Biosynthetic Analysis of Marine Myxobacterial Secondary Metabolites

（海洋性粘液細菌二次代謝産物の生合成解析）

SUN, Yuwei

Biosynthetic Analysis of Marine Myxobacterial Secondary Metabolites

SUN, Yuwei

Laboratory of Bioactive Natural Products Chemistry

Department of Applied Molecular Biosciences

Graduate School of Bioagricultural Sciences

Nagoya University

January 2016

Contents

Abbreviations

aa, amino acids

ACP, acyl carrier protein

A domain, adenylation domain

Amp, ampicillin

AT, acyl transferase

ATP, adenosine triphosphate

BLAST, basic local alignment search tool

CDD, Conserved Domain Database

C domain, condensation domain

Cm, chloramphenicol

CoA, coenzyme A

COSY, correlation spectroscopy

DH domain, dehydratase domain

DMSO, dimethyl sulfoxide

DNA, deoxyribonucleic acid

ECH, enoyl-CoA hydratase

ER, enoylreductase

ESI, electro-spray ionisation

EtOAc, ethyl acetate

EtOH, ethanol

FAD, flavin adenine dinucleotide

FT-IR, Fourier transform infrared

HC domain, heterocyclization domain

HCS, hydroxymethylglutaryl-CoA synthase

HMBC, heteronuclear multiple bond correlation

HMQC, heteronuclear multiple quantum correlation

HPLC, high performance liquid chromatography

HR, high resolution

IPTG, isopropyl-β-D-thiogalactosid

kb, kilobases

kDa, kilodalton

Kan, kanamycin

KR, ketoreductase

KS, ketosynthase

LB, Luria Bertani

LC, liquid chromatography

MeOH, methanol

Me, methyl

MO, monooxygenase

MS, mass spectrometry

MW, molecular weight

NMR, nuclear magnetic resonance

NRPS, non-ribosomal peptide synthase

Nrs, nourseothricin

OD, optical density

O-MT, *O*-methyltransferase

ORF, open reading frame

Ox domain, oxidation domain

Oxytet, oxytetracycline

PAGE, polyacrylamide gel electrophoresis

PCP, peptidyl carrier protein

PCR, polymerase chain reaction

PKS, polyketide synthase

rpm, rounds per minute

RP, reversed phase

SAM, *S*-adenosylmethionine

SDS, sodium dodecylsulfate

SNAC, *N*-acetylcysteamine

TE, thioesterase

Tet, tetracycline

TLC, thin layer chromatography

TCA cycle, tricarboxylic acid cycle

Tm, melting temperature

Tris, tris(hydroxymethyl)aminomethane

UV, ultraviolet

WT, wild type

Chapter 1. Introduction

1.1. Natural products in the discovery of antibiotics

Natural extracts of plants or animals have been used to treat various diseases for thousands of years dating back to the dawn of ancient civilizations. Since the discovery of morphine from opium, quinine from cinchona trees, caffeine from coffee beans (Figure 1-1), it has become clear that individual bioactive natural compounds play key roles in the therapeutic effects. Today, although the development of synthetic chemistry has generated a rich diversity of organic compounds, natural products continue to be important sources of leads for the development of new medicines.[1] Of all the small-molecule drugs approved between 1981 and 2010, 34% was natural products and their related structures (semisynthetic compounds or "botanical" mixtures), 16% was made by total synthesis but the pharmacophore was from a natural product.[2] Even in the first decade of 21st century, natural products field is still keeping high percentage in the drug discovery.

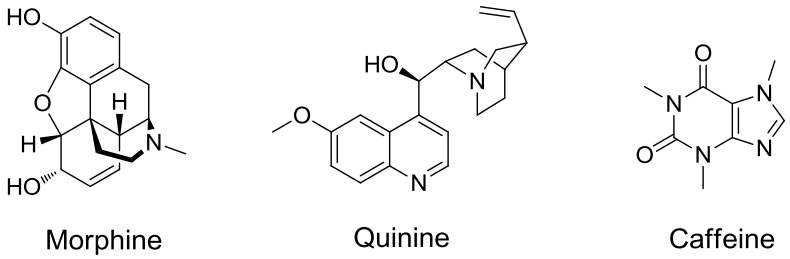


Figure 1-1. Chemical structures of some bioactive natural compounds.

The discovery of antibiotic penicillin from *Penicillium* and its clinical use revolutionized medicine in the 1940s. Looking back to the history of antibiotics, microorganism-derived secondary metabolites have served as a rich source for more than half a century. However, after the golden era from 1940s to 1960s when a large number of successful antibiotics, such as erythromycin, tetracycline and kanamycin (Figure 1-2), were isolated from actinomycetes, the pace of antibiotic discovery has slowed down and the spread of resistant bacteria is now becoming a major public health threat.[3] During the last several years, the

increasing resistant pathogen towards numerous existing antibiotics has generated an urgent need for novel bioactive lead compounds.[4] Since there are decreasing chances of finding something new from the secondary metabolome of actinomycetes due to the intense screening efforts over decades, alternative producers of antibiotics are attracting attention globally. As suggested by Donadio, *et al.*, appropriate approaches to novel antibiotic classes include exploiting microbial strains from underexplored environments, screening new microbial taxa and mining microbial genomes.[5] With the development of biotechnological techniques, those microbes that used to be thought as hard-to-culture or uncultivable have a promising future. One example is the myxobacteria.

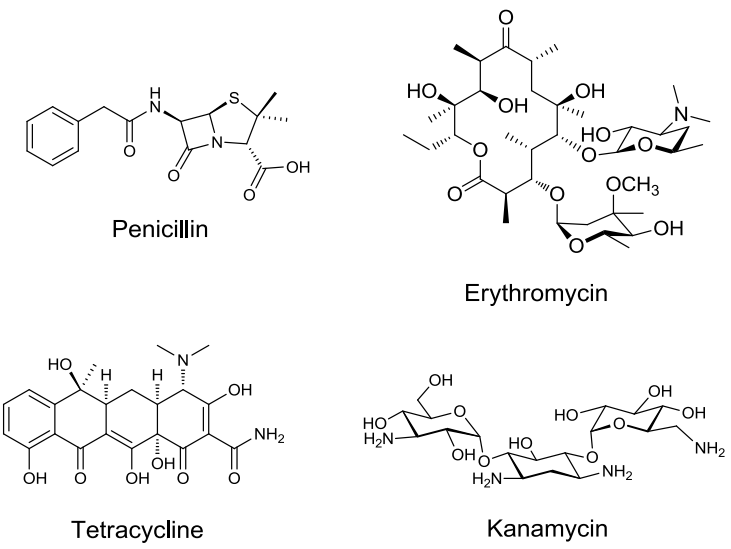


Figure 1-2. Antibiotics from actinomycetes.

1.2. Myxobacteria

Myxobacteria are rod-shaped Gram-negative bacteria belonging to the δ-group of proteobacteria.[6] They can move by gliding or creeping on solid surfaces,[7-9] and under starvation conditions, they aggregate and differentiate to form swarm and then multicellular fruiting bodies with bright color (Figure 1-3).[10, 11] This 'social' behavior, along with their ability to lyse a variety of other microorganisms using exoenzymes,[12] helps them to survive in the highly competitive environment. Currently, three suborders of *Myxococcales* (the taxonomic order of myxobacteria) have been classified as *Sorangiineae*, *Cytobacterineae* and

Nannocystineae.[13] These microorganisms ubiquitously inhabit in all climate zones, from tropical to the Antarctic region.[14] Although the genuine habitat of myxobacteria is the soil (pH between 5 and 8) with an optimum temperature of 30 °C,[14, 15] they are also found on the bark of trees, rotting wood, dung of herbivorous animals, and in aquatic habitats, as well as extreme environment including high salt concentration (> 2%),[16-20] acidic (pH 3.7) to alkaline (pH 8.8-9.2) soils[21, 22] and low temperature range of 4-9 °C.[23]

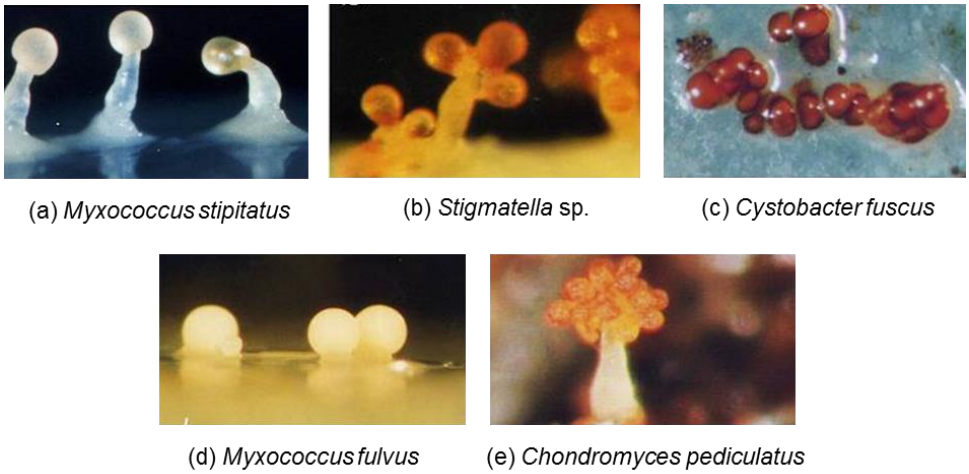


Figure 1-3. Myxobacterial fruiting bodies.[24]

Myxobacteria produce multi-class of secondary metabolites with novel scaffolds, as well as rare modes of action.[25, 26] Until 2010, a considerable number of natural products (at least 100 core structures and more than 500 derivatives) derived from myxobacteria have been discovered.[27] These small molecules are often observed to be potent antibacterial and antifungal substances, likely reflecting the competitive pressures from rival microorganisms. It is intriguing that myxobacterial secondary metabolites target sub-cellular structures which are rarely hit by other metabolites, for instance, most of them inhibit electron flow within mitochondrial respiratory chain.[25] Moreover, a much wider range of bioactivities are also reported, including anti-malarial, immunosuppressive and cytotoxic effects towards mammalian cells.[25, 27] Among all the myxobacterial products, the so called 'post-taxol' epothilones[28, 29] from *Sorangium cellulosum* are prominent examples as potent microtubule-stabilizing anticancer agents: the semi-synthetic epothilone

derivative, ixabepilone (Ixempra^{TM}) was approved by FDA in 2007 and natural epothilone B are currently in phase III clinical trials (Figure 1-4). Therefore, the myxobacteria have gained more and more attention as a promising alternative source of pharmaceutical and agricultural chemicals.

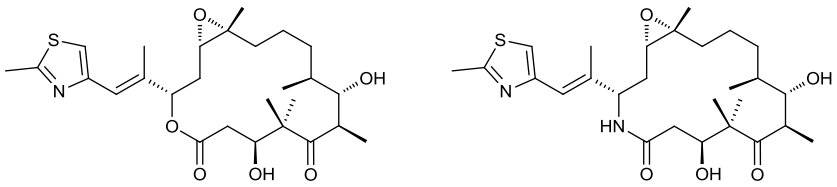


Figure 1-4. Chemical structures of epothilone B and ixabepilone.

1.3. Marine myxobacterial secondary metabolites

Myxobacteria have long been regarded as terrestrial microorganisms, while at the beginning of this century, several halophilic myxobacterial strains from marine environment came into the focus, including *Haliangium ochraceum*,[19] *Paraliomyxa miuraensis*,[30] *Enhygromyxa salina*[16, 18] and so on.[17, 31] Exploration for bioactive compounds from these rare microorganisms was soon initiated. A series of natural products, such as haliangicins,[32-34] miuraenamides,[30, 35] salimabromides,[36] salimyxins,[37] and enhygrolides[37] have hitherto been characterized (Figure 1-5).

Haliangicins were the first group of marine myxobacterial secondary metabolites isolated from a halophilic (optimal sodium chloride concentration of 2%) myxobacterium *Haliangium ochraceum* SMP-2 by Fudou *et al.* in 2001[32, 33] and by Kundim *et al.* in 2003[34]. The major compound among them, haliangicin (**1**) was reported to inhibit the growth of some fungi by interfering with the electron flow within the cytochrome *b-c1* segment in the mitochondrial respiratory chains.[32] With respect to the chemical structure, haliangicin possesses a long polyene chain bearing a unique terminal alkene with a neighboring an epoxide ring and a β-methoxyacrylate moiety, which is crucial for the antimicrobial activity.[33]

Recently, during the ongoing investigation on the secondary metabolome of *H. ochraceum* in our laboratory, a new PKS-NRPS hybrid compound, haliamide (**2**), was isolated by chance.[38] Haliamide (**2**)

also possesses a particular terminal alkene in end of the polyketide chain.

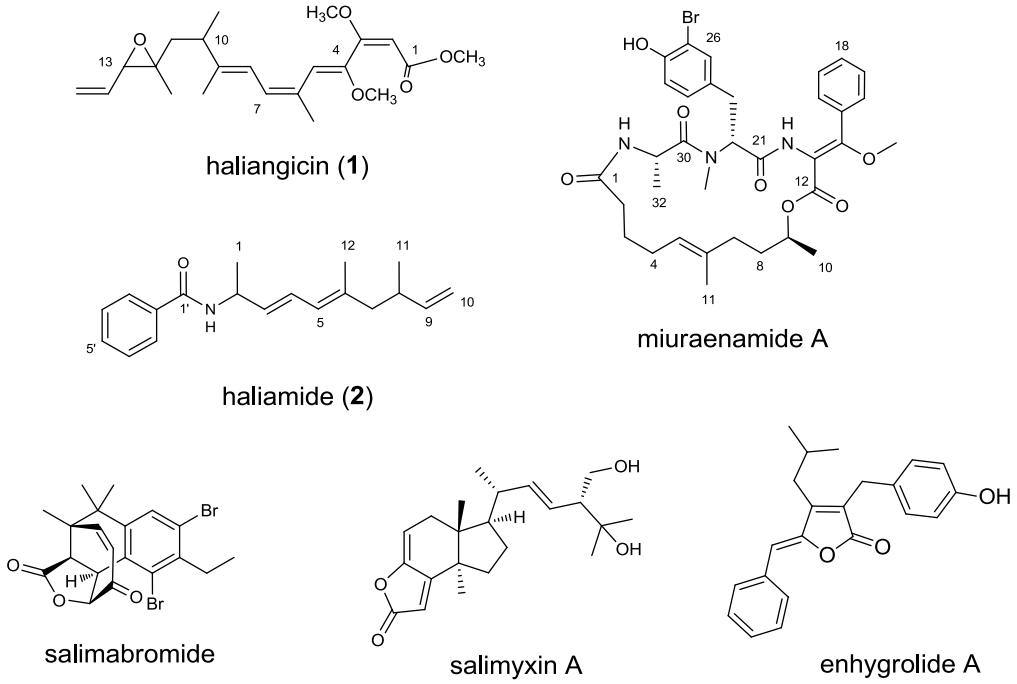


Figure 1-5. Myxobacterial products from marine isolates.

Besides haliangicin (**1**) and haliamide (**2**), our groups previously isolated a family of cyclic hybrid polyketide–depsipeptides antibiotics called miuraenamides from a slightly halophilic myxobacterium, *Paraliomyxa miuraensis* SMH-27-4, which was discovered in near-shore soil in Japan.[30, 35] The major compound, miuraenamides A, exhibited potent inhibition against the phytopathogenic fungus *Phytophthora capsici*, and moderate inhibition against some fungi and yeasts.[30] The electron transfer system of the mitochondrial respiratory chain was supposed to be the cellular target. Later chemical research revealed the importance of both the macrocyclic structure and the β-methoxyacrylate moiety from the antimicrobial activity. [35] In addition, miuraenamide A was also reported to be an actin filament stabilizer that showed potent antitumor activity.[39]

In 2013, from *Enhygromyxa salina* strains SWB005 and SWB007, salimyxin B and enhygrolide A were isolated and showed inhibitory activity toward the non-pathogenic *Arthrobacter cristallopoietes*. [37] Another compound, salimabromide (only present in strain SWB007), was isolated by *Felder et al.* and possesses a

novel carbon skeleton. [36] Its antibiotic activity was moderate with an MIC against *A. cristallopoietes* of 16 μg mL^{-1}.

Compared to the terrestrial myxobacteria, marine myxobacteria are hardly investigated systematically.[16] Although marine myxobacteria have potential to produce diverse secondary metabolites, their slow growth or difficult cultivation makes the production of antibiotics not effective. Furthermore, only a few genes express to produce secondary metabolites because most biosynthetic genes remain dormant.[40] Therefore, the studies on the fastidious marine myxobacteria remain challenging but quite attractive.

1.4. Biosynthesis of polyketides and nonribosomal peptides

The majority of the isolated myxobacterial products are polyketides (PKs), non-ribosomal polypeptides (NRPs) and their hybrids. They are known as the products biosynthesized by megasynthase polyketide synthase (PKSs) and nonribosomal peptide synthase (NRPSs). The PKSs and NRPSs are the world's biggest enzymes in size among the discovered proteins.[41] These megasynthase possess modular organizations which are composed of functional domains: each domain catalyzes a specific biosynthetic step; several related domains then join together to form units called modules.[42-44] A single PKSs or NRPSs may contain one module or multiple modules and usually, one complete biosynthetic pathway consists of several PKSs or NRPSs.

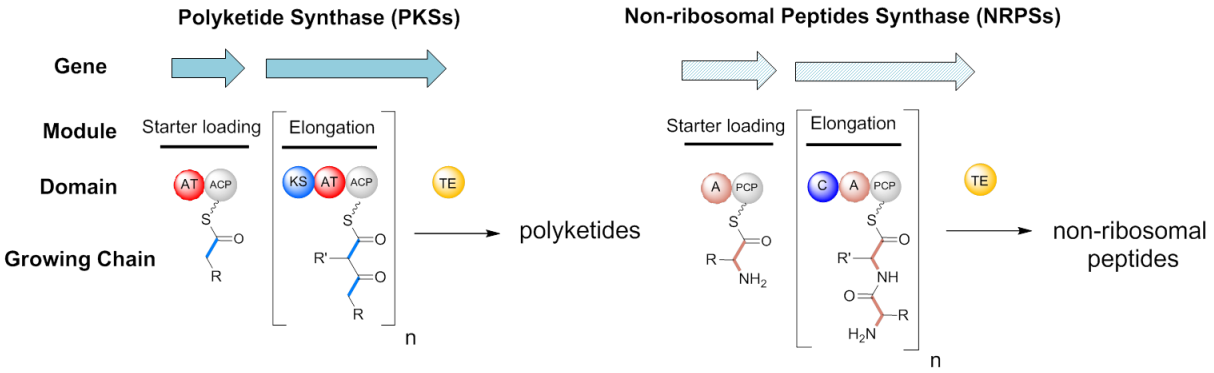


Figure 1-6. Biosynthetic strategies of polyketide synthase (PKSs) and non-ribosomal peptides synthase (NRPSs).

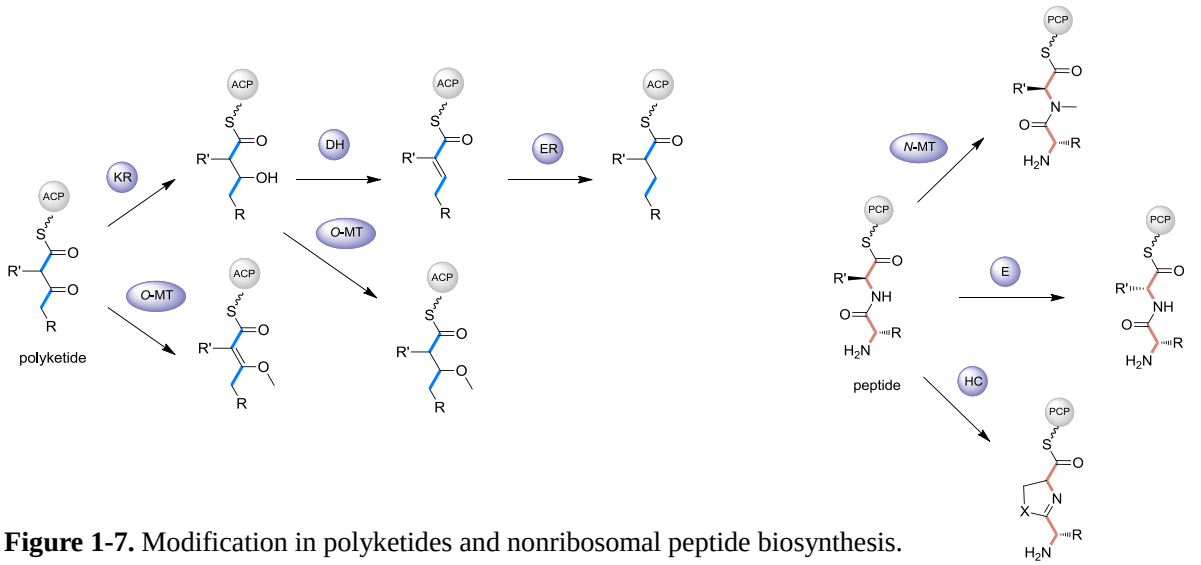


Figure 1-7. Modification in polyketides and nonribosomal peptide biosynthesis.

The PKSs and NRPSs perform highly similar biosynthesis to assemble their products with quite simple biosynthetic building blocks, short carboxylic acids and amino acids (Figure 1-6). In general, the PKSs biosynthesizes polyketide chains by successive condensation of acyl-CoA substrates and the NRPSs biosynthesizes peptides by condensation of activated amino acid units. In PKSs biosynthesis,[45-47] the acyl-CoA substrates are selected and transfer onto the acyl carrier proteins (ACP) by acyltransferase (AT) domain and then linked by ketosynthase (KS) domain through Claisen condensation. A typical PKS module minimally incorporates the three abovementioned domains, KS, AT and ACP. Other domains (if present) may modify the growing polyketide chains to generate β-hydroxyl (by ketoreductase, KR), β-methoxyl (by *O*-methyltransferase, *O*-MT), α,β-unsaturated (by dehydratase, DH) and saturated (by enoyl reductase, ER) structures (Figure 1-7). After all the elongation reactions finish, the full length intermediate is released by thioesterase (TE) domain and undergoes post-assembly tailoring to form the final product. For the biosynthesis of nonribosomal peptides,[48-51] the adenylation (A) domain of NRPSs selects and activates the amino acids; the C (condensation) domain catalyzes the peptide bond formation; the peptide carrier protein (PCP) holds the growing chains. The extending peptides may be modified by various domains such epimerization (E) domain, *N*-methyltransferase (*N*-MT) domain and heterocyclization (HC) domain (Figure 1-7). The PKSs and NRPSs may cooperate to produce products structurally contain both polyketide and non-ribosomal peptide components, which is known as the PKS-NRPS hybrids. Such hybrid systems

further increase the structural diversity of PKS-NRPS type products and therefore generate a large number of complicated compounds.[52]

With the fast-growing numbers of identified biosynthetic gene cluster being linked to their products, the last decades have seen the great progress in the understanding of PKSs and NRPSs biosynthesis. However, it has been reported that the unusual PKS/NRPS systems in myxobacteria always challenged a lot of conventional thinking of these megasynthase operations.[53, 54] The ongoing genome sequencing of myxobacterial strains suggests that their secondary metabolites are far from exhausted.[7, 27, 55] It could be expected that with the coming of myxobacterial 'genomic era', more details about the fascinating secondary metabolism of the myxobacteria as well as the molecular mechanisms behind their particular biosynthesis will be elucidated in the near future.

1.5. Genome mining

Nowadays the scientists who endeavor to discover new natural products encounter several challenges, among which the increasing rate of re-isolation of known compounds through conventional approaches has made the natural product discovery more and more inefficient. On the other hand, those metabolites with low abundance, unknown bioactivities, and those produced by cryptic gene clusters are always overlooked under the laboratory conditions.[56] To achieve a more efficient screening of novel bioactive compounds, lots of new technologies were developed to avoid the repeat isolation of existing compounds.[56-60] One of these approaches is the genome mining (Figure 1-8).[57, 61, 62]

Recent development of genome sequencing has provided easier access to the fully-sequenced microbial genome or metagenome ever than before. With the vast number of deposited sequences in hands, researchers realized that the microbial potential of producing novel products was underestimated even through long years of screening efforts. Therefore the genome mining approaches were adopted to fully uncover the biosynthetic potential of microorganisms. As defined by Bachmann *et. al.*, genome mining is "the process of technically translating secondary metabolite-encoding gene sequence data into purified molecules in tubes".[57] Generally speaking, the purpose of mining the genome of microbes is to connect the

biosynthetic gene clusters to their own products,[63] especially those that have not been identified.

The successful attempt of linking the biosynthetic genes to their molecules largely depends on the deep understanding of the existing biosynthetic system. As the key step in genome mining, comparing the sequences with the existing bioinformatics may give information on predicted function of genes, predicted substrates or predicted order of biochemistry, through which the gross structures can be estimated.[64] The follow-up efforts can be focused on the particular structural features or special biosynthetic reactions, leading to the discovery of unidentified products. Such genomics-guided approach holds the promise of speeding up the isolation of novel compounds from microbial producers.

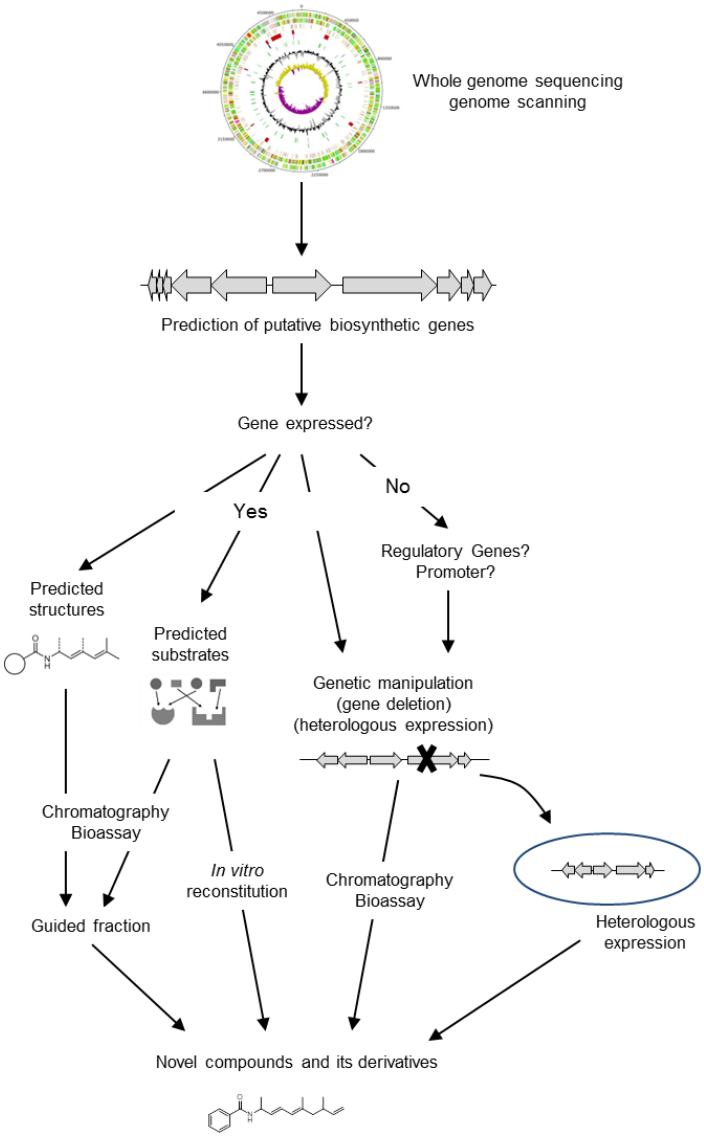


Figure 1-8. Strategies for discovery of novel products by genome mining.[65] (The scheme was modified.)

1.6. Outline of this work

Marine myxobacteria have considerable genetic potential to produce novel secondary metabolites despite their fastidious nature. In the previous work of our laboratory, the first marine myxobacterium *Haliangium ochraceum* SMP-2 was found to produce the antifungal polyketide haliangicin (**1**), which represents the first marine myxobacterial product, but its productivity is unsatisfactory and its biosynthetic machinery remains unclear. Accompanying haliangicin (**1**), an unreported PKS-NRPS hybrid haliamide (**2**) was also discovered from the secondary metabolome of *H. ochraceum*. The main topic of this study is to analyze the biosynthetic genes from the marine myxobacterium *H. ochraceum*, which are involved in the biosynthesis of these secondary metabolites.

In Chapter 2, the biosynthetic gene cluster of haliangicin (*hli*, 47.8 kbp) was analyzed and the detailed mechanism of haliangicin formation was deciphered. Taking the advantage of the heterologous expression system, I identified the biosynthetic precursor of the C-3 and C-4 position in haliangicin (**1**), achieved a highly effective production of haliangicin (10-fold greater amount and 3-fold faster in growth speed compared with the original producer) and generated some bioactive unnatural analogues of **1** through gene manipulation. A unique acyl-CoA dehydrogenase was found to catalyze an unusual γ,δ-dehydrogenation of the diketide starter unit, leading to the formation of the terminal alkene moiety of **1**. Biological evaluation of the analogues obtained through this study was performed to reveal the structure–activity relationship (SAR).

In Chapter 3, I used the published genome data of *Haliangium ochraceum* SMP-2[66] to clarify seven PKS/NRPS biosynthetic gene clusters including the biosynthetic locus for haliangicin (**1**). In a parallel work which focused on the secondary metabolic profile of *H. ochraceum*, a new PKS-NRPS hybrid compound, haliamide (**2**), was isolated and identified. The biosynthetic gene cluster of haliamide (*hla*, 21.7 kbp) was characterized through the genome mining of the producer, allowing an establishment of a model for the biosynthesis of **2**. The other five PKS/NRPS biosynthetic gene clusters were analyzed by bioinformatics tools, from which a putative biosynthetic gene cluster for aurafuron A (**3**) was discovered.

In Chapter 4, the major findings of my research were summarized.

Reference

1. Ngo LT, Okogun JI, Folk WR. 21st Century natural product research and drug development and traditional medicines. *Nat. Prod. Rep.* **30**, 584-592 (2013).

2. Newman DJ, Cragg GM. Natural Products As Sources of New Drugs over the 30 Years from 1981 to 2010. *J. of Nat. Prod.* **75**, 311-335 (2012).

3. Lewis K. Platforms for antibiotic discovery. *Nat. Rev. Drug Discov.* **12**, 371-387 (2013).

4. Pelaez F. The historical delivery of antibiotics from microbial natural products - Can history repeat? *Biochem. Pharmacol.* **71**, 981-990 (2006).

5. Donadio S, Maffioli S, Monciardini P, Sosio M, Jabes D. Antibiotic discovery in the twenty-first century: current trends and future perspectives. *J. Antibiot.* **63**, 423-430 (2010).

6. Reichenbach H. Biology of the myxobacteria: Ecology and taxonomy. In: *Myxobacteria II* (eds Dworkin M, Kaiser D). American Society for Microbiology (1993).

7. Weissman KJ, Muller R. A brief tour of myxobacterial secondary metabolism. *Bioorg. Med. Chem.* **17**, 2121-2136 (2009).

8. Gerth K, Pradella S, Perlova O, Beyer S, Muller R. Myxobacteria: proficient producers of novel natural products with various biological activities - past and future biotechnological aspects with the focus on the genus *Sorangium*. *J. Biotechnol.* **106**, 233-253 (2003).

9. Mauriello EMF, Mignot T, Yang ZM, Zusman DR. Gliding Motility Revisited: How Do the Myxobacteria Move without Flagella? *Microbiol. Mol. Biol. Rev.* **74**, 229- (2010).

10. Dworkin M. Recent advances in the social and developmental biology of the myxobacteria. *Microbiol. Rev.* **60**, 70- (1996).

11. Reichenbach H. Myxobacteria, producers of novel bioactive substances. *J. Ind. Microbiol. Biotechnol.* **27**, 149-156 (2001).

12. Berleman JE, Kirby JR. Deciphering the hunting strategy of a bacterial wolfpack. *Fems. Microbiol. Rev.* **33**, 942-957 (2009).

13. Garcia R, Gerth K, Stadler M, Dogma IJ, Muller R. Expanded phylogeny of myxobacteria and evidence

for cultivation of the 'unculturables'. *Mol. Phylogenet. Evol.* **57**, 878-887 (2010).

14. Dawid W. Biology and global distribution of myxobacteria in soils. *Fems. Microbiol. Rev.* **24**, 403-427 (2000).

15. Reichenbach H. The ecology of the myxobacteria. *Environ Microbiol* **1**, 15-21 (1999).

16. Schaberle TF, *et al.* Marine Myxobacteria as a Source of Antibiotics-Comparison of Physiology, Polyketide-Type Genes and Antibiotic Production of Three New Isolates of *Enhygromyxa salina*. *Mar. Drugs* **8**, 2466-2479 (2010).

17. Iizuka T, Jojima Y, Fudou R, Hiraishi A, Ahn JW, Yamanaka S. *Plesiocystis pacifica* gen. nov., sp nov., a marine myxobacterium that contains dihydrogenated menaquinone, isolated from the Pacific coasts of Japan. *Int. J. Syst. Evol. Microbiol.* **53**, 189-195 (2003).

18. Iizuka T, Jojima Y, Fudou R, Tokura M, Hiraishi A, Yamanaka S. *Enhygromyxa salina* gen. nov., sp nov., a slightly halophilic myxobacterium isolated from the coastal areas of Japan. *Syst. Appl. Microbiol.* **26**, 189-196 (2003).

19. Fudou R, Jojima Y, Iizuka T, Yamanaka S. *Haliangium ochraceum* gen. nov., sp nov and *Haliangium tepidum* sp nov.: Novel moderately halophilic myxobacteria isolated from coastal saline environments. *J. Gen. Appl. Microbiol.* **48**, 109-115 (2002).

20. Iizuka T, Jojima Y, Fudou R, Yamanaka S. Isolation of myxobacteria from the marine environment. *FEMS. Microbiol. Lett.* **169**, 317-322 (1998).

21. Zhang XJ, Yao Q, Cai ZP, Xie XL, Zhu HH. Isolation and Identification of Myxobacteria from Saline-Alkaline Soils in Xinjiang, China. *Plos One* **8**, 11 (2013).

22. Hook LA. Distribution of Myxobacters in Aquatic Habitats of an Alkaline Bog. *Appl. Environ. Microbiol.* **34**, 333-335 (1977).

23. Dawid W, Gallikowski CA, Hirsch P. Psychrophilic Myxobacteria from Antarctic Soils. *Polarforschung*, **58**, 271-278 (1988).

24. S. M. *The World of Microorganisms*. The Tsukuba Press (2006).

25. Weissman KJ, Muller R. Myxobacterial secondary metabolites: bioactivities and modes-of-action. *Nat.*

Prod. Rep. **27**, 1276-1295 (2010).

26. Schaberle TF, Lohr F, Schmitz A, Konig GM. Antibiotics from myxobacteria. *Nat. Prod. Rep.* **31**, 953-972 (2014).

27. Wenzel SC, Mueller R. Myxobacteria-'microbial factories' for the production of bioactive secondary metabolites. *Mol. BioSyst.* **5**, 567-574 (2009).

28. Hofle GH, Bedorf N, Steinmetz H, Schomburg D, Gerth K, Reichenbach H. Epothilone A and B - Novel 16-membered macrolides with cytotoxic activity: Isolation, crystal structure, and conformation in solution. *Angew. Chem. Int. Ed. (English)* **35**, 1567-1569 (1996).

29. Nettles JH, Li HL, Cornett B, Krahn JM, Snyder JP, Downing KH. The binding mode of epothilone A on alpha,beta-tubulin by electron crystallography. *Science* **305**, 866-869 (2004).

30. Iizuka T, *et al.* Miuraenamides A and B, novel antimicrobial cyclic depsipeptides from a new slightly halophilic myxobacterium: Taxonomy, production, and biological properties. *J. Antibiot.* **59**, 385-391 (2006).

31. Zhang J, Liu Z, Wang S, Jiang P. Characterization of a bioflocculant produced by the marine myxobacterium *Nannocystis* sp NU-2. *Appl. Microbiol. Biotechnol.* **59**, 517-522 (2002).

32. Fudou R, Iizuka T, Yamanaka S. Haliangicin, a novel antifungal metabolite produced by a marine myxobacterium 1. Fermentation and biological characteristics. *J. Antibiot.* **54**, 149-152 (2001).

33. Fudou R, Iizuka T, Sato S, Ando T, Shimba N, Yamanaka S. Haliangicin, a novel antifungal metabolite produced by a marine myxobacterium 2. Isolation and structural elucidation. *J. Antibiot.* **54**, 153-156 (2001).

34. Kundim BA, *et al.* New haliangicin isomers, potent antifungal metabolites produced by a marine myxobacterium. *J. Antibiot.* **56**, 630-638 (2003).

35. Ojika M, Inukai Y, Kito Y, Hirata M, Iizuka T, Fudou R. Miuraenamides: Antimicrobial cyclic depsipeptides isolated from a rare and slightly halophilic myxobacterium. *Chem. Asian J.* **3**, 126-133 (2008).

36. Felder S, *et al.* Salimabromide: Unexpected Chemistry from the Obligate Marine Myxobacterium

Enhygromxya salina. *Chem. Eur. J.* **19**, 9319-9324 (2013).

37. Felder S, Kehraus S, Neu E, Bierbaum G, Schaberle TF, Konig GM. Salimyxins and Enhygrolides: Antibiotic, Sponge-Related Metabolites from the Obligate Marine Myxobacterium *Enhygromyxa salina*. *ChemBioChem* **14**, 1363-1371 (2013).

38. Sato J. Search for novel secondary metabolites from marine myxobacteria. Nagoya University, Master thesis (2015).

39. Sumiya E, *et al.* Cell-Morphology Profiling of a Natural Product Library Identifies Bisebromoamide and Miuraenamide A as Actin Filament Stabilizers. *ACS Chem. Biol.* **6**, 425-431 (2011).

40. Hosaka T, *et al.* Antibacterial discovery in actinomycetes strains with mutations in RNA polymerase or ribosomal protein S12. *Nat. Biotechnol.* **27**, 462-464 (2009).

41. Weissman KJ. The structural biology of biosynthetic megaenzymes. *Nat. Chem. Biol.* **11**, 660-670 (2015).

42. Cane DE, Walsh CT. The parallel and convergent universes of polyketide synthases and nonribosomal peptide synthetases. *Chem. Biol.* **6**, R319-R325 (1999).

43. Keating TA, Walsh CT. Initiation, elongation, and termination strategies in polyketide and polypeptide antibiotic biosynthesis. *Curr. Opin. Chem. Biol.* **3**, 598-606 (1999).

44. Fischbach MA, Walsh CT. Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: Logic, machinery, and mechanisms. *Chem. Rev.* **106**, 3468-3496 (2006).

45. Weissman KJ. Polyketide biosynthesis: understanding and exploiting modularity. *Philos. Trans. R. Soc. Lond. Ser. A-Math Phys. Eng. Sci.* **362**, 2671-2690 (2004).

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

47. Khosla C. Structures and Mechanisms of Polyketide Synthases. *J. Org. Chem.* **74**, 6416-6420 (2009).

48. Marahiel MA, Stachelhaus T, Mootz HD. Modular peptide synthetases involved in nonribosomal peptide synthesis. *Chem. Rev.* **97**, 2651-2673 (1997).

49. Doyle S. Nonribosomal Peptide Synthesis. In: *Amino Acids, Peptides and Proteins in Organic*

Chemistry. Wiley-VCH Verlag GmbH & Co. KGaA (2009).

50. Schoenafinger G, A Marahiel M, Begley TP. Nonribosomal Peptides: Biosynthesis. In: *Wiley Encyclopedia of Chemical Biology*. John Wiley & Sons, Inc. (2007).

51. Schwarzer D, Finking R, Marahiel MA. Nonribosomal peptides: from genes to products. *Nat. Prod. Rep.* **20**, 275-287 (2003).

52. Walsh CT. The chemical versatility of natural-product assembly lines. *Accounts. Chem. Res.* **41**, 4-10 (2008).

53. Wenzel SC, Muller R. Myxobacterial natural product assembly lines: fascinating examples of curious biochemistry. *Nat. Prod. Rep.* **24**, 1211-1224 (2007).

54. Li Y, Mueller R. Non-modular polyketide synthases in myxobacteria. *Phytochemistry* **70**, 1850-1857 (2009).

55. Wenzel SC, Muller R. The biosynthetic potential of myxobacteria and their impact on drug discovery. *Curr. Opin. Drug. Discov. Dev.* **12**, 220-230 (2009).

56. Reen FJ, Romano S, Dobson ADW, O'Gara F. The Sound of Silence: Activating Silent Biosynthetic Gene Clusters in Marine Microorganisms. *Mar. Drugs* **13**, 4754-4783 (2015).

57. Bachmann BO, Van Lanen SG, Baltz RH. Microbial genome mining for accelerated natural products discovery: is a renaissance in the making? *J. Ind. Microbiol. Biotechnol.* **41**, 175-184 (2014).

58. Genilloud O. The re-emerging role of microbial natural products in antibiotic discovery. *Anton. Leeuw. Int. J. Gen. Mol. Microbiol.* **106**, 173-188 (2014).

59. Li K, Chung-Davidson Y-W, Bussy U, Li W. Recent Advances and Applications of Experimental Technologies in Marine Natural Product Research. *Mar. Drugs* **13**, 2694-2713 (2015).

60. Monciardini P, Iorio M, Maffioli S, Sosio M, Donadio S. Discovering new bioactive molecules from microbial sources. *Microb. Biotechnol.* **7**, 209-220 (2014).

61. Helfrich EJN, Reiter S, Piel J. Recent advances in genome-based polyketide discovery. *Curr. Opin. Biotechnol.* **29**, 107-115 (2014).

62. Zotchev SB, Sekurova ON, Katz L. Genome-based bioprospecting of microbes for new therapeutics.

Curr. Opin. Biotechnol. **23**, 941-947 (2012).

63. Walsh CT, Fischbach MA. Natural Products Version 2.0: Connecting Genes to Molecules. *J. Am. Chem. Soc.* **132**, 2469-2493 (2010).

64. Challis GL. Mining microbial genomes for new natural products and biosynthetic pathways. *Microbiol.(UK)* **154**, 1555-1569 (2008).

65. Corre C, Challis GL. New natural product biosynthetic chemistry discovered by genome mining. *Nat. Prod. Rep.* **26**, 977-986 (2009).

66. Ivanova N, *et al.* Complete genome sequence of Haliangium ochraceum type strain (SMP-2(T)). *Stand. Genomic. Sci.* **2**, 96-106 (2010).

Chapter 2. Biosynthesis of haliangicin, the first marine myxobacterial polyketide

2.1. Introduction

Natural products derived from microorganisms have a long history of use as antibiotics for clinical use. As an alternative to conventional antibiotic producers such as Streptomyces, myxobacteria have attracted considerable attention, especially since the remarkable discovery of epothilones as microtubule-stabilizing agent.[1, 2] Myxobacteria are unique Gram-negative proteobacteria, characterized by gliding and by the formation of multicellular fruiting bodies during their complicated life cycles.[3] They hold the largest genome size in bacteria and exhibit prominent abilities to produce structurally diverse bioactive secondary metabolites having novel modes of action,[4] which therefore make them a promising source of new antibiotics. Marine myxobacteria are particularly rare and difficult to handle;[5-10] consequently, their products have only recently been examined. Haliangicin (**1,** Fig. 2-1a), a long-chain polyene bearing a unique terminal epoxy alkene group and a β-methoxyacrylate pharmacophore, was isolated from a halophilic myxobacterium *Haliangium ochraceum* SMP-2 (Fig. 2-1b) and represents the first marine myxobacterial secondary metabolite.[11, 12] It was reported to be a potent fungicide that interferes with the electron flow within the cytochrome *b-c1* segment in fungal mitochondrial respiratory chains.[11]

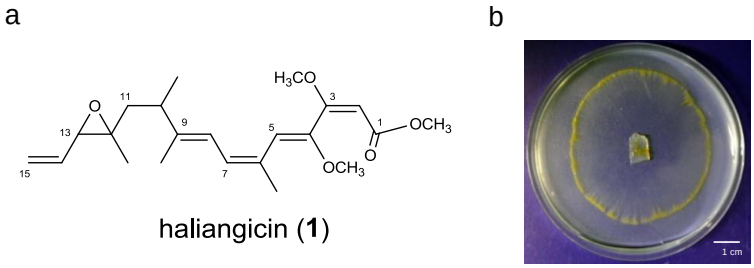


Figure 2-1. Haliangicin, the first marine myxobacterial antibiotic.

(**a**) Structure of haliangicin (**1**).

(**b**) Gliding swarms of *H. ochraceum* SMP-2 strain on an agar medium containing 2% NaCl.

Most myxobacterial metabolites are polyketides, nonribosomal peptides, or their hybrids; these are known to be biosynthesized by megasynthase polyketide synthases (PKSs) and nonribosomal peptide synthases (NRPSs). Although it has been reported that the sequences of polyketide synthase (PKS) genes in myxobacterial isolates from marine environments were highly novel (all sequences showed identity less than 70% to the known PKS sequences),[13] the researches on the secondary metabolites of marine myxobacteria have been hampered due to negative factors such as difficulty in isolation, slow growth, poor productivity and diversity of metabolites (<1 mg L^{-1} after two-week culture for haliangicin without other significant metabolites), and high tendency of cell aggregation. Furthermore, the lack of a well-developed genetic methodology applicable to the marine strains is another major obstacle. Therefore, the biosynthetic researches on the scarce marine myxobacteria are still at dawn and remain quite challenging and attractive.

In the previous studies, a heterologous expression system of haliangicin (**1**) was established to achieve an efficient production of **1** and to circumvent the difficult genetic manipulation in the native marine host (Fig. 2-2a).[14] In brief, a cosmid library of the *H. ochraceum* genome was constructed by using the *E. coli/Streptomyces* shuttle vector pDW103.[15] The resulting library, consisting of approximately 1700 cosmid clones, was then screened by colony hybridization using probes homologous to the ketosynthase (KS) genes of terrestrial myxobacteria. Two PKS-positive cosmids c7-6E and c10-11C were each modified through λ-Red recombineering[16] and integrated into the chromosome of *Myxococcus xanthus* ATCC25232, a fast-growing terrestrial myxobacterium, in a stepwise manner to reconstitute the haliangicin biosynthetic pathway. The constructed heterologous host strain was found to produce haliangicin (**1**) successfully, as revealed by HPLC (Fig. 2-2b).

In this chapter, I will describe the identification of the haliangicin biosynthetic gene cluster (*hli*) as the first representative of marine myxobacterial biosynthetic assembly line. The sequence of *hli* cluster was analyzed and the biosynthetic pathway of haliangicin was proposed. Taking the advantage of the heterologous expression host of haliangicin, I firstly identified the biosynthetic precursor of C-3 and C-4 in the haliangicin skeleton. I also performed optimization of haliangicin production by feeding of the identified biosynthetic precursors, as well as the genetic manipulation for functional analysis of

biosynthetic enzymes that involved in the haliangicin biosynthesis.

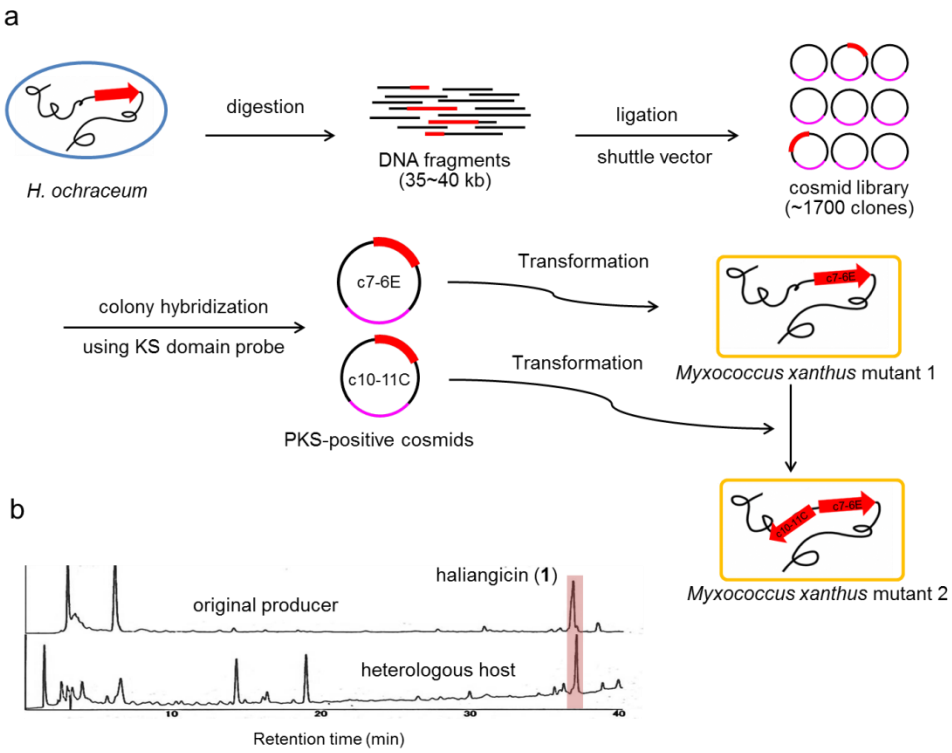


Figure 2-2. A heterologous expression system of haliangicin (**1**).

(**a**) The outline of the heterologous expression of haliangicin in *Myxococcus xanthus*.

(**b**) Production of haliangicin by the heterologous host.

