Identification of selected secondary metabolites from *Xenorhabdus* and investigation on the biosynthesis of anthraquinones from *Photorhabdus*

Dissertation

zur Erlangung des Doktorgrades

der Naturwissenschaften

vorgelegt beim Fachbereich für Biowissenschaften (15)

der Johann Wolfgang Goethe – Universität

in Frankfurt am Main

von

Qiuqin Zhou

aus Wenzhou (China)

Frankfurt 2016

(D30)

vom Fachbereich für Biowissenschaften (15) der

Johann Wolfgang Goethe - Universität als Dissertation angenommen.

Table of contents

1. Danksagung

2. Abstract

Xenorhabdus and *Photorhabdus* bacteria are gaining more and more attention as a subject of research because of their unique yet similar life cycle with nematodes and insects. This work focused on the secondary metabolites that are produced by *Xenorhabdus* and *Photorhabdus*. With the help of modern HPLC-MS methodologies and increasingly available bacterial genome sequences, the structures of unknown secondary metabolites could be elucidated and thus their biosynthesis pathways could be proposed, too.

The first paper reported 17 depsipeptides termed xentrivalpeptides produced by the bacterium *Xenorhabdus* sp. 85816. Xentrivalpeptide A could be isolated from the bacterial culture as the main component. The structure of xentrivalpeptide A was elucidated by NMR and the Marfey´s method. The remaining xentrivalpeptides were exclusively identified by feeding experiments and MS fragmentation patterns.

The second paper described the discovery and isolation of xenoamicin A from *Xenorhabdus mauleonii* DSM17908. Additionally, other xenoamicin derivatives from *Xenorhabdus doucetiae* DSM17909 were analyzed by means of feeding experiments and MS fragmentation patterns. The xenoamicin biosynthesis gene cluster was identified in *Xenorhabdus doucetiae* DSM17909.

The manuscript for publication focused on the biosynthesis of anthraquinones in *Photorhabdus luminescens.* The Type II polyketide synthase for the biosynthesis of anthraquinone derivatives was discovered in *P. luminescens* in a previous publication by the Bode group,[1] in which a partial reaction mechanism for the biosynthesis has been proposed. The manuscript reported in this thesis however elucidated the biosynthetic mechanisms in a greater detail as compared to the previous publication. Particularly, the biosynthetic mechanism was deciphered through heterologous expression of anthraquinone biosynthesis (*ant*) genes in *E. coli*. Additionally, deactivation of the genes *antG* encoding a putative CoA ligase and *antI* encoding a putative hydrolase, was performed in *P. luminescens*. Selected *ant* genes were over-expressed in *E. coli* as well as the corresponding proteins purified for *in vitro* assays. Model compounds were chemically synthesized as possible substrates of AntI and were used for *in vitro* assays. Here, it was revealed that the CoA ligase AntG played an essential role in the activation of the ACP AntF. Furthermore, a chain shortening

mechanism by the hydrolase AntI was identified and was further confirmed by *in vitro* assays using model compounds. Additionally, this chain shortening mechanism was supported by homology based structural modeling of AntI.

3. Zusammenfassung

Bakterien gehören zu den ältesten Lebewesen dieser Erde. Sie sind Einzeller und haben keinen Zellkern. Häufig können sie Kolonien bilden und in diesen zusammenleben. Mikroskopisch klein sind die meisten Bakterien und man findet sie überall. Manche Bakterien können bei Menschen gesundheitlich bedrohliche Krankheiten verursachen, wie zum Beispiel Tuberkulose. Andere Bakterien leben mit den menschlichen Zellen friedlich zusammen und profitieren gegenseitig voneinander, zum Beispiel auf der Haut, im Darm, auf der Zunge und sogar in der Lunge. Das Gleichgewicht zwischen den verschiedenen an Menschen angesiedelten Bakterien bestimmt nicht nur den Gesundheitszustand sondern auch die Stimmung der Menschen. Genauso ist es auch zwischen manchen Bakterien und anderen Lebewesen. Die Vielfalt der Bakterien macht es für Forscher interessant, sie besser zu verstehen und gleichzeitig bietet Forschung auch viele Möglichkeiten, Bakterien für uns Menschen nützlich zu machen.

Bakterien der Gattungen *Xenorhabdus* und *Photorhabdus* weisen einen einzigartigen Lebenszyklus auf, in dem sie komplexe Interaktionen mit Nematoden und Insekten eingehen. Sie leben jeweils symbiotisch in insektenpathogenen Nematoden *Steinernema* und *Heterorhabditis* bis auf der Ausnahme des Bakteriums *Photorhabdus asymbiotica*. Zusammen mit Nematoden sind die Bakterien *Xenorhabdus* und *Photorhabdus* pathogen gegen Insektenlarven. Sobald die Nematoden in Insektenlarven eindringen, treten Bakterien aus den Nematoden aus. In Insektenlarven fangen die Bakterien an sich zu vermehren und viele Sekundärmetaboliten zu produzieren. Zusammen mit anderen Toxinen und Exoenzymen sind die Sekundärmetaboliten in der Lage, Insektenlarven zu töten. Gleichzeitig unterstützen sie das Wachstum und die Vermehrung von Nematoden bis die Insektenlarven aufgegessen sind.

Diese kumulative Dissertation konzentriert sich auf die Sekundärmetaboliten, die von *Xenorhabdus* und *Photorhabdus* produziert werden. Genauer gesagt sind 17 Xentrivalpeptide von *Xenorhabdus* sp. 85816 (die erste Veröffentlichung), Xenoamicinpeptide von *Xenorhabdus mauleonii* DSM17908 und *Xenorhabdus doucetiae* DSM17909 (die zweite Veröffentlichung) und Anthrachinonderivaten (das Manuskript) von *Photorhabdus luminescens*.

Die nichtribosomale Peptidsynthase ist für die Biosynthese der Xentrivalpeptide und Xenoamicinpeptide von *Xenorhabdus* verantwortlich. Als ein riesiger Multienzymkomplex ist die nichtribosomale Peptidsynthase modular aufgebaut und nimmt die Rolle als spezifische Proteinmatrize ein, die sowohl den spezifischen Einbau der bestimmten Aminosäure kontrollieren als auch das entstandene Peptid modifizieren kann. Jeder Verlängerungsschritt beginnt mit der spezifischen Erkennung und Aktivierung einer Aminosäure durch die Adenylierungsdomäne. Anschließend wird die aktivierte Aminosäure durch einen nukleophilen Angriff auf die PCP-Domäne (*peptide carrier protein*) übertragen. Die Verlängerung des Peptides geschieht durch die Kondensation von freiem Amin der wachsenden Peptidkette mit der aktivierten Aminosäure. In manchen Fällen ist der Einbau von D-Aminosäure möglich. Sobald die Peptidkette fertig aufgebaut ist, wird sie in den meisten Fällen von der C-terminalen Thioesterase-Domäne abgespaltet. Da *Xenorhabdus* ein breites Band von Derivaten der untersuchten nichtribosomalen Peptide produziert, liegt der Schwerpunkt dieser Arbeit auf der Aufklärung von Strukturen.

Im Gegensatz dazu besteht die Arbeit über Anthrachinonderivate fast nur aus der Untersuchung und Aufklärung des Biosynthesewegs. Anthrachinonderivate gehören zu den aromatischen Polyketiden. Es ist bekannt, dass Anthrachinonderivate in Pflanzen durch Typ III Polyketidsynthase und in Bakterien und Pilzen durch Typ I / II Polyketidsynthase synthetisiert werden. In *Photorhabdus* sind Anthrachinonderivate von der klassischen Typ II Polyketidsynthase synthetisiert, die aber in Gram-Negativ Bakterien sehr selten vorkommt. Die Typ II Polyketidsynthase, die fast nur in *Streptomycetes* vorkommt, besteht aus mehreren individuellen strukturell und funktionell verschiedenen Enzymen. Die Enzyme bilden einen nicht kovalent gebunden Multienzymkomplex. Während der Verlängerung der Polyketidkette ist eine iterative Benutzung einiger Enzyme vorgesehen. In den meisten Fällen besteht das Kernstück der Typ II Polyketidsynthase aus KS_α, KS_β (Ketosynthase) und ACP (*acyl carrier protein*). Dazu kommen noch die Cyclase, Aromatase, Ketoreduktase und weitere Enzyme für die Modifizierung. Weil die Strukturen von Anthrachinonderivaten in *Photorhabdus* bereits durch frühere Arbeiten aufgeklärt wurden und die Typ II Polyketidsynthase durch die Arbeit von Brachmann et al. entdeckt und teilweise aufgeklärt wurde, ist es sinnvoll, dem Ziel zur Aufklärung des Biosynthesewegs von Anthrachinonderivaten in *Photorhabdus* zu folgen.

3. Zusammenfassung

In den ersten beiden Veröffentlichungen sind die verwendeten Methoden und die Technik zur Strukturaufklärung fast identisch. Weil jeweils ein Derivat als Hauptkomponente von *Xenorhabdus* produziert wurde, konnte die Hauptkomponente durch die semi-präparative HPLC isoliert werden. Für die Isolierung der Hauptkomponente wurden Bakterien zusammen mit Amberlite® XAD 16 kultiviert. Die hydrophoben Sekundärmetaboliten wurden in Bakterienkulturen von Amberlite® XAD-16 absorbiert und nach dem Sammeln von Amberlite® XAD-16 wieder mit Methanol aufgelöst. Das Konzentrat von Sekundärmetaboliten wurde durch Kieselgel vorläufig fraktioniert. Die Strukturaufklärung der Hauptkomponente erfolgte mit der NMR Spektroskopie und der Marfey-Methode, eine Methode zur Identifizierung von Stereochemie der eingebauten Aminosäuren. Die Marfey-Methode besteht aus den drei folgenden Schritten: Abspaltung der einzelnen Aminosäuren mit Salzsäure, Derivatisierung der frei gewordenen Aminosäure mit zwei Enantionmeren und der HPLC-Analyse der gebildeten Diastereomeren. In der Regel werden kommerziell verfügbare Aminosäuren zur Kontrolle mit der Marfey-Methode behandelt und analysiert. Die Strukturen der restlichen Peptide, die meistens nur in sehr geringer Menge produziert wurden, wurden ausschließlich mit Fütterungsexperimenten und Fragmentierungsmustern von massenspektrometrischen Experimenten aufgeklärt. Das stabile Isotop von Kohlenstoff ^{13}C, Stickstoff ^{15}N und Wasserstoff ^{2}H (auch als Deuterium bezeichnet) wurden gezielt in die Peptide eingebaut, indem entweder das Nährmedium das stabile Isotop enthielt oder die Substrate mit dem stabilen Isotop als Überschuss für die Biosynthese zur Verfügung gestellt wurden. Zusätzlich zum Ergebnis der hochauflösenden Massenspektrometrie, konnten die Zahlen von Kohlenstoffatomen und Stickstoffatomen der Peptide durch die Massenverschiebung nach dem Einbau von ^{13}C und ^{15}N bestimmt werden. Die deuterierten Aminosäuren wurden für die Feststellung der eingebauten Aminosäure verwendet. Die Peptide konnten durch MS^{n} sequenziert werden. Mit der Kombination von Fütterungsexperimenten und MS^{n}-Sequenzierung konnten die Strukturen der restlichen Peptide ohne aufwendige Isolierung lückenlos aufgeklärt werden. Weil das Genom von *Xenorhabdus doucetiae* DSM17909 schon bekannt war, konnte das Biosynthesegencluster identifiziert werden. Nachdem das Gencluster durch Geninsertion inaktiviert wurde, konnten die Xenoamicinpeptide nicht mehr produziert werden.

Ein in dieser Doktorarbeit erarbeitetes drittes Manuskript konzentriert sich auf die Biosynthese von Anthrachinonderivaten in *Photorhabdus luminescens*. In einer ersten Veröffentlichung der Bode Gruppe zur Biosynthese von Anthrachinonderivaten in *Photorhabdus luminescens*, die einen Mechanismus der Typ II Polyketid Biosynthese offenlegte, haben Brachmann *et al.* bereits einen unvollständigen Reaktionsmechanismus für die Biosynthese vorgeschlagen. Dieser Vorschlag wurde in der hier vorgelegten Arbeit experimentell molekularbiologisch, biochemisch und bioinformatisch untersucht und auf Grund der erhaltenen neuen Erkenntnisse modifiziert.

Wegen des hohen GC-Anteils vom Gencluster aus *Streptomyces*, war es nicht möglich, den Biosyntheseweg von Typ II Polyketidsynthase aus *Streptomyces* in *E. coli* heterolog zu rekonstruieren. Hier konnte zum ersten Mal gezeigt werden, dass Anthrachinon AQ-256 durch die Expression von (*ant*) Genecluster verteilt auf zwei Plasmide in *E. coli* gebildet wurde. Die Expressionsexperimente erlaubten uns eine schnelle Untersuchung von Funktion der einzelnen *ant* Gene. Mit der Unterstützung von moderner Gensynthese des codon-optimierten Gens *actI-ORF1* (kodiert Ketoreduktase RED1 aus Actinorhodin-Biosynthese) konnte ein bekannter Shunt-Produkt 4-dihydro-9-hydroxy-1-methyl-10-oxo-3-H-naptho-[2,3-c]-pyran-3-(*S*)-aceticacid (*S*-DNPA) von Actinorhodin-Biosynthese in *E. coli* produziert werden.

Weiterhin wurden Knockout-Mutanten von *Photorhabdus luminescens* durch die in-frame-Insertion oder in-frame-Deletion generiert. Die Deaktivierung vom Gen *antG* (kodiert mutmaßlich ein CoA Ligase), *antI* (kodiert mutmaßlich ein Hydrolase) oder antC (kodiert ein Cyclase) führte zum Ausbleiben von Anthrachinon AQ-256 in *Photorhabdus luminescens*. Stattdessen konnten sowohl neue als auch bekannte Shunt-Produkte identifiziert werden. Knockout-Mutanten bot uns eine zuverlässige Aussage über die Funktion des deaktivierten Gens.

Zusätzlich zu den in-vivo-Experimenten wurde versucht, alle einzelnen Ant Proteine für in-vitro-Experimente in *E. coli* zu überproduzieren und zu reinigen. Außer Phosphopantetheinyl transferase AntB, KS_α und KS_β (Ketosynthase) konnten alle Proteine einzeln löslich in *E. coli* exprimiert und mit der Ni-Affinitätschromatographie gereinigt werden. Das Problem mit KS_α und KS_β konnte durch eine Co-expression und Co-reinigung gelöst werden. Bisherige Arbeit zeigt dass AntB nicht durch andere

PPTase ersetzt werden konnte, wird die zukünftige Arbeit mit dem Protein AntB eine sehr spannende Herausforderung.

Mit den gereinigten Proteinen konnten die in-vitro-Experimente teilweise durchgeführt werden. Die Oktaketide (SEK4 und SEK4b) konnten mit Hilfe von Referenzsubstanzen identifiziert werden. Modellsubstanzen wurden als mögliche Substrate für Hydrolase AntI chemisch synthetisiert. Die in-vitro-Experimente von Hydrolase mit den Modellsubstanzen führten zur Postulierung des detaillierten Kettenabkürzungsmechanismus vom Oktaketid zum Heptaketid.

In diesem Manuskript wurde auch die ungewöhnliche Aktivierung von ACP AntF im Detail untersucht. Die Aktivierung war nur bei der Anwesenheit von Phosphopantetheinyl Transferase AntB und mutmaßlich CoA ligase AntG erfolgreich. Mit Hilfe von UTL-MALDI-TOF-MS (*ultra thin layer - matrix-assisted laser desorption ionization - time of flight mass spectrometer*) konnte das einfachgeladene *apo*-ACP mit *holo*-ACP auseinanderhalten werden. Nur im Fall mit dem aktivierten ACP (*holo*-ACP), konnten die Polyketide (SEK4 und SEK4b) in HPLC detektiert werden. Normalerweise ist die Phosphopantetheinyl Transferase allein in der Lage, ein ACP zu aktivieren. Die Phosphopantetheinyl Transferase mit breiten Aktivitäten (Sfp aus *Bacillus subtilis* oder MtaA aus *Stigmatella aurantiaca*) wurde häufig in vielen anderen Laboren für die Aktivierung von ACP verwendet. Aber Sfp oder MatA konnte ACP AntF nicht aktivieren. Wegen der vorhandenen katalytischen Domäne zur Bildung von Adenylat, wurde AntG als mutmaßlich CoA Ligase identifiziert. Die in-vitro-Experimente zeigten, dass AntG die Malonsäure oder Essigsäure mit CoA nicht ligieren konnte. Deswegen sind weitere Untersuchungen nötig, um die genaue Funktion von AntG aufzuklären.

Die katalytische Funktion von Hydrolase AntI wurde anhand der homologiebasierten Modellierung und des Docking der Modellsubstanz untersucht. Hierzu konnte die 3D-Struktur von AntI anhand DHPON Hydrolase aus *Arthrobacter nicotinovorans* dargestellt werden. Das Docking der Modellsubstanz an AntI wurde berechnet. Die resultierten katalytischen Triaden wurden mit Punktmutation analysiert. Die bioinformatischen Analysen mit den ungewöhnlichen Ketosynthasen AntD und AntE führten zur Identifizierung anderer Ant-Genclusteranalogen in anderen *Photorhabdus* Stämmen und anderen Bakterien.

4. Introduction

4.1. *Photorhabdus* and *Xenorhabdus*

4.1.1. Entomopathogenic bacteria: *Photorhabdus* and *Xenorhabdus*

Since October 2011, the human population on earth has reported to exceed over 7 billion.[2] The continual increase of the human population results in a higher demand for crops. In order to reduce yield losses and produce crops efficiently, the protection of crops from diseases plays an important role. Pesticides were widely used in agriculture but traditional pesticides are often chemically synthesized and thus are not easily degraded in soil. Accumulated pesticides in our environment are a great danger to the health of human and other animals.[3] The usage of entomopathogenic nematodes as biological control agents against pests, which can be killed more specifically, is an environmentally friendly option.[4]
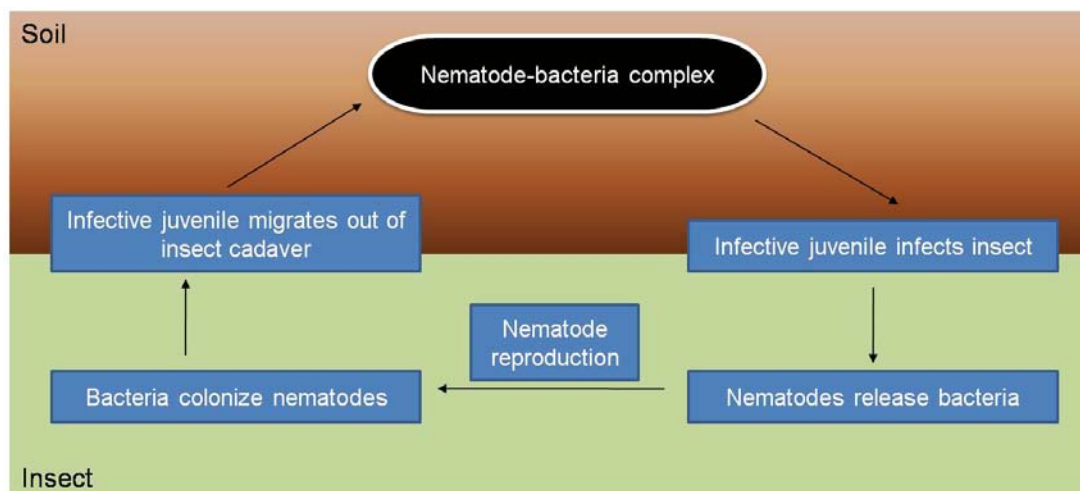
Figure 1. The lifecycle of nematode-bacteria complex (adapted from Goodrich Blair *et al.*[6]).

Bacteria of the genera *Xenorhabdus* and *Photorhabdus*, with the exception of *Photorhabdus asymbiotica,* could only be found in symbiosis with nematodes from *Steinernema spp.* and *Heterorhabditis spp.*, respectively. They live in the guts of the free-living or non-feeding form of the nematode called infective juvenile, which is also called nematode-bacteria complex (see Figure 1). The nematodes with bacteria in their guts can penetrate into the insect larvae via their mouth or anus, then migrate to the haemolymph and release the bacteria. Inside the insect larvae, the bacteria begin to proliferate and synthesize a wide range of secondary metabolites, which are described in a later chapter. The secondary metabolites together with other toxins

and hydrolytic exoenzymes are able to kill the insect larvae, but at same time they can support the nematode growth and reproduction until the insect larvae are eaten up. Consequently, nematodes are recolonized by the bacteria, leave the empty cadaver and begin to seek for a new prey.[5]

Xenorhabdus and *Photorhabdus* bacteria[7] are rod-shaped (in size ranging of 0.3-2 x 2-10 µm and 0.5-2 x 1-10 µm, respectively), Gram-negative and belong to the family Enterobacteriaceae. Their optimum reproduction temperature usually is 28 °C. *Xenorhabdus* bacteria are catalase negative, while *Photorhabdus* bacteria are catalase positive. *X. nematophila* and *P. luminescens* are the most extensively studied species.[6,8] *Xenorhabdus* and *Photorhabdus* exist in two different phases, i.e. the primary and secondary phase. Phase change from the primary to the secondary phase can take place during the cultivation of bacteria for a long time. Bacteria in primary phase are always isolated from the infective juvenile. In the secondary phase, many of the strains lose the ability of dye binding, secondary metabolites and exoenzymes production and swarming on the agar surfaces. *Photorhabdus* also loses the bioluminescence ability in secondary phase.[7]
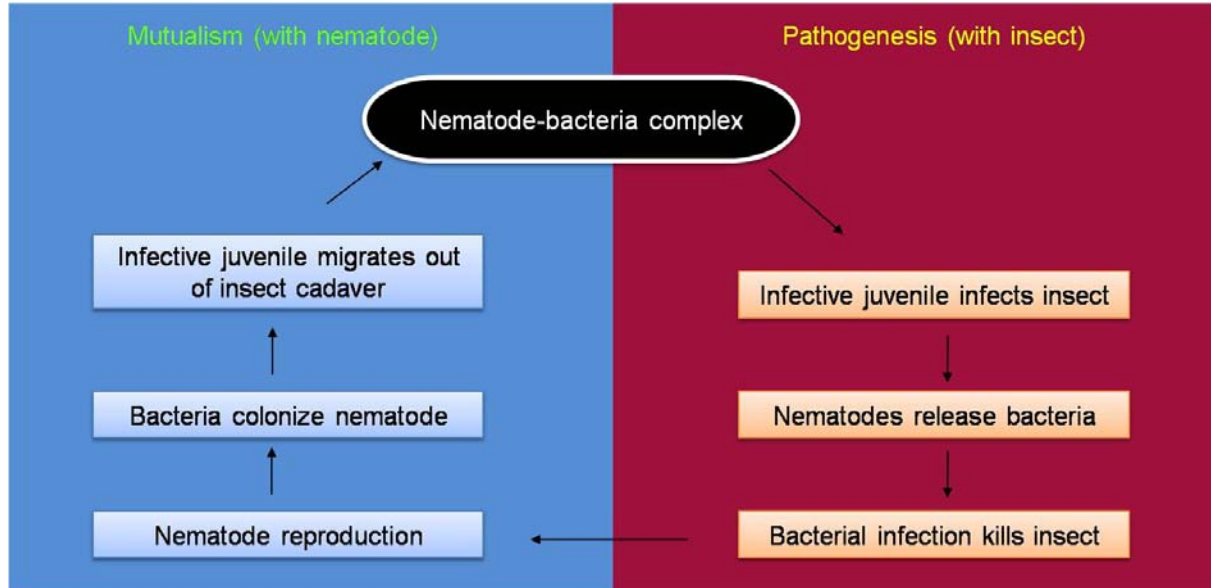
Figure 2. Mutualism and pathogenesis of *Xenorhabdus* and *Photorhabdus* (adapted from Goodrich-Blair *et al.*[5]).

In addition to the production of lethal toxins and enzymes to kill the insect larvae, a series of secondary metabolites are synthesized as signal molecules, phenoloxidase inhibitors, antibiotics against other microorganism and for other unknown functions.[9] These secondary metabolites play important roles in the symbiosis relationship with the nematode and the pathogenecity against insect larvae.[5] Moreover, they have a

potential in commercial pharmaceutical industries. Since the genome sequencing is very well-established in this 21st century, there are a number of *Photorhabdus* and *Xenorhabdus* genomes already published or will be sequenced in the future.[10-14] Decoded genomes can help us to understand the biosynthesis of secondary metabolites.

4.1.2. Secondary metabolites from *Photorhabdus* and *Xenorhabdus*

In the past years, a number of secondary metabolites (Figure 3, Figure 4 and Figure 5) produced by the entomopathogenic bacteria *Photorhabdus* and *Xenorhabdus* were identified by different methods. The standard method with the isolation and the structure characterization via NMR was found suitable for some compounds produced in a large quantity.[15] However, some compounds presumably crucial in the pathogenic and symbiotic lifestyle of *Photorhabdus* and *Xenorhabdus* could not be produced for the reason of silent and cryptic gene clusters, or are not easily detected under laboratory condition. In this case, different strategies have been used to identify these cryptic natural products. Among them were efforts to switch on silent gene clusters via promoter exchange in the host strain,[16-18] by heterologous expression of secondary metabolite genes in *E. coli*,[16,19,20] by heterologous expression of secondary metabolite genes using the yeast homologous recombination cloning methode[21] or by injection of bacteria into the insect mimicking a more natural growth condition.[22,23]

Proschak *et al.* have identified 26 simple **amide derivatives** in *X. doucetiae*, which are condensation products of different acyl moieties with phenyl ethylamine or tryptamine.[24] Among them is the compound *N*-phenylethyl-2-phenylacetamide, which has already been isolated in *X. nematophila* in an earlier publication.[25] Some of these compounds were only produced in trace amounts, not allowing their isolation from the producing strain. Therefore several compounds, whose structure was elucidated by MS-experiments, were chemically synthesized and were shown to have cytotoxic activity in bioactivity screenings.

Using several defined alternative liquid media, Theodore *et al.* have successfully switched on the gene cluster responsible for the biosynthesis of **glidobactin A** and its derivatives in *P. asymbiotica* ATCC43949.[22] They also showed that glidobactin A

and its derivatives could be found in the live cricket infected with *P. asymbiotica*. Furthermore, glidobactin A could be produced by the heterologous expression of the putative syrbactins gene cluster from *P. luminescens* in *Pseudomonas putida*.[19] Glidobactin A and its derivatives were proved to be potent protease inhibitors.

Ciche *et al.* have isolated a catecholate siderophore, named **photobactin,** from *P. luminescens*. Due to its role as a siderophore, sequestering and transferring Fe^{3+} into the bacterial cells, photobactin was required for growth under ion-limited conditions.[26] However, photobactin is not needed for *P. luminescens* to support the growth and the reproduction of its nematode host. However, purified photobactin was shown to have antibiotic activity, which has suggested its contribution in inhibiting competing bacteria.

Crawford *et al.* have achieved heterologous expression of the **rhabduscin** gene cluster in *E. coli*.[27] Rhabduscin is an amidoglycosyl- and vinyl-isonitrile-functionalized tyrosine derivative found in both *P. luminescens* and *X. nematophila* and was proven to be a highly potent inhibitor of phenoloxidase in biochemical assays. Using stimulated Raman scattering microscopy, it has been demonstrated that rhabduscin derivatives localized at the periphery of wild-type *X. nematophila* cells.

Pristinamycin II$_\alpha$,[23] previously found only in streptomycetes, was identified in extracts of *X. nematophila*-infected *Galleria mellonella* larvae. The biosynthetic gene cluster *pxn*, which is very similar to the gene cluster in *Streptomyce pristinaspiralis,* was confirmed in *X. nematophila* by inactivation of the gene cluster. It is widely accepted that typical building blocks from primary metabolite pathways, such as amino acids, fatty acids and sugars are required for secondary metabolites biosynthesis. On the other hand, building blocks from secondary metabolites pathways are not essential for the biosynthesis of primary metabolites. However, it has been found that *pxnLM* from the gene cluster responsible for the biosynthesis of the secondary metabolite pristinamycin II$_\alpha$ was also involved in the biosynthesis of primary metabolites, namely *iso*-fatty acid biosynthesis.

prexenocoumacin B

N-phenethyl-2-phenylacetamide

rhabduscin

photobactin

isopropyl stilbene

pristinamycin II$_\alpha$

nematophin

dialkylresorcinol

glidobactin A

p-hydroxyphenyl propionic acid

benylideneacetone

phurealipid A

fabclavine Ia

xenorhabdin-2

xenocyloin A

AQ-256

xenofuranone A
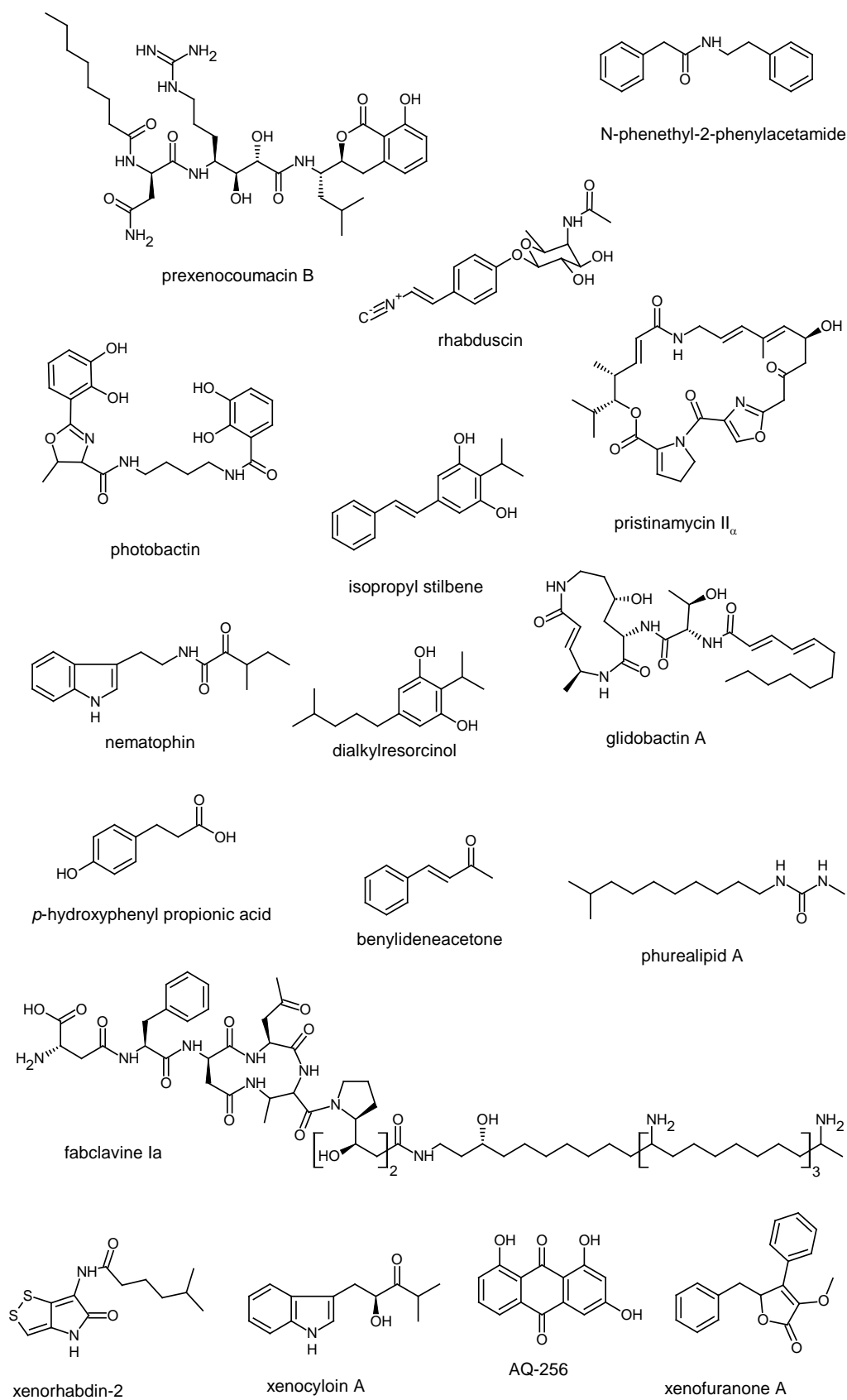
Figure 3. Selected secondary metabolites from *Photorhabdus* and *Xenorhabdus* (Part I).

In all *Photorhabdus* strains investigated,[28] the major secondary metabolites 3,5-dihydroxy-4-isopropyl-*trans*-stilbene (isopropyl **stilbene,** IPS) and 3,5-dihydroxy-4-ethyl-*trans*-stilbene (ethyl stilbene) could be found. Several research groups have isolated IPS and successfully described relevant antibiotic and phenoloxidase inhibitor activities.[29-31] An unstable epoxide-derivative of IPS was also identified in larvae of *Galleria mellonella* infected with a *P. luminescens–Heterorhabditis* association.[32] The biological role of epoxide derivative could not be elucidated. Joyce *et al.* have revealed the biosynthetic pathway of stilbene, different to that of stilbene biosynthesis in plant and also suggested that IPS is required for nematode growth and development.[33] Using the mutasynthesis and chemical synthesis, several novel stilbene derivatives could be synthesized. Two chemically synthesized derivatives were active against *Leishmania donovani*.[34]

Unlike the familiar LuxI/LuxR-type quorum sensing system using *N*-acyl homoserine lactones for bacterial communication, α-pyrones (**photopyrones**) were found in *P. luminescens* as a new type of quorum sensing molecule. The orphan LuxR-type response regulator PluR is used to detect α-pyrones as signaling molecules in *P. luminescens* and activate transcription of the *pcf* operon, leading to cell clumping.[35] Furthermore, Brameyer *et al.* have elucidated the cell-cell communication system in the human pathogen *P. asymbiotica*.[36] Both **dialkylresorcinols** and **cyclohexanediones**, produced by DarABC[37], were identified to act as signaling molecules to induce PauR-mediated gene expression, which is another LuxR-solo type response regulator.

Several simple urea lipid compounds (**phurealipids**) were identified in *P. luminescens* as an inhibitor of the insect juvenile hormone epoxide hydrolase, a key enzyme in the insect development and growth.[38] Thus, phurealipids were proposed to be a virulence factor of bacteria in insects.

Brachmann et al. have isolated **xenofuranone** A and xenofuranone B[15] from *X. szentirmaii* for the first time. It was proposed that xenofuranones were biologically derived from condensation and cyclization of two phenylpyruvate moieties followed by decarboxylation. Xenofuranone A was proposed to be the methylation product of xenofuranone B. *Aspergillus terreus* was shown to be able to produce similar furanones with the weak cytotoxicity against cancer cells.[39] However, biological functions of these compounds in *Xenorhabdus* still await their discovery.

Bacteria of the genera *Xenorhabdus* are also known for the production of yellow antibiotics named **xenorhabdins**. In 1991, McInerney *et al.* have isolated five xenorhabdins as natural products for the first time.[40] Later in 1995, Li *et al.* have isolated two additional xenorhabdins.[41] Xenorhabdins belong to dithiolopyrrolone derivatives which were also identified from the cultures of *Streptomyces clavuligerus*,[42] *Alteromonas rava*[43] and *Saccarothrix algeriensis*.[44] Recently, Bode *et al.* have identified two new xenorhabdin derivatives via a promoter exchange experiment in *X. doucetiae*.[17] With respect to the best-studied biosynthesis of a dithiolopyrrolone derivative (holomycin) in *Streptomyces clavuligerus*, Bode *et al.* have also proposed the biosynthesis of xenorhabdins in *X. doucetiae*. Dioxide derivatives of xenorhabdins were identified as xenorxides.[45] Dithiolopyrrolones have been reported to have antimicrobial and insecticidal activities as well as anticancer properties.[40,45] Dithiolopyrrolone derivatives have also been successfully synthesized chemically.[46]

Xenocyloins, previously known as indole derivatives, could be isolated from different *Xenorhabdus* strains.[28,47] Proschak *et al.* have identified the biosynthesis gene cluster of xenocyloins and elucidated the biosynthetic mechanism.[48] According to the results from *in vitro* assays, structure modeling and mutagenesis experiments, they have shown that a putative 3-ketoacyl acyl carrier protein synthase III de facto acted as an acyltransferase, transferring an acetate- or propionate-unit to the free alcohol group. Despite antibiotic activity, xenocyloins were proven to be active against insect hemocytes. Thus, xenocyloins were proposed to contribute to the overall virulence of *Xenorhabdus* against insects.[48]

Fuchs *et al.* have identified peptide-polyketide-polyamino hybrids (called **fabclavines**) and their biosynthesis gene cluster in *X. budapestensis* and *X. szentirmaii*.[49] Fabclavine Ia could be isolated from *X. budapestensis* and its structure was elucidated by NMR and MALDI-MS[(2)] analyses. Since fabclavines are active against different organisms (bacteria, fungi and other eukaryotic cells), they are considered to be able to protect infected and killed insect larvae from various food competitors.

Richardson *et al.* have isolated **anthraquinone** (AQ) AQ-256 and its derivatives from *P. luminescens* (earlier known as *X. luminescens*).[30,31] Later, Brachmann *et al.* identified the *ant* gene cluster encoding a type II PKS, which was proven to be responsible for AQ biosynthesis in *P. luminescens*.[1] Because the investigation of AQ

biosynthesis is one of the main focuses in this work, further details will be described later in this work (Chapter 4.2.4.).

In 2004, Ji *et al.* identified **benzylideneacetone** from *X. nematophila* being active against some Gram-negative bacteria.[50] In addition to benzylideneacetone, a **linear proline-tyrosine dipeptide** and an **acetylated phenylalanine-glycine-valine tripeptide** were both isolated and they were proven as phospholipase A2 inhibitors. These three compounds could also be isolated from *X. nematophila* and *P. temperata* culture broth. Additionally, *X. nematophila* can produce four additional phospholipase A2 inhibitors (**indole**, **oxidole**, *cyclo*-**proline-tyrosine dipeptide** and *p*-**hydroxyphenyl propionic acid**).[51] Phospholipase A2 is crucial for the insect's immune response and thus needs to be overcome in order to colonize the insect.

Two very small compounds, **benzaldehyde**[52] and **phthalic acid**[53], were isolated from *P. temperata*. Benzaldehyde possessed antioxidant, insecticidal and antimicrobial activities. Moreover, phthalic acid has the capacity to inhibit phenoloxidase with a consequence to suppress the insect's immune defense.

Derzelle *et al.* have identified the putative gene cluster of a **carbapenem**-like antibiotic in *P. luminescens*.[54] However, its exact structure was not characterized. Carbapenem antibiotics are members of the β-lactam family of antibiotics, which are now the most important class of antibiotics for clinical use.

Nematophin was isolated from *X. nematophila* by Li *et al.* in 1997. The metabolite showed antifungal and antibacterial activities.[55] The debate whether nematophin possesses a strong antistaphylococcal effect is yet to be verified.[56]

Xenocoumacins are major secondary metabolites produced by several *Xenorhabdus* strains. Initially, McInerney *et al.* isolated xenocoumacin-1 (Xcn1) and xenocoumacin-2 (Xcn2) with antibacterial activities.[57] Xcn1 and Xcn2 are benzopyran-1-one (isocoumarin) derivatives and they share some similarity with amicoumacins.[58] Reimer *et al.* successfully identified the gene cluster for Xcn biosynthesis in *X. nematophila* and proposed the PKS/NRPS hybrid biosynthesis mechanism.[59] Later, due to the discovery of five inactive precursor molecules, i.e. prexenocoumacins, Reimer *et al.* identified a new activation and resistance mechanism during Xcn biosynthesis, which is widespread in different bacteria.[60] This mechanism describes the transport of prexenocoumacins outside the cell and the

activation of prexenocoumacins by a peptidase, resulting in the formation of Xcn1 to kill food competitors. Additionally, a detoxification mechanism, converting Xcn1 to less toxic Xcn2, was proposed to be involved in the protection of *X. nematophila* itself.

Since *X. nematophila* is the best studied *Xenorhabdus* strain, several NRPS derived peptides were isolated and their structure and biosynthetic mechanism were elucidated. First of all, Lang *et al.* isolated **xenortides** A and B from *X. nematophila*.[56] Xenortides A and B are linear dipeptides condensed with phenethylamine and tryptamine, respectively. Later, Reimer *et al.* discovered an additional xenortide-derivative and its biosynthesis gene cluster in *X. nematophila*, which encodes a bimodular NRPS including a C-domain for the release step. Xenortides have shown activities against *Plasmodium falciparum* and *Trypanosoma brucei*.[61] Lang *et al.* have also isolated a cyclic depsipeptide named **xenematide** in *X. nematophila* with a weak insecticidal activity.[56] The further analysis of secondary metabolites produced by *X. nematophila* led to the structure elucidation of more xenematide-derivatives including their stereochemistry as well as the identification of the four-module NRPS responsible for biosynthesis.[62] In addition to xenortides, several new linear N-methylated peptides termed **rhabdopeptides** were identified by Reimer *et al.* in *X. nematophila*.[63] In their work, a gene cluster encoding NRPS responsible for their biosynthesis was identified in a promoter trap strategy. An iterative usage of one or more modules might be responsible for the different lengths of rhabdopeptides. Rhabdopeptides are active against insect hemocytes. Recently, heterologous expression of a previously uncharacterized gene cluster from *X. nematophila* in an *E. coli* strain with deletion of five transaminases led to the identification of the cyclic **xenotetrapeptide**.[64] Using this transaminase-deficient strain, the number of D-amino acids present in xenotetrapeptides could be assigned by cultivation of the *E. coli* strain in D_2O. Gualtieri *et al.* and Fuchs *et al.* have identified lysine-rich cyclic peptides termed **PAX peptides** (peptide-antimicrobial-*Xenorhabdus*), possessing strong activity against diverse pathogenic fungi, in *X. nematophila*.[65,66] Fuchs *et al.* have successfully elucidated structures of thirteen PAX peptides and identified the biosynthesis gene cluster encoding NRPS.

Among all *Photorhabdus* strains, *P. luminescens* is the most intensively studied strain. In addition to the abovementioned secondary metabolites from *P. luminescens*, several NRPS-produced peptides could be identified. Brachmann *et al.* were able to

identify the blue pigment **indigoidine** (5,5′-diamino-4,4′-dihydroxy-3,3′-diazadiphenoquinone-(2,2′)) in *P. luminescens* by a promoter exchange and in *E. coli* by heterologous expression of the biosynthesis gene *indC*.[16] However, indigoidine is not produced by the *P. luminescens* wild type. To date, the function of indigoidine as well as the regulation of its production in *P. luminescens* is still unclear. **GameXPeptides** were identified as *cyclo*-peptides using labeling experiments, MS analysis, and gene expression in transaminases mutants alone, as they could not be isolated from *P. luminescens*.[67] Until now, GameXPeptides could be produced by various methods, including heterologous expression in *E. coli*[21,64] and promoter-exchange in *Photorhabdus*.[17] Additional GameXPeptides were identified in a heterologous host during adding missing building precursors in the production media in order to mimic the environment inside insect larvae.[68] Bode *et al.* have also identified the linear peptide **mevalagmapeptide** in *P. luminescens*, similar to rhabdopeptide in *X. nematophila*.[63,67] Recently, Bode *et al.* have elucidated the structure of **kolossin A** which was only produced by the giant otherwise silent NRPS in *P. luminescens* after a promoter exchange.[18] In spite of its unusual structure, fully alternating D-/L- configuration with the exception at two amino acids, kolossin A did not show any antibiotic activity. As kolossin is not produced under the laboratory condition, it was proposed that the production of kolossin might require a specific but unknown signaling factor. One chemically synthesized isomer of kolossin showed specific activity against *Trypanosoma brucei rhodesiense*, cause of a deadly infection in Africa.

Additionally, several other NRPS-produced peptides could be identified in different *Xenorhabdus* strains. An N-formylated depsipeptide **szentiamide** was identified from *X. szentirmaii*.[69] The chemically synthesized szentiamide showed activity against insect cells and *Plasmodium falciparum*.[70] Grundmann *et al.* have successfully identified one novel depsihexapeptide **xenobactin**[71] and four depsipentapeptides **chaiyaphumines**[72] from *Xenorhabdus* sp. PB30.3 and *Xenorhabdus* PB61.4, respectively. Both classes of peptides showed satisfactory good activity against *Plasmodium falciparum*. Another six new lipodepsipeptides and a linear peptide were identified in *X. indica* and later commonly referred as **taxlllaids**.[73] The cloning by yeast homologous recombination led to the discovery of new bioactive peptides in *Xenorhabdus*. The *cyclo*-hexapeptide **ambactin** and linear acylated hexapeptides **xenolindicins** A-C were identified as a result of over-expression of the gene *ambS*

from *X. miraniensis* DSM 17902 and tree genes *xldABC* from *X. indica* DSM 17382 in *E. coli*, respectively.[21] Fuchs *et al.* have used the neutral loss fragmentation pattern for screening of arginine-rich secondary metabolites, led to the identification of **bicornutin** A1 and A2, **HCTA** and **RILXIRR** peptides.[74]

Schimming *et al.* have elucidated the structure of **pyrrolizixenamides** in *Xenorhabdus*, which belongs to pyrrolizidine alkaloids found widespread as natural products in plants.[20] The bimodular NRPS PxaA and the monooxygenases PxaB are involved in the biosynthesis of pyrrolizixenamides. It was speculated that pyrrolizixenamides might participate in the modulation of the immune response of nematodes or the suppression of the immune response during an infection process.
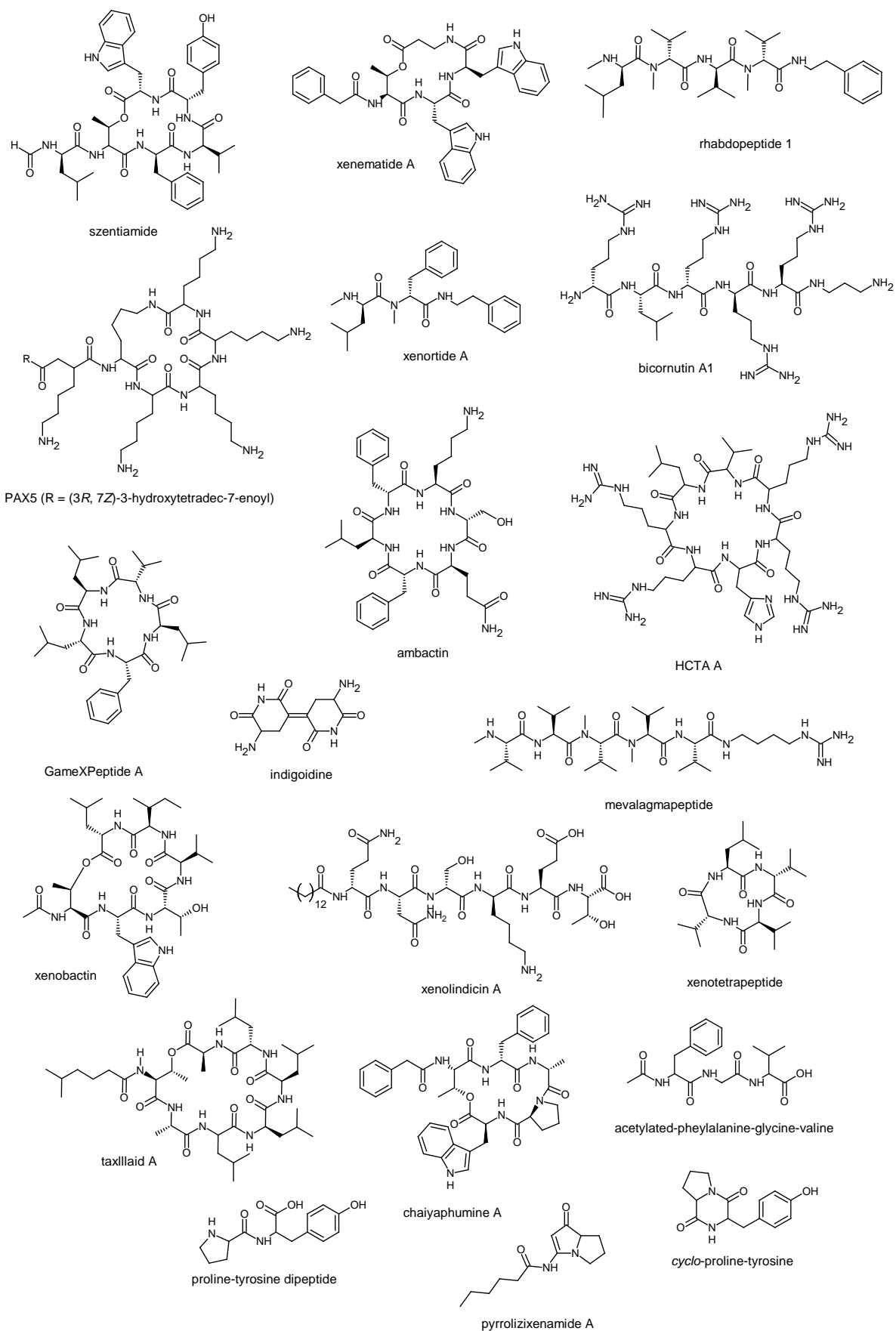
Figure 4. Selected secondary metabolites from *Photorhabdus* and *Xenorhabdus* (Part II).

In this work, more NRPS-produced peptides were identified (Figure 5) from *Xenorhabdus*. For instance, **xentrivalpeptides**[75] could be identified from *Xenorhabdus* sp. 85816 and **xenoamicins**[76] could be identified from both *X. mauleonii* DSM17908 and *X. doucetiae* DSM17909. More details are shown in subsequent chapters (chapter 6.1. and 6.2.).

	R¹	R²
xentrivalpeptide A	Bn	*i*Pr
xentrivalpeptide B	Me	*i*Pr
xentrivalpeptide C	Et	*i*Pr
xentrivalpeptide D	Pr	*i*Pr
xentrivalpeptide E	*i*Bu	*i*Pr
xentrivalpeptide F	*i*hexyl	*i*Pr
xentrivalpeptide G	*n*pentyl	*i*Pr
xentrivalpeptide H	Bn	Et
xentrivalpeptide I	Bn	*sec*Bu

	R¹
xentrivalpeptide J	Pr
xentrivalpeptide K	*i*Bu
xentrivalpeptide L	Bn
xentrivalpeptide M	*i*hexyl

	R¹	R⁴
xentrivalpeptide N	Bn	*sec*Bu
xentrivalpeptide O	Bn	Et
xentrivalpeptide P	Pr	Et

xentrivalpeptide Q

	R¹	R²	R³	R⁴
xenoamicin A	butanoyl	Me	Et	*i*Pr
xenoamicin B	butanoyl	Me	Et	*i*Bu
xenoamicin C	butanoly	Me	Et	*sec*Bu
xenoamicin D	acetoyl	Me	Et	*i*Bu/*sec*Bu
xenoamicin E	butanoyl	Me	Me	*i*Bu/*sec*Bu
xenoamicin F	butanoyl	H	Et	*i*Bu/*sec*Bu
xenoamicin G	pentanoyl	Me	Et	*i*Bu/*sec*Bu
xenoamicin H	H	Me	Et	*i*Bu/*sec*Bu

Figure 5. Selected secondary metabolites (Part III): xenomanicins and xetrivalpeptides identified in this work.

4.2. Biosynthesis of type II polyketide synthases

4.2.1. Secondary metabolites derived from type II PKS

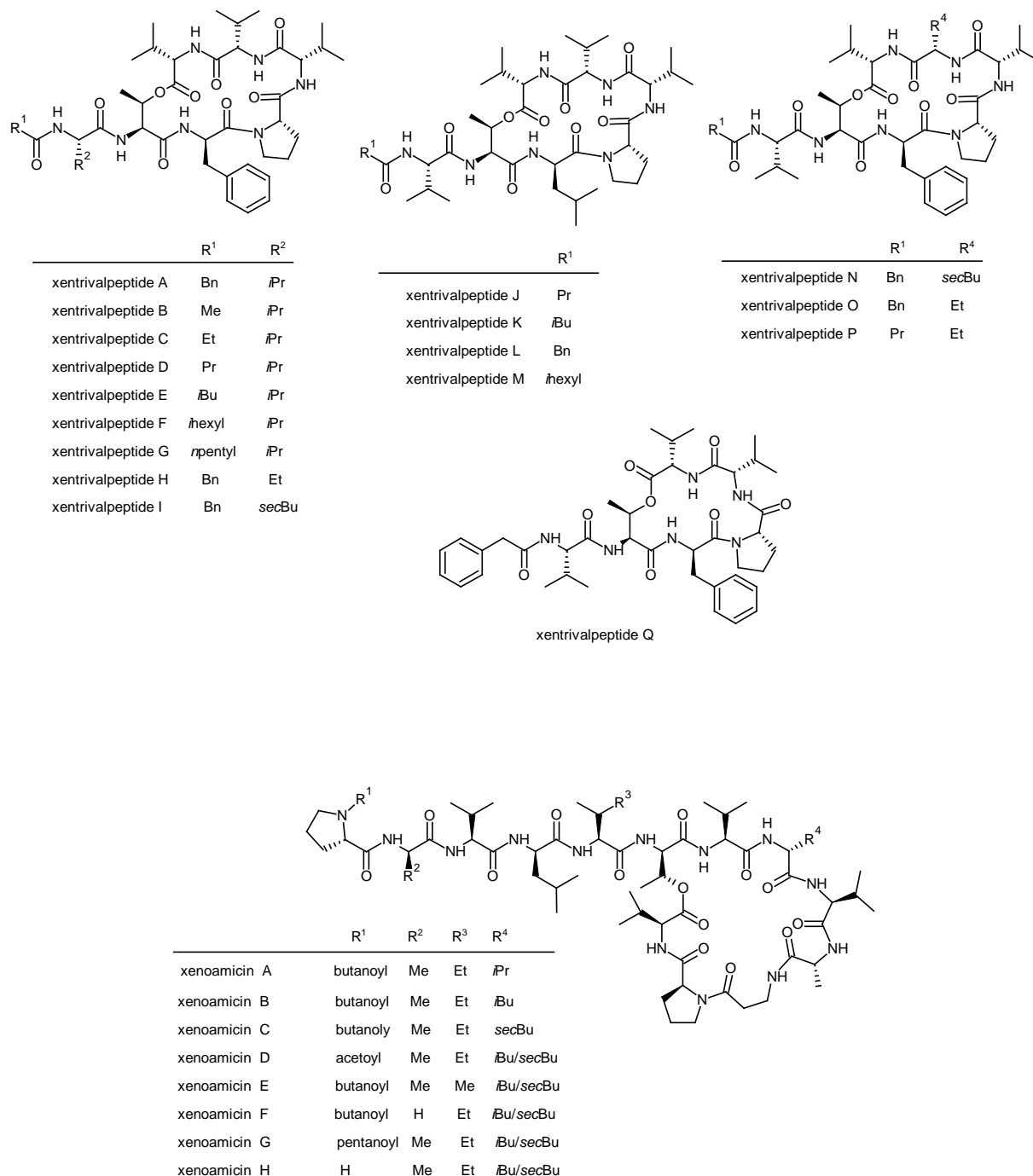Secondary metabolites produced by type II polyketide synthases, also called bacterial aromatic polyketides, are a large family of natural products widely spread in bacteria, especially in soil-borne and marine Gram-positive *Streptomyces*.[77] Although a lot of aromatic polyketides have been isolated and characterized from *Streptomyces* or environmental DNA clones[78], their biological functions, biosynthetic mechanisms and clinical applications are not always elucidated. Some of them are recognized as effective anticancer agents and antibiotics, for example doxorubicin[79] and oxytetracycline,[80] respectively.

actinorhodin
(benzisochromanequinone)

doxorubicin
(anthracycline)

oxytetracycline
(tetracycline)

landomycin A
(angucycline)

tetracenomycin C
(tetracenomycin)

benastatin A
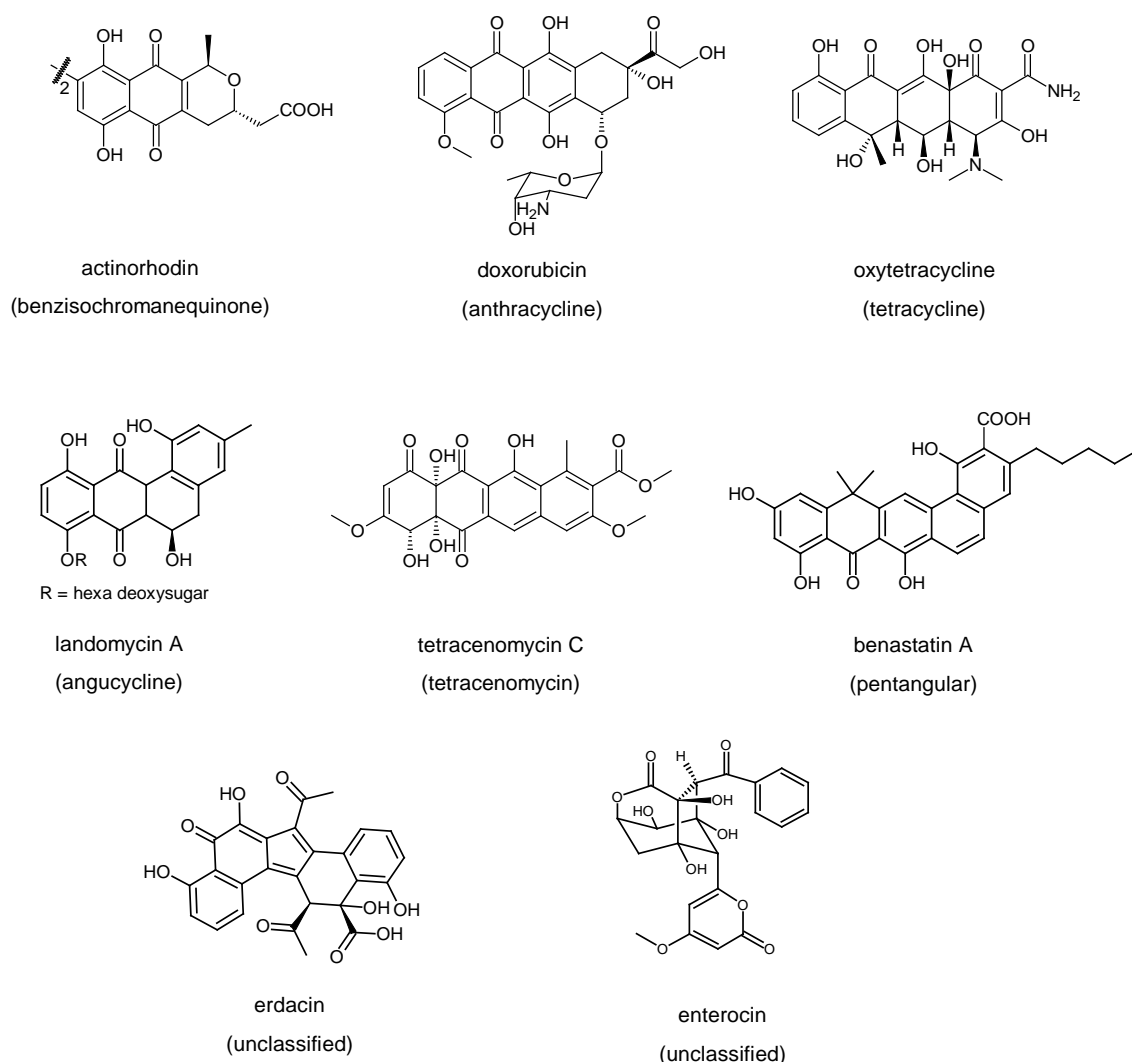(pentangular)

erdacin
(unclassified)

enterocin
(unclassified)

Figure 6. Selected representative secondary metabolites derived from type II PKS and classes (in brackets).

Since 1984, a number of gene clusters encoding type II PKS including the PKS of the model compound actinorhodin have been located and analyzed.[81,82] According to the cyclization patterns and ring topologies, most of the type II PKS-derived polyketides are classified into benzoisochromanequinones, tetracyclines, anthracyclines, tetracemomycins, angucyclines and pentangular polyphenols (see Figure 6).[83] Sometimes, type II PKS-derived polyketides are also divided into only two groups: reduced and unreduced according to the C-9 keto group (see Figure 8).[84] Recently, a phylogenetic study was used to deduce the KR regiospecifity from the amino acid sequence, leading to the prediction of cyclization pattern of type II polyketides.[85]

Benzoisochromanequinones[86] are C-9 reduced tri-aglycons with a lactonized ring. The best know representative benzoisochromanequinones is the blue pigment actinorhodin,[82] which is produced in *Streptomyces coelicolor*. The biosynthesis of actinorhodin has been extensively studied at the genetic and structural level. More details are shown later (chapter 4.2.3).

Anthracyclines[87] are linear cyclized tetracyclic aglycons with one or more deoxy- and aminodeoxy-sugars bound via O-glycosylation. In search for a less toxic anticancer agent, many anthracyclines were discovered by various research groups. Most of the anthracyclines distinguish in C-19 substitution and C-17 glycosylation patterns (numbers see Figure 8). One of the prominent anticancer drugs is doxorubicin with a single deoxysugar residue.[79] Doxorubicin is derived from one propionyl starter unit and 9 malonyl-CoAs. The tricyclic aklanonic acid methyl ester is the key intermediate, which is then cyclized to a tetracyclic aglycon. Additionally, a number of tailoring enzymes is required to obtain doxorubicin.

Tetracyclines[88] were widely used in human and animal medicine against both Gram-positive and -negative pathogens by inhibiting protein synthesis in bacteria. Tetracyclines act as protein synthesis inhibitors by binding to the 30S ribosomal subunit. As a result, many tetracycline resistant pathogens emerged, which led to a decreased effectiveness as a front-line antibiotic. Therefore modified tetracyclines were developed to overcome the resistance problem. One of the well-known tetracyclines is oxytetracycline, which was discovered in 1950 from *Streptomyces rimosus*. Oxytetracycline[80] derives from one malonamyl starter unit and 8 malonyl-CoAs. The C-9 reduced decaketide undergo cyclizations to form a tetracyclic aglycon and the subsequent C-methylation, hydroxylation, oxidation, transamination,

N-methylation and hydroxylations resulted in a highly decorated and oxidized oxytetracycline.

Angucyclines[89] derive from angularly cyclized tetracyclic aglycons. Because of the unusual oxidative skeleton rearrangements and diverse glycosylations, angucyclines exhibit diverse structural scaffolds resulted in big differences as compared to usual benz[a]anthracenes. As a result, angucyclines possess a variety of biological activities, such as inhibition of diverse enzymes, antimicrobial and anticancer activities. A representative anticancer agent is the landomycin group. Landomycins are C-9 reduced decaketides. After closure of the last ring, the spontaneous decarboxylation takes place to result in the common angucycline shunt product UWM6. Oxygenases, reductases and glycosyltransferases modify UWM6 to landomycins. Landomycins with different deoxysugar side chains are produced in *Streptomyces cyanogenus* S-136. Landomycin A with a hexa deoxysugar side chain is the major derivative.

In contrast to anthracyclines and tetracyclines, tetracenomycins derive from unreduced C-9 decaketides. Tetracenomycin C[90] discovered in *Streptomyces glaucescens* Tü49 is a cytotoxic antibiotic.

Pentangular polyphenols have a long-chain polyphenolic angular architecture with at least 24 carbons. Lackner *et al.* have used the term "pentangular polyphenols" for the first time thereby defining a new class of polyketide.[85] A model polyketide for this group is benastatin A, which has potential activities as a glutathione *S*-transferase inhibitor and as an apoptosis inducer. Benastatin A is derived from one hexanoyl starter unit and 11 malonyl-CoAs.

Not all characterized type II PKS-derived polyketides can be classified into the aforementioned groups. Numerous type II PKS-derived polyketides structurally and biosynthetically differ from the typical groups. For example, pentacyclic erdacin is biosynthesized by a type II PKS isolated from environmental DNA.[78] The enterocin without usual multi-aglycons is isolated from the marine strain *Streptomyces maritimus*.[91]

4.2.2. Type II PKS

Most of natural products mentioned above (chapter 4.2.1.) are produced by *Streptomyces*, which are Gram-positive and have genomes with high GC-content. All of these polycyclic aromatic natural products are synthesized by the type II PKS, which is related to the type II fatty acid synthase. Both comprise of a set of highly conserved discrete, monofunctional proteins. Type II fatty acid synthase are usually found in archaea, bacteria and plants.[92] The core set of proteins in type II PKS is called "minimal PKS" and includes three subunits (proteins), the ketosynthase (KS) or KS_α, the chain length factor (CLF) or KS_β, and the acyl carrier protein (ACP). The minimal PKS catalyzes the biosynthesis of the polyketide backbone.[77] In several cases, additional proteins, which are required for completion of the polyketide backbone, are included in the minimal PKS. Examples include the malonyl-CoA:ACP transferase (MCAT) and additional proteins are required for the priming reaction. MCAT, usually shared with type II fatty acid biosynthesis, is required for the transfer of the malonyl group from malonyl-CoA to the *holo*-ACP.

The smallest protein in the type II PKS is the ACP with ca. 10 kDa, which is folded as a bundle of four *α*-helices. Its small size is an advantage to investigate its modification with the mass spectrometer.[93] Each ACP has a conserved serine residue, which is indicated as –OH in the Figure 7. This serine residue has to be modified by the phosphopantetheinyl transferase (PPTase), which can transfer the 4´-phosphopantetheinyl (Ppant) group (indicated as ⩘SH in the Figure 7) from CoA to the conserved serine in the presence of Mg^{2+} ions,[94] leading to the transformation of an inactive *apo*-ACP to an active *holo*-ACP.[95] The *holo*-ACP is responsible for carrying the starter unit, the extender unit and the growing polyketide chain.

The essential PPTase is sometimes encoded within the gene cluster and sometimes independent of the gene cluster. The PPTase are divided into three different groups.[96] Usually a stand-alone PPTase is classified as the second type of PPTase, also called Sfp-type PPTase. The very well characterized PPTase Sfp from the surfactin biosynthesis pathway in *Bacillus subtilis* can activate ACP independently of their origin or affiliation to the biosynthetic logics using the thiotemplate-strategy and thus, widely used in the laboratories to activate ACP during *in vitro* assay or in heterologous hosts *in vivo*.[97]
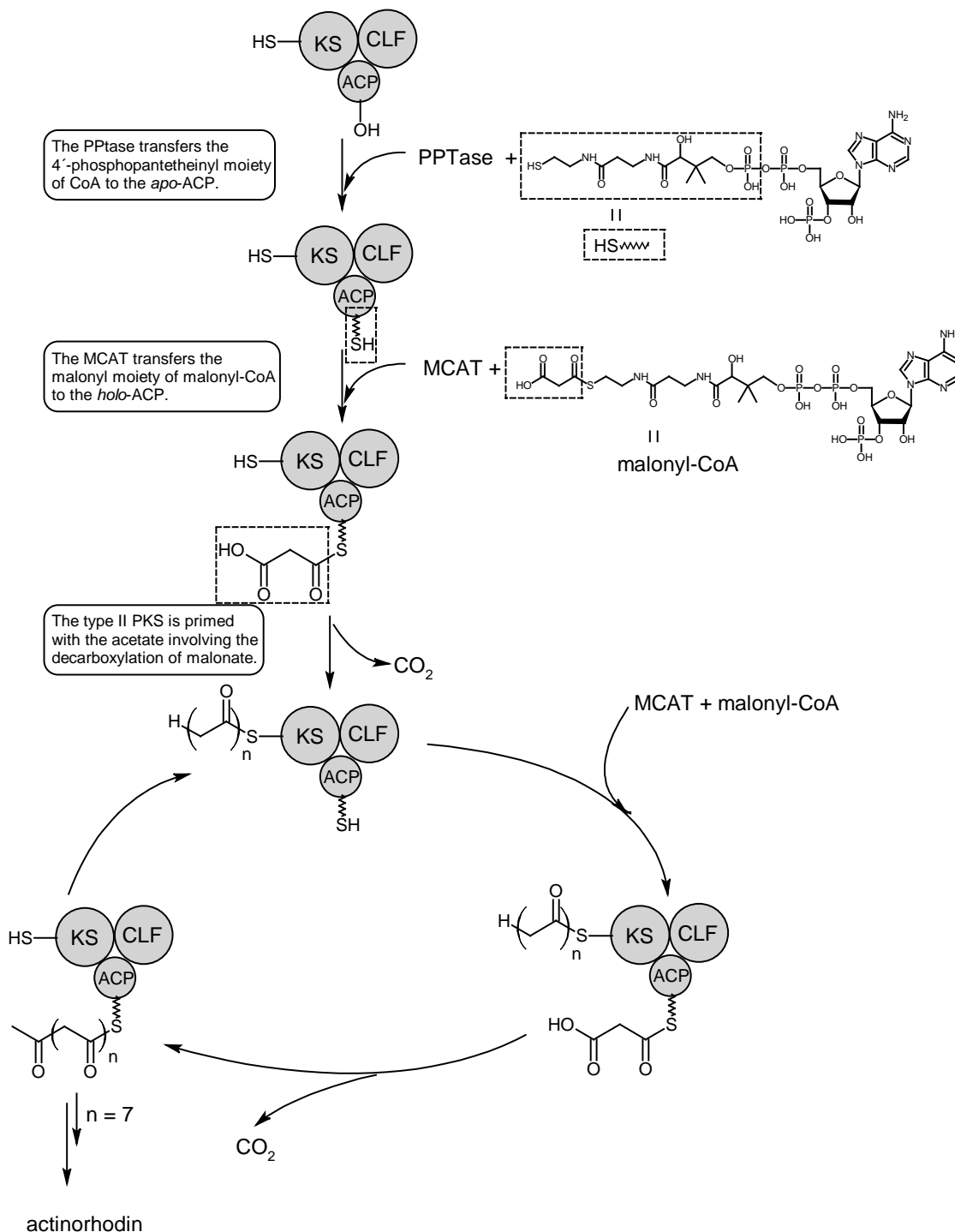
Figure 7. Biosynthesis of the type II PKS-derived polyketide backbone by the minimal PKS in case of acetate as a starter unit (modified according to Shen *et al.*[98]).

Subunits KS and CLF are able to assemble into a heterodimer, resulting in a well-protected reaction cavity. In this cavity, the malonyl-ACP is decarboxylated to form the acetyl-ACP and the KS is primed with the resulting acetyl group. Also polyketide chain elongation takes place in the cavity between KS and CLF.[99] According to an investigation of the actinorhodin minimal PKS, it was proposed that the polyketide

chain length was determined by a measuring mechanism.[100] However, a recent analysis of the KS-CLF heterodimer from the fredericamycin PKS suggested a fundamentally different mechanism, where the KS-CLF heterodimer does not exclusively control the backbone length.[101]
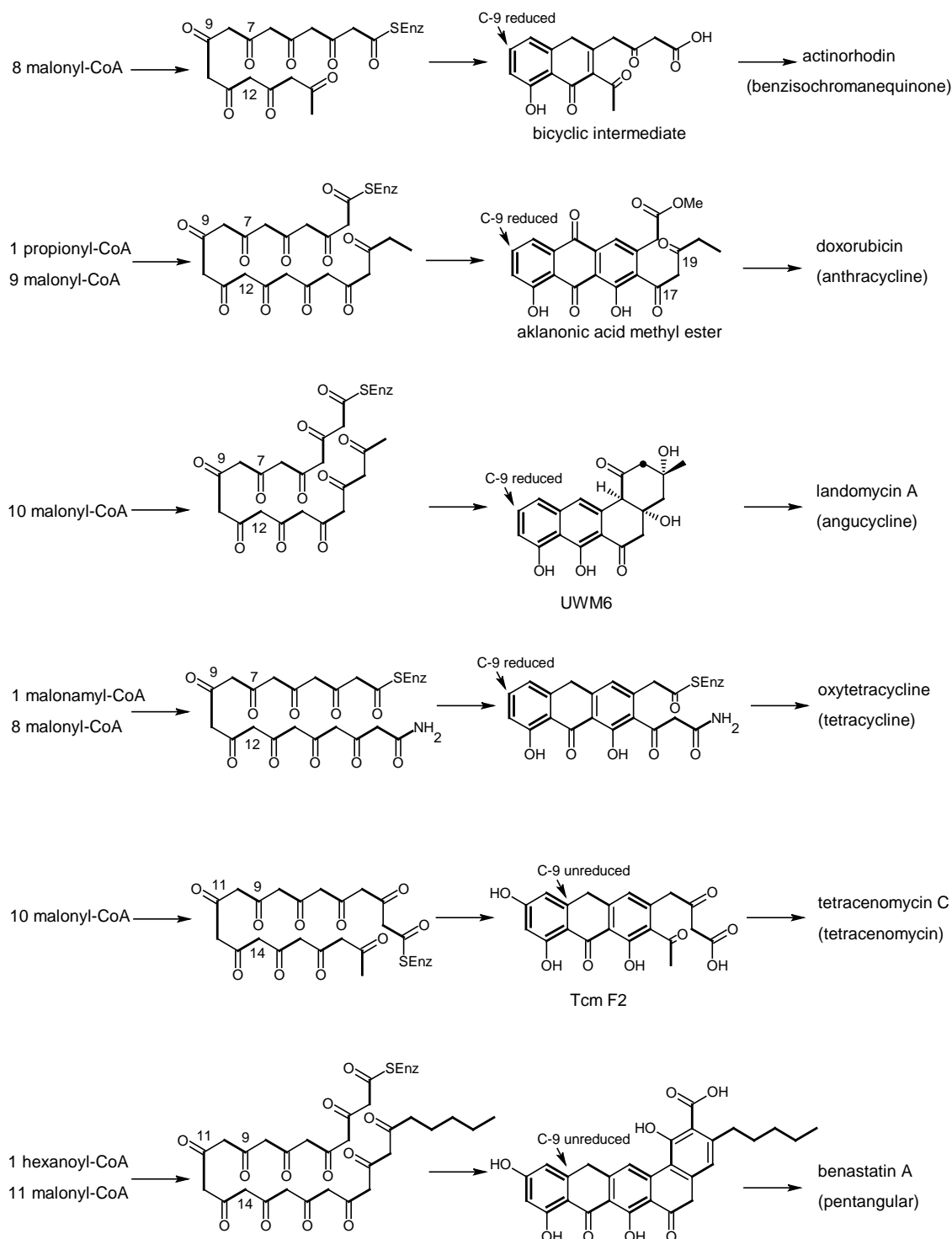

Figure 8. Cyclization patterns of different classes of type II polyketides.

The minimal PKS catalyzes the biosynthesis of the polyketide backbone during the iterative decarboxylative condensation of extender units with a starter unit (see Figure 7). To date, only malonyl-CoA has been found as extender unit; however, there are several choices of starter units: acetate, malonamate, propionate, (*iso*)butyrate, hexanoate, benzoate.[102] Three different mechanisms of the type II priming reaction have been proposed.[77] In first instances, type II PKS-derived polyketides are primed with acetate derived from the decarboxylation of malonate (see Figure 7). In the second priming mechanism, an extra set of enzymes, including a KSIII and an additional ACP, is found. The KSIII catalyses the first condensation of a non-acetate starter unit with a malonyl-CoA to form a β-ketoester, which is then attached to the additional ACP. In some cases, a set of enzymes, such as KR/DH/ER from fatty acid metabolism is used for the modification of the β-keto-*S*-ACP ester. The third priming mechanism is the direct loading of a carboxylic acid activated by an acyl-CoA ligase.

To prevent the incorrect folding and control the programmed cyclization, chain modifying enzymes including C-9 ketoreductases (KR), aromatases (ARO) and cyclases (CYC) are required. The C-9 KR reduces the C-9 keto group to an alcohol and plays an important role in the regiospecific folding of the reactive poly-β-ketone backbone. In most cases, the first ketoreduction takes places after chain elongation steps with the exception of the enterocin biosynthesis.[91] If a C-9 KR is present in the type II PKS, the polyketide backbone is cyclized with a C7-C12 pattern of the first ring (see the biosynthesis of actinorhodin, doxorubicin, landomycin A and oxytetracycline in Figure 8). In contrast, in the absence of a C-9 KR, the first ring of the polyketide backbone can be cyclized according to the C9-C14 pattern (see the biosynthesis of tetracenomycin C and benastatin A in Figure 8) or to the C7-C12 in the biosynthesis of R1128.[103] However, determining the factors of the regiospecific folding are still topics of discussion. The structure and functional analysis of the TcmN ARO/CYC,[104] the ZhuI ARO/CYC[105] and the WhiE ARO/CYC[106] suggested that the polyketide intermediate is transferred into the ARO/CYC interior pocket, which can stabilize the reactive polyketide intermediate, determine the folding pattern, and catalyze the regiospecific cyclization. Interestingly, the cyclization pattern of a unreduced poly-β-ketone produced by the actinorhodin minimal PKS could be directed to C9-C14 by the WhiE ARO/CYC or retained to C7-C12 by the ZhuI ARO/CYC.[106] Furthermore, a model with multi-enzyme complexes was suggested in the cyclization study of the

resistomycin biosynthesis.[107] A roadmap for the cyclization pattern was also drafted according to amino acid sequences.[107]

After cyclization, there are several types of proteins involved in the modification of the cyclized polyketide scaffold, which include methyltransferases, oxygenases, ketoreductases and glycosyltransferases. Most methylransferases transfer the activated methyl group from *S*-adenosyl-L-methionine (SAM) to substrates, resulting in methylated type II PKS-derived polyketides. Cytochrome P-450 monooxygenases, flavin-dependent oxygenases and anthrone oxygenase are known for the ability to modify type II PKS-derived polyketides, resulting in the quinone formation, hydroxylation, epoxidation, oxidative bond breaking and oxidative rearrangement. Glycosyltransferases can modify polyketides by using activated deoxysugars derived from nucleotide deoxysugars and adding them to the polyketide scaffolds.[77]

Generally, there are three different strategies to investigate the biosynthesis of type II PKS-derived polyketides: (1) the mutagenesis in the host strain, (2) the heterologous expression in a closely related strain, and (3) the *in vitro* analysis using purified recombinant proteins. Very often, *Streptomyces coelicolor* CH999, *Streptomyces lividans* and *Streptomyces albus* were used for the heterologous production of type II PKS-derived compounds.[108] A single *E. coli-Streptomyces* shuttle cosmid harboring the gene cluster encoding type II PKS such as pOJ446 can simplify the gene cloning and manipulation. The *in vitro* enzymatic synthesis provides an alternative option to reconstitute the biosynthetic pathway and generate more analogous derivatives, while disregarding the competing reactions taking place in the whole cell.[97] However, the availability of functional recombinant proteins makes it difficult to perform *in vitro* type II PKS biosynthesis. Yet, in contrast to the total synthesis of a type II PKS-derived polyketide in an organic chemistry lab, the *in vitro* enzymatic synthesis of a complex type II polyketide does not have problems with substitutions, stereo-selectivity and the purification of intermediates between multiple reaction steps. The total *in vitro* synthesis of enterocin[91] is one of the best representative examples.

4.2.3. Biosynthesis of actinorhodin in *Streptomyces coelicolor* A3(2)

Actinorhodin (ACT), an aromatic polyketide antibiotic, is produced by the Gram-positive bacterium *Streptomyces coelicolor* A3(2) and belongs to the benzoisochromanequinone (BIQ) class of type II PKS-derived polyketides.[82] ACT consists of two BIQ chromophores connected with each other by a C-C bond. ACT acts as a pH indicator: Red color under acidic and blue color under basic conditions. Currently there are two mainly research groups in the world (Khosla group in UK and Ichinose group in Japan) focusing on the study of ACT biosynthesis.

Initially, the gene cluster with a size of 22 kb was identified. Afterwards, early steps of ACT biosynthesis were elucidated and shunt products were identified in mutants. The following parts of ACT biosynthesis were revealed by different authors: 1) the minimal PKS encoded by *actI-ORF1-3* produces octaketide backbone; 2) the C-9 KR encoded by *act*III plays a role in the reduction of the C-9 keto group;[109] 3) the ARO encoded by *act*VII takes part in the aromatization of the first ring;[82] 4) the CYC encoded by *act*IV is involved in the aldol condensation resulting in the formation of the second ring;[82] 5) the RED1, a stereospecific ketoreductase, controls the formation of the third ring.[110] Corresponding shunt products were identified and supported the proposed ACT biosynthesis mechanism. At present, post-PKS modifying (tailoring) steps to produce the final dimeric ACT structure are main topics of research.[86,111]
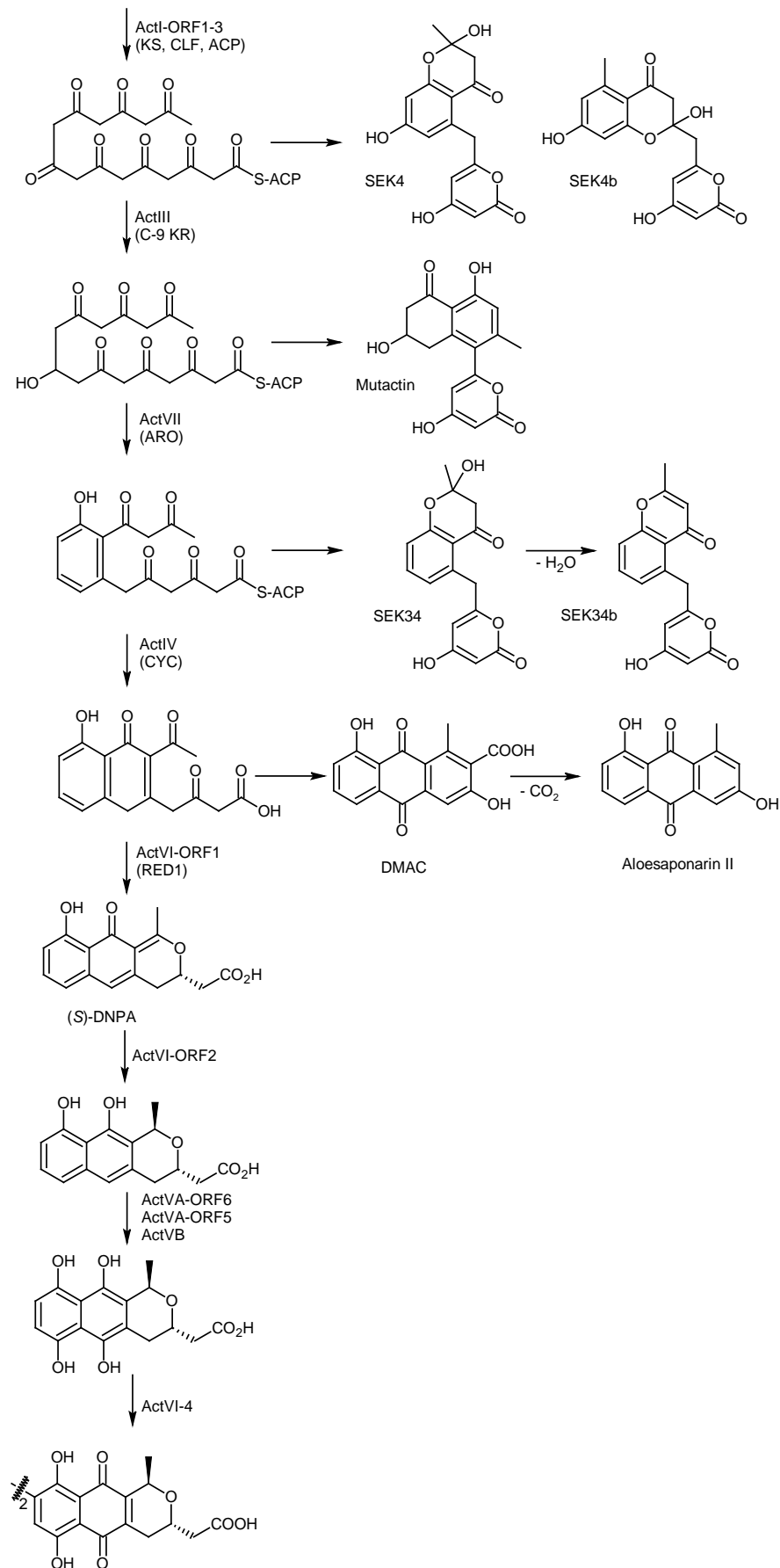
Figure 9. Biosynthesis of actinorhodin (ACT) in the Gram-positive bacterium *Streptomyces coelicolor* A3(2).

4.2.4. Biosynthesis of Anthraquinones in *Photorhabdus luminescens* TT01

On the basis of the structure, anthraquinone (IUPAC: dioxoanthracene) derives from anthracene and consists of three fused benzene rings. Usually, the term anthraquinone refers to one of the most important isomers of the anthraquinones family, 9,10-dioxoanthracene (IUPAC) (Figure 10). Anthraquinone derivatives are ubiquitous in many organisms, such as bacteria,[112] plants,[112] fungi[113] and insects.[114] In addition to the common usage as laxatives,[115,116] textile dyes[117] and bird repellents,[118] a range of biological activities including cytotoxic,[119] antifungal,[120] antimalarial[121] and antiviral[122] activities, have been found in the last years that pointed out potential applications in the future. Although the biosynthesis of anthraquinone derivatives is not well characterized at the biochemical level, two different biosynthetic pathways were identified using feeding experiments: the chorismate/o-succinylbenzoic acid pathway mostly found in higher plants[123] and the polyketide pathway common among fungi and bacteria (Figure 10).[124] PKS-derived anthraquinones can be biosynthesized by iterative type I PKS, type II PKS and type III PKS, examples for each PKS-group are asperthecin,[113] R1128[125] and hypericin,[126] respectively.
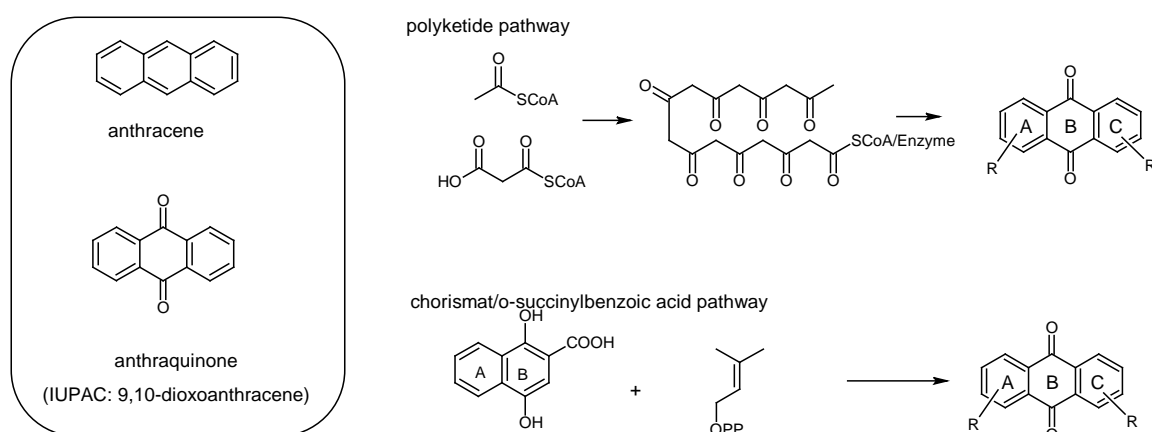

Figure 10. Biosynthetic pathway for anthraquinone derivatives identified from different organisms (modified according to Han *et al.*[123]).

In 2003, the complete genome sequence of *P. luminescens* TT01 was published,[11] which facilitated the understanding concerning the biosynthesis of secondary metabolites produced by *P. luminescens*. In 2007, Brachmann *et al.* have successfully identified the gene cluster responsible for the biosynthesis of

anthraquinones (AQs, Figure 11a) in *P. luminescens*, which mainly produces three AQ (AQ-256, AQ-270a and AQ-284a, Figure 11b).[1] In search of the biosynthetic origins of AQ atoms, feeding experiments with [1-^{13}C]acetate and [1,2-^{13}C$_2$]acetate were performed, leading to the conclusion that AQ are heptaketides (see the isotope-labeling pattern in Figure 11b).

Sequence analysis of the gene cluster revealed that AQ are synthesized by a type II PKS, which is the second example of type II PKS found in Gram-negative bacteria. A total of nine genes were found in the *ant* gene cluster and the functions of the encoded enzymes were initially deduced based on sequence homology analyses. Besides of the typical enzymes from type II PKS, including the minimal PKS AntDEF, the KR AntA, the CYC AntC and the CYC/ARO AntH, a PPTase AntB is also found among the enzymes encoded by the gene cluster. In addition, there are two uncommon genes encoding the CoA ligase AntG and the hydrolase AntI.

Brachmann investigated AQ biosynthesis by gene disruptions with an insertion plasmid or in frame deletions via conjugation of *P. luminescens* with manipulated *E. coli* s17-1 λpir strains carrying the necessary vectors. Only some mutants have been successfully generated, such as *antD*::cat, *antC*::cat,[127] *antB*::cat[127] and Δ*antH*. Two mutants (*antB*::cat and *antD*::cat) failed to produce any polyketides, indicating that the PPTase AntB and the KS AntD are involved in activation or initial reaction. Shunt products mutactin/dehydromutactin, which were identified in the dimeric octaketide ACT biosynthesis in *Streptomyces coelicolor* for the first time, were identified in the *P. luminescens* mutant Δ*antH*, leading to the need for comparing the AQ and ACT biosynthesis. In the mutant *antC*::cat, two shunt products needed to be identified. The results verified that *antD* encodes KS, a protein of the minimal PKS, while *antC* encodes a protein involved in the modification of the poly-β-ketone ketide. Surprisingly, no heptaketides could be found in any mutant and the only shunt products mutactin/dehydromutactin found were octaketides (Figure 11c).

Based on the deduced enzymatic functions of the *ant* type II PKS and the identification of mutactin/dehydromutactin, an AQ biosynthetic pathway is proposed (Figure 11) in comparison with the ACT biosynthesis. At first, the PPTase AntB activates the *apo*-ACP AntF. Then the polyketide biosynthesis is initiated by the attachment of the malonyl or acetyl starter unit to the minimal PKS (AntDEF). After the condensation with seven malonyl-CoAs, a linear octa-β-ketone intermediate is

generated. The KR AntA might be involved in the reduction of C-9 keto group, resulting in the formation of mutactin/dehydromutactin in the *ΔantH* mutant. To date, the functions of other enzymes were still under speculation, without any experimental support information, thus their specific functions have not been clearly assigned.
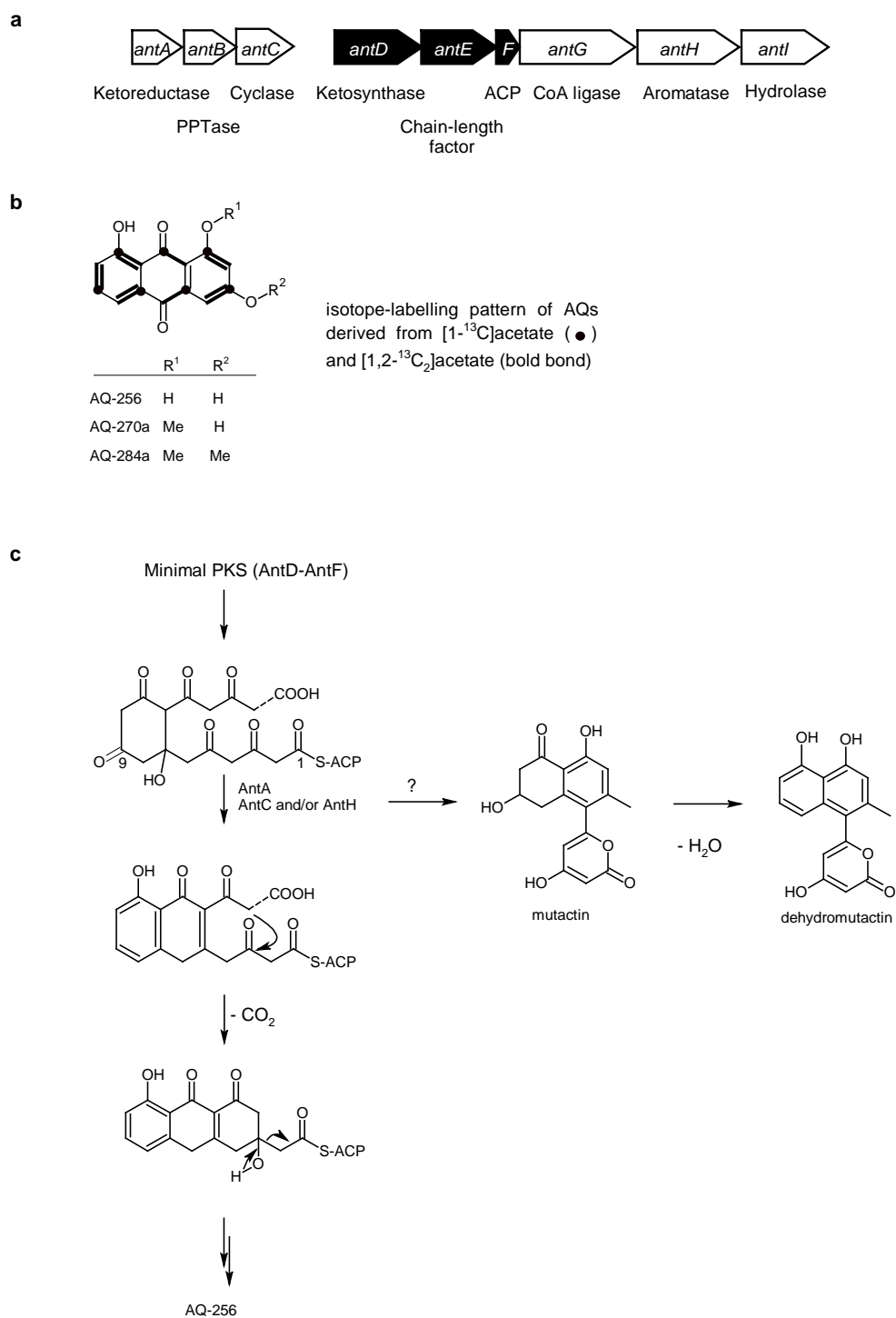

Figure 11. (a) *ant* gene cluster identified in the Gram-negative bacterium *P. luminescens*, (b) structures of the three most abundant AQs in *P. luminescens* and (c) proposed biosynthetic pathway according to the paper of Brachmann *et al.*[1] and Brachmann´s dissertation.[127]

5. Aims of this work

This work is subdivided into two parts: the first part focuses on the identification of selected secondary metabolites from *Xenorhabdus*; the second part concentrates on the biosynthesis of anthraquinones from *Photorhabdus*.

Part I: Identification of selected secondary metabolites from *Xenorhabdus*

As mentioned in the introduction (chapter 4.1.2), a lot of researchers have demonstrated that secondary metabolites from *Xenorhabdus* are potentially bioactive natural products. The biological activities of these new identified secondary metabolites are undoubtedly attractive for pharmaceutical usages. Moreover, they are also important for us to gain better understanding on the pathogenic and symbiotic lifestyle of *Xenorhabdus*. Thus, there are increasing interests to identify more of these new secondary metabolites in *Xenorhabdus*. During secondary metabolites screening in our strain collection of *Xenorhabdus*, the identification of xentrivalpeptides and xenoamicins with numerous derivatives appeared as the priority during my PhD. This work will attempt to isolate xentrivalpeptide A and xenoamicin A, which are produced as the main component by *Xenorhabdus* sp. 85816 and *Xenorhabdus mauleonii* DSM17908, respectively. Because of the presence of trace amount for most derivatives, it is rather impossible to isolate all of them. Thus, this work will intensively deal with suitable analytical methods to elucidate their structures without the need of isolation. In addition, it is indispensable to identify the gene cluster for xenoamicins in *Xenorhabdus doucetiae* DSM17909, whose genome has been sequenced and is also known to produce xenoamicins.

Part II: Investigation of the biosynthesis of anthraquinones from *Photorhabdus*

Brachmann *et al.* have successfully identified the gene cluster for the biosynthesis of anthraquinones in *P. luminescens* strain TT01.[1] Although the biosynthetic pathway of anthraquinones was partially proposed in Brachmann´s work, its plausibility should be supported by more experimental evidences and several issues are still far from understanding, such as: (1) the capability of polyketides formation with the Ant minimal PKS (AntDEF); (2) the identification of other shunt products (heptaketides or octaketides); (3) the functional distinction between the CYC AntC and the CYC/ARO AntH; (4) the activation of ACP AntF by the PPTase AntB; (5) the function of CoA ligase AntG; (6) the identification of the MCAT; and (7) the mechanism of heptaketide generation from octaketide.

The main goal of this work is to investigate the AQ biosynthesis systematically. Heterologous expression of *ant* genes in *E. coli* will be firstly used as a powerful tool to analyze the function of every single gene. To support the current understanding concerning the rare type II PKS in Gram-negative bacteria, an intensive testing of proteins via *in vitro* analyses is performed to provide a thorough insight into the biochemical mechanisms. For this purpose, model compounds will be synthesized and subsequently they serve as possible substrates for *in vitro* assays. The structural modeling of AntI would facilitate the identification of key catalytic amino acids. In addition to analysis of the anthraquinone biosynthesis in *E. coli* and *in vitro*, the approach involved deactivation of selected genes in *Photorhabdus* in order to circumvent problems that might be arising from an improper expression host or suboptimal condition during *in vitro* assay. Finally, this work attempts to prove and refine the previously proposed AQ biosynthetic pathway in order to establish a reliable hypothesis for the biosynthesis of anthraquinone in *P. luminescens*.

6. Publications and manuscript

6.1. Xentrivalpeptides A−Q: Depsipeptide Diversification in *Xenorhabdus*

Authors: Qiuqin Zhou†, Andrea Dowling‡, Heinrich Heide§, Jens Wöhnert†, Ulrich Brandt§, James Baum⊥, Richard Ffrench-Constant‡, and Helge B. Bode*†

† Institut für Molekulare Biowissenschaften, Goethe Universität Frankfurt, 60438 Frankfurt am Main, Germany

‡ Biosciences, University of Exeter in Cornwall, Tremough Campus, Penryn, Cornwall TR10 9EZ, U.K.

§ Zentrum für Biologische Chemie, Molekulare Bioenergetik, Klinikum der Goethe Universität Frankfurt, 60590 Frankfurt am Main, Germany

⊥ Monsanto Company, Chesterfield, Missouri 63017, United States

Attachments: Declaration on the contribution of the author and the paper.

6.2. Structure and Biosynthesis of Xenoamicins from Entomopathogenic *Xenorhabdus*

Authors: Qiuqin Zhou,[a]† Florian Grundmann,[a]† Marcel Kaiser,[b] Matthias Schiell,[c] Sophie Gaudriault,[d] Andreas Batzer,[c] Michael Kurz[c] and Helge B. Bode[a]

[a] Goethe-Universität Frankfurt, Department of Molecular Biotechnology, Max-von-Laue-Strasse 9, 60438 Frankfurt am Main (Germany)

[b] Swiss Tropical and Public Health Institute, Parasite Chemotherapy, Socinstrasse 57, P.O. Box, 4002 Basel (Switzerland)

[c] Sanofi R&D, Industriepark Höchst, 65926 Frankfurt am Main (Germany)

[d] INRA, UMR 1333 Laboratoire DGIMI, 34095 Montpellier (France) and Université Montpellier 2, UMR 1333 Laboratoire DGIMI, 34095 Montpellier (France)

† These authors contributed equally to this work.

Online Access:

http://onlinelibrary.wiley.com/doi/10.1002/chem.201302481/abstract;jsessionid=30EE8568DD125869FDF3CC78694479F5.f01t01

Attachments: Declaration on the contribution of the author and the paper.

6.3. Unusual start and finish of anthraquinone biosynthesis in *Photorhabdus luminescens*

43

Authors: Qiuqin Zhou, Hélène Adihou, Darko Kresovic, Kenan A. J. Bozhüyük, Helge B. Bode*

 [*] Dipl. Chem. Q. Zhou, Dr. H. Adihou, Dipl. Bioinf. D. Kresovic, Dipl. Bioinf. K. A. J. Bozhüyük, Prof. Dr. H. B. Bode , Merck Stiftungsprofessur für Molekulare Biotechnologie Fachbereich Biowissenschaften Max-von-Laue-Str. 9 60438 Frankfurt am Main (Germany)

Prof. Dr. H. B. Bode, Buchmann Institute for Molecular Life Sciences (BMLS), Goethe Universität Frankfurt, Max-von-Laue-Str. 15, 60438 Frankfurt a. M., Germany

Submitted in: *Angew. Chem*

Status: major revision including structural data on the involved proteins required

Attachments: Declaration on the contribution of the author and the manuscript.

6.4. Additional Publications (not part of this thesis)

6.4.1. Biosynthesis of the Insecticidal Xenocyloins in *Xenorhabdus bovienii*

Authors: Anna Proschak,[a] Qiuqin Zhou,[a] Tim Schöner,[a] Aunchalee Thanwisai,[a, b] Darko Kresovic,[a] Andrea Dowling,[c] Richard Ffrench-Constant,[c] Ewgenij Proschak,[d] and Helge B. Bode[a]

[a] Merck Stiftungsprofessur für Molekulare Biotechnologie, Fachbereich Biowissenschaften, Goethe Universität Frankfurt, Max-von-Laue-Strasse 9, 60438 Frankfurt am Main (Germany)

[b] Department of Microbiology and Parasiology, Faculty of Medical Science, Naresuan University, 99 Moo 9, Phitsanulok-Nakhon Sawan Road, Tha Pho, Mueang Phitsanulok, 65000 Phitsanulok (Thailand)

[c] Biosciences, University of Exeter in Cornwall, Tremough Campus, Penryn, Cornwall TR10 9EZ (UK)

[d] Institute of Pharmaceutical Chemistry, Goethe Universität Frankfurt, Max-von-Laue-Strasse 9, 60438 Frankfurt am Main (Germany)

Contribution: Determination of the absolute configuration of xenocyloins using the CD spectrometry. Design of *xclC* genes with mutations (xclC_S253A, xclC_C118A, xclC_Y283A).

Attachment: the paper.

