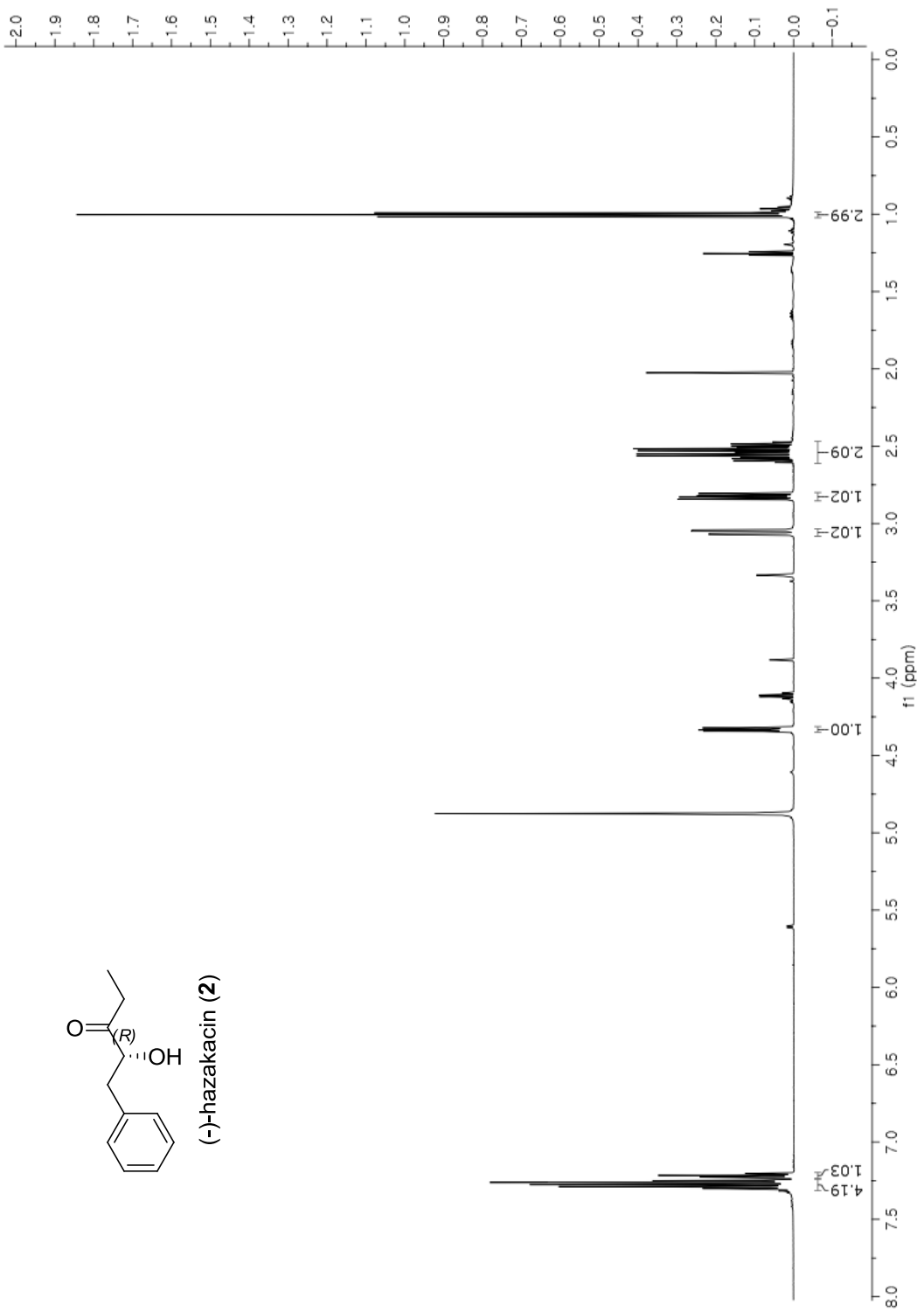


Figure A3 ^{1}H NMR spectrum of compound **2** (600 MHz, CD_3OD).