2.2. Materials and methods

2.2.1. General

Evaporation of solvents was carried out using a rotary evaporator under reduced pressure at temperatures around 35 °C. All the haliangicin-related samples were handled in dim light with aluminum foil protection, and always stored at -20 °C in benzene. Silica gel 60 F254 (0.25-μm thickness; Merck, Rahway, NJ) was used for thin-layer chromatography (TLC). Wakogel® C-300 (Wako, Japan) was used for open column chromatography. Cosmosil 75C18-OPN (Nacalai Tesque, Inc., Kyoto, Japan) was used for reversed-phase (RP) open column chromatography. Flash chromatography was carried out with a medium-pressure gradient system equipped with a Pump Module C-605 and a Pump Manager C-615 (BÜCHI, Switzerland).

HPLC was performed with a high-pressure gradient system equipped with a MD-2018 plus photodiode array detector (Jasco, Tokyo) and PU-1580 pumps. Preparative HPLC was performed on a high-pressure gradient system equipped with PU-1586 and PU-2086 pumps and a UV-1570 detector (Jasco, Tokyo). Specific rotation was measured by using a DIP-370 digital polarimeter (Jasco, Tokyo). FTIR spectra were recorded on a FT/IR-4100 Fourier transform infrared spectrometer (Jasco, Tokyo). UV spectra were recorded on a V-530 UV/Vis spectrophotometer (Jasco, Tokyo). Mass spectra (MS) were recorded on a Mariner Biospectrometry Workstation (Applied Biosystems, CA) in the positive-ESI mode. LC/ESI-TOF MS analysis was performed on an Agilent 1100 HPLC system (Agilent Technologies, CA) with a Mariner Biospectrometry Workstation (Applied Biosystems, CA). ^{1}H and ^{13}C NMR spectra were recorded at 27 ^{o}C on an Avance 400 (400 MHz for ^{1}H) spectrometer (Bruker, Rheinstetten, Germany). The chemical shifts (ppm) were referenced to the solvent ($CDCl_3$) peaks at $\delta_H = 7.26$ ppm and $\delta_C = 77.0$ ppm (residual $CHCl_3$).

All oligonucleotides were synthesized by STAR Oligo (RIKAKEN Co., Ltd., Nagoya, Japan). PCR amplifications were performed on a Takara PCR Thermal Cycler Dice Gradient TP600 (Takara, Bio. Inc.) or Gene Amp PCR System 9700 (Applied Biosystems) with GoTaq® Green Master Mix (Promega, Madison, WI) except that mentioned elsewhere. Purification of DNA fragments from agarose gel (Agarose L03, Takara Bio Inc.) electrophoresis, restriction digestion or PCR amplification was performed using Wizard® SV Gel and PCR Clean-Up System (Promega, Madison, WI). Plasmid DNA from recombinant *E. coli* was extracted by GenElute Plasmid Miniprep Kit (Sigma-Aldrich Co., MO) or Wizard® Plus SV Minipreps DNA Purification System (Promega, Madison, WI). High quality genomic DNA of *H. ochraceum* and *M. xanthus* were purified using QIAamp DNA Mini Kits (QIAGEN, Hilden, Germany) and stored at -30 ^{o}C. *M. xanthus* mutants DNA were extracted with FavorPrep Tissue Genomic DNA Extraction Mini Kit (Favorgen Biotech Corp., Taiwan, China). Small-scale and rapid purification of polyhistidine proteins was performed by MagneHis™ Ni-Particles (Promega, Madison, WI). Protein concentrations were determined by Bradford protein assay [Pierce™ Coomassie (Bradford) Protein Assay Kit, Life technologies] using bovine serum albumin (BSA) as a standard. The OD_{600} values of bacterial culture were measured on Ultraspec 2100 pro UV/visible spectrophotometer (Amersham Biosciences).

Table 2-1. Bacterial strains used in this study.

Bacterial strains	Description	Source
E. coli		
Mach 1	General cloning host	Invitrogen
XL1-Blue	General cloning host	Stratagene
SW105	λ-Red recombineering host	NCI, [21]
KRX	Expression host for recombinant proteins	Promega
NEB 5-alpha	Cloning host for Gibson Assembly cloning	NEB
M. xanthus		
ATCC25232	Wild type, heterologous host of *hli* gene cluster	Ajinomoto
c10-11C/c7-6E	Mutant harboring the entire haliangicin biosynthetic gene cluster, Kan^R, $Oxytet^R$, $Thio^R$, haliangicin (**1**), haliangicin B (**4**), haliangicin C (**5**), haliangicin D (**6**) producing	This study
Δ*hliD*	*hliD* disruptant, Kan^R, $Oxytet^R$, $Thio^R$, Nrs^R, 1-*O*-demethylhaliangicin (**7**), 1-*O*-demethyl-12,13-deoxylhaliangicin (**8**) producing	This study
Δ*hliE*	*hliE* disruptant, Kan^R, $Oxytet^R$, $Thio^R$, Nrs^R, Non-haliangicin producing	This study
Δ*hliH*	*hliH* disruptant, Kan^R, $Oxytet^R$, $Thio^R$, Nrs^R, Non-haliangicin producing	This study
Δ*hliP*	*hliP* disruptant, Kan^R, $Oxytet^R$, $Thio^R$, Nrs^R, Non-haliangicin producing	This study
Δ*hliQ*	*hliQ* disruptant, Kan^R, $Oxytet^R$, $Thio^R$, Nrs^R, Non-haliangicin producing	This study
Δ*hliR*	*hliR* disruptant, Kan^R, $Oxytet^R$, $Thio^R$, Nrs^R, 14,15-dihydrohaliangicin (**9**) producing	This study
Δ*hliU*	*hliU* disruptant, Kan^R, $Oxytet^R$, $Thio^R$, Nrs^R, 12,13-deoxylhaliangicin (**10**) producing	This study
Δ*orf22*	*orf22* disruptant, Kan^R, $Oxytet^R$, $Thio^R$, Nrs^R, haliangicin (**1**) producing	This study
H. ochraceum		
SMP-2	native haliangicin (**1**), haliangicin B (**4**), haliangicin C (**5**), haliangicin D (**6**) producer	Ajinomoto

Table 2-2. Vectors used in this study.

Plasmids	Description	Source
General vectors		
pGEM-T easy vector	General TA cloning vector, Amp^R	Promega
pCR 2.1-TOPO	General TA cloning vector, Amp^R, Kan^R	Invitrogen
pHSG398	General cloning vector, Cm^R	Takara
pET-32a	Expression vector for thioredoxin and His_6-fusion recombinant protein, Amp^R	Novagen
pNR1	Cloning template of nourseothricin resistance gene, Nrs^R	[22]
pHSG398 Infu Nrs^R	General cloning vector for construction of disruption vector, pHSG398-derived, Nrs^R under Tn5 promoter	This study
Disruption vectors	Description	Transformants
pHSG 3ECH Infu Nrs^R	pHSG398 Infu Nrs^R-derived, carrying internal fragment of *hliC* (633 bp), Nrs^R	-
pHSG GemOMT4 Infu Nrs^R	pHSG398 Infu Nrs^R-derived, carrying internal fragment of *hliD* (522 bp), Nrs^R	*M. xanthus* Δ*hliD*
pHSG 5MbLTE Infu Nrs^R	pHSG398 Infu NrsR-derived, carrying internal fragment of *hliE* (706 bp), NrsR	M. xanthus Δ*hliE*
pHSG 8FkbH Infu Nrs^R	pHSG398 Infu Nrs^R-derived, carrying internal fragment of *hliH* (750 bp), Nrs^R	*M. xanthus* Δ*hliH*
pHSG 16F1 Infu Nrs^R	pHSG398 Infu Nrs^R-derived , carrying internal fragment of *hliP* (1111 bp), Nrs^R	*M. xanthus* Δ*hliP*
pHSG 17MT Infu Nrs^R	pHSG398 Infu Nrs^R-derived, carrying internal fragment of *hliQ*(771 bp), Nrs^R	*M. xanthus* Δ*hliQ*
pHSG 18ACAD Infu Nrs^R	pHSG398 Infu Nrs^R-derived, carrying internal fragment of *hliR* (715 bp), Nrs^R	*M. xanthus* Δ*hliR*
pHSG 21Ox Infu Nrs^R	pHSG398 Infu Nrs^R-derived, carrying internal fragment of *hliU* (730 bp), Nrs^R	*M. xanthus* Δ*hliU*
pHSG 22MscS Infu Nrs^R	pHSG398 Infu Nrs^R-derived, carrying internal fragment of *orf22* (769 bp), Nrs^R	*M. xanthus* Δ*orf22*
Expression vectors	Description	Expressed protein
pET32a HliR	pET-32a-derived, expression vector for *N*-His_6-fusion HliR, Amp^R	HliR
pET32a HliD	pET-32a-derived, expression vector for *N*-His_6-fusion HliD, Amp^R	HliD

Table 2-3. Primers used in this study.

Primers	Sequence (5’ → 3’)
Amp Nrs-r	GATACGGGAGGGCTTACATCCACGGGACTTGAGAC
HliD Exp Infu-F	GCCATGGCTGATATCGGAATGCAGACGTCCGCCAAA
HliD Exp Infu-R	CAGTGGTGGTGGTGGTGGTGCTCACTTATAGGCGGCGAAC
HliE Out f	GGTGCTTTTTCACGTAGGCG
HliE Out r	GCTGAGCGTAGAAATCCCGA
HliH-Ex f	TTGTTGAAGTTTGAGCGCGG
HliH-Ex r	CCATCCGGCCACTGAGTATC
HliR Exp Infu f	GGCTGATATCGGATCCATGAGCGCAGATACAACCAAG
HliR Exp Infu r	GGTGGTGGTGCTCGATTCCATGTCGCTCGGTGTC
Kan Nrs-f	GCTATGACTGGGCACCCCTTTCTCTCCTGGAAGACT
M13f(-47)	CGCCAGGGTTTTCCCAGTCACGAC
M13r(-48)	AGCGGATAACAATTTCACACAGGA
ORF16 r4	CAGGTCTCGGTGATTTCGGC
ORF16-f1	TTTTGGTACAGTCGGTCGCA
ORF16-f3	GTTGCGTCCTGATGTATGCG
ORF16-r1	CTCCAGGTTCGCCAGTTGAT
ORF17 f1	AGATGCCAAAGTCGTCGGTT
ORF17 r1	AGCGGGTTCTTCTTGGTTGT
ORF17MT Out f	ACGCGTTTCACTCGTCGAT
ORF17MT Out r	TGCGCGGGTAGAAGGTAAAG
ORF18ACAD f	CTGGACATCTACGAGGCGTT
ORF18ACAD r	CAGCACGTGACGCGAAAATA
ORF18ACAD-Out f	ATCTTCCGCAATCAGCAGGA
ORF18ACAD-Out r	GACGTTTGTCCGCTTTCGAT
ORF19Out f	CTACGTGTCGTTGGTGGAGG
ORF19Out r	TGATGCGAAACGGAAAGGGA
ORF21Ox-In F	CGTGCAGACCTACTCGTGTT
ORF21Ox-In R	TGCGCAAGCTCTCTCTGAAA
ORF21Ox-Out F2	GAGCATCTCGGGACTTACGC
ORF21Ox-Out R	GCTGGGGGTCATAGAGAAGG
ORF22MscS Ex-f	CCCATCTCCTCGATTCACCG
ORF22MscS Ex-r	GAGGAATCGCACAGCGAAAC
ORF4-MT-EX-f1	ATGTCTGGCAAGAGCAGGTC
ORF4-MT-EX-r1	TCATGCGACTGGACGAACAA
pET32F	AACGCCAGCACATGGACAGC
pET32R	GCTAGTTATTGCTCAGCGGT
pHSG 21Ox Infu F	CCGGGGATCCTCTAGTAGGGCGGAGCTGGAGATG
pHSG 21Ox Infu R	GGCCAGTGCCAAGCTGAAGTGCCGACGATGATTGC
pHSG NrsR Infu-F	CACGTAAGAGGTTCCTGATGGTTCACGTAGTGGGC
pHSG NrsR Infu-R	CCAGGCGTTTAAGGGACTACGATACGGGAGGGCTT

(Continued)

Table 2-3. Primers used in this study. *(Continued)*

Primers	Sequence (5' → 3')
pHSG17MT Infu-f	CCGGGGATCCTCTAGGTCGTCGGTTCTGCCCTCG
pHSG17MT Infu-r	GGCCAGTGCCAAGCTGCGGGCTCGAAGGAGAAGATG
pHSG18ACAD Infu-f	CCGGGGATCCTCTAGTGGACATCTACGAGGCGTTTC
pHSG18ACAD Infu-r	GGCCAGTGCCAAGCTGCGCCGAACACGTGATTGTT
pHSG22Mscs Infu F	CCGGGGATCCTCTAGACAGCATCTCGCTGATCGTC
pHSG22Mscs Infu R	GGCCAGTGCCAAGCTCGGCTCTCCAACATCTCCTG
pHSG398 Inverse-F	CCCTTAAACGCCTGGTGCTA
pHSG398 Inverse-R	GGAACCTCTTACGTGCCGAT
pHSG3ECH Infu F	CCGGGGATCCTCTAGGGGACTGCCAGTGTTTCCTT
pHSG3ECH Infu R	GGCCAGTGCCAAGCTTGAGGTACTGATTCGCGAGC
pHSG5MbL-TE Infu-f	CCGGGGATCCTCTAGCATGATGATGCTGCCGACTG
pHSG5MbL-TE Infu-r	GGCCAGTGCCAAGCTCAATGAGTTGATCCCGCTGC
pHSG8FkbH Infu-f	CCGGGGATCCTCTAGCCGTAGGAGCCGTACTTGTC
pHSG8FkbH Infu-r	GGCCAGTGCCAAGCTTAACGACGCGGAGCAAAGAG
SP6	ATTTAGGTGACACTATAGAATAC
T7	TGTAATACGACTCACTATAGGGC

2.2.2. Bacterial strains, plasmids, primers and culture conditions

All bacterial strains, vectors, and oligonucleotides used in this study are listed in Tables 2-1~3. The marine myxobacterium *Haliangium ochraceum* SMP-2 was cultivated as previously described.[11] CTT medium[17] was used for the preculture of *Myxococcus xanthus* ATCC25232 (wild type) and all its mutants at 30 °C. Production medium[18] supplemented with 2–4% (w/v) SEPABEADS SP207 resin (Mitsubishi Chemical Co., Tokyo) and a 0.1% (v/v) trace-elements solution (TES)[19] was used for the heterologous production of haliangicin (**1**) and its related metabolites. A typical cultivation for the heterologous production was performed on a rotary shaker at 180 rpm and last for 5~7 days at 30 ^{o}C. *Escherichia coli* strains XL1-Blue (Stratagene), Mach1 (Life technologies), and NEB 5-alpha (NEB) used for cloning or propagation of plasmids were cultured in Luria-Bertani at 37 ^{o}C. The *E. coli* strain SW105 (NCI, U.S.A.) for λ-Red recombineering was cultivated in LB medium at 30 ^{o}C according to the published protocol.[16] *E. coli* KRX (Promega) as a recombinant proteins expression host was handled according to the guideline provided by the manufacturer. Antibiotics were supplemented to proper concentrations: kanamycin sulfate (Kan, 100~200 μg/mL), ampicillin (Amp, 50~100 μg/mL), oxytetracycline (Oxytet, 10 or 12.5 μg/mL), thiostrepton (Thio, 50 μg/mL) and nourseothricin (Nrs, 100 μg/mL) if necessary.

2.2.3. Heterologous production of haliangicins in *M. xanthus* c10-11C/c7-6E

The heterologous host *M. xanthus* c10-11C/c7-6E were cultivated in a preparative scale (5.75 L production medium supplemented with 2-4% (w/v) SEPABEADS SP207 absorber resin at 30 ^{o}C). The cells and resin were harvested from 4~6 days culture broth by centrifugation (6000 rpm, 5 min) and extracted with acetone (5 L × 2) at 30 ^{o}C for 30 min. After filtration, the combined filtrate were concentrated *in vacuo* and extracted with EtOAc (150 mL × 2, 100 mL× 1). The EtOAc extracts were combined and dried to afford a yellow oily extract. The extract was further partitioned between 50% MeOH and hexane/EtOAc (4/1, v/v) to yield an oily crude sample (1.61 g) from the hexane/EtOAc phase. The crude oil was chromatographed on a Hi-flash silica gel column (size L, 30g, Yamazen Co., Osaka, Japan) with a linear gradient of 0% to 20% EtOAc in hexane (40 min) at a flow rate 12 mL/min. The haliangicin-containing fraction was then applied to preparative HPLC [Develosil ODS-UG-5 (ϕ20 × 250 mm, Nomura Chemical Ltd.), 75% MeOH, 10 mL/min, 80 min, monitored at 310 nm] to obtain pure haliangicin (**1**, 7.7 mg), haliangicin B (**4**, 0.9 mg), haliangicin C (**5**, 0.3 mg) and haliangicin D (**6**, 0.4 mg) (Fig. 2-5).The NMR data of these compounds are identical to those reported.[12, 20]

2.2.4. Feeding experiments with stable-isotope-labelled precursors

The [U-$^{13}C_3$]glycerol and [1-^{13}C]bromoacetic acid used in this study was purchased from Sigma-Aldrich. [1-^{13}C]bromoacetic acid (160 mg, 1.1 mmol) was mixed with 3 mL of 3N NaOH and stirred at 50 ^{o}C for 5 hours. After acidification to pH = 6 with 7.5 mL of 1N HCl, the solution was concentrated to dryness and extracted with MeOH. The methanol extracts was further purified by silica gel chromatography to afford [1-^{13}C]glycolic acid (80 mg).

The heterologous host *M. xanthus* c10-11C/c7-6E was cultivated in three or four 3×100 mL aliquots of production medium supplemented with 2% (w/v) absorbent resin (SEPABEADS SP207) and 0.1% TES. Filter-sterilized 0.2 M solutions of the isotope-labelled compounds [1-^{13}C]glycolic acid and [U-$^{13}C_3$]glycerol in water were separately added in two portions to the bacterial culture after 48 hours and 72 hours, respectively, to give a final concentration of 4 mM. After cultivation for six or seven days in total at 30 °C and 180 rpm, labelled haliangicin was extracted, purified, and analysed by ^{13}C NMR spectroscopy (Fig. 2-7 and 2-8).

2.2.5. Optimization of haliangicin production in the *M. xanthus* heterologous host

The heterologous host *M. xanthus* c10-11C/c7-6E was cultured in the production medium. 1 M Aqueous solutions of sodium acetate, sodium propionate, and glycerol were prepared, filter-sterilized, and added to the autoclaved culture medium. L-Methionine was added before autoclaving. For higher concentrations of sodium acetate (100 mM or 200 mM), an autoclaved 5 M stock solution was added in two portions to the medium on the second and third day, respectively, to avoid inhibition of the growth of *M. xanthus* in its early phase. Glucose was prepared as 1g/mL aqueous stock solution and added to the medium on the third day. After cultivation at 30 °C and 180 rpm for the scheduled time, the resins and cell pellets were separated from the medium by centrifugation (6000 rpm, 10 min) and extracted twice with acetone. The crude extracts were filtered, evaporated to dryness, and subjected to HPLC analysis [Develosil ODS-UG-5 (ϕ 4.6 × 250 mm), λ = 290 nm; 75% MeOH–H_2O, flow rate: 1 mL/min]. The results are summarized in Fig 2-9 and 2-10.

2.2.6. Analysis of the haliangicin biosynthetic gene cluster (*hli*).

The assembled continuous contig of 63.4 kbp that putatively harbors the entire haliangicin biosynthetic gene cluster (*hli*) was annotated by various sequence-analysis tools, including BLAST, CDD,[23, 24] and antiSMASH analysis.[25, 26] The multiple alignments of amino acid sequences were generated by using the Clustal Omega program (http://www.ebi.ac.uk/Tools/msa/clustalo/) provided by the European Molecular Biology Institute (Hinxton, UK) and neighbor-joining trees of proteins were generated by Geneious Tree Builder. KS domains in *hli* were extracted and analyzed by using the NaPDoS Web tool (http://napdos.ucsd.edu/).[27]

2.2.7. Inactivation of haliangicin biosynthetic genes in the heterologous host *M. xanthus* c10-11C/c7-6E

The biosynthetic genes involved in the haliangicin biosynthesis were systematically inactivated to determine their biosynthetic functions. Disruption of the targeted *hli* genes was accomplished by integration of a disruption vector into the genome of the heterologous host *M. xanthus* c10-11C/c7-6E through single crossover recombination (Fig. 2-3).

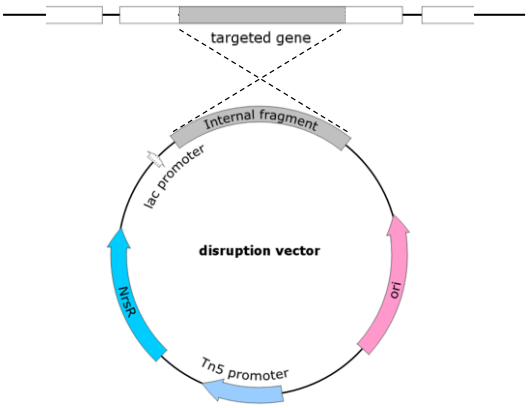


Figure 2-3. Disruption of *hli* biosynthetic genes through single crossover recombination.

For the construction of the general disruption vector pHSG398 Infu Nrs^R (Fig. 2-4), plasmid pHSG398 (Takara Bio. Co., Ltd, Tokyo) was first linearized by inverse PCR using primers pHSG398 Inverse-F/pHSG398 Inverse-R; by infusion cloning (In-Fusion® HD Cloning Kit, Clontech Laboratories, Inc., CA), the original chloramphenicol resistance gene was replaced with the Tn5 promoter (pCR 2.1 TOPO-derived, Life Technologies, U.S.) and a nourseothricin-resistance gene (*Nrs^R*) (amplified from pNR1[22] with primers Kan Nrs-f/Amp Nrs-r). Primers for the amplification of the cassette Tn5-Nrs^R were pHSG NrsR Infu-F and pHSG NrsR Infu-R.

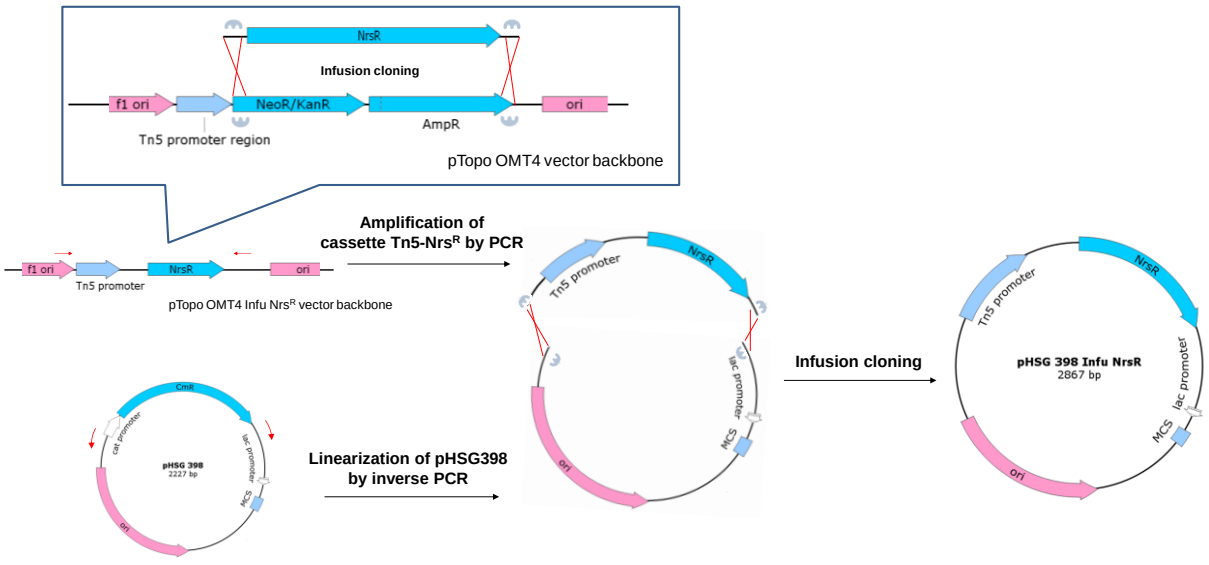


Figure 2-4. Construction of general disruption vector pHSG398 Infu Nrs^R. Nourseothricin resistance gene (*Nrs^R*) with Tn5 promoter was cloned into pHSG398 to replace the original chloramphenicol resistance gene (*Cm^R*).

The internal fragments of the target genes *hliC* (by primers pHSG3ECH Infu F/ pHSG3ECH Infu R), *hliD* (by primers ORF4-MT-f1/ORF4-MT-r1), *hliH* (by primers pHSG8FkbH Infu-f/pHSG8FkbH Infu-r), *hliP* (by primers ORF16-f1/ORF16-r1), *hliQ* (by primers pHSG17MT Infu-f/pHSG17MT Infu-r), *hliR* (by primers pHSG18ACAD Infu-f/pHSG18ACAD Infu-r), *hliU* (by primers pHSG 21Ox Infu F/pHSG 21Ox Infu R) and *orf22* (by primers pHSG22Mscs Infu F/pHSG22Mscs Infu R) were amplified from *H. ochraceum* genomic DNA and inserted into pHSG398 Infu Nrs^R through ligation or infusion cloning (In-Fusion® HD Cloning Kit, Clontech Laboratories, Inc., CA). Each disruption vector (1.0–1.5 μg) was electroporated into the haliangicin-producing host *M. xanthus* c10-11/c7-6E at 25 μF, 1.8 kV, and 200 Ω in 0.2 cm electroporation cuvettes. Transformants were selected on CTT agar plate (nourseothricin: 100 μg/mL) for about 5–7 days and then grown in 1 mL of CTT medium (kanamycin 200 μg/mL, oxytetracycline 12.5 μg/mL, nourseothricin 100 μg/mL) for an additional three days. For the mutants with proper antibiotic resistance, 20 μL of its bacterial culture was incubated at 96 ^{o}C for 15 min and then 1 μL of the suspension was directly subjected to PCR to detect integration of the corresponding plasmid into the genome. The verified mutant strains were grown in CTT medium containing 100 μg/mL nourseothricin for the isolation of fine genomic DNA, by which the disruption of targeted genes at the intended position was verified though PCR as shown in Fig. 2-16b, 2-23b, 2-26b, 2-27b, 2-28b, 2-30b, 2-34b and 2-35b.

2.2.8. Isolation of unnatural haliangicin analogs from gene-disrupted mutants of *M. xanthus*

A gene-disrupted mutant *M. xanthus* Δ*hliD* was cultured in 0.5 L of production medium supplemented with 4% (w/v) SEPABEADS SP207 absorber resin, 0.1% TES and 200 mM sodium acetate at 30 ^{o}C in a similar way to *M. xanthus* c10-11C/c7-6E. The cells and resin were harvested from 6 days culture broth by centrifugation (6000 rpm, 5 min) and extracted with acetone (0.4 L × 2) at 30 ^{o}C for 30 min. After filtration, the combined filtrate were concentrated *in vacuo* and extracted with EtOAc (100 mL × 2, 50 mL × 1). The EtOAc extracts containing 1-*O*-demthylhaliangicin (**7**) and 1-*O*-demthyl-12,13-deoxylhaliangicin (**8**) was used in section **2.2.13** and **2.2.14**.

A gene-disrupted mutant *M. xanthus* Δ*hliR* or *M. xanthus* Δ*hliU* was cultured in 3 L or 1 L of production medium. An extract (646 mg) obtained from the cultures (3 L in total) of *M. xanthus* Δ*hliR* was separated by flash column chromatography [Hi-flash L size (Yamazen, Kyoto, Japan), flow rate 12 mL/min, linear gradient of 0% → 20% EtOAc in hexane in 40 min]. The fraction (99.3 mg) eluted at 27~40 min was

subjected to preparative HPLC [Develosil ODS-UG-5 (ϕ 20 × 250 mm), Nomura Chemical Ltd.; detected at 340 nm, 75% MeOH in water, 10 mL/min] to yield 14,15-dihydrohaliangicin (**9**, 16.4 mg, t_R 42~46 min). An extract (358 mg) from the cultures (1 L in total) of *M. xanthus* Δ*hliU* was purified by flash column chromatography (Hi-flash M size, flow rate 6 mL/min, linear gradient of 0% → 15% EtOAc in hexane in 30 min). The fraction (22.3 mg) eluted at 18~26 min was subjected to preparative HPLC [Develosil ODS-UG-5 (ϕ 20 × 250 mm).; detected at 340 nm, 85% MeOH, 10 mL/min] to yield 12,13-deoxylhaliangicin (**10**, 6.4 mg, t_R 25.5~ 28 min).

14,15-dihydrohaliangicin (9): light yellow oil, $[\alpha]_D^{25}$ + 27.4 (c 0.164, MeOH); UV (MeOH) λ_{max} 246 (ε 11,600) nm; IR (film) ν_{max} 1721, 1614, 1266, 1196, 1143, 1055, 977, 924, 897 and 824 cm^{-1}; HRMS calcd. for $C_{22}H_{34}O_5Na$: 401.2299; found *m/z* 401.2298 $[M+Na]^+$. For NMR data, see Table 2-19 and Chapter 5 Appendix **5.3.**

12,13-deoxylhaliangicin (10): light yellow oil, $[\alpha]_D^{30}$ + 43.1 (c 0.13, MeOH); UV (MeOH) λ_{max} 236 (ε 15,500) nm; IR (film) ν_{max} 1721, 1672, 1649, 1614, 1265, 1197, 1143, 1054, 977, 924, 827 and 755 cm^{-1}; HRMS calcd. for $C_{22}H_{32}O_4Na$ 383.2193; found *m/z* 383.2206 $[M+Na]^+$. For NMR spectroscopic data, see Table 2-20 and Chapter 5 Appendix **5.4.**

2.2.9. Cloning and expression of HliR (acyl-CoA dehydrogenase) and HliD (*O*-methyltransferase)

HliR and HliD were amplified from the *H. ochraceum* genome by PCR using high-fidelity PrimeSTAR Max DNA polymerase (Takara Bio Inc.), using the primers HliR Exp Infu f/HliR Exp Infu r and HliD Exp Infu f/HliD Exp Infu r, respectively. Subsequently, the amplified fragments were cloned into *BamH*I/*Xho*I double-digested pET32a by using an In-Fusion® HD Cloning Kit (Clontech Laboratories, Inc., Mountain View, CA). The purified plasmids were transformed into Single Step (KRX) Competent Cells (Promega, Madison, WI). For the expression of recombinant proteins, when the OD_{600} reached 0.6, a final concentration of 0.1% rhamnose and 0.4 mM IPTG were added to induce protein expression. After induction for additional 20 h at 20 °C, the cells were harvested by centrifugation (6000 rpm, 5 min) and resuspended in 3 mL of extraction buffer (50 mM Tris-HCl, 0.4 M NaCl, pH 7.8). The resulting suspension was lysed by using a sonication homogenizer (50 W, five cycles) in the presence of benzonase nuclease (Novagen, Madison, WI) over ice for 30 min. The lysate was centrifuged (6000 rpm, 5 min) to remove insoluble cell debris. The crude protein extracts were stored at –30 °C for subsequent purification. A MagneHis™ Protein Purification System (Promega, Madison, WI) was used to purify the His_6-tagged

protein. The crude extract and purified His_6-tagged proteins were analysed by SDS-PAGE (10% Tris-HCl gel) (Fig. 2-17 and 2-33a).

2.2.10. *In vitro* enzymatic characterization of HliR

A typical HliR assay contained 500 μM flavin adenine dinucleotide (FAD), 500 μM synthesized substrate, and 1 μM recombinant HliR, in a pH 7.5 buffer (8 mM $MgSO_4$, 10 mM Tris-HCl, and 1 mM potassium phosphate; total volume: 100 μL). The assay mixture was incubated at 30 °C for 4 h ~ overnight and then quenched and extracted with EtOAc. The extracts were analysed by HPLC under the conditions indicated in Fig. 2-18, 2-20b, and 2-21.

2.2.11. Chemical synthesis of *S*-Acyl-*N*-acetylcysteamine (SNAC) derivatives

To a solution of 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (1.5 eq) and 4-dimethylaminopyridine (0.1 eq) in anhydrous dichloromethane was added the appropriate acids (1.2 ~ 1.5 eq) and *N*-acetylcysteamine (1 eq). The reaction mixture was stirred under nitrogen at room temperature overnight, then diluted with dichloromethane and washed with H_2O. The organic layer was dried, concentrated *in vacuo*, and purified by silica gel chromatography to afford the pure thioester products.

(*E*)-2-methyl-2-pentenoyl SNAC (11): colorless oil, ^{1}H NMR ($CDCl_3$, 400 MHz) δ 6.73 (qt, *J* = 1.0, 7.3 Hz, 1H), 6.07 (brs, 1H), 3.42 (q, *J* = 6.1 Hz, 2H), 3.04 (t, *J* = 6.4 Hz, 2H), 2.23 (m, 2H), 1.95 (s, 3H), 1.85 (d, *J* = 1.0 Hz, 3H), 1.08 (t, *J* = 7.6 Hz, 3H). ESI-MS *m/z* 216.1 $[M+H]^+$, 238.1 $[M+Na]^+$.

(*E*)-2-pentenoyl SNAC (12): colorless oil, ^{1}H NMR ($CDCl_3$, 400 MHz): δ 6.96 (dt, *J* =15.6, 6.4 Hz, 1H), 6.10 (d, *J* = 15.6 Hz, 1H), 6.08 (br, 1H), 3.42 (q, *J* = 6.0 Hz, 2H), 3.07 (t, *J* = 6.4 Hz, 2H), 2.22 (m, 1H), 1.94 (s, 3H), 1.06 (t, *J* = 7.2 Hz, 3H).

4-methyl-2-pentenoyl SNAC (13): colorless oil, ^{1}H NMR ($CDCl_3$, 400 MHz): δ 6.87 (dd, *J* =15.6, 6.8 Hz, 1H), 6.16 (br, 1H), 6.04 (d, *J* = 15.6 Hz, 1H), 3.42 (q, *J* = 6.4 Hz, 2H), 3.06 (t, *J* = 6.4 Hz, 2H), 2.44 (m, 1H), 1.94 (s, 3H), 1.05 (d, *J* = 6.8 Hz, 6H).

2-hexenoyl SNAC (14): colorless oil, ^{1}H NMR ($CDCl_3$, 400 MHz): δ 6.89 (dt, *J* =15.6, 7.2 Hz, 1H), 6.16 (br, 1H), 6.10 (d, *J* = 15.2 Hz), 3.43 (q, *J* = 6.0 Hz, 2H), 3.06 (t, *J* = 6.4 Hz, 2H), 2.16 (q, *J* = 7.2, 2H), 1.94 (s, 3H), 1.47 (m, 2H), 0.91 (t, *J* = 7.2 Hz, 3H).

propionyl SNAC (15): colorless oil, ^{1}H NMR ($CDCl_3$, 400 MHz) δ 5.99 (br, 1H), 3.41 (q, *J* = 6.2 Hz, 2H), 3.01 (t, *J* = 6.5 Hz, 2H), 2.58 (q, *J* =7.5 Hz, 2H), 1.95 (s, 3H), 1.17 (t, *J* = 7.4 Hz, 3H).

acryloyl SNAC (16): colorless oil, ^{1}H NMR ($CDCl_3$, 400 MHz) δ 6.34 (m, 2H), 6.01 (br, 1H), 5.72 (dd, *J* = 9.6, 1.6 Hz, 1H), 3.46 (q, *J* = 6.2 Hz, 2H), 3.12 (t, *J* = 6.4 Hz, 2H), 1.96 (s, 3H).

butyryl SNAC (17): colorless oil, ^{1}H NMR ($CDCl_3$, 400 MHz): δ 6.18 (br, 1H), 3.38 (q, *J* = 6.4 Hz, 2H), 2.98 (t, *J* = 6.4 Hz, 2H), 2.51 (t, *J* = 7.6 Hz, 2H), 1.92 (s, 3H), 1.65 (m, 2H), 0.91 (t, *J* =7.6 Hz, 3H).

valeryl SNAC (18): colorless oil, ^{1}H NMR ($CDCl_3$, 400 MHz): δ 6.02 (br, 1H), 3.40 (q, *J* = 6.0 Hz, 2H), 3.00 (t, *J* = 6.4 Hz, 2H), 2.55 (t, *J* = 7.6 Hz, 2H), 1.94 (s, 3H), 1.62 (m, 2H), 1.33 (m, 2H), 0.89 (t, *J* = 7.2 Hz, 3H).

hexanoyl SNAC (19): colorless oil, ^{1}H NMR ($CDCl_3$, 400 MHz): δ 6.17 (br, 1H), 3.39 (q, *J* = 6.0 Hz, 2H), 2.98 (t, *J* = 6.4 Hz, 2H), 2.53 (t, *J* = 7.6 Hz, 2H), 1.93 (s, 3H), 1.62 (m, 2H), 1.27 (m, 4H), 0.85 (t, *J* = 6.8 Hz, 3H).

heptanoyl SNAC (20): colorless oil, ^{1}H NMR ($CDCl_3$, 400 MHz): δ 6.29 (br, 1H), 3.36 (q, *J* = 6.4 Hz, 2H), 2.97 (t, *J* = 6.4 Hz, 2H), 2.51 (t, *J* = 7.2 Hz, 2H), 1.92 (s, 3H), 1.59 (m, 2H), 1.25 (m, 6H), 0.82 (t, *J* = 6.8 Hz, 3H).

2-methylvaleryl SNAC (21): colorless oil, ^{1}H NMR ($CDCl_3$, 400 MHz): δ 5.90 (br, 1H), 3.42 (q, *J* = 6.0 Hz, 2H), 3.00 (t, *J* = 6.4 Hz, 2H), 2.66 (m, 1H), 1.95 (s, 3H), 1.68 (m, 1H), 1.37 (m, 3H), 1.16 (d, *J* = 6.8 Hz, 3H), 0.89 (t, *J* = 7.2 Hz, 3H).

4-methylvaleryl SNAC (22): colorless oil, ^{1}H NMR ($CDCl_3$, 400 MHz): δ 6.17 (br, 1H), 3.38 (q, *J* = 6.4 Hz, 2H), 2.98 (t, *J* = 6.4 Hz, 2H), 2.53 (t, *J* = 7.6 Hz, 2H), 1.93 (s, 3H), 1.52 (m, 3H), 0.86 (d, *J* = 6.4 Hz, 6H).

2-methyl-4-pentenoyl SNAC (23): Colorless oil, ^{1}H NMR ($CDCl_3$, 400 MHz): δ 5.91 (br, 1H), 5.70 (m, 1H), 5.05 (m, 2H), 3.41 (q, *J* = 6.4 Hz, 2H), 3.01 (t, *J* = 6.4 Hz, 2H), 2.73 (q, *J* = 6.8 Hz, 1H), 2.44 (m, 1H), 2.17 (m, 1H), 1.94 (s, 3H), 1.17 (d, *J* =6.8 Hz, 3H)

4-pentenoyl SNAC (24): Colorless oil, ^{1}H NMR ($CDCl_3$, 400 MHz): δ 6.09 (br, 1H), 5.76 (m, 1H), 5.02 (m, 2H), 3.39 (q, *J* = 6.0 Hz, 2H), 3.00 (t, *J* = 6.4 Hz, 2H), 2.65 (t, *J* = 7.6 Hz, 2H), 2.38 (dt, *J* =7.2, 6.6 Hz, 2H), 1.94 (s, 3H)

cyclopent-1-enoyl SNAC (25): white needles, ^{1}H NMR ($CDCl_3$, 400 MHz): δ 6.81 (m, 1H), 6.24 (br, 1H), 3.40 (q, *J* = 6.4 Hz, 2H), 3.04 (t, *J* = 6.4 Hz, 2H), 2.53 (m, 4H), 1.94 (m, 2H), 1.93 (s, 3H).

cyclohex-1-enoyl SNAC (26): white needles, ^{1}H NMR ($CDCl_3$, 400 MHz): δ 6.98 (m, 1H), 6.11 (br, 1H), 3.41 (q, *J* = 6.4 Hz, 2H), 3.03 (t, *J* = 6.4 Hz, 2H), 2.23 (m, 4H), 1.93 (s, 3H) , 1.61 (m, 4H).

2.2.12. Bioconversion of (*E*)-2-methylpent-2-enoyl SNAC (11) and (*E*)-2-pentenoyl SNAC (12)

The HliR-overexpressing host *E. coli* KRX pET32a-HliR was cultured in five 50 mL of LB mediums for 2 days. To each 50 mL culture, (*E*)-2-methylpent-2-enoyl SNAC (**11**, 2.7 mg in 125 μL EtOH) or (*E*)-2-pentenoyl SNAC (**12**, 5 mg in 125 μL EtOH) was added twice at 6 h and 16 h after the induction was started. When the second portion was added, the cultures were transferred to 30 ^{o}C and incubated for additional 24 h. The cultures were centrifuged (6000 rpm, 5 min) and cell pellets were extracted with methanol (30 mL × 2) and the supernatants were extracted with EtOAc (50 mL × 3). The methanol extracts and EtOAc layer were combined and concentrated. The resulting residue was further subjected to silica gel column chromatography [5 g of silica gel, eluted with hexane/EtOAc (1/1, v/v)] to yield **27** (2-methylpent-2,4-dienoyl SNAC, 2.6 mg): colorless oil; UV (MeOH) λ_{max} 257 (ε 17,000), 205 (ε 11,000), 282 (sh, ε 12,000)nm; IR (film) ν_{max} 3289, 2922, 1654, 1542, 1457, 1376, 1288, 983 cm^{-1}; ^{1}H NMR ($CDCl_3$, 400 MHz) δ 7.15 (d, *J* = 11.2 Hz, 1H), 6.68 (ddd, *J* =10.0, 10.8, 16.4 Hz, 1H), 5.88 (brs, 1H), 5.65 (d, *J* = 16.4 Hz, 1H), 5.55 (d, *J* =10.0 Hz, 1H), 3.47 (q, *J* = 6.4 Hz, 2H), 3.10 (t, *J* = 6.4 Hz, 2H), 2.00 (d, *J* = 1.2 Hz, 3H), 1.97 (s, 3H); ^{13}C NMR ($CDCl_3$, 100 MHz): δ 193.9, 170.3, 137.5, 135.3, 131.8, 125.8, 39.8, 28.5, 23.3, 12.7; HRMS calcd. for $C_{10}H_{15}NO_2NaS$ 236.0716; found *m/z* 236.0720 $[M+Na]^+$; **28** (pent-2,4-dienoyl SNAC, 4.1 mg): colorless oil; UV (MeOH) λ_{max} 262 (ε 22,000), 204 (ε 9,600) nm; IR (film) ν_{max} 3288, 3081, 2930, 1656, 1591, 1550, 1289, 1254, 1181, 1121, 1032, 811 cm^{-1}; ^{1}H NMR ($CDCl_3$, 400 MHz): δ 7.21 (dd, *J* = 15.2, 10.8 Hz, 1H), 6.43 (ddd, *J* =16.8, 10.4, 10.4 Hz, 1H), 6.20 (d, *J* = 15.2 Hz, 1H), 5.92 (br, 1H), 5.70 (d, *J* = 17.2 Hz, 1H), 5.60 (d, *J* =9.6 Hz, 1H), 3.47 (dt, *J* = 6.4, 6.0 Hz, 2H), 3.12 (t, *J* = 6.4 Hz, 2H), 1.96 (s, 3H); ^{13}C NMR ($CDCl_3$, 100 MHz): δ 190.4, 170.3, 141.3, 134.4, 128.5, 127.6, 39.8, 28.4, 23.2; HRMS calcd. for $C_9H_{13}NO_2NaS$ 222.0559; found *m/z* 222.0563 $[M+Na]^+$.

2.2.13. Methyl esterification of *M. xanthus* Δ*hliD* extracts.

An ethyl acetate extract (121 mg) obtained from 500 mL of *M. xanthus* Δ*hliD* cultures was dissolved in Et_2O (5 mL) and 400 μL of trimethylsilyldiazomethane solution (2.0 M in Et_2O) was added. After stirring at room temperature for 30 min, the reaction mixture was dried under nitrogen flash and the residue was analyzed by HPLC (Develosil ODS-UG-5, ϕ 4.6 × 250 mm, 1 mL/min, detected at 290 nm, 20 → 100% methanol in H_2O in 40 min) (Fig. 2-32). The mixture was chromatographed by the same method as that for haliangicin (**1**) and 12,13-deoxylhaliangicin (**10**) to yield haliangicin (**1**, 1.0 mg) and 12,13-deoxylhaliangicin (**10**, 0.7 mg).

2.2.14. Bioconversion of the demethylated precursors of haliangicin

An EtOAc extract obtained from 500 mL of *M. xanthus* Δ*hliD* cultures was subjected to RP silica gel open column chromatography (4g of Cosmosil 75C18-OPN, eluted with 40%~100% MeOH). The 50% and 60% MeOH fractions were combined and dried to dryness *in vacuo*. The HliD-expressing host *E. coli* KRX pET32a-HliD was cultured in 50 mL of LB mediums, to which 50 mL broth of the abovementioned crude extracts containing 1-*O*-demthylhaliangicin (**7**) and 1-*O*-demthyl-12,13-deoxylhaliangicin (**8**) in 500 μL DMSO were added at 6 h after the induction was started. After the incubation at 20 ^{o}C for a total of 21 h, the cultures were centrifuged (6000 rpm, 5 min) and cell pellets were extracted with methanol (30 mL × 2) and the supernatants were extracted with EtOAc (50 mL × 3). The methanol extracts and EtOAc layer were combined and concentrated. The resulting residue was analyzed by HPLC (Develosil ODS-UG-5, ϕ 4.6 × 250 mm, 1 mL/min, detected at 290 nm, 20 → 100% acetonitrile in H_2O with 0.1% TFA in 40 min). (Fig. 2-33b).

2.2.15. Bioassay methods

2.2.15.1. Anti-oomycete activity

The phytopathogenic oomycete *Phytophthora capsici* NBRC 30696, purchased from NITE-Biological Resource Center (NBRC, Chiba, Japan), was cultured on a potato-agar medium [in 1 L, potato broth from 200 g of fresh potato, glucose (20 g), and agar (20 g)] in a 9-cm dish at 25 ^{o}C for 7 days in the dark. A piece of the colony was then inoculated on the center of a 5% V8 -agar medium [V8 vegetable juice (5 mL), agar (1.5 g), and water (95 mL)] in a 9-cm dish and incubated at 25 ^{o}C for 48 h in the dark until the colony grew to approximately 3–4 cm in diameter. A paper disc (6 mm in diameter) impregnated with a sample was placed 1 cm away from the front of the colony. After incubating for 22~24 h, the distance between the edge of the colony and the paper disc (control: 0 mm) was measured. The activity was defined as a minimum dose that induced a definite distance (0.5 mm or wider) between the colony and the disk. The results were summarized in Fig. 2-36a.

2.2.15.2. MIC assay

The microorganisms, *Candida rugosa* AJ 14513, *Bacillus subtilis* AJ 12865, and *Escherichia coli* AJ 3837, were provided by the Institute for Innovation, Ajinomoto Co., Inc. (Kanagawa, Japan). The protocols used for the assay are in the followings. For *B. subtilis* and *E. coli*, see *Methods for Dilution Antimicrobial*

Susceptibility Tests for Bacteria That Grow Aerobically; Approved Standard—Ninth Edition. CLSI document M07-A9 [ISBN 1-56238-783-9], Clinical and Laboratory Standards Institute (Wayne, PA, USA), 2012. For *C. rugosa*, see *Reference Method for Broth Dilution Antifungal Susceptibility Testing of Yeasts; Approved Standard*—Second Edition. NCCLS document M27-A2 [ISBN 1-56238-469-4], NCCLS (Wayne, PA, USA), 2002.

2.2.15.3. Cytotoxicity assay

HeLa-S3 (SC) cells were provided by the RIKEN BRC through the National Bio-Resource Project of the MEXT, Japan. The cells were cultured in Eagle's minimal essential medium (EMEM) (Wako Pure Chemical Industries, Osaka, Japan) supplemented with 10% bovine serum, 100 units/mL penicillin, and 100 μg/mL streptomycin (Thermo Fisher Scientific Inc., MA, USA). A total of 10,000 cultured cells were seeded into each well of a 96-well plate containing 99 μL of the same medium. After pre-incubation for 24 h at 37 ^{o}C in an atmosphere of 5% CO_2, a compound in 1 μL of dimethyl sulfoxide (DMSO) was added to each well, and the cells were incubated an additional 48 h. A solution (10 μL) of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) in phosphate buffer saline (PBS) (5 mg/mL) was then added to each well, and the plate was incubated for an additional 3 h. Subsequently, the medium was removed by aspiration, any generated formazan was dissolved in 100 μL of DMSO, and the absorbance was measured at 595 nm using a Multiskan FC microplate reader (Thermo Fisher Scientific). The results were summarized in Fig. 2-36b.

2.3. Results and discussion

2.3.1. Heterologous production profile of haliangicins in *M. xanthus*

To investigate the heterologous production profile of haliangicin and its stereoisomers, a large-scale cultivation of *M. xanthus* c10-11C/c7-6E was carried out. Extraction, chromatographic separation followed by NMR analysis confirmed the existence of haliangicin (**1**) and its stereoisomers, haliangicin B (**4**), C (**5**) and D (**6**) (Fig. 2-5) in the extracts of the heterologous host. This production profile is highly similar to that of *H. ochraceum*, the native producer of haliangicin (**1**). From these results, it can be concluded that the *hli* biosynthetic assembly line is successfully expressed in the heterologous host with full functions of PKSs and all tailoring modification enzymes.

