Nonribosomal Peptides Produced by *Xenorhabdus* and *Photorhabdus*

Dissertation

zur Erlangung des Doktorgrades

der Naturwissenschaften

vorgelegt beim Fachbereich für Biowissenschaften (15)

der Johann Wolfgang Goethe-Universität

in Frankfurt am Main

von

Lei Zhao

Frankfurt am Main 2020

D30

vom Fachbereich für Biowissenschaften (15) der

Johann Wolfgang Goethe-Universität als Dissertation angenommen.

Dekan: Prof. Dr. Sven Klimpel

Gutachter: Prof. Dr. Helge B. Bode

Zweitgutachter: Prof. Dr. Martin Grininger

Datum der Disputation: 12.08.2020

Table of Contents

List of Abbreviations

1D NMR	one-dimensional nuclear magnetic resonance
2D NMR	two-dimensional nuclear magnetic resonance
A domain	adenylation domain
A_{core}	N-terminal core domain of the adenylation domain
ACP	acyl carrier protein
AMP	adenosine monophosphate
antiSMASH	antibiotics and secondary metabolite analysis shell
A_{sub}	C-terminal subdomain of the adenylation domain
ATP	adenosine triphosphate
BGC	biosynthetic gene cluster
BPC	base peak chromatogram
C domain	condensation domain
C/E domain	dual condensation/epimerization domain
CLP	cyclic lipopeptide
COSY	correlation spectroscopy
C_{start} domain	starter condensation domain
C_{term} domain	terminal condensation domain
E domain	epimerization domain
ECD	electronic circular dichroism
EIC	extracted ion chromatogram
ESI	electrospray ionization
FT domain	formyltransferase domain
fTHF	N^{10}-formyltetrahydrofolate
HDAC	histone deacetylase
HMBC	heteronuclear multiple bond correlation
HPLC	high performance liquid chromatography
HRMS	high resolution mass spectrometry
HSQC	heteronuclear single quantum coherence
IC_{50}	50% inhibitory concentration
IJ	infective juvenile
LB	lysogeny broth

List of Abbreviations

m/*z*	mass-to-charge ratio
MALDI	matrix-assisted laser desorption/ionization
MIC	minimum inhibitory concentration
MS	mass spectrometry
MS^2 or MS/MS	tandem mass spectrometry
MT domain	methyltransferase domain
NMR	nuclear magnetic resonance
NP	natural product
NRP	nonribosomal peptide
NRPS	nonribosomal peptide synthetase
P$_{BAD}$	L-arabinose inducible promoter
pCEP	cluster expression plasmid
PCP	peptidyl carrier protein
PCR	polymerase chain reaction
PKS	polyketide synthase
PO	phenoloxidase
Ppan	4'-phosphopantetheinyl
PP$_i$	pyrophosphate
PPTase	phosphopantetheine transferase
proPO	prophenoloxidase
RXP	rhabdopeptide/xenortide-like peptide
SAH	*S*-adenosyl homocysteine
SAM	*S*-adenosyl methionine
SPPS	solid-phase peptide synthesis
T domain	thiolation domain
TE domain	thioesterase domain
XPPM	*Xenorhabdus* and *Photorhabdus* production medium
XU	exchange unit
XUC	exchange unit condensation domain

Summary

The application of natural products (NPs) as drugs and lead compounds has greatly improved human health over the past few decades. Despite their success, we still need to find new NPs that can be used as drugs to combat increasing drug resistance via new modes of action and to develop safer treatments with less side effects.

Entomopathogenic bacteria of *Xenorhabdus* and *Photorhabdus* that live in mutualistic symbiosis with nematodes are considered as promising producers of NPs, since more than 6.5% of their genomes are assigned to biosynthetic gene clusters (BGCs) responsible for production of secondary metabolites. The investigation on NPs from *Xenorhabdus* and *Photorhabdus* can not only provide new compounds for drug discovery but also help to understand the biochemical basis involved in mutualistic and pathogenic symbiosis of bacteria, nematode host and insect prey.

Nonribosomal peptides (NRPs) are a large class of NPs that are mainly found in bacteria and fungi. They are biosynthesized by nonribosomal peptide synthetases (NRPSs) and display diverse functions, representing more than 20 clinically used drugs. Although a large number of NRPs have been identified in *Xenorhabdus* and *Photorhabdus*, the advanced genome sequencing and bioinformatic analysis indicate that these bacteria still have many unknown NRPS-encoding gene clusters for NRP production that are worth to explore. Therefore, this thesis focuses on the discovery, biosynthesis, structure identification, and biological functions of new NRPs from *Xenorhabdus* and *Photorhabdus*.

The first publication describes the isolation and structure elucidation of seven new rhabdopeptide/xenortide-like peptides (RXPs) from *X. innexi*, incorporating putrescine or ammonia as the C-terminal amines. Bioactivity testing of these RXPs revealed potent antiprotozoal activity against the causative agents of sleeping sickness (*Trypanosoma brucei rhodesiense*) and malaria (*Plasmodium falciparum*), making them the most active RXP derivatives known to date. Biosynthetically, the initial NRPS module InxA might act iteratively with a flexible methyltransferase activity to catalyze the incorporation of the first five or six *N*-methylvaline/valine to these peptides.

Summary

The second publication focuses on the structure elucidation of seven unusual methionine-containing RXPs that were found as minor products in *E. coli* carrying the BGC *kj12ABC* from *Xenorhabdus* KJ12.1. To confirm the proposed structures from detailed HPLC-MS analysis, a solid-phase peptide synthesis (SPPS) method was developed for the synthesis of these partially methylated RXPs. These RXPs also exhibited good effects against *T. brucei rhodesiense* and *P. falciparum*, suggesting RXPs might play a role in protecting insect cadaver from soil-living protozoa to support the symbiosis with nematodes.

The third publication presents the identification of a new peptide library, named photohexapeptide library, which occurred after the biosynthetic gene *phpS* was activated in *P. asymbiotica* PB68.1 via promoter exchange. The chemical diversity of the photohexapeptides results from unusual promiscuous specificity of five out of six adenylation (A) domains being an excellent example of how to create compound libraries in nature. Furthermore, photohexapeptides enrich the family of the rare linear D-/L-peptide NPs.

The fourth publication concentrates on the structure elucidation of a new cyclohexapeptide, termed photoditritide, which was produced by *P. temperata* Meg1 after the biosynthetic gene *pdtS* was activated via promoter exchange. Photoditritide so far is the only example of a peptide from entomopathogenic bacteria that contains the uncommon amino acid homoarginine. The potent antimicrobial activity of photoditritide against *Micrococcus luteus* implies that photoditritide can protect the insect cadaver from food competitor bacteria in the complex life cycle of nematode and bacteria.

The last publication reports a new family of cyclic lipopeptides (CLPs), named phototemtides, which were obtained after the BGC *pttABC* from *P. temperata* Meg1 was heterologously expressed in *E. coli*. The gene *pttA* encodes an MbtH protein that was required for the biosynthesis of phototemtides in *E. coli*. To determine the absolute configurations of the hydroxy fatty acids, a total synthesis of the major compound phototemtide A was performed. Although the antimalarial activity of phototemtide A is only weak, it might be a starting point towards a selective *P. falciparum* compound, as it shows no activity against any other tested organisms.

Zusammenfassung

Der Einsatz von Naturstoffen als klinisch genutzte Wirkstoffe oder Leitstrukturen hat in den letzten Jahrzehnten das Leben der Menschen in vielerlei Hinsicht beeinflusst und verbessert. Trotz bereits erzielter Erfolge und gleichzeitig durch sie „ermutigt", müssen aber weitere Naturstoffe mit neuen Wirkmechanismen zur Bekämpfung verstärkt auftretender Resistenzen oder Wirkstoffe mit verbesserter Wirksamkeit bzw. Verträglichkeit gefunden werden.

Entomopathogene Bakterien der Gattung *Xenorhabdus* und *Photorhabdus* leben in mutualistischer Symbiose mit Nematoden der Gattung *Steinernema* bzw. *Heterorhabditis*. Mehr als 6,5 % ihres Genoms codieren für Biosynthesegencluster (BGC), die für die Biosynthese von Naturstoffen, sogenannter Sekundärmetabolite, verantwortlich sind. Die Analyse der von *Xenorhabdus* und *Photorhabdus* gebildeten Naturstoffe führt einerseits zur Entdeckung einer Vielzahl biologisch aktiver Verbindungen und ermöglicht andererseits ein besseres Verständnis der biochemischen Abläufe in der Bakterium-Nematoden-Symbiose und der Entomopathogenität dieses Komplexes.

Nematoden der Gattungen *Steinernema* bzw. *Heterorhabditis* haben einen komplexen Lebenszyklus: Bodenlebende, infektiöse Dauerlarven der Fadenwürmer beherbergen in ihrem Darm Bakterien der Gattung *Xenorhabdus* bzw. *Photorhabdus*. Sie dringen über den Darmkanal, das Tracheensystem oder durch Durchbrechen der Kutikula in das Insekt ein und geben ihre Bakterien in das Hemocoel ab. Aus dem infektiösen Stadium (IJ) entwickeln sich adulte, eierlegende Fadenwürmer (J4-Stadium). Die aus den Eiern geschlüpften Fadenwürmer durchlaufen den Zyklus J1-J4 etwa zwei- bis dreimal. Sind schließlich die Nährstoffe des Insektenkadavers verbraucht, entwickeln sich neue infektiöse Dauerformen, die von Bakterien besiedelt werden und zu Tausenden den Kadaver auf der Suche nach neuen Opfern verlassen.

Die symbiotischen Bakterien tragen wesentlich zur Reproduktion der Nematoden bei. Ins Hemocoel ausgeschieden replizieren sie rasch und es kommt zu einer Sepsis des Insekts. Zur Verteidigung der Nahrungsquelle und damit auch zur Sicherstellung der eigenen und der Vermehrung und Entwicklung der Nematoden produzieren sie Naturstoffe, die das Immunsystem des Insekts überwinden und Toxine, die das Insekt innerhalb von 48 Stunden töten. Der Kadaver wird von den Bakterien durch

VII

die Produktion bioaktiver Verbindungen vor Fressfeinden wie anderen Bakterien, Pilzen und Protozoen geschützt. Gleichzeitig werden auch Substanzen produziert, die die Nematodenentwicklung unterstützen. Die genauere Erforschung der in diesem komplexen Lebenszyklus wirksamen Naturstoffe führt nicht nur zur Entdeckung neuer Wirkstoffe, sondern ermöglicht auch einen tieferen Einblick und Aufklärung der biochemisch ökologischen Aspekte dieser mutualistischen und pathogenen Symbiose zwischen Bakterium, Nematode und Insekt. In den letzten zehn Jahren wurden über 40 Naturstoffklassen in *Xenorhabdus* und *Photorhabdus* identifiziert, für deren Biosynthese nicht-ribosomale Peptidsynthetasen (NRPS), Polyketidsynthasen (PKS) oder entsprechende NRPS-PKS-Hybride verantwortlich sind.

NRPS sind große multimodulare Enzymkomplexe. Entsprechende Module bewirken die Aktivierung und Verknüpfung einzelner Aminosäuren zu einer Peptidkette. Jedes Modul enthält multiple Domänen, die verschiedene enzymatische Aktivitäten katalysieren: das Initiationsmodul einer NRPS besteht mindestens aus einer Adenylierungsdomäne (A) und einer Thiolierungsdomäne (T). Elongationsmodule enthalten eine zusätzliche Kondensationsdomäne (C). Ein Elongationsmodul besteht aus C-, A- und T-Domänen. Epimerisierungs- (E), Methyltransferase- (MT) und Formyltransferase- (FT) Domänen können als zusätzliche Domänen auftreten. Die von NPRS synthetisierten nicht-ribosomalen Peptide (NRP) bilden eine große Klasse der Naturstoffe, hauptsächlich synthetisiert von Bakterien und filamentösen Pilzen. Diese Peptide sind in ihrer Struktur vielgestaltig. Neben den 20 proteinogenen Aminosäuren sind auch bis zu 100 nicht-proteinogene Bausteine zu finden. Weitere Modifizierungen wie Acylierung, Glycosylierung oder Heterozyklisierung kann durch entsprechende Domänen der NRPS die chemische Struktur des NRP bestimmen. Je nach Aktivität der Terminations-Domäne einer NRPS kann die finale Struktur der Peptide linear oder zyklisch oder von zyklisch-verzweigter Struktur sein, wie z. B. makrozyklische Lactame oder Lactone. Bis zu 30 NRP sind derzeit in klinischer Anwendung und werden als Antibiotika, Fungizide, Krebsmedikamente oder Immunsuppressiva eingesetzt.

Auch in *Photorhabdus* und *Xenorhabdus* bilden NRPS kodierende Biosynthesegencluster (BGC) die bisher größte Klasse der Naturstoffe. Diese strukturell und funktional vielfältigen Peptide sind entscheidend für den komplexen

Lebenszyklus von Bakterien und Nematoden. Durch Genomsequenzierungen der Bakterien und anschließender bioinformatischer Analyse gelingt das Auffinden zahlreicher bisher unbekannter Gencluster für die Biosynthese vielversprechender neuer NRP. Die Isolierung und Aufklärung korrespondierender Naturstoffe kann jedoch erschwert sein, wenn das Gencluster unter Laborbedingungen nur schwach exprimiert oder still ist, weil aktivierende Signale des natürlichen Ökosystems fehlen. Der Schlüssel zur Überwindung dieses Problems ist die gezielte Aktivierung der entsprechenden Gencluster. Hierzu wurden verschiedene Strategien wie der Austausch des natürlichen gegen einen induzierbaren Promotor und heterologe Expression angewendet. Nach der erfolgreichen Aktivierung der Naturstoffbiosynthese ist die Aufklärung ihrer chemischen Struktur über verschiedene analytische Methoden notwendig.

Diese Arbeit beschreibt die Entdeckung und Biosynthese neuer NRP aus *Xenorhabdus* und *Photorhabdus* durch biologische und chemische Methoden. Beginnend mit der Genomsequenzierung und bioinformatischer Analyse, erfolgte die gezielte Aktivierung der BGC durch Promotoraustausch und heterologe Expression. Die Strukturaufklärung neu entdeckter Naturstoffe gelang durch Fütterungsexperimente, HPLC-MS- und NMR Analyse, sowie die „advanced Marfey's Methode" und chemische Synthese.

Rhabdopeptid-/Xenortid-ähnliche Peptide (RXP) sind eine strukturell diverse Naturstoffklasse, die bisher nur in *Xenorhabdus* und *Photorhabdus* identifiziert wurde. Sie bestehen aus zwei bis acht Aminosäuren Valin, Phenylalanin und Leucin, wobei Valin am häufigsten eingebaut wird. Oft zeigen sie eine N-methylierung und Phenylalanin oder Tryptamin bilden das C-terminale Amin. Der RXP-BGC codiert für zwei bis vier Proteine, die für einzelne Module der NRPS kodieren. Die hohe chemische Diversität der RXP resultiert aus dem iterativen und flexiblen Gebrauch dieser Modul-Untereinheiten während der Biosynthese. Bei genauerer HPLC-MS-Analyse der *Xenorhabdus*- und *Photorhabdus*-Kulturextrakte wurden sieben neue RXP aus *X. innexi* identifiziert, die in ihrer chemischen Struktur typischerweise Putrescin oder Ammonium anstelle der bereits bekannten terminalen Amine Phenylethylamin und Tryptamin als C-terminales Amin aufweisen. Diese neuen RXP sind Thema der ersten Publikation. Im Vergleich zu bisher bekannten Derivaten dieser Naturstofffamilie zeigten diese RXP die beste antiprotozoische Aktivität gegen

Erreger der Schlafkrankheit (*Trypanosoma brucei rhodesiense*) und Malaria (*Plasmodium falciparum*). Eine Besonderheit der Biosynthese dieser RXP liegt in der iterativen Funktion des initialen NRPS-Moduls InxA, das neben einer flexiblen Methyltransferaseaktivität auch den Einbau von fünf bis sechs Molekülen *N*-Methyl-Valin/Valin katalysiert.

Als Nebenprodukte bei der heterologen Expression des BGC *kj12ABC* aus *Xenorhabdus* KJ12.1 in *E. coli* wurden sieben weitere ungewöhnliche Methionin-enthaltenden RXP entdeckt. Die zweite Publikation beschreibt die Aufklärung ihrer chemischen Struktur über detaillierte HPLC-MS-Analyse. Außerdem wurde für die Synthese dieser speziellen, partiell methylierten RXP eine einfache Festphasen-Peptidsynthese entwickelt. Im Aktivitätstest zeigten diese Peptide eine gute Wirkung gegen *Trypanosoma brucei rhodesiense* und *Plasmodium falciparum*. Diese Wirksamkeit könnte zur Verteidigung des Insektenkadavers gegen bodenlebende Protozoen und somit zum Schutz der Nematodenentwicklung und folglich der Symbiose beitragen.

Durch Genomanalyse von *P. asymbiotica* PB68.1 wurde die NRPS PhpS-identifiziert. Sie besteht aus sechs Modulen. Es wurde folglich die Synthese eines Hexapeptids erwartet, das jedoch in der Analyse der Extrakte nicht identifiziert werden konnte. Erst durch einen Promotoraustausch vor dem Gen *phpS* wurde die Biosynthese der Photohexapeptide aktiviert, die in der dritten Publikation beschrieben sind. Die chemische Diversität dieser Photohexapeptide resultiert aus einer außergewöhnlich promiskuitiven Spezifität der Adenylierungs-Domänen (A) der NRPS und stellt ein sehr gutes Beispiel für die natürliche Entstehung von Peptid-Bibliotheken dar (insgesamt können 32 verschiedene Derivate entstehen). Weiterhin bereichern die Photohexapeptide die Familie der, unter den Naturstoffen eher selten vertretenen, linearen D-/L-Peptide.

Im Genome von *P. temperata* MEG1 codiert das *pdtS*-Gen für eine bisher unbekannte NRPS bestehend aus sechs Modulen. Die Aktivierung des BGC gelang durch Promotoraustausch und die Entdeckung des Photoditritid ist Thema der vierten Publikation. Photoditritid ist das erste beschriebene Peptid eines entomopathogenen Bakteriums, das als Bausteine zwei Reste der ungewöhnlichen Aminosäure Homoarginin enthält. Photoditritid zeigt Aktivität gegen *Micrococcus luteus*, was

möglicherweise auch auf eine antimikrobielle Wirksamkeit auch im komplexen Lebenszyklus von Bakterium und Nematode schließen lässt.

Bioinformatische Analysen von *P. temperata* Meg1 führten zur Entdeckung von zwei weiteren NRPS PttBC. PttBC besteht aus insgesamt fünf Modulen mit einer C-starter Domäne (C_{start}) im Initiationsmodul von PttB. Ein weiteres Gen *pttA*, 0.8 kbp stromaufwärts kodiert für ein MbtH-Protein, dessen Aktivität für die Biosynthese der Phototemtide essentiell ist. Die Identifizierung eines entsprechenden Lipohexapeptids aus Extrakten von *P. temperata* Meg1 sowie entsprechenden Promotoraustauschmutanten (pCEP-pttA) war erfolglos. Erst durch heterologe Expression von *pttABC* gelang in *E. coli* die *pttA*-kodierte MbtH-abhängige Expression einer neuen Familie zyklischer Lipodepsipeptide (CLP). In der letzten Publikation sind die Entdeckung, Expression, Strukturaufklärung und Totalsynthese dieser neuen Phototemtide beschrieben, die somit ein exzellentes Beispiel für die Anwendung aller Methoden zur Entdeckung und Aufklärung eines Naturstoffs darstellen. Obwohl Phototemtide A nur eine schwache Aktivität gegen *Plasmodium falciparum* aufweist, könnte dies die Ausgangsverbindung für die Entwicklung eines spezifischen Wirkstoffs gegen diesen Erreger sein, da es in den getesteten Organismen keine weitere Wirkung zeigt.

1 Introduction

1.1 *Xenorhabdus* and *Photorhabdus*

Xenorhabdus and *Photorhabdus* are rod-shaped, motile, facultative anaerobic, and non-sporulating Gram-negative Gammaproteobacteria, belonging to the family of Enterobacteriaceae.[1,2] They live in a unique mutualistic symbiosis with the entomopathogenic nematodes of the genera *Steinernema* and *Heterorhabditis*, respectively.[3] The nematode−bacteria complex is highly pathogenic for a broad range of insects and has been applied to various insect pest control in organic farming owing to the advantage of this method in their timely and precise application in comparison to the synthetic insecticides.[4,5] Additionally, the complex is essential for the life cycle and reproduction of both nematodes and bacteria.[6,7] Therefore, the nematode−bacteria complex is well known as a model system for the investigation of mutualistic and pathogenic symbiosis.[8–10]

The nematodes of *Steinernema* and *Heterorhabditis* share a quite similar life cycle.[6] The basic principle of the nematode−bacteria life cycle, as shown in Figure 1, has been known for a long time.[11] In nature, the bacteria are carried as symbionts in the intestine of the free-living infective juvenile (IJ) stage of the nematodes.[6] The IJ, a modified third stage juvenile (J3), represents a non-feeding, non-developing, and non-reproductive stable "dauer" form of the nematode that can survive for long periods of time in the soil before entering an insect prey.[6,12] The symbionts are able to persist for many weeks within the IJs until becoming pathogenic when released within the insect hemocoel.[6] Once the IJs enter the prey through natural openings or insect cuticles, the bacteria are released by the IJs that now undergo a complex life cycle involving four different juvenile stages (J1−J4), starting from the IJ stage and subsequently developing into the reproducing egg-laying adults via the J4 stage.[6,12] This life cycle is usually repeated 2−3 times.[6] When the food source is depleted, a new generation of IJs receive the recolonization with *Xenorhabdus* or *Photorhabdus* before they exit the cadaver and start their search for a new insect prey.[6]
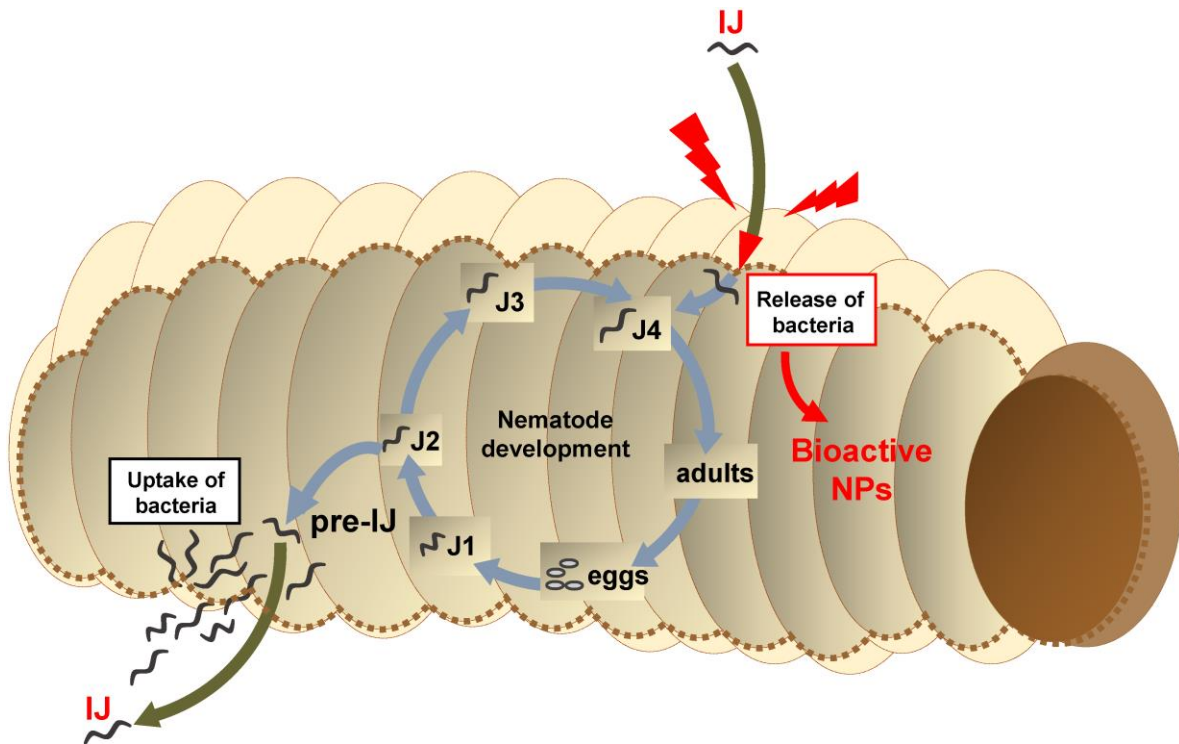
Figure 1. Schematic overview of the life cycle of the entomopathogenic nematode−bacteria complex. The infective juvenile (IJ) of the nematode carries the bacteria and invades the insect. After entering the insect, the bacteria are released by the IJ. The bacteria replicate and kill the insect. The IJ recovers into reproducing egg-laying adults via J4 juveniles. From their eggs, the cycle starts via J1−J4 and is repeated about 2−3 times. When the food source is depleted, the pre-IJ develops and is recolonized by the bacteria. They finally exit the exploited insect cadaver (grey) as new IJ generation in search of a new prey. Adapted from[12].

The symbiotic bacteria of *Xenorhabdus* and *Photorhabdus* are vital for the life cycle and reproduction of the entomopathogenic nematodes.[6,7] Released into the insect hemocoel, the bacteria replicate rapidly. To ensure the survival of bacteria and their nematode host, the bacteria start to produce a large number of toxins and secondary metabolites to conquer the immune system of the insect and also kill the prey within 48 h.[13–15] After the death of the insect, the bacteria produce compounds to protect the prey cadaver from food competitors, like bacteria, fungi, protozoa and natural predators.[15,16] In the meantime, the bacteria secret lytic enzymes to break down the insect cadaver which provides nutrient as well as an ideal environment for the growth and development of bacteria and nematodes.[10,17]

1.2 Natural Products from *Xenorhabdus* and *Photorhabdus*

The entomopathogenic lifestyle of the nematodes and bacteria makes *Xenorhabdus* and *Photorhabdus* a promising source of new potentially bioactive natural products (NPs) that are desperately needed for drug discovery and development.[14,15] Moreover, Genome-sequencing projects have revealed that more than 6.5% of the overall genome sequence can be assigned to biosynthetic gene clusters (BGCs) of secondary metabolites.[18] Although in nature, *Xenorhabdus* and *Photorhabdus* have only been isolated from infected insects or their nematode hosts,[19] it is of great fortune for NP scientists and microbiologists that the bacteria can also be cultivated without their host in the laboratory. All these advantages have attracted NP scientists' attention for exploring new bioactive NPs from *Xenorhabdus* and *Photorhabdus*. During the past few decades, more than 40 classes of NPs have been reported from several members of *Xenorhabdus* and *Photorhabdus* (Table 1). Some of them have been identified to play important roles in the complex life cycle of nematodes and bacteria. According to the life cycle, the functions of several selected NPs (Figure 2) will be described in the following paragraphs.
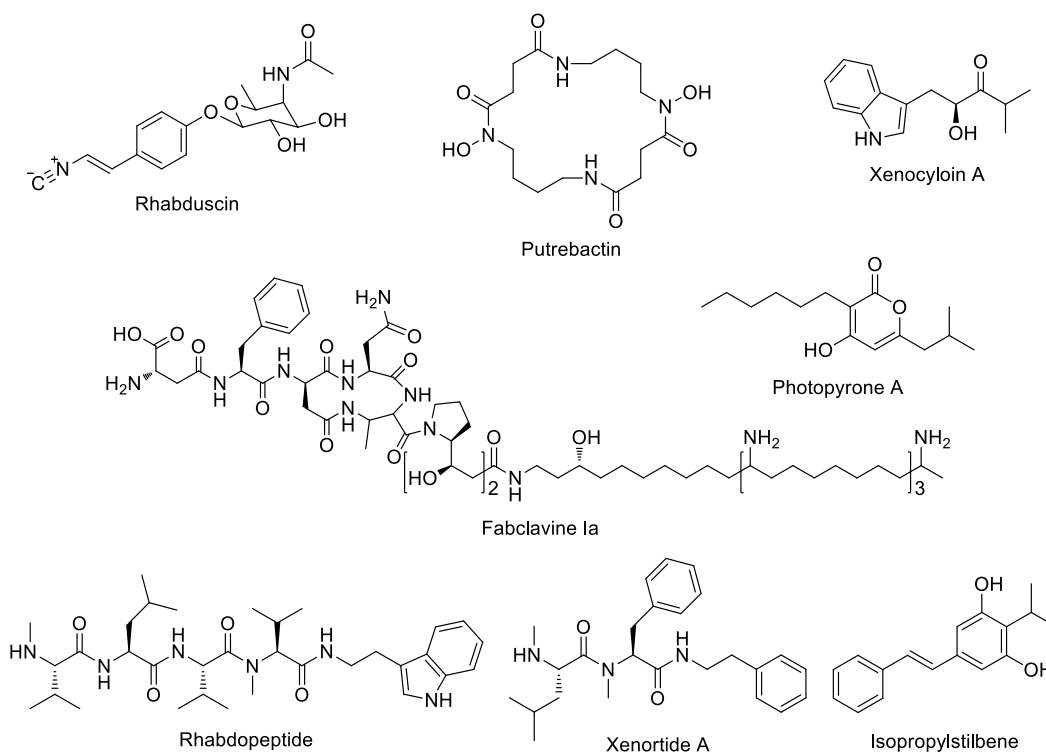
Figure 2. Structures of selected natural products from *Xenorhabdus* and *Photorhabdus*.

After entering the insect, the nematode−bacteria pair must first overcome the extensive insect immune system, including cellular immune response and humoral immune response. The cellular immune response comprises phagocytosis, nodulation and encapsulation and the humoral response involves the expression of antimicrobial peptides as well as the proteolytic cascade of prophenoloxidase (proPO) that results in formation of melanized nodules and toxic reactive compounds.[20] Bacterial peptidoglycan fragments and lipopolysaccharide are bound by recognition molecules in the insect hemolymph activating the proteolytic cascade of the proPO.[20,21] These proteins mediate an immune response that finally activates the phenoloxidase (PO) that is responsible for melanization and encapsulation of the pathogen.[22] *Xenorhabdus* and *Photorhabdus* produce a number of compounds specifically inhibiting several of the insect immune pathways. Rhabduscin, a tyrosine derivative functionalized with a 4-acetylamino-4-deoxyglycosyl and an isocyano group, is produced in almost all *Xenorhabdus* and *Photorhabdus* strains.[23] It was found to decorate the bacterial cell surface and to inhibit the PO activity, thereby protecting the bacterial cell.[24] Besides, a hydroxamate siderophore, putrebactin, was identified from *X. budapestensis* recently.[25] Hydroxamate siderophores are known to be potent inhibitors of histone deacetylases (HDACs).[26] HDAC complexes are a key activation mechanism that transcriptionally regulate the production of antimicrobial peptides in eukaryotes like insects.[27] Repression of HDACs would therefore significantly hamper the ability of the insect to mount an effective antimicrobial peptide-based immune response.[28]

To kill the insect prey, *Xenorhabdus* and *Photorhabdus* produce several bioactive NPs involved in insect pathogenicity. The indole derivatives xenocyloins, originally isolated from *X. bovienii*,[29,30] are also identified in many other *Xenorhabdus* strains.[23] In *X. bovienii*, they were produced mainly by the more virulent primary variant and showed activity against insect hemocytes, hence contributing to the overall virulence of *X. bovienii* against insects.[30] *P. luminescens* and *P. temperata* produce a family of α-pyrone moiety containing small molecules, the photopyrones, involved in a new cell−cell communication system.[31] This system consists of endogenously produced photopyrones as signaling molecules and PluR, a LuxR-like receptor.[31,32] The binding of photopyrones to PluR leads to the expression of the *Photorhabdus* clumping factor operon (*pcfABCDEF*) that causes clumping of cells and results in insect toxicity.[31,32]

Following insect killing, *Xenorhabdus* and *Photorhabdus* produce NPs to protect the nutrient-rich insect cadaver from food competitors, such as bacteria, fungi, and protozoa, to allow the nematodes to continue their life cycle. Fabclavines, an unusual class of peptide−polyketide−polyamine hybrids, were isolated and characterized from *X. budapestensis* and *X. szentirmaii*.[33] As they show broad-spectrum activity against bacteria, fungi, as well as other eukaryotic cells, they might act as "protection factors" against all kinds of food competitors in the soil or the gut of the insect prey.[33] Rhabdopeptide/xenortide-like peptides (RXPs), a unique class of nonribosomally derived linear peptides, are widespread and structurally diverse in *Xenorhabdus* and *Photorhabdus*.[34] They are prominently produced during insect infection and are most abundant after insect death.[35] The activity of RXPs against hemocytes confirmed the involvement of RXPs in pathogenesis against insects.[35] Furthermore, RXPs are active against different protozoa.[34,36] Therefore, the biological diversity of RXPs might serve as a multipotent arsenal that facilitates *Xenorhabdus* and *Photorhabdus* strains to conquer different insect prey and food competitors.[15]

Xenorhabdus and *Photorhabdus* also produce NPs to support nematode development. Isopropylstilbene has been identified in all *Photorhabdus* strains investigated so far.[37] It is a cross-kingdom signal required for the recovery of the nematode IJ stage into other developmental stages.[37] Besides, isopropylstilbene has antibacterial and antifungal activities, indicating its role in protecting the insect cadaver from microbial saprophytes living in the soil.[38,39] Furthermore, isopropylstilbene has a function as a PO inhibitor to counter host immune reactions.[40] Thereby, it might play an essential role in both mutualism and pathogenicity and is thus a central building block in the interplay between bacteria, nematodes, and insect larvae.[37]

Biosynthetically, most NPs from *Xenorhabdus* and *Photorhabdus* are derived from nonribosomal peptide synthetases (NRPSs), polyketide synthases (PKSs), or hybrids thereof. As summarized in Table 1, the largest class of NP BGCs so far identified in *Xenorhabdus* and *Photorhabdus* are NRPSs that perform nonribosomal peptide (NRP) synthesis. These structurally and functionally diverse peptides are crucial to the complex life cycle of nematode and bacteria. The advanced genome sequencing combined with genome mining indicates that the bacteria of *Xenorhabdus* and

Photorhabdus still have a large number of unknown NRPS-encoding gene clusters for peptide production that need to be explored.

Table 1. Summary of natural product classes identified from the original strains of *Xenorhabdus* or *Photorhabdus* and their biosynthetic gene types

Natural products	Source strain	Type	Refs
Ambactin	*X. miraniensis*	NRPS	[41]
Anthraquinones	*P. luminescens*	PKS	[39,42]
Benzylideneacetone	*X. nematophila*	unknown	[43]
Bicornutins	*X. budapestensis*	NRPS	[44]
Chaiyaphumines	*Xenorhabdus* sp. PB61.4	(most likely) NRPS	[45]
Dihydrophenylalanines	*P. luminescens*	other	[46]
Fabclavines	*X. budapestensis*, *X. szentirmaii*	NRPS−PKS	[33]
GameXPeptides	*P. luminescens*	NRPS	[47]
Glidobactins	*P. asymbiotica*	NRPS−PKS	[48]
HCTA-peptides	*X. miraniensis*, *Xenorhabdus* sp. XPB 63.3	NRPS	[44]
Hydroxamate siderophores	*X. szentirmaii*, *X. budapestensis* *Photorhabdus* KK1.3	siderophore	[25]
Indigoidine	*P. luminescens*	resorcinol	[49]
Isopropylstilbene	*P. luminescens*	resorcinol	[37]
Kolossins	*P. luminescens*	NRPS	[50]
Lumiquinone	*P. luminescens*	PKS	[51]
Lumizinones	*P. luminescens*	unknown	[52]
Mevalagmapeptides	*P. luminescens*	NRPS	[53]
N-acetyl dipeptides	*P. luminescens*	unknown	[52]
Odilorhabdins	*X. nematophila*	NRPS	[54]
Oxindole	*X. nematophila*	unknown	[55]
PAX-peptides	*X. nematophila*	NRPS	[56,57]
Pepteridines	*P. luminescens*	NRPS−pteridine	[58]

Table 1 (*Contd.*)

| Natural products	Source strain	Type	Refs
Phenazines	*X. szentirmaii*	phenazine	[59]
Photolose	*P. luminescens*	NRPS−carbohydrate	[60]
Photopyrones	*P. luminescens*, *P. temperata*	other	[31,32]
Phurealipids	*P. luminescens*	other	[61]
Pyrrolizixenamides	*X. stockiae*	NRPS	[62]
Rhabdopeptides/Xenortides	*X. nematophila*, *X. cabanillasii*	NRPS	[35,36,63]
Rhabdoplanins	*X. bovienii*	unknown	[64]
Rhabduscin	*P. luminescens*, *X. nematophila*	other	[24]
RILXIRRpeptides	*X. indica*	NRPS	[44]
Simple amides	*X. nematophilus,* *X. doucetiae*	other	[65,66]
Szentiamides	*X. szentirmaii*	NRPS	[67]
Taxlllaids	*X. indica*	NRPS	[68]
Tilivallines	*X. eapokensis*	NRPS	[69]
Xefoampeptides	*X. beddingii*	NRPS	[23]
Xenematides	*X. nematophila*	NRPS	[63,70]
Xenoamicins	*X. doucetiae*, *X. mauleonii*	NRPS	[71]
Xenobactin	*Xenorhabdus* sp. PB30.3	unknown	[72]
Xenocoumacins	*X. nematophila*	NRPS−PKS	[73,74]
Xenocyloins	*X. bovienii*	other	[29,30]
Xenofuranones	*X. szentirmaii*	unknown	[75]
Xenolindicins	*X. indica*	NRPS	[41]
Xenorhabdins	*X. doucetiae*	NRPS	[53,76]
Xenotetrapeptide	*X. nematophila*	NRPS	[77]
Xentrivalpeptides	*Xenorhabdus* sp. 85816	unknown	[78]

1.3 Nonribosomal Peptide Synthetases

Nature has evolved several solutions for the directed polymerization of amino acids into peptides under the formation of amide bonds. Ribosomal synthesis is a fundamental process for the synthesis of peptides. NRP synthesis performed by dedicated NRPSs is another major alternative biosynthesis pathway that is mainly found in bacteria and fungi.[79] NRPs are a large class of structurally diverse NPs, which contain not only the 20 proteinogenic amino acids but also hundreds of different building blocks in comparison with peptides and proteins produced by ribosomal synthesis.[80] Further modifications, such as methylation, acylation, glycosylation, heterocyclic ring formation, and epimerization, can be conducted by the corresponding enzymes that are commonly embedded in NRPSs.[81] Moreover, the released peptides can be linear, cyclic, or branched-cyclic leading to macrocyclic lactams or lactones.[81] The enormous structural diversity confers the bioactive diversity of NRPs for many biological targets. The usefulness of nearly 30 clinically used drugs from NRPs is evident.[79] Their predominant use is as antibacterials (vancomycin), followed by antitumor drugs (bleomycin), antifungals (enniatin), animal food additives (bacitracin), immunosuppressants (cyclosporine), or in obstetrics (ergometrine), and pain treatment (ergotamine).[79]

The study of NRP synthesis started in the early 1960s when researchers investigated how certain cyclic peptides containing D-amino acids were synthesized by *Bacillus* species.[82] In the past few decades, biochemical and structural biology studies have gained mechanistic insights into NRPSs, the highly complex enzymatic assembly line of NRPs.[79] Although peptides synthesized by NRPSs are highly diverse in structure, most of them share a common mode of synthesis, the multienzyme thiotemplate mechanism.[80,83] NRPSs are large multienzyme machineries consisting of modules, in which each module performs the activation and coupling of a single amino acid to the growing peptide chain. Each module comprises multiple domains that catalyze different enzymatic activities. Initial modules consist of at least one adenylation (A) domain and one thiolation (T) domain; elongation modules contain in addition a condensation (C) domain. Therefore, the domain sequence for a minimal initial module is A-T and for a minimal elongation module C-A-T.[84] The C, A and T domains in a minimal elongation module are called core domains. The essential enzymatic

activities for one complete elongation in NRP synthesis are illustrated in Figure 3 and described in more detail in the following sections.
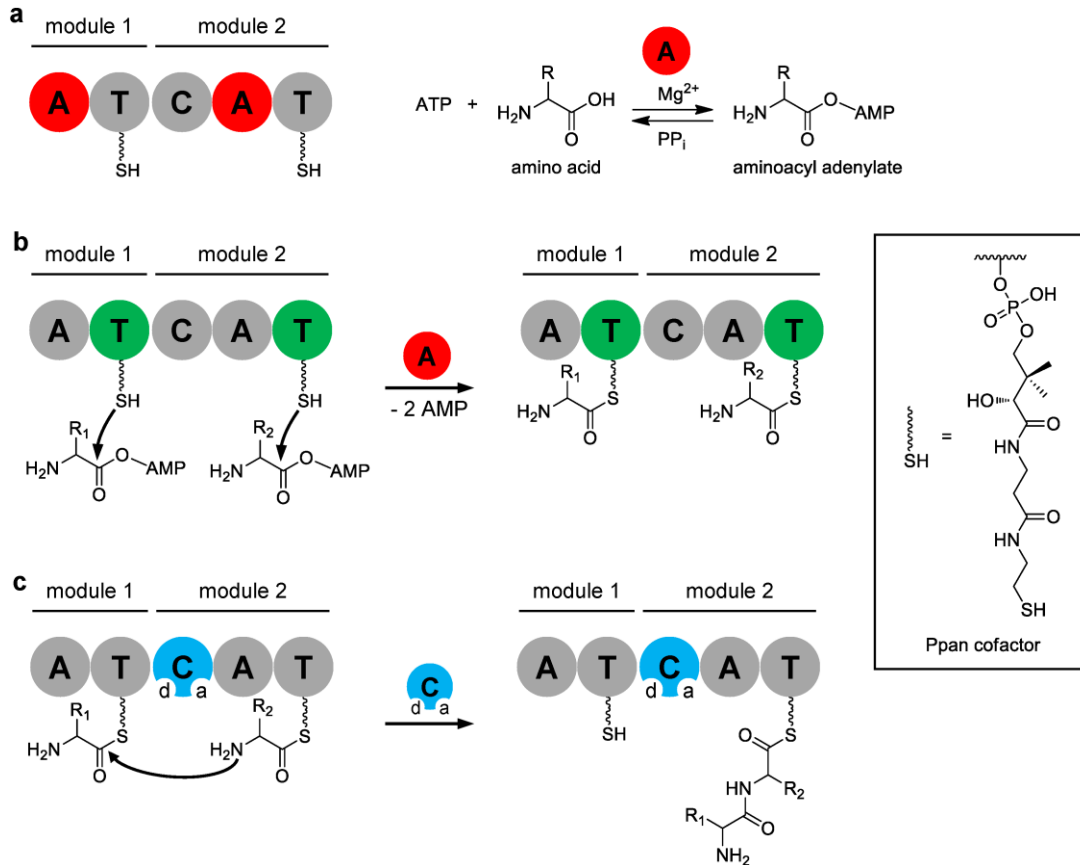
Figure 3. Essential enzymatic activities of NRPS core domains. Domains in action are highlighted. (a) Amino acid recognition and activation with ATP by the A domain; (b) Covalent attachment of the activated aminoacyl adenylate onto the free thiol group of the T domain-bound Ppan cofactor; (c) Peptide elongation by the C domain, which catalyzes an attack of the nucleophilic amine of the acceptor substrate onto the electrophilic thioester of the donor substrate. The acceptor site of the C domain is indicated by a, the donor site by d. Domains: A: adenylation, T: thiolation, C: condensation; ATP: adenosine triphosphate; AMP: adenosine monophosphate; PP$_i$: pyrophosphate; Ppan: 4'-phosphopantetheinyl. Adapted from[80].

1.3.1 Core Domains

Adenylation domains

The adenylation (A) domain (ca. 60 kDa) belongs to the ANL (Acyl-CoA synthetases, NRPS adenylation domains, and Luciferase enzymes) superfamily of adenylating

enzymes, which catalyze two partial reactions: the initial adenylation of a carboxylate to form an acyl-AMP intermediate, followed by a second partial reaction, the formation of a thioester.[85] Likewise, in NRP synthesis, the selection and activation of amino acid substrates as the initial step are fulfilled by the A domain in a two-step chemical reaction.[79,80] First, the A domain catalyzes the formation of an aminoacyl adenylate intermediate by $Mg^{2+}-ATP$ consumption and the release of pyrophosphate (PP_i) (Figure 3a).[80] Second, the activated aminoacyl-AMP is converted into a covalently bound thioester by a nucleophilic attack of the free thiol group of the T domain-bound 4'-phosphopantetheinyl (Ppan) cofactor (Figure 3b).[80] As the aminoacyl-AMP is prone to nonproductive hydrolysis, the A domain protects the high-energy intermediate from bulk water.[79] Thus, A domains can be described as a gatekeeper in NRP synthesis.

A domains consist of an approximately 50 kDa N-terminal core domain (A_{core}) and an approximately 10 kDa C-terminal subdomain (A_{sub}), which are flexibly linked by a hinge region of about five residues.[79] The specific binding of an amino acid and $Mg^{2+}-ATP$ occurs within the A_{core} domain close to the $A_{core}-A_{sub}$ interface.[79] Although all A domains only show 30−60% sequence identity, they share 10 highly conserved "core motifs" (A1−A10).[86] These consensus motifs play important structural and functional roles and have led to the establishment of a specificity-conferring code for A domains.[87–89] Bioinformatic algorithms have been developed for the prediction of potential A domain substrates and, thus, the structures of NRPS-derived peptides by genome mining approaches.[90,91] However, it should be noted that the accuracy of the predictions varies depending on the amino acid.[92] Furthermore, relaxed substrate specificity of A domains have been observed in several NRPS systems, which might be a strategy used by NRP-producing organisms to increase NP diversity within a single NRPS.[70,93,94]

Various NRP BGCs contain small MbtH-like protein encoding genes. These small proteins are usually encoded near the NRPS genes or gene clusters. Genetic inactivation experiments have indicated that MbtH-like proteins are often required for the efficient production of NRPs.[95,96] Strikingly, MbtH-like protein paralogues from other NRPS gene clusters within the genome can partially complement each other such that NRP production is not compromised.[95] Several biochemical studies have shown that MbtH-like proteins interact with A domains where they stimulate

adenylation reactions.[97–99] To date, the mechanism of MbtH-like proteins as well as their role in NRP synthesis in general still remains vague, but their importance in future biotechnological approaches is already evident.[79]

Thiolation domains

The thiolation (T) domain (ca. 10 kDa), also referred to as peptidyl carrier protein (PCP) domain, is the smallest core domain in NRPS. The T domain is located C-terminal to the A domain and is the only NRPS domain without an autonomous catalytic activity.[80] Conformational flexibility is an inherent trait as well as an essential requirement for the communication and choreography of NRPS domains.[79] The T domain can be regarded as the paradigm for this functional flexibility.[79] The active *holo* form of the T domain with its Ppan cofactor can be described as the flexible robot arm of the NRPS assembly line that covalently sequesters and transfers the amino acyl-/peptidyl-thioester intermediates to all the catalytic centers that are required for modification, condensation, or release.[79,100]

Structural studies of the T domains revealed that the T domain is a distorted four-helix bundle with an extended loop between the first two helices.[101,102] The highly conserved serine residue (GxxS core motif), the site of cofactor binding, is located at the interface between this loop and the second helix.[101] The active *holo* form of the T domain is derived from its *apo* form with the conserved serine residue being posttranslationally modified by a phosphopantetheine transferase (PPTase).[101] As PPTase-mediated installment of Ppan cofactor is a functional prerequisite, PPTases play an essential role in NRP synthesis.[103,104]

Condensation domains

The condensation (C) domain (ca. 50 kDa) belongs to the superfamily of chloramphenicol acetyltransferases[80] and is located in between the upstream T and downstream A domain. As the name suggests, a C domain catalyzes condensation reaction of two activated and T domain-bound substrates by forming an amide bond. It mediates the nucleophilic attack of the downstream T domain-bound acceptor amino acid using its free α-amino group on the α-carboxyl group of the upstream T domain-bound donor amino acid or peptide for peptide elongation (Figure 3c).[80,105] The C domains can be differentiated according to their stereoselectivity: $^L C_L$ domains

promote the coupling of two L-amino acids, whereas ^DC_L domains catalyze the condensation of an upstream D-amino acid and a downstream L-amino acid.[106]

The C domain consists of two subdomains, an N-terminal and a C-terminal subdomain.[107] The N-terminal part of the C domain is called donor site and the C-terminal part acceptor site. Structural study revealed that the C domain is a V-shaped pseudodimer of two subdomains that form a central cleft at their interface.[107] Sequence alignments, along with biochemical characterizations of C domains, suggested a highly conserved HHxxxDG motif (the so-called "His motif") that is involved in catalytic activity of C domains.[105] The His-motif is located at the interface of both subdomains; a solvent channel runs through the domains to allow access to the His motif.[107] The second histidine in this motif has been proposed to act as a general base to promote nucleophilic attack of the α-amino group on the thioester[108,109] or to stabilize the tetrahedral transition state.[110,111] To establish the binding between two T domain-bound substrates, the donor and acceptor Ppan arms have to penetrate from opposite sides to reach the conserved active-site motif HHxxxDG of the C domain.[79,84] The crystal structure of the internal T-C domain showed that C and T domains are connected by an 18 residue linker and domain-domain interactions are mediated via hydrogen bonds, salt bridges and hydrophobic patches.[110]

Interestingly, it has been suggested that C domains play the role of a second selectivity filter during NRP synthesis.[112] In the case of incorrect substrate selection by the A domain, a second proofreading at the C domain minimizes the error rate of an NRPS.[79] The substrate selectivity testing of C domain donor and acceptor sites with chemically activated amino acids revealed that the acceptor site of the C domain can act as its major selectivity filter according to the side chain and chirality of amino acids.[112] Due to the lack of substrate-bound C domain structures, efforts to deduce a specificity-conferring code equivalent to that of A domains have been proved challenging.[79]

In addition to "standard" C domains, there is a separate phylogenetic subtype of C domains, termed starter condensation (C_{start}) domains. A C_{start} domain is located in the NRPS initiation module, where it catalyzes the acylation of the first incorporated amino acid.[113] A C_{start} domain transfers an acyl carrier protein (ACP)-bound fatty acid from fatty acid biosynthesis to the first amino acid[114,115] and can be found in the

NRPSs of lipopeptides, such as anikasin,[116] daptomycin,[117] and taxlllaids.[68] The crystal structure of the C_{start} domain revealed no major structural differences from internal C domains, with the exception of N- and C-terminal subdomains that are in a more closed conformation.[118]

1.3.2 Editing Domains

The versatile functions of NRPs owe to their diverse structures that arise from the ability of NRPSs to incorporate various building blocks as well as to install additional modifications during and after NRP assembly.[79] While the amino acid is covalently tethered onto the T domain, several editing domains can carry out the modifications.[80] These modifications often confer beneficial properties to NRPs that are important for biological activity.[80]

Epimerization domains

The majority of NRPs incorporated with D-amino acids obtain beneficial properties, for example, predetermining bioactive conformations or lowering their proteolytic susceptibility.[79] Although some A domains are capable of activating D-amino acids which are provided by cytosolic racemases, for example, in the cyclosporine synthetase,[119,120] most stereochemical transformations are performed *in situ* on T-domain-tethered substrates by NRPS-integrated epimerization (E) domains (ca. 50 kDa).[121] These specialized domains are typically embedded into the NRPS assembly line between T and $^D C_L$ domains and share conserved sequence motifs as well as the same overall protein fold with their evolutionary ancestor C domains.[111,113] The enzymes promote epimerization of the α-carbon of the T domain-bound L-amino acids or C-terminal amino acids of the growing polypeptide to afford a racemate of L- and D-amino acid.[122] Specific incorporation of only the D-amino acid into the growing peptide chain is ensured by the downstream C domain.[112]

In addition to E and C domains, there are dual condensation/epimerization (C/E) domains (ca. 51 kDa) with epimerization and $^D C_L$ activity,[123] thereby explaining D-configured building blocks in NRPs despite the absence of E domains in the corresponding NRPS modules.[50] The usual module sequence in the assembly line is C/E-A-T. The dual C/E domain catalyzes the epimerization of the L-amino acid incorporated by the upstream module into its D configuration and subsequently promotes condensation of the downstream T domain-bound acceptor amino acid with

the D-amino acid.[86] Strikingly, dual C/E domains lack the catalytically important glutamate residue of E domains,[113,123] which suggests that these bifunctional catalysts have evolved a distinct mechanism that allows two reactions in the same catalytic pocket.[79] Dual C/E domains harbor a second HHxxxD motif in helix a1,[113] which covers the active site His residue and hence may modulate its catalytic properties.[79]

Methyltransferase domains

Some NRPs such as cyclosporine,[124] pristinamycin[125] and rhabdopetides[34] exhibit N-methylation of backbone amino groups that is by far the most abundant methylation type. The N-methylation during NRP biosynthesis represents a stable modification that has a significant impact on polarity, hydrogen-bonding capabilities, proteolytic stability, and conformational freedom.[126] The majority of N-methylations are catalyzed by methyltransferase (MT) domains (ca. 45 kDa), which are integrated into the A domains.[79] The most frequent insertion point is the region between the core motifs A8 and A9 located in the flexible A_{sub} hinge domain, but an integration between the core motifs A2 and A3 of the A_{core} domain has also been reported.[127] The MT domain catalyzes the transfer of the S-methyl group from the cosubstrate S-adenosyl methionine (SAM) to the backbone amino group of the T domain-bound substrate by releasing S-adenosyl homocysteine (SAH) as a by-product.[79]

Formyltransferase domains

Similar to the transfer of fatty acids to the N-terminal substrates there are domains that transfer a formyl residue to the activated substrates of the first A domains. These domains are termed formyltransferase (FT) domains and are commonly located in the initiation modules with the module sequence FT-A-T, as observed in the synthesis of several NRPs, such as linear gramicidin,[128,129] anabaenopeptilide,[130] and kolossin.[50] The N-terminal valine residue of linear gramicidin decorated with a formyl group is the most prominent and well-characterized example of N-formylation in NRP biosynthesis.[129] The attachment of the formyl group to valine is catalyzed by the small FT domain (ca. 20 kDa) located at the N-terminus of the initiation module of linear gramicidin synthetase (LgrA), in the presence of the cofactor N^{10}-formyltetrahydrofolate (fTHF).[129] The formyl group is essential for linear gramicidin synthesis to proceed and for the important antibacterial activity.[131]

Further Modifications

Besides epimerization, methylation and formylation, further modifications can be introduced into the NRPs by optional domains, including cyclization (Cy), oxidase (Ox), ketoacyl reductase (KR), and monooxygenases (MOx) domains. For further information concerning these modification domains and the underlying principles please refer to the reviews.[79,80]

1.3.3 Peptide Release Domains

Once synthesized, the NRP reaches the final T domain as a thioester and has to be cleaved off from the NRPS assembly line in order to release the mature product and to regenerate the multienzyme complex for the next catalytic cycle.[132] Nature employs several different domains to catalyze the mature peptide release;[79] two most widely used domains are described in detail in the next subsections.

Thioesterase domains

The most common way to release the covalently bound final peptide from the NRPS assembly line is via thioesterase (TE) domains (ca. 30 kDa), particularly present in most bacteria NRPSs.[133,134] TE domains are located C-terminally to the T domain of the termination module[80] and belong to the α/β-hydrolase superfamily.[135] A conserved catalytic triad, composed of three amino acid residues in the binding pocket, is responsible for the enzymatic acitivity.[136] In the catalytic reaction, the active site serine residue, or in some cases cysteine, attacks the N-terminal peptidyl-T-domain thioester and thus temporarily anchors the oligopeptide through an intermediate ester bond to the TE domain.[79,137] For peptide release from the TE domain, there are two major routes.[135] One is hydrolysis in which water acts as the nucleophile to attack the carbonyl carbon of (peptidyl)acyl-O-TE, and a linear product is released. Another major route of product release is intramolecular cyclization. In this mechanism, a hydroxyl or an amino group of the (peptidyl)acyl chain acts as an internal nucleophile to attack the carbonyl carbon of (peptidyl)acyl-O-TE, and a cyclic product is released. TE domains provide a source of diversity in cyclization, as a variety of groups can be the nucleophile in the reaction, for example, the N-terminal backbone amino group of peptides (head-to-tail cyclization),[138] side chain hydroxyl or amino group of the peptides or fatty acids (branched cyclic molecule).[132,139,140]

Terminal condensation domains

Alternatively, peptide release can be catalyzed by terminal condensation (C_{term}) domains instead of TE domains. There are two different modes of action of such C_{term} domains.[141] The first type of C_{term} domain mediated peptide release is most often encountered in fungal NRPS systems,[142] which use the C_{term} domains to catalyze the condensation between the N-terminal and C-terminal residue of the peptide chain resulting in a cyclic final product, for example, cyclosporine A[119] and apicidin.[143] The other type of C_{term} domains release the final peptide by amination of the C-terminal residue of the peptide resulting in a linear final product.[35] Hereby, the C_{term} domains catalyze the formation of a terminal amide bond between the T domain-bound linear peptide and an amine that is freely available in the cytosol.

RXP NRPSs in bacteria of *Xenorhabdus* and *Photorhabdus* harbor the second type of C_{term} domains.[34] They are involved in the condensation reaction of amines such as phenylethylamine, tryptamine, or tyramine to release the final peptide; these natural amines might be derived from the decarboxylation of the three aromatic amino acids phenylalanine, tryptophan and tyrosine, respectively.[34,35] There are two structurally different C_{term} domains in RXP NRPSs. Besides the C_{term} domain as the final domain at the end of the module as in *X. nematophila*,[35] a stand-alone C domain is observed as the last module in *Xenorhabdus* KJ12.1.[34]

Structural data of a C_{term} domain that catalyzes the peptide release via cyclization from TqaA is available.[144] Analogous to the standard C domains, the catalytic activity of the TqaA C_{term} domain is dependent on the His residue.[79] Furthermore, the X-ray structures of the TqaA C_{term} domain revealed the same overall protein fold with canonical C domains, but a more compact arrangement of the two subdomains to each other, thereby blocking the non-functional acceptor site and corresponding solvent channel.[144]

Other release mechanisms

In addition to the above mentioned domains that act to release the mature peptides from ultimate T domains in NRP biosynthesis, several alternative release mechanisms, for example, terminal reductase (R) domain-mediated release mechanisms, R domain- or TE domain-mediated Dieckmann-type cyclization, terminal BtrH-like domain-mediated peptide release, and product release with

unknown mechanism, have been observed. For more information, please refer to the reviews.[79,135]

1.3.4 Types of NRPSs

The NRPS systems can be simply classified into three groups, linear NRPSs (type A), iterative NRPSs (type B), and nonlinear NRPSs (type C), according to their biosynthetic logic.[81] These groups are briefly introduced in the following subsections.

Type A − linear NRPSs

As mentioned before, NRP synthesis by an NRPS is often compared to a molecular assembly line, where each NRPS module incorporates one amino acid into the elongating peptide chain, based on the collinearity rule. This rule is true for linear NRPSs, which are also called type A NRPSs[81] and represent the most common form of NRPSs. Examples for this type A NRPSs are the tyrocidine,[145] surfactin,[146] and cyclosporine synthetases.[119]

Type B − iterative NRPSs

Some NRPS machineries reuse dedicated modules multiple times during a full biosynthetic cycle.[79] These iterative NRPSs are called type B NRPSs[81] and they are less common than type A NRPSs. In many cases, they cause molecular symmetry of the products, for example, enterobactin, a cyclic trimer of dihydroxybenzoylserine units.[147] Mechanistically, the iterative biosynthesis affords a repetitive utilization of modules and requires a so-called "waiting position" for the reaction intermediates.[79] This storage function of product intermediates can be either fulfilled by T domains (e.g. T3 of bassianolide synthetase) or by terminal TE domains (e.g. in gramicidin S synthetase).[79,148,149] A reasonable explanation why nature makes use of an iterative mode may be that large and energetically more costly NRPSs with their high susceptibility to misfolding and proteolysis are preferably avoided.[79]

Type C − nonlinear NRPSs

A further variation of the iterative mechanism is the nonlinear mode (type C),[81] which relies on the fact that one domain, not an entire module, is used more than once during NRP biosynthesis.[79] This type C often has one unusual C, A or T core domain;[150] for example, in mannopeptimycin assembly in which one A domain

provides the respective substrate not only for its intramodular T2 domain but also for the T3 domain of the C-terminal module.[151] Due to the complexity of domain interplay, more precise distinctions can be performed only with progress in understanding the underlying biosynthetic logic in nonlinear NRPSs.[81]

1.4 Discovery of Nonribosomal Peptides

As mentioned earlier, the highly complex assembly line of NRPSs produce structurally and functionally diverse NRPs, which are widely used in medicine for the treatment of a range of clinical diseases. Due to the urgent need of new drugs against the increasing drug resistance, NRPs will continue to act as invaluable source of lead compounds for drug discovery.

However, traditional bioactivity-based screening of extracts from cultures of microorganisms that are recognized as prolific producers of bioactive NPs has yielded disappointing returns due to the tendency to rediscover already known metabolites.[152,153] Fortunately, NRPS gene clusters are comparatively easy to identify because of their large multidomain organization by modern genome mining approaches.[79] The large and complex products of NRPSs are also easy to predict by a variety of bioinformatic tools, which potentially allow rapid assessment of product novelty.[154] Thereby, microbial genome sequencing, in combination with bioinformatic analysis, is one exciting approach to identify new NRP BGCs probably also producing novel and bioactive NRPs. Recent advances in genome sequencing and bioinformatics have revealed that many microorganisms have a far greater potential to produce specialized metabolites than it has been discovered by classic screening-based methods.[153]

Nevertheless, the potential of microbes to produce such novel NPs has been hampered by the fact that many specialized metabolite BGCs are silent or expressed at very low levels under standard laboratory conditions due to their tight control in response to either direct or indirect environmental signals.[155,156] Hence activating the expression of silent BGCs is the key to achieve the full potential of genomics-driven approaches for the discovery of novel microbial NPs.[157] To reach this important goal, many strategies have been developed.

Following activation of a BGC, an essential stage in NP discovery is the elucidation of their structures that is necessary for detailed biological activity studies and structure−activity relationship investigations. The structure elucidation often comes along with compound isolation, which usually can be achieved by chromatographic and spectroscopic methods.

A typical workflow for genomics-driven discovery of NPs, most useful for NRPs derived from gene clusters expressing NRPSs, is illustrated in Figure 4. The identification and activation of novel NRPSs, coupled with the isolation and characterization of their metabolic products, currently represents one of the most promising avenues for the discovery of new NRPs.[153] The methods used for pathway-specific activation of NRPSs and structure elucidation of their products, which have been addressed in this thesis, are described in the next sections. Most methods can also be applied to other classes of BGCs and their products. For more methods, including new developments of analytical technologies, please refer to recent reviews.[153,158–160]
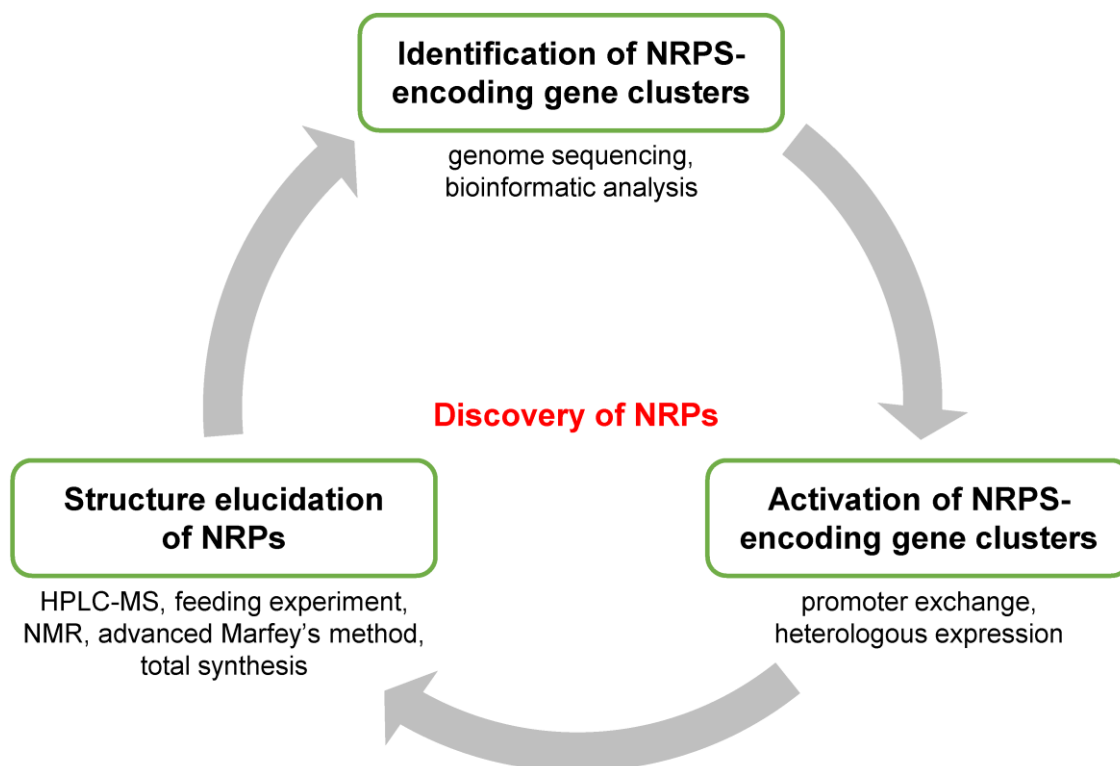
Figure 4. Overview of the workflow for genomics-driven discovery of NPRs. NRPS-encoding gene clusters are first identified from genomic sequences by bioinformatics. Silent NRPS gene clusters are then activated by different activation strategies. Structures of new NRPs are elucidated by using a combination of different analytical methods. The obtained data can further guide the next round of genome mining. Adapted from[158].

1.4.1 Activation of Silent NRPS Gene Clusters

Promoter exchange

Promoter exchange strategies usually involve the replacement of the natural promoter of a gene cluster of interest with a strong constitutive promoter or an inducible promoter whose regulation is well understood.[53,153] This method can be applied in the native host and takes advantage of the intact elements in BGCs, for example, precursor supply and end-product export.[158] As a pathway-specific approach, such strategy is particularly useful for BGCs when information pertaining to the structural novelty of their products can be gleaned from predictive bioinformatic analysis.[153]

As an example, promoter exchange has been used to activate a silent BGC in *Burkholderia pseudomallei*, in which a constitutive promoter was introduced upstream of a putative operon encoding a hybrid PKS−NRPS biosynthetic pathway, triggering expression of the operon and inducing the production of burkholderic acid that may function as a signaling molecule involved in quorum sensing.[161] Promoter exchange has also been used to activate silent BGCs in *Photorhabdus* and *Xenorhabdus* through exchanging the natural promoter against the well-known arabinose-inducible promoter P_{BAD}.[53,138] The inducible systems allow the analysis of induced versus non-induced conditions and thus can mimic an "overproducing mutant" and a "knock-out mutant" in single strain grown under two different conditions, leading to a more reliable identification of NPs, such as GameXPeptides, xenoamicins, and indigoidine.[53] Very recently, by exchanging the native promoter of a desired BGC against an inducible promoter in deletion mutants of the gene encoding the global regulator Hfq, almost exclusive production of the corresponding NPs from the targeted BGCs in *Photorhabdus*, *Xenorhabdus* and *Pseudomonas* was observed including the production of several new NPRs derived from previously uncharacterized NRPSs.[162] This approach can act as an efficient tool for specialized metabolite production enabling direct bioactivity testing.[162]

Heterologous expression

Heterologous expression is an alternative pathway-specific approach for BGC activation. In contrast to a promoter exchange, this method involves the expression of an intact BGC in a heterologous host. The tight regulation of silent BGCs can be

relieved by either direct cloning or refactoring using well-characterized promoters.[158] Moreover, heterologous hosts possess significant growth advantages over most native hosts and therefore make activation of BGCs even from unculturable organisms possible.[158]

The strategy of heterologous expression has been used, for example, to activate a silent large NRPS gene cluster (67 kbp) from the marine actinomycete *Saccharomonospora* sp. CNQ-490 by transformation-associated recombination cloning and transferring it into the model heterologous host *Streptomyces coelicolor*, resulting in the discovery of two novel lipopeptide antibiotics, taromycin A and B.[163,164] In addition, a cryptic NRPS gene cluster from *Paenibacillus lavae* was cloned by RecET recombination system and activated by heterologous expression in *E. coli*, resulting in the production of the novel compound sevadicin.[165]

It is important to note that correct NRP biosynthesis might be impossible, if it is dependent on essential building blocks that cannot be synthesized by the heterologous host. For example, expression of the GameXPeptide producing NRPS from *P. luminescens* in *E. coli* results in the production of derivatives without *p*-aminophenylalanine, as this building block is specifically produced only by *Photorhabdus* in infected insects.[77,166,167]

1.4.2 Structure Elucidation of NRPs

Bioinformatic analysis

Thanks to a deep understanding of NRPS biosynthetic principles and significant advance of bioinformatics, structure elucidation efforts to the large and complex NRPs can directly or indirectly benefit from the utilization of genomic information that can be now obtained quickly and cheap, especially for bacteria.[154] The program antiSMASH (antibiotics and secondary metabolite analysis shell) is probably the most widely used bioinformatic tool for identifying and analyzing BGCs in genome sequences now.[91] It is worth noting that structural information of NRPs from antiSMASH is sometimes incomplete. Although substrate specificities of NRPS modules can be analyzed, the algorithm does not take nonlinearity, module-skipping, cyclizations, or modifications into account.[154] Nonetheless, the output from antiSMASH and its embedded tools still provide a wealth of structural information.[154]

One application of genomics followed by bioinformatic analysis is to assess compound novelty. If a gene cluster for an unknown metabolite is sufficiently similar to another gene cluster for a known compound, expensive scale-up and spectrometric efforts can be avoided; conversely, one can reasonably expect a new structure if the gene cluster is not found within the databases. In *Thermobifida fusca*, for instance, an NRPS gene cluster showing sufficiently divergent from known clusters was identified, and the compound fuscachelin was predicted to be novel based on the sequence-inferred NRPS module amino acid substrate specificity.[168]

Furthermore, it is possible to predict the structures of some NRPs because the NRPSs are organized in a multimodular fashion with the order of the modules frequently aligning with the building block order in the corresponding products, a concept known as "collinearity".[154] Here, the NRPS module substrate specificity, as discussed previously, can be predicted. For example, the pyrazinone NPs of NRPS origin, aureusimines, were found in *Staphylococcus aureus* via genome mining and their structures were predicted reasonably-well from their NRPS as composed of tyrosine and valine.[169]

Besides, genomic information can be useful in verifying that the spectrometrically derived assignments are correct, or at least consistent with biosynthetic logic, particularly for NRPs that consist of a large number of repeated or similar amino acid residues with overlapping spectral properties. An example is teixobactin, an antibiotic depsipeptide recently discovered from *Eleftheria terrae*; the bioinformatically identified two-gene NRPS cluster was found to consist of eleven modules whose predicted specificities exactly matched the elucidated structure.[170]

Moreover, gene sequence-derived knowledge of biosynthetic pathways can be used to confirm or predict amino acid stereochemical configuration. NRPS modules typically employ L-amino acids as substrates and D-amino acids are incorporated based on the presence of E or dual C/E domains. In the example of the 13-residue depsipeptides xenoamicins isolated from *X. mauleonii*, the complete determination of amino acid configurations depended on the genomic analysis of the closely-related, already-sequenced *X. doucetiae*.[71] However, caution must be taken with the *in silico* predictions, since not all E or C/E domains are active, such as in the case of myxochromides.[171] Additionally, D-amino acids may be directly utilized by certain NRPSs, presumably epimerized by enzymes elsewhere in the genome and thus not

easily discernable through bioinformatics.[120,154] Thus, bioinformatic predictions are best trusted when they can be supported by some spectroscopic or synthetic evidences.[154]

To gain absolute certainty of the NRP structure, several analytical methods, including HPLC-MS analysis, feeding experiment, NMR analysis, and the advanced Marfey's method, sometimes supplemented with total synthesis, are still required and discussed in more detail in the following subsections.

HPLC-MS analysis

Mass spectrometry (MS) measures the mass-to-charge ratio (m/z) of individual molecules present as charged ions. The equipment is composed of three principal parts: ionization source, mass analyzer and detector. The usability of MS especially in life sciences was changed fundamentally with the invention of soft ionization strategies like matrix-assisted laser desorption/ionization (MALDI)[172] and electrospray ionization (ESI).[173] MALDI is often employed for the analysis of proteins, peptides and low-mass molecules, while ESI is now widely used for the study of a vast majority of polar to medium polar small molecules.[159] The most eminent advantage of MS is the high sensitivity in comparison with other methods. Furthermore, MS can be coupled to chromatographic separation prior to ionization and detection, which largely increases the scope of usage. Today, MS is used as a standard analytical technique for molecular structure elucidation.

In the process of NRP identification, high performance liquid chromatography coupled to mass spectrometry (HPLC-MS) is a powerful analytical technique. Firstly, a compound can quickly be discarded as non-novel before scale-up or purification is attempted, using accurate mass from high resolution mass spectrometry (HRMS) on biological samples. Meanwhile, the molecular formula of a new compound can generally (and even unambiguously) be obtained by its HRMS data, which is usually the initial step of structure elucidation.[174,175] Furthermore, the structural information of NRPs, including amino acid composition and sequence, can be investigated by tandem mass spectrometry (MS^2) approaches. When HPLC separation is coupled to MS^2, retention time combined with fragmentation patterns and relative abundances of product ions can be used for the identification of an unknown compound by comparison with that of an authentic standard. This method is especially useful for the identification of minor products that are difficult to isolate but can be chemically

synthesized for structural confirmation. The linear peptide xenortide D, for instance, was produced in only trace amount in *X. nematophila*; the structure of xenortide D was deduced by combining detailed HPLC-MS analysis with feeding experiments and was confirmed by chemical synthesis.[36] Additionally, the structural information of peptides obtained from MS^2 experiments can directly be used to identify the corresponding BGCs, because, as already described, peptides are all derived from genetically-encoded modular pathways. Such an MS-guided genome mining approach is known as natural product peptidogenomics that connects the chemotypes of peptide natural products to their BGCs by iteratively matching *de novo* tandem MS structures to genomics-based structures following biosynthetic logic.[176]

Another promising application of HPLC-MS is molecular networking analysis that inspects a suite of MS^2 spectra obtained on a whole extract and clusters the detected analytes by spectral similarity. In this way, it is possible to get information on related metabolites in an extract and provide rapid access to NP congeners or biosynthetic intermediates.[154] Visualization of the generated molecular networks also provides a broad overview of the metabolic composition of an extract.[154] This method was recently applied with extracts from *Xenorhabdus* and *Photorhabdus*, resulting in the discovery of a large number of previously unidentified metabolite classes, including the NRPS-derived xefoampeptides.[23]

Although with these excellent features, the use of HPLC-MS-based structure elucidation is limited concerning the complex structural properties of many NRPs including incorporation of nonproteinogenic amino acids and internal cross-links of different parts of one compound. Thus by HPLC-MS analysis alone it is impossible to unambiguously identify a molecule and consequently one must rely on complementary information.[159]

Feeding experiment

As shown above for the identification of xenortide D, feeding experiments, also termed as labeling experiments, are an important supplementary approach for NRP structure identification that is always followed by HPLC-MS analysis. First, feeding experiments can be used to determine the molecular formula of a NP.[47] When an extract from standard growth medium is analyzed by HPLC-MS, the sum formula of the desired peak can be predicted from HRMS data. Nevertheless, it is often difficult to find the correct formula due to several possibilities within a given ppm range

(commonly ≤5 ppm) without ultrahigh resolution MS systems. When the NP-producing microorganism is grown in fully labeled ^{13}C or ^{15}N medium, the correct sum formula can be rapidly identified by HPLC-MS analysis, as the numbers of nitrogen and carbon atoms can be easily determined from the mass shifts of the desired compound peak. Feeding experiments can also be used to identify building blocks of NPs such as amino acids.[47] Here, labeled precursors with ^{13}C, ^{2}H, or ^{15}N atoms are usually added to a producing culture and their incorporation confirms the involvement of the precursors in the NP biosynthesis. Hence HPLC-MS-based detection allows for the identification of precursor incorporation by a mass shift to higher mass. However, as not all possible precursors are available for a reasonable price, one can also add a precursor with natural abundance (^{12}C, ^{14}N) to a culture grown in a fully labeled ^{13}C or ^{15}N medium. Precursor incorporation in such an inverse labeling experiment is easily visible by a shift to lower mass.

Also, the determination of the amino acid configurations can be performed by feeding experiments followed by HPLC-MS analysis.[47,77] During most NRP synthesis, E domains or mixed C/E domains are responsible for the transformation of L-amino acids into their D-forms. By feeding an amino acid with a ^{2}H label at the α-carbon position to a transaminase mutant or a transaminase-deficient heterologous host carrying the NRPS gene cluster, the presence of an active E domain can be probed indirectly by HPLC-MS analysis. If the label is found in the compound mass (higher mass), the amino acid at this position must have the L configuration due to a missing or nonfunctional E domain. If the label is lost, then it has been exchanged against ^{1}H from the culture medium by means of E-domain-catalyzed racemization, and the amino acid at this position is confirmed to have the D configuration. These feeding approaches have been widely used for the identification of NRPs from microorganisms, including *Xenorhabdus* and *Photorhabdus*.[53,138,177,178]

NMR analysis

Nuclear magnetic resonance (NMR) spectroscopy measures the chemical shifts (ppm) of pure compounds that are usually isolated by using chromatography techniques. As the chemical shift depends on the exact electronic environment surrounding the nucleus, chemical shifts can be diagnostic for a specific metabolite. In contrast to MS, NMR spectroscopy is a rather insensitive but universal detection method and thereby is seen as the gold standard for structure elucidation of NPs.

To elucidate an unknown NP, one-dimensional (1D) and two-dimensional (2D) NMR experiments are always carried out following the determination of the molecular formula using HRMS. ^{1}H and ^{13}C spectra separately record chemical shifts of proton and carbon atoms in a molecule and belong to 1D NMR. They are often measured prior to 2D NMR experiments. In the case of NRPs, 1D NMR often exhibits the characteristic signals of a peptide, illustrating amide NH protons and α-amino protons in ^{1}H spectrum along with carbonyl carbons in ^{13}C spectrum.[179]

The acquisition of 2D NMR data, which is plotted like a topographic map with cross-peaks on the map indicating linked nuclei, plays a crucial role in structure elucidation. The most frequently used set of 2D NMR experiments are correlation spectroscopy (COSY), heteronuclear single quantum coherence (HSQC), and heteronuclear multiple bond correlation (HMBC).[159,180] The COSY spectrum usually reveals homonuclear correlations between vicinal hydrogens separated by three bonds ($^3J_{HH}$). This makes it possible to identify the neighbor carbon atoms connected by a chemical bond. The HSQC spectrum shows heteronuclear correlations that arise as a result of $^1J_{CH}$ couplings between ^{13}C nuclei and protons attached to the corresponding atoms. This allows one to detect all CH, CH_2 and CH_3 groups with chemical shift assignment. The HMBC spectrum reveals heteronuclear correlations between ^{1}H and ^{13}C nuclei separated by two or three chemical bonds, sometimes in conjugated systems, four bonds, but the direct one bond correlations are suppressed. One can use the HMBC spectrum to establish the connectivity of building blocks such as amino acids in NRPs. In principle, combining all 1D and 2D NMR data allows the complete assembly of a structure. However, as structure elucidation of NPs is a complicated process, it is not surprising that different scientists may come to different structures from the same initial data; many structural misassignments have been reported in a series of reviews.[181–183] This is not uncommon for NRPs due to their complex structures and large number of repeated or similar building blocks, as it has been observed for microginins, peptides isolated from *Microcystis aeruginosa*.[184] A in-depth analysis of the NMR and HRMS/MS data of the metabolites proved that previously published structures were wrong, showing an inversion of two amino acids in the structure. Thus, to unambiguously identify an NRP, NMR spectroscopy should be carefully examined and a combination with other analytical methods such as bioinformatics, MS and total synthesis can be necessary.

Advanced Marfey's method

Having determined the planar structure of a given NP, the next essential step is often configurational assignment, given that the configuration is of utmost importance for biological activity. The most extensively used approach for configuration determination of amino acids in new peptides is the advanced Marfey's method,[72,185] which is also employed for confirmation of bioinformatically predicted configurations of the amino acids in NRPs.[138] This method usually starts with the hydrolysis of the peptide with 6 M HCl aqueous solution. The hydrolysate is separately derivatized with Marfey's reagents like Nα-(5-fluoro-2,4-dinitrophenyl)-L-leucinamide (L-FDLA) or Nα-(5-fluoro-2,4-dinitrophenyl)-D-leucinamide (D-FDLA). The analysis of the derivatized amino acids is carried out by HPLC-MS and the configurations are determined by the elution order. If necessary, the commercial standard amino acids are prepared as reference. However, in some cases, the advanced Marfey's method is limited because of the presence of several repeated amino acids. For example, the amino acid configurations of xenoamicin A were determined by the advanced Marfey's method, but only partial information was provided, as the compound contains four L- and one D-valine residues and the stereochemical assignments to the corresponding valine residues were not possible at this stage.[71] Therefore, bioinformatic analysis and total synthesis can act as important alternative methods.

Total synthesis

As mentioned before, total synthesis is an important approach for structural confirmation of NRPs that are predicted by bioinformatics and HPLC-MS analysis. Meanwhile, the synthesis is a key cornerstone in the structural revision of NRPs and often provides sufficient amount of compounds for further bioactivity testing.

