

**JUSTUS-LIEBIG-UNIVERSITY GIESSEN**

Faculty 09 - Agricultural Sciences, Nutritional Sciences, and Environmental Management

**Microfluidic-Based Miniaturization as well as  
Genomics- and Metabolomics-Guided  
Prioritization of Bacterial Producer Strains  
Leverages the Numbers Game to Discover and  
Characterize Bioactive Natural Products**

**DISSERTATION**

(Dr. rer. nat.)

**Stephan BRINKMANN**

Giessen, April 2022

**1<sup>st</sup> Reviewer: Prof. Dr. Till F. Schäberle**

Department of Natural Product Research

Institute for Insect Biotechnology

Justus-Liebig-University Giessen

**2<sup>nd</sup> Reviewer: Prof. Dr. Markus Nett**

Technical Biology

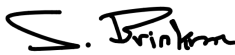
Department of Biochemical and Chemical Engineering

TU Dortmund University

TO MY PARENTS

**Erklärung gemäß der Promotionsordnung des Fachbereichs 09 vom 07. Juli 2004 § 17 (2),  
geändert am 29. Mai 2019**

Ich erkläre: Ich habe die vorgelegte Dissertation selbständig und ohne unerlaubte fremde Hilfe und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht. Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis“ niedergelegt sind, eingehalten.



---

Stephan Brinkmann

Gießen, 23.04.2022

Die vorliegende Arbeit wurde von Dezember 2017 bis August 2021 in der Abteilung Bioressourcen des Fraunhofer-Instituts für Molekularbiologie und Angewandte Ökologie (IME-BR) in Gießen (Institutsleitung: Prof. Dr. Andreas Vilcinskas) angefertigt. Die Arbeit wurde in der Arbeitsgruppe Naturstoffforschung unter der Leitung von Dr. Jens Glaeser (12/2017-09/2018) und Prof. Dr. Till F. Schäberle (10/2018-07/2021) im Rahmen des Public Private Partnerships „Sanofi-Fraunhofer Natural Products Center of Excellence“ (12/2017-09/2018) und dessen Nachfolgers, dem „Fraunhofer-Evotec Natural Products Center of Excellence“ (10/2018-12/2019), unter der Projektleitung von Prof. Dr. Peter E. Hammann und Dr. Jens Glaeser durchgeführt. Die Promotion wurde in diesem Zeitraum vom hessischen Ministerium für Wissenschaft und Kunst (HMWK) über die Landes-Offensive zur Entwicklung Wissenschaftlich-ökonomischer Exzellenz (LOEWE) für das LOEWE-Zentrum für Insektenbiotechnologie & Bioressourcen finanziert.

# PUBLICATIONS

---

## Part of this work

M. Oberpaul<sup>†</sup>, **S. Brinkmann**<sup>†</sup>, M. Marner, S. Mihajlovic, B. Leis, M. A. Patras, C. Hartwig, A. Vilcinskas, P. E. Hammann, T. F. Schäberle, M. Spohn and J. Glaeser (Epub 2021, **2022**) Combination of high-throughput microfluidics and FACS technologies to leverage the numbers game in natural product discovery. *Microb. Biotechnol.* **15**: 415-430, DOI 10.1111/1751-7915.13872

**S. Brinkmann**, M. Oberpaul, J. Glaeser and T. F. Schäberle (**2021**) Two-step generation of monodisperse agarose-solidified double emulsions (w/w/o) excluding an inner oil barrier. *MethodsX* **8**: 101565, DOI 10.1016/j.mex.2021.101565

**S. Brinkmann**, M. Kurz, M. A. Patras, C. Hartwig, M. Marner, B. Leis, A. Billion, Y. Kleiner, A. Bauer, L. Toti, C. Pöverlein, P. E. Hammann, A. Vilcinskas, J. Glaeser, M. S. Spohn and T. F. Schäberle (**2022**) Genomic and Chemical Decryption of the Bacteroidetes Phylum for Its Potential to Biosynthesize Natural Products. *Microbiol. Spectr.*, DOI 10.1128/spectrum.02479-21

M.-K. Bill<sup>†</sup>, **S. Brinkmann**<sup>†</sup>, M. Oberpaul, M. A. Patras, B. Leis, M. Marner, M.-P. Maitre, P. E. Hammann, A. Vilcinskas, S. M. M. Schuler and T. F. Schäberle (**2021**) Novel Glycerophospholipid, Lipo- and N-Acyl Amino Acids from Bacteroidetes: Isolation, Structure Elucidation and Bioactivity. *Molecules* **26**: 5195, DOI 10.3390/molecules26175195

**S. Brinkmann**<sup>†</sup>, S. Semmler<sup>†</sup>, C. Kersten, M. A. Patras, M. Kurz, N. Fuchs, S. J. Hammerschmidt, J. Legac, P. E. Hammann, A. Vilcinskas, P. J. Rosenthal, T. Schirmeister, A. Bauer and T. F. Schäberle (**2022**) Identification, Characterization and Synthesis of Natural Parasitic Cysteine Protease Inhibitors – Pentacitidins are More Potent Falcitidin Analogs. *ACS Chem. Biol.* **17**: 576-589, DOI 10.1021/acscchembio.1c00861

**S. Brinkmann**, M. S. Spohn and T. F. Schäberle, Bioactive natural products from Bacteroidetes. (**2022**) *Nat. Prod. Rep.*, DOI 10.1039/D1NP00072A

---

## Not part of this work

M. Oberpaul, M. Spohn, **S. Brinkmann**, S. Mihajlovic, M. Marner, M. A. Patras, L. Toti, M. Kurz, A. Vilcinskas, J. Glaeser and T. F. Schäberle (**2022**) *Trichoderma*-Derived Pentapeptides from the Infected Nest Mycobiome of the Subterranean Termite *Coptotermes testaceus*. *ChemBioChem*, DOI 10.1002/cbic.202100698

<sup>†</sup> These authors contributed equally to this work.

# CONTENTS

---

PUBLICATIONS.....	VI
ABSTRACT.....	- 8 -
INTRODUCTION.....	- 10 -
1. Natural product research .....	- 11 -
1.1. Microfluidic-based anti-infective discovery approaches.....	- 12 -
1.2. Review – Bioactive natural products from Bacteroidetes.....	- 13 -
2. Aim of this work.....	- 35 -
CHAPTER 1 – MICROFLUIDICS .....	- 36 -
Publication I – Combination of high-throughput microfluidics and FACS technologies to leverage the numbers game in natural product discovery.....	- 37 -
Publication II – Two-step generation of monodisperse agarose-solidified double emulsions (w/w/o) excluding an inner oil barrier .....	- 54 -
CHAPTER 2 – BACTEROIDETES.....	- 65 -
Publication III – Genomic and Chemical Decryption of the Bacteroidetes Phylum for Its Potential to Biosynthesize Natural Products.....	- 66 -
Publication IV – Novel Glycerophospholipid, Lipo- and N-acyl Amino Acids from Bacteroidetes: Isolation, Structure Elucidation and Bioactivity.....	- 87 -
Publication V – Identification, Characterization and Synthesis of Natural Parasitic Cysteine Protease Inhibitors – Pentacididins are More Potent Falcitidin Analogs.....	- 106 -
DISCUSSION .....	- 121 -
Droplet-based microfluidics approaches for anti-infective NP discovery.....	- 121 -
Bacteroidetes – A yet underexplored source of natural product research.....	- 125 -
CONCLUSION and OUTLOOK .....	- 128 -
REFERENCES .....	- 130 -
ACKNOWLEDGMENTS .....	- 135 -
COPYRIGHT PERMISSION .....	- 136 -

## ABSTRACT

---

Over the past few decades, about two-thirds of the small molecule drugs in application are based on microbial natural products (NPs) and their analogs. The undisputed success notwithstanding, the continuous increase of antimicrobial resistance and a high rate of rediscovery of NPs applying traditional cultivation and screening methods pose a serious threat to modern medicine and agriculture. Modern high-throughput microfluidic-based approaches as well as omics-based technologies are proven complementary and beneficial alternatives in order to discover novel NP scaffolds and combat this crisis.

In this study, droplet-based microfluidic approaches have been developed and implemented for high-throughput microbial cultivation and screening. First, more than a thousand axenic cultures were retrieved from a soil sample and their antimicrobial activities were evaluated. The cultured strains were affiliated to five phyla (57 genera) including one member of the Acidobacteria phylum (genus *Edaphobacter*). Based on this droplet-based microfluidic approach and with the aim of setting up a more complex, multi compartment, and high-throughput microfluidic assay, a technical method to encase the previously generated agarose-solidified droplets with a further agarose shell has been successfully developed and validated. In parallel, the traditional activity-based discovery of NPs has been supplemented with genomics and metabolomics technologies to study the entire phylum of the Bacteroidetes – a taxonomic branch that has not been intensively explored in relation to the production of NPs. Computational prediction tools (i.e., antiSMASH, BiG-SCAPE) were used to determine the biosynthetic gene cluster (BGC) amount, type, distribution, as well as diversity to rate the biosynthetic potential encoded within the genomes of 600 Bacteroidetes strains. It revealed a genetic repertoire comparable towards classical NP producer taxa of a few genera, rather than a general ability of the entire phylum-members. A metabolomics analysis of the outstanding genus *Chitinophaga* (25 strains) confirmed the genomic-based prediction by revealing a chemical space not able to be assigned to any microbial NP identified today, except of the antimalarial compound Falcitidin. Based on this datasets, several bioactive natural products have been isolated, characterized and their structures have been elucidated by performing



extensive UHPLC-MS/MS and NMR experiments supplemented by Marfey's analysis. While the novel cyclic lipodepsipeptides chitinopeptins A–D exhibit broad antimicrobial activity, the lipoamino acids are active against *Moraxella catarrhalis* FH6810, and the identified pentacididins (linear pentapeptides with a C-terminal aldehyde) act as parasitic cysteine protease inhibitors with an additional low-micromolar inhibition of  $\alpha$ -chymotrypsin. Putative non-ribosomal peptide synthetase BGCs corresponding to the structural features of the chitinopeptins and the pentacididins have been discovered and similar BGCs have been found in further *Chitinophaga* genomes.

This work contributes to the field of NP research as it states the enormous potential of microfluidic-based high-throughput cultivation and screening assays by establishing and advancing complex agarose-solidified droplet-based microfluidic approaches. Furthermore, omics and isolation studies performed as part of this work highlight the phylum Bacteroidetes as an exciting bioresource capable of producing novel NP scaffolds, many of which await discovery.



## INTRODUCTION

---

In the early 20<sup>th</sup> century the discovery of the first antibacterial substances, e.g., Prontosil, a sulphonamide drug, and penicillin, represented a breakthrough in therapeutic treatment of bacterial infections that previously frequently resulted in the patient's death (Domagk, 1935; Chain *et al.*, 1940). The success of these small molecules promoted the discovery and development of further anti-infective compounds, especially microbe-derived natural products (NPs). Many of the nowadays known anti-infective compound classes were found during the golden age of antibiotics up on the 1970s. Ended by constant rediscovery and high developing costs, the golden age has been followed by a diminishing number of approved anti-infective agents for application to this day. Evidently, this development represents a contradiction to the constant accumulation of non-treatable, multi, and even pan-resistant pathogens with socioeconomic consequences for human health and global food supply security (Fair and Tor, 2014; Talebi Bezmin Abadi *et al.*, 2019; Lewis, 2020; Murray *et al.*, 2022). However, since about two-thirds of the new chemical entities approved by the US Food and Drug Administration from 1981 to 2019 are associated to NPs (Newman and Cragg, 2020), the relevance for the discovery and constant supply of NPs for the development of new therapeutic agents remains high (Atanasov *et al.*, 2021). Fortunately, the renaissance of NP research in both academia and industry has been witnessed by advances in high-throughput screening and sequencing, analytical technologies as well as synthetic biology, and bioinformatics (Li and Vederas, 2009; Bérdy, 2012; Wohlleben *et al.*, 2016; Wright, 2019).

## 1. Natural product research

NPs are specialized chemical compounds produced in nature and commonly referred to as 'secondary metabolites'. These low molecular weight molecules are usually taxonomically restricted compounds not directly involved in cell growth or reproduction with specialized functions. Their characteristics are shaped by evolutionary adaptation to specific environments. They serve as intra- and inter-species communication and differentiation effectors, as metal transporting agents or as weapons used against environmental competitors (Demain and Fang, 2000; Davies, 2013). Nature possesses a vast chemical diversity of unique molecules that are customized and adjusted for interactions with their respective molecular target (Firn and Jones, 2003). NP scaffolds are diverse, target-affine, and specific, and are therefore an essential source for drug discovery (Hong, 2011). Historically, most of the anti-infective NPs originate from intensive studies of rather easily to cultivate and prolific microbial NP producer, i.e. the Actinobacteria, Firmicutes, Proteobacteria and Myxobacteria, covering a limited taxonomical diversity (Monciardini *et al.*, 2014; Challinor and Bode, 2015). In contrast, technical and scientific progress today predicts a connection between phylogenetic and genomic divergence and chemical diversity (Achtman and Wagner, 2008; Medema *et al.*, 2014; Monciardini *et al.*, 2014; Tracanna *et al.*, 2017; Hoffmann *et al.*, 2018). Along with a not yet-exhausted genetic potential of traditional prolific microbial NP producers such as *Streptomyces* and *Bacillus* (Zhao and Kuipers, 2016; Adamek *et al.*, 2019; Belknap *et al.*, 2020), biosynthetic richness is found across various branches of the tree of life (Cimermancic *et al.*, 2014; Challinor and Bode, 2015; Tracanna *et al.*, 2017). Thus, investigation efforts exploiting the genetic and metabolic capabilities of underexplored, or even prior uncultivable microbial taxa, revealed a hidden potential for further unknown NPs (Brady *et al.*, 2009; Piel, 2011; Wilson *et al.*, 2014; Challinor and Bode, 2015; Harvey *et al.*, 2015; Dejong *et al.*, 2016; Hoffmann *et al.*, 2018).

In the next decades, scientists will continue to discover new classes of anti-infectives from the emphasized almost boundless chemical diversity of NPs that is predicted in the large diversity of biosynthetic gene clusters (BGCs), found in microbial genomes within all branches of the tree of life.

### 1.1. Microfluidic-based anti-infective discovery approaches

Accessibility of taxonomically broad and diverse microbial species is crucial for NP research, especially those that are underrepresented in this research area or even uncultured today. So far, the five phyla – Actinobacteria, Bacteroidetes, Cyanobacteria, Firmicutes, and Proteobacteria – represent 95% of all cultivated and published species. The remaining 5% are representatives of more than 20 phyla (Keller and Zengler, 2004).

Traditional cultivation approaches are limited in the throughput and the cultivability of most of these microorganisms *in vitro*. The development of miniaturized cultivation methods paved the way to increase the number of culturable microorganisms. Miniature diffusion chambers of various sizes were designed to retrieve environmental microbial colonies *in situ* or to rapidly identify cultivation conditions for target cultivation of representatives of previously uncultured microbes. Several successfully *in situ* studies have shown that such devices are capable of cultivating new species likely because diffusion provides them with their naturally occurring growth factors. (Manome *et al.*, 2001; Bollmann *et al.*, 2007; Ingham *et al.*, 2007; Nichols *et al.*, 2010; Ma *et al.*, 2014b; Ma *et al.*, 2014a; Tandogan *et al.*, 2014). One of these new species, the rare Proteobacteria *Eleftheria terrae*, was even found to produce the novel antibiotic teixobactin (Ling *et al.*, 2015).

Other miniaturization approaches focus on even smaller compartments applying droplet-based microfluidics (Kaminski *et al.*, 2016; Price and Paegel, 2016). Droplet-based systems create discrete volumes ( $\mu\text{L}$  to  $\text{fL}$ ) with the use of immiscible phases (Teh *et al.*, 2008; Seemann *et al.*, 2012). In recent years, simplified systems were commercialized and thus supported the development of e.g., high-throughput cultivation approaches (Vincent *et al.*, 2010; Nge *et al.*, 2013; Volpatti and Yetisen, 2014; Jiang *et al.*, 2016; Cao *et al.*, 2017; Terekhov *et al.*, 2017; Mahler *et al.*, 2018; Villa *et al.*, 2020; Watterson *et al.*, 2020; Mahler *et al.*, 2021). Downscaling and compartmentalization are advantages of such systems that allow to retrieve a few thousand of axenic microbial cultures in a single run (Zengler *et al.*, 2002; Akselband *et al.*, 2006; Theberge *et al.*, 2010; Mahler *et al.*, 2018; Mahler *et al.*, 2021). The extreme small-scale droplets allow cells to shape their microenvironment towards their specific/individual needs and avoid growth competition (Martin *et al.*, 2003; Keller and Zengler, 2004; Ben-Dov *et al.*, 2009; Boedicker *et al.*,

2009; Ishii *et al.*, 2010; Vincent *et al.*, 2010; Stewart, 2012; Boitard *et al.*, 2015; Jiang *et al.*, 2016; Cao *et al.*, 2017; Terekhov *et al.*, 2017; Villa *et al.*, 2020; Watterson *et al.*, 2020). In general, these systems increased the throughput as well as the diversity, and therefore, the probability to retrieve microorganisms uncultured before (Hengoju *et al.*, 2020a).

## 1.2. Review – Bioactive natural products from Bacteroidetes

The Bacteroidetes phylum is represented by Gram-stain-negative, chemo-organotrophic, non-spore forming rod shaped bacteria that are widely distributed in all environments. Most of them are easy to cultivate, which allows them to be intensively studied under laboratory conditions (Thomas *et al.*, 2011; Mendes *et al.*, 2013). Interestingly, there is an attributed high genetic potential for the production of NPs from Bacteroidetes species (Borsetto *et al.*, 2019), while the number of biologically active compounds is low. Hence, we reviewed the literature for bioactive molecules from Bacteroidetes to highlight this phylum as an underexplored microbial source in the field of NP research.

Natural Product  
Reports



REVIEW

[View Article Online](#)  
[View Journal](#)



## Bioactive natural products from Bacteroidetes

Cite this: DOI: 10.1039/d1np00072a

Stephan Brinkmann, <sup>a</sup> Marius S. Spohn <sup>a</sup> and Till F. Schäberle \*<sup>abc</sup>

**Copyright:** This article is Open Access (CC BY 3.0).

### **Contribution:**

Stephan Brinkmann conceived and performed the literature search on NPs produced by Bacteroidetes species. He prepared all the figures for the manuscript. Stephan Brinkmann, Marius S. Spohn and Till F. Schäberle wrote the draft of the manuscript as well as the following revisions.



## Bioactive natural products from Bacteroidetes

Cite this: DOI: 10.1039/d1np00072a

Stephan Brinkmann, <sup>a</sup> Marius S. Spohn <sup>a</sup> and Till F. Schäberle <sup>\*abc</sup>

Covering: up to end of January 2022

Received 1st November 2021

DOI: 10.1039/d1np00072a

rsc.li/npr

Bacteria representing the phylum Bacteroidetes produce a diverse range of natural products, including polyketides, peptides and lactams. Here, we discuss unique aspects of the bioactive compounds discovered thus far, and the corresponding biosynthetic pathways if known, providing a comprehensive overview of the Bacteroidetes as a natural product reservoir.

1	<b>Introduction</b>	6.1.4	<b>Formadicins A–D</b>
2	<b>Macrolides</b>	6.2	<b>Cephems</b>
2.1	<b>Elansolids</b>	6.2.1	<b>Desacetoxycephalosporin C</b>
2.2	<b>YL-02905-A and YL-02905-B</b>	6.2.2	<b>7a-Formylaminocephalosporins (SQ 28516 and SQ 28517)</b>
3	<b>Peptides</b>	6.2.3	<b>Chitinovorins A–D</b>
3.1	<b>Cyclic depsipeptides</b>	7	<b>Additional Bacteroidetes natural products</b>
3.1.1	<b>(Iso)pedopeptins</b>	7.1	<b>Pigments</b>
3.1.2	<b>Chitinopeptins A–D</b>	7.2	<b>Siderophores</b>
3.1.3	<b>Katanosins A and B</b>	7.3	<b>Lipids</b>
3.1.4	<b>YM-47141 and YM-47142</b>	7.4	<b>Others</b>
3.2	<b>Short linear peptides</b>	8	<b>Concluding remarks</b>
3.2.1	<b>TAN-1057 A–D</b>	9	<b>Conflicts of interest</b>
3.2.2	<b>Falcitidin and pentacitidins A and B</b>	10	<b>Acknowledgements</b>
3.3	<b>Ribosomally synthesized and post-translationally modified peptides (RiPPs)</b>	11	<b>References</b>
3.3.1	<b>Lantipeptides</b>		
3.3.1.1	<b>Pinensins A and B</b>		
3.3.1.2	<b>PedA peptides</b>		
3.3.2	<b>Bacteriocin-like peptide FR901451</b>		
4	<b>Polyketide – non-ribosomal peptide hybrids</b>		
4.1	<b>Ariakemicins A and B</b>		
5	<b>Quinolines</b>		
5.1	<b>G1499-2</b>		
5.2	<b>Marinoquinolines and marinoazepinones</b>		
6	<b>Lactams</b>		
6.1	<b>Monobactams</b>		
6.1.1	<b>SQ 28332 (also known as PB-5582-A)</b>		
6.1.2	<b>SQ 28502 and SQ 28503</b>		
6.1.3	<b>PB-5266 A–C</b>		

## 1 Introduction

Antimicrobial resistance (AMR) is a serious threat to modern medicine and agriculture.<sup>1</sup> Bacteria and fungi resistant to many or most antimicrobials have dire socioeconomic consequences for human health and global food security.<sup>2</sup> The threat of AMR is exacerbated by the limited number of candidate antibiotics entering the development pipeline and thus the declining number of approvals for new anti-infective drugs.<sup>3</sup>

The AMR crisis is being addressed by efforts to discover natural products that can serve as new anti-infective leads, and microbes are a rich source of such compounds.<sup>4</sup> However, natural product research is hindered by high rediscovery rates, especially in classical producers such as Actinobacteria, Firmicutes and Myxobacteria – despite the strong genetic potential for the biosynthesis of novel molecules.<sup>5</sup> This limitation can be circumvented by focusing on underexplored bacterial taxa.

Members of the phylum Bacteroidetes are well-known, abundant and widely distributed sources of natural products, and many members are easy to cultivate. They colonize all types of

<sup>a</sup>Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Branch for Bioresources, 35392 Giessen, Germany. E-mail: stephan.brinkmann@ime.fraunhofer.de; marius.spohn@ime.fraunhofer.de; till.f.schaerberle@agrar.uni-giessen.de

<sup>b</sup>Institute for Insect Biotechnology, Justus Liebig University of Giessen, 35392 Giessen, Germany

<sup>c</sup>German Centre for Infection Research (DZIF), Partner Site Giessen-Marburg-Langen, Giessen, Germany

habitats, including soil, ocean, freshwater, and the gastrointestinal tracts of animals.<sup>6</sup> The phylum comprises six classes of gram-negative, non-spore forming, chemo-organotrophic bacteria. Gastrointestinal tracts harbor species from the mostly anaerobic class *Bacteroidia*, whereas classes *Chitinophagia*, *Cytophagia*, *Flavobacteriia*, *Saprospiria* and *Sphingobacteriia* are mainly found in the environment.<sup>7</sup> All of them are non-motile (or motile by gliding)<sup>8</sup> bacteria that can degrade polymeric organic matter.<sup>6</sup> A few are pathogenic<sup>9</sup> or endosymbiotic.<sup>10</sup> Their promoter structures feature a unique consensus sequence<sup>11</sup> recognized by the core RNA polymerase and an unusual primary sigma factor.<sup>11,12</sup> The strong genetic potential for the biosynthesis of natural products<sup>13,14</sup> is enriched in a few genera (e.g., *Chitinophaga* and *Pedobacter*) and less abundant in the anaerobic class *Bacteroidia*.<sup>14</sup> However, the number of compounds isolated thus far is rather low. Known products are structurally diverse and some are potential antibiotic leads, such as the (iso)pedopeptins that can

inhibit drug-resistant gram-negative bacteria including World Health Organization (WHO) top-priority carbapenem-resistant pathogens, and the dipeptide TAN-1057, with activity against methicillin-resistant *Staphylococcus aureus* (MRSA) strains.

In this review, we comprehensively discuss the bioactivity and biosynthesis of natural products that have already been isolated from Bacteroidetes.

## 2 Macrolides

Macrolides are hydrophobic polyketides characterized by a large macrocyclic lactone ring decorated with variable side chains and groups. One of the better-known macrolides is erythromycin,<sup>15</sup> which has bacteriostatic activity based on the inhibition of bacterial protein biosynthesis.<sup>16</sup> Macrolides also show antifungal, antiparasitic, antiviral, cytotoxic and immunosuppressive activities.<sup>17</sup>



*Stephan Brinkmann received his bachelor's degree in Biotechnology and Instrumentation Engineering from the Bielefeld University of Applied Sciences in 2014 and his master's degree in Molecular Biotechnology from the University Bielefeld in 2017. In his following PhD thesis, he developed microfluidic-based high-throughput approaches for the cultivation and screening of microorganisms for their potential*

*to produce bioactive natural products and also analyzed the Bacteroidetes phylum for its potential to biosynthesize natural products using omics technologies at the Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Branch for Bioresources, in Giessen, Germany.*



*Marius Spohn received his master's degree in Microbiology in 2012 from the University of Tübingen. After his PhD studies focusing on the investigation of regulatory circuit's controlling microbial natural products biosynthesis and a short post-doctoral position he joined in 2016 the Fraunhofer Institute for Molecular Biology and Applied Ecology (IME). Working as the group leader for molec-*

*ular microbiology, he continues his research on microbial natural products in joint projects with partners from pharma-, agro- and veterinary industry.*

### 2.1 Elansolids

The elansolids are unique macrolides produced by *Chitinophaga sancti* FxGBF13 (DSM 21134)<sup>18–20</sup> and *Chitinophaga pinensis* DSM 2588.<sup>21</sup> They share a bicyclo[4.3.0]nonane core originating from an intramolecular Diels–Alder cycloaddition (Fig. 1).<sup>22</sup>

Elansolids A1 (1) and A2 (2) are atropisomers featuring a 19-membered macrolactone ring. They are derived from the precursor elansolid A3 (3) by lactonization *via* a intramolecular Michael addition of the quinone methide group.<sup>20,22</sup> The side products elansolids B1–3 (4–6) and D1–2 (7–8) are instead derived from the highly reactive quinone methide moiety in elansolid A3.<sup>20,21</sup> The cultivation of *C. sancti* DSM 21134 in complex media containing different soy peptones, or synthetic medium supplemented with anthranilic acid, yielded elansolid C1 (9).<sup>23</sup> This is derived *via* a nucleophilic addition of the anthranilic acid amino group to the *p*-quinone methide carbon of elansolid A3.<sup>23</sup>



*Till F. Schäberle studied biology and at the age of 29, he received his PhD from the University of Tübingen, Germany at the Microbiology/Biotechnology department. In 2009, he moved to the University of Bonn, Germany where he worked as a postdoc and subsequently was heading his own working group within the Institute for Pharmaceutical Biology. Currently, he is professor at the Justus-Liebig-*

*University Giessen and department head at the Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Branch Bioresources. His natural product (NP) research groups applies genetic, biochemical, bioinformatic and chemical analytical methods for NP biosynthesis/discovery.*

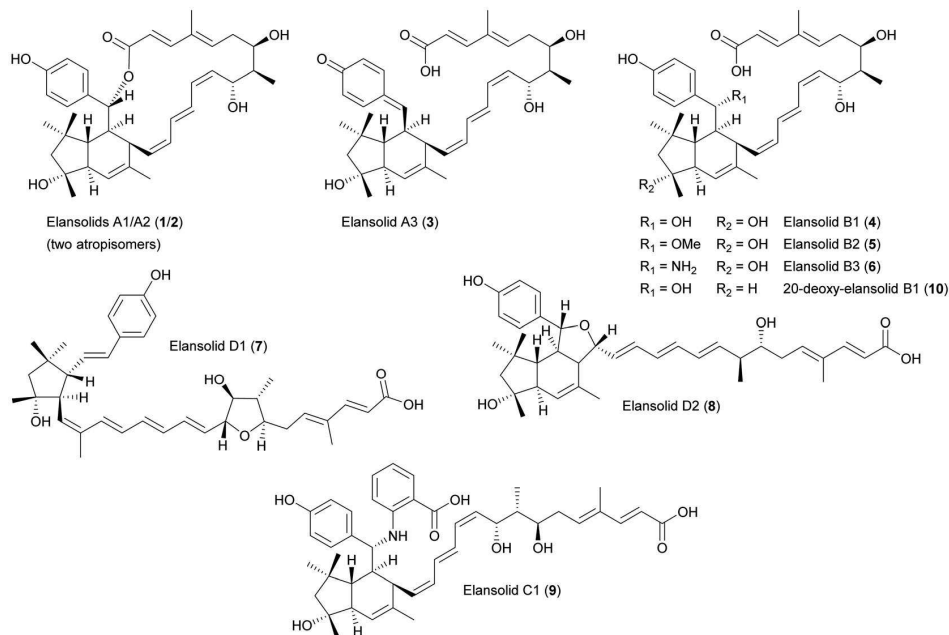


Fig. 1 Chemical structures of elansolids (1–10).

The elansolids are produced by *trans*-acyltransferase (AT) polyketide synthases (PKSs) (Fig. 2A).<sup>21,22</sup> The biosynthetic gene cluster (BGC) of *C. sancti* was identified by scanning a cosmid library.<sup>22</sup> In parallel, a closely related set of genes (84.6% nucleotide sequence identity) was identified in *C. pinensis* DSM 2588, facilitating the isolation of elansolid D2 (8) and the deduction of its biosynthetic pathway.<sup>21</sup> Elansolid A3 is probably the mature pathway product (Fig. 2B).<sup>21,22</sup> The two *trans*-AT type I polyketide BGCs consist of either 18 or 19 genes (*ela/els*

PKS) with their assembly line including six multimodular PKS proteins and two *trans*-ATs.<sup>21,22</sup> We also screened 600 publicly available Bacteroidetes genomes *in silico* to examine the sequential and compositional similarity of their BGCs, revealing *Chitinophaga* sp. YR627 as a third strain carrying a BGC highly similar to the known elansolid BGCs.<sup>14</sup> Interestingly, the comparison of all three BGCs suggests the presence of an additional gene encoding a major facilitator superfamily 1 protein, which may transport elansolids across membranes.