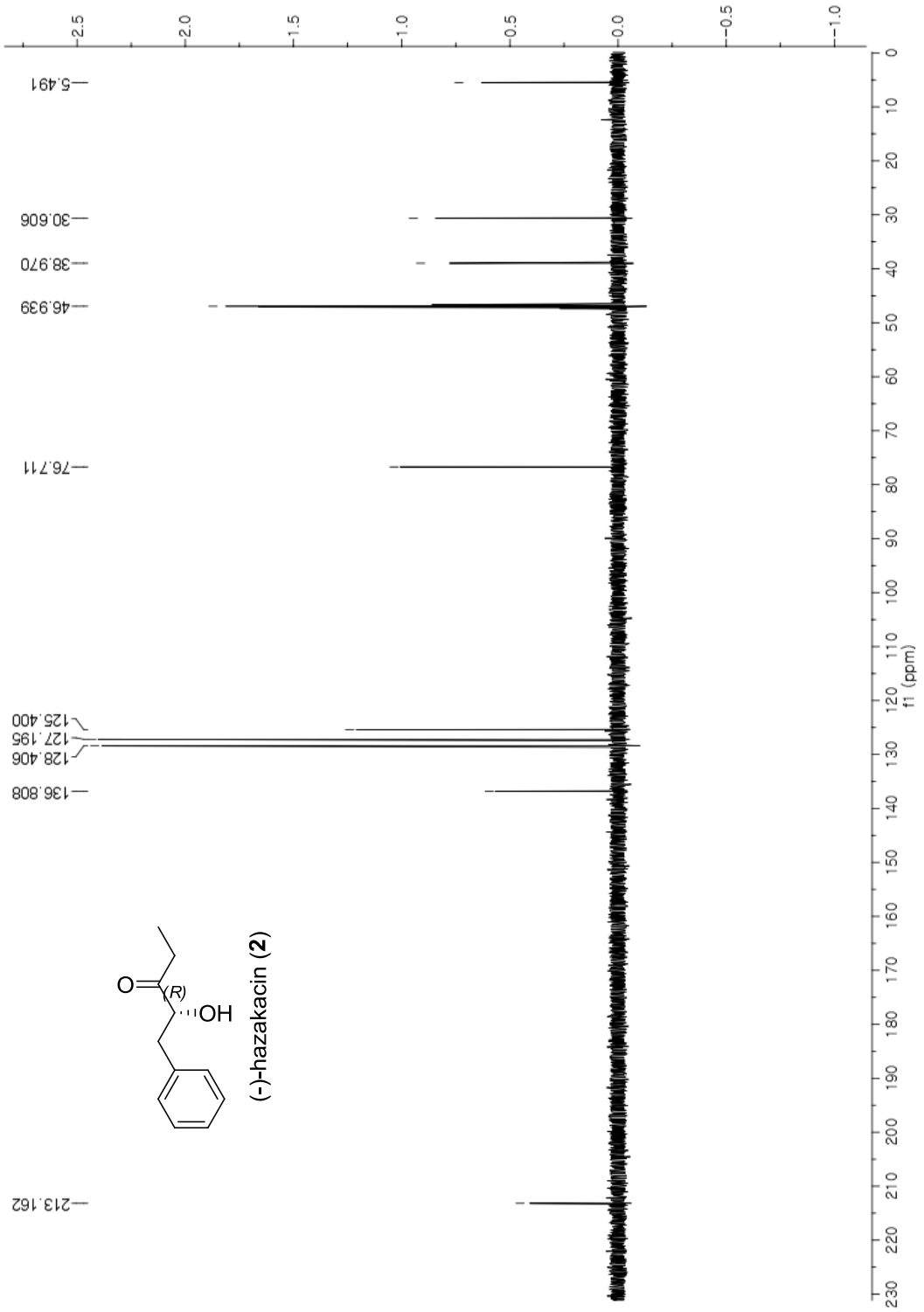


Figure A4 ^{13}C NMR spectrum of compound **2** (150 MHz, CD_3OD).