6.4.2. Simple "On-Demand" Production of Bioactive Natural Products

Authors: Edna Bode,[a] Alexander O. Brachmann,[a] Carsten Kegler,[a] Rukayye Simsek,[b] Christina Dauth,[a] Qiuqin Zhou,[a] Marcel Kaiser,[c] Petra Klemmt,[b] and Helge B. Bode[a, d]

[a] Merck Stiftungsprofessur für Molekulare Biotechnologie, Fachbereich Biowissenschaften, Goethe Universität Frankfurt, Max-von-Laue-Strasse 9, 60438 Frankfurt am Main (Germany)

[b] Institute for Cell Biology and Neuroscience, Molecular Cell Biology and Human Genetics, Goethe University, Max-von-Laue-Strasse 13, 60438 Frankfurt (Main) (Germany)

[c] Swiss Tropical and Public Health Institute, Parasite Chemotherapy, Socinstrasse 57, 4002 Basel (Switzerland)

[d] Buchmann Institute for Molecular Life Sciences (BMLS), Goethe Universität Frankfurt, Max-von-Laue-Strasse 15, 60438 Frankfurt am Main (Germany)

Contribution: Isolation and identification of a new xenorhabdin derivative.

Attachment: the paper.

7. Discussion

This work has successfully elucidated the structure of two peptides families from *Xenorhabdus*, i.e. xentrivalpeptides and xenoamicins. However, their biological functions are still unclear. The findings from this work proposed that A- and C-domains in the NRPS possess broad amino acid specificities, resulting in peptides with similar structures. The similar phenomenon could also be found in many other peptides synthesized by NRPS in *Xenorhabdus* and *Photorhabdus* (see Chapter 4.1). This might help widening the range of biological activities. However, the production of many similar peptides in one culture causes a great challenge, which has already been discussed in publications.[75,76] In the first part of discussion (chapter 7.1), the biosynthesis of xentrivalpeptides and the thioesterase in the biosynthesis of xenoamicins are discussed in detail.

This work aimed to elucidate the AQ biosynthesis in *Photorhabdus*. Based on results from the heterologous expression of *ant* genes in *E. coli*, the *antG* inactivation in *Photorhabdus* and *in vitro* assays, it was proposed that the CoA ligase AntG is involved in the biosynthesis of *holo*-ACP (*holo*-AntF). Additionally, this work has elucidated the chain shortening mechanism by the hydrolase AntI based on the findings extracted from *in vitro* experiments using model compounds and the structural modeling of AntI. In chapter 7.2, the AQ biosynthesis with focus on unfinished tasks and experimental problems appeared in this work was discussed.

7.1. Depsipeptides from *Xenorhabdus*: Xentrivalpeptides and Xenoamicins

Xentrivalpeptides

Seventeen depsipeptides termed xentrivalpeptides A-Q (Figure 5) could be identified from *Xenorhabdus* sp. 85816 in this work. They belong to the most diverse depsipeptides produced in *Xenorhabdus*. To date, the genome of *Xenorhabdus* sp. 85816 remains unknown.
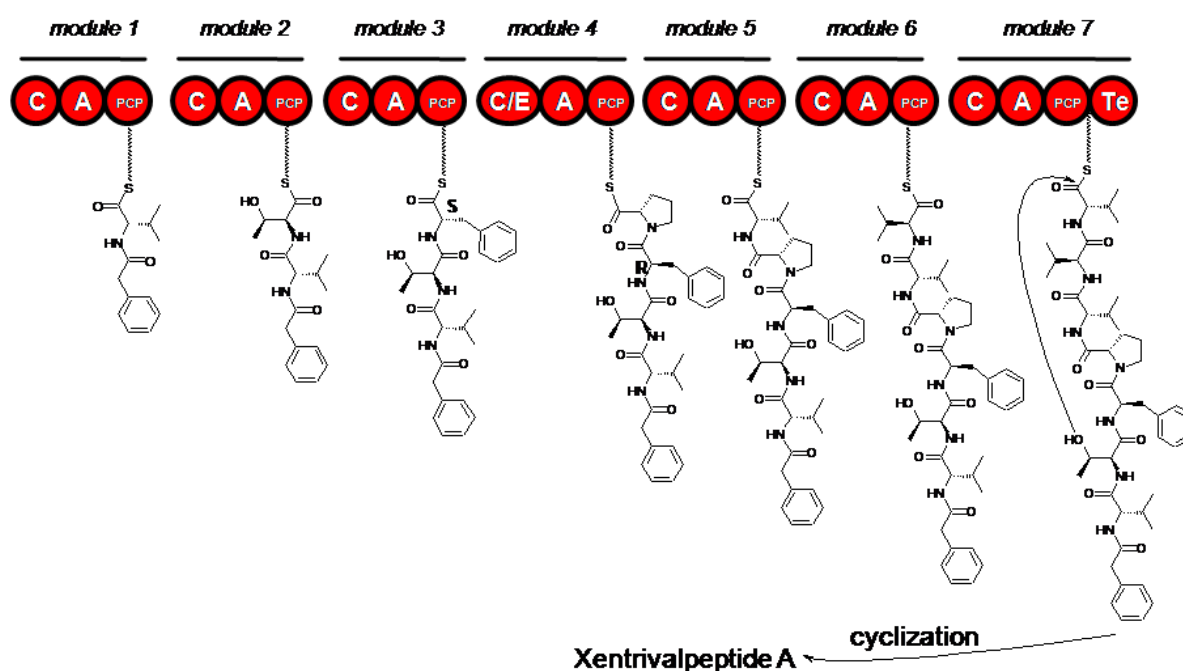

Figure 12. Proposed NRPS responsible for the biosynthesis of xentrivalpeptide A. Xentrivalpeptides B-P might be produced by the same gene cluster.

According to their structures, a gene cluster encoding NRPS might be responsible for the biosynthesis of xentrivalpeptides. The components of this gene cluster can be predicted in detail. This prediction might help to locate the gene cluster encoding in the genome faster and easier in the future work. A NRPS comprising seven modules was predicted to be responsible for the biosynthesis of xentrivalpeptides. As xentrivalpeptide A is produced as the dominant peptide, amino acid specificities of A- and C-domains are predicted as Val-Thr-Phe-Pro-Val-Val-Val according to the structure of xentrivalpeptide A (Figure 12). Considering the large structural diversity of xentrivalpeptides, relaxed A-domain specificities should be taken into account. Furthermore, as a phenyl acetyl moiety is incorporated as starter unit in xentrivalpeptide A, a set of proteins, which might not be included in the gene cluster, but might be required for the biosynthesis and the activation of the phenyl acetate

might be encoded in the genome. The C/E dual domain in module 4 is proposed to be responsible for the epimerization to build the single D-amino acid (phenylalanine in xentrivalpeptide A).[128] The D-amino acid has been reported to significantly change the 3D structure of peptides and stability against the peptidase.[128] Thus, the investigation of the 3D structure of xentrivalpeptides might be the key point to reveal their relevant biological functions.

Xentrivalpeptide Q is produced by *Xenorhabdus* sp. 85816 in a very small amount. Different from xentrivalpeptides A-P, xentrivalpeptide Q has a valine residue missing in the ring structure. It represents NRPS that can produce two classes of peptides (7 amino acids and 6 amino acids) in *Xenorhabdus*. In addition to relaxed amino acid specificities resulting in such diverse peptide structures (Figure 5), we assume that a module-skipping mechanism (skipping from module 5 to module 7) might be responsible for the biosynthesis of xentrivalpeptide Q (Figure 13). Both strategies contribute to the structural diversity of xentrivalpeptides. The unusual module-skipping was only found in a few NRPS. In 2006, the biosynthesis of the pentapeptides myxochromides in myxobacteria involves a module-skipping process.[129] Besides, peptaibol NRPS was shown to produce both 11- and 14-residue peptaibol peptides.[130]
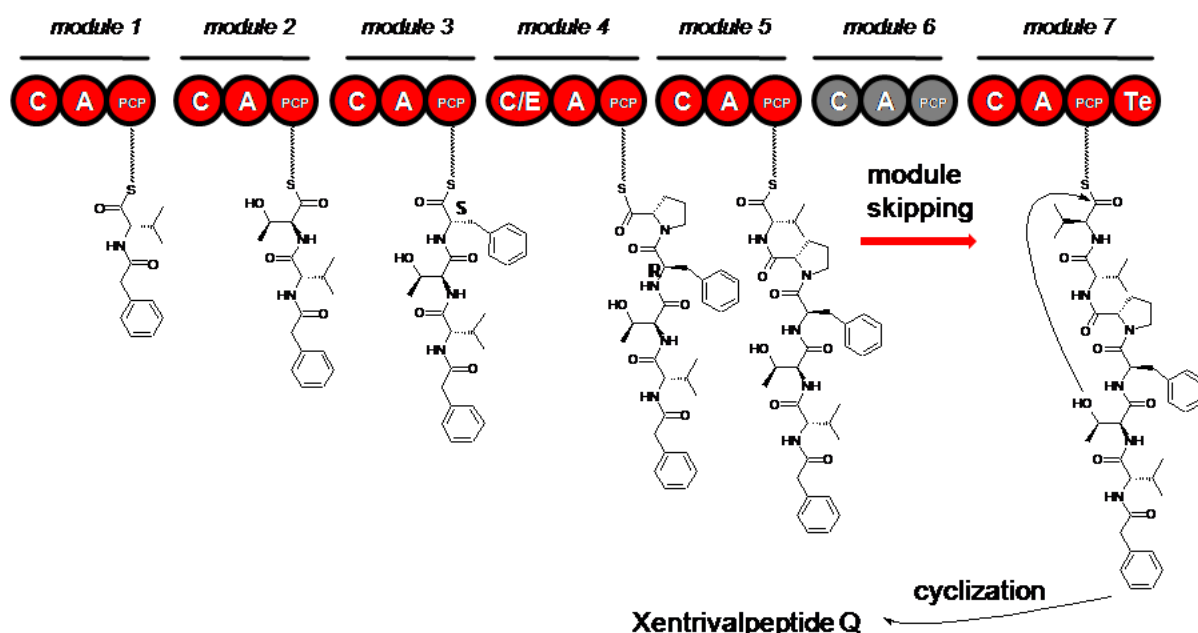

Figure 13. Proposed model for the biosynthesis of xentrivalpeptide Q. Grey domains are presumably inactive. Module skipping: the peptide chain is transfer from the PCP domain of module 5 to the PCP domain of module 7. The absolute configuration of amino acids for xentrivalpeptide Q is proposed according to xentrivalpeptide A.

In the screening for new secondary metabolites produced by *Xenorhabdus*, xentrivalpeptides could also be found in several other *Xenorhabdus* strains: *Xenorhabdus* sp. 86788, *Xenorhabdus* sp. 86789, *Xenorhabdus* sp. 85808, *Xenorhabdus* KK7.4, *Xenorhabdus* sp. 19.4, *Xenorhabdus* sp. 9.5, *Xenorhabdus* PB 49.1, *Xenorhabdus* PB 26.2, *Xenorhabdus* 26.5 and *Xenorhabdus* PB 32.4. As an even larger structural diversity among xentrivalpeptides might be possible, our next task is to isolate and identify xentrivalpeptides from other *Xenorhabdus* strains. The phenomenon that a set of related but structurally different peptides are produced, could be found in several other peptides from *Photorhabdus* and *Xenorhabdus*: rhabdopeptides,[63] GameXPeptides,[67] taxlllaids,[73] chaiyaphumines,[72] xenematides,[56] xenortides[56] and xenoamicins.[76] The impact of production of a set of derivatives on the biological activity has yet to be solved.

Xenoamicins

Xenoamicins, acylated depsipeptides, were identified in this work and more xenoamicin derivatives have been isolated and characterized in different *Xenorhabdus* strains by F. Grundmann[131] and A. Linck (Poster in VAAM-Workshop 2015, information kindly provided by A. Linck) from our research group. According to their structural properties (such as the number of Thr in the peptide and the peptide chain length), xenoamicin derivatives are divided into four subclasses (xenoamicin I-IV, Figure 14 according to Linck´s unpublished results). Recently, Bode *et al.* could exchange the promoter of the xenoamicin gene cluster in *X. doucetiae*, resulting in the overproduction of xenoamicins (xenoamicin C as the main compound) as white needles in the culture.[17] In the future, the promoter exchange approach might help us identify the xenoamicin gene cluster in other *Xenorhabdus* strains, thus allow us to obtain sufficient amounts of xenoamicin derivatives for structure elucidation and bioactivity testing. Using the yeast homologous recombination cloning, Schimming *et al.* have attempted the heterologous expression of the xenoamicin gene cluster in *E. coli*, but no xenoamicins could be detected.[21] It was suspected that some unknown factors (proteins, precursors) might be missing or mutations might have occurred during the PCR.
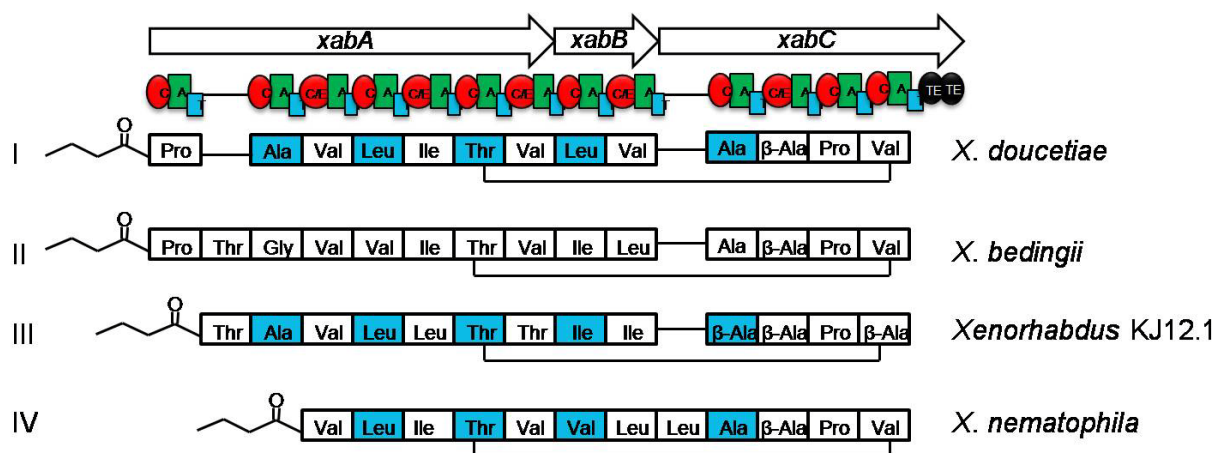
Figure 14. Gene cluster, module architecture and four xenoamicin subclasses I-IV with selected xenoamicins (blue marked amino acids are D-amino acids) in corresponding *Xenorhabdus* strains. (kindly provided by A. Linck, revision of the Xab gene cluster consisting only three genes *xabABC* from Linck´s unpublished results).

During the analysis of the genome database, the possible gene cluster encoding NRPS could be located in the genome of *X. doucetiae* DSM17909. In most cases, the final step of the biosynthesis of depsipeptides is a TE domain catalyzed cyclization of linear peptides. However, the sequence analysis of the *xab* gene cluster revealed an unusual tandem of two distinct TE domains with catalytic triade. This is in contrast to the well-studied single TE domain at the end of the gene cluster encoding NRPS.[132] Recently, Bode *et al.* have successfully exchanged the promoter of the Xab gene cluster whereby xenoamicin C was biosynthesized as the main compound in an extremely high amount.[17] Nevertheless, no linear xenoamicins were detected. The tandem TE might ensure the complete cyclization in the release step. In contrast, the same promoter exchange of the Gxp gene cluster encoding a single TE domain resulted in a partially linear GameXPeptides. In the future, the characterization of individual TE domains might provide more details about their respective cyclization-efficiency and their tolerance towards diverse linear peptide precursors. The assumption, that tandem TE can improve the cyclization efficiency, could be proved by the investigation of tandem TE in the arthrofactin biosynthesis.[133] However, the biochemical investigation of tandem TE in the lysobactin biosynthesis showed that LybB TE1 exclusively catalyzed the formation of a macrocyclic structure, whereas LybB TE2 deacylated unwanted substrates.[134]

Non-proteogenic amino acids

The non-proteinogenic amino acid β-alanine was incorporated in xenoamicins. This non-proteogenic amino acid could be also found In the cytotoxic theonellapeptolide IIIa from the sponge *Lamellomorpha strongylata*[135], and in the antiangiogenic destruxins from the entomopathogenic fungus *Metarhizium anisopliae.*[136] The contribution of β-alanine to the specific structural conformation of xenoamicins has already discussed in Grundmann´s dissertation.[131] In his work, it was discussed that a β-alanine could increase the flexibility and could reduce the tension in the structure. Furthermore, A. Linck showed that the incorporation of the second β-alanine in xenoamicin III provides an interesting point to understand the mechanism of the module evolution in the biosynthesis of xenoamicin biosynthesis in different *Xenorhabdus* strains. In the *xab* gene cluster, several dual condensation/epimerization (C/E) domains were discovered. As the result, the amino acid residues incorporated directly before the C/E domains are usually epimerized. The incorporation of D-amino acids could be verified by the advanced Marfey´s analysis. Peptides with D-amino acids not only give rise to the structural diversity in three dimensions but also provide structural stability against L-specific proteases.[128]

The non-proteinogenic amino acid α-aminobutyric acid was identified in xentrivalpeptide H, O and P. It was confirmed by feeding experiments with L-α-aminobutyric acid and subsequent MS fragmentation experiments. With respect to a majority of organisms, α-aminobutyric acid is biosynthesized via threonine.[137] First, threonine dehydratase converts threonine to α-ketobutyric acid (Figure 15). Subsequently, the amination of α-ketobutyric acid leads to the formation of α-aminobutyric acid. It is of future research interest to prove the biosynthesis of α-aminobutyric acid in *Xenorhabdus*.
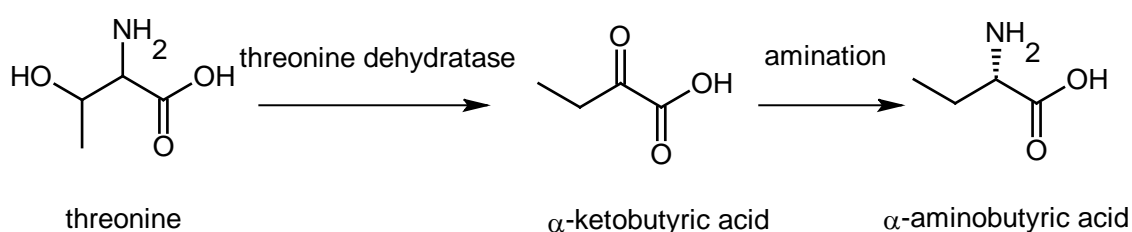

threonine   α-ketobutyric acid   α-aminobutyric acid

Figure 15. Biosynthesis of the non-proteinogenic amino acid α-aminobutyric acid in most organisms.[137]

Secondary metabolites in the future

Nowadays, the databank of microbial genome sequences is expanding quickly. With the help of bioinformatic programs, especially antismash (antibiotics and secondary metabolite analysis shell,[138] http://antismash.secondarymetabolites.org), NRPS, PKS and NRPS-PKS hybrid gene clusters of uncharacterized secondary metabolites could be efficiently located. The bioinformatic results support us as initial points to find new NP, polyketides and NP-polyketide hybrids.

7.2. Biosynthesis of Anthraquinones in *Photorhabdus*

The genus *Photorhabdus* is divided into three species: *P. luminescens*, *P. temperata* and *P. asymbiotica*. The *ant* gene cluster could be found in all *P. luminescens* strains and *P. temperata* strains with known genome sequences. However, the *ant* gene cluster is absent from insect and human pathogenic *P. asymbiotica*.[139]

Anthraquinone in *Photorhabdus* and actinorhodin in *Streptomyces*

The type II PKS responsible for the AQ biosynthesis in *Photorhabdus* is the second example of a type II PKS in Gram-negative bacteria.[140] In this work, the putative CoA ligase AntG with unknown function and the rare C-C hydrolase AntI were intensively investigated by heterologous expression of *ant* genes in *E. coli*, the inactivation of *ant* genes in *Photorhabdus*, *in vitro* assays, and structure modeling of AntI. Besides the differences in priming step and tailoring steps, there are many similarities between AQ biosynthesis and actinorhodin biosynthesis in *Streptomyces coelicolor A3(2)*.[82] Based on the result that the co-expression of *antA-H* with the gene encoding RED1 from the actinorhodin biosynthesis led to the production of (*S*)-DNPA, it was proposed that the bicyclic intermediate was identical in both biosynthesis pathways (Figure 16). However, results from the *in vitro* assays with model compounds in this work suggested that an ACP-bound intermediate in the AQ biosynthesis while an ACP-free intermediate is used in actinorhodin biosynthesis.[110] Thus, further studies are required to confirm that AntI catalyzes reactions on an ACP-bound intermediate.
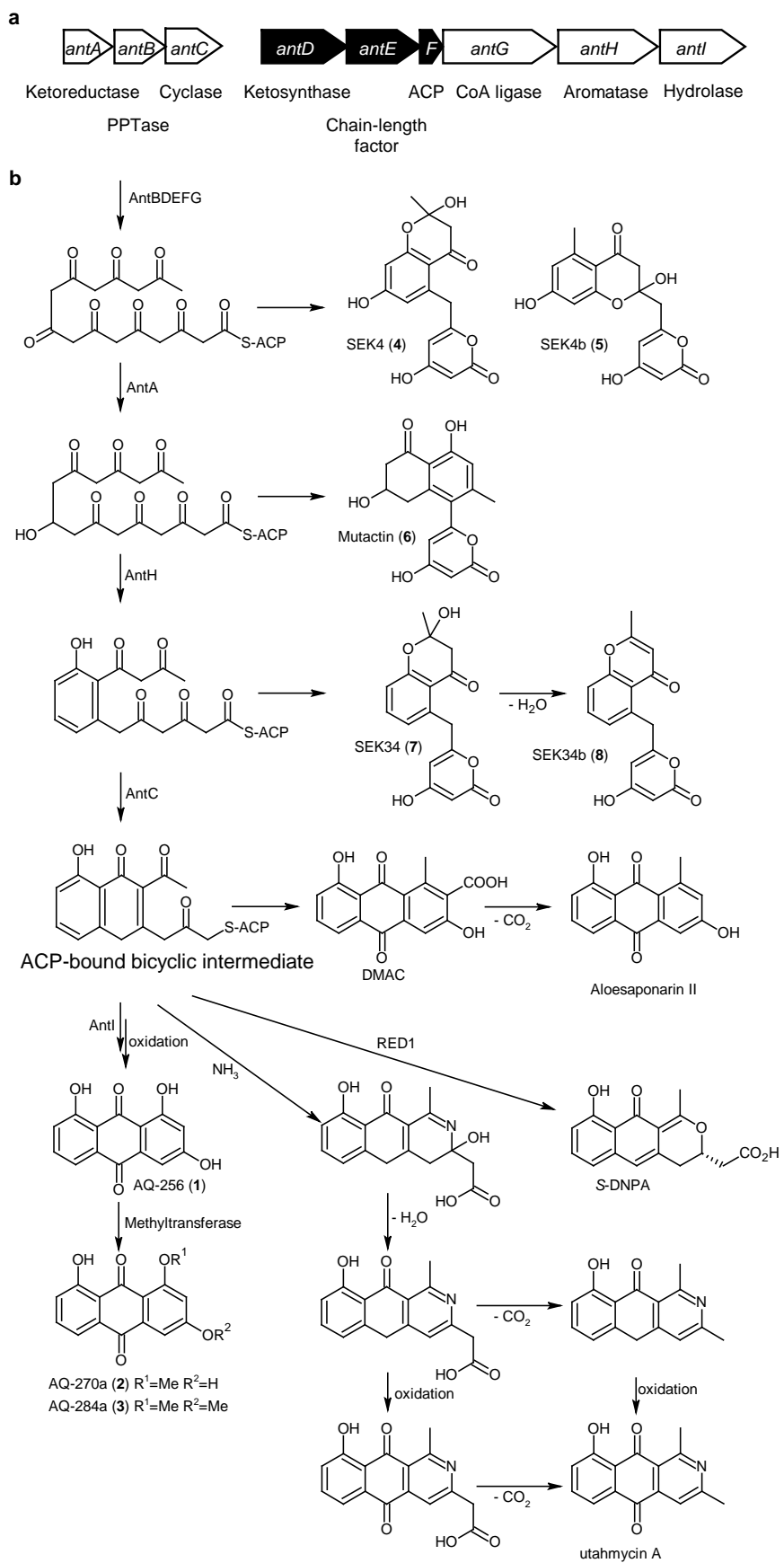
Figure 16. Proposed AQ biosynthesis in *Photorhabdus*. The ACP-bound bicyclic intermediate was proposed to be the substrate for the hydrolase AntI.

Protein expression and purification

The *in vitro* analysis[97] is widely used in the laboratory to investigate reaction mechanisms and substrate specificities of individual proteins. Usually, the greatest challenge is to get a sufficient amount of active proteins. This is because proteins are often expressed as inactive protein aggregates, called inclusion bodies. As inclusion bodies normally contain intact proteins with non-native conformations, various refolding procedures can be used to dissociate, refold and recover active proteins.[141] Different affinity fusion and chromatography systems are appropriate techniques to purify active soluble proteins.[142] However, refolding experiments and the selection of different affinity fusion systems are proved to be rather time-consuming trial-and error methods.

In this work, Duet™ expression systems from Novagen were chosen for expression of single gene or co-expression of different proteins in *E. coli* BL21(DE3). In addition to purified proteins in the submitted manuscript, His$_6$-AntA and His$_6$-AntH and His$_6$-AntC-His$_6$ were successfully purified (unpublished data, Figure 17).
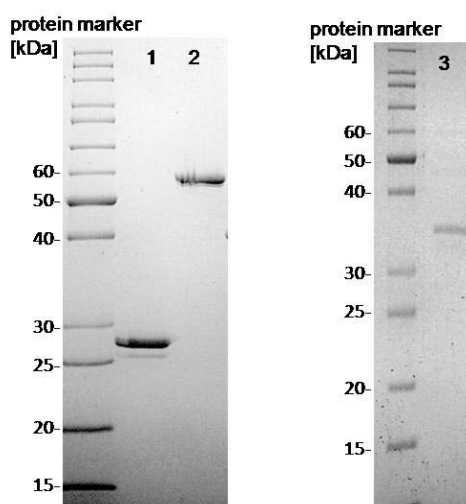

Figure 17. SDS-Page gel of purified recombinant His$_6$-tagged proteins (unpublished data). His$_6$-AntA (29.0 kDa, lane 1). His$_6$-AntH (57.1 kDa, lane 2). His$_6$-AntC-His$_6$ (34.1, lane 3).

Currently, type II PKS is intensively studied with focus on the function of a single protein. It is important to consider protein interactions but this possibility is often ignored in previous studies. The first strong protein interaction was discovered between KS and CLF, which led to the formation of a heterodimer.[99] Using coaffinity chromatography and docking simulations, Castaldo *et al.* have also showed that protein interactions are extended, which include the MCAT.[143]

In this work, heterologous over-expression of the individual genes encoding AntD (KS) and AntE (CLF) in *E. coli* has led to the formation of inclusion bodies. After numerous attempts, AntD and AntE could be successfully co-expressed and co-purified in one step as a soluble protein complex. The protein AntD is designed with a N-terminal His$_6$-fusion, while the protein AntE does not carry a His$_6$-fusion. The fact that AntE could still be co-purified with His$_6$-AntD indicated a strong interaction between AntD and AntE.
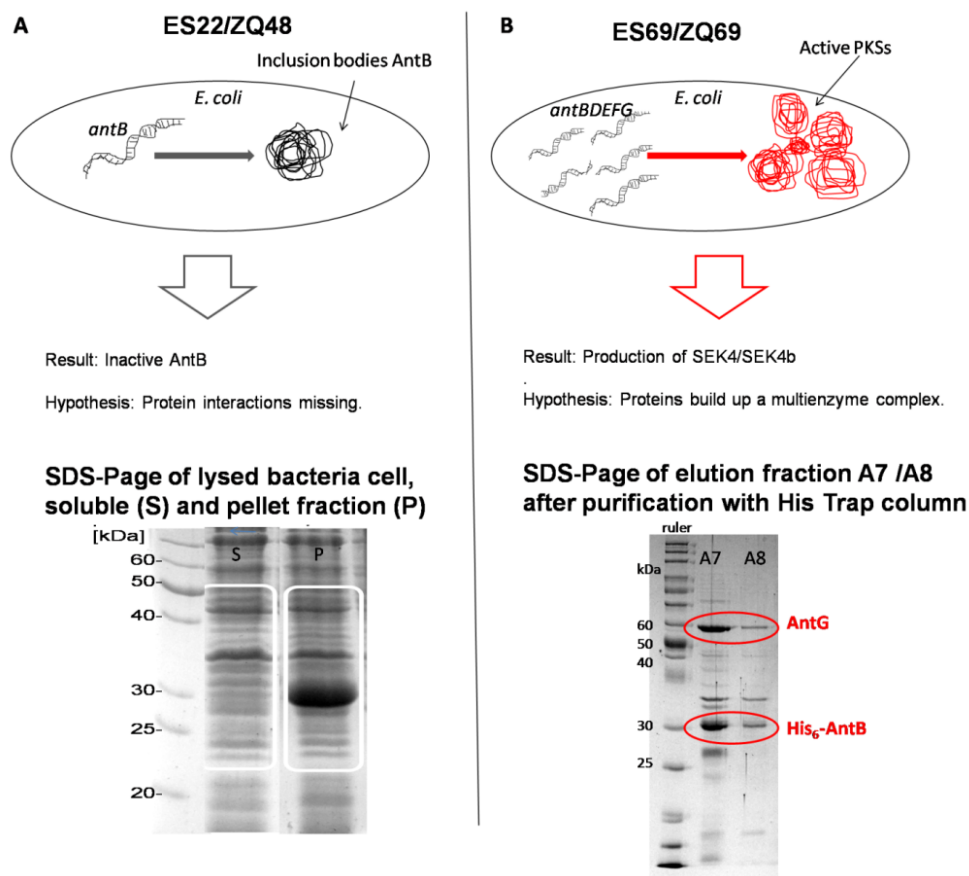

Figure 18. (A) Sole production of AntB using *E. coli* strain ES22 containing the expression-plasmid ZQ48. Plasmid ZQ48 is build from pCOLA Duet vector by insertion of the *antB* gene at the EcoRI/PstI restriction site. (B) Expression AntB together with AntDEFG using *E. coli* strain ES69 containing the expression-plasmid ZQ69. Plasmid ZQ69 is build from the plasmid ZQ48 by insertion of the *antDEFG* genes at the BglII/Acc65I restriction site. Results: (A) AntB was produced as inactive inclusion bodies. (B) Based on the identification of the shunt products SEK4/SEK4b, AntB and AntDEFG were proposed to be produced as active proteins. After the standard purification step, His$_6$-AntB, AntG and some contaminants could be visualized on SDS-Page (B).

The bottleneck of the *in vitro* assays in the AQ biosynthesis is the PPTase AntB, which could be expressed in *E. coli* only as inclusion bodies. Even though the single protein AntB is inactive expressed in *E. coli*, shunt products could be detected in *E.*

coli with genes encoding AntBDEFG. This result demonstrated again the importance of protein interactions, which were proposed to play a crucial role in the stabilization of AntB in the *E. coli* environment (Figure 18). The analysis of the soluble und pellet fraction on the SDS-Page in Figure 18A showed that His_6-AntB was only found in the pellet fraction when it was expressed alone. Subsequently, the co-expression of *antB* with *antDEFG* led to the production of SEK4/SEK4b, indicating that AntBDEFG were produced as active proteins. In particular, AntG could be co-purified with His_6-AntB using the standard purification process (unpublished data, Figure 18B). Some additional contaminants could be also detected on SDS-Page. Therefore, it should be noted that further analyses (peptide mass fingerprinting) are required to verify the nature of the assigned protein bands. For future work, it can be tested if AntB could be co-expressed with AntG or AntDEG.
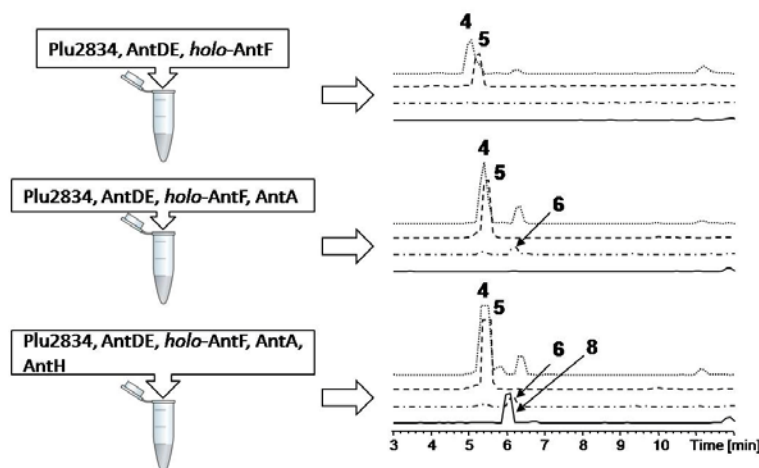
Currently the Ant proteins are crystallized for X-ray analysis in collaboration with the Groll group (Technical University of Munich) alone and in different combinations in order to identify important protein-protein interactions.

In vitro assays

We are interested in reconstruction of the AQ biosynthesis *in vitro*. The first part of the AQ biosynthesis was successfully reproduced using *in vitro* assay in this work, resulting in the production of shunt products SEK4/SEK4b. The following part of the AQ biosynthesis with the KR AntA and the ARO AntH was successfully performed as well, resulting in the production of the shunt products mutactin and SEK34b (unpublished data, Figure 19a). In the future, the first challenge (Figure 19b) of *in vitro* assays is attempting to form the second ring of the polyketide with the CYC AntC. Because of the absence of ammonia, which was proposed to be added into the ACP-bound bicyclic intermediate, azaquinone derivatives might not be produced. For this purpose, ammonia should be considered to be added in the reaction mixture. As a result, shunt products DMAC and aloesaponarin II might be produced and a color change is expected to take place in the test tube. The second challenge is to recover heptaketide AQ-256 from the octaketide intermediate with the help of the hydrolase AntI. The last challenge is to synthesize the methylated AQ using

methyltransferase(s), which have been identified in the *Photorhabdus* genome in our lab recently.
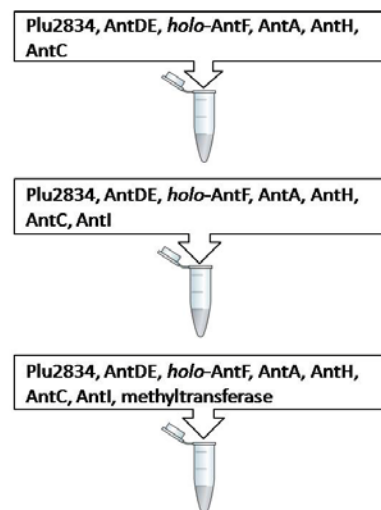

Figure 19. *In vitro* assays performed in this work (a) and challenge in the future (b). Shown are EIC for SEK4 (4, dotted line, m/z 301 [M+H]$^+$), SEK4b (5, dashed line, m/z 319 [M+H]$^+$), mutactin (6, dash-dotted line, m/z 303 [M+H]$^+$) and SEK34b (8, m/z 285 [M+H]$^+$). Structures of shunt products (4, 5, 6 and 8) see Figure 16.

In vitro assays with multi-proteins are challenging due to the necessity of adjusting protein ratios. In an optimal reaction, the last processing protein should be able to completely convert the intermediate to the next shunt product. Improper ratio of the involved proteins might be responsible for the formation of shunt products in different reaction stages. This might be the reason why shunt products SEK4 and SEK4b could be detected in the presence of the KR AntA and the ARO AntC. It should be stated that many factors could influence results from *in vitro* assays. Thus, each assay already performed in this work was supported by a negative control assay. The buffer used for *in vitro* assays was 400 mM Tris-HCl with a pH value of 7.5 at 4 °C. In the future, different buffers of a wide range of pH value are expected to produce contrasting results. Another challenge is to keep all proteins involved in the *in vitro* assays active during storage time. During this work, it appeared that the protein Plu2834 became inactive after storage at -78 °C, leading to many unsuccessful experiments. Although it has happened only once, the storage should be optimized in future work. Moreover, shunt products from *in vitro* assays should be isolated for further detailed characterization.

Heterologous expression of type II PKS in *E. coli*

Hitherto, most type II PKS was found in Gram-positive bacteria of the genus *Streptomyces*. Thus, it is reasonable to reproduce type II PKS-derived polyketides by the heterologous expression in closely related *Streptomyces* hosts.[108] However, the workhorse *E. coli* with lower cost, faster growth rate and well-established genetic manipulation strategy, is still preferred as the host for the heterologous expression.[144] This work pioneered the first heterologous expression of the type II PKS in *E. coli*. Furthermore, co-expression of *ant* genes and the codon optimized *actI-ORF1** gene encoding the KR RED1 involved in the actinorhodin biosynthesis has led to the production of (*S*)-DNPA in *E. coli* for the first time. Gene optimization for the heterologous host using GeneOptimizer® and gene synthesizing using GeneArt® Gene Synthesis (Life Technologies™) might help other type II PKS with the heterologous expression in *E. coli*. Generally, this strategy can be applied for the biosynthesis in different organisms in order to save time and cost.

C-C Hydrolase AntI

In this work, we have demonstrated a chain shortening mechanism catalyzed by the hydrolase AntI. In the absence of AntI, *P. luminescens* and *E. coli* containing *antA-H* genes could only produce octaketides. Thus, AntI was proposed to be involved in the C-C bond cleavage reaction. The chain shortening mechanism could also be confirmed by the *in vitro* assays using synthesized model compounds. Additionally, the 3D-structural modeling of AntI based on the known and structurally related DHPON structure from *Arthrobacter nicotinovorans* supported the shortening mechanism. However, it remained unclear whether the 3D crystal structure of AntI will lead to the same possible catalytic triade consisting of Ser245, His355 and Asp326. According to the modeled structure of AntI, Trp101 and Arg283 showed additional possible interactions with the docked ligand. A further mutagenesis of AntI (AntI_Trp101 and AntI_Arg283) might provide more information concerning whether Trp101 and Arg283 are indispensable for the catalytic reaction.
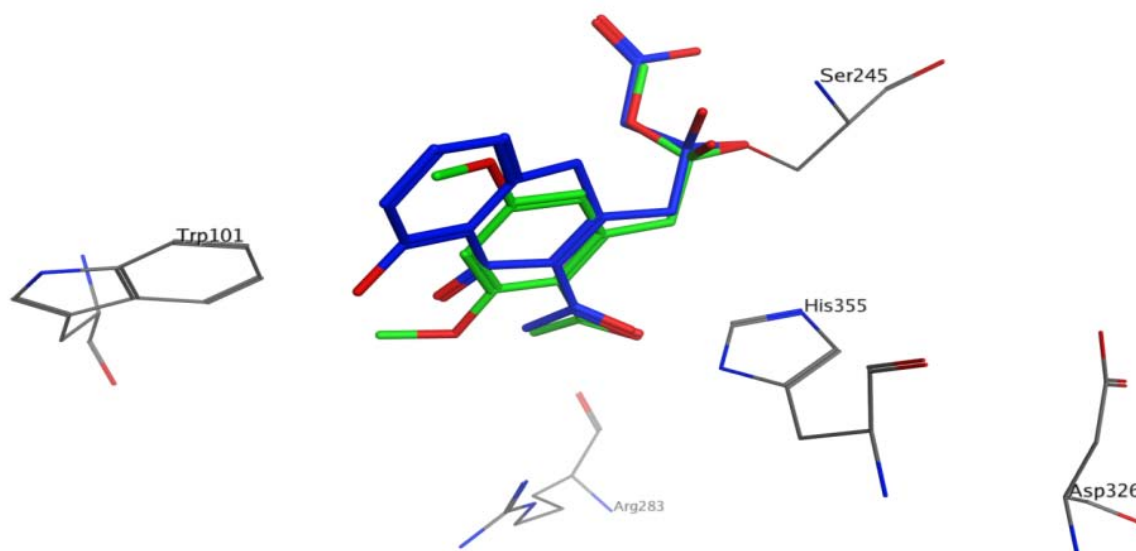
Figure 20. Catalytic triade of the hydrolase AntI and amino acids Trip101 and Arg283 showing possible interaction with the docked ligand.

AntI from the AQ biosynthesis in *P. luminescens* with the C-C breaking function belonged to the α/β-fold hydrolase family with a catalytic triad. Activities of α/β-fold hydrolase enzymes are divided into different categories: C-N, C-O, C-S, C-Halogen and C-C bond breaking. The peptidases (C-N breaking), esterases (C-O breaking) and thioesterases (C-S breaking) are the most well-known α/β-fold hydrolases in biosynthesis. Until now, the C-C breaking mechanism was intensively studied only in the degradation of aromatic compounds.[145] Although the C-C bond breaking hydrolases are not well investigated, they have been identified in many other biosynthetic pathways. The hydrolase PhlG from *Pseudomonas fluorescens* was proven to catalyze the C-C bond cleavage using a zinc ion cofactor.[146] The unusual Pks1 TE from *Colletotrichum lagenarium* could catalyze the Claisen cyclization and the deacetylation.[147] The murine fumarylacetoacetate hydrolase was able to catalyze the hydrolytic cleavage of fumarylacetoacetate, yielding fumarate and acetoacetate.[148] Therefore, a thorough elucidation of the 3D structure of AntI might expand our current understanding on the C-C bond breaking hydrolases.

Reference List

1. Brachmann, A. O.; Joyce, S. A.; Jenke-Kodama, H.; Schwar, G.; Clarke, D. J.; Bode, H. B. A type II polyketide synthase is responsible for anthraquinone biosynthesis in *Photorhabdus luminescens. Chembiochem.* **2007,** *8* (14), 1721-1728.

2. State of World Population 2011: people and possibilities in a world of 7 billion. 2011. http://foweb.unfpa.org/SWP2011/reports/EN-SWOP2011-FINAL.pdf.

3. Low, F.; Lin, H. M.; Gerrard, J. A.; Cressey, P. J.; Shaw, I. C. Ranking the risk of pesticide dietary intake. *Pest. Manag. Sci.* **2004,** *60* (9), 842-848.

4. Gaugler, R.; Kaya, H. K. (1990) *Entomopathogenic nematodes in biological control;* CRC Press, Inc.

5. Goodrich-Blair, H.; Clarke, D. J. Mutualism and pathogenesis in *Xenorhabdus* and *Photorhabdus*: two roads to the same destination. *Mol. Microbiol.* **2007,** *64* (2), 260-268.

6. Herbert, E. E.; Goodrich-Blair, H. Friend and foe: the two faces of *Xenorhabdus nematophila. Nature Reviews Microbiology* **2007,** *5* (8), 634-646.

7. Gaugler, R. (2002) Entomopathogenic nematology. CABI Publishing.

8. Watson, R. J.; Millichap, P.; Joyce, S. A.; Reynolds, S.; Clarke, D. J. The role of iron uptake in pathogenicity and symbiosis in *Photorhabdus luminescens* TT01. *Bmc Microbiology* **2010,** *10*.

9. Bode, H. B. Entomopathogenic bacteria as a source of secondary metabolites. *Curr. Opin. Chem. Biol.* **2009,** *13* (2), 224-230.

10. Latreille, P.; Norton, S.; Goldman, B. S.; Henkhaus, J.; Miller, N.; Barbazuk, B.; Bode, H. B.; Darby, C.; Du, Z.; Forst, S.; Gaudriault, S.; Goodner, B.; Goodrich-

Blair, H.; Slater, S. Optical mapping as a routine tool for bacterial genome sequence finishing. *Bmc Genomics* **2007,** *8*, 321.

11. Duchaud, E.; Rusniok, C.; Frangeul, L.; Buchrieser, C.; Givaudan, A.; Taourit, S.; Bocs, S.; Boursaux-Eude, C.; Chandler, M.; Charles, J. F.; Dassa, E.; Derose, R.; Derzelle, S.; Freyssinet, G.; Gaudriault, S.; Medigue, C.; Lanois, A.; Powell, K.; Siguier, P.; Vincent, R.; Wingate, V.; Zouine, M.; Glaser, P.; Boemare, N.; Danchin, A.; Kunst, F. The genome sequence of the entomopathogenic bacterium *Photorhabdus luminescens. Nat Biotechnol.* **2003,** *21* (11), 1307-1313.

12. Ghazal, S.; Hurst, S. G.; Morris, K.; Abebe-Akele, F.; Thomas, W. K.; Badr, U. M.; Hussein, M. A.; AbouZaied, M. A.; Khalil, K. M.; Tisa, L. S. Draft Genome Sequence of *Photorhabdus luminescens* Strain BA1, an Entomopathogenic bacterium isolated from nematodes found in Egypt. *Genome Announc.* **2014,** *2* (2).

13. 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. Draft genome sequence of *Photorhabdus temperata* strain Meg1, an entomopathogenic bacterium isolated from *Heterorhabditis megidis* nematodes. *Genome Announc.* **2014,** *2* (6).

14. Lanois, A.; Ogier, J. C.; Gouzy, J.; Laroui, C.; Rouy, Z.; Givaudan, A.; Gaudriault, S. Draft genome sequence and annotation of the entomopathogenic bacterium *Xenorhabdus nematophila* strain F1. *Genome Announc.* **2013,** *1* (3).

15. Brachmann, A. O.; Forst, S.; Furgani, G. M.; Fodor, A.; Bode, H. B. Xenofuranones A and B: phenylpyruvate dimers from *Xenorhabdus szentirmaii. Journal of Natural Products* **2006,** *69* (12), 1830-1832.

16. 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.* **2011**.

17. 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* (7), 1115-1119.

18. 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. Structure elucidation and activity of Kolossin A, the D-/L-pentadecapeptide product of a giant nonribosomal peptide synthetase. *Angew. Chem Int. Ed Engl.* **2015,** *54* (35), 10352-10355.

19. Dudnik, A.; Bigler, L.; Dudler, R. Heterologous expression of a *Photorhabdus luminescens* syrbactin-like gene cluster results in production of the potent proteasome inhibitor glidobactin A. *Microbiol. Res.* **2013,** *168* (2), 73-76.

20. Schimming, O.; Challinor, V. L.; Tobias, N. J.; Adihou, H.; Grun, P.; Poschel, L.; Richter, C.; Schwalbe, H.; Bode, H. B. Structure, biosynthesis, and occurrence of bacterial pyrrolizidine alkaloids. *Angew. Chem Int. Ed Engl.* **2015,** *54* (43), 12702-12705.

21. 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* (9), 1290-1294.

22. 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* (11), 2007-2011.

23. Brachmann, A. O.; Reimer, D.; Lorenzen, W.; Augusto, A. E.; Kopp, Y.; Piel, J.; Bode, H. B. Reciprocal cross talk between fatty acid and antibiotic biosynthesis in a nematode symbiont. *Angew. Chem. Int. Ed Engl.* **2012,** *51* (48), 12086-12089.

24. Proschak, A.; Schultz, K.; Herrmann, J.; Dowling, A. J.; Brachmann, A. O.; ffrench-Constant, R.; Muller, R.; Bode, H. B. Cytotoxic fatty acid amides from *Xenorhabdus. Chembiochem.* **2011,** *12* (13), 2011-2015.

25. Paik, S.; Park, Y. H.; Suh, S. I.; 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* (4), 372-374.

26. Ciche, T. A.; Blackburn, M.; Carney, J. R.; Ensign, J. C. Photobactin: a catechol siderophore produced by *Photorhabdus luminescens*, an entomopathogen mutually associated with *Heterorhabditis bacteriophora* NC1 nematodes. *Appl. Environ. Microbiol.* **2003,** *69* (8), 4706-4713.

27. Crawford, J. M.; Portmann, C.; Zhang, X.; Roeffaers, M. B.; Clardy, J. Small molecule perimeter defense in entomopathogenic bacteria. *Proc. Natl. Acad. Sci. U. S. A* **2012,** *109* (27), 10821-10826.

28. 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* (3), 589-597.

29. 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. An antibiotic produced by an insect-pathogenic bacterium suppresses host defenses through phenoloxidase inhibition. *Proc. Natl. Acad. Sci. U. S. A* **2007,** *104* (7), 2419-2424.

30. 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* (12), 4329-4333.

31. 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* (6), 1602-1605.

32. Hu, K.; Li, J.; Li, B.; Webster, J. M.; Chen, G. A novel antimicrobial epoxide isolated from larval *Galleria mellonella* infected by the nematode symbiont, *Photorhabdus luminescens* (Enterobacteriaceae). *Bioorg. Med. Chem.* **2006,** *14* (13), 4677-4681.

33. Joyce, S. A.; Brachmann, A. O.; Glazer, I.; Lango, L.; Schwar, G.; Clarke, D. J.; Bode, H. B. Bacterial biosynthesis of a multipotent stilbene. *Angew. Chem. Int. Ed Engl.* **2008,** *47* (10), 1942-1945.

34. Kronenwerth, M.; Brachmann, A. O.; Kaiser, M.; Bode, H. B. Bioactive derivatives of isopropylstilbene from mutasynthesis and chemical synthesis. *Chembiochem.* **2014,** *15* (18), 2689-2691.

35. 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* (9), 573-578.

36. Brameyer, S.; Kresovic, D.; Bode, H. B.; Heermann, R. Dialkylresorcinols as bacterial signaling molecules. *Proc. Natl. Acad. Sci. U. S. A* **2015,** *112* (2), 572-577.

37. Fuchs, S. W.; Bozhuyuk, K. A.; Kresovic, D.; Grundmann, F.; Dill, V.; Brachmann, A. O.; Waterfield, N. R.; Bode, H. B. Formation of 1,3-cyclohexanediones and resorcinols catalyzed by a widely occurring ketosynthase. *Angew. Chem Int. Ed Engl.* **2013,** *52* (15), 4108-4112.

38. 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.; Phan, K. L.; Thanwisai, A.; Bozhuyuk, K. A.; Chantratita, N.; Gotz, F.; Waterfield, N. R.; Vilcinskas, A.; Stelzer, E. H.; Goodrich-Blair, H.; Hammock, B. D.; Bode, H. B. A *Photorhabdus* natural product inhibits insect juvenile hormone epoxide hydrolase. *Chembiochem.* **2015,** *16* (5), 766-771.

39. Parvatkar, R. R.; D'Souza, C.; Tripathi, A.; Naik, C. G. Aspernolides A and B, butenolides from a marine-derived fungus *Aspergillus terreus. Phytochemistry* **2009,** *70* (1), 128-132.

40. McInerney, B. V.; Gregson, R. P.; Lacey, M. J.; Akhurst, R. J.; Lyons, G. R.; Rhodes, S. H.; Smith, D. R.; 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.

41.   Li, J.; Chen, G.; Webster, J. M.; Czyzewska, E. Antimicrobial metabolites from a bacterial symbiont. *J. Nat Prod* **1995,** *58,* 1081-1086.

42.   Li, B.; Walsh, C. T. *Streptomyces clavuligerus* HlmI is an intramolecular disulfide-forming dithiol oxidase in holomycin biosynthesis. *Biochemistry* **2011,** *50,* 4615-4622.

43.   Shiozawa, H.; Shimada, A.; Takahashi, S. Thiomarinols D, E, F and G, new hybrid antimicrobial antibiotics produced by a marine bacterium; isolation, structure, and antimicrobial activity. *J. Antibiot. (Tokyo)* **1997,** *50* (5), 449-452.

44.   Bouras, N.; Mathieu, F.; Sabaou, N.; Lebrihi, A. Effect of amino acids containing sulfur on dithiolopyrrolone antibiotic productions by *Saccharothrix algeriensis* NRRL B-24137. *J. Appl. Microbiol.* **2006,** *100,* 390-397.

45.   Li, J.; Hu, K.; Webster, J. M. Antibiotics from *Xenorhabdus* spp. and *Photorhabdus* spp. (Enterobacteriaceae) (Review). *Chemistry of Heterocyclic Compounds* **1998,** *34,* 1331-1339.

46.   Hjelmgaard, T.; Givskov, M.; Nielsen, J. Expedient total synthesis of pyrrothine natural products and analogs. *Org. Biomol. Chem.* **2007,** *5* (2), 344-348.

47.   Li, J.; Chen, G.; Webster, J. M.; Czyzewska, E. Antimicrobial metabolites from a bacterial symbiont. *J. Nat Prod* **1995,** *58,* 1081-1086.

48.   Proschak, A.; Zhou, Q.; Schoner, 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* (3), 369-372.

49.   Fuchs, S. W.; Grundmann, F.; Kurz, M.; Kaiser, M.; Bode, H. B. Fabclavines: bioactive peptide-polyketide-polyamino hybrids from *Xenorhabdus*. *Chembiochem.* **2014,** *15* (4), 512-516.

50.   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* (2), 241-248.

51. Seo, S.; Lee, S.; Hong, Y.; Kim, Y. Phospholipase A2 inhibitors synthesized by two entomopathogenic bacteria, *Xenorhabdus nematophila* and *Photorhabdus temperata* subsp. *temperata*. *Appl. Environ. Microbiol.* **2012,** *78* (11), 3816-3823.

52. Ullah, I.; Khan, A. L.; Ali, L.; Khan, A. R.; Waqas, M.; Hussain, J.; Lee, I. J.; Shin, J. H. Benzaldehyde as an insecticidal, antimicrobial, and antioxidant compound produced by *Photorhabdus temperata* M1021. *J. Microbiol.* **2015,** *53* (2), 127-133.

53. Ullah, I.; Khan, A. L.; Ali, L.; Khan, A. R.; Waqas, M.; Lee, I. J.; Shin, J. H. An insecticidal compound produced by an insect-pathogenic bacterium suppresses host defenses through phenoloxidase inhibition. *Molecules.* **2014,** *19* (12), 20913-20928.

54. Derzelle, S.; Duchaud, E.; Kunst, F.; Danchin, A.; Bertin, P. Identification, characterization, and regulation of a cluster of genes involved in carbapenem biosynthesis in *Photorhabdus luminescens*. *Appl. Environ. Microbiol.* **2002,** *68* (8), 3780-3789.

55. Li, J.; Chen, G.; Webster, J. M. Nematophin, a novel antimicrobial substance produced by *Xenorhabdus nematophilus* (Enterobactereaceae). *Can. J. Microbiol.* **1997,** *43* (8), 770-773.

56. 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* (6), 1074-1077.

57. 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.

58. Pinchuk, I. V.; Bressollier, P.; Sorokulova, I. B.; Verneuil, B.; Urdaci, M. C. Amicoumacin antibiotic production and genetic diversity of *Bacillus subtilis* strains isolated from different habitats. *Res. Microbiol.* **2002,** *153* (5), 269-276.

59. Reimer, D.; Luxenburger, E.; Brachmann, A. O.; Bode, H. B. A new type of pyrrolidine biosynthesis is involved in the late steps of xenocoumacin production in *Xenorhabdus nematophila*. *ChemBioChem.* **2009,** *10* (12), 1997-2001.

60. Reimer, D.; Pos, K. M.; Thines, M.; Grun, P.; Bode, H. B. A natural prodrug activation mechanism in nonribosomal peptide synthesis. *Nat Chem Biol* **2011,** *7* (12), 888-890.

61. Reimer, D.; Nollmann, F. I.; Schultz, K.; Kaiser, M.; Bode, H. B. Xenortide biosynthesis by entomopathogenic *Xenorhabdus nematophila*. *J. Nat Prod* **2014**.

62. Crawford, J. M.; Portmann, C.; Kontnik, R.; Walsh, C. T.; Clardy, J. NRPS substrate promiscuity diversifies the xenematides. *Org. Lett.* **2011,** *13* (19), 5144-5147.

63. 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* (15), 1991-1997.

64. 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* (6), 826-828.

65. Gualtieri, M.; Aumelas, A.; Thaler, J. O. Identification of a new antimicrobial lysine-rich cyclolipopeptide family from *Xenorhabdus nematophila*. *J. Antibiot. (Tokyo)* **2009,** *62* (6), 295-302.

66. 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*. *Organic & Biomolecular Chemistry* **2011,** *9* (9), 3130-3132.

67. Bode, H. B.; Reimer, D.; Fuchs, S. W.; Kirchner, F.; Dauth, C.; Kegler, C.; Lorenzen, W.; Brachmann, A. O.; Grun, P. Determination of the absolute configuration of peptide natural products by using stable isotope labeling and mass spectrometry. *Chem. Eur. J.* **2012,** *18* (8), 2342-2348.

68. 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* (2), 205-208.

69. Ohlendorf B, S. S. W. J. I. J. Szentiamide, an N-formylated cyclic depsipeptide from *Xenorhabdus szentirmaii* DSM 16338T. *Nat Prod Commun.* **2011,** *6*, 1247-1250.

70. Nollmann, F. I.; Dowling, A.; Kaiser, M.; Deckmann, K.; Grösch, S.; ffrench-Constant, R.; Bode, H. B. Synthesis of szentiamide, a depsipeptide from entopathogenic *Xenorhabdus szentirmaii* with activity against *Plamodium falciparum. Beilstein Journal of Organic Chemistry* **2012**.

71. 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 Advances* **2013,** *3* (44), 22072-22077.

72. 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* (4), 779-783.

73. Kronenwerth, M.; Bozhuyuk, K. A.; Kahnt, A. S.; Steinhilber, D.; Gaudriault, S.; Kaiser, M.; Bode, H. B. Characterisation of taxlllaids A-G; natural products from *Xenorhabdus indica. Chemistry* **2014,** *20* (52), 17478-17487.

74. 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* (16), 6948-6955.

75. Zhou, Q.; Dowling, A.; Heide, H.; Wohnert, J.; Brandt, U.; Baum, J.; ffrench-Constant, R.; Bode, H. B. Xentrivalpeptides A-Q: depsipeptide diversification in *Xenorhabdus. J. Nat Prod* **2012,** *75* (10), 1717-1722.

76. 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*. *Chemistry* **2013,** *19* (49), 16772-16779.

77. Hertweck, C.; Luzhetskyy, A.; Rebets, Y.; Bechthold, A. Type II polyketide synthases: gaining a deeper insight into enzymatic teamwork. *Natural Product Reports* **2007,** *24* (1), 162-190.

78. King, R. W.; Bauer, J. D.; Brady, S. F. An environmental DNA-derived type II polyketide biosynthetic pathway encodes the biosynthesis of the pentacyclic polyketide erdacin. *Angew. Chem. Int. Ed Engl.* **2009,** *48* (34), 6257-6261.

79. Hutchinson, C. R.; Colombo, A. L. Genetic engineering of doxorubicin production in *Streptomyces peucetius*: a review. *J. Ind. Microbiol. Biotechnol.* **1999,** *23* (1), 647-652.

80. Pickens, L. B.; Tang, Y. Oxytetracycline biosynthesis. *J. Biol Chem.* **2010,** *285* (36), 27509-27515.

81. Malpartida, F.; Hopwood, D. A. Molecular cloning of the whole biosynthetic pathway of a *Streptomyces* antibiotic and its expression in a heterologous host. *Nature* **1984,** *309* (5967), 462-464.

82. McDaniel, R.; Ebert-Khosla, S.; Hopwood, D. A.; Khosla, C. Engineered biosynthesis of novel polyketides. *Science* **1993,** *262* (5139), 1546-1550.

83. Zhou, H.; Li, Y.; Tang, Y. Cyclization of aromatic polyketides from bacteria and fungi. *Nat Prod Rep.* **2010,** *27* (6), 839-868.

84. Rawlings, B. J. Biosynthesis of polyketides (other than actinomycete macrolides). *Natural Product Reports* **1999,** *16* (4), 425-484.

85. Lackner, G.; Schenk, A.; Xu, Z.; Reinhardt, K.; Yunt, Z. S.; Piel, J.; Hertweck, C. Biosynthesis of pentangular polyphenols: deductions from the benastatin and griseorhodin pathways. *J. Am. Chem. Soc.* **2007,** *129* (30), 9306-9312.

86. Taguchi, T.; Yabe, M.; Odaki, H.; Shinozaki, M.; Metsa-Ketela, M.; Arai, T.; Okamoto, S.; Ichinose, K. Biosynthetic conclusions from the functional

dissection of oxygenases for biosynthesis of actinorhodin and related *Streptomyces* antibiotics. *Chem Biol* **2013,** *20* (4), 510-520.

87. Metsa-Ketela, M.; Niemi, J.; Mantsala, P.; Schneider, G. Anthracycline biosynthesis: genes, enzymes and mechanisms. *Top Curr Chem* **2008,** *282*, 101-140.

88. Pickens, L. B.; Tang, Y. Decoding and engineering tetracycline biosynthesis. *Metab Eng* **2009,** *11* (2), 69-75.

89. Kharel, M. K.; Pahari, P.; Lian, H.; Rohr, J. Enzymatic total synthesis of rabelomycin, an angucycline group antibiotic. *Org. Lett.* **2010,** *12* (12), 2814-2817.

90. Hutchinson, C. R. Biosynthetic studies of daunorubicin and tetracenomycin C. *Chem Rev* **1997,** *97* (7), 2525-2536.

91. Cheng, Q.; Xiang, L.; Izumikawa, M.; Meluzzi, D.; Moore, B. S. Enzymatic total synthesis of enterocin polyketides. *Nature Chemical Biology* **2007,** *3* (9), 557-558.

92. White, S. W.; Zheng, J.; Zhang, Y. M.; Rock The structural biology of type II fatty acid biosynthesis. *Annu. Rev Biochem.* **2005,** *74*, 791-831.

93. Crosby, J.; Byrom, K. J.; Hitchman, T. S.; Cox, R. J.; Crump, M. P.; Findlow, I. S. C.; Bibb, M. J.; Simpson, T. J. Acylation of *Streptomyces* type II polyketide synthase acyl carrier proteins. *Febs Letters* **1998,** *433* (1-2), 132-138.

94. 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. *Chemistry & Biology* **1996,** *3* (11), 923-936.

95. Evans, S. E.; Williams, C.; Arthur, C. J.; Burston, S. G.; Simpson, T. J.; Crosby, J.; Crump, M. P. An ACP structural switch: conformational differences between the *apo* and *holo* forms of the actinorhodin polyketide synthase acyl carrier protein. *Chembiochem.* **2008,** *9* (15), 2424-2432.

96. Dall'Aglio, P.; Arthur, C. J.; Williams, C.; Vasilakis, K.; Maple, H. J.; Crosby, J.; Crump, M. P.; Hadfield, A. T. Analysis of *Streptomyces coelicolor* phosphopantetheinyl transferase, AcpS, reveals the basis for relaxed substrate specificity. *Biochemistry* **2011,** *50* (25), 5704-5717.

97. Zhang, W.; Tang, Y. *In vitro* analysis of type II polyketide synthase. *Methods Enzymol.* **2009,** *459*, 367-393.

98. Shen, B. Biosynthesis of aromatic polyketides. *Topic in Current Chemistry* **2000,** *209*.

99. Keatinge-Clay, A.; Maltby, D. A.; Medzihradszky, K. F.; Khosla, C.; Stroud, R. M. An antibiotic factory caught in action. *Nature Structural & Molecular Biology* **2004,** *11* (9), 888-893.

100. Nicholson, T. P.; Winfield, C.; Westcott, J.; Crosby, J.; Simpson, T. J.; Cox, R. J. First *in vitro* directed biosynthesis of new compounds by a minimal type II polyketide synthase: evidence for the mechanism of chain length determination. *Chemical Communications* **2003,** (6), 686-687.

101. Szu, P. H.; Govindarajan, S.; Meehan, M. J.; Das, A.; Nguyen, D. D.; Dorrestein, P. C.; Minshull, J.; Khosla, C. Analysis of the ketosynthase-chain length factor heterodimer from the fredericamycin polyketide synthase. *Chem. Biol* **2011,** *18* (8), 1021-1031.

102. Moore, B. S.; Hertweck, C. Biosynthesis and attachment of novel bacterial polyketide synthase starter units. *Natural Product Reports* **2002,** *19* (1), 70-99.

103. Zhou, H.; Li, Y.; Tang, Y. Cyclization of aromatic polyketides from bacteria and fungi. *Nat. Prod. Rep.* **2010,** *27* (6), 839-868.

104. Ames, B. D.; Korman, T. P.; Zhang, W.; Smith, P.; Vu, T.; Tang, Y.; Tsai, S. C. Crystal structure and functional analysis of tetracenomycin ARO/CYC: implications for cyclization specificity of aromatic polyketides. *Proc. Natl. Acad. Sci. U. S. A* **2008,** *105* (14), 5349-5354.

105. Ames, B. D.; Lee, M. Y.; Moody, C.; Zhang, W.; Tang, Y.; Tsai, S. C. Structural and biochemical characterization of ZhuI aromatase/cyclase from the R1128 polyketide pathway. *Biochemistry* **2011,** *50* (39), 8392-8406.

106. Lee, M. Y.; Ames, B. D.; Tsai, S. C. Insight into the molecular basis of aromatic polyketide cyclization: crystal structure and *in vitro* characterization of WhiE-ORFVI. *Biochemistry* **2012,** *51* (14), 3079-3091.

107. Fritzsche, K.; Ishida, K.; Hertweck, C. Orchestration of discoid polyketide cyclization in the resistornycin pathway. *Journal of the American Chemical Society* **2008,** *130* (26), 8307-8316.

108. Fujii, I. Heterologous expression systems for polyketide synthases. *Nat Prod Rep.* **2009,** *26* (2), 155-169.

109. Fu, H.; Ebertkhosla, S.; Hopwood, D. A.; Khosla, C. Engineered biosynthesis of novel polyketides - dissection of the catalytic specificity of the Act ketoreductase. *Journal of the American Chemical Society* **1994,** *116* (10), 4166-4170.

110. Itoh, T.; Taguchi, T.; Kimberley, M. R.; Booker-Milburn, K. I.; Stephenson, G. R.; Ebizuka, Y.; Ichinose, K. Actinorhodin biosynthesis: structural requirements for post-PKS tailoring intermediates revealed by functional analysis of ActVI-ORF1 reductase. *Biochemistry* **2007,** *46* (27), 8181-8188.

111. Metsa-Ketela, M.; Oja, T.; Taguchi, T.; Okamoto, S.; Ichinose, K. Biosynthesis of pyranonaphthoquinone polyketides reveals diverse strategies for enzymatic carbon-carbon bond formation. *Curr Opin. Chem Biol* **2013,** *17* (4), 562-570.

112. Bringmann, G.; Irmer, A.; Feineis, D.; Gulder, T. A.; Fiedler, H. P. Convergence in the biosynthesis of acetogenic natural products from plants, fungi, and bacteria. *Phytochemistry* **2009,** *70* (15-16), 1776-1786.

113. Li, Y. R.; Chooi, Y. H.; Sheng, Y. W.; Valentine, J. S.; Tang, Y. Comparative characterization of fungal anthracenone and naphthacenedione biosynthetic pathways reveals an alpha-hydroxylation-dependent claisen-like cyclization

catalyzed by a dimanganese thioesterase. *Journal of the American Chemical Society* **2011,** *133* (39), 15773-15785.

114. Pankewitz, F.; Hilker, M. Polyketides in insects: ecological role of these widespread chemicals and evolutionary aspects of their biogenesis. *Biol Rev Camb. Philos. Soc.* **2008,** *83* (2), 209-226.

115. Sendelbach, L. E. A review of the toxicity and carcinogenicity of anthraquinone derivatives. *Toxicology* **1989,** *57* (3), 227-240.

116. van Gorkom, B. A.; de Vries, E. G.; Karrenbeld, A.; Kleibeuker, J. H. Review article: anthranoid laxatives and their potential carcinogenic effects. *Aliment Pharmacol Ther* **1999,** *13* (4), 443-452.

117. Gilbert (nee Stocker), K. G.; Cooke, D. T. Dyes from plants: Past usage, present understanding and potential. *Plant Growth Regulation* **2001,** *34*, 57-69.

118. Avery, M. L.; Humphrey, J. S.; Primus, T. M.; Decker, D. G.; McGrane, A. P. Anthraquinone protects rice seed from birds. *Crop Protection* **1998,** *17*, 225-230.

119. Cichewicz, R. H.; Zhang, Y.; Seeram, N. P.; Nair, M. G. Inhibition of human tumor cell proliferation by novel anthraquinones from daylilies. *Life Sci.* **2004,** *74* (14), 1791-1799.

120. Cota, B. B.; de Oliveira, A. B.; Guimaraes, K. G.; Mendonca, M. P.; de Souza Filho, J. D.; Braga, F. C. Chemistry and antifungal activity of *Xyris* species (Xyridaceae): a new anthraquinone from *Xyris pilosa*. *Biochemical Systematics and Ecology* **2004,** *32*, 391-397.

121. Eyong, K. O.; Folefoc, G. N.; Kuete, V.; Beng, V. P.; Krohn, K.; Hussain, H.; Nkengfack, A. E.; Saeftel, M.; Sarite, S. R.; Hoerauf, A. Newbouldiaquinone A: a naphthoquinone-anthraquinone ether coupled pigment, as a potential antimicrobial and antimalarial agent from *Newbouldia laevis*. *Phytochemistry* **2006,** *67* (6), 605-609.

122. Barnard, D. L.; Huffman, J. H.; Morris, J. L.; Wood, S. G.; Hughes, B. G.; Sidwell, R. W. Evaluation of the antiviral activity of anthraquinones, anthrones

and anthraquinone derivatives against human cytomegalovirus. *Antiviral Res.* **1992,** *17* (1), 63-77.

123.   Han, Y.; Van der Heijden, R.; Verpoorte, R. Biosynthesis of anthraquinones in cell cultures of the Rubiaceae. *Plant Cell, Tissue and Organ Culture* **2001,** *67*, 201-220.

124.   Bringmann, G.; Irmer, A. Acetogenic anthraquinones: biosynthetic convergence and chemical evidence of enzymatic cooperation in nature. *Phytochem Rev* **2008,** *7* (499), 511.

125.   Ames, B. D.; Lee, M. Y.; Moody, C.; Zhang, W. J.; Tang, Y.; Tsai, S. C. Structural and biochemical characterization of Zhul aromatase/cyclase from the R1128 polyketide pathway. *Biochemistry* **2011,** *50* (39), 8392-8406.

126.   Karppinen, K.; Hokkanen, J.; Mattila, S.; Neubauer, P.; Hohtola, A. Octaketide-producing type III polyketide synthase from *Hypericum perforatum* is expressed in dark glands accumulating hypericins. *FEBS J.* **2008,** *275* (17), 4329-4342.

127.   Brachmann, A. O. Dissertation "Isolation and identification of natural products and biosynthetic pathways from *Photorhabdus* and *Xenorhabdus"*. 2009.

128.   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* (11), 1189-1200.

129.   Wenzel, S. C.; Meiser, P.; Binz, T. M.; Mahmud, T.; Muller, R. Nonribosomal peptide biosynthesis: point mutations and module skipping lead to chemical diversity. *Angew. Chem Int. Ed Engl.* **2006,** *45* (14), 2296-2301.

130.   Degenkolb, T.; Karimi, A. R.; Dieckmann, R.; Neuhof, T.; Baker, S. E.; Druzhinina, I. S.; Kubicek, C. P.; Bruckner, H.; von, D. H. The production of multiple small peptaibol families by single 14-module Peptide synthetases in *Trichoderma/Hypocrea*. *Chem Biodivers.* **2012,** *9* (3), 499-535.

131.   Grundmann, F. Dissertation "Natural products from entomopathogenic bacteria". 2013.

132. Sieber, S. A.; Marahiel, M. A. Molecular mechanisms underlying nonribosomal peptide synthesis: approaches to new antibiotics. *Chem Rev* **2005,** *105* (2), 715-738.

133. Roongsawang, N.; Washio, K.; Morikawa, M. *In vivo* characterization of tandem C-terminal thioesterase domains in arthrofactin synthetase. *Chembiochem.* **2007,** *8* (5), 501-512.

134. Hou, J.; Robbel, L.; Marahiel, M. A. Identification and characterization of the lysobactin biosynthetic gene cluster reveals mechanistic insights into an unusual termination module architecture. *Chem Biol* **2011,** *18* (5), 655-664.

135. Li, S.; Dumdei, E. J.; Blunt, J. W.; Munro, M. H.; Robinson, W. T.; Pannell, L. K. Theonellapeptolide IIIe, a new cyclic peptolide from the New Zealand deep water sponge, *Lamellomorpha strongylata*. *J. Nat Prod* **1998,** *61* (6), 724-728.

136. Dornetshuber-Fleiss, R.; Heffeter, P.; Mohr, T.; Hazemi, P.; Kryeziu, K.; Seger, C.; Berger, W.; Lemmens-Gruber, R. Destruxins: fungal-derived cyclohexadepsipeptides with multifaceted anticancer and antiangiogenic activities. *Biochem. Pharmacol* **2013,** *86* (3), 361-377.

137. Zhang, K.; Li, H.; Cho, K. M.; Liao, J. C. Expanding metabolism for total biosynthesis of the nonnatural amino acid L-homoalanine. *Proc. Natl. Acad. Sci. U. S. A* **2010,** *107* (14), 6234-6239.

138. Medema, M. H.; Blin, K.; Cimermancic, P.; de, J., V; Zakrzewski, P.; Fischbach, M. A.; Weber, T.; Takano, E.; Breitling, R. antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res.* **2011,** *39* (Web Server issue), W339-W346.

139. Wilkinson, P.; Waterfield, N. R.; Crossman, L.; Corton, C.; Sanchez-Contreras, M.; Vlisidou, I.; Barron, A.; Bignell, A.; Clark, L.; Ormond, D.; Mayho, M.; Bason, N.; Smith, F.; Simmonds, M.; Churcher, C.; Harris, D.; Thompson, N. R.; Quail, M.; Parkhill, J.; Ffrench-Constant, R. H. Comparative genomics of the emerging human pathogen *Photorhabdus asymbiotica* with the insect pathogen *Photorhabdus luminescens. BMC. Genomics* **2009,** *10*, 302.

140.   Sandmann, A.; Dickschat, J.; Jenke-Kodama, H.; Kunze, B.; Dittmann, E.; Muller, R. A Type II polyketide synthase from the gram-negative bacterium *Stigmatella aurantiaca* is involved in Aurachin alkaloid biosynthesis. *Angew. Chem Int. Ed Engl.* **2007,** *46* (15), 2712-2716.

141.   Yamaguchi, H.; Miyazaki, M. Refolding techniques for recovering biologically active recombinant proteins from inclusion bodies. *Biomolecules.* **2014,** *4* (1), 235-251.

142.   Nilsson, J.; Stahl, S.; Lundeberg, J.; Uhlen, M.; Nygren, P. A. Affinity fusion strategies for detection, purification, and immobilization of recombinant proteins. *Protein Expr. Purif.* **1997,** *11* (1), 1-16.

143. Castaldo, G.; Zucko, J.; Heidelberger, S.; Vujaklija, D.; Hranueli, D.; Cullum, J.; Wattana-Amorn, P.; Crump, M. P.; Crosby, J.; Long, P. F. Proposed arrangement of proteins forming a bacterial type II polyketide synthase. *Chem Biol* **2008,** *15* (11), 1156-1165.

144.   Francis, D. M.; Page, R. Strategies to optimize protein expression in *E. coli*. *Curr Protoc. Protein Sci.* **2010,** *Chapter 5*, Unit-29.

145.   Khajamohiddin, S.; Repalle, E. R.; Pinjari, A. B.; Merrick, M.; Siddavattam, D. Biodegradation of aromatic compounds: an overview of meta-fission product hydrolases. *Crit Rev Microbiol.* **2008,** *34* (1), 13-31.

146.   He, Y. X.; Huang, L.; Xue, Y.; Fei, X.; Teng, Y. B.; Rubin-Pitel, S. B.; Zhao, H.; Zhou, C. Z. Crystal structure and computational analyses provide insights into the catalytic mechanism of 2,4-diacetylphloroglucinol hydrolase PhlG from *Pseudomonas fluorescens. J. Biol Chem* **2010,** *285* (7), 4603-4611.

147.   Vagstad, A. L.; Hill, E. A.; Labonte, J. W.; Townsend, C. A. Characterization of a fungal thioesterase having Claisen cyclase and deacetylase activities in melanin biosynthesis. *Chem Biol* **2012,** *19* (12), 1525-1534.

148.   Timm, D. E.; Mueller, H. A.; Bhanumoorthy, P.; Harp, J. M.; Bunick, G. J. Crystal structure and mechanism of a carbon-carbon bond hydrolase. *Structure.* **1999,** *7* (9), 1023-1033.

9. Table of abbreviations

AQ	anthraquinone
PKS	polyketide synthase
KR	ketoreductase
KS	ketosynthase
CLF	chain length factor
MCAT	malonyl-CoA:ACP transferase
ACP	acyl carrier protein
Ppant	4´-phosphopantetheinyl
PPTase	phosphopantetheinyl transferase
ARO	aromatase
CYC	cyclase
ACT	actinorhodin
BIQ	benzoisochromanequinone
ER	enoylreductase
DH	dehydratase
NRPS	nonribosomal peptide-synthetase
NP	NRPS derived peptide
A	adenylation domain
PCP	peptidyl carrier Protein
C	condensation domain
SAM	*S*-adenosyl-L-methionine
TE	thioesterase
C/E	condensation/epimerization dual domain
HPLC	high performance liquid chromatography
NMR	nuclear magnetic resonance
MS	mass spectrometry
MALDI	matrix-assisted laser desorption ionization
PCR	polymerase chain reaction
SDS-Page	sodium dodecyl sulfate polyacrylamide gel electrophoresis
P.	*Photorhabdus*
X.	*Xenorhabdus*
E. coli	*Escherichia coli*
CoA	coenzyme A
IPS	isopropyl stilbene
Xcn1	xenocoumacin-1
Xcn2	xenocoumacin-2
(*S*)-DNPA	4-dihydro-9-hydroxy-1-methyl-10-oxo-3-H-naptho-[2,3-c]-pyran-3-(*S*)-aceticacid
DMAC	3,8-dihydroxy-1-methylanthraquinone-2-carboxylic acid

10. Record of conferences and list of publications

Oral presentation:

Zhou, Q., Bode, B. H. Investigation of the type II polyketide synthase from Gram-negative bacteria *Photorhabdus luminescence* TT01. Vereinigung für Allgemeine und Angewandte Mikrobiologie (VAAM) Jahrestagung, 18-21 March 2012, Tübingen (Germany).

Poster presentations and short oral presentations

Zhou, Q.; Bode, B. H. Biosynthesis of Anthraquinones in *Photorhabdus luminescence*. Vereinigung für Allgemeine und Angewandte Mikrobiologie (VAAM) International Workshop "Biology of Bacteria Producing Natural Products", September 2010, Tübingen (Germany).

Zhou, Q.; Bode, B. H. Analysis of Anthraquinones Biosynthesis in *Photorhabdus*. Vereinigung für Allgemeine und Angewandte Mikrobiologie (VAAM) Jahrestagung, April 2011, Karlsruhe (Germany).

Zhou, Q.; Bode, B. H. Isolation and identification of PATA-Peptides produced by *Xenorhabdus*. Vereinigung für Allgemeine und Angewandte Mikrobiologie (VAAM) International Workshop "Biology of Bacterial Producers of Natural Compounds", September, 2011 Bonn (Germany).

Zhou, Q.; Adihou, H.; Kresovic, D.; Bozhüyük, K. A. J.;Bode, H.B. Biosynthesis of anthraquinones in *Photorhabdus luminescens*. Vereinigung für Allgemeine und Angewandte Mikrobiologie (VAAM) International Workshop "Biology of Bacterial Producers of Natural Compounds", September 2015, Frankfurt (Germany).

List of Publications:

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.

Proschak, A.; **Zhou, Q.**; Schoner, 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 (3)*, 369-372.

Zhou, Q.; Dowling, A.; Heide, H.; Wohnert, J.; Brandt, U.; Baum, J.; ffrench-Constant, R.; Bode, H. B. Xentrivalpeptides A-Q: depsipeptide diversification in *Xenorhabdus*. *J. Nat Prod* **2012**, *75 (10)*, 1717-1722.

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*. *Chemistry* **2013**, *19 (49)*, 16772-16779.

Manuscript for publication

Zhou, Q.; Adihou, H.; Kresovic D.; Bozhüyük K. A. J.; Bode H. B. Unusual start and finish of anthraquinone biosynthesis in Photorhabdus luminescens. **2016** submitted. status: major revision

11. Attachment: declaration on the contribution of the author, publications and manuscript

11.1. Xentrivalpeptides A−Q: depsipeptide diversification in *Xenorhabdus*

11.2. Structure and biosynthesis of xenoamicins from entomopathogenic *Xenorhabdus*

11.3. Unusual start and finish of anthraquinone biosynthesis in *Photorhabdus luminescens*

11.4. Biosynthesis of the insecticidal xenocyloins in *Xenorhabdus bovienii*

11.5. Simple "on-demand" production of bioactive natural products

11.1. Xentrivalpeptides A−Q: Depsipeptide Diversification in *Xenorhabdus*

85

Authors: Qiuqin Zhou†, Andrea Dowling‡, Heinrich Heide§, Jens Wöhnert†, Ulrich Brandt§, James Baum⊥, Richard Ffrench-Constant‡, and Helge B. Bode*†

† Institut für Molekulare Biowissenschaften, Goethe Universität Frankfurt, 60438 Frankfurt am Main, Germany

‡ Biosciences, University of Exeter in Cornwall, Tremough Campus, Penryn, Cornwall TR10 9EZ, U.K.

§ Zentrum für Biologische Chemie, Molekulare Bioenergetik, Klinikum der Goethe Universität Frankfurt, 60590 Frankfurt am Main, Germany

⊥ Monsanto Company, Chesterfield, Missouri 63017, United States

Attachments: Declaration on the contribution of the author and the paper.

Erklärung zu den Autorenanteilen

an der Publikation: "Xentrivalpeptides A−Q: Depsipeptide Diversification in *Xenorhabdus"*

Status: published in *J. Nat. Prod.*, **2012**, *75 (10)*, 1717–1722

Beteiligte Autoren: Qiuqin Zhou (QZ), Andrea Dowling (AD), Heinrich Heide (HH), Jens Wöhnert (JW), Ulrich Brandt, James Baum, Richard Ffrench-Constant, Helge B. Bode (HBB)

Was hat der Promovierende bzw. was haben die Koautoren beigetragen?

(1) zu Entwicklung und Planung

Promovierender: 50 %
Co-Autor HBB: 50 %

(2) zur Durchführung der einzelnen Untersuchungen und Experimente

Promovierender: 84 %, alle Experimente außer HRESI-MS, der NMR Messung, dem Bioaktivitätstest, der Messung von Schmelzpunkt und UV/IR
Co-Autor HH: 10 %, HRESI-MS
Co-Autor JW: 3 % NMR Messung und Messung von Schmelzpunkt und UV/IR
Co-Autor AD: 3 %, Bioaktivitätstest

(3) zur Erstellung der Datensammlung und Abbildungen

Promovierender: 95 %, ausschließlich außer dem Bioaktivitätstest
Co-Autor AD: 5 %, Bioaktivitätstest

(4) zur Analyse und Interpretation der Daten

Promovierender: 95 %, ausschließlich außer dem Bioaktivitätstest
Co-Autor AD: 5 %, Bioaktivitätstest

(5) zum Verfassen des Manuskripts

Promovierender: 65 %
Co-Autor HBB: 35 %

Unterschrift Promovend: _____ Datum/Ort: _____

Zustimmende Bestätigungen der oben genannten Angaben

Unterschrift Betreuer: _____ Datum/Ort: _____

Xentrivalpeptides A−Q: Depsipeptide Diversification in *Xenorhabdus*

Qiuqin Zhou,[†] Andrea Dowling,[‡] Heinrich Heide,[§] Jens Wöhnert,[†] Ulrich Brandt,[§] James Baum,[⊥] Richard ffrench-Constant,[‡] and Helge B. Bode*,[†]

[†]Institut für Molekulare Biowissenschaften, Goethe Universität Frankfurt, 60438 Frankfurt am Main, Germany

[‡]Biosciences, University of Exeter in Cornwall, Tremough Campus, Penryn, Cornwall TR10 9EZ, U.K.

[§]Zentrum für Biologische Chemie, Molekulare Bioenergetik, Klinikum der Goethe Universität Frankfurt, 60590 Frankfurt am Main, Germany

[⊥]Monsanto Company, Chesterfield, Missouri 63017, United States

S *Supporting Information*

ABSTRACT: Seventeen depsipeptides, xentrivalpeptides A−Q (**1**−**17**), have been identified from an entomopathogenic *Xenorhabdus* sp. Whereas the structure of xentrivalpeptide A (**1**) was determined after its isolation by NMR spectroscopy and the advanced Marfey's method, the structures of all other derivatives were determined using a combination of stable isotope labeling and detailed MS analysis.

RESULTS AND DISCUSSION
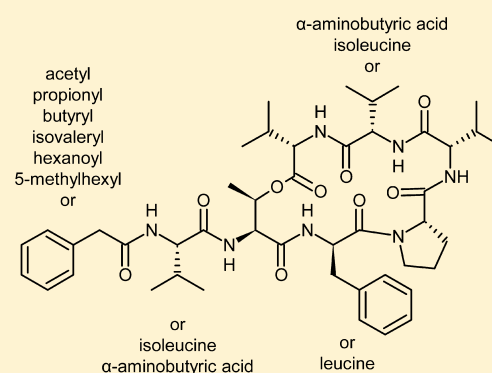
Entomopathogenic bacteria of the genera *Xenorhabdus* and *Photorhabdus* live in symbiosis with nematodes of the genera *Steinernema* and *Heterorhabditis*, respectively, and together with them are able to infect and kill different insect larvae.[1−4] Although protein toxins are mainly responsible for larval mortality,[5,6] small molecules that are toxic to insect cells have also been identified.[7] It has been postulated that these bacteria produce antibiotics to protect the insect cadaver from competitors living in the soil, and compounds showing these activities have been identified recently.[7−9] Besides several small molecules such as isopropylstilbenes,[10,11] anthraquinones,[11,12] and xenofuranones,[13] recent work has shown that several peptides are produced by these bacteria. Among them are the highly polar PAX peptides,[14] the GameXPeptides,[15] and the depsipeptides xenematide[16,17] and szentiamide.[18] Of these, the PAX peptides,[14,19] GameXPeptides,[15] and xenematides[16,17] are produced as mixtures of several derivatives. Thus, *Xenorhabdus* and *Photorhabdus* have proven to be a rich source of bioactive natural products, and our goal is to identify additional bioactive natural products from them. Here we describe the identification and structure elucidation of the xentrivalpeptides, depsipeptides that show a much higher chemical diversity than the xentiamides[16,17] and szentiamide,[18] the two known depsipeptide classes from *Xenorhabdus*.

RESULTS AND DISCUSSION

During our search for new secondary metabolites from *Xenorhabdus* bacteria, we identified compound **1** with *m/z* 860.5 in extracts of *Xenorhabdus* sp. 85816 obtained from the

Monsanto Company when grown in LB medium with 2% Amberlite XAD-16. From a 5 L culture grown under the same conditions, 22 mg of **1** was isolated from the XAD extract using preparative HPLC/MS. The molecular formula of **1** was determined from HRESIMS analysis (*m/z* 860.4880) as $C_{46}H_{65}N_7O_9$ (Table S1, Figure S1), and the structure of **1** was determined by detailed 1D (^{1}H,^{13}C) and 2D (COSY, HSQC, HMBC, TOCSY) NMR experiments (Table 1). Eight different spin systems were identified on the basis of coupling constants in the ^{1}H NMR spectrum and COSY data (Figure 1a). Several signals around 1 ppm were assigned to methyl groups from four valines and one threonine, and in addition, two phenyl groups were identified on the basis of the typical chemical shifts of δ_H 7.20−7.30 with total integration of 10 protons and typical δ_C shifts between 127 and 138 in the ^{13}C NMR spectrum. Each spin system could be connected with one of eight quaternary carbons in the range δ_C 169.5−174.9 by HMBC correlations (Figure 1b). In this way, the linear sequence derived from the connections of α-H with quaternary carbons and the connection between the β-H in Thr (3) and the quaternary carbon in Val (7) established the structure of the depsipeptide with a six-membered ring (Figure 1b).

The absolute configurations of the amino acids were determined using the advanced Marfey's method,[20] showing that only phenylalanine has the D-configuration (Table S2). We

Table 1. NMR Spectroscopic Data (500 MHz (^{1}H), 125 MHz (^{13}C) in CD$_3$OD) of 1, δ in ppm^a

subunit	position	δ_C	δ_H, mult (J in Hz)
PA	1	174.2	
	2	43.6	3.60, d (1.0)
	3	137.0	
	4	130.3^b	7.29, m
	5	129.7	7.30, m
	6	128.2^c	7.25, m
L-Val(1)	C=O	172.8	
	α	60.2	4.25, d (7.3)
	β	31.7	2.10, m
	γ	19.8	0.94, d (6.7)
	δ	18.6^d	0.92, d (6.8)
L-Thr(2)	C=O	169.5	
	α	55.0	4.79, d (3.8)
	β	69.9	5.09, m
	γ	14.6	1.14, d (2.4)
D-Phe(3)	C=O	171.3	
	α	55.3	4.94, dd (5.2, 9.9)
	β1	40.5	2.91, dd (9.8, 12.6)
	β2	40.5	3.24, dd (5.0, 12.5)
	γ	137.0	
	δ	130.6^b	7.27–7.29
	ε	129.7	7.27–7.29
	ζ	127.9^c	7.20–7.24
L-Pro(4)	C=O	174.9	
	α	62.4	4.09, dd (4.9, 8.7)
	β1	30.7	1.79, m
	β2	30.7	1.89, m
	γ1	25.5	1.47, m
	γ2	25.5	1.79, m
	δ1	48.6	2.68, m
	δ2	48.6	3.44, m
L-Val(5)	C=O	173.6	
	α	65.5	3.64, d (11.3)
	β	30.8	2.31, m
	γ	20.3	0.89, d (6.4)
	δ	20.1	0.98, d (6.4)
L-Val(6)	C=O	173.5	
	α	58.0	4.42, d (5.9)
	β	33.6	2.04, m
	γ	18.7^d	0.93, d (7.0)
	δ	19.7	0.99, d (7.0)
L-Val(7)	C=O	172.0	
	α	62.4	3.94, d (7.7)
	β	30.2	2.07, m
	γ	19.8	1.13, d (6.7)
	δ	18.9	1.04, d (6.8)

^{a}PA: phenyl acetyl; d: doublet; m: multiplet. $^{b-d}$Assignments are interchangeable.
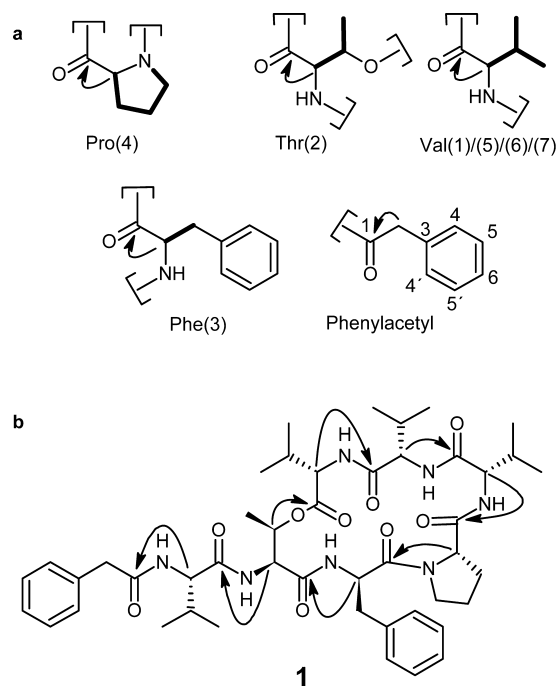
Figure 1. Subunits (a) with selected COSY (bold lines) and HMBC correlations (arrows) in xentrivalpeptide A (**1**) (b). Amino acid numbering from N- to C-terminus.

Table 2. Data for Xentrivalpeptides A–Q (1–17)

no.	compound	t_R/min	amount^a	fragment B^b (m/z)
1	A	10.4	100.0	625.3
2	B	6.4	1.1	
3	C	7.6	<1.0	
4	D	8.8	3.0	
5	E	9.8	2.1	
6	F	12.7	<1.0	
7	G	11.5	<0.1	
8	H	9.8	<1.0	
9	I	11.3	3.5	
10	J	8.2	<1.0	591.3
11	K	9.2	<1.0	
12	L	9.8	2.1	
13	M	12.2	<0.1	
14	N	12.3	3.7	639.3
15	O	9.8	<1.0	611.3
16	P	7.6	<0.1	
17	Q	8.5	<0.1	526.3^c

^{a}Amount relative to **1** calculated from peak areas in HPLC/MS. ^{b}For details on fragment B see Figure 2. ^{c}**17** has a smaller ring than the other derivatives (see structure below and Figure S7). Retention time (t_R).

analyzed strain 85816 for other derivatives of this class of depsipeptides and could indeed identify several, which we named xentrivalpeptides because they were isolated from a _Xenorhabdus_ strain and all contained at least three (tri) valines. All xentrivalpeptides showed similarity to **1** based on the MS fragmentation patterns (Table 2, Table S1). We performed extensive MS fragmentation experiments with **1** (Table S1, Figure S1), which revealed a general fragmentation pattern for this class of peptides (Figure 2) used for the structure elucidation of the other xentrivalpeptides. Additionally, we

applied a combination of labeling and detailed MS analysis to strain 85816, as previously described, to differentiate isobaric building blocks such as leucine and isoleucine.[15] From the results of labeling experiments using deuterated or fluorinated building blocks in LB medium, and an inverse labeling experiment adding nonlabeled ^{12}C building blocks to a culture grown in fully ^{13}C-labeled medium, the building blocks of **1** could be confirmed (Figure 3). For example, labeling with _p_-fluorophenylalanine or _p_-fluorophenylacetic acid in LB medium showed the expected +18 Da shift to m/z 878.5 (Figure 3b and c), and labeling with phenylalanine in a ^{13}C-labeled culture
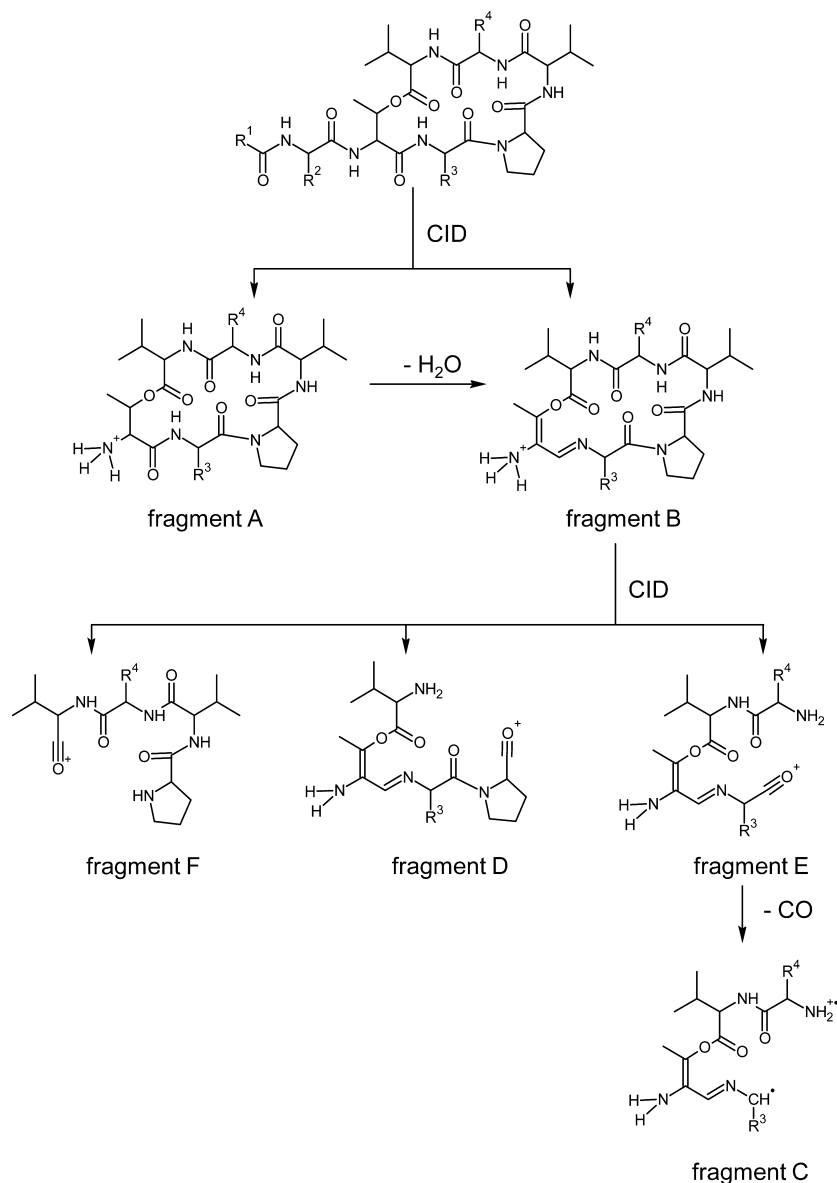
Figure 2. General fragmentation pathway of xentrivalpeptides A–P (**1**–**16**) and the proposed fragment structures. For fragment data see Table S1.

showed the expected shift to a lower mass due to the incorporation of nine and eight carbons from incorporated phenylalanine and phenylacetic acid, respectively (Figure 3j). Moreover, from cultivation in ^{13}C and ^{15}N medium the number of carbon and nitrogen atoms of **1** could be determined easily, thus affording the correct molecular formula as shown previously[15] (Figure 3f and g).

The analysis of these labeling experiments followed by MS fragmentation experiments (Figures S2–S5) allowed the structure elucidation of xentrivalpeptides B–P (**2**–**16**). The absolute configurations of the amino acids were not determined experimentally due to their low production titer (Table 2), but are assumed to follow that of **1**.

The differences between **1**–**9** were only in the side chain, as shown by an identical ring fragment B of m/z 625.3 (Figure 2, Table S1) and confirmed by labeling experiments showing the presence of the same ring amino acids as in **1** (Figure S2). Xentrivalpeptides A–G (**1**–**7**) differed from each other at R^1 (acyl moiety) only. The nature of R^1 in these derivatives was unambiguously determined from labeling experiments: D_6-

propionic acid was incorporated into **3**, as detected from a mass shift of +5 Da from m/z 798.5 to 803.5, ^{12}C-butyric acid was incorporated in a ^{13}C-labeled culture into **4**, as shown by a shift of −4 Da from m/z 854.5 (U-^{13}C-labeled **4**) to m/z 850.5, and five carbons from the ^{12}C-leucine-derived isovaleryl unit were incorporated into **5** and **6** in ^{13}C medium due to the presence of isovaleryl (in **5**) or 5-methylhexyl acyl groups (in **6**), resulting in a −5 Da shift in both compounds. Additionally, a sufficient amount of **4** was isolated allowing full characterization via NMR spectroscopy (Table S3, Figure S6). The presence of an acetyl moiety in **2** was deduced from the HRESIMS data, and the valine labeling results showed that one valine is present in the side chain (amino acid two), thus securing the acetyl group as the acyl moiety. The presence of a hexanoyl moiety in **7**, which was produced only in very minor amounts (Table 2), was proposed, as neither four carbons from valine (as isobutyryl) nor leucine were incorporated into the acyl chain and assuming that the acyl moiety is not branched (data not shown). For **8** and **9** the incorporations of α-aminobutyric acid and isoleucine in ^{13}C medium were detected at amino acid 1
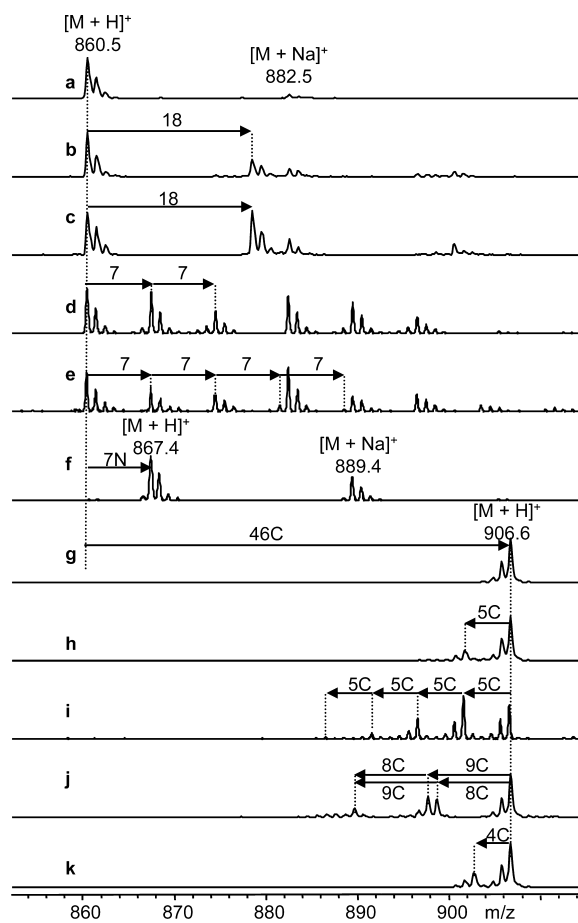
Figure 3. MS spectra from labeling experiments of xentrivalpeptide A (**1**). LB medium (^{12}C, a), LB medium with *p*-fluoro-DL-phenylalanine (^{12}C + *p*-F-Phe, b), with *p*-fluorophenylacetic acid (^{12}C + *p*-F-PAA, c), with additional L-[2,3,3,5,5′,6,6′,7-^{2}H]phenylalanine (^{12}C + ^{2}H$_8$-Phe, d), with DL-[2,3,4,4,4,5,5,5-^{2}H$_8$]valine (^{12}C + ^{2}H$_8$-Val, e), ^{15}N medium (^{15}N, f), ^{13}C medium (^{13}C, g), ^{13}C medium with L-proline (^{13}C + Pro, h), with L-valine (^{13}C + Val, i), with L-phenylalanine (^{13}C + Phe, j), and with L-threonine (^{13}C + Thr, k), respectively.
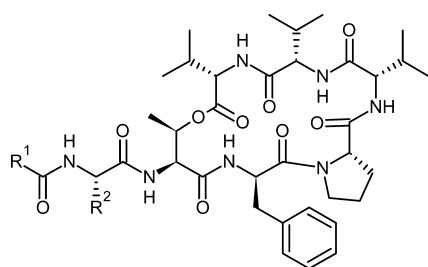
	R^1	R^2
xentrivalpeptide A (**1**)	Bn	*i*Pr
xentrivalpeptide B (**2**)	Me	*i*Pr
xentrivalpeptide C (**3**)	Et	*i*Pr
xentrivalpeptide D (**4**)	Pr	*i*Pr
xentrivalpeptide E (**5**)	*i*Bu	*i*Pr
xentrivalpeptide F (**6**)	*i*hexyl	*i*Pr
xentrivalpeptide G (**7**)	*n*pentyl	*i*Pr
xentrivalpeptide H (**8**)	Bn	Et
xentrivalpeptide I (**9**)	Bn	*sec*Bu

instead of the usual valine due to the mass shifts of 4 and 6 Da, respectively, and an unchanged fragment B for both compounds (Figure S2). As isoleucine is incorporated into **9**, the configuration in the isoleucine side chain is proposed to be *S*, as for the natural amino acid. The R^1 group was determined to be phenylacetyl in **8** and **9** from the 8 Da mass shifts from ^{12}C-phenylalanine in the ^{13}C medium and incorporation of *p*-fluorophenylacetic acid (Figure S2).
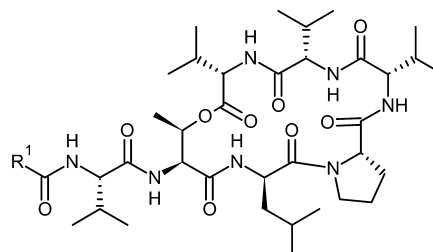
	R^1
xentrivalpeptide J (**10**)	Pr
xentrivalpeptide K (**11**)	*i*Bu
xentrivalpeptide L (**12**)	Bn
xentrivalpeptide M (**13**)	*i*hexyl

Fragmentation of xentrivalpeptides J–M (**10**–**13**) revealed a fragment B with *m/z* 591.3 (Table S1). The difference of −34 Da compared to fragment B of **1** points to a leucine/isoleucine instead of a phenylalanine, which was confirmed by the incorporation of D$_{10}$-leucine into **10**–**13** (Figure S3) and no incorporation of phenylalanine into fragment B in these compounds. Moreover, MS fragmentation proved that **10**–**12** showed an identical fragmentation of fragment B with fragments D and E showing the mass shift of 34 Da. As valine incorporation could be observed in the side chain for **10**–**13**, these compounds again differed only in the acyl moiety, whose nature was assigned from the incorporation of butyric acid in **10**, leucine-derived isovalerate incorporation in **11** and **13**, and the incorporation of 4-fluorophenylacetic acid or phenylalanine in **12** (Figure S3), as described above for **4**, **5**, and **1**, respectively.