Chemical synthesis of peptides consists of liquid- and solid-phase synthesis. Originally, the synthesis was performed in solution. The liquid-phase peptide synthesis is highly flexible with respect to the chemistry of coupling and the combination of the peptide building blocks. However, this technique is used for the synthesis of small peptides composed of only a few amino acid residues. In addition, it is time-consuming to separate the peptide intermediates or products from impurities and side products.

Peptide synthesis has been revolutionized after the emergence of solid-phase peptide synthesis (SPPS). In contrast to naturally occurring peptide synthesis, the sequence elongation starts with the C-terminal amino acid immobilized on an insoluble support via a cleavable linker. Coupling reagents and excess amino acids can be washed away, thereby minimizing purification issues. Additionally, the individual reaction steps can be easily automated, leading to a better time management. These principles make SPPS simple but productive. As an example, kolossin A, a large pentadecapeptide was produced in *P. luminescens* after the NRPS-encoding gene was activated by promoter exchange.[50] However, the trace amount of the product did not allow its isolation for NMR experiments but the possible structures, including the amino acid configurations, were deduced by bioinformatics, feeding experiments and detailed HPLC-MS analysis. Out of all possible structures that were independently synthesized using SPPS and compared with the natural product using HPLC-MS, the correct one was successfully identified. The synthesis also provided enough amount of kolossin A for biological screenings. Another example is similanamid, a cyclic hexapeptide isolated from the marine sponge-associated fungus *Aspergillus similanensis* KUFA 0013.[186] The total synthesis of the proposed structure was achieved by solid-phase synthesis of a linear precursor and liquid-phase macrolactamization. The NMR spectra of the synthetic final product were not identical to those of the isolated material. Thereby, the conclusion was that similanamide is identical to PF1171C, a previously reported diastereomeric hexapeptide.

Summing up the abovementioned considerations, although the finding of the NRPS gene clusters and the bioinformatic analysis can be done quickly, the process of structure elucidation is still the most time-consuming and rate-limiting step. There is no single approach that is able to unambiguously identify an NRP, hence the structure elucidation of NRPs must rely on a combination of several methods.

1.5 Aim of This Thesis

Based on the introduced knowledge, the aim of this thesis was to search for and characterize new NRPs from *Xenorhabdus* and *Photorhabdus* strains by using a combination of biological and chemical methods.

RXPs are a class of structurally diverse NRPs exclusively found in *Xenorhabdus* and *Photorhabdus*.[34] These derivatives are composed of two to eight amino acids with valine as the most abundant one (followed by phenylalanine and leucine) that are often *N*-methylated and frequently incorporate phenylethylamine or tryptamine as the C-terminal amines.[34] Generally, RXP BGCs encode two to four single-module-subunit NRPSs; the chemical diversity of RXPs results from a combination of iterative and flexible use of these single-module-subunit NRPSs.[34] During a more-detailed HPLC-MS analysis of the extracts from *Xenorhabdus* and *Photorhabdus* strains, several new RXPs showing different C-terminal amines (putrescine and ammonia) instead of phenylethylamine and tryptamine were found in a *X. innexi* wild-type strain. Therefore, the first goal of the study was to investigate these new RXPs. The results, including structure elucidation, antiprotozoal activity, and proposed biosynthesis of these RXPs, are described in chapter 2.1. Additionally, several new methionine-containing RXPs were found as minor products during an in-depth investigation on RXP production in *E. coli* carrying *kj12ABC* from *Xenorhabdus* KJ12.1. Their structure elucidation and chemical synthesis, as well as antiprotozoal activity, are presented in chapter 2.2.

Then, genome mining of *P. asymbiotica* PB68.1 showed that the stain encodes an NRPS named PhpS. A detailed analysis revealed that PhpS comprises six modules and was thus expected to produce a hexapeptide. Since no such peptide could be identified in *P. asymbiotica* PB68.1, a promoter exchange approach was applied to *phpS*, resulting in the production of a library of new peptides, named photohexapeptides. The identification and biosynthesis of this library are described in chapter 2.3.

Next, in the genome of *P. temperata* Meg1, the *pdtS* gene was predicted to encode an unknown NRPS consisting of six modules. PdtS was hence expected to produce a hexapeptide. As no such peptide could be observed in *P. temperata* Meg1 when the strain was cultivated in the laboratory in different media such as lysogeny broth (LB) or Sf-900 II SFM medium, a promoter exchange approach was applied to activate *pdtS*, resulting in the production of photoditritide, a new cyclohexapeptide containing

two uncommon homoarginine residues. The structure elucidation, biosynthesis, as well as antimicrobial and antiprotozoal activity of photoditritide, are reported in chapter 2.4.

Last, bioinformatic analysis showed that *P. temperata* Meg1 encodes a BGC with two NRPSs, termed PttBC. Detailed analysis identified that PttBC consist of five modules, including a C_{start} domain in the initiation module of PttB. Thereby, PttBC were expected to produce a lipopeptide containing five amino acids. Notably, one MbtH-encoding gene, termed *pttA*, is located 0.8 kbp upstream from *pttB*. The MbtH proteins have been proposed to play an important role in stimulating adenylation reactions that are required for biosynthesis of some NRPs.[96,187] Nevertheless, no such peptide could be identified in extracts from either *P. temperata* Meg1 wild-type strain or *P. temperata* Meg1 pCEP-*pttA* promoter exchange mutant under standard laboratory conditions. Therefore, heterologous expression was applied to activate *pttABC*. Our effort expressing *pttBC* together with *pttA* in *E. coli* resulted in successful production of a new family of cyclic lipodepsipeptides, named phototemtides. Indeed, *E. coli* was found incapable of producing phototemtides without MbtH-encoding gene *pttA*. To identify phototemtides, a combined use of different analytical methods, including HPLC-MS analysis, feeding experiments, NMR analysis, the advanced Marfey's method, and total synthesis, were carried out. The results, along with antiprotozoal activity of the major compound phototemtide A, are presented in chapter 2.5.

2 Publications

2.1 Rhabdopeptide/Xenortide-like Peptides from *Xenorhabdus innexi* with Terminal Amines Showing Potent Antiprotozoal Activity

Authors:

Lei Zhao,[1,2] Marcel Kaiser,[3,4] and Helge B. Bode*[,1,5]

[1]Molecular Biotechnology, Department of Biosciences, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

[2]Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, 210014 Nanjing, China

[3]Parasite Chemotherapy, Swiss Tropical and Public Health Institute, 4051 Basel, Switzerland

[4]University of Basel, 4003 Basel, Switzerland

[5]Buchmann Institute for Molecular Life Sciences (BMLS), Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

*Corresponding author

Online access:

https://pubs.acs.org/doi/10.1021/acs.orglett.8b01975

Attachments:

Declaration on the contribution of the authors and the publication including supporting information

Summary

Parasitic diseases, such as malaria, sleeping sickness, Chagas disease, and leishmaniasis caused by protozoa, are a major public health problem worldwide, mainly affecting hundreds of millions of the poorest populations in tropical and subtropical countries. Because of the low return on investments, only a few antiprotozoal drugs are available. Unfortunately, most of these drugs are decades old and have one or more limitations, including high cost, toxicity, and drug resistance. Thus, there is an urgent need for the development of novel, effective, inexpensive, and safe drugs.

Rhabdopeptide/xenortide-like peptides (RXPs), which represent a unique class of peptides exclusively found in *Xenorhabdus* and *Photorhabdus*, are active against different protozoa. These RXPs are composed of two to eight amino acids (valine, phenylalanine, and leucine) that are often *N*-methylated and show phenylethylamine or tryptamine as the typical C-terminal amines for most RXPs. During a more-detailed HPLC-MS analysis of different *Xenorhabdus* and *Photorhabdus* strains, several new RXPs were identified showing putrescine or ammonia as the C-terminal amines from *X. innexi*.

In this study, seven new RXPs with putrescine or ammonia as the C-terminal amines were isolated from *X. innexi*. Their structures were elucidated by HRMS and NMR experiments. The absolute configurations of the amino acids were determined according to the advanced Marfey's method. Bioactivity testing of these RXPs revealed potent antiprotozoal activity against the causative agents of sleeping sickness (*Trypanosoma brucei rhodesiense*) and malaria (*Plasmodium falciparum*), with IC_{50} values of 0.07−6.25 and 0.091−3.16 μM, respectively, making them the most active RXP derivatives known to date.

Furthermore, bioinformatic analysis showed that the biosynthetic gene cluster (BGC) of RXPs from *X. innexi* encodes three monomodular nonribosomal peptide synthetases (NRPSs), namely, InxA, InxB, and InxC. Based on the chemical structures, combined with amino acid specificity for adenylation (A) domains, the biosynthesis for these RXPs was proposed. InxA might act iteratively with a flexible methyltransferase activity to catalyze the incorporation of the first five or six *N*-methylvalines/valines to these peptides.

active against *Trypanosoma brucei rhodesiense* (IC_{50} 0.07–6.25 μM) and *Plasmodium falciparum* (IC_{50} 0.091–3.16 μM)
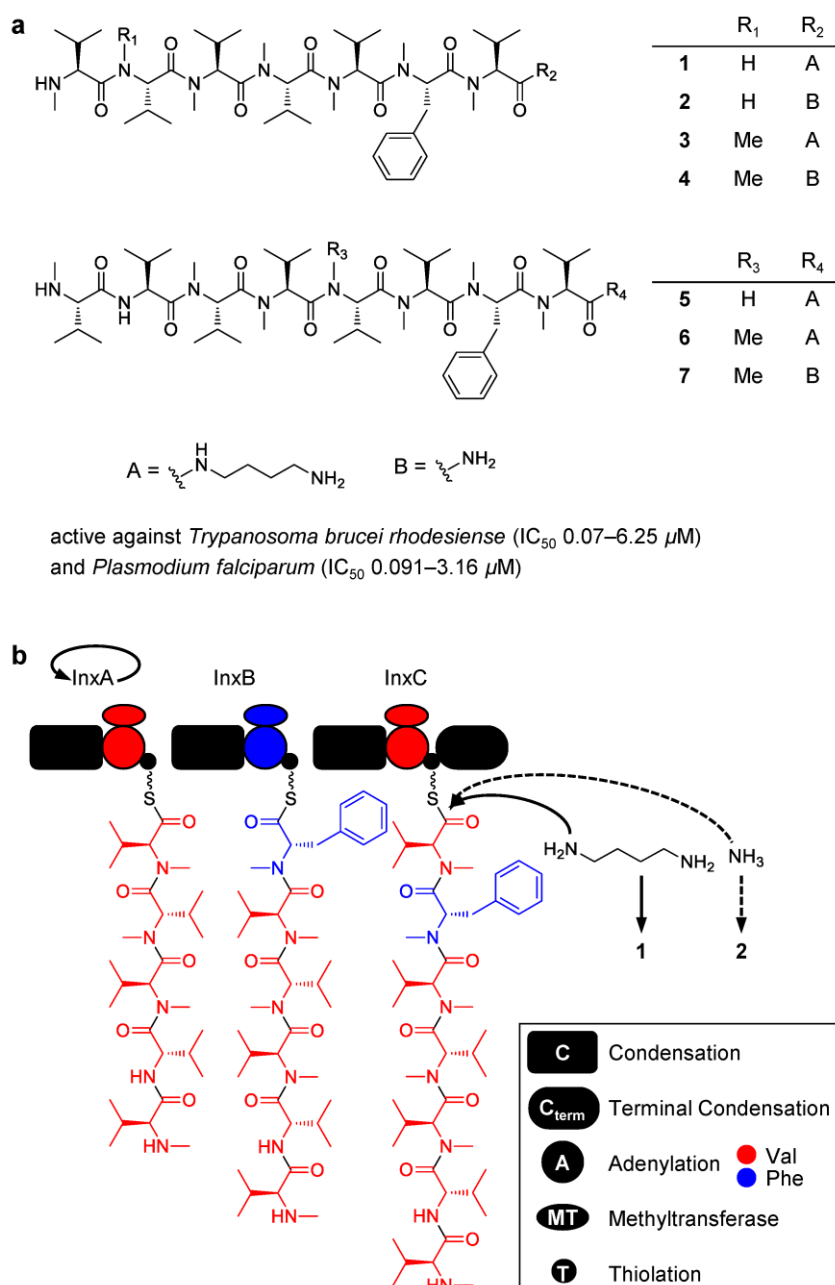
Figure 5. Graphical summary of the publication "Rhabdopeptide/Xenortide-like Peptides from *Xenorhabdus innexi* with Terminal Amines Showing Potent Antiprotozoal Activity". (a) Chemical structures of new RXPs (**1**−**7**) with putrescine or ammonia as the C-terminal amines and their bioactivity against protozoan parasites; (b) Proposed biosynthesis as shown for **1** and **2** as examples. InxA might act iteratively with a flexible methyltransferase activity to catalyze the incorporation of the first five *N*-methylvalines/valines to **1** and **2**.

2.2 Methionine-Containing Rhabdopeptide/Xenortide-like Peptides from Heterologous Expression of the Biosynthetic Gene Cluster *kj12ABC* in *Escherichia coli*

Authors:

Lei Zhao,[1,2] Xiaofeng Cai,[1] Marcel Kaiser,[3,4] and Helge B. Bode*[,1,5]

[1]Molecular Biotechnology, Department of Biosciences, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

[2]Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, 210014 Nanjing, China

[3]Parasite Chemotherapy, Swiss Tropical and Public Health Institute, 4051 Basel, Switzerland

[4]University of Basel, 4003 Basel, Switzerland

[5]Buchmann Institute for Molecular Life Sciences (BMLS), Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

*Corresponding author

**Online access:**

https://pubs.acs.org/doi/10.1021/acs.jnatprod.8b00425

Attachments:

Declaration on the contribution of the authors and the publication including supporting information

Summary

RXPs are a class of chemically diverse nonribosomal peptides (NRPs) exclusively found in *Xenorhabdus* and *Photorhabdus*. More than 60 RXPs have been identified in 51 different wild-type strains by detailed HPLC-MS analysis. So far, only the amino acids valine, phenylalanine, and leucine have been observed in the previously identified RXPs.

The RXP BGC from *Xenorhabdus* KJ12.1, named *kj12ABC*, encoding three monomodular NRPS subunits Kj12A, Kj12B, and Kj12C, is the simplest and best-characterized system. Chemical diversity of the RXPs in *Xenorhabdus* KJ12.1 results from a combination of iterative and flexible use of Kj12ABC. Previous studies have shown that heterologous expression of *kj12ABC* in *E. coli* led to the same production of RXPs as found in the wild-type strain of *Xenorhabdus* KJ12.1. However, during an in-depth investigation on RXP production in *E. coli* carrying *kj12ABC*, we found seven new methionine-containing RXPs as minor products besides the valine-only-containing main compounds.

In this publication, the structures of seven sulfur-containing RXPs were initially suggested by HRMS data and detailed analysis of their MS^2 fragmentation patterns. To confirm the proposed structures, these RXPs were synthesized by using a solid-phase peptide synthesis (SPPS) method developed for the synthesis of partially methylated RXPs. As RXPs are active against protozoa, the bioactivity of these methionine-containing RXPs as well as two valine-only-containing main products was evaluated. All tested compounds displayed good effects against *T. brucei rhodesiense* (IC_{50} 1.4−12.3 *μ*M) and *P. falciparum* (IC_{50} 2.3−10.9 *μ*M).

It is worth noting that methionine-containing RXPs were only detected in the heterologous *E. coli* strain, not in *Xenorhabdus* KJ12.1 wild-type strain. Furthermore, the expression of *kj12B* and *kj12C* (without *kj12A*) did not produce the methionine-containing RXPs. These results suggest that *E. coli* background and Kj12A might contribute to the formation of methionine-containing RXPs by a still unknown mechanism.
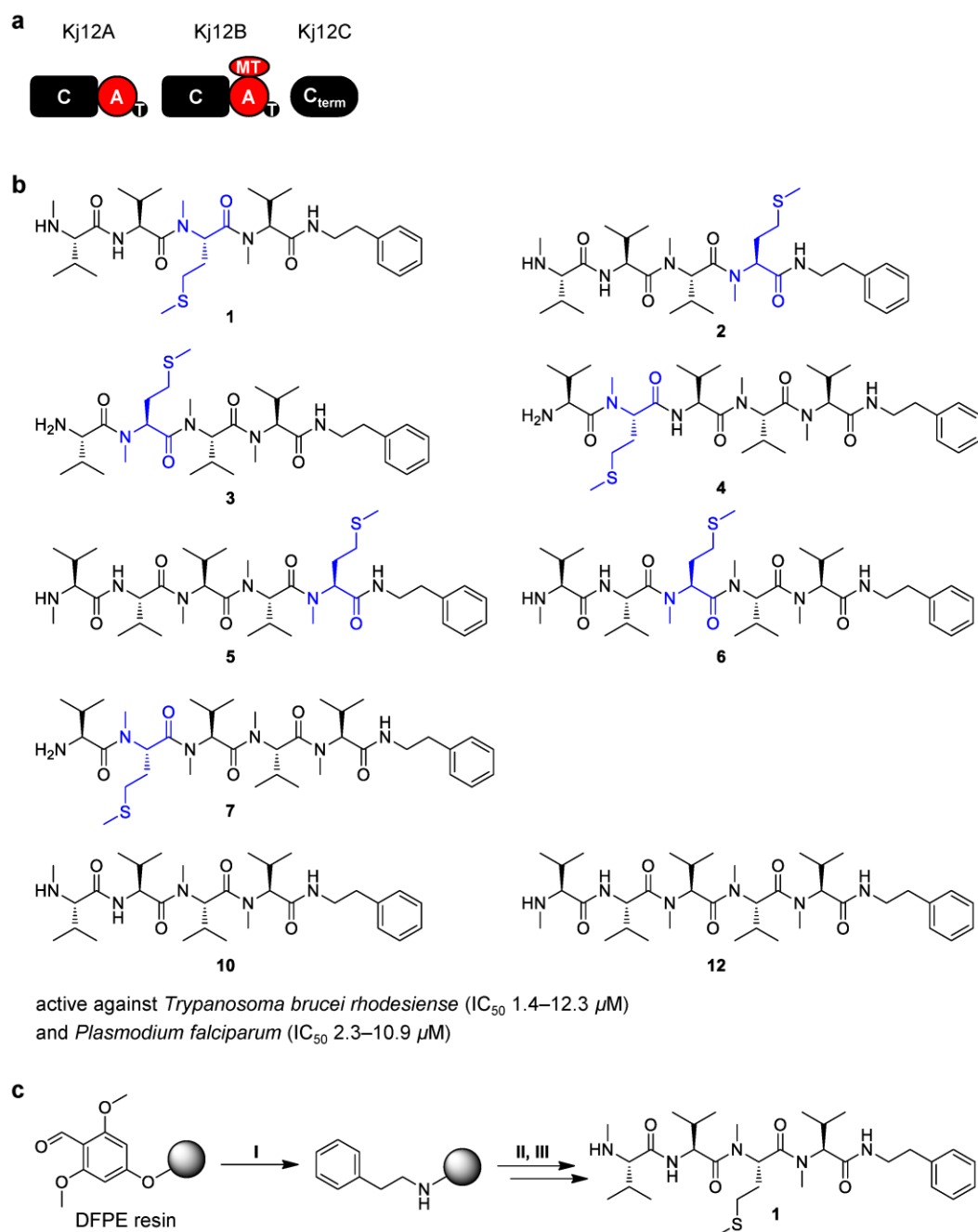
Figure 6. Graphical summary of the publication "Methionine-Containing Rhabdopeptide/Xenortide-like Peptides from Heterologous Expression of the Biosynthetic Gene Cluster *kj12ABC* in *Escherichia coli*". (a) Kj12ABC from *Xenorhabdus* KJ12.1. Domains: C: condensation, A: adenylation, T: thiolation, MT: methyltransferase, Cterm: terminal condensation; (b) Structures of new methionine-containing RXPs (**1–7**) and two valine-only-containing main RXPs (**10** and **12**) with their bioactivity against protozoan parasites; (c) Solid-phase peptide synthesis (SPPS), shown for **1** as an example. I. attachment of the C-terminal amine phenylethylamine on DFPE resin; II. stepwise coupling of amino acids to peptide sequence using standard Fmoc chemistry mediated by different coupling reagents; III. cleavage of peptide from resin.

2.3 Production of a Photohexapeptide Library from Entomopathogenic *Photorhabdus asymbiotica* PB68.1

Authors:

<u>Lei Zhao</u>[1,2] and Helge B. Bode*[,1,3]

[1]Molecular Biotechnology, Department of Biosciences, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

[2]Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, 210014 Nanjing, China

[3]Buchmann Institute for Molecular Life Sciences (BMLS), Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

*Corresponding author

Online access:

https://pubs.rsc.org/ko/content/articlehtml/2019/ob/c9ob01489f

Attachments:

Declaration on the contribution of the authors and the publication including supporting information

Summary

Natural product (NP) compound libraries have been observed in various organisms, including bacteria, fungi, insects and plants. The production of such chemically diverse libraries often is the result of the catalytic promiscuity of specialized metabolic enzymes that synthesize compounds not immediately required for the survival of the host organisms, but nonetheless contribute to the reproductive success of populations in complex ecosystems.

NRPs are well known as an important source of therapeutic drugs. The assembly of these peptides is catalyzed by NRPSs that are composed of modules, each often consisting of three domains: condensation (C), adenylation (A) and thiolation (T) domains, to carry out one cycle of amino acid chain extension. Several modules act as a conveyor belt for the stepwise incorporation of amino acids usually in a collinear fashion. In rare cases, the high promiscuity of A domains and the iterative use of modules or domains result in the generation of peptide libraries.

This publication describes a new NRP compound library, photohexapeptide library, which was generated in *Photorhabdus asymbiotica* PB68.1 after the silent NRPS-encoding gene *phpS* was activated via promoter exchange. Peptide structures, including the absolute configurations of the amino acids, were determined by using a combination of bioinformatic analysis and isotopic labeling experiments followed by detailed HPLC-MS analysis. Additionally, their structures were confirmed by chemical synthesis and NMR after preparative isolation. The chemical diversity of the photohexapeptides results from promiscuous A domain specificity. Although relaxed substrate specificity of A domains has been shown in several other NRPS systems resulting in the generation of peptide diversity, PhpS is unusual since it possesses five out of six A domains which can utilize two different amino acids, leading to a very large number of products. Moreover, photohexapeptides enrich the family of the rare linear D-/L-peptide NPs. Although they did not show antibacterial and antifungal activity, further studies will pursue the biological functions of photohexapeptides, especially in the complex life cycle of *P. asymbiotica* PB68.1, including bacteria, nematode host and insect prey.
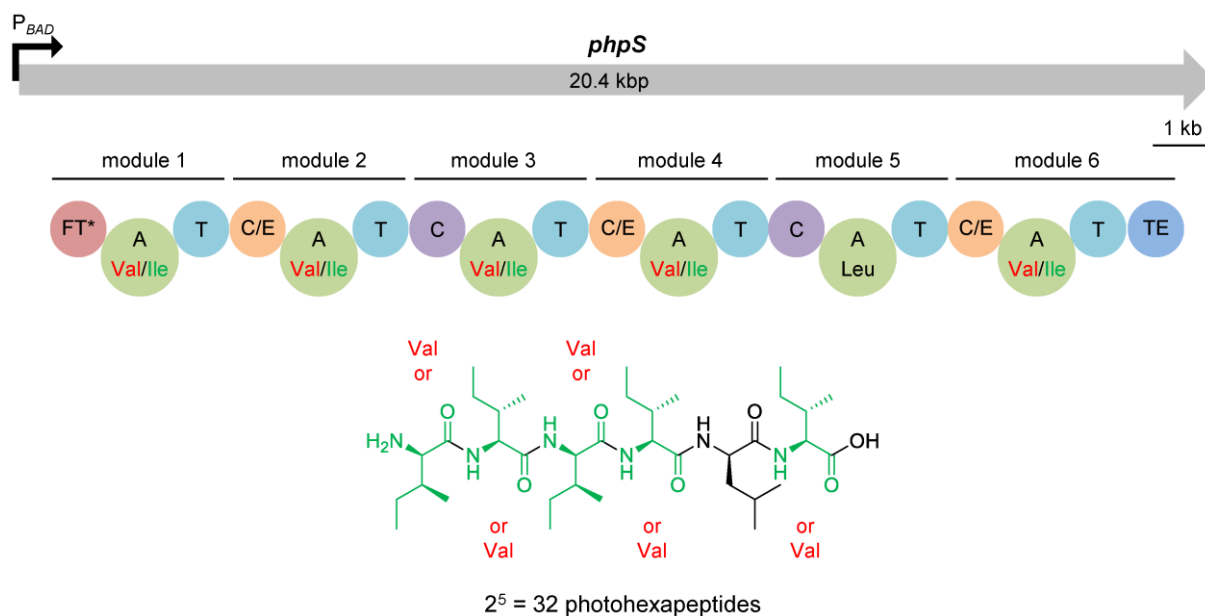
Figure 7. Graphical summary of the publication "Production of a Photohexapeptide Library from Entomopathogenic *Photorhabdus asymbiotica* PB68.1". The NRPS-encoding gene *phpS* with the domain organization corresponding to the production of photohexapeptide library. Domains: FT: formyltransferase (*nonfunctional), A: adenylation, T: thiolation, C/E: dual condensation/epimerization, C: condensation, TE: thioesterase. The chemical diversity of the photohexapeptides results from promiscuous A domain specificity that can utilize the two different amino acids valine (Val) or isoleucine (Ile).

45

2.4 Structure, Biosynthesis, and Bioactivity of Photoditritide from *Photorhabdus temperata* Meg1

Authors:

Lei Zhao,[1,2] Ryan Musumba Awori,[1] Marcel Kaiser,[3] Jonathan Groß,[4] Till Opatz,[4] and Helge B. Bode*[,1,5]

[1]Molecular Biotechnology, Department of Biosciences, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

[2]Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, 210014 Nanjing, China

[3]Swiss Tropical and Public Health Institute, 4051 Basel, Switzerland

[4]Institute of Organic Chemistry, Johannes Gutenberg University Mainz, 55128 Mainz, Germany

[5]Buchmann Institute for Molecular Life Sciences (BMLS), Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

*Corresponding author

Online access:

https://pubs.acs.org/doi/10.1021/acs.jnatprod.9b00932

Attachments:

Declaration on the contribution of the authors and the publication including supporting information

Summary

Bacteria of the genera *Xenorhabdus* and *Photorhabdus* are a rich source of bioactive NPs due to their unique niche. An effort to search for additional NPs from *Xenorhabdus* and *Photorhabdus* strains resulted in the identification of photoditritide. Photoditritide is a new cyclohexapeptide containing two uncommon homoarginine residues. This publication reports the discovery, structure elucidation, biosynthesis, and bioactivity of photoditritide.

Photoditritide was produced by *Photorhabdus temperata* Meg1 after the silent NRPS-encoding gene *pdtS* was activated via promoter exchange. Its structure was elucidated by HRMS, NMR and isotopic labeling experiments. The absolute configurations of the amino acids were determined according to the advanced Marfey's method in combination with bioinformatic analysis. Bioactivity testing of photoditritide revealed potent antimicrobial activity against *Micrococcus luteus* with an MIC value of 3.0 μM and weak antiprotozoal activity against *T. brucei rhodesiense* with an IC$_{50}$ value of 13 μM. Additionally, the biosynthetic pathway of photoditritide was proposed. Although production of the nonproteinogenic amino acid homoarginine-containing peptides have been reported in several marine organisms, photoditritide is so far the only example of a peptide from entomopathogenic bacteria that contains homoarginine.
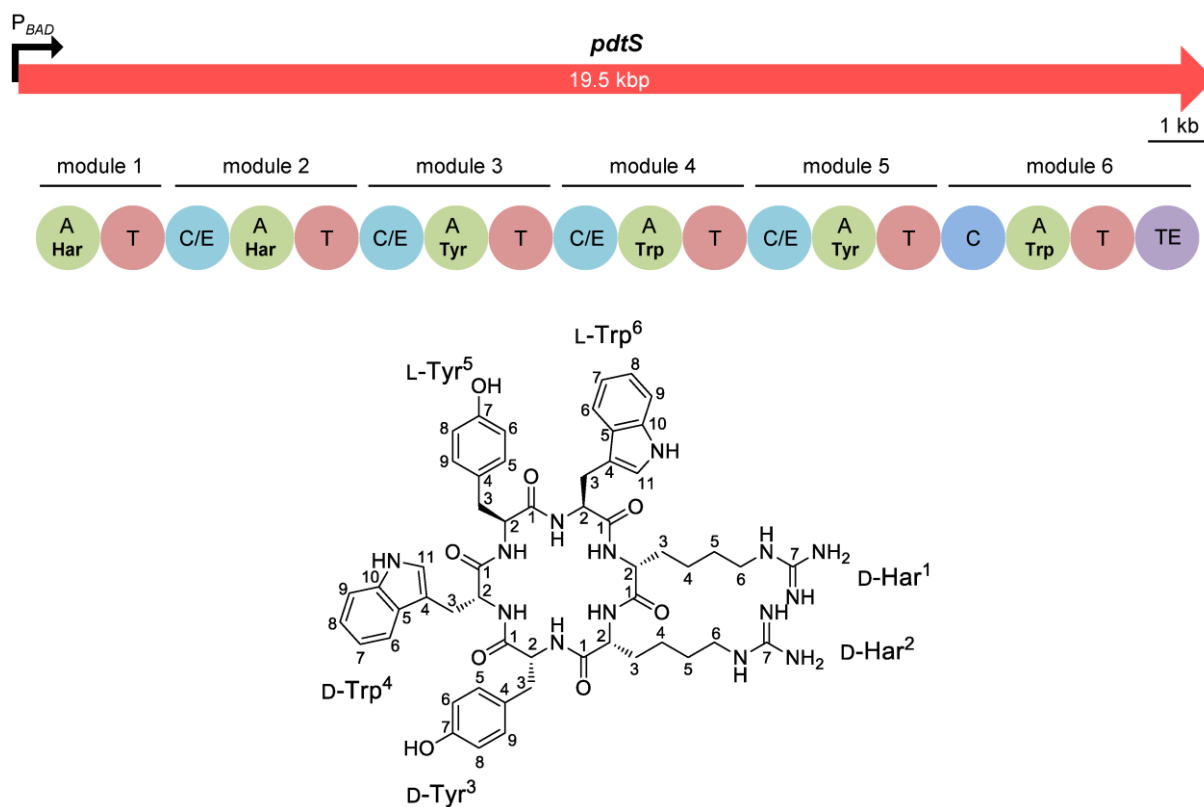
Figure 8. Graphical summary of the publication "Structure, Biosynthesis, and Bioactivity of Photoditritide from *Photorhabdus temperata* Meg1". The NRPS-encoding gene *pdtS* with the domain organization corresponding to the production of photoditritide. Har: homoarginine; Domains: A: adenylation, T: thiolation, C/E: dual condensation/epimerization, C: condensation, TE: thioesterase.

2.5 Phototemtide A, a Cyclic Lipopeptide Heterologously Expressed from *Photorhabdus temperata* Meg1, Shows Selective Antiprotozoal Activity

Authors:

Lei Zhao,[1,2] Tien Duy Vo,[1] Marcel Kaiser,[3] and Helge B. Bode*[,1,4]

[1]Molecular Biotechnology, Department of Biosciences, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

[2]Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, 210014 Nanjing, China

[3]Swiss Tropical and Public Health Institute, 4002 Basel, Switzerland

[4]Buchmann Institute for Molecular Life Sciences (BMLS), Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

*Corresponding author

Online access:

https://onlinelibrary.wiley.com/doi/abs/10.1002/cbic.201900665

Attachments:

Declaration on the contribution of the authors and the publication including supporting information

Summary

Cyclic lipopeptides (CLPs) are a class of structurally diverse NPs with various bioactivities, such as antibacterial, antifungal, antiprotozoal, and cytotoxic activities. They are mainly produced by a wide variety of microorganisms, including cyanobacteria, bacteria, actinobacteria, and fungi. Daptomycin, isolated from *Streptomyces roseoporus*, is the first clinically used CLP antibiotic with a new structural type and unique mechanism of action. As one of the few newly approved antibiotics, the recent success of daptomycin highlights the evolving role of CLPs as important pharmaceutical lead compounds. During our investigation to search for new NPs from *Xenorhabdus* and *Photorhabdus* strains, we found a new family of CLPs, named phototemtides. This publication describes the discovery, structure elucidation, biosynthesis, total synthesis, and bioactivity of the major compound phototemtide A. In addition, three minor derivatives, phototemtides B−D, were identified by detailed HPLC-MS analysis.

Phototemtides were identified from *P. temperata* Meg1 after the silent NRPS-encoding gene cluster *pttBC* was activated together with the MbtH-encoding gene *pttA* by heterologous expression in *E. coli*. The MbtH protein PttA was required for the biosynthesis of phototemtides in *E. coli*. The major compound phototemtide A was isolated and its structure was elucidated by HRMS and NMR experiments. The absolute configurations of the amino acids were determined by using the advanced Marfey's method in combination with bioinformatic analysis. The stereochemistry of 3-hydroxyoctanoic acid (3-HOA) was assigned by comparison of phototemtide A containing (*R*)- and (*S*)-3-HOA moiety, respectively, derived by total synthesis. The three minor derivatives phototemtides B−D were identified by a detailed analysis of their MS^2 fragmentation patterns. Phototemtide A showed weak antimalarial activity with an IC$_{50}$ value of 9.8 μM. Nevertheless, it might be a starting point towards a selective *P. falciparum* compound, since it shows no activity against any other tested organisms. With an efficient approach of total synthesis in hand, further investigation could focus on structure–activity relationships and subsequent in vivo experiments of this new family of CLPs. This work is also the first example of the importance of MbtH for peptide production in *Photorhabdus*.
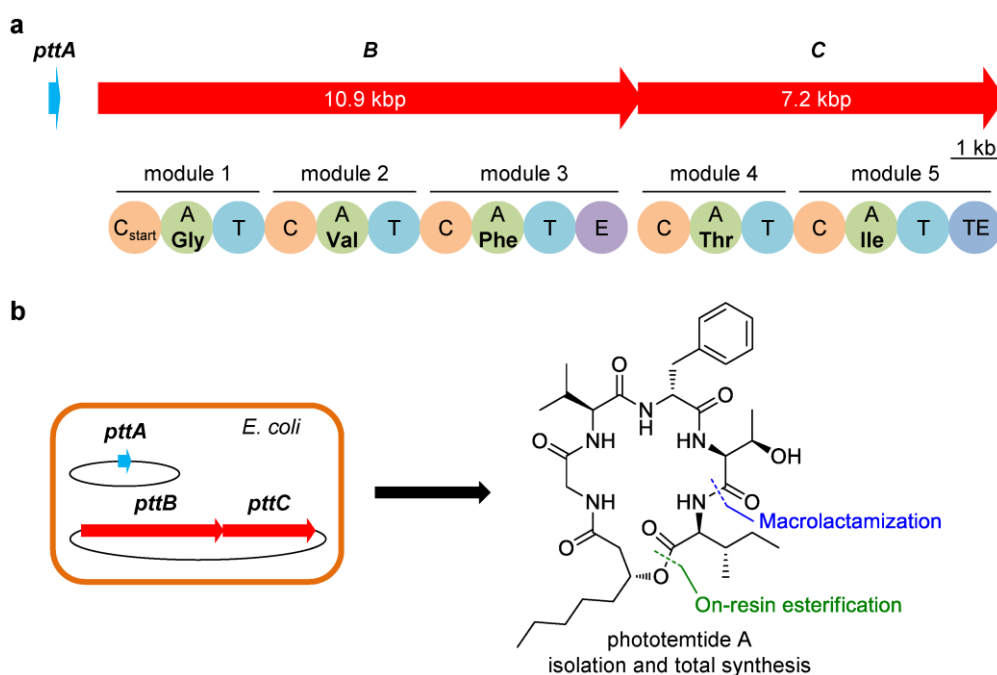
Figure 9. Graphical summary of the publication "Phototemtide A, a Cyclic Lipopeptide Heterologously Expressed from *Photorhabdus temperata* Meg1, Shows Selective Antiprotozoal Activity". (a) The NRPS-encoding gene cluster *pttABC* with the domain organization of PttBC. Domains: C_{start}: starter condensation, A: adenylation, T: thiolation, C: condensation, E: epimerization, TE: thioesterase; (b) Production of phototemtide A by heterologous expression of *pttABC* in *E. coli* and the key steps for total synthesis of phototemtide A.

3 Discussion and Outlook

By employing a combination of biological and chemical methods, four classes of nonribosomal peptides (NRPs) were unambiguously identified from *Xenorhabdus* and *Photorhabdus* strains. They comprise C-terminal amine containing linear peptides rhabdopeptide/xenortide-like peptides (RXPs, chapter 2.1 and 2.2), linear peptides photohexapeptides (chapter 2.3), cyclic peptide photoditritide (chapter 2.4), and cyclic lipopeptides (CLPs) phototemtides (chapter 2.5). In the following sections these peptides are discussed in some points regarding their biosynthesis, structure identification, and biological functions. Also, the future research work is proposed.

3.1 Rhabdopeptide/Xenortide-like Peptides

RXPs are a class of structurally diverse NRPs and widespread in *Xenorhabdus* and *Photorhabdus*.[34] The newly identified RXPs in this thesis from *X. innexi* and *Xenorhabdus* KJ12.1 further enrich the diversity of RXPs. They show polar putrescine or ammonia as the C-terminal amines and methionine as amino acid building blocks, which are distinctive characteristics in comparison with previously reported RXPs showing non-polar phenylethylamine or tryptamine as the main C-terminal amines and hydrophobic valine, phenylalanine or leucine as amino acid building blocks for most RXPs in most other strains.

Biosynthetically, both RXP biosynthesis gene clusters (BGCs) from *X. innexi* and *Xenorhabdus* KJ12.1 encode three monomodular nonribosomal peptide synthetases (NRPSs), namely, InxABC and Kj12ABC, respectively. However, against the collinearity rule, four to eight amino acids with optional *N*-methylation are present in these new RXPs, thus concluding an iterative use and a flexible methyltransferase activity of the NRPSs in these RXP biosynthesis. A recent study suggested that the high flexibility of RXP NRPSs might be a result of promiscuous docking domains mediating protein−protein interactions between the different NRPS subunits.[188] Meanwhile, the incorporation of putrescine or ammonia instead of most frequent phenylethylamine or tryptamine as the C-terminal amines of RXPs from *X. innexi* indicates a high flexibility of the C-terminal condensation (C_{term}) domains. The discovery of the methionine-containing RXPs after heterologous expression of the Kj12ABC from *Xenorhabdus* KJ12.1 in *E. coli* suggests that the adenylation (A) domain specificity might alter during RXP biosynthesis in a heterologous host. One

possibility might be the formation of protein complexes between the NRPSs and unknown proteins that alter the specificity of the A domains as described for MbtH proteins.[79,189] However, overexpression of MbtH from *E. coli* and *Xenorhabdus* KJ12.1 did not change the production levels of methionine-containing RXPs in the *E. coli* system, and therefore another yet unknown mechanism must be responsible for their production.

Functionally, these new RXP derivatives, particularly RXPs showing putrescine or ammonia as the C-terminal amines, exhibited strong effects against protozoan parasites. Considering the unique niche of *Xenorhabdus* and *Photorhabdus*, RXPs might play a role in protecting the insect cadaver from soil-living protozoa to support the symbiosis with nematodes.

Structurally, RXPs show different methylation patterns of the amide NH ranging from nonmethylation over partial methylation to permethylation. The synthetic strategy for nonmethylated and permethylated RXPs has been developed most recently by using a solid-phase peptide synthesis (SPPS) method.[190] Based on this strategy, an effective SPPS method for the synthesis of partially methylated RXPs was developed within this thesis. According to the structure−activity relationships of the RXPs, it is quite obvious that the terminal amines and amino acid building blocks have a strong influence on the bioactivity of RXPs (putrescine > ammonia, valine > methionine against *Trypanosoma brucei rhodesiense* and *Plasmodium falciparum*). With these efficient synthesis strategies in hand, it is possible to optimize RXPs in future with respect to their amino acid building blocks and terminal amines in view of their potent antiprotozoal activity for drug discovery, especially against the causative agents of sleeping sickness (*T. brucei rhodesiense*) and malaria (*P. falciparum*).

3.2 Photohexapeptides

A natural product (NP) compound library can be simply defined as a collection of a class of chemically diverse NPs derived from organisms. The generation of NP compound libraries is regarded as the biotic responses of organisms to challenging and changing environmental pressures.[191] Another explanation for such libraries might be against different targets in different organisms, which would be analogous to high-throughput screening procedures carried out during medical drug development.[192] In this thesis, the new NRP library of photohexapeptides, potentially containing up to 32 derivatives, was identified from *P. asymbiotica* PB68.1 pCEP-

phpS mutant. This library was produced by the single NRPS PhpS due to its promiscuous A domain specificity. The unusual PhpS possesses five (A1−A4 and A6), from in total six A domains, which can utilize two different amino acids valine or isoleucine, leading to the large number of products. The difference of the predicted A domain specificity between A5 domain for phenylalanine (actually incorporating its similar amino acid leucine) and the other five A domains all for valine (actually incorporating valine or its similar amino acid isoleucine) might explain the result. Thereby the chemical diversity of the photohexapeptides is an excellent example of how to create compound libraries in nature.

Moreover, the alternating C/E and C domains in PhpS places photohexapeptides among the family of the rare linear D-/L-peptide NPs. This class of peptides feature a propensity for β-helical structures in membranes and hence ion channels.[50,71] The well-studied example is the nonribosomally antibiotic gramicidin D, which requires two molecules to span a membrane.[193]

It is worth noting that the PhpS contains a formyltransferase (FT) domain as the first domain of the initiation module. However, no formylated products were found in photohexapeptides. During our in-depth investigation on *phpS*, we found a formamidase-encoding gene (*ppb6_04102*) located 1.3 kbp upstream from *phpS*. Thus we postulated that the formyl group might be removed by the formamidase during the biosynthesis of photohexapeptides. To test the hypothesis, the PhpS alone was expressed in *E. coli* without Ppb6_04102, but still no formylated products except photohexapeptides were detected. This suggests that the FT domain might be inactive in PhpS. As FT domains often contain a conserved Asn-His-Asp catalytic triad,[194] we compared the FT domain of the PhpS with previously reported active FT domains, but found PhpS to have all three described key catalytic residues. From sequence alignment, also all residues required for binding of N^{10}-formyltetrahydrofolate (fTHF, as the formyl donor) are present, suggesting some other reason why this FT domain is inactive.

Although the major photohexapeptide did not show antibacterial and antifungal activity, it displayed a moderate antiprotozoal activity against *P. falciparum* (IC_{50} 7.9 μM, unpublished data) according to the most recent bioactivity testing. Considering the importance of NP compound libraries for the host organisms, further studies will

pursue the biological functions of the photohexapeptide library, especially in the complex life cycle of bacteria, nematode and insect.

3.3 Photoditritide

Photoditritide is a new cyclic peptide identified from *P. temperata* Meg1 after the NRPS-encoding gene *pdtS* was activated via promoter exchange. The most prominent structure feature is that it contains two nonproteinogenic amino acid homoarginine (Har) residues. The incorporation of nonproteinogenic amino acids by NRPSs is an important contribution to the structural diversity of NRPs in comparison with peptides produced by ribosomal synthesis which often contain only the 20 proteinogenic amino acids.

Har is a rare amino acid. Therefore, Har-containing peptides are uncommon in microorganisms, except for the occurrence in several marine organisms such as sponges,[195] cyanobacteria,[196] and marinederived actinomycetes.[197] Biosynthesis of Har has been elucidated in *P. luminescens* recently (unpublished data). The gene *plu0158* encodes a putative amidinotransferase, which has 68% identity to the amidinotransferase of *Pseudomonas syringae phaseolicola* that catalyzes the transfer of a guanidinium group from arginine to lysine to produce Har.[198] Its homologous gene (*MEG1_RS12600*) can also be found in *P. temperata* Meg1. Thereby, the gene encoding protein MEG1_RS12600 might be responsible for Har biosynthesis in *P. temperata* Meg1, which can be tested by promoter exchange and feeding experiments in future.

Bioactivity testing of photoditritide revealed potent antimicrobial activity against *Micrococcus luteus*, suggesting that photoditritide can protect nutrient-rich insect cadaver from food competitor bacteria in the complex life cycle of nematode and bacteria.

3.4 Phototemtides

CLPs are a class of chemically diverse NPs and can act as important pharmaceutical lead compounds.[116,199] Phototemtides, a new family of CLPs, were heterologously expressed from *P. temperata* Meg1. The identification of phototemtides is a great methodological example comprising all the approaches used for NRP discovery in this thesis, including bioinformatic analysis, activation of silent NRPS gene clusters, HPLC-MS analysis, feeding experiments, NMR analysis, the advanced Marfey's

method, and total synthesis. The structures of phototemtides were unequivocally elucidated by the combined use of all these methods.

Bioactivity testing showed a selective antiparasitic activity of the major compound phototemtide A against the causative agent of malaria *P. falciparum*, but no activity against any other tested organisms, including protozoan parasites (*T. brucei rhodesiense*, *T. cruzi*, and *Leishmania donovani*), Gram-positive bacteria (*M. luteus*), fungi (*Saccharomyces cerevisiae*), and mammalian L6 cells. Thereby it might be a starting point toward a compound that shows selective bioactivity against *P. falciparum*, although the antimalarial activity of phototemtide A is only weak. With an efficient approach of total synthesis in hand, future work could focus on structure–activity relationship studies and subsequent in vivo experiments of this new family of CLPs.

NRPS engineering to create new peptides especially where chemical synthesis fails entirely is of great interest. Based on NRPSs from *Xenorhabdus* and *Photorhabdus*, two new strategies for NRPS engineering that use exchange unit (XU) concept[200] and exchange unit condensation domain (XUC) concept[201] have been developed recently. These newly identified NRPSs in this thesis will also provide a new source for future engineering purposes, particularly for unusual building block containing peptides, cyclic peptides, or cyclic lipopeptides.

4 References

(1) Thomas, G. M.; Poinar Jr, G. O. *Xenorhabdus* gen. nov., a Genus of Entomopathogenic, Nematophilic Bacteria of the Family Enterobacteriaceae. *Int. J. Syst. Evol. Microbiol.* **1979**, *29*, 352−360.

(2) Fischer-Le Saux, M.; Viallard, V.; Brunel, B.; Normand, P.; Boemare, N. E. Polyphasic Classification of the Genus *Photorhabdus* and Proposal of New Taxa: *P. luminescens* subsp. *luminescens* subsp. nov., *P. luminescens* subsp. *akhurstii* subsp. nov., *P. luminescens* subsp. *laumondii* subsp. nov., *P. temperata* sp. nov., *P. temperata* subsp. *temperata* subsp. nov. and *P. asymbiotica* sp. nov. *Int. J. Syst. Evol. Microbiol.* **1999**, *49*, 1645−1656.

(3) Thanwisai, A.; Tandhavanant, S.; Saiprom, N.; Waterfield, N. R.; Long, P. K.; Bode, H. B.; Peacock, S. J.; Chantratita, N. Diversity of *Xenorhabdus* and *Photorhabdus* spp. and Their Symbiotic Entomopathogenic Nematodes from Thailand. *PLoS One* **2012**, *7*, e43835.

(4) Ehlers, R. U. Entomopathogenic Nematodes in the European Biocontrol Market. *Commun. Agric. Appl. Biol. Sci.* **2003**, *68,* 3−16.

(5) Lacey, L. A.; Grzywacz, D.; Shapiro-Ilan, D. I.; Frutos, R.; Brownbridge, M.; Goettel, M. S. Insect Pathogens as Biological Control Agents: Back to the Future. *J. Invertebr. Pathol.* **2015**, *132*, 1−41.

(6) Ciche, T. A.; Darby, C.; Ehlers, R.-U.; Forst, S.; Goodrich-Blair, H. Dangerous Liaisons: The Symbiosis of Entomopathogenic Nematodes and Bacteria. *Biol. Control* **2006**, *38*, 22−46.

(7) Bennett, H. P. J.; Clarke, D. J. The PbgPE Operon in *Photorhabdus luminescens* Is Required for Pathogenicity and Symbiosis. *J. Bacteriol.* **2005**, *187,* 77−84.

(8) Forst, S.; Dowds, B.; Boemare, N.; Stackebrandt, E. *Xenorhabdus* and *Photorhabdus* spp.: Bugs That Kill Bugs. *Annu. Rev. Microbiol.* **1997**, *51*, 47−72.

(9) Waterfield, N. R.; Ciche, T.; Clarke, D. *Photorhabdus* and a Host of Hosts. *Annu. Rev. Microbiol.* **2009**, *63,* 557−574.

(10) Clarke, D. J. *Photorhabdus*: A Model for the Analysis of Pathogenicity and

Mutualism. *Cell. Microbiol.* **2008**, *10*, 2159−2167.

(11) Poinar Jr, G. O.; Thomas, G.; Haygood, M.; Nealson, K. H. Growth and Luminescence of the Symbolic Bacteria Associated with the Terrestrial Nematode, *Heterorphabditis bacteriophora. Soil Biol. Biochem.* **1980**, *12*, 5−10.

(12) Bode, E. Structure, Function and Biosynthesis of Natural Products from *Xenorhabdus doucetiae*, Goethe University Frankfurt, 2017.

(13) ffrench-Constant R. H.; Dowling, A.; Waterfield, N. R. Insecticidal Toxins from *Photorhabdus* Bacteria and Their Potential Use in Agriculture. *Toxicon* **2007**, *49*, 436−451.

(14) Vizcaino, M. I.; Guo, X.; Crawford, J. M. Merging Chemical Ecology with Bacterial Genome Mining for Secondary Metabolite Discovery. *J. Ind. Microbiol. Biotechnol.* **2014**, *41*, 285−299.

(15) Shi, Y.-M.; Bode, H. B. Chemical Language and Warfare of Bacterial Natural Products in Bacteria−Nematode−Insect Interactions. *Nat. Prod. Rep.* **2018**, *35*, 309−335.

(16) Zhou, X.; Kaya, H. K.; Heungens, K.; Goodrich-Blair, H. Response of Ants to a Deterrent Factor(s) Produced by the Symbiotic Bacteria of Entomopathogenic Nematodes. *Appl. Environ. Microbiol.* **2002**, *68*, 6202−6209.

(17) Goodrich-Blair, H.; Clarke, D. J. Mutualism and Pathogenesis in *Xenorhabdus* and *Photorhabdus*: Two Roads to the Same Destination. *Mol. Microbiol.* **2007**, *64*, 260−268.

(18) Bode, H. B. Entomopathogenic Bacteria as a Source of Secondary Metabolites. *Curr. Opin. Chem. Biol.* **2009**, *13*, 224−230.

(19) Poinar Jr, G. O.; Grewal, P. S. History of Entomopathogenic Nematology. *J. Nematol.* **2012**, *44*, 153−161.

(20) Casanova-Torres, Á.; Goodrich-Blair, H. Immune Signaling and Antimicrobial Peptide Expression in Lepidoptera. *Insects* **2013**, *4*, 320−338.

(21) Cerenius, L.; Lee, B. L.; Söderhäll, K. The ProPO-System: Pros and Cons for Its Role in Invertebrate Immunity. *Trends Immunol.* **2008**, *29*, 263−271.

(22) Eleftherianos, I.; Revenis, C. Role and Importance of Phenoloxidase in Insect Hemostasis. *J. Innate Immun.* **2011**, *3*, 28−33.

(23)   Tobias, N. J.; Wolff, H.; Djahanschiri, B.; Grundmann, F.; Kronenwerth, M.; Shi, Y.-M.; Simonyi, S.; Grün, P.; Shapiro-Ilan, D.; Pidot, S. J.; Stinear, T. P.; Ebersberger, I.; Bode, H. B. Natural Product Diversity Associated with the Nematode Symbionts *Photorhabdus* and *Xenorhabdus*. *Nat. Microbiol.* **2017**, *2*, 1676−1685.

(24)   Crawford, J. M.; Portmann, C.; Zhang, X.; Roeffaers, M. B. J.; Clardy, J. Small Molecule Perimeter Defense in Entomopathogenic Bacteria. *Proc. Natl. Acad. Sci.* **2012**, *109*, 10821−10826.

(25)   Hirschmann, M.; Grundmann, F.; Bode, H. B. Identification and Occurrence of the Hydroxamate Siderophores Aerobactin, Putrebactin, Avaroferrin and Ochrobactin C as Virulence Factors from Entomopathogenic Bacteria. *Environ. Microbiol.* **2017**, *19*, 4080−4090.

(26)   Griffith, D. M.; Szőcs, B.; Keogh, T.; Suponitsky, K. Y.; Farkas, E.; Buglyó, P.; Marmion, C. J. Suberoylanilide Hydroxamic Acid, a Potent Histone Deacetylase Inhibitor; Its X-Ray Crystal Structure and Solid State and Solution Studies of Its Zn (II), Ni (II), Cu (II) and Fe (III) Complexes. *J. Inorg. Biochem.* **2011**, *105*, 763−769.

(27)   Mukherjee, K.; Fischer, R.; Vilcinskas, A. Histone Acetylation Mediates Epigenetic Regulation of Transcriptional Reprogramming in Insects during Metamorphosis, Wounding and Infection. *Front. Zool.* **2012**, *9*, 25.

(28)   Tobias, N. J.; Shi, Y.-M.; Bode, H. B. Refining the Natural Product Repertoire in Entomopathogenic Bacteria. *Trends Microbiol.* **2018**, *26*, 833−840.

(29)   Paul, V. J.; Frautschy, S.; Fenical, W.; Nealson, K. H. Antibiotics in Microbial Ecology: Isolation and Structure Assignment of Several New Antibacterial Compounds from the Insect-Symbiotic Bacteria *Xenorhabdus* Spp. *J. Chem. Ecol.* **1981**, *7*, 589−597.

(30)   Proschak, A.; Zhou, Q.; Schöner, T.; Thanwisai, A.; Kresovic, D.; Dowling, A.; Ffrench-Constant, R.; Proschak, E.; Bode, H. B. Biosynthesis of the Insecticidal Xenocyloins in *Xenorhabdus bovienii*. *Chembiochem* **2014**, *15*, 369−372.

(31)   Brachmann, A. O.; Brameyer, S.; Kresovic, D.; Hitkova, I.; Kopp, Y.; Manske, C.; Schubert, K.; Bode, H. B.; Heermann, R. Pyrones as Bacterial Signaling Molecules. *Nat. Chem. Biol.* **2013**, *9*, 573−578.

References

(32) Kresovic, D.; Schempp, F.; Cheikh-Ali, Z.; Bode, H. B. A Novel and Widespread Class of Ketosynthase Is Responsible for the Head-to-Head Condensation of Two Acyl Moieties in Bacterial Pyrone Biosynthesis. *Beilstein J. Org. Chem.* **2015**, *11*, 1412−1417.

(33) Fuchs, S. W.; Grundmann, F.; Kurz, M.; Kaiser, M.; Bode, H. B. Fabclavines: Bioactive Peptide−Polyketide−Polyamino Hybrids from *Xenorhabdus.* *ChemBioChem* **2014**, *15*, 512−516.

(34) Cai, X.; Nowak, S.; Wesche, F.; Bischoff, I.; Kaiser, M.; Fürst, R.; Bode, H. B. Entomopathogenic Bacteria Use Multiple Mechanisms for Bioactive Peptide Library Design. *Nat. Chem.* **2017**, *9*, 379−386.

(35) Reimer, D.; Cowles, K. N.; Proschak, A.; Nollmann, F. I.; Dowling, A. J.; Kaiser, M.; ffrench-Constant, R.; Goodrich-Blair, H.; Bode, H. B. Rhabdopeptides as Insect-specific Virulence Factors from Entomopathogenic Bacteria. *ChemBioChem* **2013**, *14*, 1991−1997.

(36) Reimer, D.; Nollmann, F. I.; Schultz, K.; Kaiser, M.; Bode, H. B. Xenortide Biosynthesis by Entomopathogenic *Xenorhabdus nematophila. J. Nat. Prod.* **2014**, *77*, 1976−1980.

(37) Joyce, S. A.; Brachmann, A. O.; Glazer, I.; Lango, L.; Schwär, G.; Clarke, D. J.; Bode, H. B. Bacterial Biosynthesis of a Multipotent Stilbene. *Angew. Chemie Int. Ed.* **2008**, *47*, 1942−1945.

(38) Richardson, W. H.; Schmidt, T. M.; Nealson, K. H. Identification of an Anthraquinone Pigment and a Hydroxystilbene Antibiotic from *Xenorhabdus luminescens. Appl. Environ. Microbiol.* **1988**, *54*, 1602−1605.

(39) Li, J.; Chen, G.; Wu, H.; Webster, J. M. Identification of Two Pigments and a Hydroxystilbene Antibiotic from *Photorhabdus luminescens. Appl. Environ. Microbiol.* **1995**, *61*, 4329−4333.

(40) Eleftherianos, I.; Boundy, S.; Joyce, S. A.; Aslam, S.; Marshall, J. W.; Cox, R. J.; Simpson, T. J.; Clarke, D. J.; Reynolds, S. E. An Antibiotic Produced by an Insect-Pathogenic Bacterium Suppresses Host Defenses through Phenoloxidase Inhibition. *Proc. Natl. Acad. Sci.* **2007**, *104*, 2419−2424.

(41) Schimming, O.; Fleischhacker, F.; Nollmann, F. I.; Bode, H. B. Yeast

Homologous Recombination Cloning Leading to the Novel Peptides Ambactin and Xenolindicin. *Chembiochem* **2014**, *15*, 1290−1294.

(42)   Brachmann, A. O.; Joyce, S. A.; Jenke-Kodama, H.; Schwär, G.; Clarke, D. J.; Bode, H. B. A Type II Polyketide Synthase Is Responsible for Anthraquinone Biosynthesis in *Photorhabdus luminescens*. *Chembiochem* **2007**, *8*, 1721−1728.

(43)   Ji, D.; Yi, Y.; Kang, G.-H.; Choi, Y.-H.; Kim, P.; Baek, N.-I.; Kim, Y. Identification of an Antibacterial Compound, Benzylideneacetone, from *Xenorhabdus nematophila* against Major Plant-Pathogenic Bacteria. *FEMS Microbiol. Lett.* **2004**, *239*, 241−248.

(44)   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.* **2012**, *84*, 6948−6955.

(45)   Grundmann, F.; Kaiser, M.; Schiell, M.; Batzer, A.; Kurz, M.; Thanwisai, A.; Chantratita, N.; Bode, H. B. Antiparasitic Chaiyaphumines from Entomopathogenic *Xenorhabdus* sp. PB61. 4. *J. Nat. Prod.* **2014**, *77*, 779−783.

(46)   Crawford, J. M.; Mahlstedt, S. A.; Malcolmson, S. J.; Clardy, J.; Walsh, C. T. Dihydrophenylalanine: A Prephenate-Derived *Photorhabdus luminescens* Antibiotic and Intermediate in Dihydrostilbene Biosynthesis. *Chem. Biol.* **2011**, *18*, 1102−1112.

(47)   Bode, H. B.; Reimer, D.; Fuchs, S. W.; Kirchner, F.; Dauth, C.; Kegler, C.; Lorenzen, W.; Brachmann, A. O.; Grün, P. Determination of the Absolute Configuration of Peptide Natural Products by Using Stable Isotope Labeling and Mass Spectrometry. *Chem. Eur. J.* **2012**, *18*, 2342−2348.

(48)   Theodore, C. M.; King, J. B.; You, J.; Cichewicz, R. H. Production of Cytotoxic Glidobactins/Luminmycins by *Photorhabdus asymbiotica* in Liquid Media and Live Crickets. *J. Nat. Prod.* **2012**, *75*, 2007−2011.

(49)   Brachmann, A. O.; Kirchner, F.; Kegler, C.; Kinski, S. C.; Schmitt, I.; Bode, H. B. Triggering the Production of the Cryptic Blue Pigment Indigoidine from *Photorhabdus luminescens. J. Biotechnol.* **2012**, *157*, 96−99.

(50) Bode, H. B.; Brachmann, A. O.; Jadhav, K. B.; Seyfarth, L.; Dauth, C.; Fuchs, S. W.; Kaiser, M.; Waterfield, N. R.; Sack, H.; Heinemann, S. H. Structure Elucidation and Activity of Kolossin A, the D-/L-Pentadecapeptide Product of a Giant Nonribosomal Peptide Synthetase. *Angew. Chemie Int. Ed.* **2015**, *54*, 10352−10355.

(51) Park, H. B.; Crawford, J. M. Lumiquinone A, an α-Aminomalonate-Derived Aminobenzoquinone from *Photorhabdus luminescens*. *J. Nat. Prod.* **2015**, *78*, 1437−1441.

(52) Park, H. B.; Crawford, J. M. Pyrazinone Protease Inhibitor Metabolites from *Photorhabdus luminescens*. *J. Antibiot.* **2016**, *69*, 616−621.

(53) Bode, E.; Brachmann, A. O.; Kegler, C.; Simsek, R.; Dauth, C.; Zhou, Q.; Kaiser, M.; Klemmt, P.; Bode, H. B. Simple "On-demand" Production of Bioactive Natural Products. *ChemBioChem* **2015**, *16*, 1115−1119.

(54) Pantel, L.; Florin, T.; Dobosz-Bartoszek, M.; Racine, E.; Sarciaux, M.; Serri, M.; Houard, J.; Campagne, J.-M.; De Figueiredo, R. M.; Midrier, C.; Gaudriault, S.; Givaudan, A.; Lanois, A.; Forst, S.; Aumelas, A.; Cotteaux-Lautard, C.; Bolla, J.-M.; Lundberg, C. V.; Huseby, D. L.; Hughes, D.; Villain-Guillot, P.; Mankin, A. S.; Polikanov, Y. S.; Gualtieri, M. Odilorhabdins, Antibacterial Agents That Cause Miscoding by Binding at a New Ribosomal Site. *Mol. Cell* **2018**, *70*, 83−94.

(55) Eom, S.; Park, Y.; Kim, Y. Sequential Immunosuppressive Activities of Bacterial Secondary Metabolites from the Entomopahogenic Bacterium *Xenorhabdus nematophila*. *J. Microbiol.* **2014**, *52*, 161−168.

(56) Gualtieri, M.; Aumelas, A.; Thaler, J.-O. Identification of a New Antimicrobial Lysine-Rich Cyclolipopeptide Family from *Xenorhabdus nematophila*. *J. Antibiot.* **2009**, *62*, 295−302.

(57) Fuchs, S. W.; Proschak, A.; Jaskolla, T. W.; Karas, M.; Bode, H. B. Structure Elucidation and Biosynthesis of Lysine-Rich Cyclic Peptides in *Xenorhabdus nematophila*. *Org. Biomol. Chem.* **2011**, *9*, 3130−3132.

(58) Park, H. B.; Perez, C. E.; Barber, K. W.; Rinehart, J.; Crawford, J. M. Genome Mining Unearths a Hybrid Nonribosomal Peptide Synthetase-like-Pteridine Synthase Biosynthetic Gene Cluster. *Elife* **2017**, *6*, e25229.

(59) Shi, Y.-M.; Brachmann, A. O.; Westphalen, M. A.; Neubacher, N.; Tobias, N. J.; Bode, H. B. Dual Phenazine Gene Clusters Enable Diversification during Biosynthesis. *Nat. Chem. Biol.* **2019**, *15*, 331−339.

(60) Perez, C. E.; Crawford, J. M. Characterization of a Hybrid Nonribosomal Peptide−Carbohydrate Biosynthetic Pathway in *Photorhabdus luminescens*. *Biochemistry* **2019**, *58*, 1131−1140.

(61) Nollmann, F. I.; Heinrich, A. K.; Brachmann, A. O.; Morisseau, C.; Mukherjee, K.; Casanova-Torres, Á. M.; Strobl, F.; Kleinhans, D.; Kinski, S.; Schultz, K.; Beeton, M. L.; Kaiser, M.; Chu, Y. Y.; Ke, L. P.; Thanwisai, A.; Bozhüyük, K. A. J.; Chantratita, N.; Götz, F.; Waterfield, N. R.; Vilcinskas, A.; Stelzer, E. H. K.; Goodrich-Blair, H.; Hammock, B. D.; Bode, H. B. A *Photorhabdus* Natural Product Inhibits Insect Juvenile Hormone Epoxide Hydrolase. *ChemBioChem* **2015**, *16*, 766−771.

(62) Schimming, O.; Challinor, V. L.; Tobias, N. J.; Adihou, H.; Grün, P.; Pöschel, L.; Richter, C.; Schwalbe, H.; Bode, H. B. Structure, Biosynthesis, and Occurrence of Bacterial Pyrrolizidine Alkaloids. *Angew. Chemie Int. Ed.* **2015**, *54*, 12702−12705.

(63) Lang, G.; Kalvelage, T.; Peters, A.; Wiese, J.; Imhoff, J. F. Linear and Cyclic Peptides from the Entomopathogenic Bacterium *Xenorhabdus nematophilus*. *J. Nat. Prod.* **2008**, *71*, 1074−1077.

(64) Oh, J.; Kim, N. Y.; Chen, H.; Palm, N. W.; Crawford, J. M. An Ugi-like Biosynthetic Pathway Encodes Bombesin Receptor Subtype-3 Agonists. *J. Am. Chem. Soc.* **2019**, *141*, 16271−16278.

(65) Paik, S.; Park, Y. H.; Suh, S. II; Kim, H. S.; Lee, I. S.; Park, M. K.; Lee, C. S.; Park, S. H. Unusual Cytotoxic Phenethylamides from *Xenorhabdus nematophilus*. *Bull. Korean Chem. Soc.* **2001**, *22*, 372−374.

(66) Bode, E.; He, Y.; Vo, T. D.; Schultz, R.; Kaiser, M.; Bode, H. B. Biosynthesis and Function of Simple Amides in *Xenorhabdus doucetiae*. *Environ. Microbiol.* **2017**, *19*, 4564−4575.

(67) Ohlendorf, B.; Simon, S.; Wiese, J.; Imhoff, J. F. Szentiamide, an *N*-Formylated Cyclic Depsipeptide from *Xenorhabdus szentirmaii* DSM 16338T. *Nat. Prod. Commun.* **2011**, *6*, 1247−1250.

(68) Kronenwerth, M.; Bozhüyük, K. A. J.; Kahnt, A. S.; Steinhilber, D.; Gaudriault, S.; Kaiser, M.; Bode, H. B. Characterisation of Taxllaids A−G; Natural Products from *Xenorhabdus indica*. *Chem. Eur. J.* **2014**, *20*, 17478−17487.

(69) Wolff, H.; Bode, H. B. The Benzodiazepine-like Natural Product Tilivalline Is Produced by the Entomopathogenic Bacterium *Xenorhabdus eapokensis*. *PLoS One* **2018**, *13*, e0194297.

(70) Crawford, J. M.; Portmann, C.; Kontnik, R.; Walsh, C. T.; Clardy, J. NRPS Substrate Promiscuity Diversifies the Xenematides. *Org. Lett.* **2011**, *13*, 5144−5147.

(71) Zhou, Q.; Grundmann, F.; Kaiser, M.; Schiell, M.; Gaudriault, S.; Batzer, A.; Kurz, M.; Bode, H. B. Structure and Biosynthesis of Xenoamicins from Entomopathogenic *Xenorhabdus*. *Chem. Eur. J.* **2013**, *19*, 16772−16779.

(72) Grundmann, F.; Kaiser, M.; Kurz, M.; Schiell, M.; Batzer, A.; Bode, H. B. Structure Determination of the Bioactive Depsipeptide Xenobactin from *Xenorhabdus* sp. PB30. 3. *RSC Adv.* **2013**, *3*, 22072−22077.

(73) McInerney, B. V; Taylor, W. C.; Lacey, M. J.; Akhurst, R. J.; Gregson, R. P. Biologically Active Metabolites from *Xenorhabdus* spp., Part 2. Benzopyran-1-One Derivatives with Gastroprotective Activity. *J. Nat. Prod.* **1991**, *54*, 785−795.

(74) Park, D.; Ciezki, K.; Van Der Hoeven, R.; Singh, S.; Reimer, D.; Bode, H. B.; Forst, S. Genetic Analysis of Xenocoumacin Antibiotic Production in the Mutualistic Bacterium *Xenorhabdus nematophila*. *Mol. Microbiol.* **2009**, *73*, 938−949.

(75) Brachmann, A. O.; Forst, S.; Furgani, G. M.; Fodor, A.; Bode, H. B. Xenofuranones A and B: Phenylpyruvate Dimers from *Xenorhabdus szentirmaii*. *J. Nat. Prod.* **2006**, *69*, 1830−1832.

(76) McInerney, B. V; Gregson, R. P.; Lacey, M. J.; Akhurst, R. J.; Lyons, G. R.; Rhodes, S. H.; Smith, D. R. J.; Engelhardt, L. M.; White, A. H. Biologically Active Metabolites from *Xenorhabdus* spp., Part 1. Dithiolopyrrolone Derivatives with Antibiotic Activity. *J. Nat. Prod.* **1991**, *54*, 774−784.

(77) Kegler, C.; Nollmann, F. I.; Ahrendt, T.; Fleischhacker, F.; Bode, E.; Bode, H.

B. Rapid Determination of the Amino Acid Configuration of Xenotetrapeptide. *Chembiochem* **2014**, *15*, 826−828.

(78) Zhou, Q.; Dowling, A.; Heide, H.; Wöhnert, J.; Brandt, U.; Baum, J.; Ffrench-Constant, R.; Bode, H. B. Xentrivalpeptides A−Q: Depsipeptide Diversification in *Xenorhabdus*. *J. Nat. Prod.* **2012**, *75*, 1717−1722.

(79) Süssmuth, R. D.; Mainz, A. Nonribosomal Peptide Synthesis—Principles and Prospects. *Angew. Chemie Int. Ed.* **2017**, *56*, 3770−3821.

(80) Sieber, S. A.; Marahiel, M. A. Molecular Mechanisms Underlying Nonribosomal Peptide Synthesis: Approaches to New Antibiotics. *Chem. Rev.* **2005**, *105*, 715−738.

(81) Mootz, H. D.; Schwarzer, D.; Marahiel, M. A. Ways of Assembling Complex Natural Products on Modular Nonribosomal Peptide Synthetases. *Chembiochem* **2002**, *3*, 490−504.

(82) Mach, B.; Reich, E.; Tatum, E. L. Separation of the Biosynthesis of the Antibiotic Polypeptide Tyrocidine from Protein Biosynthesis. *Proc. Natl. Acad. Sci.* **1963**, *50*, 175−181.

(83) Lipmann, F. Bacterial Production of Antibiotic Polypeptides by Thiol-Linked Synthesis on Protein Templates. *Adv. Microb. Physiol.* **1981**, *21*, 227−266.

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

(85) Gulick, A. M. Conformational Dynamics in the Acyl-CoA Synthetases, Adenylation Domains of Non-ribosomal Peptide Synthetases, and Firefly Luciferase. *ACS Chem. Biol.* **2009**, *4*, 811−827.

(86) Bozhüyük, K. A. J. Reprogramming Non-ribosomal Peptide Synthetases, Goethe University Frankfurt, 2016.

(87) Conti, E.; Stachelhaus, T.; Marahiel, M. A.; Brick, P. Structural Basis for the Activation of Phenylalanine in the Non-ribosomal Biosynthesis of Gramicidin S. *EMBO J.* **1997**, *16*, 4174−4183.

(88) Lee, T. V.; Johnson, L. J.; Johnson, R. D.; Koulman, A.; Lane, G. A.; Lott, J. S.; Arcus, V. L. Structure of a Eukaryotic Nonribosomal Peptide Synthetase

Adenylation Domain That Activates a Large Hydroxamate Amino Acid in Siderophore Biosynthesis. *J. Biol. Chem.* **2010**, *285*, 2415−2427.

(89) Challis, G. L.; Ravel, J.; Townsend, C. A. Predictive, Structure-Based Model of Amino Acid Recognition by Nonribosomal Peptide Synthetase Adenylation Domains. *Chem. Biol.* **2000**, *7*, 211−224.

(90) Skinnider, M. A.; Dejong, C. A.; Rees, P. N.; Johnston, C. W.; Li, H.; Webster, A. L. H.; Wyatt, M. A.; Magarvey, N. A. Genomes to Natural Products Prediction Informatics for Secondary Metabolomes (PRISM). *Nucleic Acids Res.* **2015**, *43*, 9645−9662.

(91) Blin, K.; Shaw, S.; Steinke, K.; Villebro, R.; Ziemert, N.; Lee, S. Y.; Medema, M. H.; Weber, T. AntiSMASH 5.0: Updates to the Secondary Metabolite Genome Mining Pipeline. *Nucleic Acids Res.* **2019**, *47*, W81−W87.

(92) Röttig, M.; Medema, M. H.; Blin, K.; Weber, T.; Rausch, C.; Kohlbacher, O. NRPSpredictor2—a Web Server for Predicting NRPS Adenylation Domain Specificity. *Nucleic Acids Res.* **2011**, *39*, W362−W367.

(93) Meyer, S.; Kehr, J.-C.; Mainz, A.; Dehm, D.; Petras, D.; Süssmuth, R. D.; Dittmann, E. Biochemical Dissection of the Natural Diversification of Microcystin Provides Lessons for Synthetic Biology of NRPS. *Cell Chem. Biol.* **2016**, *23*, 462−471.

(94) Mudalungu, C. M.; von Törne, W. J.; Voigt, K.; Rückert, C.; Schmitz, S.; Sekurova, O. N.; Zotchev, S. B.; Süssmuth, R. D. Noursamycins, Chlorinated Cyclohexapeptides Identified from Molecular Networking of *Streptomyces noursei* NTR-SR4. *J. Nat. Prod.* **2019**, *82*, 1478−1486.

(95) Wolpert, M.; Gust, B.; Kammerer, B.; Heide, L. Effects of Deletions of *mbtH*-like Genes on Clorobiocin Biosynthesis in *Streptomyces coelicolor*. *Microbiology* **2007**, *153*, 1413−1423.

(96) Liu, J.; Wang, B.; Li, H.; Xie, Y.; Li, Q.; Qin, X.; Zhang, X.; Ju, J. Biosynthesis of the Anti-Infective Marformycins Featuring Pre-NRPS Assembly Line *N*-Formylation and *O*-Methylation and Post-Assembly Line *C*-Hydroxylation Chemistries. *Org. Lett.* **2015**, *17*, 1509−1512.

(97) 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* **2010**, *49*, 8815−8817.

(98) Zhang, W.; Heemstra Jr, J. R.; Walsh, C. T.; Imker, H. J. Activation of the Pacidamycin PacL Adenylation Domain by MbtH-like Proteins. *Biochemistry* **2010**, *49,* 9946−9947.

(99) Boll, B.; Taubitz, T.; Heide, L. Role of MbtH-like Proteins in the Adenylation of Tyrosine during Aminocoumarin and Vancomycin Biosynthesis. *J. Biol. Chem.* **2011**, *286*, 36281−36290.

(100) Kittilä, T.; Mollo, A.; Charkoudian, L. K.; Cryle, M. J. New Structural Data Reveal the Motion of Carrier Proteins in Nonribosomal Peptide Synthesis. *Angew. Chemie Int. Ed.* **2016**, *55,* 9834−9840.

(101) Weber, T.; Baumgartner, R.; Renner, C.; Marahiel, M. A.; Holak, T. A. Solution Structure of PCP, a Prototype for the Peptidyl Carrier Domains of Modular Peptide Synthetases. *Structure* **2000**, *8*, 407−418.

(102) Koglin, A.; Mofid, M. R.; Löhr, F.; Schäfer, B.; Rogov, V. V; Blum, M.-M.; Mittag, T.; Marahiel, M. A.; Bernhard, F.; Dötsch, V. Conformational Switches Modulate Protein Interactions in Peptide Antibiotic Synthetases. *Science* **2006**, *312*, 273−276.

(103) 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.* **1996**, *3*, 923−936.

(104) Beld, J.; Sonnenschein, E. C.; Vickery, C. R.; Noel, J. P.; Burkart, M. D. The Phosphopantetheinyl Transferases: Catalysis of a Post-Translational Modification Crucial for Life. *Nat. Prod. Rep.* **2014**, 31, 61−108.

(105) Stachelhaus, T.; Mootz, H. D.; Bergendahl, V.; Marahiel, M. A. Peptide Bond Formation in Nonribosomal Peptide Biosynthesis Catalytic Role of the Condensation Domain. *J. Biol. Chem.* **1998**, *273*, 22773−22781.

(106) Clugston, S. L.; Sieber, S. A.; Marahiel, M. A.; Walsh, C. T. Chirality of Peptide Bond-Forming Condensation Domains in Nonribosomal Peptide Synthetases: The C_5 Domain of Tyrocidine Synthetase Is a ^DC_L Catalyst. *Biochemistry* **2003**, *42*, 12095−12104.

(107) Keating, T. A.; Marshall, C. G.; Walsh, C. T.; Keating, A. E. The Structure of VibH Represents Nonribosomal Peptide Synthetase Condensation, Cyclization and Epimerization Domains. *Nat. Struct. Biol.* **2002**, *9*, 522−526.

(108) Bergendahl, V.; Linne, U.; Marahiel, M. A. Mutational Analysis of the C-domain in Nonribosomal Peptide Synthesis. *Eur. J. Biochem.* **2002**, *269*, 620−629.

(109) Roche, E. D.; Walsh, C. T. Dissection of the EntF Condensation Domain Boundary and Active Site Residues in Nonribosomal Peptide Synthesis. *Biochemistry* **2003**, *42*, 1334−1344.

(110) Samel, S. A.; Schoenafinger, G.; Knappe, T. A.; Marahiel, M. A.; Essen, L.-O. Structural and Functional Insights into a Peptide Bond-Forming Bidomain from a Nonribosomal Peptide Synthetase. *Structure* **2007**, *15*, 781−792.

(111) Samel, S. A.; Czodrowski, P.; Essen, L.-O. Structure of the Epimerization Domain of Tyrocidine Synthetase A. *Acta Crystallogr. D* **2014**, *70*, 1442−1452.

(112) Belshaw, P. J.; Walsh, C. T.; Stachelhaus, T. Aminoacyl-CoAs as Probes of Condensation Domain Selectivity in Nonribosomal Peptide Synthesis. *Science* **1999**, *284*, 486−489.

(113) Rausch, C.; Hoof, I.; Weber, T.; Wohlleben, W.; Huson, D. H. Phylogenetic Analysis of Condensation Domains in NRPS Sheds Light on Their Functional Evolution. *BMC Evol. Biol.* **2007**, *7*, 78.

(114) Kopp, F.; Linne, U.; Oberthür, M.; Marahiel, M. A. Harnessing the Chemical Activation Inherent to Carrier Protein-Bound Thioesters for the Characterization of Lipopeptide Fatty Acid Tailoring Enzymes. *J. Am. Chem. Soc.* **2008**, *130*, 2656−2666.

(115) Kraas, F. I.; Giessen, T. W.; Marahiel, M. A. Exploring the Mechanism of Lipid Transfer during Biosynthesis of the Acidic Lipopeptide Antibiotic CDA. *FEBS Lett.* **2012**, *586*, 283−288.

(116) Götze, S.; Herbst-Irmer, R.; Klapper, M.; Görls, H.; Schneider, K. R. A.; Barnett, R.; Burks, T.; Neu, U.; Stallforth, P. Structure, Biosynthesis, and Biological Activity of the Cyclic Lipopeptide Anikasin. *ACS Chem. Biol.* **2017**, *12*, 2498−2502.

(117) Robbel, L.; Marahiel, M. A. Daptomycin, a Bacterial Lipopeptide Synthesized

by a Nonribosomal Machinery. *J. Biol. Chem.* **2010**, *285*, 27501−27508.

(118) Bloudoff, K.; Rodionov, D.; Schmeing, T. M. Crystal Structures of the First Condensation Domain of CDA Synthetase Suggest Conformational Changes during the Synthetic Cycle of Nonribosomal Peptide Synthetases. *J. Mol. Biol.* **2013**, *425*, 3137−3150.

(119) Weber, G.; Schörgendorfer, K.; Schneider-Scherzer, E.; Leitner, E. The Peptide Synthetase Catalyzing Cyclosporine Production in *Tolypocladium niveum* Is Encoded by a Giant 45.8-Kilobase Open Reading Frame. *Curr. Genet.* **1994**, *26*, 120−125.

(120) Hoffmann, K.; Schneider-Scherzer, E.; Kleinkauf, H.; Zocher, R. Purification and Characterization of Eucaryotic Alanine Racemase Acting as Key Enzyme in Cyclosporin Biosynthesis. *J. Biol. Chem.* **1994**, *269*, 12710−12714.

(121) Stein, T.; Kluge, B.; Vater, J.; Franke, P.; Otto, A.; Wittmann-Liebold, B. Gramicidin S Synthetase 1 (Phenylalanine Racemase), a Prototype of Amino Acid Racemases Containing the Cofactor 4'-Phosphopantetheine. *Biochemistry* **1995**, *34*, 4633−4642.

(122) Stindl, A.; Keller, U. Epimerization of the D-Valine Portion in the Biosynthesis of Actinomycin D. *Biochemistry* **1994**, *33*, 9358−9364.

(123) Balibar, C. J.; Vaillancourt, F. H.; Walsh, C. T. Generation of D Amino Acid Residues in Assembly of Arthrofactin by Dual Condensation/Epimerization Domains. *Chem. Biol.* **2005**, *12*, 1189−1200.

(124) Velkov, T.; Lawen, A. Mapping and Molecular Modeling of *S*-Adenosyl-L-Methionine Binding Sites in *N*-Methyltransferase Domains of the Multifunctional Polypeptide Cyclosporin Synthetase. *J. Biol. Chem.* **2003**, *278*, 1137−1148.