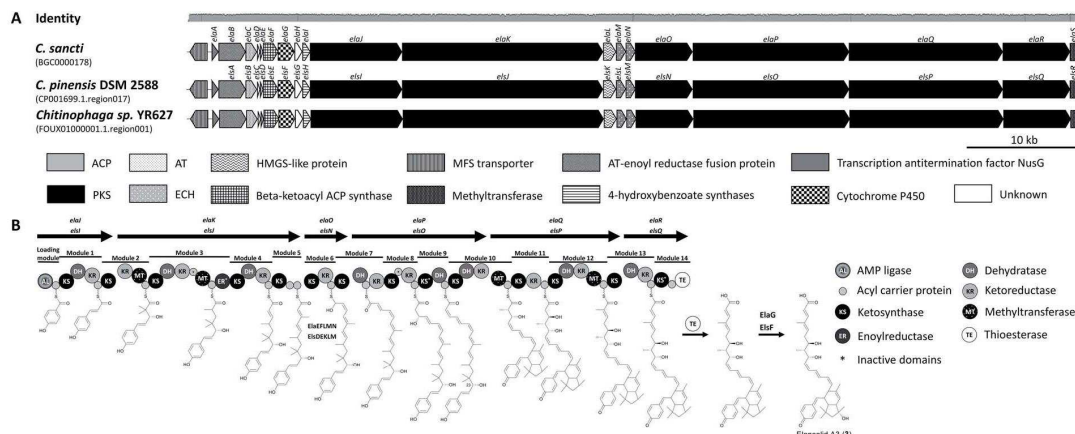


Fig. 2 Elansolid biosynthetic gene clusters (MIBIG BGC0000178) and pathway. (A) Elansolid BCCs found in genomes of three *Chitinophaga* strains. Identity represents the nucleotide alignment of all five BGCs using MAFFT.<sup>26</sup> (B) Model for elansolid biosynthesis.<sup>22</sup> The genes are indicated by arrows. The genes are denominated differently in the two source publications, so both names are given. The timing of the dehydration centered at C23 and subsequent IMDA reaction cannot be deduced with certainty. ACP, acyl carrier protein; PKS, polyketide synthase; AT, acyltransferase; MFS, major facilitator superfamily; HMGS, hydroxymethylglutaryl-CoA synthase; ECH, enoyl-CoA hydratase/isomerase.



Major facilitator superfamily proteins are membrane efflux pumps that contribute to the transport of various substances, including macrolides.<sup>24</sup>

Elansolids A1 (1) and A2 (2) are atropisomers, but only the latter is active against gram-positive bacteria, with minimal inhibitory concentrations (MICs) of 0.2–64  $\mu\text{g mL}^{-1}$ . This product also showed moderate cytotoxicity against L929 mouse fibroblasts, B104 rat cells, SH-SY5Y human neuroblastoma cells, and HeLa cells.<sup>18–21</sup> Furthermore, the highly reactive *p*-quinone methide elansolid A3 (3) can be stabilized as an ammonium salt, which showed antibacterial and cytotoxic activity identical to elansolid A2 (2).<sup>20</sup> This reactivity was used to synthesize an elansolid library from crude extracts *via* Michael-type conjugate addition, using 21 different nucleophiles. None of these precursor-directed derivatives was more potent than elansolids A2 (2) and A3 (3), but all were less cytotoxic than elansolid A2 (2).<sup>23</sup>

Elansolids A2 (2) and C1 (9) inhibit the secretion-deficient *Escherichia coli*  $\Delta\text{tolC}$  mutant when combined with the membrane-permeabilizing cationic cyclic peptide polymyxin B nonapeptide (PMBN) at 2 or 8  $\mu\text{g mL}^{-1}$ ,<sup>25</sup> indicating that a molecular target is also present in gram-negative bacteria but protection is achieved by its outer membrane.

The total chemical synthesis of elansolid B1 (4)<sup>26</sup> was followed by the second-generation total synthesis of the same compound and the first synthesis of elansolid B2 (5).<sup>27</sup> Further studies to elucidate the biosynthesis of elansolids<sup>28,29</sup> revealed a further polyketide, 20-deoxy-elansolid B1 (10), with activity similar to elansolid A2 (2), showing that the hydroxyl group at C20 is not essential for antibacterial activity.<sup>28</sup>

## 2.2 YL-02905-A and YL-02905-B

Two *Cytophaga* strains were found to produce four macrolactones with activity against *S. aureus*.<sup>30–32</sup> First reported in a patent,<sup>30</sup> the two 22-membered macrolide ring antibiotics YL-02905-A (also named YM-32890-B, 11) and YL-02905-B (also

named YM-32890-A, 12) were isolated from *Cytophaga* sp. YL-02905S.<sup>31</sup> Their natural isomers (13 and 14) complement this group.<sup>32</sup> The four macrolactones differ in the aliphatic chain attached to the macrolactone ring at C21 (Fig. 3). Accordingly, molecules 11 and 13 possess a conjugated tetrene system whereas 12 and 14 contain a conjugated triene system, separated by a methylene carbon at C28.<sup>31,32</sup> Compound 12 undergoes tautomerization to form 11 (ref. 32) and is generally unstable under light and aerobic conditions.<sup>31</sup> Only compounds 12 and 14 displayed activity against *S. aureus* KCTC 1927 with MICs of  $\sim 10$ – $20 \mu\text{g mL}^{-1}$ .<sup>32</sup> These data reflect the importance of the conjugation pattern of the double bonds along the aliphatic chain.<sup>32</sup> The biosynthesis or total synthesis of these compounds has yet to be reported.

## 3 Peptides

Peptide natural products are valuable sources of anti-infective leads and are either synthesized ribosomally or assembled by multi-modular non-ribosomal peptide synthetases (NRPSs), allowing the incorporation of non-proteinogenic amino acids. Optional (multi)cyclization and extensive post-translational modifications result in remarkable structural diversity.<sup>33</sup>

### 3.1 Cyclic depsipeptides

More than 1300 diverse natural cyclic depsipeptides have been described, and are characterized by the presence of ester and amide bonds that contribute to a range of biological properties.<sup>34,35</sup> Most cyclic depsipeptides are synthesized by modular NRPSs,<sup>36</sup> but some (*e.g.*, microviridins) are synthesized ribosomally.<sup>37</sup>

**3.1.1 (Iso)pedopeptins.** Natural products that inhibit lipopolysaccharide (LPS) binding to the receptor CD14 are useful leads against sepsis and septic shock. Screening for such products led to the discovery of pedopeptins A–C (15–17, patented as B-12489 A–C; Fig. 4).<sup>38–40</sup> These cyclic lipodepsipeptides (CLPs) were isolated from the culture broth of *Pedobacter* sp. SANK 72003.<sup>39,40</sup> They share three cationic amino acids – two diaminopropionic acid (Dap) and one diaminobutyric acid (Dab) residues – among a constituent of nine, including 2-amino-2-butyric acid (dehydrobutyrine, Dhb) and  $\beta$ -hydroxy valine ( $\beta$ -OH-Val), as well as an aliphatic chain.<sup>40,41</sup> UHPLC-MS/MS analysis revealed that similar cyclic and linear peptides are also produced by several other *Pedobacter* strains.<sup>42</sup> Detailed investigation of the strain *Pedobacter cryoconitis* UP508 led to the discovery of isopedopeptins A–H (18–25; Fig. 4).<sup>42–44</sup> These are similar to the pedopeptins, but Leu-2 and Phe are exchanged, Thr replaces Leu-1, and the acyl moieties are distinct.<sup>44</sup>

Pedopeptins A (15), B (16) and C (17) inhibit the binding of LPS to CD14 with half-maximal inhibitory concentrations ( $\text{IC}_{50}$ ) of 20, 11 and 47 nM, respectively. LPS-induced cytokine release in cell-based *in vitro* assays revealed dose-dependent inhibition, with  $\text{IC}_{50}$  values of 0.08–0.33  $\mu\text{M}$  ( $\sim 1$ – $10 \mu\text{g mL}^{-1}$ ). However, pedopeptins A and B were also cytotoxic at 100 and 30  $\mu\text{g mL}^{-1}$ , respectively. Antimicrobial testing of all three peptides revealed MICs of 2–4  $\mu\text{g mL}^{-1}$  against two *E. coli* strains. Interestingly, only

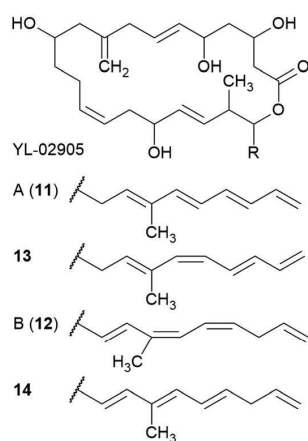


Fig. 3 Chemical structures of YL-02905 A (11) and B (12) as well as their structural isomers 13 and 14.

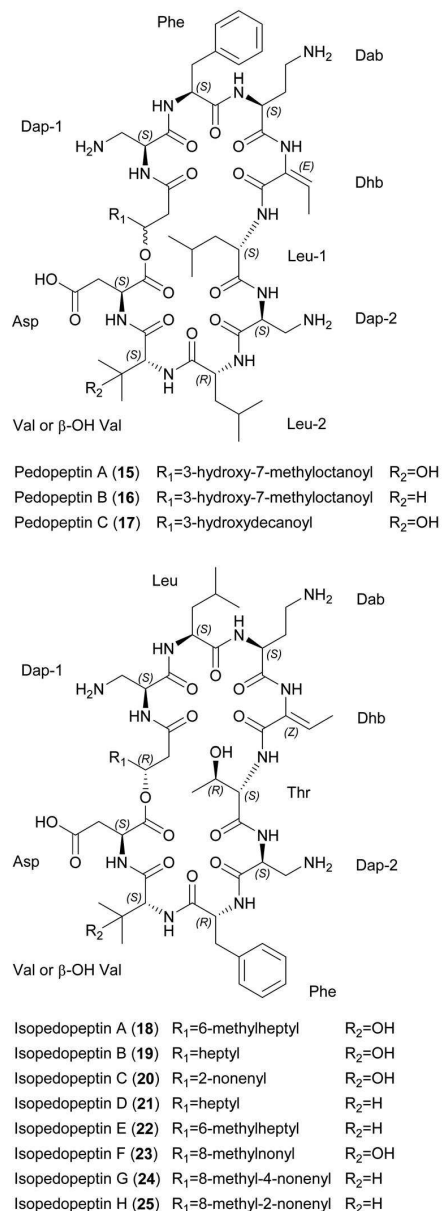


Fig. 4 Chemical structures of pedopeptins A–C (15–17) and isopedopeptins A–H (18–25). Dap, diaminopropionic acid; Dab, diaminobutyric acid; Dhb, dehydrobutyrine;  $\beta\text{-OH-Val}$ ,  $\beta$ -hydroxy valine.

pedopeptin B showed strong activity ( $4\ \mu\text{g mL}^{-1}$ ) against two gram-positive bacteria: *S. aureus* ATCC 6538P and *S. epidermidis* ATCC 14990.<sup>39</sup> All isopedopeptins show activity against a broad panel of gram-negative bacteria, including carbapenem-resistant strains on the WHO top-priority list. Isopedopeptin B (19) is the most potent representative, with MICs as low as  $1\ \mu\text{g mL}^{-1}$ . In contrast to the pedopeptins, isopedopeptins show minimal activity against gram-positive bacteria (e.g., MIC =  $16\ \mu\text{g mL}^{-1}$  against *Bacillus cereus* CCUG7414), and none against fungi.

Furthermore, isopedopeptins are cytotoxic against HepG2 liver cells ( $\text{IC}_{50} = 9\text{--}50\ \mu\text{M}$  or  $10\text{--}56\ \mu\text{g mL}^{-1}$ ) and are hemolytic (0.8–52.3%). Isopedopeptin B was therefore the most promising candidate due to its potent antibacterial activity and acceptable cytotoxicity.<sup>44</sup>

The mode of action (MoA) of the cationic pedopeptins is thought to involve interactions with lipid layers due to structural similarities with other cationic cyclic peptides such as polymyxin B and colistin, which are known to interact with anionic LPS.<sup>39,45</sup> Interestingly, the polycationic isopedopeptin B (19) showed activity against colistin-resistant strains of *Acinetobacter baumannii*, *E. coli*, and *Klebsiella pneumoniae* with a very low frequency of resistance ( $<3 \times 10^{-9}$ ).<sup>44</sup> Using an *in vitro* bacterial membrane system based on unilamellar liposomes prepared from *E. coli* phospholipid extracts, isopedopeptin F (23) was shown to cause membrane leakage with a half-maximal effective concentration ( $\text{EC}_{50}$ ) of  $2\ \mu\text{M}$ , a similar concentration to other cyclic peptides such as polymyxin B ( $\text{EC}_{50} = 1\ \mu\text{M}$ ) and LL-37 ( $\text{EC}_{50} = 0.6\ \mu\text{M}$ ).<sup>44</sup>

A putative BGC for the pedopeptins has been identified in the pedopeptin producer strain *Pedobacter lusitanus* NL19 (Fig. 5).<sup>46</sup> The NRPS is composed of a large and a small subunit. Genes encoding diaminobutyrate-2-oxoglutarate transaminase and oxygenase (cytochrome P450 family) are found nearby, and may be responsible for the supply of Dab as well as Val hydroxylation. The expression of both NRPS genes correlates with the production of pedopeptins. Furthermore, within this BGC, the C5 domain shows similarity to C-domains that modify incorporated amino acids. In this case, it should catalyze the dehydration of Thr (incorporated by A-domain 4) to Dhb. Genes encoding glutamate synthase and dehydrogenase were also found in the genome, and may be responsible for Dap biosynthesis.<sup>46</sup> A BGC for the isopedopeptins has yet to be described. However, the discovery of additional (iso)pedopeptin analogs in several other *Pedobacter* extracts indicates the existence of related BGCs, suggesting that small changes in A-domain specificity and/or module rearrangement within the pedopeptin BGC may facilitate the biosynthesis of various natural analogs, including the isopedopeptins.<sup>42</sup> The total synthesis of these cyclic peptides has not yet been reported.

**3.1.2 Chitinopeptins A–D.** The chitinopeptins are CLPs produced by *Chitinophaga* species. Chitinopeptins A (26) and B

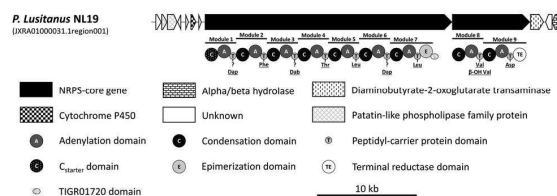


Fig. 5 Pedopeptin biosynthetic gene cluster. Non-ribosomal peptide synthetase biosynthetic gene cluster and assembly line responsible for the biosynthesis of pedopeptins. *In silico* predicted amino acids are indicated in the first row and those confirmed by NMR spectroscopy are underlined. Dap, diaminopropionic acid; Dab, diaminobutyric acid;  $\beta\text{-OH Val}$ ,  $\beta$ -hydroxy valine.

(27) were isolated from *C. eiseniae* DSM 22224, whereas chitinopeptins C1, C2, D1 and D2 (28–31) were isolated from *C. flava* KCTC 62435. They share many non-proteinogenic amino acids such as Dap and *N*-methyl-Val, and a high degree of  $\beta$ -hydroxylation (Asp, Phe and Ile residues). The tetradecalipodepsipeptides 26 (2,9-dimethyl-3-amino decanoic acid) and 27 (3-amino-9-methyl decanoic acid) differ in the fatty acid residue attached to the peptide core. In contrast, the pentadecalipodepsipeptides (28–31) carry a 3-hydroxy-9-methyl decanoic acid residue and an additional Dap residue inserted between the fatty acid and the cyclized peptide chain, and the Leu-2 present in molecules 26 and 27 is replaced with Ile (28 and 29) or Val (30 and 31). The

constitution of the Dap-2 residue (peptide bond between the  $\alpha$  or  $\beta$  amino groups and the carbonyl group of Lys) distinguishes between molecules 28 and 29 as well as between molecules 30 and 31 (Fig. 6).<sup>14</sup>

Chitinopeptins inhibit gram-negative and gram-positive bacteria such as *Moraxella catarrhalis* (MIC = 2  $\mu\text{g mL}^{-1}$ ) and *Bacillus subtilis* (MIC = 4  $\mu\text{g mL}^{-1}$ ) as well as the yeast *Candida albicans* (MICs  $\geq$  4  $\mu\text{g mL}^{-1}$ ). MICs of  $\geq$  16  $\mu\text{g mL}^{-1}$  were observed against *A. baumannii*, *Micrococcus luteus* and *Zysoseptoria tritici*. Interestingly, the chitinopeptins bind iron, which in the case of chitinopeptin A (26) reduces but does not abolish its activity.<sup>14</sup>

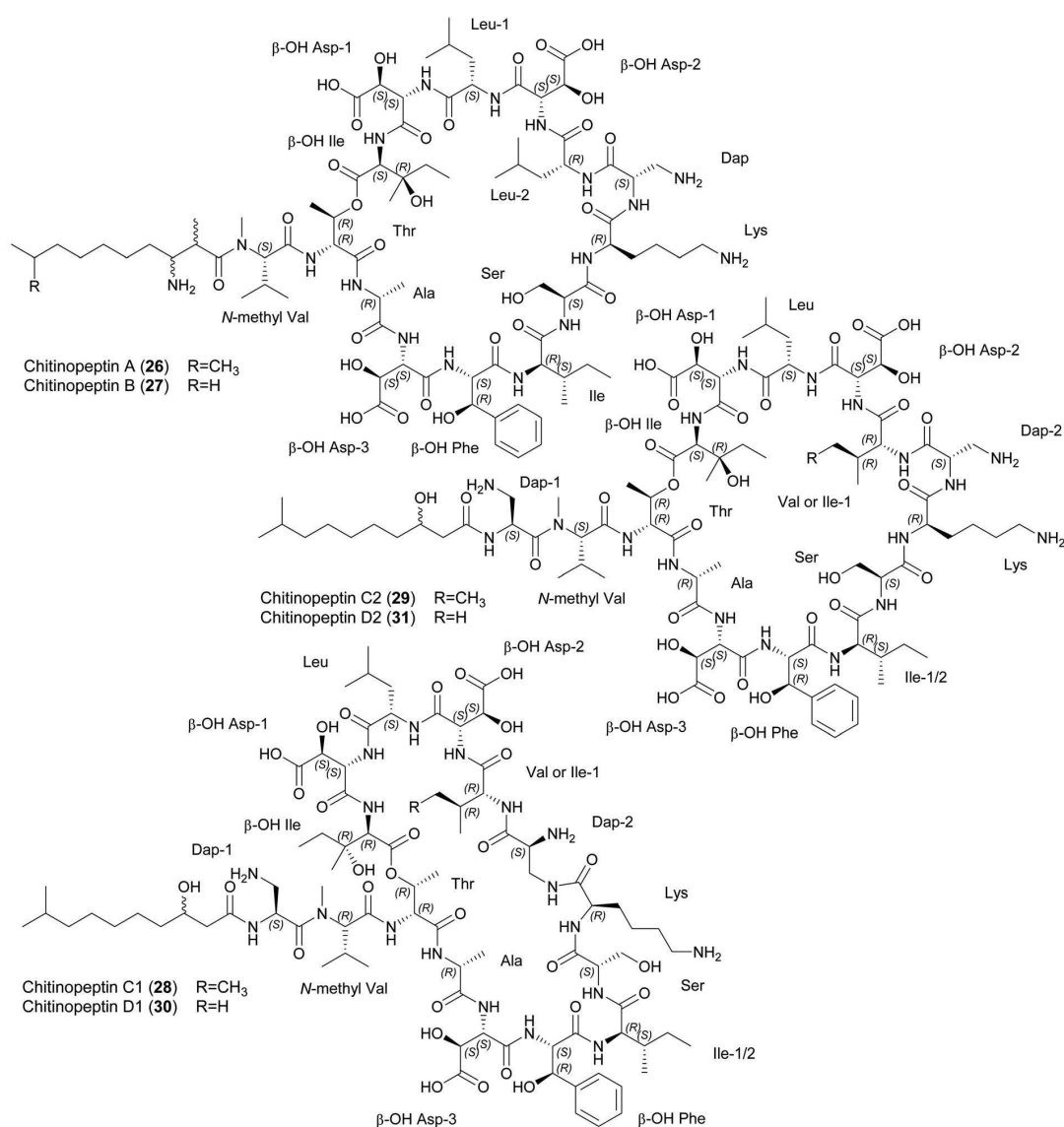


Fig. 6 Chemical structures of chitinopeptins A–D (26–31).  $\beta$ -OH Asp,  $\beta$ -hydroxy aspartic acid;  $\beta$ -OH Phe,  $\beta$ -hydroxy phenylalanine; Dap, diaminopropionic acid;  $\beta$ -OH Ile,  $\beta$ -hydroxy isoleucine.

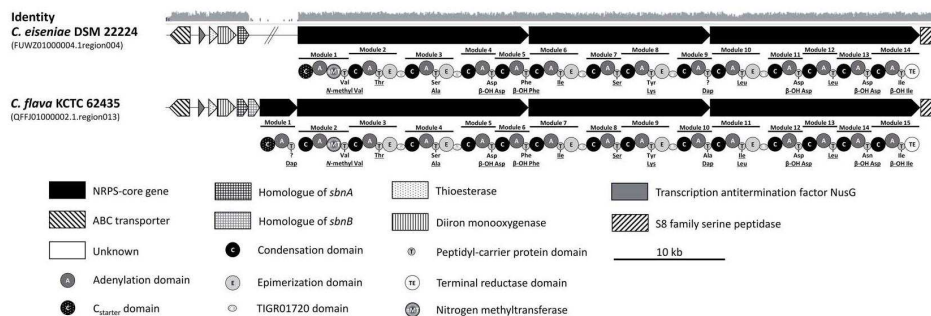


Fig. 7 Chitinopeptin biosynthetic gene clusters. Non-ribosomal peptide synthetase biosynthetic gene clusters and assembly lines responsible for the biosynthesis of chitinopeptins. *In silico* predicted amino acids are indicated and those confirmed by NMR spectroscopy are underlined.  $\beta$ -OH Asp,  $\beta$ -hydroxy aspartic acid;  $\beta$ -OH Phe,  $\beta$ -hydroxy phenylalanine; Dap, diaminopropionic acid;  $\beta$ -OH Ile,  $\beta$ -hydroxy isoleucine.

Putative NRPS-BGCs have been identified for the synthesis of these CLPs (Fig. 7). The number and predicted substrate specificity of the A-domains, the overall composition of the NRPS assembly line, and genes encoding enzymes responsible for precursor supply and post-assembly modifications matched the structural composition of the chitinopeptins. A BGC-similarity networking approach revealed the presence of homologous BGCs in *C. oryzziterrae* JCM16595 and *C. niastensis* DSM 24859.<sup>14</sup> The total synthesis of chitinopeptins has yet to be disclosed. The synthesis of the  $\beta$ -hydroxylated amino acids Asp, Phe and Ile has been described as an attempt to elucidate the absolute configuration of these amino acids in all chitinopeptin structures.<sup>14</sup>

**3.1.3 Katanosins A and B.** Two groups independently elucidated the structure of the same cyclic depsipeptide antibiotic while searching for novel cell wall inhibitors. This molecule was isolated from the producer strain *Cytophaga* PBJ-5356 and was named katanosin B (32)<sup>47,48</sup> but was also isolated

from *Lysobacter* sp. ATCC 53042 and was named lysobactin.<sup>49–51</sup> *Cytophaga* PBJ-5356 also produces the analog katanosin A (33), which is almost identical to katanosin B but has a Val residue replaced with Ile.<sup>47,48</sup> Both compounds consist of 11 amino acids, six of which are non-proteinogenic, including  $\beta$ -hydroxylated Phe, Asn and Leu (Fig. 8).<sup>48,51</sup> They show moderate activity against some gram-negative bacteria and strong activity against a broad panel of gram-positive bacteria including MRSA (MIC = 0.39  $\mu\text{g mL}^{-1}$ ) and vancomycin-resistant enterococci (MIC = 0.78  $\mu\text{g mL}^{-1}$ ).<sup>47,49,50</sup>

Both molecules target peptidoglycan synthesis. To gain insight into the MoA, staphylococcal transglycosylation and its preceding reactions were studied by incorporating [<sup>14</sup>C]GlcNAc into *S. aureus* peptidoglycan in whole cells, as well as nascent peptidoglycan formation *in vitro*. Both peptides inhibit nascent peptidoglycan synthesis, but only katanosin B also inhibits lipid intermediate formation.<sup>52</sup> Lipid II was identified as the cellular target of katanosin B.<sup>53</sup> The MoA of katanosins A and B differs from that of vancomycin and the peptides show a low frequency of resistance, making them suitable for further development. Accordingly, katanosin B and its method of production have been patented.<sup>54</sup>

A cloning study with a *Lysobacter* sp. ATCC 53042 genomic library revealed a partial NRPS gene (4.6 kb) as part of the BGC.<sup>55</sup> The complete BGC containing two multi-modular NRPSs (*lybA* and *lybB*) was discovered by genome sequencing (Fig. 9).<sup>56</sup> The termination module showed a tandem thioesterase architecture and both thioesterases were biochemically characterized *in vitro*. Peptide cyclization and simultaneous release is only mediated by the penultimate thioesterase. The second thioesterase fulfills a proofreading function involving the deacylation of miss-primed peptidyl carrier proteins.<sup>56</sup> Synthetic access was established based on modular, solution-phase approaches<sup>57</sup> and by solid-phase peptide syntheses.<sup>58</sup>

**3.1.4 YM-47141 and YM-47142.** The two macrocyclic depsipeptides YM-47141 (34) and YM-47142 (35), carrying a hydrated vicinal tricarbonyl moiety, were isolated from *Flexibacter* sp. Q17897.<sup>59,60</sup> Their structure comprises a peptide chain containing the non-proteinogenic amino acid 2,3-dioxo-4-amino-6-methylheptanoic acid (Dah), which has not been found

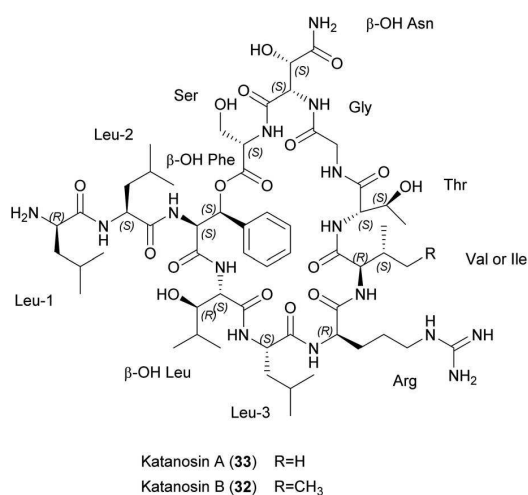


Fig. 8 Chemical structures of katanosins A (33) and B (32).  $\beta$ -OH Phe,  $\beta$ -hydroxy phenylalanine;  $\beta$ -OH Asn,  $\beta$ -hydroxy asparagine;  $\beta$ -OH Leu,  $\beta$ -hydroxy leucine.

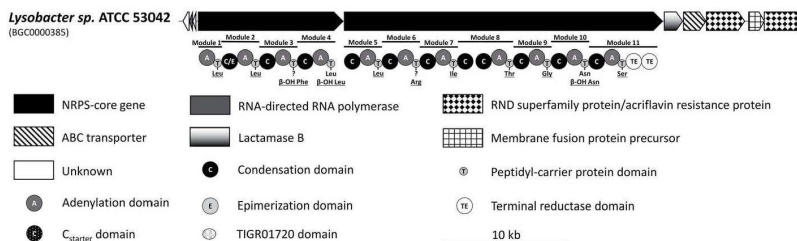


Fig. 9 Lysobactin/katanosin biosynthetic gene cluster (MIBiG BGC0000385). Non-ribosomal peptide synthetase biosynthetic gene cluster and assembly line responsible for the biosynthesis of lysobactin/katanosin. *In silico* predicted amino acids are indicated and those confirmed by NMR spectroscopy are underlined.  $\beta$ -OH Phe,  $\beta$ -hydroxy phenylalanine;  $\beta$ -OH Leu,  $\beta$ -hydroxy leucine;  $\beta$ -OH Asn,  $\beta$ -hydroxy asparagine.

before in microbial natural products. Five of the six remaining amino acids were found to have the *L*-configuration, in addition to *D*-Ala. The isovaleryl moiety in YM-47141 (**34**) is replaced by a phenylacetyl group in YM-47142 (**35**) (Fig. 10).<sup>60</sup> Total synthesis of both compounds conformed the *L*-configuration of Dah.<sup>61</sup> Both compounds inhibited human leukocyte elastase (HLE) with  $IC_{50}$  values of 15 and 30  $\mu$ M, respectively.<sup>59</sup> YM-47141 (**34**) was also tested against porcine pancreatic elastase (PPE;  $IC_{50} = 0.39 \mu$ M), and was shown to inhibit other serine proteases, such as  $\alpha$ -chymotrypsin and cathepsin G.<sup>62</sup> The BGC for these compounds has yet to be reported.

### 3.2 Short linear peptides

Short linear peptides also display remarkably diverse biological activities with biomedical and cosmeceutical applications.<sup>63</sup> They tend to have a short *in vivo* half-life<sup>64</sup> and many are immunogenic,<sup>65</sup> but generally they are highly selective, can cross biological membranes, and cost-effective synthesis has been reported.<sup>63</sup>

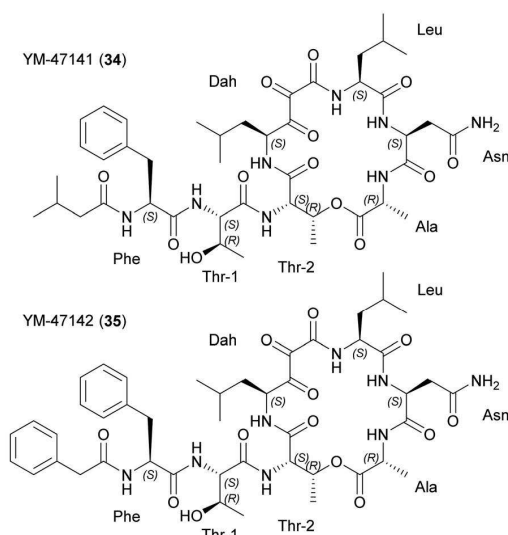


Fig. 10 Chemical structures of YM-47141 (**34**) and YM-47142 (**35**). Dah, 2,3-dioxo-4-amino-6-methylheptanoic acid.