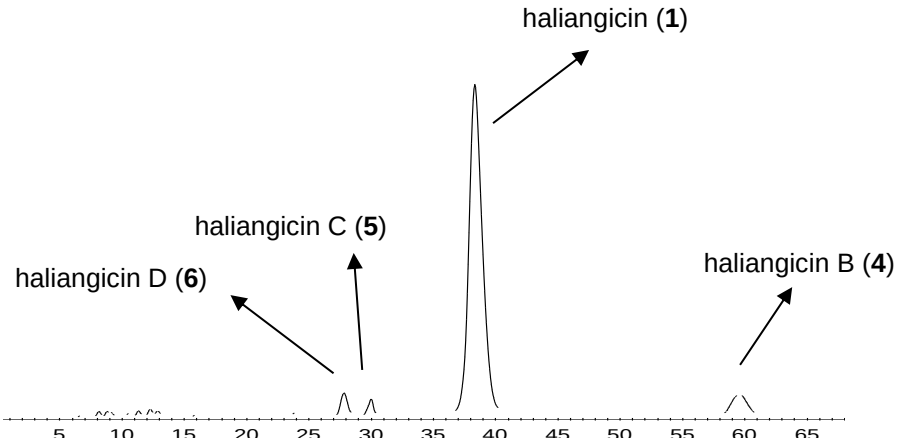


Figure 2-5. A preparative HPLC chart of haliangicin-containing fraction from *M. xanthus* c10-11C/c7-6E.

2.3.2. Identification of the biosynthetic precursor of C-3 and C-4 in haliangicin

According the previous works, the biosynthetic precursors of the most parts of haliangicin have been clarified by feeding experiments, showing two intact acetate units, one acetate-derived methyl branch (9-Me), four propionate units and three *S*-adenosylmethionine (SAM)-originated *O*-methyl carbons (Fig. 2-6). However, the origin of C-3 and C-4 remains unclear because neither acetate nor propionate-enhanced signals were observed in earlier studies. This finding indicates that an unusual biosynthetic precursor may be involved in the chain elongation of C-3 and C-4.

The C_2 unit of C-3 and C-4 bearing vicinal oxygen atoms (referred as a "glycolate" extender) deviates from polyketide rude. There are several similar examples of this rare "glycolate" extension unit in natural products, such as ansamitocin P-3,[28, 29] soraphen A,[30, 31] FK-520,[32] and aflastatin.[33] Earlier studies on these

compounds provided two possible precursors of this unique unit. In the case of aflastatin, glycolic acid was reported to be a biosynthetic origin in spite of its low incorporation,[33] while the other cases revealed glycerate as a key intermediate. After testing both candidates, we excluded the former possibility because the addition of synthesized [1-^{13}C]glycolic acid showed no isotope enrichment over the natural abundance (Fig. 2-7). Subsequently, the later candidate [U-$^{13}C_3$]glycerol labeled C-3 and C-4 as doublets due to direct ^{13}C-^{13}C coupling, accompanied by additional enhancements of acetate-derived carbon signals of C-1~2, C-7~8 and 9-Me (Fig. 2-8). Such labeling pattern might be explained by the metabolism of glycerate via glycolysis that finally reach acetyl-CoA as described in other "glycolate" unit-containing compounds.

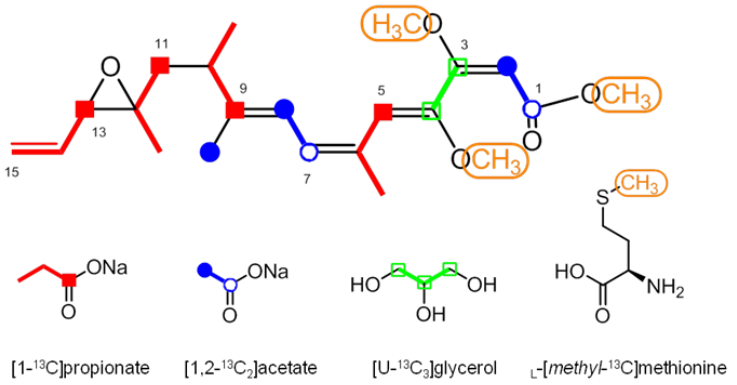


Figure 2-6. Biosynthetic building blocks of haliangicin (**1**) deduced from feeding experiments using labeled precursors.

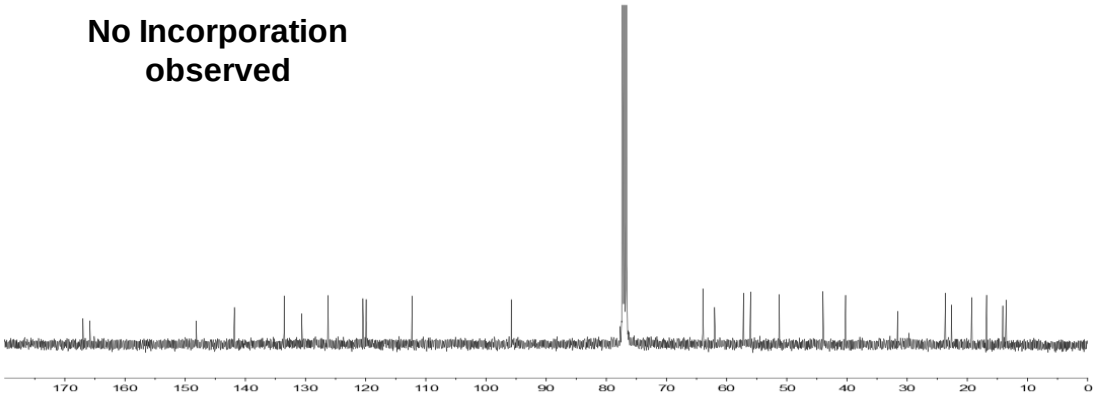


Figure 2-7. ^{13}C NMR spectrum (100 MHz) of [1-^{13}C]glycolic acid-feeding haliangicin in $CDCl_3$.

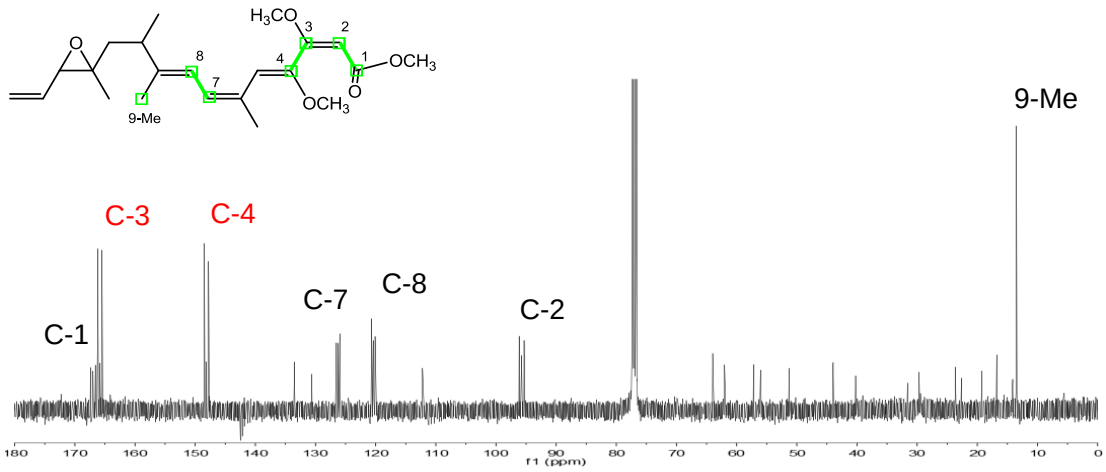


Figure 2-8. ^{13}C NMR spectrum (100 MHz) of [U-$^{13}C_3$]glycerol-labeled haliangicin in $CDCl_3$.

2.3.3. Optimization of haliangicin productivity in *M. xanthus*

Acting as the carbon sources, all the four biosynthetic precursors (sodium acetate, sodium propionate, $_L$-methionine and glycerol) were tested for their effect on haliangicin productivity (Fig. 2-9). The haliangicin titer increased significantly when sodium acetate was added at a concentration higher than 50 mM and finally reached 11.0 ± 2.1 mg L^{-1} at 200 mM of sodium acetate. The efficiency of haliangicin production was 10-times higher in quantity and 3-times shorter in culture period than that achieved with the native producer *H. ochraceum*.

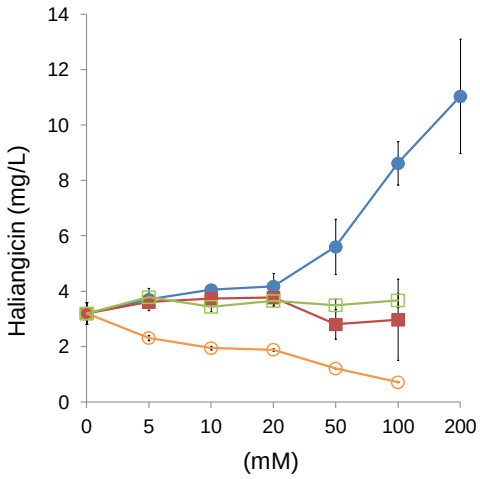


Figure 2-9. Optimization of the production of haliangicin (**1**) by feeding the appropriate biosynthetic precursors (red: sodium propionate; blue: sodium acetate; green: glycerol; orange: $_L$-methionine).

I also investigated another carbon source, glucose, because it may link to various biosynthetic building blocks through tricarboxylic acid (TCA) cycle. 10% Glucose boosted haliangicin productivity to 10.1 ± 2.4 mg/L that almost comparable to the effect of sodium acetate (Fig. 2-10), though it took a relatively longer time (> 10 days), possibly due to the time or energy consuming processes for glycolysis.

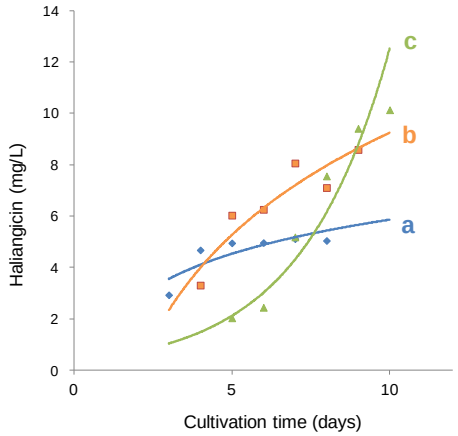


Figure 2-10. Time course of haliangicin productivity in production medium supplemented with glucose.

(**a**) 6% glucose in 4 ~ 8 days.

(**b**) 8% glucose in 5 ~ 9 days.

(**c**) 10% glucose in 6 ~ 10 days.

2.3.4. Analysis of haliangicin biosynthetic gene cluster (*hli*)

The putative haliangicin biosynthetic gene cluster (*hli*) spans 48 kb (GC content of 67.2%), containing 21 open reading frames (ORFs) (Fig. 2-11). All the proposed functions of the proteins encoded by these genes were summarized in Table 2-4.

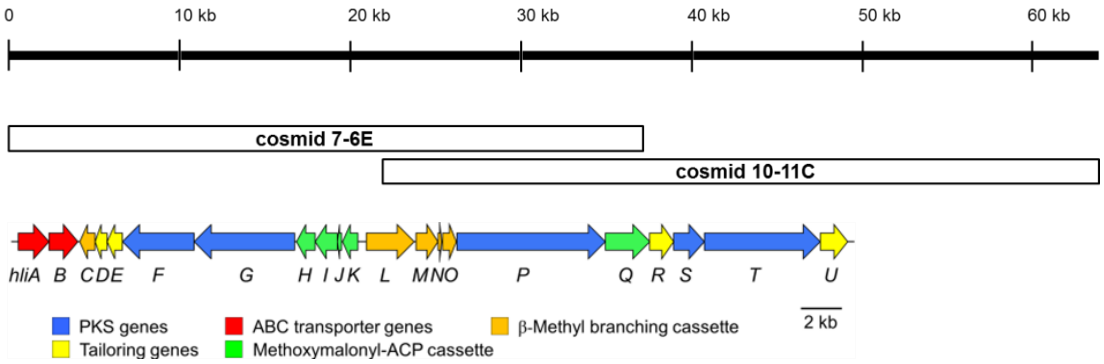


Figure 2-11. Genetic organization of the haliangicin biosynthetic gene cluster (*hli*).

According to the biosynthetic functions, genes in *hli* locus could be divided into several groups (Fig. 2-11). Five type I PKS genes (*hliF, G, P, S, T*) encoding 6 modules are predicted to be responsible for the biosynthesis of haliangicin skeleton (section **2.3.6** and **2.3.8**). Annotation of domains in these PKSs is summarized in Table 2-5. The β-methyl branching cassette incorporating 9-Me consists of five genes, *hliL, M, N, O,* and *C* (Table 2-6 and section **2.3.9**). Genes involved in methoxymalonyl-ACP biosynthesis include *hliQ, H, I, J* and *K* (Table 2-7 and section **2.3.10**). Other proteins were identified to be an acyl-CoA dehydrogenase (ACAD) HliR involved in the starter dehydrogenation (section **2.3.7**) and several post-assembly tailoring enzymes including a standalone *O*-methyltransferase (HliD, section **2.3.12**), a metallo-β-lactamase type thioesterase (HliE, section **2.3.11**) and an epoxidase (HliU, section **2.3.12**). A gene (*orf22*, section **2.3.13**) located downstream to *hliU* shows a similarity to mechanosensitive ion channel protein-encoding genes, which was thought to be unrelated to the haliangicin biosynthesis.

Unlike typical type-I PKS biosynthetic architectures, the *hli* locus displays several unusual characteristics. First, the gene arrangement is rather disorganized. Not only the PKS genes, but also genes of β-methyl branching cassette and methoxymalonyl-ACP cassette are also not tightly clustered together (Fig. 2-11). These aberrant architectures might be a reflection that the evolutionary integration of haliangicin biosynthetic genes is not fully completed. Second, the number of domains in *hli* appears to be deficient. From a typical biosynthetic view point, haliangicin biosynthesis requires at least six PKS elongation steps following a starter loading step to generate the heptaketide backbone. Therefore, six KS domains in the elongation modules are needed but only five can be found (section **2.3.4.1**). The number of AT domains in *hli* also shows a discrepancy to the haliangicin structure (section **2.3.4.2**). There is an apparently obsolete AT in HliP, indicating an unusual AT-deficient feature that might link to β-methyl branching biosynthesis. Besides, the seven integrated ACPs in PKSs and two free-standing ACPs (HliJ, HliN) all possess the conserved serine residues for phosphopantetheine binding (section **2.3.4.6**) but still seem to be unable to fit exactly the observed structure.

Table 2-4. Proposed functions of the proteins involved in the haliangicin biosynthesis.

Protein	Length (aa)	Proposed function	Top-hit homologues (Origin)	Similarity/ identity (%)	Accession No.
HliA	571	ABC transporter	ABC transporter protein (*Sorangium cellulosum*)	58/34	CCE88383
HliB	609	ABC transporter	ABC transporter protein (*Sorangium cellulosum*)	44/24	CCE88383
HliC	256	Enoyl-CoA hydratase	enoyl-CoA hydratase (*Nostoc punctiforme* PCC 73102)	74/53	YP_001866568
HliD	217	*O*-methyltransferase	JerF (*Sorangium cellulosum*)	62/45	ABK32292
HliE	303	Metallo-β-lactamase type TE	β-lactamase (*Rhodothermus marinus* SG0.5JP17-172)	67/49	YP_004824011
HliF	1372	PKS (KS-AT-*O*-MT-ACP)	MtaF (*Stigmatella aurantiaca* DW4/3-1)	62/45	YP_003953631
HliG	1966	PKS (KS-AT-DH-KR-ACP)	MxaD (*Stigmatella aurantiaca*)	57/42	AAK57188
HliH	367	FkbH-like protein	HAD superfamily phosphatase (*Sorangium cellulosum*)	78/62	CCE88385
HliI	411	Acyl-CoA dehydrogenase	acyl-CoA dehydrogenase I (*Sorangium cellulosum*)	73/57	CCE88384
HliJ	83	ACP	hypothetical protein (*Paenibacillus dendritiformis*)	76/57	WP_006675272
HliK	287	3-hydroxyacyl-CoA dehydrogenase	hydroxyacyl-CoA dehydrogenase (*Sorangium cellulosum*)	79/66	CCE88386
HliL	950	PKS (KS-AT)	polyketide synthase (*Chondromyces crocatus*)	61/48	CBD77738
HliM	407	HMG-CoA synthase	HMG-CoA synthase (*Nostoc punctiforme* PCC 73102)	65/48	YP_001866566
HliN	83	ACP	acyl carrier protein (*Nostoc* sp. 'Peltigera membranacea cyanobiont')	81/56	AGH69805
HliO	277	Enoyl-CoA hydratase	enoyl-CoA hydratase/isomerase (*Nostoc punctiforme* PCC 73102)	58/42	YP_001866567
HliP	2890	PKS (KS-ACP-KS-AT-DH-KR-ACP)	polyketide synthase (*Myxococcus xanthus* DK 1622)	55/42	YP_632696
HliQ	861	AT-ACP-*O*-MT	putative methoxymalonyl-CoA synthase (*Sorangium cellulosum*)	59/44	AAK19884
HliR	392	Acyl-CoA dehydrogenase	acyl-CoA dehydrogenase (*Herbidospora cretacea*)	69/55	WP_030450088
HliS	626	AT-ACP	MxaF (*Stigmatella aurantiaca*)	57/43	AAK57190
HliT	2225	PKS (KS-AT-DH-ER-KR-ACP)	polyketide synthase (*Sorangium cellulosum*)	61/44	CCE88378
HliU	463	Epoxidase	FAD-binding monooxygenase (*Geodermatophilus obscurus* DSM 43160)	56/42	YP_003408356

Table 2-5. PKS in *hli* for the biosynthesis of haliangicin skeleton.

Gene	Protein	Length (aa)	Type	Domain position
hliF	HliF	1372	Type I PKS	KS (36-458), AT (552-847), *O*-MT (942-1195), ACP (1280-1349)
hliG	HliG	1966	Type I PKS	KS (37-460), AT (594-896), DH (979-1145), KR (1544-1722), ACP (1834-1899)
hliP	HliP	2890	Type I PKS	KS (12-449), ACP (850-923), KS (953-1380), AT (1498-1804), DH (1874-2042), KR (2431-2623), ACP (2733-2803)
hliS	HliS	626	Type I PKS	AT (98-412), ACP (502-571)
hliT	HliT	2225	Type I PKS	KS (36-459), AT (564-853), DH (927-1092), ER (1501-1808), KR (1832-2010), ACP (2111-2181)

Table 2-6. β-methyl branching cassette in *hli* and its homologies.

Product	KS^S	ACP	HCS	ECH	ECH	AT	Tandem ACP	Identity of HCS to HliM
haliangicin	HliL*[1]	HliN	HliM	HliO	HliC	HliL	-	100%
crocacin[34]	CroD*[1]	CroH	CroE	CroF	CroG	CroD	-	50%
virginiamycin M[35]	VirB	*[2]	VirC	VirD	VirE	VirI	VirA	37%
myxovirescin[36]	TaK	TaB TaE	TaC TaF	TaX	TaY	TaV	-	34% 35%
corallopyronin A[37]	CorD	CorC	CorE	CorF	CorG	CorA	CorJ/K/L	34%
mupirocin[38]	MupG	Macp15	MupH	MupJ	MupK	MmpIII	MmpI	34%
batumin[39]	BatB	BatA	BatC	BatD	BatE	BatH	Bat2*[3]	34%

*1 KS domains does not feature the cysteine-to-serine mutation in typical KS^S

*2 An ACP can be found between VirA and VirB but did not present in the original report

*3 Tandem ACPs present in triplets

Table 2-7. Methoxymalonyl-ACP biosynthesis-involving proteins in *hli* and their homologies.

	HliQ	HliH	HliI	HliJ	HliK
Proposed function	AT-ACP-*O*-MT	FkbH-like protein	Acyl-CoA dehydrogenase	ACP	3-hydroxyacyl-CoA dehydrogenase
Identity to *pel* biosynthesis[40]	PelG (48%)	PeK (64%)	PelJ (57%)	PelI (54%)	PelH (66%)
Identity to *sor* biosynthesis[31]	SorC (44%)	-	SorE (55%)	-	SorD (59%)
Identity to *fkb* biosynthesis[32]	FkbG (29%)	FkbH (47%)	FkbI (35%)	FkbJ (52%)	FkbK (38%)

Accession: PelG (CCE88382.1), PelK (CCE88385.1), PelJ (CCE88384.1), PelI (CCE88387.1), PelH (CCE88386.1), SorC (AAK19884.1), SorE (AAK19892.1), SorD (AAK19885.1), FkbG (AAF86386.1), FkbH (AAF86387.1), FkbI (AAF86388.1), FkbJ (AAF86389.1), FkbK (AAF86390.1)

2.3.4.1. KS domains

The sequence of whole *hli* was analyzed by Natural Product Domain Seeker (NaPDoS,[41] http://npdomainseeker.ucsd.edu/) to reveal six ketosynthase (KS) domains encoded in five PKSs. All the KS domains were classified as *cis*-AT modular types (Fig. 2-12).

The results of phylogenetic analysis were also supported by multiple alignments of amino sequences of the KS domains. All *hli* KS domains possess CHH in triad active residues[42] required for both decarboxylation and C-C condensation (Table 2-8). Lack of a crucial β-keto synthase (KS^S) responsible for decarboxylation in β-methyl branching cassette is unexpected. HliL_KS was assumed to be an alternative for the sake of its close location to other β-methyl branching genes and its corresponding malonate-specific AT. HliL is highly similar to a bifunctional protein CroD in *Chondromyces crocatus*,[34] in which the KS domain also does not feature the Cys-to-Ser mutation (Table 2-9).

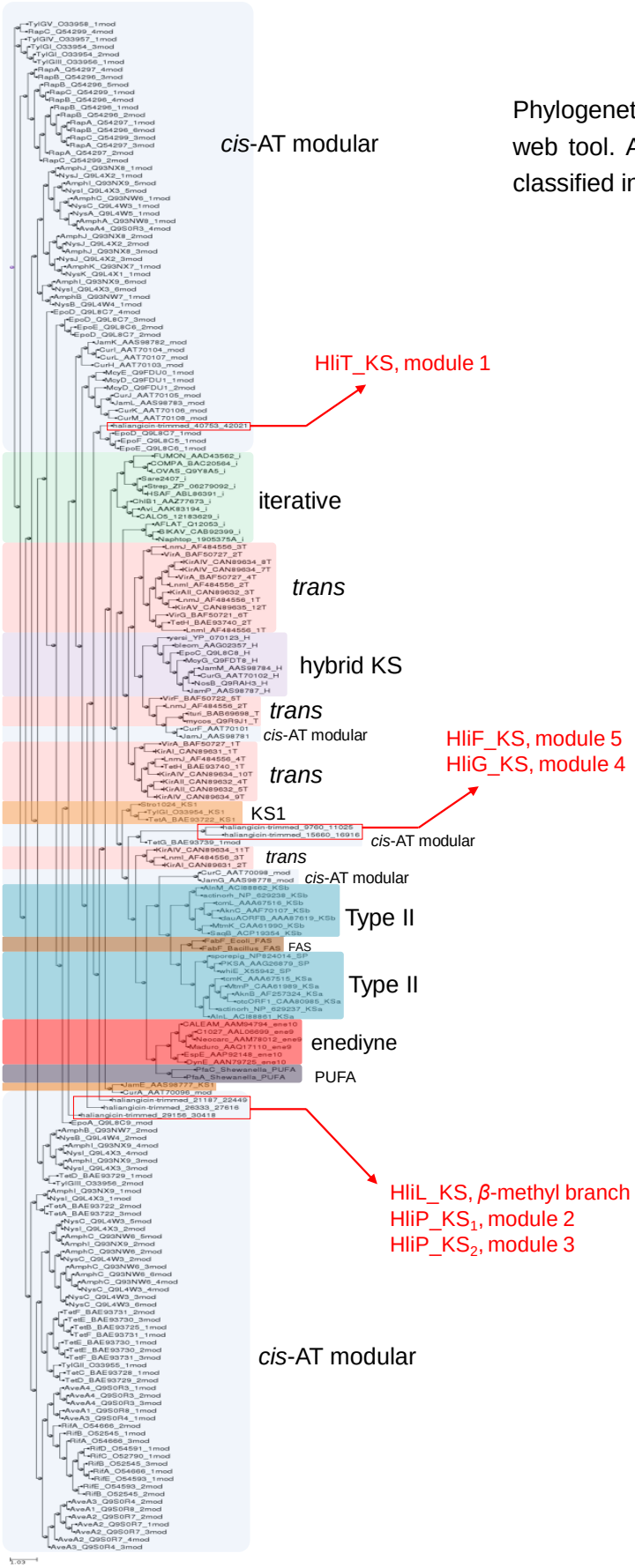
Table 2-8. Consensus sequences of KS domains in *hli* (CHH triad active residues are show in red).

KS Domain	DxxCSSxL								H									H								
HliF_KS	**D**	T	A	**C**	**S**	**S**	S	**L**	Y	V	E	T	**H**	G	T	G	T	T	N	M	G	**H**	A	E	T	A
HliG_KS	**D**	T	A	**C**	**S**	**S**	S	**L**	Y	I	E	A	**H**	G	T	G	T	T	N	F	G	**H**	L	E	A	A
HliL_KS	**D**	T	A	**C**	**S**	**S**	S	**L**	Y	I	E	A	**H**	G	T	G	T	S	N	I	G	**H**	L	E	S	A
$HliP\_KS_1$	**D**	T	G	**C**	**S**	A	S	**L**	Y	I	E	A	**H**	G	V	G	T	P	Y	I	G	**H**	L	E	T	A
$HliP\_KS_2$	**D**	T	A	**C**	**S**	**S**	S	**L**	Y	L	E	A	**H**	G	T	G	T	T	N	I	G	**H**	S	E	S	A
HliT_KS	**D**	T	A	**C**	**S**	**S**	G	**L**	L	L	E	A	**H**	G	S	G	T	P	N	L	G	**H**	M	E	S	A

Table 2-9. Alignments of consensus sequences of KS^s domain in β-methyl branching cassette.

KS Domain	GGASASGx								H									H								
HliL	D	T	**A**	C	S	**S**	S	L	Y	I	E	A	**H**	G	T	G	T	S	N	I	G	**H**	L	E	S	A
CroD	D	T	**A**	C	S	**S**	S	L	Y	I	E	A	**H**	G	T	G	T	T	N	L	G	**H**	T	E	A	A
VirB	**G**	**G**	**A**	**S**	**A**	**S**	**G**	T	Y	V	N	A	**H**	G	T	G	S	P	L	T	G	**H**	C	L	S	A
TaK	**G**	**G**	**A**	**S**	**A**	**S**	**G**	L	Y	V	N	P	**H**	G	S	G	S	S	I	T	G	**H**	G	L	S	S
CorD	**G**	**G**	**A**	**S**	**A**	**S**	**G**	N	Y	V	N	A	**H**	G	S	A	S	G	W	T	G	**H**	C	L	G	A
MupG	**G**	A	**A**	**S**	**A**	**S**	**G**	Q	Y	I	N	P	**H**	G	S	G	S	S	L	I	G	**H**	G	L	A	S
BatB	**G**	**G**	**A**	**S**	**A**	**S**	**G**	Q	Y	I	N	P	**H**	G	T	G	S	S	L	I	G	**H**	G	L	S	A

Accession: CroD (AIR74913.1), VirB (BAF50726.1), TaK (CAC85720.1), CorD (ADI59526.1), MupG (AAM12921.1), BatB (ADD82943.1)



Phylogenetic tree generated by NaPDoS web tool. All the KS domains in *hli* are classified into the modular type.

Figure 2-12. Phylogenetic analysis of KS domains in *hli* genes.

2.3.4.2. AT domains

All acyltransferase (AT) domains in *hli* are considered to be active based on their consensus motifs of Gx**S**xG (S92, catalytic residue),[43] except for HliP-AT_1 which is obviously degraded (truncated from around 100 amino acids) (Table 2-10). A complementary *trans*-acting AT to HliP_AT_1 was not found in the *hli* gene cluster, indicating a special AT-deficient feature in the haliangicin biosynthesis. This kind of AT-lacking PKS biosynthesis has already been reported in several other metabolites such as neocarzilin[44] and FK228.[45]

Among the functional AT domains, HliS_AT is assumed to be a starter-recognizing acyltransferase in the light of a glutamine at residue 117 instead of an arginine (Table 2-10). The important role of conserved Arg117 in discriminating between the CoA esters of mono- and di-carboxylic acid has been proved in various natural product biosynthesis[43] and therefore, a mutant of R117 to Q117 strongly supports HliS_AT as an AT domain recruiting mono-carboxylic acid in the starter-loading module. Considering the size of haliangicin structure and the number of AT domain, the substrate for HliS_AT is most likely to be a dipropionyl-CoA ester, (*E*)-2-methylpent-2-enoyl CoA rather than propionyl-CoA. We proved that its neighboring acyl-CoA dehydrogenase HliR specifically dehydrogenates (*E*)-2-methylpent-2-enoyl CoA but not propionyl-CoA through *in vitro* enzymatic assay (section **2.3.7**).

Alignments of acyltransferase (AT) domains in *hli* also allowed us to predict *in silico* the AT substrate specificity in each module (Table 2-10). Most of the extender units observed in haliangicin show good agreements with predicated substrate specificity: Phe_{200} in HliF_AT/HliL_AT support malonyl-CoA as substrates and Ser_{200} in HliP_AT_2/HliT_AT support methylmalonyl-CoA. Other than malonyl-CoA and methylmalonyl-CoA, HliG_AT specific for methoxymalonate gives a methylmalonyl-CoA characteristic motif (VASH). Alignments of HliG_AT with other methoxymalonyl-ACP specific AT[28, 32, 40, 46] (Table 2-11) indicated two kinds of amino acid in residue 200, serine or glycine (but not phenylalanine). It is still hard to give a general pattern characteristic for methoxymalonyl-ACP in residue 200 due to the rare reports of this extender. The discrimination between methoxymalonate/methylmalonate is predicted to be governed by more residues other than position 200. HliQ_AT which is supposed to recruit a 1,3-biphosphoglycerate shows distinct active residue in H11 and Y91 (Table 2-12). These mutations are also observed in SorC_AT

and PelG_AT that are functionally identical. This phenomenon indicates that such mutations might not be occasional, although more examples are expected to prove this theory.

Table 2-10. Conserved residues of AT domains in *hli* and prediction of substrate specificities.

AT domain	11	63	90	91	92	93	94	117	198	199	200	201	231	250	255	Predicted	Observed
	Q	Q	G	H	S	x	G	**R**	x	A	**x**	H	x	x	V		
HliF_AT	Q	Q	G	H	S	L	G	**R**	Q	A	**F**	H	N	Q	V	malonate	malonate
HliG_AT	Q	Q	G	H	S	M	G	**R**	V	A	**S**	H	T	N	V	methylmalonate	methoxymalonate
HliL_AT	Q	Q	G	H	S	V	G	**R**	H	A	**F**	H	C	H	V	malonate	malonate
HliP_AT$_1$*	Q	Q	G	D	R	V	G	E	-	-	-	-	-	-	-	-	malonate
HliP_AT$_2$	Q	Q	G	H	S	M	G	**R**	V	A	**S**	H	T	N	V	methylmalonate	methylmalonate
HliQ_AT	H	Q	G	Y	S	L	G	**R**	H	A	**F**	H	N	H	V	-	1,3-biphosphoglycerate
HliS_AT	Q	Q	G	H	S	L	G	**Q**	V	A	**G**	H	T	N	V	propionate	di-propionate
HliT_AT	Q	Q	G	H	S	M	G	**R**	V	A	**S**	H	T	N	V	methylmalonate	Methylmalonate

* HliP-AT$_1$ was supposed to be inactive and complemented by one *trans*-acting AT incorporating malonate, although such *trans*-AT cannot be found within the boundary of *hli*.

Table 2-11. Conserved residues of methoxymalonyl-ACP specific AT domains.

AT domain	11	63	90	91	92	93	94	117	198	199	200	201	255
	Q	Q	G	H	S	x	G	R	x	A	x	H	V
HliG-AT	Q	Q	G	H	S	M	G	R	V	A	**S**	H	V
PelA-AT$_2$	Q	Q	G	H	S	M	G	R	V	A	**S**	H	I
CndD-AT	Q	H	G	H	S	V	G	R	V	A	**S**	H	L
AsmB-AT$_1$	H	Q	G	H	S	Q	G	R	F	A	**G**	H	V
FkbA-AT$_1$	L	P	G	H	S	L	G	R	H	A	**G**	H	T
FkbA-AT$_2$	Q	Q	G	H	S	L	G	R	H	A	**G**	H	A

Accession: PelA (CAN97008.1), CndD (CAQ43078.1), AsmB (AAM54076.1), FkbA (AAF86396.1)

Table 2-12. Conserved residues of AT domains in SorC-like multifunctional enzymes.

AT domain	11	63	90	91	92	93	94	117	198	199	200	201	255
	Q	Q	G	H	S	x	G	R	x	A	x	H	V
HliQ_AT	**H**	Q	G	**Y**	S	L	G	R	H	A	F	H	V
PelG_AT	**H**	Q	G	**Y**	S	L	G	R	H	A	F	H	V
SorC_AT	**H**	Q	G	**Y**	S	L	G	R	H	A	F	H	V

Accession: PelG (CCE88382.1), SorC (AAK19884.1)

2.3.4.3. DH domains

Three dehydratase (DH) domains were found in HliG, HliP and HliT. Proposed biosynthetic route suggested them to be necessary for generating the observed polyene structure and thus should all be functionally active. HliP_DH and HliT_DH both contain consensus motif of LxxHxxxGxxxxP,[47] while HliG_DH bears a mutation of glycine to aspartic acid (Table 2-13). Replacement of this active glycine might be insignificant as it also happens in the biosynthesis of other bacterial secondary metabolites like microsclerodermin (MscA_DH).[48]

Table 2-13. Conserved sequences of DH domains in *hli*.

DH Domain	Active consensus (**L**xx**H**xxx**G**xxxx**P**)												
HliG_DH	**L**	E	D	**H**	T	V	Q	D	M	A	V	V	**P**
HliP_DH	**L**	G	D	**H**	R	V	G	**G**	A	V	V	F	**P**
HliT_DH	**L**	G	D	**H**	R	V	Q	**G**	A	V	L	A	**P**

2.3.4.4. KR domains

Accompanied with three DH domains, there are also three corresponding functional ketoreductase (KR) domains encoded in HliG, HliP and HliT. HliP_KR and HliT_KR both exhibit conserved motifs consistent with GxGxxGxxxA in active consensus 1 and LxS(G)RxG in consensus 2. In the case of HliG_KR, a glycine to serine mutation was observed in consensus 1 (Table 2-14). This substitution might also not affect the ketoreduction activity considering the proposed biosynthesis.

As for the analysis of KR subtype that may be related to the stereochemistry of ketone-derived hydroxyl, LDD motif in region 93-95 together with the additional Pro144, Asn148 indicate B type whereas Trp 141 supports A type.[49, 50] According to the alignments of KR amino acid sequence in *hli*, we can classify HliT_KR and HliG_KR as B type, although HliG_KR shows VDD (close to LDD) motif and Ser148 (Table 2-14). HliP_KR contains distinct residues in LDD motif (VSG) and do not have N148, nevertheless it also do not show Trp141 specific for A type and other A type characteristics, so the classification of HliP_KR remains unclear. The detailed stereogenic mechanism govern by KR domain has not been clearly unveiled up to now.

Table 2-14. Conserved sequences of KR domains in *hli* and classification of stereochemistry subtype.

KR Domain	Active consensus 1										Active consensus 2						Stereochemistry						Type
	GxGxxGxxxA										**LxS(G)RxG**						93	94	95	141	144	148	
HliG_KR	**G**	T	**S**	G	I	**G**	L	E	L	**A**	**L**	A	**S**	**R**	S	**G**	**V**	**D**	**D**	**L**	**P**	**S**	B
HliP_KR	**G**	L	**G**	G	I	**G**	A	E	V	**A**	**L**	L	**S**	**R**	S	**G**	**V**	S	G	**F**	**P**	**S**	?
HliT_KR	**G**	C	**G**	G	L	**G**	L	S	V	**A**	**L**	L	**G**	**R**	S	**G**	**L**	**D**	**D**	**I**	**P**	**N**	B

2.3.4.5. ER domain

Alignment of HliT_ER in module 1 with other enoyl reductase (ER) domains in several homologous gene clusters indicated that it hold GGVGxxAxQxA motif necessary for NADPH binding (Table 2-15). Thus, module 1 which contains functional DH-ER-KR domains is considered to be a fully reduced elongation cycle responsible for the single bond formation of C-10~C-11. Furthermore, HliT_ER is also responsible for the stereochemistry of 10-Me. Residue Trp52 is thought to be a key indicator of 2*S* conformation[49] but not the sole governing factor since lots of exceptions were reported. Anyway, the absence of Trp52 in HliT_ER might support an *R* configuration at C-10.

Table 2-15. Conserved sequences of ER domains (residue52 is highlighted in red).

ER Domain	NADPH binding site **GGVG**xx**A**x**Q**x**A**											Consensus motif (region 39-53)														
HliT_ER	**G**	**G**	**V**	**G**	L	A	**A**	V	**Q**	M	**A**	G	V	N	F	L	D	V	L	R	A	L	D	V	V	S
PelD_ER	**G**	**G**	**V**	**G**	L	S	**A**	I	**Q**	W	**A**	G	L	N	F	L	D	V	L	L	A	L	G	M	L	P
PelF_ER	**G**	**G**	**V**	**G**	L	S	**A**	I	**Q**	W	**A**	G	L	N	F	R	D	V	L	L	A	L	G	V	L	P
CndA_ER	**G**	**G**	**V**	**G**	L	A	**A**	I	**Q**	I	**A**	S	L	N	F	L	D	V	L	S	A	M	G	V	R	P
StiF_ER	**G**	**G**	**V**	**G**	M	A	**A**	V	**Q**	L	**A**	G	L	N	F	R	D	V	L	N	A	L	G	M	Y	P
JamL_ER	**G**	**G**	**V**	**G**	Q	A	**A**	V	**Q**	I	**A**	G	L	N	F	R	D	V	L	N	A	L	G	L	L	K

Accession: PelD (CCE88376.1), PelF (CCE88378.1), CndA (CAQ43075.1), StiF (CAD19090.1), JamL (AAS98783.1)

2.3.4.6. ACP domains

The *hli* locus harbors nine standard ACPs, regardless of PKS-integrated ACPs or standalone ACPs. All of these ACPs possess the conserved motif of GxxS, especially the serine residues for phosphopantetheine binding (Table 2-16).

Table 2-16. Conserved motif of ACP domains in *hli* (phosphopantetheine binding sites are in red).

ACP Domain	Active consensus (GxxS)				
HliF_ACP	L	**G**	L	D	**S**
HliG_ACP	L	**G**	F	D	**S**
HliJ	M	**G**	F	V	**S**
HliN	Y	**G**	A	S	**S**
HliP_ACP$_1$	A	**G**	M	S	**S**
HliP_ACP$_2$	L	**G**	I	D	**S**
HliQ_ACP	L	**G**	G	N	**S**
HliS_ACP	L	**G**	V	D	**S**
HliT_ACP	L	**G**	L	D	**S**

2.3.4.7. *O*-MT domains

Three methyltransferases present in *hli* that are responsible for the formation of 1-OCH_3, 3-OCH_3 and

4-OCH_3. Alignment of their amino sequences revealed SAM Co-factor binding motif for *O*-methyltransferase (DxGxGxG), although HliF_MT and HliQ_MT lost the second or the third glycine residues (Table 2-17). In consideration of the proposed biosynthetic pathway of haliangicin, these mutations should not affect the function of the *O*-methyltransferase

Table 2-17. Conserved sequences of *O*-MT domains.

MT Domain	Active consensus (**D**x**G**x**G**x**G**)						
HliD	**D**	V	**G**	V	**G**	S	**G**
HliF_MT	**D**	I	**G**	R	K	K	-
HliQ_MT	**D**	V	**G**	S	N	I	**G**

2.3.4.8. HCS

Alignment of HCS proteins indicated that HliM and its closest homology CroE both showed CHT catalytic triad rather than CHN (Table 2-18). Considering the unusual β-keto synthases (HliL/CroD) present in these two β-branching cassettes, the N→T mutation might represent a novel pattern of conserved residues in HCS which is coupled with atypical β-keto synthase.

Table 2-18. Conserved sequences of HCS homologues.

HCS			**C**					**H**					**N**			
HliM	H	A	**C**	Y	G	I	Y	**H**	A	P	L	G	T	A	Y
CroE	H	A	**C**	Y	G	I	Y	**H**	V	P	I	G	T	S	Y
TaF	Q	A	**C**	Y	A	A	M	**H**	T	P	V	G	**N**	L	C
MupH	Q	A	**C**	Y	S	A	F	**H**	T	P	V	G	**N**	I	M
TaC	S	A	**C**	Y	S	A	F	**H**	T	P	V	G	**N**	I	M
BatC	Q	A	**C**	Y	S	T	F	**H**	T	P	V	G	**N**	I	M
VirC	Q	A	**C**	Y	S	A	M	**H**	T	P	V	G	**N**	I	Y
CorE	Q	A	**C**	Y	G	A	F	**H**	T	P	V	G	**N**	L	Y

Accession: CroE (AIR74914.1), TaF (ABF92623.1), MupH (AAM12922.1), TaC (ABF90176.1), BatC (ADD82944.1), VirC (BAF50725.1), CorE (ADI59527.1)

2.3.5. The proposed haliangicin biosynthetic machinery

In the previous works, a conventional approach to the genetic modification of haliangicin pathway in *H. ochraceum* was adopted to unveil the biosynthetic mechanism of haliangicin. Disappointingly but not unexpectedly, direct mutagenesis by the targeted inactivation of PKS genes was unsuccessful in the slow-growing and delicate marine myxobacterial host.[14] While, in contrast to the technically intractable *H. ochraceum*, the heterologous expression mutant *M. xanthus* c10-11C/c7-6E provided great benefits for easier genetic manipulation of the haliangicin biosynthetic gene cluster (All *M. xanthus* mutants are listed in Table 2-1). Based on the knowledge of biosynthetic building blocks and sequence analysis of the *hli* gene cluster, I proposed the haliangicin biosynthetic machinery as Fig. 2-13 depicted.

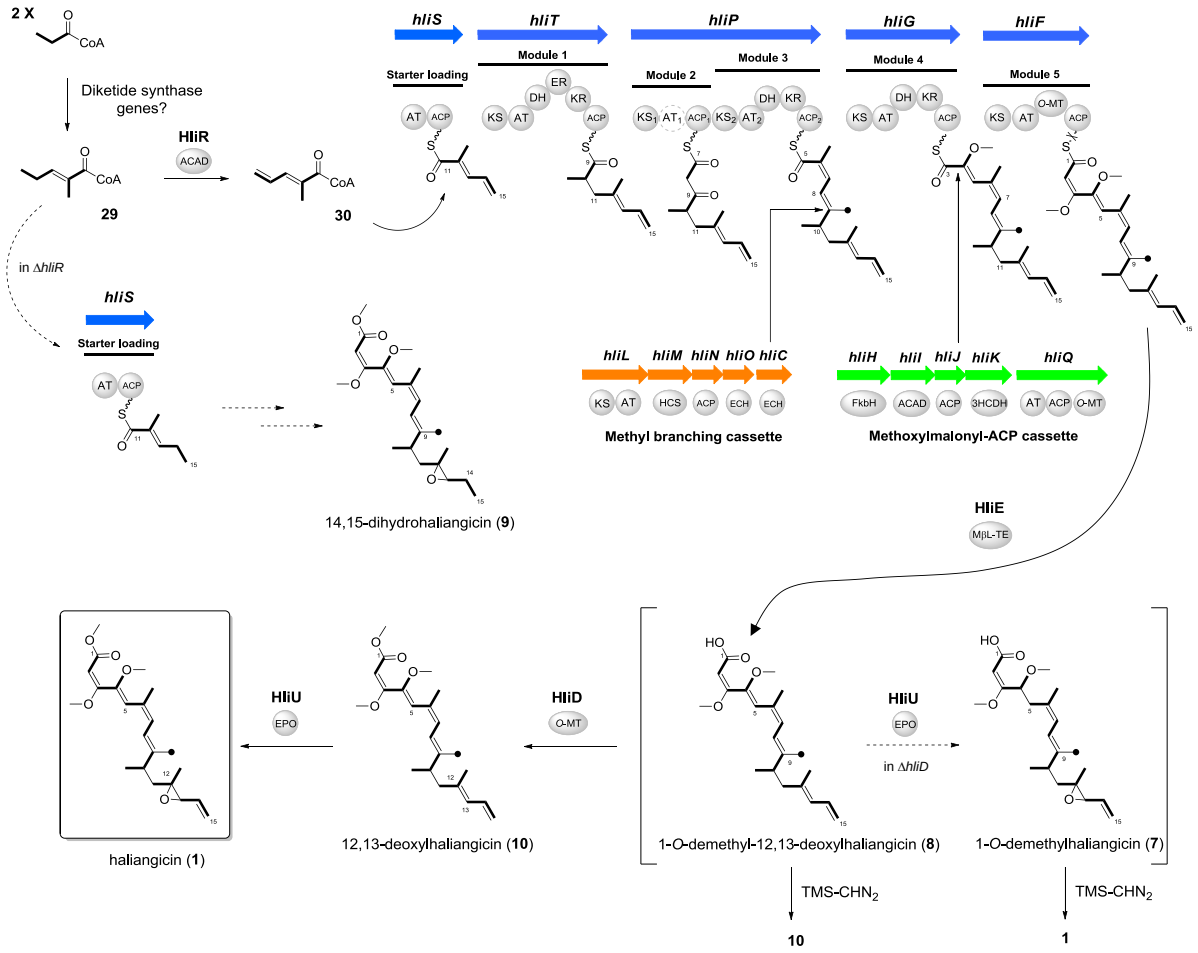


Figure 2-13. A proposed biosynthesis pathway for haliangicin (**1**). AT_1 in module 2 is inactive, but no complementary *trans*-AT is found in *hli*. The dashed arrows indicate the unnatural pathways in the

gene-disrupted strains of the heterologous host *M. xanthus* c10-11C/c7-6E. *Abbreviations*: ACAD, acyl-CoA dehydrogenase; ACP, acyl carrier protein; AT, acyl transferase; DH, dehydratase; ECH, enoyl-CoA hydratase; EPO, epoxidase; ER, enoyl reductase; FkbH, FkbH-like protein; HCS, HMG-CoA synthase; KR, ketoreductase; KS, ketosynthase; MβL-TE, metallo β-lactamase-type thioesterase; O-MT, O-methyltransferase; 3HCDH, 3-hydroxyacyl-CoA dehydrogenase.

From the viewpoint of conventional PKS biosynthesis, the construction of the heptaketide backbone of haliangicin (**1)** is likely to require six elongation steps following a C3 starter-loading step, but the *hli* machinery possesses only five elongation modules (modules 1–5 in Fig. 2-13). An iterative function of any particular PKS is unlikely, because all the KS domains in *hli* are classified as modular not iterative (Fig. 2-12). One possible solution to this problem of a missing module might be the direct use of a diketide (C_6) starter, such as 2-methylpent-2-enoyl-CoA (**29**) (Fig. 2-13 and more details in section **2.3.6**), which is suggested by a putative diketide synthase found in the genomes of *H. ochraceum* and *M. xanthus* (section **2.3.6** and Fig 2-15).

The diketide precursor **29** might undergo γ,δ-dehydrogenation by HliR, a particular acyl-CoA dehydrogenase, to form 2-methylpent-2,4-dienoyl CoA (**30**), which might serve as a starter for the formation of the unique 14,15-terminal olefin of haliangicin (**1**). A detailed functional analysis of HliR will be described in section **2.3.7**. The diene starter **30**, prepared by HliR, is next loaded into the typical starter-loading module encoded by *hliS*, and the five PKS-mediated elongation steps take place sequentially to generate the haliangicin backbone (Fig. 2-13). Sequence analysis (section **2.3.4**) showed that the domain arrangements of the PKSs corresponded well with the chemical structure of haliangicin (**1**). The module 2-mediated elongation step is followed by the formation of β-methyl branch at the 9-position by the methyl-branching cassette[26, 27] (HliC and HliL–O) (section **2.3.9**). Module 4 incorporates a glycolate extender, biosynthesized by the methoxymalonyl-ACP cassette[24, 28] (HliH–K and HliQ) (section **2.3.10**), to construct the vicinal oxygen functionality at C3–C4. Finally, the nascent carboxylic acid released by HliE (section **2.3.11**), which is homologous to a metallo-β-lactamase-type thioesterase[29, 30], undergoes two

post-assembly modification steps: *O*-methylation at the carboxyl terminus by HliD and epoxidation at the diene terminus by HliU (Fig. 2-13, section **2.3.12**).

2.3.6. Proposed initiation of the haliangicin biosynthesis

The haliangicin biosynthesis requires a starter loading step followed by six elongation reactions to generate the heptaketide backbone, therefore I first predicted a biosynthetic route that utilizes propionyl-CoA as a starter and followed by six elongation modules (Fig. 2-14a). This pathway needs one iterative KS and four typical KSs, while NaPDoS analysis indicated that all the KS domains in *hli* were classified into the modular type rather than the iterative ones (Fig. 2-12). One possible solution to the "missing module" is the direct utilization of a C_6 dipropionyl precursor such as 2-methyl-2-pentenoyl CoA (**29**) (Fig. 2-14b). Biosynthesis starting from such diketide needs only five typical KSs which exactly match the gene architecture.

