Doctoral Dissertation

Studies on Natural Products from Untapped Bacterial Resources

(未開拓細菌類を起源とする天然物に関する研究)

Jabal Rahmat Haedar

Laboratory of Natural Product Chemistry

Graduate School of Life Science, Hokkaido University

March 2022

博士学位論文

Studies on Natural Products from Untapped Bacterial Resources

(未開拓細菌類を起源とする天然物に関する研究)

Jabal Rahmat Haedar

北海道大学大学院 生命科学院

生命科学専攻　　生命医薬科学コース

天然物化学研究室

2022年3月

List of Contents

Abbreviation and Symbols

α	Alfa
β	Beta
γ	Gamma
AA	Amino Acid
A domain	Adenylation Domain
ACP Domain	Acyl Carrier Protein Domain
Ala	Alanine
AMP	Adenosine 5'monophosphate
AT	Acetyltransferase
BC	*Bacillus subtillis*
BGC	Biosynthetic Gene Cluster
BM	*Burkholderia multivorans*
C Domain	Condensation Domain
Ca. Entotheonella	*Candidatus* Entotheonella
CAS	Chrome Azurol S
COSY	^{1}H-^{1}H Correlated Spectroscopy
$CDCl_3$	Chloroform-*d*
CI	Chemical Ionization
1D NMR	One Dimension Nuclear Magnetic Resonance
2D NMR	Two Dimension Nuclear Magnetic Resonance
DNA	Deoxyribonucleic Acid
DMEM	Dulbecco's modified Eagles's Medium
DMOX	Dimethyloxazoline
E Domain	Epimerization Domain
EDTA	Ethylenediamine tetraacetic acid

EI	Electron Ionization
ESI	Electrospray Ionization
EC	*Escherichia coli*
FAME	Fatty Acid Methyl Ester
FAAL	Fatty acyl-AMP ligase
FDLA	5-fluoro-2,4-dinitrophenyl Leucinamide
FDA	Flavin Adenine Dinucleotide
ESI-MS	Electrospray Ionization – Mass Spectrometry
Ga(III)	Gallium (III)
GC/MS	Gas Chromatography – Mass Spectroscopy
HCl	Hydrochloric Acid
HMBC	^{1}H-^{13}C Heteronuclear Multiple-Bond Correlation
HPLC	High Pressure Liquid Chromatography
HRESIMS	High Resolution Electrospray Ionization – Mass Spectrometry
HSQC	^{1}H-^{13}C Heteronuclear Single Quantum Coherence
Ile	Isoleucine
Kb	Kilobase
Kg	Kilogram
KS Domain	Ketosynthase Domain
KR Domain	Ketoreducatse Domain
KR	*Kocuria rhizophila*
1 L	1 Liter
Leu	Leucine
LC-MS	Liquid Chromatography – Mass Spectrometry
PKS	Polyketide Synthase
7 M	7 Molar

Mb	Megabase
MeOAc	Methoxy Acetyl
MeLeu	Methyl Leucine
MeVal	Methyl Valine
MeAla	Methyl Alanine
MeIle	Methyl Isoleucine
Me-β-Ala	Methyl β-Alanine
MALDI	Matrix-assisted Laser Desorption/Ionization
MS	Mass Spetrum
MNP	Marine Natural Product
MRSA	Methicillin-resistant *Staphylococcus aureus*
MT Domain	Methyltransferase Domain
NDM	Nutrient-Deprived Medium
NMR	Nuclear Magnetic Resonance
NP	Natural Product
NRPS	Non-Ribosomal Peptide Synthase
ODS	Octadecylsilyl
OTU	Operation Taxonomic Unit
PCP Domain	Peptidyl Carrier Protein Domain
PCR	Polymerase Chain Reaction
PGPR	Plant Growth-Promoting Rhizobacteria
Rf	Retention Factor
RNA	Ribonucleic Acid
RiPPs	Ribosomally-synthesized and Post-translationally modified Peptides
ROESY	Rotating-Frame Overhauser Effect Spectroscopy
T. swinhoei	*Theonella swinhoei*

TE Domain	Thioesterase Domain
TFA	Trifluoroacetic Acid
TLC	Thin Layer Chromatography
Thr	Threonine
UV	Ultraviolet
Val	Valine
WT	Wild Type

Chapter 1
Background and Introduction

1.1 Pharmacologically Important Natural Products

Natural products (NPs) are defined as naturally occurred small molecules produced by any organisms, yet their original functions are mainly idiopathic. One of the growing theories suggested that production of NPs represents chemical defense to deter predators and competitors.[1,2] Intriguingly, these specialized metabolites often exhibit potent biological activities, including antibacterial, cytotoxic, and antiparasitic. Therefore, they are beneficial to be developed as therapeutic agents for treating human diseases. Such potent biological properties are conferred by their unique and complex structures that have evolved in nature for centuries.[2–5] Until today, approximately 60% of clinically used therapeutic agents are derived from NPs and their derivatives, suggesting NPs are essential resources of drug discoveries and developments.[6]

NPs-based medicines have mainly been developed from plants due to the long history of empirical herbal application for traditional medication. A notable example of drug discovery inspired by traditional medicine is the antimalarial agent, artemisinin from *Artemisia annua*, known as sweet wormwood, native to Asia (Figure 1.1).[7] Another example of potential NPs with remarkably complex structure was the anti-cancer agent vincristine isolated from Madagascar periwinkle *Catharantus roseus* (Figure 1.1).[8,9] Through extensive screening of thousands of plant extracts, various plant-derived NPs have successfully been developed for clinical purposes. This is best exemplified by paclitaxel from the bark of the Pacific yew *Taxus brevifolia* (Figure 1.1) as one of the most successful anti-cancer agents.[10,11] NPs not only contribute to drug discovery, but also they can have other essential health benefits, such as dietary supplement and natural dye as shown by curcumin from *curcuma longa* (Figure 1.1).[12]

From the middle of the last century, microbes and marine animals had been gaining power as NPs producers. In 1929, a serendipitous discovery of penicillin G from the filamentous fungus *Penicillium chrysogenum* (Figure 1.1) by Fleming had marked the era for NP explorations from microorganisms.[13] An intensive investigation on anti-infectious agents from other microbial resources, extensively on genus *Streptomyces*, had led to discovery of the first potent anti-tuberculosis drug, streptomycin from *Streptomyces griseus* (Figure 1.1).[14] Additionally, another prolific source of structurally diverse molecules is organisms, especially

invertebrates and their associated microorganisms, inhabiting marine environment. Starting from the mid of 1970, the report on NPs from marine environment had drastically increased resulting thousands of compounds with new scaffolds that have never been observed from other sources.[15–21] A number of marine natural products and their derivatives have been clinically approved as therapeutic options, demonstrating marine environment as a promising source of novel natural products.

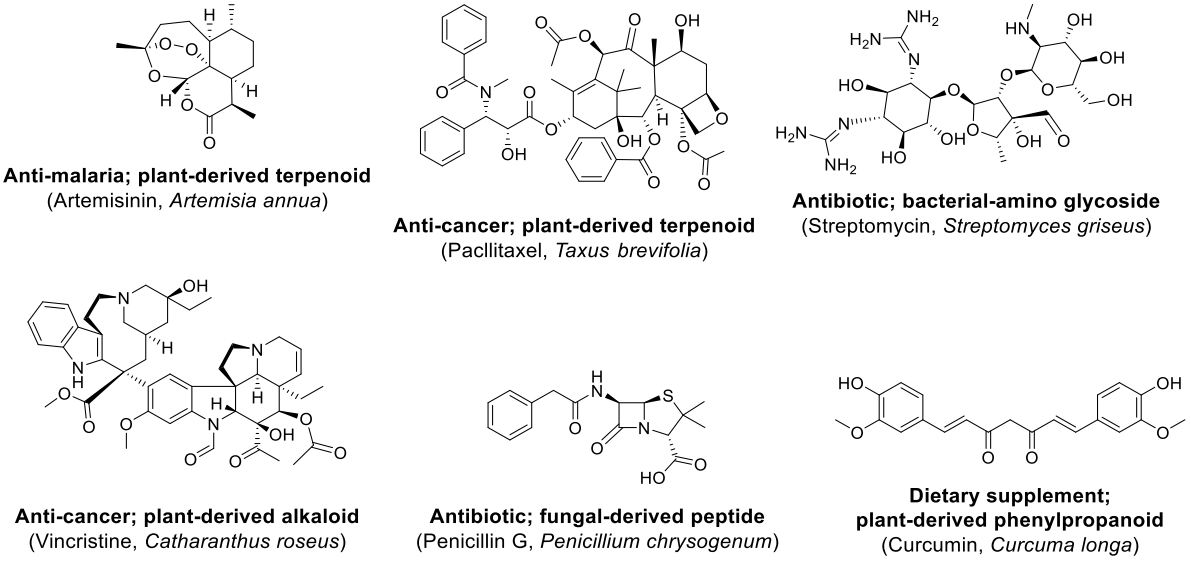


Figure 1.1 Selected NP-derived drugs. Their biological properties and compound classes are typed in bold.

1.2 Natural Products Derived from Marine Environment

More than 70 % of the earth's surface is covered by ocean, suggesting the huge potential of marine environment as reservoirs of diverse NPs accumulated in organisms. Marine ecosystems are home to highly divers organisms in the world. It is found that 32 out of 33 existing animal phyla live in marine habitats. These highlight the great potential of marine biodiversity as rich sources of NPs with potent chemotherapeutic activities.[22] The exploration of marine NPs began around the middle of 1970s, which has contributed to the discovery of an outstanding number of new compounds.[23,24] So far, it has been estimated that over 15,000 NPs reported from various marine organisms.[15,16,19–21]

Sponguridin and spongothymidine isolated from marine sponges *Tectitethya crypta* have inspired the discovery of the first nucleoside type medicines Cytarabine (Ara-C) and vidarabine (Ara-A) (Figure 1.2).[25] Another notable example of marine-derived NP that has successful been developed as a clinically used drug is eribulin mesylate (Figure 1.2), which

was chemically synthesized based on the core-structure of halichondrin B reported from marine sponge *Haliclondria okadai* sp.[26] and *Lyssodendoryx* sp.[27] This compound is medically used to treat metastatic breast cancer[28] and liposarcoma[29]. Another example includes trabectedin, a complex alkaloid isolated from colonial tunicate *Ecteinascidia turbinate.*[30] This compound, also known as ecteinascidin 743, has been approved as an anti-cancer agent under the trade name of Yondelis® for treatment of soft tissue sarcomas (STS)[31] and relapsed ovarian cancer in conjunction with doxorubicin.[32] Furthermore, bryostatin isolated from *Bugula neritina*[33] represents the only marine bryozoan-derived NPs currently under clinical trials for Alzheimer's disease.[34]

Table 1.1 Some of the NPs derived from marine environments that already approved and under clinical trials.

Compound Name	Chemical Class	Isolated Source	Target Molecule	Activity	Status	Ref.
Cytarabine	Nucleoside	Sponge	DNA polymerase	Anti-cancer	Approved	35
Vidarabine	Nucleoside	Sponge	Viral DNA polymerase	Anti-viral	Approved	36
Ziconotide	Peptide	Cone Snail	N-Type Ca channel	Chronic Pain	Approved	37,38
Omega-3-carboxylic acid	Fatty acid	Fish	Triglyceride synthesizing enzymes	Hypertriglyceridemia	Approved	39
Eribulin mesylate	Macrolide	Sponge	Microtubules	Anti-cancer	Approved	40
Trabectedin	Alkaloid	Tunicate	Minor groove of DNA	Anti-cancer	Approved	41
Salinosporamide	Beta-lactone	Bacterium	20S proteasome	Anti-cancer	Phase III	42
Tetrodotoxin	Guanidinium alkaloid	Pufferfish	Sodium Channel	Chronic Pain	Phase III	43
Plinabulin	Diketopiperazine	Fungus	Microtubules	Anti-cancer	Phase III	44
Bryostatin	Macrolide	Bryozoan	Protein kinase C	Alzheimer's Disease	Phase II	34

While most of the approved drugs derived from marine environment belong to anti-cancer agents, several marine NPs without anti-cancer activities, such as ziconotide and tetrodotoxin, have been approved as clinically used medicines or currently under clinical trials (Table 1.1). Ziconotide represents the first marine derived NP approved for reducing chronic pain currently marketed under the trade name of Prialt®. This non-narcotic analgesic peptide is a component of a cocktail venoms used by a cone snail *Conus* to paralyze its prey.[45] Another marine pain-reducing compounds is tetrodotoxin (Figure 1.2), a potent neurotoxin originally isolated from the Japanese pufferfish "fugu".[43,46,47] The ability of this toxin to selectively

blocks voltage-gated sodium channel makes it suitable to be developed as a therapeutic agent for pain management.[43] Omega-3-carboxylic acid from fish is another example of non-anti-cancer marine drug (Figure 1.2) applied for treatment of severe hypertriglyceridemia.[39]

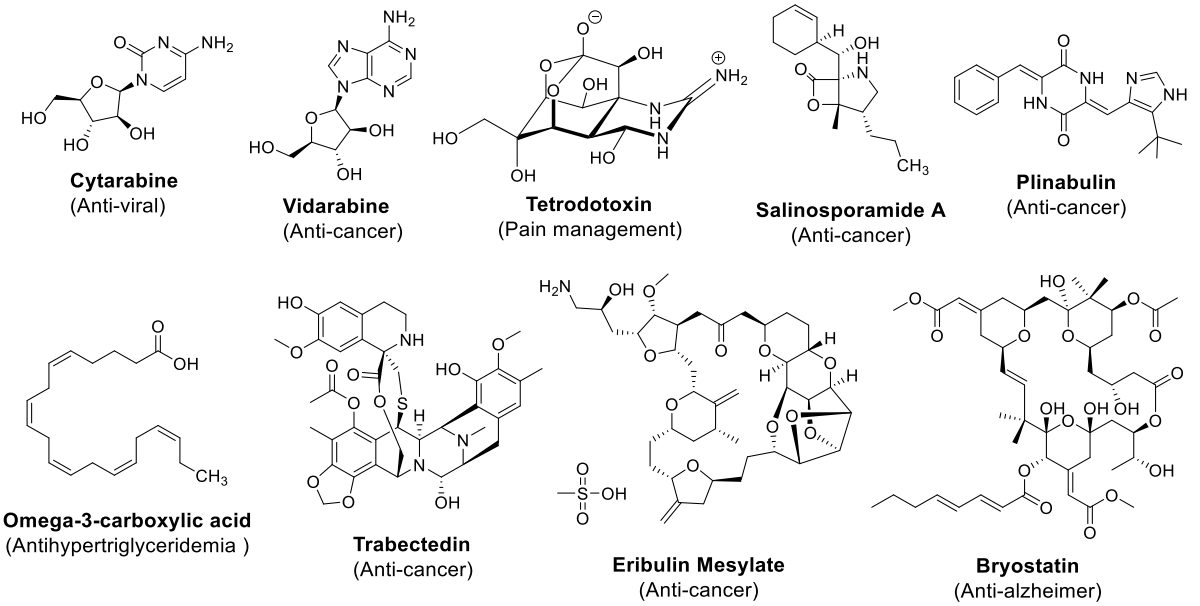


Figure 1.2 Several approved and under clinical trial compounds derived from marine natural products.

Marine microorganisms within groups of bacteria, cyanobacteria, fungi and dinoflagellates also contribute as the main sources of NPs.[15,18,19,21] Salinosporamide A and plinabulin represent two marine microbial compounds potentially developed as anti-cancer drugs (Table 1.1). Pionering work of Fenical and colleagues had led to the isolation of salinosporamide A from an obligate marine bacteria *Salinospora tropica*.[42,48] This potent 20S proteasome inhibitor is currently under phase III clinical trials as an anti-cancer agent. Plinabulin is a marine fungal-derived diketopiperazine derivative recently entered phase III clinical trials for non-small cell lung cancer by targeting distinct microtubule of cancer cells.[44]

For the last few decades, it has been increasing evidences that many NPs from marine macro-organisms are actually produced by symbiotic bacteria.[49] In 1996, Faulkner and colleagues proposed that several compounds reported from the marine sponge *Theonella swinhoei* are produced by the sponge-associated filamentous heterotrophic bacteria.[50] Recent development of DNA sequencing technologies supported by advances in analytical chemistry instruments have contributed greatly to our understanding on the importance roles of marine symbiotic bacteria as NP producers. For instance, using meta-omics techniques, Sherman and colleagues successfully confirmed the tunicate's symbiont *Candidatus* Endoecteinascidia

frumentensis as the actual producer of trabectedin.[51–53] In addition, bryostatin from bryozoan *Bugula neritina* was proposed to be produced by the as-yet uncultured symbiont bacteria "*Candidatus* Endobugula sertula" based on localization studies using CARD-FISH.[54–59]

By metagenome sequencing and single-cell analyses combined with functional studies, Piel and colleagues showed that almost all 40 NPs previously reported from marine sponge *T. swinhoei* were attributed to as-yet uncultivated bacterial symbionts "*Candidatus* Enthotheonella sp." as the producers.[60,61] These reports are emphasizing the great potential of uncultured symbiotic bacteria associated with marine invertebrates. Therefore, accessing these promising yet untapped marine bacterial resources by any approaches would lead to discovery of bioactive small molecules.

1.3 Natural Products from Terrestrial Microorganisms

NPs from microorganisms have become interesting targets for drug discovery since the accidental discovery of penicillin G from the filamentous fungi *Penicillium chrysogenum* (Figure 1.1) by Alexander Fleming in 1928. For 20 years from discovery of penicillin as a potent anti-infectious agent is known as the golden era of antibiotics. The intensive investigations of NPs were focused not only on filamentous fungi but also on other potential bacteria. A notable example is the bacterial genus *Streptomyces* identified as the prolific producer of streptomycin (Figure 1.1),[14] a potent antibiotic clinically used to treat a number of bacterial infections including tuberculosis,[62] endocarditis,[63,64] and brucellosis[65]. Following these compounds, thousands of new anti-infectious agents have been reported from terrestrial microorganisms, demonstrating the potential of microorganisms as prolific source of bioactive small molecules.

Discovery of anti-infectious agents was the main goal at that time following the successful of penicillin G and streptomycin. Consequently, many classes of antibiotics, representing here by the macrolide erythromycin and the glycopeptide vancomycin (Figure 1.3), have been approved for clinical purposes.[13] Erythromycin, for instance, was isolated from *Streptomyces erythreus*[66] and has been approved as a therapeutic agent for treating gram-positive and gram-negative infections including syphilis,[67] whooping cough,[68] respiratory[69] and gastrointestinal infections[70]. Biosynthesis of this compound is catalyzed by a polyketide synthase (PKS) followed by glycosylation.[71] Vancomycin is a complex glycopeptide firstly reported from *Amycolatopsis orientalis*,[72] which is known as one of the essential antibiotics to treat methicillin-resistant *Staphylococcus aureus* (MRSA). Biosynthesis of this complex

glycopeptide in *Amycolatopsis orientalis* is mediated by 7 modules of non-ribosomal peptide synthase (NRPS) assembly lines, followed by a set of distinct tailoring enzymes responsible for glycosylation and oxidation to generate the final product vancomycin.[73,74] Amphotericin B from the soil bacterium *Streptomyces nodosus* represents anti-fungal NPs (Figure 1.3) that has been approved for life-threatening fungal infections caused by *Aspergillus.*[75,76] This compound is biosynthesized by a huge and complex PKS system consisting of 19 modules and several tailoring enzymes.[77]

Many NPs derived from terrestrial microorganisms have been evaluated for anti-cancer activities, exemplified by bleomycin and daunorubicin (Figure 1.3). Bleomycin is a glycopeptide isolated from *Streptomyces verticillus* and has been clinically used for treatment of several cancer diseases including, testicular, ovarian, and Hodgkin's lymphoma.[78] Bleomycin biosynthesis was proposed to be catalyzed by a large PKS and NRPS hybrid through the assembly of ten simple building blocks into a linear peptide followed by glycosylation steps.[79] Daunorubicin is a member of anthracycline type compounds along with doxorubicin, epirubicin, and idarubicin.[80] These compounds inhibit progression of cancer cell by DNA intercalation, thereby interfering DNA and RNA production.[81] Daunorubicin is biosynthesized by a type II PKS systems in *Streptomyces peucetius* to produce the anthracycline backbone followed by glycosylation.[82]

Further examples of potent NPs from terrestrial microorganisms are rapamycin, ivermectin, and lipstatin (Figure 1.3). Rapamycin was originally isolated as an antifungal agent from *Streptomyces rapamycinicus*. However, this compound has been approved for clinically used drugs as anti-proliferative and immunosuppressive agent.[83] Recently, this compound was also reported to prolong life span of diverse species.[84] In addition, ivermectin is an FDA-approved antiparasitic drug to treat disease caused by roundworm and ectoparasites.[85] This compound was isolated from *Streptomyces avermitilis.*[86] Finally, lipstatin is the first anti-obesity drug derived from a microbial NP. This compound was reported from *Streptomyces toxytricini* to have activity as an inhibitor for a pancreatic lipase.[87] Interestingly, lipstatin is biosynthesized through unique condensation of three fatty acids to form a β-lactone ring and three fatty acyl moieties.[88]

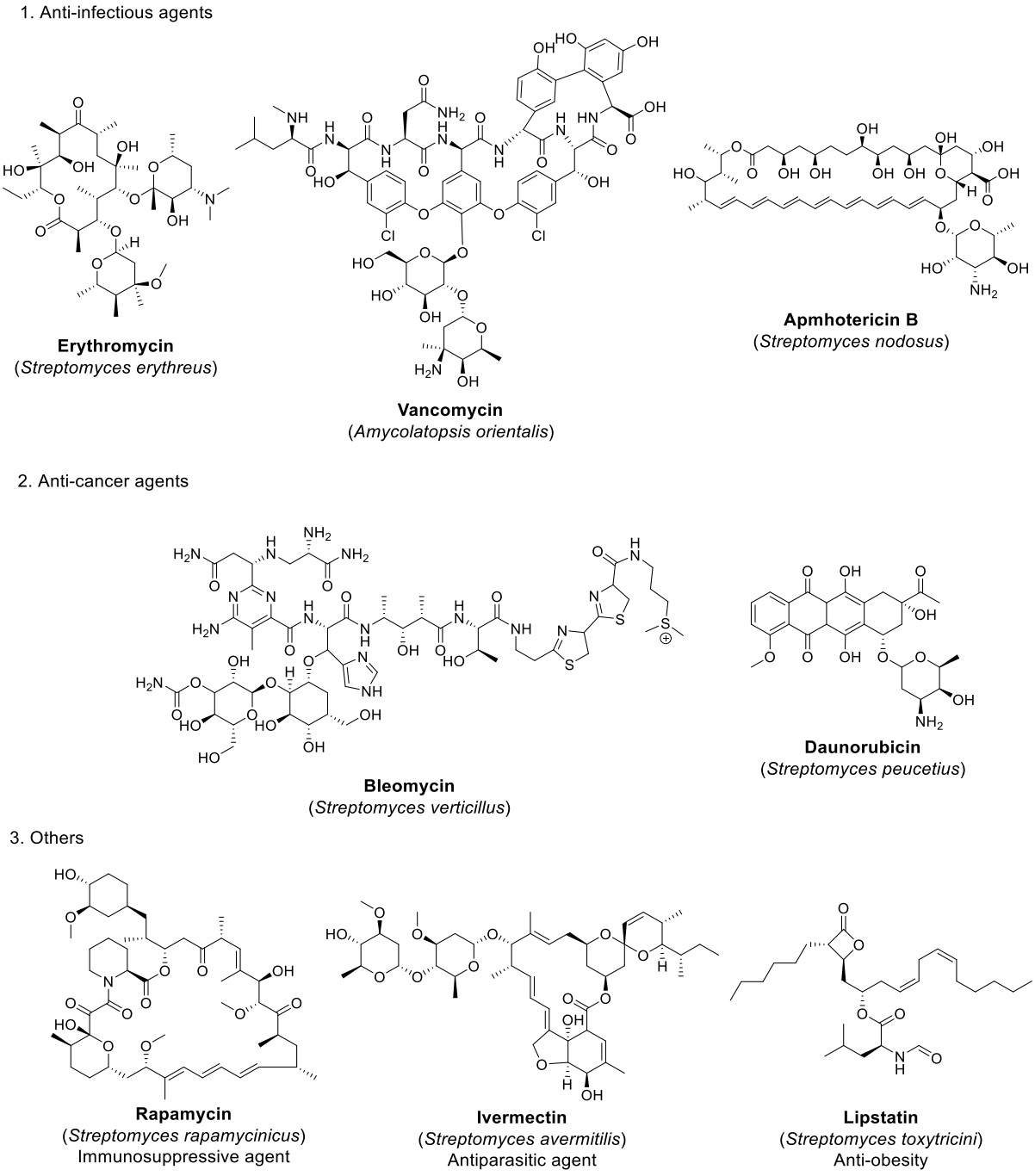


Figure 1.3 Several approved NPs derived from terrestrial microorganisms.

This recent post-genomic era has been marked with the accessibility toward the huge accumulation of genome sequence of many organisms deposited in GenBank. This allows intensive genome mining using various bioinformatic tools to identify microbial genes encoding for secondary metabolites. For example, through whole genome sequence analysis, it was found that the rapamycin producer *Streptomyces rapamycinus* has 52 NP biosynthetic gene clusters spanned for 3.09 megabase (Mb).[89] This indicates that 24.3% of total genome

12.70 Mb are dedicated to secondary metabolite productions. Recent development of molecular biology techniques such as metagenomic approaches and DNA sequencing technologies enables identification of enormous biosynthetic gene clusters (BGCs) in uncultivated microorganisms inhabiting diverse soil ecosystem.[90–92] Therefore, cross-interdisciplinary collaborations between different scientific fields are crucial to access the bioactive small molecules encoded on biosynthetic potential genes.

1.4 Summary and Research Goal

Natural products (NPs) are one of the main sources of drug discovery and development. It contributed to approximately 60% of all approved drugs for clinical applications.[6] Generally, NPs with chemotherapeutic potential are characterized by their diverse and complex structures useful for biological activities. Among a large number of NPs discovered from plants, marine organisms, and microorganisms, many of them have been clinically approved for medical therapeutics. However, there are increasing and urgent needs for new bioactive agent to treat various new diseases.

Analysis of genomic information derived from marine invertebrates-associated microorganisms and soil microbes revealed the presence of enormous number of biosynthetic gene clusters (BGCs) encoding for the production of new bioactive molecules, suggesting that the exploration on these untapped resources remains widely open. Furthermore, recent genetic studies have indicated that many BGCs detected in marine invertebrates are actually located in symbiotic bacteria.[60,61] These suggests that targeting these untapped microbial strains from both marine and soil ecosystems would lead to discovery of novel bioactive natural products.

This study covered two research project exploring NPs from untapped microorganisms in marine invertebrate and soil ecosystem. The first project targets novel NPs from potential symbiotic bacteria associated with marine sponge *Theonella swinhoei.* The second project focuses on isolation of potential soil microbes by employing *in situ* cultivation approach. Investigating the isolated potential strains through metabolite analysis led to the discovery of two groups of siderophores.

Chapter 2

Natural Products from Marine Sponge *Theonella Swinhoei*

2.1 Introduction

Marine sponge has been sitting in the sea floor for a very long period of time. It is suggested through sterane biomarkers analyses that this organism has been already existed from Cryogenian period (calc. 720–635 million years ago).[93] Marine sponges are sessile invertebrates, and therefore they rely on cytotoxic compounds present in their body's tissues to defend themselves from predators, such as fishes.[94] There has recently been increasing evidence that bacteria symbiotically associated with marine sponges are the actual producers of such defensive compounds.[50] Marine sponges pump large volumes of seawater, allowing microbes to enter the labyrinth-like aquiferous system located in the mesohyl layer of the sponge body (Figure 2.1).[95] Some microbial cells establish symbiotic relationship as a part of the sponge-specific symbionts,[96] and may contribute to chemical defense in the sponge host by releasing bioactive compounds.

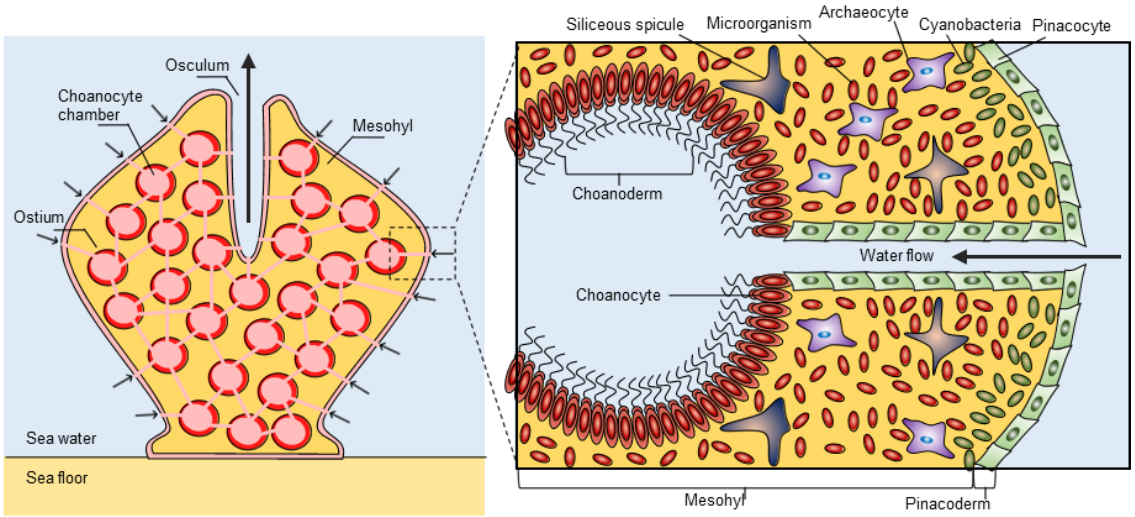


Figure 2.1 a. Typical anatomy of a demosponge; **b.** Internal structure of an aquiferous system inside sponge's mesohyl tissue.

The symbiotic bacteria are well distributed in all sponge tissue. In general, the outer light-exposed layer of the sponge's body is mostly inhabited by the photosynthetic bacteria, cyanobacteria, that usually confer coloration of the host.[97] Heterotrophic and autotrophic bacteria typically inhabit the internal mesophyll parts of the sponge's body. A number of

studies using culture-independent techniques showed the microbial diversity of more than 100 operation taxonomic units (OTUs) within *T. swinhoei*, mainly dominated by the phyla Actinobacteria, Acidobacteria, Chloroflexi, Cyanobacteria, Poribacteria, and Proteobacteria.[98,99] The filamentous heterotrophic bacteria "*Candidatus* Enthotheonella spp." that belong to a newly proposed candidate phylum Tectomicrobia were found widely distributed in *Theonella* sp. living in geographically distant locations.[100–102]

Marine sponges have been widely known as a prolific source of a diverse array of structurally complex compounds with potent biological activities and unique properties. Up to date there are more than four thousand natural products were reported from marine sponges, suggesting that they are prolific producers of biologically active compounds.[103] The first marine sponges-derived that were approved as antiviral and chemotherapeutic agents are vidarabine (Ara-A) and cytarabine (Ara-C) (Figure 2.2). Both compounds are derivatives from spongouridin and spongothymidine isolated from sponge *Tectitethya crypta*.[25] The discovery of these compounds have marked the era of the exploration of marine natural products (MNPs). Several marine sponges-derived natural products have been developed and subjected to clinical trials. One of the most notable examples is eribulin mesylate (Figure 2.2), a truncated version of halichondrin B reported from marine sponge *Halicondria okadaii*.[26] This compound is medically used to treat metastatic breast cancer[28] and liposarcoma.[29]

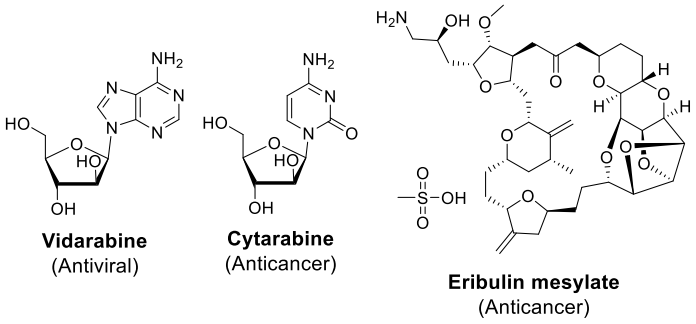


Figure 2.2 FDA-approved drugs from marine sponge

A number of recent studies by Piel and colleagues[61,100,104,105] showed that many bioactive secondary metabolites reported from *T. swinhoei* (Figure 2.3) are biosynthetically made by polyketides synthases (PKSs), non-ribosomal peptide synthases (NRPSs), PKS/NRPS hybrid systems. The encoding biosynthetic genes for a metabolite are found to be clustered inside the genome of the bacterial producer. Interestingly, by utilizing genomic and metabolomic analysis Wilson *et. al*. reported that many natural products previously isolated from *T. swinhoei* Y chemotype such as polytheonamide,[106] nazumamide,[107] konbamide,[108]

cyclotheonamide,[109] keramamide,[110] and onnamide[111] are produced by *Candidatus* Entotheonella factor TSY1.[100] Moreover, using similar approach, Mori *et. al.* also successfully validated that misakinolide[112] and theonellamide[113] from *T. swinhoei* WA chemotype are in fact produced by *Ca.* Entotheonella serta TSW1.[61] Further genomic analysis of identified Entotheonella strains revealed a large amount of biosynthetic gene cluster (BGCs) that encoding for potentially novel natural products. These findings demonstrated that *Ca.* Entotheonella is one of the promising yet untapped bacterial resources. Unfortunately, a way to explore natural products directly from *Ca.* Entotheonella spp. has not been established since the cultivation of these potential symbionts remains elusive.

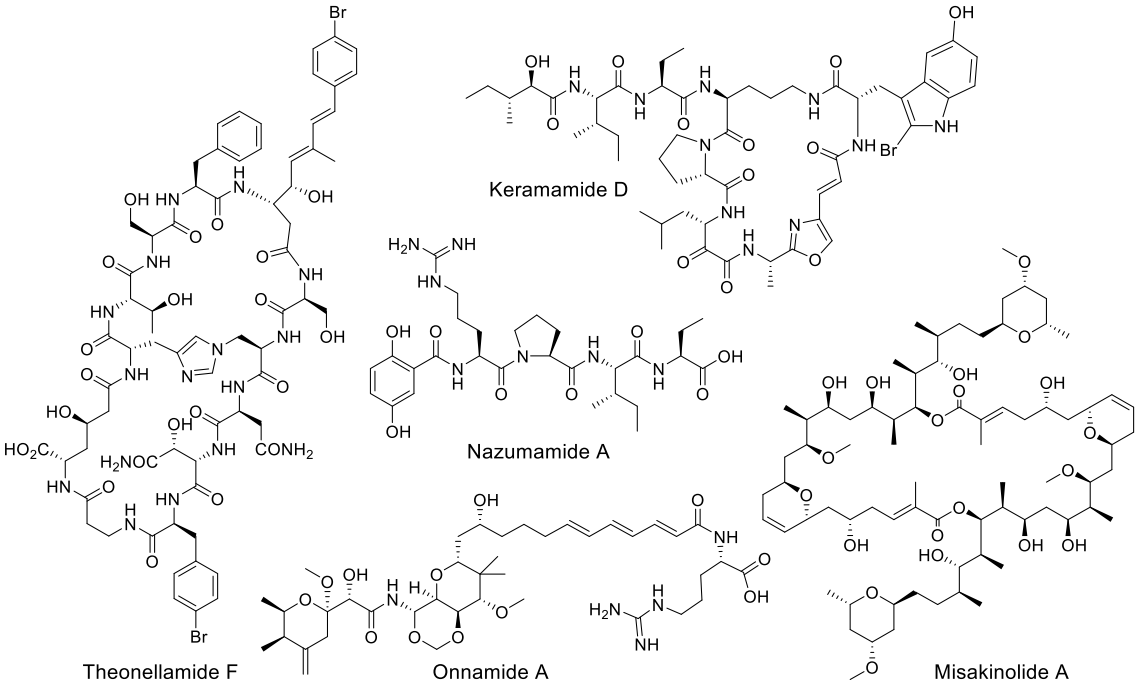
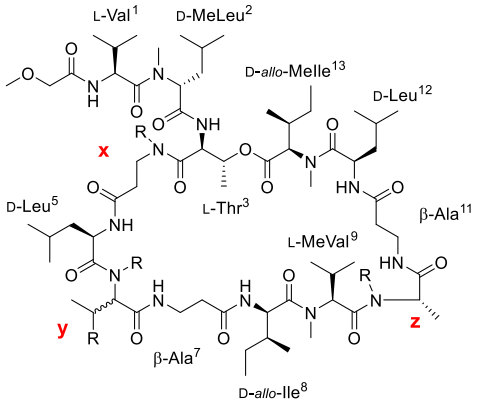


Figure 2.3 Representative natural products reported from *T. swinhoei*

Therefore, to access those potentially novel compounds, in this study, we analyzed and isolated several new members of theonellapeptolide family from marine sponge *T. swinhoei*, the host organisms of promising uncultured symbiotic bacteria *Ca*. Entotheonella spp. Here, the isolation, characterization, and investigation of biological activity of the newly isolated compounds were described. In addition, based on the biosynthetic logic in conjunction with genetic evidence, we proposed *Ca*. Entotheonella sp. as the producer of theonellapeptolides (Figure 2.4).



	x	y	z
1	(β-Ala^4)	(D-MeVal^6)	(L-Ala^{10})
2	(β-Ala^4)	(D-Val^6)	(L-Ala^{10})
3	(Me-β-Ala^4)	(L-MeVal^6)	(L-Ala^{10})
4	(β-Ala^4)	(L-MeIle^6)	(L-Ala^{10})
5	(Me-β-Ala^4)	(L-MeIle^6)	(L-Ala^{10})
6	(β-Ala^4)	(L-MeIle^6)	(L-MeAla^{10})

Figure 2.4 New members (**1**-**3**) and known (**4**-**6**) theonellapeptolide-type compounds from *T. swinhoei.*

2.2 Result

2.2.1 Isolation of Natural Products from Marine Sponge *Theonella swinhoei*

The isolation of natural products from marine sponge inhabited by promising uncultured symbiotic bacteria *Ca.* Entotheonella sp. has been done according to the workflow (Figure S2.1). Approximately 1.0 kg (wet weight) of marine sponge *T. swinhoei* was extracted three times with 1 L of methanol and concentrated *in vacuo*. Preliminary obtained crude extract was subject to TLC analysis developed by chloroform : methanol (9:1) as mobile phase. After treatment with Dragendorff's reagent, it showed three brown spots after treatment with Rf values of 0.31, 0.49, and 0.71, respectively (Figure S2.1), suggesting the presence of relatively non-polar nitrogen-containing compounds such as alkaloids and/or peptides. Further LC-MS analysis of the crude extract indicated the presence of several peaks that correspond to new molecular weight based on MarinLit database, a database dedicated to marine natural products research. These findings had motivated us to isolate the natural products from *T. swinhoei* guided by the obtained new molecular weight.

To purify the target compounds, initially, the methanolic crude extract was partitioned between water and ethyl acetate. The ethyl acetate soluble fraction was concentrated and subsequently subjected to a gel-filtration chromatography to separate the compounds based on their molecular weight (range 100-4000 Da). The methanol-swollen beads (Sephadex LH-20) act to trap small molecules into their pores, leaving the high molecular weight compound eluted in the earlier fractions.[114] Out of the 160 obtained fractions, we found that the target compounds were present in the fraction number 62 to 78 (16 fractions) after being verified by TLC with Dragendorff's reagent treatment. The fractions containing target compounds were pooled and

concentrated *in vacuo*. Subsequently, the obtained residue was further subjected to open column chromatography packed with silica gel using a step-gradient of chloroform and methanol as mobile phase. Guided by ESI-MS, the fractions containing target molecular weight were further purified through a reversed-phase HPLC equipped with Cosmosil $5C_{18}$-MS-II column (10 mm I.D. × 250 mm) using 75% aqueous acetonitrile (0.05 % TFA) as mobile phase to afford compounds **1** (4.0 mg), **2** (9.0 mg), **3** (3.0 mg), **4** (40.0 mg) and **5** (8.0 mg). In addition, another major compound from other fractions was successfully purified as **6** (12 mg).

2.2.2 Structure Elucidation of the Isolated Compounds

According to their NMR and tandem mass analyses, it was found that three of the isolated compounds (**1**-**3**) are novel, whereas the other (**4**-**6**) were previously described as theonellapeptolide IId, IIe, and Id, respectively, cyclic depsipeptides from same sponge species (Figure 2.4).[115–117] On the other hand, ^{1}H-NMR spectra of newly isolated compounds measured in chloroform-*d* ($CDCl_3$) showed several broad *N*-methyl protons (δ_H 2.80 to 3.20), amide protons (δ_H 6.90 to 7.80), and methoxy acetyl group protons (δ_H 3.50 to 3.90), strongly indicated that these compounds are new members of theonellapeptolide-type compound. Unfortunately, due to broadening signals in some key regions, it was challenging to determine their planar structures solely based on the NMR spectra. Furthermore, COSY correlation spectra indicated the presence of several conformers within the purified compounds (Figure S2.2). Therefore, alternatively, their amino acid sequences were initially determined by combination of tandem mass (MS2) and MS/MS/MS (MS3) analyses followed by confirmation using NMR analysis.

There are many ionization techniques for peptide fragmentation, including electron ionization (EI), chemical ionization (CI), electrospray ionization (ESI), and matrix-assisted laser desorption/ionization (MALDI). Fragmentation in a protonated peptide occurs through amide bond cleavage. If the charge is retained by the C-terminus fragment, a protonated amino acid (y_1) or peptide (y_n) will be formed. If the charge remains on the N-terminus fragment, a neutral amino acid is eliminated and b_n ions are formed.[118] Accordingly, a series of b-ions allows determination of the N-terminal sequence. Likewise, a series of y-ions enables sequence determination at the C-terminus (Figure 2.5).[119]

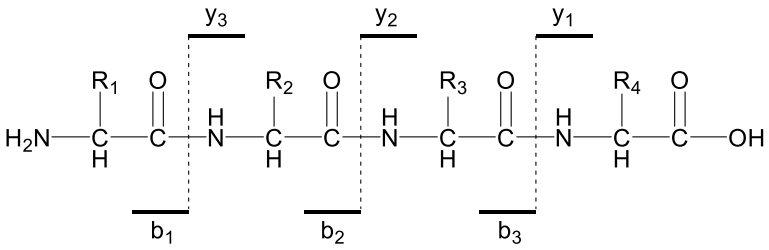


Figure 2.5 Standard nomenclature of fragment ions derived from a peptide.

However, characterization of a cyclic peptide is a great analytical challenge due to no well-defined termini to anchor the assignment of the amino acid sequence. In addition, the ring-opening may occur at several peptide backbone positions, meaning that all the fragment ions are not referenced to a single terminal position. To address this issue, *seco*-acid methyl ester peptides or a linearized form of the purified compounds were generated by reaction with 7M ammonia solution in methanol as shown in Figure 2.6. Determination of amino acid sequences of the newly isolated compounds **1**-**3** has been done by subjecting the obtained ring-opened product to MS2 and MS3 analyses.

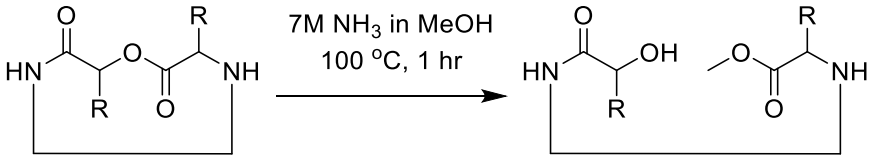


Figure 2.6 Ring opening reaction strategy of a cyclic-depsipeptide compound

2.2.2.a. Determination of planar structure of compound **1**-**2**

Based on HRESIMS data, the molecular formula of compound **1** is $C_{68}H_{121}O_{16}N_{13}$, corresponding to the loss of one methylene group of theonellapeptolide IId (**4**) ($C_{69}H_{123}O_{16}N_{13}$).[116] MS2 analysis of the *seco*-acid methyl ester of **1a** showed three predominant fragment ion peaks in the spectrum (Figure 2.7; S2.3-4). Two high intense fragment ions at *m/z* 299 and *m/z* 584 designated as b_2 and b_5 ions, respectively, which were also detected in the fragmentation of ring-opened peptide (**4a**) of **4** (Figure S2.3), suggesting that the predominant b_2 and b_5 fragment ions were corresponding to the N-terminal moieties, *N*-methoxyacetyldipeptide and *N*-methoxyacetylpentapeptide, respectively. The third high intense fragment ion peaks were different between **1a**, and **4a**, and appeared at *m/z* 825.62, and *m/z* 839.57, which were assignable to the remaining C-terminus sequences of **1a** and **4a**, respectively and the 14 mass unit differences suggested the loss of methyl or methylene group at these units (Figure 2.7). We then implemented the structure confirmation of the individual

fragment ions (b_2, b_5 and y_8) based on the MS3 fragmentation profiles in conjunction with the NMR data of the intact compound.

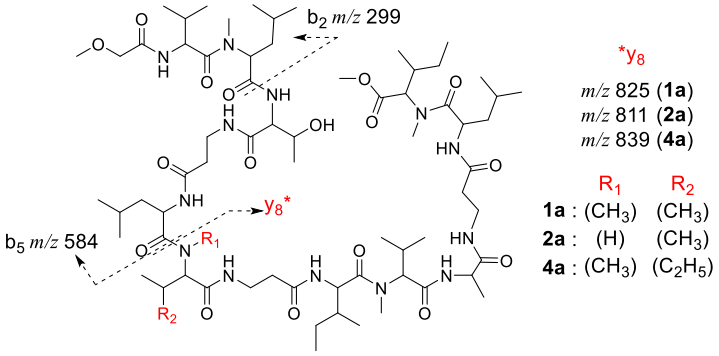


Figure 2.7 MS2 analyses of the *seco*-acid methyl ester peptides of **1**, **2**, and **4**.

To confirm the amino acid sequence of the b_2 ion NMR analysis was carefully conducted as listed in Table S2.11-S2.13 (Figure 2.8). The presence of an ester linkage between the C terminus of the peptide to the hydroxyl group of Thr^3 was confirmed by HMBC correlation showing the cross-peak between the β-CH of Thr^3 and the carbonyl carbon of a *N*-MeIle^{13}. The HMBC spectrum of **1** showed correlation of the Hα of Val^1 (δ_H 4.72) to the amide carbonyl position of the methoxy acetyl (MeO-Ac) unit (δ_C 170.10) (Figure 2.8, Table S2.11), indicating that the N-terminus was capped by MeO-Ac moiety. The presence of a methoxy group on the b_2 fragment ion of **1a** was supported by its typical ^{1}H and ^{13}C chemical shifts (δ_H 3.59 and δ_C 70.20). Furthermore, the Hα of *N*-MeLeu^2 (δ_H 5.08) was connected to the amide carbonyl at δ_C 173.10 assigned to Val^1 in **1**. This was supported by the MS3 profile of the b_2 ion showing the presence of an ion peak at *m/z* 172.07 corresponding with the loss of 127 mass units due to the removal of *N*-MeLeu^2 residue (Figure S2.7). The combined tandem MS and HMBC data established the amino acid sequence of the b_2 ion (*m/z* 299) as MeO-Ac-Val^1-*N*-MeLeu^2.

The MS3 profile of the b_5 fragment ion (Figure S2.6) indicated the presence of smaller fragment ions at *m/z* 471 (corresponding to the loss of Leu^5), at *m/z* 400 (due to the additional removal of β-Ala^4), and at *m/z* 299 (due to the additional loss of Thr^3). This was supported by the 2D NMR spectra of **1** showing correlations from the Hα of Thr^3 (δ_H 4.39) to the carbonyl signal of *N*-MeLeu^2 at δ_C 171.20, and additionally the Hα of β-Ala^4 to the carbonyl of Thr^3 at δ_C 168.20. The presence of a β-Ala at the residue position 4 was confirmed by the COSY and HMBC analysis showing ^{1}H-^{1}H COSY correlation between the NH at δ_H 7.25 and Hα of β-Ala^4 (δ_H 3.47 and δ_H 3.26) and subsequently ^{1}H-^{13}C correlation of the Hα of β-Ala^4 to the

carbonyl of Thr[3] (Figure 2.8). The combined tandem MS data of b_2 and COSY/HMBC analyses suggests that the b_5 fragment ion at *m/z* 584 covers the N-terminal sequence, MeO-Ac-Val[1]-*N*-MeLeu[2]-Thr[3]-β-Ala[4]-Leu[5].

The MS3 profile of the y_8-ions derived from **1a**, and **4a** showed similar fragments generated toward C-termini (from β-Ala[7] to *N*-MeIle[13]) and different by 14 mass unit toward N-termini, implying that the variation among these derivatives occurred at amino acid residue position 6. The fragmentation pattern of the fragment ion y_8 of **1a** in conjunction with its COSY and HMBC data (Figure 2.8; S2.5) established the extended amino acid sequence to C-terminus as *N*-MeVal[6]-β-Ala[7]-Ile[8]-*N*-MeVal[9]-Ala[10]-β-Ala[11]-Leu[12]-*N*-MeIle[13]. The presence of *N*-MeVal[6] adjacent to the Leu at position 5 is supported by ROESY correlations between Hα of Leu[5] and *N*-methyl proton of *N*-MeVal[6], and between Hγ of Leu[5] and Hβ-of *N*-MeVal[6].

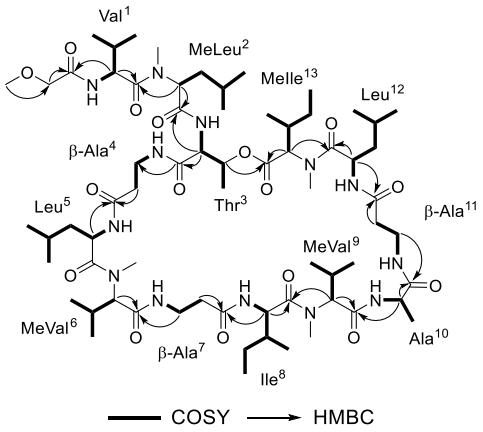


Figure 2.8 Analysis of COSY and HMBC key correlations for compound **1**

The molecular formula of compound **2** was assigned as $C_{67}H_{119}O_{16}N_{13}$ according to the HRESIMS data suggested less one methylene group compared to that of **1**. The ^1H and ^{13}C NMR spectra of **2** showed high similarity to those of **1**, with several overlapped signals of methylated nitrogen (δ_H 2.7-3.3 ppm) and amide protons (δ_H 7.0-7.6 ppm). Unfortunately, we were unable to assign several key correlations in its 2D spectra due to the broadening signals and scarce amount. Tandem mass of a methanolysis product of **2**, named **2a**, showed the presence of three predominant fragment ions including b_2 (*m/z* 299.19), b_5 (*m/z* 584.42), and y_8 (*m/z* 811.63), which are mimicking the fragmentation pattern profile to those of **1a** and **4a** (Figure 2.7; S2.3,S2.8). Interestingly, the fragment ion y_8 of **2a** (*m/z* 811.63) showed difference by 14 mass units to y_8 of **1a** (*m/z* 825.59) indicating that the variation of amino acid residue is located at this fragment ion.

Fragment ion b_5 of **2a** is calculated to cover *N*-methoxyacetylheptapeptide as suggested by MS3 analyses. Fragmentation of b_5 generated several smaller ions at *m/z* 471, 400, 299 indicating loss of Leu^5, β-Ala^4, and Thr^3, respectively (Figure S2.10). In addition, fragment ion b_2 (*m/z* 299.19) is proposed to cover MeO-Ac, Val^1 , and Me-Leu^2 moiety similar to **1a** and **4a**. The presence of MeO-Ac in b_2 is supported by the HMBC analysis of **2** that showed correlations of a methoxy proton at δ_H 3.59 to a methylene carbon at δ_C 71.5 that adjacent to a carbonyl amide at δ_C 170.1. Moreover, correlation of an Hα (δ_H 4.72) of Val^1 to the previous carbonyl amide suggesting that it is located next to MeO-Ac moiety. Based on these data, we proposed the amino acid sequence of **2a** as MeO-Ac-Val^1-MeLeu^2-Thr^3-β-Ala^4-and Leu^5. Furthermore, the MS3 spectrum of y_8 (Figure S2.9, Table S2.4) displayed several fragment ions that implying gradual loss of MeIle^{13} to MeVal^9. Combining the fragment ions generated from MS2 and MS3 analyses of **2a** in addition to several data from complicated NMR spectra of intact compound, allowed us to determine the amino acid sequence of y_8 of **2a** as Val^6-β-Ala^7-Ile^8-*N*-MeVal^9-Ala^{10}-β-Ala^{11}-Leu^{12} and *N*-MeIle^{13} (Figure S2.9). Thus, compound **2** was determined to have amino acid residue at position 6 as Val instead of MeVal in **1** at the same position.

The HRESIMS data showed that compound **3** ($C_{69}H_{123}O_{16}N_{13}$) is less one methylene (CH_2) group from theonellapeptolide IIe **5** ($C_{70}H_{125}O_{16}N_{13}$)[117]. Tandem MS analysis of the *seco*-acid methyl ester peptide of **3**, designated as **3a**, showed three high intense fragment ions, namely b_2 at *m/z* 299.19, b_4 at *m/z* 400.27, and y_{10} at *m/z* 1023.75 (Figure S2.11; Table S2.5). The fragmentation profile of **3a** was apparently similar with that of the ring-opened derivative of known compound **5** (**5a**); b_2 *m/z* 299.18, b_4 *m/z* 400.18, and b_{10} *m/z* 1037.73 (Figure 2.9, S2.17; Table S2.10). The b_4 fragment ions of **3a** and **5a** at *m/z* 400.10 were assigned as MeO-Ac-Val^1-*N*-MeLeu^2-Thr^3 in order from N-terminus (Figure S2.13, S2.19). The y_{10} fragment ion peak *m/z* 1023.7 was subsequently subjected to MS3 analysis (Figure S2.12; Table S2.6), allowing us to assign it as *N*-Me-β-Ala^4-Leu^5-*N*-MeVal^6-β-Ala^7-Ile^8-*N*-MeVal^9-Ala^{10}-β-Ala^{11}-Leu^{12}-*N*-MeIle^{13} in order from the N-terminus. The y_{10} fragment ion of **3a** (*m/z* 1023.73) differs from that of **5a** (*m/z* 1037.74) only in the substitution of *N*-MeVal^6 for *N*-MeIle^6 (Figure 2.9).

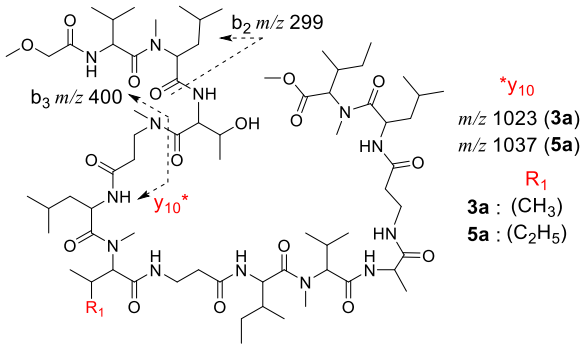


Figure 2.9 MS2 analyses of the *seco*-acid methyl ester peptides of **3** and **5**.

2.2.2.c. Determination of absolute configuration of **1**-**3**

Upon determining the amino acid sequences of the isolated compounds, we deduced the configurations of the amino acid residues by means of Marfey's analysis.[120] Compound **1**-**3** consist of several duplicated amino acids such as methyl valine (duplicated in **1** and **3**), valine (duplicated in **2**), and leucine (duplicated in **1**-**3**). Hence, to unambiguously determine the configuration of each residue, we performed partial hydrolysis of **1**-**3** to obtain a peptide fragment containing one of each duplicated residue (Figure 2.10) followed by Marfey's analysis and compare the results that obtained from the totally hydrolyzed original compound.

Targeted fragment peptide of **1** corresponding to N-terminus, has been purified from the partial hydrolysate of **1** after treatment with TFA 30% for 1 hour at 100 °C and designated as fragment methoxy acetyl hexapeptide **1b** as confirmed by MS2 analysis (Figure S2.38). Subsequent Marfey's analysis of **1b** enabled us to deduce the absolute configuration of the first five residues at the N-terminus as L-Val[1], D-*N*-MeLeu[2], L-Thr[3], D-Leu[5], and D-MeVal[6], respectively. In addition, the Marfey's analysis of the total hydrolysate of **1** (Figure S2.39) exhibited only one residue of D-Leu and enantiomeric mixture of D and L-*N*-MeVal. Comparison of data obtained from **1b** and **1** allowed us to determine all absolute configurations including D-*N*- MeVal[6] and L-*N*- MeVal[9] in **1** (Figure 2.10 B).

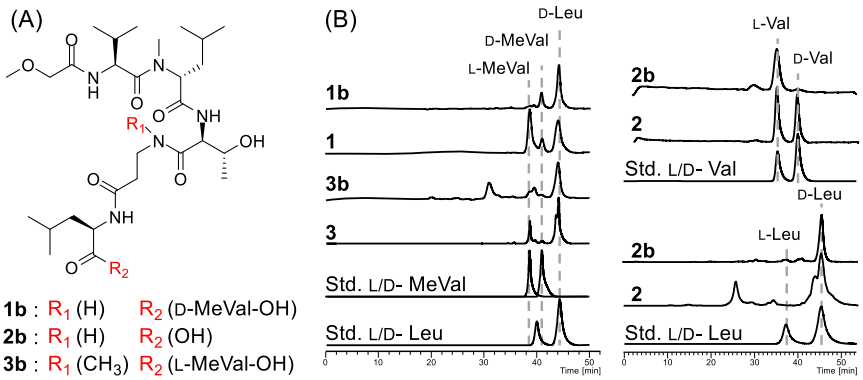


Figure 2.10 (A) Isolated fragment peptide to determine amino acid configuration of repeated residue. (B) Determination of absolute configuration of repeated amino acids.

Using similar strategy, the pentapeptide **2b** (Figure S2.40A) and the intact compound **2** were subjected to Marfey's analysis (Figure S2.40B and S2.41). As the result, the absolute configuration of the N-terminal four amino acid residues has been determined as L-Val[1], D-*N*-MeLeu[2], L-Thr[3], and D-Leu[5]. In addition, the absolute configuration of Val[6] and Leu[12] was deduced to be D-configuration based on the comparison of the obtained Marfey's analysis data of **2b** and **2** (Figure 2.10 B). Thus, using this strategy allowed us to determine all absolute configurations of amino acid residues, including duplicated valine and leucine, in **2**.

Finally, a hexapeptide **3b** was purified from partial hydrolysate of **3** as confirmed by tandem MS analysis (Figure S2.42A). Subjecting **3b** to Marfey's analysis enabled determination of the absolute configuration of its amino acid residues as L-Val[1], D-*N*-MeLeu[2], L-Thr[3], and D-Leu[5] (Figure S2.42B). It was difficult to determine the absolute configuration of *N*-MeVal in this step due to a low peak-resolution generated from small amount of sample (Figure 2.10 B). However, Marfey's analysis of the total hydrolysate of **3** (Figure S2.43) showed only one residue of D-Leu and one of L-*N*-MeVal, suggesting D-configurations of both Leu residues and L-configurations of both *N*-MeVal residues (Figure 2.10 B). Therefore, the absolute configuration of amino acid residues in 3 is determined as L-Val[1], D-*N*-MeLeu[2], L-Thr[3], D-Leu[5], L-MeVal[6], D-*allo*-Ile[8], L-MeVal[9], L-MeAla[10], D-Leu[12], D- *allo*-MeVal[13] (Figure 2.11).

In summary, based on the Marfey's analysis results described above, we assigned the identity and absolute configuration of the amino acid at the residue position 6 as D-MeVal in **1**, D-Val in **2**, and L-MeVal in **3**. These data placed **1** and **2** as the first members of theonellapeptolide family to have a valine residue with D configuration at the position 6. We therefore propose **1** and **2** as new theonellapeptolide analogues, designated here as

theonellapeptolides IIb and IIa. Furthermore, the presence of a L-*N*-MeVal at the residue position 6 and a *N*-Me-β-alanine at the residue position 4 in **3** places this compound as a new analogue designated here as theonellapeptolide IIc (Figure 2.11).

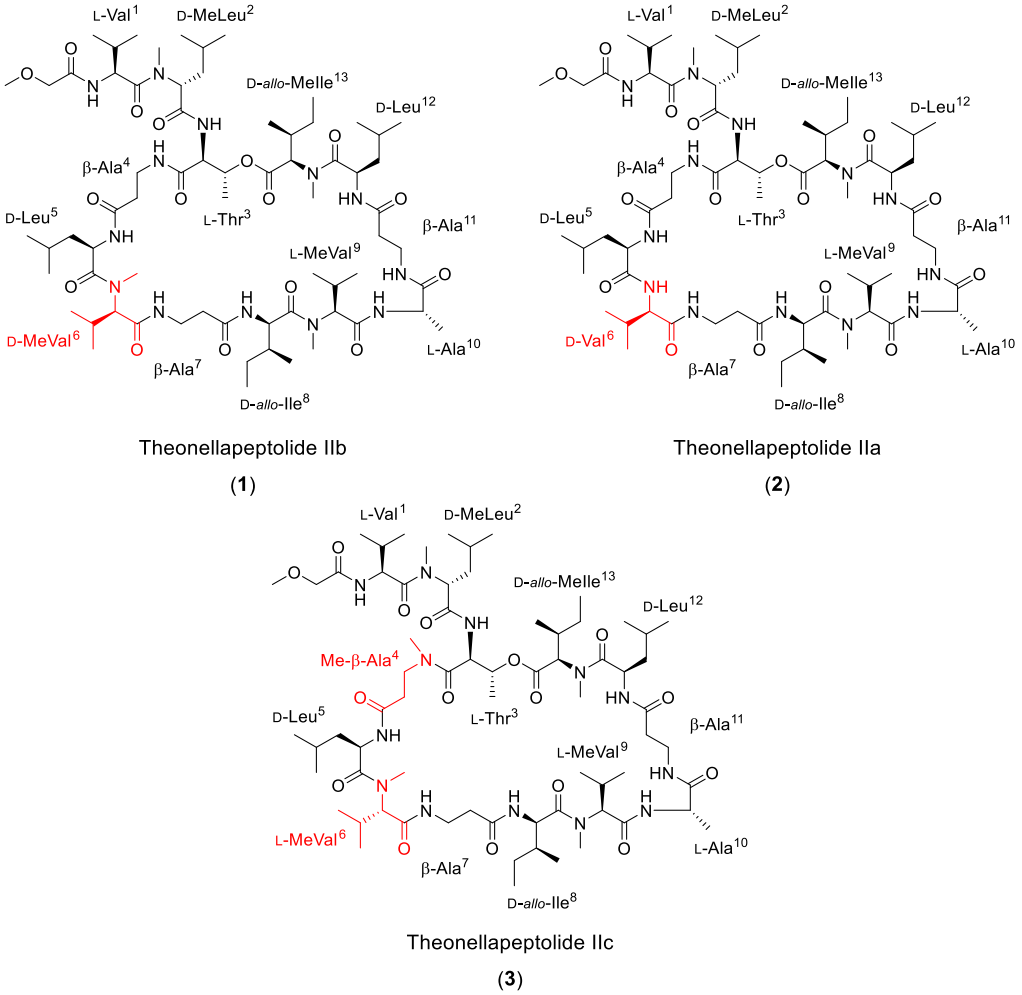


Figure 2.11 Newly isolated and characterized members of theonellapeptolide family (**1**-**3**). The modified amino acid residues are highlighted in red.

2.2.3 Biological Activity of the Isolated Compounds

There are many interesting biological properties that frequently associated with cyclic peptides due to their conformational rigidity.[121] In addition, the presence of unique amino acid residues, methylated nitrogen, and acyl group modifications at N or C terminal have been proven to give superior biological activity. For instance, cyclosporin A is a cyclic peptide that has been therapeutically used as an immunosuppressant agent for organ transplantation[122] and rheumatoid arthritis.[123] This compound is consisted of eleven amino acid residues, two of them are unique residues (butenyl-methyl-threonine and L-alpha-aminobutyric acid), a single D-amino acid residue (D-alanine), and intriguingly, seven of its peptide bonds are methylated.[124]

Detailed analysis of its structure activity relationship revealed the roles of each residue as well as the *N*-methylation to the activity.[125] Several cyclic peptides with similar characteristics also exhibited potent biological activities such as nematocidal activity from omphalotin,[126] cytotoxic activity from odoamide,[127] and antibacterial activity from cycloheptadepsipetide.[128]

Several biological activities associated with theonellapeptolide-type compounds were previously investigated. The first identified theonellapeptolide I series, including Ia, Ib, Ic, Id (**6**), and Ie, have been reported to inhibit development of the fertilized egg of sea urchin *Hemicentrotus pulcherrimus* at concentration 2, 2, 2, 50, and 10 μg/ml, respectively,[115] suggesting anti-mitotic property of these compounds.[129] In addition, theonellapeptolide Ib, Ic, Id (**6**), and Ie also showed moderate cytotoxic activity against mouse lymphocytic leukemia cell, L1210, with IC_{50}; 1.6, 1.3, 2.4, and 1.4 μg/ml, respectively.[130] Antibacterial and moderate cytotoxic activities have been observed when a methylsulfinyl acetyl group blocked their N-termini as in sulfinyltheonellapeptolide Ia and sulfinyltheonellapeptolide Id.[131,132] Among the members of theonellapeptolide I series, Id (**6**) exhibited ionophoretic activity for transport of Na^+, K^+, and Ca^{2+}, while Ie preferentially showed transport activity of Na^+ and K^+. Interestingly, a *seco*-acid methyl ester peptide of Id did not show ionophoretic activity against all tested metals indicating that the macrolactone structure is essential to exert the activity.[133] Moreover, a demethylated nitrogen at residue 10 (Ala^{10}) version of Id, named theonellapeptolide IId (**4**), showed a very weak ionophoretic activity.[116] In contrast, the immunomodulatory activity of IId (**4**) displayed significantly more potent compared to Id (**6**).[117] It is a remarkable structure activity relationship of these compounds that with small variation at *N*-methylation affected their biological and physical properties. Finally, a group of theonellapeptolide III series reported from distinct marine sponge showed modest cytotoxic activity against leukemia cancer cell line P388.[134]

The newly isolated compounds **1**-**3** and known **4**-**6** were subjected to antibacterial and cytotoxic activity. The antibacterial activity of **1**-**6** has been investigated against gram-negative *Escherichia coli* JW5503 and gram-positive *Bacillus cereus* NBRC 15305 and *Kocuria* rhizophila NBRC 12708. Unfortunately, none of the tested compounds showed any antibacterial activity even in the highest given concentration (66 μg/mL) (Table S2.14). In addition, three cancer cell lines were used to examine the antiproliferative activity including human pancreatic cancer (MIA PaCa-2), breast cancer (MCF-7), and liver cancer (HepG2). Compounds **1**-**6** exhibited moderate to no antiproliferative activity (Table S2.14).

In addition, all compounds were also tested for their anti-austerity property. Principally, this assay targets the ability of certain cancer line to survive under nutrient deficient environment.[135] Generally, tumor cells utilize angiogenesis to accommodate required oxygen, glucose, nutrients, and growth factors for progression.[136] However, pancreatic cancer has widely known to be a hypovascular tumor meaning that it is incapable to obtain sufficient nutrients to grow.[135] Consequently, these cancer cells alter their metabolic system to acquire tolerance of nutrient starvation to survive and promote progression. Pancreatic cancer cells demonstrate a remarkable tolerance even in severe nutrient deprivation, leading to difficult eradication of these tumor cells.[135] Therefore, discovery of chemical agents to eliminate the cancer cells' tolerance in nutrient starvation becomes an important strategy in anti-cancer drug development.[135,137]

The study of underlying mechanisms of the survival and proliferation ability of pancreatic cancer cells is still ongoing. Nevertheless, the anti-austerity approach to identify potent agents has been established.[137] Briefly, the cancer cells were grown two different media, (i) one in standard cultivation medium (DMEM, Dulbecco's modified Eagles's Medium), (ii) and another one in a nutrient-deprived medium (NDM). Subsequently, both culture cells were treated by several dilution of interested agents followed incubation and determination of survived cells. The compounds that showed activity only in NDM, presented in preferential cytotoxicity (PC_{50}), are considered to be anti-austerity agents. Following this strategy, several efforts have been made including using chemical synthesis approach [138–142], screening of natural products from plant [143–147], and compounds from microorganism [137,148].

The anti-austerity property of isolated compounds **1**-**6** has been investigated against human pancreatic cancer line (MIA PaCa-2), breast cancer line (MCF-7), and liver cancer line (HepG2). While the tested compounds were inactive against breast and liver cancer line, they exhibited moderate to strong activity toward pancreatic cancer cell line MIA PaCa-2 (Table S2.14; Figure 2.12). Among them compound **5** displayed the most potent activity with PC_{50} : 3.50 μM in NDM while its IC_{50} in DMEM is 40.04 μM. Strong anti-austerity activity also demonstrated by compounds **3**, **4** and **6** with PC_{50} : 10.0, 7.8, and 8.2 μM, respectively, whereas moderate activity was observed for compounds **1** and **2** (Figure 2.12). The structure-activity relationship of the tested compounds revealed that the combination of methyl β-alanine and methyl isoleucine at the positions 4 and 6 of **6** increased the activity (PC_{50} 3.5 μM). In contrast, the absence of both residues, as in **1** and **2**, resulted in moderate activity (PC_{50} 33.4 and 42.8 μM) in MIA PaCa-2 cells. Furthermore, in the presence of one of them, the activity

of **3**–**5** increased (PC_{50} 10.0, 7.8 and 8.2 μM, respectively) but was lower than that of **6** (Table S2.14; Figure 2.12). These data suggested that the presence of methyl β-alanine-4 and/or methyl isoleucine-6 is important for the activity of theonellapeptolide type compounds as anti-cancer candidate under nutrient starvation conditions.

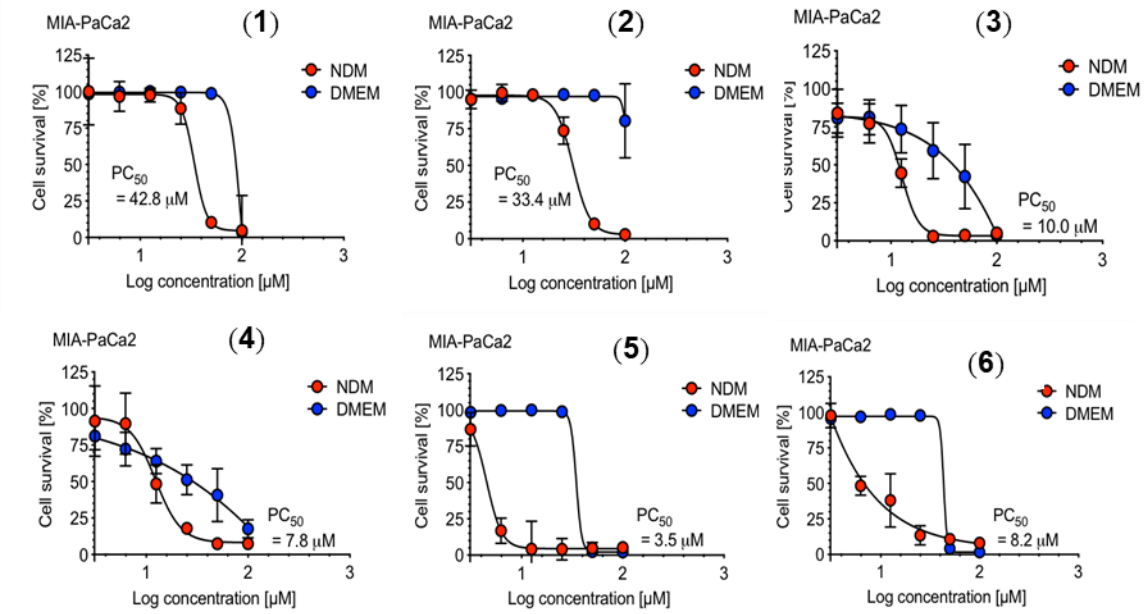


Figure 2.12 Cytotoxic and anti-austerity activity of the new compounds against pancreatic cancer cell line MIA PaCa2.