Figure A5-A13　インターネット公表出来ないために削除

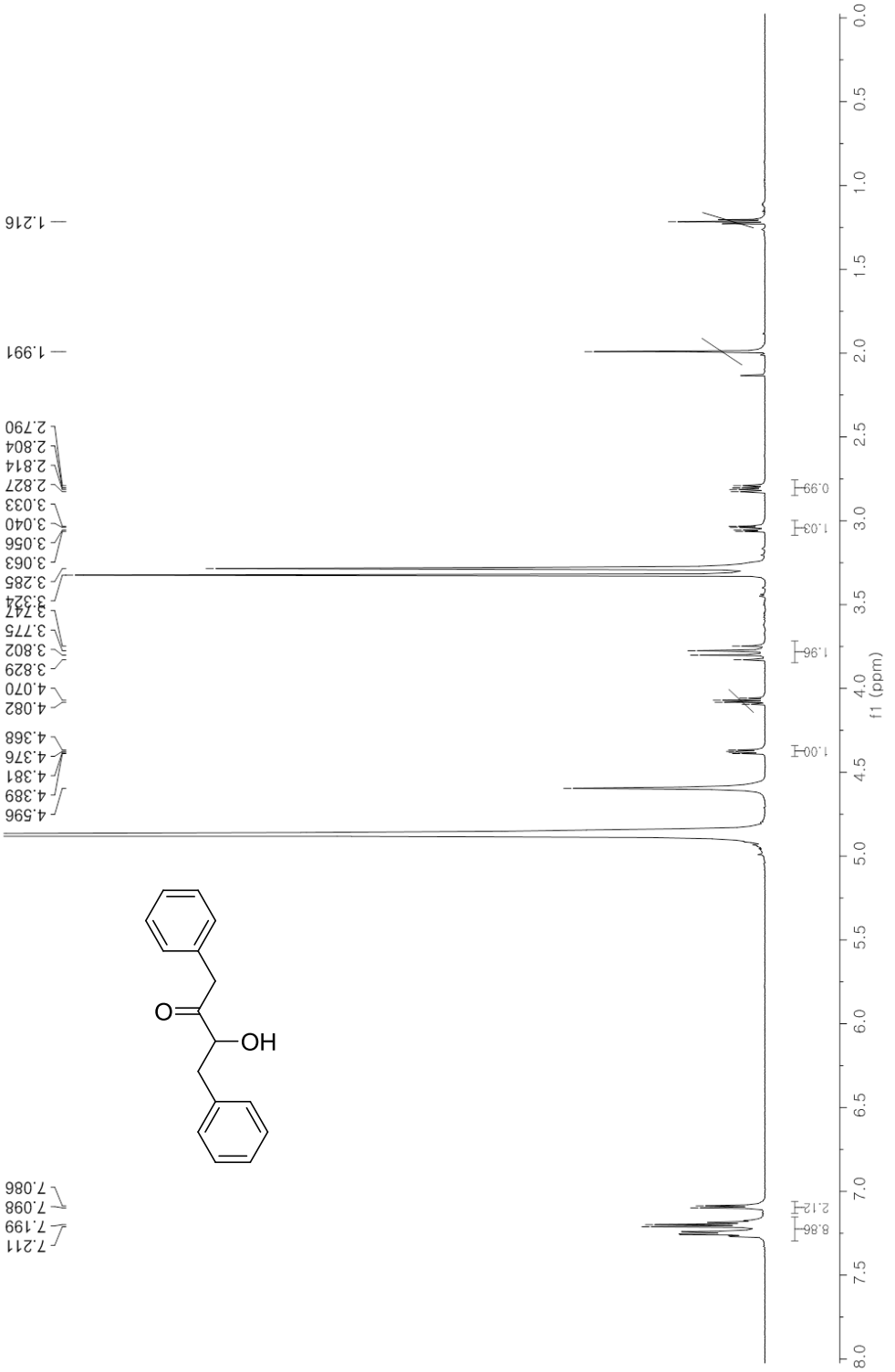


Figure A5 ^{1}H NMR spectrum of Thzk0150 minor product, **1-5** (600 MHz, CD_3OD). The peaks with a diagonal line indicate ethyl acetate impurities.

