Study on the Polyketides Produced by Marine-derived *Streptomyces* sp. NPS554

Toyama Prefectural University

Tao Zhou
Mar 2016

Contents

CHAPTER 1

INTRODUCTION

Natural products refer to the organic compounds isolated from natural sources that are produced from the pathways of metabolisms. In the field of organic chemistry, the definition of natural products is often further restricted to secondary metabolites. Unlike primary metabolites such as amino acids, nucleotides, lipids and carbohydrates, which are basic cellular components and required for life, secondary metabolites are dispensable and not absolutely required for survival. Nevertheless, the variety of functions that secondary metabolites possess provides the producers evolutionary advantages. Though not all of the secondary metabolites can be explained how they benefit their producers yet, some of these compounds are produced for obvious purposes such as pheromones that act as social signaling molecules, agents that solubilize and transport nutrients, and cytotoxic materials used as “chemical warfare” against competitors, prey, and predators.

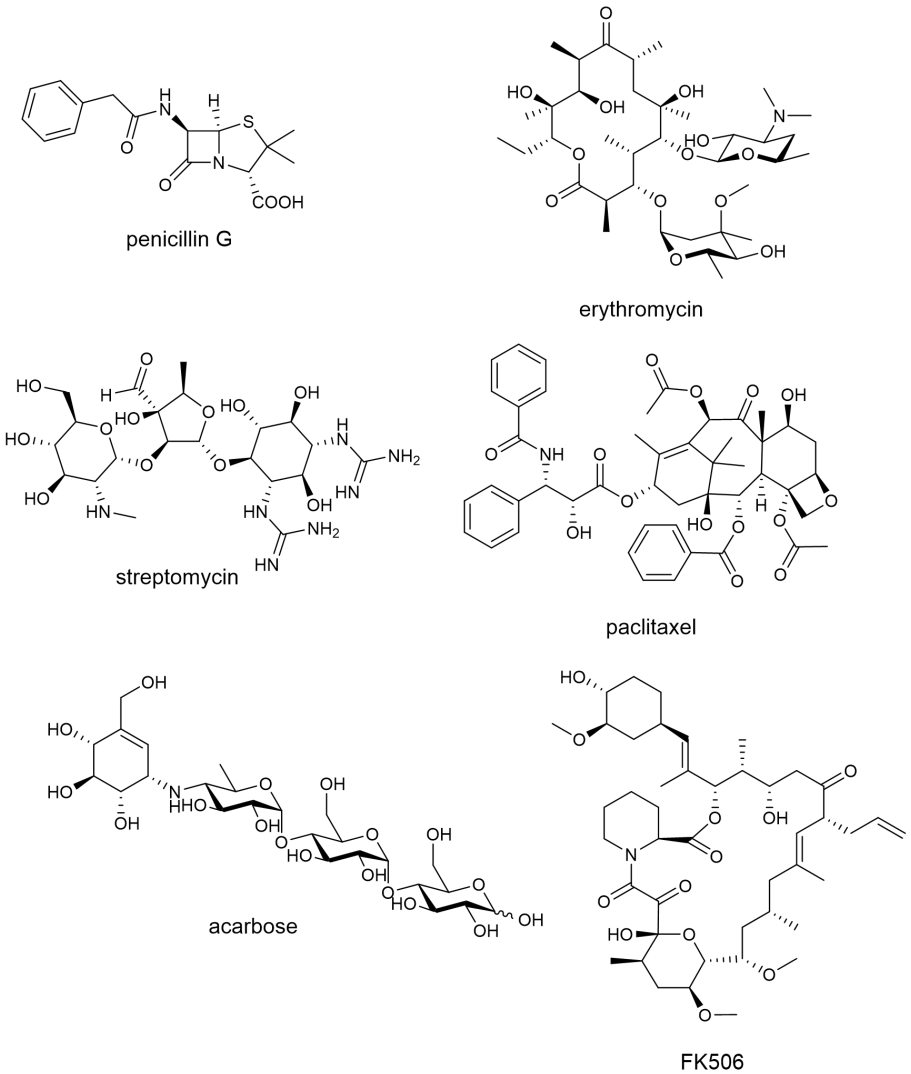


Figure 1-1. Representative natural products with pharmaceutical value.

We have used these bioactive compounds as traditional medicines for hundreds of years, even not knowing the active components and their potential. Discovery and subsequent clinical success of application of penicillin prompted a large-scale search for other environmental microorganisms that might produce anti-infective natural products: the total number of discovered compounds kept increasing exponentially in the following decades, and by the end of 2002 about 22000 bioactive

secondary metabolites were reported.[1] These natural products are widely used in drug design and development, not only as antibiotics such as penicillin, erythromycin, and streptomycin, but also for various biological activities: acarbose, an inhibitor of α-glucosidase used to treat type 2 diabetes mellitus; FK506, an immunosuppressive drug lowering the risk of organ rejection after allogeneic organ transplant; paclitaxel, interfering with the normal breakdown of microtubules during cell division that used to treat cancer (Figure 1-1).

1-1 Resources of Natural Products

The source of natural products covers prokaryotes and eukaryotes, which are almost all the living organisms, and among which actinomycetes have made magnificent contribution in producing compounds with pharmaceutical potential: about half of the naturally generated antibiotics were found from actinomycetes, and 70% of them are from *Streptomyces* (Table 1-1).[1]

Table 1-1. Number of antibiotics from microorganisms.[1]

	Antibiotics	Bioactive	Total
Bacteria	2900	900	3800
Actinomycetales			
Streptomyces	6550	1080	7630
Rare actinomycetes	2250	220	2470
Fungi	4900	3700	8600
Total	16500	6000	22500

Streptomyces is the largest genus of Actinobacteria with over 500 known species belonging to the family Streptomycetaceae. *Streptomyces* are Gram-positive and grow in various environments, with a filamentous form resembling fungi. The formation of hyphae and their differentiation into spores is unique among Gram-positives. The most interesting property of *Streptomyces* is the ability of producing diverse secondary metabolites.[2-4] The discovery of antibiotics derived from *Streptomyces* began with streptothricin found in 1942, and in the following decades, over 7500 bioactive compounds have been found from this genus, such as erythromycin (*S. fradiae*), neomycin (*S. griseus*), tetracycline (*S. rimosus*), vancomycin (*S. orientalis*), daptomycin (S. *rosesporus*), rifamycin (*S. mediterranei*), chloramphenicol (*S. venezuelae*), promycin (*S. alboniger*), lincomycin (*S. lincolnensis*), cefoxin (*S. lactamdurans*), etc (Figure 1-2).[5]

We have been benefited from natural product research since the discovery and medicinal use of antibiotics in the 1950s, and in the following decades the average lifetime of the population increased significantly with many infectious diseases almost disappeared or controllable. The upcoming technological improvements of isolation and structure determination methods increased the number of newly isolated metabolites, from which a lot of new applications in non-medical areas were developed. However, after the 'golden age' from the sixties, the scientific field has changed continuously, and by the nineties, these changes became significant. The efficiency of antibiotic research decreased, productivity of classical screening methods failed and rediscovery became more frequent than ever. Only a few new chemical structures were discovered, and the majority of new compounds were analogues of known compounds.[6,7] The continuously increasing research costs and less promising leads have resulted in the decreasing activity and productivity of

pharmaceutical industry: the introduction of new compounds into therapy and the marketing of new products have significantly decreased from 20-30 new drugs per decade to 3-4 newly marketed drugs over the past years.[8]

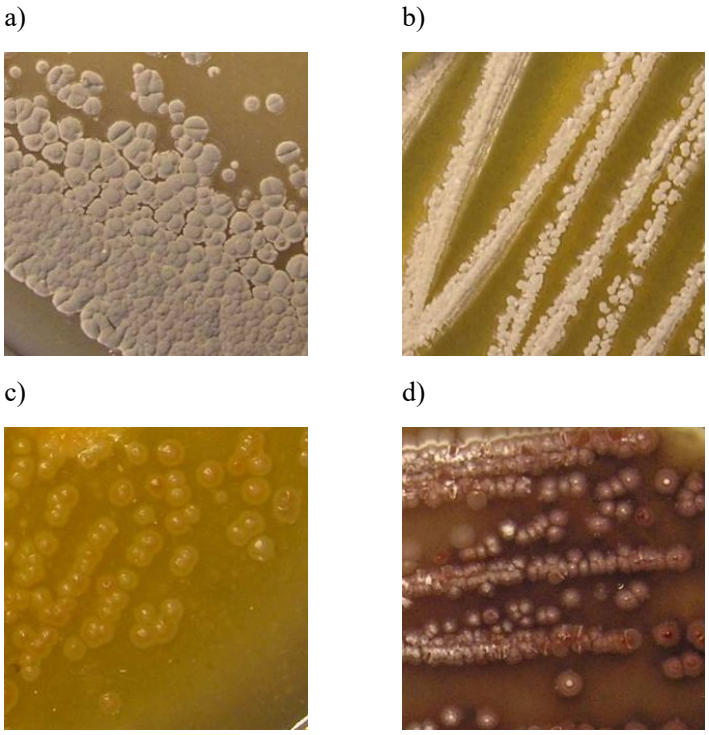


Figure 1-2. Antibiotic-producing *Streptomyces* strains.[9]
a) *S. fradiae* b) *S. rimosus* c) *S. griseus* d) *S. coelicolor*

On the other hand, combinatorial synthesis that fast developed in 1990s provides methods of preparing a large number of compounds in a single process. This strategy was once believed to be capable of replacing natural products as main resource for drug development. However, despite the increased speed of synthesis, the combinatorial synthesis has not yielded any real increase in the number of lead optimization candidates or drugs. This is probably because compared with natural products, randomly synthesized compounds lack structure variety and biological meaning. Therefore, it is believed that natural products are still the most promising source for new drug leads.[10]

From a pharmaceutical point of view, natural products are still important resource for screening bioactive compounds. Based on the statistical numbers,[11,12] the total number of marketed drugs used in human therapy is estimated to be ~3500 compounds, representing less than 0.01% of all known chemical compounds. Approximately 50% of drugs are direct natural products or derivatives of natural product scaffolds, which represent a high success rate. Of all the antimicrobial-antitumor drugs, natural product-derived drugs represent the majority. The percentage of natural antibiotic drugs among all known natural products is 0.6%, and for microbial products it is 1.6%. Approximately half (47%) of the microbial metabolites (~33000 compounds) exhibit some kind of biological activity, including antibiotic and 'other' effects, and ~40% (~28000 metabolites) are conventional antibiotics (Table 1-2).

Table 1-2. Approximate number of known synthetic compounds, all natural products and microbial products.

	Synthetic chemical compounds	Natural products	Microbial products
Numbers	8-10 millions	~500000	~70000
Drugs	2000-2500	1200-1300	450-500
Percentage	0.005	0.6	1.6

However, in contrast to this situation, the need for new bioactive compounds is never decreasing. In fact, the situation is very serious. Today, more than 70% of pathogenic bacteria are resistant to most antibiotics on the market. Reappearing 'old' pathogens such as mycobacteria and along with the new emerging ones will potentially threaten public health and bring human society back to pre-antibiotic era. The mortality of some multi-resistant infections has reached to 50%-80%. Over two million fatalities per year are due to bacterial infections. Close to two billion people carry *Mycobacterium tuberculosis*, and according to the WHO, by 2020, about 35 million people will die of tuberculosis. Malaria and HIV account for 300 million illnesses and more than five million deaths each year.

It seems that the research of natural products is trapped in a difficult position: it is quite possible that valuable discoveries are hidden deeply within the mine of natural products, but for some economic and scientific reasons the development is not prosperous. To stop the decline of microbial metabolite research and increase the effectiveness of the discoveries, the following two directions are probably the best choices for researchers:[13] unexplored new species and screening of microbial genome for new secondary metabolites.

1-2 Unexplored Microbial Species

Unexplored new microbial species are important resources for discovery of novel secondary metabolites. Since new microbial species are genetically different from the known ones, it is possible that they may carry different gene clusters for production of secondary metabolites. According to the 16S rRNA gene that is used as the standard for classification and identification of microbes, the newly isolated microbe strains can be compared within public database such as NCBI. The strains with gene similarity below 97% are very possible to be new species, and also good candidates for screening of novel natural products.

Acquiring new microbial species is an important way to approach novel products. Although a large number of actinomycetes have been isolated and screened from soil in the past few decades, recent efficiency of discovery of new metabolites from terrestrial actinomycetes has decreased.[14] Thus, it is a crucial moment that new groups of actinomycetes from unexplored habitats should be developed as sources of novel secondary metabolites, such as marine and physically or geochemically extreme environments.

The diversity of life in the terrestrial environment is extraordinary, but the greatest biodiversity occurs in the oceans.[15] The oceans cover 70% of the Earth's surface and harbor most of the planet's biodiversity. Although marine plants and invertebrates have received considerable attention as a resource for natural product discovery, the microbiological component of this diversity remains relatively unexplored. Marine sources, such as deep-sea sediments, from the seashore mud to the depths of 10000 meters are rich sources of microbes as same as soil samples.

Indeed, the marine environment is a basically untapped source of novel actinomycete diversity,[16,17] and therefore, of new metabolites.[18,19]

Although the exploitation of marine actinomycetes as a source for discovery of novel secondary metabolites is at an early stage, numerous novel metabolites have been isolated in the past few years.[20] Among them, a few compounds such as sporolides, salinosporamide A, lodopyridone, arenimycin, marinomycins and proximicins (Figue 1-3) are of particular interest due to their rarity and potent and diverse bioactivity.

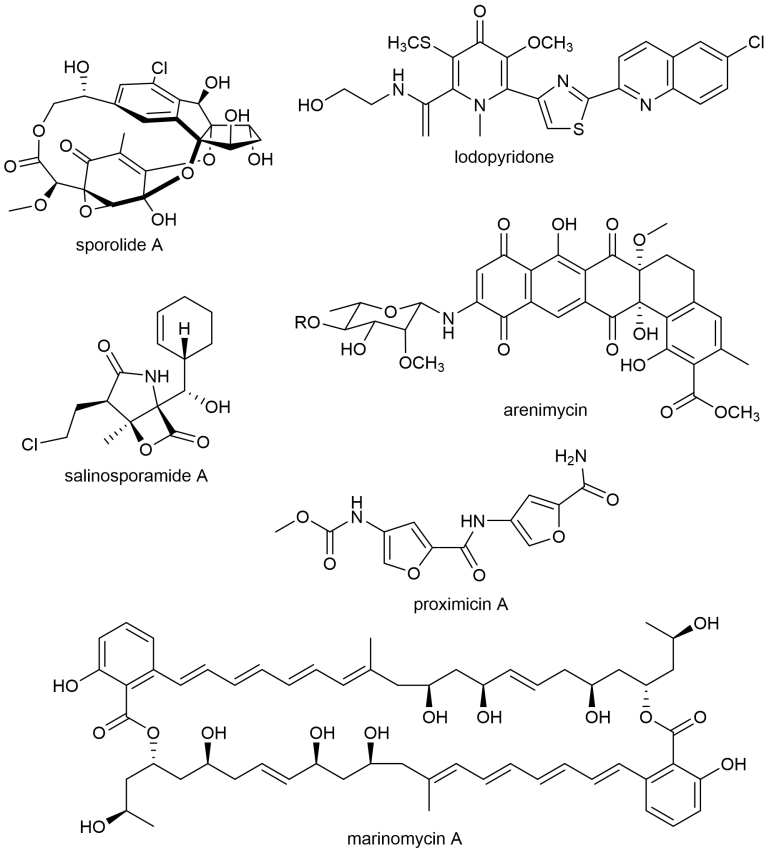


Figure 1-3. Some new secondary metabolites produced by marine actinomycetes.

Sporolides are novel polycyclic macrolides from the obligate marine bacterium *Salinispora tropica*, which is found in ocean sediment, that are composed of a chlorinated cyclopenta[a]indene ring and a cyclohexenone moiety.[21] Salinosporamide A is a novel bicyclic β-lactone γ-lactam isolated from an obligate marine actinomycete, *Salinispora tropica*.[22,23] This compound is a potent proteasome inhibitor being studied as a potential anticancer agent.[24] Lodopyridone is a unique alkaloid produced by a marine *Saccharomonospora* sp. isolated from marine sediment collected at the mouth of the La Jolla Submarine Canyon.[25] Lodopyridone possesses activity against the human colon adenocarcinoma cell line HCT-116 with an IC_{50} of 3.6 μM. Arenimycin is a new antibiotic belonging to the benzo[α]naphthacene quinone class produced by the obligate marine actinomycete, *Salinispora arenicola*.[26] Arenimycin has effective antibacterial activity against rifampin- and methicillin-resistant *Staphylococcus aureus* and exhibits potent antimicrobial activities against drug-resistant *Staphylococci* and other Gram-positive human pathogens. Marinomycins were isolated from a marine actinomycete, *Marinispora* sp., which are unusual

macrodiolides composed of dimeric 2-hydroxy-6-alkenyl-benzoic acid lactones with conjugated tetraene-pentahydroxy polyketide chains.[27] Marinomycins show significant antimicrobial activities against drug resistant bacterial pathogens and demonstrate impressive and selective cancer cell cytotoxicities against six melanoma cell lines. Proximicins were isolated from the marine *Verrucosispora* strain MG-37 with characteristic structural element of 4-amino-furan-2-carboxylic acid.[28] Proximicins show a weak antibacterial activity but a strong cytostatic effect to various human tumor cell lines.

Considering the immense biological diversity in the sea, it is increasingly recognized that a large number of novel chemical entities exists in the oceans. As marine microorganisms, particularly actinomycetes, have evolved with great genomic and metabolic diversity,[29] more efforts should be focused on exploring marine actinomycetes as a source of novel secondary metabolites.

1-3 Polyketides are Structurally Diverse

According to the characterization of biosynthetic pathways, microbial secondary metabolites can be divided into different types, such as polyketides, polypeptides, terpenoids, and glycosides. Polyketides are structurally characterized as modified fatty acids or aromatic compounds, which are widely used in pharmaceutical products, such as clarithromycin, an effective antibiotic against pylori infection; avermectin for treatment of parasitic worms; tacrolimus as an immunosuppressive drug. Polyketides are structurally special for their high diversity in functional groups, chiral centers, and ring systems, which are considered as a leading resource for discovery of pharmaceutical compounds.

The variety of bioactivity of secondary metabolites essentially rises their diversity in chemical structures. Polyketides represent a highly diverse group of natural products with chemical diversity, and carbon skeletons of which can be basically classified into polyphenols, macrolides, polyenes, enediynes, and polyethers. From a pharmacological point of view, polyketide is a wealthy source of novel therapeutics:[30] antibiotics, immunosuppressants, antiparasitics, cholesterol-lowering, and antitumoral agents, particularly.[31] It is astonishing that the vast structural and functional diversity is derived from the controlled assembly of some of the simplest biosynthetic building blocks: acetate and propionate. In general, polyketides are constructed by repetitive decarboxylative Claisen thioester condensations of an activated acyl starter unit with (methyl)malonyl-CoA-derived extender units.[32] Usually, this process involves a β-ketoacylsynthase (KS), an optional (malonyl)acyl transferase (MAT/AT), and a phosphopantethienylated acyl carrier protein (ACP) or coenzyme A (CoA) which fixes the growing chain. At the end of each elongation step, the β-oxo functionality is processed by a set of optional reductive steps, such as ketoreductase (KR), dehydratase (DH), and an enoyl reductase (ER), which give rise to a complex pattern of functionalization of polyketides. (Figure 1-4)

From the aspect of biosynthetic architecture and the pattern of processing biosynthesis of the enzymatic assembly lines, PKSs can be classified into different types (Table 1-3).[33,34] Type I refers to large multifunctional enzymes composed of linearly arranged and covalently linked catalytic domains, whereas the term type II indicates individual enzymes produced discretedly and usually monofunctional. Furthermore, a third group of multifunctional enzymes of the chalcone synthase type is denoted as type III PKS.

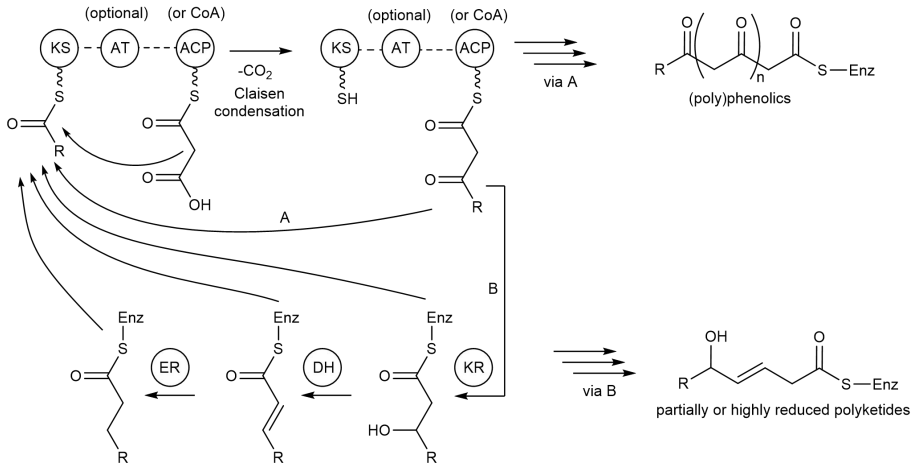


Figure 1-4. Basic mechanisms involved in polyketide biosynthesis.

Table 1-3. Representative types of PKSs.

PKS type	Building blocks	Organisms
Modular type I (non-iterative); subtypes:cis-AT, trans-AT	ACP, various extender units; (in situ methylation possible)	bacteria (protists)
Iterative type I subtypes: NR-, PR-, HR-PKS	ACP, only malonyl-CoA extenders (in situ methylation possible)	mainly fungi, some bacteria
(Iterative) type II	ACP, only malonyl-CoA extenders	exclusively bacteria
(Iterative) type III	Acyl-CoA, only malonyl-CoA extenders	mainly plants, some bacteria and fungi
PKS-NRPS hybrid	ACP, malonyl-CoA, amino acids	bacteria (modular) fungi (iterative)

Apart from the type I to III classification by the enzymes or enzyme complexes, the PKSs are further categorized as iterative or non-iterative, that is, whether the KS domain catalyzes more than one round of elongation. Non-iterative type I PKSs, such as the archetypal erythromycin PKS, 6-deoxyerythronolide (6-dEB) synthase (DEBS) are giant multimodular megasynthases which are mainly found in prokaryotes (Figure 1-5).[35] It is only recently that non-iterative PKSs have been reported in protozoans, such as dinoflagellates.[36-38] In non-iterative PKSs, KS, AT, and ACP domains, along with optional β-keto processing domains constitute a module. Generally, each module is responsible for only a single elongation cycle. Therefore, the number of the modules is consistent with the number of extension cycles executed by the PKS, and the presence of KR, DH, and ER domains determines the degree of β-keto processing.[39-41]

Iterative type I PKSs is a distinctive character of fungal polyketide biosynthesis, such as the lovastatin synthase (Figure 1-6).[42,43] Although the chain elongation is processed as an iterative way, the degree of reduction in each extension unit can be different, because the KR, DH, ER, and even methyl transferase (MT) domains are optionally applied in every elongation step.

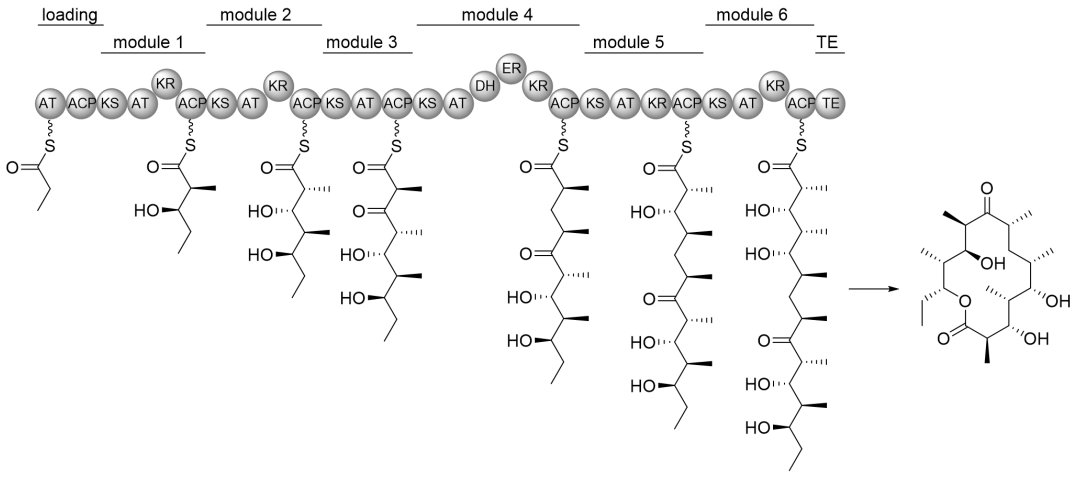


Figure 1-5. The deoxyerythronolide-B-synthase (DEBS) required for erythromycin biosynthesis as an example of a modular type I PKS.

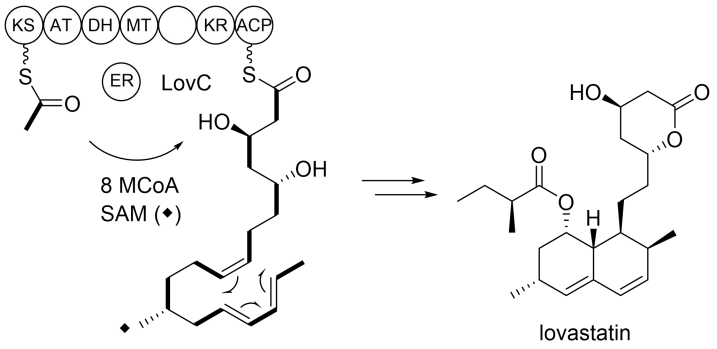


Figure 1-6. Iterative PKSs involved in the biosynthesis of lovastatin.
SAM: *S*-adenosylmethionine.

Type II PKSs are mainly found in actinomycetes, and only two examples from Gram-negative bacteria are known.[44,45] In type II PKS, the enzymes for polyketide assembly are expressed from a distinct gene and compose a minimal set of PKS, such as doxorubicin (Figure 1-6). This system is composed by two ketosynthase units (KS_α and KS_β), an ACP that serves in tethering the growing chain, and additional PKS subunits, including ketoreductases, cyclases (CYC), and aromatases (ARO).[46]

It is impressive that based on the sophisticated logic of PKSs, numerous structurally complicated compounds are synthesized. Therefore in this thesis, I selected polyketides as target compounds for study.

1-4 Screening Methodology

There are in principle two approaches of screening for new bioactive compounds from natural resource, which are referred to ‘culture-dependent’ and ‘culture-independent’ methods. Literally, in a ‘culture-dependent’ screening process, the cultivation of the candidate microbial strains have to be finished first, while in a ‘culture-independent’ screening it is not necessary.

1-4-1 'Culture-Dependent' Screening

The conventional biological screening can be classified to 'culture-dependent' screening, which traces biological activity for isolation. This method has worked well in history and led to the discovery of thousands of natural products. By using biological screening, various bioactive compounds were found in our laboratory, such as xanthoepocin, arisostatins, novobiocin, watasemycins, kosinostatin, and TPU-0037-A (Figure 1-7).[47-52] This screening has been proved successful in the past fifty years, especially in the discovery of antibiotics. The compounds isolated by this method are guaranteed with certain biological activities, which is an obvious advantage. However, the effectiveness of this method seems to be getting unsatisfactory with higher rate of re-discovery than before. It is mostly because that in this method, little structural information is known before the pure compound is obtained. Since the cultivation and isolation processes are time-consuming and fund-costing, continuous re-discoveries of new compounds are unacceptable for the researchers.

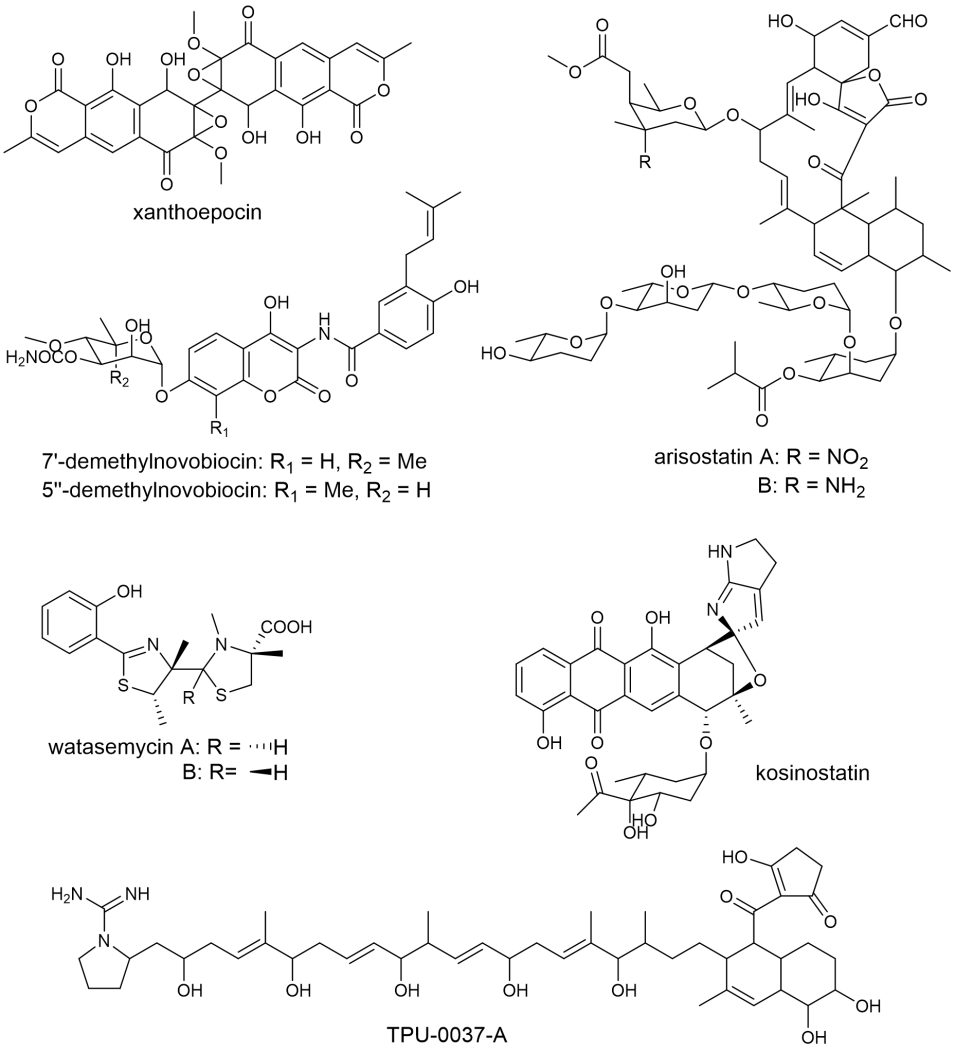


Figure 1-7. Bioactive compounds found by biological screening in our laboratory.

In order to obtain higher efficiency in discovering new compounds, it is proposed that if certain spectroscopic properties of a compound are available before isolation, the screening can be very much fastened. This is the basic principle of another 'culture-dependent' screening, which is defined as spectroscopic screening. The development of HPLC technology has made it possible to fast separate compounds from a mixture by their retention times. The down-flow can be attached to an equipment for testing their spectroscopic properties, such as UV, MS, or even NMR. In this strategy, each compound can be described by its retention time and spectroscopic properties, and these records can be used to build up a library of natural products. By comparing with the known compounds in this library, the screening process can be more efficient to select potent new compounds before isolation.

Although this method is advantageous for fast screening for new compounds, it has to be mentioned that the isolated compounds are not guaranteed for activities. The activity of the isolated compounds is unknown until the compounds were purified and tested by different assays, and sometimes no strong activities are found within them. Nevertheless, this strategy is still a powerful method that leads to the discovery of a large amount of novel natural products. Using spectroscopic screening our laboratory has successfully discovered many new microbial secondary metabolites, such as campechic acids, abyssomicin I, and jomthonic acids.[53-56]

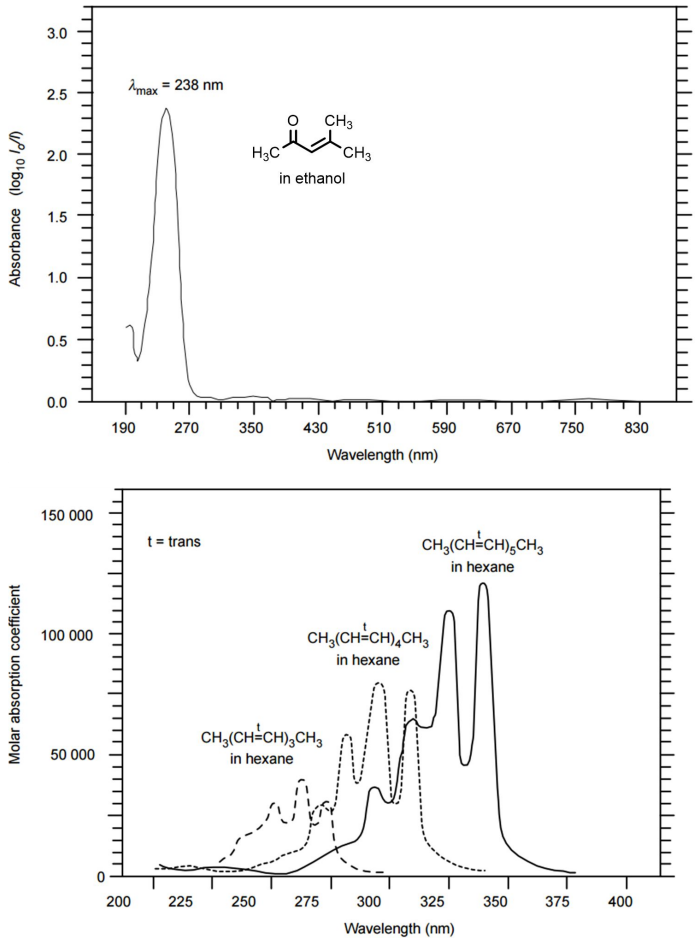


Figure 1-8. Comparison of UV absorption wavelength and conjugated system.[57]

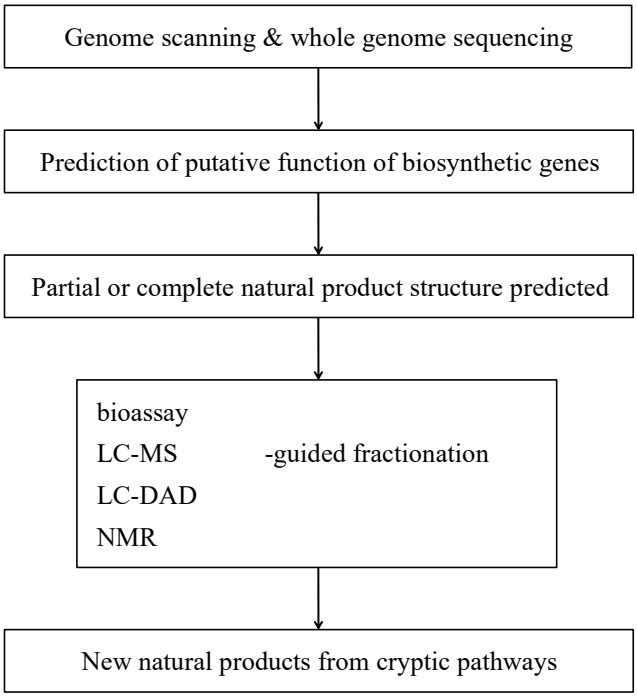
In this thesis I used UV-spectroscopic screening, because the UV absorption of an organic compound is closely related to its structure. Among the several classes of natural products, polyketides have a great chemical diversity.[58,59] In particular, compounds synthesized by bacterial type I polyketide synthases (PKSs) are attractive as a source of drug discovery because these enzymes utilize various starter and extender units to build highly branched and functionalized carbon skeletons. In contrast, type II and type III PKSs incorporate malonates as a sole substrate for chain elongation, which essentially finalize the formation of polyaromatic cores.[60-62] Most of the metabolites produced by type I PKS, except for polyenes, contain relatively small conjugated systems that display UV absorption bands around 230-300 nm, distinguishable from metabolites containing highly conjugated systems (Figure 1-8). Therefore this strategy can fast distinguish type I PKS products from other compounds.

1-4-2 'Culture-Independent' Screening

Since the relationships between polyketide compounds and gene clusters are revealed as described above, it is possible to predict the structure of the compounds from genomic information, which is the basic concept of genome mining. In this way, the screening process is mostly based on the gene sequences, which requires no cultivation of the candidate microbial strains. Therefore it is defined as 'culture-independent' strategy in this thesis.

Analyses of microbial genome sequences have revealed numerous examples of biosynthetic gene clusters with the potential to direct the production of novel, structurally complex natural products. Although microbes have great potential of producing novel compounds, it seems to be impossible to discover all of them only by 'culture-dependent' screening, because the production of secondary metabolites is controlled by gene expression regulation, which is further effected by environment. Therefore, in a 'culture-dependent' screening, one should try as many culture conditions as possible to acquire higher possibility of natural product gene expression, which is ineffective and uneconomical. It is always possible that during the 'culture-dependent' screening, some products may escape from being found. To obtain overall perspective of microbial secondary metabolites, genome mining seems to be the best choice.

The initiation of large scale microbial whole genome sequencing projects in the mid 1990s triggered a revolution in the genetics and biochemistry of natural product biosynthesis. Analyses of microbial genome sequences currently in the publicly-accessible databases have revealed numerous examples of gene clusters encoding enzymes similar to those known to be involved in the biosynthesis of many important natural products.[63] Many of these gene clusters are hypothesized to assemble novel natural products, while some others have been suggested or shown to direct assembly of known compounds via previously unexplored mechanisms. Numerous novel metabolites have been discovered as the products of such so-called 'cryptic' or 'orphan' biosynthetic gene clusters.[64-66] The approach for discovery of the metabolic products of cryptic biosynthetic gene clusters is shown in Scheme 1-1.



Scheme 1-1. Strategies for discovery of novel natural products by genome mining.

Genome mining is an important prospect of the process of secondary metabolite discovery, which has the theoretical potential to eliminate chance-dependence from secondary metabolite discovery. In comparison to the historical 'grind and find' mode of natural product discovery, the genome mining methods will be defined by the degree to which they unleash secondary metabolic gene clusters within a given system and identify encoded metabolites. In recent years, genome mining has been facilitated by the development of next generation sequencing technology.[67] Moreover, automated bio-informatics platforms now actualized the semi-automated prediction of natural products encoded by secondary metabolic blueprints.[68,69]

The importance of genome mining is more than its potential in the process of secondary metabolite discovery. Understanding the connection between metabolites and the gene sequences that encode them, can provide insight into the basic biology of producing organisms. As a relatively recent example, progress in understanding the biosynthesis of ribosomally synthesized and post-translationally modified peptides (RiPPs) has started a trend of identification of gene clusters encoding this previously barely understood class of compounds, and created an entire new category of genome mining and synthetic biology efforts.[70] There are undoubtedly many such uninvestigated systems for secondary metabolites that create new domains for biosynthesis pathways.

Many investigations of cryptic biosynthetic gene clusters have revealed that new enzymology is involved in assembly of the novel metabolic products of the clusters. Such discoveries are contributing significantly to increasing our understanding of the mechanisms for natural product biosynthesis and how such mechanisms bring about the extraordinary and still-expanding structural diversity of natural products.