Xentrivalpeptide N (**14**, C$_{47}$H$_{67}$N$_7$O$_9$) showed a fragment B (Table S1) with *m/z* 639.3, indicating the presence of leucine/isoleucine instead of valine in the ring, as was confirmed by labeling with isoleucine in fully ^{13}C-labeled medium and by the expected mass shift of 6 Da in fragment B (Figure S4). The position of the isoleucine was readily identified from MS fragmentation experiments that confirmed the 14 Da mass shift only in fragments C, E, and F (Table S1, Figure 2), which is consistent with isoleucine (2*S*,3*S* configuration as in **9**) as amino acid 6. All other feeding experiments confirmed that **14** is otherwise identical to **1**.

Fragment B of xentrivalpeptides O (**15**) and P (**16**) showed *m/z* 611.3 (Table S1), indicating the loss of a methyl group compared to **1**. Thus the presence of α-aminobutyric acid instead of valine was proposed. This was confirmed in the respective feeding experiment (Figure S5). Fragments C, E, and F showed the −14 Da mass shift compared to **1**, thus again confirming amino acid 6 as the variable position (Table S1). Whereas the structure of **15** could be confirmed from labeling experiments and MS fragmentation experiments (Figure S5), labeling of **16** was very weak. However, as no labeling of leucine or isoleucine was observed and butyrate is an abundant starting
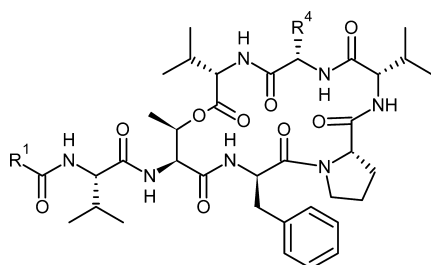
	R[1]	R[4]
xentrivalpeptide N (14)	Bn	*sec*Bu
xentrivalpeptide O (15)	Bn	Et
xentrivalpeptide P (16)	Pr	Et

unit for xentrivalpeptides (Table 2), we postulate that **16** has a butyryl unit as found in **4**.

The molecular formula of $C_{41}H_{56}N_6O_8$ for xentrivalpeptide Q (**17**) as determined by HRESIMS indicated the loss of one valine moiety compared to **1**. Fragmentation for **17** differed significantly from that of **1**−**16**. Fragments A′ and B′ with m/z 544.3 and 526.3, which correspond to fragments A and B in **1**−**16**, showed the expected loss of one valine moiety due to the loss of 99 Da. Additional fragments confirmed that only one valine in the ring is missing (Figure S7). Thus, **17** was confirmed to be a ring-contracted derivative of **1**.
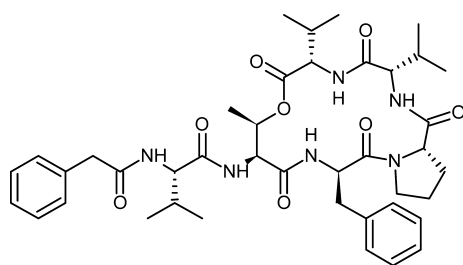
xentrivalpeptide Q (17)

The xentrivalpeptides are the third and by far the most diverse class of depsipeptides isolated from *Xenorhabdus*[8,16−18] and thus a nice example of natural combinatorial biochemistry, which in addition to the N-acyl variability is in part due to relaxed amino acid specificities in the corresponding adenylation (A) and condensation (C) domains of the nonribosomal peptide synthetase (NRPS) responsible for biosynthesis.[21−23] Whereas this was sometimes interpreted as a "mistake" of the respective enzymes in the past, it might in fact be beneficial for the producing organism to generate a large chemical diversity with a minimal set of building blocks and enzymes.[24] As these derivatives might also have different biological activities, the resulting compounds might allow *Xenorhabdus* to kill several different insect larvae as well as to protect them against different food competitors. However, so far no biological activity has been detected for **1** in standard bioactivity tests (antibiotic Gram-negative or Gram-positive, antifungal, cytotoxic against eukaryotic cells). Whereas biological activity for other *Xenorhabdus*-derived compounds has been detected at low concentrations,[7] only a high concentration (100 μg/mL) of **1** led to actin ruffling, pointing to the cytoskeleton as the target for the xentrivalpeptides (Figure S8), as will be investigated in the future in more detail.

Because the xentrivalpeptides lack any unusual amino acids, their synthesis should be facile using solid-phase methods.[8]

Therefore, current efforts in our group are directed toward synthesizing the xentrivalpeptides to verify the proposed absolute configuration and to broaden the scope of biological testing.

■ EXPERIMENTAL SECTION

General Experimental Procedures. Melting points were determined with a MPM-H2 melting point meter by Schropp Gerätetechnik and are uncorrected. Optical rotations were measured on a Perkin-Elmer 341 polarimeter. UV spectra were recorded with a GE NanoVue Plus photometer. IR spectra were obtained with a Perkin-Elmer Spectrum Two spectrometer. NMR spectra were recorded with a Bruker AV 400 spectrometer using deuterated methanol as solvent. Collision-induced dissociation (CID) was performed on the ion trap in the amaZon X in positive ion mode, and HRESIMS analysis was carried out using an LTQ Orbitrap (Thermo Fisher). Preparative HPLC was performed using an Waters autopurification system.

Strain Cultivation and Extract Preparation. *Xenorhabdus* sp. 85816 was identified as *X. stockiae* based on its *recA* sequence (Genbank accession number JX485977) and was cultivated in LB medium at 30 °C. For extract analysis, 20 mL of LB medium was inoculated with 1% overnight culture, and 2% Amberlite XAD-16 was added. After cultivation for 72 h, the XAD beads were collected, washed once with 5 mL of H_2O, and extracted with 20 mL of MeOH. To isolate the peptides, XAD beads from a 5 L culture were extracted with 200 mL of MeOH three times. The extract was fractionated on silica gel with $CHCl_3$ and MeOH with a linear gradient of 5% to 40% MeOH. The fractions were analyzed by HPLC-MS, and fractions eluted with ∼20% MeOH contained the xentrivalpeptides. The enriched fraction containing xentrivalpeptides was used for HPLC purification.

Feeding Experiments in Nonlabeled Medium. The cell pellet from a 1 mL overnight *Xenorhabdus* sp. 85816 culture was washed once and resuspended with 1 mL of LB medium. The 5 mL feeding culture was inoculated with a 1% (50 μL) solution of washed cells. After incubation for 6 h at 30 °C, 200 rpm, 50 μL of a stock solution (100 mM) of substrate (p-fluoro-DL-phenylalanine, p-fluorophenyl-acetic acid, L-[2,3,3,5,5′,6,6′,7-$^2H_8$]phenylalanine, DL-[2,3,4,4,4,5,5,5-$^2H_8$]valine, L-[2,3,3,4,5,5,5,5′,5′,5′-$^2H_{10}$]leucine, or [U-$^2H_6$]propionic acid) was added. Two further feedings of substrate were carried out after 24 and 48 h to a final concentration of 3 mM. Cultures were harvested after 72 h of incubation by extraction with 5 mL of EtOAc. The extracts were evaporated to dryness and dissolved in 500 μL of MeOH. Diluted solutions were analyzed by HPLC-MS, and control cultivation was carried out without feeding.

Feeding Experiments in Labeled Medium. The cell pellet from a 1 mL overnight *Xenorhabdus* sp. 85816 culture was washed once and resuspended with 1 mL of ISOGRO-^{13}C medium. ISOGRO-^{13}C medium was prepared with 1 g of ISOGRO-^{13}C powder (Sigma-Aldrich), 1.8 g/L K_2HPO_4, 1.4 g/L KH_2PO_4, and 11.1 mg/L $CaCl_2 \cdot H_2O$ in 100 mL of H_2O. The feeding culture was started by inoculation of a 1% (50 μL) solution of washed cells in 5 mL of ISOGRO-^{13}C medium. After incubation for 6 h at 30 °C, 200 rpm, 50 μL of a stock solution (100 mM) of substrate (L-α-aminobutyric acid, butyric acid, L-α-valine, L-α-leucine, L-α-proline, L-α-threonine, L-α-phenylalanine, or L-α-isoleucine) was added. Two further feedings of every substrate were carried out after 24 and 48 h of incubation to a final concentration of 3 mM. Cultures were harvested after 72 h of incubation by extraction with 5 mL of EtOAc. The extracts were evaporated to dryness and dissolved in 500 μL of MeOH. Diluted solutions were analyzed by HPLC-MS. Control cultivation was carried out without feeding. *Xenorhabdus* sp. 85816 was also cultivated in ^{15}N-labeled medium without feeding of substrate. The ^{15}N-labeled medium was prepared in the same manner as the ^{13}C-labeled medium.

HPLC and Mass Spectrometry. Analysis of the extracts was carried out on an Ultimate 3000 LC system from Dionex, coupled to an amaZon X electrospray ionization mass spectrometer from Bruker Daltonics. Peptides were separated on a C_{18} column (Acquity UPLC

BEH, 1.7 μm 2.1 × 50 mm, flow rate 0.6 mL/min, Waters). Acetonitrile−H$_2$O containing 0.1% HCOOH was used as the mobile phase under a linear gradient from 35% to 55% CH$_3$CN over 11 min. The relative amount of derivatives was obtained by comparison of the peak areas to the peak area of **1**. For preparative purification an Xbridge column (Waters, OBD, 5 μm, 19 × 150 mm) was used with CH$_3$CN−H$_2$O containing 0.1% HCOOH as the mobile phase with a linear gradient from 35% to 55% CH$_3$CN over 22 min for the separation.

Xentrivalpeptide A (1): colorless solid; mp 140 °C; $[\alpha]^{20}_D$ −33 (*c* 0.36; CHCl$_3$); UV (MeOH/H$_2$O) λ_{max} (log ε) 218 (4.34), 258 (3.82); IR ν_{max} 3280, 2967, 1756, 1633, 1531, 1454, 1157 cm^{-1}; for NMR data see Table 1; HRESIMS *m/z* 860.4880 [M + H]$^+$ (calcd for C$_{46}$H$_{66}$N$_7$O$_9$, 860.4922).

Determination of the Absolute Amino Acid Configurations (ref 20). Approximately 0.5 mg of **1** was hydrolyzed with 0.8 mL of 6 M HCl in an ACE high-pressure tube at 110 °C for 16 h. The hydrolysate was evaporated to dryness and resuspended in 100 μL of H$_2$O. To each half-portion (50 μL) were added 10 μL of 1 M NaHCO$_3$ and 100 μL of 1% FDLA N_α-(5-fluoro-2,4-dinitrophenyl)-L-leucinamide or D-leucinamide (L-FDLA or D-FDLA, solution in acetone), respectively. The reaction vials were closed and placed in a water bath at 40 °C for 1 h. After that, the reactions were cooled to room temperature, quenched with 10 μL of 1 M HCl, and evaporated to dryness. The residue was dissolved in 400 μL of MeOH. The analyses of L- and LD-FDLA-derivatized amino acids were carried out with LC-MS. Acetonitrile−H$_2$O containing 0.1% HCOOH was used as solvent with a linear gradient from 20% to 60% CH$_3$CN over 34 min.

Bioactivity Tests. Xentrivalpeptide A (**1**) was tested against Gram-negative (*Escherichia coli, Pseudomonas aeruginosa, Enterococcus faecalis*) and Gram-positive bacteria (*Staphylococcus aureus, S. epidermidis, Micrococcus luteus, Bacillus subtilis*), the yeast *Saccharomyces cerevisiae*, the eukaryotic cell lines L-929 (mouse connective tissue fibroblast; ACC 2) and HL-60 (human acute myeloid leukemia; ACC 3), and *Galleria mellonella* hemocytes as described previously.[7]

ASSOCIATED CONTENT

Ⓢ Supporting Information
HRMS and fragmentation data for all compounds, NMR data for **1** and **4**, and activity of **1** against *G. mellonella* hemocytes. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

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

Notes
The authors declare no competing financial interest.

ACKNOWLEDGMENTS

REFERENCES

(1) Bode, H. B. *Curr. Opin. Chem. Biol.* **2009**, *13*, 224−230.
(2) Goodrich-Blair, H.; Clarke, D. J. *Mol. Microbiol.* **2007**, *64*, 260−268.
(3) Herbert, E. E.; Goodrich-Blair, H. *Nat. Rev. Microbiol.* **2007**, *5*, 634−646.
(4) Waterfield, N. R.; Ciche, T.; Clarke, D. *Annu. Rev. Microbiol.* **2009**, *63*, 557−574.
(5) Bowen, D.; Rocheleau, T. A.; Blackburn, M.; Andreev, O.; Golubeva, E.; Bhartia, R.; ffrench-Constant, R. H. *Science* **1998**, *280*, 2129−2132.
(6) ffrench-Constant, R. H.; Dowling, A.; Waterfield, N. R. *Toxicon* **2007**, *49*, 436−451.
(7) Proschak, A.; Schultz, K.; Herrmann, J.; Dowling, A. J.; Brachmann, A. O.; ffrench-Constant, R.; Müller, R.; Bode, H. B. *ChemBioChem* **2011**, *12*, 2011−2015.
(8) Nollmann, F. I.; Dowling, A.; Kaiser, M.; Deckmann, K.; Grosch, S.; ffrench-Constant, R.; Bode, H. B. *Beilstein J. Org Chem* **2012**, *8*, 528−533.
(9) Park, D.; Ciezki, K.; van der, H. R.; Singh, S.; Reimer, D.; Bode, H. B.; Forst, S. *Mol. Microbiol.* **2009**, *73*, 938−949.
(10) Joyce, S. A.; Lango, L.; Clarke, D. J. *Adv. Appl. Microbiol.* **2011**, *76*, 1−25.
(11) Richardson, W. H.; Schmidt, T. M.; Nealson, K. H. *Appl. Environ. Microbiol.* **1988**, *54*, 1602−1605.
(12) Brachmann, A. O.; Joyce, S. A.; Jenke-Kodama, H.; Schwär, G.; Clarke, D. J.; Bode, H. B. *ChemBioChem* **2007**, *8*, 1721−1728.
(13) Brachmann, A. O.; Forst, S.; Furgani, G. M.; Fodor, A.; Bode, H. B. *J. Nat. Prod.* **2006**, *69*, 1830−1832.
(14) Gualtieri, M.; Aumelas, A.; Thaler, J. O. *J. Antibiot.* **2009**, *62*, 295−302.
(15) 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.
(16) Crawford, J. M.; Portmann, C.; Kontnik, R.; Walsh, C. T.; Clardy, J. *Org. Lett.* **2011**, *13*, 5144−5147.
(17) Lang, G.; Kalvelage, T.; Peters, A.; Wiese, J.; Imhoff, J. F. *J. Nat. Prod.* **2008**, *71*, 1074−1077.
(18) Ohlendorf, B.; Simon, S.; Wiese, J.; Imhoff, J. F. *Nat. Prod. Commun.* **2011**, *6*, 1247−1250.
(19) Fuchs, S. W.; Proschak, A.; Jaskolla, T. W.; Karas, M.; Bode, H. B. *Org. Biomol. Chem.* **2011**, *9*, 3130−3132.
(20) Fujii, K.; Ikai, Y.; Oka, H.; Suzuki, M.; Harada, K. *Anal. Chem.* **1997**, *69*, 5146−5151.
(21) Finking, R.; Marahiel, M. A. *Annu. Rev. Microbiol.* **2004**, *58*, 453−488.
(22) Koglin, A.; Walsh, C. T. *Nat. Prod. Rep.* **2009**, *26*, 987−1000.
(23) Sieber, S. A.; Marahiel, M. A. *Chem. Rev.* **2005**, *105*, 715−738.
(24) Firn, R. D.; Jones, C. G. *Nat. Prod. Rep.* **2003**, *20*, 382−391.

Supplementary Material for

Xentrivalpeptides A-Q: Depsipeptide diversification in *Xenorhabdus*

Qiuqin Zhou,[†] Andrea Dowling,[‡] Heinrich Heide,[§] Jens Wöhnert,[†] Ulrich Brandt,[§] James Baum,[⊥] Richard ffrench-Constant,[‡] and Helge B. Bode[†, *]

†Institut für Molekulare Biowissenschaften, Goethe Universität Frankfurt, 60438 Frankfurt am Main, Germany
‡Biosciences, University of Exeter in Cornwall, Tremough Campus, Penryn, Cornwall TR10 9EZ, UK
§Zentrum für Biologische Chemie, Molekulare Bioenergetik, Klinikum der Goethe Universität Frankfurt, 60590 Frankfurt am Main, Germany
⊥Monsanto Company, Chesterfield, MO 63017, USA

1

Table S1. Results from HRESIMS (exp.) and the resulting molecular formula with the tolerances in ppm. **a)** protonated xentrivalpeptides and the two main MS[2] fragments. **b)** MS[3] fragments using fragment B as precursor. Fragment A was fragmented to fragment B by losing a water molecule. Fragmentation of **17** was shown in Figure S8.

a

Nr.	molecular formula $[M + H]^+$	exp.	theo.	Δ ppm	molecular formula fragment A	m/z exp.	Δ ppm	molecular formula fragment B	m/z exp.	Δ ppm
1	$C_{46}H_{66}O_9N_7^+$	860.4880	860.4922	4.2	$C_{33}H_{51}O_7N_6^+$	643.3773	6.0	$C_{33}H_{49}O_6N_6^+$	625.3671	6.0
2	$C_{40}H_{62}O_9N_7^+$	784.4580	784.4609	3.1	$C_{33}H_{51}O_7N_6^+$	643.3780	7.8	$C_{33}H_{49}O_6N_6^+$	625.3659	7.8
3	$C_{41}H_{64}O_9N_7^+$	798.4742	798.4766	2.2	$C_{33}H_{51}O_7N_6^+$	643.3773	6.5	$C_{33}H_{49}O_6N_6^+$	625.3678	6.5
4	$C_{42}H_{66}O_9N_7^+$	812.4817	812.4922	0.1	$C_{33}H_{51}O_7N_6^+$	643.3782	4.7	$C_{33}H_{49}O_6N_6^+$	625.3679	4.7
5	$C_{43}H_{68}O_9N_7^+$	826.5049	826.5079	2.9	$C_{33}H_{51}O_7N_6^+$	643.3783	4.6	$C_{33}H_{49}O_6N_6^+$	625.3679	4.6
6	$C_{45}H_{72}O_9N_7^+$	854.5364	854.5392	2.6	$C_{33}H_{51}O_7N_6^+$	643.3766	6.9	$C_{33}H_{49}O_6N_6^+$	625.3665	6.9
7	$C_{44}H_{70}O_9N_7^+$	840.5211	840.5235	2.0	$C_{33}H_{51}O_7N_6^+$	643.3743	11.0	$C_{33}H_{49}O_6N_6^+$	625.3675	5.3
8	$C_{45}H_{64}O_9N_7^+$	846.4729	846.4766	3.7	$C_{33}H_{51}O_7N_6^+$	643.3768	7.2	$C_{33}H_{49}O_6N_6^+$	625.3674	5.5
9	$C_{47}H_{68}O_9N_7^+$	874.5057	874.5079	1.8	$C_{33}H_{51}O_7N_6^+$	643.3749	8.8	$C_{33}H_{49}O_6N_6^+$	625.3653	8.8
10	$C_{39}H_{68}O_9N_7^+$	778.5041	778.5079	4.1	$C_{30}H_{53}O_7N_6^+$	609.3951	3.8	$C_{30}H_{51}O_6N_6^+$	591.3842	3.8
11	$C_{40}H_{70}O_9N_7^+$	792.5214	792.5235	2.0	$C_{30}H_{53}O_7N_6^+$	609.3939	6.3	$C_{30}H_{51}O_6N_6^+$	591.3828	6.3
12	$C_{43}H_{68}O_9N_7^+$	826.5049	826.5079	2.9	$C_{30}H_{53}O_7N_6^+$	609.3945	5.2	$C_{30}H_{51}O_6N_6^+$	591.3834	5.2
13	$C_{42}H_{74}O_9N_7^+$	820.5531	820.5548	1.4	$C_{30}H_{53}O_7N_6^+$	609.3939	6.0	$C_{30}H_{51}O_6N_6^+$	591.3829	6.0
14	$C_{47}H_{68}O_9N_7^+$	874.5057	874.5079	1.8	$C_{34}H_{53}O_7N_6^+$	657.3923	6.9	$C_{34}H_{51}O_6N_6^+$	639.3820	6.9
15	$C_{45}H_{64}O_9N_7^+$	846.4729	846.4766	3.7	$C_{32}H_{49}O_7N_6^+$	629.3613	7.1	$C_{32}H_{47}O_6N_6^+$	611.3523	4.7
16	$C_{41}H_{64}O_9N_7^+$	798.4742	798.4766	2.2	$C_{33}H_{51}O_7N_6^+$	629.3607	8.1	$C_{32}H_{47}O_6N_6^+$	611.3527	4.0
17	$C_{41}H_{57}O_8N_6^+$	761.4197	761.4238	4.6	fragment A´ in figure S7			fragment B´ in figure S7		

2

b

Nr.	precursor	fragment C			fragment D			fragment E			fragment F		
		chemical formula	m/z exp.	Δ ppm	chemical formula	m/z exp.	Δ ppm	chemical formula	m/z exp.	Δ ppm	chemical formula	m/z exp.	Δ ppm
1	625.3	$C_{22}H_{33}O_3N_4^+$	401.2533	3.5	$C_{23}H_{31}O_4N_4^+$	427.2315	5.8	$C_{23}H_{33}O_4N_4^+$	429.2468	6.5	$C_{20}H_{35}O_4N_4^+$	395.2647	0.6
2	625.3	$C_{22}H_{33}O_3N_4^+$	401.2545	0.6	$C_{23}H_{31}O_4N_4^+$	427.2326	3.3	$C_{23}H_{33}O_4N_4^+$	429.2448	11.2	$C_{20}H_{35}O_4N_4^+$	395.2673	5.0
3	625.3	$C_{22}H_{33}O_3N_4^+$	401.2533	3.6	$C_{23}H_{31}O_4N_4^+$	427.2323	3.9	$C_{23}H_{33}O_4N_4^+$	429.2465	7.3	$C_{20}H_{35}O_4N_4^+$	395.2634	4.7
4	625.3	$C_{22}H_{33}O_3N_4^+$	401.2527	5.1	$C_{23}H_{31}O_4N_4^+$	427.2332	1.8	$C_{23}H_{33}O_4N_4^+$	429.2458	8.9	$C_{20}H_{35}O_4N_4^+$	395.2641	3.1
5	625.3	$C_{22}H_{33}O_3N_4^+$	401.2535	3.0	$C_{23}H_{31}O_4N_4^+$	427.2291	11.5	$C_{23}H_{33}O_4N_4^+$	429.2467	6.8	$C_{20}H_{35}O_4N_4^+$	395.2663	2.5
6	625.3	$C_{22}H_{33}O_3N_4^+$	401.2522	6.2	$C_{23}H_{31}O_4N_4^+$	427.2326	3.2	$C_{23}H_{33}O_4N_4^+$	429.2479	4.1	$C_{20}H_{35}O_4N_4^+$	395.2640	3.2
7	625.3	$C_{22}H_{33}O_3N_4^+$	401.2527	5.0	$C_{23}H_{31}O_4N_4^+$	427.2325	3.6	$C_{23}H_{33}O_4N_4^+$	429.2471	5.9	$C_{20}H_{35}O_4N_4^+$	395.2639	3.6
8	625.3	$C_{22}H_{33}O_3N_4^+$	401.2513	8.4	$C_{23}H_{31}O_4N_4^+$	427.2304	8.3	$C_{23}H_{33}O_4N_4^+$	429.2521	5.8	$C_{20}H_{35}O_4N_4^+$	395.2661	1.9
9	625.3	$C_{22}H_{33}O_3N_4^+$	401.2507	0.1	$C_{23}H_{31}O_4N_4^+$	427.2338	0.5	$C_{23}H_{33}O_4N_4^+$	429.2495	0.4	$C_{20}H_{35}O_4N_4^+$	395.2666	3.3
10	591.3	*	*	*	$C_{20}H_{33}O_4N_4^+$	393.2478	4.7	$C_{20}H_{35}O_4N_4^+$	395.2643	2.4	$C_{20}H_{35}O_4N_4^+$	395.2643	2.4
11	591.5	*	*	*	$C_{20}H_{33}O_4N_4^+$	393.2479	4.4	$C_{20}H_{35}O_4N_4^+$	395.2640	3.1	$C_{20}H_{35}O_4N_4^+$	395.2640	3.1
12	591.3	*	*	*	$C_{20}H_{33}O_4N_4^+$	393.2491	3.8	$C_{20}H_{35}O_4N_4^+$	395.2645	1.9	$C_{20}H_{35}O_4N_4^+$	395.2645	1.9
14	639.3	$C_{23}H_{35}O_3N_4^+$	415.2699	1.9	$C_{23}H_{31}O_4N_4^+$	427.2337	0.7	$C_{24}H_{35}O_4N_4^+$	443.2646	1.5	$C_{21}H_{37}O_4N_4^+$	409.2804	1.3
15	611.3	$C_{21}H_{31}O_3N_4^+$	387.2367	5.7	$C_{23}H_{31}O_4N_4^+$	427.2296	10.2	$C_{22}H_{31}O_4N_4^+$	415.2326	3.2	$C_{19}H_{33}O_4N_4^+$	381.2477	5.0

* not detected; MS^3 fragments of **13** and **16** were too weak.

3

Table S2. Results from the advanced Marfey´s method: retention times (t_R, minute) of the derivatized amino acids from the hydrolyzed **1** detected at negative ion mode by HPLC-MS. Commercial L-amino acids were used as standard. Only phenylalanine in **1** has D-configuration.

	m/z [M-H]$^-$	t_R, L-FDLA[a]	t_R, LD-FDLA[a]
Pro(4)	408.3	**14.7**	**14.7**
			17.2
Thr(2)	412.3	**11.8**	**11.8**
			15.9
Val(1)/(5)/(6)/(7)	410.3	**17.1**	**17.1**
			23.0
Phe(3)	458.3	**25.0**	20.0
			25.0

[a]: Marfey´s reagents (N$_\alpha$-(5-fluoro-2,4-dinitrophenyl)-L-leucinamide, abbr. L-FDLA; N$_\alpha$-(5-fluoro-2,4-dinitrophenyl)-D-leucinamide, abbr. D-FDLA); mixture of L-FDLA and D-FDLA, abbr. LD-FDLA.

Table S3. NMR Spectroscopic Data (500 MHz (^{1}H), 125 MHz (^{13}C) in CDCl$_3$) of **4**, δ in ppm. Amino acid numbering from N- to C-terminus.

subunit	position	δ$_C$	δ$_H$, mult (J in Hz)
Butyryl	1	175.8	
	2a	38.4	2.66, m
	2b	38.4	2.43, m
	3	18.7	1.82, m
	4	19.6^a	1.03, t (7.4)
L-Val(1)	C=O	172.4	
	α	61.4	4.16, m
	β	29.8	2.25, m
	γ	19.6^a	0.97, d (6.6)
	δ	20.1	0.95, d (6.4)
	NH		6.39, d (5.8)
L-Thr(2)	C=O	169.9	
	α	54.9	4.70, d (9.9)
	β	69.2	5.27, m
	γ	18.4	1.33, d (6.5)
	NH		7.32. d (10.0)
D-Phe(3)	C=O	173.2	
	α	54.9	4.35, m
	β	37.5	3.16, m
	γ	135.7	
	δ	129.4	7.26, m
	ε	128.3	7.28, m
	ζ	127.3	7.21, m
	NH		7.36, m
L-Pro (4)	C=O	171.7	
	α	61.6	4.14, m
	β1	29.3	1.78, m
	β2	29.3	1.91, m
	γ1	24.4	1.50, m
	γ2	24.4	1.68, m
	δ1	47.4	2.47, m
	δ2	47.4	3.68, m
L-Val(5)	C=O	173.6	
	α	57.6	4.47, m
	β	28.3	2.50, m
	γ	19.6^a	0.87, d, (7.0)
	δ	16.4	0.73, d, (6.9)
	NH		6.19, d, (10.1)
L-Val(6)	C=O	174.4	
	α	67.6	3.15, m
	β	27.2	2.90, m
	γ	17.9	1.02, d (6.8)
	δ	19.6^a	0.98, d (6.8)
	NH		8.04, d (6.9)
L-Val(7)	C=O	169.4	
	α	57.9	4.00, t (7.5)
	β	27.6	2.14, m
	γ	19.5	1.12, d (6.9)
	δ	18.5	1.11, d (6.7)
	NH		8.80, bs

d: doublet; m: multiplet; ^{a}: same chemical shift

5
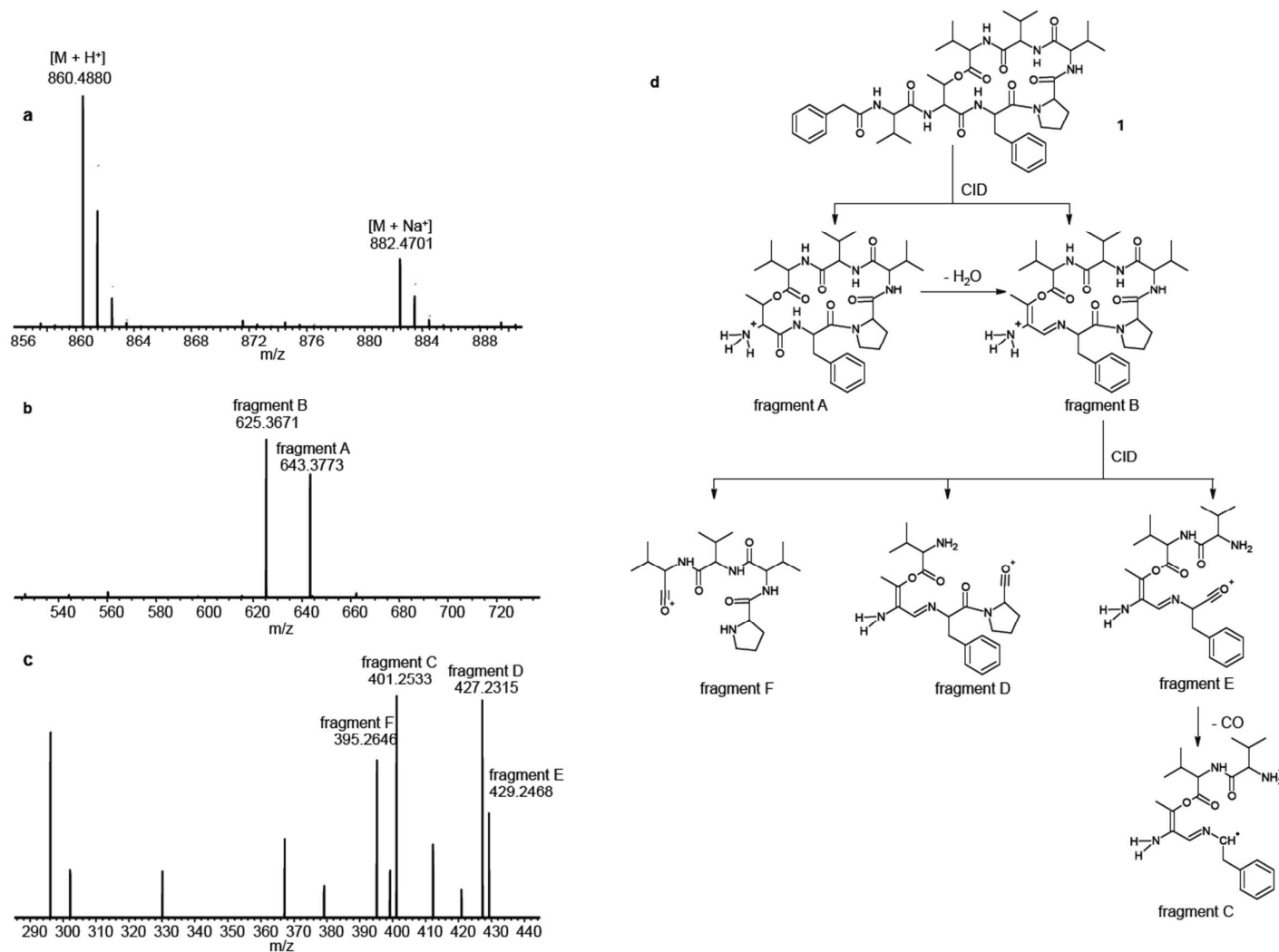
Figure S1. a) HRESIMS of **1**. **b)** MS^2-spectra of **1** after collision induced dissociation (CID) of [M+H]$^+$. **c)** MS^3-spectra of fragment B in selected rang. **d)** Schema for fragmentation path of **1** and the proposed fragment structures. The fragments A-F (Table S1) were selected for the structure elucidation.

6

1 (R¹ = g, R² = c)

	[M +H]⁺
^{12}C	860.5
^{12}C + 1 × ^{2}H$_8$-Phe	867.5
^{12}C + 2 × ^{2}H$_8$-Phe	874.5
^{12}C + 1 × ^{2}H$_8$-Val	867.5
^{12}C + 1 × p-F-PAA	878.5
^{13}C	906.5
^{13}C + 1 × Phe	898.5, 897.5
^{13}C + 2 × Phe	889.5
^{13}C + 1 × Val	901.5

↓ CID

fragment B	m/z
^{12}C	625.3
^{12}C + 1 × ^{2}H$_8$-Phe	625.3, 632.3
^{12}C + 2 × ^{2}H$_8$-Phe	632.3
^{12}C + 1 × ^{2}H$_8$-Val	625.3, 632.3
^{13}C	658.3
^{13}C + 1 × Phe	649.3, 658.3
^{13}C + 2 × Phe	649.3
^{13}C + 1 × Val	653.3, 658.3

2 (R¹ = a, R² = c)

	[M +H]⁺
^{12}C	784.5
^{12}C + 1 × ^{2}H$_8$-Val	791.5
^{13}C	824.5
^{13}C + 1 × Val	819.5

↓ CID

fragment B	m/z
^{12}C	625.3
^{12}C + 1 × ^{2}H$_8$-Val	625.3, 632.3
^{13}C	658.3,
^{13}C + 1 × Val	653.3, 658.3

3 (R¹ = b, R² = c)

	[M +H]⁺
^{12}C	798.5
^{12}C + 1 × ^{2}H$_8$-Val	805.5
^{12}C + 1 × ^{2}H$_6$-propionic acid	803.5
^{13}C	839.5
^{13}C + 1 × Val	834.5

↓ CID

fragment B	m/z
^{12}C	625.3
^{12}C + 1 × ^{2}H$_8$-Val	625.3, 632.3
^{13}C	658.3
^{13}C + 1 × Val	653.3, 658.3

4 (R¹ = e, R² = c)

	[M +H]⁺
^{12}C	812.5
^{12}C + 1 × ^{2}H$_8$-Phe	819.5
^{12}C + 1 × ^{2}H$_8$-Val	819.5
^{12}C + 2 × ^{2}H$_8$-Val	826.5
^{13}C	854.5
^{13}C + 1 × Phe	845.5
^{13}C + 1 × Val	849.5
^{13}C + 2 × Val	844.5
^{13}C + 1 × Butyric acid	850.5

↓ CID

fragment B	m/z
^{12}C	625.3
^{12}C + 1 × ^{2}H$_8$-Phe	632.3
^{12}C + 1 × ^{2}H$_8$-Val	625.3, 632.3
^{12}C + 2 × ^{2}H$_8$-Val	632.3, 639.3
^{13}C	658.3
^{13}C + 1 × Phe	649.3
^{13}C + 1 × Val	653.3, 658.3
^{13}C + 2 × Val	648.3, 653.3
^{13}C + 1 × Butyric acid	658.3

5 (R¹ = d, R² = c)

	[M +H]⁺
^{12}C	826.5
^{12}C + 1 × ^{2}H$_8$-Phe	833.5
^{12}C + 1 × ^{2}H$_{10}$-Leu	835.5
^{12}C + 1 × ^{2}H$_8$-Val	833.5
^{13}C	869.5
^{13}C + 1 × Phe	860.5
^{13}C + 1 × Leu	864.5
^{13}C + 1 × Val	864.5

↓ CID

fragment B	m/z
^{12}C	625.3
^{12}C + 1 × ^{2}H$_8$-Phe	632.3
^{12}C + 1 × ^{2}H$_{10}$-Leu	625.3
^{12}C + 1 × ^{2}H$_8$-Val	625.3, 632.3
^{13}C	658.3
^{13}C + 1 × Phe	649.3
^{13}C + 1 × Leu	658.3
^{13}C + 1 × Val	653.3, 658.3

6 (R¹ = h, R² = c)

	[M +H]⁺
^{12}C	854.5
^{12}C + 1 × ^{2}H$_8$-Phe	861.5
^{12}C + 1 × ^{2}H$_{10}$-Leu	863.5
^{12}C + 1 × ^{2}H$_8$-Val	861.5
^{13}C	899.5
^{13}C + 1 × Phe	890.5
^{13}C + 1 × Leu	894.5
^{13}C + 1 × Val	894.5

↓ CID

fragment B	m/z
^{12}C	625.3
^{12}C + 1 × ^{2}H$_8$-Phe	632.3
^{12}C + 1 × ^{2}H$_{10}$-Leu	625.3
^{12}C + 1 × ^{2}H$_8$-Val	625.3, 632.3
^{13}C	658.3
^{13}C + 1 × Phe	649.3
^{13}C + 1 × Val	653.3, 658.3

8 (R¹ = g, R² = b)

	[M +H]⁺
^{12}C	846.5
^{12}C + 1 × p-F-PAA	864.5
^{13}C	891.5
^{13}C + 1 × Phe	883.5, 882.5
^{13}C + 1 × Val	886.5
^{13}C + 1 × aminobutyric acid	887.5

↓ CID

fragment B	m/z
^{12}C	625.3
^{13}C	658.3
^{13}C + 1 × Phe	649.3, 658.3
^{13}C + 1 × Val	653.3
^{13}C + 1 × aminobutyric acid	658.3

9 (R¹ = g, R² = f)

	[M +H]⁺
^{12}C	874.5
^{12}C + 1 × ^{2}H$_8$-Phe	881.5
^{12}C + 1 × ^{2}H$_8$-Val	881.5
^{12}C + 1 × p-F-PAA	892.5
^{13}C	921.5
^{13}C + 1 × Phe	912.5, 913.5
^{13}C + 2 × Phe	904.5
^{13}C + 1 × Ile	915.5
^{13}C + 1 × Val	916.5

↓ CID

fragment B	m/z
^{12}C	625.3
^{12}C + 1 × ^{2}H$_8$-Phe	625.3, 632.3
^{12}C + 1 × ^{2}H$_8$-Val	625.3, 632.3
^{13}C	658.3
^{13}C + 1 × Phe	649.3, 658.3
^{13}C + 2 × Phe	649.3
^{13}C + 1 × Ile	658.3
^{13}C + 1 × Val	653.3, 658.3

Figure S2. Selected results of feeding experiments from **1-6**, **8**, and **9** which have the same ring structure. Proline and threonine were incorporated in all xentrivalpeptides (data not shown). Structures of R¹ and R² were shown in figure 4. **Feeding experiments in nonlabled medium** (LB medium, ^{12}C): incorporation of one L[2,3,5,5′,6,6′,7-^{2}H$_8$]phenylalanine (^{12}C + 1 × ^{2}H$_8$-Phe), incorporation of two L[2,3,5,5′,6,6′,7-^{2}H]phenylalanine (^{12}C + 2 × ^{2}H$_8$-Phe), incorporation of one DL[2,3,4,4,4,5,5,5-^{2}H$_8$]valine (^{12}C + 1 × ^{2}H$_8$-Val), incorporation of two DL[2,3,4,4,4,5,5,5-^{2}H$_8$]valine (^{12}C + 2 × ^{2}H$_8$-Val), incorporation of one L[2,3,3,4,5,5,5,5′,5′,5′-^{2}H$_{10}$]leucine (^{12}C + 1 × Leu), incorporation of one p-phenylacetic acid (^{12}C + 1 × p-F-PAA), incorporation of one [U-^{2}H$_6$]propionic acid (^{12}C + 1 × ^{2}H$_6$-propionic acid). **Feeding experiments in labled medium** (ISOGRO®-^{13}C growth medium, ^{13}C): incorporation of one L-α-amino acid (^{13}C + 1 × amino

7
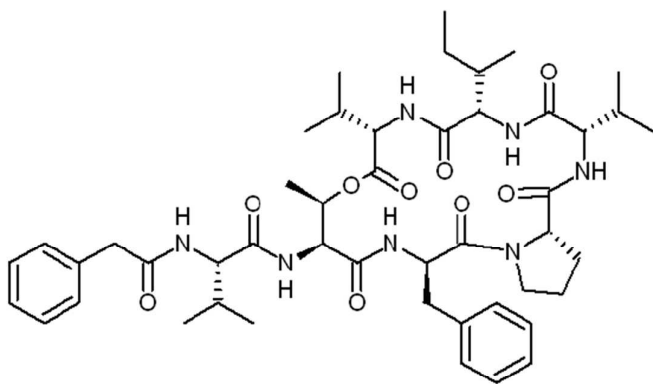
acid), incorporation of two L-α-amino acids (^{13}C + 2 × amino acids), incorporation of one butyric acid (^{13}C + 1 × butyric acid). CID: collision induced dissociation. All xentrivalpeptides were fragmented to fragment A and B (Figure S1).
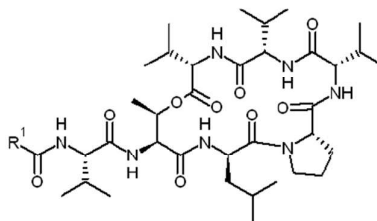
10 (R¹ = e)	[M +H]⁺	11 (R¹ = d)	[M +H]⁺	12 (R¹ = g)	[M +H]⁺	13 (R¹ = h)	[M +H]⁺
^{12}C	778.5	^{12}C	792.5	^{12}C	826.5	^{12}C	820.5
^{12}C + 1 × ^{2}H$_{10}$-Leu	787.5	^{12}C + 1 × ^{2}H$_{10}$-Leu	801.5	^{12}C + 1 × ^{2}H$_{10}$-Leu	835.5	^{12}C + 1 × ^{2}H$_{10}$-Leu	829.5
^{12}C + 1 × ^{2}H$_8$-Val	785.5	^{12}C + 1 × ^{2}H$_8$-Val	799.5	^{12}C + 1 × ^{2}H$_8$-Val	833.5		
^{12}C + 2 × ^{2}H$_8$-Val	792.5			^{12}C + 1 × ^{2}H$_8$-Phe	833.5		
				^{12}C + 1 × p-F-PAA	844.5		
^{13}C	817.5	^{13}C	832.5	^{13}C	869.5		
^{13}C + 1 × Leu	811.5	^{13}C + 1 × Leu	826.5, 827.5	^{13}C + 1 × Leu	863.5		
^{13}C + 1 × Val	812.5	^{13}C + 1 × Val	827.5	^{13}C + 1 × Val	864.5		
^{13}C + 2 × Val	807.5	^{13}C + 2 × Val	822.5	^{13}C + 2 × Val	859.5		
^{13}C + 1 × Butyric acid	813.5			^{13}C + 1 × Phe	861.5		

CID | CID | CID | CID

fragment B	m/z	fragment B	m/z	fragment B	m/z	fragment B	m/z
^{12}C	591.3	^{12}C	591.3	^{12}C	591.3	^{12}C	591.3
^{12}C + 1 × ^{2}H$_{10}$-Leu	600.3	^{12}C + 1 × ^{2}H$_{10}$-Leu	591.3, 600.3	^{12}C + 1 × ^{2}H$_{10}$-Leu	600.3	^{12}C + 1 × ^{2}H$_{10}$-Leu	600.3, 609.3
^{12}C + 1 × ^{2}H$_8$-Val	591.3, 598.3	^{12}C + 1 × ^{2}H$_8$-Val	591.3, 598.3	^{12}C + 1 × ^{2}H$_8$-Val	591.3, 598.3		
^{12}C + 2 × ^{2}H$_8$-Val	598.3, 605.3						
^{13}C	621.3	^{13}C	621.3	^{13}C	621.3		
^{13}C + 1 × Val	621.3, 616.3	^{13}C + 1 × Leu	621.3, 615.3	^{13}C + 1 × Leu	615.3		
^{13}C + 2 × Val	616.3, 611.3	^{13}C + 1 × Val	621.3, 616.3	^{13}C + 1 × Val	621.3, 616.3		
^{13}C + 1 × Butyric acid	621.3	^{13}C + 2 × Val	616.3, 611.3	^{13}C + 2 × Val	616.3, 611.3		
				^{13}C + 1 × Phe	621.3		

Figure S3. Selected results of feeding experiments for **10-13**, which have the same ring structure fragment B. Abbreviations see Figure S2.

14	[M+H]$^+$
^{12}C	874.5
^{12}C + 1 × ^{2}H$_8$-Phe	881.5
^{12}C + 1 × ^{2}H$_8$-Val	881.5
^{12}C + 1 × *p*-F-PAA	892.5
^{13}C	921.5
^{13}C + 1 × Phe	912.5, 913.5
^{13}C + 2 × Phe	904.5
^{13}C + 1 × Ile	915.5
^{13}C + 1 × Val	916.5

CID

fragment B	m/z
^{12}C	639.3
^{12}C + 1 × ^{2}H$_8$-Phe	639.3, 646.3
^{12}C + 1 × ^{2}H$_8$-Val	639.3, 646.3
^{13}C	673.3
^{13}C + 1 × Phe	664.3, 673.3
^{13}C + 2 × Phe	664.3
^{13}C + 1 × Ile	667.3
^{13}C + 1 × Val	668.3, 673.3

Figure S4. Selected results of feeding experiments for **14**. Abbreviations see Figure S2.

15	[M+H]+
^{12}C	846.5
^{12}C + 1 × p-F-PAA	864.5
^{13}C	891.5
^{13}C + 1 × Phe	883.5, 882.5
^{13}C + 1 × Val	886.5
^{13}C + 1 × aminobutyric acid	887.5

CID

fragment B	m/z
^{12}C	611.3
^{13}C	643.3
^{13}C + 1 × Phe	634.3, 643.3
^{13}C + 1 × Val	638.3, 643.3
^{13}C + 1 × aminobutyric acid	639.3

Figure S5. Selected results of feeding experiments for **15**. Abbreviations see Figure S2. Fragment B of **15** and **16** indicated the same ring structure (Table S1). But, labeling of **16** was too weak for its characterization.

a

Pro(4)   Thr(2)   Val(1)/(5)/(6)/(7)

Phe(3)   Butyryl

b
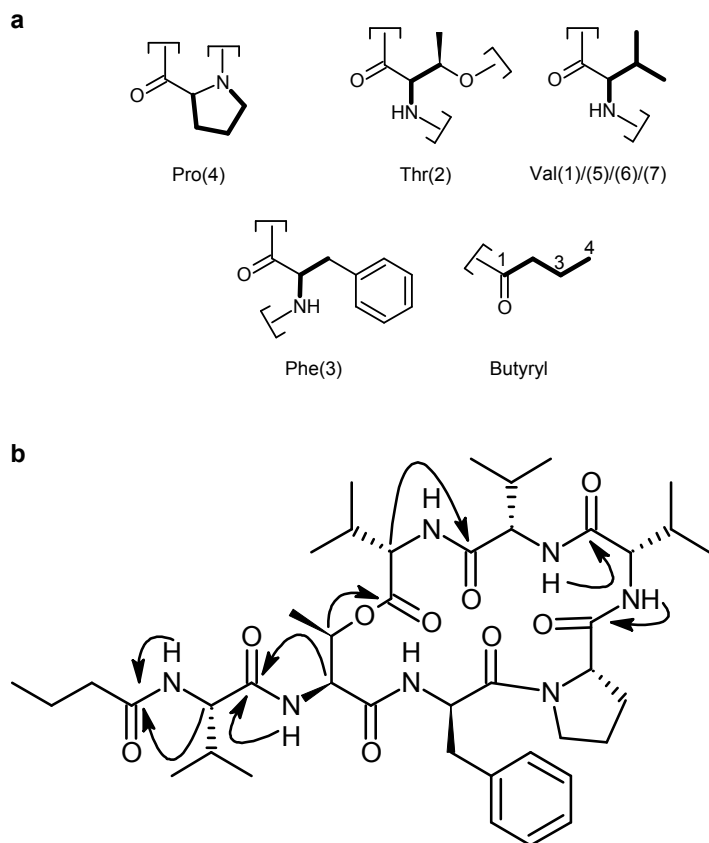
Figure S6. a) Subunits and selected COSY correlations (bond lines) for **4**. **b**) Selected HMBC correlations (arrows from ^{1}H to ^{13}C) between subunits of **4**. Amino acid numbering from N- to C-terminus.

11

103

a

fragment A´
544.3105

[M + H]⁺ - Val
662.3524

fragment B´
526.2996

fragment H
461.2737

fragment I
448.2207

fragment G
466.2307

fragment K
301.1539

fragment J
314.2073

b

fragment G

− H₂O

fragment I

− Phe

fragment K

fragment A´

− H₂O

fragment B´

fragment J

− Phe
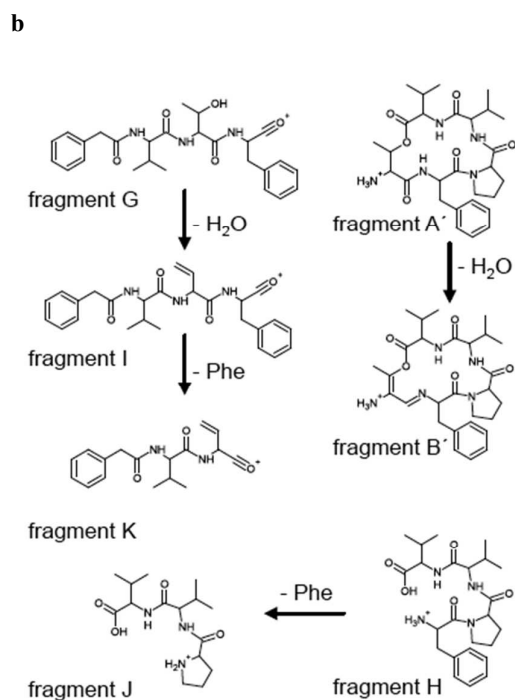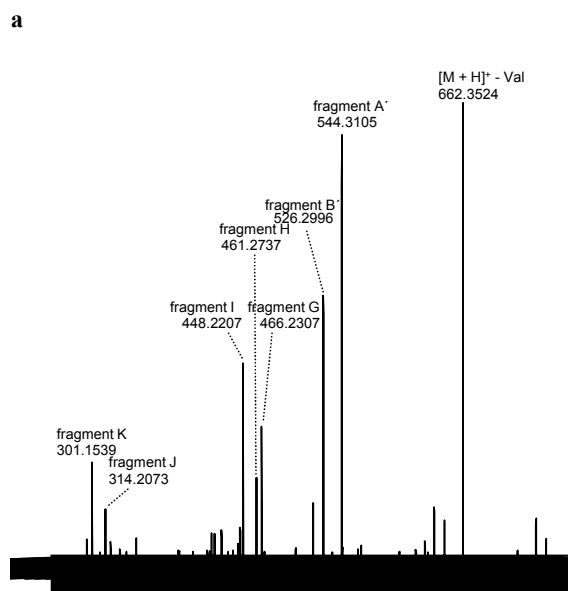
fragment H

c

fragment	formula	m/z exp.	Δppm
[M + H]⁺ -Val	$C_{26}H_{48}O_7N_5^+$	622.3524	3.7
fragment A´	$C_{28}H_{42}O_6N_5^+$	544.3105	5.4
fragment B´	$C_{28}H_{40}O_5N_5^+$	526.2996	5.4
fragment G	$C_{26}H_{32}O_5N_3^+$	466.2307	6.2
fragment H	$C_{24}H_{37}O_5N_4^+$	461.2737	4.6
fragment I	$C_{26}H_{30}O_4N_3^+$	448.2207	5.2
fragment J	$C_{15}H_{28}N_3O_4^+$	314.2073	4.0
fragment K	$C_{17}H_{21}N_2O_3^+$	301.1539	2.1

Figure S7. Structure elucidation of **17**. **a**) HR-ESI MS² fragments. **b**) Proposed fragment structures. **c**) Molecular formula of the fragments.
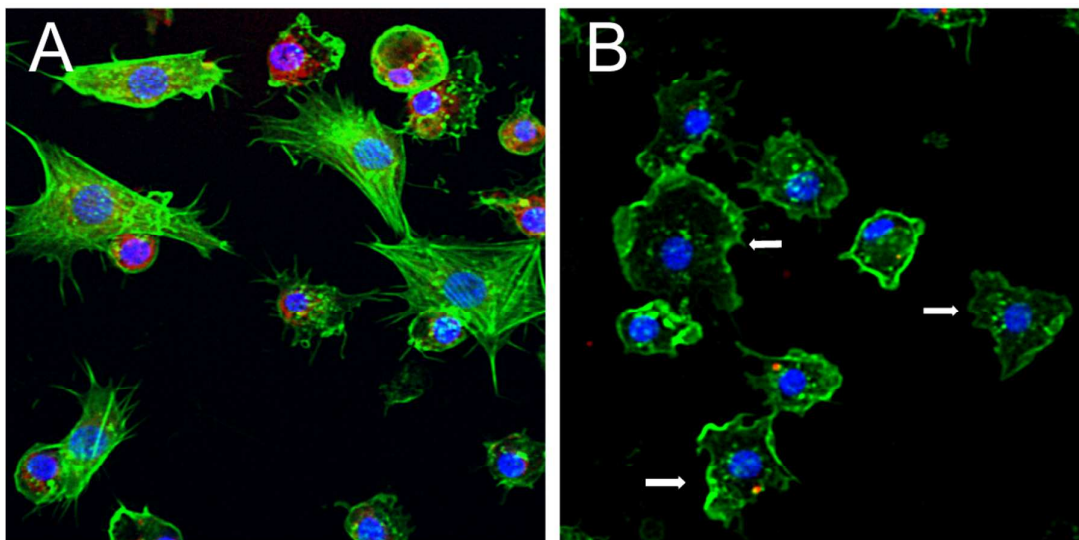
Figure S8. *Galleria mellonella* hemocyte monolayers were incubated for 4h with (A) 2% DMSO in Graces Insect Medium as a control, and (B) 100 µg/ml of **1**. White arrows indicate regions of actin cytoskeletal ruffling caused by the compound. Green = FITC-phalloidin labeled f-actin, Red = polarized mitochondria, Blue = nuclei.

13

^{1}H-NMR of **1**

14

^{13}C-NMR of **1**

15

COSY of **1**

16

HSQC of **1**

17

HMBC of **1**

18

^{1}H-NMR of **4**

19

^{13}C-NMR of **4**

ppm (t1)

20

COSY of **4**

21

HSQC of **4**

HMBC of 4

11.2. Structure and Biosynthesis of Xenoamicins from Entomopathogenic *Xenorhabdus*

Authors: Qiuqin Zhou,[a]† Florian Grundmann,[a]† Marcel Kaiser,[b] Matthias Schiell,[c] Sophie Gaudriault,[d] Andreas Batzer,[c] Michael Kurz[c] and Helge B. Bode[a]

[a] Goethe-Universität Frankfurt, Department of Molecular Biotechnology, Max-von-Laue-Strasse 9, 60438 Frankfurt am Main (Germany)

[b] Swiss Tropical and Public Health Institute, Parasite Chemotherapy, Socinstrasse 57, P.O. Box, 4002 Basel (Switzerland)

[c] Sanofi R&D, Industriepark Höchst, 65926 Frankfurt am Main (Germany)

[d] INRA, UMR 1333 Laboratoire DGIMI, 34095 Montpellier (France) and Université Montpellier 2, UMR 1333 Laboratoire DGIMI, 34095 Montpellier (France)

† These authors contributed equally to this work.

Attachments: Declaration on the contribution of the author and the paper.

Erklärung zu den Autorenanteilen

an der Publikation: "Structure and Biosynthesis of Xenoamicins from Entomopathogenic *Xenorhabdus"*

Status: published in *Chem. Eur. J.*, **2013**, *19,* 16772–16779.

Beteiligte Autoren: Qiuqin Zhou (QZ), Florian Grundmann (FG), Marcel Kaiser (MK), Matthias Schiell (MS), Sophie Gaudriault , Andreas Batzer (AB), Michael Kurz (MK) and Helge B. Bode (HBB)

Was hat der Promovierende bzw. was haben die Koautoren beigetragen?

(1) zu Entwicklung und Planung

Promovierender: 25 %

Co-Autor FG: 25 %

Co-Autor HBB: 45 %

Co-Autor AB: 5 %

(2) zur Durchführung der einzelnen Untersuchungen und Experimente

Promovierender: 47 %, MS-Experimente, Fütterungsexperimente, Knockout-Mutant und Marfey-Experiment

Co-Autor FG: 47%, Isolierung

Co-Autor MK: 3 %, Bioaktivitätstest

Co-Autor MS: 3 % Wartung der Isolierungsanlagen

(3) zur Erstellung der Datensammlung und Abbildungen

Promovierender: 48 %, Abbildung und Text zu MS-Experimenten, Fütterungsexperimenten, Biosyntheseaufklärung und Marfey-Experiment

Co-Autor FG: 48 %, Zusammentragen aller Daten und Schreiben der Veröffentlichung

Co-Autor MK: 4 %, Bioaktivitätstest

(4) zur Analyse und Interpretation der Daten

Promovierender: 48 %, MS-Daten Xenoamcine A-C, Fütterungsexperimenten, Biosyntheseaufklärung und Marfey-Experiment

Co-Autor FG: 48 %, Auswertung von NMR, Auswertung von MS Daten Xenoamicine D-H

Co-Autor MK: 4 %, Auswertung von Bioaktivitätstest

(5) zum Verfassen des Manuskripts

Promovierender: 35 %

Co-Autor FG: 35 %

Co-Autor HBB: 30 %

Unterschrift Promovend: _____ Datum/Ort: _____

Zustimmende Bestätigungen der oben genannten Angaben

Unterschrift Betreuer: _____ Datum/Ort: _____

Structure and Biosynthesis of Xenoamicins from Entomopathogenic *Xenorhabdus*

Qiuqin Zhou,[a] Florian Grundmann,[a] Marcel Kaiser,[b] Matthias Schiell,[c] Sophie Gaudriault,[d] Andreas Batzer,[c] Michael Kurz,[c] and Helge B. Bode*[a]

Abstract: During the search for novel natural products from entomopathogenic *Xenorhabdus doucetiae* DSM17909 and *X. mauleonii* DSM17908 novel peptides named xenoamicins were identified in addition to the already known antibiotics xenocoumacin and xenorhabdin. Xenoamicins are acylated tridecadepsipeptides consisting of mainly hydrophobic amino acids. The main derivative xenoamicin A (**1**) was isolated from *X. mauleonii* DSM17908, and its structure elucidated by detailed 1D and 2D NMR experiments. Detailed MS experiments, also in combination with labeling experiments, confirmed the determined structure and allowed structure elucidation of additional derivatives. Moreover, the xenoamicin biosynthesis gene cluster was identified and analyzed in *X. doucetiae* DSM17909, and its participation in xenoamicin biosynthesis was confirmed by mutagenesis. Advanced Marfey's analysis of **1** showed that the absolute configuration of the amino acids is in agreement with the predicted stereochemistry deduced from the nonribosomal peptide synthetase XabABCD. Biological testing revealed activity of **1** against *Plasmodium falciparum* and other neglected tropical diseases but no antibacterial activity.

Keywords: biosynthesis · natural products · peptides · structure elucidation

Introduction

Natural products have always been a great source of novel bioactive compounds, in particular in the field of anti-infectives. As the number of multiresistant pathogens is increasing, it is necessary to continue research in the field of natural products,[1] and thus novel approaches for the identification of natural products from well-established or novel natural-product producers are required. Additionally, novel molecular approaches also allow the activation of previously silent gene clusters or heterologous expression.[2,3] Among novel natural-product producers, bacteria of the genus *Xenorhabdus* have shown promise.[4–6] The bacteria live in symbiosis with nematodes of the genus *Steinernema* and are released from the nematode gut when the nematode infects an insect. Besides insecticidal protein toxins *Xenorhabdus* also produces small molecules that are involved in insect virulence or act as antibiotics. Examples are the benzylideneacetone,[7] rhabduscin,[8] xenocoumacins,[9] xenorhabdins,[10] xenofuranones,[11] nematophin,[12] PAX-Peptides,[13] GameXPeptides[14] and even depsipeptides, such as xenematides,[15] szentiamides,[15] and xentrival peptides,[16] which show antibiotic, antifungal, or cytotoxic activity. As several of the compounds may be released to protect the insect cadaver from food competitors living in the soil, they may also be potential anti-infectives. Here we describe the structure and biosynthesis of a novel class of large hydrophobic depsipeptides named xenoamicins from *Xenorhabdus* showing activity against *Plasmodium* and *Trypanosoma*.

[a] Q. Zhou,+ F. Grundmann,+ Prof. Dr. H. B. Bode
Goethe-Universität Frankfurt
Department of Molecular Biotechnology
Max-von-Laue-Strasse 9, 60438 Frankfurt am Main (Germany)
E-mail: h.bode@bio.uni-frankfurt.de

[b] M. Kaiser
Swiss Tropical and Public Health Institute, Parasite Chemotherapy
Socinstrasse 57, P.O. Box,
4002 Basel (Switzerland)

[c] M. Schiell, Dr. A. Batzer, Dr. M. Kurz
Sanofi R&D, Industriepark Höchst
65926 Frankfurt am Main (Germany)

[d] Dr. S. Gaudriault
INRA, UMR 1333 Laboratoire DGIMI, 34095 Montpellier (France)
and
Université Montpellier 2, UMR 1333 Laboratoire DGIMI
34095 Montpellier (France)

[+] These authors contributed equally to this work.

💻 Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/chem.201302481.

Results and Discussion

During our screening for novel natural products by UPLC-ESI-MS of XAD-16 extracts[17] from *Xenorhabdus* cultures in standard LB medium we identified a new family of natural products. According to characteristic neutral losses for amino acids such as valine, leucine, and isoleucine in their MS^2 fragmentation patterns, these compounds were identified as peptides.[18] Their molecular formulas were determined on the basis of MALDI-Orbitrap-MS data (Table S1 of the Supporting Information) revealing $C_{64}H_{109}N_{13}O_{15}$ for xenoamicin A (**1**; *m/z* 1300.822 [*M*+H]+, Δppm 0.5) as the

main product in *X. mauleonii* DSM17908. The determination of the molecular formula was facilitated and confirmed by the identification of the correct number of carbon and nitrogen atoms from labeling experiments with strain DSM 17908 in standard growth medium, ^{13}C-labeled medium, and ^{15}N-labeled medium (Figure S1 of the Supporting Information), as described previously.[14, 16] Compound **1** (29.2 mg) was subsequently isolated from strain DSM17908 by preparative chromatography with HPLC-UV/ELSD in a three step protocol (Figures S2–S4 of the Supporting Information). The first two chromatographic steps were performed on C_{18} columns with different particle sizes under acidic conditions. In the third step the selectivity was changed through a higher pH value of the eluent. Due to the lack of a chromophore, **1** shows only weak absorption in the UV spectrum even at 215 nm. Despite several peaks with a strong absorption at 215 nm in the extract of *X. mauleonii* DSM17908, **1** could easily be identified with the universal signal of the evaporative light scattering detector (ELSD). Due to the relatively large amount of **1** in the extract of DSM17908 a large signal in the ELSD but not in the UV spectra could be observed, whereas other compounds in this extract with a chromophore gave large signals in both spectra (Figure S2 of the Supporting Information). Thus, the ELSD signal gives a very good indication of the actual amount of the individual compounds in the extract, independent of the physicochemical properties, and therefore ELSD is especially valuable for the detection of peptides without any chromophore.[19]

The structure of **1** was determined by comprehensive NMR studies. In most standard solvents, such as [D$_6$]DMSO, CD$_3$OD, or CDCl$_3$, NMR spectra of poor quality were obtained. The resonances of the amide protons in particular appeared as very broad signals (Figures S5, S6 and Annex 1 of the Supporting Information). However, in [D$_6$]benzene at 293 K well-resolved NMR spectra with reasonable dispersion of the amide signals were obtained (Figure S6 of the Supporting Information). The assignment of all proton and carbon resonances was carried out with various 2D NMR techniques, including DQF-COSY, TOCSY, ROESY, multiplicity-edited HSQC, HSQC-TOCSY, and HMBC spectra (Table 1, Annexes 1–20 in the Supporting Information). The assignment was hampered by the presence of two sets of signals in a ratio of 1:1 resulting from proline *cis/trans* conformers (see below). Primary analysis of the spin systems in the TOCSY and HSQC-TOCSY spectra revealed the presence of 13 amino acid residues: five valines, one leucine, one isoleucine, two alanines, two prolines, one threonine, and one β-alanine residue. In addition a butyric acid moiety was observed. The two sets of signals are caused by different orientations of the amide bond between the β-alanine residue and one of the proline residues. In the *cis* conformer an intensive ROE correlation between proline Hα and β-alanine Hβ is observed. In the the *trans* conformer such an intensive ROE should appear between proline Hδ and β-alanine Hβ. Unfortunately, this ROE cannot be assigned, because of peak overlaps between proline Hδ and

the Hβ of β-alanine. The presence of a *cis* and a *trans* conformer is also confirmed by the difference in ^{13}C chemical shifts of Cβ and Cγ of proline [$\Delta(\beta\gamma) = \delta(^{13}C\beta - ^{13}C\gamma)$].[20] The differences for the *cis* and *trans* conformers are $\Delta(\beta\gamma) = 8.83$ ppm and $\Delta(\beta\gamma) = 5.17$ ppm, respectively, which are in agreement with literature data.[20] Due to overlap of the carbon resonances in the carbonyl region and line broadening of some amide protons, the HMBC spectrum was not sufficient for the sequential assignment (see Annex 3 of the Supporting Information). Therefore, the peptide sequence was determined by analysis of ROE correlations which resulted in a depsipeptide structure with ring formation between the C-terminal valine residue and the threonine side chain at position 7 (Figure 1). The ring closure could be confirmed by an ROE between threonine Hβ and valine 14 Hα in both conformers, whereby the ROE of the *cis* conformer is only weak (Annex 13 of the Supporting Information). All other sequential NH/Hα ROE correlations of both conformers were observed with the exception of the correlation between valine 10 and alanine 11 in both conformers. Nevertheless, the connection of valine 10 and alanine 11 was in accordance with the molecular formula and was shown by MS fragmentation data (see below). The assignment of the single spin systems to the corresponding conformer was confirmed by additional long-range ROE correlations between valine 10 NH and isoleucine 6 Hα. Moreover, a long-range ROE correlation between threonine Hβ and alanine 11 Hα confirmed the assignment of the *trans* conformer. The N-terminal proline residue is acylated with butyric acid, as could be concluded from an ROE correlation between Hα of the acid and Hα of proline in position 2. Unfortunately, the alanine NH in position 11 of the *trans* conformer could not be detected due to very broad signals.

The configuration of the amino acids in **1** was elucidated by using the advanced Marfey's method[16, 21, 22] leading to the identification of the L configuration for isoleucine and both proline residues. The D configuration was revealed for leucine, *allo*-threonine, and both alanine residues. The ratio of L to D valine was found to be 1:4. Therefore, four valine residues in **1** may have the D configuration and one valine residue has the L configuration (Figure 2, Table S2 of the Supporting Information). However, the stereochemical assignment of the valine residues was not possible at this stage.

To verify and support the structure elucidation by NMR spectroscopy and to enable the structural elucidation of further derivatives in *X. doucetiae* DSM17909, MS-based structure elucidation was additionally performed for **1** in DSM1908. To confirm the building blocks, standard ^{12}C-L-amino acids were added to bacterial cultures growing in a reversed labeling experiment in fully ^{13}C labeled medium (Figure S7 of the Supporting Information), as described previously.[14, 16] Incorporation could be detected by HPLC-ESI-MS as shifts to lower masses of *m/z* 1.5 for β-alanine, 2 for threonine, 2.5 for proline, 3 for isoleucine, 3 for leucine, 1.5 for alanine, and 2.5 for valine, related to half of the numbers of carbon atoms incorporated due to the presence of doubly charged ions.
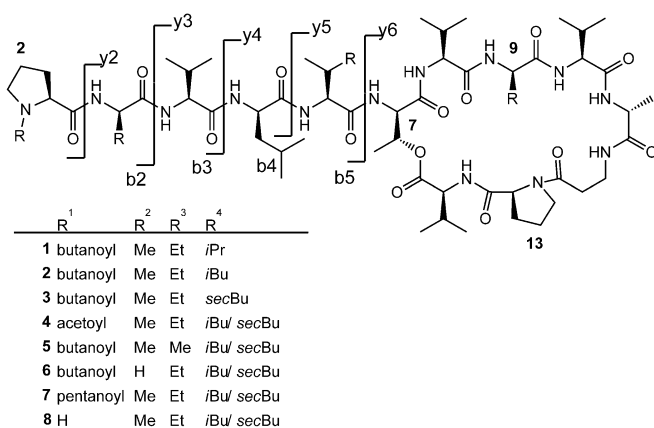
Table 1. NMR spectroscopic data of conformer mixture (600 MHz (^{1}H), 150 MHz (^{13}C) in [D$_6$]benzene) of xenoamicin A (**1**), δ in ppm. Spin system numbering from acylated N- to C-terminus. Only key sequential ROESY correlations are listed. Pos. = Position.

Amino acid	Pos.	trans Conformer			cis Conformer		
		δ_C	δ_H	ROESY	δ_C	δ_H	ROESY
1-BA	α	36.46	2.34	2-α	36.46	2.34	2-α
	α	36.46	1.95		36.46	1.95	
	β	18.63	1.90		18.63	1.90	
	β		1.85			1.85	
	γ	14.23	0.98		14.23	0.98	
2-Pro	α	60.30	4.58	1-α, 3-NH	60.30	4.58	1-α, 3-NH
	β	29.54	1.82	3-NH	29.54	1.82	3-NH
	β		1.47	3-NH		1.47	3-NH
	γ	24.96	2.08		24.96	2.08	
	γ		1.29			1.29	
	δ	47.13	3.28		47.13	3.28	
	δ		2.73			2.73	
3-Ala[a]	NH		8.71	2-α, 2-β		8.71	2-α, 2-β
	α	48.06[b]	5.01	4-NH	48.01[b]	5.03	4-NH
	β	16.56	1.69		16.56	1.67	
4-Val[a]	NH		7.82	3-α		7.82	3-α
	α	60.04	4.84	5-NH	60.08	4.78	5-NH
	β	30.70	2.25		30.70	2.25	
	γ	19.63	1.10		19.64	1.06	
	γ	19.45	1.04		19.38	1.00	
5-Leu	NH		9.22	4-α		9.23	4-α
	α	50.76	5.14	6-NH	50.93	5.13	6-NH
	β	40.34	2.03		40.34	2.00	
	γ	25.56[b]	1.97		25.50[b]	1.97	
	δ	22.59	1.09		22.59	1.09	
	δ	23.45	1.10		23.45	1.10	
6-Ile	NH		8.70	5-α		8.74	5-α
	α	58.05	5.00	7-NH, 7-β, 10-NH	58.01	4.92	7-NH, 7-β, 10-NH
	β	34.91	2.34		34.83	2.34	
	γ1	9.75	0.96		9.75	0.94	
	γ2	25.44	1.72		25.39	1.72	
	γ2		1.52			1.52	
	δ	15.91	1.14		15.87	1.12	
7-Thr	NH		9.60	6-α		9.54	6-α
	α	56.71	5.18	8-NH	56.14	5.15	8-NH
	β	70.03	6.05	6-α, 11-α, 14-α	71.82	5.82	6-α, 14-α
	γ	17.66	1.47		17.24	1.49	
8-Val	NH		9.03	7-α, 7-β		8.84	7-α, 7-β
	α	61.88	4.62	9-NH	61.54	4.56	9-NH
	β	29.36	2.57		29.54	2.56	
	γ	19.72	1.37		19.71	1.33	
	γ	19.55	1.27		19.93	1.28	
9-Val	NH		8.97	8-α		8.85	8-α
	α	60.82	4.76	10-NH	60.57	4.79	10-NH
	β	30.45[b]	2.55		30.49[b]	2.56	
	γ	19.99	1.24		19.91	1.25	
	γ	19.99	1.24		19.91	1.25	
10-Val	NH		9.15	6-α, 9-α		9.01	6-α, 9-α
	α	60.30	4.90		58.64	5.49	
	β	29.68	2.56		31.07	2.38	
	γ	19.44[b]	1.23		19.50[b]	1.23	
	γ	19.36	1.18		19.09	1.17	
11-Ala	NH					9.35	
	α	49.93	5.45	7-β, 12-NH	50.31	5.16	12-NH
	β	17.24	1.75		19.29	1.58	
12-β-Ala	NH		7.97	11-α		6.70	11-α
	α	35.66	4.20		35.79	4.42	13-α
	α		3.38			4.26	
	β	34.80	2.84		35.13	3.00	
	β		2.18			2.35	13-α

Table 1. (Continued)

Amino acid	Pos.	trans Conformer			cis Conformer		
		δ_C	δ_H	ROESY	δ_C	δ_H	ROESY
13-Pro	α	63.19	4.47		61.47	4.26	12-α, 12-β
	β	30.21	2.17		31.39	1.76	
	β		1.73			1.66	
	γ	25.04	1.40		22.56	1.26	
	γ		1.68			1.11	
	δ	46.57	3.20	14-NH	46.48	3.54	14-NH
	δ		2.81	14-NH		3.22	14-NH
14-Val	NH		6.60	7-β, 13-δ		6.65	13-δ
	α	57.70	4.69	7-β	56.14	4.51	7-β
	β	32.77	2.27		32.05	2.17	
	γ	18.83	0.97		18.81	0.66	
	γ	18.33	0.96		16.98	0.63	

[a] Interchangeable spin systems in the different conformers at the same position. [b] Interchangeable positions in the spin systems of the different conformers at the same position.
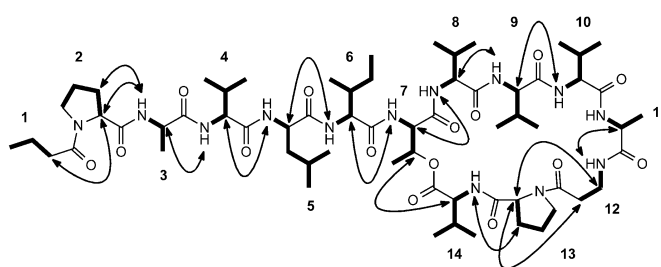
Figure 1. Connectivities in xenoamicin A (**1**), as determined by COSY and TOCSY (bold lines). Arrows indicate ROESY correlations used for the sequential assignment. Additionally the numeration of the spin systems is shown.

	R¹	R²	R³	R⁴
1	butanoyl	Me	Et	iPr
2	butanoyl	Me	Et	iBu
3	butanoyl	Me	Et	secBu
4	acetoyl	Me	Et	iBu/secBu
5	butanoyl	Me	Me	iBu/secBu
6	butanoyl	H	Et	iBu/secBu
7	pentanoyl	Me	Et	iBu/secBu
8	H	Me	Et	iBu/secBu

Figure 2. Structures of xenoamicins A–H (**1–8**) indicating b2–b5 and y2–y5 fragmentation ions of the MS^2 spectra. No differentiation between leucine or isoleucine at position 9 (R^4) could be obtained for **4–8**.

The more intense doubly charged ions were used for fragmentation and showed mainly b2–b5 and y2–y6 ions (Figure 2, Figure S8 and Table S3 of the Supporting Information), as xenoamicins exhibited a characteristic collision-induced dissociation (CID) MS^2 fragmentation pattern. However, as the ring stayed intact in the CID experiment, the sequence of the ring could not be determined. There-
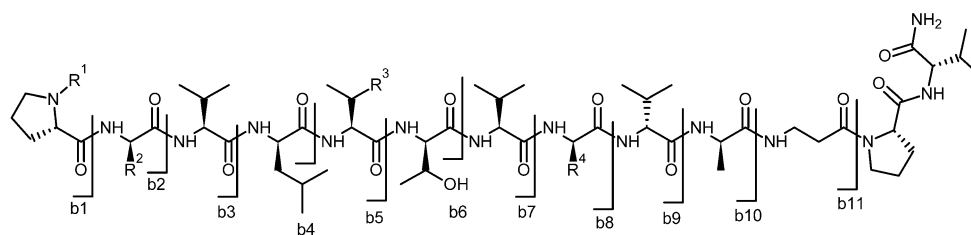
Figure 3. Structures of linear xenoamicins A–C (**1–3**) after reaction with NH$_3$. Described b1–b11 ions could be observed in MS^n fragmentations. For R^1–R^4 see Figure 2.
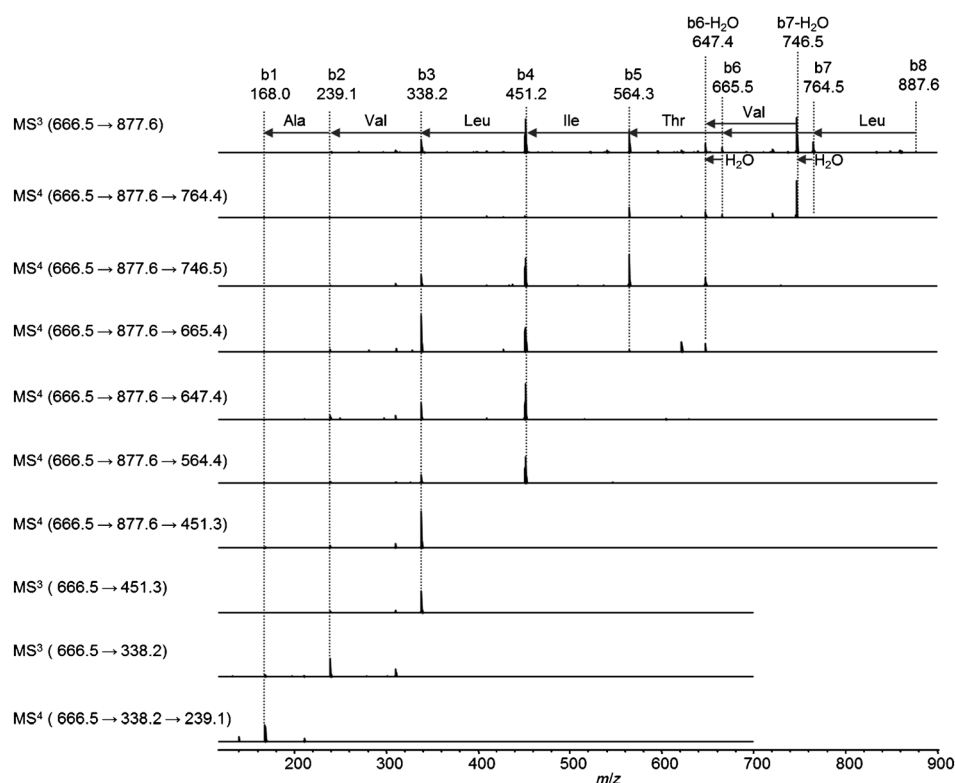

Figure 4. MS^3 and MS^4 spectra of linear xenoamicin B (**2**) after reaction with NH$_3$. According to MS^3 and MS^4 spectra, b2–b5 and b8 ions could be assigned in MS^2 spectra. Assignment of b9–b11 ions were proposed due to complex MS^3 and MS^4 fragmentation.

porting Information). Both show m/z 1314.8 $[M+H]^+$ indicating the presence of an additional methyl group in both compounds. The location of this additional methyl group in **2** and **3** could be identified as position 9 due to the difference between b7 and b8 ions (Figures 3 and 4; Figure S11 of the Supporting Information). In **1** the difference between the b7 and b8 ions of 99 Da represents the neutral loss of a valine building block, whereas the difference in **2** and **3** of 113 Da indicates leucine or isoleucine, respectively. Isobaric leucine and isoleucine could be distinguished by analyzing the results from inverse feeding experiments (Figure S7 of the Supporting Information). By feeding ^{12}C-leucine to a culture in ^{13}C-enriched media two shifts of 3 Da from the double charged species could be seen for **2** and show the incorporation of two leucine residues instead of one leucine residue in **1** and **3**. For **3** two shifts for isoleucine could be observed instead of one shift in **1** and **2** after the feeding of ^{12}C-isoleucine. The extracts of the inverse feeding experiments were also hydrolyzed with 28% aqueous NH$_3$ solution. However, the intensity of the b6 and b7 ions of the ^{13}C-peptides with incorporated ^{12}C-leucine or ^{12}C-isoleucine was too low for detection (Table S4 of the Supporting Information). In summary, MS^4 fragmentation and CID MS^2 fragmentation of the linearized peptides together with inverse feeding experiments were sufficient for the complete sequential structural elucidation of **1–3**.

Additionally, xenoamicins D–H (**4–8**) could be detected in DSM17909, but not in DSM17908, as doubly charged species ($[M+2H]^{2+}$, m/z 644.0, 651.0, 651.0, 665.0), although these were only produced in trace amounts (Table S1 of the

fore, xenoamicin-containing extracts were hydrolyzed with 28% aqueous NH$_3$ solution, which resulted in positive mass shifts of 8.5 Da and 9 Da in doubly charged species for the addition of NH$_3$ (17 Da; resulting in the linear amide) or for the addition of water (18 Da; resulting in the linear acid), respectively (Figure 3, Table S4 of the Supporting Information).[23] Based on CID MS^2 of the linearized peptide (Figure S9 of the Supporting Information) and MS^n experiments (Figure S10 of the Supporting Information), b1–b8 ions could be identified and also used for determination of the amino acid sequence of the ring, which confirmed the sequence of **1** as determined by NMR spectroscopy. The same analyses were performed for compounds **2** and **3**, which appear in *X. mauleonii* DSM17908 and *X. doucetiae* DSM17909 (Figures S8 and S9, Tables S3 and S4 of the Sup-

Supporting Information). Their structure could only partially be determined from their b2–b5 and y2–y6 ions in MS^2 spectra (Figure S12 and Table S5 of the Supporting Information). Determination of the ring amino acids of **4–8** was not possible. However, as variability in the ring was only observed at position 9 in compounds **1–3**, this may also be the variable position for compounds **4–8**, although no differentiation between leucine and isoleucine is possible. As described above, the differentiation of leucine and isoleucine can only be achieved by reversed feeding experiments and subsequent linearization if the structure elucidation is performed by MS. However, the amounts of **4–8** were not sufficient for further characterization.

Nevertheless, **5** and **6** show variability in positions 6 and 3, respectively. Furthermore, two additional N-terminal acyl moieties, acetate and pentanoate, were observed in **4** and **8**. Experimentally proven characterization of the ring amino acids is not essential for elucidation of these variabilities.

The b2 ion of **4** with a mass of 211 Da revealed that acetate is acylated with the N-terminal proline residue. The neutral loss of proline is 97 Da and that of alanine 71 Da. Therefore the remaining component of the b2 ions had a mass of 43 Da indicating acetate. By comparison of the b2 ions with 211 Da and the y2 ions with 1147 Da, alanine in position 3 is confirmed. Valine in position 4, leucine in position 5, and isoleucine in position 6 could be predicted by the mass shifts of 99, 113, and 113 Da between the y3, y4, y5, and y6 ions, respectively. Leucine and isoleucine cannot be distinguished by their mass, but it can be assumed that leucine is incorporated at position 5 and isoleucine at position 6, as in **1–3**. However, CID MS^2 fragmentation was not sufficient for the analysis of the ring. Therefore, it had to be assumed that the ring consists of the same or very similar building blocks as in **1–3**. However, because the y6 ions of **4** showed a mass of 751 Da, like for the y6 ions of **2** and **3**, we propose that leucine or isoleucine was incorporated at position 9, as in **2** and **3**, and that all other amino acids in the ring were as in **1–3**.

The CID MS^2 fragmentation revealed that compound **5** has alanine in position 3, valine in position 4, leucine in position 5, and valine in position 6 due to the shifts of 71, 99, 113, and 99 Da between the y2, y3, y4, y5, and y6 ions, respectively. Therefore this compound shows variability in position 6 due to the incorporation of valine instead of isoleucine. The b2 ions with a mass of 239 Da revealed a butanoyl acid unit in position 1. The obtained y6 ions for compound **5** exhibited the same size as for compound **4**, **6**, **7**, and **8**, thus the structure shown in Figure 2 resulted.

In compound **6** glycine instead of alanine is incorporated in position 3, as revealed by the mass shift of 57 Da between the y2 and y3 ions. According to the shifts between y3, y4, y5, and y6 ions, valine, leucine, and isoleucine are incorporated in positions 4, 5, and 6, respectively. The loss of 14 Da of the b2 ions in comparison with the b2 ions of **1**, **2**, **3**, and **5** resulted from incorporation of glycine instead of alanine.

Compound **7** showed the same mass shifts as **2–4**. Therefore, the same amino acid sequence can be assumed. How-

ever, the b2 ions with a mass of 253 Da showed that **7** contained a pentanoyl moiety as the difference of the y2 and the y3 ions indicated an alanine in position 3. When the mass of alanine and proline were subtracted from the mass of the b2 ions the mass of a pentanoyl moiety is left.

The MS^2 spectrum of **8** is more complex. Therefore the MS^2 spectrum of the sodium adducts of **8** is additionally presented (Figure S12 of the Supporting Information). The y2–y6 ions showed that **8** exhibited the same amino acid sequence and ring size as **2–4** and **7**. Taking the molecular mass of 1244 Da into account, it can be assumed that the N-terminus of **8** is not acylated.

To fully assign the absolute configuration of xenoamicin, which was not completely possible from the NMR data and amino acid analysis, the biosynthesis gene cluster was identified and analyzed. As the genome of *X. mauleonii* DSM17908 is not available yet, the already finished genome sequence of strain DSM17909 (originally strain FRM16) (http://www.genoscope.cns.fr/agc/microscope/home/index.php) was searched for a biosynthesis gene cluster encoding nonribosomal peptide synthetases (NRPS) involved in xenoamicin biosynthesis (*xab*) by using the antismash software tool.[24–26] The only candidate that fits the predicted biosynthesis gene cluster encodes five nonribosomal peptidases XabABCD and an aspartic acid decarboxylase XabE putatively involved in the formation of β-alanine.[14,27,28]

To prove that this gene cluster is indeed involved in xenoamicin biosynthesis, the gene *xabB* encoding the second NRPS was disrupted by plasmid insertion. Comparison between the *xabB:cat* mutant and the wild type showed complete loss of **1–8** in the mutant, whereas all other natural products such as such as xenorhabdins and xenocoumacins were still produced (Figures S13 and S14 of the Supporting Information).

Overall 13 modules and all domains for the biosynthesis of the xenoamicins could be identified (Figure 5, Table S6 of the Supporting Information). The presence of an adenylation (A) domain specific for β-alanine was readily detectable due to the differences in the specificity conferring code (Table S7 of the Supporting Information).[29–31] Analysis of the condensation (C) domains[32] revealed the presence of dual condensation/epimerization (C/E) domains[33] in modules 3, 5, 7, 9, and 11. Thus, it can be assumed that the amino acids incorporated in the previous modules 2, 4, 6, 8, and 10 are epimerized, which is in accordance with the results from the advanced Marfey's analysis (Table S2 of the Supporting Information). For example, exclusively L-proline and D-alanine were detected in the Marfey's analysis and only C and C/E domains were identified following the modules responsible for the incorporation of proline (modules 1 and 12) and alanine (modules 2 and 10), respectively. Furthermore, **1–3** were identical in both strains according to the MS analysis of the linearized peptides as well as the retention times of cyclic or linear peptides. Thus, the absolute configuration of all amino acids in xenoamicins can be predicted accordingly and this also allows the stereochemical assignment for the valine moieties to be made (Figure 2).

Figure 5. Proposed biosynthesis of xenoamicin A (**1**) in *X. doucetiae* DSM 17909, showing the nonribosomal peptide synthetases XabABCD involved in its biosynthesisis. Domains and modules are labeled: adenylation domain (A), peptidyl carrier protein domain (PCP), condensation domain (C), condensation/epimerization domain (C/E), thioesterase domain (Te).

With alternating L and D configuration in the peptide chain, xenoamicins belong to the class of peptides which can form β helices in membranes and hence ion channels. The best-studied example is the antibiotic gramicidin D,[34] which is also formed nonribosomally and requires two molecules to span a membrane. Recently, a novel class of ribosomally made alternating D/L peptides, namely, proteusins, which has strong cytotoxic activity and polytheonamide as most prominent member, has been identified and its biosynthesis elucidated.[35]

In contrast to these D/L peptides, **1** most likely cannot form such a β-helix due to the depsipeptide structure. Nevertheless, the presence of such a large number of hydrophobic amino acids and a N-terminal acyl moiety in **1** may indicate an interaction with the membrane. Hence, the bioactivity of xenoamicin A (**1**) was investigated. Whereas no antibacterial or antifungal activity could be observed in an agar diffusion assay with *E. coli* DH10B, *B. subtilis*, *M. luteus*, *P. aeruginosa*, *S. cerevisiae*, and *C. albicans*,[36] good activity against *Plasmodium falciparum* NF 54 (IC$_{50}$ of 2.35 μg mL^{-1}) and *Trypanosoma brucei* rhodesiense STIB900 (IC$_{50}$ of 6.41 μg mL^{-1}) was observed. Much weaker activity against *Trypanosoma cruci* Tulahuen C4 (IC$_{50}$ of 30.5 μg mL^{-1}), *L. donnovanni* MHOM-ET-67L82 (IC$_{50}$ of 50.1 μg mL^{-1}), and mammalian L6 cells (IC$_{50}$ of 68.5 μg mL^{-1}) indicated a specific target.[37] However, further analysis is needed to identify this target and also to elucidate the natural function of this novel class of compounds in the complex life cycle of *Xenorhabdus* with its nematode host and insect prey.

Experimental Section

Xenoamicin A (1): White powder; ^{1}H NMR (600 MHz, [D$_6$]benzene, 25 °C): see Table 1; ^{13}C NMR (150 MHz, [D$_6$]benzene, 25 °C, see Table 1; MS/MS (70 eV) and HRMS (MALDI): see Table S1 of the Supporting Information.

Cultivation and extraction: *X. doucetiae* DSM17909 and *X. mauleonii* DSM17908 were always cultivated at 30 °C, 200 rpm. For inoculation, 1 % overnight preculture and 2 % XAD-16 in LB medium were used. For analytical and preparative HPLC, LB medium was used (20 mL and 12 L, respectively). XAD-16 resin beads were collected with a sieve after 72 h, washed with a small amount of water, and extracted with methanol to yield the crude extract (9.5 g) after evaporation from the large-scale cultivation.

Feeding experiments: The strains DSM17909 and DSM17908 were cultivated at 30 °C overnight. The cell pellet of the overnight culture was washed once with ISOGRO-^{13}C medium (1 mL) and dissolved again in ^{13}C medium (1 mL). ISOGRO-^{13}C medium was prepared with ISO-GRO-^{13}C powder (1 g), K$_2$HPO$_4$ (1.8 g L^{-1}), KH$_2$PO$_4$ (1.4 g L^{-1}), and CaCl$_2$·x H$_2$O (11.1 mg L^{-1}) dissolved in H$_2$O (100 mL). The feeding culture in ^{13}C medium (5 mL) was inoculated with washed overnight culture (50 μL). After incubation for 6 h at 30 °C, a stock solution (50 μL, 100 mM) of the respective amino acid (β-alanine, L-α-threonine, L-α-proline, L-α-isoleucine, L-α-leucine, L-α-alanine, or L-α-valine) was added. Another two additions were carried out at 24 and 48 h. The cultures were extracted with ethyl acetate (5 mL) after incubation for 72 h. The extract was then prepared for HPLC-MS analyses.[14]

Hydrolysis of the extract: The extract was treated with ammonium hydroxide solution (Sigma-Aldrich, 28 %) to open the ring in the depsipeptides.[23] Hydrolysis was started with methanol extract solution (50 μL) and ammonium hydroxide solution (500 μL). After incubation for 2 h at 40 °C, the reaction mixture was neutralized with HCl (3 M) solution. After evaporation of the solvent, the sample was then prepared for HPLC-MS analyses.

HPLC and mass spectrometric analysis of the extract: Analysis of the extracts was carried out by using an Ultimate 3000 LC system from Dionex, coupled to an amaZon X electrospray ionization mass spectrometer from Bruker Daltonics. Peptides were separated on a C18 Column (ACQITY UPLC BEH, 1.7 μm, 2.1×50 mm, flow rate 0.6 mL min^{-1}, Waters). Acetonitrile/water containing 0.1% HCOOH was used as mobile phase under a linear gradient from 40–50% acetonitrile over 17.5 min for the separation. Collision-induced dissociation (CID) was performed on the ion trap in the amaZon X in positive mode with a scan range of m/z 100–1500.[17] HR-MALDI-MS data were obtained with a MALDI LTQ Orbitrap XL from Thermo Fisher Scientific equipped with a 337 nm nitrogen laser. The extract was diluted in acetonitrile for MALDI-MS analysis. 0.2 μL of the extract solution and 0.5 μL of a 20 mM 4-cloro-α-cyanocinnamic acid (ClCCA) solution (70% acetonitrile) were spotted on the MALDI target and air-dried. The following instrument parameters were used: laser energy, 25 μJ; automatic gain control, on; auto spectrum filter, off; resolution, 100 000; plate motion, survey CPS. Mass spectra were internally calibrated with calibration mixture 2 (Applied Biosystems, Sequazyme peptide mass standard kits). The molecular formulas were calculated by using the Qual Browser software V. 2.0.7 (Thermo Fisher Scientific).[18]

Advanced Marfey's method: Compound **1** (ca. 0.5 mg) was hydrolyzed with HCl (6 M, 0.8 mL) in a high-pressure Ace tube at 110°C for 16 h. The hydrolysate was evaporated to dryness and resuspended in H$_2$O (100 μL). To 50 μL of this solution NaHCO$_3$ (10 μL, 1 M) and N_α-(5-fluoro-2,4-dinitrophenyl)-L-leucinamide or -D-leucinamide (L-FDLA or D-FDLA, 100 μL of a 1% solution in acetone) were added. The reaction vials were covered with aluminum foil and placed in a water bath at 40°C for 1 h. Then the reaction mixtures were cooled to room temperature, quenched with HCl (10 μL, 1 M), and evaporated to dryness.[21,22] The residue was dissolved in methanol (400 μL). The analysis of L- and D-FDLA-modified amino acids was carried out by LC-MS, as mentioned above for the analysis of the extract. Acetonitrile/water containing 0.1% HCOOH was used as mobile phase under a linear gradient from 20–60% acetonitrile over 34 min for the separation. At same time, the commercial standard amino acids were prepared as reference.

Isolation: For isolation of **1**, the XAD extract (9.5 g) of *X. doucetiae* DSM 17909 was dissolved in 50 mL of DMSO/MeOH/iPrOH (7:2:1). The resulting suspension was centrifuged at 13.3×10^3 rpm for 4 min and the pellet was discarded after further analysis. The following preparative HPLC setup was used: 4 mL per injection by using a Gilson 231 XL sample injector with a sandwich injection method, a Gilson 402 syringe pump, a Gilson 215 liquid handler and two Varian Prep Star SD-1 pumps with a flow rate of 95 mL min^{-1} and an acetonitrile/water gradient from 0–1 min 5%, 1–16 min 5–95%, 16–19 min 95%, a Gilson 306 pump for the addition of 10% HCOOH buffer with a flow rate of 2.5 mL min^{-1}, a Dionex P580 pump for the makeup flow of 1 mL min^{-1} acetonitrile, a Jasco UV 975 recording at 215 nm, an MRA Active Splitter, a Varian 380LC ELSD, and a Luna C18 10 μm 50×50 column from Phenomenex (for the chromatogram, see Figure S2 of the Supporting Information). The obtained fractions were freeze-dried. The fractions containing **1** were combined and purified in a second step on the same instrument setup, but with a flow rate of 75 mL min^{-1} and an acetonitrile/water gradient from 0–1 min 45%, 1–16 min 45–65%, 16–19 min 65–95% and a Luna C18 5 μm 30×75 column. The buffer for the second step was changed to 3.33% trifluoroacetic acid buffer, which was added at a flow rate of 2.5 mL min^{-1} (Figure S3 of the Supporting Information). In a third purification step the selectivity of the chromatographic system was changed through the pH value. Therefore, an aqueous solution of ammonium acetate (50 g L^{-1}) was used as buffer. A flow rate of 75 mL min^{-1} with an acetonitrile/water gradient from 0–1 min 5%, 1–16 min 5–95%, 16–19 min 95% and a Luna C18 5 μm 30×75 column (Figure S4 of the Supporting Information) were used.

NMR spectroscopy: NMR spectra were recorded on a Bruker AVANCE 600 spectrometer operating at a ^{1}H frequency of 600.2 MHz and a ^{13}C frequency of 150.9 MHz. The instrument was equipped with a 5 mm TCI probe head. The final experiments were carried out with a sample of **1** (10 mg) dissolved in [D$_6$]benzene (600 μL) at 293 K. For structure eluci-

dation and complete assignment of proton and carbon resonances, 1D ^{1}H, 1D ^{13}C, DQF-COSY, ROESY (mixing time 150 ms, spinlock field 2 kHz), multiplicity-edited HSQC, and HMBC spectra were acquired. ^{1}H chemical shifts were referenced to the solvent signals (^{1}H: 7.15 ppm, ^{13}C: 128.0 ppm).

Two-dimensional homonuclear experiments (DQF-COSY, TOCSY, and ROESY) were performed with a spectral width of 11 ppm. Spectra were recorded with 1024 increments in t_1 and 4096 complex data points in t_2. For each t_1 value 4 (DQF-COSY), 8 (TOCSY), or 16 (ROESY) transients were averaged.

For the HSQC spectrum 1024 increments with 3072 complex data points in t_2 were collected with sweep widths of 8 ppm in the proton and 90 ppm in the carbon dimension. For each t_1 value four transients were averaged. For the HSQC-TOCSY spectrum 1024 increments with 3072 complex data points in t_2 were collected with a sweep width of 11 ppm in the proton and 80 ppm in the carbon dimension. For each t_1 value 32 transients were averaged, and a mixing time of 80 ms was used for the TOCSY transfer. The HMBC spectrum was acquired with a sweep width of 11 ppm in the proton and 200 ppm in the carbon dimension by using a defocusing delay of 62 ms (optimized for coupling constants of 8 Hz). A total of 32 transients were averaged for each of 1024 increments in t_1, and 4096 complex points in t_2.

Identification and verification of the gene cluster: The proposed gene cluster responsible for the biosynthesis of xenoamicins from *Xenorhabdus doucetiae* DSM17909 (originally strain FRM16) was analyzed by using PKS/NRPS Analysis Web-site (http://nrps.igs.umaryland.edu/nrps/).[24] To confirm the gene cluster, gene *xabB* was disrupted by insertion of a plasmid. *Xenorhabdus doucetiae* DSM17909 was naturally resistant to ampicillin, which was used to select *Xenorhabdus* after conjugation. Primer pair 5′-GAACTG*GCATGC*GGAAATTGAGGCGCAAC and 5′-ACAAGA*GAGCTC*CCTTCCTGAAGCGGTG were used to amplify parts of gene *xabB*, which were then digested and cloned into pDS132 vector and transformed into *E. coli* S17 (λpir). After conjugation, primer pair 5′-GCCCGATATCCTGTCATTG and 5′-ACATGTG-GAATTGTGAGCGG was used to verify the insertion of the plasmid in the mutants. The detailed method was used by us before.[38] The extract from the mutant was prepared as for the wild type, except for chloramphenicol in the LB medium for the mutant.