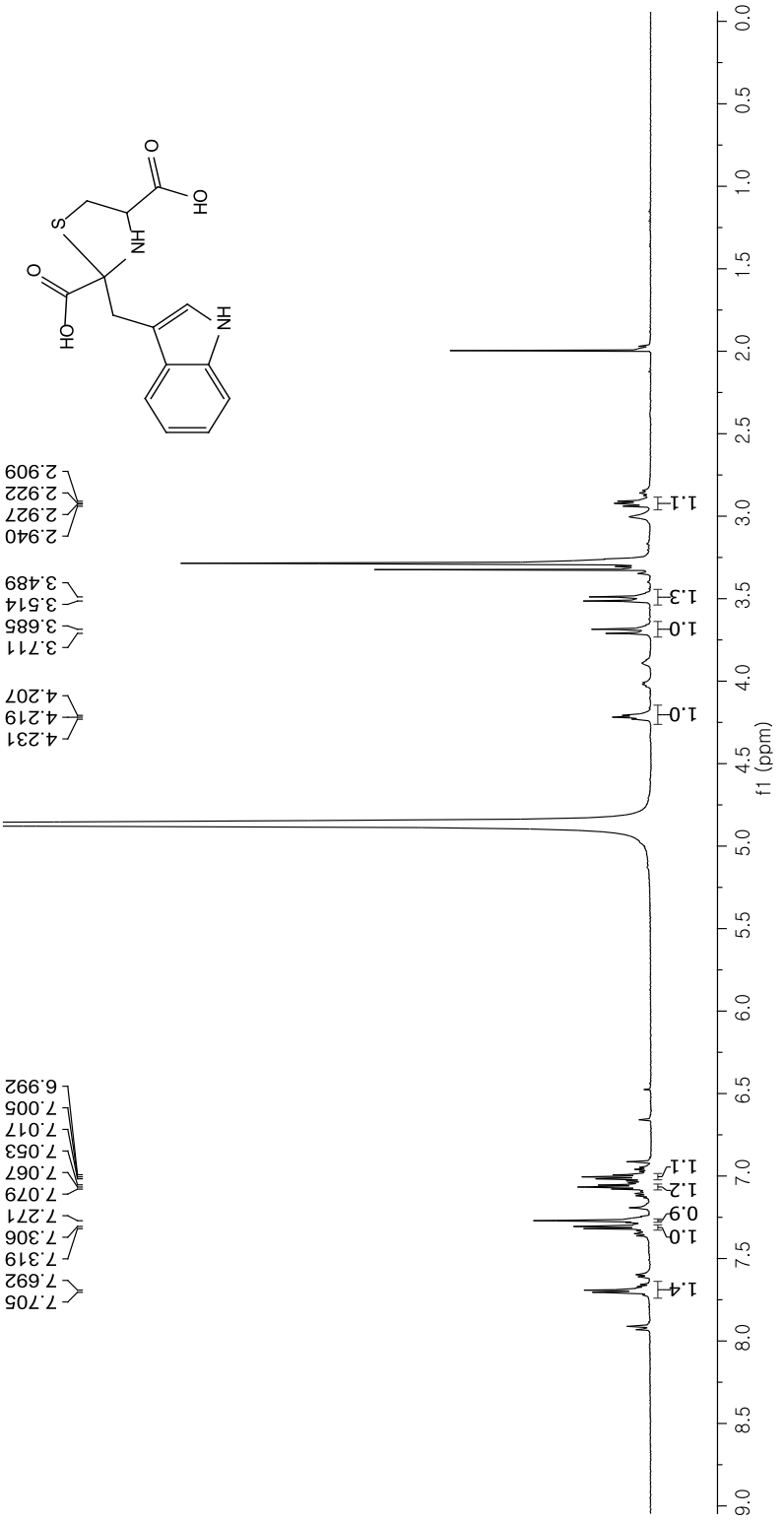


Figure A6 ^{1}H NMR spectrum of non-enzymatically generated cyclic compound (MW 306) (CD_3OD, 600 MHz)

REFERENCE

Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990). "Basic local alignment search tool." J Mol Biol **215**(3): 403-410.

Ansari, M. Z., G. Yadav, R. S. Gokhale and D. Mohanty (2004). "NRPS-PKS: a knowledge-based resource for analysis of NRPS/PKS megasynthases." Nucleic Acids Res **32**(Web Server issue): W405-413.

Aronoff, M. R., N. A. Bourjaily and K. A. Miller (2010). "Concise, protecting group free total syntheses of (+)-sattabacin and (+)-4-hydroxysattabacin." Tetrahedron Lett **51**(49): 6375-6377.

Ausubel, F. M. (1999). Short protocols in molecular biology : a compendium of methods from Current protocols in molecular biology. New York ; Chichester, Wiley.

Aziz, R. K., D. Bartels, A. A. Best, M. DeJongh, T. Disz, R. A. Edwards, K. Formsma, S. Gerdes, E. M. Glass, M. Kubal, F. Meyer, G. J. Olsen, R. Olson, A. L. Osterman, R. A. Overbeek, L. K. McNeil, D. Paarmann, T. Paczian, B. Parrello, G. D. Pusch, C. Reich, R. Stevens, O. Vassieva, V. Vonstein, A. Wilke and O. Zagnitko (2008). "The RAST Server: rapid annotations using subsystems technology." BMC Genomics **9**: 75.

Bachmann, B. O. and J. Ravel (2009). "Chapter 8. Methods for in silico prediction of microbial polyketide and nonribosomal peptide biosynthetic pathways from DNA sequence data." Methods Enzymol **458**: 181-217.

Balskus, E. P. and C. T. Walsh (2008). "Investigating the initial steps in the biosynthesis of cyanobacterial sunscreen scytonemin." J Am Chem Soc **130**(46): 15260-15261.

Balskus, E. P. and C. T. Walsh (2009). "An Enzymatic Cyclopentyl[b]indole Formation Involved in Scytonemin Biosynthesis." J Am Chem Soc **131**(41): 14648-+.

Bao, B., P. Zhang, Y. Lee, J. Hong, C. O. Lee and J. H. Jung (2007). "Monoindole alkaloids from a marine sponge Spongosorites sp." Mar Drugs **5**(2): 31-39.

Birnboim, H. C. and J. Doly (1979). "A rapid alkaline extraction procedure for screening recombinant plasmid DNA." Nucleic Acids Res **7**(6): 1513-1523.

Blin, K., M. H. Medema, D. Kazempour, M. A. Fischbach, R. Breitling, E. Takano and T. Weber (2013). "antiSMASH 2.0--a versatile platform for genome mining of secondary metabolite producers." Nucleic Acids Res **41**(Web Server issue): W204-212.

Boone, D. R., Y. Liu, Z. J. Zhao, D. L. Balkwill, G. R. Drake, T. O. Stevens and H. C. Aldrich (1995). "Bacillus infernus sp. nov., an Fe(III)- and Mn(IV)-reducing anaerobe from the deep terrestrial subsurface." Int J Syst Bacteriol **45**(3): 441-448.

Browne, L. M., K. L. Conn, W. A. Ayer and J. P. Tewari (1991). "The Camalexins - New Phytoalexins Produced in the Leaves of Camelina-Sativa (Cruciferae)." Tetrahedron **47**(24): 3909-3914.

Bult, C. J., O. White, G. J. Olsen, L. Zhou, R. D. Fleischmann, G. G. Sutton, J. A. Blake, L. M. FitzGerald, R. A. Clayton, J. D. Gocayne, A. R. Kerlavage, B. A. Dougherty, J. F. Tomb, M. D. Adams, C. I. Reich, R. Overbeek, E. F. Kirkness, K. G. Weinstock, J. M. Merrick, A. Glodek, J. L. Scott, N. S. Geoghagen and J. C. Venter (1996). "Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii." Science **273**(5278): 1058-1073.

Challis, G. L. and J. Ravel (2000). "Coelichelin, a new peptide siderophore encoded by the Streptomyces coelicolor genome: structure prediction from the sequence of its non-ribosomal peptide synthetase." FEMS Microbiol Lett **187**(2): 111-114.