References

1 Bérdy J. *J. Antibiot*. **2005**, 58, 1-26.
2 Omura S, Ikeda H, Ishikawa J, *et al*. *PNAS*, **2001**, 98, 12215-12220.
3 Khan ST, Komaki H, Motohashi K, *et al*. *Environ. Microbiol*. **2011**, 13, 391-403.
4 Patzer S, Braun V. *J. Bacteriol*. **2010**, 192, 426-435.
5 Procópio RE, Silva IR, Martins MK, *et al*. *Braz. J. Infect. Dis*. **2012**, 16, 466-471.
6 Demain AL, Sanchez S. *J. Antibiot.* **2009**, 62, 5-16.
7 Fernandes P. *Nat. Biotechnol*. **2006**, 24, 1497-1503.
8 Bérdy J. *J. Antibiot.* **2012**, 65, 385-395.
9 Available from http://lv-microbcollect.lviv.ua/en/
10 Feher M, Schmidt J. *J. Chem. Inf. Comput. Sci*. **2003**, 43, 218-227.
11 *Handbook of Antibiotic Compounds.*
12 *Dictionary of Natural Products*.
13 Baltz RH. *Curr. Opin. Pharmacol*. **2008**, 8, 557-563.
14 Fenical W, Jensen PR. *Nat. Chem. Biol*. **2006**, 2, 666-673.
15 Donia M, Hamann MT. *Lancet. Infect. Dis*. **2003**, 3, 338-348.
16 Stach JE, Maldonado LA, Ward AC, *et al. Environ. Microbiol*. **2003**, 5, 828-841.
17 Magarvey NA, Keller JM, Bernan V, *et al. Appl. Environ. Microbiol*. **2004**, 70, 7520-7529.
18 Bull AT, Stach JE, Ward AC, *et al. Antonie Van Leeuwenhoek*. **2005**, 87, 65-79.
19 Fiedler HP, Bruntner C, Bull AT, *et al. Antonie Van Leeuwenhoek*. **2005**, 87 37-42.
20 Lam KS. *Curr. Opin. Microbiol*. **2006**, 9, 245-251.
21 Buchanan GO, Williams PG, Feling RH, *et al. Org. Lett*. **2005**, 7, 2731-2734.
22 Feling RH, Buchanan GO, Mincer TJ, *et al. Angew. Chem. Int. Ed. Engl*. **2003**, 42, 355-357.
23 Jensen PR, Williams PG, Oh DC, *et al. Appl. Environ. Microbiol*. **2007**, 7, 1146-1152.
24 Chauhan D, Catley L, Li G, *et al. Cancer Cell*. **2005**, 8, 407-419.
25 Maloney KN, MacMillan JB, Kauffman CA, *et al. Org. Lett*. **2009**, 11, 5422;
26 Asolkar RN, Kirkland TN, Jensen PR, *et al. J. Antibiot*. **2010**, 63, 37-39.
27 Kwon HC, Kauffman CA, Jensen PR, *et al. J. Am. Chem. Soc*. **2006**, 128, 1622-1632.
28 Fiedler HP, Bruntner C, Riedlinger J, *et al. J. Antibiot*. **2008**, 61, 158-163.
29 Jensen PR. *J. Ind. Microbiol. Biotechnol*. **2010**, 37, 219-224.
30 Hertweck C. *Angew. Chem. Int. Ed. Engl*. **2009**, 48, 4688-4716.
31 David O'Hagan. *The Polyketide Metabolites*, Ellis Horwood Series in Organic Chemistry, **1991**.
32 Rawlings BJ. *Nat. Prod. Rep*. **1998**, 15, 275-308.
33 Hopwood DA. *Chem. Rev*. **1997**, 97, 2465-2498.
34 Staunton J, Weissman KJ. *Nat. Prod. Rep*. **2001**, 18, 380-416.
35 Rawlings BJ. *Nat. Prod. Rep*. **2001**, 18, 190-227.
36 Zhu G, LaGier MJ, Stejskal F, *et al. Gene*. **2002**, 18, 79-89.
37 Snyder RV, Gibbs PD, Palacios A, *et al. Mar. Biotechnol.* **2003**, 5, 1-12.
38 Monroe EA, Van Dolah FM. *Protist*, **2008**, 159, 471-482
39 Cane DE, Walsh CT. *Chem. Biol*. **1999**, 6, R319-R325.
40 Walsh CT. *Science*. **2004**, 19, 1805-1810.
41 Fischbach MA, Walsh CT. *Chem. Rev*. **2006**, 106, 3468-3496.
42 Schümann J, Hertweck C. *J. Biotechnol*. **2006**, 124, 690-703.

43 Cox RJ. *Org. Biomol. Chem*. **2007**, 5, 2010-2026.

44 Sandmann A, Dickschat J, Jenke-Kodama H, *et al*. *Angew. Chem. Int. Ed. Engl.* **2007**, 46, 2712-2716.

45 Brachmann AO, Joyce SA, Jenke-Kodama H, *et al*. *Chembiochem*. **2007**, 24, 1721-1728.

46 Rawlings BJ. *Nat. Prod. Rep*. **1999**, 16, 425-484.

47 Igarashi Y, Kuwamori Y, Takagi K, *et al. J. Antibiot*. **2000**, 53, 928-933.

48 Furumai T, Takagi K, Igarashi Y, *et al. J. Antibiot*. **2000**, 53, 227-232.

49 Sasaki T, Igarashi Y, Saito N, *et al. J. Antibiot*. **2001**, 54, 441-447.

50 Sasaki T, Igarashi Y, Saito N, *et al. J. Antibiot*. **2002**, 55, 249-255.

51 Furumai T, Igarashi Y, Higuchi H, *et al. J. Antibiot*. **2002**, 55, 128-133.

52 Furumai T, Eto K, Sasaki T, *et al. J. Antibiot*. **2002**, 55, 873-880.

53 Igarashi Y, Yu L, Miyanaga S, *et al. J. Nat. Prod*. **2010**, 73, 1943-1946.

54 Igarashi Y, Yu L, Ikeda M, *et al. J. Nat. Prod*. **2012**, 75, 986-990.

55 Yu L, Oikawa T, Kitani S, *et al. J. Antibiot*. **2014**, 65, 345-347.

56 Yu L, Trujillo ME, Miyanaga S, *et al. J. Nat. Prod*. **2014**, 77, 976-982.

57 Ben Faust. *Modern Chemical Techniques : An Essential Reference for Students and Teachers.* **1997**.

58 Shen B. *Curr. Opin. Chem. Biol*. **2003**, 7, 285-295.

59 Staunton J, Weissman K. *J. Nat. Prod. Rep*. **2001**, 18, 380-416.

60 Dalby SM, Paterson I. *Curr. Opin. Drug. Devel*. **2010**, 13, 777-794.

61 Jenke-Kodama H, Dittmann E. *Nat. Prod. Rep*. **2009**, 26, 874-883.

62 Van Lanen SG, Shen B. *Curr. Opin. Drug. Devel*. **2008**, 11, 186-195.

63 Donadio S, Monciardini P, Sosio M. *Nat. Prod. Rep*. **2007**, 24, 1073-1109.

64 Challis GL. *J. Med. Chem*. **2008**, 51, 2618-2628.

65 Bode HB, Müller R. *Angew. Chem. Int. Ed*. **2005**, 44, 6828-6846.

66 Gross H. *Appl. Microbiol. Biotechnol*. **2007**, 75, 267-277.

67 Didelot X, Bowden R, Wilson DJ, *et al. Nat. Rev. Genet*. **2012**, 13, 601-612.

68 Blin K, Medema MH, Kazempour D, *et al. Nucleic Acids Res*. **2013**, 41, W204-212.

69 Boddy CN. *J. Ind. Microbiol. Biotechnol*. **2014**, 41, 443-450.

70 Maksimov MO, Link AJ. *J. Ind. Microbiol. Biotechnol*. **2014**, 41, 333-344.

CHAPTER 2

Akaeolide, a Carbocyclic Polyketide from Marine-Derived *Streptomyces* sp. NPS554

2-1 Background

Because of the declining efficiency in terrestrial microbial metabolite research in recent years, marine is attracting more attention for a vast reservoir of biological and chemical diversity than ever.[1] In particular, marine *Streptomyces*, widely spread in oceanic ecosystems, has been recognized as an emerging source of new bioactive compounds.[2] In this study, a set of marine-derived actinomycetes supplied by Nippon Suisan Kaisha Ltd. was used to verify the first hypothesis that new species produce new compounds. These strains have relatively low 16S rRNA gene sequence similarity to known species, thereby being defined as a potential producer of new chemical entities as discussed in Chapter 1 (Table 2-1).

Table 2-1. Strains provided by Nippon Suisan Kaisha Ltd. with low gene similarity (<98%).

16S rDNA similarity (%)	strain	collection site	year
92	*Thermomonospora* sp. NPS595	Kochi, -28 m	2007
92	*Thermomonospora* sp. NPS954	Kochi, -28 m	2007
94	*Sacchropolyspora* sp. NPS2051	Kochi, -20 m	2007
96	*Sacchropolyspora* sp. NPS569	Kochi (sediment)	2006
96	*Streptomyces* sp. NPS568	Kochi, -3 m	2007
96	*Streptomyces* sp. NPS857	Kochi, -18 m	2007
96	*Streptomyces* sp. NPS601	Kochi, -23 m	2007
96	*Streptomyces* sp. NPS2057	Fukui, -9 m	2007
96	*Streptomyces* sp. NPS643	Kochi, -18 m	2007
96	*Streptomyces* sp. NPS2059	Fukui, -13 m	2007
97	*Streptomyces* sp. NPS553	Nobeoka, -38 m	2006
97	*Streptomyces* sp. NPS561	Okinawa (sponge)	2006
97	*Streptomyces* sp. NPS562	Nobeoka, -16 m	2006
97	*Streptomyces* sp. NPS554	Nobeoka, -38 m	2006
97	*Streptomyces* sp. NPS606	Nobeoka, -20 m	2006
97	*Streptomyces* sp. NPS544	Kochi, -15 m	2007
97	*Gordonia* sp. NPS641	Kochi, -17 m	2007
97	*Nonomuraea* sp. NPS642	Kochi (sponge)	2007
97	*Streptomyces* sp. NPS610	Ehime (sponge)	2007
97	*Actinomodura* sp. NPS664	Kochi, -13 m	2007
97	*Streptomyces* sp. NPS644	Fukui, -11 m	2007
97	*Nocardiopsis* sp. NPS640	Kochi, -20 m	2007
97	*Pseudonocardia* sp. NPS639	Kochi, -2 m	2007
97	*Nonomuraea* sp. NPS2070	Kochi, -21 m	2007
97	*Sacchropolyspora* sp. NPS2072	Kochi (sponge)	2007
97	*Nocardiopsis* sp. NPS887	Kochi, -26 m	2007
97	*Catellatospora* sp. NPS2073	Kagoshima, -52 m	2007

In this study, I designed the screening program to use HPLC/UV-based spectroscopic analysis targeted to structurally novel type I PKS products produced from a group of bacterial species.[3-5] From the actinomycete strains listed in Table 2-1, *Streptomyces* sp. NPS554 was thus selected as a candidate strain through spectroscopic screening using HPLC/DAD (diode array detector) analysis in combination with our in-house UV database dereplication (Figure 2-1). This strain was originally isolated from a sediment sample collected at a depth of 38 m near Miyazaki Harbor (previously known as Akae Harbor), Miyazaki, Japan. In the solvent extract of strain NPS554 cultured in A16 medium, three distinguishable peaks were detected at 9.2 min, 16.1 min, and 23.4 min (Figure 2-2). The UV-vis spectrum of each peak showed absorption maximum at 242 nm for 9.2 min peak, 330 nm for 16.1 min peak, and 240 nm for 23.4 min peak (Figure 2-3). It is already known that this strain produces lorneic acids A and B, linear polyketides bearing a benzene ring (Figure 2-4).[6] According to the UV spectral comparison and LC/MS analysis, the peak at 23.4 min was assigned as lorneic acid A.

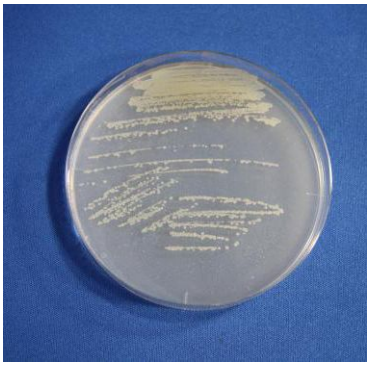


Figure 2-1. *Streptomyces* sp. NPS554 growing on Bn-2 medium.

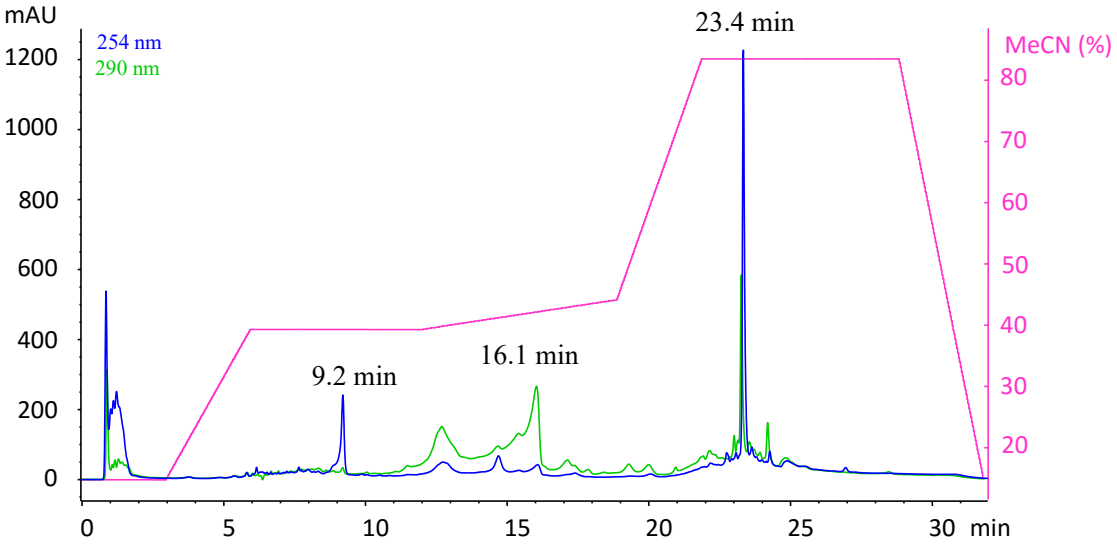


Figure 2-2. HPLC analysis of crude extract of strain NPS554.

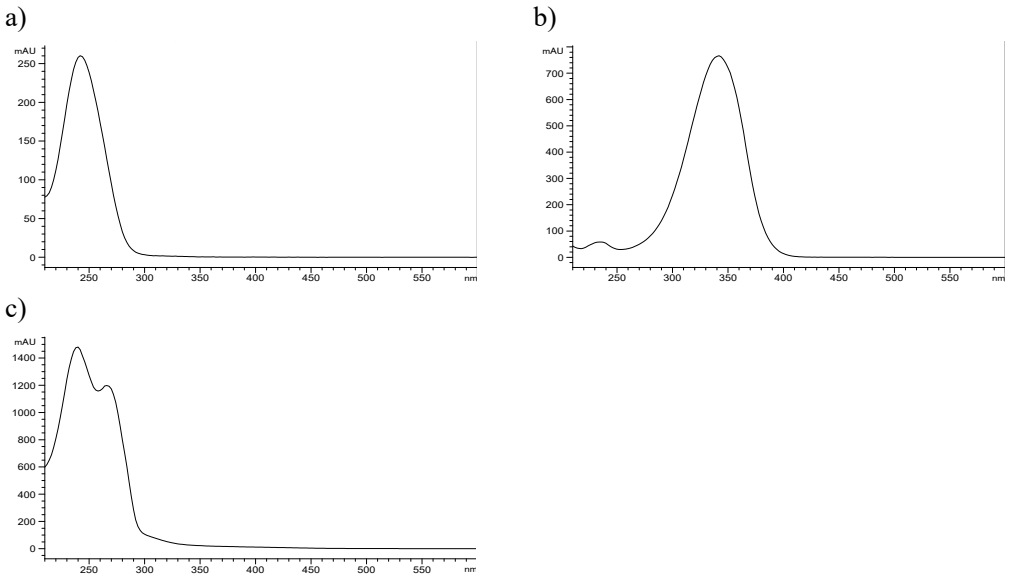


Figure 2-3. UV spectrum of the peak at a) 9.2 min; b) 16.1 min; c) 23.4 min.

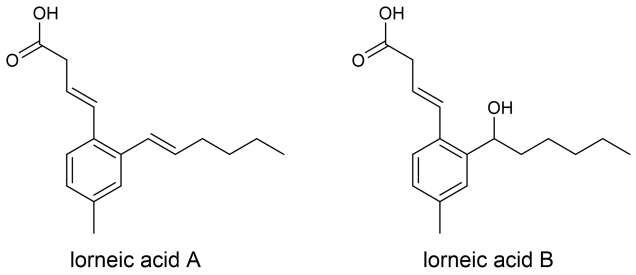


Figure 2-4. Structures of lorneic acids A and B.

Of the two unknown components with UV absorption maximum at 242 nm (9.2 min) and 330 nm (16.1 min), both of which are not present in our in-house natural product database, the compound with UV absorption at 330 nm was proposed to have a polyene structure while the one at 242 nm could possibly be a polyketide derived from type I PKS, as discussed in Chapter 1. Therefore, I selected the compound with UV absorption maximum at 242 nm as my target compound. The extract was subjected to HPLC/UV-guided purification which resulted in the isolation of akaeolide (**1**), a carbocyclic polyketide with novel carbon framework (Figure 2-5). In this chapter, the isolation, structure determination, and biological activity of **1** are described.

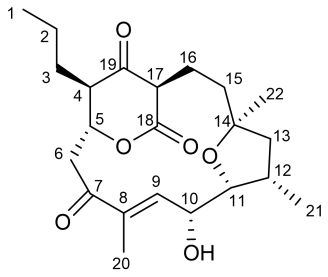


Figure 2-5. Structure of akaeolide (**1**).

2-2 Results and Discussion

2-2-1 Fermentation and Isolation

Since *Streptomyces* sp. NPS554 was isolated from marine environment, it was first tested by mediums with different salt concentration to acquire the best culture condition. Bn-2 medium was used for solid agar cultivation, and V-22 was used for broth. Both mediums were prepared by different concentration of sea water, including 0%, 30%, 60%, and 100%, and the photos were taken after 72 hours for agar culture and 48 hours for liquid culture (Figure 2-6). This strain grew well in all mediums with no obvious difference in growing speed or appearance according to salinity, except for Bn-2 fresh water agar medium culture, which showed slower growing speed than the other three.

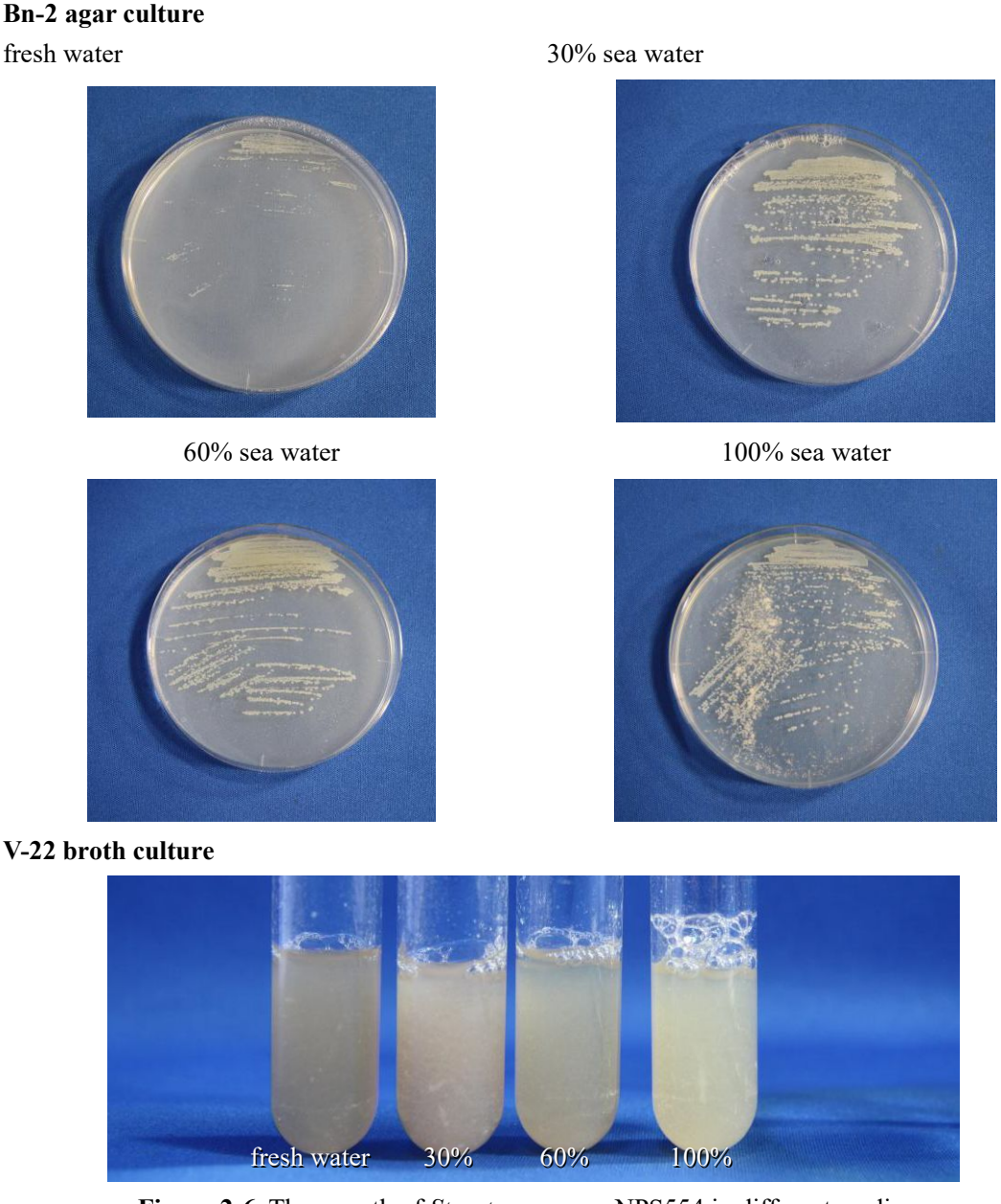


Figure 2-6. The growth of *Streptomyces* sp. NPS554 in different medium.

To check the production under different salinity, *Streptomyces* sp. NPS554 was cultured in A-16 medium prepared by fresh water and sea water respectively. The BuOH extract of the culture was analyzed by HPLC (Figure 2-7), which indicates that the production of this strain is more active in the medium prepared by sea water than fresh water. Considering the growth and product conditions, in the following experiments of this study, all mediums were prepared by sea water.

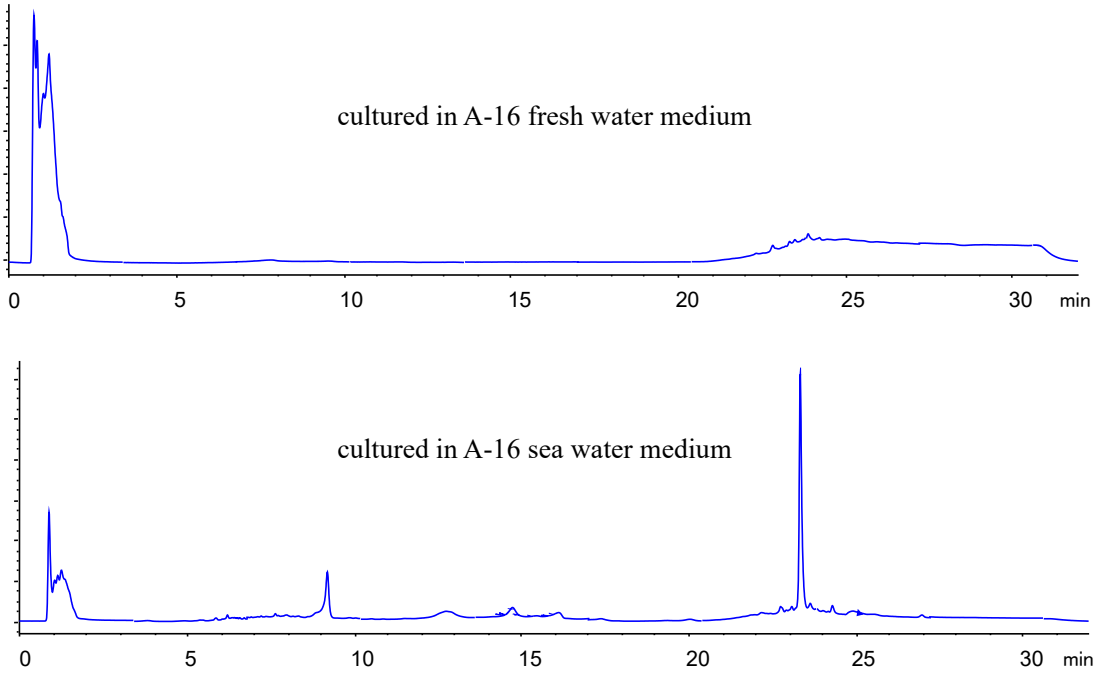
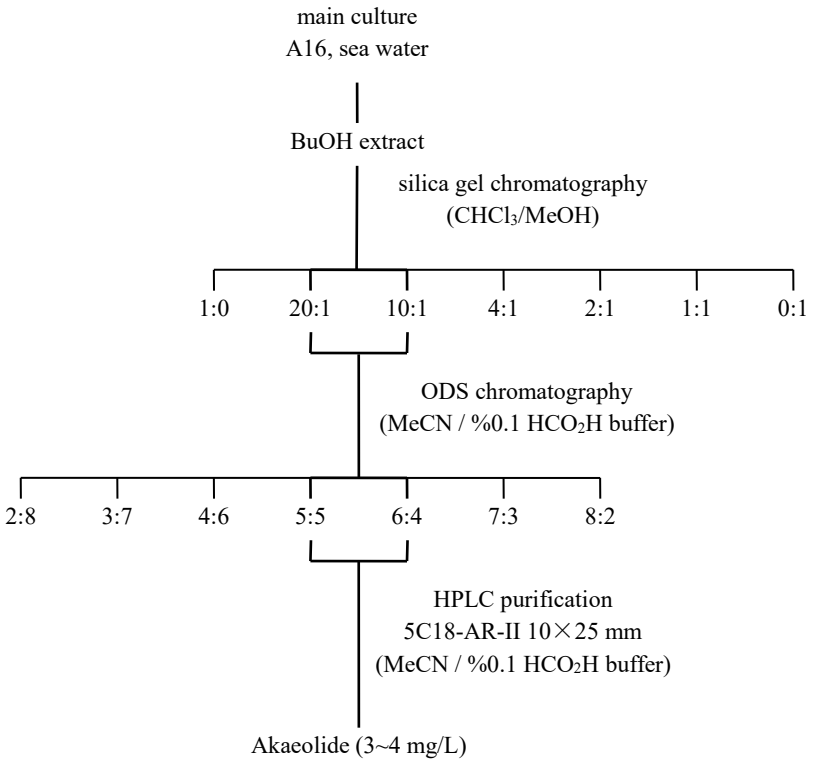


Figure 2-7. HPLC results of products from strain NPS554 cultured in different salinity. (detection at 254 nm wavelength)



Scheme 2-1. Isolation of akaeolide.

To screen for new compounds from *Streptomyces* sp. NPS554, it was first seed cultured in V-22 medium and then be transferred into A-16M medium for production. The whole culture broth was extracted with 1-butanol and the crude extract was subjected to silica gel and ODS column chromatography. The HPLC purification gave akaeolide (**1**) with the yield of 3~4 mg from 1 L culture (Scheme 2-1).

2-2-2 Structure Determination of Akaeolide

2-2-2-1 Planar Structure

Akaeolide (**1**) was isolated as an optically active colorless solid ($[\alpha]_D$ +33, *c* 0.10, $CHCl_3$). The molecular formula of **1** was determined as $C_{22}H_{32}O_6$ by high-resolution ESITOFMS analysis that showed a pseudomolecular ion at *m/z* 391.2127 $[M-H]^-$ (Δ +0.1 mmu, calcd for $C_{22}H_{31}O_6$). The existence of carbonyl groups were confirmed by IR spectrum, which displayed the absorption bands at 1712 and 1660 cm^{-1}. An absorption band of bathochromic shift at 393 nm was observed on the UV spectrum measured in alkaline solution but not in acidic or neutral solution, suggesting the presence of a phenolic or enolic functional group.[7]

From the ^1H and ^{13}C NMR spectra along with the HSQC analysis, the presence of 22 carbons were confirmed, which accounts for three oxygen-bearing quaternary sp^2 carbons, one quaternary sp^2 carbon, one sp^2 methine, one oxygenated quaternary sp^3 carbon, six sp^3 methylene, six sp^3 methine (three are oxygenenated), and four methyl groups. The degree of unsaturation was calculated as seven, whereas three carbonyls and one C-C double bond accounted four equivalently. Therefore **1** must possess three rings to satisfy the molecular formula.

Interpretation of 2D NMR spectroscopic data allowed the assemble of three partial structures: a *β*-keto-*δ*-lactone, an *α*,*β*-unsaturated ketone, and a tetrahydrofuran unit (Figure 2-8, Table 2-2). A four-carbon fragment, deduced from a series of COSY correlations from the methyl proton triplet H-1 to a methine H-4, was joined with an another COSY-defined fragment H-5/H-6 by HMBC correlations from H-5 and H-6 to C-4, H-4 to C-6, and H-3 to C-5 to provide a six-carbon fragment with an oxygen substitution at C-5. A COSY-cross peak was observed between H-15 and H-16, and the latter protons were further correlated to H-17. This three-carbon fragment was expanded to include two carbonyl carbons C-18 (δ_C 169.7) and C-19 (δ_C 203.7), each connecting at C-17 on the basis of HMBC correlations from H-16 and H-17 to these carbons. This fragment and the aforementioned six-carbon fragment were joined on the basis of HMBC correlations from H-5 to C-18 and H-5 and H-3 to C-19, establishing a *β*-keto-*δ*-lactone substructure bearing three aliphatic substituents.

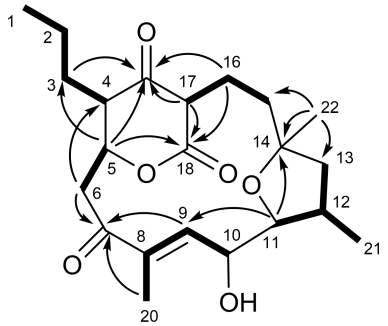


Figure 2-8. COSY and key HMBC correlations for **1**.

Table 2-2. ^{1}H and ^{13}C NMR Data for Akaeolide (**1**) in $CDCl_3$.

position	δ_C[a]	δ_H mult (*J* in Hz)[b]	HMBC[b,c]
1	13.6, CH_3	0.91, t (7.0)	2, 3
2	19.9, CH_2	1.38[d], m	
3	32.2, CH_2	1.56, m	1, 2, 4, 5, 19
		1.53, m	1, 2, 4, 5, 19
4	50.50.3, CH	2.49[d], m	3, 6, 19
5	76.7, CH	4.85, dd (12.0, 5.5)	3, 4, 6, 18, 19
6	41.5, CH_2	3.77, dd (12.0, 12.0)	4, 5, 7
		2.43, dd(12.0, 5.5)	5, 7, 8
7	200.1, qC		
8	136.2, qC		
9	146.4, CH	6.51, d (8.5)	7, 8, 10, 11, 20
10	68.3, CH	4.64, d (8.5)	8, 9
11	82.2, CH	3.66, d (7.0)	9, 10, 14, 21
12	35.8, CH	2.53, m	13, 21
13	49.3, CH_2	2.14, dd (12.5, 9.0)	11, 12, 15, 21, 22
		1.42[d], m	
14	83.0, qC		
15	31.8, CH_2	2.49[d], m	22
		1.37[d], m	
16	20.2, CH_2	2.50[d], m	14, 15, 17, 18, 19
		1.71,m	14, 17, 18
17	54.2, CH	4.59, dd (9.5, 1.5)	15, 16, 18, 19
18	169.7, qC		
19	203.7, qC		
20	11.8, CH_3	1.83, s	7, 8, 9
21	18.0, CH_3	1.19, d (7.5)	11, 12, 13
22	25.3, CH_3	1.43, s	13, 14, 15

[a] Recorded at 125 MHz.

[b] Recorded at 500 MHz.

[c] HMBC correlations are from proton(s) stated to the indicated carbon.

[d] Overlapping signals.

Structure elucidation of the second partial structure was started from the vinylic methyl proton singlet H-20 that showed HMBC correlations to the carbonyl carbon C-7 (δ_C 200.1) and olefinic carbons C-8 and C-9. The proton attached to this latter carbon C-9 showed a COSY correlation to the oxymethine proton H-10, establishing this partial structure as an *α,β*-unsaturated ketone with a methyl and an oxygenated methine substitutions at the *α*- and *β*-positions, respectively. An HMBC correlation from H-6 to C-7 allowed this fragment to be connected to the *δ*-lactone unit through C-6.

The tetrahydrofuran unit was elucidated starting from the tertiary methyl proton singlet (H-22) that showed HMBC correlations to C-13, C-14, and C-15. The proton attached to C-13 showed a COSY correlation to a methine proton H-12, which was in turn correlated to a methyl proton

doublet (H-21) and an oxygenated methine (H-11). Finally, an HMBC correlation from H-11 to C-14 established a tetrahydrofuran ring being connected to the δ-lactone unit at C-15. Further HMBC correlations observed from H-10 and H-11 to C-9 and from the olefinic proton attached to this carbon back to C-10 and C-11 completed the structure assignment of **1**.

The planar structure of **1** revealed that the methine proton H-17 flanked by two carbonyl carbons can be dissociated giving the enolized isomer. This explains that the ^{1}H NMR spectrum of **1** in $CDCl_3$ displayed the resonances for at least three distinct isomers, and during the overnight NMR measurement at 27 ˚C, these isomers converged into one dominant isomer although small peaks for minor isomers were still remaining (Figure S1). Further detailed analysis on tautomerizaion using NMR revealed that the enol-form was the major isomer in pyridine-d_5 (Figure 2-9). In fact more than three isomers were recognized in the ^{1}H and ^{13}C-NMR spectra of **1**. In addition to the two keto-enol tautomers defined in this study, other isomeric forms possibly including conformational isomers may exist.

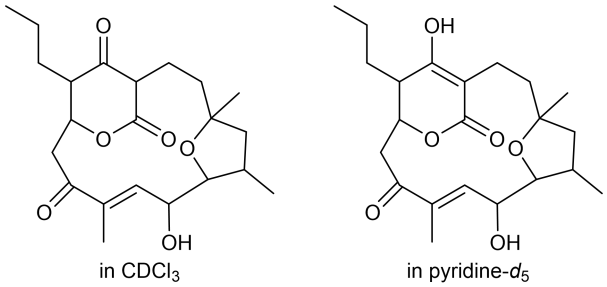


Figure 2-9. Structures of dominant tautomers of akaeolide (**1**) in $CDCl_3$ and pyridine-d_5.

2-2-2-2 Absolute Configuration

In order to establish the relative configuration of **1**, crystallization was attempted in various solvents but failed to give crystalline solid. As the tautomeric characteristic of this molecule was likely affecting negatively the crystal formation, the configuration of the enolizable carbon C-17 was fixed by chlorinating this carbon by the treatment with *N*-chlorosuccinimide in CH_2Cl_2.[8,9] The reaction completed in 10 min at room temperature to give a derivative (**2**) selectively chlorinated at C-17 as a single product. After purification by flash chromatography, **2** was crystallized from a mixture of diisopropyl ether and methanol to give plate crystals suitable to X-ray crystallographic analysis. On the basis of the diffraction anisotropy of the chlorine atom the absolute configurations of all seven asymmetric centers in **2** were determined as 4*R*, 5*R*, 10*R*, 11*R*, 12*S*, 14*S*, 17*R*, defining the absolute configuration of **1** except for the chlorinated carbon C-17 (Figures 2-10 and 2-11).

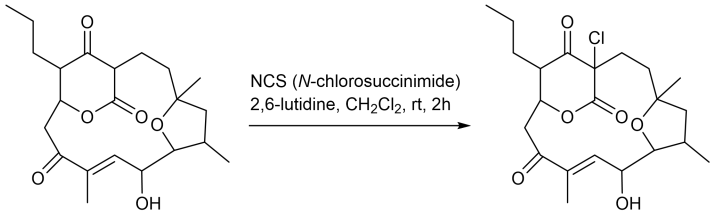


Figure 2-10. Chlorination of akaeolide (**1**) to yield 17-chloroakaeolide (**2**).

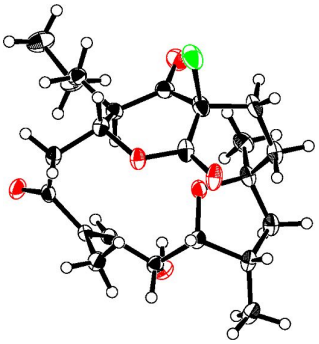


Figure 2-11. ORTEP drawing of crystal structure of **2**.

The relative configuration at C-17 of **1** in $CDCl_3$ was determined by analyzing NOESY spectrum that gave cross peaks for H-17/H-6 (δ_H 3.77), H-17/H-9, H-6 (δ_H 3.77)/H-9. These correlations allowed the placement of the H-17 methine proton and the carbon chain branched at C-5 on the same side of the *δ*-lactone ring, establishing the C-17 configuration of **1** as *R* (Figure 2-12).

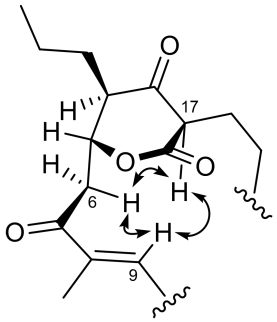


Figure 2-12. Key NOEs observed for **1** in $CDCl_3$.

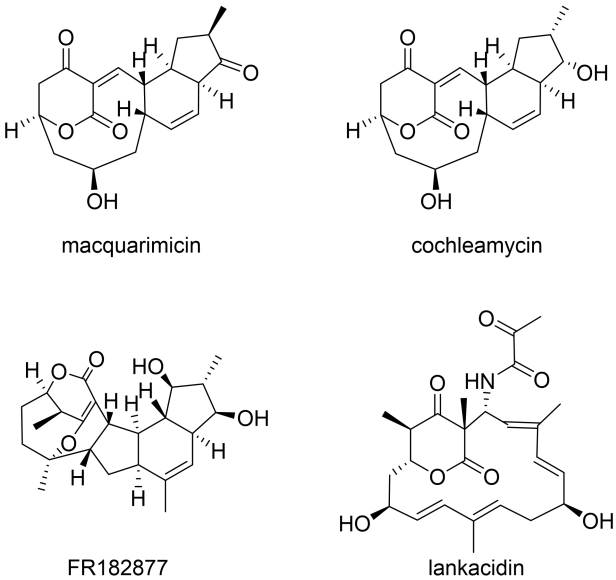


Figure 2-13. Structures of macquarimicin, cochleamycin, and lankacidin.

Akaeolide (**1**) is featured by its 15-membered carbocyclic structure functionalized with a five-membered cyclic ether and a *β*-keto-*δ*-lactone unit. As I searched for structurally related compounds of **1** in SciFinder database, mangromicin B was found to have a similar planar structure with compound **1**, which is recorded in a patent.[10] However, according to the opposite optical rotation (mangromicin B: $[\alpha]_D$ -24, *c* 0.1, MeOH vs **1**: $[\alpha]_D$ +35, *c* 0.1, MeOH) and different NMR spectral data of mangromicin B, I concluded that akaeolide and mangromicin B are different compounds.

Besides mangromicin B, only limited number of carbocyclic polyketide compounds were reported to have a *β*-keto-*δ*-lactone unit in their structures, including macquarimicin, cochleamycin, FR182877, and lankacidin (Figure 2-13).[11-15] These compounds might have a close biosynthetic relationship to **1**.

2-2-3 Bioactivity

Limited biological testing has shown that **1** is active against *Micrococcus luteus* with an MIC value of 25 *μ*g/mL but is inactive to *Escherichia coli* and *Candida albicans*. In addition, **1** displayed modest cytotoxicity to 3Y1 rat fibroblasts with an IC_{50} value of 8.5 ± 1.5 *μ*M. **1** was not active in adipocyte differentiation assay[16] and also did not show inhibitory effects on *Staphylococcus aureus* and *Enterococcus faecalis* quorum sensing signaling.[17]

2-3 Experimental Section

General experimental procedures. Optical rotation was measured using a JASCO DIP-3000 polarimeter. UV spectrum was recorded on a Hitachi U-3210 spectrophotometer. IR spectrum was measured on a Perkin Elmer Spectrum 100. NMR spectra were obtained on a Bruker AVANCE 500 spectrometer, using the signals of the residual solvent protons (δ 7.26 for $CDCl_3$; δ 7.22 for pyridine-d_5) and carbons (δ 77.0 for $CDCl_3$; δ 123.9 for pyridine-d_5) as internal standards. High-resolution ESITOFMS was recorded on a Bruker microTOF focus. X-ray crystallographic analysis was performed on a Rigaku VariMax-DW with RAPID system. Silica gel 60-C18 (Nakalai Tesque 250-350 mesh) was used for ODS column chromatography. HPLC separation was performed using a Cosmosil 5C18-AR-II (Nacalai Tesque Inc., 20 × 250 mm) with a photodiode array detector.