(125) de Crécy-Lagard, V.; Blanc, V.; Gil, P.; Naudin, L.; Lorenzon, S.; Famechon, A.; Bamas-Jacques, N.; Crouzet, J.; Thibaut, D. Pristinamycin I Biosynthesis in *Streptomyces pristinaespiralis*: Molecular Characterization of the First Two Structural Peptide Synthetase Genes. *J. Bacteriol.* **1997**, *179*, 705−713.

(126) Chatterjee, J.; Rechenmacher, F.; Kessler, H. *N*-methylation of Peptides and Proteins: An Important Element for Modulating Biological Functions. *Angew. Chemie Int. Ed.* **2013**, *52*, 254−269.

(127) Labby, K. J.; Watsula, S. G.; Garneau-Tsodikova, S. Interrupted Adenylation Domains: Unique Bifunctional Enzymes Involved in Nonribosomal Peptide Biosynthesis. *Nat. Prod. Rep.* **2015**, *32*, 641−653.

(128) Kessler, N.; Schuhmann, H.; Morneweg, S.; Linne, U.; Marahiel, M. A. The Linear Pentadecapeptide Gramicidin Is Assembled by Four Multimodular Nonribosomal Peptide Synthetases That Comprise 16 Modules with 56 Catalytic Domains. *J. Biol. Chem.* **2004**, *279*, 7413−7419.

(129) Schoenafinger, G.; Schracke, N.; Linne, U.; Marahiel, M. A. Formylation Domain: An Essential Modifying Enzyme for the Nonribosomal Biosynthesis of Linear Gramicidin. *J. Am. Chem. Soc.* **2006**, *128*, 7406−7407.

(130) Rouhiainen, L.; Paulin, L.; Suomalainen, S.; Hyytiäinen, H.; Buikema, W.; Haselkorn, R.; Sivonen, K. Genes Encoding Synthetases of Cyclic Depsipeptides, Anabaenopeptilides, in *Anabaena* Strain 90. *Mol. Microbiol.* **2000**, *37*, 156−167.

(131) Wallace, B. A. Common Structural Features in Gramicidin and Other Ion Channels. *Bioessays* **2000**, *22*, 227−234.

(132) Kopp, F.; Marahiel, M. A. Macrocyclization Strategies in Polyketide and Nonribosomal Peptide Biosynthesis. *Nat. Prod. Rep.* **2007**, *24*, 735−749.

(133) Galea, C. A.; Han, M.; Zhu, Y.; Roberts, K.; Wang, J.; Thompson, P. E.; Velkov, T. Characterization of the Polymyxin D Synthetase Biosynthetic Cluster and Product Profile of *Paenibacillus polymyxa* ATCC 10401. *J. Nat. Prod.* **2017**, *80*, 1264−1274.

(134) Zhao, L.; Bode, H. B. Production of a Photohexapeptide Library from Entomopathogenic *Photorhabdus asymbiotica* PB68.1. *Org. Biomol. Chem.* **2019**, *17*, 7858−7862.

(135) Du, L.; Lou, L. PKS and NRPS Release Mechanisms. *Nat. Prod. Rep.* **2010**, *27*, 255−278.

(136) Tseng, C. C.; Bruner, S. D.; Kohli, R. M.; Marahiel, M. A.; Walsh, C. T.; Sieber, S. A. Characterization of the Surfactin Synthetase C-Terminal Thioesterase Domain as a Cyclic Depsipeptide Synthase. *Biochemistry* **2002**, *41*, 13350−13359.

(137) Samel, S. A.; Wagner, B.; Marahiel, M. A.; Essen, L.-O. The Thioesterase Domain of the Fengycin Biosynthesis Cluster: A Structural Base for the Macrocyclization of a Non-ribosomal Lipopeptide. *J. Mol. Biol.* **2006**, *359*, 876−889.

(138) Zhao, L.; Awori, R. M.; Kaiser, M.; Groß, J.; Opatz, T.; Bode, H. B. Structure, Biosynthesis, and Bioactivity of Photoditritide from *Photorhabdus temperata* Meg1. *J. Nat. Prod.* **2019**, *82*, 3499−3503.

(139) Dose, B.; Niehs, S. P.; Scherlach, K.; Flórez, L. V; Kaltenpoth, M.; Hertweck, C. Unexpected Bacterial Origin of the Antibiotic Icosalide: Two-Tailed Depsipeptide Assembly in Multifarious *Burkholderia* Symbionts. *ACS Chem. Biol.* **2018**, *13*, 2414−2420.

(140) Sun, C.; Yang, Z.; Zhang, C.; Liu, Z.; He, J.; Liu, Q.; Zhang, T.; Ju, J.; Ma, J. Genome Mining of *Streptomyces atratus* SCSIO ZH16: Discovery of Atratumycin and Identification of Its Biosynthetic Gene Cluster. *Org. Lett.* **2019**, *21*, 1453−1457.

(141) Nowak, S. Characterisation and Structural Assignments of Monomodular Non-ribosomal Peptide Synthetases from the Entomopathogenic Bacterium *Xenorhabdus* KJ12.1, Goethe University Frankfurt, 2018.

(142) Gao, X.; Haynes, S. W.; Ames, B. D.; Wang, P.; Vien, L. P.; Walsh, C. T.; Tang, Y. Cyclization of Fungal Nonribosomal Peptides by a Terminal Condensation-like Domain. *Nat. Chem. Biol.* **2012**, *8*, 823−830.

(143) Jin, J.; Lee, S.; Lee, J.; Baek, S.; Kim, J.; Yun, S.; Park, S.; Kang, S.; Lee, Y. Functional Characterization and Manipulation of the Apicidin Biosynthetic Pathway in *Fusarium semitectum*. *Mol. Microbiol.* **2010**, *76*, 456−466.

(144) Zhang, J.; Liu, N.; Cacho, R. A.; Gong, Z.; Liu, Z.; Qin, W.; Tang, C.; Tang, Y.; Zhou, J. Structural Basis of Nonribosomal Peptide Macrocyclization in Fungi. *Nat. Chem. Biol.* **2016**, *12*, 1001−1003.

(145) Mootz, H. D.; Marahiel, M. A. The Tyrocidine Biosynthesis Operon of *Bacillus brevis*: Complete Nucleotide Sequence and Biochemical Characterization of Functional Internal Adenylation Domains. *J. Bacteriol.* **1997**, *179,* 6843−6850.

(146) Cosmina, P.; Rodriguez, F.; de Ferra, F.; Grandi, G.; Perego, M.; Venema, G.;

van Sinderen, D. Sequence and Analysis of the Genetic Locus Responsible for Surfactin Synthesis in *Bacillus subtilis*. *Mol. Microbiol.* **1993**, *8*, 821−831.

(147) Gehring, A. M.; Mori, I.; Walsh, C. T. Reconstitution and Characterization of the *Escherichia coli* Enterobactin Synthetase from EntB, EntE, and EntF. *Biochemistry* **1998**, *37*, 2648−2659.

(148) Xu, Y.; Orozco, R.; Wijeratne, E. M. K.; Espinosa-Artiles, P.; Gunatilaka, A. A. L.; Stock, S. P.; Molnár, I. Biosynthesis of the Cyclooligomer Depsipeptide Bassianolide, an Insecticidal Virulence Factor of *Beauveria bassiana*. *Fungal Genet. Biol.* **2009**, *46,* 353−364.

(149) Hoyer, K. M.; Mahlert, C.; Marahiel, M. A. The Iterative Gramicidin S Thioesterase Catalyzes Peptide Ligation and Cyclization. *Chem. Biol.* **2007**, *14,* 13−22.

(150) Ali, H.; Ries, M. I.; Lankhorst, P. P.; van der Hoeven, R. A. M.; Schouten, O. L.; Noga, M.; Hankemeier, T.; van Peij, N. N. M. E.; Bovenberg, R. A. L.; Vreeken, R. J.; Driessen, A. J. M. A Non-canonical NRPS Is Involved in the Synthesis of Fungisporin and Related Hydrophobic Cyclic Tetrapeptides in *Penicillium chrysogenum*. *PLoS One* **2014**, *9*, e98212.

(151) Magarvey, N. A.; Haltli, B.; He, M.; Greenstein, M.; Hucul, J. A. Biosynthetic Pathway for Mannopeptimycins, Lipoglycopeptide Antibiotics Active against Drug-Resistant Gram-Positive Pathogens. *Antimicrob. Agents Chemother.* **2006**, *50*, 2167−2177.

(152) Tulp, M.; Bohlin, L. Rediscovery of Known Natural Compounds: Nuisance or Goldmine? *Bioorg. Med. Chem.* **2005**, *13*, 5274−5282.

(153) Rutledge, P. J.; Challis, G. L. Discovery of Microbial Natural Products by Activation of Silent Biosynthetic Gene Clusters. *Nat. Rev. Microbiol.* **2015**, *13*, 509−523.

(154) Tietz, J. I.; Mitchell, D. A. Using Genomics for Natural Product Structure Elucidation. *Curr. Top. Med. Chem.* **2016**, *16*, 1645−1694.

(155) van Wezel, G. P.; McDowall, K. J. The Regulation of the Secondary Metabolism of *Streptomyces*: New Links and Experimental Advances. *Nat. Prod. Rep.* **2011**, *28*, 1311−1333.

(156) Brakhage, A. A. Regulation of Fungal Secondary Metabolism. *Nat. Rev. Microbiol.* **2013**, *11*, 21−32.

(157) Chiang, Y.-M.; Chang, S.-L.; Oakley, B. R.; Wang, C. C. C. Recent Advances in Awakening Silent Biosynthetic Gene Clusters and Linking Orphan Clusters to Natural Products in Microorganisms. *Curr. Opin. Chem. Biol.* **2011**, *15*, 137−143.

(158) Ren, H.; Wang, B.; Zhao, H. Breaking the Silence: New Strategies for Discovering Novel Natural Products. *Curr. Opin. Biotechnol.* **2017**, *48*, 21−27.

(159) Dias, D. A.; Jones, O. A. H.; Beale, D. J.; Boughton, B. A.; Benheim, D.; Kouremenos, K. A.; Wolfender, J.-L.; Wishart, D. S. Current and Future Perspectives on the Structural Identification of Small Molecules in Biological Systems. *Metabolites* **2016**, *6*, 46.

(160) Linington, R. G.; Kubanek, J.; Luesch, H. New Methods for Isolation and Structure Determination of Natural Products. *Nat. Prod. Rep.* **2019**, *36*, 942−943.

(161) Franke, J.; Ishida, K.; Hertweck, C. Genomics-Driven Discovery of Burkholderic Acid, a Noncanonical, Cryptic Polyketide from Human Pathogenic *Burkholderia* Species. *Angew. Chemie Int. Ed.* **2012**, *51*, 11611−11615.

(162) Bode, E. ; Heinrich, A. K.; Hirschmann, M.; Abebew, D.; Shi, Y.-N.; Vo, T. D.; Wesche, F.; Shi, Y.-M.; Grün, P.; Simonyi, S.; Keller, N.; Engel, Y.; Wenski, S.; Bennet, R.; Beyer, S.; Bischoff, I.; Buaya, A.; Brandt, S.; Cakmak, I.; Cimen, H.; Eckstein, S.; Frank, D.; Fürst, R.; Gand, M.; Geisslinger, G.; Hazir, S.; Henke, M.; Heermann, R.; Lecaudey, V.; Schäfer, W.; Schiffmann, S.; Schüffler, A.; Schwenk, R.; Skaljac, M.; Thines, E.; Thines, M.; Ulshöfer, T.; Vilcinskas, A.; Wichelhaus, T. A.; Bode, H. B. Promoter Activation in Δ*hfq* Mutants as an Efficient Tool for Specialized Metabolite Production Enabling Direct Bioactivity Testing. *Angew. Chemie Int. Ed.* **2019**, *58*, 18957−18963.

(163) Yamanaka, K.; Reynolds, K. A.; Kersten, R. D.; Ryan, K. S.; Gonzalez, D. J.; Nizet, V.; Dorrestein, P. C.; Moore, B. S. Direct Cloning and Refactoring of a Silent Lipopeptide Biosynthetic Gene Cluster Yields the Antibiotic Taromycin A. *Proc. Natl. Acad. Sci.* **2014**, *111*, 1957−1962.

(164) Reynolds, K. A.; Luhavaya, H.; Li, J.; Dahesh, S.; Nizet, V.; Yamanaka, K.;

Moore, B. S. Isolation and Structure Elucidation of Lipopeptide Antibiotic Taromycin B from the Activated Taromycin Biosynthetic Gene Cluster. *J. Antibiot.* **2018**, *71*, 333−338.

(165) Tang, Y.; Frewert, S.; Harmrolfs, K.; Herrmann, J.; Karmann, L.; Kazmaier, U.; Xia, L.; Zhang, Y.; Müller, R. Heterologous Expression of an Orphan NRPS Gene Cluster from *Paenibacillus larvae* in *Escherichia coli* Revealed Production of Sevadicin. *J. Biotechnol.* **2015**, *194*, 112−114.

(166) Fu, J.; Bian, X.; Hu, S.; Wang, H.; Huang, F.; Seibert, P. M.; Plaza, A.; Xia, L.; Müller, R.; Stewart, A. F.; Zhang, Y. Full-Length RecE Enhances Linear-Linear Homologous Recombination and Facilitates Direct Cloning for Bioprospecting. *Nat. Biotechnol.* **2012**, *30*, 440−446.

(167) Nollmann, F. I.; Dauth, C.; Mulley, G.; Kegler, C.; Kaiser, M.; Waterfield, N. R.; Bode, H. B. Insect-Specific Production of New GameXPeptides in *Photorhabdus luminescens* TTO1, Widespread Natural Products in Entomopathogenic Bacteria. *ChemBioChem* **2015**, *16*, 205−208.

(168) Dimise, E. J.; Widboom, P. F.; Bruner, S. D. Structure Elucidation and Biosynthesis of Fuscachelins, Peptide Siderophores from the Moderate Thermophile *Thermobifida fusca*. *Proc. Natl. Acad. Sci.* **2008**, *105*, 15311−15316.

(169) Wyatt, M. A.; Wang, W.; Roux, C. M.; Beasley, F. C.; Heinrichs, D. E.; Dunman, P. M.; Magarvey, N. A. *Staphylococcus aureus* Nonribosomal Peptide Secondary Metabolites Regulate Virulence. *Science* **2010**, *329*, 294−296.

(170) 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* **2015**, *517*, 455−459.

(171) Wenzel, S. C.; Kunze, B.; Höfle, G.; Silakowski, B.; Scharfe, M.; Blöcker, H.; Müller, R. Structure and Biosynthesis of Myxochromides $S_{1–3}$ in *Stigmatella aurantiaca*: Evidence for an Iterative Bacterial Type I Polyketide Synthase and for Module Skipping in Nonribosomal Peptide Biosynthesis. *ChemBioChem* **2005**, *6*, 375−385.

(172) Karas, M.; Hillenkamp, F. Laser Desorption Ionization of Proteins with Molecular Masses Exceeding 10,000 Daltons. *Anal. Chem.* **1988**, *60*, 2299−2301.

(173) Yamashita, M.; Fenn, J. B. Electrospray Ion Source. Another Variation on the Free-Jet Theme. *J. Phys. Chem.* **1984**, *88*, 4451−4459.

(174) Koomsiri, W.; Inahashi, Y.; Leetanasaksakul, K.; Shiomi, K.; Takahashi, Y.; Ōmura, S.; Samborskyy, M.; Leadlay, P. F.; Wattana-Amorn, P.; Thamchaipenet, A.; Nakashima, T. Sarpeptins A and B, Lipopeptides Produced by *Streptomyces* sp. KO-7888 Overexpressing a Specific SARP Regulator. *J. Nat. Prod.* **2019**, *82*, 2144−2151.

(175) Zhang, Z.-Y.; Ma, N.; Tao, L.-J.; Gong, X.-Y.; Ye, W.-C.; Wang, L. Linear Peptides Containing D-Leucine with Neuroprotective Activities from the Leech *Whitmania pigra* Whitman. *J. Nat. Prod.* **2019**, *82*, 2349−2353.

(176) Kersten, R. D.; Yang, Y.-L.; Xu, Y.; Cimermancic, P.; Nam, S.-J.; Fenical, W.; Fischbach, M. A.; Moore, B. S.; Dorrestein, P. C. A Mass Spectrometry-guided Genome Mining Approach for Natural Product Peptidogenomics. *Nat. Chem. Biol.* **2011**, *7*, 794−802.

(177) Silber, J.; Ohlendorf, B.; Labes, A.; Näther, C.; Imhoff, J. F. Calcaripeptides A−C, Cyclodepsipeptides from a *Calcarisporium* Strain. *J. Nat. Prod.* **2013**, *76*, 1461−1467.

(178) Klitgaard, A.; Nielsen, J. B.; Frandsen, R. J. N.; Andersen, M. R.; Nielsen, K. F. Combining Stable Isotope Labeling and Molecular Networking for Biosynthetic Pathway Characterization. *Anal. Chem.* **2015**, *87*, 6520−6526.

(179) Oh, D.-C.; Kauffman, C. A.; Jensen, P. R.; Fenical, W. Induced Production of Emericellamides A and B from the Marine-Derived Fungus *Emericella* sp. in Competing Co-culture. *J. Nat. Prod.* **2007**, *70*, 515−520.

(180) Elyashberg, M. Identification and Structure Elucidation by NMR Spectroscopy. *TrAC Trends Anal. Chem.* **2015**, *69*, 88−97.

(181) Maier, M. E. Structural Revisions of Natural Products by Total Synthesis. *Nat. Prod. Rep.* **2009**, *26*, 1105−1124.

(182) Suyama, T. L.; Gerwick, W. H.; McPhail, K. L. Survey of Marine Natural

Product Structure Revisions: A Synergy of Spectroscopy and Chemical Synthesis. *Bioorg. Med. Chem.* **2011**, *19*, 6675−6701.

(183) Chhetri, B. K.; Lavoie, S.; Sweeney-Jones, A. M.; Kubanek, J. Recent Trends in the Structural Revision of Natural Products. *Nat. Prod. Rep.* **2018**, *35*, 514−531.

(184) Calabro, K.; Genta-Jouve, G.; Thomas, O. P. Structure Revision of Microginins 674 and 690 from the Cultured Cyanobacterium *Microcystis aeruginosa*. *J. Nat. Prod.* **2019**, *82*, 1040−1044.

(185) Fujii, K.; Ikai, Y.; Oka, H.; Suzuki, M.; Harada, K. A Nonempirical Method Using LC/MS for Determination of the Absolute Configuration of Constituent Amino Acids in a Peptide: Combination of Marfey's Method with Mass Spectrometry and Its Practical Application. *Anal. Chem.* **1997**, *69,* 5146−5151.

(186) Masuda, Y.; Tanaka, R.; Ganesan, A.; Doi, T. Structure Revision of Similanamide to PF1171C by Total Synthesis. *J. Nat. Prod.* **2015**, *78*, 2286−2291.

(187) Baltz, R. H. Function of MbtH Homologs in Nonribosomal Peptide Biosynthesis and Applications in Secondary Metabolite Discovery. *J. Ind. Microbiol. Biotechnol.* **2011**, *38*, 1747−1760.

(188) Hacker, C.; Cai, X.; Kegler, C.; Zhao, L.; Weickhmann, A. K.; Wurm, J. P.; Bode, H. B.; Wöhnert, J. Structure-Based Redesign of Docking Domain Interactions Modulates the Product Spectrum of a Rhabdopeptide-Synthesizing NRPS. *Nat. Commun.* **2018**, *9*, 4366.

(189) Crüsemann, M.; Kohlhaas, C.; Piel, J. Evolution-Guided Engineering of Nonribosomal Peptide Synthetase Adenylation Domains. *Chem. Sci.* **2013**, *4*, 1041−1045.

(190) Wesche, F.; Adihou, H.; Kaiser, A.; Wurglics, M.; Schubert-Zsilavecz, M.; Kaiser, M.; Bode, H. B. Combined Approach of Backbone Amide Linking and On-Resin *N*-Methylation for the Synthesis of Bioactive and Metabolically Stable Peptides. *J. Med. Chem.* **2018**, *61*, 3930−3938.

(191) Weng, J.-K.; Noel, J. P. The Remarkable Pliability and Promiscuity of Specialized Metabolism. *Cold Spring Harbor Symp. Quant. Biol.* **2012**, *77*,

309−320.

(192) Firn, R. D.; Jones, C. G. Natural Products − a Simple Model to Explain Chemical Diversity. *Nat. Prod. Rep.* **2003**, *20*, 382−391.

(193) Burkhart, B. M.; Gassman, R. M.; Langs, D. A.; Pangborn, W. A.; Duax, W. L.; Pletnev, V. Gramicidin D Conformation, Dynamics and Membrane Ion Transport. *Pept. Sci.* **1999**, *51*, 129−144.

(194) Williams, G. J.; Breazeale, S. D.; Raetz, C. R. H.; Naismith, J. H. Structure and Function of Both Domains of ArnA, a Dual Function Decarboxylase and a Formyltransferase, Involved in 4-Amino-4-deoxy-L-arabinose Biosynthesis. *J. Biol. Chem.* **2005**, *280*, 23000−23008.

(195) Bonnington, L. S.; Tanaka, J.; Higa, T.; Kimura, J.; Yoshimura, Y.; Nakao, Y.; Yoshida, W. Y.; Scheuer, P. J. Cupolamide A: A Cytotoxic Cyclic Heptapeptide from Two Samples of the Sponge *Theonella cupola*. *J. Org. Chem.* **1997**, *62*, 7765−7767.

(196) Saito, K.; Konno, A.; Ishii, H.; Saito, H.; Nishida, F.; Abe, T.; Chen, C. Nodularin-Har: A New Nodularin from *Nodularia*. *J. Nat. Prod.* **2001**, *64*, 139−141.

(197) Cha, J. W.; Park, J.-S.; Sim, T.; Nam, S.-J.; Kwon, H. C.; Del Valle, J. R.; Fenical, W. Structure Assignment of Lucentamycin E and Revision of the Olefin Geometries of the Marine-Derived Lucentamycins. *J. Nat. Prod.* **2012**, *75*, 1648−1651.

(198) Hernández-Guzmán, G.; Alvarez-Morales, A. Isolation and Characterization of the Gene Coding for the Amidinotransferase Involved in the Biosynthesis of Phaseolotoxin in *Pseudomonas syringae* pv. *phaseolicola*. *Mol. plant-microbe Interact.* **2001**, *14*, 545−554.

(199) Vilhena, C.; Bettencourt, A. Daptomycin: A Review of Properties, Clinical Use, Drug Delivery and Resistance. *Mini Rev. Med. Chem.* **2012**, *12*, 202−209.

(200) Bozhüyük, K. A. J.; Fleischhacker, F.; Linck, A.; Wesche, F.; Tietze, A.; Niesert, C.-P.; Bode, H. B. De Novo Design and Engineering of Non-ribosomal Peptide Synthetases. *Nat. Chem.* **2018**, *10*, 275−281.

(201) Bozhüyük, K. A. J.; Linck, A.; Tietze, A.; Kranz, J.; Wesche, F.; Nowak, S.;

Fleischhacker, F.; Shi, Y.-N.; Grün, P.; Bode, H. B. Modification and de Novo Design of Non-ribosomal Peptide Synthetases Using Specific Assembly Points within Condensation Domains. *Nat. Chem.* **2019**, *11*, 653−661.

5 Attachments

5.1 Rhabdopeptide/Xenortide-like Peptides from *Xenorhabdus innexi* with Terminal Amines Showing Potent Antiprotozoal Activity

Authors:

Lei Zhao,[1,2] Marcel Kaiser,[3,4] and Helge B. Bode*[1,5]

[1]Molecular Biotechnology, Department of Biosciences, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

[2]Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, 210014 Nanjing, China

[3]Parasite Chemotherapy, Swiss Tropical and Public Health Institute, 4051 Basel, Switzerland

[4]University of Basel, 4003 Basel, Switzerland

[5]Buchmann Institute for Molecular Life Sciences (BMLS), Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

*Corresponding author

Declaration on the contribution of the authors

Publication: Rhabdopeptide/Xenortide-like Peptides from *Xenorhabdus innexi* with Terminal Amines Showing Potent Antiprotozoal Activity

Status: published

Journal: Organic Letters

Authors: <u>Lei Zhao (LZ)</u>, Marcel Kaiser (MK), Helge B. Bode (HBB)

What did the doctoral candidate or the co-authors contribute individually to the dissertation?

(1) Development and planning

<u>LZ (50%)</u>, HBB (50%)

(2) Performance of individual research and experiments

<u>LZ (90%): strain cultivation, extract preparation, HPLC-MS analysis, compound isolation, NMR measurement, configuration determination, biosynthesis analysis;</u> MK (10%): bioactivity testing

(3) Collection of data and preparation of figures

<u>LZ (90%): collection of NMR and HPLC-MS data, preparation of all figures;</u> MK (10%): collection of bioactivity data

(4) Analysis and interpretation of data

<u>LZ (90%): analysis of NMR and HPLC-MS data, interpretation of bioactivity data;</u> HBB (10%): interpretation of bioactivity data

(5) Preparation of manuscript

<u>LZ (80%)</u>, HBB (20%)

Herewith approving the indications above

_____
date/place

_____
signature doctoral candidate

_____
date/place

_____
signature supervisor

_____
date/place

_____
If necessary, signature corresponding author

Rhabdopeptide/Xenortide-like Peptides from *Xenorhabdus innexi* with Terminal Amines Showing Potent Antiprotozoal Activity

Lei Zhao,[†,‡] Marcel Kaiser,[§,⊥] and Helge B. Bode*[,†,‖]

[†]Molekulare Biotechnologie, Fachbereich Biowissenschaften, Goethe Universität Frankfurt, 60438 Frankfurt am Main, Germany

[‡]Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, 210014 Nanjing, China

[§]Parasite Chemotherapy, Swiss Tropical and Public Health Institute, 4051 Basel, Switzerland

[⊥]University of Basel, 4003 Basel, Switzerland

[‖]Buchmann Institute for Molecular Life Sciences (BMLS), Goethe Universität Frankfurt, 60438 Frankfurt am Main, Germany

S *Supporting Information*

ABSTRACT: Seven new rhabdopeptide/xenortide-like peptides (RXPs) (**1−7**) with putrescine or ammonia as the C-terminal amines were isolated from *Xenorhabdus innexi* DSM 16336. Their chemical structures were elucidated by high-resolution mass spectroscopy (HR-MS) and one-dimensional (1D) and two-dimensional (2D) NMR. They were evaluated for their activities against protozoan parasites and cytotoxicity against rat skeletal myoblasts (L6 cells). All tested compounds exhibited strong effects against *Trypanosoma brucei rhodesiense* and *Plasmodium falciparum*, with IC_{50} values of 0.07−6.25 and 0.091−3.16 μM, respectively, making them the most active RXP derivatives known to date.
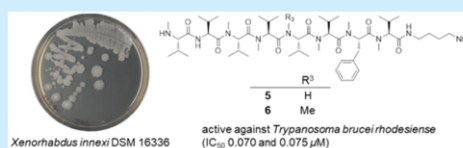
Parasitic diseases, such as malaria, sleeping sickness, Chagas disease, and leishmaniasis, are a major public health
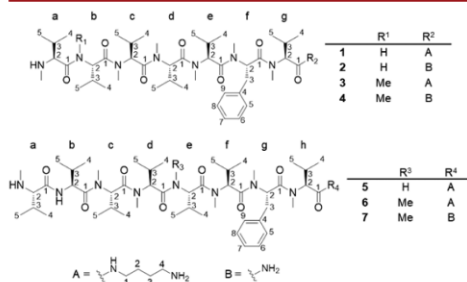
Figure 1. Chemical structures of compounds **1−7**.

problem worldwide, especially affecting hundreds of millions of the poorest populations in tropical and subtropical countries.[1,2] Because of the low return on investments, only a few antiparasitic drugs are available. Unfortunately, most of these drugs including chloroquine (malaria), melarsoprol (sleeping sickness), benznidazole (Chagas disease), and miltefosine (leishmaniasis) are decades old and have one or more limitations, such as high cost, toxicity, and drug resistance.[2,3] Thus, there is an urgent need for the development of novel, effective, inexpensive, and safe drugs.

Entomopathogenic bacteria of the genera *Xenorhabdus* and *Photorhabdus* are a rich source of bioactive natural products.[4,5]

It is suggested that these natural products have a specialized function, including the protection of the insect cadaver against food competitors such as bacteria, fungi, and soil-living protozoa, to support the symbiosis with nematodes of the genus *Steinernema* and *Heterorhabditis*.[4,5] Rhabdopeptide/xenortide-like peptides (RXPs), which represent a unique class of peptides exclusively found in *Xenorhabdus* and *Photorhabdus*,[6−10] are probably responsible for defending protozoan competitors, as shown in our previous study.[8−10] A large number of RXPs have been identified in different *Xenorhabdus* and *Photorhabdus* wild-type strains by HPLC-MS analysis from our in-house strain collection.[10] These derivatives are composed of two to eight amino acids (valine, phenylalanine, and leucine) that are often *N*-methylated and show phenylethylamine or tryptamine as the main C-terminal amines.[10] During more-detailed HPLC-MS analysis of different *Xenorhabdus* and *Photorhabdus* strains, we found several new RXPs showing different C-terminal amine fragmentation pattern of 88 or 17 mass units in *X. innexi* DSM 16336, instead of 121 (phenylethylamine) and 160 (tryptamine) mass units that are typical for most RXPs in most other strains. Here, we describe the isolation and structure elucidation of seven new RXPs (**1−7**) (see Figure 1) with putrescine or ammonia as the C-terminal amines from *X. innexi* DSM 16336. Moreover, their bioactivities against protozoan parasites and cytotoxicity against mammalian L6 cells were evaluated.

5116

Table 1. ^{1}H (500 MHz) and ^{13}C (125 MHz) NMR Data of 1−4 in Methanol-d_4 (δ in ppm, J in Hz)

subunit	position	1		2		3		4	
		δ_C	δ_H	δ_C	δ_H	δ_C	δ_H	δ_C	δ_H
a	1	169.9		169.8		171.1		171.9	
	2	68.9	3.49, d (5.1)	68.9	3.48, d (5.1)	65.9	4.13, d (3.5)	66.0	4.03, d (2.4)
	3	31.9	2.13−2.01, m	32.0	2.13−2.02, m	31.4	2.13−2.00, m	31.5	2.07−1.98, m
	4	18.9	0.99, overlap	18.8	0.99, overlap	17.1	0.99, overlap	17.1	0.98, overlap
	5	18.8	0.99, overlap	18.8	0.99, overlap	19.9	1.13, d (7.0)	20.1	1.12, d (7.0)
	HN−CH$_3$	33.8	2.55, s	33.8	2.56, s	34.1	2.55, s	34.3	2.50, s
b	1	174.2		174.2		172.1		172.1	
	2	56.5	4.76, d (7.5)	56.5	4.76, d (7.4)	60.3	5.28, d (10.8)	60.3	5.28, d (10.8)
	3	31.7	2.13−2.01, m	31.7	2.13−2.02, m	28.4	2.42−2.26, m	28.3	2.42−2.26, m
	4	20.1	0.99, overlap	20.1	0.99, overlap	19.8	0.92, d (6.5)	19.9	0.92, d (6.5)
	5	18.6	0.99, overlap	18.6	0.99, overlap	19.0	0.89, overlap	19.0	0.89, d (6.6)
	N−CH$_3$					31.1	3.05, s	31.0	3.05, s
c	1	172.0		172.0		171.9		171.9	
	2	59.9	5.15, d (10.8)	59.9	5.15, d (10.8)	59.8	5.16, d (10.8)	59.8	5.16, d (10.8)
	3	28.6	2.35−2.30, m	28.6	2.36−2.24, m	28.6	2.42−2.26, m	28.6	2.42−2.26, m
	4	20.0	0.82, overlap	20.0	0.82, overlap	19.9	0.82, overlap	20.0	0.82, overlap
	5	19.0	0.77, overlap	19.0	0.78, overlap	18.5	0.76, overlap	18.6	0.76, overlap
	N−CH$_3$	31.3	3.15, s	31.3	3.15, s	31.3	3.07, s	31.4	3.07, s
d	1	171.2		171.2		171.2		171.2	
	2	59.8	4.99, d (10.7)	59.8	4.99, d (10.7)	59.9	4.99, d (10.7)	59.9	4.99, d (10.7)
	3	28.5	2.35−2.30, m	28.5	2.36−2.24, m	28.5	2.42−2.26, m	28.5	2.42−2.26, m
	4	20.5	0.88, overlap	20.5	0.89, d (6.4)	20.5	0.89, overlap	20.5	0.89, d (6.6)
	5	18.6	0.77, overlap	18.6	0.78, overlap	18.7	0.76, overlap	18.6	0.76, overlap
	N−CH$_3$	31.3	2.96, s	31.3	2.96, s	31.4	2.97, s	31.3	2.98, s
e	1	171.1		171.1		171.1		171.1	
	2	60.2	5.03, d (10.6)	60.2	5.03, d (10.7)	60.2	5.03, d (10.6)	60.2	5.04, d (10.6)
	3	28.3	2.28−2.14, m	28.3	2.24−2.13, m	28.3	2.26−2.13, m	28.3	2.25−2.14, m
	4	20.3	0.82, overlap	20.3	0.82, overlap	20.3	0.82, overlap	20.3	0.82, overlap
	5	18.3	0.67, d (6.8)	18.3	0.67, d (6.8)	18.2	0.67, d (6.8)	18.2	0.67, d (6.8)
	N−CH$_3$	30.5	2.29, s	30.5	2.29, s	30.5	2.29, s	30.5	2.29, s
f	1	173.0		173.1		173.0		173.1	
	2	55.5	5.94, dd (9.9, 5.9)	55.5	5.95, dd (10.1, 5.8)	55.5	5.95, dd (10.0, 5.7)	55.5	5.96, dd (10.1, 5.8)
	3	36.1	3.13−3.06, m	36.1	3.13−3.03, m	36.1	3.08−3.01, m	36.1	3.17−3.03, m
	4	138.3		138.3		138.3		138.3	
	5	130.8	7.24, overlap	130.7	7.24, overlap	130.8	7.24, overlap	130.7	7.24, overlap
	6	129.8	7.28, overlap	129.8	7.28, overlap	129.8	7.28, overlap	129.8	7.28, overlap
	7	128.2	7.26, overlap	128.2	7.26, overlap	128.2	7.26, overlap	128.2	7.26, overlap
	8	129.8	7.28, overlap	129.8	7.28, overlap	129.8	7.28, overlap	129.8	7.28, overlap
	9	130.8	7.24, overlap	130.7	7.24, overlap	130.8	7.24, overlap	130.7	7.24, overlap
	N−CH$_3$	31.5	2.98, s	31.6	3.00, s	31.5	2.98, s	31.6	3.00, s
g	1	172.2		174.3		172.2		174.3	
	2	64.3	4.56, d (11.0)	63.6	4.63, d (11.0)	64.3	4.56, d (11.0)	63.6	4.64, d (11.0)
	3	27.9	2.28−2.14, m	27.7	2.24−2.13, m	27.9	2.26−2.13, m	27.7	2.25−2.14, m
	4	19.4	0.82, overlap	19.2	0.80, overlap	19.4	0.82, overlap	19.2	0.80, overlap
	5	20.0	0.94, d (6.4)	20.1	0.99, overlap	20.0	0.95, overlap	20.1	0.98, overlap
	N−CH$_3$	31.6	3.03, s	31.5	3.01, s	31.6	3.03, s	31.5	3.01, s
A/B	1	39.6	3.28−3.22, m			39.6	3.28−3.21, m		
	2	27.5	1.60−1.48, m			27.5	1.60−1.52, m		
	3	26.1	1.69−1.60, m			26.1	1.69−1.60, m		
	4	40.4	2.95−2.91, m			40.4	2.97−2.91, m		
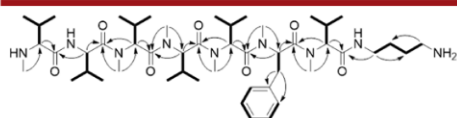
Figure 2. COSY (bold) and key HMBC (arrow) correlations of **1**.

To isolate these new RXPs, *X. innexi* DSM 16336 was cultivated in a lysogeny broth (LB) medium with 2% Amberlite XAD-16. From 4 L of cultures, seven RXPs (**1−7**) were isolated from the XAD-16 extracts by using Sephadex LH-20 chromatography, followed by preparative HPLC.

Compound **1** was obtained as a white solid. The molecular formula of **1** was determined to be $C_{49}H_{87}N_9O_7$ from HR-MS analysis (see Table S1 in the Supporting Information). Its

Table 2. ^{1}H (500 MHz) and ^{13}C (125 MHz) NMR Data of 5-7 in Methanol-d_4 (δ in ppm, J in Hz)

subunit	position	5 δ_C	5 δ_H	6 δ_C	6 δ_H	7 δ_C	7 δ_H
a	1	170.6		170.1		171.4	
	2	69.2	3.41, d (5.5)	69.0	3.46, d (5.3)	69.6	3.33, submerged
	3	32.1	2.13–1.96, m	32.0	2.13–2.01, m	32.2	2.11–1.98, m
	4	19.0	1.01, overlap	19.0	1.01, overlap	19.1	0.99, overlap
	5	19.0	0.99, overlap	18.9	0.99, overlap	19.0	0.99, overlap
	HN–CH$_3$	34.0	2.52, s	33.9	2.54, s	34.2	2.49, s
b	1	174.3		174.3		174.3	
	2	56.4	4.74, d (7.7)	56.5	4.76, d (7.5)	56.4	4.76, d (7.5)
	3	31.8	2.13–1.96, m	31.7	2.13–2.01, m	31.7	2.11–1.98, m
	4	20.0	1.00, overlap	20.0	1.00, overlap	20.0	0.99, overlap
	5	18.7	0.98, overlap	18.6	1.00, overlap	18.6	0.99, overlap
c	1	173.0		172.3		172.3	
	2	59.8	5.20, d (10.8)	59.9	5.19, d (10.8)	59.8	5.19, d (10.8)
	3	28.8	2.37–2.26, m	28.6	2.37–2.29, m	28.6	2.37–2.24
	4	20.0	0.89, overlap	20.0	0.88, overlap	20.1	0.88, overlap
	5	19.1	0.77, d (6.8)	19.0	0.77, d (6.7)	19.0	0.77, overlap
	N–CH$_3$	31.3	3.15, s	31.3	3.14, s	31.3	3.14, s
d	1	171.8		171.8		171.8	
	2	63.7	4.64, d (11.1)	59.8	5.17, d (10.8)	59.8	5.17, d (10.8)
	3	27.9	2.27–2.13, m	28.6	2.37–2.29, m	28.6	2.37–2.24, m
	4	19.4	0.75, overlap	18.6	0.76, d (6.7)	18.6	0.76, overlap
	5	19.8	0.84, overlap	20.0	0.83, overlap	20.0	0,84, overlap
	N–CH$_3$	31.9	3.10, s	31.4	3.04, s	31.4	3.04, s
e	1	173.8		171.2		171.2	
	2	55.8	4.49, d (7.0)	59.9	4.98, d (10.7)	59.9	4.99, d (10.7)
	3	31.5	1.88–1.81, m	28.5	2.37–2.29, m	28.5	2.37–2.24, m
	4	20.4	0.86, overlap	20.5	0.87, overlap	20.5	0.88, overlap
	5	18.2	0.86, overlap	18.5	0.79, overlap	18.5	0.79, overlap
	N–CH$_3$			31.3	2.96, s	31.3	2.97, s
f	1	171.2		171.1		171.1	
	2	60.2	5.02, d (10.7)	60.2	5.03, d (10.6)	60.2	5.03, d (10.7)
	3	28.3	2.27–2.13, m	28.3	2.27–2.13, m	28.3	2.24–2.14, m
	4	20.2	0.83, overlap	20.3	0.83, overlap	20.3	0.83, overlap
	5	18.5	0.70, overlap	18.2	0.67, d (6.8)	18.2	0.67, d (6.8)
	N–CH$_3$	30.6	2.51, s	30.5	2.29, s	30.5	2.29, s
g	1	173.0		173.0		173.1	
	2	55.2	5.95, dd (9.8, 6.5)	55.5	5.94, dd (10.0, 5.9)	55.5	5.95, dd (10.1, 5.7)
	3	36.1	3.17–3.00, m	36.1	3.13–3.06, m	36.1	3.12–3.03, m
	4	138.2		138.3		138.3	
	5	130.8	7.24, overlap	130.8	7.24, overlap	130.7	7.24, overlap
	6	129.7	7.28, overlap	129.8	7.28, overlap	129.8	7.28, overlap
	7	128.2	7.26, overlap	128.2	7.26, overlap	128.2	7.26, overlap
	8	129.7	7.28, overlap	129.8	7.28, overlap	129.8	7.28, overlap
	9	130.8	7.24, overlap	130.8	7.24, overlap	130.7	7.24, overlap
	N–CH$_3$	31.5	2.98, s	31.5	2.98, s	31.6	3.00, s
h	1	172.2		172.2		174.3	
	2	64.3	4.56, d (11.1)	64.3	4.56, d (11.0)	63.6	4.64, d (11.0)
	3	27.9	2.27–2.13, m	27.9	2.27–2.13, m	27.7	2.24–2.14, m
	4	19.4	0.81, overlap	19.4	0.81, overlap	19.2	0.81, overlap
	5	20.0	0.94, overlap	20.0	0.95, d (6.5)	20.0	0.99, overlap
	N–CH$_3$	31.6	3.03, s	31.6	3.04, s	31.5	3.01, s
A/B	1	39.6	3.27–3.20, m	39.6	3.28–3.22, m		
	2	27.5	1.60–1.51, m	27.5	1.60–1.53, m		
	3	26.1	1.69–1.60, m	26.1	1.69–1.60, m		
	4	40.4	2.97–2.90, m	40.4	2.95–2.91, m		

structure was elucidated based on detailed 1D (^{1}H and ^{13}C) and 2D (^{1}H–^{1}H-COSY, ^{1}H–^{13}C-HSQC, and ^{1}H–^{13}C-HMBC) NMR experiments (see Table 1). Seven carbonyl carbon signals with chemical shifts of δ_C 169.9−174.2 were found in the ^{13}C NMR spectrum, and each was assigned to a single amino acid building block (numbered as a−g from the
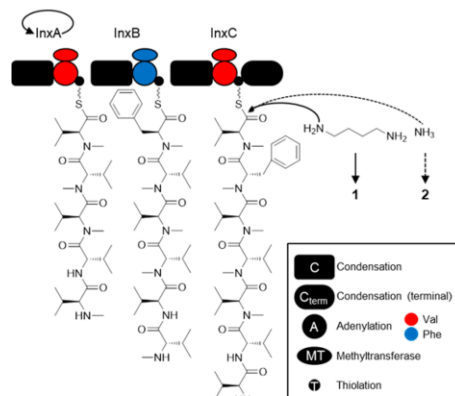
Figure 3. Proposed biosynthesis as shown for **1** and **2** as examples.

N- to the C-terminus) of the peptide. In the ^{1}H NMR spectrum, 12 methyl groups in the range of δ_H 0.67−0.99 ppm are characteristic for the isopropyl groups of six valines, as confirmed by HSQC and COSY data (see Figure 2). One phenyl group was identified on the basis of the typical chemical shifts of δ_H 7.24−7.28 with total integration of five protons in the ^{1}H NMR spectrum and typical chemical shifts of δ_C 128.2−138.3 in the ^{13}C NMR spectrum. Six singlets in the range of δ_H 2.29−3.15 with integration of three protons for each correspond to N-methyl groups of six amino acid building blocks. In addition, putrescine was identified as the C-terminal amine of **1** based on the signals of two N-methylene groups (δ_H 3.28−3.22 and δ_C 39.6, δ_H 2.95−2.91 and δ_C 40.4) and two methylene groups (δ_H 1.60−1.48 and δ_C 27.5, δ_H 1.69−1.60 and δ_C 26.1) combining COSY data (see Figure 2). The connectivity of these building blocks was determined by HMBC correlations (see Figure 2).

Compounds **2**−**7** were obtained as white solids. Their molecular formulas were determined to be $C_{45}H_{78}N_8O_7$, $C_{50}H_{89}N_9O_7$, $C_{46}H_{80}N_8O_7$, $C_{54}H_{96}N_{10}O_8$, $C_{55}H_{98}N_{10}O_8$, and $C_{51}H_{89}N_9O_8$, respectively, based on HR-MS analysis (see Table S1 in the Supporting Information). NMR data of **2**−**7** (see Tables 1 and 2) revealed their similar constitution to **1**. The only differences are the number of incorporated valines, the degree of N-methylation of valines, and the terminal

amines (putrescine for **1**, **3**, **5**, and **6**, ammonia for **2**, **4**, and **7**). Detailed analysis of MS^2 fragmentation pattern of **1**−**7** confirmed their structures (see Figure S1 in the Supporting Information).

To determine the absolute configurations of all amino acids, **1**−**7** were hydrolyzed and analyzed according to the advanced Marfey's method.[8] Thereby, the absolute configurations of all valines and phenylalanines in **1**−**7** were identified to be L (see Figure S2 in the Supporting Information).

RXP biosynthesis gene clusters (RXP-BGCs) from *Xenorhabdus* and *Photorhabdus* show high homology.[10] Detailed bioinformatic analysis showed the RXP-BGC from *X. innexi* encodes three monomodular nonribosomal peptide synthetases (NRPSs), namely, InxA, InxB, and InxC, that are all composed of a condensation (C), a adenylation (A), a methyltransferase (MT), and a thiolation (T) domain, and an additional C-terminal C domain in InxC (see Figure 3).[10] The amino acid specificity for A domains was predicted to activate valine for InxA and InxC, and phenylalanine for InxB.[10] Since no separate epimerization domain or dual condensation/epimerization domain was detected in InxABC, all amino acids in the RXPs might have L configuration, as was confirmed by the advanced Marfey's method (see Figure S2 in the Supporting Information). Previous work has shown that the chemical diversity of the RXPs results from a combination of iterative and flexible use of the monomodular NRPSs.[8,10] Similar situation might also exist in the biosynthesis of **1**−**7**, since there are 7 or 8 amino acids in **1**−**7**, yet only three modules. Based on the chemical structures of **1**−**7**, combined with amino acid specificity for A domains, the biosynthesis for **1**−**7** was proposed (see Figure 3). InxA might act iteratively with flexible methyltransferase activity, to catalyze the incorporation of first five or six N-methylvaline/valine to peptides **1**−**7** as previously described for Kj12B from *Xenorhabdus* KJ12.1 in RXP biosynthesis.[10]

The bioactivities of **1**−**7** against the causative agents of sleeping sickness (*Trypanosoma brucei rhodesiense*), Chagas disease (*Trypanosoma cruzi*), leishmaniasis (*Leishmania donovani*), and malaria (*Plasmodium falciparum*) were evaluated (see Table 3). All tested compounds showed strong activities against *T. brucei rhodesiense* and *P. falciparum*, with IC$_{50}$ values of 0.07−6.25 and 0.091−3.16 μM, respectively. Further analysis of the structure−activity relationships of the compounds revealed that the putrescine compounds **1**, **3**, **5**, and **6** are more active than the ammonia compounds **2**, **4**, and **7**, and, in fact, **5** and **6** showed an activity of ∼70 nM against *T.*

Table 3. Bioactivities of 1−7 against Different Protozoan Parasites and Their Cytotoxic Activities against L6 Cells[a]

RXP	IC$_{50}$ (μM)				
	T.b.r.	*T.c.*	*L.d.*	*P.f.*	L6
1	0.35 ± 0.02	25.54 ± 3.33	56.77 ± 1.64	0.71 ± 0.37	19.03 ± 0.00
2	2.24 ± 0.04	21.70 ± 1.90	>100	2.59 ± 0.61	30.36 ± 5.46
3	0.13 ± 0.03	16.00 ± 1.88	53.90 ± 0.86	0.23 ± 0.07	6.47 ± 0.14
4	6.25 ± 0.62	45.67 ± 5.20	28.47 ± 6.22	3.16 ± 0.60	34.36 ± 6.83
5	0.070 ± 0.010	16.18 ± 2.07	76.28 ± 0.00	0.25 ± 0.00	16.63 ± 0.44
6	0.075 ± 0.002	5.49 ± 0.82	55.28 ± 0.39	0.091 ± 0.012	5.72 ± 0.08
7	1.73 ± 0.06	4.31 ± 0.51	53.49 ± 3.82	0.87 ± 0.12	6.01 ± 0.30
ref	0.018 ± 0.007	2.84 ± 0.00	0.31 ± 0.07	0.006 ± 0.000	0.010 ± 0.000

[a]Abbreviations: RXP, rhabdopeptide/xenortide-like peptide; *T.b.r.*, *T. brucei rhodesiense*; *T.c.*, *T. cruzi*; *L.d.*, *L. donovani*; *P.f.*, *P. falciparum*. The positive reference is different for each target organism: melarsoprol for *T.b.r.*, benznidazole for *T.c.*, miltefosine for *L.d.*, chloroquine for *P.f.*, and podophyllotoxin for L6 cells.

brucei rhodesiense. Compounds **6** and **7** also exhibited good activities (IC_{50} 5.49 and 4.31 μM) against *T. cruzi.*

Compounds **1**−**7** were tested for their cytotoxic activities against rat skeletal myoblasts (L6 cells) (see Table 3). Although the putrescine compounds are more toxic than the corresponding ammonia compounds, which might weaken putrescine compounds as lead compounds, **5**, in particular, showed a very good selectivity index of >230 (IC_{50} L6 cells/ IC_{50} T.b.r.) against *T. brucei rhodesiense* and can be regarded as a screening hit.[11] Moreover, with efficient synthesis strategies in hand,[12] one also might now even further optimize RXPs, with respect to their terminal amines, which obviously can also be polar, as we show in this work.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.orglett.8b01975.

> Experimental section; HR-MS data, HPLC-MS analysis, determination of amino acid configuration and NMR spectra of **1**−**7**; NMR data of **2** in DMSO-d_6 (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: h.bode@bio.uni-frankfurt.de.

ORCID Ⓓ

Helge B. Bode: 0000-0001-6048-5909

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Mclean, A. R. D.; Ataide, R.; Simpson, J. A.; Beeson, J. G.; Fowkes, F. J. I. *Parasitology* **2015**, *142*, 999−1015.

(2) Cavalli, A.; Bolognesi, M. L. *J. Med. Chem.* **2009**, *52*, 7339−7359.

(3) Salas, P. F.; Herrmann, C.; Orvig, C. *Chem. Rev.* **2013**, *113*, 3450−3492.

(4) Bode, H. B. *Curr. Opin. Chem. Biol.* **2009**, *13*, 224−230.

(5) Shi, Y. M.; Bode, H. B. *Nat. Prod. Rep.* **2018**, *35*, 309−335.

(6) Lang, G.; Kalvelage, T.; Peters, A.; Wiese, J.; Imhoff, J. F. *J. Nat. Prod.* **2008**, *71*, 1074−1077.

(7) Crawford, J. M.; Portmann, C.; Kontnik, R.; Walsh, C. T.; Clardy, J. *Org. Lett.* **2011**, *13*, 5144−5147.

(8) Reimer, D.; Cowles, K. N.; Proschak, A.; Nollmann, F. I.; Dowling, A. J.; Kaiser, M.; Ffrench-Constant, R.; Goodrich-Blair, H.; Bode, H. B. *ChemBioChem* **2013**, *14*, 1991−1997.

(9) Reimer, D.; Nollmann, F. I.; Schultz, K.; Kaiser, M.; Bode, H. B. *J. Nat. Prod.* **2014**, *77*, 1976−1980.

(10) Cai, X.; Nowak, S.; Wesche, F.; Bischoff, I.; Kaiser, M.; Fürst, R.; Bode, H. B. *Nat. Chem.* **2017**, *9*, 379−386.

(11) Pink, R.; Hudson, A.; Mouriès, M. A.; Bendig, M. *Nat. Rev. Drug Discovery* **2005**, *4*, 727−740.

(12) Wesche, F.; Adihou, H.; Kaiser, A.; Wurglics, M.; Schubert-Zsilavecz, M.; Kaiser, M.; Bode, H. B. *J. Med. Chem.* **2018**, *61*, 3930−3938.

Supporting Information

Rhabdopeptide/Xenortide-like Peptides from *Xenorhabdus innexi* with Terminal Amines Showing Potent Antiprotozoal Activity

Lei Zhao,[1,2] Marcel Kaiser,[3,4] and Helge B. Bode*[1,5]

[1]Molecular Biotechnology, Department of Biosciences, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

[2]Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, 210014 Nanjing, China

[3]Parasite Chemotherapy, Swiss Tropical and Public Health Institute, 4051 Basel, Switzerland

[4]University of Basel, 4003 Basel, Switzerland

[5]Buchmann Institute for Molecular Life Sciences (BMLS), Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

*Corresponding author

Experimental Section

General experimental procedures

NMR spectra were recorded on a Bruker AV 500 spectrometer at 500 MHz (^{1}H) and 125 MHz (^{13}C) using deuterated MeOH or DMSO as solvent. HPLC-ESI-MS analysis was carried out on a Dionex UltiMate 3000 system coupled to a Bruker AmaZonX mass spectrometer. Compounds were eluted on an ACQUITY UPLC BEH C_{18} column (130 Å, 2.1 mm × 100 mm, 1.7 μm) using a CH_3CN/H_2O gradient containing 0.1% HCOOH (5−95% CH_3CN for 16 min) at a flow rate of 0.6 mL/min. A scan range of 100−1200 *m/z* in positive ionization mode was performed to detect compounds. HPLC-HR-ESI-MS were recorded on a Dionex UltiMate 3000 system coupled to a Bruker Impact II QTOF mass spectrometer using an ACQUITY UPLC BEH C_{18} RP Column (130 Å, 2.1 mm × 50 mm, 1.7 μm) and a same gradient at a flow rate of 0.4 mL/min for 16 min.

Strain cultivation

Xenorhabdus innexi DSM 16336 was cultivated in a LB medium at 30 °C. For HPLC-MS analysis, 100 μL of overnight culture was inoculated to 10 mL of fresh LB medium containing 2% of Amberlite XAD-16. The cultures were grown at 30 °C and 200 rpm. After 72 h, the XAD beads were separated and extracted with 10 mL of MeOH for 1 h. For isolation of RXPs, 4 L of cultures were cultivated using above culture conditions.

Extraction and purification

After 72 h incubation, the XAD-16 beads from 4 L of cultures were extracted with 300 mL of MeOH three times. The solvent was removed to obtain the MeOH extract. The crude extract was dissolved in H_2O and extracted with EtOAc. The EtOAc fraction (865.7 mg) was fractionated by Sephadex LH-20 chromatography using MeOH as the eluent. The fractions were analyzed by HPLC-MS, and the enriched fractions containing RXPs were used for further purification. The purification was carried out on an Agilent preparative HPLC system and a Cholester column (10 ID × 250 mm). The compounds were eluted by using a CH_3CN/H_2O gradient containing 0.1% HCOOH (25−45% CH_3CN for 40 min) at a flow rate of 3 mL/min to yield **1** (19.4 mg), **2** (21.5 mg), **3** (4.7 mg), **4** (3.0 mg), **5** (3.3 mg), **6** (7.2 mg), and **7** (3.0 mg).

92

Determination of amino acid configurations

Amino acid configurations of **1**−**7** were determined using the advanced Marfey's method as described previously.[1,2]

Bioactivity tests

Compounds **1**−**7** were tested against the parasites *Trypanosoma brucei rhodesiense* STIB900, *Trypanosoma cruzi* Tulahuen C4, *Leishmania donovani* MHOM-ET-67/L82, and *Plasmodium falciparum* NF54 in duplicates as described previously.[3] Their cytotoxic activities against the rat skeletal myoblasts (L6 cells) were conducted also in duplicates as described previously.[3]

92

Supplementary Tables

Table S1. HR-MS data of **1–7**

compound	sum formula	m/z [M+2H]$^{2+}$		m/z [M+H]$^+$		Δppm
		calcd	found	calcd	found	
1	$C_{49}H_{87}N_9O_7$	457.8437	457.8432			1.1
2	$C_{45}H_{78}N_8O_7$			843.6066	843.6032	4.0
3	$C_{50}H_{89}N_9O_7$	464.8515	464.8511			0.9
4	$C_{46}H_{80}N_8O_7$			857.6223	857.6183	4.7
5	$C_{54}H_{96}N_{10}O_8$	507.3779	507.3774			1.0
6	$C_{55}H_{98}N_{10}O_8$	514.3857	514.3850			1.4
7	$C_{51}H_{89}N_9O_8$			956.6907	956.6837	7.3

S4

Table S2. ¹H (600 MHz) and ¹³C (150 MHz) NMR Data of **2** in DMSO-d_6 (δ in ppm, J in Hz)

subunit	position	δ_C	δ_H
N-Me-L-Val (a)	1	173.0 (C)	
	2	69.4 (CH)	2.71, d (5.9)
	3	30.9 (CH)	1.78−1.68, m
	4	18.7 (CH₃)	0.84, overlap
	5	18.7 (CH₃)	0.84, overlap
	N-CH₃	34.7	2.15, s
L-Val (b)	1	172.4 (C)	
	2	53.6 (CH)	4.60, d (8.3)
	3	29.8 (CH)	1.99−1.90, m
	4	18.1 (CH₃)	0.86, overlap
	5	19.0 (CH₃)	0.86, overlap
	NH		8.07, d (8.6)
N-Me-L-Val (c)	1	169.5 (C)	
	2	57.4 (CH)	5.02, d (10.7)
	3	26.5 (CH)	2.23−2.12, m
	4	19.4 (CH₃)	0.81, overlap
	5	19.3 (CH₃)	0.72, overlap
	N-CH₃	29.9	3.01, s
N-Me-L-Val (d)	1	168.6 (C)	
	2	57.3 (CH)	4.87, d (11.3)
	3	26.4 (CH)	2.23−2.12, m
	4	19.1 (CH₃)	0.72, overlap
	5	17.7 (CH₃)	0.66, overlap
	N-CH₃	29.8	2.81, s
N-Me-L-Val (e)	1	168.3 (C)	
	2	57.8 (CH)	4.89, d (11.1)
	3	26.2 (CH)	2.12−2.02, m
	4	19.3 (CH₃)	0.73, overlap
	5	17.0 (CH₃)	0.55, d (6.5)
	N-CH₃	28.7	2.09, s
N-Me-L-Phe (f)	1	170.1 (C)	
	2	53.1 (CH)	5.85, d (7.7)
	3	34.2 (CH₂)	2.99, d (7.8)
	4	137.2 (C)	
	5	129.3 (CH)	7.22, overlap
	6	128.0 (CH)	7.24, overlap
	7	126.3 (CH)	7.22, overlap
	8	128.0 (CH)	7.24, overlap
	9	129.3 (CH)	7.22, overlap
	N-CH₃	29.7	2.83, s
N-Me-L-Val (g)	1	171.3 (C)	
	2	61.2 (CH)	4.56, d (10.9)
	3	26.2 (CH)	2.12−2.02, m
	4	19.6 (CH₃)	0.90, d (6.3)
	5	18.4 (CH₃)	0.70, overlap
	N-CH₃	30.3	2.91, s
NH₂ (B)			7.31, s; 7.06, s

S5

94

Supplementary Figures

1 mV-V-mV-mV-mV-mF-mV-Put

m/z 457.77 [M + 2H]$^{2+}$
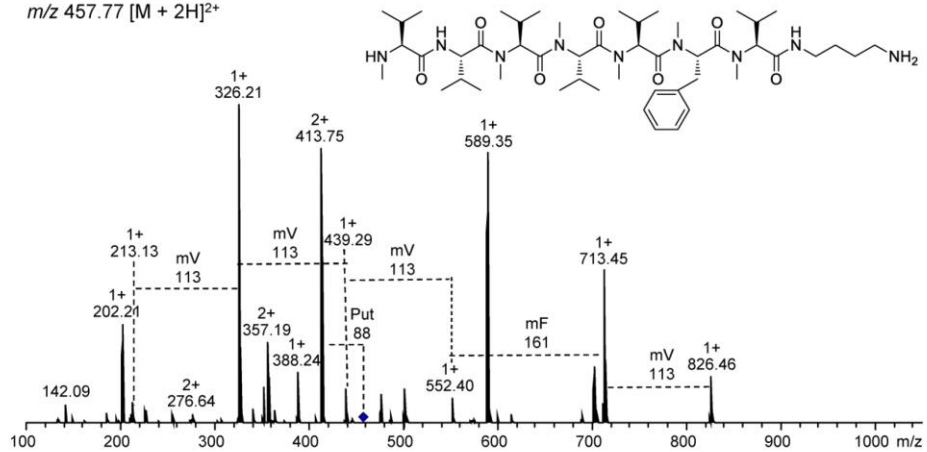
2 mV-V-mV-mV-mV-mF-mV-NH$_3$

m/z 843.32 [M + H]$^{+}$

S6
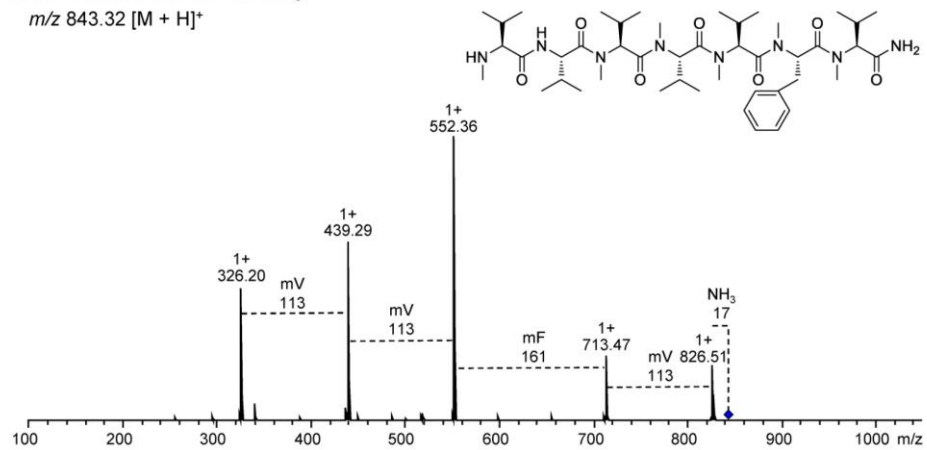
3 mV-mV-mV-mV-mV-mF-mV-Put

m/z 464.79 [M + 2H]$^{2+}$
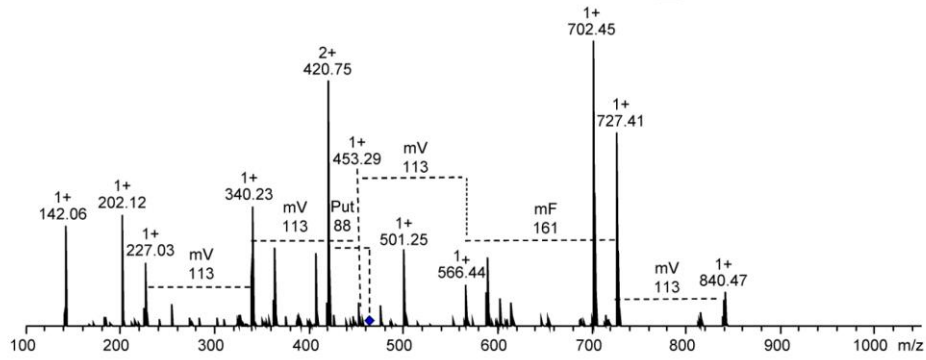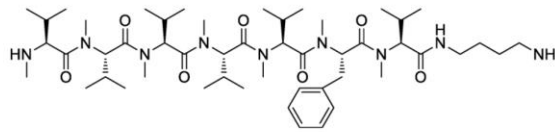
4 mV-mV-mV-mV-mV-mF-mV-NH$_3$

m/z 857.39 [M + H]$^+$

S7
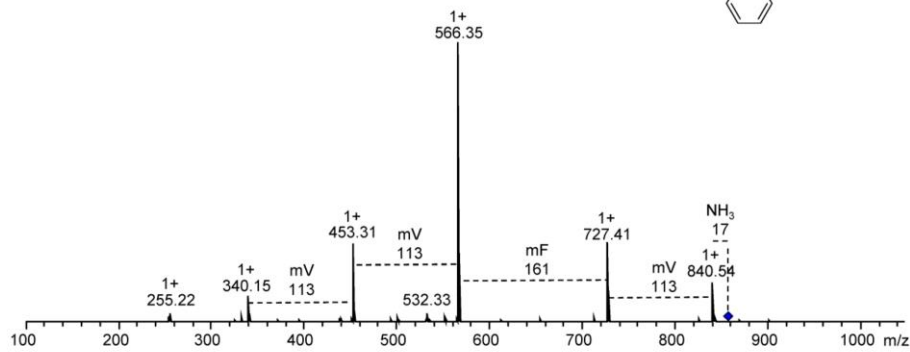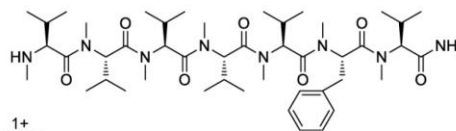
5 mV-V-mV-mV-V-mV-mF-mV-Put

m/z 507.32 [M + 2H]$^{2+}$
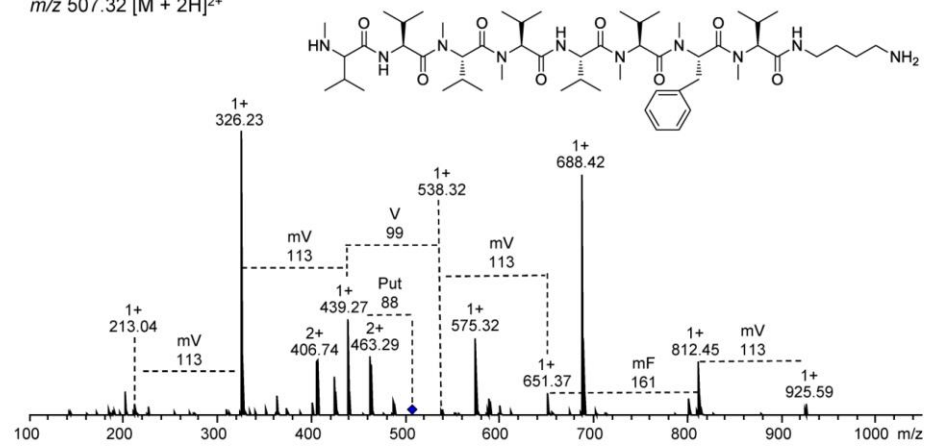
6 mV-V-mV-mV-mV-mV-mF-mV-Put

m/z 514.31 [M + 2H]$^{2+}$

S8
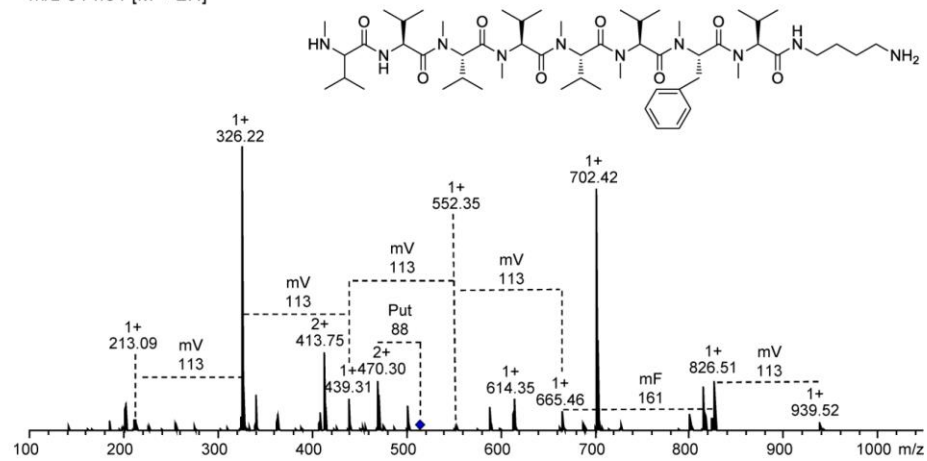
7 mV-V-mV-mV-mV-mV-mF-mV-NH$_3$
m/z 956.43 [M + H]$^+$
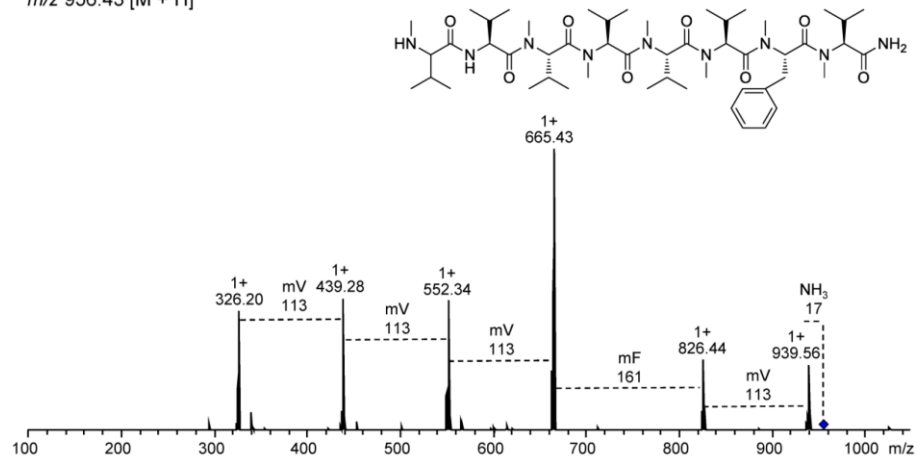
Figure S1. MS^2 fragmentation pattern of **1**−**7**.

S9
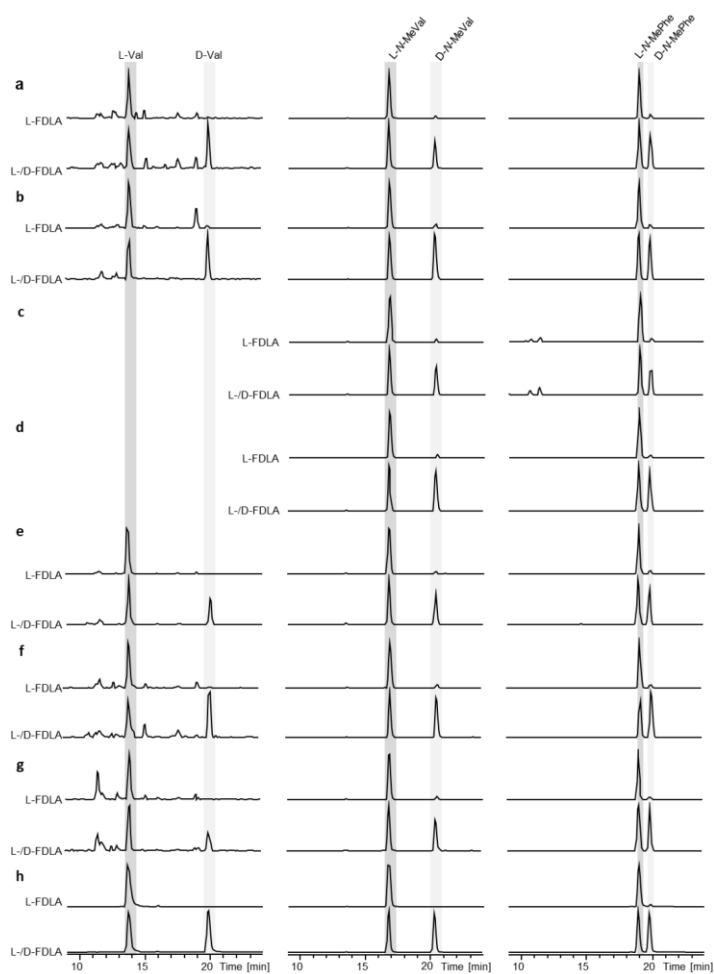
Figure S2. Determination of amino acid configuration for **1−7** using the advanced Marfey's method. (a)-(g) HPLC-MS analysis of hydrolyzed **1−7** and derivatized with L-FDLA (upper) and L-/D-FDLA (lower) in comparison to (h) standard amino acids. Depicted are extracted ion chromatogram (EIC) traces for valine (Val, *m/z* 412 [M+H]⁺), *N*-methylvaline (*N*-MeVal, *m/z* 426 [M+H]⁺), and *N*-methylphenylalanine (*N*-MePhe, *m/z* 474 [M+H]⁺). The configuration is determined by the elution order and L-FDLA derivatized amino acids elute prior to its D-enantiomer.

Figure S3. ¹H NMR (500 MHz, methanol-d_4) spectrum of **1**.

Figure S4. ¹³C NMR (125 MHz, methanol-d_4) spectrum of **1**.
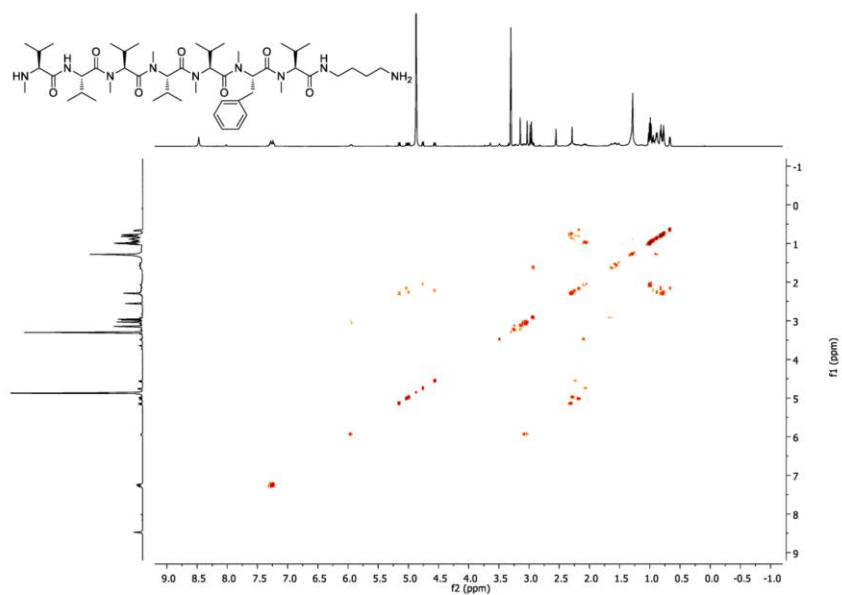
Figure S5. COSY (methanol-*d*₄) spectrum of **1**.
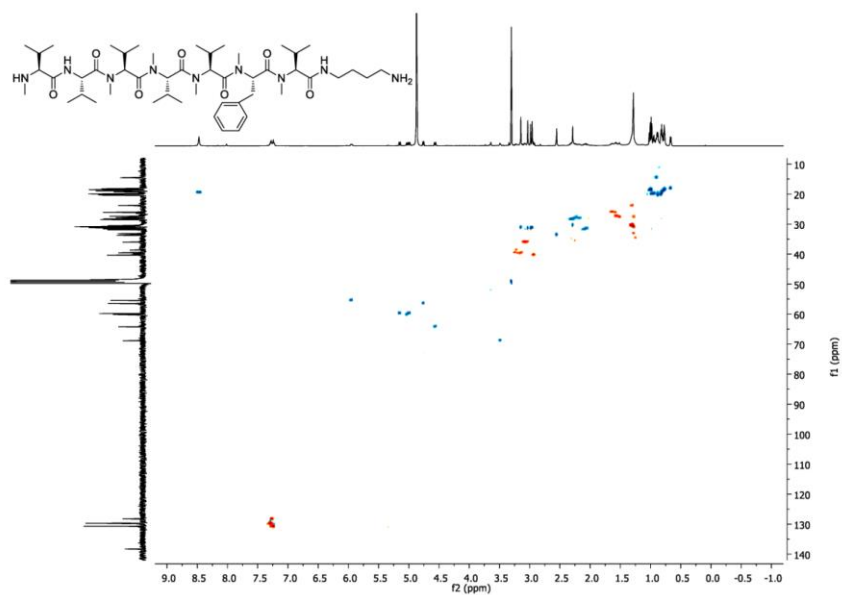
Figure S6. HSQC (methanol-*d*₄) spectrum of **1**.

S12

101
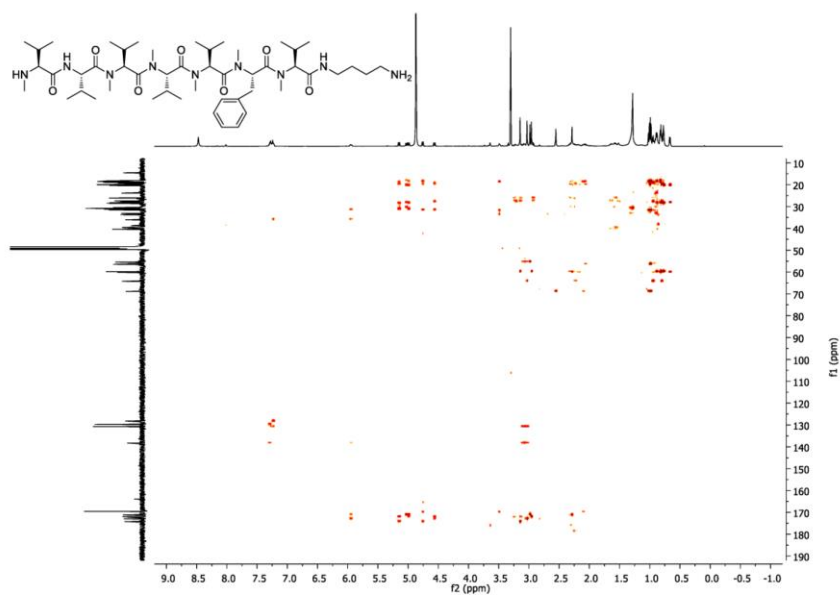
Figure S7. HMBC (methanol-d_4) spectrum of **1**.
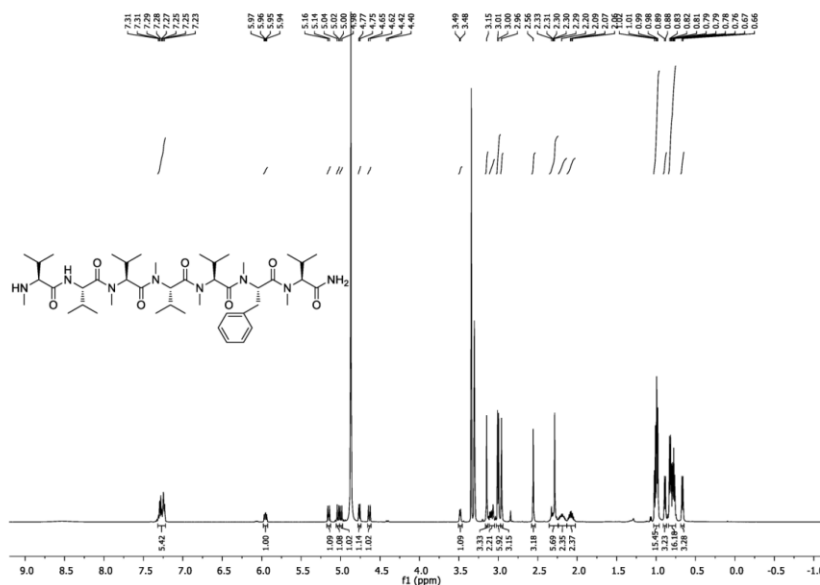
Figure S8. ^{1}H NMR(500 MHz, methanol-d_4) spectrum of **2**.
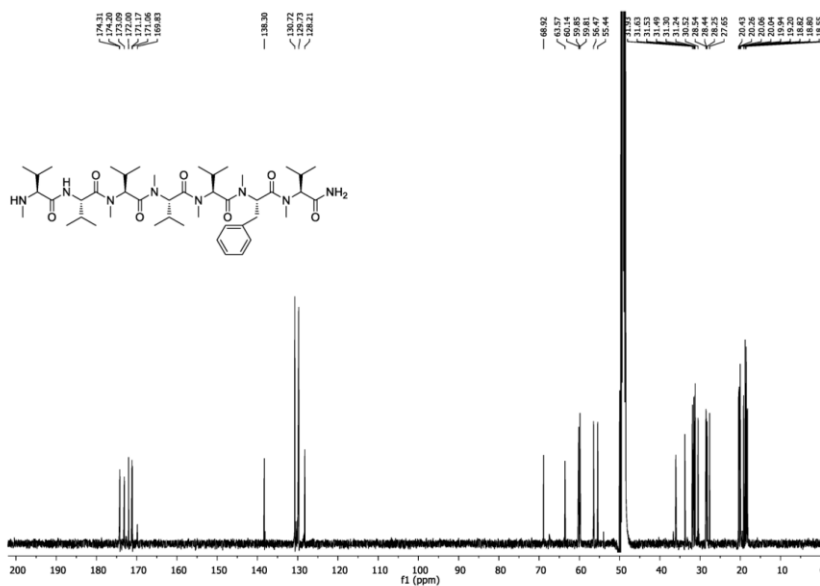
Figure S9. ^{13}C NMR (125 MHz, methanol-d_4) spectrum of **2**.
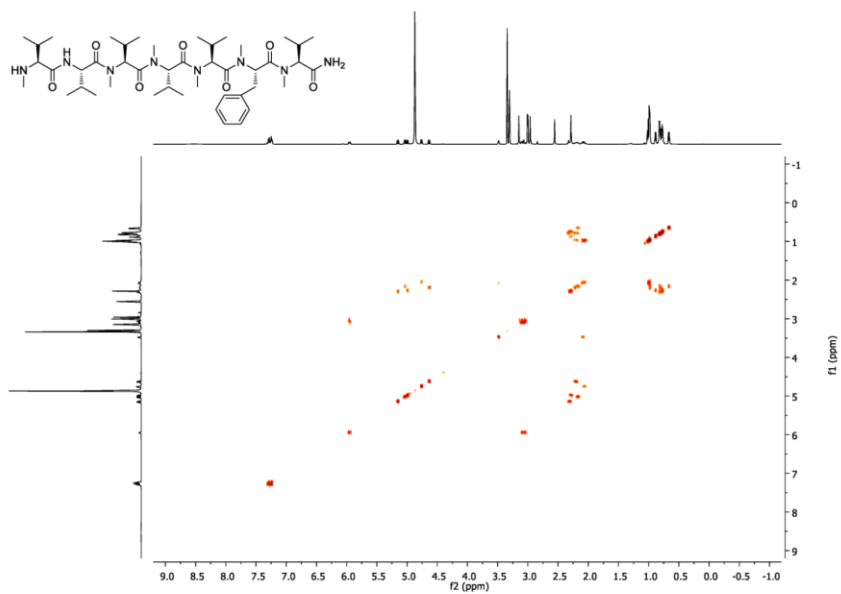
Figure S10. COSY (methanol-*d*₄) spectrum of **2**.
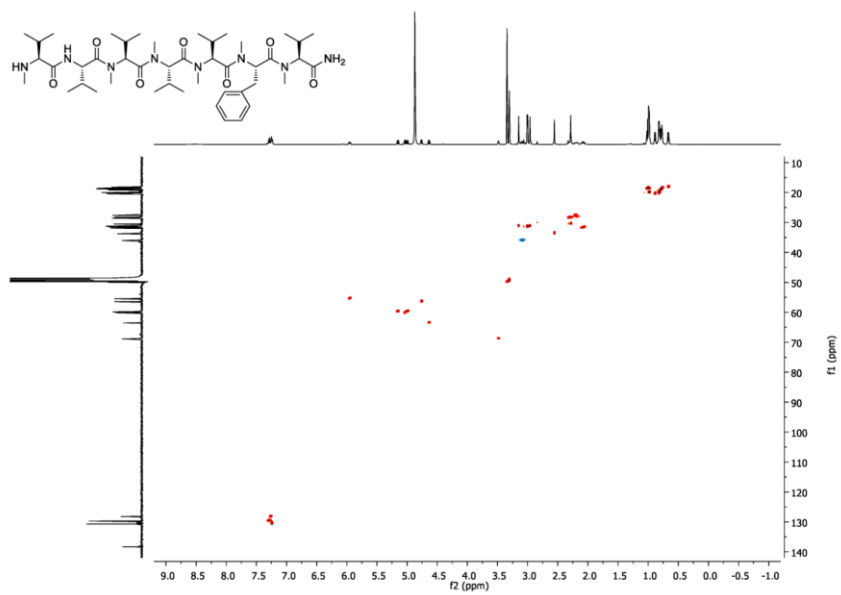
Figure S11. HSQC (methanol-*d*₄) spectrum of **2**.