**3.2.1 TAN-1057 A–D.** The dipeptide-like metabolites TAN-1057 A (**36**) and B (**37**) (Fig. 11) are diastereomers first described by researchers at Takeda Chemical Industries Ltd.<sup>66–68</sup> They are produced by *Flexibacter* sp. PK-74 and PK-176,<sup>67</sup> along with diastereomers TAN-1057 C (**38**) and D (**39**). TAN-1057 A and B spontaneously epimerize in aqueous solution, whereas TAN-1057 C and D only transform into TAN-1057 A and B, respectively.<sup>66,68</sup> All four compounds show strong activity against staphylococci, including MRSA strains, but they are less active against gram-negative bacteria.<sup>67</sup> Subcutaneous treatment of MRSA-infected mice revealed an effective dose ( $ED_{50}$ ) of  $\geq 0.026 \text{ mg kg}^{-1}$ , which is more potent than vancomycin ( $ED_{50} = 2.3 \text{ mg kg}^{-1}$ ). However, both TAN-1057 A and B are cytotoxic, with median lethal doses ( $LD_{50}$ ) in mice of 50  $\text{mg kg}^{-1}$  (intraperitoneal) and 100  $\text{mg kg}^{-1}$  (intravenous).<sup>67</sup> This has been addressed by total synthesis to investigate the structure–activity relationship (SAR).<sup>69</sup> A synthetic derivative with a methyl group on the proximal guanidine nitrogen was less potent (10-fold lower MIC) but the  $ED_{100}$  was  $< 0.5 \text{ mg kg}^{-1}$  (intravenous).<sup>70</sup>

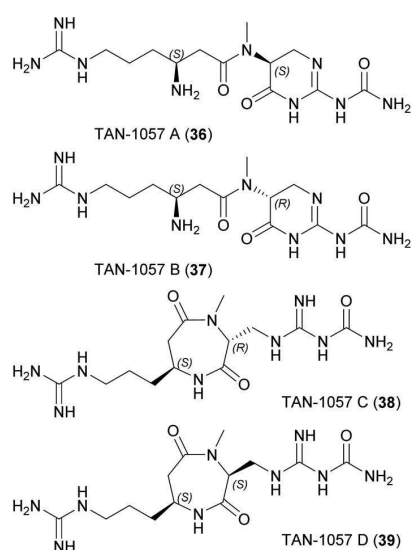


Fig. 11 Chemical structures of TAN-1057 A–D (**36–39**).

MoA studies with the TAN-1057 series showed a negative effect on protein synthesis at the elongation step in whole-cell and cell-free translation assays.<sup>67,71</sup> More detailed analysis revealed the target was the peptidyl transferase of the 50S ribosomal subunit, but the binding site appears to differ from that of other inhibitors.<sup>67,71,72</sup> Resistance studies indicated a rapid response based on changes in the transport of cellular dipeptides, followed by ribosomal alterations later on.<sup>71,73</sup> A BGC congruent with these dipeptide-like metabolites has yet to be identified.

**3.2.2 Falcitidin and pentacitidins A and B.** The linear *N*-acyl tetrapeptide falcitidin (**40**) was isolated from *Chitinophaga* sp. Y23, and features an unusual C-terminal amidated proline residue (Fig. 12).<sup>74</sup> Together with its synthetic derivatives, it represents a new class of cysteine protease inhibitors. In contrast to classical inhibitors, a reactive group such as an aldehyde that binds the active site Cys residue has not been identified.<sup>74,75</sup> Metabolic networking analysis revealed the presence of more than 30 natural analogs of falcitidin, including further falcitidin-like tetrapeptides.<sup>76</sup> They are produced by various *Chitinophaga* strains and two of these analogs were isolated from *Chitinophaga eiseniae* DSM 22224. Pentacitidins A (**41**) and B (**42**) are falcitidin-like pentapeptides with Phe as an additional C-terminal amino acid with an aldehyde group (phenylalaninal, H-Phe-al; Fig. 12). NMR spectroscopy revealed that both pentacitidins were purified as diastereomers for the Phe residue.<sup>76</sup> Interestingly, a pentapeptide and its corresponding falcitidin tetrapeptide analog with a C-terminal amidated proline were detected by UHPLC-MS, following the NMR analysis of pure pentacitidins, suggesting that falcitidin-like tetrapeptides are the degradation products of the mature natural products – the pentacitidins.<sup>76</sup> Further studies are needed to test this hypothesis and elucidate the degradation mechanism.

Falcitidin displays falcipain-2 enzyme activity *in vitro* ( $IC_{50} = 6 \mu M$ )<sup>74</sup> but lacks whole-cell activity against *Plasmodium falciparum* 3D7 ( $IC_{50} > 10 \mu M$ ).<sup>75</sup> A liquid-phase synthesis approach produced an *N*-acyl trifluoromethyl derivative with a hydrophobic *N*-trityl group on the imidazole residue. In contrast to the isolated natural product, this molecule exhibits whole-cell

antimalarial activity ( $IC_{50} = 0.14 \mu M$ ).<sup>75</sup> Pentacitidins were produced by solid-phase peptide synthesis.<sup>76</sup> Aldehydes are electrophilic warheads that form covalent-reversible hemithioacetal adducts with catalytic Cys residues and covalent-reversible hemiacetal adducts with Ser residues. Testing pentacitidins against human cysteine proteases (cathepsins B and L) and related proteases from parasites (falcipain-2/3 and rhodesain) revealed selective inhibition of the parasitic enzymes with sub-micromolar  $IC_{50}$  values. In addition, the natural products were active against  $\alpha$ -chymotrypsin. Docking studies showed favorable distances between the catalytic Cys residue and the pentacitidin aldehyde in its *L*-configuration.<sup>76</sup>

A pentacitidin BGC was tentatively identified, consisting of a single NRPS gene (18.5 kb) with five A-domains and a C-starter unit (Fig. 13).<sup>76</sup> It is congruent with the pentapeptide structure, strengthening the degradation hypothesis. The BGC contains a terminal reductase domain probably involved in the formation of the C-terminal aldehyde and alcohol analogs *via* reductive release. These alcohol analogs of pentacitidins were detected by molecular networking analysis and synthesized as pentacitidin intermediates, but showed no bioactivity.<sup>76</sup> Similar BGCs were found in the genomes of four other *Chitinophaga* strains.<sup>76</sup>

### 3.3 Ribosomally synthesized and post-translationally modified peptides (RiPPs)

There are more than 20 subclasses of RiPPs, and scientific interest in these peptides has increased in recent decades.<sup>77</sup> A few RiPPs have already been isolated from Bacteroidetes, and many more can be expected due to development of new genome mining tools for RiPP discovery.<sup>78</sup>

**3.3.1 Lantipeptides.** Lantipeptides are RiPPs containing the thioether cross-linked amino acids lanthionine and/or methylanthionine. Most are produced by gram-positive bacteria, including the natural food preservative nisin from *Lactococcus* species.<sup>79</sup> Only a few such molecules have been isolated from gram-negative bacteria but genome analysis of the Bacteroidetes has revealed abundant putative lantipeptide BGCs.<sup>78</sup>

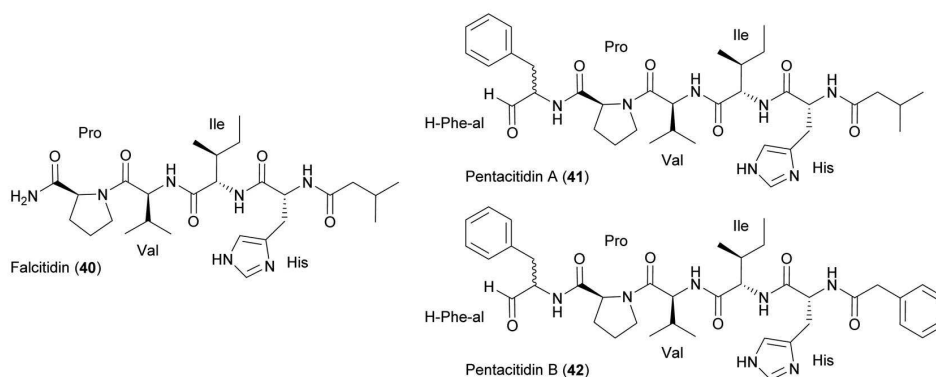


Fig. 12 Chemical structures of falcitidin (**40**) as well as pentacitidins A (**41**) and B (**42**).

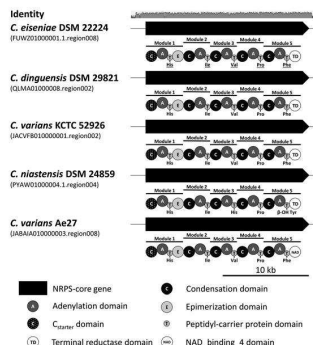


Fig. 13 Pentacitidin biosynthetic gene clusters. Non-ribosomal peptide synthetase biosynthetic gene clusters and assembly lines responsible for the biosynthesis of pentacitidins. *In silico* predicted amino acids are indicated and those confirmed by NMR spectroscopy are underlined. Identity represents the nucleotide alignment of all five BGCs using MAFFT.<sup>156</sup>  $\beta$ -OH Tyr,  $\beta$ -hydroxy tyrosine.

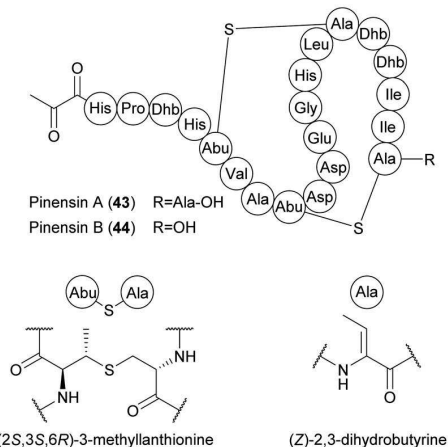


Fig. 14 Chemical structures of pinensins A (43) and B (44). Dhb, dehydrobutyrine; Abu,  $\alpha$ -aminobutyric acid.

3.3.1.1 *Pinensins A and B*. Pinensins A (43) and B (44) are produced by *C. pinensis* DSM 28390. These two cyclic lantipeptides contain 22 and 21 amino acids, respectively, with a rare N-terminal pyruvic acid. Both incorporate two 3-methyl-

lantionine-bridged ring systems with an unusual (2*S*,3*S*,6*R*)-stereochemical configuration (Fig. 14).<sup>80</sup>

As a 1 : 1 mixture, pinensins A and B exhibit broad activity against filamentous fungi and yeast ( $MIC > 5 \mu g mL^{-1}$ ) but only weak antibacterial activity. These lantipeptides also inhibit the growth of higher eukaryotic cells ( $IC_{50} > 10 \mu M$ ). The pinensin BGC predicted *in silico* consists of at least five genes encoding the lantipeptide precursor peptide (PinA), two dehydratases (PinB1 and PinB2), one lantionine synthetase (PinC) and PinT, required for export and cleavage of the precursor peptide (Fig. 15A). An unusual feature of pinensin biosynthesis is the split dehydratase, with functional domains catalyzing glutamylation and elimination encoded by separate open reading frames.<sup>78,80</sup> The precursor peptide PinA has 22 residues like pinensin A, suggesting that pinensin A is converted into pinensin B by the post-translational removal of the terminal Ala residue, but the enzyme responsible remains elusive.<sup>80</sup> Pinensin-like BGCs have also been identified in the genomes of other *Chitinophaga*,<sup>14</sup> *Cryseobacterium*, *Elizabethkingia*, *Pedobacter* and *Sinomicrobium* strains.<sup>78</sup> Predicted variations in the amino acid sequence of the core peptides indicate a yet undiscovered structural diversity of pinensin-type lantipeptides. A total synthesis approach for these peptides has not been reported.

3.3.1.2 *Peda peptides*. The computational prediction of precursor peptide sequences revealed several BGCs that may provide novel lantipeptides.<sup>78</sup> A BGC from *Pedobacter lusitanus* NL19 encoding two precursor peptides (Fig. 15B) could not be elicited in the native producer, so a heterologous expression strategy was implemented to gain access to these compounds. Co-expression of precursor peptide genes *pedA15.1* and *pedA15.2* in *E. coli*, together with their modifying enzymes, led to the discovery of the Peda15 peptides.<sup>81</sup> Only the structure of Peda15.1 (45) was elucidated in detail by NMR analysis (Fig. 16). This contains a novel 14-membered bis-lantionine ring system in the  $\mu$ -lantionine stereochemical configuration. The 4-fold or 5-fold dehydrated peptide carries three lantionine bridges between Dha7 and Cys16 nested with the Cys8–Dha15 ring, and a final ring between Dha18 and Cys21. This lantipeptide ring pattern has not been described in other lantipeptides. Peda15.1 carries further dehydrated but non-cyclized amino acids as well as a Ser that escaped dehydration in the heterologous host. The biosynthetic machinery of *E. coli* thus allows the production of modified peptides, but the heterologous products may not be

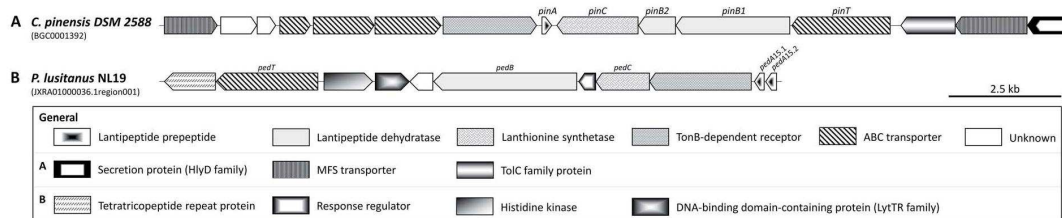


Fig. 15 Biosynthetic gene clusters of the ribosomally synthesized and post-translationally modified peptides pinensin (MIBiG BGC0001392) (A) and pedA15.1 (B).

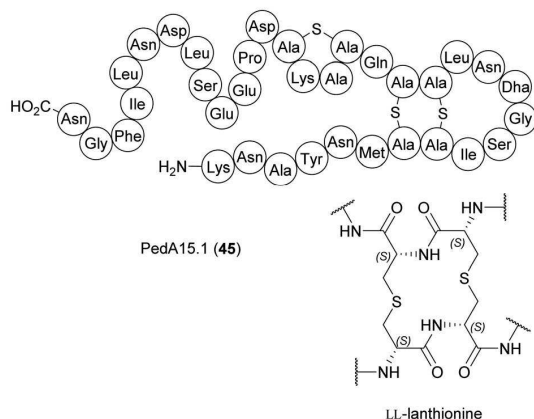


Fig. 16 Chemical structure of PedA15.1 (45). Dha, dehydroalanine.

the same as those synthesized in the native host, potentially explaining why PedA15.1 and PedA15.2 lack antibacterial and antifungal activity.<sup>81</sup> As for the pinensins, the total synthesis of PedA15.1 has not been described.

**3.3.2 Bacteriocin-like peptide FR901451.** The elastase inhibitor FR901451 (46) was isolated from *Flexibacter* sp. 758.<sup>82</sup> The tri-macrocyclic architecture of this patented compound<sup>83</sup> (Fig. 17) is based on 11 proteinogenic L-amino acids that are cyclized *via* two lactone bonds and one amide bond.<sup>82,84</sup> A substrate-competitive mode of inhibition was observed against PPE and HLE, with *in vitro* IC<sub>50</sub> values of 0.27 and 0.23  $\mu\text{M}$ , respectively.<sup>82</sup> In hamsters, FR901451 prevented HLE-induced lung hemorrhage *in vivo* and modulated PPE-induced changes to respiratory mechanics.<sup>85</sup> Crystallographic studies confirmed that FR901451 interacts with the substrate-binding site of PPE.<sup>84</sup> The BGC is unknown and total synthesis has not been reported for this compound. However, a BGC has been reported for the similar protease inhibitors microviridins B and J. A ribosomally synthesized precursor peptide is modified by two cyclases related to GRASP-like ligases that introduce two ester bonds and one amide bond<sup>37</sup>

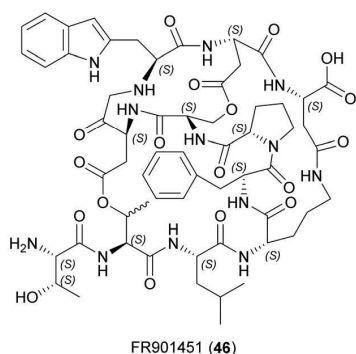


Fig. 17 Chemical structure of FR901451 (46).

## 4 Polyketide – non-ribosomal peptide hybrids

Modular PKS and NRPS systems present analogous assembly lines with similar mechanisms used to incorporate building blocks in the nascent chain. Genes encoding enzymes involved in either polyketide or non-ribosomal peptide synthesis can be combined, resulting in hybrid natural products that greatly expand the potential chemical diversity.<sup>86</sup> One example is the immunosuppressant rapamycin, with a pipercolate group in its polyketide skeleton.<sup>87</sup> Several of these complex polyketide–amino acid/peptide hybrids, featuring both acyl and aminoacyl building blocks, have been described in Bacteroidetes.

### 4.1 Ariakemicins A and B

Ariakemicins A (47) and B (48) were isolated from a marine gliding bacterium of the genus *Rapidithrix* and were described as an inseparable mixture.<sup>88</sup> The linear structures carry on one side an aromatic moiety substituted with one hydroxyl and one methoxy group, and on the other a methylated oxazol moiety decorated with a primary amide. The two compounds differ in the location of the double bond, which is found in position  $\Delta^{12}$  and  $\Delta^{11}$  in ariakemicins A and B, respectively (Fig. 18). The ariakemicin mixture was shown to inhibit gram-positive bacteria, with MICs ranging from 0.46  $\mu\text{g mL}^{-1}$  against *S. aureus* IFO12732 to 83  $\mu\text{g mL}^{-1}$  against *B. subtilis* IFO3134, but was inactive against gram-negative bacteria and *C. albicans*. Cytotoxicity was observed against human lung cancer cell line A549 and baby hamster kidney (BHK) cells with IC<sub>50</sub> values of 25 and 15  $\mu\text{g mL}^{-1}$ , respectively.<sup>88</sup> We were able to dereplicate ariakemicin A in extracts from *Rapidithrix thailandica* s80 (ref. 89) and subsequent isolation yielded pure ariakemicin A (47). This was active against *E. coli* ATCC25922  $\Delta\text{tolC}$  (our unpublished results), indicating that a molecular target for this compound is also present in gram-negative bacteria. One drawback of the ariakemicins is their rapid decomposition and loss of activity.<sup>88</sup> This may also explain why these compounds have not yet been synthesized.

## 5 Quinolines

The heterocyclic aromatic compound quinoline was first extracted from coal tar in the 18<sup>th</sup> century, and its structural motif can be found in many plant alkaloids, including quinine. However, there are also prominent bacterial examples, such as the antifungal compound pyrrolnitrin in *Pseudomonas* spp.<sup>90</sup>

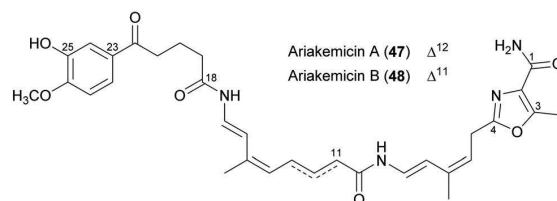


Fig. 18 Chemical structures of ariakemicins A (47) and B (48).



This aminophenylpyrrole alkaloid became a lead structure for the development of foliar fungicides. Additional structural variants have been isolated mainly from marine bacteria of the order *Cytophagales*.

### 5.1 G1499-2

The first quinoline isolated from the genus *Cytophaga* was 3-methyl-2-(2-pentylidencyclopropyl)-4-quinolinone, also known as G1499-2 (**49**) (Fig. 19).<sup>91</sup> It has weak activity against a few bacteria, such as *Flavobacterium* sp. 1980, a marine species, and no toxicity in mice. The BGC is unknown and total synthesis has not been reported for this compound.

### 5.2 Marinoquinolines and marinoazepinones

Several natural products featuring a rarely described 3*H*-pyrrolo [2,3-*c*]quinoline ring system have also been identified (Fig. 20), beginning with marinoquinoline A (**50**), which was isolated from *Rapiditythrix thailandica*.<sup>92</sup> This product is an acetylcholinesterase inhibitor with an  $IC_{50}$  value of 4.9  $\mu$ M.<sup>93</sup> Five derivatives known as marinoquinolines B–F (**51–55**) are produced by *Ohtaekwangia kribbensis*.<sup>94</sup> All five showed toxicity against three cancer cell lines in the low micromolar range (0.3–8.0  $\mu$ g mL<sup>-1</sup>) and two (**51** and **55**) showed potent activity against *P. falciparum* K1 ( $IC_{50}$  = 1.8 and 1.7  $\mu$ M, respectively).

Several more marine Bacteroidetes were isolated by cultivation in seawater-based media, leading to the discovery of marinoquinolines G–K (**56–60**).<sup>95</sup> Furthermore, the prolific producer strain *Mooreia alkaloidigena* CNX-216<sup>T</sup> synthesizes the structurally-related marinoazepinones A (**61**) and B (**62**), as well as other alkaloids.

To investigate quinoline biosynthesis in Bacteroidetes, the BGCs predicted *in silico* were expressed in *E. coli* along with other key enzymes for *in vitro* assays.<sup>96</sup> This revealed that only a single enzyme-catalyzed reaction is essential, in which tryptophan (**63**) is converted to 3-(2'-aminophenyl)-pyrrole (**64**), which is another potent acetylcholinesterase inhibitor already isolated from a *Rapiditythrix* strain.<sup>97</sup> This active intermediate can be coupled to a variety of aldehydes *via* a spontaneous Pictet–Spengler-like reaction under physiological conditions, producing diverse products using a minimal complement of enzymes (Fig. 21).

The synthesis of a pyrroloquinoline system has been achieved using several approaches,<sup>98–102</sup> including the synthesis of four different marinoquinolines *via* Heck–Matsuda arylation and Pictet–Spengler cyclisation.<sup>100</sup> Brønsted acid-promoted arene-ynamide cyclization was applied to produce high yields

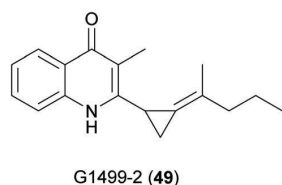


Fig. 19 Chemical structure of G1499-2 (**49**).

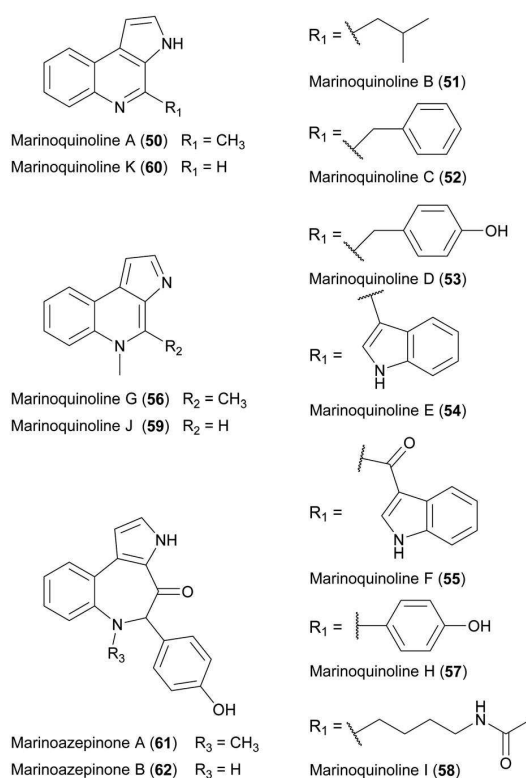


Fig. 20 Chemical structures of marinoquinolines A–K (**50–60**) as well as marinoazepinones A (**61**) and B (**62**).

of marinoquinolines A (**50**) and C (**52**).<sup>101</sup> More recently, the total synthesis of marinoquinoline A was achieved in four steps starting with the palladium-catalyzed Ullmann coupling of *o*-bromonitrobenzene to iodinated pyrrole.<sup>102</sup> Also, an intramolecular Diels–Alder oxazole cycloaddition reaction was used to produce marinoquinoline A from oxazole with an overall yield of 12%.<sup>99</sup>

## 6 Lactams

The clinical introduction of penicillin in the 1940s is a milestone in medical history. Even today,  $\beta$ -lactam antibiotics are

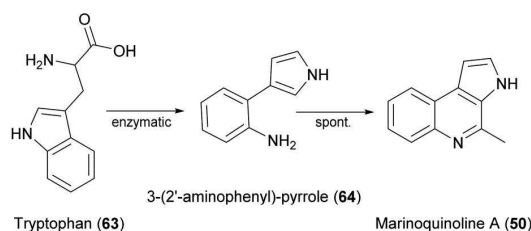


Fig. 21 Tryptophan (**63**) is enzymatically converted to 3-(2'-aminophenyl)-pyrrole (**64**). This intermediate reacts spontaneously with a variety of aldehydes, resulting in marinoquinolines such as marinoquinoline A (**50**).

## Review

ranked among the most important antimicrobials. Following the penicillins, further  $\beta$ -lactam antibiotic classes (e.g., cephalosporins, carbapenems and clavams) were discovered in fungi and actinomycetes. By expanding the search to include Proteobacteria and Bacteroidetes, the global  $\beta$ -lactam discovery effort culminated in the description of the monobactams in the early 1980s.<sup>103</sup> All reports on lactams from Bacteroidetes date back to a short period between 1982 and 1987, so there is not yet any information about the genetic or metabolic basis of such compounds and their total synthesis has not been reported.

## 6.1 Monobactams

**6.1.1 SQ 28332 (also known as PB-5582-A).** The Squibb Institute for Medical Research used a  $\beta$ -lactamase reporter assay to detect compounds with intact  $\beta$ -lactam rings regardless of their antibacterial activity. Many new  $\beta$ -lactams were identified in this manner, including several monobactams that lack an annulated ring system.<sup>103</sup> SQ 28332 or PB-5582-A (**65**) was the first  $\beta$ -lactam isolated from a Bacteroidetes strain (*Flexibacter* sp. ATCC 35208). It is a desmethoxy monobactam, in which the terminal  $\beta$ -lactam ring is attached to a branched tetrapeptide comprising D-alanine, N-methyl-L-serine, glycine and L-2,3-diaminopropionic acid, with an N-terminal D-glyceric acid (Fig. 22). This compound showed weak antimicrobial activity against gram-positive bacteria (MIC = 6–25  $\mu\text{g mL}^{-1}$ ) and no activity against gram-negative bacteria. Enzymatic analysis confirmed resistance against narrow-spectrum  $\beta$ -lactamases (P99 and TEM-2) but slight susceptibility to extended-spectrum  $\beta$ -lactamase K1. SQ 28332 is a weak inhibitor of the P99 cephalosporinase and does not inhibit TEM-2 or K1.<sup>104</sup>

**6.1.2 SQ 28502 and SQ 28503.** The Squibb screening assay also identified desmethoxy monobactams decorated with an oligopeptide side chain (SQ 28502 and SQ 28503), which were isolated from *Flexibacter* sp. ATCC 35103. These compounds are more complex than SQ 28332 and their molecular weight exceeds 1400,<sup>105</sup> but their precise structures are not yet known. Both compounds resist hydrolysis by  $\beta$ -lactamases TEM-2, K1 and P99 and indeed strongly inhibited P99 in a time- and concentration-dependent manner (TEM-2 and K1 were not inhibited). SQ 28502 and SQ 28503 showed weak activity against both gram-positive and gram-negative bacteria with SQ 28503 showing the greatest potency.<sup>105</sup>

**6.1.3 PB-5266 A–C.** Another series of desmethoxy monobactams was isolated from *Cytophaga* sp. ATCC 43843.<sup>106,107</sup> The three compounds PB-5266 A–C (**66–68**, Fig. 23) are structurally similar to SQ 28332 (**65**), but the terminal  $\beta$ -lactam ring is attached to a dipeptide, decorated with an N-terminal D-glyceric acid. All three structures also contain a dehydroasparagine

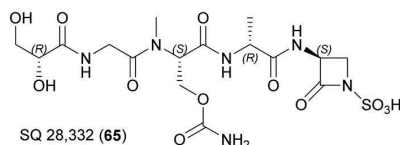


Fig. 22 Chemical structure of SQ 28332 (**63**).

## Natural Product Reports

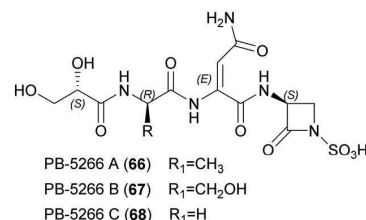


Fig. 23 Chemical structures of PB-5266 A–C (**64–66**).

residue. Structural variation can be observed in the second amino acid of the dipeptide, which is an Ala in **66**, Ser in **67** and Gly in **68**. The three compounds weakly inhibit an *E. coli* mutant with increased sensitivity to  $\beta$ -lactam antibiotics.<sup>106</sup>

**6.1.4 Formadicins A–D.** Formadicins A–D (**69–72**, Fig. 24) are produced by *Flexibacter alginoquefaciens* sp. nov. YK-49.<sup>108</sup> Similar to the nocardicidin-type monocyclic  $\beta$ -lactams, the  $\beta$ -lactam ring of the formadicins includes a phenylacetic acid group, contrasting with formadicins from the other Bacteroidetes monobactams (which are N-sulfonated). In addition, formadicins A (**69**) and B (**70**) are glycosylated  $\beta$ -lactams, carrying a  $\beta$ -D-glucuronic acid. This sugar, attached to the phenylacetic acid, is not present in formadicins C (**71**) or D (**72**). The formadicins also contain a rare formylamino group, from which they get their name. This group is bound directly to the  $\beta$ -lactam ring in formadicins A (**69**) and C (**71**) but to the side chain in formadicins B (**70**) and D (**72**).<sup>109</sup> Similar substitutions on the  $\beta$ -lactam ring confer resistance to hydrolysis by  $\beta$ -lactamase, as exemplified by the methoxylated cephamycins. Accordingly, formadicins A (**69**) and C (**71**) were strongly resistant to hydrolysis by penicillinases and cephalosporinases, but formadicins B (**70**) and D (**72**) were not.<sup>108</sup> The formadicins showed antibacterial activity in a narrow spectrum, particularly against *Pseudomonas*, *Proteus* and *Alcaligenes* species. Formacidin C (**71**) was the most potent (MIC  $\geq 3.13 \mu\text{g mL}^{-1}$ ). Competitive binding with labeled benzylpenicillin revealed that formadicins A (**69**) and C (**71**) bind with stronger affinity than the other formadicins to *Pseudomonas aeruginosa* penicillin-binding proteins 1a and 1b and offered greater protection to mice against infections with *E. coli* and *Proteus vulgaris*.<sup>109</sup>

## 6.2 Cepheids

**6.2.1 Desacetoxycephalosporin C.** Desacetoxycephalosporin C (**73**, Fig. 25) is an intermediate in the

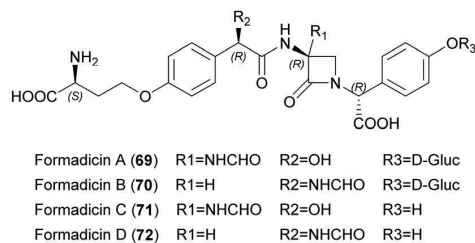


Fig. 24 Chemical structures of formadicins A–D (**67–70**).