Isolation of akaeolide (1). Strain NPS554 cultured on a Bn-2 slant [soluble starch 0.5%, glucose 0.5%, meat extract (Kyokuto Pharmaceutical Industrial Co., Ltd.) 0.1%, yeast extract (Difco Laboratories) 0.1%, NZ-case (Wako Chemicals USA, Inc.) 0.2%, NaCl 0.2%, $CaCO_3$ 0.1%, agar 1.5%] was inoculated into 500 mL K-1 flasks each containing 100 mL of the V-22 seed medium consisting of soluble starch 1%, glucose 0.5%, NZ-case 0.3%, yeast extract 0.2%, Tryptone (Difco Laboratories) 0.5%, K_2HPO_4 0.1%, $MgSO_4 \cdot 7H_2O$ 0.05%, and $CaCO_3$ 0.3% (pH 7.0) in natural seawater. The flasks were placed on a rotary shaker (200 rpm) at 30 ˚C for 4 days. The seed culture (3 mL) was transferred into 500 mL K-1 flasks each containing 100 mL of the A-16 production medium consisting of glucose 2%, Pharmamedia (Traders Protein) 1%, and $CaCO_3$ 0.5% in natural seawater. The inoculated flasks were placed on a rotary shaker (200 rpm) at 30 ˚C for 6 days. At the end of the fermentation period, 100 mL of 1-butanol was added to each flask,

and the flasks were allowed to shake for 1 h. The mixture was centrifuged at 5000 rpm for 10 min, and the organic layer was separated from the aqueous layer containing the mycelium. Evaporation of the solvent gave 7.6 g of extract from 2.3 L of culture. The crude extract (7.6 g) was subjected to silica gel column chromatography with a step gradient of $CHCl_3$/MeOH (1:0, 20:1, 10:1, 4:1, 2:1, 1:1, and 0:1 v/v). Fraction 2 ($CHCl_3$/MeOH = 20:1) was concentrated to provide 0.22 g of brown solid, which was further purified by repeated reverse phase preparative HPLC using a Cosmosil 5C18-AR-II column (Nacalai Tesque Inc., 10 × 250 mm) with MeCN in 0.1% HCO_2H (35:65, flow rate 4 mL/min), followed by evaporation and extraction with EtOAc to give akaeolide (**1**, 9.7 mg, t_R = 19.7 min).

HPLC analysis condition for broth extract.
Instrument: Agilent HP-1100 HPLC system
Column: Microsorb-MV^{TM} (Rainin Instrument Company Inc.) C-18 3 μm (4.6 × 75 mm)
Solvent: MeCN-0.15% KH_2PO_4 buffer (pH 3.5)
MeCN: 0-3 min: 15-15%, 3-6 min: 15-40%, 6-12 min: 40-40%, 12-19 min: 40-45%, 19-22 min: 45-85%, 22-29 min: 85-85%, 29-32 min: 85-15%
Flow rate: 1.2 mL/min
Absorbance read by photodiode array detector at 200-600 nm
Akaeolide (1): colorless amorphous solid; $[\alpha]_D^{23}$ +35 (*c* 0.10, MeOH); $[\alpha]_D^{22}$ +33 (*c* 0.10, $CHCl_3$); UV (MeOH) λ_{max} (log ε) 234 (4.61), 283 (4.31) nm; (0.01 N HCl-MeOH) 235 (4.62); (0.01 N NaOH-MeOH) 234 (4.54), 283 (4.61), 393 (4.32); IR (ATR) ν_{max} 1712, 1660 cm^{-1}; For ^{1}H and ^{13}C NMR data, see Table 1; high-resolution ESITOFMS *m/z* 391.2127 $[M-H]^-$ (calcd for $C_{22}H_{31}O_6$, *m/z* 391.2126).

Chlorination of 1 to yield 17-chloroakaeolide (2). To a solution of **1** (5.2 mg, 13.3 μmol) in CH_2Cl_2 (1 mL) were added 2,6-lutidine (0.82 μL, 7.1 μmol) and *N*-chlorosuccinimide (3.6 mg, 26.8 μmol) at room temperature. After stirring for 2 h, the reaction mixture was diluted with diethyl ether (2.5 mL) and washed with sat. NaCl solution (1 mL). The organic layer was separated and concentrated in vacuo to give 5.8 mg of crude residue. Purification by silica gel chromatography (*n*-hexane/EtOAc = 10:1~1:1) yielded 17-chloroakaeolide (**2**, 2.2 mg, 39% yield) as a white powder. A portion of **2** was recrystallized from a mixture of CH_2Cl_2 and diisopropyl ether to afford plate crystals for X-ray crystallographic analysis.

17-Chloroakaeolide (2): colorless plates; mp >150 ˚C (dec); $[\alpha]_D^{22}$ -118 (*c* 0.10, MeOH); UV (MeOH) λ_{max} (log ε) 235 (3.77) nm; IR ν_{max} 3459, 1722 cm^{-1}; for ^{1}H and ^{13}C NMR data, see Table S2; high-resolution ESITOFMS *m/z* 449.1701 $[M+Na]^+$ (calcd for $C_{22}H_{31}O_6{}^{35}ClNa$, *m/z* 449.1701).

Antimicrobial assay. Antimicrobial assay was carried out according to the procedures previously described.[18]

Cytotoxicity evaluation. The cytotoxicity of **1** against 3Y1 rat embryonic fibroblasts was evaluated by a microculture-scale colorimetric assay using 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) as a bioreduction dye. Cells were maintained in a 5%

CO_2-95% air atmosphere at 37 ˚C in low glucose-type (1 g/L) Dulbecco's modified Eagle medium (DMEM) supplemented with 10%v/v fetal bovine serum, 584 mg/L L-glutamine, 20 mg/mL gentamycin sulfate, and 10^5 unit/L penicillin, 100 mg/L streptomycin sulfate, and 250 μg/L amphotericin B. To each well of a 96-well culture plate was seeded a 100 μL of cell suspension prepared at a density of 6×10^3 cells/mL. After 24 h-incubation to settle the cells, serial half-log dilutions of drugs in DMEM medium were made in a separate culture plate, and 100 μL aliquots from this were added to the microcultures in triplicate. Along with **1** was tested doxorubicin hydrochloride as a positive control and dilutions of vehicle (DMSO) in DMEM as a negative control. After 72 h-incubation, 100 μL of 1 mg/mL MTT solution in PBS without Ca^{2+} and Mg^{2+} was added to each well, and the culture were incubated for 4 h. Medium was then removed by aspiration and dyed cells were solubilized with 100 μL DMSO to measure the absorbance of the solution at 540 nm. The IC_{50} value was deduced by fitting a regression curve on growth inhibitory ratios plotted along the axis of logarithmic drug concentration.

References

1 Debbab A, Aly AH, Lin WH. *Microb. Biotechnol*. **2010**, 3, 544-563.

2 Manivasagan P, Venkatesan J, Sivakumar K, *et al*. *Microbial. Res.* **2014**, 169, 262-278.

3 Igarashi Y, Yu L, Miyanaga S, *et al*. *J. Nat. Prod*. **2010**, 73, 1943-1946.

4 Igarashi Y, Kim Y, In Y, *et al*. *Org. Lett*. **2010**, 12, 3402-3405.

5 Igarashi Y, Asano D, Furihata K, *et al*. *Tetrahedron Lett*. **2012**, 53, 654-656.

6 Iwata F, Sato S, Mukai T, et al. *J. Nat. Prod.* **2009**, 72, 2046-2048.

7 Silverstein RM, Bassler GC, Morrill TC. *Spectrometric Identification of Organic Compounds*, 5th ed. John Wiley & Sons, Inc.:New York, 1991.

8 Rahn N, Kalesse M. *Angew. Chem. Int. Ed*. **2008**, 47, 597-599.

9 Hoffman, RV, Weiner WS, Maslouh N. *J. Org. Chem*. **2001**, 66, 5790-5795.

10 Omura S, Takahashi Y, Nakashima T, *et al*. Patent WO2013031239 A8. Dec 12, 2013.

11 Hochlowski JE, Mullally MM, Henry R, *et al*. *J. Antibiot*. **1995**, 48, 467-470.

12 Shindo K, Matsuoka M, Kawai H. *J. Antibiot*. **1996**, 49, 241-243.

13 Shindo K, Iijima H, Kawai H. *J. Antibiot*. **1996**, 49, 244-248.

14 Sato B, Muramatsu H, Miyauchi M, *et al*. *J. Antibiot*. **2000**, 53, 123-130.

15 Uramoto M, Otake N, Ogawa Y, *et al*. *Tetrahedron Lett*. **1969**, 2249-2254.

16 Kunimasa K, Kuranuki S, Matsuura N, *et al. Med. Chem. Lett*. **2009**, 19, 2062-2064.

17 Desouky SE, Nishiguchi K, Zendo T, *et al. Biosci. Biotechnol. Biochem*. **2013**, 77, 923-927.

18 Igarashi Y, Yu L, Miyanaga S, *et al. J. Nat. Prod*. **2010**, 73, 1943-1946.

2-4 Spectral Data

Table of Contents

Table S1. NMR data for akaeolide (**1**) (pyridine-d_5)

position	δ_C[a]	δ_H mult (J in Hz)[b]	HMBC[b,c]
1	14.3, CH_3	0.66, t (7.3)	2, 3
2	20.3, CH_2	1.35, m	1
		1.19, m	1
3	35.8, CH_2	1.68, m	2, 4, 5
4	43.3, CH	2.58[e], m	2, 3, 6
5	77.0, CH	4.93, dd (11.7, 4.8)	3, 4, 6, 18, 19
6	43.2, CH_2	4.33, dd (11.6, 11.7)	4, 5, 7
		2.31, dd (11.6, 4.8)	5, 7, 8
7	203.7, qC		
8	134.2, qC		
9	150.5, CH	7.28, dd (8.4, 0.7)	7, 20
10	69.6, CH	4.96, d (8.4)	8, 9
11	84.0, CH	4.20, dd (8.3, 1.4)	10, 21
12	36.8, CH	2.64[e], m	13, 14, 21
13	50.1, CH_2	2.04, dd (11.9, 8.1)	11, 12, 15, 21
		1.62[e], m	12, 14, 15, 21, 22
14	81.5, qC		
15	33.9, CH_2	3.13, ddd (13.3, 13.0, 3.0)	14, 17
		1.47[e], m	16, 17
16	20.6, CH_2	3.43, ddd (13.0, 12.6, 3.2)	14, 17, 18, 19
		2.62[e], m	14, 15, 17, 18, 19
17	104.0, qC		
18	165.8[d], qC		
19	165.8[d], qC		
20	12.3, CH_3	2.08, d (0.7)	7, 8, 9
21	17.1, CH_3	1.48, d (7.1)	11, 12, 13
22	25.6, CH_3	1.62, s	13, 14, 15

[a] Recorded at 125 MHz, referenced to the solvent carbon (123.9 ppm).

[b] Recorded at 500 MHz, referenced to the residual solvent proton (7.22 ppm).

[c] HMBC correlations are from proton(s) stated to the indicated carbon.

[d] Interchangeable.

[e] Overlapping signals.

Table S2. NMR data for 17-chloroakaeolide (**2**) ($CDCl_3$)

position	δ_C[a]	δ_H mult (*J* in Hz)[b]	HMBC[b,c]
1	14.3[d], CH_3	0.95, t (7.3)	2, 3
2	18.7, CH_2	1.34, m	
		1.29, m	
3	28.6, CH_2	1.93, m	
		1.62, m	
4	46.6, CH	2.81, m	2, 3, 5, 6, 19
5	75.2, CH	5.00, ddd (10.8, 5.1, 3.1)	7
6	43.2, CH_2	3.31, dd (14.6, 3.1)	4, 5, 7
		2.79, dd (14.6, 5.1)	7, 8, 18
7	202.9, qC		
8	138.7, qC		
9	140.4, CH	5.78, dd (6.2, 1.3)	7, 20
10	68.8, CH	4.41, br.t (7.3)	
11	83.9, CH	3.79, d (8.9)	10, 12, 13, 14, 21
12	34.5, CH	2.57, m	
13	48.9, CH_2	1.82, dd (12.2, 7.5)	11, 12, 14, 15
		1.48, dd (12.2, 12.0)	12, 14, 15, 21, 22
14	82.8, qC		
15	34.5, CH_2	2.42, m	14, 16
		1.42, ddd (14.6, 5.6, 3.6)	17
16	29.8, CH_2	3.08, ddd (14.7, 11.2, 5.6)	15, 17, 18, 19
		2.27, ddd (14.7, 6.1, 3.6)	14, 15, 17, 19
17	104.0, qC		
18	164.3, qC		
19	196.2, qC		
20	14.3[d], CH_3	1.98, s	7, 8, 9
21	14.3[d], CH_3	1.17, d (7.0)	11, 12, 13
22	24.9, CH_3	1.19, s	13, 14, 15

[a] Recorded at 125 MHz, referenced to the solvent carbon (123.9 ppm).

[b] Recorded at 500 MHz, referenced to the residual solvent proton (7.22 ppm).

[c] HMBC correlations are from proton(s) stated to the indicated carbon.

[d] Interchangeable.

Figure S1. ^{1}H NMR Spectrum of Akaeolide (**1**) at 500 MHz in $CDCl_3$

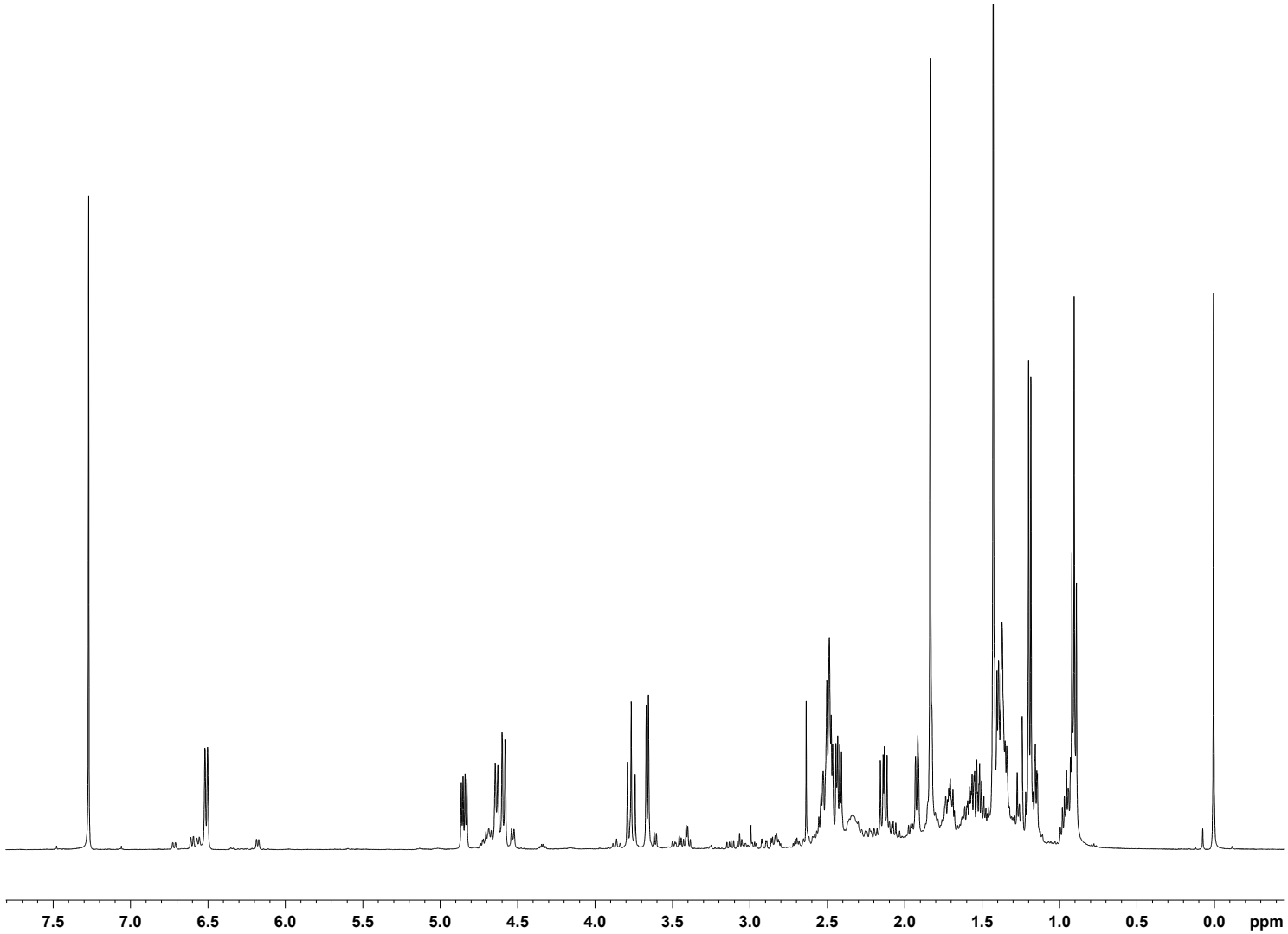


Figure S2. ^{13}C NMR Spectrum of Akaeolide (**1**) at 125 MHz in $CDCl_3$

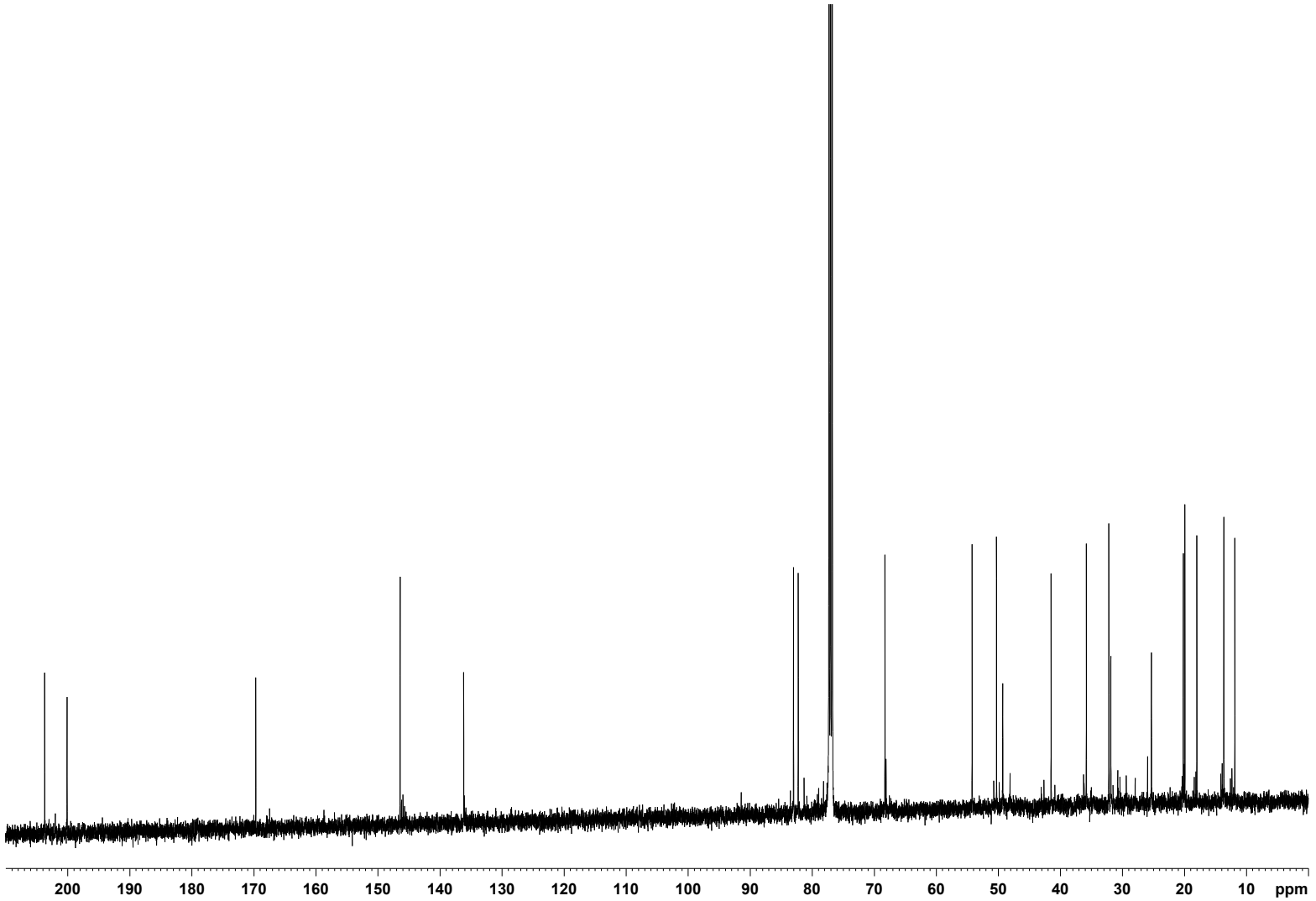


Figure S3. ^{1}H-^{1}H COSY Spectrum of Akaeolide (**1**) at 500 MHz in $CDCl_3$

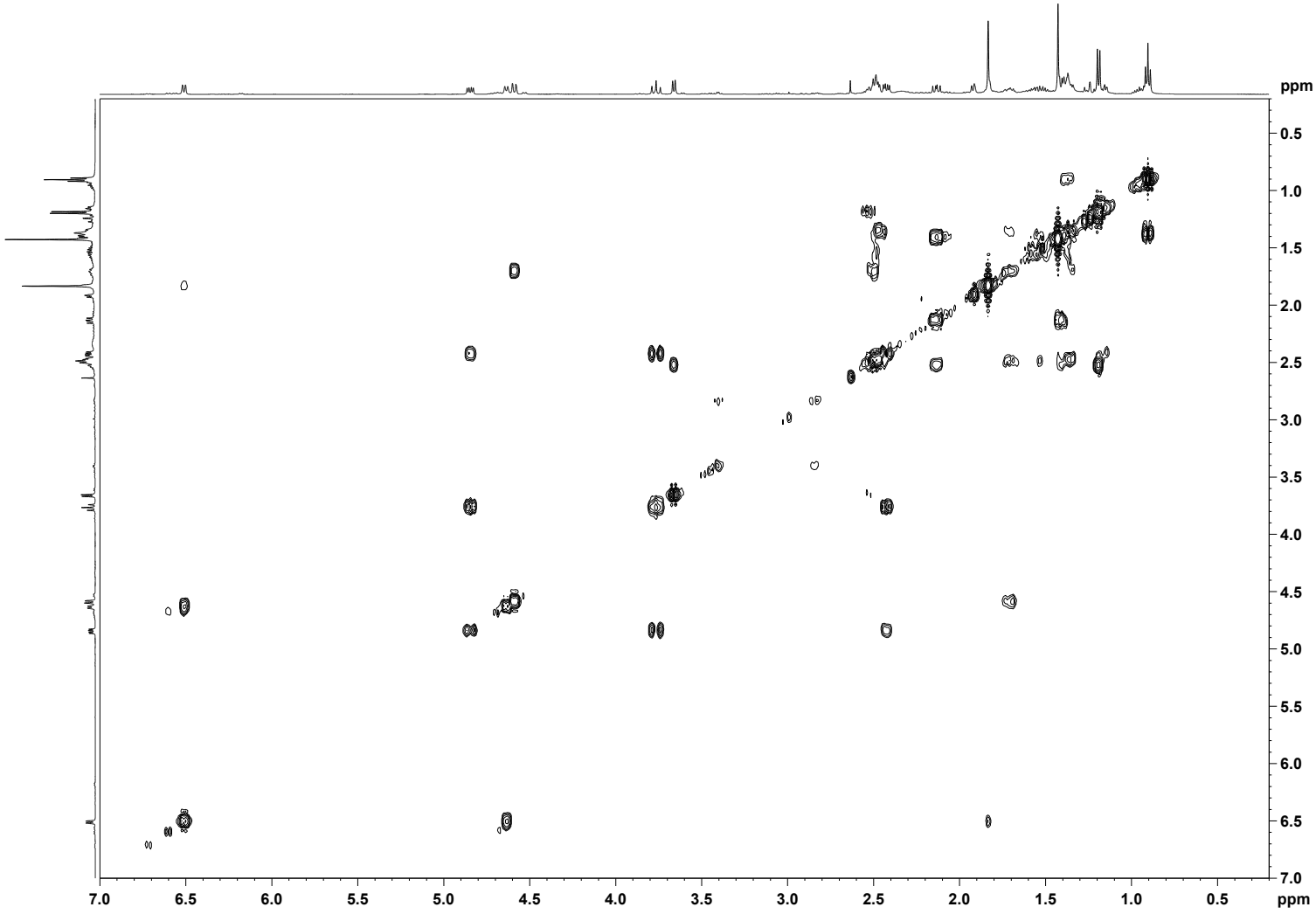


Figure S4. HSQC Spectrum of Akaeolide (**1**) at 500 MHz in $CDCl_3$

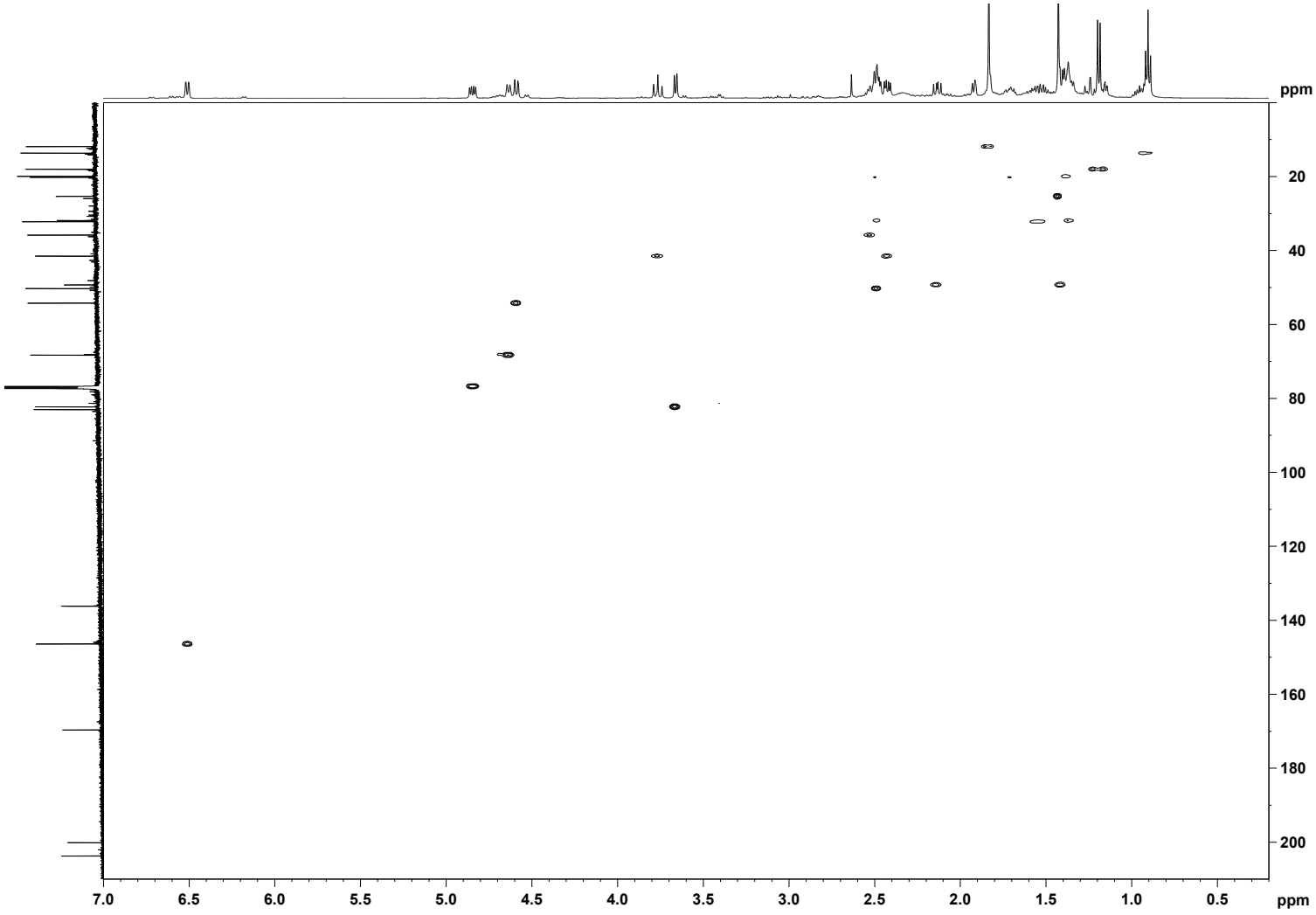


Figure S5. HMBC Spectrum of Akaeolide (**1**) at 500 MHz in $CDCl_3$

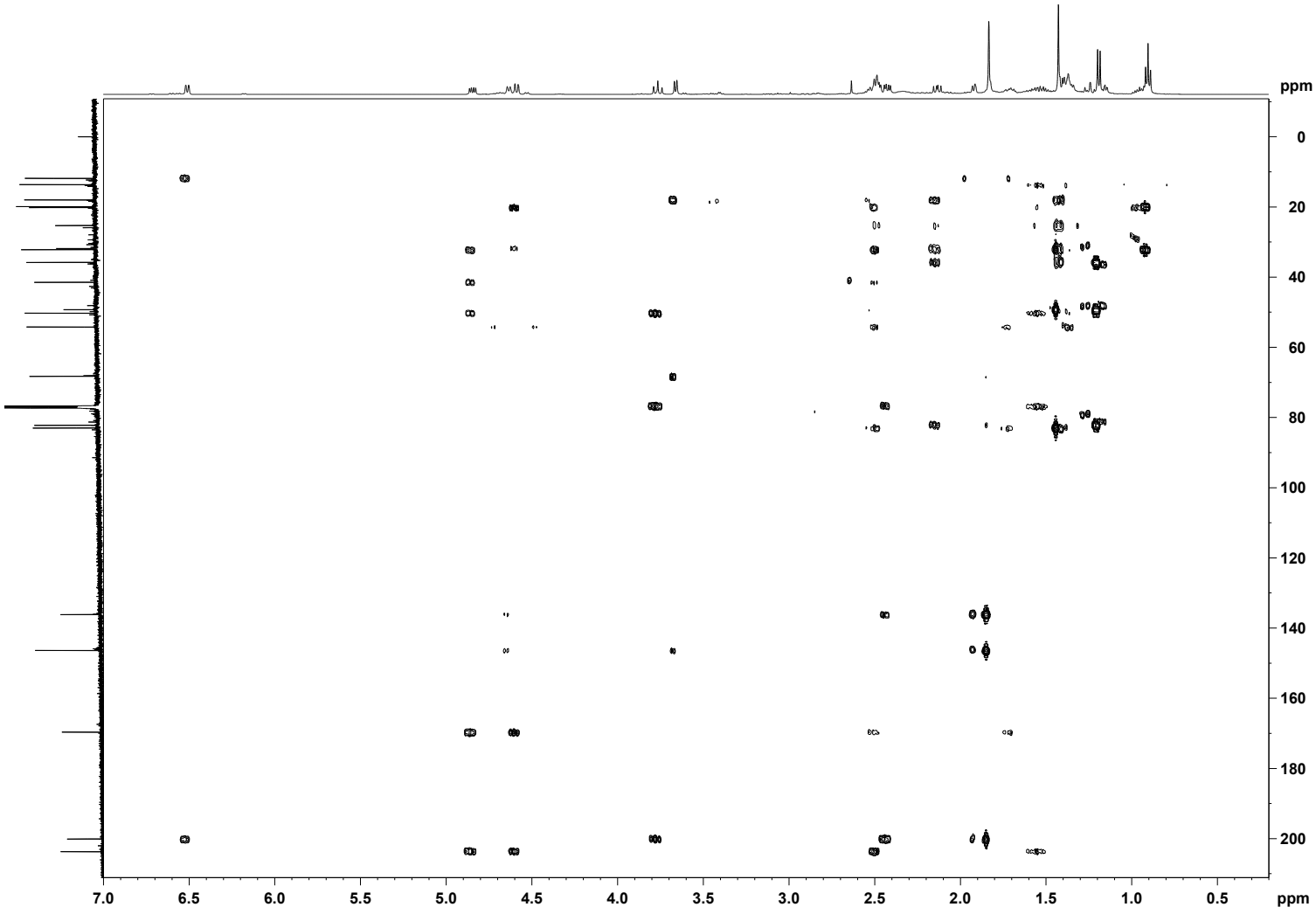


Figure S6. NOESY Spectrum of Akaeolide (**1**) at 500 MHz in $CDCl_3$

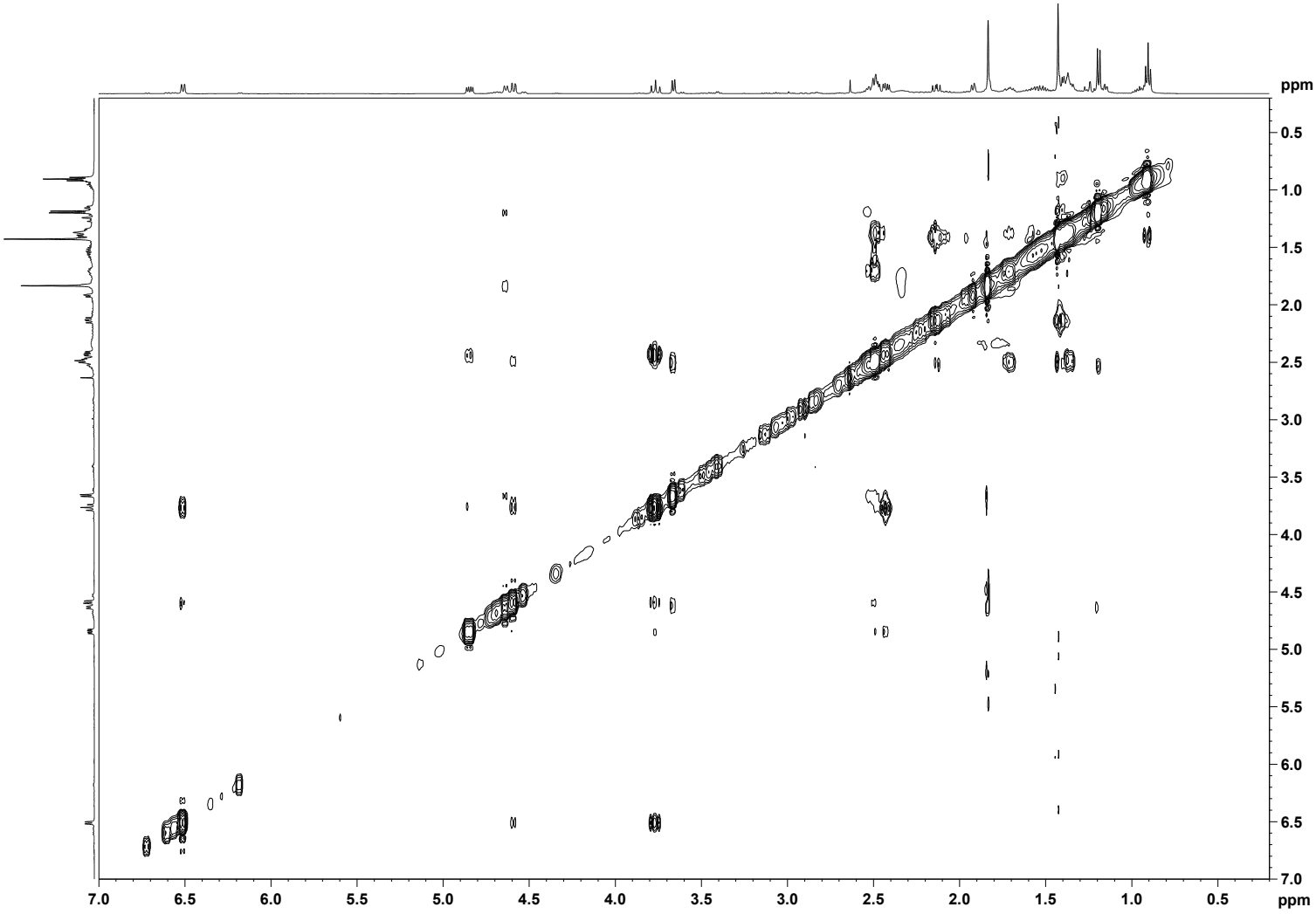


Figure S7. ^{1}H NMR Spectrum of Akaeolide (**1**) at 500 MHz in pyridine-d_5

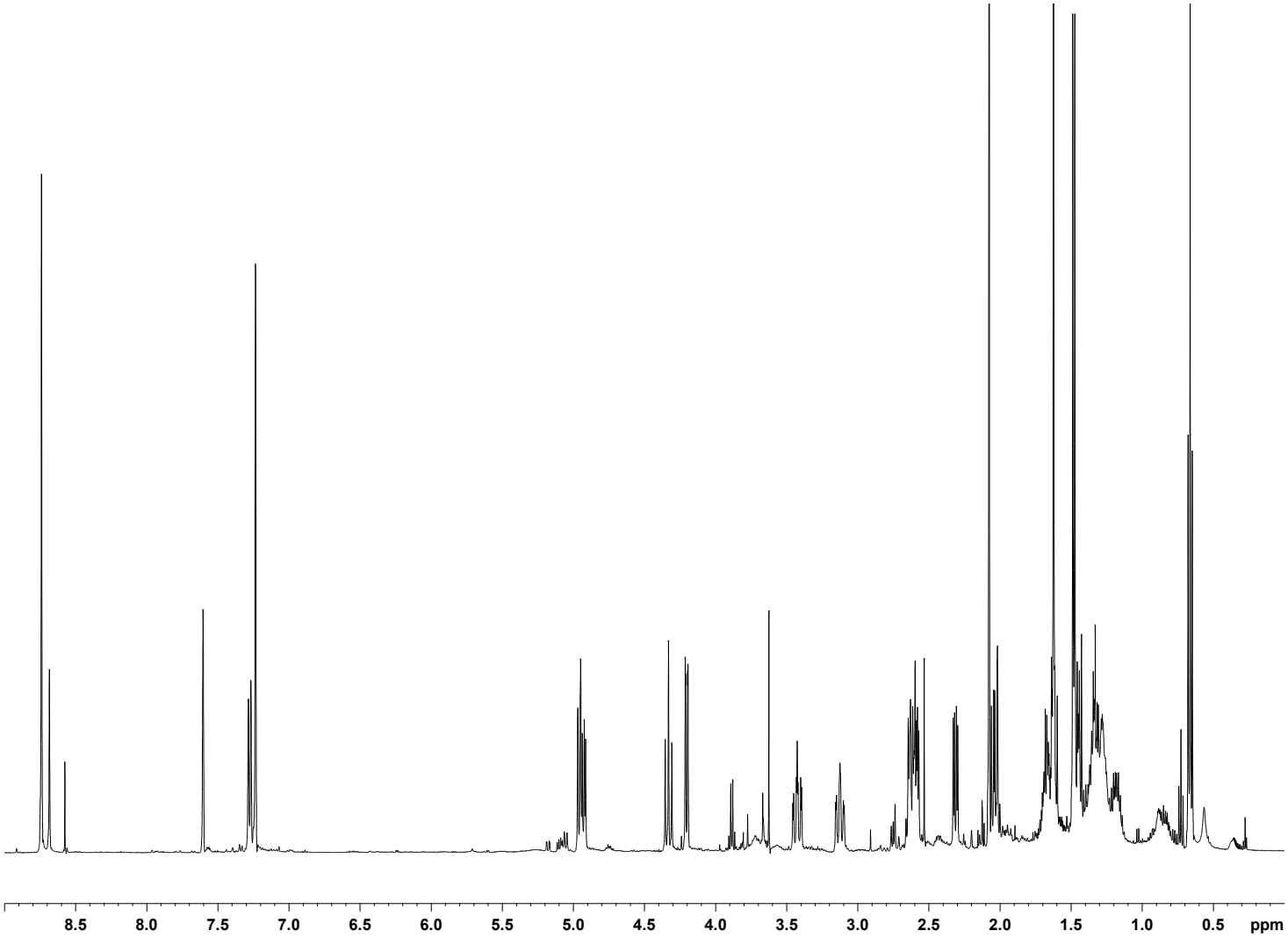


Figure S8. ^{13}C NMR Spectrum of Akaeolide (**1**) at 125 MHz in pyridine-d_5

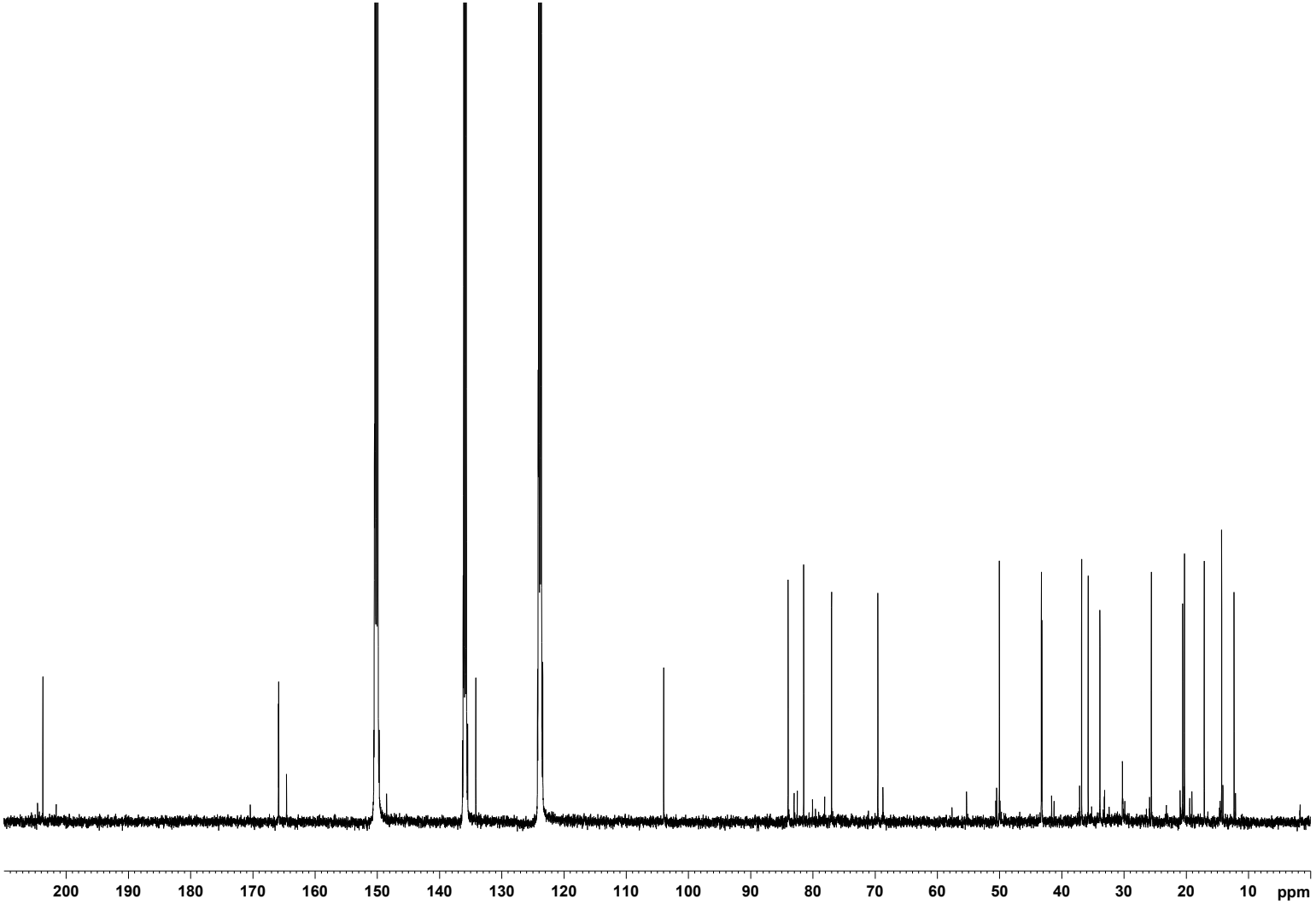


Figure S9. ^{1}H-^{1}H COSY Spectrum of Akaeolide (**1**) at 500 MHz in pyridine-d_5