The anti-austerity property of theonellapeptolide-type of compounds was further examined in a parallel real-time time-lapse microscopy. The purpose of the experiment was to obtain live evidence of the effect of the tested compound against MIA PaCa-2 in nutrient starvation. Compound **4** was selected for the experiment considering the availability of the pure sample. The cultivated cells in NDM were treated with 25 μM of **4** or untreated followed incubation in an CO_2 incubator equipped with a real-time microscopy system. The cell imaging system was set to capture a frame every 10-minutes in three different locations of a plate for 24 hours. As the result, compound **4** induced cell death within 4 hours and a complete eradication after 8 hours (Figure 2.13). On the other hand, untreated cells cultivated under same condition in NDM were able to survive for 24 hours. This result demonstrated the live evidence of the cytotoxic activity of theonellapeptolide-type compounds against MIA PaCa-2 under nutrient starvation.

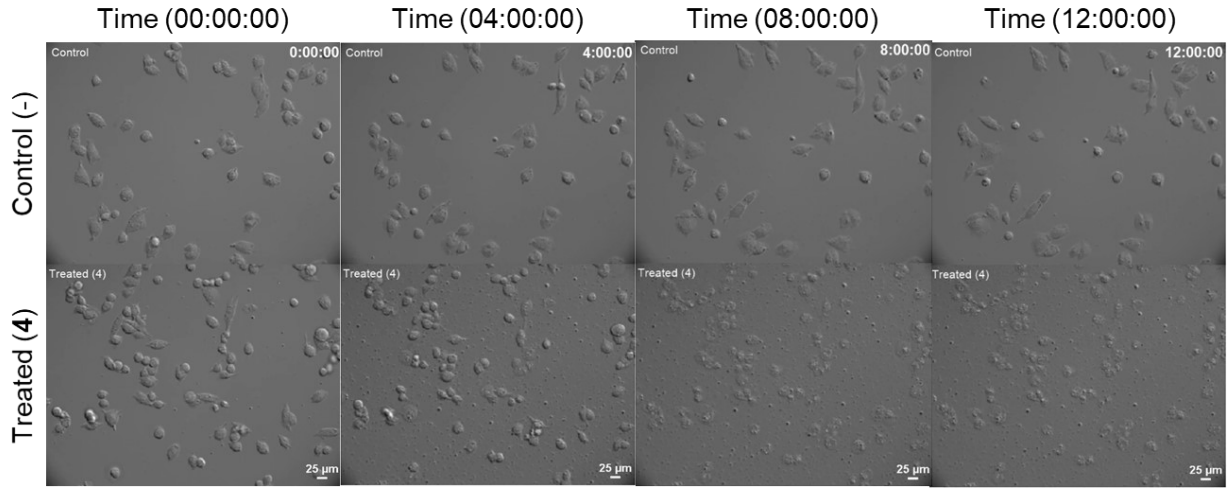


Figure 2.13 Real-time time-lapse microscopy showing antiproliferative effect of **4** (below) compared to control negative (up).

2.2.4 Biosynthetic Pathway of Theonellapeptolides

The newly isolated compounds **1**-**3** and other members of theonellapeptolide-type compounds are cyclic tridecapeptide.[115–117,130–132] Interestingly, they show some unique structural characteristics in their moieties such as β-amino acids, D-amino acids, and *N*-methylated amino acid residues. In addition, their N-terminus is capped by some variations of acyl groups such as methoxy acetyl,[115,116,130] methylsulfinyl acetyl,[131,132] and acetyl moieties.[132] Moreover, modifications among the isolated members generally are occurred at amino acid moieties. Based on these observations, we postulated a hybrid nonribosomal peptide synthase (NRPS) and polyketide synthase (PKS) system as the responsible pathway to produce theonellapeptolides. Other compounds which possess similar characteristics have been previously elucidated as a product of NRPS system as in cyclosporin,[149] or NRPS/PKS hybrid system as in chondramide.[150] Intriguingly, biosynthetic pathway of a series of cyclic peptides, omphalotins, possess a highly *N*-methylated amino acid residues from a basidiomycete *Omphalotus olearius*[151,152] has been reported to be produced by ribosomally-synthesized and post-translationally modified peptides (RiPPs).[153] Furthermore, a RiPPs system was also reported to be responsible for production of polytheonamides,[154] a group of peptides that are comprised of many D-amino acids and their N-terminus capped by an acetyl group.[106] However, up to date, there is no report that cyclic peptides haboring β-amino acid specifically β-alanine come from RiPPs biosynthetic machinery. Therefore, based on these considerations, in this study, we postulated that the biosynthetic pathway of theonellapeptolides is a hybrid NRPS/PKS system.

Following our postulated pathway (Figure 2.14), the first module in theonellapeptolide's biosynthesis is comprised of an acyltransferase (AT) and acyl carrier protein (ACP) domains. This AT domain is proposed to select glycolyl-CoA for the first building block and transferred it to the adjacent ACP domain.[3] Alternatively, it also selects acetyl-CoA found in one congener of theonellapeptolide Ie.[132] The methylation of the glycolic moiety generally found in theonellapeptolides can be occurred after loading the substrate into ACP domain, or prior to loading step as an unique substrate. The second to the final modules are consisted of domain organization following NRPS systems such condensation domain (C), adenylation domain (A), methyltransferase domain (MT) peptide carrier protein (PCP) and epimerization domain (E). The A domains at each module are responsible to recognize and adenylate an amino acid residue and transferred it to their adjacent PCP domain.[155] Interestingly, some of the A domains in this system (A-6; A-8; A13) are able to accept adenylate substrates with similar characteristics such as branched amino acid valine, leucine, and isoleucine resulting in variation of isolated final products.

One of the unique characteristics of theonellapeptolide is the presence of several *N*-methylated amino acid residue. Therefore, to accommodate the methylation, MT domains are postulated occur in several modules including module 3, 5, 7, 10 and 11. These MT domains are of interest since several of the reported members[115–117,130,132] showed *N*-methylated and demethylated at the same position suggesting unique methylation behavior. Further study on these genes will be required to elucidate the factors and conditions that effect this unique behavior. Another accessory gene that postulated to have role in the theonellapeptolide's biosynthesis is epimerization or E domain. This domain is responsible to epimerize a loaded L-amino acid in PCP domain prior to incorporation with adjacent substrate.[155] Once all the substrates have already be loaded and modified in their corresponding PCP domains, C domains incorporate each of the substrates into their adjacent substrates in a stepwise manner from module 1 to the last module.[155] Finally, the completed linear peptide in the PCP domain at last module will be hydrolyzed and further cyclized by a TE domain that commonly presence at the C terminus of the last module in NRPS system.[155]

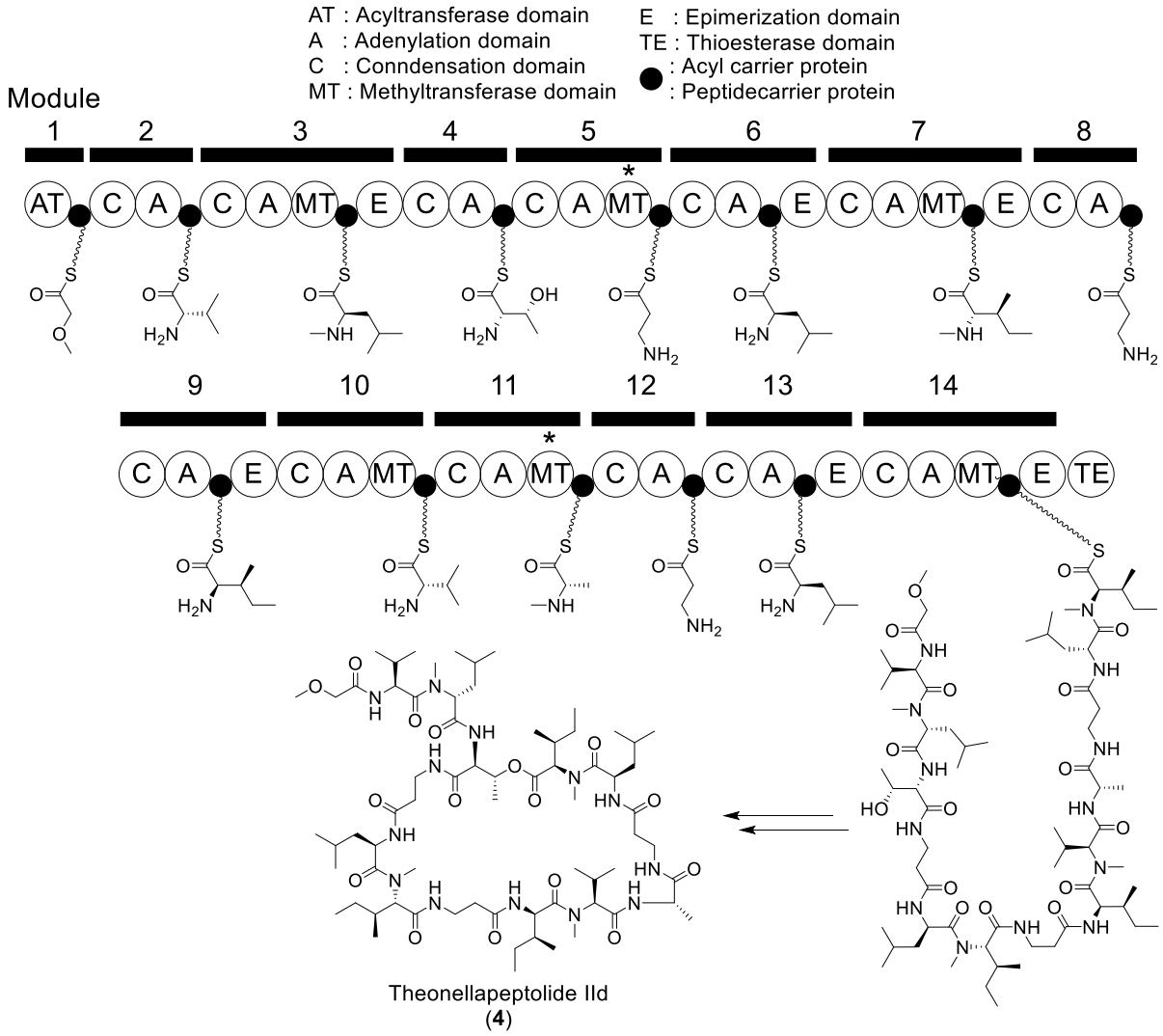


Figure 2.14 Postulated biosynthetic pathway of theonellapeptolides. Unique on-off domains are noted in asterisk.

The postulated biosynthesis of theonellapeptolides suggested three A domains that select specific β-alanine as their substrate. The presence of three β-alanine residues is unique to the all reported theonellapeptolide-type compounds.[115–117,130–132,134] Therefore, according to these considerations, subsequently, we try to identify the producer of these compounds that inhabited marine sponge *Theonella swinhoei.* Recent genomic analysis of Entotheonella sp. from *T. swinhoei* W chemotype termed as *Ca*. Entotheonella serta TSW1 indicated *Ca*. Entotheonella serta TSW1 harbored hundreds of secondary metabolite related genes including a large number of A domains.[61] Among them, there are three A domains are predicted to have substrate specificity for β-alanine. Anticipated that these genes might involve in the biosynthesis of theonellapeptolides, we tried to detect them in Entotheonella cells prepared from our *T. swinhoei* specimens. As a result, we successfully confirmed the presence of both

genes in our samples by PCR detection (Figure 2.15) and subsequent DNA sequencing (Figure S2.44, S2.45), suggesting the Entotheonella as the producer of theonellapeptolides.

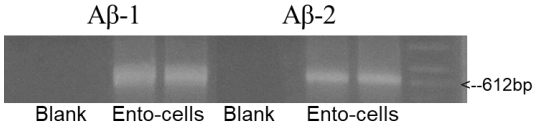


Figure 2.15 Detection of A domains for β-alanine in Entotheonella cells from *T. swinhoei* containing theonellapeptolides.

2.3 Conclusion

We have identified three new members (**1**-**3**) of theonellapeptolide family from marine sponge *Theonella swinhoei* inhabited by a promising yet uncultured symbiotic bacteria *Ca.* Entotheonella spp. In addition, three known compounds (**4**-**6**)were also isolated through a series of chromatographic techniques. All the isolated compounds showed cytotoxic activity against pancreatic cancer cell line MIA PaCa-2 in a nutrient starvation condition, suggesting anti-austerity property as a new biological activity associated to theonellapeptolide-type compounds. Among them, compound **6** possessing combination of methyl β-alanine[4] and L-methyl isoleucine[6] displayed the most potent activity. In addition, we postulated the biosynthetic pathway of theonellapeptolides according to the hybrid polyketide synthase (PKS) and non-ribosomal peptide synthase (NRPS) systems. Furthermore, by considering the postulated biosynthesis pathway, we suggested *Ca*. Entotheonella sp. as the actual producer of theonellapeptolides. Based on our findings in this study demonstrated the potential of untapped symbiotic bacteria from unique host organisms such as marine sponge.

2.4 AuthorContributions

This sub-project has been done in the support by collaborations with several researchers and their contributions are as follows : Jabal R. Haedar, Ismail, Subehan, Toshiyuki Wakimoto managed marine sponge collections. Jabal R. Haedar, Ayumu Enomoto, Toshiyuki Wakimoto contributed to the experimental design. Jabal R. Haedar performed isolation and characterization of compounds. Jabal R. Haedar, Agustinus R. Uria, Toshiyuki Wakimoto designed biosynthetic and metagenomic experiment. Dya Fita Dibwe performed cytotoxic and anti-austerity assay of the isolated compounds.

Chapter 3
Natural Products from Ichip-domesticated Bacteria

3.1 Introduction

The majority of soil biomass is consisted of microorganisms. They have a critical role in many geochemical cycles, including carbon and nitrogen cycles.[156,157] Additionally, they also form a symbiotic relationship with plant by releasing and metabolizing plant-growth hormones.[158] Inversely, plant provides space and nutrient such as organic acid, amino acids, sugar in the form of exudate.[159,160] To obtain these resources, microorganisms engage in antagonistic interactions with their neighbors. The interaction always involves production of toxin that can impair or kill the competitor.[161] Such interaction had inspired many scientists to discover numerous of bioactive small molecules from soil bacteria. Up until today, more than a half of microbial-derived approved drugs are reported from a large family of soil-dwelling bacteria. Several notable bioactive small molecules that widely used for medical purposes including streptomycin (antibacterial),[14] amphotericin B (antifungal),[76] daunorubicin (anti-cancer),[162] ivermectin (antiparasitic),[86] and rapamycin (immunosuppressant)[163] (Figure 3.1).

In the last few decades, explorations of microbial natural products largely focused on family of Actinobacteria, particularly the most prolific genus *Streptomyces*. However, over exploitations toward these genera had led to re-discovery of many already reported compounds. Alternatively, other potential of soil-derived microbial producers from diverse phyla such as those belong to Firmicutes (e.g. *Bacillus*)[164,165] and Proteobacteria (e.g. *Myxobacteria*,[166] and *Burkholderia*[167]) have been extensively studied. On the other hand, the advancement of sequencing technology nowadays, have shed light on the greater potency of soil bacteria.[90,91] Generally, sequencing of entire genetic materials recovered from environmental samples or also known as metagenomic study is one of the powerful approaches to describe the microbial diversity[168] and discover genes that probably involve in production of bioactive small molecules.[169] Interestingly, combinations of metagenomic and meta-transcriptomic studies by Crits-Christoph *et al* in 2018 revealed enormous of potential biosynthetic gene clusters (BGCs) encoded for small molecules associated to a rare group of bacteria that have never been explored.[92] According to their meta-transcriptomic analysis, these BGCs are actively expressed in the soil biomass.[92] Thus, this study indicated that there are many potential rare bacteria living in soil ecosystem that cannot be recovered using a standard cultivation method.

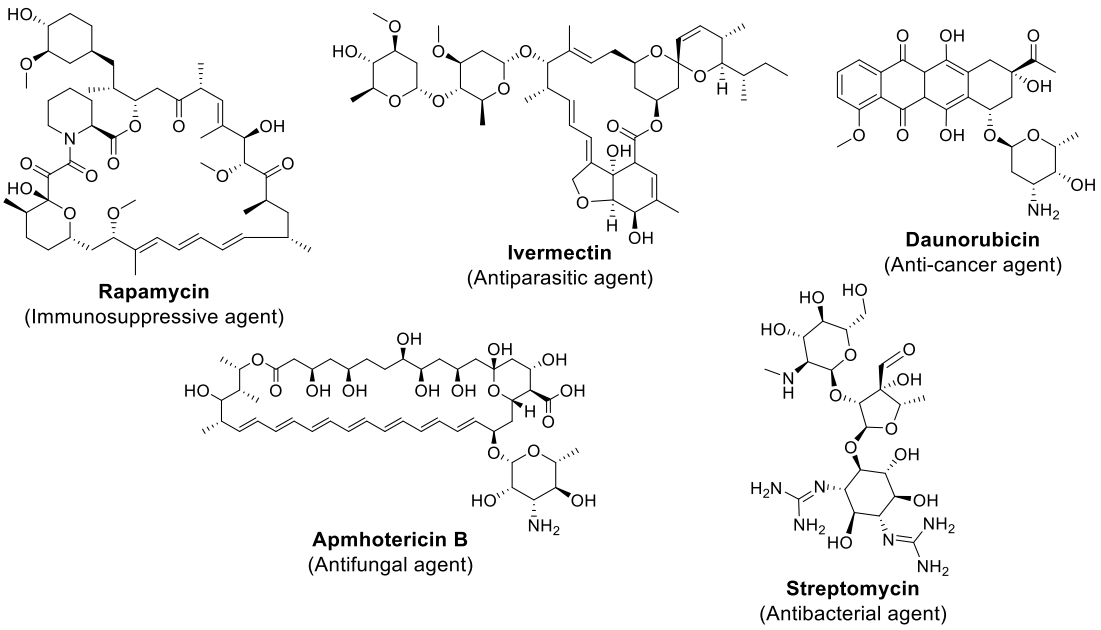


Figure 3.1. Some selected compounds isolated from soil-dwellings bacteria that already approved for medical purposes.

A lot of strategies have been devised to improve the accessibility of microorganisms in various environmental sources including soil ecosystem.[170–173] Recently an *in situ* cultivation approach was launched by considering that microorganisms rely on complex biotic and abiotic factors in their habitats to support their growth.[174–176] In this cultivation method, microbial cells from environment are placed in a diffusion chamber that covered by a pair of membrane filters prior to returning them at the original habitats (Figure 3.2).[174,176] *In situ* cultivation allows microorganism to experience and receive the essential factors including micronutrients and growth factors from neighboring organisms.[175] Additionally, inside the diffusion chamber, the microbial cells are trapped in a standard agar, allowing domestication process which is essential for downstream application and analysis.[175] Application of *in situ* cultivation using diffusion chamber shows superiority compared to the traditional method in phyla distribution and the number of novel isolated bacterial strains.[174]

In 2015, a biotechnology-based company called NovoBiotic lead by Ling and colleagues had successfully isolated a previously unculturable bacterial strain named *Eleftheria terrae* by employing ichip, a small and high throughput version of diffusion chamber.[177] Metabolite investigation of this rare strain led to discovery of teixobactin (Figure 3.2), a new class of antibiotic that effectively kill multi-drug resistance gram-positive bacterial pathogens.[177] Additionally, the application ichip *in situ* cultivation on a marine sponge, a unique habitat and widely known as a prolific source of natural products, had permitted the

purification of a new *N*-acyl tyrosine compound from a putatively new bacterial species, *Alteromonas* sp. RKMC-009 (Figure 3.2).[178] This approach also successfully explained that siderophore is one of the largely unknown growth factors that support the survival of previously unculturable bacteria (Figure 3.2).[179]

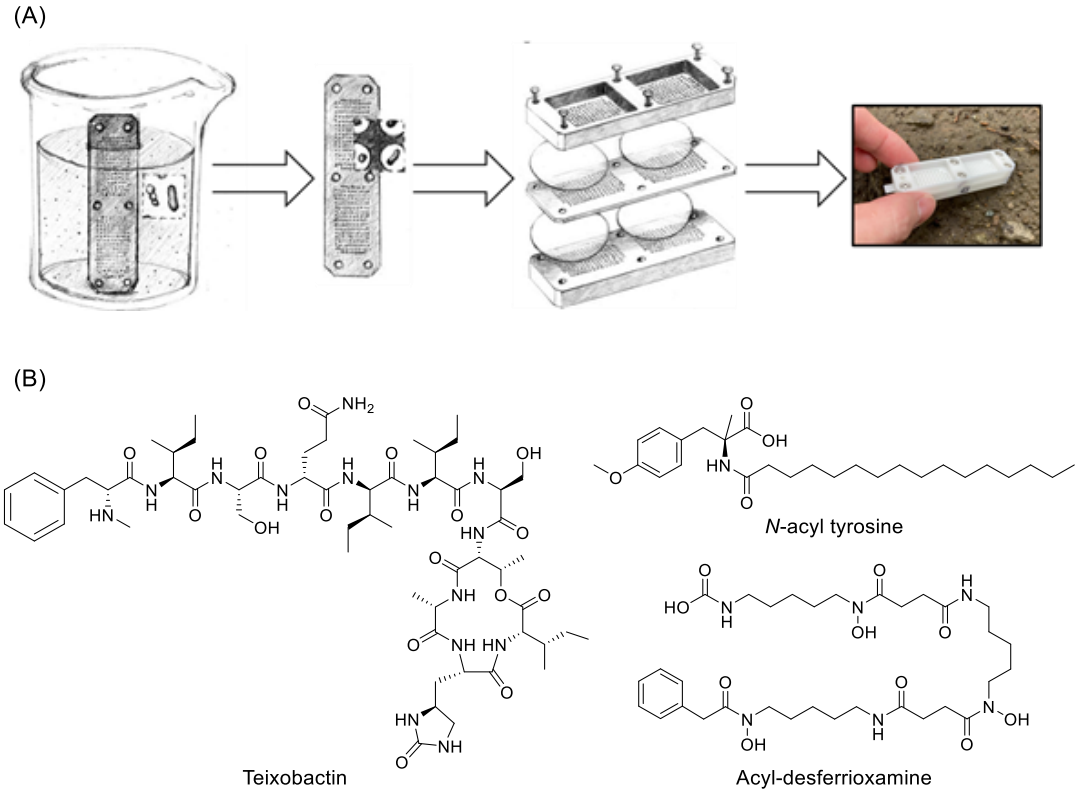


Figure 3.2. (A) The sequel process of ichip-based *in situ* cultivation approach, showing loading of microbial cells into the array until returning into the original habitat. (Adopted from *Nichols et al*. 2010). (B) Reported compounds from strains that isolated using ichip-based in situ cultivation approach.

Therefore, in this study, by employing ichip devices for *in situ* cultivation at several ecosystems, we aim to recover as many as possible bacterial strains and investigate their potential for production of novel natural products.

3.2 Result

3.2.1 Utilization of Ichip to Domesticate Environmental Bacteria.

The ichip-based cultivation approach has been applied at several habitats including soil, mud, and river stream around the Sapporo campus of Hokkaido University. In total, sixteen devices were incubated in the environments for two weeks to a month. Upon incubation, the incubated devices were brought to laboratory for bacterial retrieval and domestication.

Successful domesticated bacteria in standard laboratory cultivation method were identified by sequencing analysis of 16s rRNA, a gate keeper gene of prokaryotes widely used for phylogenetics.[180] As a result, 1046 bacterial strains were successfully domesticated and identified in this study (Table S3.1). Phyla distribution analysis revealed that around 75% of domesticated bacteria are belong to Proteobacteria (Figure 3.3). Three of six classes within this phylum (α-proteobacteria, β-proteobacteria, and γ-proteobacteria) were dominantly found in our samples. Furthermore, novelty analysis of the domesticated bacterial strains indicated as much as 25 of domesticated bacterial strains are at least new bacterial species according to their similarity to the closest strains (Figure 3.3).[174]

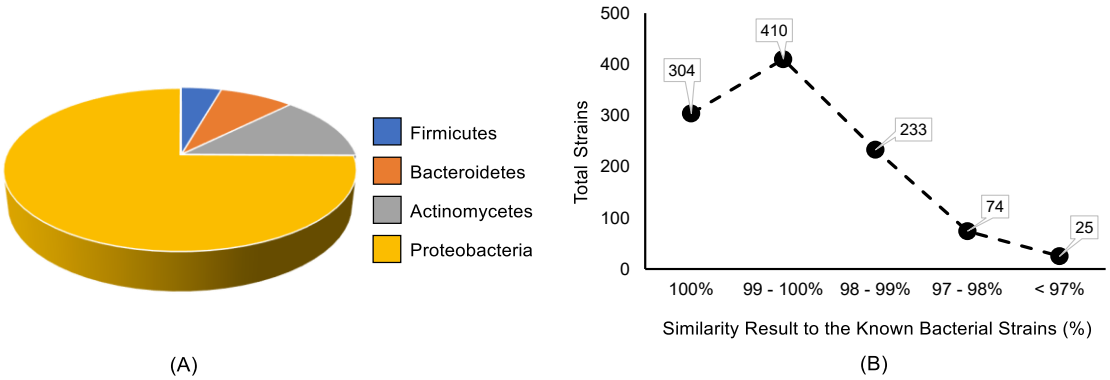


Figure 3.3. (A) Phyla distribution and (B) Novelty analysis of ichip-domesticated bacterial strains.

The successfully domesticated bacterial strains were subsequently subjected to antibacterial screening against gram-negative bacteria (*Escherichia coli*) and gram-positive bacteria (*Bacillus cereus*). Briefly, the seed-culture of domesticated bacterial strains were inoculated to rich nutrient medium (LB) or a defined nutrients medium (M9 minimal medium) in a 96-well polypropylene DeepWell plate and incubated with shake for seven days. Subsequently, the cultures were spun down, and the obtained supernatants were transferred to a new 96-microplate for antibacterial assay. After an overnight incubation, the growth of the tested bacteria was examined using a plate reader for UV absorbance at 600 nm (Figure 3.4). The strains that exhibited growth inhibition ≥ 50% of the tested bacteria were analyzed for their reproducibility. Interestingly, a bacterial strain identified as *Variovorax* sp. H002 showed growth inhibition against *Escherichia coli* exclusively when it was cultivated with M9 minimal media but inactive with rich nutrients LB media. Some other strains also showed similar results, however due to lack of reproducibility, we selected *Variovorax* sp. H002 to explore its potential for production of novel natural products.

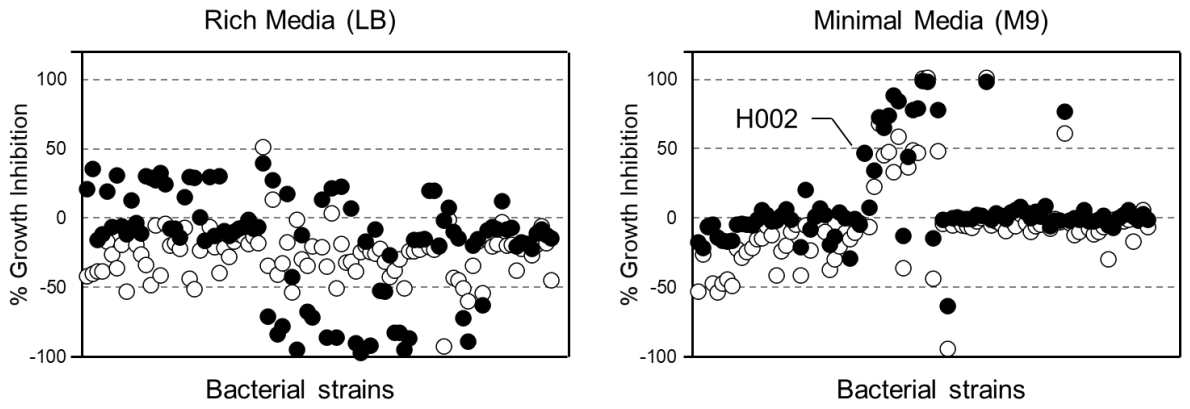


Figure 3.4. Antibacterial screening of ichip-domesticated bacterial strains cultivated under Luria-Bertani (LB) media or M9 minimal media. Growth inhibition represented by dot in black against gram-negative (*Escherichia coli*); and dot in white against gram-positive (*Bacillus cereus*).

The application of M9 minimal medium or other equivalent media has been widely used for natural product explorations from microorganisms, exemplified by pacifibactin,[181] crochelin,[182] and fabrubactin.[183] Generally, the isolated compounds cultivated under this defined medium displayed metal chelating activity. Understandably since M9 minimal media is only comprised of salt and glucose as the main carbon source. Therefore, to survive in this condition, microorganisms excreted special metabolites such as siderophore to acquire iron that is essential for their optimal growth.[184]

Variovorax sp. is a gram-negative bacterium and member of plant growth-promoting rhizobacteria (PGRP).[185–187] It is also known to have ecological role for degradation of some dangerous substances such as isoprene,[188] linuron[189] and 3,3'-thiodipropionate.[190] Nonetheless, the natural product potential of these bacteria is poorly studied. So far, only four group of compounds have been reported,[191–193] suggesting that the exploration of natural products from this genus rare bacteria remains widely open. Therefore, we further screened the other ichip-domesticated *Variovorax* strains in our collection for production of siderophore-type compounds using chrome azurol S (CAS) assay.[194] This assay is based on the ferric ion binding competition between CAS and the other chelator (Experimental Section). The positive result represented by the coloration change from deep blue of complex CAS-ferric ion to pale pink. The result can be quantified by measuring the resulting absorbance at UV 630 nm. As much as sixteen in-house *Variovorax* spp. strains were subjected to CAS assay screening after cultivation under combinations of several rich and defined nutrient media. Several *Variovorax* strains including, *Variovorax* sp. H002, C010, M072, B014, and F001exhibited strong ferric

ion chelating activity after cultivation with modified M9 minimal media (Figure 3.5). Among them, we selected *Variovorax* sp. H002 and *Variovorax* sp. B014 for investigation of novel natural products.

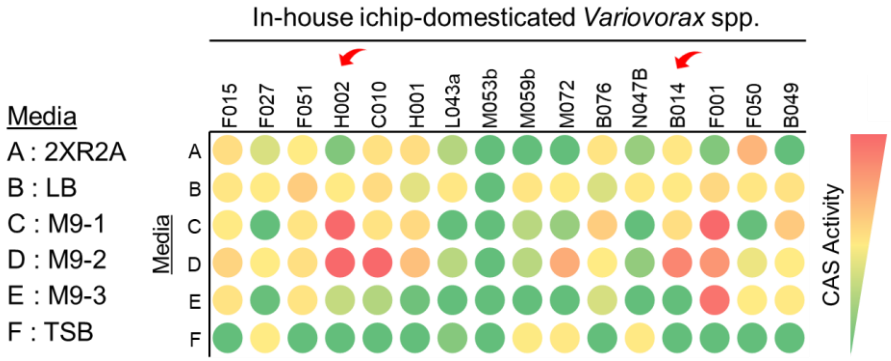


Figure 3.5. CAS assay-based screening of in-house *Variovorax* collections.

3.2.2 Isolation and Characterization of Compounds Derived from *Variovorax* sp. H002 (7-11)

To purify compounds from *Variovorax* sp. H002, the overnight seed culture was inoculated into a fresh modified M9 media (MM9) with resin HP-20, a polysterene/divinylbenzene matrix widely used to catch small molecules, and incubated with shake at 30 °C for one week. After cultivation, the resins were filtrated and washed with distilled water followed by elution with excessive methanol. The methanolic fraction was concentrated under vacuo to obtain crude extract. The obtained crude extract was subjected to gel filtration chromatography packed with LH-20 and eluted with methanol. Finally, the fractions containing active compounds were further purified with semi-preparative HPLC to yield six compounds, **7** (3 mg), **8** (2 mg), **9** (5 mg), **10** (20 mg), and **11** (10 mg).

The molecular formulas of the isolated compounds were assigned based on their HRESIMS as (**7**; *m/z* 1044.5603 $[M-H]^-$ calc. $C_{45}H_{76}O_{17}N_{11}$), (**8**; *m/z* 534.78371 $[M-2H]^{2-}$ calc. $C_{47}H_{79}O_{17}N_{11}$), (**9**; *m/z* 1098.60710 $[M-H]^-$ calc. $C_{49}H_{84}O_{17}N_{11}$), (**10**; *m/z* 535.79122 $[M-2H]^{2-}$ calc. $C_{47}H_{81}O_{17}N_{11}$), and (**11**; *m/z* 1100.62177 $[M-H]^-$ calc. $C_{49}H_{86}O_{17}N_{11}$). In addition, their ^1H NMR showed similar characteristics each other, suggesting that they are a group of analog compounds. To elucidate their planar structure, compound **10** was selected for a complete 1D and 2D NMR analyses due to the availability of pure compound. The ^1H NMR spectrumof **10**, recorded in methanol-d_4, showed several typical Hα at δ_H (3.5-4.6 ppm) and high intensity of methylene signals, indicating a lipo-peptide nature (Figure S3.1). Analyses of ^1H-^1H COSY and ^1H-^{13}C HSQC spectra demonstrated the presence of 7 spin systems including

two ornithine residues, a proline residue, a serine residue, a β-hydroxy aspartic acid, a 3-hydroxy-2-methylheptanoic acid and a fatty acyl moiety (Table S3.2). HMBC correlations of two different delta protons (Hδ) of ornithine residues to overlapped carbonyl carbons of acetyl moieties at δ_C 170.3 ppm, suggesting that both ornithine residues are acetylated at the distal *N* to form *N*(5)-hydroxyl-*N*(5)-acetyl ornithine. The hydroxyl group attached to *N*(5)-acetyl ornithine was supported by tandem mass analysis. A correlation of Hα to a carbon of hydroxylated methine at δ_C 70.1 ppm adjacent to a carboxylate moiety confirmed the presence of a β-hydroxyl aspartic acid. Furthermore, a guanidine moiety was found to be attached to a unique γ amino acid residue as inferred from HMBC correlation of a carbon at 156.5 ppm that normally observed from this moiety.[195–197] The amino acid sequence of **10** was determined by HMBC correlation analysis of Hα to carbonyl amide of the adjacent residue as depicted in Figure 3.6.

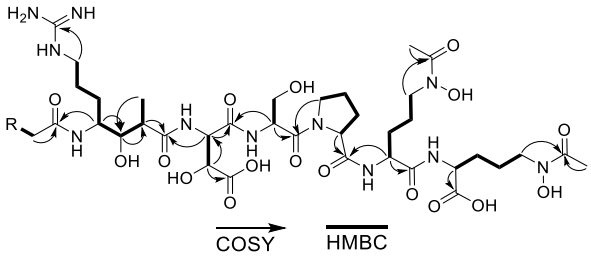


Figure 3.6. Partial planar structure of compound **10** showing COSY correlations with bold bonds and HMBC with arrows.

We then implemented tandem mass analyses of **10** alongside with other isolated compounds to confirm their amino acid sequences. Tandem mass spectrum of **10** showed fragment ions at *m/z* 884, 712, 615, 528, and 397, demonstrating gradual loss of a *N*(5)-hydroxyl-*N*(5)-acetyl ornithine, a *N*(5)-hydroxyl-*N*(5)-acetyl ornithine, a proline, a serine, and a β-hydroxy aspartic acid, respectively. The last fragment ion at *m/z* 397 was calculated to cover the 4-amino-7-guanidino-3-hydroxy-2-methylheptanoic acid and a fatty acyl moiety (Figure S3.8). Tandem mass analyses in conjunction with HRESIMS and NMR data of other isolated compounds suggested modification at fatty acyl moieties. Therefore, we subsequently determined their fatty acid moieties by gas chromatography/mass spectrometry (GC/MS). Gas chromatography analysis is a common approach to identify the composition of fatty acids in a wide variety of samples.[198] This approach is also widely applied for characterization of fatty acid moiety in bacterial lipopeptides.[199–202] Derivatization step prior to GC/MS analysis has been previously proved by several studies to enhance the effectiveness of fragment ion

detection leading to easier determination of the structure of fatty acids.[203–206] Therefore, we initially subjected the fatty acid methyl esters (FAME) derived from **7**-**10** to GC/MS analysis. Analyses of the obtained GC/MS spectra of **10** and **11** allowed us to deduce their fatty acid moieties as a 12-carbon saturated, dodecanoic acid and a 14-carbon saturated , tetradecanoic acid. (Figure 3.7). This is also supported by the fact that their molecular formulas are different for an ethylene group (C_2H_4). Unfortunately, FAMEs derived from **7**-**9** were not efficiently detected by GC/MS due to scarce amounts. In addition, based on the HRESIMS data, the molecular formula of **8** and **9** showed one degree of unsaturation suggesting the presence of a double bond in their structures. Hence, to identify the position of the olefinic bond as well as to enhance ionization, we treated the free fatty acids of **7**-**9** with 2-amino-2-methyl-propanol to form a 4,4-dimethyloxazoline (DMOX) derivative.[207] Based on the obtained spectra, the fatty acid moieties in **7**-**9** were established as decanoic acid, 9-dodecenoic acid, and 7-tetradecenoic acid, respectively (Figure 3.7; Figure S3.10-13). Taken together the obtained data from NMR, MS/MS, and GC/MS analyses, we proposed the planar structures of **7**-**11** as depicted in Figure 3.7. Database search of the proposed planar structures suggested compound **7**-**9** are novel, whereas the identified major compounds **10** and **11** are known variochelins from *Variovorax boronicumulans* reported by Kurth *et. al*.[192]

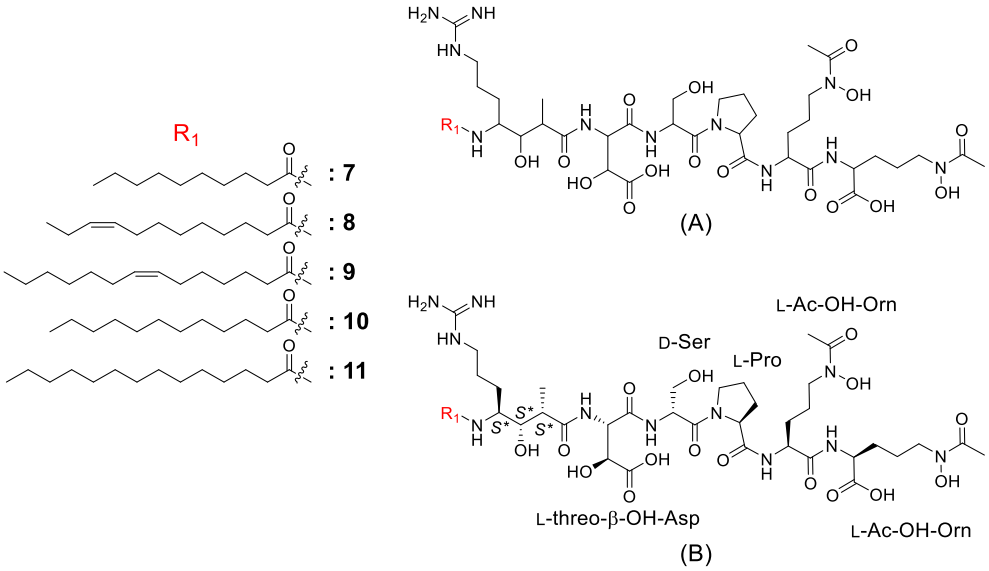


Figure 3.7. Proposed planar structure(A) and absolute configuration of compounds **7**-**11** (B).

The absolute configurations of isolated compounds were determined by application of Marfey's method.[120] Briefly, adequate amounts of compounds were subjected to acid hydrolysis by treatment with 6 M HCl at 110 °C for overnight. The concentrated hydrolysates were derivatized with *N*-(5-fluoro-2,4-dinitrophenyl)-L-leucinamide (L-FDLA) and subjected

to LC/MS analysis. Based on the obtained Marfey's analysis data (Figure 3.7), the absolute configurations of amino acids in **7-11** showed identical result to that of reported compounds.[192] Determination of absolute configuration of the unique γ -amino acid residue using Marfey's analysis was unsuccessful after several attempts to release the arginine moiety. However, by employing 1D NOE experiment, the configurations of stereocenters at γ-amino acid residue (4-amino-7-guanidino-3-hydroxy-2-methylheptanoic acid) are previously proposed as (*2S**, *3S**, and *4S**).[192]

3.2.3 Biosynthesis Pathway of 7-11

Compounds **7-11** are linear peptides possessing several unique amino acid residues including *N*(5)-acetyl-*N*(5)-hydroxyornithine, β-hydroxy aspartic acid, and γ-amino acid residues. These modified residues are of interest to study their biogenesis. Therefore, we sequenced the whole genome of *Variovorax* sp. H002 using hybrid next generation and third generation sequencing technology. The obtained short and long sequence reads were assembled using hybrid SPAdes algorithm.[208] Subsequently, to identify the putative biosynthetic gene cluster, the draft genome of *Variovorax* sp. H002 was submitted to antiSMASH bacterial version 6.0.[209] As a result, a candidate region DNA sequence spanned for 41-kb named *varH002* is proposed to involve in the biosynthesis of compound **7-11**. The putative biosynthetic pathway of known variochelins has been previously identified from genome of *Variovorax boronicumulans.*[192] The proposed biosynthetic gene cluster of **7-11** is provided in Figure 3.8 and their annotation in Table 3.1.

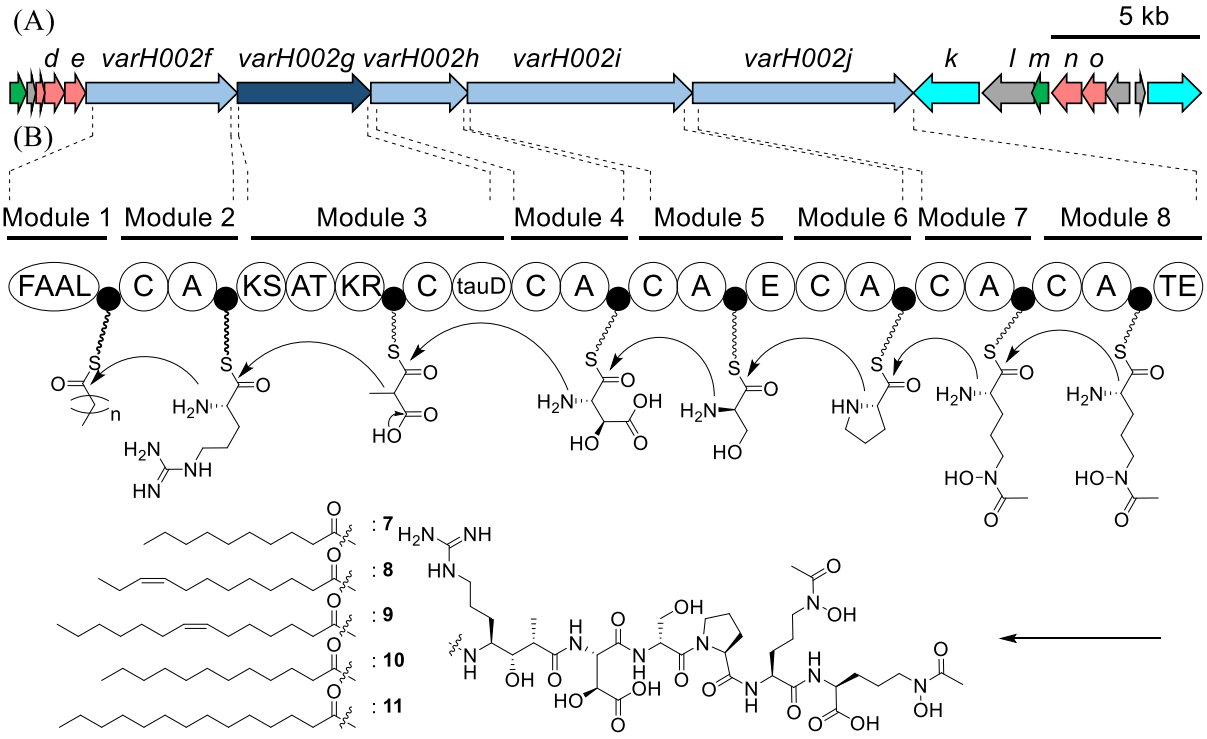


Figure 3.8. (A) The biosynthetic gene cluster of **7-11**; (B) The modular assembly and proposed biosynthesis of **7-11**. Domain annotation: FAAL, fatty acyl-AMP ligase; ●, (acyl carrier protein or peptidyl carrier protein); C, condensation; A, adenylation; KS, ketosynthase; AT, acyltransferase; KR, ketoreductase; TauD, hydroxylase; E, epimerization; TE, thioesterase

Table 3.1 Annotation of biosynthetic genes of *varH002*

Gene	Length (AA)	Proposed function (domain architecture)
varH002a	177	DNA-directed RNA polymerase sigma-70 factor
varH002b	81	Anti-FecI sigma factor FecR
varH002c	82	MbtH domain-containing protein
varH002d	248	Type II thioesterase
varH002e	233	4'-phosphopantetheinyl transferase
varH002f	1754	Nonribosomal peptide synthase (FAAL-ACP-C-A-PCP)
varH002g	2351	Polyketide synthase (KS-AT-KR-ACP-TauD)
varH002h	1034	Nonribosomal peptide synthase (C-A-PCP)
varH002i	2587	Nonribosomal peptide synthase (C-A-PCP-E-C-A-PCP)
varH002j	2458	Nonribosomal peptide synthase (C-A-PCP-C-A-PCP-TE)
varH002k	815	TonB-dependent receptor
varH002l	342	Anti-FecI sigma factor FecR
varH002m	192	RNA polymerase subunit sigma-24
varH002n	439	Ornithine *N*-monooxygenase
varH002o	368	*N*-hydroxyornithine acetyltransferase
varH002p	261	Ferric iron reductase
varH002q	77	Anti-FecI sigma factor FecR
varH002r	559	Peptide transporter

The *varH002* gene cluster is comprised of five core biosynthetic genes that are further divided into eight modules following hybrid NRPS and PKS system. Prediction of substrate specificity of each adenylation (A) domain has been done using *in silico* PKS/NRPS prediction tool[210] and the result is provided at Table 3.2. The A domain at the second module is predicted to recognize an arginine residue prior to incorporation into peptidyl carrier protein (PCP). Therefore, the presence of a guanidino-containing γ-amino acid residue in **7-11** can be rationalized as the product of a decarboxylative Claisen condensation of arginine and methylmalonate by the adjacent PKS module. Subsequently, a keto reductase (KR) domain in the PKS module will reduce the β -carbonyl of arginine into a β-hydroxyl group. Interestingly, we can observe the presence of C domain and a Taurine deoxygenase (TauD) at the C terminus of the PKS module. It is proposed that these domains are responsible for hydroxylation of aspartic acid at the next module.[211] The other interesting features of *varH002* gene cluster are the presence of two accessory genes (*varH002n and o)* annotated as *N*-monooxygenase and *N*-hydroxy acetyltransferase. Both genes are proposed to be responsible in decorating ornithine residue to form *N*(5)-acetyl-*N*(5)-hydroxyornithine.[212]

Table 3.2 Substrate prediction of adenylation (A) domains in *varH002* gene cluster.

Domain	Consensus	Prediction
A2	D A E D L G A I	Arg
A4	D L T K V G H V	Asp
A5	D V W H V S L I	Ser
A6	D V Q F L A Q V	Pro
A7	D G E C T G G I	Orn
A8	D G E C T G G I	Orn

The biosynthesis of **7-11** is started with the activation of fatty acyl moiety by FAAL domain and transferred it into adjacent ACP domain. Subsequently, loaded fatty acid will be incorporated with the activated substrate in the adjacent module by C domain, following collinear logic of assembly line enzymology in a stepwise manner until reach the last module of the gene cluster.[155] Once the final product completed, a TE domain at the C terminus of the last module will release the linear lipopeptide. Hence, variation among the isolated compounds is occurred at the early step of the biosynthesis, where FAAL domain recognizes at least three (C10, C12, and C14) saturated fatty acids and (C12 and C14) unsaturated fatty acids. To confirm the role of *varH002* in the biosynthesis of **7-11**, a mutant *Variovorax* sp. H002 strain lacking *varH002g* (ΔH002::*varH002g*) was generated by employing homologous recombination technique and confirmed using PCR detection (Figure S3.16). The metabolite

profile of the mutant strain after cultivation under iron deprived medium was compared to the wild type. As the result, the production of major compounds **10** and **11** in mutant strain is abolished while a contrast profile was observed from wild type, suggesting that the gene cluster *varH002* involve in the biosynthesis of **7-11** (Figure 3.9).

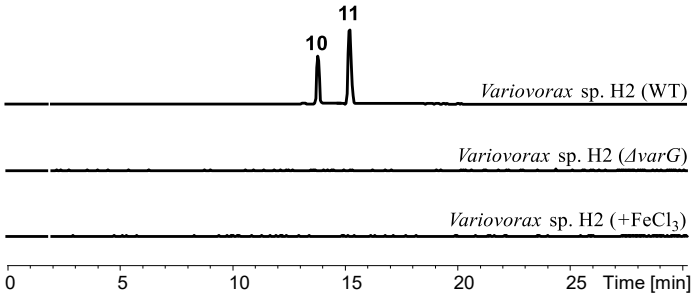


Figure 3.9 Detection of compounds **10** and **11** in *Variovorax* sp. H002 wild type (line 1), *varH002g* null mutant (line 2), and *Variovorax* sp. H002 wild type with additional $FeCl_3$ in cultivation media(line 3).

3.2.4 Isolation and Characterization of Compounds Derived from *Variovorax* sp. B014 (12-16)

The natural product isolation from ichip-domesticated *Variovorax* sp. B014 was initiated by the production of metabolites under M9 minimal media containing lack of ferric ion. After cultivation, the supernatant was subjected to a ODS open column followed by elution with gradual concentration of water and methanol to yield several fractions that were monitored by CAS assay activity. The fractions containing active molecules were concentrated in vacuo to obtain a semi-crude extract. Subsequently, it was subjected to several rounds of HPLC purification to furnish five compounds **12** (10 mg), **13** (6 mg), **14** (6 mg), **15** (12 mg), and **16** (8 mg).

The ^{1}H NMR spectra of **13**, recorded in several deuterated solvents (Figure S3.17), showed broad signals in some key regions, implying a challenge in structure elucidation using the intact compound. Previous studies suggested the advantage of utilization of complex metal-natural products in addressing such issues.[193,213,214] Therefore, the structure elucidation of **13** has been conducted with a form **13**-Ga(III) complex. To do that, initially, metal contaminants were removed from the intact compound **13** by treatment with 8-hydroxy quinone for overnight. The 8-hydroxy quinone-metal complexes were separated to afford a metal-free **13**. Subsequently, addition of Ga(III) sulfate to the obtained metal-free **13** generated **13**-Ga(III) complex molecules. The ^{1}H NMR spectrum of **13**-Ga(III) showed improved signals suitable for structure elucidation (Figure S3.18). Analyses of ^{1}H-^{1}H COSY correlations of **13**-Ga(III),

recoded in methanol-d_3, indicated the presence of several spin systems including, two ornithine residues, a serine residue, a threonine residue, a 3-hydroxybutyric acid, a 4-amino-3-hydroxy-heptanoic acid, and a fatty acyl moiety (Table S3.3). In addition, ^{1}H-^{13}C HSQC spectra showed the presence of two aldehyde groups (Figure S3.20). Based on HMBC analyses showed correlation of a methylene proton at position 5' (δ_H 3.50) of ornithine to an aldehyde group (δ_C 154.4), suggesting the presence of *N*(5)-formyl-*N*(5)-hydroxyornithine. The other aldehyde group was observed to connect with the γ -amino acid to form *N*(7)-formyl-*N*(7)-hydroxy-4-amino-3-hydroxy-heptanoic acid. Interestingly, the carbonyl carbon of 3-hydroxybutyric acid at δ_C 162.2 was attached to another ornithine residue to form a *N*(5)-3-hydroxybutyric acyl-*N*(5)-hydroxyornithine. Thus, the fragment structures of **13** are comprised of a serine residue, a threonine residue, a *N*(5)-formyl-*N*(5)-hydroxyornithine residue, a *N*(7)-formyl-*N*(7)-hydroxy-4-amino-3-hydroxy-heptanoic acid, a *N*(5)-3-hydroxybutyric acyl-*N*(5)-hydroxyornithine and a fatty acyl moiety. The molecular formula **13**-Ga(III) was assigned based on HRESIMS data, *m/z* 985.4013 $[M+H]^+$ calculated for $C_{40}H_{68}N_8O_{16}Ga$, suggesting the fatty acyl moiety in **13**-Ga(III) as C10-saturated, decanoic acid. Unfortunately, attempt to determine the sequence of **13**-Ga(III) based on HMBC analysis was failed due to lack of correlation signals from Hα and amide protons to carbonyl carbon, although a correlation between carbonyl carbon of fatty acyl moiety (δ_C 176.4) and Hγ proton of γ-amino acid moiety (δ_H 3.96) was detected.

Alternatively, we performed tandem mass analysis of the intact compound **13** to determine the location of each fragment structure. However, MS/MS spectrum of **13** only showed two fragment ions at *m/z* 476.24 and 563.26 indicating loss of a serine residue (Figure S3.22). Based on this observation, we suspected an occurrence of a cyclic fragment in **13**. It was supported by the low-field chemical shift of the oxymethine proton of hydroxybutyrate unit at δ_H 4.98 and its carbon at δ_C 71.6, suggesting the presence of an ester lactone bridge of the cyclic fragment. Consequently, we speculated that methanolysis product of **13** will give an improved MS/MS spectrum. Therefore, the intact compound **13** was then treated with 7 M ammonia solution in methanol prior to implemented to tandem mass analysis. The successful generation of *seco*-acid methyl ester peptide of **13** was confirmed by the increment 32 mass units from the intact compound (Figure S3.23). Expectedly, the MS/MS spectrum of ring-opened compound of **13** allowed us to establish the position of each fragment in **13**. Taking together all the obtained data, we proposed the planar structure of **13** as depicted in Figure 3.11

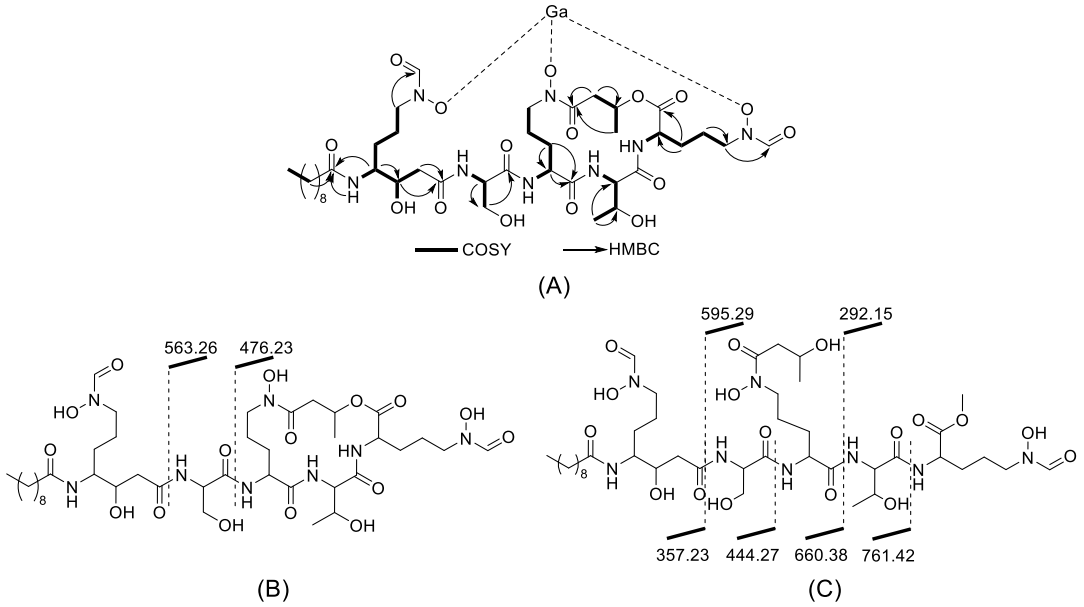


Figure 3.10 The proposed planar structure of **13**. (A) COSY and HMBC key correlations; (B) Detected fragment ions derived from MS/MS analysis of **13**; (C) Detected fragment ions derived from MS/MS analysis of a ring-opened product of **13**.

The molecular weight of **12** (HRESIMS; *m/z* 891.4666 $[M+H]^+$ calc. $C_{38}H_{66}N_8O_{16}$) is 28 mass units smaller than that of **13** (HRESIMS; *m/z* 919.4989 $[M+H]^+$ calc. $C_{40}H_{70}N_8O_{16}$), suggesting that they are different by one ethylene (C_2H_4) group. Based on this data in conjucntion with ^{1}H NMR data, we concluded that **12** is an analog of **13** possesing C8-saturated, octanoic acid as its fatty acid moiety. In addition, the HRESIMS of **14** (*m/z* 937.5107 $[M+H]^+$ calc. $C_{40}H_{72}N_8O_{17}$) indicated an addition of H_2O to **13**, suggesting that **14** is a linear analog of **13**. This suggestion was supported by MS/MS analysis of **14** showing a clear fragmentation pattern (Figures S3.24). Tandem mass analysis of **16** generated similar fragment ions to the *seco*-acid methyl ester peptide of **13**, suggesting **16** as an analog of **14** with an additional of methyl group in the ornithine residue at the C-terminus (Figures S3.26). However, it was difficult to determine its position using tandem mass analysis. Therefore, by analyzing its HSQC spectrum, we deduced that **16** is a methyl ester analog of **14** (Figures S3.27). Finally, tandem mass analysis of **15** generated similar fragment ions toward C terminus compared to that of **16**, but different by 28 mass units toward N terminus suggesting their modification at fatty acid moieties (Figure3.11; Figure S3.28). Database search of proposed planar structures suggested all the isolated compounds from *Variovorax* sp. B014 are novel natural products.

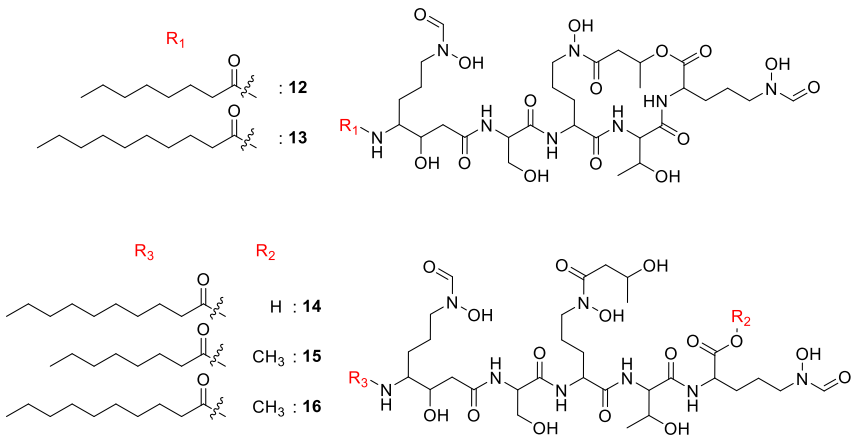


Figure 3.11 The proposed planar structure of **12-16**.

Absolute configurations of the isolated compounds were determined by application of Marfey's method. Briefly, adequate amounts of compounds were subjected to acid hydrolysis by treatment with 6 M HCl at 110 °C for overnight. The concentrated hydrolysates were derivatized with *N*-(5-fluoro-2,4-dinitrophenyl)-L-leucinamide (L-FDLA) and subjected to LC/MS analysis. This analysis established the serine and threonine residue to be in L configuration. Marfey's method also revealed the presence of L and D-ornithine residue in the isolated compounds. The positions of each ornithine were suggested by analysis of domain organization in biosynthetic gene cluster of **12-16**. The configuration of Hγ in γ-amino acid (*N*(7)-formyl-*N*(7)-hydroxy-4-amino-3-hydroxy-heptanoic acid) was proposed by the similar approach as well. Finally, the configurations of two stereocenters were unable to be determined in this study (Figure 3.12)

3.2.5 Biosynthesis Pathway of 12-16

Compounds **12-16** are characterized by the presence of three modified ornithine residues. One of them is occurred as a product of decarboxylative Claisen condensation with malonyl, which is further modified by a β-ketoreductase to form a γ-amino acid, a *N*(7)-formyl-*N*(7)-hydroxy-4-amino-3-hydroxy-heptanoic acid. This suggests the involvement a PKS module in the NRPS assembly line, which resembles the biosynthetic pathway of the previously isolated compounds **7-11**. Therefore, to confirm this assumption, we sequenced and annotated the entire genome of *Variovorax* sp. B014 by combining next and third-generation sequencing technologies. Putative gene clusters of secondary metabolites were identified by submitting the obtained draft genome to a web-based bioinformatic tool called antiSMASH 6.0.[209] As the result, eight biosynthetic gene loci were predicted to be related to secondary metabolisms. One of them, the third locus, was putatively proposed as the target biosynthetic

gene cluster based on the presence of one PKS module in the middle of NRPS assembly line. In addition, three out of five predicted substrates using PKS/NRPS prediction tool[210] were consistent with the amino acid compositions of the isolated compounds (Table 3.3). Thus, the proposed biosynthetic gene cluster of **12-16** was identified as *varB014*.

Table 3.3 Substrate prediction of adenylation (A) domains in *varB014* gene cluster.

Domain	Consensus	Prediction
A2	D V W N I G L I	Thr
A4	D V W H V S L I	Ser
A5	D G E G S G G V	Orn
A6	D F W N I G M V	Thr
A7	D V W N I G L I	Thr

The *varB014* is comprised of five core biosynthetic genes as well as several accessory genes spanned for 36-kb contiguous DNA (Figure 3.12). Each of the involved genes (*varB014a-p*) was manually annotated as provided in Table 3.4. One of the accessory genes (*varB014l*) is annotated as a member of FAD-dependent monooxygenase family. Some of the members of this tailoring enzyme have been experimentally proved to hydroxylate *N*(5)-ornithine in vicibactin[212] and coelichelin.[215] Therefore, we proposed that encoded enzyme VarB014l is responsible for the synthesis of a key intermediate of three modified ornithine in **12-16** (Figure S3.29). This intermediate is further conjugated with a 3-hydroxybutyric acid by VarB014m, an acetyltransferase, to provide a final substrate for the A domain at module 5. Additionally, the hydroxylated ornithine serves as a substrate for the subsequent N^{10}-formyl-tetrahydrofolate-dependent (N^{10}-fH_4F) VarB014n-catalyzed formylation reaction,[216] providing the substrate for A domain at modules 2 and 7 (Figure 3.12; Figure S3.29).

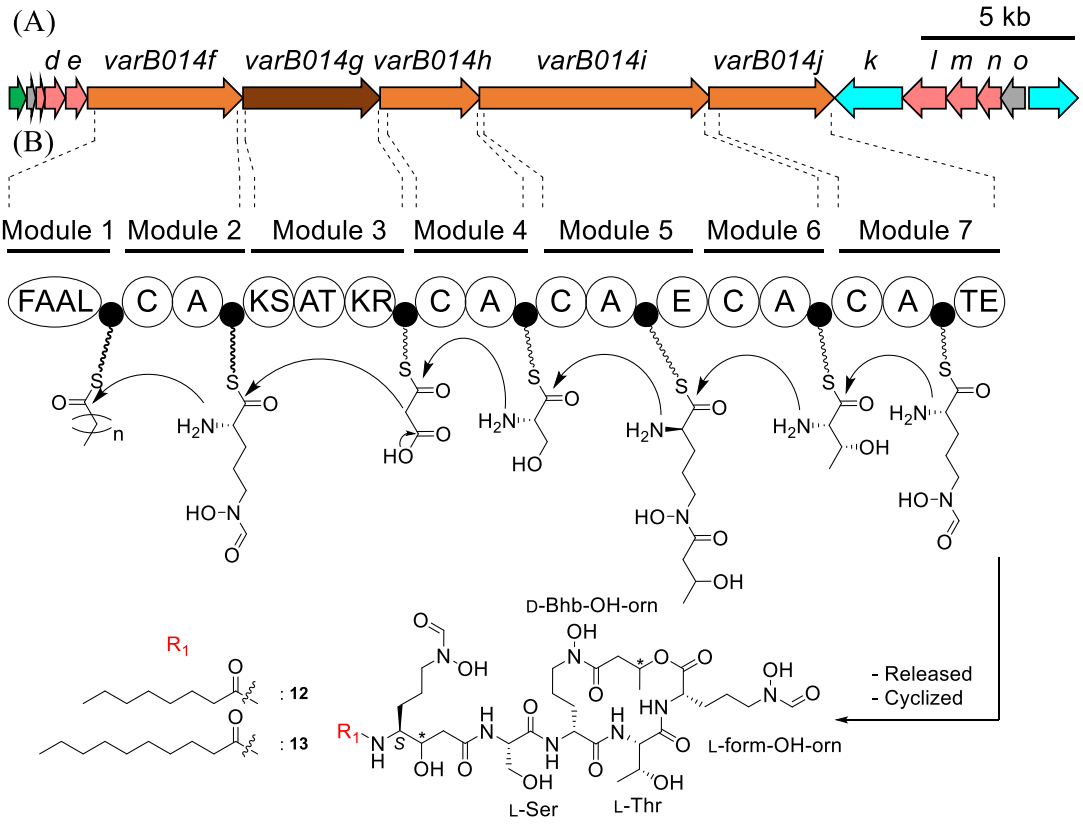


Figure 3.12 (A) The biosynthetic gene cluster of **12-16**; (B) The modular assembly and proposed biosynthesis of **12-16**. Domain annotation: FAAL, fatty acyl-AMP ligase; ●, (acyl carrier protein or peptidyl carrier protein); C, condensation; A, adenylation; KS, ketosynthase; AT, acyltransferase; KR, ketoreductase; E, epimerization; TE, thioesterase

Table 3.4 Annotation of biosynthetic genes of *varB014*

Gene	Length (AA)	Proposed function (domain architecture)
varB014a	181	DNA-directed RNA polymerase sigma-70 factor
varB014b	67	Anti-FecI sigma factor FecR
varB014c	84	MbtH domain-containing protein
varB014d	245	Type II thioesterase
varB014e	230	4'-phosphopantetheinyl transferase
varB014f	1762	Nonribosomal peptide synthase (FAAL-ACP-C-A-PCP)
varB014g	1526	Polyketide synthase (KS-AT-KR-ACP)
varB014h	1110	Nonribosomal peptide synthase (C-A-PCP)
varB014i	2621	Nonribosomal peptide synthase (C-A-PCP-E-C-A-PCP)
varB014j	1362	Nonribosomal peptide synthase (C-A-PCP-TE)
varB014k	733	TonB-dependent receptor
varB014l	441	Ornithine *N*-monooxygenase
varB014m	344	*N*-hydroxyornithine acyltransferase
varB014n	283	*N*-hydroxyornithine transformylase
varB014o	242	Ferric iron reductase
varB014p	573	Peptide transporter

The encoded seven PKS and NRPS modules by *varB014f-j* are proposed to assemble the molecular backbone of **12-16** (Figure 3.12). Hence, the biosynthetic pathway of **12-16** is started with the activation of C8 or C10 saturated fatty acid by FAAL domain in the first module. Subsequently, these fatty acyl moieties are incorporated into a linear oligomer by successive condensation steps following the collinear logic of modular assembly line enzymology.[155] The PKS module encoding by *varB014g* contains the typical ketosynthase (KS), acyltransferase (AT), β-ketoreductase (KR), and acyl carrier protein (ACP), confirming the generation of a γ-amino acid, *N*(7)-formyl-*N*(7)-hydroxy-4-amino-3-hydroxy-heptanoic acid in **12-16**. All the encoded NRPS modules harbor complete set of condensation (C), adenylation(A), and peptidyl carrier protein (PCP) domains. In addition, the presence of an epimerization (E) domain at module 5 suggested a stereochemical inversion of the installed *N*(5)-3-hydroxybutyric acyl-*N*(5)-hydroxyornithine, establishing the final D-configured product. Consequently, the configurations of another ornithine and Hγ of γ-amino acid residues were determined as L and *S*, due to the absence of E domain in the corresponding modules. Finally, once the chain elongation is completed, a TE domain at the last module releases linear peptides **14-16** or further catalyzes cyclization as found in **12** and **13** (Figure 3.12)

The gene cluster *varB014* also encodes other important accessory proteins that are essential for NRPS and PKS enzymes. An MbtH-like protein encoded by *varB014c* is proposed to involve in the substrate activation in NRPS systems,[217] whereas the type II thioesterase encoded by *varB014d* is proposed to act as an proof-reading system for the assembly line by removing the aberrant intermediates.[218] Additionally, the *varB014e* encoded 4'-phosphopantetheinyl transferase is reported to be essential in the conversion of ACP and PCP domains from the inactive *apo* protein into the active *holo* forms.[219] The remaining enzymes are likely involved in siderophore transport and gene expression regulators. Finally, generation of a mutant strain of *Variovorax* sp. B014, to correlate the gene cluster *varB014* with the production of **12-16**, is underway.