Challis, G. L., J. Ravel and C. A. Townsend (2000). "Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains." Chem Biol **7**(3): 211-224.

Chan, Y. A., M. T. Boyne, 2nd, A. M. Podevels, A. K. Klimowicz, J. Handelsman, N. L. Kelleher and M. G. Thomas (2006). "Hydroxymalonyl-acyl carrier protein (ACP) and aminomalonyl-ACP are two additional type I polyketide synthase extender units." Proc Natl Acad Sci U S A **103**(39): 14349-14354.

Culioli, G., A. Ortalo-Magne, R. Valls, C. Hellio, A. S. Clare and L. Piovetti (2008). "Antifouling activity of meroditerpenoids from the marine brown alga Halidrys siliquosa." J Nat Prod **71**(7): 1121-1126.

Cutignano, A., G. Villani and A. Fontana (2012). "One metabolite, two pathways: convergence of polypropionate biosynthesis in fungi and marine molluscs." Org Lett **14**(4): 992-995.

Damianakos, H., G. Sotiroudis and I. Chinou (2013). "Pyrrolizidine Alkaloids from Onosma erecta." J Nat Prod.

Davies-Coleman, M. T. and M. J. Garson (1998). "Marine polypropionates." Nat Prod Rep **15**(5): 477-493.

Dimise, E. J., P. F. Widboom and S. D. Bruner (2008). "Structure elucidation and biosynthesis of fuscachelins, peptide siderophores from the moderate thermophile Thermobifida fusca." Proc Natl Acad Sci U S A **105**(40): 15311-15316.

Donadio, S., Monciardini, P., and Sosio, M. (2007). "Polyketide synthases and nonribosomal peptide synthetases: the emerging view from bacterial genomics." Nat Prod Rep **24**: 1073-1109.

Du, L. and L. Lou (2010). "PKS and NRPS release mechanisms." Nat Prod Rep **27**(2): 255-278.

Emmert, E. A., A. K. Klimowicz, M. G. Thomas and J. Handelsman (2004). "Genetics of zwittermicin a production by Bacillus cereus." Appl Environ Microbiol **70**(1): 104-113.

Erickson, K. L., K. R. Gustafson, D. J. Milanowski, L. K. Pannell, J. R. Klose and M. R. Boyd (2003). "Myriastramides A-C, new modified cyclic peptides from the Philippines marine sponge Myriastra clavosa." Tetrahedron **59**(51): 10231-10238.

Fenical, W. (1993). "Chemical Studies of Marine-Bacteria - Developing a New Resource." Chem Rev **93**(5): 1673-1683.

Fischbach, M. A. and C. T. Walsh (2006). "Assembly-line enzymology for polyketide and nonribosomal Peptide antibiotics: logic, machinery, and mechanisms." Chem Rev **106**(8): 3468-3496.

Gehring, A. M., E. DeMoll, J. D. Fetherston, I. Mori, G. F. Mayhew, F. R. Blattner, C. T. Walsh and R. D. Perry (1998). "Iron acquisition in plague: modular logic in enzymatic biogenesis of yersiniabactin by Yersinia pestis." Chem Biol **5**(10): 573-586.

Gondry, M., L. Sauguet, P. Belin, R. Thai, R. Amouroux, C. Tellier, K. Tuphile, M. Jacquet, S. Braud, M. Courcon, C. Masson, S. Dubois, S. Lautru, A. Lecoq, S. Hashimoto, R. Genet and J. L. Pernodet (2009). "Cyclodipeptide synthases are a family of tRNA-dependent peptide bond-forming enzymes." Nat Chem Biol **5**(6): 414-420.

Gruenewald, S., H. D. Mootz, P. Stehmeier and T. Stachelhaus (2004). "In vivo production of artificial nonribosomal peptide products in the heterologous host *Escherichia coli*." Appl Environ Microbiol **70**(6): 3282-3291.

Hertweck, C. (2009). "The biosynthetic logic of polyketide diversity." Angew Chem Int Ed Engl **48**(26): 4688-4716.

Huang, S. X., E. Powell, S. R. Rajski, L. X. Zhao, C. L. Jiang, Y. Duan, W. Xu and B. Shen (2010). "Discovery and total synthesis of a new estrogen receptor heterodimerizing actinopolymorphol A from Actinopolymorpha rutilus." Org Lett **12**(15): 3525-3527.

Ishikawa, J. and K. Hotta (1999). "FramePlot: a new implementation of the frame analysis for predicting protein-coding regions in bacterial DNA with a high G + C content." FEMS Microbiol Lett **174**(2): 251-253.

Jeong, S. Y., K. Ishida, Y. Ito, S. Okada and M. Murakami (2003). "Bacillamide, a novel algicide from the marine bacterium, Bacillus sp SY-1, against the harmful dinoflagellate, Cochlodinium polykrikoides." Tetrahedron Lett **44**(43): 8005-8007.