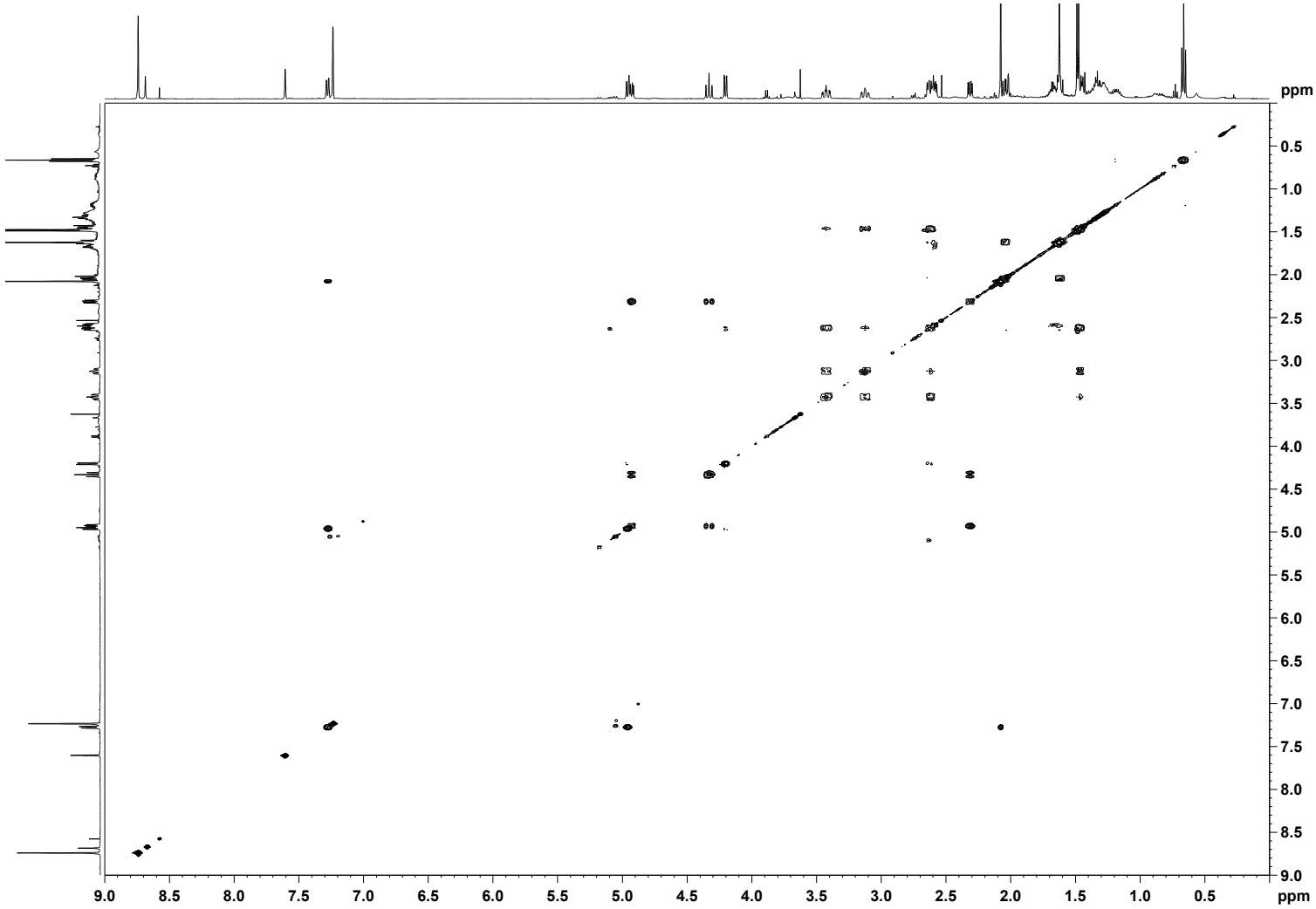


Figure S10. HSQC Spectrum of Akaeolide (**1**) at 500 MHz in pyridine-d_5

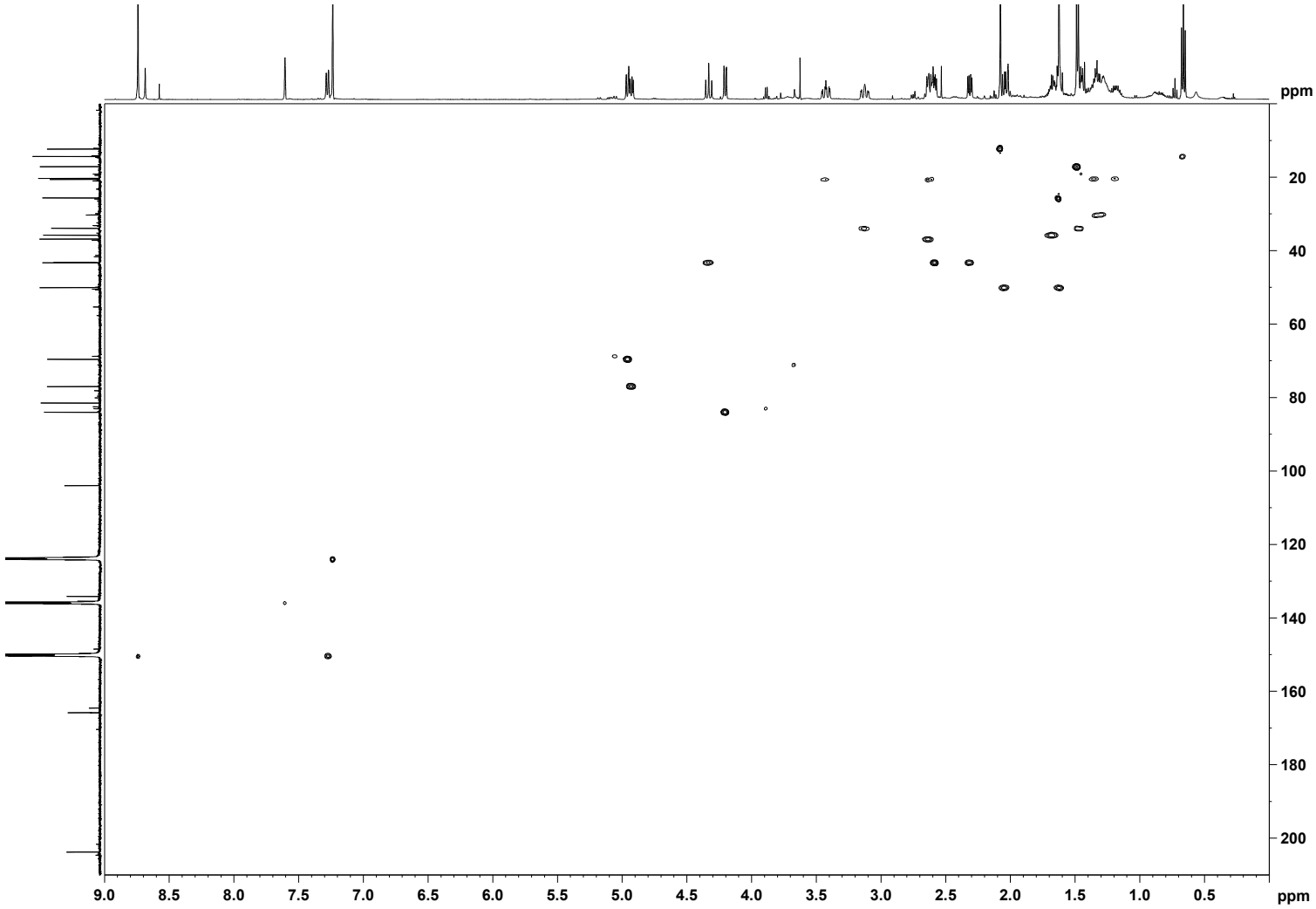


Figure S11. HMBC Spectrum of Akaeolide (**1**) at 500 MHz in pyridine-d_5

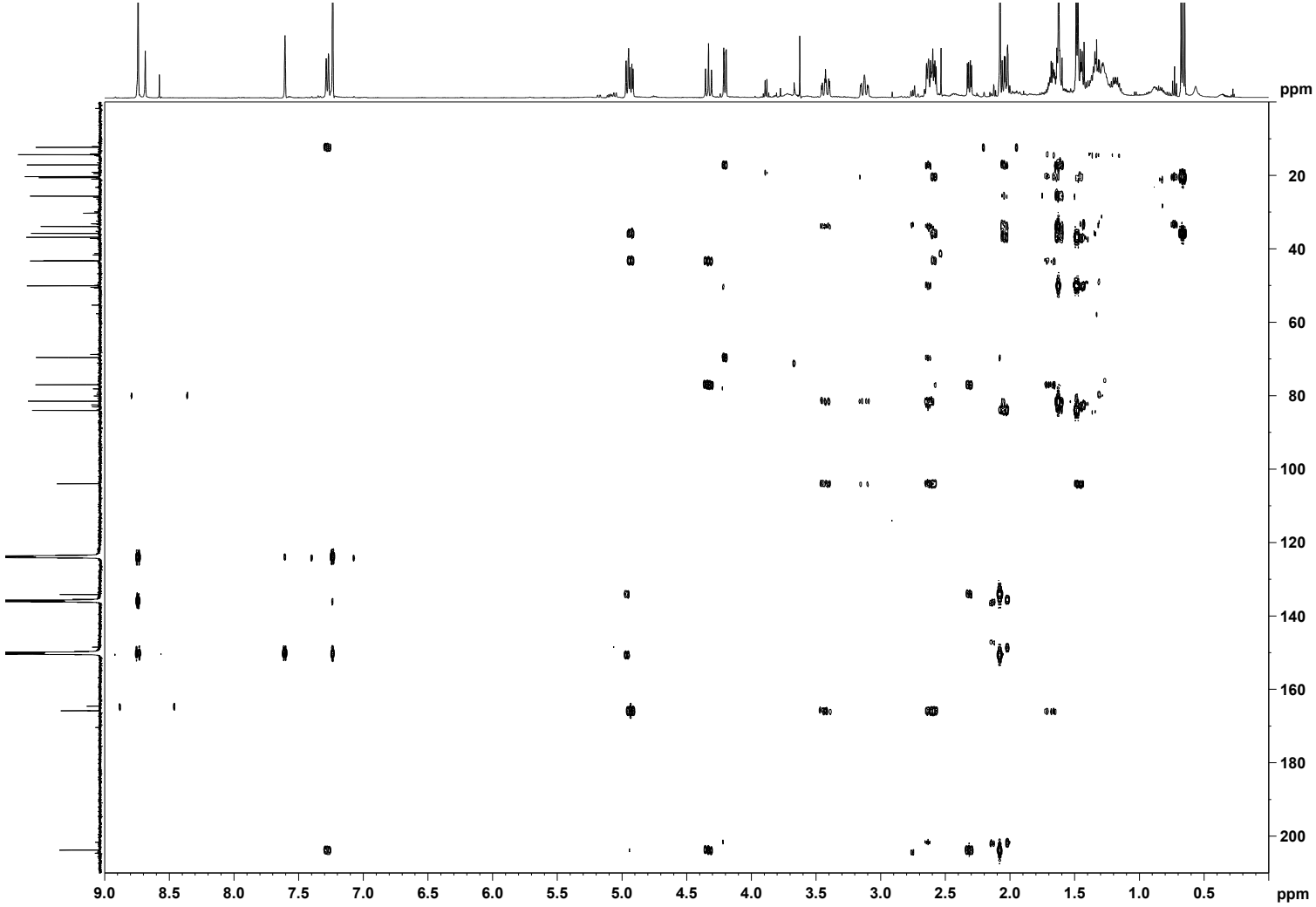


Figure S12. NOESY Spectrum of Akaeolide (**1**) at 500 MHz in pyridine-d_5

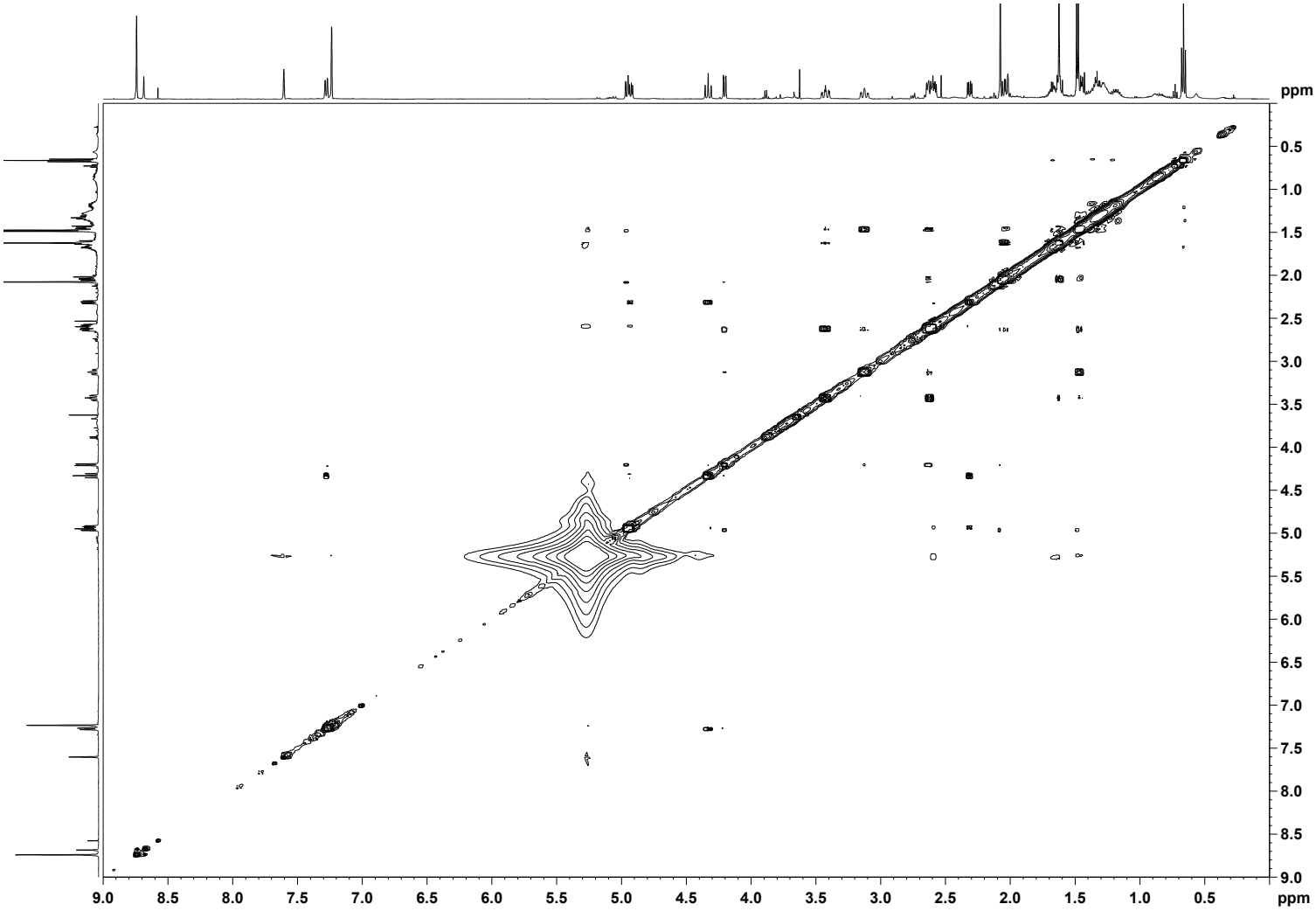
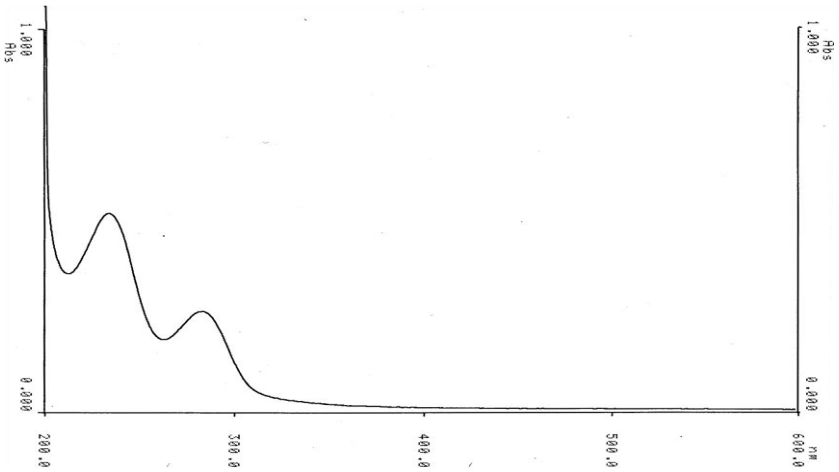
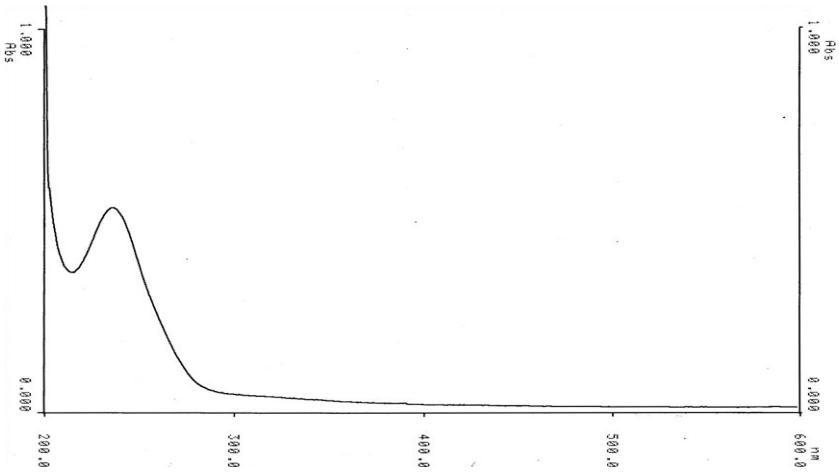


Figure S13. UV Spectra of Akaeolide (**1**)

(A) MeOH



(B) 0.01N HCl-MeOH



(C) 0.01N NaOH-MeOH

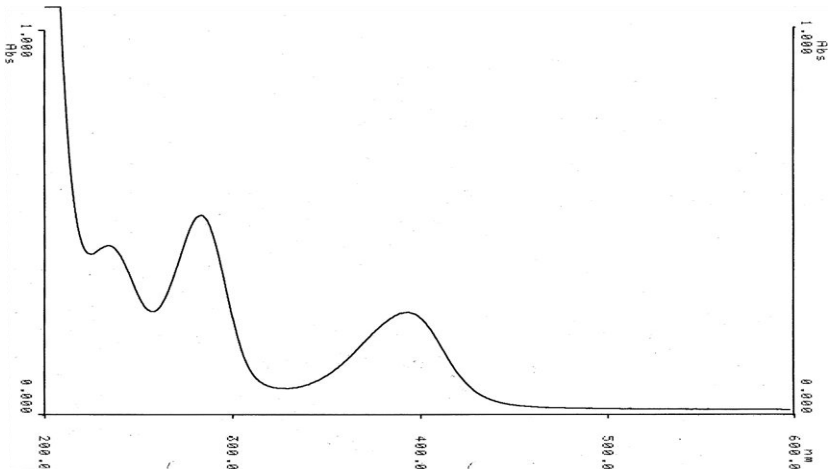


Figure S14. IR Spectrum of Akaeolide (**1**)

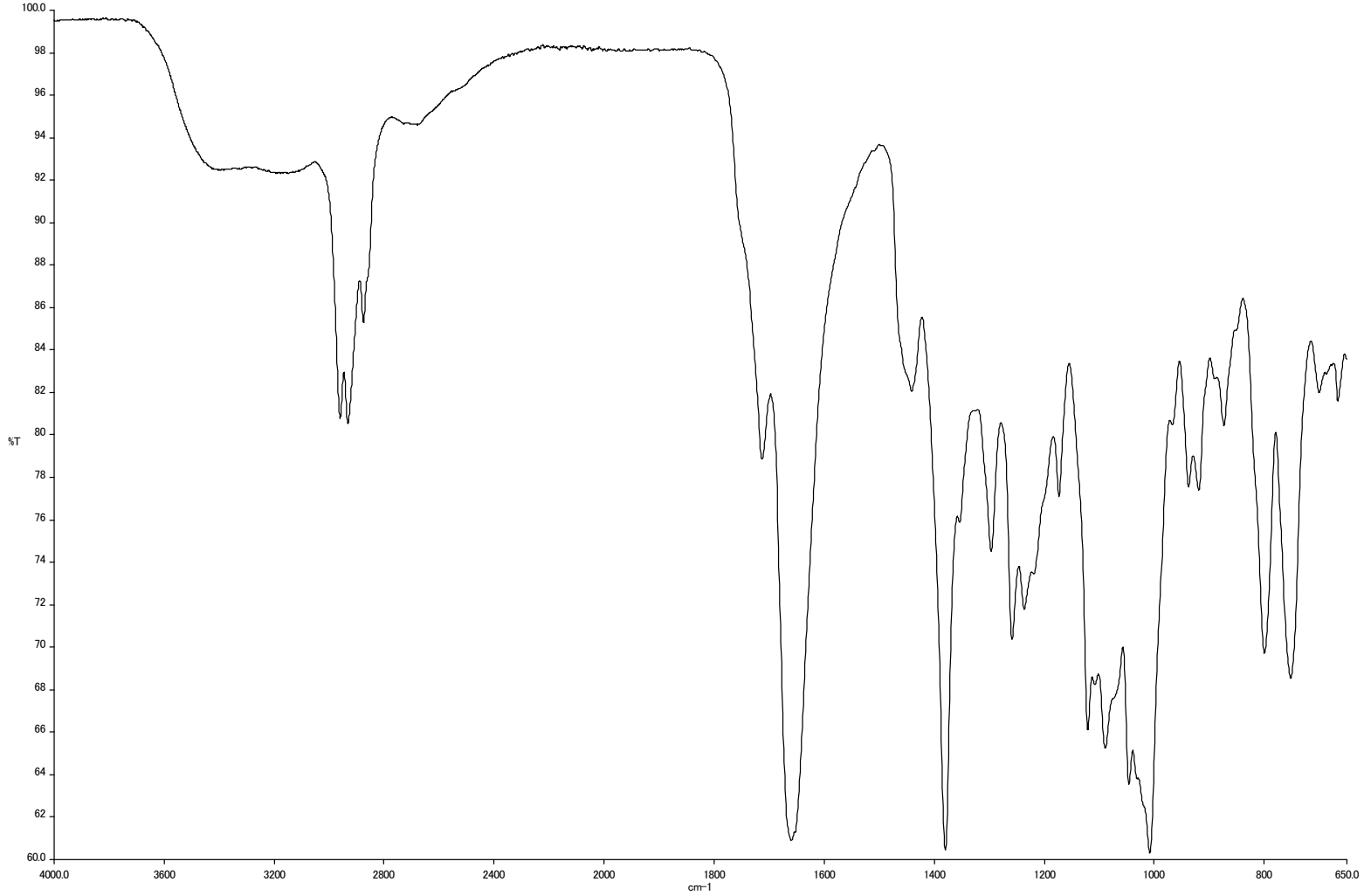


Figure S15. ^{1}H NMR Spectrum of Chloroakaeolide (**2**) at 500 MHz in $CDCl_3$

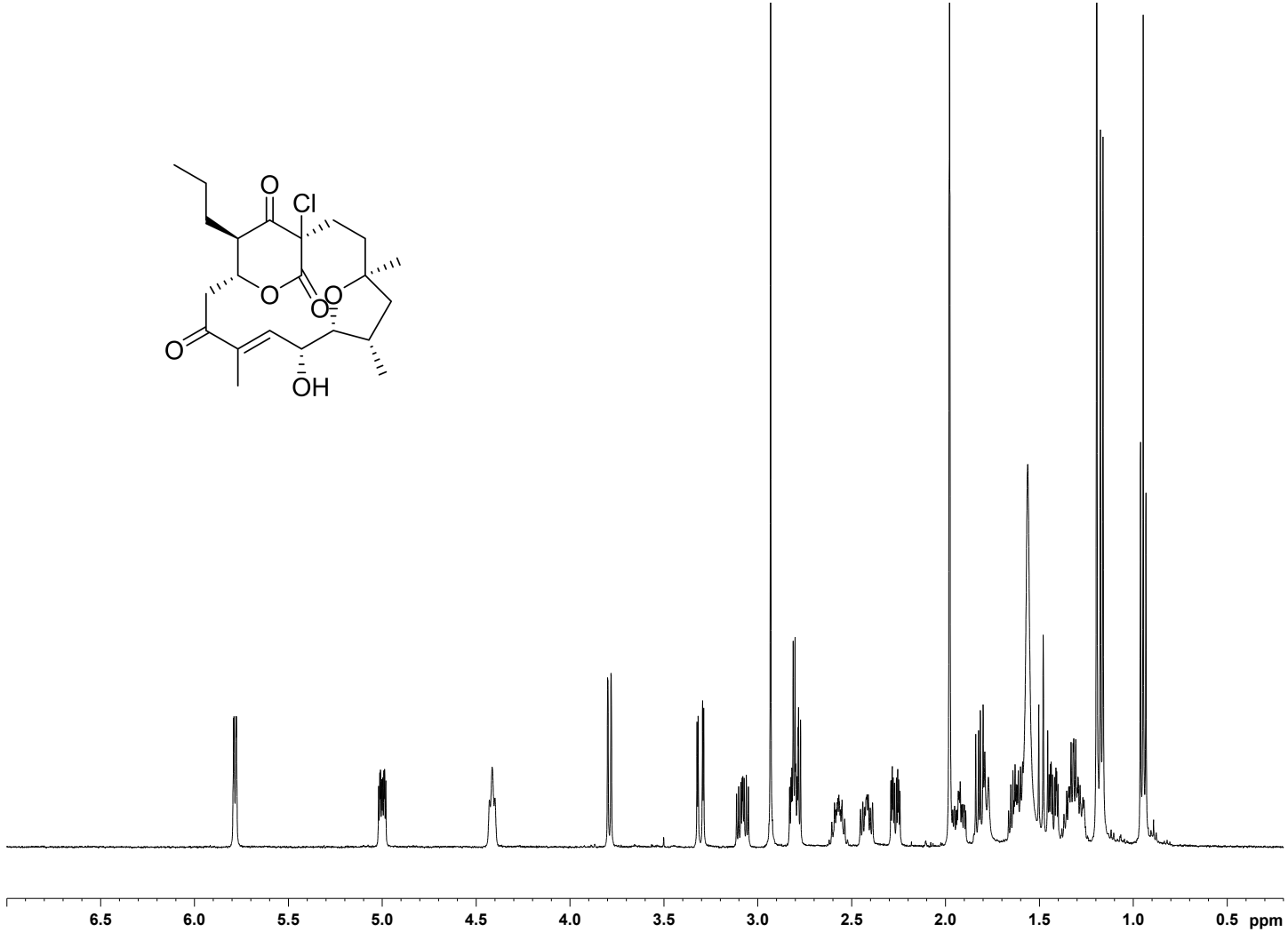


Figure S16. ^{13}C NMR Spectrum of Chloroakaeolide (**2**) at 125 MHz in $CDCl_3$

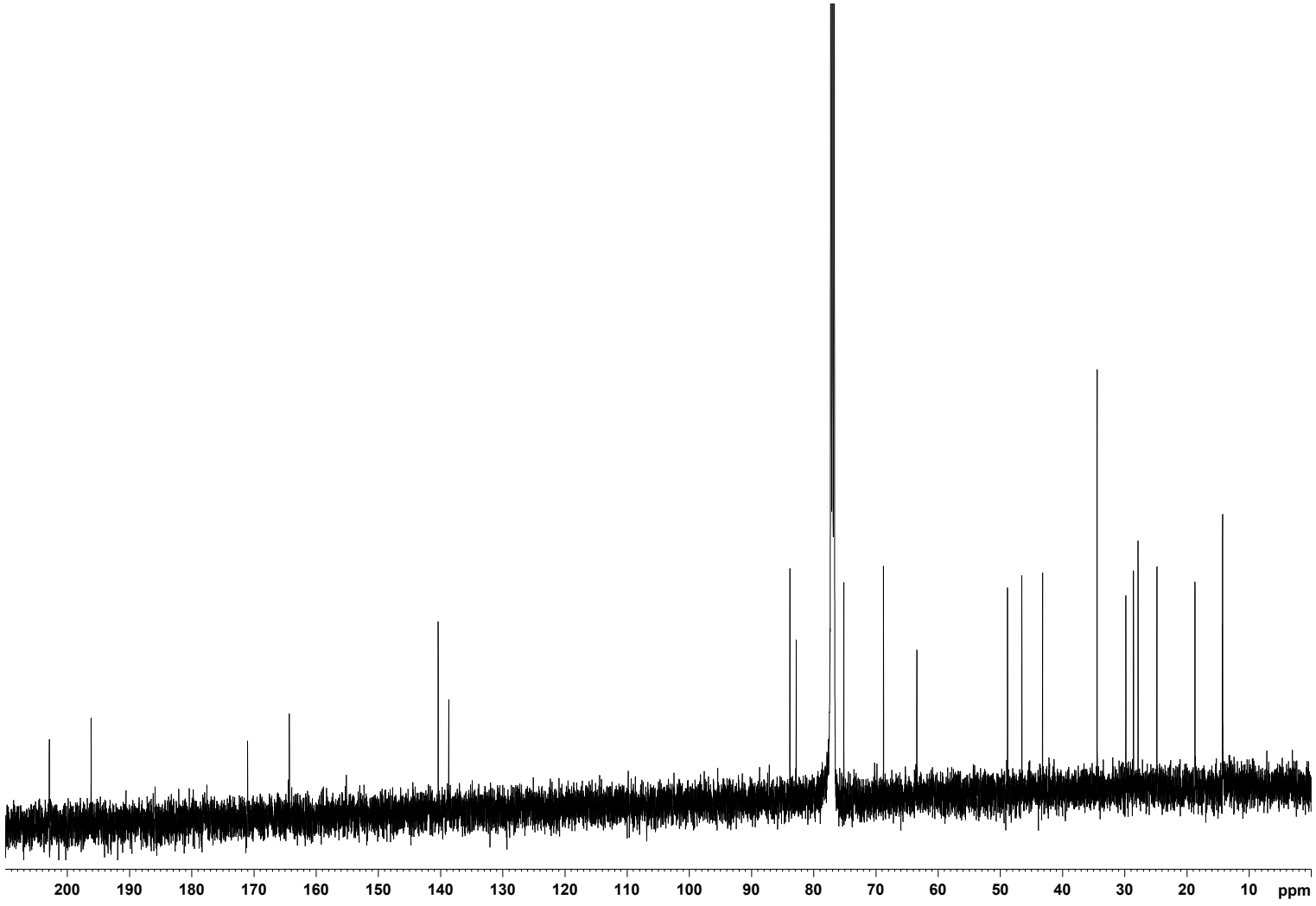


Figure S17. ^{1}H-^{1}H COSY Spectrum of Chloroakaeolide (**2**) at 500 MHz in $CDCl_3$