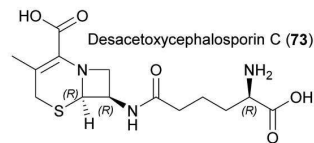


Fig. 25 Chemical structure of desacetoxycephalosporin C (71).

synthesis of cephalosporin C produced by several fungi and *Streptomyces* strains. Its synthesis requires the enzyme deacetoxycephalosporin-C synthase, which converts penicillins into cephalosporins.<sup>110</sup> This intermediate was also found in the culture broth of *Flavobacterium* sp. SC 12,154, revealing that Bacteroidetes can also synthesize more complex  $\beta$ -lactams, including those with annulated ring systems.<sup>111</sup>

**6.2.2 7a-Formylaminocephalosporins (SQ 28516 and SQ 28517).** *Flavobacterium* sp. SC 12154 also produces the 7a-substituted cephalosporins SQ 28516 (74) and SQ 28517 (75).<sup>112</sup> These  $\beta$ -lactam ring substitutions increase resistance to  $\beta$ -lactamase hydrolysis. The  $\beta$ -lactam ring in these compounds also features a rare formylamino structure (Fig. 26) like that found in the monocyclic formadecins (69 and 71). SQ 28516 and SQ 28517 are structurally unstable, so they were isolated as acetylated derivatives. The proposed structure of SQ 28516 carries a carboxylic group, whereas the minor fermentation product SQ 28517 is cationic at acidic pH and probably carries an amide instead. SQ 28516 showed no antibacterial activity and its ability to inhibit  $\beta$ -lactamase has not been reported.

**6.2.3 Chitinovorins A–D.** Chitinovorins are 7a-formylaminocephems in which an  $\alpha$ -amino adipic acid group is attached to the  $\beta$ -lactam ring. Chitinovorins A–D (76–79) differ in the nature of the substitutions in their six-membered dihydrothiazine ring (Fig. 27). Chitinovorin C (78) is the simplest, featuring a single hydroxymethyl group. In the other chitinovorins, this is expanded by an ester-linked peptide. In chitinovorin A, the side chain consists of a guanidine-containing building block and an Ala residue, whereas in chitinovorin B this expands to two Ala residues (77).<sup>113</sup> Chitinovorin D (79) is the largest, with two additional N-terminal constituents.<sup>114</sup> Chitinovorins A (76) and B (77) showed very weak antimicrobial activity ( $\text{MIC} = 25\text{--}50 \mu\text{g mL}^{-1}$ ),<sup>113</sup> and this was even lower for chitinovorins C (78) and D (79).<sup>114</sup>

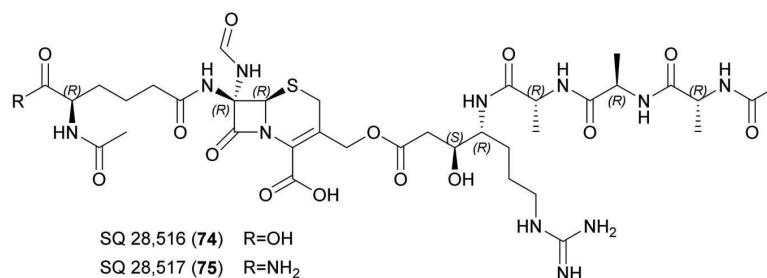


Fig. 26 Chemical structures of SQ 28516 (72) and SQ 28517 (73).

## 7 Additional Bacteroidetes natural products

### 7.1 Pigments

The best-known Bacteroidetes pigments are flexirubin-like compounds,<sup>115</sup> which are phylogenetically conserved and were used as chemotaxonomic markers for the phylum (formerly known as the *Cytophaga–Flavobacterium–Bacteroides* group) upon their discovery in *Chitinophaga filiformis* (formerly *Flexibacter elegans* Fx e1).<sup>116</sup> The total synthesis<sup>117</sup> and biosynthesis<sup>118</sup> of these pigments has been reviewed.<sup>115</sup> Flexirubin analogs have been evaluated for their antioxidant activity.<sup>119</sup> Other pigments, found in many Bacteroidetes, include carotenoids such as saxoxyanthin, myxol, zeaxanthin, flexixanthin and deoxyflexixanthin, some of which also possess antioxidant activity.<sup>120</sup> The flexixanthin BGC has been identified.<sup>121</sup>

### 7.2 Siderophores

Siderophores are iron-chelating compounds. Hydroxamate siderophores (tenacibactins A–D and bisucaberin B) have been isolated from *Tenacibaculum* spp.,<sup>122,123</sup> and fulvivirgamides from *Fulvivirga* sp. W222.<sup>124</sup> The bisucaberin B and fulvivirgamid BGCs have been experimentally validated.<sup>123–125</sup>

### 7.3 Lipids

*Chryseobacterium* spp. produce sulfonolipids (sulfobacin A and B) that are von Willebrand factor (vWF) receptor antagonists<sup>126</sup> and potential anti-inflammatory<sup>127</sup> and anticarcinogenic<sup>128</sup> leads. Flavocristamides A and B (whereby flavocristamide B is structurally identical to the before mentioned sulfobacin A) were isolated from a marine *Flavobacterium* species and were shown to inhibit DNA polymerase  $\alpha$ .<sup>129</sup> Other sulfonolipids (the capnoids capnine and *N*-acylcapnine)<sup>130</sup> were shown to be required for gliding motility in *Capnocytophaga* spp.<sup>131</sup> The *N*-acetylated sulfonolipids RIF-1 and RIF-2 (rosette inducing factors) and IOR-1 (inhibitor of rosettes) isolated from *Algoriphagus machipongonensis* were shown to influence morphogenesis in the choanoflagellate *Salpingoeca rosetta*.<sup>132</sup> The total synthesis<sup>133</sup> and biosynthesis<sup>134</sup> of most of these sulfonolipids has been investigated.

Other than sulfonolipids, the major lipids isolated from Bacteroidetes include lipoamino acids (LAAs), phospholipids

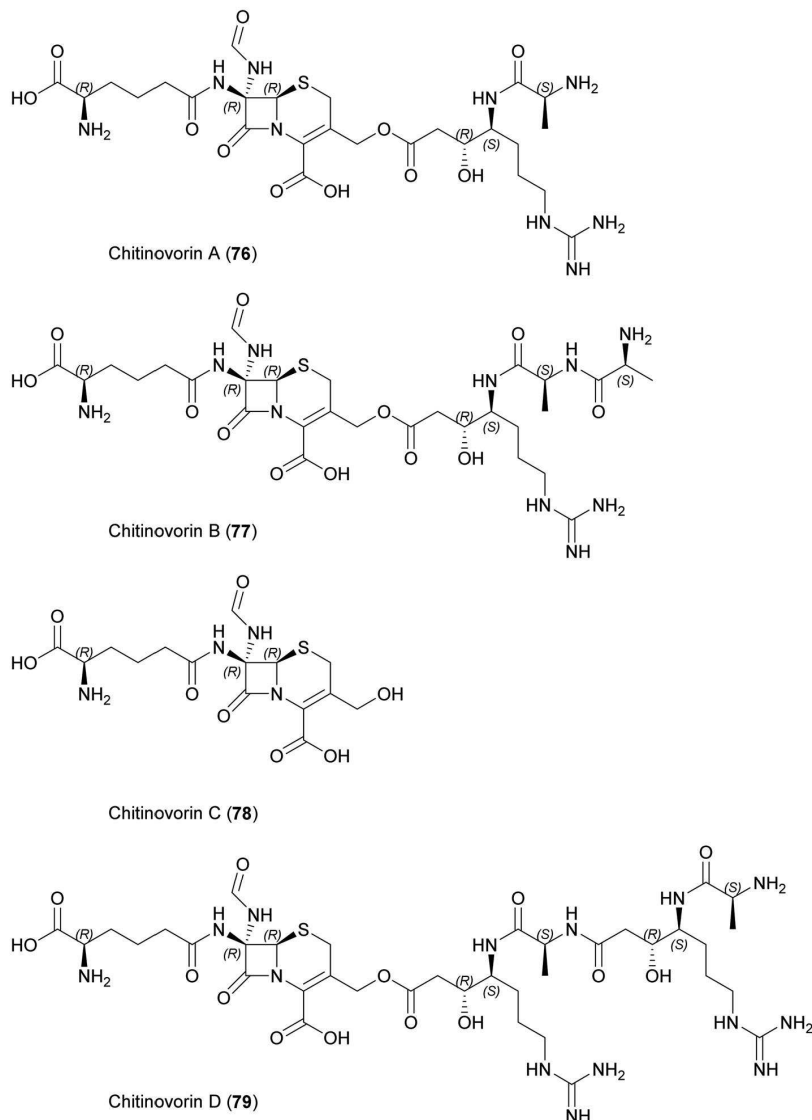


Fig. 27 Chemical structures of chitinovorins A–D (74–77).

and *N*-acyl amino acids (NAAAs). LAAs consist of a single de-esterified fatty acid or two fatty acids linked by esterification at C3 with different degrees of unsaturation. The hydrophobic fatty acid part is linked *via* an amide to Gly, Ser and ornithine, or a combination of two or three of these amino acids. The physicochemical properties and bioactivities differ significantly between stereoisomers in many cases, including WB-3559 A–D;<sup>135</sup> the topostins B553, B567 (also known as cytolipin), D640 and D654 (also known as WB-3559D),<sup>136–138</sup> and flavolipin<sup>139,140</sup> (also known as lipid 654 (ref. 141 and 142)).<sup>143</sup> In terms of biological properties, these LAAs act as mammalian DNA topoisomerase I inhibitors,<sup>136,137</sup> macrophage activators,<sup>144</sup> promoters of hemagglutination,<sup>140</sup> bacterial virulence

factors,<sup>145,146</sup> *N*-type calcium channel blockers,<sup>147</sup> and ligands for Toll-like receptor 2.<sup>141,142,148</sup> Accordingly, they are linked to the development of two chronic inflammatory diseases (periodontitis and atherosclerosis<sup>146</sup>), multiple sclerosis,<sup>149</sup> as well as showing antimicrobial activities.<sup>150</sup>

#### 7.4 Others

Other diverse compounds produced by Bacteroidetes include indole analogs. For example, *Cytophaga* strain AM13.1 produces 2,5-bis(3-indolylmethyl)pyrazine and pharacine, a natural *p*-cyclophane.<sup>151</sup> *Cytophaga johnsonae* AJ 12589 produces resorcinin, which stimulates the growth of NIH 3T3 mouse fibroblasts.<sup>152</sup> Furthermore, four neoverrucosane diterpenoids were

identified as products of *Saprosira grandis* ATCC 23116,<sup>153</sup> whereas the algal morphogen thallusin is produced by a *Zobellia* strain and has been synthesized.<sup>154</sup>

## 8 Concluding remarks

The phylum Bacteroidetes produces a wide range of metabolites that highlight its status as a valuable resource for natural products research. Although the isolation of interesting compounds from Bacteroidetes traces back to the golden era of natural product discovery (1950–1960), the total number of chemical entities is still remarkably low.

The recent identification of novel natural products is exemplified by the isopedopeptins and chitinopeptins, supporting the hypothesis that the natural product diversity in this phylum has yet to be exploited comprehensively. Of outstanding significance are the isopedopeptins, which show potent activity against gram-negative bacteria on the WHO's top-priority list, a low frequency of resistance, and acceptable levels of cytotoxicity, making them promising new candidate antibiotics.<sup>38,44</sup> The  $\beta$ -lactams are a more historical (but still extremely relevant) class of natural products that were isolated from Bacteroidetes in the 1980s. Their clinical use as antibiotics that inhibit bacterial penicillin-binding proteins and also act as  $\beta$ -lactamase inhibitors with the potential to break AMR is of great importance to international healthcare. Resistance to  $\beta$ -lactam antibiotics and fluoroquinolones is a major societal burden that accounts for more than 70% of all deaths attributed to AMR.<sup>1</sup> The  $\beta$ -lactams are still an essential drug class and a cornerstone of the current antibiotics market. Therefore, we must develop new solutions to counteract the steady loss of existing products. The naturally evolved strategy of combining  $\beta$ -lactam antibiotics with a  $\beta$ -lactamase inhibitor sharing the same warhead is therefore crucial to address the AMR crisis. There is a constant need for the development new  $\beta$ -lactamase inhibitors that maintain or even expand the antibacterial range of our clinical  $\beta$ -lactam/ $\beta$ -lactamase inhibitor combinations.<sup>155</sup> Although not directly pursued as a therapeutic strategy, we also consider  $\beta$ -lactams from the Bacteroidetes as very much part of this success story in learning from naturally evolved concepts and translating microbial diversity (and the corresponding structural and functional diversity of natural products) into life-saving drugs. Accordingly, the Bacteroidetes join other  $\beta$ -lactam producers such as the Actinobacteria, Proteobacteria and fungi.

Even compounds not followed up in the past such, as the highly active but unstable ariakemicin, may be suitable as candidates for repurposing strategies that aim to generate more applicable structures based on a natural scaffold. Following this approach, our recent omics studies revealed the immense genomic potential of this phylum that corresponds to uncharted chemical space, which differs significantly from the more widely investigated producer taxa.<sup>14</sup> In our opinion, this pronounced and unique biosynthetic capacity offers the potential for groundbreaking new compounds in the near future. Considering their historically proven role in the discovery of structurally diverse natural products, the Bacteroidetes have not received the attention they deserve during the

latest renaissance in natural products research. It is therefore time to reconsider the phylum Bacteroidetes as a source for the discovery of interesting and valuable natural products to fill the future drug development pipeline.

## 9 Conflicts of interest

There are no conflicts to declare.

## 10 Acknowledgements

The authors would like to acknowledge the financial support of their original work by the Hessmann State Ministry of Higher Education, Research and the Arts (HMWK) *via* the state initiative for the development of scientific and economic excellence for the LOEWE Center for Insect Biotechnology and Bioresources, and the German Centre for Infection Research (DZIF).

## 11 References

- 1 C. J. L. Murray, K. S. Ikuta, F. Sharara, L. Swetschinski, G. Robles Aguilar, A. Gray, C. Han, C. Bisignano, P. Rao, E. Wool, S. C. Johnson, A. J. Browne, M. G. Chipeta, F. Fell, S. Hackett, G. Haines-Woodhouse, B. H. Kashef Hamadani, E. A. P. Kumaran, B. McManigal, R. Agarwal, S. Akech, S. Albertson, J. Amuasi, J. Andrews, A. Aravkin, E. Ashley, F. Bailey, S. Baker, B. Basnyat, A. Bekker, R. Bender, A. Bethou, J. Bielicki, S. Boonkasidecha, J. Bukosia, C. Carvalheiro, C. Castañeda-Orjuela, V. Chansamouth, S. Chaurasia, S. Chiurchiù, F. Chowdhury, A. J. Cook, B. Cooper, T. R. Cressey, E. Criollo-Mora, M. Cunningham, S. Darboe, N. P. J. Day, M. de Luca, K. Dokova, A. Dramowski, S. J. Dunachie, T. Eckmanns, D. Eibach, A. Emami, N. Feasey, N. Fisher-Pearson, K. Forrest, D. Garrett, P. Gastmeier, A. Z. Giref, R. C. Greer, V. Gupta, S. Haller, A. Haselbeck, S. I. Hay, M. Holm, S. Hopkins, K. C. Iregbu, J. Jacobs, D. Jarovsky, F. Javanmardi, M. Khorana, N. Kissoon, E. Kobeissi, T. Kostyanov, F. Krapp, R. Krumkamp, A. Kumar, H. H. Kyu, C. Lim, D. Limmathurotsakul, M. J. Loftus, M. Lunn, J. Ma, N. Mturi, T. Munera-Huertas, P. Musicha, M. M. Mussi-Pinhata, T. Nakamura, R. Nanavati, S. Nangia, P. Newton, C. Ngoun, A. Novotney, D. Nwakanma, C. W. Obiero, A. Olivas-Martinez, P. Olliaro, E. Ooko, E. Ortiz-Brizuela, A. Y. Peleg, C. Perrone, N. Plakkal, A. Ponce-de-Leon, M. Raad, T. Ramdin, A. Riddell, T. Roberts, J. V. Robotham, A. Roca, K. E. Rudd, N. Russell, J. Schnall, J. A. G. Scott, M. Shivamallappa, J. Sifuentes-Osornio, N. Steenkeste, A. J. Stewardson, T. Stoeva, N. Tasak, A. Thaiprakong, G. Thwaites, C. Turner, P. Turner, H. R. van Doorn, S. Velaphi, A. Vongpradith, H. Vu, T. Walsh, S. Waner, T. Wangrangsamakul, T. Wozniak, P. Zheng, B. Sartorius, A. D. Lopez, A. Stergachis, C. Moore, C. Dolecek and M. Naghavi, *Lancet*, 2022, **399**, 629–655.
- 2 A. Talebi Bezmin Abadi, A. A. Rizvanov, T. Haertlé and N. L. Blatt, *J. Bionanosci.*, 2019, **9**, 778–788.

- 3 T. F. Schäberle and I. M. Hack, *Trends Microbiol.*, 2014, **22**, 165–167.
- 4 D. J. Newman and G. M. Cragg, *J. Nat. Prod.*, 2020, **83**, 770–803.
- 5 N. Ziemert, M. Alanjary and T. Weber, *Nat. Prod. Rep.*, 2016, **33**, 988–1005.
- 6 F. Thomas, J.-H. Hehemann, E. Rebuffet, M. Czjzek and G. Michel, *Front. Microbiol.*, 2011, **2**, 93.
- 7 R. L. Hahnke, J. P. Meier-Kolthoff, M. García-López, S. Mukherjee, M. Huntemann, N. N. Ivanova, T. Woyke, N. C. Kyrpides, H.-P. Klenk and M. Göker, *Front. Microbiol.*, 2016, **7**, 2003.
- 8 M. J. McBride and Y. Zhu, *J. Bacteriol.*, 2013, **195**, 270–278.
- 9 M. J. McBride, in *The Prokaryotes. Other Major Lineages of Bacteria and The Archaea*, ed. E. F. DeLong, S. Lory, E. Stackebrandt and E. Rosenberg, Springer, Berlin, Heidelberg, 2014, pp. 643–676.
- 10 J. W. Clark and S. Kambhampati, *Mol. Phylogenet. Evol.*, 2003, **26**, 82–88.
- 11 D. P. Bayley, E. R. Rocha and C. J. Smith, *FEMS Microbiol. Lett.*, 2000, **193**, 149–154.
- 12 (a) D. Vingadassalom, A. Kolb, C. Mayer, T. Rybkine, E. Collatz and I. Podglajen, *Mol. Microbiol.*, 2005, **56**, 888–902; (b) S. Chen, M. Bagdasarian, M. G. Kaufman, A. K. Bates and E. D. Walker, *J. Bacteriol.*, 2007, **189**, 5108–5118.
- 13 C. Borsetto, G. C. A. Amos, U. N. da Rocha, A. L. Mitchell, R. D. Finn, R. F. Laidi, C. Vallin, D. A. Pearce, K. K. Newsham and E. M. H. Wellington, *Microbiome*, 2019, **7**, 78.
- 14 S. Brinkmann, M. Kurz, M. A. Patras, C. Hartwig, M. Marner, B. Leis, A. Billion, Y. Kleiner, A. Bauer, L. Toti, C. Pöverlein, P. E. Hammann, A. Vilcinskas, J. Glaeser, M. S. Spohn and T. F. Schäberle, *bioRxiv*, 2021, DOI: 10.1101/2021.07.30.454449.
- 15 (a) J. M. McGUIRE, R. L. BUNCH, R. C. ANDERSON, H. E. BOAZ, E. H. FLYNN, H. M. POWELL and J. W. SMITH, *Antibiot. Chemother.*, 1952, **2**, 281–283; (b) C. W. Pettinga, W. M. Stark and F. R. van Abeele, *J. Am. Chem. Soc.*, 1954, **76**, 569–571.
- 16 D. Jelić and R. Antolović, *Antibiotics*, 2016, **5**, 29.
- 17 K. D. Lenz, K. E. Klosterman, H. Mukundan and J. Z. Kubicek-Sutherland, *Toxins*, 2021, **13**.
- 18 G. Klaus, H. Steinmetz and G. Höfle, *European patent.*, EP2093212A1, 2008.
- 19 H. Steinmetz, K. Gerth, R. Jansen, N. Schläger, R. Dehn, S. Reinecke, A. Kirschning and R. Müller, *Angewandte Chemie (International ed. in English)*, 2011, **50**, 532–536.
- 20 R. Jansen, K. Gerth, H. Steinmetz, S. Reinecke, W. Kessler, A. Kirschning and R. Müller, *Chem. - Eur. J.*, 2011, **17**, 7739–7744.
- 21 R. Teta, M. Gurgui, E. J. N. Helfrich, S. Künne, A. Schneider, G. van Echten-Deckert, A. Mangoni and J. Piel, *Chembiochem*, 2010, **11**, 2506–2512.
- 22 R. Dehn, Y. Katsuyama, A. Weber, K. Gerth, R. Jansen, H. Steinmetz, G. Höfle, R. Müller and A. Kirschning, *Angewandte Chemie (International ed. in English)*, 2011, **50**, 3882–3887.
- 23 H. Steinmetz, W. Zander, M. A. M. Shushni, R. Jansen, K. Gerth, R. Dehn, G. Dräger, A. Kirschning and R. Müller, *Chembiochem*, 2012, **13**, 1813–1817.
- 24 S. Kumar, G. He, P. Kakarla, U. Shrestha, K. C. Ranjana, I. Ranaweera, T. M. Willmon, S. R. Barr, A. J. Hernandez and M. F. Varela, *Infect. Disord. - Drug Targets*, 2016, **16**, 28–43.
- 25 A. Beckmann, S. Hüttel, V. Schmitt, R. Müller and M. Stadler, *Microb. Cell Factories*, 2017, **16**, 143.
- 26 A. Weber, R. Dehn, N. Schläger, B. Dieter and A. Kirschning, *Org. Lett.*, 2014, **16**, 568–571.
- 27 L.-L. Wang and A. Kirschning, *Beilstein J. Org. Chem.*, 2017, **13**, 1280–1287.
- 28 L. L. Wang, D. Candito, G. Dräger, J. Herrmann, R. Müller and A. Kirschning, *Chem. - Eur. J.*, 2017, **23**, 5291–5298.
- 29 L.-L. Wang, D. Candito, G. Dräger and A. Kirschning, *Eur. J. Org. Chem.*, 2017, **2017**, 5582–5591.
- 30 Yamanouchi Pharmaceutical Co. Ltd, JP06340651, 1994.
- 31 K. Kamigiri, T. Tokunaga, T. Sugawara, K. Nagai, M. Shibasaki, B. Setiawan, R. M. Rantiatmodjo, M. Morioka and K. Suzuki, *J. Antibiot.*, 1997, **50**, 556–561.
- 32 N.-G. Jung and B.-T. Kim, *J. Korean Chem. Soc.*, 2013, **57**, 416–419.
- 33 T. Dang and R. D. Süßmuth, *Acc. Chem. Res.*, 2017, **50**, 1566–1576.
- 34 L. Taevernier, E. Wynendaele, B. Gevaert and B. de Spiegeleer, *Curr. Protein Pept. Sci.*, 2017, **18**, 425–452.
- 35 S. Sivanathan and J. Scherckenbeck, *Molecules*, 2014, **19**, 12368–12420.
- 36 D. A. Alonzo and T. M. Schmeing, *Protein Sci.*, 2020, **29**, 2316–2347.
- 37 N. Ziemert, K. Ishida, A. Liaimer, C. Hertweck and E. Dittmann, *Angewandte Chemie (International ed. in English)*, 2008, **47**, 7756–7759.
- 38 N. Mutsuo, H. Yuki, A. Osamu, K. Nahojiyu, K. Shiho and F. Daisuke, JP2005200324(A), 2005.
- 39 S. Kozuma, Y. Hirota-Takahata, D. Fukuda, N. Kuraya, M. Nakajima and O. Ando, *J. Antibiot.*, 2014, **67**, 237–242.
- 40 Y. Hirota-Takahata, S. Kozuma, N. Kuraya, D. Fukuda, M. Nakajima and O. Ando, *J. Antibiot.*, 2014, **67**, 243–251.
- 41 A. Broberg, C. Nord, J. J. Levenfors, J. Bjerketorp, B. Guss and B. Öberg, *Amino Acids*, 2021, **53**, 323–331.
- 42 J. Bjerketorp, J. J. Levenfors, C. Nord, B. Guss, B. Öberg and A. Broberg, *Front. Microbiol.*, 2021, **12**, 642829.
- 43 B. Öberg, A. Broberg, B. Guss, J. Levenfors, J. Bjerketorp and C. Nord, WO2020/046190A1, 2020.
- 44 C. Nord, J. Bjerketorp, J. J. Levenfors, S. Cao, A. A. Strömstedt, B. Guss, R. Larsson, D. Hughes, B. Öberg and A. Broberg, *ACS Chem. Biol.*, 2020, **15**, 2937–2944.
- 45 M. J. Trimble, P. Mlynářčík, M. Kolář and R. E. W. Hancock, *Cold Spring Harbor Perspect. Med.*, 2016, **6**, a025288.
- 46 C. Covas, B. Almeida, A. C. Esteves, J. Lourenço, P. Domingues, T. Caetano and S. Mendo, *New Biotechnol.*, 2021, **60**, 62–71.