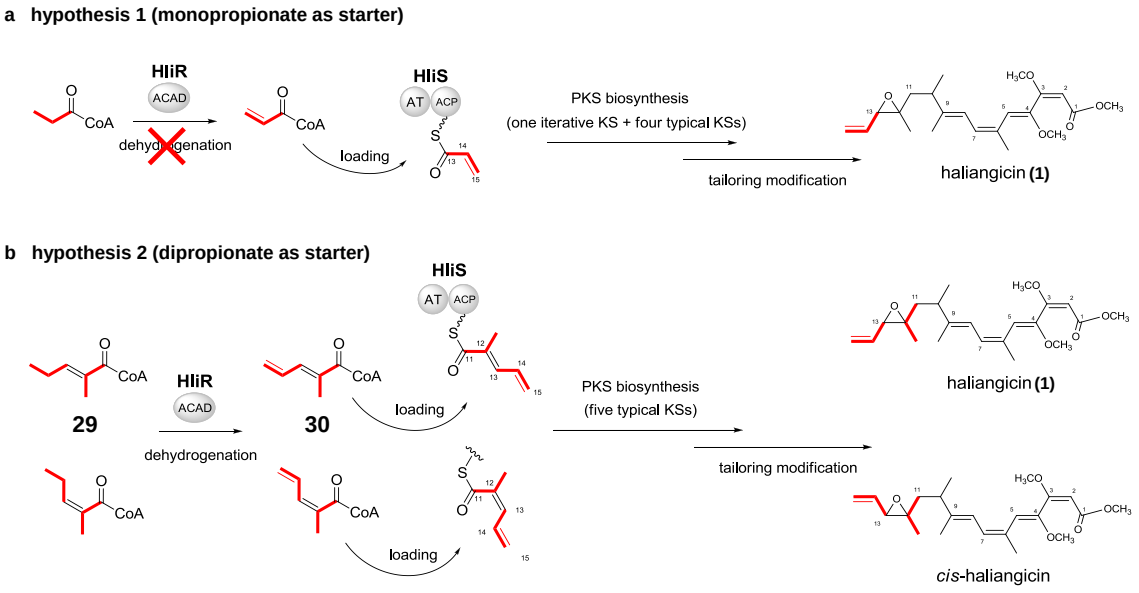


Figure 2-14. Proposed starter biosynthesis of haliangicin (**1**).

(**a**) Utilization of propionyl-CoA as starter (hypothesis 1). One iterative KS and four typical KSs are required to generate the heptaketide backbone. However, sequence analysis of KSs indicated no iterative KS in the *hli* gene cluster (Fig 2-12). Functional characterization of HliR also does not support hypothesis 1

because propionyl-SNAC is not accepted by HliR (section **2.3.7**).

(**b**) Utilization of dipropionyl-CoA as starter (hypothesis 2). Only five typical KSs are required. Functional characterization of HliR supports hypothesis 2 because HliR specifically catalyzes the dehydrogenation of the 2-methylpent-2-enoyl-CoA (**29**) (section **2.3.7**).

Although CoA esters of short branched-chain carboxylic acids like isobutyrate, isovalerate, and 2-methylbutyrate have already been reported to serve as starter units for bacterial metabolites,[51] a six-carbon unit composed of two propionates derived from carboxylate has never been reported. The condensation of two propionyl-CoA molecules with concomitant reduction and dehydration may result in such enone starter, while this process needs a complete cycle of PKS elongation consisting of KS, AT, ACP, DH and KR domains, which is not detected within the region of *hli*. Searching the genome of *H. ochraceum* and the heterologous host *M. xanthus* led to the discovery of two diketide synthase in their respective genomes (Fig. 2-15)

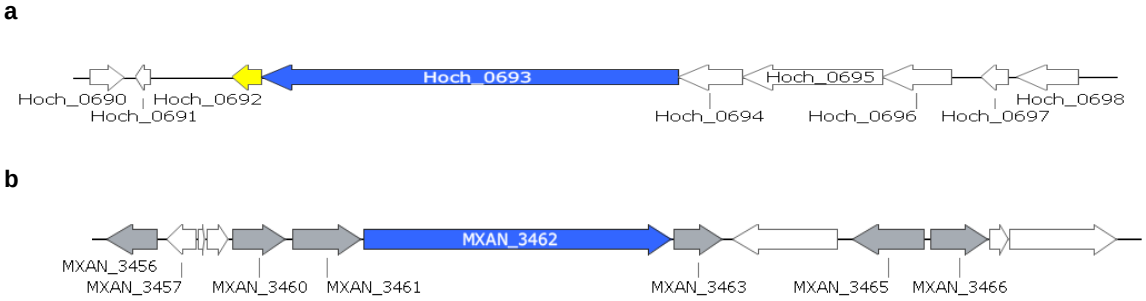


Figure 2-15. The putative diketide synthases for the biosynthetic formation of the starters.

(**a**) The putative diketide synthase harboring in the genome of *H. ochraceum*. Hoch_0693 (1984 amino acids) encodes four domains: KS, AT, KR and DH, which are just fit for one round of elongation reaction to generate an α,β-unsaturated diketide. The successive Hoch_0692 encodes an acyl-CoA thioesterase which might be involved in the formation of CoA ester of the diketide.

(**b**) The putative diketide synthase harboring in the genome of *M. xanthus*. MXAN_3462 (2216 amino acids) highly resembles Hoch_0693 and consists of five domains: acyl-CoA thioesterase, AT, KS, KR and DH.

The roles of its accompanying enzymes, including MXAN_3456 (radical SAM domain-containing protein), MXAN_3460 (oxidoreductase), MXAN_3461 (short chain dehydrogenase/reductase oxidoreductase), MXAN_3461 (fatty acid desaturase), MXAN_3465 (histidine ammonia-lyase) and MXAN_3466 (aspartate kinase) remain unclear.

It should be noticed that the stereochemistry of the α,β-double bond in the diketide starter dominates the final configuration of 12,13-epoxyl in the haliangicin structure (Fig. 2-14b). Haliangicin is accompanied by *cis*-haliangicin (12,13-*cis*-epoxy isomer) in a ratio of 3:2 in the native producer *H. ochraceum*,[20] whereas the heterologous host *M. xanthus* c10-11C/c7-6E strain almost solely produces haliangicin with no *cis*-haliangicin.[14] I suggested that in *H. ochraceum*, the dipropionyl starters present as *trans/cis* isomers in a ratio of 3:2, while the *M. xanthus* heterologous host predominately provides the *trans*-form. This difference might a reasonable explanation for why haliangicin is solely produced in *M. xanthus* host.

The diketide starter undergoes a rare γ,δ-dehydrogenation by the HliR (detail in section **2.3.7**) and then activated and loaded by HliS. HliS is a didomain protein showing a classical starter loading module organization (AT-ACP), which resembles the loading module of the well-studied 6-deoxyerythronolide B synthase 1 (DEBS 1)[52, 53] and avermectin synthase 1 (AVES 1).[54] A mutation Arg_{117} to Gln_{117} in the AT domain of HliS was observed (Table 2-10). Residue 117 of AT domain is characterized to be crucial to distinguish monocarboxylic acids (usually as starters) and dicarboxylic acids (usually as elongation units).[43] The glutamine in this residue strongly supports HliS as a starter-loading acyltransferase (AT_L).

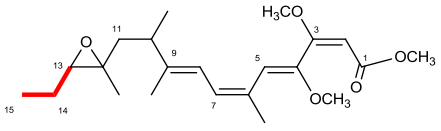
2.3.7. HliR, a unique acyl-CoA dehydrogenase involved in terminal alkene formation

An acyl-CoA dehydrogenase encoding gene *hliR* located upstream to *hliS* (the starter-loading module) is supposed to be responsible for the dehydrogenation of the starter, leading to the formation of the terminal olefin moiety at C-14 and C-15 in haliangicin (**1**). Since terminal olefins infrequently occur in natural products, for example, curacin A,[55] tautomycetin[56] and kalkitoxin,[57] they have been attracting lots of interest to unveil the biosynthetic mechanism behind them. Gu, *et al.* reported a unique

O-sulfonation-mediated chain termination leading to the terminal alkene by the sulfotransferase (ST)-TE domain of CurM in the curacin A biosynthesis.[58] Another unusual mechanism was elucidated as a decarboxylative dehydration reaction catalyzed by TtnC, D, and F in the installation of the terminal diene in tautomycetin.[59] However, haliangicin seems to undergo neither pathway because no sequences homologous to CurM or TtnC/D/F were found in the haliangicin biosynthetic machinery. Referring to other biosynthetic routes generating similar structures such as allyl side chain, I noticed that HliR is a good candidate for a dehydrogenase that generates the C-14/C-15 double bond. Disruption of *hliR* in the heterologous host *M. xanthus* through the single crossover recombination indeed abolished the production of haliangicin and generated a novel compound, 14,15-dihydrohaliangicin (**9**, Table 2-19 and Chapter 5 Appendix **5.3**), demonstrating the direct involvement of HliR in the terminal alkene formation (Fig. 2-16).

Table 2-19. ^{1}H and ^{13}C NMR data for 14,15-dihydrohaliangicin (**9**) in $CDCl_3$ (400 MHz).

	14,15-dihydrohaliangicin (**9**)	
Position	δ_H (mult, *J* in Hz)	δ_C
1	-	167.0
2	5.25 (s)	95.7
3	-	165.8
4	-	148.1
5	5.78 (s)	112.2
6	-	130.5
7	6.03 (d, 11.2)	126.3
8	6.11 (d, 11.2)	120.3
9	-	142.0
10	2.33 (m)	40.4
11	1.21 (dd, 9.2, 13.6), 1.95 (dd, 6.0, 13.6)	44.3
12	-	60.2
13	2.68 (t, 6.4)	65.1
14	1.48 (m), 1.60 (m)	22.0
15	1.01 (t, 7.6)	10.5
1-OMe	3.67 (s)	51.2
3-OMe	3.72 (s)	56.0
4-OMe	3.55 (s)	57.2
6-Me	2.09 (s)	23.6
9-Me	1.69 (s)	13.5
10-Me	1.00 (d, 6.8)	19.2
12-Me	1.23 (s)	16.6



14,15-dihyrohaliangicin (**9**)

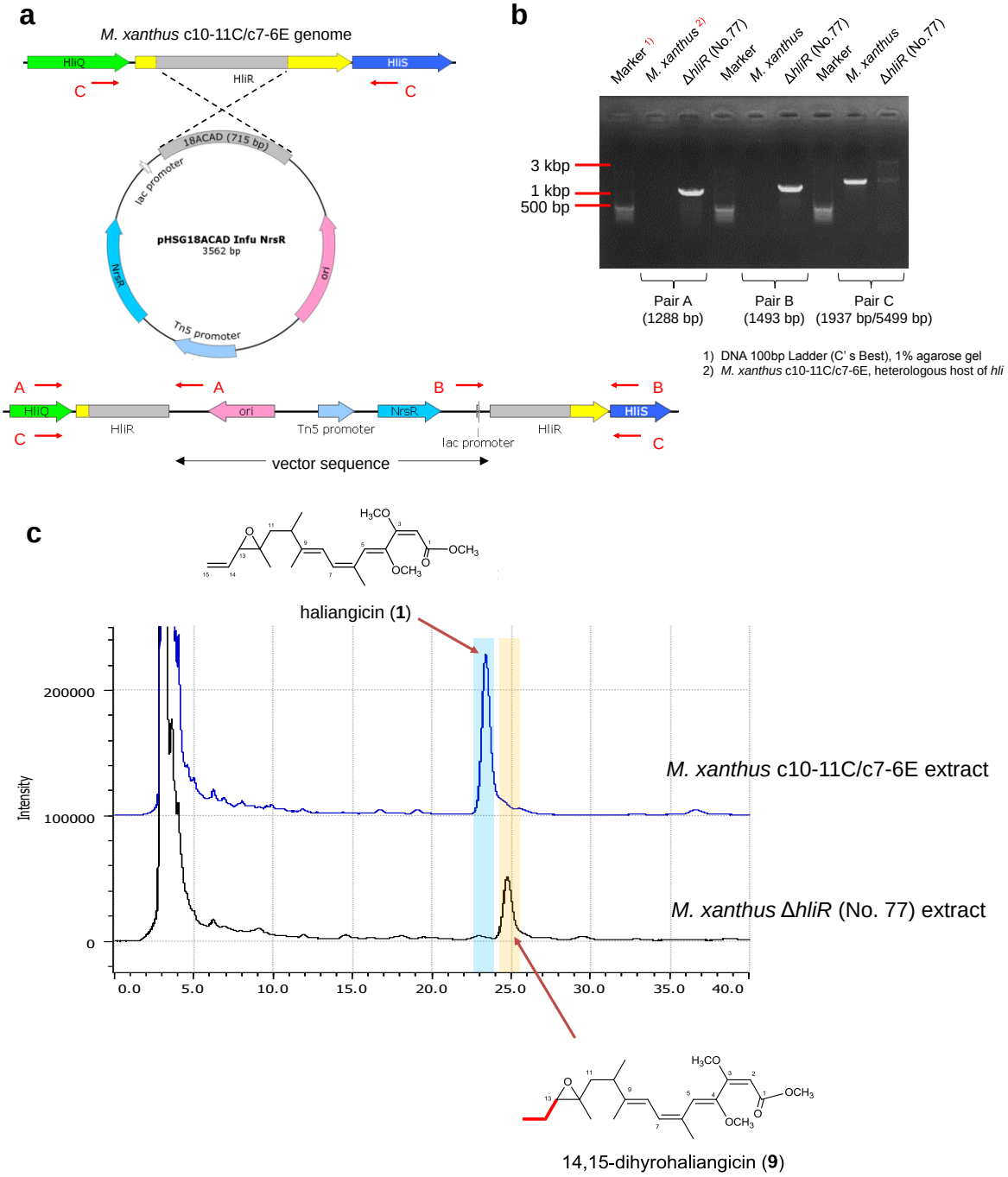


Figure 2-16. Inactivation of HliR (acyl-CoA dehydrogenase).

(**a**) Disruption of *hliR* via single crossover homologous recombination.

(**b**) Verification of Δ*hliR* mutant by PCR [Pair A: M13f(-47)/ORF18ACAD-Out f; Pair B: M13r(-48)/ORF18ACAD-Out r; Pair C: ORF18ACAD-Out f/ORF18ACAD-Out r].

(**c**) Production profile of *M. xanthus* Δ*hliR*. Production of 14,15-dihydrohaliangicin (**9**, t_R = 24.8 min) was observed.

Next, I focused on determining of the exact timing of the dehydrogenation step during the haliangicin biosynthesis by characterizing *in vitro* enzymatic activity of HliR. Recombinant HliR was expressed in *E. coli* as a soluble protein with *N*-His_6-tags and purified (Figure 2-17). Incubation of 14,15-dihydrohaliangicin (**9**) with recombinant HliR in the presence of flavin adenine dinucleotide (FAD) resulted no production of dehydrogenated products, excluding the possibility of dehydrogenation at C-14/C-15 position as a post-PKS modification step (Figure 2-18).

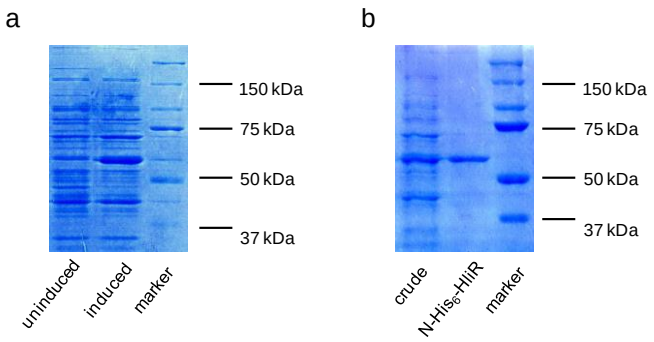


Figure 2-17. SDS-PAGE (10% Tris-HCl gel) of recombinant HliR. Marker: Precision Plus Protein™ Dual Color Standards (Bio-Rad)

(**a**) Soluble protein extracted from *E. coli* KRX pET32a HliR with/without induction of 0.1% rhamnose and 0.4 mM IPTG.

(**b**) Purified N-His_6-HliR.

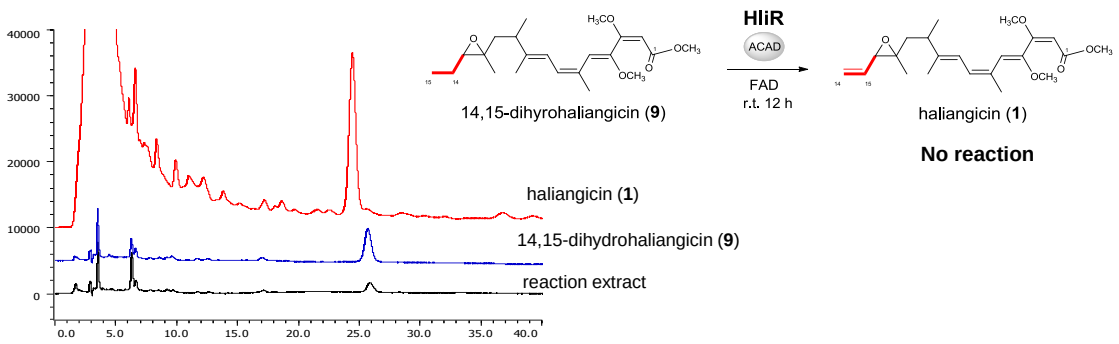


Figure 2-18. HPLC analysis of HliR enzymatic reaction mixture using 14,15-dihydrohaliangicin (**9**) as substrate. HPLC condition: Develosil ODS-UG-5 (ϕ 4.6 mm × 250 mm), 75% methanol, flow rate 1.0

mL/min, UV detection at 290 nm. Formation of haliangicin (**1**) was not detected.

A detailed sequence analysis (Fig. 2-19) deduced that HliR was a unique short chain acyl-CoA dehydrogenase phylogenetically close to TcsD of *Streptomyces kanamyceticus* and *Streptomyces tacrolimicus* (approximately 50% identity at protein sequence level).[60] TcsD was identified as a unique acyl-ACP dehydrogenase catalyzing the desaturation of a five-carbon extender, propylmalonyl-ACP, to allylmalony-ACP in the side chain biosynthesis of FK506. In analogy to this particular biosynthetic pathway, I hypothesized that HliR dehydrogenates the terminal ethyl moiety of diketide starter 2-methylpent-2-enoyl-CoA (**29**), which results in an olefin end (C-14/C-15) of haliangicin (Fig 2-13 and 2-14). To prove this hypothesis, a thio ester mimic of **29**, (*E*)-2-methylpent-2-enoyl SNAC (**11**) was synthesized as a substrate and evaluated its reactivity to recombinant HliR (Fig. 2-20a). According to the HPLC analysis, **11** was entirely converted into the corresponding dehydro product **27** by HliR in the presence of FAD (Fig. 2-20b). Scale-up enzymatic reaction was achieved by feeding of **11** to the recombinant HliR-expressing *E. coli* under the induction by IPTG and Rhamnose. Despite of its relative instablity due to diene structure, **27** was obtained from *E. coli* extracts and confirmed as the dehydrogenated product of **11** using comprehensive spectroscopic analysis including ^{1}H, ^{13}C NMR and HRMS (for spectra, see Chapter 5 Appendix **5.5**). From these results, I concluded that the real substrate for HliR is possibly a diketide acyl CoA ester **29** during the early stage of haliangicin biosynthesis.

The substrate specificity of HliR was further evaluated by using 14 additional short thio ester mimics, revealing that HliR is highly specific to α,β-unsaturated short acyl compounds, but does not consume saturated or γ,δ-unsaturated compounds (Fig. 2-20 and 2-21). Especially, HliR does not accept the "mono-ketide" mimic propionyl SNAC (**15**) as its substrate. No dehydrogenation product with identical retention time to acryloyl SNAC (**16**) was observed in the enzymatic reaction. On the other hand, among the substrates tested, pent-2-enoyl SNAC (**12**), like **11**, was also completely converted into **28** (for spectra, see Chapter 5 Appendix **5.6**), whereas 4-methylpent-2-enoyl SNAC (**13**) and hex-2-enoyl SNAC (**14**) underwent partial dehydrogenation (Fig. 2-20b). These results suggested that the presence of substituents at

the γ- or δ-position might hinder the dehydrogenation reaction from the enzyme. In addition, unlike TcsD, which demands acyl carrier protein-tethered substrates, HliR can utilize free small thio esters that mimic acyl-CoA substrates, offering a new strategy for preparing terminal alkenes.

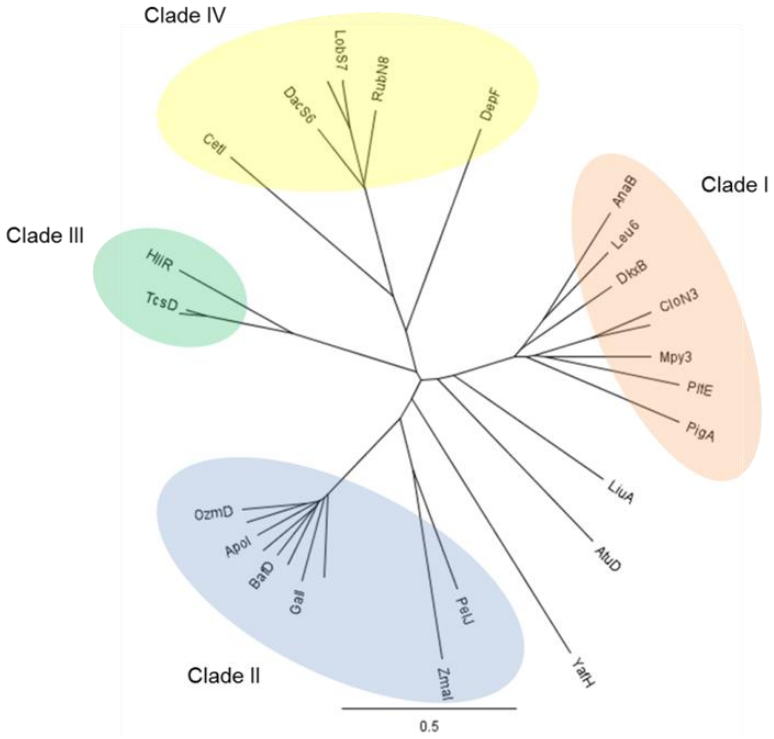


Figure 2-19. Phylogenetic tree of bacterial acyl-CoA dehydrogenase (ACAD)-related oxidoreductases obtained by using the NJ method. Clade I: L-prolyl-*S*-PCP dehydrogenase; Clade II: methoxymalonyl-ACP biosynthesis involved oxygenase; Clade III: propylmalonyl-ACP dehydrogenase; Clade IV: nitrosugar biosynthesis-related nitrososynthase. HliR is classified in Clade III with TcsD of *S. kanamyceticus* and *S. tacrolimicus*.[60] Accession No: AnaB (ACR33074.1), ApoI (AEP40922.1), AtuD (AEV61767.1), BafD (ADC79623.1), CetI (ACH85569.1), CloN3 (AAN65232.1), DacS6 (AFU65918.1), DepF (ABP57750.1), DkxB (CAQ34915.1), GalI (ADE22333.1), Leu6 (ADZ24990.1), LiuA (BAN53548.1), LobS7 (AGI99501.1), Mpy3 (AFP87520.1), OzmD (ABA39084.2), PelJ (CCE88384.1), PigA (CAH55646.1), PltE (AAD24879.1), RubN8 (CAI94696.1), TcsD (ADU56239.1), YafH (NP_414756.2), ZmaI (AAR87758.1).

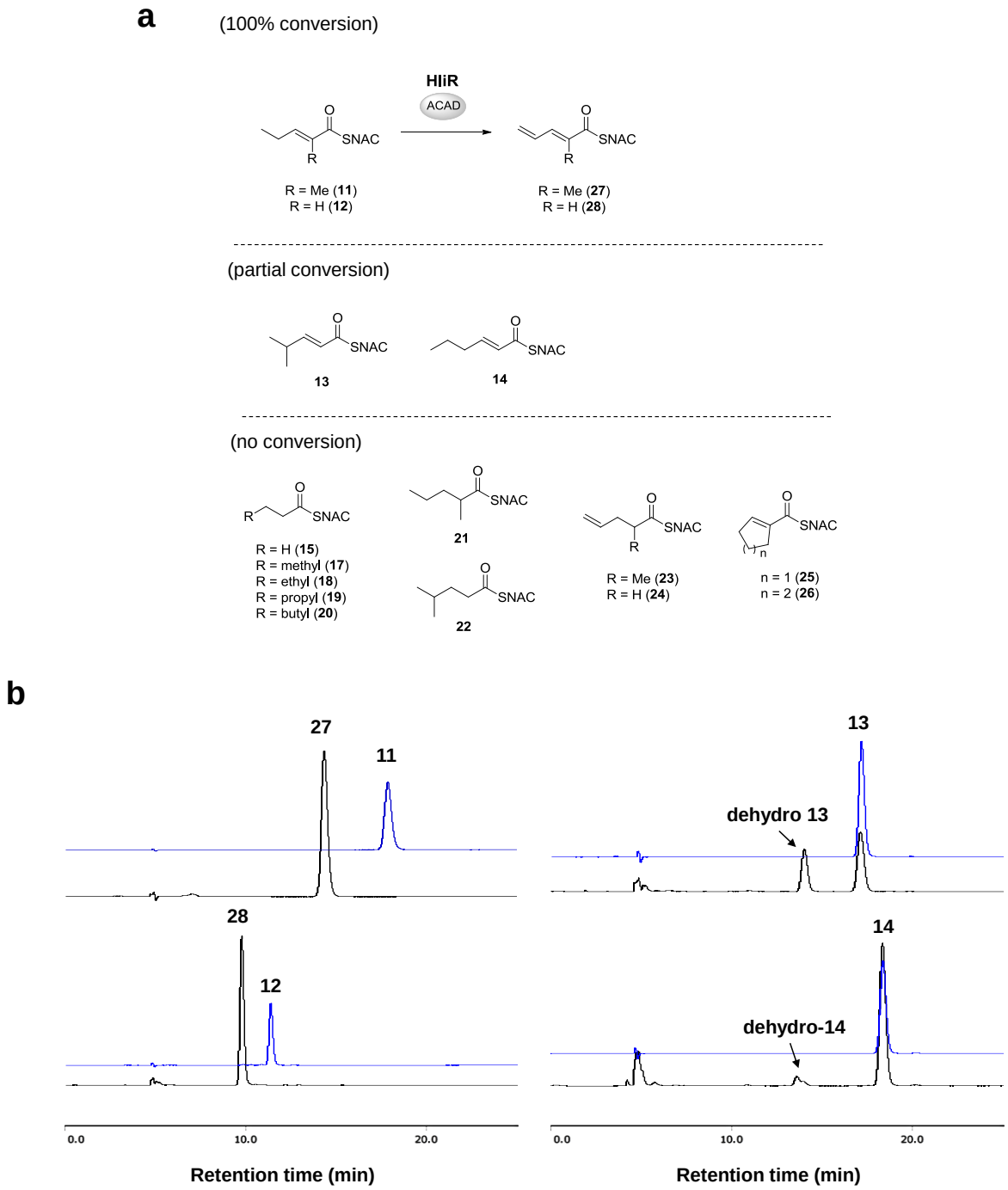


Figure 2-20. Functional analyses of HliR, an acyl-CoA dehydrogenase.

(**a**) Substrate specificity of recombinant acyl-CoA dehydrogenase HliR.

(**b**) *In vitro* dehydrogenation of acyl-SNAC compounds by HliR. HPLC analyses of the enzymatic reaction mixtures (black lines) with the substrates **11**–**14** (blue lines).

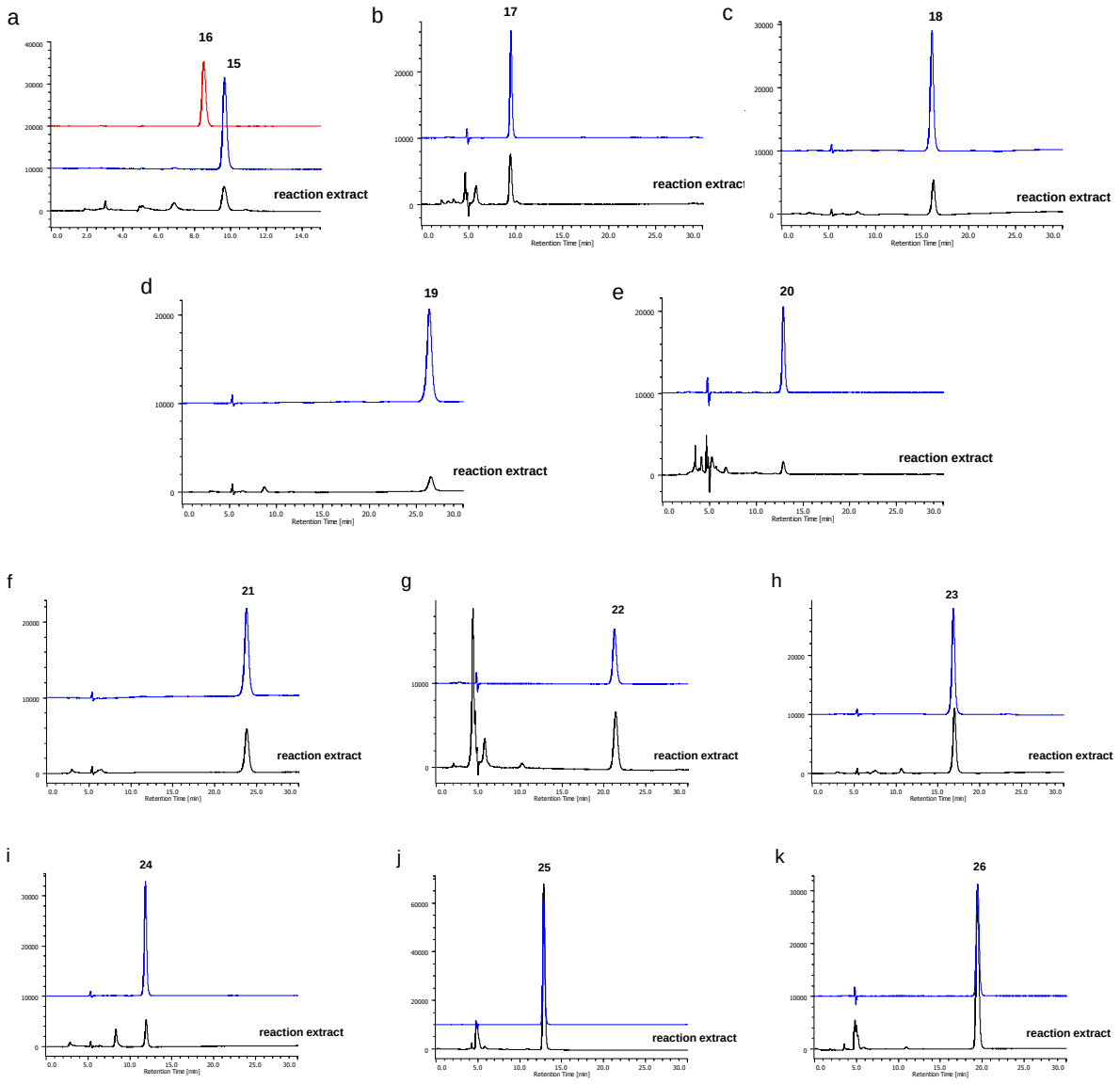


Figure 2-21. HPLC analyses of the HliR enzymatic reactions using, propionyl SNAC (**15**, **a**), butyryl SNAC (**17**, **b**), valeryl SNAC (**18**, **c**), hexanoyl SNAC (**19**, **d**), heptanoyl SNAC (**20**, **e**), 2-methylvaleryl SNAC (**21**, **f**), 4-methylvaleryl SNAC (**22**, **g**), 2-methyl-4-pentenoyl SNAC (**23**, **h**) 4-pentenoyl SNAC (**24**, **i**), , cyclo-1-pentenoyl SNAC (**25**, **j**), cyclo-1-hexenoyl SNAC (**26**, **k**). Formation of dehydro products were not detected. HPLC condition: Develosil ODS-UG-5 (ϕ 4.6 × 250 mm); 50% methanol, flow rate 0.6 mL/min, and UV detection at 239 nm for **a**; 60% methanol, flow rate 0.6 mL/min, and UV detection at 254 nm for **b**-**d**, **f**-**k**; 75% methanol, flow rate 0.6 mL/min, and UV detection at 254 nm for **e**.

2.3.8. PKS for chain elongation

After the starter loading step, the haliangicin biosynthesis continues with five PKS-mediated elongation steps to generate the haliangicin backbone (Fig. 2-13). Four PKS megasynthases (HliF, G, P and T) are thought to be responsible for the backbone construction. Among the four PKSs, HliP is the largest and encodes two modules (Module 2 and 3) that catalyze the condensation of C-7~C-8 unit and C-5~C-6 unit to the growing chain (Fig. 2-22). Module 2 (the upstream of HliP) has an abbreviated organization that consists of one KS, one ACP and one AT domain relic supposed to be inactive (Table 2-10). This kind of AT-missing module organization was found in some of the polyketides with β-branched methyl groups such as corallopyronin A,[37] elansolid,[61] and calyculin,[62] in which the AT deficiency is complemented by *trans*-acting AT. However, there is no obvious *trans*-acting AT in the *hli* gene cluster and a phylogenetic analysis of KS clades also supported *hli* to be rather a *cis*-AT system. Without further genetic information at present, it is difficult to explain how the AT-missing domain activity is fulfilled.

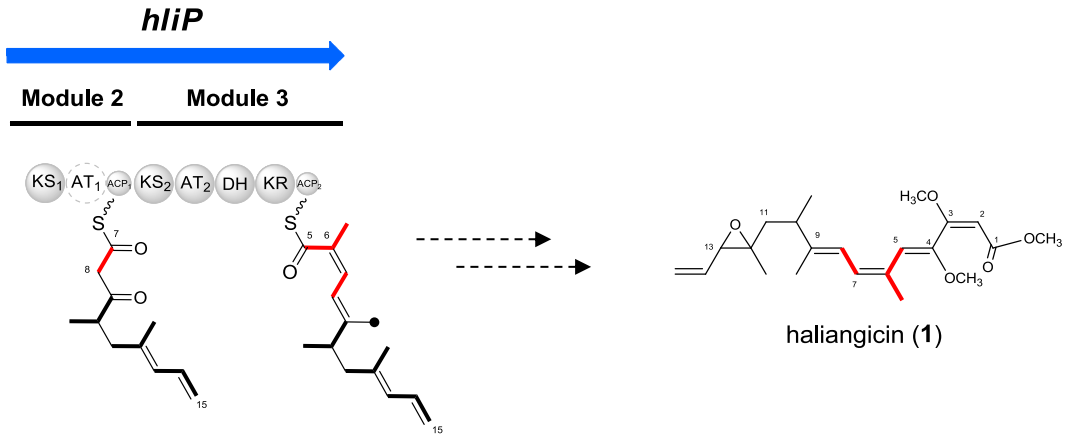


Figure 2-22. HliP (polyketide synthase) involved in the construction of haliangicin carbon backbone.

Disruption of HliP through single crossover recombination in the heterologous host of *M. xanthus* abolished the haliangicin production (Fig. 2-23). This result firmly verified the assignment of *hli* locus for the haliangicin biosynthesis.

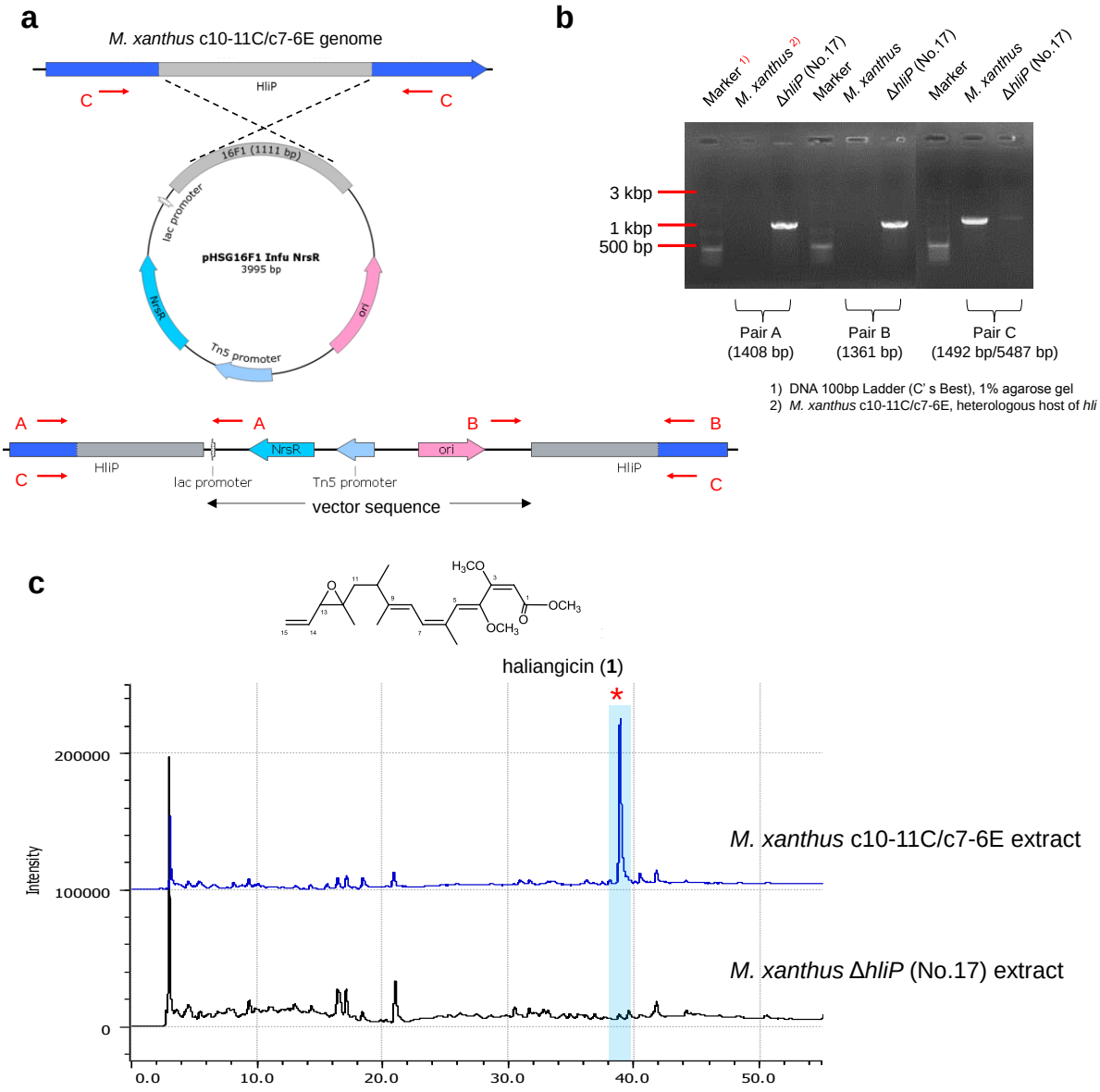


Figure 2-23. Inactivation of HliP (PKS).

(**a**) Disruption of *hliP* via single crossover homologous recombination.

(**b**) Verification of Δ*hliP* mutant by PCR [Pair A: M13r(-48)/ORF16-f3; Pair B: M13f(-47)/ORF16 r4; Pair C: ORF16-f3/ORF16 r4].

(**c**) Production profile of *M. xanthus* Δ*hliP*. Production of haliangicin (**1**) was abolished.

2.3.9. β-methyl branching biochemistry in the haliangicin biosynthesis

After the elongation step by module 2, the PKS biosynthesis once shift to the β-methyl branching biochemistry to install 9-Me at C-9 (Fig. 2-13, 2-24). The installation of the acetate-derived 9-Me involves a set of highly conserved genes called "β-methyl branching cassette". Referring to previous reports on this gene cassette, the core biosynthetic machinery required at least one HMG-CoA synthase (HCS), one β-keto synthase with a Cys-to-Ser mutation (KS^S), one free-standing ACP and two enoyl-CoA hydratases (ECH).[63] Although the proteins similar to these proteins can be found as HliM (HCS), HliN (ACP), HliO (ECH_1) and HliC (ECH_2) in *hli*, to our perplexity, *hli* lacks a β-keto synthase (KS^S), which is critical for decarboxylation. This is quite rare because KS^S presents in most natural products assembly lines harboring the "β-methyl branching cassette" (Table 2-6). I hypothesized that the KS domain in the adjacent HliL complements this KS^S. Its coupling AT (HliL_AT) in favor of a malonate (the origin of 9-Me) also supported HliL to be part of the β-methyl branching cassette. However, the KS domain of HliL appears to be a typical KS as it retains the conserved triad active residues (CHH) (Table 2-9). Muller, *et al.* recently reported a similar bifunctional protein CroD in *Chondromyces crocatus*, in which the KS domain also does not feature the Cys-to-Ser mutation.[64] CroD is predicted to be a relic of a former PKS but now takes part in the construction of β-branching structure in crocacin.

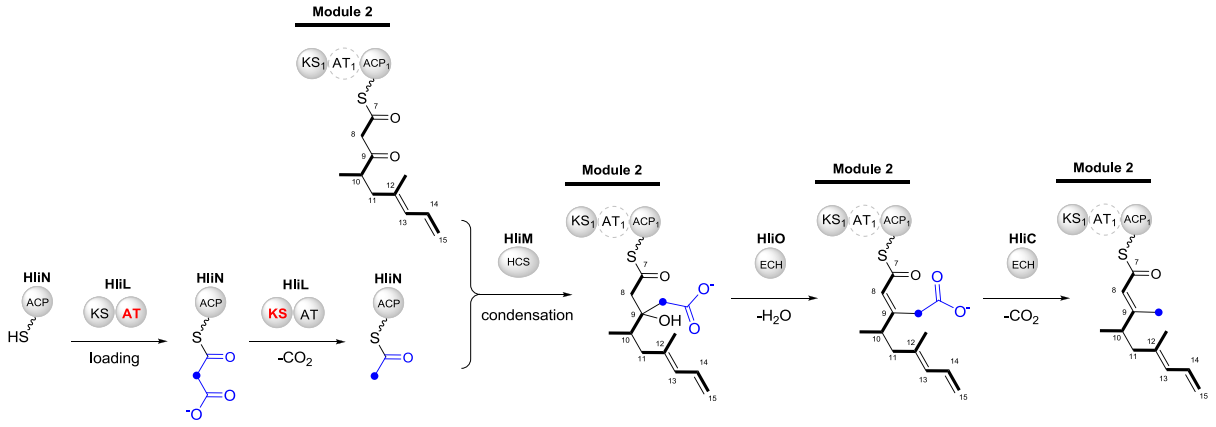


Figure 2-24. Proposed β-methyl branching biochemistry in the haliangicin biosynthesis.

2.3.10. Methoxymalonyl-ACP biosynthesis in the haliangicin biosynthesis

Module 4 incorporates a methoxymalonate extender, which is likely biosynthesized by the methoxymalonyl-ACP cassette (HliH ~ K and HliQ). The incorporation of this particular building block into the C-3 and C-4 positions of the haliangicin skeleton was clarified by feeding studies using [U-$^{13}C_3$]glycerol as described in section **2.3.2**. According to the known glycolate biosynthetic pathways such as in pellasoren,[40] soraphen A, FK-520 and ansamitocin, I predicted the methoxymalonate pathway as follows: 1,3-biphosphoglycerate,[29] an key intermediate in glycolysis, is activated and loaded to ACP by an AT domain in HliQ or a FkbH-like protein HliH. The ACP-bound 1,3-biphosphoglycerate intermediate then undergoes two oxidation steps by two discrete acyl-CoA dehydrogenases (HliI and HliK) followed by methylation by a *O*-methyltransferase (*O*-MT domain of HliQ) to afford methoxymalonyl-ACP (Fig. 2-25). Similar to pellasoren, a multifunctional protein HliQ (AT-ACP-*O*-MT) leads to several functionally redundant proteins. The AT domain of HliQ and HliH are both supposed to recognize 1,3-biphosphoglycerate and load it to an ACP (the ACP of HliQ or the freestanding HliI, which also present in duplicate).

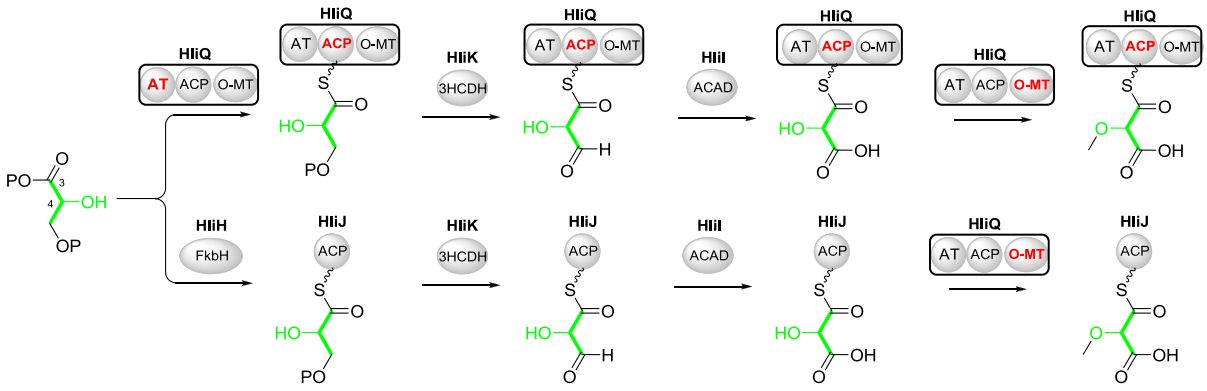


Figure 2-25. Proposed biosynthetic pathway of methoxymalonyl-ACP in haliangicin biosynthesis.

In the pellasoren biosynthesis, the role of each protein of *pelG~ K* is unclear due to their operon organization. Unlikely to *pelG ~ K*, *hliQ* and *H ~ K* are divided into two operons, which might facilitate

gene inactivation, providing a further insight into this unique gene arrangement. *HliH* and *hliQ* were individually disrupted in the heterologous host *M. xanthus* c10-11C/c7-6E through the homologous recombination with a disruption vector, revealing that both Δ*hliH* and Δ*hliQ* mutants stopped the production of haliangicin (**1**) (Fig. 2-26 and 2-27). These results again revealed the complexity of PKS biosynthesis concerning the "glycolate" extender unit, which still requires further investigations.

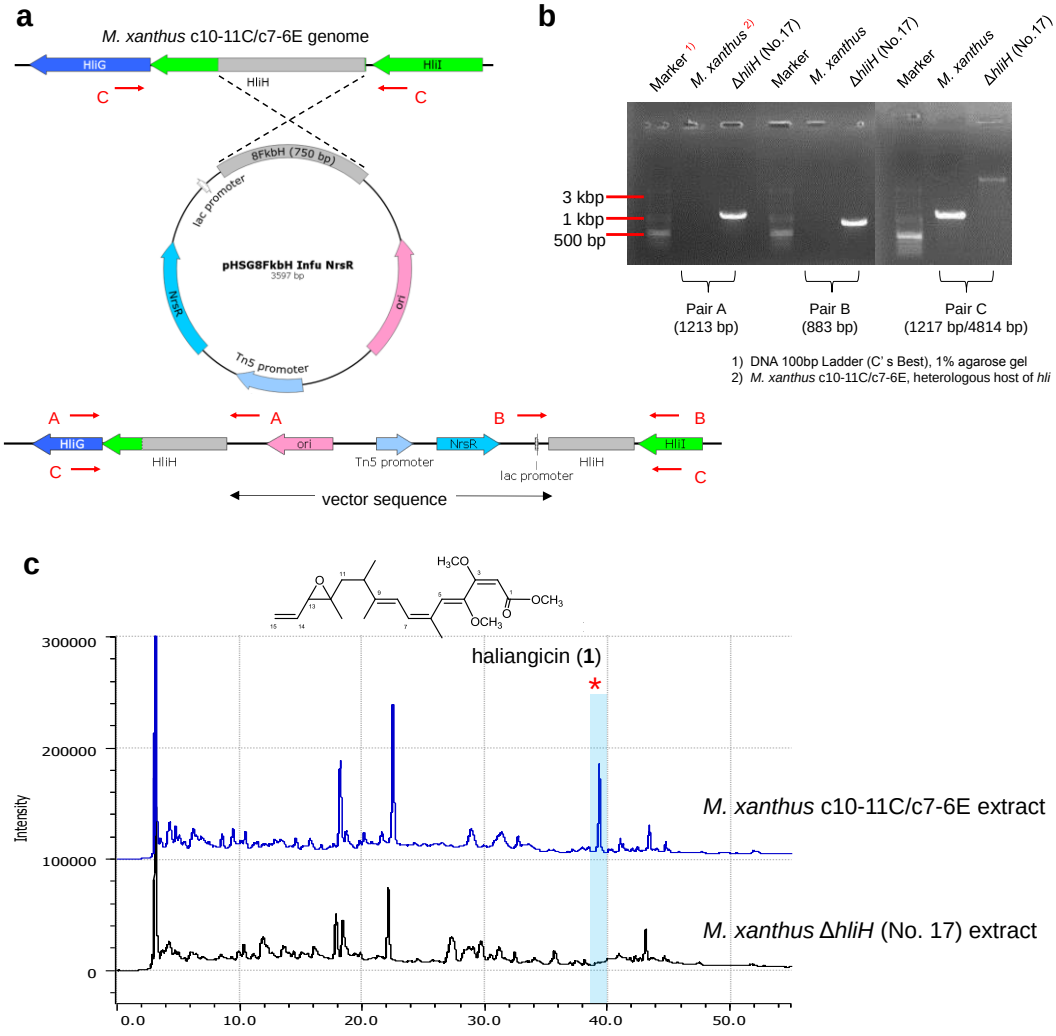


Figure 2-26. Inactivation of HliH (FkbH-like protein).

(**a**) Disruption of *hliH* via single crossover homologous recombination.

(**b**) Verification of Δ*hliH* mutant by PCR [Pair A: M13f(-47)/HliH-Ex f; Pair B: M13r(-48)/HliH-Ex r; Pair C: HliH-Ex f/HliH-Ex r].

(**c**) Production profile of *M. xanthus* Δ*hliH*. Production of haliangicin (**1**) was abolished.