S15

104
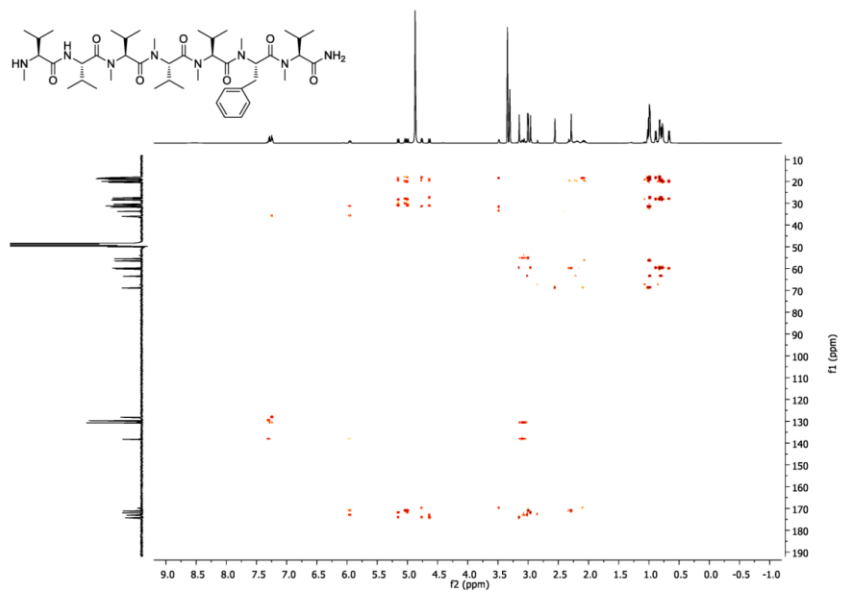
Figure S12. HMBC (methanol-*d*₄) spectrum of **2**.

Figure S13. ^{1}H NMR (600 MHz, DMSO-d_6) spectrum of **2**.

Figure S14. ^{13}C NMR (150 MHz, DMSO-d_6) spectrum of **2**.
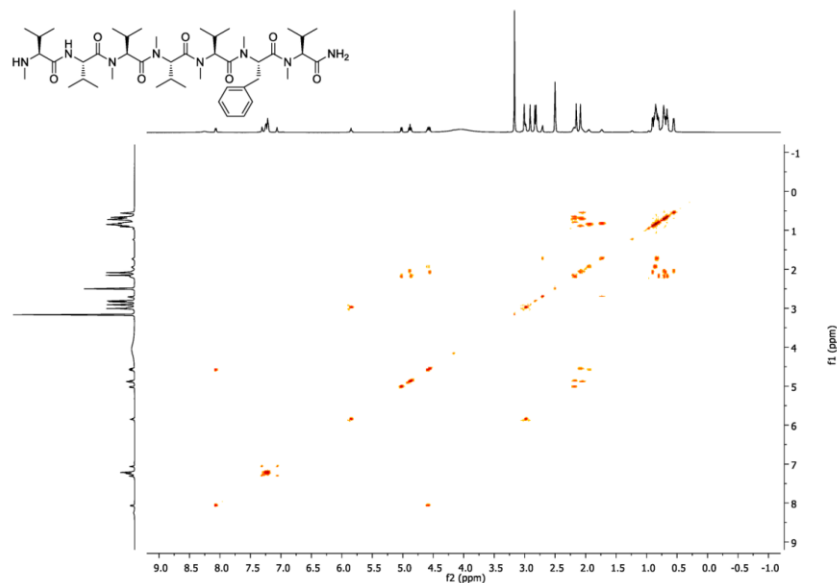
Figure S15. COSY (DMSO-*d*6) spectrum of **2**.
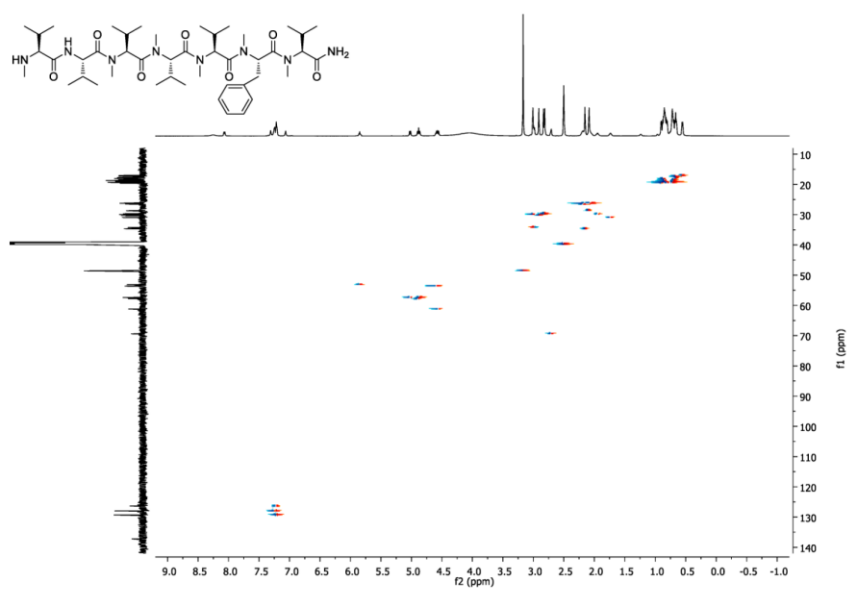
Figure S16. HSQC (DMSO-*d*6) spectrum of **2**.

S18

107
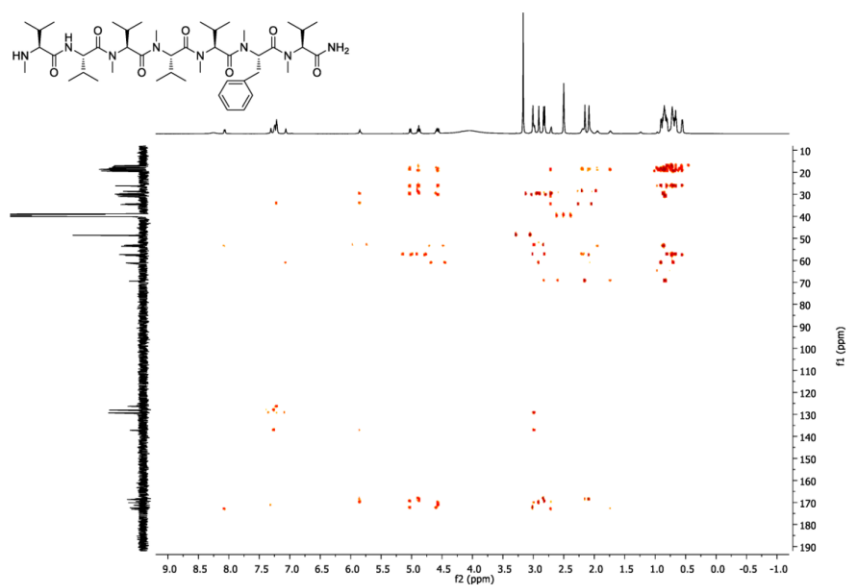
Figure S17. HMBC (DMSO-d_6) spectrum of **2**.

Figure S18. ^{1}H NMR (500 MHz, methanol-d_4) spectrum of **3**.

Figure S19. ^{13}C NMR (125 MHz, methanol-d_4) spectrum of **3**.

S20

109
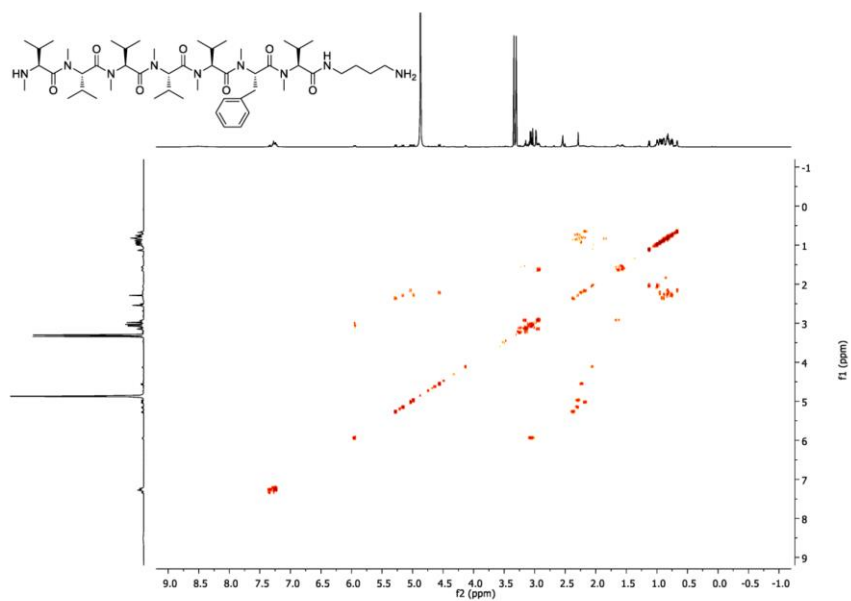
Figure S20. COSY (methanol-*d*4) spectrum of **3**.
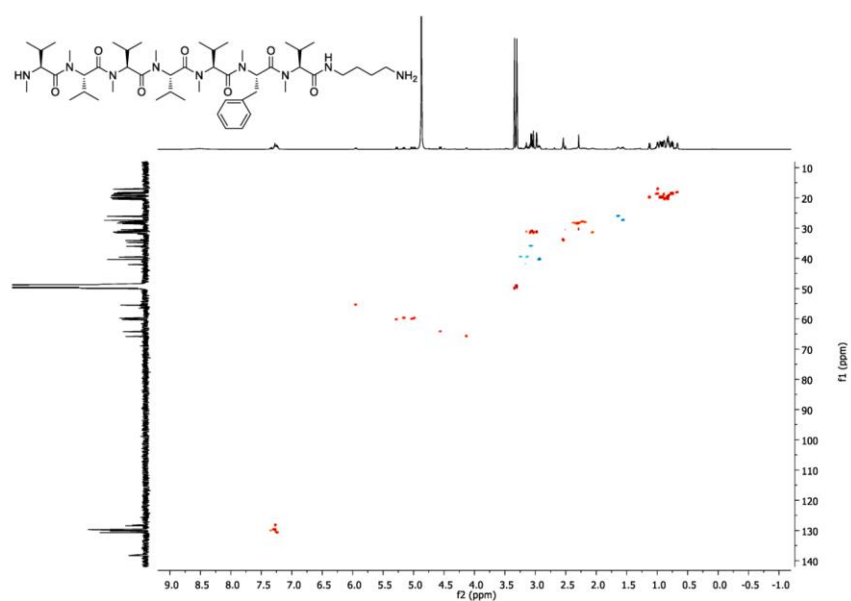
Figure S21. HSQC (methanol-*d*4) spectrum of **3**.

S21

110
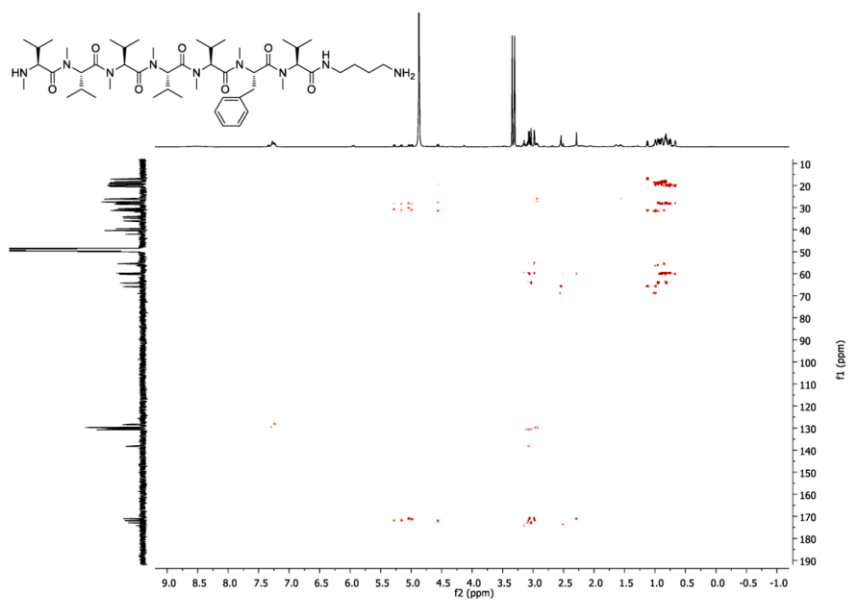
Figure S22. HMBC (methanol-*d*₄) spectrum of **3**.

Figure S23. ¹H NMR (500 MHz, methanol-d_4) spectrum of **4**.

Figure S24. ¹³C NMR (125 MHz, methanol-d_4) spectrum of **4**.

S23
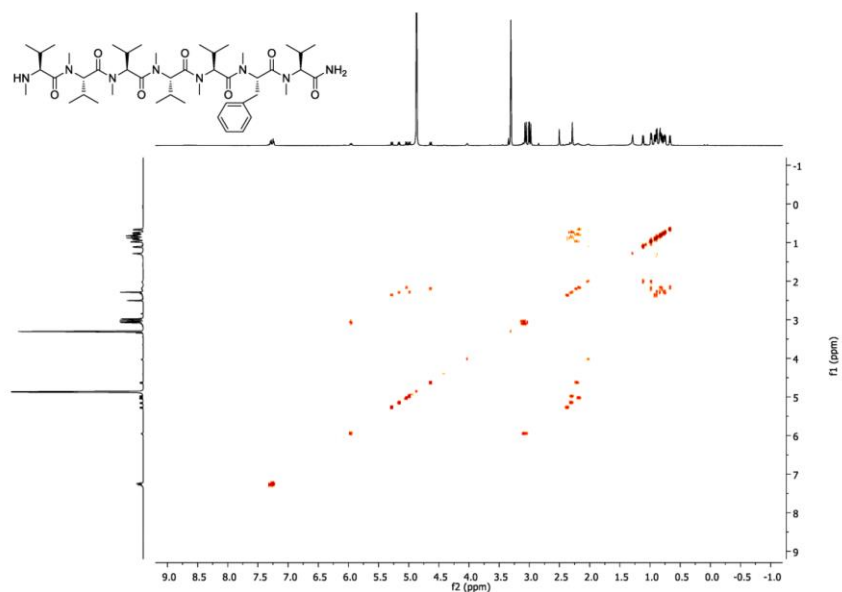
Figure S25. COSY (methanol-*d*4) spectrum of **4**.
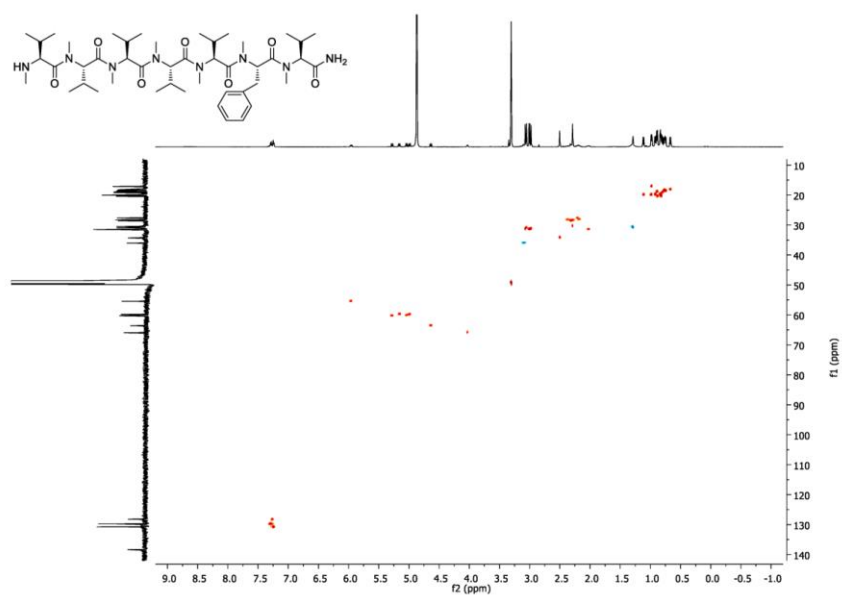
Figure S26. HSQC (methanol-*d*4) spectrum of **4**.

S24

113
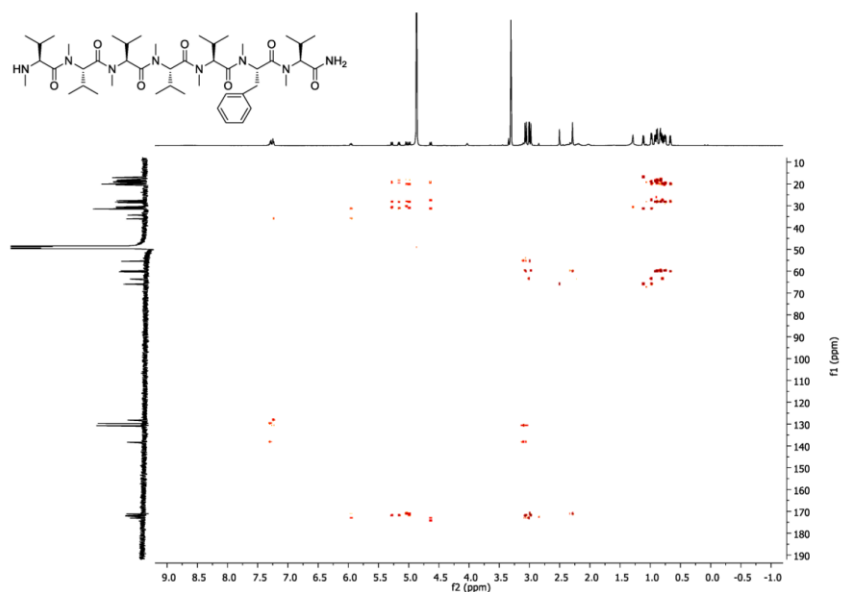
Figure S27. HMBC (methanol-d_4) spectrum of **4**.

Figure S28. ^{1}H NMR (500 MHz, methanol-d_4) spectrum of **5**.

Figure S29. ^{13}C NMR (125 MHz, methanol-d_4) spectrum of **5**.
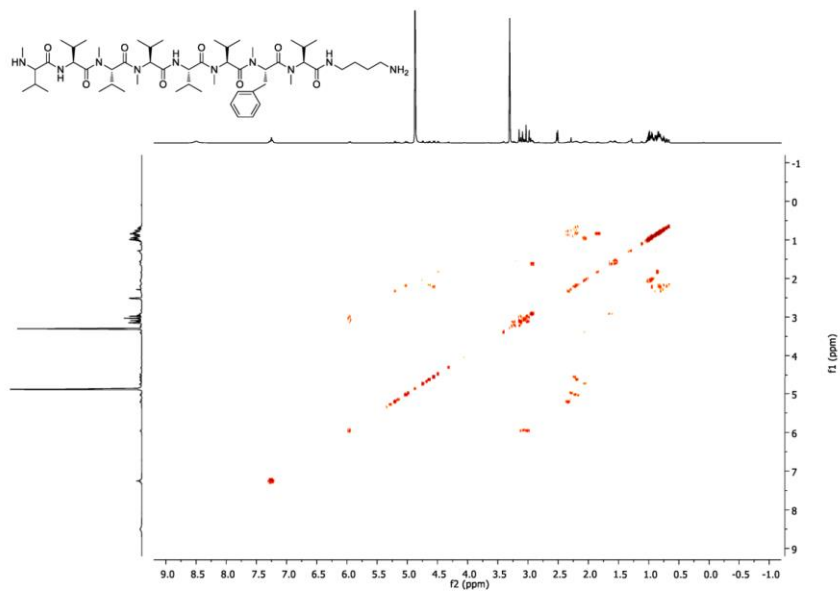
Figure S30. COSY (methanol-*d*₄) spectrum of **5**.
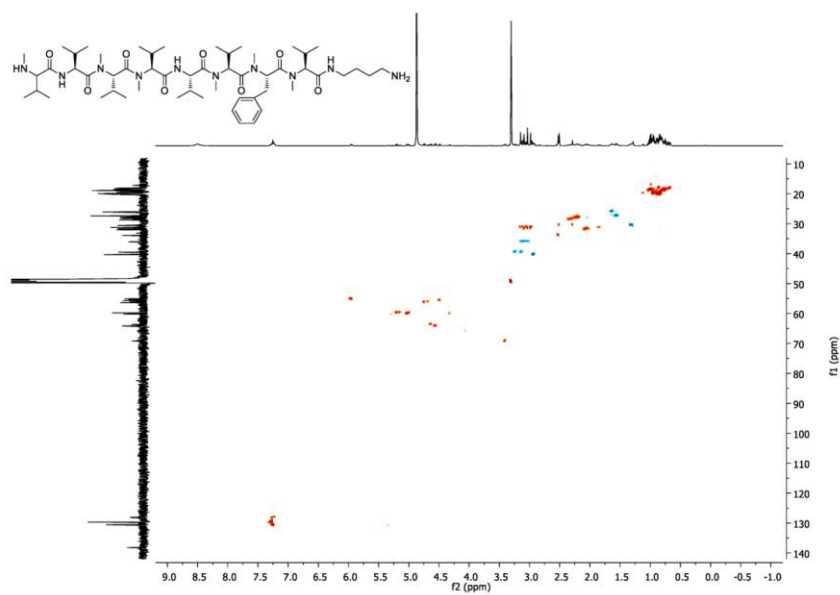
Figure S31. HSQC (methanol-*d*₄) spectrum of **5**.

S27

116
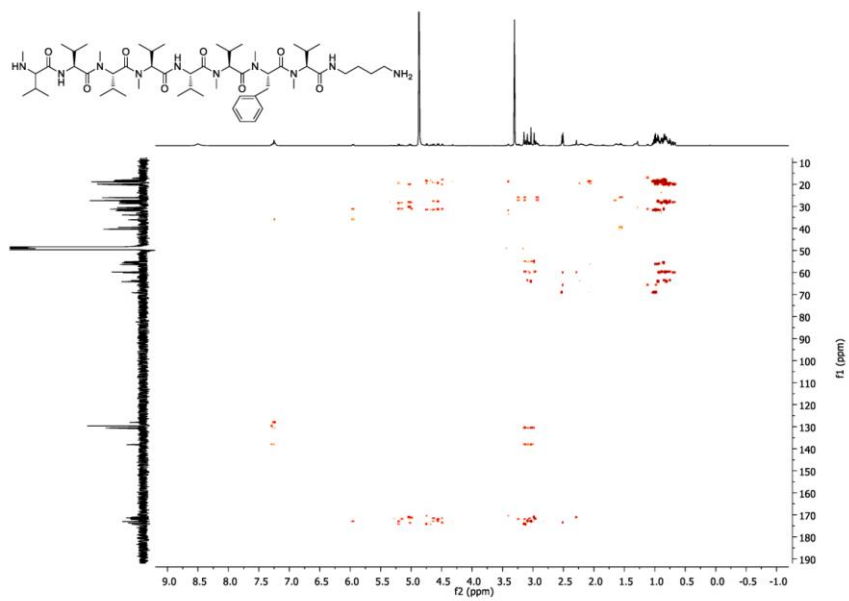
Figure S32. HMBC (methanol-d_4) spectrum of **5**.
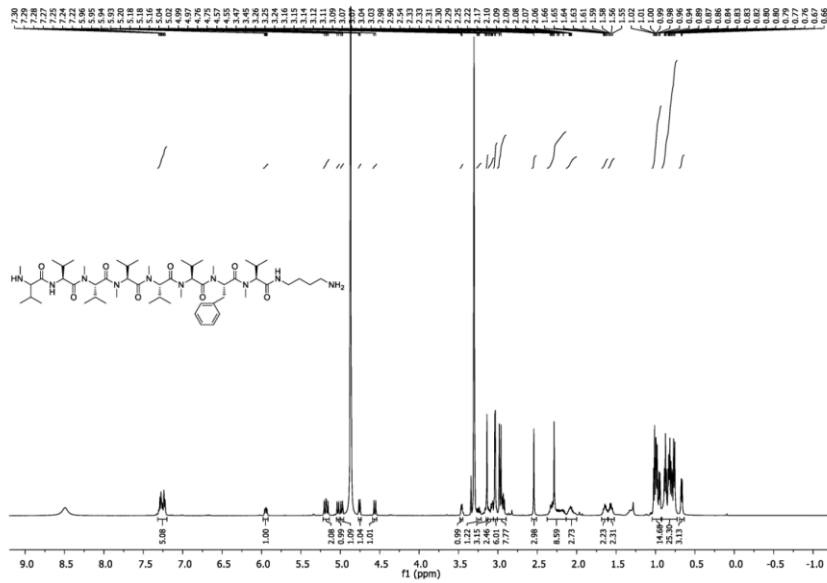
Figure S33. ^{1}H NMR (methanol-d_4, 500 MHz) spectrum of **6**.
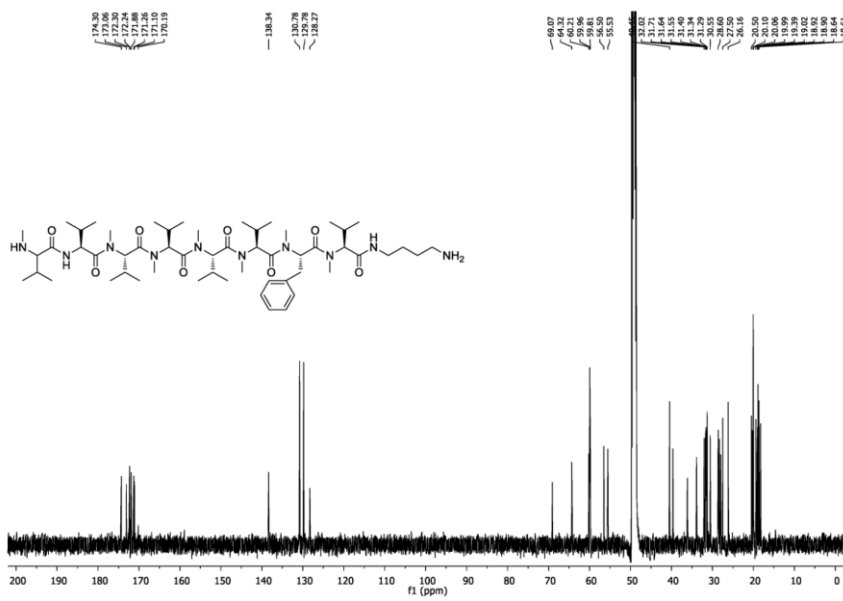
Figure S34. ^{13}C NMR (methanol-d_4, 125 MHz) spectrum of **6**.

S29
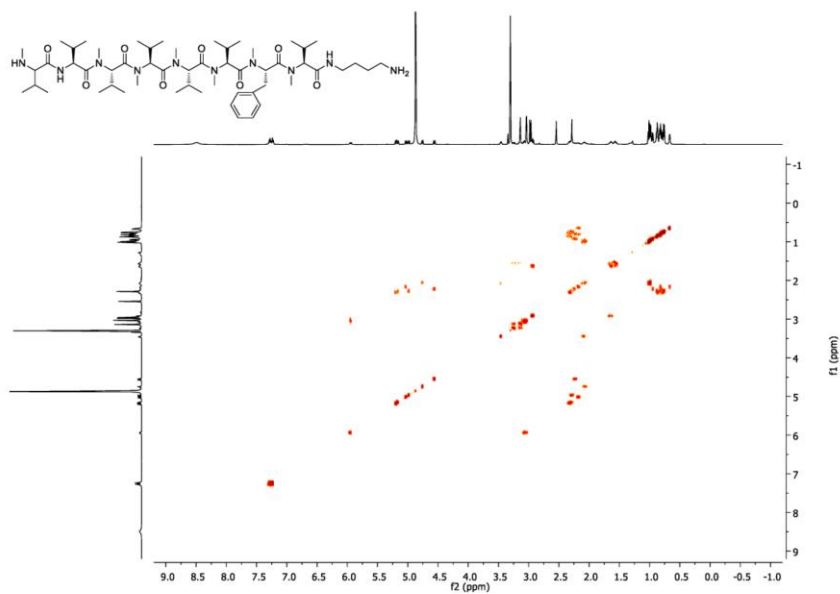
Figure S35. COSY (methanol-d_4) spectrum of **6**.
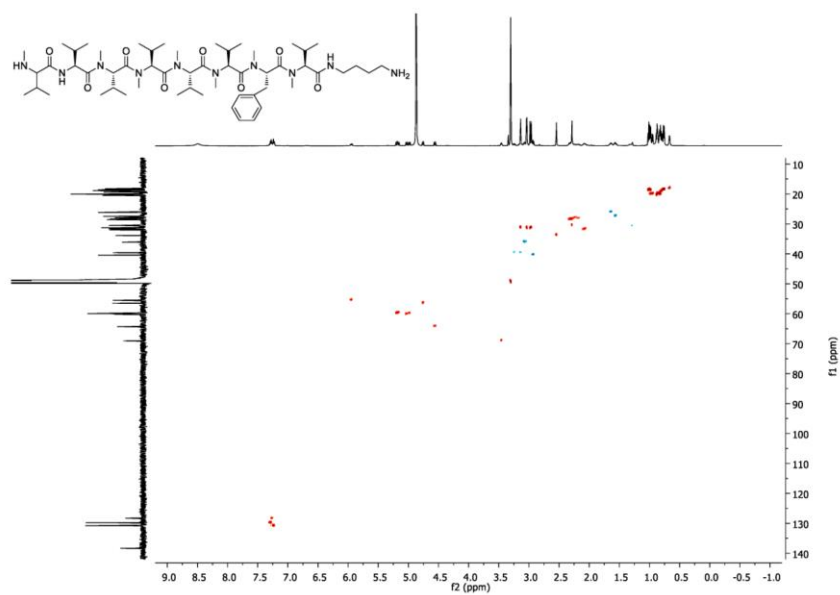
Figure S36. HSQC (methanol-d_4) spectrum of **6**.

S30

119
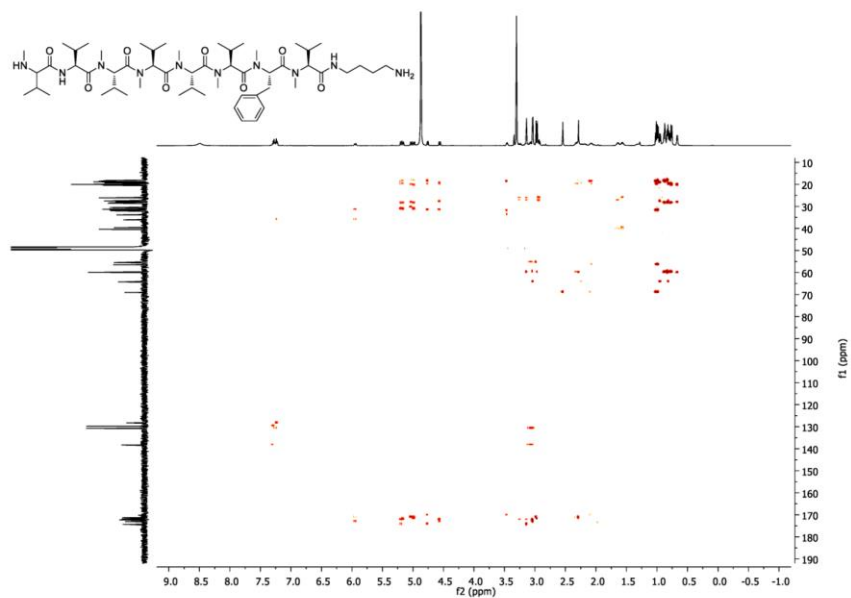
Figure S37. HMBC (methanol-d_4) spectrum of **6**.
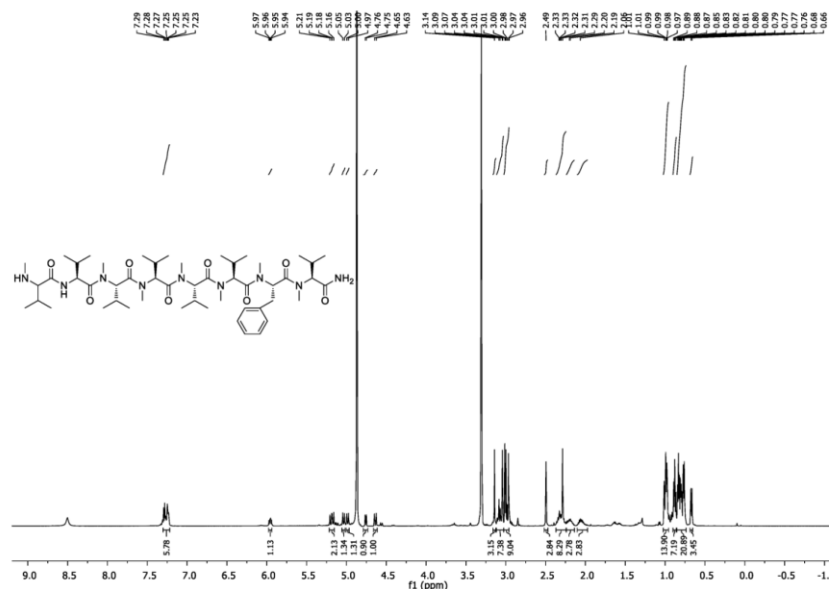
Figure S38. ¹H NMR (500 MHz, methanol-*d₄*) spectrum of **7**.
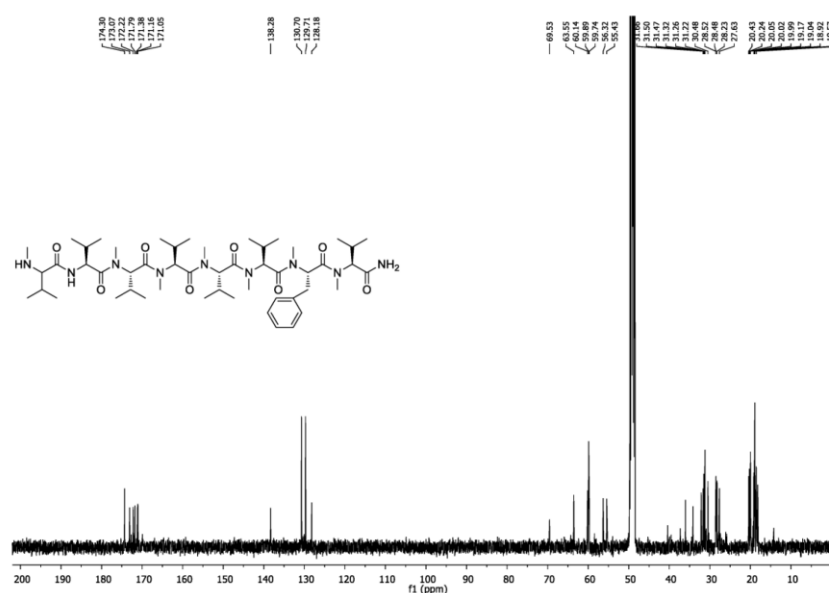
Figure S39. ¹³C NMR (125 MHz, methanol-*d₄*) spectrum of **7**.

S32

121
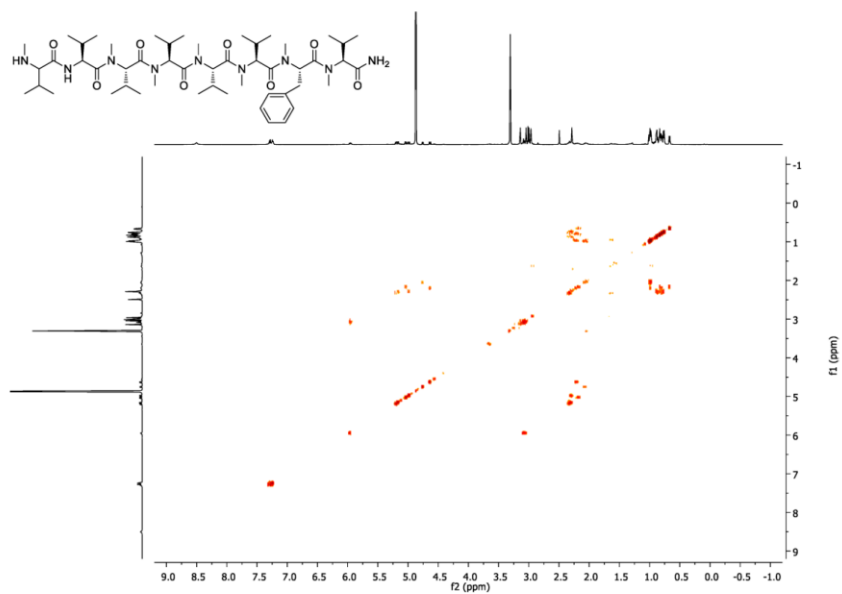
Figure S40. COSY (methanol-*d*4) spectrum of **7**.
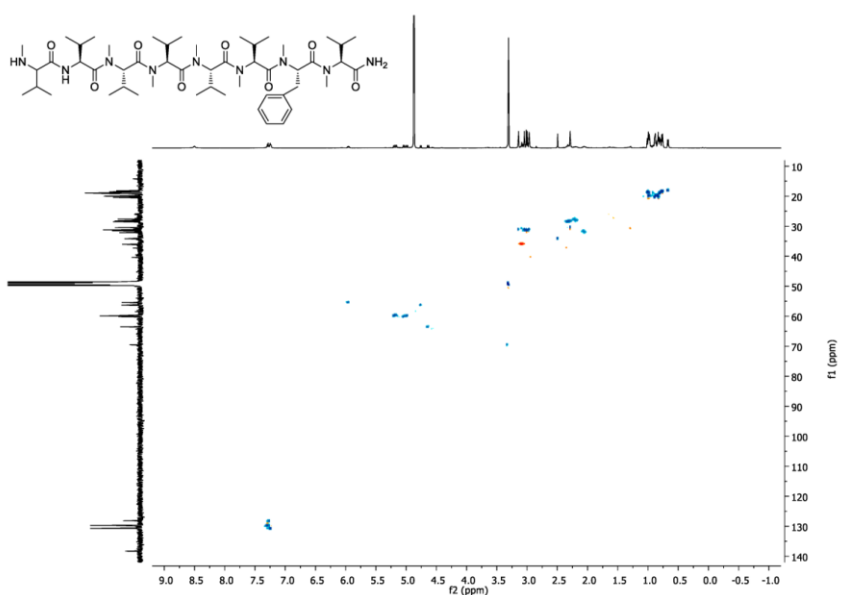
Figure S41. HSQC (methanol-*d*4) spectrum of **7**.

S33

122
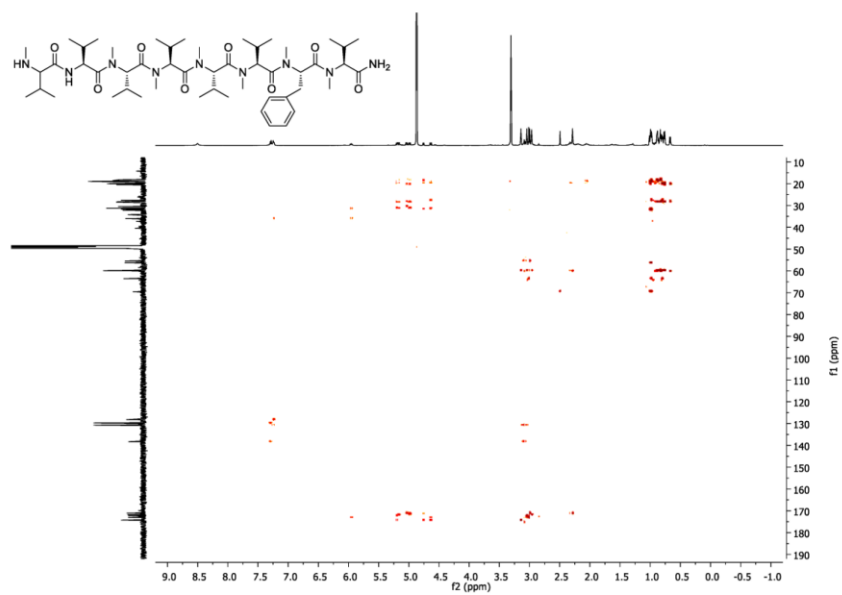
Figure S42. HMBC (methanol-*d*4) spectrum of **7**.

References

(1) Reimer, D.; Cowles, K. N.; Proschak, A.; Nollmann, F. I.; Dowling, A. J.; Kaiser, M.; Ffrench-Constant, R.; Goodrich-Blair, H.; Bode, H. B. *ChemBiochem* **2013**, *14*, 1991–1997.

(2) Grundmann, F.; Kaiser, M.; Kurz, M.; Schiell, M.; Batzer, A.; Bode, H. B. *RSC Adv.* **2013**, *3*, 22072–22077.

(3) Orhan, I.; Sener, B.; Kaiser, M.; Brun, R.; Tasdemir, D. *Mar. Drugs* **2010**, *8*, 47–58.

5.2 Methionine-Containing Rhabdopeptide/Xenortide-like Peptides from Heterologous Expression of the Biosynthetic Gene Cluster *kj12ABC* in *Escherichia coli*

Authors:

Lei Zhao,[1,2] Xiaofeng Cai,[1] Marcel Kaiser,[3,4] and Helge B. Bode*[,1,5]

[1]Molecular Biotechnology, Department of Biosciences, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

[2]Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, 210014 Nanjing, China

[3]Parasite Chemotherapy, Swiss Tropical and Public Health Institute, 4051 Basel, Switzerland

[4]University of Basel, 4003 Basel, Switzerland

[5]Buchmann Institute for Molecular Life Sciences (BMLS), Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

*Corresponding author

Online access:

https://pubs.acs.org/doi/10.1021/acs.jnatprod.8b00425

Declaration on the contribution of the authors

Publication: Methionine-Containing Rhabdopeptide/Xenortide-like Peptides from Heterologous Expression of the Biosynthetic Gene Cluster *kj12ABC* in *Escherichia coli*

Status: published

Journal: Journal of Natural Products

Authors: <u>Lei Zhao (LZ)</u>, Xiaofeng Cai (XC), Marcel Kaiser (MK), Helge B. Bode (HBB)

What did the doctoral candidate or the co-authors contribute individually to the dissertation?

(1) Development and planning

<u>LZ (40%)</u>, XC (20%), HBB (40%)

(2) Performance of individual research and experiments

<u>LZ (60%): strain construction, extract preparation, HPLC-MS analysis, chemical synthesis, NMR measurement, product quantification;</u> XC (30%): strain construction, HPLC-MS analysis; MK (10%): bioactivity testing

(3) Collection of data and preparation of figures

<u>LZ (60%): collection of NMR data, collection of HPLC-MS data including quantification data, preparation of figures;</u> XC (30%): collection of HPLC-MS data, preparation of figures; MK (10%): collection of bioactivity data

(4) Analysis and interpretation of data

<u>LZ (70%): analysis of NMR data, analysis of HPLC-MS data including quantification data, interpretation of bioactivity data;</u> XC (30%): analysis of HPLC-MS data

(5) Preparation of manuscript

<u>LZ (80%)</u>, HBB (20%)

Herewith approving the indications above

_____   _____
date/place   signature doctoral candidate

_____   _____
date/place   signature supervisor

_____   _____
date/place   If necessary, signature corresponding author

Methionine-Containing Rhabdopeptide/Xenortide-like Peptides from Heterologous Expression of the Biosynthetic Gene Cluster *kj12ABC* in *Escherichia coli*

Lei Zhao,[†,‡] Xiaofeng Cai,[†] Marcel Kaiser,[§,⊥] and Helge B. Bode[*,†,‖]

[†]Molekulare Biotechnologie, Fachbereich Biowissenschaften, Goethe Universität Frankfurt, 60438 Frankfurt am Main, Germany

[‡]Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, 210014 Nanjing, China

[§]Parasite Chemotherapy, Swiss Tropical and Public Health Institute, 4051 Basel, Switzerland

[⊥]University of Basel, 4003 Basel, Switzerland

[‖]Buchmann Institute for Molecular Life Sciences (BMLS), Goethe Universität Frankfurt, 60438 Frankfurt am Main, Germany

Ⓢ *Supporting Information*

ABSTRACT: Seven new methionine-containing rhabdopeptide/xenortide-like peptides (1−7) were identified from *Escherichia coli* expressing the rhabdopeptide/xenortide-like peptide biosynthetic gene cluster *kj12ABC* from *Xenorhabdus* KJ12.1. Their structures were elucidated by detailed HPLC-HR-MS/MS analysis and confirmed by chemical synthesis. Bioactivity tests of these first rhabdopeptide/xenortide-like peptide derivatives (2−7) showing methionine building blocks compared to the usually found derivatives containing exclusively hydrophobic amino acids such as valine, leucine, or phenylalanine revealed good activities of 2−7 against protozoan parasites and no cytotoxicity against mammalian L6 cells.
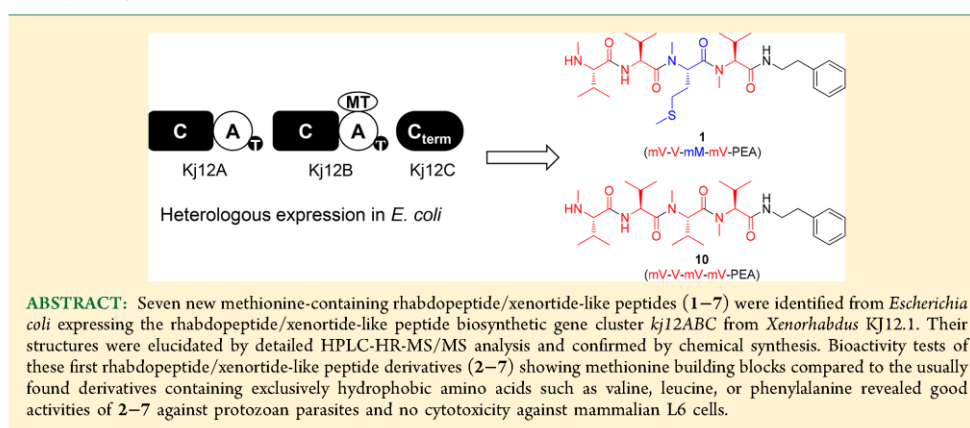
Rhabdopeptide/xenortide-like peptides are a unique class of peptides derived from non-ribosomal peptide synthetases (NRPSs) exclusively found in entomopathogenic bacteria of the genera *Xenorhabdus* and *Photorhabdus* that live in symbiosis with nematodes of the genera *Steinernema* and *Heterorhabditis*, respectively.[1−6] More than 60 rhabdopeptide/xenortide-like peptides have been identified in 51 different wild-type strains by detailed HPLC-MS analysis from our in-house strain collection.[5] These derivatives are composed of two to eight amino acids that are often *N*-methylated and show phenylethylamine, tryptamine, or other amines at the C-terminus.[5] For the amino acids previously identified in rhabdopeptide/xenortide-like peptides, only valine, leucine, and phenylalanine have been observed.[1−5] Previous studies showed that rhabdopeptide/xenortide-like peptides are insect-specific virulence factors, which probably contribute to insect killing and also protect the insect cadaver from food competitors.[3,7]

Rhabdopeptide/xenortide-like peptide biosynthesis gene clusters (RXP-BGCs) from *Xenorhabdus* and *Photorhabdus* show high homology.[5] The RXP-BGC from *Xenorhabdus* KJ12.1, named as *kj12ABC*, is the simplest and best-characterized system.[5] It encodes three monomodular NRPSs, namely Kj12A, Kj12B, and Kj12C.[5] Kj12A and Kj12B are all composed of a condensation (C), an adenylation (A), and a thiolation (T) domain, and an additional methyltransferase (MT) domain in Kj12B; Kj12C is a stand-alone terminal C domain (Figure 1a).[5] Chemical diversity of the rhabdopeptide/xenortide-like peptides in *Xenorhabdus* KJ12.1 results from a combination of iterative and flexible use of Kj12ABC.[5] Previous studies have shown that heterologous expression of *kj12ABC* in *Escherichia coli* led to the same production of rhabdopeptide/xenortide-like peptides as found in the wild-type strain of *Xenorhabdus* KJ12.1.[5] However, during our in-depth investigation on rhabdopeptide/xenortide-like peptide production in *E. coli* carrying *kj12ABC*, we found seven new methionine-containing rhabdopeptide/xenortide-like peptides as minor products (Figure 1b,c, 1−7) besides the valine-only-containing main compounds (Figure 1b,c, 8−12).[5]

Attachments
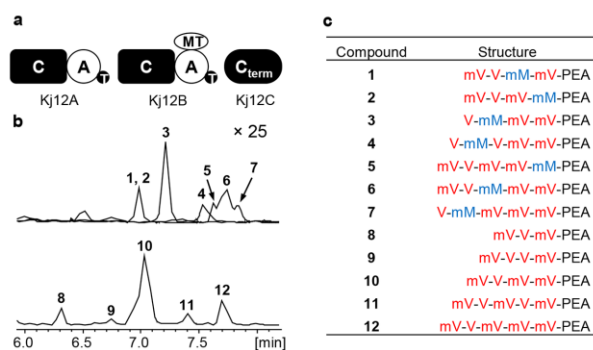
Figure 1. Rhabdopeptide/xenortide-like peptides from heterologous expression of Kj12ABC in *E. coli*. (a) Kj12ABC from *Xenorhabdus* KJ12.1: condensation (C), adenylation (A), methyltransferase (MT), thiolation (T), and terminal condensation (C_{term}) domains. (b) HPLC-HR-MS chromatograms for the produced rhabdopeptide/xenortide-like peptides **1–12**. Extracted ion chromatograms (EICs) for methionine-containing rhabdopeptide/xenortide-like peptides **1–7** (intensity increased 25-fold). Base-peak chromatogram (BPC) for valine-only-containing main compounds **8–12**. (c) Structures of **1–12**. Abbreviations: *N*-methylated valine (mV), valine (V), *N*-methylated methionine (mM), and phenylethylamine (PEA).
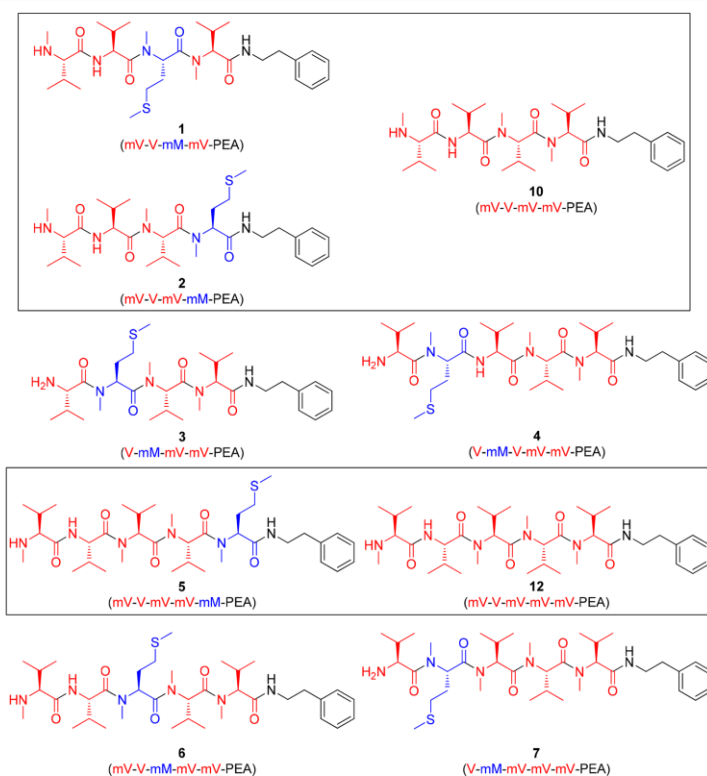
Figure 2. Structures of compounds **1–7**, **10**, and **12**.

2293

128

Scheme 1. Stepwise Solid-Phase Synthesis, Shown for 1 as an Example[a]
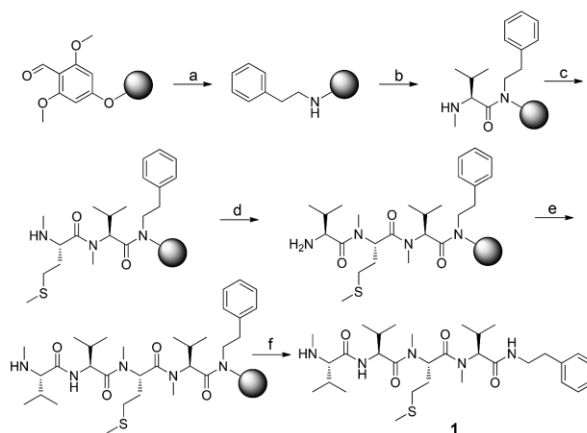
[a]Reagents and conditions: (a) PEA (10 equiv), DMF/MeOH/AcOH (80:19:1), NaBH$_3$CN (10 equiv), 60 °C, overnight; (b) Fmoc-*N*-Me-Val-OH (10 equiv), HATU (10 equiv), HOAt (10 equiv), DIPEA (20 equiv), DMF, overnight, then 20% piperidine/DMF; (c) Fmoc-*N*-Me-Met-OH (3 equiv), BTC (1.15 equiv), collidine (10 equiv), DIPEA (8 equiv), THF, 2 h, then 20% piperidine/DMF; (d) Fmoc-Val-OH (3 equiv), HATU (3 equiv), HOAt (3 equiv), DIPEA (6 equiv), DMF, overnight, then 20% piperidine/DMF; (e) Fmoc-*N*-Me-Val-OH (3 equiv), BTC (1.15 equiv), collidine (10 equiv), DIPEA (8 equiv), THF, 2 h, then 20% piperidine/DMF; (f) TFA/TIS/water (95:2.5:2.5), 2 h.

Here, we describe the structure elucidation of 1−7 by HPLC-HR-MS/MS analysis and chemical synthesis, as well as biological activities of 2−7 against protozoan parasites and cytotoxicity against mammalian L6 cells.

HPLC-HR-MS/MS analysis of crude extract from *E. coli* carrying *kj12ABC* allowed determination of the molecular formulas of seven sulfur-containing derivatives, 1−7 (Table S1). Detailed analysis of their MS^2 fragmentation pattern (Figure S1) and comparison with corresponding valine-substituted compounds (Figure S2) allowed the identification of their amino acid composition and sequence together with the C-terminal amine. Based on these preliminary results, we suggested the structures of 1−7 as shown in Figure 2. In order to confirm the proposed structures, compounds 1−7 were synthesized by using a solid-phase peptide synthesis method developed for the synthesis of partly methylated rhabdopeptide/xenortide-like peptides (Scheme 1)[8,9] and compared to natural 1−7 regarding retention time and MS^2 fragmentation pattern (Figure S3).

Additionally, we synthesized two rhabdopeptide/xenortide-like peptides (10 and 12) which were produced as main products in both heterologous *E. coli* and *Xenorhabdus* KJ12.1 wild-type strains. They show *N*-methylated valine instead of the corresponding *N*-methylated methionine of 1, 2, and 5 (Figure 2), respectively. Compounds 1−7 are produced only in trace amounts in comparison with valine-incorporating rhabdopeptide/xenortide-like peptides; for example, compounds 1 and 2 are produced in a total of only 0.3 mg/L compared to their corresponding valine derivative 10 (22.4 mg/L), as deduced from quantitative analysis using the synthesized derivatives as standards.

Compounds 2−7, 10, and 12 were tested against various protozoal diseases, namely the causative agents of sleeping sickness (*Trypanosoma brucei rhodesiense*), Chagas disease (*Trypanosoma cruzi*), leishmaniasis (*Leishmania donovani*), and malaria (*Plasmodium falciparum*) (Table 1).[10] All compounds displayed

Table 1. Bioactivities of 2−7, 10, and 12 against Different Protozoa and Cytotoxicity (IC$_{50}$ in μM)[a]

compd	*T. brucei rhodesiense*	*T. cruzi*	*L. donovani*	*P. falciparum*	L6
2	12.3	75.0	>100	10.9	>100
3	8.5	62.9	>100	7.9	75.2
4	7.3	14.2	18.8	3.2	70.8
5	4.3	53.0	>100	6.6	71.2
6	5.7	25.4	>100	4.1	31.3
7	2.7	12.6	92.3	2.3	23.9
10	10.5	69.4	>100	5.3	89.1
12	1.4	26.2	>100	4.6	56.2
reference	0.02	2.8	0.3	0.006	0.01

[a]The positive reference is different for each target organism: melarsoprol for *T. brucei rhodesiense*, benznidazole for *T. cruzi*, miltefosin for *L. donovani*, chloroquine for *P. falciparum*, and podophyllotoxin for mammalian L6 cells.

good effects against *T. brucei rhodesiense* (IC$_{50}$ = 1.4−12.3 μM) and *P. falciparum* (IC$_{50}$ = 2.3−10.9 μM), with compounds 7 and 12 being the most potent against *P. falciparum* (IC$_{50}$ = 2.3 μM) and *T. brucei rhodesiense* (IC$_{50}$ = 1.4 μM), respectively, whereas all compounds showed weak or no activities against *T. cruzi* and *L. donovani*. Interestingly, valine rhabdopeptide/xenortide-like peptide 12 is more active than the corresponding methionine rhabdopeptide/xenortide-like peptide 5 against *T. brucei rhodesiense*, *P. falciparum*, and *T. cruzi*. All compounds exhibited no cytotoxicity (IC$_{50}$ > 10 μM) toward the mammalian L6 cells, indicating that these compounds have a high selectivity against parasitic protozoa.

It is worth noting that methionine-containing rhabdopeptide/xenortide-like peptides can only be detected in the heterologous *E. coli* DH10B MtaA strain, but not in *Xenorhabdus* KJ12.1 wild-type strain (Figure S4). Furthermore, the expression with *kj12B* and *kj12C* alone without *kj12A* did not produce

methionine-containing rhabdopeptide/xenortide-like peptides (Figure S4). These results suggested that *E. coli* DH10B MtaA background and Kj12A might contribute to the formation of methionine-containing rhabdopeptide/xenortide-like peptides by an unknown mechanism. One possibility might be the formation of protein complexes between the NRPS and unknown proteins that alter the specificity of the adenylation domains as described for MbtH proteins.[11,12] However, overexpression of MbtH from *E. coli* DH10B MtaA and *Xenorhabdus* KJ12.1 did not change the levels of methionine-containing rhabdopeptide/xenortide-like peptides in the *E. coli* DH10B MtaA system (Figure S5), and therefore another yet unknown mechanism must be responsible for their production.

EXPERIMENTAL SECTION

General Experimental Procedures. NMR spectra were recorded on a Bruker AV 500 spectrometer at 500 MHz (^{1}H) and 125 MHz (^{13}C). HPLC-ESI-MS analysis was performed on a Dionex UltiMate 3000 system coupled to a Bruker AmaZonX mass spectrometer. Crude extracts and standard compounds (5 μL injection volume) were eluted on an ACQUITY UPLC BEH C$_{18}$ column (130 Å, 2.1 mm × 100 mm, 1.7 μm) using a gradient from 5% to 95% aqueous acetonitrile containing 0.1% formic acid at a flow rate of 0.6 mL/min for 16 min. HPLC-HR-MS analysis (5 μL injection volume) was recorded on a Dionex UltiMate 3000 system coupled to a Bruker Impact II QTOF mass spectrometer. An ACQUITY UPLC BEH C$_{18}$ column (130 Å, 2.1 mm × 50 mm, 1.7 μm) and the same gradient were used to elute the crude extracts and compounds at a flow rate of 0.4 mL/min for 16 min. Positive mode with scan range from 100 to 1200 m/z was used to detect rhabdopeptide/xenortide-like peptides.

Strain Construction and Cultivation. Heterologous *E. coli* strains carrying RXP-BGC of *Xenorhabdus* KJ12.1 (*kj12ABC*) or carrying *kj12BC* without *kj12A* were constructed as described previously.[5] MbtH-overproducing strains were constructed by transferring MbtH-encoding plasmids into heterologous *E. coli* strains already carrying *kj12ABC*. For extract analysis, 10 mL of lysogeny broth (LB) medium containing appropriate antibiotics was inoculated with 0.1% of overnight culture, and 0.1% L-arabinose (from a 25% stock solution), 1 mM phenylethylamine (PEA), and 2% Amberlite XAD-16 resin were added. The cultures were grown at 30 °C and shaken at 200 rpm. After 24 h, the cultures were harvested. XAD-16 beads were separated and extracted with 10 mL of methanol for 1 h. The extracts were evaporated to dryness on a rotary evaporator. Before HPLC-MS analysis, the residue was re-dissolved in 1 mL of methanol and diluted 10 times with methanol.

Chemical Synthesis. The synthesis was performed manually using stepwise solid-phase peptide synthesis (SPPS) method.[8,9] For a schematic overview and detailed steps, see Scheme 1 and the Supporting Information.

Bioactivity Tests. Compounds **2**−**7**, **10**, and **12** were tested against the parasites *Trypanosoma brucei rhodesiense* STIB900, *Trypanosoma cruzi* Tulahuen C4, *Leishmania donovani* MHOM-ET-67/L82, and *Plasmodium falciparum* NF54 as described previously.[10] Cytotoxicity against rat skeletal myoblasts (L6 cells) was evaluated as described previously.[10] IC$_{50}$ (50% inhibitory concentration) values of compounds against these cells were calculated.

Quantitative Analysis. Quantitative analysis of **1**, **2**, and **10** produced in the heterologous *E. coli* DH10B MtaA strain was carried out as described previously.[13]

ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jnatprod.8b00425.

Chemical synthesis, HR-MS data, NMR data, HPLC-MS analysis, and NMR spectra of **1**−**7**, **10**, and **12**; bacterial strains, plasmids, and primers used, including Figures S1−S50 and Tables S1−S13 (PDF)

AUTHOR INFORMATION

Corresponding Author
*Tel.: +49 69 798 29557. Fax: +49 69 798 29527. E-mail: h.bode@bio.uni-frankfurt.de.

ORCID
Xiaofeng Cai: 0000-0001-7274-017X
Helge B. Bode: 0000-0001-6048-5909

Notes
The authors declare no competing financial interest.

REFERENCES

(1) Lang, G.; Kalvelage, T.; Peters, A.; Wiese, J.; Imhoff, J. F. *J. Nat. Prod.* **2008**, *71*, 1074−1077.
(2) Crawford, J. M.; Portmann, C.; Kontnik, R.; Walsh, C. T.; Clardy, J. *Org. Lett.* **2011**, *13*, 5144−5147.
(3) Reimer, D.; Cowles, K. N.; Proschak, A.; Nollmann, F. I.; Dowling, A. J.; Kaiser, M.; Constant, R. F.; Goodrich-Blair, H.; Bode, H. B. *ChemBioChem* **2013**, *14*, 1991−1997.
(4) Reimer, D.; Nollmann, F. I.; Schultz, K.; Kaiser, M.; Bode, H. B. *J. Nat. Prod.* **2014**, *77*, 1976−1980.
(5) Cai, X.; Nowak, S.; Wesche, F.; Bischoff, I.; Kaiser, M.; Fürst, R.; Bode, H. B. *Nat. Chem.* **2017**, *9*, 379−386.
(6) Thanwisai, A.; Tandhavanant, S.; Saiprom, N.; Waterfield, N. R.; Long, P. K.; Bode, H. B.; Peacock, S. J.; Chantratita, N. *PLoS One* **2012**, *7*, e43835.
(7) Bode, H. B. *Curr. Opin. Chem. Biol.* **2009**, *13*, 224−230.
(8) Wesche, F.; Adihou, H.; Kaiser, A.; Wurglics, M.; Schubert-Zsilavecz, M.; Kaiser, M.; Bode, H. B. *J. Med. Chem.* **2018**, *61*, 3930−3938.
(9) Sable, G. A.; Park, J.; Kim, H.; Lim, S.; Jang, S.; Lim, D. *Eur. J. Org. Chem.* **2015**, *2015*, 7043−7052.
(10) Orhan, I.; Şener, B.; Kaiser, M.; Brun, R.; Tasdemir, D. *Mar. Drugs* **2010**, *8*, 47−58.
(11) Süssmuth, R. D.; Mainz, A. *Angew. Chem., Int. Ed.* **2017**, *56*, 3770−3821.
(12) Crüsemann, M.; Kohlhaas, C.; Piel, J. *Chem. Sci.* **2013**, *4*, 1041−1045.
(13) Cai, X.; Challinor, V. L.; Zhao, L.; Reimer, D.; Adihou, H.; Grün, P.; Kaiser, M.; Bode, H. B. *Org. Lett.* **2017**, *19*, 806−809.

Supporting Information

Methionine-Containing Rhabdopeptide/Xenortide-like Peptides from Heterologous Expression of the Biosynthetic Gene Cluster *kj12ABC* in *Escherichia coli*

Lei Zhao,[1,2] Xiaofeng Cai,[1] Marcel Kaiser,[3,4] and Helge B. Bode*[1,5]

[1]Molecular Biotechnology, Department of Biosciences, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

[2]Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, 210014 Nanjing, China

[3]Parasite Chemotherapy, Swiss Tropical and Public Health Institute, 4051 Basel, Switzerland

[4]University of Basel, 4003 Basel, Switzerland

[5]Buchmann Institute for Molecular Life Sciences (BMLS), Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

*Corresponding author

Supplementary Methods

Chemical synthesis

The synthesis was performed manually using stepwise solid phase peptide synthesis (SPPS) method.[1,2] For a schematic overview see Scheme 1. Synthesis of **1** was shown as an example. At step **a**, the attachment of the C-terminal amine phenylethylamine (PEA) on the 2-(3,5-dimethoxy-4-formylphenoxy)ethyl (DFPE) resin was carried out. A mixture of PEA (126 μL, 1.0 mmol, 10 eq.) in 1.4 mL N,N-dimethylformamide (DMF)/MeOH/AcOH (80:19:1), NaBH$_3$CN (62.8 mg, 1.0 mmol, 10 eq.) and DFPE resin (95.2 mg, 0.1 mmol, 1 eq.) were placed in a 2 mL Eppendorf tube and incubated in a thermoshaker at 60 °C overnight. The resin was filtered, and subsequently washed with DMF (5 ×) and dichloromethane (DCM) (5 ×), and dried. At step **b**, acylation of N-Me-L-Val was conducted. A solution of Fmoc-N-Me-Val-OH (353.4 mg, 1.0 mmol, 10 eq.), N-[(dimethylamino)-1H-1,2,3-triazolo[4,5-b]-pyridino-1-ylmethylene]-N-methylmethanamini-um hexafluorophosphate (HATU) (381 mg, 1.0 mmol, 10 eq), 1-hydroxy-7-azabenzotriazole (HOAt) (136 mg, 1.0 mmol, 10 eq) and N,N-diisopropylethylamine (DIPEA) (0.34 mL, 2.0 mmol, 20 eq.) in 2.0 mL DMF was added to resin. The resulting mixture was incubated in a plastic reactor vessel equipped with a Teflon frit at room temperature overnight. The resin was washed with DMF (5 ×) and DCM (5 ×) and treated with 20% piperidine in DMF (3 × 10 min, 2 mL) to remove the Fmoc protecting group. Afterwards the resin was washed with DCM (5 ×) and dried. At step **c**, N-Me-L-Met was coupled to peptide sequence. The coupling of Fmoc-N-Me-Met-OH was mediated by using the efficient and rapid BTC coupling reagent. The dried peptidyl resin (25 μmol) was swollen with dry tetrahydrofuran (THF) (1 mL) for 15 min; meantime, in a separate flask, bis-(trichloromethyl)carbonate (BTC) (8.5 mg, 28.3 μmol, 1.15 eq.) was dissolved in dry THF (0.7 mL), and the Fmoc-N-Me-Met-OH (28.9 mg, 75 μmol, 3 eq.) was added to it, which resulted in a clear amino acid solution. Collidine (34 μL, 250 μmol, 10 eq.) was added to this clear solution, and a precipitate was immediately formed. This precipitate was added to the resin beads, which were pre-mixed with DIPEA (35 μL, 200 μmol, 8 eq.), and the whole reaction mixture was shaken at room temperature for 2 h. The resin was washed with DMF (5 ×) and DCM (5 ×) and treated with 20% piperidine in DMF (3 × 10 min, 2 mL) to remove the Fmoc protecting group. Afterwards the resin was washed with

S2

132

DCM (5 ×) and dried. At step **d**, The L-Val was coupled by using the HATU/HOAt coupling reagent. The dried peptidyl resin (25 μmol) was swollen in DMF. A solution of Fmoc-Val-OH (25.5 mg, 75 μmol, 3 eq.), HATU (28.6 mg, 75 μmol, 3 eq.), HOAt (10.2 mg, 75 μmol, 3 eq.) and DIPEA (25.5 μL, 150 μmol, 6 eq.) in 0.5 mL DMF was added to resin and shaken at room temperature overnight. The resin was washed with DMF (5 ×) and DCM (5 ×) and treated with 20% piperidine in DMF (3 × 10 min, 2 mL) to remove the Fmoc protecting group. Afterwards the resin was washed with DCM (5 ×) and dried. At step **e**, *N*-Me-L-Val was coupled to peptide sequence using the same method as step **c**. At the final step **f**, the peptide was cleaved from the resin. A total of 1 mL trifluoroacetic acid (TFA)/triisopropylsilane (TIS)/water (95:2.5:2.5) was added to the peptidyl resin (25 μmol) and the mixture was agitated for at least 2 h at room temperature. The resin was removed by filtration and washed twice with TFA. The solution was concentrated *in vacuo*. The residue was purified by Agilent HPLC system. The structures of pure compounds were confirmed by HR-MS, 1D and 2D NMR.