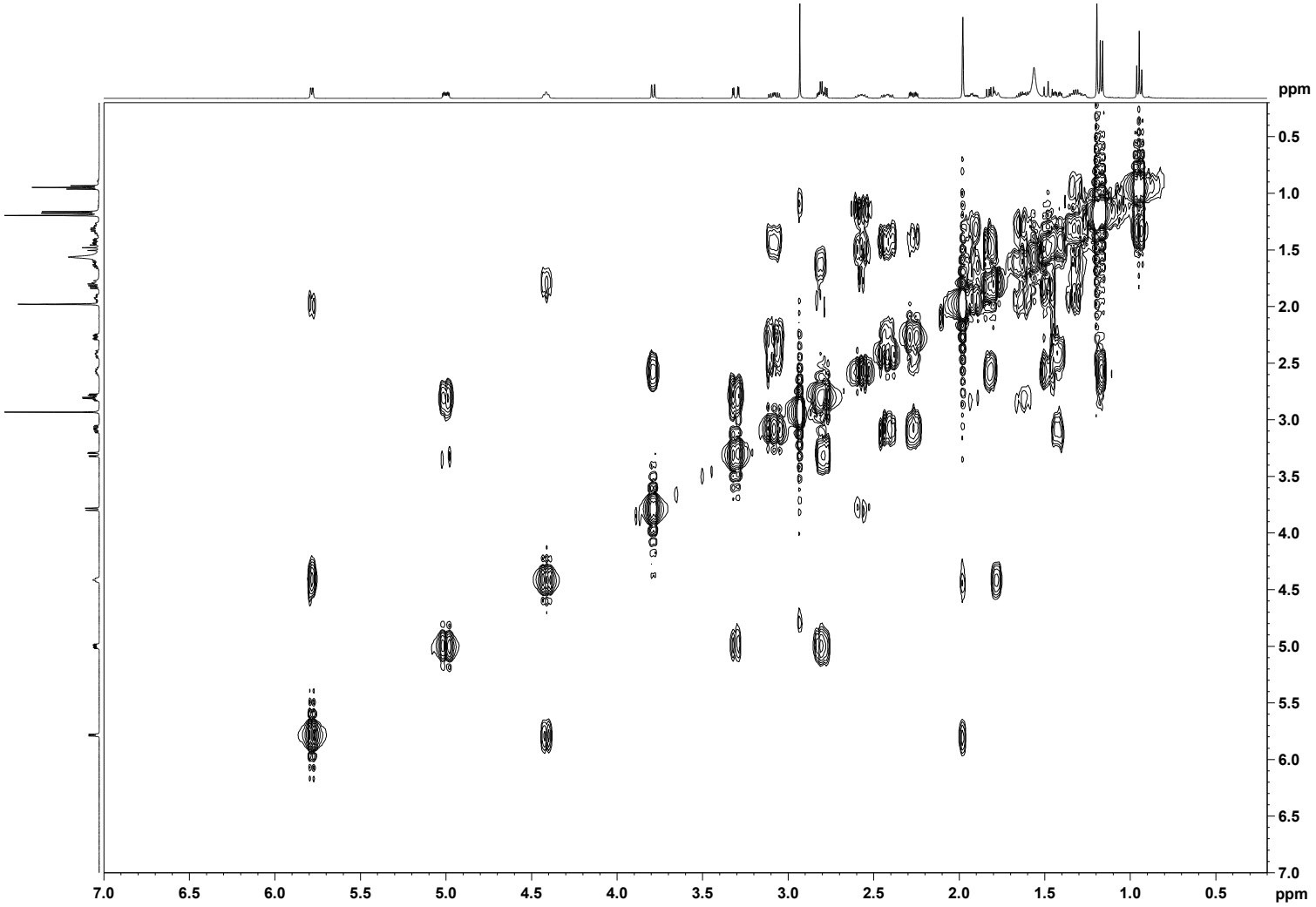


Figure S18. HSQC Spectrum of Chloroakaeolide (**2**) at 500 MHz in $CDCl_3$

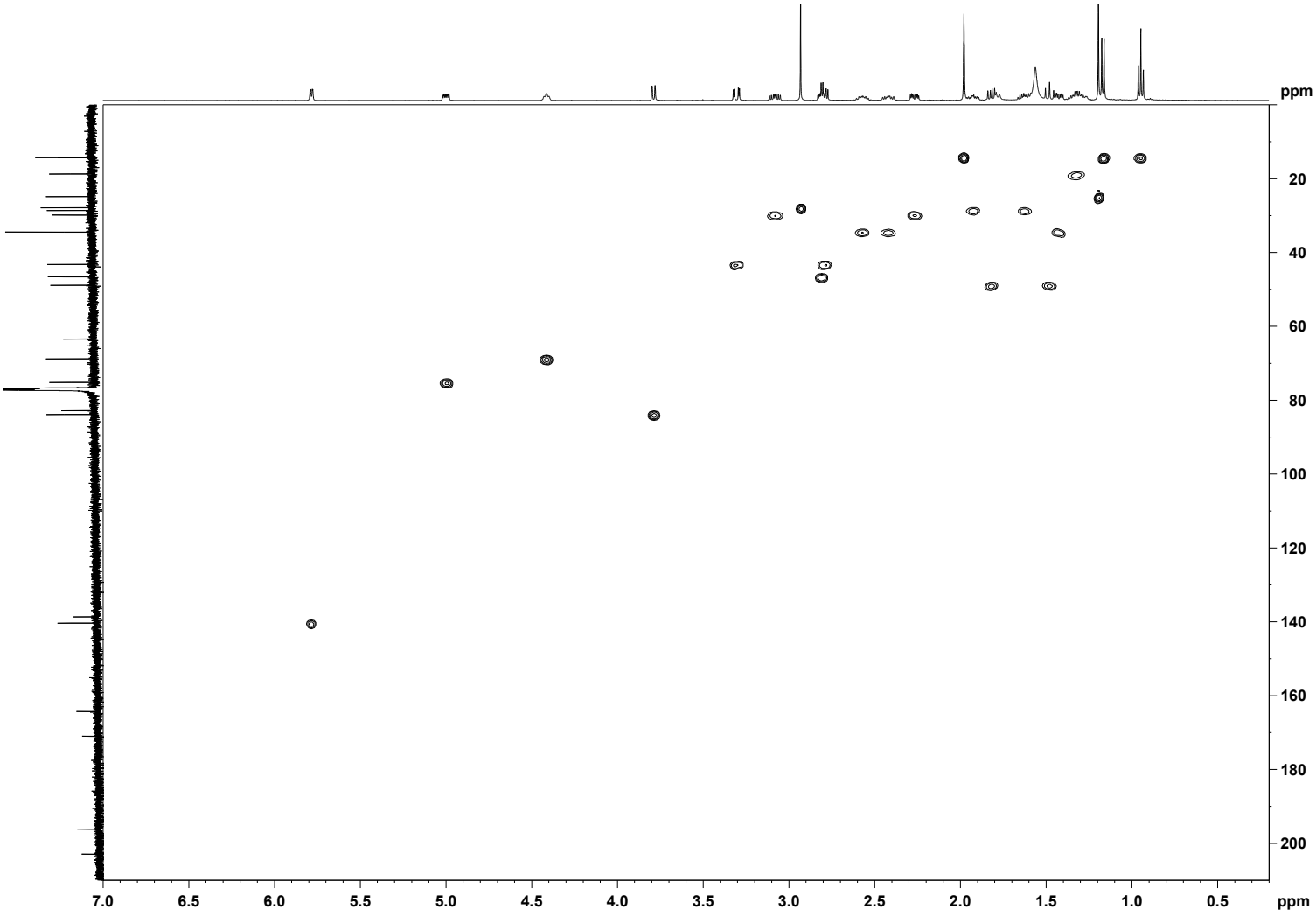


Figure S19. HMBC Spectrum of Chloroakaeolide (**2**) at 500 MHz in $CDCl_3$

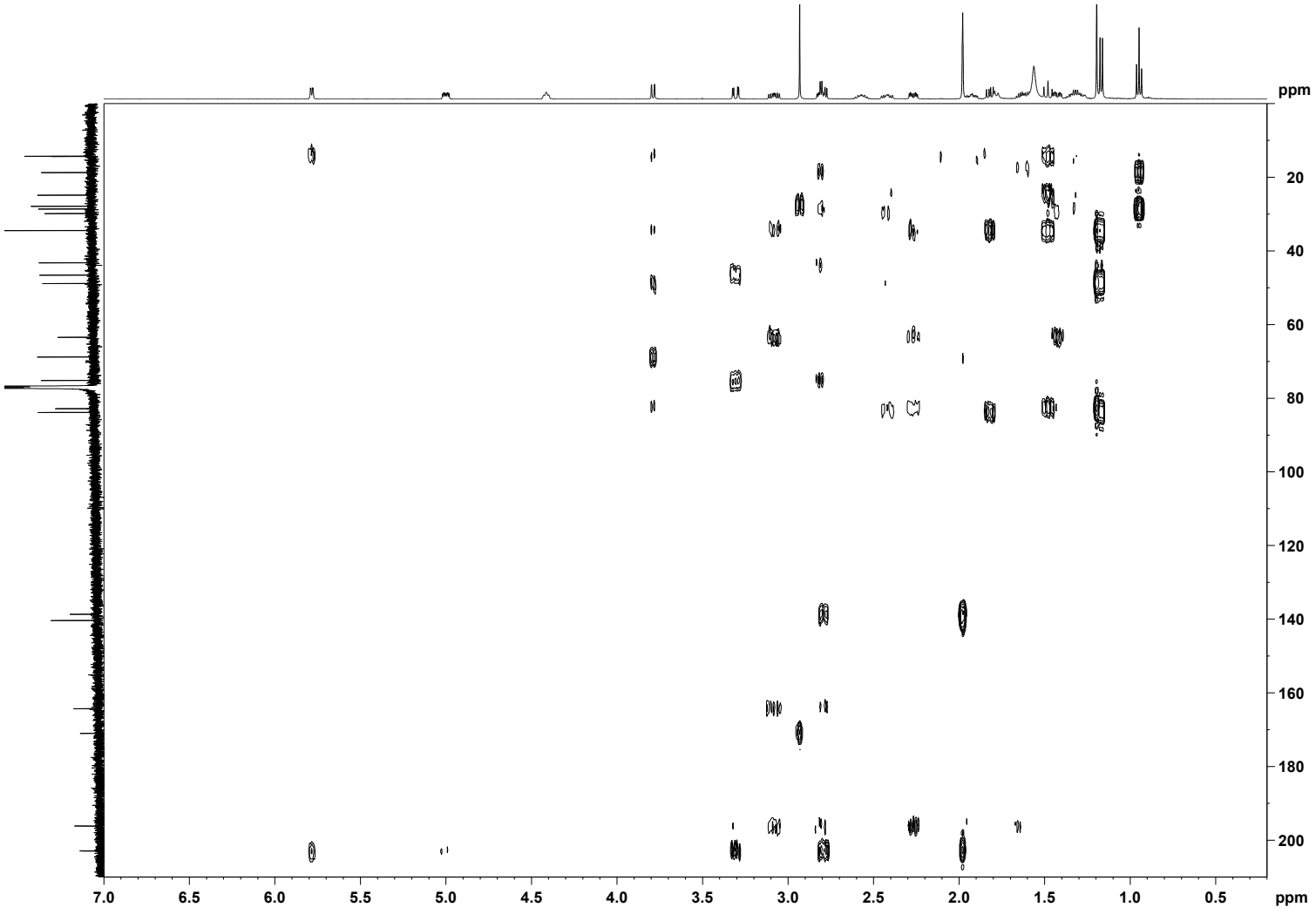
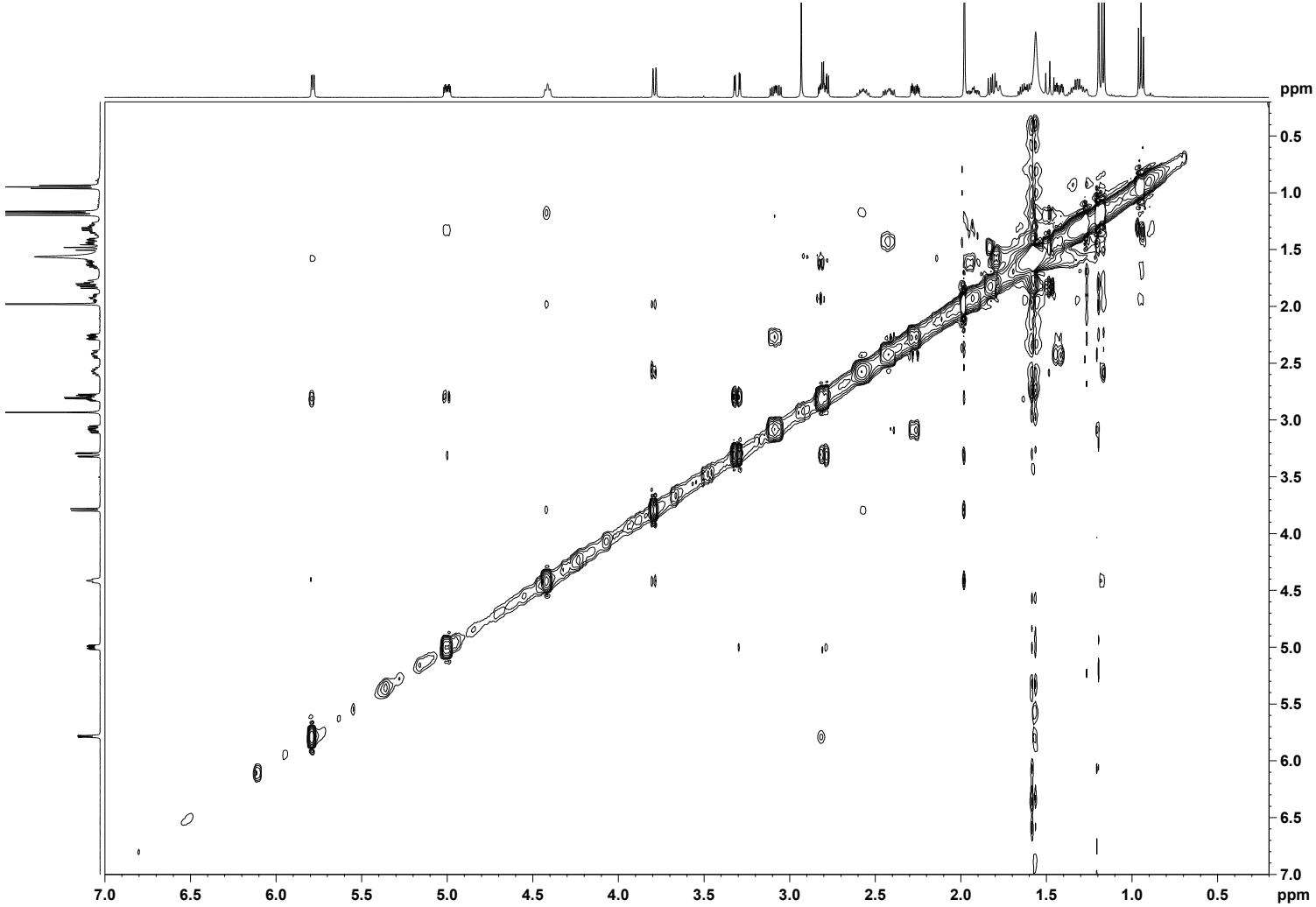


Figure S20. NOESY Spectrum of Chloroakaeolide (**2**) at 500 MHz in $CDCl_3$



CHAPTER 3

Biosynthesis of Akaeolide and Lorneic Acids and Annotation of Type I Polyketide Synthase Gene Clusters in the Genome of *Streptomyces* sp. NPS554

3-1 Background

As described in Chapter 2, *Streptomyces* sp. NPS554 is capable to produce two types of polyketides: akaeolide (**1**) and lorneic acids. Both of these compounds are structurally specific.

Akaeolide (**1**) is featured by its 15-membered carbocyclic structure functionalized with a five-membered cyclic ether and a *β*-keto-*δ*-lactone unit. Usually, in macrolide-type polyketide compounds, the carbon chain is cyclized through an ester bond to form a ring structure (Figure 3-1). However, in akaeolide (**1**), the ring system is composed only by carbon atoms, which is quite uncommon in polyketides. Only several compounds might have close biosynthetic relationships to **1**, including lankacidin,[1] cochleamycin,[2,3] and macquarimicin.[4] In lankacidin biosynthesis, the polyketide chain extension begins from a glycine-derived C-18/C-17 unit and ends with the terminating propionate unit C-19/C-2/C-1 (Figure 3-2). The C-C bond between C-2 and C-18 is thought to be formed by a nucleophilic addition of the activated methine carbon C-2 to the *N*-acyl imine carbon C-18 with the catalysis of an amine oxidase.[5] Biosynthetic origin of cochleamycin was elucidated by incorporation of ^{13}C-labeled precursors.[6] Starting from the C-5/C-6 acetate unit, the chain extension terminates with a C-4/C-3 acetate unit accompanied by the *δ*-lactone formation. The additional C-C bond formation is proposed to arise from the oxidation of C-5 methyl carbon to the aldehyde and the following intramolecular aldol reaction (Figure 3-3). However, unlike lankacidin, this biosynthetic pathway is not supported by any enzymatic or genomic evidence. Similar to these biosynthetic pathways, an extra C-C bond formation must take place for the construction of the macrocyclic structure of **1** in addition to the regular C-C bond formation by PKS.

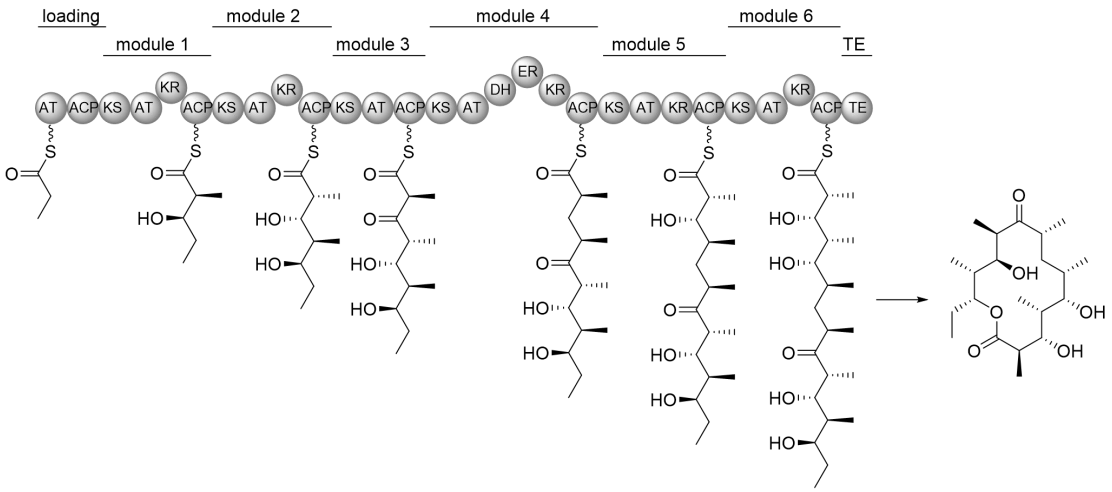


Figure 3-1. Biosynthesis of deoxyerythronolide-B-synthase (DEBS) required for erythromycin as an example of a macrolide type I PKS.

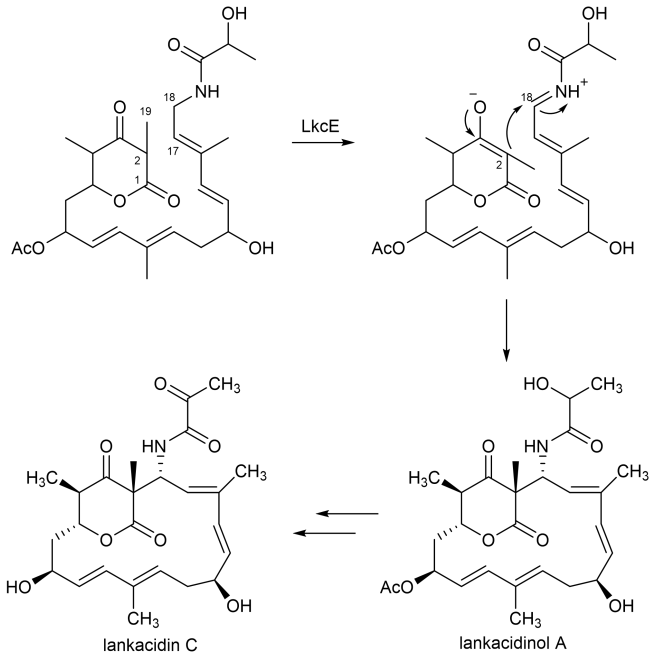


Figure 3-2. Proposed biosynthetic pathway of lankacidin.

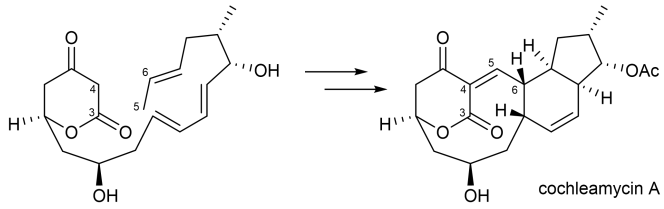


Figure 3-3. Proposed biosynthetic pathway of cochleamycin.

Lorneic acid A (**3**) is structurally characterized by a linear fatty acid chain bearing a benzene ring in the middle of the structure. Only limited natural products are known to have such molecular architecture, including lorneic acids B, C and D, lorneamides A and B, and BE52211 analogs (Figure 3-4).[21,22] Although there are many polyketides bearing aromatic rings, most of them are found in type II PKSs (Figure 3-5). Even for the type I PKS compounds bearing aromatic ring, their resources are only known from starter units with aromatic ring, but not directly synthesized by PKS. Therefore, the compounds like lorneic acids (Figure 3-4) are biosynthetically special type I PKS.

Since both akaeolide (**1**) and lorneic acid A (**3**) have biosynthetically unique features and their biosynthesis remains unknown, I next attempted to elucidate the biosynthetic pathway of these two types of polyketides.

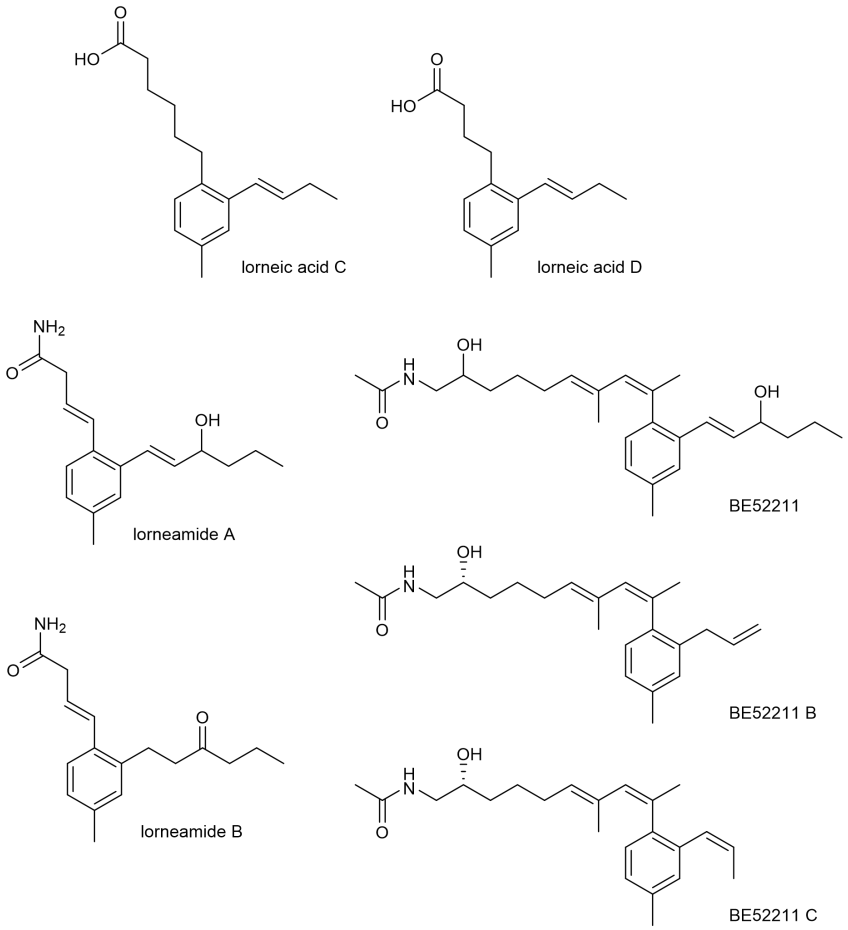


Figure 3-4. Structures of lorneic acids C and D, lorneamide A and B, and BE52211 analogs.

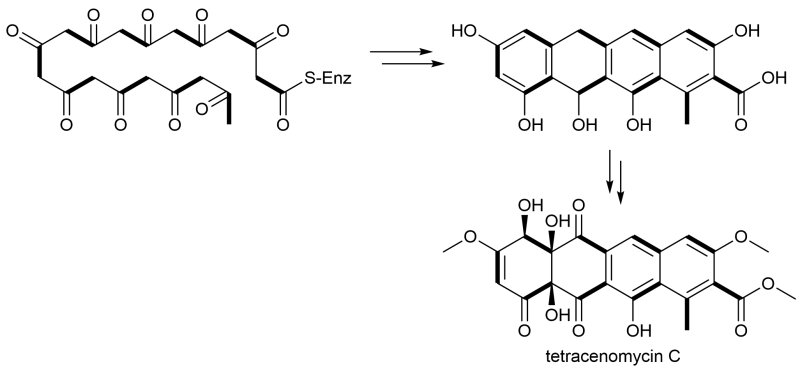


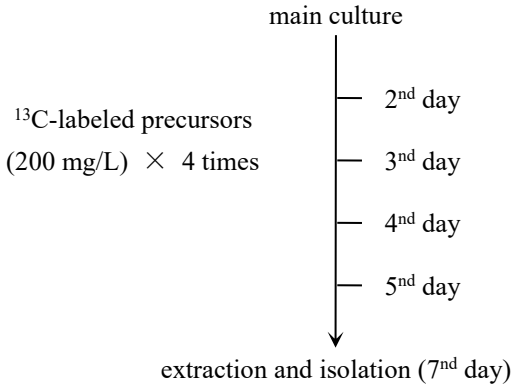
Figure 3-5. Aromatic rings synthesized by type II PKS.

3-2 Results and Discussion

3-2-1 Incorporation of ^{13}C-Labeled Precursors

Inspection of the carbon connectivity and the position of carbon branches suggested that **1** was synthesized through the polyketide pathway. In order to elucidate the biosynthetic origin and incorporation pattern, strain NPS554 was cultured in the presence of plausible biosynthetic precursors labeled with carbon-13, namely [1-^{13}C]acetate, [2-^{13}C]acetate, and [1-^{13}C]propionate,

which could be incorporated into the polyketide backbone via acyl CoA carboxylation (Scheme 3-1). According to the previous study described in Chapter 2, **1** exists as a mixture of several tautomeric isomers in NMR solvents caused by the enolization at C-17, consequently giving multiple ^{13}C signals for each carbon. This was undesirable to the quantification of carbon intensity; therefore, the purified ^{13}C-labeled **1** was converted to the chlorinated derivative **2** which could not undergo isomerization (Figure 2-9).



Scheme 3-1. Feeding culture with ^{13}C-labeled precursors.

The relative enrichment of each carbon by the incorporation of ^{13}C-labeled precursors was determined by ^{13}C NMR measurement (Table 3-1). Enrichments at C-5, C-9, C-15, C-18, and C-19 were observed by the feeding of [1-^{13}C]acetate while C-3, C-7, C-11, and C-13 were highly enriched by [1-^{13}C]propionate feeding (Table 6, Figures S1 and S2). C-12 and C-15 were overlapped at 34.7 ppm in the ^{13}C NMR spectrum, but the signal enhancement was ascribed to the incorporation of [1-^{13}C]acetate into C-15 because the three-carbon fragment C-21/C-12/C-11 was derived from a propionate as proven by [1-^{13}C]propionate incorporation into C-11 (Figure 3-6).

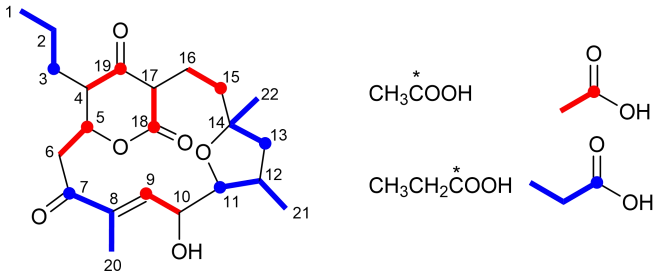


Figure 3-6. Incorporation of ^{13}C-labeled precursors into akaeolide (**1**).

C-16 and C-17 were labeled by [2-^{13}C]acetate and C-15 and C-18 by [1-^{13}C]acetate (Figure 3-6). This labeling pattern is inconsistent with normal polyketide chain elongation but could be explained by incorporation of succinate, which can be labeled by ^{13}C-labeled acetates in TCA cycle (Figure 3-7).[7,8]

Table 3-1. Incorporation of ^{13}C-labeled precursors into 17-chloroakaeolide (**2**).

Position	δ_C	Relative Enrichments$^{\alpha}$		
		[1-^{13}C]-acetate	[2-^{13}C]-acetate	[1-^{13}C]-propionate
1	14.5	1.0	0.9	0.8
2	19.0	1.0	1.0	1.0
3	28.8	1.0	1.7	**32.9**
4	46.8	1.0	**4.4**	1.0
5	75.4	**5.0**	1.1	0.8
6	43.5	1.1	**4.8**	0.9
7	203.1	1.2	2.1	**40.3**
8	138.9	0.9	0.8	0.6
9	140.6	**4.6**	1.0	0.8
10	69.0	1.0	**4.5**	0.6
11	84.1	1.0	1.6	**30.2**
12	34.7	**2.9**	1.0	0.9
13	49.1	1.0	1.6	**31.6**
14	83.1	1.0	0.9	0.9
15	34.7	**2.9**	1.0	0.9
16	30.1	1.0	**4.3**	1.0
17	63.7	1.0	**3.8**	1.1
18	164.5	**4.8**	0.9	0.8
19	196.4	**4.4**	1.0	1.1
20	14.5	1.0	0.9	0.8
21	25.1	1.0	0.8	1.0
22	14.6	1.0	0.9	0.8

$^{\alpha}$ ^{13}C signal intensity of each peak in the labeled 17-chloroakaeolide divided by that of the corresponding signal in the unlabeled, normalized to give an enrichment ratio of 17-chloroakaeolide for enriched peak. The numbers in bold type indicate ^{13}C-enriched atoms from ^{13}C-labeled precursors.

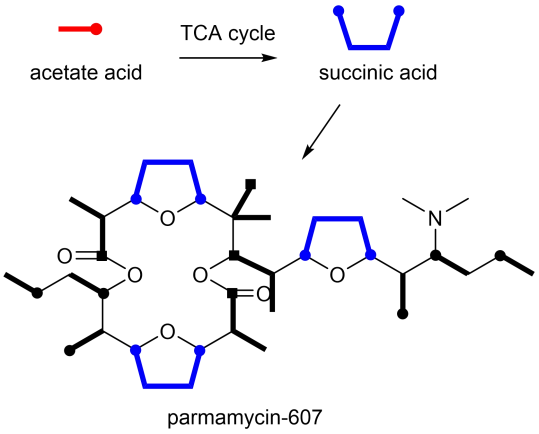


Figure 3-7. Acetate-labeled succinic acid incorporation into polyketide.

Table 3-2. Incorporation of ^{13}C-labeled precursors into lorneic acid A (**3**).

Position	δ_C	relative enrichments[α]	
		[1-^{13}C]-acetate	[1-^{13}C]-propionate
1	177.7	**7.3**	0.9
2	38.3	1.0	1.0
3	121.8	**6.7**	1.0
4	131.9	1.0	1.0
5	131.7	**6.5**	0.8
6	126.4	1.0	1.0
7	127.8	1.0	**42.6**
8	137.3	0.9	0.7
9	127.0	**6.5**	1.1
10	136.1	0.8	0.8
11	127.4	**6.1**	1.1
12	133.7	0.9	1.0
13	33.0	**6.4**	1.1
14	31.6	1.0	1.0
15	22.3	**6.3**	1.0
16	13.9	1.0	1.0
17	21.2	1.0	1.0

[α] ^{13}C signal intensity of each peak in the labeled lorneic acid A divided by that of the corresponding signal in the unlabeled, normalized to give an enrichment ratio of lorneic acid A for enriched peak. The numbers in bold type indicate ^{13}C-enriched atoms from ^{13}C-labeled precursors.

However, the feeding of 1,4-$^{13}C_2$-succinic acid gave no signal enhancement for C-15 and C-18 of akaeolide (Figure S2). Therefore, C-16 and C-17 should be connected after the polyketide chain assembly.

Biosynthetic origin of lorneic acid A (**3**) was similarly elucidated by feeding experiments with [1-^{13}C]acetate and [1-^{13}C]propionate. The relative enrichment of each carbon by the incorporation of ^{13}C-labeled precursors is summarized in Table 3-2. Enrichments at C-1, C-3, C-5, C-9, C-11, C-13, and C-15 were observed with the incorporation of [1-^{13}C]acetate, while only C-7 was labeled by [1-^{13}C]propionate feeding (Figure S3). This labeling pattern obviously indicates that the carbon chain extension begins with C-16/C-15 acetate, followed by the condensation of three malonates, one methylmalonate, and three malonates, and terminates at C-1 (Figure 3-8).

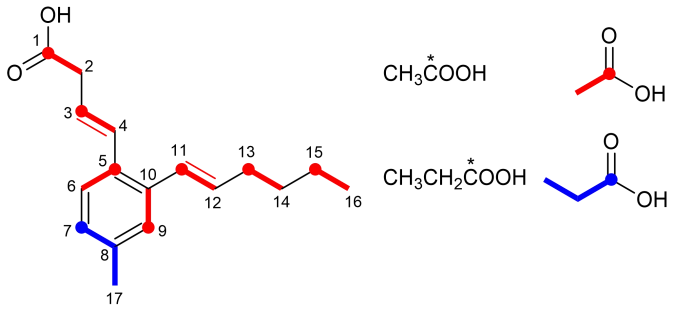


Figure 3-8. Incorporation of ^{13}C-labeled precursors into lorneic acid A (**3**).

3-2-2 Genome Analysis and Annotation of Biosynthetic Genes

Although akaeolide (**1**) and lorneic acid A (**3**) are synthesized through the polyketide pathway, the following two questions are yet not answered. How the extra C-C bond is formed in akaeolide biosynthesis? How the benzene ring of lorneic acids is constructed from a linear precursor? To get insight into more detailed biosynthetic mechanism, whole genome shotgun sequencing was carried out for *Streptomyces* sp. NPS554 to identify the biosynthetic gene clusters. At least twelve modular type I polyketide synthase (type I PKS) gene clusters are present in the genome (Table S1). Eight clusters from #1 to #8 were completely sequenced, but the remaining four clusters, from #9 to #12 were partial, likely divided into two or more scaffolds in the draft genome sequences.

Since the polyketide backbones of akaeolide and lorneic acids are constructed from one acetate starter unit and seven extender units, each cluster should have eight PKS modules. Among the PKS gene clusters present in this strain, clusters #1 and #6 are comprised of one loading module and seven extension modules, totally eight PKS modules. On the basis of predicted substrates for acyltransferase (AT) domains and the presence or absence of ketoreductase (KR), dehydratase (DH), and enoylreductase (ER) domains, clusters #1 and #6 could be assigned to lorneic acid and akaeolide biosynthetic clusters, respectively (Figure 3-9).

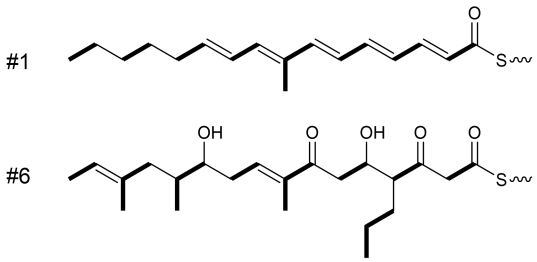


Figure 3-9. Predicted structures of linear products of gene clusters #1 and #6.

3-2-2-1 Biosynthetic Gene Cluster for Akaeolide

Gene organization of the biosynthetic gene cluster for akaeolide (**1**) and proposed functions of the PKS genes and the neighboring genes are summarized in Table 3-3. This gene cluster consists of four PKS genes with eight modules (Figure 3-10a). AT domains in loading module (LM), module 3 (m3), m5, and m7 have signature amino-acid residues specific to malonyl-CoA (C_2), while those of m1, m2, and m4 have that for methylmalonyl-CoA (C_3).[9,10] The signature amino-acid residue of m6 predicted its substrate as an alkylmalonyl-CoA, specifically propylmalonyl-CoA (C_5) in this case as suggested by the structure of the product.[11-13] These annotation results indicate that the PKSs assemble the polyketide carbon chain by sequential incorporation of C_2, C_3, C_3, C_2, C_3, C_2, C_5, and C_2 units. In combination with the results from ^{13}C-labeling experiments, it could be proposed that the chain extension starts from C-16/C-15 acetate unit, followed by a sequence of condensation with two methylmalonates, one malonate, one methylmalonate, one malonate, and one propylmalonate, and terminates at C-18 by incorporation of one malonate (Figure 3-6). In multimodular type I PKS pathways, combination of DH, ER, and KR domains in each module regulates the oxidation level of carbons in the polyketide chain.[14] This PKS cluster has four DH/KR domains and one DH/ER/KR domain, corresponding to four double bonds and one saturated methylene formation, respectively, but the actual product has only two double bonds. DH domains in m3 and m6 are likely not functional (represented by 'dh' in Figure 3-11). A quite similar gene cluster is present in the genome of *Lechevalieria aerocolonigenes* NBRC 13195T, which will be discussed later (Figure 3-10b).

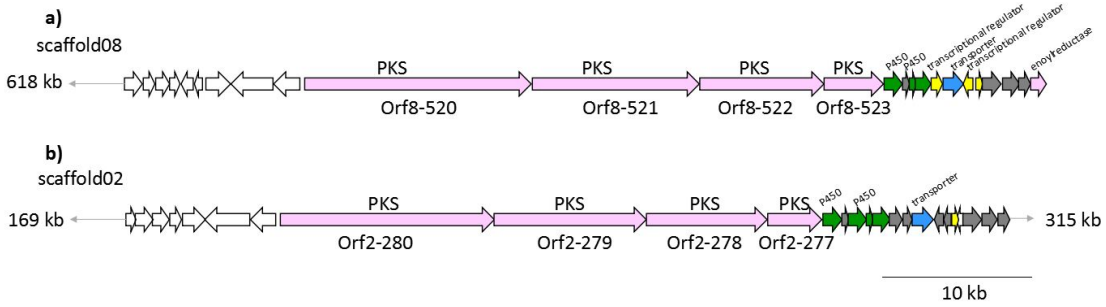


Figure 3-10. Gene organization of the akaeolide biosynthetic gene cluster of *Streptomyces* sp. NPS554 (**a**) and its orthologous gene cluster in *Lechevalieria aerocolonigenes* NBRC 13195^T (**b**).

Table 3-3. Proposed functions of type I PKS genes for akaeolide biosynthesis and neighboring genes.

Orf8-	Size [a]	Proposed Function	Protein Homolog, *Origin*, Accession Number	% [b]
511	413	ligase	ligase, *Streptomyces* sp. NRRL S-118, WP_031071281	86/92
512	242	glutamine amidotransferase	glutamin amidotransferase, *Streptomycetaceae* bacterium MP113-05, EST22401	86/91
513	315	unknown	membrane protein, *S. Hygroscopicus*, WP_030828871	76/83
514	248	type II thioesterase	thioesterase, *S. mobaraensis*, EME98724	64/76
515 [c]	294	methyltransferase	methyltransferase tpe 12, *S. bingchenggensis*, ADI12483	70/82
516 [c]	180	unknown	GCN5 family acetyltransferase, *Streptomyces* sp. NRRL S-87, WP_030200822	68/78
517	532	Acyl CoA carboxylase α-subunit	carboxyl transferase, *S. Rimosus*, ELQ79111	82/89
518 [c]	947	transcriptional regulator	hypothetical protein, *Lechevalieria aerocolonigenes*, WP_030473578	45/57
519 [c]	532	oxidoreductase	hypothetical protein, *L. aerocolonigenes*, WP_030473579	64/76
520	4917	type I PKS	putative type I polyketide synthase, *S. bingchenggensis*, ADI04502	53/64
521	3601	type I PKS	polyketide synthase, *S. rapamycinicus*, AGP57755	53/64
522	2771	type I PKS	type I polyketide synthase, *S. flaveolus*, ACY06287	52/63
523	1322	type I PKS	hypothetical protein, partial, *S. novaecaesareae*, WP_033330653	52/61
524	408	cytochrome P450	cytochrome P450, *L. aerocolonigenes*, WP_030470230	72/84
525	115	cyclase	hypothetical protein, *L. aerocolonigenes*, WP_030470229	63/79
526	78	ferredoxin	hypothetical protein BN6_14320, *Saccharothrix espanaensis* DSM 44229, CCH28755	61/73
527	402	cytochrome P450	cytochrome P450, *L. aerocolonigenes*, WP_030470226	73/84
528	229	transcriptional regulator	NmrA family transcriptional regulator, *L. aerocolonigenes*, WP_030470224	76/84
529	498	transporter	Puromycin resistance protein pur8, partial, *L. aerocolonigenes*, WP_030470223	66/76
530 [c]	216	transcriptional regulator	NmrA family transcriptional regulator, *L. aerocolonigenes*, WP_030470222	70/82

531	109	transcriptional regulator	HxlR family transcriptional regulator, *L. aerocolonigenes*, WP_030470221	69/83
532	457	enoyl-CoA reductase/carboylase	NADPH: quinone reductase, *L. aerocolonigenes*, WP_030470219	79/87
533	340	3-oxoacyl ACP synthase	3-oxoacyl-ACP synthase, *S. Olivaceus*, WP_031035846	70/82
534	286	3-hydroxyacyl CoA dehydrogenase	3-hydroxybutyryl-CoA dehydrogenase, *L. aerocolonigenes*, WP_030470217	63/80
535	339	enoylreductase	putative succinate-semialdehyde dehydrogenase (acetylating), *S. afghaniensis*, EPJ38274	68/77
536	>213 [d]	transposase	transposase, *Streptomyces* sp. NRRL B-24484, WP_030264850	78/87

[a] in aa; [b] identity/similarity; [c] encoded in complementary strand; [d] located at scaffold terminal.

Based on the results obtained from gene analysis (Table 3-3), biosynthetic pathway for akaeolide is proposed (Figure 3-11). First, the polyketide backbone is synthesized by four type I PKSs, Orf8-520 to -523 containing all the catalytic domains necessary for the construction of the carbon framework. According to the KR fingerprints rule (Figure 3-12, quotation marks on *R* and *S* represents the conventional *RS* nomenclature system adapted only for this purpose. The definition is different from the IUPAC nomenclature rule: the γ position is given the lowest priority after the hydrogen when discussing the chirality at the β position[15]), both of the KR domains in modules 2 and 6 belong to B1-type, suggesting that the absolute configuration of hydroxyl groups at C-5 and C-11 is '*S*' and that of the propyl group at C-4 is '*R*'.[16] The absolute configuration of the methyl group at C-12 was suggested to be '*S*' on the basis of amino acid alignment in the ER domain of module 2.[17] These predictions were completely consistent with the stereochemistry of the actual natural product. After the release from the enzyme along with lactone formation, the putative linear product might undergo cyclic ether formation although it was unable to assign any genes responsible for this cyclization reaction only from the annotation data. Two cytochromes P450 would be involved in hydroxylation at C-10 and oxidation of the methyl group at C-16 to the aldehyde. The next step could be carbon-carbon bond formation between the activated methylene C-17 and the aldehyde at C-16. This reaction would be catalyzed by a putative cyclase Orf8-525 that was predicted to have a similar function to SnoaL on the basis of domain homology by InterPro.[18] SnoaL is a polyketide cyclase that catalyzes aldol condensation in the biosynthesis of nogalamycin, an aromatic polyketide produced by *Streptomyces* (Figure 3-13).[19,20] After dehydration from the aldol product, Orf8-535 coding enoylreductase would act to reduce the conjugated C-C double bond between C-16 and C-17. These functional assignments are mostly reasonable to explain the biosynthesis of akaeolide (**1**).

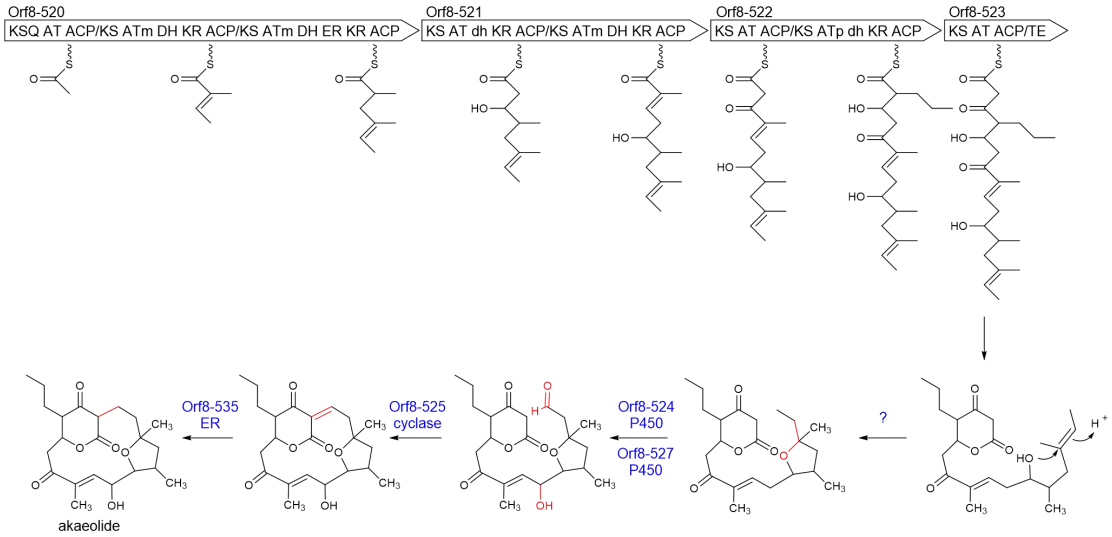


Figure 3-11. PKS domain organization and plausible biosynthetic pathway for akaeolide.