- 47 J. Shoji, H. Hino, K. Matsumoto, T. Hattori, T. Yoshida, S. Matsuura and E. Kondo, *J. Antibiot.*, 1988, **41**, 713–718.
- 48 T. Kato, H. Hino, Y. Terui, J. Kikuchi and J. Shoji, *J. Antibiot.*, 1988, **41**, 719–725.
- 49 J. O'Sullivan, J. E. McCullough, A. A. Tymiak, D. R. Kirsch, W. H. Trejo and P. A. Principe, *J. Antibiot.*, 1988, **41**, 1740–1744.
- 50 D. P. Bonner, J. O'Sullivan, S. K. Tanaka, J. M. Clark and R. R. Whitney, *J. Antibiot.*, 1988, **41**, 1745–1751.
- 51 A. A. Tymiak, T. J. McCormick and S. E. Unger, *J. Org. Chem.*, 1989, **54**, 1149–1157.
- 52 H. Maki, K. Miura and Y. Yamano, *Antimicrob. Agents Chemother.*, 2001, **45**, 1823–1827.
- 53 W. Lee, K. Schaefer, Y. Qiao, V. Srisuknimit, H. Steinmetz, R. Müller, D. Kahne and S. Walker, *J. Am. Chem. Soc.*, 2016, **138**, 100–103.
- 54 S. Anlauf, M.-A. Bruening, N. A. Brunner, R. Endermann, C. Fuerstner, E. Hartmann, J. Koebberling, J. Ragot, G. Schiffer, J. Schuhmacher, N. Svenstrup, J. Telser and F. von Nussbaum, WO2004099239A1, 2004.
- 55 F. Bernhard, G. Demel, K. Soltani, H. V. Döhren and V. Blinov, *DNA Sequence*, 1996, **6**, 319–330.
- 56 J. Hou, L. Robbel and M. A. Marahiel, *Chem. Biol.*, 2011, **18**, 655–664.
- 57 (a) F. von Nussbaum, S. Anlauf, J. Benet-Buchholz, D. Häbich, J. Köbberling, L. Musza, J. Telser, H. Rübsamen-Waigmann and N. A. Brunner, *Angewandte Chemie (International ed. in English)*, 2007, **46**, 2039–2042; (b) A. Guzman-Martinez, R. Lamer and M. S. VanNieuwenhze, *J. Am. Chem. Soc.*, 2007, **129**, 6017–6021.
- 58 (a) B. J. Egner and M. Bradley, *Tetrahedron*, 1997, **53**, 14021–14030; (b) E. A. Hall, E. Kuru and M. S. VanNieuwenhze, *Org. Lett.*, 2012, **14**, 2730–2733.
- 59 K. Yasumuro, Y. Suzuki, M. Shibazaki, K. Teramura, K. Abe and M. Orita, *J. Antibiot.*, 1995, **48**, 1425–1429.
- 60 M. Orita, K. Yasumuro, K. Kokubo, M. Shimizu, K. Abe, T. Tokunaga and H. Kaniwa, *J. Antibiot.*, 1995, **48**, 1430–1434.
- 61 H. H. Wasserman, J.-H. Chen and M. Xia, *J. Am. Chem. Soc.*, 1999, **121**, 1401–1402.
- 62 K. Teramura, K. Yasumuro and K. Abe, *J. Enzym. Inhib.*, 1996, **11**, 33–38.
- 63 V. Apostolopoulos, J. Bojarska, T.-T. Chai, S. Elnagdy, K. Kaczmarek, J. Matsoukas, R. New, K. Parang, O. P. Lopez, H. Parhiz, C. O. Perera, M. Pickholz, M. Remko, M. Saviano, M. Skwarczynski, Y. Tang, W. M. Wolf, T. Yoshiya, J. Zabrocki, P. Zielenkiewicz, M. AlKhazindar, V. Barriga, K. Kelaidonis, E. M. Sarasia and I. Toth, *Molecules*, 2021, **26**, 430.
- 64 A. Henninot, J. C. Collins and J. M. Nuss, *J. Med. Chem.*, 2018, **61**, 1382–1414.
- 65 Y. A. Haggag, *BJSTR*, 2018, **8**, 6659–6662.
- 66 H. Ono, Y. Funabashi, S. Harada, *European Pat.*, EP0339596, 1989.
- 67 N. Katayama, S. Fukusumi, Y. Funabashi, T. Iwahi and H. Ono, *J. Antibiot.*, 1993, **46**, 606–613.
- 68 Y. Funabashi, S. Tsubotani, K. Koyama, N. Katayama and S. Harada, *Tetrahedron*, 1993, **49**, 13–28.
- 69 (a) C. Yuan and R. M. Williams, *J. Am. Chem. Soc.*, 1997, **119**, 11777–11784; (b) R. M. Williams, C. Yuan, V. J. Lee and S. Chamberland, *J. Antibiot.*, 1998, **51**, 189–201; (c) V. V. Sokolov, S. I. Kozhushkov, S. Nikolskaya, V. N. Belov, M. Es-Sayed and A. de Meijere, *Eur. J. Org. Chem.*, 1998, **1998**, 777–783; (d) P. Lin and A. Ganesan, *Synthesis*, 2000, **2000**, 2127–2130; (e) N. Aguilar and J. Krüger, *Molecules*, 2002, **7**, 469–474; (f) M. Brands, R. Endermann, R. Gahlmann, J. Krüger and S. Raddatz, *Bioorg. Med. Chem. Lett.*, 2003, **13**, 241–245; (g) M. Brands, Y. C. Grande, R. Endermann, R. Gahlmann, J. Krüger and S. Raddatz, *Bioorg. Med. Chem. Lett.*, 2003, **13**, 2641–2645; (h) V. N. Belov, V. V. Sokolov, B. D. Zlatopolskiy and A. de Meijere, *Eur. J. Org. Chem.*, 2011, **2011**, 4093–4097.
- 70 M. Brands, R. Endermann, R. Gahlmann, J. Krüger, S. Raddatz, J. Stoltefuss, V. N. Belov, S. Nizamov, V. V. Sokolov and A. de Meijere, *J. Med. Chem.*, 2002, **45**, 4246–4253.
- 71 N. Böddeker, G. Bahador, C. Gibbs, E. Mabery, J. Wolf, L. Xu and J. Watson, *RNA*, 2002, **8**, 1120–1128.
- 72 W. S. Champney, J. Pelt and C. L. Tober, *Curr. Microbiol.*, 2001, **43**, 340–345.
- 73 E. Limburg, R. Gahlmann, H.-P. Kroll and D. Beyer, *Antimicrob. Agents Chemother.*, 2004, **48**, 619–622.
- 74 B. Somanadhan, S. R. Kotturi, C. Yan Leong, R. P. Glover, Y. Huang, H. Flotow, A. D. Buss, M. J. Lear and M. S. Butler, *J. Antibiot.*, 2013, **66**, 259–264.
- 75 S. R. Kotturi, B. Somanadhan, J.-H. Chng, K. S.-W. Tan, M. S. Butler and M. J. Lear, *Tetrahedron Lett.*, 2014, **55**, 1949–1951.
- 76 S. Brinkmann, S. Semmler, C. Kersten, M. A. Patras, M. Kurz, N. Fuchs, S. J. Hammerschmidt, J. Legac, P. E. Hammann, A. Vilcinskas, P. J. Rosenthal, T. Schirmeister, A. Bauer and T. F. Schäberle, *ACS Chem. Biol.*, 2022, DOI: 10.1021/acscembio.1c00861.
- 77 M. Montalbán-López, T. A. Scott, S. Ramesh, I. R. Rahman, A. J. van Heel, J. H. Viel, V. Bandarian, E. Dittmann, O. Genilloud, Y. Goto, M. J. Grande Burgos, C. Hill, S. Kim, J. Koehnke, J. A. Latham, A. J. Link, B. Martínez, S. K. Nair, Y. Nicolet, S. Rebuffat, H.-G. Sahl, D. Sareen, E. W. Schmidt, L. Schmitt, K. Severinov, R. D. Süßmuth, A. W. Truman, H. Wang, J.-K. Weng, G. P. van Wezel, Q. Zhang, J. Zhong, J. Piel, D. A. Mitchell, O. P. Kuipers and W. A. van der Donk, *Nat. Prod. Rep.*, 2021, **38**, 130–239.
- 78 T. Caetano, W. van der Donk and S. Mendo, *Microbiol. Res.*, 2020, **235**, 126441.
- 79 J. Delves-Broughton, P. Blackburn, R. J. Evans and J. Hugenholtz, *Antonie Leeuwenhoek*, 1996, **69**, 193–202.
- 80 K. I. Mohr, C. Volz, R. Jansen, V. Wray, J. Hoffmann, S. Bernecker, J. Wink, K. Gerth, M. Stadler and R. Müller, *Angewandte Chemie (International ed. in English)*, 2015, **54**, 11254–11258.
- 81 I. R. Bothwell, T. Caetano, R. Sarkisian, S. Mendo and W. A. van der Donk, *ACS Chem. Biol.*, 2021, **16**, 1019–1029.

- 82 T. Fujita, H. Hatanaka, K. Hayashi, N. Shigematsu, S. Takase, M. Okamoto, M. Okuhara, K. Shimatani and A. Satoh, *J. Antibiot.*, 1994, **47**, 1359–1364.
- 83 H. Hatanaka, S. Takase, T. Fujita, M. Okamoto and M. Okuhara, WO9302203A1, 1992.
- 84 T. Kinoshita, T. Kitatani, M. Warizaya and T. Tada, *Acta Crystallogr. F*, 2005, **61**, 808–811.
- 85 T. Fujita, Y. Shinguh, A. Yamazaki, K. Nakahara, M. Okamoto and M. Okuhara, *J. Antibiot.*, 1994, **47**, 1365–1368.
- 86 A. Miyanaga, F. Kudo and T. Eguchi, *Nat. Prod. Rep.*, 2018, **35**, 1185–1209.
- 87 Y. J. Yoo, H. Kim, S. R. Park and Y. J. Yoon, *J. Ind. Microbiol. Biotechnol.*, 2017, **44**, 537–553.
- 88 N. Oku, K. Adachi, S. Matsuda, H. Kasai, A. Takatsuki and Y. Shizuri, *Org. Lett.*, 2008, **10**, 2481–2484.
- 89 L. Linares-Otoya, V. Linares-Otoya, L. Armas-Mantilla, C. Blanco-Olano, M. Crüsemann, M. L. Ganoza-Yupanqui, J. Campos-Florian, G. M. König and T. F. Schäberle, *Mar. Drugs*, 2017, **15**, 308.
- 90 K. H. van Pée and J. M. Ligon, *Nat. Prod. Rep.*, 2000, **17**, 157–164.
- 91 J. R. Evans, E. J. Napier and R. A. Fletton, *J. Antibiot.*, 1978, **31**, 952–958.
- 92 (a) P. Srisukchayakul, C. Suwanachart, Y. Sangnoi, A. Kanjana-opas, S. Hosoya, A. Yokota and V. Arunpairojana, *Int. J. Syst. Evol. Microbiol.*, 2007, **57**, 2275–2279; (b) A. Kanjana-opas, S. Panphon, H.-K. Fun and S. Chantrapromma, *Acta Crystallogr. E*, 2006, **62**, o2728–o2730.
- 93 Y. Sangnoi, O. Sakulkeo, S. Yuenyongsawad, A. Kanjana-opas, K. Ingkaninan, A. Plubrukarn and K. Suwanborirux, *Mar. Drugs*, 2008, **6**, 578–586.
- 94 P. W. Okanya, K. I. Mohr, K. Gerth, R. Jansen and R. Müller, *J. Nat. Prod.*, 2011, **74**, 603–608.
- 95 E. J. Choi, S.-J. Nam, L. Paul, D. Beatty, C. A. Kauffman, P. R. Jensen and W. Fenical, *Chem. Biol.*, 2015, **22**, 1270–1279.
- 96 L. Linares-Otoya, Y. Liu, V. Linares-Otoya, L. Armas-Mantilla, M. Crüsemann, M. L. Ganoza-Yupanqui, J. Campos-Florian, G. M. König and T. F. Schäberle, *ACS Chem. Biol.*, 2019, **14**, 176–181.
- 97 Y. Sangnoi, O. Sakulkeo, S. Yuenyongsawad, A. Kanjana-opas, K. Ingkaninan, A. Plubrukarn and K. Suwanborirux, *Mar. Drugs*, 2008, **6**, 578–586.
- 98 (a) X. Ma, Y. Vo, M. G. Banwell and A. C. Willis, *Asian J. Org. Chem.*, 2012, **1**, 160–165; (b) J. P. Mahajan, Y. R. Suryawanshi and S. B. Mhaske, *Org. Lett.*, 2012, **14**, 5804–5807; (c) L. Ni, Z. Li, F. Wu, J. Xu, X. Wu, L. Kong and H. Yao, *Tetrahedron Lett.*, 2012, **53**, 1271–1274; (d) B. Patel and S. Hilton, *Synlett*, 2014, **26**, 79–83; (e) B. Bolte, C. S. Bryan, P. P. Sharp, S. Sayyahi, C. Rihouey, A. Kendrick, P. Lan, M. G. Banwell, C. J. Jackson, N. J. Fraser, A. C. Willis and J. S. Ward, *J. Org. Chem.*, 2020, **85**, 650–663.
- 99 M. Osano, D. P. Jhaveri and P. Wipf, *Org. Lett.*, 2020, **22**, 2215–2219.
- 100 C. S. Schwalm and C. R. D. Correia, *Tetrahedron Lett.*, 2012, **53**, 4836–4840.
- 101 Y. Yamaoka, T. Yoshida, M. Shinozaki, K.-i. Yamada and K. Takasu, *J. Org. Chem.*, 2015, **80**, 957–964.
- 102 F. Khan, M. Dlugosch, X. Liu and M. G. Banwell, *Acc. Chem. Res.*, 2018, **51**, 1784–1795.
- 103 W. L. Parker, J. O'Sullivan and R. B. Sykes, in *Advances in Applied Microbiology*, Elsevier, 1986, vol. 31, pp. 181–205.
- 104 P. D. Singh, J. H. Johnson, P. C. Ward, J. S. Wells, W. H. Trejo and R. B. Sykes, *J. Antibiot.*, 1983, **36**, 1245–1251.
- 105 R. Cooper, K. Bush, P. A. Principe, W. H. Trejo, J. S. Wells and R. B. Sykes, *J. Antibiot.*, 1983, **36**, 1252–1257.
- 106 T. Kato, H. Hino, J. Shoji, K. Matsumoto, T. Tanimoto, T. Hattori, K. Hirooka and E. Kondo, *J. Antibiot.*, 1987, **40**, 135–138.
- 107 T. Kato, H. Hino, Y. Terui, J. Nishikawa, Y. Nakagawa, Y. Ikenishi and J. Shoji, *J. Antibiot.*, 1987, **40**, 139–144.
- 108 N. Katayama, Y. Nozaki, K. Okonogi, H. Ono, S. Harada and H. Okazaki, *J. Antibiot.*, 1985, **38**, 1117–1127.
- 109 T. Hida, S. Tsubotani, N. Katayama, H. Okazaki and S. Harada, *J. Antibiot.*, 1985, **38**, 1128–1140.
- 110 C. E. Higgins, R. L. Hamill, T. H. Sands, M. M. Hoehn and N. E. Davis, *J. Antibiot.*, 1974, **27**, 298–300.
- 111 P. D. Singh, P. C. Ward, J. S. Wells, C. M. Ricca, W. H. Trejo, P. A. Principe and R. B. Sykes, *J. Antibiot.*, 1982, **35**, 1397–1399.
- 112 P. D. Singh, M. G. Young, J. H. Johnson, C. M. Cimarusti and R. B. Sykes, *J. Antibiot.*, 1984, **37**, 773–780.
- 113 J. Shoji, T. Kato, R. Sakazaki, W. Nagata, Y. Terui, Y. Nakagawa, M. Shiro, K. Matsumoto, T. Hattori and T. Yoshida, *J. Antibiot.*, 1984, **37**, 1486–1490.
- 114 J. Shoji, R. Sakazaki, T. Kato, Y. Terui, K. Matsumoto, T. Tanimoto, T. Hattori, K. Hirooka and E. Kondo, *J. Antibiot.*, 1985, **38**, 538–540.
- 115 C. A. Aruldass, L. Dufossé and W. A. Ahmad, *J. Clean. Prod.*, 2018, **180**, 168–182.
- 116 (a) H. Achenbach, *Arch. Microbiol.*, 1974, **101**, 131–144; (b) H. Achenbach, W. Kohl, H. Reichenbach and H. Kleinig, *Tetrahedron Lett.*, 1974, **15**, 2555–2556.
- 117 H. Achenbach and J. Witzke, *Angew. Chem., Int. Ed. Engl.*, 1977, **16**, 191–192.
- 118 (a) H. Achenbach, A. Böttger, W. Kohl, E. Fautz and H. Reichenbach, *Phytochemistry*, 1979, **18**, 961–963; (b) E. Fautz and H. Reichenbach, *Phytochemistry*, 1979, **18**, 957–959; (c) H. Achenbach, A. Böttger-Vetter, E. Fautz and H. Reichenbach, *Arch. Microbiol.*, 1982, **132**, 241–244; (d) M. J. McBride, G. Xie, E. C. Martens, A. Lapidus, B. Henrissat, R. G. Rhodes, E. Goltsman, W. Wang, J. Xu, D. W. Hunnicutt, A. M. Staroscik, T. R. Hoover, Y.-Q. Cheng and J. L. Stein, *Appl. Environ. Microbiol.*, 2009, **75**, 6864–6875; (e) S. W. Fuchs, K. A. J. Bozhüyük, D. Kresovic, F. Grundmann, V. Dill, A. O. Brachmann, N. R. Waterfield and H. B. Bode, *Angew. Chem., Int. Ed. Engl.*, 2013, **52**, 4108–4112; (f) T. A. Schöner, S. W. Fuchs, C. Schönau and H. B. Bode, *J. Microb. Biotechnol.*, 2014, **7**, 232–241.



- 119 (a) M. E. P. Jiménez, C. M. B. Pinilla, E. Rodrigues and A. Brandelli, *Nat. Prod. Res.*, 2019, **33**, 1541–1549; (b) A. Mogadem, M. A. Almamary, N. A. Mahat, K. Jemon, W. A. Ahmad and I. Ali, *Molecules*, 2021, **26**.
- 120 (a) D. L. Fox and R. A. Lewin, *Can. J. Microbiol.*, 1963, **9**, 753–768; (b) A. J. Aasen and S. L. Jensen, *Acta Chem. Scand.*, 1966, **20**, 811–819; (c) A. J. Aasen and S. L. Jensen, *Acta Chem. Scand.*, 1966, **20**, 1970–1988; (d) A. J. Aasen and S. L. Jensen, *Acta Chem. Scand.*, 1966, **20**, 2322–2324; (e) A. J. Aasen, S. Liaaen-Jensen, G. Borch, A. F. Andresen, W. B. Pearson and V. Meisalo, *Acta Chem. Scand.*, 1972, **26**, 404–405; (f) K. Shindo, K. Kikuta, A. Suzuki, A. Katsuta, H. Kasai, M. Yasumoto-Hirose, Y. Matsuo, N. Misawa and S. Takaichi, *Appl. Microbiol. Biotechnol.*, 2007, **74**, 1350–1357; (g) K. Shindo and N. Misawa, *Mar. Drugs*, 2014, **12**, 1690–1698.
- 121 L. Tao, H. Yao, H. Kasai, N. Misawa and Q. Cheng, *Mol. Genet. Genomics*, 2006, **276**, 79–86.
- 122 J.-H. Jang, K. Kanoh, K. Adachi, S. Matsuda and Y. Shizuri, *J. Nat. Prod.*, 2007, **70**, 563–566.
- 123 M. J. Fujita, K. Nakano and R. Sakai, *Molecules*, 2013, **18**, 3917–3926.
- 124 Z.-J. Wang, H. Zhou, G. Zhong, L. Huo, Y.-J. Tang, Y. Zhang and X. Bian, *Org. Lett.*, 2020, **22**, 939–943.
- 125 M. J. Fujita, Y. Goto and R. Sakai, *Mar. Drugs*, 2018, **16**, 342.
- 126 (a) T. Kamiyama, T. Umino, Y. Iteazono, Y. Nakamura, T. Satoh and K. Yokose, *J. Antibiot.*, 1995, **48**, 929–936; (b) T. Kamiyama, T. Umino, T. Satoh, S. Sawairi, M. Shirane, S. Ohshima and K. Yokose, *J. Antibiot.*, 1995, **48**, 924–928.
- 127 J. Maeda, M. Nishida, H. Takikawa, H. Yoshida, T. Azuma, M. Yoshida and Y. Mizushima, *Int. J. Mol. Med.*, 2010, **26**, 751–758.
- 128 P. N. Chaudhari, K. S. Wani, B. L. Chaudhari and S. B. Chincholkar, *Appl. Biochem. Biotechnol.*, 2009, **158**, 231–241.
- 129 J. 'i. Kohayashi, S. Mikami, H. Shigemori, T. Takao, Y. Shimonishi, S. Izuta and S. Yoshida, *Tetrahedron*, 1995, **51**, 10487–10490.
- 130 (a) W. Godchaux and E. R. Leadbetter, *J. Bacteriol.*, 1980, **144**, 592–602; (b) D. R. Abbanat, W. Godchaux and E. R. Leadbetter, *Arch. Microbiol.*, 1988, **149**, 358–364.
- 131 (a) D. R. Abbanat, E. R. Leadbetter, W. Godchaux and A. Escher, *Nature*, 1986, **324**, 367–369; (b) W. Godchaux and E. R. Leadbetter, *J. Bacteriol.*, 1983, **153**, 1238–1246; (c) W. Godchaux and E. R. Leadbetter, *J. Biol. Chem.*, 1984, **259**, 2982–2990.
- 132 (a) R. A. Alegado, L. W. Brown, S. Cao, R. K. Dermenjian, R. Zuzow, S. R. Fairelough, J. Clardy and N. King, *Elife*, 2012, **1**, e00013; (b) A. M. Cantley, A. Woznica, C. Beemelmans, N. King and J. Clardy, *J. Am. Chem. Soc.*, 2016, **138**, 4326–4329; (c) A. Woznica, A. M. Cantley, C. Beemelmans, E. Freinkman, J. Clardy and N. King, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 7894–7899; (d) C. Beemelmans, A. Woznica, R. A. Alegado, A. M. Cantley, N. King and J. Clardy, *J. Am. Chem. Soc.*, 2014, **136**, 10210–10213.
- 133 (a) N. Irako and T. Shioiri, *Tetrahedron Lett.*, 1998, **39**, 5793–5796; (b) N. Irako and T. Shioiri, *Tetrahedron Lett.*, 1998, **39**, 5797–5798; (c) H. Takikawa, S.-e. Muto, D. Nozawa, A. Kayo and K. Mori, *Tetrahedron Lett.*, 1998, **39**, 6931–6934; (d) H. Takikawa, D. Nozawa, A. Kayo, S.-e. Muto and K. Mori, *J. Chem. Soc., Perkin Trans. 1*, 1999, **1**, 2467–2477; (e) T. Shioiri and N. Irako, *Tetrahedron*, 2000, **56**, 9129–9142; (f) O. Labeeuw, P. Phansavath and J.-P. Genêt, *Tetrahedron Lett.*, 2003, **44**, 6383–6386; (g) O. Labeeuw, P. Phansavath and J.-P. Genêt, *Tetrahedron: Asymmetry*, 2004, **15**, 1899–1908; (h) P. Gupta, S. V. Naidu and P. Kumar, *Tetrahedron Lett.*, 2004, **45**, 9641–9643; (i) A. Sharma, S. Gamre and S. Chattopadhyay, *Tetrahedron Lett.*, 2007, **48**, 3705–3707.
- 134 (a) D. R. Abbanat, W. Godchaux, G. Polychroniou and E. R. Leadbetter, *Biochem. Biophys. Res. Commun.*, 1985, **130**, 873–878; (b) R. H. White, *J. Bacteriol.*, 1984, **159**, 42–46; (c) M. Á. Vences-Guzmán, R. Peña-Miller, N. A. Hidalgo-Aguilar, M. L. Vences-Guzmán, Z. Guan and C. Sohlenkamp, *Environ. Microbiol.*, 2021, **23**, 2448–2460.
- 135 (a) I. Uchida, K. Yoshida, Y. Kawai, S. Takase, Y. Itoh, H. Tanaka, M. Kohsaka and H. Imanaka, *J. Antibiot.*, 1985, **38**, 1476–1486; (b) K. Yoshida, M. Iwami, Y. Umehara, M. Nishikawa, I. Uchida, M. Kohsaka, H. Aoki and H. Imanaka, *J. Antibiot.*, 1985, **38**, 1469–1475.
- 136 K. Suzuki, H. Yamaguchi, S. Miyazaki, K. Nagai, S. Watanabe, T. Saito, K. Ishii, M. Hanada, T. Sekine and Y. Ikegami, *J. Antibiot.*, 1990, **43**, 154–157.
- 137 Y. Ikegami, N. Takeuchi, M. Hanada, Y. Hasegawa, K. Ishii, T. Andoh, T. Sato, K. Suzuki, H. Yamaguchi and S. Miyazaki, *J. Antibiot.*, 1990, **43**, 158–162.
- 138 (a) T. Nemoto, M. Ojika, Y. Takahata, T. Andoh and Y. Sakagami, *Tetrahedron*, 1998, **54**, 2683–2690; (b) T. Shioiri, Y. Terao, N. Irako and T. Aoyama, *Tetrahedron*, 1998, **54**, 15701–15710.
- 139 M. Shiozaki, N. Deguchi, T. Mochizuki, T. Wakabayashi, T. Ishikawa, H. Haruyama, Y. Kawai and M. Nishijima, *Tetrahedron*, 1998, **54**, 11861–11876.
- 140 Y. Kawai, I. Yano and K. Kaneda, *Eur. J. Biochem.*, 1988, **171**, 73–80.
- 141 R. B. Clark, J. L. Cervantes, M. W. Maciejewski, V. Farrokhi, R. Nemati, X. Yao, E. Anstadt, M. Fujiwara, K. T. Wright, C. Riddle, C. J. La Vake, J. C. Salazar, S. Finegold and F. C. Nichols, *Infect. Immun.*, 2013, **81**, 3479–3489.
- 142 Y.-H. Wang, R. Nemati, E. Anstadt, Y. Liu, Y. Son, Q. Zhu, X. Yao, R. B. Clark, D. W. Rowe and F. C. Nichols, *Bone*, 2015, **81**, 654–661.
- 143 R. Nemati, C. Dietz, E. J. Anstadt, J. Cervantes, Y. Liu, F. E. Dewhirst, R. B. Clark, S. Finegold, J. J. Gallagher, M. B. Smith, X. Yao and F. C. Nichols, *J. Lipid Res.*, 2017, **58**, 1999–2007.
- 144 Y. Kawai and K. Akagawa, *Infect. Immun.*, 1989, **57**, 2086–2091.
- 145 O. Teng, C. K. E. Ang and X. L. Guan, *Front. Immunol.*, 2017, **8**, 1836.
- 146 I. Olsen and F. C. Nichols, *Infect. Immun.*, 2018, **86**, e00035-18.

## Review

## Natural Product Reports

- 147 T. Morishita, A. Sato, M. Hisamoto, T. Oda, K. Matsuda, A. Ishii and K. Kodama, *J. Antibiot.*, 1997, **50**, 457–468.
- 148 (a) C. S. Mirucki, M. Abedi, J. Jiang, Q. Zhu, Y.-H. Wang, K. E. Safavi, R. B. Clark and F. C. Nichols, *J. Endod.*, 2014, **40**, 1342–1348; (b) F. C. Nichols, W. J. Housley, C. A. O'Connor, T. Manning, S. Wu and R. B. Clark, *Am. J. Surg. Pathol.*, 2009, **175**, 2430–2438; (c) K. Gomi, K. Kawasaki, Y. Kawai, M. Shiozaki and M. Nishijima, *J. Immunol.*, 2002, **168**, 2939–2943.
- 149 V. Farrokhi, R. Nemat, F. C. Nichols, X. Yao, E. Anstadt, M. Fujiwara, J. Grady, D. Wakefield, W. Castro, J. Donaldson and R. B. Clark, *Clin. Transl. Immunol.*, 2013, **2**, e8.
- 150 (a) Y. K.-H. Schneider, K. Ø Hansen, J. Isaksson, S. Ullsten, E. H Hansen and J. Hammer Andersen, *Molecules*, 2019, **24**, 3991; (b) M.-K. Bill, S. Brinkmann, M. Oberpaul, M. A. Patras, B. Leis, M. Marner, M.-P. Maitre, P. E. Hammann, A. Vilcinskas, S. M. M. Schuler and T. F. Schäberle, *Molecules*, 2021, **26**, 5195.
- 151 M. Shaaban, R. P. Maskey, I. Wagner-Döbler and H. Laatsch, *J. Nat. Prod.*, 2002, **65**, 1660–1663.
- 152 S. Imai, K. Fujioka, K. Furihata, R. Fudo, S. Yamanaka and H. Seto, *J. Antibiot.*, 1993, **46**, 1319–1322.
- 153 A. Spyere, D. C. Rowley, P. R. Jensen and W. Fenical, *J. Nat. Prod.*, 2003, **66**, 818–822.
- 154 (a) Y. Matsuo, H. Imagawa, M. Nishizawa and Y. Shizuri, *Science*, 2005, **307**, 1598; (b) H. Yamamoto, Y. Takagi, N. Yamasaki, T. Mitsuyama, Y. Kasai, H. Imagawa, Y. Kinoshita, N. Oka and M. Hiraoka, *Tetrahedron*, 2018, **74**, 7173–7178.
- 155 (a) K. M. Papp-Wallace and R. A. Bonomo, *Infect. Dis. Clin.*, 2016, **30**, 441–464; (b) S. Andrei, G. Droc and G. Stefan, *Discoveries*, 2019, **7**, e102.
- 156 K. Katoh and D. M. Standley, *Mol. Biol. Evol.*, 2013, **30**, 772–780.

## 2. Aim of this work

The aim of this work is to develop and establish a microbial cultivation approach based on agarose-solidified droplet-microfluidics. This technology will then be implemented into a cultivation and screening platform used to search for novel anti-infective NPs. Therefore, starting from diverse environmental samples, axenic cultures should be retrieved and domesticated for *in vitro* growth, enabling classical bioactivity screening of the resulting extracts.

Building on this microfluidics method, it should be developed into a more complex two compartment system to potentially allow new high-throughput screening approaches in the future.

Furthermore, is aimed to investigate the potential of the microbial phylum Bacteroidetes to biosynthesize valuable anti-infective NPs.

## CHAPTER 1 – MICROFLUIDICS

---

Nowadays the cultivability of microorganisms at laboratory conditions is indispensable for the discovery of NPs that serve as future medical lead structures. Hereby, the taxonomical broad and diverse accessibility to this valuable resource is of crucial importance. Microfluidic-based approaches have proven to be applicable to culture microorganisms including previously uncultured species.

In frame of this work, a microbial cultivation approach based on droplet-microfluidics should be established and implemented into a cultivation and screening platform used to search for novel anti-infective NPs. The first goal was to retrieve and screen a high taxonomical diversity of axenic microbial cultures from environmental samples, including representatives of species not yet explored. The following publication – *“Combination of high-throughput microfluidics and FACS technologies to leverage the numbers game in natural product discovery”* – depicts the potential and feasibility of the agarose-solidified droplet-microfluidic cultivation approach established herein. In addition to this miniaturized cultivation approach, which is based on simple water-in-oil (w/o) emulsions, the second goal was to further develop the approach towards a more complex two compartment system based on agarose-solidified double emulsions (w/w/o). The scaled-down, high-throughput screening of environmental microorganisms for anti-infective NPs at this miniaturized droplet-scale could be a possible application thereof. The second microfluidics publication – *“Two-step generation of monodisperse agarose-solidified double emulsions (w/w/o) excluding an inner oil barrier”* – represents the developed microfluidics method – the first of its kind – to enable the generation of agarose-solidified double emulsions (w/w/o).

## Publication I

# microbial biotechnology

Open Access

## Combination of high-throughput microfluidics and FACS technologies to leverage the numbers game in natural product discovery

Markus Oberpaul,<sup>1,†</sup>  Stephan Brinkmann,<sup>1,†</sup>   
Michael Marner,<sup>1</sup>  Sanja Mihajlovic,<sup>1</sup>   
Benedikt Leis,<sup>1</sup> Maria A. Patras,<sup>1</sup>   
Christoph Hartwig,<sup>1</sup>  Andreas Vilcinskas,<sup>1,2</sup>   
Peter E. Hammann,<sup>3</sup> Till F. Schäberle<sup>1,2,4,\*</sup>   
Marius Spohn<sup>1,\*\*</sup>  and Jens Glaeser<sup>3,\*\*\*</sup>

<sup>1</sup>Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Branch for Bioresources, Giessen, 35392, Germany.

<sup>2</sup>Institute for Insect Biotechnology, Justus-Liebig-University-Giessen, Giessen, 35392, Germany.

<sup>3</sup>Evotec International GmbH, Göttingen, 37079, Germany.

<sup>4</sup>German Center for Infection Research (DZIF), Partner Site Giessen-Marburg-Langen, Giessen, 35392, Germany.

**Copyright:** This article is Open Access (CC BY 4.0).

**Supporting Information** (not included in this dissertation file):

<https://sfamjournals.onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2F1751-7915.13872&file=mbt213872-sup-0001-supinfo.docx>

### **Contribution:**

Markus Oberpaul and Stephan Brinkmann conceived, designed and performed the microfluidics and FACS workflow to retrieve axenic cultures from environmental samples, including the upstream process of cell sorting, cultivation and cryopreservation for long-term storage. Both of them established and validated the microbial cultivation experiments at droplet-scale, and have been involved in, as well as conceived all other experiments, analyzed the data and prepared all the figures for the manuscript. Markus Oberpaul, Stephan Brinkmann and Marius S. Spohn wrote the draft of the manuscript as well as the following revisions.