3.2.6 Biological Activity of Isolated Compounds

All isolated peptide-type compounds in this study showed similar characteristics in their moieties including the presence of *N*(5)-acyl-*N*(5)-hydroxy ornithine, γ-amino acid, and their N-termini were capped by a fatty acid moiety. The production of *N*(5)-acyl-*N*(5)-hydroxy ornithine residue is catalyzed by two step enzymatic reactions. Initially a key intermediate, *N*(5)-hydroxy ornithine, is catalyzed by an FAD-dependent monooxygenase followed by acylation by acyltransferase, acetyltransferase, and transformylase.[212,215,216] These enzymes are

responsible for formation of a functional hydroxamate group that essential in metal coordination. The hydroxamate functional group is widely reported as metalloenzyme protease inhibitor[220] that lead to anti-cancer activity[221] and anti-parasitic activity.[222,223] In addition, naturally occurring hydroxamates that produced by microorganisms are often associated with metal chelation (particularly iron) activity.[224,225]

Iron is one of the essential minerals that is required for many biological processes and metabolisms. Microorganisms fully rely on the availability of soluble iron in their environment. Cellular uptake of iron by microorganisms is restricted to its physiologically most relevant species, which are Fe^{2+} (ferrous ion) and Fe^{3+} (ferric ion).[184] In the environment, iron mostly occurred as a complex with oxide and water to form $Fe_2O_3 \cdot nH_2O$, a stable complex in the presence of water and oxygen, resulting the availability of soluble Fe^{3+} (ferric ion) in the environment is extremely low. Therefore, microorganisms produce specialized metabolites for acquisition of extracellular ferric ion called siderophores.[184] Once the secreted siderophores captured ferric ion, the uptake of Fe-siderophore complex will be mediated by TonB.[226] Based on their structural characteristics, siderophore is categorized into i) hydroxamate-type such as desferrioxamine,[224] ii) catecholate-type such as enterobactin,[227] iii) phenolate-type such as yersiniabactin,[228] iv) carboxylate-type such as staphyloferrin,[229] and vi) mixed-type such as aerobactin (Figure 3.13).[230]

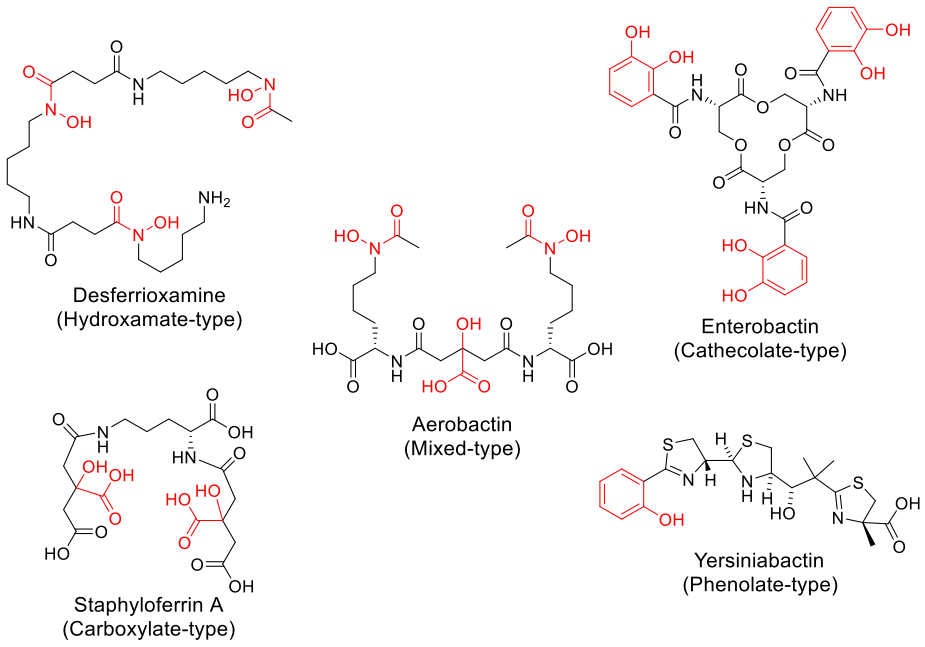


Figure 3.13 Type of siderophores and their examples, the functional groups are highlighted in red.

Based on these observations, we eager to confirm role of the isolated compounds as siderophores. Initially, Ferric ion chelating activity of the isolated compounds was examined by chrome azurol-S (CAS) assay following the procedure by Schwyn and Neilands with a minor modification (Experimental Section).[231] The compounds were diluted in a microplate containing 100 µl of miliQ water. Subsequently, equal volumes of freshly prepared CAS solutions were added to the diluted tested compounds. Principally, binding competition between CAS and siderophore to available ferric ion was representatively displayed by the intensity of the blue color generated from complex of CAS-ferric ion. Therefore, reduction of intensity of the blue color and generation a pale pink solution suggested ferric ion chelating activity. Accordingly, the provided values represented the minimal concentration that still showed chelating activity (Table 3.5). Isolated compounds from *Variovorax* sp. H002 exhibited chelating activity similar to EDTA as a positive control at 7.81 µM. Additionally, isolated compounds from *Variovorax* sp. B014 showed chelating activity at minimal concentration of 15.6 µM, a slightly lower compared to the EDTA.

Table 3.5 Antibacterial activity and ferric ion chelating activity of isolated compounds. MIC (Minimum inhibitory concentration); (EC: *Escherichia coli*); (BM: *Burkholderia multivorans*); (BC: *Bacillus subtillis*); (KR: *Kocuria rhizophila*)

Comp.	Antibacterial Activity MIC (µM)				Fe^{3+} chelating Activity (µM)	Comp.	Antibacterial Activity MIC (µM)				Fe^{3+} chelating Activity (µM)
	EC	*BM*	*BC*	*KR*			*EC*	*BM*	*BC*	*KR*	
7	9.67	17.0	>63	>63	7.81	**12**	>74	8.77	>74	>74	15.6
8	n.t.	n.t.	n.t.	n.t.	n.t	**13**	>71	8.49	>71	>71	15.6
9	10.6	17.4	>60	>60	7.81	**14**	>70	>70	>70	>70	15.6
10	5.73	9.30	>61	>61	7.81	**15**	>71	>71	>71	>71	15.6
11	4.06	13.2	>59	>59	7.81	**16**	>69	>69	>69	>69	n.t.

Additionally, we investigated the metabolite production of *Variovorax* sp. H002 in both rich and lack of ferric ion M9 minimal media. As a result, the production of two major compounds **10** and **11** is abolished in ferric ion rich condition, whereas these compounds are readily detected in ferric ion starvation conditions (Figure 3.14). This indicates that productions of **10** and **11** are related to the availability of ferric ion in the cultivation media. We further investigated the minimal concentration of $FeCl_3$ that affect production of these compounds. Hence, we cultivated *Variovorax* sp. H002 in four different ferric ion concentrations including 0, 0.1, 0.5 and 5 µM. After cultivation, the metabolite profiles of each culture were examined by quantitative HPLC analysis using purified **11** as an external standard. The maximum

production of **11** was observed in the absence of additional ferric ion in cultivation media, whereas it slightly reduced in the presence of 0.1 µM of ferric ion (Figure 3.14). In addition, the production of **11** was significantly reduced when 0.5 µM ferric ion was added to the cultivation media. Furthermore, according to the obtained data, *Variovorax* sp. H002 refrained to produce **11** in the cultivation media containing 5 µM of ferric ion (figure 3.14.)

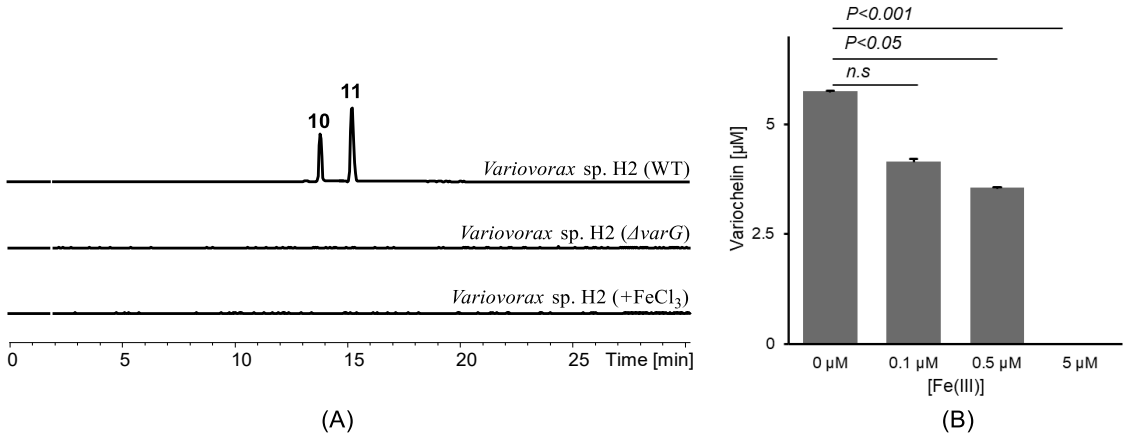


Figure 3.14 (A) Metabolite profile of *Variovorax* sp. H002 wild type in the presence of ferric ion (line 3); (B) The effect of various concentration of ferric ion to the production of **11**

The isolated compounds were investigated for their antibacterial activity against gram negative-bacteria (*Escherichia coli* JW5503 and *Bukholderia multivorans* NBRC 102086) and gram-positive bacteria (*Bacillus cereus* NBRC 15305 and *Kocuria rhizophila* NBRC 12708). The result demonstrated antibacterial activity of the compounds isolated from *Variovorax* sp. H002 against gram-negative bacteria but inactive against gram-positive bacteria (Table 3.5). In addition, compound **12** and **13** isolated from *Variovorax* sp. B014 also showed antibacterial activity against *Burkholderia multivorans* but inactive against the other tested bacterial strains. In contrast, compound **14**-**16** were inactive against all tested bacterial strains (Table 3.5), suggesting antibacterial property preferred for cyclic compounds. Moreover, the complex Ga(III)-containing compounds of **12** and **13** were inactive, indicating that the activity of **12** and **13** are due to cellular iron depletion inhibiting the grow of *Burkholderia multivorans* (Figure S3.30).

3.3 Conclusion

Application of ichip-based cultivation approach has led to the domestication of a thousand bacterial strain after *in situ* incubation at several locations around the Sapporo campus of Hokkaido University. Among them, as much as 25 domesticated strains are categorized at

least new microbial species. Two *Variovorax* spp. strains (H002 and B014) were selected based on the antibacterial and CAS assay screening result. The metabolites that responsible for their activity were successfully identified and characterized. Isolated compounds **7**-**16** are siderophores characterized by the presence of hydroxamate moieties installed in their ornithine residues. Compounds **7**-**11** showed ferric ion binding activity similar to the positive control (EDTA) whereas a little bit lower activity displayed by compound **12**-**16**. In addition, compound **7**-**13** exhibited antibacterial activity against gram-negative *Burkholderia multivorans*. Our results suggested that even in soil ecosystem around us can actually become a source of novel natural products. Additionally, investigation of two rare bacteria had permitted us to isolate eight novel compounds, demonstrating the great potential of soil microbes.

3.4 Author Contributions

This sub-project has been done in the support by collaborations with several researchers and their contributions are as follows : Jabal R. Haedar, Masachika Takaoka, Kumiko Imachi, Shota Tomimoto, Takahiro Jomori, Agustinus R. Uria, Kenichi Matsuda, Masachika Takada isolated environmental bacterial strains using ichip. Jabal R. Haedar, Aya Yoshimura, Toshiyuki Wakimoto contributed to the experimental design. Jabal R. Haedar performed isolation and characterization compounds. Jabal R. Haedar, Aya Yoshimura, Kenichi Matsuda performed whole genome sequencing and downstream analysis. Jabal R. Haedar, Aya Yoshimura, Kenichi Matsuda, Toshiyuki Wakimoto evaluated the data.

Chapter 4

Discussion and Summary

Natural product (NP) and its derivatives have made a major contribution to drug discovery and development. Thousands of NPs have been clinically approved to treat various human diseases.[6] However, new therapeutic agents are still required to treat cancer and infectious diseases. In addition, recent exploration of novel NPs is challenged by high-rate rediscovery of previously reported molecules. On the other hand, metagenomic analysis of many marine invertebrates and soil ecosystems revealed enormous of biosynthetic gene clusters (BGCs) encoded for new bioactive small molecules.[61,92,100,169] Furthermore, these potential BGCs are associated with diverse rare bacteria that have never been investigated before.[61,92] Thus, targeting these untapped microbial strains from marine and soil ecosystems would lead to discovery of novel natural products.

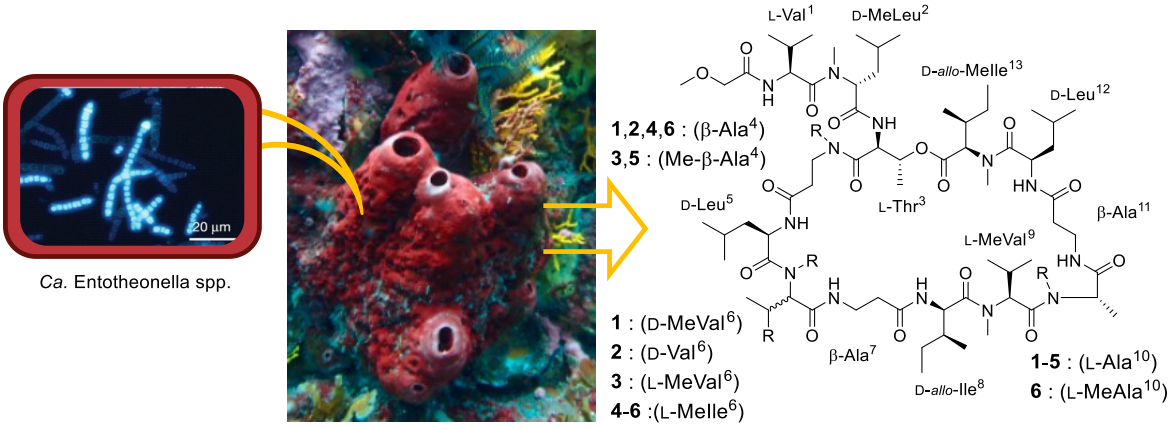


Figure 4.1 Exploration of NPs from marine sponge *T. swinhoei* inhabited by promising yet uncultured bacteria *Ca.* Entotheonella spp. (The fluorescence imaging of *Ca.* Entotheonella spp. is adopted from *Wilson et al*. 2014).

Therefore, in the first project, we targeted novel NPs from the talented bacterial symbiont associated with marine sponge *Theonella swinhoei*. This marine invertebrate is well-known as a prolific producer of diverse array of complex bioactive compounds.[15–21] Additionally, recent genomic analysis suggested that those active molecules are actually encoded in the genome of yet uncultured symbiotic bacteria named *Candidatus* Entotheonella spp.[61,100] The effort to grow these bacteria under standard laboratory conditions remains elusive. Nevertheless, the produced secondary metabolites are readily detected in the host organisms. Thus, initially we investigated novel NPs from *T. swinhoei*. Guided by unique

molecular weight, we succeeded in isolating three new members of theonellapeptolide-type compounds (**1**-**3**) along with three known compounds (**4**-**6**). The molecular structures of isolated compounds were elucidated by combination of 2D NMR and tandem mass analyses and subsequent applications Marfey's method to determine their absolute configurations. Accordingly, all the newly isolated compounds possess valine (in **1**) or methyl valine (in **2** and **3**) residue at amino acid residue position 6 while methyl isoleucine present in known **4**-**6**. Additionally, resemble to **5**, new compound **3** has methyl β-alanine at residue 4 discriminated it with **1** and **2** (Figure 4.1).

Bioactivity investigations of all compounds (**1**-**6**) have been done with several biological systems including, antibacterial, cytotoxic, and anti-austerity activities. While none of them showed antibacterial and cytotoxic activities, compounds (**1**-**6**) inhibited the grow of pancreatic cancer cell MIA PaCa-2 in nutrient starvation conditions (anti-austerity activity), demonstrating a new biological property of marine-derived peptide compounds.[135] Among the tested theonellapeptolide members, the most potent activity displayed by **5**, suggesting the importance of methyl β-alanine and methyl isoleucine residues at position 4 and 6 in theonellapeptolide as an anti-austerity agent. In addition, in a real-time experiment suggested theonellapeptolide (represented by **4**) trigger cell necrotic in 4 hours and full eradication after 8 hours incubation in nutrient starvation condition. Moreover, theonellapeptolides (represented by **4** and **6**) also prevented colony formation of pancreatic cancer cell MIA PaCa-2 in a rich medium, suggesting that these compounds are able to inhibit cell metastasis as one of the characteristics of pancreatic cancers.

The second project aimed to explore the potential bacterial strains from soil ecosystems. Despite soil microbes have been extensively studied for NPs, metagenomic analysis of many soil habitats suggested enormous of BGCs that associated with divers untapped microorganisms.[92] Employing *in situ* cultivation approach using a small device called ichip enabled to domesticate 1046 bacterial strains.[174] These successfully domesticated bacterial strains were screened based on their antibacterial activity after being cultivated under rich (LB) and define (M9) media. As the result, we selected strain *Variovorax* sp. H002 that showed antibacterial activity exclusively after cultivation under M9 minimal media. This media is widely used for productions of small molecules having metal chelator activity from diverse microorganisms.[181–183] Therefore, we further screened all the *Variovorax* strains in our collection using chrome azurol S (CAS) assay for detection of metal chelator (particularly iron)

producer.[231] Accordingly, *Variovorax* sp. H002 and *Variovorax* sp. B014 were selected for investigation of novel NPs.

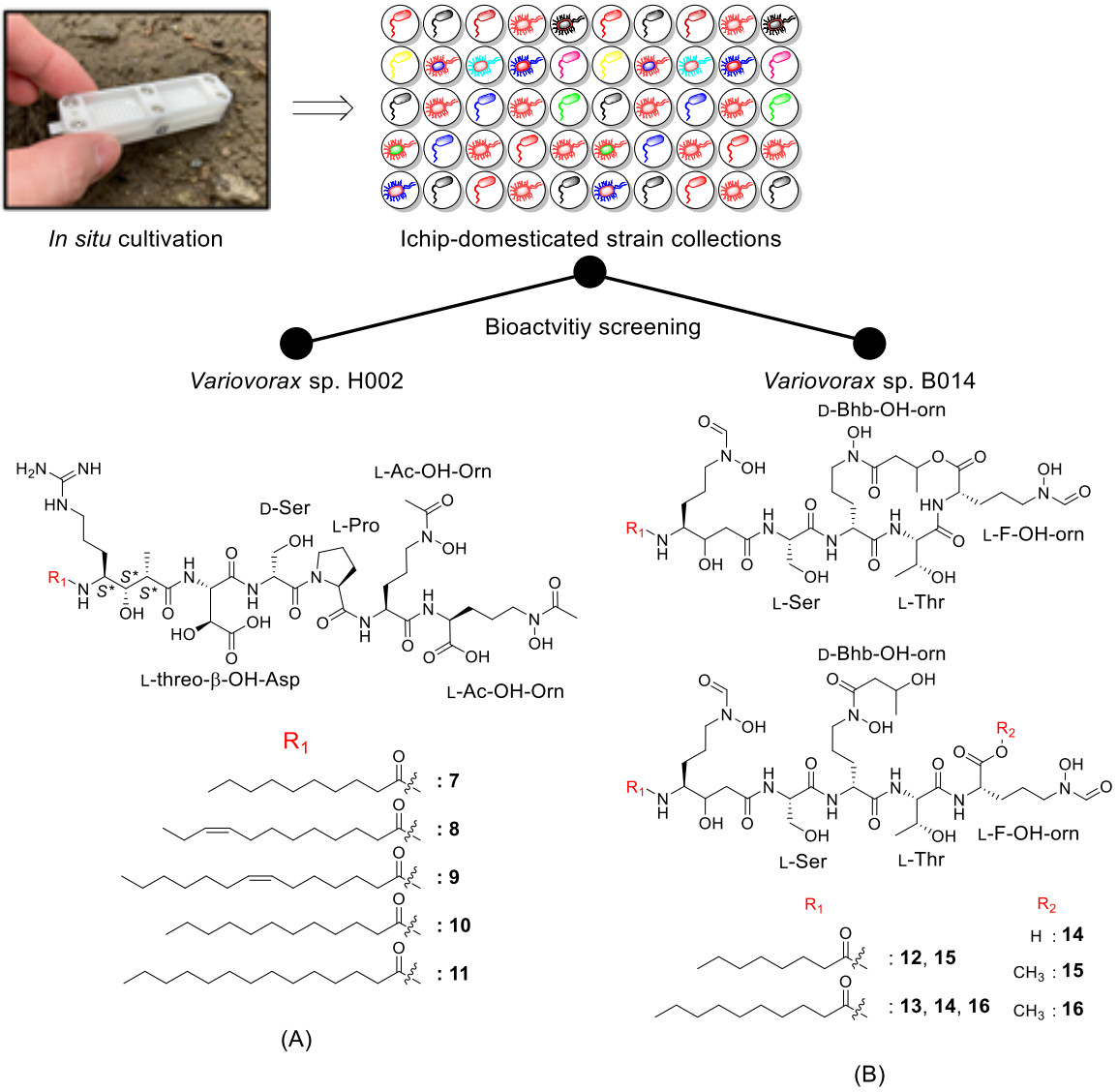


Figure 4.2 Exploration of NPs from ichip-domesticated bacterial strains; (A) Characterized NPs from *Variovorax* sp. H002 and (B) Characterized NPs from *Variovorax* sp. B014.

Five compounds (**7**-**11**) were successfully isolated from strain *Variovorax* sp. H002 as well as five compounds (**12**-**16**) from *Variovorax* sp. B014. Their molecular structures were elucidated by combination of 2D NMR and tandem mass analyses. Compounds **7**-**11** are lipopeptide possessing six amino acid residues and their N termini are blocked by fatty acid moiety. Among the amino acid residues, two *N*(5)-acetyl-*N*(5)-hydroxy ornithine, a β-hydroxy aspartic acid, and a γ-amino acid are installed in **7**-**11**. GC/MS analyses of derivatized free fatty acids from **7**-**11** indicated they different in fatty acid moieties as shown in Figure 4.2A. In addition, **12**-**13** are cyclic compounds having five amino acid residues and fatty acyl

moieties blocking their N termini. These compounds are characterized by the presence of three highly modified ornithine residues including, one *N*(5)-formyl-*N*(5)-hydroxy ornithine, one *N*(5)-3-hydroxybutyricacyl-*N*(5)-hydroxy ornithine, and the last one is in the form of γ-amino acid, *N*(7)-formyl-*N*(7)-hydroxy-4-amino-3-hydroxy-heptanoic acid residue. The main difference between these compounds is in their fatty acid moieties where a C-8 saturated in **12** and a C-10 saturated in **13** (Figure 4.2B). Moreover, compound **14**-**16** are linear peptides having identical amino acid residue with those of **12** and **13**. Compounds **14** and **16** are linear version of **13**, with additional methyl group attached to the carboxylic group at the last ornithine residue in **16**. Whereas compound **15** is analog of **16** having a C-8 saturated fatty acid blocking its N terminus (Figure 4.2B).

Compounds **7**-**16** shared common features with several group hydroxamates installed on each of their ornithine residues that might be responsible for iron coordination as in several known siderophores.[181,182,184,232] Expectedly, the productions of **10** and **11** are significantly decreased after cultivation of *Variovorax* sp. H002 under M9 minimal medium fortified 0.5 μM ferric ions and completely abolished in the presence of 5 μM of ferric ions, suggesting that these compounds are siderophores, produced to acquire ferric ion from environment. Therefore, their ferric ion chelating properties were investigated by using CAS assay. As a result, isolated compounds from *Variovorax* sp. H002 exhibited chelating activity similar to EDTA as a positive control at 7.81 μM. Additionally, isolated compounds from *Variovorax* sp. B014 showed chelating activity at minimal concentration of 15.6 μM, a slightly lower compared to the EDTA. Moreover, antibacterial activity against gram negative *Escherichia coli* and *Burkholderia multivorans* were exhibited by compounds **7**-**16**. Compounds **12** and **13** showed antibacterial activity against *Burkholderia multivorans* while compound **14**-**16** were inactive, suggesting antibacterial property preferred for cyclic compounds.

In this study, we have identified three novel natural products from marine invertebrate inhabited by promising *Ca.* Entotheonella spp. as well as eight novel compounds from two rare bacterial strains from soil ecosystem domesticated by *in situ* cultivation. All the novel molecules showed biological properties as anti-austerity or metal chelating activities. Thus, our findings demonstrated that untapped and rare microorganisms from both marine and terrestrial habitats are promising sources of novel natural products. In addition, our genomic analysis of several ichip-domesticated bacterial strains revealed a large number of biosynthetic gene clusters encoding for bioactive small molecules. These promising strains are currently being intensively studied for novel natural production.

Supplementary Information for Chapter 2

Table of Contents:

Material and Methods

General Procedures Optical rotations were measured on a Jasco P-1030 polarimeter. Infrared spectra were measured on a Jasco FT/IR 4100. NMR spectra were recorded on a JEOL ECX 500 (500 MHz) or a Bruker DRX (500 MHz) spectrometer. Chemical shifts are denoted in δ (ppm) relative to residual solvent peaks as internal standard ($CDCl_3$, ^{1}H δ 7.24, ^{13}C δ 77.0). ESI-MS spectra were recorded on a Thermo Scientific Exactive mass spectrometer or a SHIMADZU LCMS-2020 spectrometer. High performance liquid chromatography (HPLC) experiments were performed with a SHIMADZU HPLC system equipped with a LC-20AD intelligent pump. LC-MS experiments were performed with amaZon SL-NPC (Bruker Daltonics). All reagents were used as supplied unless otherwise stated.

Animal Material Marine Sponge Theonella swinhoei was collected by scuba diving in Kodingareng Keke Island near Makassar, Indonesia in August 2015. The specimen was immediately frozen after collection and kept until processed.

Extraction and Isolation Approximately 1.0 kg (wet weight) of marine sponge Theonella swinhoei was extracted three times with 1 L methanol and concentrated under vacuo. The crude methanolic extract (13.0 g) was partitioned between ethyl acetate (300 mL x 3 times) and water (300 mL) to afford ethyl acetate soluble material (6.0 g). Subsequently, non-polar portion (ethyl acetate soluble material) was subjected to gel filtration chromatography (Sephadex LH-20) with methanol as mobile phase to give 160 fractions. Furthermore, fractions containing target compound (fr.62-78) were pooled and concentrated under vacuo. The residue (2.70 g) was further separated by silica gel open column chromatography [Silica gel 60N (spherical, neutral, 40–50 μm), Kanto Chemical Co., Inc.] using step gradient of mobile phase chloroform : methanol; [0%, 2.5%,5%, 7.25%, 10% (% methanol)], and washed with chloroform : methanol : water (4:5:1) to afford 112 fractions. Finally, fractions containing target compounds were purified through high pressure liquid chromatography equipped with Cosmosil® $5C_{18}$-MS-II (10 mm ID × 250 mm) using 75% acetonitrile in water with additional 0.05% TFA to yield compounds **1** (4.0 mg), **2** (9.0 mg), **3** (3.0 mg), **4** (40.0 mg), and **5** (8.0 mg). In addition, a major theonellapeptolide Id (**6**) was also purified from different fractions and its spectra was compared to the previously reported [9, 19].

Theonellapeptolide IIb (**1**): Colorless solid; $[\alpha]^{25}_{D}$ −22.9622 (c 0.31, MeOH); ^{1}H and ^{13}C NMR data are shown in Table S2.11; HRESIMS *m/z* 1398.89294 $[M + Na]^{+}$ (calcd. for $C_{68}H_{121}N_{13}O_{16}Na$, 1398.89519). IR: ν_{max} 3300, 2960, 1630, 1520 cm^{-1}.

Theonellapeptolide IIa (**2**): Colorless solid; $[\alpha]^{22}_{D}$ −35.500 (c 0.7, MeOH); ^{1}H and ^{13}C NMR data are shown in Table S2.12; HRESIMS *m/z* 1384.88153 $[M + Na]^{+}$ (calcd. for $C_{67}H_{119}N_{13}O_{16}Na$, 1384.87954). IR: ν_{max} 3300, 2960, 1630, 1530 cm^{-1}.

Theonellapeptolide IIc (**3**): Colorless solid; $[\alpha]^{23}_{D}$ −55.1241 (c 1.45, MeOH); ^{1}H and ^{13}C NMR data are shown in Table S2.13; HRESIMS *m/z* 1412.91000 $[M + Na]^{+}$ (calcd. for $C_{69}H_{123}N_{13}O_{16}Na$, 1412.91084). IR: ν_{max} 3300, 2960, 1630, 1520 cm^{-1}.

Determination of Amino Acid Sequence by Tandem MS Analysis Amino acid sequence of compounds **1**-**5** was determined by means of tandem mass analysis of *seco*-acid methyl ester peptides. Each compound (100 µg) was treated with ammonia solution 7N in methanol at 100 °C for 1 hour and concentrated in vacuo. The obtain products were subjected to LC-MS/MS analysis equipped with Inertsil ODS-4 (3.0 ID × 150 mm) using a MeCN-H_2O gradient 10% to 100% over 10 minutes and keeping 100% MeCN for 10 min, with 0.1% formic acid.

Total Hydrolysis As much as 100 µg of compounds **1**-**5** was hydrolyzed with 300 µL of HCl 6M at 110 °C for 24 hours. The hydrolysates were concentrated under inert gas and used for starting materials for amino acid analysis by Marfey's method.

Partial Hydrolysis As much as 0.5 mg of compounds **1**-**5** was treated with 300 µL of 30% TFA for 40 minutes at 110 °C in sealed tube. The residues were concentrated, and the target fragments were isolated through preparative HPLC equipped with Cosmosil® $5C_{18}$-AR using a MeCN-H_2O gradient from 10% to 100% over 70 minutes and hold 100% MeCN for 15 minutes. After confirmation using tandem mass analysis, the target fragments were further hydrolyzed with 300 µL of HCl 6M at 110 °C for 24 hours. The hydrolysates were concentrated under inert gas and used for starting materials for amino acid analysis by Marfey's method.

Amino Acid Analysis by Marfey's Method The obtained partial and total hydrolysates were individually mixed with H_2O (160 µL), saturated $NaHCO_3$ (200 µL), and 1% FDAA in acetone (160 µL). The mixture was heated at 40 °C for 30 minutes. Subsequently, the reaction was quenched with 1M HCl (200 µL) and concentrated under vacuo. The residue was dissolved with DMSO (400 µL) and passed through membrane filter before injecting to LC-MS, Cosmosil® $5C_{18}$-MS-II (2 mm ID × 100 mm) (oven 40 °C ; flowrate 0.2 ml $min^{-}1$) using gradient program from 10% to 60% MeCN in water (formic acid 0.1%) for 45 minutes.

Antibacterial Activity Compounds **1**-**6** were tested for antibacterial activity. The assay for all compounds has been done in triplicate according to the Clinical Laboratory Standards Institute testing standard in a 96-well plate microbroth dilution assay. Compounds were tested against gram negative-bacteria (*Escherichia coli* JW5503) and gram-positive bacteria (*Bacillus cereus* NBRC 15305 and *Kocuria rhizophila* NBRC 12708). After incubation for 22 h (30 °C), the obtained optical density was recorded using plate reader at 600 nm to determine percent growth inhibition.

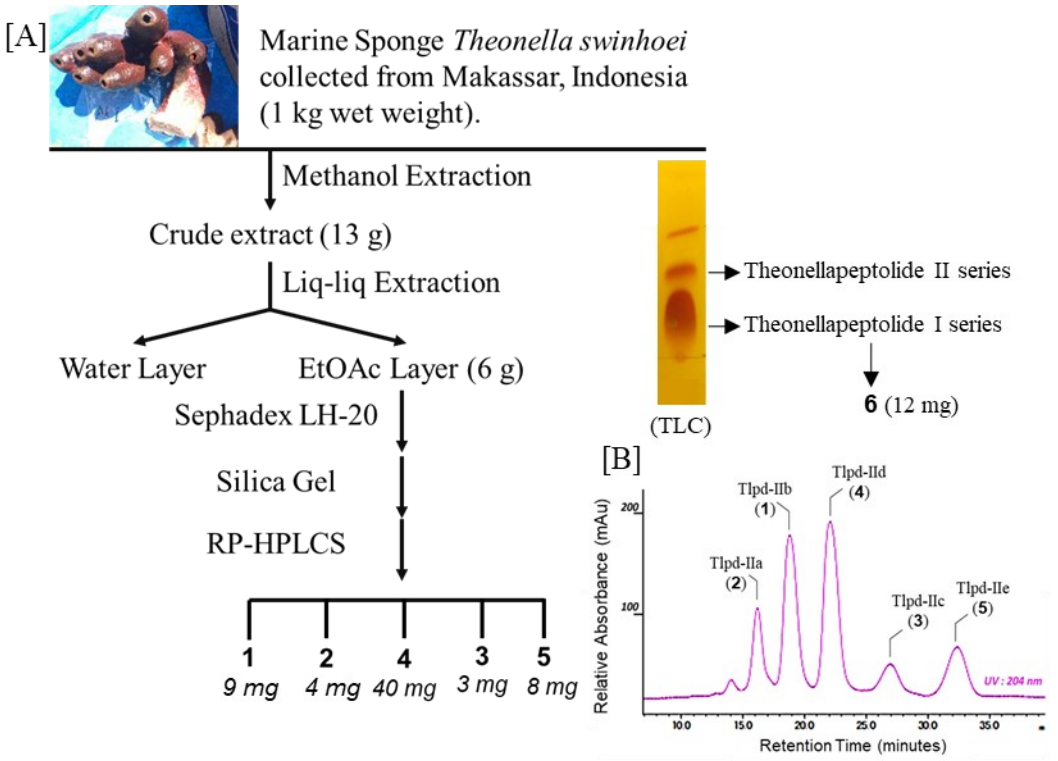


Figure S2.1 Isolation and purification of five theonellapeptolide-type compounds (**1**, **2**, **3**, **4**, and **5**) from Kondingarengan *Theonella swinhoei*. Thin Layer Chromatography (TLC) analysis of the sponge methanol extract using dragendorff's reagent showed three brown spots. Secondary metabolites corresponding to the second spot were purified, resulting in the isolation of five different compounds [A] HPLC profile of five purified compounds **1**-**5**, in which two of them (**4** and **5**) are known theonellapeptolide IId and IIe, respectively. [B] Three new analogues (IIa, IIb, and IIc) were subsequently characterized to determine their chemical structures. In addition, a major theonellapeptolide Id (**6**) was purified from the first spot.

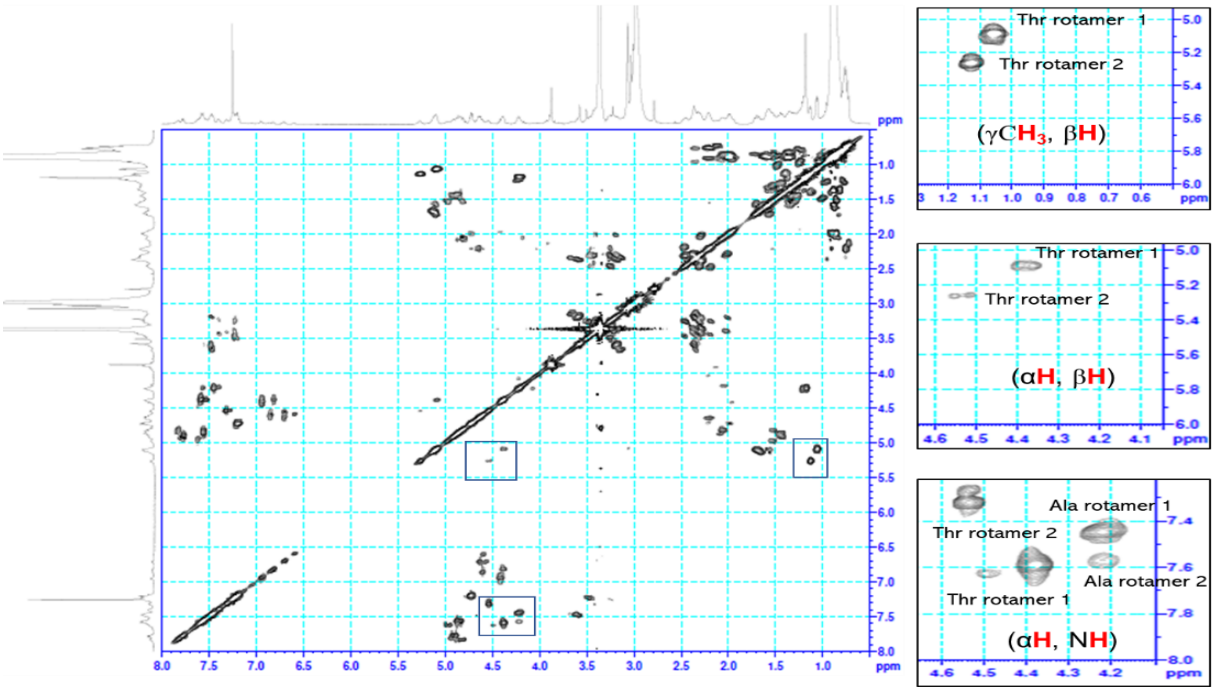


Figure S2.2 ^{1}H-^{1}H COSY spectrum (500 MHz) of compound **1** in $CDCl_3$ showing threonine and alanine conformers.

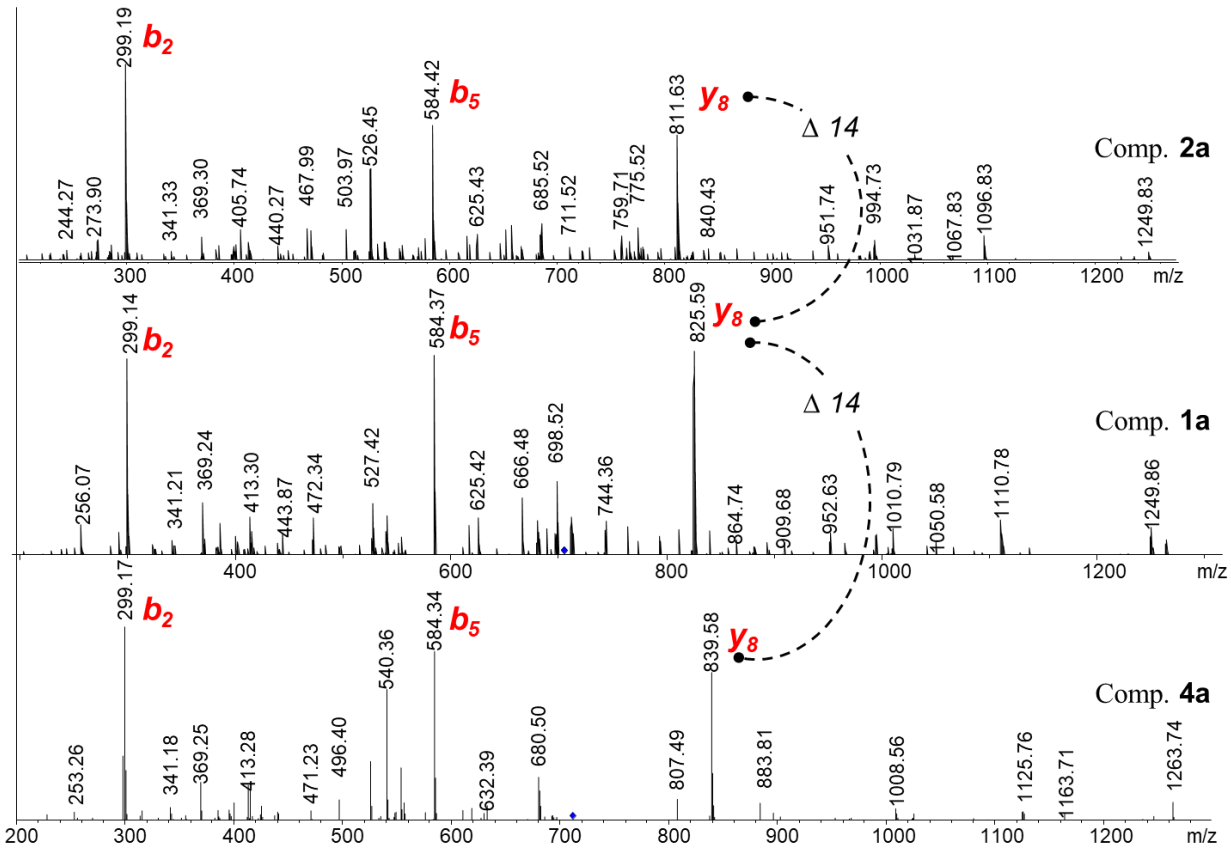


Figure S2.3 Tandem MS spectra of the *seco*-acid methyl ester peptides **1a** and **2a** compared with originally from known theonellapeptolide IId (**4a**).

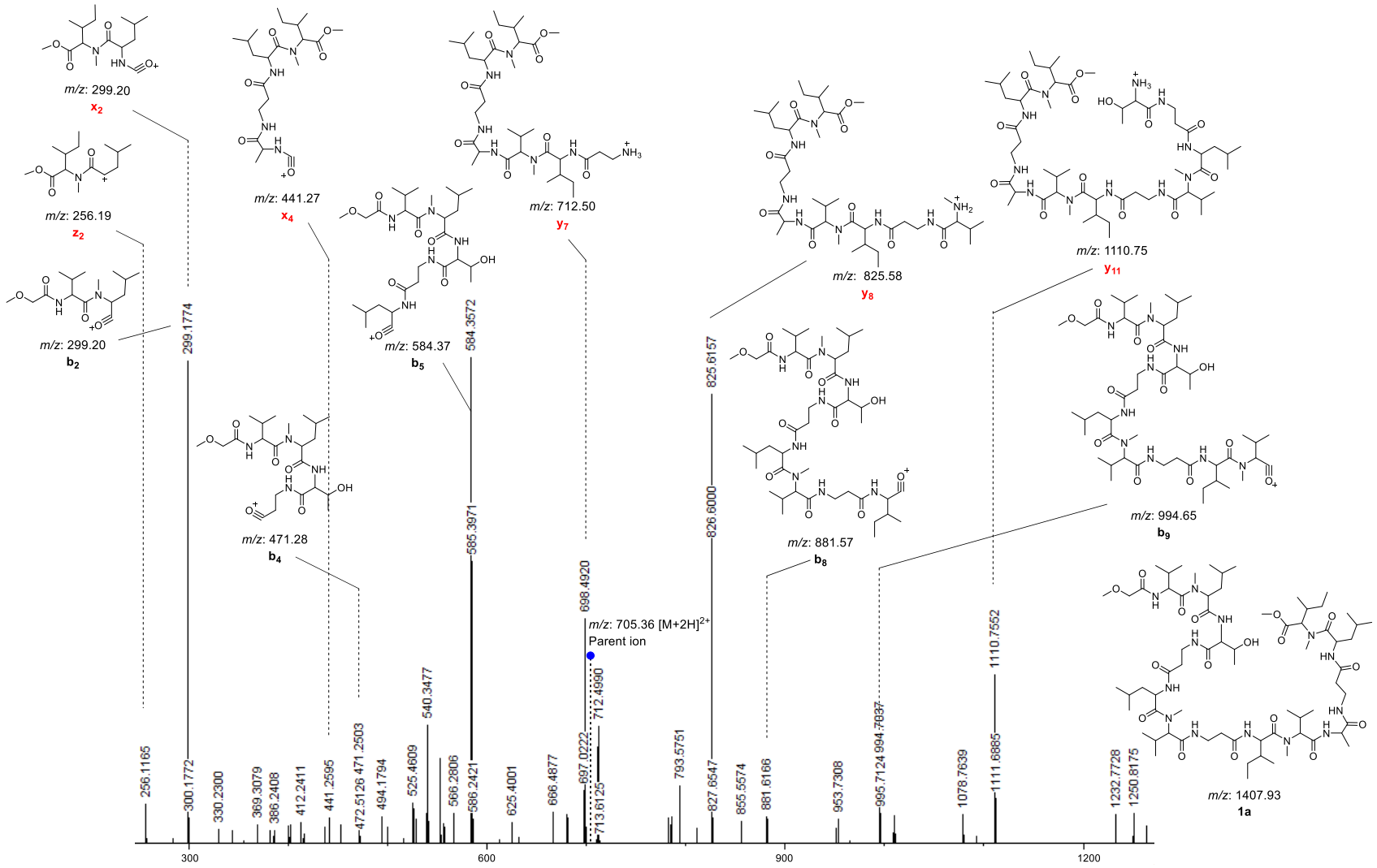


Figure S2.4 MS/MS spectrum of *seco*-acid methyl ester peptide **1a** *m/z* 705.36 $[M+2H]^{2+}$, (red colored letter denoted fragment ions generated toward C-terminus, black colored letter denoted fragment ions toward N-terminus). The other detected fragment ions are listed in Table S2.1.

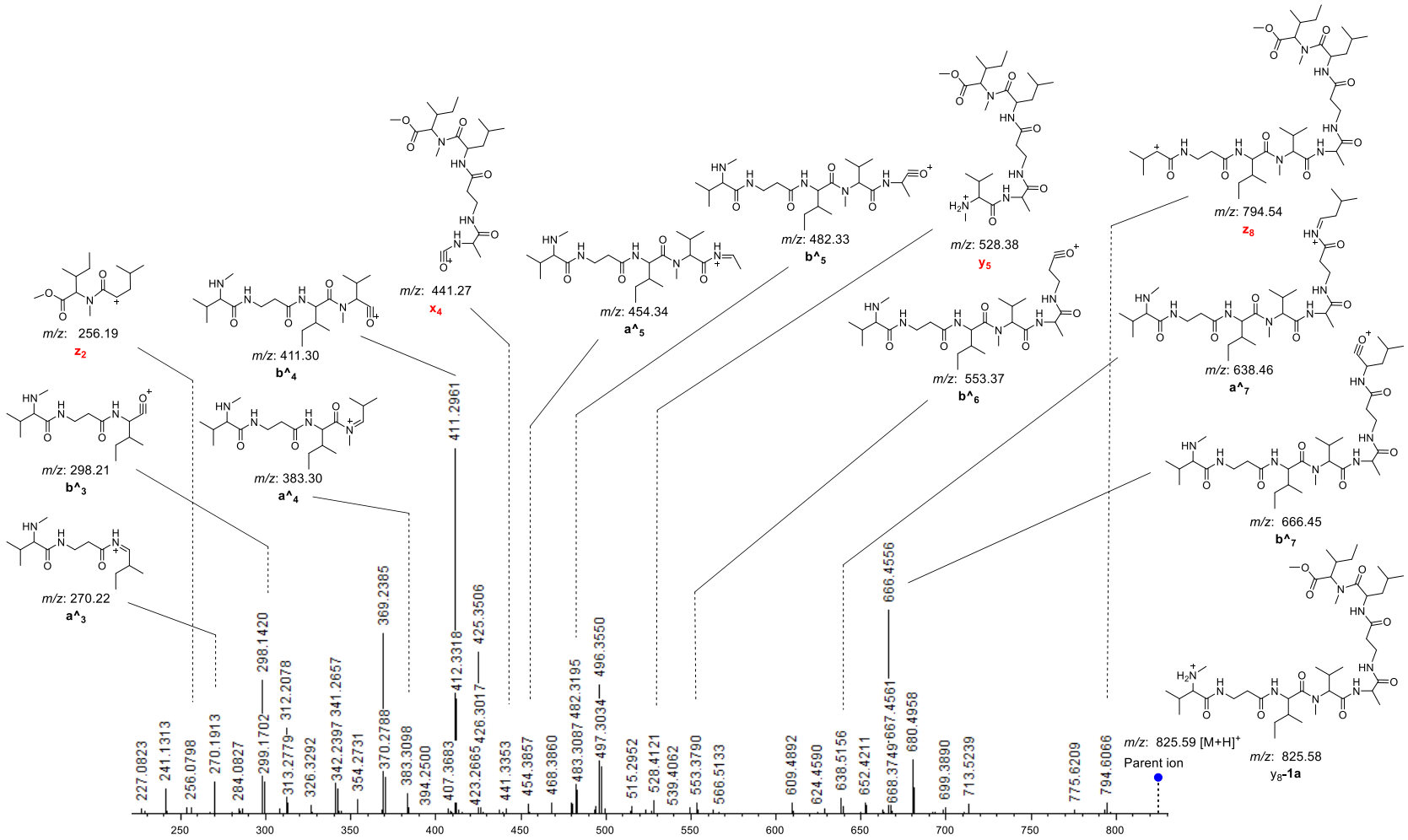


Figure S2.5 MS/MS/MS spectrum of fragment ion y_8 of **1a**, *m/z* 825.59 $[M+H]^+$ (red colored letter denoted fragment ions generated toward C-terminus, black colored letter denoted fragment ions toward N-terminus). The other detected fragment ions are listed in Table S2.2.

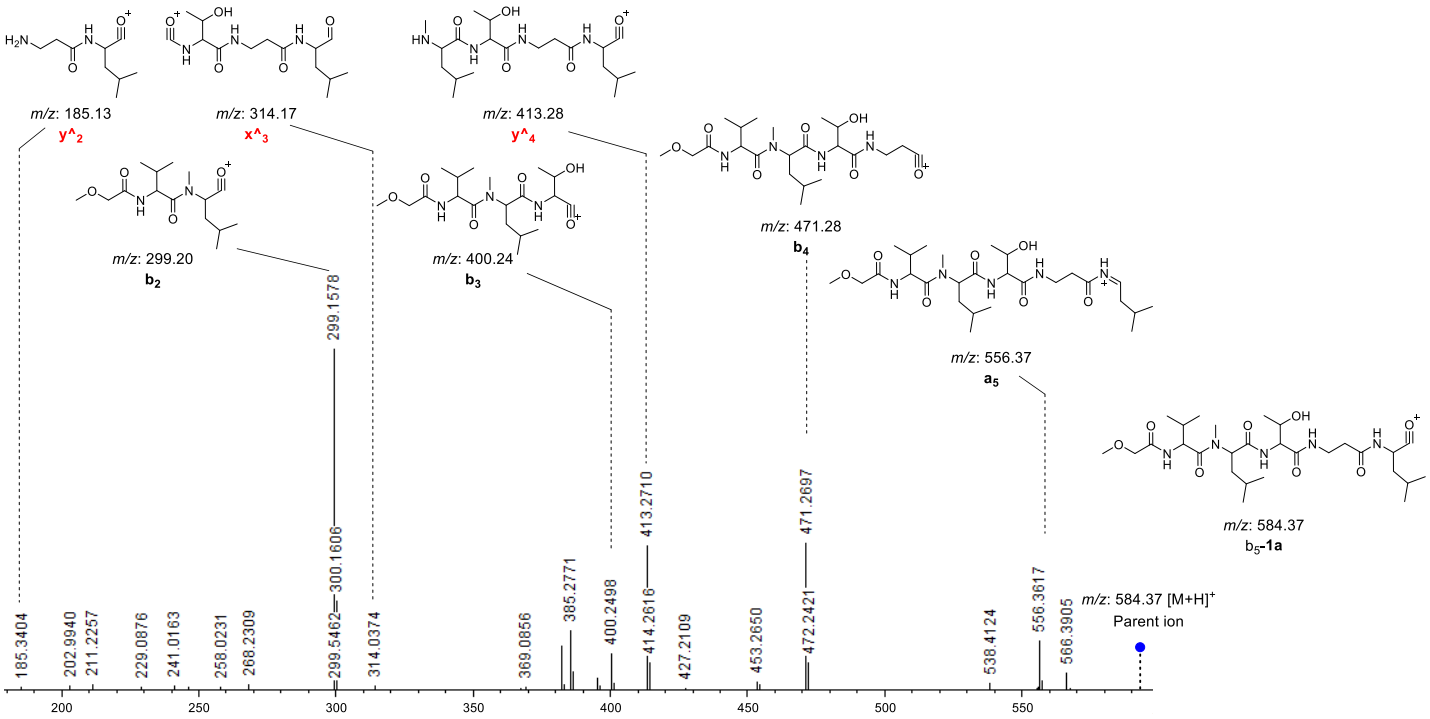


Figure S2.6 MS/MS/MS spectrum of fragment ion b_5 of **1a**, *m/z* 584.37 $[M+H]^+$ (red colored letter denoted fragment ions generated toward C-terminus; black colored letter denoted fragment ions toward N-terminus).

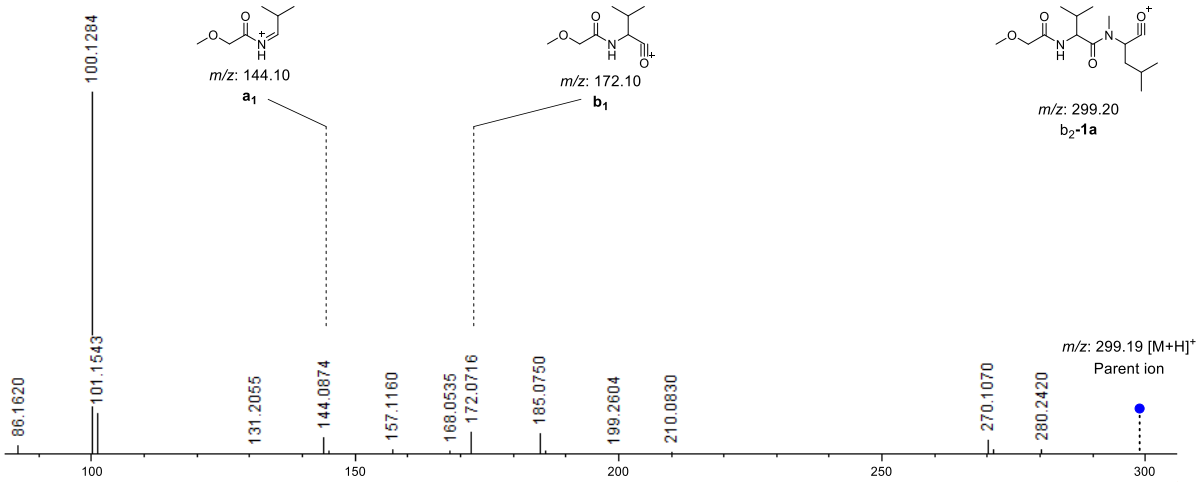


Figure S7 MS/MS/MS spectrum of fragment ion b_2 of **1a**, *m/z* 299.19 $[M+H]^+$ (fragment ions toward N-terminus).

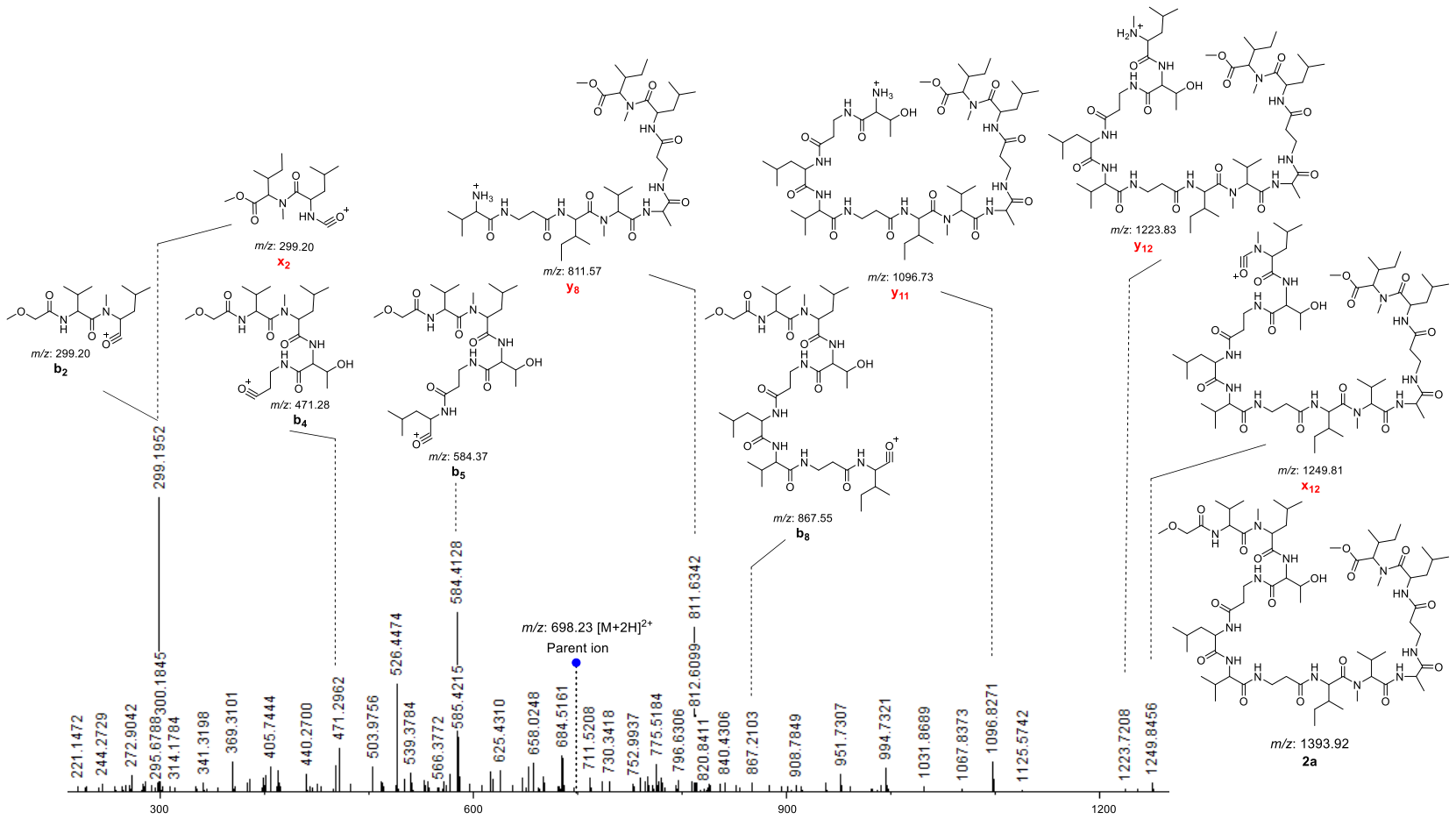


Figure S2.8 MS/MS spectrum of *seco*-acid methyl ester peptide **2a**, *m/z* 698.23 $[M+2H]^{2+}$, (red colored letter denoted fragment ions generated toward C-terminus, black colored letter denoted fragment ions toward N-terminus). The other detected fragment ions are listed in Table S2.3.

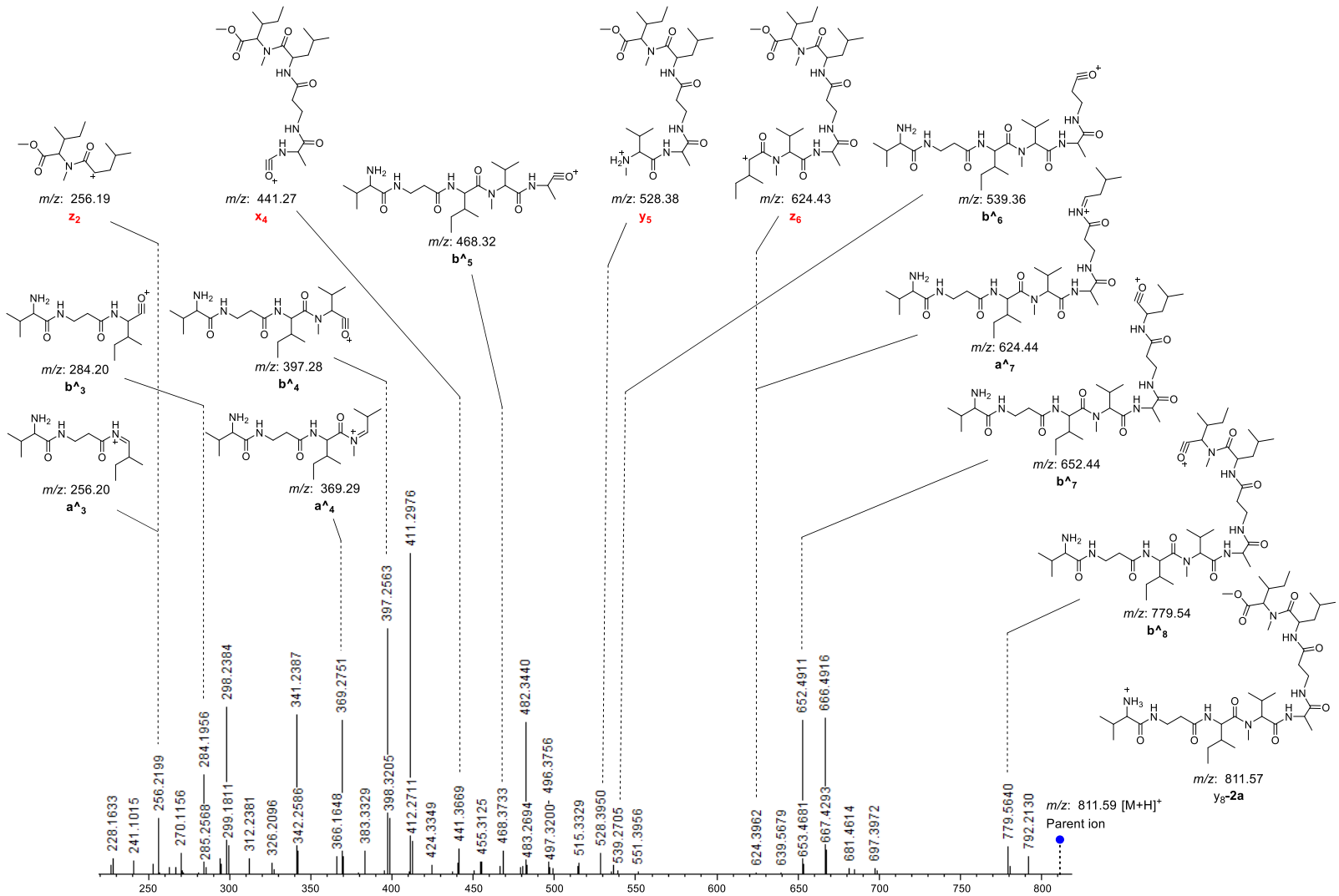


Figure S2.9 MS/MS/MS spectrum of fragment ion y_8 of **2a**, *m/z* 811.59 $[M+H]^+$, (red colored letter denoted fragment ions generated toward C-terminus, black colored letter denoted fragment ions toward N-terminus). The other obtained fragment ions are listed in Table S2.4.

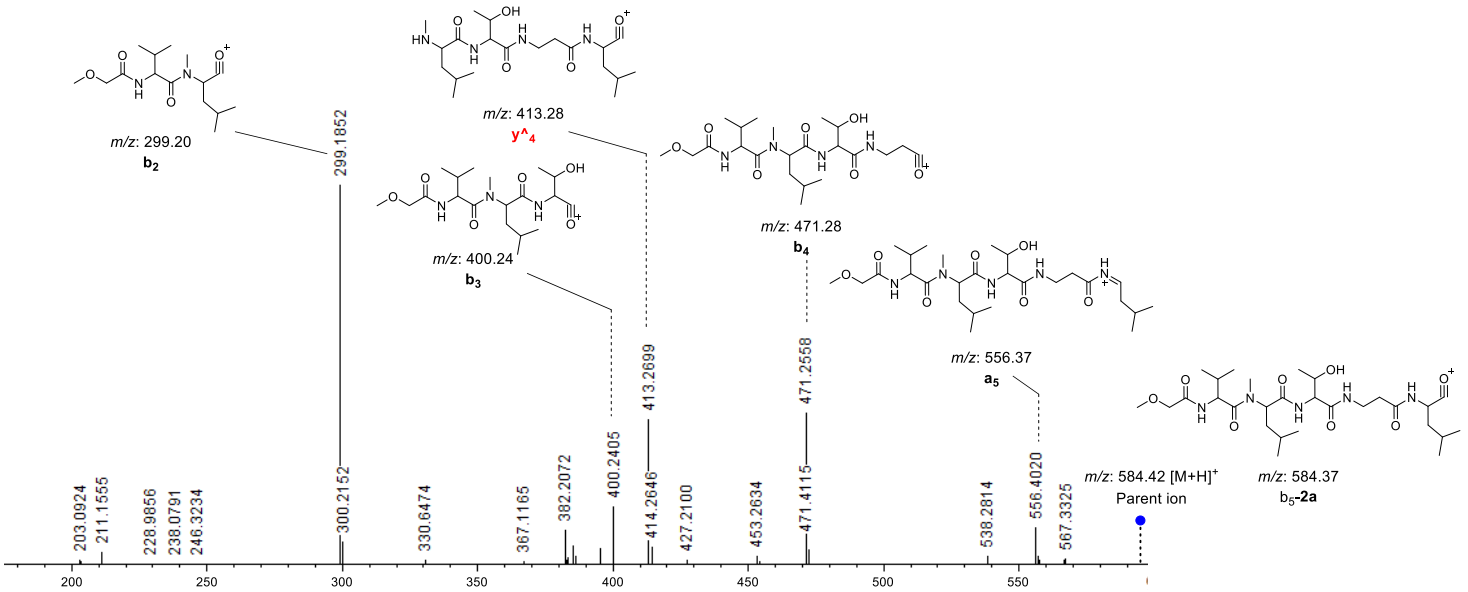


Figure S2.10 MS/MS/MS spectrum of fragment ion b5 of **2a**, *m/z* 584.42 $[M+H]^+$, (red colored letter denoted fragment ions generated toward C-terminus, black colored letter denoted fragment ions toward N-terminus).

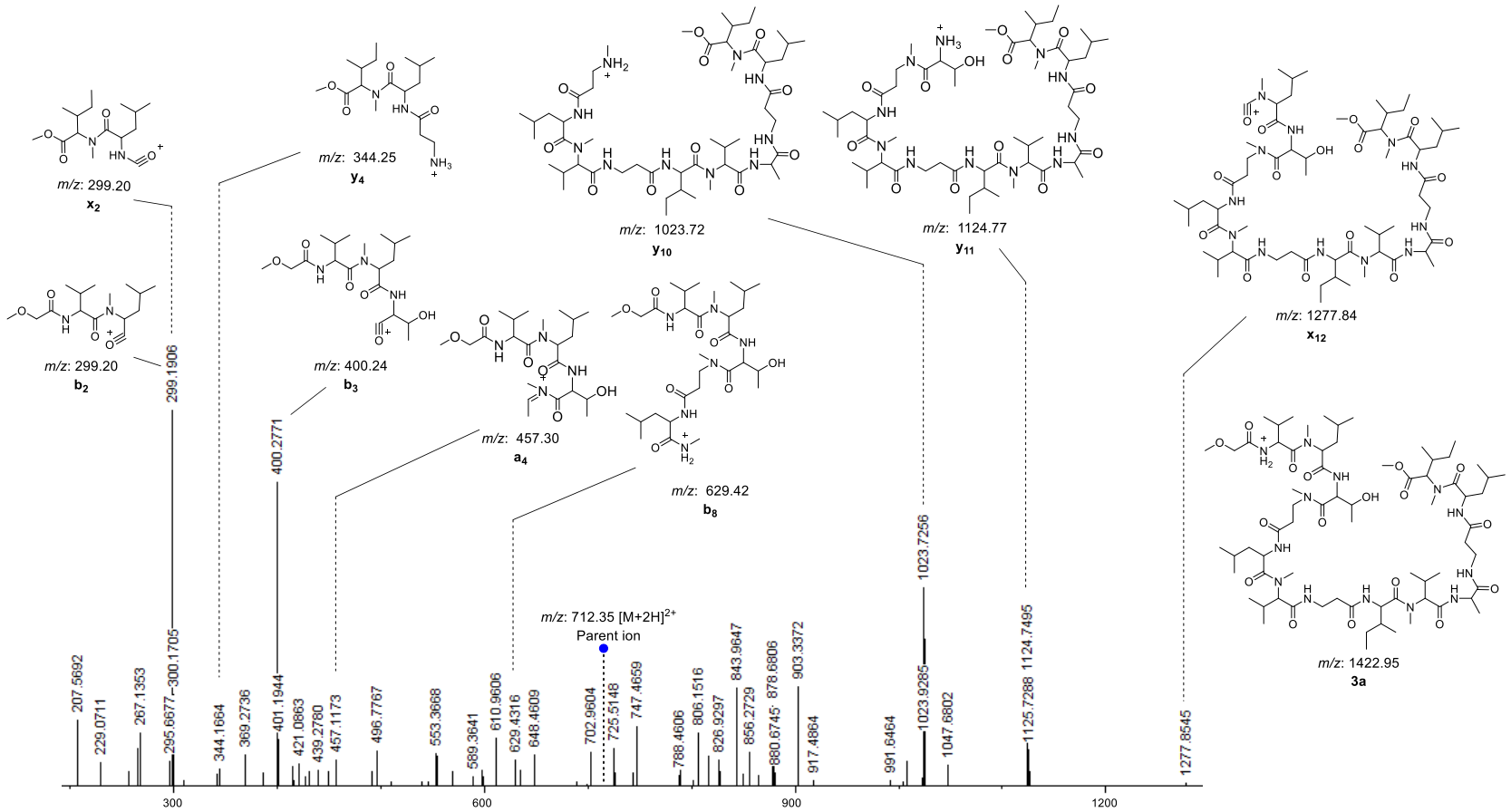


Figure S2.11 MS/MS spectrum of *seco*-acid methyl ester peptide **3a**, *m/z* 712.35 $[M+2H]^{2+}$, (red colored letter denoted fragment ions generated toward C-terminus, black colored letter denoted fragment ions toward N-terminus). The other detected fragment ions are listed in Table S2.5.

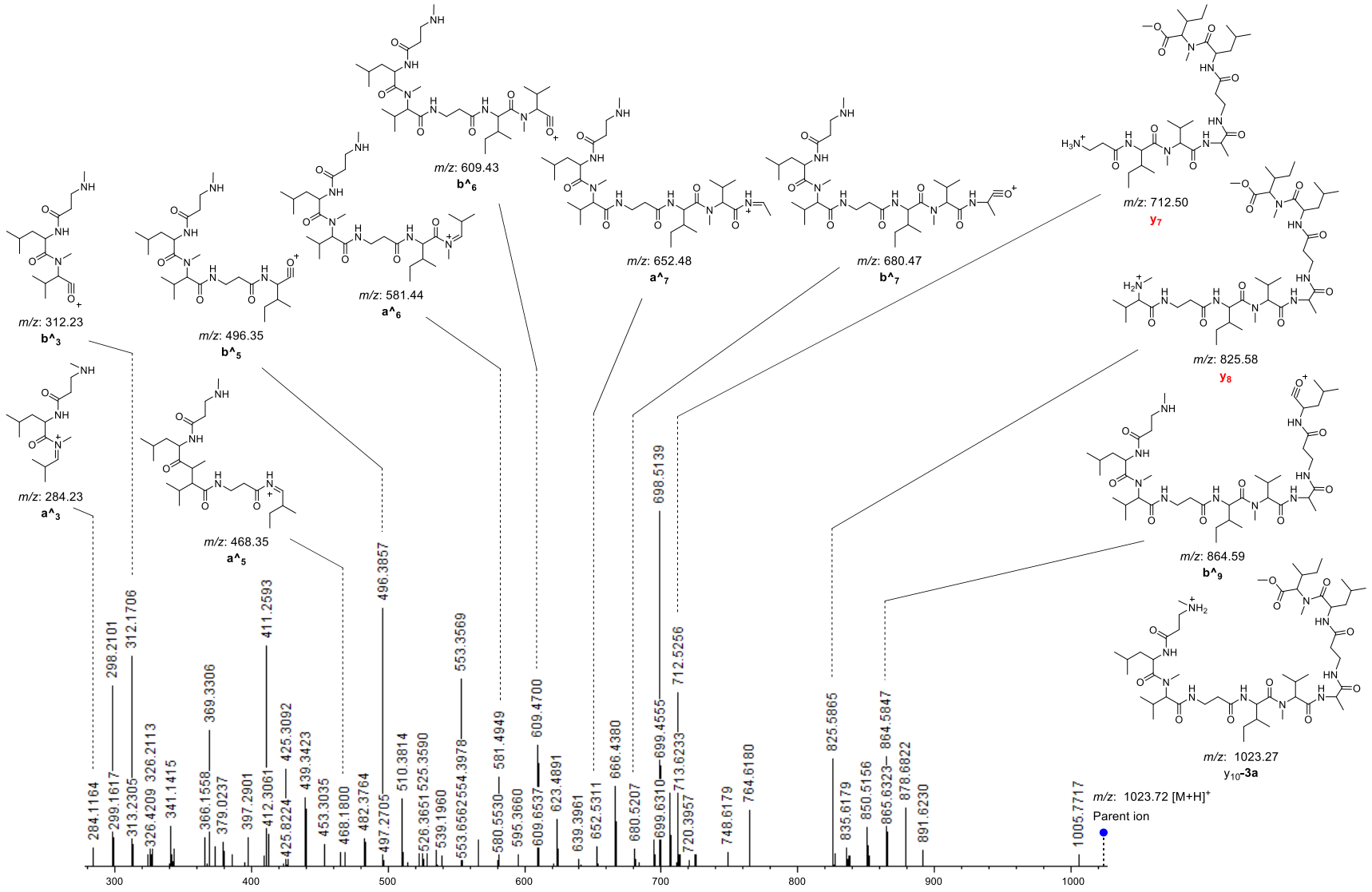


Figure S2.12 MS/MS/MS spectrum of fragment ion y_{10} of **3a**, *m/z* 1023.27 $[M+H]^+$, (red colored letter denoted fragment ions generated toward C-terminus, black colored letter denoted fragment ions toward N-terminus). The other detected fragment ions are listed in Table S2.6.