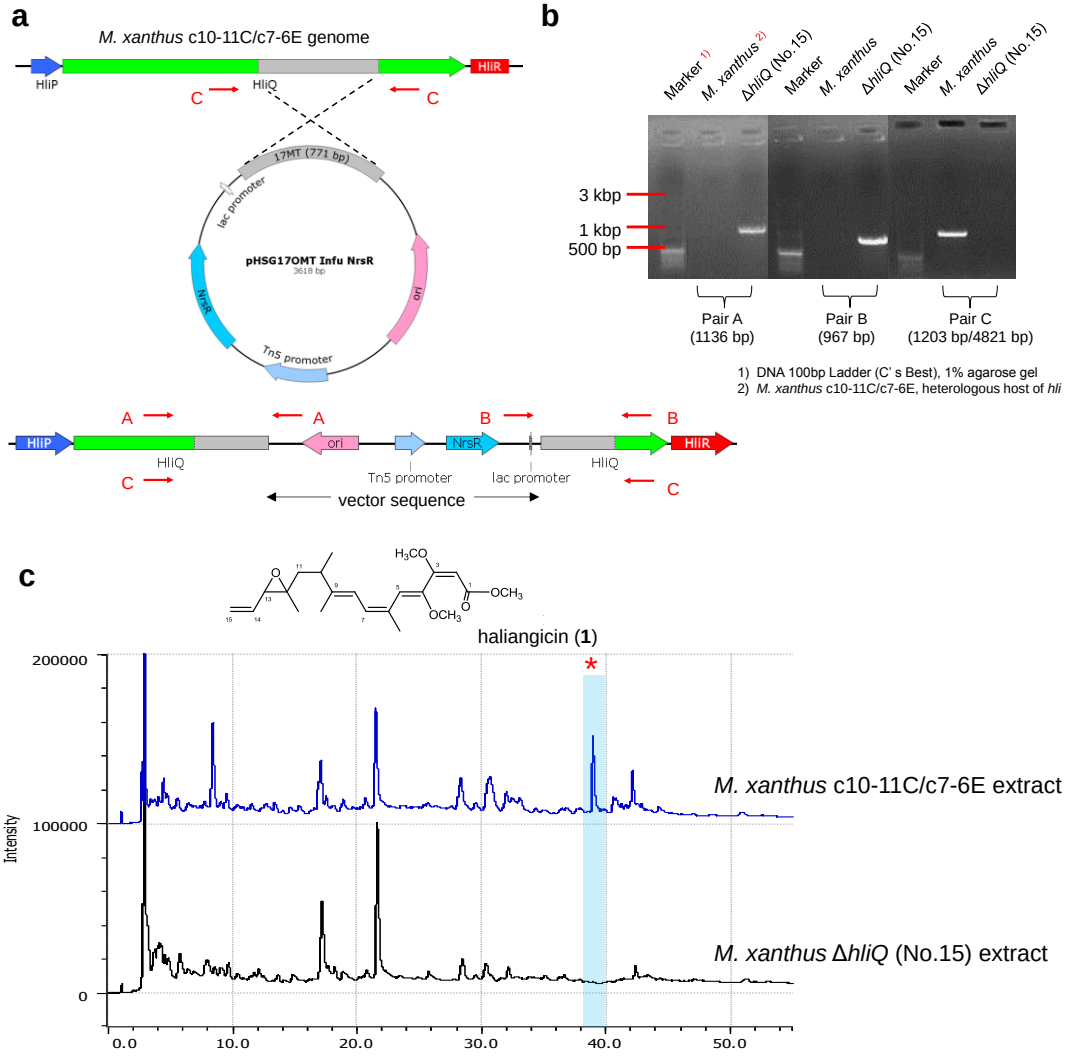


Figure 2-27. Inactivation of HliQ (methoxymalonyl-ACP synthase).

(**a**) Disruption of *hliQ* via single crossover homologous recombination.

(**b**) Verification of Δ*hliQ* mutant by PCR [Pair A: M13f(-47)/ORF17MT Out f; Pair B: M13r(-48)/ORF17MT Out r; Pair C: ORF17MT Out f/ORF17MT Out r].

(**c**) Production profile of *M. xanthus* Δ*hliQ*. Production of haliangicin (**1**) was abolished.

2.3.11. Termination of the carbon elongation by HliE

The absence of thioesterase (TE) domain in the last module (module 5) suggested the termination of the carbon elongation steps by HliE, a metallo-β-lactamase-type thioesterase (MβL-TE). Sequence analysis of HliE revealed a close homology to a group of functionally unknown metallo-β-lactamase family proteins.

Unlike several studied cases of MβL-TEs (*i.e.* ACTE,[65] FccB[66]) that catalyze the release of an aromatic polyketide, HliE might represent a novel type of MβL-TE which hydrolyzes the long chain ketoacyl thioester. Within the boundary of myxobacterial biosynthesis, MβL-TE was discovered for the first time.

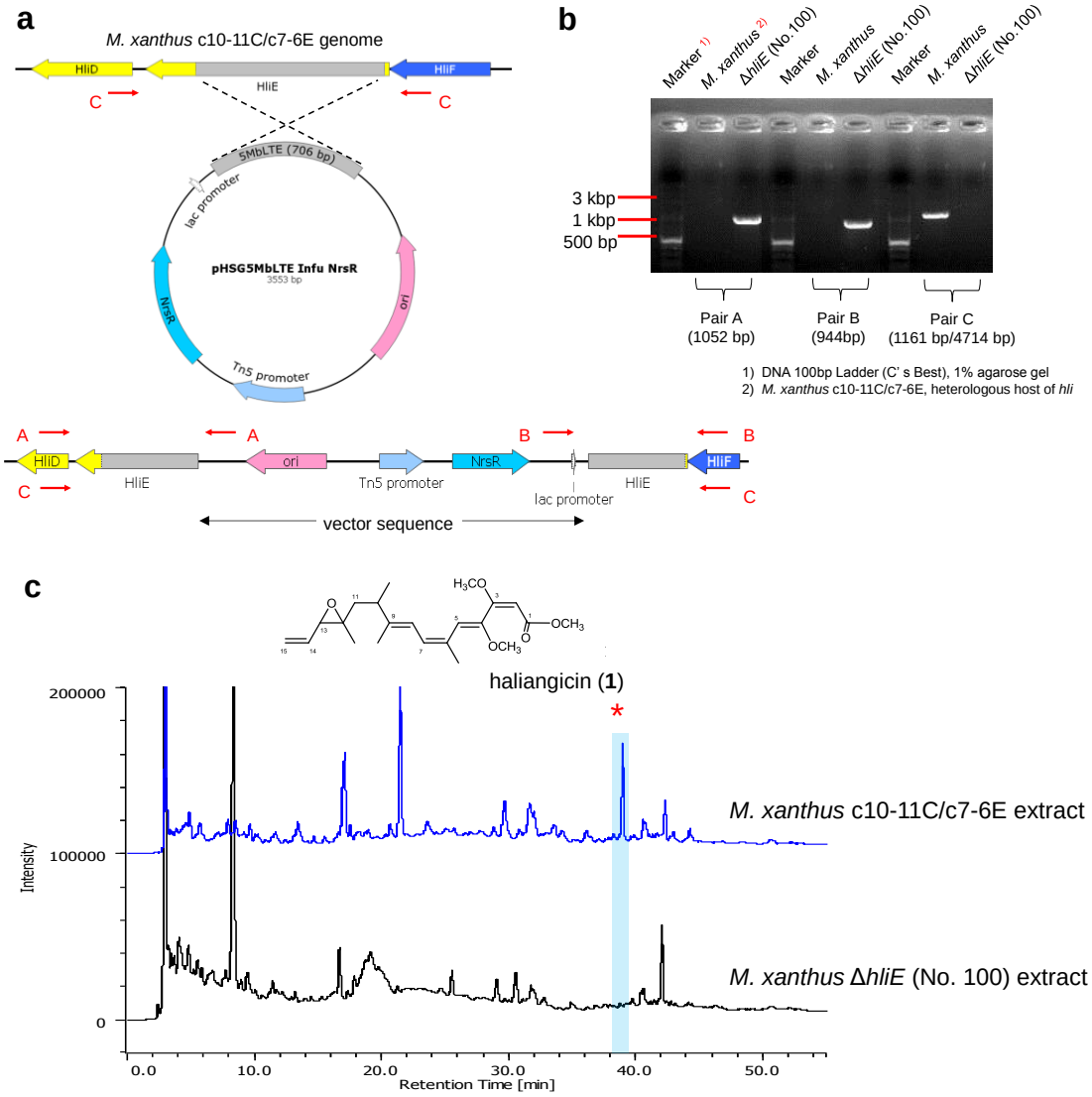


Figure 2-28. Inactivation of HliE (metallo-β-lactamase-type thioesterase).

(**a**) Disruption of *hliE* via single crossover homologous recombination.

(**b**) Verification of Δ*hliE* mutant by PCR [Pair A: M13f(-47)/HliE Out f; Pair B: M13r(-48)/HliE Out r; Pair C: HliE Out f/HliE Out r]

(**c**) Production profile of *M. xanthus* Δ*hliE*. Production of haliangicin (**1**) was abolished.

2.3.12. Post PKS modification in the haliangicin biosynthesis

The nascent carboxylic acid released by HliE then undergoes two steps of post-assembly modifications, *O*-methylation at the carboxyl terminus and epoxidation at the diene terminus, by HliD and HliU, respectively (Fig. 2-13 and 2-29).

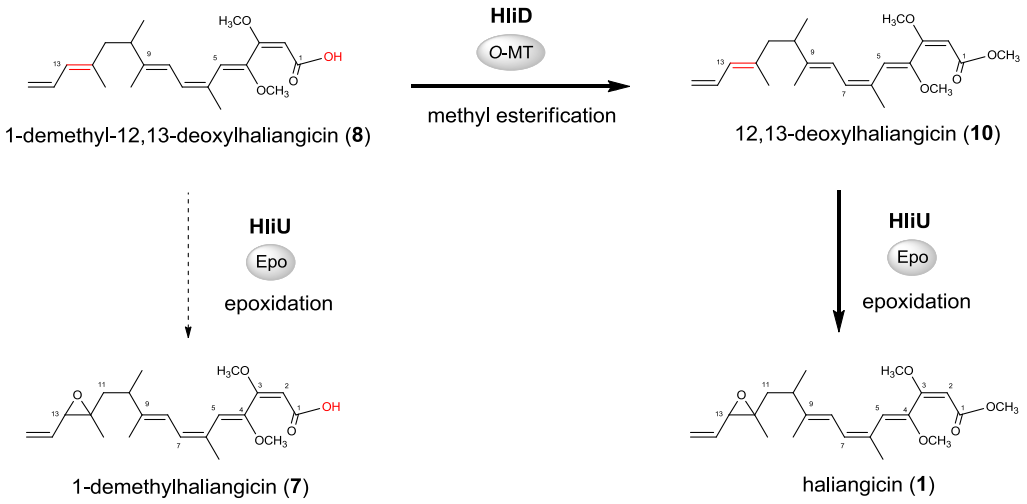


Figure 2-29. Proposed post-PKS modification of haliangicin. The bold arrows indicate the natural biosynthetic pathway that methyl esterification is prior to epoxidation. The dashed arrows show the "unnatural" modification steps due substrate promiscuity of HliU (epoxidase).

Inactivation of *hliD* (Fig. 2-30) in the *M. xanthus* heterologous expression mutant led to the production of two demethylated precursors of haliangicin, 1-*O*-demethylhaliangicin (**7**), 1-*O*-demethyl-12,13-deoxylhaliangicin (**8**) as confirmed by LC-MS analysis (Fig. 2-31). However, because of their labile nature, these acids could not be purified for NMR analyses; consequently, their structures were chemically confirmed by methyl esterification of a crude extract from Δ*hliD* mutant with trimethylsilyldiazomethane, leading to the production of haliangicin (**1**) and 12,13-deoxylhaliangicin (**10**) (Fig. 2-32). Feeding of the crude extract from Δ*hliD* mutant to the HliD-expressing *E. coli* strains (Fig. 2-33a) also led to the production of haliangicin (**1**) and 12,13-deoxylhaliangicin (**10**) as expected (Fig. 2-33b).

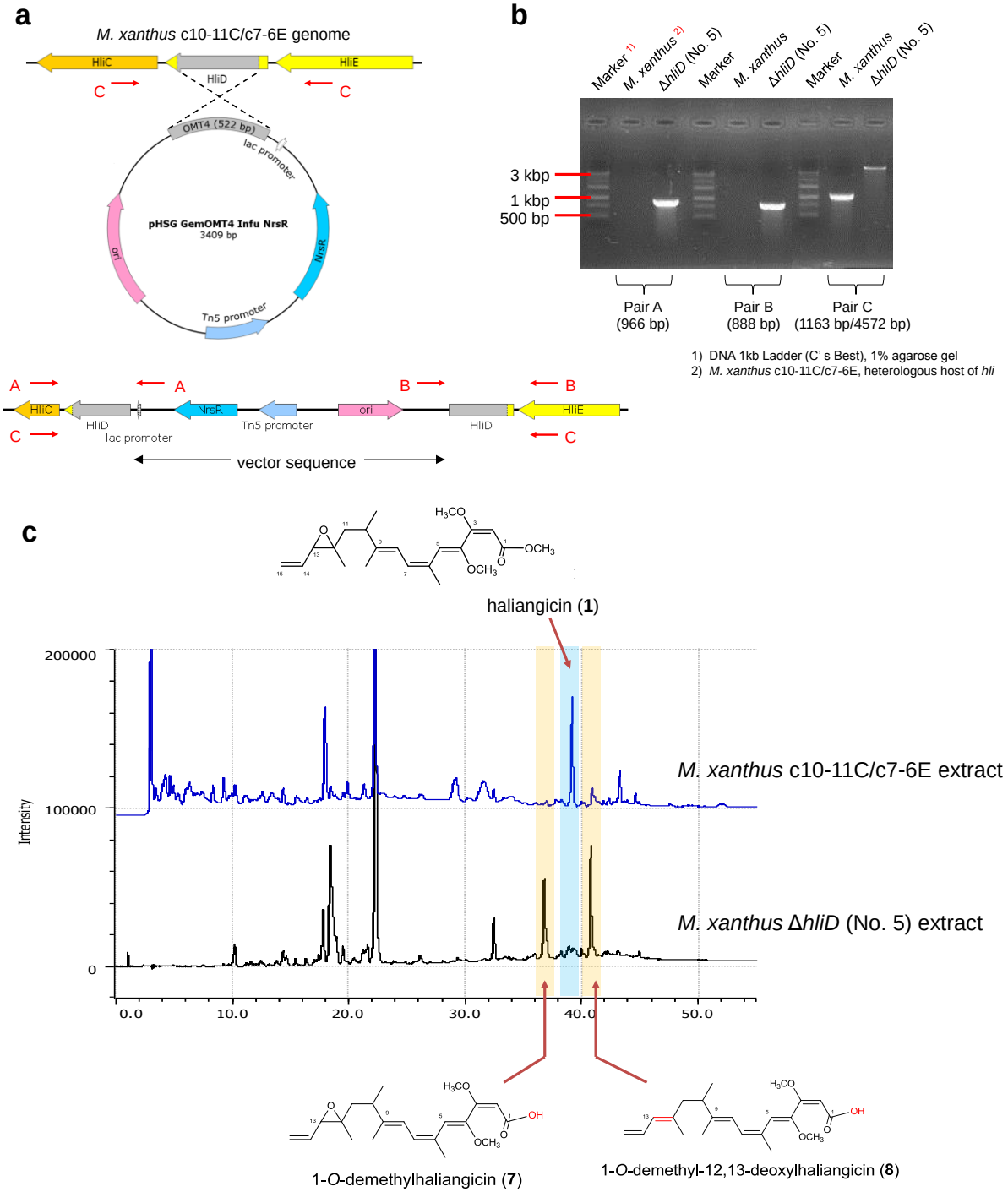


Figure 2-30. Inactivation of HliD (*O*-methyltransferase).

(**a**) Disruption of *hliD* via single crossover homologous recombination.

(**b**) Verification of Δ*hliD* mutant by PCR [Pair A: M13r(-48)/ ORF4-MT-Ex-f1; Pair B: M13f(-47)/ ORF4-MT-Ex-r1; Pair C: ORF4-MT-Ex-f1/ ORF4-MT-Ex-r1].

(**c**) Production profile of *M. xanthus* Δ*hliD*. Production of 1-*O*-demethylhaliangicin (**7**, R_t = 36.5 min) and 1-*O*-demethyl-12,13-deoxylhaliangicin (**8**, R_t = 41 min) were observed.

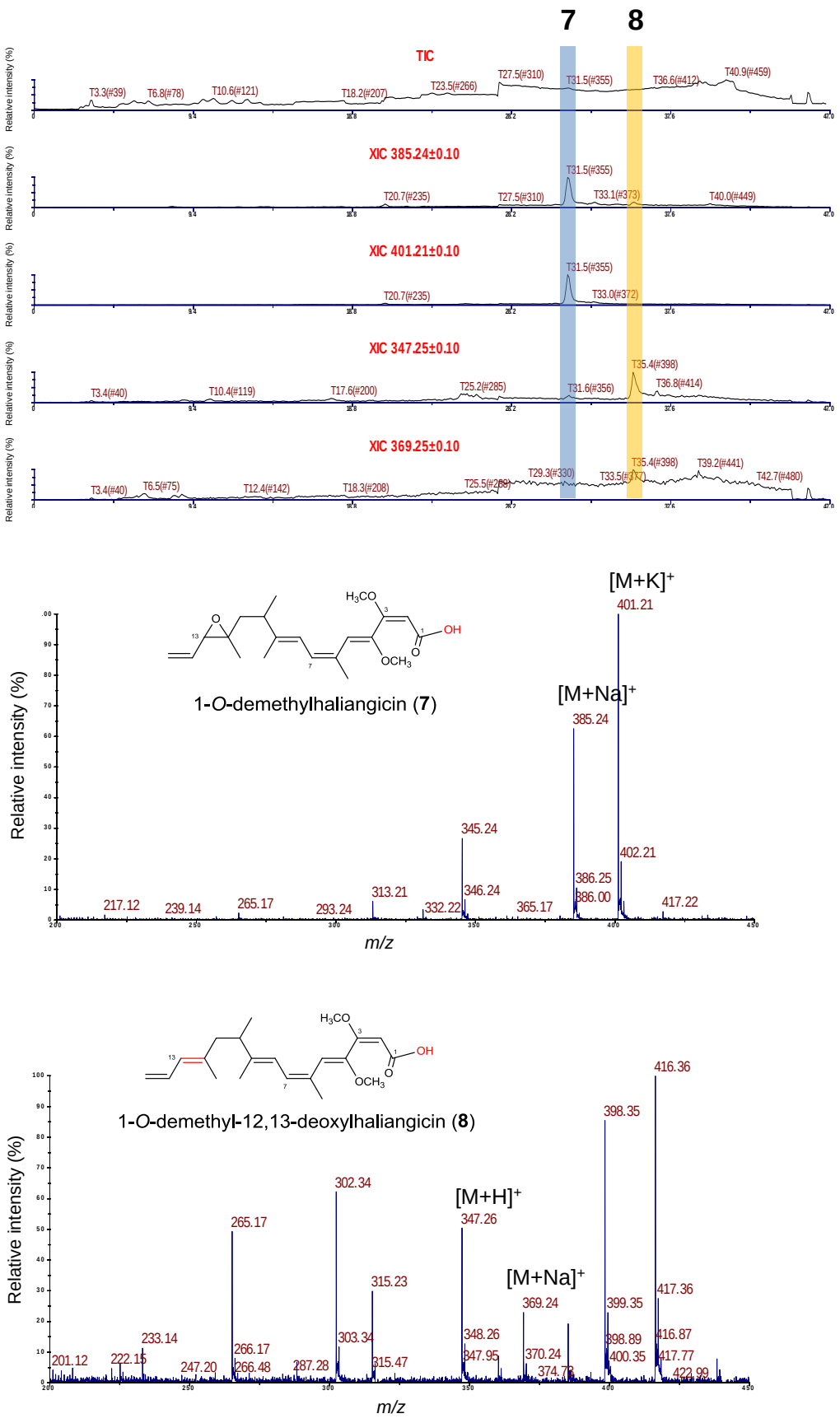


Figure 2-31. LC-MS analysis of *M. xanthus* Δ*hliD* extract.

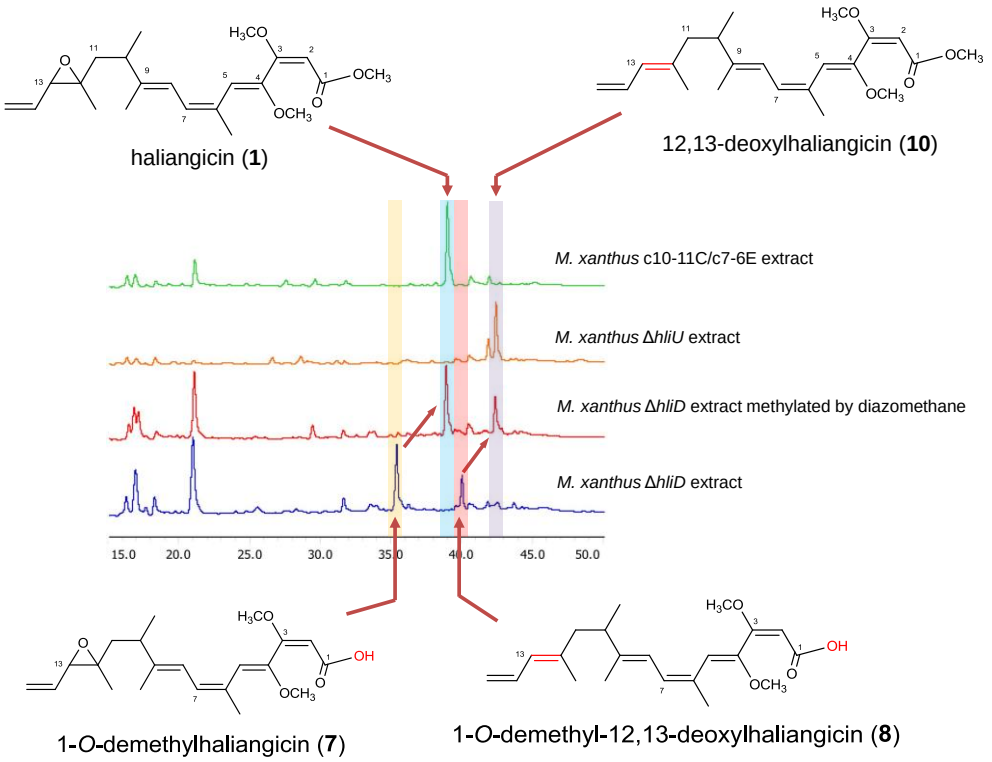


Figure 2-32. HPLC analysis of *M. xanthus* Δ*hliD* extracts before (blue) and after (red) methyl esterification by diazomethane, comparing with authentic sample of 12,13-deoxylhaliangicin (**10**, orange) and haliangicin (**1**, green).

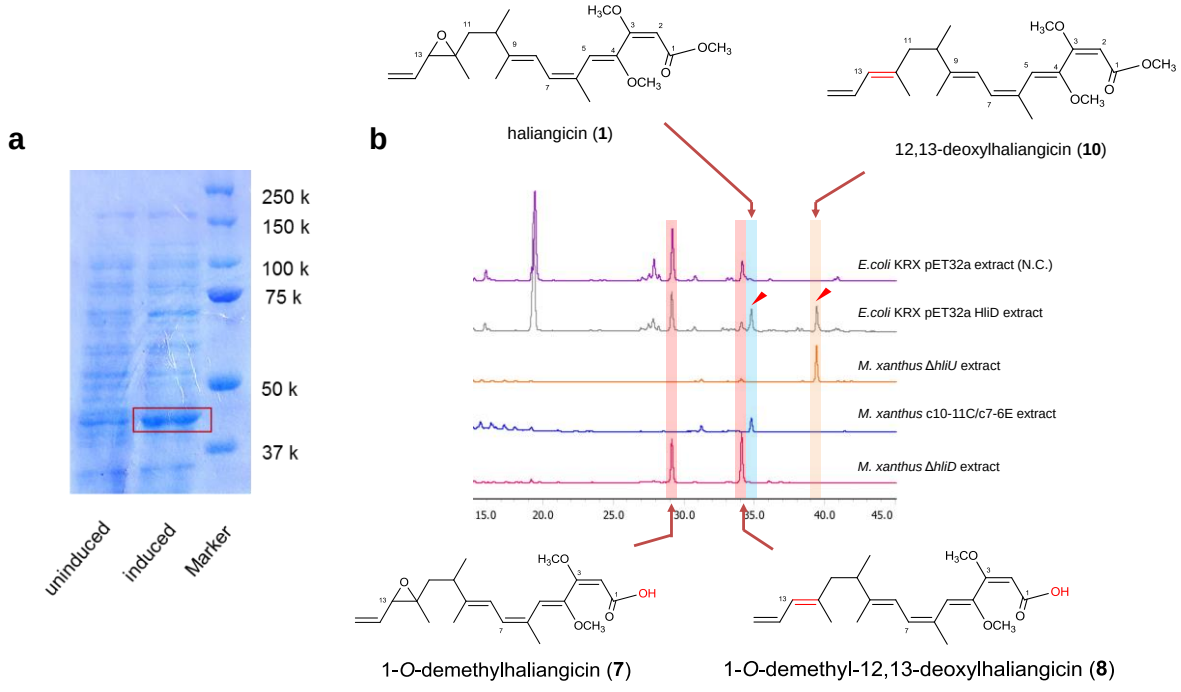


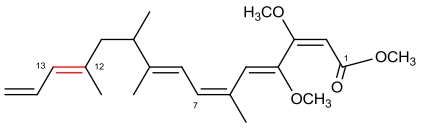
Figure 2-33. Bioconversion of the demethylated precursors of haliangicin. (**a**) Expression of recombinant

HliD; (**b**)HPLC analysis of HliD-expressing *E. coli* extracts fed with demethylated precursors (grey), comparing with authentic sample of 12,13-deoxylhaliangicin (**10**, orange) and haliangicin (**1**, blue) and negative control (*E. coli* harboring empty vector, purple).

The final post-PKS modification, epoxidation, may occur at the terminal diene moiety of the precursor 12,13-deoxylhaliangicin (**10**). HliU was supposed to be the tailoring enzyme for this epoxidation step. Disruption of *hliU* actually accumulated the diene precursor 12,13-deoxylhaliangicin (**10**) as expected (Fig. 2-34, Table 2-20 and Chapter 5 Appendix **5.4**). It is interesting that HliU also accept the demethylated product as it substrate. The production of 1-*O*-demethylhaliangicin (**7**) in Δ*hliD* mutant was supposed to be resulted from the substrate promiscuity of HliU, by which part of the demethylated precursor **8** still can be epoxidated (Fig. 2-13 and 2-30).

Table 2-20. ^{1}H and ^{13}C NMR data for 12,13-deoxylhaliangicin (**10**) in $CDCl_3$ (400 MHz).

	12,13-deoxylhaliangicin (**10**)	
Position	δ_H (mult, *J* in Hz)	δ_C
1	-	167.1
2	5.25 (s)	95.7
3	-	165.9
4	-	148.0
5	5.80 (s)	112.5
6	-	130.1
7	6.04 (d, 11.2)	126.5
8	6.09 (d, 12.0)	120.3
9	-	142.5
10	2.41 (m)	41.5
11	1.99 (dd, 8.4, 13.2), 2.18 (dd, 6.8, 13.6)	46.0
12	-	138.2
13	5.81 (d, 10.4)	127.0
14	6.54 (td, 10.4, 16.8)	133.4
15	4.96 (dd, 1.6, 10.4), 5.07 (dd, 1.6, 16.8)	114.6
1-OMe	3.68 (s)	51.3
3-OMe	3.73 (s)	56.0
4-OMe	3.55 (s)	57.2
6-Me	2.09 (s)	23.6
9-Me	1.70 (s)	13.4
10-Me	0.96 (d, 6.8)	18.9
12-Me	1.73 (s)	16.5



12,13-deoxylhaliangicin (**10**)

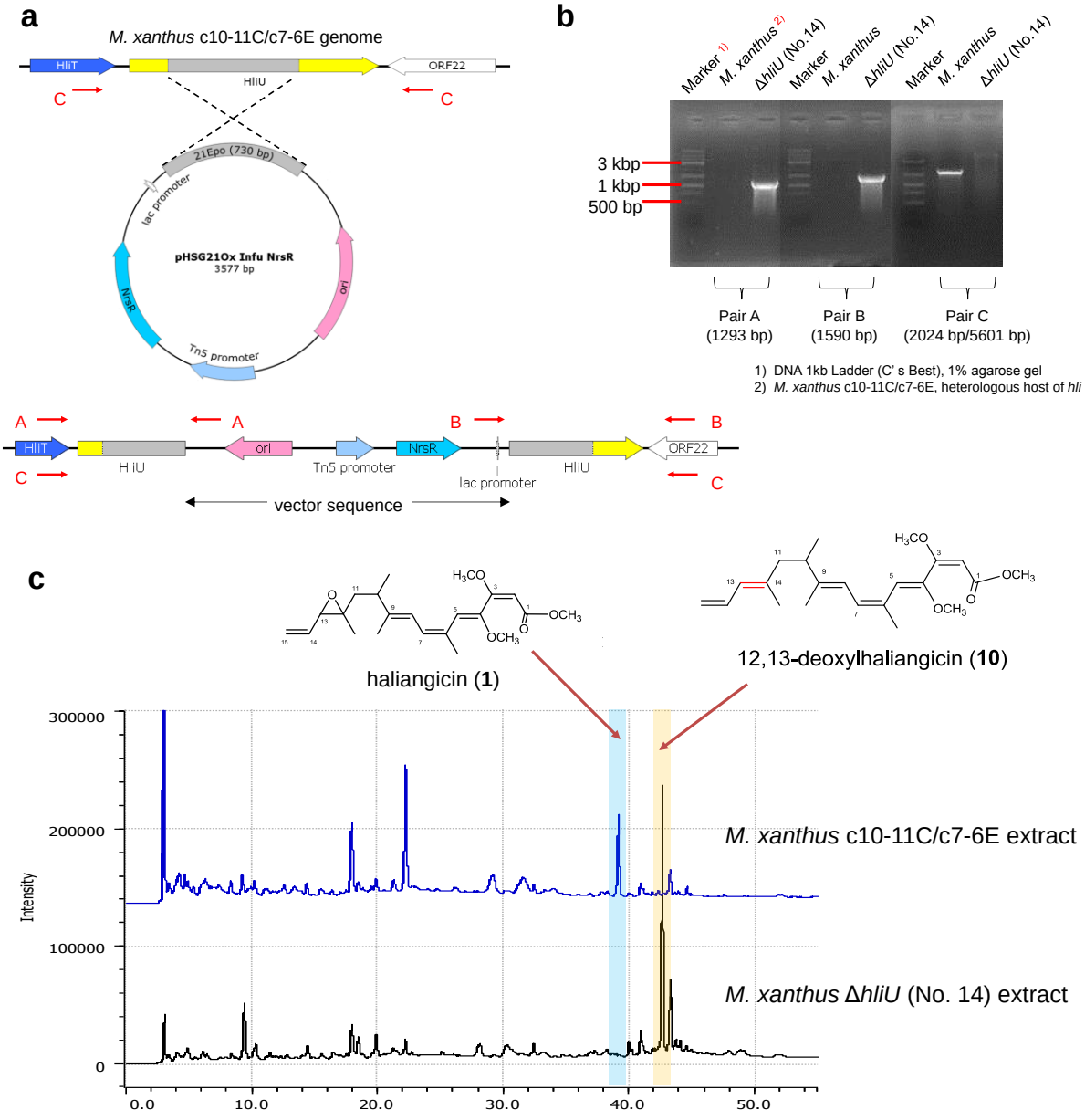


Figure 2-34. Inactivation of HliU (epoxidase).

(**a**) Disruption of *hliU* via single crossover homologous recombination.

(**b**) Verification of Δ*hliU* mutant by PCR [Pair A: M13f(-47)/ORF21Ox-Out F2; Pair B: M13r(-48)/ORF21Ox-Out R; Pair C: ORF21Ox-Out F2/ORF21Ox-Out R].

(**c**) Production profile of *M. xanthus* Δ*hliU*. Production of 12,13-deoxylhaliangicin (**10**, t_R = 43 min) was observed.

2.3.13. Orf 22

Orf22 downstream to *hliU* shows a similarity to the mechanosensitive ion channel protein-encoding genes. Orf22 knockout strain still produces haliangicin (**1**) (Fig. 2-35), which indicated it to be unrelated to the haliangicin biosynthesis.

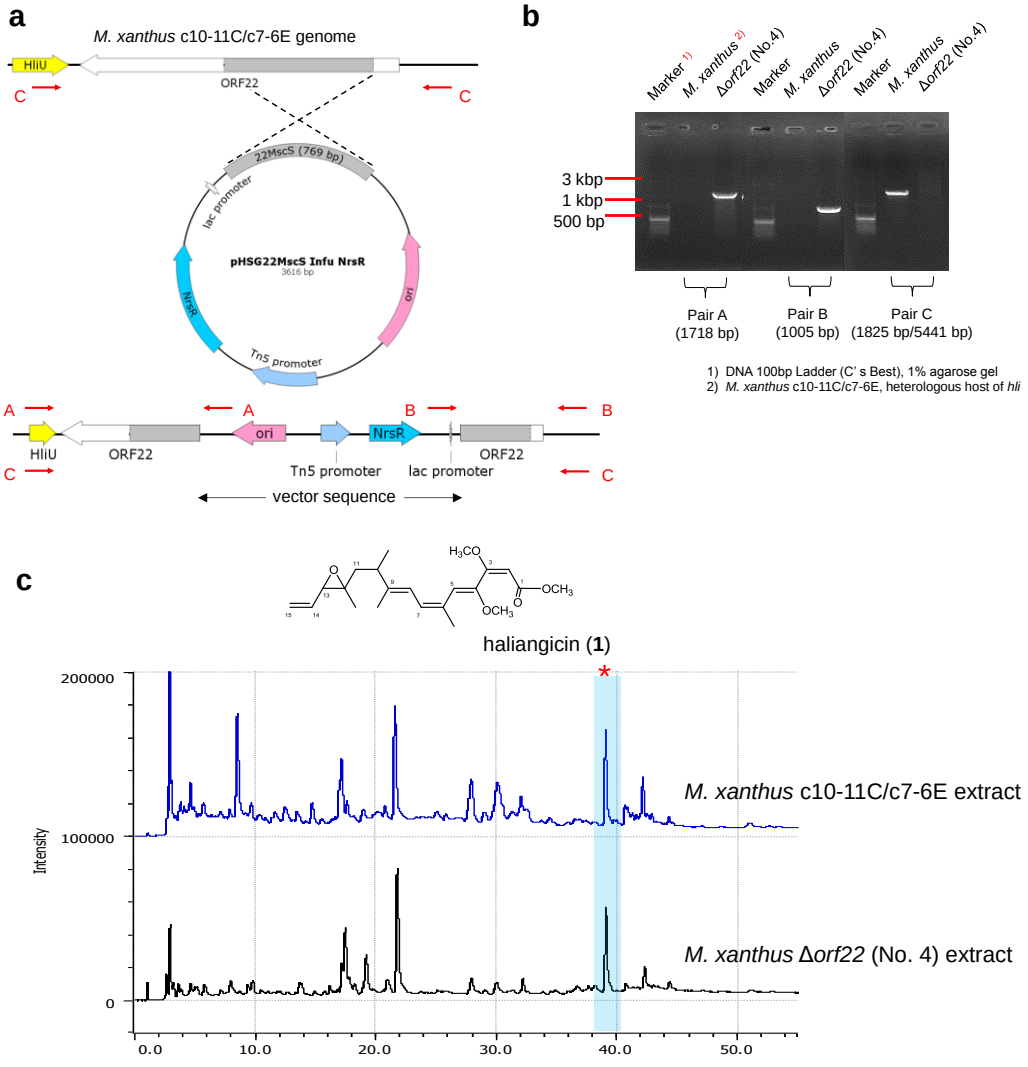


Figure 2-35. Inactivation of Orf22 (mechanosensitive ion channel).

(**a**) Disruption of *orf22* via single crossover homologous recombination.

(**b**) Verification of Δ*orf22* mutant by PCR [Pair A: M13f(-47)/ORF22MscS Ex-f; Pair B: M13r(-48)/ORF22MscS Ex-r; Pair C: ORF22MscS Ex-f/ORF22MscS Ex-r].

(**c**) Production profile of *M. xanthus* Δ*orf22*. Production of haliangicin (**1**) was still observed.

2.3.14. Biological activities of haliangicin analogues.

Haliangicin (**1**) is known to show potent inhibitory activity against the plant pathogenic oomycete *Phytophthora capsici*; in this activity, the terminal β-methoxyacrylate moiety serves as the pharmacophore.[20] Because the unnatural analogues **9** and **10** obtained in this study are modified at the terminus opposite the β-methoxyacrylate, additional biological information might be obtained. In a test using *P. capsici*, analogue **9**, which is saturated at the 14,15-olefin, showed the same activity as **1** and the versatile fungicide metalaxyl, whereas analogue **10**, lacking the 12,13-epoxide moiety, was 30-fold less active than the other analogues (Table 2-21), suggesting that the epoxide plays a key role in this activity. It is noteworthy that these compounds showed potent cytotoxicity against a tumour cell line and that analogue **10** was more active than the natural compound **1**. These compounds showed no inhibitory effects on other microorganisms tested.

Table 2-21. Biological evaluation of haliangicin (**1**) and its unnatural analogues (**9** and **10**).

Organisms	**1**	**9**	**10**	PC
Phytophthora capsici (μg disk^{-1})[a]	0.1	0.1	3.0	0.1
Candida rugosa (MIC, μg mL^{-1})[b]	>32	>32	>32	0.13
Bacillus subtilis (MIC, μg mL^{-1})[b]	>32	>32	>32	0.08
Escherichia coli (MIC, μg mL^{-1})[b]	>32	>32	>32	8.0
HeLa S_3 cells (IC_{50}, nM)[c]	41	55	17	8.6

[a] Disk diffusion test indicating minimum doses for a definite inhibition zone. Positive control (PC): metalaxyl. [b] MIC test. PC: amphotericin B for *C. rugosa* and ampicillin for *B. subtilis* and *E. coli*. [c] MTT test. PC: paclitaxel.

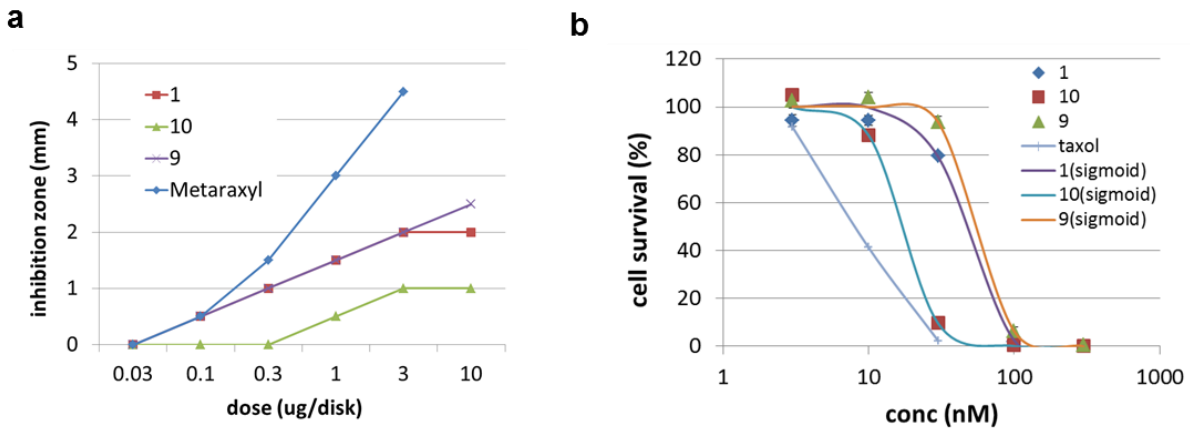


Figure 2-36. Biological activities of haliangicin analogues.

(**a**) Anti-oomycete activities against *Phytophthora capsici*.

(**b**) Cytotoxicity against HeLa-S3 cells, markers indicate average values with SE, and lines indicate theoretical sigmoid curves.

2.4. Summary

Marine myxobacterial biosynthetic machinery was an unexplored field due to the fastidious nature of marine myxobacteria (slow growth, low secondary metabolism, stickiness, etc.). In this chapter, my works on the haliangicin biosynthetic assembly line provide a first insight into the marine myxobacterial biosynthesis. The utilization of the heterologous expression system of haliangicin (**1**) not only improved the productivity (10-fold greater amount and 3-fold faster in growth speed compared with the original producer), but also generated novel unnatural analogues (compound **7**, **8**, **9** and **10**), one of which was a more potent cytotoxin (12,13-deoxylhaliangicin, **10**) than the natural compound **1**. In addition, the easier access to the biosynthesis machinery of haliangicin (**1**) in a more culturable organism also led to the functional characterization of a unique dehydrogenase that introduces the terminal alkene moiety of **1**, thereby widening our understanding of the biosynthesis mechanism underlying the formation of terminal alkene groups.

The *hli* locus exhibits some atypical features that have not been observed in the terrestrial myxobacterial biosynthetic machinery. The relatively disorganized gene architecture and the deficient number of module

in haliangicin biosynthesis once more suggest the complexity of myxobacterial PKS system. Genomic information on the native producer *Haliangium ochraceum* suggests the presence of six other latent biosynthetic gene clusters, which may prompts us to analyze them by applying the present expertise to produce novel metabolites that are not obtained by cultivation of the native marine myxobacterium.

Reference

1. Hofle GH, Bedorf N, Steinmetz H, Schomburg D, Gerth K, Reichenbach H. Epothilone A and B - Novel 16-membered macrolides with cytotoxic activity: Isolation, crystal structure, and conformation in solution. *Angew. Chem. Int. Edit.* **35**, 1567-1569 (1996).

2. Nettles JH, Li HL, Cornett B, Krahn JM, Snyder JP, Downing KH. The binding mode of epothilone A on alpha,beta-tubulin by electron crystallography. *Science* **305**, 866-869 (2004).

3. Weissman KJ, Muller R. A brief tour of myxobacterial secondary metabolism. *Bioorg. Med. Chem.* **17**, 2121-2136 (2009).

4. Wenzel SC, Muller R. Myxobacteria-'microbial factories' for the production of bioactive secondary metabolites. *Mol. Biosyst.* **5**, 567-574 (2009).

5. Iizuka T, Jojima Y, Fudou R, Tokura M, Hiraishi A, Yamanaka S. *Enhygromyxa salina* gen. nov., sp nov., a slightly halophilic myxobacterium isolated from the coastal areas of Japan. *Syst. Appl. Microbiol.* **26**, 189-196 (2003).

6. Iizuka T, Jojima Y, Hayakawa A, Fujii T, Yamanaka S, Fudou R. *Pseudenhygromyxa salsuginis* gen. nov., sp nov., a myxobacterium isolated from an estuarine marsh. *Int. J. Syst. Evol. Microbiol.* **63**, 1360-1369 (2013).

7. Fudou R, Jojima Y, Iizuka T, Yamanaka S. *Haliangium ochraceum* gen. nov., sp nov and *Haliangium tepidum* sp nov.: Novel moderately halophilic myxobacteria isolated from coastal saline environments. *J. Gen. Appl. Microbiol.* **48**, 109-115 (2002).

8. Schaeberle TF, *et al.* Marine Myxobacteria as a Source of Antibiotics-Comparison of Physiology, Polyketide-Type Genes and Antibiotic Production of Three New Isolates of *Enhygromyxa salina*. *Mar. Drugs* **8**, 2466-2479 (2010).

9. Felder S, *et al.* Salimabromide: Unexpected Chemistry from the Obligate Marine Myxobacterium *Enhygromxya salina*. *Chem. Eur. J.* **19**, 9319-9324 (2013).

10. Felder S, Kehraus S, Neu E, Bierbaum G, Schaeberle TF, Koenig GM. Salimyxins and Enhygrolides: Antibiotic, Sponge-Related Metabolites from the Obligate Marine Myxobacterium *Enhygromyxa salina*.

Chembiochem **14**, 1363-1371 (2013).

11. Fudou R, Iizuka T, Yamanaka S. Haliangicin, a novel antifungal metabolite produced by a marine myxobacterium 1. Fermentation and biological characteristics. *J. Antibiot.* **54**, 149-152 (2001).
12. Fudou R, Iizuka T, Sato S, Ando T, Shimba N, Yamanaka S. Haliangicin, a novel antifungal metabolite produced by a marine myxobacterium 2. Isolation and structural elucidation. *J. Antibiot.* **54**, 153-156 (2001).
13. Komaki H, *et al.* PCR Detection of Type I Polyketide Synthase Genes in Myxobacteria. *Appl. Environ. Microbiol.* **74**, 5571-5574 (2008).
14. Suzuki A. Studies on the biosynthetic genes of myxobacterial antibiotics. Master thesis. Nagoya University (2008).
15. Hyun CG, Kim SS, Sohng JK, Hahn JJ, Kim JW, Suh JW. An efficient approach for cloning the dNDP-glucose synthase gene from actinomycetes and its application in Streptomyces spectabilis, a spectinomycin producer. *FEMS Microbiol. Lett.* **183**, 183-189 (2000).
16. Sharan SK, Thomason LC, Kuznetsov SG, Court DL. Recombineering: a homologous recombination-based method of genetic engineering. *Nat. Protoc*. **4**, 206-223 (2009).
17. Hodgkin J, Kaiser D. Cell-to-cell Stimulation of Movement in Nonmotile Mutants of Myxococcus. *Proc. Natl. Acad. Sci. U. S. A.* **74**, 2938-2942 (1977).
18. Ojika M, *et al.* Cystothiazoles A and B, new bithiazole-type antibiotics from the myxobacterium *Cystobacter fuscus*. *J. Antibiot.* **51**, 275-281 (1998).
19. Reichenbach H, Dworkin M. In: *The prokaryotes* (eds Balows A, Truper HG, Dworkin M, Harder W, Schleifer K). 2nd edn. Springer (1992).
20. Kundim BA, *et al.* New haliangicin isomers, potent antifungal metabolites produced by a marine myxobacterium. *J. Antibiot.* **56**, 630-638 (2003).
21. Warming S, Costantino N, Court DL, Jenkins NA, Copeland NG. Simple and highly efficient BAC recombineering using galK selection. *Nucleic Acids Res.* **33**, e36 (2005).
22. Malonek S, Rojas MC, Hedden P, Hopkins P, Tudzynski B. Restoration of gibberellin production in

Fusarium proliferatum by functional complementation of enzymatic blocks. *Appl. Environ. Microbiol.* **71**, 6014-6025 (2005).

23. Marchler-Bauer A, *et al.* CDD: NCBI's conserved domain database. *Nucleic Acids Res.* **43**, D222-D226 (2015).

24. Marchler-Bauer A, *et al.* CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res.* **39**, D225-D229 (2011).

25. Blin K, *et al.* antiSMASH 2.0-a versatile platform for genome mining of secondary metabolite producers. *Nucleic Acids Res.* **41**, W204-W212 (2013).

26. Weber T, *et al.* antiSMASH 3.0-a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res.* **43**, W237-W243 (2015).

27. Ziemert N, Podell S, Penn K, Badger JH, Allen E, Jensen PR. The Natural Product Domain Seeker NaPDoS: A Phylogeny Based Bioinformatic Tool to Classify Secondary Metabolite Gene Diversity. *Plos One* **7**, 9 (2012).

28. Yu TW, *et al.* The biosynthetic gene cluster of the maytansinoid antitumor agent ansamitocin from *Actinosynnema pretiosum*. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 7968-7973 (2002).

29. Wenzel SC, *et al.* On the biosynthetic origin of methoxymalonyl-acyl carrier protein, the substrate for incorporation of "glycolate" units into ansamitocin and soraphen A. *J. Am. Chem. Soc.* **128**, 14325-14336 (2006).

30. Hill AM, Harris JP, Siskos AP. Investigation of glycerol incorporation into soraphen A. *Chem. Comm.*, 2361-2362 (1998).

31. Ligon J, *et al.* Characterization of the biosynthetic gene cluster for the antifungal polyketide soraphen A from *Sorangium cellulosum* So ce26. *Gene* **285**, 257-267 (2002).

32. Wu K, Chung L, Revill WP, Katz L, Reeves CD. The FK520 gene cluster of *Streptomyces hygroscopicus* var. ascomyceticus (ATCC 14891) contains genes for biosynthesis of unusual polyketide extender units. *Gene* **251**, 81-90 (2000).

33. Ono M, *et al.* Structures and biosynthesis of Aflastatins: Novel inhibitors of aflatoxin production by

Aspergillus parasiticus. *J. Antibiot.* **51**, 1019-1028 (1998).