Supplementary Tables

Table S1. HR-MS data of natural and synthetic **1–7**, **10** and **12**

| compound	sum formula	m/z [M + H]$^+$ calcd.	natural		synthetic	
			m/z [M + H]$^+$ found	Δppm	m/z [M + H]$^+$ found	Δppm
1	$C_{31}H_{53}N_5O_4S$	592.3891	592.3887	0.7	592.3878	2.2
2	$C_{31}H_{53}N_5O_4S$	592.3891	592.3887	0.7	592.3886	0.8
3	$C_{31}H_{53}N_5O_4S$	592.3891	592.3882	1.5	592.3885	1.0
4	$C_{36}H_{62}N_6O_5S$	691.4575	691.4559	2.3	691.4564	1.6
5	$C_{37}H_{64}N_6O_5S$	705.4732	705.4717	2.1	705.4717	2.1
6	$C_{37}H_{64}N_6O_5S$	705.4732	705.4717	2.1	705.4723	1.2
7	$C_{37}H_{64}N_6O_5S$	705.4732	705.4727	0.7	705.4723	1.2
10	$C_{31}H_{53}N_5O_4$	560.4170	560.4164	1.2	560.4165	0.9
12	$C_{37}H_{64}N_6O_5$	673.5011	673.5000	1.7	673.5003	1.2

Table S2. ^{1}H (500 MHz) and ^{13}C (125 MHz) NMR Data of **1** in Methanol-d_4 (δ in ppm, J in Hz)
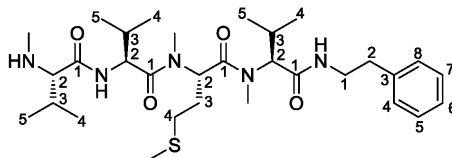
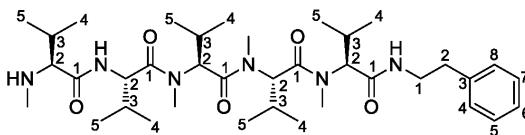| subunit	position	δ_C	δ_H
N-Me-L-Val	1	168.1 (C)	
	2	68.2 (CH)	3.68, d (5.5)
	3	31.7 (CH)	2.24−2.14, m
	4	18.8 (CH$_3$)	1.07, d (6.3)
	5	18.7 (CH$_3$)	1.03, d (6.9)
	N-CH$_3$	33.3	2.65, s
L-Val	1	173.8 (C)	
	2	56.9 (CH)	4.70, d (7.6)
	3	31.6 (CH)	2.13−2.04, m
	4	18.6 (CH$_3$)	1.01, d (6.8)
	5	19.9 (CH$_3$)	0.99, d (6.8)
N-Me-L-Met	1	172.8 (C)	
	2	53.8 (CH)	5.64, dd (8.4, 6.1)
	3	29.5 (CH$_2$)	2.04−1.96, 1.96−1.88, m
	4	31.4 (CH$_2$)	2.39, t (7.2)
	S-CH$_3$	15.4	2.07, s
	N-CH$_3$	31.7	3.11, s
N-Me-L-Val	1	171.7 (C)	
	2	64.0 (CH)	4.53, d (11.1)
	3	27.6 (CH)	2.24−2.14, m
	4	19.9 (CH$_3$)	0.87, d (6.4)
	5	19.2 (CH$_3$)	0.76, d (6.6)
	N-CH$_3$	31.5	2.95, s
PEA	1	41.6 (CH$_2$)	3.54−3.47, 3.41−3.33, m
	2	36.5 (CH$_2$)	2.84−2.72, m
	3	140.5 (C)	
	4	130.0 (CH)	7.19, overlap
	5	129.7 (CH)	7.27, overlap
	6	127.6 (CH)	7.19, overlap
	7	129.7 (CH)	7.27, overlap
	8	130.0 (CH)	7.19, overlap

Table S3. ^{1}H (500 MHz) and ^{13}C (125 MHz) NMR Data of **2** in Methanol-d_4 (δ in ppm, J in Hz)
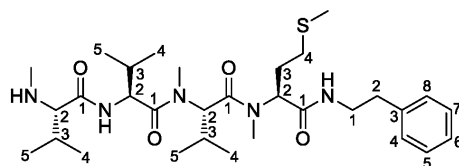
subunit	position	δ_C	δ_H
N-Me-L-Val	1	168.0 (C)	
	2	68.2 (CH)	3.67, d (5.4)
	3	31.7 (CH)	2.20−2.12, m
	4	18.7 (CH$_3$)	1.03, d (7.0)
	5	18.5 (CH$_3$)	1.00, overlap
	N-CH$_3$	33.3 (CH$_3$)	2.63, s
L-Val	1	174.1 (C)	
	2	56.7 (CH)	4.72, d (7.9)
	3	31.8 (CH)	2.11−2.01, m
	4	18.8 (CH$_3$)	1.00, overlap
	5	19.8 (CH$_3$)	0.95, d (6.8)
N-Me-L-Val	1	173.0 (C)	
	2	59.7 (CH)	5.18, d (10.8)
	3	28.7 (CH)	2.45−2.25, m
	4	19.9 (CH$_3$)	0.88, d (6.4)
	5	19.2 (CH$_3$)	0.79, d (6.6)
	N-CH$_3$	31.4 (CH$_3$)	3.12, s
N-Me-L-Met	1	172.1 (C)	
	2	57.3 (CH)	5.09, dd (9.7, 5.8)
	3	29.3 (CH$_2$)	2.11−2.01, 1.92−1.82, m
	4	31.5 (CH$_2$)	2.29, t (7.4)
	S-CH$_3$	15.4 (CH$_3$)	2.04, s
	N-CH$_3$	32.2 (CH$_3$)	2.97, s
PEA	1	41.9 (CH$_2$)	3.42, t (7.2)
	2	36.6 (CH$_2$)	2.77, t (7.2)
	3	140.5 (C)	
	4	130.0 (CH)	7.19, overlap
	5	129.7 (CH)	7.26, overlap
	6	127.6 (CH)	7.19, overlap
	7	129.7 (CH)	7.26, overlap
	8	130.0 (CH)	7.19, overlap

Table S4. ^{1}H (500 MHz) and ^{13}C (125 MHz) NMR Data of **3** in Methanol-d_4 (δ in ppm, J in Hz)
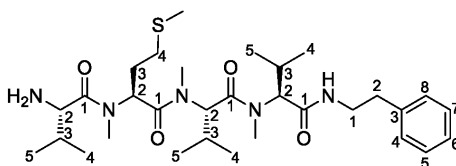
subunit	position	δ_C	δ_H
L-Val	1	171.0 (C)	
	2	57.1 (CH)	4.27, d (4.5)
	3	30.9 (CH)	2.24–2.16, m
	4	19.6 (CH$_3$)	1.11, d (7.0)
	5	16.9 (CH$_3$)	0.99, d (6.9)
N-Me-L-Met	1	172.6 (C)	
	2	53.7 (CH)	5.73, t (7.3)
	3	29.0 (CH$_2$)	2.16–2.09, 2.00–1.91, m
	4	31.4 (CH$_2$)	2.54–2.41, m
	S-CH$_3$	15.2 (CH$_3$)	2.07, s
	N-CH$_3$	31.7 (CH$_3$)	3.08, s
N-Me-L-Val	1	172.5 (C)	
	2	60.1 (CH)	5.09, d (10.8)
	3	28.7 (CH)	2.35–2.27, m
	4	20.0 (CH$_3$)	0.88, d (6.5)
	5	19.0 (CH$_3$)	0.79, d (6.7)
	N-CH$_3$	31.1 (CH$_3$)	3.03, s
N-Me-L-Val	1	171.8 (C)	
	2	63.9 (CH)	4.57, d (11.1)
	3	27.8 (CH)	2.24–2.16, m
	4	19.9 (CH$_3$)	0.86, d (6.6)
	5	19.2 (CH$_3$)	0.76, d (6.7)
	N-CH$_3$	31.5 (CH$_3$)	2.99, s
PEA	1	41.7 (CH$_2$)	3.50–3.35, m
	2	36.6 (CH$_2$)	2.79–2.74, m
	3	140.4 (C)	
	4	129.9 (CH)	7.19, overlap
	5	129.7 (CH)	7.26, overlap
	6	127.6 (CH)	7.19, overlap
	7	129.7 (CH)	7.26, overlap
	8	129.9 (CH)	7.19, overlap

S7

Table S5. ^{1}H (500 MHz) and ^{13}C (125 MHz) NMR Data of **4** in Methanol-d_4 (δ in ppm, J in Hz)
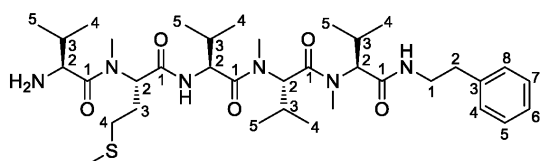
subunit	position	δ_C	δ_H
L-Val	1	171.5 (C)	
	2	57.1 (CH)	4.22, d (5.0)
	3	31.1 (CH)	2.25–2.13, m
	4	19.5 (CH$_3$)	1.11, d (7.0)
	5	17.1 (CH$_3$)	1.01, d (6.9)
N-Me-L-Met	1	172.0 (C)	
	2	57.5 (CH)	5.20, d (7.1)
	3	29.7 (CH$_2$)	2.25–2.13, 2.00–1.91, m
	4	31.4 (CH$_2$)	2.43, t (7.4)
	S-CH$_3$	15.4 (CH$_3$)	2.09, s
	N-CH$_3$	32.0 (CH$_3$)	3.09, s
L-Val	1	174.6 (C)	
	2	56.5 (CH)	4.60, d (8.4)
	3	31.7 (CH)	2.08–2.00, m
	4	18.8 (CH$_3$)	0.93, overlap
	5	19.9 (CH$_3$)	0.91, overlap
N-Me-L-Val	1	172.8 (C)	
	2	59.8 (CH)	5.18, d (10.7)
	3	29.0 (CH)	2.36–2.27, m
	4	19.9 (CH$_3$)	0.87, overlap
	5	18.9 (CH$_3$)	0.81, d (6.7)
	N-CH$_3$	31.4 (CH$_3$)	3.11, s
N-Me-L-Val	1	171.8 (C)	
	2	63.8 (CH)	4.58, d (11.1)
	3	27.7 (CH)	2.25–2.13, m
	4	19.8 (CH$_3$)	0.87, overlap
	5	19.0 (CH$_3$)	0.74, d (6.7)
	N-CH$_3$	31.6	3.03, s
PEA	1	41.7 (CH$_2$)	3.49–3.36, m
	2	36.7 (CH$_2$)	2.77, t (7.2)
	3	140.4 (C)	
	4	129.9 (CH)	7.19, overlap
	5	129.7 (CH)	7.26, overlap
	6	127.6 (CH)	7.19, overlap
	7	129.7 (CH)	7.26, overlap
	8	129.9 (CH)	7.19, overlap

S8

Table S6. ^{1}H (500 MHz) and ^{13}C (125 MHz) NMR Data of **5** in Methanol-d_4 (δ in ppm, J in Hz)
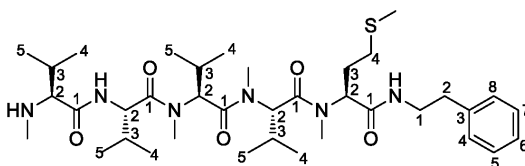
subunit	position	δ_C	δ_H
N-Me-L-Val	1	168.0 (C)	
	2	68.3 (CH)	3.68, d (5.3)
	3	31.7 (CH)	2.21–2.12, m
	4	18.7 (CH$_3$)	1.05, d (7.0)
	5	18.5 (CH$_3$)	1.02, d (6.9)
	N-CH$_3$	33.3 (CH$_3$)	2.65, s
L-Val	1	174.1 (C)	
	2	56.7 (CH)	4.76, d (7.6)
	3	31.7 (CH)	2.11–2.03, m
	4	18.7 (CH$_3$)	1.02, d (6.9)
	5	19.9 (CH$_3$)	0.99, d (6.7)
N-Me-L-Val	1	172.3 (C)	
	2	59.8 (CH)	5.20, d (10.8)
	3	28.7 (CH)	2.37–2.25, m
	4	20.0 (CH$_3$)	0.88, d (6.4)
	5	19.1 (CH$_3$)	0.81, overlap
	N-CH$_3$	31.3 (CH$_3$)	3.16, s
N-Me-L-Val	1	172.8 (C)	
	2	59.9 (CH)	5.17, d (10.8)
	3	28.7 (CH)	2.37–2.25, m
	4	20.1 (CH$_3$)	0.92, d (6.4)
	5	18.8 (CH$_3$)	0.79, overlap
	N-CH$_3$	31.5 (CH$_3$)	3.02, s
N-Me-L-Met	1	172.2 (C)	
	2	57.5 (CH)	5.10, dd (9.8, 5.6)
	3	29.2 (CH$_2$)	2.11–2.03, 1.96–1.80, m
	4	31.5 (CH$_2$)	2.37–2.25, m
	S-CH$_3$	15.4 (CH$_3$)	2.06, s
	N-CH$_3$	32.2 (CH$_3$)	2.96, s
PEA	1	42.0 (CH$_2$)	3.46–3.41, m
	2	36.6 (CH$_2$)	2.79, dd (8.8, 5.5)
	3	140.5 (C)	
	4	130.0 (CH)	7.20, overlap
	5	129.7 (CH)	7.28, overlap
	6	127.6 (CH)	7.20, overlap
	7	129.7 (CH)	7.28, overlap
	8	130.0 (CH)	7.20, overlap

Table S7. ^{1}H (500 MHz) and ^{13}C (125 MHz) NMR Data of **6** in Methanol-d_4 (δ in ppm, J in Hz)
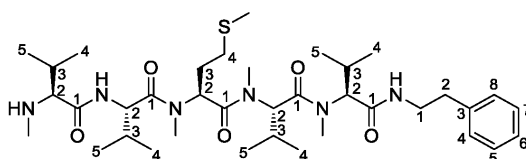
subunit	position	δ_C	δ_H
N-Me-L-Val	1	169.1 (C)	
	2	68.6 (CH)	3.55, d (5.6)
	3	31.9 (CH)	2.24–2.01, m
	4	19.0 (CH$_3$)	1.04, overlap
	5	18.8 (CH$_3$)	1.02, overlap
	N-CH$_3$	33.5 (CH$_3$)	2.60, s
L-Val	1	174.0 (C)	
	2	56.8 (CH)	4.70, d (8.0)
	3	31.6 (CH)	2.24–2.01, m
	4	18.8 (CH$_3$)	1.02, overlap
	5	19.8 (CH$_3$)	0.99, overlap
N-Me-L-Met	1	172.6 (C)	
	2	53.3 (CH)	5.76, t (7.3)
	3	29.2 (CH$_2$)	2.24–2.01, 1.93–1.84, m
	4	31.3 (CH$_2$)	2.47–2.36, m
	S-CH$_3$	15.2 (CH$_3$)	2.03, s
	N-CH$_3$	31.8 (CH$_3$)	3.17, s
N-Me-L-Val	1	172.5 (C)	
	2	60.1 (CH)	5.10, d (10.7)
	3	28.7 (CH)	2.36–2.27, m
	4	20.1 (CH$_3$)	0.89, d (6.4)
	5	18.9 (CH$_3$)	0.80, d (6.8)
	N-CH$_3$	31.1 (CH$_3$)	3.01, s
N-Me-L-Val	1	171.8 (C)	
	2	63.9 (CH)	4.57, d (11.1)
	3	27.8 (CH)	2.24–2.01, m
	4	19.9 (CH$_3$)	0.86, d (6.5)
	5	19.2 (CH$_3$)	0.76, d (6.6)
	N-CH$_3$	31.5 (CH$_3$)	2.99, s
PEA	1	41.7 (CH$_2$)	3.51–3.35, m
	2	36.7 (CH$_2$)	2.80–2.72, m
	3	140.4 (C)	
	4	129.9 (CH)	7.19, overlap
	5	129.7 (CH)	7.26, overlap
	6	127.6 (CH)	7.19, overlap
	7	129.7 (CH)	7.26, overlap
	8	129.9 (CH)	7.19, overlap

Table S8. [1]H (500 MHz) and ^{13}C (125 MHz) NMR Data of **7** in Methanol-d_4 (δ in ppm, J in Hz)
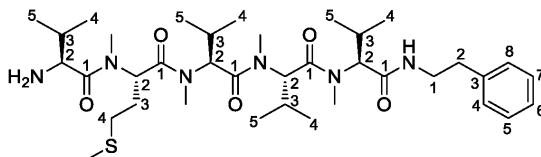
subunit	position	δ_C	δ_H
L-Val	1	171.1 (C)	
	2	57.1 (CH)	4.26, d (4.5)
	3	30.9 (CH)	2.24–2.04, m
	4	19.6 (CH$_3$)	1.11, d (7.0)
	5	16.9 (CH$_3$)	0.99, d (6.9)
N-Me-L-Met	1	172.7 (C)	
	2	53.7 (CH)	5.74, t (7.3)
	3	28.9 (CH$_2$)	2.24–2.04, 2.02–1.93, m
	4	31.4 (CH$_2$)	2.55–2.42, m
	S-CH$_3$	15.2 (CH$_3$)	2.08, s
	N-CH$_3$	31.7 (CH$_3$)	3.09, s
N-Me-L-Val	1	172.0 (C)	
	2	60.1 (CH)	5.12, d (10.8)
	3	28.5 (CH)	2.37–2.27, m
	4	20.0 (CH$_3$)	0.88, overlap
	5	18.9 (CH$_3$)	0.79, overlap
	N-CH$_3$	31.0 (CH$_3$)	3.05, s
N-Me-L-Val	1	172.6 (C)	
	2	59.9 (CH)	5.15, d (10.8)
	3	28.9 (CH)	2.37–2.27, m
	4	20.0 (CH$_3$)	0.88, overlap
	5	18.9 (CH$_3$)	0.79, overlap
	N-CH$_3$	31.3 (CH$_3$)	2.99, s
N-Me-L-Val	1	171.8 (C)	
	2	63.8 (CH)	4.57, d (11.1)
	3	27.7 (CH)	2.24–2.04, m
	4	19.9 (CH$_3$)	0.86, overlap
	5	18.9 (CH$_3$)	0.73, d (6.7)
	N-CH$_3$	31.6 (CH$_3$)	3.02, s
PEA	1	41.7 (CH$_2$)	3.50–3.35, m
	2	36.7 (CH$_2$)	2.80–2.72, m
	3	140.4 (C)	
	4	129.9 (CH)	7.19, overlap
	5	129.7 (CH)	7.26, overlap
	6	127.6 (CH)	7.19, overlap
	7	129.7 (CH)	7.26, overlap
	8	129.9 (CH)	7.19, overlap

Table S9. ^{1}H (500 MHz) and ^{13}C (125 MHz) NMR Data of **10** in Methanol-d_4 (δ in ppm, J in Hz)
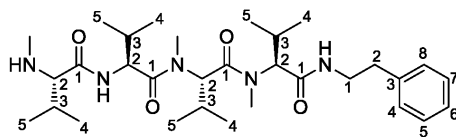
subunit	position	δ_C	δ_H
N-Me-L-Val	1	168.8 (C)	
	2	68.6 (CH)	3.58, d (5.4)
	3	31.8 (CH)	2.16−2.09, m
	4	18.8 (CH$_3$)	1.03, d (6.5)
	5	18.7 (CH$_3$)	1.00, overlap
	N-CH$_3$	33.5 (CH$_3$)	2.60, s
L-Val	1	174.2 (C)	
	2	56.6 (CH)	4.73, d (7.9)
	3	31.8 (CH)	2.09−2.02, m
	4	18.8 (CH$_3$)	1.00, overlap
	5	19.8 (CH$_3$)	0.95, d (6.8)
N-Me-L-Val	1	172.9 (C)	
	2	59.6 (CH)	5.19, d (10.8)
	3	28.9 (CH)	2.35−2.27, m
	4	19.8 (CH$_3$)	0.86, d (5.9)
	5	19.2 (CH$_3$)	0.80, d (6.7)
	N-CH$_3$	31.4 (CH$_3$)	3.15, s
N-Me-L-Val	1	171.8 (C)	
	2	63.8 (CH)	4.58, d (11.1)
	3	27.7 (CH)	2.24−2.16, m
	4	19.8 (CH$_3$)	0.86, d (5.9)
	5	19.0 (CH$_3$)	0.74, d (6.7)
	N-CH$_3$	31.7 (CH$_3$)	3.05, s
PEA	1	41.7 (CH$_2$)	3.48−3.35, m
	2	36.7 (CH$_2$)	2.76, t (7.3)
	3	140.4 (C)	
	4	129.9 (CH)	7.19, overlap
	5	129.7 (CH)	7.26, overlap
	6	127.6 (CH)	7.19, overlap
	7	129.7 (CH)	7.26, overlap
	8	129.9 (CH)	7.19, overlap

Table S10. [1]H (500 MHz) and ^{13}C (125 MHz) NMR Data of **12** in Methanol-d_4 (δ in ppm, J in Hz)

subunit	position	δ_C	δ_H
N-Me-L-Val	1	168.6 (C)	
	2	68.5 (CH)	3.61, d (5.3)
	3	31.8 (CH)	2.24–2.12, m
	4	18.7 (CH$_3$)	1.01, overlap
	5	18.6 (CH$_3$)	1.01, overlap
	N-CH$_3$	33.5 (CH$_3$)	2.61, s
L-Val	1	174.2 (C)	
	2	56.6 (CH)	4.76, d (7.7)
	3	31.7 (CH)	2.11–2.03, m
	4	18.8 (CH$_3$)	1.01, overlap
	5	20.0 (CH$_3$)	0.98, overlap
N-Me-L-Val	1	172.4 (C)	
	2	59.8 (CH)	5.20, d (10.8)
	3	28.7 (CH)	2.37–2.26, m
	4	19.9 (CH$_3$)	0.86, overlap
	5	19.0 (CH$_3$)	0.78, overlap
	N-CH$_3$	31.3 (CH$_3$)	3.15, s
N-Me-L-Val	1	172.7 (C)	
	2	59.7 (CH)	5.18, d (10.8)
	3	28.9 (CH)	2.37–2.26, m
	4	19.9 (CH$_3$)	0.86, overlap
	5	19.0 (CH$_3$)	0.78, overlap
	N-CH$_3$	31.5 (CH$_3$)	3.04, s
N-Me-L-Val	1	171.8 (C)	
	2	63.8 (CH)	4.57, d (11.1)
	3	27.7 (CH)	2.24–2.12, m
	4	19.9 (CH$_3$)	0.86, overlap
	5	19.1 (CH$_3$)	0.74, d (6.66)
	N-CH$_3$	31.6 (CH$_3$)	3.04, s
PEA	1	41.7 (CH$_2$)	3.49–3.35, m
	2	36.7 (CH$_2$)	2.76, m
	3	140.4 (C)	
	4	129.9 (CH)	7.19, overlap
	5	129.7 (CH)	7.26, overlap
	6	127.6 (CH)	7.19, overlap
	7	129.7 (CH)	7.26, overlap
	8	129.9 (CH)	7.19, overlap

S13

Table S11. Bacterial strains used in this study

strain	relevant genotype	reference
E. coli DH10B MtaA	F– mcrA, Δ(*mrr-hsd*RMS-*mcr*BC), Φ80*lacZ*ΔM15, ΔlacX74, *rec*A1, *end*A1, *ara*D139, Δ(*ara leu*)7697, *gal*U, *gal*K, *rps*L, *nup*G, λ–, *ent*D::*mtaA*	3,4
Xenorhabdus KJ12.1	wild type	3

Table S12. Plasmids used in this study

plasmid	description	reference
pCOLA-ara-tacI	3,345 bp, modified from pCOLA_tacI/I that contains arabinose-inducible promotor and kanamycin resistance gene (Km^R)	3
pCDF-ara-tacI	3,404 bp, modified from pCDF_tacI/I that contains arabinose-inducible promoter and spectinomycin resistance gene (Sm^R)	3
pCX3	16,107 bp, *kj12ABC* gene cluster from *Xenorhabdus* KJ12.1 genomic DNA assembled into pCOLA-ara-tacI, Km^R	3
pLZ59	3566 bp, *mbtH* gene from *Xenorhabdus* KJ12.1 genomic DNA assembled into pCDF-ara-tacI, Sm^R	this work
pLZ60	3566 bp, *mbtH* gene from *E. coli* DH10B MtaA genomic DNA assembled into pCDF-ara-tacI, Sm^R	this work

Table S13. Primers used in this study

primer	sequence (5'-3')	targeting DNA fragment	plasmid
XC252-Fw	AATTCCATGGAACAATTAACCGGAAATG	*mbtH* from *Xenorhabdus* KJ12.1 (218 bp)	
XC252-Rv	ATGATTAATTGTTAGTGCATATCAGTCTG CTTTTTAG		
XC253-Fw	AGACTGATATGCACTAACAATTAATCATC GGCTCGTATAATG	pCDF-ara-tacI vector backbone (3403 bp)	pLZ59
XC253-Fw	ATTTCCGGTTAATTGTTCCATGGAATTCC TCCTGTTAGC		
LZ_160	ATGGCATTCAGTAATCCCTTCGATG	*mbtH* from *E. coli* DH10B MtaA (219 bp)	
LZ_161	TCATTGTGCCTCCTGCAACTG		
LZ_162	AATTTTACCCAGTTGCAGGAGGCACAAT GACAATTAATCATCGGCTCGTATAATG	pCDF-ara-tacI vector backbone (3427 bp)	pLZ60
LZ_163	TGCGGATCATCGAAGGGATTACTGAATG CCATGGAATTCCTCCTGTTAGCCC		

Supplementary Figures

1 mV-V-mM-mV-PEA (---)
2 mV-V-mV-mM-PEA (---)
m/z 592.3887 [M + H]⁺

3 V-mM-mV-mV-PEA
m/z 592.3882 [M + H]⁺

4 V-mM-V-mV-mV-PEA
m/z 691.4559 [M + H]⁺

5 mV-V-mV-mV-mM-PEA
m/z 705.4717 [M + H]⁺

S15
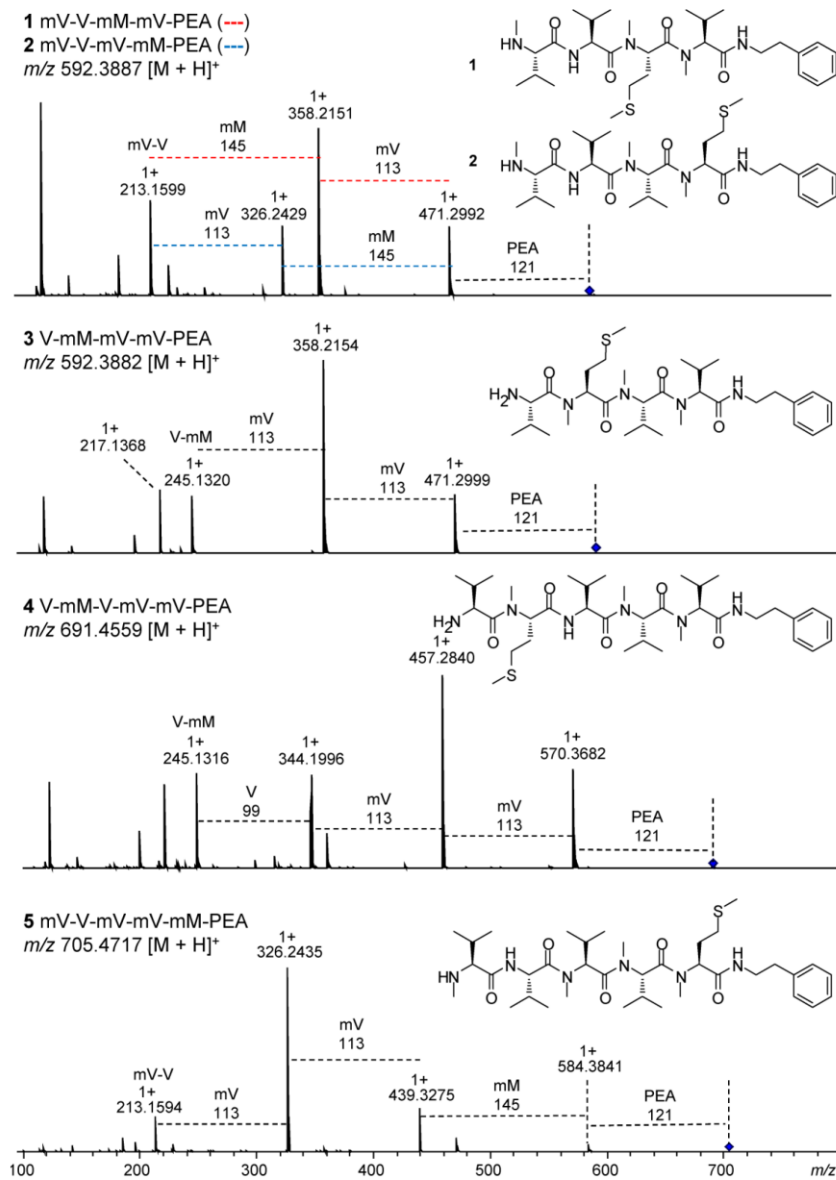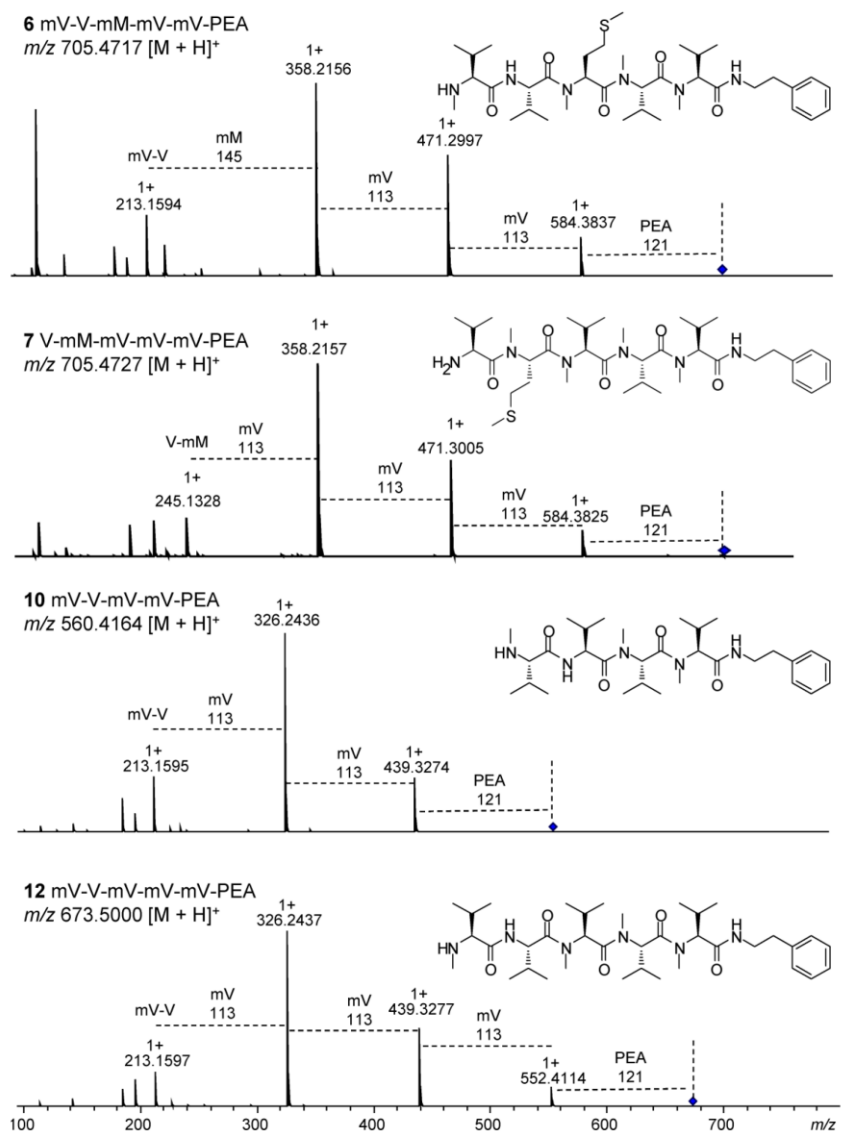
Figure S1. MS^2 fragmentation patterns of natural **1−7**, **10** and **12**.

S16

146
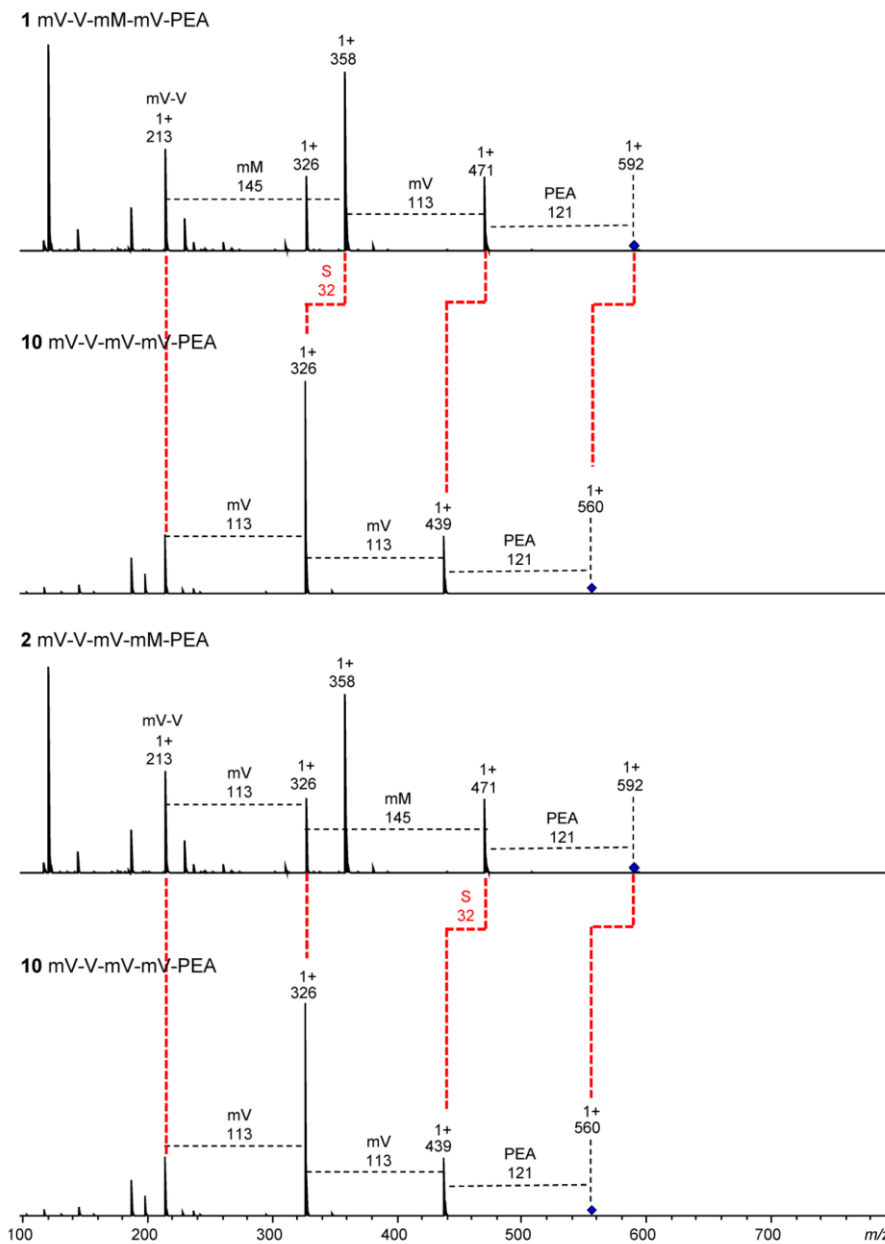
S17
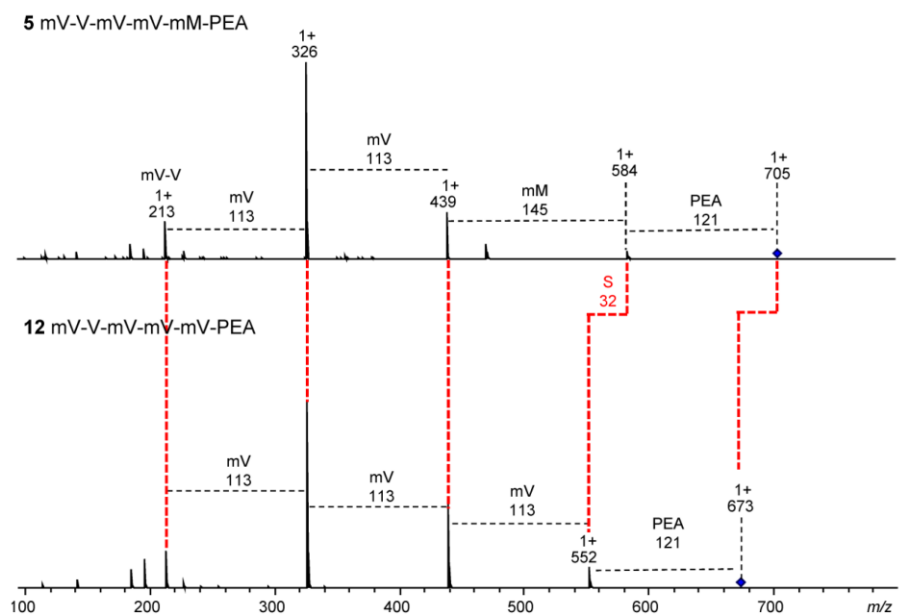
Figure S2. Comparative MS^2 fragmentation patterns of methionine-containing compounds **1**, **2** and **5** with corresponding valine-substituted compounds **10** and **12**.

S18

148

a *m/z* 592 [M + H]⁺

150

b *m/z* 592 [M + H]⁺

c *m/z* 691 [M + H]+

natural **4**

synthetic **4**

natural **4**

synthetic **4**

S21

S22
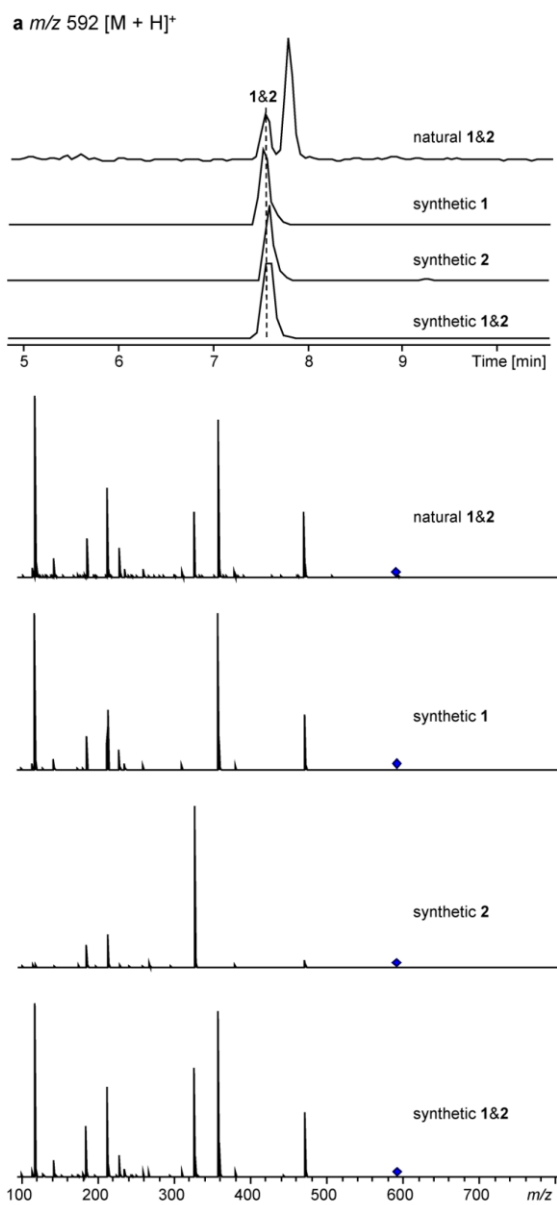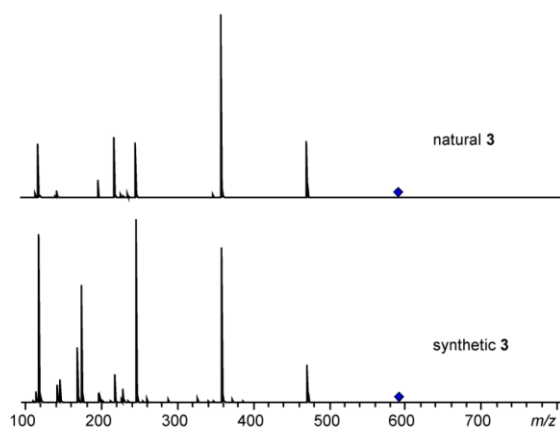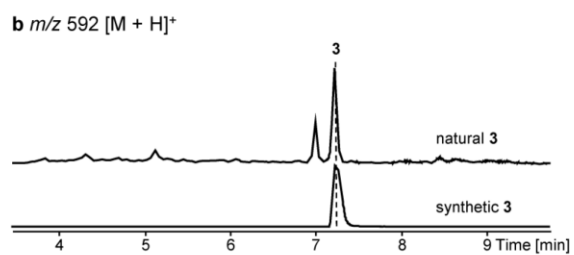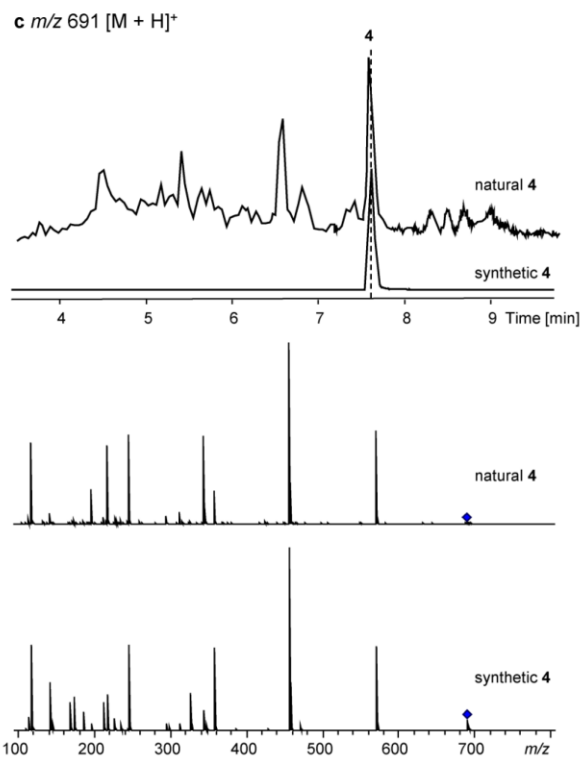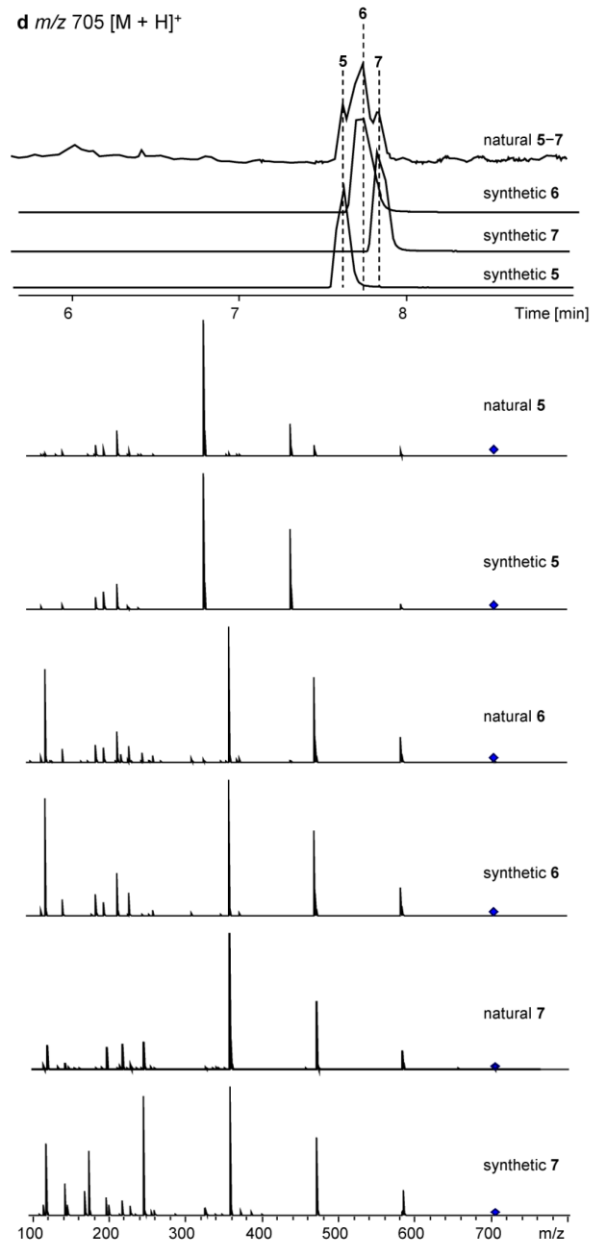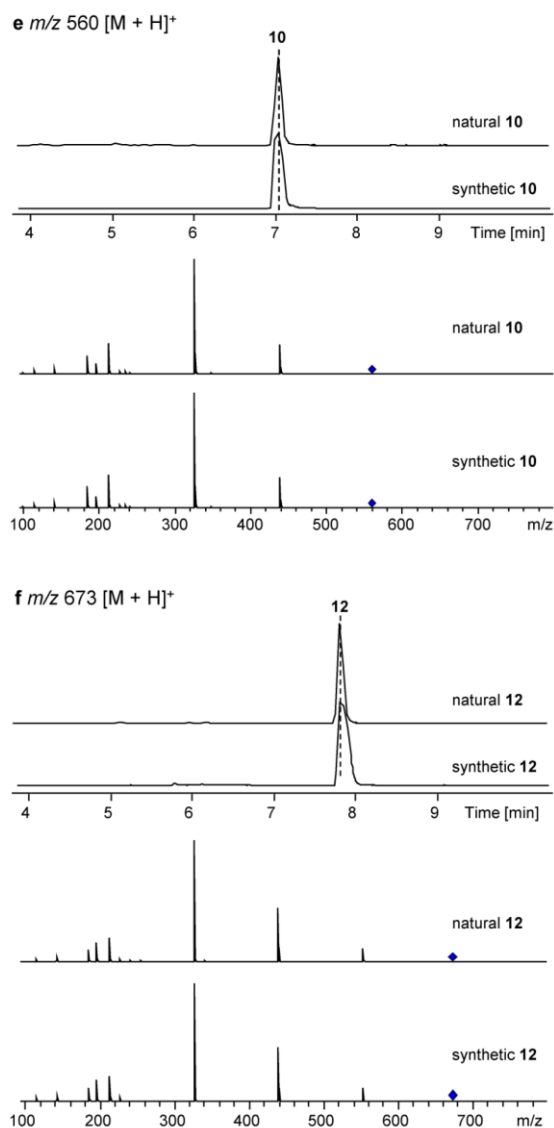
Figure S3. Extracted ion chromatograms (EICs) and the corresponding MS^2 fragmentation patterns of natural and synthetic **1−7**, **10** and **12** by HPLC-HR-ESI-MS/MS analysis. (a) *m/z* 592 [M + H]$^+$ for compounds **1** and **2** (EIC from HPLC-ESI-MS analysis). (b) *m/z* 592 [M + H]$^+$ for compound **3**; (c) *m/z* 691 [M + H]$^+$ for compound **4**; (d) *m/z* 705 [M + H]$^+$ for compounds **5−7**; (e) *m/z* 560 [M + H]$^+$ for compound **10**; (f) *m/z* 673 [M + H]$^+$ for compound **12**.
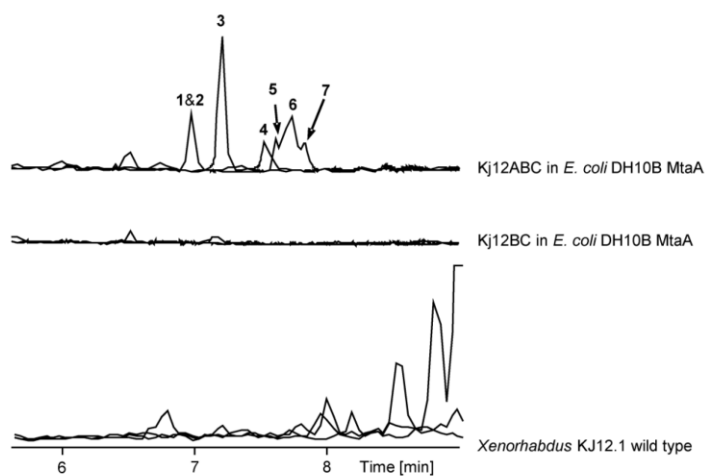
Figure S4. Extracted ion chromatograms (EICs) for methionine-containing rhabdopeptide/xenortide-like peptides (**1**−**7**) produced from the expression of Kj12ABC and Kj12BC in *E. coli* DH10B MtaA and wild type strains by HPLC-HR-ESI-MS analysis.
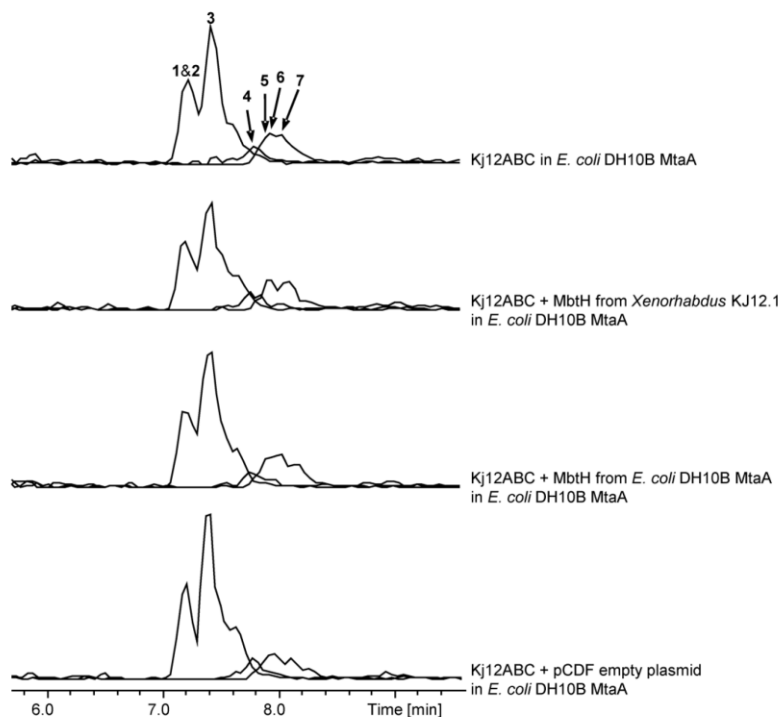
Figure S5. Extracted ion chromatograms (EICs) for methionine-containing rhabdopeptide/xenortide-like peptides (**1−7**) produced from the overexpression of *mbtH* of *Xenorhabdus* KJ12.1 and *E. coli* DH10B MtaA in *E. coli* DH10B MtaA strain by HPLC-ESI-MS analysis. All chromatograms were drawn to the same scale.
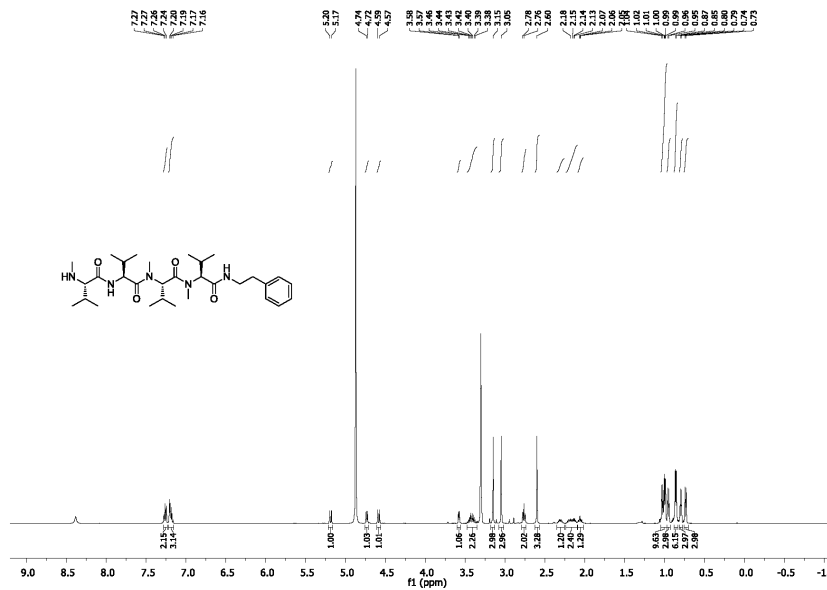
S25

155
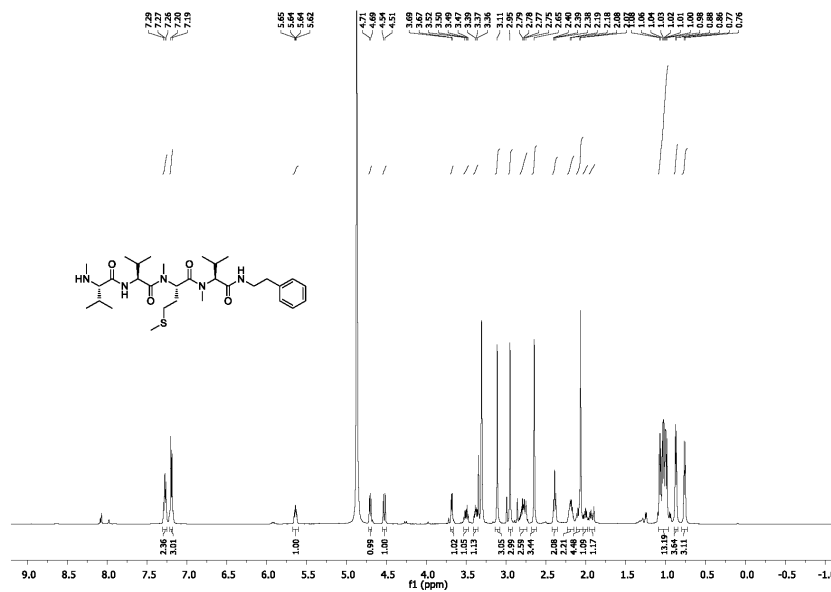
Figure S6. ¹H NMR (500 MHz, methanol-*d*₄) spectrum of **1**.
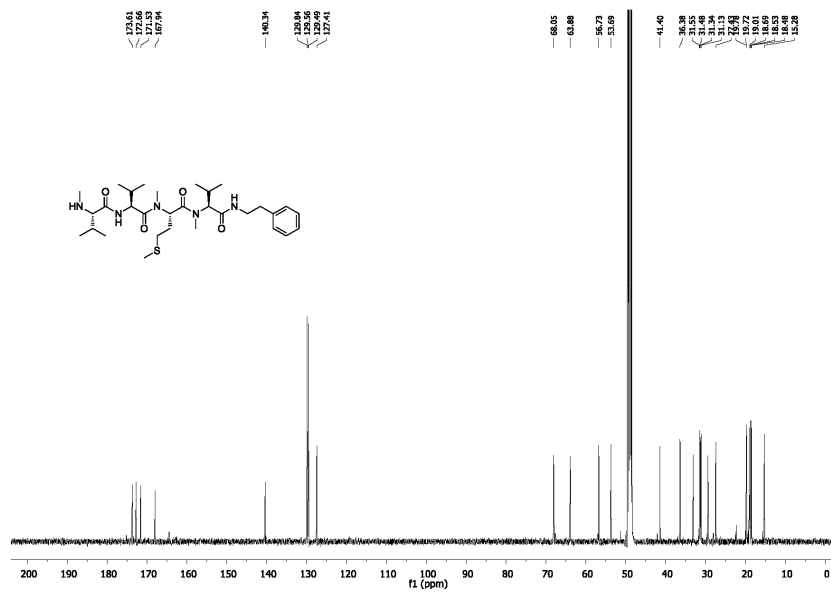
Figure S7. ¹³C NMR (125 MHz, methanol-*d*₄) spectrum of **1**.
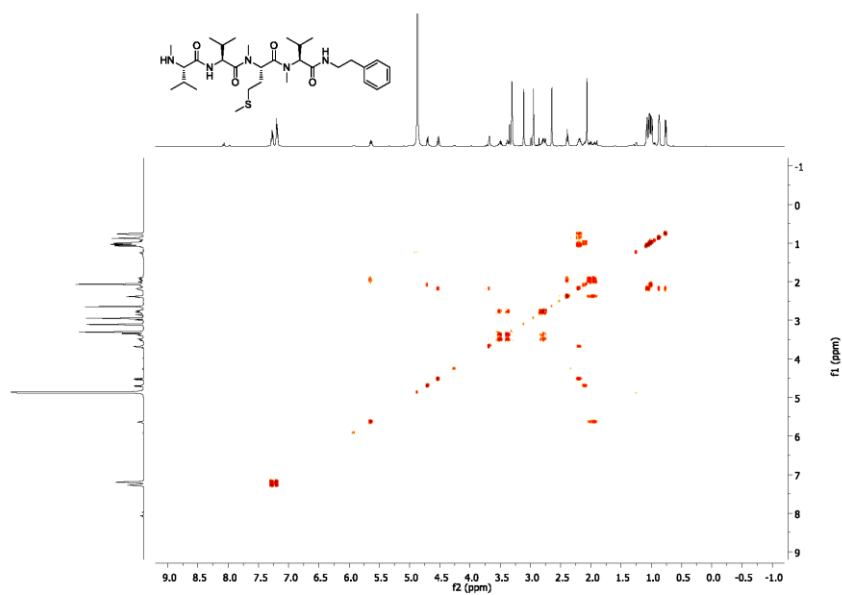
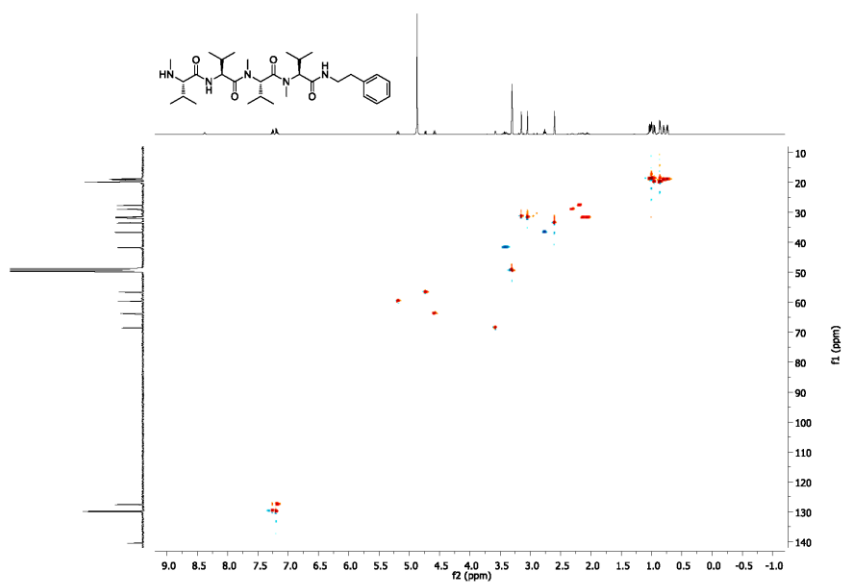**Figure S8.** COSY (methanol-d_4) spectrum of **1**.
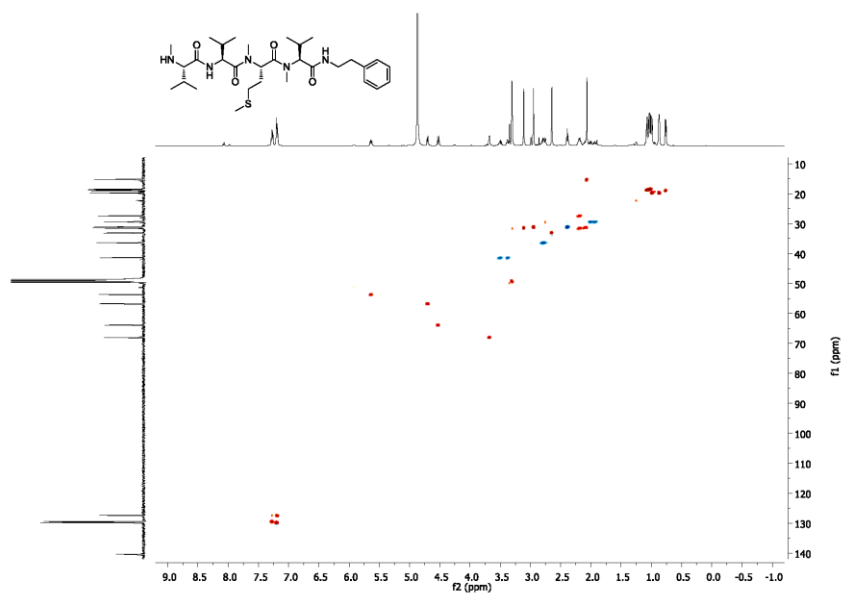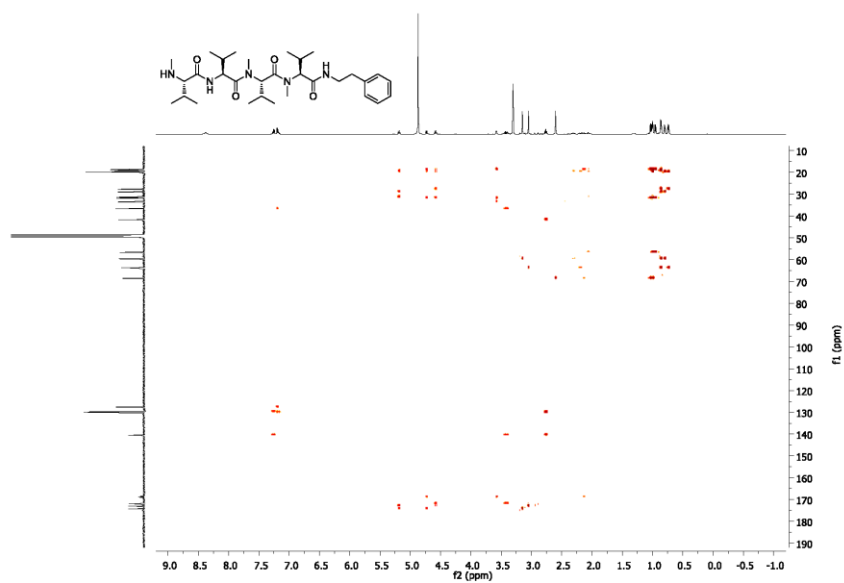

Figure S9. HSQC (methanol-d_4) spectrum of **1**.

S27

157
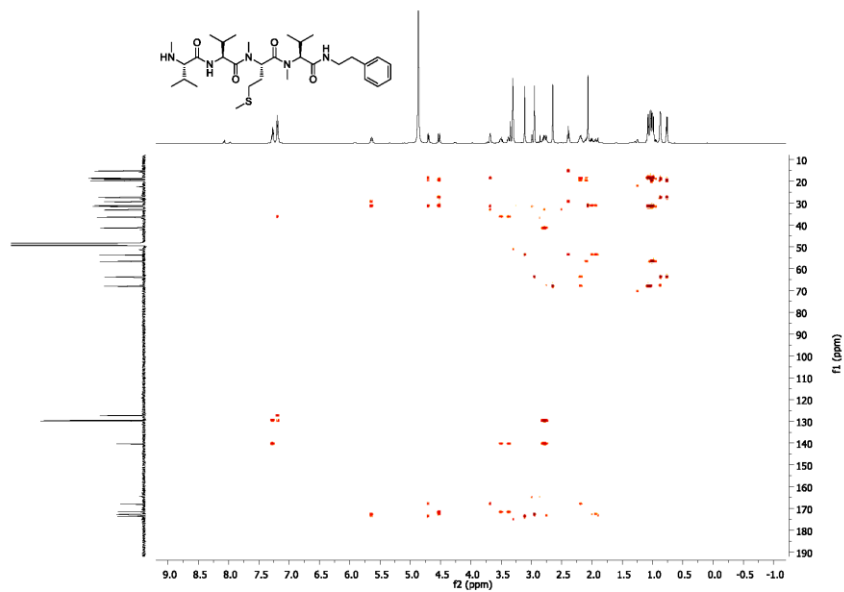
Figure S10. HMBC (methanol-d_4) spectrum of **1**.
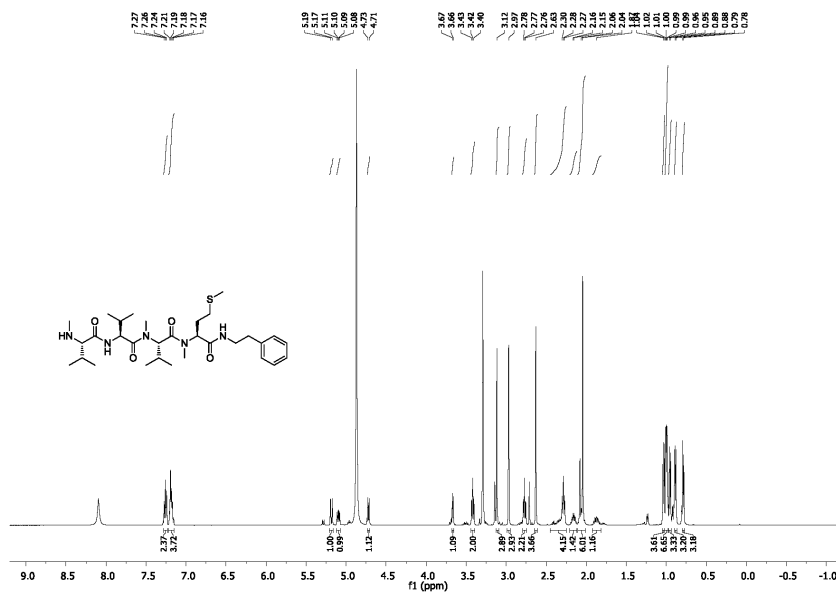
Figure S11. ^{1}H NMR (500 MHz, methanol-d_4) spectrum of **2**.
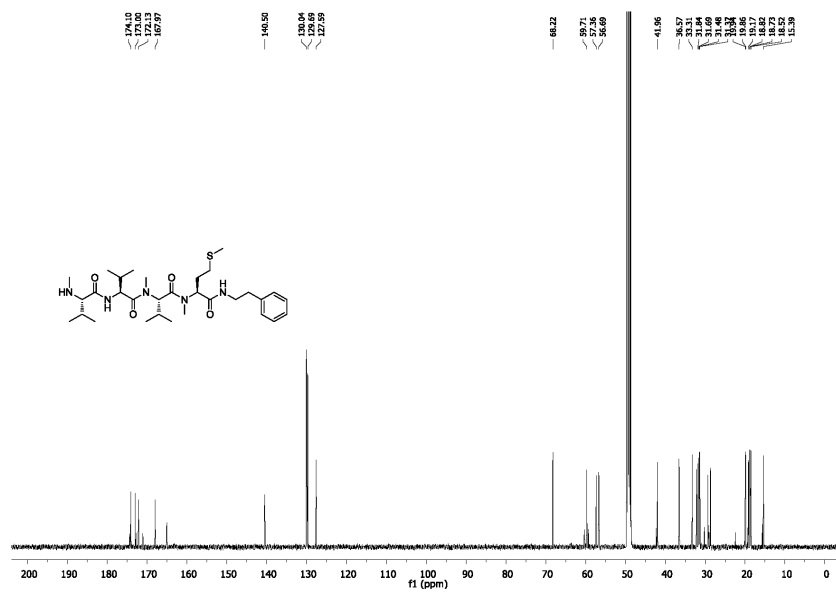
Figure S12. ^{13}C NMR (500 MHz, methanol-d_4) spectrum of **2**.

S29

159
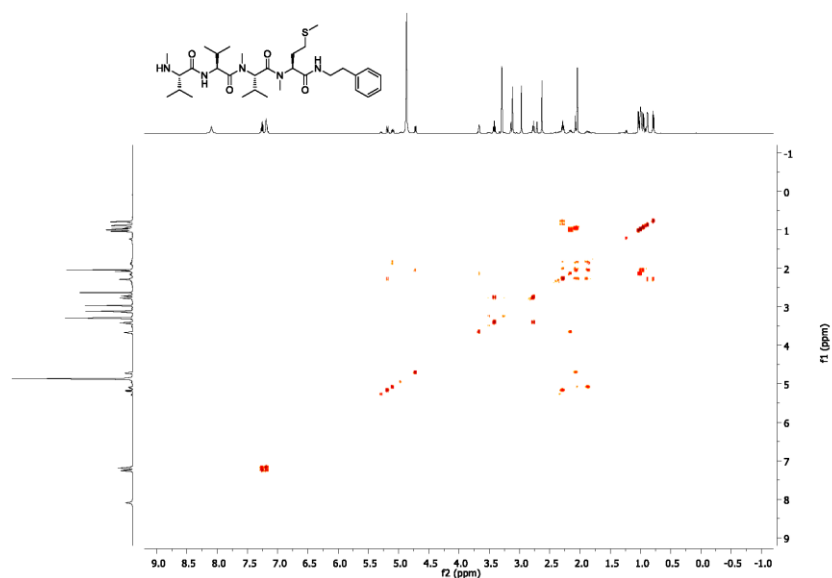
Figure S13. COSY (methanol-*d₄*) spectrum of **2**.
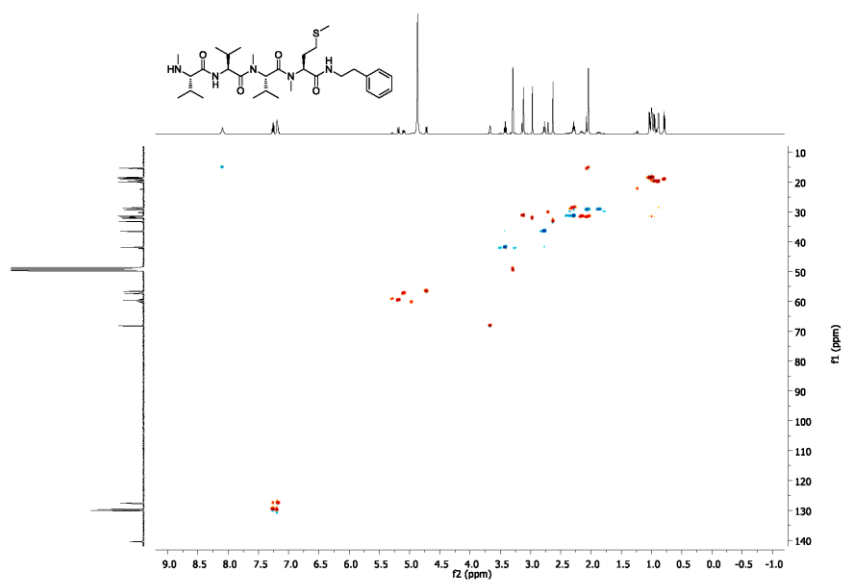
Figure S14. HSQC (methanol-*d₄*) spectrum of **2**.

S30

160
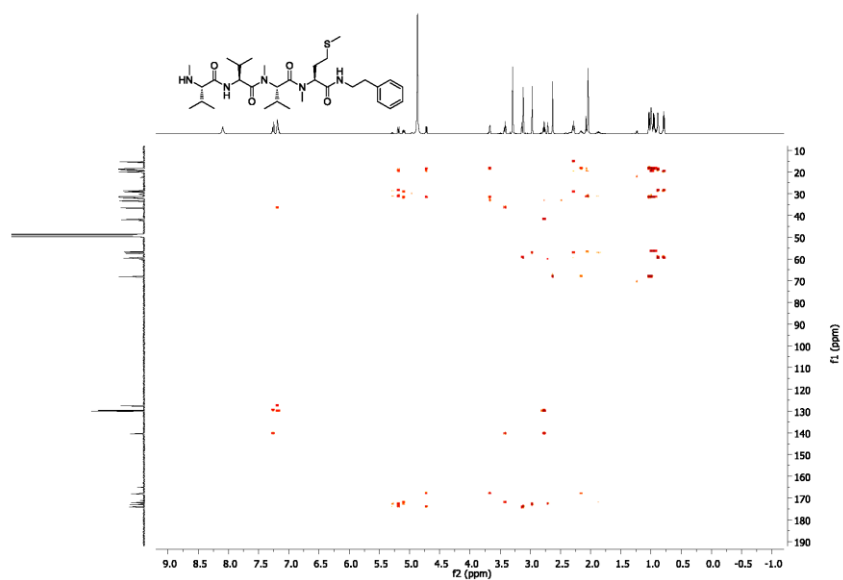
Figure S15. HMBC (methanol-*d*₄) spectrum of **2**.

Figure S16. ^{1}H NMR (500 MHz, methanol-d_4) spectrum of **3**.

Figure S17. ^{13}C NMR (125 MHz, methanol-d_4) spectrum of **3**.

S32

162
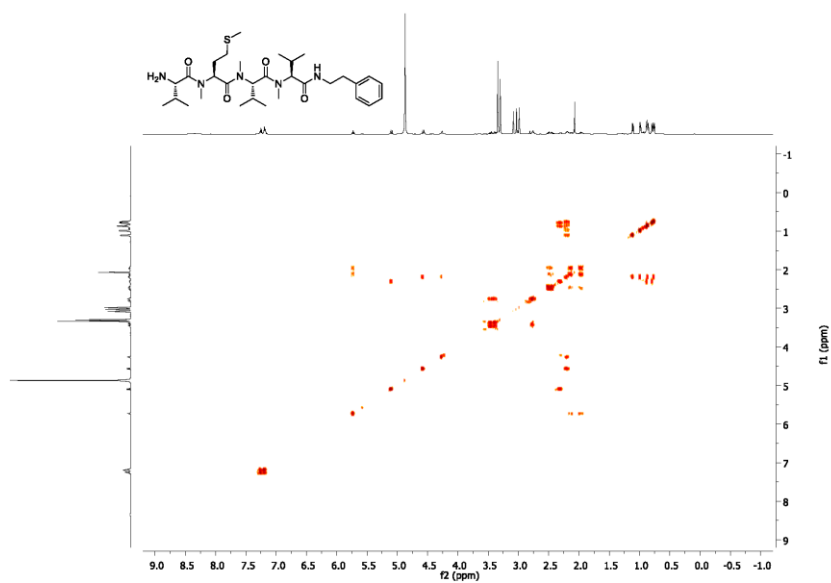
Figure S18. COSY (methanol-*d*4) spectrum of **3**.
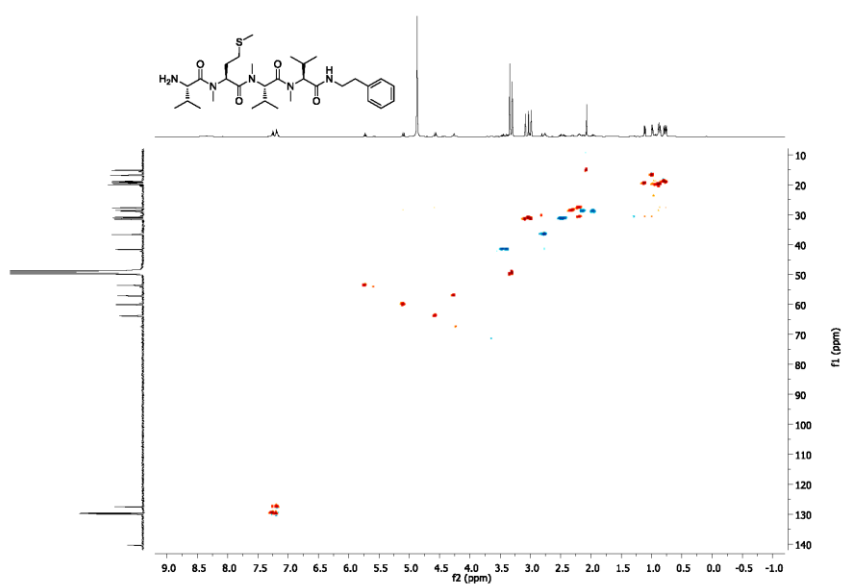
Figure S19. HSQC (methanol-*d*4) spectrum of **3**.

S33

163
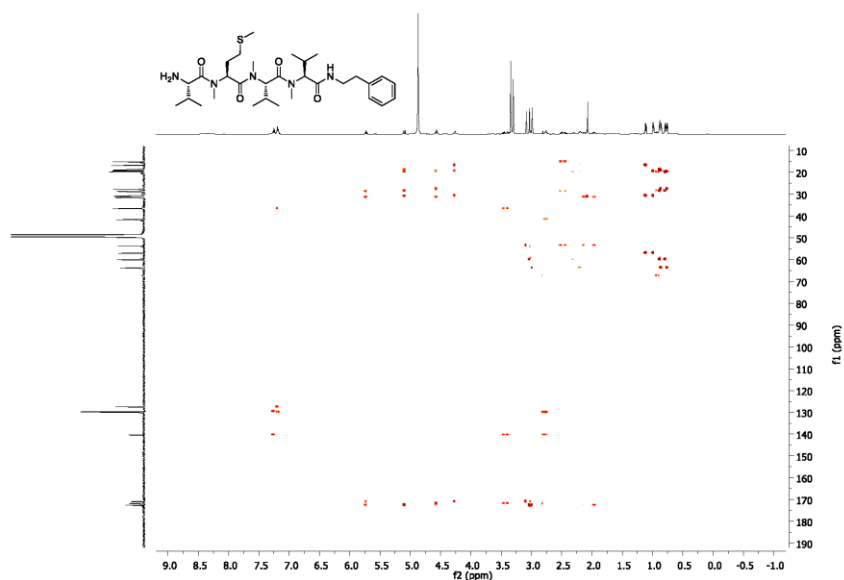
Figure S20. HMBC (methanol-*d*₄) spectrum of **3**.

Figure S21. ^{1}H NMR (500 MHz, methanol-d_4) spectrum of **4**.

Figure S22. ^{13}C NMR (125 MHz, methanol-d_4) spectrum of **4**.
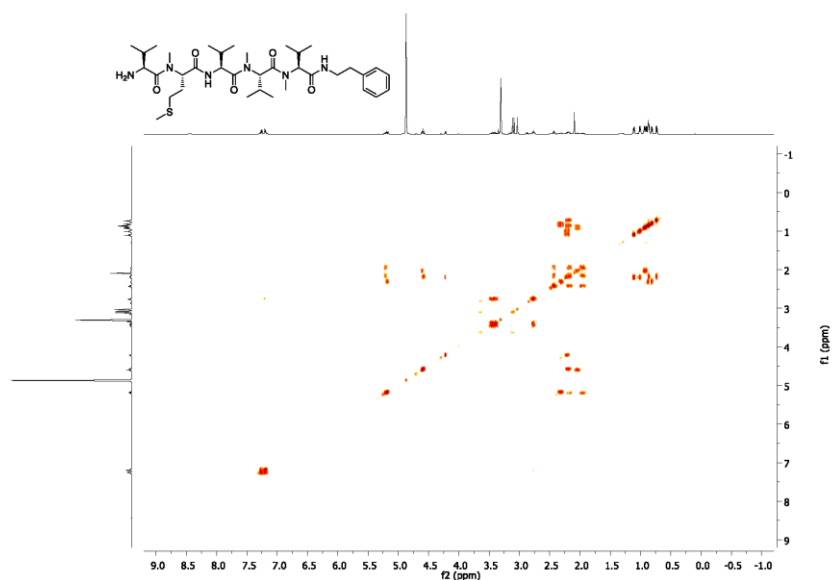
Figure S23. COSY (methanol-*d*₄) spectrum of **4**.
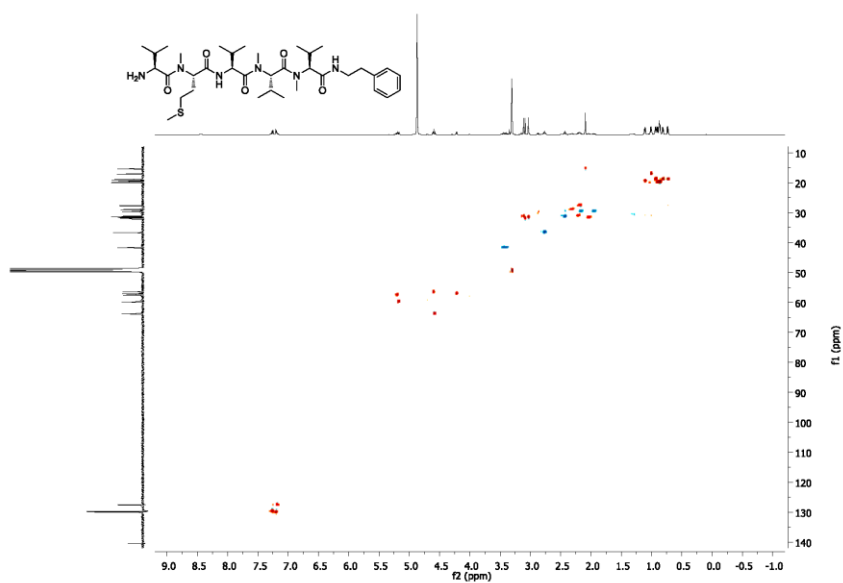
Figure S24. HSQC (methanol-*d*₄) spectrum of **4**.

S36

166
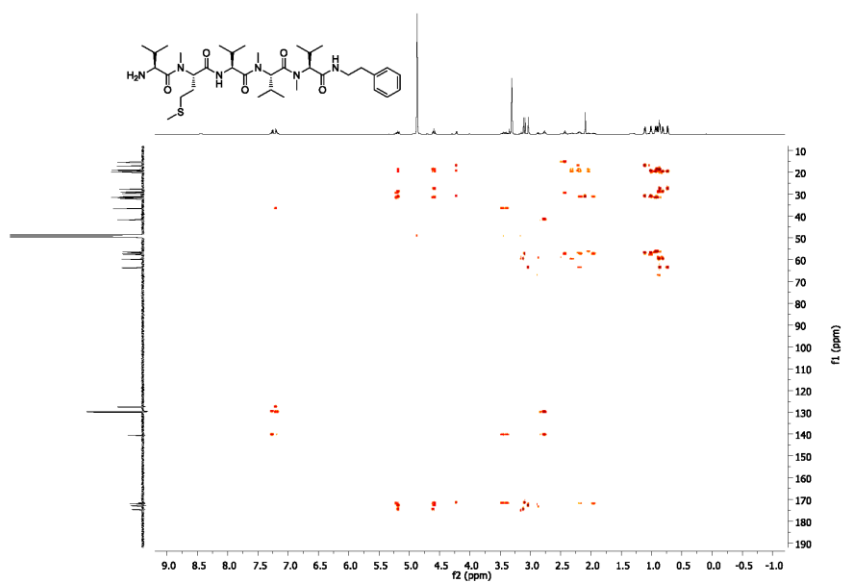
Figure S25. HMBC (methanol-d_4) spectrum of **4**.
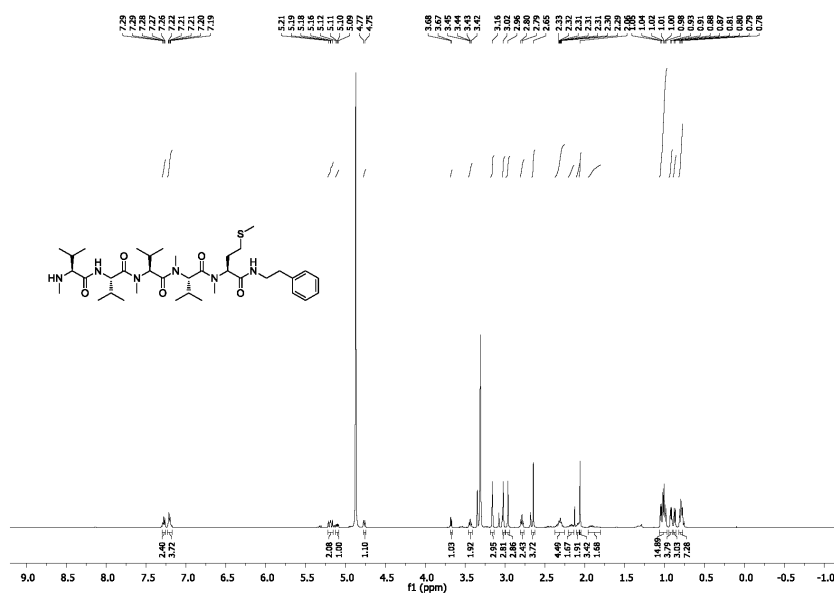
Figure S26. ¹H NMR (500 MHz, methanol-*d₄*) spectrum of **5**.

Figure S27. ¹³C NMR (125 MHz, methanol-*d₄*) spectrum of **5**.

S38

168
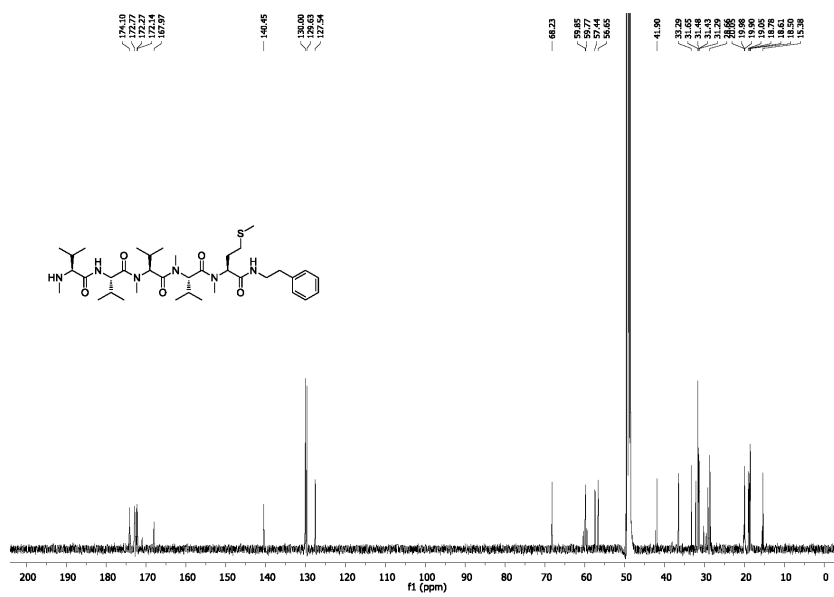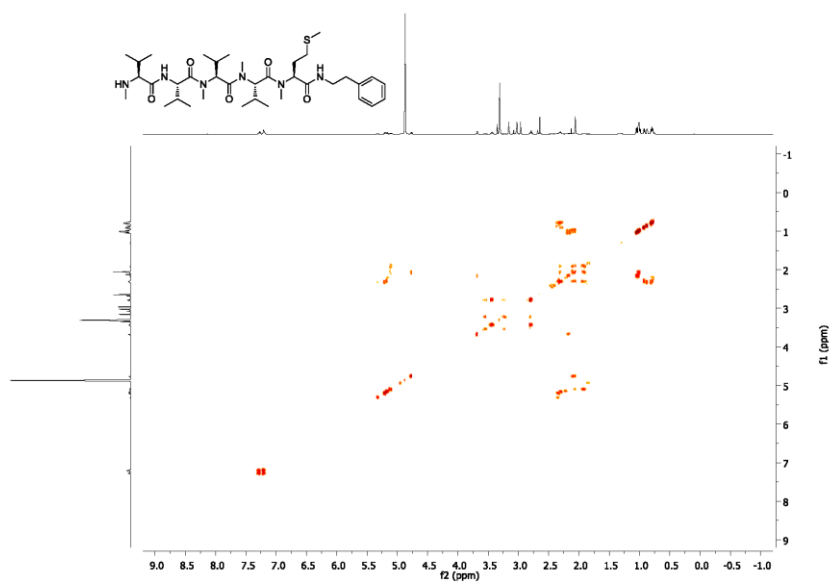
Figure S28. COSY (methanol-*d*4) spectrum of **5**.
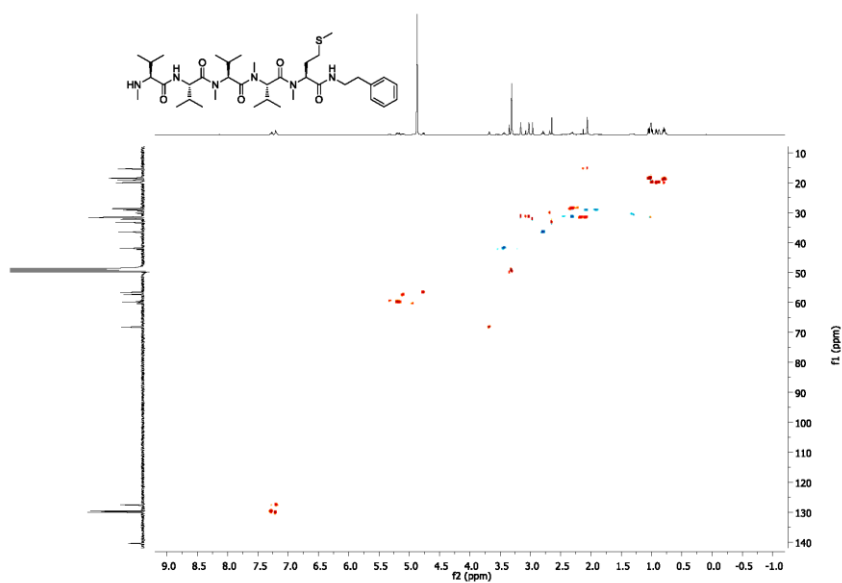
Figure S29. HSQC (methanol-*d*4) spectrum of **5**.

S39

169
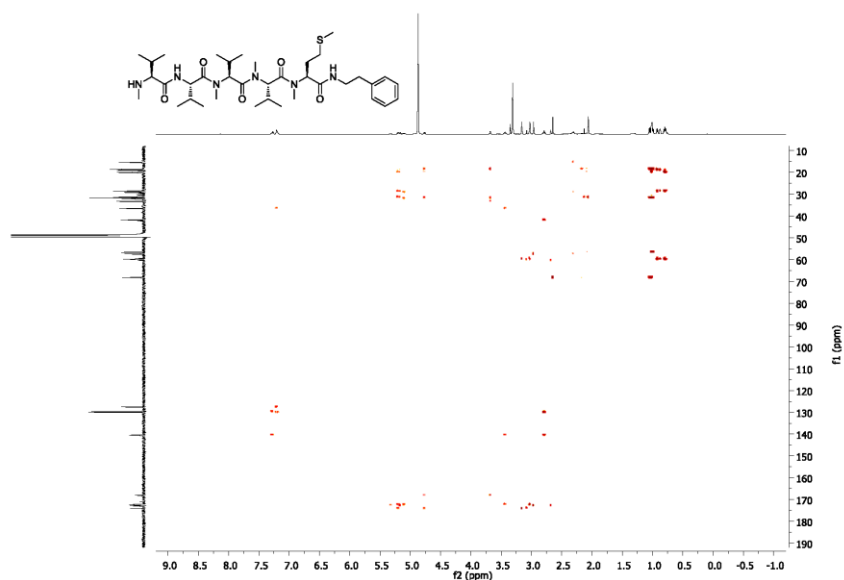
Figure S30. HMBC (methanol-d_4) spectrum of **5**.

Figure S31. ^{1}H NMR (500 MHz, methanol-d_4) spectrum of **6**.