Jez, J. M., J. L. Ferrer, M. E. Bowman, R. A. Dixon and J. P. Noel (2000). "Dissection of malonyl-coenzyme A decarboxylation from polyketide formation in the reaction mechanism of a plant polyketide synthase." Biochemistry **39**(5): 890-902.

Kawahara, T., J. H. Hwang, M. Izumikawa, J. Hashimoto, M. Takagi and K. Shin-ya (2012). "JBIR-129 and -139, cytotoxic 34-membered polyol macrolides of microbial origin." J Nat Prod **75**(10): 1814-1818.

Khan, S. T., H. Komaki, K. Motohashi, I. Kozone, A. Mukai, M. Takagi and K. Shin-ya (2011). "Streptomyces associated with a marine sponge Haliclona sp.; biosynthetic genes for secondary metabolites and products." Environ Microbiol **13**(2): 391-403.

Kieser, T., Bibb, M.J., Buttner, M.J., Chater, K.F. and Hopwood, D.A. (2000) Practical Streptomyces Genetics: John Innes Foundation, Norwich Research Park, Colney, Norwich NR4 7UH, UK

Kolrkmaz, C. A., E. E. Hames-Kocabas, A. Uzel and E. Bedir (2008). "Tryptamine derived amides with thiazole ring system from Thermoactinomyces strain TA66-2." Magn Reson Chem **46**(1): 80-83.

Kreutzer, M. F. and M. Nett (2012). "Genomics-driven discovery of taiwachelin, a lipopeptide siderophore from Cupriavidus taiwanensis." Org Biomol Chem **10**(47): 9338-9343.

Lampis, G., D. Deidda, C. Maullu, M. A. Madeddu, R. Pompei, F. Delle Monachie and G. Satta (1995). "Sattabacins and sattazolins: new biologically active compounds with antiviral properties extracted from a Bacillus sp." J Antibiot (Tokyo) **48**(9): 967-972.

Lautru, S., R. J. Deeth, L. M. Bailey and G. L. Challis (2005). "Discovery of a new peptide natural product by Streptomyces coelicolor genome mining." Nat Chem Biol **1**(5): 265-269.

Ledent, P., C. Duez, M. Vanhove, A. Lejeune, E. Fonze, P. Charlier, F. Rhazi-Filali, I. Thamm, G. Guillaume, B. Samyn, B. Devreese, J. Van Beeumen, J. Lamotte-Brasseur and J. M. Frere (1997). "Unexpected influence of a C-terminal-fused His-tag on the processing of an enzyme and on the kinetic and folding parameters." FEBS Lett **413**(2): 194-196.

Lin, Z., L. Marett, R. W. Hughen, M. Flores, I. Forteza, M. A. Ammon, G. P. Concepcion, S. Espino, B. M. Olivera, G. Rosenberg, M. G. Haygood, A. R. Light and E. W. Schmidt (2013). "Neuroactive diol and acyloin metabolites from cone snail-associated bacteria." Bioorg Med Chem Lett **23**(17): 4867-4869.

Merz, A., M. C. Yee, H. Szadkowski, G. Pappenberger, A. Crameri, W. P. Stemmer, C. Yanofsky and K. Kirschner (2000). "Improving the catalytic activity of a thermophilic enzyme at low temperatures." Biochemistry **39**(5): 880-889.

Ochman, H., J. G. Lawrence and E. A. Groisman (2000). "Lateral gene transfer and the nature of bacterial innovation." Nature **405**(6784): 299-304.

Ohnishi, Y., H. Yamazaki, J. Y. Kato, A. Tomono and S. Horinouchi (2005). "AdpA, a central transcriptional regulator in the A-factor regulatory cascade that leads to morphological development and secondary metabolism in Streptomyces griseus." Biosci Biotechnol Biochem **69**(3): 431-439.

Okesli, A., L. E. Cooper, E. J. Fogle and W. A. van der Donk (2011). "Nine post-translational modifications during the biosynthesis of cinnamycin." J Am Chem Soc **133**(34): 13753-13760.

Osorio, C. R., S. Juiz-Rio and M. L. Lemos (2006). "A siderophore biosynthesis gene cluster from the fish pathogen Photobacterium damselae subsp. piscicida is structurally and functionally related to the Yersinia high-pathogenicity island." Microbiology **152**(Pt 11): 3327-3341.

Pang, S. S., R. G. Duggleby and L. W. Guddat (2002). "Crystal structure of yeast acetohydroxyacid synthase: a target for herbicidal inhibitors." J Mol Biol **317**(2): 249-262.

Pang, S. S., R. G. Duggleby, R. L. Schowen and L. W. Guddat (2004). "The crystal structures of Klebsiella pneumoniae acetolactate synthase with enzyme-bound cofactor and with an unusual intermediate." J Biol Chem **279**(3): 2242-2253.

Powell, E., S. X. Huang, Y. Xu, S. R. Rajski, Y. Wang, N. Peters, S. Guo, H. E. Xu, F. M. Hoffmann, B. Shen and W. Xu (2010). "Identification and characterization of a novel estrogenic ligand actinopolymorphol A." Biochem Pharmacol **80**(8): 1221-1229.