# microbial biotechnology

Open Access

## Combination of high-throughput microfluidics and FACS technologies to leverage the numbers game in natural product discovery

Markus Oberpaul,<sup>1,†</sup>  Stephan Brinkmann,<sup>1,†</sup>   
 Michael Marner,<sup>1</sup>  Sanja Mihajlovic,<sup>1</sup>   
 Benedikt Leis,<sup>1</sup>  Maria A. Patras,<sup>1</sup>   
 Christoph Hartwig,<sup>1</sup>  Andreas Vilcinskas,<sup>1,2</sup>   
 Peter E. Hammann,<sup>3</sup>  Till F. Schäberle<sup>1,2,4,\*</sup>   
 Marius Spohn<sup>1,\*\*</sup>  and Jens Glaeser<sup>3,\*\*\*</sup>

<sup>1</sup>Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Branch for Bioresources, Giessen, 35392, Germany.

<sup>2</sup>Institute for Insect Biotechnology, Justus-Liebig-University-Giessen, Giessen, 35392, Germany.

<sup>3</sup>Evotec International GmbH, Göttingen, 37079, Germany.

<sup>4</sup>German Center for Infection Research (DZIF), Partner Site Giessen-Marburg-Langen, Giessen, 35392, Germany.

### Summary

High-throughput platforms facilitating screening campaigns of environmental samples are needed to discover new products of natural origin counteracting the spreading of antimicrobial resistances constantly threatening human and agricultural health. We applied a combination of droplet microfluidics and fluorescence-activated cell sorting (FACS)-based technologies to access and assess a microbial environmental sample. The cultivation performance of our microfluidics workflow was evaluated in respect

to the utilized cultivation media by Illumina amplicon sequencing of a pool of millions of droplets, respectively. This enabled the rational selection of a growth medium supporting the isolation of microbial diversity from soil (five phyla affiliated to 57 genera) including a member of the acidobacterial subgroup 1 (genus *Edaphobacter*). In a second phase, the entire diversity covered by 1071 cultures was used for an arrayed bioprospecting campaign, resulting in > 6000 extracts tested against human pathogens and agricultural pests. After redundancy curation by using a combinatorial chemical and genomic fingerprinting approach, we assigned the causative agents present in the extracts. Utilizing UHPLC-QTOF-MS/MS-guided fractionation and microplate-based screening assays in combination with molecular networking the production of bioactive ionophorous macrotetrolides, phospholipids, the cyclic lipopeptides massetolides E, F, H and serratamolide A and many derivatives thereof was shown.

### Introduction

Continuously increasing levels of drug resistance and the consequent loss of existing drug and control agents for the treatment of infections poses an enormous threat to human health care systems and plant disease management strategies (Lewis, 2013; Lakemeyer *et al.*, 2018). This crisis has to be counteracted by continued innovation in discovery campaigns to increase the probability of finding new anti-infective lead structures (Schäberle and Hack, 2014; Tacconelli *et al.*, 2018; Theuretzbacher *et al.*, 2020).

A rich source for new chemical entities are microbial-derived natural products (NPs), which always has been a major inspiration for development of drugs and control agents (Newman and Cragg, 2020). NP evolution brought forth a vast diversity of unique molecules that are optimized for interactions with their respective molecular targets; the latter a plethora of biological macromolecules themselves (Firm and Jones, 2003; Bon and Waldmann, 2010; Hong, 2011). Fortunately, the predicted potential for chemical variety encoded within the microbial diversity is yet only scarcely exploited. This is

Received 2 May, 2021; revised 17 May, 2021; accepted 6 June, 2021.

For correspondence. \*E-mail Till.F.Schaeberle@agrar.uni-giessen.de; Tel. +49 641 972-19140; Fax +49 641 972-19499. \*\*E-mail marius.spohn@ime.fraunhofer.de; Tel. +49 641 972-19189; Fax +49 641 972-19499. \*\*\*E-mail Jens.Glaeser@evotec.com; Tel. +49 (0)151 25565835; Fax +49 641 972-19499.

<sup>†</sup>These authors contributed equally to this work.

*Microbial Biotechnology* (2022) 15(2), 415–430  
 doi:10.1111/1751-7915.13872

### Funding information

This work was financially supported by the Hessen State Ministry of Higher Education, Research and the Arts (HMWK) via the state initiative for the development of scientific and economic excellence for the LOEWE Center for Insect Biotechnology and Bioresources. Sanofi-Aventis Deutschland GmbH and Evotec International GmbH funded this work in the framework of the Sanofi-Fraunhofer Natural Products Center and its follow-up, the Fraunhofer-Evotec Natural Products Center of Excellence.

© 2021 The Authors. *Microbial Biotechnology* published by Society for Applied Microbiology and John Wiley & Sons Ltd. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

416 M. Oberpaul *et al.*

mainly due to the fact that major NP discovery efforts of past decades had been focusing on rather easily cultivable microorganisms and consequently on a limited phylogenetic diversity (Monciardini *et al.*, 2014). Along with the likely also not yet-exhausted potential of classical NP producers, for example *Streptomyces* spp. (Adamek *et al.*, 2019; Belknap *et al.*, 2020), additional phylogenetic branches of the bacterial kingdom represent a promising source for discovery of novelty (Gross and Loper, 2009; Panthee *et al.*, 2016; Tracanna *et al.*, 2017). Thus, driven by the dogma that phylogenetic and genomic divergence translates directly into chemical diversity (Medema *et al.*, 2014; Monciardini *et al.*, 2014), greater effort is demanded to bring a broader diversity of genera and families in culture (Hoffmann *et al.*, 2018; Nicault *et al.*, 2020) while directly exploiting their metabolic capabilities. The still valid functionality of the methodology to valorize bacterial derived NPs was recently shown by the development and market introduction of fenpicoxamid, derived from the antifungal compound UK-2A produced by *Streptomyces* sp. 517-02. This NP provides a new target site for the control of *Zyoseptoria tritici* (SEPTTR), the causative agent of Septoria tritici Blotch (STB) (Butler and Paterson, 2020) by inhibiting the mitochondrial complex III at a target site distinguished from the one attacked by the strobilurin class (Owen *et al.*, 2017). The trend towards the valorization and application of naturally derived fungicides continues to rise, predicted to obtain further attention in the next decades (Umetsu and Shirai, 2020). In antibacterial research, success in discovery of novel NPs with unique modes of action was recently mainly achieved with the isolation of novel anti-infective lead structures from rare Proteobacteria, particularly darobactin (from *Photorhabdus* sp.), a heptapeptide antibiotic targeting the gram-negative outer membrane protein *BamA* (Imai *et al.*, 2019) and teixobactin (from *Eleftheria* sp.), which was made accessible by a novel cultivation technique (Ling *et al.*, 2015).

The basis for all such findings is the fundamental accessibility to a broad, thus diverse phylogenetic space of the bacterial kingdom. The challenging task to bring an additional layer of today's microbial dark matter into culture is mainly a numbers game that has to be approached by the development and implementation of new and efficient methods (Lok, 2015). Microbial cultivation strategies, with their common theme in mimicking environmental conditions in laboratory settings (Kaeberlein, 2002), require massive miniaturization and brute force cultivation techniques while still considering the microorganisms' metabolic needs (Keller and Zengler, 2004). Advanced miniaturization approaches are microfluidics-based strategies, posing a rapidly emerging technology for cultivation of microbial diversity that were

also already utilized in the field of NP discovery (Zinchenko *et al.*, 2014; Mahler *et al.*, 2015; Terekhov *et al.*, 2017). Droplet microfluidics is based on converging aqueous and oil phases in a laminar flow with the addition of surfactants as stabilizing agents. Pressurized in microchannels with a diameter ranging from 30 to 500  $\mu\text{m}$ , this ultimately results in aqueous droplets in nl to pl scale (Leman *et al.*, 2015). The production of millions of droplets per hour was improved over years, and simplified systems are nowadays on the market (Nge *et al.*, 2013; Volpatti and Yetisen, 2014). Their combination with further technologies, such as fluorescence-activated cell sorting (FACS), enables the direct identification of desired events and their arrayed sorting (Zinchenko *et al.*, 2014). An advantageous value of these technologies is the downscaling and compartmentalization which allows single cells to be physically separated into distinct vessels while maintaining the overall microbial complexity (Zengler *et al.*, 2002; Theberge *et al.*, 2010). This separation is mandatory to avoid growth competition between different species while also allowing single cells to utilize and shape their microenvironment according to their specific needs and at their own pace (Keller and Zengler, 2004; Boitard *et al.*, 2015). The pico-litre scaled droplets display an environmental mimicry themselves, since the extremely small scale enables even single cells to adapt their environment by, for example accumulating self-mediating growth factors eventually breaking microbial dormancy (Boedicker *et al.*, 2009; Ishii *et al.*, 2010; Stewart, 2012). In principal, high-throughput microfluidic-based platforms increase the microbial cultures in amount but also diversity and consequently the probability to cultivate and identify also underexplored microorganisms (Akselband *et al.*, 2006; Baret *et al.*, 2009).

In this study, we succeeded in isolating an extended taxonomic diversity and accessing microorganisms considered as under- or even unexploited for NP discovery. We implemented a biphasic workflow consisting of (i) an efficient cultivation of microbial diversity using microfluidics and FACS-based technologies, followed by (ii) a miniaturized bioactivity-guided NP discovery process on the obtained isolates focused on the two devastating pathogens, that is *Mycobacterium tuberculosis* (MTB) and SEPTTR. Our here described cultivation platform is based on long-term stable agarose-solidified microdroplets ( $\sim 40 \mu\text{m}$  in size, volume of  $\sim 33 \text{ pl}$ ). Droplets were generated up to rates of  $\sim 1.3 \text{ kHz}$ ; thereby, exceeding the dimensions for an application to be affiliated as ultra-high throughput (Payne *et al.*, 2020). At a cell distribution of  $\lambda 0.1$ , this set up allowed the parallel encapsulation of  $\sim 500\,000 \text{ cells h}^{-1}$  with a statistical probability below 0.5% to obtain co-cultures. In total, we brought 1071 microorganisms into culture, from which

~ 74% could be identified by 16S rRNA gene sequencing. The isolated bacteria belong to five different phyla affiliated to 57 different genera. Besides classical NP producing taxa a representative of fastidious Acidobacteria and certain proteobacterial genera which are under-represented in public strain libraries (e.g. *Luteibacter* and *Variovorax*) were brought into culture. The entire obtained microbial diversity was eventually integrated into a bioactivity-guided NP discovery process. After redundancy curation of initially active extracts, the organic extracts of *Erwinia*, *Pseudomonas* and *Streptomyces* which most strongly inhibited the growth of MTB and/or SEPTTR were followed up by bioassay- and UHPLC-QTOF-MS/MS-guided fractionation. Complemented by extensive metabolomic analysis via molecular networking, this led to the identification of serratomolides, massetolides, phospholipids and macrotetrolides as the bioactivity causing agents.

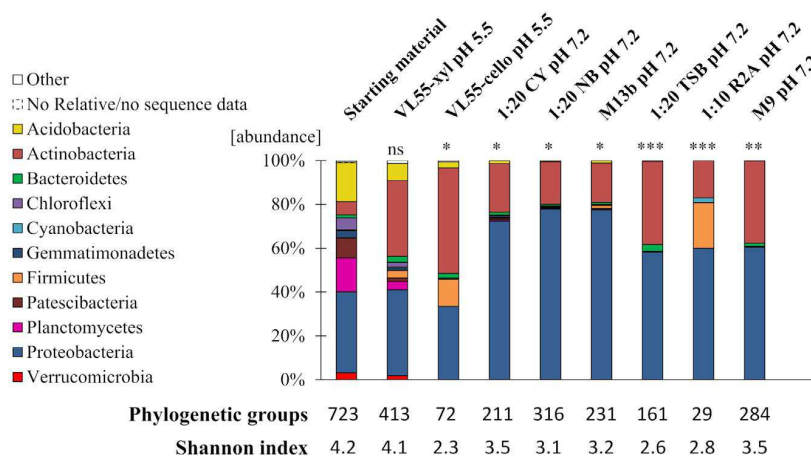
## Results and discussion

### Cultivation of microbial diversity using microfluidics and FACS

In this study (platform scheme see Fig. S2), we aimed to bring microbial diversity into culture to expand our strain collection (Fox, 2014) and to exploit the bacterial proportion in a bioactivity-guided NP discovery programme. Known for their highly diverse bacterial communities (Delgado-Baquerizo *et al.*, 2016), we selected a combined soil sample as starting material for our study.

Upfront the main droplet cultivation and screening campaign, we examined the general cultivation success

using our microfluidics platform and explicitly the specific impact of the chosen growth media on the bacterial community. Specifically adapted towards soil and previously shown to be suitable to transfer a high diversity of soil microorganisms into culture, we selected ISEM (Nguyen *et al.*, 2018) as benchmark medium for this study. Considering the importance of using buffered media as well as mimicking neutral and acidic soils for the cultivation success (Overmann *et al.*, 2017), we buffered the media to pH 7.2 and pH 5.5 (Tovar *et al.*, 2020). In order to identify an additional media suitable to access a broad microbial diversity of the bioresource used in this study, we encapsulated the retrieved microorganisms ( $\lambda 0.1$ ) in eight different media (VL55-xyl, VL55-cello, 1:20 CY, 1:20 NB, M13b, 1:20 TSB, 1:10 R2A and M9). Then, we comparatively determined the cultivation success by Illumina amplicon sequencing of the sorted droplet populations among themselves and towards the starting material. The sequencing was conducted on DNA directly isolated from the eight different droplet populations. Illumina amplicon sequencing of isolated DNA yielded in total 1 947 118 classified sequences, of which 0.04% could not be classified and were defined as *No Relative*. The data revealed that cultivation in VL55-xyl resulted in the overall highest diversity on genus level resolution (Shannon index [SI]: 4.1), followed by M9 (SI: 3.5) and 1:20 CY (SI: 3.5) (Fig. 1). Compared to the starting material, the microbial diversity covered by VL55-xyl was not significantly reduced ( $P > 0.5$ ). In contrast, the exchange of the sole C-source xylan by cellobiose led to a significant reduction in diversity ( $P < 0.01$ ) and a lower SI (SI: 2.3). While showing an



**Fig. 1.** Microbial diversity grown in agarose-solidified droplets by Illumina amplicon sequencing proves VL55-xyl as best cultivation media to obtain highest diversity. Isolated environmental cells (starting material) were encapsulated in droplets ( $\lambda 0.1$ ) and incubated at 28°C for seven days before total DNA was directly extracted from the droplet populations and sequenced with Illumina 300-bp paired-end 16S V3-V4 amplicon next-generation sequencing, respectively. VL55-xyl outperformed with the highest number of phylogenetic groups (491) and Shannon index on genus level of 4.1. A significant difference between media sample medians was observed on phylum level ( $P > 0.005$ ). Dunn's post hoc test was performed on all samples on phylum level compared to the starting material ( $***P < 0.001$ ;  $**P < 0.01$ ;  $*P < 0.1$ ; ns: not significant).



418 M. Oberpaul *et al.*

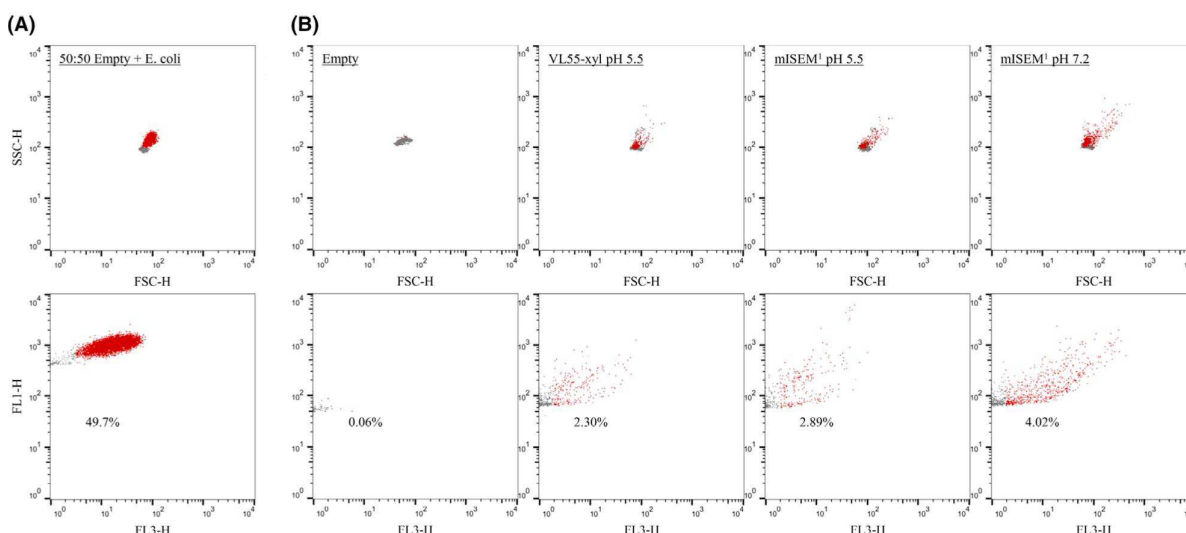
overall lower diversity, the media (1:20 CY, 1:20 NB, M13b, 1:20 TSB and 1:10 R2A), which contained complex components such as yeast or peptone showed an increased proportion of Proteobacteria (up to 77.7%) in comparison with the VL55-based media (up to 39.2%) and the starting material (36.7%). A similar diversity and an increased proportion of Proteobacteria were observed by using M9. While 1:10 R2A appears to be suitable to enrich Firmicutes (20.8% vs 0.06% in starting material), all other media buffered at pH 7.2 do not show a remarkable difference within their cultivated microbial community ( $P > 0.7$ ; Fig. 1).

Overall, VL55-xyl with its complex C-source and its pH favourable for acidophilic and acid-tolerant bacteria was shown to be the most suitable medium for our soil material and selected it to be included in all subsequent experiments. Considering the observed effects of the Illumina analysis, particularly the impact of the provided C-source and the altered pH profile, we supplemented ISEM with a variation of six complex C-sources including xylan (mISEM<sup>1</sup>) and buffered it to pH 5.5 and pH 7.2.

In this study, we applied a micron-scaled cultivation platform to receive axenic cultures from a complex biore-source, adapted for the purpose to also cultivate slow-growing microorganisms. This necessitates the analysis of droplets in high-throughput fashioned manner and the reliable discrimination between droplets harbouring grown microorganisms with manifold morphologies and the

empty droplet background. Therefore, we used a FACS approach utilizing the carbocyanine dye DiOC<sub>2</sub>(3) to stain microorganisms with a membrane potential grown inside the agarose-solidified droplets. DiOC<sub>2</sub>(3) is a potentiometric probe that exhibited green fluorescence in all bacterial cells; however, the fluorescence shifts towards red emission as the dye molecules self-associate at the higher cytosolic concentrations caused by larger membrane potentials (Shapiro, 2000; Biener *et al.*, 2017). To show the feasibility of using DiOC<sub>2</sub>(3) for this process step, a mixed droplet population containing 50% droplets harbouring grown microcolonies of *E. coli* ( $\lambda$ 10) incubated for 4 h at 37°C and 50% empty droplets was stained and subsequently measured using FACS while analysing the ratio of green (FL1-H) and red (FL3-H) fluorescence. The FL1-H to FL3-H scatter exposed 49.7% of all analysed events in the defined gate, congruent with the adjusted 50.0% of the droplet population containing grown *E. coli* microcolonies (Fig. 2A). This verified the applicability of DiOC<sub>2</sub>(3) as a fluorescent dye for staining microorganisms grown in agarose-solidified droplets. Based on these distinguishable populations, the FACS settings (e.g. sorting gates for events of interest) were defined and later on applied for the cultivation campaign.

The cell suspension from the used soil sample retrieved using nycodenz gradient centrifugation had an estimated living cell count of  $\sim 4.2 \times 10^6$  cells ml<sup>-1</sup> determined via FACS [considering live/dead ratio ( $\sim 70:30$ ; Fig. S3)].

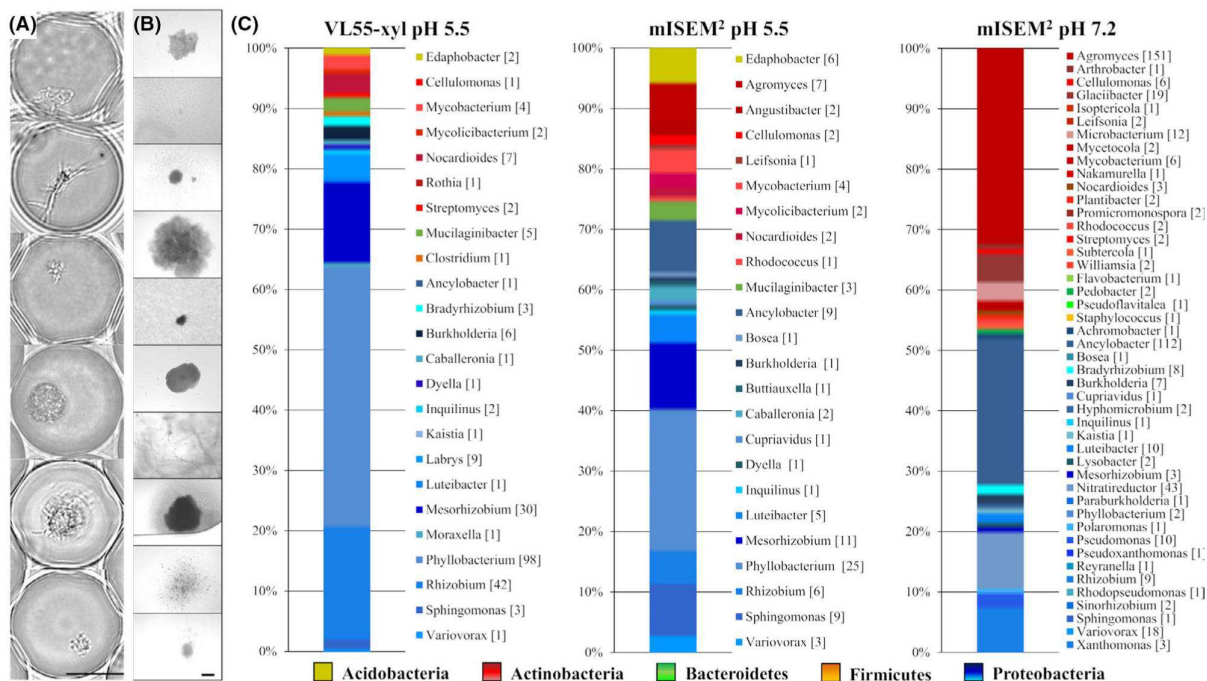


**Fig. 2.** Sorting of microcolonies grown in agarose-solidified droplets stained using DiOC<sub>2</sub>(3).

A. Differentiation of a 50:50 mixture of empty droplets and those containing grown microcolonies of *E. coli* ATCC 25922 ( $\lambda$ 10). 15 000 events were analysed by FACS considering granularity (SSC-H) to size (FSC-H) and green (FL1-H) to red fluorescence (FL3-H) after staining. B. Sorting of droplets containing grown microcolonies of environmental soil microorganisms ( $\lambda$ 0.1) in VL55-xyl pH 5.5, mISEM<sup>1</sup> pH 5.5 and mISEM<sup>1</sup> pH 7.2 (10 000 events displayed). The background noise was 0.06% (empty = control without cells encapsulated). After seven days of incubation at 28°C, grown microcolonies as determined by membrane potential dye were sorted and recovered. Red dots = droplets containing fluorescently labelled cells; grey dots = empty droplets.

This was taken into consideration, while cells were encapsulated in different media ( $\lambda 0.1$  each) and subsequently incubated at 28°C, respectively. The microbial growth within droplets was frequently monitored over time by microscopy, exemplarily showing the processing of microorganisms with various morphologies (Fig. 3A). After seven days of incubation, the samples were sorted by using DiOC<sub>2</sub>(3) staining. We could affiliate 2.30% (VL55-xyl pH 5.5), 2.89% (mISEM<sup>2</sup> pH 5.5) and 4.02% (mISEM<sup>2</sup> pH 7.2) of all evaluated events to droplets containing grown microcolonies (Fig. 2B). These events were sorted and distributed into 384-well MTPs, and microbial growth was again monitored by turbidimetry and microscopy, respectively (Fig. 3B). After automatic transfer of grown cultures into 96-deep well MTPs and further incubation for up to 14 days, this up-scaling procedure provided 1071 cultures in total. Sample aliquots of each culture were taken to generate cryo-conserved cultures and cell lysates for taxonomic identification by 16S rRNA gene sequencing (partial sequencing using primer 1492R). This revealed the processing of 57 different genera affiliated to five different phyla (Fig. 3C and Table S1). Comparing the cultivation,

success of mISEM<sup>2</sup> at pH 5.5 and pH 7.2 clearly shows the worth of modulating this parameter to increase the cultivated diversity. Together, 55 genera were isolated using mISEM<sup>2</sup>, while only 15 of them could be found under both pH conditions. Medium mISEM<sup>2</sup> pH 7.2 led to the successful cultivation and affiliation of 46 genera (462 strains, SI: 2.4) while 31 genera were unique to this condition. In contrast, at pH 5.5 a total of 24 genera (108 strains, SI: 2.7) and 9 unique ones were identified. While use of acidified media led to an overall decreased amount of genera, this condition showed a higher diversity and expanded the cultivation success towards a fifth phyla (Acidobacteria) not present at the neutral pH set-up. Within both media comprising pH 5.5, we found eight representatives of the Acidobacteria, all belonging to the genus *Edaphobacter*. A representative thereof was recovered on plate and incorporated into our strain collection. As judged on nearly full-length 16S rRNA sequence comparison, this particular strain, namely FHG110552, is phylogenetically affiliated to the subgroup 1 of Acidobacteria and most closely related to *Edaphobacter modestus* Jbg-1T (~ 98.8% similarity) isolated from alpine and forest soils (Koch *et al.*, 2008;



**Fig. 3.** Successful cultivation of microbial diversity after incubation of solidified-agarose droplets verified by 16S rRNA gene sequencing (1492R primer) and microscopy.

A. Examples of grown microcolonies in agarose-solidified droplets were microscopically examined. Scale bar: 20  $\mu$ m.

B. Inverse microscopy of exemplary 384-well MTPs containing arrayed droplets and liquid medium. Different morphologies from microcolonies, which grew out of droplets are indicated, representing a section of the final retrieved and conserved bacterial diversity. Scale bar: 40  $\mu$ m.

C. In total, 1071 MOs affiliated to 5 phyla were brought into culture, whereas ~ 74% were initially identified by 16S rRNA gene sequencing (partial sequence using primer 1492R, mISEM<sup>2</sup> pH 7.2: 462, mISEM<sup>2</sup> pH 5.5: 106, VL55-xyl pH 5.5: 225). Numbers in square brackets indicate the amount of 16S rRNA gene sequencing affiliations. For further information and data used for the figure caption, see Table S1.

420 M. Oberpaul *et al.*

Fig. S4). The phylum Acidobacteria is postulated to be a promising bioresource for the field of NP discovery based on its predicted biosynthetic potential encoded in their genomes (Kielak *et al.*, 2016; Crits-Christoph *et al.*, 2018). Though, the description of Acidobacteria NPs is yet limited to hopanoids (Damsté *et al.*, 2017).

On phylum level, a major difference between the neutral and acidified media is the ratio between Actinobacteria and Proteobacteria, which is at pH 7.2 shifted towards the Actinobacteria, while towards Proteobacteria at pH 5.5. The same is seen in VL55-xyl covering in total 23 genera (223 strains, SI: 1.9) including *Edaphobacter* but also unique genera such as *Labrys* and *Rothia*, not covered by mISEM<sup>2</sup>. In total, 17 Actinobacteria genera were brought into culture using mISEM<sup>2</sup> pH 7.2, whereas 8 genera in mISEM<sup>2</sup> pH 5.5 and 6 genera in VL55-xyl. This is particularly biased by the genus *Agromyces*, which was strongly enriched in mISEM<sup>2</sup> pH 7.2 (151 strains) while present in low abundance at pH 5.5 (7 strains).

Within the Proteobacteria, such an opposite pH-dependent shift occurs within the class of Alphaproteobacteria. At pH 7.2, the genus *Ancylobacter* represents almost half of all Proteobacteria (112 strains) while being significantly less abundant at pH 5.5. In contrast, the genus *Phyllobacterium* represents > 30% of the Proteobacteria at acidic condition while < 1% at neutral pH. In total, the isolated Proteobacteria proportion comprises member of 33 different genera belonging to the classes of Alpha-, Beta- and Gammaproteobacteria and include also representatives of nine scarcely cultured genera according to our definition (namely *Ancylobacter*, *Buttiauxella*, *Inquilinus*, *Kaistia*, *Labrys*, *Luteibacter*, *Polaromonas*, *Reyranella* and *Variovorax*).

#### Bioactivity-guided bioprospecting process

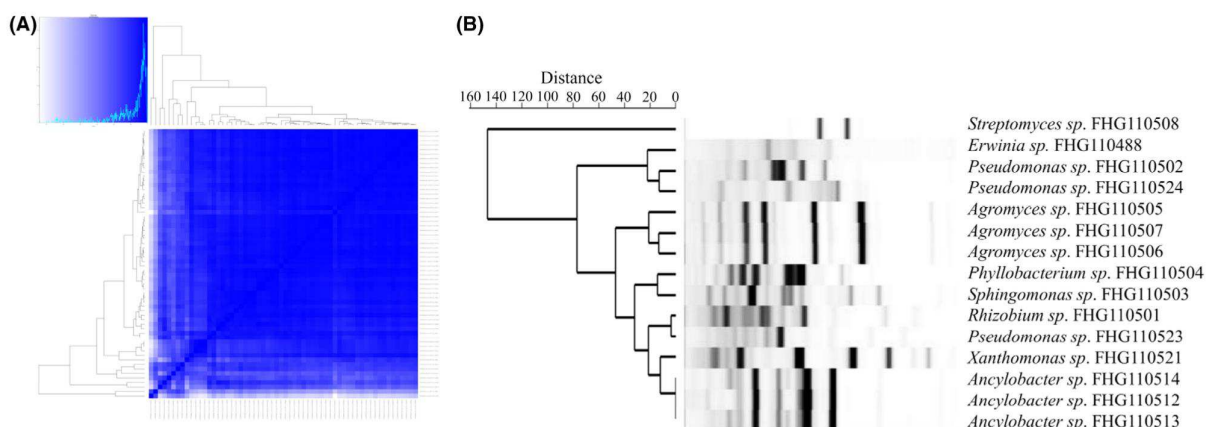
To prospect the here cultured microbial diversity for NPs with desired bioactivity, we integrated all arrayed microbes in a downscaled bioactivity-guided NP discovery process. Each culture was grown in its respective isolation medium and in addition in ISP2 and BSM for four and seven days. This process aimed to more sophisticatedly provoke expression of the microbial secondary metabolite production potential. Using 96-deep well Duetz systems and 1 ml culture volumes, 6426 organic extracts were generated and used for bioactivity screenings and metabolome analysis.