Figure S32. ^{13}C NMR (125 MHz, methanol-d_4) spectrum of **6**.
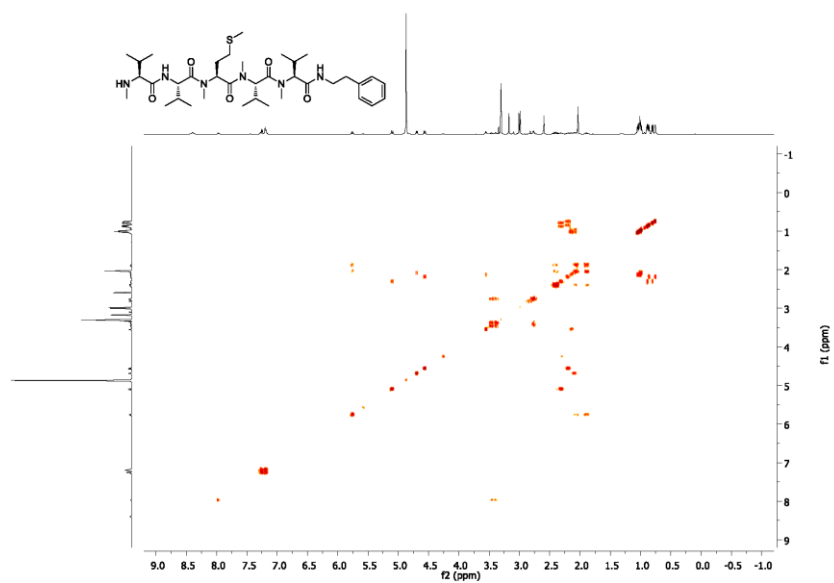
Figure S33. COSY (methanol-*d*₄) spectrum of **6**.
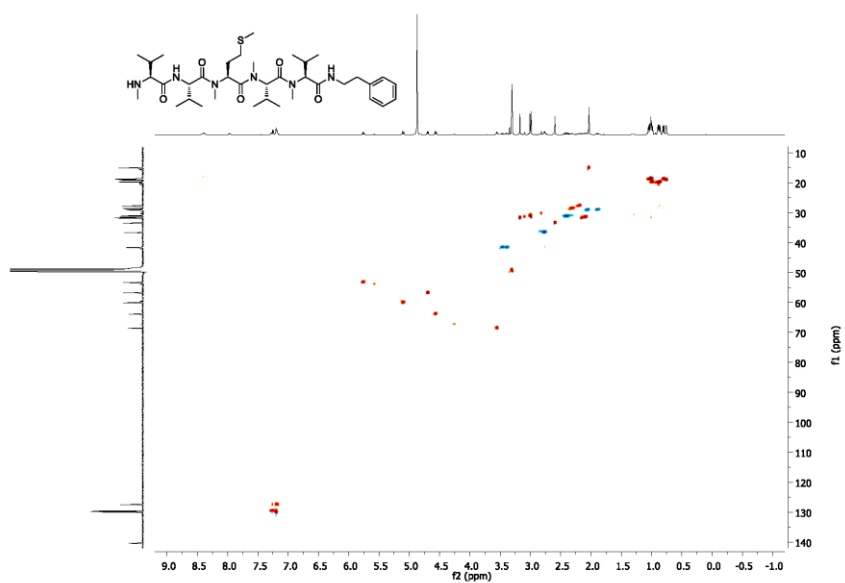
Figure S34. HSQC (methanol-*d*₄) spectrum of **6**.

S42

172
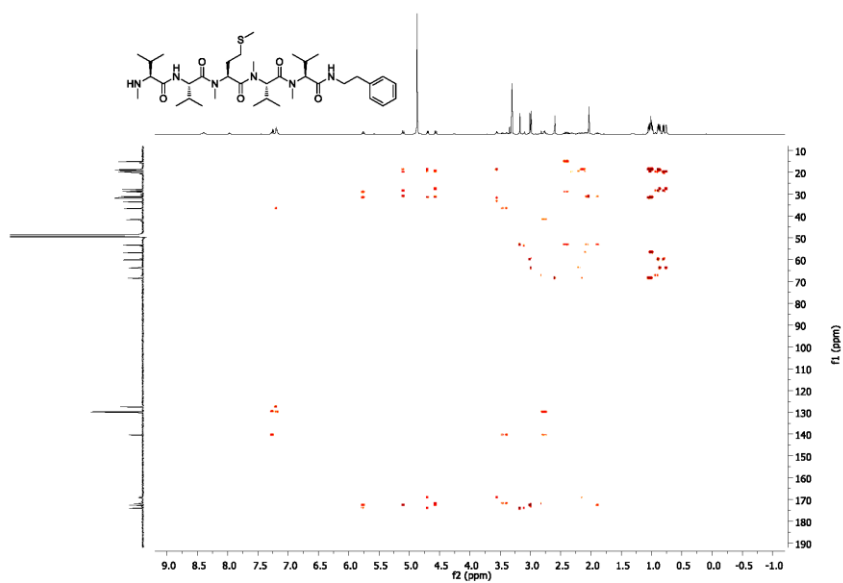
Figure S35. HMBC (methanol-d_4) spectrum of **6**.
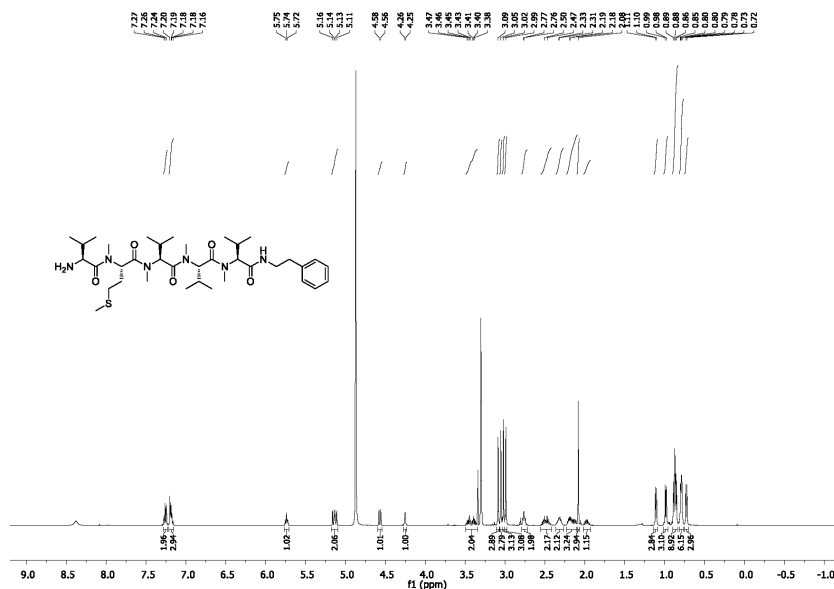
Figure S36. ^{1}H NMR (500 MHz, methanol-d_4) spectrum of **7**.
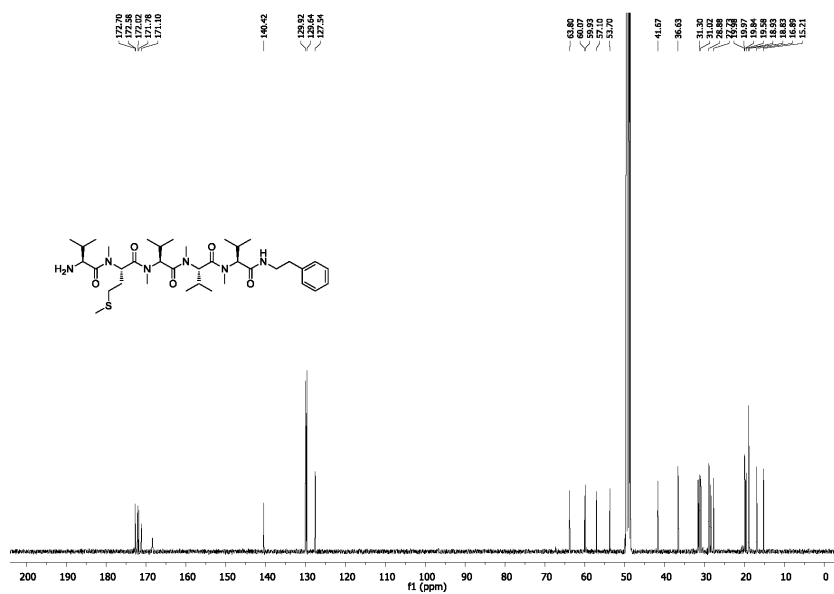
Figure S37. ^{13}C NMR (125 MHz, methanol-d_4) spectrum of **7**.
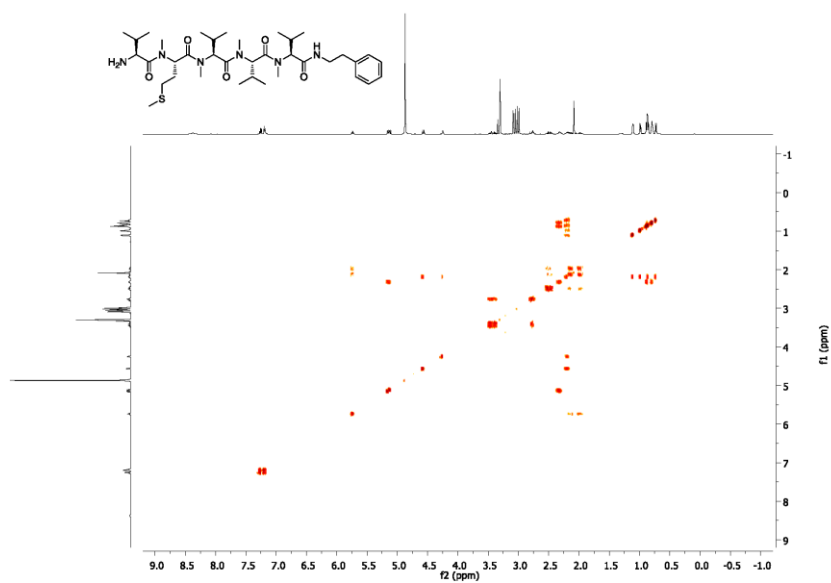
Figure S38. COSY (methanol-*d*4) spectrum of **7**.
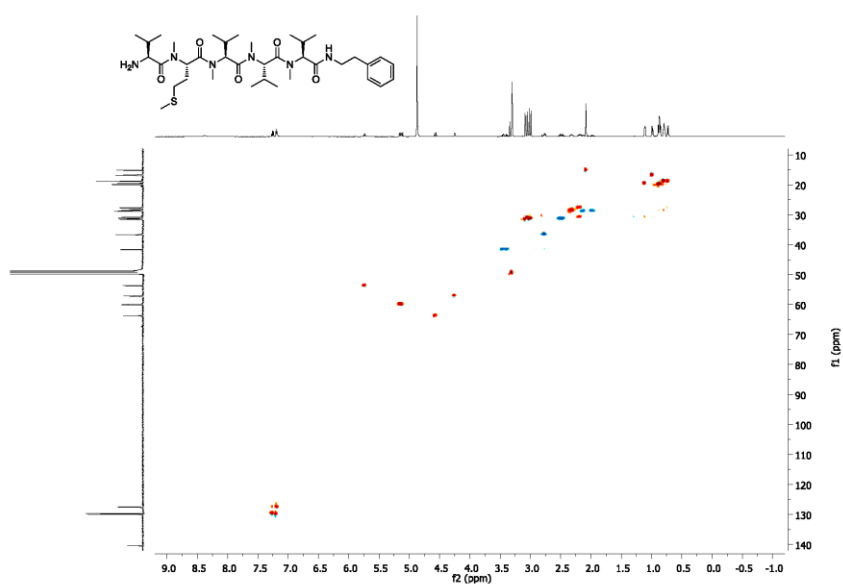
Figure S39. HSQC (methanol-*d*4) spectrum of **7**.

S45

175
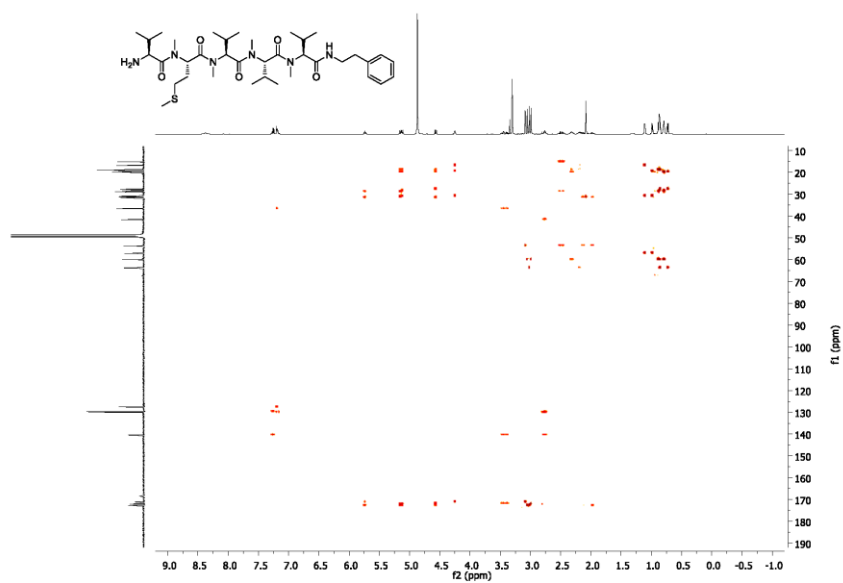
Figure S40. HMBC (methanol-d_4) spectrum of **7**.

Figure S41. ^{1}H NMR (500 MHz, methanol-d_4) spectrum of **10**.
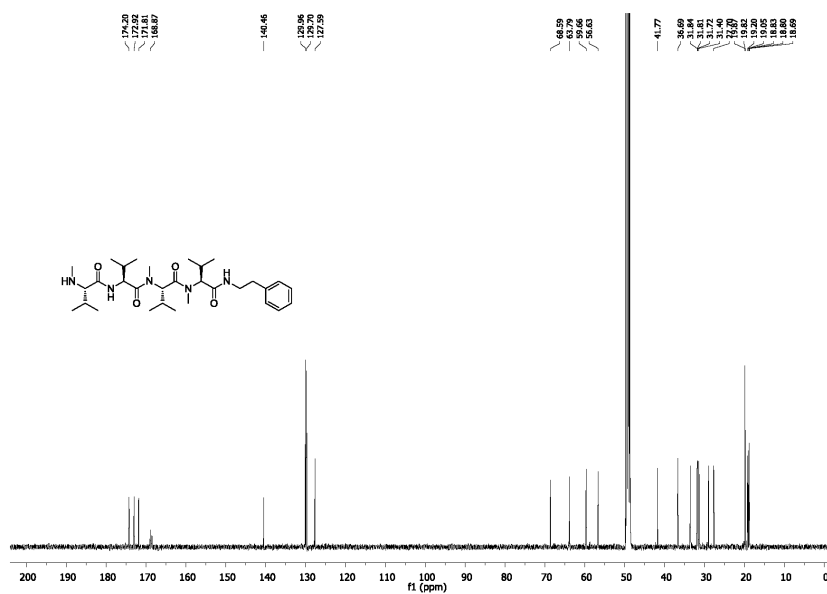
Figure S42. ^{13}C NMR (125 MHz, methanol-d_4) spectrum of **10**.
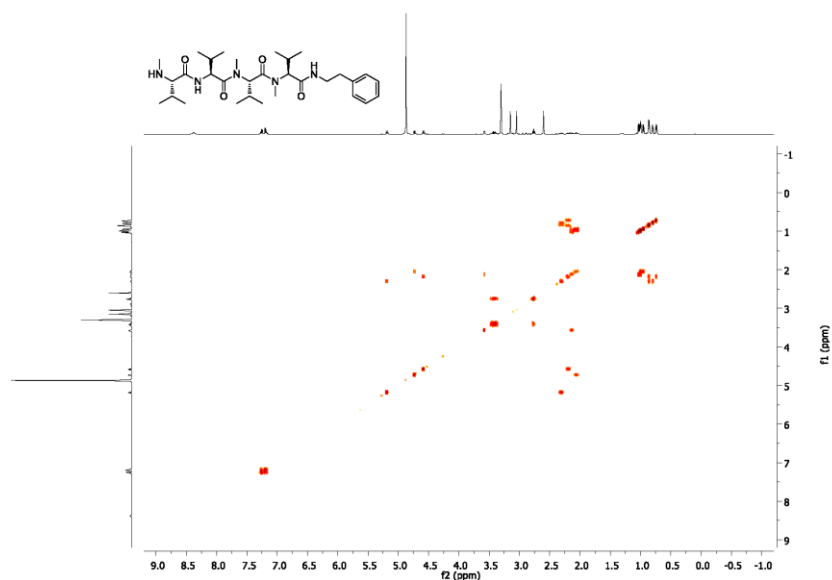
Figure S43. COSY (methanol-d_4) spectrum of **10**.

Figure S44. HSQC (methanol-d_4) spectrum of **10**.

S48

178

Figure S45. HMBC (methanol-d_4) spectrum of **10**.
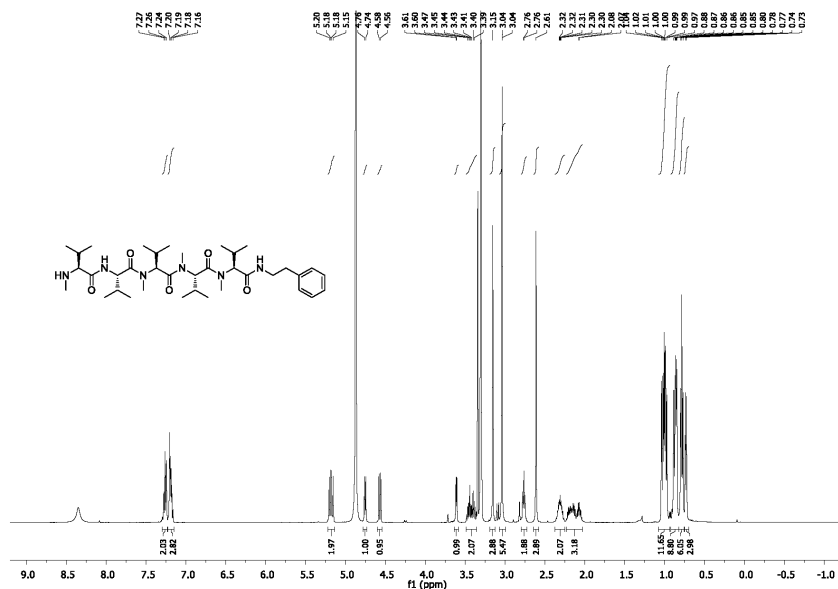
Figure S46. ^{1}H NMR (500 MHz, methanol-d_4) spectrum of **12**.
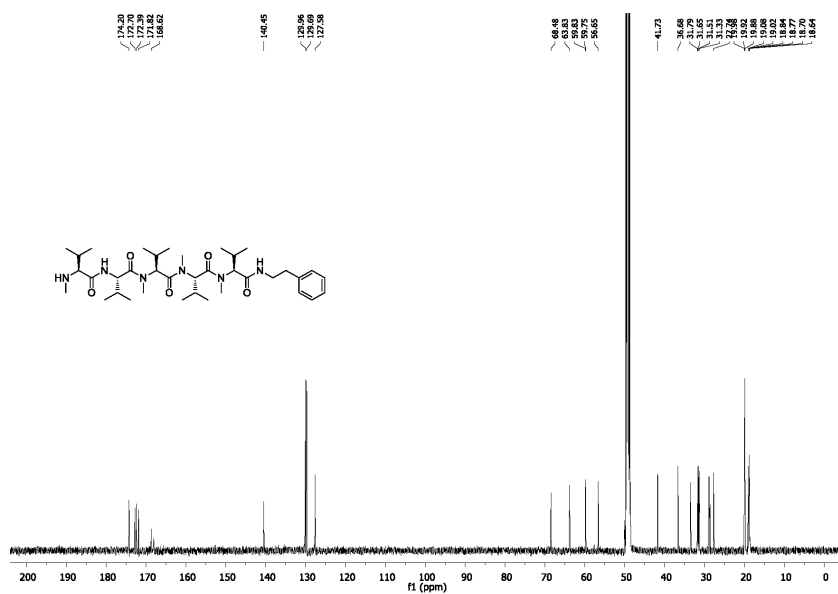
Figure S47. ^{13}C NMR (125 MHz, methanol-d_4) spectrum of **12**.

Figure S48. COSY (methanol-*d*₄) spectrum of **12**.

Figure S49. HSQC (methanol-*d*₄) spectrum of **12**.

S51
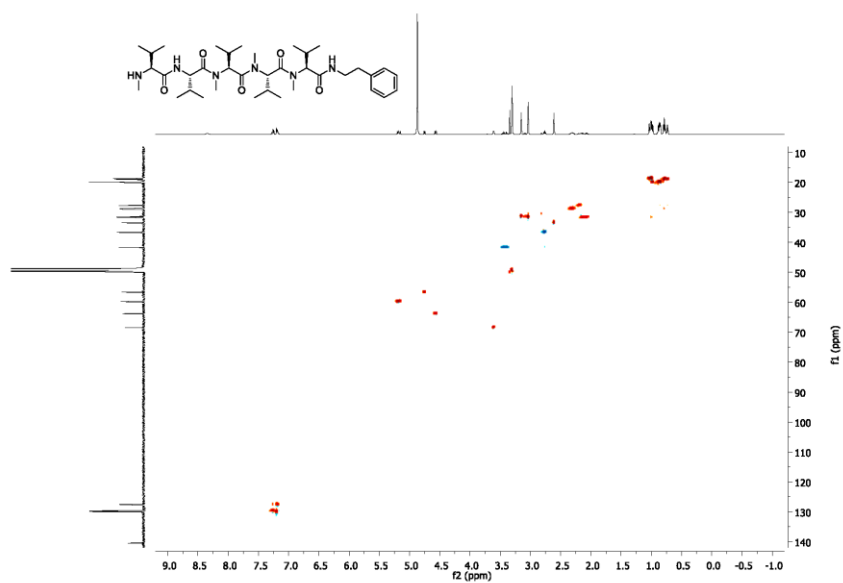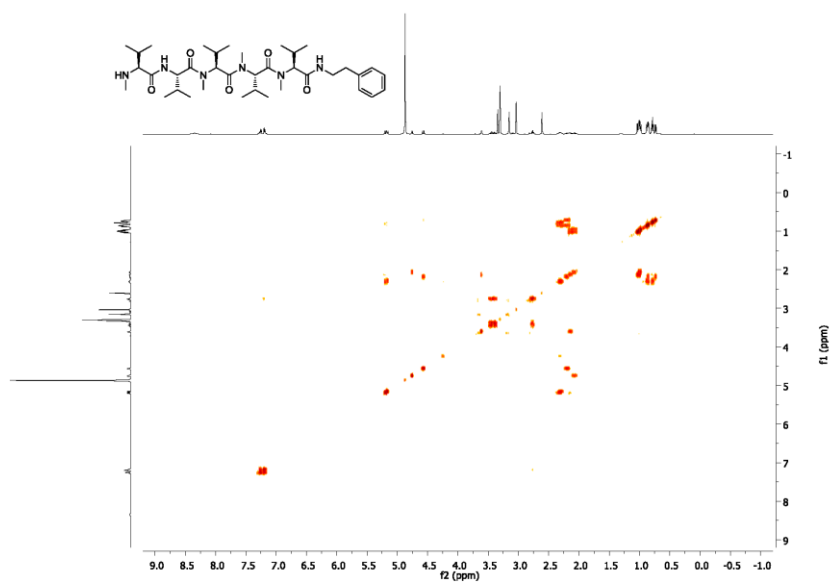
181
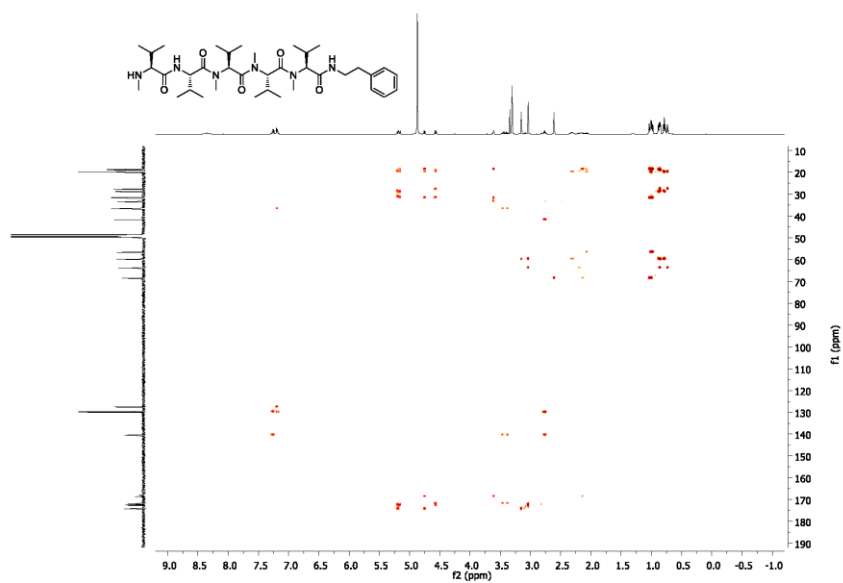
Figure S50. HMBC (methanol-d_4) spectrum of **12**.

References

(1) Wesche, F.; Adihou, H.; Kaiser, A.; Wurglics, M.; Schubert-Zsilavecz, M.; Kaiser, M.; Bode, H. B. *J. Med. Chem.* **2018**, *61*, 3930–3938.

(2) Sable, G. A.; Park, J.; Kim, H.; Lim, S. J.; Jang, S.; Lim, D. *Eur. J. Org. Chem.* **2015**, 7043–7052.

(3) Cai, X.; Nowak, S.; Wesche, F.; Bischoff, I.; Kaiser, M.; Fürst, R.; Bode, H. B. *Nat. Chem.* **2017**, *9*, 379–386.

(4) Schimming O., Fleischhacker F., Nollmann F. I., Bode H. B. *Chembiochem* **2014**, *15*, 1290–1294.

5.3 Production of a Photohexapeptide Library from Entomopathogenic *Photorhabdus asymbiotica* PB68.1

Authors:

Lei Zhao[1,2] and Helge B. Bode*[,1,3]

[1]Molecular Biotechnology, Department of Biosciences, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

[2]Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, 210014 Nanjing, China

[3]Buchmann Institute for Molecular Life Sciences (BMLS), Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

*Corresponding author

Online access:

https://pubs.rsc.org/ko/content/articlehtml/2019/ob/c9ob01489f

Declaration on the contribution of the authors

Publication: Production of a Photohexapeptide Library from Entomopathogenic *Photorhabdus asymbiotica* PB68.1

Status: published

Journal: Organic & Biomolecular Chemistry

Authors: <u>Lei Zhao (LZ)</u>, Helge B. Bode (HBB)

What did the doctoral candidate or the co-authors contribute individually to the dissertation?

(1) Development and planning

<u>LZ (50%)</u>, HBB (50%)

(2) Performance of individual research and experiments

<u>LZ (100%): strain construction, extract preparation, HPLC-MS analysis, labeling experiments, compound isolation, chemical synthesis, NMR measurement, product quantification, bioactivity testing</u>

(3) Collection of data and preparation of figures

<u>LZ (80%): collection of NMR data, collection of HPLC-MS data including quantification data, preparation of figures;</u> HBB (20%): preparation of figures

(4) Analysis and interpretation of data

<u>LZ (70%): analysis of NMR data, analysis of HPLC-MS data including quantification data, interpretation of bioactivity;</u> HBB (30%): analysis of HPLC-MS data

(5) Preparation of manuscript

<u>LZ (80%)</u>, HBB (20%)

Herewith approving the indications above

_____   _____
date/place   signature doctoral candidate

_____   _____
date/place   signature supervisor

_____   _____
date/place   If necessary, signature corresponding author

COMMUNICATION

Check for updates

Production of a photohexapeptide library from entomopathogenic *Photorhabdus asymbiotica* PB68.1†

Lei Zhao [a,b] and Helge B. Bode [*a,c]

A new natural product compound library, photohexapeptide library, was identified from entomopathogenic *Photorhabdus asymbiotica* PB68.1 after the NRPS-encoding gene *phpS* was activated *via* promoter exchange. Peptide structures, including the absolute configurations of amino acids, were determined by using a combination of bioinformatics analysis and isotopic labelling experiments followed by detailed HPLC-MS analysis. Additionally, their structures were confirmed by chemical synthesis and NMR after preparative isolation. The chemical diversity of the photohexapeptides results from promiscuous adenylation domain specificity being an excellent example of how to create libraries in nature.

Natural product compound libraries have been observed in various organisms including bacteria,[1] fungi,[2] insects[3] and plants.[4] The production of such chemically diverse libraries often is the result of the catalytic promiscuity of specialized metabolic enzymes that synthesize compounds not immediately required for the survival of the host organisms, but nonetheless contribute to the reproductive success of populations in complex ecosystems.[5] Therefore, the generation of compound libraries is generally regarded as the biotic responses of organisms to challenging and changing environmental pressures.[5] Another explanation for such libraries might be that they represent a "snapshot" of ongoing compound development for a specific ecological application. This "Screening Hypothesis" would be analogous to high-throughput screening programmes carried out during medical drug development.[6]

Nonribosomally derived peptides are well known as an important source of therapeutic drugs such as vancomycin, cyclosporine and daptomycin. The assembly of these peptides is catalyzed by large multifunctional protein complexes termed nonribosomal peptide synthetases (NRPSs). NRPSs are composed of modules, each often consisting of condensation (C), adenylation (A) and thiolation (T) domains together to carry out one cycle of amino acid chain extension.[7] Several modules act as a conveyor belt for the stepwise incorporation of amino acids usually in a collinear fashion.[1] On rare occasions, the high promiscuity of A domains[8] and the iterative use of modules[2] or domains[9] result in the generation of peptide libraries.

During our search for new natural products from entomopathogenic bacteria of the genera *Xenorhabdus* and *Photorhabdus* that live in symbiosis with nematodes,[10] we found a library of new peptides, named photohexapeptides, after their biosynthetic gene *phpS* was activated in *Photorhabdus asymbiotica* PB68.1. Here, we describe the identification of photohexapeptides by using a combination of bioinformatics analysis, promoter exchange, isotopic labelling experiments, HPLC-MS analysis, NMR, and chemical synthesis.

In the genome of *P. asymbiotica* PB68.1, the *phpS* (photohexapeptide synthetase) gene (*ppb6_04101*, 20.4 kb, GenBank MN099046.1) was identified encoding an unknown NRPS. Detailed analysis showed that PhpS consists of six modules with overall 19 domains (Fig. 1) and PhpS was thus expected to produce a hexapeptide. Since no such peptide could be identified in *P. asymbiotica* PB68.1, a promoter exchange approach was applied to *phpS* as previously described[11] that mimics an "overproducing" or a "knock-out" mutant of *P. asymbiotica* PB68.1 with or without L-arabinose induction, respectively, thus resulting in the detection of two separate new compounds (**1** and **1-OMe**) and four clusters of new compounds (**2–5**) based on HPLC-MS analysis (Fig. 2).

Detailed HPLC-MS analysis showed that the MS² fragmentation patterns of these new compounds are composed of 99 and 113 Da mass units, indicating valine, leucine and/or isoleucine as building blocks of these peptides. Compound **1-OMe**, for instance, contains six 113 Da mass units (Fig. 3a),

*a*Molecular Biotechnology, Department of Biosciences, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany. E-mail: h.bode@bio.uni-frankfurt.de
*b*Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, 210014 Nanjing, China
*c*Buchmann Institute for Molecular Life Sciences (BMLS), Goethe University Frankfurt, 60438 Frankfurt am Main, Germany
†Electronic supplementary information (ESI) available: Experimental procedures and supplementary tables and figures. See DOI: 10.1039/c9ob01489f
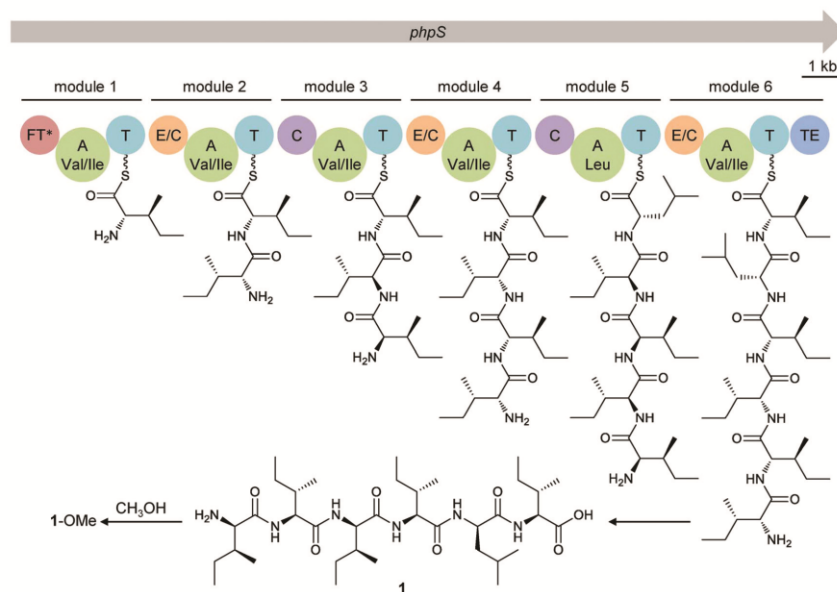
7858 | *Org. Biomol. Chem.*, 2019, **17**, 7858–7862

This journal is © The Royal Society of Chemistry 2019

Organic & Biomolecular Chemistry

Fig. 1 Domain organization of the NRPS PhpS corresponding to the production of 1 and 1-OMe as examples. Domains: FT: formyl transferase (*nonfunctional), A: adenylation, T: thiolation, E/C: dual epimerization/condensation, C: condensation, TE: thioesterase.
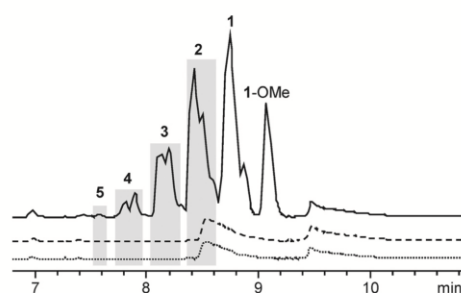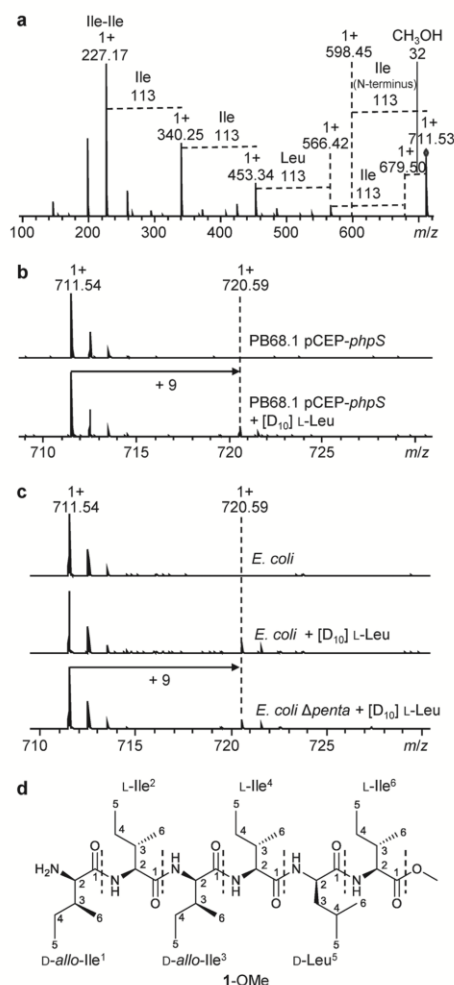

Fig. 2 HPLC-MS analysis of XAD extracts (with methanol) from *P. asymbiotica* PB68.1 wild type (dotted line) and *P. asymbiotica* PB68.1 pCEP-*phpS* without (dashed line) and with (continuous line) L-arabinose induction. Base-peak chromatograms (BPCs) are shown.

suggesting that six leucine and/or isoleucine residues are present in **1**-OMe. In order to differentiate between leucine or isoleucine, labelling experiments in lysogeny broth (LB) medium with deuterated leucine were performed revealing the incorporation of only one leucine moiety (Fig. 3b). In addition, a mass unit of 32 Da found at the C-terminus (Fig. 3a) suggests **1**-OMe to be a methyl ester.

To determine the position of the only leucine and confirm the methyl ester in **1**-OMe, compound **1**-OMe was isolated due

to its good resolution, and its structure was subsequently elucidated based on detailed NMR experiments (Table 1, Fig. 4). Thereby, compound **1**-OMe was confirmed to be a methyl ester with leucine residue at the fifth position, corresponding to the predicted specificity of the A5 domain. The A5 domain shows a specificity for phenylalanine when compared to the other five A domains that all show identical specificity codes for valine incorporation (Table S1†).[12,13]

To address the absolute configurations of the amino acids in **1**-OMe, the dual epimerization/condensation (E/C) domains, which transform the initially bound L-amino acids into their D-form,[14,15] were identified in PhpS using antiSMASH (Fig. 1).[13] The alternating E/C and C domains in PhpS suggest **1**-OMe with the configuration sequence DLDLDL. Therefore, the configuration of leucine at the fifth position is suggested to be D as was confirmed by labelling experiments with deuterated L-[D₁₀]leucine. When L-[D₁₀]leucine was fed to a transaminase-deficient *E. coli* strain expressing *phpS*, a mass shift of 10 Da would be indicative of the incorporation of L-leucine while the observed 9 Da mass shift suggests an epimerization and the loss of the deuterium at the α-carbon (Fig. 3c) as shown previously.[16] It should be noted that in the case of isoleucine, the E/C domain can only convert L-isoleucine to its diastereomer, D-*allo*-isoleucine, and not to its enantiomer, D-isoleucine, since epimerization would only affect the α-carbon, leaving the β-methyl centre unchanged.[14,15,17] Therefore, compound **1**-OMe features two D-

This journal is © The Royal Society of Chemistry 2019

Org. Biomol. Chem., 2019, **17**, 7858–7862 | **7859**

Communication

Fig. 3 Structure identification of 1-OMe using labelling experiments combined with HPLC-MS analysis. (a) MS^2 fragmentation pattern of 1-OMe; (b) MS data of 1-OMe from labelling experiments with/without [D$_{10}$] L-Leu fed to *P. asymbiotica* PB68.1 pCEP-*phpS* in LB medium; (c) MS data of 1-OMe from labelling experiments with/without [D$_{10}$] L-Leu fed to *E. coli* or transaminase-deficient *E. coli* Δ*penta* expressing *phpS* in LB medium; (d) structure of 1-OMe.
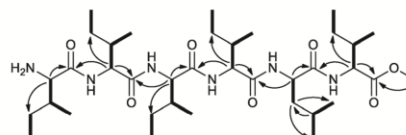
Table 1 ^{1}H (500 MHz) and ^{13}C (125 MHz) NMR data of 1-OMe in methanol-d$_4$ (δ in ppm)

Subunit^a	Position	δ_C	δ_H, mult. (J in Hz)
Ile^1	1	170.9 (C)	
	2	58.9 (CH)	3.81, d (5.0)
	3	38.3 (CH)	1.99, m
	4	27.1 (CH$_2$)	1.53, 1.30, m
	5	12.0 (CH$_3$)	0.99, overlap
	6	14.1 (CH$_3$)	0.99, overlap
Ile^2	1	173.9 (C)	
	2	59.7 (CH)	4.31, d (7.7)
	3	37.8 (CH)	1.89, m
	4	26.1 (CH$_2$)	1.41, 1.22, m
	5	12.1 (CH$_3$)	0.96, overlap
	6	16.2 (CH$_3$)	0.96, overlap
Ile^{3,b}	1	173.6 (C)	
	2	59.7 (CH)	4.24, d (7.1)
	3	38.0 (CH)	1.91, m
	4	26.1 (CH$_2$)	1.41, 1.22, m
	5	11.3 (CH$_3$)	0.93, overlap
	6	16.1 (CH$_3$)	0.93, overlap
Ile^{4,b}	1	173.8 (C)	
	2	58.3 (CH)	4.42, d (5.3)
	3	38.1 (CH)	1.98, m
	4	27.7 (CH$_2$)	1.40, 1.23, m
	5	11.5 (CH$_3$)	0.92, overlap
	6	15.3 (CH$_3$)	0.92, overlap
Leu^5	1	174.7 (C)	
	2	53.2 (CH)	4.47, dd (9.2, 5.8)
	3	42.0 (CH$_2$)	1.6, m
	4	26.1 (CH)	1.65, m
	5	23.5 (CH$_3$)	0.96, overlap
	6	22.0 (CH$_3$)	0.92, overlap
Ile^6	1	173.6 (C)	
	2	58.5 (CH)	4.35, d (6.3)
	3	38.3 (CH)	1.93, m
	4	26.6 (CH$_2$)	1.42, 1.23, m
	5	11.7 (CH$_3$)	0.92, overlap
	6	16.1 (CH$_3$)	0.92, overlap
O–CH$_3$		52.7	3.70, s

^{a} Amino acid numbering from the N- to C-terminus. ^{b} Each other can be exchanged.

Fig. 4 COSY (bold) and key HMBC (arrow) correlations of 1-OMe.

allo-isoleucines at the first and third positions, and the structure of **1**-OMe is shown in Fig. 3d.

Similarly, the structure of **1** was deduced from detailed HPLC-MS analysis and labelling experiments (Fig. S1†). As in **1**-OMe, only one leucine is incorporated into **1** (Fig. S1b†), therefore we hypothesized that leucine of **1** is also at the fifth position with D-configuration. To confirm the proposed struc-

ture, including its absolute configuration as DLDLDL, compound **1** was synthesized by using a solid-phase peptide synthesis method (Fig. S2†)[18] and compared to natural **1** regarding retention time and MS^2 fragmentation patterns (Fig. S3†). Based on these efforts, the structure of **1** was unambiguously elucidated as shown in Fig. 1. Compound **1**-OMe is the methyl ester of **1**, which might be derived from the esterification of **1** with methanol during sample work-up. Accordingly, no **1**-OMe was detected when the culture supernatants and samples were treated with acetonitrile instead of methanol (Fig. S4†).

The major challenges were to identify the compounds in clusters **2–5** due to serious peak overlapping. Nonetheless, we were able to figure out the numbers of incorporated valine, leucine and isoleucine by labelling experiments combined with HPLC-MS analysis. These results suggest that only one leucine is incorporated into these compounds with one valine and four isoleucines in cluster **2**, two valines and three isoleucines in cluster **3**, three valines and two isoleucines in cluster **4**, and four valines and one isoleucine in cluster **5** (Fig. S5†). Compound **6** consisting of five valines and one leucine without isoleucine was also detected by HPLC-MS, albeit in a much lower amount, and was confirmed by chemical synthesis (Fig. S6†). Although these compounds in clusters **2–5** have a similar retention time, they might show differences in the fragmentation patterns by HPLC-MS analysis. Applying this method we obtained the sequence of thirteen compounds with valine/isoleucine incorporation varying from one to four, only by considering the major fragmentation patterns (Fig. S7†). Consequently, we guessed there are many more than identified peptides since A1–A4 and A6 domains are flexible for valine or isoleucine. Therefore, in total, we would expect 32 derivatives (photohexapeptide-1 to 32, Table 2) from simple mathematical calculations of $2 \times 2 \times 2 \times 2 \times 1 \times 2$ with one to five

valines or isoleucines, respectively. To prove our hypothesis, cluster **2** was selected as an example and analysed in detail. All five expected compounds (**2a–e**) in cluster **2** were synthesized and their mixture was compared to natural compounds (Fig. S8†). These results suggest that a peptide library containing 32 derivatives was produced by NRPS PhpS from *P. asymbiotica* PB68.1 due to the promiscuous A domain specificity.

The synthesis also provided a sufficient amount of peptides for bioactivity testing but no activity against Gram-positive bacteria (*Micrococcus luteus*) or fungi (*Saccharomyces cerevisiae*) was observed for **1** or **1**-OMe. The synthetic standard **1** additionally enabled the quantification of all produced peptides in the promoter exchange mutant based on HPLC-MS, resulting in up to 210 mg L^{-1} of all photohexapeptides.

It is worth noting that the NRPS PhpS contains a formyl transferase (FT) domain as the first domain of the FT-A-T initiation module (Fig. 1). The formyl group seems to be important for the bioactivity of some nonribosomal peptides. For example, it is essential for the pore forming activity of gramicidin D.[19,20] However, no formylated products at the first amino acid position were found in photohexapeptides. During our in-depth investigation on *phpS*, we found a formamidase-encoding gene (*ppb6_04102*, 1.0 kb) located 1.3 kb upstream from *phpS* (Fig. S9†). Thus we postulated that the formyl group might be removed by the formamidase during the biosynthesis of photohexapeptides. To test the hypothesis, we expressed PhpS alone in *E. coli* without Ppb6_04102, but still no formylated products except photohexapeptides were detected (Fig. S10†). This suggests that the FT domain might be inactive in PhpS. As FT domains often contain a conserved Asn-His-Asp catalytic triad,[21] we compared the FT domain of PhpS with previously reported active FT domains, but found PhpS to have these three key catalytic residues (data not shown). From sequence alignment, also all residues required for binding of N^{10}-formyltetrahydrofolate (fTHF, as the formyl donor) are present, suggesting other reasons why this FT domain is inactive.

In summary, a library of photohexapeptides was generated in *P. asymbiotica* PB68.1 after the silent gene *phpS* was activated *via* promoter exchange. The structures of photohexapeptides were unequivocally elucidated by using a combination of biological and chemical methods. Although relaxed substrate specificity of A domains has been shown in several other NRPS systems resulting in the generation of peptide diversity such as microcystins,[8] xenematides,[22] and noursamycins,[23] PhpS is unusual since it possesses five out of six A domains which can utilize two different amino acids, leading to a very large number of products. Moreover, photohexapeptides enrich the family of rare linear D-/L-peptide natural products, including kolossin A,[24] gramicidin A,[25] and feglymycin.[26] Further studies will pursue the biological functions of photohexapeptides, especially in the complex life cycle of *P. asymbiotica* PB68.1 including bacteria, nematode hosts and insect prey[27] although they did not show antibacterial and antifungal activity.

Table 2 Structure sequences of photohexapeptides

m/z [M + H]$^+$	Compd name	No.	Amino acid sequence (D-L-D-L-D-L)	Confirmed by synthesis
697.52	PHP^a-1	1	aIle^b-Ile-aIle-Ile-Leu-Ile	√
683.50	PHP-2	2a	aIle-Ile-aIle-Ile-Leu-Val	√
	PHP-3	2b	aIle-Ile-aIle-Val-Leu-Ile	√
	PHP-4	2c	aIle-Ile-Val-Ile-Leu-Ile	√
	PHP-5	2d	aIle-Val-aIle-Ile-Leu-Ile	√
	PHP-6	2e	Val-Ile-aIle-Ile-Leu-Ile	√
669.49	PHP-7	3a	aIle-Ile-aIle-Val-Leu-Val	
	PHP-8	3b	aIle-Ile-Val-Ile-Leu-Val	
	PHP-9	3c	aIle-Ile-Val-Val-Leu-Ile	
	PHP-10	3d	aIle-Val-aIle-Ile-Leu-Val	
	PHP-11	3e	aIle-Val-aIle-Val-Leu-Ile	
	PHP-12	3f	aIle-Val-Val-Ile-Leu-Ile	
	PHP-13	3g	Val-Ile-aIle-Ile-Leu-Val	
	PHP-14	3h	Val-Ile-aIle-Val-Leu-Ile	
	PHP-15	3i	Val-Ile-Val-Ile-Leu-Ile	
	PHP-16	3j	Val-Val-aIle-Ile-Leu-Ile	
655.47	PHP-17	4a	aIle-Ile-Val-Val-Leu-Val	
	PHP-18	4b	aIle-Val-aIle-Val-Leu-Val	
	PHP-19	4c	aIle-Val-Val-Ile-Leu-Val	
	PHP-20	4d	aIle-Val-Val-Val-Leu-Ile	
	PHP-21	4e	Val-Ile-aIle-Val-Leu-Val	
	PHP-22	4f	Val-Ile-Val-Ile-Leu-Val	
	PHP-23	4g	Val-Ile-Val-Val-Leu-Ile	
	PHP-24	4h	Val-Val-aIle-Ile-Leu-Val	
	PHP-25	4i	Val-Val-aIle-Val-Leu-Ile	
	PHP-26	4j	Val-Val-Val-Ile-Leu-Ile	
641.46	PHP-27	5a	aIle-Val-Val-Val-Leu-Val	
	PHP-28	5b	Val-Ile-Val-Val-Leu-Val	
	PHP-29	5c	Val-Val-aIle-Val-Leu-Val	
	PHP-30	5d	Val-Val-Val-Ile-Leu-Val	
	PHP-31	5e	Val-Val-Val-Val-Leu-Ile	
627.44	PHP-32	6	Val-Val-Val-Val-Leu-Val	√

^{a} PHP: Photohexapeptide. ^{b} aIle: *allo*-Isoleucine.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Notes and references

1 X. Cai, S. Nowak, F. Wesche, I. Bischoff, M. Kaiser, R. Fürst and H. B. Bode, *Nat. Chem.*, 2017, **9**, 379.
2 Y. Xu, R. Orozco, E. M. K. Wijeratne, A. A. L. Gunatilaka, S. P. Stock and I. Molnár, *Chem. Biol.*, 2008, **15**, 898.
3 F. C. Schröder, J. J. Farmer, A. B. Attygalle, S. R. Smedley, T. Eisner and J. Meinwald, *Science*, 1998, **281**, 428.
4 J.-K. Weng, R. N. Philippe and J. P. Noel, *Science*, 2012, **336**, 1667.
5 J.-K. Weng and J. P. Noel, *Cold Spring Harbor Symp. Quant. Biol.*, 2012, **77**, 309.
6 R. D. Firn and C. G. Jones, *Nat. Prod. Rep.*, 2003, **20**, 382.
7 H. D. Mootz, D. Schwarzer and M. A. Marahiel, *ChemBioChem*, 2002, **3**, 490.
8 S. Meyer, J.-C. Kehr, A. Mainz, D. Dehm, D. Petras, R. D. Süssmuth and E. Dittmann, *Cell Chem. Biol.*, 2016, **23**, 462.
9 M. Juguet, S. Lautru, F.-X. Francou, Š. Nezbedová, P. Leblond, M. Gondry and J.-L. Pernodet, *Chem. Biol.*, 2009, **16**, 421.
10 A. Thanwisai, S. Tandhavanant, N. Saiprom, N. R. Waterfield, P. K. Long, H. B. Bode, S. J. Peacock and N. Chantratita, *PLoS One*, 2012, **7**, e43835.
11 E. Bode, A. O. Brachmann, C. Kegler, R. Simsek, C. Dauth, Q. Zhou, M. Kaiser, P. Klemmt and H. B. Bode, *ChemBioChem*, 2015, **16**, 1115.
12 T. Stachelhaus, H. D. Mootz and M. A. Marahiel, *Chem. Biol.*, 1999, **6**, 493.
13 K. Blin, T. Wolf, M. G. Chevrette, X. Lu, C. J. Schwalen, S. A. Kautsar, H. G. Suarez Duran, E. L. C. de Los Santos, H. U. Kim, M. Nave, J. S. Dickschat, D. A. Mitchell, E. Shelest, R. Breitling, E. Takano, S. Y. Lee, T. Weber and M. H. Medema, *Nucleic Acids Res.*, 2017, **45**, W36.
14 T. Stachelhaus and C. T. Walsh, *Biochemistry*, 2000, **39**, 5775.
15 H. B. Bode, D. Reimer, S. W. Fuchs, F. Kirchner, C. Dauth, C. Kegler, W. Lorenzen, A. O. Brachmann and P. Grün, *Chem. – Eur. J.*, 2012, **18**, 2342.
16 C. Kegler, F. I. Nollmann, T. Ahrendt, F. Fleischhacker, E. Bode and H. B. Bode, *ChemBioChem*, 2014, **15**, 826.
17 J. N. deGruyter and W. A. Maio, *Org. Lett.*, 2014, **16**, 5196.
18 C. Hacker, X. Cai, C. Kegler, L. Zhao, A. K. Weickhmann, J. P. Wurm, H. B. Bode and J. Wöhnert, *Nat. Commun.*, 2018, **9**, 4366.
19 J. S. Morrow, W. R. Veatch and L. Stryer, *J. Mol. Biol.*, 1979, **132**, 733.
20 B. M. Burkhart, R. M. Gassman, D. A. Langs, W. A. Pangborn, W. L. Duax and V. Pletnev, *Pept. Sci.*, 1999, **51**, 129.
21 G. J. Williams, S. D. Breazeale, C. R. H. Raetz and J. H. Naismith, *J. Biol. Chem.*, 2005, **280**, 23000.
22 J. M. Crawford, C. Portmann, R. Kontnik, C. T. Walsh and J. Clardy, *Org. Lett.*, 2011, **13**, 5144.
23 C. M. Mudalungu, W. J. von Törne, K. Voigt, C. Rückert, S. Schmitz, O. N. Sekurova, S. B. Zotchev and R. D. Süssmuth, *J. Nat. Prod.*, 2019, **82**, 1478.
24 H. B. Bode, A. O. Brachmann, K. B. Jadhav, L. Seyfarth, C. Dauth, S. W. Fuchs, M. Kaiser, N. R. Waterfield, H. Sack, S. H. Heinemann and H. D. Arndt, *Angew. Chem., Int. Ed.*, 2015, **54**, 10352.
25 N. Kessler, H. Schuhmann, S. Morneweg, U. Linne and M. A. Marahiel, *J. Biol. Chem.*, 2004, **279**, 7413.
26 M. Gonsior, A. Mühlenweg, M. Tietzmann, S. Rausch, A. Poch and R. D. Süssmuth, *ChemBioChem*, 2015, **16**, 2610.
27 Y.-M. Shi and H. B. Bode, *Nat. Prod. Rep.*, 2018, **35**, 309.

7862 | *Org. Biomol. Chem.*, 2019, **17**, 7858–7862

This journal is © The Royal Society of Chemistry 2019

191

Supporting Information

Production of a photohexapeptide library from entomopathogenic *Photorhabdus asymbiotica* PB68.1

Lei Zhao[1,2] and Helge B. Bode*[1,3]

[1]Molecular Biotechnology, Department of Biosciences, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

[2]Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, 210014 Nanjing, China

[3]Buchmann Institute for Molecular Life Sciences (BMLS), Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

*Corresponding author

Experimental Procedures

Strain construction

Construction of promoter exchange mutant of *Photorhabdus asymbiotica* PB68.1 was carried out as described previously.[1] Briefly, pCEP (cluster expression plasmid, pLZ39) carrying the first 696 bp of *phpS* gene (*ppb6_04101*) from *P. asymbiotica* PB68.1 was constructed by using Gibson cloning and was transformed into *E. coli* S17-λpir. For the conjugation, *P. asymbiotica* PB68.1 was mated with *E. coli* S17-λpir carrying pLZ39. Both strains were grown in lysogeny broth (LB) medium with 50 μg/mL kanamycin added to *E. coli* S17-λpir. The cells were harvested and washed three times with fresh LB medium after OD_{600} of 0.5−0.7. Subsequently, *E. coli* S17-λpir and *P. asymbiotica* PB68.1 strains were mixed on a LB agar plate in a ratio of 1:3 and incubated at 37°C for 3 hours followed by incubation at 30°C overnight. The next day, the bacterial cell layer was harvested and resuspended in fresh LB medium. Serial dilutions were spread out on selective LB agar plates with kanamycin and ampicillin (100 μg/mL), and incubated at 30°C for 2−3 days. The genotype of individual clones was verified by PCR.

For construction of heterologous *E. coli* strains expressing *phpS*, functional plasmid (pLZ51) was achieved by using yeast cloning[2] and verified by enzyme digestion, and transformed into *E. coli* DH10B MtaA and *E. coli* DH10B MtaA Δ*penta*, respectively. Individual clones were analyzed by HPLC-MS for the photohexapeptide production.

Strain cultivation and culture extraction

100 μL of overnight culture of photohexapeptide production strains were inoculated into 10 mL liquid LB medium containing appropriate antibiotics, and 0.1% L-arabinose (from a 25% stock solution) and 2% Amberlite XAD-16 resin were added. The cultures were grown at 30 °C on a rotary shaker. After 72 h, the cultures were harvested, and XAD-16 beads were separated and extracted with 10 mL methanol for 1 h. Subsequently, the extracts were analyzed by HPLC-MS.

HPLC-MS analysis

HPLC-MS analysis was performed on a Dionex UltiMate 3000 system coupled to a Bruker Impact II QTOF mass spectrometer. Crude extracts and synthetic compounds in 5 μL injection volume were eluted on an ACQUITY UPLC BEH C_{18} column (130 Å, 2.1

S2

mm × 50 mm, 1.7 μm) using a gradient from 5% to 95% (or 5% to 60%) acetonitrile water solution containing 0.1% formic acid at a flow rate of 0.4 mL/min for 16 min. Positive mode with scan range from 100 to 1200 *m/z* was used to detect photohexapeptides.

Labeling experiments

The strain cultivation, culture extraction, and HPLC-MS analysis for labeling experiments were carried out as described above. The production culture was additionally fed with L-[D$_{10}$] leucine and L-[D$_8$] valine, respectively, in LB medium.

Compound isolation

Compound **1**-OMe was isolated from 1 L culture of promoter exchange mutant of *P. asymbiotica* PB68.1. The XAD extract was washed with EtOAc and fractionated by Sephadex LH-20 chromatography using MeOH as the eluent. The enriched fractions containing **1**-OMe were collected and used for further purification. The purification was carried out on an Agilent semi-preparative HPLC system. The compounds were eluted on a Cholester column (10 mm ID × 250 mm) by using an isocratic 35% acetonitrile water solution containing 0.1% formic acid for 40 min at a flow rate of 3 mL/min to yield **1**-OMe (2.5 mg).

Chemical synthesis

The selected photohexapeptides were synthesized by using a solid-phase peptide synthesis method as described previously.[3] For a schematic overview see Fig. S2 showing the synthesis of **1** as an example. Step **I** was loading of the isoleucine on the 2-chlorotrityl chloride (2-CTC) resin. A solution of Fmoc-Ile-OH (212 mg, 0.6 mmol, 3 eq.) and *N,N*-diisopropylethylamine (DIPEA, 0.31 mL, 1.8 mmol, 9 eq.) in 4 mL dry dichloromethane (DCM) was placed in a plastic reactor vessel filled with 2-CTC resin (125 mg, 0.2 mmol, 1.0 eq.). The resulting mixture was incubated at room temperature overnight. The remaining free binding sites were capped upon incubating twice with DCM/CH$_3$OH/DIPEA (80:15:5) for 10 min at room temperature. The resin was washed several times with dimethylformamide (DMF), CH$_3$OH, and DCM and treated with 20% piperidine in DMF (3 × 10 min) to remove the Fmoc-protecting group. The combined filtrates were used to determine the actual loading of the resin at $\lambda_{301\,nm}$. Afterwards, the resin was washed with DCM and dried. Step **II** is solid-phase peptide synthesis. The

linear sequence was synthesized on the preloaded Ile-2-CTC resin on a 25 μmol scale with a Syro Wave peptide synthesizer by using standard Fmoc chemistry. The resin was placed in a plastic reactor vessel with a Teflon frit and an amount of 6 eq. of amino acid derivatives (Fmoc-D-Leu-OH, Fmoc-Ile-OH, Fmoc-D-*allo*-Ile-OH, 0.2 M) were activated in situ at room temperature with 6 eq. of *O*-(6-Chlorobenzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium hexafluorophosphate (HCTU, 0.6 M) in DMF in the presence of 12 eq. DIPEA (2.4 M) in NMP for 50 min. Fmoc-protecting groups were removed with a solution of 40% piperidine in DMF for 5 min and the deprotection step was repeated for another 10 min with 20% piperidine in DMF. After each coupling and deprotection step, the resin was washed with NMP. After the addition of the final residue, the resin was washed with NMP, DMF, and DCM and dried. Step **III** is cleavage of peptide from the resin. A total of 1 mL 95% trifluoroacetic acid and 2.5% triisopropylsilane in water were added to the peptidyl resin (25 μmol) and the mixture was agitated for at least 2 h at room temperature. The resin was removed by filtration and washed twice with TFA. The solution was concentrated *in vacuo*. The peptide was purified by Agilent HPLC system to give **1**. The structure of **1** was confirmed by HR-MS and NMR.

Bioactivity testing

Bioactivity of **1** and **1**-OMe against *Micrococcus luteus* and *Saccharomyces cerevisiae* was tested using filter paper discs loaded with MeOH-dissolved compounds (1 mg/mL) on the prepared plates which were incubated at 30 °C for 2 days.

Quantitative analysis

The amount of all produced photohexapeptides in promoter exchange mutant of *P. asymbiotica* PB68.1 was deduced from quantitative analysis using the synthesized **1** as standard as described previously.[4]

S4

Supplementary Tables

Table S1. A domain specificity prediction of PhpS

A domain	Stachelhaus sequence	most likely amino acid predicted	amino acid detected
A1	DAFWIGATFK	Val	Val/Ile
A2	DAFWIGATFK	Val	Val/Ile
A3	DAFWIGATFK	Val	Val/Ile
A4	DAFWIGATFK	Val	Val/Ile
A5	DAWCIGAVCK	Phe	Leu
A6	DAFWIGATFK	Val	Val/Ile

Table S2. Bacterial strains used in this study

strain	genotype	reference
E. coli S17-λpir	Tp[r] Sm[r] *recA thi hsdR* RP4-2-Tc::Mu-Km::Tn7, λ*pir*	[5]
E. coli DH10B MtaA	F-*mcrA*, Δ(*mrr-hsd*RMS-*mcr*BC), Φ80*lacZ*ΔM15, Δ*lac*X74, *rec*A1, *end*A1, *ara*D139, Δ(*ara leu*)7697, *gal*U, *gal*K, *rps*L, *nup*G, λ–, *ent*D::*mtaA*	[2]
E. coli DH10B MtaA Δ*penta*	F-*mcrA*, Δ(*mrr-hsd*RMS-*mcr*BC), Φ80*lacZ*ΔM15, Δ*lac*X74, *rec*A1, *end*A1, *ara*D139, Δ(*ara leu*)7697, *gal*U, *gal*K, *rps*L, *nup*G, λ–, *mtaA*, Δ*penta* (Δ*aspC* Δ*ilvE* Δ*tyrB* Δ*avtA* Δ*ybfQ*), Cm[R]	[6]
P. asymbiotica PB68.1	wild type	[7]

Table S3. Plasmids used in this study

plasmid	genotype/description	reference
pCEP	pDS132 based vector, R6Kγ ori, oriT, *araC*, *ara*BAD promoter, Km[R]	[1]
pFF1	2μ ori, G418[R], PBAD promoter, pCOLA ori, MCS, Ypet-Flag, kan[R]	[8]
pLZ39	5489 bp, first 696 bp of *phpS* gene from PB68.1 genomic DNA assembled into pCEP, Km[R]	this work
pLZ51	27505 bp, *phpS* gene from PB68.1 genomic DNA assembled into pFF1, Km[R]	this work

Table S4. Primers used in this study

primer	sequence (5'-3')	targeting DNA fragment	plasmid
LZ_88	TTTGGGCTAACAGGAGGCTAGCAT ATGAAAGATAGCATTGCTAACGC	first 696 bp of *phpS* gene from PB68.1 (696 bp)	
LZ_89	TCTGCAGAGCTCGAGCATGCACATT TGCGAGAGATCTTGTGGATG		pLZ39
pCEP_fw_gib	ATGTGCATGCTCGAGCTC	pCEP vector backbone (4841 bp)	
pCEP_rv_gib	ATGCTAGCCTCCTGTTAGC		
V_pCEP_fw	GCTATGCCATAGCATTTTTATCCAT AAG	insert verification from pLZ39 (882 bp)	
V_pCEP_rv	ACATGTGGAATTGTGAGCGG		
V_pCEP_fw	GCTATGCCATAGCATTTTTATCCAT AAG	conjugation verification from PB68.1_pCEP-*phpS* (1342 bp)	
LZ_90	TCTGCAACAACATCTGCAATG		
LZ_126	ATCGCAACTCTCTACTGTTTCTCCA TACCCGTTTTTTTGGGCTAACAGGA GGAATTCCATGAAAGATAGCATTGC TAACGC	fragment I of *phpS* gene from PB68.1 (7314 bp)	
LZ_127	ATCAGTATGCTGACAGACCAGC		
LZ_128	AGTACCTGATGTGGAGAAACAAC	fragment II of *phpS* gene from PB68.1 (7240 bp)	pLZ51
LZ_129	TGAGATGGTCAACCCAGGTAG		
LZ_130	ACAGGTGGTAGAAATTGTCCAAC	fragment III of *phpS* gene from PB68.1 (6272 bp)	
LZ_131	TCTTCACCTTTGCTCATGAACTCGC CAGAACCAGCAGCGGAGCCAGCG GATCCGGCGCGCCTCACTGAACTT CTCCATCACC		

Table S5. HR-MS data of isolated and synthetic compounds

compound	sum formula	calcd. [M + H]$^+$	found [M + H]$^+$	Δppm
isolated **1**-OMe	$C_{37}H_{71}N_6O_7$	711.5379	711.5366	1.7
synthetic **1**	$C_{36}H_{69}N_6O_7$	697.5222	697.5205	2.4
synthetic **2a**	$C_{35}H_{67}N_6O_7$	683.5066	683.5050	2.3
synthetic **2b**	$C_{35}H_{67}N_6O_7$	683.5066	683.5056	1.4
synthetic **2c**	$C_{35}H_{67}N_6O_7$	683.5066	683.5056	1.5
synthetic **2d**	$C_{35}H_{67}N_6O_7$	683.5066	683.5055	1.5
synthetic **2e**	$C_{35}H_{67}N_6O_7$	683.5066	683.5054	1.7
synthetic **6**	$C_{31}H_{59}N_6O_7$	627.4440	627.4426	2.2

Supplementary Figures
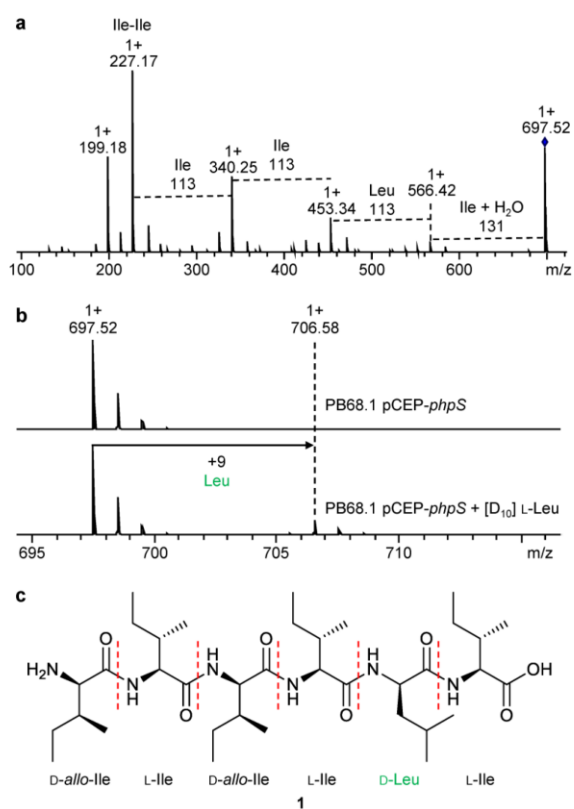
Fig. S1 Structure identification of **1** using labeling experiments combined with HPLC-MS analysis. (a) MS^2 fragmentation pattern of **1**; (b) MS data of **1** from labeling experiments with/without [D$_{10}$] L-Leu fed to *P. asymbiotica* PB68.1 pCEP-*phpS* in LB medium; (c) Structure of **1**.

S7
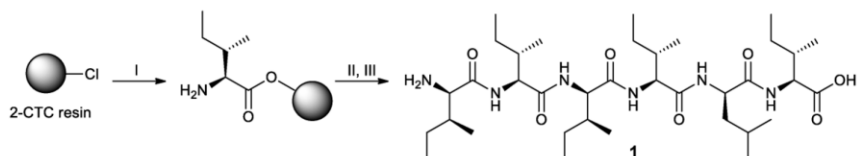
Fig. S2 Solid-phase peptide synthesis, shown for **1** as an example. I. Fmoc-Ile-OH, DIPEA, DCM, overnight, DCM/MeOH/DIPEA (80:15:5), then 20% piperidine/DMF; II. Fmoc-AA-OH, HCTU, DIPEA, DMF, NMP, 50 min, then 40% piperidine/DMF, 20% piperidine/DMF; III. TFA/TIS/Water (95:2.5:2.5), 2 h.
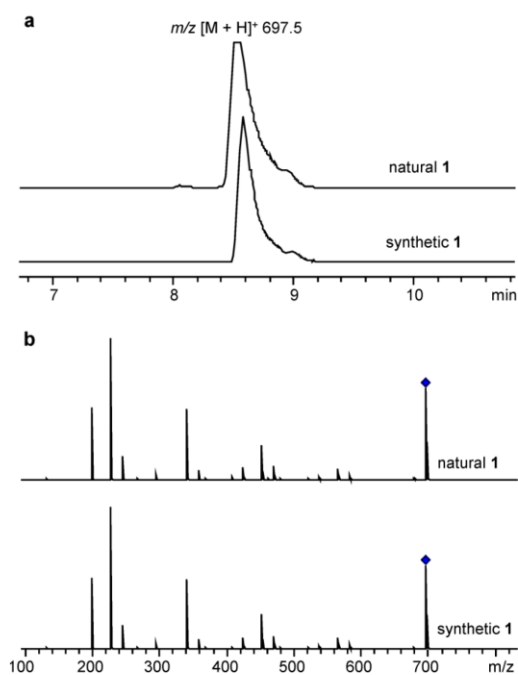
Fig. S3 Comparison of natural and synthetic 1 by HPLC-MS analysis. (a) Extracted ion chromatograms (EICs) of natural and synthetic **1**; (b) MS^2 fragmentation patterns of natural and synthetic **1**.
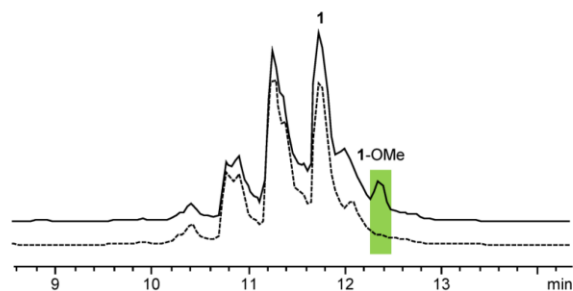
Fig. S4 HPLC-MS analysis of extracts from *P. asymbiotica* PB68.1 pCEP-*phpS*. XAD extract was treated with methanol (continuous line) and supernatant was treated with acetonitrile (1:1) (dotted line), respectively. Base-peak chromatograms (BPCs) are shown.

Cluster **2**: *m/z* [M + H]⁺ 683.50, 1 Leu + 1 Val + 4 Ile

Cluster **3**: *m/z* [M + H]⁺ 669.49, 1 Leu + 2 Val + 3 Ile

S10

Cluster **4**: *m/z* [M + H]⁺ 655.47, 1 Leu + 3 Val + 2 Ile
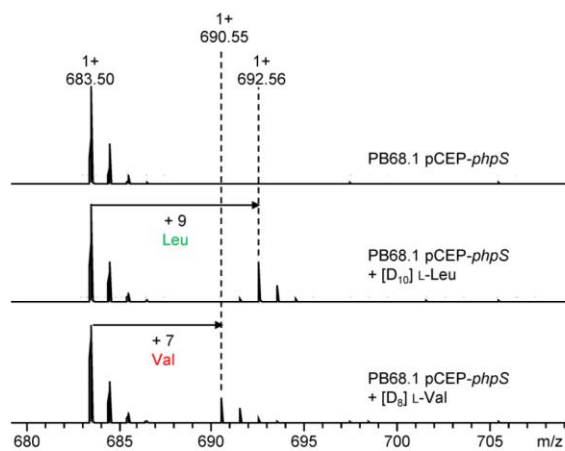
Cluster **5**: *m/z* [M + H]⁺ 641.46, 1 Leu + 4 Val + 1 Ile
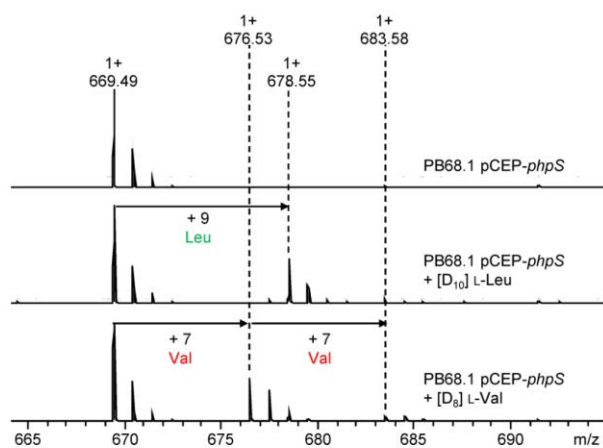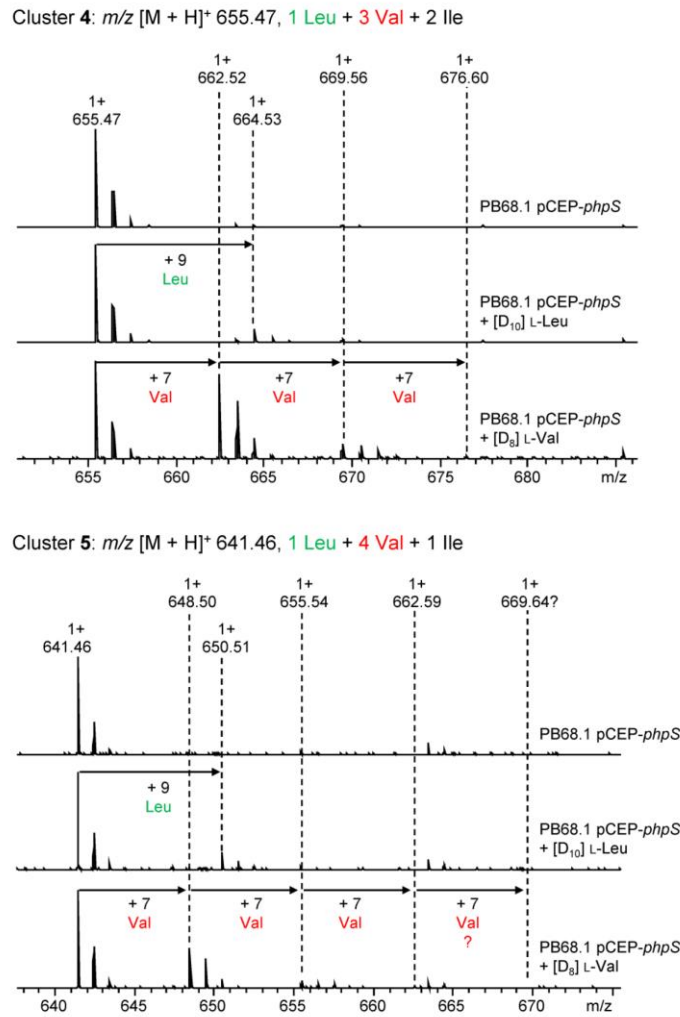
Fig. S5 MS data of compounds in clusters **2**−**5** from labeling experiments with/without [D₁₀] L-Leu or [D₈] L-Val fed to *P. asymbiotica* PB68.1 pCEP-*phpS* in LB medium.

S11
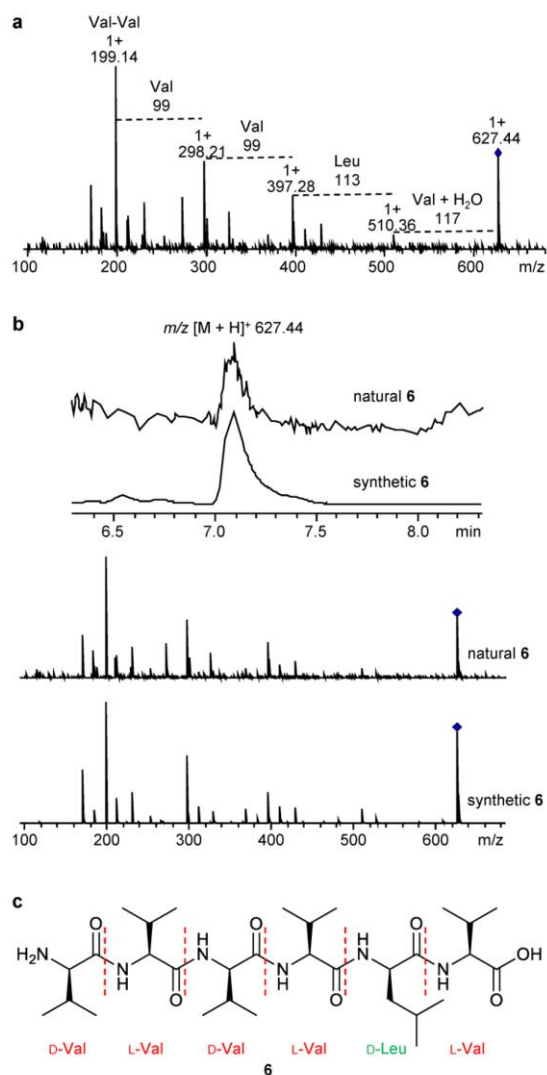
Fig. S6 Structure identification of **6** by HPLC-MS analysis and chemical synthesis. (a) MS^2 fragmentation pattern of **6**; (b) Comparison of natural and synthetic **6**; (c) Structure of **6**.

S12

Cluster **2**: *m/z* [M + H]⁺ 683.50, 1 Leu + 1 Val + 4 Ile

RT = 8.33 min

RT = 8.41 min

RT = 8.54 min

Val-Ile
1+
213.16

1+
683.50

1+
584.43

Ile
113

1+
326.24

Ile
113

Val (N-terminus)
99

1+
185.16

Leu
113

1+
439.32

Ile + H₂O
131

1+
552.41

D-Val   L-Ile   D-allo-Ile   L-Ile   D-Leu   L-Ile

2e

Cluster **3**: *m/z* [M + H]⁺ 669.49, 1 Leu + 2 Val + 3 Ile

RT = 8.08 min

Ile-Ile
1+
227.17

1+
185.16

Val
99

1+
326.24

Val
99

1+
669.49

1+
425.31

Leu
113

1+
538.39

Ile + H₂O
131

D-allo-Ile   L-Ile   D-Val   L-Val   D-Leu   L-Ile

3c

S14

205

RT = 8.18 min

Ile-Val
1+
213.16

1+
669.49

1+
556.40
Ile (N-terminus)
113

1+
185.16

Ile
113

1+
326.24

Val
99

1+
425.31

Leu
113

1+
538.39

Ile + H₂O
131

D-*allo*-Ile   L-Val   D-*allo*-Ile   L-Val   D-Leu   L-Ile

3e

RT = 8.23 min

Val-Ile
1+
213.16

1+
669.49

1+
570.42
Val (N-terminus)
99

1+
185.16

Val
99

1+
312.23

Ile
113

1+
425.31

Leu
113

1+
538.39

Ile + H₂O
131

D-Val   L-Ile   D-Val   L-Ile   D-Leu   L-Ile

3i

S15

RT = 8.25 min

Val-Ile
1+
213.16

Ile
113

1+
326.24

Val
99 425.31
1+

Leu
113 538.39
1+

1+
570.42

Val (N-terminus)
99

Ile + H₂O
131

1+
669.49

D-Val   L-Ile   D-*allo*-Ile   L-Val   D-Leu   L-Ile

3h

Cluster **4**: *m/z* [M + H]⁺ 655.47, 1 Leu + 3 Val + 2 Ile

RT = 7.85 min

Ile-Val
1+
213.16

1+
185.16

Val
99 312.22
1+

Ile
113 425.31
1+

Leu
113

1+
538.39

1+
542.39

Ile (N-terminus)
113

Val + H₂O
117

1+
655.47

D-*allo*-Ile   L-Val   D-Val   L-Ile   D-Leu   L-Val

4c

S16

RT = 7.89 min

Cluster **5**: *m/z* [M + H]⁺ 641.46, 1 Leu + 4 Val + 1 Ile

RT = 7.51 min

S17

RT = 7.57 min

Val-Ile
1+
213.16

Val 312.23
1+

99

Val 411.30
1+

99

Leu
113

1+
542.39
Val (N-terminus)
99

1+
524.38 Val + H₂O
117

1+
641.46

D-Val   L-Ile   D-Val   L-Val   D-Leu   L-Val

5b

RT = 7.59 min

Val-Val
1+
199.14

1+
171.15

Ile
113

1+
312.23

Val
99

1+
411.30

Leu
113

Val + H₂O
117

1+
524.38

1+
641.46

D-Val   L-Val   D-*allo*-Ile   L-Val   D-Leu   L-Val

5c

Fig. S7 Structure identification of thirteen compounds in cluster **2−5** by HPLC-MS analysis.