[1] H. B. Bode, *Angew. Chem.* **2009**, *121*, 6512–6514; *Angew. Chem. Int. Ed.* **2009**, *48*, 6394–6396.

[2] C. Corre, G. Challis, *Chem. Biol.* **2007**, *14*, 7.

[3] L. Song, F. Barona-Gomez, C. Corre, *J. Am. Chem. Soc.* **2006**, *128*, 14754–14755.

[4] H. B. Bode, *Curr. Opin. Chem. Biol.* **2009**, *13*, 224–230.

[5] H. Goodrich-Blair, *Curr. Opin. Microbiol.* **2007**, *10*, 225–230.

[6] E. E. Herbert, H. Goodrich-Blair, *Nat. Rev. Microbiol.* **2007**, *5*, 634–646.

[7] D. Ji, Y. Yi, G. H. Kang, Y. H. Choi, P. Kim, N. I. Baek, Y. Kim, *FEMS Microbiol Lett.* **2004**, *239*, 241–248.

[8] J. Crawford, C. Portmann, X. Zhang, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 10821–10826.

[9] B. McInerney, W. Taylor, M. Lacey, R. Akhurst, R. Gregson, *J. Nat. Prod.* **1991**, *54*, 785–795.

[10] T. Hjelmgaard, M. Givskov, J. Nielsen, *Org. Biomol. Chem.* **2007**, *5*, 344–348.

125

[11] A. O. Brachmann, S. Forst, G. M. Furgani, A. Fodor, H. B. Bode, *J. Nat. Prod.* **2006**, *69*, 1830–1832.

[12] J. Li, G. Chen, J. M. Webster, *Can. J. Microbiol.* **1997**, *43*, 770–773.

[13] S. W. Fuchs, A. Proschak, T. W. Jaskolla, M. Karas, H. B. Bode, *Org. Biomol. Chem.* **2011**, *9*, 3130–3132.

[14] H. B. Bode, D. Reimer, S. W. Fuchs, F. Kirchner, C. Dauth, C. Kegler, W. Lorenzen, A. O. Brachmann, P. Grün, *Chem. Eur. J.* **2012**, *18*, 2342–2348.

[15] G. Lang, T. Kalvelage, A. Peters, J. Wiese, J. F. Imhoff, *J. Nat. Prod.* **2008**, *71*, 1074–1077.

[16] 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.

[17] F. Grundmann, V. Dill, A. Dowling, A. Thanwisai, E. Bode, N. Chantratita, R. ffrench-Constant, H. B. Bode, *Beilstein J. Org. Chem.* **2012**, *8*, 749–752.

[18] S. W. Fuchs, C. C. Sachs, C. Kegler, F. I. Nollmann, M. Karas, H. B. Bode, *Anal. Chem.* **2012**, *84*, 6948–6955.

[19] C. R. Mitchell, Y. Bao, N. J. Benz, S. Zhang, *J. Chromatogr. B* **2009**, *877*, 4133–4139.

[20] M. Schubert, D. Labudde, H. Oschkinat, P. Schmieder, *J. Biomol. NMR* **2002**, *24*, 149–154.

[21] K. Fujii, Y. Ikai, H. Oka, M. Suzuki, K. Harada, *Anal. Chem.* **1997**, *69*, 5146–5151.

[22] K. Fujii, T. Shimoya, Y. Ikai, H. Oka, K. Harada, *Tetrahedron Lett.* **1998**, *39*, 2579–2582.

[23] I. Vallet-Gely, A. Novikov, L. Augusto, P. Liehl, G. Bolbach, M. Pechy-Tarr, P. Cosson, C. Keel, M. Caroff, B. Lemaitre, *Appl. Environ. Microbiol.* **2010**, *76*, 910–921.

[24] B. O. Bachmann, J. Ravel, *Meth. Enzymol.* **2009**, *458*, 181–217.

[25] M. H. Medema, K. Blin, P. Cimermancic, V. de Jager, P. Zakrzewski, M. A. Fischbach, T. Weber, E. Takano, R. Breitling, *Nucleic Acids Res.* **2011**, *39*, W339–W346.

[26] K. Blin, M. H. Medema, D. Kazempour, M. A. Fischbach, R. Breitling, E. Takano, T. Weber, *Nucleic Acids Res.* **2013**, *41*, W204–W212.

[27] S. F. Altschul, T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, D. J. Lipman, *Nucleic Acids Res.* **1997**, *25*, 3389–3402.

[28] T. Weber, K. J. Laiple, E. K. Pross, A. Textor, S. Grond, K. Welzel, S. Pelzer, A. Vente, W. Wohlleben, *Chem. Biol.* **2008**, *15*, 175–188.

[29] M. Rottig, M. H. Medema, K. Blin, T. Weber, C. Rausch, O. Kohlbacher, *Nucleic Acids Res.* **2011**, *39*, W362–W367.

[30] C. Rausch, T. Weber, O. Kohlbacher, W. Wohlleben, D. H. Huson, *Nucleic Acids Res.* **2005**, *33*, 5799–5808.

[31] T. Stachelhaus, H. D. Mootz, M. A. Marahiel, *Chem. Biol.* **1999**, *6*, 493–505.

[32] D. Konz, M. A. Marahiel, *Chem. Biol.* **1999**, *6*, R39–48.

[33] C. J. Balibar, F. H. Vaillancourt, C. T. Walsh, *Chem. Biol.* **2005**, *12*, 1189–1200.

[34] B. M. Burkhart, R. M. Gassman, D. A. Langs, W. A. Pangborn, W. L. Duax, V. Pletnev, *Biopolymers* **1999**, *51*, 129–144.

[35] M. F. Freeman, C. Gurgui, M. J. Helf, B. I. Morinaka, A. R. Uria, N. J. Oldham, S. Hahl, S. Matsunaga, J. Piel, *Science* **2012**, *338*, 387–390.

[36] D. Velasco, M. M. del Tomas, M. Cartelle, A. Beceiro, A. Perez, F. Molina, R. Moure, R. Villanueva, G. Bou, *J. Antimicrob. Chemother.* **2005**, *55*, 379–382.

[37] F. I. Nollmann, F. I. Nollmann, A. Dowling, M. Kaiser, K. Deckmann, S. Grösch, R. ffrench-Constant, H. B. Bode, H. B. Bode, *Beilstein J. Org. Chem.* **2012**, *8*, 528–533.

[38] A. O. Brachmann, S. A. Joyce, H. Jenke-Kodama, G. Schwär, D. J. Clarke, H. B. Bode, *ChemBioChem* **2007**, *8*, 1721–1728.

CHEMISTRY
A EUROPEAN JOURNAL

Supporting Information

Structure and Biosynthesis of Xenoamicins from Entomopathogenic *Xenorhabdus*

Qiuqin Zhou,[a] Florian Grundmann,[a] Marcel Kaiser,[b] Matthias Schiell,[c]
Sophie Gaudriault,[d] Andreas Batzer,[c] Michael Kurz,[c] and Helge B. Bode*[a]

chem_201302481_sm_miscellaneous_information.pdf

127

Content

2

Table S1. Sum formula for xenoamicines A-H (**1**-**8**) determined by HR-MALDI-MS analysis.

	sum formula	m/z of [M+H]$^+$ exp.	theo.	Δppm	m/z of [M+2H]$^{2+}$ theo.
1	$C_{64}H_{110}O_{15}N_{13}^+$	1300.8232	1300.8239	0.5	650.9156
2	$C_{65}H_{112}O_{15}N_{13}^+$	1314.8388	1314.8395	0.6	657.9237
3	$C_{65}H_{112}O_{15}N_{13}^+$	1314.8388	1314.8395	0.6	657.9237
4	$C_{63}H_{108}O_{15}N_{13}^+$	1286.8077	1286.8082	0.4	643.9081
5	$C_{64}H_{110}O_{15}N_{13}^+$	1300.8232	1300.8239	0.5	650.9156
6	$C_{64}H_{110}O_{15}N_{13}^+$	1300.8232	1300.8239	0.5	650.9156
7	$C_{66}H_{114}O_{15}N_{13}^+$	1328.8544	1328.8552	0.6	664.9315
8	$C_{61}H_{106}O_{14}N_{13}^+$	1244.7972	1244.7977	0.4	622.9028

Table S2. Stereochemistry of amino acid building blocks of **1** obtained from advanced Marfey`s method analyzed by HPLC-ESI-MS.

amino acid	m/z [M-H]$^+$	$t_{R,}$ L-FDLA	$t_{R,}$ LD-FDLA
L-proline	410.2	12.5	12.5
			14.9
D-alanine	384.1	15.7	12.3
			15.7
D-leucine	426.2	24	17.3
			24
L-isoleucine	426.2	17	17
			23.7
D-*allo*-Thr	414.2	11.8	10.4
			11.8
L-valine	412.2	14.8	14.8
D-valine		21	21

3

Table S3. CID-MS^2 fragments of xenoamicines A–C (**1**-**3**) and labeled **1** using doubly charged species for fragmentation.

	[M+2H]$^{2+}$	b2	b3	b4	b5	y2	y3	y4	y5	y6
1	C$_{64}$H$_{111}$N$_{13}$O$_{15}$	C$_{12}$H$_{19}$N$_2$O$_3$	C$_{17}$H$_{28}$N$_3$O$_4$	C$_{23}$H$_{39}$N$_4$O$_5$	C$_{29}$H$_{50}$N$_5$O$_6$	C$_{55}$H$_{97}$N$_{12}$O$_{13}$	C$_{52}$H$_{92}$N$_{11}$O$_{12}$	C$_{47}$H$_{83}$N$_{10}$O$_{11}$	C$_{41}$H$_{72}$N$_9$O$_{10}$	C$_{35}$H$_{61}$N$_8$O$_9$
	651	239.1	338.2	451.3	564.4	1133.6	1062.6	963.7	850.6	737.5
1 (^{13}C)	^{13}C$_{64}$H$_{111}$N$_{13}$O$_{15}$	^{13}C$_{12}$H$_{19}$N$_2$O$_3$	^{13}C$_{17}$H$_{28}$N$_3$O$_4$	^{13}C$_{23}$H$_{39}$N$_4$O$_5$	^{13}C$_{29}$H$_{50}$N$_5$O$_6$	^{13}C$_{55}$H$_{97}$N$_{12}$O$_{13}$	^{13}C$_{52}$H$_{92}$N$_{11}$O$_{12}$	^{13}C$_{47}$H$_{83}$N$_{10}$O$_{11}$	^{13}C$_{41}$H$_{72}$N$_9$O$_{10}$	^{13}C$_{35}$H$_{61}$N$_8$O$_9$
	683	251.1	355.2	474.3	593.4	1188.9	1114.9	1010.8	891.7	772.6
1 (^{13}C + Leu)	^{13}C$_{58}$C$_6$H$_{111}$N$_{13}$O$_{15}$	^{13}C$_{12}$H$_{19}$N$_2$O$_3$	^{13}C$_{17}$H$_{28}$N$_3$O$_4$	^{13}C$_{17}$C$_6$H$_{39}$N$_4$O$_5$	^{13}C$_{23}$C$_6$H$_{50}$N$_5$O$_6$	^{13}C$_{49}$C$_6$H$_{97}$N$_{12}$O$_{13}$	^{13}C$_{46}$C$_6$H$_{92}$N$_{11}$O$_{12}$	^{13}C$_{41}$C$_6$H$_{83}$N$_{10}$O$_{11}$	^{13}C$_{41}$H$_{72}$N$_9$O$_{10}$	^{13}C$_{35}$H$_{61}$N$_8$O$_9$
	680	251.1	355.2	468.3	587.4	1182.8	1108.9	1004.7	891.7	772.6
1 (^{13}C + Ile)	^{13}C$_{58}$C$_6$H$_{111}$N$_{13}$O$_{15}$	^{13}C$_{12}$H$_{19}$N$_2$O$_3$	^{13}C$_{17}$H$_{28}$N$_3$O$_4$	^{13}C$_{23}$H$_{39}$N$_4$O$_5$	^{13}C$_{23}$C$_6$H$_{50}$N$_5$O$_6$	^{13}C$_{49}$C$_6$H$_{97}$N$_{12}$O$_{13}$	^{13}C$_{46}$C$_6$H$_{92}$N$_{11}$O$_{12}$	^{13}C$_{41}$C$_6$H$_{83}$N$_{10}$O$_{11}$	^{13}C$_{35}$C$_6$H$_{72}$N$_9$O$_{10}$	^{13}C$_{35}$H$_{61}$N$_8$O$_9$
	680	251.1	355.2	474.3	587.4	1182.8	1108.9	1004.7	885.7	772.6
2	C$_{65}$H$_{113}$N$_{13}$O$_{15}$	C$_{12}$H$_{19}$N$_2$O$_3$	C$_{17}$H$_{28}$N$_3$O$_4$	C$_{23}$H$_{39}$N$_4$O$_5$	C$_{29}$H$_{50}$N$_5$O$_6$	C$_{56}$H$_{99}$N$_{12}$O$_{13}$	C$_{53}$H$_{94}$N$_{11}$O$_{12}$	C$_{48}$H$_{85}$N$_{10}$O$_{11}$	C$_{42}$H$_{74}$N$_9$O$_{10}$	C$_{36}$H$_{63}$N$_8$O$_9$
	658	239.1	338.2	451.3	564.4	1147.7	1077.6	977.6	864.5	751.4
3	C$_{65}$H$_{113}$N$_{13}$O$_{15}$	C$_{12}$H$_{19}$N$_2$O$_3$	C$_{17}$H$_{28}$N$_3$O$_4$	C$_{23}$H$_{39}$N$_4$O$_5$	C$_{29}$H$_{50}$N$_5$O$_6$	C$_{56}$H$_{99}$N$_{12}$O$_{13}$	C$_{53}$H$_{94}$N$_{11}$O$_{12}$	C$_{48}$H$_{85}$N$_{10}$O$_{11}$	C$_{42}$H$_{74}$N$_9$O$_{10}$	C$_{36}$H$_{63}$N$_8$O$_9$
	658	239.1	338.2	451.3	564.4	1147.7	1077.6	977.6	864.5	751.4

Table S4. Comparison of CID-MS^2 fragments of doubly charged **1**–**3**, ^{13}C labeled peptides and peptides from inverse feeding experiments with ^{12}C amino acids in ^{13}C media after reaction with NH_3 for linearization.

	[M+2H]$^{2+}$	b1	b2	b3	b4	b5	b6 – H$_2$O	b6	b7 – H$_2$O	b7	b8	b9	b10	b11
1 + NH$_3$	$C_{64}H_{114}N_{14}O_{15}$	$C_9H_{14}NO_2$	$C_{12}H_{19}N_2O_3$	$C_{17}H_{28}N_3O_4$	$C_{23}H_{39}N_4O_5$	$C_{29}H_{50}N_5O_6$	$C_{33}H_{55}N_6O_7$	$C_{33}H_{57}N_6O_8$	$C_{38}H_{64}N_7O_8$	$C_{38}H_{66}N_7O_9$	$C_{43}H_{75}N_8O_{10}$	$C_{48}H_{84}N_9O_{11}$	$C_{51}H_{89}N_{10}O_{12}$	$C_{54}H_{94}N_{11}O_{13}$
	659.5	168.1	239.1	338.2	451.3	564.4	647.5	665.5	746.5	764.4	863.5	962.6	1033.6	1104.7
2 + NH$_3$	$C_{65}H_{116}N_{14}O_{15}$	$C_9H_{14}NO_2$	$C_{12}H_{19}N_2O_3$	$C_{17}H_{28}N_3O_4$	$C_{23}H_{39}N_4O_5$	$C_{29}H_{50}N_5O_6$	$C_{33}H_{55}N_6O_7$	$C_{33}H_{57}N_6O_8$	$C_{38}H_{64}N_7O_8$	$C_{38}H_{66}N_7O_9$	$C_{44}H_{77}N_8O_{10}$	$C_{49}H_{86}N_9O_{11}$	$C_{52}H_{91}N_{10}O_{12}$	$C_{55}H_{96}N_{11}O_{13}$
	666.5	168	239.1	338.2	451.3	564.4	647.5	665.5	746.5	764.4	877.6	976.7	1047.7	1118.8
2 + NH$_3$ (13C)	$^{13}C_{65}H_{116}N_{14}O_{15}$		$^{13}C_{12}H_{19}N_2O_3$	$^{13}C_{17}H_{28}N_3O_4$	$^{13}C_{23}H_{39}N_4O_5$	$^{13}C_{29}H_{50}N_5O_6$					$^{13}C_{44}H_{77}N_8O_{10}$	$^{13}C_{49}H_{86}N_9O_{11}$	$^{13}C_{52}H_{91}N_{10}O_{12}$	
	699		251.1	355.2	474.4	593.5					921.8	1025.8	1099.9	
2 + NH$_3$ (13C + Leu)	$^{13}C_{53}C_{12}H_{116}N_{14}O_{15}$		$^{13}C_{12}H_{19}N_2O_3$	$^{13}C_{17}H_{28}N_3O_4$	$^{13}C_{17}C_6H_{39}N_4O_5$	$^{13}C_{23}C_6H_{50}N_5O_6$					$^{13}C_{32}C_{12}H_{77}N_8O_{10}$	$^{13}C_{37}C_{12}H_{86}N_9O_{11}$	$^{13}C_{40}C_{12}H_{91}N_{10}O_{12}$	
	693		251.1	355.2	468.4	587.5					909.7	1013.8	1087.8	
2 + NH$_3$ (13C + Ile)	$^{13}C_{59}C_6H_{116}N_{14}O_{15}$		$^{13}C_{12}H_{19}N_2O_3$	$^{13}C_{17}H_{28}N_3O_4$	$^{13}C_{17}C_6H_{39}N_4O_5$	$^{13}C_{23}C_6H_{50}N_5O_6$					$^{13}C_{38}C_6H_{77}N_8O_{10}$	$^{13}C_{43}C_6H_{86}N_9O_{11}$	$^{13}C_{46}C_6H_{91}N_{10}O_{12}$	
	696		251.1	355.2	474.4	587.5					915.7	1019.8	1093.9	
3 + NH$_3$	$C_{65}H_{116}N_{14}O_{15}$	$C_9H_{14}NO_2$	$C_{12}H_{19}N_2O_3$	$C_{17}H_{28}N_3O_4$	$C_{23}H_{39}N_4O_5$	$C_{29}H_{50}N_5O_6$	$C_{33}H_{55}N_6O_7$	$C_{33}H_{57}N_6O_8$	$C_{38}H_{64}N_7O_8$	$C_{38}H_{66}N_7O_9$	$C_{44}H_{77}N_8O_{10}$	$C_{49}H_{86}N_9O_{11}$	$C_{52}H_{91}N_{10}O_{12}$	$C_{55}H_{96}N_{11}O_{13}$
	666.5	168	239.1	338.2	451.3	564.4	647.5	665.5	746.5	764.4	877.6	976.7	1047.7	1118.8
3 + NH$_3$ (13C)	$^{13}C_{65}H_{116}N_{14}O_{15}$		$^{13}C_{12}H_{19}N_2O_3$	$^{13}C_{17}H_{28}N_3O_4$	$^{13}C_{23}H_{39}N_4O_5$	$^{13}C_{29}H_{50}N_5O_6$					$^{13}C_{44}H_{77}N_8O_{10}$	$^{13}C_{49}H_{86}N_9O_{11}$		
	699		251.1	355.2	474.4	593.5					921.8	1025.8		
3 + NH$_3$ (13C + Leu)	$^{13}C_{59}C_6H_{116}N_{14}O_{15}$		$^{13}C_{12}H_{19}N_2O_3$	$^{13}C_{17}H_{28}N_3O_4$	$^{13}C_{17}C_6H_{39}N_4O_5$	$^{13}C_{23}C_6H_{50}N_5O_6$					$^{13}C_{38}C_6H_{77}N_8O_{10}$	$^{13}C_{43}C_6H_{86}N_9O_{11}$	$^{13}C_{46}C_6H_{91}N_{10}O_{12}$	
	696		251.1	355.2	468.4	587.5					915.7	1019.8	1093.9	
3 + NH$_3$ (13C + Ile)	$^{13}C_{59}C_6H_{116}N_{14}O_{15}$		$^{13}C_{12}H_{19}N_2O_3$	$^{13}C_{17}H_{28}N_3O_4$	$^{13}C_{17}C_6H_{39}N_4O_5$	$^{13}C_{23}C_6H_{50}N_5O_6$					$^{13}C_{32}C_{12}H_{77}N_8O_{10}$	$^{13}C_{37}C_{12}H_{86}N_9O_{11}$	$^{13}C_{40}C_{12}H_{91}N_{10}O_{12}$	
	693		251.1	355.2	474.4	587.5					909.7	1013.8	1087.8	

5

Table S5. b2 – b5 and y2 – y6 ions from MS^2 spectra for the elucidation of compounds **4–8**.

	[M+H]$^+$	[M+2H]$^{++}$	b2	b3	b4	b5	y2	y3	y4	y5	y6
4	$C_{63}H_{108}N_{13}O_{15}$		$C_{10}H_{15}N_2O_3$	$C_{15}H_{24}N_3O_4$	$C_{21}H_{35}N_4O_5$	$C_{27}H_{46}N_5O_6$	$C_{56}H_{99}N_{12}O_{13}$	$C_{53}H_{94}N_{11}O_{12}$	$C_{48}H_{85}N_{10}O_{11}$	$C_{42}H_{74}N_9O_{10}$	$C_{36}H_{63}N_8O_9$
	1287	644	211.0	310.1	423.2	537.2	1147.7	1076.6	977.7	864.5	751.4
5	$C_{63}H_{108}N_{13}O_{15}$		$C_{12}H_{19}N_2O_3$	$C_{17}N_{28}N_3O_4$	$C_{23}H_{39}N_4O_5$	$C_{28}H_{48}N_5O_6$	$C_{55}H_{97}N_{12}O_{13}$	$C_{52}H_{92}N_{11}O_{12}$	$C_{47}H_{82}N_{10}O_{11}$	$C_{42}H_{74}N_9O_{10}$	$C_{36}H_{63}N_8O_9$
	1301	651	239.0	338.1	451.2	550.3	1134.6	1063.7	963.6	850.5	751.4
6	$C_{63}H_{108}N_{13}O_{15}$		$C_{11}H_{17}N_2O_3$	$C_{16}H_{26}N_3O_4$	$C_{22}H_{37}N_4O_5$	$C_{28}H_{48}N_5O_6$	$C_{55}H_{97}N_{12}O_{13}$	$C_{53}H_{94}N_{11}O_{12}$	$C_{48}H_{85}N_{10}O_{11}$	$C_{42}H_{63}N_8O_9$	$C_{36}H_{63}N_8O_9$
	1301	651	225.0	324.1	437.2	550.2	1133.7	1076.6	977.6	864.5	751.4
7	$C_{63}H_{108}N_{13}O_{15}$		$C_{13}H_{21}N_2O_3$	$C_{18}H_{30}N_3O_4$	$C_{24}H_{41}N_4O_5$	$C_{30}H_{52}N_5O_6$	$C_{56}H_{99}N_{12}O_{13}$	$C_{53}H_{94}N_{11}O_{12}$	$C_{48}H_{85}N_{10}O_{11}$	$C_{42}H_{74}N_9O_{10}$	$C_{36}H_{63}N_8O_9$
	1329	665	253.0	352.1	465.2	578.3	1147.7	1076.6	977.6	864.5	751.4
8	$C_{63}H_{108}N_{13}O_{15}$			$C_{13}H_{22}N_3O_3$	$C_{19}H_{33}N_4O_4$	$C_{25}H_{44}N_5O_5$			$C_{48}H_{85}N_{10}O_{11}$	$C_{42}H_{74}N_9O_{10}$	$C_{36}H_{63}N_8O_9$
	1245	623		268	381.2	494.3			977.6	864.5	751.4

Table S6. Identified genes *xabA–xabE* responsible for xenoamicine biosynthesis in strain DSM17909 and their closest homologue. C = condensation domain; A = adenylation domain; T = transferase domain; TE = thioesterase domain.

gene	presumed function	module	domain	position	closest protein homologue	identity/ positivity(%)	origin	accession no.
xabA	NRPS	1	C	36-482	putative Ornithine racemase	72/79	X. bovienii SS-2004	YP_003467869.1
		1	A	494-1027				
		1	T	1039-1103				
xabB	NRPS	2	C	1-388	XNC1_2040	60/67	*X. nematophila* HGB081	YP_003712280.1
		2	A	398-915				
		2	T	927-990				
		3	C/E	1057-1473				
		3	A	1481-2006				
		3	T	2018-2082				
		4	C	2099-2541				
		4	A	2553-3079				
		4	T	3091-3154				
		5	C/E	3221-3645				
		5	A	3657-4183				
		5	T	4195-4259				
		6	C	4276-4717				
		6	A	4724-5262				
		6	T	5276-5337				
		7	C/E	5404-5812				
		7	A	5820-6345				
		7	T	6357-6421				
xabC	NRPS	8	C	42-489	XNC1_2039	71/81	*X. nematophila* HGB081	YP_003712279.1
		8	A	505-1032				
		8	T	1049-1112				
		9	C/E	1179-1595				
		9	A	1607-2132				
		9	T	2140-2204				
xabD	NRPS	10	C	36-473	XNC1_2038	61/68	*X. nematophila* HGB081	YP_003712278.1
		10	A	501-1031				
		10	T	1053-1116				
		11	C/E	1186-1609				
		11	A	1634-2159				
		11	T	2171-2235				
		12	C	2252-2689				
		12	A	2703-3236				
		12	T	3248-3312				
		13	C	3329-3766				
		13	A	3778-4303				
		13	T	4311-4374				
		13	TE	4399-4654				
		13	TE	4674-4925				
xabE	aspartate decarboxylase				XNC1_2037	81/89	*X. nematophila* ATCC 19061	YP_003712277

Table S7. A-domain analysis performed by NRPS/PKS Predictor[1] of XabABCD from DSM 17909.

A domain	Residues in the binding pocket							Prediction
1	D	V	Q	F	I	A	X	X	NosA-M3-mePro
2	D	I	F	N	N	A	L	I	McyA-M2-Gly
3	D	A	W	W	L	G	G	T	TycC-M4-Val
4	D	A	W	L	L	G	A	V	CssA-M8-Leu
5	D	A	W	F	L	G	M	T	McyB-M1-Leu
6	D	F	W	N	X	X	M	V	NO HIT
7	D	A	W	W	L	G	G	T	TycC-M4-Val
8	D	A	W	F	L	G	M	T	McyB-M1-Leu
9	D	A	W	W	L	G	G	T	TycC-M4-Val
10	D	L	Y	N	N	A	L	–	McyA-M2-Gly
11	V	D	T	V	V	S	F	G	NO HIT
12	D	V	Q	F	I	A	X	X	NosA-M3-mePro
13	D	A	W	W	L	G	G	T	TycC-M4-Val
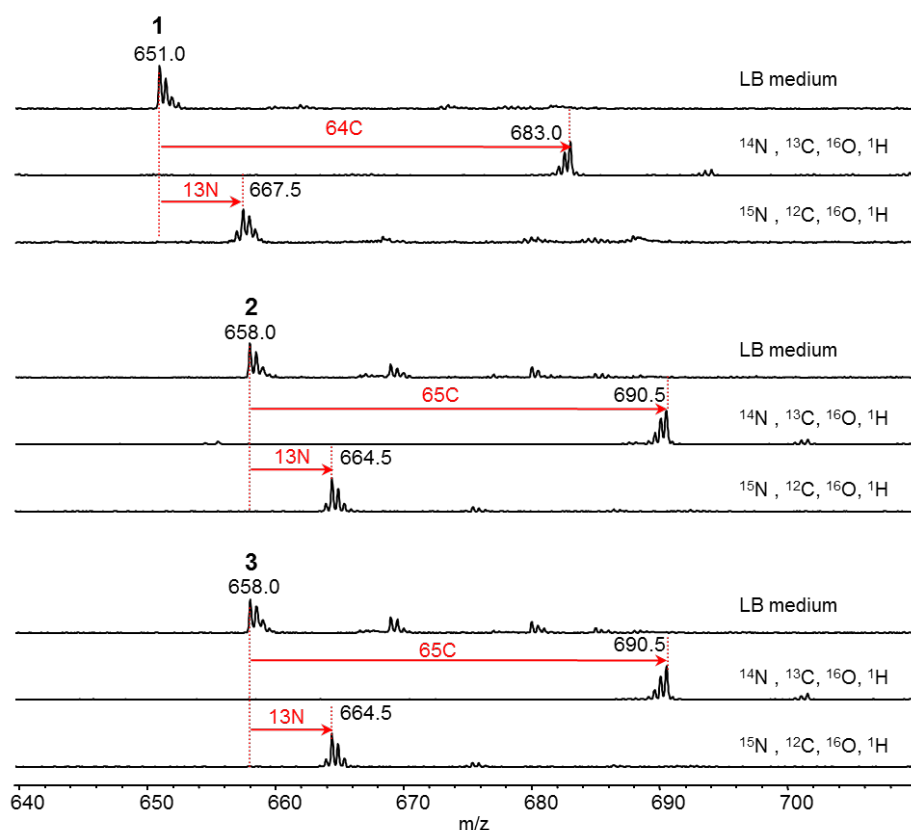
Figure S1. Determination of the number of carbon and nitrogen atoms for **1-3** by cultivation of DSM17908 and DSM17909 in standard growth medium (LB medium), ^{15}N labeled medium (^{15}N, ^{12}C, ^{16}O, ^{1}H), or ^{13}C labeled medium (^{14}N, ^{13}C, ^{16}O, ^{1}H). Compounds **1-3** were observed in MS as doubly charged species.
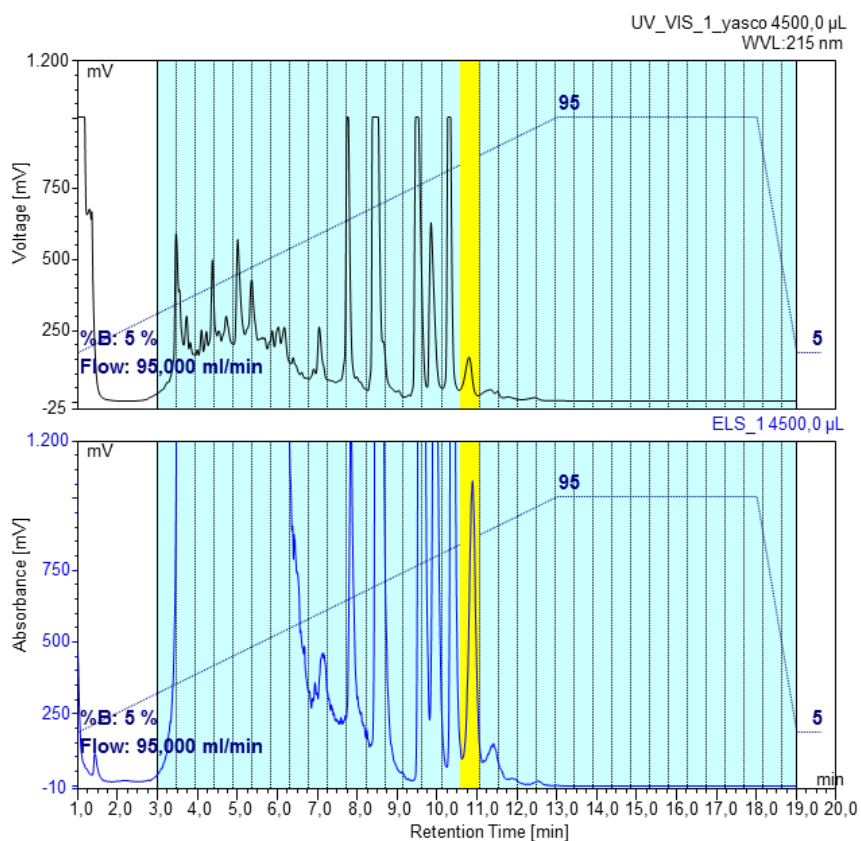
Figure S2. Fractionation of the crude extract; the peak at 10.8 min with low intensity at 215 nm in the UV spectra and high intensity within the ELSD spectra contained **1**. Fraction in yellow was taken for further purification.

9
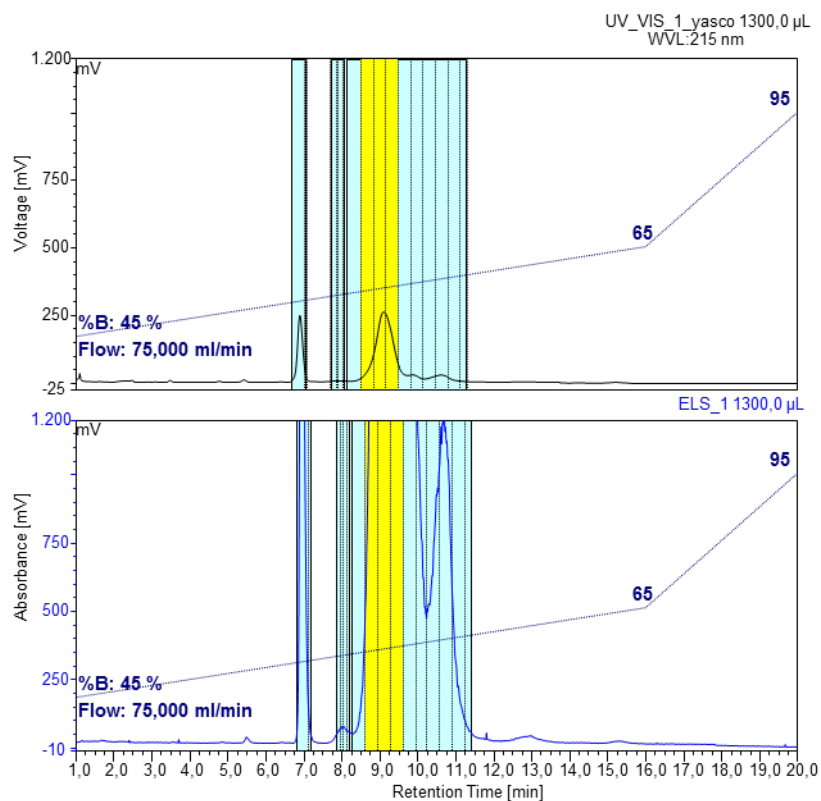
Figure S3. Second chromatographic step: Purification of the fraction containing **1**. Yellow fractions containing **1** were used for further purification.
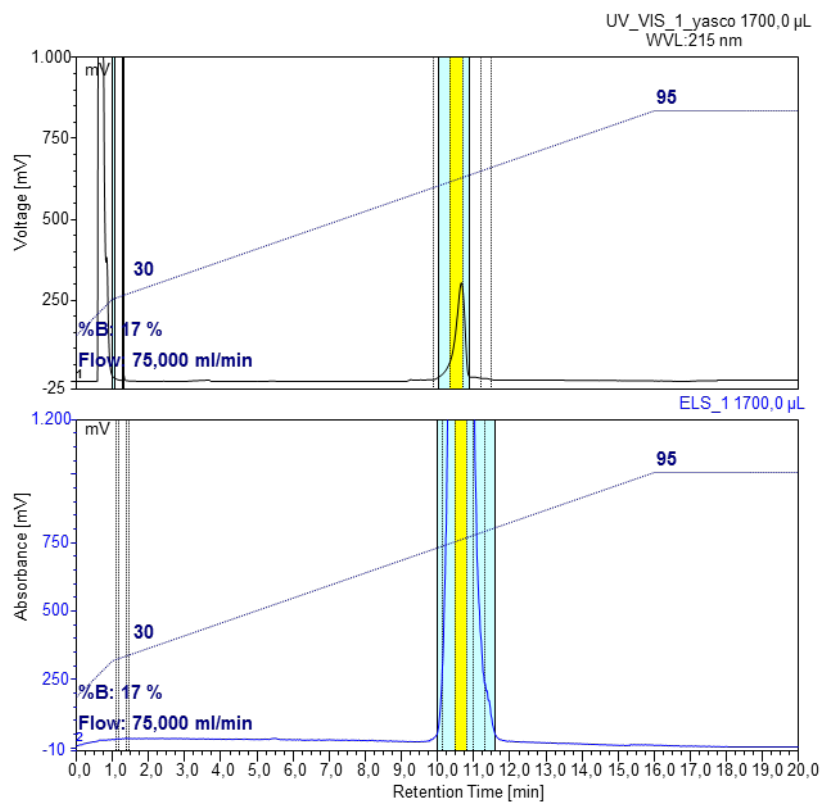
Figure S4. Third chromatographic step: Yellow fraction contained pure **1**.
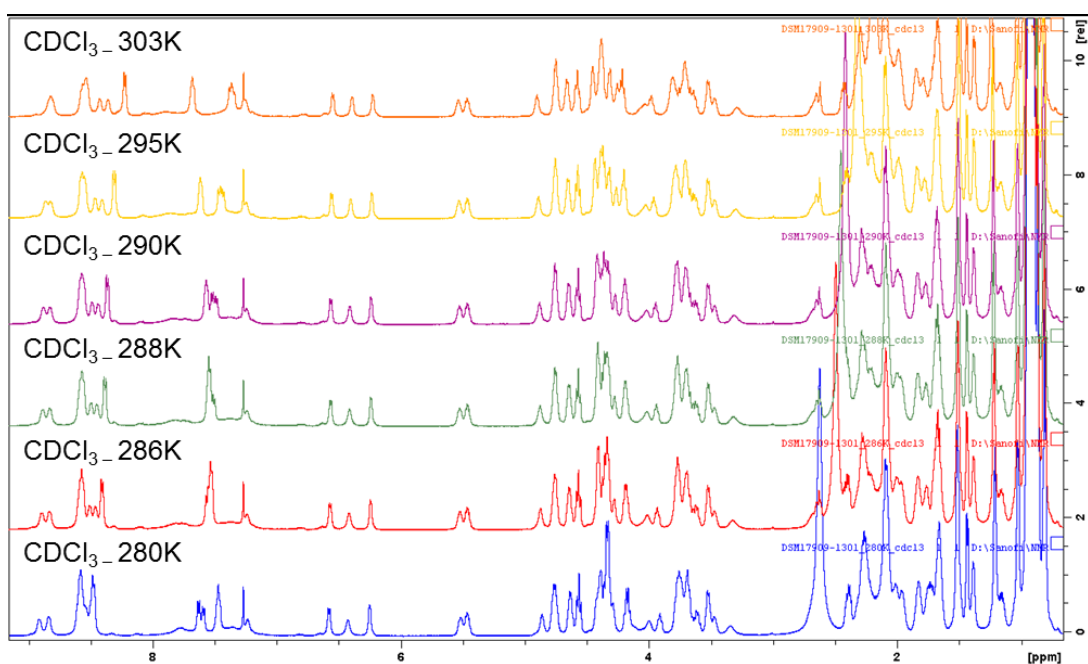
Figure S5. Stacked ^{1}H NMR spectra of **1** in CDCl$_3$ at different temperatures, mixture of conformers, 600MHz. Temperature changing did not improve the peak overlapping sufficiently.
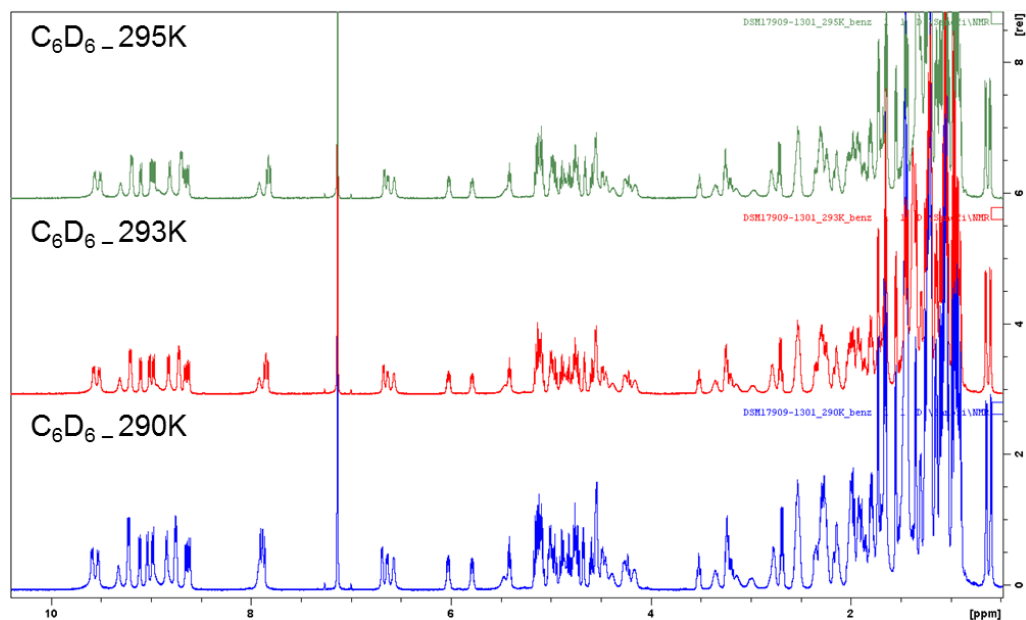
Figure S6. Stacked ^{1}H NMR spectra of **1** in d$_6$-benzene at different temperatures, mixture of conformers, 600MHz. Best peak separation was obtained at 293K.

11
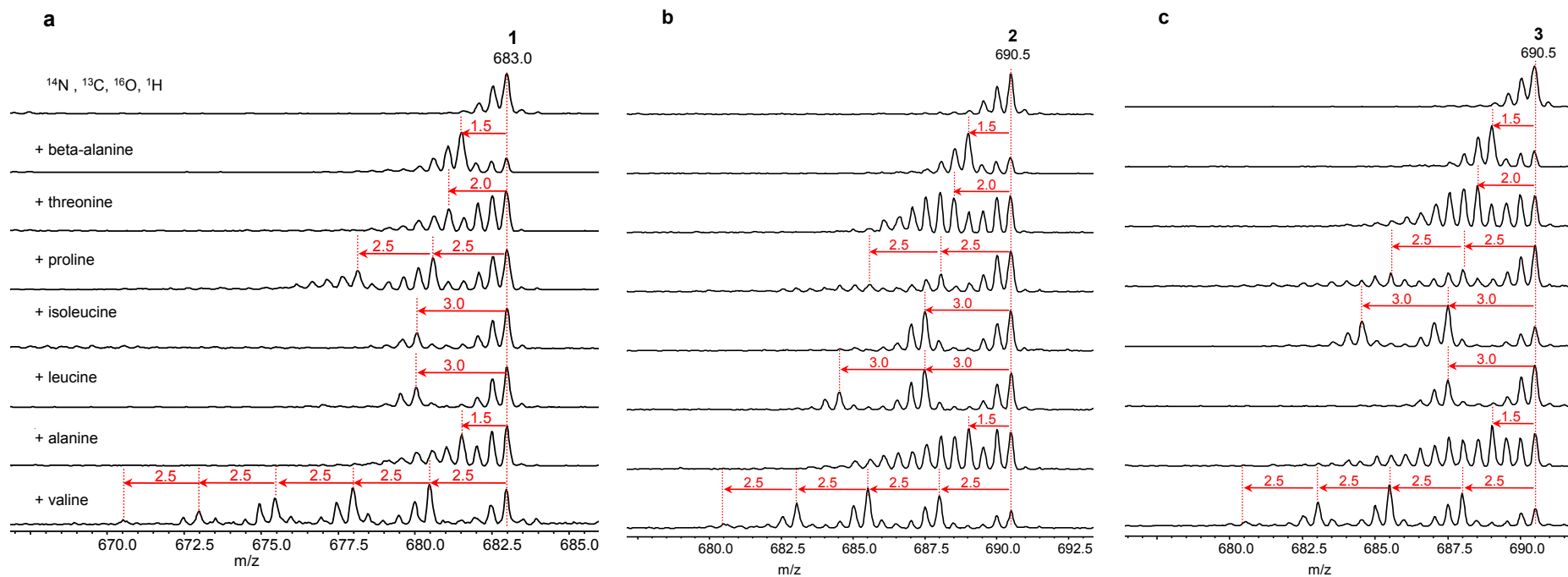
Figure S7. Determination of amino acid building blocks of **1** (a), **2** (b) and **3** (c) resulted from cultivation of strains DSM17908 (**1**) and DSM17909 (**2** and **3**) in ^{13}C labeled medium (^{14}N, ^{13}C, ^{16}O, ^{1}H) or with addition of different standard L-amino acids. Compounds **1**-**3** were observed in MS as doubly charged species.
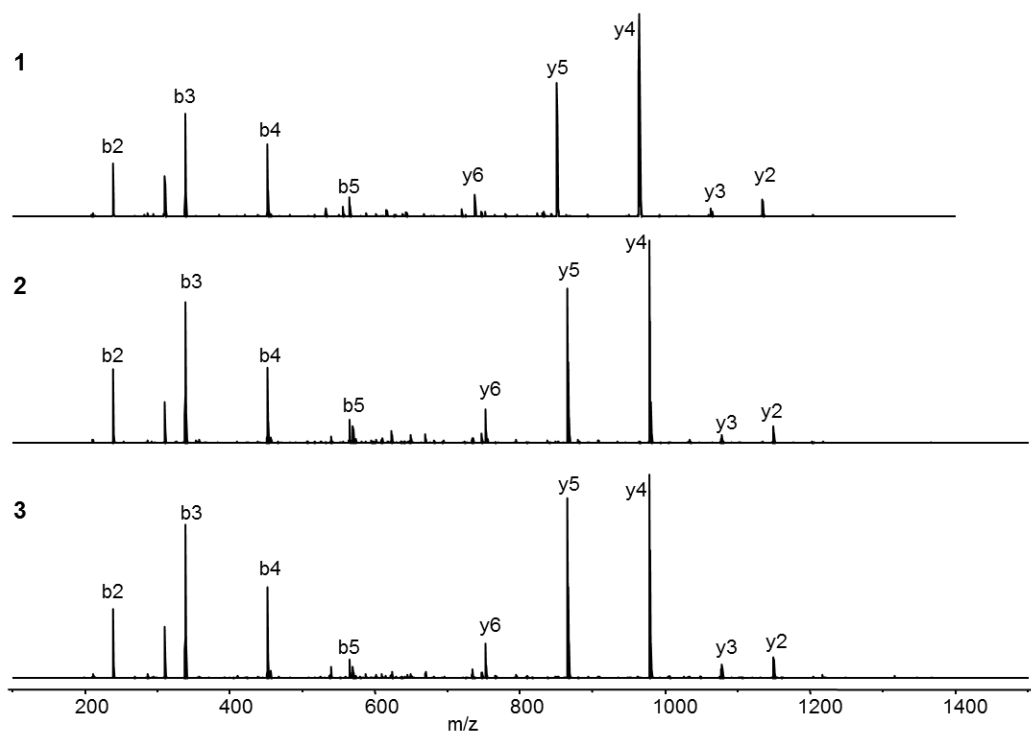
Figure S8. MS^2 spectra of **1**-**3** with assignment of the main peaks as b2 – b5 and y2 – y6 ions. Doubly charged species [M + 2H]$^{2+}$ were isolated for fragmentation.
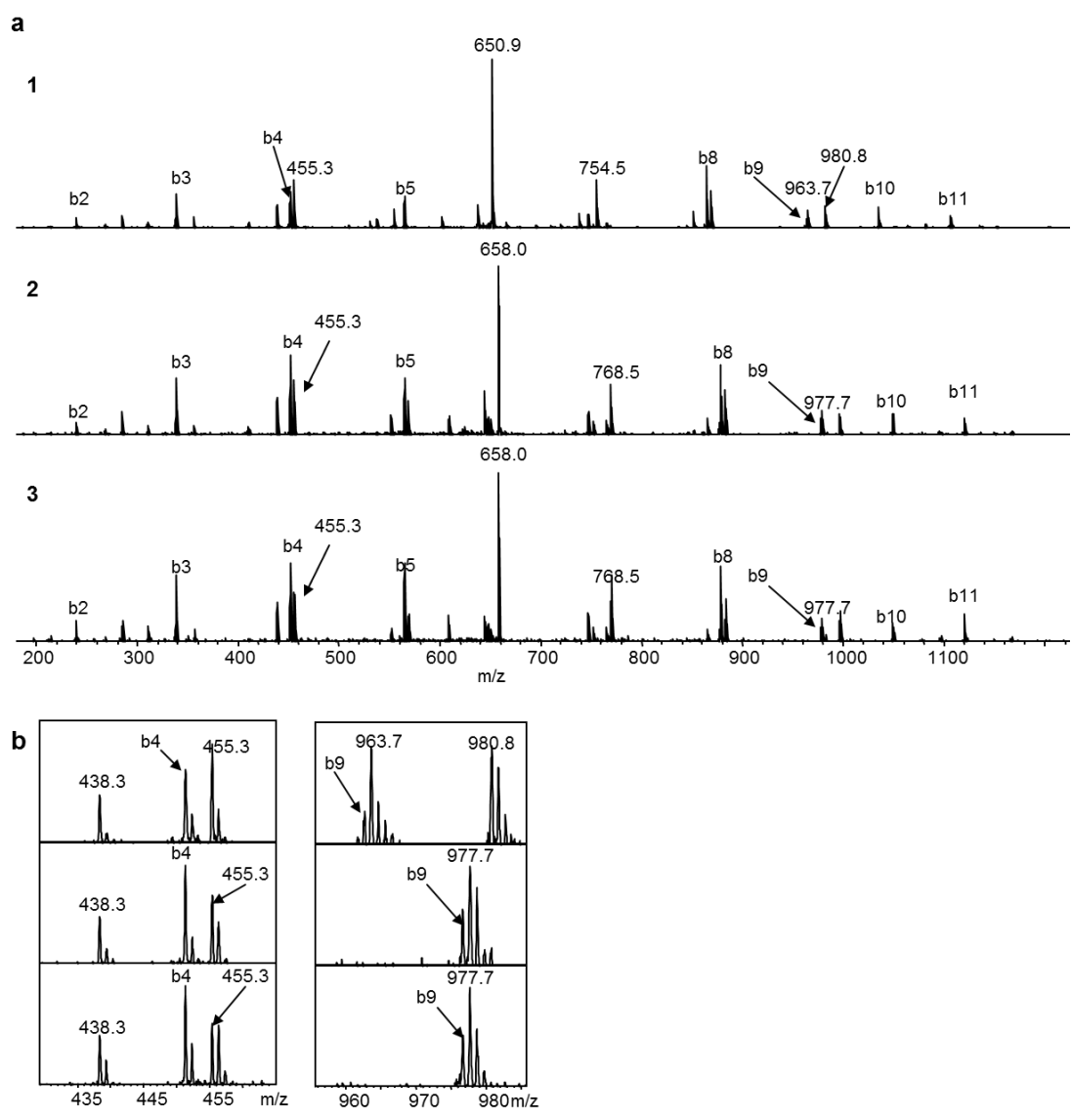
Figure S9. b2 – b11 fragments in MS^2 spectra of linear xenoamicines A–C (**1**–**3**), linear peptides were obtained by amidation with NH$_3$ (a); enlargement of MS^2 spectra (b).
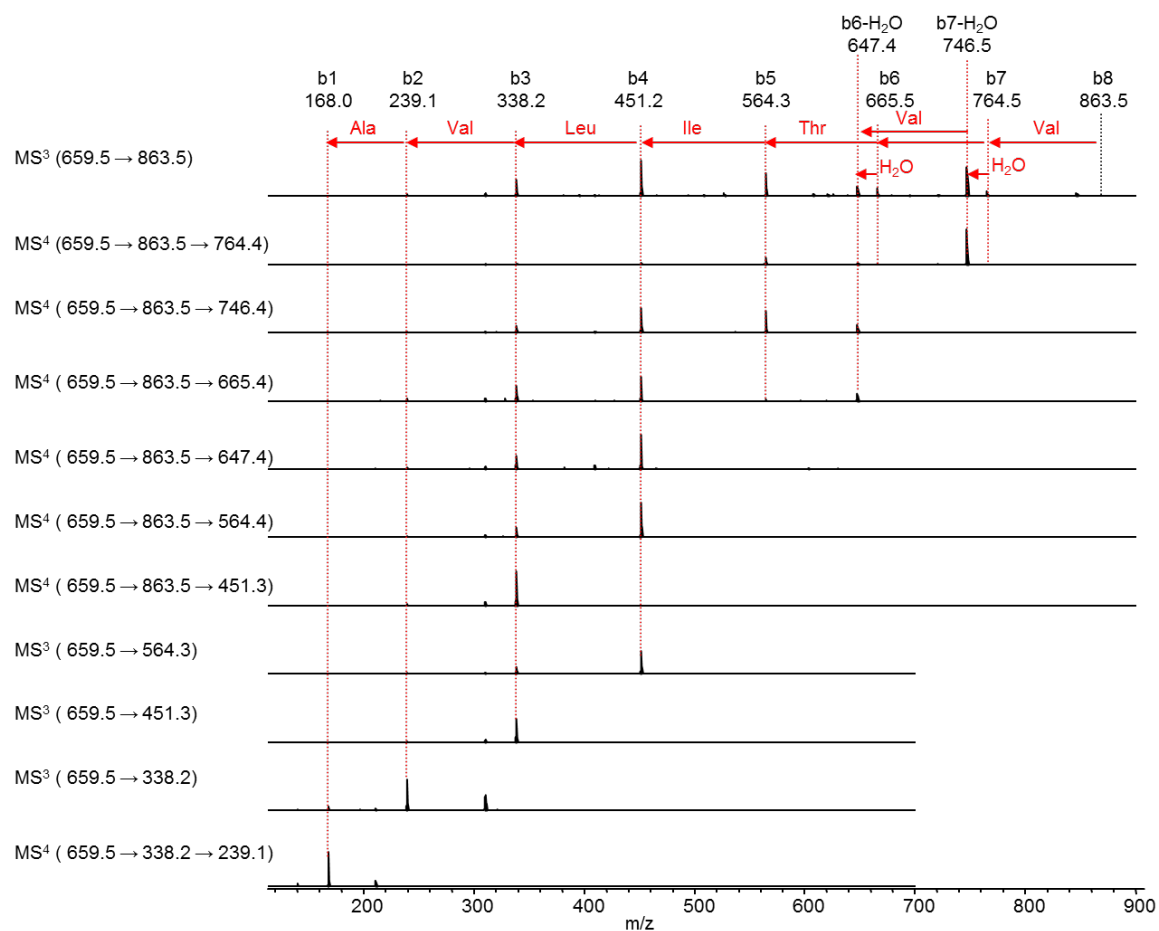
Figure S10. MS^3 and MS^4 spectra of linearized (after reaction with NH_3) **1** starting with the doubly charged species [M+2H]$^{2+}$ (666.5 m/z). Illustration of b1 – b8 fragments.
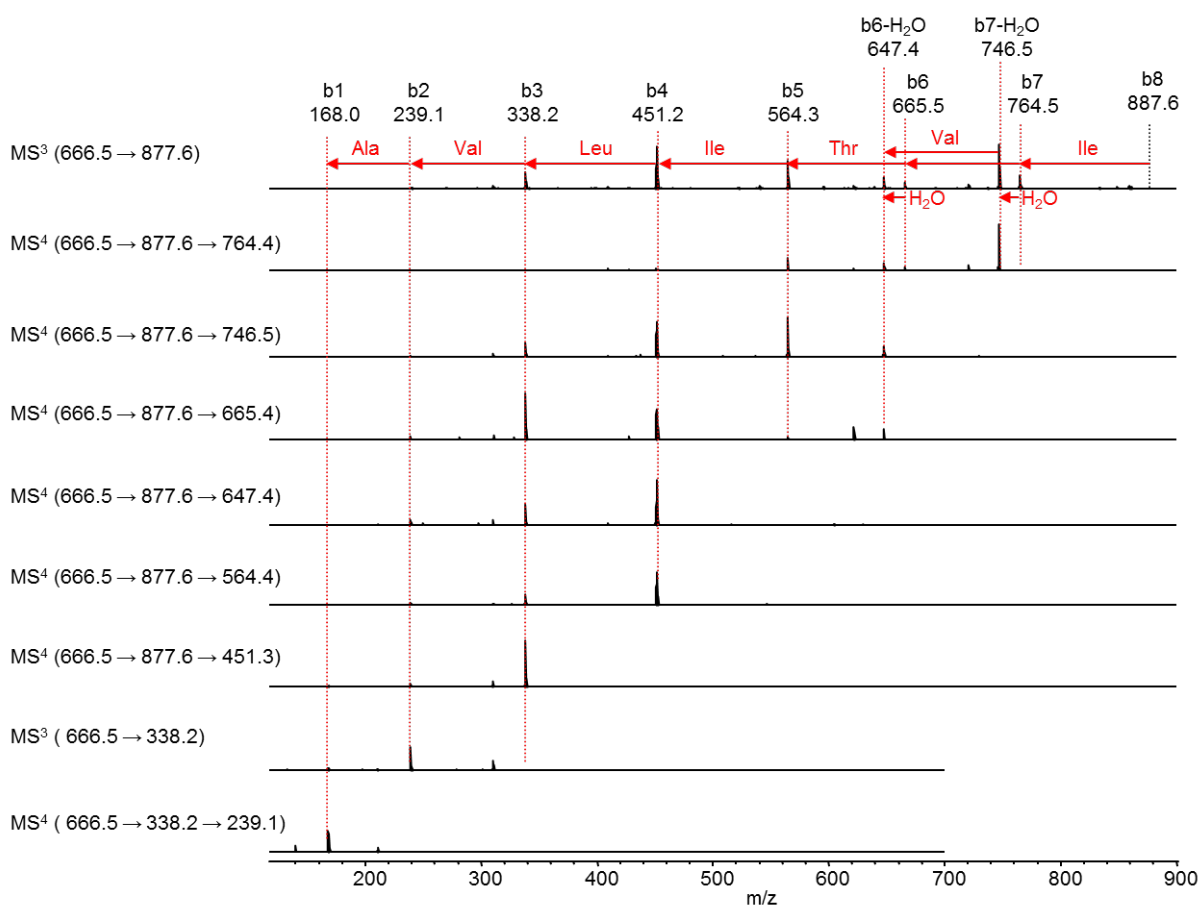
Figure S11. MS[3] and MS[4] spectra of linearized (after reaction with NH_3) **3**. Starting with the doubly charged species $[M+2H]^{2+}$ (666.5 m/z). Illustration of b1 – b8 fragments.
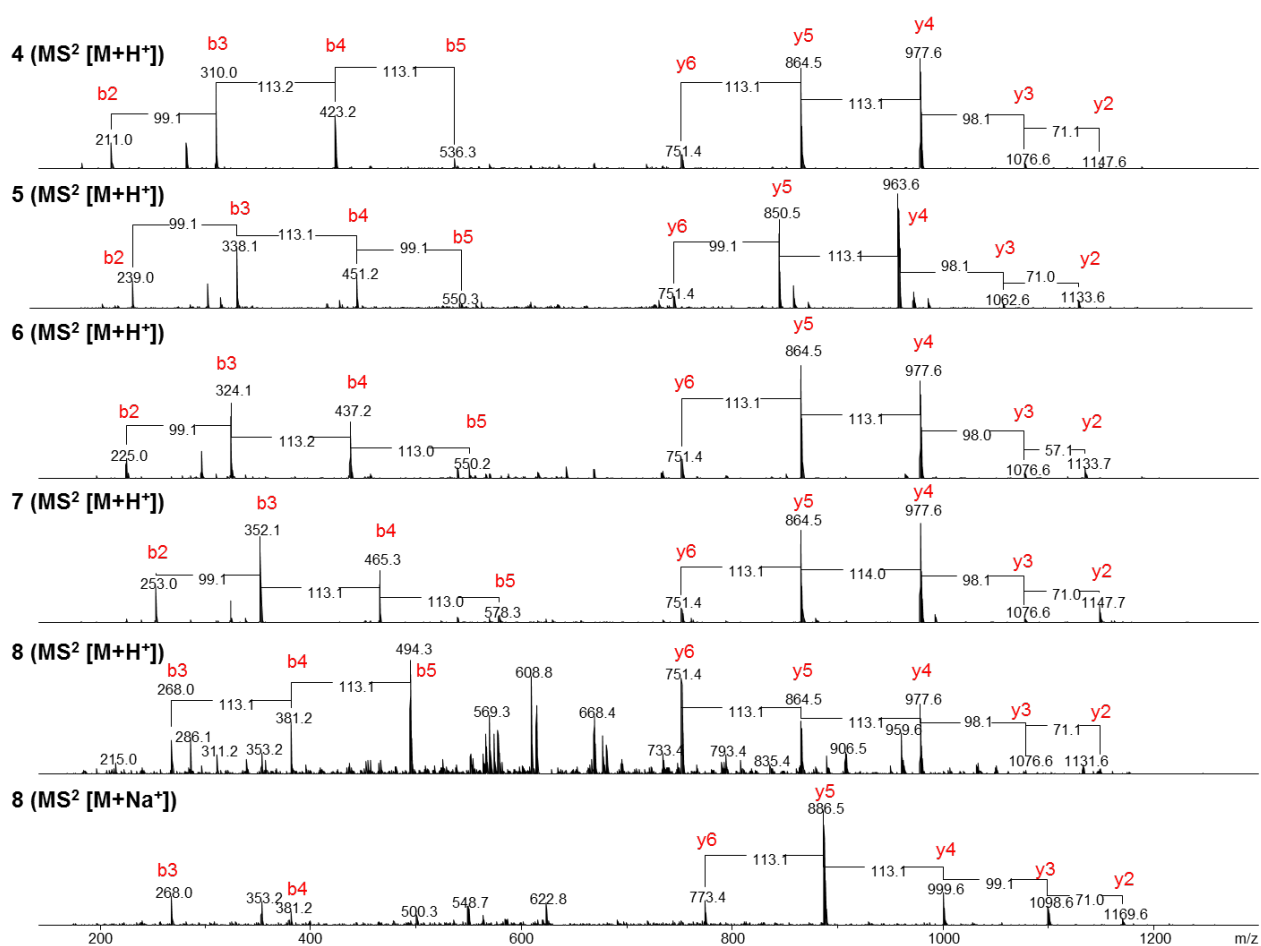
Figure S12. Structures and MS^2 spectra of xenoamicines D–H (**4**–**8**) present only in DSM 17909 with b2 – b5 and y2 – y6 ions. The neutral loss of 99 Da between y5 and y5 ions of **5** and the neutral loss of 57 Da between y2 and y3 ions of **6** show the variability in these positions. In **5** glycine instead of alanine is incorporated and in **6** valine instead of isoleucine is incorporated. The b2 ions of **4** and **7** reveal that the chain length of the acid, which is acylated with N- terminal proline, is different from other derivatives. The MS^2 spectrum of **8** is more complex. Therefore the MS^2 spectrum of the sodium adduct of **8** is additionally presented. y2 – y6 ions show that **8** has the same amino acid sequence and ring size like **2** and **3**. Therefore the N-terminus of **8** is not acylated.

17
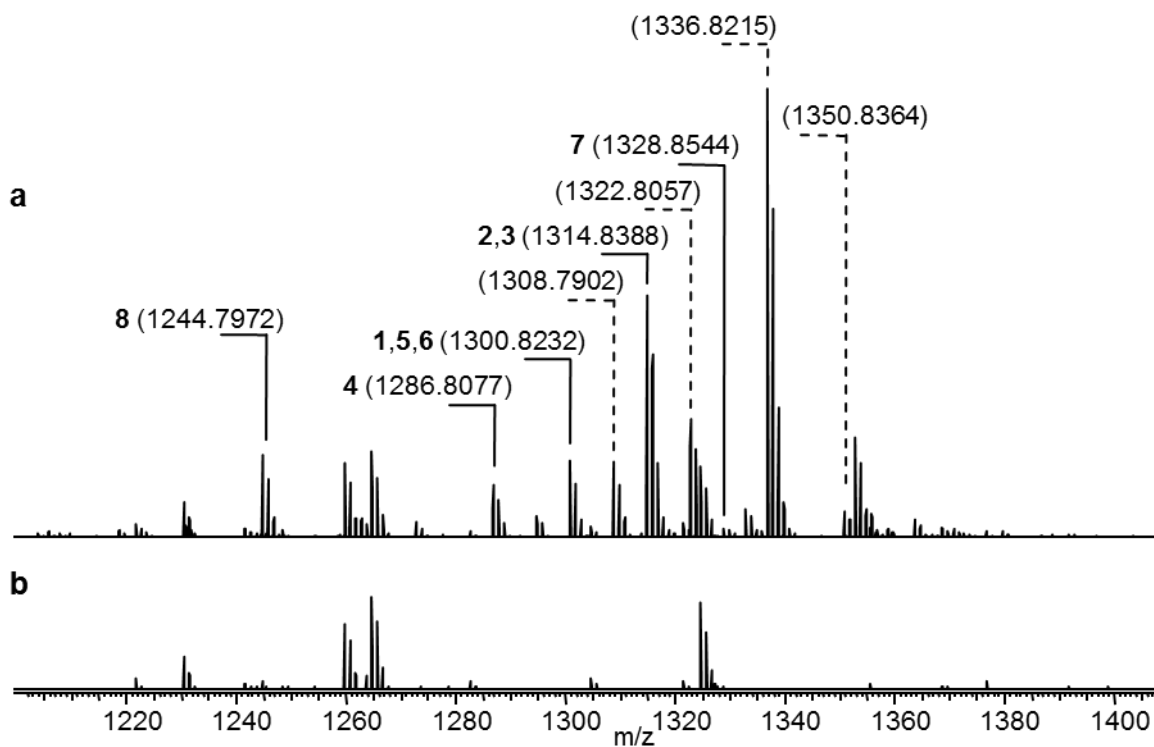
Figure S13. MALDI-Orbitrap spectra of DSM17909 wild type (a) and *xabB*::cat (b) knockout mutant. Highlighted monoisotopic peaks are xenoamicines protonated species [M+H$^+$], whereas other charged species ([M+Na]$^+$ and [M+K]$^+$) are not highlighted or shown with dashed lines.
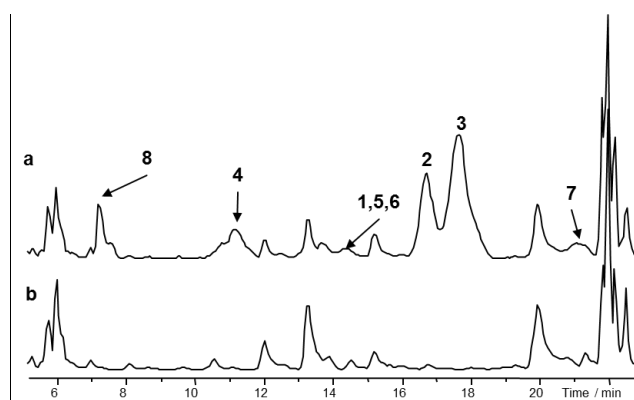
Figure S14. HPLC/MS analysis (base peak chromatogram, 300-800 m/z) of DSM17909 wild type (a) and knockout mutant *xabB*::cat (b). All xenoamicines are lost in the mutant.

a   Dual C/E Domains

```
                   IxxxxxxxPGGxxNVQDIYPLAPLQEGILxHHxxxxxGDxY
    XabB C/E 3     EIDAIAEATPDGVRNIQDIYPLAPLQEGILYHHLRQAQGDTYLLQSLLAFES
    XabB C/E 5     EIDTIVAHIPDGAGNVQDIYPLAPLQSGILFHHLLQTQGDDYLLHIMLAFAT
    XabB C/E 7     EIDAIVATVPGGAANVQDIYPLSPLQEGILFHHLLQTQGDNYLLQSLLAFDH
    XabC C/E 9     EIDAIIATLPGGASNIQDIYPLAPLQAGILFHYLMQAQGDTYLLQSLLDFDT
    XabD C/E 11    EIDAIIETIPSGVSNVQDIYPLAPLQEGIFFFQHLLQAQGDDYLLQSLFAFDS
```

b   C Domains

      core C1   core C2   core C3

```
                SxAQxRLWxL              RHExLRTxF            xHHxITDGWS
    XabA C 1    FSLSSTQQVVW-LDQY        VVYRHDALRLRLIST      GDAITLGLIAEEIANT
    XabB C 2    ----------------        -VGPPWSPGFPFFLG      LLTLHHIITDGWSLGV
    XabB C 4    LPLSFAQQRLWFLSQI        LIARHEILRTRFVLV      LLTQHHIISDGWSIGV
    XabB C 6    LPLSFAQQRLWFLGQL        LVARQEILRTRFVLV      LITQHHIISDGWSIGV
    XabC C 8    LPLSFAQQRLWFLSQI        LIARHESLRTRFVLV      LLTLHHIITDGWSTGV
    XabD C 10   VPLSFAQQRLWFLNQL        LIARHESLRTRFRLI      LLTQHHIISDGWSIGI
    XabD C 12   LPLSFAQQRLWFLAQL        LVSRHESLRTRFVSI      LLTLHHIITDGWSLGV
    XabD C 13   LPLSFAQQRLWFLGQL        LVARHESLRTRFVSV      LLTLHHIITDGWSLGV
```

      core C4   core C5   core C6

```
                YxDYAVW                 IGFxVNTLxxR          HQDxPFE
    XabA C 1    DHAYLHSQRYQQD           KRTIGMFSSVIPVGITV    CYKRQRLPIAEIN
    XabB C 2    PIQYADYAVWQHE           EGLIGFFVNTLALRMTC    AYAHQDLPFEQVV
    XabB C 4    PIQYADYAVWQRE           EPLIGFFVNTLALRVTL    AYAHQDLPFEQVV
    XabB C 6    PIQYADYAVWQRD           EGMLGFFVNTLALRVTL    AYTHQDLPFEQVV
    XabC C 8    PIQYADYAVWQHE           EGLMGFFVNTLALRVTF    AYAHQDLPFEQVV
    XabD C 10   PVQYADYAVWQRE           EGLIGFVVNTLALRVTF    AYAHQDLPFEQVV
    XabD C 12   PIQYADYAVWQH            EGLIGFFTNTLALRVRC    AYAHQDLPFEQVV
    XabD C 13   PIQYADYAVWQHE           EGLIGFFVNTLALRMTC    AYAHQDLPFEQVV
```

Figure S15. ClustalW-alignment of dual C/E- (a) and C- (b) domains from the gene cluster of xenoamicines highlighting consensus sequences. [2] [3]
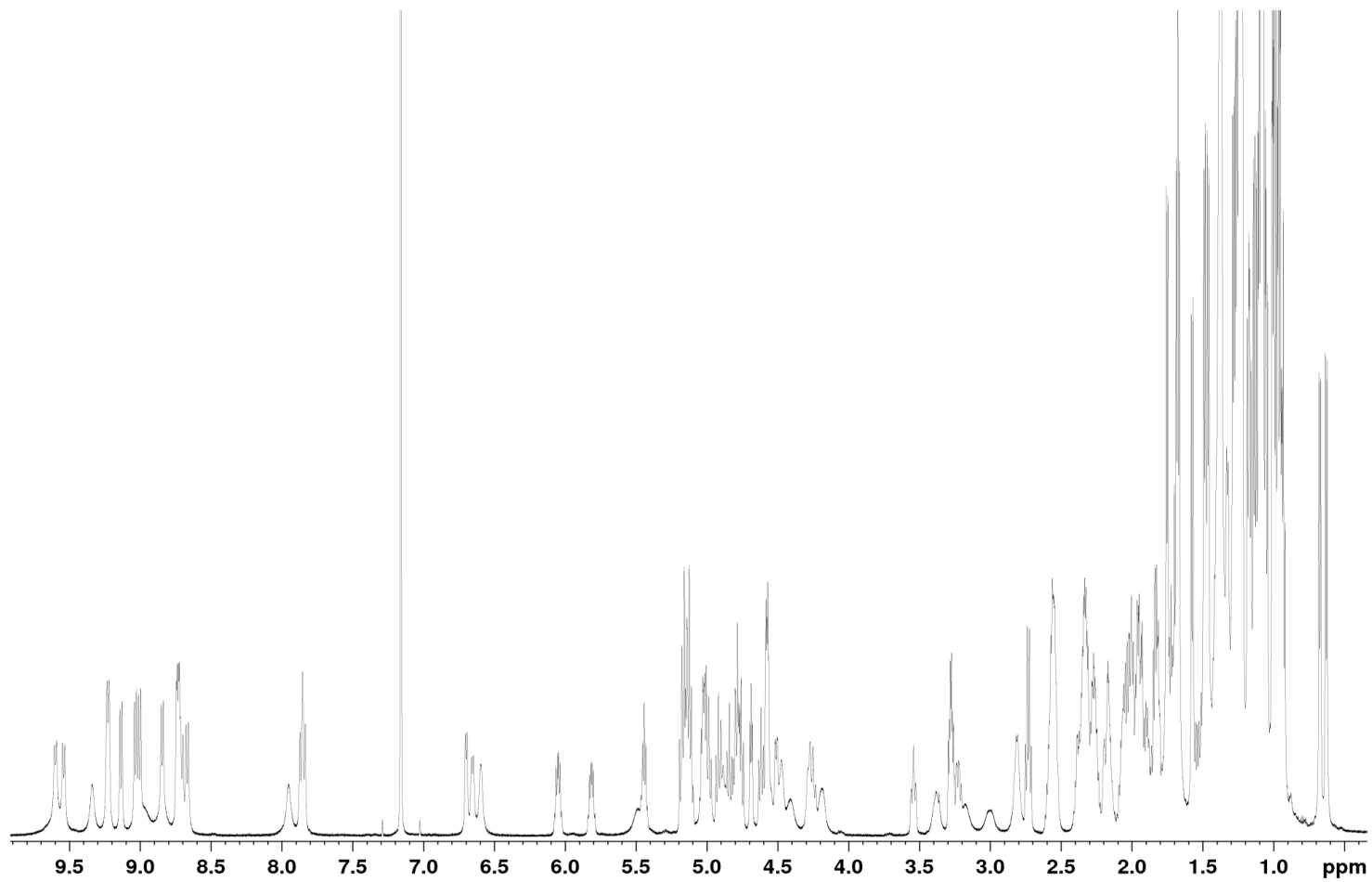
Annex 1. ^{1}H NMR spectra of **1**, d$_6$-benzene, 293K, mixture of conformers, 600MH

20
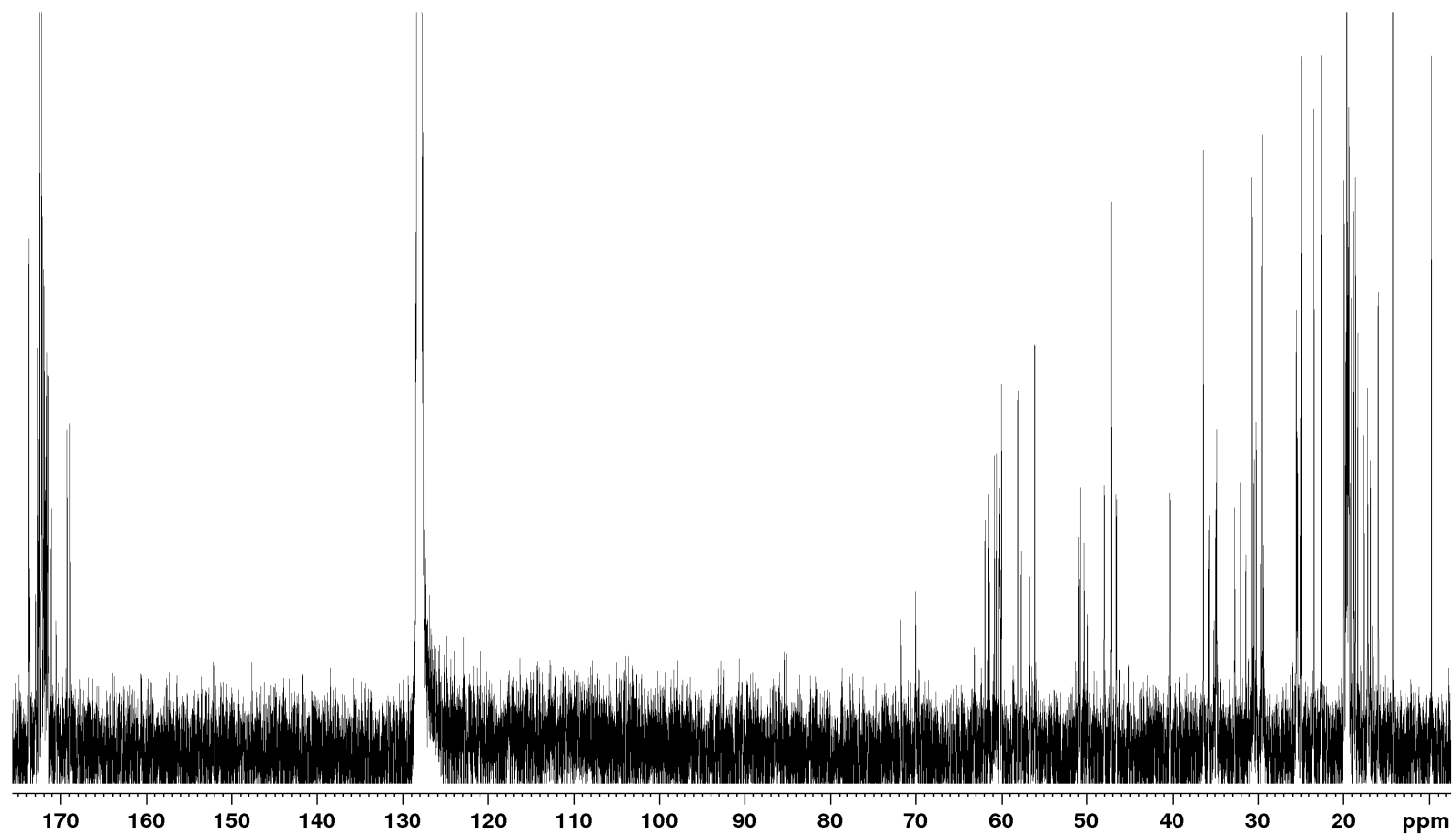
Annex 2. ^{13}C NMR spectra of **1**, d$_6$-benzene, 293K, 600MHz, mixture of conformers.

21
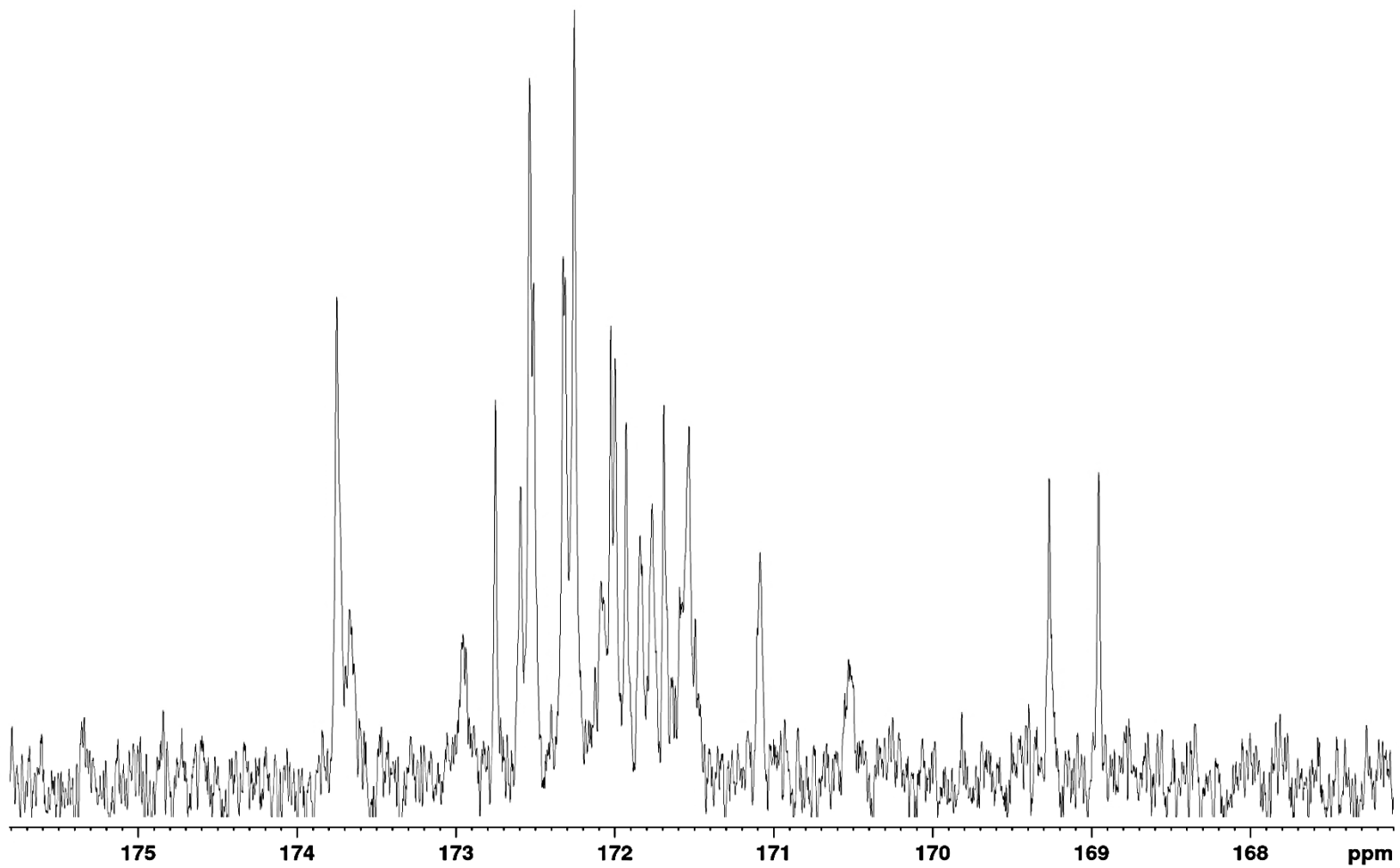
Annex 3. ^{13}C NMR spectra of **1**, d$_6$-benzene, 293K, 600MHz, mixture of conformers, zoom of the carbonyl region.
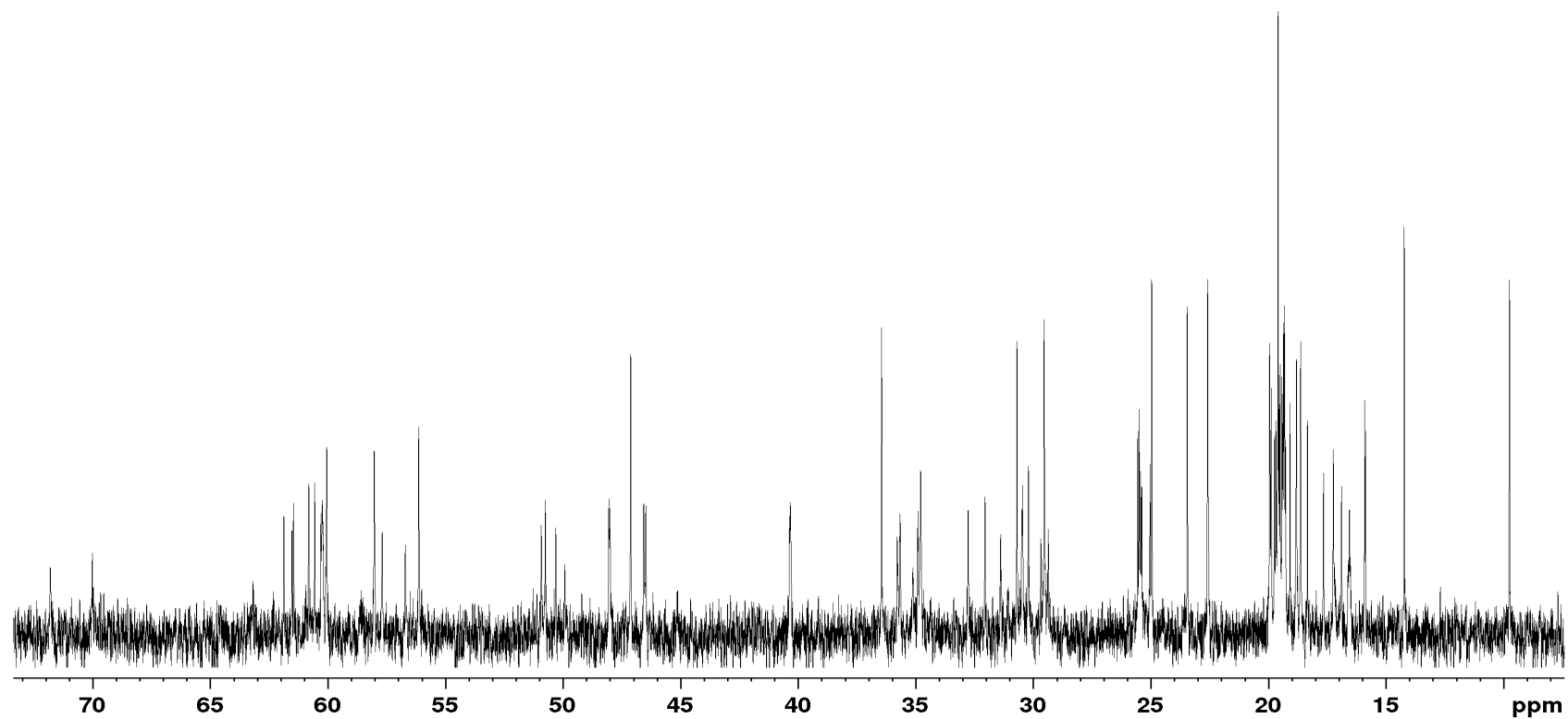
Annex 4. ^{13}C NMR spectra of **1**, d$_6$-benzene, 293K, 600MHz, mixture of conformers, zoom of the aliphatic region.

23
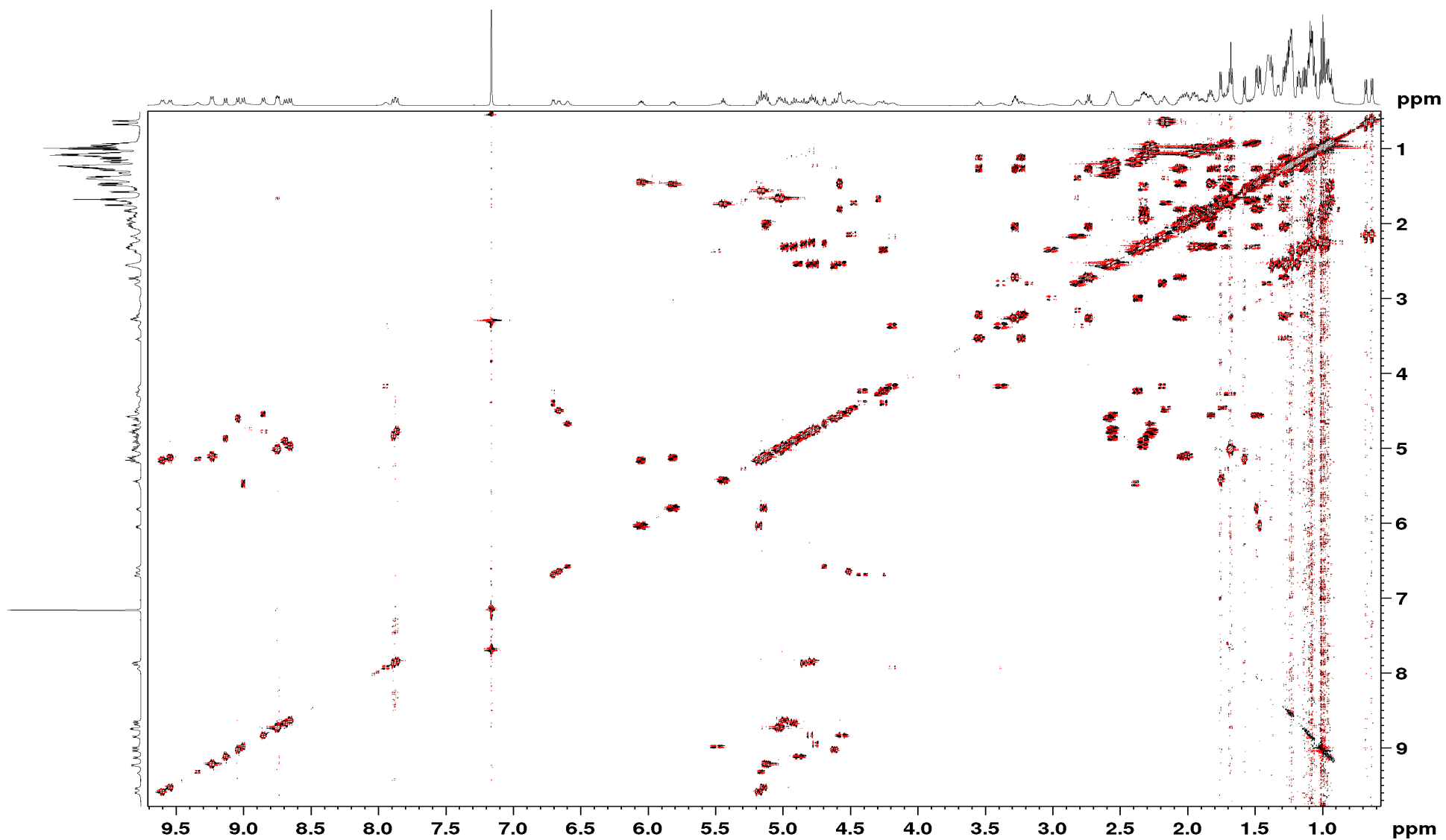
Annex 5. ^{1}H-^{1}H COSY NMR spectra of **1**, d$_6$-benzene, 293K, 600MHz, mixture of conformers.
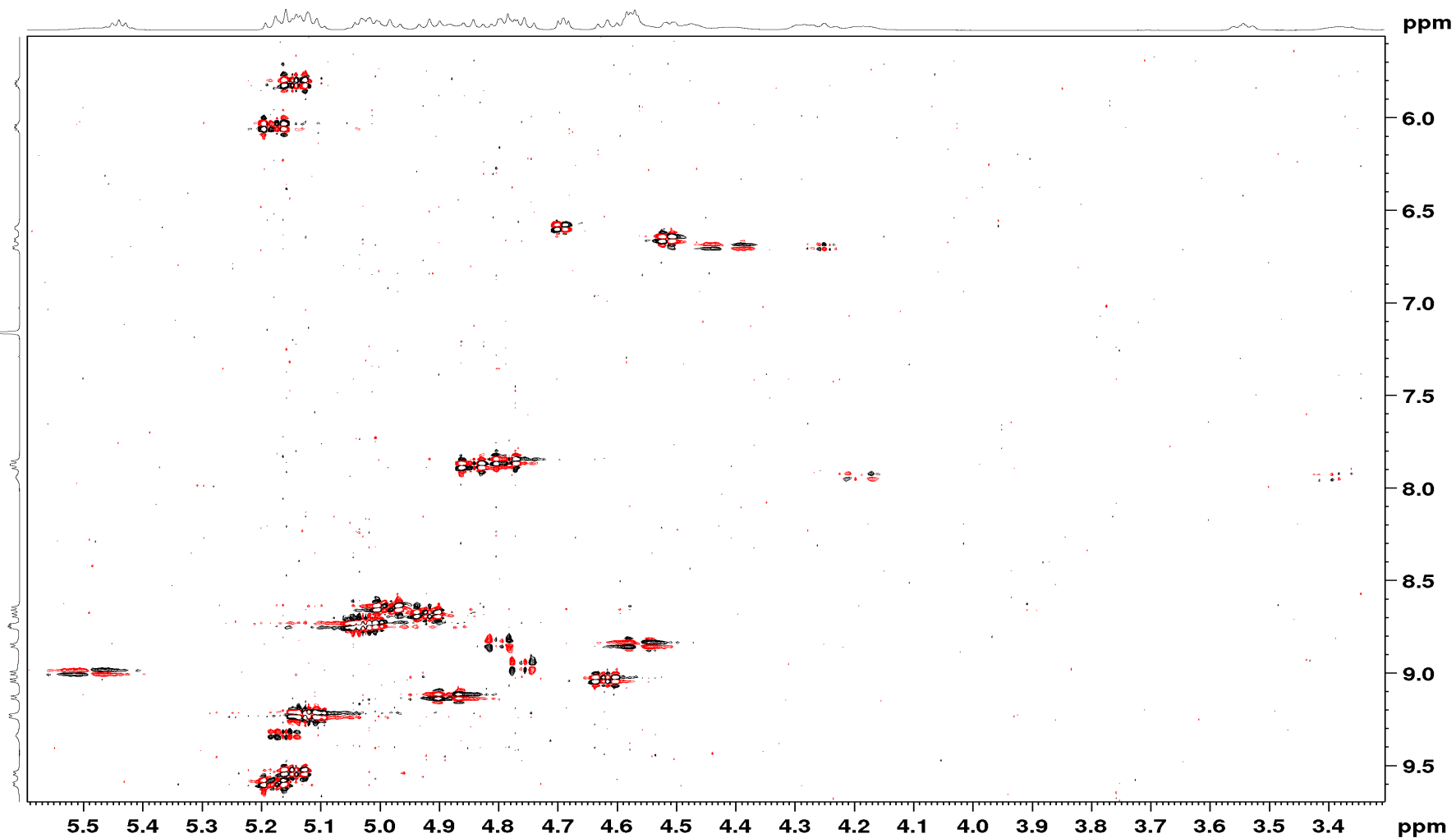
Annex 6. ^{1}H-^{1}H COSY NMR spectra of **1**, d$_6$-benzene, 293K, 600MHz, mixture of conformers, zoom 1.

25
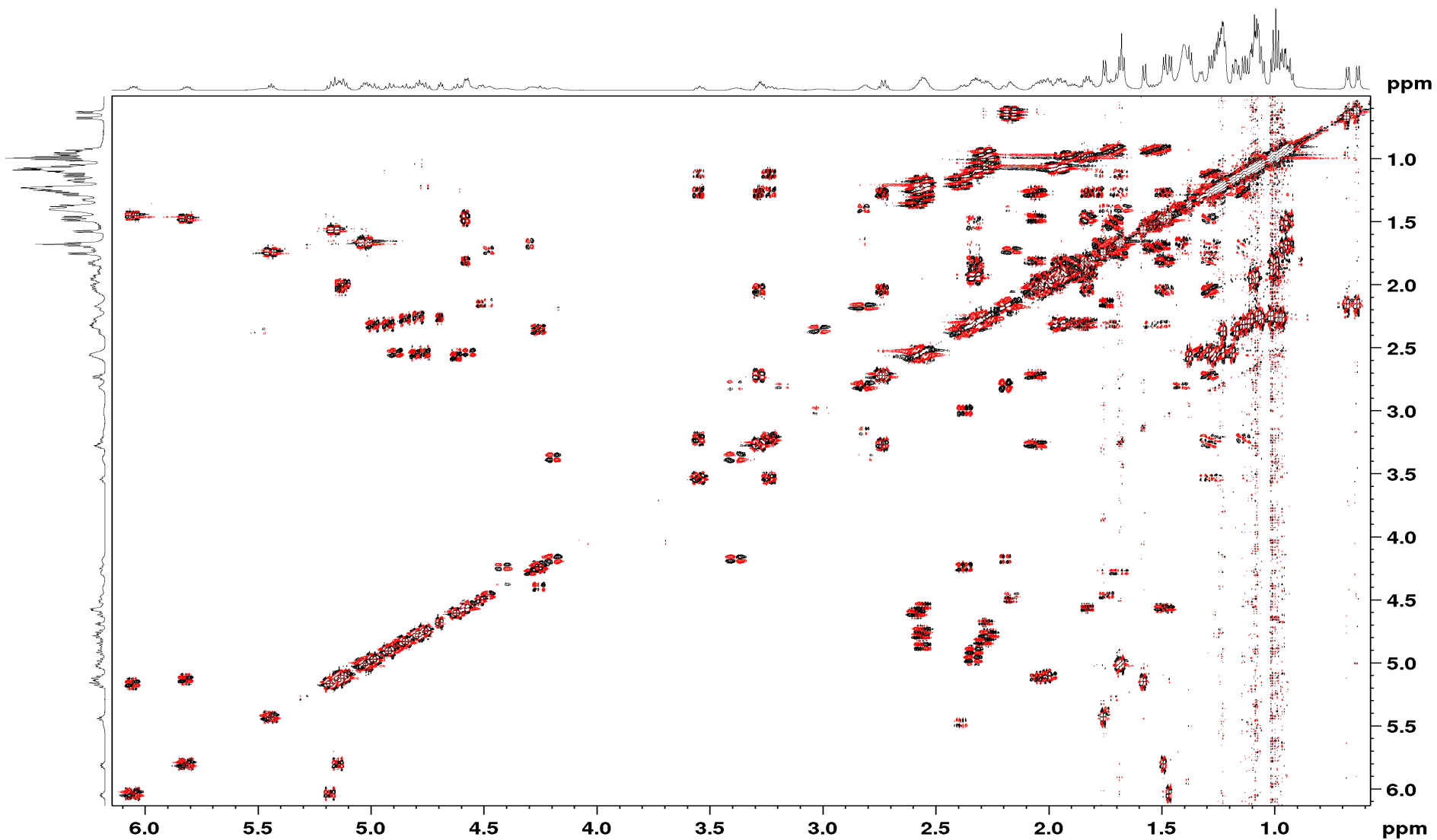
Annex 7. ^{1}H-^{1}H COSY NMR spectra of **1**, d$_{6}$-benzene, 293K, 600MHz, mixture of conformers, zoom 2.
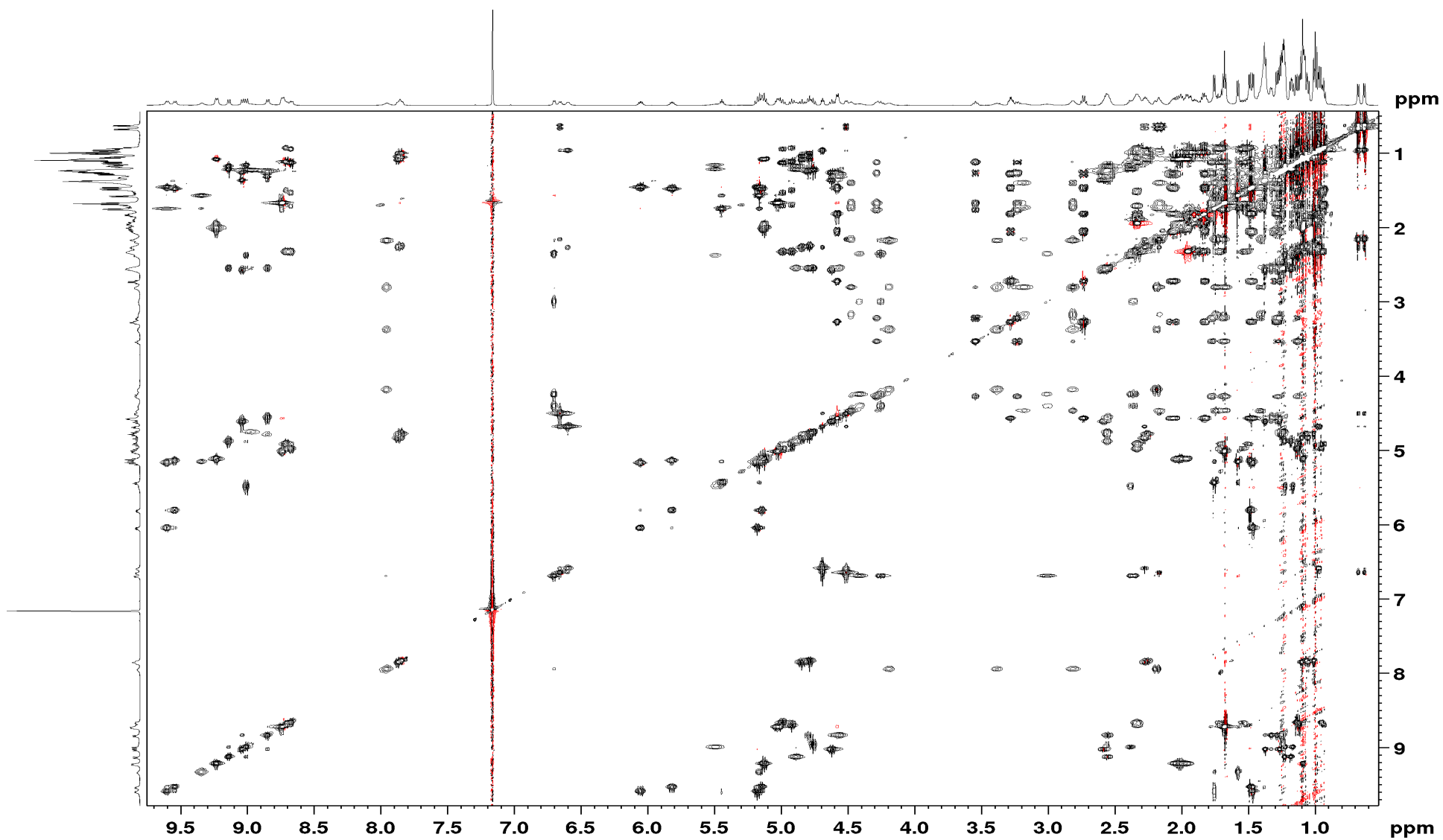
Annex 8. ^{1}H-^{1}H TOCSY NMR spectra of **1**, d$_6$-benzene, 293K, 600MHz, mixture of conformers.