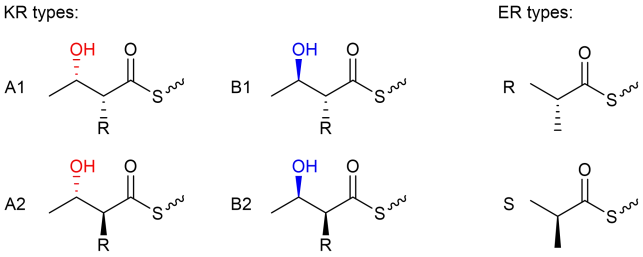


Figure 3-12. Different types of KR and ER domains in PKS predicting stereochemistry of products.

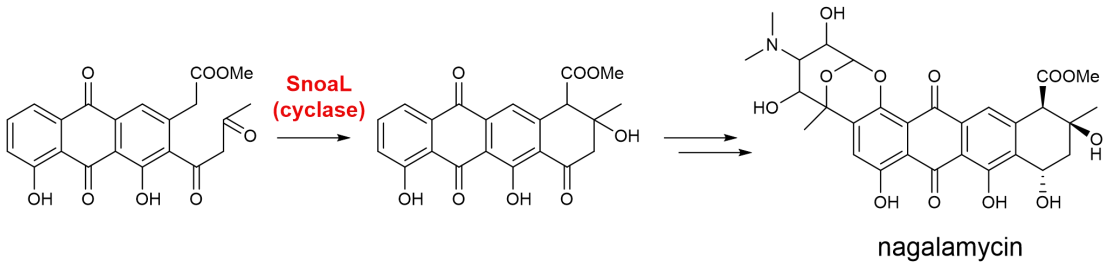


Figure 3-13. Aldol condensation catalyzed by SnoaL in the biosynthesis of nogalamycin.

3-2-2-2 Biosynthetic Gene Cluster for Lorneic Acids

Since the gene cluster for lorneic acids biosynthesis were assigned, the biosynthetic pathway is further proposed from genomic information.

Gene organization of the biosynthetic gene cluster for lorneic acids is shown in Figure 3-14 and proposed functions of the PKS genes and the neighboring genes are summarized in Table 3-4. The PKS gene cluster is composed of one loading module and seven extension modules (Figure 3-15). Signature amino-acid residues of AT domains predicted the substrate specificity: acetyl-CoA (C_2) would be loaded onto LM, malonyl-CoA (C_2) onto m1, m2, m3, m5, m6, and m7, and methylmalonyl-CoA (C_3) onto m4. Therefore, the polyketide carbon chain would be assembled by

sequential incorporation of C_2, C_2, C_2, C_2, C_3, C_2, C_2, and C_2 units, consistent with the results obtained from the ^{13}C-labeling experiments (Figure 3-8). The PKS cluster contains three DH/ER/KR domains in m1, m2, and m4 and four DH/KR domains in m3, m5, m6, and m7, corresponding to three saturated methylene carbons and four double bonds, respectively. However, the ER domain in m4 might be unfunctional if a putative intermediate A (Figure 3-15) were generated as described below.[23]

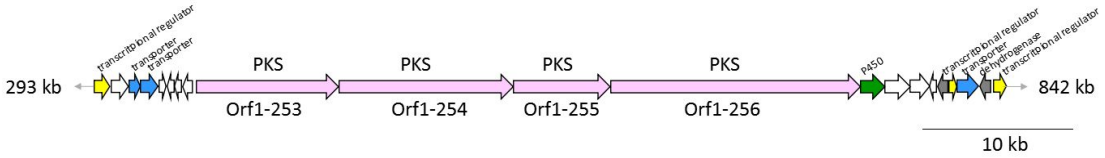


Figure 3-14. Gene organization of the biosynthetic gene cluster for lorneic acids.

Table 3-4. Proposed functions of type I PKS genes for lorneic acid biosynthesis and neighboring genes.

Orf1-	Size [a]	Proposed Function	Protein Homolog, *Origin*, Accession Number	% [b]
245	372	transcriptional regulator	transcriptional regulator, *Cellulosimicrobium cellulans*, WP_024839151	80/89
246	376	unknown	hypothetical protein, *C. cellulans*, WP_024839150	76/84
247	255	ABC transporter	sugar ABC transporter ATP-binding protein, *C. cellulans*, WP_024839149	79/86
248	418	ABC transporter	ABC transporter permease, *C. cellulans*, WP_024839148	74/82
249	138	unknown	hypothetical protein AOR_1_82014, *Aspergillus oryzae*, XP_003189833	50/60
250 [c]	185	unknown	hypothetical protein, *Streptomyces* sp. CNQ766, WP_018834862	24/43
251 [c]	158	unknown	hypothetical protein, *Jiangella gansuensis*, WP_026874523	56/66
252 [c]	239	unknown	hypothetical protein, *Actinopolymorpha alba*, WP_020576671	60/70
253	3128	type I PKS	hypothetical protein, *Nocardia* sp. BMG51109, WP_024802962	59/70
254	3893	type I PKS	hypothetical protein, *Nocardia* sp. BMG51109, WP_024802962	58/69
255	2176	type I PKS	hypothetical protein, *Nocardia* sp. BMG51109, WP_024802960	68/78
256	5477	type I PKS	hypothetical protein, *Nocardia* sp. BMG51109, WP_024802963	62/72
257	534	cytochrome P450	cytochrome P450, *Micromonospora* sp. ATCC 39149, EEP70569	73/84

258	588	unknown	hypothetical protein (glycosidase), *Streptomyces* sp. CNY243, WP_018851787	79/87
259	435	unknown	beta-lactamase, *Catenulispora acidiphila*, ACU74739	80/87
260 [c]	118	unknown	PDZ and LIM domain protein 1, partial, *Columba livia*, EMC79925	31/52
261 [c]	254	methyltransferase	Methyltransferase, *S. avermitilis*, BAC68374	84/89
262	148	transcriptional regulator	Rrf2 family transcriptional regulator, *Nocardia* sp. BMG51109, WP_024804268	78/86
263	486	transporter	MFS transporter, *Nocardia* sp. BMG51109, WP_024804269	68/80
264 [c]	250	Short-chain dehydrogenase	short-chain dehydrogenase, *A. alba*, WP_020577481	92/94
265	331	transcriptional regulator	AraC family transcriptional regulator, *A. alba*, WP_020577482	92/94

[a] in aa; [b] identity/similarity; [c] encoded in complementary strand.

Annotation of the domain function for the PKS genes indicated the formation of a linear polyolefinic acid starting from an acetate to which six malonates and one methylmalonate are condensed (Figure 2-15). This polyene intermediate tethered to the PKS would have all-*trans* double bonds because the KR domain in module 5 is B-type.[24] After being released from the enzyme, the intermediate would undergo *cis/trans*-double bond isomerization at C-6/C-7 and cyclization. Orf1-257 located downstream of the PKS genes shows high homology to cytochromes P450, suggesting the involvement of P450-mediated monooxygenation in the aromatization process from the polyolefinic precursor (**A**) and the introduction of a hydroxyl group at C-11. Epoxidation of **A** by Orf1-257 would give an epoxide (**B**), which could undergo aromatization through an unknown pathway to provide lorneic acid B (**4**) (Figure 2-15). Further dehydration might afford lorneic acid A (**3**). This pathway is likely plausible since the addition of one oxygen atom to the intermediate **A** ($C_{17}H_{24}O_2$ + O) can give rise to the intermediate **B** that has the same molecular formula as that of lorneic acid B (**4**, $C_{17}H_{24}O_3$).

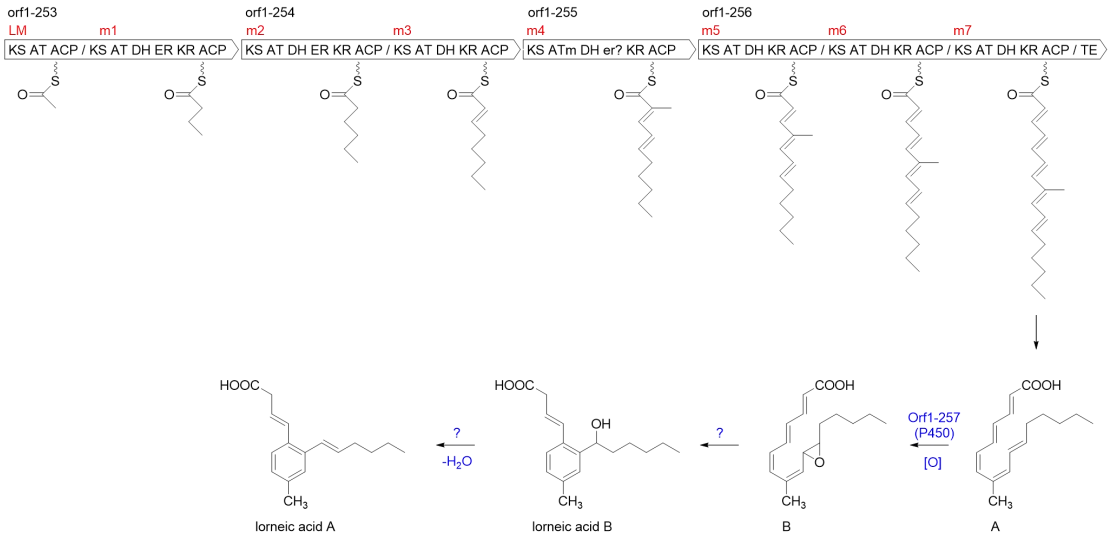


Figure 3-15. Proposed biosynthetic pathway for lorneic acids.

3-2-2-3 Distribution of Biosynthetic Gene Clusters for Akaeolide and Lorneic Acid in Other Strains

Uncommon structural features of akaeolide (**1**) and lorneic acids (**3** and **4**) prompted me to investigate the distribution of related biosynthetic genes in other organisms. Homologues of PKS genes for akaeolide and lorneic acid biosynthesis were searched in GenBank DNA database by using BLAST.[25] An orthologous gene cluster to akaeolide cluster is likely present in *L. aerocolonigenes* NRRL B-3298T (Table 3-3) but its presence is unclear because the cluster was divided into several contigs (contig25.1, contig88.1, contig59.1 etc.) and not completely sequenced. Therefore, genome sequencing of *L. aerocolonigenes* NBRC 13195T (the same strain as NRRL B-3298T, deposited at NBRC, Japan) was undertaken to obtain the complete gene sequence. The PKS domain organization of akaeolide cluster orthologue in *L. aerocolonigenes* NBRC 13195T was in good agreement with that of the akaeolide cluster in *Streptomyces* sp. NPS554 (data not shown) as well as the synteny of the cluster (Figure 3-10b). This orthologous cluster present in *L. aerocolonigenes* might be responsible for the production of mangromicins, congeners of akaeolide, isolated from *L. aerocolonigenes* K10-0216.[26] *Nocardia* sp. BMG51109 is the only organism that has the orthologous genes to the PKS genes for lorneic acids (Table 3-4) but no secondary metabolites are known from this strain.

3-2-2-4 Other Type I PKS Gene Clusters in Streptomyces sp. NPS554

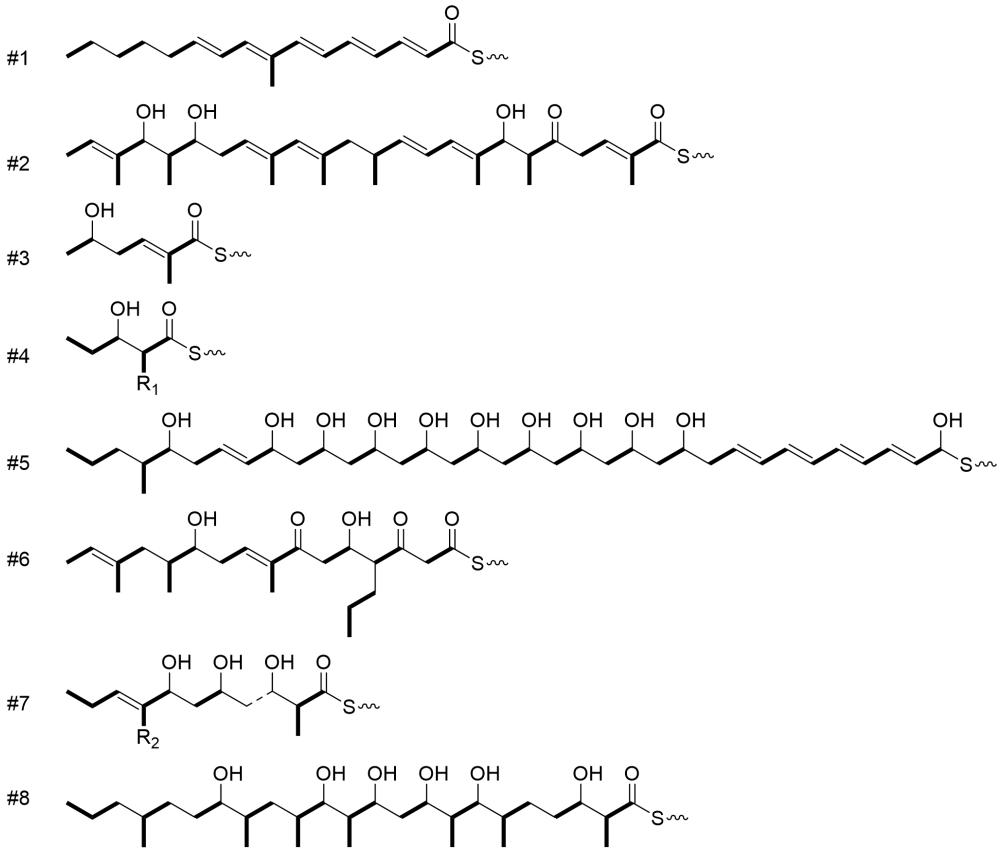


Figure 3-16. Predicted structures of linear products of orphan type I PKS gene clusters in *Streptomyces* sp. NPS554.

In addition to akaeolide and lorneic acid clusters, eight orphan type I PKS gene clusters are present in *Streptomyces* sp. strain NPS554 (Table S1). As for the six clusters completely sequenced, chemical structures of the linear products were predicted on the basis of domain organization (Figure 3-16) and stereochemistry of methyl and hydroxyl groups was predicted (Table S1). Clusters #2, #5, and #8 were predicted to yield linear polyketides comprising of 24-, 35- and 23-carbon chain length, respectively, while the products of clusters #3 and #4 are much smaller, consisting of less than ten carbons. The structure of the product from cluster #7 was not completely predicted since one of the PKS modules lacked an AT domain. The side chains R_1 and R_2 could not be assigned only from the signature amino acid-sequences. Structures of the products from clusters #9 to #12 could not be predicted because these clusters were not completely sequenced.

3-3 Materials and Methods

General Experimental Procedures. Sodium [1-^{13}C]acetate, sodium [2-^{13}C]acetate, sodium [1-^{13}C]propionate, and [1,4-$^{13}C_2$]succinic acid were purchased from Cambridge Isotope Laboratories, Inc. (Andover, MA, USA). ^{13}C NMR spectra were obtained on a Bruker AVANCE 500 spectrometer (Bruker BioSpin K.K., Yokohama, Japan). Cosmosil 75C18-PREP (Nakalai Tesque, Inc., Kyoto, Japan, 70 μm) was used for ODS column chromatography. HPLC separation was performed using a Cosmosil 5C18-AR-II (Nacalai Tesque Inc., 20 × 250 mm, Kyoto, Japan) with a photodiode array detector.

Isolation of Akaeolide (1) and Lorneic Acid A (3). Strain NPS554 cultured on a Bn-2 slant (soluble starch 0.5%, glucose 0.5%, meat extract (Kyokuto Pharmaceutical Industrial Co., Ltd., Tokyo, Japan) 0.1%, yeast extract (Difco Laboratories, Surrey, United Kingdom) 0.1%, NZ-case (Wako Pure Chemical Industries, Ltd., Osaka, Japan) 0.2%, NaCl 0.2%, $CaCO_3$ 0.1%, agar 1.5%) was inoculated into 500 mL K-1 flasks each containing 100 mL of the V-22 seed medium consisting of soluble starch 1%, glucose 0.5%, NZ-case 0.3%, yeast extract 0.2%, Tryptone (Difco Laboratories) 0.5%, K_2HPO_4 0.1%, $MgSO_4 \cdot 7H_2O$ 0.05%, and $CaCO_3$ 0.3% (pH 7.0) in natural seawater. The flasks were placed on a rotary shaker (200 rpm) at 30 ˚C for 4 days. The seed culture (3 mL) was transferred into 500 mL K-1 flasks each containing 100 mL of the A-16 production medium consisting of glucose 2%, Pharmamedia (Traders Protein) 1%, and $CaCO_3$ 0.5% in natural seawater. The inoculated flasks were placed on a rotary shaker (200 rpm) at 30 ˚C for 6 days. At the end of the fermentation period, 100 mL of 1-butanol was added to each flask, which was allowed to shake for 1 h. The mixture was centrifuged at 5000 rpm for 10 min and the organic layer was separated from the aqueous layer containing the mycelium. After evaporation of the solvent, the crude extract was subjected to silica gel column chromatography with a step gradient of $CHCl_3$/MeOH (1:0, 20:1, 10:1, 4:1, 2:1, 1:1, and 0:1 v/v). Fractions 2 and 3 ($CHCl_3$/MeOH = 20:1 and 10:1) were concentrated to provide a brown solid, which was further purified by repeated reverse phase preparative HPLC using a Cosmosil 5C18-AR-II column (Nacalai Tesque Inc., 10 × 250 mm) with MeCN in 0.1% HCO_2H (35:65, flow rate 4 mL/min), followed by evaporation and extraction with EtOAc to give akaeolide (**1**, 9.7 mg/L) and lorneic acid A (**3**, 13.5 mg/L).

Chlorination of 1 to Yield 17-Chloroakaeolide (2). To a solution of 1 (5.2 mg, 13.3 μmol) in CH_2Cl_2 (1 mL) were added 2,6-lutidine (0.82 μL, 7.1 μmol) and *N*-chlorosuccinimide (3.6 mg, 27 μmol) at room temperature. After stirring for 2 h, the reaction mixture was diluted with diethyl ether (2.5 mL) and washed with brine (1 mL). The organic layer was separated and concentrated *in vacuo* to give 5.8 mg of crude residue. Purification by silica gel chromatography (n-hexane/EtOAc=10:1~1:1) yielded 17-chloroakaeolide (**2**, 2.2 mg, 39% yield) as a white powder.

Preparation of ^{13}C-Labeled Akaeolide (1). The inoculation, cultivation, extraction, and purification were performed in the same manner as described above. Addition of ^{13}C-labeled precursors (1 mL solution/flask) was initiated at 48 h after inoculation and periodically carried out every 24 h for four times. After further incubation for 24 h, the cultures were extracted with 1-butanol.

(1) Sodium [1-^{13}C]acetate: After feeding of sodium [1-^{13}C]acetate (total 800 mg; 20 mg × 10 flasks × 4 days), 9.7 mg of ^{13}C-labeled **1** was obtained from 1 L of culture. The ^{13}C NMR spectrum showed enriched signals at δ 34.7, 75.4, 140.6, 164.5 and 196.4.

(2) Sodium [2-^{13}C]acetate: After feeding of sodium [2-^{13}C]acetate (total 800 mg; 20 mg × 10 flasks × 4 days), 9.2 mg of ^{13}C-labeled **1** was obtained from 1 L of culture. The ^{13}C NMR spectrum showed enriched signals at δ 30.1, 43.5, 46.8, 63.7 and 69.0.

(3) Sodium [1-^{13}C]propionate: After feeding of sodium [1-^{13}C]propionate (total 800 mg; 20 mg × 10 flasks × 4 days), 8.7 mg of ^{13}C-labeled **1** was obtained from 1 L of culture. The ^{13}C NMR spectrum showed enriched signals at δ 28.8, 49.1, 84.1 and 203.1.

Preparation of ^{13}C-Labeled Lorneic Acid A (3). The inoculation, cultivation, extraction, and purification were performed in the same manner as described above. Addition of ^{13}C-labeled precursors (1 mL solution/flask) was initiated at 48 h after inoculation and periodically carried out every 24 h for 4 times. After further incubation for 24 h the cultures were extracted with 1-butanol.

(1) Sodium [1-^{13}C]acetate: After feeding of sodium [1-^{13}C]acetate (total 800 mg; 20 mg × 10 flasks × 4 days), 13.5 mg of ^{13}C-labeled **3** was obtained from 1 L of culture. The ^{13}C NMR spectrum showed enriched signals at δ 22.3, 33.0, 121.8, 127.0, 127.4, 131.7 and 177.7.

(2) Sodium [1-^{13}C]propionate: After feeding of sodium [1-^{13}C]propionate (total 800 mg; 20 mg × 10 flasks × 4 days), 8.5 mg of ^{13}C-labeled **3** was obtained from 1 L of culture. The ^{13}C NMR spectrum showed enriched signals at δ 127.8.

Whole Genome Shotgun Sequencing. Whole genomes of *Streptomyces* sp. NPS554 (=NBRC 109706) and *L. aerocolonigenes* NBRC 13195T were read using a combined strategy of shotgun sequencing with GS FLX+ (Roche; 58 Mb sequences and 7.4-fold coverage for NPS554, 96 Mb sequences and 9.0-fold coverage for NBRC 13195) and paired-end sequencing with MiSeq (Illumina; 705 Mb sequences and 91-fold coverage for NPS554, 770 Mb sequences and 71-fold coverage for NBRC 13195). These reads were assembled using a Newbler v2.6 software, and subsequently finished using a GenoFinisher software,[27] which led to a final assembly of 26 and 55 scaffold sequences of >500 bp each for strains NPS554 and NBRC 13195, respectively. The total size of the NPS554 assembly was 7,736,999 bp, with a G+C content of 71.7%, while that of

NBRC 13195 one was 10,698,154 bp, with a G+C content of 68.9%. The draft genome sequences of *Streptomyces* sp. NPS554 and *L. aerocolonigenes* NBRC 13195T were deposited in GenBank/ENA/DDBJ under the accession numbers BBOM00000000 and BBOJ00000000, respectively.

Annotation of PKS Gene Clusters. The resulting scaffold sequences were submitted to the in-house auto-annotation pipeline. Coding sequences (CDSs) were predicted by Prodigal v2.6[28] and searched for domains related to polyketide synthase (PKS) genes such as ketosynthase (KS) domain using the SMART and PFAM domain databases. PKS gene clusters and their domain organizations were determined as reported.[29] Gene functions were assigned according to BLAST searches conducted using the NCBI BlastP program[25] against the non-redundant protein sequences (nr) database. Substrates of each acyltransferase (AT) domains were predicted based on substrate-specific signature amino-acid sequences.[9,10]

References

1 Uramoto M, Otake N, Ogawa Y, *et al. Tetrahedron Lett*. **1969**, 2249-2254.

2 Shindo K, Matsuoka M, Kawai H. *J. Antibiot*. **1996**, 49, 241-243.

3 Shindo K, Iijima H, Kawai H. *J. Antibiot*. **1996**, 49, 244-248.

4 Hochlowski JE, Mullally MM, Henry R, *et al. J. Antibiot*. **1995**, 48, 467-470.

5 Arakawa K, Sugino F, Kodama K, *et al. Chem. Biol*. **2005**, 12, 249-259.

6 Shindo K, Sakakibara M, Kawai H, *et al. J. Antibiot*. **1996**, 49, 249-252.

7 Hashimoto M, Komatsu H, Kozone I, *et al. Biosci. Biotechnol*. **2005**, 69 , 315-320.

8 Jones CA, Sidebottom PJ, Cannell RJP, *et al*. *J. Antibiot*. **1992**, *45*, 1492-1498.

9 Del Vecchio F, Petkovic H, Kendrew SG, *et al. J. Ind. Microbiol. Biotechnol*. **2003**, 30, 489-494.

10 Kakavas SJ, Katz L, Stassi D. *J. Bacteriol*. **1997**, 179, 7515-7522.

11 Eustaquio AS, McGlinchey RP, Liu Y, *et al. Proc. Natl. Acad. Sci. USA*. **2009**, 106, 12295-12300.

12 Liu Y, Hazzard C, Eustáquio AS, *et al. J. Am. Chem. Soc*. **2009**, 131, 10376-10377.

13 Goranovic D, Kosec G, Mrak P, *et al. J. Biol. Chem*. **2010**, 285, 14292-14300.

14 Fischbach MA, Walsh CT. *Chem. Rev*. **2006**, 106, 3468-3469.

15 Caffrey P. *Chem. Biol*. **2005**, 12, 1060-1062.

16 Keatinge-Clay A. *Chem. Biol*. **2007**, 14, 898-908.

17 Kwan DH, Sun Y, Schulz F, *et al. Chem. Biol*. **2008**, 15, 1231-1240.

18 InterProScan Sequence Search.Available online: http://www.ebi.ac.uk/interpro/search/ sequence- search (accessed on 1 December 2014).

19 Torkkell S, Kunnari T, Palmu K, *et al. Antimicrob. Agents Chemother*. **2000**, 44, 396-399.

20 Sultana A, Kallio P, Jansson A, *et al. EMBO J*. **2004**, 23, 1911-1921.

21 Robert JC, Colin S, Ernest L, *et al. J. Nat. Prod*. **2000**, 63, 1682-1683.

22 Natsuki KK, Shinji O, Emi O, *et al. J. Nat. Prod*. **2004**, 67, 85-87.

23 Ichikawa N, Sasagawa M, Yamamoto M, *et al. Nucleic Acids Res*. **2013**, 41, D408-D414.

24 Alhamadsheh MM, Palaniappan N, DasChouduri S, *et al. J. Am. Chem. Soc*. **2007**, 129, 1910-1911.

25 BLAST Assembled Genomes. Available online: http://blast.ncbi.nlm.nih.gov/ (accessed on 1 December 2014).

26 Nakashima T, Iwatsuki M, Ochiai J, *et al. J. Antibiot*. **2014**, 67, 253-260.

27 Ohtsubo Y, Maruyama F, Mitsui H, *et al. J. Bacteriol.* **2012**, 194, 6970-6971.

28 Hyatt D, Chen GL, Losascio PF, *et al. BMC Bioinformatics*. **2010**, 11, 119.

29 Komaki H, Ichikawa N, Oguchi A, *et al. J. Gen. Appl. Microbiol*. **2012**, 58, 363-372.

3-4 Spectral Data and Genomic Information

Table of Contents

Figure S1. ^{13}C NMR spectra of 17-chloroakaeolide (**2**) labeled with [1-^{13}C]acetate (**A**) and [2-^{13}C]acetate (**B**) (100 MHz, $CDCl_3$).

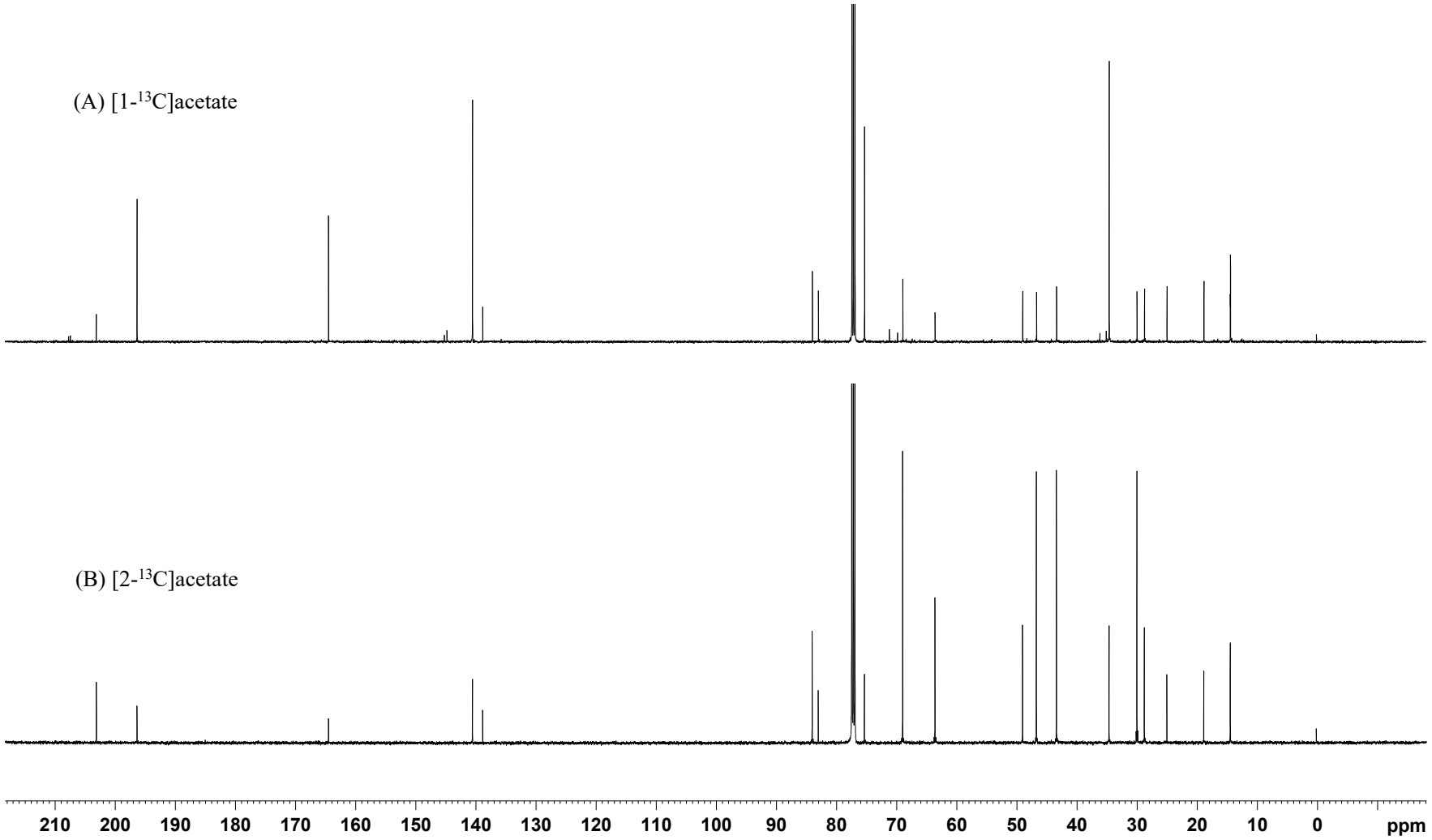


Figure S2. ^{13}C NMR spectra of 17-chloroakaeolide (**2**) labeled with [1-^{13}C]propionate (**A**) and [1,4-$^{13}C_2$]succinate (**B**) (100 MHz, $CDCl_3$).

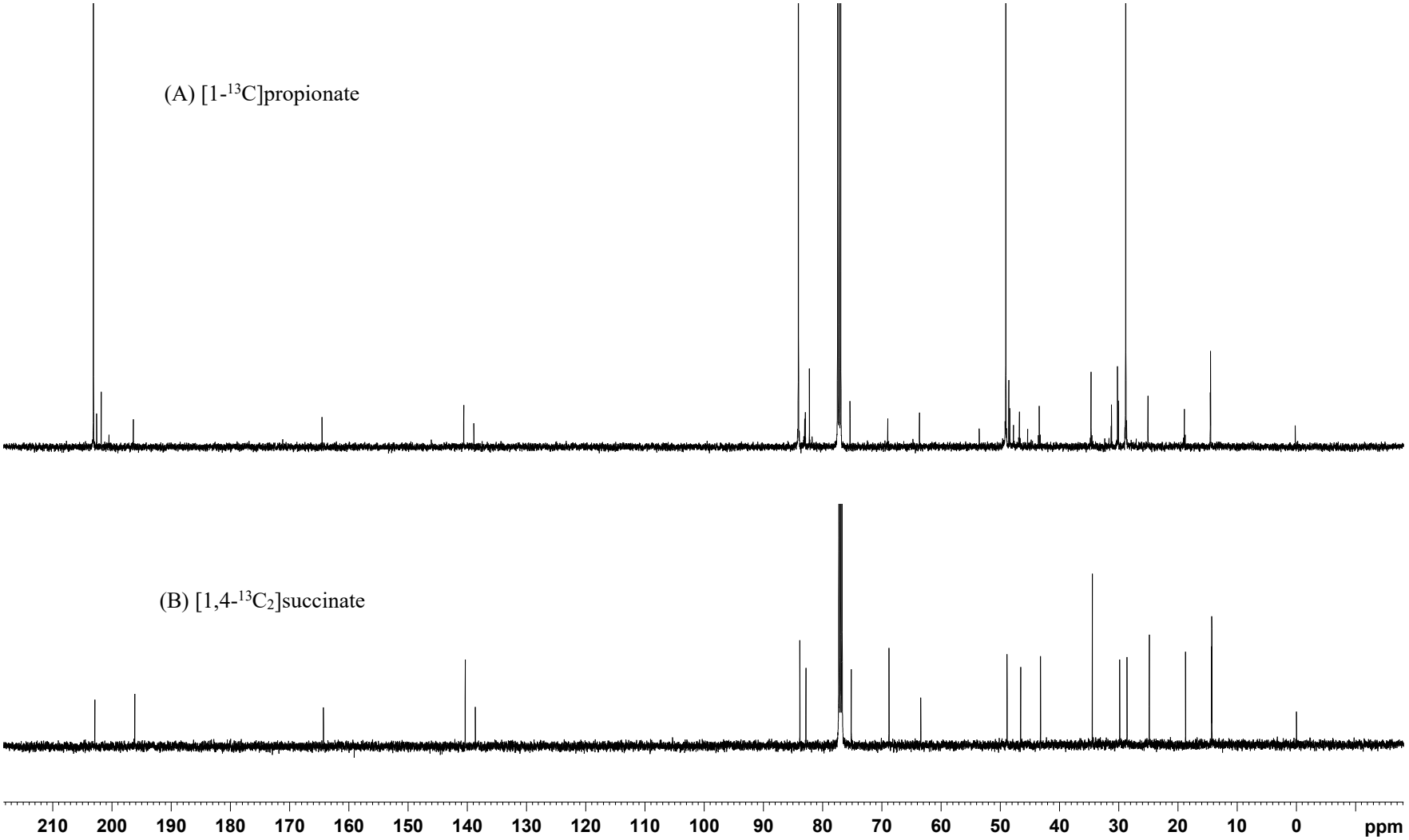


Figure S3. ^{13}C NMR spectra of lorneic acid A (**3**) labeled with [1-^{13}C]acetate (**A**) and [1-^{13}C]propionate (**B**) (100 MHz, CDCl3).