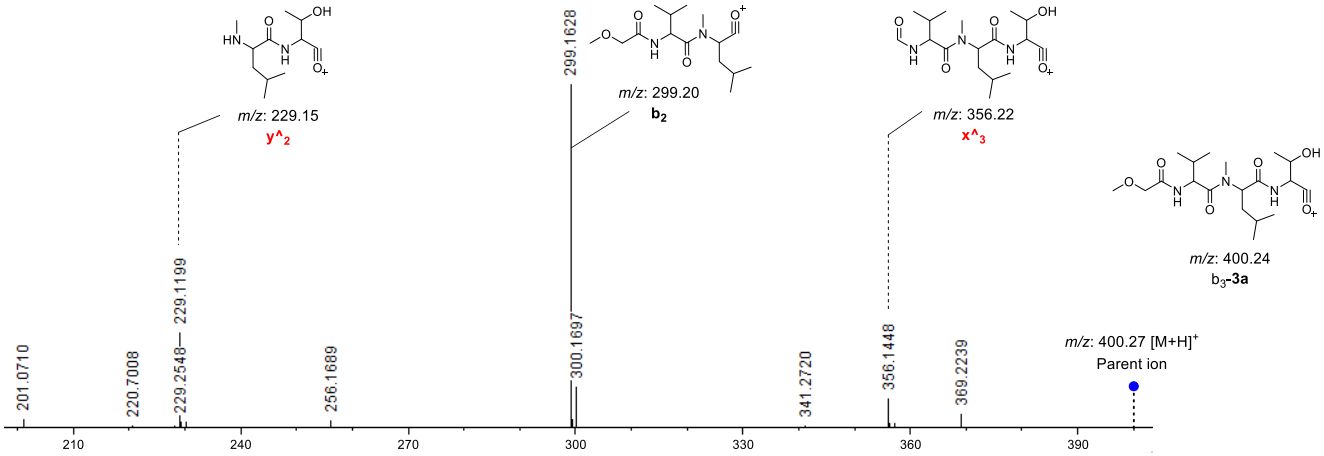


Figure S2.13 MS/MS/MS spectrum of fragment ion b_3 of **3a**, *m/z* 400.42 $[M+H]^+$, (red colored letter denoted fragment ions generated toward C-terminus, black colored letter denoted fragment ions toward N-terminus).

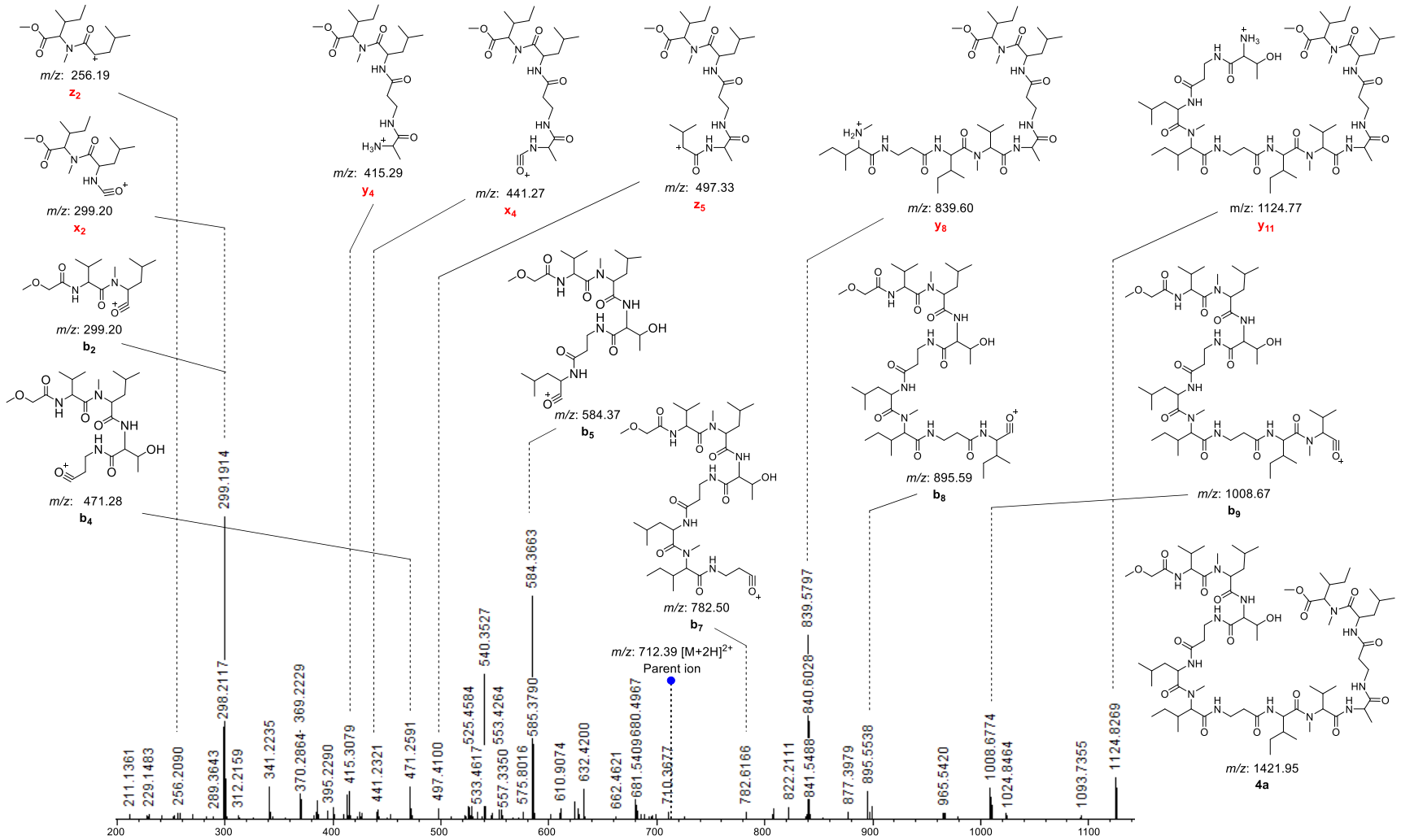


Figure S2.14 MS/MS spectrum of *seco*-acid methyl ester peptide **4a**, *m/z* 712.39 $[M+2H]^{2+}$, (red colored letter denoted fragment ions generated toward C-terminus, black colored letter denoted fragment ions toward N-terminus). The other detected fragment ions are listed in Table S2.7.

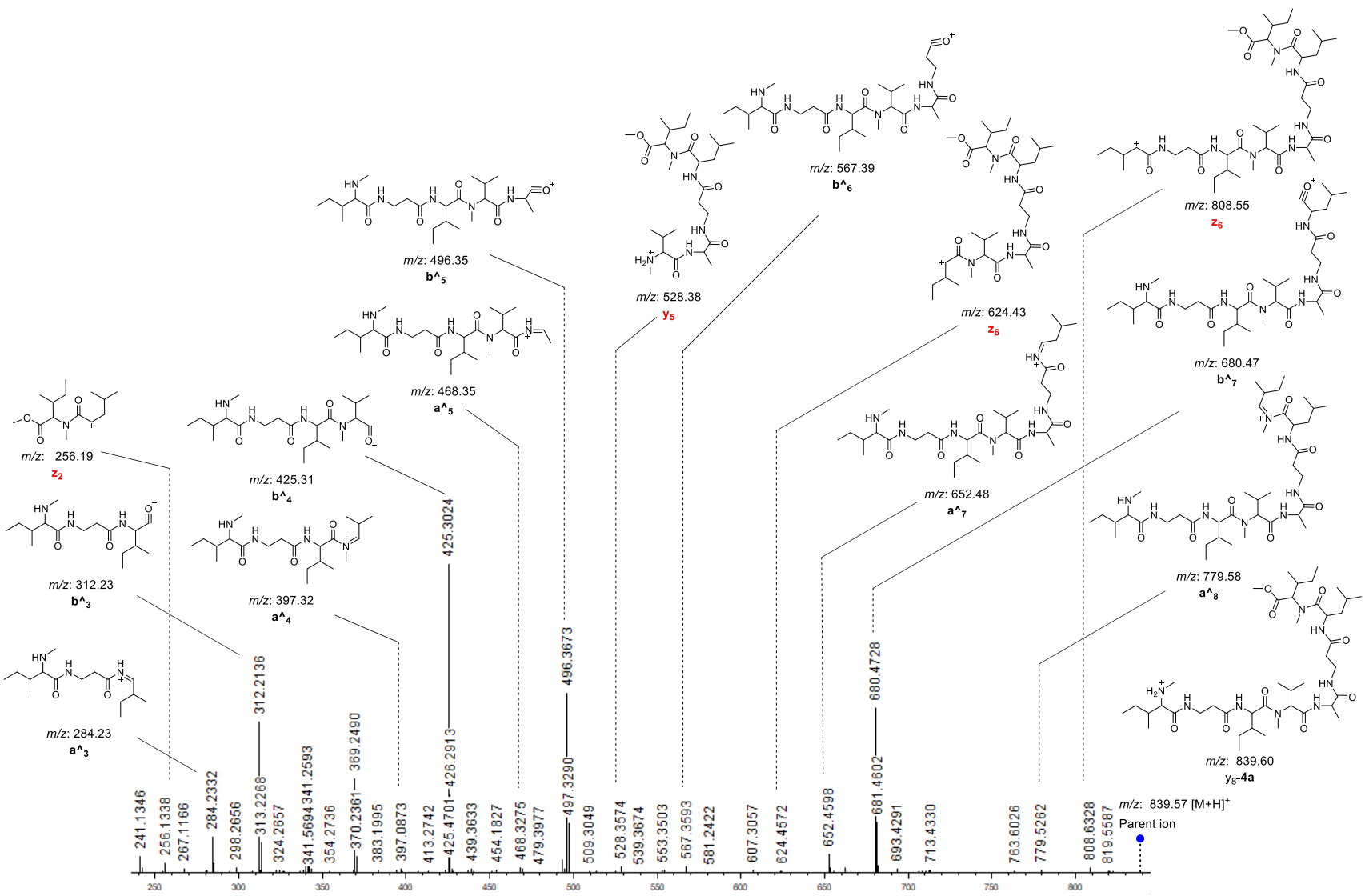


Figure S2.15 MS/MS/MS spectrum of fragment ion y_8 of **4a**, *m/z* 839.57 $[M+H]^+$, (red colored letter denoted fragment ions generated toward C-terminus, black colored letter denoted fragment ions toward N-terminus). The other detected fragment ions are listed in Table S2.8.

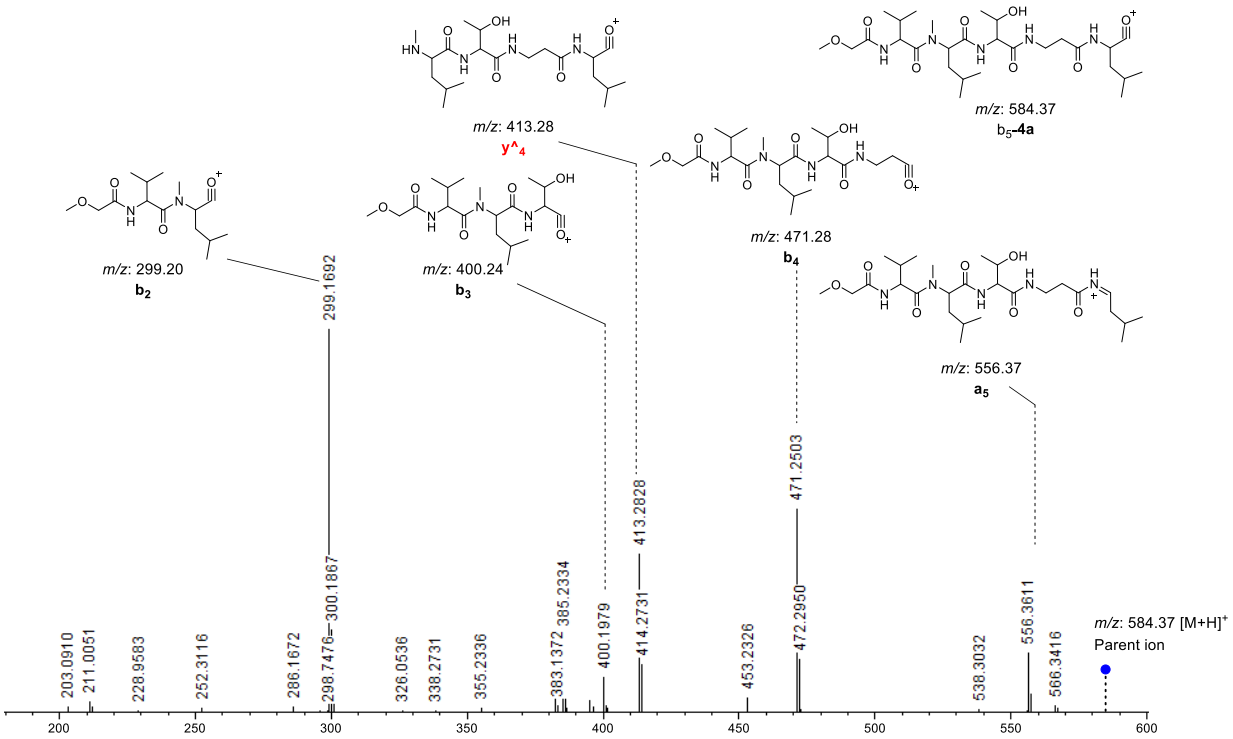


Figure S2.16 MS/MS/MS spectrum of fragment ion b_5 of **4a**, *m/z* 584.37 $[M+H]^+$, (red colored letter denoted fragment ions generated toward C-terminus, black colored letter denoted fragment ions toward N-terminus).

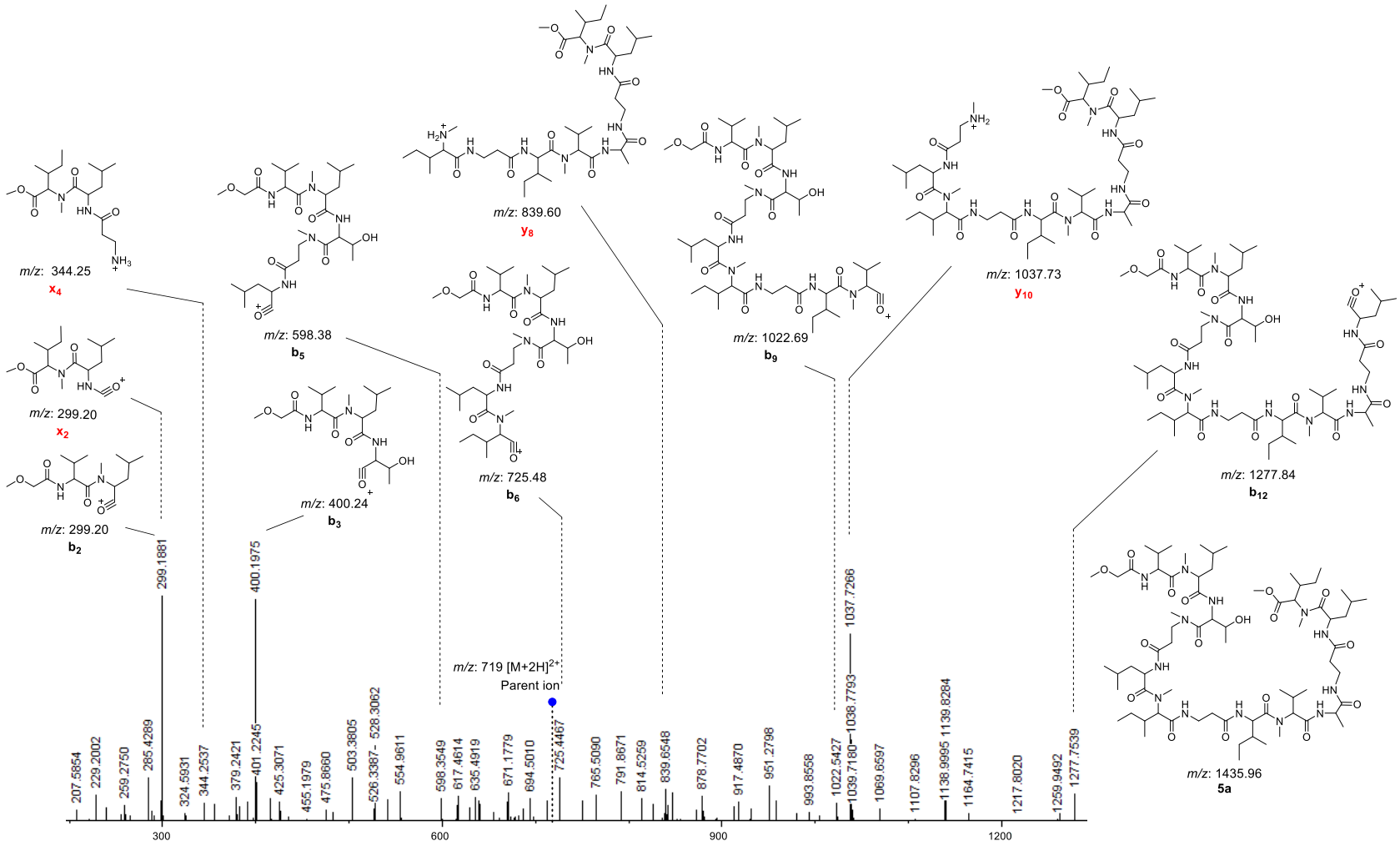


Figure S2.17 MS/MS spectrum of *seco*-acid methyl ester peptide **5a**, *m/z* 719 $[M+2H]^{2+}$, (red colored letter denoted fragment ions generated toward C-terminus, black colored letter denoted fragment ions toward N-terminus). The other detected fragment ions are listed in Table S2.9.

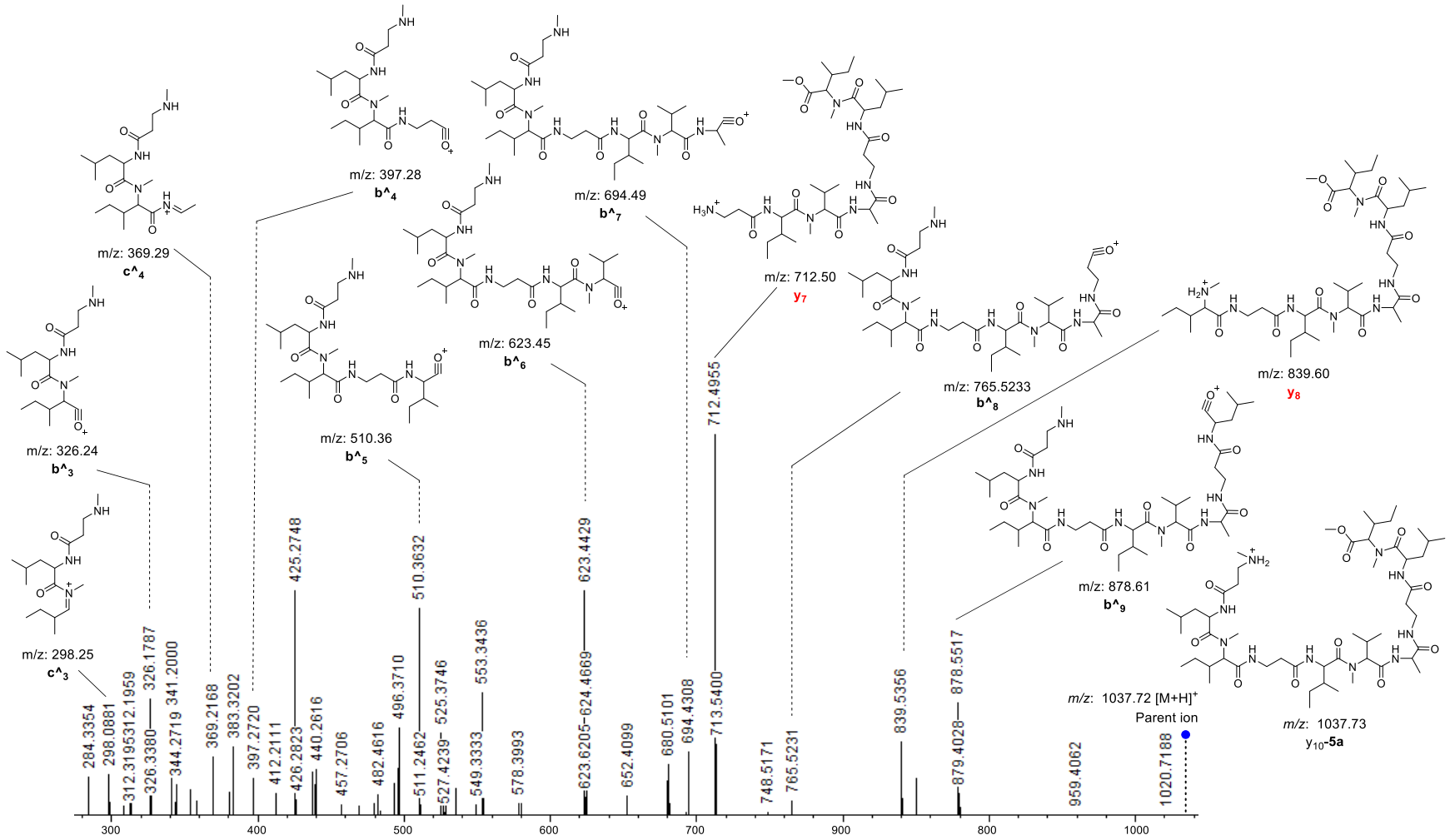


Figure S2.18 MS/MS/MS spectrum of fragment ion y_{10} of **5a**, *m/z* 1037.72 $[M+H]^+$, (red colored letter denoted fragment ions generated toward C-terminus, black colored letter denoted fragment ions toward N-terminus). The other detected fragment ions are listed in Table S2.10.

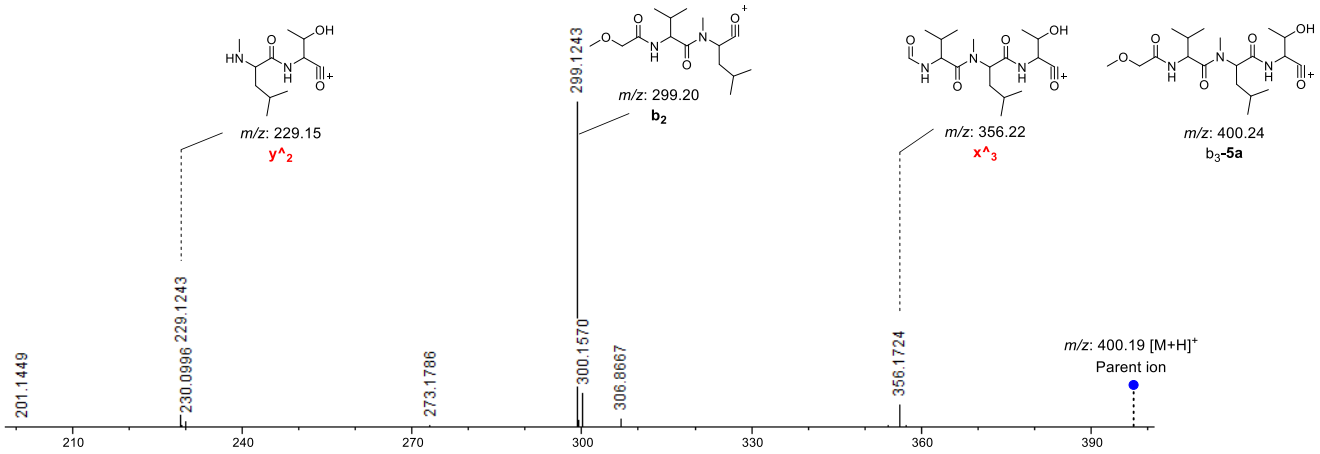


Figure S2.19 MS/MS/MS spectrum of fragment ion b3 of **5a**, *m/z* 400.19 $[M+H]^+$, (red colored letter denoted fragment ions generated toward C-terminus, black colored letter denoted fragment ions toward N-terminus).

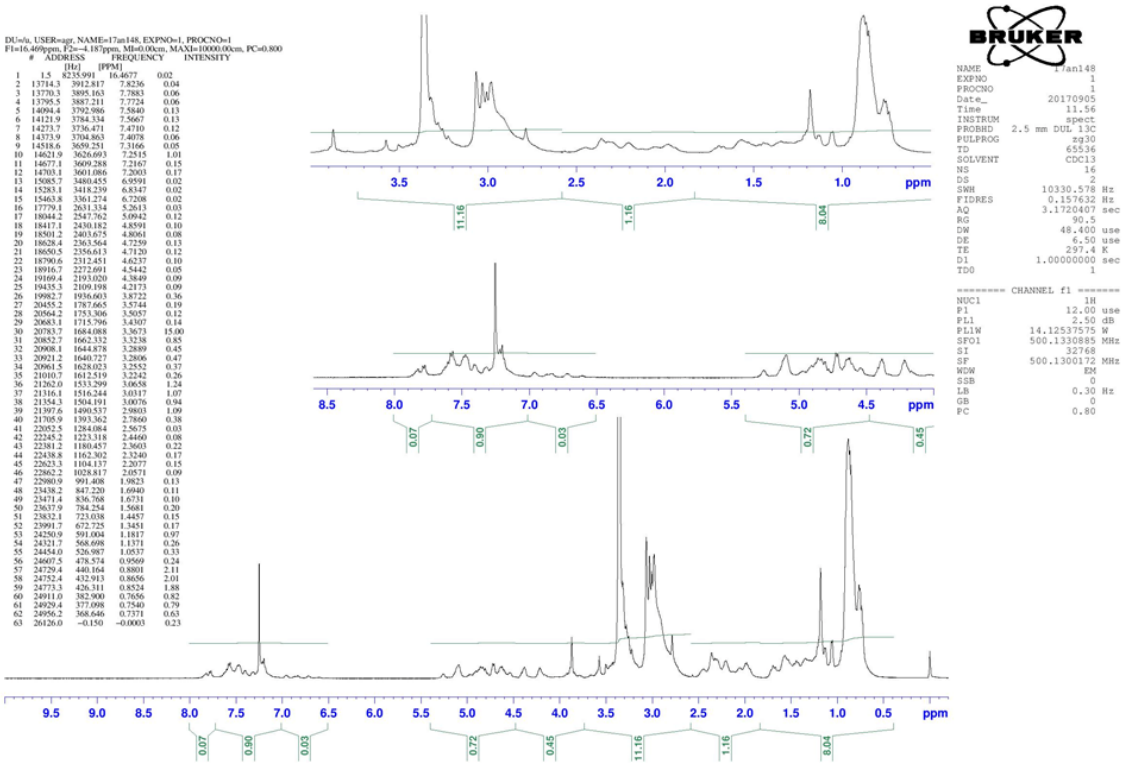


Figure S2.20 ^{1}H NMR spectrum ($CDCl_3$, 500 MHz) of compound **1**

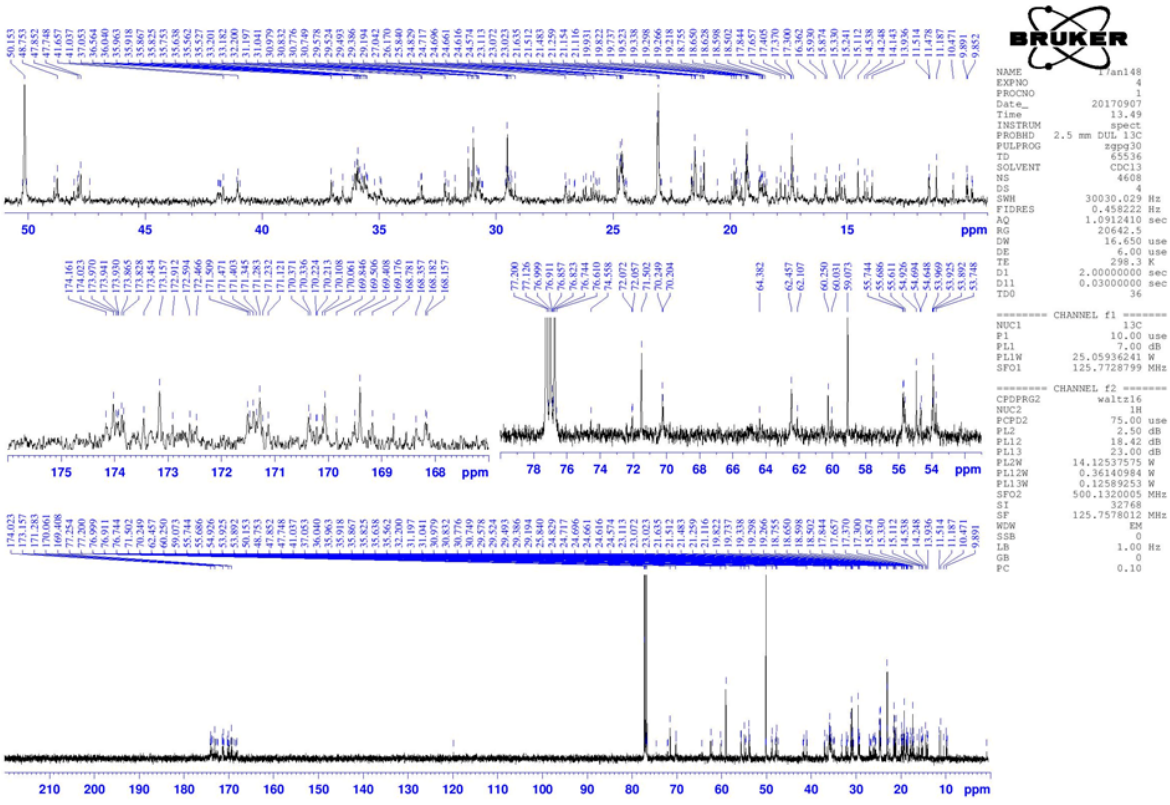


Figure S2.21 ^{13}C NMR spectrum ($CDCl_3$, 500 MHz) of compound **1**

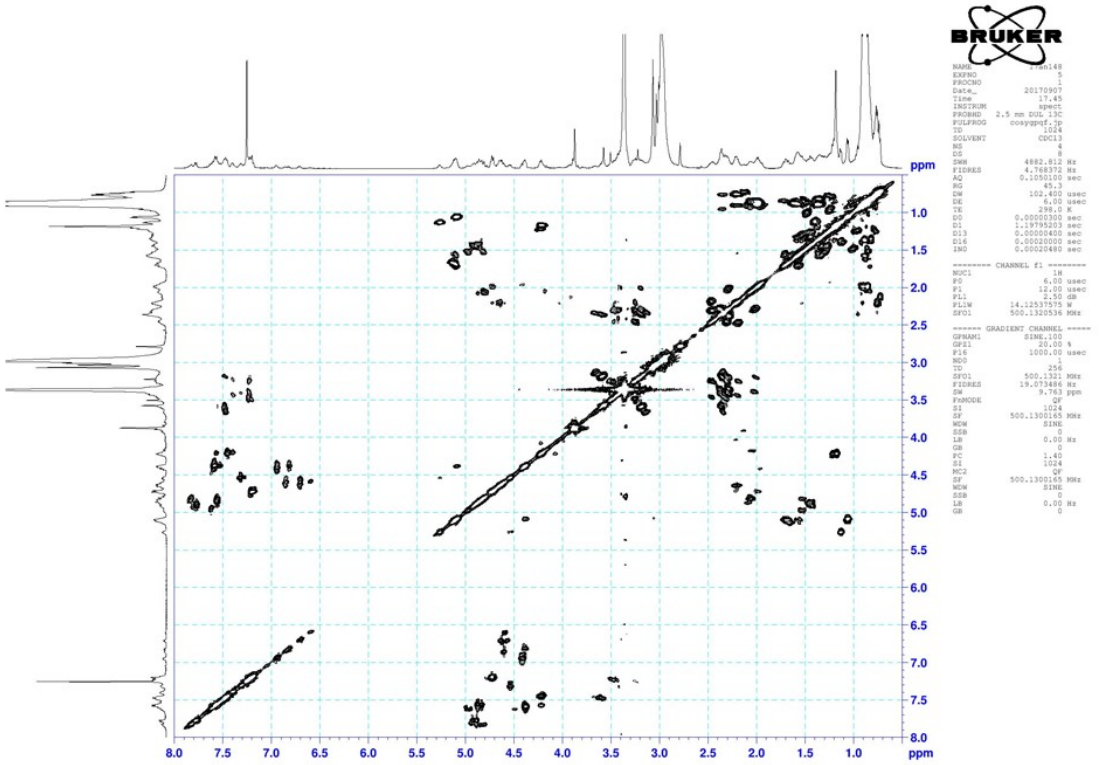


Figure S2.22 COSY NMR spectrum ($CDCl_3$, 500 MHz) of compound **1**

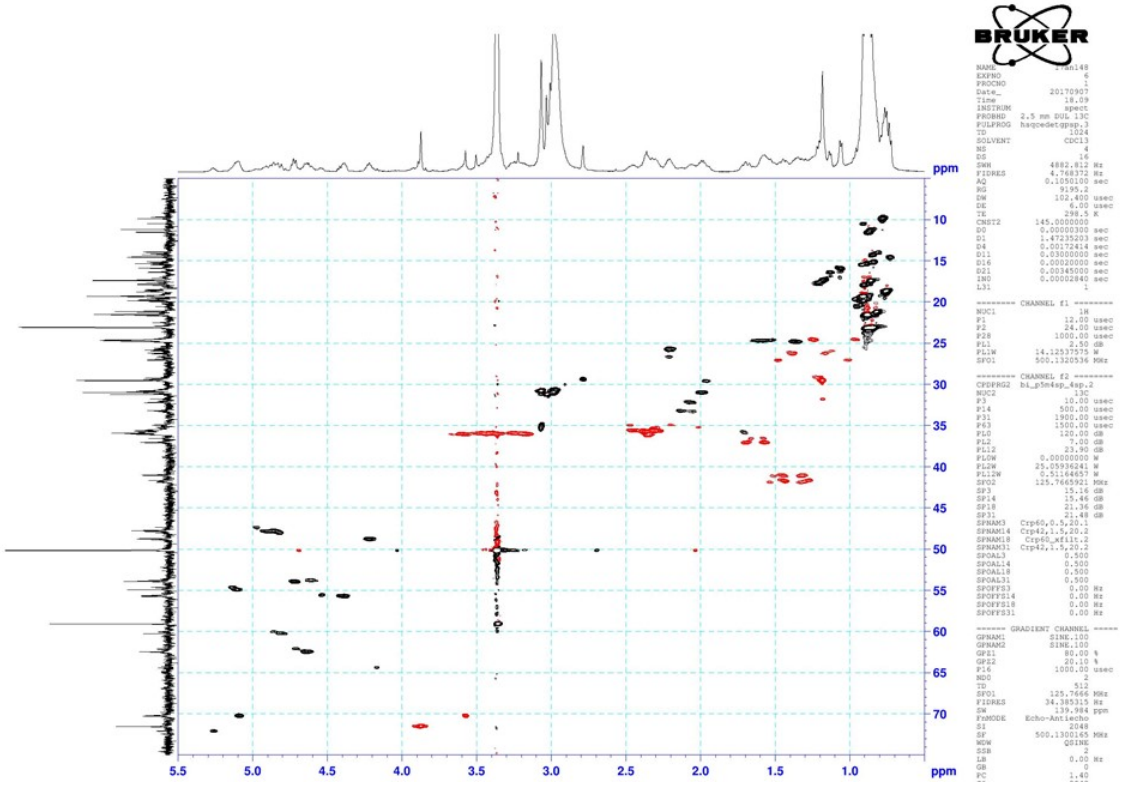


Figure S23 HSQC NMR spectrum ($CDCl_3$, 500 MHz) of compound **1**

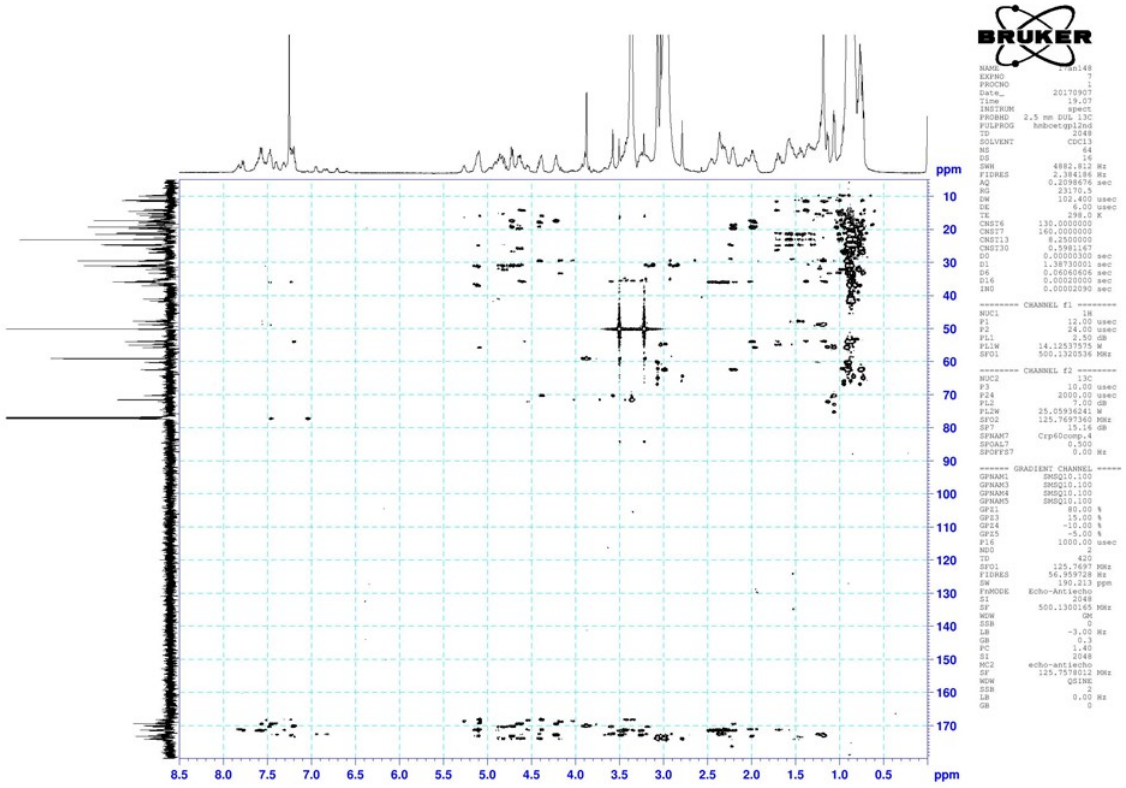


Figure S2.24 HMBC NMR spectrum ($CDCl_3$, 500 MHz) of compound **1**

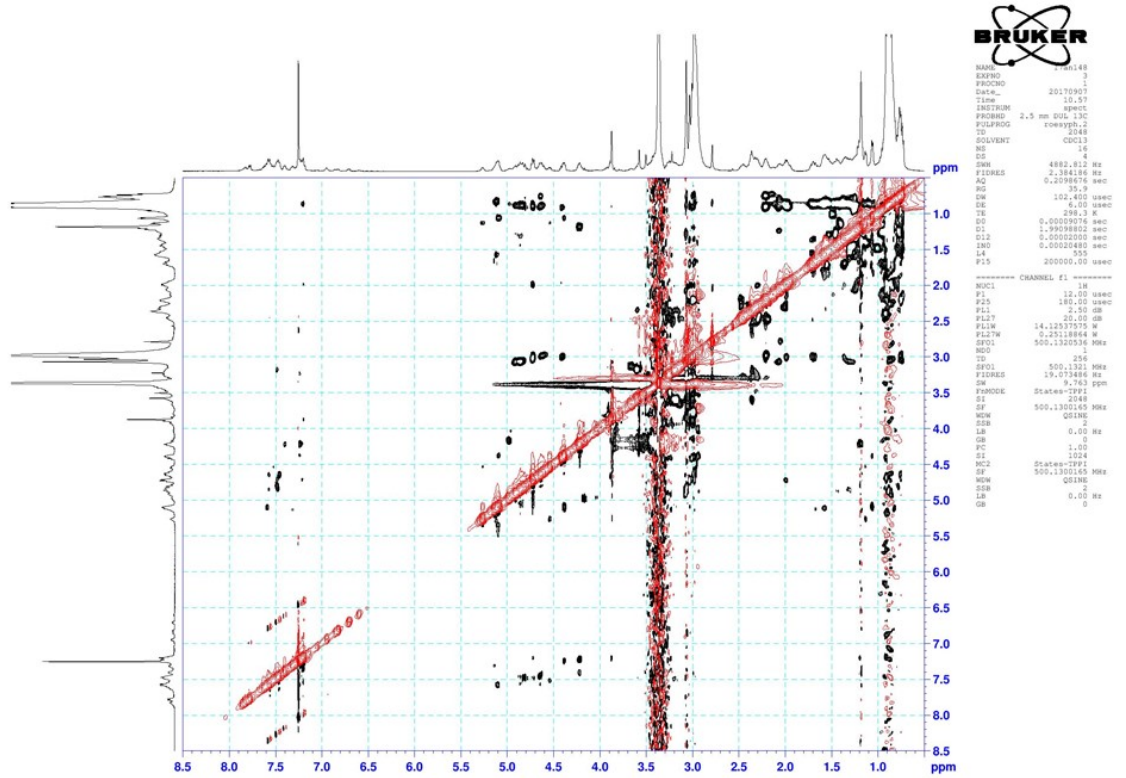


Figure S2.25 ROESY NMR spectrum ($CDCl_3$, 500 MHz) of compound **1**

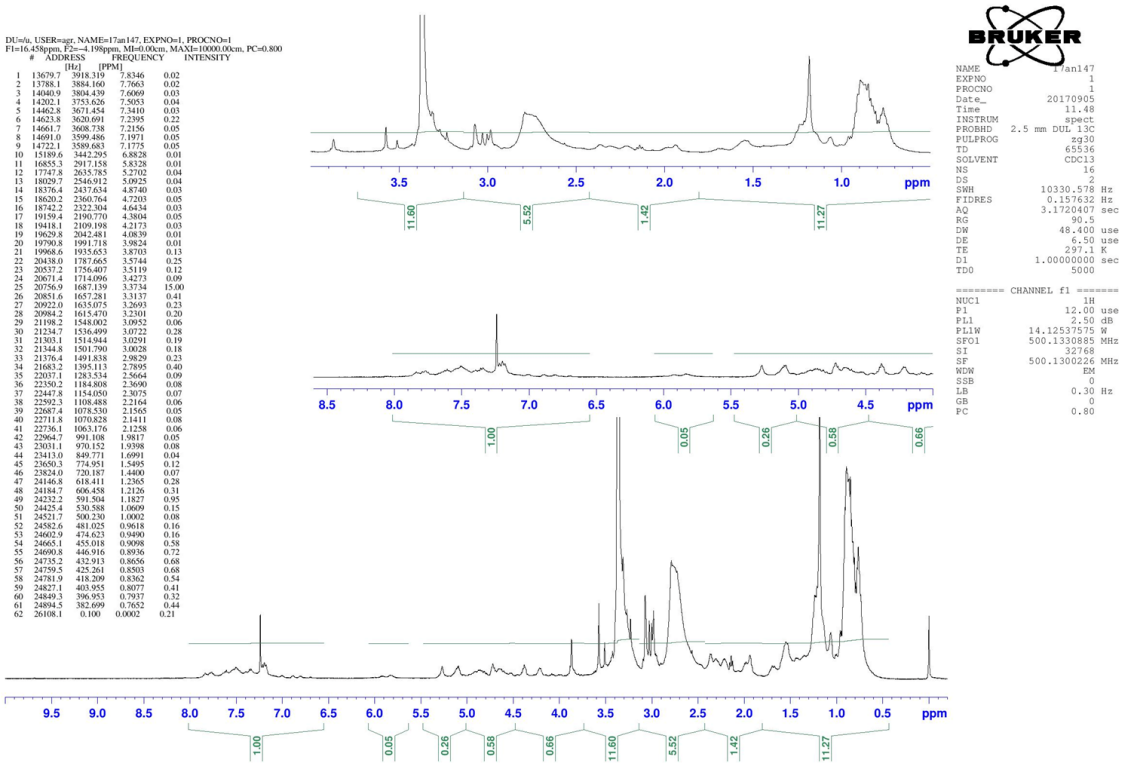


Figure S2.26 ^{1}H NMR spectrum ($CDCl_3$, 500 MHz) of compound **2**

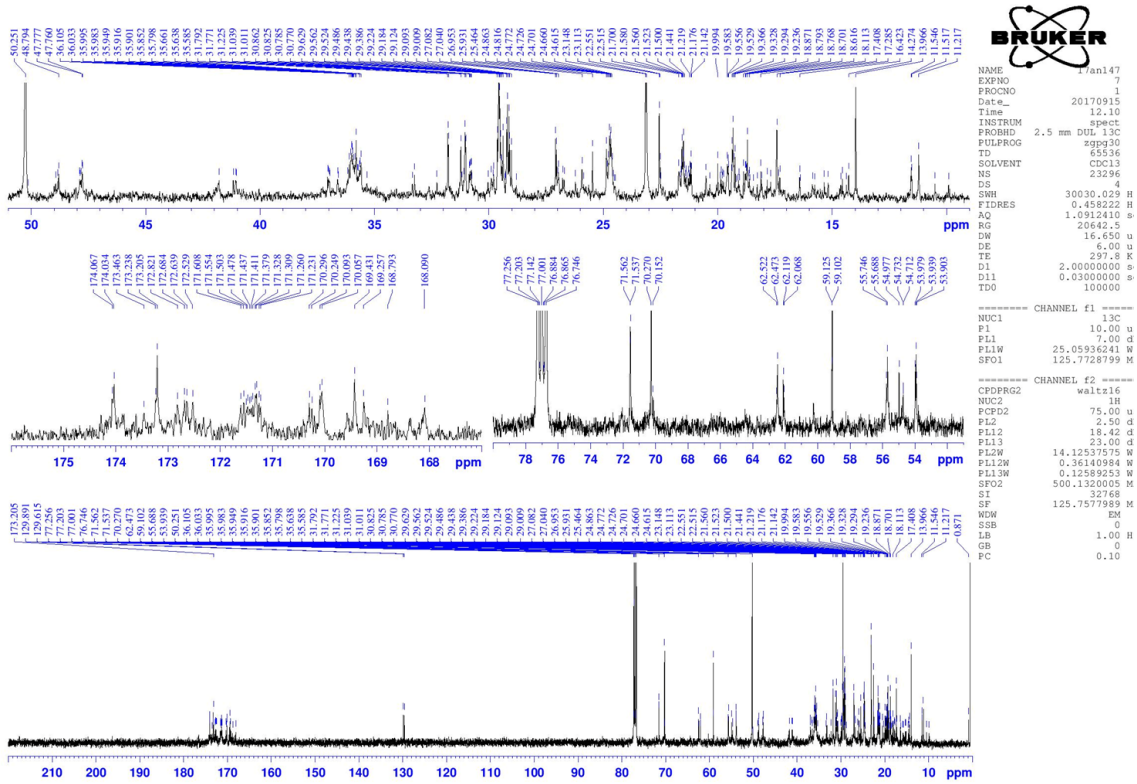


Figure S2.27 ^{13}C NMR spectrum ($CDCl_3$, 500 MHz) of compound **2**

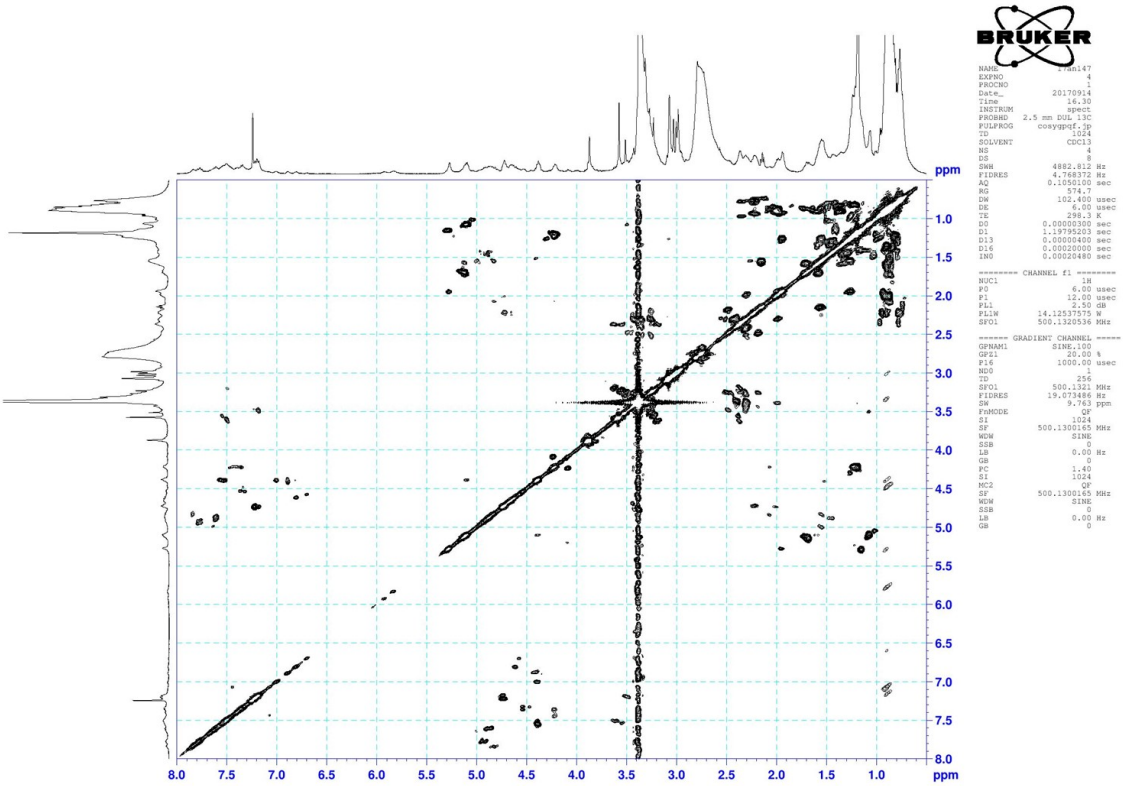


Figure S2.28 COSY NMR spectrum ($CDCl_3$, 500 MHz) of compound **2**

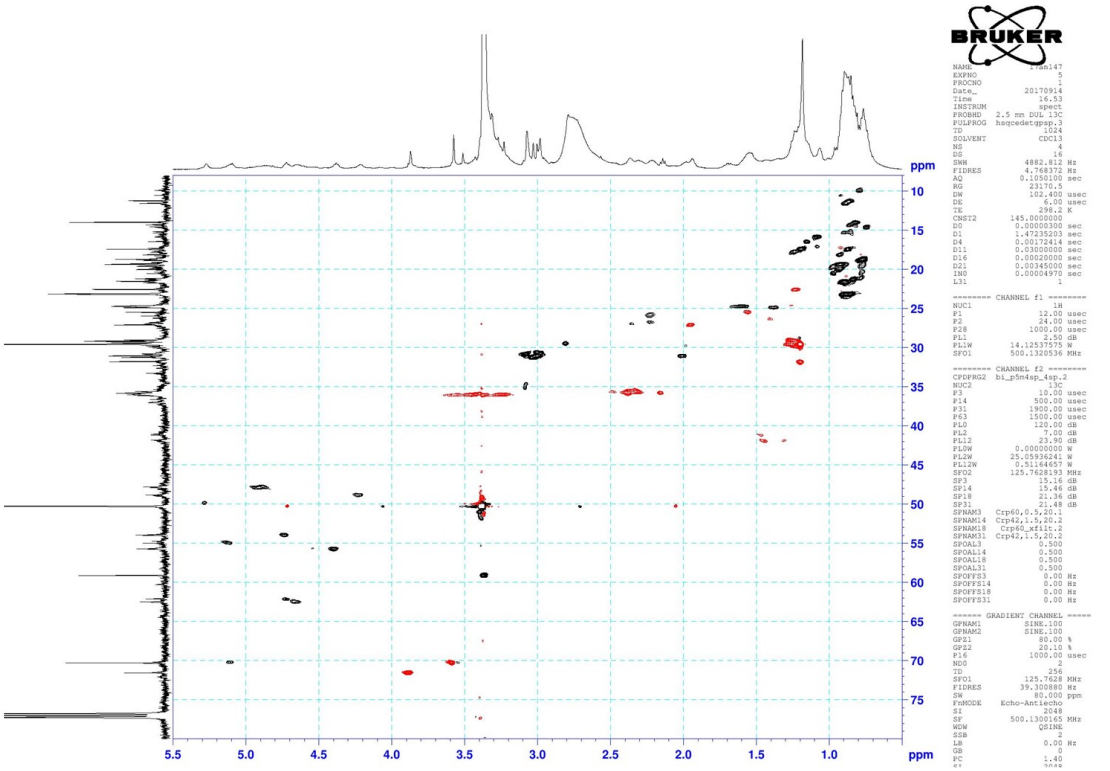


Figure S2.29 HSQC NMR spectrum ($CDCl_3$, 500 MHz) of compound **2**

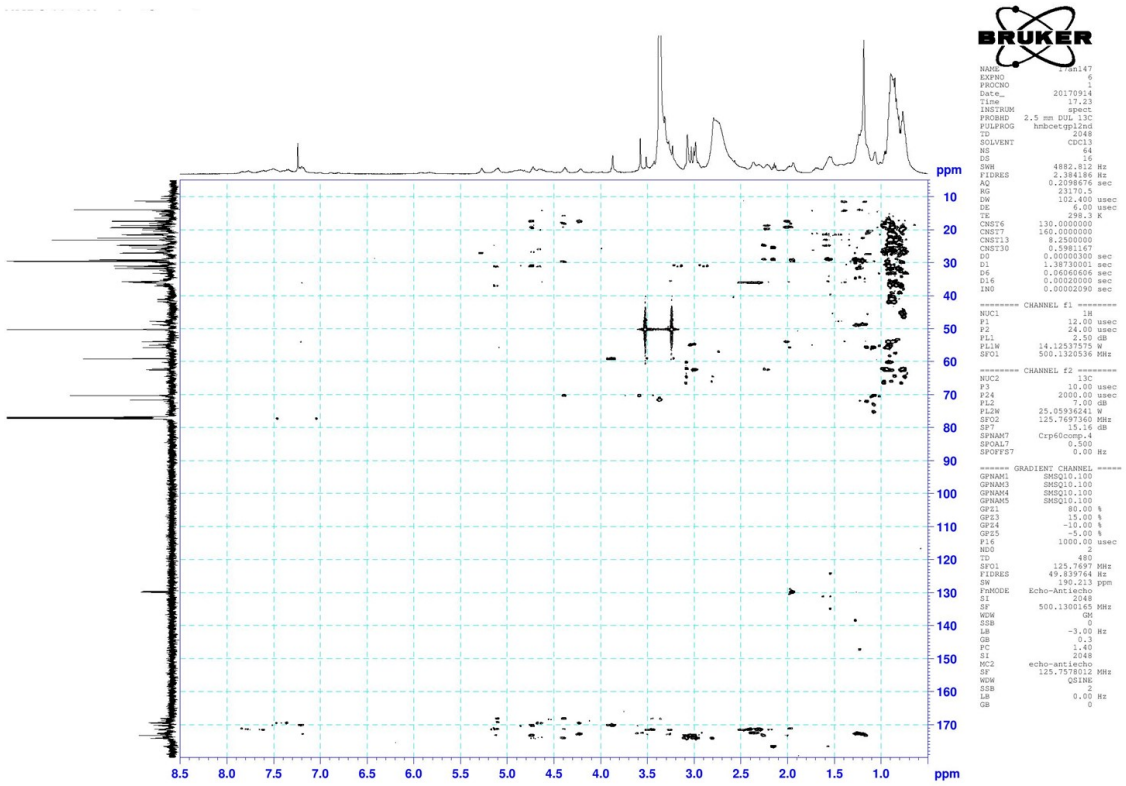


Figure S2.30 HMBC NMR spectrum ($CDCl_3$, 500 MHz) of compound **2**

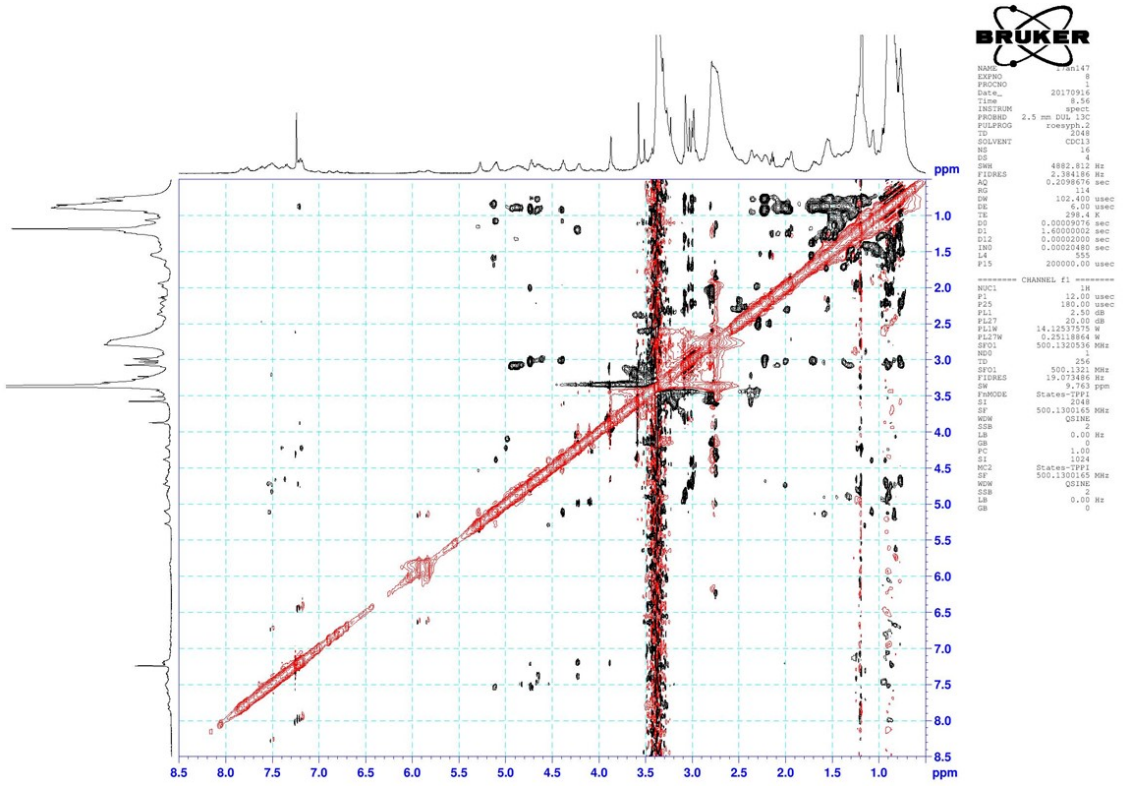


Figure S2.31 ROESY NMR spectrum ($CDCl_3$, 500 MHz) of compound **2**

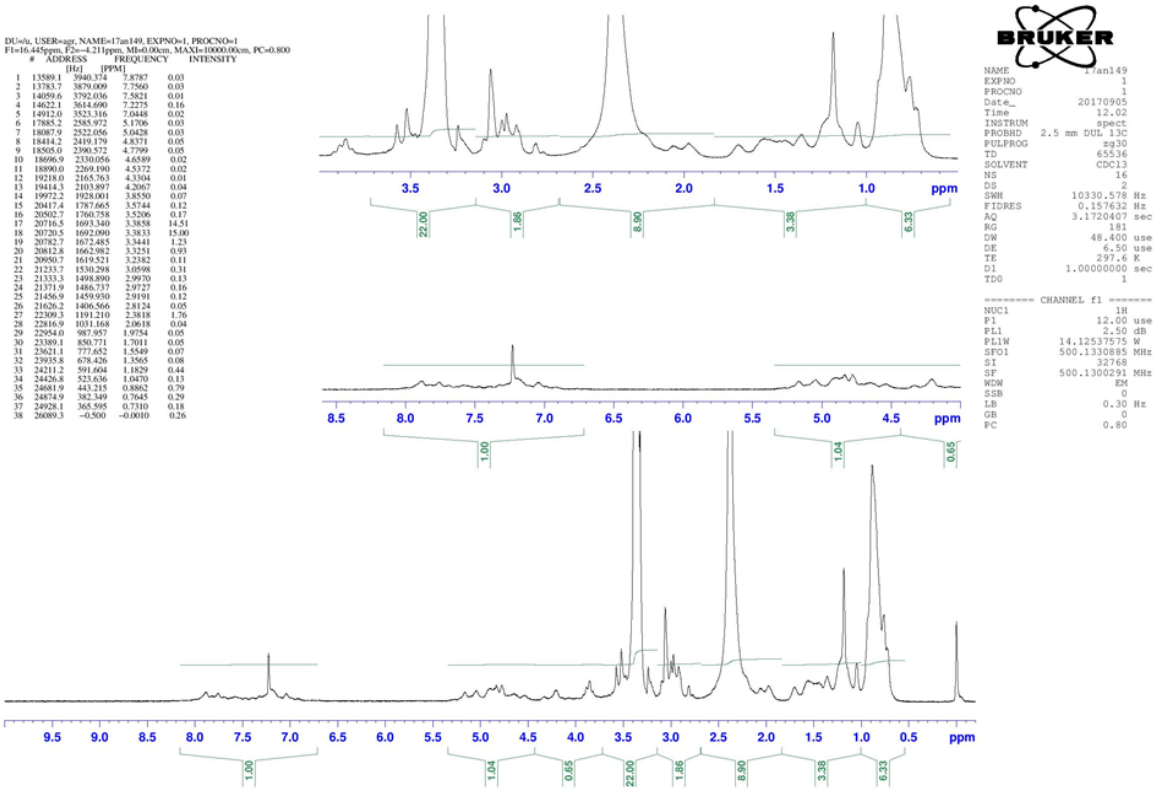


Figure S2.32 ^{1}H NMR spectrum ($CDCl_3$, 500 MHz) of compound **3**

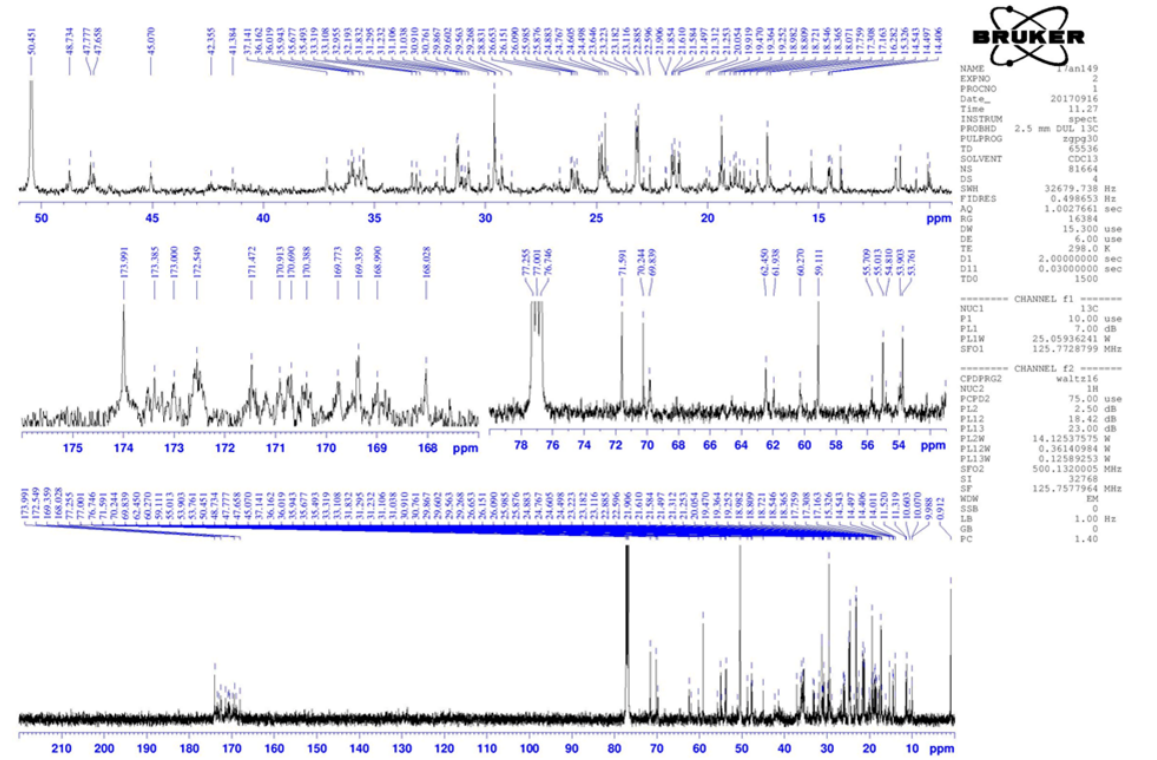


Figure S2.33 ^{13}C NMR spectrum ($CDCl_3$, 500 MHz) of compound **3**

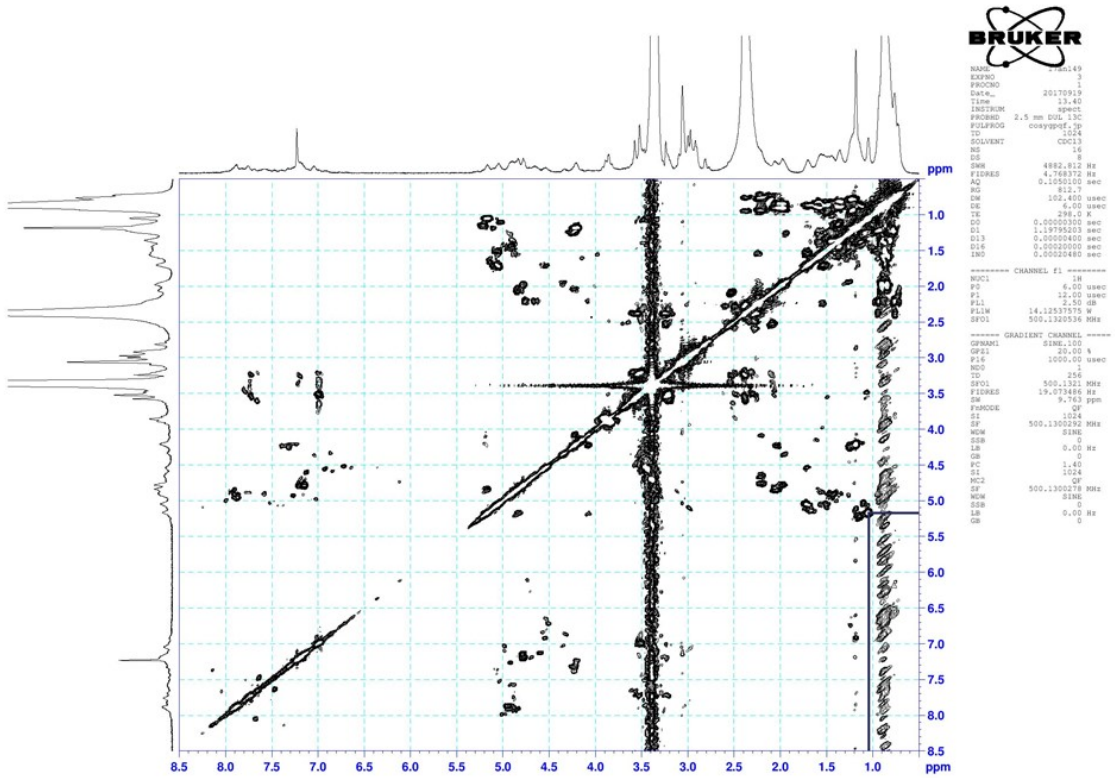


Figure S2.34 COSY NMR spectrum ($CDCl_3$, 500 MHz) of compound **3**

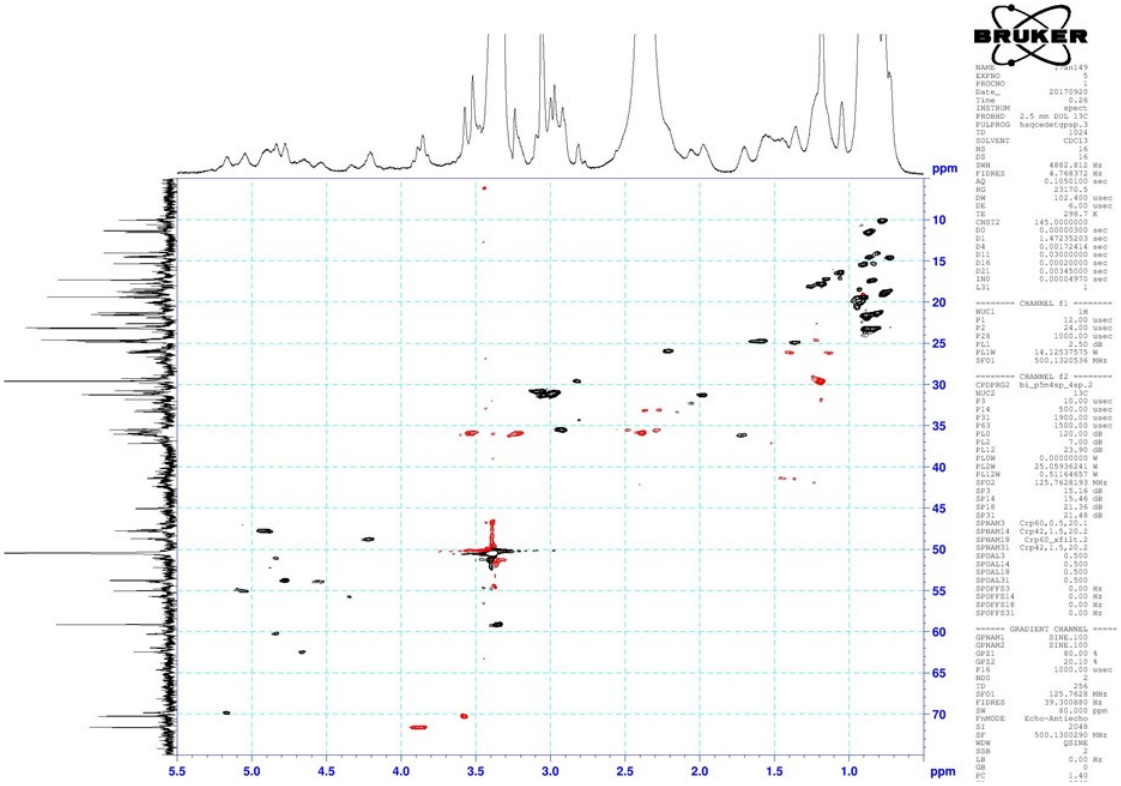


Figure S2.35 HSQC NMR spectrum ($CDCl_3$, 500 MHz) of compound **3**

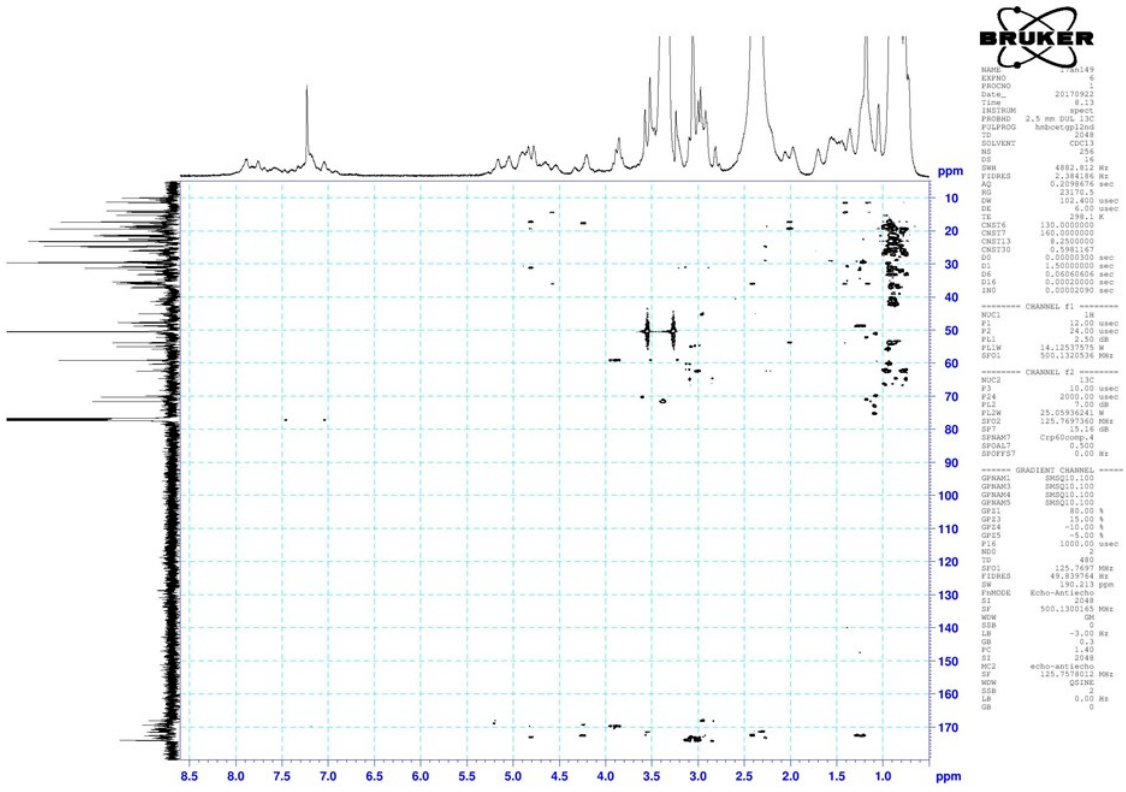


Figure S2.36 HMBC NMR spectrum ($CDCl_3$, 500 MHz) of compound **3**

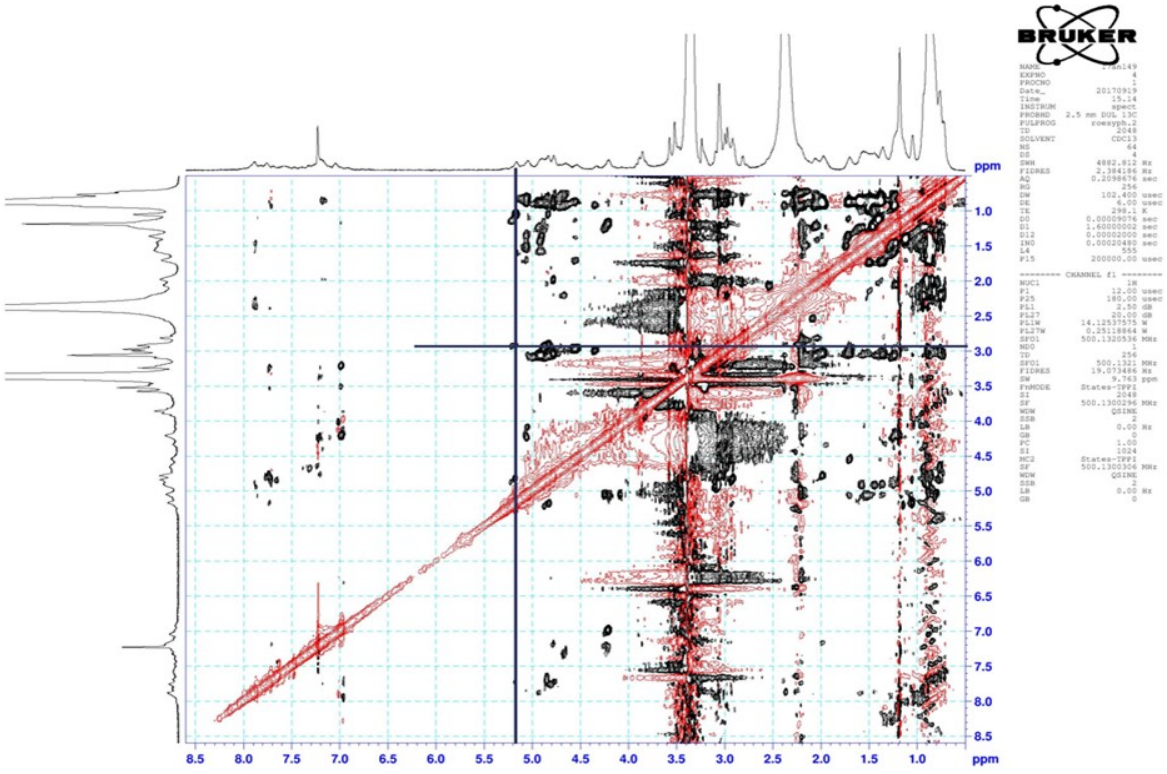


Figure S2.37 ROESY NMR spectrum ($CDCl_3$, 500 MHz) of compound **3**

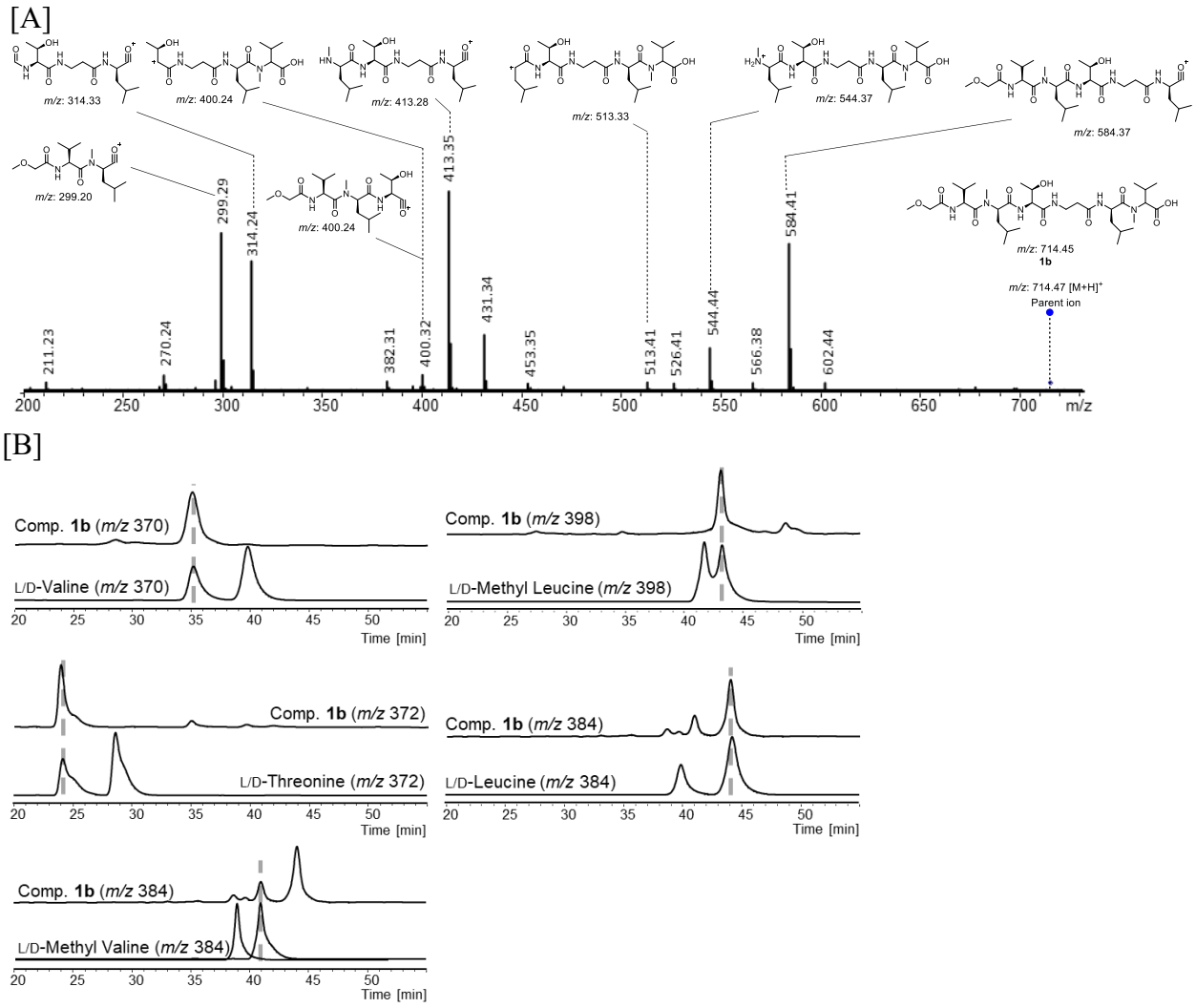


Figure S2.38 Tandem MS analysis of the targeted fragment **1b** derived from the N-terminal part of compound **1** [A], followed with Marfey's analysis to determine its absolute configuration [B].