27
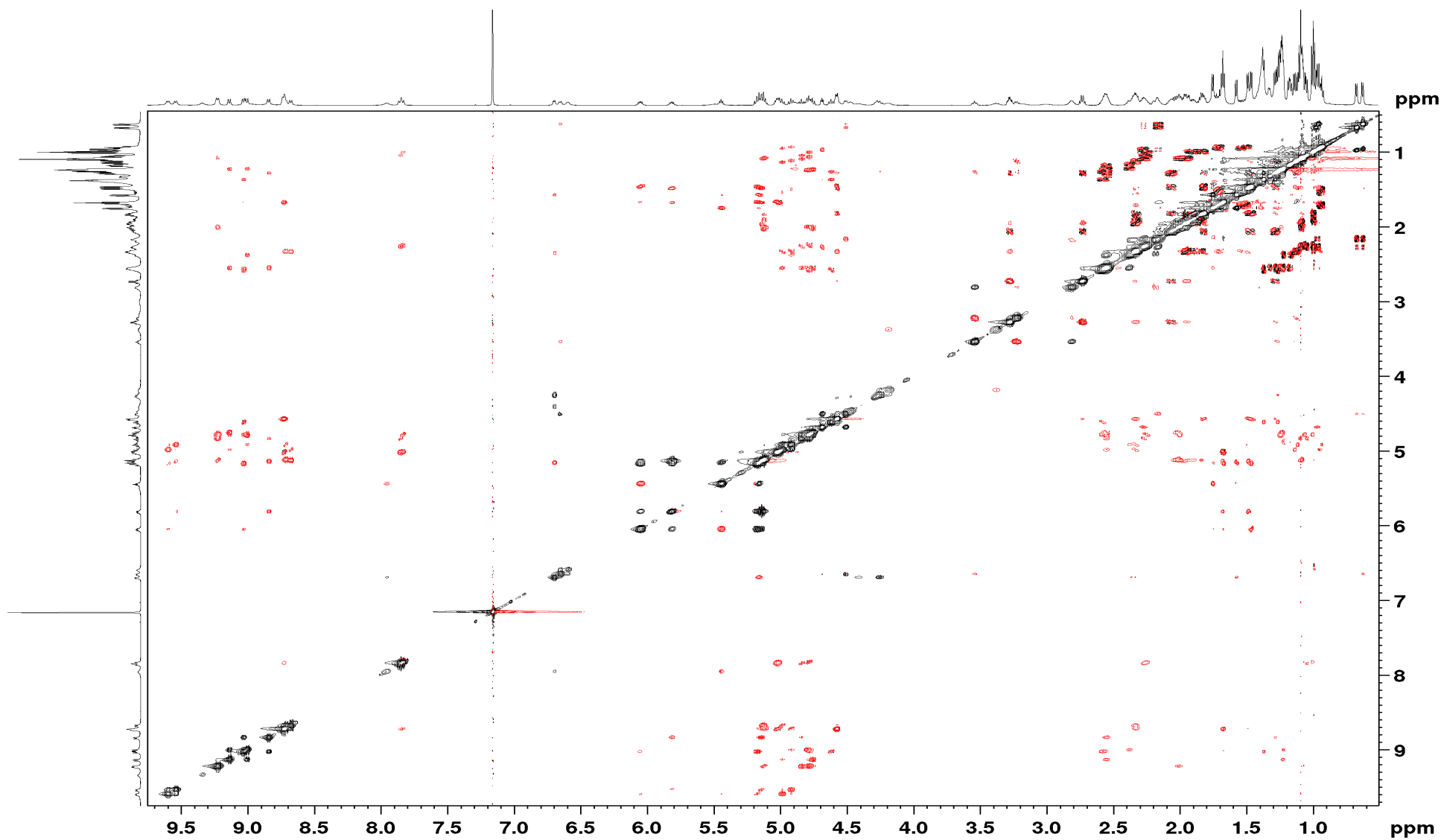
Annex 9. ^{1}H-^{1}H ROESY NMR spectra of **1**, d$_6$-benzene, 293K, 600MHz, mixture of conformers.
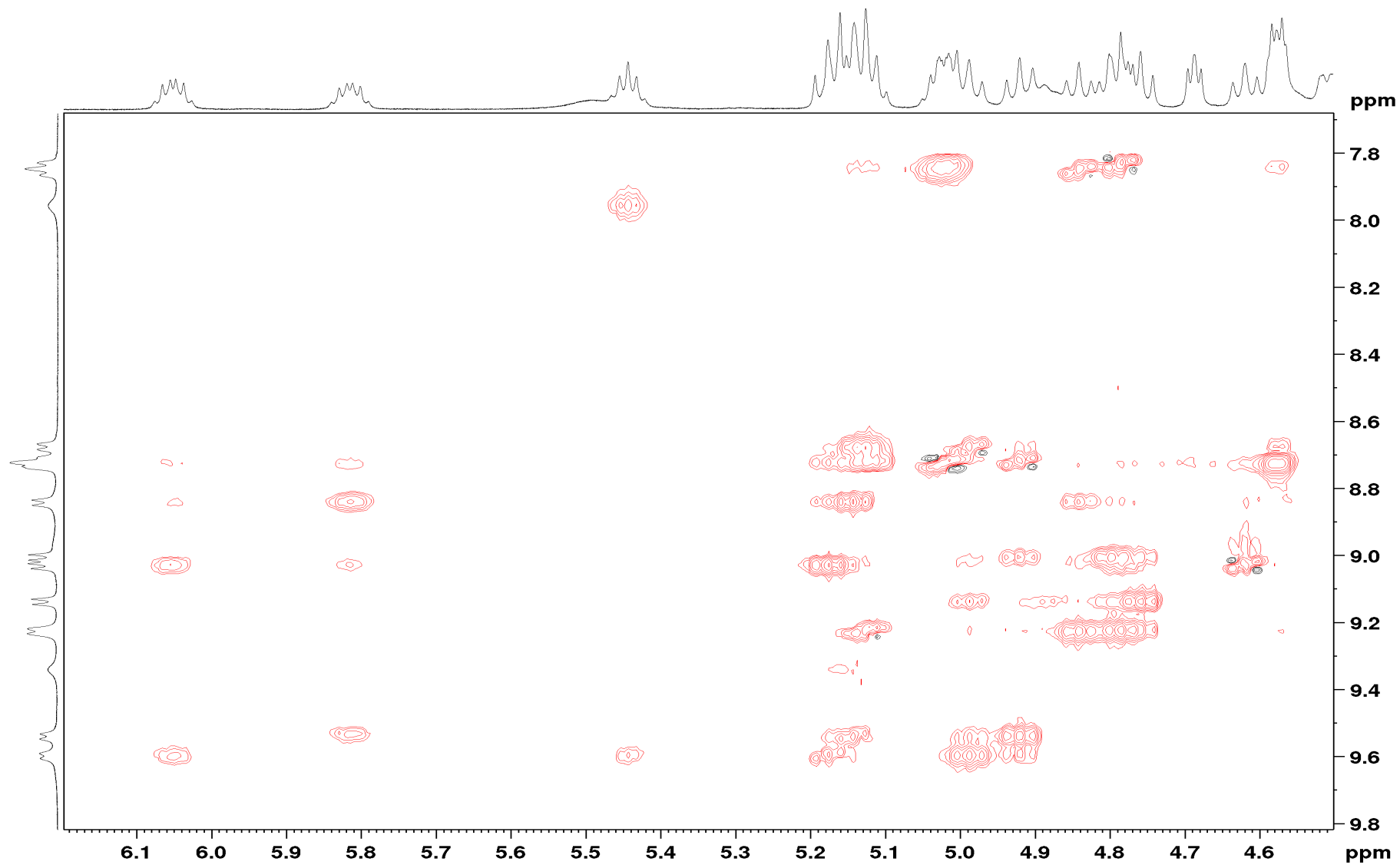
Annex 10. ^{1}H-^{1}H ROESY NMR spectra of **1**, d$_6$-benzene, 293K, 600MHz, mixture of conformers, zoom 1.
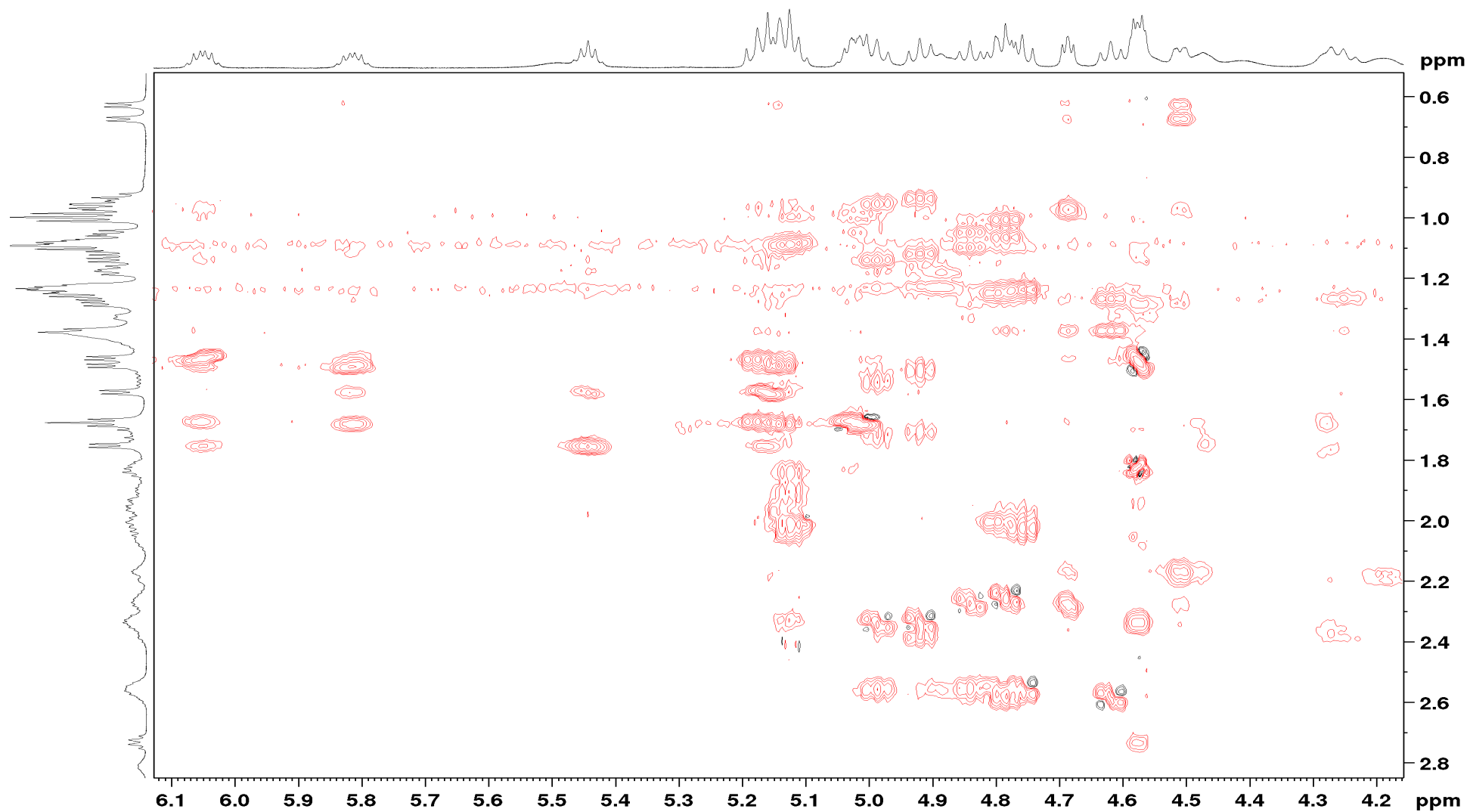
Annex 11. ^{1}H-^{1}H ROESY NMR spectra of **1**, d$_6$-benzene, 293K, 600MHz, mixture of conformers, zoom 2.
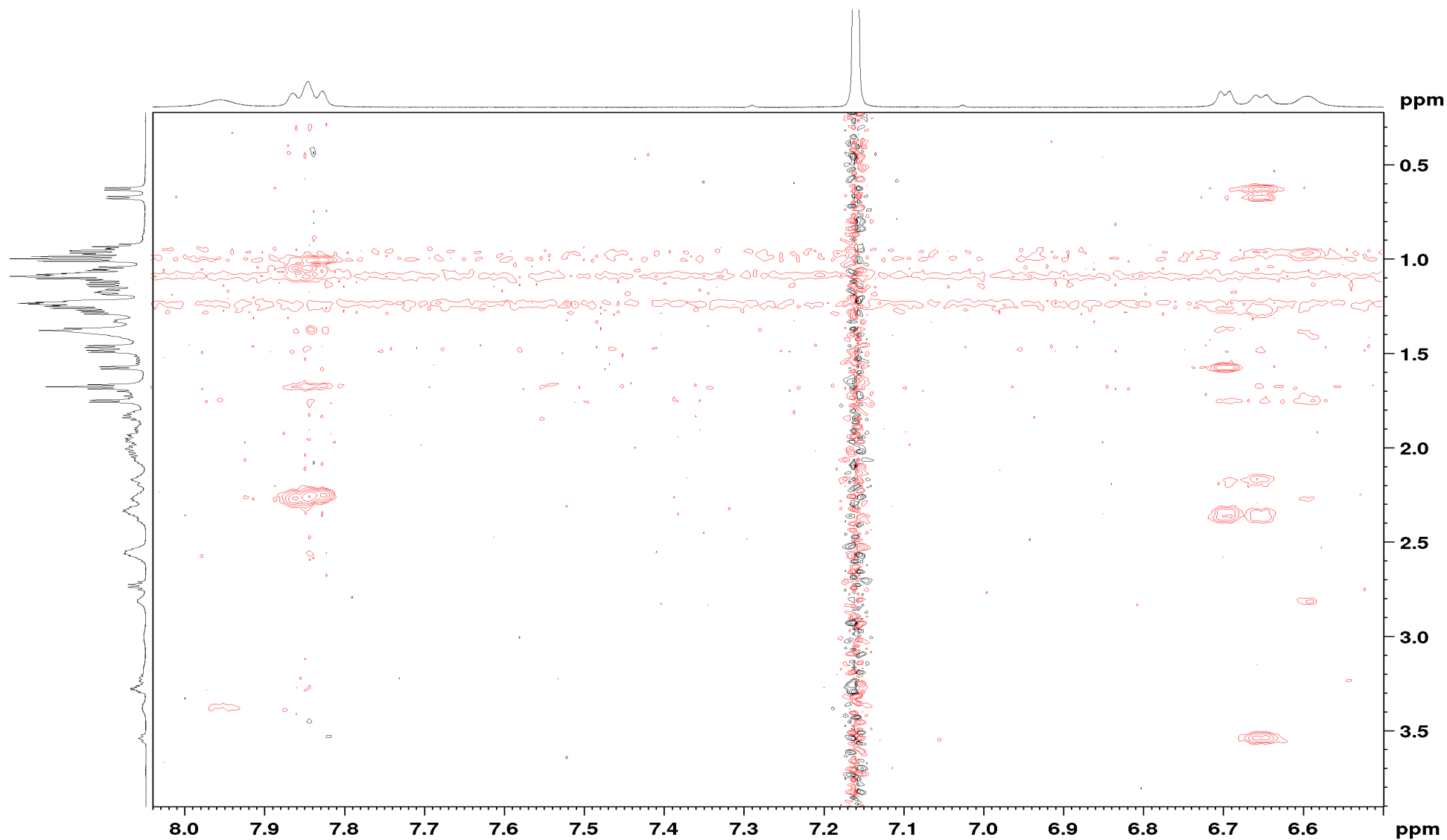
Annex 12. ^{1}H-^{1}H ROESY NMR spectra of **1**, d$_6$-benzene, 293K, 600MHz, mixture of conformers, zoom 3.

31
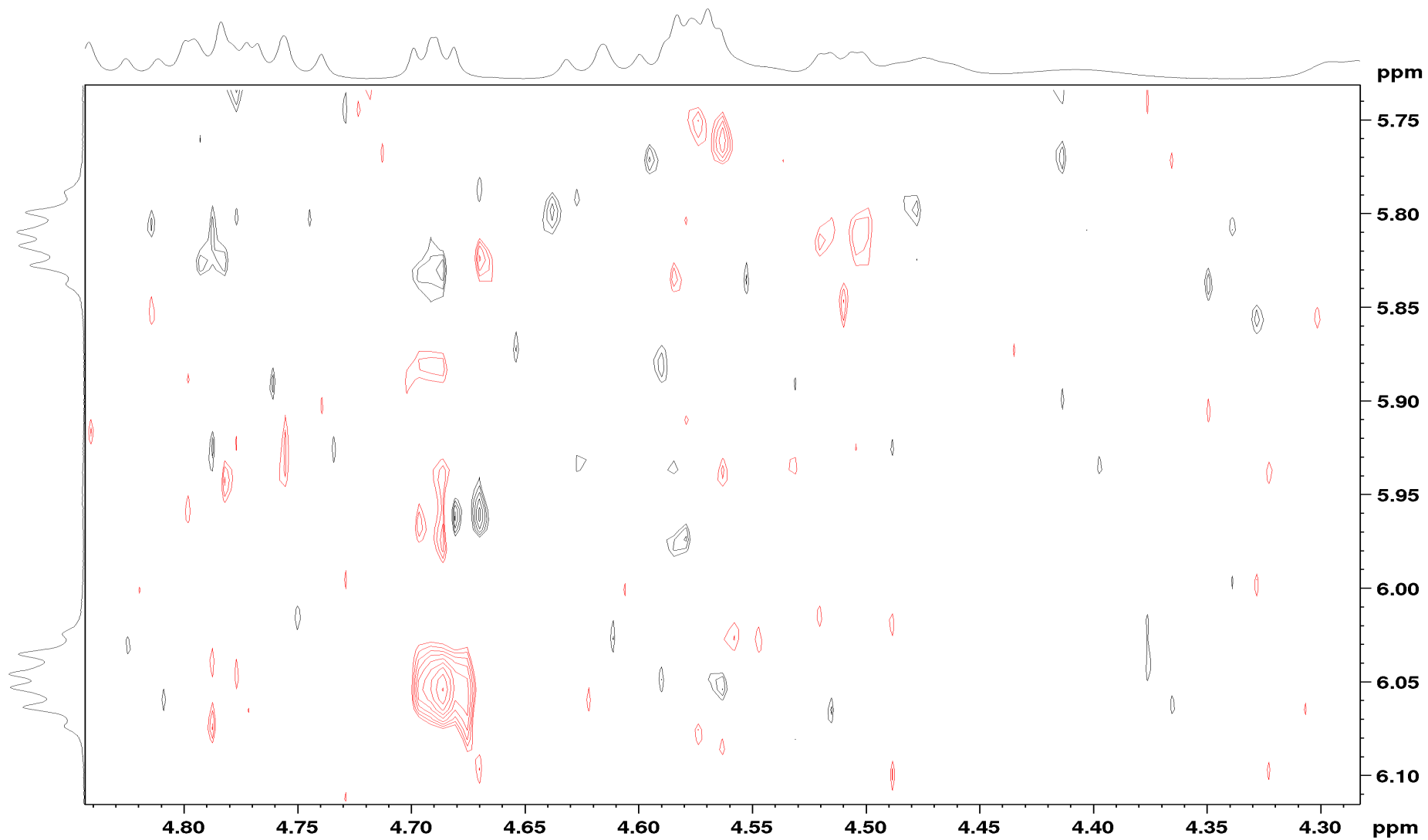
Annex 13. ^{1}H-^{1}H ROESY NMR spectra of **1**, d$_6$-benzene, 293K, 600MHz, mixture of conformers, zoom 4.
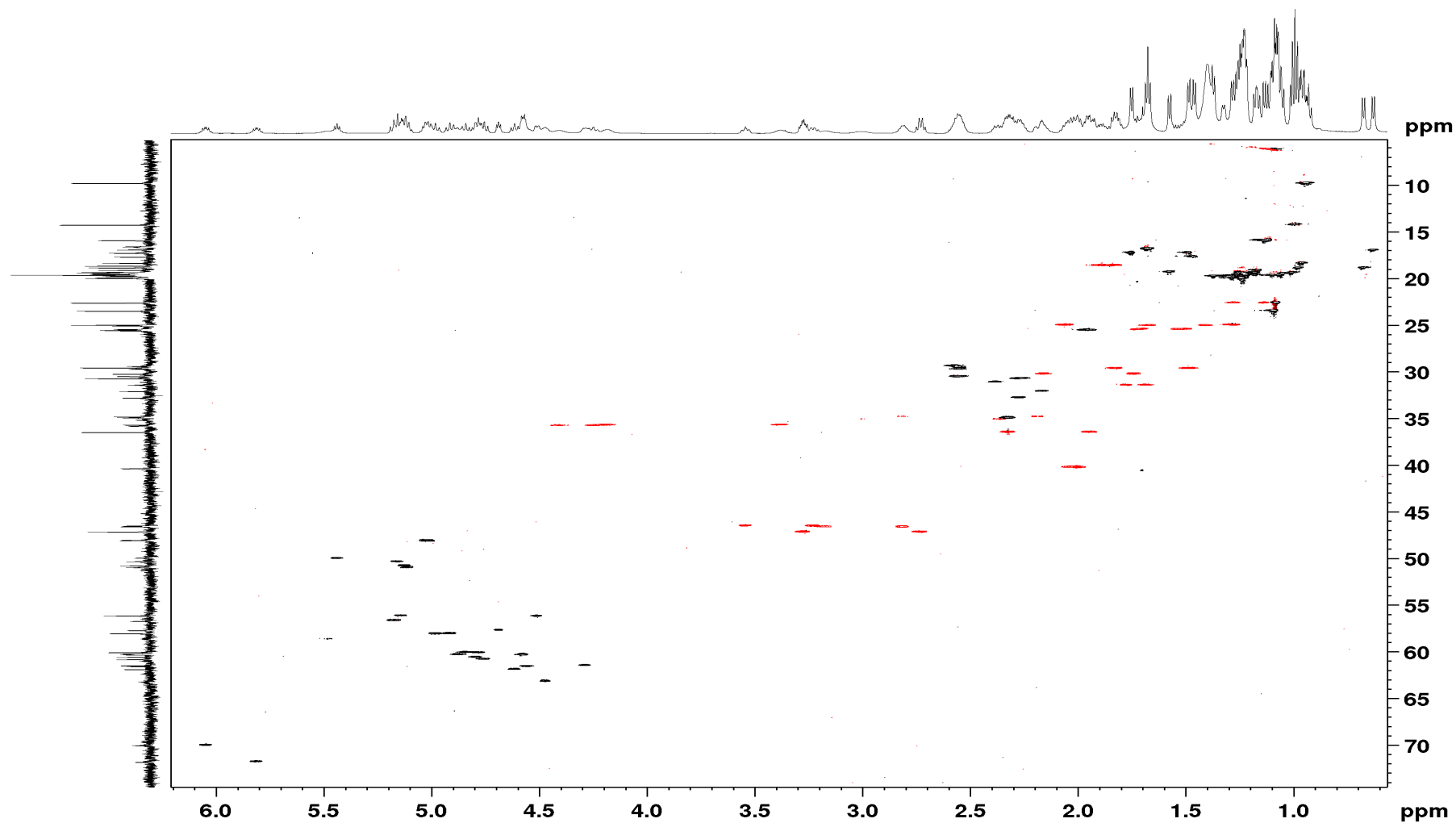
Annex 14. ^{1}H-^{13}C HSQC NMR spectra of **1**, d$_6$-benzene, 293K, 600MHz, mixture of conformers.

33
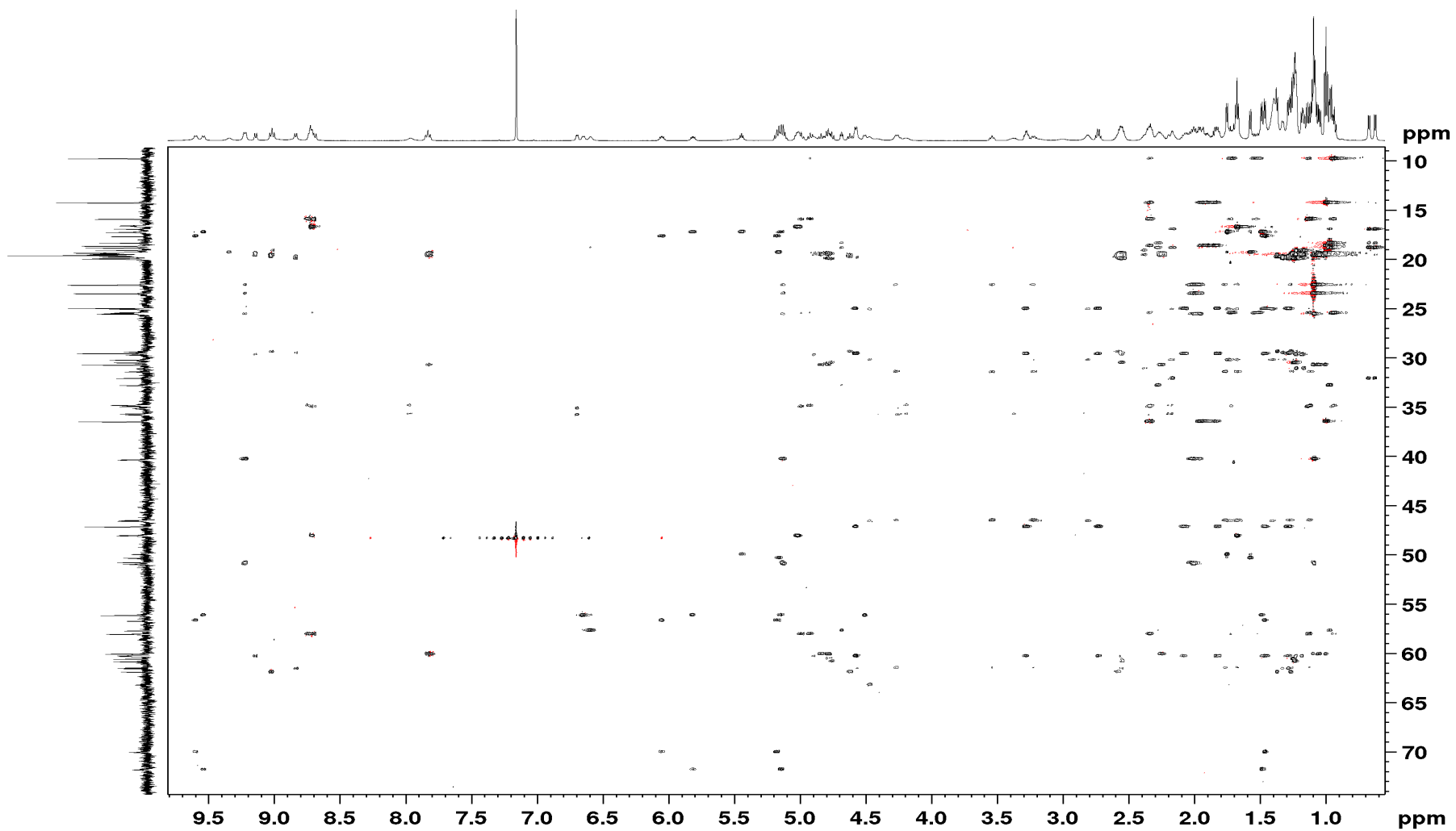
Annex 15. ^{1}H-^{13}C HSQC-TOCSY NMR spectra of **1**, d$_6$-benzene, 293K, 600MHz, mixture of conformers.
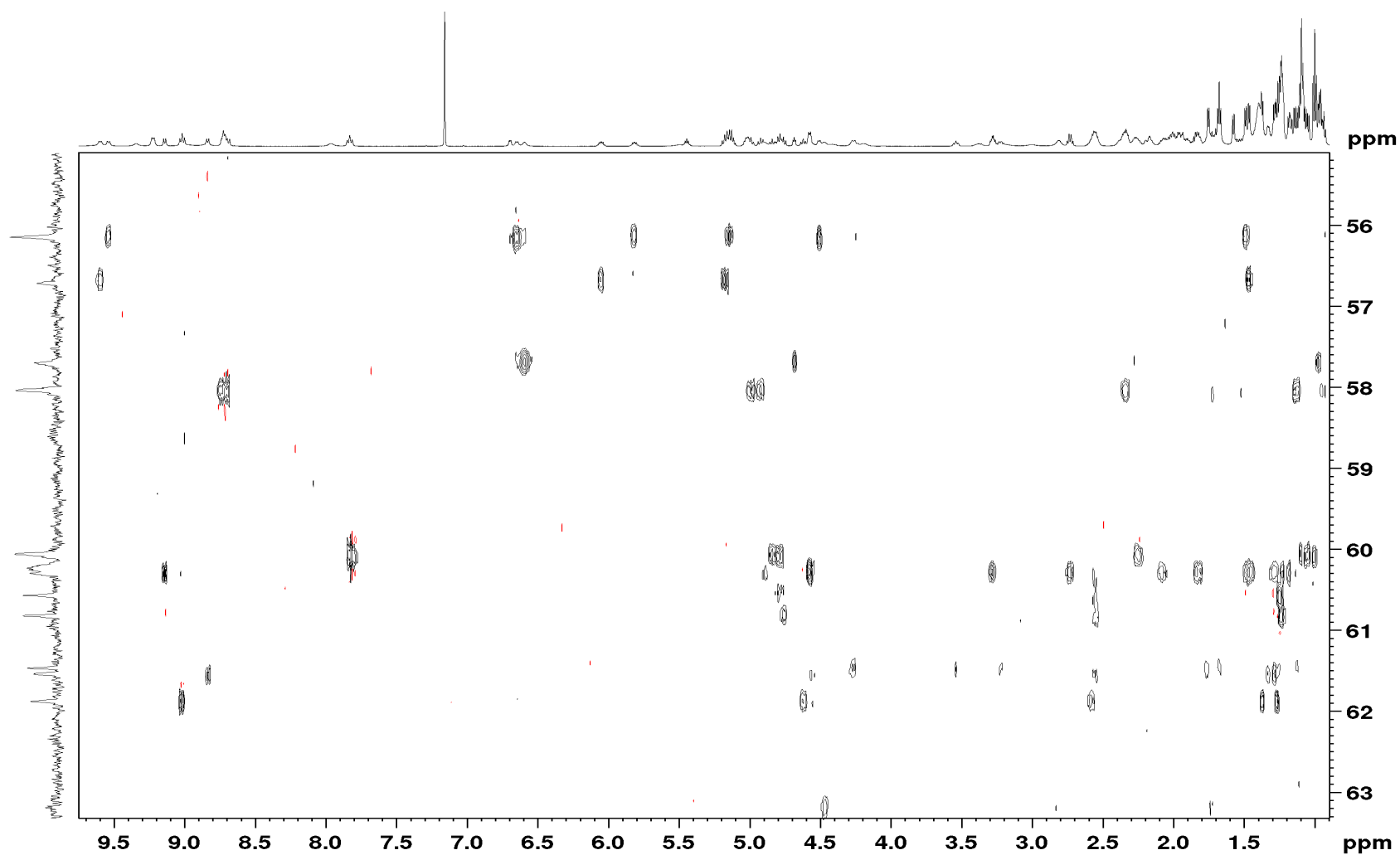
Annex 16. ^{1}H-^{13}C HSQC-TOCSY NMR spectra of **1**, d$_6$-benzene, 293K, 600MHz, mixture of conformers, zoom 1.

35
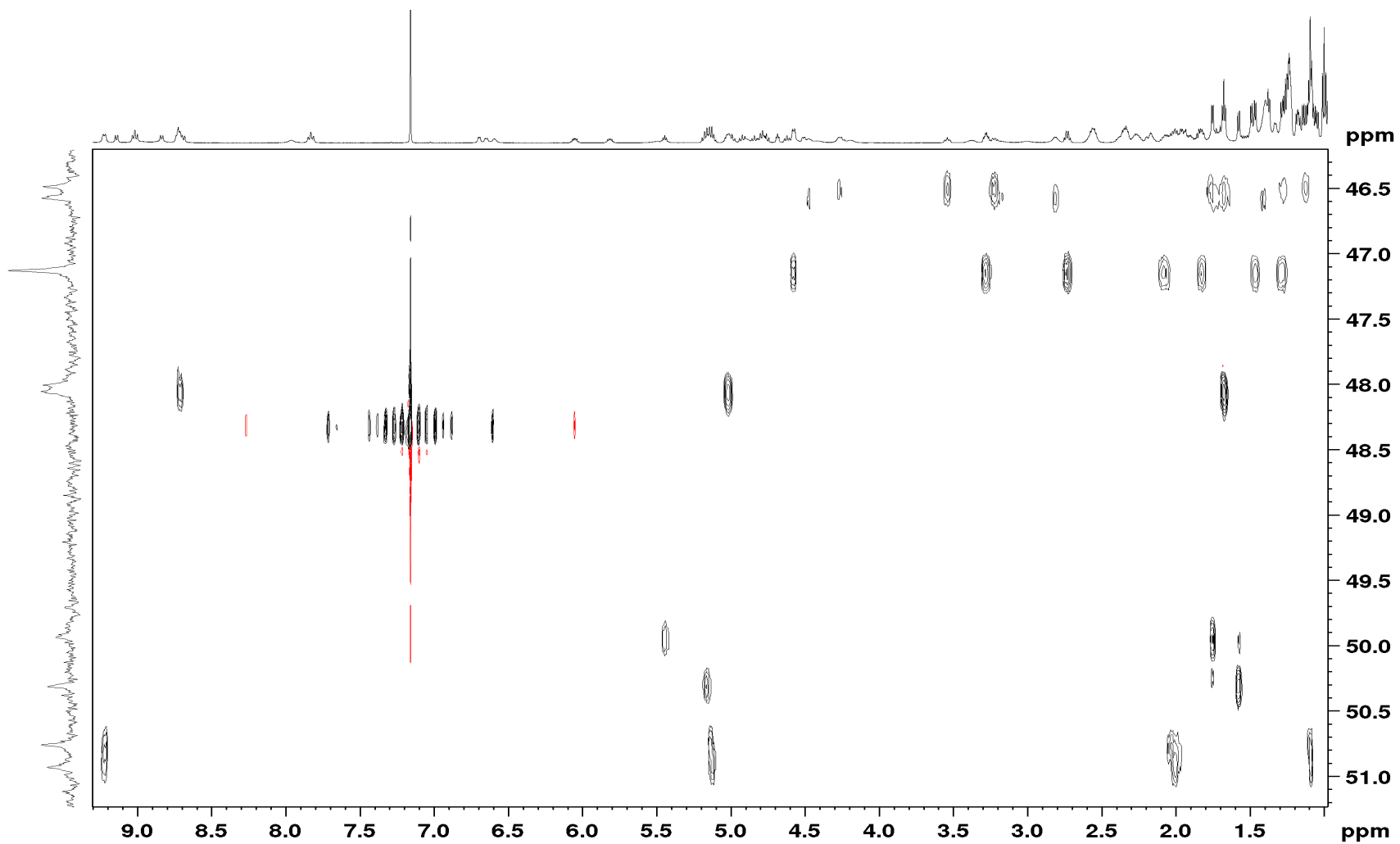
Annex 17. ^{1}H-^{13}C HSQC-TOCSY NMR spectra of **1**, d$_6$-benzene, 293K, 600MHz, mixture of conformers, zoom 2.
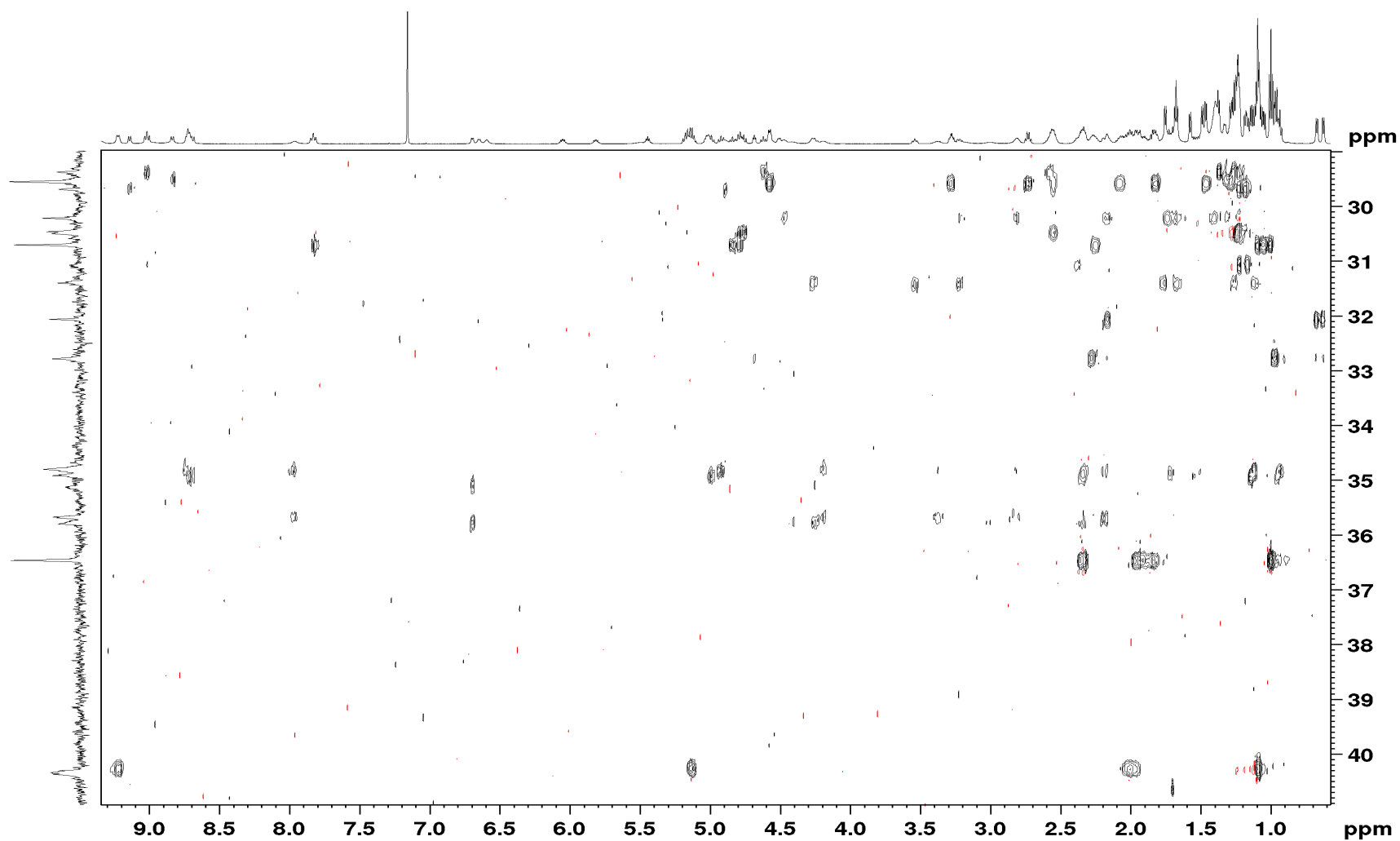
Annex 18. ^{1}H-^{13}C HSQC-TOCSY NMR spectra of **1**, d$_6$-benzene, 293K, 600MHz, mixture of conformers, zoom 3.

37
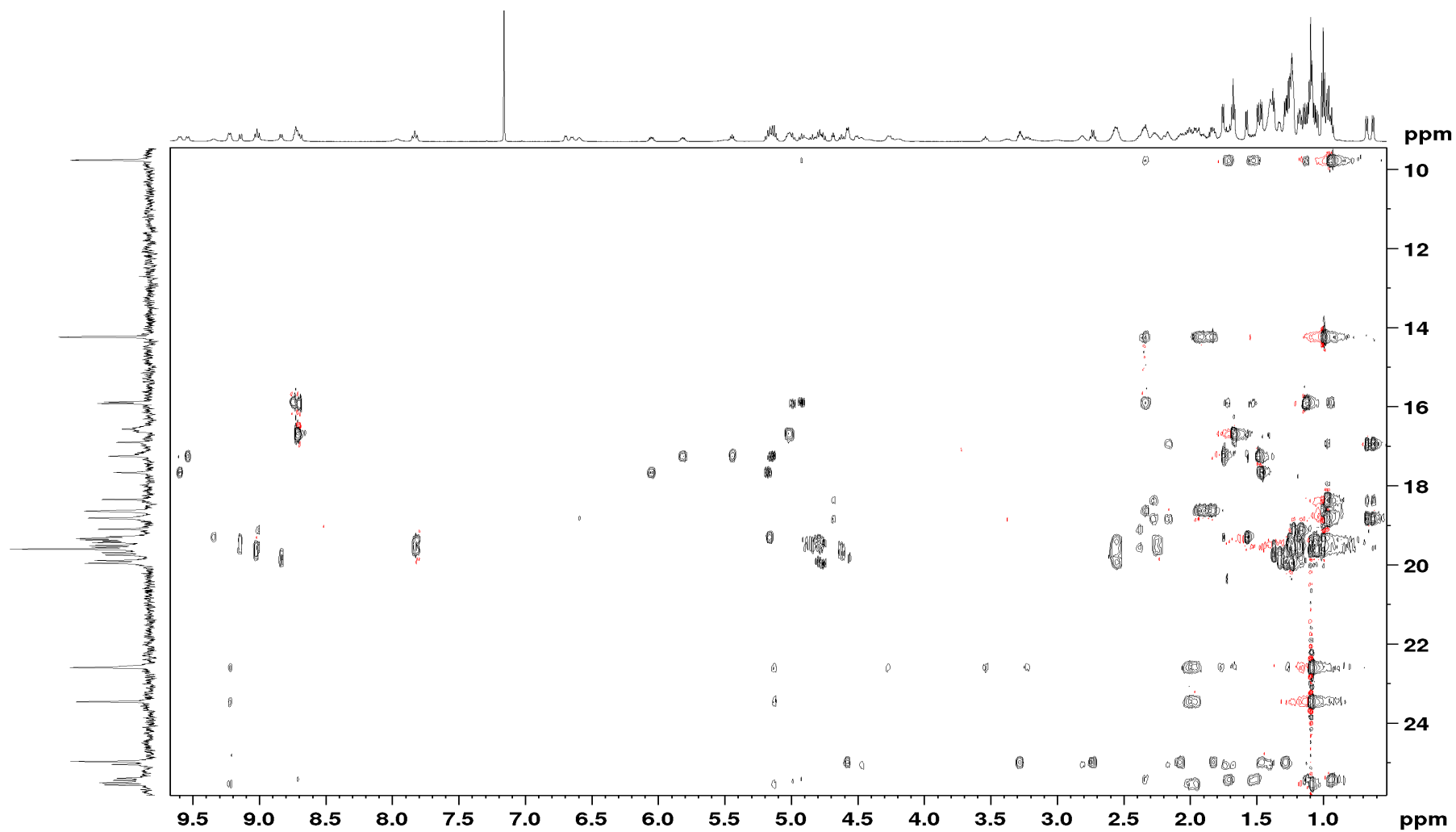
Annex 19. ^{1}H-^{13}C HSQC-TOCSY NMR spectra of **1**, d$_6$-benzene, 293K, 600MHz, mixture of conformers, zoom 4.
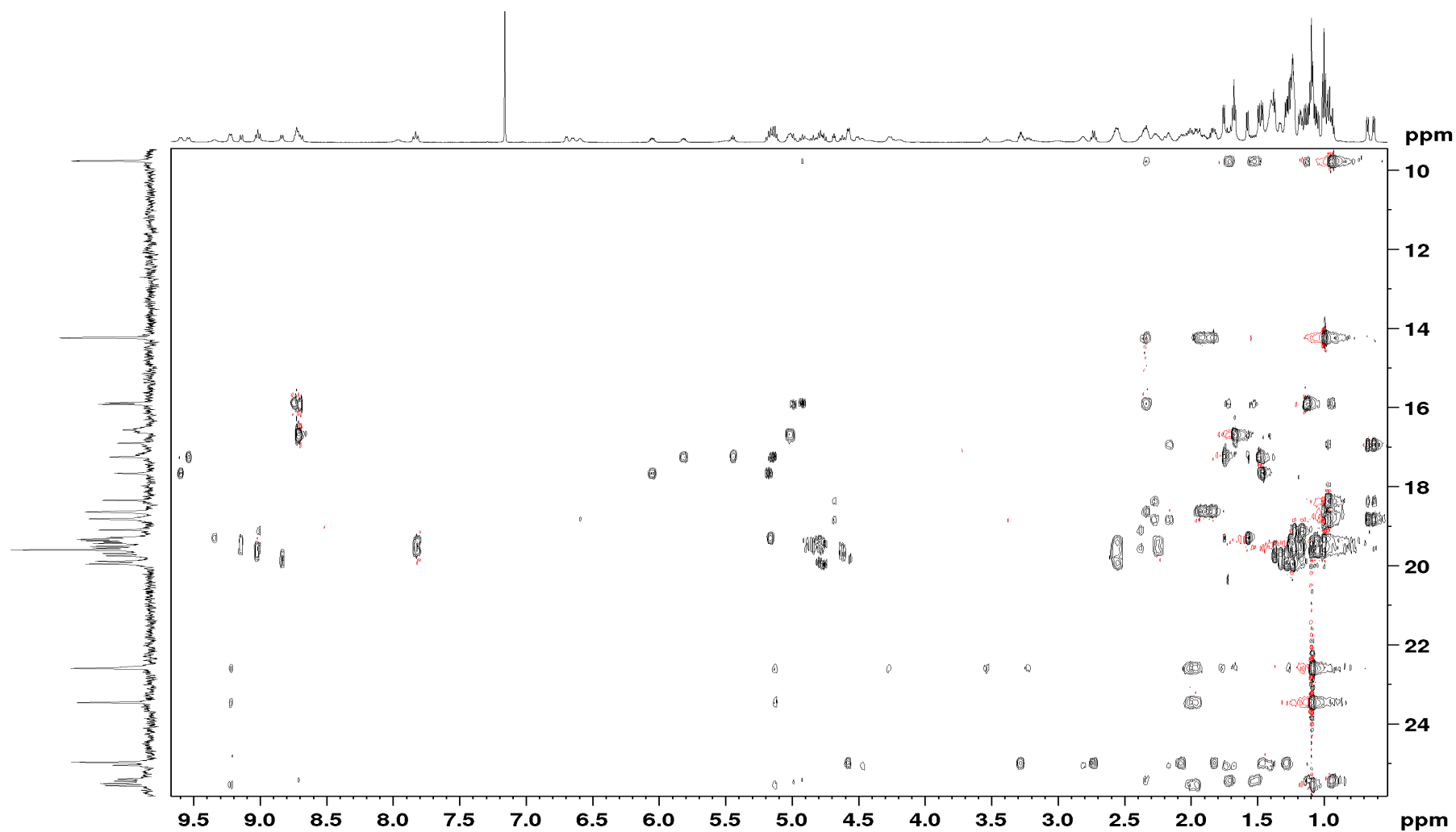
Annex 20. ^{1}H-^{13}C HSQC-TOCSY NMR spectra of **1**, d$_6$-benzene, 293K, 600MHz, mixture of conformers, zoom 5.

39

References

[1]   B. O. Bachmann, J. Ravel, *Meth. Enzymol.* **2009**, *458*, 181-217.

[2]   D. Konz, M. A. Marahiel, *Chem. Biol.* **1999**, *6*, R39-48.

[3]   C. J. Balibar, F. H. Vaillancourt, C. T. Walsh, *Chem. Biol.* **2005**, *12*, 1189-1200.

40

11.3. Unusual start and finish of anthraquinone biosynthesis in *Photorhabdus luminescens*

Authors: Qiuqin Zhou, Hélène Adihou, Darko Kresovic, Kenan A. J. Bozhüyük, Helge B. Bode*

 [*] Dipl. Chem. Q. Zhou, Dr. H. Adihou, Dipl. Bioinf. D. Kresovic, Dipl. Bioinf. K. A. J. Bozhüyük, Prof. Dr. H. B. Bode , Merck Stiftungsprofessur für Molekulare Biotechnologie Fachbereich Biowissenschaften Max-von-Laue-Str. 9 60438 Frankfurt am Main (Germany)

Prof. Dr. H. B. Bode, Buchmann Institute for Molecular Life Sciences (BMLS), Goethe Universität Frankfurt, Max-von-Laue-Str. 15, 60438 Frankfurt a. M., Germany

Submitted in: *Angew. Chem*

Status: major revision including structural data on the involved proteins required

Attachments: Declaration on the contribution of the author and the manuscript.

Erklärung zu den Autorenanteilen

an der Publikation: "Unusual start and finish of anthraquinone biosynthesis in *Photorhabdus luminescens"*

Status: in revision

Beteiligte Autoren: Qiuqin Zhou (QZ), Hélène Adihou (HA), Darko Kresovic (DK), Kenan A. J. Bozhüyük (KAJB), Helge B. Bode (HBB)

Was hat der Promovierende bzw. was haben die Koautoren beigetragen?

(1) zu Entwicklung und Planung

Promovierender: 35 %

Co-Autoren HA, DK und KAJB: 10 %

Co-Autor HBB: 55 %

(2) zur Durchführung der einzelnen Untersuchungen und Experimente

Promovierender: 75 %, alle Experimente außer den Experimenten von Co-Autoren HA, DK und KAJB (siehe unten)

Co-Autor HA: 10 %, Synthese von Modellsubstanzen, in vitro-Experimente mit AntI

Co-Autoren DK und KAJB: 15 %, Homologiemodelle, Phylogenetischer Baum und Alignment

(3) zur Erstellung der Datensammlung und Abbildungen

Promovierender: 75 %, alle Daten außer den Daten von Co-Autoren HA, DK und KAJB (siehe unten)

Co-Autor HA: 10 %, Synthese von Modellsubstanzen, in vitro-Experimente mit AntI

Co-Autoren DK und KAJB: 15 %, Homologiemodelle, Phylogenetischer Baum und Alignment

(4) zur Analyse und Interpretation der Daten

Promovierender: 75 %, alle Analyse außer den Experimenten von Co-Autoren HA, DK und KAJB (siehe unten)

Co-Autor HA: 10 %, Synthese von Modellsubstanzen, in vitro-Experimente mit AntI

Co-Autoren DK und KAJB: 15 %, Homologiemodelle, Phylogenetischer Baum und Alignment

(5) zum Verfassen des Manuskripts

Promovierender: 40 %

Co-Autoren HA, DK und BKAJ: 10 %

Co-Autor HBB: 50 %

Unterschrift Promovend: _____ Datum/Ort: _____

Zustimmende Bestätigungen der oben genannten Angaben

Unterschrift Betreuer: _____ Datum/Ort: _____

Unusual start and finish of anthraquinone biosynthesis in *Photorhabdus luminescens***

Qiuqin Zhou, Hélène Adihou, Darko Kresovic, Kenan A. J. Bozhüyük, Helge B. Bode*

[*]   Dipl. Chem. Q. Zhou, Dr. H. Adihou, Dipl. Bioinf. D. Kresovic, Dipl. Bioinf. K. A. J. Bozhüyük, Prof. Dr. H. B. Bode
        Merck Stiftungsprofessur für Molekulare Biotechnologie
        Fachbereich Biowissenschaften
        Max-von-Laue-Str. 9
        60438 Frankfurt am Main (Germany)
        E-mail: h.bode@bio.uni-frankfurt.de
        Homepage: http://www.uni-frankfurt.de/fb/fb15/institute/inst-3-mol-biowiss/AK-Bode

        Prof. Dr. H. B. Bode, Buchmann Institute for Molecular Life Sciences (BMLS), Goethe Universität Frankfurt, Max-von-Laue-Str. 15, 60438 Frankfurt a. M., Germany

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

Abstract

Polycyclic aromatic compounds derived from type II polyketide synthases (PKS) are typical natural products from Gram-positive *Streptomyces*, where their biosynthesis has been studied in detail. The only known example for a classical type II PKS biosynthesis in Gram-negative bacteria was found in *Photorhabdus luminescens* involved in the anthraquinone biosynthesis. A detailed analysis of this biosynthesis using heterologous expression of the *ant* gene cluster, analysis of various mutants and *in vitro* reconstitution of the biosynthesis has revealed that the putative CoA ligase AntG is required for initiation of the polyketide biosynthesis and the hydrolase AntI is responsible for the terminal polyketide shortening from an initial octaketide to the final heptaketide anthraquinone.

Polyketides are one major class of natural products with several of them have become important therapeutics. Their structural diversity is derived from only a few building blocks that are assembled by different types of polyketide synthases (PKS):[1,2] The modular type I PKS similar to type I fatty acid synthases (FAS) represent multifunctional giant enzymes responsible for the production of compounds like macrolides including erythromycin or candicin both clinically used.[3] Type II PKS are related to type II FAS and are constitute of single proteins that act in an iterative manner resulting in the production of polycyclic aromatic compounds like tetracycline or doxorubicin.[4,5] Type III PKS are single enzymes acting iterative to form stilbenes, chalcones or pyrones.[6]

Whereas subclasses of type I and type III PKS are known from bacteria, fungi and plants (for type III PKS),[1] type II PKS are well-known from Gram-positive *Streptomyces* and other actinomycetes.[4] We have previously identified a biosynthesis gene cluster encoding a type II PKS in the entomopathogenic bacterium *Photorhabdus luminescens*, which is responsible for the biosynthesis of anthraquinones (AQ)[7] and besides aurachin from *Stigmatella aurantiaca*[8] is only the second example of a type II PKS derived product from Gram-negative bacteria. In contrast to aurachin it can regarded as a "typical" PKS product since AQ derivatives are also known from other bacteria having type II PKS systems.

Here we describe the heterologous expression of the AQ biosynthesis gene cluster in *E. coli* and the requirements for the initial biosynthesis steps requiring a CoA ligase

for activation of the ACP and the terminal biosynthesis steps requiring a hydrolase involved in the shortening of the intermediary octaketide to the final AQ heptaketide.

Results and Discussion

The biosynthesis gene cluster involved in AQ biosynthesis (*antABCDEFGHI*, Fig. 1a) was heterologously expressed in *E. coli*. Whereas inactivation of *antA*, *antH* and *antC* encoding the typical type II PKS enzymes ketoreductase, aromatase and cyclase led to the expected accumulation of well-known shunt products from the ACP-bound intermediates (Fig. 1b, Fig. S1, Table S4) as identified by HPLC/MS and through comparison to standards (Fig. S2, Table S5), production of AQ-256 (**1**) required the expression of all *ant* genes (Fig. 2). Additionally, the known octaketides 3,8-dihydroxy-1-methyl-anthraquinone-2-carboxylic acid (DMAC, **9**), aleosaponarin (**10**) and utahmycin A (**12**)[9] were detected in addition to three other compounds (**11**, **14**, **15**) whose structures were predicted based on MS and labelling experiments (Fig. S3) as well as from their postulated biosynthesis (Fig. 1b). Similarly, disruption of *antI* in *P. luminescens* resulted in the loss of production of all AQ derivatives **1**-**3** but led to the production of **11**. However, in contrast to the *E. coli* experiment, **9** and **10** were not detected (Fig. S4).

As **9**-**15** are octaketides and formation of the AQ heptaketides was postulated to involve AntI, an *E. coli* strain expressing all *ant* genes except *antI* was analysed showing the loss of production of **1** (Fig. 2c). Therefore, from the structures of **9**-**15** an intermediate **16** can be proposed that is either the substrate for AntI resulting in the formation of **1** or that can bind NH_3 resulting in **11**, which is further transformed to **12**-**15** (Fig. 1b). In order to prove **16** as intermediate, which was also proposed as intermediate in the biosynthesis of actinorhodin, we expressed the codon optimized *actVI-ORF1* encoding the ketoreductase RED1 from the actinorhodin biosynthesis[10] in *E. coli* and could indeed show the production of 4-dihydro-9-hydroxy-1-methyl-10-oxo-3-H-naptho-[2,3-c]-pyran-3-(S)-aceticacid (*S*-DNPA, **17**) when *antA-H* were coexpressed (Fig. S5). Previous experiments with RED1 in the actinorhodin biosynthesis suggest a free acid as substrate[10] that could also lead to the formation of **11**-**15**.

AntI is a hydrolase with similarity to 2,6-dihydroxy-*pseudo*-oxynicotine hydrolase (DHPON),[11] a C-C bond hydrolase from *Arthrobacter nicotinovorans*, catalyzing a similar chain shortening as can be proposed for the catalytic mechanism of AntI (Fig.

171

3a). In order to test the proposed mechanism, AntI was heterologously produced in *E. coli* and used for *in vitro* experiments with model compounds **18**-**21** of the key intermediate **16** (Fig. 3b) synthesized chemically (Fig. S6). Formation of **22** from these substrates indicated that AntI is required for C-C bond cleavage but not for the intramolecular cyclisation and subsequent formation of the third ring: The overall polyketide shortening might require the ACP-bound intermediate **16** and not the free acid since there was more **22** formed when the methyl ester **18** was used as substrate compared to the free acid **19** (Fig. 3b). Formation of the third ring was dependent only on an activated carboxylic acid (methyl ester in **20**) but independent of AntI since **22** was also formed from **20** without any enzyme. Molecular modeling of AntI based on the DHPON structure (Fig. S7) allowed the prediction of the binding pocket and a catalytic mechanism with a catalytic triad (Fig. 3c) that was further confirmed by mutagenesis of *antI* (Fig. S8). Here, either D326 is not part of the catalytic triad but D327 or D327 can complement the loss of D326 but not vice versa. Interestingly, also in enzymes with a similar postulated mechanism the conserved Asp often has additional Asp nearby (Fig. S9). Additionally, in the modeled structure two tunnels are visible that might allow access of the ACP-bound **16** (Fig. S10). A similar C-C bond cleavage has been proposed in the biosynthesis of 1,3,5,8-tetrahydroxynapthaline derived melanin and thus the respective enzymes Ayg1p[12] (Fig. S11) and WdYG1[13] (Fig. S12) were also modeled confirming the proposed mechanism that was furthermore supported by the similarity of all proteins.

From these results the following mechanism is proposed for AntI and related enzymes (Fig. 3a): First, the activated serine residue attacks the beta-keto group of the polyketide intermediate, which might still be ACP bound, resulting in a tetrahedral oxyanion as transition state that might eliminate the terminal and probably still enzyme bound C_2 fragment resulting in an enzyme bound heptaketide intermediate. Second, the acyl enzyme intermediate reacts with the enolate at the second side chain (probably) activated by AntI, leading to recovery of the enzyme and formation of the third ring. The fact that no C-C cleavage of the more simple substrate 3-oxo-4-naphthylbutyric acid (ONBA) or the respective N-acetylcysteaminethioester (ONB-SNAC) (Fig. S6b) could be detected might indicate that without the second side chain it might not be possible to release AntI from the acyl enzyme intermediate. In organic synthesis, the second step known also as Claisen condensation is a common

method to form a C-C bond between one ester and another carbonyl under basic condition.[14] The last two steps including aromatization and oxidation might then take place spontaneously.

With the function of AntI elucidated we moved forward to reveal the role of AntG showing similarity to CoA ligases. From the previous results a function in the early biosynthesis was most likely. Thus, the genes encoding the minimal PKS *antDEF* were expressed with and without the phosphopantetheinyl transferase (PPTase) *antB* and CoA ligase *antG* (Fig. 4a). Surprisingly, both *antB* and *antG* were required for the production of SEK4 (**4**) and SEK4b (**5**) and both proteins were also required for the phosphopantetheinylation of *apo*-ACP AntF into its *holo*-form (Fig. 4b). The interaction between the ACP AntF and the PPTase AntB seems to be highly specific since no polyketides were produced upon exchange of *antB* against *mtaA* or *sfp* (Fig. 4a) and no *holo*-AntF could be observed when AntB was exchanged against Sfp or MtaA, two known PPTases with broad carrier protein activity (Fig. S13).

Additionally, coexpression of *antDE* with genes encoding ActI-ORF3 (ACP from the actinorhodin biosynthesis)[15] or RemC (ACP from the resistomycin biosynthesis)[16] did not result in polyketide production indicating also a specific interaction between AntF and AntDE. *In vitro* data with purified proteins (Fig. S14 and S15, Table S8) revealed that AntG is not required for the biosynthesis once *holo*-AntF is present, indicating a function in the generation of *holo*-AntF.

Although AntG looks like a functional CoA ligase with all known sequence motifs (Fig. S16),[17] no CoA ligase activity could be observed with acetate or malonate as putative substrates involved in AQ biosynthesis (Fig. S17). Therefore the current hypothesis is that AntG might have lost its original (CoA ligase) function and now functions as a chaperone to stabilize a protein-protein interaction important during the biosynthesis. However, additional research is needed to prove this hypothesis.

In summary we have revealed the mechanism of AntI catalyzed polyketide shortening and have shown the important role of the CoA ligase AntG in the initiation of the AQ biosynthesis in *Photorhabdus*. To our knowledge, AQs from *Photorhabdus* and aurachins[8] from the myxobacterium *Stigmatella aurantiaca* are the only two type II PKS derived natural products from Gram-negative bacteria and both show unique biosynthesis mechanisms. Regarding *Photorhabdus*, all *ant* gene clusters are highly

173

homologous indicating similar mechanisms in all strains (Fig. S18a). Moreover, *ant-*like gene clusters also encoding homologs of the unusual KS$_\beta$ AntE characteristic for the *ant* cluster[7] have also been detected in Gram-positive *Streptococcus* and *Lactobacillus* strains associated with the human oral cavity as well in a *Hoeflea* sp., a Gram-negative alphaproteobacterium from the Baltic sea (Fig. S18b). Although none of these bacteria has been described as polyketide producer, the natural products derived from these gene clusters might play an important ecological function. Finally, as a phylogenetic analysis of the two ketosynthases AntD and AntE also shows similarity to other unusual ketosynthases like RemC from *Streptomyces* (Fig. S19 and S20), the full elucidation of the AQ biochemistry in *Photorhabdus* might also strengthen our knowledge regarding the biosynthesis of the medically important type II PKS derived products from actinomycetes in general. Here the simple expression of the *ant* genes in *E. coli*[18] opens up a new avenue for the investigation of the underlying principles exemplarily shown here for the coexpression of AntABCDEFGH and RED1.

11.3. Unusual start and finish of anthraquinone biosynthesis in Photorhabdus luminescens
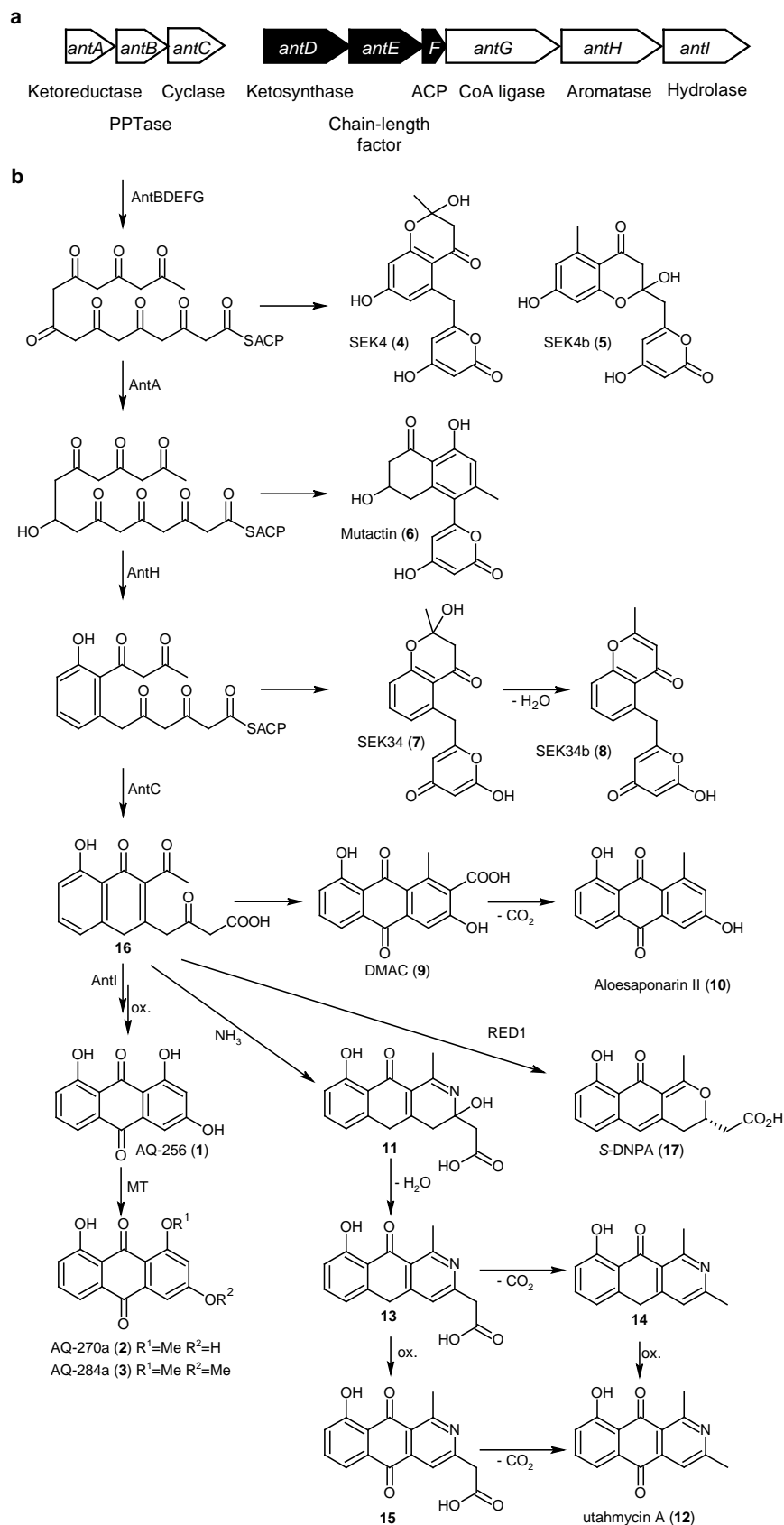
Fig. 1. Biosynthesis gene cluster *antABCDEFGHI* for the production of anthraquinones from *P. luminescens* (**a**) and proposed biosynthesis of all identified compounds (**b**).
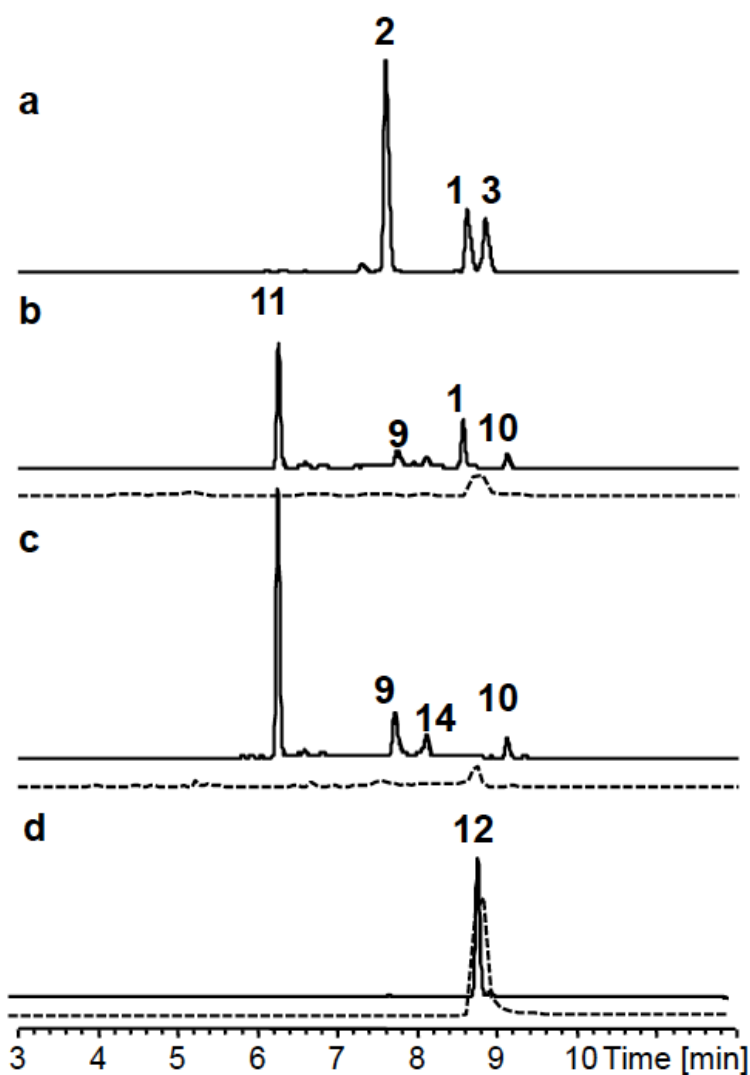
Fig. 2. HPLC/UV analysis (420 nm) of *P. luminescens* (**a**), *E. coli* expressing *antABCDEFGHI* (**b**), *E. coli* expressing *antABCDEFGH* (**c**) and utahmycin A (**12**) standard (**d**). As only small amounts of **12** were detected in **b** and **c**, an extracted ion chromatogram of its characteristic signal (*m/z* 254 [M+H]$^+$) is also shown (dashed line).
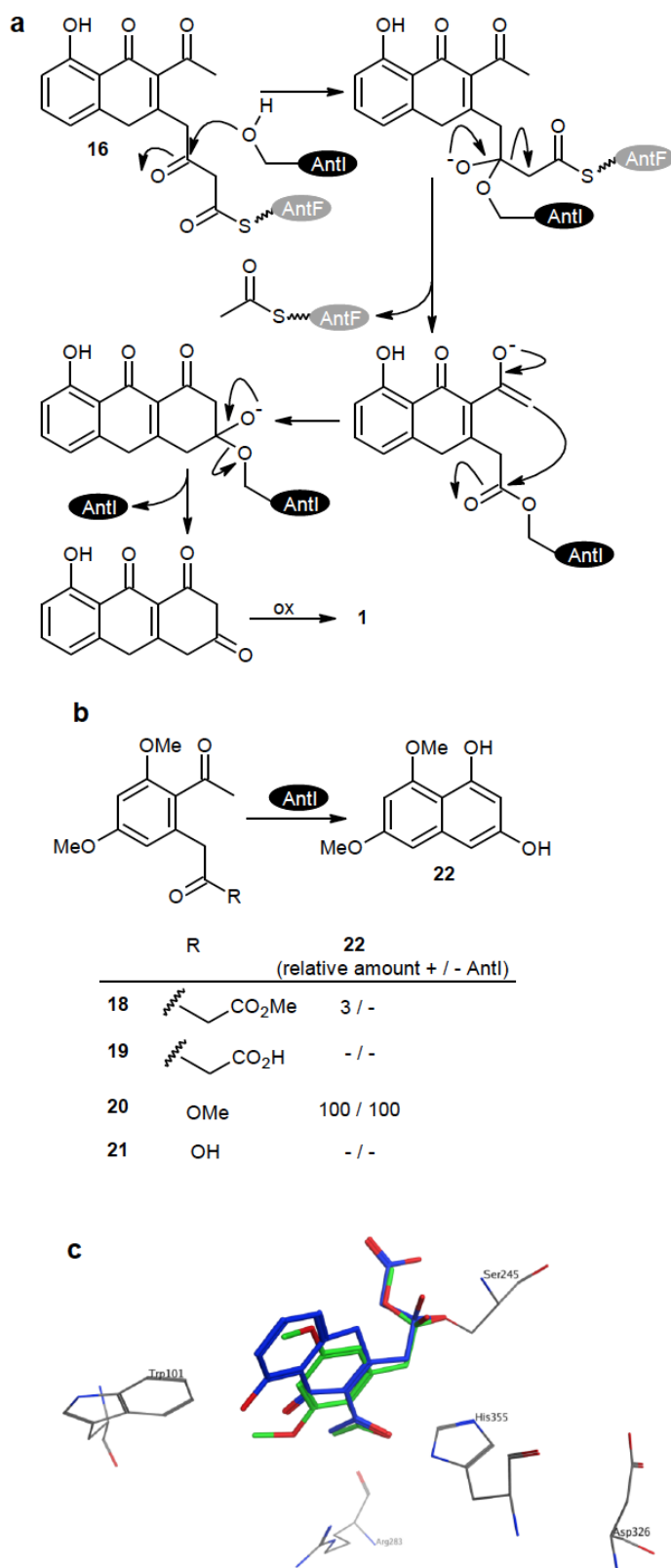
Fig. 3. Proposed mechanism for polyketide shortening catalysed by AntI (**a**), AntI catalysed conversion of model compounds **18-21** (**b**), and representation of the AntI binding pocket with docked **16** (blue) and **20** (green) covalently bound to Ser245 showing the proposed catalytic triadeSer245, His355 and Asp326 (**c**).

9
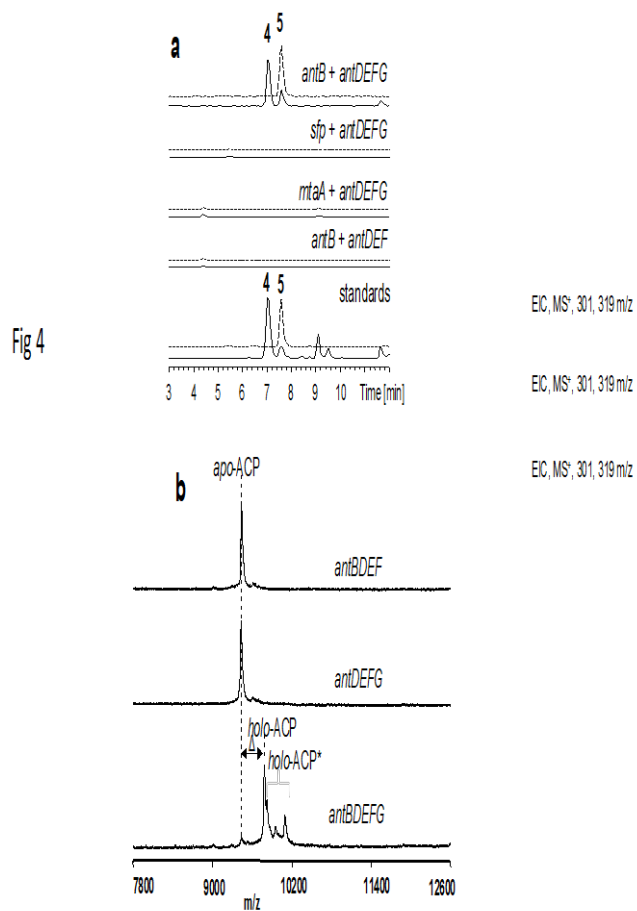
Fig. 4. a) HPLC/MS analysis of *E. coli* expressing different PPTases and *ant* genes. Shown are EICs for **4** (*m/z* 301 [M+H]⁺) and **5** (dashed line, *m/z* 319 [M+H]⁺). **b**) UTL-MALDI-TOF MS results from lysed *E. coli* cells expressing different *ant* genes.

References

[1]   C. Hertweck, *Angew. Chem. Int. Ed.* **2009**, *48*, 4688–4716.

[2]   J. Staunton, K. J. Weissman, *Nat Prod Rep* **2001**, *18*, 380–416.

[3]   C. Khosla, D. Herschlag, D. E. Cane, C. T. Walsh, *Biochemistry* **2014**, *53*, 2875–2883.

[4]   C. Hertweck, A. Luzhetskyy, Y. Rebets, A. Bechthold, *Nat Prod Rep* **2007**, *24*, 162.

[5]   J. Zhan, *Curr Top Med Chem* **2009**, *9*, 1958–1610.

[6]   M. B. Austin, J. P. Noel, *Nat Prod Rep* **2003**, *20*, 79–110.

[7]   A. O. Brachmann, S. A. Joyce, H. Jenke-Kodama, G. Schwär, D. J. Clarke, H. B. Bode, *ChemBioChem* **2007**, *8*, 1721–1728.

[8]   A. Sandmann, J. Dickschat, H. Jenke-Kodama, B. Kunze, E. Dittmann, R. Müller, *Angew. Chem. Int. Ed.* **2007**, *46*, 2712–2716.

[9]   J. D. Bauer, R. W. King, S. F. Brady, *J. Nat. Prod.* **2010**, *73*, 976–979.

[10]   T. Itoh, T. Taguchi, M. R. Kimberley, K. I. Booker-Milburn, G. R. Stephenson, Y. Ebizuka, K. Ichinose, *Biochemistry* **2007**, *46*, 8181–8188.

[11]   C. Schleberger, P. Sachelaru, R. Brandsch, G. E. Schulz, *J Mol Biol* **2007**, *367*, 409–418.

[12]   I. Fujii, Y. Yasuoka, H.-F. Tsai, Y. C. Chang, K. J. Kwon-Chung, Y. Ebizuka, *J Biol Chem* **2004**, *279*, 44613–44620.

[13]   M. H. Wheeler, D. Abramczyk, L. S. Puckhaber, M. Naruse, Y. Ebizuka, I. Fujii, P. J. Szaniszlo, *Eukaryotic Cell* **2008**, *7*, 1699–1711.

[14]   S. Akai, K. Kakiguchi, Y. Nakamura, I. Kuriwaki, T. Dohi, S. Harada, O. Kubo, N. Morita, Y. Kita, *Angew. Chem. Int. Ed.* **2007**, *46*, 7458–7461.

[15]   H. Fu, S. Ebert-Khosla, D. A. Hopwood, *J. Am. Chem. Soc.* **1994**, *116*, 4166–4170.

[16]   K. Jakobi, C. Hertweck, *J. Am. Chem. Soc.* **2004**, *126*, 2298–2299.

[17]   K.-H. Chang, H. Xiang, D. Dunaway-Mariano, *Biochemistry* **1997**, *36*, 15650–15659.

[18]   I. Fujii, *Nat Prod Rep* **2009**, *26*, 155.

Supplementary Information

Material and methods

Cultivation of strains. All *E. coli* DH10B strains (Table S3), *E. coli* BL21 (DE3) without any overexpression plasmids and *Photorhabdus* strains (Table S6) used in this study were cultivated in liquid or solid Luria-Bertani (LB, pH 7.0) medium at 30°C. All *E. coli* BL21 (DE3) strains with different overexpression plasmids (Table S3) were cultivated in a standard cultivation condition for protein expression (method described below). The cultivation condition for construction of *Photorhabdus luminescens* mutants was described below. Chloramphenicol (34 µg mL^{-1}), ampicillin (100 µg mL^{-1}), kanamycin (50 µg mL^{-1}), streptomycin (50 µg mL^{-1}) and rifampicin (50 µg mL^{-1}) were used for the selection of strains with corresponding resistant markers. For cultivation of *E. coli* strain ES15 in ^{15}N or ^{13}C labeled medium, ISOGRO-^{15}N or ISOGRO-^{13}C medium was prepared according to manufacturer´s instructions (Sigma-Aldrich). Isopropyl-β-D-1-thiogalactopyranoside (IPTG, 0.1 mM) was used to induce the heterologous expression of *ant* genes in *E. coli* BL21 (DE3).

General methods in molecular biology. All methods used in molecular biology were conducted according to standard procedures and manufacturer´s instructions. All oligonucleotides (primers) obtained from Sigma-Aldrich were listed in Table S1. All Plasmids constructed in this work were confirmed by sequencing at the SeqIT GmbH (Germany, Kaiserslautern) also listed in Table S2. Polymerase chain reactions (PCRs) were performed using the phusion high-fidelity polymerase (Thermo-scientific). DNA isolation was performed with GeneJET™ Gel Extraction Kit (Fermentas). Plasmid isolation was performed with GeneJET™ Plasmid Miniprep Kit (Fermentas). Transformation of plasmids into *E. coli* was carried out using electroporation protocol for *E. coli* in an electroporation cuvette with a width of 1 mm (1250V, 25 µF, 200 Ω). Genomic DNA was isolated using Gentra® Puregene® kit (Qiagen) according to the protocol for Gram-negative bacteria. Plasmids ZQ80-82 were constructed using an artificial gene (synthesized by Life Technologies™) flanked with restriction sited *Eco*RI/*Pst*I.

Construction of *E. coli* strains with different combinations of *ant* genes. Two strategies were used to combine different *ant* genes in this work. The first one is the

combination of pJET1.2/blunt and pSU18 based plasmids. The second one is the combination of pCOLA Duet-1, pACYC Duet-1 and pCDF Duet-1 based plasmids. All *E. coli* strains used in this work (with the respective plasmids) were listed in Table S3. Plasmids derived from pJET1.2/blunt and pSU18 vector were transformed into *E. coli* DH10B for heterologous expression of *ant* genes. Plasmids derived from Duet-vectors were transformed into *E. coli* BL21 (DE3) for IPTG induced heterologous expression of *ant* genes.

Heterologous expression of genes in *E. coli* DH10B: for analysis of produced polyketides. First of all, different DNA fragments were amplified using primers in Table S1, resulting in *antB*, *antABC*, *antAB*, *antBC* PCR products. The PCR products were then digested with restriction enzymes *Nde*I/*Sac*I and subsequently ligated into pSU18 vector, generating plasmids (Table S2) ZQ6, ZQ10, ZQ11 and ZQ16, respectively. The plasmid pSUsfp was cloned using the same method. Later, *antDEF*, *antDEFG*, *antDEFGH* and *antDEFGHI* genes were amplified with the primers in Table S1. The resultedPCR products were directly ligated into pJET1.2/blunt, generating the plasmids ZQ1, ZQ9, ZQ12 and ZQ13 in Table S2. The plasmid pSUmtaA was described previously.[1] All generated plasmids were selectively transformed into *E. coli* DH10B for heterologous expression, yielding strains ES1-17 in Table S3 and S4. The polyketides found in the bacteria culture were listed in Table S4.

Heterologous expression of genes in *E. coli* BL21 (DE3): for protein purification. In this work, proteins in Table S7 were overexpressed and purified for *in vitro* assays.

The gene *plu2834* was amplified using primers in Table S1. The PCR product was then digested with restriction enzymes *Pst*I/*Eco*RI and subsequently ligated into pCOLA Duet-1 vector, generating plasmid ZQ37 (Table S2). Plasmids ZQ40 (encoding AntI), ZQ46 (encoding *apo*-AntF), ZQ47 (encoding AntG) and ZQ48 (encoding AntB) were constructed using the same procedure. Plasmid pET-MatB was kindly provided by the Tang lab.[2] All plasmids were transformed into *E. coli* BL21 (DE3) for N-terminal His$_6$-tagged protein expression.

To purify proteins AntD and AntE, two separated plasmids were constructed as described above for Plu2834. But, N-terminal His$_6$-tagged AntD and N-terminal His$_6$-

2

tagged AntE could be only expressed as inclusion bodies in *E. coli* (data not shown in this work). According to previous work, it was proposed that AntD and AntE build up a heterodimer. Thus, a new plasmid ZQ32 (Table S2) encoding both genes *antDE* was constructed. First, the gene *antE* was amplified by PCR and digested with restriction enzymes *Nde*I/*Acc*65I and ligated into pCOLADuet-1 vector, resulting in the plasmid pCOLADuet-1_AntE. After that, PCR product of the gene *antD* was digested with restriction enzymes *Eco*RI/*Pst*I and ligated into digested pCOLADuet-1_AntE, resulting in plasmid ZQ32. Primers used here are listed in Table S1. At the end, plasmid ZQ32 was transformed into *E. coli* BL21 (DE3) for coexpression of N-terminal His$_6$-tagged AntE and natural AntD. The heterodimer could be co-purified at one step using standard purification method.

Because the PPTase AntB could not be purified in this work, *in vitro* activation of *apo*-ACP could not be realized. Based on the result that polyketides SEK4/SEK4b could be produced by coexpression of AntBDEFG, plasmid ZQ65 (Table S2) was constructed in two steps. A PCR product of the gene *antF* was digested with restriction enzymes *Pst*I/*Eco*RI and ligated into pACYCDuet-1 vector, resulting in the plasmid pACYCDuet-1_AntF. Then, PCR product of the genes *antDEFG* was digested with restriction enzymes BglII/Acc65I and ligated into pACYCDuet-1_AntF, yielding plasmid ZQ65. Primers used here are listed in Table S1. Plasmids ZQ65 and ZQ48 (described above) were transformed into *E. coli* BL21 (DE3) for coexpression of AntDBEFG, resulting in strain ES23 for purification ofN-terminal His$_6$-tagged AntF. His$_6$-tagged AntF was proved by MALDI-TOF as activated His$_6$-*holo*-ACP.

Heterologous expression of genes in *E. coli* BL21 (DE3): for mutagenesis of AntI. ZQ40_S245A, ZQ40_D326A, ZQ40_D327A and ZQ40_H355A harboring the mutated gene *antI**, were generated (Table S2). The site-directed mutation was achieved by ligation of two amplified PCR fragments, using the designed primer pair binding around the mutated region (Table S1) and a primer pair binding at ColA *ori* (pCOLA Duet-1).The plasmid ZQ40 harboring the original *antI* gene was used as PCR template. The primer pair binding at ColA *ori* is 5`-[Phos]GTGGATTTAGATATCGAGAGTC-3´ and 5´-TAATTCTCAGTTACCAATGGC-3´. For coexpression, plasmid ZQ62 encoding AntDEFGH and plasmid ZQ76 encoding AntABC were constructed using primers in Table S1. Transformation of plasmids ZQ40, ZQ62 and ZQ76 into *E. coli* BL21 (DE3) led to stran ES53.

Transformation of plasmids encoding mutated AntI* into *E. coli* BL21 (DE3) containing plasmids ZQ62 and ZQ76 led to strains ES59, ES61, ES62 and ES64 in Table S3. HPLC-MS analysis is shown in Figure S8.

Heterologous expression of genes in *E. coli* BL21 (DE3): for coexpression with RED1. RED1 encoded originally by *actVI-ORF1* in *Streptomyces coelicolor*A3(2) was investigated in order to find out whether the polyketide intermediate**16** produced by AntA-H can serve as a substrate for the stereospecific ketoredution.[3] For this experiment, a gene encoding RED1 optimized for expression in *E. coli* was synthesized (Life Technologies™). The artificial gene flanked with restriction sites *Eco*RI/*Pst*I was digested and ligated into pCOLA Duet-1 vector, resulted in plasmid ZQ80 (Table S2). Plasmid ZQ62 encoding AntDEFGH and Plasmid ZQ76 encoding AntABC (described above) were used for coexpression with plasmid ZQ80 in *E. coli*, resulting in strain ES65 in Table S3.

DNA sequence of the synthesized artificial gene encoding ketoreductase RED1 from the actinorhodin biosynthesis (*actVI-ORF1***):
^{EcoRI}*GAATTC*CATGAGCACCGTTACCGTTATTGGTGCAGGCACCATTGGTCTGGG
TTGGATTAACCTGTTTAGCGCACGTGGTCTGACCGTTCGTGTTAATAGCCGTCG
TCCGGATGTTCGTCGTGTTGTTCATGAAGCACTGGAACTGTTTAGTCCGGGTCG
TGTTGATGAACTGGCAGCACGTATTGAATATGAACCGGATGTGGGTCGTGCAGT
TGCCGGTGCAGATGTTGTTAGCGAAAATGCACCGGATGATCTGCCGCTGAAAC
AGCGTCTGTTTGCAGAAATTGGTGCCGCAGCACCGGATCATGCACTGGTTCTGA
GCAGCACCAGCAAACTGCTGCCGGATGAACTGAGCCGTGATATGCCTGGTCCT
GGTCGTCTGGTTGTTGCACATCCGTTTAATCCGCCTCATATTGTTCCGCTGGTT
GAAGTTGTTCGTGGTGAACGTACCGATCCGGAAGCAGTGGAACGTACCCTGGC
ATTTCTGGCAAGCGTTGGTCGTACACCGGTTGTTGTGCGTCGTGCACTGCCTG
GTTTTGCAGCAAATCGTCTGCAAAGCGCACTGCTGCGTGAAAGCATTCATCTGG
TTCTGGAAGGTGTTGTTACCGTTGAAGAACTGGATCGTATTGTTACCGATAGTAT
TGGCCTGCGTTGGAGCACAATTGGTCCGTTTCATGCATTTCATCTGGGTGGTGG
TCCGGGTGGTCTGCGTAAATGGCTGGAACATCTGGGTAGCGGTCTGGAACAGG
GTTGGCGTGGTCTGGGCCAGCCTGCACTGACACCGCAGGCAGTTGAAGCCCT
GGTTGCACAGACCGAAGCAGCATATGGTCATCGTCCGTATGCAGAACTGGTTC

GTGATCGTGATGATCGTCATCTGGCCGTTCTGGCAGCCCTGGAACGCACCGAA
CAGCCGCAAGAAGAAACCAAATAA*CTGCAG*^{PstI}

Heterologous expression of genes in *E. coli* BL21 (DE3): Coexpression with RemC or ActI-ORF3. RemC (ACP from the resistomycin biosynthesis)[4] and ActI-ORF3 (ACP from the actinorhodin biosynthesis) were coexpressed with AntDE and Sfp in *E. coli* BL21 (DE3). For this purpose, artificial genes *actI-ORF3*** and *remC*** were synthesized by the company Life Technologie™ after codon optimization for *E. coli* expression. Genes *antDE* were amplified using primers in Table 1. The PCR fragment was digested with *Bgl*II/*Acc*65I and ligated into pACYC Duet-1 vector, resulting in pACYC Duet_antDE. The artificial genes were digested with *Eco*RI/*Pst*I and ligated into pACYC Duet_antDE, resulting in plasmid ZQ81 and ZQ82 (Table S3). Plasmid ZQ78 encoding for PPTase Sfp was cloned using pCDFDuet-1. For coexpression, plasmids ZQ81 and ZQ82 were transformed into *E. coli* BL21 (DE3) containing ZQ78, resulting in *E. coli* strain ES67 and ES68, respectively. The strains were cultivated as described for protein expression in small amount. After induction with IPTG, there were none polyketide products detected.

DNA sequence of the synthesized gene encoding ACP from the actinorhodin biosynthesis (*actI-ORF3***):
^{EcoRI}*GAATTC*CATGGCAACCCTGCTGACCACCGATGATCTGCGTCGTGCACTGGT
TGAATGTGCCGGTGAAACCGATGGCACCGATCTGAGCGGTGATTTTCTGGATCT
GCGTTTTGAAGATATTGGCTATGATAGCCTGGCACTGATGGAAACCGCAGCACG
TCTGGAAAGCCGTTATGGTGTTAGCATTCCGGATGATGTTGCAGGTCGTGTTGA
TACACCGCGTGAACTGCTGGATCTGATTAATGGTGCACTGGCAGAAGCAGCATA
A*CTGCAG*^{PstI}

DNA sequence of the synthesized gene encoding ACP from the resistomycin biosynthesis (*remC***):
^{EcoRI}*GAATTC*CATGGATGTTCGTGAACAGGTTTTTGCAGTTATGCGTGGTGTTGGT
TTTGATGGTGATGAACTGGCAGATGGTAAACGTCTGGCAGAAGATCTGGATATT
GATAGCACCGAACTGGTTGAAATTGTTGTTGCACTGGAACAGCATTTTGATATTA
CCGTTGATGCAGATGCAGAAGGTGGTTTTACCACCGTTGGTGATCTGGTTGGTT
GTGTTGGTCGTCTGCTGAGCGCAGGTAGCGCAGCACGTGGTTAA*CTGCAG*^{PstI}

Heterologous expression of genes in *E. coli* BL21 (DE3): for analysis of AntF by UTL-MALDI-TOF MS. Plasmids ZQ56 and ZQ61 were constructed for overexpression of *antDEFG* and *antDEF* genes in *E. coli*, respectively (Table S2). For this reason, *antDEFG* and *antDEF* genes were amplified using primers in Table S1. The resulted PCR products were digested with restriction enzymes *Eco*RI/*Pst*I and ligated into pACYCDuet-1 vector, yielding plasmids ZQ56 and ZQ61, respectively. Transformation of plasmids ZQ56 and plasmids ZQ48 (previously described) into *E. coli* BL21 (DE3) resulted in strain ES24. Transformation of plasmids ZQ61 and plasmids ZQ48 into *E. coli* BL21 (DE3) resulted in strain ES25. Transformation of plasmid ZQ56 into *E. coli* BL21 (DE3) resulted in strain ES26. All three *E. coli* strains were used to analyze whether AntF (*apo*-ACP) is transformed to *holo*-ACP.

Construction of *P. luminescens* mutants. Insertion mutants (*antI*::cat and *antG*::cat) and the deletion mutant (Δ*antC*) were constructed as described previously.[5]

Analysis of polyketides in bacteria culture by HPLC-MS. For production analysis, extracts from 10 ml of culture were prepared. Usually, 1 % overnight pre-culture was used for inoculation and 2 % AMBERLITE™ XAD-16 (Sigma Aldrich) was added at the same time for the absorption of hydrophobic secondary metabolites. AMBERLITE™ XAD-16 resin beads were collected with a sieve (or filter paper) after 72 h, washed with small amount of water and extracted with 20 mL of methanol. The methanol extract was dried and dissolved in 1 mL methanol, which could be then diluted (1:10-dilution) for the analysis in HPLC-MS. Additionally, the remaining bacteria culture without AMBERLITE™ XAD-16 and bacteria cells was extracted using 10 mL ethyl acetate/acetic acid (99:1). After remove of all the solvent, the dried extract was dissolved in 1 mL methanol, which could be then diluted (1:10-dilution) for the analysis in HPLC-MS.

Analysis of extracts was performed on Dionex ULTIMATE 3000 HPLC system with a photodiode array detector in the range of 200-600 nm and an Acquity UPLC BEH C18 column (1.7 µm, Waters), which is coupled to Bruker AmaZon X iontrap mass

6

spectrometer using electrospray ionization at positive and negative mode. Solvent A: water with 0.1 % formic acid and solvent B: acetonitrile with 0.1 % formic acid. Gradient A: 5 % solvent B for 2 min, increasing to 95 % solvent B in 12 min, 95 % solvent B for 3 min. Gradient B: 10 % solvent B for 4 min, increasing to 25 % in 9 min, increasing to 95 % solvent B in 1 min, 95 % solvent B for 3 min. Flow rate: 0.6 mL min^{-1}. Gradient A was used as standard gradient for most analysis in this work. Gradient B was used for separation of **4**, **5**, **6** and **7** (Table S5).

Analysis of ACP activation in *E. coli* with UTL-MALDI-TOF MS. Three *E. coli* strains ES24-26 (Table S3) were used for heterologous expression of *ant* genes in *E. coli*. After all three strains were cultivated with the standard protocol for protein expression, the *E. coli* cells were harvested by centrifugation. The cells were then treated with 10 μg/ml lysozyme and 2.5 U/ml benzonase in PBS buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4, pH = 7.4). The cell lysates were ready for the ultra thin layer (UTL) preparation. For the preparation of the UTL substrate, 10 μL of the thin layer substrate solution, consisting of 1 part of saturated CHCA (66 % ACN, 0.1 % TFA) and 3 parts of isopropanol, was spread on the plate with the side of a pipette tip. The dried plate was then gently wiped with a soft tissue. The cell lysates were mixed with the matrix solution (1:10, usually 0.5 μL cell lysate with 4.5 μL matrix solution), consisting of saturated CHCA in 3 parts of formic acid and 2 parts of isopropanol. Shortly after, 0.5 μL of the lysate matrix solution was spotted onto the ultra thin layer.[6]

Protein expression. For small scale of protein expression, *E. coli* BL21 (DE3) strains with overexpression plasmids (strains ES18-ES65 in Table S3) were cultivated in 10 mL LB with appropriate antibiotics at 30 °C to an OD_{600} of 0.6-0.8, using 1% overnight culture for inoculation. After the cell cooled on ice for 15 min, protein expression was induced with 0.1 mM IPTG at 16 °C for 16 h. For comparison, a same culture was cultivated as described above, but without addition of IPTG. For large scale of protein expression, the culture volume was adjusted to 250 mL in a 1L flask.

Protein purification. The *E. coli* BL21 (DE3) strains ES18 (for protein His$_6$-Plu2834), ES19 (for protein His$_6$-AntG), ES20 (His$_6$-*apo*-AntF), ES21 (for heterodimer proteins

His$_6$-AntD-AntE), ES22 (for protein His$_6$-AntB), ES23 (for protein His$_6$-*holo*-AntF) and ES41 (for protein His$_6$-MatB) were cultivated as described above. The bacterialcells were harvested, washed once with PBS buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na$_2$HPO4, 2 mM KH$_2$PO4, pH = 7.4) and lysed with 10 µg/ml lysozyme in binding buffer (500 mM NaCl, 20 mM imidazol, 20 mM Tris HCl, pH = 7.5 at 4 °C). Benzonase (2.5 U/ml) was used to digest the DNA and lower the viscosity of the solution. The cell debris was separated from the protein solution by centrifugation at 10000 g. The proteins with His$_6$-tag were purified using ÄKTAexplorer™ (GE Healthcare) equipped with a1 mL His Trap™ HP nickel affinity column at 4°C. The binding buffer (described above) and elution buffer (500 mM NaCl, 500 mM imidazol, 20 mM Tris HCl, pH = 7.5 at 4 °C) were used for the purification. Unfortunately, His$_6$-AntB could only been overexpressed as inclusion bodies after many attempts using several different conditions (not shown here). Purified protein Cpin_1856 was kindly provided by Dr. S. Fuchs. Purified proteins Sfp and MtaA were kindly provided by Dr. C. Kegler.

Proteins solutions were concentrated using Amicon® Ultra filters (Merck Millipore) according to manufacturer´s instructions. SDS-Page was used to check the protein size and the purity (Figure S14). UTL-MALDI-TOF MS was used to analysis the purified AntF (*apo* or *holo*). Protein concentrations were determined by a NanoVue™ Plus Spectrophotometer (GE Healthcare).

Biochemical assays. Hydrolase AntI activity assays were performed with synthesized model compounds **18**, **19**, **20**, **21**, ONBA and ONB-SNAC. The organic syntheses of **18**, **19**, **20** and **21** were described below and in Figure S6. Compounds ONBA and ONB-SNAC were synthesized as described.[3]The methyl esters (**19** and **20**) were hydrolyzed in situ with 250 µM NaOH solution with a final concentration of 2 mM for synthetic analogs. The assays with the four model compounds (2.5 µM, **18**, **19**, **20** and **21**) were realized with 250 µM His$_6$-AntI in Tris-HCl (pH 8).The assays with the other two model compounds (1 mM ONBA and 1 mM ONB-SNAC) were performed in 20 µL Tris-HCl buffer (100 mM, pH 7.5) with 1 µM His$_6$-AntI at room temperature overnight. Because of the low solubility, ONB-SNAC partially precipitated from the assay. ONBA was completely soluble under the assay solutions. The same assays were performed withoutHis$_6$-AntI for the negative

8

controls. To stop the reactions, 200 µL ACN were added. After centrifugation, the clear solutions were analyzed by HPLC/MS analysis quantifying the formation of **22** (m/z 221 [M+H]$^+$). Results showed that model compounds **19**, **21**(Figure 3), ONBA and ONB-SNAC could be not converted by His$_6$-AntI.

In order to study the biosynthesis of polyketide products **4** and **5** using purified minimal PKS proteins, a typical reaction was performed in 50 µl Tris-HCl buffer solution (100 mM, pH 7.5). To stop the reactions, 200 µL of ACN were added. After centrifugation, the clear solutions were used for HPLC/MS analysis. The other components used for assays and results are listed in Table S8. The results were also shown in Figure S15 as chromatogram.

Assays of activation of His$_6$-*apo*-AntF with PPTase (Sfp or MtaA) were performed in 20 µl Tris-HCl (100 mM, pH = 7.5) buffer solutions containing 1 µMHis$_6$-*apo*-AntF, 0.1 µM PPTase, 5 mM CoA and 1mM MgSO$_4$. The assays were incubated at room temperature for 2 h. The assay solutions were directly used for mixing with matrix solution (for UTL-MALDI-TOF MS described above). The protein Cpin_1856 (*apo*-ACP) was used as a positive control. The results were shown in Figure S13.

The CoA ligase activities of His$_6$-AntG (0.5 µM) were test with sodium acetate/sodium malonate (50 mM) with coenzyme A (10 mM) in the present of cofactors MgCl$_2$ (5 mM) and ATP (5 mM). The assays were performed in 20 µL Tris-HCl (100 mM, pH 7.5) buffer solutions at room temperature for 2h.The malonyl CoA ligase MatB was used as a positive control.[2] To stop the reactions, 200 µL ACN were added in to reaction tubes, which were then centrifuged. The clear solutions were used for HPLC/MS analysis. The results are shown in Figure S17.

Synthesis of 20. 145 mg of methyl 2-(3,5-dimethoxyphenyl)acetate were dissolved in 5 ml of acetic acid, then 540 µl of acetic anhydride and 3 drops of HClO$_4$ were added to the mixture. A yellow precipitate appeared. The mixture was heated to 35°C for 7 min. The completion of the reaction was monitored by TLC. The acid was quenched with ice and 20 ml of saturated NaHCO$_3$. The aqueous layer was extracted 3 times with 10 ml of Et$_2$O. The organic layers were combined and extracted with 20 ml of brine and then dried over Na$_2$SO$_4$. The solvents were removed under vacuum.

The crude product was purified by flash chromatography with a gradient from 40-80 % EtOAc in hexane over 10 column volumes. After evaporation of the solvents, 120 mg of **20** were obtained as yellow oil. ^{1}H NMR (500 MHz, CDCl$_3$) δ 6.31 (dd, J = 30.1, 2.1 Hz, 2H), 3.76 (s, 3H), 3.74 (s, 3H), 3.62 (s, 2H), 3.61 (s, 3H), 2.46 (s, 3H).^{13}C NMR (126 MHz, CDCl$_3$) δ 203.77, 171.77, 161.61, 159.45, 134.97, 123.64, 108.27, 97.50, 55.62, 55.42, 51.97, 39.10, 32.24.HRMS (ESI) Calcd for C$_{13}$H$_{17}$O$_5$: 253.1070 [M+H]$^+$, Found: 253.1089 [M+H]$^+$.