S19
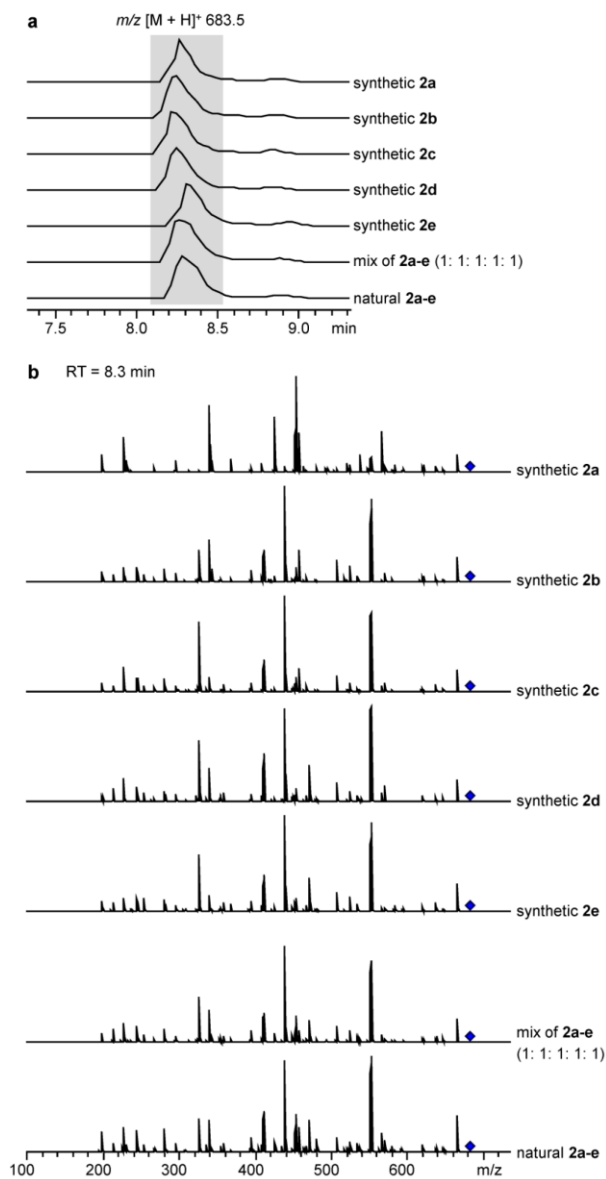
Fig. S8 Comparison of natural and synthetic **2a-e** by HPLC-MS analysis. (a) Extracted ion chromatograms (EICs) of natural and synthetic **2a-e**; (b) MS^2 fragmentation patterns of natural and synthetic **2a-e** at 8.3 min.

S20

211
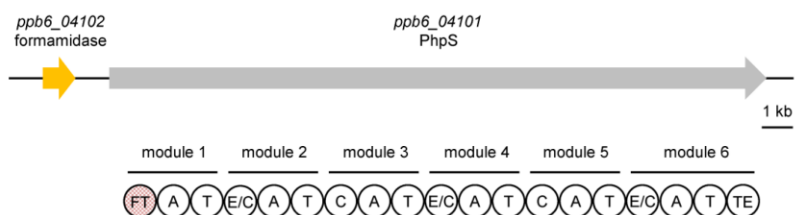
Fig. S9 Formamidase and PhpS NRPS encoding genes with domain organization. Domains: FT: formyl transferase, A: adenylation, T: thiolation, E/C: dual epimerization/condensation, C: condensation, TE: thioesterase.
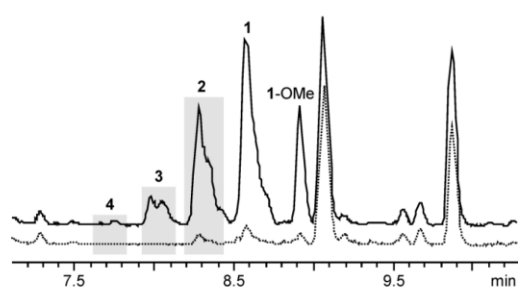
Fig. S10 HPLC-MS analysis of methanol extracts from *E. coli* expressing *phpS* with (continuous line) and without (dotted line) arabinose induction, respectively. Base-peak chromatograms (BPCs) are shown.

S21
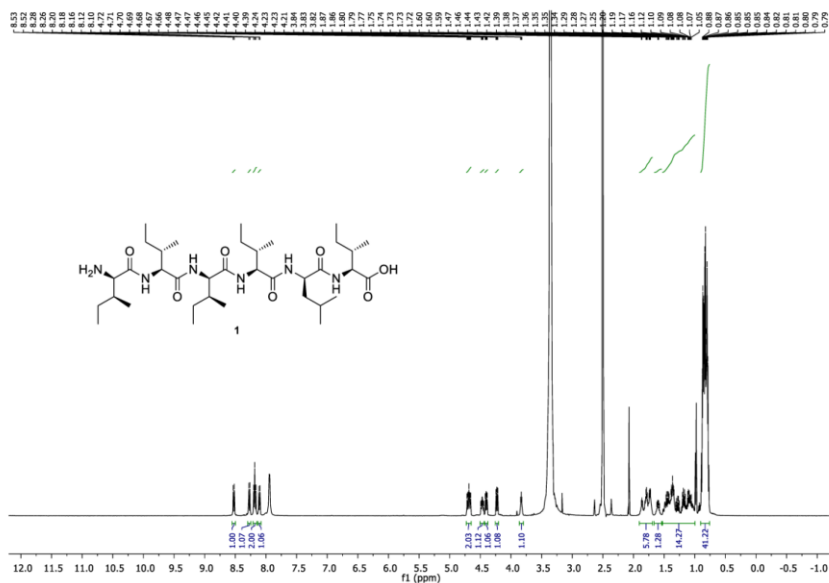
Fig. S11 ^{1}H NMR (500 MHz, DMSO-d_6) spectrum of synthetic **1**.
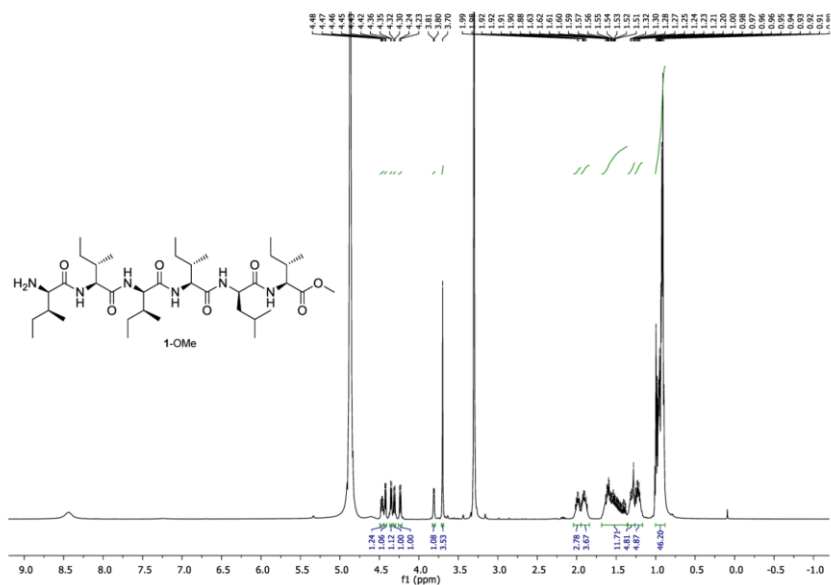
Fig. S12 ^{1}H NMR (500 MHz, methanol-d_4) spectrum of isolated **1**-OMe.
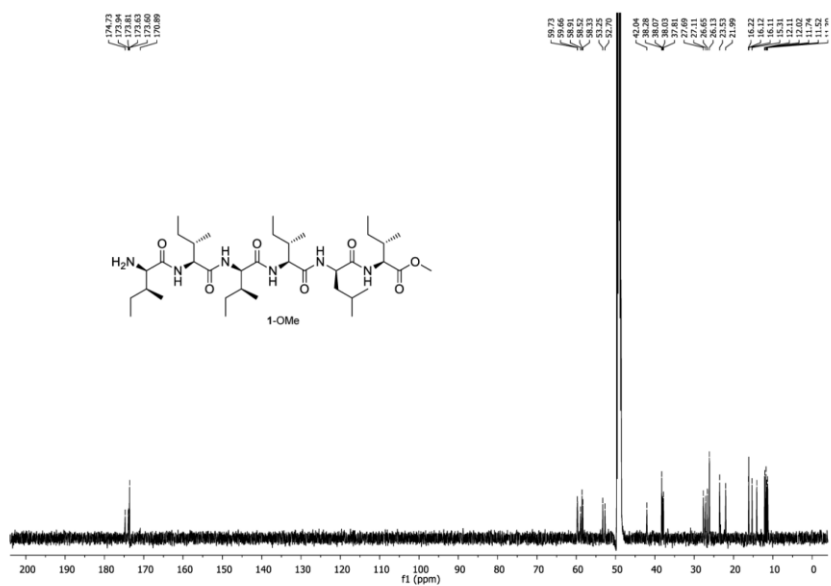
Fig. S13 ^{13}C NMR (125 MHz, methanol-d_4) spectrum of isolated **1**-OMe.

S23

214
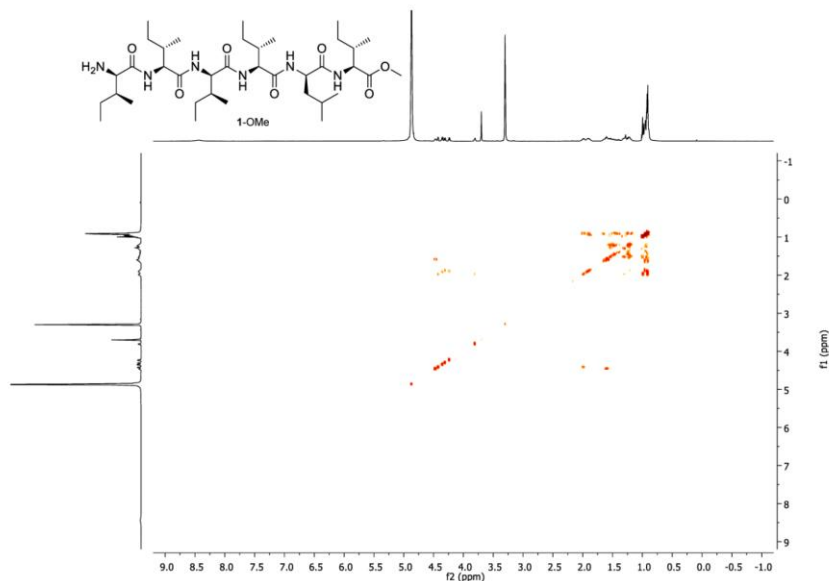
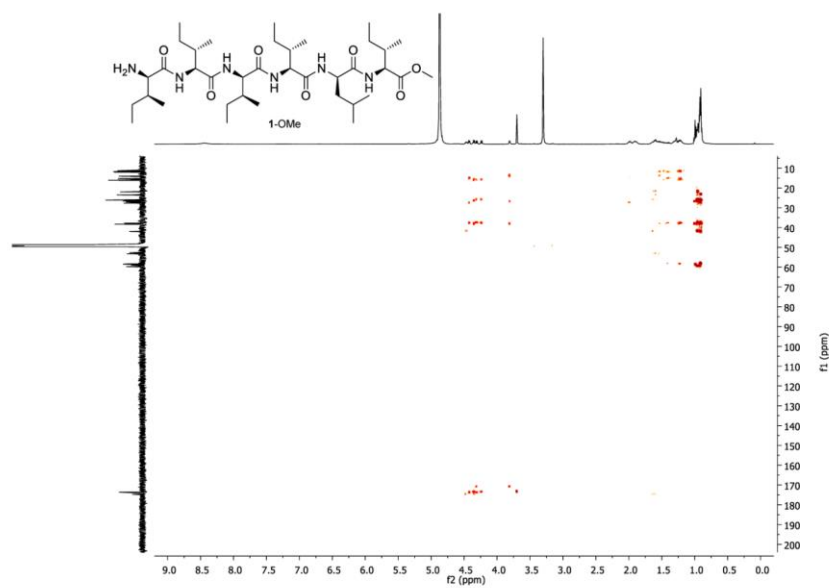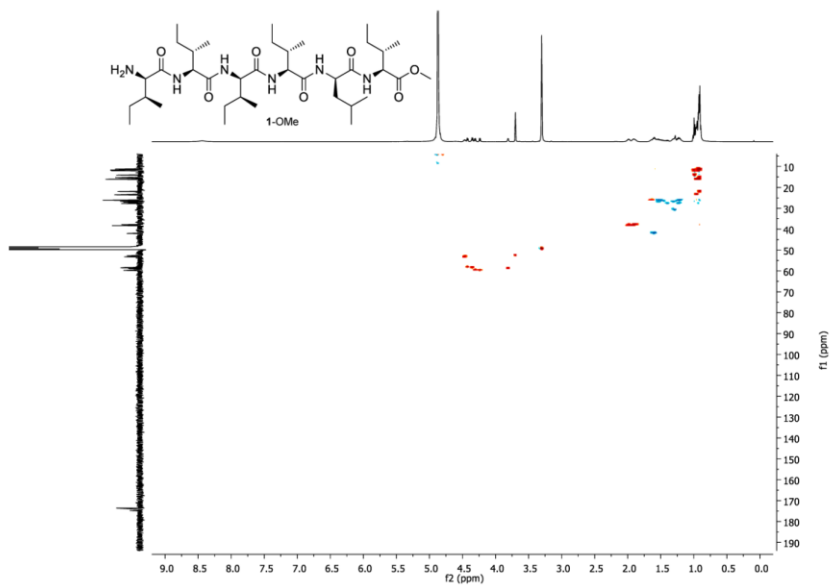**Fig. S14** COSY (methanol-d_4) spectrum of isolated **1**-OMe.

Fig. S15 HMBC (methanol-d_4) spectrum of isolated **1**-OMe.

Fig. S16 HSQC (methanol-d_4) spectrum of isolated **1**-OMe.

References

1 E. Bode, A. O. Brachmann, C. Kegler, R. Simsek, C. Dauth, Q. Zhou, M. Kaiser, P. Klemmt and H. B. Bode, *ChemBioChem*, 2015, **16**, 1115.

2 O. Schimming, F. Fleischhacker, F. I. Nollmann and H. B. Bode, *Chembiochem*, 2014, **15**, 1290.

3 C. Hacker, X. Cai, C. Kegler, L. Zhao, A. K. Weickhmann, J. P. Wurm, H. B. Bode and J. Wöhnert, *Nat. Commun.*, 2018, **9**, 4366.

4 X. Cai, V. L. Challinor, L. Zhao, D. Reimer, H. Adihou, P. Grün, M. Kaiser and H. B. Bode, *Org. Lett.*, 2017, **19**, 806.

5 S. R, P. U and Puhler A, *Nat. Biotechnol.*, 1983, **1**, 784.

6 C. Kegler, F. I. Nollmann, T. Ahrendt, F. Fleischhacker, E. Bode and H. B. Bode, *Chembiochem*, 2014, **15**, 826.

7 A. Thanwisai, S. Tandhavanant, N. Saiprom, N. R. Waterfield, P. K. Long, H. B. Bode, S. J. Peacock and N. Chantratita, *PLoS One*, 2012, **7**, e43835.

8 K. A. J. Bozhüyük, F. Fleischhacker, A. Linck, F. Wesche, A. Tietze, C.-P. Niesert and H. B. Bode, *Nat. Chem.*, 2018, **10**, 275.

5.4 Structure, Biosynthesis, and Bioactivity of Photoditritide from *Photorhabdus temperata* Meg1

Authors:

Lei Zhao,[1,2] Ryan Musumba Awori,[1] Marcel Kaiser,[3] Jonathan Groß,[4] Till Opatz,[4] and Helge B. Bode*,[1,5]

[1]Molecular Biotechnology, Department of Biosciences, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

[2]Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, 210014 Nanjing, China

[3]Swiss Tropical and Public Health Institute, 4051 Basel, Switzerland

[4]Institute of Organic Chemistry, Johannes Gutenberg University Mainz, 55128 Mainz, Germany

[5]Buchmann Institute for Molecular Life Sciences (BMLS), Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

*Corresponding author

Online access:

https://pubs.acs.org/doi/10.1021/acs.jnatprod.9b00932

Declaration on the contribution of the authors

Publication: Structure, Biosynthesis, and Bioactivity of Photoditritide from *Photorhabdus temperata* Meg1

Status: published

Journal: Journal of Natural Products

Authors: <u>Lei Zhao (LZ)</u>, Ryan Musumba Awori (RMA), Marcel Kaiser (MK), Jonathan Groß (JG), Till Opatz (TO), Helge B. Bode (HBB)

What did the doctoral candidate or the co-authors contribute individually to the dissertation?

(1) Development and planning

<u>LZ (60%)</u>, HBB (40%)

(2) Performance of individual research and experiments

<u>LZ (75%)</u>: strain construction, extract preparation, HPLC-MS analysis, labeling experiments, compound isolation, configuration determination, NMR measurement; RMA (10%): bioactivity testing; MK (10%): bioactivity testing; JG and TO (5%): measurement of optical rotation and ECD

(3) Collection of data and preparation of figures

<u>LZ (70%)</u>: collection of NMR and HPLC-MS data, preparation of figures; RMA (10%): collection of bioactivity data; MK (10%): collection of bioactivity data; JG and TO (10%): collection of optical rotation and ECD data, preparation of figure

(4) Analysis and interpretation of data

<u>LZ (100%)</u>: analysis of NMR and HPLC-MS data, interpretation of bioactivity data

(5) Preparation of manuscript

<u>LZ (80%)</u>, HBB (20%)

Herewith approving the indications above

date/place	signature doctoral candidate
date/place	signature supervisor
date/place	If necessary, signature corresponding author

JOURNAL OF NATURAL PRODUCTS

Note

Structure, Biosynthesis, and Bioactivity of Photoditritide from *Photorhabdus temperata* Meg1

Lei Zhao,[†,‡] Ryan Musumba Awori,[†] Marcel Kaiser,[§] Jonathan Groß,[⊥] Till Opatz,[⊥] and Helge B. Bode[*,†,‖]

[†]Molecular Biotechnology, Department of Biosciences, and [‖]Buchmann Institute for Molecular Life Sciences (BMLS), Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

[‡]Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, 210014 Nanjing, China

[§]Swiss Tropical and Public Health Institute, 4051 Basel, Switzerland

[⊥]Institute of Organic Chemistry, Johannes Gutenberg University Mainz, 55128 Mainz, Germany

S *Supporting Information*

ABSTRACT: A new cyclic peptide photoditritide (**1**), containing two rare amino acid D-homoarginine residues, was isolated from *Photorhabdus temperata* Meg1 after the nonribosomal peptide synthetase encoding gene *pdtS* was activated via promoter exchange. The structure of **1** was elucidated by HR-MS and NMR experiments. The absolute configurations of amino acids were determined according to the advanced Marfey's method after hydrolysis of **1**. Bioactivity testing of **1** revealed potent antimicrobial activity against *Micrococcus luteus* with an MIC value of 3.0 μM and weak antiprotozoal activity against *Trypanosoma brucei rhodesiense* with an IC$_{50}$ value of 13 μM. Additionally, the biosynthetic pathway of **1** was also proposed.
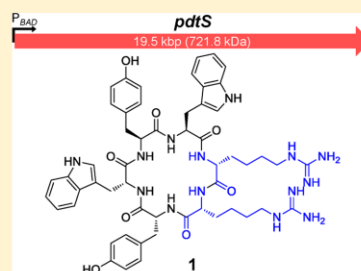
*P*hotorhabdus is a genus of Gram-negative entomopatho-genic bacteria that live in symbiosis with nematodes of the genus *Heterorhabditis*.[1,2] The nematode–bacterium complex is highly pathogenic for a broad range of insects and is well-known as a model system for the investigation of mutualistic and pathogenic symbiosis.[2–4] The *Heterorhabditis* nematode releases *Photorhabdus* bacteria from its gut into the insect hemocoel after insect infection.[2] The bacteria replicate and produce a large number of protein toxins and secondary metabolites to kill the insect and protect the prey cadaver for nutrition against soil-living food competitors, including bacteria, fungi, and protozoa.[5,6] The unique niche makes *Photorhabdus* a rich source of bioactive natural products,[6–8] and several compounds, such as isopropylstilbene,[9] phureali-pids,[10] kolossins,[11] photolose,[12] lumizinones,[13] rhabduscins,[14] and photopyrones[15] with antibiotic,[16] insecticidal,[10] antipro-tozoal,[11] immune regulation,[12] protease inhibitory,[13] virulence factor,[14] and signaling activity,[15] respectively, have been identified recently.

Natural products have played and will continue to play a highly significant role in the drug discovery and development process.[17] Our efforts to search for additional natural products from promising *Photorhabdus* strains resulted in the identi-fication of photoditritide (**1**), a new cyclohexapeptide containing two uncommon amino acid homoarginine residues, after its biosynthetic gene *pdtS* was activated in *Photorhabdus*

temperata Meg1. Here, we report the discovery, structure elucidation, biosynthesis, and bioactivity of **1**.
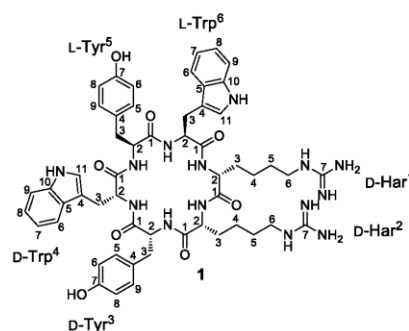
In the genome of *P. temperata* Meg1, in total 25 biosynthesis gene clusters of secondary metabolites have been predicted by antiSMASH.[18] The *pdtS* (photoditritide synthetase) gene (*MEG1_RS04325*, 19.5 kbp, accession number NZ_JGVH01000009) was identified encoding an unknown

3499

Note
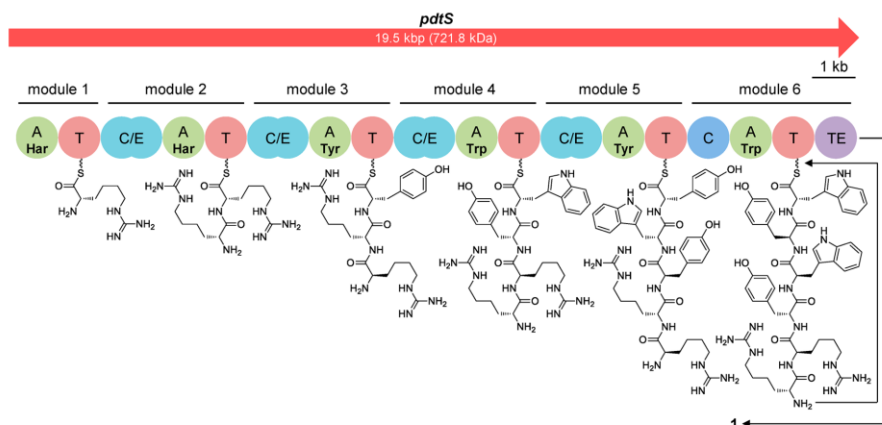
Figure 1. Domain organization of the PdtS and proposed biosynthesis of photoditritide (**1**). Domains: A: adenylation, T: thiolation, C/E: dual condensation/epimerization, C: condensation, TE: thioesterase.

Table 1. ^{1}H (500 MHz) and ^{13}C (125 MHz) NMR Data of 1 in DMSO-d_6 (δ in ppm)

subunit	position	δ_C, type	δ_H (J in Hz)	subunit	position	δ_C, type	δ_H (J in Hz)
Har^1	1	170.3, C			3b		2.91, overlap
	2	52.2, CH	4.23, dd (13.4, 6.5)		4	108.8, C	
	2-NH		7.23, d (7.0)		5	127.9, C	
	3a	32.1, CH$_2$	1.52, m		6	118.5, CH	7.05, overlap
	3b		1.45, m		7	118.1, CH	6.83, t (7.4)
	4	22.0, CH$_2$	1.12, m		8	120.6, CH	6.99, overlap
	5	28.1, CH$_2$	1.40, m		9	111.0, CH	7.26, d (8.1)
	6	40.4, CH$_2$	2.97, overlap		10	135.8, C	
	6-NH		8.45, brs		10-NH		10.79, s
	7	157.2, C			11	123.8, CH	6.48, s
Har^2	1	171.7, C		Tyr^5	1	171.2, C	
	2	53.0, CH	3.99, m		2	57.4, CH	3.83, m
	2-NH		8.36, submerge		2-NH		8.27, d (4.5)
	3	29.9, CH$_2$	1.30, m		3a	36.0, CH$_2$	2.43, m
	4a	22.0, CH$_2$	0.96, m		3b		2.35, m
	4b		0.71, m		4	127.8, C	
	5	28.1, CH$_2$	1.30, m		5	130.0, CH	6.79, d (8.2)
	6	40.2, CH$_2$	2.91, overlap		6	114.9, CH	6.53, d (8.2)
	6-NH		8.51, brs		7	155.9, C	
	7	157.2, C			8	114.9, CH	6.53, d (8.2)
Tyr^3	1	171.3, C			9	130.0, CH	6.79, d (8.2)
	2	56.3, CH	4.17, m	Trp^6	1	170.4, C	
	2-NH		8.70, d (7.6)		2	53.3, CH	4.45, dd (12.6, 9.5)
	3a	36.5, CH$_2$	3.09, brd (12.4)		2-NH		8.03, d (7.7)
	3b		2.61, t (12.7)		3a	25.9, CH$_2$	3.34, overlap
	4	128.4, C			3b		2.97, overlap
	5	130.0, CH	7.04, overlap		4	110.8, C	
	6	115.0, CH	6.65, d (8.1)		5	127.3, C	
	7	156.0, C			6	118.2, CH	7.49, d (7.8)
	8	115.0, CH	6.65, d (8.1)		7	118.3, CH	6.99, overlap
	9	130.0, CH	7.04, overlap		8	120.9, CH	7.07, overlap
Trp^4	1	170.9, C			9	111.3, CH	7.31, overlap
	2	54.4, CH	4.57, m		10	136.1, C	
	2-NH		7.31, overlap		10-NH		10.85, s
	3a	27.3, CH$_2$	3.32, overlap		11	123.2, CH	7.07, overlap

nonribosomal peptide synthetase consisting of six modules with overall 18 domains (Figure 1). PdtS was hence expected to produce a hexapeptide. As no such peptide could be identified in *P. temperata* Meg1 when the strain was cultivated in the lab in different media such as lysogeny broth (LB) or Sf-900 II SFM medium, a promoter exchange approach was applied to activate *pdtS* as previously described,[19] thus resulting in the detection of one new natural product based on HPLC-MS analysis (Figure S1).

To isolate **1**, the promoter exchange mutant *P. temperata* Meg1 pCEP-*pdtS* was grown in *Xenorhabdus* and *Photorhabdus* production medium (XPPM), an amino acid rich medium developed from M9 medium,[20] which gives a clear background for compound isolation from *Xenorhabdus* and *Photorhabdus* strains. From 4 L of cultures, 12 mg of **1** was isolated from the XAD-16 extracts by using Sephadex LH-20 chromatography, followed by semipreparative HPLC.

Compound **1** was obtained as a brown powder. The molecular formula of **1** was determined to be $C_{54}H_{66}N_{14}O_8$ by its HRESIMS data combined with isotopic labeling experiments as previously described (Figure S2),[21] indicating 29 degrees of unsaturation. Detailed analysis of 1D and 2D NMR data of **1** (Table 1, Figure 2) implied the presence of six amino
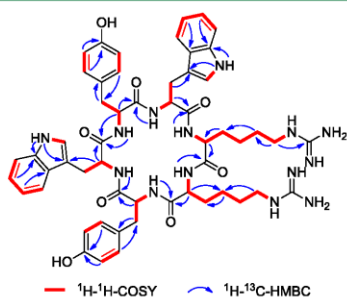

Figure 2. COSY and key HMBC correlations of **1**.

acids, two homoarginines (Har), two tyrosines (Tyr), and two tryptophans (Trp); and therefore, **1** was named as photoditritide. As nonproteinogenic amino acids like homoarginine are rarely found in natural products from entomopathogenic bacteria, a reverse labeling experiment with homoarginine was performed in ^{13}C medium verifying the incorporation of two homoarginines (Figure S3). Compound **1** was further identified to be a monocyclic peptide on the basis of the remaining 1 degrees of unsaturation because the six amino acids accounted for 28 of the 29 degrees of unsaturation. The connectivity of the six amino acid building blocks was established by HMBC (Figure 2) and was confirmed by the assembly order of predicted amino acids from the adenylation (A) domain specificity conferring codes (Table S1). Thereby, compound **1** was unequivocally elucidated to be a cyclic hexapeptide.

To address the absolute configurations of the amino acids in **1**, four dual condensation/epimerization (C/E) domains were identified in PdtS by antiSMASH (Figure 1).[22] C/E domains have been shown to transform the initially bound L-amino acids into their D-form,[21,23] suggesting **1** with the configuration sequence DDDDLL in its biosynthesis order. To confirm the configurations for homoarginine (D), tyrosine (D), tryptophan

(D), tyrosine (L), and tryptophan (L), compound **1** was hydrolyzed with 6 M HCl supplemented with 2.5% triisopropylsilane to prevent degradation of tryptophan.[24] The hydrolysate was analyzed according to the advanced Marfey's method,[25] and the finding was in accordance with the results predicted above (Figure S4).

Finally, the antimicrobial and antiprotozoal activity of **1** was investigated in view of the ecological function of *Photorhabdus* (Table 2). Compound **1** exhibited potent activity against

Table 2. Bioactivity of 1 against Different Microorganisms (MIC in μM), Protozoan Parasites (IC_{50} in μM), and Mammalian L6 Cells (IC_{50} in μM)

species	1	control[a]
M. luteus	3.0	4.2
Escherichia coli	24	0.60
Candida lusitaniae	>96	0.42
T. brucei rhodesiense	13	0.005
Trypanosoma cruzi	71	2.1
Leishmania donovani	>100	0.73
Plasmodium falciparum	27	0.009
mammalian L6 cells	83	0.007

[a]The positive control is different for each target organism: ampicillin for *M. luteus*, chloramphenicol for *E. coli*, amphotericin B for *C. lusitaniae*, melarsoprol for *T. brucei rhodesiense*, benznidazole for *T. cruzi*, miltefosine for *L. donovani*, chloroquine for *P. falciparum*, and podophyllotoxin for L6 cells.

Gram-positive bacterium *Micrococcus luteus* with an MIC value of 3.0 μM and weak activity against the causative agent of sleeping sickness *Trypanosoma brucei rhodesiense* with an IC_{50} value of 13 μM, but no activity against other tested organisms and no cytotoxic activity against mammalian L6 cells.

In summary, a new cyclic peptide photoditritide was unequivocally identified from *P. temperata* Meg1 by using a combination of bioinformatics analysis, promoter exchange, isotopic labeling experiments, HPLC-MS analysis, and NMR. Although nonproteinogenic amino acid homoarginine containing peptides have been reported in several marine organisms such as a sponge,[26] cyanobacteria,[27] and marine-derived actinomycetes;[28] photoditritide so far is the only example of a peptide from entomopathogenic bacteria that contains homoarginine.

■ EXPERIMENTAL SECTION

General Experimental Procedures. Optical rotation was measured on a PerkinElmer 241 polarimeter. ECD spectrum was obtained on a JASCO J-815 spectropolarimeter. ^{1}H, ^{13}C, COSY, HSQC, and HMBC NMR spectra were recorded on a Bruker AV 500 spectrometer at 500 MHz (^{1}H) and 125 MHz (^{13}C) using DMSO-d_6 as solvent; the ^{1}H and ^{13}C NMR chemical shifts were referenced to the solvent peaks at δ_H 2.50 and δ_C 39.52 for DMSO-d_6. HPLC-ESI-MS analysis was recorded on a Dionex UltiMate 3000 system coupled to a Bruker AmaZonX mass spectrometer; the MeOH extracts were eluted on an ACQUITY UPLC BEH C_{18} column (130 Å, 2.1 mm × 100 mm, 1.7 μm) using a gradient from 5% to 95% MeCN/H_2O solution containing 0.1% formic acid at a flow rate of 0.6 mL/min for 16 min. HPLC-HR-ESI-MS analysis was performed on a Dionex UltiMate 3000 system coupled to a Bruker Impact II QTOF mass spectrometer; the MeOH extracts were eluted on an ACQUITY UPLC BEH C_{18} column (130 Å, 2.1 mm × 50 mm, 1.7 μm) using a gradient from 5% to 95% MeCN/H_2O solution containing 0.1% formic acid at a flow rate of 0.4 mL/min for 16 min. Positive mode with scan range from 100 to 1200 m/z for HPLC-ESI-MS and HPLC-

HR-ESI-MS analysis was used to detect photoditritide (**1**). Semi-preparative HPLC was conducted on an Agilent 1260 system with a ZORBAX Eclipse XDB-C$_{18}$ column (9.4 mm × 250 mm, 5.0 μm).

Strain Construction. The construction of the promoter exchange mutant *P. temperata* Meg1 pCEP-*pdtS* was carried out as described previously.[19] Briefly, the plasmid pLZ46 carrying the first 650 bp of *pdtS* from *P. temperata* Meg1 was constructed by using Gibson cloning and was transformed into *E. coli* ST18. For conjugation, *P. temperata* Meg1 and *E. coli* ST18 were grown in LB medium. Kanamycin (50 μg/mL) and δ-aminolevulinic acid (ALA, 50 μg/mL) were added to *E. coli* ST18. After OD$_{600}$ 0.5–0.7, the cells were harvested and washed three times with fresh LB medium. Subsequently, *E. coli* ST18 and *P. temperata* Meg1 were mixed on a LB agar plate containing ALA in a ratio of 1:3 and were incubated at 37 °C for 3 h followed by growth at 30 °C overnight. The next day, the bacterial cell layer was harvested and resuspended in fresh LB medium. Serial dilutions were spread out on selective LB agar plates with kanamycin and were incubated at 30 °C for 2–3 days. The genotype of individual clones was verified by PCR.

Production Cultivation. 100 μL of overnight cultures of *P. temperata* Meg1 pCEP-*pdtS* were inoculated into 10 mL fresh liquid LB medium with kanamycin (50 μg/mL), 0.1% L-arabinose (from a 25% stock solution), and 2% Amberlite XAD-16 resin added. The cultures were grown at 30 °C on a rotary shaker at 200 rpm and were harvested after 3 days. The XAD-16 beads were filtered and extracted with 10 mL MeOH for 1 h to give extracts for HPLC-MS analysis.

Labeling Experiments. Labeling experiments were performed in 5 mL fully ^{13}C labeled medium as described previously.[21] Culture conditions, extract preparation, and HPLC-MS analysis were carried out as described above. To confirm the incorporation of homoarginine, L-homoarginine was supplied to *P. temperata* Meg1 pCEP-*pdtS* in ^{13}C medium.

Compound Isolation. To isolate **1**, the XAD-16 beads from 4 L XPPM cultures of *P. temperata* Meg1 pCEP-*pdtS* were extracted with MeOH. The XAD-16 extract was fractionated by Sephadex LH-20 chromatography using MeOH as the eluent. The enriched fractions containing **1** were collected and further purified by semipreparative HPLC with a gradient 20% to 40% MeCN/H$_2$O solution containing 0.1% formic acid at a flow rate of 3 mL/min for 20 min to yield **1** (12 mg).

Photoditritide (**1**): brown, amorphous solid; $[\alpha]^{25}_D$ − 1490.9 (*c* 0.022, MeOH); ECD (*c* 5.3 × 10^{-5} M, MeOH) λ_{max} ($\Delta\varepsilon$) 232 (35.25), 214 (7.42), and 202 (−11.60) nm; ^{1}H and ^{13}C NMR data, Table 1; HRESIMS *m/z* 520.2669 [M + 2H]$^{2+}$ (calcd for C$_{54}$H$_{66}$N$_{14}$O$_8$, 520.2667; Δppm 0.5).

Configuration Determination. The configurations of amino acids in **1** were determined by the advanced Marfey's method as described previously.[24,25] Approximately 1 mg of peptide was dissolved in 200 μL MeOH and was hydrolyzed with 800 μL HCl (6 M) in an Ace high-pressure tube at 110 °C for 1 h. To prevent the degradation of tryptophan, 2.5% triisopropylsilane was added. The hydrolysate was evaporated to dryness and was redissolved in 100 μL H$_2$O. To each half portion (45 μL) were added 10 μL 1 M NaHCO$_3$ and 80 μL 1% Nα-(5-fluoro-2,4-dinitrophenyl)-L-leucinamide or Nα-(5-fluoro-2,4-dinitrophenyl)-D-leucinamide (L-FDLA or D-FDLA, dissolved in acetone). The brown reaction vials were incubated in a water bath at 40 °C for 1 h. After that the reactions were cooled to room temperature, quenched with 10 μL 1 M HCl, and evaporated to dryness. The residue was dissolved in 400 μL MeOH and was analyzed by HPLC-MS.

Antimicrobial Activity Testing. Antimicrobial activity was tested against representative Gram-positive (*Micrococcus luteus*), Gram-negative (*Escherichia coli* OP50), and yeast (*Candida lusitaniae* DSM 70102) strains. For antibacterial activity, a microdilution broth assay modified from the Clinical and Laboratory Standards Institute (CLSI) protocol[29] was carried out in round-bottom 96-well plates with a final inoculum concentration of 1.3 × 10^5 CFU/mL, incubation time of 20 h at 30 °C, and final culture volume of 100 μL. A 2-fold dilution series was created for both **1** and the positive control ampicillin or chloramphenicol with a concentration range between

100 μg/mL and 0.02 μg/mL. RPMI 1640 medium supplemented with 10% LB medium was used instead of Muller Hinton broth to ensure the detection of antibacterial activity.[30] For antifungal activity, a microdilution broth assay modified from the European Committee for Antimicrobial Susceptibility Testing (EUCAST) protocol[31] was performed in round-bottom 96-well plates with a final inoculum concentration 2.8 × 10^4 CFU/mL, incubation period of 49 h at 30 °C in YPD medium, and final culture volume of 100 μL. A 2-fold dilution series was created for both **1** and the positive control amphotericin B with a concentration range between 100 μg/mL and 0.02 μg/mL. For both assays, minimum inhibitory concentration (MIC) was defined as the lowest antibiotic concentration that inhibits visible growth of the microorganisms.

Antiprotozoal Activity and Mammalian Cell Cytotoxicity Testing. Bioactivity testing of **1** against the parasites *Trypanosoma brucei rhodesiense* STIB900, *Trypanosoma cruzi* Tulahuen C4, *Leishmania donovani* MHOM-ET-67/L82, and *Plasmodium falciparum* NF54 was performed as described previously.[32] Cytotoxicity of **1** against rat skeletal myoblasts (L6 cells) was evaluated as described previously.[32] IC$_{50}$ (50% inhibitory concentration) values against these cells were calculated.

■ ASSOCIATED CONTENT

ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jnatprod.9b00932.

　　Adenylation domain specificity prediction of PdtS; bacterial strains, plasmids, and primers used; PdtS and its nearby proteins with predicted function; HPLC-MS analysis; ECD and NMR spectra for **1** (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: h.bode@bio.uni-frankfurt.de.

ORCID Ⓞ

Till Opatz: 0000-0002-3266-4050

Helge B. Bode: 0000-0001-6048-5909

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Thanwisai, A.; Tandhavanant, S.; Saiprom, N.; Waterfield, N. R.; Long, P. K.; Bode, H. B.; Peacock, S. J.; Chantratita, N. *PLoS One* **2012**, *7*, No. e43835.

(2) Forst, S.; Dowds, B.; Boemare, N.; Stackebrandt, E. *Annu. Rev. Microbiol.* **1997**, *51*, 47−72.

(3) Ehlers, R. U. *Commun. Agric. Appl. Biol. Sci.* **2003**, *68*, 3−16.

(4) Mulley, G.; Beeton, M. L.; Wilkinson, P.; Vlisidou, I.; Ockendon-Powell, N.; Hapeshi, A.; Tobias, N. J.; Nollmann, F. I.; Bode, H. B.; Van Den Elsen, J.; Ffrench-Constant, R. H.; Waterfield, N. R. *PLoS One* **2015**, *10*, No. e0144937.

(5) Ffrench-Constant, R. H.; Dowling, A.; Waterfield, N. R. *Toxicon* **2007**, *49*, 436−451.

(6) Shi, Y.-M.; Bode, H. B. *Nat. Prod. Rep.* **2018**, *35*, 309−335.

(7) Bode, H. B. *Curr. Opin. Chem. Biol.* **2009**, *13*, 224−230.

(8) Vizcaino, M. I.; Guo, X.; Crawford, J. M. *J. Ind. Microbiol. Biotechnol.* **2014**, *41*, 285−299.

(9) Kronenwerth, M.; Brachmann, A. O.; Kaiser, M.; Bode, H. B. *ChemBioChem* **2014**, *15*, 2689−2691.

(10) Nollmann, F. I.; Heinrich, A. K.; Brachmann, A. O.; Morisseau, C.; Mukherjee, K.; Casanova-Torres, A. M.; Strobl, F.; Kleinhans, D.; Kinski, S.; Schultz, K.; Beeton, M. L.; Kaiser, M.; Chu, Y.-Y.; Ke, L. P.; Thanwisai, A.; Bozhüyük, K. A. J.; Chantratita, N.; Götz, F.; Waterfield, N. R.; Vilcinskas, A.; Stelzer, E. H. K.; Goodrich-Blair, H.; Hammock, B. D.; Bode, H. B. *ChemBioChem* **2015**, *16*, 766−771.

(11) Bode, H. B.; Brachmann, A. O.; Jadhav, K. B.; Seyfarth, L.; Dauth, C.; Fuchs, S. W.; Kaiser, M.; Waterfield, N. R.; Sack, H.; Heinemann, S. H.; Arndt, H.-D. *Angew. Chem., Int. Ed.* **2015**, *54*, 10352−10355.

(12) Perez, C. E.; Crawford, J. M. *Biochemistry* **2019**, *58*, 1131−1140.

(13) Park, H. B.; Crawford, J. M. *J. Antibiot.* **2016**, *69*, 616−621.

(14) Crawford, J. M.; Portmann, C.; Zhang, X.; Roeffaers, M. B. J.; Clardy, J. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 10821−10826.

(15) Brachmann, A. O.; Brameyer, S.; Kresovic, D.; Hitkova, I.; Kopp, Y.; Manske, C.; Schubert, K.; Bode, H. B.; Heermann, R. *Nat. Chem. Biol.* **2013**, *9*, 573−578.

(16) Eleftherianos, I.; Boundy, S.; Joyce, S. A.; Aslam, S.; Marshall, J. W.; Cox, R. J.; Simpson, T. J.; Clarke, D. J.; Ffrench-Constant, R. H.; Reynolds, S. E. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 2419−2424.

(17) Newman, D. J.; Cragg, G. M. *J. Nat. Prod.* **2016**, *79*, 629−661.

(18) Blin, K.; Wolf, T.; Chevrette, M. G.; Lu, X.; Schwalen, C. J.; Kautsar, S. A.; Suarez Duran, H. G.; de Los Santos, E. L. C.; Kim, H. U.; Nave, M.; Dickschat, J. S.; Mitchell, D. A.; Shelest, E.; Breitling, R.; Takano, E.; Lee, S. Y.; Weber, T.; Medema, M. H. *Nucleic Acids Res.* **2017**, *45*, W36−W41.

(19) Bode, E.; Brachmann, A. O.; Kegler, C.; Simsek, R.; Dauth, C.; Zhou, Q.; Kaiser, M.; Klemmt, P.; Bode, H. B. *ChemBioChem* **2015**, *16*, 1115−1119.

(20) Bode, E. et al.; *Angew. Chem., Int. Ed.* **2019**, DOI: 10.1002/anie.201910563.

(21) Bode, H. B.; Reimer, D.; Fuchs, S. W.; Kirchner, F.; Dauth, C.; Kegler, C.; Lorenzen, W.; Brachmann, A. O.; Grün, P. *Chem. - Eur. J.* **2012**, *18*, 2342−2348.

(22) Blin, K.; Shaw, S.; Steinke, K.; Villebro, R.; Ziemert, N.; Lee, S. Y.; Medema, M. H.; Weber, T. *Nucleic Acids Res.* **2019**, *47*, W81−W87.

(23) Stachelhaus, T.; Walsh, C. T. *Biochemistry* **2000**, *39*, 5775−5787.

(24) Grundmann, F.; Kaiser, M.; Kurz, M.; Schiell, M.; Batzer, A.; Bode, H. B. *RSC Adv.* **2013**, *3*, 22072−22077.

(25) Fujii, K.; Ikai, Y.; Oka, H.; Suzuki, M.; Harada, K. *Anal. Chem.* **1997**, *69*, 5146−5151.

(26) Bonnington, L. S.; Tanaka, J.; Higa, T.; Kimura, J.; Yoshimura, Y.; Nakao, Y.; Yoshida, W. Y.; Scheuer, P. J. *J. Org. Chem.* **1997**, *62*, 7765−7767.

(27) Saito, K.; Konno, A.; Ishii, H.; Saito, H.; Nishida, F.; Abe, T.; Chen, C. *J. Nat. Prod.* **2001**, *64*, 139−141.

(28) Cha, J. W.; Park, J.-S.; Sim, T.; Nam, S.-J.; Kwon, H. C.; Del Valle, J. R.; Fenical, W. *J. Nat. Prod.* **2012**, *75*, 1648−1651.

(29) Patel, J. B.; Cockerill, F. R., III; Bradford, P. A.; Eliopoulos, G. M.; Hindler, J. A.; Jenkins, S. G.; Lewis, J. S., II; Limbago, B.; Miller, L. A.; Nicolau, D. P.; Powell, M.; Swenson, J. M.; Traczewski, M. M.; Turmidge, J. D.; Weinstein, M. P.; Zimmer, B. L. *Methods for Dilution Antimicrobial Susceptibility Tests for Bacteria That Grow Aerobically; Approved Standard M07-A10*; Clinical and Laboratory Standards Institute: Wayne, PA, USA, 2015; Vol 35, No 2.

(30) Kumaraswamy, M.; Lin, L.; Olson, J.; Sun, C.-F.; Nonejuie, P.; Corriden, R.; Döhrmann, S.; Ali, S. R.; Amaro, D.; Rohde, M.; Pogliano, J.; Sakoulas, G.; Nizet, V. *J. Antimicrob. Chemother.* **2016**, *71*, 1264−1269.

(31) Rodriguez-Tudela, J. L.; Arendrup, M. C.; Barchiesi, F.; Bille, J.; Chryssanthou, E.; Cuenca-Estrella, M.; Dannaoui, E.; Denning, D. W.; Donnelly, J. P.; Dromer, F.; Fegeler, W.; Lass-Florl, C.; Moore, C.; Richardson, M.; Sandven, P.; Velegraki, A.; Verweij, P. *Clin. Microbiol. Infect.* **2008**, *14*, 398−405.

(32) Orhan, I.; Şener, B.; Kaiser, M.; Brun, R.; Tasdemir, D. *Mar. Drugs* **2010**, *8*, 47−58.

Supporting Information

Structure, Biosynthesis, and Bioactivity of Photoditritide from *Photorhabdus temperata* Meg1

Lei Zhao,[1,2] Ryan Musumba Awori,[1] Marcel Kaiser,[3] Jonathan Groß,[4] Till Opatz,[4] and Helge B. Bode*[1,5]

[1]Molecular Biotechnology, Department of Biosciences, and [5]Buchmann Institute for Molecular Life Sciences (BMLS), Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

[2]Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, 210014 Nanjing, China

[3]Swiss Tropical and Public Health Institute, 4051 Basel, Switzerland

[4]Institute of Organic Chemistry, Johannes Gutenberg University Mainz, 55128 Mainz, Germany

*Corresponding author

Supplementary Tables

Table S1. A domain specificity prediction of PdtS

A domain	Stachelhaus sequence	most likely amino acid predicted	amino acid detected
A1	DVESIGGVTK	Lys	Har
A2	DVESIGGVTK	Lys	Har
A3	DASFIADVCK	Tyr	Tyr
A4	DVQCIGDVCK	Tyr	Trp
A5	DASFIADVCK	Tyr	Tyr
A6	DASLVGDVCK	Tyr	Trp

Table S2. Bacterial strains used in this study

strain	genotype	reference
E. coli ST18	*E. coli* S17 λpir Δ*hemA*	1
P. temperata Meg1	wild type	2

Table S3. Plasmids used in this study

plasmid	genotype/description	reference
pCEP	pDS132 based vector, R6Kγ ori, oriT, *ara*C, *ara*BAD promoter, Km^R	3
pLZ46	5443 bp, first 650 bp of *pdtS* gene from Meg1 genomic DNA assembled into pCEP, Km^R	this work

Table S4. Primers used in this study

primer	sequence (5'-3')	targeting DNA fragment	plasmid
LZ_112	TTTGGGCTAACAGGAGGCTAGCATATGA AAGATAGCATTGCTGAAGC	first 650 bp of *pdtS* from Meg1 (650 bp)	
LZ_113	TCTGCAGAGCTCGAGCATGCACATCAC GGTGTGTCATCTTGTATCA		pLZ46
pCEP_fw_gib	ATGTGCATGCTCGAGCTC	pCEP vector backbone (4841 bp)	
pCEP_rv_gib	ATGCTAGCCTCCTGTTAGC		
V_pCEP_fw	GCTATGCCATAGCATTTTTATCCATAAG	insert verification from pLZ46 (836 bp)	
V_pCEP_rv	ACATGTGGAATTGTGAGCGG		
V_pCEP_fw	GCTATGCCATAGCATTTTTATCCATAAG	conjugation verification from Meg1 pCEP-*pdtS* (1050 bp)	
LZ_114	ACTGAGCATCTGCTCTTGATAC		

S2

227

Table S5. PdtS and its nearby proteins with predicted function
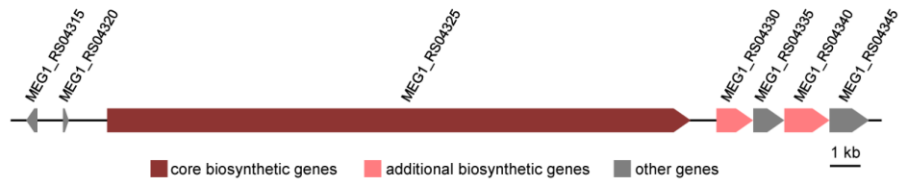
| locus tag	predicted function
MEG1_RS04315	hypothetical protein
MEG1_RS04320	hypothetical protein
MEG1_RS04325 (PdtS)	non-ribosomal peptide synthase
MEG1_RS04330	aspartate aminotransferase family protein
MEG1_RS04335	arginine *N*-succinyltransferase
MEG1_RS04340	succinylglutamate-semialdehyde dehydrogenase
MEG1_RS04345	*N*-succinylarginine dihydrolase

S3

Supplementary Figures
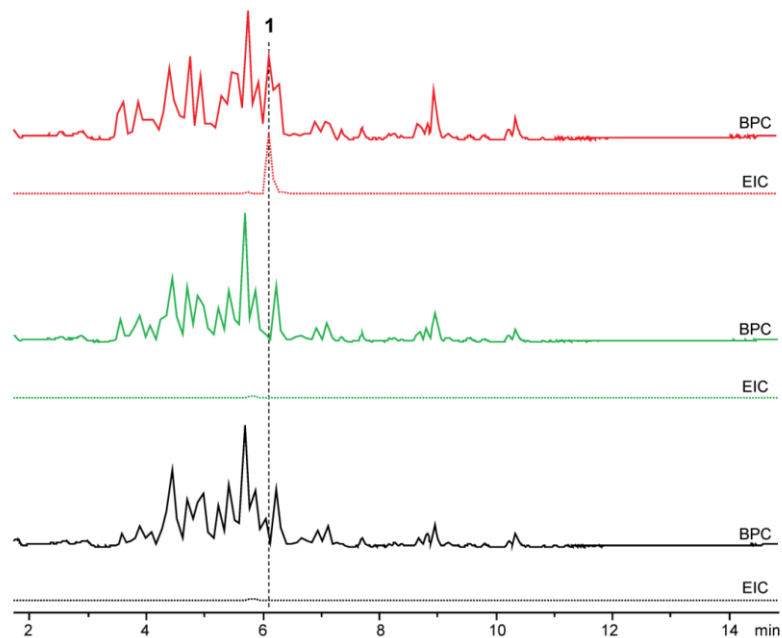
Figure S1. HPLC-MS analysis of XAD extracts from *P. temperata* Meg1 wild type (black line) and *P. temperata* Meg1 pCEP-*pdtS* without (green line) and with (red line) L-arabinose induction in LB medium. Base-peak chromatograms (BPCs) and extracted ion chromatograms (EICs) for **1** are shown, respectively.
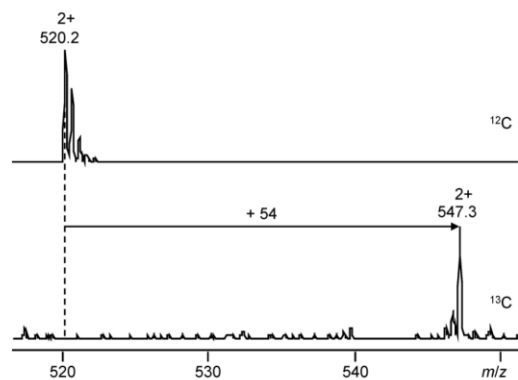
Figure S2. Determination of the number of carbon atoms for **1** by growing *P. temperata* Meg1 pCEP-*pdtS* in standard ^{12}C medium and fully ^{13}C labeled medium, respectively.
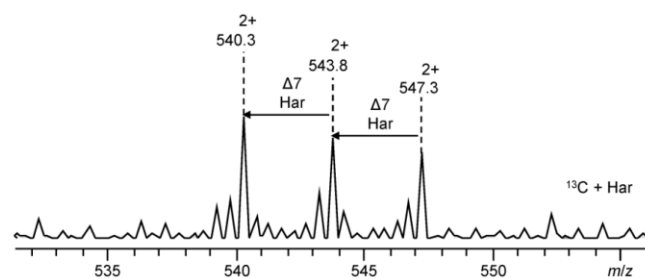
Figure S3. HPLC-MS analysis of **1** from reverse labeling experiment with L-homoarginine fed to *P. temperata* Meg1 pCEP-*pdtS* in fully ^{13}C labeled medium.
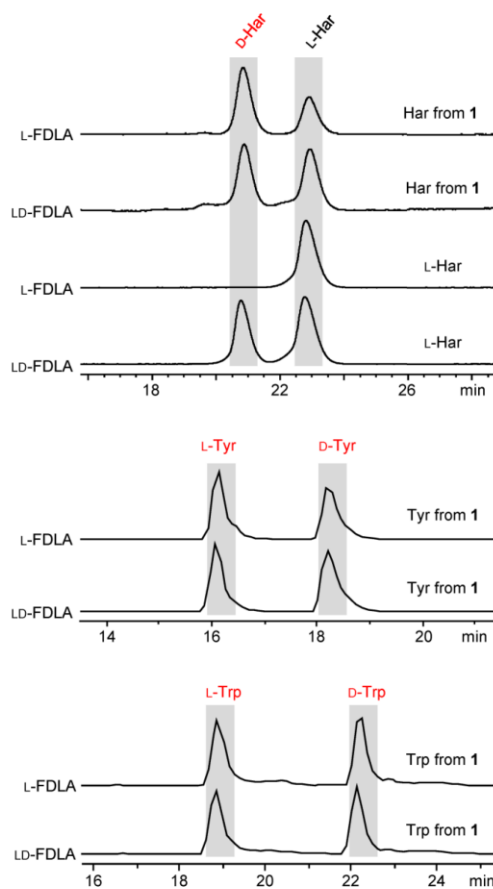
S5

Figure S4. Configuration determination of amino acids in **1** using the advanced Marfey's method. HPLC-MS analysis of hydrolyzed **1** and derivatized with L-FDLA and LD-FDLA. Depicted are extracted ion chromatogram (EIC) traces for homoarginine (Har, m/z 483 [M + H]$^+$), tyrosine (Tyr, m/z 476 [M + H]$^+$), and tryptophan (Trp, m/z 499 [M + H]$^+$). The configurations are determined by the elution order: L-FDLA derivatized tyrosine and tryptophan elute prior to its D-enantiomer, while D-FDLA derivatized homoarginine elutes prior to its L-enantiomer. The small amount of L-Har might be from racemization of D-Har during acid hydrolysis of **1**.

S6
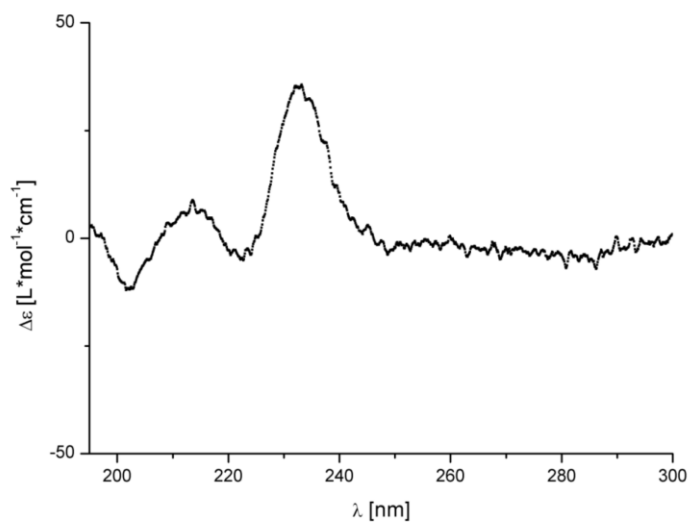
Figure S5. ECD spectrum of **1**.
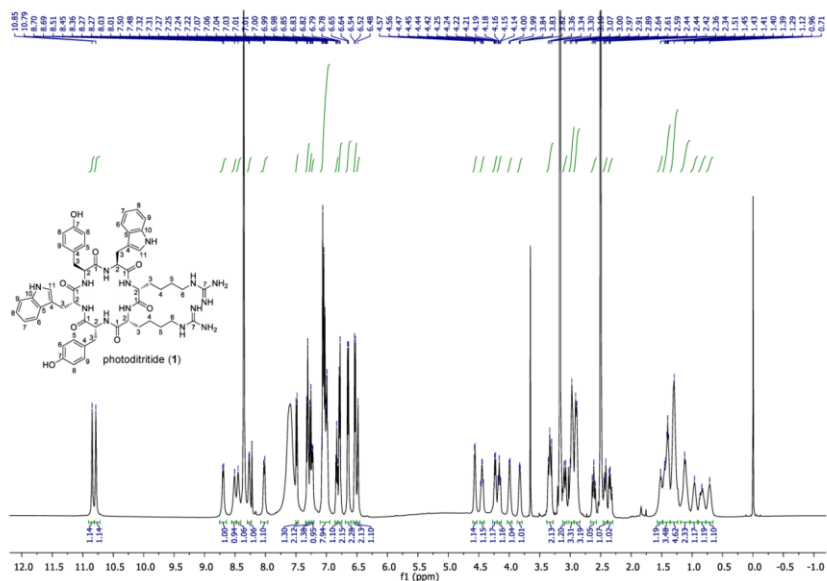
Figure S6. ^{1}H NMR (500 MHz, DMSO-d_6) spectrum of **1**.
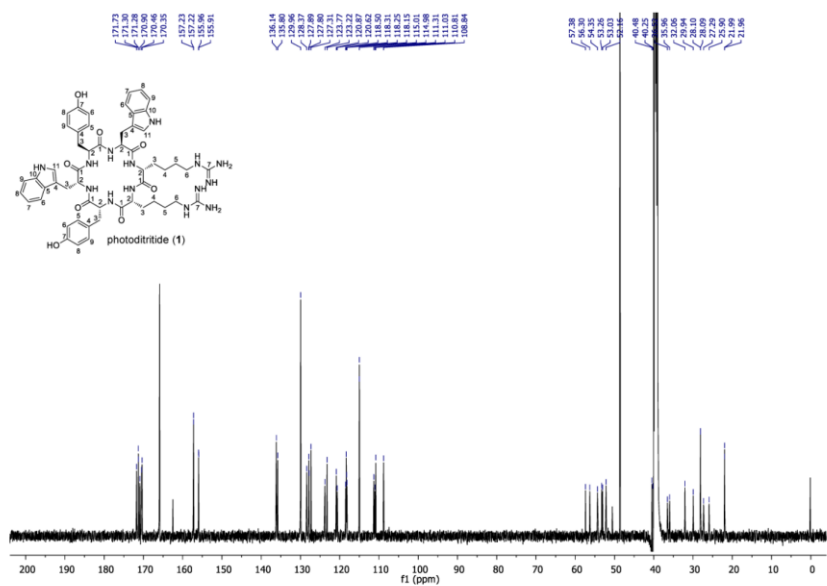
Figure S7. ^{13}C NMR (125 MHz, DMSO-d_6) spectrum of **1**.
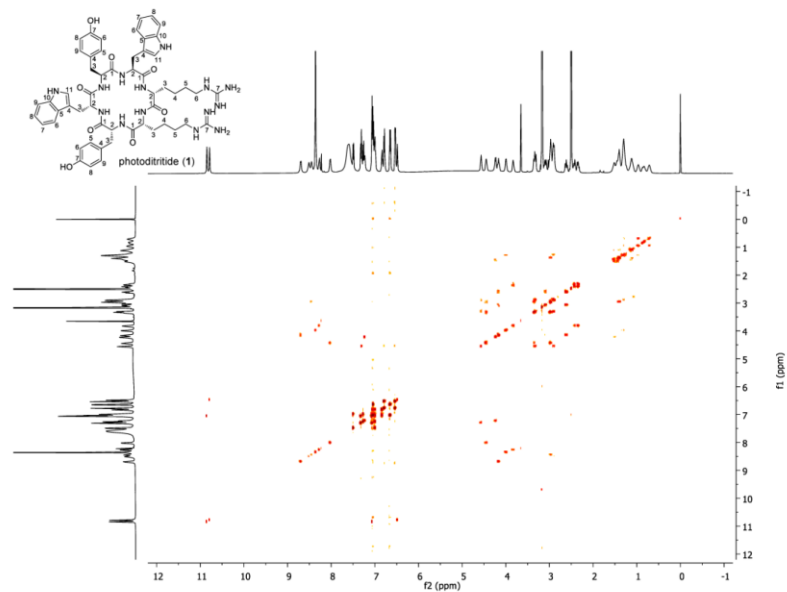
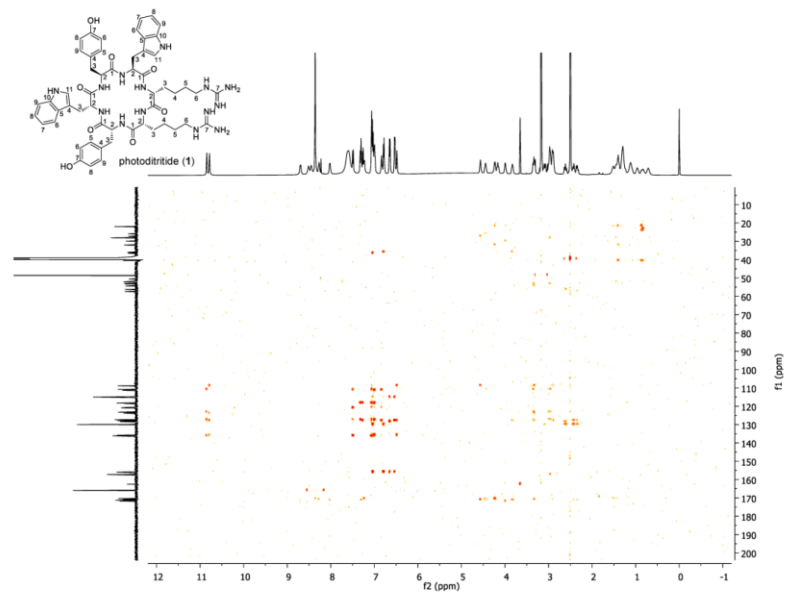**Figure S8.** COSY (DMSO-*d*6) spectrum of **1**.

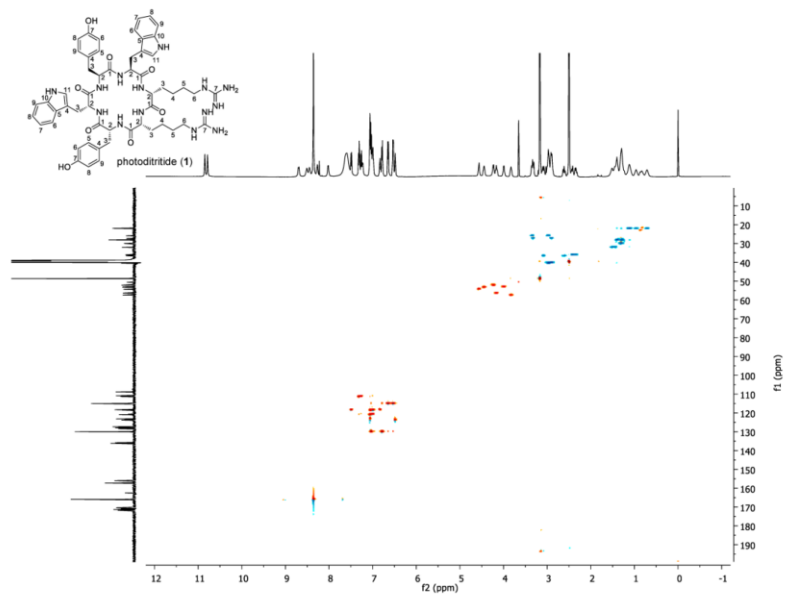**Figure S9.** HMBC (DMSO-*d*6) spectrum of **1**.

S9

Figure S10. HSQC (DMSO-d_6) spectrum of **1**.

References

(1) Thoma, S.; Schobert, M. *FEMS Microbiol. Lett.* **2009**, *294*, 127–132.

(2) Hurst, S. G.; Ghazal, S.; Morris, K.; Abebe-Akele, F.; Thomas, W. K.; Badr, U. M.; Hussein, M. A.; AbouZaied, M. A.; Khalil, K. M.; Tisa, L. S. *Genome Announc.* **2014**, *2*, e01273–14.

(3) Bode, E.; Brachmann, A. O.; Kegler, C.; Simsek, R.; Dauth, C.; Zhou, Q.; Kaiser, M.; Klemmt, P.; Bode, H. B. *ChemBioChem* **2015**, *16*, 1115–1119.

S11

5.5 Phototemtide A, a Cyclic Lipopeptide Heterologously Expressed from *Photorhabdus temperata* Meg1, Shows Selective Antiprotozoal Activity

Authors:

Lei Zhao,[1,2] Tien Duy Vo,[1] Marcel Kaiser,[3] and Helge B. Bode*[1,4]

[1]Molecular Biotechnology, Department of Biosciences, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

[2]Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, 210014 Nanjing, China

[3]Swiss Tropical and Public Health Institute, 4002 Basel, Switzerland

[4]Buchmann Institute for Molecular Life Sciences (BMLS), Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

*Corresponding author

Declaration on the contribution of the authors

Publication: Phototemtide A, a Cyclic Lipopeptide Heterologously Expressed from *Photorhabdus temperata* Meg1, Shows Selective Antiprotozoal Activity

Status: published

Journal: ChemBioChem

Authors: <u>Lei Zhao (LZ)</u>, Tien Duy Vo (TDV), Marcel Kaiser (MK), Helge B. Bode (HBB)

What did the doctoral candidate or the co-authors contribute individually to the dissertation?

(1) Development and planning

<u>LZ (60%)</u>, HBB (40%)

(2) Performance of individual research and experiments

<u>LZ (70%): strain construction, extract preparation, HPLC-MS analysis, compound isolation, NMR measurement, configuration determination, chemical synthesis, bioactivity testing;</u> TDV (20%): chemical synthesis, HPLC-MS analysis, NMR measurement; MK (10%): bioactivity testing

(3) Collection of data and preparation of figures

<u>LZ (70%): collection of NMR and HPLC-MS data, preparation of figures;</u> TDV (20%): collection of NMR data, preparation of figure; MK (10%): collection of bioactivity data

(4) Analysis and interpretation of data

<u>LZ (80%): analysis of NMR and HPLC-MS data, interpretation of bioactivity data;</u> TDV (20%): analysis of NMR data, interpretation of bioactivity data

(5) Preparation of manuscript

<u>LZ (60%)</u>, TDV (20%), HBB (20%)

Herewith approving the indications above

_____   _____
date/place   signature doctoral candidate

_____   _____
date/place   signature supervisor

_____   _____
date/place   If necessary, signature corresponding author

Phototemtide A, a Cyclic Lipopeptide Heterologously Expressed from *Photorhabdus temperata* Meg1, Shows Selective Antiprotozoal Activity

Lei Zhao,[a, b] Tien Duy Vo,[a] Marcel Kaiser,[c] and Helge B. Bode*[a, d]

A new cyclic lipopeptide, phototemtide A (**1**), was isolated from *Escherichia coli* expressing the biosynthetic gene cluster *pttABC* from *Photorhabdus temperata* Meg1. The structure of **1** was elucidated by HR-ESI-MS and NMR experiments. The absolute configurations of amino acids and 3-hydroxyoctanoic acid in **1** were determined by using the advanced Marfey's method and comparison after total synthesis of **1**, respectively. Additionally, three new minor derivatives, phototemtides B–D (**2–4**), were identified by detailed HPLC–MS analysis. Phototemtide A (**1**) showed weak antiprotozoal activity against *Plasmodium falciparum*, with an IC$_{50}$ value of 9.8 μM. The biosynthesis of phototemtides A–D (**1–4**) was also proposed.

Cyclic lipopeptides (CLPs) are a class of structurally diverse natural products mainly produced by a wide variety of microorganisms, including cyanobacteria,[1] bacteria,[2] actinobacteria,[3] and fungi.[4] They are generally composed of a fatty acid tail linked to the N terminus of a short oligopeptide and the C terminus of the oligopeptide forms a lactone or lactam with a hydroxy, phenol, or amino functional group of the side chains of peptide or part of the lipid moiety.[5] These compounds can be considered as amphiphiles due to the existence of hydrophobic lipid tails and hydrophilic amino acids or peptides, which endow them with ideal biosurfactant properties.[6] Although the biological activity of CLPs is often reduced to these biosur-

factant properties,[5] they also displayed potent and selective antibacterial,[7] antifungal,[8] antiprotozoal,[9] and cytotoxic activities.[10] Daptomycin, isolated from *Streptomyces roseoporus*, is the first clinically used CLP antibiotic with a new structural type and unique mechanism of action.[11] It was approved by FDA in 2003 for the nontopical treatment of complicated skin and skin structure infections caused by Gram-positive pathogens.[11] As one of the few newly approved antibiotics, the recent success of daptomycin highlights the evolving role of CLPs as important pharmaceutical lead compounds.[5]

The entomopathogenic bacteria of the genera *Photorhabdus* and *Xenorhabdus* that live in symbiosis with nematodes are a rich source of bioactive natural products for killing the insect, nematode development, and protecting the insect cadaver against food competitors.[12–15] Several CLPs, such as xenematides,[16] xefoampeptides,[17] chaiyaphumines,[9] taxllaids,[18] and xentrivalpeptides,[19] have been identified previously. During our further investigation on new natural products from *Photorhabdus* and *Xenorhabdus* strains, we found a new family of CLPs, named phototemtides (Scheme 1), after their biosynthetic gene cluster (BGC) *pttABC* from *Photorhabdus temperata* Meg1 was heterologously expressed in *Escherichia coli*. Here,

[a] L. Zhao, T. D. Vo, Prof. Dr. H. B. Bode
Molecular Biotechnology, Department of Biosciences
Goethe University Frankfurt
Max-von-Laue-Strasse 9, 60438 Frankfurt am Main (Germany)
E-mail: h.bode@bio.uni-frankfurt.de

[b] L. Zhao
Institute of Botany, Jiangsu Province and Chinese Academy of Sciences
QianHuHouCun 1, 210014 Nanjing (China)

[c] M. Kaiser
Swiss Tropical and Public Health Institute
Socinstrasse 57, 4002 Basel (Switzerland)

[d] Prof. Dr. H. B. Bode
Buchmann Institute for Molecular Life Sciences (BMLS)
Goethe University Frankfurt
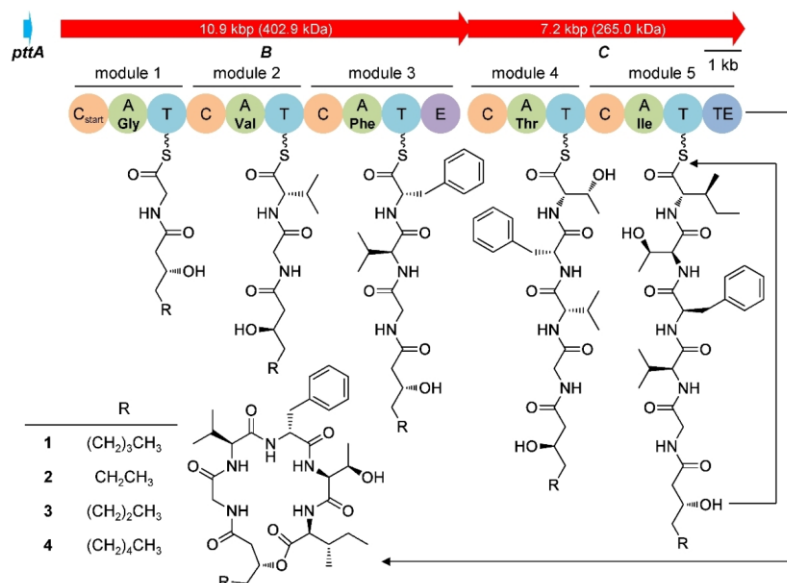Max-von-Laue-Strasse 15, 60438 Frankfurt am Main (Germany)

This article is part of a Special Collection on Microbial Biosynthesis and Interactions. To view the complete collection, visit our homepage
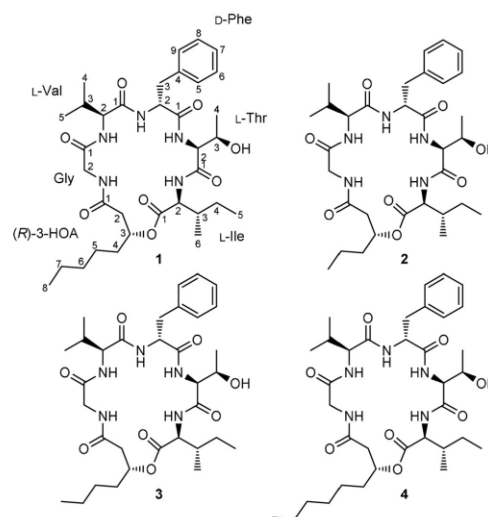
Scheme 1. Chemical structures of phototemtides A–D (**1–4**).

Figure 1. Biosynthetic gene cluster and proposed biosynthesis of phototemtides A–D (**1–4**). Domains: C$_{start}$: starter condensation, C: condensation, A: adenylation, T: thiolation, E: epimerization, TE: thioesterase.

we report the discovery, structural elucidation, biosynthesis, total synthesis, and bioactivity of the major compound phototemtide A (**1**). In addition, three minor derivatives, phototemtides B–D (**2–4**), were identified by detailed HPLC–MS analysis.

Genome mining of *P. temperata* Meg1 (accession No. NZ_JGVH01000010) showed that the strain encodes a BGC with two nonribosomal peptide synthetases (NRPSs), termed PttB (MEG1_RS04970) and PttC (MEG1_RS04975). Detailed analysis identified that PttBC consist of five modules with overall 17 domains, including a starter condensation (C$_{start}$) domain in the initiation module of PttB (Figure 1), which could load a fatty acid to the N terminus of oligopeptide as shown in many lipopeptides, such as anikasin,[5] daptomycin,[20] and taxlllaids.[18] PttBC was thus expected to produce a lipodepsipeptide containing five amino acids. However, no such peptide could be identified in extracts of *P. temperata* Meg1 under standard laboratory conditions. Heterologous expression of an intact BGC in surrogate host has been proven a good strategy to bypass the endogenous regulatory control and activate silent pathways for the production of many interesting natural products,[21] therefore this approach was applied to activate *pttBC* in well-characterized *E. coli*. Notably, one MbtH-encoding gene, termed *pttA* (*MEG1_RS04960*, 195 bp) is located 0.8 kb upstream from *pttB*. The MbtH proteins have been proposed to play an important role in stimulating adenylation reactions that are required for biosynthesis of some nonribosomal peptides.[22,23] Our effort expressing *pttBC* together with *pttA* in *E. coli* resulted in successful production of phototemtides A–D (Figure S1 in the Supporting Information). Indeed, *E. coli* was found incapable of

producing phototemtides without MbtH-encoding gene *pttA* (Figure S1).

To identify the structures of **1–4**, the major compound phototemtide A (**1**) was isolated as a white solid (4 mg) from the XAD-16 extracts of 4 L cultures of *E. coli* expressing *pttABC*. The molecular formula of **1** was determined to be C$_{34}$H$_{53}$N$_5$O$_8$ by its HR-ESI-MS data (*m/z* 660.3984 [*M*+H]$^+$, calcd for C$_{34}$H$_{54}$N$_5$O$_8$, 660.3967; Δppm 2.5), indicating eleven degrees of unsaturation. Its structure was subsequently elucidated based on detailed 1D and 2D NMR experiments (Table 1, Figure 2). The ^{1}H NMR spectrum of **1** exhibited characteristics of a typical lipopeptide, illustrating five amide NH protons (δ_H = 8.75, 8.06, 7.91, 7.74, 7.53 ppm), six α-amino protons (δ_H = 4.50, 4.17, 4.08, 4.04, 3.94, 3.44 ppm), one ester carbinol proton (δ_H = 4.95 ppm), and an alkyl chain (δ_H = 1.60–0.80 ppm). In the ^{13}C NMR spectrum, six carbonyl carbon signals (δ_C = 171.4, 170.9, 170.8, 170.1, 169.5, 168.4 ppm) with five nitrogen-bearing carbon signals (δ_C = 57.8, 57.4, 57.1, 54.6, 42.8 ppm), and two oxygenated sp^3 carbon signals (δ_C = 71.9, 66.1 ppm) were observed. In addition, one phenyl group was identified on the basis of the typical chemical shifts of δ_H = 7.30–7.13 ppm with total integration of five protons in the ^{1}H NMR spectrum and typical chemical shifts of δ_C = 137.9–126.2 ppm in the ^{13}C NMR spectrum. Because six carbonyl carbons and one phenyl ring accounted for ten of the eleven degrees of unsaturation, compound **1** should be a monocyclic peptide. Combining ^{1}H,^{1}H COSY and ^{1}H,^{13}C HSQC NMR data, six partial structures in **1** were constructed as glycine, valine, phenylalanine, threonine, isoleucine, and 3-hydroxyoctanoic acid (3-HOA). The connectiv-

240

Table 1. ^{1}H (500 MHz) and ^{13}C (125 MHz) NMR data for **1** in [D$_6$]DMSO (δ in ppm).

Subunit	Position	δ_C, type	δ_H (J [Hz])	Subunit	Position	δ_C, type	δ_H (J [Hz])
Gly	1	168.4, C		Thr	1	170.1, C	
	2a	42.8, CH$_2$	3.94, overlap		2	57.4, CH	4.17, m
	2b		3.44, dd (16.4, 3.9)		3	66.1, CH	3.94, overlap
	2-NH		7.91, t (5.2)		4	19.1, CH$_3$	0.83, overlap
Val	1	171.4, C			2-NH		7.53, d (8.8)
	2	57.8, CH	4.08, dd (11.9, 5.4)	Ile	1	170.8, C	
	3	30.0, CH	1.71, td (13.6, 6.8)		2	57.1, CH	4.04, dd (9.4, 4.2)
	4	18.7, CH$_3$	0.72, d (6.7)		3	35.9, CH	1.80, m
	5	18.6, CH$_3$	0.50, d (6.7)		4a	24.9, CH$_2$	1.47, m
	2-NH		7.74, d (8.8)		4b		1.18, m
Phe	1	170.9, C			5	11.1, CH$_3$	0.85, overlap
	2	54.6, CH	4.50, m		6	15.3, CH$_3$	0.87, overlap
	3a	35.3, CH$_2$	3.02, dd (13.8, 5.4)		2-NH		8.06, d (6.8)
	3b		2.81, dd (13.8, 10.1)	3-HOA	1	169.5, C	
	4	137.9, C			2	40.4, CH$_2$	2.38, m
	5	129.2, CH	7.28, d (7.2)		3	71.9, CH	4.95, m
	6	128.0, CH	7.23, t (7.5)		4	33.3, CH$_2$	1.54, m
	7	126.2, CH	7.15, t (7.2)		5	24.5, CH$_2$	1.24, m
	8	128.0, CH	7.23, t (7.5)		6	30.8, CH$_2$	1.24, m
	9	129.2, CH	7.28, d (7.2)		7	22.0, CH$_2$	1.24, m
	2-NH		8.75, d (6.9)		8	13.8, CH$_3$	0.85, overlap
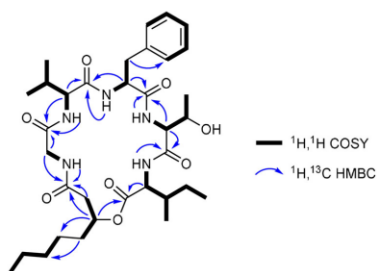
Figure 2. COSY and key HMBC correlations of **1**.

— ^{1}H,^{1}H COSY

⌒ ^{1}H,^{13}C HMBC

ity of these building blocks was established by ^{1}H,^{13}C HMBC. The composition and sequence of the amino acids also perfectly match the prediction of adenylation (A) domain specificity except A5 domain predicting Val but showing Ile (Table S1); the examples of an amino acid different from that predicted in the BGC analysis are not uncommon.[24,25] Thereby, compound **1** was unequivocally elucidated to be a CLP as depicted in Scheme 1, showing fatty acid involved in a lactone ring formation, as it is rare in *Photorhabdus* and *Xenorhabdus* strains and so far has only been observed for the xefoampeptides.[17]

Analysis of MS^2 fragmentation pattern of **1** further confirmed its structure (Figure S2). Ester bond of **1** was first cleaved to give [M+H]$^+$ precursor ion at m/z 660 according to McLafferty rearrangement.[26] Subsequently, the peak at m/z 642 was obtained after the loss of H$_2$O from m/z 660. Further peaks at m/z 529, 428, 281, 182, and 125 indicated the consecutive losses of Ile, Thr, Phe, Val, and Gly, which is identical with amino acid composition and sequence of **1**. With these fragmentation regularities in hand, the minor derivatives **2–4** were identified by detailed analysis of their MS^2 fragmentation pat-

terns (Figure S2). They display similar constitution to **1** but are different in length of fatty acid chains (Scheme 1). As *E. coli* is not capable of producing iso- or anteiso-branched fatty acids due to a missing branched-chain keto acid dehydrogenase complex,[27] we only expect linear fatty acids. Considering that odd-numbered fatty acid containing natural products are uncommon in *E. coli*, a labeling experiment in lysogeny broth (LB) medium with deuterated propionic acid fed as the biosynthetic precursors of odd-chain fatty acids[28] was performed (Figure S3), revealing the incorporation of odd-numbered fatty acid building block (C9) in **4**. That is to say, *E. coli* DH10B MtaA used in this study is capable of synthesizing odd-numbered fatty acids when cultivated in LB medium, albeit in trace amounts.