34. Muller S, *et al.* Biosynthesis of Crocacin Involves an Unusual Hydrolytic Release Domain Showing Similarity to Condensation Domains. *Chem. Biol*. **21**, 855-865 (2014).
35. Pulsawat N, Kitani S, Nihira T. Characterization of biosynthetic gene cluster for the production of virginiamycin M, a streptogramin type A antibiotic, in *Streptomyces virginiae*. *Gene* **393**, 31-42 (2007).
36. Simunovic V, Zapp J, Rachid S, Krug D, Meiser P, Muller R. Myxovirescin A biosynthesis is directed by hybrid polyketide synthases/nonribosomal peptide synthetase, 3-hydroxy-3-methylglutaryl-CoA synthases, and trans-acting acyltransferases. *Chembiochem* **7**, 1206-1220 (2006).
37. Erol O, *et al.* Biosynthesis of the Myxobacterial Antibiotic Corallopyronin A. *Chembiochem* **11**, 1253-1265 (2010).
38. Gurney R, Thomas CM. Mupirocin: biosynthesis, special features and applications of an antibiotic from a Gram-negative bacterium. *Appl. Microbiol. Biotechnol.* **90**, 11-21 (2011).
39. Mattheus W, *et al.* Isolation and Purification of a New Kalimantacin/Batumin-Related Polyketide Antibiotic and Elucidation of Its Biosynthesis Gene Cluster. *Chem. Biol.* **17**, 149-159 (2010).
40. Jahns C, *et al.* Pellasoren: Structure Elucidation, Biosynthesis, and Total Synthesis of a Cytotoxic Secondary Metabolite from *Sorangium cellulosum*. *Angew. Chem. Int. Edit.* **51**, 5239-5243 (2012).
41. Ziemert N, Podell S, Penn K, Badger JH, Allen E, Jensen PR. The Natural Product Domain Seeker NaPDoS: A Phylogeny Based Bioinformatic Tool to Classify Secondary Metabolite Gene Diversity. *PLoS One* **7**, (2012).
42. Zhang YM, Hurlbert J, White SW, Rock CO. Roles of the active site water, histidine 303, and phenylalanine 396 in the catalytic mechanism of the elongation condensing enzyme of *Streptococcus pneumoniae*. *J. Biol. Chem.* **281**, 17390-17399 (2006).
43. Yadav G, Gokhale RS, Mohanty B. Computational approach for prediction of domain organization and substrate specificity of modular polyketide synthases. *J. Mol. Biol.* **328**, 335-363 (2003).
44. Otsuka M, Ichinose K, Fujii I, Ebizuka Y. Cloning, sequencing, and functional analysis of an iterative type I polyketide synthase gene cluster for biosynthesis of the antitumor chlorinated polyenone

neocarzilin in "*Streptomyces carzinostaticus*". *Antimicro. Agents Chemother.* **48**, 3468-3476 (2004).

45. Potharla VY, Wesener SR, Cheng YQ. New Insights into the Genetic Organization of the FK228 Biosynthetic Gene Cluster in Chromobacterium violaceum No. 968. *Appl Environ Microbiol* **77**, 1508-1511 (2011).

46. Rachid S, Scharfe M, Blocker H, Weissman KJ, Muller R. Unusual Chemistry in the Biosynthesis of the Antibiotic Chondrochlorens. *Chem. Biol.* **16**, 70-81 (2009).

47. Bevitt DJ, Cortes J, Haydock SF, Leadlay PF. 6-Deoxyerythronolide-B synthase-2 from *Saccharopolyspora erythraea* - Cloning of the Structural Gene, Sequence-analysis and Inferred Domain-structure of the Multifunctional Enzyme. *Eur. J. Biochem.* **204**, 39-49 (1992).

48. Hoffmann T, Muller S, Nadmid S, Garcia R, Muller R. Microsclerodermins from Terrestrial Myxobacteria: An Intriguing Biosynthesis Likely Connected to a Sponge Symbiont. *J. Am. Chem. Soc.* **135**, 16904-16911 (2013).

49. Kwan DH, Schulz F. The Stereochemistry of Complex Polyketide Biosynthesis by Modular Polyketide Synthases. *Molecules* **16**, 6092-6115 (2011).

50. Caffrey P. Conserved amino acid residues correlating with ketoreductase stereospecificity in modular polyketicle synthases. *Chembiochem* **4**, 654-657 (2003).

51. Moore BS, Hertweck C. Biosynthesis and attachment of novel bacterial polyketide synthase starter units. *Nat. Prod. Rep*. **19**, 70-99 (2002).

52. Pereda A, Summers RG, Stassi DL, Ruan XA, Katz L. The loading domain of the erythromycin polyketide synthase is not essential for erythromycin biosynthesis in *Saccharopolyspora erythraea*. *Microbiol.-(UK)* **144**, 543-553 (1998).

53. Lau J, Cane DE, Khosla C. Substrate specificity of the loading didomain of the erythromycin polyketide synthase. *Biochem.* **39**, 10514-10520 (2000).

54. Ikeda H, Nonomiya T, Usami M, Ohta T, Omura S. Organization of the biosynthetic gene cluster for the polyketide anthelmintic macrolide avermectin in *Streptomyces avermitilis*. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 9509-9514 (1999).

55. Chang ZX, *et al.* Biosynthetic pathway and gene cluster analysis of curacin A, an antitubulin natural product from the tropical marine cyanobacterium *Lyngbya majuscula*. *J. Nat. Prod.* **67**, 1356-1367 (2004).

56. Ubukata M, Cheng XC, Uzawa J, Isono K. Biosynthesis of the dialkylmaleic anhydride-containing antibiotics, tautomycin and tautomycetin. *J. Chem. Soc.-Perkin Transactions 1*, 2399-2404 (1995).

57. Wu M, *et al.* Structure, synthesis, and biological properties of kalkitoxin, a novel neurotoxin from the marine cyanobacterium *Lyngbya majuscula*. *J. Am. Chem. Soc.* **122**, 12041-12042 (2000).

58. Gu LC, *et al.* Polyketide Decarboxylative Chain Termination Preceded by *O*-Sulfonation in Curacin A Biosynthesis. *J. Am. Chem. Soc.* **131**, 16033-+ (2009).

59. Li WL, Luo YG, Ju JH, Rajski SR, Osada H, Shent B. Characterization of the Tautomycetin Biosynthetic Gene Cluster from *Streptomyces griseochromogenes* Provides New Insight into Dialkylmaleic Anhydride Biosynthesis. *J. Nat. Prod.* **72**, 450-459 (2009).

60. Mo S, *et al.* Biosynthesis of the Allylmalonyl-CoA Extender Unit for the FK506 Polyketide Synthase Proceeds through a Dedicated Polyketide Synthase and Facilitates the Mutasynthesis of Analogues. *J. Am. Chem. Soc.* **133**, 976-985 (2011).

61. Dehn R, *et al.* Molecular Basis of Elansolid Biosynthesis: Evidence for an Unprecedented Quinone Methide Initiated Intramolecular Diels-Alder Cycloaddition/Macrolactonization. *Angew. Chem. Int. Edit.* **50**, 3882-3887 (2011).

62. Wakimoto T, *et al.* Calyculin biogenesis from a pyrophosphate protoxin produced by a sponge symbiont. *Nat. Chem. Biol.* **10**, 648-U193 (2014).

63. Calderone CT. Isoprenoid-like alkylations in polyketide biosynthesis. *Nat. Prod. Rep.* **25**, 845-853 (2008).

64. Muller S, *et al.* Biosynthesis of crocacin involves an unusual hydrolytic release domain showing similarity to condensation domains. *Chem. Biol.* **21**, 855-865 (2014).

65. Awakawa T, *et al.* Physically Discrete beta-Lactamase-Type Thioesterase Catalyzes Product Release in Atrochrysone Synthesis by Iterative Type I Polyketide Synthase. *Chem. Biol.* **16**, 613-623 (2009).

66. Konig CC, *et al.* Bacterium Induces Cryptic Meroterpenoid Pathway in the Pathogenic Fungus *Aspergillus fumigatus*. *Chembiochem* **14**, 938-942 (2013).

Chapter 3. Exploration of other PKS/NRPS biosynthetic gene clusters in the genome of the marine myxobacterium *Haliangium ochraceum*

3.1. Introduction

Myxobacteria are unique Gram-negative bacteria characterized by gliding on solid surfaces and formation of multicellular fruiting body.[1-3] Although long been regarded as terrestrial microorganisms, several halophilic myxobacterial strains have been noticed recently[4-8] and found to be excellent producers of novel secondary metabolites, such as polyketides (PKs), non-ribosomal peptides (NRPs) and their hybrids. A recent report suggested that the polyketide synthase (PKSs) gene sequences in marine-derived myxobacterial isolates are highly novel (all of the sequences were less than 70% similar to the known ones[9]). However, studies of the secondary metabolites of marine myxobacteria have been hampered by unfavorable factors such as difficulties in the isolation from environment, slow growth rates, strong tendency of cell aggregation and poor metabolite productivity. Up to now, there are only a few groups of antibiotics that have been identified from the marine environment, including haliangicins[10-12], miuraenamides[13, 14], salimabromides[15], salimyxins and enhygrolides[16].

Haliangium ochraceum SMP-2 is the first marine myxobacterium, from which a potent antifungal agent haliangicin was discovered.[10, 11, 17] Since the isolation in 2002 of *H. ochraceum* and its only identified metabolite haliangicin, there have been no further reports on its secondary metabolome probably due to the fastidious nature of this microorganism. Recently, during the ongoing investigation on the secondary metabolome of *H. ochraceum* in our laboratory, a new PKS-NRPS hybrid compound, haliamide (**2**), was isolated and identified (Fig. 3-1a).[18] The biosynthetic building blocks of haliamide (**2**) were identified by feeding experiments, which revealed one benzoate, one alanine, two propionates, one acetate and one acetate-derived terminal methylene. (Fig. 3-1b)

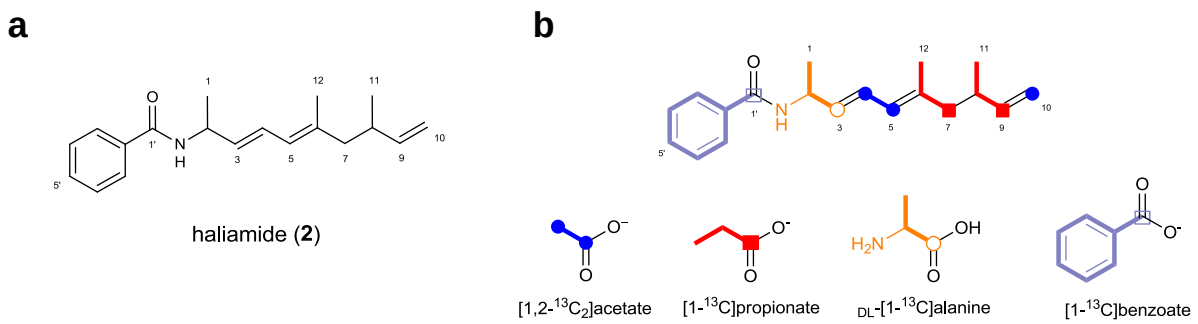


Figure 3-1. Haliamide, a new PKS-NRPS hybrid from the *Haliangium ochraceum*.

(**a**) Structure of haliamide (**2**).

(**b**) Biosynthetic building blocks of haliamide (**2**) deduced from feeding experiments.

In 2010, the complete genome sequence of *Haliangium ochraceum* SMP-2T was published by Ivanova *et al*., showing 9.45 Mbp lengths with a 69.5% GC content.[19] It represents the first complete genome sequence of a member of the myxococcales suborder *Nannocystineae*. Accessing to the genome of *H. ochraceum* provided much invaluable information on the biosynthetic genes it encodes, which may greatly facilitate the studies of the secondary metabolites in this bacterium.

In this chapter, I performed the analyses of the biosynthetic genes in *H. ochraceum* with emphasis on the PKS/NRPS biosynthetic loci. Mining of the genome of *H. ochraceum* led to the discovery of six PKS/NRPS biosynthetic gene clusters as well as the haliangicin biosynthetic gene cluster (*hli*). Among these clusters, the biosynthetic gene cluster of haliamide (*hla*, 21.7 kbp) was characterized, allowing an establishment of a model for the haliamide biosynthesis. The remaining five PKS/NRPS biosynthetic gene clusters for other metabolites were analyzed by bioinformatics tools, from which a putative biosynthetic gene cluster for aurafuron A (**3**) was discovered.

3.2. Materials and methods

3.2.1. General

Flash chromatography was carried out with a medium-pressure gradient system equipped with a Pump Module C-605 and a Pump Manager C-615 (BÜCHI, Switzerland). Preparative HPLC was performed on a high-pressure gradient system equipped with PU-2087 plus pumps and a UV-2075 plus detector (Jasco, Tokyo). ^{1}H and ^{13}C NMR spectra were recorded at 27 °C on an Avance 400 (400 MHz for ^{1}H) spectrometer (Bruker, Rheinstetten, Germany). Chemical shifts (ppm) were referenced to the solvent ($CDCl_3$) peaks at δ_H 7.26 ppm (residual $CHCl_3$) and δ_C 77.0 ppm.

3.2.2. Analysis of the PKS/NRPS biosynthetic gene cluster in the genome of *H. ochraceum*

The genome data of *H. ochraceum* SMP-2^{T} (Accession No. NC_013440.1[19]) was analyzed by antiSMASH[20, 21] (http://antismash.secondarymetabolites.org/). The sequences for the putative *hla* and *auf* biosynthetic gene clusters were annotated by BLAST and CDD[22]. The multiple alignments of amino acid sequences were generated by Clustal Omega program provided by EMBL.

3.2.3. Isolation of haliamide (2)

The marine myxobacterium *Haliangium ochraceum* SMP-2 was cultivated in 2 L flasks containing 500 mL of the production medium at 30 °C and 180 rpm for two weeks as previously reported.[11] The cells and resin were harvested from 15 L culture broth by filtration and extracted with methanol at 30 °C for 60 min on a horizontal shaker (120 rpm). The combined methanolic extracts were concentrated *in vacuo*. The resulting crude extract was further partitioned between water and EtOAc, and the organic layer was concentrated to yield an oily sample. The crude oil was chromatographed on a Hi-Flash silica gel column (size L, 30 g, Yamazen Co., Osaka, Japan) with a linear gradient of 10%-50% EtOAc in hexane (40 min) at 12 mL/min. The fractions (350 mg) eluted with 18%~25% EtOAc in hexane were combined and subjected to preparative

HPLC [Develosil ODS-HG-5 (ϕ 20 × 250 mm, Nomura Chemical Ltd.), 60% MeCN, 8 mL/min, monitored at 220 nm] to obtain haliamide (**2**, 5.6 mg, t_R =45-50 min). The NMR data was identical to that reported.[18]

3.2.4. Bioassay

3.2.4.1. Cytotoxicity assay

HeLa-S3 (SC) cells were provided by the RIKEN BRC through the National Bio-Resource Project of the MEXT, Japan. The cells were cultured in Eagle's minimal essential medium (EMEM) (Wako Pure Chemical Industries, Osaka, Japan) supplemented with 10% bovine serum, 100 units/mL penicillin, and 100 μg/mL streptomycin (Thermo Fisher Scientific Inc., MA, USA). A total of 10,000 cultured cells were seeded into each well of a 96-well plate containing 99 μL of the same medium. After pre-incubation for 24 h at 37 °C in an atmosphere of 5% CO_2, a compound (haliamide or paclitaxel as a positive control) in 1 μL of dimethyl sulfoxide (DMSO), or only 1 μL of DMSO as control, was added to each well, and the cells were incubated an additional 48 h. A solution (10 μL) of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) in phosphate buffer saline (PBS) (5 mg/mL) was then added to each well, and the plate was incubated for an additional 3 h. Subsequently, the medium was removed by aspiration, any generated formazan was dissolved in 100 μL of DMSO, and the absorbance was measured at 595 nm using a Multiskan FC microplate reader (Thermo Fisher Scientific). Values are means ± standard error (n = 4).

3.2.4.2. Anti-oomycete activity

The phytopathogenic oomycete *Phytophthora capsici* NBRC 30696, purchased from NITE-Biological Resource Center (NBRC, Chiba, Japan), was cultured on a potato-agar medium [in 1 L, potato broth from 200 g of fresh potato, glucose (20 g), and agar (20 g)] in a 9-cm dish at 25 °C for 7 days in the dark. A piece of the colony was then inoculated on the center of a 5% V8-agar medium [V8 vegetable juice (5 mL), agar (1.5 g), and water (95 mL)] in a 9-cm dish and incubated at 25 °C for 48 h in the dark until the colony grew to approximately 4 cm in diameter. A paper disc (6 mm in diameter) impregnated with a sample was placed 1 cm away from the front of the colony. After incubating for 22~24 h, the distance between the edge

of the colony and the paper disc (control: 0 mm) was measured. The activity was defined as a minimum dose that induced a definite distance (0.5 mm or wider) between the colony and the disk. The tested doses were 1, 3, 10, and 30 μg/disk.

3.2.4.3. MIC assay

The microorganisms, *Candida rugosa* AJ 14513, *Bacillus subtilis* AJ 12865, and *Escherichia coli* AJ 3837, were provided by the Institute for Innovation, Ajinomoto Co., Inc. (Kanagawa, Japan). The protocols used for the assay are in the followings. For *B. subtilis* and *E. coli*, see *Methods for Dilution Antimicrobial Susceptibility Tests for Bacteria That Grow Aerobically; Approved Standard*—Ninth Edition. CLSI document M07-A9 [ISBN 1-56238-783-9], Clinical and Laboratory Standards Institute (Wayne, PA, USA), 2012. For *C. rugosa*, see *Reference Method for Broth Dilution Antifungal Susceptibility Testing of Yeasts; Approved Standard*—Second Edition. NCCLS document M27-A2 [ISBN 1-56238-469-4], NCCLS (Wayne, PA, USA), 2002. The tested concentrations were 1, 2, 4, 8, 16, 32, 64, and 128 μg/mL.

3.3. Results and discussion

3.3.1. Discovery of the PKS/NPRS biosynthetic loci in the genome of *H. ochraceum*

The genome data of *H. ochraceum* SMP-2^{T} (Accession No. NC_013440.1[19]) was subjected to antiSMASH analysis[20, 21] with default parameter. A total of 25 secondary metabolite clusters were detected, including two PKSs, three NRPSs, three PKS-NRPS hybrids, three for terpenes, five for bacteriocin, one for ectoine, one for lassopeptide, three for lantipeptide and four for unknown type of metabolites. In this work, I selected 7 of the 8 PKS/NRPS clusters (one NRPS is thought to be too small to biosynthesize any secondary metabolites and therefore be excluded) as the research subjects. The locations of these seven clusters in the genome of *H. ochraceum* are showed in Fig. 3-2 and their features are summarized in Table 3-1.

Among detected PKS/NRPS type clusters, one PKS cluster (location at 1.50 - 1.55 Mbp, 47.8 kp, Fig. 3-2 and Table 3-1) is identical to our sequenced haliangicin biosynthetic gene cluster (*hli*, Chapter 2). The remaining six clusters listed in Table 3-1 show GC contents of around 65-75%, which are typical in the myxobacteria.

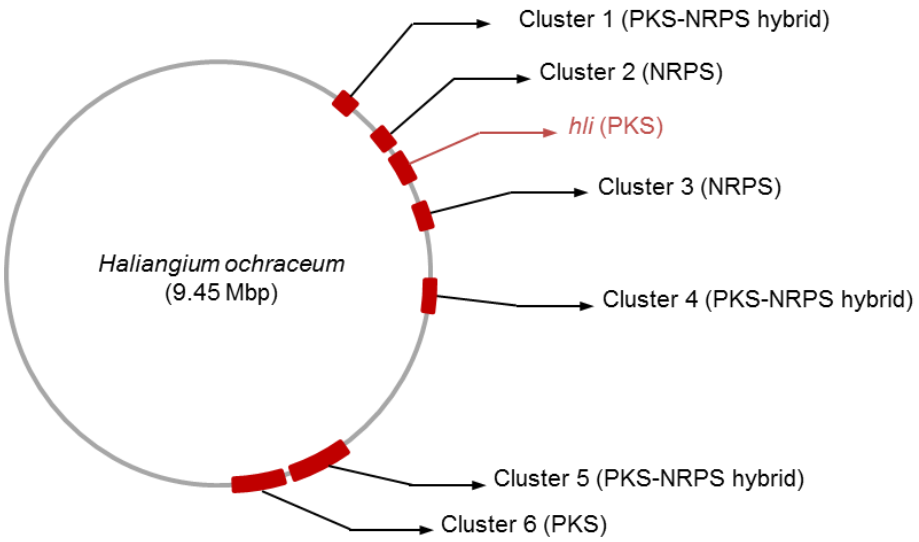


Figure 3-2. The PKS/NRPS biosynthetic loci in the genome of *H. ochraceum*.

Table 3-1. Summary of the PKS/NRPS biosynthetic gene clusters in *H. ochraceum.*

Cluster	Type	Location (Mbp)	Length (kbp)	GC content (%)
1	PKS-NRPS hybrid	1.03 – 1.05	24.7	73
2	NRPS	1.38 – 1.40	24.5	71
3	NRPS	1.81 – 1.84	30.1	72
4	PKS-NRPS hybrid	2.38 – 2.42	39.9	71
5	PKS-NRPS hybrid	4.01 – 4.07	60.1	67
6	PKS	4.08 – 4.13	52.6	77
hli	PKS	1.50 - 1.55	47.8	67

3.3.2. Analysis the PKS/NPRS biosynthetic gene clusters in the genome of *H. ochraceum*

3.3.2.1. Cluster 1

Cluster 1 is a PKS-NRPS hybrid which spans 25 kbp with high GC content of 73%. It harbors one PKS-NRPS hybrid-encoding gene (*hoch_0798*), one PKS-encoding gene (*hoch_0799*) and a hypothetical gene (*hoch_0800*) (Fig. 3-3 and Table 3-2). The product produced by this cluster is predicted to be a PKS-NRPS hybrid compound which consists of one amino acid unit and four acetate/propionate units. This cluster was characterized as the haliamide biosynthetic gene cluster (*hla*) which will be discussed later in section **3.3.3** and **3.3.4**.

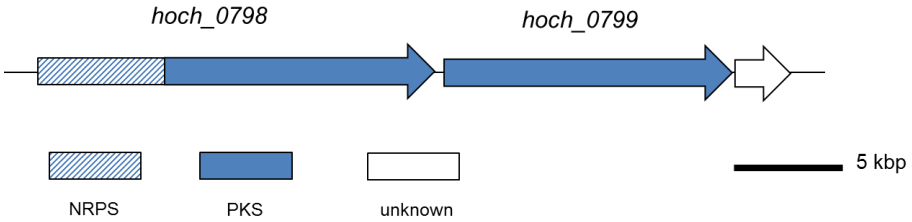


Figure 3-3. Genetic organization of Cluster 1.

Table 3-2. Summary of Cluster 1.

Cluster	Type	ORF	Module	PKS	NRPS	PKS-NRPS	Modification enzyme
1	PKS-NRPS	3	5	1	0	1	0

3.3.2.2. Cluster 2

Cluster 2 is a NRPS cluster which also spans 25 kbp with high GC content of 71%. One large NRPS gene (*hoch_1077*) encoding 4 modules were found, accompanied by eight modification enzyme-encoding genes downstream to this NRPS gene (Fig. 3-4 and Table 3-3). The modification enzymes include acyl transferase, aminotransferase, pyruvate carboxyltransferase and methyltransferase. The product of this gene cluster is predicted to be a non-ribosomal peptide consisting of five amino units.

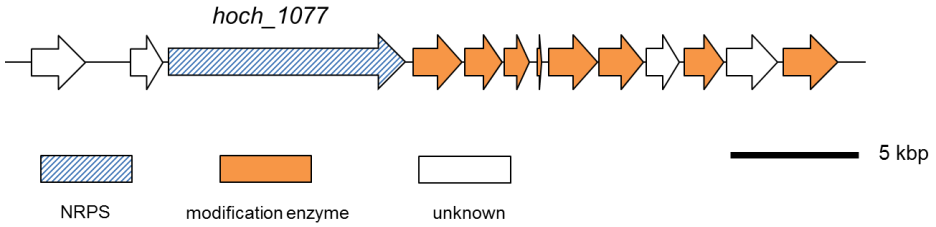


Figure 3-4. Genetic organization of Cluster 2.

Table 3-3. Summary of Cluster 2.

Cluster	Type	ORF	Module	PKS	NRPS	PKS-NRPS	Modification enzyme
2	NRPS	13	4	0	1	0	8

3.3.2.3. Cluster 3

Cluster 3 consists of one NRPS gene (*hoch_1367*) and several modification enzyme-encoding genes (Fig. 3-5 and Table 3-4). *Hoch_1367* encodes six NRPS modules, and therefore is proposed to be responsible for a non-ribosomal peptide with seven amino units. Surrounding modification enzymes include some oxidoreductases and a methyltransferase.

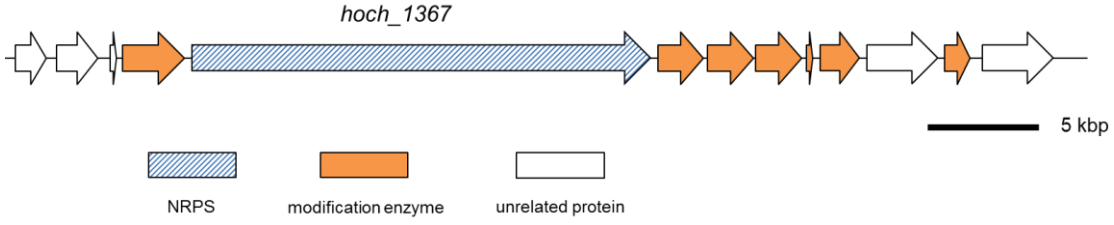


Figure 3-5. Genetic organization of Cluster 3.

Table 3-4. Summary of Cluster 3.

Cluster	Type	ORF	Module	PKS	NRPS	PKS-NRPS	Modification enzyme
3	NRPS	13	6	0	1	0	7

3.3.2.4. Cluster 4

Cluster 4 consists of two NRPS gene (*hoch_1748*, *hoch_1749*) and one PKS-NRPS hybrid encoding gene (*hoch_1750*) (Fig. 3-6 and Table 3-5). There are nine modules in this cluster, which are thought to be able to biosynthesize a PKS-NRPS hybrid compound that possesses a PKS-derived starter and a peptide moiety. The downstream modification gene (*hoch_1746*) encodes an AMP-dependent ligase which may participate in the starter biosynthesis.

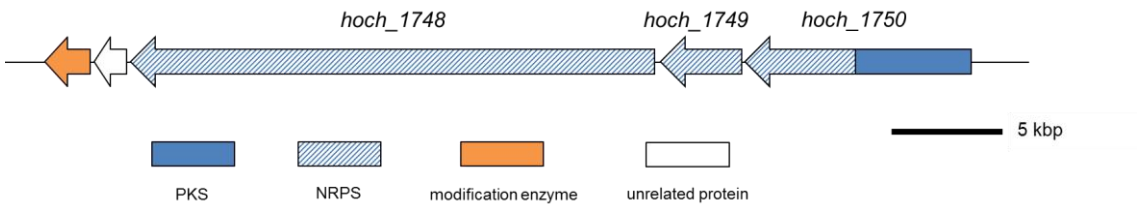


Figure 3-6. Genetic organization of Cluster 4.

Table 3-5. Summary of Cluster 4.

Cluster	Type	ORF	Module	PKS	NRPS	PKS-NRPS	Modification enzyme
4	PKS-NRPS	5	9	0	2	1	1

3.3.2.5. Cluster 5

The PKS-NRPS type cluster 5 consists of four PKS genes (*hoch_2953*, *hoch_2954*, *hoch_2955*, *hoch_2956*), one NRPS gene (*hoch_2952*) and one PKS-NRPS hybrid encoding gene (*hoch_2957*), spanning a region of 60 kbp (Fig. 3-7 and Table 3-6). Several modification enzyme-encoding genes are found, among which one methoxymalonyl-ACP cassette shows a high similarity to the methoxymalonyl-ACP cassette in the *hli* cluster (Chapter 2, section **2.3.10**). Due to the presence of this

gene cassette, a dioxyl structure in the compound produced by cluster 5 can be predicted.

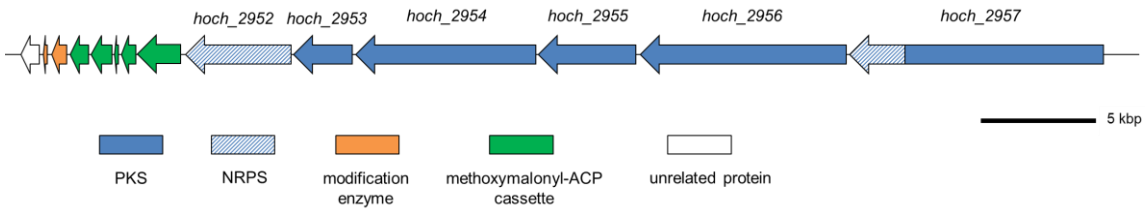


Figure 3-7. Genetic organization of Cluster 5.

Table 3-6. Summary of Cluster 5.

Cluster	Type	ORF	Module	PKS	NRPS	PKS-NRPS	Modification enzyme
5	PKS-NRPS	14	10	4	1	1	2 1 methoxymalonyl-ACP cassette

3.3.2.6. Cluster 6

Cluster 6 is a pure PKS system in which six PKS genes (*hoch_2968, hoch_2969, hoch_2970, hoch_2971, hoch_2972, hoch_2974*) encode 8 modules (Fig. 3-8 and Table 3-7). Three cytochrome P450 (*hoch_2966, hoch_2967, hoch_2973*) and one monooxygenase (*hoch_2975*) are proposed to be involved in modification reactions. Sequence analysis of this cluster suggested it to be homologous to the aurafuron biosynthetic gene cluster (*auf*), which will be describe in section **3.3.7**.

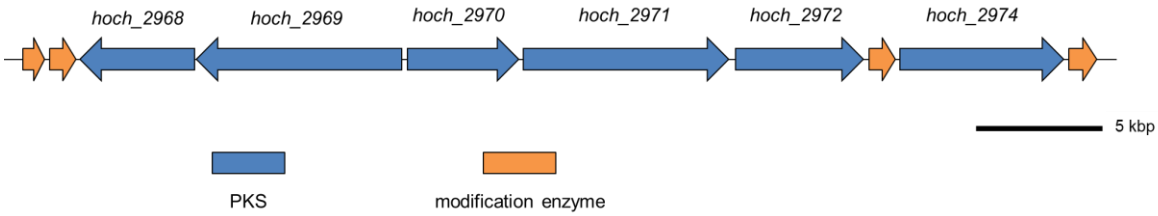


Figure 3-8. Genetic organization of Cluster 6.

Table 3-7. Summary of Cluster 6.

Cluster	Type	ORF	Module	PKS	NRPS	PKS-NRPS	Modification enzyme
6	PKS	10	8	6	0	0	4

3.3.3. Analysis of the haliamide biosynthetic gene cluster (*hla*)

The chemical structure of haliamide (**2**) and its biosynthetic precursors deduced from the feeding experiments (Fig. 3-1) suggest it to be a hybrid compound biosynthesized by PKS and NRPS. The *in silico* analysis of the genome data of *H. ochraceum* using antiSMASH[20, 21] revealed a candidate gene cluster (Cluster 1), which possesses one NRPS module followed by four PKS modules and is in a good agreement with the assembly of the biosynthetic units in **2**. The putative haliamide gene cluster (*hla*) consists of one NRPS/PKS hybrid gene (*hlaA*, GenBank accession No. KU516823) and one PKS gene (*hlaB*, GenBank accession No. KU516824), spanning a region of 21.7 kbp (Fig. 3-9).

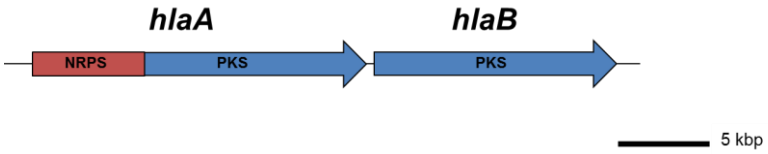


Figure 3-9. Genetic organization of the haliamide biosynthetic gene cluster (*hla*, 21.7 kbp).

Domain annotation of *hlaA* and *hlaB* by antiSMASH (Fig. 3-10 and Table 3-8) revealed 1 NRPS and 2 PKS modules in *hlaA* and 2 PKS modules in *hlaB*. While considering the structure of haliamide (**2**), several domains of PKSs are thought to be missing. These domains are AT domains in module 2 and 5, DH in module 4 and KR domain in module 5 (Fig. 3-10). For the missing AT domains, a standalone acyltransferase (Hoch_5652) was found in the genome of *H. ochraceum*, which harbors the APFH motif that suggests this enzyme to be malonate-specific. Such substrate specificity well fits for the malonate incorporation by modules 2 and 5, as demonstrated by the feeding experiments (Figure 3-1b). On the hand, no candidates of trans-acting DH and KR domains were found. I supposed that some oxidoreductases downstream to the *hla* cluster, e. g., cytochrome P450 (Hoch_0804, Hoch_0812, Hoch_0813 and Hoch_0814) and amine oxidase (Hoch_0806), may complement these missing domains, though the real roles of these proteins remain unclear.

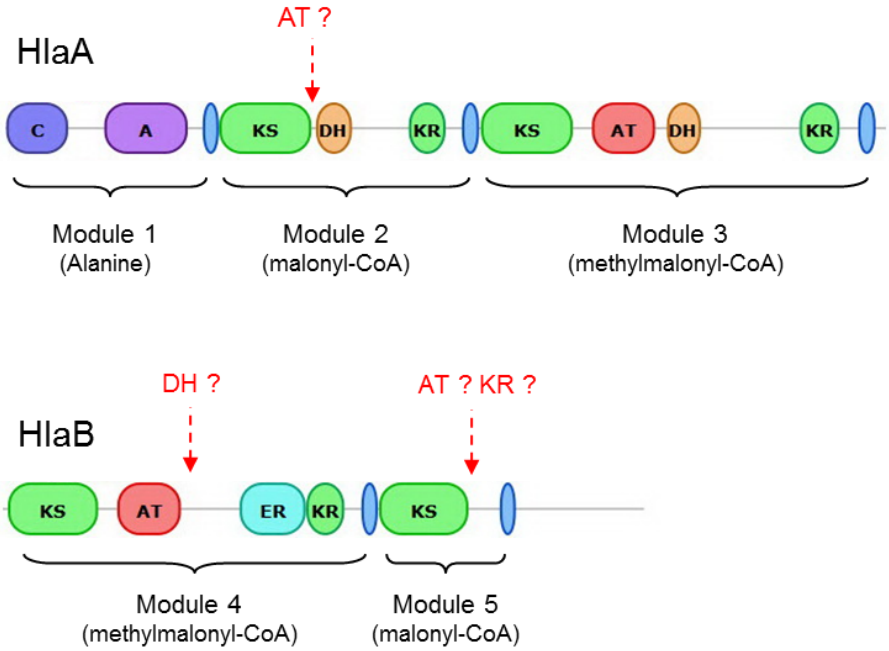


Figure 3-10. Detailed annotation of PKS and NRPS by antiSMASH 3.0 analysis.

Table 3-8. PKS/NRPS in *hla* for the biosynthesis of the haliamide skeleton.

Gene	Protein	Length(aa)	Type	Domain position
hlaA	HlaA	4183	NRPS-type I PKS	C (11-293), A (477-861), PCP (944-1009), KS_1 (1023-1448), DH_1 (1482-1641), KR_1 (1923-2086), ACP_1 (2175-2245), KS_2 (2266-2690), AT (2793-3081), DH_2 (3148-3299), KR_2 (3779-3956), ACP_2 (4057-4128)
hlaB	HlaB	3045	Type I PKS	KS_1 (31-447), AT (550-840), ER (1131-1433), KR (1447-1616), ACP_1 (1709-1775), KS_2 (1793-2205), ACP_2 (2366-2430)

Another unusual feature in the PKS modules is that the AT domain of module 4 indicates a malonate (acetate unit)-specific conserved motif HAFH[23] despite the fact that this module incorporated one methylmalonate (propionate) unit as shown in the feeding experiments (Fig. 3-1b). Such inconsistencies are not common in the typical bacterial PKS biosynthesis but sometimes observed in myxobacteria. To get a deeper understanding of this unusual AT domain, a phylogenetic tree of AT domains of type I PKS was generated by using several sequences from actinobacteria and proteobacteria (Fig. 3-11). The AT domain of module 4 was found to be phylogenetically close to the AT domain of EpoC (Fig. 3-11), a promiscuous AT

loading both malonate and methylmalonate in the epothilone biosynthetic pathway from terrestrial myxobacterium *Sorangium cellulosum*.[24]

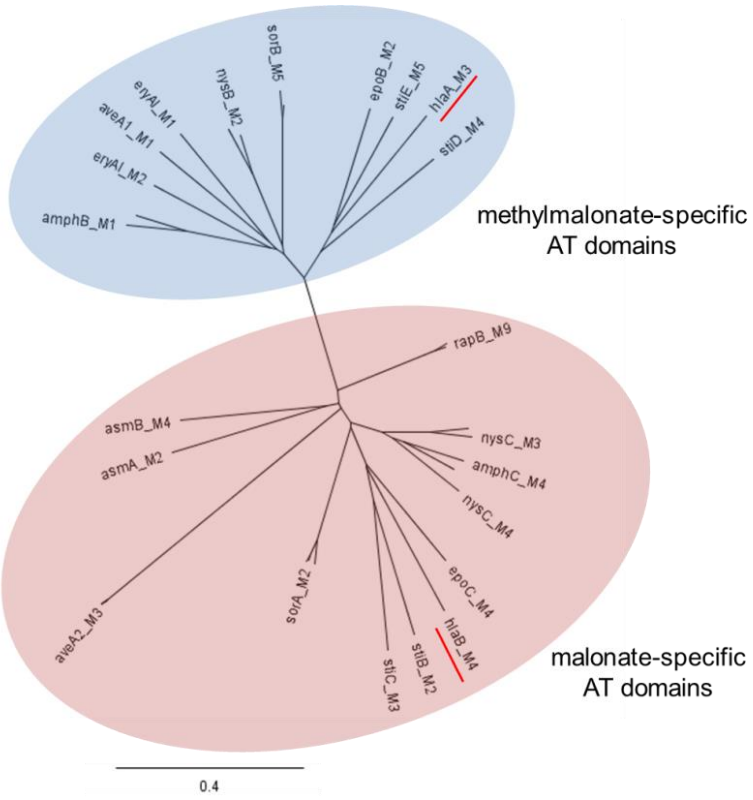


Figure 3-11. Phylogenetic tree of AT domains of type I PKS from several bacterial metabolites. Thirty three sequences from actinobacteria and proteobacteria were used, including AT domains in the amphotericin B (*amph*), ansamitocin (*asm*), avermectin (*ave*), epothilone (*epo*), erythromycin (*ery*), haliamide (*hla*), nystatin (*nys*), rapamycin (*rap*), soraphen (*sor*) and stigmatellin (*sti*) biosynthetic gene cluster. The AT domains from *hla* cluster are indicated by red underline. The AT domain of module 3 in *hlaA* is grouped in the methylmalonate specific clade, while AT domain of module 4 in *hlaB* is grouped in the malonate specific clade.

3.3.4. Proposed biosynthetic mechanism of haliamide (2)

The biosynthesis of haliamide (**2**) was proposed to begin with benzoyl CoA. The benzoate unit as a biosynthetic precursor is not very common and has been reported for just few natural products such as soraphen A[25] and enterocin.[26] Although the loading mechanism of this starter remains unclear due to the absence of the loading module in the *hla* gene cluster, the direct condensation of the benzoyl CoA starter and the forthcoming alanine unit on module 1 could be possible as a similar starter loading mechanism was recently reported in the biosynthesis of macyranones.[27] The C domain in HlaA may catalyze the formation of the amide bond. The subsequent chain elongation reactions take place by four PKS modules (Module 2–5), which incorporate two malonyl-CoAs and two methylmalonyl-CoAs to construct the polyketide backbone of **2** (Fig. 3-12a).

After integrating the last biosynthetic block (acetate unit) by module 5, the biosynthesis undergoes a particular chain termination by decarboxylation, leading to the formation of a terminal olefin of the haliamide molecule, which may be mediated by the sulfotransferase (ST)-thioesterase (TE) domains encoded in the C-terminus of HlaB (Fig. 3-12b). These domains are homologous to CurM TE and CurM ST (identity of 37% and 41%, respectively) of curacin A,[28] in which β-keto group is reduced by KR domain and the resulting β-hydroxyl intermediate is sulfated by ST, followed by TE-catalyzed hydrolysis coupled with decarboxylative elimination to give a terminal olefin.[29] The haliamide biosynthesis was predicted to be terminated in a similar way, though a KR domain to generate β-hydroxyl group is absent in module 5 (Fig. 3-12b). This is the second report of ST-TE mediated decarboxylative termination leading to the terminal alkene formation in the biosynthesis of natural products. It is interesting that another metabolite from *H. ochraceum*, haliangicin (**1**), also possesses a terminal alkene. However, the formation of the terminal alkene in haliangicin (**1**) has been proved to involve an acyl-CoA dehydrogenase in the early stage of biosynthesis (Chapter 2, section **2.3.7**), which is a totally different way from the ST-TE mediated alkene formation.

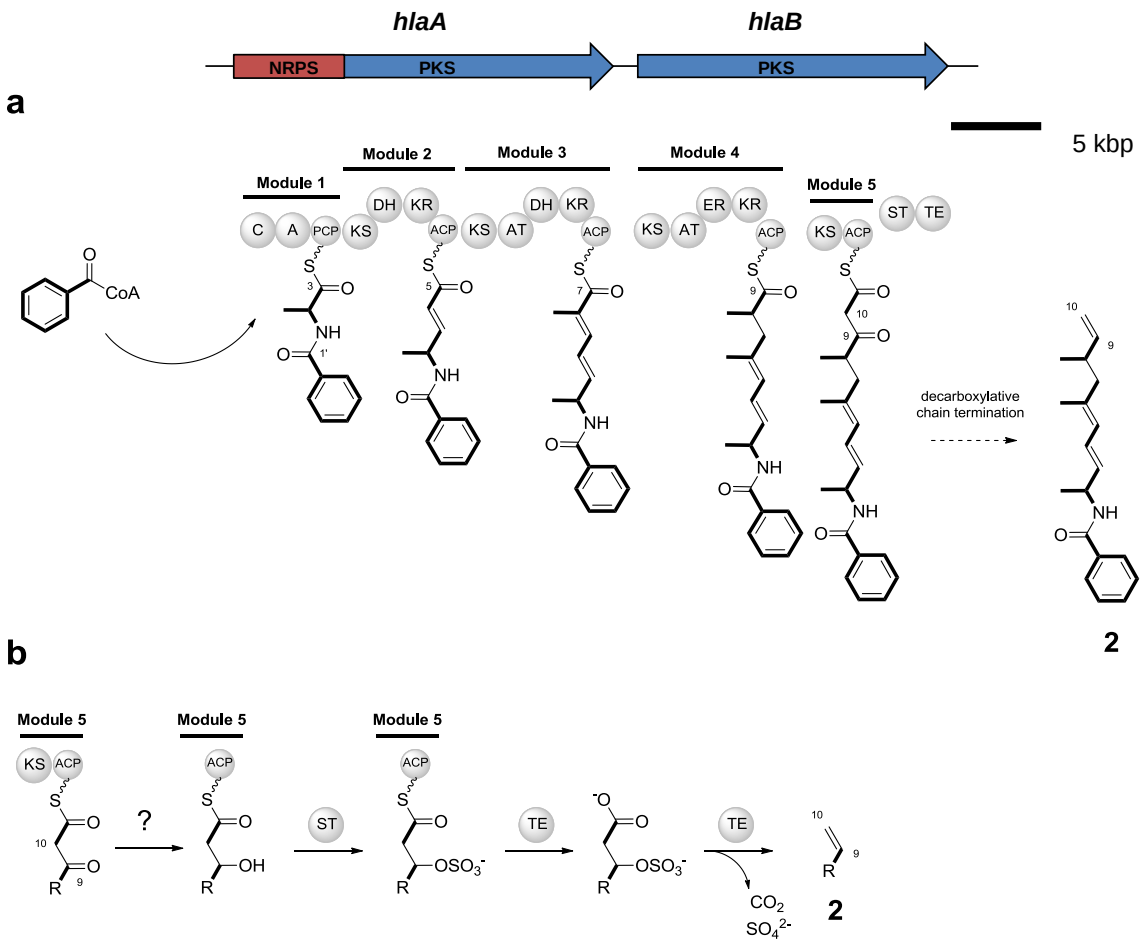


Figure 3-12. Proposed biosynthetic machinery of haliamide (**2**).

(**a**) Genetic organization of the haliamide biosynthetic gene cluster (*hla*, 21.7 kbp) and a biosynthetic pathway for **2**.

(**b**) Decarboxylative chain termination at the final step leading to the formation of the terminal olefin at C-9 and C-10. *Abbreviations:* A, adenylation domain; ACP, acyl carrier protein; AT, acyl transferase; C, condensation domain; DH, dehydratase; ER, enoyl reductase; KR, ketoreductase; KS, ketosynthase; PCP, peptide carrier protein; ST, sulfotransferase; TE, thioesterase.

3.3.5. Isolation of haliamide (2)

The marine myxobacterium *H. ochraceum* SMP-2 showed an optimal NaCl demand of 2% (w/v) for its growth and needs approximately two weeks for the maximal production of secondary metabolites. The bacterial cells from a total of 15 L cultures were extracted with methanol. After solvent partitioning of the

methanol extract, an EtOAc fraction was chromatographed on silica gel followed by reverse-phase HPLC to give 5.6 mg of haliamide (**2**) as well as haliangicin.

3.3.6. Bioactivities of haliamide (2)

The cytotoxic effect of the isolated compound haliamide (**2**) was evaluated by MTT assay using a tumor cell line (HeLa-S_3). The IC_{50} was determined to be 12 μM (a positive control, paclitaxel: IC_{50} = 8.9 nM, Fig. 3-13). In the antimicrobial tests, haliamide (**2**) did not show inhibitory activities against fungi (*Phytophthora capsici*: minimum inhibition dose > 30 mg/disk; *Candida rugosa*: MIC > 128 mg/mL), gram negative (*Escherichia coli*: MIC > 128 mg/mL) and gram positive bacteria (*Bacillus subtilis*: MIC > 128 mg/mL).

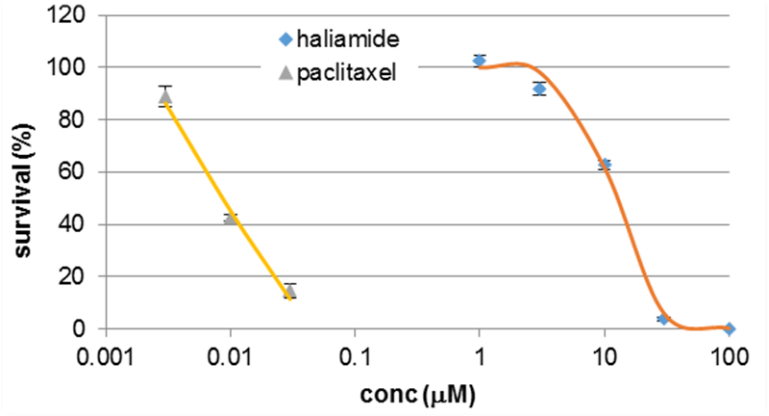


Figure 3-13. Cytotoxicity of haliamide (**2**) against HeLa-S3 cells. Markers indicate average values with SE (n = 4), and the lines indicate theoretical sigmoid curves [$y = 100/(1+e^{-a(x-b)})$, x: log(μM)]. Paclitaxel is a positive control (IC_{50} = 8.9 nM).

3.3.7. Putative aurafuron biosynthetic gene cluster (*auf*)

Gene annotation of the cluster 6 by antiSMASH, BLAST and CDD suggested it to be a pure PKS cluster. Considering the highly similar genetic architecture to the reported biosynthetic gene cluster (*auf*) for aurafuron A (**3**) from *Stigmatella aurantiaca* [30, 31](Fig. 3-14), I supposed this cluster to be a putative *auf*

locus. The genes in the cluster 6 and their identities to the *auf* counterpart are summarized in Table 3-9. So far, there has been no report of aurafuron derivatives from the marine myxobacterium *H. ochraceum*.

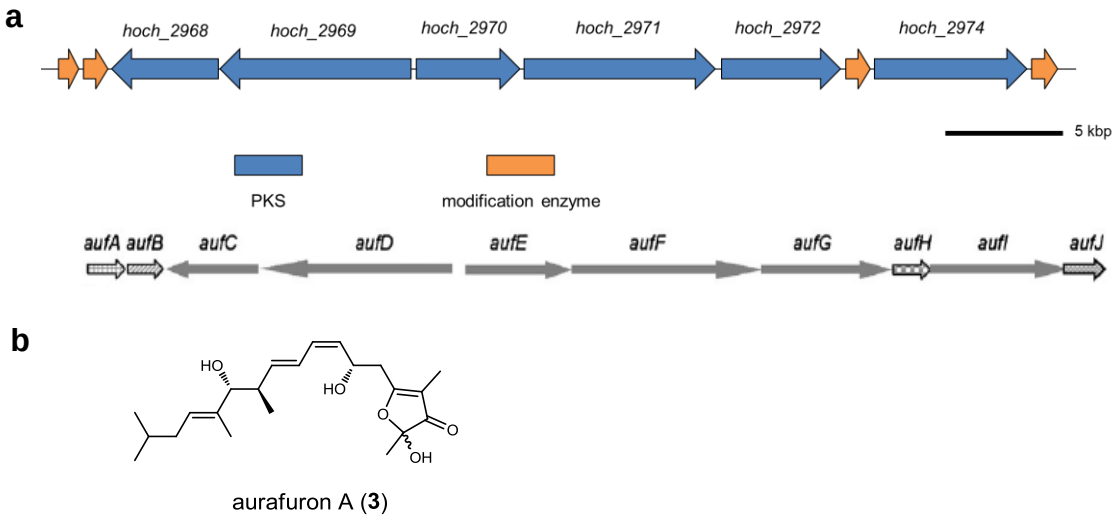


Figure 3-14. Putative aurafuron biosynthetic gene cluster (*auf*).

(**a**) Comparison of the genetic architecture of cluster 6 and the reported aurafuron biosynthetic gene cluster (*auf*) from *S. aurantiaca*.

(**b**) Structure of aurafuron A (**3**).

Table 3-9. Comparison of cluster 6 and the aurafuron biosynthetic gene cluster (*auf*).