Synthesis of 24. 565 mg of methyl ester (2.7 mmol, 1 equiv.) were dissolved in 8.5 ml of dry toluene under nitrogene. The mixture was then cooled down to -78°C and 2.8 ml of Me$_3$Al in hexane (2M, 2.8 mmol, 1.03 eq.) was added. 1.55 g of tert-butyl((1-methoxyvinyl)oxy)dimethylsilane (7.57 mmol, 2.83 equiv.) were dissolved in 2 ml of toluene. This solution was slowly added to the reaction mixture. The mixture was warmed to 0°C and stirred for 7 hours. The mixture was quenched with 100 ml of saturated aqueous solution of NaHCO$_3$, and then extracted 3 times with 30 ml of EtOAc. The organic layers were combined und washed once with brine, then dried over Na$_2$SO$_4$. The solvent removed under vacuum and the residue oil was purified by flash chromatography using a gradient from 5-40 % of EtOAc in hexane over 10 column volumes. The two diasteroisomers of **24** were isolated, 250 mg of diastereoisomer A (22% of yield) and 120 mg of diastereoisomer B (11%) were obtained both as yellow oil. Isomer A:^{1}H NMR (250 MHz, CDCl$_3$) δ 6.44 (d, J = 2.3 Hz, 2H), 6.21 (t, J = 2.3 Hz, 1H), 3.64 (s, 6H), 3.54 (s, 3H), 3.20 (s, 3H), 3.02 (dd, J = 33.6, 13.1 Hz, 2H), 2.44 (dd, J = 45.5, 14.2 Hz, 2H), 0.78 (s, 9H), 0.02 (d, J = 9.1 Hz, 6H).Isomer B:^{1}H NMR (250 MHz, CDCl$_3$) δ 6.44 (d, J = 2.3 Hz, 2H), 6.21 (t, J = 2.2 Hz, 1H), 3.64 (s, 6H), 3.54 (s, 3H), 3.20 (s, 3H), 3.02 (dd, J = 33.6, 13.1 Hz, 2H), 2.44 (dd, J = 45.4, 14.2 Hz, 2H), 0.78 (s, 10H), 0.02 (d, J = 9.1 Hz, 6H).^{13}C NMR (63 MHz, CDCl$_3$) δ 170.33, 160.30, 139.51, 108.96, 100.80, 98.73, 55.28, 51.44, 49.05, 44.50, 41.52, 25.86, 18.34, -2.67, -2.90.

Synthesis of 25. 160 mg of **24** was dissolved in 16 ml of MeOH, and 8 ml of HCl 6M were added at 0°C. The mixture was stirred 30 min at 0°C then the ice bath was removed and the mixture was stirred for 1 hour at room temperature. The mixture was quenched with 100 ml of NaHCO$_3$ and extracted three times with 25 ml of EtOAc. The organic layers were combined, washed with brine and dried over

Na$_2$SO4. The solvents were removed under vacuum and the crude was purified by flash chromatography. The expected compound **25** (63 mg)was obtained as yellow oil with a yield of 60%.^{1}H NMR (500 MHz, CDCl$_3$) δ 6.37 (d, *J* = 2.2 Hz, 1H), 6.34 (d, *J* = 2.2 Hz, 2H), 3.77 (s, 6H), 3.73 (s, 2H), 3.71 (s, 3H), 3.46 (s, 2H).^{13}C NMR (126 MHz, CDCl$_3$) δ 197.78, 165.06, 158.59, 132.71, 105.02, 104.81, 96.86, 52.80, 52.77, 49.84, 47.84, 45.18.

Synthesis of 18. 30 mg of **25** were dissolved in 0.190 ml of acetic acid, then 180 µl of acetic anhydride and two drops of HClO$_4$ were added to the mixture. The mixture color turned to orange. It was heated to 35°C for 5 min. The completion of the reaction was monitored by TLC. The acid was quenched with ice and 20 ml of saturated NaHCO3. The aqueous layer was extracted 3 times with 10 ml of Et$_2$O. The organic layers were combined and washed with 10 ml of brine and then dried over Na$_2$SO$_4$. The solvents were removed under vacuum. The crude product was purified by flash chromatography with a gradient from 40-80 % EtOAc in hexane over 10 column volumes. After evaporation of the solvents, 12 mg of **18** (yield: 36%) were obtained as a slight purple solid. ^{1}H NMR (500 MHz, CDCl$_3$) δ 6.35 (d, *J* = 2.2 Hz, 1H), 6.23 (d, *J* = 2.2 Hz, 1H), 3.78 (s, 3H), 3.77 (s, 2H), 3.75 (s, 3H), 3.67 (s, 3H), 3.51 (s, 2H), 2.41 (s, 3H).^{13}C NMR (126 MHz, CDCl$_3$) δ203.00, 199.83, 167.92, 162.03, 160.14, 135.60, 122.76, 108.83, 97.69, 55.62, 55.46, 52.31, 48.59, 48.27, 32.29.HRMS (ESI) Calcd for C$_{15}$H$_{19}$O$_6$: 295.1176 [M+H]$^+$, Found: 295.1184[M+H]$^+$.

Homology modeling of the hydrolase AntI. The protein sequences of AntI, Ayg1p and WdYG1 were used as queries for BLASTP[7] searches in the PDB[8], to identify the most similar available structure in the PDB. This resulted in the identification of the DHPON hydrolase from *Arthrobacter nicotinovorans*.[9]This template structure was used to create a sequence alignment applying the ClustalW algorithm[10]. The homology models were generated using the Homology Modelling Tool integrated in MOE 2012.10 (Molecular Operating Environment; Chemical Computing Group Inc., Montreal, Canada) and the ClustalW sequence alignment was imported. A series of ten models was created, for further processing the one with the highest packing quality score was chosen and energy minimized applying the AMBER12EHT (integrated in MOE) force field. All figures showing protein structures in this work, were created using MOE.

Docking calculations of Antl. Protein-ligand docking calculations were carried out using the program GOLD (version 5.2)[11] using the empirical scoring function for advanced protein-ligand docking CHEMPLP[12]. For each docking study the result with the highest docking score is shown in this work.

Multiple sequence alignments of hydrolases. All performed multiple sequence alignments were calculated using the ClustalW-algorithm[10] applying the standard parameters.

Phylogenetic analysis. A PHYML[13] tree (50 bootstraps) was calculated using a ClustalW alignment (gap opening: 10; gap extension: 0.1), which was generated using the collected ketosynthases. For visualization and calculation of the alignment as well as the PHYML tree the Geneious software (Biomatters Ltd., NewZealand) was used.

Table S1. Primers used for the plasmid construction and genotype verification. Restriction sites are underlined. Inserted nucleotides before start codon are marked in bold.

plasmid	primer 5` to 3`	vector
pSUsfp	AGCCGA<u>CATATG</u>AAGATTTACGGAATTT	pSU18
	CCGAAT<u>GGATCC</u>TTATAAAAGCTCTTCGTACG	
ZQ1	ATAATTAGAATTCGATGATAATAAATAACAGAAATG	pJET1.2/blunt
	ATTAGCCTGCAGTTAATTTTTATCGTTTAAAC	
ZQ6	TACTAA<u>CATATG</u>GACGATATTTCTTTATCATCTGATT	pSU18
	GTCAAT<u>GTCGAC</u>TATTACTCATCTTTGTTCCTTATAATCTCG	
ZQ9	ACTTGCAGATAACGATTTTC	pJET1.2/blunt
	ACATTCCTGGCCATTTAT	
ZQ10	TAGTTC<u>CATATG</u>AAATATGCCTTTATTACCGG	pSU18
	CTTGAA<u>GTCGAC</u>CCATTGGGTATATGAAATCTCTTT	
ZQ11	TAGTTC<u>CATATG</u>AAATATGCCTTTATTACCGG	pSU18
	TTCTCA<u>GTCGAC</u>TTCCCAAAAATCACAATCTATAGG	
ZQ12	ACTTGCAGATAACGATTTTC	pJET1.2/blunt
	CTATTGGGTTTATTTTTATTATTCATCT	
ZQ13	ACTTGCAGATAACGATTTTC	pJET1.2/blunt
	TTACCATCGCGATGTATATT	
ZQ16	TACTAA<u>CATATG</u>GACGATATTTCTTTATCATCTGATT	pSU18
	CTTGAA<u>GTCGAC</u>CCATTGGGTATATGAAATCTCTTT	
ZQ32	AGTAGA<u>CATATG</u>CGTAAAAGAGTTGTCGTTACC	pCOLADuet-1
	TTTAGA<u>GGTACC</u>TTAAATAGCTGAAAAACTCAACGC	
	TAATTA<u>GAATTC</u>**C**GTGATAATAAATAACAGAAATGAATCTCAACC	
	CACCAA<u>CTGCAG</u>CCGGTAACGACAACTCTTTTACG	
ZQ37	ACATGT<u>GAATTC</u>**C**ATGTCTGAATTTGCAATGGTATTT	pCOLADuet-1
	AATCTC<u>CTGCAG</u>ATTGCCCTCAATTATTTTCCA	
ZQ40	AAAATA<u>GAATTC</u>**C**ATGAATAATAAAAATAAACCCAATAGA	pCOLADuet-1
	GTATAA<u>CTGCAG</u>TCAATTAACCTTTTTTATAGCCA	
ZQ40_S245A	[Phos]CTTTTGGTGGTTATTTTGC	pCOLADuet-1
	CAATTCCTAAGAAACAAAGAAG	
ZQ40_D326A	[Phos]GAAAGTGAAAAAATTAGATCAAC	pCOLADuet-1
	GCGATCTGAAATATATCATCTAA	
ZQ40_D327A	[Phos]CTATATTTCAGATCGATAAAGTG	pCOLADuet-1
	CATCTAATTCACCATGAAC	
ZQ40_H355A	[Phos]CTGTTTGCTTAAATAAAATAAACG	pCOLADuet-1
	CAGCCTCTGATTCATAACATAA	
ZQ46	AAAAAT<u>GAATTC</u>**C**ATGAATAATCATCCAGAAGTAAA	pCOLADuet-1
	TTAATT<u>CTGCAG</u>CAGTTAATTTTTATCGTTTAAACTT	
ZQ47	AAACAG<u>GAATTC</u>**C**ATGAAACTAATCTCTATGTTGTTACA	pCOLADuet-1
	TCTACA<u>CTGCAG</u>ATTCATTATTGATTCCTCAATG	
ZQ48	TACTAA<u>GAATTC</u>**C**ATGGACGATATTTCTTTATCAT	pCOLADuet-1
	AGTCAA<u>CTGCAG</u>ATATTACTCATCTTTGTTCCTTATAAT	
ZQ56	TAATTA<u>GAATTC</u>**C**GTGATAATAAATAACAGAAATGAATCTCAACC	pACYCDuet-1
	TCTACA<u>CTGCAG</u>ATTCATTATTGATTCCTCAATG	

ZQ61	TAATTA<u>GAATTC</u>**C**GTGATAATAAATAACAGAAATGAATCTCAACC	pACYCDuet-1
	TTAATT<u>CTGCAG</u>CAGTTAATTTTTATCGTTTAAACTT	
ZQ62	TAATTA<u>GAATTC</u>CGTGATAATAAATAACAGAAATGAATCTCAACC	pACYCDuet-1
	TCCTTT<u>CTGCAG</u>GTTACTAAATACGAGTGTCTAACCACT	
ZQ65	TAATTA<u>AGATCT</u>**A**GTGATAATAAATAACAGAAATGAATCT	pACYCDuet-1
	TCTACA<u>GGTACC</u>ATTCATTATTGATTCCTCAATG	
	AAAAAT<u>GAATTC</u>**C**ATGAATAATCATCCAGAAGTAAA	
	TTAATT<u>CTGCAG</u>CAGTTAATTTTTATCGTTTAAACTT	
ZQ76	TAGTTC<u>GAATTC</u>**C**ATGAAATATGCCTTTATTACCGG	pCDFDuet-1
	ATGACA<u>CTGCAG</u>TTATTATAATATTGCGACCACTC	
ZQ78	AGCCGA<u>CATATG</u>AAGATTTACGGAATTT	pCDF Duet-1
	CCGAAT<u>GGTACC</u>TTATAAAAGCTCTTCGTACG	
ZQ80	No primers, more information in method section.	pCOLADuet-1
ZQ81	AAAAAT<u>GAATTC</u>**C**ATGAATAATCATCCAGAAGTAAA	pACYC Duet-1
	TTAATT<u>CTGCAG</u>CAGTTAATTTTTATCGTTTAAACTT	
ZQ82	AAAAATGAATTCCATGAATAATCATCCAGAAGTAAA	pACYC Duet-1
	TTAATTCTGCAGCAGTTAATTTTTATCGTTTAAACTT	
pDS_plu4186	GGTCAA<u>GCATGC</u>GTGGGTGATAGCTATATTAATATCG	pDS132
	TTCACT<u>GAGCTC</u>CCCAATCTGAAACTTGTATCAT	
pDS_plu4188	AACTTA<u>GCATGC</u>CTCCGCAATCTATTGCTAAC	pDS132
	CGAGAT<u>GAGCTC</u>CCAGTGGCAAACCACTC	
pDS_plu4192	[5]	pDS132

primers for verification PCR	
vP_plu4186_Fw	GTGATTCAGTAAAAGTCATTTATAATG
vP_plu4186_Rv	GCCAGTTAATACCTCAGCAG
vP_plu4188_Fw	GCGCTTTAGTAATCAAGGTC
vP_plu4188_Rv	GCTGAGAATTGATTTTAATTACG
vP_plu4192_Fw	TACCTTATGGATTTCAAGATGC
vP_plu4192_Rv	AACTCTTTGTTATTGCCATCAC
pDS132fw	GATCGATCCTCTAGAGTCGACCT
pDS132rv	ACATGTGGAATTGTGAGCGG

Table S2. Plasmids used in this work. *: gene *antI* with selected mutation. **: codon optimized for *E. coli* expression.

plasmid	genotype	reference
pJET1.2/blunt	pMB1ori, Ap^r	Fermentas
pSU18	P15A ori, Cm^r, *lacZ*promoter	[14]
pACYCDuet-1	P15A ori, Cm^r, T7*lac*promoter	Novagen
pCOLADuet-1	ColAori, Km^r, T7*lac* promoter	Novagen
pCDF Duet-1	CDF ori, Sm^r, T7*lac* promoter	Novagen
pDS132	*pir*dependent, Cm^r, *oriT, oriV, sacB*	[15]
pSUsfp	P15A ori, Cm^r, *sfp, lacZ*promoter	this work
pSUmtaA	P15A ori, Cm^r, *mtaA, lacZ*promoter	[1]
ZQ1	pMB1ori, (Ap^r), *antDEF* with native promoter	thiswork
ZQ6	P15A ori, Cm^r, *antB, lacZ*promoter	this work
ZQ9	pMB1ori, (Ap^r), *antDEFG* with native promoter	this work
ZQ10	P15A ori, Cm^r, *antABC, lacZ*promoter	this work
ZQ11	P15A ori, Cm^r, *antAB, lacZ*promoter	this work
ZQ12	pMB1ori, (Ap^r), *antDEFGH* with native promoter	this work
ZQ13	pMB1ori, (Ap^r), *antDEFGHI* with native promoter	thiswork
ZQ16	P15A ori, Cm^r, *antBC, lacZ*promoter	this work
ZQ32	ColAori, Km^r, T7*lac* promoter, *antD, antE*	this work
ZQ37	ColA ori, Km^r, T7*lac* promoter, *plu2834*	this work
ZQ40	ColA ori, Km^r, T7*lac* promoter, *antI*	this work
ZQ40_S245A	ColA ori, Km^r, T7*lac* promoter, *antI**	this work
ZQ40_D326A	ColA ori, Km^r, T7*lac* promoter, *antI**	this work
ZQ40_D327A	ColA ori, Km^r, T7*lac* promoter, *antI**	this work
ZQ40_H355A	ColA ori, Km^r, T7*lac* promoter, *antI**	thiswork
ZQ46	ColAori, Km^r, T7*lac* promoter, *antF*	this work
ZQ47	ColAori, Km^r, T7*lac* promoter, *antG*	this work
ZQ48	ColAori, Km^r, T7*lac* promoter, *antB*	this work
ZQ56	P15A ori, Cm^r, T7*lac* promoter, *antDEFG*	this work
ZQ61	P15A ori, Cm^r, T7*lac* promoter, *antDEF*	this work
ZQ62	P15A ori, Cm^r, T7*lac* promoter, *antDEFGH*	this work
ZQ65	P15A ori, Cm^r, T7*lac* promoter, *antF, antDEFG*	this work
ZQ76	CDF ori, Sm^r, T7*lac* promoter, *antABC*	this work
ZQ78	CDF ori, Sm^r, T7*lac* promoter, *sfp*	this work
ZQ80	ColA ori, Km^r, T7*lac* promoter, *actVI-ORF1***	this work
ZQ81	P15A ori, Cm^r, T7*lac* promoter, *antDE, remC***	this work
ZQ82	P15A ori, Cm^r, T7*lac* promoter, *antDE, actI-ORF3***	this work
pET-MatB	pBR322 ori, Ap^r, T7*lac* promoter, *matB*	Tang lab[2]
pDS_plu4186	*pir* dependent, Cm^r, *oriT, oriV, sacB,* partial *plu4186*	thiswork
pDS_4188	*pir* dependent, Cm^r, *oriT, oriV, sacB,* partial *plu4188*	this work
pDS_plu4192	*pir* dependent, Cm^r, *oriT, oriV, sacB,* partial *plu4192*	this work

strain	genotype	reference
Table S3. *E. coli* strains (ES) strains used in this work.		
E. coli DH10B		[16]
E. coli BL21 (DE3)		Novagen
E. coli s17-1 λpir		[17]
ES1	DH10B::ZQ9, pSUsfp	this work
ES2	DH10B::ZQ9, pSUmtaA	this work
ES3	DH10B::ZQ1, pSUsfp	this work
ES4	DH10B::ZQ1, pSUmtaA	this work
ES5	DH10B::ZQ1, ZQ6	this work
ES6	DH10B::ZQ1	this work
ES7	DH10B::ZQ9	this work
ES8	DH10B::ZQ9, ZQ6	this work
ES9	DH10B::ZQ12, ZQ16	this work
ES10	DH10B::ZQ13, ZQ16	this work
ES11	DH10B::ZQ9, ZQ16	this work
ES12	DH10B::ZQ9, ZQ11	this work
ES13	DH10B::ZQ9, ZQ10	this work
ES14	DH10B::ZQ12, ZQ11	this work
ES15	DH10B::ZQ13, ZQ11	this work
ES16	DH10B::ZQ12, ZQ10	this work
ES17	DH10B::ZQ13, ZQ10	this work
ES18	BL21 (DE3)::ZQ37	this work
ES19	BL21 (DE3)::ZQ47	this work
ES20	BL21 (DE3)::ZQ46	this work
ES21	BL21 (DE3)::ZQ32	this work
ES22	BL21 (DE3)::ZQ48	thiswork
ES23	BL21 (DE3)::ZQ65, ZQ48	this work
ES24	BL21 (DE3)::ZQ56, ZQ48	this work
ES25	BL21 (DE3)::ZQ61, ZQ48	this work
ES26	BL21 (DE3)::ZQ56	this work
ES36	BL21 (DE3)::ZQ40	this work
ES41	BL21 (DE3)::pET-MatB	Tang Lab[2]
ES53	BL21 (DE3)::ZQ62, ZQ76, ZQ40	this work
ES59	BL21 (DE3)::ZQ62, ZQ76, ZQ40_S245A	this work
ES61	BL21 (DE3)::ZQ62, ZQ76, ZQ40_D326A	this work
ES62	BL21 (DE3)::ZQ62, ZQ76, ZQ40_D327A	this work
ES64	BL21 (DE3)::ZQ62, ZQ76, ZQ40_H355A	this work
ES65	BL21 (DE3)::ZQ62, ZQ76, ZQ80	this work
ES67	BL21 (DE3)::ZQ81, ZQ78	this work

195

16

ES68	BL21 (DE3)::ZQ82, ZQ78	this work
ES27	s17-1 λpir::pDS_plu4186	this work
ES28	s17-1 λpir::pDS_4188	this work
ES29	s17-1 λ pir::pDS_4192	this work

Table S4. *E. coli* strains with coexpressed proteins and major polyketide products.

Strain	Proteins	Major products
ES1	AntDEFG, Sfp	no polyketide
ES2	AntDEFG, MtaA	no polyketide
ES3	AntDEF, Sfp	no polyketide
ES4	AntDEF, MtaA	no polyketide
ES5	AntBDEF	no polyketide
ES6	AntDEF	no polyketide
ES7	AntDEFG	no polyketide
ES8	AntBDEFG	**4, 5**
ES9	AntB-H	**4, 5**
ES10	AntB-I	**4, 5**
ES11	AntB-G	**4, 5**
ES12	AntABDEFG	**6**
ES13	AntA-G	**6**
ES14	AntABDEFGH	**8**
ES15	AntABDEFGHI	**8**
ES16	AntA-H	**11**, **9**, **14**, **12**
ES17	AntA-I	**1**, **11**, **9**

	[M+H]⁺	[M-H]⁻	RT^A / min	RT^B /min	identified with
SEK4 (**4**)	301^C		5.3	6.5	authentic standard (Tsai lab), UV
SEK4b (**5**)	319		5.4	7.1	authentic standard (Tsai lab), UV
mutactin (**6**)	303		6.3	10.2	authentic standard (Δ*antC* mutant), UV
SEK34 (**7**)	285^D		6.0	9.5	UV
SEK34b (**8**)	285		6.8		UV
DMAC (**9**)		297	7.8		isolationand NMR, UV
(**11**)	302		6.2		UV, labeled media
(**14**)	240		8.2		labeled media
utahmycin A (**12**)	254		8.4		authenticstandard (Brady lab)[18]
AQ-256 (**1**)		255	8.7		standard (TT01 wild type), UV

Table S5. Identification of polyketides using HPLC-UV-MS or NMR.

A: HPLC gradient A; B: HPLC gradient B (described above); C: SEK4-H_2O; D: SEK34-H_2O

TT01 strain	Genotype	Major products	Reference
Table S6. *P.luminescens* TT01 wild type and mutants with genotypes and major products			
TT01	*P. luminescence* wild type	**1**, **2**, **3**	[19]
TT01 *antI*::cat	*antI* inactivated by plasmid insertion	**11**	this work
TT01 *antG*::cat	*antG* inactivated by plasmid insertion	nopolyketide	this work
TT01 Δ*antC*	*antC* deletion	**7**, **8**	this work
TT01 Δ*antH*	*antH* deletion	**6**	[5]
TT01 *antD*::cat	*antD* inactivated by plasmid insertion	nopolyketide	[5]

Proteins	Size (kDa)	Function	Source organism	Affinity tag	Strain	Purified
Table S7. List of recombinant proteins, size, function, source organism, affinity tag and expression plasmid.						
His$_6$-AntD-AntE	47.5-40.7	KS-CLF	TT01	N-terminal His$_6$-tag at AntD / no tag at AntE	ES21	yes
His$_6$-*apo*-AntF	11.1	*apo*-ACP	TT01	N-terminal His$_6$-tag	ES20	yes
His$_6$-*holo*-AntF	11.4	*holo*-ACP	TT01	N-terminal His$_6$-tag	ES23	yes
His$_6$-AntG	60.2	CoA ligase	TT01	N-terminal His$_6$-tag	ES19	yes
His$_6$-Plu2834	35.0	MCAT	TT01	N-terminal His$_6$-tag	ES18	yes
His$_6$-AntI	46.0	Hydrolyse	TT01	N-terminal His$_6$-tag	ES36	yes
His$_6$-MatB	60.0	MatB	*R. trifolii*	N-terminal His$_6$-tag	ES41	yes
His$_6$-AntB	29.5	PPTase	TT01	N-terminal His$_6$-tag	ES22	no
Cpin_1856	10.2	*apo*-ACP	*C. pinensis*	no tag	provided by Dr. S. Fuchs	
Sfp	26.0	PPTase	*B. subtilis*	no tag	provided by Dr. C. Kegler	
MtaA	31.6	PPTase	*S. aurantiaca*	no tag	provided by Dr. C. Kegler	

assay	a	b	c	d	e	f	g	h	i	j
Table S8. Summary of components used for *in vitro* minimal PKS assays and results from HPLC-MS.										
His$_6$-AntD-AntE (0.1 µM)	-	✓	✓	✓	✓	✓	✓	-	✓	✓
His$_6$-Plu2834 (0.1 µM)	-	✓	✓	✓	✓	✓	-	✓	✓	✓
His$_6$-AntG (0.1 µM)	-	✓	✓	✓	-	-	-	-	-	-
His$_6$-*holo*-AntF (1 µM)	✓	-	✓	✓	✓	-	✓	✓	✓	✓
His$_6$-*apo*-AntF (1 µM)	-	✓	-	-	-	-	-	-	-	-
malonyl-CoA (4 mM)	-	✓	✓	✓	✓	✓	✓	✓	-	✓
acetyl-CoA (4 mM)	-	✓	✓	-	-	-	-	-	-	-
Tris-HCl Buffer (50 mM, pH = 7.5)	✓	✓	✓	✓	✓	✓	✓	✓	✓	-
total volume / µL	50	50	50	50	50	50	50	50	50	50
SEK4 (**4**)/SEK4b (**5**)	-	-	✓	✓	✓	-	-	-	-	-

Table S9. Accession numbers for the 44 Hydrolases which were used for the generation of the multiple sequence alignment (Figure S9).

	Protein	Organism	Accession number
1	A1O7_05046	*Cladophialophora yegresii* CBS 114405	XP_007757246
2	G647_03620	*Cladophialophora carrionii* CBS 160.54	XP_008726187
3	A1O5_07014	*Cladophialophora psammophila* CBS 110553	XP_007745793
4	A1O9_08991	*Exophiala aquamarina* CBS 119918	KEF54549
5	A1O1_02354	*Capronia coronata* CBS 617.96	XP_007721455
6	A1O3_00818	*Capronia epimyces* CBS 606.96	XP_007729158
7	WdYG1	*Wangiella dermatitidis*	AY667610
8	HMPREF1541_01613	*Cyphellophora europaea* CBS 101466	XP_008714193
9	M438DRAFT_409345	*Aureobasidium pullulans* EXF-150	KEQ79600
10	W97_07417	*Coniosporium apollinis* CBS 100218	XP_007783237
11	EPUS_03003	*Endocarpon pusillum* Z 07020	XP_007801136
12	AO090005000332	*Aspergillus oryzae* RIB 40	BAE55407
13	Ao3042_06339	*Aspergillus oryzae* 3.042	EIT77502
14	AOR_1_584174	*Aspergillus oryzae* RIB 40	XP_001817409
15	AFLA_075640	*Aspergillus flavus* NRRL 3357	XP_002372481
16	NFIA_092970	*Neosartorya fischeri* NRRL 181	XP_001261232
17	Ayg1p	*Aspergillus fumigatus*	FJ406474
18	ACLA_076460	*Aspergillus clavatus* NRRL 1	XP_001276032
19	PDIG_54110	*Penicillium digitatum* PHI 26	EKV10923
20	PDE_09452	*Penicillium oxalicum* 114-2	EPS34488
21	EURHEDRAFT_548125	*Aspergillus ruber* CBS 135680	EYE98107
22	ANI_1_740124	*Aspergillus niger* CBS 513.88	XP_001401158
23	G205_09318	*Arthrobacter* sp. SJCon	ELT44872
24	WP_035752425	*Arthrobacter* sp. SJCon	WP_035752425
25	DHPON hydrolase	*Arthrobacter nicotinovorans*	2JBW
26	EP51_26660	*Rhodococcus opacus*	AII08016
27	WP_037230080	*Rhodococcus wratislaviensis*	WP_037230080
28	WP_012689950	*Rhodococcus opacus*	WP_012689950
29	WP_005250115	*Rhodococcus opacus*	WP_005250115
30	WP_032407992	*Rhodococcus fascians*	WP_032407992
31	WP_037817419	*Streptomyces* sp. NRRL F-3213	WP_037817419
32	CF8_0352	*Nocardioides* sp. CF8	EON25515
33	WP_036491922	*Nocardioides* sp. *CF8*	WP_036491922
34	WP_023044674	*Photorhabdus temperata*	WP_023044674
35	WP_036840784	*Photorhabdus temperata*	WP_036840784
36	B738_17427	*Photorhabdus temperate* subsp. *temperata* M1021	EQB99488
37	WP_036844637	*Photorhabdus temperata*	WP_036844637
38	WP_036779573	*Photorhabdus luminescens*	WP_036779573

39	AntI	*Photorhabdus luminescens*	WP_011148290
40	WP_015179936	*Crinalium epipsammum*	WP_015179936
41	WP_015205658	*Cylindrospermum stagnale*	WP_015205658
42	WP_006540205	*Frankia* sp. EUN1f	WP_006540205
43	WP_037569324	*Streptacidiphilus oryzae*	WP_037569324
44	WP_009739949	*Frankia* sp. QA3	WP_009739949

Table S10. Ketosynthases used for the phylogenetic tree. The sequences are ordered clockwise according to their location in the respective branches.

	Protein	**Organism**	**Accession number**
1	AuaD	*S. aurantiaca*	CAL48956.1
	KS type I PKS		
2	Plu1885	*P. luminescens*	NP_929153
3	NanA8	*S. nanchangensis*	AAP42874
4	EryAII	*S. erythraea*	YP_001102990
5	TylGI KSQ	*S. fradiae*	AAB66504
6	MerA	*S. violaceusniger*	ABJ97437
7	TamAI	*S.* sp. 3079	ADC79637
8	OleAI KSQ	*S. antibioticus*	AAF82408
9	HedT	*S. griseoruber*	AAP85336
	Closest BLAST-P hits for XclF[20]		
10	3-Oxoacyl-ACP synthase	*R. blandensis*	WP_008043745.1
11	3-Oxoacyl-ACP synthase	*X. nematophila*	YP_003714026.1
12	3-Oxoacyl-ACP synthase	*X. nematophila*	WP_010848687.1
13	3-Oxoacyl-ACP synthase	*M.* sp. PE36	WP_006034384.1
14	3-Oxoacyl-ACP synthase	*P. profundum*	YP_132684.1
15	3-Oxoacyl-ACP synthase	*P. damselae*	WP_005305524.1
16	3-Oxoacyl-ACP synthase	*P.* sp. AK15	WP_007465048.1
17	3-Oxoacyl-ACP synthase	*P. leiognathi*	WP_008989540.1
18	3-Oxoacyl-ACP synthase	*P.* sp. SKA34	WP_006644045.1
19	3-Oxoacyl-ACP synthase	*P. angustum*	WP_005364526.1
	FabB		
20	cpin1855	*C. pinensis*	YP_003121552
21	Dfer_1997	*D. fermentans*	YP_003086385
22	FabB	*A. pleuropneumoniae*	ZP_00134992
23	FabB	*C.* sp. 30_2	ZP_04562837
24	NP_416826	*E. coli*	NP_416826
25	FabB	*S. boydii*	YP_001881145
26	AuaC	*S. aurantiaca*	CAL48955.1
	FabF		
27	FabF	*S. avermitilis*	BAC70003
28	FabF	*T. thermophilus*	YP_143679
29	FabF	*N. punctiforme*	YP_001867862
30	NP_344945	*S. pneumoniae*	NP_344945
31	FabF	*B. subtilis*	NP_389016
32	NP_645683	*S. aureus*	NP_645683
33	FabF	*P. luminescens*	NP_930065
34	FabF	*E. albertii*	ZP_02902779.1
35	FabF	*E. coli*	NP_287229
36	NP_415613	*E. coli*	NP_415613
	Type II PKS KS a		
37	Lactobacillus oris KSa	*L. oris*	WP_003712532.1

38	Streptococcus GMD4S KSa	*S. GMD4S*	WP_000883312.1
39	RemA	*S. resistomycificus*	WP_030043016.1
40	AntD	*P. luminescens*	NP_931374
41	Hoeflea KSa	*Hoeflea* sp. IMCC20628	WP_047031056.1
42	EncA	*S. maritimus*	AAF81728
43	ActiB	*S. coelicolor* A3(2)	SCO5088
44	NcnA	*S. arenae*	AAD20267
45	TcmK	*S. davawensis*	CCK26894
46	SimA1	*S. antibioticus*	AAK06784
47	OvmP	*S. antibioticus*	CAG14965.1
48	Pd2A	*S.* sp. WP 4669	AAO65362.1
	Type II PKS KS b		
49	Lactobacillus oris KSb	*L. oris*	WP_003712534.1
50	Streptococcus GMD4S KSb	*S. GMD4S*	WP_000213426.1
51	AntE	*P. luminescens*	WP_011148294.1
52	Hoeflea KSb	*Hoeflea* sp. IMCC20628	WP_047031057.1
53	RemB	*S. resistomycificus*	WP_030043017.1
54	SimA2	*S. antibioticus*	AF324838_4
55	OvmK	*S. antibioticus*	WP_030791003.1
56	Pd2B	*S.* sp. WP 4669	AAO65363.1
57	TcmL	*S. glaucescens*	AAA67516
58	EncB	*S. maritimus*	AAF81729
59	ActIA	*S. coelicolor* A3(2)	SCO5087
60	NcnB	*S. arenae*	AAD20268
	DarB		
61	O3I_37171	*N. brasiliensis*	ZP_09843377
62	M446_0174	*M.* sp. 4-46	YP_001767187
63	cpin6850	*C. pinensis*	YP_003126452
64	BFO_3187	*T. forsythia*	YP_005015826
65	NiasoDRAFT_0547	*N. soli*	ZP_09632794
66	Mucpa_6793	*M. paludis*	ZP_09618305
67	Oweho_0889	*O. hongkongensis*	YP_004988545
68	CHU_0390	*C. hutchinsonii*	YP_677020
69	Fluta_1447	*F. taffensis*	YP_004344279
70	Dfer_5797	*D. fermentans*	YP_003090150
71	BZARG_2045	*B. argentinensis*	ZP_08820341
72	Lacal_2074	*L.* sp. 5H-3-7-4	YP_004580348
73	Aeqsu_0932	*A. sublithincola*	YP_006417450
74	Zobellia_2074	*Z. galactanivorans*	YP_004736513
75	Lbys_1508	*L. byssophila*	YP_003997574
76	HMPREF0204_10987	*C. gleum*	ZP_07085127
77	PMI13_02465	*C.* sp. CF314	ZP_10726507
78	HMPREF0156_01383	*B.* taxon 274 str. F0058	ZP_06983320
79	HMPREF9071_0527	*C.* taxon 338 str. F0234	ZP_08201061
80	CAPGI0001_0843	*C. gingivalis*	ZP_04056582
81	HMPREF1154_2288	*C.* sp. CM59	ZP_10880679
82	HMPREF1320_1701	*C.* taxon 335 str. F0486	EJF37460
83	HMPREF1321_1154	*C.* taxon 412 str. F0487	ZP_10366882
84	CAPSP0001_1216	*C. sputigena*	ZP_03390203
85	Coch_0547	*C. ochracea*	YP_003140666
86	HMPREF1319_0374	*C. ochracea*	EJF43732
87	HMPREF1977_1456	*C. ochracea*	ZP_07866642
88	Weevi_1554	*W. virosa*	YP_004238832.1
89	HMPREF9716_01579	*M. odoratimimus*	EKB07937
90	Myrod_1723	*M. odoratus*	ZP_09672239
91	HMPREF9711_01694	*M. odoratimimus*	EKB04829
92	HMPREF9712_01161	*M. odoratimimus*	ZP_09523568
93	Fcol_11845	*F. columnare*	YP_004942963
94	FP2279	*F. psychrophilum*	YP_001297136
95	PMI10_02641	*F.* sp. CF136	ZP_10730768
96	FF52_12311	*F.* sp. F52	ZP_10481912
97	Fjoh_1102	*F. johnsoniae*	YP_001193454

98	FJSC11DRAFT_3961	*F.* sp. JSC-11	ZP_08987753
99	MICAG_1820011	*M. aeruginosa*	CCI22605
100	DP1817	*D. psychrophila*	YP_065553
101	DaAHT2_1139	*D. alkaliphilus*	YP_003690456
102	MldDRAFT_4065	delta proteobacterium MLMS-1	ZP_01289639
103	CBGD1_514	*S. gotlandica*	ZP_05070248
104	SMGD1_1386	*S. gotlandica*	EHP29910
105	Sdel_2118	*S. deleyianum*	YP_003305165
106	Sulba_2257	*S. barnesii*	YP_006405107
107	Arnit_2310	*A. nitrofigilis*	YP_003656468
108	HMPREF9401_0244	*A. butzleri*	ZP_07890833
109	Hbal_2902	*H. baltica*	YP_003061270
110	ParcA3_010100003428	*P. arctica*	ZP_10280196
111	PspoU_010100018642	*P. spongiae*	ZP_10300425
112	PSJM300_17945	*P. stutzeri*	AFN79642
113	MDS_0597	*P. mendocina*	YP_004378380
114	Psefu_0435	*P. fulva*	YP_004472512
115	Plu2164	*P. luminescens*	NP_929424
116	PA-RVA6-3077	*P. asymbiotica*	CAR66906
117	PAU_02401	*P. asymbiotica*	YP_003041237
118	PchlO6_4243	*P. chlororaphis*	ZP_10172862
119	DarB	*P. chlororaphis*	AAN18032
120	Pchl3084_3967	*P. chlororaphis*	EJL05977
121	PMI20_00702	*P.* sp. GM17	ZP_10707840
122	Daro_2368	*D. aromatica*	YP_285574
123	azo0292 DarB	*A.* sp. BH72	YP_931796
124	Rfer_3974	*R. ferrireducens*	YP_525203
125	Slit_0359	*S. lithotrophicus*	YP_003522988
126	PMI12_02025	*V.* sp. CF313	ZP_10567997
127	Vapar_3389	*V. paradoxus*	YP_002945272
128	Varpa_2231	*V. paradoxus*	YP_004154548
129	COI_2002	*M. haemolytica*	ZP_05992665
130	COK_0379	*M. haemolytica*	ZP_05988513
131	HMPREF9417_0595	*H. parainfluenzae*	ZP_08147854
132	HMPREF9952_1824	*H. pittmaniae*	ZP_08755481
133	HMPREF9064_0174	*A.segnis*	ZP_07888807
134	ATCC33389_0196	*A. aphrophilus*	EGY32238
135	NT05HA_1737	*A. aphrophilus*	YP_003008155
136	HMPREF9335_01583	*A. aphrophilus*	EHB89432
137	GCWU000324_02596	*K. oralis*	ZP_04603113
138	EIKCOROL_00456	*E. corrodens*	ZP_03712789
139	HMPREF9371_1043	*N. shayeganii*	ZP_08886538
140	HMPREF9370_1914	*N. wadsworthii*	ZP_08940206
141	NEIFLAOT_02523	*N. flavescens*	ZP_03720660
142	HMPREF0604_01363	*N. mucosa*	ZP_07993739
143	NEIFL0001_0036	*N. flavescens*	ZP_04757628
144	NEISUBOT_03200	*N. subflava*	ZP_05983976
145	NEISICOT_02133	*N. sicca*	ZP_05318975
146	HMPREF9418_1128	*N. macacae*	ZP_08684521
147	HMPREF1051_1749	*N. sicca*	EIG27057
148	HMPREF1028_00835	*N.* sp. GT4A_CT1	ZP_08888860
149	HMPREF9016_01947	*N.* taxon 014 str. F0314	ZP_06980826
	PpyS homologues[21]		
150	3-Oxoacyl-ACP synthase	*B.* sp. CCGE1001	YP_004230959
151	3-Oxoacyl-ACP synthase	*B. phenoliruptrix* BR3459a	YP_006793509
152	3-Oxoacyl-ACP synthase	*B.* sp. CCGE1003	YP_003910175
153	3-Oxoacyl-ACP synthase	*B. phytofirmans* PsJN	YP_001889944
154	3-Oxoacyl-ACP synthase	*C. fritschii*	WP_016876568
155	3-Oxoacyl-ACP synthase	*A.* sp. PCC 7108	WP_016949109
156	PpyS	*P. luminescens* TT01	AGO97060
157	PyrS	*P.* sp. GM30	WP_007967127
158	3-Oxoacyl-ACP synthase	*X. bovienii*	WP_012989958.1

159	3-Oxoacyl-ACP synthase	*X. nematophila*	WP_013184973.1
	Closest BLAST-P hits for XclC		
160	3-Oxoacyl-ACP synthase	*C. acetobutylicum*	NP_347450.1
161	3-Oxoacyl-ACP synthase	*P. lactis*	WP_007130623.1
162	3-Oxoacyl-ACP synthase	*B. thuringiensis*	YP_006930640.1
163	3-Oxoacyl-ACP synthase	*B.* sp. 1NLA3E	YP_007911827.1
164	3-Oxoacyl-ACP synthase	*O. scapharcae*	WP_010098042.1
165	3-Oxoacyl-ACP synthase	*P. polymyxa*	YP_003947618.1
166	3-Oxoacyl-ACP synthase	*P. polymyxa*	YP_003871436.1
167	3-Oxoacyl-ACP synthase	*P.* sp. Aloe-11	WP_007431139.1
168	3-Oxoacyl-ACP synthase	*P. terrae*	YP_005077926.1
169	3-Oxoacyl-ACP synthase	*P. peoriae*	WP_010345468.1
	FabH		
170	CorB	*C. coralloides*	ADI59524
171	Myxopyronin ketosynthase	*M. fulvus*	AGS77282
172	FabHB	*B. subtilis*	NP_388898
173	FabH	*N. punctiforme*	YP_001865657
174	3-oxoacyl-ACP synthase	*B. subtilis*	NP_389015.1
175	FabH	*A. fabrum*	NP_354198
176	FabH	*P. luminescens*	NP_930069
177	FabH	*E. coli*	NP_287225
178	FabH	*S. griseus*	YP_001826619
179	FabH	*S. echinatus*	AAV84077
180	NP_626634	*S. coelicolor* A3(2)	NP_626634
181	FabH	*S. avermitilis*	BAC73499
182	Q54206	*S. glaucescens*	Q54206
183	FdmS	*S. griseus*	AAQ08929
184	CAM58805_S._sp._BenQ	*S.* sp. A2991200	CAM58805
185	ZhuH 1MZJ	*S.* sp. R1128	AAG30195
186	FrnI	*S. roseofulvus*	AAC18104
187	AlnI	*S.* sp CM020	ACI88883
188	OleA	*X. campestris* pv. *campestris*	3S21_A
	KS type III PKS		
189	Chs-like	*R. baltica*	NP_868579
190	BPS (PLN03172)	*H. androsaemum*	Q8SAS8
191	CHS H. (PLN03173)	*H. androsaemum*	Q9FUB7
192	CHS9	*M. sativa*	AAA02827
193	STS	*P. quinquefolia*	AAM21773
194	BAS	*R. palmatum*	AAK82824
195	bpsA	*B. subtilis* str. 168	NP_390087
196	MXAN_6639	*M. xanthus*	YP_634756
197	PKS10	*M. tuberculosis*	NP_216176
198	PKS11	*M. tuberculosis*	NP_216181
199	Cpz6 Capramyzin ketosynthase	*Streptomyces* sp. MK730–62F2	
200	Germicidin synthase	*Streptomyces coelicolor*	3V7I_A
201	RppA S	*S. antibioticus*	BAB91443
202	RppA	*S. avermitilis*	NP_828307
203	RppB	*S. antibioticus*	BAB91444
	Closest BLAST-P hits for XclB		
204	3-Oxoacyl-ACP synthase III	*B.* sp. EniD312	WP_009111263.1
205	3-Oxoacyl-ACP synthase III	*A. nasoniae*	CBA73264.1
206	3-Oxoacyl-ACP synthase III	*P. carotovorum*	WP_010301235.1
207	3-Oxoacyl-ACP synthase III	*P. pacifica*	WP_006975318.1
208	3-Oxoacyl-ACP synthase III	*C. stagnale*	YP_007317906.1
209	3-Oxoacyl-ACP synthase III	*N. punctiforme*	YP_001865628.1
210	3-Oxoacyl-ACP synthase III	*R.* sp. PCC 7116	YP_007056099
211	3-Oxoacyl-ACP synthase III	*S. cyanosphaera*	YP_007130807.1
212	3-Oxoacyl-ACP synthase III	*Calothrix* sp. PCC 6303	YP_007138278
213	3-Oxoacyl-ACP synthase III	*N. punctiforme*	YP_001868566.1
214	3-Oxoacyl-ACP synthase III	*R.*sp. PCC 7116	YP_007057764.1
	KS adjacent to XclA		

homologues

215	3-Oxoacyl-ACP synthase	*C.* sp. PCC 7822	YP_003899922.1
216	3-Oxoacyl-ACP synthase	*N. punctiforme*	YP_001865657.1
217	3-Oxoacyl-ACP synthase	*A. cylindrica*	YP_007155727.1
	ChlB6; CerJ; KSIII DpsC-like		
218	ChlB6	*S. antibioticus*	AAZ77679
219	CerJ	*S. tendae*	AEI91069
220	CosE	*S. olindensis*	ABC00733
221	DpsC	*S. peucetius*	AAA65208
222	AknE2	*S.* sp. SPB74	ZP_04991255.1
223	AknE2	*S. galilaeus*	AAF70109
224	BAB72048	*S. galilaeus*	BAB72048
225	PokM2	*S. diastatochromogenes*	ACN64832
226	CalO4	*S. aurantiaca*	ZP_01462124
227	FabH	*S. erythraea*	YP_001107471
228	NdasDRAFT_3133	*N. dassonvillei*	ZP_04334033.1
229	ChlB3	*S. antibioticus*	AAZ77676
230	CalO4	*M. echinospora*	AAM70354
231	AviN	*S. viridochromogenes*	AAK83178
232	PlaP2	*S.* sp. Tu6071	ABB69750
233	CouN2	*S. rishiriensis*	AAG29787
234	CloN2	*S. roseochromogenes*	AAN65231
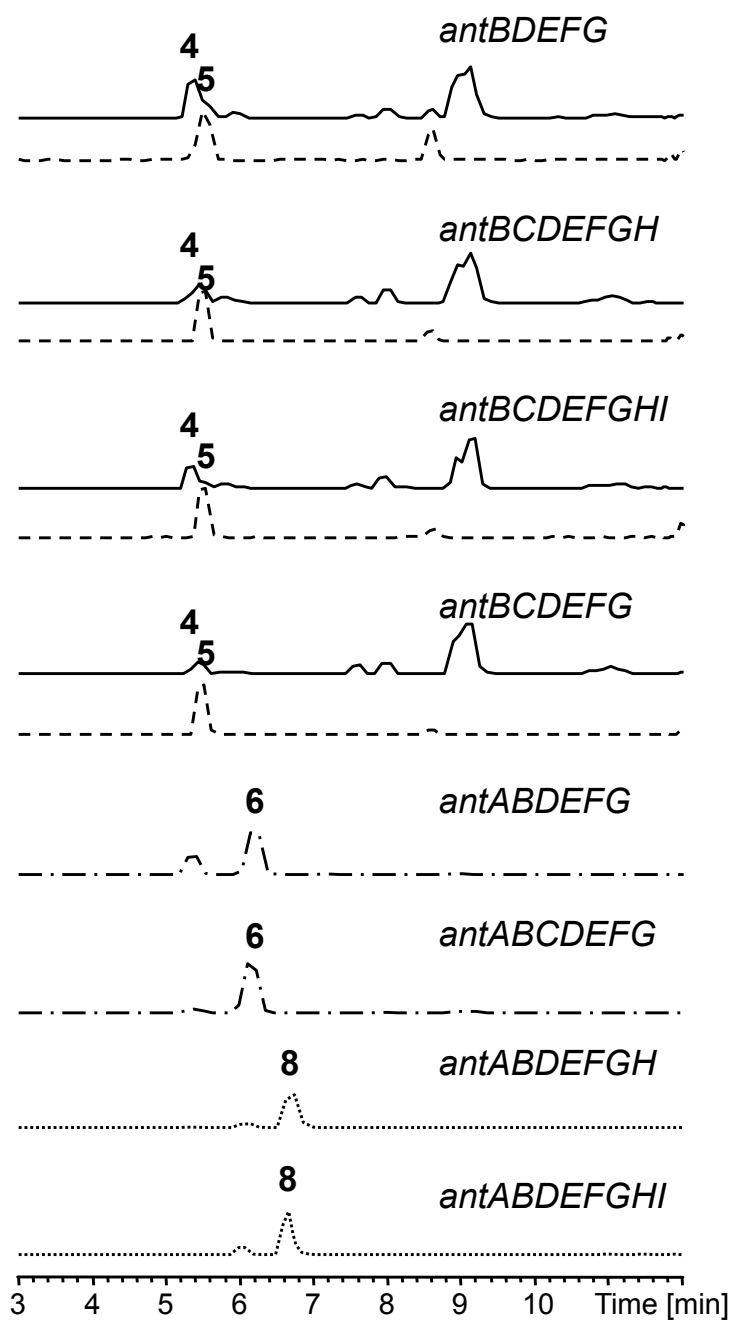
Figure S1. HPLC/MS of *E. coli* strains expressing different *ant* genes. Shown are EICs for **4** (continuous line, m/z 301 [M+H]$^+$), **5** (dashed line, m/z 319 [M+H]$^+$), **6** (broken line, m/z 303 [M+H]$^+$) and **8** (dotted line, m/z 285 [M+H]$^+$).
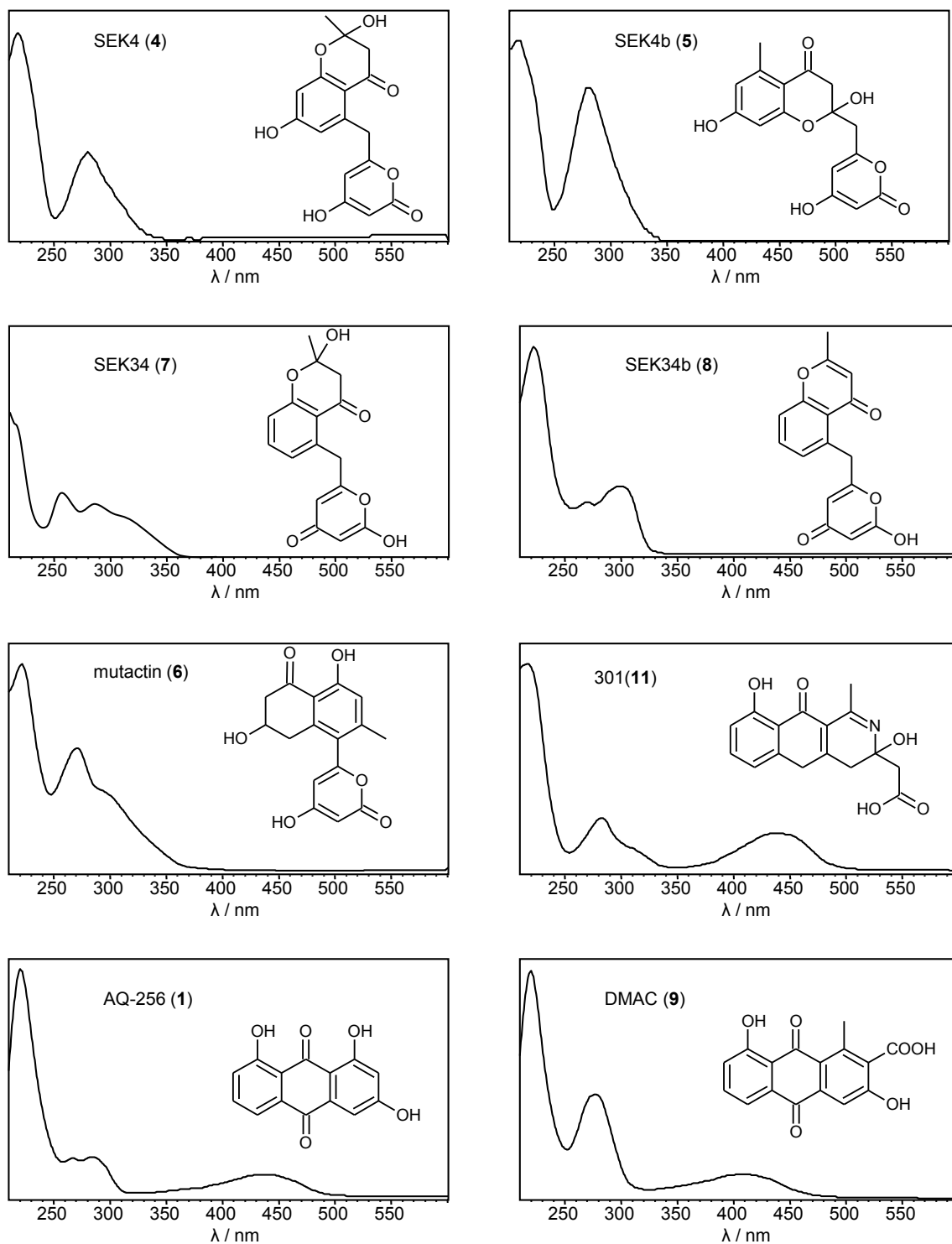
Figure S2. UV spectra (λ: 210-600 nm) and structures of major shunt products produced by *E. coli* DH10B with *ant* genes.
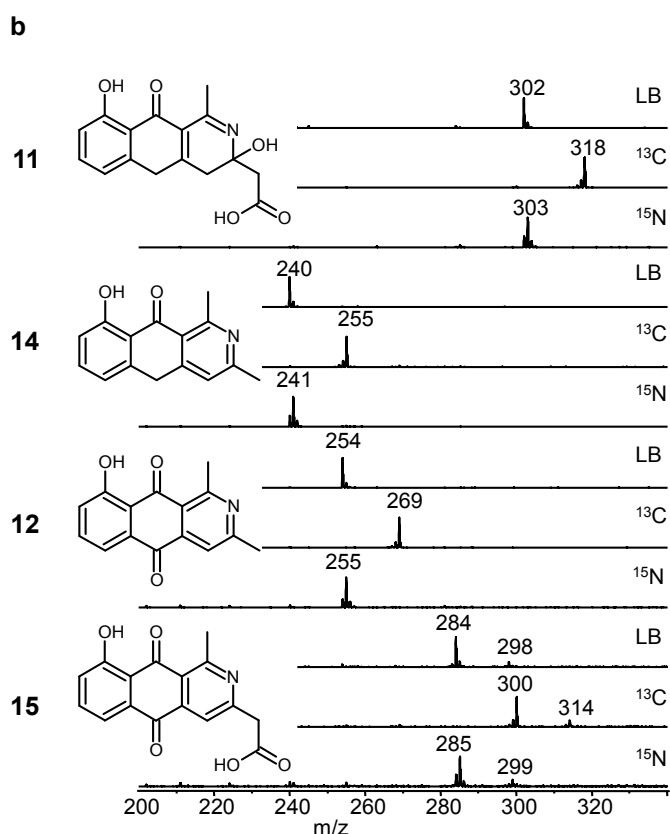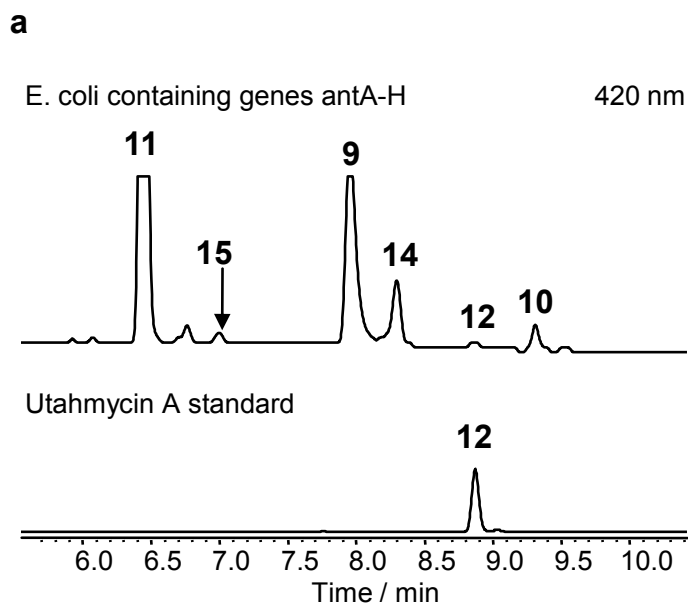
Figure S3. (**a**) HPLC-UV analyses (420 nm) of extracts of heterologous expression of genes antA-antH in *E. coli* DH10B (upper) and authentic Utahmycin A standard (down). (**b**) Determination of the number of carbon and nitrogen atoms for **11**, **14**, **12** and **15** by the way of cultivation of the *E. coli* strain with genes *antA-AntH* in standard growth medium (LB medium, ^{14}N , ^{12}C, ^{16}O, ^{1}H), ^{15}N labeled medium (^{15}N , ^{12}C, ^{16}O, ^{1}H), or ^{13}C labeled medium (^{14}N , ^{13}C, ^{16}O, ^{1}H).
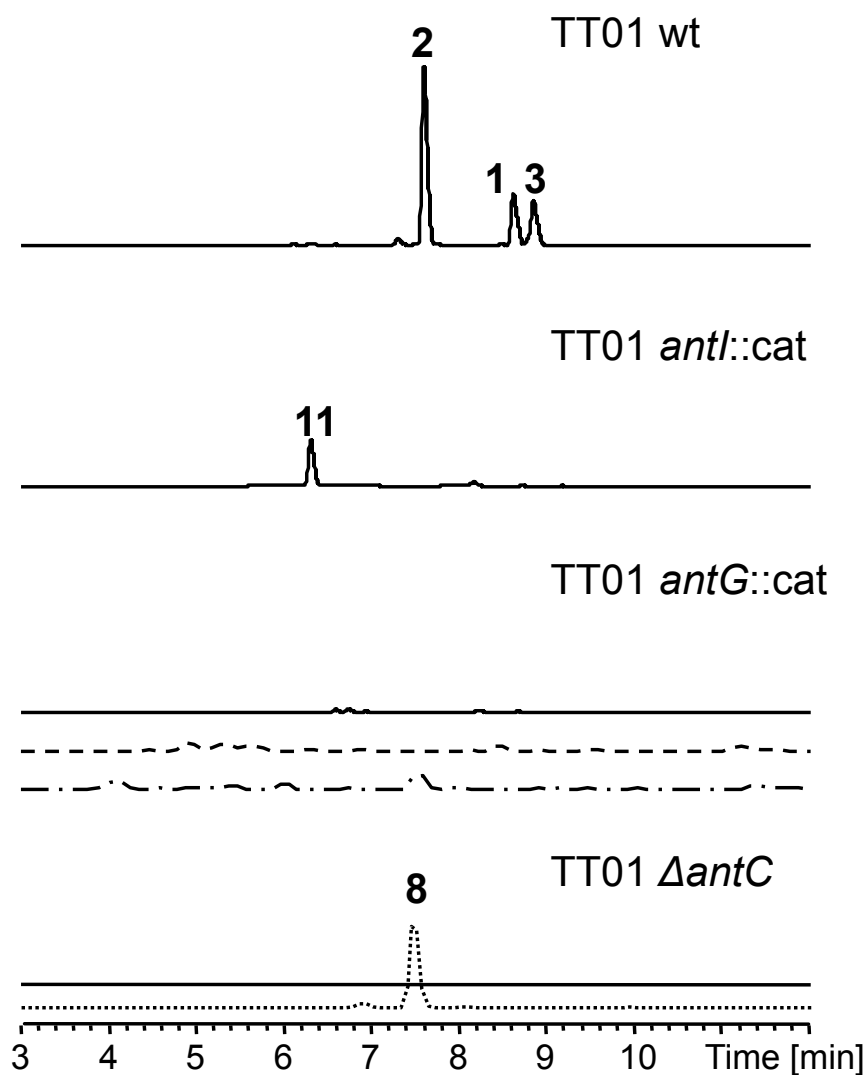
Figure S4. HPLC-UV analyses (420 nm) of extracts of *P. luminescens* TT01 wild type, mutant *antI*::cat, mutant *antG*::cat and mutant *ΔantC*. In all mutants, **1**-**3** were not produced. A new shunt product **11** could be identified in the mutant *antI*::cat. In the mutant *antG*::cat, searching with extracted ion chromatogram (EIC) showed there were also no shunt products **4** (dashed line, m/z 301 [M+H]$^+$) and **5** (broken line, m/z 319 [M+H]$^+$) produced. In the mutant *ΔantC*, shunt product **8** could be identified with EIC (dotted line, m/z 285 [M+H]$^+$). See also Table S6.

30
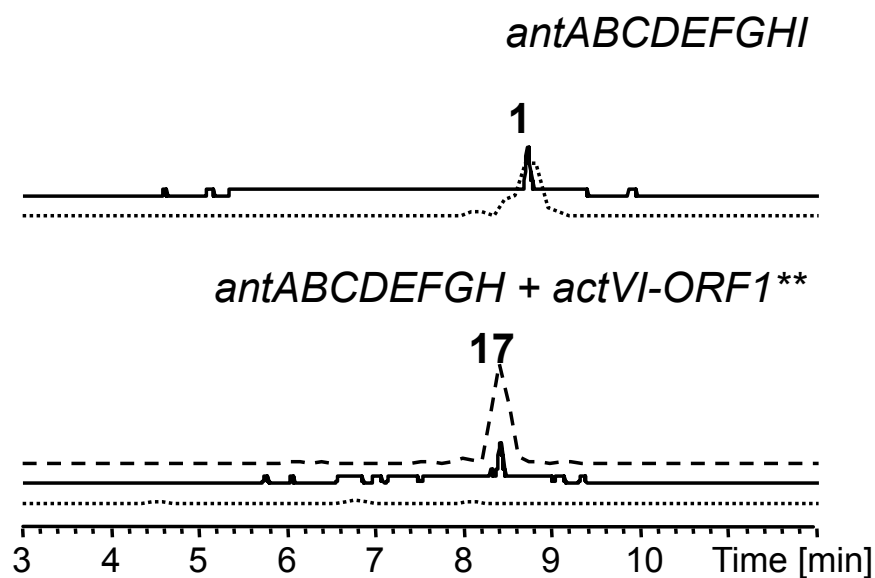
Figure S5. HPLC analyses of *E. coli* expressing *antABCDEFGHI* and *antABCDEFGH + actVI-ORF1*** (encoding RED1). Besides UV traces (420 nm), EIC for **1** (dotted line, *m/z* 255 [M-H]⁻) and **17** (dashed line, *m/z* 287 [M+H]⁺) are shown.
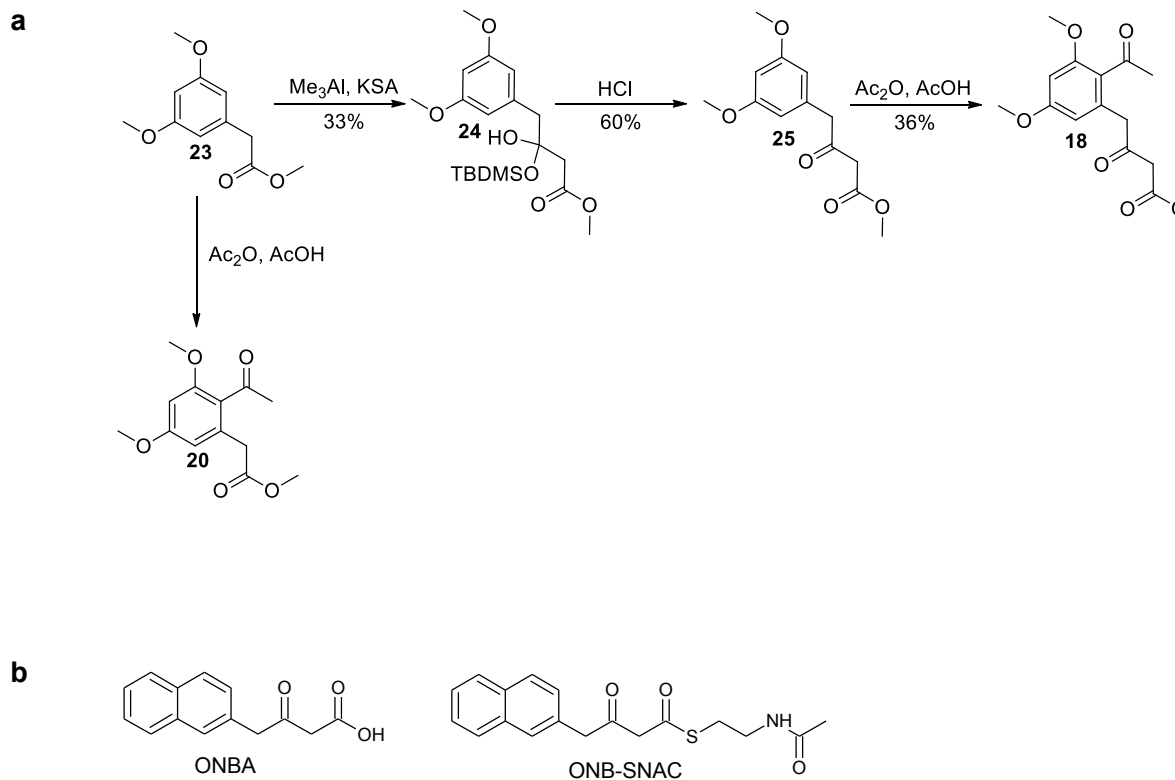
Figure S6. (**a**) Synthesis of model compounds **18** and **20**. (**b**) Structures of ONBA and ONB-SNAC. KSA: tert-butyl((1-methoxyvinyl)oxy)dimethylsilane; Me$_3$Al: trimethylaluminium; AcOH: acetic acid; Ac$_2$O: acetic anhydride.
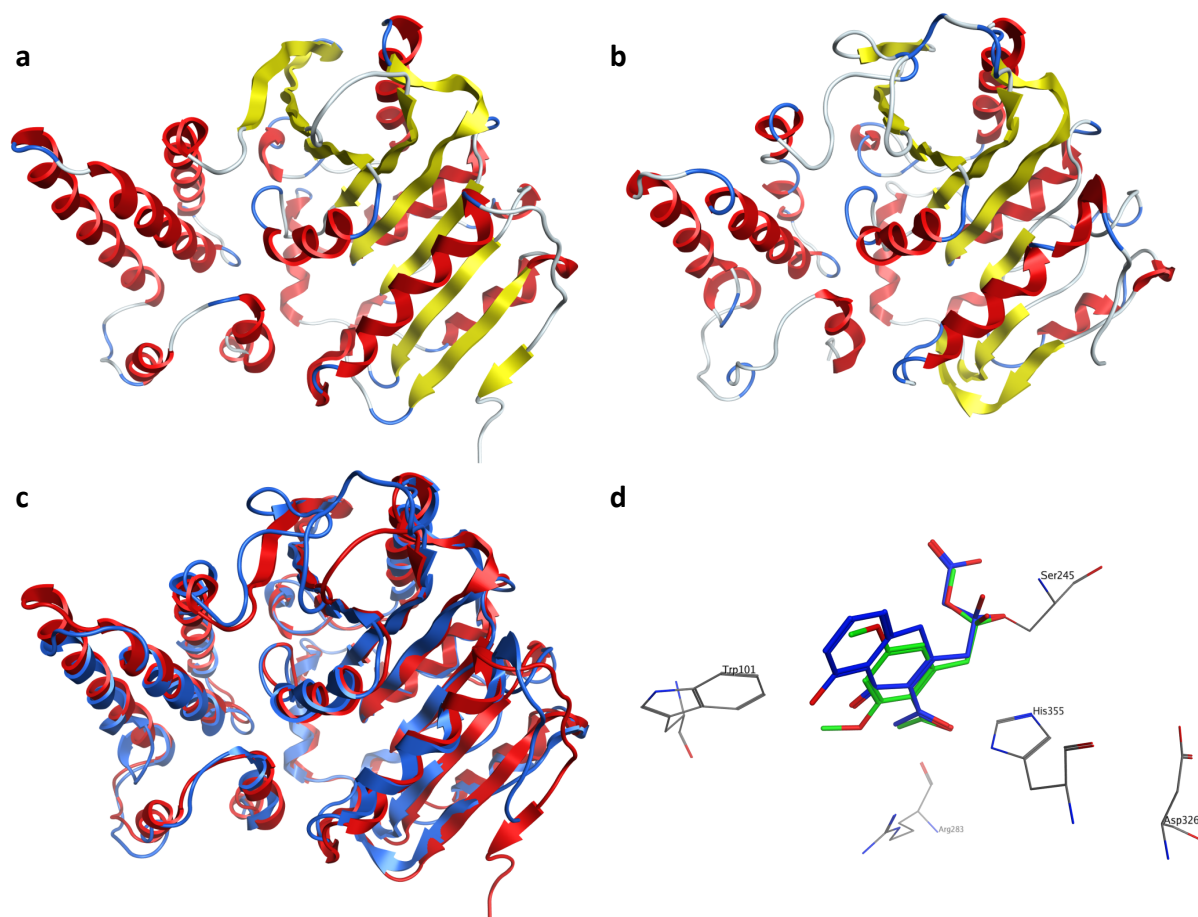
Figure S7. 3D Modeling of AntI. Tertiary structure of DHPON (2JBW) from *Arthrobacter nicotinovorans* (**a**) and the modeled structure of AntI from *Photorhabdus luminescens* subsp. *laumondii* TT01 (**b**). To determine the quality of the homology model both structures were superposed (DHPON (red) and modeled AntI (blue)) (**c**). This superposition showed only little structural deviations and a calculated root-mean-square deviation (RMSD) of 1.8 Å. The covalent docking of **16** (blue) and **20** (green) to Ser245 (**d**) revealed possible interactions between Trp101 and Arg283 to the docked ligand. Also shown is the possible catalytic triad of AntI consisting of Ser245, His355 and Asp326.
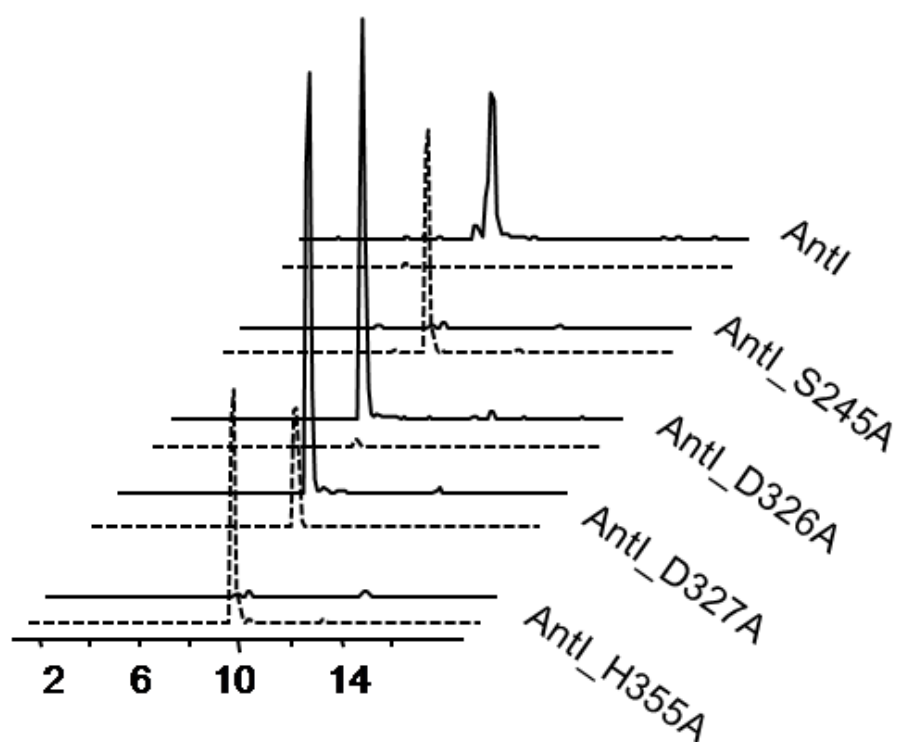
Figure S8. HPLC/MS analyses of *E. coli* expressing *antABCDEFGH* with different variants of AntI showing that S245A and H355A resulted in the complete loss of production of **1** and production of **10** instead. No change in production was observed for the D326A variant, whereas the D327A mutant showed the production of both **1** and **10**. Shown are EICs for **1** (m/z 255 [M-H]⁻) and **10** (dashed line, m/z 253 [M-H]⁻)

34

```
AntI   ACYHWAEFMYFSDRSR-----KIQLREYIRSCFLSSIKYS----DLLVDHQYIVVDKFHMPFFLIFPK   155
WdYG1  AVYRISRFPTPRSEKQ----KYAWRKGCEVFYKGAALMEYPIKEVRIPHKHGIEGEGDVVPVNFLLPP   169
Ayg1p  VVYRISRFPYVDITKPNSIKRVAFERQKQAYLKATSLWTQPIREVTVPHTYRTGNDGAHIPIYIRTPA   167
2JBW   LCAQYAQFLWFDERRQ-----KGQARKVELYQKAAPLLSP------PAERHELVVDGIPMPVYVRIPE   147

AntI   GYKEEENHPLPCVILSNGLDSMTEIEILSLAEFFLGKNMAVAIFDGPGQGINLGKSPIAIDMELYVSS   223
WdYG1  NAS--ETSPVPCVLIITGLDGYRT-ELAVWQQGWRSKGVATVIAEIPGTGDSPALRQDPTSPDRQWSS   234
Ayg1p  GAD--QSNPVPIVLIMTGLDGYRP-DNSQRTHEILARGWAAVVAEIPGTADCPADPADPASPDRLWDS   232
2JBW   GP-----GPHPAVIMLGGLESTKE-ESFQMENLVLDRGMATATFDGPGQGEMFEYKRIAGDYEKYTSA   209

AntI   IVKLLEDDARINSNLLCFLGISFGGYFALRVAQRIGDKFCCIVNLSGGPEIAEFDKLPRRLKEDFQFA   291
WdYG1  VLDWIESQKAVDSKKIVAWGFSTGGYYALRMAHTHKDRLLATISLGGGA-HHMFDREWLEHANKLEYP   301
Ayg1p  VLSYLDQRPELNTAKMVVWGLSAGGYYAIRAAHTHRDRLLGAIAHGPGC-HYYLDPEWLAKVNDHEYP   299
2JBW   VVDLLTKLEAIRNDAIGVLGRSLGGNYALKSAA-CEPRLAACISWGGFSDLDYWDLETPLTKESWKYV   276

AntI   F--------------MQDNS-HMQSIFDEIKLDISLPCKTKVFTVHGELDDIFQIDKVKKLDQLWGD   343
WdYG1  FD-LSNTLAYKFGYPDLESFIEESS-KFSLLNDGTLQKPCTKVLLVNGNDDEIFPIDDMFVSLENGQP   367
Ayg1p  FE-ITAAWATKHGYKTVEEFVAGAQKKFSLVETGIVDQPSCRLLLLNGVDDGVVPIEDCLVLFEHGSP   366
2JBW   SKV------------DTLEEARLHVHAALETRDVLSQIACPTYILHGVHDEVPLSFVDTVLELVPAE   331

AntI   NHQLLCYESEAHVCLNKINEYMIQ   367
WdYG1  KLA-RMVKGKKHMGE---PESFSI   387
Ayg1p  KEG-RFYKGLPHMGY---PNSLPV   386
2JBW   HLNLVVEKDGDHCCHNLGIRPRLE   355
```

Figure S9. Multiple sequence alignment of hydrolase sequences. DHPON (2JBW) from *Arthrobacter nicotinovorans,* WdYG1 from *Wangiella dermatitidis,* Ayg1p from *Aspergillus fumigatus* and AntI from *Photorhabdus luminescens* subsp. *laumondii* TT01. Here shown is a part of the multiple alignment which originally is consisting of 44 hydrolase sequences. Highlighted are conserved amino acids that share a similarity of more than 80% in grey and amino acids predicted to form the catalytic triad are marked purple. The shown C-terminal part is responsible for the formation of the conserved α/β-hydrolase fold.
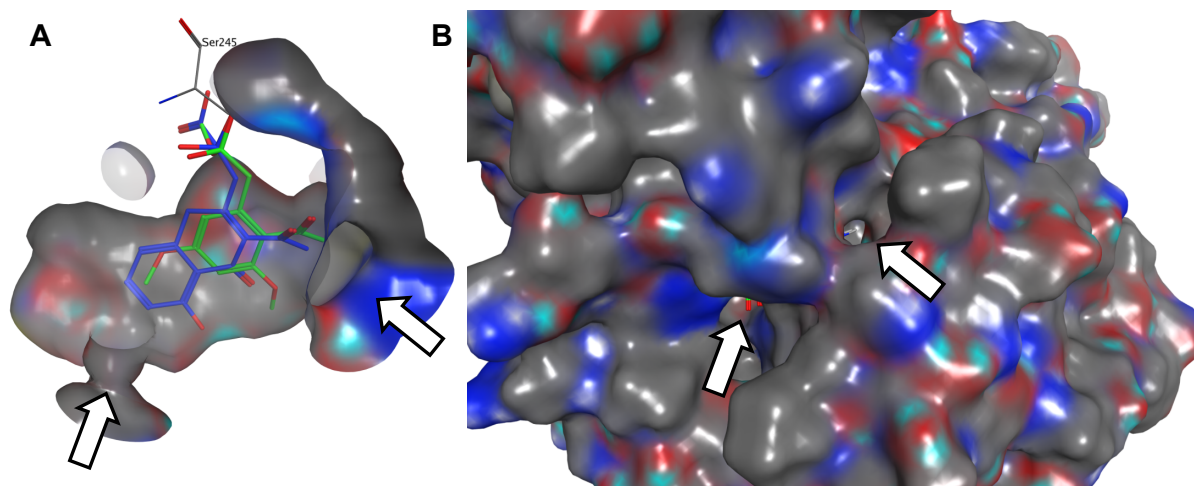
Figure S10. A detailed view of the proposed AntI binding pocket with **16** (blue) and **18** (green) covalently bound to Ser245 (**A**). The calculated surface of the binding pocket is shown, indicating that two possible tunnels are being formed. The overall calculated AntI surface is also shown (B), indicating access of the two tunnels to the outer milieu.
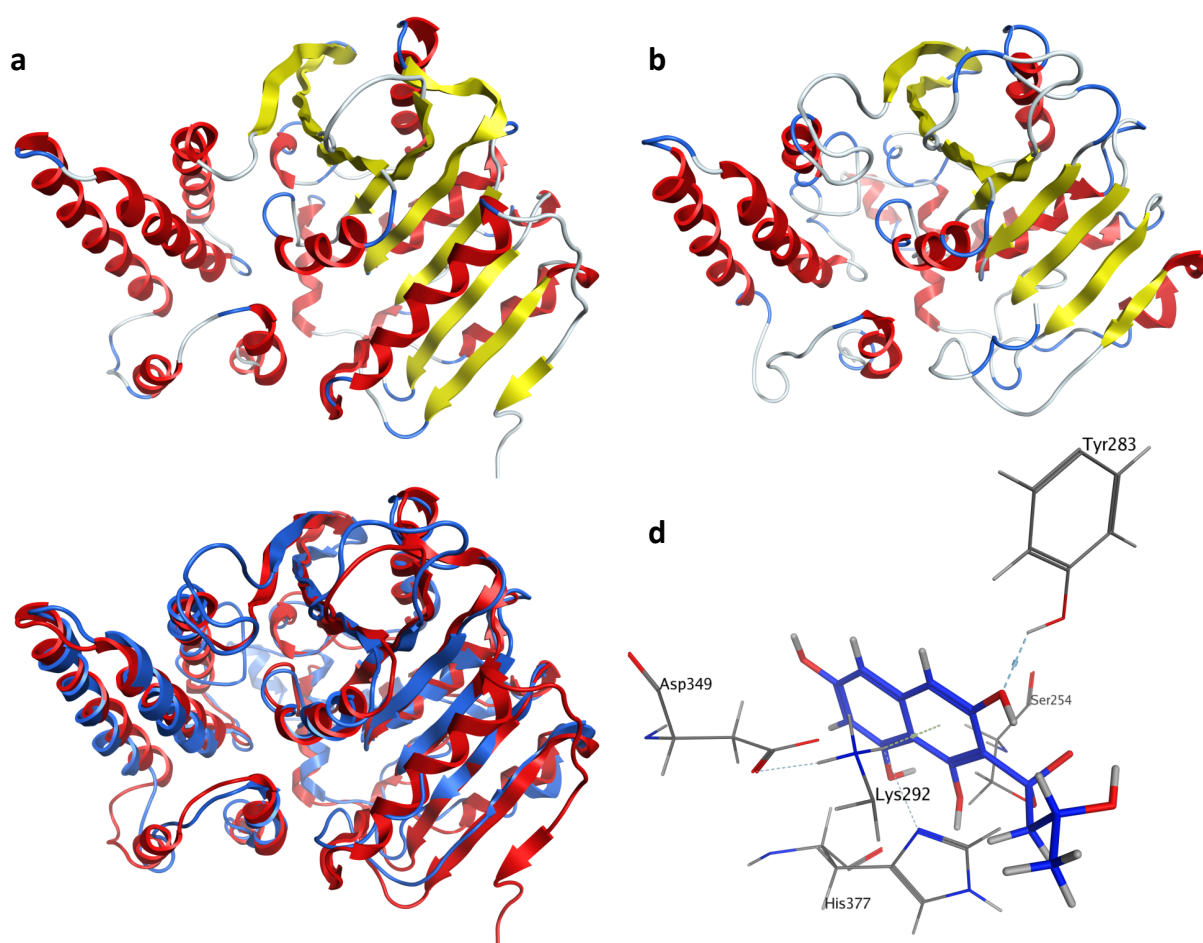
Figure S11. 3D Modeling of Ayg1p. Tertiary structure of DHPON (2JBW) from *Arthrobacter nicotinovorans* (**a**) and the modeled structure of Ayg1p from *Aspergillus fumigatus*(**b**). To determine the quality of the homology model both structures were superposed (DHPON (red) and modeled Ayg1p (blue)) (**c**). This superposition showed only little structural deviations and a calculated root-mean-square deviation (RMSD) of 1.8 Å. The covalent docking of the heptaketide napthopyrone YWA1 (blue) to Ser254 (**d**) revealed possible interactions between Tyr283 and Lys292 to the docked ligand. Also shown is the possible catalytic triad of Ayg1p consisting of Ser254, His377 and Asp349.
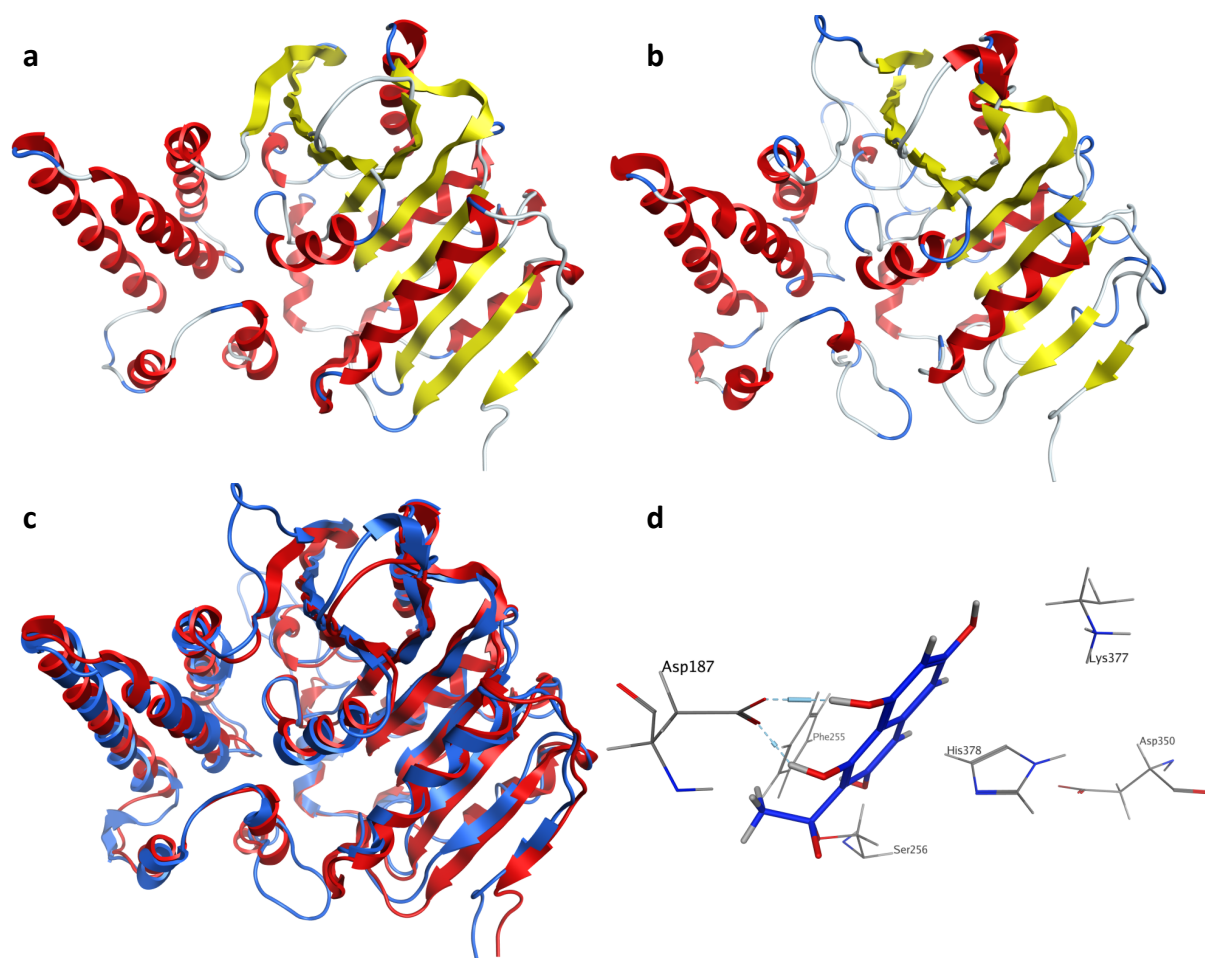
Figure S12. 3D Modeling of WdYG1. Tertiary structure of DHPON (2JBW) from *Arthrobacter nicotinovorans* (**a**) and the modeled structure of WdYG1 from *Wangiella dermatitidis* (**b**). To determine the quality of the homology model both structures were superposed (DHPON (red) and modeled WdYG1 (blue)) (**c**). This superposition showed only little structural deviations and a calculated root-mean-square deviation (RMSD) of 1.7 Å. The covalent docking of 2-acetyl-1,3,6,8-tetrahydroxy naphthalene (blue) to Ser256 (**d**) revealed possible interactions between Asp187, Phe255 and Lys377 to the docked ligand. Also shown is the possible catalytic triad of WdYG1 consisting of Ser256, His378 and Asp350.
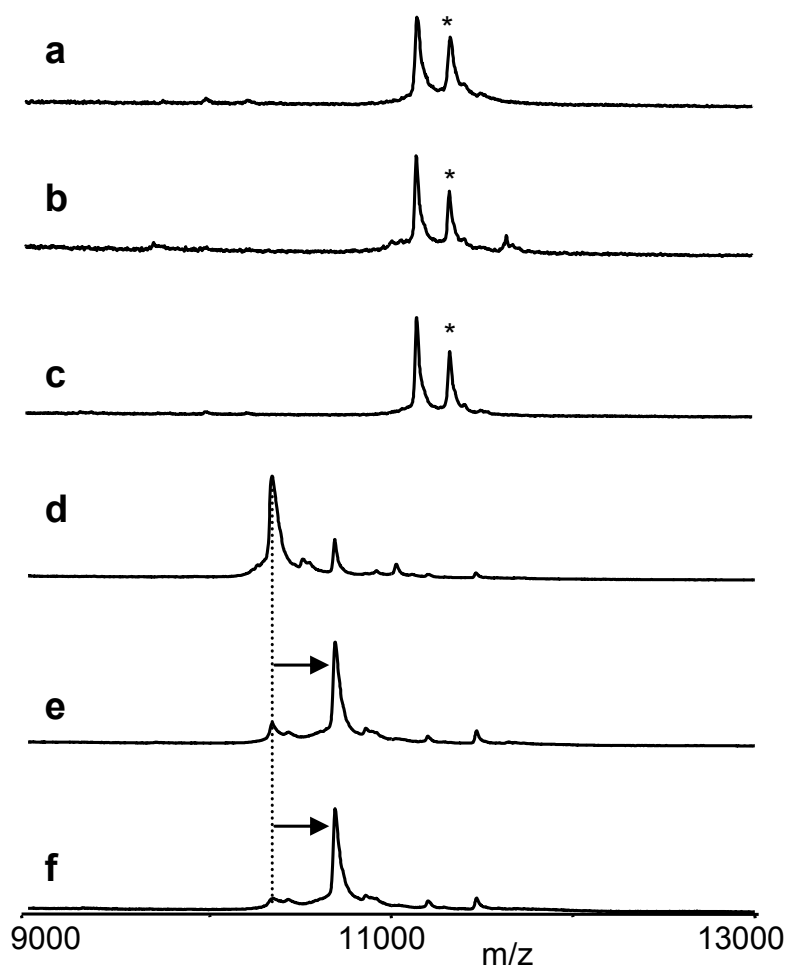
Figure S13. UTL-MALDI-TOF results from *in vitro* assays, activation of His6-*apo*-AntF with the purified protein Sfp and MtaA. (**a**) His6-*apo*-AntF. (**b**) His6-*apo*-AntF + MtaA. (**c**) His6-*apo*-AntF + Sfp. (**d**) Cpin_1856. (**e**) Cpin_1856 + MtaA. (**f**) Cpin_1856 + Sfp. Cpin_1856, used as a positive control, is an ACP from *Chitinophaga pinensis*. The peaks marked with star (*) was gluconylated His6-*apo*-ACP.[22]
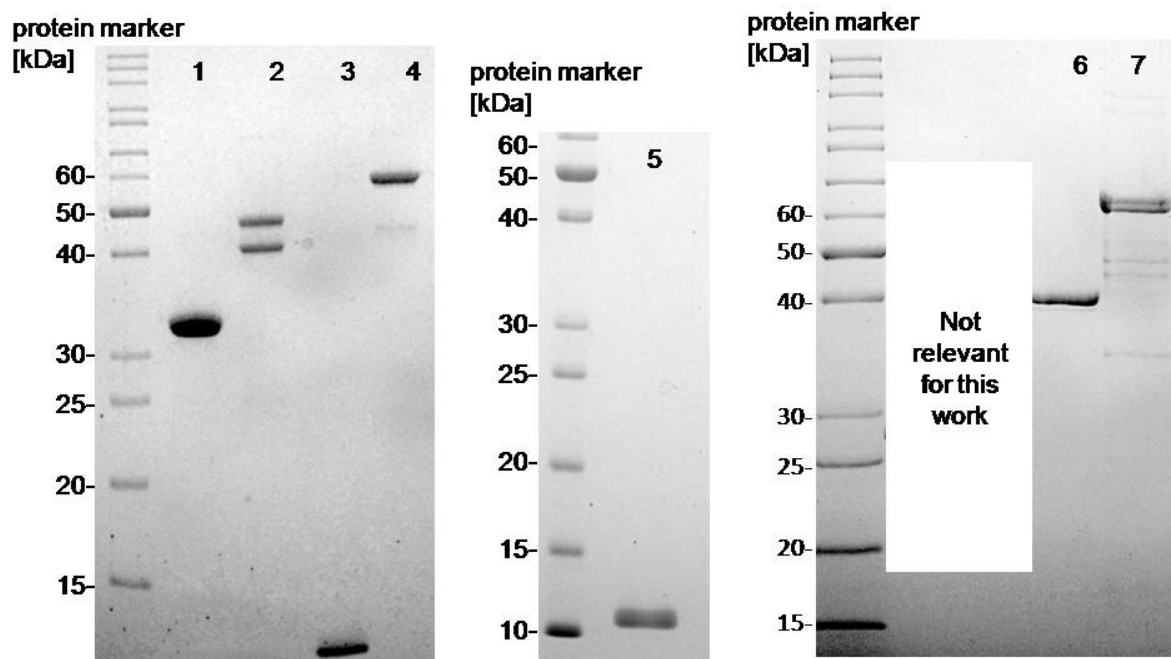
Figure S14. SDS-Page gel (15%) of purified recombinant His$_6$-tagged protein stained with Coomassie blue. His$_6$-Plu2834 (35.0 kDa, lane 1). His$_6$-AntD-AntE (47.5 kDa, 40.7 kDa, lane 2). His$_6$-*holo*-AntF (11.4 kDa, lane 3). His$_6$-AntG (60.2 kDa, lane 4). His$_6$-*apo*-AntF (11.1 kDa, lane 5). His$_6$-AntI (46.0 kDa, lane 6). His6-MatB (60.0 kDa, lane 7).
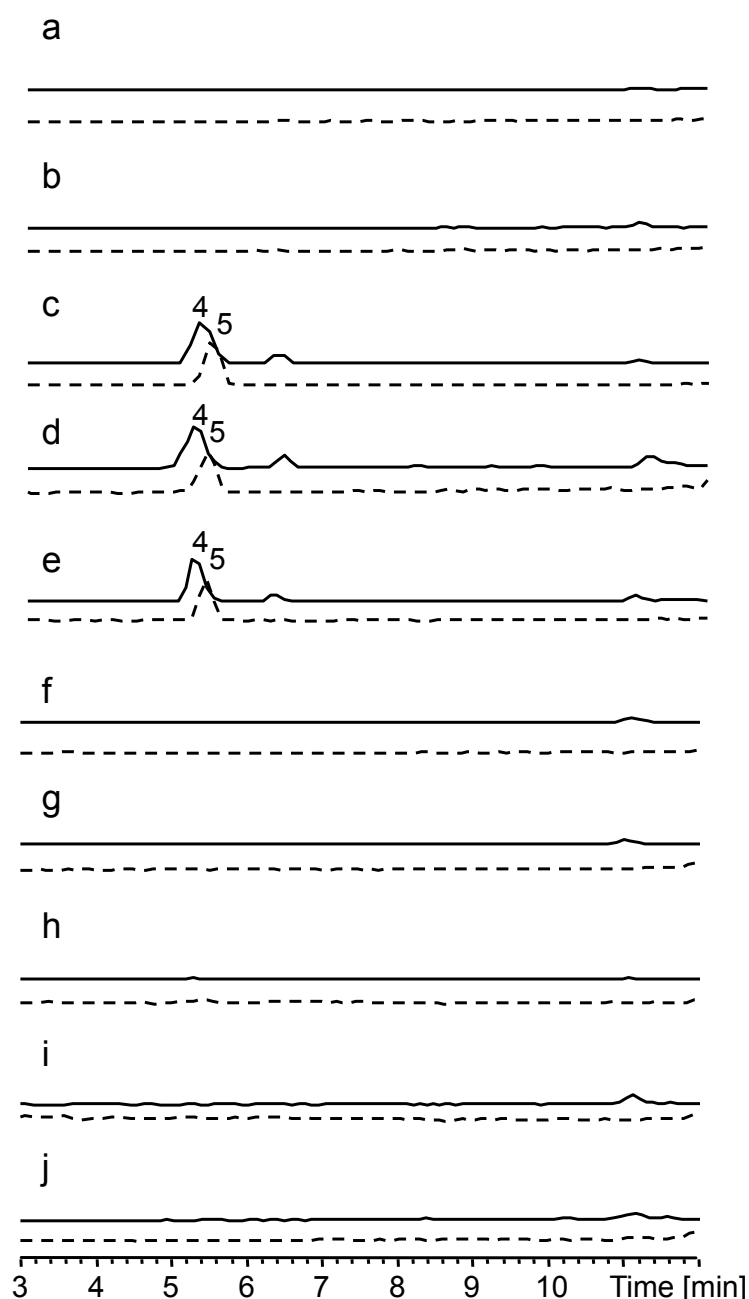
Figure S15. HPLC/MS data of *in vitro* biosynthesis of **4** and **5**. Shown are EICs for **4** (m/z 301 [M+H]$^+$) and **5** (dashed line, m/z 319 [M+H]$^+$). **a**: *holo*-AntF + Buffer; **b**: AntDE + Plu2834 + AntG + *apo*-AntF + malonyl-CoA + acetyl-CoA + Buffer; **c**: AntDE + Plu2834 + AntG + *holo*-ACP + malonyl-CoA + acetyl-CoA + Buffer; **d**: AntDE + Plu2834 + AntG + *holo*-ACP + malonyl-CoA + Buffer; **e**: AntDE + Plu2834 + *holo*-ACP + malonyl-CoA + Buffer; **f**: AntDE + Plu2834 + malonyl-CoA + Buffer; **g**: AntDE + *holo*-ACP + malonyl-CoA + Buffer; **h**: Plu2834 + *holo*-ACP + malonyl-CoA + Buffer; **i**: AntDE + Plu2834 + *holo*-ACP + Buffer; **j**: AntDE + Plu2834 + *holo*-ACP + malonyl-CoA. Details about the assay condition are shown in Table S8.

Figure S16. Alignment of adenylate forming domain, class I superfamily representatives (top).[23] This family includes acyl- and aryl-CoA ligases, as well as the adenylation domain of nonribosomal peptide synthetases and firefly luciferases. The AMP binding site is yellow shaded, CoA binding site is grey shaded and the acyl-activating enzyme consensus motif is shaded purple. AntG is marked in red. For details about the included sequences see the table at the bottom.
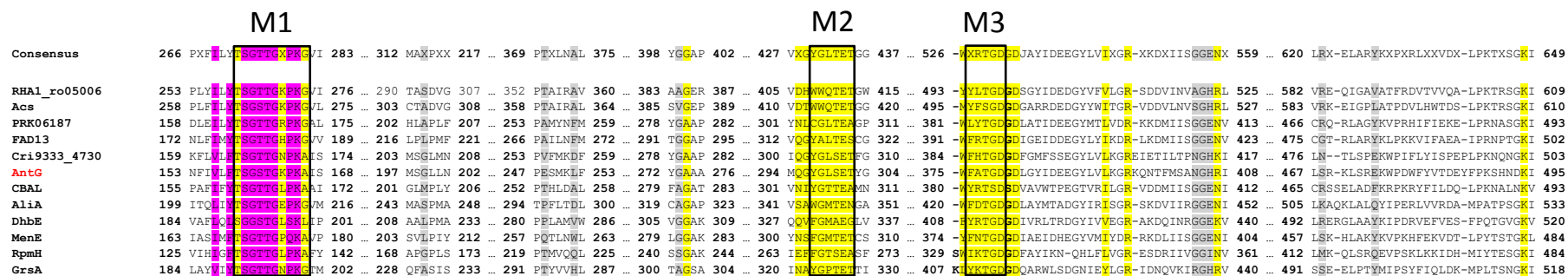
Protein	Accession	Description	Organism
RHA1_ro05006	gi|111021973|ref|YP_704945.1|	propionate--CoA ligase	Rhodococcus jostii RHA1
Acs	gi|31616027|pdb|1PG3|A	Acetyl CoA Synthetase	Salmonella enterica
PRK06187	gi|653070309|ref|WP_027321487.1|	long-chain fatty acid--CoA ligase	Bacillus sp. URHB0009
FAD13	gi|374977594|pdb|3R44|A	Fatty Acyl Coa Synthetase	Mycobacterium Tuberculosis
Cri9333_4730	gi|504992837|ref|WP_015179939.1|	o-succinylbenzoate--CoA ligase	Crinalium epipsammum
AntG	gi|499461147|ref|WP_011148292.1|	hypothetical protein	Photorhabdus luminescens
CBAL	gi|197725159|pdb|3CW8|X	4-chlorobenzoyl-coa Ligase/synthetase	Alcaligenes sp.
AliA	gi|2190573|gb|AAC23919.1|	cyclohex-1-ene-1-carboxylate CoA ligase	Rhodopseudomonas palustris CGA009
DhbE	gi|16080251|ref|NP_391078.1|	2,3-dihydroxybenzoate-AMP ligase	Bacillus subtilis subsp. subtilis str. 168
MenE	gi|446270509|ref|WP_000348364.1|	2-succinylbenzoate--CoA ligase	Staphylococcus aureus
RpmH	gi|78099285|sp|Q5HRH4.1|VRAA_STAEQ	Acyl-CoA synthetase	Staphylococcus epidermidis RP62A
PheA/GrsA	gi|3318718|pdb|1AMU|A	Phenylalanine Activating Domain Of Gramicidin Synthetase 1	Brevibacillus brevis

221

Figure S17. HPLC analysis of the CoA ligase activities, using His$_6$-MatB and His$_6$-AntG. MatB (a malonyl CoA ligase from *Rizobium trifolii*). EIC with *m/z* 810 [M+H]$^+$ for acetyl-CoA (**a**, **c**, **e**) and EIC with *m/z* 854 [M+H]$^+$ for malonyl-CoA (**b**, **d**, **f**). The results showed that His6-AntG is neither able to synthesize malonyl-CoA nor acetyl-CoA. His6-MatB was used as a positive control.