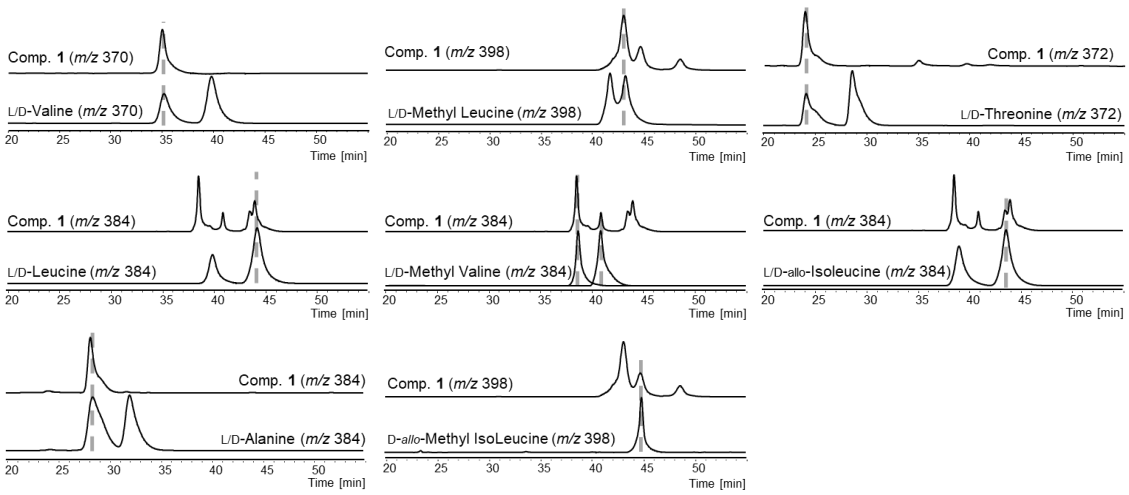
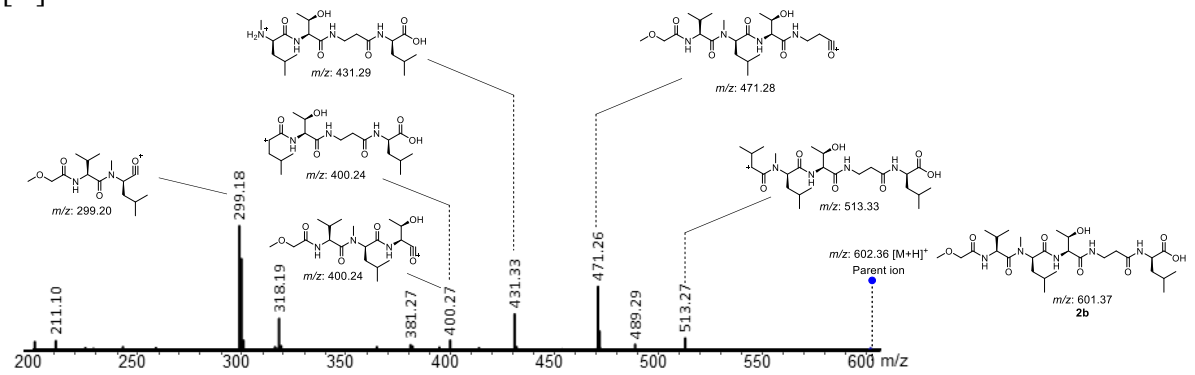


Figure S2.39 Marfey's analysis of the total hydrolysate of compound **1** in comparison with the amino acid standards.

[A]



[B]

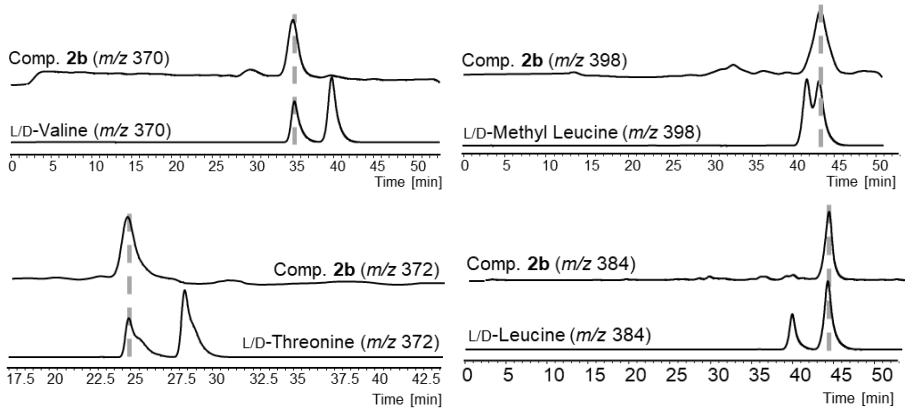


Figure S2.40 Tandem MS analysis of the targeted fragment **2b** derived from the N-terminal part of compound **2** [A], followed with Marfey's analysis to determine its absolute configuration [B].

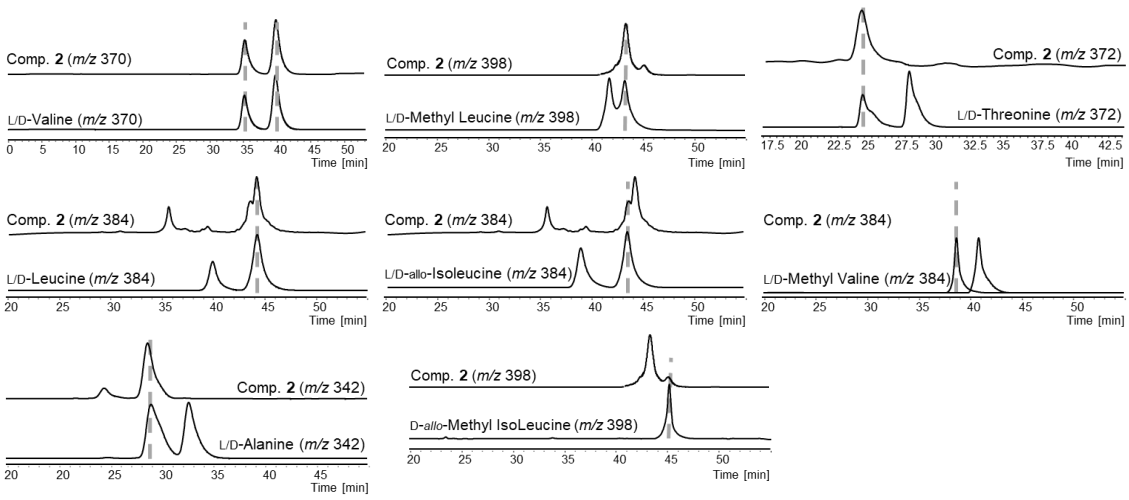
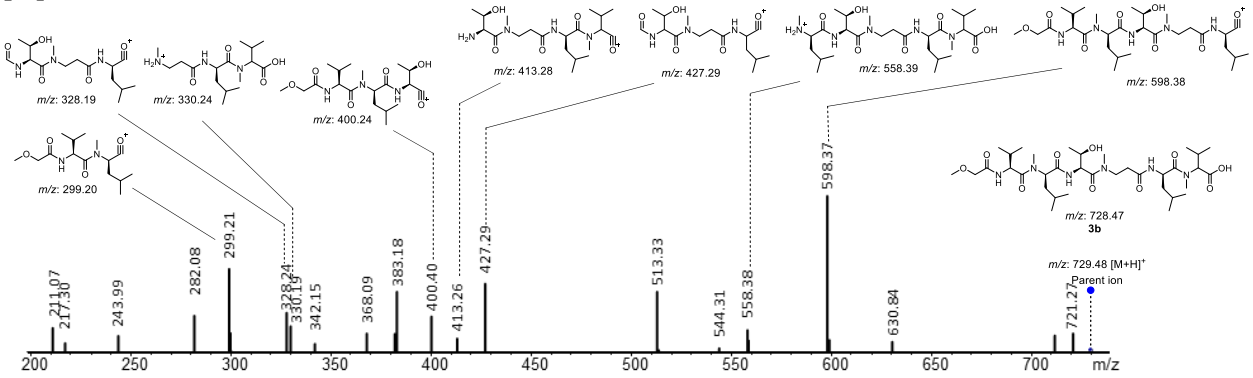


Figure S2.41 Marfey's analysis of the total hydrolysate of compound **2** in comparison with the amino acid standards.

[A]



[B]

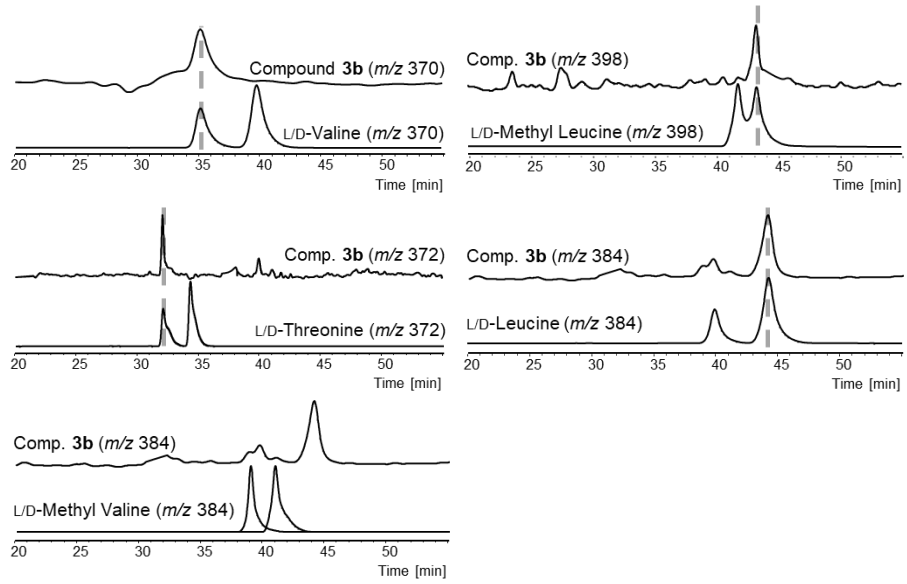


Figure S2.42 Tandem MS analysis of the targeted fragment **3b** derived from the N-terminal part of compound **3** [A], followed with Marfey's analysis to determine its absolute configuration [B].

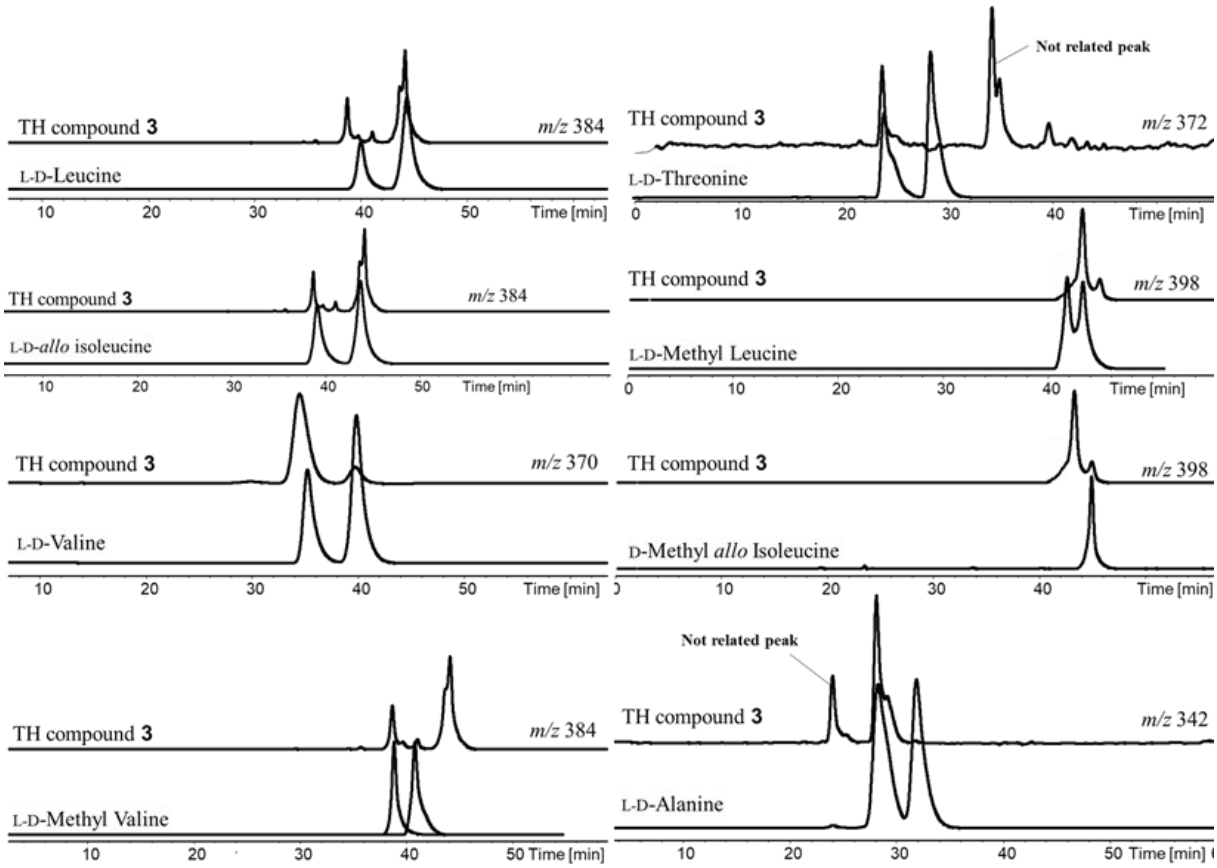


Figure S2.43 Marfey's analysis of the total hydrolysate of compound **3** in comparison with the amino acid standards.

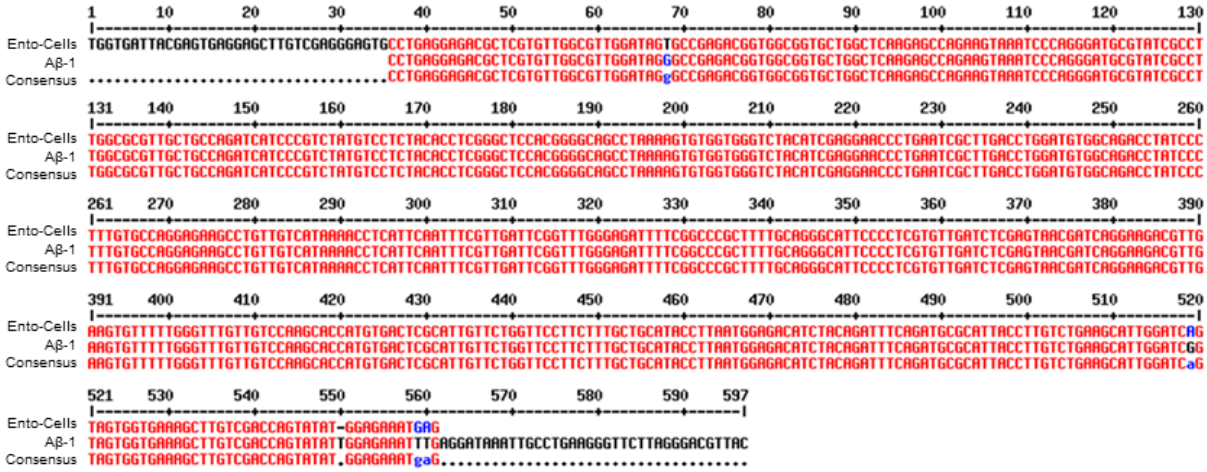


Figure S2.44 Sequence of amplified gene derived from Entotheonella cells from this study compared to database sequence of the first A domains from *Ca*. Entotheonella serta TSW1

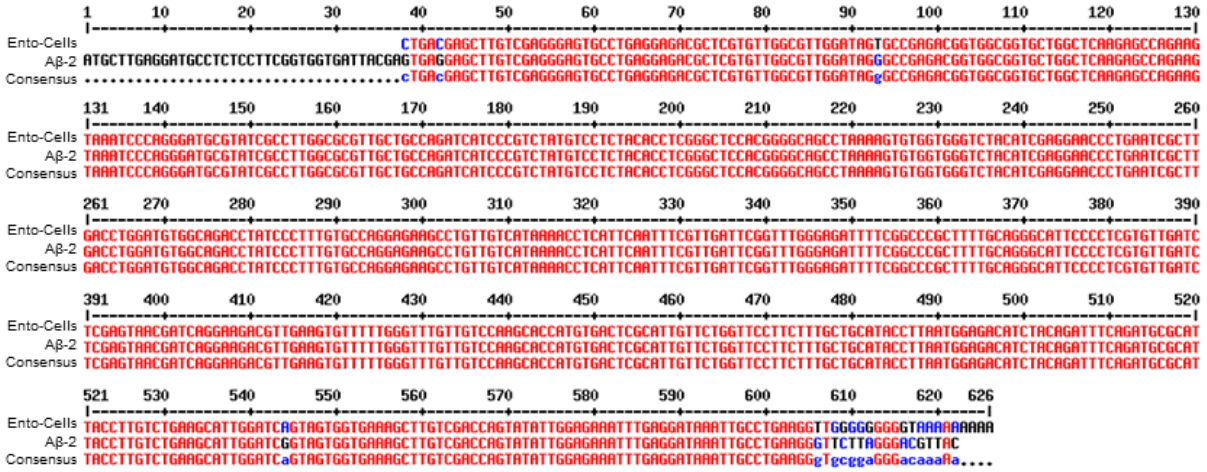
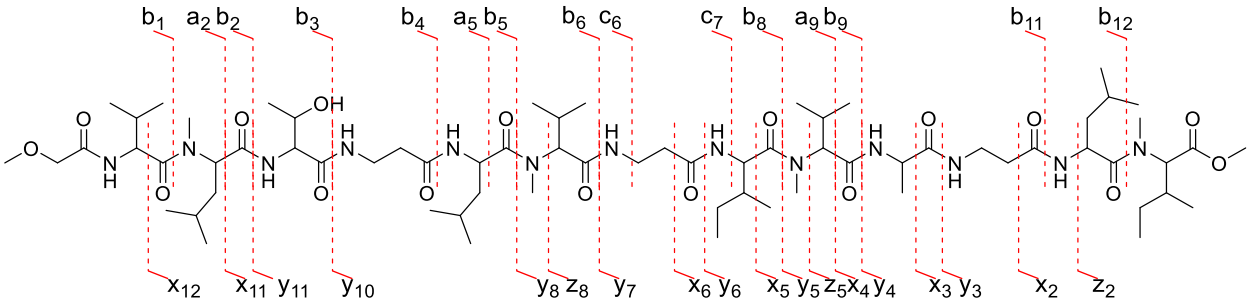


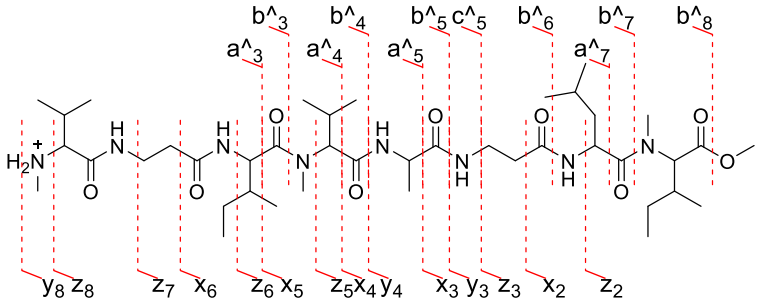
Figure S2.45 Sequence of amplified gene derived from Entotheonella cells from this study compared to database sequence of the first B domains from *Ca*. Entotheonella serta TSW1



1a

Table S2.1 List of obtained fragment ions from MS/MS analysis of **1a**

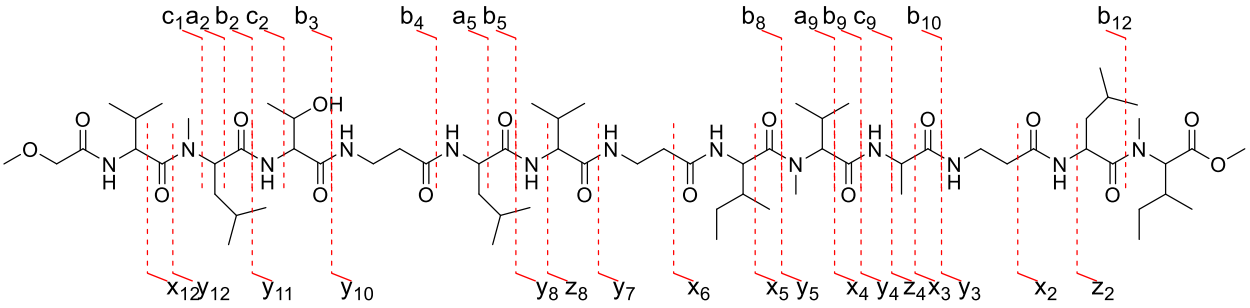
Ion	Calc. (Da)	Obs. (Da)	Δdelta (Da)	Ion	Calc. (Da)	Obs. (Da)	Δdelta (Da)
b1	172.0968	172.0716	0.0253	z2	256.1907	256.1165	0.0742
a2	271.2016	271.1479	0.0537	x2+H2O	317.1965	317.2708	0.0742
b2+H2O	317.1965	317.2708	0.0742	x2	299.1965	299.1774	0.0192
b2	299.1965	299.1774	0.0192	y3	344.2544	344.3039	0.0495
b3	400.2442	400.0228	0.2214	x3	370.2337	370.2658	0.0322
b4	471.2813	471.2503	0.0310	z4+H2O	416.2650	416.3188	0.0539
a5	556.3705	556.3348	0.0356	y4	415.2915	415.2365	0.0550
b5	584.3654	584.3572	0.0082	x4+H2O	459.2708	459.2392	0.0316
b6	697.4495	697.0222	0.4273	x4	441.2708	441.2595	0.0112
c6	714.4760	714.6461	0.1701	z5+H2O	515.3334	515.9225	0.5891
b7+H2O	786.4866	786.0298	0.4568	z5	497.3334	497.3955	0.0622
c7	785.5131	785.9608	0.4477	y5	528.3756	528.4362	0.0607
b8	881.5706	881.6166	0.0460	x5+H2O	572.3548	572.3600	0.0052
a9	966.6598	966.5709	0.0889	x5	554.3548	554.5071	0.1523
b9	994.6547	994.7037	0.0490	z6+H2O	642.4331	642.0222	0.4109
b11	1136.7289	1136.7332	0.0042	y6	641.4596	641.3369	0.1227
b12	1249.8130	1249.8666	0.0536	x6	667.4389	667.4990	0.0601
				z7+H2O	713.4702	713.6125	0.1423
				y7	712.4967	712.4990	0.0023
				z8	794.5386	794.5554	0.0168
				y8	825.5805	825.6157	0.0352
				z9+H2O	939.6383	939.3075	0.3308
				z10+H2O	1010.6754	1010.5914	0.0841
				y10	1009.7020	1009.6266	0.0754
				z11+H2O	1111.7231	1111.6885	0.0346
				y11	1110.7497	1110.7552	0.0056
				x11	1136.7289	1136.7332	0.0042
				x12	1263.8286	1263.8693	0.0406



y_8-**1a**

Table S2.2 List of obtained fragment ions from MS/MS/MS analysis of y_8-**1a**

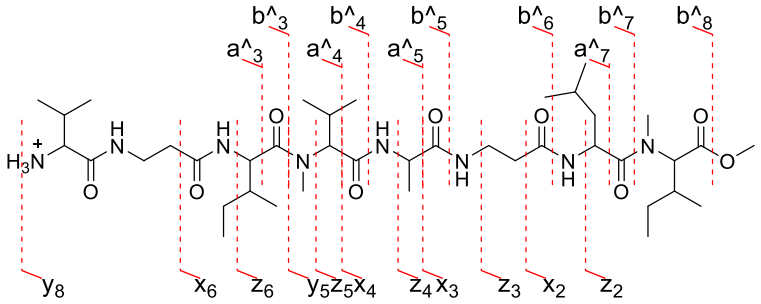
Ion	Calc. (Da)	Obs. (Da)	Δdelta (Da)	Ion	Calc. (Da)	Obs. (Da)	Δdelta (Da)
a^3	270.2176	270.1913	0.0263	z2	256.1907	256.0798	0.1109
b^3	298.2125	298.1420	0.0705	x2	299.1965	299.1702	0.0263
a^4	383.3017	383.3098	0.0081	z3	327.2278	327.2412	0.0134
b^4	411.2966	411.2961	0.0005	y3	344.2544	344.2126	0.0418
a^5	454.3388	454.3857	0.0469	x3	370.2337	370.2788	0.0452
b^5	482.3337	482.3195	0.0142	y4	415.2915	415.2700	0.0215
c^5	499.3603	499.3214	0.0388	x4	441.2708	441.3353	0.0645
b^6	553.3708	553.3790	0.0082	z5+H2O	515.3334	515.2952	0.0381
a^7	638.4600	638.5156	0.0556	z5	497.3334	497.3034	0.0300
b^7	666.4549	666.4556	0.0007	y5	528.3756	528.4121	0.0365
b^8	793.5546	793.5571	0.0025	x5	554.3548	554.3976	0.0428
				z6	624.4331	624.4590	0.0260
				x6	667.4389	667.4561	0.0172
				z8	794.5386	794.6066	0.0680



2a

Table S2.3 List of obtained fragment ions from MS/MS analysis of **2a**

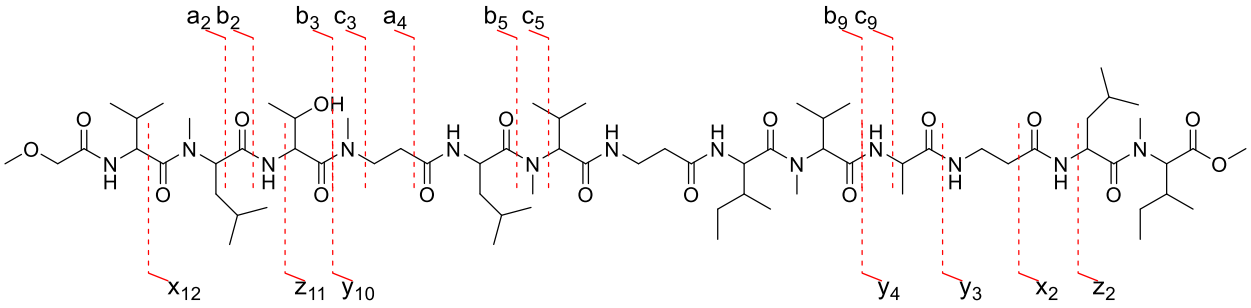
Ion	Calc. (Da)	Obs. (Da)	Δdelta (Da)	Ion	Calc. (Da)	Obs. (Da)	Δdelta (Da)
c1	203.1390	203.0924	0.0467	z2	256.1907	256.1545	0.0362
a2	271.2016	271.1439	0.0578	x2	299.1965	299.2136	0.0171
b2	299.1965	299.2136	0.0171	y3	344.2544	344.3035	0.0491
c2	316.2231	316.2105	0.0125	x3	370.2337	370.2604	0.0268
b3	400.2442	400.2528	0.0086	z4	398.2650	398.2693	0.0044
b4	471.2813	471.3384	0.0570	y4	415.2915	415.3503	0.0588
a5	556.3705	556.3973	0.0268	x4	441.2708	441.3018	0.0310
b5	584.3654	584.3900	0.0246	y5	528.3756	528.4907	0.1151
b5 + H2O	602.3654	602.4204	0.0550	x5	554.3548	554.3870	0.0322
a6 + H2O	673.4389	673.5240	0.0851	x6	667.4389	667.5293	0.0904
b8	867.5550	867.7025	0.1475	y7	712.4967	712.5329	0.0362
a9	952.6441	952.8411	0.1969	z8	794.5386	794.6181	0.0795
b9	980.6391	980.8881	0.2491	y8	811.5652	811.6312	0.0660
c9	997.6656	997.3965	0.2690	x8+H2O	855.5444	855.4921	0.0523
b10	1051.6762	1051.7500	0.0738	z10+H2O	996.6598	996.7625	0.1027
a11+H2O	1112.7184	1112.7456	0.0272	y10	995.6863	995.7158	0.0295
b12	1235.7973	1235.7866	0.0107	z11+H2O	1097.7075	1097.8020	0.0945
				y11	1096.7340	1096.8569	0.1229
				y12	1223.8337	1223.7208	0.1129
				x12	1249.8130	1249.8978	0.0848



y_8-**2a**

Table S2.4 List of obtained fragment ions from MS/MS/MS analysis of y_8-**2a**

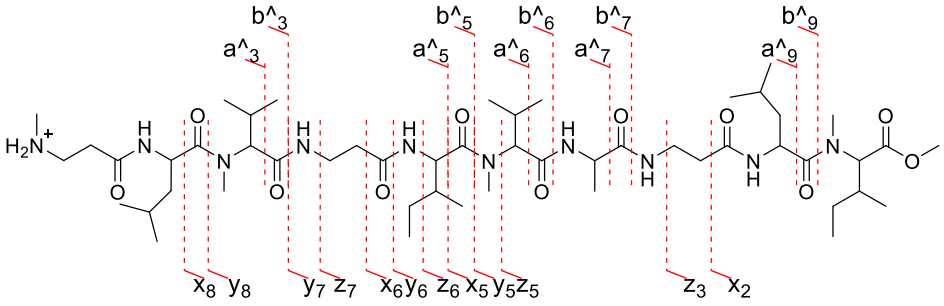
Ion	Calc. (Da)	Obs. (Da)	Δdelta (Da)	Ion	Calc. (Da)	Obs. (Da)	Δdelta (Da)
a^3	256.2020	256.2199	0.0180	z2	256.1907	256.2199	0.0292
b^3	284.1969	284.1956	0.0013	x2	299.1965	299.1811	0.0155
a^4	369.2860	369.2751	0.0109	z3	327.2278	327.2722	0.0444
b^4	397.2809	397.2563	0.0247	x3	370.2337	370.2709	0.0373
a^5	440.3231	440.2814	0.0418	z4	398.2650	398.3205	0.0556
b^5	468.3181	468.3733	0.0553	x4	441.2708	441.3669	0.0962
b^6	539.3552	539.2705	0.0847	z5	497.3336	497.3200	0.0136
a^7	624.4443	624.3962	0.0481	y5	528.3756	528.3950	0.0195
b^7	652.4392	652.4911	0.0519	z6	624.4331	624.3962	0.0368
b^8	779.5389	779.5640	0.0250	x6	667.4389	667.4293	0.0096



3a

Table S2.5 List of obtained fragment ions from MS/MS analysis of **3a**

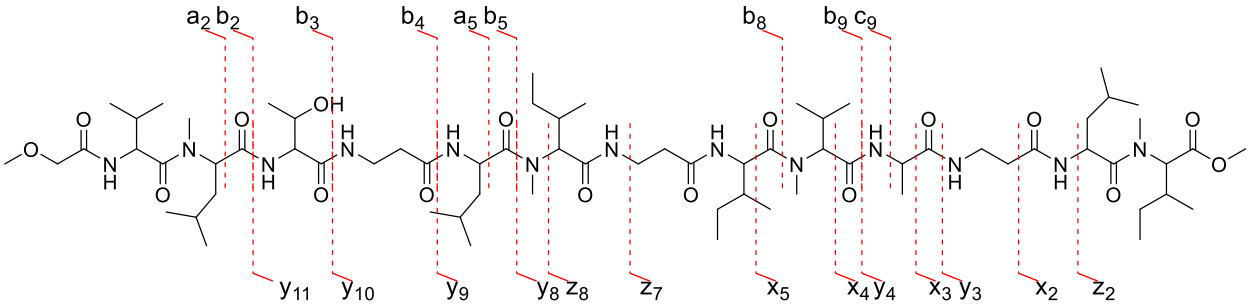
Ion	Calc. (Da)	Obs. (Da)	Δdelta (Da)	Ion	Calc. (Da)	Obs. (Da)	Δdelta (Da)
a2	271.2016	271.1167	0.0850	z2	256.1907	256.2096	0.0189
b2	299.1965	299.1906	0.0060	x2	299.1965	299.1906	0.0060
b3	400.2442	400.2771	0.0329	y3	344.2544	344.1664	0.0880
c3+H2O	449.3864	449.9526	0.5662	z4+H2O	416.2650	416.5485	0.2836
c3	431.3864	431.1473	0.2391	y4	415.2915	415.2530	0.0385
a4	457.3021	457.1173	0.1848	y8+H2O	843.5808	843.9647	0.3839
b5	598.3810	598.4725	0.0915	y10	1023.7176	1023.7256	0.0080
c5	629.4232	629.4316	0.0083	z11+H2O	1125.7388	1125.7288	0.0100
b9	1008.6704	1008.6068	0.0635	y11	1124.7653	1124.7495	0.0158
c9	1025.6969	1025.7659	0.0690	x12	1277.8443	1277.8545	0.0102



y_{10}-**3a**

Table S2.6 List of obtained fragment ions from MS/MS/MS analysis of y_{10}-**3a**.

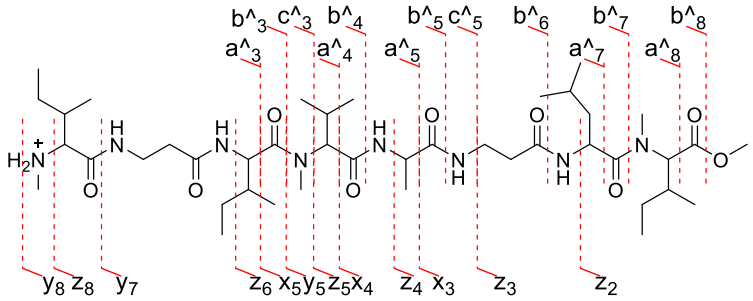
Ion	Calc. (Da)	Obs. (Da)	Δdelta (Da)	Ion	Calc. (Da)	Obs. (Da)	Δdelta (Da)
a^3	284.2333	284.1164	0.1169	x2	299.1965	299.1617	0.0348
b^3	312.2282	312.1706	0.0576	z3	327.2278	327.1898	0.0381
a^4+H2O	373.2704	373.2407	0.0296	z5	497.3334	497.2705	0.0629
a^5	468.3544	468.1800	0.1744	y5	528.3756	528.3364	0.0392
b^5+H2O	514.3494	514.2982	0.0511	x5	554.3548	554.3978	0.0430
b^5	496.3494	496.3857	0.0364	z6	624.4331	624.4921	0.0590
a^6	581.4385	581.4949	0.0564	y6	641.4596	641.5109	0.0512
b^6	609.4334	609.4700	0.0366	x6	667.4389	667.4294	0.0094
a^7	652.4756	652.5311	0.0555	z7+H2O	713.4702	713.3705	0.0996
b^7+H2O	698.4705	698.5139	0.0434	z7	695.4702	695.4602	0.0100
b^7	680.4705	680.5207	0.0502	y7	712.4967	712.5256	0.0289
a^9	836.5968	836.5648	0.0320	y8	825.5808	825.5865	0.0057
b^9	864.5917	864.5847	0.0070	x8	851.5601	851.4989	0.0612



4a

Table S2.7 List of obtained fragment ions from MS/MS analysis of **4a**

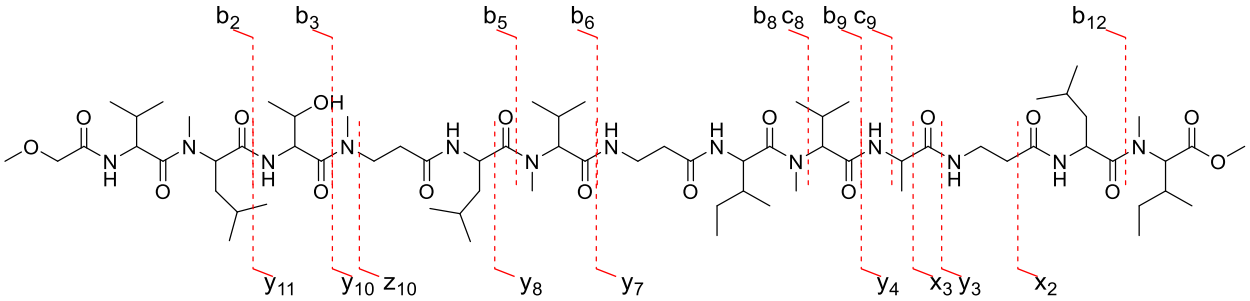
Ion	Calc. (Da)	Obs. (Da)	Δdelta (Da)	Ion	Calc. (Da)	Obs. (Da)	Δdelta (Da)
a2	271.2016	271.1837	0.0180	z2	256.1907	256.2704	0.0797
b2	299.1965	299.1750	0.0215	x2	299.1965	299.1750	0.0215
b3	400.2442	400.2280	0.0162	y3	344.2544	344.2978	0.0434
b4	471.2813	471.2309	0.0505	x3	370.2337	370.2358	0.0022
a5	556.3705	556.3392	0.0312	z4+H2O	416.2650	416.3665	0.1016
b5	584.3654	584.3403	0.0251	y4	415.2915	415.3169	0.0254
b8	895.5863	895.8677	0.2814	x4	441.2708	441.3416	0.0709
b9	1008.6704	1008.5639	0.1064	y5+H2O	546.3756	546.8662	0.4907
c9	1025.6969	1025.7136	0.0167	x5	554.3548	554.3998	0.0450
				x6+H2O	685.4389	685.5997	0.1608
				z7	695.4702	695.7993	0.3291
				z8	808.5543	808.5745	0.0203
				y8	839.5965	839.5953	0.0012
				x8+H2O	883.5757	883.8066	0.2309
				y9	952.6805	952.6453	0.0353
				z10+H2O	1024.6911	1024.6700	0.0210
				y10	1023.7176	1023.6793	0.0384
				z11+H2O	1125.7388	1125.7618	0.0231
				y11	1124.7653	1124.8483	0.0830



y_8-**4a**

Table S2.8 List of obtained fragment ions from MS/MS/MS analysis of y_8-**4a**

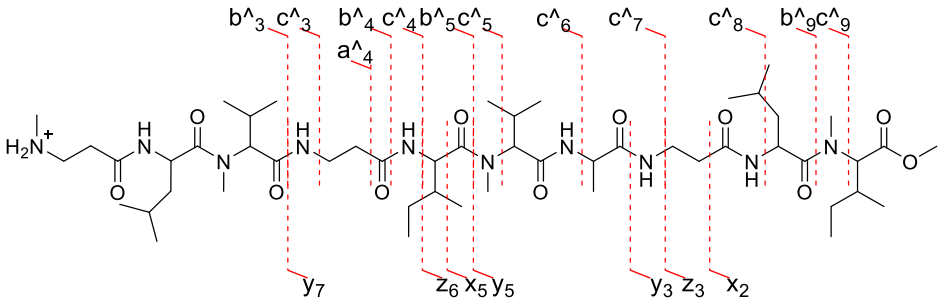
Ion	Calc. (Da)	Obs. (Da)	Δdelta (Da)	Ion	Calc. (Da)	Obs. (Da)	Δdelta (Da)
a^3	284.2333	284.2332	0.0000	z2	256.1907	256.1338	0.0569
b^3	312.2282	312.2136	0.0146	z3	327.2278	327.2726	0.0448
c^3	343.2704	343.1992	0.0712	x3	370.2337	370.2361	0.0025
a^4	397.3173	397.3780	0.0607	z4	398.2650	398.2657	0.0007
b^4	425.3122	425.3024	0.0098	x4	441.2708	441.3000	0.0292
a^5	468.3544	468.3275	0.0270	z5	497.3334	497.3290	0.0044
b^5	496.3494	496.3673	0.0179	y5	528.3756	528.3574	0.0181
c^5	513.3759	513.3659	0.0100	x5	554.3548	554.4479	0.0930
b^6	567.3865	567.3593	0.0272	z6+H2O	642.4331	642.5131	0.0800
a^7	652.4756	652.4598	0.0158	z6	624.4331	624.4572	0.0241
b^7	680.4705	680.4728	0.0023	z7+H2O	713.4702	713.4330	0.0372
a^8	779.5753	779.5262	0.0491	y7	712.4967	712.4854	0.0114
b^8	807.5702	807.5375	0.0328	z8	808.5543	808.6328	0.0786



5a

Table S2.9 List of obtained fragment ions from MS/MS analysis of **5a**

Ion	Calc. (Da)	Obs. (Da)	Δdelta (Da)	Ion	Calc. (Da)	Obs. (Da)	Δdelta (Da)
c1+H2O	221.1390	221.0938	0.0453	y2+H2O	291.2173	291.7729	0.5557
b2	299.1965	299.1843	0.0123	x2	299.1965	299.1843	0.0123
c2+H2O	334.2231	334.4922	0.2692	y3	344.2544	344.2374	0.0170
b3	400.2442	400.1840	0.0602	x3	370.2337	370.3042	0.0705
b4+H2O	503.2970	503.3360	0.0390	y4	415.2915	415.3901	0.0986
a5+H2O	588.3861	588.3979	0.0118	y7	712.4967	712.4849	0.0118
b5	598.3810	598.3754	0.0057	y8	839.5965	839.5500	0.0465
b6	725.4808	725.4555	0.0252	z10	1006.6911	1006.7667	0.0756
c6+H2O	760.5073	760.4808	0.0265	y10	1037.7333	1037.7349	0.0016
b8	909.6019	909.5986	0.0034	z11+H2O	1139.7810	1139.8339	0.0529
c8	940.6441	940.6146	0.0296	y11	1138.7810	1138.7908	0.0098
b9	1022.6860	1022.6556	0.0304				
c9	1039.7125	1039.7324	0.0199				
b12	1277.8443	1277.9799	0.1356				



y_{10}-**5a**

Table S2.10 List of obtained fragment ions from MS/MS/MS analysis of y_{10}-**5a**

Ion	Calc. (Da)	Obs. (Da)	Δdelta (Da)	Ion	Calc. (Da)	Obs. (Da)	Δdelta (Da)
b^3	299.2562	299.1168	0.1394	x2	299.1965	299.1168	0.0797
c^3	327.2511	327.2304	0.0207	z3	327.2278	327.2304	0.0025
a^4	344.2776	344.2719	0.0057	y3	344.2544	344.2719	0.0175
b^4	369.2860	369.2168	0.0692	y5	528.3756	528.3298	0.0457
c^4	397.2809	397.2720	0.0089	x5	554.3548	554.3840	0.0292
b^5	482.3701	482.4616	0.0916	z6	624.4331	624.4669	0.0338
c^5+H2O	528.3650	528.3298	0.0352	z7+H2O	713.4702	713.5400	0.0698
c^5	510.3650	510.3632	0.0018	y7	712.4967	712.4955	0.0012
c^6	623.4491	623.4429	0.0062				
c^7+H2O	712.4862	712.4955	0.0094				
c^7	694.4862	694.4308	0.0554				
c^8	765.5233	765.5231	0.0002				
b^9	850.6124	850.6707	0.0583				
c^9	878.6074	878.5517	0.0557				

Table S2.11 NMR data (in $CDCl_3$, 500 MHz) of compound **1**

Pos.	δ_C, type		δ_H (mult.)	Pos.	δ_C, type		δ_H (mult.)	Pos.	δ_C, type		δ_H (mult.)
Methoxy acetyl				D-Leucine[5]				C2	62.3, CH		4.65 (m)
C1	170.1, C			C1	n.r*., C			C3	26.7, CH		2.22 (m)
C2	71.5, CH_2		3.89 (m)	C2	47.7, CH		4.89 (m)	C4	18.6, CH_3		0.75 (m)
C3	70.2, CH_3		3.59 (m)	C3	41.0, CH_2	a:	1.32 (m)	C5	19.6, CH_3		0.90 (m)
L-Valine[1]						b:	1.44 (m)	NMe	31.0, CH_3		2.99 (m)
C1	173.1, C			C4	24.6, CH		1.49 (m)	L-Alanine[10]			
C2	53.9, CH		4.72 (m)	C5	23.0, CH_3		0.85 (m)	C1	172.7, C		
C3	30.9, CH		1.99 (m)	C6	23.1, CH_3		0.88 (m)	C2	48.75, CH		4.22 (m)
C4	19.3, CH_3		0.90 (m)	NH			7.57 (m)	C3	17.2.0, CH_3		1.15 (m)
C5	17.31, CH_3		0.85 (m)	D-Methyl Valine[6]				NH			7.45 (m)
NH			7.20 (m)	C1	170.4, C			β-Alanine[11]			
D-Methyl Leucine[2]				C2	62.1, CH		4.72 (m)	C1	171.4, C		
C1	171.2, C			C3	25.7, CH		2.15 (m)	C2	34.9, CH_2	a:	2.38 (m)
C2	54.9, CH		5.08 (m)	C4	19.8, CH_3		0.91 (m)			b:	2.29 (m)
C3	37.1, CH_2	a:	1.58 (m)	C5	18.8, CH_3		0.75 (m)	C3	36.2, CH_2	a:	3.43 (m)
		b:	1.72 (m)	NMe	31.3, CH_3		3.01 (m)			b:	3.35 (m)
C4	24.8, CH		1.35 (m)	β-Alanine[7]				NH			7.52 (m)
C5	21.1, CH_3		0.82 (m)	C1	172.5, C			D-Leucine[12]			
C6	21.5, CH_3		0.87 (m)	C2	35.8, CH_2	a:	2.36 (m)	C1	173.1, C		
NMe	30.8, CH_3		3.05 (m)			b:	n.r.	C2	47.8, CH		4.92 (m)
L-Threonine[3]				C3	36.0, CH_3	a :	3.60 (m)	C3	41.6, CH_2	a:	1.28 (m)
C1	168.2, C					b:	3.19 (m)			b:	1.43 (m)
C2	55.7, CH		4.39 (m)	NH			7.52 (m)	C4	24.6, CH		1.50 (m)
C3	70.2, CH		5.08 (m)	D-*allo* Isoleucine[8]				C5	23.1, CH_3		0.83 (m)
C4	15.9, CH_3		1.05 (m)	C1	173.7, C			C6	23.2, CH_3		0.89 (m)
NH			7.59 (m)	C2	60.2, CH		4.80 (m)	NH			7.78 (m)
				C3	24.8, CH		2.05 (m)	D-Methyl *allo* Isoleucine[13]			
β-Alanine[4]				C4	24.6, CH_2	a:	1.24 (m)	C1	173.09, C		
C1	171.3, C					b:	0.96 (m)	C2	54.7, CH		5.15 (m)
C2	35.8, CH_2	a:	2.46 (m)	C5	15.2, CH_3		0.91 (m)	C3	36.6, CH		1.71 (m)
		b:	2.19 (m)	C6	15.1, CH_3		0.83 (m)	C4	26.2, CH_2	a:	1.38 (m)
C3	36.0, CH_2	a:	3.47 (m)	NH			7.83 (m)			b:	1.16 (m)
		b:	3.26 (m)	L-Methyl Valine[9]				C5	17.9, CH_3		0.88 (m)
NH			7.25 (m)	C1	169.5, C			C6	17.2, CH_3		0.85 (m)
								NMe	30.6, CH_3		2.98 (m)

*n.r (not resolved)

Table S2.12 NMR data (in $CDCl_3$, 500 MHz) of compound **2**

Pos.	δ_C, type		δ_H (mult.)	Pos.	δ_C, type		δ_H (mult.)	Pos.	δ_C, type		δ_H (mult.)
Methoxy acetyl				D-Leucine[5]				C2	62.6, CH		4.65 (m)
C1	170.1, C			C1	n.r., C			C3	25.8, CH		2.21 (m)
C2	71.5, CH_2		3.85 (m)	C2	47.9, CH		4.93 (m)	C4	18.7, CH_3		0.78 (m)
C3	70.2, CH_3		3.59 (m)	C3	41.2, CH_2	a:	1.48 (m)	C5	19.8, CH_3		0.95 (m)
L-Valine[1]						b:	1.46 (m)	NMe	30.9, CH_3		2.98 (m)
C1	173.2, C			C4	n.r., CH		n.r.	L-Alanine[10]			
C2	62.2, CH		4.72 (m)	C5	20.0, CH_3		0.90 (m)	C1	172.7, C		
C3	26.8, CH		2.22 (m)	C6	21.7, CH_3		0.88 (m)	C2	48.7, CH		4.22 (m)
C4	18.7, CH_3		0.78 (m)	NH			7.61 (m)	C3	17.2, CH_3		1.15 (m)
C5	19.8, CH_3		0.93 (m)	D-Valine[6]				NH			7.45 (m)
NH			7.18 (m)	C1	170.4, C			β-Alanine[11]			
D-Methyl Leucine[2]				C2	54.0, CH		4.71 (m)	C1	171.5 , C		
C1	n.r., C			C3	31.1, CH		2.00 (m)	C2	35.8, CH_2	a:	2.48 (m)
C2	55.0, CH		5.10 (m)	C4	17.5, CH_3		0.85 (m)			b:	2.19 (m)
C3	37.0, CH_2	a:	2.01 (m)	C5	19.3, CH_3		0.91 (m)	C3	36.0, CH_2	a:	3.52 (m)
		b:	1.39 (m)	NH			7.22 (m)			b:	3.26 (m)
C4	n.r. CH		n.r.	β-Alanine[7]				NH			7.55 (m)
C5	20.4, CH_3		0.83 (m)	C1	172.5, C			D-Leucine[12]			
C6	19.4, CH_3		0.90 (m)	C2	35.8, CH_2	a:	2.38 (m)	C1	n.r, C		
NMe	30.8, CH_3		3.04 (m)			b:	2.42 (m)	C2	47.8, CH		4.88 (m)
L-Threonine[3]				C3	36.0, CH_2	a:	3.59 (m)	C3	41.9, CH_2	a:	1.28 (m)
C1	168.1, C					b:	3.39 (m)			b:	1.43 (m)
C2	55.9, CH		4.38 (m)	NH			7.50 (m)	C4	n.r., CH		n.r
C3	70.2, CH		5.10 (m)	D-*allo* Isoleucine[8]				C5	21.3, CH_3		0.83 (m)
C4	17.2, CH_3		1.07 (m)	C1	n.r., C			C6	20.3, CH_3		0.85 (m)
NH			7.55 (m)	C2	53.9, CH		4.75 (m)	NH			7.25 (m)
				C3	n.r. CH		2.15 (m)	D-Methyl *allo* Isoleucine[13]			
β-Alanine[4]				C4	25.4, CH_2	a:	1.54 (m)	C1	n.r., C		
C1	171.4, C					b:	n.r.	C2	54.8, CH		5.15 (m)
C2	35.8, CH_2	a:	2.40 (m)	C5	21.2, CH_3		0.78 (m)	C3	n.r., CH		1.71 (m)
		b:	2.32 (m)	C6	23.2, CH_3		0.93 (m)	C4	24.7, CH_2	a:	1.38 (m)
C3	36.0, CH_2	a:	3.43 (m)	NH			7.25 (m)			b:	1.16 (m)
		b:	3.37 (m)	L-Methyl Valine[9]				C5	21.4, CH_3		0.92 (m)
NH			7.20 (m)	C1	169.4, C			C6	23.3, CH_3		0.82 (m)
								NMe	31.2, CH_3		2.99 (m)

*n.r (not resolved)

Table S2.13 NMR data (in $CDCl_3$, 500 MHz) of compound **3**

Pos.	δ_C, type		δ_H (mult.)	Pos.	δ_C, type		δ_H (mult.)	Pos.	δ_C, type		δ_H (mult.)
Methoxy acetyl				D-Leucine[5]				C2	62.4, CH		4.66 (m)
C1	170.1, C			C1	n.r., C			C3	25.9, CH		2.21 (m)
C2	71.5, CH_2		3.89 (m)	C2	n.r., CH		4.85 (m)	C4	18.8, CH_3		0.78 (m)
C3	70.2, CH_3		3.59 (m)	C3	41.3, CH_2	a:	1.35 (m)	C5	20.3, CH_3		0.96 (m)
L-Valine[1]						b:	1.47 (m)	NMe	n.r., CH_3		3.05 (m)
C1	n.r., C			C4	n.r., CH		1.99 (m)	L-Alanine[10]			
C2	53.6, CH		4.78 (m)	C5	21.6, CH_3		0.88 (m)	C1	n.r., C		
C3	26.0, CH		1.99 (m)	C6	20.2, CH_3		0.93 (m)	C2	48.8, CH		4.22 (m)
C4	18.7, CH_3		0.77 (m)	NH			7.12 (m)	C3	17.8, CH_3		1.25 (m)
C5	19.4, CH_3		0.95 (m)	L-Metyl Valine[6]				NH			7.31 (m)
NH			7.19 (m)	C1	n.r., C			β-Alanine[11]			
D-Methyl Leucine[2]				C2	53.7, CH		4.77 (m)	C1	n.r., C		
C1	171.2, C			C3	31.2, CH		2.01 (m)	C2	35.7, CH_2	a:	2.37 (m)
C2	55.2, CH		5.10 (m)	C4	21.6, CH_3		0.88 (m)			b:	n.r.
C3	37.0, CH_2	a:	1.52 (m)	C5	20.4, CH_3		0.95 (m)	C3	35.9, CH_2	a:	3.23 (m)
		b:	1.43 (m)	NMe	n.r., CH_3		3.07 (m)			b:	3.57 (m)
C4	n.r., CH		1.72 (m)	β-Alanine[7]				NH			7.20 (m)
C5	21.1, CH_3		0.81 (m)	C1	n.r., C			D-Leucine[12]			
C6	21.5, CH_3		0.92 (m)	C2	35.8, CH_2	a:	2.30 (m)	C1	n.r., C		
NMe	30.8, CH_3		n.r.			b:	2.48 (m)	C2	n.r., CH		4.92 (m)
L-Threonine[3]				C3	35.9, CH_2	a:	3.50 (m)	C3	42.0, CH_2	a:	1.22 (m)
C1	n.r., C					b:	3.20 (m)			b:	1.45 (m)
C2	55.1, CH		4.84 (m)	NH			7.19 (m)	C4	n.r., CH		1.59 (m)
C3	69.8, CH		5.17 (m)	D-*allo* Isoleucine[8]				C5	21.3, CH_3		0.86 (m)
C4	17.0, CH_3		1.07 (m)	C1	n.r., C			C6	20.3, CH_3		0.91 (m)
NH			7.75 (m)	C2	60.3, CH		4.84 (m)	NH			7.58 (m)
				C3	n.r., CH		2.26 (m)	D-Methyl *allo* Isoleucine[13]			
Methyl β-Alanine[4]				C4	24.5, CH_2	a:	1.22 (m)	C1	n.r., C		
C1	171.3, C					b:	1.45 (m)	C2	53.8, CH		4.57 (m)
C2	35.8, CH_2	a:	2.48 (m)	C5	18.4, CH_3		0.93 (m)	C3	36.15, CH		1.71 (m)
		b:	2.30 (m)	C6	17.4, CH_3		0.85 (m)	C4	26.1, CH_2	a:	1.41 (m)
C3	35.9, CH_2	a:	3.50 (m)	NH			7.89 (m)			b:	1.15 (m)
		b:	3.20 (m)	L-Methyl Valine[9]				C5	20.20, CH_3		0.92 (m)
NMe	n.r., CH_3		2.95 (m)	C1	n.r., C			C6	23.30, CH_3		0.87 (m)
								NMe	n.r., CH_3		3.08 (m)

*n.r (not resolved)

Table S14 Biological activity of the isolated compounds.

Comp.	MIC (μM)*			IC_{50} (μM) in DMEM*				PC_{50} (μM) in NDM*			
	EC	BC	KR	Panc2	HepG2	MIA PaCa-2	MCF 7	Panc2	HepG2	MIA PaCa-2	MCF 7
1	>66	>66	>66	>100	>100	>100	>100	>100	>100	33.4	>100
2	>66	>66	>66	>100	>100	>100	>100	>100	>100	42.7	>100
3	>66	>66	>66	>73.2	>73.2	>73.2	>73.2	>100	>100	9.9	>100
4	>66	>66	>66	>70.9	>70.9	>70.9	>70.9	>100	>100	7.8	>100
5	>66	>66	>66	>38.2	>38.2	>38.2	>38.2	>100	>100	8.2	>100
6	>66	>66	>66	>40.0	>40.0	>40.0	>40.0	>100	>100	3.5	>100

*MIC (Minimum Inhibitory Concentration); PC (Preferential Cytotoxic); IC (Inhibitory Concentration).

Supplementary Information for Chapter 3

Table of Contents:

Material and Methods

General Procedures Optical rotations were measured on a Jasco P-1030 polarimeter. Infrared spectra were measured on a Jasco FT/IR 4100. NMR spectra were recorded on a JEOL ECX 500 (500 MHz) or JEOL ECZ 400 (400 MHz) spectrometer. ESI-MS spectra were recorded on a Thermo Scientific Exactive mass spectrometer or a SHIMADZU LCMS-2020 spectrometer. High performance liquid chromatography (HPLC) experiments were performed with a SHIMADZU HPLC system equipped with a LC-20AD intelligent pump. LC-MS experiments were performed with amaZon SL-NPC (Bruker Daltonics). GC/MS experiments were performed on SHIMADZU-QP2010 Ultra. All reagents were used as supplied unless otherwise stated.

Ichip-based Cultivation Method

Preparation of Environmental Sample. Adequate environmental samples (soils, mud, and stream water) were resuspended with 20 ml of sterilized miliQ water following vigorous mixed with vortex for 10 min to detach any cell from the soil debris. These mixtures were used as stock solutions for subsequent dilution steps. The prepared stock solutions were diluted until into 1:10^4 times dilution with sterilized miliQ water. Another dilution step was performed in a low melting point (LMP) agar fortified with 0.1-time diluted R2A media. This sample was kept at 30 °C water-bath to prevent solidification.

Implantation of Bacterial Cells into Ichip. The middle array of a sterilized ichip containing hundreds of 1 mm diameter through-holes was dipped into the prepared mixture of bacterial cells and LMP agar. Subsequently, two pairs of membranes (0.02 μm) were used to cover the array in both sides before re-assembling with the bottom and upper part of the ichip device. The assembled device was tightened with screws to avoid contaminations. After completing the assembly process, the devices were returned to their original sampling sites and incubated for about two weeks or more.

Retrieval of Ichip-domesticated Bacterial Strains. After *in-situ* incubation, the ichip devices were collected from the incubation sites and directly processed. Initially, the devices were cleaned by carefully rinsing with miliQ water to remove any attached debris and disassembled inside a sterile hood. Subsequently, the plug agar inside the middle array of the device was

individually retrieved using a sterilized paper clip into prepared 0.1-time diluted R2A plate. The plates were incubated at 30 °C until colony formed. Further bacterial isolation steps were performed if several different strains were observed.

Antibacterial Screening To prepare the sample solutions, as much as 2 µl of ichip strains library were inoculated to 120 µl of 2x R2A in 96 well round bottom and incubated with shake at 30°C for overnight. The seed culture (20 µl) was then used for inoculation into 450 µl Modified M9 media or LB media in a 96 deep-well plate and incubated at 30 °C for one week. Upon incubation, the culture was centrifuged 3500 r.p.m. for 30 minutes and 100 µl of supernatant was transferred to a 96 well flat bottom. Bacterial tests were prepared by inoculating an overnight culture into a fresh MHB media to final OD_{600nm} of 0.1. The MHB media contain bacterial test was added into previously prepared sample solutions followed by incubation at 30 °C (without shake) for 24 hours. After incubation, the cell growth was examined by measuring OD_{600nm} in a plate reader. Ichip strains that inhibit the growth of bacterial tests more than 50% are considered to be potential strains.

CAS assay screening Chrome Azurol S (CAS) assay was used to screen potential strains for producing siderophore. To prepare the sample solutions, as much as 2 µl of ichip strains library was inoculated to 120 µl of 2x R2A in 96 well round bottom and incubated with shake at 30 °C for overnight. The seed culture (20 µl) was then used for inoculation into 450 µl modified minimal M9 media (or other) in a 96 deep-well plate and incubated at 30 °C for one week. Upon incubation, the culture was centrifuged 3500 r.p.m. for 30 minutes and 100 µl of supernatant was transferred to a 96 well flat bottom. CAS solution was made of three solutions (A, B, and C). Solution A was prepared by dissolving 21.9 mg of HDTMA in 25 ml MiliQ water; Solution B was prepared by combining 1.5 ml of 1 mM $FeCl_3.6H_2O$ in 10 mM HCl and 7.5 ml of 2 mM CAS in water; and Solution C was prepared by dissolving 9.76 g of MES in 50 ml MiliQ water and adjusting pH to 5.6 with 50% KOH. While stirring, 9 ml of solution B was added to 25 ml of solution A slowly followed by addition 50 ml of solution C and 16 ml of MiliQ water was added to bring the total volume of 100 ml. Immediately before use 86.4 mg of 5-sulfosalicylic acid was added to the above mixture and stirred for a few seconds. To perform the assay, into previously prepared sample solutions, 100 µl of CAS solution was added followed by briefly mixing with Micromixer and incubation for 1 hour in dark at room temperature. After incubation, UV 630 nm absorption of reaction mixtures was checked by

plate reader. To analyze the data, UV 630 nm absorption of blank was compared with bacterial strains. Bacterial strains that showed score more than 1.1 are considered to be potential strains.