In first line, the extracts were screened for their antimicrobial activity against *M. smegmatis*, the opportunistic microbial pathogen *S. aureus* and for their antifungal activity against *Z. tritici*, *C. albicans* and *A. flavus*. Extracts that showed growth inhibitory properties underwent a subsequent diversity assessment by cosine similarity comparison of features detected within each corresponding

UHPLC-UHR-MS chromatogram. This quality control step was included to rule out extensive analysis of redundant samples, unavoidably appearing in every microbial isolation project. After screening the methanolic extracts, 64 cultures were assigned to bioactivity by either inhibiting the growth of *M. smegmatis* and/or the fungal indicator strains. The cosine similarity of their associated UHPLC-UHR-MS chromatograms were determined. After excluding samples with less than 50 features, the remaining 60 samples clustered into eight distinct groups applying the cosine similarity threshold 0.9 (Fig. 4A, Table S2). From the most dominant group, we recovered *Penicillium* sp. (FHG110518). With no focus on fungal metabolites in this project, we decided to discontinue the work on this abundant strain causing the major fractions of observed bioactivities (~ 70%). In order to now identify the bacterial strains in this dataset, we complemented the redundancy curation process by genotyping the cultures using BOX-PCR (Fig. 4B). By superimposing the patterns of both grouping technologies, we identified and subsequently recovered strains of the genera *Pseudomonas* [FHG110502, FHG110523 and FHG110524 (each unique)], *Xanthomonas* (FHG110521), *Rhizobium* (FHG110501), *Sphingomonas* (FHG110503), *Phyllobacterium* (FHG110504), *Ancylobacter* [FHG110512, FHG110513, FHG110514 (equal)], *Agromyces* [FHG110505, FHG110506, FHG110507 (equal)], *Erwinia* (FHG110488) and *Streptomyces* (FHG110508).

In order to identify the causative NP beyond the observed crude extract activities, the cultivation of the unique strains was scaled-up to 50 ml volumes. The resulting culture broths were extracted by methanol and reconstituted in DMSO for bioactivity screenings. The extracts from the three strains *Pseudomonas* sp. (FHG110502), *Erwinia* sp. (FHG110488) and *Streptomyces* sp. (FHG110508) exhibited the severest growth inhibitory potency and were selected for further analysis. Their crude extract complexity was strongly reduced by their separation into 159 fractions. After rescreening all fractions, the compounds present in the active fractions were identified by their exact masses, characteristic UV absorption spectra and fragmentation signature.

*Erwinia* sp. FHG110488. Based on the primary activity of the crude extract against *M. smegmatis* and *Z. tritici*, the corresponding extract was fractionated and rescreened. Compound 1 with a *m/z* of 515.3329 [M + H]<sup>+</sup>, corresponding to the molecular formula C<sub>26</sub>H<sub>46</sub>N<sub>2</sub>O<sub>8</sub>, was dereplicated within an active fraction (Fig. S5). On the basis of the molecular formula and the observed MS/MS fragmentation pattern, the bioactive compound was dereplicated as the known serratamolide A (Wasserman *et al.*, 1961; Dwivedi *et al.*, 2008). Using MS/MS networking, six additional derivatives were



**Fig. 4.** Metabolic grouping of extracts based on the cosine similarity of their associated UHPLC-UHR-MS chromatograms and the respective genetic fingerprinting by BOX-PCR applied to reduce the redundancy of active strains.

(A) Metabolic clustering of extracts of all bioactive strains was analysed via LC-MS. The cosine similarities between samples were calculated. Cosine similarity values for strain-condition pairs were extracted. Samples were sorted according to clustering results, and pairwise similarities were used to define metabolic groups. If the pairwise similarity between two subsequent clustered samples was at the threshold 0.9 or higher, they were assigned to one metabolic group. Figure caption can be found in Table S2.

(B) Repetitive genome sequences of all strains initially found to be bioactive were targeted by BOX-PCR to identify unique strains.

[Correction added on 01 July 2021, after first online publication: Figure 4 legend has been amended for clarity in this version]

identified (measured **2**  $m/z$  541.3480  $[M + H]^+$ ,  $C_{28}H_{48}N_2O_8$ , theoretical  $m/z$  541.3483; **3**  $m/z$  543.3635  $[M + H]^+$ ,  $C_{28}H_{50}N_2O_8$ , theoretical  $m/z$  543.3639; **4**  $m/z$  533.3430  $[M + H]^+$ ,  $C_{26}H_{48}N_2O_9$ , theoretical  $m/z$  533.3432; **5**  $m/z$  547.3588  $[M + H]^+$ ,  $C_{27}H_{50}N_2O_9$ , theoretical  $m/z$  547.3589; **6**  $m/z$  575.3896  $[M + H]^+$ ,  $C_{29}H_{54}N_2O_9$ , theoretical  $m/z$  575.3902; **7**  $m/z$  573.3745  $[M + H]^+$ ,  $C_{29}H_{52}N_2O_9$ , theoretical  $m/z$  573.3745 (Fig. S6). Serratamolide A is a biosurfactant with plant protecting properties (Thies *et al.*, 2014), additionally alleviating uptake of substances due to its amphiphilic wetting effect, lowering surface and interfacial tensions (Mulligan, 2005). It is known to show antioomycetous activity with particular efficacy against *Pythium ultimum* and *Phytophthora parasitica* (Strobel *et al.*, 2005), a further devastating pest with lack in effective control measures under investigation (Pöppel *et al.*, 2014).

*Pseudomonas sp. FHG110502*. The antimycobacterial activity caused by *Pseudomonas* strain FHG110502 was associated to three cyclo-lipo-nonapeptide **8–10** (**8**  $m/z$  1112.6811  $[M + H]^+$ ,  $C_{53}H_{93}N_9O_{16}$ , theoretical  $m/z$  1112.6813; **9**  $m/z$  1126.6966  $[M + H]^+$ ,  $C_{54}H_{95}N_9O_{16}$ , theoretical  $m/z$  1126.6969; and **10**  $m/z$  1154.7288  $[M + H]^+$ ,  $C_{56}H_{99}N_9O_{16}$ , theoretical  $m/z$  1154.7282) perfectly matching the MS/MS fragmentation signature of massetolide E, F and H, respectively. Massetolide E and F differ in AA9 by incorporation of valine and leucine, respectively. Massetolide H differs in the fatty acid chain. The hydroxydecanoic acid residue of E and F is replaced by a hydroxylauric acid residue (Fig. S7). At

position AA9, the known antimycobacterial massetolide family is known to be prone to natural variations, with indications in a general correlation between greater lipophilicity and increased potency (Gerard *et al.*, 1997).

Furthermore, three phospholipids dereplicated as lyso-palmitoyl-phosphoethanolamine (**11**), palmitoleoyl-palmitoyl-phosphoethanolamine (**12**) and palmitoleoyl-oleoyl-phosphoethanolamine (**13**) (measured **11**  $m/z$  454.2931  $[M + H]^+$ ,  $C_{21}H_{44}N_1O_7P_1$ , theoretical  $m/z$  454.2928; **12**  $m/z$  690.5073  $[M + H]^+$ ,  $C_{37}H_{72}N_1O_8P_1$ , theoretical  $m/z$  690.5068; and **13**  $m/z$  716.5237  $[M + H]^+$ ,  $C_{39}H_{74}N_1O_8P_1$ , theoretical  $m/z$  716.5224), respectively, were assigned the other active fractions (Fig. S8). The bioactivity of phospholipids against several indicator strains has been reported, for example bacillosin as an antifungal active phospholipid isolated from *Bacillus subtilis* (Tamehiro *et al.*, 2002).

*Streptomyces sp. FHG110508*. A classical NP producer organism that was isolated during our process is *Streptomyces* FHG110508. Its organic extracts showed pronounced activity against *S. aureus* and the fungal indicator strains. Dereplication of active fractions revealed the presence of the macrotretrolides nonactin, monactin, dinactin and macrotretrolide G (Fig. S9, **14–17**) (Phillies, 1975; Řezanka *et al.*, 2010). Additionally, the presence of macrotretrolide D (**18**) was shown by molecular networking (Fig. S10). Ionophore antibiotics of the macrotretrolide family are beyond the most commonly observed bioactive metabolites from actinomycetes passing in various screening disciplines having antibacterial, antifungal,

422 M. Oberpaul *et al.*

antiprotozoan, antiparasitic, insecticidal and acaricidal activity (Zizika, 1998).

**Conclusive remarks.** Our study represents a combination of applied microfluidics and FACS technologies for the cultivation of microorganisms in agarose-solidified droplets. The combination of streamlined high-throughput technologies facilitated the generation and analysis of thousands of droplets within seconds. This paved the way to a fast isolation and characterization of diverse axenic cultures, followed by downstream identification of bioactive natural products. There are no limits on a particular bioresource, organism types or cultivation conditions (media, nutrient limitation, temperature, incubation method,  $\pm$ oxygen supply, etc.), since especially all parameters towards cultivation in solidified droplets are highly customizable. Microfluidic platforms are already today a core element to access and assess microbial diversity from environmental samples. Their efficient application on a diversity of bioresources will be of steadily increasing importance for various research disciplines, with particular (but not exclusive) regard to the field of environmental microbiology.

### Experimental procedures

#### *Sampling procedure of forest soil*

Forrest soil samples (*soil*) were taken from the 'Nature Conservation Area' Bergwerkswald at 50.564032 N 8.672555 E (Hasenkoeppel, Giessen, Germany). Five samples within a radius of ten metres (max. 10 cm in depth) including rhizosphere, lichen, humus soil and sandy loam were pooled. Samples were taken in December 2018 (Illumina approach) and January 2019 (cultivation campaign). Microorganisms were retrieved by Nycodenz density gradient centrifugation (Barra Caracciolo *et al.*, 2005; Oberpaul *et al.*, 2020). All samples were directly chilled and stored at 4°C until processing for the cultivation or at -50°C for Illumina amplicon sequencing.

#### *Encapsulation of microorganisms using microfluidic devices*

The cell concentration of retrieved microorganisms was determined using the Bacteria Counting Kit (B7277, Invitrogen, Carlsbad, CA, USA) on a FACSCalibur (BD Bioscience, San Jose, CA, USA) according to the manufacturer's protocol. The live/dead ratio was estimated using the LIVE/DEAD BacLight Bacterial Viability and Counting Kit (L7007, Invitrogen).

The commercially available  $\mu$ Encapsulator microfluidics set-up, described in detail elsewhere (Caballero-Aguilara *et al.*, 2021), consisting of three pulse-free pressure pumps (Mitos P-Pump), a temperature control unit (TCU-100),

diverse microfluidic chips, polytetrafluoroethylene (PTFE) tubing (1/16" OD, 0.25 mm ID and 0.5 mm ID) and connectors were purchased from Dolomite Microfluidics, a brand of Blacktrace Holdings Ltd (Royston, UK). A high-speed CMOS camera PL-D721CU (Navitar, Rochester, NY, USA) on a stereomicroscope Stemi SV 11 (Carl Zeiss, Oberkochen, Germany) equipped with a halogen light source KL 2500 LCD (Schott AG, Mainz, Germany) was used to image microfluidic operations. Samples for encapsulation were loaded onto the  $\mu$ Encapsulator Sample Reservoir Chip. Droplet generation was realized on a fluorophilic 50  $\mu$ m  $\mu$ Encapsulator 2 Reagent Droplet Chip with an applied water to water to oil ratio of 1:1:20 at 30°C for ~ 30 min (Video S1). Pico-Surf 1 (2% (w/w) in FC-40) (Sphere Fluidics, Cambridge, UK) was used as the continuous oil phase. Both aqueous phases consisted of cultivation media with adjusted environmental or *E.coli* cell concentrations to obtain target cell distributions following Poisson (Fig. S1; Collins *et al.*, 2015). All liquids except cell suspensions were filtered through a 0.2  $\mu$ m CA syringe filter (Coming, Coming, NY, USA) to prevent blocking of the chip. Prior chip loading, media were mixed 50:50 with liquid, pre-warmed 3% (w/v) SeaPlaque agarose (Lonza, Basel, Switzerland) in water. The droplets were collected in 1.5 ml reaction tubes and cooled at 4°C for 10 min to facilitate gelling of the agarose. Thereafter, droplets were incubated at 28°C in a humidity chamber for seven days for the Illumina amplicon sequencing and for the cultivation experiments. Neubauer chambers (0.1 mm depth, Paul Marienfeld GmbH KG, Lauda-Königshofen, Germany) were used for imaging of stationary droplets on a fluorescence microscope DM2000 LED equipped with a DFC450 C camera and LAS V4.7 software (Leica Microsystems, Wetzlar, Germany) for picture analysis.

#### *Cell staining and droplet sorting via FACS*

Analysis and sorting of droplets were performed using a FACSCalibur. Surrounding oil was de-emulsified upfront fluorescence staining and FACS analysis using Pico-Break 1 (Dolomite Microfluidics) added in a ratio of 1:200 to a sample diluted with  $1 \times$  PBS to adjust towards an analysis of 2000 events  $s^{-1}$ . Fluorescence labelling of microorganisms was achieved by adding 30  $\mu$ M DiOC<sub>2</sub>(3) (B34950, Invitrogen) to the samples and an incubation time of 5 min. Sorting via FACS (laser ex: 488 nm, em: 530  $\pm$  30 nm [FL1-H] and 670  $\pm$  30 nm [FL3-H]) was carried out using the single-cell mode and  $1 \times$  PBS as sheath fluid.

#### *Cultivation media*

Media used for the growth of microorganisms in droplets, and the subsequent Illumina 16S V3-V4 gene amplicon

sequencing were as follows: VL55 supplemented with 0.05% (w/v) xylan (VL55-*xyl*) or 0.05% (w/v) cellobiose (VL55-*cello*), including selenite–tungstate solution and trace elements SL-10 (Oberpaul *et al.*, 2020) and buffered to pH 5.5. In addition, 1:10 diluted Reasoner's 2A medium (1:10 R2A, DMSZ medium 830), Minimal medium M9 (DMSZ medium 382), 1:20 diluted Casitone-Yeast medium (1:20 CY, DMSZ medium 67), 1:20 diluted Nutrient Broth (1:20 NB, DMSZ medium 1), M13 with 200 mg l<sup>-1</sup> ampicillin (M13b; (Wiegand *et al.*, 2019)) and 1:20 diluted Tryptic Soy Broth medium (1:20 TSB, Sigma-Aldrich, St. Louis, MO, USA) were used.

For the cultivation campaign of microbial diversity using microfluidics and FACS, either VL55-*xyl* or modified intensive soil extract medium (ISEM) (Nguyen *et al.*, 2018) were used. ISEM was modified by addition of 1% (w/v) C-source solution consisting of arabinogalactan, D-(+)-cellobiose, D-(+)-melezitose, xylan, galacto-D-mannan from *Ceratonia siliqua* and N-acetylglucosamine (25 mg l<sup>-1</sup> of each in deionized water) (mISEM<sup>1</sup>). For the cultivation of sorted and arrayed events in 384-well microplates (MTPs), we exchanged the soil extract with 0.1 g l<sup>-1</sup> yeast extract (Oxoid) and 0.1 g l<sup>-1</sup> casamino acids (Difco), 0.1 g l<sup>-1</sup> proteose peptone (Roth, Karlsruhe, Germany) (mISEM<sup>2</sup>). Trace elements SL-10 and selenite-tungstate solution (each 2 ml l<sup>-1</sup>) were added to all media used for the cultivation experiments.

For the bioprospecting campaign, cultures were fermented in small-scaled 96-well Duetz system (Duetz *et al.*, 2000) (Adolf Kühner AG, Birsfelden, Switzerland) using VL55-*xyl*, ISP2 (DSMZ medium 987), basal salt medium (BSM) supplemented with glycerol (Marner *et al.*, 2020) and mISEM<sup>2</sup>, adjusted to pH 5.5 and 7.2.

#### *Illumina amplicon sequencing and statistical analysis*

In order to evaluate the relative cultivation success, we encapsulated the retrieved environmental microorganisms ( $\lambda$ 0.1) considering the live/dead ratio in eight different media and incubated the droplets for seven days at 28°C. Environmental DNA was extracted using the NucleoSpin<sup>®</sup> Soil DNA purification kit (Macherey Nagel, Düren, Germany) according to manufacturer's protocol. The bacterial community composition of samples was assessed by Illumina 300-bp paired-end 16S V3-V4 amplicon next-generation sequencing using the degenerate primer pair 341F (3'-CCTACGGGNGGC WGCAG-5') and 785R (3'-GACTACHVGGGTATCTAA KCC-5'). Sequencing was performed by LGC Genomics GmbH (Berlin, Germany) on a MiSeq (Illumina, San Diego, CA, USA), and data evaluation was supported by the SILVAngs pipeline (SILVA SSU Ref dataset; release 132; <http://www.arb-silva.de>; SILVA

Incremental Aligner (SINA SINA v1.2.10 for ARB SVN [revision 21008] Ribocon GmbH, Bremen, Germany)) and BLASTN v2.2.30+ as previously described (Oberpaul *et al.*, 2020). Phylogenetic groups with a relative abundance < 0.001% of total were excluded. Statistical calculations were done by using PAST v4.03 (Hammer *et al.*, 2001)) including the several-sample one-way ANOVA test on ranks (Daniel, 1990) followed by a Dunn's post hoc test (Dunn, 1964) to judge on significances among samples.

#### *Microbial cultivation and extract preparation*

Sorted droplets were concentrated using a 0.5  $\mu$ m CA membrane filter (Whatman plc, Little Chalfont, UK) and a laboratory vacuum filtration system (Sartorius AG, Göttingen, Germany). Concentrated droplets were recovered in the respective growth medium (VL55-*xyl*, mISEM<sup>2</sup> pH 5.5 or pH 7.2) and arrayed into 384-well MTPs with a distribution probability of approx. 0.25 droplets per well using a Matrix Wellmate microplate dispenser (Thermo Fisher Scientific, Waltham, MA, USA). After incubation as static cultures at 28°C in a humidity chamber for up to seven days, growth detection was assessed via turbidimetry (OD) at 600 nm using a Wallac 1420 Victor2 Microplate Reader (Perkin Elmer, Waltham, MA, USA) and by microscopy using a Zeiss Axiovert 200M (Carl Zeiss Microscopy GmbH, Jena, Germany) equipped with an SPOT RT Monochrome 2.1.1 camera (Diagnostic Instruments, Sterling Heights, MI, USA). All cultures exceeding our defined turbidity threshold (determined based on media control blanks present on each processed microplate) were automatically transferred into 96-deep well MTPs pre-filled with media (Corning, New York, NY, USA) using a Precision XS liquid handling system (BioTek Instruments GmbH, Bad Friedrichshall, Germany). These plates were incubated using the Duetz system at 28°C for seven days, shaking at 220 rpm, and 2.5 cm deflection. Sample aliquots were taken for DNA preparation (for 16S rRNA gene sequencing) and to generate a cryo-conserved culture for long-term conservation at -80°C.

For preparation of cryo-conserved samples, 70% (v/v) glycerol and 5% (v/v) DMSO were filled into 96-deep well MTPs and mixed with grown culture broths (ratio 2:3) using a VIAFLO 384 (Integra Biosciences, Biebertal, Germany).

Cryo-conserved strains were recovered by agar plating and integrated into our strain collection. Bioprospecting campaigns were conducted either in 1 ml culture volume using 96-well Duetz systems or in 50 ml culture volume using 300 ml Erlenmeyer flasks, respectively. Incubation of Erlenmeyer flasks occurred at 28°C using an RC-406 orbital shaker (Infors, Bottmingen, Switzerland) with

424 M. Oberpaul *et al.*

5 cm deflection at 180 rpm for four and seven days. Fermentations were stopped by cooling microbial cultures and medium controls to  $-50^{\circ}\text{C}$ . Frozen samples were lyophilized using a delta 2-24 LSCplus (Martin Christ Gefriertrocknungsanlagen GmbH, Osterode am Harz, Germany), and an equal volume of methanol (related to the culture volume) was added. Extracts were further concentrated *in vacuo* (50-fold in relation to the culture volume) for UHPLC-QTOF-MS/MS analysis. For screening purposes, extract aliquots were reconstituted in DMSO (100-fold in relation to the culture broth).

#### *Phylogenetic classification and genotyping of isolated strains*

DNA extraction was carried out by transferring 200  $\mu\text{l}$  of each grown culture supernatant into Collection Microtubes (Qiagen, Hilden, Germany) containing three zirconia beads (2.3 mm, Carl Roth, Karlsruhe, Germany). Cells were disrupted (twice, 30 Hz for 1 min) by using a TissueLyser II (Qiagen). Plates were centrifuged at  $4000 \times g$  for 5 min, incubated at  $70^{\circ}\text{C}$  for 45 min and again centrifuged. The supernatants were transferred into fresh 96-V-bottom plates and used as template for 16S rRNA gene amplification following the PCR protocol described by (Kämpfer *et al.*, 2014) using the primer pair E8F (5'-GAGTTTGATCCTGGCTCAG-3') and 1492R (5'-AGAGTTTGATCCTGGCTCAG-3') and for 18S rRNA gene amplification using the primer pair NS1 (5'-GTAGTCATATGCTTGTCTC-3') and FR1 (5'-AICCATTCAATCGGTAIT-3') following the protocol described by (Panzer *et al.*, 2015). Genera belonging to Proteobacteria were defined as rarely cultured if less than ten representatives are recorded in the List of Prokaryotic names with Standing in Nomenclature (Parte, 2018, 2018; <https://lpsn.dsmz.de/> access date: 01/2021).

All cultures affiliated to Acidobacteria were propagated on buffered R2A pH 5.5 agar to receive higher culture densities. All recovered nearly full-length 16S and rRNA sequences were affiliated to the most similar sequences of type strains using BLASTN (version: BLAST+ 2.11.0; <http://blast.ncbi.nlm.nih.gov/Blast.cgi>) with the NCBI Reference Sequence Database (version: RefSeq Release 202; <https://www.ncbi.nlm.nih.gov/refseq/>). All recovered nearly full-length 18S rRNA sequences were affiliated to the most similar sequences of type strains using the pairwise alignment tool from MycoBank ([https://www.mycobank.org/page/Pairwise\\_alignment](https://www.mycobank.org/page/Pairwise_alignment)). Representatives of different subgroups of Acidobacteria were included to visualize the phylogenetic relationship towards isolated Acidobacteria of this study. Therefore, multiple sequence alignment was done by CLUSTALW. Using the maximum-likelihood method, a phylogenetic tree was calculated in

MEGA v7.0.26 (<https://www.megasoftware.net>) under the Tamura–Nei model (Kumar *et al.*, 2016) performing 1000 bootstrap replications. Graphical modifications and annotations were made with iTOL v5.6.3 (<https://itol.embl.de/>; Letunic and Bork, 2019).

BOX-A1R-based repetitive extragenic palindromic sequence PCR (5'-CTACGGCAAGGCGACGCTGACG-3') was used to amplify repetitive elements to perform molecular genotyping (Koeuth *et al.*, 1995). Genomic fingerprinting patterns (BOX-pattern) were analysed by LabChip GX Touch HT using DNA 5K Assay (Cat. No. CLS760675, PerkinElmer, Waltham, MA, USA) and GEL-COMPAR II software version 6.5 (Applied Maths, Sint-Martens-Latem, Belgium) for data interpretation. A hierarchical cluster of the obtained data was calculated based on Dice similarity matrix data by applying Ward's method (Ward, 1963) to discriminate genotypic redundancies within pure cultures.

#### *Bioassays*

Crude extract screening was done by microplate broth dilution assays in final assay conc. 0.25-fold, 0.5-fold and onefold (in relation to the culture volume) against pathogenic bacteria and fungi (*Staphylococcus aureus* ATCC25923, *Mycobacterium smegmatis* ATCC 607, *Aspergillus flavus* ATCC 9170, *Zygomycetia* campaign *ici* Roberge in Desmazières MUCL 45407 and *Candida albicans* FH 2173).

To evaluate the growth inhibitory effect of microbial extracts, a seeding cell suspension of the indicator strains was prepared from pre-cultures or previously prepared spore solutions: for *S. aureus*, an overnight culture ( $37^{\circ}\text{C}$ , 18 h, 180 rpm) was diluted to  $2 \times 10^4$  cells  $\text{ml}^{-1}$  in cation adjusted Mueller Hinton II medium (MHII). *M. smegmatis* was cultured in brain-heart infusion medium (BD) supplemented with 1% (w/v) Tween-80 (Sigma) for 48 h, before the assay cell concentration was adjusted ( $1 \times 10^5$  cells  $\text{ml}^{-1}$ ). The pre-culture of *C. albicans* was incubated for 48 h at  $28^{\circ}\text{C}$  before a  $1 \times 10^5$  cells  $\text{ml}^{-1}$  suspension was prepared in MHII. Spore solutions of *A. flavus* and *Z. tritici* were diluted to  $1 \times 10^5$  spores  $\text{ml}^{-1}$ . A dilution series of gentamycin, isoniazid or nystatin was used as positive control and cell suspensions without extract or antibiotic as negative control. Test plates were incubated in the dark ( $37^{\circ}\text{C}$ , 180 rpm, 85% rH) for 18 h (*A. flavus*, *S. aureus*), 48 h (*C. albicans*, *M. smegmatis*) and 72 h at  $25^{\circ}\text{C}$  (*Z. tritici*). End-point detection towards optical density was determined using a LUMIstar<sup>®</sup> Omega (BMG Labtech GmbH, Ortenberg, Germany) by measuring the turbidity at 600 nm or by ATP quantification using BacTiter-Glo<sup>™</sup> according to manufacturer's protocol (Promega Corporation, Fitchburg, WI, USA).

### Mass spectrometry

All mass spectrometry experiments were performed on a 1290 UHPLC system (Agilent, Santa Clara, CA, USA) equipped with DAD, ELSD and maXis II™ (Bruker, Billerica, MA, USA) ESI-QTOF-UHRMS with the gradient: 0 min: 95% A; 0.30 min: 95% A; 18.00 min: 4.75% A; 18.10 min: 0% A; 22.50 min: 0% A; 22.60 min: 95% A; 25.00 min: 95% A (A: H<sub>2</sub>O, 0.1% formic acid (FA); B: acetonitrile, 0.1% FA; flow: 600 µl min<sup>-1</sup>). Column oven temperature: 45°C. Column: Acquity UPLC BEH C18 1.7 µm (2.1 × 100 mm) with Acquity UPLC BEH C18 1.7 µm VanGuard Pre-Column (2.1 × 5 mm). Injection volume was either 1 or 2 µl.

### Cosine similarity calculations

Data processing was performed with Data Analysis 4.4 (Bruker, Billerica, MA, USA) using recalibration with sodium formate. *RecalculateLinespectra* with threshold 10 000 and subsequent *FindMolecularFeatures* (0.5–25 min, S/N = 0, correlation coefficient threshold = 0.7, minimum compound length = 8 spectra, smoothing width = 2) was performed. Bucketing was performed using *ProfileAnalysis* 2.3 (Bruker, Billerica, MA, USA) (30–1080 s, 100–1600 m/z, Advanced Bucketing with 12 s, 5 ppm, no transformation, Bucketing basis = H<sup>+</sup>). Samples with less than 50 features were excluded from further analysis. The generated bucket table was subsequently used as input for analysis via R (version 3.6.0) (R Core Team, 2020) with libraries *readr* (<https://CRAN.R-project.org/package=readr>), *coop* (<https://cran.r-project.org/package=coop>), *gplots* (<https://CRAN.R-project.org/package=gplots>), *data.table* (<https://CRAN.R-project.org/package=data.table>), and *parallelDist* (<https://CRAN.R-project.org/package=parallelDist>). The cosine similarities (dot product of vectors) between samples were calculated. Samples were sorted according to clustering results, and pairwise similarities were extracted and subsequently used to define metabolic groups. The script used in this publication was deposited in a GitHub repository (Hartwig, 2020). If the pairwise similarity between two subsequent clustered samples is at the threshold 0.9 or higher, they were assigned to one metabolic group.

### Dereplication of bioactive extracts using UHPLC-UHR MS/MS

Crude extracts showing at least 70% growth inhibition were considered as bioactive and were subjected to microfractionation. Five or ten microliter of concentrated methanolic extracts (50-fold relative to the culture volume) were partitioned into 159 fractions (~ 7 s each) in a 384-well MTP. UHPLC-UHR-MS analysis was performed

(settings see Section Mass spectrometry). For microfractionation, 90% of the flow was collected with a custom made fraction collector (Zinsser–Analytic, Eschborn, Germany) while the rest was analysed in MS/MS mode in maXis II™. Collision induced dissociation was performed at 28.0–35.05 eV using argon at 10<sup>-2</sup> mbar. Additionally to the chromatographically separated 159 fractions, crude extract was applied as fraction 160 in the same amount as the injection volume serving as positive control. All extracts were evaporated in a GeneVac HT-12 (SP Industries Warminster, PA, USA) and were rescreened against the same test strain. Correlation between bioactive fractions and the corresponding MS data of analytes detected in those fractions was performed. Dereplication was facilitated by comparison of mass to charge ratios, retention time and fragmentation signatures with our in-house reference database containing ~ 1700 structurally characterized microbial metabolites at the time of data processing. Molecular formula assignment was done manually for all compounds present in the active fractions, allowing a mass accuracy tolerance of ± 2 ppm. Annotation of the MS/MS spectra was performed manually for all the compounds present in active fractions, whenever no hits were found in the in-house compound database. Molecular formula searches were performed on AntiBase 2017 (Laatsch, 2017) and Dictionary of Natural Products (<http://dnp.chemnetbase.com/faces/chemical/ChemicalSearch.xhtml>; accessed on Nov 16, 2020).