Protein	Accession	Length (aa)	Proposed functions	Auf homology (identities %)
Hoch_2966	ACY15478.1	475	cytochrome P450	AufA (74%)
Hoch_2967	ACY15479.1	434	cytochrome P450	AufB (74%)
Hoch_2968	ACY15480.1	1909	6-deoxyerythronolide-B synthase	AufC (66%)
Hoch_2969	ACY15481.1	3427	acyl transferase	AufD (67%)
Hoch_2970	ACY15482.1	1841	β-ketoacyl synthase	AufE (61%)
Hoch_2971	ACY15483.1	3424	3-xxoacyl-(acyl-carrier-protein) reductase, 6-deoxyerythronolide-B synthase	AufF (68%)
Hoch_2972	ACY15484.1	2136	6-deoxyerythronolide-B synthase	AufG (58%)
Hoch_2973	ACY15485.1	467	cytochrome P450	AufH (73%)
Hoch_2974	ACY15486.1	2719	6-deoxyerythronolide-B synthase, (acyl-carrier-protein) *S*-malonyltransferase	AufI (62%)
Hoch_2975	ACY15487.1	506	monooxygenase FAD-binding protein	AufJ (62%)

3.4. Summary

In the present study, the genome mining of the marine myxobacterium *H. ochraceum* revealed six PKS/NRPS gene clusters, as well as the haliangicin biosynthetic gene cluster (*hli*). Cluster 1 was characterized as the biosynthetic gene cluster responsible for haliamide (**2**), a new PKS/NRPS hybrid type metabolite discovered from the halophilic producer *H. ochraceum*. The biosynthetic pathway of **2** was proposed, which suggested a unique terminal alkene formation via decarboxylation. Genomic analysis also provided a putative biosynthetic gene cluster highly similar to the aurafuron biosynthetic gene cluster (*auf*). The detected PKS/NRPS gene clusters prompt us to discover additional unknown metabolites of this hard-to-culture marine myxobacterium in the future.

Reference

1. Reichenbach H. Myxobacteria, producers of novel bioactive substances. *J. Ind. Microbiol. Biotechnol.* **27**, 149-156 (2001).

2. Weissman KJ, Muller R. A brief tour of myxobacterial secondary metabolism. *Bioorg Med Chem* **17**, 2121-2136 (2009).

3. Wenzel SC, Muller R. Myxobacteria-'microbial factories' for the production of bioactive secondary metabolites. *Mol. Biosyst.* **5**, 567-574 (2009).

4. Iizuka T, Jojima Y, Fudou R, Hiraishi A, Ahn JW, Yamanaka S. *Plesiocystis pacifica* gen. nov., sp nov., a marine myxobacterium that contains dihydrogenated menaquinone, isolated from the Pacific coasts of Japan. *Int. J. Syst. Evol. Microbiol.* **53**, 189-195 (2003).

5. Iizuka T, Jojima Y, Fudou R, Tokura M, Hiraishi A, Yamanaka S. *Enhygromyxa salina* gen. nov., sp nov., a slightly halophilic myxobacterium isolated from the coastal areas of Japan. *Syst. Appl. Microbiol.* **26**, 189-196 (2003).

6. Iizuka T, Jojima Y, Hayakawa A, Fujii T, Yamanaka S, Fudou R. *Pseudenhygromyxa salsuginis* gen. nov., sp nov., a myxobacterium isolated from an estuarine marsh. *Int. J. Syst. Evol. Microbiol.* **63**, 1360-1369 (2013).

7. Schaberle TF, *et al.* Marine Myxobacteria as a Source of Antibiotics-Comparison of Physiology, Polyketide-Type Genes and Antibiotic Production of Three New Isolates of *Enhygromyxa salina*. *Mar. Drugs* **8**, 2466-2479 (2010).

8. Still PC, Johnson TA, Theodore CM, Loveridge ST, Crews P. Scrutinizing the Scaffolds of Marine Biosynthetics from Different Source Organisms: Gram-Negative Cultured Bacterial Products Enter Center Stage. *J. Nat. Prod.* **77**, 690-702 (2014).

9. Komaki H, *et al.* PCR detection of type I polyketide synthase genes in myxobacteria. *Appl. Environ. Microbiol.* **74**, 5571-5574 (2008).

10. Fudou R, Iizuka T, Sato S, Ando T, Shimba N, Yamanaka S. Haliangicin, a novel antifungal metabolite produced by a marine myxobacterium 2. Isolation and structural elucidation. *J. Antibiot.* **54**, 153-156

(2001).

11. Fudou R, Iizuka T, Yamanaka S. Haliangicin, a novel antifungal metabolite produced by a marine myxobacterium 1. Fermentation and biological characteristics. *J. Antibiot.* **54**, 149-152 (2001).
12. Kundim BA, *et al.* New haliangicin isomers, potent antifungal metabolites produced by a marine myxobacterium. *J. Antibiot.* **56**, 630-638 (2003).
13. Iizuka T, *et al.* Miuraenamides A and B, novel antimicrobial cyclic depsipeptides from a new slightly halophilic myxobacterium: Taxonomy, production, and biological properties. *J. Antibiot.* **59**, 385-391 (2006).
14. Ojika M, Inukai Y, Kito Y, Hirata M, Iizuka T, Fudou R. Miuraenamides: Antimicrobial cyclic depsipeptides isolated from a rare and slightly halophilic myxobacterium. *Chem. Asian. J.* **3**, 126-133 (2008).
15. Felder S, *et al.* Salimabromide: Unexpected Chemistry from the Obligate Marine Myxobacterium Enhygromxyasalina. *Chem. Eur. J.* **19**, 9319-9324 (2013).
16. Felder S, Kehraus S, Neu E, Bierbaum G, Schaberle TF, Konig GM. Salimyxins and Enhygrolides: Antibiotic, Sponge-Related Metabolites from the Obligate Marine Myxobacterium *Enhygromyxa salina*. *ChemBioChem* **14**, 1363-1371 (2013).
17. Fudou R, Jojima Y, Iizuka T, Yamanaka S. *Haliangium ochraceum* gen. nov., sp nov and Haliangium tepidum sp nov.: Novel moderately halophilic myxobacteria isolated from coastal saline environments. *J. Gen. Appl. Microbiol.* **48**, 109-115 (2002).
18. Sato J. Search for novel secondary metabolites from marine myxobacteria. Mster thesis. Nagoya University (2015).
19. Ivanova N, *et al.* Complete genome sequence of Haliangium ochraceum type strain (SMP-2^{T}). *Stand. Genomic. Sci.* **2**, 96-106 (2010).
20. Blin K, *et al.* antiSMASH 2.0-a versatile platform for genome mining of secondary metabolite producers. *Nucleic Acids Res.* **41**, W204-W212 (2013).
21. Weber T, *et al.* antiSMASH 3.0-a comprehensive resource for the genome mining of biosynthetic gene

clusters. *Nucleic Acids Res.* **43**, W237-W243 (2015).

22. Marchler-Bauer A, *et al.* CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res.* **39**, D225-D229 (2011).

23. Yadav G, Gokhale RS, Mohanty B. Computational approach for prediction of domain organization and substrate specificity of modular polyketide synthases. *J. Mol. Biol.* **328**, 335-363 (2003).

24. Petkovic H, *et al.* Substrate specificity of the acyl transferase domains of EpoC from the epothilone polyketide synthase. *Org. Biomol. Chem.* **6**, 500-506 (2008).

25. Ligon J, *et al.* Characterization of the biosynthetic gene cluster for the antifungal polyketide soraphen A from *Sorangium cellulosum* So ce26. *Gene* **285**, 257-267 (2002).

26. Kalaitzis JA, *et al.* Policing starter unit selection of the enterocin type II polyketide synthase by the type II thioesterase EncL. *Bioorg. Med. Chem.* **19**, 6633-6638 (2011).

27. Keller L, Plaza A, Dubiella C, Groll M, Kaiser M, Muller R. Macyranones: Structure, Biosynthesis, and Binding Mode of an Unprecedented Epoxyketone that Targets the 20S Proteasome. *J. Am. Chem. Soc.* **137**, 8121-8130 (2015).

28. Chang ZX, *et al.* Biosynthetic pathway and gene cluster analysis of curacin A, an antitubulin natural product from the tropical marine cyanobacterium *Lyngbya majuscula*. *J. Nat. Prod.* **67**, 1356-1367 (2004).

29. Gu LC, *et al.* Polyketide Decarboxylative Chain Termination Preceded by *O*-Sulfonation in Curacin A Biosynthesis. *J. Am. Chem. Soc.* **131**, 16033- (2009).

30. Frank B, Wenzel SC, Bode HB, Scharfe M, Blocker H, Muller R. From genetic diversity to metabolic unity: Studies on the biosynthesis of aurafurones and aurafuron-like structures in myxobacteria and streptomycetes. *J. Mol. Biol.* **374**, 24-38 (2007).

31. Kunze B, Reichenbach H, Muller R, Hofle G. Aurafuron A and B, new bioactive polyketides from *Stigmatella aurantiaca* and *Archangium gephyra* (Myxobacteria) - Fermentation, isolation, physico-chemical properties, structure and biological activity. *J. Antibiot.* **58**, 244-251 (2005).

Chapter 4 Conclusion

The work present in this thesis focused on the biosynthetic genes of marine myxobacteria and their secondary metabolites. Marine myxobacteria are rare and hard-to-culture microorganisms, but they genetically harbor high potential to produce novel antibiotics, which thus makes them a promising source of new leads for drug development. The firstly discovered marine myxobacterium *Haliangium ochraceum* produces a potent antifungal polyketide haliangicin (**1**) in a low productivity. In the previous study, its biosynthetic gene cluster was cloned and heterologously expressed in a fast growing terrestrial myxobacterium. The biosynthetic building blocks were partially identified, while the origin of C-3 and C-4 position remained unclear. Moreover, detailed biosynthetic machinery of haliangicin (**1**) was unknown, which became a major subject in the present work. In Chapter 2, the detailed biosynthetic machinery of haliangicin was deciphered through the analyses of its biosynthetic genes. A highly efficient production of haliangicin (**1**) and generation of some unnatural haliangicin analogues were also achieved in the work described in Chapter 2.In Chapter 3, mining the genome of *H. ochraceum* revealed several biosynthetic gene clusters for other secondary metabolites which are of interest. In the previous study, a thorough exam of the secondary metabolome of *H. ochraceum* led to the isolation of a new PKS-NRPS hybrid product, haliamide (**2**). Analyses of the detected gene clusters in the genome of *H. ochraceum* suggested a candidate that in a good agreement with the structure of **2**. A model for the haliamide biosynthesis was established. The remaining five PKS/NRPS biosynthetic gene clusters for other metabolites were also analyzed, from which putative biosynthetic gene cluster (*auf*) for aurafuron A (**3**) was discovered.

4.1. Biosynthetic machinery of haliangicin (1)

Based on the sequence analysis, the biosynthetic gene cluster of **1** (*hli*, 47.8 kbp) was found to consist of five polyketide synthases (PKSs, in blue) and various tailoring enzymes (in grey) (Fig. 4-1) and the biosynthetic mechanism was proposed (Fig. 4-2). The mechanism of the starter and the post-PKS modifications was elucidated through gene inactivation and *in vitro* functional analysis of tailoring

enzymes. Disruption of the methyltransferase (HliD), acyl-CoA dehydrogenase (HliR) and epoxidase (HliU) provided new "unnatural" products **7**, **8**, **9** and **10**. HliR was found to be a unique γ,δ-dehydrogenase that specific to short chain α,β-unsaturated acyl compounds (Fig. 4-3). In addition, 10-fold improvement of haliangicin productivity was achieved by using the heterologous expression system.

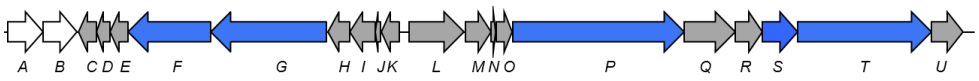


Figure 4-1. Genetic organization of the haliangicin biosynthetic gene cluster (*hli*).

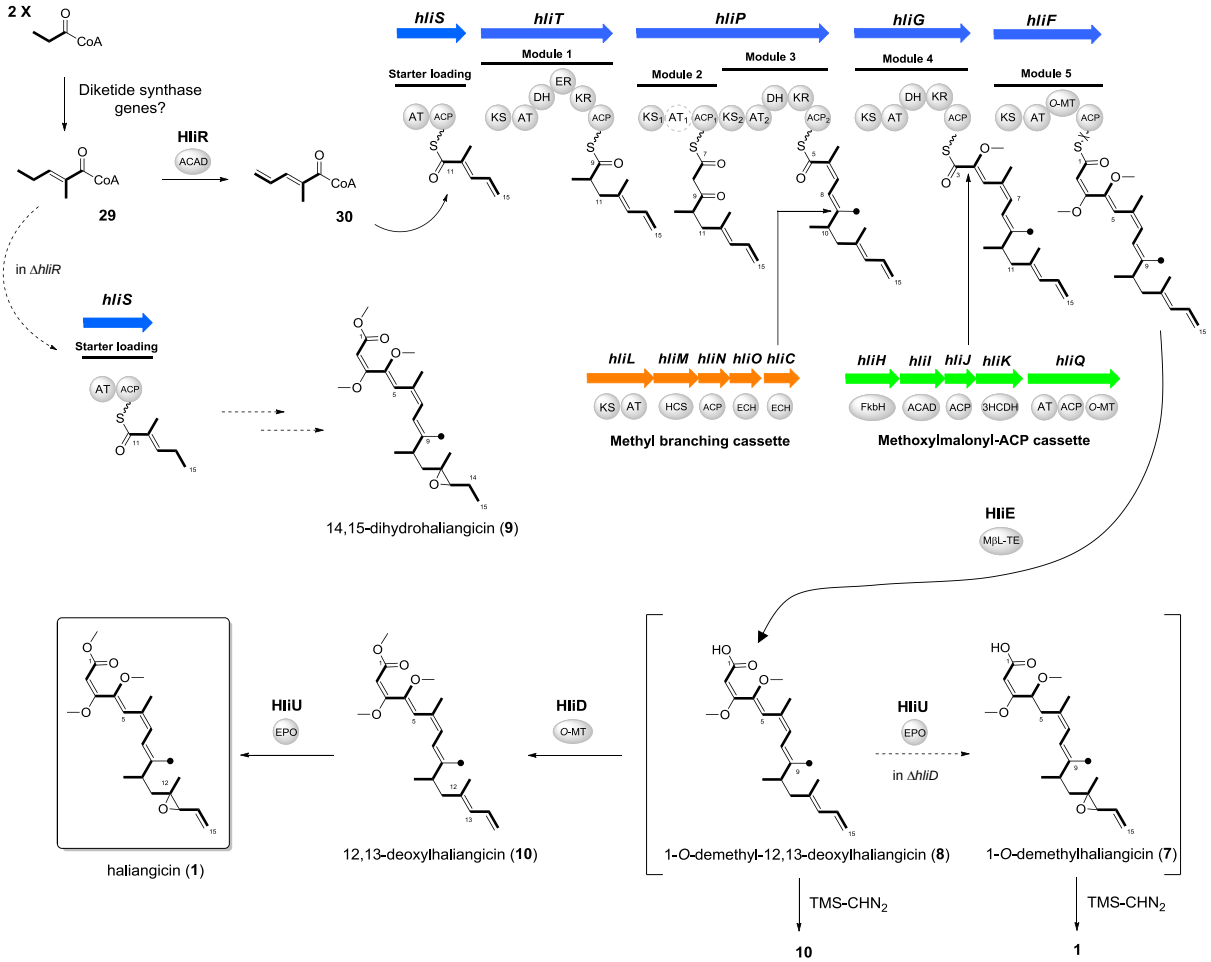


Figure 4-2. A proposed biosynthetic pathway for haliangicin.

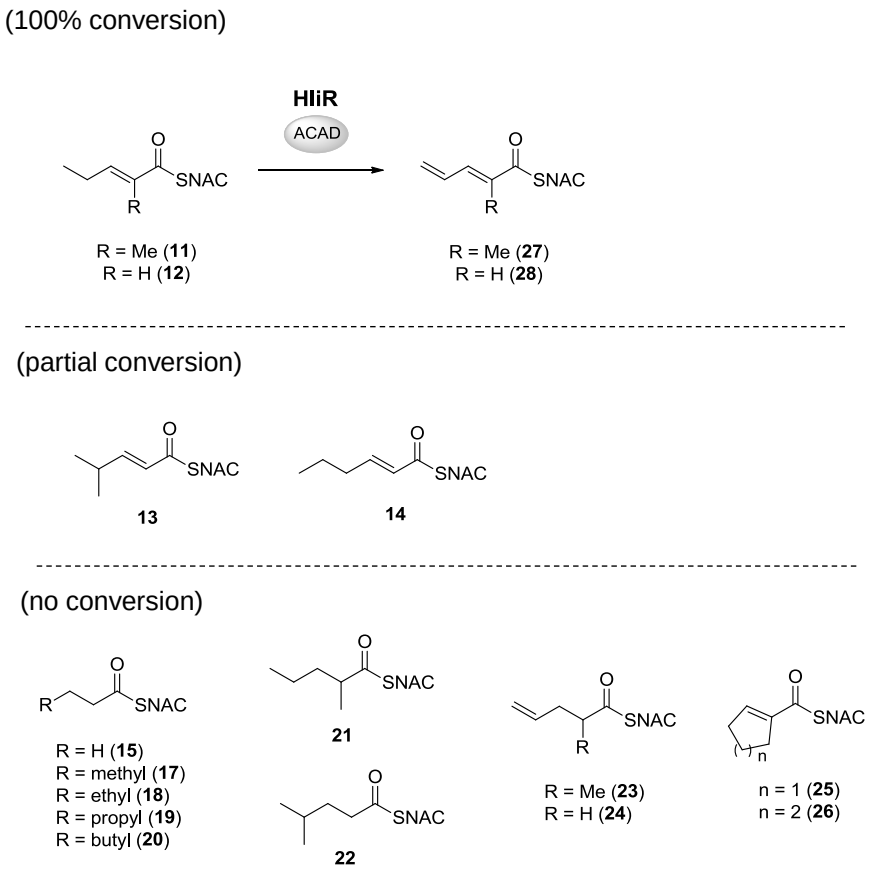


Figure 4-3. Substrate specificity of recombinant acyl-CoA dehydrogenase HliR.

4.2. Analyses of biosynthetic genes for haliamide (2) and other secondary metabolites

Haliamide (**2**) is another metabolite previously isolated from *H. ochraceum*. Based on the genome analysis by antiSMASH (Fig. 4-4), a PKS-NRPS hybrid gene cluster was identified as the genes responsible for the haliamide biosynthesis. Access to the biosynthetic gene cluster of haliamide (*hla*, 21.7 kbp) allowed an establishment of a model for the haliamide biosynthesis (Fig. 4-5a). The sulfotransferase (ST)-thioesterase (TE) domains encoded in *hlaB* appears to be responsible for the terminal alkene formation via decarboxylation (Fig. 4-5b).

In addition, I also found a PKS gene cluster similar to that for aurafuron A (**3**), a known polyketide from the terrestrial myxobacterium *Stigmatella aurantiaca*. Other four PKS/NRPS gene clusters for unknown secondary metabolites could be analyzed by a similar methodology in the future.

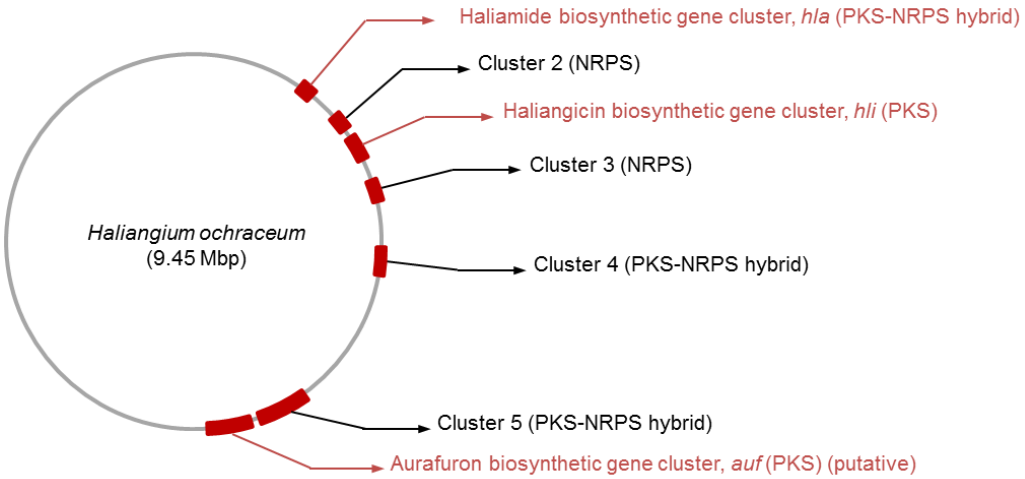


Figure 4-4. The PKS/NRPS loci in the genome of *H. ochraceum* detected by antiSMASH analysis.

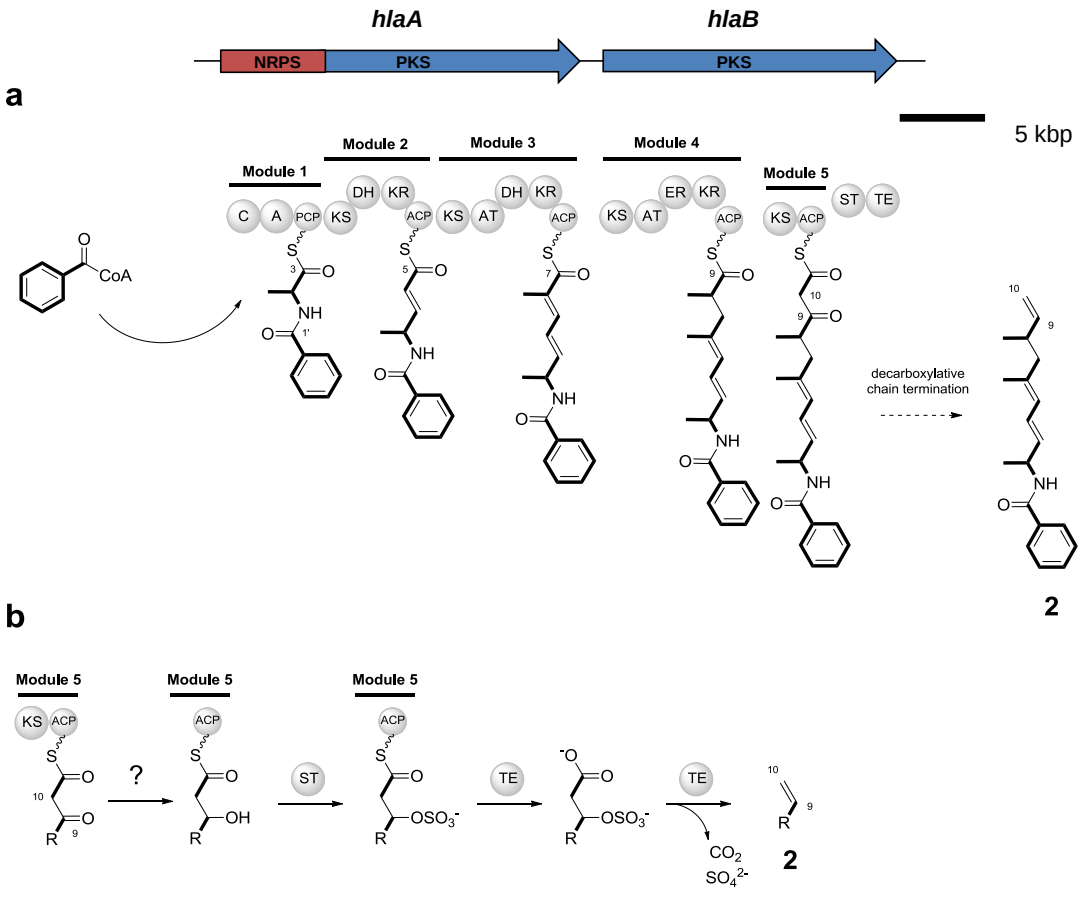


Figure 4-5. Proposed biosynthetic machinery of haliamide (**2**).

(**a**) Genetic organization of the haliamide biosynthetic gene cluster (*hla*) and a biosynthetic pathway for **2**.

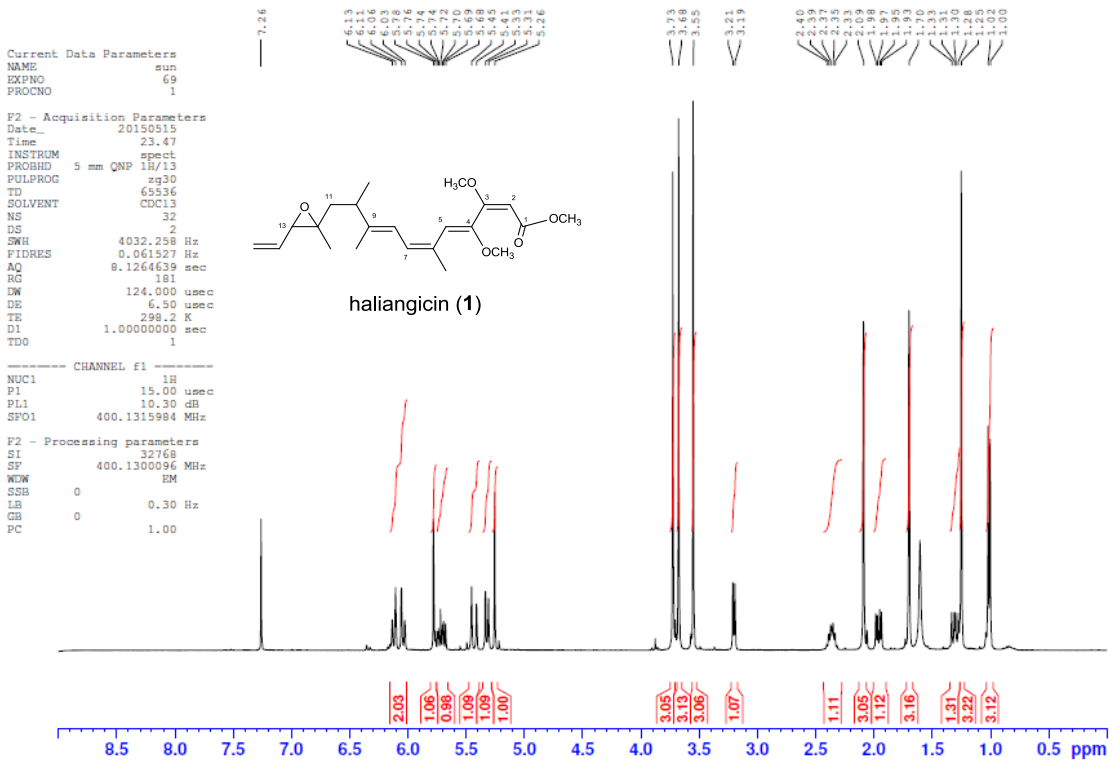
(**b**) Decarboxylative chain termination at the final step leading to the formation of the terminal olefin.

Chapter 5. Appendix

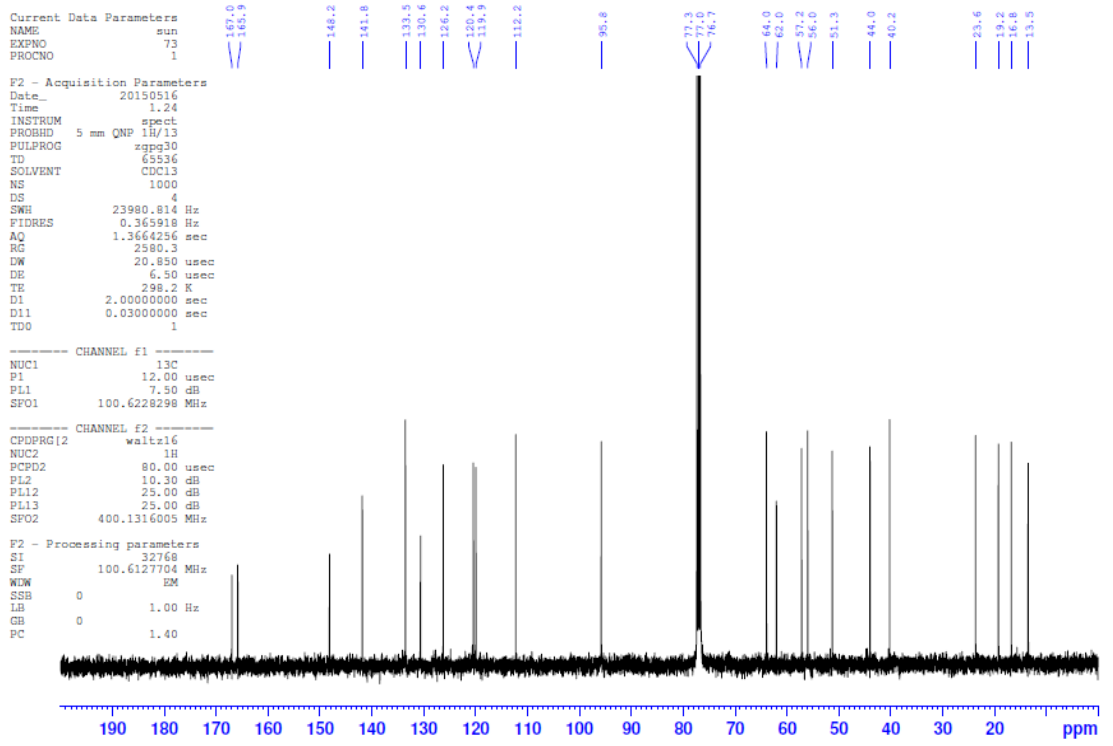
5.1. NMR spectra of haliangicin (1).

(**a**) ^{1}H NMR (400 MHz, $CDCl_3$); (**b**) ^{13}C NMR (100 MHz, $CDCl_3$).

a



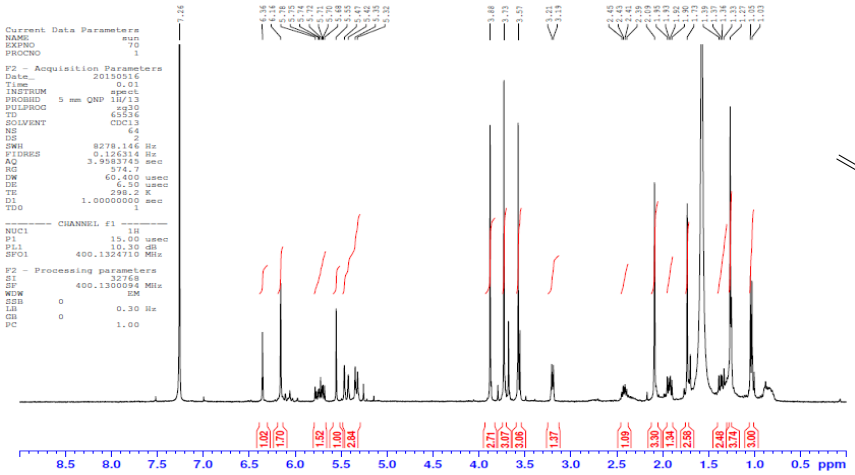
b

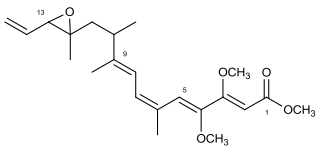


5.2. ^{1}H NMR spectra of haliangicin B (4), haliangicin C (5) and haliangicin D (6).

(**a**) ^{1}H NMR spectrum (400 MHz) of haliangicin B (**4**) in $CDCl_3$; (**b**) ^{1}H NMR spectrum (400 MHz) of haliangicin C (**5**) in $CDCl_3$; (**c**) ^{1}H NMR spectrum (400 MHz) of haliangicin C (**6**) in $CDCl_3$.

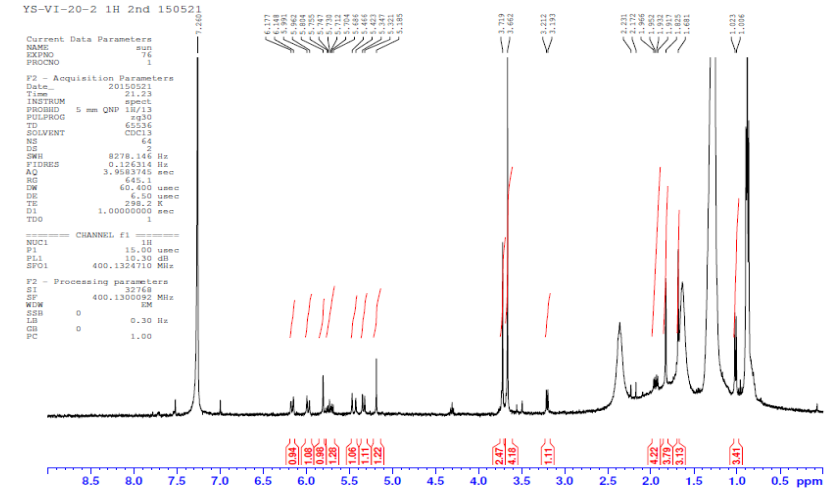
a

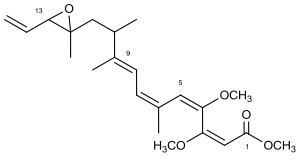




haliangicin B (**4**)

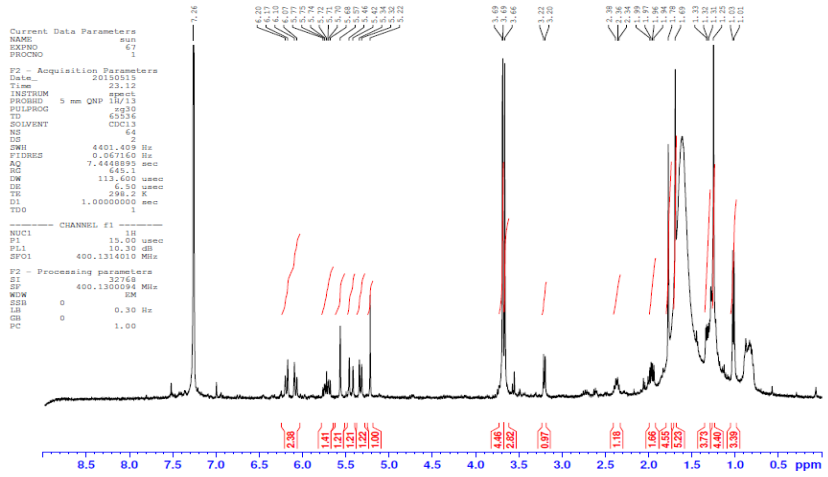
b

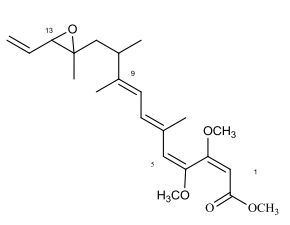




haliangicin C (**5**)

c



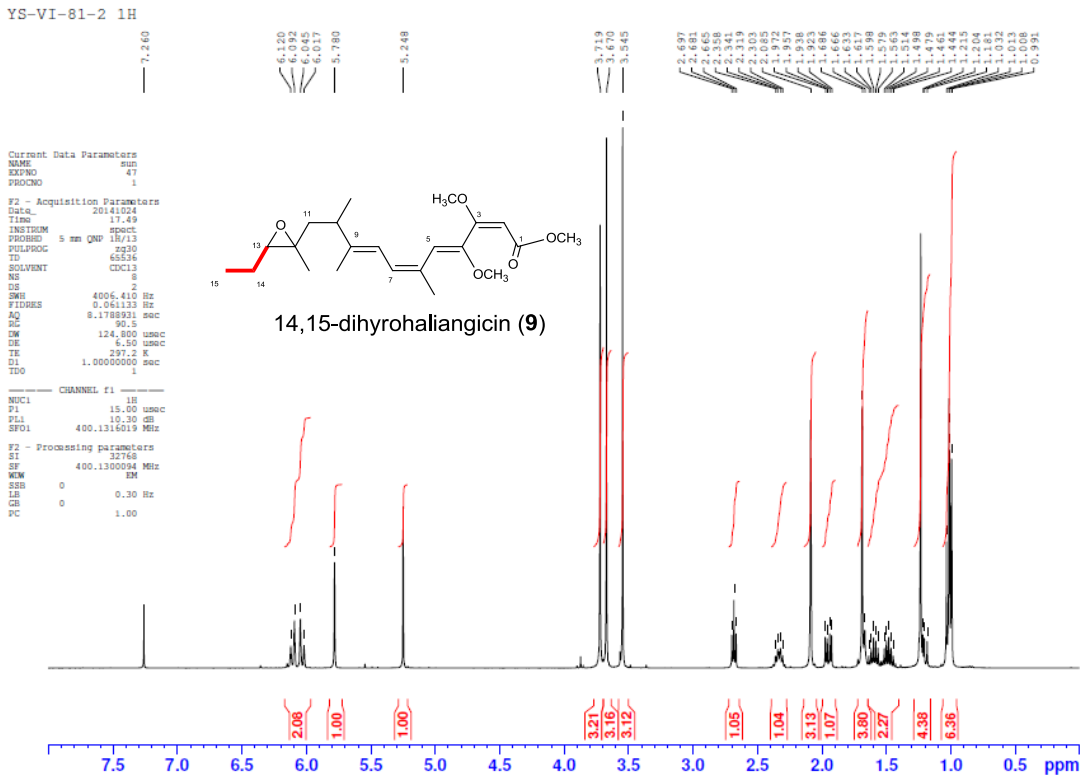


haliangicin D (**6**)

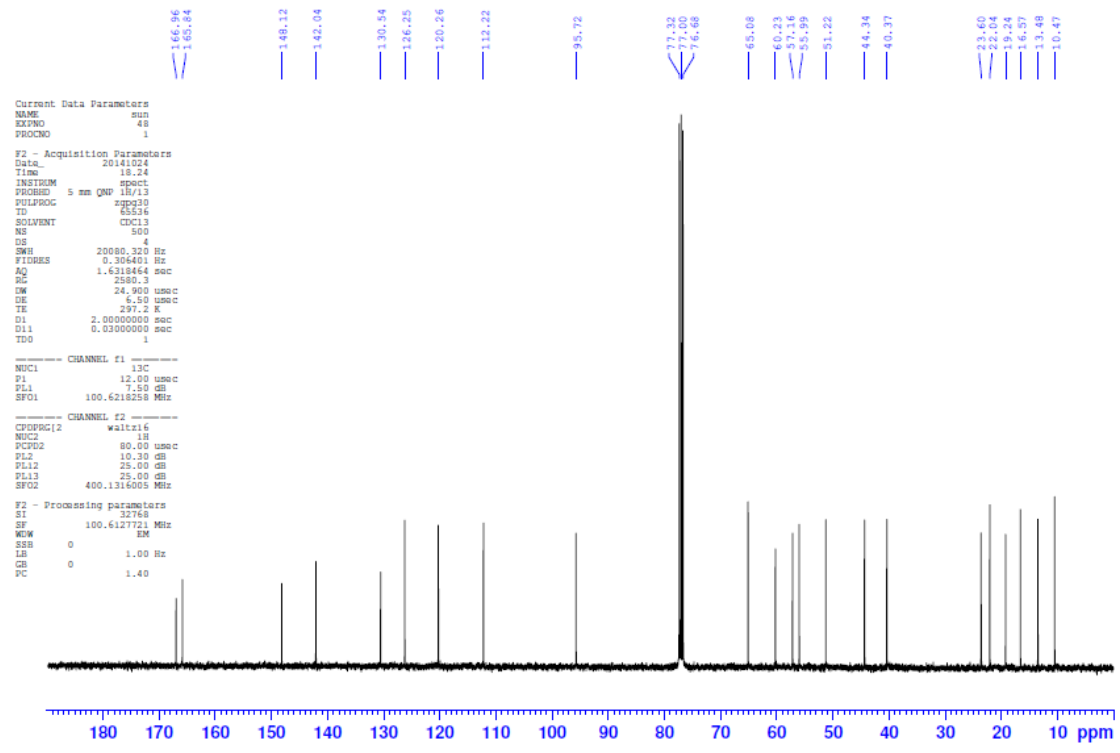
5.3. Spectra of 14,15-dihydrohaliangicin (9).

(**a**) ^{1}H NMR (400 MHz, $CDCl_3$); (**b**) ^{13}C NMR (100 MHz, $CDCl_3$); (**c**) HMQC; (**d**) NOESY; (**e**) UV; (**f**) IR; (**g**) ESI-MS.

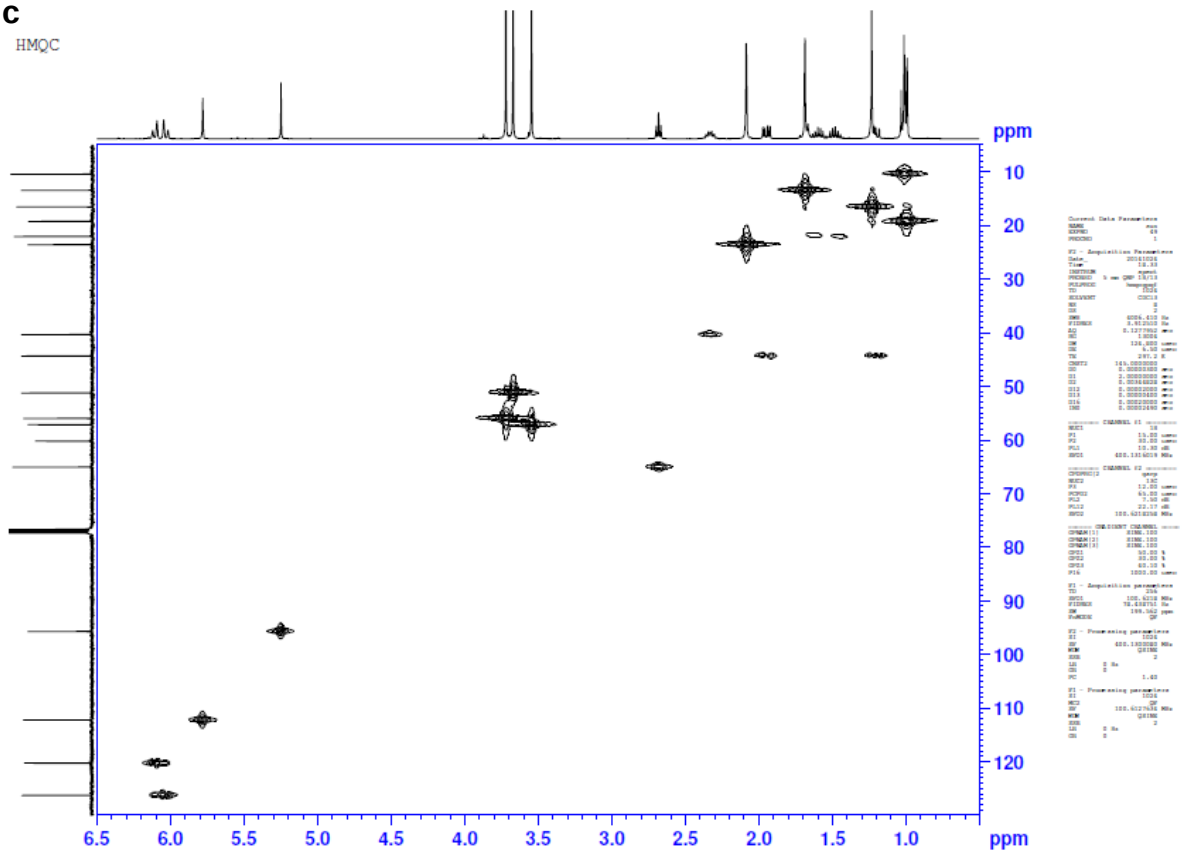
a



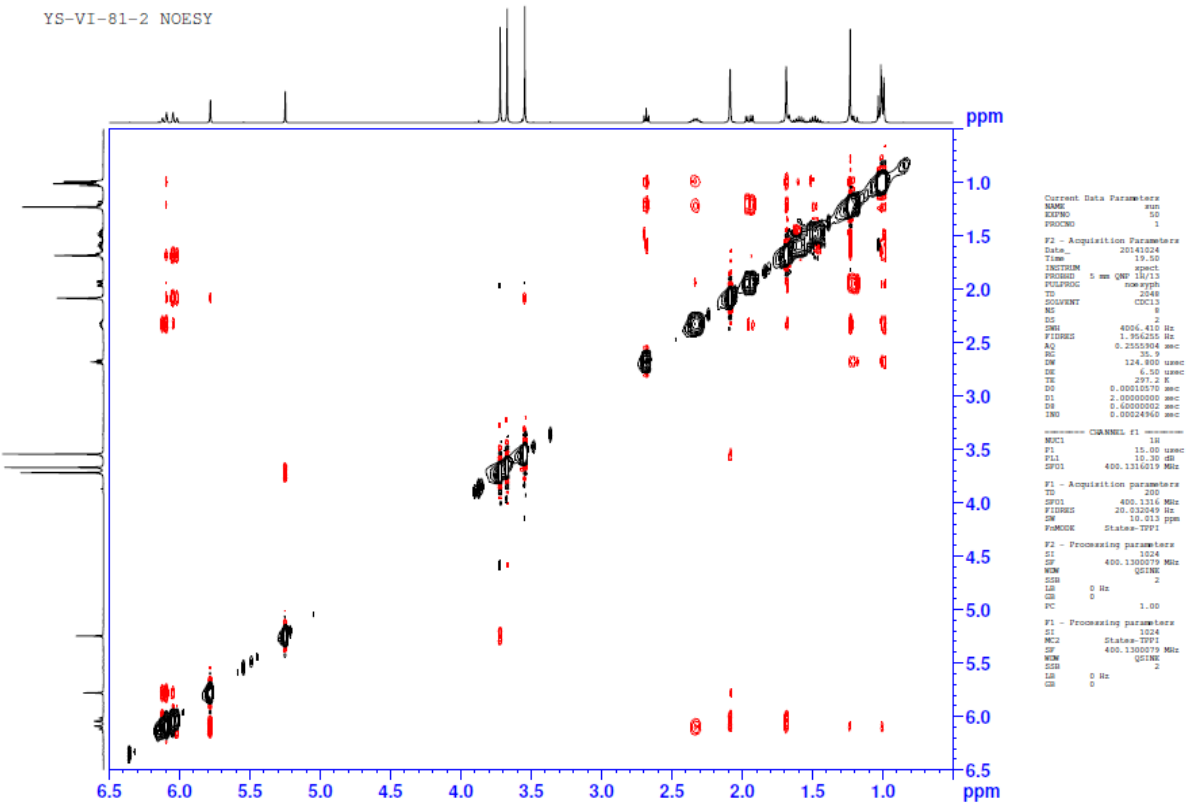
b



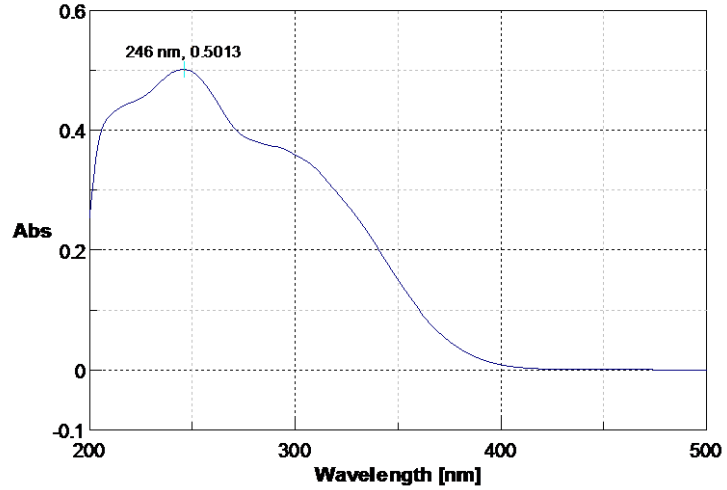
c


HMQC
ppm
10
20
30
40
50
60
70
80
90
100
110
120
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
ppm

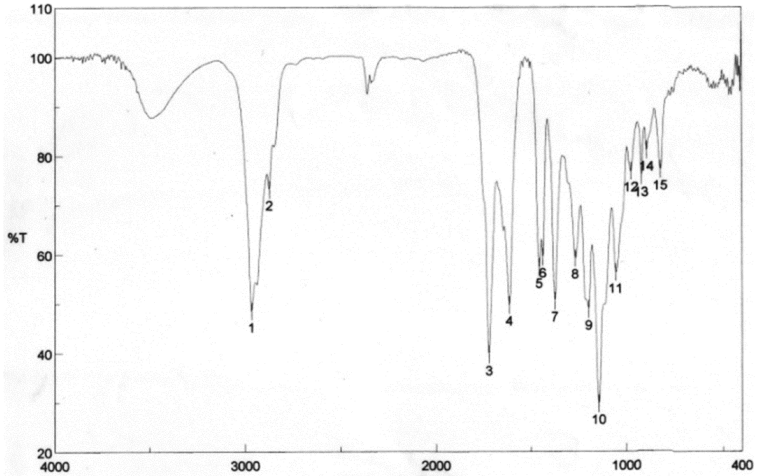
d


YS-VI-81-2 NOESY
ppm
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
ppm

e



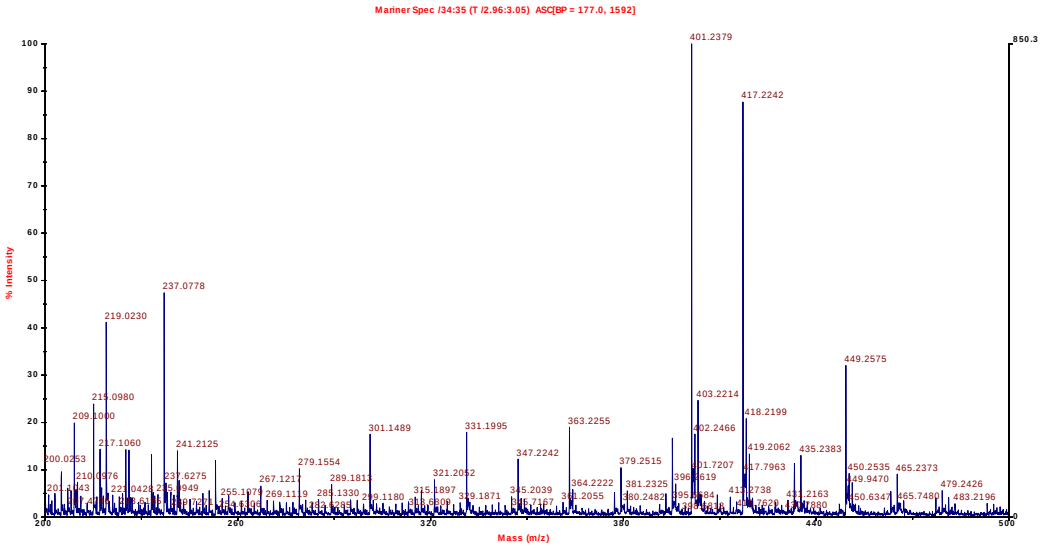
f



Results of Peak Find

No.	Position	Intensity	No.	Position	Intensity	No.	Position	Intensity	No.	Position	Intensity
1	2967.91	48.4289	2	2876.31	73.1488	3	1721.16	39.6228	4	1614.13	49.7053
5	1456.96	57.284	6	1437.67	59.4077	7	1374.03	50.5085	8	1266.04	59.1661
9	1195.65	48.8217	10	1142.62	29.7048	11	1054.87	56.2141	12	976.769	76.7474
13	923.736	75.6548	14	896.737	80.8557	15	824.42	77.0239			