Metabolite Production and Isolation Natural Products from *Variovorax* sp. H002 The overnight seed culture of ichip-domesticated *Variovorax* sp. H002 was inoculated into 100 ml fresh modified minimal M9 media (M9 salt, 1% glucose, 1% casamino acid) with additional HP20 resins in 40 baffled flasks. The cultures were incubated with shake at 30 °C for seven days. After cultivation, the cultures were filtered out with cotton leaving the HP20 resins, which were packed to a column for chromatography. The packed resins were washed with water several times followed by elution with excessive methanol. The collected methanol fraction was then concentrated under vacuo to obtain a crude methanolic extract. Subsequently, the crude extract was subjected to gel-filtration chromatography packed with Sephadex LH-20® and methanol as mobile phase. The collected fractions were monitored by ESI-MS. Fractions containing target compounds were concentrated and subjected to a semi-preparative HPLC equipped with Cosmosil® $5C_{18}$-MS-II (10 mm ID × 250 mm) using gradient systems from 30% to 70% acetonitrile for 30 minutes to yield compounds **7** (3 mg), **8** (2 mg), **9** (5 mg), **10** (20 mg), and **11** (10 mg).

Preparation of Fatty Acid Methyl Ester (FAME) Derivatives Adequate amount of **7-11** was hydrolyzed with 6 M HCl at 110 °C for overnight. The obtained hydrolysates were freeze dried and resuspended with 100 µl of methanol followed addition of 100 µl of trimethylsilyldiazomethane. The reactions were kept at room temperature for one hour. The reaction mixtures were concentrated under vacuo and dissolved with acetone prior to subjecting into GC/MS, Zebron ZB-WAX (30 m × 0.25 mm ID × 0.25 µm), using He as gas source with flow rate 1.41 mL min^{-1}. The initial temperature (120 °C) was hold for one minute follow gradient to 230 °C at a rate of 10 °C min^{-1} and hold for six minutes.

Preparation of 2-alkenyl-4,4-Dimethyloxazoline (DMOX) Fatty Acid Derivatives Adequate amount of **7-11** was hydrolyzed with 6 M HCl at 110 °C for overnight. The obtained hydrolysates were treated with 250 µl of 2-amino-2-methyl propanol in screw-capped vials. After flushing with N_2, the reactions were sealed and heated in 180 °C for 18 hours. After completed, the reactions were cooled down and resuspended with 3 ml of dichloromethane followed by washing with three times with1 ml of water. The organic layer was dried with anhydrous Na_2SO_4 and concentrated under vacuo. The samples were redissolved with acetone prior to subjecting into GC/MS with aforementioned system.

Determination of Amino Acid Configuration by Marfey's Method. 100 µg of samples were acid-hydrolyzed with 6 M HCl at 110 °C for overnight. The obtained hydrolysates were freeze dried and mixed with H_2O (160 µL), saturated $NaHCO_3$ (200 µL), and 1% L-FDLA in acetone (160 µL). The mixtures were heated at 40 °C for 30 minutes. Subsequently, the reaction was quenched with 1M HCl (200 µL) and concentrated under vacuo. The residue was dissolved with DMSO (400 µL) and passed through membrane filter before injecting to LC-MS, Cosmosil® $5C_{18}$-MS-II (2 mm ID × 100 mm) (oven 40 °C ; flowrate 0.2 ml min^{-1}) using gradient program from 10% to 60% MeCN in water (with additional formic acid 0.1%) for 45 minutes.

Genome DNA extraction. To extract genome DNA from interest strain especially gram-negative bacteria, an overnight culture from single colony bacteria was harvested by spinning down at 8,000 r.p.m. for 10 minutes. Obtained bacterial cell-pellet was resuspended with 800 µl of EDTA-saline solution and transferred to a 2 ml tube. Cell wall lysis was performed by adding 10 µl of 300 µg/ml lysozyme followed by brief vortexing and incubating at 37 °C for 15-45 minutes in water bath. While incubation, the suspension was briefly vortexed every 15 minutes. Subsequently, to disturb the cell membranes, 80 µl of 25 % w/v of SDS-solution was added followed by mixing thoroughly until the viscosity increased and incubated at 65 °C for 10 minutes. During this incubation, the suspension was vortexed one time. The suspension was spun down briefly and 250 µl of 5 M NaCl solution was added followed by brief vortexing. Insoluble cell debris and protein contamination was removed by, at least, three rounds of 400 µl of phenol:chloroform:isoamyl alcohol (25:24:1) extraction. Finally, 100 µl of 3 M sodium acetate pH 5 and 700 µl of isopropanol alcohol was added to the 1 ml of contamination-free suspension, followed by mixing and centrifugation at 15,000 rpm for 30 minutes at 4 °C. The resulting DNA pellet was washed two rounds with pre-cold 70% ethanol to remove trace amount of chemical contamination and air-dried for about one minute to let all the ethanol evaporate. The dried DNA pellet was dissolved with nuclease free water and stored at 4 °C for temporary or -30 °C for long term storage.

Genome Sequencing The library preparation and sequencing experiments employing third generation sequencing technology, Oxford Nanopore Technology, have been done following a protocol from ligation sequencing kit SQK-LSK109 from Oxford Nanopore Technology. The obtained reads were trimmed with Porechop algorithm to yield refined raw reads from long read sequencing technology. Additionally, short read sequencing experiments using next

generation sequencing technology have been done by collaboration with a third party. The obtained reads from short and long read sequencing technologies were assembled using hybrid de novo assembly with SPAdes algorithm.

Construction of *varH002g* null mutant strain A plasmid construct for homologous recombination was prepared by amplifying the upstream and downstream region of *varH002g* using pairs of listed primers. Additionally, a chloramphenicol resistant gene was used to replace the targeted *varH002g* gene. All the amplified fragments were assembled following Gibson assembly protocol in a pRED backbone. The assembled fragment was introduced to *Escherichia coli* DH5α for plasmid production. The resulted plasmid pRED-Δ*varH002g* and a conjugative pGM160 were digested with HindIII and ligated, followed by transformation into the *E. coli* S17-1 λ pir. The obtained *E. coli* S17-1 was used to transform pGM160-Δ*varH002g* into *Variovorax* sp. H002. Both *Variovorax* sp. H002 and *E. coli* S17-1 were cultured overnight in 2xR2A and LB at 30 °C. The overnight cultures were washed with 2xR2A three times, and the pellets were resuspended with 500 µl of 2xR2a. The OD_{600} values of donor (*E. coli* S17-1) and recipient cells (*Variovorax* sp. H002) were set to 2.0 and mixed in 1:1 ratio. The mixtures were spotted to 2xR2A followed by incubation at 30 °C for two days. The colonies were scraped up and spread on 2xR2A agar plate with chloramphenicol (30 µg/ml) and kanamycin (50 µg/ml). The plates were incubated at 30 °C until colony formed. Several conjugants were picked and checked by PCR to confirm the insertion site.

Metabolite Production and Isolation Natural Products from *Variovorax* sp. B014 The overnight culture of *Variovorax* sp. B014 was inoculated onto 500 ml baffled flask containing 100 ml fresh modified minimal M9 media (M9 salt, 1% glucose, and 1% casamino acid). The cultures were incubated with shake at 30 °C for a week. After cultivation, the cultures were spun down and subsequent subjected to a pre-washed ODS column with methanol and water. The retain metabolites were washed with water followed by elution of increasing methanol concentration in water. The fractions containing target compounds judged by ESI-MS analysis were concentrated under vacuo to obtain a semi-crude extract. This extract was further purified with a semi-preparative HPLC equipped with Cosmosil® $5C_{18}$-MS-II (10 mm ID × 250 mm) using gradient systems from 20% to 70% acetonitrile for 30 minutes to yield compounds **12** (10 mg), **13** (6 mg), **14** (6 mg), **15** (12 mg), and **16** (8 mg).

Preparation of complex 13-Ga(III) The mixture of compound **13** dissolved in water and 8-hydroxy quinoline in methanol was stirred at room temperature for overnight. The organic solvent was then evaporated under vacuo followed by extraction with equal volume of dichloromethane for three times. The water layer was then freeze-dried followed by resuspension with phosphate buffer solution containing Ga(III) sulfate. The mixture was stirred at room temperature for overnight. Subsequently, the mixture was extracted with n-butanol for three times. The obtained organic layer was concentrated under vacuo and subjected to semi-preparative HPLC equipped with Cosmosil® $5C_{18}$-MS-II (10 mm ID × 250 mm) using 35% acetonitrile in water for 30 minutes to yield compound **13**-Ga(III), HRESIMS; *m/z* 985.4013 $[M+H]^+$ Calc. $C_{40}H_{68}N_8O_{16}Ga$.

Methanolysis of 13 As much as 1 mg of **13** was stirred with 200 µl of 7 M ammonium solution in methanol at 100 °C for one hour. The mixture was concentrated under vacuo and redissolved with small methanol. Subsequently, the mixture was subjected to HPLC equipped with Cosmosil® $5C_{18}$-MS-II (10 mm ID × 250 mm) using gradient systems from 20% to 70% acetonitrile for 30 minutes to yield a *seco*-acid methyl ester product of **13**, HRESIMS; *m/z* 951.52 $[M+H]^+$ Calc. $C_{41}H_{74}N_8O_{17}$.

Antibacterial Activity The assay for all compounds has been done in triplicate according to the Clinical Laboratory Standards Institute testing standard in a 96-well plate microbroth dilution assay. Compounds were tested against gram negative-bacteria (*Escherichia coli* JW5503 and *Bukholderia multivorans* NBRC 102086) and gram-positive bacteria (*Bacillus cereus* NBRC 15305 and *Kocuria rhizophila* NBRC 12708). After incubation for 22 h (30 °C), the obtained optical density was recorded using plate reader at 600 nm to determine percent growth inhibition.

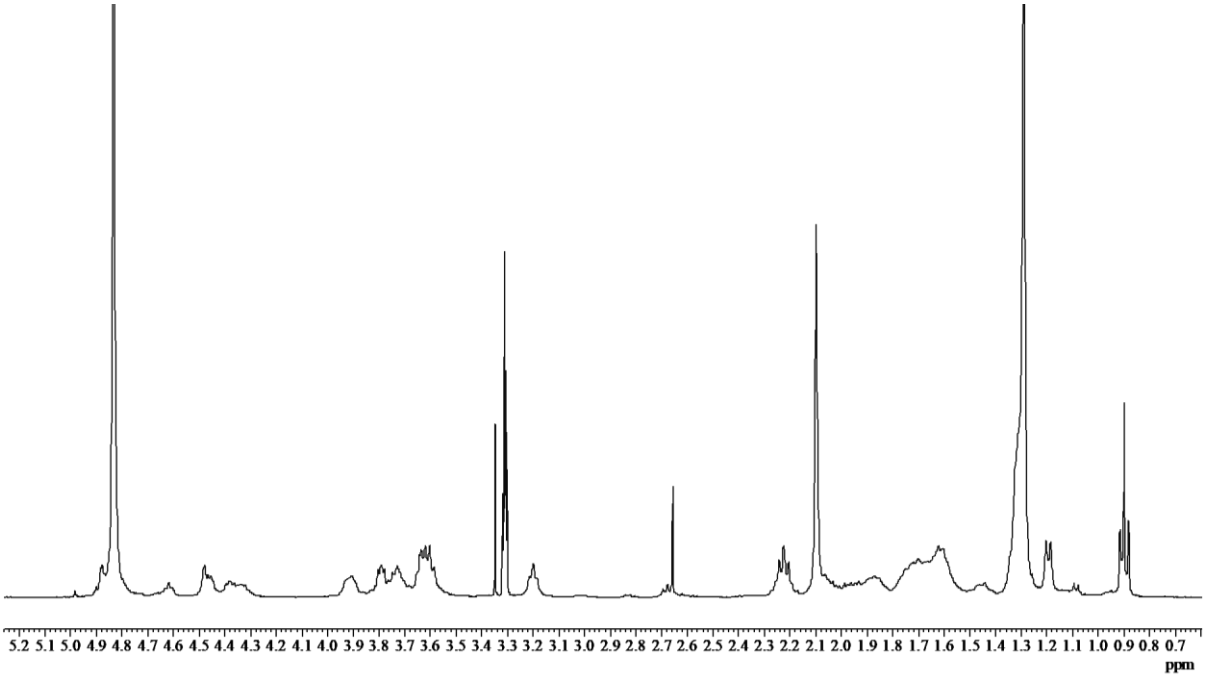


Figure S3.1 ^{1}H NMR Spectrum of **10** (methanol-d_4, 400 MHz)

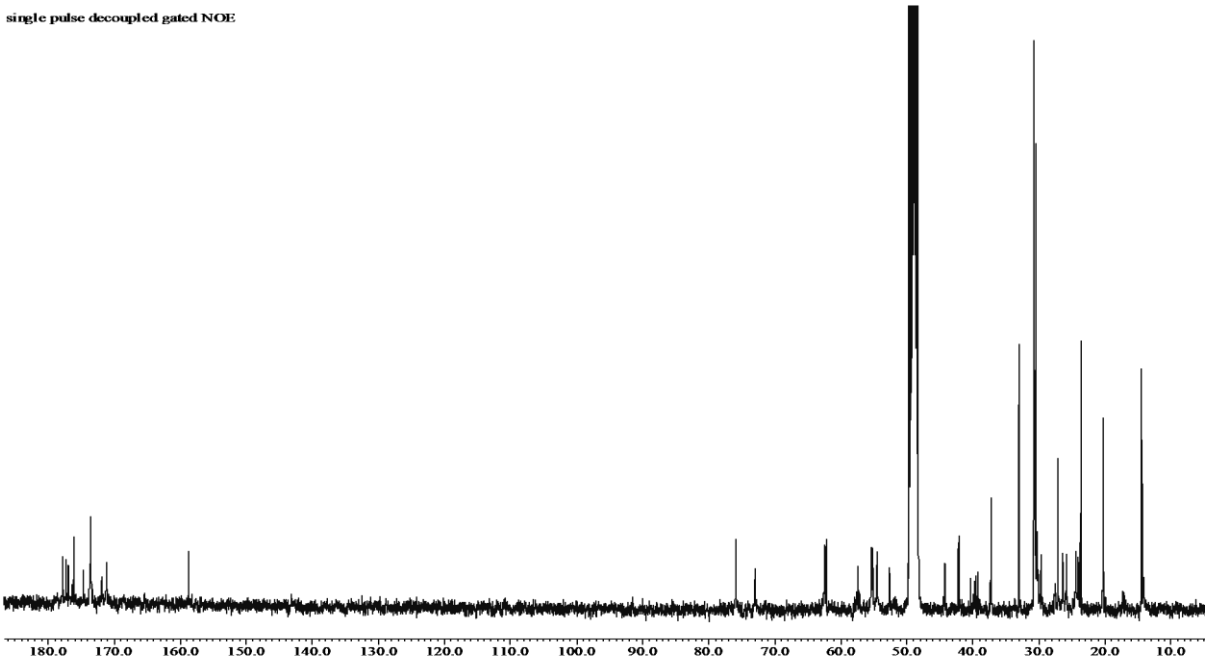


Figure S3.2 ^{13}C NMR Spectrum of **10** (methanol-d_4, 400 MHz)

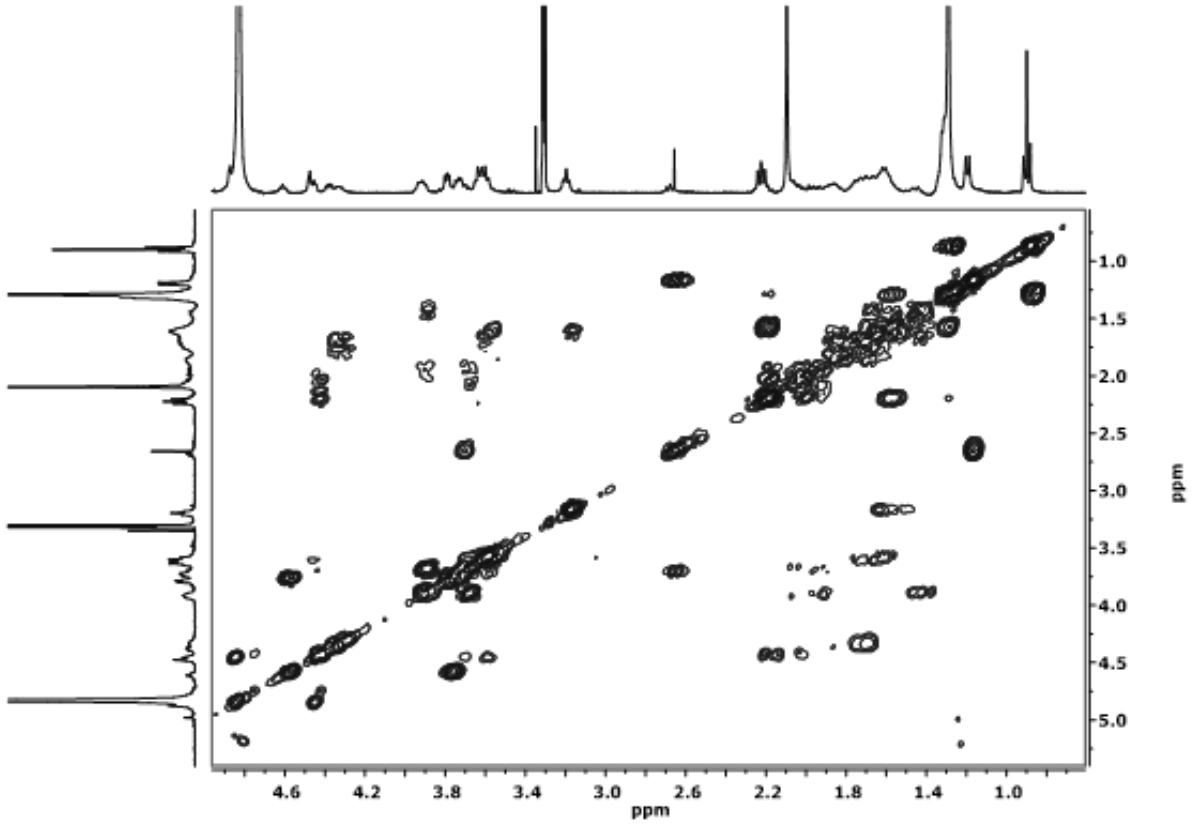


Figure S3.3 COSY Spectrum of **10** (methanol-d_4, 400 MHz)

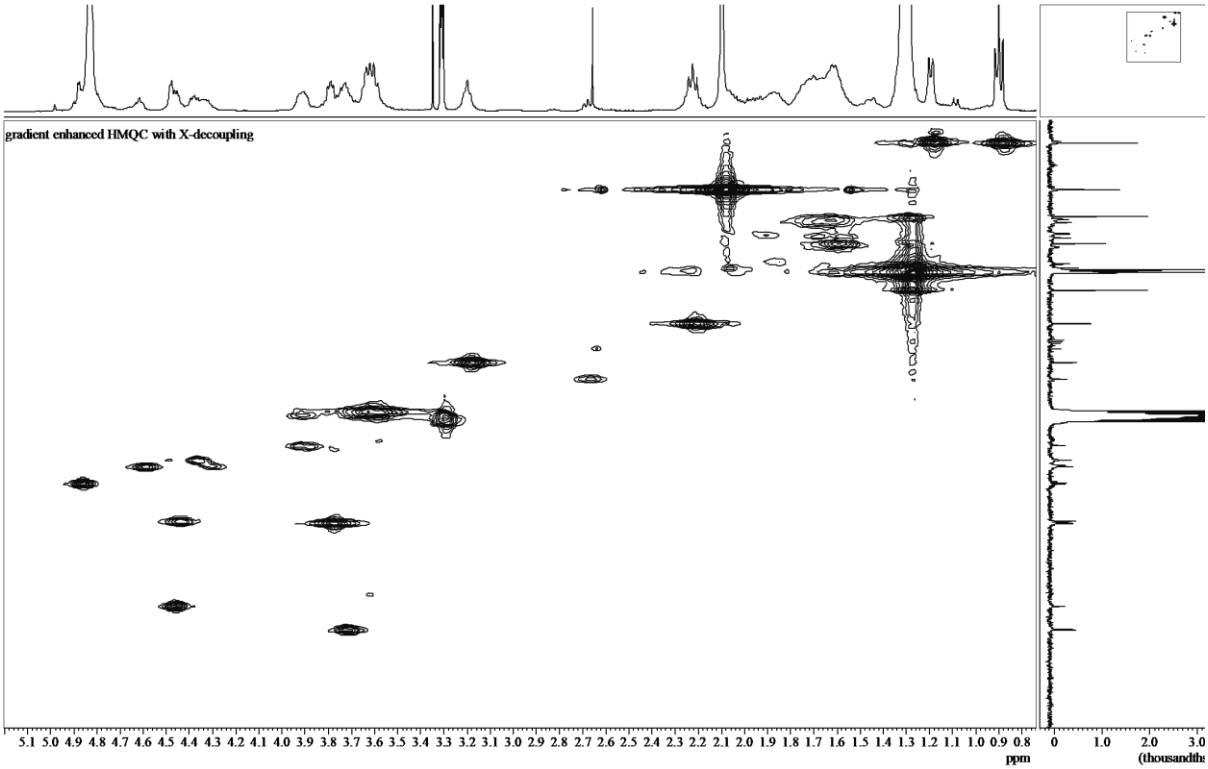


Figure S3.4 HSQC Spectrum of **10** (methanol-d_4, 400 MHz)

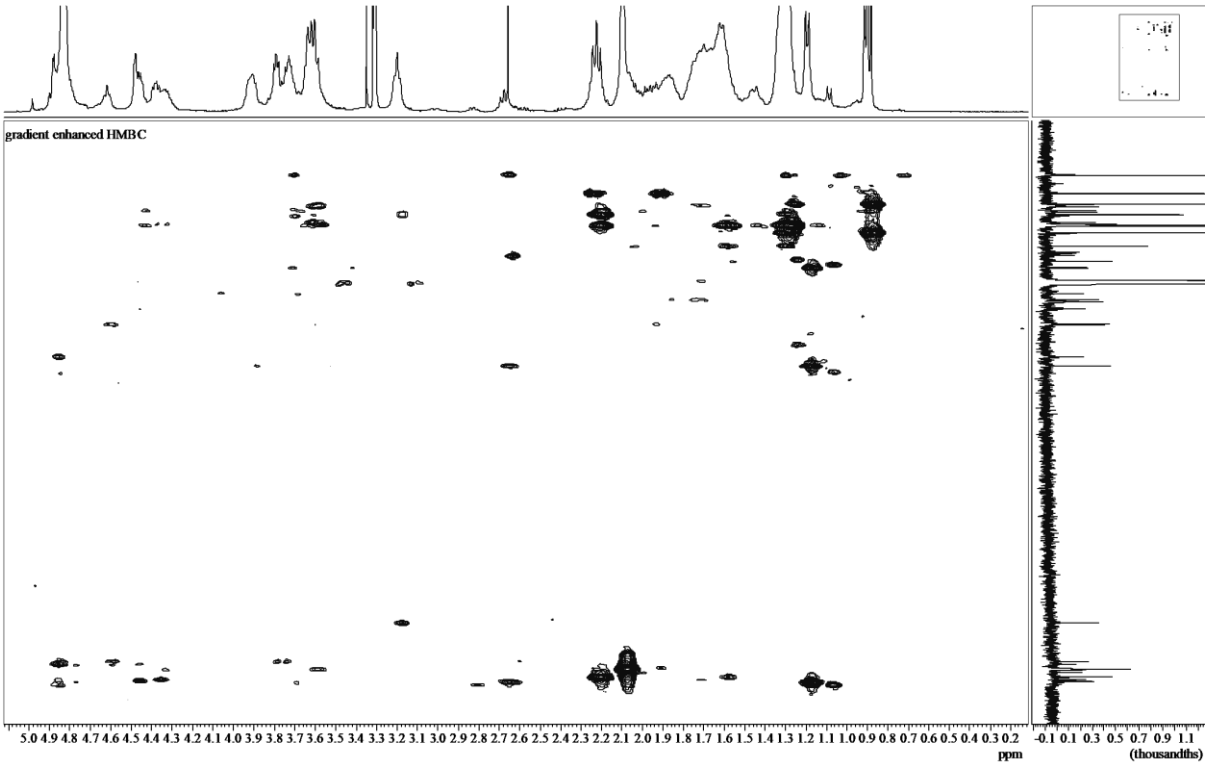
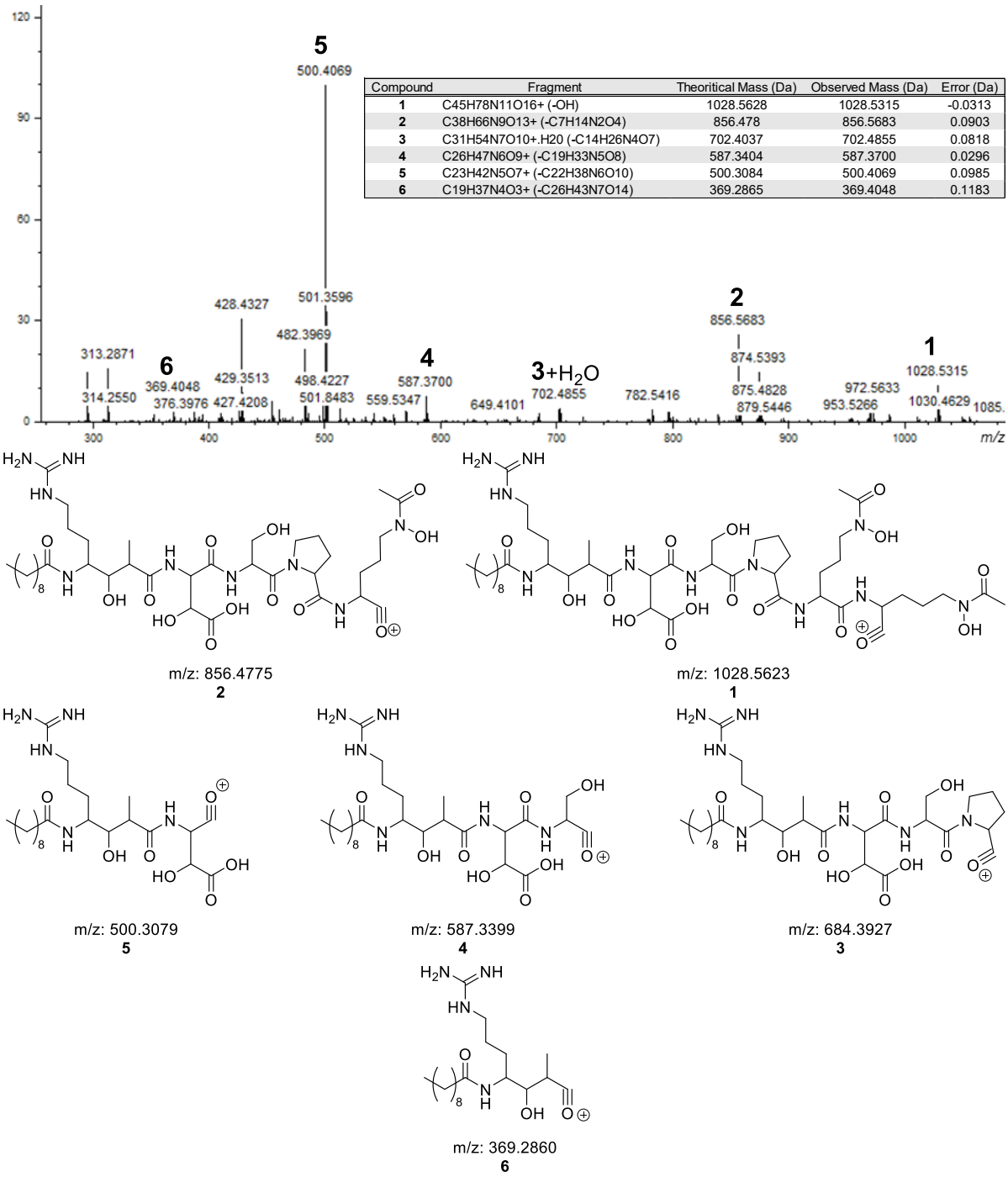
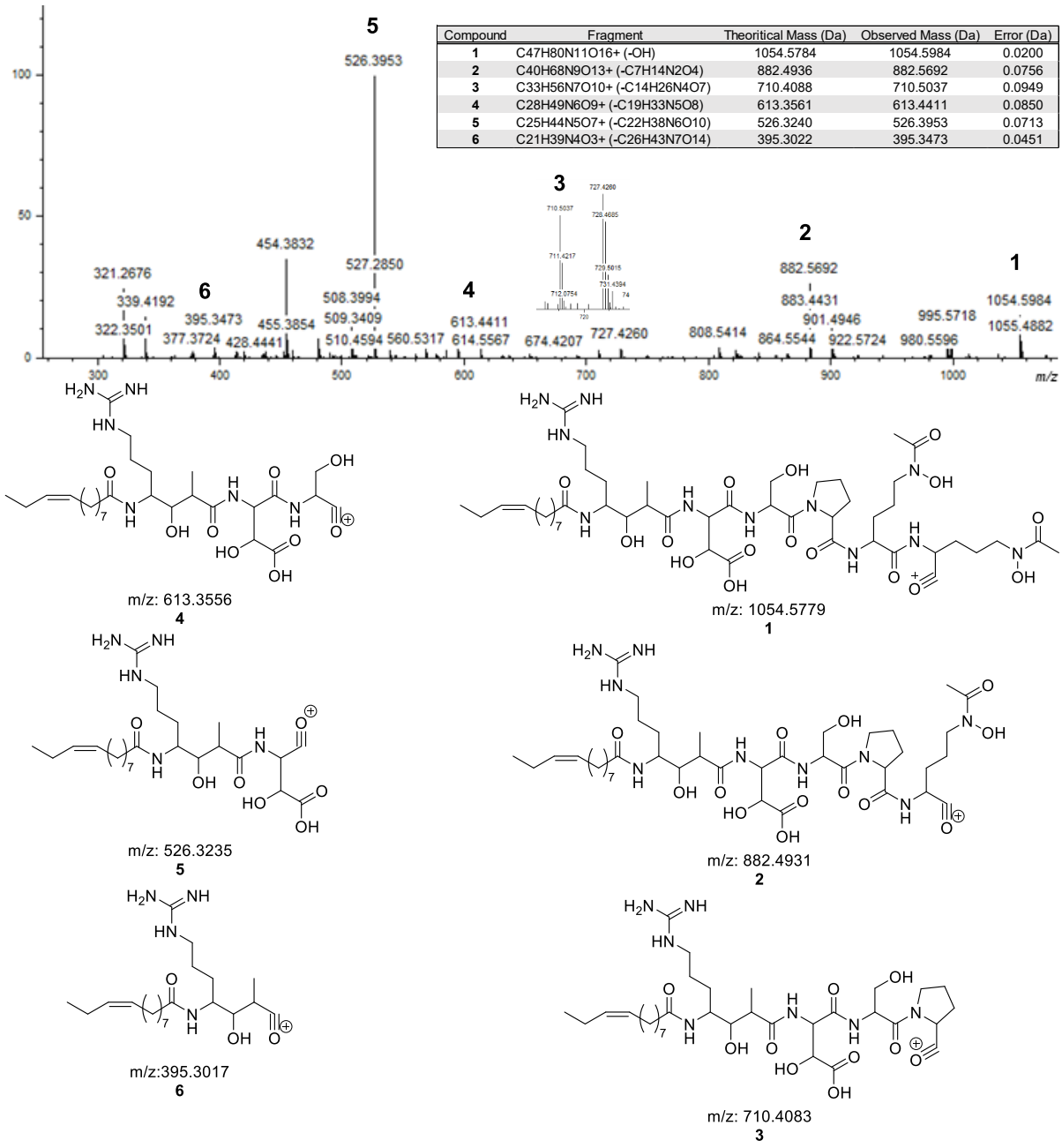


Figure S3.5 HMBC Spectrum of **10** (methanol-d_4, 400 MHz)



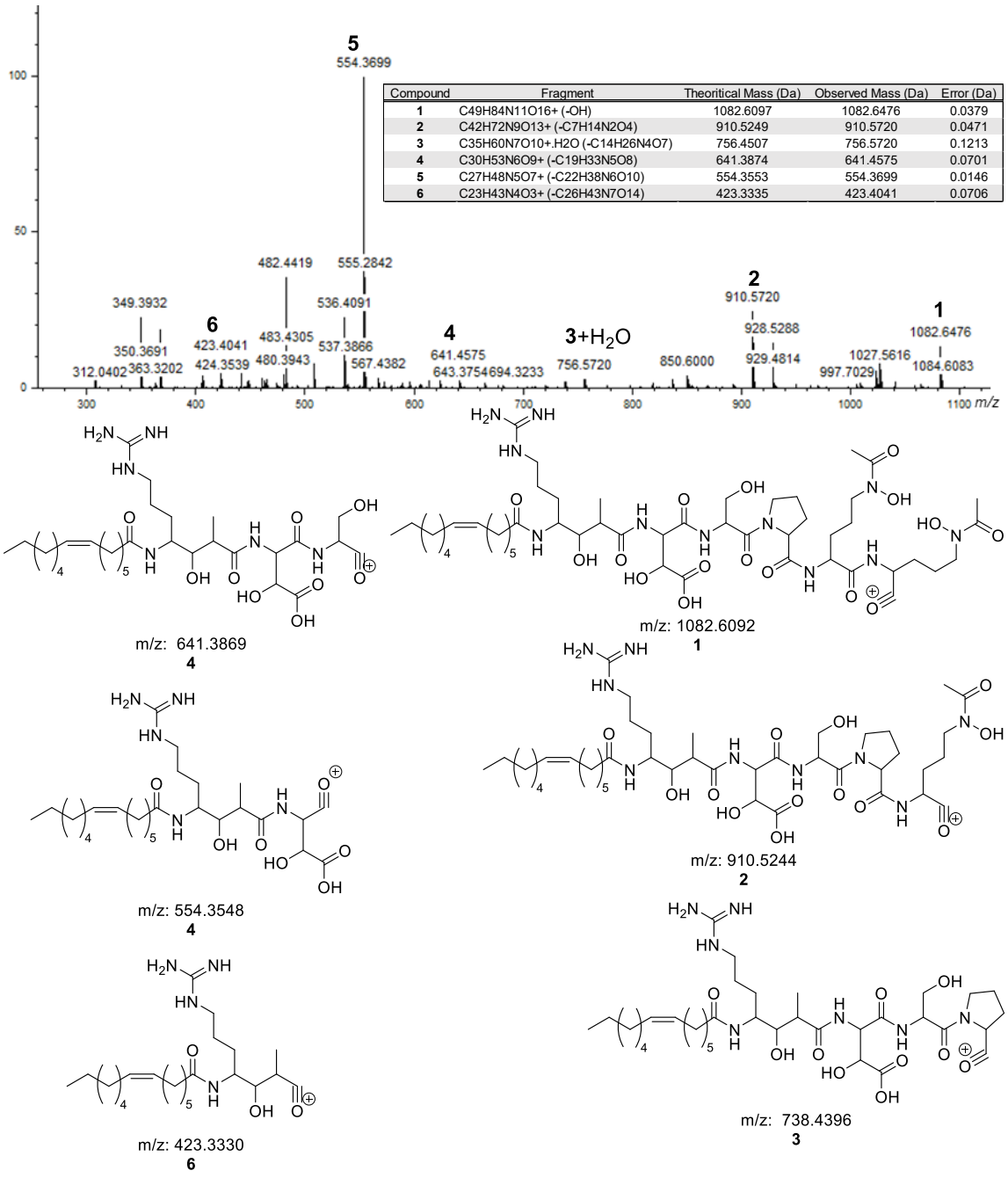
Compound	Fragment	Theoritical Mass (Da)	Observed Mass (Da)	Error (Da)
1	C45H78N11O16+ (-OH)	1028.5628	1028.5315	-0.0313
2	C38H66N9O13+ (-C7H14N2O4)	856.478	856.5683	0.0903
3	C31H54N7O10+.H20 (-C14H26N4O7)	702.4037	702.4855	0.0818
4	C26H47N6O9+ (-C19H33N5O8)	587.3404	587.3700	0.0296
5	C23H42N5O7+ (-C22H38N6O10)	500.3084	500.4069	0.0985
6	C19H37N4O3+ (-C26H43N7O14)	369.2865	369.4048	0.1183

Figure S3.6 MS/MS analysis of compound **7**



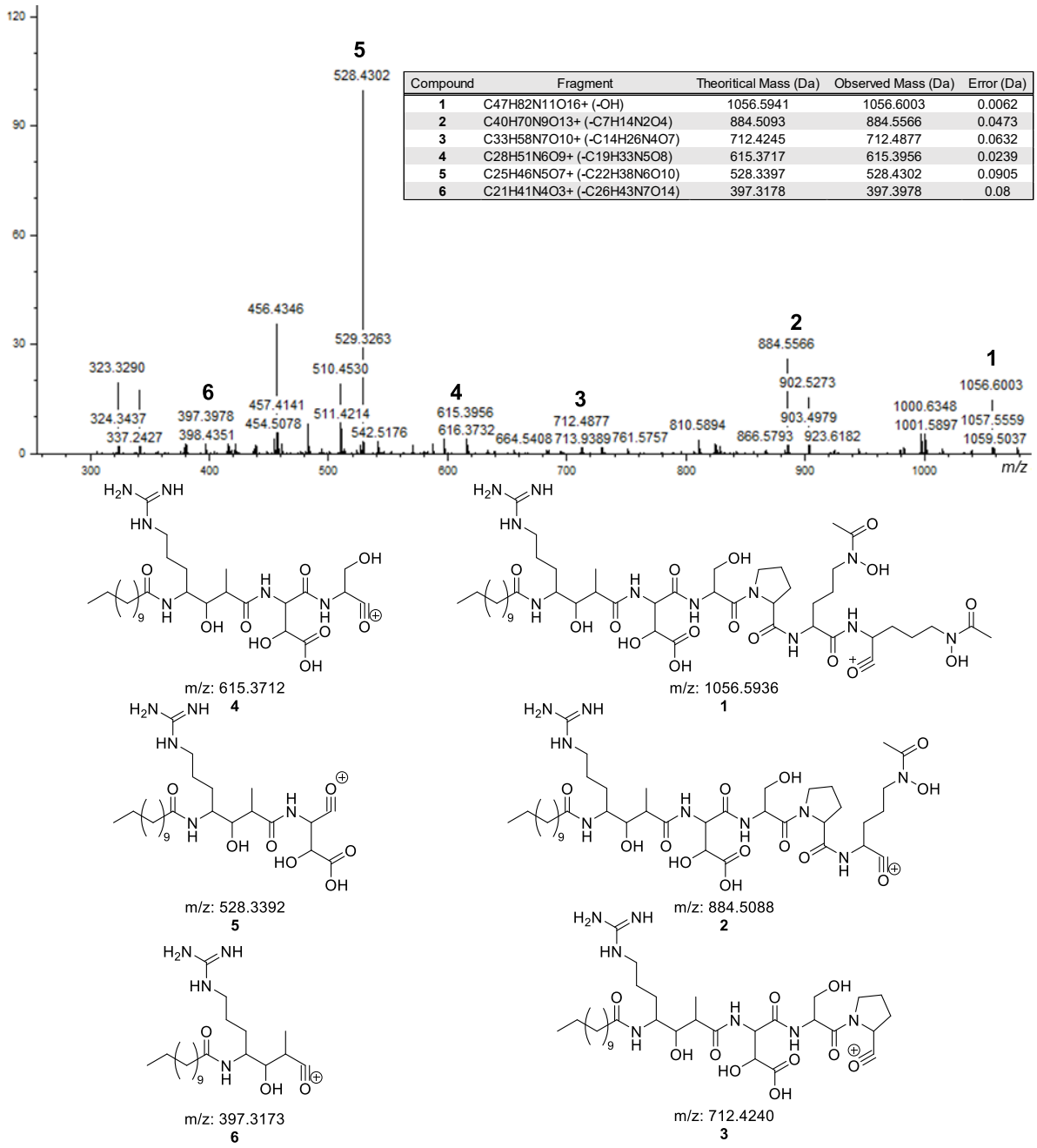
Compound	Fragment	Theoritical Mass (Da)	Observed Mass (Da)	Error (Da)
1	C47H80N11O16+ (-OH)	1054.5784	1054.5984	0.0200
2	C40H68N9O13+ (-C7H14N2O4)	882.4936	882.5692	0.0756
3	C33H56N7O10+ (-C14H26N4O7)	710.4088	710.5037	0.0949
4	C28H49N6O9+ (-C19H33N5O8)	613.3561	613.4411	0.0850
5	C25H44N5O7+ (-C22H38N6O10)	526.3240	526.3953	0.0713
6	C21H39N4O3+ (-C26H43N7O14)	395.3022	395.3473	0.0451

Figure S3.7 MS/MS analysis of compound **8**



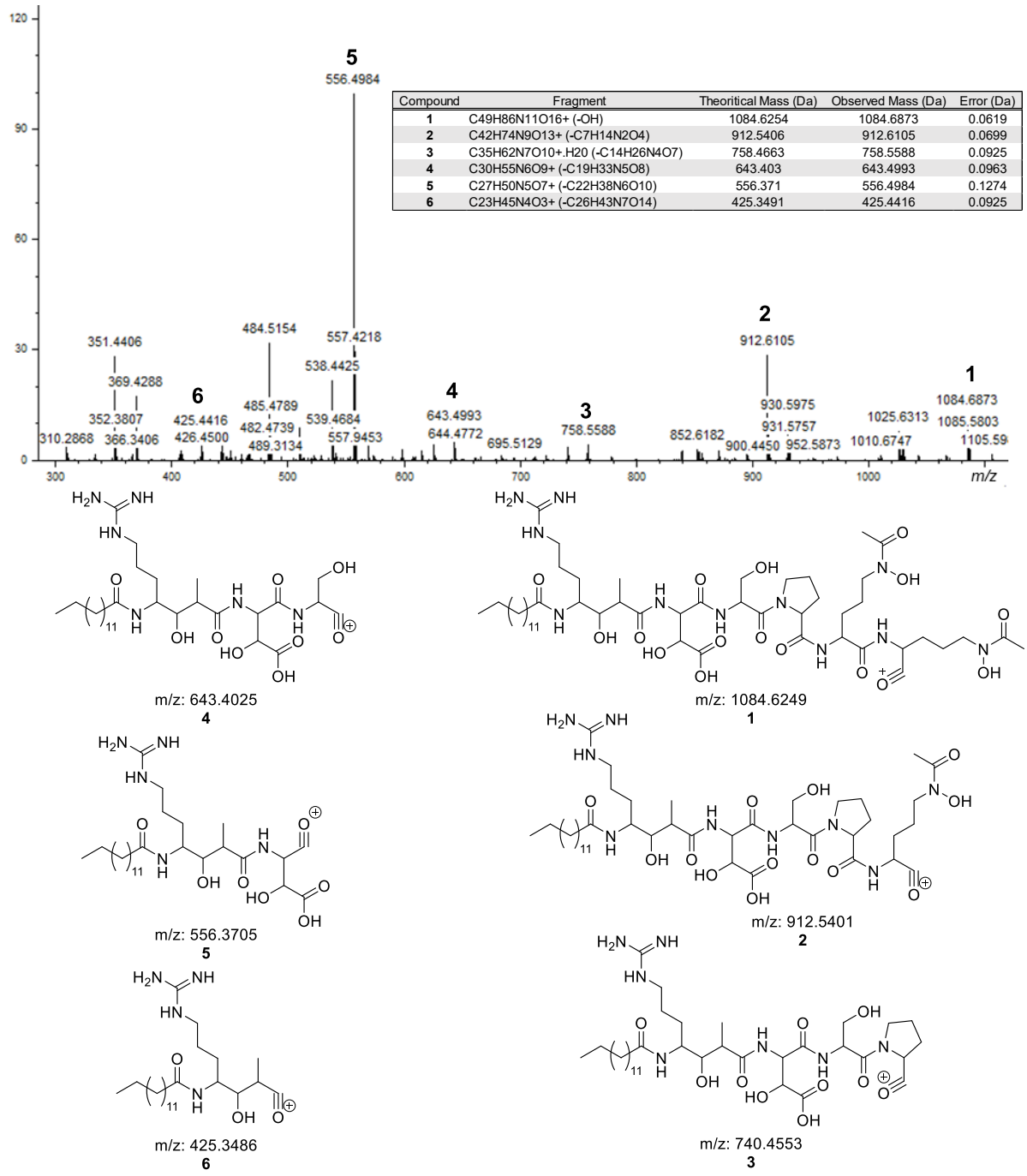
Compound	Fragment	Theoritical Mass (Da)	Observed Mass (Da)	Error (Da)
1	C49H84N11O16+ (-OH)	1082.6097	1082.6476	0.0379
2	C42H72N9O13+ (-C7H14N2O4)	910.5249	910.5720	0.0471
3	C35H60N7O10+.H2O (-C14H26N4O7)	756.4507	756.5720	0.1213
4	C30H53N6O9+ (-C19H33N5O8)	641.3874	641.4575	0.0701
5	C27H48N5O7+ (-C22H38N6O10)	554.3553	554.3699	0.0146
6	C23H43N4O3+ (-C26H43N7O14)	423.3335	423.4041	0.0706

Figure S3.8 MS/MS analysis of compound **9**



Compound	Fragment	Theoritical Mass (Da)	Observed Mass (Da)	Error (Da)
1	C47H82N11O16+ (-OH)	1056.5941	1056.6003	0.0062
2	C40H70N9O13+ (-C7H14N2O4)	884.5093	884.5566	0.0473
3	C33H58N7O10+ (-C14H26N4O7)	712.4245	712.4877	0.0632
4	C28H51N6O9+ (-C19H33N5O8)	615.3717	615.3956	0.0239
5	C25H46N5O7+ (-C22H38N6O10)	528.3397	528.4302	0.0905
6	C21H41N4O3+ (-C26H43N7O14)	397.3178	397.3978	0.08

Figure S3.9 MS/MS analysis of compound **10**



Compound	Fragment	Theoritical Mass (Da)	Observed Mass (Da)	Error (Da)
1	C49H86N11O16+ (-OH)	1084.6254	1084.6873	0.0619
2	C42H74N9O13+ (-C7H14N2O4)	912.5406	912.6105	0.0699
3	C35H62N7O10+.H2O (-C14H26N4O7)	758.4663	758.5588	0.0925
4	C30H55N6O9+ (-C19H33N5O8)	643.403	643.4993	0.0963
5	C27H50N5O7+ (-C22H38N6O10)	556.371	556.4984	0.1274
6	C23H45N4O3+ (-C26H43N7O14)	425.3491	425.4416	0.0925

Figure S3.10 MS/MS analysis of compound **11**

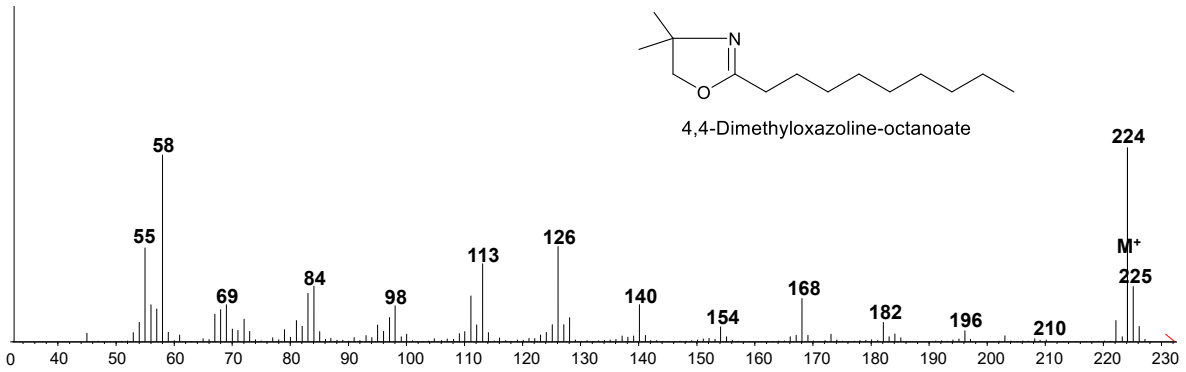


Figure S3.11 GC/MS analysis of DMOX derivatized fatty acid moiety of **7**

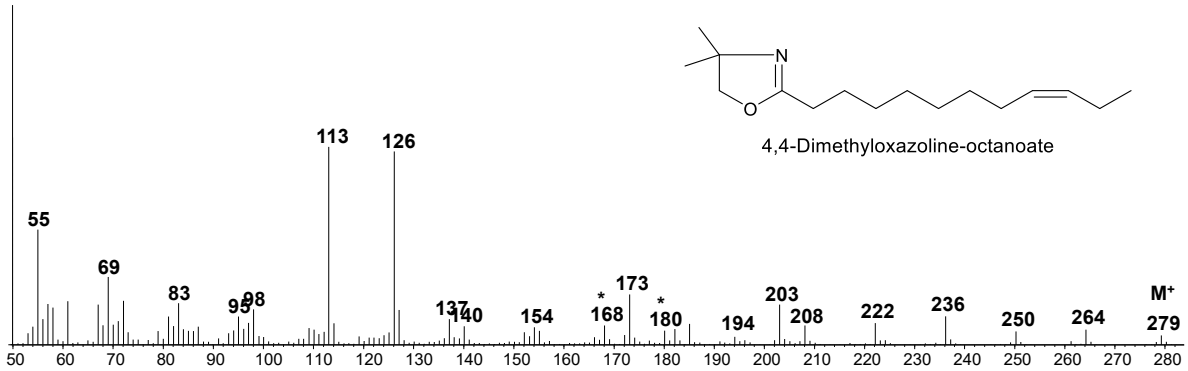


Figure S3.12 GC/MS analysis of DMOX derivatized fatty acid moiety of **8**

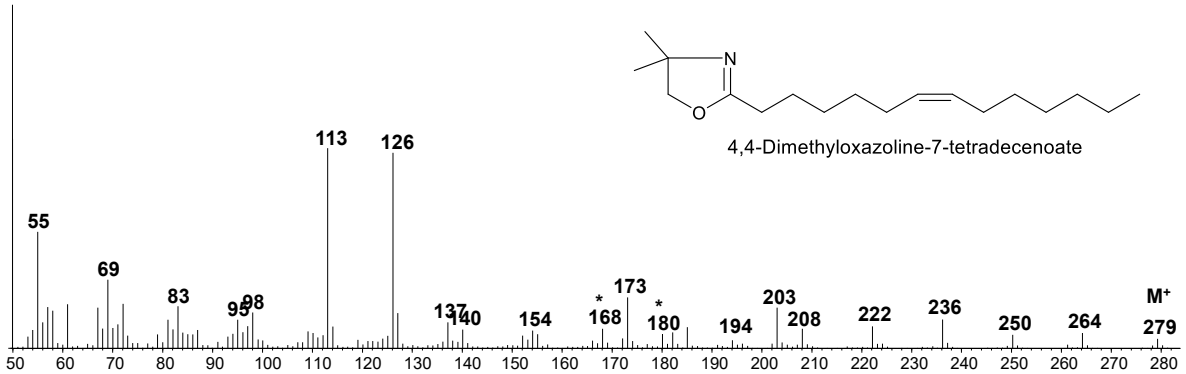


Figure S3.13 GC/MS analysis of DMOX derivatized fatty acid moiety of **9**

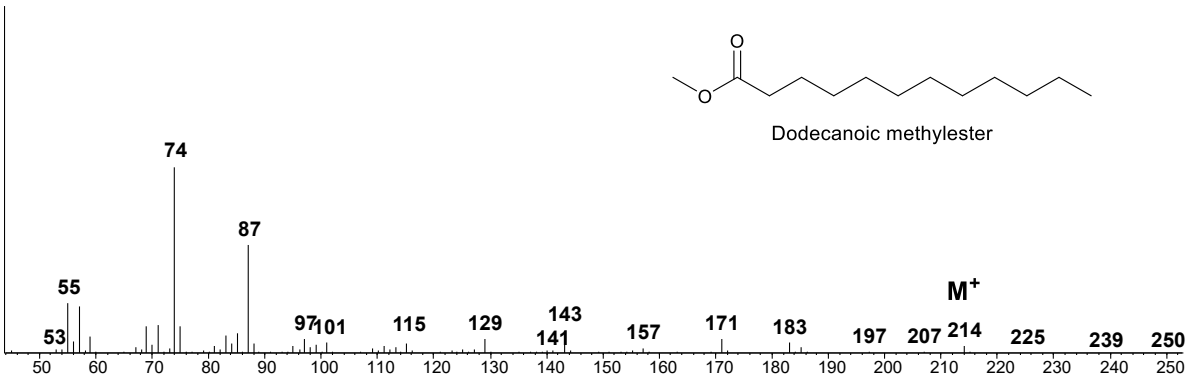


Figure S3.14 GC/MS analysis of FAME derivatized fatty acid moiety of **10**

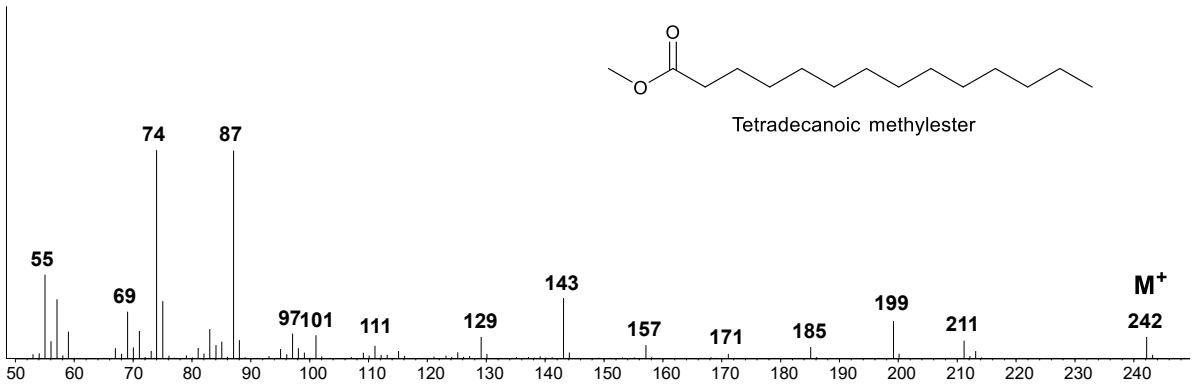


Figure S3.15 GC/MS analysis of FAME derivatized fatty acid moiety of **11**

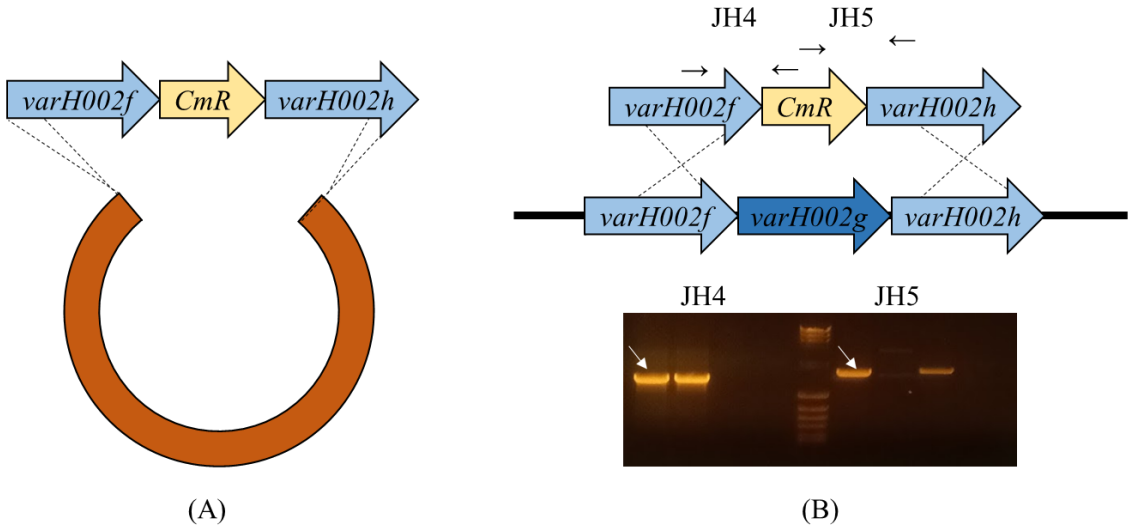


Figure S3.17 Reconstruction of null *varH002g* of *Variovorax* sp. H002; (A) Reconstruction of plasmid pGM160::pJH2; and (B) Homologous recombination and successful confirmation by PCR amplification.

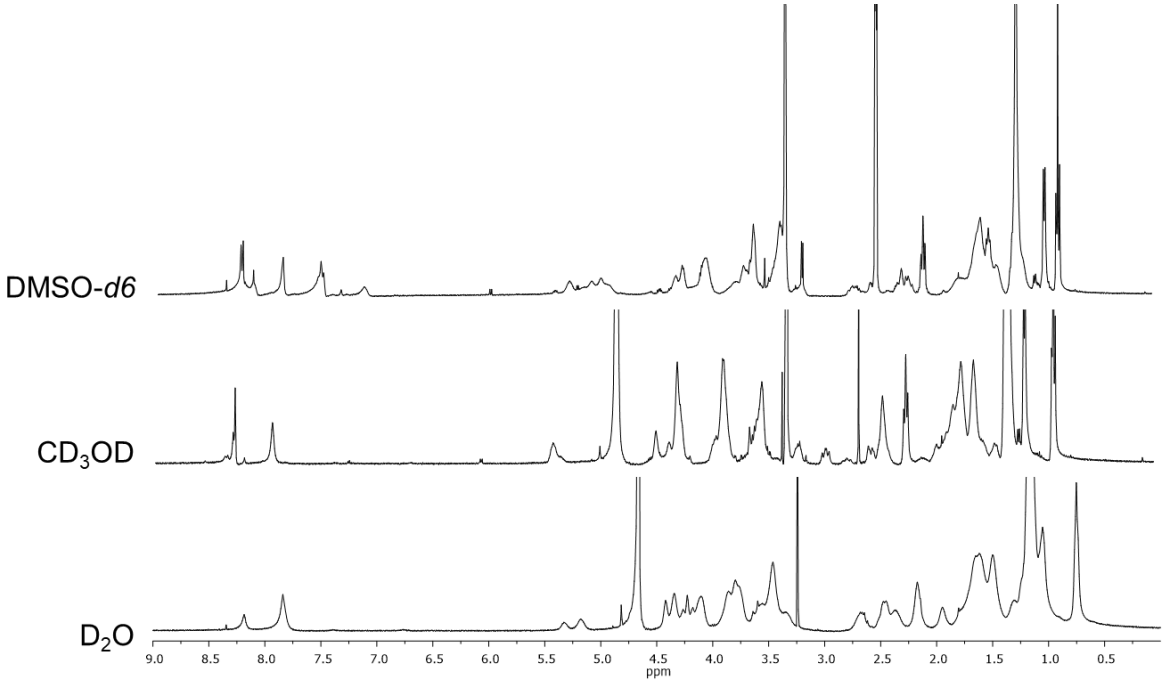


Figure S3.18 ^{1}H NMR Spectra of **13** in several deuterated solvent

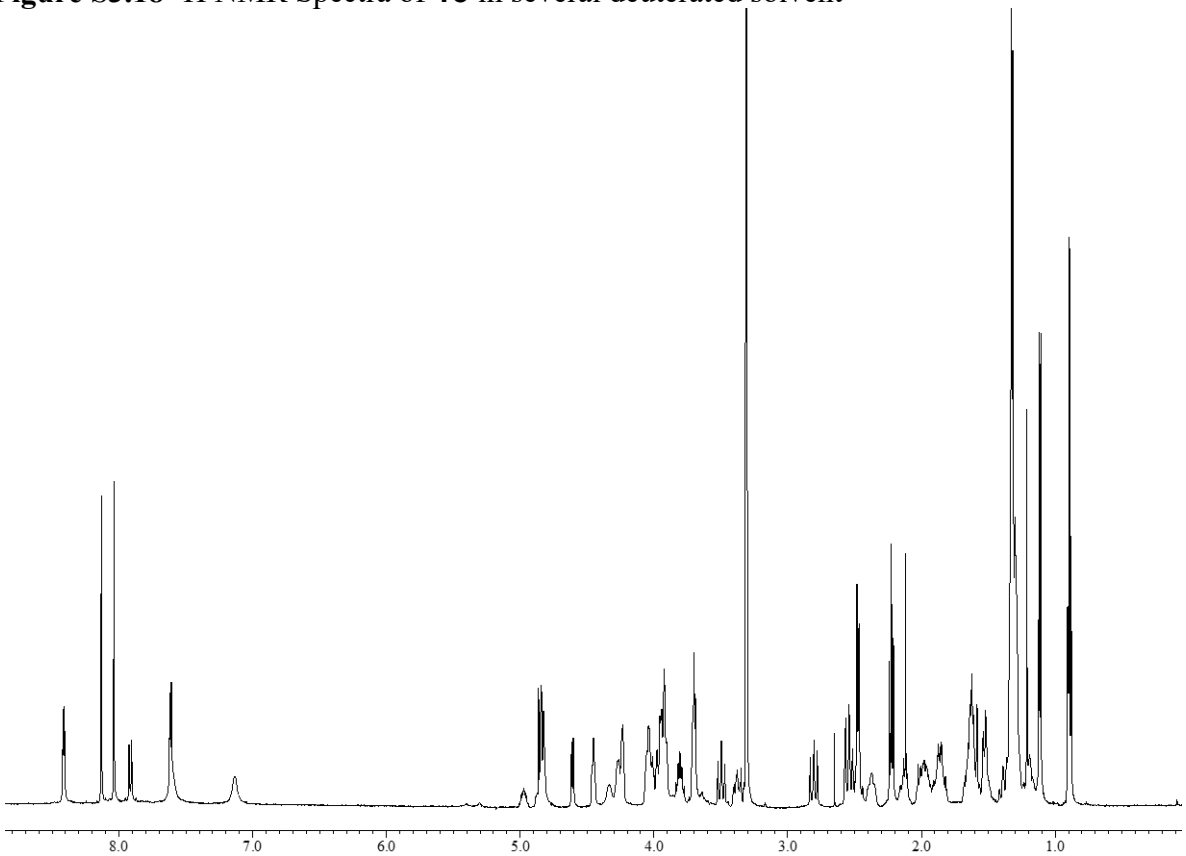


Figure S3.19 ^{1}H NMR Spectrum of **13**-Ga(III) (methanol-d_3, 500 MHz)

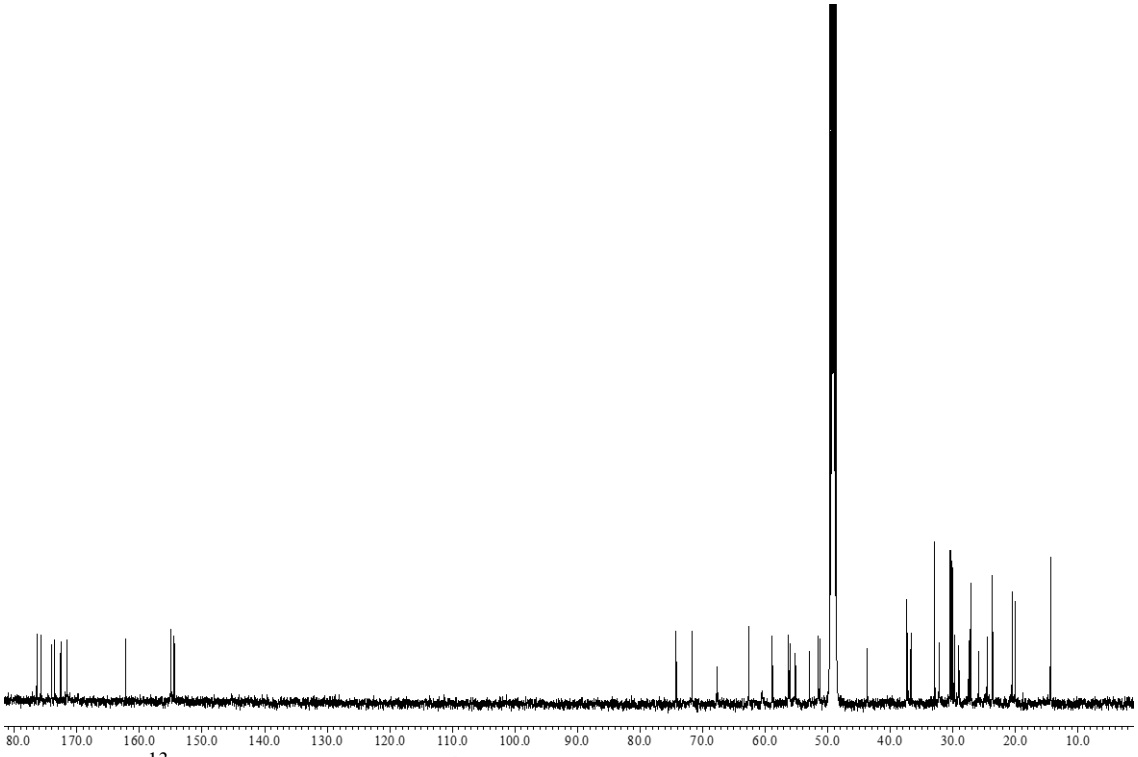


Figure S3.20 ^{13}C NMR Spectrum of **13**-Ga(III) (methanol-d_3, 500 MHz)

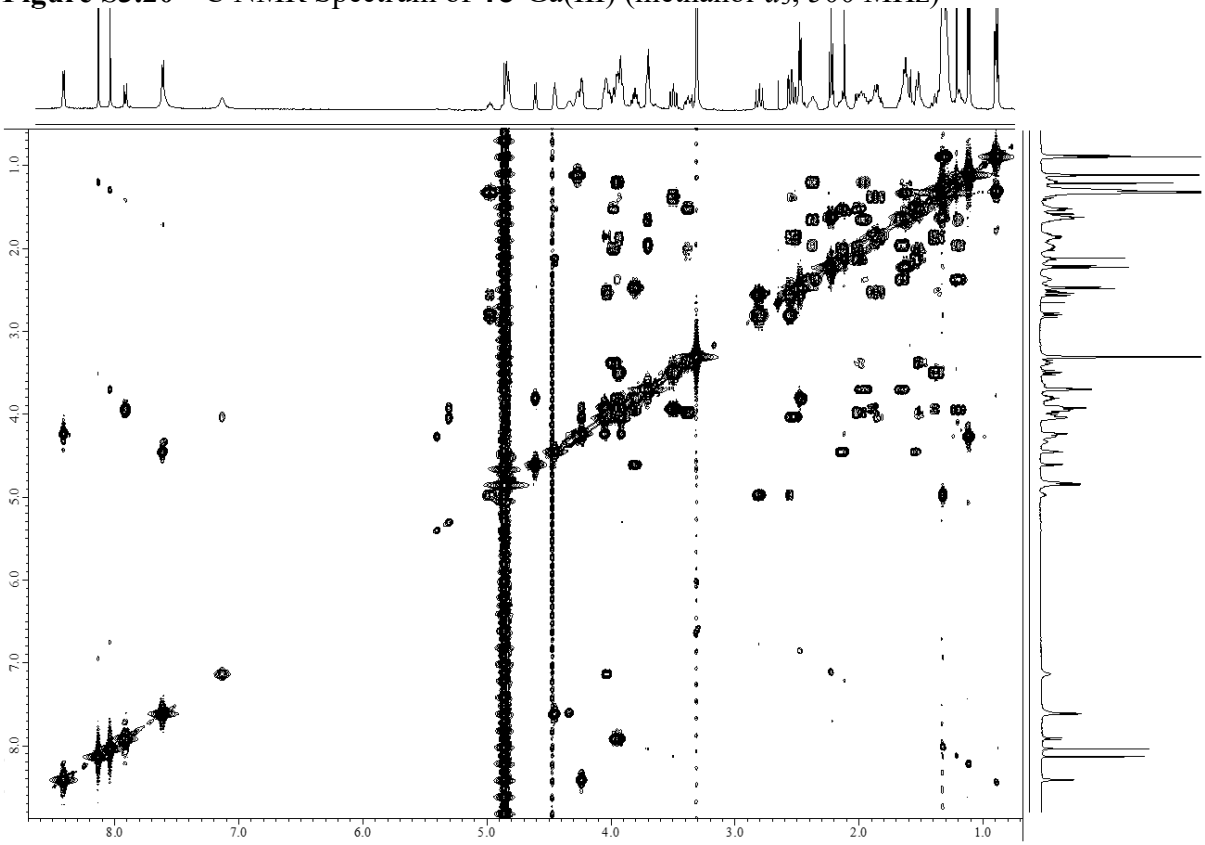


Figure S3.21 COSY NMR Spectrum of **13**-Ga(III) (methanol-d_3, 500 MHz)

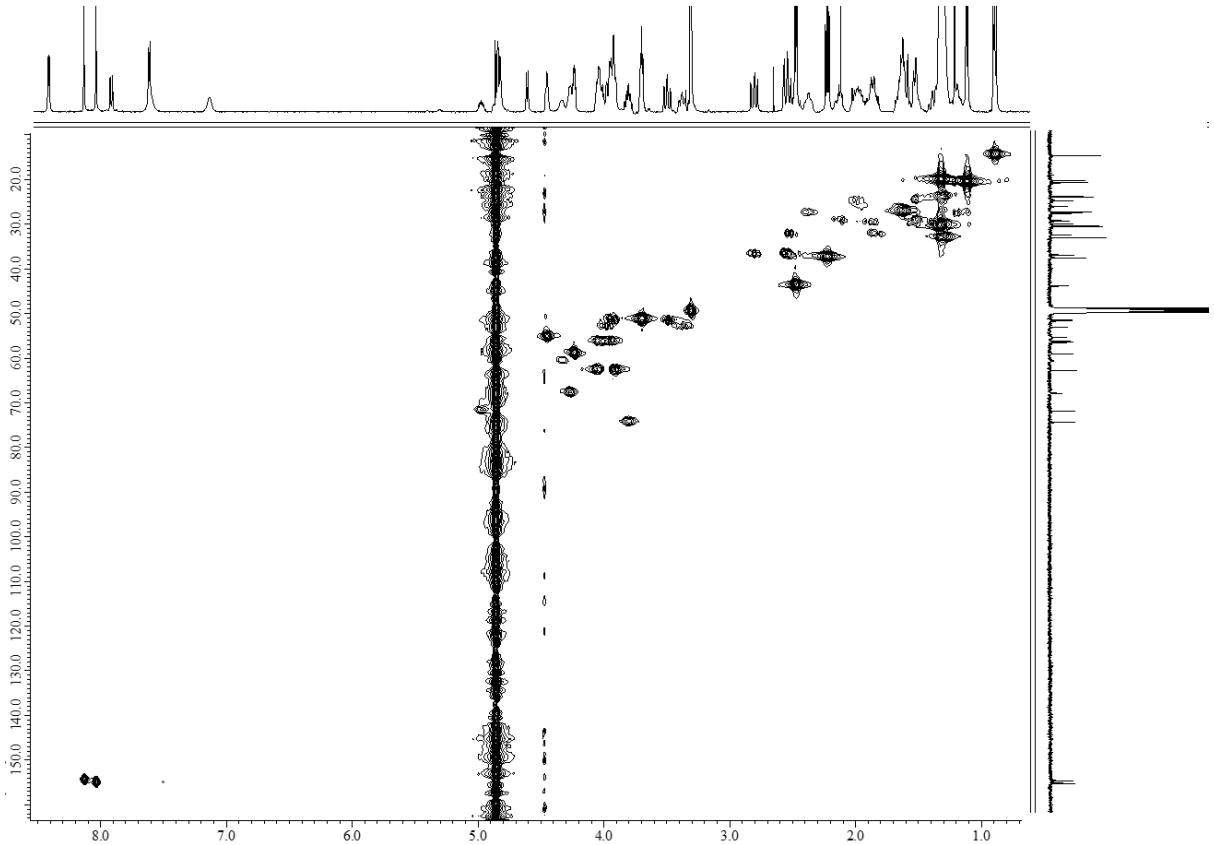


Figure S3.22 HSQC NMR Spectrum of **13**-Ga(III) (methanol-d_3, 500 MHz)