Rajakumari, S. and G. Daum (2010). "Janus-faced enzymes yeast Tgl3p and Tgl5p catalyze lipase and acyltransferase reactions." Mol Biol Cell **21**(4): 501-510.

Rauhut, T., P. Spiteller, W. Eisenreich, M. Spiteller and E. Glawischnig (2008). "Biosynthetic origin of BE-10988 in Streptomyces sp. BA10988." J Org Chem **73**(14): 5279-5286.

Rausch, C., T. Weber, O. Kohlbacher, W. Wohlleben and D. H. Huson (2005). "Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs)." Nucleic Acids Res **33**(18): 5799-5808.

Romano, A., D. Vitullo, M. Senatore, G. Lima and V. Lanzotti (2013). "Antifungal Cyclic Lipopeptides from Bacillus amyloliquefaciens Strain BO5A." J Nat Prod **76**(11): 2019-2025.

Rottig, M., M. H. Medema, K. Blin, T. Weber, C. Rausch and O. Kohlbacher (2011). "NRPSpredictor2--a web server for predicting NRPS adenylation domain specificity." Nucleic Acids Res **39**(Web Server issue): W362-367.

Roy, R. S., A. M. Gehring, J. C. Milne, P. J. Belshaw and C. T. Walsh (1999). "Thiazole and oxazole peptides: biosynthesis and molecular machinery." Nat Prod Rep **16**(2): 249-263.

Sauguet, L., M. Moutiez, Y. Li, P. Belin, J. Seguin, M. H. Le Du, R. Thai, C. Masson, M. Fonvielle, J. L. Pernodet, J. B. Charbonnier and M. Gondry (2011). "Cyclodipeptide synthases, a family of class-I aminoacyl-tRNA synthetase-like enzymes involved in non-ribosomal peptide synthesis." Nucleic Acids Res **39**(10): 4475-4489.

Schwarzer, D., H. D. Mootz and M. A. Marahiel (2001). "Exploring the impact of different thioesterase domains for the design of hybrid peptide synthetases." Chem Biol **8**(10): 997-1010.

Smith, S. and S. C. Tsai (2007). "The type I fatty acid and polyketide synthases: a tale of two megasynthases." Nat Prod Rep **24**(5): 1041-1072.

Song, J., M. Clagett-Dame, R. E. Peterson, M. E. Hahn, W. M. Westler, R. R. Sicinski and H. F. DeLuca (2002). "A ligand for the aryl hydrocarbon receptor isolated from lung." Proc Natl Acad Sci U S A **99**(23): 14694-14699.

Stachelhaus, T., H. D. Mootz and M. A. Marahiel (1999). "The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases." Chem Biol **6**(8): 493-505.

Strieker, M., F. Kopp, C. Mahlert, L. O. Essen and M. A. Marahiel (2007). "Mechanistic and structural basis of stereospecific Cbeta-hydroxylation in calcium-dependent antibiotic, a daptomycin-type lipopeptide." ACS Chem Biol **2**(3): 187-196.

Sunazuka, T., T. Hirose, Z. M. Tian, R. Uchida, K. Shiomi, Y. Harigaya and S. Omura (1997). "Syntheses and absolute structures of novel protein farnesyltransferase inhibitors, kurasoins A and B." J Antibiot (Tokyo) **50**(5): 453-455.

Tsai, S. C. and B. D. Ames (2009). "Structural enzymology of polyketide synthases." Methods Enzymol **459**: 17-47.

Uchida, R., K. Shiomi, J. Inokoshi, R. Masuma, T. Kawakubo, H. Tanaka, Y. Iwai and S. Omura (1996). "Kurasoins A and B, new protein farnesyltransferase inhibitors produced by Paecilomyces sp. FO-3684. I. Producing strain, fermentation, isolation, and biological activities." J Antibiot (Tokyo) **49**(9): 932-934.

Uchida, R., K. Shiomi, T. Sunazuka, J. Inokoshi, A. Nishizawa, T. Hirose, H. Tanaka, Y. Iwai and S. Omura (1996). "Kurasoins A and B, new protein farnesyltransferase inhibitors produced by Paecilomyces sp. FO-3684. II. Structure elucidation and total synthesis." J Antibiot (Tokyo) **49**(9): 886-889.

Walsh, C. T. and T. A. Wencewicz (2013). "Prospects for new antibiotics: a molecule-centered perspective." J Antibiot (Tokyo).

Wang, J. G., P. K. Lee, Y. H. Dong, S. S. Pang, R. G. Duggleby, Z. M. Li and L. W. Guddat (2009). "Crystal structures of two novel sulfonylurea herbicides in complex with Arabidopsis thaliana acetohydroxyacid synthase." FEBS J **276**(5): 1282-1290.

Weiberth, F. J. and S. S. Hall (1985). "Tandem Alkylation-Reduction - Highly Stereoselective Synthesis of (E)-1-Hydroxymethyl Methyl Propenyl Ethers from Aldehydes Using 1-Lithio-1-Methoxyallene." J Org Chem **50**(25): 5308-5314.