222
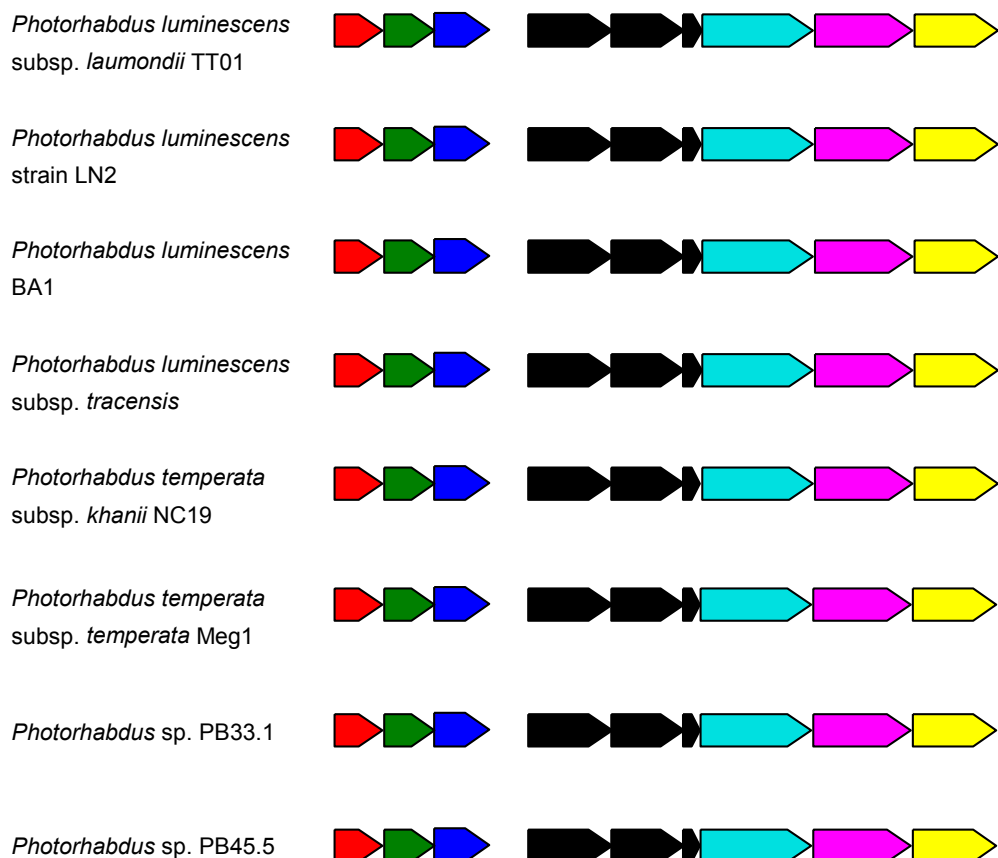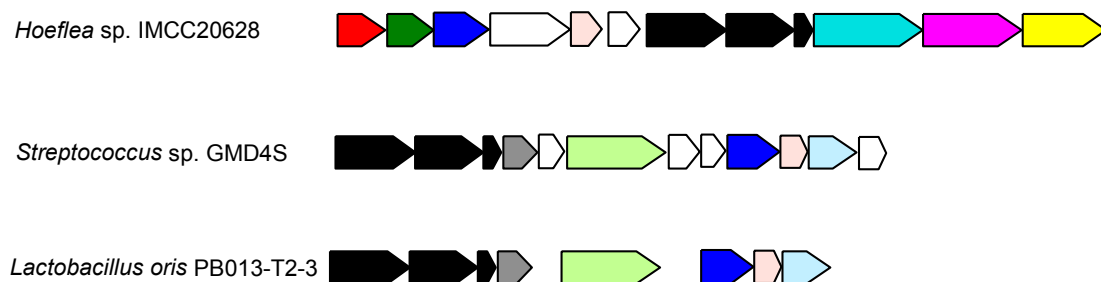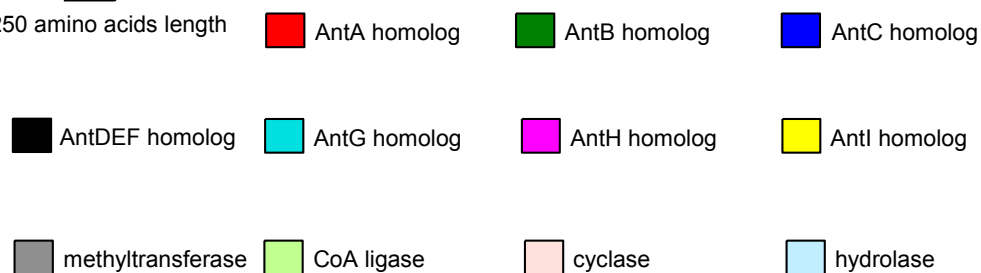
Figure S18. Biosynthesis gene clusters related to *ant* from *P. luminescens* from other *Photorhabdus* strains (a) and other bacteria (b).
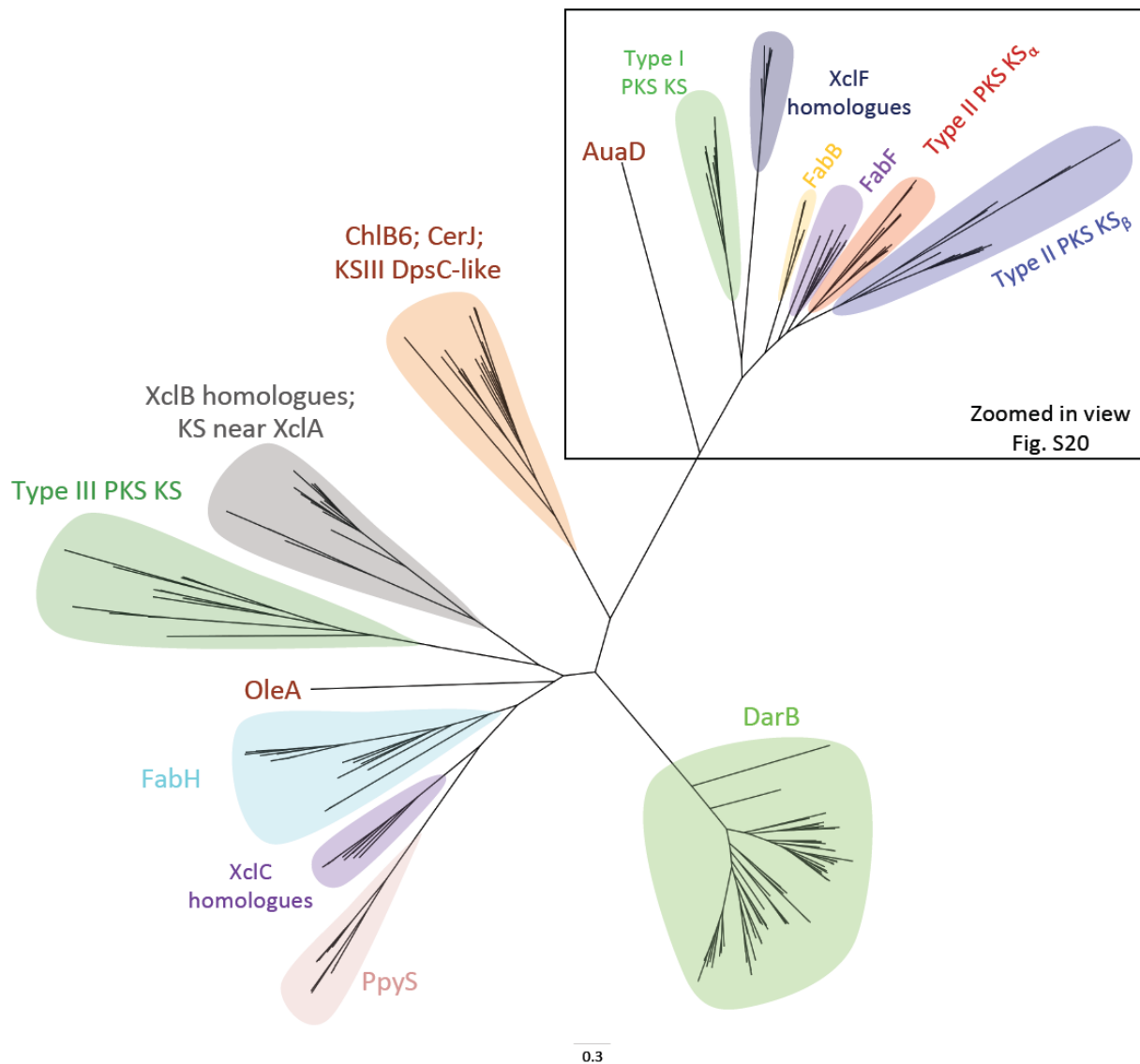
Figure S19. Phylogenetic tree (PHYML) composed of AntE and AntD, its homologues and other known ketosynthases. Table S10 lists all ketosynthases shown in this tree. The scale bar indicates the degree of divergence as substitutions per site.
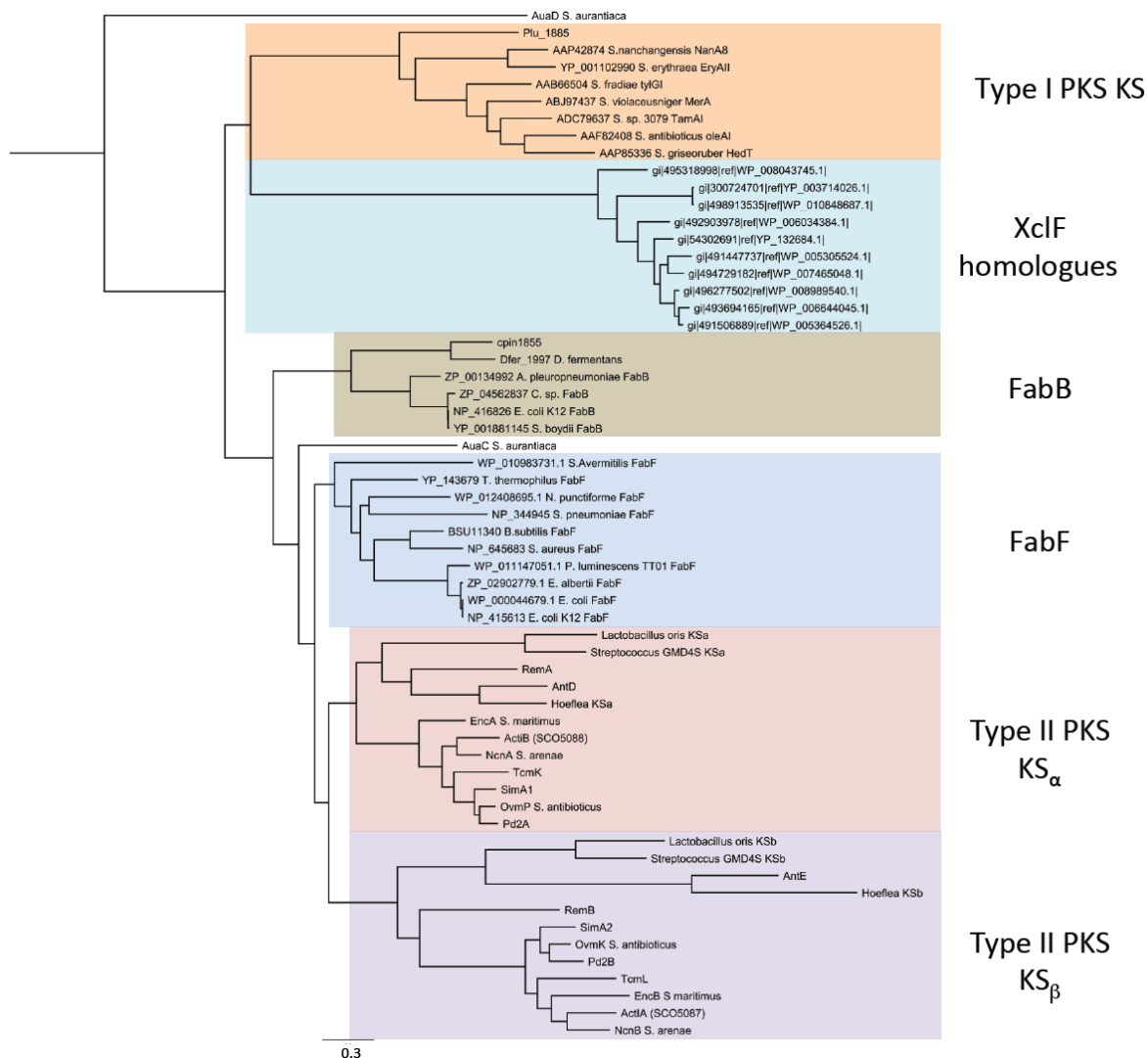
Figure S20. A zoomed in view of the top part of the ketosynthase phylogenetic tree. Comprising known AntE, AntD, its homologues and other ketosynthase α and ketosynthase β.

NMR spectra

226

47

Parameter	Value
Origin	Bruker BioSpin GmbH
Solvent	CDCl3
Temperature	300.0
Pulse Sequence	zg30
Number of Scans	32
Receiver Gain	51
Relaxation Delay	0.1000
Pulse Width	10.6000
Acquisition Time	3.9998
Spectrometer Frequency	250.13
Spectral Width	4990.0
Lowest Frequency	-1036.1
Nucleus	1H
Acquired Size	19959
Spectral Size	262144

Parameter	Value
Origin	Bruker BioSpin GmbH
Solvent	CDCl3
Temperature	300.0
Pulse Sequence	zg30
Number of Scans	32
Receiver Gain	51
Relaxation Delay	0.1000
Pulse Width	10.6000
Acquisition Time	3.9998
Spectrometer Frequency	250.13
Spectral Width	4990.0
Lowest Frequency	-1035.9
Nucleus	1H
Acquired Size	19959
Spectral Size	262144

227

48

Parameter	Value
Origin	Bruker BioSpin GmbH
Solvent	CDCl3
Temperature	300.0
Pulse Sequence	zgpg30
Number of Scans	4096
Receiver Gain	11585
Relaxation Delay	0.5000
Pulse Width	10.8000
Acquisition Time	1.0421
Spectrometer Frequency	62.90
Spectral Width	15723.3
Lowest Frequency	-1572.1
Nucleus	13C
Acquired Size	16384
Spectral Size	131072

170.33
160.30
139.51
108.96
100.80
98.73
55.28
51.44
49.05
44.50
41.52
25.86
18.34
-2.67
-2.90

f1 (ppm)

228

49

Parameter	Value
Origin	Bruker BioSpin GmbH
Solvent	CDCl3
Temperature	298.2
Pulse Sequence	zg30_ns
Number of Scans	32
Receiver Gain	45
Relaxation Delay	0.1000
Pulse Width	11.2000
Acquisition Time	3.9999
Spectrometer Frequency	500.18
Spectral Width	10000.0
Lowest Frequency	-2044.7
Nucleus	1H
Acquired Size	39999
Spectral Size	131072
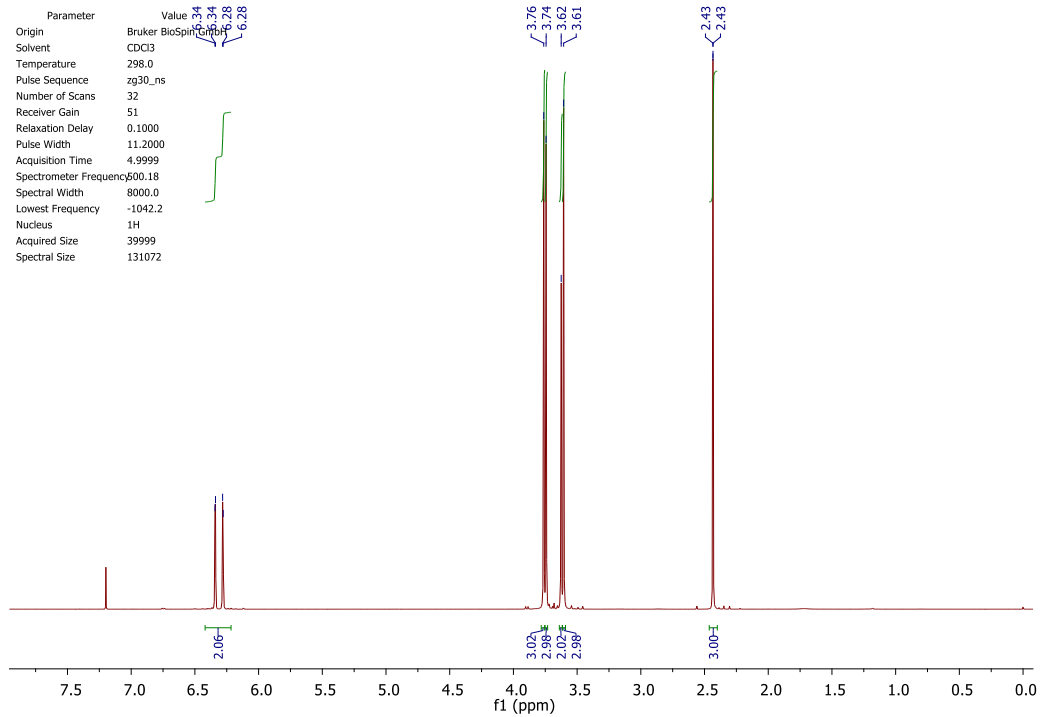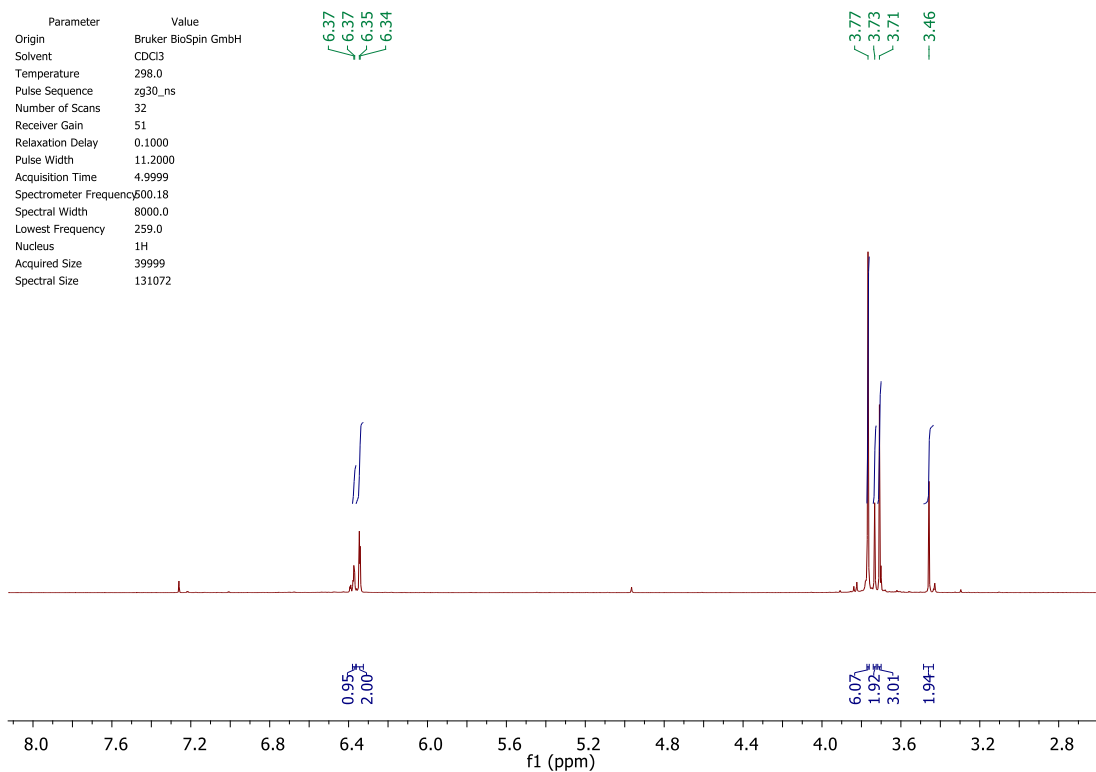
Parameter	Value
Origin	Bruker BioSpin GmbH
Solvent	CDCl3
Temperature	298.0
Pulse Sequence	deptqsp
Number of Scans	1024
Receiver Gain	2050
Relaxation Delay	2.0000
Pulse Width	10.0000
Acquisition Time	1.0486
Spectrometer Frequency	125.77
Spectral Width	31250.0
Lowest Frequency	-3125.0
Nucleus	13C
Acquired Size	32768
Spectral Size	65536

230

51
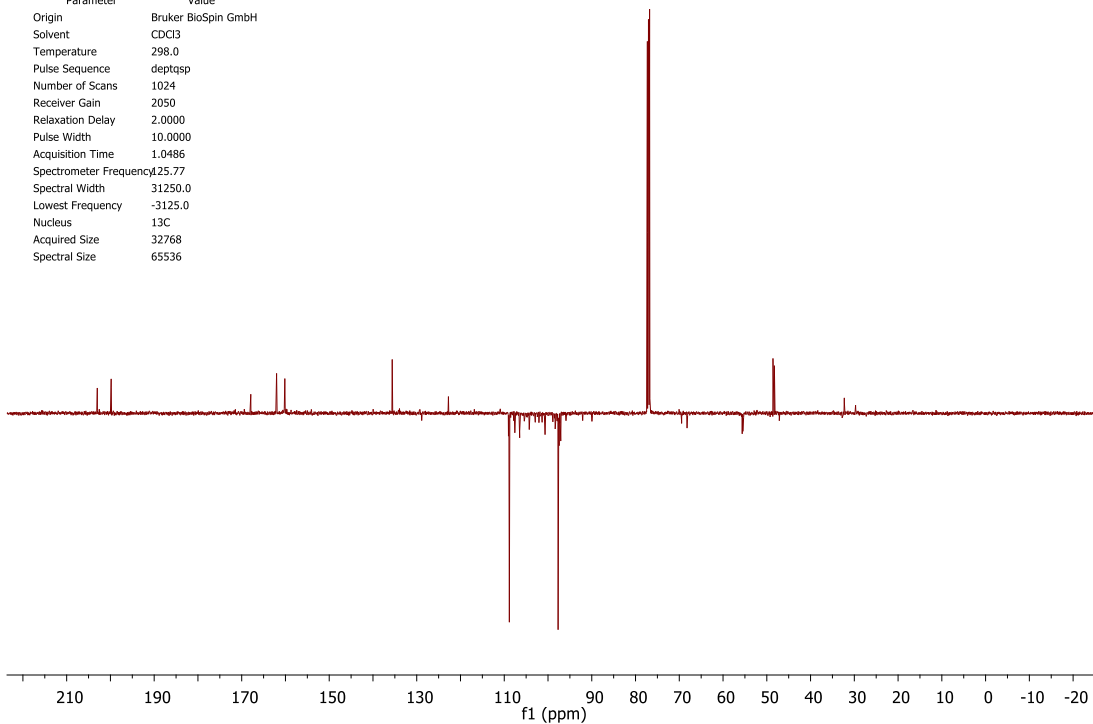
References

[1]   N. Gaitatzis, A. Hans, R. Müller, S. Beyer, *Journal of Biochemistry* **2001**, *129*, 119–124.

[2]   J. H. An, G. Y. Lee, J. W. Jung, W. Lee, Y. S. Kim, *Biochem. J.* **1999**, *344 Pt 1*, 159–166.

[3]   T. Itoh, T. Taguchi, M. R. Kimberley, K. I. Booker-Milburn, G. R. Stephenson, Y. Ebizuka, K. Ichinose, *Biochemistry* **2007**, *46*, 8181–8188.

[4]   K. Jakobi, C. Hertweck, *J. Am. Chem. Soc.* **2004**, *126*, 2298–2299.

[5]   A. O. Brachmann, S. A. Joyce, H. Jenke-Kodama, G. Schwär, D. J. Clarke, H. B. Bode, *ChemBioChem* **2007**, *8*, 1721–1728.

[6]   D. Fenyo, Q. Wang, J. A. DeGrasse, J. C. Padovan, M. Cadene, B. T. Chait, *JoVE* **2007**, 192.

[7]   S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, *Journal of Molecular Biology* **1990**, *215*, 403–410.

[8]   H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucleic Acids Research* **2000**, *28*, 235–242.

[9]   G. S. Kachalova, G. P. Bourenkov, T. Mengesdorf, S. Schenk, H. R. Maun, M. Burghammer, C. Riekel, K. Decker, H. D. Bartunik, *Journal of Molecular Biology* **2010**, *396*, 785–799.

[10]   M. A. Larkin, G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, et al., *Bioinformatics* **2007**, *23*, 2947–2948.

[11]   G. Jones, P. Willett, R. C. Glen, A. R. Leach, R. Taylor, *Journal of Molecular Biology* **1997**, *267*, 727–748.

[12]   O. Korb, T. Stützle, T. E. Exner, *J. Chem. Inf. Model.* **2009**, *49*, 84–96.

[13]   S. Guindon, O. Gascuel, *Systematic Biology* **2003**, *52*, 696–704.

[14]   B. Bartolomé, Y. Jubete, E. Martínez, F. de la Cruz, *Gene* **1991**, *102*, 75–78.

[15]   N. Philippe, J.-P. Alcaraz, E. Coursange, J. Geiselmann, D. Schneider, *Plasmid* **2004**, *51*, 246–255.

[16]   S. G. Grant, J. Jessee, F. R. Bloom, D. Hanahan, *Proceedings of the National Academy of Sciences* **1990**, *87*, 4645–4649.

[17]   L. D. Simon, B. Randolph, N. Irwin, G. Binkowski, *Proceedings of the National Academy of Sciences* **1983**, *80*, 2059–2062.

[18]   J. D. Bauer, R. W. King, S. F. Brady, *J. Nat. Prod.* **2010**, *73*, 976–979.

[19]   E. Duchaud, C. Rusniok, L. Frangeul, C. Buchrieser, A. Givaudan, S. Taourit, S. Bocs, C. Boursaux-Eude, M. Chandler, J.-F. Charles, et al., *Nat Biotechnol* **2003**, *21*, 1307–1313.

[20]   A. Proschak, Q. Zhou, T. Schöner, A. Thanwisai, D. Kresovic, A. Dowling, R. ffrench-Constant, E. Proschak, H. B. Bode, *ChemBioChem* **2014**, *15*, 369–372.

[21]   D. Kresovic, F. Schempp, Z. Cheikh-Ali, H. B. Bode, *Beilstein J. Org. Chem.* **2015**, *11*, 1412–1417.

[22]   K. F. Geoghegan, H. B. F. Dixon, P. J. Rosner, L. R. Hoth, A. J. Lanzetti, K. A. Borzilleri, E. S. Marr, L. H. Pezzullo, L. B. Martin, P. K. LeMotte, et al., *Analytical Biochemistry* **1999**, *267*, 169–184.

[23]   K.-H. Chang, H. Xiang, D. Dunaway-Mariano, *Biochemistry* **1997**, *36*, 15650–15659.

11.4. Biosynthesis of the Insecticidal Xenocyloins in *Xenorhabdus bovienii*

Authors: Anna Proschak,[a] Qiuqin Zhou,[a] Tim Schöner,[a] Aunchalee Thanwisai,[a, b] Darko Kresovic,[a] Andrea Dowling,[c] Richard Ffrench-Constant,[c] Ewgenij Proschak,[d] and Helge B. Bode[a]

[a] Merck Stiftungsprofessur für Molekulare Biotechnologie, Fachbereich Biowissenschaften, Goethe Universität Frankfurt, Max-von-Laue-Strasse 9, 60438 Frankfurt am Main (Germany)

[b] Department of Microbiology and Parasiology, Faculty of Medical Science, Naresuan University, 99 Moo 9, Phitsanulok-Nakhon Sawan Road, Tha Pho, Mueang Phitsanulok, 65000 Phitsanulok (Thailand)

[c] Biosciences, University of Exeter in Cornwall, Tremough Campus, Penryn, Cornwall TR10 9EZ (UK)

[d] Institute of Pharmaceutical Chemistry, Goethe Universität Frankfurt, Max-von-Laue-Strasse 9, 60438 Frankfurt am Main (Germany)

Contribution: Determination of the absolute configuration of xenocyloins using the CD spectrometry. Design of *xclC* genes with mutations (xclC_S253A, xclC_C118A, xclC_Y283A).

Attachment: the paper.

Biosynthesis of the Insecticidal Xenocyloins in *Xenorhabdus bovienii*

Anna Proschak,[a] Qiuqin Zhou,[a] Tim Schöner,[a] Aunchalee Thanwisai,[a, b] Darko Kresovic,[a] Andrea Dowling,[c] Richard ffrench-Constant,[c] Ewgenij Proschak,[d] and Helge B. Bode*[a]

The biosynthesis gene cluster for the production of xenocyloins was identified in the entomopathogenic bacterium *Xenorhabdus bovienii* SS-2004, and their biosynthesis was elucidated by heterologous expression and in vitro characterization of the enzymes. XclA is an *S*-selective ThDP-dependent acyloin-like condensation enzyme, and XclB and XclC are examples of the still-rare acylating ketosynthases that catalyze the acylation of the XclA-derived initial xenocyloins with acetyl-, propionyl-, or malonyl-CoA, thereby resulting in the formation of further xenocyloin derivatives. All xenocyloins were produced mainly by the more virulent primary variant of *X. bovienii* and showed activity against insect hemocytes thus contributing to the overall virulence of *X. bovienii* against insects.

Here we describe the identification of the xenocyloin biosynthesis gene cluster and elucidation of the biosynthesis pathway, including the rare formation of *S*-hydroxyketones. Feeding experiments were performed to identify the precursors of xenocyloin biosynthesis. In this study, XclC, a homologue of 3-ketoacyl acyl carrier protein (ACP) synthase III (KAS III), was shown to act as an acyltransferase; it is only the second example of such an acylating ketosynthase (KS). We postulate the mechanism of acetyl transfer to xenocyloins based on in vitro assays, homology modeling, and simulated docking experiments, as well as site-directed mutagenesis of XclC. Furthermore, we show that xenocyloins are insecticidal compounds active against insect hemocytes.

Introduction

Xenorhabdus species are Gram-negative bacteria that live in symbiotic association with soil-dwelling nematodes of the genus *Steinernema*. Once the nematode infects an insect larva, bacteria are released and start to proliferate in the hemocoel of the insect and to produce natural products that inhibit the insect immune system and protect their food source from other microorganisms.[1–4] In our analysis of natural products from *Xenorhabdus bovienii*, the genome of which was recently sequenced,[5] structurally simple indole compounds were identified and described as potent antibiotics.[6–8] These compounds were named "xenocyloins" because of the postulated acyloin condensation involved in their biosynthesis.

[a] A. Proschak, Q. Zhou, T. Schöner, Dr. A. Thanwisai, D. Kresovic,
Prof. Dr. H. B. Bode
Merck Stiftungsprofessur für Molekulare Biotechnologie
Fachbereich Biowissenschaften, Göthe Universität Frankfurt
Max-von-Laue-Strasse 9, 60438 Frankfurt am Main (Germany)
E-mail: h.bode@bio.uni-frankfurt.de

[b] Dr. A. Thanwisai
Department of Microbiology and Parasiology
Faculty of Medical Science, Naresuan University
99 Moo 9, Phitsanulok-Nakhon Sawan Road, Tha Pho
Mueang Phitsanulok, 65000 Phitsanulok (Thailand)

[c] Dr. A. Dowling, Prof. Dr. R. ffrench-Constant
Biosciences, University of Exeter in Cornwall
Tremough Campus, Penryn, Cornwall TR10 9EZ (UK)

[d] Prof. Dr. E. Proschak
Institute of Pharmaceutical Chemistry, Göthe Universität Frankfurt
Max-von-Laue-Strasse 9, 60438 Frankfurt am Main (Germany)

Supporting information for this article is available on the WWW under
http://dx.doi.org/10.1002/cbic.201300694.

Results and Discussion

Analysis of *X. bovienii* SS-2004 for the production of natural products revealed known indole derivatives **1**–**4** as the main compounds (Scheme 1).[9] Additionally, a novel derivative (**5**) with a propionyl side chain was isolated, and its structure was elucidated by detailed 1D and 2D NMR methods (see Table S1 in the Supporting Information, Annexes I–V) and detailed MS analysis (Table 1). The absolute configuration of **1** was determined by circular dichroism to reveal the *S*-configuration (Figure S1), which is rare in nature; thus, the configuration of the hydroxyl groups of all compounds was assumed to be *S*.[10, 11]

In order to identify the biosynthesis gene cluster responsible for the production of these compounds, the genome of *X. bovienii* SS-2004 was searched for genes encoding thiamine diphosphate (ThDP)-dependent acetolactate-synthase-like proteins, as these have been described to be involved in the postulated acyloin condensation.[12] One such gene was part of a three-gene operon and could not be found in related species such as *Xenorhabdus nematophila* HGB081 and *Photorhabdus luminescens* TT01. A BLASTP analysis of the proteins encoded by these three genes (XBJ1_3901, XBJ1_3900, and XBJ1_3898) and their neighboring genes allowed their annotation (Figure 1A, Table S2). Heterologous expression of the acetolactate-synthase-encoding gene *xclA* (XBJ1_3901) in *Escherichia coli* led to the expected production of **1** and **2** (Figure 1C). As this

Scheme 1. Structures of xenocyloins A–E (**1**–**5**) produced by *X. bovienii* SS-2004.

Xenocyloin	R¹	R²
A	**1**	H	Me
B	**2**	H	Et
C	**3**	Ac	Me
D	**4**	Ac	Et
E	**5**	Pr	Et

Table 1. Xenocyloin A–E (**1**–**5**) from *X. bovienii* SS-2004 and their amounts relative to **2** (100%), and the labeling results from heterologously expressed *xclABCDEF* in *E. coli* grown in ^{13}C medium. Incorporation is indicated by a tick (✓); no insertion is shown with a dash (–).

Compound	1	2	3	4	5
relative amount [%]	38	100	7	40	4
M_w [Da]	231	245	273	287	301
HRESI-MS [*m/z*]	232.1331	246.1484	274.1433	288.1592	302.1330
molecular formula	$C_{14}H_{17}NO_2$	$C_{15}H_{19}NO_2$	$C_{16}H_{19}NO_3$	$C_{17}H_{21}NO_3$	$C_{18}H_{23}NO_3$
^{13}C [*m/z*] ([*M*+H]$^+$; [*M*+Na]$^+$)	246; 268	261; 283	290; 312	305; 327	320; 342
Reverse feeding in ^{13}C medium with ^{12}C precursors					
L-tryptophan	✓	✓	✓	✓	✓
L-valine	✓	–	✓	–	–
L-isoleucine	–	✓	–	✓	✓
indole-3-pyruvic acid	✓	✓	✓	✓	✓
indole acetaldehyde	✓	✓	✓	✓	✓
3-methyl-2-oxobutyric acid	✓	–	✓	–	–
3-methyl-2-oxovaleric acid	–	✓	–	✓	✓

Figure 2). The unstable indole acetaldehyde could substitute for indole-3-pyruvate (as has been shown by Müller and co-workers);[13] no incorporation of indole acetic acid, 2-metylbutyrylaldehyde, or 2-methylbutyric acid was observed (data not shown). Subsequent in vitro formation of **1** or **2** by XclA (upon incubation of indole-3-pyruvate with 3-methyl-2-oxobutanoate or 3-methyl-2-oxovalerate, respectively; Figure 3) confirmed the acyloin condensation mechanism (Figure S2), as has been proposed for the biosynthesis of scytonemin,[14] including formation of *S*-hydroxyketone as the first intermediate in the biosynthesis. The proposed short-lived carboxylated precursor of **1** (Figure S2) was not detected in the XclA reaction, probably because of its instability, as previously reported.[14]

Surprisingly, deletion of the predicted esterase-encoding gene *xclD* (XBJ1_3897) did not affect production of **3**–**5** (data not shown), and expression of *xclABC* was sufficient to produce **1**–**5** (Figure 1D). As XclB (XBJ1_3900) and XclC (XBJ1_3898) show similarity to 3-oxoacyl ACP synthase III (KS III, an enzyme usually involved in C—C bond formation), both enzymes were overexpressed in *E. coli* (Figure S3), purified, and characterized
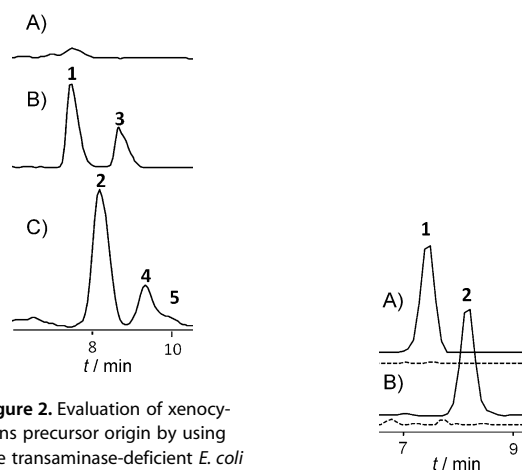

Figure 1. A) Xenocyloin biosynthesis gene cluster. Genes: *xclA*, acetolactate synthase (ALS); *xclB*, *xclC*, and *xclF*, three genes for 3-oxoacyl ACP synthase III (KS III); *xclD*, esterase; and *xclE,* alcohol dehydrogenase (ADH). B) HPLC/MS analysis of xenocyloin production in *X. bovienii*, and after heterologous expression of C) *xclA* and D) *xclABC* in *E. coli*.

not only confirmed the involvement of the cluster in the biosynthesis but also highlighted acyloin condensation as an essential step in the biosynthesis, **1**–**5** were named xenocyloin A–E, respectively. Expression of *xclABCDEF* in *E. coli* resulted in the production of **1**–**5**, thus clearly showing that the acylating activity is also encoded by one of the genes (data not shown). Moreover, labeling experiment in a transaminase-deficient *E. coli* strain also allowed elucidation of the underlying biochemistry. Tryptophan and indole-3-pyruvate are incorporated in all xenocyloins. Valine and 3-methyl-2-oxobutanoate are required for the biosynthesis of **1** and **3**, whereas isoleucine and 3-methyl-2-oxovalerate are incorporated into **2** and **4** (Table 1,

Figure 2. Evaluation of xenocyloins precursor origin by using the transaminase-deficient *E. coli* DL39 strain. Extracted ion chromatograms (EIC) for xenocyloins A–E (**1**–**5**) from heterologously expressed *xclABCDEF* in *E. coli* DL39: A) without supplementation, B) supplemented with indole-3-pyruvic acid (1 mM) and 3-methyl-2-oxobutanoic acid (3 mM), and C) supplemented with indole-3-pyruvic acid (1 mM) and 3-methyl-2-oxovaleric acid (1 mM).
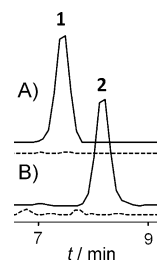

Figure 3. In vitro assay for determination of XclA activity. Condensation of indole-3-pyruvic acid with A) 3-methyl-2-oxobutyric acid and B) 3-methyl-2-oxovaleric acid resulted in the formation of **1** (*m/z* 254 [*M*+Na]$^+$) and **2** (*m/z* 268 [*M*+Na]$^+$), respectively. Dashed lines represent control experiments (without XclA).
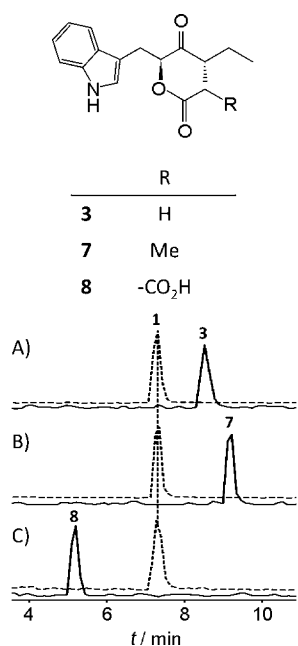
Figure 4. HPLC/MS analysis of the O-acylation reaction of **1** without (broken line) and with XclC (continuous line) in the presence of A) acetyl-CoA, B) propionyl-CoA, and C) XclB and XclC with malonyl-CoA (formation of **3**, **7**, and **8**, respectively).

in vitro with **1** as substrate. XclC was able to transfer acetyl and propionyl units from the respective CoA esters to **1** thereby resulting in formation of **3** and the novel xenocyloin **7**. Although no activity of XclB alone was detected with these substrates, XclB and XclC together were able to use malonyl-CoA as substrate to form xenocylon **8** (Figure 4). Thus, a heterodimer of XclBC might be required for malonylation, similarly to the heterodimerisation of KSα and KSβ in the type II polyketide synthase mechanism.[15,16]

In order to get more insights into the underlying mechanism, XclC was modeled by using FabH from *Stapyloccocus aureus* (PDB ID: 1ZOW) as template as this has the highest resolution (2.0 Å; Figure S4);[17] this allowed prediction of the catalytic triad as Cys118, Ser253, and Tyr283 (Figure 5 A). Simulated docking experiments with substrates **2** and acetyl-CoA (Figure 5 B) allowed prediction of a catalytic mechanism: binding of acetyl-CoA to the active site, acetyl transfer to Cys118, binding of **2** to the active site, then activation of the hydroxyl group by Tyr283 and Ser253, thus allowing nucleophilic attack of the activated hydroxyl group of **2** to the Cys118 bound acetyl moiety.[18] Separate exchange of the three active site residues to alanine resulted in severe (S253A) or complete loss (C118A and Y283A) of **3** when **1** was used as the substrate (Figure S5). Comparison of XclC with CerJ, the only proven acylating KS (involved in malonyl or methylmalonyl transfer during cervimycin biosynthesis), showed only very low se-
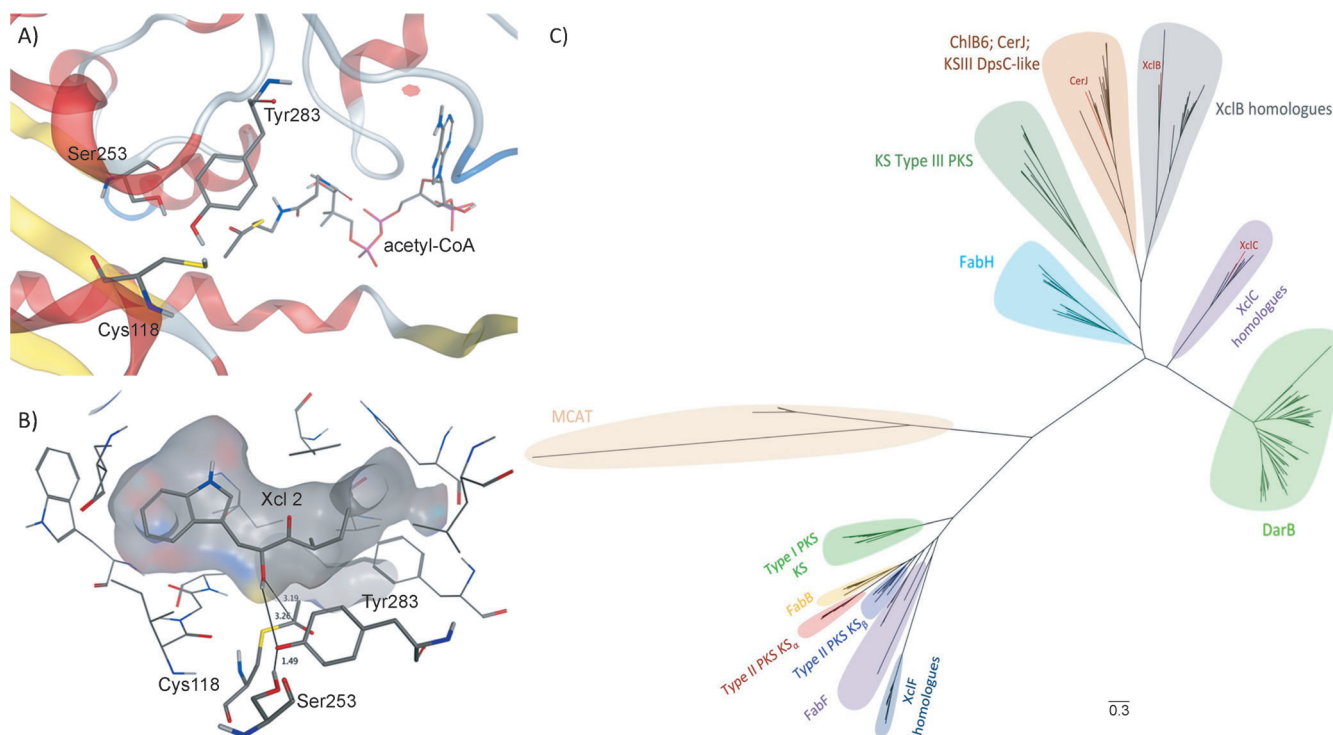
Figure 5. Residues of XclC proposed to be involved in the transfer of acetyl-CoA onto **2**. Side chains of the XclC catalytic triad (Cys118, Ser253, and Tyr283) are close to each other. A) Acetyl-CoA (coordinates from FabH in complex with acetyl-CoA, PDB ID: 3IL4) is near Cys118, thus allowing transfer of the acetyl residue to the thiol moiety. B) Compound **2** was docked into the binding site of XclC with the acetylated Cys118 and deprotonated Tyr283. The second step of the reaction is activation of the hydroxy group of **2** by Tyr283 (negative charge stabilized by Ser253); the activated hydroxy group of **2** then performs the nucleophilic attack at the carbonyl moiety of the thioester at Cys118. C) Phylogenetic tree (PhyML, http://www.atgc-montpellier.fr/phyml/) of different KSs; malonyl-CoA ACP transacylase (MCAT) as outgroup (details in Table S3). The branches for XclB, XclC, and CerJ are highlighted in red. The scale bar indicates the degree of divergence as substitutions per sequence position.

236

quence identity, including differences in the catalytic triad.[19] A phylogenetic analysis confirmed that XclC and CerJ are only distantly related and showed that XclB and XclC as well as their closest homologues (identified by BLASTP searches) form separate clades of novel KS families, thus implying that such acylating KSs might be more widespread (Figure 5C, Table S3). Interestingly, database searches revealed several *xclC* homologues close to *xclA* homologues, thus indicating that xenocylcin-like compounds and their biosynthesis mechanisms (including XclA and XclC homologues) might also be more widespread in bacterial genera (Table S3).

As **1**–**4** have been described as potent antibiotics,[9] all derivatives were tested against different Gram-positive and Gram-negative bacteria: no activity was observed at 1 mg mL^{-1}, consistent with recent reports by Nguyen et al.[20] However, **2** and **4** were active against insect hemocytes from *Galleria mellonella* (EC$_{50}$ 69 and 33 µg mL^{-1}, respectively). Microscopic analysis indicated that both compounds might interfere with the actin skeleton and thus they might contribute to the overall insecticidal activity of *X. bovienii* (Figure S6).[21] Consistent with this theory, much higher amounts of **1**–**5** are produced in the more virulent primary form of *X. bovienii* SS-2004 compared to the production in its less virulent secondary form (Figure S7).[22]

In summary, we have identified the biosynthesis gene cluster for the xenocyloins in *X. bovienii* SS-2004. Xenocyloin biosynthesis involves XclA as an *S*-selective ThDP-dependent acyloin-like condensation enzyme and two ketosynthases (XclB and XclC) that are responsible for O-acylation and thus formation of the more active xenocyloin derivatives. Because of their activity against insect hemocytes, xenocyloins might contribute to the overall virulence of *X. bovienii* against insects.

Acknowledgements

[1] S. Forst, B. Dowds, N. Boemare, E. Stackebrandt, *Annu. Rev. Microbiol.* **1997**, *51*, 47–72.
[2] H. Goodrich-Blair, D. J. Clarke, *Mol. Microbiol.* **2007**, *64*, 260–268.
[3] I. Vallet-Gely, B. Lemaitre, F. Boccard, *Nat. Rev. Microbiol.* **2008**, *6*, 302–313.
[4] H. B. Bode, *Curr. Opin. Chem. Biol.* **2009**, *13*, 224–230.
[5] J. M. Chaston, G. Suen, S. L. Tucker, A. W. Andersen, A. Bhasin, E. Bode, H. B. Bode, A. O. Brachmann, C. E. Cowles, K. N. Cowles, C. Darby, L. de Léon, K. Drace, Z. Du, A. Givaudan, E. E. Herbert Tran, K. A. Jewell, J. J. Knack, K. C. Krasomil-Osterfeld, R. Kukor, et al., *PLoS One* **2011**, *6*, e27909.
[6] L. Sundar, F. N. Chang, *J. Gen. Microbiol.* **1993**, *139*, 3139–3148.
[7] J. Li, G. Chen, J. M. Webster, E. Czyzewska, *J. Nat. Prod.* **1995**, *58*, 1081–1086.
[8] K. K. Ng, J. M. Webster, *Can. J. Plant Pathol.* **1997**, *19*, 125–132.
[9] V. J. Paul, S. Frautschy, W. Fenical, K. H. Nealson, *J. Chem. Ecol.* **1981**, *7*, 589–597.
[10] E. P. Balskus, C. T. Walsh, *J. Am. Chem. Soc.* **2009**, *131*, 14648–14649.
[11] D. Gocke, L. Walter, E. Gauchenova, G. Kolter, M. Knoll, C. L. Berthold, G. Schneider, J. Pleiss, M. Müller, M. Pohl, *ChemBioChem* **2008**, *9*, 406–412.
[12] M. Müller, D. Gocke, M. Pohl, *FEBS J.* **2009**, *276*, 2894–2904.
[13] D. Gocke, C. L. Nguyen, M. Pohl, T. Stillger, L. Walter, M. Müller, *Adv. Synth. Catal.* **2007**, *349*, 1425–1435.
[14] E. P. Balskus, C. T. Walsh, *J. Am. Chem. Soc.* **2008**, *130*, 15260–15261.
[15] C. Bisang, P. F. Long, J. Cortes, J. Westcott, J. Crosby, A.-L. Matharu, R. J. Cox, T. J. Simpson, J. Staunton, P. F. Leadlay, *Nature* **1999**, *401*, 502–505.
[16] A. Witkowski, A. K. Joshi, Y. Lindqvist, S. Smith, *Biochemistry* **1999**, *38*, 11643–11650.
[17] X. Qiu, A. E. Choudhry, C. A. Janson, M. Grooms, R. A. Daines, J. T. Lonsdale, S. S. Khandekar, *Protein Sci.* **2005**, *14*, 2087–2094.
[18] A. J. Powell, J. A. Read, M. J. Banfield, F. Gunn-Moore, S. D. Yan, J. Lustbader, A. R. Stern, D. M. Stern, R. L. Brady, *J. Mol. Biol.* **2000**, *303*, 311–327.
[19] T. Bretschneider, G. Zocher, M. Unger, K. Scherlach, T. Stehle, C. Hertweck, *Nat. Chem. Biol.* **2012**, *8*, 154–161.
[20] S. T. Nguyen, M. M. Butler, L. Varady, N. P. Peet, T. L. Bowlin, *Bioorg. Med. Chem. Lett.* **2010**, *20*, 5739–5742.
[21] S. Seo, S. Lee, Y. Hong, Y. Kim, *Appl. Environ. Microbiol.* **2012**, *78*, 3816–3823.
[22] D. R. Sugar, K. E. Murfin, J. M. Chaston, A. W. Andersen, G. R. Richards, L. de Léon, J. A. Baum, W. P. Clinton, S. Forst, B. S. Goldman, K. C. Krasomil-Osterfeld, S. Slater, S. P. Stock, H. Goodrich-Blair, *Environ. Microbiol.* **2012**, *14*, 924–939.

CHEMBIOCHEM

Supporting Information

Biosynthesis of the Insecticidal Xenocyloins in *Xenorhabdus bovienii*

Anna Proschak,[a] Qiuqin Zhou,[a] Tim Schöner,[a] Aunchalee Thanwisai,[a, b] Darko Kresovic,[a] Andrea Dowling,[c] Richard ffrench-Constant,[c] Ewgenij Proschak,[d] and Helge B. Bode*[a]

cbic_201300694_sm_miscellaneous_information.pdf

Experimental section

Bacteria strain, culture conditions and extract preparation. *Xenorhabdus bovienii* SS-2004[1] was cultivated at 30 °C and 200 rpm on a rotary shaker in 100 ml Erlenmeyer flasks containing 10 ml of Luria-Bertani (LB) broth (pH 7.0) and 2% (v/v) XAD-16 resin (Sigma-Aldrich, Deisenhofen, Germany). For fermentation this culture was inoculated with 0.1% (v/v) of a 24 h preculture. The bacterial cultures were harvested after 3 days. XAD beads were separated from the supernatant by sieving. XAD-16 resin was extracted with MeOH (10 ml) and the methanol extract was concentrated to dryness and dissolved in 1.0 ml MeOH for HPLC/MS analysis.

Cloning experiments and heterologous expression. Construction of insertions into different vectors: DNA isolation, plasmid preparation, restriction digest, PCR, gel electrophoresis and ligation reactions were performed according to standard methods.[2] PCR experiments were conducted with Phusion Polymerase (Finnzymes, Espoo, Finnland) and used according to the manufacturer´s instructions. Plasmid constructions with pJET1.2 vector and transformation into *E. coli* DH10B and DL39 strains: For construction of mutants with the complete cluster we amplified a PCR product (9.056 bp) of the desired gens (*xclABCDEF*). The PCR was performed using Phusion polymerase and primers XB_ALS-fw-NotI (5'-ATAT<u>GCGGCCGC</u>AGATAATAGTCTATGCTTTAA-3') and XB_ALS-rv-XhoI (5'-ATAT<u>CTCGAG</u>ACGAAACAATGTCTATTTTC-3') containing restriction sides *NotI* and *XhoI* (underlined). For construction of mutants with desired insertions of genes (*xclABC*, 4.570 bp) and (*xclAB*, 3.440 bp) we amplified products using XB_ALS- fw-NotI as forward primer and xbjABC-rv (5'-ATGCGAAAAGGCCGGACAGC-3') and xbjAB-rv (5'-AAGTTGG AGGATGCCAAATT CC-3') as reverse primer, respectively. PCR products were subjected to gel electrophoresis, and bands of appropriate size were purified with a QIAquick PCR Purification Kit (Qiagen, Valencia, USA). PCR product was blunt end ligated into pJET1.2/Blunt using Clone JET1.2™ PCR Cloning Kit (Fermentas, St. Leon-Rot, Germany). The resulting expression vectors (pJET1.2_*xclA*, pJET1.2_*xclABC* and pJET1.2_*xclABCDEF*) were introduced into *E. coli* DH10B strain. Additionally the plasmid construct pJET1.2_*xclABCDEF* was transformed into transaminase-deficient *E. coli* DL39 strain. Colonies were selected by growth on LB agar containing ampicillin.

Plasmid-constructions with pCOLADuet™-1/ pACYCDuet™-1 vector and transformation into *E. coli* BL21 Star (DE3) strain: For biosynthesis studies and protein expression we amplified following PCR products with primers containing restriction sides *BamHI* and *NotI*: following primer pairs were used:

XcyA_BamHI-fw (5'-ACTC<u>GGATCC</u>GATGTCAGTTAATATTGCGAAAGC-3')
XcyA_NotI-rv (5'-ACTC<u>GCGGCCGC</u>TTATTTTCCATCCAGCTCTTCC-3');
XcyB_BamHI-fw (5'-<u>ACTCGGATCC</u>GATGGAAAATAACATCTCAATTATTTCG-3')
XcyB_NotI-rv (5'-ACTC<u>GCGGCCGC</u>TTAAATATTTTGAGTCACACCACTAG-3');

1

XcyC_BamHI-fw (5'- ACTC<u>GGATCC</u> GATGACAAGAGTAGCGATTGC-3')

XcyC_NotI-rv (5'- ACTC<u>GCGGCCGC</u>TTAATATTGCATTAATACAGCGATAC-3').

Furthermore, these primers were used to amplify PCR products including *xclAB* and *xclBC* genes. After purification of the PCR products and digestion with *Bam*HI and *Not*I (Fermentas) the fragments were cloned into the pCOLADuet™-1 (Novagen) vector with kanamycin resistance. The purified PCR product of *xclA* gene was cloned into pACYCDuet™-1 vector (Novagen) with chloramphenicol resistance. The resulting plasmids were electroporated into *E. coli* BL 21 Star (DE3) strain. Colonies were selected by growth on LB agar containing kanamycin and chloramphenicol, respectively.

Reverse feeding experiments. Non-labelled precursors were fed to culture of transformed *E. coli* or *X. bovienii* strains grown in fully labelled ^{13}C medium.[3] The incorporation of precursors can be seen as shift to lower masses, dependent upon the number of carbon atoms introduced by the precursor. Bacteria were grown in 50 ml Erlenmeyer flasks containing 5 ml of ISOGRO-^{13}C (Sigma–Aldrich) medium also containing K_2HPO_4 (10 mM), KH_2PO_4 (10 mM), $MgSO_4$ x $7H_2O$ (8 mM) and $CaCl_2$ x H_2O (90 mM). Feeding cultures were inoculated with 1% of a 24 h preculture grown in LB prior to this washed with ISOGRO-^{13}C medium. Furthermore 2% (v/v) of XAD-16 was added. Supplementation was carried out with ^{12}C–precursors L-tryptophan, L-isoleucine, L-valine, (±)-3-methyl-2-oxovaleric acid sodium salt, 3-methyl-2-oxobutanoic acid sodium salt, indole-3-pyruvic acid, 2-methylbutyraldehyde, isobutyraldehyde, (±)-2-methylbutyric acid, isobutyric acid, and 3-indole acetic acid. All precursors were obtained from Sigma-Aldrich expect (1*H*-indol-3-yl)-acetaldehyde which was freshly synthetized as described previously.[4] Labelling experiments in the transaminase-deficient *E. coli* DL39 strain carrying *xclABCDEF* was carried out in ^{12}C-LB medium using the following precursor combinations: 3-methyl-2-oxovaleric acid and indole-3-pyruvic acid, 3-methyl-2-oxobutanoic acid and indole-3-pyruvic acid. The feeding solutions were added at 6, 24 and 48 h of incubation in three equal portions leading to a final concentration of 3 mM. All cultures were incubated at 30 °C on a rotary shaker at 180 rpm. After 72 h XAD-16 beads were separated from the culture and extracted with MeOH (2 x 5ml), evaporated to dryness and dissolved in 1 ml MeOH. HPLC/MS analyses were performed with diluted methanol extracts.

Protein expression and purification of XclA, XclB, and XclC. *E. coli* BL21 Star (DE3) strains carrying the expression vectors pCOLA-XclA, pCOLA-XclB, pCOLA-XcyC were used. Protein production was achieved using an auto-induction medium.[5] A single colony of bacteria was inoculated in 25 ml LB medium in 100 ml flask that contained 25 µl kanamycin (30 mg/ml). The preculture was shaken at 180 rpm and 30 °C overnight. 1% of the overnight preculture was transferred to a 1 L baffled Erlenmeyer flask with 250 ml of auto-induction medium containing 1 ml kanamycin (30 mg/ml). 250 ml auto-induction medium are composed of LB medium (231 ml), 0.25 ml of 1M $MgSO_4$, 5 ml of 50x 5052 (glycerol 25 g, glucose 2.5 g, α-lactose 10 g, distilled water 73 ml), 12.5 ml

of 20x NPS ($(NH_4)_2SO_4$ 6.6 g, KH_2PO_4 13.6 g, Na2HPO4 $2H_2O$ 17.8 g and distilled water 100 ml), adjust pH to 6.75) and kanamycin (25 mg/ml). The culture was shaken at 180 rpm at 30°C for 3h. At an OD_{600} 0.6-1.0 the culture was cooled down on ice for 30 minutes to 4 °C. Subsequently it was shaken at 16°C for about 18 h until obtaining an OD_{600} 12-15. The culture was centrifuged at 4°C, 10000 g for 30 min to obtain the pellet. All further work was done on ice. The pellet of each flask was dissolved in 12 ml of resuspension buffer (50 mM Tris HCl pH 8.0, 0.5 M NaCl) that contained 50 µl of proteinase inhibitor. Pellets were pooled together and sonicated. After cell disruption the soluble fraction was centrifuged at 4°C, 20.000 g for 45 minutes, the supernatant was transferred into new tube for protein purification and 40 µl of Benzonase® Nuclease (Novogen) (2.5U/µl) was added to the supernatant and incubated at room temperature for 10 minutes.

Protein purification was performed on a Äkta purifier (GE Healthcare, Sweden). For each protein purification, the lysate of 0.25 l of *E. coli* culture (sample volume 10 ml; filtered through a 0.45-µm syringe filter) was loaded on a 1-ml Ni Sepharose High Performance column HisTrap^TM HP column (GE Healthcare, Sweden). The proteins of interest were eluted in a step gradient with 10% (5CV), 20 % (5CV), 40 % (5CV) and 100 % of elution buffer, respectively. As binding/elution buffers the following solutions were used: 500 mM NaCl, 20 mM 500 mM Imidazol, 50 mM Tris HCl in destilled water. The purified protein was concentrated with an Amicon® Ultra 4 ml (Millipore, Germany) tube and desalting and buffer exchange was performed with PD-10 Desalting columns (GE Healthcare, USA), both according to standard protocols. Storage buffer for XclA consists of 100 mM NaCl, 50 mM Tris HCl, 1 mM EDTA, 20 % glycerol in distilled water. Storage buffer for XclB and XclC are composed of 25 mM Tris HCl, 1 mM EDTA, 1mM DTT, 20% glycerol in distilled water. All buffers were cooled down to 4 °C, adjust to pH 7.5, filtered and stored at 4°C. Purified protein fractions were used for SDS-PAGE (12%).

Enzyme activity assays

Acyloin condensation. The enzyme activity was assayed by acyloin condensation of indole pyruvate with 3-methyl-2 oxovaleric acid and 3-methyl-2-oxobutyric acid. The assay master mix consisted of purified enzyme XclA 25µl (2 µg/µl), potassium phosphate buffer (100 µl of 500 mM, pH 6.8) $MgCl_2$ (30 µl, 100mM) thiamine pyrophosphate (TPP, 20 µl, 50 mM), water (250 µl) and substrates: indole-3-pyruvic acid, 3-methyl-2 oxovaleric acid and 3-methyl-2-oxobutyric acid. The control consisted of the assay mastermix without enzyme but water. All samples were shaken at 800 rpm and 28°C overnight. The reaction was terminated by addition of aqueous HCl (1 M) and methanol.

Acyltransferase reaction. The activity of XclB and XclC was performed with **1** previously isolated and acetyl-, propionyl-, malonyl- and palmitoyl-CoA respectively. CoA substrates were diluted in buffer (50 mM Tris HCl and 100 mM KCl pH=7.5). 50 µl of substrate **1** (10 mM, diluted in methanol with 2 drops of DMSO) were mixed each with CoA-substrates (50 µl, 500 µM) water and XclB and

3

XclC respectively. The control consisted of the assay mastermix without enzyme. All samples were shaken at 800 rpm and 37°C overnight. The reaction was terminated by addition of aqueous HCl (1 M) and methanol. All samples were filtered and analysed by HPLC/MS.

***In vivo* assay with mutants (Ser253Ala, Tyr283Ala, Cys118Ala).** pCOLADuet™-1 plasmids with mutations in the *xclC* gene resulting in amino acid exchanges at Ser253Ala, Tyr283Ala, Cys118Ala in XclC were transformed into *E. coli* BL21 (DE3) Star, respectively. The control experiment was performed with *E. coli* strains only carrying pCOLADuet™-1 and pCOLADuet_XcyC-WT plasmid. 20 ml LB (1% kanamycin) were inoculated with the strains and were grown over night at 30°C and 800 rpm. 20 mL of the auto induction medium were inoculated with the overnight cultures (1:100) and cultivated until OD_{600}=0.5 at 37°C. 995 µl of each culture was mixed with 5 µl of **1** (100 mM) in a 2 ml Eppendorf tube which lid was transfixed for better aeration. The cultures were incubated at 900 rpm and 37°C for 20 h. The 1 ml probes were mixed with 1 ml methanol, concentrated until dryness and resuspended in 300 µl of methanol. The methanolic extract was diluted, filtered and analysed by HPLC. As controls methanolic solutions of **1** and **3** were analyzed.

Homology modelling. The protein sequence of XcyC was used as query for a BLASTP search [6] in the PDB [7] using default parameters. Four non-redundant proteins were selected for the homology modeling step: 2EBD (seq. identity 22%; E-value 7e-19), 4DFE (seq. identity 25%; E-value 5e-16), 1ZOW (seq. identity 23% ; E-value 6e-14), and 1UB7 (seq. identity 25% ; E-value 4e-11). These four template structures were used for multiple sequence alignment using the CLUSTALW algorithm [8] and its default parameters. The obtained alignment was employed to generate a homology model using Homology Modelling Tool integrated in MOE (Molecular Operating Environment; Chemical Computing Group Inc., Montreal, Canada). The energy of the obtained model was minimized in three steps using MMFF94x force field [9] and steepest descent optimization with 0.1 kcal/mol/Å^2 gradient: 1. Only the positions of hydrogen atoms were optimized, the positions of the heavy atoms remained fixed; 2. The side chain atoms were optimized; all backbone heavy atoms remained fixed; 3. Global optimization of all atoms. The structure of FabH (PDB code 1ZOW) was used for the generation of figures due to the highest resolution (2.0 Å).

HPLS/MS analysis. The HPLC/MS analysis of methanol extracts were performed on an Dionex UltiMate 3000 HPLC system with diode array detector coupled to a amaZon ionentrap mass spectrometer and an UPLC BEH C18/1.7 µm, 2.1 x 50 mm column (Waters). Separation was achieved with eluents: MeCN/0.1% formic acid and H_2O/0.1% formic acid and gradient range from 5 to 95% in 22 min at a flow rate of 0.6 mL/min.

4

Isolation. Xenocyloins were isolated from a 5 L culture grown in LB media with XAD-16. The cultures were shaken at 200 rpm and 30°C for 72h The XAD-16 beads were separated from supernatant extracted with MeOH, evaporated till dryness and dissolved in MeOH. Compound isolation was performed with an autopurification system (Waters) equipped with a model 3100 mass detector (Waters). The culture extract was applied to an XBridge C18 column (5μm, 21 x 150 mm, Waters). Purification was achieved with 0.1% formic acid in MeOH and 0.1% aq. formic acid; gradient, 20–95% over 30 min; flow rate, 17 mL min^{-1}. After purification we obtained slightly yellow oily compounds.

NMR analysis. NMR experiments were carried out in CDCl$_3$ on a Bruker AV400 with resonance frequencies 400 MHz for ^{1}H and ^{13}C measurements, respectively.

CD Analysis. CD spectra were recorded on a Jasco J-810 spectrometer in a 0.1 cm cuvette in methanol. The CD spectrometry allowed us to determine the absolute configuration from the direction of the Cotton effect.[10] In the experiments, the ellipticities θ in mdeg was recorded from 230-400 nm, using a concentration of 0.004 mol/L*0.1 cm cuvette. The molar ellipticities [θ] in deg.cm^{2}.dmol^{-1} and the difference in molar absorption coefficient Δε could be calculated using following equation:

$$\Delta\varepsilon = \frac{[\theta]}{3300} \text{ and } [\theta] = \frac{0.1 * \theta}{c * l}$$

Molar absorption coefficient Δε is shown in the relevant wavelength from 260-320 nm (Fig. S1). The maximum Δε for **1** were observed at λ = 291 nm, showing the same maximum absorption as the UV-Vis spectra. The positive Cotton effect indicated that compound **1** carries *S*-configuration at the stereo-center.

Phylogenetic analysis. A PHYML tree (200 bootstraps) was calculated using a ClustalW alignment (gap opening: 10; gap extension: 0.1), which was performed with the collected ketosynthases. For visualization and calculation of the alignment as well as the PHYML tree the Geneious software (Biomatters Ltd., New-Zeeland) was used.[11],[8]

5
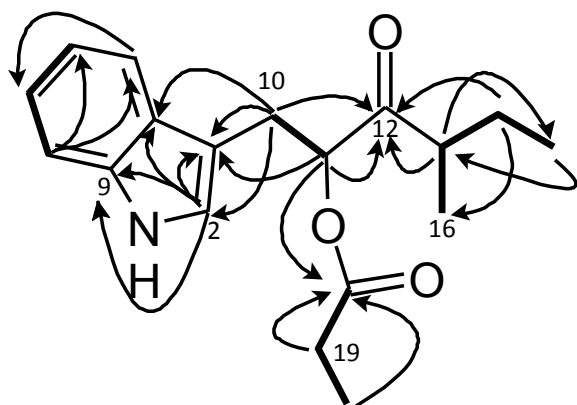
Table S1. ^{13}C, ^{1}H NMR data (400 MHz) for **5** in $CDCl_3$. COSY correlations are shown in bold and HMBC correlations are shown with arrows.

	δ_C (ppm)	δ_H (ppm)	Multiplicity (J [Hz])
1		8.03	s, NH, 1H
2	123.02	7.07	s, 1H
5	118.9	7.66	d, 1H (7.8)
6	122.3	7.22	t, 1H (7.1)
7	120.3	7.15	t, 1H (7.1)
8	111.1	7.37	d, 1H (8.1)
10	26.6	3.31	dd, 1H (4.6, 15.1)
10	26.6	3.20	dd, 1H (8.2, 15.2)
11	77.6	5.48	q, 1H (4.7)
12	210.0		
13	44.2	2.64	m, 2H
14	26.3	1.75	m, 1H
14	26.3	1.39	m, 1H
15	11.5	0.88	t, 3H (7.5)
16	15.0	0.99	d, 3H (6.8)
18	173.9		
19	27.7	2.37	m, 1H
20	9.0	1.10	t, 3H (7.5)

6

Table S2. Postulated xenocyloin biosynthesis gene cluster (*xcl*) in *X. bovienii* SS-2004 and proposed protein functions and similarity to the closest homologue.

Protein	original gene name	Size [aa]	Proposed Function	Origin	Closest homologue Identities/ Positives [%]	Accession number
XclA	Xbj1_3901	565	Acetolactate synthase large subunit	*P. carotovorum* subsp. *carotovorum* WPP14	50/67	ZP_03832235.1
XclB	Xbj1_3900	343	3-oxoacyl-[acyl-carrier-protein] synthase III	*P. carotovorum* subsp. *carotovorum* WPP14	46/66	ZP_03832234.1
XclC	Xbj1_3898	326	3-oxoacyl-[acyl-carrier-protein] synthase III	*Paenibacillus lactis* 154	55/73	ZP_09002357
XclD	Xbj1_3897	288	putative esterase	*P. luminescens* subsp. *laumondii* TTO1	89/92	NP_931508
XclE	Xbj1_3896	369	Alcohol dehydrogenase classe III	*P. luminescens* subsp. *laumondii* TTO1	97/98	NP_931507
XclF	Xbj1_3894	558	Putative 3-oxoacyl-[acyl-carrier-protein] synthase III	*X. nematophila* ATCC19061	75/86	YP_003714026

7

Table S3. Ketosynthases used for the phylogenetic tree (Fig. 5c). The sequences are ordered according to their location in the respective branches. XclB homologes encoded in an operon and/or adjacent to *xclA* homologues are highlighted in green.

	Protein	Organism	Accession number
	Closest BLAST-P hits for XclB		
1	3-Oxoacyl-ACP synthase III	B. sp. EniD312	WP_009111263.1
2	XclB	X. bovienii SS-2004	
3	3-Oxoacyl-ACP synthase III	A. nasoniae	CBA73264.1
4	3-Oxoacyl-ACP synthase III	P. carotovorum	WP_010301235.1
5	3-Oxoacyl-ACP synthase III	P. pacifica	WP_006975318.1
6	3-Oxoacyl-ACP synthase III	C. stagnale	YP_007317906.1
7	3-Oxoacyl-ACP synthase III	N. punctiforme	YP_001865628.1
8	3-Oxoacyl-ACP synthase III	R. sp. PCC 7116	YP_007056099.1
9	3-Oxoacyl-ACP synthase III	S. cyanosphaera	YP_007130807.1
10	3-Oxoacyl-ACP synthase III	C. sp. PCC 6303	YP_007138278.1
11	3-Oxoacyl-ACP synthase III	N. punctiforme	YP_001868566.1
12	3-Oxoacyl-ACP synthase III	N. punctiforme	YP_001865657.1
13	3-Oxoacyl-ACP synthase III	A. cylindrica	YP_007155727.1
14	3-Oxoacyl-ACP synthase III	C. sp. PCC 7822	YP_003899922.1
15	3-Oxoacyl-ACP synthase III	R. sp. PCC 7116	YP_007057764.1
	Closest BLAST-P hits for XclC		
16	3-Oxoacyl-ACP synthase	C. acetobutylicum	NP_347450.1
17	3-Oxoacyl-ACP synthase	P. lactis	WP_007130623.1
18	XclC	X. bovienii SS-2004	
19	3-Oxoacyl-ACP synthase	B. thuringiensis	YP_006930640.1
20	3-Oxoacyl-ACP synthase	B. sp. 1NLA3E	YP_007911827.1
21	3-Oxoacyl-ACP synthase	O. scapharcae	WP_010098042.1
22	3-Oxoacyl-ACP synthase	P. polymyxa	YP_003947618.1
23	3-Oxoacyl-ACP synthase	P. polymyxa	YP_003871436.1
24	3-Oxoacyl-ACP synthase	P. sp. Aloe-11	WP_007431139.1
25	3-Oxoacyl-ACP synthase	P. terrae	YP_005077926.1
26	3-Oxoacyl-ACP synthase	P. peoriae	WP_010345468.1
	Closest BLAST-P hits for XclF		
27	3-Oxoacyl-ACP synthase	X. nematophila	WP_010848687.1
28	3-Oxoacyl-ACP synthase	X. nematophila	YP_003714026.1
29	3-Oxoacyl-ACP synthase	P. leiognathi	WP_008989540.1
30	3-Oxoacyl-ACP synthase	P. sp. AK15	WP_007465048.1
31	3-Oxoacyl-ACP synthase	P. angustum	WP_005364526.1
32	3-Oxoacyl-ACP synthase	M. sp. PE36	WP_006034384.1
33	3-Oxoacyl-ACP synthase	R. blandensis	WP_008043745.1
34	3-Oxoacyl-ACP synthase	P. damselae	WP_005305524.1
35	3-Oxoacyl-ACP synthase	P. sp. SKA34	WP_006644045.1
36	3-Oxoacyl-ACP synthase	P. profundum	YP_132684.1
	MCAT		
37	malonyl CoA-ACP transacylase	X. bovienii	YP_003468748.1

38	ACP S-malonyltransferase	X. bovienii	YP_003468584.1
39	malonyl-CoA-ACP transacylase	X. nematophila	YP_003712936.1
40	acyltransferase	X. nematophila	YP_003712007.1
41	polyketide synthase	X. nematophila	YP_003711953.1
42	malonyl CoA-ACP transacylase	P. luminescens	CAE15208.1
43	malonyl CoA-ACP transacylase	P. aeruginosa	WP_015649132.1
44	malonyl CoA-ACP transacylase	P. aeruginosa	YP_001347565.1
45	malonyl CoA-ACP transacylase	P. aeruginosa	NP_251658.1
46	malonyl CoA-ACP transacylase	P. aeruginosa	YP_790205.1
47	malonyl CoA-ACP transacylase	P. asymbiotica	YP_003040539.1
48	malonyl CoA-ACP transacylase	P. asymbiotica	CAQ83795
49	malonyl CoA-ACP transacylase	E. coli	NP_415610.1
50	malonyl CoA-ACP transacylase	E. coli	EDU66705.1
51	malonyl CoA-ACP transacylase	E. coli	YP_001457936.1

CerJ; KSIII DpsC-like KS

52	CerJ	S. tendae	AEI91069
53	ChlB3	S. antibioticus	AAZ77676
54	ChlB6	S. antibioticus	AAZ77679
55	CosE	S. olindensis	ABC00733
56	DpsC	S. peucetius	AAA65208
57	AknE2	S. sp. SPB74	ZP_04991255.1
58	AknE2	S. galilaeus	AAF70109
59	BAB72048	S. galilaeus	BAB72048
60	CalO4	S. aurantiaca	ZP_01462124
61	PokM2	S. diastatochromogenes	ACN64832
62	FabH	S. erythraea	YP_001107471
63	NdasDRAFT_3133	N. dassonvillei	ZP_04334033.1
64	AviN	S. viridochromogenes	AAK83178
65	CalO4	M. echinospora	AAM70354
66	PlaP2	S. sp. Tu6071	ABB69750
67	CouN2	S. rishiriensis	AAG29787
68	CloN2	S. roseochromogenes	AAN65231

KS type III PKS

69	RppA S	S. antibioticus	BAB91443
70	RppA	S. avermitilis	NP_828307
71	RppB	S. antibioticus	BAB91444
72	MXAN_6639	M. xanthus	YP_634756
73	PKS10	M. tuberculosis	NP_216176
74	PKS11	M. tuberculosis	NP_216181
75	CHS H. (PLN03173)	H. androsaemum	Q9FUB7
76	Chs-like	R. baltica	NP_868579
77	CHS9	M. sativa	AAA02827
78	BPS (PLN03172)	H. androsaemum	Q8SAS8
79	STS	P. quinquefolia	AAM21773
80	BAS	R. palmatum	AAK82824

FabH

81	FabHA	B. subtilis	NP_389015

9

82	FabHB	B. subtilis	NP_388898
83	FabH	N. punctiforme	YP_001865657
84	FabH	P. luminescens	NP_930069
85	FabH	A. fabrum	NP_354198
86	FabH	E. coli	NP_287225
87	3-oxoacyl-ACP synthase	B. subtilis	NP_389015.1
88	FabH	S. griseus	YP_001826619
89	FabH	S. echinatus	AAV84077
90	NP_626634	S. coelicolor A3(2)	NP_626634
91	FabH	S. avermitilis	BAC73499
92	Q54206	S. glaucescens	Q54206
93	FdmS	S. griseus	AAQ08929
94	CAM58805_S._sp._BenQ	S. sp. A2991200	CAM58805
95	ZhuH 1MZJ	S. sp. R1128	AAG30195
96	FrnI	S. roseofulvus	AAC18104
97	AlnI	S. sp CM020	ACI88883

KS type I PKS

98	Plu1885	P. luminescens	NP_929153
99	NanA8	S. nanchangensis	AAP42874
100	EryAII	S. erythraea	YP_001102990
101	TylGI KSQ	S. fradiae	AAB66504
102	MerA	S. violaceusniger	ABJ97437
103	TamAI	S. sp. 3079	ADC79637
104	OleAI KSQ	S. antibioticus	AAF82408
105	HedT	S. griseoruber	AAP85336

FabB

106	AntD (Plu4191)	P. luminescens	NP_931374
107	EncA	S. maritimus	AAF81728
108	ActiB	S. coelicolor A3(2)	SCO5088
109	NcnA	S. arenae	AAD20267
110	TcmK	S. davawensis	CCK26894
111	SimA1	S. antibioticus	AAK06784

Type II PKS KS alpha

112	TcmL	S. glaucescens	AAA67516
113	SimA2	S. antibioticus	AF324838_4
114	EncB	S. maritimus	AAF81729
115	ActIA	S. coelicolor A3(2)	SCO5087
116	NcnB	S. arenae	AAD20268

Type II PKS KS beta

117	FabF	P. luminescens	NP_930065
118	FabF	E. coli	NP_287229
119	FabF	B. subtilis	NP_389016
120	FabF	N. punctiforme	YP_001867862
121	FabF	T. thermophilus	YP_143679
122	NP_344945	S. pneumoniae	NP_344945
123	NP_645683	S. aureus	NP_645683

10

124	FabF	E. albertii	ZP_02902779.1
125	NP_415613	E. coli	NP_415613
126	FabF	S. avermitilis	BAC70003

FabF

127	FabF	M. sp. 4-46	YP_001771620
128	FabF	C. pinensis	ACU62401
129	cpin1855	C. pinensis	YP_003121552
130	Dfer_1997	D. fermentans	YP_003086385
131	FabB	A. pleuropneumoniae	ZP_00134992
132	FabB	C. sp. 30_2	ZP_04562837
133	NP_416826	E. coli	NP_416826
134	FabB	S. boydii	YP_001881145

DarB

135	Aeqsu_0932	A. sublithincola	YP_006417450
136	Arnit_2310	A. nitrofigilis	YP_003656468
137	ATCC33389_0196	A. aphrophilus	EGY32238
138	azo0292 DarB	A. sp. BH72	YP_931796
139	Slit_0359	S. lithotrophicus	YP_003522988
140	SMGD1_1386	S. gotlandica	EHP29910
141	BFO_3187	T. forsythia	YP_005015826
142	Weevi_1554	W. virosa	YP_004238832.1
143	BZARG_2045	B. argentinensis	ZP_08820341
144	Zobellia_2074	Z. galactanivorans	YP_004736513
145	CAPGI0001_0843	C. gingivalis	ZP_04056582
146	CAPSP0001_1216	C. sputigena	ZP_03390203
147	CBGD1_514	S. gotlandica	ZP_05070248
148	Varpa_2231	V. paradoxus	YP_004154548
149	Vapar_3389	V. paradoxus	YP_002945272
150	Plu2164	P. luminescens	NP_929424
151	PspoU_010100018642	P. spongiae	ZP_10300425
152	CHU_0390	C. hutchinsonii	YP_677020
153	Coch_0547	C. ochracea	YP_003140666
154	COI_2002	M. haemolytica	ZP_05992665
155	COK_0379	M. haemolytica	ZP_05988513
156	Sulba_2257	S. barnesii	YP_006405107
157	cpin6850	C. pinensis	YP_003126452
158	DaAHT2_1139	D. alkaliphilus	YP_003690456
159	DarB	P. chlororaphis	AAN18032
160	Daro_2368	D. aromatica	YP_285574
161	Sdel_2118	S. deleyianum	YP_003305165
162	Dfer_5797	D. fermentans	YP_003090150
163	DP1817	D. psychrophila	YP_065553
164	EIKCOROL_00456	E. corrodens	ZP_03712789
165	Fcol_11845	F. columnare	YP_004942963
166	FF52_12311	F. sp. F52	ZP_10481912
167	Fjoh_1102	F. johnsoniae	YP_001193454
168	FJSC11DRAFT_3961	F. sp. JSC-11	ZP_08987753
169	Fluta_1447	F. taffensis	YP_004344279

11

170	FP2279	F. psychrophilum	YP_001297136
171	GCWU000324_02596	K. oralis	ZP_04603113
172	Hbal_2902	H. baltica	YP_003061270
173	HMPREF0156_01383	B. taxon 274 str. F0058	ZP_06983320
174	HMPREF0204_10987	C. gleum	ZP_07085127
175	HMPREF0604_01363	N. mucosa	ZP_07993739
176	HMPREF1028_00835	N. sp. GT4A_CT1	ZP_08888860
177	HMPREF1051_1749	N. sicca	EIG27057
178	HMPREF1154_2288	C. sp. CM59	ZP_10880679
179	HMPREF1319_0374	C. ochracea	EJF43732
180	HMPREF1320_1701	C. taxon 335 str. F0486	EJF37460
181	HMPREF1321_1154	C. taxon 412 str. F0487	ZP_10366882
182	HMPREF1977_1456	C. ochracea	ZP_07866642
183	HMPREF9016_01947	N. taxon 014 str. F0314	ZP_06980826
184	HMPREF9064_0174	A.segnis	ZP_07888807
185	HMPREF9071_0527	C. taxon 338 str. F0234	ZP_08201061
186	HMPREF9335_01583	A. aphrophilus	EHB89432
187	HMPREF9370_1914	N. wadsworthii	ZP_08940206
188	HMPREF9371_1043	N. shayeganii	ZP_08886538
189	HMPREF9401_0244	A. butzleri	ZP_07890833
190	HMPREF9417_0595	H. parainfluenzae	ZP_08147854
191	HMPREF9418_1128	N. macacae	ZP_08684521
192	HMPREF9711_01694	M. odoratimimus	EKB04829
193	HMPREF9712_01161	M. odoratimimus	ZP_09523568
194	HMPREF9716_01579	M. odoratimimus	EKB07937
195	HMPREF9952_1824	H. pittmaniae	ZP_08755481
196	Lacal_2074	L. sp. 5H-3-7-4	YP_004580348
197	Lbys_1508	L. byssophila	YP_003997574
198	M446_0174	M. sp. 4-46	YP_001767187
199	MDS_0597	P. mendocina	YP_004378380
200	MICAG_1820011	M. aeruginosa	CCI22605
201	MldDRAFT_4065	delta proteobacterium MLMS-1	ZP_01289639
202	Mucpa_6793	M. paludis	ZP_09618305
203	Myrod_1723	M. odoratus	ZP_09672239
204	Rfer_3974	R. ferrireducens	YP_525203
205	NEIFL0001_0036	N. flavescens	ZP_04757628
206	NEIFLAOT_02523	N. flavescens	ZP_03720660
207	NEISICOT_02133	N. sicca	ZP_05318975
208	NEISUBOT_03200	N. subflava	ZP_05983976
209	NiasoDRAFT_0547	N. soli	ZP_09632794
210	NP_645683	S. aureus	NP_645683
211	NT05HA_1737	A. aphrophilus	YP_003008155
212	O3I_37171	N. brasiliensis	ZP_09843377
213	Oweho_0889	O. hongkongensis	YP_004988545
214	PA-RVA6-3077	P. asymbiotica	CAR66906
215	ParcA3_010100003428	P. arctica	ZP_10280196
216	PAU_02401	P. asymbiotica	YP_003041237
217	Pchl3084_3967	P. chlororaphis	EJL05977
218	PchlO6_4243	P. chlororaphis	ZP_10172862
219	PMI10_02641	F. sp. CF136	ZP_10730768

12

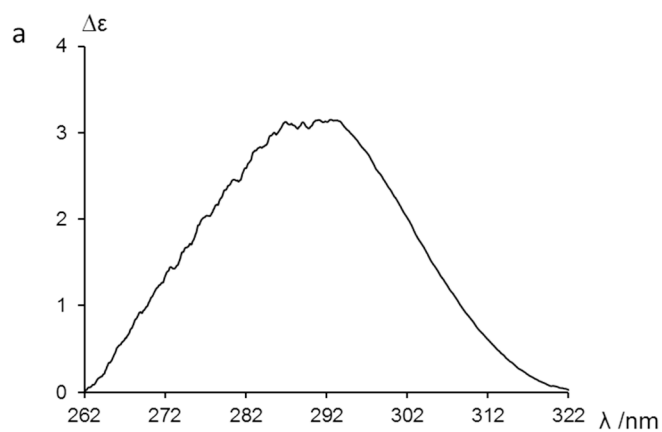| 222 | PMI12_02025 | V. sp. CF313 | ZP_10567997 |
223	PMI13_02465	C. sp. CF314	ZP_10726507
224	PMI20_00702	P. sp. GM17	ZP_10707840
225	Psefu_0435	P. fulva	YP_004472512
226	PSJM300_17945	P. stutzeri	AFN79642

Figure S1. CD-spectrum of **1** (a) indicate a positive Cotton effect at 291 nm with Δε + 3.0. The positive value of the optical rotation indicates *S*-configuration.
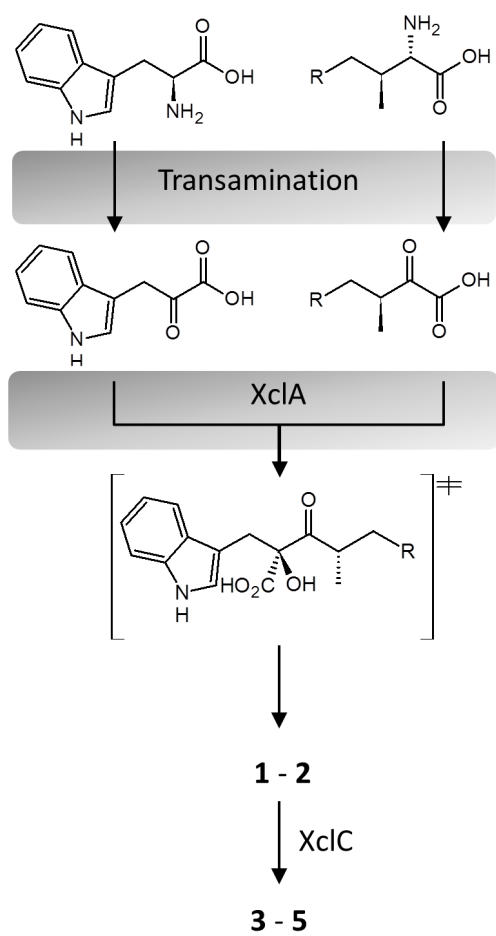
13

Figure S2. Proposed biosynthesis pathway of xenocyloins A-E (**1-5**).
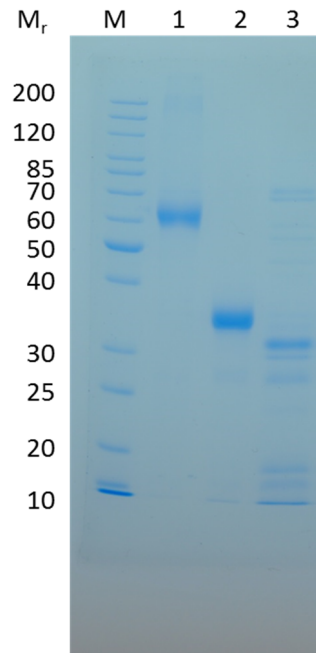
Figure S3. SDS-PAGE (12 %) showing purified His-tagged proteins (His)$_6$-XclA (62 kDa) (lane 1), (His)$_6$-XclB (38 kDa) (lane 2) and (His)$_6$-XclC (36 kDa) (lane 3).

15
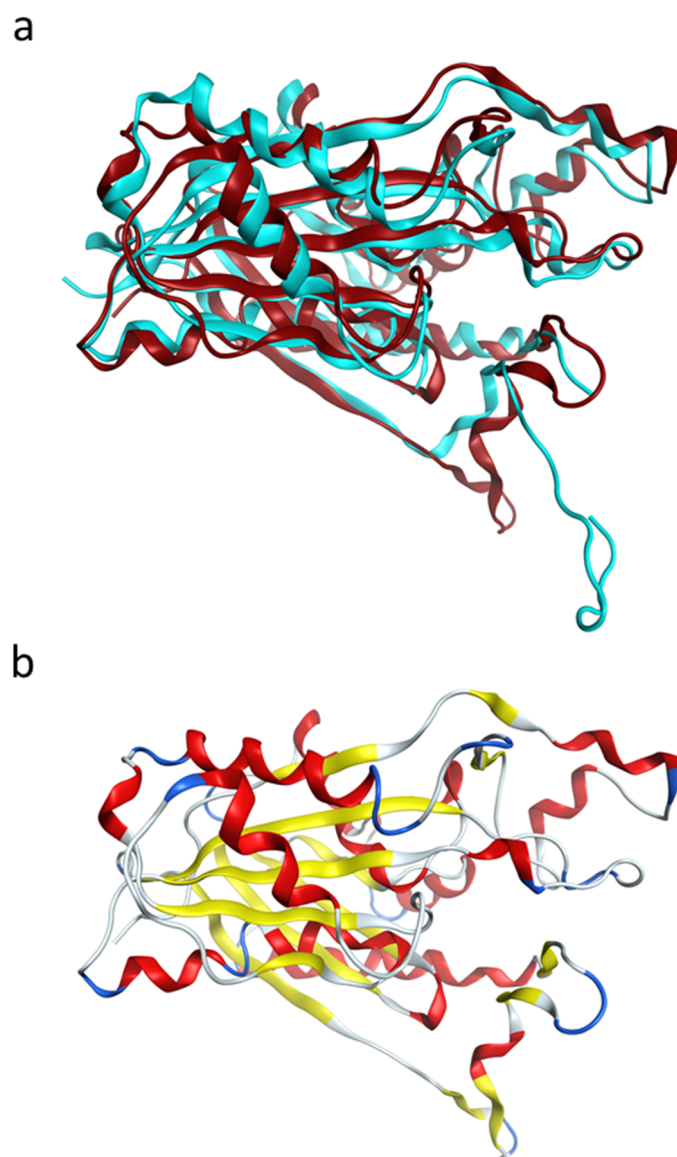
Figure S4. Superposition of the homology model of XclC (red) with the template structure (FabH, PDB code: 1ZOW, cyan) (a) and cartoon representation of the homology model of XclC (b). α-Helices (red), β-sheets (yellow) and turns (blue) are shown.
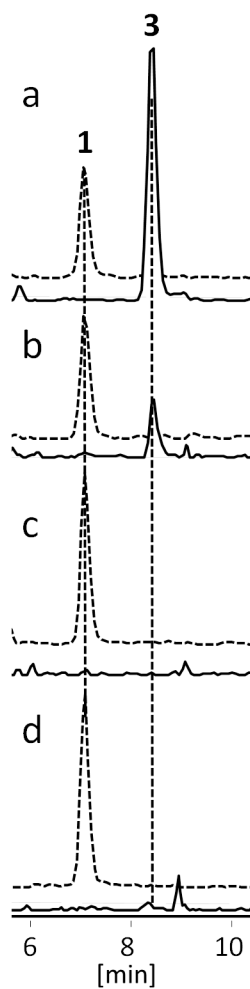
Figure S5. *In vivo* acetylation assay of wt XclC and XclC mutants. Extracted ion chromatograms (EIC) of **1** (m/z 254 [M+Na]$^+$) (dashed line) and **3** (m/z 296 [M+Na]$^+$) (continuous line) in *E. coli* culture extracts transformed with pCOLA_XclC_WT (a), pCOLA_XclC_S253A (b), pCOLA_XclC_C118A (c) and pCOLA_XclC_Y283A (d).

17
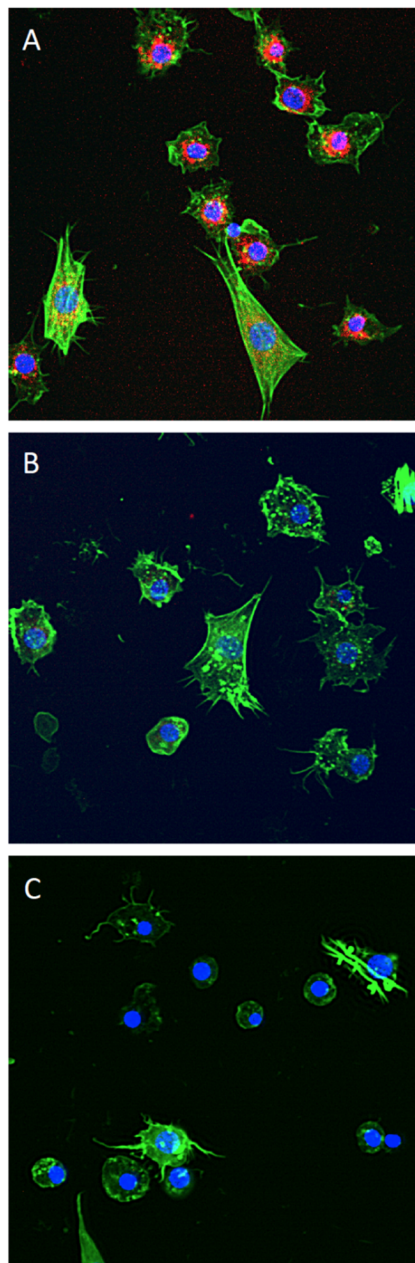
Figure S6. Hemocyte morphology after treatment with **2** and **4**. A) Control hemocytes incubated with 2% DMSO, B) hemocytes treated with 100 µg mL^{-1} of **2** and C) of **4**. Actin polymerisation is strongly reduced with **2** and especially **4** and only a few viable mitochondria remain present in the treated cells. Nucleus (blue), filamentous actin cytoskeleton (green), viable mitochondria (red).

18
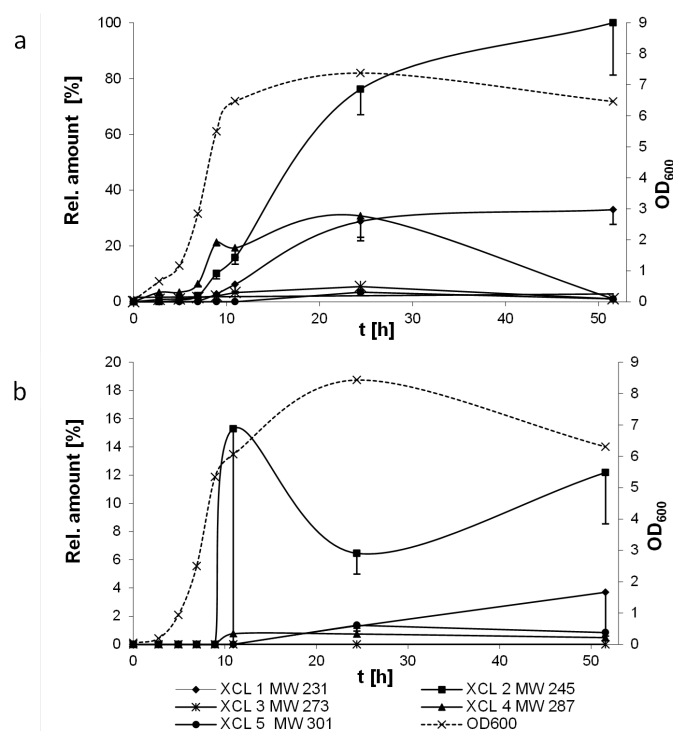
Figure S7. Xenocyloin production in the primary orange form F1 (a) and the secondary cream colored form F2 (b) of *X. bovienii* SS-2004 strain.[12] 100% refers to the maximum production of the main compound **2**. The experiments were carried out in triplicates.