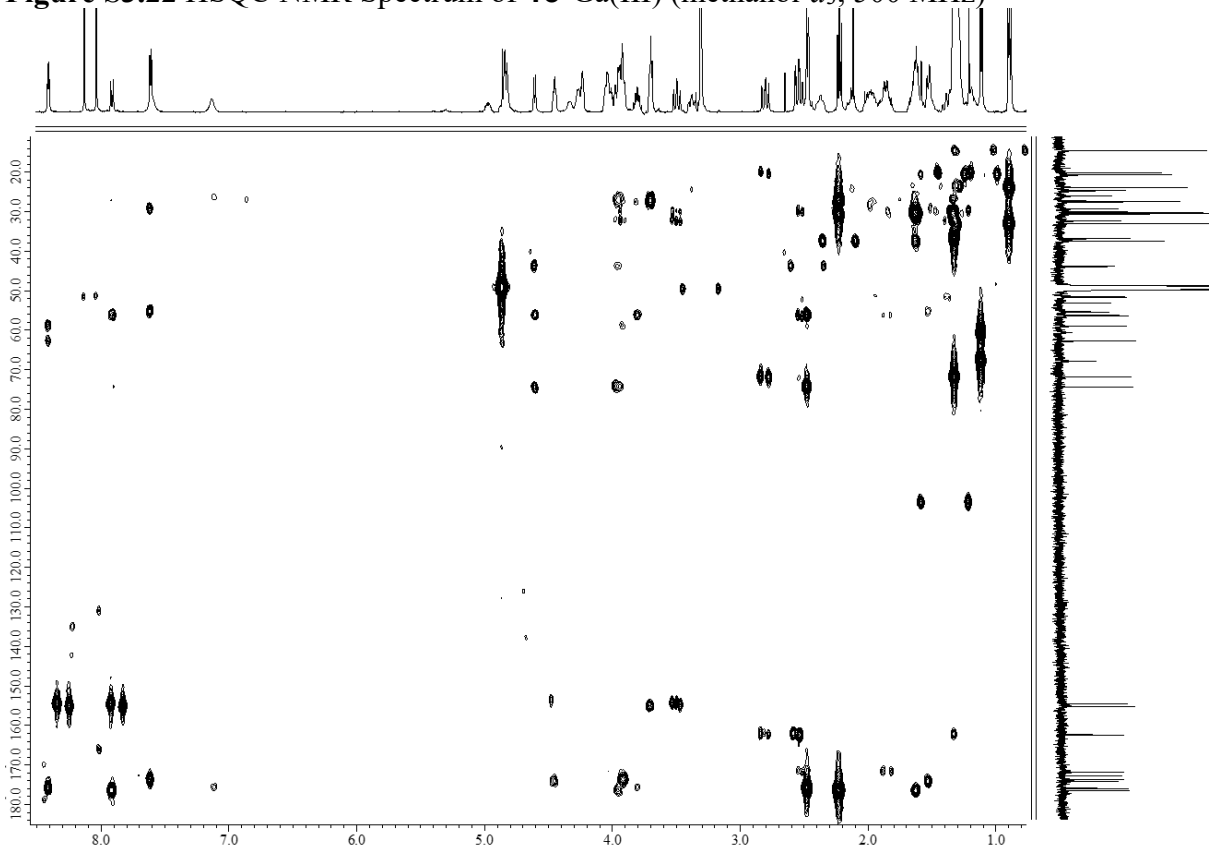


Figure S3.23 HMBC NMR Spectrum of **13**-Ga(III) (methanol-d_3, 500 MHz)

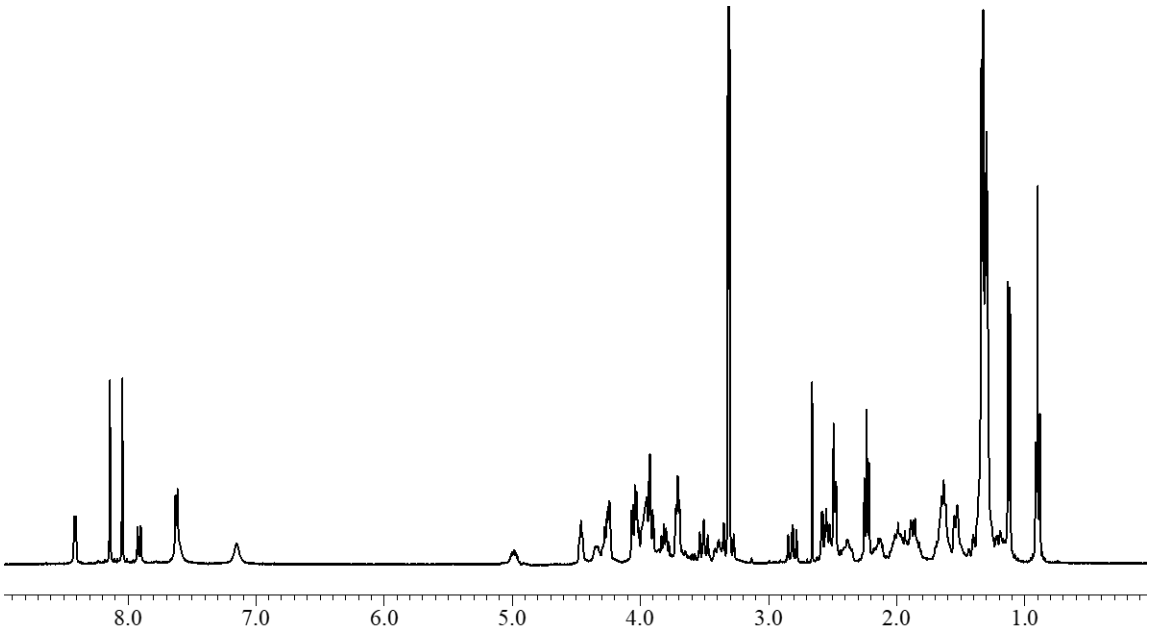


Figure S3.24 ^{1}H NMR Spectrum of **12**-Ga(III) (methanol-d_3, 400 MHz)

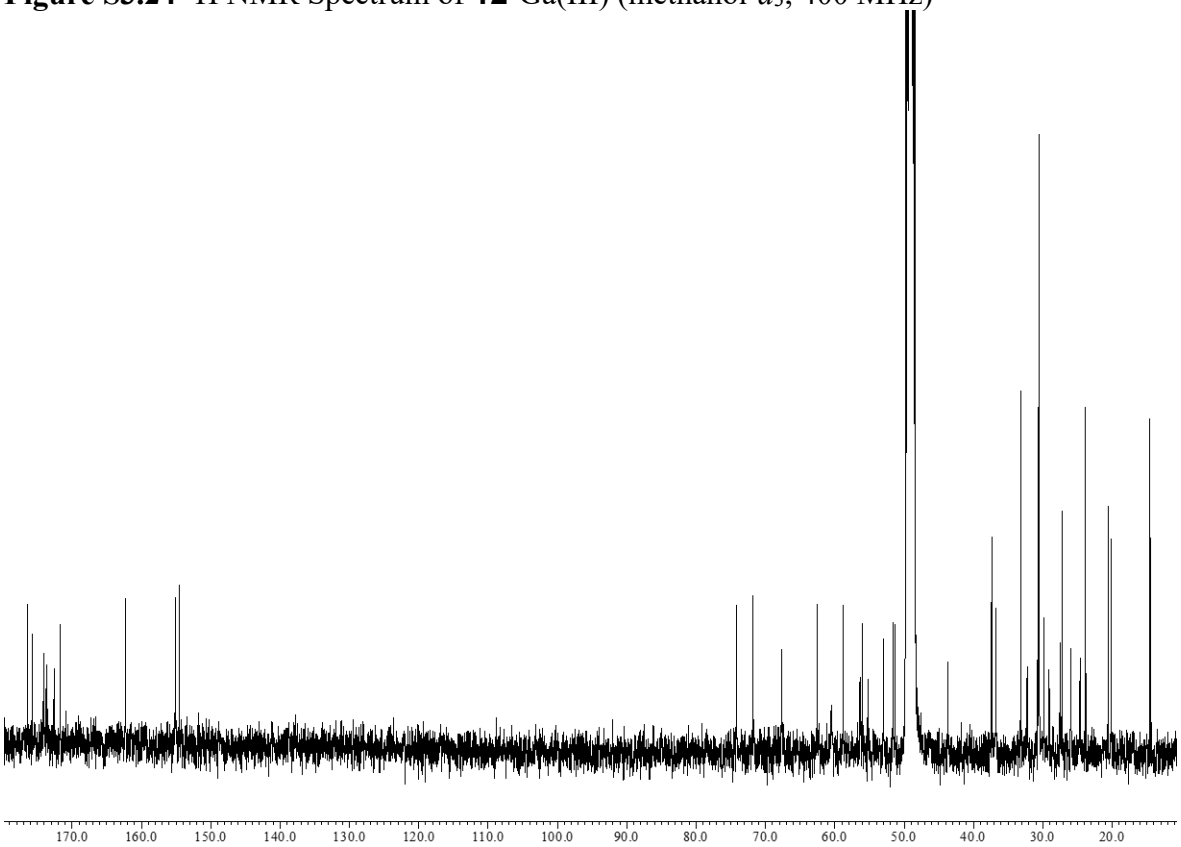


Figure S3.25 ^{13}C NMR Spectrum of **12**-Ga(III) (methanol-d_3, 400 MHz)

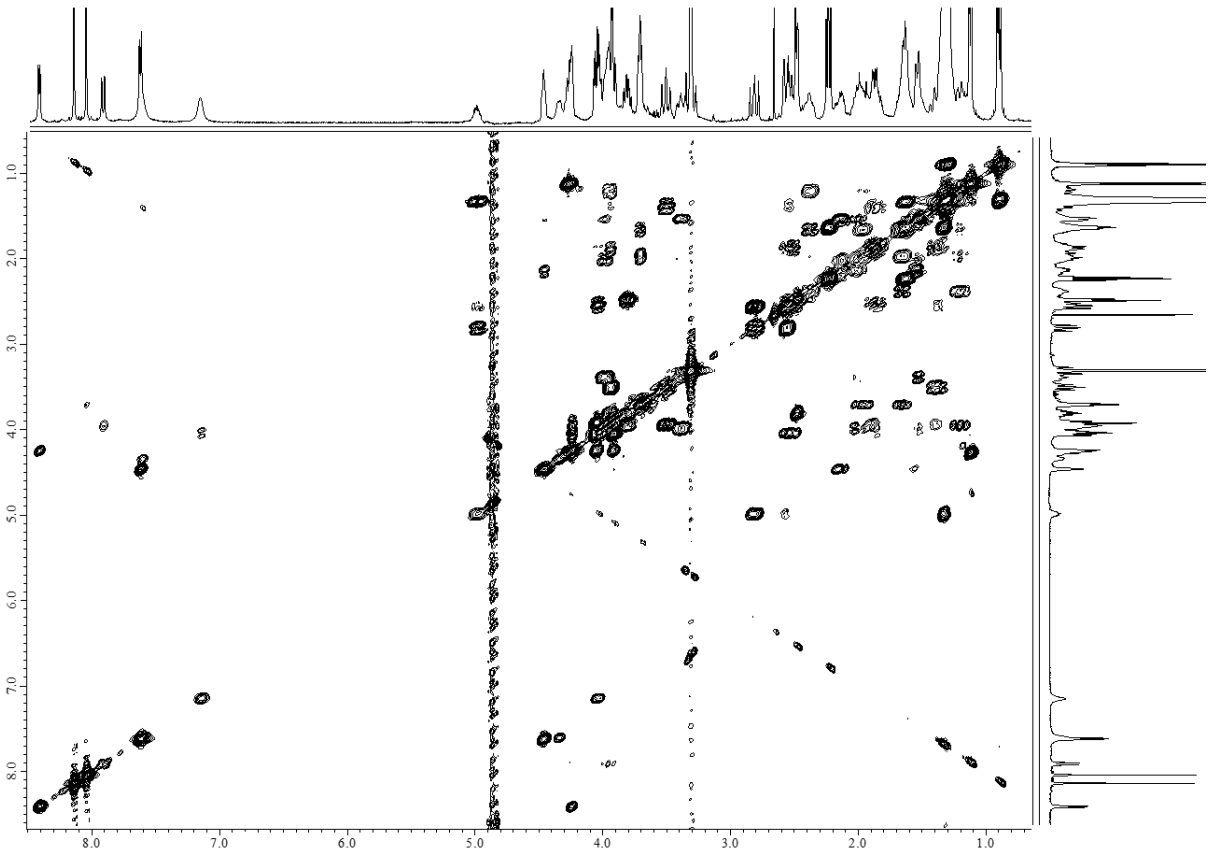


Figure S3.26 COSY NMR Spectrum of **12**-Ga(III) (methanol-d_3, 400 MHz)

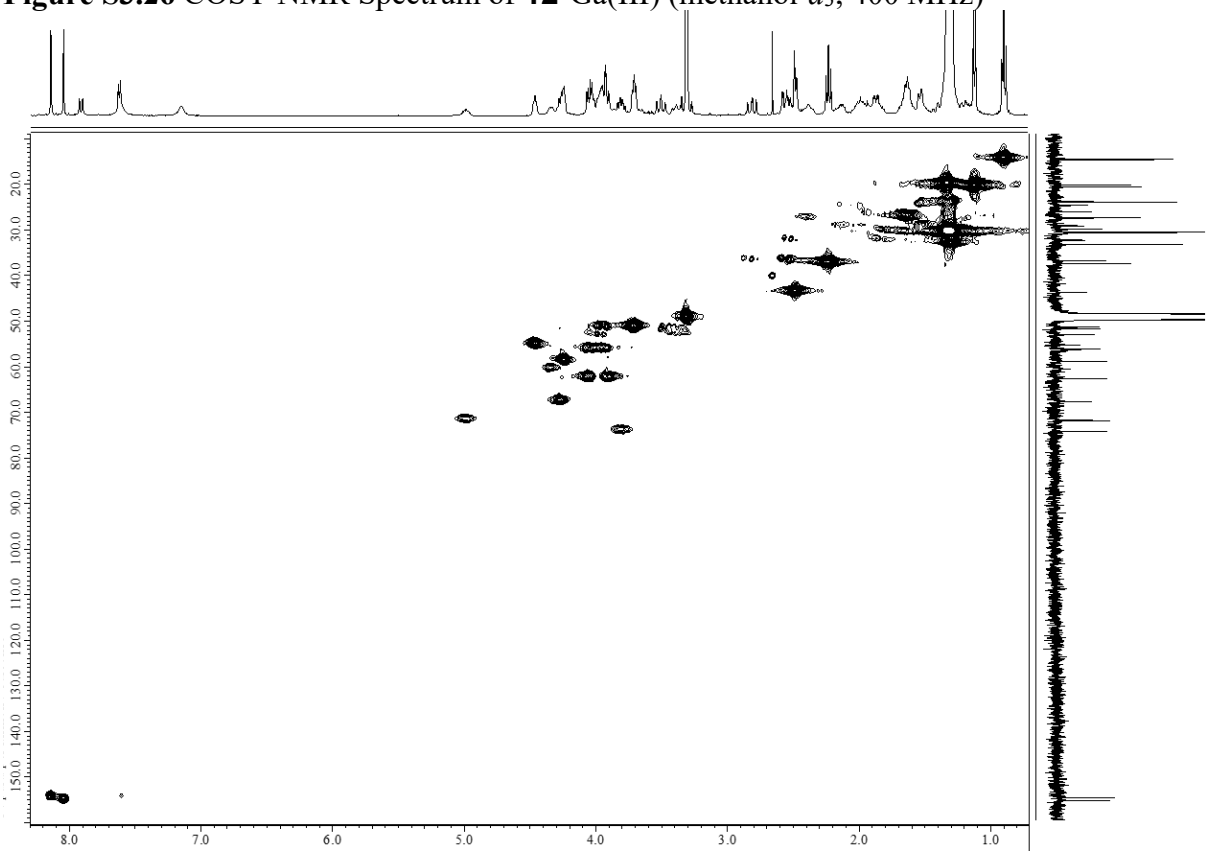


Figure S3.27 HSQC NMR Spectrum of **12**-Ga(III) (methanol-d_3, 400 MHz)

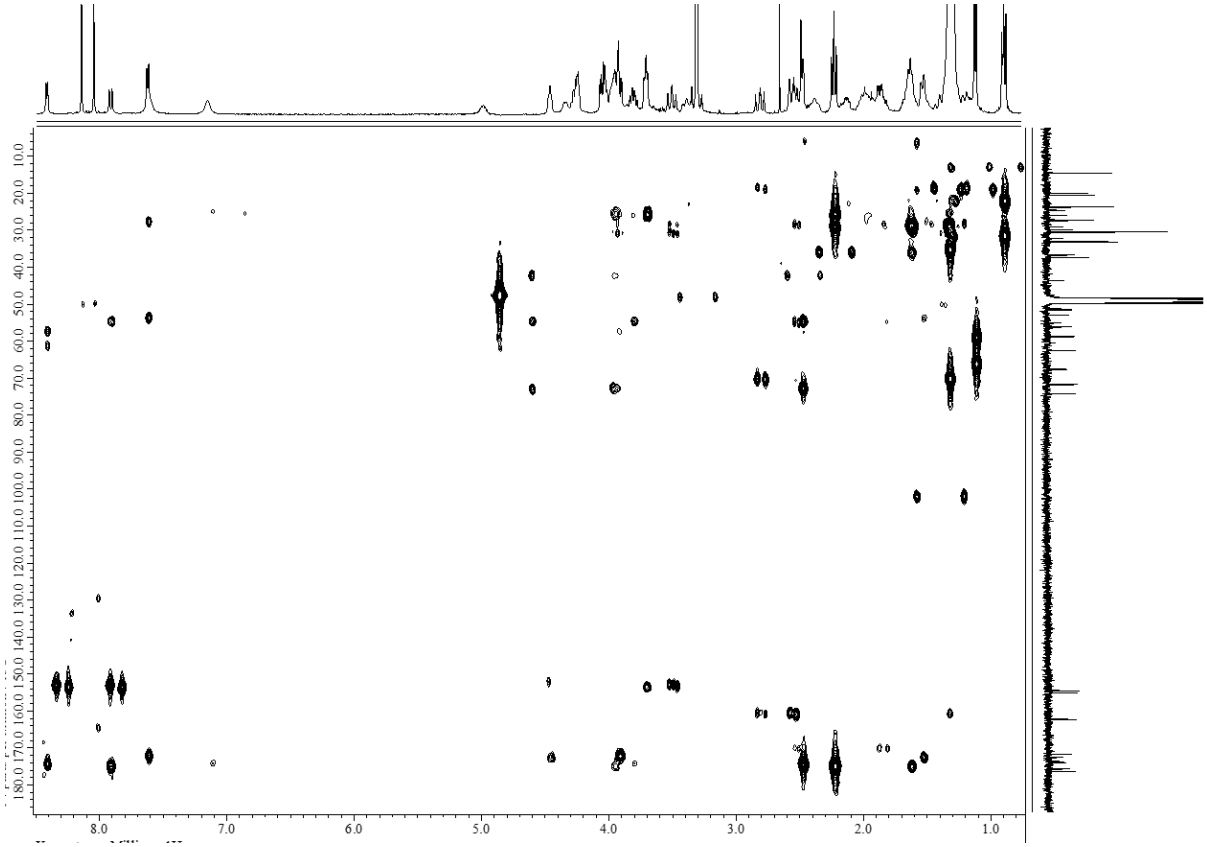


Figure S3.28 HMBC NMR Spectrum of **12**-Ga(III) (methanol-d_4, 400 MHz)

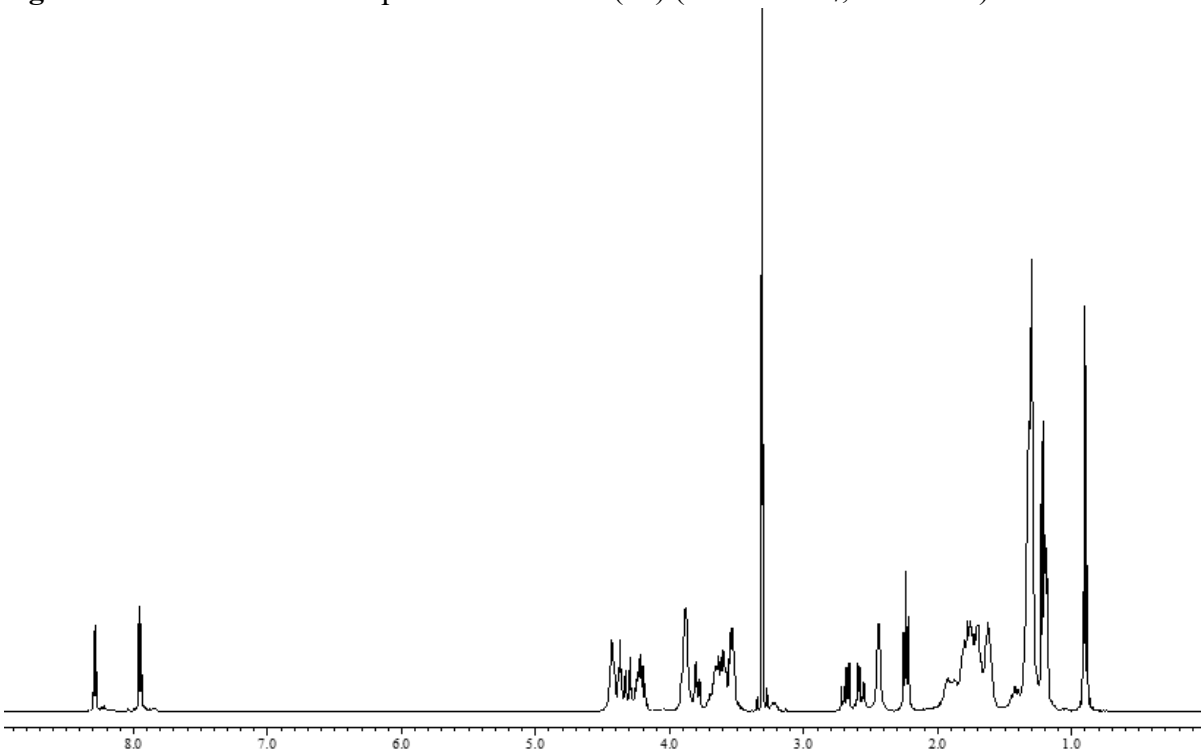


Figure S3.29 ^{1}H NMR Spectrum of **14** (methanol-d_4, 400 MHz)

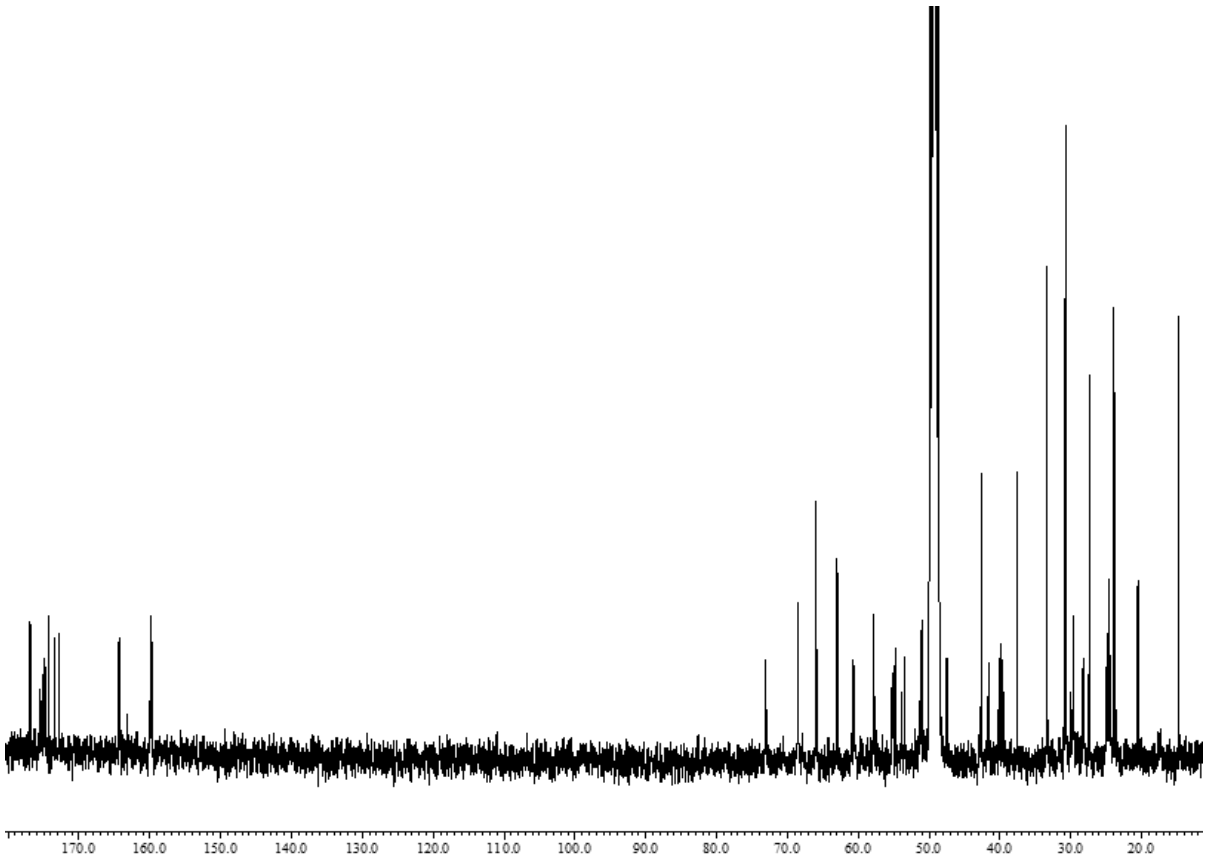


Figure S3.30 ^{13}C NMR Spectrum of **14** (methanol-d_4, 400 MHz)

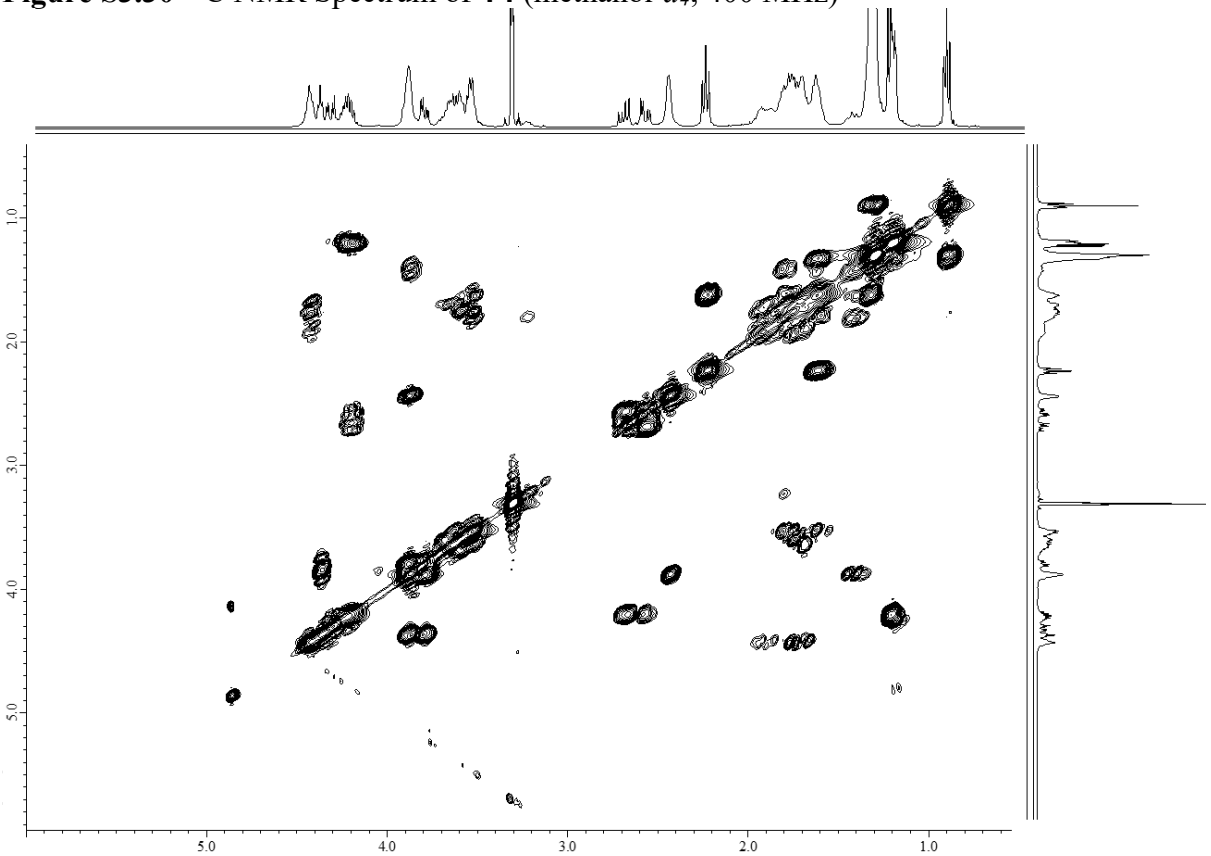


Figure S3.31 COSY NMR Spectrum of **14** (methanol-d_4, 400 MHz)

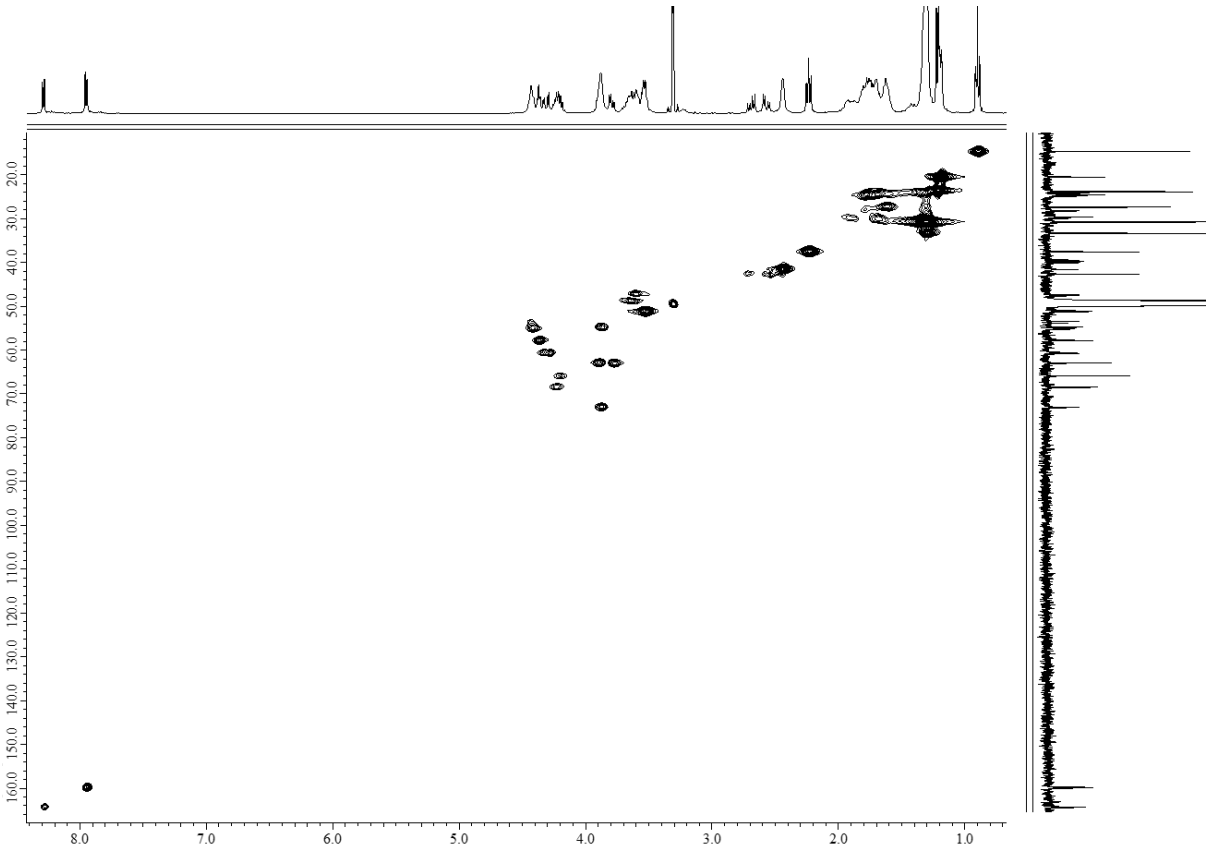


Figure S3.32 HSQC NMR Spectrum of **14** (methanol-d_4, 400 MHz)

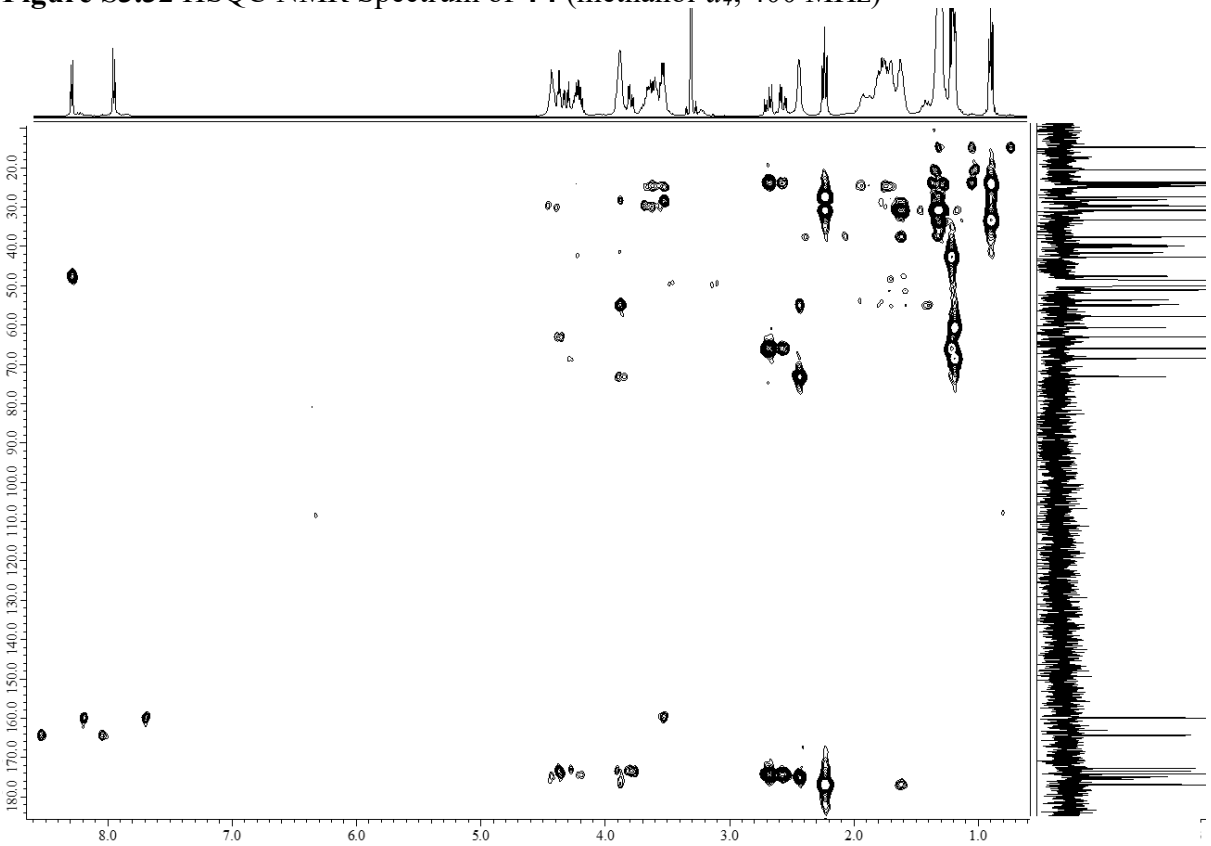


Figure S3.33 HMBC NMR Spectrum of **14** (methanol-d_4, 400 MHz)

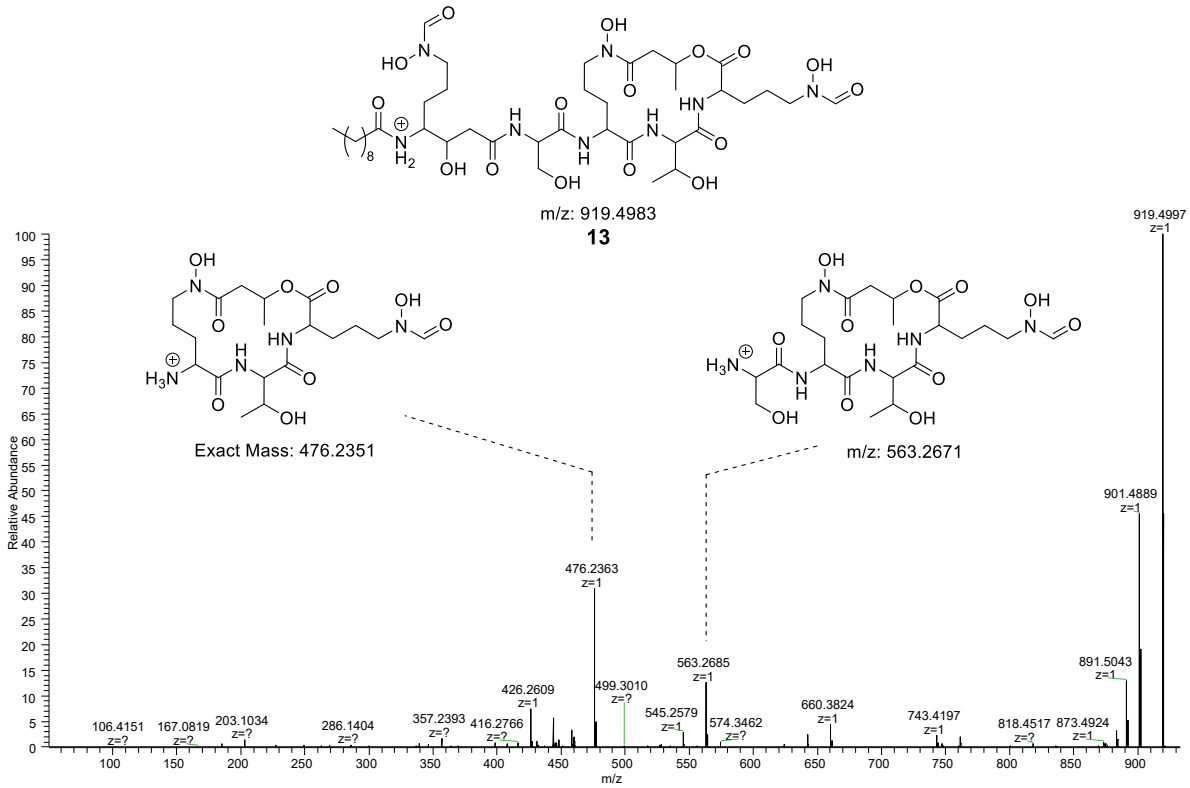


Figure S3.34 MS/MS Analysis of Compound **13**

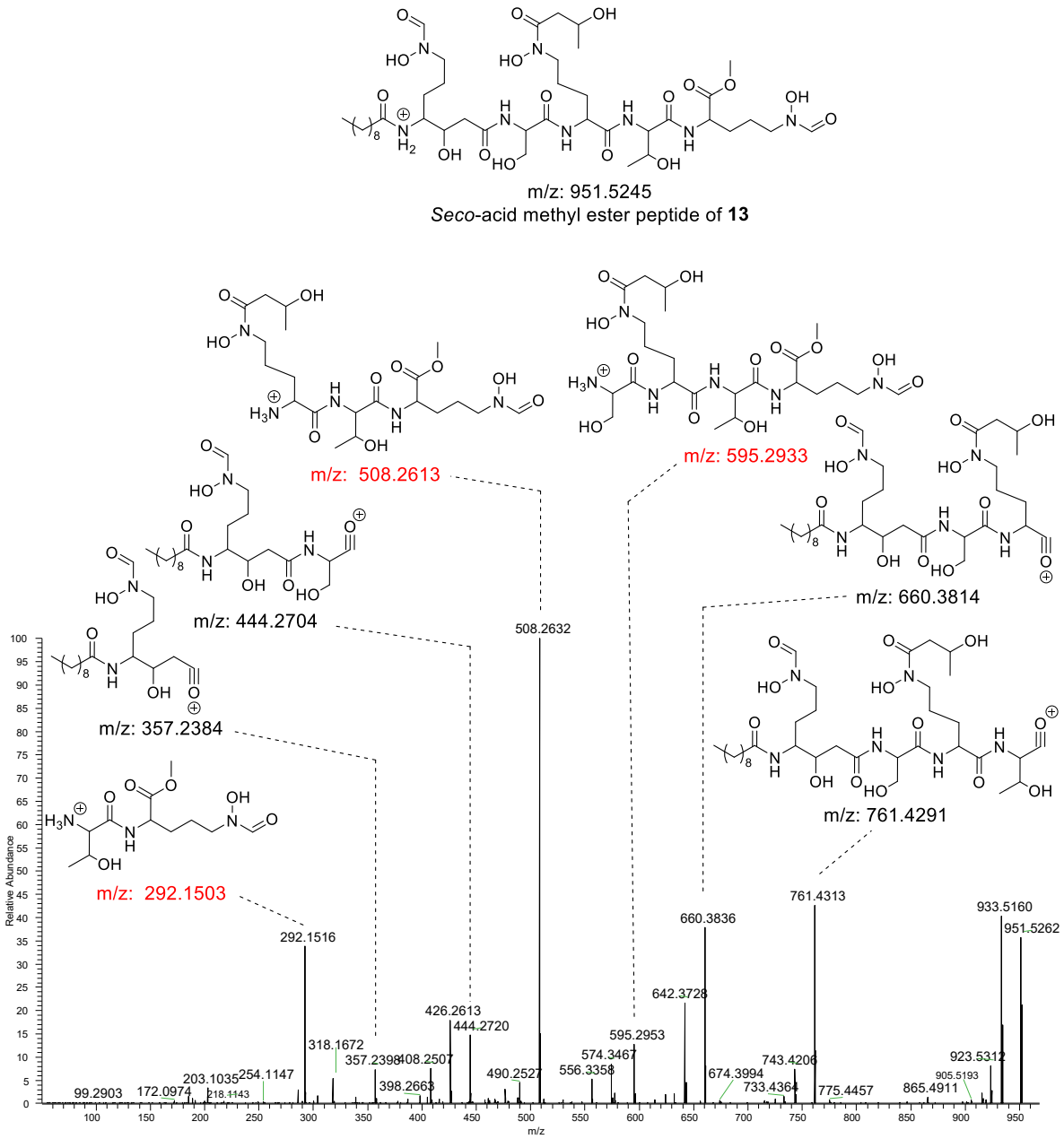


Figure S3.35 MS/MS Analysis of Compound MS/MS Analysis of Compound *seco*-acid methyl ester of **13**, red letter representing fragment ions generated toward C-terminus.

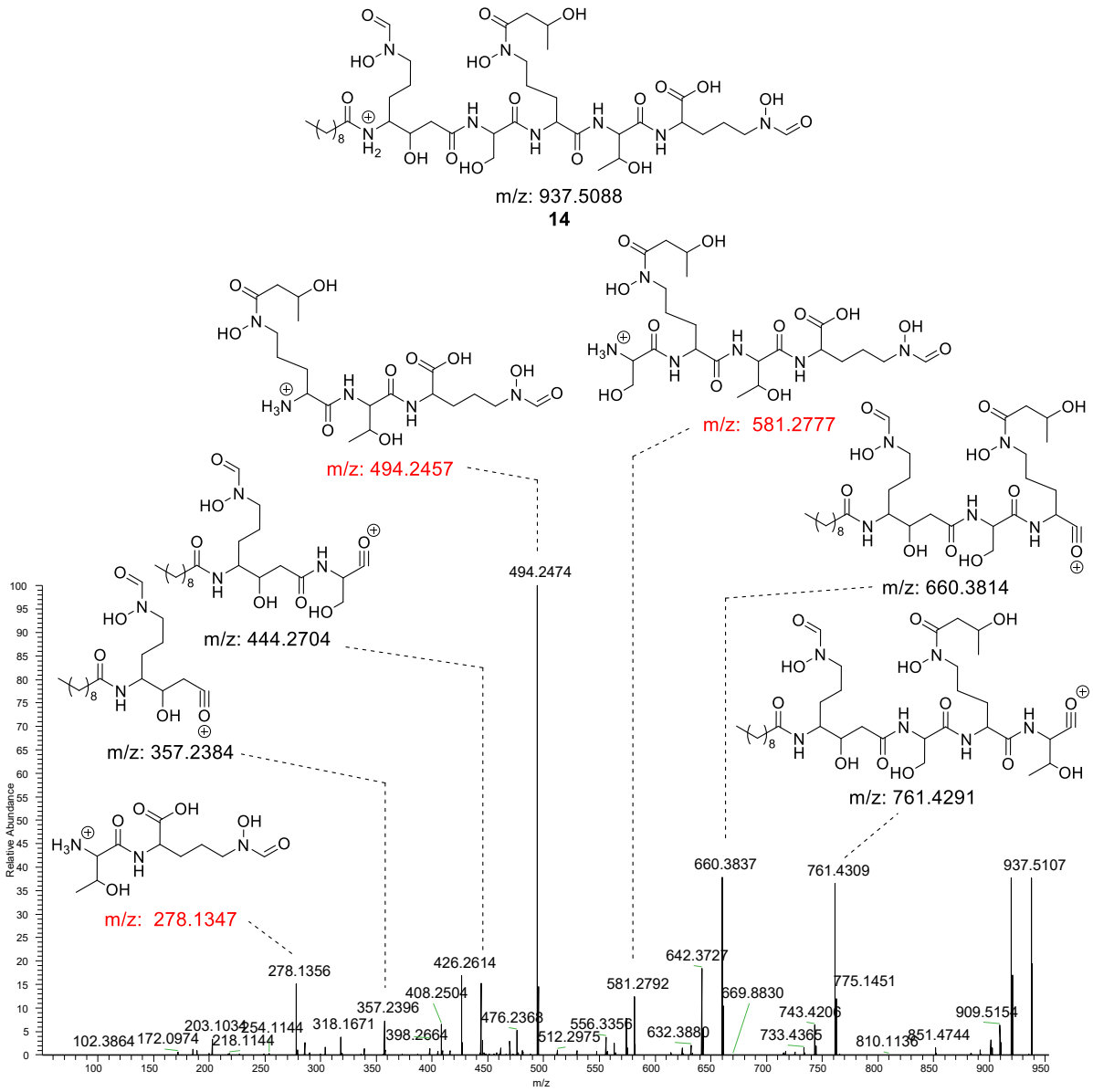


Figure S3.36 MS/MS Analysis of Compound **14**, red letter representing fragment ions generated toward C-terminus.

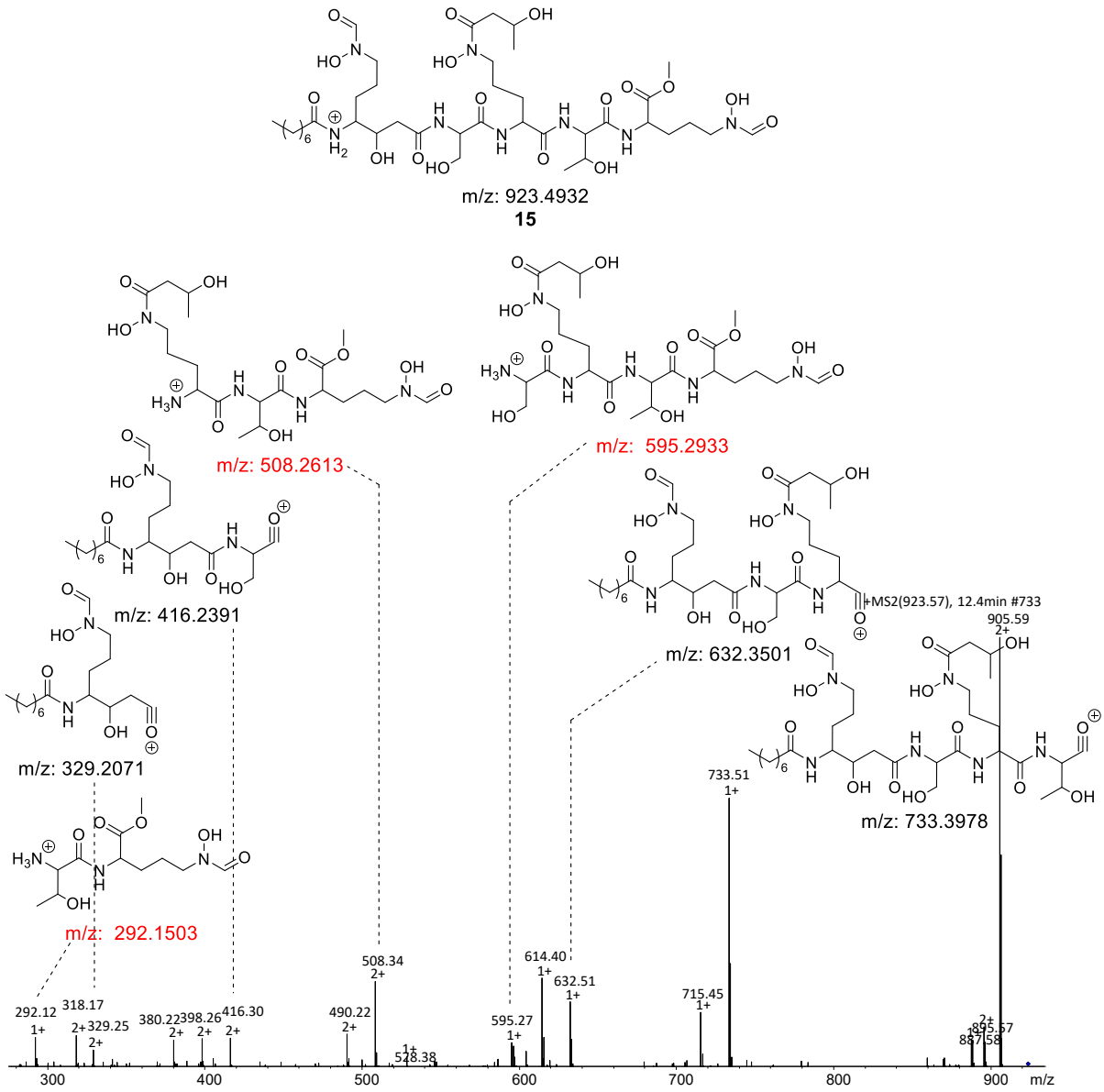


Figure S3.37 MS/MS Analysis of Compound **15**, red letter representing fragment ions generated toward C-terminus.

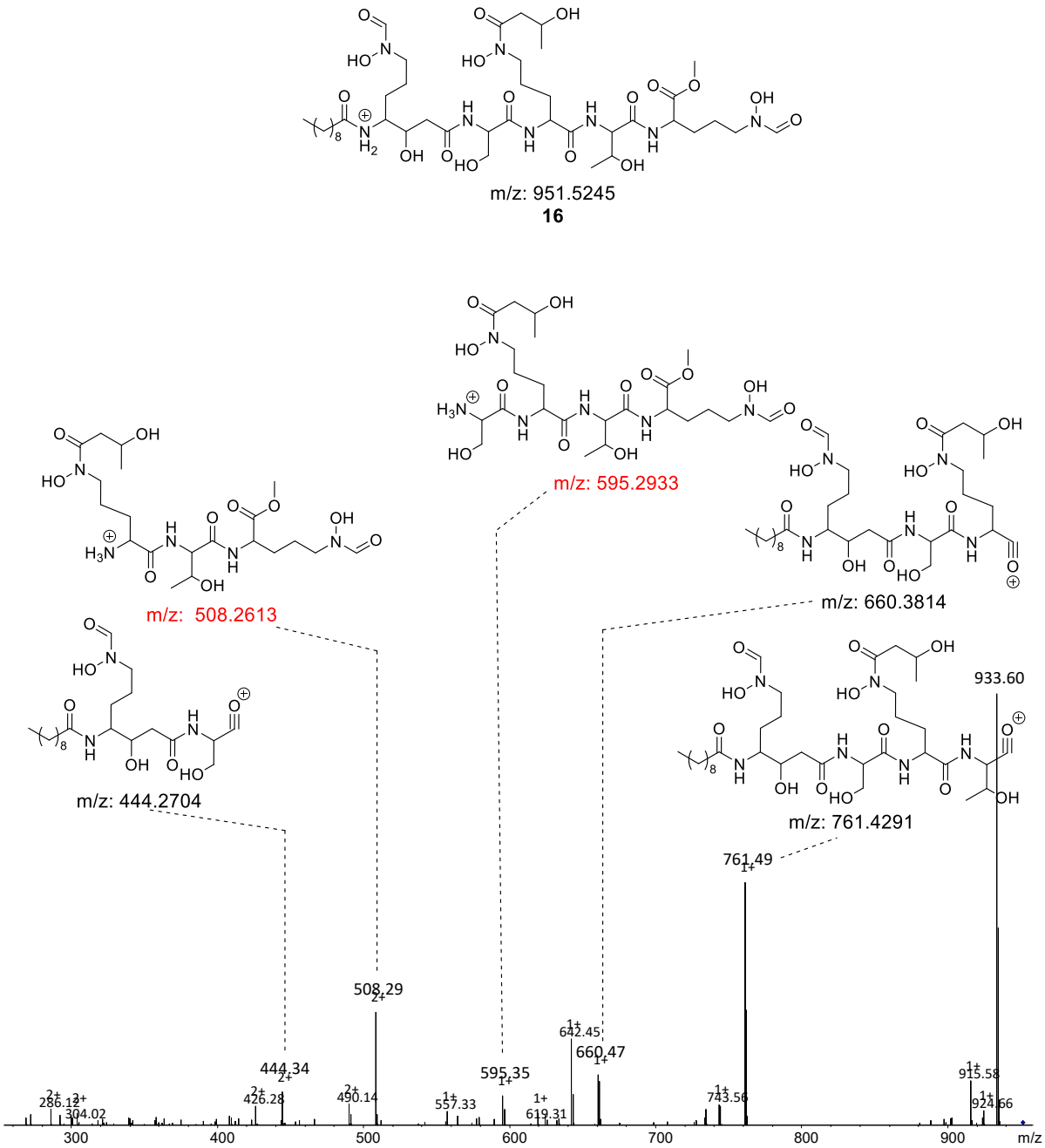


Figure S3.38 MS/MS Analysis of Compound **16**, red letter representing fragment ions generated toward C-terminus.

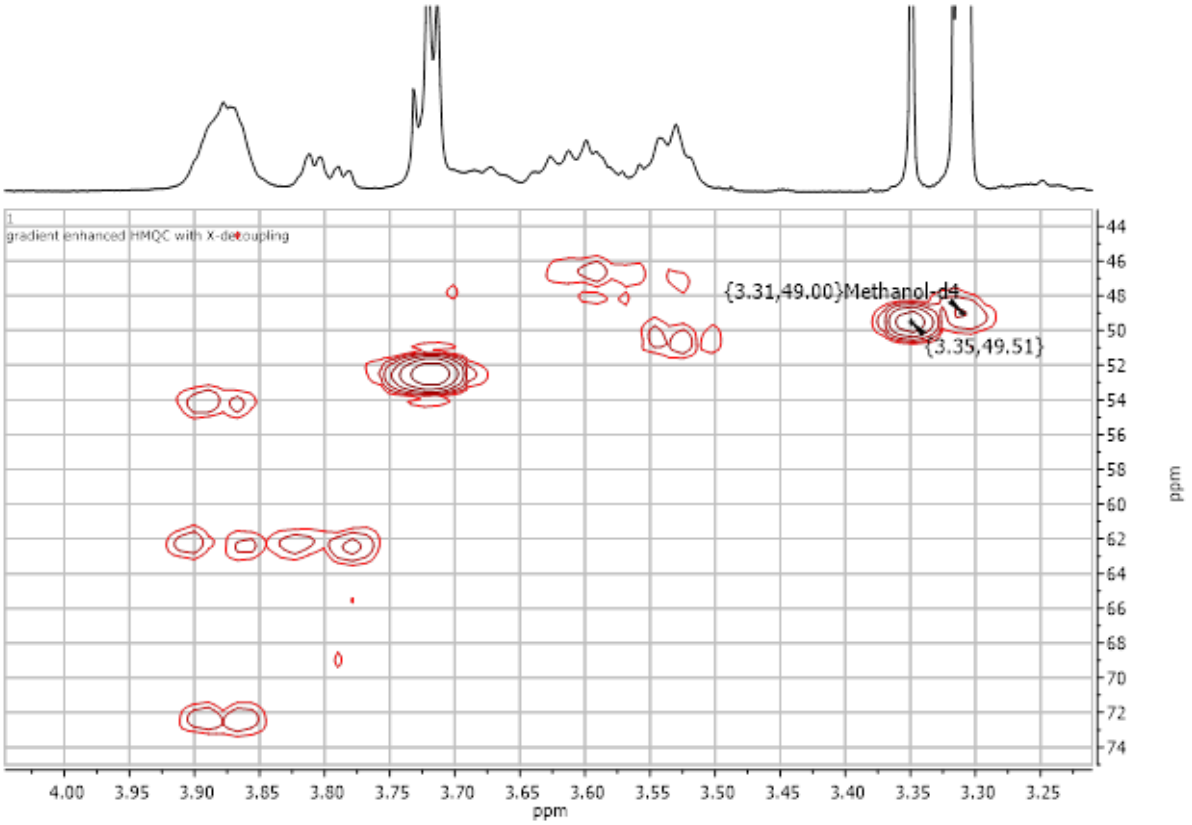


Figure S4.39 HSQC NMR spectrum of **16**.

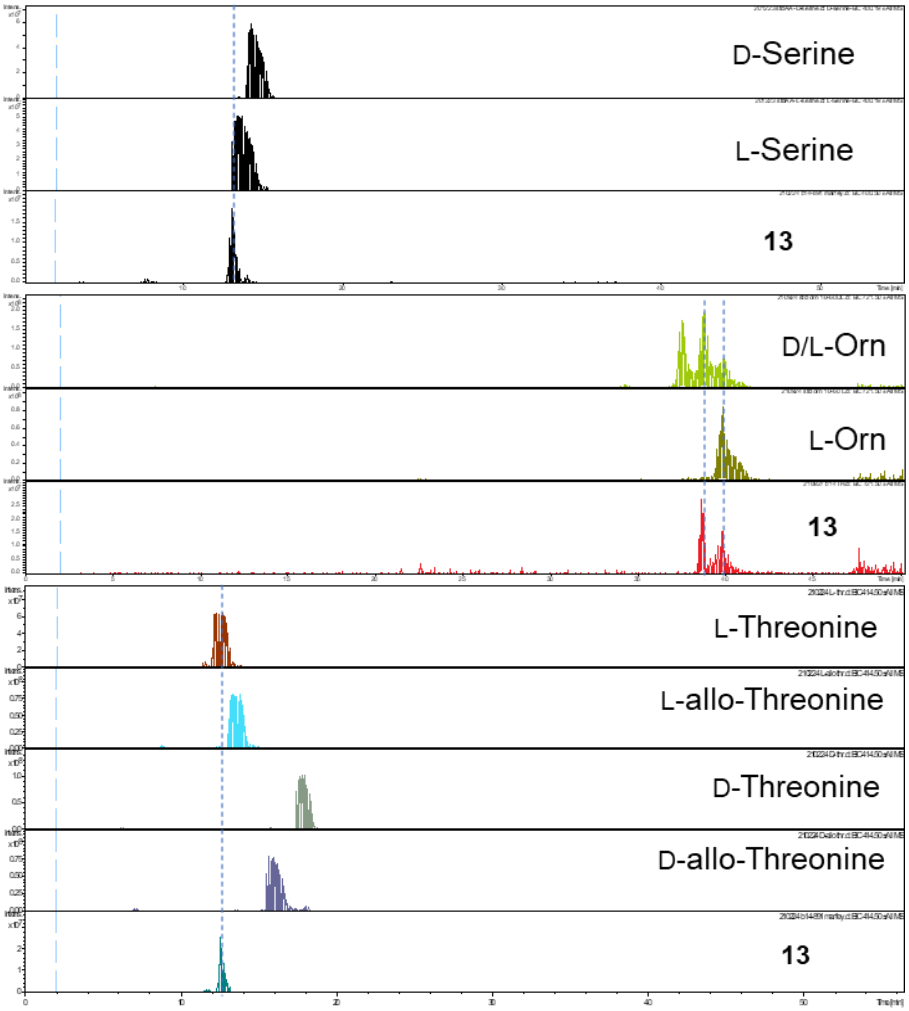


Figure S4.40 Marfey's analysis of **13**.

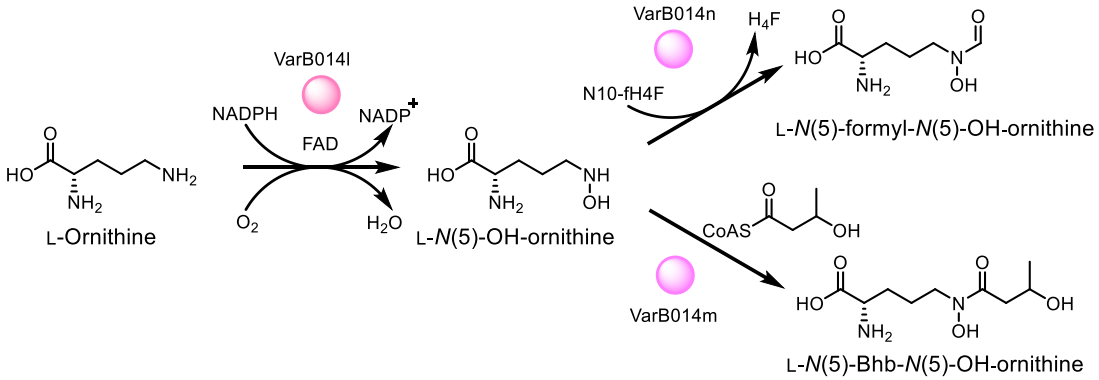


Figure S4.41 Biogenesis of *N*(5)-Fh-Orn and *N*(5)-Bhb-Orn

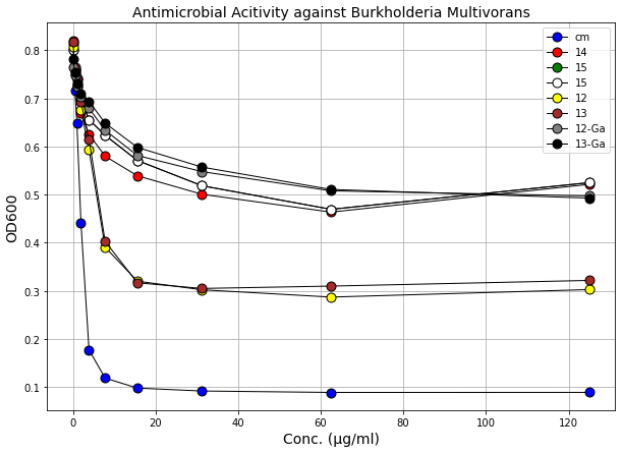


Figure S4.42 Antibacterial activity of **13**-Ga(III) against *B. multivorans*

Table S3.1 Summary of ichip-domesticated bacterial strains

Ichip	Habitats	Strains	Ichip	Habitats	Strains
B	Soil	86	J	Soil	24
C	Soil	39	K	Soil	42
D	River stream	92	L	Soil	87
E	River stream	40	M	Soil	83
F	Soil	120	N	Soil	72
G	River stream	75	O	Soil	20
H	Soil	11	P	Sponge	177
I	Soil	22	Q	Mangrove mud	56
Total					1046

Table S3.2 NMR data of **10** (methanol-*d*4, 400 MHz)

Pos.	δ_C, type		δ_H (mult.)	Pos.	δ_C, type		δ_H (mult.)
1. *N*(5)-Ac-*N*(5)-OH ornithine				5. β-hydroxyaspartic acid			
C1	173.4, C			C1	168.5, C		
C2	51.6, CH		4.17 (m)	C2	55.0, CH		4.61 (m)
C3	29.0, CH_2	α	1.70 (m)	C3	70.1, CH		3.79 (m)
		β	1.56 (m)	C4	173.0, C		
C4	22.9, CH_2		1.53 (m)	6. γ-amino acid*			
C5	46.5, CH_2		3.48 (m)	C1	175.5, C		
C6	170.3, C			C2	41.8, CH		2.70 (m)
C7	20.3, CH_3		2.09 (s)	C3	73.3, CH		3.74 (m)
2. *N*(5)-Ac-*N*(5)-OH ornithine				C4	49.5, CH		3.92 (m)
C1	171.4, C			C5	27.5, CH_2		1.22 (m)
C2	51.6, CH		4.35 (m)	C6	25.1, CH_2	α	1.44 (m)
C3	29.8, CH_2	α	1.61 (m)			β	1.36 (m)
		β	1.50 (m)	C7	40.8, CH_2	α	3.17 (m)
C4	22.8, CH_2		1.53 (m)			β	3.02 (m)
C5	46.8, CH_2		3.48 (m)	C8	156.5, C=N		
C6	170.3 C			C9	11.3, CH_3		1.19 (d)
C7	20.3, CH_3		2.10 (s)	7. Dodecanoic Acid			
3. Proline				C1	172.1, C		
C1	171.3, C			C2	35.5, CH_2		2.22 (m)
C2	59.7, CH		4.46 (m)	C3	25.4, CH_2		1.31 (m)
C3	29.8, CH_2	α	2.23 (m)	C4	28.7, CH_2		1.31 (m)
		β	2.04 (m)	C5	28.8, CH_2		1.31 (m)
C4	24.1, CH_2		1.48 (m)	C6	29.0, CH_2		1.31 (m)
C5	47.0, CH_2		3.48 (m)	C7	29.1, CH_2		1.31 (m)
4. Serine				C8	29.0, CH_2		1.31 (m)
C1	168.5, C			C9	29.0, CH_2		1.31 (m)
C2	52.8, CH		4.48 (m)	C10	31.3, CH_2		1.31 (m)
C3	61.8, CH_2	α	3.61 (m)	C11	22.1, CH_2		1.31 (m)
		β	3.72 (m)	C12	14.0, CH_3		0.90 (t)
*4-amino-7-guanidino-3-hydroxy-2-methylheptanoic acid							

Table S3.3 List of primers used for preparation of *varH002g* null mutant.