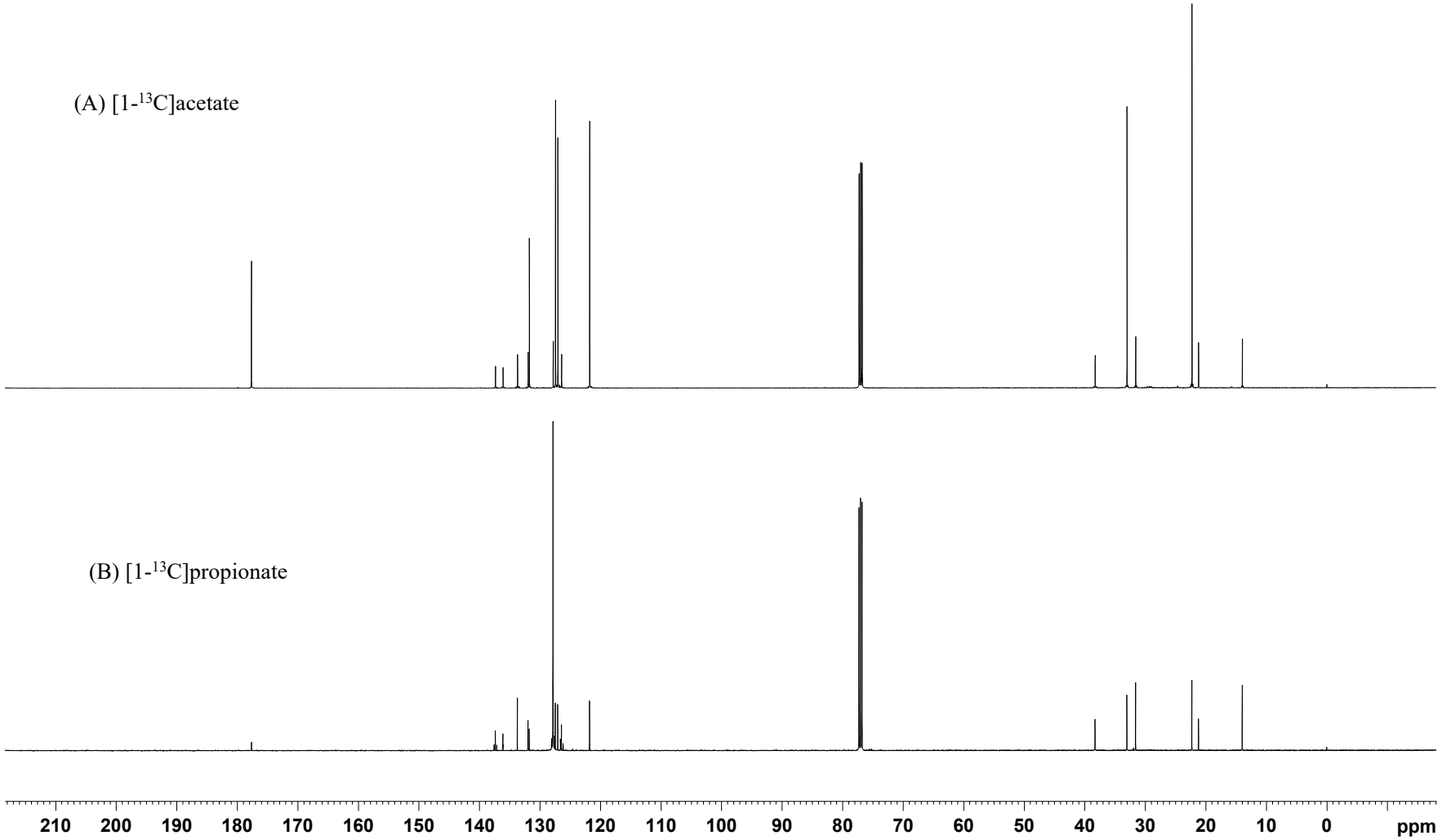


Table S1. Multimodular type I PKS gene clusters in *Streptomyces* sp. NPS554 genome

Cluster	Module	Orf **	Domain Organization ***
#1	8	1-253	KS/AT/ACP-KS/AT/DH/ER$_S$/KR$_{B1}$/ACP
		1-254	KS/AT/DH/ER$_S$/KR$_{B1}$/ACP-KS/AT/DH/KR$_{B1}$/ACP
		1-255	KS/AT/DH/ER$_S$/KR$_{B1}$/ACP
		1-256	KS/AT/DH/KR$_{B1}$/ACP-KS/AT/DH/KR$_{B1}$/ACP-KS/AT/DH/KR$_{B1}$/ACP/TE
#2	12	4-545	KS/AT/ACP-KS/ATm/DH/KR$_{B1}$/ACP-KS/ATm/KR$_{B2}$/ACP-KS/AT/KR$_{A1}$/ACP-KS/ATm/DH/KR$_{B1}$/ACP
		4-544	KS/ATm/DH/KR$_{B1}$/ACP-KS/ATm/DH/ER$_R$/KR$_{B1}$/ACP-KS/AT/DH/KR$_{B1}$/ACP
		4-543	KS/ATm/DH/KR$_{B1}$/ACP-KS/ATm/KR$_{B1}$/ACP
		4-542	KS/AT/ACP-KS/ATm/DH/KR$_{B1}$/ACP/TE
#3	3	5-51	KS/AT/ACP-KS/AT/KR$_{B1}$/ACP
		5-52	KS/ATm/DH/KR$_{B1}$/ACP/TE
#4	2	5-363	KS/ATm/ACP-KR$_?$
		5-364	KS/ATx
		5-365	ACP
#5	17	8-406	KS/ATm/ACP-KS/ATm/DH/ER$_R$/KR$_{B1}$/ACP-KS/AT/DH/KR$_{B1}$/ACP
		8-407	KS/AT/DH/KR$_{B1}$/ACP-KS/AT/DH/KR$_?$/ACP-KS/AT/KR$_?$/ACP-KS/AT/KR$_{A1}$/ACP
		8-408	KS/AT/KR$_?$/ACP-KS/AT/KR$_{A1}$/ACP-KS/AT/KR$_{A1}$/ACP-KS/AT/KR$_{A1}$/ACP-KS/AT/KR$_{A1}$/ACP-KS/AT/DH/KR$_{B1}$/ACP
		8-409	KS/AT/DH/KR$_{B1}$/ACP-KS/AT/DH/KR$_{B1}$/ACP-KS/AT/DH/KR$_{B1}$/ACP-KS/AT/DH/KR$_{B1}$/ACP/TE
#6	8	8-520	KS/AT/ACP-KS/ATm/DH/KR$_{B1}$/ACP-KS/ATm/DH/ER$_S$/KR$_{B1}$/ACP
		8-521	KS/AT/DH/KR$_{B1}$/ACP-KS/ATm/DH/KR$_{B1}$/ACP
		8-522	KS/AT/ACP-KS/ATp/DH/KR$_{B1}$/ACP
		8-523	KS/AT/ACP/TE

#7	5	9-219	KS/ATm/ACP-KS/ATx/DH/KR_{B1}/ACP
		9-220	KS/AT/$KR_{?}$/ACP
		9-221	KS/$KR_{?}$/ACP-KS/ATm/$KR_{?}$/ACP
#8	9	13-34	KS/ATm/ACP-KS/ATm/DH/ER_{R}/KR_{B1}/ACP
		13-31	KS/AT/DH/ER_{S}/KR_{B1}/ACP
		13-30	KS/ATm/KR_{A2}/ACP
		13-29	KS/ATm/DH/ER_{S}/KR_{B1}/ACP-KS/ATm/KR_{A1}/ACP-KS/AT/KR_{A1}/ACP
		13-28	KS/ATm/KR_{B2}/ACP
		13-27	KS/ATm/KR_{B1}/ACP-KS/AT/DH/ER_{R}/KR_{B1}/ACP
		13-25	KS/ATm/KR_{B2}/ACP-KS
#9 *	>13	5-1	KS/AT/DH/ER/KR/ACP-KS/AT/KR/ACP-KS/ATm/KR/ACP-KS/AT/DH/KR/ACP
		5-2	KS/AT/DH/KR/ACP-KS/AT/DH/KR/ACP-KS/AT/DH/ER/KR/ACP-KS/AT/DH/ER/KR/ACP KS/AT/DH/KR/ACP
		5-3	KS/AT/KR/ACP-KS/ATx/DH/ACP
		5-4	KS/ATm/DH/ER/KR/ACP
		5-5	KS/ATm/DH/ER/KR/ACP/TE
#10 *	>17	5-412	ACP-KS/AT/KR/ACP-KS/AT/KR/ACP-KS/AT/DH/KR/ACP-KS/AT/DH/KR/ACP
		5-411	KS/AT/KR/ACP-KS/ATm/KR/ACP-KS/AT/KR/ACP-KS/AT/KR/ACP/TE
		5-409	KS/AT/DH/ER/KR/ACP-KS/ATx/KR/ACP-KS/ATm/DH/KR/ACP
		5-408	KS/AT/KR/ACP
		5-407	KS/AT/KR/ACP-KS/AT/KR/ACP-KS/AT/KR/ACP-KS/AT/DH/KR/ACP-KS/AT/DH/KR/ACP
#11 *	>3	13-1	KS/AT/ACP-KS/AT/DH/ER/KR/ACP-KS/AT/DH/KR/ACP-
#12 *	>19	14-1	ATm/DH/KR/ACP-KS/ATm/KR/ACP-KS/AT/DH/ER/KR/ACP-KS/ATm/DH/KR/ACP-KS/AT/KR/ACP-KS/AT/KR/ACP
		14-2	KS/ATm/DH/ER/KR/ACP-KS/ATm/KR/ACP-KS/AT/KR/ACP-KS/AT/KR/ACP-KS/ATm/DH/KR/ACP-KS/AT/KR/ACP
		14-3	KS/AT/KR/ACP-KS/AT/DH/KR/ACP-KS/AT/DH/KR/ACP-KS/AT/DH/KR/ACP-KS/AT/DH/KR/ACP-KS/AT/KR/ACP-KS /AT/KR/ACP

PKS gene clusters only with a single PKS module and short PKS gene fragments are not included. * #9 to #12, partial, probably divided into multiple scaffolds; might be joined to form one cluster. ** scaffold numbers are shown before hyphen in the orf numbers. KS, ketosynthase domain; AT, acyltransferase domain for malonyl-CoA; ATm, AT for methylmalonyl-CoA; ATp, AT for propylmalonyl-CoA; ATx, AT whose substrate is unpredictable; DH, dehydratase domain; ER, enoylreductatse domain; KR, ketoreductase domain; ACP, acyl carrier protein domain; TE, thioesterase domain. *** KR is classified into A1-, A2-, B1-, B2-, C1-, C2-type, or unknown (?) based on KR fingerprints.[13] ER is classified into *R*- or *S*-type.[15]

CHAPTER 4

Bioinformatics-Inspired Isolation of Akaemycin

4-1 Background

As described in Chapter 3, eight orphan type I PKS (polyketide synthase) gene clusters were found to be responsible for production of unknown metabolites (Figure 3-16, Chapter 3) and their structures were predicted. Among these structures, two of them are responsible for the synthesis of lorneic acid A (#1) and akaeolide (#6). Within the remaining six products, three of them have short carbon-chains (#3, #4, and #7) with less than ten carbons, which are proposed to be smaller molecules; on the other hand, the remaining three products (#2, #5, and #8) have longer carbon-chains with 24, 35, and 23 carbons respectively. The maximum UV absorption for these three larger molecules is proposed to be 240 nm (#2), 330 nm (#5), and 200 nm (#8). Based on this information, the HPLC data of crude extract from the culture of NPS554 strain was re-examined, because there was an unknown component with UV maximum absorption at 330 nm, as discussed in Chapter 2. This compound is very likely to be the same product of #5 in genome-based prediction. As any identical or similar compounds are not present in SciFinder database, further HPLC/UV-guided isolation was undertaken, which led to the discovery of a new macrolide designated akaemycin (**4**, Figure 4-1). The structure of **5** was identical with the structure predicted by annotation of PKS gene functions. Herein, the isolation, structural determination, and biological properties of **4** are described.

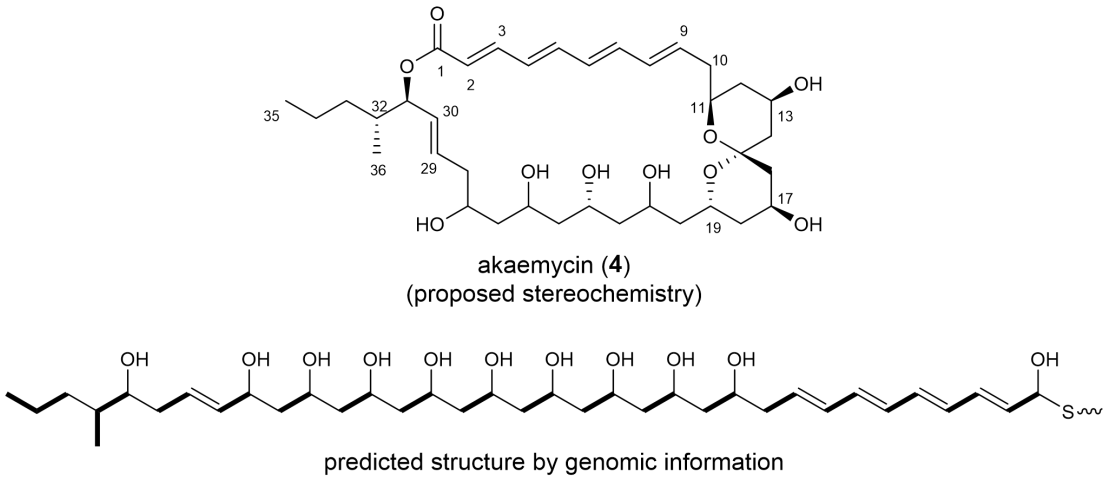


Figure 4-1. Structures of akaemycin (**4**, proposed stereochemistry) and predicted structure of linear product from gene clusters.

4-2 Results and Discussion

4-2-1 Fermentation and Isolation

Streptomyces sp. NPS554 was cultured in seawater-based A-16 liquid medium. HPLC-DAD analysis of the butanol extract of the culture broth indicated the peaks for akaeolide and lorneic acid A and unknown peaks showing UV absorption band around 330 nm attributable to the polyene structure (Figures 4-2 and 4-3). UV-guided purification from the extract using normal-phase and reverse-phase chromatography techniques led to the isolation of akaemycin (**5**, Scheme 4-1).

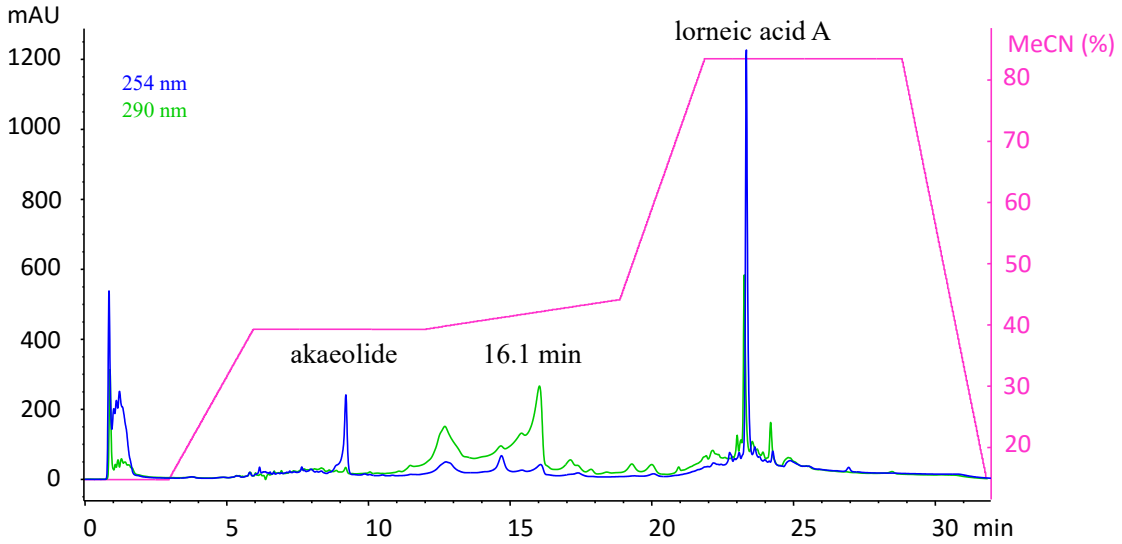
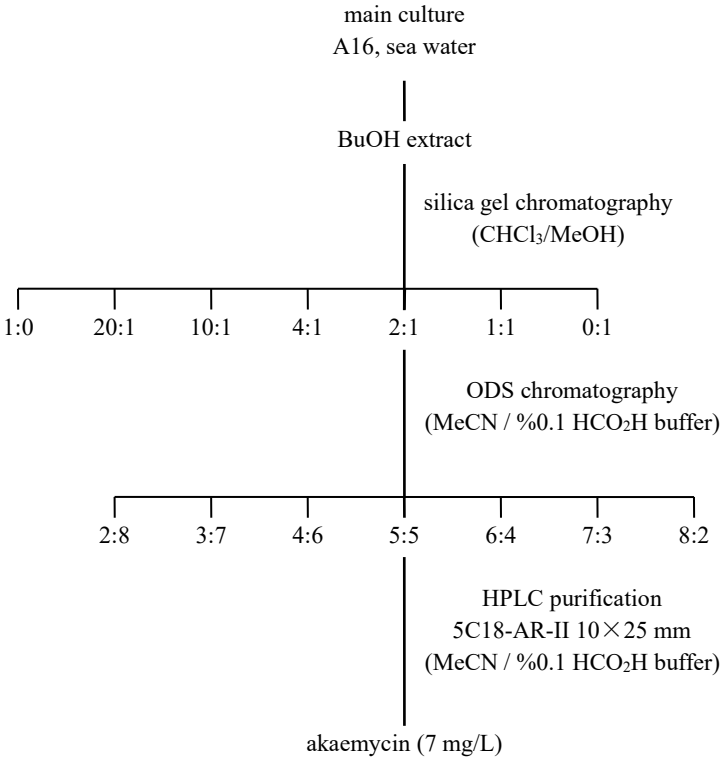


Figure 4-2. HPLC analysis of crude extract of strain NPS554.



Scheme 4-1. Isolation of akaemycin.

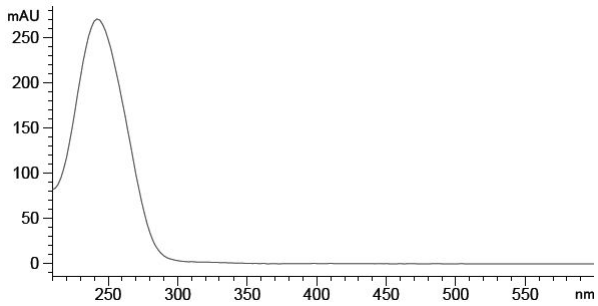


Figure 4-3. UV spectrum of the peak at 16.1 min.

4-2-2 Structure Determination of Akaemycin

4-2-2-1 Planar Structure

Akaemycin (**4**) was obtained as a light brown amorphous solid ($[\alpha]_D^{23}$ -114, *c* 0.10, MeOH). The molecular formula of **4** was established as $C_{36}H_{56}O_{10}$ on the basis of HR-ESITOFMS that showed a pseudomolecular ion peak at *m/z* 647.3795 $[M-H]^-$ (647.3795 calculated for $C_{36}H_{55}O_{10}$). Nine degrees of unsaturation were suggested by the molecular formula. The UV spectrum of **4** in methanol exhibited absorption maxima at 237 and 330 nm. The IR spectrum displayed the absorption bands at 3374, 2931, 1689, 1596, and 1251 cm^{-1}, indicating the presence of hydroxy and carbonyl groups and a conjugated C-C double bond system. The ^1H and ^{13}C NMR spectra of **4** contained one carbonyl, ten olefinic, nine oxymethine, one acetal, and fifteen aliphatic carbons (Table 4-1). These spectral features suggested that **4** was a typical type I PKS-derived polyketide. Ten olefinic and one carboxylic carbons accounted for six within the total nine degrees of unsaturation, implying that three-ring structures were present in **4**.

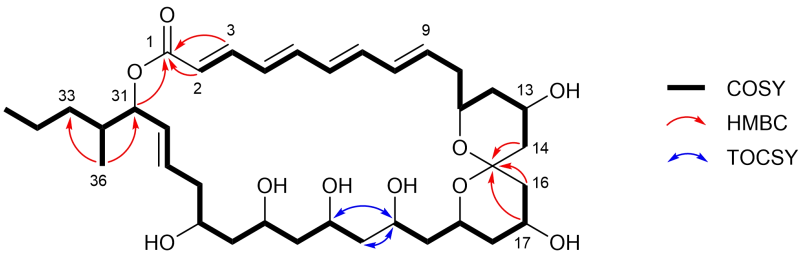


Figure 4-2. ^1H-^1H COSY, TOCSY, and selected HMBC correlations of **4**.

Planar structure of **5** was determined by analyzing 1D- and 2D-NMR spectral data (Figure 4-2). ^1H-^1H COSY and TOCSY correlations showed a long spin system including a conjugated tetraene from C-2 to C-9 and two oxymethine protons at C-11 (δ_C 71.7) and C-13 (δ_C 64.9). Another long spin system from C-16 to C-35 was also established to include six oxymethines at C-17 (δ_C 66.1), C-19 (δ_C 62.7), C-21 (δ_C 66.1), C-23 (δ_C 65.9), C-25 (δ_C 69.8), and C-27 (δ_C 72.8), one methyl group at C-36 (δ_C 15.2), and one double bond connection between C-29 (δ_C 130.5) and C-30 (δ_C 128.3). HMBC correlations from H-14 (δ_H 1.32, 1.99), H-16 (δ_H 1.74, 1.86), and H-17 (δ_H 4.05) to an sp3 carbon C-15 (δ_C 100.8) connected the above-mentioned two spin systems, C-2 to C-14 and C-16 to C-35, through the acetal carbon. Additionally, HMBC correlations from H-2, H-3, and

H-31 (δ_H 5.28) to C-1 (δ_C 168.0) established the carbon backbone from C-1 to C-35 as well as a macrocyclic structure being closed by an ester linkage between C-1 and C-31.

Table 4-1. ^{1}H and ^{13}C-NMR and key HMBC data of compound **4** in methanol-d_4

Position	δ_H, mult (*J*, Hz)[a]	δ_C[b]	HMBC
1		168.0	
2	5.85, d (15.2)	121.1	1, 4
3	7.43, dd (11.3, 15.2)	147.1	1, 2, 4, 5
4	6.44, dd (11.3, 14.7)	130.0	3, 5, 6
5	6.69, dd (10.8, 14.7)	143.9	3, 6, 7
6	6.35, dd (10.8, 14.7)	131.4	7
7	6.48, dd (10.6, 14.7)	139.5	5, 6, 8, 9
8	6.31, dd (10.6, 14.9)	133.6	6, 7
9	6.03, dt (14.9, 7.4)	136.8	7, 10, 11
10	2.32, m[c]	39.9	8, 9, 11
11	3.69, m[c]	71.7	
12	1.23, 1.96, m[c]	41.6	10, 11, 13, 14
13	4.02, m[c]	64.9	15
14	1.32, 1.99, m[c]	45.4	12, 13, 15
15		100.8	
16	1.74, 1.86, m[c]	41.3	15, 17
17	4.05, m[c]	66.1	15, 18
18	1.41, 1.83, m[c]	37.5	17, 19, 20
19	3.89, m[c]	62.7	
20	1.78, m[c]	44.2	19, 21
21	3.95, m[c]	66.1	
22	1.47, m[c]	44.9	
23	4.15, m[c]	65.9	
24	1.53, m[c]	47.6	23, 25
25	4.12, m[c]	69.8	
26	1.57, 1.69, m[c]	43.1	25
27	3.98, m[c]	72.8	
28	2.33, 2.42	40.8	27, 29, 30
29	5.66, dd (4.1, 15.4)	130.5	30, 31
30	5.84, dd (6.7, 15.4)	128.3	1, 29, 31
31	5.28, t (4.2)	78.0	1,29, 30, 32, 36
32	1.90, m[c]	37.9	
33	1.18, 1.47, m[c]	36.1	31, 32, 34, 35
34	1.34, 1.48, m[c]	21.4	33, 35
35	0.95, t (6.9)	14.7	33, 34
36	0.91, d (6.9)	15.2	31, 32, 33

[a] Recorded at 500 MHz

[b] Recorded at 125 MHz

[c] Overlapping signals.

The remaining two rings could be attributed to the bicyclic acetal moiety. According to the NMR data and the molecular formula, **4** should have eight sp^3 oxymethines assignable to six secondary hydroxy groups and two ether-connected oxymethines. In order to distinguish the locations of free hydroxy groups and ether-bonded oxygens, ^{13}C chemical shifts of **4** measured in CD_3OD and CD_3OH were compared. The overlaid HSQC spectrum of **4** is shown in Figure 4-3 displaying chemical shift change for six oxymethine carbons (C-13, C-17, C-21, C-23, C-25, and C-27) caused by exchange of alcoholic proton with deuteron. This clearly indicated that the remaining two oxymethines C-11 and C-19 were ether-bonded. The configurations of the olefinic bonds were determined as all *E* based on their large $^3J_{HH}$ coupling constants near 15 Hz (Table 4-1). This conclusion was further confirmed by the ROESY correlations as depicted in Figure 4-4.

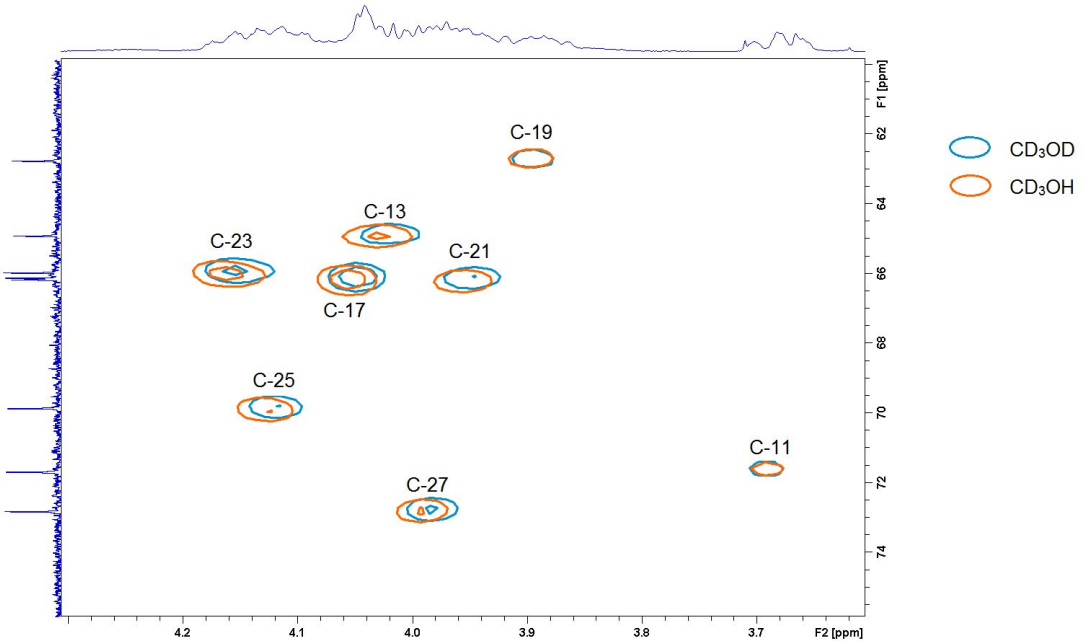


Figure 4-3. Partial HSQC spectrum for the oxymethine region measured in CD_3OD and CD_3OH.

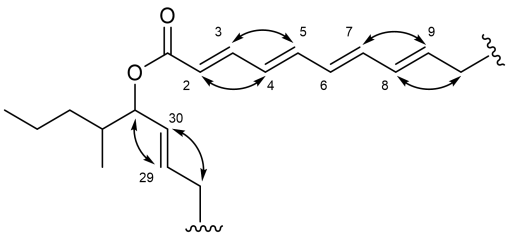


Figure 4-4. Key ROESY correlations of **4**.

4-2-2-2 Relative Configuration

The relative configuration of the spiro-*bis*-tetrahydropyran was analyzed by inspecting the ROESY spectrum (Figure 4-5). The correlation among H-11, H-13, and H-12b suggested the 1,3-diaxial relationship of H-11 and H-13. A strong correlation for H-11/H-19, a weak correlation for H-11/H-20, and no correlation for H-11/H-16 or H-11/H-17 indicated that the acetal oxygen atom in ring B was oriented axially to ring A and H-19 was also axial-oriented in ring B. ROESY

correlation was observed for H-19/H-18b but not for H-19/H-18a, indicating the axial orientation of H-18a in the opposite side to H-19. Correlations for H-17/H-18a and H-17/H-18b suggested the equatorial orientation of H-17. Thus, the relative configuration of the bicyclic acetal moiety was determined as 11*S**, 13*R**, 15*R**, 17*S**, and 19*S**.

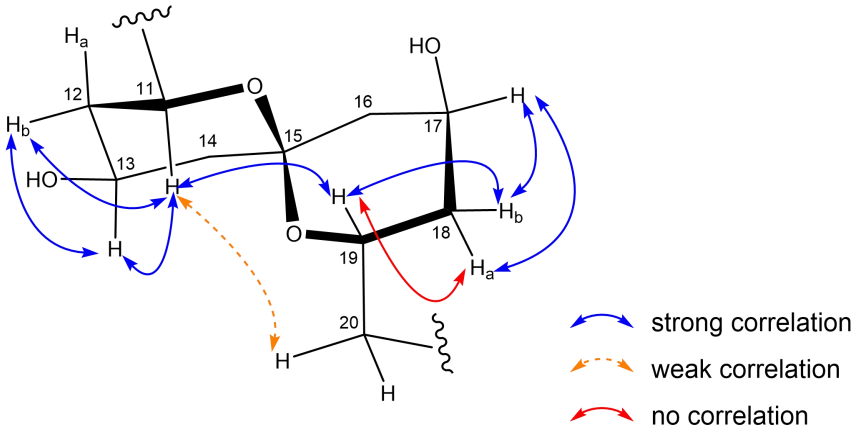


Figure 4-5. Key ROESY correlations of spiro-*bis*-tetrahydropyran moiety of **4**.

4-2-2-3 Stereochemistry of Akaemycin Predicted from Genomic Information

PKS domain organization for akaemycin biosynthesis and plausible biosynthetic pathway are illustrated in Figure 4-6. The PKS gene cluster contains four genes with 17 modules. Signature amino-acid residues of AT domains suggested that loading module (LM) and module 1 (m1) take methylmalonyl CoA as a substrate and the remaining fifteen modules (m2 to m16) incorporate malonyl CoA. Therefore, this PKS cluster should assemble a linear polyketide comprising two C_3 and fifteen C_2 units (Table S1, Chapter 3). All the domains required for polyketide decoration (i.e. KR (ketoreductase), DH (dehydratase), and ER (enoylreductase) domains) are present although a KR domain in module 10 (m10) is likely non-functional. The C-15 keto group would be intact because this functionality is necessary for spiroacetal formation. In the polyketide biosynthesis by type I PKS, absolute configurations of secondary hydroxy groups derived from acyl keto groups and methyl branches derived from methylmalonates are predictable on the basis of gene sequences for KR and ER domains, respectively (Figure 3-12 in Chapter 3).[1,2] The A1-type KR domains in modules 6, 8, 9, 10, and 11 were indicative of the "*R*" configurations for C-13, C-15, C-17, C-19, and C-23 chiral centers (Figure 4-7), while the B1-type KR domains in modules 2 and 12 suggested "*S*" configuration for C-11 and C-31 (Quotation marks on *R* and *S* represents the conventional *RS* nomenclature system as described in Chapter 3). The methyl group at C-32 was assigned "*R*" on the basis of amino acid alignment in the ER domain of module 1. The predicted configurations for the spiro-ketal ring system were consistent with the ROESY analysis (Figure 4-5). The unknown type of KR domains in modules 4, 5, and 7 left the configurations at C-21, C-25, and C-27 undetermined, giving the bioinformatics-based proposed absolute configuration of **4** as 11*S*, 13*R*, 15*R*, 17*S*, 19*S*, 23*S*, 31*R*, and 32*R*.

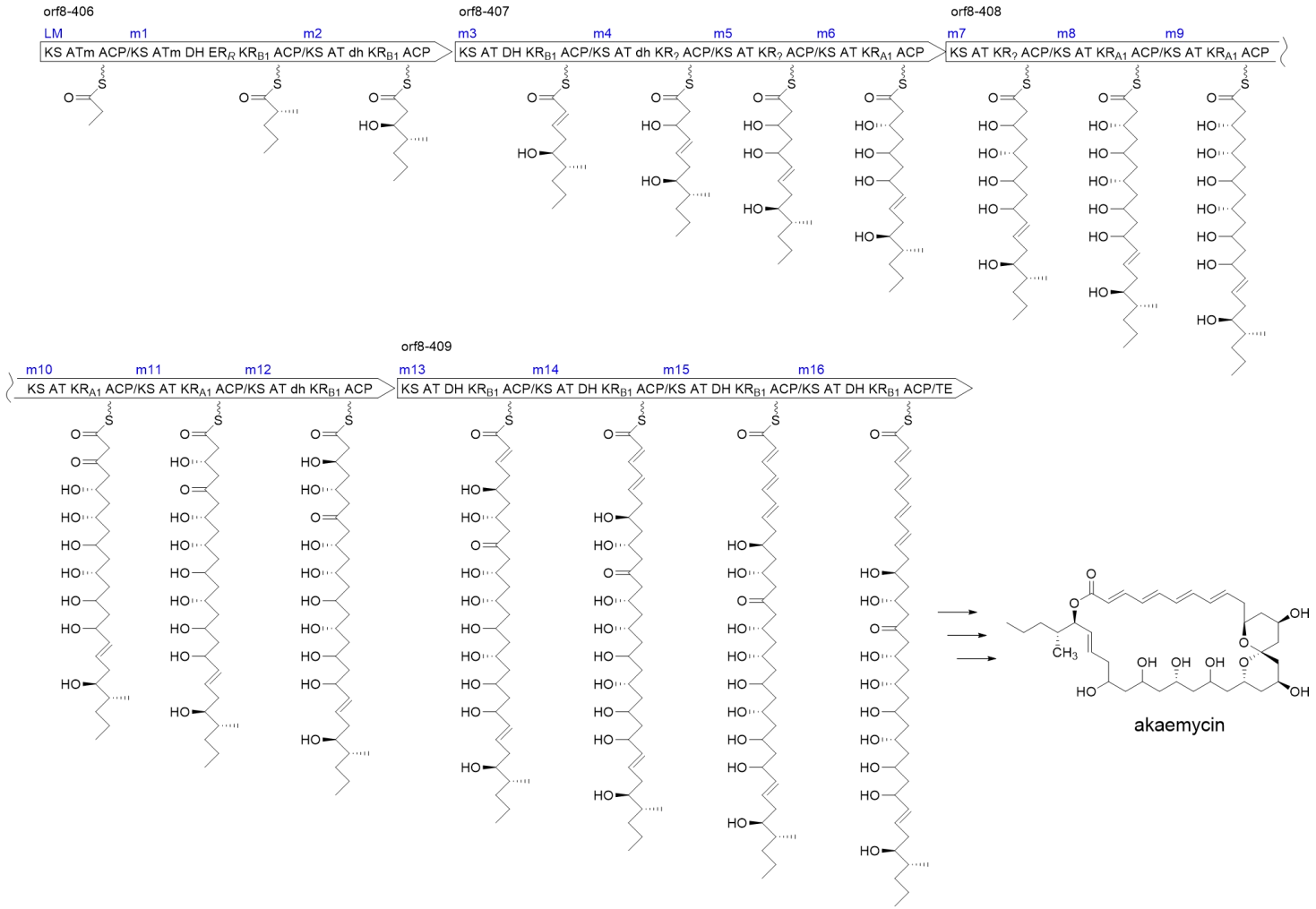


Figure 4-6. PKS domain organization and proposed biosynthetic pathway for **4**.

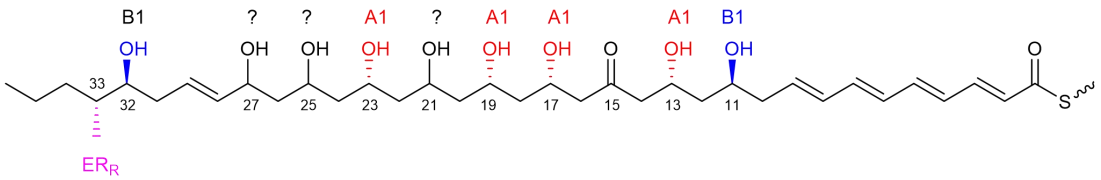


Figure 4-7. Linear structure of **4** with predicted stereochemistry.

Among the known natural products, the most structurally similar compound to akaemycin is marinisporolides (Figure 4-8).[3] Akaemycin and marinisporolide share the polyolefinic and polyhydroxy moieties and a bicyclic spiro-ketal unit as common structural units. Marinisporolides are the secondary metabolites of *Marinispora* strain that was also collected from marine environment. Since no similar compounds have been obtained from terrestrial microorganisms, these compounds are likely marine-specific. No similar biosynthetic gene clusters have been also registered in gene databases to date.

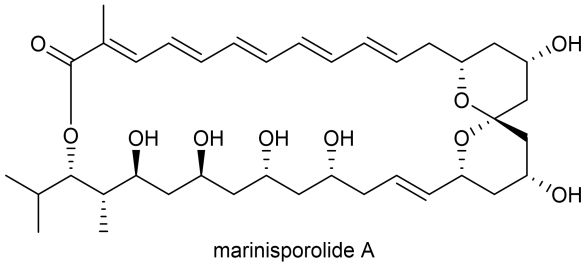


Figure 4-8. Structure of marinisporolide A.

4-2-3 Bioactivity

Limited biological testing revealed that **4** had antimicrobial activities against *Kocuria rhizophila* (previously termed as *Micrococcus luteus*) with an MIC value of 9.6 μg/mL but is inactive to *Escherichia coli*, *Candida albicans*, and *Kluyveromyces fragilis*. Paper disk test (20 μg / 6 mm diameter) on solid agar has shown weak antifungal activity against *Penicillium chrysogenum* (7 mm inhibition zone) and *Rhizopus oryzae* (8 mm inhibition zone). Cytotoxicity was not detected at 10 μM against melanoma A375 cells.

4-3 Experimental Section

General experimental procedures. Optical rotation was measured using a JASCO DIP-3000 polarimeter. UV spectrum was recorded on a Hitachi U-3210 spectrophotometer. IR spectrum was measured on a Perkin Elmer Spectrum 100. NMR spectra were obtained on a Bruker AVANCE 500 spectrometer, using the signals of the residual solvent protons (δ 3.31) and carbons (δ 44.9) as internal standards. High-resolution ESITOFMS was recorded on a Bruker microTOF focus. Silica gel 60-C18 (Nacalai Tesque 250-350 mesh) was used for ODS column chromatography. HPLC separation was performed using a Cosmosil 5C18-AR-II (Nacalai Tesque Inc., 20 × 250 mm) with a photodiode array detector.

Fermentation. Strain NPS554 cultured on a Bn-2 slant [soluble starch 0.5%, glucose 0.5%, meat extract (Kyokuto Pharmaceutical Industrial Co., Ltd.) 0.1%, yeast extract (Difco Laboratories) 0.1%, NZ-case (Wako Chemicals USA, Inc.) 0.2%, NaCl 0.2%, $CaCO_3$ 0.1%, agar 1.5%] was inoculated into 500 mL K-1 flasks each containing 100 mL of the V-22 seed medium consisting of soluble starch 1%, glucose 0.5%, NZ-case 0.3%, yeast extract 0.2%, Tryptone (Difco Laboratories) 0.5%, K_2HPO_4 0.1%, $MgSO_4 \cdot 7H_2O$ 0.05%, and $CaCO_3$ 0.3% (pH 7.0) in natural seawater. The flasks were placed on a rotary shaker (200 rpm) at 30 ˚C for 4 days. The seed culture (3 mL) was transferred into 500 mL K-1 flasks each containing 100 mL of the A-16 production medium consisting of glucose 2%, Pharmamedia (Traders Protein) 1%, and $CaCO_3$ 0.5% in natural seawater. The inoculated flasks were placed on a rotary shaker (200 rpm) at 30 ˚C for 6 days.

Extraction and isolation. At the end of the fermentation period, 100 mL of 1-butanol was added to each flask, and they were allowed to shake for 1 h. The mixture was centrifuged at 5000 rpm for 10 min, and the organic layer was separated from the aqueous layer containing the mycelium. The crude extract was subjected to silica gel column chromatography with a step gradient of $CHCl_3$/MeOH (1:0, 20:1, 10:1, 4:1, 2:1, 1:1, and 0:1 v/v). Fraction 2 ($CHCl_3$/MeOH=20:1) was concentrated to provide a brown solid, which was further purified by repeated reverse phase preparative HPLC using a Cosmosil 5C18-AR-II column (Nacalai Tesque Inc., 10 × 250 mm) with MeCN in 0.1% HCO_2H (35:65, flow rate 4 mL/min), followed by evaporation and extraction with EtOAc to give akaemycin (7.1 mg/L).

Akaemycin (4). Yellow amorphous solid; $[\alpha]_D^{23}$ -114 (*c* 0.10, MeOH); UV (MeOH) λ_{max} (log ε) 330.8 (4.44), 237.2 (3.80) nm; ^1H and ^{13}C NMR data, see Table 10; HR-ESITOFMS $[M-H]^-$ 647.3795 (calcd for $C_{36}H_{55}O_{10}$, 647.3795), $[M+Na]^+$ 671.3774 (calcd for $C_{36}H_{56}O_{10}Na$, 671.3771).

Biological assays. Antimicrobial assay tested in broth medium was carried out according to the procedures previously described.[4] In paper disk assays, akaemycin (**4**) was diluted to 2 mg/mL with MeOH, and 10 L aliquots were loaded on paper disks with 6 mm diameter, which were left until completely dried. A loop of the test organism, suspended in a small amount of sterilized water, was mixed with liquefied agar medium precooled to nearly body temperature, and the inoculated medium was quickly poured into a sterile plastic dish. The medium was composed of 0.5% yeast extract, 1.0% Tryptone, 1.0% NaCl, 0.5% glucose, and 1.5% agar. After the agar was solidified, the drug-loaded disks were placed on the plate, and the test cultures were incubated at 32 ˚C for a day or two until the diameters of inhibitory haloes turned measurable.