References

[1] R. J. Akhurst, N. E. Boemare, *J. Gen. Microbiol.* **1988**, *134*, 1835–1845.
[2] D. Reimer, E. Luxenburger, A. O. Brachmann, H. B. Bode, *Chembiochem* **2009**, *10*, 1997–2001.
[3] A. Proschak, K. Schultz, J. Herrmann, A. J. Dowling, A. O. Brachmann, R. ffrench-Constant, R. Müller, H. B. Bode, *Chembiochem* **2011**, *12*, 2011–2015.
[4] G. Revelant, S. Dunand, S. Hesse, G. Kirsch, *Synthesis* **2011**, *2011*, 2935–2940.
[5] F. W. Studier, *Protein Expr. Purif.* **2005**, *41*, 207–234.
[6] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, *J. Mol. Biol.* **1990**, *215*, 403–410.
[7] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucleic Acids Res.* **2000**, *28*, 235–242.
[8] M. A. Larkin, G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, *Bioinformatics* **2007**, *23*, 2947–2948.
[9] T. A. Halgren, *J. Comput. Chem.* **1996**, *17*, 490–519.
[10] G. Snatzke, *Angew. Chem.* **1968**, *80*, 15–26.
[11] S. Guindon, O. Gascuel, *Syst. Biol.* **2003**, *52*, 696–704.
[12] D. R. Sugar, K. E. Murfin, J. M. Chaston, A. W. Andersen, G. R. Richards, L. deLéon, J. A. Baum, W. P. Clinton, S. Forst, B. S. Goldman, et al., *Environ. Microbiol.* **2012**, *14*, 924–939.
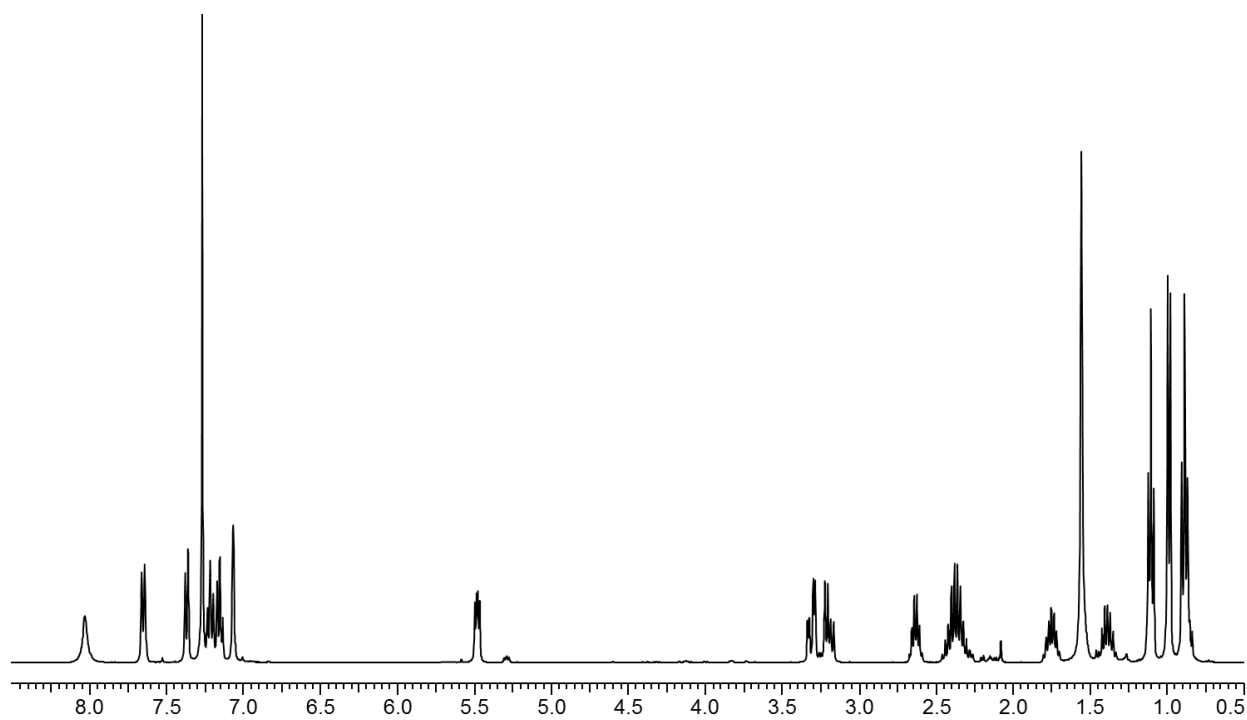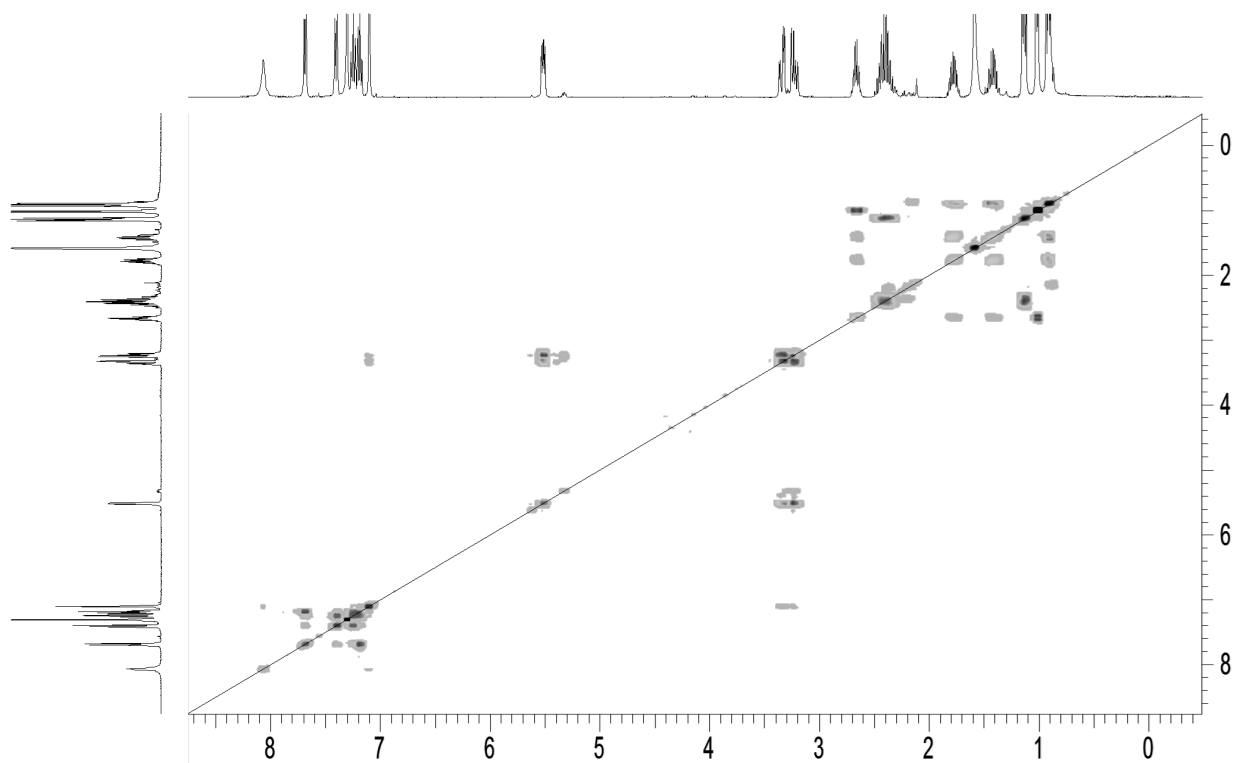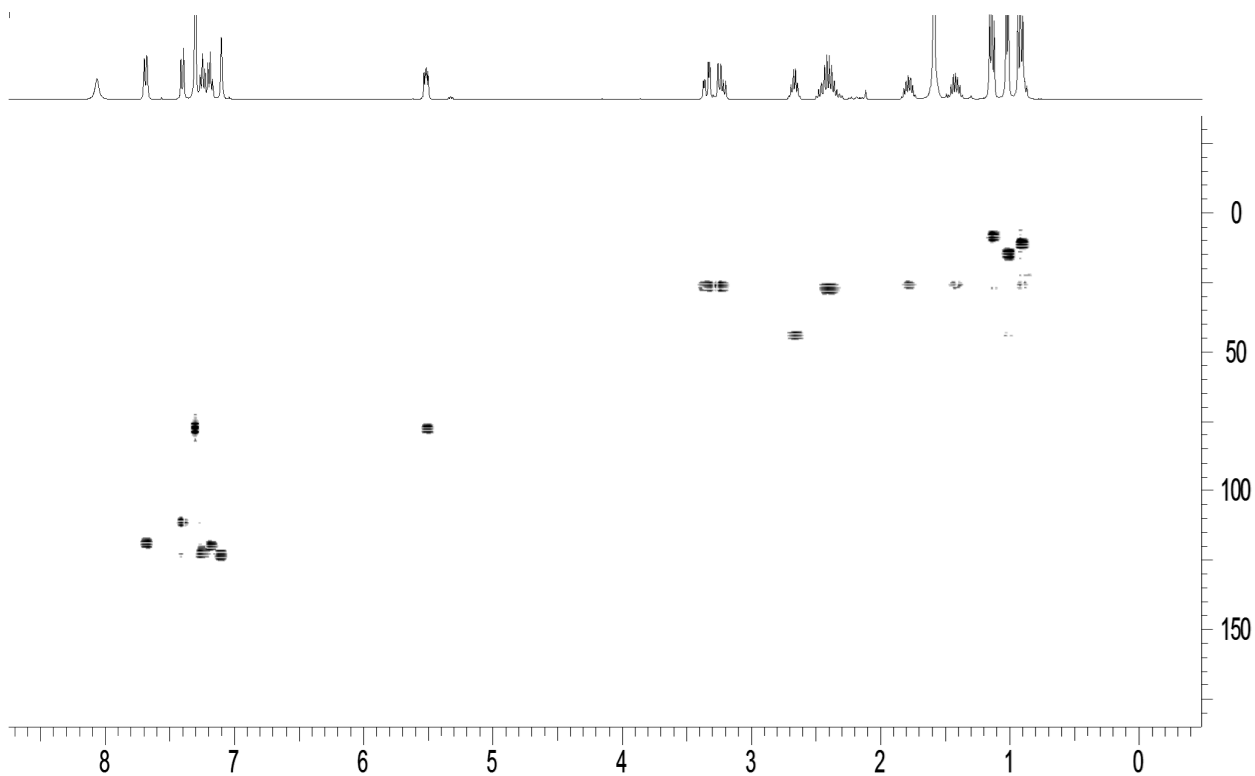
Annex I: [1]H-spectrum of compound **5**.
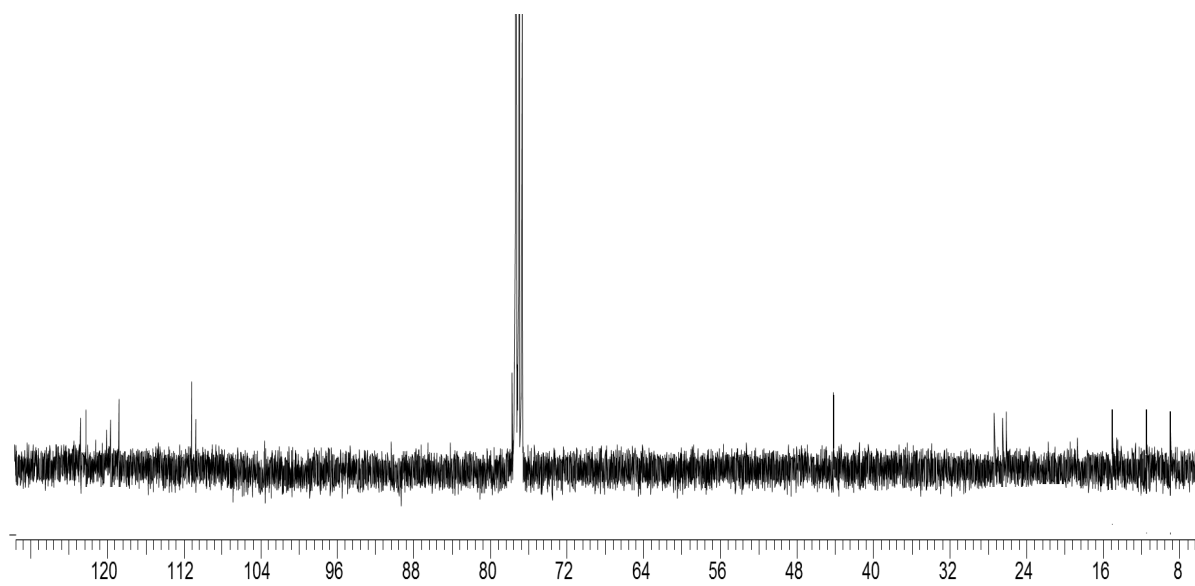
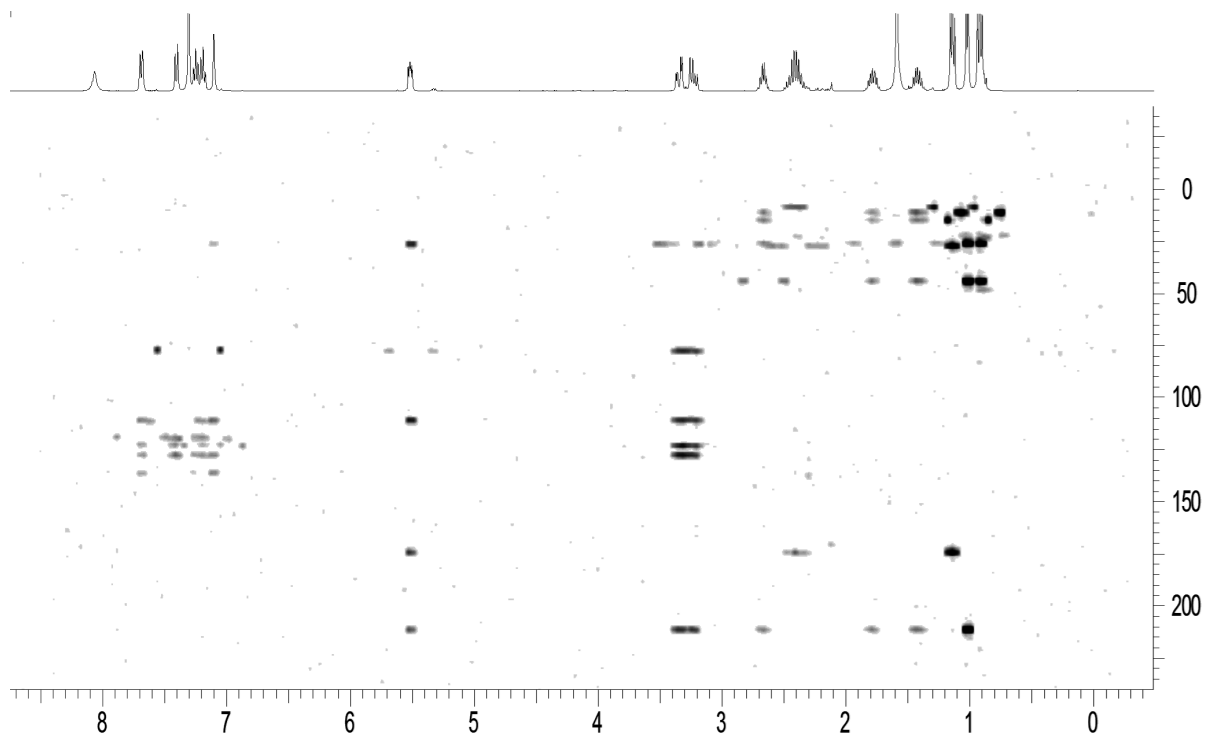**Annex II:** [13]C Spectrum of compound **5**.

20

Annex III: ^{1}H-^{1}H-COSY spectrum of compound **5**.

Annex IV: ^{1}H-^{13}C-HSQC spectrum of compound **5.**

21

Annex V: ^{1}H-^{13}C-HMBC spectrum of compound **5**.

11.5. Simple "On-Demand" Production of Bioactive Natural Products

Authors: Edna Bode,[a] Alexander O. Brachmann,[a] Carsten Kegler,[a] Rukayye Simsek,[b] Christina Dauth,[a] Qiuqin Zhou,[a] Marcel Kaiser,[c] Petra Klemmt,[b] and Helge B. Bode[a, d]

[a] Merck Stiftungsprofessur für Molekulare Biotechnologie, Fachbereich Biowissenschaften, Goethe Universität Frankfurt, Max-von-Laue-Strasse 9, 60438 Frankfurt am Main (Germany)

[b] Institute for Cell Biology and Neuroscience, Molecular Cell Biology and Human Genetics, Goethe University, Max-von-Laue-Strasse 13, 60438 Frankfurt (Main) (Germany)

[c] Swiss Tropical and Public Health Institute, Parasite Chemotherapy, Socinstrasse 57, 4002 Basel (Switzerland)

[d] Buchmann Institute for Molecular Life Sciences (BMLS), Goethe Universität Frankfurt, Max-von-Laue-Strasse 15, 60438 Frankfurt am Main (Germany)

Contribution: Isolation and identification of a new xenorhabdin derivative.

Attachment: the paper.

Simple "On-Demand" Production of Bioactive Natural Products

Edna Bode,[a] Alexander O. Brachmann,[a] Carsten Kegler,[a] Rukayye Simsek,[b] Christina Dauth,[a] Qiuqin Zhou,[a] Marcel Kaiser,[c] Petra Klemmt,[b] and Helge B. Bode*[a, d]

Exchange of the native promoter to the arabinose-inducible promoter P_{BAD} was established in entomopathogenic bacteria to silence and/or activate gene clusters involved in natural product biosynthesis. This allowed the "on-demand" produc- tion of GameXPeptides, xenoamicins, and the blue pigment in- digoidine. The gene clusters for the novel "mevalagmapep- tides" and the highly toxic xenorhabdins were identified by this approach.

Introduction

The introduction of natural products and derivatives thereof into medicine as antibacterial, antifungal, and anticancer com- pounds has greatly improved human health over the last 60 years.[1,2] Despite this success, the natural-product pipeline is running dry, as natural-product research is very expensive and time-consuming and (in summary) not as profitable as the development of medications for chronic diseases or life-style drugs.[3] This is especially dramatic as resistance is a natural and ancient trait that will always result in anti-infectives be- coming useless over time.[4]

Over the past ten years natural-product research has greatly benefited from progress in sequencing technology and analyti- cal chemistry.[5–7] Various genome-sequencing projects have revealed that natural-product producers have the capacity to produce many more compound classes than were previously known, and mass spectrometry-based analysis of well-studied natural-product producers has shown that several additional compounds (and classes) are usually produced; these have been overlooked because of analytical limitations.[8–10] The

combination of this knowledge with new molecular tools allows heterologous expression of formerly silent or cryptic biosynthesis gene clusters, thereby leading to the identification of even more natural products.[11,12] Deletion or induction of regulatory proteins identified in genome sequences also re- sults in the production of novel compounds.[6,13] A third strat- egy is the exchange of the natural promoter of a gene cluster of interest to a strong constitutive promoter or an inducible promoter whose regulation is well understood.[14–17]

We previously used the last strategy to induce the produc- tion of the blue pigment indigoidine from *Photorhabdus lumi- nescens*.[14] As this pigment is easily detectable by the naked eye but its biosynthesis is regulated by unknown factors, it represented an ideal model for such approaches. Indigoidine was produced after exchange of the natural promoter to either the strong constitutive promoter *rpsM* or to *cipB*, which is acti- vated in the stationary phase.

Here we describe the simple application of an inducible system by using the well-known P_{BAD} promoter for the produc- tion of several natural products from the entomopathogenic bacteria *P. luminescens* and *Xenorhabdus doucetiae*.[18,19] Induci- ble systems allow the analysis of induced versus non-induced conditions, and thus can mimic a "knock-out" and an "over- producing mutant" in a single strain grown under two different conditions leading to a more reliable identification of natural products.

Results and Discussion

For activation of secondary metabolite biosynthesis gene clus- ters, a pCEP (cluster expression plasmid) vector was construct- ed based on the integrative plasmid pDS132 and the expres- sion plasmid pBAD30 (Figure S1 in the Supporting Informa- tion). The pCEP derivatives carried the first 300–600 bp of the desired biosynthesis gene, thus enabling the integration of the plasmid at the start site of the target biosynthesis gene. Ho- mologous integration results in an exchange of the respective promoter (Scheme 1, Figure S2). When this system was tested

[a] *E. Bode,+ Dr. A. O. Brachmann,+ Dr. C. Kegler, C. Dauth, Q. Zhou,*
     Prof. Dr. H. B. Bode
     Merck Stiftungsprofessur für Molekulare Biotechnologie
     Fachbereich Biowissenschaften, Goethe Universität Frankfurt
     Max-von-Laue-Strasse 9, 60438 Frankfurt am Main (Germany)
     E-mail: h.bode@bio.uni-frankfurt.de
     Homepage: http://www.bio.uni-frankfurt.de/48050101/ak-bode

[b] *R. Simsek, Dr. P. Klemmt*
     Institute for Cell Biology and Neuroscience
     Molecular Cell Biology and Human Genetics, Goethe University
     Max-von-Laue-Strasse 13, 60438 Frankfurt (Main) (Germany)

[c] *M. Kaiser*
     Swiss Tropical and Public Health Institute, Parasite Chemotherapy
     Socinstrasse 57, 4002 Basel (Switzerland)

[d] *Prof. Dr. H. B. Bode*
     Buchmann Institute for Molecular Life Sciences (BMLS)
     Goethe Universität Frankfurt
     Max-von-Laue-Strasse 15, 60438 Frankfurt am Main (Germany)

[+] *These authors contributed equally to this work.*

     Supporting information for this article is available on the WWW under
     http://dx.doi.org/10.1002/cbic.201500094.
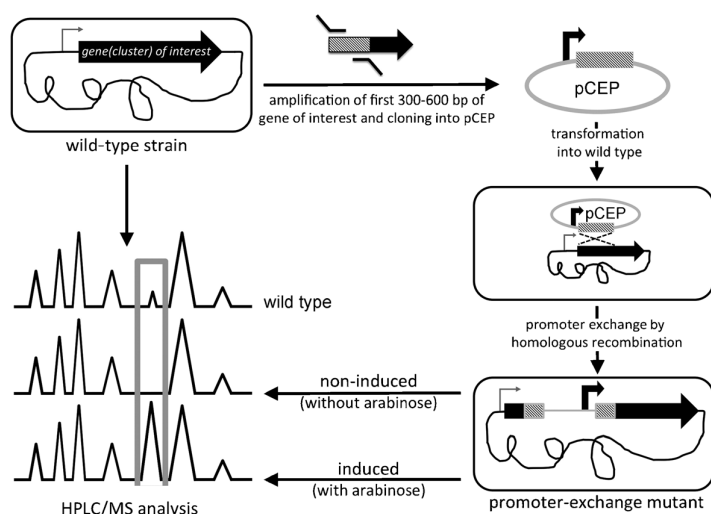
Scheme 1. The promoter-exchange approach. The first 300–600 bp of the gene of interest are amplified by PCR and cloned into plasmid pCEP, which is then transformed into the strain carrying the respective gene. Homologous recombination of the non-replicating plasmid results in the formation of a strain in which the expression of the full-length gene is not driven by the natural promoter (gray arrow) but by the introduced inducible promoter (black arrow; here P_{BAD}). The resulting promoter-exchange mutant can be selected based on the pCEP-encoded resistance gene. As the new promoter is tightly controlled and shows no activity without the inducer (here arabinose), the non-induced strain behaves like a knock-out mutant (no production of the compound of interest). With inducer, overexpression of the desired gene is achieved, thus resulting in an overproducing mutant (relative to the wild-type strain).
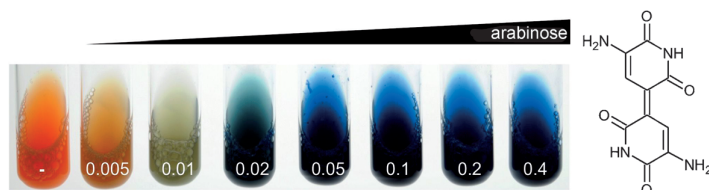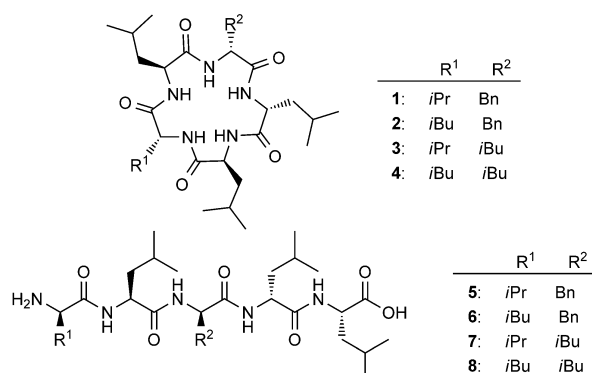

Figure 1. Indigoidine structure and production in *P. luminescens* by pCEP induction (P_{BAD}–indC) with L-arabinose (values shown are percentage concentrations).

with *indC* (encoding the non-ribosomal peptide synthetase (NRPS) responsible for indigoidine biosynthesis),[14] indigoidine production was observed in the *P. luminescens* pCEP-indC strain, proportional to the amount of the L-arabinose (P_{BAD} inducer; 0.005–0.05%; Figure 1). Similarly, plasmid pCEP-gxpS (Figure S2 A) was constructed and introduced into *P. luminescens* thereby leading to the production of GameXPeptides A–D (**1**–**4**;[20] Table S3). Interestingly linear GameXPeptides (**5**–**8**) were also detected, thus indicating that the speed of the peptide synthesis might be too fast for the correct thioesterase-catalyzed cyclization, and probably faster hydrolytic release. The structures of these linear compounds were elucidated by MS, and all structures were confirmed by solid-phase peptide synthesis (Figure S3). Similarly, cyclic and linear products were also observed when the strong constitutive promoter rpsM was used (P_{rpsM}–gxpS): even higher production rates were detected (up to 15.5 and 12.8 mg L^{-1} for **1** and **3**; Table S4).

In general, the promoter-exchange approach resulted in a tenfold increase in production for all compounds; there was no significant difference between the strong inducible and constitutive promoters. However, the great advantage of inducible promoters (Figure 2 A) is that without the inducer (here L-arabinose) essentially no production takes place, or it is below the detection level (similar to gene disruption by plasmid insertion; compare Figure 2 A and B). Thus, in cases in which the promoter is tightly regulated, as in the case of P_{BAD}, the inducible promoter allows analysis of the equivalent of a knock-out mutant (without the arabinose inducer) or overexpression (with inducer) from the same strain.

Next, we exchanged the promoter of a unknown biosynthesis gene cluster in *P. luminescens* by introducing plasmid pCEP-map (Figure S4, Table S5); this cluster has similarity to that for the previously described insecticidal rhabdopeptides in *Xenorhabdus nematophila*.[21] Without arabinose the loss of a polar compound **9** was observed, relative to wild-type; it was strongly produced upon addition of arabinose, together with a second compound (**10**; Figure 2 C). Isolation of **9** by preparative HPLC/MS allowed its structure elucidation by 1D (^{1}H,^{13}C) and 2D (COSY, HSQC, HMBC) NMR experiments (Supporting Information). The compound was named mevalagmapeptide A, as it is composed of N-methyl valine and valine connected to an agmatine moiety at the C terminus (Table S6). As no epimerization domain is present in the NRPS MapABC (Figure S4) the absolute configurations of all amino acids are predicted to be L, as confirmed by advanced Marfey's analysis (data not shown). Mevalagmapeptide B (**10**) differed from **9** only in the terminal amine, as concluded from the fragmentation pattern and HRMS data. Putrescine was the terminal amine, as has previously been found in the biosynthesis of bicornutins from *Xenorhabdus budapestensis*.[10]

	R^1	R^2
1:	*i*Pr	Bn
2:	*i*Bu	Bn
3:	*i*Pr	*i*Bu
4:	*i*Bu	*i*Bu

	R^1	R^2
5:	*i*Pr	Bn
6:	*i*Bu	Bn
7:	*i*Pr	*i*Bu
8:	*i*Bu	*i*Bu

We also applied the promoter exchange approach to *X. doucetiae* DSM 17909 (GenBank entry: FO704550.1) for the biosyn-
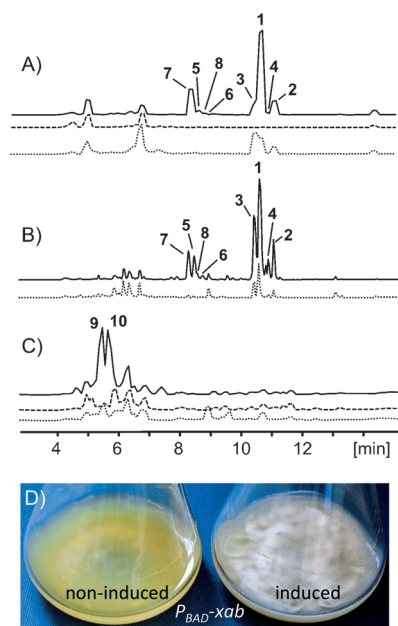
Figure 2. A)–C) HPLC/MS chromatograms of *P. luminescens* with arabinose induction or constitutive overexpression (continuous lines), without arabinose induction (dashed lines), and wild-type (dotted line). A) P_{BAD}-*gxpS*, B) P_{rpsM}-*gxpS*, C) P_{BAD}-*map*. EIC for the respective compounds are shown for A), and base-peak chromatograms are shown for (B) (*m/z* 500–650) and (C) (*m/z* 100–1000). EIC analysis for the specific *m/z* ranges in (B) and (C) show no production of compounds without arabinose induction; thus the P_{BAD} system is tightly regulated. All chromatograms are drawn to the same scale. D) Xenoamicin production in *X. doucetiae* based on P_{BAD}-*xab*. Xenoamicin (10 mg L^{-1} in the induced culture) is visible as white crystals floating on top of the culture.

9: R = ![structure] NH$_2$, 10: R = H
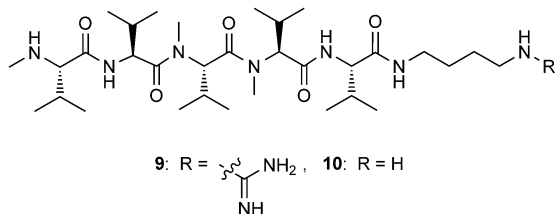
thesis gene clusters involved in GameXPeptide (pCEP-gxpS) and xenoamicin[22] production (pCEP-xab; Figure S5). In the latter case, the production was so good that white needles containing xenoamicin C (**11**) as the main compound were observed in the culture (Figure 2 D), whereas no production was visible in the non-induced culture. Additionally, we exchanged the promoter of a biosynthesis gene cluster (*xrdA–xrdJ*;
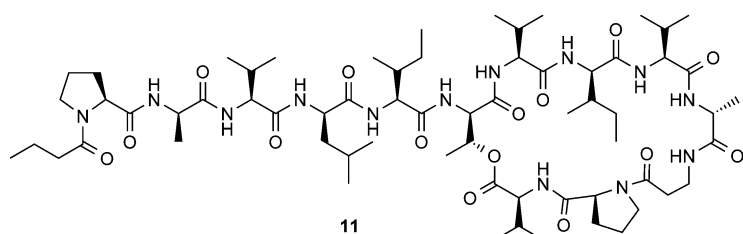

11

Figure 3) with similarity to the holomycin biosynthesis gene cluster in *Streptomyces clavuligerus*,[23] and indeed without arabinose the loss of yellow compounds was observed. Upon the addition of arabinose, the production of these yellow compounds was restored, and further derivatives were also detected (Figure 3). Isolation and structure elucidation of two main compounds allowed their identification as xenorhabdin-2 (**12**) and -5 (**13**), based on NMR data comparison.[24] The structures of the additional derivatives **14**–**17** (including the new derivatives **16** and **17**) were elucidated according to the characteristic dithiolopyrrolone fragment, which allowed simple differentiation between non-methylated (*m/z* 173.8) and *N*-methylated core structures (*m/z* 187.8; Figure S6). This was confirmed by the incorporation of labeled methionine and the presence of a leucine-derived iso-acyl moiety, as determined from the incorporation of deuterated leucine (Figure S7).

The gene clusters for dithiolopyrrolone biosynthesis have been identified in *S. clavuligerus*[23] and *Yersinia ruckeri*,[25] as well as the related thiomarinol derivatives in *Pseudoalteromonas* species.[26–28] Biochemically, the best-characterized pathway is that of holomycin from *S. clavuligerus*, with the enzymatically characterized acyltransferase HlmA and the dithiol oxidase HlmI that is responsible for disulfide bond formation.[29] The *xrd* biosynthesis gene cluster in *X. doucetiae* is very similar (Table S7) but without an HlmI homologue. The homologue might be encoded elsewhere in the genome but it could not be found using a homology search.

Dithiolopyrrolones have been described as potent antibiotics against MRSA strains, most likely by targeting RNA polymerase.[27,30,31] Although xenorhabdins might indeed function as antibiotics in the bacterium–nematode–insect relationship (killing other insect-associated bacteria and thus food competitors), we thought that they might also target eukaryotic cells, as recently suggested for pyrroloformamide.[32] Therefore, we tested **12** and **13** (as well as **9**) against the causative agents of tropical diseases as well as for their cytotoxic effects against L6 (rat skeletal muscle) cells (Table 1).[33] Both xenorhabdins were highly toxic against all tested organisms but their very high cytotoxicity against mammalian cells makes therapeutic application very unlikely. Additionally, **13** was analyzed for its influence on metabolic activity (WST-1) and DNA synthesis (BrdU incorporation). The IC$_{50}$ values were 4.0 and 1.5 µM, respectively (Figure S8). Moreover, treatment with **13** affected MCF7 (human breast cancer) cell adhesion and morphology in a dose-dependent manner. Cells were subjected to **13** at different concentrations (around IC$_{50}$) for 24 h followed by immunofluorescent staining; this revealed altered cell morphology (changes in the distribution of the actin cytoskeleton and mitochondria; Figure S8).

Conclusions

In summary, we used arabinose-inducible promoters in entomopathogenic bacteria for the overproduction of desired natural products, thus enabling their isolation and structure elucidation. The advantage of this system is the tightly regulated loss (without inducer)
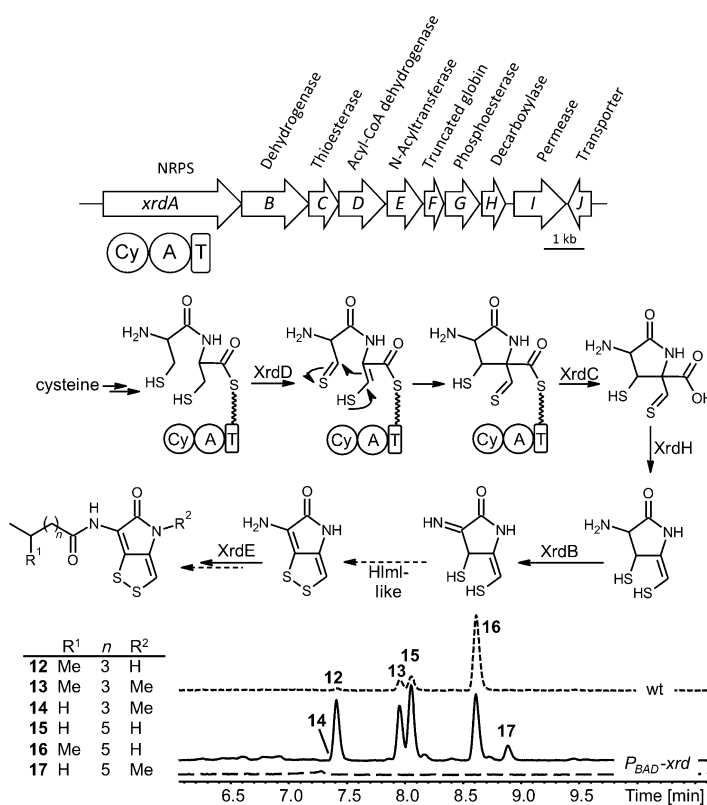
Figure 3. Biosynthesis gene cluster for xenorhabdin biosynthesis (top), proposed xenorhabdin biosynthesis (middle), and HPLC chromatograms (380–420 nm) of *X. doucetiae*: wild-type (wt), P_{BAD}-xrdA non-induced (dashed), and induced (continuous line).

tions (often regarded as "silent" under the selected conditions) will increase our knowledge about natural products and their biological functions, as we have shown here for different compound classes including the xenorhabdins. Interestingly, xenorhabdins are not only highly active antibiotics but also show high activity against eukaryotic cells. Regarding the ecological niche of the bacterium, this might be an efficient strategy to "kill two birds with one stone" with potent bioactivity (competing bacteria and eukaryotes as well as the insect prey).

Acknowledgements

Table 1. Bioactivity of **9**, **12,** and **13** against different protozoa and mammalian cells.

	IC$_{50}$ [µM]			
	9	**12**	**13**	Ref.[a]
Trypanosoma brucei rhodesiense STIB900	129.7	0.021	0.010	0.01
Trypanosoma cruzci Tulahuen C4	118.0	0.218	0.483	1.73
Leishmania donovani	60.7	0.412	0.412	0.460
Plasmodium falciparum NF 54	38.4	4.36	1.84	0.006
Rat L6 cells	>150	0.197	0.107	0.02

[a] Reference compounds (positive controls): melarsoprol for *T. brucei rhodesiense*, benznidazole for *T. cruzi*, miltefosin for *L. donovani*, chloroquine for *P. falciparum* NF 54, and podophyllotoxin for L6 cells.

or overproduction (with inducer) of the desired compound in a single strain. This approach is superior to heterologous expression of desired gene clusters, as the biosynthetic capacity to produce all building blocks for the production of the desired compound is available in the original producer. Thus it ensures that "real" natural products are produced rather than non-natural derivatives as a consequence of missing precursors (e.g., amino acids or fatty acids).

The wide applicability of this approach for the activation of gene clusters that are not active under standard growth condi-

[1] D. J. Newman, G. M. Cragg, *J. Nat. Prod.* **2012**, *75*, 311–335.
[2] G. M. Cragg, D. J. Newman, *Biochim. Biophys. Acta Gen. Subj.* **2013**, *1830*, 3670–3695.
[3] F. von Nussbaum, M. Brands, B. Hinzen, S. Weigand, D. Häbich, *Angew. Chem. Int. Ed.* **2006**, *45*, 5072–5129; *Angew. Chem.* **2006**, *118*, 5194–5254.
[4] V. M. D'Costa, C. E. King, L. Kalan, M. Morar, W. W. L. Sung, C. Schwarz, D. Froese, G. Zazula, F. Calmels, R. Debruyne, G. B. Golding, H. N. Poinar, G. D. Wright, *Nature* **2011**, *477*, 457–461.
[5] H. B. Bode, R. Müller, *Angew. Chem. Int. Ed.* **2005**, *44*, 6828–6846; *Angew. Chem.* **2005**, *117*, 6988–7007.
[6] J. Piel, *Annu. Rev. Microbiol.* **2011**, *65*, 431–453.
[7] A. O. Brachmann, H. B. Bode, *Adv. Biochem. Eng./Biotechnol.* **2013**, *135*, 123–155.
[8] R. D. Kersten, P. C. Dorrestein, *ACS Chem. Biol.* **2009**, *4*, 599–601.
[9] R. D. Kersten, Y.-L. Yang, Y. Xu, P. Cimermancic, S.-J. Nam, W. Fenical, M. A. Fischbach, B. S. Moore, P. C. Dorrestein, *Nat. Chem. Biol.* **2011**, *7*, 794–802.
[10] S. W. Fuchs, C. C. Sachs, C. Kegler, F. I. Nollmann, M. Karas, H. B. Bode, *Anal. Chem.* **2012**, *84*, 6948–6955.
[11] O. Schimming, F. Fleischhacker, F. I. Nollmann, H. B. Bode, *ChemBioChem* **2014**, *15*, 1290–1294.
[12] K. Yamanaka, K. A. Reynolds, R. D. Kersten, K. S. Ryan, D. J. Gonzalez, V. Nizet, P. C. Dorrestein, B. S. Moore, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 1957–1962.
[13] A. A. Brakhage, V. Schroeckh, *Fungal Genet. Biol.* **2011**, *48*, 15–22.
[14] A. O. Brachmann, F. Kirchner, C. Kegler, S. C. Kinski, I. Schmitt, H. B. Bode, *J. Biotechnol.* **2012**, *157*, 96–99.
[15] J. B. Biggins, X. Liu, Z. Feng, S. F. Brady, *J. Am. Chem. Soc.* **2011**, *133*, 1638–1641.
[16] J. B. Biggins, H.-S. Kang, M. A. Ternei, D. DeShazer, S. F. Brady, *J. Am. Chem. Soc.* **2014**, *136*, 9484–9490.
[17] J. Kennedy, G. Turner, *Mol. Gen. Genet.* **1996**, *253*, 189–197.
[18] E. Duchaud, C. Rusniok, L. Frangeul, C. Buchrieser, A. Givaudan, S. Taourit, S. Bocs, C. Boursaux-Eude, M. Chandler, J.-F. Charles, E. Dassa, R.

Derose, S. Derzelle, G. Freyssinet, S. Gaudriault, C. Médigue, A. Lanois, K. Powell, P. Siguier, R. Vincent, et al., *Nat. Biotechnol.* **2003**, *21*, 1307–1313.

[19] P. Tailliez, S. Pagès, N. Ginibre, N. Boemare, *Int. J. Syst. Evol. Microbiol.* **2006**, *56*, 2805–2818.

[20] H. B. Bode, D. Reimer, S. W. Fuchs, F. Kirchner, C. Dauth, C. Kegler, W. Lorenzen, A. O. Brachmann, P. Grün, *Chem. Eur. J.* **2012**, *18*, 2342–2348.

[21] D. Reimer, K. N. Cowles, A. Proschak, F. I. Nollmann, A. J. Dowling, M. Kaiser, R. ffrench-Constant, H. Goodrich-Blair, H. B. Bode, *ChemBioChem* **2013**, *14*, 1991–1997.

[22] Q. Zhou, F. Grundmann, M. Kaiser, M. Schiell, S. Gaudriault, A. Batzer, M. Kurz, H. B. Bode, *Chem. Eur. J.* **2013**, *19*, 16772–16779.

[23] B. Li, C. T. Walsh, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 19731–19735.

[24] B. V. McInerney, R. P. Gregson, M. J. Lacey, R. J. Akhurst, G. R. Lyons, S. H. Rhodes, D. R. J. Smith, L. M. Engelhardt, A. H. White, *J. Nat. Prod.* **1991**, *54*, 774–784.

[25] Z. Qin, A. T. Baker, A. Raab, S. Huang, T. Wang, Y. Yu, M. Jaspars, C. J. Secombes, H. Deng, *J. Biol. Chem.* **2013**, *288*, 14688–14697.

[26] D. Fukuda, A. S. Haines, Z. Song, A. C. Murphy, J. Hothersall, E. R. Stephens, R. Gurney, R. J. Cox, J. Crosby, C. L. Willis, T. J. Simpson, C. M. Thomas, *PLoS ONE* **2011**, *6*, e18031.

[27] A. C. Murphy, D. Fukuda, Z. Song, J. Hothersall, R. J. Cox, C. L. Willis, C. M. Thomas, T. J. Simpson, *Angew. Chem. Int. Ed.* **2011**, *50*, 3271–3274; *Angew. Chem.* **2011**, *123*, 3329–3332.

[28] A. C. Murphy, S.-S. Gao, L.-C. Han, S. Carobene, D. Fukuda, Z. Song, J. Hothersall, R. J. Cox, J. Crosby, M. P. Crump, C. M. Thomas, C. L. Willis, T. J. Simpson, *Chem. Sci.* **2014**, *5*, 397.

[29] B. Li, C. T. Walsh, *Biochemistry* **2011**, *50*, 4615–4622.

[30] Z. Qin, S. Huang, Y. Yu, H. Deng, *Mar. Drugs* **2013**, *11*, 3970–3997.

[31] B. Li, W. J. Wever, C. T. Walsh, A. A. Bowers, *Nat. Prod. Rep.* **2014**, *31*, 905–923.

[32] P. A. Abreu, T. S. Sousa, P. C. Jimenez, D. V. Wilke, D. D. Rocha, H. P. S. Freitas, O. D. L. Pessoa, J. J. La Clair, L. V. Costa-Lotufo, *ChemBioChem* **2014**, *15*, 501–506.

[33] F. Grundmann, M. Kaiser, M. Schiell, A. Batzer, M. Kurz, A. Thanwisai, N. Chantratita, H. B. Bode, *J. Nat. Prod.* **2014**, *77*, 779–783.

266

CHEMBIOCHEM

Supporting Information

Simple "On-Demand" Production of Bioactive Natural Products

Edna Bode,[a] Alexander O. Brachmann,[a] Carsten Kegler,[a] Rukayye Simsek,[b] Christina Dauth,[a] Qiuqin Zhou,[a] Marcel Kaiser,[c] Petra Klemmt,[b] and Helge B. Bode*[a, d]

cbic_201500094_sm_miscellaneous_information.pdf

Supporting information

Materials and methods

Cultivation of strains:

Photorhabdus luminescens ssp. *laumondii* TT01[1] (rifampicin resistant strain), *Xenorhabdus doucetiae* DSM 17909[2] (ampicillin resistant strain) and promoter exchange mutants thereof were routinely cultivated in Luria-Bertani (LB) broth (pH 7.0) at 30°C and 200 rpm on rotary shaker and on LB agar at 30°C. Appropriate antibiotics were added to LB liquid and agar cultures when necessary at following concentrations: ampicillin 100 µg/ml, rifampicin 50 µg/ml, chloramphenicol 34 µg/ml, and kanamycin 50 µg/ml . *E. coli* S17-1λ*pir* (Tp[r] Sm[r] *recA thi hsdR* RP4-2-Tc::Mu-Km::Tn7, λ *pir* phage lysogen) was used for conjugation. All production cultures of promoter exchange mutants were inoculated with a 24 h preculture of the same medium (0.1 %, v/v).

All feeding experiments were applied in either ISOGRO® ^{15}N medium, ISOGRO® ^{13}C medium, ISOGRO® ^{13}C medium supplemented with ^{12}C amino acids or LB medium supplemented with deuterated amino acids as described prieviously.[3-5] In general, production cultures were grown in 50 mL Erlenmeyer flasks containing 5 mL LB broth and 2 % Amberlite® XAD-16. The recombined *ara*BAD-promoter strains were induced with 0.2 % L-arabinose after three hours and destined cultures were also supplemented with amino acids to a final concentration of 2 mM. The XAD-16 resin was harvested after two days and extracted with methanol. The crude extracts were analyzed by means of MALDI-MS and HPLC-MS (Bruker AmaZon).

Construction of pCEP:

The vector pCEP was derived from fusing the vector backbones of pDS132[6] and pBAD30[7]. For pBAD30 a fragment of 1.3 kbp was amplified by PCR with primers pBAD_CEP_Eco147I_Fw and pBAD_CEP_XbaI_Rv, including the coding sequences for *araC* and *ara*BAD promoter. In addition, restriction sites *Nde*I, *Pae*I, *Sac*I, *Pst*I, *Xho*I, *Smi*I, *Pme*I, *Ehe*I, *Xba*I and a Shine Dalgarno box were introduced by Primer pBAD_CEP_XbaI_Rv (Figure and Table). The fragment was subcloned into vector pJET1.2 (Thermo/Fermentas) and subsequently digested via restriction sites *Xba*I and *Eco*147I. The vector pDS132 was digested in same way yielding a fragment of 5.8 kbp, including origin of replication R6Kγ, *traI*, origin of transfer (oriT) and chloramphenicol resistance (*cat*). The obtained fragments were recombined by

ligation and the resulting conjugatable vector pCEP-Cm was introduced into *E. coli* S17-1 λ*pir* by electroporation. The verification of vector pCEP-Cm (5887 bp) was tested by digestion with *Bam*HI, yielding three fragments of 637 bp, 1724 bp and 3526 bp. For construction of pCEP-Km, the chloramphenicol resistance gene was exchanged against a kanamycin resistance gene using Gibson cloning (New England Biolabs).

Construction of pCEP, pCK, and pCK_rpsM vectors for *Photorhabdus* and *Xenorhabdus* promoter mutants.

All pCEP promoter exchange vectors were constructed by using the genomic DNA sequence of *P. luminescens*, initiating at the start codon of the respective gene and comprising 530 to 750 bp. The corresponding forward primers where artificially complemented to constitute, together with the start codon (ATG), an *Nde*I (CATATG) restriction site. The reverse primers were artificially complemented with either a *Sac*I or *Pst*I restriction site. In the following: primers Prom_IndC_Fw and Prom_IndC_Rv were used to amplify fragment *plu2186*, primers CEP_Agm_Rhabdopep_Fw and CEP_Agm_Rhabdopep_Rv to amplify fragment *plu0897*, primers CEP_plu3263_NdeI and CEP_plu3263_SacI to amplify fragment *plu3263* from *P. luminescens*, and primer EB_XDV3_110151_PstI_fw and EB_XDV3_110151_XbaI_rv to amplify fragment *gxpS,* EB_XDV3_70458_PstI_fw and EB_XDV3_70458_XbaI_rv to amplify fragment xabA, and EB_XDV3_10499_PstI_fw and EB_XDV3_10499_XbaI_rv to amplify fragment xrdA from *X. doucetiae* (Table S1). All PCR amplicons were subcloned into vector pJET1.2 (Thermo/Fermentas) and subsequently digested and cloned into the vector pCEP via restriction sites *Nde*I and either *Sac*I or *Pst*I. The resulting pCEP vectors are listed in Table S2. The pCK_rpsM vector[8] was constructed in a similar way using primers 3263_NdeI_for and 3263_MluI_rev to amplify the fragment *plu3263* of *P. luminescens*.

Conjugation

P. luminescens and *X. doucetiae* were mated with *E. coli* S17-1λ*pir* carrying the respective promoter exchange vector. Both strains were grown in LB broth to an OD_{600} of 0.6 to 0.7 and the cells were washed once with fresh LB. Subsequently, the donor and recipient strain were mixed on a LB agar plate in a ratio of 1:3 and

269

incubated at 37°C for 3 hours followed by incubation at 30°C overnight. The next day, the bacterial cell layer was harvested with help of an inoculating loop and resuspended in fresh LB broth. Serial dilutions were spread out on selective LB agar plates with rifampicin (alternatively ampicillin for *Xenorhabdus*) and chloramphenicol and incubated at 30°C for 2 days. Individual clones were analyzed by means of HPLC-MS and the genotype was verified for all mutants by using plasmid and genome specific oligonucleotides as described previously.

MS-Analysis, quantification and compound isolation:

MS analysis and structure elucidation based on labelling experiments and MS analysis was performed as described previously.[4,5] Briefly, L-[*methyl*-^{2}H$_3$]methionine and L-[^{2}H$_{10}$]leucine were added to growing cultures of *X. doucetiae* P_{BAD}::*xrd* in LB with 2% XAD-16 at the time of arabinose induction (6 h after inoculation) at 2 mM and cultivation was continued for 2 d before harvest and extraction of the XAD resin.

For quantification cultures were grown in triplicates with XAD-16 and after 48 h, the XAD was harvested by sieving and extracted with MeOH as described previously. Extracts were quantified by means of HPLC-MS (Bruker AmaZon) against isolated or synthesized standards.

For the isolation of **12** and **13**, *X. doucetiae* P_{BAD}-*xrdA* was grown in 5 x 5L Erlenmeyer flasks containing 1L of LB medium each with 2% XAD-16. Arabinose (0.2%) was added 6h after inoculation and cultivation was continued for 48 h. After harvest of the XAD-16, it was extracted with MeOH and the obtained crude extract after evaporation of the solvent was subjected to preparative HPLC/MS (Waters Autopurification system) resulting in pure **12** and **13** as described previously for other compounds. Mevalagmapeptide A (**9**) was obtained using the same protocol from *P. luminescens* P_{BAD}-*map*.

Solid phase peptide synthesis (SPPS) of the linear pentapeptides (5-8): Anchoring Fmoc-L-Leu-OH on 2-chlorotrityl (2-ClTrt) resin. Fmoc-L-Leu-OH was loaded on 2-chlorotrityl chloride resin using DIPEA in DCM following a procedure of Barlos et al.[9] The loading of the resin was determined with 2% DBU in DMF according to Gude et al.[10]

Microwave-assisted SPPS couplings. The linear pentapeptides (**5-8**) were built up by microwave-assisted SPPS using a *Discovery* microwave system from *CEM*.

Fmoc-L-Leu-Wang (200 mmol) was successively coupled with the corresponding Fmoc protected amino acids (Fmoc-D-Leu-OH, Fmoc-D-Phe-OH, Fmoc-D-(p-*NH*Boc)-Phe-OH, Fmoc-L-Leu-OH or Fmoc-D-Val-OH) using the following single coupling protocol:

step	reagents	MW conditions	rinsing
1	20 % piperidine in DMF	75°C, 35 W, 30 s	1x DMF
2	20 % piperidine in DMF	75°C, 35 W, 3 min	1x DMF
3	20 % piperidine in DMF	75°C, 35 W, 3 min	4x DMF
4	Fmoc-AA-OH (0.2 M in DMF, 6.0 eq) HBTU (0.5 M in DMF, 5.0 eq) DIPEA (2 M in NMP, 10 eq)	75°C, 25 W (Leu, Phe) or 20 W (Val), 10 min	3x DMF

Resin-Cleavage. Prior to resin cleavage the resin-bound pentapeptides were Fmoc-deprotected applying steps 1-3 of the single coupling protocol, washed with DMF (6x) and DCM (6x) and dried under vacuum. For **5**-**8** the Wang resin was cleaved with TFA 95% (6 ml) under microwave irradiation (30°C, 20 W, 18 min). The TFA solution was evaporated with a light air stream over night and the oily residue was evaporated three times with DCM. The crude linear pentapeptides **5**-**8** were received quantitatively as pure white or slightly yellow foams without need for further purification.

Bioinformatic analysis of biosynthesis gene clusters

Biosynthesis gene clusters for GameXPeptide and xenoamicin have been described previously. For the identification of the biosynthesis gene clusters involved in mevalagmapeptide biosynthesis and the production of natural products in *X. doucetiae*, an antiSMASH analysis[11,12] of the respective genomes (*X. doucetiae*: http://www.ncbi.nlm.nih.gov.proxy.ub.uni-frankfurt.de/nuccore/FO704550.1) was performed and proteins and domains encoded by *mapABC* and *xrdABCDEFGHIJ* were analyzed by BLAST-P (Table S5 and S7).

Metabolic activity and DNA synthesis assay

Subconfluent human MCF7 cells (ATCC, HTB-22) were serum-starved overnight and dissociated with 1×trypsin-EDTA, followed by inhibition with DMEM containing 10% FCS and 1% PenStrep (DMEM complete). Cells were washed in PBS and seeded in

96-well plates at a density of 2×10^4 cells per well in 100 µl DMEM for 20 h prior addition of **13** or DMSO (vehicle control) for a further 24 h. DNA synthesis rate and metabolic activity were determined by the use of 5-bromo-2′-deoxyuridine (BrdU) incorporation with ELISA (Roche) and the colorimetric WST-1 based cell proliferation assay (Roche) following the manufacturer's instructions. The optical density measurements of BrdU and WST-1 measurements were expressed as percentage relative to the DNA synthesis and metabolic activity of MCF7 cells treated with the respective DMSO concentration.

Immunofluorescence staining

Subconfluent human MCF7 cells (ATCC, HTB-22) were serum-starved overnight and dissociated with 1×trypsin-EDTA, followed by inhibition with DMEM containing 10% FCS and 1% PenStrep (DMEM complete). Cells were washed in PBS and seeded onto 13-mm diameter glass coverslips at a density of 2×10^4 cells per well and treated as above. MitoTracker® Red CMXRos (200 nM) was added for the last 45min of culture prior fixation with 4% PFA and counterstaining with ActinGreen™ 488 ReadyProbes® Reagent and DAPI. Coverslips were mounted with Mowiol and analysed with a Keyence Biorevio BZ9000.

References

[1]   E. Duchaud, C. Rusniok, L. Frangeul, C. Buchrieser, A. Givaudan, S. Taourit, S. Bocs, C. Boursaux-Eude, M. Chandler, J.-F. Charles, et al., *Nat Biotechnol* **2003**, *21*, 1307–1313.
[2]   P. Tailliez, S. Pages, N. Ginibre, N. Boemare, *Int J Sys Evol Microbiol* **2006**, *56*, 2805–2818.
[3]   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.
[4]   H. B. Bode, D. Reimer, S. W. Fuchs, F. Kirchner, C. Dauth, C. Kegler, W. Lorenzen, A. O. Brachmann, P. Grün, *Chem. Eur. J.* **2012**, *18*, 2342–2348.
[5]   D. Reimer, K. M. Pos, M. Thines, P. Grün, H. B. Bode, *Nat. Chem. Biol.* **2011**, *7*, 888–890.
[6]   N. Philippe, J.-P. Alcaraz, E. Coursange, J. Geiselmann, D. Schneider, *Plasmid* **2004**, *51*, 246–255.
[7]   L. M. Guzman, D. Belin, M. J. Carson, J. Beckwith, *J. Bacteriol.* **1995**, *177*, 4121–4130.
[8]   A. O. Brachmann, F. Kirchner, C. Kegler, S. C. Kinski, I. Schmitt, H. B. Bode, *J. Biotechnol.* **2012**, *157*, 96–99.
[9]   K. Barlos, D. Gatos, S. Kapolos, C. Poulos, W. Schafer, W. Q. Yao, *Int J Pept Protein Res* **1991**, *38*, 555–561.
[10]   M. Gude, J. Ryf, P. D. White, *Lett Pept Sci* **2002**, *9*, 203–206.
[11]   K. Blin, M. H. Medema, D. Kazempour, M. A. Fischbach, R. Breitling, E. Takano, T. Weber, *Nucleic Acids Research* **2013**, *41*, W204–12.
[12]   M. H. Medema, K. Blin, P. Cimermancic, V. de Jager, P. Zakrzewski, M. A. Fischbach, T. Weber, E. Takano, R. Breitling, *Nucleic Acids Research* **2011**, *39*, W339–46.

Table S1. Oligonucleotides. Restriction sites used for cloning are underlined.

Oligonucleotide	5´to 3´Sequence
pBAD_CEP_Eco147I_Fw	AGGCCTATCGATGCATAATGTGCCTGTC
pBAD_CEP_XbaI_Rv	TCTAGAGGCGCCGTTTAAACATTTAAATCTGCAGAGCTCGAGCATGCACATATGCTAGCCTCCTGTTAGCCCAAAAAAACGGGTATGGAG
Prom_IndC_Fw	CATATGTTAGAAAATAATATTACACAGATTTC
Prom_IndC_Rv	GAGCTCCATTTGATTTACAATACTGTGATGTTC
CEP_Agm_Rhabdopep_Fw	CATATGAAAAATGCAGTGCAAATTGTGAATGA
CEP_Agm_Rhabdopep_Rv	CTGCAGACCAAGTTATCTTCTTCTTTACTGAATCTT
CEP_plu3263_NdeI	CATATGAAAGATAGCATGGCTAAAAAGGAAAT
CEP_plu3263_SacI	GAGCTCAGTAAGTCAGGATTAAAACTCTCTTCCGC
V_pCEP_Fw	GCTATGCCATAGCATTTTTATCCATAAG
V_CEP_Plu_Rv	TACATTCTGAGCCTGTTGCG
V_CEP_Ind_Rv	TAATTTGGTTGTCAGCGCTTCTC
V_CEP_AgmR_Rv	GAGGAATACAGAGCGCGACC
3263_NdeI_for	AACCATATGAAAGATAGCATGGCTAAAAAGGA
3263_MluI_rev	ACGCGTCAATAACTGCTGAGGTGGGTTG
EB_XDV3_110151_PstI_fw	GATCCTGCAGATGAAAGATAGCAGGGCTAA
EB_XDV3_110151_XbaI_rv	GATCTCTAGATGGTTTGTGTATCGGTGAG
EB_XDV3_70458_PstI_fw	GATCCTGCAGATGCCTATGTCATGCAATGGTA
EB_XDV3_70458_XbaI_rv	GATCTCTAGAATGATGACAACAGAACTGCCAG
EB_XDV3_10499_PstI_fw	GATCCTGCAGATGCAGCTTTTTCTCAAAGC
EB_XDV3_10499_XbaI_rv	GATCTCTAGACTATGAATATCGACAAACATCA

6

Table S2. Plasmids constructed and used.

Plasmid	Genotype	Source / Reference
pBAD30	p15a ori, Ap[r], *araC*, *ara*BAD promoter	[7]
pDS132	R6Kγ ori, *sacB*, Cm[r]	[6]
pCK_rpsM	R6Kγ ori, oriT, Cm[r,], *rpsM* promoter	[8]
pCEP-Cm	R6Kγ ori, oriT, Cm[r,], *araC*, *ara*BAD promoter	This work
pCEP-Km	R6Kγ ori, oriT, Km[r,], *araC*, *ara*BAD promoter	This work
pCEP_ind	pCEP, *indC´* (*plu2186*)	This work
pCEP_map	pCEP-Cm, *plu0897´*	This work
pCEP_gxp	pCEP-Cm, *plu3263´*	This work
pCK_rpsM_gxp	pCK, *plu3263´*	This work
pCEP_gxp	pCEP-Km, *XDv3_11015*	This work
pCEP_xab	pCEP-Km, *XDv3_70458*	This work
pCEP_xrd	pCEP-Km, *XDv3_10499*	This work

Table S3. HR-MS data including sum formula of the natural compounds **1**-**17**. D-amino acids are written in lowercase and italics.

compound	sum formular [H$^+$]	m/z calc. [Da]	m/z exp. [Da]	Δppm
GameXPeptide A (**1**)	C$_{32}$H$_{52}$O$_5$N$_5$	586.3968	586.3959	-0.9
GameXPeptide B (**2**)	C$_{33}$H$_{54}$O$_5$N$_5$	600.4125	600.4121	0.3
GameXPeptide C (**3**)	C$_{29}$H$_{54}$O$_5$N$_5$	552.4125	552.4113	-1.2
GameXPeptide D (**4**)	C$_{30}$H$_{56}$O$_5$N$_5$	566.4281	566.4270	-1.1
v-L-*f*-*l*-L (**5**)	C$_{32}$H$_{54}$O$_6$N$_5$	604.4074	604.4075	1.1
l-L-*f*-*l*-L (**6**)	C$_{33}$H$_{56}$O$_6$N$_5$	618.4231	618.4231	0.9
v-L-*l*-*l*-L (**7**)	C$_{29}$H$_{56}$O$_6$N$_5$	570.4231	570.4227	0.4
l-L-*l*-*l*-L (**8**)	C$_{30}$H$_{58}$O$_6$N$_5$	584.4387	584.4386	0.7
Mevalagmapeptide A (**9**)	C$_{33}$H$_{66}$N$_9$O$_5$	668.5182	668.5180	-0.3
Mevalagmapeptide B (**10**)	C$_{32}$H$_{64}$N$_7$O$_5$	626.4963	626.4962	-0.01
Xenoamicin C (**11**)	C$_{65}$H$_{112}$N$_{13}$O$_{15}$	1314.8395	1314.8388	0.6
12/14	C$_{12}$H$_{17}$N$_2$O$_2$S$_2$	285.0726	285.0722	1.4
13/15	C$_{13}$H$_{19}$N$_2$O$_2$S$_2$	299.0882	299.0871	4.0
16/17	C$_{14}$H$_{21}$N$_2$O$_2$S$_2$	313.1039	313.1029	3.1

8

Table S4. Absolute quantification (in mg/L) of GameXPeptides in *P. luminescens* TT01 (wt) and *P. luminescens* TT01::pCK_rpsM_plu3263. Arabinose induction led to slightly lower production.

compound	production wt (± SD)	rpsM_plu3263 (± SD)	ratio rpsM/wt
1	1.8 (0.19)	15.5 (0.96)	8.7
2	0.44 (0.085)	2.9 (0.20)	6.6
3	1.6 (0.25)	12.8 (0.51)	8.3
4	0.19 (0.022)	1.1 (0.11)	5.8
5	0.025 (0.0035)	0.40 (0.062)	16.5
6	0.0069 (0.00070)	0.11 (0.035)	16.5
7	0.045 (0.0077)	0.70 (0.084)	15.5
8	0.010 (0.0027)	0.18 (0.048)	17.9

Table S5. A-domain specificity of MapABC prediction using NRPSpredictor2.

A-domain	gene locus-tag	small-cluster	most likely single amino acid prediction	Stachelhaus code
A1	plu0897	pro	pro	DVQFIAHVVK
A2	plu0898	gly,ala	ala	DLYNNALTYK
A3	plu0899	val,leu,ile,abu,iva	val	DAWWLGGTFK

Table S6. NMR data of mevalagmapeptide A (**9**). Selected COSY (in bold) and HMBC (arrows) correlations are shown. All valine and N-methylvaline residues have been determined previously as L.
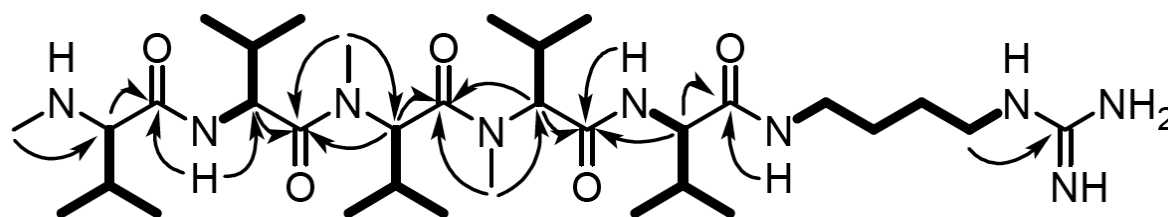
		5	
subunit	position	δ_C	δ_H, mult. (*J* in Hz)
N-Me-L-Val[1]	1	173.6	
	2	69.7	2.64, d (6.1)
	3	31.0	1.71, oktett (6.7)
	4	19.1	0.83, d (6.5)
	5	19.1	0.83, d (6.5)
	HN-CH$_3$	34.8	2.13, s
L-Val[2]	1	172.4	
	2	53.5	4.58, t (8.5)
	3	29.9	1.94, m
	4	18.2	0.86, d (6.7)
	5	18.2	0.86, d (6.7)
	NH		8.01, d (8.8)
N-Me-L-Val[3]	1	170.4	
	2	57.4	5.07, d (10.6)
	3	26.6	2.22, m
	4	17.9	0.68, d (6.4)
	5	19.0	0.80, d (6.7)
	N-CH$_3$	30.2	3.06, s
N-Me-L-Val[4]	1	169.2	
	2	61.4	4.66, d (11.3)
	3	25.6	2.10, m
	4	17.9	0.65, d (6.4)
	5	19.1	0.82, d (6.7)
	N-CH$_3$	30.1	2.94, s
L-Val[5]	1	170.6	
	2	57.9	4.47, d (8.91)
	3	30.3	1.96, m
	4	17.9	0.74, d (6.7)
	5	19.0	0.78, d (6.7)
	NH		7.52, d (9.1)
Agm[6]	1	37.9	3.05, m
	2	29.9	1.92, m
	3	25.8	1.42, m
	4	40.2	3.02, m
	4-NH		7.93, br t (5.0)
	5	157.4	
	5-NH$_2$		7.75, br s
	5=NH		8.75, br s

Table S7. Proteins of the *xrd* cluster, their deduced function, and closest homologues.

Protein name	locus tag		Closest homologue		
	XDD1_	Deduced function	Organism	Identity/ Similarity (%)	Protein locus tag
XrdA	0770	NRPS	*X. bovienii sp. kraussei* Becker Underwood	77/86	CHD25798.1
XrdB	0769	Dehydrogenase	*X. bovienii sp. kraussei* Becker Underwood	82/88	CHD25797.1
XrdC	0768	Thioesterase	*X. bovienii sp. putauvense*	79/88	CDG98999.1
XrdD	0767	Acyl-CoA dehydrogenase	*X. bovienii sp. kraussei* Becker Underwood	85/93	CDH25795.1
XrdE	0766	N-Acyltransferase	*X. bovienii sp. kraussei* Quebec	62/78	CDH21592.1
XrdF	0765	Globin-like family	*X. bovienii*	90/93	YP_003467654.1
XrdG	0764	Phosphoesterase	*X. bovienii sp. putauvense*	80/88	CDG98994.1
XrdH	0763	Phophopantothenoyl-cysteine decarboxylase	*X. bovienii sp. intermedium*	82/89	CDH31302.1
XrdI	0762	Permease	*X. szentirmaii*	87/92	CDL83376.1
XrdJ	0761	Inner membrane protein YaaH	*X. bovienii sp. kraussei* Quebec	95/95	CDH21597.1

CTGACGCTTTTTATCGCAACTCTCTACTGTTTCTCCATACCCGTTTTTTTGGGCTAACAGGAGGCTAGCATATGTGCATGCTCGAGCTCTGCAG
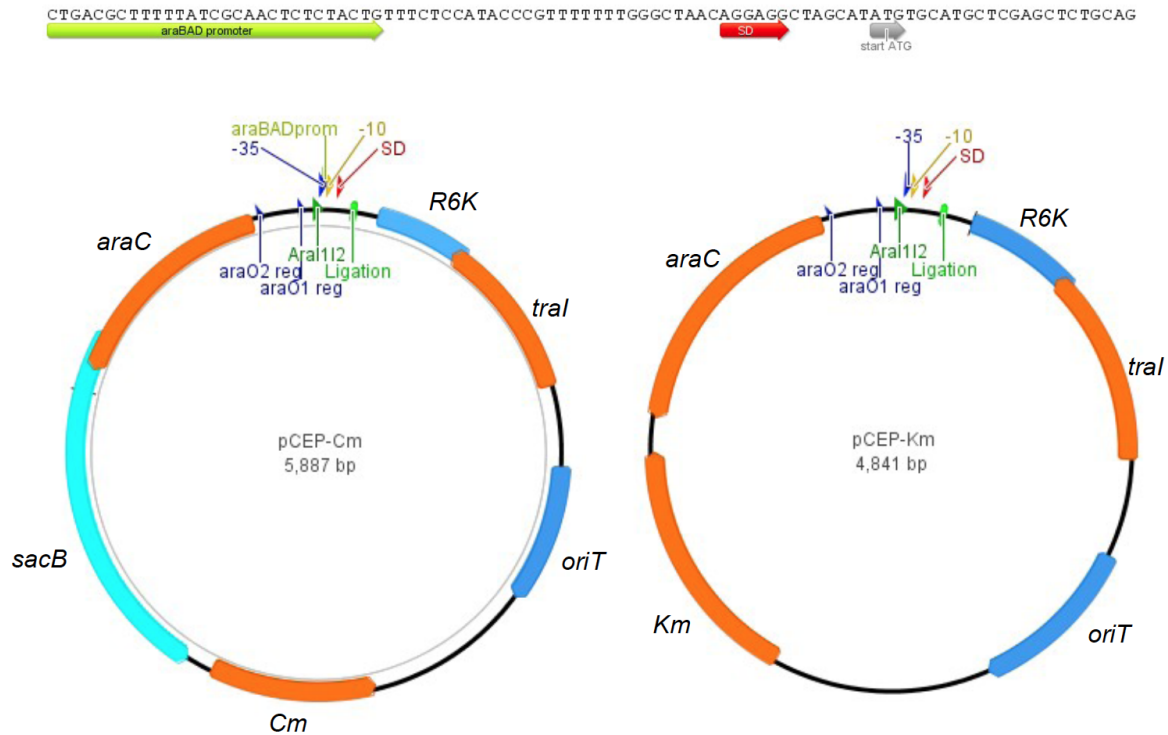
Figure S1. Vector map of pCEP-Cm and pCEP-Km and enlarged region of the multiple cloning site. The *ara*BAD promoter is indicated as a green arrow and the Shine-Dalgarno box as a red arrow and start codon as a grey arrow.
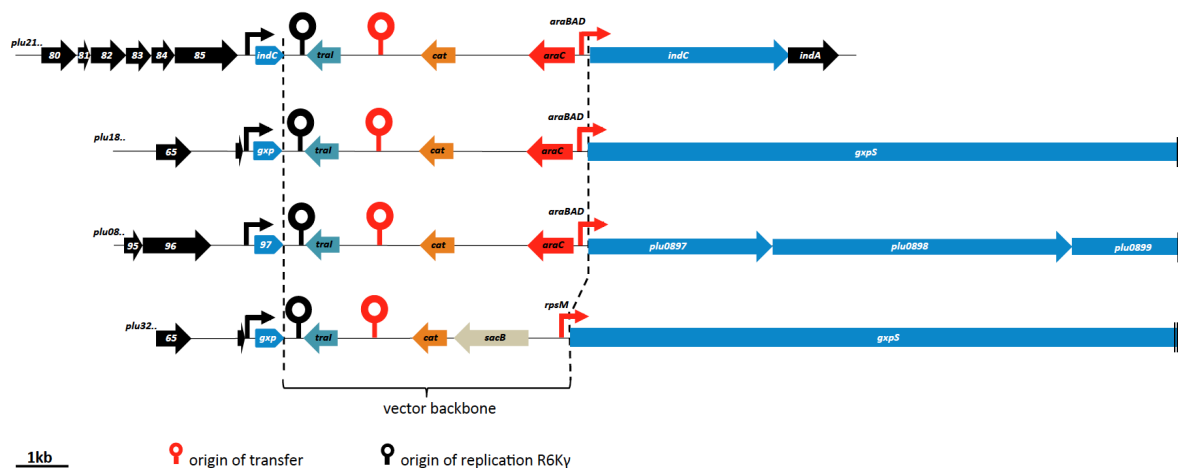
Figure S2. Genotypes of *Photorhabdus luminescens* promoter exchange mutants. Arrows depict promoter regions of the native (black) and the artificial (red) promoters P_{BAD} and P_{rpsM}, respectively.
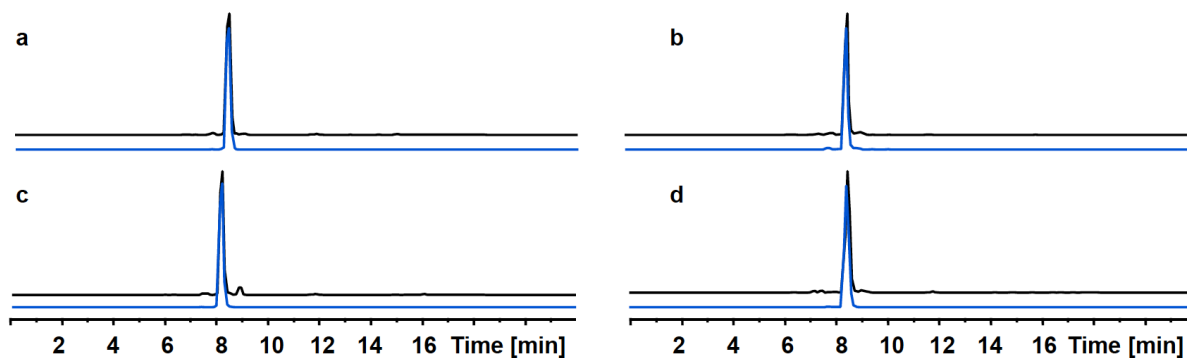
Figure S3. HPLC/MS analysis of the synthesized linear pentapeptides vLfIL **5** (**a**), lLfIL **6** (**b**), vLIIL **7** (**c**), ILIIL **8** (**d**). Depicted are the basepeak chromatograms (black lines) and extracted ion chromatograms (blue lines) for m/z 604.4 (**a**), 618.4 (**b**), 570.4 (**c**), 584.4 (**d**) in the positive mode. All compounds showed identical retention times compared to the natural compounds.
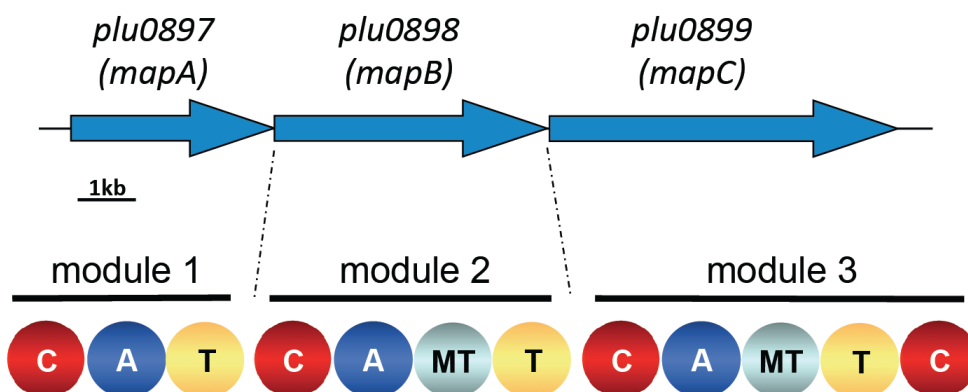
Figure S4. Biosynthesis gene cluster involved in mevalagmapeptide production and schematic representation of NRPS domains. Condensation (C), adenylation (A) and thiolation (T) domain. Modules 2 and 3 harbor additional methyltransferase domains (MT) (conserved domains for S-adenosylmethionine-dependent methyltransferases were identified with help of NCBI BLAST-P analysis). The boundaries of the methyltransferases were determined by adopting NCBI bl2seq blastp using the amino acid sequence of Plu0897 (CAT) aligned against Plu0898 (CAMTT) and Plu0899 (CAMTTC), respectively. The blast results revealed that both methyltransferases are nested between the conserved adenylation domain motifs $A_{1\text{-}8}$ and $A_{9\text{-}10}$. Similarly, both methyltransferases harbor a highly conserved GxG amino acid sequence, like it is invariably encountered in motif I of N-methyltransferases.[19]

13
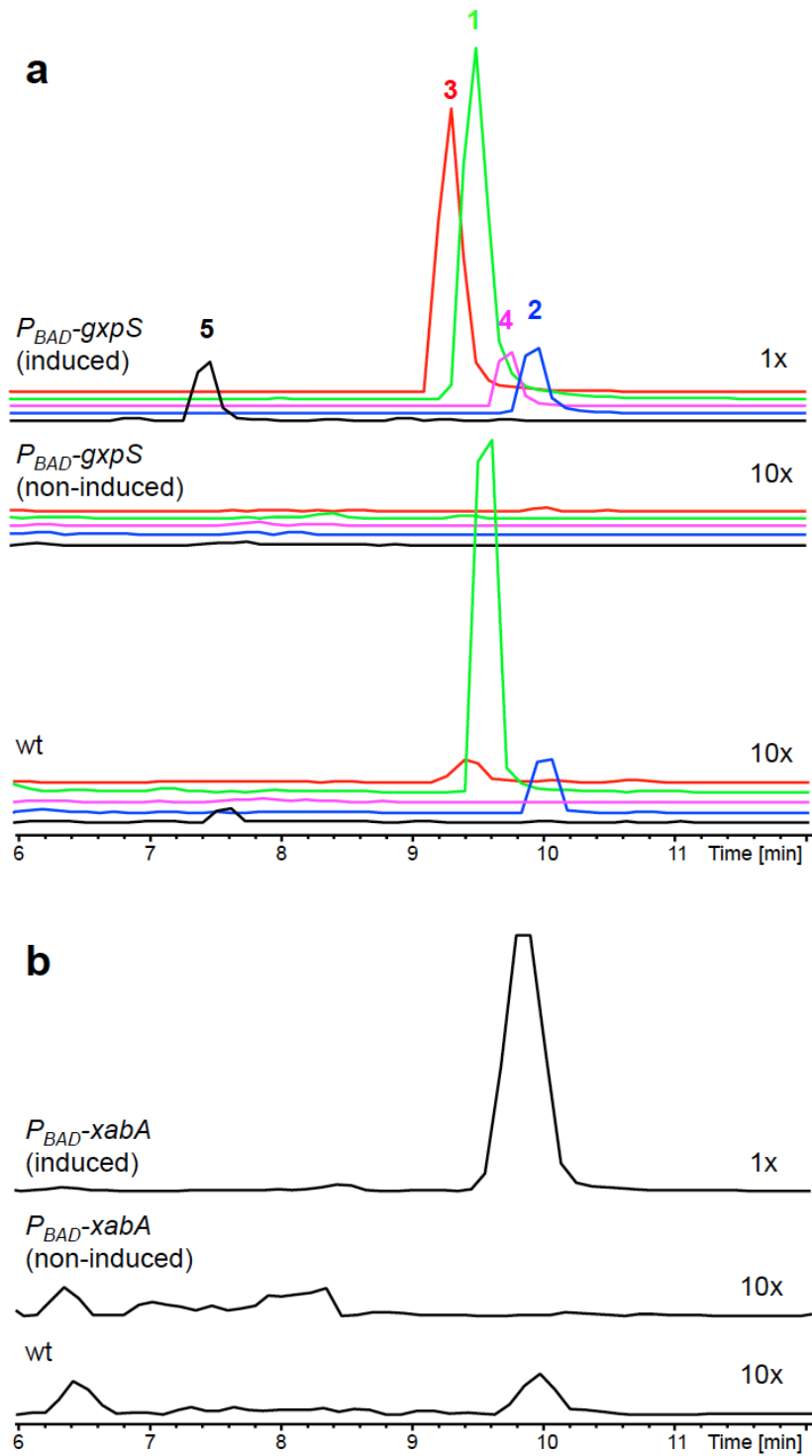
Figure S5. HPLC/MS analysis of promoter exchange mutants of GameXPeptide (a) and xenoamicin production (b) in *X. doucetiae*. Note that non-induced and wt chromatograms are shown 10x increased.
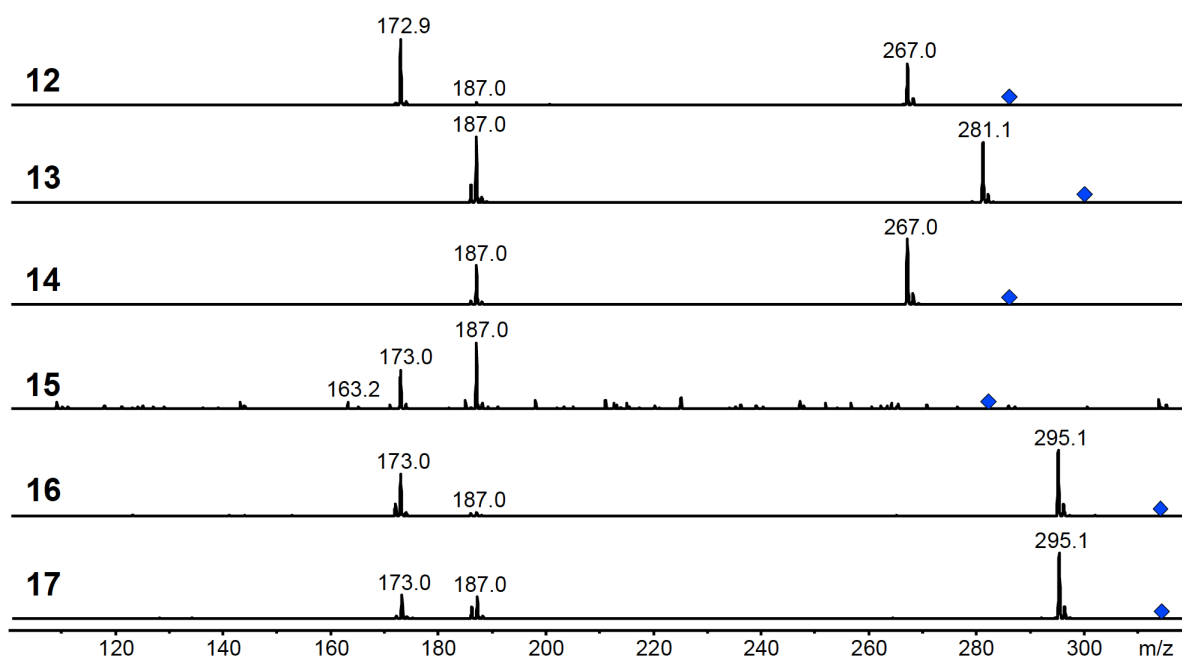
Figure S6. MS-MS analysis of xenorhabdins. Due to their small amounts, MS-MS spectra for **15** and **17** could only be obtained in mixture with the isobaric compound **13** and **16**, respectively.
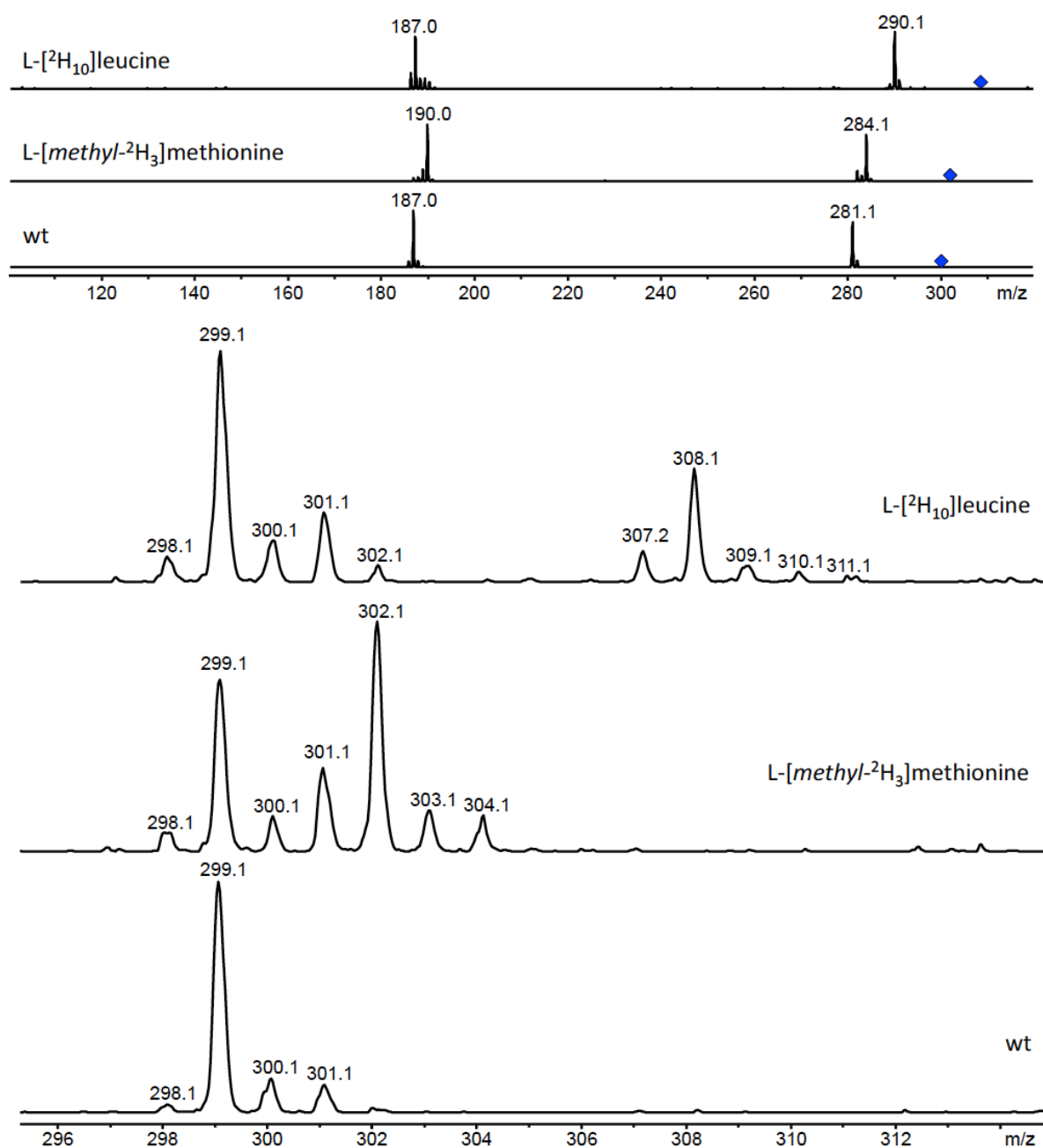
Figure S7. MS analysis of **13** derived from labelling experiments with L-[*methyl-*^{2}H$_3$]methionine and L-[^{2}H$_{10}$]leucine. MS-MS spectra (top) and detailed view on parent ions (bottom).
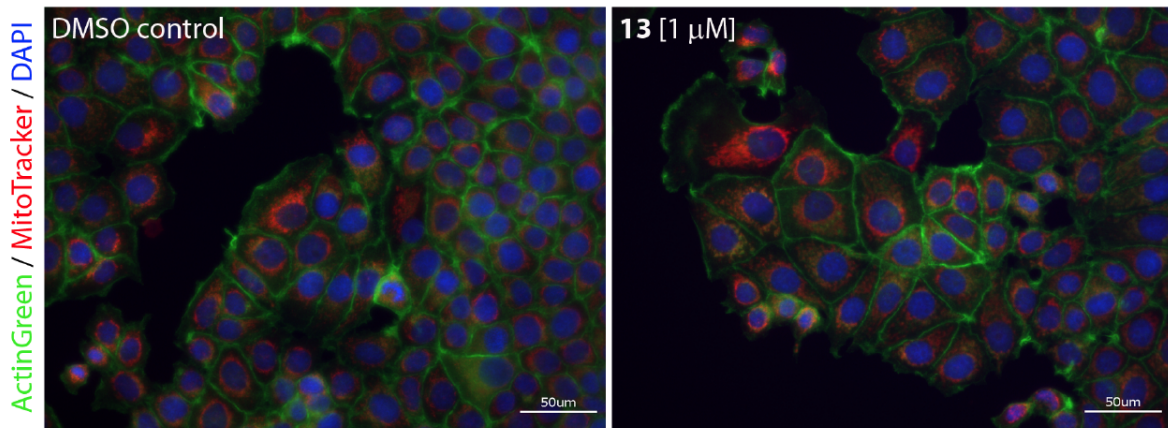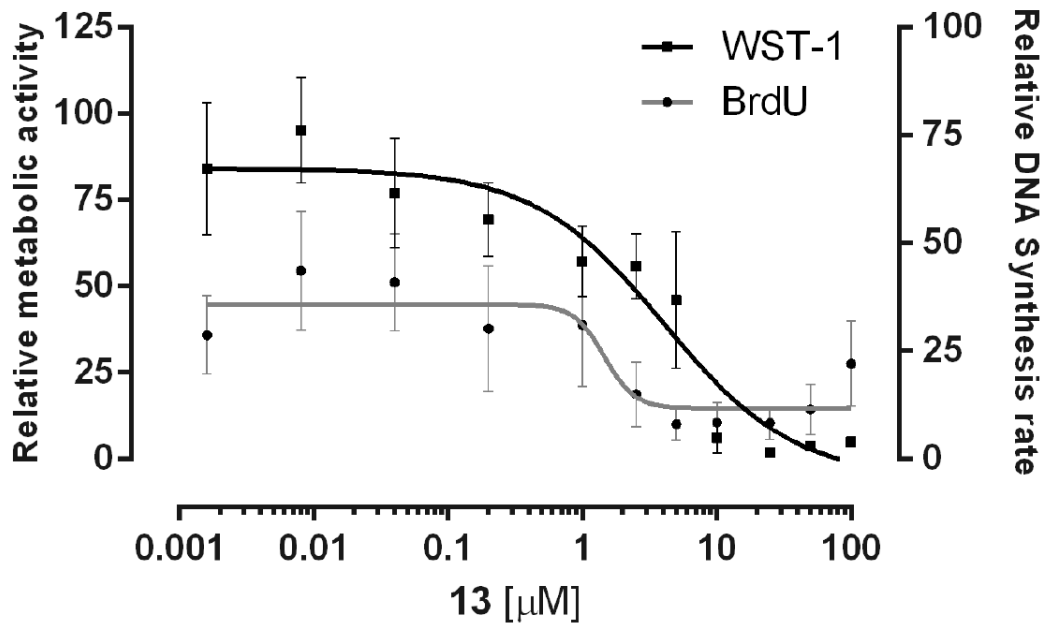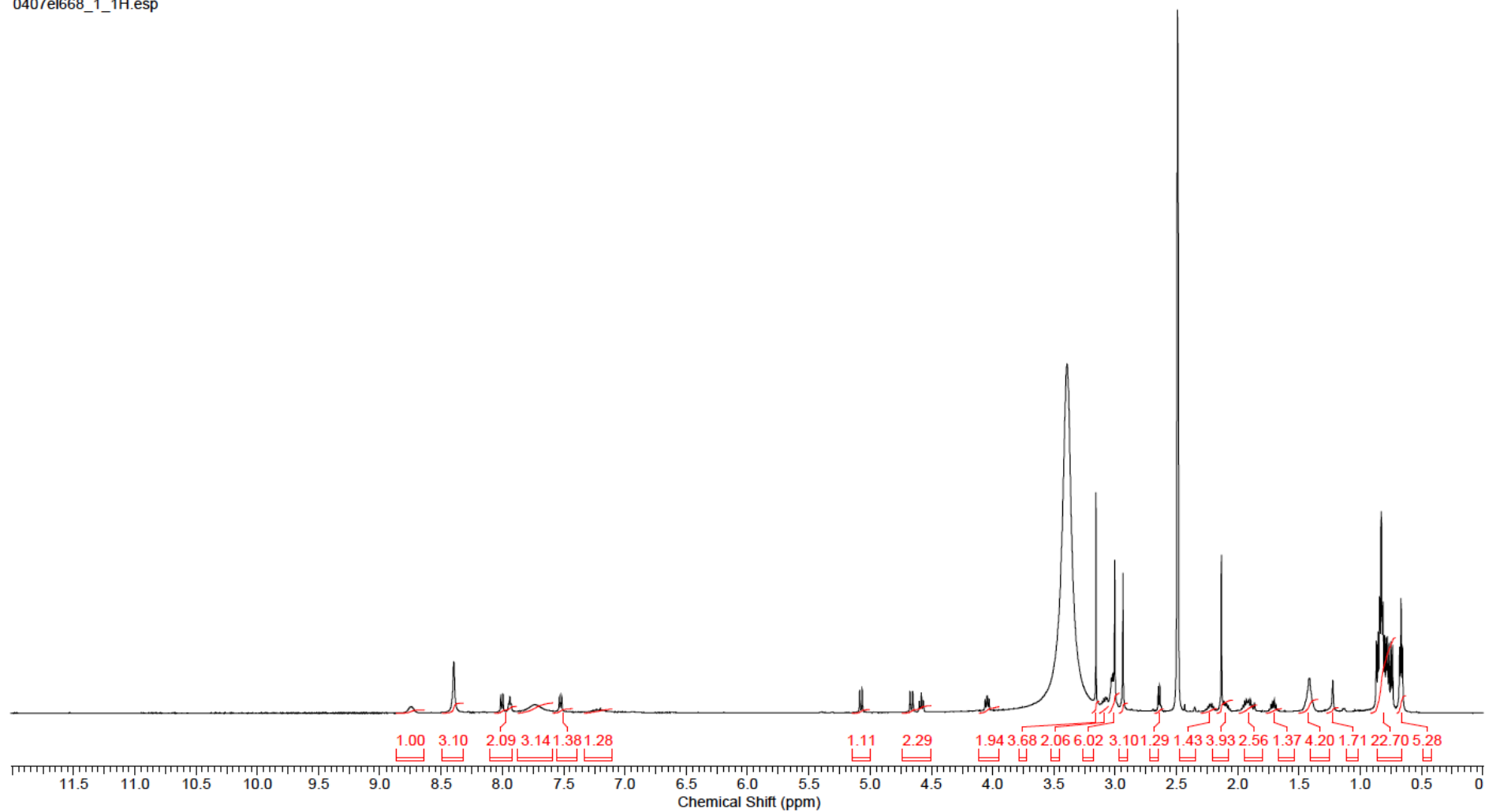
Figure S8. Bioactivity data (metabolic activity [WST-1] and DNA synthesis [BrdU]) and immune fluorescence both on MCF7 cells after 24 h of treatment with compound **13**.

^{1}H NMR (d$_6$-DMSO) of **9**

0407el668_1_1H.esp

Chemical Shift (ppm)

285

18

COSY (d$_6$-DMSO) of **9**

0407el668_1_COSY.esp
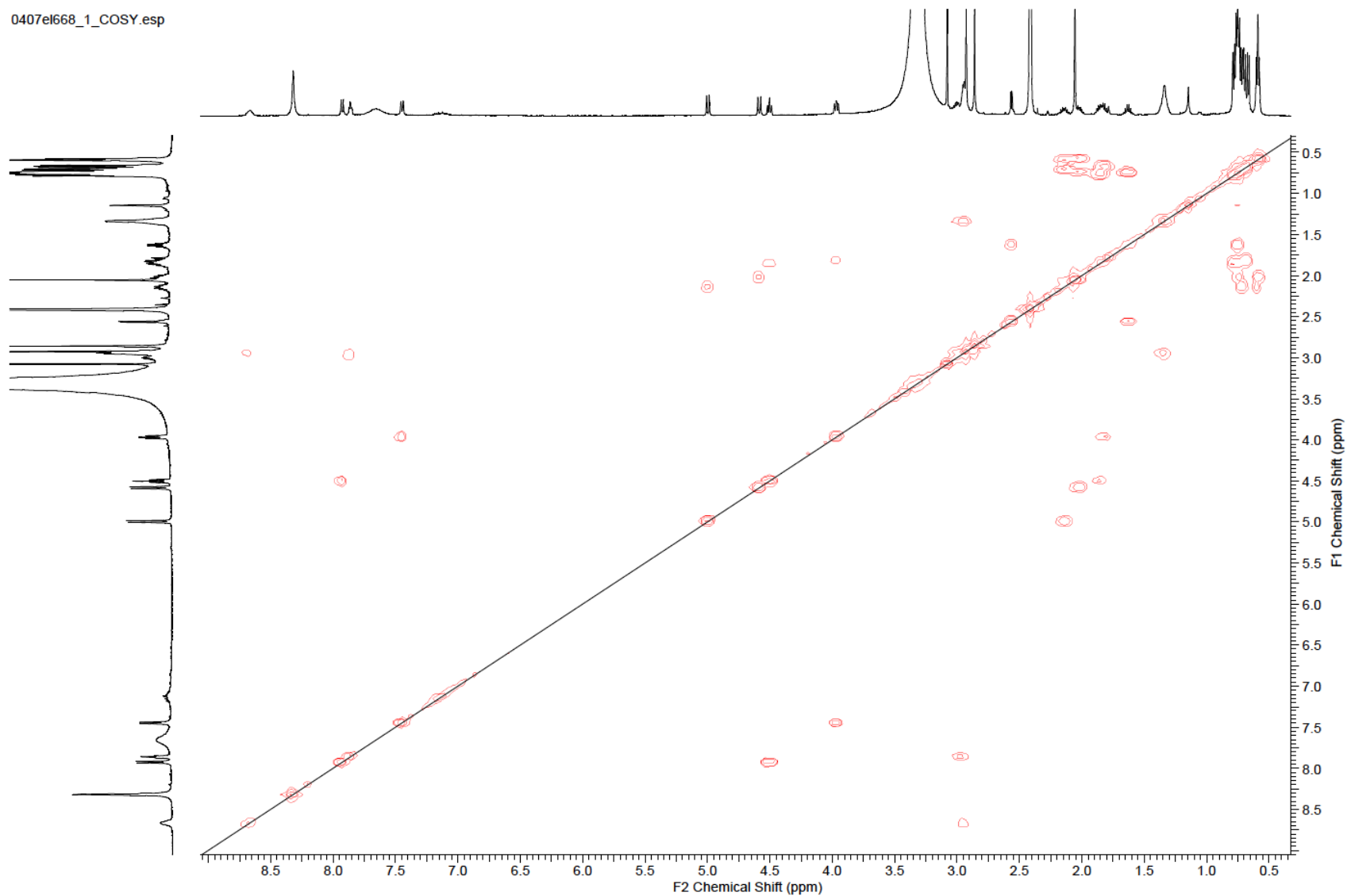
HSQC (d₆-DMSO) of **9**

0407el668_1_HSQC.esp

20

0407el668_1_HMBC.esp

21
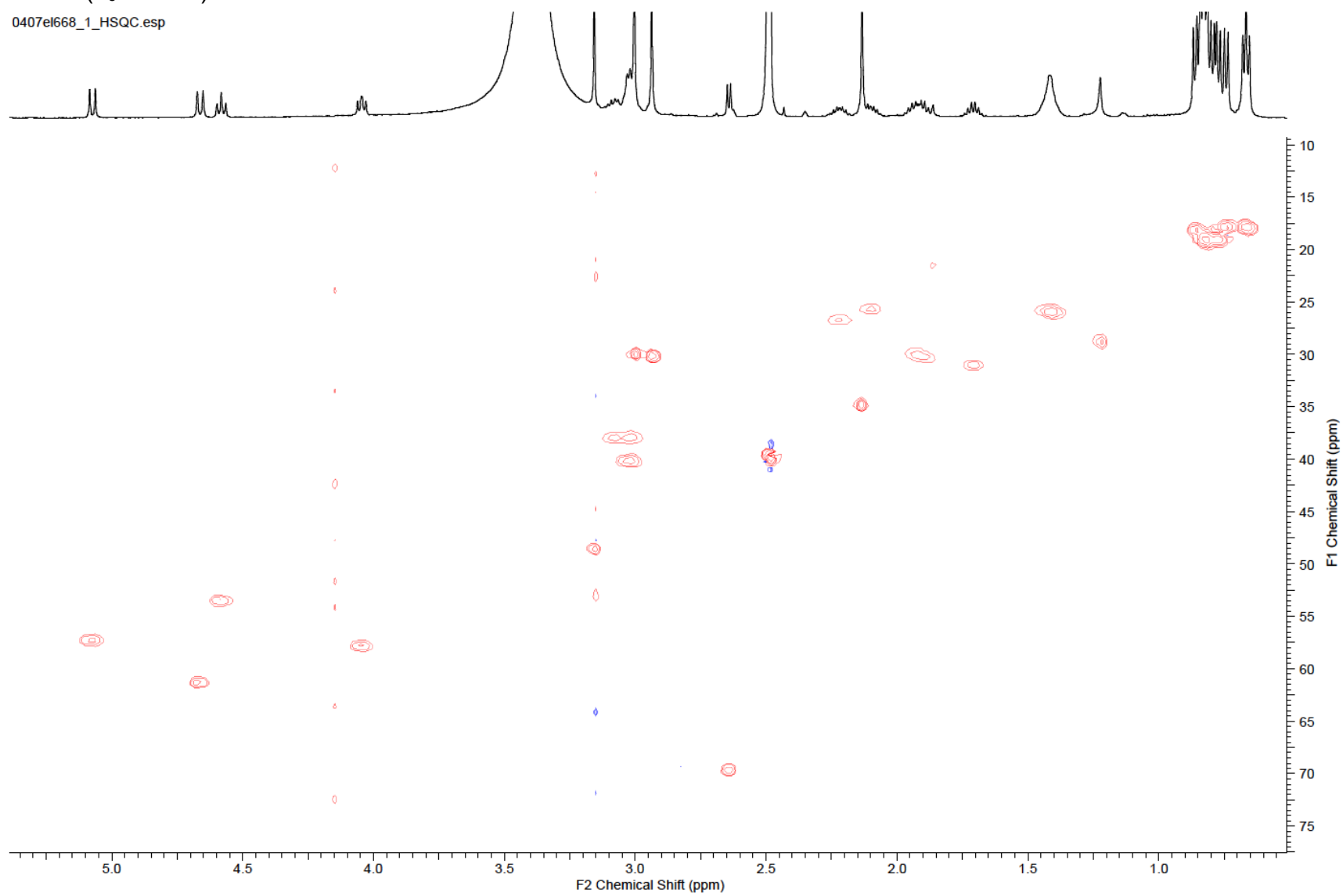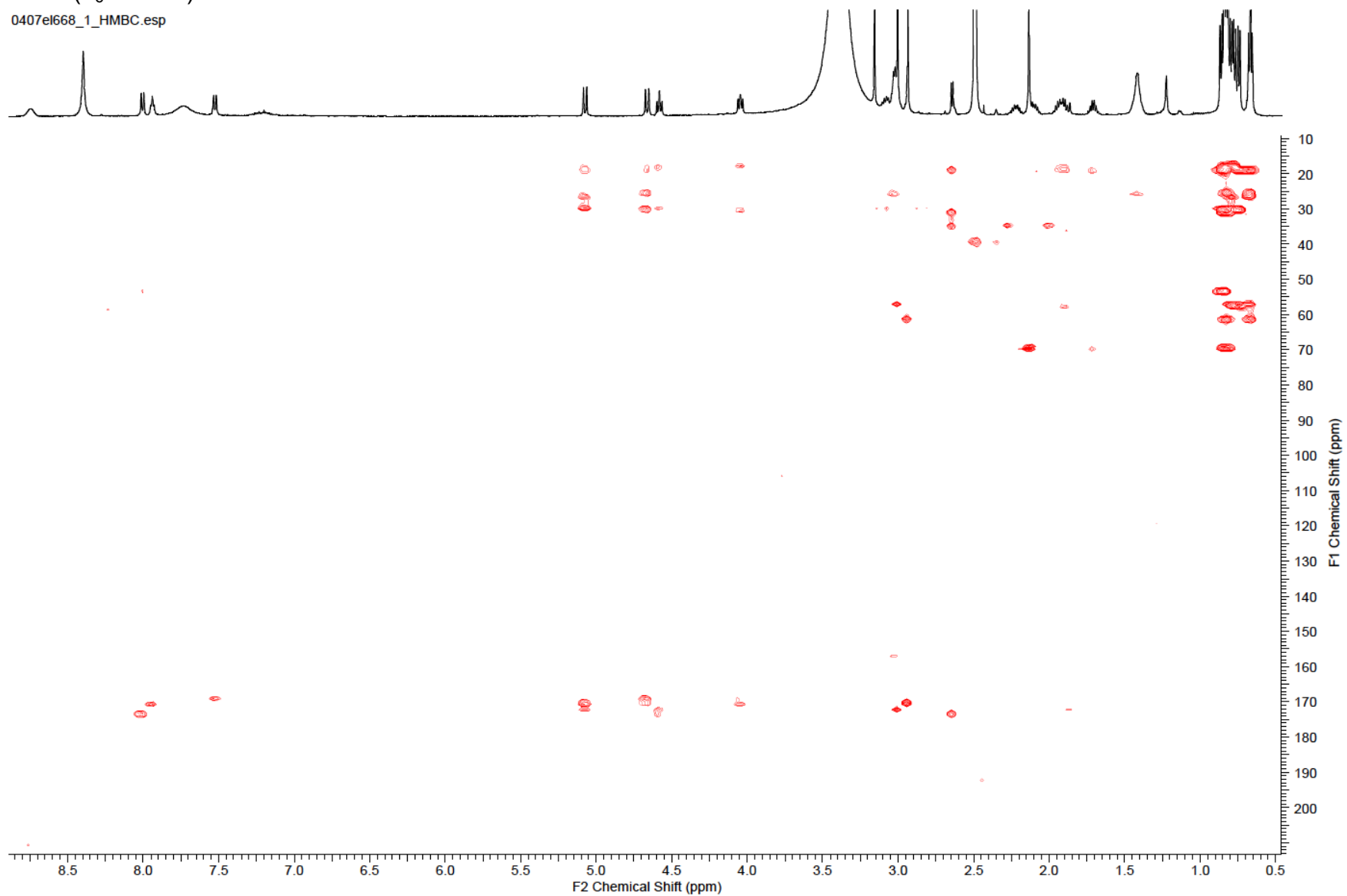
12. Lebenslauf

Name:	Qiuqin Zhou (Chin. Name: 周秋琴)
Geburtsdatum:	01.12.1982
Geburtsort:	Wenzhou in China
Nationalität:	Chinesisch
Familienstand	Verheiratet, zwei Kinder
Adresse:	Edith-Stein-Str. 7, 68766 Hockenheim
Email:	zhouqiuqin@hotmail.com

5 / 2009 – jetzt	Promotionsstudium an der Goethe Universität Frankfurt
10 / 2015 – 2 / 2016	Wissenschaftliche Mitarbeiterin im Max Rubner-Institut
5 / 2009 – 5 / 2012	Wissenschaftliche Mitarbeiterin an der Goethe Universität Frankfurt
3 / 2004 – 12 / 2008	Studium der Chemie an der Philipps-Universität Marburg
4 / 2003 – 2 / 2004	Deutschkurs an der Philipps-Universität Marburg
9 / 2001 - 4 / 2003	Studium der Chemie auf Lehramt an der pädagogischen Universität Peking (Beijing Normal University)
8 / 2001	Schulabschluss an der OhHai ZhongXue in Wenzhou (China)