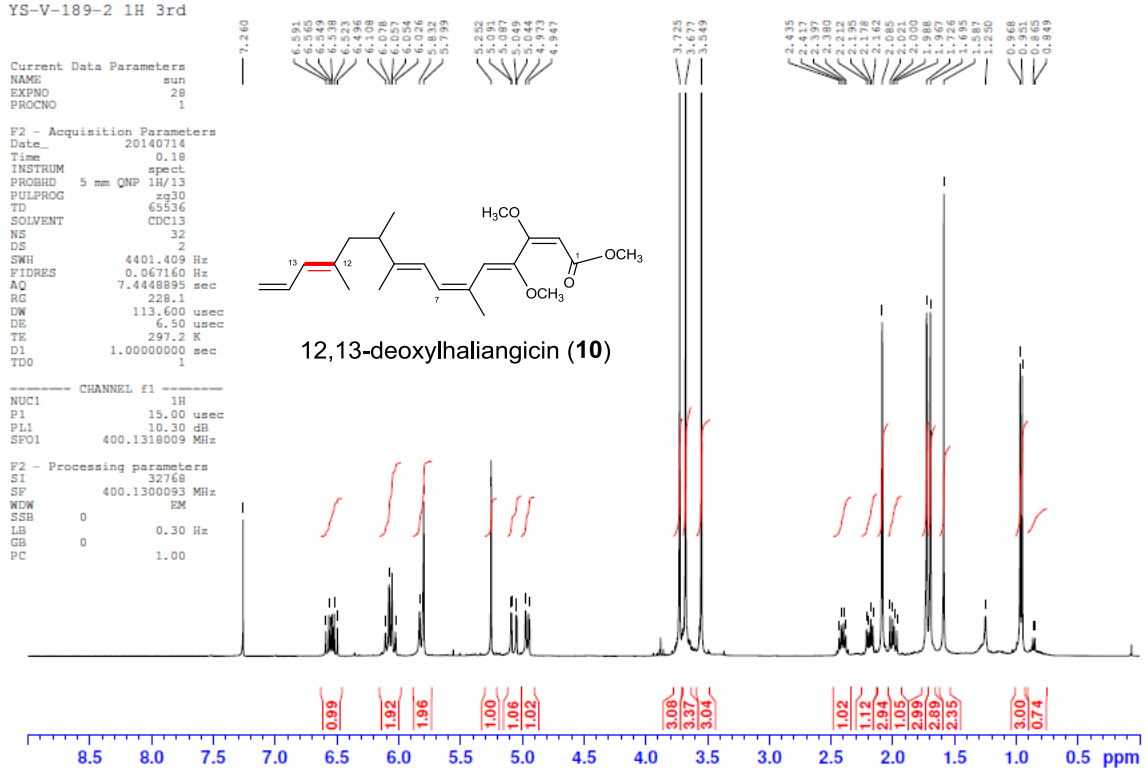
g



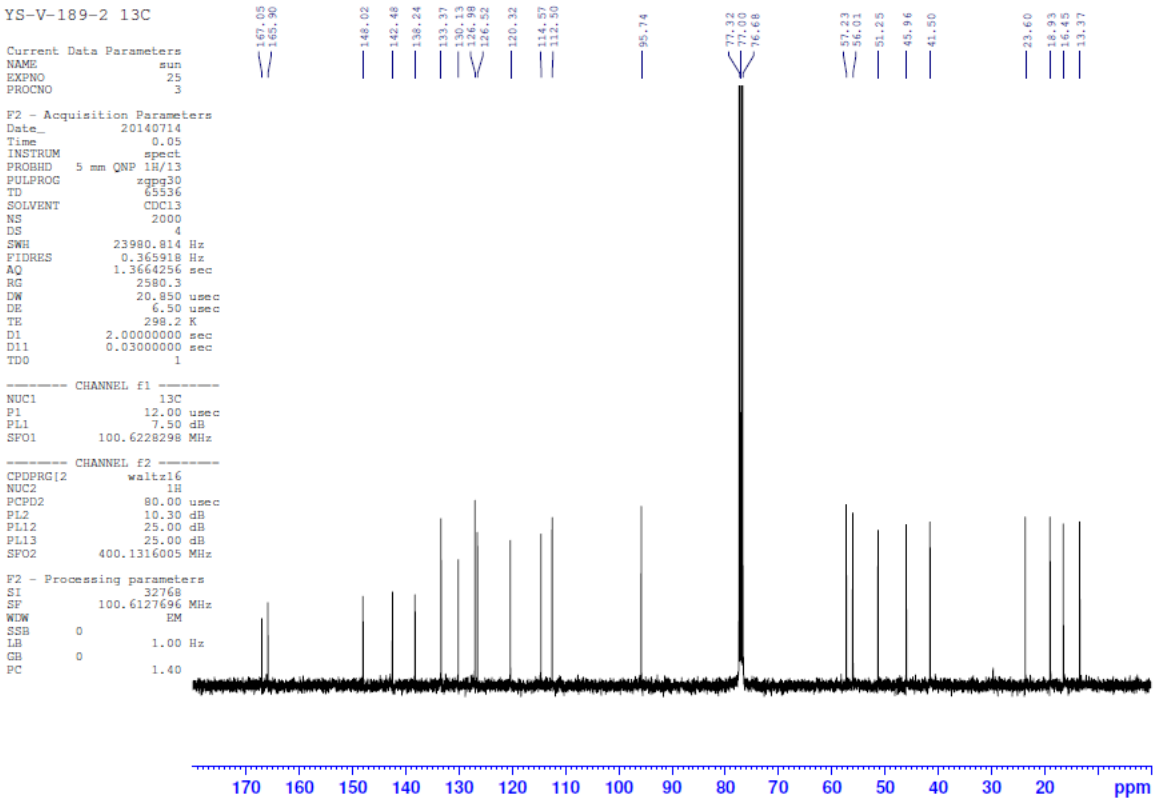
5.4. Spectra of 12,13-deoxylhaliangicin (10).

(**a**) ^{1}H NMR (400 MHz, $CDCl_3$); (**b**) ^{13}C NMR (100 MHz, $CDCl_3$); (**c**) HMQC; (**d**) NOESY; (**e**) UV; (**f**) IR; (**g**) ESI-MS.

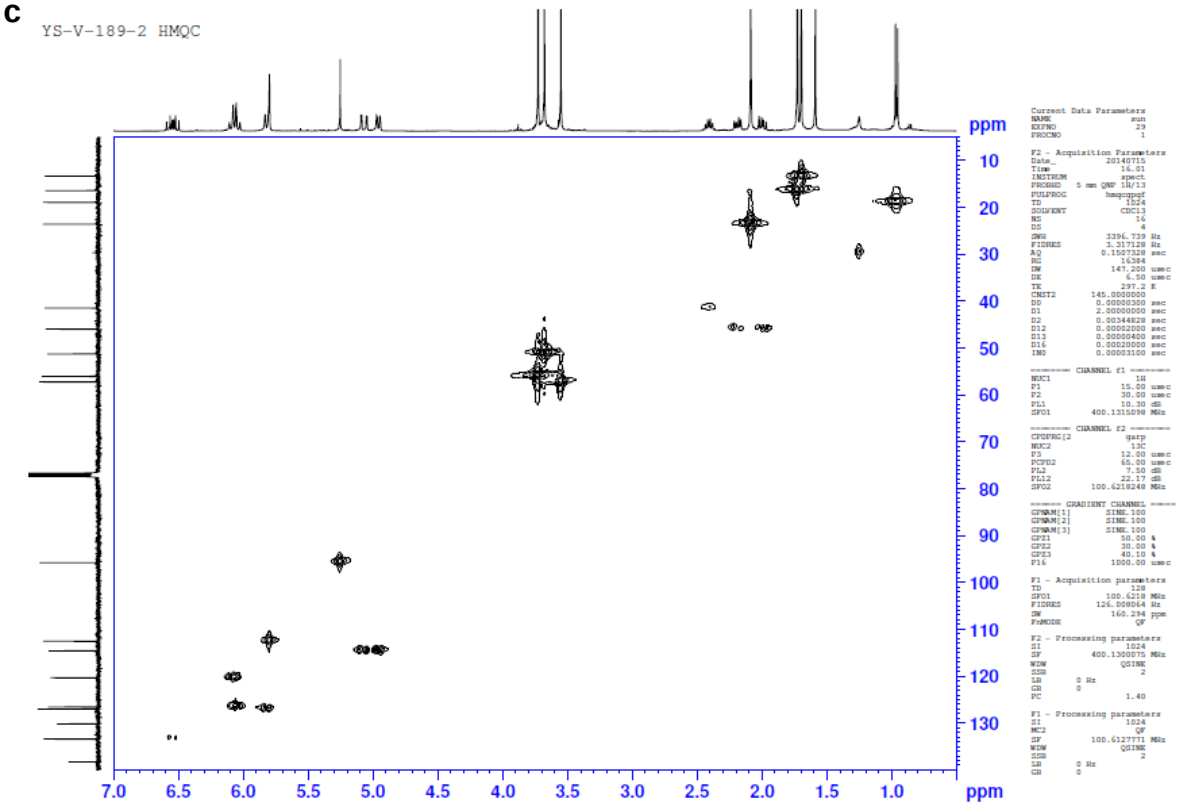
a



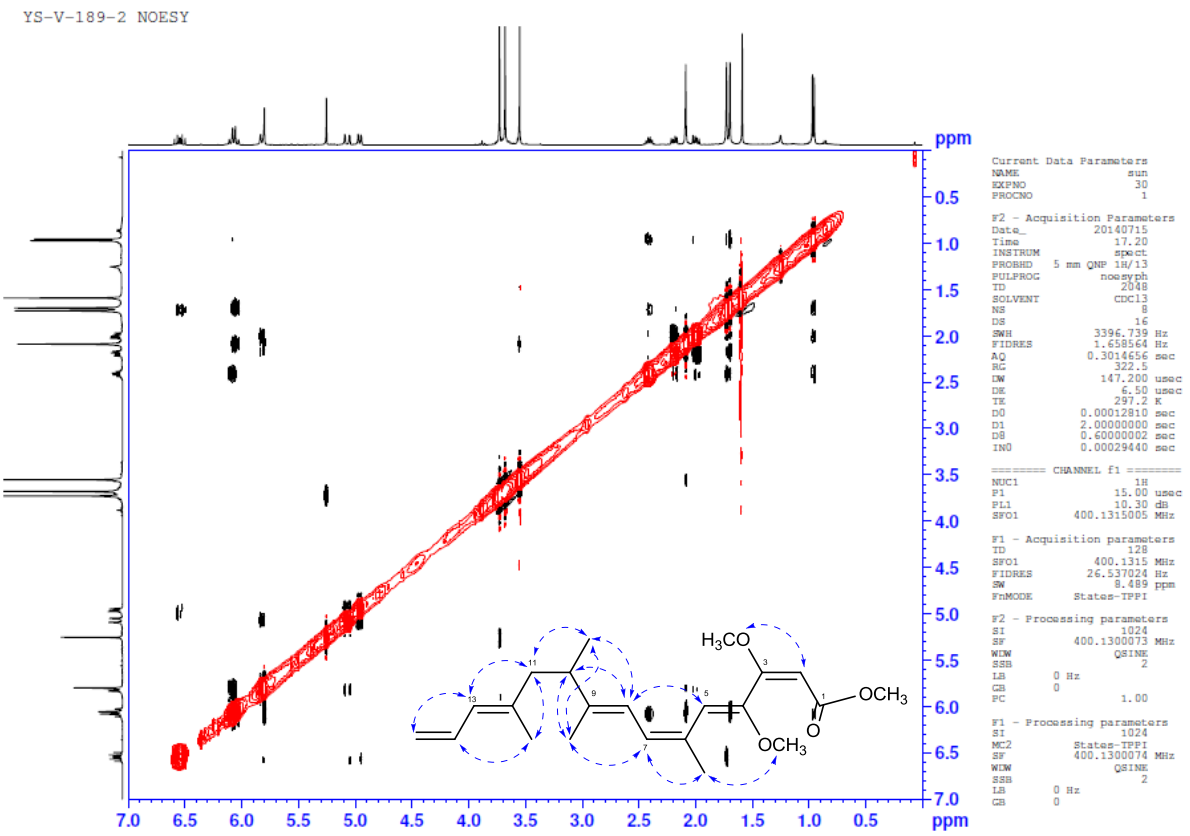
b



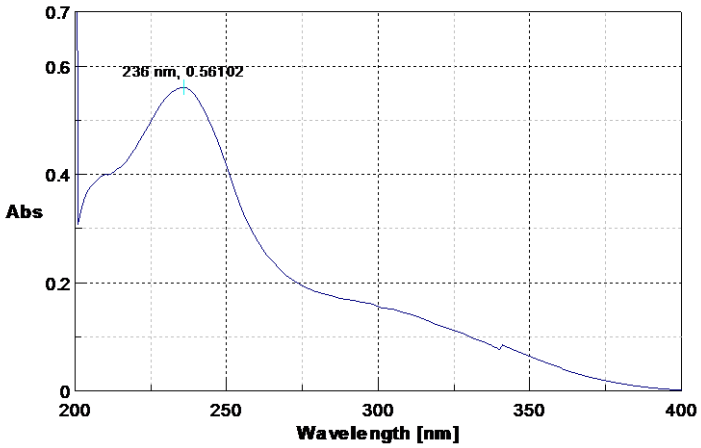
c


YS-V-189-2 HMQC
ppm
7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0
10 20 30 40 50 60 70 80 90 100 110 120 130

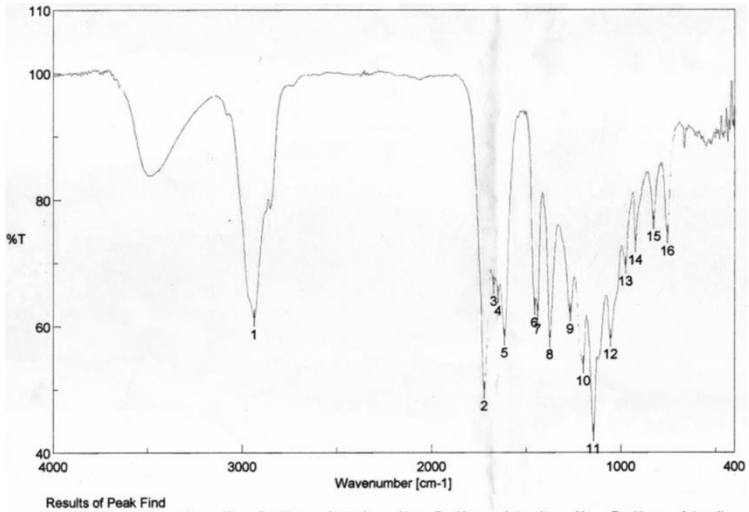
d


YS-V-189-2 NOESY
ppm
7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5
0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0 5.5 6.0 6.5 7.0
H3CO
OCH3
OCH3
O

e



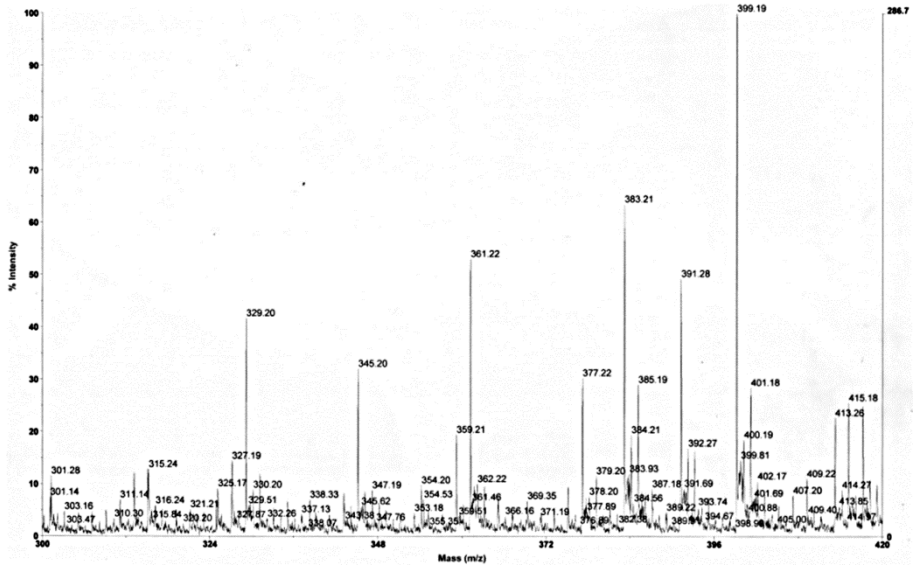
f



Results of Peak Find

No.	Position	Intensity	No.	Position	Intensity	No.	Position	Intensity	No.	Position	Intensity
1	2937.06	61.2488	2	1721.16	49.9209	3	1671.98	66.4737	4	1648.84	64.9475
5	1614.13	58.2573	6	1454.06	63.1251	7	1437.67	62.0902	8	1373.07	58.0122
9	1265.07	62.0927	10	1196.61	53.818	11	1142.62	43.1098	12	1053.91	58.0824
13	976.769	69.6259	14	923.736	72.9976	15	827.312	76.701	16	754.995	74.3712

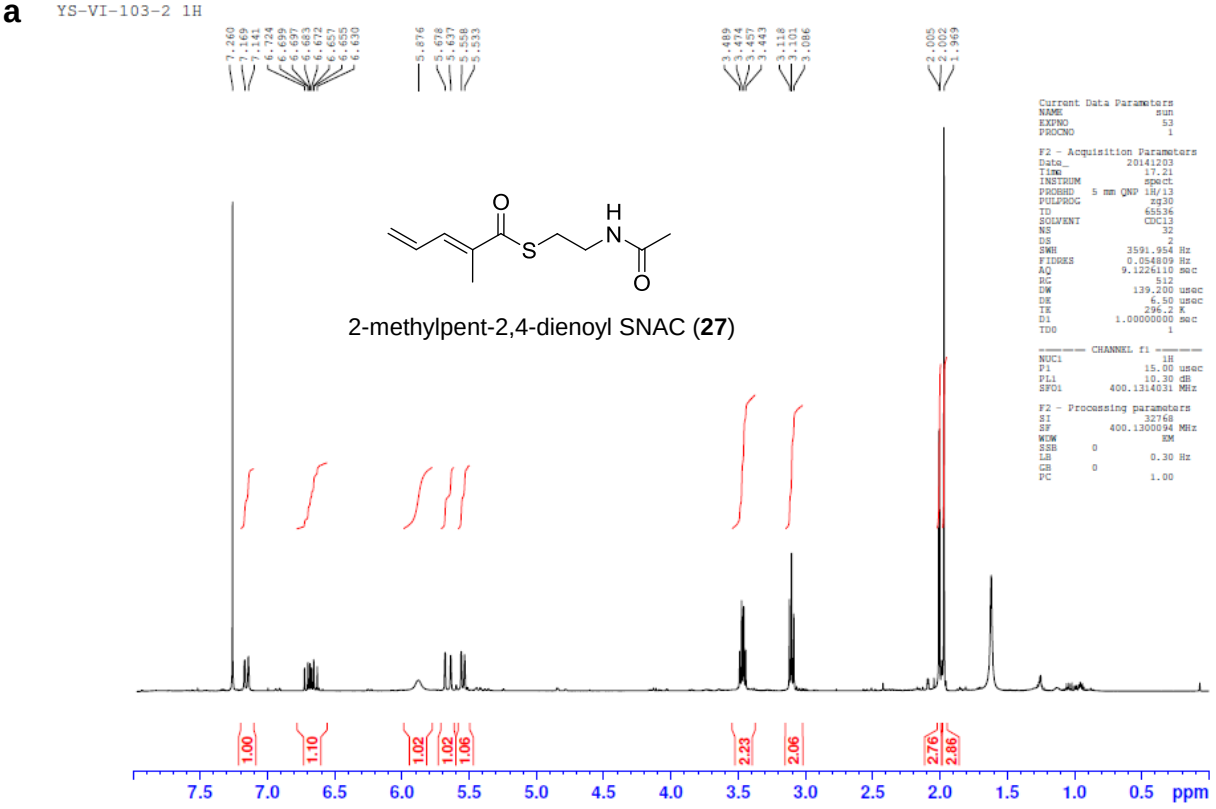
g



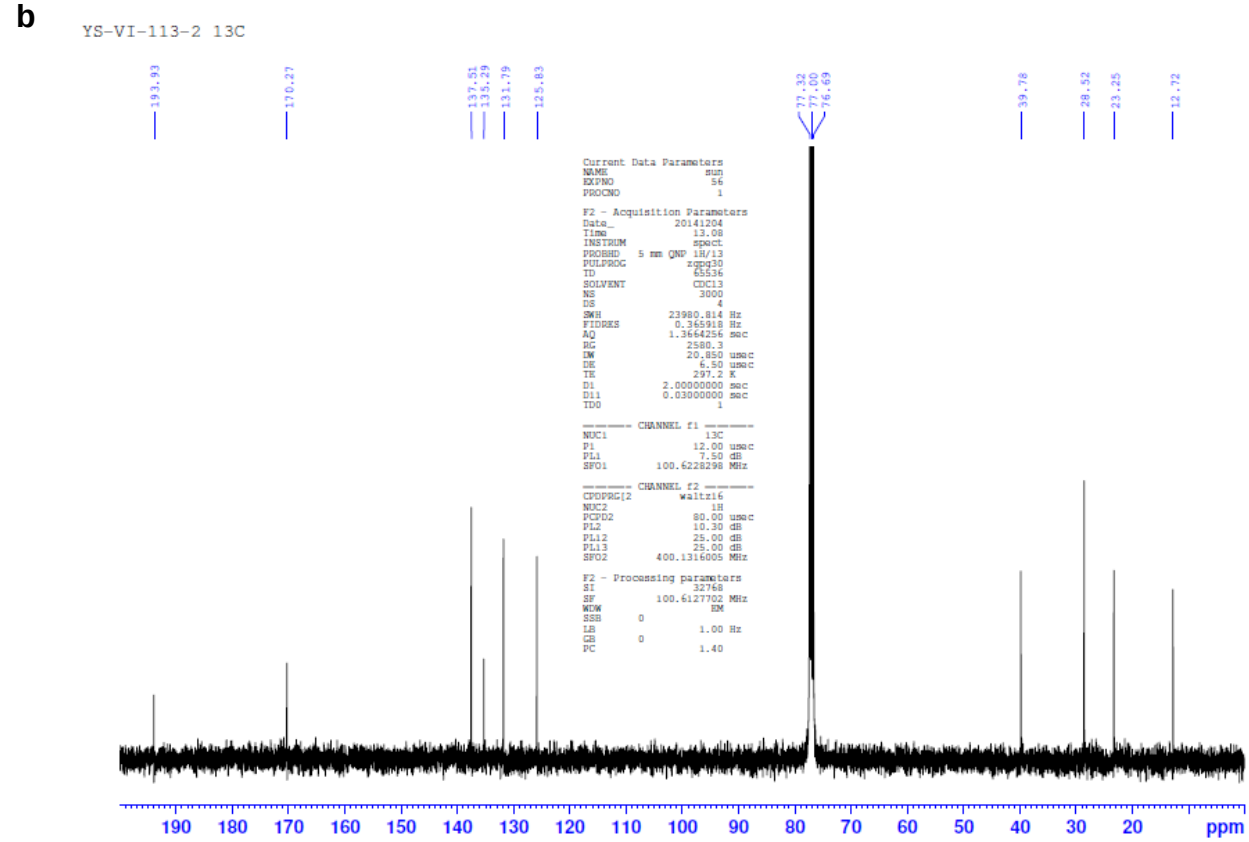
5.5. Spectra of 2-methylpent-2,4-dienoyl SNAC (27).

(**a**) ^{1}H NMR (400 MHz, $CDCl_3$); (**b**) ^{13}C NMR (100 MHz, $CDCl_3$); (**c**) UV; (**d**) IR; (**e**) ESI-MS.

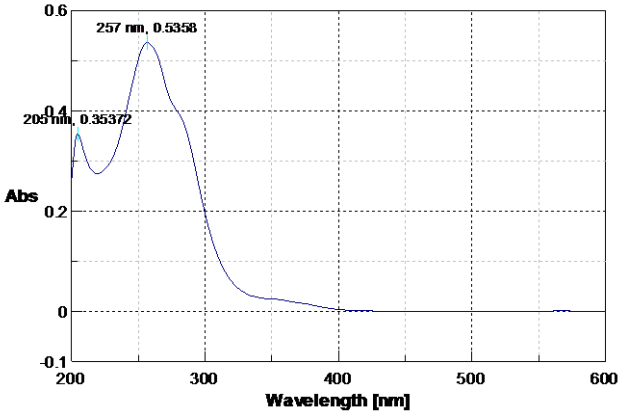
a



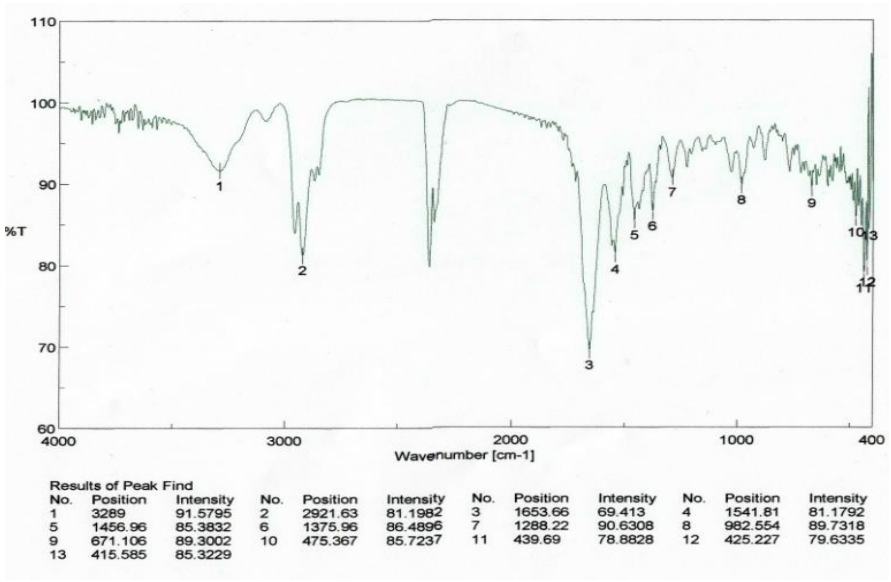
b



c



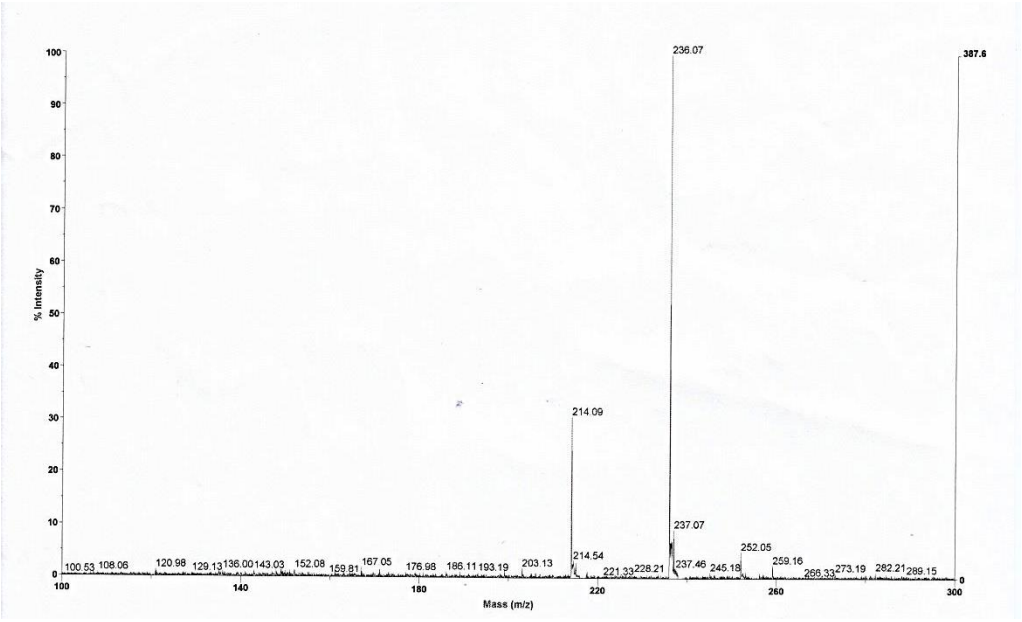
d



Results of Peak Find

No.	Position	Intensity	No.	Position	Intensity	No.	Position	Intensity	No.	Position	Intensity
1	3289	91.5795	2	2921.63	81.1982	3	1653.66	69.413	4	1541.81	81.1792
5	1456.96	85.3832	6	1375.96	86.4896	7	1288.22	90.6308	8	982.554	89.7318
9	671.106	89.3002	10	475.367	85.7237	11	439.69	78.8828	12	425.227	79.6335
13	415.585	85.3229									

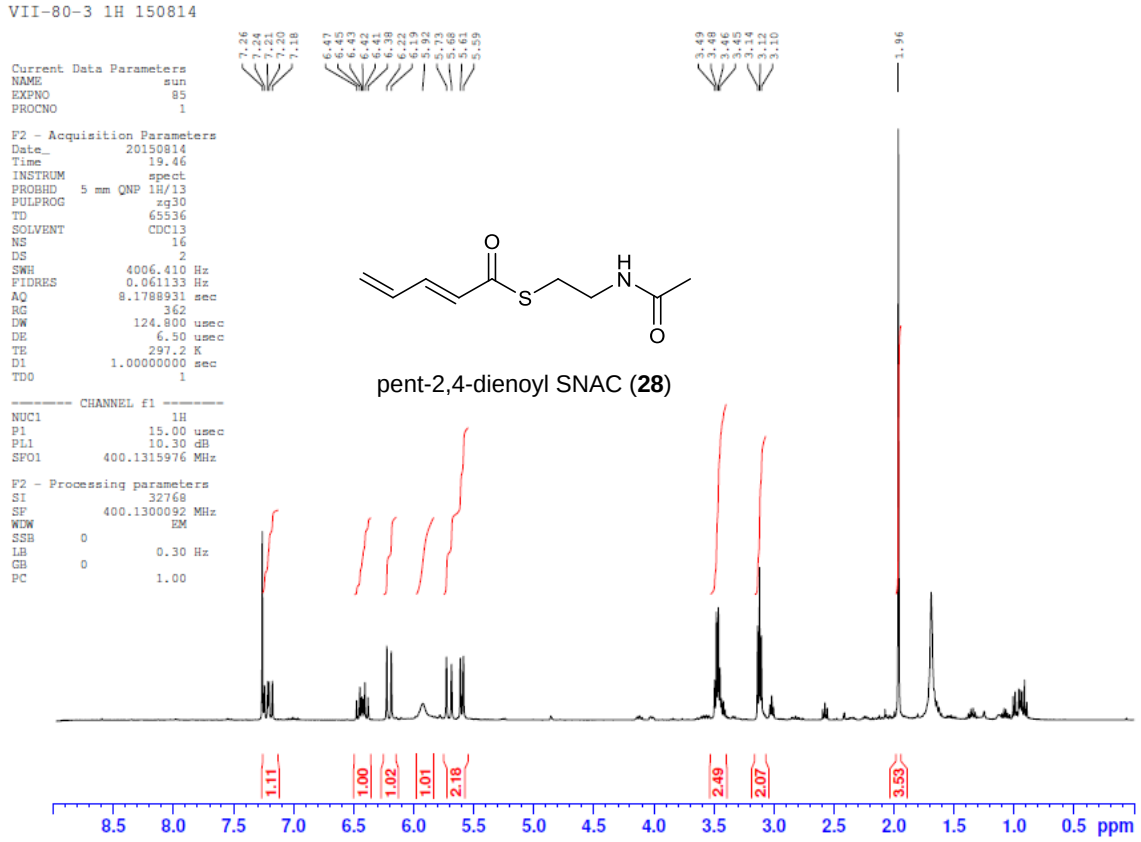
e



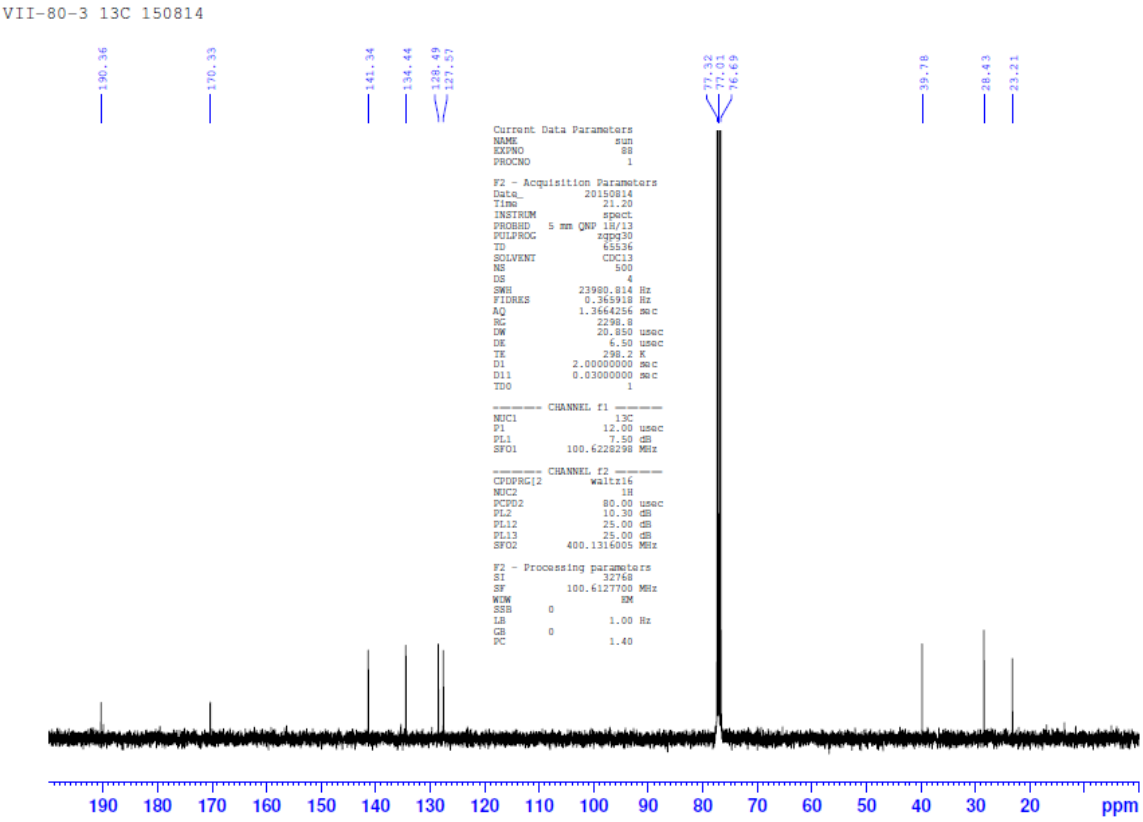
5.6. Spectra of pent-2,4-dienoyl SNAC (28).

(**a**) ^{1}H NMR (400 MHz, $CDCl_3$); (**b**) ^{13}C NMR (100 MHz, $CDCl_3$); (**c**) UV; (**d**) IR; (**e**) ESI-MS.

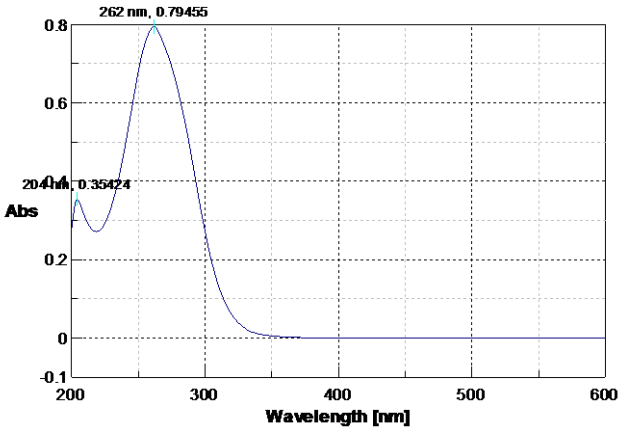
a



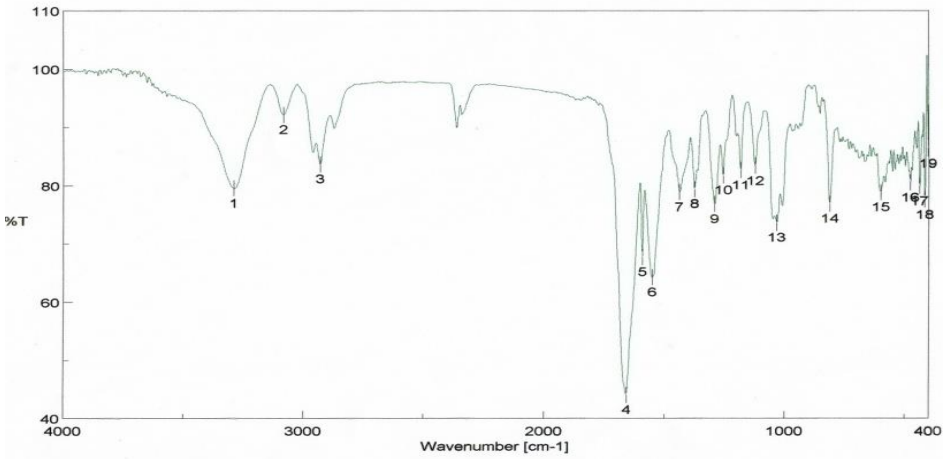
b



c



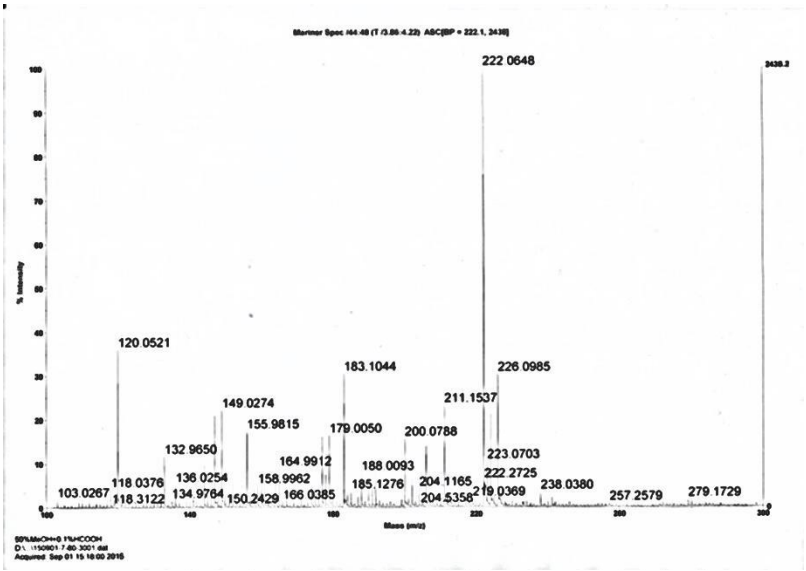
d



Results of Peak Find

No.	Position	Intensity	No.	Position	Intensity	No.	Position	Intensity	No.	Position	Intensity
1	3288.04	79.4597	2	3080.73	92.1127	3	2930.31	83.6381	4	1655.59	43.9675
5	1590.99	67.6686	6	1550.49	64.1995	7	1436.71	78.817	8	1374.03	79.4209
9	1289.18	76.7579	10	1254.47	81.7138	11	1181.19	82.6247	12	1121.4	83.3508
13	1031.73	73.6046	14	810.92	76.8619	15	598.789	78.8038	16	476.331	80.4674
17	437.762	79.8511	18	416.549	77.4858	19	403.05	86.3234			

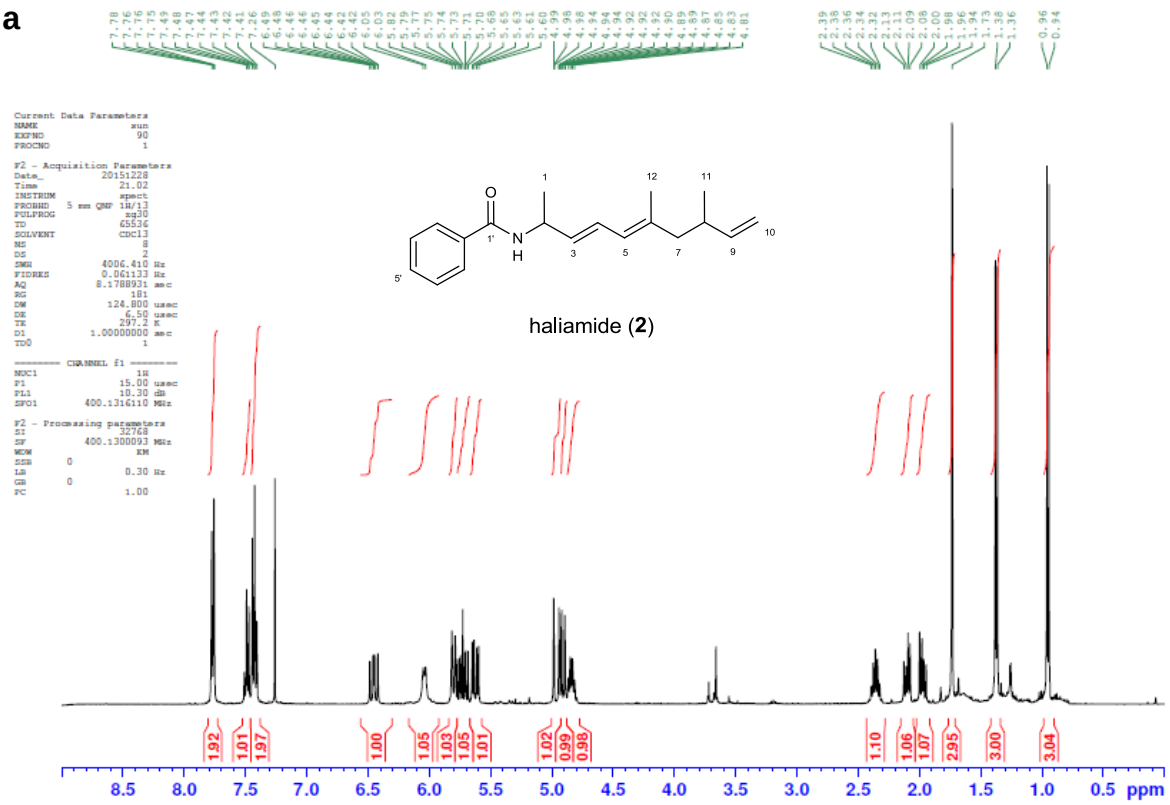
e



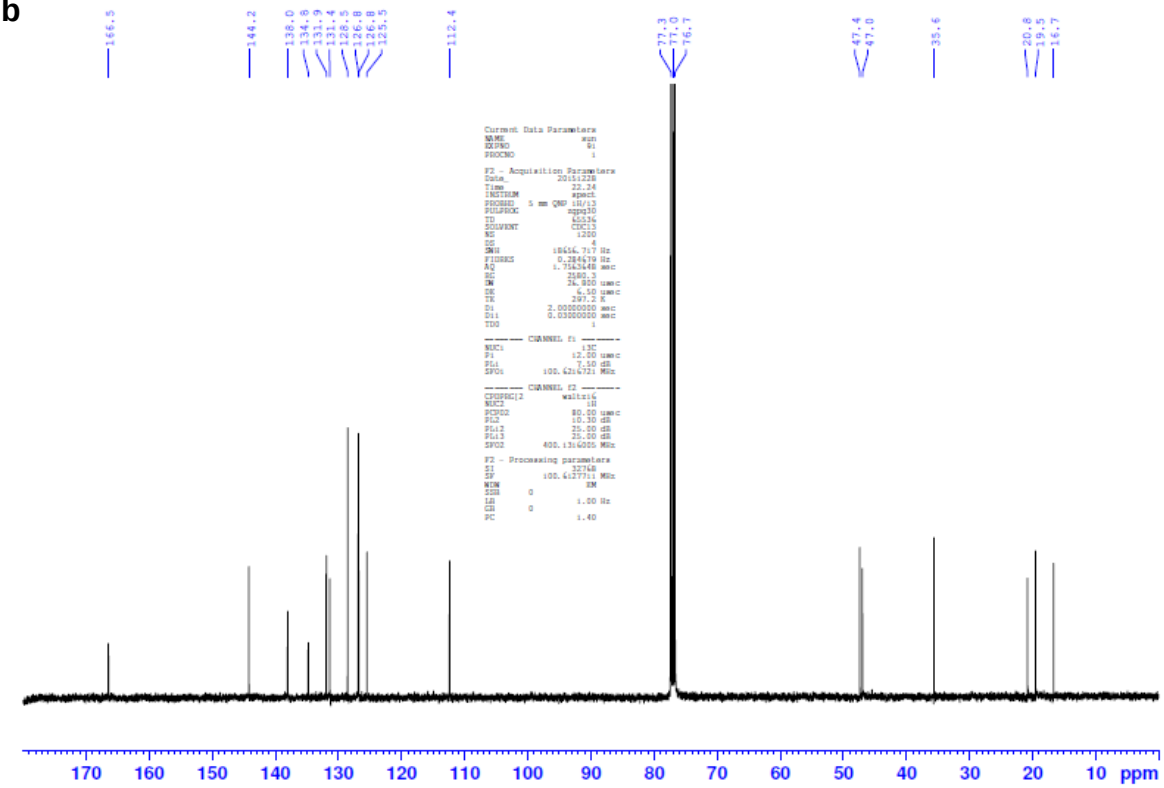
5.7. NMR spectra of haliamide (2).

(**a**) ^{1}H NMR (400 MHz, $CDCl_3$); (**b**) ^{13}C NMR (100 MHz, $CDCl_3$).

a



b



Acknowledgements

The work in this thesis was carried out at the Laboratory of Bioactive Natural Products Chemistry and Molecular Function Modelling, Department of Applied Molecular Biosciences, Graduate School of Bioagricultural Sciences, Nagoya University, Japan. During my entire Ph.D. course and four years life in Japan, there are so many people to whom I would like to express sincerest gratitude.

First, I wish to express the sincerest gratitude to my supervisor, Professor Makoto Ojika, of Graduate School of Bioagricultural Sciences, Nagoya University, for his expert advice and constant encouragement throughout my study. I am very grateful to him for teaching me invaluable knowledge in the field of natural products chemistry and all the skills concerning the chromatographic and spectroscopic techniques. I am also grateful to him for his kind and patient reviewing of my manuscripts in the final stage of my work.

I am deeply thankful to the scholarship provided by the China State-Funded Postgraduates Overseas Study Program from the Chinese Scholarship Council (CSC), which financially supported me to complete Ph.D. course.

I am also sincerely grateful to professors Takashi Tsuge and Hitoshi Mori of Graduate School of Bioagricultural Sciences, Nagoya University, for their kind discussion about the molecular cloning techniques, their kind share of plasmid DNA and permission of using an electroporation system. This study would not be finished without their support. I am also thankful to Dr. Yuichi Oba, Dr. Tatsuhiko Kondo and Dr. Yu Nakagawa in my laboratory for providing invaluable comments on my work and specific instruction of experiments.

I am very grateful to Dr. Zhiyang Feng, Mr. Akira Suzuki and Mr. Seishi Miyano for their excellent previous work on the biosynthesis of haliangicin, which made a solid foundation of my

present research. I am also thankful to Mr. Tomohiko Tomura, Mr. Junichi Sato, Mr. Daiki Hazeyama and Mr. Satoshi Yamazaki for their help and collaboration of bioassay, genetic research and isolation of important compounds.

I particular thanks Dr. Ryosuke Fudou and Takashi Iizuka from Ajinomoto Co. Inc. for their supply of myxobacterium *Haliangium ochraceum*, *Myxococcus xanthus* and suggestion on the fermentation of these microbes, which greatly helped me completing my experiments.

I am very thankful to all my lab mates, who taught me Japanese and gave me useful suggestion on my life in Japan. I will never forget the days full of the memory with all my Japanese friends.

Last, I would like to express my deepest gratitude to my parents, my girlfriend and all my friends in China, who supported me, encouraged me and patiently waited long years for me. I will cherish your kindness in the following years.

January 2016

Yuwei Sun