Widdick, D. A., H. M. Dodd, P. Barraille, J. White, T. H. Stein, K. F. Chater, M. J. Gasson and M. J. Bibb (2003). "Cloning and engineering of the cinnamycin biosynthetic gene cluster from Streptomyces cinnamoneus cinnamoneus DSM 40005." Proc Natl Acad Sci U S A **100**(7): 4316-4321.

Wiegel, J. and L. G. Ljungdahl (1986). "The Importance of Thermophilic Bacteria in Biotechnology." Crit Rev Biotech **3**(1): 39-108.

Wilson, Z. E. and M. A. Brimble (2009). "Molecules derived from the extremes of life." Nat Prod Rep **26**(1): 44-71.

Yabe, S., Y. Aiba, Y. Sakai, M. Hazaka and A. Yokota (2010). "A life cycle of branched aerial mycelium- and multiple budding spore-forming bacterium Thermosporothrix hazakensis belonging to the phylum Chloroflexi." J Gen Appl Microbiol **56**(2): 137-141.

Yabe, S., Y. Aiba, Y. Sakai, M. Hazaka and A. Yokota (2010). "Thermosporothrix hazakensis gen. nov., sp. nov., isolated from compost, description of Thermosporotrichaceae fam. nov. within the class Ktedonobacteria Cavaletti et al. 2007 and emended description of the class Ktedonobacteria." Int J Syst Evol Microbiol **60**(Pt 8): 1794-1801.

Zazopoulos, E., K. Huang, A. Staffa, W. Liu, B. O. Bachmann, K. Nonaka, J. Ahlert, J. S. Thorson, B. Shen and C. M. Farnet (2003). "A genomics-guided approach for discovering and expressing cryptic metabolic pathways." Nat Biotechnol **21**(2): 187-190.

Zendo, T., J. Nakayama, K. Fujita and K. Sonomoto (2008). "Bacteriocin detection by liquid chromatography/mass spectrometry for rapid identification." J Appl Microbiol **104**(2): 499-507.

Zhao, C., Y. Luo, C. Song, Z. Liu, S. Chen, Z. Yu and M. Sun (2007). "Identification of three Zwittermicin A biosynthesis-related genes from Bacillus thuringiensis subsp. kurstaki strain YBT-1520." Arch Microbiol **187**(4): 313-319.

Zou, X. W., S. C. Liu, Z. H. Zheng, H. Zhang, X. L. Chen, X. Z. Liu and E. W. Li (2011). "Two New Imidazolone-Containing Alkaloids and Further Metabolites from the Ascomycete Fungus Tricladium sp." Chem Biodiversity **8**(10): 1914-1920.

Zuther, K., P. Mayser, U. Hettwer, W. Wu, P. Spiteller, B. L. Kindler, P. Karlovsky, C. W. Basse and J. Schirawski (2008). "The tryptophan aminotransferase Tam1 catalyses the single biosynthetic step for tryptophan-dependent pigment synthesis in Ustilago maydis." Mol Microbiol **68**(1): 152-172.

ACKNOWLEDGEMENTS

It was meaningful and precious time for me to pursue the Ph.D for last three years in the University of Tokyo, and I do believe my dissertation would not exist without such a sincere help and cooperation the wonderful people have provided. Thanks to everyone who has supported and led me to complete my doctorate course successfully. In particular, I would like to express my appreciation and acknowledgement to:

Dr. Makoto Nishiyama, Professor of the Laboratory of Cell Biotechnology (CBT), Biotechnology Research Center (BRC), the University of Tokyo for his supervision and keen consideration during this course. I would convey my deep gratitude to him to allow me to join this lab and conduct my research.

Dr. Tomohisa Kuzuyama, Associate Professor of the Lab of CBT, BRC, the University of Tokyo for sparing no efforts for his guidance, advice, and encouragement during my graduate studies, which are the key elements to conduct my research. In particular, I would appreciate his supervising the whole process of my project with expertise guidance and inspiring ideas as well as careful proof-reading of my work.

Dr. Takeo Tomida, Assistant Professor of the Lab of CBT, BRC, the University of Tokyo for his kind guidance of a laboratorial life.

Dr. Shuhei Yabe, Kennan Eisei Kogyo Co., Ltd. for affording me the opportunity to study this research by providing *Thermosporothrix hazakensis* SK20-1^{T}.

Dr. Kazuo Shin-ya, National Institute of Advanced Industrial Science and Technology (AIST), Dr. Noritaka Kagaya, Technology Research Association for Next-Generation Natural Product Chemistry, Dr. Junko Hashimoto, Japan Biological Informatics Consortium (JBIC), and Dr. Miho Izumikawa, JBIC for kindly conducting bioactivity assays.

I also express my gratitude to every single member in the Lab of CBT both past and present ever. Their help and mutual understanding among are the successful factors to enable me to accomplish this course for last three years, indeed. Thanks for experimental techniques as well as many laughs.

Finally, I express my cordial gratitude to my family, especially to my wife Yoonjung Chang for her permanent encouragement and trust to me.