References

1 Keatinge-Clay AT. *Chem. Biol*. **2007**, 14, 898-908.

2 Kwan DH, Sun Y, Schulz F, *et al. Chem. Biol*. **2008**, 15, 1231-1240.

3 Kwon HC, Kauffman CA, Jensen PR, *et al. J. Org. Chem*. **2009**, 74, 675-684.

4 Igarashi Y, Yu L, Miyanaga S, *et al. J. Nat. Prod*. **2010**, 73, 1943-1946.

4-4 Spectral Data

Table of Contents

Figure S1. ^{1}H NMR spectrum of akaemycin (**4**) at 500 MHz in methanol-d_4

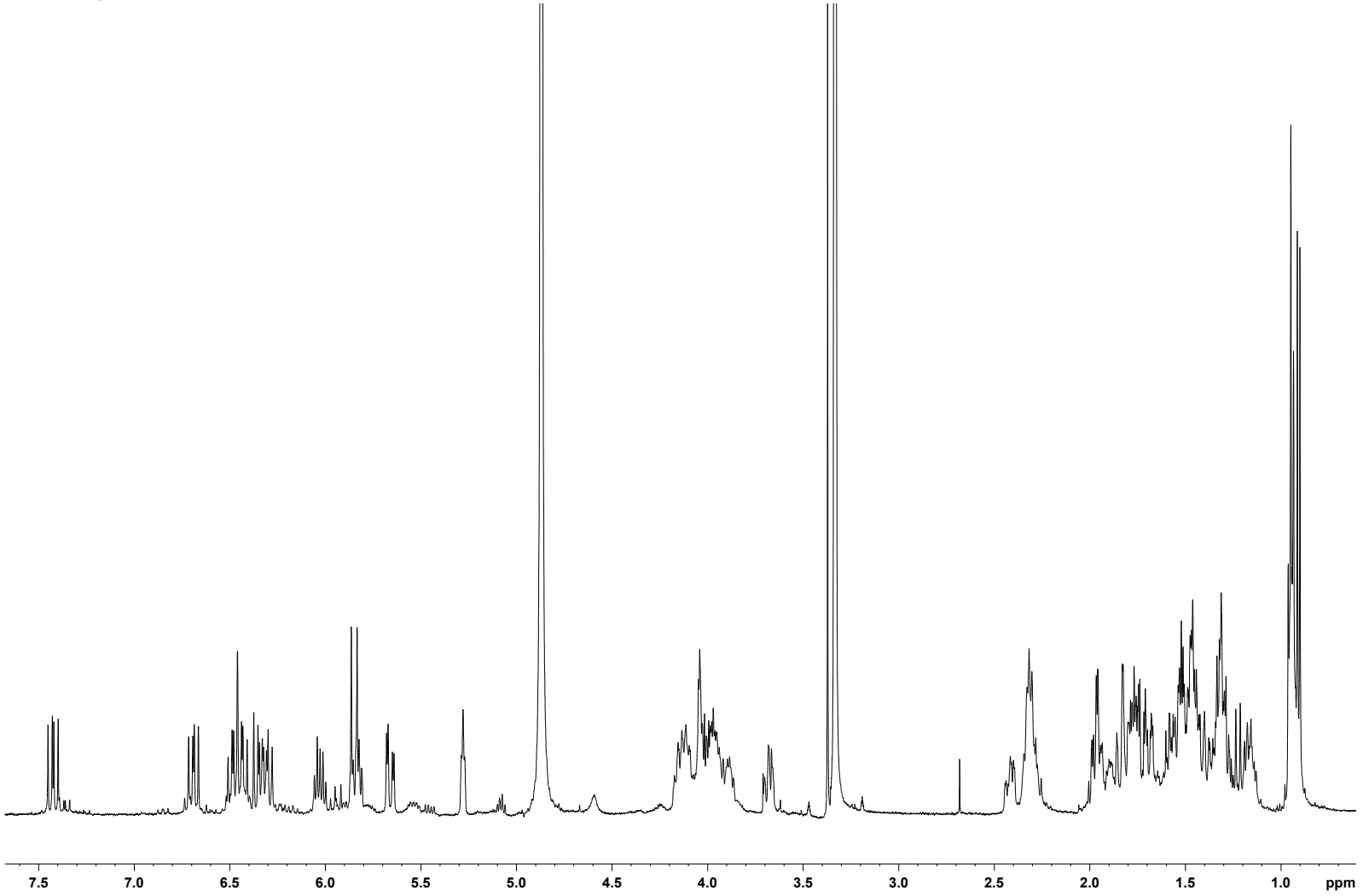


Figure S2. ^{13}C NMR spectrum of **4** at 125 MHz in methanol-d_4

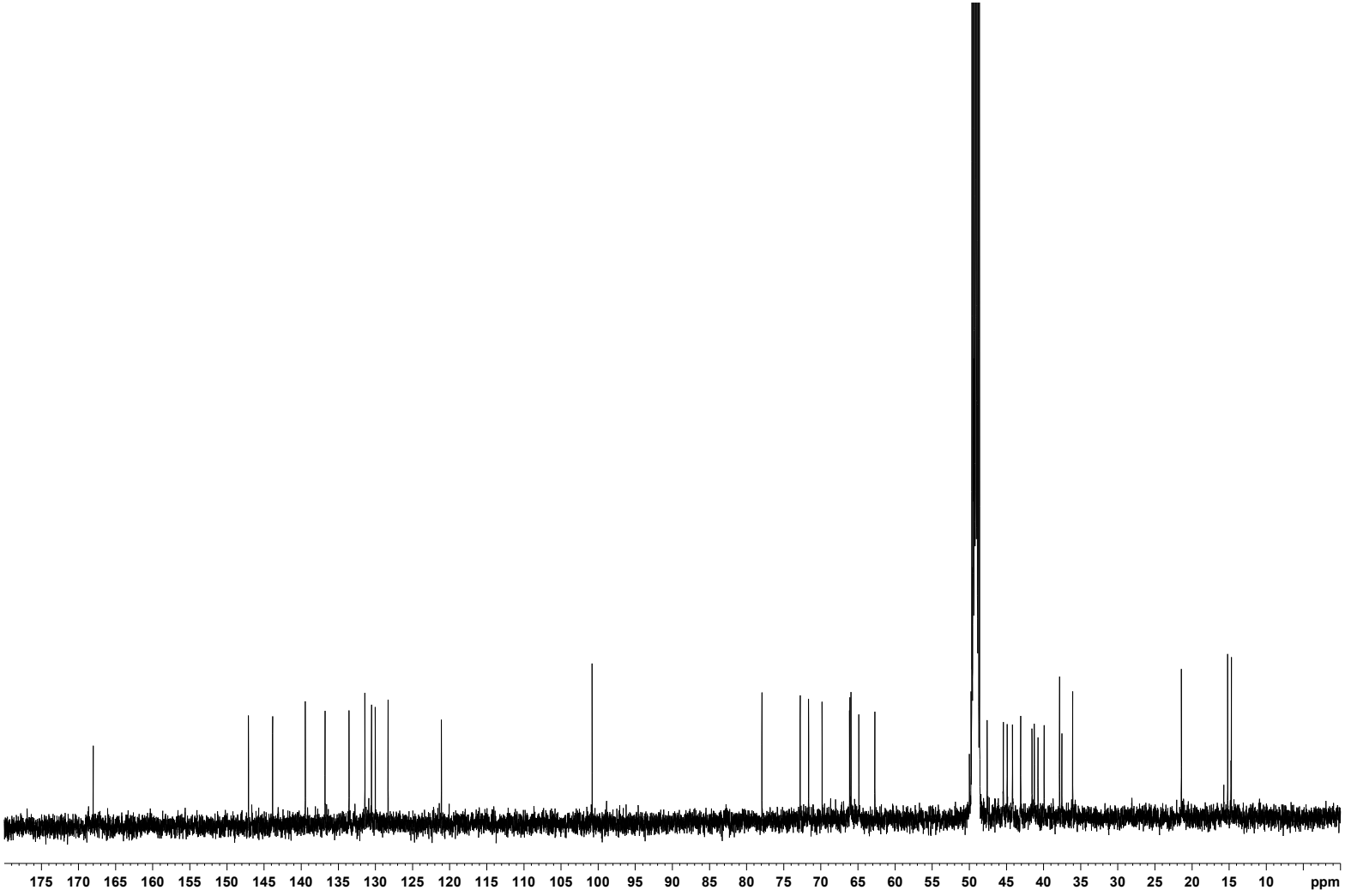


Figure S3. ^{1}H-^{1}H COSY spectrum of **4** at 500 MHz in methanol-d_4

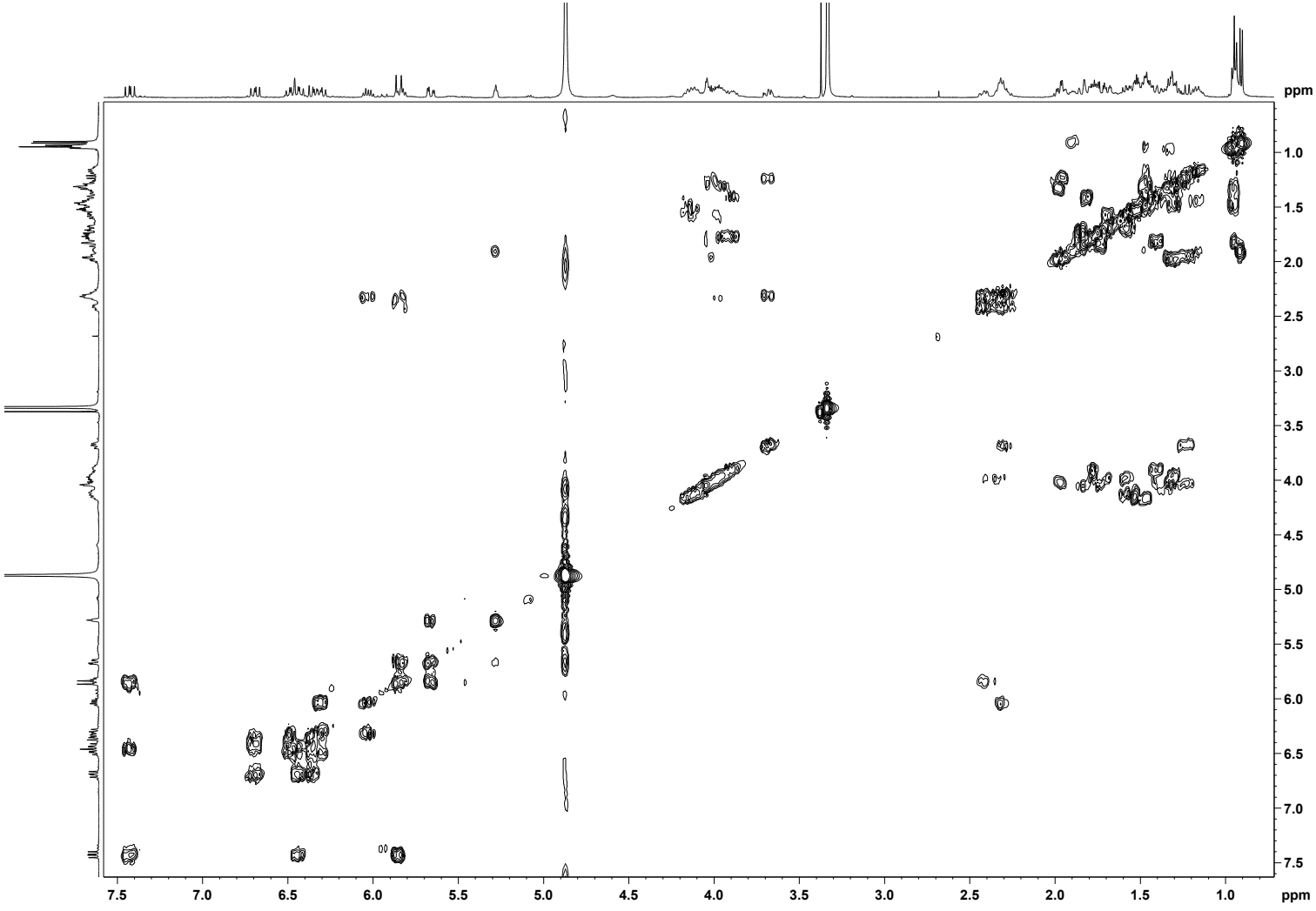


Figure S4. HSQC spectrum of **4** at 500 MHz in methanol-d_4

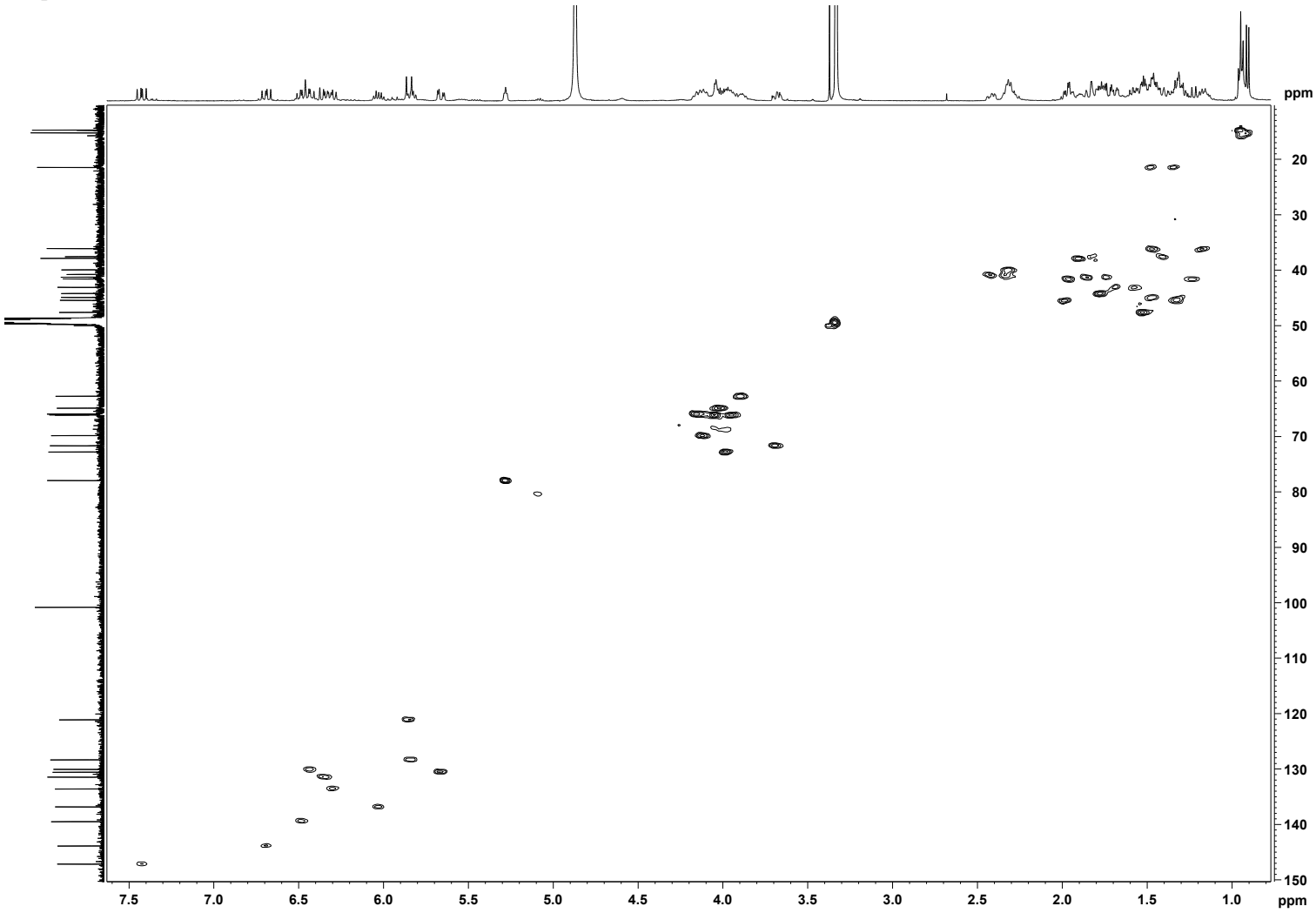


Figure S5. HMBC spectrum of **4** at 500 MHz in methanol-d_4

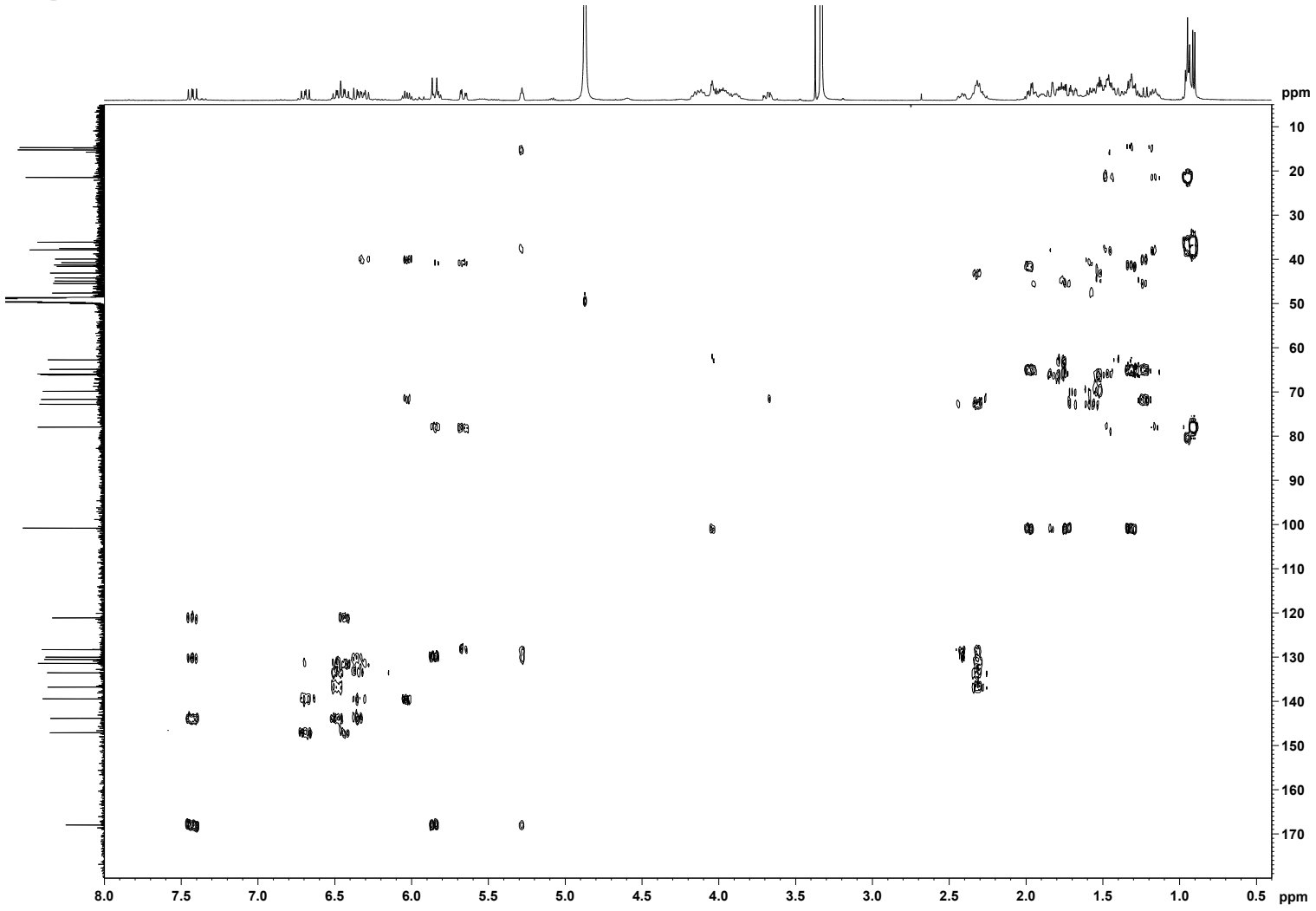


Figure S6. ROESY spectrum of **4** at 500 MHz in methanol-d_4

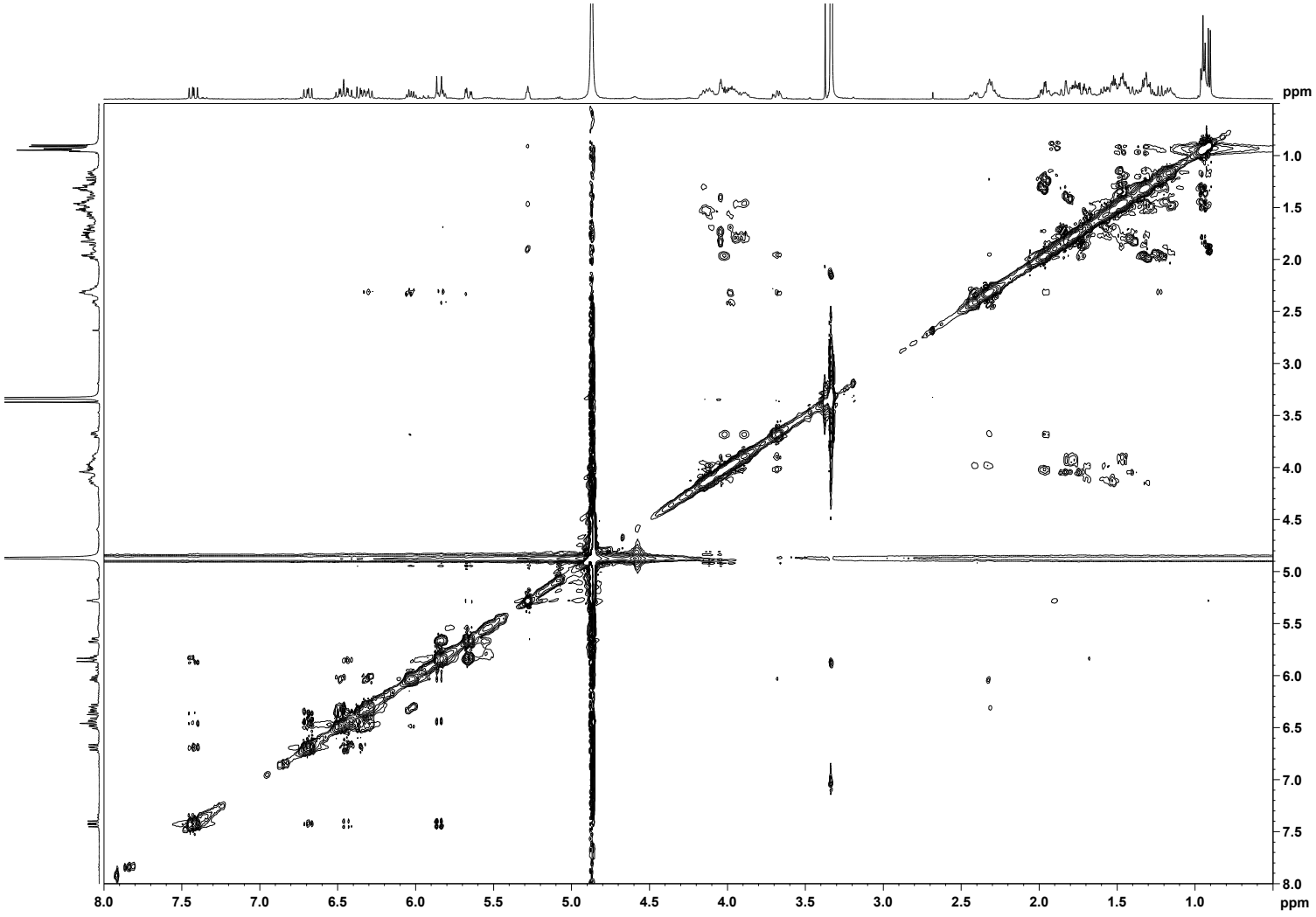


Figure S7. ^{1}H-^{1}H TCOSY spectrum of **4** at 500 MHz in methanol-d_4

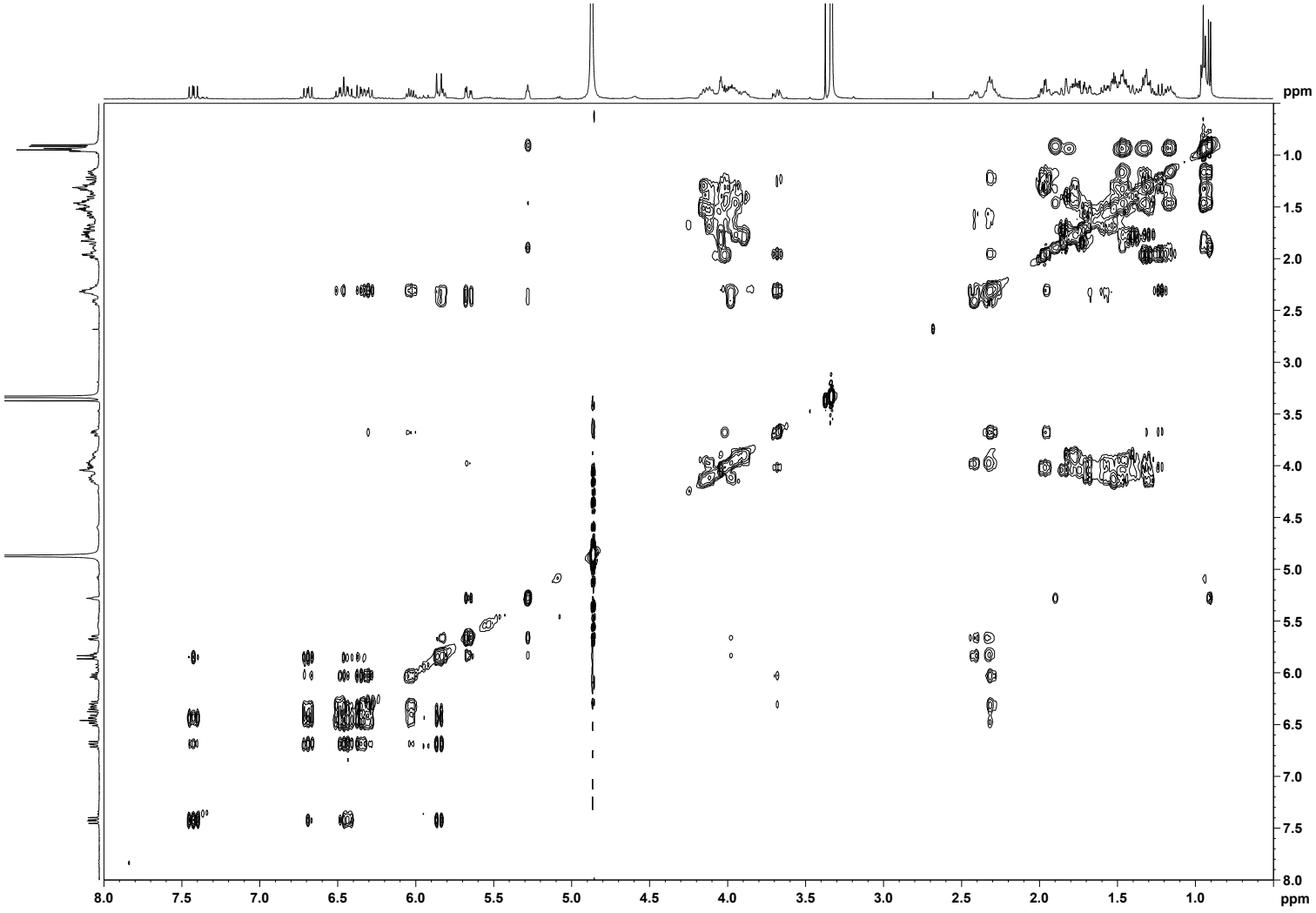


Figure S8. UV spectrum of **4** in methanol

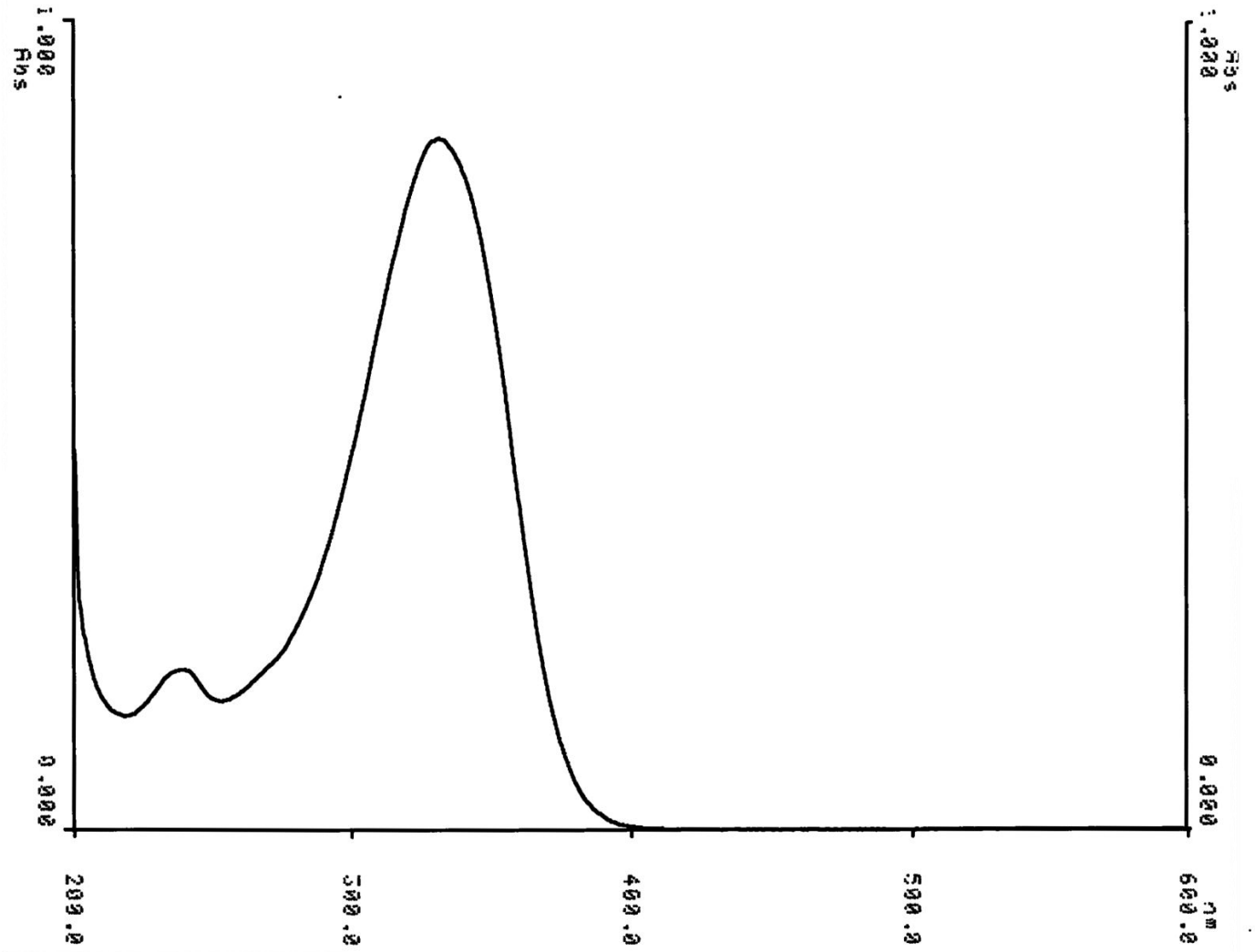
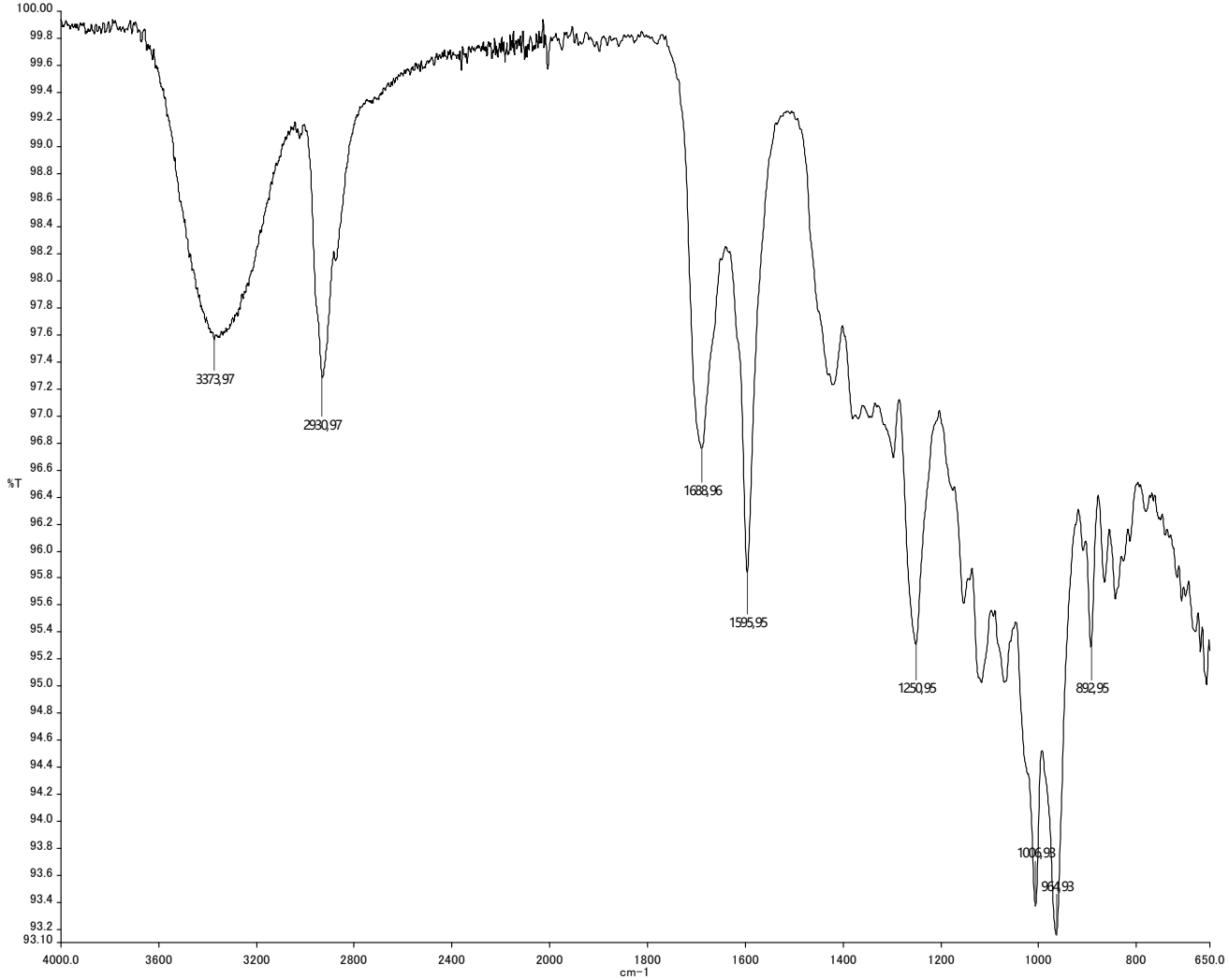


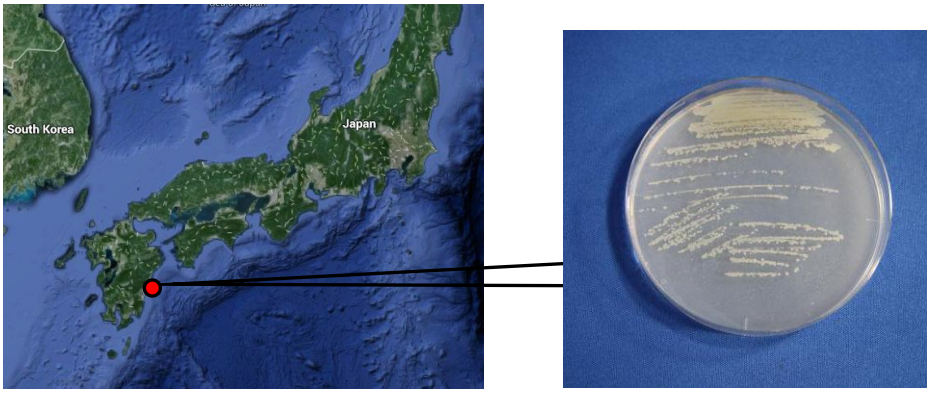
Figure S9. IR spectrum of **4** (ATR)



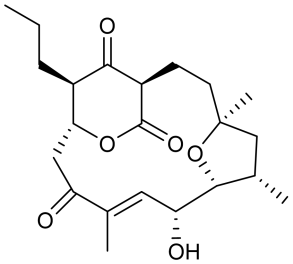
CHAPTER 5

CONCLUSION

In this study, as detailed discussed in Chapters 2 to 4, from marine-derived *Streptomyces* sp. NPS554 isolated from Miyazaki Harbor, by UV-spectroscopic screening akaeolide was discovered, and by genome mining based on the genomic information, akaemycin was discovered.

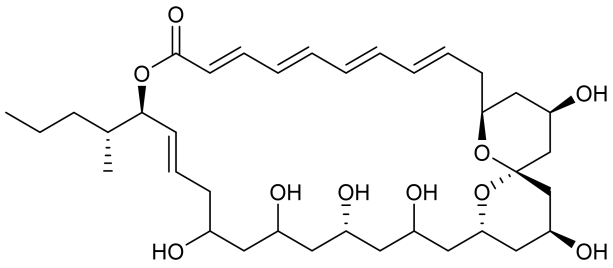


In the screening program for novel compounds from 27 marine bacteria with low gene similarity, *Streptomyces* sp. NPS554 was selected as a candidate strain. The extract of culture broth was analyzed by HPLC-DAD (diode-array detector) system, which led to the isolation of akaeolide, a novel carbocyclic polyketide compound. The planar structure of akaeolide was elucidated by spectroscopic analysis, and the absolute configuration was determined by X-ray crystallography of 17-chloroakaeolide and NOESY correlation of akaeolide. Akaeolide shows activity against *Micrococcus luteus* with an MIC value of 25 μg/mL and modest cytotoxicity to 3Y1 rat fibroblasts with an IC_{50} value of 8.5 ± 1.5 μM.



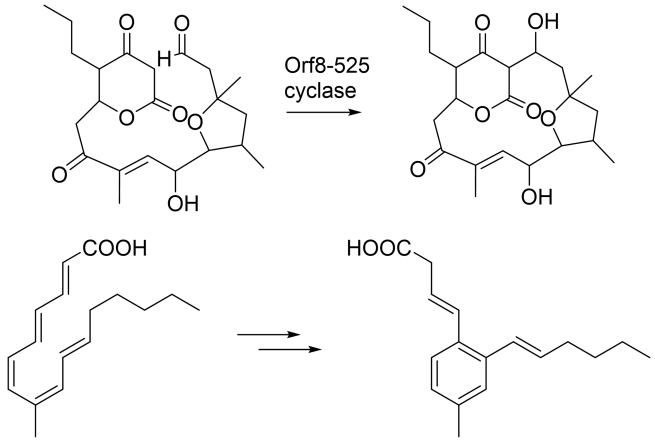
akaeolide (**1**)

The genome mining of this strain led to the identification of 12 PKS gene clusters, and the products were structurally predicted. Besides akaeolide and lorneic acids, there are still several other candidate novel compounds. The UV-spectroscopic screening for the predicted structure led to the isolation of akaemycin, a new macrolide compound. The planar structure of akaemycin was determined by spectroscopic analysis, which is consistent with the predicted structure. The NOESY correlations show partial relative configurations, and genomic information proposed the absolute stereochemistry of akaemycin. Akaemycin shows antimicrobial activities against *Kocuria rhizophila* (previously termed as *Micrococcus luteus*) with an MIC value of 9.6 μg/mL, and weak antifungal activity against *Penicillium chrysogenum* (7 mm inhibition zone) and *Rhizopus oryzae* (8 mm inhibition zone).



akaemycin (**5**)

Both these compounds have novel chemical structures with no similar analogue compounds reported, which should be difficult to find only by biological screening. According to the ^{13}C-labeled precursor feeding culture and in silico analysis of biosynthetic gene clusters, the biosynthetic origin of akaeolide and lorneic acid A was discovered and the biosynthesis pathway was proposed. These pathways include unreported C-C bond formation reactions.



The biosynthesis of akaeolide and lorneic acids was further elucidated by the following ^{13}C isotope precursor feeding experiment. The results show that they are typical type I polyketides biosynthesized from acetic and propionic units. With this information, from the sequenced gene of *Streptomyces* sp. NPS554, the gene clusters of akaeolide and lorneic acids were assigned. The biosynthesis of akaeolide and lorneic acids is proposed from genomic information. In akaeolide, the carbocyclic backbone is enclosed by an aldol condensation catalyzed by predicted cyclase with domain homology to SnoaL, and in lorneic acids, the formation of benzene ring is synthesized from a polyene intermediate. Such mechanism has not been reported before this study.

In summary, from marine-derived *Streptomyces* sp. NPS554, by using UV-spectroscopic screening akaeolide was discovered, and by genome mining based on the genomic information from gene clusters for secondary metabolite, akaemycin was discovered. These two new compounds are difficult to be found only by bioactive screening. They are both characterized to have high structural novelty and few similar analogues. According to the isotope labeled experiments and in silico biosynthetic gene analysis, the biosynthetic origin of akaeolide and lorneic acid A was determined, and their plausible biosynthetic pathways were proposed. They both include previously unreported carbon-carbon connections. The three polyketides produced by strain NPS554, akaeolide, akaemycin, and lorneic acid A are all structurally unique compounds,

and none of the biosynthesis gene clusters are reported or found in the DNA database before this study. These results strongly confirmed the hypothesis mentioned in Chapter 1 that, taxonomically new microbial species are the suitable resource for exploring new natural products.

Finally, according to the results in this research, it can be concluded that, from previously unexplored new microbial species, combining multiple methods such as chemical screening and genome mining, new compounds with novel carbon backbones can be discovered.

Acknowledgements

I wish to express my sincere gratitude to Professor Yasuhiro Igarashi, Toyama Prefectural University, for his valuable guidance, encouragement and generous support throughout the course of this work.

I also wish to express my sincere thanks to Assistant Professor Naoya Oku, Toyama Prefectural University, for his valuable advice and warm encouragement throughout this work.

Acknowledgements are also made to co-suthors: Dr. S. Sato (Nippon Suisan Kaisha, Ltd.), Dr. T. Matsumoto (Rigaku Corporation), Dr. H. Komaki (NBRC), Dr. N. Ichikawa (NBRC), Dr. L. Yu (Zhejiang University of Science & Technology).

Special thanks are due to Dr. T. Okuda (Tamagawa University), Ms. Y. Sudoh (Tamagawa University), Dr. M. Imoto at (Keio University) for their much helpful collaboration.

I greatly appreciate all the members of the laboratory for Microbial Engineering, Biotechnology Center, Toyama Prefectural University.

Finally, I am deeply grateful to my parents and my dear wife for their encouragement and support.

Publications

1. Akaeolide, a carbocyclic polyketide from marine-derived *Streptomyces*
Yasuhiro Igarashi, Tao Zhou, Seizo Sato, Takashi Matsumoto, Linkai Yu, and Naoya Oku
Organic Letters 15: 5678-5681 (2013)

2. Biosynthesis of akaeolide and lorneic acids and annotation of type I polyketide synthase gene clusters in the genome of *Streptomyces* sp. NPS554
Tao Zhou, Hisayuki Komaki, Natsuko Ichikawa, Akira Hosoyama, Seizo Sato, and Yasuhiro Igarashi
Marine Drugs 13: 581-596 (2015)

3. Bioinformatics-inspired isolation of akaemycin, a new macrolide from marine *Streptomyces* sp
Tao Zhou, Seizo Sato, Hisayuki Komaki, and Yasuhiro Igarashi
Current Biotechnology 4: 249-253 (2016)