To determine the absolute configurations of amino acids, compound **1** was hydrolyzed with 6 M HCl as previously described[29] and was analyzed according to the advanced Marfey's method,[30] resulting in the absolute configurations for Val (L), Phe (D), Thr (L), and Ile (L; Figure S4). As nonproteinogenic amino acids L-*allo*-Thr and L-*allo*-Ile are relatively uncommon in natural peptides produced by organisms,[31,32] compound **1** was more likely to feature L-Thr and L-Ile, which was confirmed by subsequent synthesis of **1**. The D-Phe also matches the presence of epimerization (E) domain within module 3. The absolute configurations of amino acids in **2–4** were assumed to be the same as those of **1** because of their common biosynthesis origin.

However, the information regarding the stereochemistry of 3-OH fatty acid was still missing. To address this problem, we conducted the total synthesis of two epimeric **1a** and **1b** containing (R)- and (S)-3-HOA moiety, respectively (Scheme 2). Due to lack of commercial availability, the synthetic work started with preparation of the chiral 3-HOA using samarium(II) iodide (SmI$_2$) mediated Reformatsky-type reaction of (R)- or (S)-4-
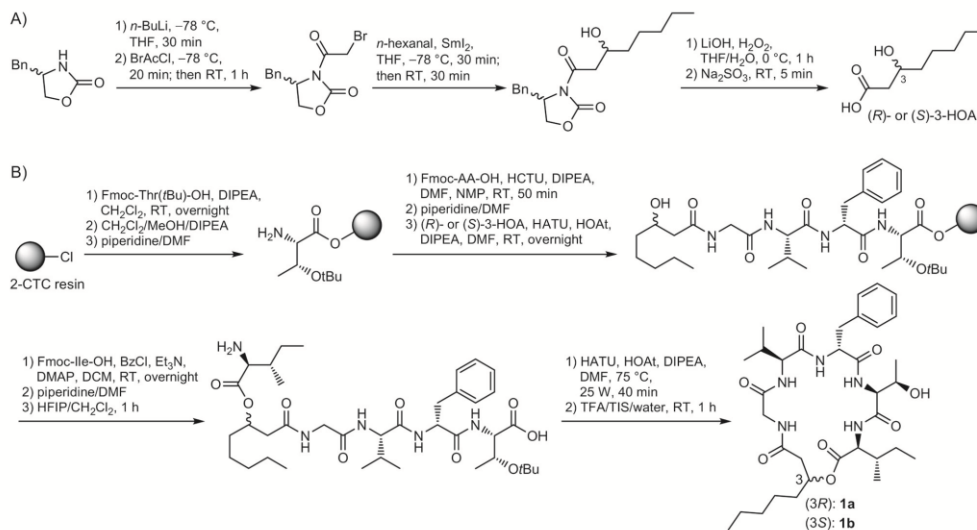
Scheme 2. Total synthesis of epimeric **1a** and **1b**. A) Synthesis of chiral 3-HOA, B) synthesis of epimeric **1a** and **1b**.

benzyl-3-bromoacetyl-2-oxazolidinone with *n*-hexanal, followed by hydrolysis with lithium hydroxide and hydrogen peroxide in aqueous tetrahydrofuran.[33] The next step was to construct the polymer-bound tetrapeptide H$_2$N-Gly-Val-Phe-Thr(*t*Bu)-O-resin by employing standard Fmoc solid-phase peptide synthesis (SPPS) using a 2-chlorotrityl chloride (2-CTC) resin. Following this, the chiral lipid tails were coupled to the N terminus of the tetrapeptide chain by using 1-[bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate (HATU), 1-hydroxy-7-azabenzotriazole (HOAt), and *N*,*N*-diisopropylethylamine (DIPEA) in dimethylformamide (DMF). The key on-resin esterification of 3-HOA was then achieved with Fmoc-Ile-OH using a modified Yamaguchi esterification conditions with benzoyl chloride (BzCl), triethylamine (Et$_3$N), 4-(dimethylamino)pyridine (DMAP) in dichloromethane (DCM) after several attempts.[34] Subsequently, the Fmoc-protecting group was removed via treatment with piperidine in DMF, and the peptide was cleaved from the resin with hexafluoroisopropanol (HFIP) in DCM in order to preserve the side-chain protecting group of Thr.[35] Without purification, we next performed the key macrolactamization step in solution with HATU, HOAt, and DIPEA in DMF, assisted by microwave irradiation.[35] Finally, the remaining side-chain protecting group of peptide was removed by using trifluoroacetic acid (TFA), followed by purification using semipreparative HPLC to give **1a** and **1b** in 6.1 and 1.8% overall yield based on initially actual loading of the resin, respectively. With lipid tail isomers **1a** and **1b** available, we assigned the absolute configuration of 3-HOA in natural **1** to be *R* by comparison of the retention time with synthetic standards since natural **1** showed the identical retention time with synthetic **1a** containing (*R*)-3-HOA moiety, while **1b** containing (*S*)-3-HOA moiety eluted prior to **1/1a** (Figure S5). The absolute

configurations of 3-OH fatty acids in **2–4** were assumed to be the same with that of **1**. Besides, the total synthesis also identified L-Thr and L-Ile, not L-*allo*-Thr and L-*allo*-Ile incorporated in **1**.

Bioactivity testing revealed weak antiparasitic activity of **1** against the causative agent of malaria *Plasmodium falciparum* (IC$_{50}$ = 9.8 μM, Table 2), but no activity against Gram-positive bacteria (*Micrococcus luteus*) or fungi (*Saccharomyces cerevisiae*), and no cytotoxic activity against mammalian L6 cells.

Table 2. Bioactivity of **1** against different protozoan parasites and mammalian L6 cells.

Species	IC$_{50}$ [μM]	
	1	Control[a]
Trypanosoma brucei rhodesiense	62	0.005
Trypanosoma cruzi	83	2.1
Leishmania donovani	>100	0.73
P. falciparum	9.8	0.009
mammalian L6 cells	>100	0.007

[a] The control is different for each target organism: melarsoprol for *T. brucei rhodesiense*, benznidazole for *T. cruzi*, miltefosine for *L. donovani*, chloroquine for *P. falciparum*, and podophyllotoxin for L6 cells.

In summary, a new antimalarial CLP phototemtide A (**1**), with three minor derivatives, phototemtides B-D (**2–4**), was identified from entomopathogenic *P. temperata* Meg1 after the silent NRPS-encoding gene cluster *pttABC* was activated by heterologous expression in *E. coli*. The structural elucidation was accomplished by combining spectroscopic analysis and chemical methods including a total synthesis of phototemtide A. Recent-

242

ly, new peptide drugs have emerged due to the low toxicity, easy synthesis, rapid elimination, and less side effects,[36] but there is currently no approved peptide alternative for the treatment of malaria. The number of marketed active ingredients is limited and most of them face challenges, such as newly observed resistance[37,38] and unpleasant side effects.[39,40] Although the bioactivity of 1 against *P. falciparum* is only weak, it might be a starting point toward a selective *P. falciparum* compound, as it shows no activity against any other tested organisms. With an efficient approach of total synthesis in hand, further investigation could focus on structure–activity relationships and subsequent in vivo experiments of this new family of CLPs. This work is also the first example of the importance of MbtH for peptide production in *Photorhabdus* and we are currently studying the role of PttA in the original producer and related strains.

Acknowledgements

Conflict of Interest

The authors declare no conflict of interest.

[1] S. Luo, A. Krunic, H.-S. Kang, W.-L. Chen, J. L. Woodard, J. R. Fuchs, S. M. Swanson, J. Orjala, *J. Nat. Prod.* **2014**, *77*, 1871–1880.
[2] G. Jahanshah, Q. Yan, H. Gerhardt, Z. Pataj, M. Lämmerhofer, I. Pianet, M. Josten, H.-G. Sahl, M. W. Silby, J. E. Loper, H. Gross, *J. Nat. Prod.* **2019**, *82*, 301–308.
[3] Z. Yu, S. Vodanovic-Jankovic, M. Kron, B. Shen, *Org. Lett.* **2012**, *14*, 4946–4949.
[4] F. J. Ortíz-López, M. C. Monteiro, V. González-Menéndez, J. R. Tormo, O. Genilloud, G. F. Bills, F. Vicente, C. Zhang, T. Roemer, S. B. Singh, E. Reyes, *J. Nat. Prod.* **2015**, *78*, 468–475.
[5] S. Götze, R. Herbst-Irmer, M. Klapper, H. Görls, K. R. A. Schneider, R. Barnett, T. Burks, U. Neu, P. Stallforth, *ACS Chem. Biol.* **2017**, *12*, 2498–2502.
[6] M. Inès, G. Dhouha, *Peptides* **2015**, *71*, 100–112.
[7] T. P. Wyche, Y. Hou, E. Vazquez-Rivera, D. Braun, T. S. Bugni, *J. Nat. Prod.* **2012**, *75*, 735–740.
[8] T. Neuhof, P. Schmieder, K. Preussel, R. Dieckmann, H. Pham, F. Bartl, H. von Döhren, *J. Nat. Prod.* **2005**, *68*, 695–700.
[9] F. Grundmann, M. Kaiser, M. Schiell, A. Batzer, M. Kurz, A. Thanwisai, N. Chantratita, H. B. Bode, *J. Nat. Prod.* **2014**, *77*, 779–783.
[10] K. C. Tan, T. Wakimoto, I. Abe, *Org. Lett.* **2014**, *16*, 3256–3259.
[11] C. Vilhena, A. Bettencourt, *Mini-Rev. Med. Chem.* **2012**, *12*, 202–209.
[12] J. M. Crawford, J. Clardy, *Chem. Commun.* **2011**, *47*, 7559–7566.
[13] M. I. Vizcaino, X. Guo, J. M. Crawford, *J. Ind. Microbiol. Biotechnol.* **2014**, *41*, 285–299.
[14] H. B. Bode, *Curr. Opin. Chem. Biol.* **2009**, *13*, 224–230.
[15] Y.-M. Shi, H. B. Bode, *Nat. Prod. Rep.* **2018**, *35*, 309–335.
[16] J. M. Crawford, C. Portmann, R. Kontnik, C. T. Walsh, J. Clardy, *Org. Lett.* **2011**, *13*, 5144–5147.
[17] N. J. Tobias, H. Wolff, B. Djahanschiri, F. Grundmann, M. Kronenwerth, Y.-M. Shi, S. Simonyi, P. Grün, D. Shapiro-Ilan, S. J. Pidot, T. P. Stinear, I. Ebersberger, H. B. Bode, *Nat. Microbiol.* **2017**, *2*, 1676–1685.
[18] M. Kronenwerth, K. A. J. Bozhüyük, A. S. Kahnt, D. Steinhilber, S. Gaudriault, M. Kaiser, H. B. Bode, *Chem. Eur. J.* **2014**, *20*, 17478–17487.
[19] Q. Zhou, A. Dowling, H. Heide, J. Wöhnert, U. Brandt, J. Baum, R. Ffrench-Constant, H. B. Bode, *J. Nat. Prod.* **2012**, *75*, 1717–1722.
[20] L. Robbel, M. A. Marahiel, *J. Biol. Chem.* **2010**, *285*, 27501–27508.
[21] S. H. Liu, W. Wang, K. B. Wang, B. Zhang, W. Li, J. Shi, R. H. Jiao, R. X. Tan, H. M. Ge, *Org. Lett.* **2019**, *21*, 3785–3788.
[22] R. H. Baltz, *J. Ind. Microbiol. Biotechnol.* **2011**, *38*, 1747–1760.
[23] J. Liu, B. Wang, H. Li, Y. Xie, Q. Li, X. Qin, X. Zhang, J. Ju, *Org. Lett.* **2015**, *17*, 1509–1512.
[24] L. Zhao, H. B. Bode, *Org. Biomol. Chem.* **2019**, *17*, 7858–7862.
[25] L. Zhao, R. M. Awori, M. Kaiser, J. Groß, T. Opatz, H. B. Bode, *J. Nat. Prod.* **2019**, *82*, 3499–3503.
[26] S.-Z. Yang, D.-Z. Wei, B.-Z. Mu, *J. Biochem. Biophys. Methods* **2006**, *68*, 69–74.
[27] A. O. Brachmann, D. Reimer, W. Lorenzen, E. Augusto Alonso, Y. Kopp, J. Piel, H. B. Bode, *Angew. Chem. Int. Ed.* **2012**, *51*, 12086–12089; *Angew. Chem.* **2012**, *124*, 12252–12255.
[28] M. Pfeuffer, A. Jaudszus, *Adv. Nutr.* **2016**, *7*, 730–734.
[29] F. Grundmann, M. Kaiser, M. Kurz, M. Schiell, A. Batzer, H. B. Bode, *RSC Adv.* **2013**, *3*, 22072–22077.
[30] K. Fujii, Y. Ikai, H. Oka, M. Suzuki, K. Harada, *Anal. Chem.* **1997**, *69*, 5146–5151.
[31] A. A. Tymiak, T. J. McCormick, S. E. Unger, *J. Org. Chem.* **1989**, *54*, 1149–1157.
[32] Y. Song, Q. Li, X. Liu, Y. Chen, Y. Zhang, A. Sun, W. Zhang, J. Zhang, J. Ju, *J. Nat. Prod.* **2014**, *77*, 1937–1941.
[33] S. Fukuzawa, H. Matsuzawa, S. Yoshimitsu, *J. Org. Chem.* **2000**, *65*, 1702–1706.
[34] K. Hung, P. W. R. Harris, A. M. Heapy, M. A. Brimble, *Org. Biomol. Chem.* **2011**, *9*, 236–242.
[35] F. I. Nollmann, A. Dowling, M. Kaiser, K. Deckmann, S. Grösch, R. Ffrench-Constant, H. B. Bode, *Beilstein J. Org. Chem.* **2012**, *8*, 528–533.
[36] Q.-S. Du, N.-Z. Xie, R.-B. Huang, *Med. Chem.* **2015**, *11*, 235–247.
[37] C. V. Plowe, *Curr. Top. Microbiol. Immunol.* **2005**, *295*, 55–79.
[38] A.-C. Uhlemann, S. Krishna, *Curr. Top. Microbiol. Immunol.* **2005**, *295*, 39–53.
[39] T. Weinke, M. Trautmann, T. Held, G. Weber, D. Eichenlaub, K. Fleischer, W. Kern, H. D. Pohle, *Am. J. Trop. Med. Hyg.* **1991**, *45*, 86–91.
[40] R. W. Steketee, J. J. Wirima, L. Slutsker, C. O. Khoromana, D. L. Heymann, J. G. Breman, *Am. J. Trop. Med. Hyg.* **1996**, *55*, 50–56.

COMMUNICATIONS

L. Zhao, T. D. Vo, M. Kaiser, H. B. Bode*

■■ – ■■

Phototemtide A, a Cyclic Lipopeptide Heterologously Expressed from *Photorhabdus temperata* Meg1, Shows Selective Antiprotozoal Activity
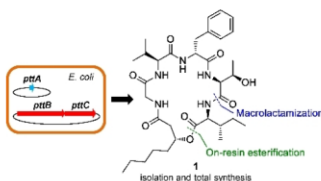
Cyclic lipopeptides (CLPs) are important pharmaceutical lead compounds. We report a new antimalarial CLP phototemtide A with three minor derivatives, phototemtides B–D, which are generated after the silent NRPS-encoding gene cluster *pttABC* from entomopathogenic *P. temperata* Meg1 was activated by heterologous expression in *E. coli*. The structures of these new CLPs were unequivocally identified by combining biological and chemical methods.

↖↖ These are not the final page numbers!

Supporting Information

Phototemtide A, a Cyclic Lipopeptide Heterologously Expressed from *Photorhabdus temperata* Meg1, Shows Selective Antiprotozoal Activity

Lei Zhao,[1,2] Tien Duy Vo,[1] Marcel Kaiser,[3] and Helge B. Bode*[1,4]

[1]Molecular Biotechnology, Department of Biosciences, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

[2]Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, 210014 Nanjing, China

[3]Swiss Tropical and Public Health Institute, 4002 Basel, Switzerland

[4]Buchmann Institute for Molecular Life Sciences (BMLS), Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

*Corresponding author

Experimental Procedures

Strain Construction

For construction of heterologous *Escherichia coli* expressing *pttABC* (*MEG1_RS04960*, *MEG1_RS04970*, and *MEG1_RS04975*) or *pttBC*, plasmids pLZ57 and pLZ58 were generated by yeast cloning[1] and Gibson cloning, respectively, and were verified by restriction enzyme digest. The correct pLZ57 was transformed into *E. coli* DH10B MtaA for coexpression with or without pLZ58. Individual colonies were analyzed by HPLC-MS for the phototemtide production.

Strain Cultivation and Culture Extraction

100 μL of overnight cultures of heterologous production *E. coli* were inoculated into 10 mL fresh liquid lysogeny broth (LB) medium supplemented with appropriate antibiotics (50 μg/mL kanamycin and spectinomycin), 0.1% L-arabinose (from a 25% stock solution), and 2% Amberlite XAD-16 resin. The cultures were grown at 30 °C and were shaken on a rotary shaker at 200 rpm. After 3 days, the cultures were harvested, and XAD-16 beads were separated and extracted with 10 mL methanol for 1 h. Subsequently, the extracts were analyzed by HPLC-MS.

HPLC-MS Analysis

HPLC-HR-ESI-MS analysis was carried out on a Dionex UltiMate 3000 system coupled to a Bruker Impact II QTOF mass spectrometer. Compounds were eluted on an ACQUITY UPLC BEH C$_{18}$ column (130 Å, 2.1 mm × 50 mm, 1.7 μm) using a gradient from 5% to 95% acetonitrile water solution containing 0.1% formic acid at a flow rate of 0.4 mL/min for 16 min. HPLC-ESI-MS analysis was recorded on a Dionex UltiMate 3000 system coupled to a Bruker AmaZonX mass spectrometer. Compounds were eluted on an ACQUITY UPLC BEH C$_{18}$ column (130 Å, 2.1 mm × 100 mm, 1.7 μm) using a gradient from 5% to 95% acetonitrile water solution containing 0.1% formic acid at a flow rate of 0.6 mL/min for 16 min. Positive mode with scan range from 100 to 1200 *m/z* was used to detect phototemtides.

Compound Isolation

To isolate phototemtide A (**1**), the XAD-16 beads from 4 L cultures of *E. coli* expressing *pttABC* were extracted with ethyl acetate. The purification of **1** was performed on an

Agilent semipreparative HPLC system with an Eclipse XDB-C$_{18}$ column (9.4 mm × 250 mm, 5 μm) using a gradient 50% to 55% acetonitrile water solution containing 0.1% formic acid for 20 min at a flow rate of 3 mL/min to afford **1** (4 mg).

Configuration Determination of Amino Acids

Amino acid configurations in **1** were determined using the advanced Marfey's method as described previously.[2,3] Approximately 1 mg of peptide was dissolved in 200 μL MeOH and was hydrolyzed with 800 μL HCl (6 M) in an Ace high-pressure tube at 110 ℃ for 1 h. The hydrolysate was evaporated to dryness and redissolved in 100 μL H$_2$O. To each a half portion (45 μL) were added 10 μL 1M NaHCO$_3$ and 80 μL 1% FDLA Nα-(5-fluoro-2,4-dinitrophenyl)-L-leucinamide or D-leucinamide (L-FDLA or D-FDLA, dissolved in acetone). The brown reaction vials were incubated in a water bath at 40 ℃ for 1 h. After that the reactions were cooled to room temperature, quenched with 10 μL 1 M HCl and evaporated to dryness. The residue was dissolved in 400 μL MeOH. The analysis of the L- and LD-FDLA-derivatized amino acids was carried out with HPLC-MS.

Synthesis of Chiral 3-Hydroxyoctanoic Acids

The synthesis was carried out by employing key samarium(II) iodide (SmI$_2$) mediated asymmetric Reformatsky-type reaction of chiral 4-benzyl-3-bromoacetyl-2-oxazolidinone with *n*-hexanal.[4]

Step a - Preparation of chiral 4-benzyl-3-bromoacetyl-2-oxazolidinone. In a two-neck round-bottom flask containing a magnetic stirring bar were charged (*R*)- or (*S*)-4-benzyl-2-oxazolidinone (352 mg, 3.0 mmol) and tetrahydrofuran (THF, 12 mL) under a slight pressure of nitrogen. The flask was cooled in a dry ice–isopropanol bath (−78 °C), and a *n*-hexane solution of *n*-butyllithium (*n*-BuLi, 1.6 M, 2.1 mL, 3.3 mmol) was then added using a syringe through the septum with magnetic stirring. After 30 min, bromoacetyl chloride (BrAcCl, 482 mg, 3.1 mmol) was slowly added to the mixture at the same temperature over a period of 20 min. When the addition was completed, the dry ice bath was removed, and the mixture was allowed to warm to room temperature and stirred for an additional 1 h. The reaction was quenched with saturated aqueous potassium hydrogen phosphate (K$_2$HPO$_4$, 3 mL), and the solution was then extracted with three 3 mL portions of diethyl ether. The combined extracts were dried over magnesium sulfate (MgSO$_4$), and the solvent was removed on a rotary evaporator

leaving an orange liquid. The crude product was purified by flash silica gel column chromatography to give (R)- or (S)-4-benzyl-3-bromoacetyl-2-oxazolidinone as a yellow solid (~300 mg, 33% yield).

Step b - SmI2 mediated Reformatsky-type reaction. All reactions were carried out in a nitrogen atmosphere using a Schlenk tube with standard techniques for air-sensitive materials. A mixture of *n*-hexanal (124 μL, 1.0 mmol) and (R)- or (S)-4-benzyl-3-bromoacetyl-2-oxazolidinone (299 mg, 1.0 mmol) in THF (2 mL) was slowly dropwise injected to a THF solution of SmI2 (0.1 M, 30 mL, 3.0 mmol) at –78 °C. After stirring at –78 °C for 0.5 h, the dry ice bath was removed, and the mixture was allowed to warm to room temperature and was stirred for an additional 0.5 h. The reaction was quenched with hydrochloric acid (HCl, 0.1 M, 25 mL), and the aqueous phase was extracted with three 20 mL portions of diethyl ether. The organic layer was washed with saturated aqueous sodium thiosulfate ($Na_2S_2O_3$), and then dried over $MgSO_4$. The solvent was removed under reduced pressure, and the residue was purified by flash silica gel column chromatography to afford (R)-4-Benzyl-3-(R)-3-hydroxyoctanoyl-2-oxazolidinone or (S)-4-Benzyl-3-(S)-3-hydroxyoctanoyl-2-oxazolidinone as a white solid (~195 mg, 60% yield).

Step c - Hydrolysis of the chiral 4-benzyl-3-3-hydroxyoctanoyl-2-oxazolidinone. In a round-bottom flask containing a magnetic stirring bar was charged (R)-4-Benzyl-3-(R)-3-hydroxyoctanoyl-2-oxazolidinone or (S)-4-Benzyl-3-(S)-3-hydroxyoctanoyl-2-oxazolidinone (192 mg, 0.6 mmol). A solution of THF in H_2O (4:1, 3 mL) was added to the flask. The flask was cooled in an ice-water bath, and a solution of hydrogen peroxide (H_2O_2) (30% in H_2O, 273 μL, 2.4 mmol) and a solution of lithium hydroxide (LiOH, 24 mg, 1.0 mmol) in H_2O (1.2 mL) were added via syringe. The resulting mixture was stirred for 1 h at 0 °C. The septum was then removed, and a solution of sodium sulfite (Na_2SO_3, 300 mg, 2.4 mmol) in H_2O (1.8 mL) was added. The THF was removed on rotary evaporator and the remaining solution was extracted three times with dichloromethane (DCM, 3.6 mL) to remove the free oxazolidinone. The aqueous layer containing the carboxylate salt of the desired product was cooled in an ice bath and was brought to pH 1 with 6 M HCl. The aqueous solution was then extracted five times with ethyl acetate (2.4 mL). The combined ethyl acetate layers were dried over $MgSO_4$ and the solvent was removed on

a rotary evaporator to give (*R*)- or (*S*)-3-hydroxyoctanoic acid (3-HOA) as a colorless oil (~85 mg, 89 % yield). Their structures were confirmed by ^{1}H NMR.

(*R*)-3-Hydroxyoctanoic acid: colorless oil; ^{1}H NMR (250 MHz, CDCl$_3$): δ 4.03 (dt, *J* = 8.4, 3.6 Hz, 1H), 2.59 (dd, *J* = 16.6, 3.5 Hz, 1H), 2.47 (dd, *J* = 16.6, 8.6 Hz, 1H), 1.39 (m, 8H), 0.90 (t, *J* = 6.4 Hz, 3H).

(*S*)-3-Hydroxyoctanoic acid: colorless oil; ^{1}H NMR (250 MHz, CDCl$_3$): δ 4.07 (m, 1H), 2.59 (dd, *J* = 16.5, 3.5 Hz, 1H), 2.48 (dd, *J* = 16.4, 8.8 Hz, 1H), 1.40 (m, 8H), 0.88 (t, *J* = 6.4 Hz, 3H).

Synthesis of Epimeric Peptides

The synthesis was accomplished by combining solid and solution phase synthesis.[5,6]

Step a - Loading of Thr(tBu) on the 2-chlorotrityl chloride (2-CTC) resin. A solution of Fmoc-Thr(tBu)-OH (119 mg, 0.3 mmol, 3 equiv) and *N*,*N*-diisopropylethylamine (DIPEA, 153 μL, 0.9 mmol, 9 equiv) in DCM (1.5 mL) was placed in a plastic reactor vessel filled with 2-CTC resin (63 mg, 0.1 mmol, 1.0 equiv). The resulting mixture was incubated at room temperature overnight. The remaining free binding sites were capped upon incubating twice with DCM/MeOH/DIPEA (80:15:5) for 10 min at room temperature. The resin was washed several times with dimethylformamide (DMF), MeOH, and DCM, and treated with 20% piperidine in DMF (3 × 10 min) to remove the Fmoc-protecting group. The combined filtrates were used to determine the actual loading of the resin at λ_{301} nm. Afterwards, the resin was washed with DCM and dried.

Step b - Solid-phase peptide synthesis (SPPS). The linear sequence was synthesized twice on the 2-CTC resin preloaded with L-Thr(tBu) on a 25 μmol scale with a Syro Wave peptide synthesizer by using standard Fmoc chemistry. The resin was placed in a plastic reactor vessel with a Teflon frit and an amount of 6 equiv of amino acid derivatives (Fmoc-D-Phe-OH, Fmoc-Val-OH, Fmoc-Gly-OH, 0.2 M) were activated in situ at room temperature with *O*-(6-Chlorobenzotriazol-1-yl)-*N*,*N*,*N'*,*N'*-tetramethyluronium hexafluorophosphate (HCTU, 0.6 M, 6 equiv) in DMF in the presence of DIPEA (2.4 M, 12 equiv) in *N*-methyl-2-pyrrolidone (NMP) for 50 min. Fmoc-protecting groups were removed with 40% piperidine in DMF for 5 min and the deprotection step was repeated for another 10 min with 20% piperidine in DMF. After each coupling and deprotection

step, the resin was washed with NMP. After the addition of the final amino acid residue, the resin was washed with NMP, DMF, and DCM and dried.

Step c - Attachment of the chiral 3-HOA on the N-terminus of polymer-bound tetrapeptide. A solution of (*R*)- or (*S*)-3-HOA (20 mg, 0.125 mmol, 5 equiv), 1-[bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate (HATU) (48 mg, 0.125 mmol, 5 equiv), 1-hydroxy-7-azabenzotriazole (HOAt) (17 mg, 0.125 mmol, 5 equiv) and DIPEA (43 μL, 0.25 mmol, 10 equiv) in DMF (1.5 mL) was added to resin. The resulting mixture was incubated in a plastic reactor vessel equipped with a Teflon frit at room temperature overnight. The resin was washed with DMF and DCM, and was treated with 40% piperidine in DMF for 5 min and then 20% piperidine in DMF for 10 min at room temperature to remove the Fmoc-protecting group. Afterwards the resin was washed with DCM and dried.

Step d - Esterification and cleavage from the resin. For the esterification of chiral 3-HOA with Fmoc-Ile-OH, resin was added with Fmoc-Ile-OH (177 mg, 0.5 mmol, 20 equiv), benzoyl chloride (BzCl, 58 μL, 0.5 mmol, 20 equiv), triethylamine (Et$_3$N, 140 μL, 1.0 mmol, 40 equiv), and 4-(dimethylamino)pyridine (DMAP, 1.2 mg, 0.1 mmol, 0.4 equiv) in DCM (3 mL) and was stirred at room temperature overnight. After reaction, the Fmoc-protecting group was removed by using 40% piperidine in DMF for 5 min and then 20% piperidine in DMF for 10 min at room temperature. The protected branched peptide was cleaved with 20% hexafluoroisopropanol (HFIP) in DCM at room temperature for 1 h.

Step e - Cyclization and full deprotection. The peptide was cyclized in solution assisted by microwave irradiation (40 min, 25 W, 75 °C) by using HATU (19 mg, 50 μmol, 2 equiv), HOAt (6.8 mg, 50 μmol, 2 equiv), and DIPEA (17 μL, 0.1 mmol, 4 equiv) in DMF (c = 1 mM). The cyclized product was fully deprotected by incubation with 95% trifluoroacetic acid (TFA) and 2.5% triisopropylsilane (TIPS) in H$_2$O at room temperature for 1 h. The peptide was purified by Agilent semipreparative HPLC system to give **1a** and **1b** in (1.0 mg) 6.1% and (0.3 mg) 1.8% overall yield based on initially actual loading of the resin, respectively. Their structures were confirmed by HR-ESI-MS and ^{1}H NMR.

Compound **1a**: white, amorphous solid; HR-ESI-MS *m/z* 660.3956 [M + H]$^+$ (calcd for C$_{34}$H$_{54}$N$_5$O$_8$, 660.3967, Δppm 1.7); ^{1}H NMR (500 MHz, DMSO-*d*$_6$): δ 8.73 (d, *J* = 7.0 Hz,

1H), 8.05 (d, J = 7.0 Hz, 1H), 7.90 (m, 1H), 7.73 (d, J = 8.8 Hz, 1H), 7.52 (d, J = 8.8 Hz, 1H), 7.28 (d, J = 7.2 Hz, 2H), 7.23 (m, 2H), 7.15 (t, J = 7.2 Hz, 1H), 4.95 (m, 1H), 4.50 (m, 1H), 4.17 (dd, J = 8.8, 4.3 Hz, 1H), 4.08 (overlap, 1H), 4.04 (overlap, 1H), 3.94 (overlap, 2H), 3.45 (dd, J = 16.4, 4.3 Hz, 1H), 3.02 (dd, J = 13.8, 5.4 Hz, 1H), 2.81 (dd, J = 13.8, 10.1 Hz, 1H), 2.38 (m, 2H), 1.80 (m, 1H), 1.72 (m, 1H), 1.54 (m, 2H), 1.47 (m, 1H), 1.24 (m, 6H), 1.18 (m, 1H), 0.87 (overlap, 3H), 0.85 (overlap, 6H), 0.83 (overlap, 3H), 0.72 (d, J = 6.8 Hz, 3H), 0.51 (d, J = 6.7 Hz, 3H).

Compound **1b**: white, amorphous solid; HR-ESI-MS m/z 660.3957 [M + H]$^+$ (calcd for C$_{34}$H$_{54}$N$_5$O$_8$, 660.3967, Δppm 1.6); ^{1}H NMR (500 MHz, DMSO-d_6): δ 8.70 (d, J = 6.6 Hz, 1H), 7.96 (d, J = 8.8 Hz, 1H), 7.91 (d, J = 8.8 Hz, 1H), 7.82 (t, J = 5.6 Hz, 1H), 7.47 (d, J = 6.7 Hz, 1H), 7.28 (d, J = 7.2 Hz, 2H), 7.23 (t, J = 7.6 Hz, 2H), 7.16 (t, J = 6.7 Hz, 1H), 5.05 (d, J = 5.3 Hz, 1H), 4.47 (m, 1H), 4.32 (dd, J = 8.7, 5.3 Hz, 1H), 4.17 (dd, J = 8.8, 3.0 Hz, 1H), 4.10 (m, 1H), 4.02 (t, J = 7.2 Hz, 1H), 3.73 (dd, J = 16.7, 5.5 Hz, 1H), 3.66 (dd, J = 17.1, 5.2 Hz, 1H), 3.07 (dd, J = 13.9, 5.6 Hz, 1H), 2.78 (dd, J = 13.9, 10.0 Hz, 1H), 1.99 (m, 2H), 1.73 (m, 1H), 1.54 (m, 2H), 1.46 (m, 1H), 1.24 (m, 6H), 1.18 (m, 1H), 0.85 (overlap, 12H), 0.72 (d, J = 6.8 Hz, 3H), 0.52 (d, J = 6.7 Hz, 3H).

Bioactivity Testing

Bioactivity testing of **1** against the parasites *Trypanosoma brucei rhodesiense* STIB900, *Trypanosoma cruzi* Tulahuen C4, *Leishmania donovani* MHOM-ET-67/L82, and *Plasmodium falciparum* NF54 was carried out as described previously.[7] Cytotoxicity of **1** against rat skeletal myoblasts (L6 cells) was evaluated as described previously.[7] Bioactivity of **1** against *Micrococcus luteus* and *Saccharomyces cerevisiae* was tested using filter paper discs loaded with MeOH-dissolved compound (1 mg/mL) on the prepared plates which were incubated at 30 °C for 2 days.

Supplementary Tables

Table S1. A domain specificity prediction of PttABC

A domain	Stachelhaus sequence	most likely amino acid predicted	amino acid detected
A1	DILQIGVIWK	Gly	Gly
A2	DAYWLGGTFK	Val	Val
A3	DAWTVAAVCK	Phe	Phe
A4	DFWNVGMVHK	Thr	Thr
A5	DAYFFGITYK	Val	Ile

Table S2. HR-MS data of natural and synthetic compounds

compound	sum formula	calcd. $[M + H]^+$	found $[M + H]^+$	Δppm
natural **1**	$C_{34}H_{53}N_5O_8$	660.3967	660.3984	2.5
natural **2**	$C_{32}H_{49}N_5O_8$	632.3654	632.3659	0.8
natural **3**	$C_{33}H_{51}N_5O_8$	646.3810	646.3825	2.3
natural **4**	$C_{35}H_{55}N_5O_8$	674.4123	674.4135	1.8
synthetic **1a**	$C_{34}H_{53}N_5O_8$	660.3967	660.3956	1.7
synthetic **1b**	$C_{34}H_{53}N_5O_8$	660.3967	660.3957	1.6

Table S3. Bacterial strains used in this study

strain	genotype	reference
E. coli DH10B MtaA	F-*mcr*A, Δ(*mrr-hsd*RMS-*mcr*BC), Φ80*lac*ZΔM15, Δ*lac*X74, *rec*A1, *end*A1, *ara*D139, Δ(*ara leu*)7697, *gal*U, *gal*K, *rps*L, *nup*G, λ–, *ent*D::*mtaA*	1
P. temperata Meg1	wild type	8

Table S4. Plasmids used in this study

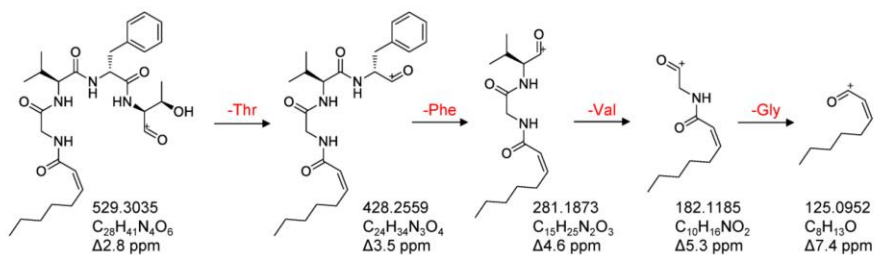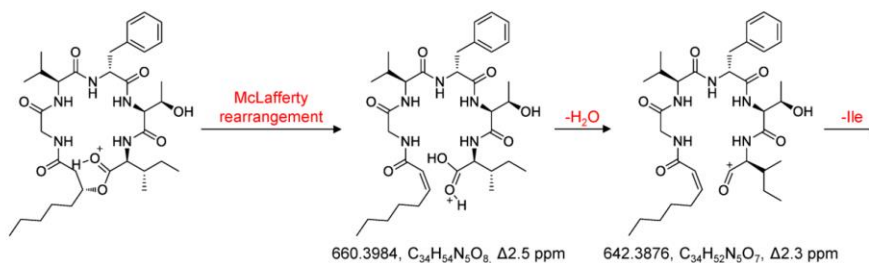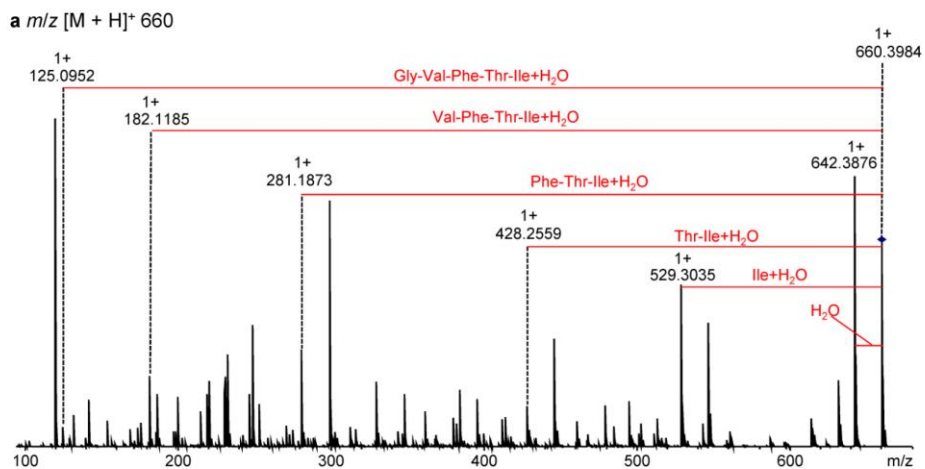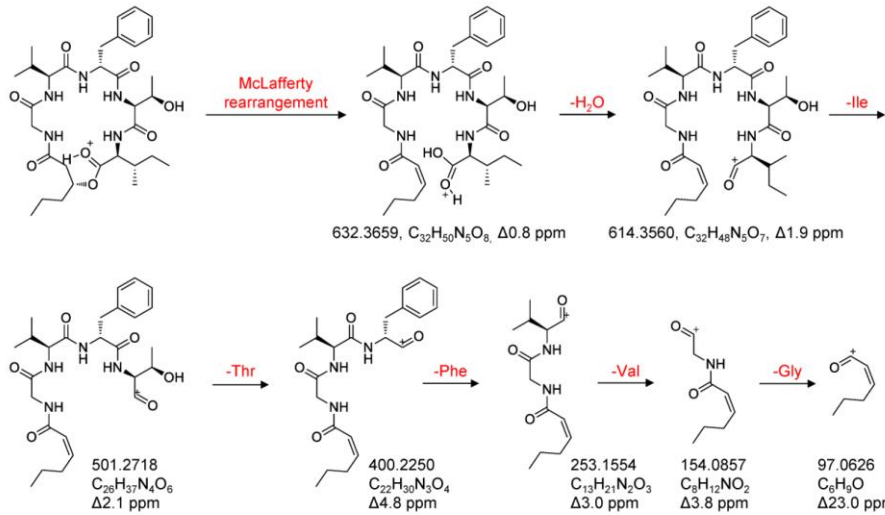| plasmid	genotype/description	reference
pFF1	2μ ori, G418^R, pBAD promoter, pCOLA ori, MCS, Ypet-Flag, Km^R	9
pCDF-ara-tacI	3404 bp, modified from pCDF_tacI/I that contains arabinose-inducible promoter, Sm^R	Bode lab
pLZ57	25113 bp, *pttBC* from Meg1 genomic DNA assembled into pFF1, Km^R	this work
pLZ58	3560 bp, *pttA* from Meg1 genomic DNA assembled into pCDF, Sm^R	this work

Table S5. Primers used in this study

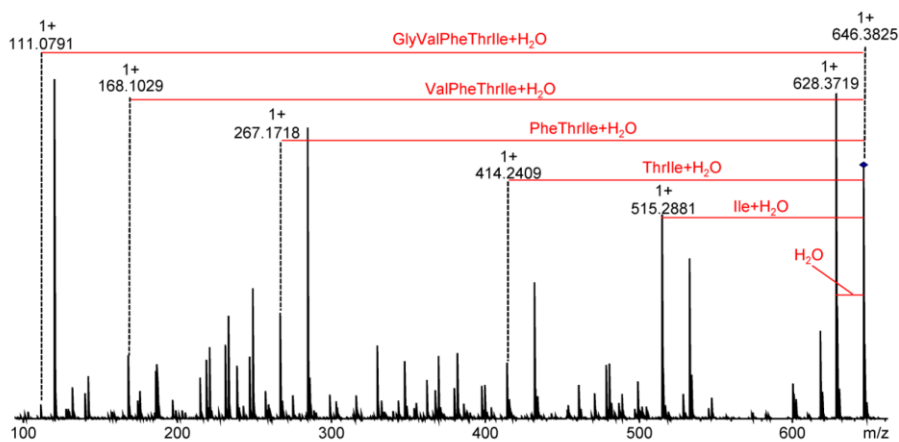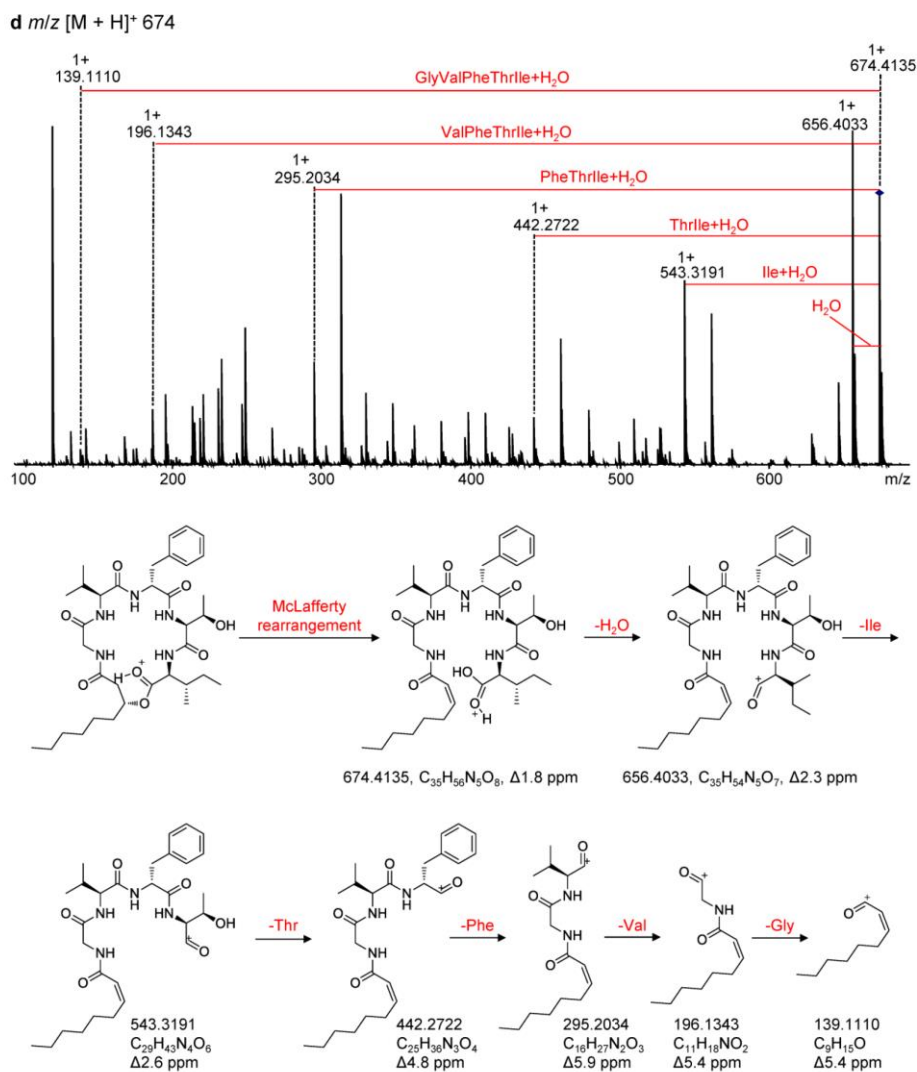| primer	sequence (5'-3')	targeting DNA fragment	plasmid
LZ_152	TCGCAACTCTCTACTGTTTCTCCATACC CGTTTTTTTGGGCTAACAGGAGGAATTC CATGAATAAGGATGGATATT	fragment I of *pttBC* from Meg1 (6727 bp)	
LZ_153	AGTAGAGGAGCAAGTGATGC		
LZ_154	TGTTCCATGTTGAACCGAATC	fragment II of *pttBC* from Meg1 (7557 bp)	pLZ57
LZ_155	AATGCTTCATCATTGAGCGAC		
LZ_156	AGCAGAGAATGTGGCTGTTAG		
LZ_157	TCTTCACCTTTGCTCATGAACTCGCCAG AACCAGCAGCGGAGCCAGCGGATCCG GCGCGCCTTACTCAATGCTGCTGTCTG	fragment III of *pttBC* from Meg1 (3917 bp)	
LZ_148	TTGAGCAATCCATTTGAAAATGAATC	*pttA* from Meg1 (195 bp)	
LZ_158	TCACTGCAAACTTTTTAGCCTC		
LZ_159	AGATATGAGGCTAAAAAGTTTGCAGTGA CAATTAATCATCGGCTCGTATAATG	pCDF-ara-tacI vector backbone (3423 bp)	pLZ58
LZ_151	AAGAGATTCATTTTCAAATGGATTGCTC AAGGAATTCCTCCTGTTAGCCC		

S9

253

Supplementary Figures
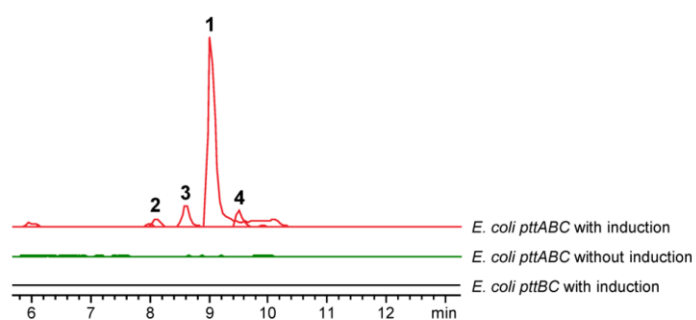
Figure S1. HPLC-MS analysis of XAD-16 extracts from *E. coli* expressing pttABC with and without arabinose induction and *E. coli* expressing pttBC with arabinose induction. Extracted ion chromatograms (EICs) for **1−4** are shown.

a *m/z* [M + H]$^+$ 660

660.3984, C$_{34}$H$_{54}$N$_5$O$_8$, Δ2.5 ppm

642.3876, C$_{34}$H$_{52}$N$_5$O$_7$, Δ2.3 ppm

529.3035
C$_{28}$H$_{41}$N$_4$O$_6$
Δ2.8 ppm

428.2559
C$_{24}$H$_{34}$N$_3$O$_4$
Δ3.5 ppm

281.1873
C$_{15}$H$_{25}$N$_2$O$_3$
Δ4.6 ppm

182.1185
C$_{10}$H$_{16}$NO$_2$
Δ5.3 ppm

125.0952
C$_8$H$_{13}$O
Δ7.4 ppm

S11

255

b *m/z* [M + H]⁺ 632

S12

c m/z $[M + H]^+$ 646

d *m/z* [M + H]⁺ 674
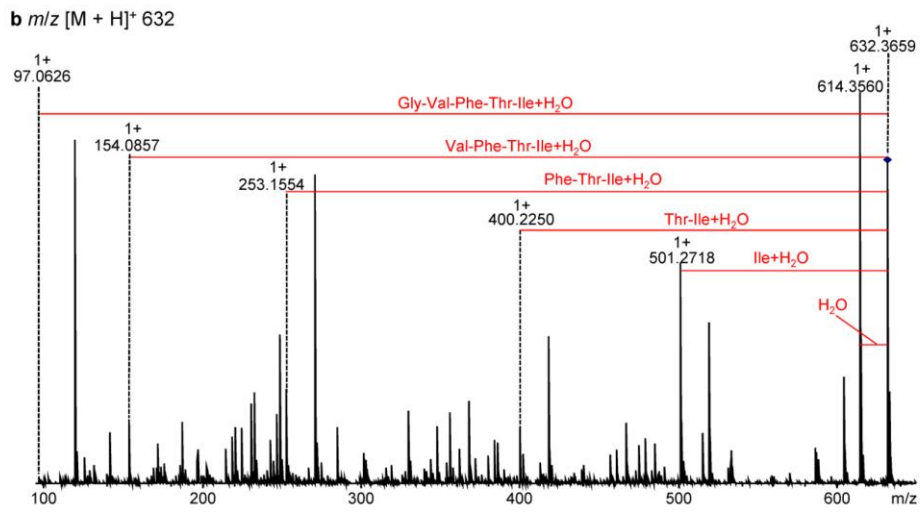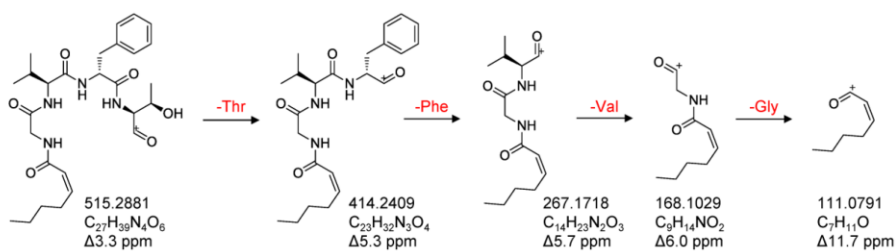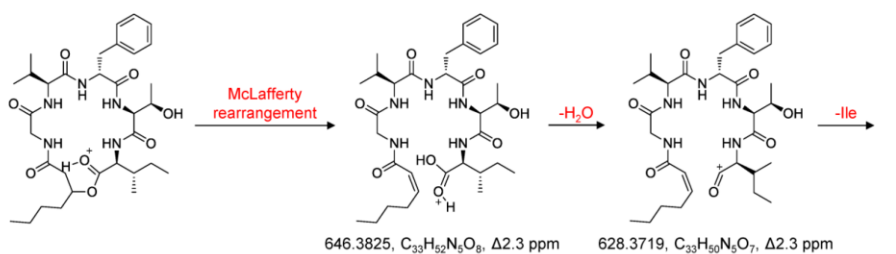
Figure S2. Structure identification of **1**−**4** by detailed HPLC-MS analysis. MS² spectra and fragmentation pathways with proposed fragment structures for **1** (a), **2** (b), **3** (c), and **4** (d).

S14

258
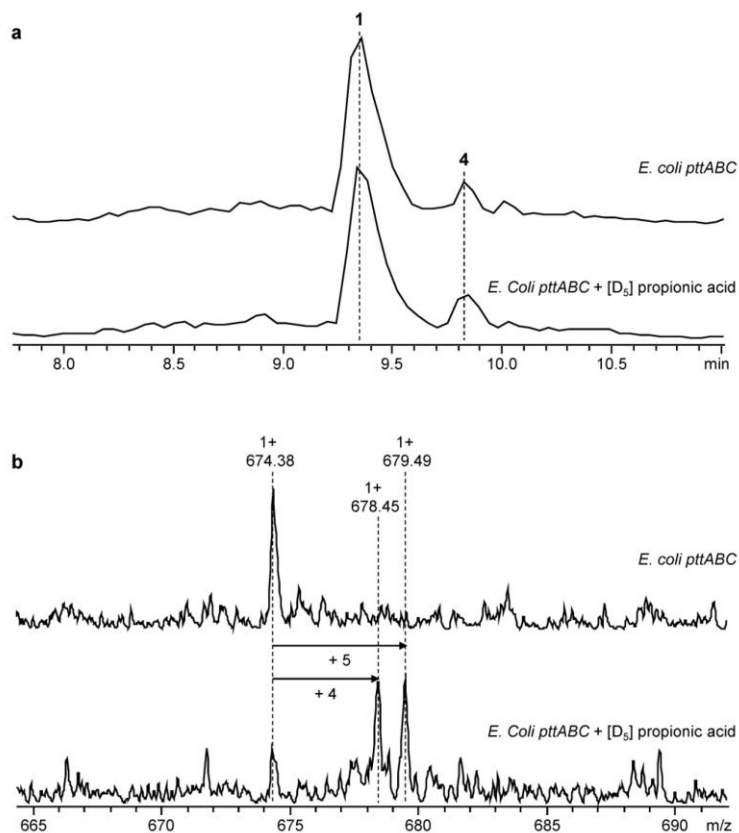
Figure S3. HPLC-MS analysis of XAD-16 extracts from labeling experiments with or without [D5] propionic acid fed to *E. coli* expressing *pttABC* in LB medium. (a) Base-peak chromatograms (BPCs) for **1** and **4**; (b) Mass shifts for **4**.
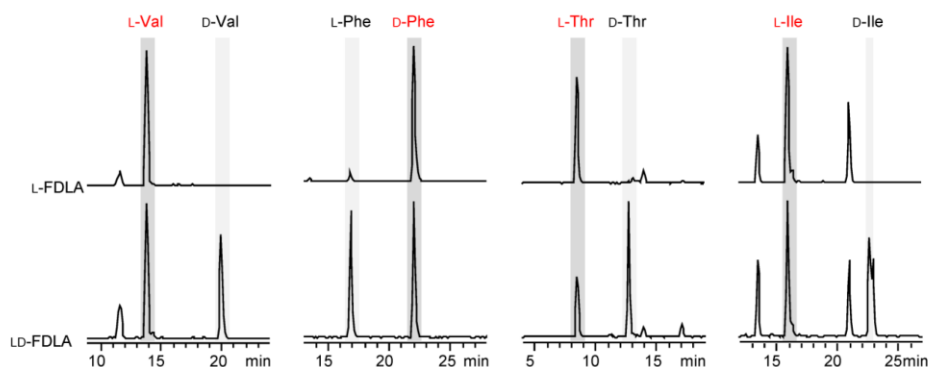
S15

Figure S4. Configuration determination of amino acids in **1** using the advanced Marfey's method. HPLC-MS analysis of hydrolyzed **1** and derivatized with L-FDLA (upper) and LD-FDLA (lower). Depicted are EIC traces for valine (Val, *m/z* 412 [M + H]⁺), phenylalanine (Phe, *m/z* 460 [M + H]⁺), threonine (Thr, *m/z* 414 [M + H]⁺), and isoleucine (Ile, *m/z* 426 [M + H]⁺). The configurations are determined by the elution order and L-FDLA derivatized amino acids elute prior to its D-enantiomer.

S16

260
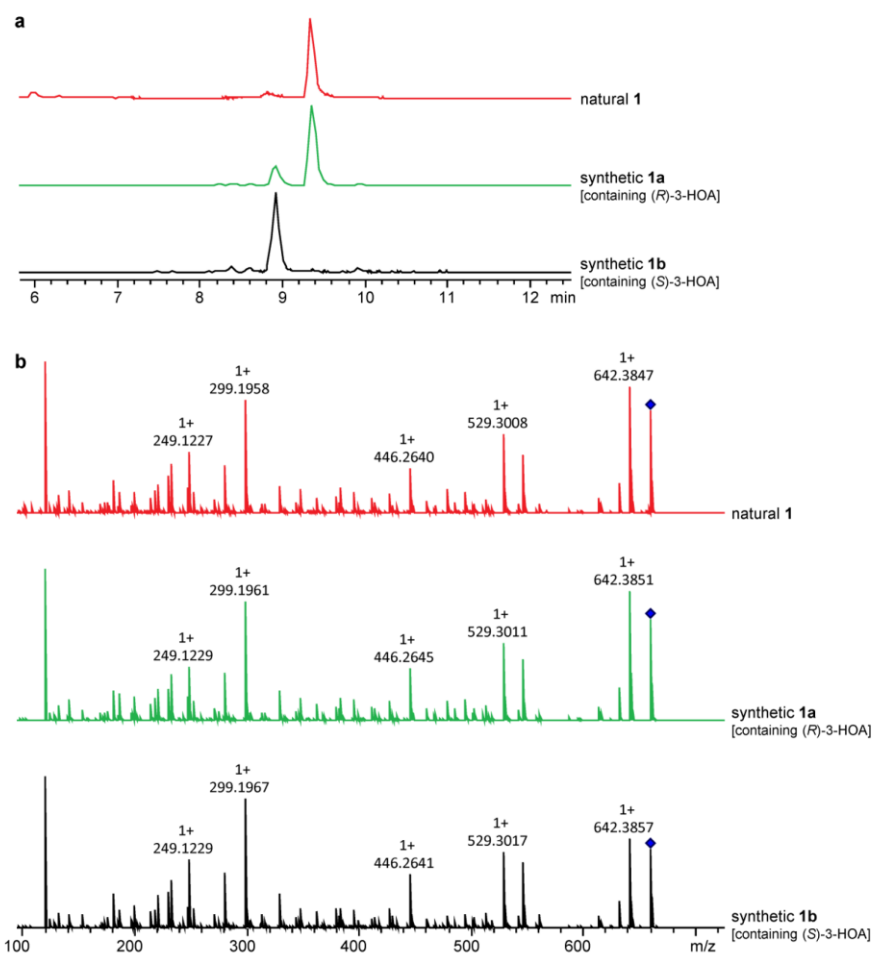
Figure S5. Comparison of natural **1** and synthetic **1a** and **1b** by HPLC-MS analysis. (a) EICs for natural **1** and synthetic **1a** and **1b**; (b) MS^2 fragmentation patterns for natural **1** and synthetic **1a** and **1b**.
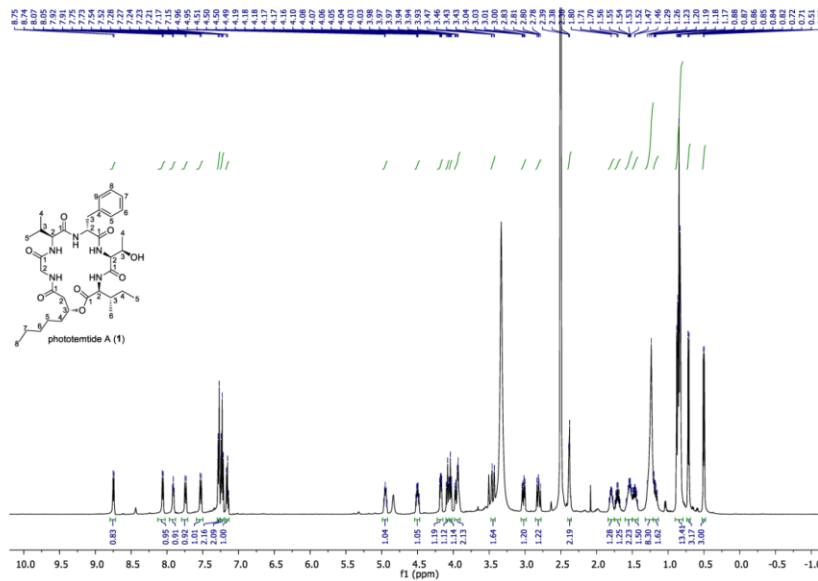
S17

261

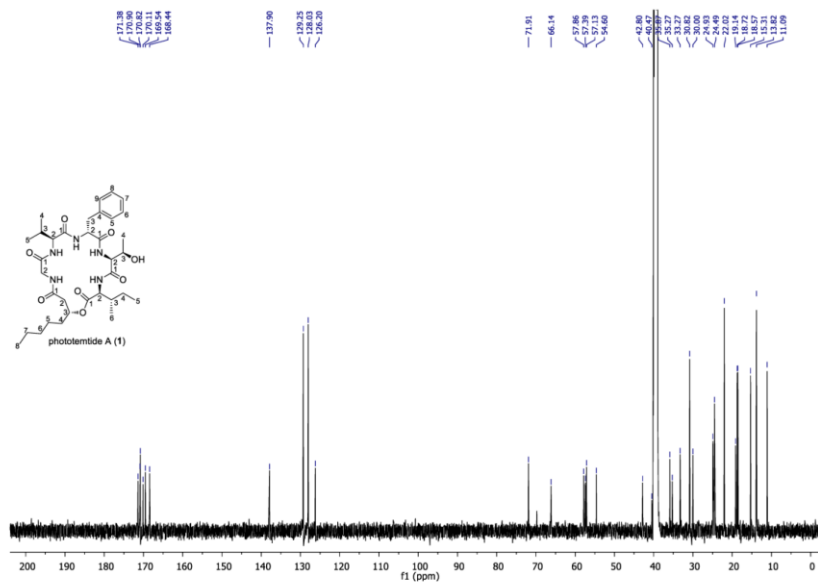**Figure S6.** ^{1}H NMR (500 MHz, DMSO-d_6) spectrum of nature **1**.

Figure S7. ^{13}C NMR (125 MHz, DMSO-d_6) spectrum of natural **1**.

S18

262
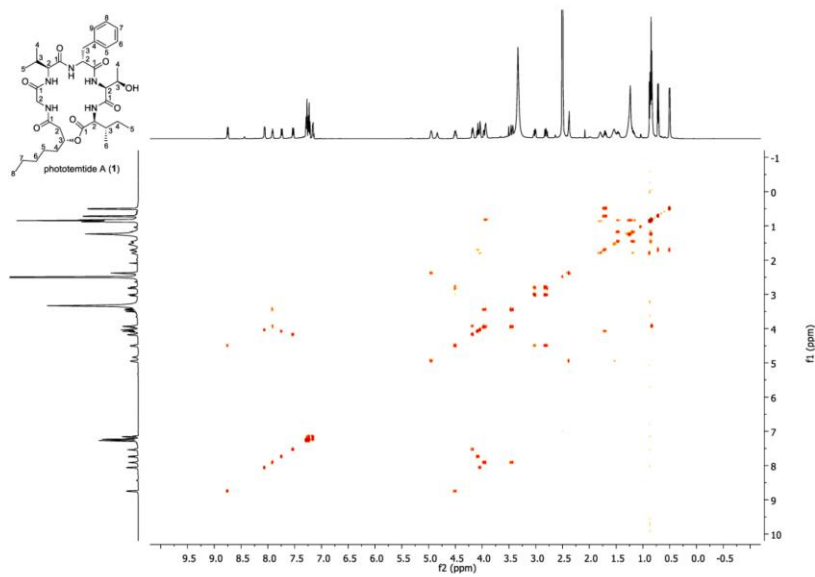
Figure S8. COSY (DMSO-*d*6) spectrum of natural **1**.

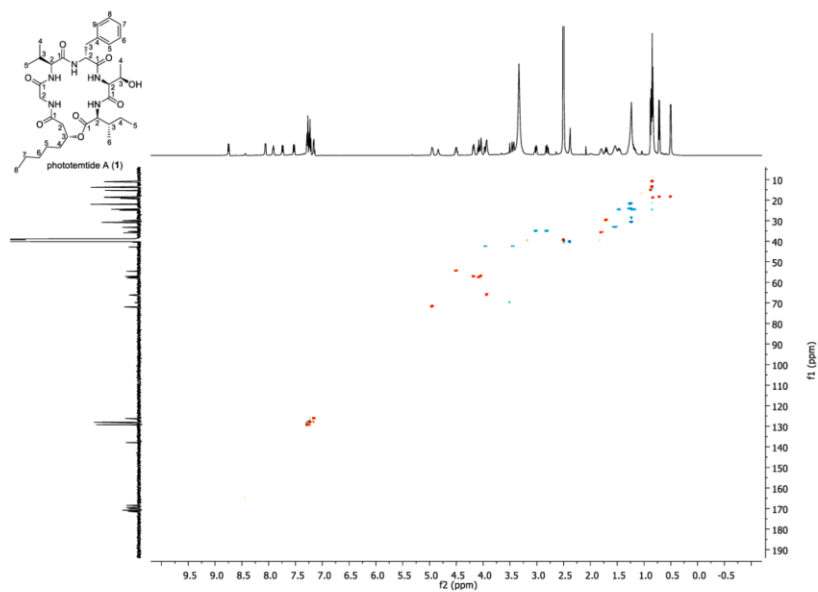**Figure S9.** HMBC (DMSO-*d*6) spectrum of natural **1**.

S19

263

Figure S10. HSQC (DMSO-d_6) spectrum of natural **1**.

S20

264
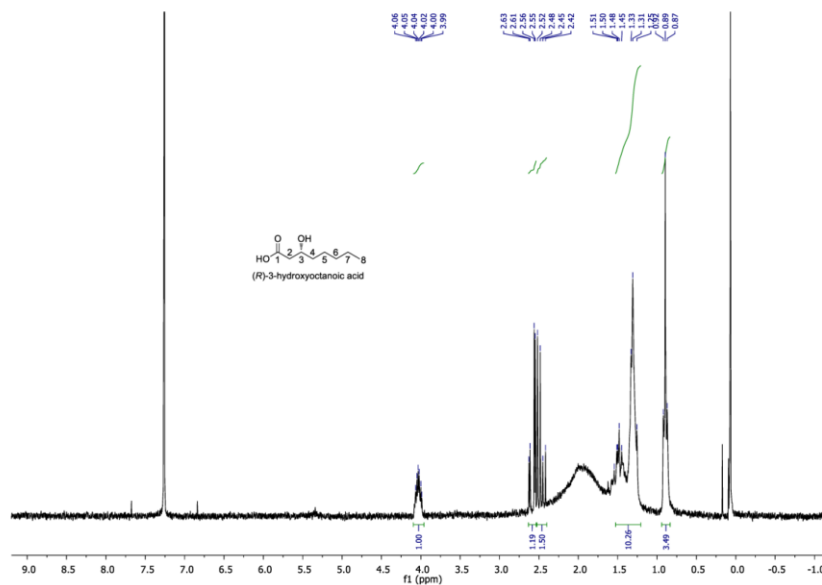
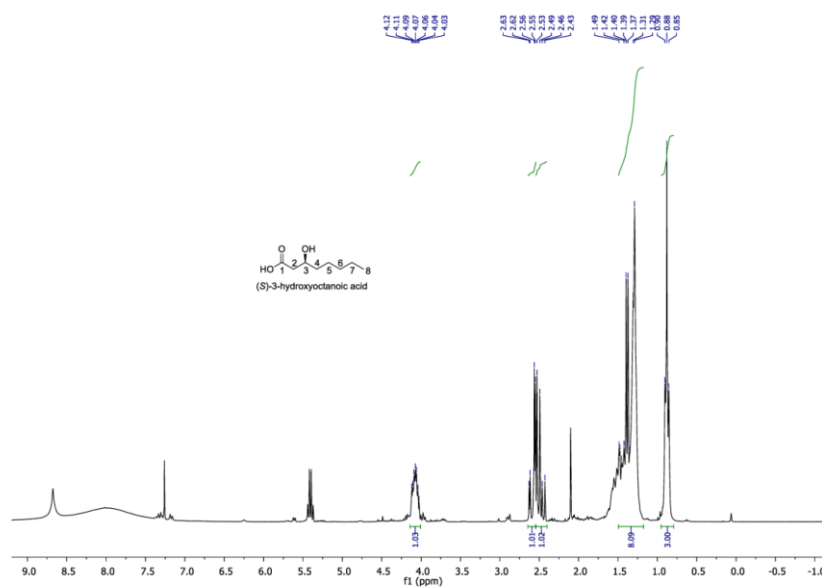**Figure S11.** ^{1}H NMR (250 MHz, CDCl$_3$) spectrum of (*R*)-3-hydroxyoctanoic acid.

Figure S12. ^{1}H NMR (250 MHz, CDCl$_3$) spectrum of (*S*)-3-hydroxyoctanoic acid.
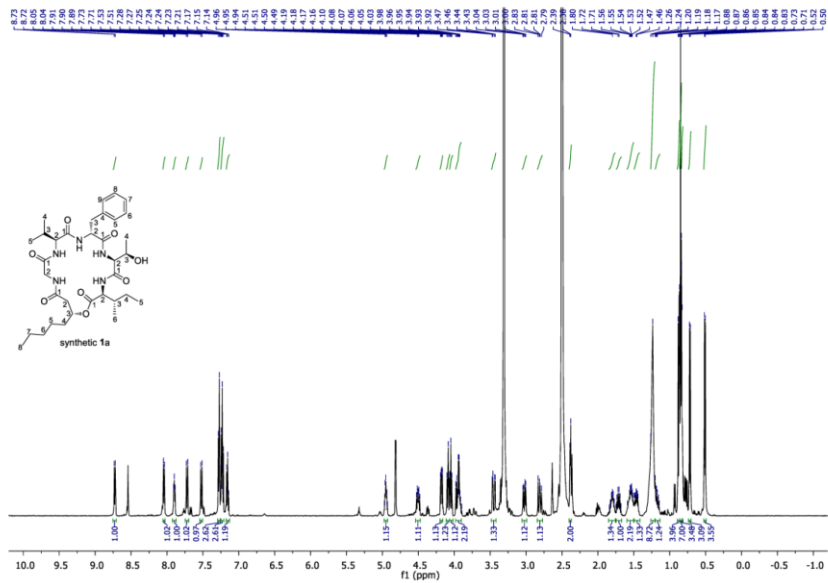
S21

265

Figure S13. ^{1}H NMR (500 MHz, DMSO-d_6) spectrum of synthetic **1a**.
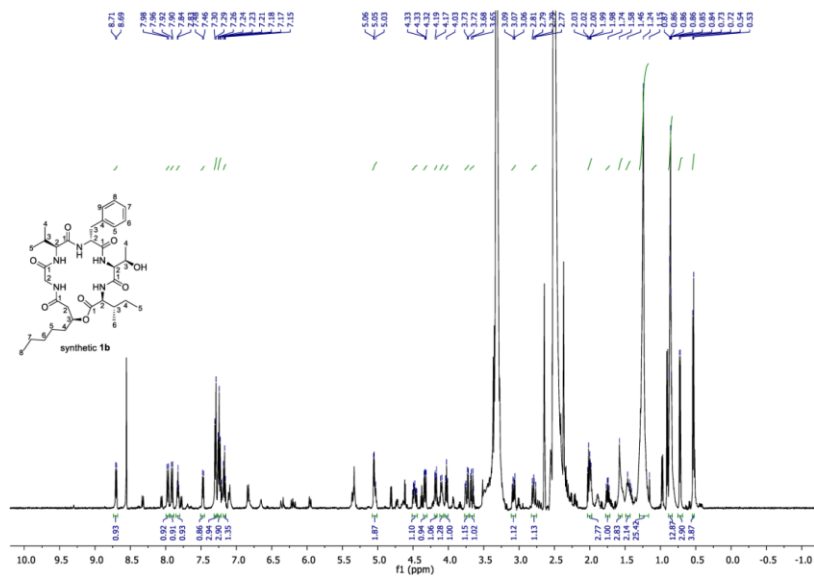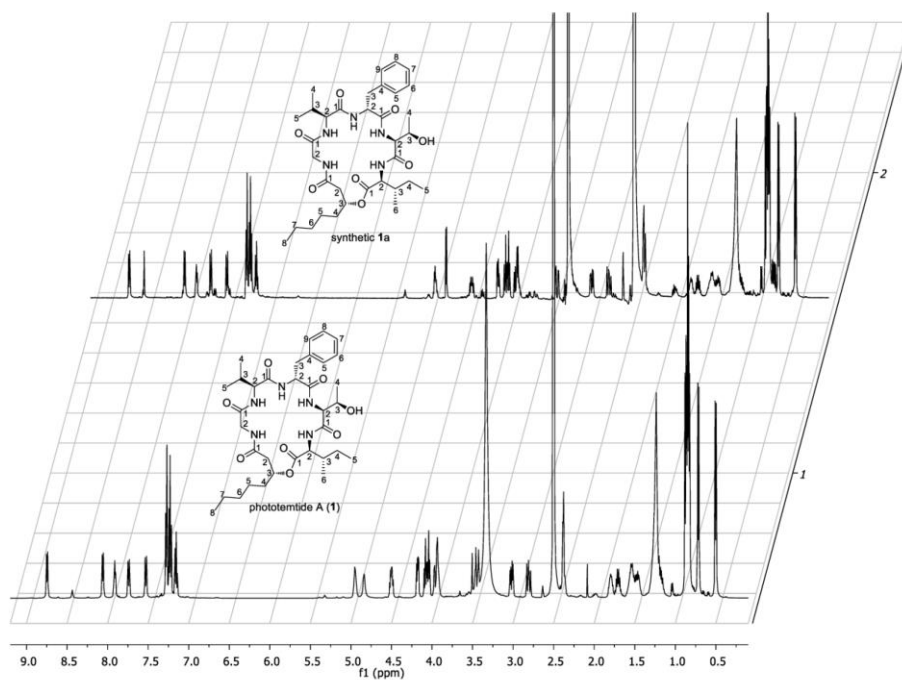

Figure S14. ^{1}H NMR (500 MHz, DMSO-d_6) spectrum of synthetic **1b**.

Figure S15. Comparison of ^{1}H NMR (500 MHz, DMSO-d_6) spectra of natural **1** (bottom) and synthetic **1a** (top).

S23

References

(1) Schimming, O.; Fleischhacker, F.; Nollmann, F. I.; Bode, H. B. *Chembiochem* **2014**, *15*, 1290.

(2) Grundmann, F.; Kaiser, M.; Kurz, M.; Schiell, M.; Batzer, A.; Bode, H. B. *RSC Adv.* **2013**, *3*, 22072.

(3) Fujii, K.; Ikai, Y.; Oka, H.; Suzuki, M.; Harada, K. *Anal. Chem.* **1997**, *69*, 5146.

(4) Fukuzawa, S.; Matsuzawa, H.; Yoshimitsu, S. *J. Org. Chem.* **2000**, *65*, 1702.

(5) Hung, K.; Harris, P. W. R.; Heapy, A. M.; Brimble, M. A. *Org. Biomol. Chem.* **2011**, *9*, 236.

(6) Nollmann, F. I.; Dowling, A.; Kaiser, M.; Deckmann, K.; Grösch, S.; Ffrench-Constant, R.; Bode, H. B. *Beilstein J. Org. Chem.* **2012**, *8*, 528.

(7) Orhan, I.; Şener, B.; Kaiser, M.; Brun, R.; Tasdemir, D. *Mar. Drugs* **2010**, *8*, 47.

(8) Hurst IV, S. G.; Ghazal, S.; Morris, K.; Abebe-Akele, F.; Thomas, W. K.; Badr, U. M.; Hussein, M. A.; AbouZaied, M. A.; Khalil, K. M.; Tisa, L. S. *Genome Announc.* **2014**, *2*, e01273-14.

(9) Bozhüyük, K. A. J.; Fleischhacker, F.; Linck, A.; Wesche, F.; Tietze, A.; Niesert, C.-P.; Bode, H. B. *Nat. Chem.* **2018**, *10*, 275.

6 Curriculum Vitae

Personal Data

Lei Zhao

Date of Birth: 10.02.1989

Place of Birth: Shaanxi

Nationality: Chinese

E-Mail: l.zhao89@hotmail.com

Professional Experience

Since 01/2016	**PhD Student** Group of Prof. Dr. Helge B. Bode Goethe University Frankfurt, Germany
Since 09/2018	**Research Associate** Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, China
09/2015−08/2018	**Research Assistant** Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, China

Education

Since 01/2016	**Doctoral Candidate** Group of Prof. Dr. Helge B. Bode Goethe University Frankfurt, Germany
09/2012−07/2015	**Master of Science (Botany)** Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, China
09/2008−07/2012	**Bachelor of Science (Biotechnology)** Northwest University, China

7 List of Publications

Publications which are part of this thesis

<u>Lei Zhao</u>, Marcel Kaiser, and Helge B. Bode. Rhabdopeptide/Xenortide-like Peptides from *Xenorhabdus innexi* with Terminal Amines Showing Potent Antiprotozoal Activity. *Organic Letters* **2018**, *20*, 5116−5120.

<u>Lei Zhao</u>, Xiaofeng Cai, Marcel Kaiser, and Helge B. Bode. Methionine-Containing Rhabdopeptide/Xenortide-like Peptides from Heterologous Expression of the Biosynthetic Gene Cluster *kj12ABC* in *Escherichia coli. Journal of Natural Products* **2018**, *81*, 2292−2295.

<u>Lei Zhao</u> and Helge B. Bode. Production of a Photohexapeptide Library from Entomopathogenic *Photorhabdus asymbiotica* PB68.1. *Organic & Biomolecular Chemistry* **2019**, *17*, 7858−7862.

<u>Lei Zhao</u>, Ryan Musumba Awori, Marcel Kaiser, Jonathan Groß, Till Opatz, and Helge B. Bode. Structure, Biosynthesis, and Bioactivity of Photoditritide from *Photorhabdus temperata* Meg1. *Journal of Natural Products* **2019**, *82*, 3499−3503.

<u>Lei Zhao</u>, Tien Duy Vo, Marcel Kaiser, and Helge B. Bode. Phototemtide A, a Cyclic Lipopeptide Heterologously Expressed from *Photorhabdus temperata* Meg1, Shows Selective Antiprotozoal Activity. *ChemBioChem* **2020**, *21*, 1−6.

Publications which are not part of this thesis

Xiaofeng Cai, Victoria L. Challinor, <u>Lei Zhao</u>, Daniela Reimer, Hélène Adihou, Peter Grün, Marcel Kaiser, and Helge B. Bode. Biosynthesis of the Antibiotic Nematophin and Its Elongated Derivatives in Entomopathogenic Bacteria. *Organic Letters* **2017**, *19*, 806−809.

Carolin Hacker, Xiaofeng Cai, Carsten Kegler, <u>Lei Zhao</u>, Anna Katharina Weickhmann, Jan Philip Wurm, and Helge B. Bode, Jens Wöhnert. Structure-Based Redesign of Docking Domain Interactions Modulates the Product Spectrum of a Rhabdopeptide-Synthesizing NRPS. *Nature Communications* **2018**, *9*, 4366.

Xiaofeng Cai, <u>Lei Zhao</u>, and Helge B. Bode. Reprogramming Promiscuous Nonribosomal Peptide Synthetases for Production of Specific Peptides. *Organic Letters* **2019**, *21*, 2116−2120.

8 Record of Conferences

VAAM Workshop on the Biology of Bacteria Producing Natural Products

Date: 31.08.−02.09.2018

Place: Frankfurt am Main, Germany

Contribution: Poster presentation

Title: Structure and Biosynthesis of Rhabdopeptide/Xenortide-like Peptides in *Xenorhabdus innexi* and *Xenorhasbdus cabanillasii*

VAAM Workshop on the Biology of Microorganisms Producing Natural Products

Date: 15.−17.09.2019

Place: Jena, Germany

Contribution: Poster presentation

Title: Structure Elucidation and Biosynthesis of Nonribosomally Derived Peptides from Entomopathogenic *Photorhabdus* strains

9 Acknowledgements

At the end of this thesis, I would like to thank all the people who were involved in this work.

First of all, I would like to thank my supervisor Prof. Dr. Helge B. Bode for giving me the opportunity to work on a quite diverse topic in Germany. I really appreciate it that he has supervised, motivated, inspired and supported me with great patience during my entire doctoral period.

I thank Prof. Dr. Martin Grininger, my second reviewer, who was willing to survey this work without hesitation. Thanks a lot for his time.

I would particularly like to thank Dr. Edna Bode and Gina Grammbitter for proofreading this work. Their constructive comments and elimination of mistakes were very helpful to improve this work. I also thank Dr. Edna Bode for translating the summary of this thesis into German.

I especially thank Peter Grün for helping me with my initial laboratory work, Dr. Xiaofeng Cai for teaching me the methods of molecular biology, and Dr. Frank Wesche for bringing me into the field of solid-phase peptide synthesis.

Furthermore, I would like to thank all the co-authors of the publications for the good cooperation.

I would also like to thank the MS teams for measuring my samples.

A special thanks go to my office 0.02 for the pleasant company in the past four years.

I also thank the whole working group with all current and former members for the constructive cooperation and the friendly working atmosphere.

Last but not least, I would like to thank my family for their unconditional support on the way to the completion of this work, whose trust and love always strengthen me.

10 Erklärung und Versicherung

Erklärung

Ich erkläre hiermit, dass ich mich bisher keiner Doktorprüfung im Mathematisch-Naturwissenschaftlichen Bereich unterzogen habe.

Frankfurt am Main, den   Lei Zhao

Versicherung

Ich erkläre hiermit, dass ich die vorgelegte Dissertation über

"Nonribosomal Peptides Produced by *Xenorhabdus* and *Photorhabdus*"

selbstständig angefertigt und mich anderer Hilfsmittel als der in ihr angegebenen nicht bedient habe, insbesondere, dass alle Entlehnungen aus anderen Schriften mit Angabe der betreffenden Schrift gekennzeichnet sind.

Ich versichere, die Grundsätze der guten wissenschaftlichen Praxis beachtet, und nicht die Hilfe einer kommerziellen Promotionsvermittlung in Anspruch genommen zu haben.

Frankfurt am Main, den   Lei Zhao