Primers id	Sequences 5'→3'	Size (kb)
Upstream-fw	gccaaaagttggcccagggcTCAGAGGGTTCTTTCCGTGGG	1
Upstream-rv	tcttcctggcagataagcttCGTGAACGCCTACGGCCC	
Downstream-fw	cggatgaatgggataagcttTGGGCGGCGCCTCGACGA	1
Downstream-rv	ttttaaggcagttattggtgATGACGACCATGACCGACCACTTCGAC	
CmR-fw	CACCAATAACTGCCTTAAAAAAATTACG	0.819
CmR-rv	GCCCTGGGCCAACTTTTG	
pRED-fw	AAGCTTATCTGCCAGGAAG	1.625
pRED-rv	AAGCTTATCCCATTCATCC	

Table S3.4 NMR data **13**-Ga(III) (methanol-*d*3, 500 MHz)

Pos.	**δ_C, type**		**δ_H (J in Hz)**	**Pos.**	**δ_C, type**		**δ_H (mult.)**
1. Decanoic Acid				4. *N*(5)-hydroxy ornithine			
C1	176.4, C			C1	174.0, C		
C2	37.3, CH		2.20, t (7.5)	C2	55.1, CH		4.45, (m)
C3	26.7, CH_2		1.62, m	C3	29.0, CH_2	α	2.12, (m)
C4	23.6, CH_2		1.34, m			β	1.52, (m)
C5	23.6, CH_2		1.32, m	C4	24.5, CH_2	α	1.97, (m)
C6	23.6, CH_2		1.32, m			β	1.53, (m)
C7	23.6, CH_2		1.32, m	C5	52.9, CH_2	α	3.38, (m)
C8	23.6, CH_2		1.32, m			β	3.96, (m)
C9	23.6, CH_2		1.32, m	C2-NH			7.62, d (6.2)
C10	14.4, CH_3		0.87, t (7.0)	5. Hydroxybutyrate			
2. γ-amino acid*				C1	162.2, C		
C1	175.7, C			C2	36.7, CH_2	α	2.79, dd
C2	43.7, CH_2		2.46, (m)			β	2.56, dd
C3	74.2, CH		3.81, (m)	C3	71.6, CH		4.98, (m)
C4	56.0, CH		3.96, (m)	C4	20.5, CH_3		1.32, d (6.1)
C5	27.4, CH_2	α	2.37, (m)	6. Threonine			
		β	1.19, (m)	C1	172.5, C		
C6	25.9, CH_2	α	1.95, (m)	C2	60.7, CH		4.34, (m)
		β	1.65, (m)	C3	67.6, CH		4.27, (m)
C7	51.2, CH_2		3.70, (m)	C4	20.0, CH_3		1.11, d (6.4)
C8	155.0, CH		8.04, (s)	C2-NH			7.54, bd
C4-NH			7.89, d (9.5)	7. *N*(5)-formyl-*N*(5)-hydroxy ornithine			
C3-OH			4.61, d (7.0)	C1	171.6, C		
3. Serine				C2	56.2, CH		4.04, (m)
C1	173.6, C			C3	32.2, CH_2	α	2.51, (m)
C2	58.8, CH		4.23, (m)			β	1.85, (m)
C3	62.6 CH_2		4.01, (m)	C4	29.7, CH_2	α	1.87, (m)
			3.92, (m)			β	1.38, (m)
C2-NH			8.39, d(5.4)	C5	51.5, CH_2	α	3.50, t (12.7)
						β	3.98, (m)
				C6	154.4, CH		8.13, s
				C2-NH			7.11, bd
*4-amino-7-N-hydroxyformyl-3-hydroxy-heptanoic acid							

Table S3.5 NMR data **12**-Ga(III) (methanol-*d*4, 400 MHz)

Pos.	δ_C, type		δ_H (J in Hz)	Pos.	δ_C, type		δ_H (mult.)
1. Decanoic Acid				4. *N*(5)-hydroxy ornithine			
C1	176.3, C			C1	174.0, C		
C2	37.3, CH		2.23, t (7.4)	C2	55.1, CH		4.46, (m)
C3	33.1, CH_2		1.33, (m)	C3	29.0, CH_2	α	2.13, (m)
C4	30.7, CH_2		1.33, (m)			β	1.52, (m)
C5	30.4, CH_2		1.33, (m)	C4	24.6, CH_2	α	1.97, (m)
C6	27.5, CH_2		1.33, (m)			β	1.53, (m)
C7	23.8, CH_2		1.33, (m)	C5	52.9, CH_2	α	3.39, (m)
C8	14.5, CH_3		0.90, (6.9)			β	3.96, (m)
				C2-NH			7.62, d (6.2)
				5. Hydroxybutyrate			
2. γ-amino acid*				C1	162.2, C		
C1	175.7, C			C2	36.7, CH_2	α	2.81, dd (15.0, 11.6)
C2	43.7, CH_2		2.48, (m)			β	2.56, dd (14.8, 3.4) d
C3	74.1, CH		3.80, (m)	C3	71.7, CH		4.99, (m)
C4	56.0, CH		3.96, (m)	C4	20.5, CH_3		1.30, (m)
C5	27.2, CH_2	α	2.38, (m)	6. Threonine			
		β	1.19, (m)	C1	172.6, C		
C6	25.9, CH_2	α	1.95, (m)	C2	60.4, CH		4.34, (m)
		β	1.65, (m)	C3	67.6, CH		4.28, (m)
C7	51.3, CH_2		3.71, (m)	C4	20.1, CH_3		1.12, d (6.4)
C8	155.1, CH		8.04, (s)	C2-NH			7.15, bd
C4-NH			7.89, d (9.4)	7. *N*(5)-formyl-*N*(5)-hydroxy ornithine			
C3-OH			n.r.	C1	171.6, C		
3. Serine				C2	56.2, CH		4.06, (m)
C1	173.6, C			C3	32.2, CH_2	α	2.52, (m)
C2	58.8, CH		4.25, (m)			β	1.85, (m)
C3	62.5 CH_2		4.04, (m)	C4	29.8, CH_2	α	1.87, (m)
			3.92, (m)			β	1.38, (m)
C2-NH			8.41, d (5.4)	C5	51.6, CH_2	α	3.51, (m)
						β	3.98, (m)
				C6	154.5, CH		8.14, s
				C2-NH			7.11 bd
*4-amino-7-N-hydroxyformyl-3-hydroxy-heptanoic acid							

Table S3.6 NMR data **14** (methanol-*d*4, 400 MHz)

Pos.	δ_C, type		δ_H (J in Hz)	Pos.	δ_C, type		δ_H (mult.)
1. Decanoic Acid				4. *N*(5)-hydroxy ornithine			
C1	176.9, C			C1	175.4, C		
C2	37.5, CH		2.23, t (7.4)	C2	54.9, CH		4.41, (m)
C3	26.7, CH_2		1.62, m	C3	29.0, CH_2	α	1.57-1.83, obs.
C4	23.6, CH_2		1.34, m			β	1.57-1.83, obs.
C5	23.6, CH_2		1.32, m	C4	24.6, CH_2	α	1.57-1.83, obs.
C6	23.6, CH_2		1.32, m			β	1.57-1.83, obs.
C7	23.6, CH_2		1.32, m	C5	48.7, CH_2	α	3.63, (m)
C8	14.5, CH_3		0.90, (6.9)			β	-
				C2-NH			-
				5. Hydroxybutyrate			
2. γ-amino acid*				C1	174.2, C		
C1	175.0, C			C2	42.6, CH_2	α	2.68, dd
C2	41.6, CH_2		2.48, (m)			β	2.56, dd
C3	73.0, CH		3.80, (m)	C3	65.9, CH		4.20, (m)
C4	54.7, CH		3.96, (m)	C4	20.5, CH_3		1.18, (m)
C5	27.2, CH_2	α	1.57-1.83, obs.	6. Threonine			
		β	1.57-1.83, obs.	C1	172.8, C		
C6	25.9, CH_2	α	1.57-1.83, obs.	C2	60.7, CH		4.33, (m)
		β	1.57-1.83, obs.	C3	68.4, CH		4.23, (m)
C7	47.6, CH_2		3.58, (m)	C4	23.7, CH_3		1.20, (m)
C8	164.3, CH		8.27, (s)	C2-NH			-
C4-NH			-	7. *N*(5)-formyl-*N*(5)-hydroxy ornithine			
C3-OH			-	C1	174.8, C		
3. Serine				C2	60.5, CH		4.29, (m)
C1	173.4, C			C3	28.3, CH_2	α	1.57-1.83, obs.
C2	57.8, CH		4.36, (m)			β	1.57-1.83, obs.
C3	62.9 CH_2		3.77, (m)	C4	24.8, CH_2	α	1.57-1.83, obs.
			3.88, (m)			β	1.57-1.83, obs.
C2-NH			-	C5	51.0, CH_2	α	3.52, (m)
				C6	159.8, CH		7.93, s
				C2-NH			-

*4-amino-7-N-hydroxyformyl-3-hydroxy-heptanoic acid; obs. (observed)

References

1. Bennett, R. N. & Wallsgrove, R. M. Secondary Metabolites in Plant Defense Mechanisms. *New Phytol.* **127**, 617–633 (1994).
2. Wink, M. Evolution of Secondary Metabolites from an Ecological and Molecular Phylogenetic Perspective. *Phytochemistry.* **64**, 3–19 (2003).
3. Nivina, A., Yuet, K. P., Hsu, J. & Khosla, C. Evolution and Diversity of Assembly-line Polyketide Synthases. *Chem. Rev.* **119**, 12524–12547 (2019).
4. Theobald, S., Vesth, T.C., Rendsvig, J.K., Nielsen, K.F., Riley, R., de Abreu, L.M., Salamov, A., Frisvad, J.C., Larsen, T.O., Andersen, M.R., & Hoof, J.B. Uncovering Secondary Metabolite Evolution and Biosynthesis Using Gene Cluster Networks and Genetic Dereplication. *Sci. Rep.* **8**, 17957 (2018).
5. Wink, M. Evolution of Secondary Metabolism in Plants. in *Ecological Biochemistry* 38–48 (John Wiley & Sons, Ltd, 2014).
6. Newman, D. J. & Cragg, G. M. Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. *J. Nat. Prod.* **83**, 770–803 (2020).
7. Miller, L. H. & Su, X. Artemisinin: Discovery from the Chinese Herbal Garden. *Cell.* **146**, 855–858 (2011).
8. Martino, E., Casamassima, G., Castiglione, S., Cellupica, E., Pantalone, S., Papagni, F., Rui, M., Siciliano, A.M., & Collina, S. Vinca Alkaloids and Analogues as Anti-cancer Agents: Looking Back, Peering Ahead. *Bioorganic Med. Chem. Lett.* **28**, 2816–2826 (2018).
9. Alam, M. M., Naeem, M., Khan, M. M. A. & Uddin, M. Vincristine and Vinblastine Anticancer Catharanthus Alkaloids: Pharmacological Applications and Strategies for Yield Improvement. in *Catharanthus roseus: Current Research and Future Prospects* (eds. Naeem, M., Aftab, T. & Khan, M. M. A.) 277–307 (Springer International Publishing, 2017).
10. Gallego-Jara, J., Lozano-Terol, G., Sola-Martínez, R. A., Cánovas-Díaz, M. & de Diego Puente, T. A Compressive Review about Taxol®: History and Future Challenges. *Molecules*. **25**, 5986 (2020).
11. Wani, M. C. & Horwitz, S. B. Nature as a Remarkable Chemist: A Personal Story of The Discovery and Development of Taxol. *Anti-Cancer Drugs.* **25**, 482–487 (2014).

12. Hosen, Md. D., Rabbi, Md. F., Raihan, Md. A. & Al Mamun, Md. A., Effect of turmeric dye and biomordants on knitted cotton fabric coloration: A promising alternative to metallic mordanting. *Cleaner Engineering and Technology* **3**, 100124 (2021).

13. Hutchings, M. I., Truman, A. W. & Wilkinson, B. Antibiotics: Past, Present and Future. *Curr. Opin. Microbiol.* **51**, 72–80 (2019).

14. Kuehl, F. A., Peck, R. L., Hoffhine, C. E. & Folkers, K. Streptomyces Antibiotics. XVIII. Structure of Streptomycin. *J. Am. Chem. Soc.* **70**, 2325–2330 (1948).

15. Faulkner, D. J. Marine Natural Products. *Nat. Prod. Rep.* **17**, 7–55 (2000).

16. Faulkner, D. J. Marine Natural Products. *Nat. Prod. Rep.* **18**, 1R-49R (2001).

17. Blunt, J. W., Copp, B. R., Keyzers, R. A., Munro, M. H. G. & Prinsep, M. R. Marine Natural Products. *Nat. Prod. Rep.* **33**, 382–431 (2016).

18. Blunt, J. W., Copp, B. R., Keyzers, R. A., Munro, M. H. G. & Prinsep, M. R. Marine Natural Products. *Nat. Prod. Rep.* **34**, 235–294 (2017).

19. Blunt, J.W., Carroll, A.R., Copp, B.R., Davis, R.A., Keyzers, R.A., & Prinsep, M.R. Marine Natural Products. *Nat. Prod. Rep.* **35**, 8–53 (2018).

20. Mehbub, M. F., Lei, J., Franco, C. & Zhang, W. Marine Sponge Derived Natural Products between 2001 and 2010: Trends and Opportunities for Discovery of Bioactives. *Mar. Drugs.* **12**, 4539–4577 (2014).

21. R. Carroll, A., R. Copp, B., A. Davis, R., A. Keyzers, R. & R. Prinsep, M. Marine natural products. *Nat. Prod. Rep.* **38**, 362–413 (2021).

22. Altmann, K.-H. Drugs from the Oceans: Marine Natural Products as Leads for Drug Discovery. *CHIMIA.* **71**, 646–652 (2017).

23. Hu, G.-P., Yuan, J., Sun, L., She, Z.-G., Wu, J.-H., Lan, X.-J., Zhu, X., Lin, Y.-C., & Chen, S.-P. Statistical Research on Marine Natural Products Based on Data Obtained between 1985 and 2008. *Mar. Drugs.* **9**, 514–525 (2011).

24. Hu, Y., Chen, J., Hu, G., Yu, J., Zhu, X., Lin, Y., Chen, S., Yuan, J., Hu, Y., Chen, J., Hu, G., Yu, J., Zhu, X., Lin, Y., Chen, S., & Yuan, J. Statistical Research on the Bioactivity of New Marine Natural Products Discovered during the 28 Years from 1985 to 2012. *Mar. Drugs.* **13**, 202–221 (2015).

25. Bergmann, W. & Feeney, R. J. The Isolation of a New Thymine Pentoside from Sponges. *J. Am. Chem. Soc.* **72**, 2809–2810 (1950).

26. Hirata, Y. & Uemura, D. Halichondrins - Antitumor Polyether Macrolides from a Marine Sponge. *Pure Appl. Chem.* **58**, 701–710 (1986).

27. Hickford, S. J. H., Blunt, J. W. & Munro, M. H. G. Antitumour polyether macrolides: Four new halichondrins from the New Zealand deep-water marine sponge *Lissodendoryx* sp. *Bio. Med. Chem.* **17**, 2199–2203 (2009).

28. Kaufman, P.A., Awada, A., Twelves, C., Yelle, L., Perez, E.A., Velikova, G., Olivo, M.S., He, Y., Dutcus, C.E., & Cortes, J. Phase III Open-Label Randomized Study of Eribulin Mesylate Versus Capecitabine in Patients with Locally Advanced or Metastatic Breast Cancer Previously Treated with An Anthracycline and A Taxane. *J.C.O.* **33**, 594–601 (2015).

29. Schöffski, P., Chawla, S., Maki, R.G., Italiano, A., Gelderblom, H., Choy, E., Grignani, G., Camargo, V., Bauer, S., Rha, S.Y., Blay, J.-Y., Hohenberger, P., D'Adamo, D., Guo, M., Chmielowski, B., Le Cesne, A., Demetri, G.D., & Patel, S.R. Eribulin Versus Dacarbazine in Previously Treated Patients with Advanced Liposarcoma or Leiomyosarcoma: A Randomised, Open-label, Multicentre, Phase 3 Trial. *The Lancet.* **387**, 1629–1637 (2016).

30. Rinehart, K.L., Holt, T.G., Fregeau, N.L., Stroh, J.G., Keifer, P.A., Sun, F., Li, L.H., & Martin, D.G. Ecteinascidins 729, 743, 745, 759A, 759B, and 770: Potent Antitumor Agents from the Caribbean Tunicate *Ecteinascidia turbinata*. *J. Org. Chem.* **55**, 4512–4515 (1990).

31. Brodowicz, T. Trabectedin in Soft Tissue Sarcomas. *Future Oncol.* **10**, 1–5 (2014).

32. Teplinsky, E. & Herzog, T. J. The Efficacy of Trabectedin in Treating Ovarian Cancer. *Expert Opin. Pharmacother.* **18**, 313–323 (2017).

33. Pettit, G.R., Herald, C.L., Doubek, D.L., Herald, D.L., Arnold, E., & Clardy, J. Isolation and Structure of Bryostatin 1. *J. Am. Chem. Soc.* **104**, 6846–6848 (1982).

34. Nelson, T.J., Sun, M.-K., Lim, C., Sen, A., Khan, T., Chirila, F.V., & Alkon, D.L. Bryostatin Effects on Cognitive Function and PKCε in Alzheimer's Disease Phase IIa and Expanded Access Trials. *J. Alzheimer's Dis.* **58**, 521–535 (2017).

35. Renis, H. E. Antiviral Activity of Cytarabine in Herpesvirus–Infected Rats. *Antimicrob. Agents and Chemother.* **4**, 439–444 (1973).

36. Lee, W. W., Benitez, A., Goodman, L. & Baker, B. R. Potential Anticancer Agents.1 Xl. Synthesis of the B-Anomer of 9-(D-Arabinofuranosyl)-Adenine. *J. Am. Chem. Soc.* **82**, 2648–2649 (1960).

37. McIntosh, M., Cruz, L. J., Hunkapiller, M. W., Gray, W. R. & Olivera, B. M. Isolation and Structure of a Peptide Toxin from the Marine Snail *Conus magus*. *Arch. Biochem. Biophys.* **218**, 329–334 (1982).

38. Olivera, B.M., Cruz, L.J., De Santos, V., LeCheminant, G., Griffin, D., Zeikus, R., McIntosh, J.M., Galyean, R., Varga, J. Neuronal Calcium Channel Antagonists. Discrimination between Calcium Channel Subtypes Using ω-Conotoxin from *Conus magus* Venom. *Biochemistry.* **26**, 2086–2090 (1987).

39. Blair, H. A. & Dhillon, S. Omega-3 Carboxylic Acids (Epanova®): A Review of Its Use in Patients with Severe Hypertriglyceridemia. *Am. J. Cardiovasc. Drugs.* **14**, 393–400 (2014).

40. Towle, M.J., Salvato, K.A., Budrow, J., Wels, B.F., Kuznetsov, G., Aalfs, K.K., Welsh, S., Zheng, W., Seletsky, B.M., Palme, M.H., Habgood, G.J., Singer, L.A., DiPietro, L.V., Wang, Y., Chen, J.J., Quincy, D.A., Davis, A., Yoshimatsu, K., Kishi, Y., Yu, M.J., & Littlefield, B.A. In Vitro and In Vivo Anticancer Activities of Synthetic Macrocyclic Ketone Analogues of Halichondrin B. *Cancer Res.* **61**, 1013–1021 (2001).

41. D'Incalci, M., Badri, N., Galmarini, C. M. & Allavena, P. Trabectedin, a Drug Acting on Both Cancer Cells and the Tumour Microenvironment. *Br. J. Cancer.* **111**, 646–650 (2014).

42. Gulder, T. A. M. & Moore, B. S. Salinosporamide Natural Products: Potent 20 S Proteasome Inhibitors as Promising Cancer Chemotherapeutics. *Angew. Chem. Int. Ed.* **49**, 9346–9367 (2010).

43. Nieto, F.R., Cobos, E.J., Tejada, M.Á., Sánchez-Fernández, C., González-Cano, R., & Cendán, C.M. Tetrodotoxin (TTX) as a Therapeutic Agent for Pain. *Mar. Drugs.* **10**, 281–305 (2012).

44. Natoli, M., Herzig, P., Pishali Bejestani, E., Buchi, M., Ritschard, R., Lloyd, G.K., Mohanlal, R., Tonra, J.R., Huang, L., Heinzelmann, V., Trüb, M., Zippelius, A., & Kashyap, A.S. Plinabulin, a Distinct Microtubule-Targeting Chemotherapy, Promotes M1-Like Macrophage Polarization and Anti-tumor Immunity. *Front. Oncol.* **11**, 644608 (2021).

45. Safavi-Hemami, H., Brogan, S. E. & Olivera, B. M. Pain Therapeutics from Cone Snail Venoms: From Ziconotide to Novel Non-opioid Pathways. *J. Proteomics.* **190**, 12–20 (2019).

46. Lago, J., Rodríguez, L. P., Blanco, L., Vieites, J. M. & Cabado, A. G. Tetrodotoxin, an Extremely Potent Marine Neurotoxin: Distribution, Toxicity, Origin and Therapeutical Uses. *Mar. Drugs.* **13**, 6384–6406 (2015).

47. Suehiro, M. Historical Review on Chemical and Medical Studies of Globefish Toxin before World War II. *Yakushigaku Zasshi.* **29**, 428–434 (1994).

48. Feling, R.H., Buchanan, G.O., Mincer, T.J., Kauffman, C.A., Jensen, P.R., & Fenical, W. Salinosporamide A: A Highly Cytotoxic Proteasome Inhibitor from a Novel Microbial Source, a Marine Bacterium of the New Genus Salinospora. *Angew. Chem. Int. Ed.* **42**, 355–357 (2003).

49. Newman, D. J. Predominately Uncultured Microbes as Sources of Bioactive Agents. *Front. Microbiol.* **7**, 1832 (2016).

50. Bewley, C. A., Holland, N. D. & Faulkner, D. J. Two Classes of Metabolites from *Theonella swinhoei* are Localized in Distinct Populations of Bacterial Symbionts. *Experientia.* **52**, 716–722 (1996).

51. Moss, C., Green, D.H., Pérez, B., Velasco, A., Henríquez, R., & McKenzie, J.D. Intracellular Bacteria Associated with the Ascidian *Ecteinascidia turbinata*: Phylogenetic and *in situ* Hybridisation Analysis. *Mar. Biol.* **143**, 99–110 (2003).

52. Rath, C. M. Janto, B., Earl, J., Ahmed, A., Hu, F.Z., Hiller, L., Dahlgren, M., Kreft, R., Yu, F., Wolff, J.J., Kweon, H.K., Christiansen, M.A., Håkansson, K., Williams, R.M., Ehrlich, G.D., & Sherman, D.H. Meta-omic Characterization of the Marine Invertebrate Microbial Consortium That Produces the Chemotherapeutic Natural Product ET-743. *ACS Chem. Biol.* **6**, 1244–1256 (2011).

53. Pérez-Matos, A. E., Rosado, W. & Govind, N. S. Bacterial Diversity Associated with the Caribbean Tunicate *Ecteinascidia turbinata*. *Anton. Leeuw. Int. J. G.* **92**, 155–164 (2007).

54. Haygood, M. G. & Davidson, S. K. Small-subunit rRNA Genes and *in situ* Hybridization with Oligonucleotides Specific for the Bacterial Symbionts in the Larvae of the Bryozoan *Bugula neritina* and Proposal of '*Candidatus* Endobugula sertula'. *Appl. Environ. Microbiol.* **63**, 4612–4616 (1997).

55. Davidson, S. K. & Haygood, M. G. Identification of Sibling Species of the Bryozoan *Bugula neritina* That Produce Different Anticancer Bryostatins and Harbor Distinct Strains of the Bacterial Symbiont '*Candidatus* Endobugula sertula'. *Biol. Bull.* **196**, 273–280 (1999).

56. Lopanik, N., Lindquist, N. & Targett, N. Potent Cytotoxins Produced by a Microbial Symbiont Protect Host Larvae from Predation. *Oecologia*. **139**, 131–139 (2004).

57. Sharp, K. H., Davidson, S. K. & Haygood, M. G. Localization of '*Candidatus* Endobugula sertula' and the Bryostatins Throughout the Life Cycle of the Bryozoan *Bugula neritina*. *ISME J.* **1**, 693–702 (2007).

58. Lopanik, N. B., Shields, J.A., Buchholz, T.J., Rath, C.M., Hothersall, J., Haygood, M.G., Håkansson, K., Thomas, C.M., & Sherman, D.H. In Vivo and In Vitro Trans-Acylation by BryP, the Putative Bryostatin Pathway Acyltransferase Derived from an Uncultured Marine Symbiont. *Chem. Biol.* **15**, 1175–1186 (2008).

59. Davidson, S. K., Allen, S. W., Lim, G. E., Anderson, C. M. & Haygood, M. G. Evidence for the Biosynthesis of Bryostatins by the Bacterial Symbiont "*Candidatus* Endobugula sertula" of the Bryozoan *Bugula neritina*. *Appl. Environ. Microbiol.* **67**, 4531–4537 (2001).

60. Wilson, M. C. & Piel, J. Metagenomic Approaches for Exploiting Uncultivated Bacteria as a Resource for Novel Biosynthetic Enzymology. *Chem. Biol.* **20**, 636–647 (2013).

61. Mori, T., Cahn, J.K.B., Wilson, M.C., Meoded, R.A., Wiebach, V., Martinez, A.F.C., Helfrich, E.J.N., Albersmeier, A., Wibberg, D., Dätwyler, S., Keren, R., Lavy, A., Rückert, C., Ilan, M., Kalinowski, J., Matsunaga, S., Takeyama, H., & Piel, J. Single-bacterial Genomics Validates Rich and Varied Specialized Metabolism of Uncultivated Entotheonella Sponge Symbionts. *Proc. Natl. Acad. Sci. U.S.A.* . **115**, 1718–1723 (2018).

62. Hinshaw, H. C., Pyle, M. M. & Feldman, W. H. Streptomycin in Tuberculosis. *Am. J. Med.* **2**, 429–435 (1947).

63. Wilson, W. R., Wilkowske, C. J., Wright, A. J., Sande, M. A. & Geraci, J. E. Treatment of Streptomycin-susceptible and Streptomycin-resistant *Enterococcal* endocarditis. *Ann. Intern. Med.* **100**, 816–823 (1984).

64. Spiegel, C. A. & Huycke, M. Endocarditis Due to Streptomycin-susceptible *Enterococcus faecalis* with High-level Gentamicin Resistance. *Arch. Intern. Med.* **149**, 1873–1875 (1989).

65. Spink, W. W., Hall, W. H., Shaffer, J. & Braude, A. I. Treatment of Brucellosis with Streptomycin and a Sulfonamide Drug. *JAMA* **139**, 352–356 (1949).

66. Wiley, P. F., Gerzon, K., Flynn, E. H., Sigal, M. V. & Quarck, U. C. Erythromycin. V. Isolation and Structure of Degradation Products. *J. Am. Chem. Soc.* **77**, 3677–3678 (1955).

67. Montgomery, C. H. & Knox, J. M. Erythromycin in the Treatment of Early Syphilis: A Preliminary Report. *Arch. Dermatol.* **80**, 205–209 (1959).

68. Bergquist, S.-O., Bernander, S., Dahnsjo, H. & Sundelof, B. O. Erythromycin in the Treatment of Pertussis: A Study of Bacteriologic and Clinical Effects. *J. Pediatr. Infecti. Dis.* **6**, 458–460 (1987).

69. Gould, J. C. Erythromycin in Respiratory Tract Infection. *Scott. Med. J.* **22**, 355–359 (1977).

70. Hawkyard, C. V. & Koerner, R. J. The Use of Erythromycin as a Gastrointestinal Prokinetic Agent in Adult Critical Care: Benefits Versus Risks. *J. Antimicrob. Chemother.* **59**, 347–358 (2007).

71. Zhang, H., Wang, Y., Wu, J., Skalina, K. & Pfeifer, B. A. Complete Biosynthesis of Erythromycin A and Designed Analogs Using E. coli as a Heterologous Host. *Chem. Biol.* **17**, 1232–1240 (2010).

72. Griffith, R. S. Introduction to Vancomycin. *Rev. Infect. Dis.* **3**, 200–204 (1981).

73. Choroba, O. W., Williams, D. H. & Spencer, J. B. Biosynthesis of the Vancomycin Group of Antibiotics: Involvement of an Unusual Dioxygenase in the Pathway to (*S*)-4-Hydroxyphenylglycine. *J. Am. Chem. Soc.* **122**, 5389–5390 (2000).

74. Yim, G., Thaker, M. N., Koteva, K. & Wright, G. Glycopeptide Antibiotic Biosynthesis. *J. Antibiot.* **67**, 31–41 (2014).

75. Mesa-Arango, A. C., Scorzoni, L. & Zaragoza, O. It Only Takes One to Do Many Jobs: Amphotericin B as Antifungal and Immunomodulatory Drug. *Front. Microbiol.* **3**, 286 (2012).

76. Torrado, J. J., Espada, R., Ballesteros, M. P. & Torrado-Santiago, S. Amphotericin B Formulations and Drug Targeting. *J. Pharm. Sci.* **97**, 2405–2425 (2008).

77. Caffrey, P., Lynch, S., Flood, E., Finnan, S. & Oliynyk, M. Amphotericin Biosynthesis in *Streptomyces nodosus*: Deductions from Analysis of Polyketide Synthase and Late Genes. *Chem. Biol.* **8**, 713–723 (2001).

78. Blum, R. H., Carter, S. K. & Agre, K. A Clinical Review of Bleomycin—A New Antineoplastic Agent. *Cancer.* **31**, 903–914 (1973).

79. Du, L., Sánchez, C., Chen, M., Edwards, D. J. & Shen, B. The Biosynthetic Gene Cluster for the Antitumor Drug Bleomycin from *Streptomyces verticillus* ATCC15003 Supporting Functional Interactions between Nonribosomal Peptide Synthetases and a Polyketide Synthase. *Chem. Biol.* **7**, 623–642 (2000).

80. Martins-Teixeira, M. B. & Carvalho, I. Antitumour Anthracyclines: Progress and Perspectives. *Chem. Med. Chem.* **15**, 933–948 (2020).

81. Yang, F., Kemp, C. J. & Henikoff, S. Anthracyclines Induce Double-strand DNA Breaks at Active Gene Promoters. *Mutat. Res-Fund. Mol. M.* **773**, 9–15 (2015).

82. Fujii, I. & Ebizuka, Y. Anthracycline Biosynthesis in *Streptomyces galilaeus*. *Chem. Rev.* **97**, 2511–2524 (1997).

83. Li, J., Kim, S. G. & Blenis, J. Rapamycin: One Drug, Many Effects. *Cell Metab.* **19**, 373–379 (2014).

84. Schinaman, J. M., Rana, A., Ja, W. W., Clark, R. I. & Walker, D. W. Rapamycin Modulates Tissue Aging and Lifespan Independently of the Gut Microbiota in Drosophila. *Sci. Rep.* **9**, 7824 (2019).

85. Goa, K. L., McTavish, D. & Clissold, S. P. Ivermectin. *Drugs* **42**, 640–658 (1991).

86. Campbell, W. C., Fisher, M. H., Stapley, E. O., Albers-Schönberg, G. & Jacob, T. A. Ivermectin: A Potent New Antiparasitic Agent. *Science.* **221**, 823–828 (1983).

87. Weibel, E. K., Hadvary, P., Hochuli, E., Kupfer, E. & Lengsfeld, H. Lipstatin, an Inhibitor of Pancreatic Lipase, produced by *Streptomyces toxytricini I*. Producing Organism, Fermentation, Isolation and Biological Activity. *J. Antibiot.* **40**, 1081–1085 (1987).

88. Bai, T., Zhang, D., Lin, S., Long, Q., Wang, Y., Ou, H., Kang, Q., Deng, Z., Liu, W., & Tao, M. Operon for Biosynthesis of Lipstatin, the Beta-Lactone Inhibitor of Human Pancreatic Lipase. *Appl. Environ. Microbiol.* **80**, 7473–7483 (2014).

89. Baranasic, D., Gacesa, R., Starcevic, A., Zucko, J., Blažič, M., Horvat, M., Gjuračić, K., Fujs, Š., Hranueli, D., Kosec, G., Cullum, J., & Petković, H. Draft Genome Sequence of *Streptomyces rapamycinicus* Strain NRRL 5491, the Producer of the Immunosuppressant Rapamycin. *Genome Announc.* **4**, e00581-13 (2013).

90. Goodwin, S., McPherson, J. D. & McCombie, W. R. Coming of Age: Ten Years of Next-Generation Sequencing Technologies. *Nat. Rev. Genet.* **17**, 333–351 (2016).

91. Amarasinghe, S. L., Su, S., Dong, X., Zappia, L., Ritchie, M.E., & Gouil, Q. Opportunities and Challenges in Long-Read Sequencing Data Analysis. *Genome Biol.* **21**, 30 (2020).

92. Crits-Christoph, A., Diamond, S., Butterfield, C. N., Thomas, B. C. & Banfield, J. F. Novel Soil Bacteria Possess Diverse Genes for Secondary Metabolite Biosynthesis. *Nature.* **558**, 440–444 (2018).

93. Zumberge, J. A., Love, G.D., Cárdenas, P., Sperling, E.A., Gunasekera, S., Rohrssen, M., Grosjean, E., Grotzinger, J.P., & Summons, R.E. Demosponge Steroid Biomarker 26-Methylstigmastane Provides Evidence for Neoproterozoic Animals. *Nat. Ecol. Evol.* **2**, 1709–1714 (2018).

94. Helber, S. B., Hoeijmakers, D. J. J., Muhando, C. A., Rohde, S. & Schupp, P. J. Sponge Chemical Defenses are a Possible Mechanism for Increasing Sponge Abundance on Reefs in Zanzibar. *PLoS one.* **13**, e0197617(2018).

95. Hentschel, U., Piel, J., Degnan, S. M. & Taylor, M. W. Genomic Insights into the Marine Sponge Microbiome. *Nat. Rev. Microbiol.* **10**, 641–654 (2012).

96. Hill, M., Hill, A., Lopez, N. & Harriott, O. Sponge-specific Bacterial Symbionts in the Caribbean Sponge, *Chondrilla nucula* (Demospongiae, Chondrosida). *Mar. Biol.* **148**, 1221–1230 (2006).

97. Usher, K. M., Kuo, J., Fromont, J. & Sutton, D. C. Vertical Transmission of Cyanobacterial Symbionts in the Marine Sponge *Chondrilla australiensis* (Demospongiae). *Hydrobiologia.* **461**, 9–13 (2001).

98. Hentschel, U., Hopke, J., Horn, M., Friedrich, A.B., Wagner, M., Hacker, J., & Moore, B.S. Molecular Evidence for a Uniform Microbial Community in Sponges from Different Oceans. *Appl. Environ. Microbiol.* **68**, 4431–4440 (2002).

99. Schmitt, S., Tsai, P., Bell, J., Fromont, J., Ilan, M., Lindquist, N., Perez, T., Rodrigo, A., Schupp, P.J., Vacelet, J., Webster, N., Hentschel, U., & Taylor, M.W. Assessing the

Complex Sponge Microbiota: Core, Variable and Species-Specific Bacterial Communities in Marine Sponges. *ISME J.* **6**, 564–576 (2012).

100. Wilson, M. C., Mori, T., Rückert, C., Uria, A.R., Helf, M.J., Takada, K., Gernert, C., Steffens, U.A.E., Heycke, N., Schmitt, S., Rinke, C., Helfrich, E.J.N., Brachmann, A.O., Gurgui, C., Wakimoto, T., Kracht, M., Crüsemann, M., Hentschel, U., Abe, I., Matsunaga, S., Kalinowski, J., Takeyama, H., & Piel, J. An Environmental Bacterial Taxon with a Large and Distinct Metabolic Repertoire. *Nature*. **506**, 58–62 (2014).

101. Schmidt, E. W., Obraztsova, A. Y., Davidson, S. K., Faulkner, D. J. & Haygood, M. G. Identification of the Antifungal Peptide-Containing Symbiont of the Marine Sponge *Theonella swinhoei* as a Novel δ-Proteobacterium, "*Candidatus* Entotheonella palauensis". *Mar. Biol.* **136**, 969–977 (2000).

102. Keren, R., Lavy, A., Mayzel, B. & Ilan, M. Culturable Associated-Bacteria of the Sponge *Theonella swinhoei* Show Tolerance to High Arsenic Concentrations. *Front. Microbiol.* **6**, 1-11 (2015).

103. Mehbub, M. F., Lei, J., Franco, C., & Zhang, W. Marine Sponge Derived Natural Products between 2001 and 2010: Trends and Opportunities for Discovery of Bioactives. *Mar. Drugs.* **12**, 4539–4577 (2014).

104. Piel, J., Hui, D., Wen, G., Butzke, D., Platzer, M., Fusetani, N., & Matsunaga, S., Antitumor Polyketide Biosynthesis by an Uncultivated Bacterial Symbiont of the Marine Sponge *Theonella swinhoei*. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 16222–16227 (2004).

105. Ueoka, R., Uria, A.R., Reiter, S., Mori, T., Karbaum, P., Peters, E.E., Helfrich, E.J.N., Morinaka, B.I., Gugger, M., Takeyama, H., Matsunaga, S., & Piel, J., Metabolic and Evolutionary Origin of Actin-Binding Polyketides from Diverse Organisms. *Nat. Chem. Biol.* **11**, 705–712 (2015).

106. Hamada, T., Matsunaga, S., Yano, G. & Fusetani, N. Polytheonamides A and B, Highly Cytotoxic, Linear Polypeptides with Unprecedented Structural Features, from the Marine Sponge, Theonella swinhoei. *J. Am. Chem. Soc.* **127**, 110–118 (2005).

107. Fusetani, N., Nakao, Y. & Matsunaga, S. Nazumamide A, a Thrombin-Inhibitory Tetrapeptide, from a Marine Sponge, *Theonella* sp. *Tetrahedron Lett.* **32**, 7073–7074 (1991).

108. Kobayashi, J., Sato, M., Murayama, T., Ishibashi, M., Wälchi, M.R., Kanai, M., Shoji, J., & Ohizumi, Y. Konbamide, a Novel Peptide with Calmoduiin Antagonistic Activity

from the Okinawan Marine Sponge *Theonella* sp. *J. Chem. Soc., Chem. Commun.* 1050–1052 (1991).

109. Fusetani, N., Matsunaga, S., Matsumoto, H. & Takebayashi, Y. Bioactive Marine Metabolites. 33. Cyclotheonamides, Potent Thrombin Inhibitors, from a Marine Sponge *Theonella* sp. *J. Am. Chem. Soc.* **112**, 7053–7054 (1990).

110. Kobayashi, J., Sato, M., Ishibashi, M., Shigemori, H., Nakamura, T. & Ohizumi, Y. Keramamide A, a Novel Peptide from the Okinawan Marine Sponge *Theonella* sp. *J. Chem. Soc., Perkin Trans.* **1** 2609–2611 (1991).

111. Sakemi, S., Ichiba, T., Kohmoto, S., Saucy, G. & Higa, . Isolation and Structure Elucidation of Onnamide A, a New Bioactive Metabolite of a Marine Sponge, *Theonella* sp. *J. Am. Chem. Soc.* **110**, 4851–4853 (1988).

112. Kato, Y., Fusetani, N., Matsunaga, S., Hashimoto, K., Sakai, R., Higa, T. & Kashman, Y. Antitumor Macrodiolides Isolated from a Marine Sponge *Theonella* sp.: Structure Revision of Misakinolide A. *Tetrahedron Lett.* **28**, 6225–6228 (1987).

113. Matsunaga, S., Fusetani, N., Hashimoto, K. & Walchli, M. Theonellamide F. A Novel Antifungal Bicyclic Peptide from a Marine Sponge *Theonella* sp. *J. Am. Chem. Soc.* **111**, 2582–2588 (1989).

114. Salituro, G. M. & Dufresne, C. Isolation by Low-Pressure Column Chromatography. in *Natural Products Isolation* (ed. Cannell, R. J. P.) 111–140 (Humana Press, 1998).

115. Kitagawa, I., Kobayashi, M., Lee, N.K., Shibuya, H., Kawata, Y. & Sakiyama, F. Structure of Theonellapeptolide Id, a New Bioactive Peptolide from an Okikawan Marine Sponge, *Theonella* sp. (Theonelliae). *Chem. Pharm. Bull.* **34**, 2664–2667 (1986).

116. Kobayashi, M., Kanzaki, K., Katayama, S., Ohashi, K., Okada, H. & Ikegami, S. Kitagawa, I. Marine Natural Products. XXXIII. Theonellapeptolide IId, a New Tridecapeptide Lactone from the Okinawan Marine Sponge *Theonella swinhoei*. *Chem. Pharm. Bull.* **42**, 1410–1415 (1994).

117. Roy, M. C., Ohtani, I.I., Ichiba, T., Tanaka, J., Satari, R. & Higa, T. New Cyclic Peptides from the Indonesian Sponge *Theonella swinhoei*. *Tetrahedron.* **56**, 9079–9092 (2000).

118. Harrison, A. G. & Young, A. B. Fragmentation of Protonated Oligoalanines: Amide Bond Cleavage and Beyond. *J. Am. Soc. Mass Spectrom*. **15**, 1810–1819 (2004).

119. Roepstorff, P. & Fohlman, J. Letter to the Editors. *Biomed. Mass Spectrom.* **11**, 601–601 (1984).

120. Marfey, P. Determination of D-Amino Acids. II. Use of a Bifunctional Reagent, 1,5-difluoro-2,4-dinitrobenzene. *Carlsberg Res. Commun.* **49**, 591–596 (1984).

121. Rezai, T., Yu, B., Millhauser, G. L., Jacobson, M. P. & Lokey, R. S. Testing the Conformational Hypothesis of Passive Membrane Permeability Using Synthetic Cyclic Peptide Diastereomers. *J. Am. Chem. Soc.* **128**, 2510–2511 (2006).

122. Calne, R. Y., Thiru, S., Mcmaster, P., Craddock, G.N., White, D.J.G., Evans, D.B., Dunn, D.C., Pentlow, B.D. & Rolles, K. Cyclosporin a in Patients Receiving Renal Allografts from Cadaver Donors. *The Lancet.* **312**, 1323–1327 (1978).

123. Dijkmans, B. A. C., van Rijthoven, A. W. A. M., Thè, H. S. G., Boers, M. & Cats, A. Cyclosporine in Rheumatoid Arthritis. *Semin. Arthritis Rheum.* **22**, 30–36 (1992).

124. Rüegger, A., Kuhn, M., Lichti, H., Loosli, H.-R., Huguenin, R., Quiquerez, C. & Wartburg, A. Cyclosporin A, a Peptide Metabolite from *Tvichoderrna polysporum* (LINK ex PERS.) *Rifai*, with a Remarkable Immunosuppressive Activity. *Helv. Chim. Acta.* **59**, 1075–1092 (1976).

125. Loor, F., Tiberghien, F., Wenandy, T., Didier, A. & Traber, R. Cyclosporins: Structure−Activity Relationships for the Inhibition of the Human MDR1 P-Glycoprotein ABC Transporter. *J. Med. Chem.* **45**, 4598–4612 (2002).

126. Mayer, A., Anke, H. & Sterner, O. Omphalotin, A New Cyclic Peptide with Potent Nematicidal Activity from Omphalotus Olearius I. Fermentation and Biological Activity. *Nat. Prod. Let.* **10**, 25–32 (1997).

127. Sueyoshi, K., Kaneda, M., Sumimoto, S., Oishi, S., Fujii, N., Suenaga, K. & Teruya, T. Odoamide, a Cytotoxic Cyclodepsipeptide from the Marine Cyanobacterium *Okeania* sp. *Tetrahedron.* **72**, 5472–5478 (2016).

128. Zhou, X., Huang, H., Li, J., Song, Y., Jiang, R., Liu, J., Zhang, S., Hua, Y. & Ju, J. New Anti-infective Cycloheptadepsipeptide Congeners and Absolute Stereochemistry from the Deep Sea-derived *Streptomyces drozdowiczii* SCSIO 10141. *Tetrahedron.* **70**, 7795–7801 (2014).

129. Semenova, M. N., Kiselyov, A. & Semenov, V. V. Sea Urchin Embryo as a Model Organism for the Rapid Functional Screening of Tubulin Modulators. *Bio. Tech.* **40**, 765–774 (2006).

130. Kobayashi, M., Lee, N. K., Shibuya, H., Momose, T. & Kitagawa, I. Marine Natural Products. XXVI. Biologically Active Tridecapeptide Lactones from the Okinawan Marine Sponge *Theonella swinhoei* (Theonellidae).(2).Structures of Theonellapeptolides Ia, Ib, Ic, and Ie. *Chem. Pharm. Bull.* **39**, 1177–1184 (1991).

131. Sinisi, A., Calcinai, B., Cerrano, C., Dien, H.A., Zampella, A., D'Amore, C., Renga, B., Fiorucci, S. & Taglialatela-Scafati, O. New Tridecapeptides of the Theonellapeptolide Family from the Indonesian Sponge *Theonella swinhoei*. *Beilstein J. Org. Chem.* **9**, 1643–1651 (2013).

132. Tsuda, M., Shimbo, K., Kubota, T., Mikami, Y. & Kobayashi, J. Two Theonellapeptolide Congeners from Marine Sponge *Theonella* sp. *Tetrahedron.* **55**, 10305–10314 (1999).

133. Kitagawa, I., Ohashi, K., Kawanishi, H., Shibuya, H., Shinkai, K. & Akedo, H. Ionophoretic Activities of Oligopeptide Lactones and Resin-Glycosides in Human Erytherocytes. *Chem. Pharm. Bull.* **37**, 1679–1681 (1989).

134. Li, S., Dumdei, E.J., Blunt, J.W., Munro, M.H.G., Robinson, W.T. & Pannell, L.K. Theonellapeptolide IIIe, a New Cyclic Peptolide from the New Zealand Deep Water Sponge, *Lamellomorpha strongylata*. *J. Nat. Prod.* **61**, 724–728 (1998).

135. Izuishi, K., Kato, K., Ogura, T., Kinoshita, T. & Esumi, H. Remarkable Tolerance of Tumor Cells to Nutrient Deprivation: Possible New Biochemical Target for Cancer Therapy. *Cancer Res.* **60**, 6201–6207 (2000).

136. Carmeliet, P. & Jain, R. K. Angiogenesis in cancer and other diseases. *Nature.* **407**, 249–257 (2000).

137. Lu, J., Kunimoto, S., Yamazaki, Y., Kaminishi, M. & Esumi, H. Kigamicin D, a Novel Anticancer Agent Based on a New Anti-Austerity Strategy Targeting Cancer Cells' Tolerance to Nutrient Starvation. *Cancer Sci.* **95**, 547–552 (2004).

138. Awale, S., Okada, T., Dibwe, D.F., Maruyama, T., Takahara, S., Okada, T., Endo, S. & Toyooka, N. Design and Synthesis of Functionalized Coumarins as Potential Anti-austerity Agents that Eliminates Cancer Cells' Tolerance to Nutrition Starvation. *Bio. Med. Chem. Lett.* **29**, 1779–1784 (2019).

139. Devji, T., Reddy, C., Woo, C., Awale, S., Kadota, S. & Carrico-Moniz, D. Pancreatic Anticancer Activity of a Novel Geranylgeranylated Coumarin Derivative. *Bio. Med. Chem. Lett.* **21**, 5770–5773 (2011).

140. Esumi, H., Lu, J., Kurashima, Y. & Hanaoka, T. Antitumor Activity of Pyrvinium Pamoate, 6-(dimethylamino)-2-[2-(2,5-dimethyl-1-phenyl-1H-pyrrol-3-yl)ethenyl]-1-methyl-Quinolinium Pamoate Salt, Showing Preferential Cytotoxicity during Glucose Starvation. *Cancer Sci.* **95**, 685–690 (2004).

141. Farley, C. M., Dibwe, D.F., Ueda, J., Hall, E.A., Awale, S. & Magolan, J. Evaluation of Synthetic Coumarins for Antiausterity Cytotoxicity Against Pancreatic Cancers. *Bio. Med. Chem. Lett.* **26**, 1471–1474 (2016).

142. Jun, M., Bacay, A. F., Moyer, J., Webb, A. & Carrico-Moniz, D. Synthesis and Biological Evaluation of Isoprenylated Coumarins as Potential Anti-pancreatic Cancer Agents. *Bio. Med. Chem. Lett.* **24**, 4654–4658 (2014).

143. Awale, S., Dibwe, D.F., Balachandran, C., Fayez, S., Feineis, D., Lombe, B.K. & Bringmann, G., Ancistrolikokine E3, a 5,8′-Coupled Naphthylisoquinoline Alkaloid, Eliminates the Tolerance of Cancer Cells to Nutrition Starvation by Inhibition of the Akt/mTOR/Autophagy Signaling Pathway. *J. Nat. Prod.* **81**, 2282–2291 (2018).

144. Fayez, S., Feineis, D., Aké Assi, L., Kaiser, M., Brun, R., Awale, S. & Bringmann, Ancistrobrevines E-J and Related Naphthylisoquinoline Alkaloids from the West African Liana *Ancistrocladus abbreviatus* with Inhibitory Activities Against *Plasmodium falciparum* and PANC-1 Human Pancreatic Cancer Cells. *Fitoterapia.* **131**, 245–259 (2018).

145. Kavatsurwa, S. M., Lombe, B.K., Feineis, D., Dibwe, D.F., Maharaj, V., Awale, S. & Bringmann, G. Ancistroyafungines A-D, 5,8′- and 5,1′-coupled Naphthylisoquinoline Alkaloids from a Congolese *Ancistrocladus* Species, with Antiausterity Activities Against Human PANC-1 Pancreatic Cancer Cells. *Fitoterapia.* **130**, 6–16 (2018).

146. Lombe, B. K., Feineis, D., Mudogo, V., Brun, R., Awale, S. & Bringmann, G. Michellamines A6 and A7, and Further Mono- and Dimeric Naphthylisoquinoline Alkaloids from a Congolese *Ancistrocladus liana* and Their Antiausterity Activities Against Pancreatic Cancer Cells. *RSC Adv.* **8**, 5243–5254 (2018).

147. Magolan, J., Adams, N.B.P., Onozuka, H., Hungerford, N.L., Esumi, H. & Coster, M.J. Synthesis and Evaluation of Anticancer Natural Product Analogues Based on Angelmarin: Targeting the Tolerance towards Nutrient Deprivation. *Chem. Med. Chem.* **7**, 766–770 (2012).

148. Abdel-Naime, W. A., Kimishima, A., Setiawan, A., Fahim, J.R., Fouad, M.A., Kamel, M.S. & Arai, M. Mitochondrial Targeting in an Anti-Austerity Approach Involving Bioactive Metabolites Isolated from the Marine-Derived Fungus Aspergillus sp. *Mar. Drugs.* **18**, 555 (2020).

149. Velkov, T., Horne, J., Scanlon, M.J., Capuano, B., Yuriev, E. & Lawen, A., Characterization of the N-Methyltransferase Activities of the Multifunctional Polypeptide Cyclosporin Synthetase. *Chem. Biol.* **18**, 464–475 (2011).

150. Rachid, S., Krug, D., Kunze, B., Kochems, I., Scharfe, M., Zabriskie, T.M., Blöcker, H. & Müller, R. Molecular and Biochemical Studies of Chondramide Formation—Highly Cytotoxic Natural Products from *Chondromyces crocatus* Cm c5. *Chem. Biol.* **13**, 667–681 (2006).

151. Mayer, A., Anke, H. & Sterner, O. Omphalotin, A New Cyclic Peptide with Potent Nematicidal Activity from Omphalotus Olearius I. Fermentation and Biological Activity. *Nat. Prod. Lett.* **10**, 25–32 (1997).

152. Sterner, O., Etzel, W., Mayer, A. & Anke, H. Omphalotin, A New Cyclic Peptide with Potent Nematicidal Activity from Omphalotus Olearius II. Isolation and Structure Determination. *Nat. Prod. Lett.* **10**, 33–38 (1997).

153. Ramm, S., Krawczyk, B., Mühlenweg, A., Poch, A., Mösker, E. & Süssmuth, R.D. A Self-Sacrificing N-Methyltransferase is the Precursor of the Fungal Natural Product Omphalotin. *Angew. Chem. Int. Ed.* **56**, 9994–9997 (2017).

154. Freeman, M. F., Gurgui, C., Helf, M.J., Morinaka, B.I., Uria, A.R., Oldham, N.J., Sahl, H.-G., Matsunaga, S. & Piel, J., Metagenome Mining Reveals Polytheonamides as Posttranslationally Modified Ribosomal Peptides. *Science.* **338**, 387–390 (2012).

155. Marahiel, M. A., Stachelhaus, T. & Mootz, H. D. Modular Peptide Synthetases Involved in Nonribosomal Peptide Synthesis. *Chem. Rev.* **97**, 2651–2674 (1997).

156. Gougoulias, C., Clark, J. M. & Shaw, L. J. The Role of Soil Microbes in the Global Carbon Cycle: Tracking the Below-ground Microbial Processing of Plant-derived Carbon for Manipulating Carbon Dynamics in Agricultural Systems. *J. Sci. Food Agric.* **94**, 2362–2371 (2014).

157. Pajares, S. & Bohannan, B. J. M. Ecology of Nitrogen Fixing, Nitrifying, and Denitrifying Microorganisms in Tropical Forest Soils. *Front. Microbiol.* **7**, 1045 (2016).

158. Kudoyarova, G., Arkhipova, T., Korshunova, T., Bakaeva, M., Loginov, O. & Dodd, I.C. Phytohormone Mediation of Interactions Between Plants and Non-Symbiotic Growth Promoting Bacteria Under Edaphic Stresses. *Front. Plant Sci.* **10**, 1368 (2019).

159. Hayat, S., Faraz, A. & Faizan, M. Root Exudates: Composition and Impact on Plant–Microbe Interaction. in *Biofilms in Plant and Soil Health* 179–193 (John Wiley & Sons, Ltd, 2017).

160. Olanrewaju, O. S., Ayangbenro, A. S., Glick, B. R. & Babalola, O. O. Plant Health: Feedback Effect of Root Exudates-Rhizobiome Interactions. *Appl. Microbiol. Biotechnol.* **103**, 1155–1166 (2019).

161. Hibbing, M. E., Fuqua, C., Parsek, M. R. & Peterson, S. B. Bacterial Competition: Surviving and Thriving in the Microbial Jungle. *Nat. Rev. Microbiol.* **8**, 15–25 (2010).

162. Cassinelli, G., Rivola, G., Ruggieri, D., Arcamone, F., Grein, A., Merli, S., Spalla, C., Casazza, A.M., Marco, A.D. & Pratesi, G. New Anthracycline Glycosides: 4-O-Demethyl-11-Deoxydoxorubicin and Analogues from *Streptomyces peucetius* Var. Aureus. *J. Antibiot.* **35**, 176–183 (1982).

163. Baroja-Mazo, A., Revilla-Nuin, B., Ramírez, P. & Pons, J. A. Immunosuppressive Potency of Mechanistic Target of Rapamycin Inhibitors in Solid-organ Transplantation. *World J. Transplant.* **6**, 183–192 (2016).

164. Wang, T., Liang, Y., Wu, M., Chen, Z., Lin, J. & Yang, L. Natural Products from *Bacillus subtilis* with Antimicrobial Properties. *Chin. J. Chem. Eng.* **23**, 744–754 (2015).

165. Kaspar, F., Neubauer, P. & Gimpel, M. Bioactive Secondary Metabolites from *Bacillus subtilis*: A Comprehensive Review. *J. Nat. Prod.* **82**, 2038–2053 (2019).

166. Herrmann, J., Fayad, A. A. & Müller, R. Natural Products from Myxobacteria: Novel Metabolites and Bioactivities. *Nat. Prod. Rep.* **34**, 135–160 (2017).

167. Kunakom, S. & Eustáquio, A. S. *Burkholderia* as a Source of Natural Products. *J. Nat. Prod.* **82**, 2018–2037 (2019).

168. Segata, N., Waldron, L., Ballarini, A., Narasimhan, V., Jousson, O. & Huttenhower, C. Metagenomic Microbial Community Profiling Using Unique Clade-specific Marker Genes. *Nat. Methods.* **9**, 811–814 (2012).

169. Robinson, S. L., Piel, J. & Sunagawa, S. A Roadmap for Metagenomic Enzyme Discovery. *Nat. Prod. Rep.* **38**, 1994–2023 (2021).

170. Aoi, Y., Kinoshita, T., Hata, T., Ohta, H., Obokata, H. & Tsuneda, S., Hollow-Fiber Membrane Chamber as a Device for In Situ Environmental Cultivation. *App. Environ. Microbiol.* **75**, 3826–3833 (2009).

171. Connon, S. A. & Giovannoni, S. J. High-Throughput Methods for Culturing Microorganisms in Very-Low-Nutrient Media Yield Diverse New Marine Isolates. *App. Environ. Microbiol.* **68**, 3878–3885 (2002).

172. Ferrari, B. C., Binnerup, S. J. & Gillings, M. Microcolony Cultivation on a Soil Substrate Membrane System Selects for Previously Uncultured Soil Bacteria *App. Environ. Microbiol.* **71**, 8714–8720 (2005).

173. Ferrari, B. C. & Gillings, M. R. Cultivation of Fastidious Bacteria by Viability Staining and Micromanipulation in a Soil Substrate Membrane System. *App. Environ. Microbiol.* **75**, 3352–3354 (2009).

174. Nichols, D., Cahoon, N., Trakhtenberg, E.M., Pham, L., Mehta, A., Belanger, A., Kanigan, T., Lewis, K. & Epstein, S.S. Use of Ichip for High-Throughput *In situ* Cultivation of "Uncultivable" Microbial Species. *Appl. Environ. Microbiol.* **76**, 2445–2450 (2010).

175. Kaeberlein, T., Lewis, K. & Epstein, S. S. Isolating 'Uncultivable' Microorganisms in Pure Culture in a Simulated Natural Environment. *Science.* **296**, 1127–1129 (2002).

176. Berdy, B., Spoering, A. L., Ling, L. L. & Epstein, S. S. *In situ* Cultivation of Previously Uncultivable Microorganisms Using the Ichip. *Nat. Prot.* **12**, 2232–2242 (2017).

177. Ling, L. L., Schneider, T., Peoples, A.J., Spoering, A.L., Engels, I., Conlon, B.P., Mueller, A., Schäberle, T.F., Hughes, D.E., Epstein, S., Jones, M., Lazarides, L., Steadman, V.A., Cohen, D.R., Felix, C.R., Fetterman, K.A., Millett, W.P., Nitti, A.G., Zullo, A.M., Chen, C., Lewis, K. A New Antibiotic Kills Pathogens without Detectable Resistance. *Nature.* **517**, 455–459 (2015).

178. MacIntyre, L. W., Charles, M. J., Haltli, B. A., Marchbank, D. H. & Kerr, R. G. An Ichip-Domesticated Sponge Bacterium Produces an N-Acyltyrosine Bearing an α-Methyl Substituent. *Org. Lett.* **21**, 7768–7771 (2019).

179. D'Onofrio, A., Crawford, J.M., Stewart, E.J., Witt, K., Gavrish, E., Epstein, S., Clardy, J. & Lewis, K. Siderophores from Neighboring Organisms Promote the Growth of Uncultured Bacteria. *Chem. Biol.* **17**, 254–264 (2010).

180. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci.* U. S. A. **82,** 4576–4579 (1990).

181. Hardy, C. D. & Butler, A. Ambiguity of NRPS Structure Predictions: Four Bidentate Chelating Groups in the Siderophore Pacifibactin. *J. Nat. Prod.* **82**, 990–997 (2019).

182. Baars, O., Zhang, X., Gibson, M.I., Stone, A.T., Morel, F.M.M. & Seyedsayamdost, M.R. Crochelins: Siderophores with an Unprecedented Iron-Chelating Moiety from the Nitrogen-Fixing Bacterium *Azotobacter chroococcum*. *Angew. Chem. Int. Ed.* **57**, 536–541 (2018).

183. Vinnik, V., Zhang, F., Park, H., Cook, T.B., Throckmorton, K., Pfleger, B.F., Bugni, T.S. & Thomas, M.G. Structural and Biosynthetic Analysis of the Fabrubactins, Unusual Siderophores from *Agrobacterium fabrum* Strain C58. *ACS Chem. Biol.* **16**, 125–135 (2021).

184. Miethke, M. & Marahiel, M. A. Siderophore-Based Iron Acquisition and Pathogen Control. *Microbiol. Mol. Biol. Rev.* **71**, 413–451 (2007).

185. Sun, S.-L., Yang, W.-L., Fang, W.-W., Zhao, Y.-X., Guo, L. & Dai, Y.-J. The Plant Growth-Promoting Rhizobacterium *Variovorax boronicumulans* CGMCC 4969 Regulates the Level of Indole-3-Acetic Acid Synthesized from Indole-3-Acetonitrile. *Appl. Environ. Microbiol.* **84**, e00298-18 (2018).

186. Garcia Teijeiro, R., Belimov, A. A. & Dodd, I. C. Microbial Inoculum Development for Ameliorating Crop Drought Stress: A Case Study of *Variovorax paradoxus* 5C-2. *N. Biotechnol.* **56**, 103–113 (2020).

187. Kim, Y.-N., Khan, M. A., Kang, S.-M., Hamayun, M. & Lee, and I.-J. Enhancement of Drought-Stress Tolerance of *Brassica oleracea* var. italica L. by Newly Isolated *Variovorax* sp. YNA59. *J. Microbiol. Biotechnol*. **30**, 1500–1509 (2020).

188. Larke-Mejía, N. L., Crombie, A. T., Pratscher, J., McGenity, T. J. & Murrell, J. C. Novel Isoprene-Degrading Proteobacteria from Soil and Leaves Identified by Cultivation and Metagenomics Analysis of Stable Isotope Probing Experiments. *Front. Microbiol.* **10**, 2700 (2019).

189. Dhakar, K., Zarecki, R., van Bommel, D., Knossow, N., Medina, S., Öztürk, B., Aly, R., Eizenberg, H., Ronen, Z. & Freilich, S. Strategies for Enhancing in vitro Degradation of

Linuron by *Variovorax* sp. Strain SRS 16 Under the Guidance of Metabolic Modeling. *Front. Bioeng. Biotechnol.* **9**, 191 (2021).

190. Heine, V., Meinert-Berning, C., Lück, J., Mikowsky, N., Voigt, B., Riedel, K. & Steinbüchel, A. The Catabolism of 3,3'-thiodipropionic Acid in *Variovorax paradoxus* Strain TBEA6: A Proteomic Analysis. *PLOS ONE.* **14**, e0211876 (2019).

191. Johnston, C. W., Skinnider, M.A., Wyatt, M.A., Li, X., Ranieri, M.R.M., Yang, L., Zechel, D.L., Ma, B. & Magarvey, N.A. An Automated Genomes-to-Natural Products Platform (GNP) for the Discovery of Modular Natural Products. *Nat. Comm.* **6**, 8421 (2015).

192. Kurth, C., Schieferdecker, S., Athanasopoulou, K., Seccareccia, I. & Nett, M. Variochelins, Lipopeptide Siderophores from *Variovorax boronicumulans* Discovered by Genome Mining. *J. Nat. Prod.* **79**, 865–872 (2016).

193. Robertson, A. W., McCarville, N.G., MacIntyre, L.W., Correa, H., Haltli, B., Marchbank, D.H. & Kerr, R.G. Isolation of Imaqobactin, an Amphiphilic Siderophore from the Arctic Marine Bacterium *Variovorax* Species RKJM285. *J. Nat. Prod.* **81**, 858–865 (2018).

194. Himpsl, S. D. & Mobley, H. L. T. Siderophore Detection Using Chrome Azurol S and Cross-Feeding Assays. *Methods Mol. Biol.* **2021**, 97–108 (2019).

195. Fuchs, S. W., Sachs, C.C., Kegler, C., Nollmann, F.I., Karas, M. & Bode, H.B. Neutral Loss Fragmentation Pattern Based Screening for Arginine-Rich Natural Products in *Xenorhabdus* and *Photorhabdus*. *Anal. Chem.* **84**, 6948–6955 (2012).

196. Bosello, M., Zeyadi, M., Kraas, F.I., Linne, U., Xie, X. & Marahiel, M.A., Structural Characterization of the Heterobactin Siderophores from *Rhodococcus erythropolis* PR4 and Elucidation of Their Biosynthetic Machinery. *J. Nat. Prod.* **76**, 2282–2290 (2013).

197. Johnston, C. W., Wyatt, M.A., Li, X., Ibrahim, A., Shuster, J., Southam, G. & Magarvey, N.A. Gold Biomineralization by a Metallophore from a Gold-associated Microbe. *Nat. Chem. Biol.* **9**, 241–243 (2013).

198. Chiu, H.-H. & Kuo, C.-H. Gas Chromatography-Mass Spectrometry-based Analytical Strategies for Fatty Acid Analysis in Biological Samples. *J. Food Drug Anal.* **28**, 60–73 (2020).

199. Zhao, Y., Yang, S. Z. & Mu, B. Z. Quantitative Analyses of the Isoforms of Surfactin Produced by *Bacillus subtilis* HSO 121 Using GC-MS. *Anal. Sci.* **28**, 789–793 (2012).

200. Yakimov, M. M., Timmis, K. N., Wray, V. & Fredrickson, H. L. Characterization of a New Lipopeptide Surfactant Produced by Thermotolerant and Halotolerant Subsurface *Bacillus licheniformis* BAS50. *App. Environ. Microbiol.* **61**, 1706–1713 (1995).

201. Lyu, Y., Zhang, T., Dou, B., Li, G., Ma, C. & Li, Y. A Lipopeptide Biosurfactant from *Bacillus* sp. Lv13 and Their Combined Effects on Biodesulfurization of Dibenzothiophene. *RSC Adv.* **8**, 38787–38791 (2018).

202. Yang, S.-Z., Wei, D.-Z. & Mu, B.-Z. Determination of the Structure of the Fatty Acid Chain in a Cyclic Lipopeptide Using GC–MS. *J. Biochem. Biophysic. Met.* **70**, 519–523 (2007).

203. Quehenberger, O., Armando, A. M. & Dennis, E. A. High Sensitivity Quantitative Lipidomics Analysis of Fatty Acids in Biological Samples by Gas Chromatography–Mass Spectrometry. *Biochim. Biophys. Acta Mol. Cell Biol. Lipids.* **1811**, 648–656 (2011).

204. Åkesson-Nilsson, G. & Wesén, C. Structural Characterization of 5,6-dichlorotetradecanoic Acid, an Isolated Metabolite of 9,10-dichlorooctadecanoic Acid, by Studying Picolinyl Esters, Pyrrolidides and Methyl Esters with Electron Ionization Mass Spectrometry. *J. Mass Spec.* **39**, 1313–1320 (2004).

205. Kawashima, H. & Ohnishi, M. Novel Heneicosadienoic and Tricosadienoic Acid Isomers in Ovaries of Marine Archaeogastropods. *Lipids* **47**, 827–833 (2012).

206. Christie, W. W., Dobson, G. & Adlof, R. O. A Practical Guide to the Isolation, Analysis and Identification of Conjugated Linoleic Acid. *Lipids* **42**, 1073–1084 (2007).

207. Lamberto, M. & Ackman, R. G. Positional Isomerization of trans-3-Hexadecenoic Acid Employing 2-Amino-2-methyl-propanol as a Derivatizing Agent for Ethylenic Bond Location by Gas-Chromatography/Mass Spectrometry. *Anal. Biochem.* **230**, 224–228 (1995).

208. Antipov, D., Korobeynikov, A., McLean, J. S. & Pevzner, P. A. HybridSPAdes: An Algorithm for Hybrid Assembly of Short and Long Reads. *Bioinformatics.* **32**, 1009–1015 (2016).

209. Blin, K., Shaw, S., Kloosterman, A.M., Charlop-Powers, Z., van Wezel, G.P., Medema, M.H. & Weber, T. antiSMASH 6.0: Improving Cluster Detection and Comparison Capabilities. *Nucleic Acids Res.* **49**, W29–W35 (2021).

210. Bachmann, B. O. & Ravel, J. Chapter 8 Methods for In Silico Prediction of Microbial Polyketide and Nonribosomal Peptide Biosynthetic Pathways from DNA Sequence Data. in *Methods in Enzymology* vol. 458 181–217 (Academic Press, 2009).

211. Reitz, Z. L., Hardy, C. D., Suk, J., Bouvet, J. & Butler, A. Genomic Analysis of Siderophore β-hydroxylases Reveals Divergent Stereocontrol and Expands the Condensation Domain Family. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 19805–19814 (2019).

212. Heemstra, J. R., Walsh, C. T. & Sattely, E. S. Enzymatic Tailoring of Ornithine in the Biosynthesis of the Rhizobium Cyclic Trihydroxamate Siderophore Vicibactin. *J. Am. Chem. Soc.* **131**, 15317–15329 (2009).

213. Stephan, H., Freund, S., Beck, W., Jung, G., Meyer, J.-M. & Winkelmann, G. Ornibactins- A New Family of Siderophores from *Pseudomonas*. *Biometals.* **6**, 93–100 (1993).

214. Lautru, S., Deeth, R. J., Bailey, L. M. & Challis, G. L. Discovery of a New Peptide Natural Product by *Streptomyces coelicolor* Genome Mining. *Nat. Chem. Biol.* **1**, 265–269 (2005).

215. Pohlmann, V. & A. Marahiel, M. δ-Amino Group Hydroxylation of L-Ornithine during Coelichelin Biosynthesis. *Org. Biomol. Chem.* **6**, 1843–1848 (2008).

216. Bosello, M., Mielcarek, A., Giessen, T. W. & Marahiel, M. A. An Enzymatic Pathway for the Biosynthesis of the Formylhydroxyornithine Required for Rhodochelin Iron Coordination. *Biochemistry.* **51**, 3059–3066 (2012).

217. Felnagle, E. A., Barkei, J.J., Park, H., Podevels, A.M., McMahon, M.D., Drott, D.W. & Thomas, M.G. MbtH-Like Proteins as Integral Components of Bacterial Nonribosomal Peptide Synthetases. *Biochemistry.* **49**, 8815–8817 (2010).

218. Kotowska, M. & Pawlik, K. Roles of Type II Thioesterases and Their Application for Secondary Metabolite Yield Improvement. *Appl. Microbiol. Biotechnol.* **98**, 7735–7746 (2014).

219. Lambalot, R. H., Gehring, A.M., Flugel, R.S., Zuber, P., LaCelle, M., Marahiel, M.A., Reid, R., Khosla, C. & Walsh, C.T. A New Enzyme Superfamily — the Phosphopantetheinyl Transferases. *Chem. Biol.* **3**, 923–936 (1996).

220. Brown, P. D. Matrix Metalloproteinase Inhibitors: A Novel Class of Anticancer Agents. *Adv. Enzym. Regul.* **35**, 293–301 (1995).

221. Zhu, Z., Mazzola, R., Sinning, L., McKittrick, B., Niu, X., Lundell, D., Sun, J., Orth, P., Guo, Z., Madison, V., Ingram, R. & Beyer, B.M. Discovery of Novel Hydroxamates as Highly Potent Tumor Necrosis Factor-α Converting Enzyme Inhibitors: Part I—Discovery of Two Binding Modes. *J. Med. Chem.* **51**, 725–736 (2008).

222. Flipo, M., Beghyn, T., Leroux, V., Florent, I., Deprez, B.P. & Deprez-Poulain, R.F. Novel Selective Inhibitors of the Zinc Plasmodial Aminopeptidase PfA-M1 as Potential Antimalarial Agents. *J. Med. Chem.* **50**, 1322–1334 (2007).

223. Chen, Y., Lopez-Sanchez, M., Savoy, D.N., Billadeau, D.D., Dow, G.S. & Kozikowski, A.P. A Series of Potent and Selective, Triazolylphenyl-Based Histone Deacetylases Inhibitors with Activity against Pancreatic Cancer Cells and *Plasmodium falciparum*. *J. Med. Chem.* **51**, 3437–3448 (2008).

224. Keberle, H. The Biochemistry of Desferrioxamine and its Relation to Iron Metabolism. *Ann. N. Y. Acad. Sci.* **119**, 758–768 (1964).

225. Vraspir, J. M., Holt, P. D. & Butler, A. Identification of New Members within Suites of Amphiphilic Marine Siderophores. *Biometals.* **24**, 85–92 (2011).

226. Noinaj, N., Guillier, M., Barnard, T. J. & Buchanan, S. K. TonB-Dependent Transporters: Regulation, Structure, and Function. *Annu. Rev. Microbiol.* **64**, 43–60 (2010).

227. Karpishin, T. B. & Raymond, K. N. The First Structural Characterization of a Metal–Enterobactin Complex: [V(enterobactin)]2^-. *Angew. Chem. Int. Ed.* **31**, 466–468 (1992).

228. Miller, M. C., Parkin, S., Fetherston, J. D., Perry, R. D. & DeMoll, E. Crystal Structure of Ferric-Yersiniabactin, a Virulence Factor of *Yersinia pestis*. *J. Inorg. Biochem.* **100**, 1495–1500 (2006).

229. Konetschny-Rapp, S., Jung, G., Meiwes, J. & Zähner, H. Staphyloferrin A: a Structurally New Siderophore from *Staphylococci*. *Eur. J. Biochem.* **191**, 65–74 (1990).

230. Lorenzo, V. de, Bindereif, A., Paw, B. H. & Neilands, J. B. Aerobactin Biosynthesis and Transport Genes of Plasmid ColV-K30 in *Escherichia coli* K-12. *J. Bacteriol.* **165**, 570–578 (1986).

231. Schwyn, B. & Neilands, J. B. Universal Chemical Assay for the Detection and Determination of Siderophores. *Anal. Biochem.* **160**, 47–56 (1987).

232. Neilands, J. B. Siderophores: Structure and Function of Microbial Iron Transport Compounds. *J. Biol. Chem.* **270**, 26723–26726 (1995).

Acknowledgements

Firstly, I would like to thank my adviser, Prof. Toshiyuki Wakimoto, for his guidance and support throughout my graduate studies. The mentorship and laboratory environment he provided allowed me to develop research independence to acquire skillsets that importance for my future career. I would also like to thank Dr. Kenichi Matsuda, Dr. Aya Yoshimura, Dr. Agustinus R. Uria, for their continued support, insight, and guidance throughout my thesis project as well as for my committee members Prof. Satoshi Ichikawa and Dr. Fumika Yakushiji for critical input and valuable discussion. I would like also to appreciate all the collaborator Dr. Dya Fita Dibwe, Dr. Subehan, Ismail, and Masachika Takada for assistance to complete the project requirement.

Secondly, I would like to express my gratitude to the former laboratory staffs and members, Dr. Egami, Dr. Kuranaga, Dr. Tan, Enomoto, Shiroyama, Takatsuki, Takaoka, Ito, Sano, Ozawa, Yamamoto, Suzuki, Fukuba, Hayashi, Toda, Tachikawa, Abe, Akiyama, and Jomori for all support and memories from the beginning of my studies in this laboratory.

Finally, my great appreciation to INPEX scholarship foundation and Toshimi Otsuka scholarship foundation for supporting me financially. I would also like to thank all of my friends along with my wonderful family for providing constant support and encouragement throughout my study. Whether it was being understanding of my feelings, physical absence, lending an ear, or just giving me an occasional word of encouragement. I would like to thank my personal support system Fitri Anggareni D.N.H for can always spare a special time to discuss anything.