### Molecular networking

Based on published protocols, molecular networking with a cosine similarity cut-off of > 0.7 was performed (Yang *et al.*, 2013; Allard *et al.*, 2016). The tool MSConvert (ProteoWizard package32) was used to convert the raw data (\*.d files) into plain text files (\*.mgf), wherein all detected fragment ions are expressed as a list of mass/intensity value pairs sorted according to their parent ions (peak picking: vendor MS level = 1–2; threshold type = absolute intensity, value = 1000, orientation = most-intense). Sharing at least six fragments (tolerance Δppm 0.05) with at least one partner ion those ions were included in the final network (Riyanti *et al.*, 2020). Known NPs were highlighted by including deposited compounds from the *in silico* fragmented (Allen *et al.*, 2015) commercial database AntiBase 2017 (Laatsch, 2017) and our *in-house* reference compound MS/MS database. The data were visualized with CYTOSCAPE v3.6.0 (Shannon, 2003) as described elsewhere (Marnier *et al.*, 2020).

### Acknowledgements

The authors would like to thank Dr. Michael Mourez, Dr. Luigi Toti, Dr. Jochen Fracowiak and Mai Ngyuen for



426 M. Oberpaul *et al.*

helpful discussions. We thank Mareike Rothenheber, Christine Wehr, Jennifer Kuhn, Nadine Zucchetto, Sina Abdo, Regina Zweigert, Veronika Eckert, and Kirsten-Susann Bommersheim for their technical support during the experiments.

#### Authors contributions

MO, SB, JG, and MS conceived and designed the experiments. MO, SB, MS, SM, BL, MM, CH, MP performed the experiments. MO, SB, MS, SM, MM, CH and MP analysed the data. MO, SB, MS and TFS drafted the first manuscript. TFS, JG and MS revised the manuscript. AV and PH initiated the public-private partnership between Fraunhofer and Sanofi (later Evotec). All authors accepted the final version of the manuscript.

#### Conflict of interest

The authors declare no conflict of interest. The authors declare no competing financial interest.

#### Data availability statement

The Illumina amplicon sequencing data supporting the findings of this study are openly available under the BioProject PRJNA699730 (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA699730>).

#### References

- Adamek, M., Alanjary, M., and Ziemert, N. (2019) Applied evolution: phylogeny-based approaches in natural products research. *Nat Prod Rep* **36**: 1295–1312.
- Akselband, Y., Cabral, C., Castor, T.P., Chikarmane, H.M., and McGrath, P. (2006) Enrichment of slow-growing marine microorganisms from mixed cultures using gel microdrop (GMD) growth assay and fluorescence-activated cell sorting. *J Exp Mar Biol Ecol* **329**: 196–205.
- Allard, P.-M., Péresse, T., Bisson, J., Gindro, K., Marcourt, L., van Pham, C., *et al.* (2016) Integration of molecular networking and in-silico MS/MS fragmentation for natural products dereplication. *Anal Chem* **88**: 3317–3323.
- Allen, F., Greiner, R., and Wishart, D. (2015) Competitive fragmentation modeling of ESI-MS/MS spectra for putative metabolite identification. *Metabolomics* **11**: 98–110.
- Baret, J.-C., Miller, O.J., Taly, V., Ryckelynck, M., El-Harrak, A., Frenz, L., *et al.* (2009) Fluorescence-activated droplet sorting (FADS): efficient microfluidic cell sorting based on enzymatic activity. *Lab Chip* **9**: 1850.
- Barra Caracciolo, A., Grenni, P., Cupo, C., and Rossetti, S. (2005) In situ analysis of native microbial communities in complex samples with high particulate loads. *FEMS Microbiol Lett* **253**: 55–58.
- Belknap, K.C., Park, C.J., Barth, B.M., and Andam, C.P. (2020) Genome mining of biosynthetic and chemotherapeutic gene clusters in *Streptomyces* bacteria. *Sci Rep* **10**: 1–9.
- Biener, G., Masson-Meyers, D.S., Bumah, V.V., Hussey, G., Stoneman, M.R., Enwemeka, C.S., and Raicu, V. (2017) Blue/violet laser inactivates methicillin-resistant *Staphylococcus aureus* by altering its transmembrane potential. *J Photochem Photobiol B Biol* **170**: 118–124.
- Boedicker, J.Q., Vincent, M.E., and Ismagilov, R.F. (2009) Microfluidic confinement of single cells of bacteria in small volumes initiates high-density behavior of quorum sensing and growth and reveals its variability. *Angew Chem Int Ed Engl* **48**: 5908–5911.
- Boitard, L., Cottinet, D., Bremond, N., Baudry, J., and Bibette, J. (2015) Growing microbes in millifluidic droplets. *Eng Life Sci* **15**: 318–326.
- Bon, R.S., and Waldmann, H. (2010) Bioactivity-guided navigation of chemical space. *Acc Chem Res* **43**: 1103–1114.
- Butler, M.S., and Paterson, D.L. (2020) Antibiotics in the clinical pipeline in October 2019. *J Antibiot* **73**: 329–364.
- Caballero-Aguilar, L.M., Duchib, S., Quigley, A., Onofriolob, C., Di Bellab, C., and Moultona, S.E. (2021) Microencapsulation of growth factors by microfluidic system. *MethodsX* **8**: 101324.
- Collins, D.J., Neild, A., deMello, A., Liu, A.Q., and Ai, Y. (2015) The Poisson distribution and beyond: methods for microfluidic droplet production and single cell encapsulation. *Lab Chip* **15**: 3439–3459.
- Crits-Christoph, A., Diamond, S., Butterfield, C.N., Thomas, B.C., and Banfield, J.F. (2018) Novel soil bacteria possess diverse genes for secondary metabolite biosynthesis. *Nature* **558**: 440–444.
- Damsté, J.S.S., Rijpstra, W.I.C., Dedysh, S.N., Foesel, B.U., and Villanueva, L. (2017) Pheno- and Genotyping of Hopanoid Production in Acidobacteria. *Front Microbiol* **8**: 968.
- Daniel, W.W. (1990) *Applied nonparametric statistics*. PWS-KENT Pub. University of Michigan. URL <https://books.google.de/books?id=0hPvAAAAMAAJ>.
- Delgado-Baquerizo, M., Maestre, F.T., Reich, P.B., Jeffries, T.C., Gaitan, J.J., Encinar, D., *et al.* (2016) Microbial diversity drives multifunctionality in terrestrial ecosystems. *Nat Commun* **7**: 10541.
- Duetz, W.A., Rüedi, L., Hermann, R., O'Connor, K., Büchs, J., and Witholt, B. (2000) Methods for intense aeration, growth, storage, and replication of bacterial strains in microtiter plates. *Appl Environ Microbiol* **66**: 2641–2646.
- Dunn, O.J. (1964) Multiple comparisons using rank sums. *Technometrics* **6**: 241–252.
- Dwivedi, D., Jansen, R., Molinari, G., Nimtz, M., Johri, B.N., and Wray, V. (2008) Antimycobacterial serratamolides and diacyl peptoglycosamine derivatives from *Serratia* sp. *J Nat Prod* **71**: 637–641.
- Firm, R.D., and Jones, C.G. (2003) Natural products - a simple model to explain chemical diversity. *Nat Prod Rep* **20**: 382–391.
- Fox, J.L. (2014) Fraunhofer to mine Sanofi microbial collection. *Nat Biotechnol* **32**: 305.
- Gerard, J., Lloyd, R., Barsby, T., Haden, P., Kelly, M.T., and Andersen, R.J. (1997) Massetolides A-H, antimycobacterial cyclic depsipeptides produced by two Pseudomonads isolated from marine habitats. *J Nat Prod* **60**: 223–229.

## Microfluidics-FACS bioprospecting platform 427

- Gross, H., and Loper, J.E. (2009) Genomics of secondary metabolite production by *Pseudomonas* spp. *Nat Prod Rep* **26**: 1408.
- Hammer, Ø., Harper, D.A.T., and Ryan, P.D. (2001) PAST: Paleontological statistics software package for education and data analysis. *Palaeontol Electron* **4**: 1–9.
- Hartwig, C. (2020) christoph-hartwig-ime-br/cosine-V2: cosine-v2. *Zenodo*.
- Hoffmann, T., Krug, D., Bozkurt, N., Duddela, S., Jansen, R., Garcia, R., *et al.* (2018) Correlating chemical diversity with taxonomic distance for discovery of natural products in myxobacteria. *Nat Commun* **9**: 803.
- Hong, J. (2011) Role of natural product diversity in chemical biology. *Curr Opin Chem Biol* **15**: 350–354.
- Imai, Y.u., Meyer, K.J., Iinishi, A., Favre-Godal, Q., Green, R., Manuse, S., *et al.* (2019) A new antibiotic selectively kills Gram-negative pathogens. *Nature* **576**: 459–464.
- Ishii, S., Tago, K., and Senoo, K. (2010) Single-cell analysis and isolation for microbiology and biotechnology: methods and applications. *Appl Microbiol Biotechnol* **86**: 1281–1292.
- Kaeberlein, T. (2002) Isolating “Uncultivable” microorganisms in pure culture in a simulated natural environment. *Science* **296**: 1127–1129.
- Kämpfer, P., Dott, W., Martin, K., and Glaeser, S.P. (2014) *Rhodococcus defluvi* sp. nov., isolated from wastewater of a bioreactor and formal proposal to reclassify [*Corynebacterium hoagii*] and *Rhodococcus equi* as *Rhodococcus hoagii* comb. nov. *Int J Syst Evol Microbiol* **64**(Pt 3): 755–761.
- Keller, M., and Zengler, K. (2004) Tapping into microbial diversity. *Nat Rev Microbiol* **2**: 141–150.
- Kielak, A.M., Barreto, C.C., Kowalchuk, G.A., van Veen, J.A., and Kuramae, E.E. (2016) The ecology of Acidobacteria: moving beyond genes and genomes. *Front Microbiol* **7**: 744.
- Koch, I.H., Gich, F., Dunfield, P.F., and Overmann, J. (2008) *Edaphobacter modestus* gen. nov., sp. nov., and *Edaphobacter aggregans* sp. nov., acidobacteria isolated from alpine and forest soils. *Int J Syst Evol Microbiol* **58**: 1114–1122.
- Koeth, T., Versalovic, J., and Lupski, J.R. (1995) Differential subsequence conservation of interspersed repetitive *Streptococcus pneumoniae* BOX elements in diverse bacteria. *Genome Res* **5**: 408–418.
- Kumar, S., Stecher, G., and Tamura, K. (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* **33**: 1870–1874.
- Laatsch, H. (2017) *AntiBase: The Natural Compound Identifier*. Weinheim, Germany: Wiley-VCH.
- Lakemeyer, M., Zhao, W., Mandl, F.A., Hammann, P., and Sieber, S.A. (2018) Über bisherige Denkweisen hinaus - neue Wirkstoffe zur Überwindung der Antibiotika-Krise. *Angew Chem* **130**: 14642–14682.
- Leman, M., Abouakil, F., Griffiths, A.D., and Tabeling, P. (2015) Droplet-based microfluidics at the femtolitre scale. *Lab Chip* **15**: 753–765.
- Letunic, I., and Bork, P. (2019) Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* **47**: W256–W259.
- Lewis, K. (2013) Platforms for antibiotic discovery. *Nat Rev Drug Discov* **12**: 371–387.
- Ling, L.L., Schneider, T., Peoples, A.J., Spoering, A.L., Engels, I., Conlon, B.P., *et al.* (2015) A new antibiotic kills pathogens without detectable resistance. *Nature* **517**: 455–459.
- Lok, C. (2015) Mining the microbial dark matter. *Nature* **522**: 270–273.
- Mahler, L., Tovar, M., Weber, T., Brandes, S., Rudolph, M.M., Ehgartner, J., *et al.* (2015) Enhanced and homogeneous oxygen availability during incubation of microfluidic droplets. *RSC Adv* **5**: 101871–101878.
- Marner, M., Patras, M.A., Kurz, M., Zubeil, F., Förster, F., Schuler, S., *et al.* (2020) Molecular networking-guided discovery and characterization of stechlisins, a group of cyclic lipopeptides from a *Pseudomonas* sp. *J Nat Prod* **83**: 2607–2617.
- Medema, M.H., Cimermancic, P., Sali, A., Takano, E., and Fischbach, M.A. (2014) A systematic computational analysis of biosynthetic gene cluster evolution: lessons for engineering biosynthesis. *PLoS Comput Biol* **10**: e1004016.
- Monciardini, P., Iorio, M., Maffioli, S., Sosio, M., and Donadio, S. (2014) Discovering new bioactive molecules from microbial sources. *Microb Biotechnol* **7**: 209–220.
- Mulligan, C.N. (2005) Environmental applications for biosurfactants. *Environ Pollut* **133**: 183–198.
- Newman, D.J., and Cragg, G.M. (2020) Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *J Nat Prod* **83**: 770–803.
- Nge, P.N., Rogers, C.I., and Woolley, A.T. (2013) Advances in microfluidic materials, functions, integration, and applications. *Chem Rev* **113**: 2550–2583.
- Nguyen, T.M., Seo, C., Ji, M., Paik, M.-J., Myung, S.-W., and Kim, J. (2018) Effective soil extraction method for cultivating previously uncultured soil bacteria. *Appl Environ Microbiol* **84**: e01145-18.
- Nicault, M., Tidjani, A.-R., Gauthier, A., Dumarcay, S., Gelhaye, E., Bontemps, C., and Leblond, P. (2020) Mining the biosynthetic potential for specialized metabolism of a *Streptomyces* soil community. *Antibiotics* **9**: 271.
- Oberpaul, M., Zumkeller, C.M., Culver, T., Spohn, M., Mihajlovic, S., Leis, B., *et al.* (2020) High-throughput cultivation for the selective isolation of Acidobacteria from termite nests. *Front Microbiol* **11**: 597628.
- Overmann, J., Abt, B., and Sikorski, J. (2017) Present and future of culturing bacteria. *Annu Rev Microbiol* **71**: 711–730.
- Owen, W.J., Yao, C., Myung, K., Kemmitt, G., Leader, A., Meyer, K.G., *et al.* (2017) Biological characterization of fenpicoxamid, a new fungicide with utility in cereals and other crops. *Pest Manag Sci* **73**: 2005–2016.
- Panthee, S., Hamamoto, H., Paudel, A., and Sekimizu, K. (2016) *Lysobacter* species: a potential source of novel antibiotics. *Arch Microbiol* **198**: 839–845.
- Panzer, K., Yilmaz, P., Weiß, M., Reich, L., Richter, M., Wiese, J., *et al.* (2015) Identification of habitat-specific biomes of aquatic fungal communities using a comprehensive nearly full-length 18S rRNA dataset enriched with contextual data. *PLoS One* **10**: e0134377.

428 M. Oberpaul *et al.*

- Parte, A.C. (2018) LPSN - List of Prokaryotic names with Standing in Nomenclature (bacterio.net), 20 years on. *Int J Syst Evol Microbiol* **68**: 1825–1829.
- Payne, E.M., Holland-Moritz, D.A., Sun, S., and Kennedy, R.T. (2020) High-throughput screening by droplet microfluidics: perspective into key challenges and future prospects. *Lab Chip* **20**: 2247–2262.
- Phillies, G. (1975) Nonactin, monactin, dinactin, trinactin, and tetranactin. A Raman Spectroscopic Study. *Biopolymers* **14**: 2311–2327.
- Pöppel, A.-K., Koch, A., Kogel, K.-H., Vogel, H., Kollwe, C., Wiesner, J., and Vilcinskis, A. (2014) Lucimycin, an antifungal peptide from the therapeutic maggot of the common green bottle fly *Lucilia sericata*. *Biol Chem* **395**: 649–656.
- R Core Team (2020) *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Řezanka, T., Prell, A., Spížek, J., and Sigler, K. (2010) Pilot-plant cultivation of *Streptomyces griseus* producing homologues of nonactin by precursor-directed biosynthesis and their identification by LC/MS-ESI. *J Antibiot* **63**: 524–529.
- Riyanti, Mamer, M., Hartwig, C., Patras, M., Wodi, S., Rieuwpassa, F., *et al.* (2020) Sustainable low-volume analysis of environmental samples by Semi-Automated Prioritization of Extracts for Natural Product Research (SeaPEPR). *Mar Drugs* **18**: 649.
- Schäberle, T.F., and Hack, I.M. (2014) Overcoming the current deadlock in antibiotic research. *Trends Microbiol* **22**: 165–167.
- Shannon, P. (2003) Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**: 2498–2504.
- Shapiro, H.M. (2000) Membrane potential estimation by flow cytometry. *Methods* **21**: 271–279.
- Stewart, E.J. (2012) Growing unculturable bacteria. *J Bacteriol* **194**: 4151–4160.
- Strobel, G.A., Morrison, S.L., and Cassella, M. (2005) Protecting plants from oomycete pathogens by treatment with compositions containing serratomolide and oocycin A from *Serratia marcescens*. Patent No.: US6926892B2.
- Taconelli, E., Carrara, E., Savoldi, A., Harbarth, S., Mendelson, M., Monnet, D.L., *et al.* (2018) Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. *Lancet Infect Dis* **18**: 318–327.
- Tamehiro, N., Okamoto-Hosoya, Y., Okamoto, S., Ubukata, M., Hamada, M., Naganawa, H., and Ochi, K. (2002) Bacilysoicin, a Novel phospholipid antibiotic produced by *Bacillus subtilis* 168. *Antimicrob Agents Chemother* **46**: 315–320.
- Terekhov, S.S., Smirnov, I.V., Stepanova, A.V., Bobik, T.V., Mokrushina, Y.A., Ponomarenko, N.A., *et al.* (2017) Microfluidic droplet platform for ultrahigh-throughput single-cell screening of biodiversity. *Proc Natl Acad Sci USA* **114**: 2550–2555.
- Theberge, A.B., Courtois, F., Schaerli, Y., Fischlechner, M., Abell, C., Hollfelder, F., and Huck, W.T. (2010) Microdroplets in microfluidics: an evolving platform for discoveries in chemistry and biology. *Angew Chem Int Ed Engl* **49**: 5846–5868.
- Theuretzbacher, U., Outtersson, K., Engel, A., and Karlén, A. (2020) The global preclinical antibacterial pipeline. *Nat Rev Microbiol* **18**: 275–285.
- Thies, S., Santiago-Schübel, B., Kovačić, F., Rosenau, F., Hausmann, R., and Jaeger, K.-E. (2014) Heterologous production of the lipopeptide biosurfactant serrawettin W1 in *Escherichia coli*. *J Biotechnol* **181**: 27–30.
- Tovar, M., Mahler, L., Buchheim, S., Roth, M., and Rosenbaum, M.A. (2020) Monitoring and external control of pH in microfluidic droplets during microbial culturing. *Microb Cell Fact* **19**: 16.
- Tracanna, V., de Jong, A., Medema, M.H., and Kuipers, O.P. (2017) Mining prokaryotes for antimicrobial compounds: from diversity to function. *FEMS Microbiol Rev* **41**: 417–429.
- Umetsu, N., and Shirai, Y. (2020) Development of novel pesticides in the 21st century. *J Pestic Sci* **45**: 54–74.
- Volpatti, L.R., and Yetisen, A.K. (2014) Commercialization of microfluidic devices. *Trends Biotechnol* **32**: 347–350.
- Ward, J.H. (1963) Hierarchical grouping to optimize an objective function. *J Am Stat Assoc* **58**: 236–244.
- Wasserman, H.H., Keggi, J.J., and McKeon, J.E. (1961) Serratamolide, a metabolic product of *Serratia*. *J Am Chem Soc* **83**: 4107–4108.
- Wiegand, S., Jogler, M., Boedeker, C., Pinto, D., Vollmers, J., Rivas-Marin, E., *et al.* (2019) Cultivation and functional characterization of 79 planctomycetes uncovers their unique biology. *Nat Microbiol* **5**: 126–140.
- Yang, J.Y., Sanchez, L.M., Rath, C.M., Liu, X., Boudreau, P.D., Bruns, N., *et al.* (2013) Molecular networking as a dereplication strategy. *J Nat Prod* **76**: 1686–1699.
- Zengler, K., Toledo, G., Rappe, M., Elkins, J., Mathur, E.J., Short, J.M., and Keller, M. (2002) Cultivating the uncultured. *Proc Natl Acad Sci USA* **99**: 15681–15686.
- Zinchenko, A., Devenish, S.R., Kintsjes, B., Colin, P.Y., Fischlechner, M., and Hollfelder, F. (2014) One in a million: flow cytometric sorting of single cell-lysate assays in monodisperse picolitre double emulsion droplets for directed evolution. *Anal Chem* **86**: 2526–2533.
- Zizika, Z. (1998) Biological effects of macrotretolide antibiotics and nonactin acids. *Folia Microbiol* **43**: 7–14.

### Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Fig. S1.** Poisson distribution of  $\lambda 0.1$  and  $\lambda 10$  showing the probability of the amount of cells encapsulated per droplet.

**Fig. S2.** Work scheme of this study showing the microfluidics workflow, the subsequent cultivation, screening and dereplication

**Fig. S3.** Live/dead staining using the LIVE/DEAD BacLight Bacterial Viability and Counting Kit (L7007, Invitrogen). Manufacturer's protocol was applied on the cells retrieved by nycodenz density gradient centrifugation (A: SYTO 9, B: propidium iodide, C: merged) directly after bacterial isolation. Exemplary pictures are shown, ten independent stains were done and considered for the calculation. Live:dead ratio was estimated resulting in  $\sim 70:30 \pm 6.7\%$ .

**Fig. S4.** Phylogenetic classification of FHG110511 within the phylum Acidobacteria clustering into subgroup 1. The tree is based on a ClustalW alignment of available 16S rRNA gene sequences from the ref\_seq database between positions 113 and 1357 [based on *Escherichia coli* 16S rRNA gene numbering (Brosius et al., 1978)] from the most similar sequences to the isolated strains, and also includes representatives of Acidobacteria subgroups 1, 3, 4, 6, 7, 8, and 10. The tree was calculated using MEGA v7.0.26 with the maximum-likelihood method and GTR-Gamma model. Circles on the tree branches indicate values of 1000 bootstrap replicates with a bootstrap support of more than 50%. Subgroup affiliations are indicated by colors. The new isolate is indicated by a black arrow. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site.

**Fig. S5.** (A) Assay read-out of  $\mu$ -fractionation plates of strain FHG110488 against *M. smegmatis* ATCC 607. Fractions are numbered and those causing at least 70% rel. growth inhibition were considered “active” and marked red. Column 1: medium control; Column 2+3: antibiotic standard (isoniazid); Column 4: growth control. Area AH05-AH24 top: 5  $\mu$ l injection volume; Area AH05-AH24 bottom: 10  $\mu$ l injection; Crude: crude extract as a control. (B) Overlaid Base peak Chromatogram (grey), fraction collector analog signal (light blue bars) and extracted ion chromatogram of  $m/z$  515.3329 $\pm$ 0.005 [M+H]<sup>+</sup> (**1**, red) of the 50 fold concentrated extract (in MeOH) with 5  $\mu$ l injection volume. (C) UV and MS spectrum of fractions 84-87. (D) MS/MS fragmentation of the precursor ion at  $m/z$  515.3330 [M+H]<sup>+</sup> (dereplicated as Serratamolide A, displayed in red) with manual annotation of the neutral losses.

**Fig. S6.** (A) MS2-network of “active” extract of FHG110488 against *Septoria tritici* MUCL45407 focusing on the cluster representing all seven detected serratamolide derivatives and their literature known structures (dots of parent ions found as hits in our internal database or AntiBase are marked in gold). (B) Overlaid Base peak Chromatogram (grey) and extracted ion chromatograms of serratamolides **1–7** (**1**  $m/z$  515.3327 [M+H]<sup>+</sup>, C<sub>26</sub>H<sub>47</sub>N<sub>2</sub>O<sub>8</sub><sup>+</sup> (red); **2**  $m/z$  541.3483 [M+H]<sup>+</sup>, C<sub>28</sub>H<sub>49</sub>N<sub>2</sub>O<sub>8</sub><sup>+</sup> (blue); **3**  $m/z$  543.3640 [M+H]<sup>+</sup>, C<sub>28</sub>H<sub>51</sub>N<sub>2</sub>O<sub>8</sub><sup>+</sup> (black); **4**  $m/z$  533.3433 [M+H]<sup>+</sup>, C<sub>26</sub>H<sub>49</sub>N<sub>2</sub>O<sub>9</sub><sup>+</sup> (cyan); **5**  $m/z$  547.3598 [M+H]<sup>+</sup>, C<sub>27</sub>H<sub>51</sub>N<sub>2</sub>O<sub>9</sub><sup>+</sup> (yellow); **6**  $m/z$  575.3902 [M+H]<sup>+</sup>, C<sub>29</sub>H<sub>55</sub>N<sub>2</sub>O<sub>9</sub><sup>+</sup> (light green); **7**  $m/z$  573.3746 [M+H]<sup>+</sup>, C<sub>29</sub>H<sub>53</sub>N<sub>2</sub>O<sub>9</sub><sup>+</sup> (dark green)) of the 50 fold concentrated extract (in MeOH) with 5  $\mu$ l injection volume. (C) MS/MS fragmentation of the precursor ions **1-7**.

**Fig. S7.** (A) Assay read-out of  $\mu$ -fractionation plate of strain FHG110502 against *Mycobacterium smegmatis* ATCC 607. Fractions are numbered and those causing at least 70% rel. growth inhibition were considered “active” and marked red. Column 1: medium control; Column 2+3: antibiotic standard (isoniazid); Column 4: growth control. Area AH05-AH24: 2  $\mu$ l injection volume; Area IP05-IP24: 5  $\mu$ l injection; Crude: crude extract as a control. (B) Overlaid Base peak Chromatograms (grey), Fraction collector analog signals (light blue bars) and extracted ion chromatograms of  $m/z$  1112.6814 $\pm$ 0.005 [M+H]<sup>+</sup> (**8**, red) with corresponding  $m/z$  556.8446 $\pm$ 0.005 [M+2H]<sup>2+</sup> (green),  $m/z$  1126.6973 $\pm$ 0.005 [M+H]<sup>+</sup> (**9**, yellow) with corresponding  $m/z$  563.8524  $\pm$  0.005 [M+2H]<sup>2+</sup> (blue), and  $m/z$  1154.7288 $\pm$ 0.005 [M+H]<sup>+</sup> (**10**,

purple) with corresponding  $m/z$  577.8680 $\pm$ 0.005 [M+2H]<sup>2+</sup> (magenta) of the 50 fold concentrated extract (in MeOH) with 5  $\mu$ l injection volume. (C) UV and MS spectrum of fractions 105–106 (left), 108–118 (middle) and 116 (right). (D) MS/MS fragmentation of the precursor ion at  $m/z$  1112.6814 [M+H]<sup>+</sup>,  $m/z$  1126.6973 [M+H]<sup>+</sup>, and  $m/z$  1154.7288 [M+H]<sup>+</sup> (dereplicated as massetolide E, massetolide F and massetolide H, respectively), manual annotation of the neutral losses and proposed structures of the fragment ions at  $m/z$  284.2229 and  $m/z$  312.2533. (E) Structures of all three dereplicated compounds **8–10**.

**Fig. S8.** (A) Assay read-out of  $\mu$ -fractionation plate of strain FHG110502 against *Mycobacterium smegmatis* ATCC 607. Fractions are numbered and those causing at least 70% rel. growth inhibition were considered “active” and marked red. Column 1: medium control; Column 2+3: antibiotic standard (isoniazid); Column 4: growth control. Area AH05-AH24: 2  $\mu$ l injection volume; Area IP05-IP24: 5  $\mu$ l injection; Crude: crude extract as a control. (B) Overlaid Base peak Chromatograms (grey), Fraction collector analog signals (light blue bars) and extracted ion chromatograms of  $m/z$  454.2931  $\pm$  0.005 [M+H]<sup>+</sup> (**11**, red),  $m/z$  690.5073  $\pm$  0.005 [M+H]<sup>+</sup> (**12**, blue) and  $m/z$  716.5237  $\pm$  0.005 [M+H]<sup>+</sup> (**13**, green) of the 50 fold concentrated extract (in MeOH) with 5  $\mu$ l injection volume. (C) UV and MS spectrum of fractions 93–94 (left) and 130–136 (right). (D) MS/MS fragmentation of the precursor ion at  $m/z$  454.2931 [M+H]<sup>+</sup>,  $m/z$  690.5073 [M+H]<sup>+</sup>, and  $m/z$  716.5237 [M+H]<sup>+</sup> (dereplicated as lyso-palmitoyl-phosphoethanolamine, palmitoleoyl-palmitoyl-phosphoethanolamine and palmitoleoyl-oleoyl-phosphoethanolamine, respectively), manual annotation of the neutral losses and proposed structures of the fragment ions. (E) Putative structures of all three dereplicated compounds **11–13**.

**Fig. S9.** (A) Assay read-out of  $\mu$ -fractionation plate of strain FHG110508 against *Staphylococcus aureus* ATCC 25923. Fractions are numbered and those causing at least 70% rel. growth inhibition were considered “active” and marked red. Column 1: medium control; Column 2+3: antibiotic standard (gentamycin); Column 4: growth control. Area AH05-AH24: 2  $\mu$ l injection volume; Area IP05-IP24: 5  $\mu$ l injection; Crude: crude extract as a control. (B) Overlaid Base peak Chromatograms (grey), Fraction collector analog signals (light blue bars) and extracted ion chromatograms of  $m/z$  737.4475 $\pm$ 0.005 [M+H]<sup>+</sup> (**14**, red),  $m/z$  751.4636 $\pm$ 0.005 [M+H]<sup>+</sup> (**15**, green),  $m/z$  765.4794 $\pm$ 0.005 [M+H]<sup>+</sup> (**16**, blue), and  $m/z$  779.4956 $\pm$ 0.005 [M+H]<sup>+</sup> (**17**, yellow) of the 50 fold concentrated extract (in MeOH) with 5  $\mu$ l injection volume. (C) UV and MS spectrum of fractions 120, 124, 128 and 131-132 (from left to right). (D) MS/MS fragmentation of the precursor ion at  $m/z$  737.4462 [M+H]<sup>+</sup>,  $m/z$  751.4618 [M+H]<sup>+</sup>,  $m/z$  765.4779 [M+H]<sup>+</sup>, and  $m/z$  779.4933 [M+H]<sup>+</sup> (dereplicated as nonactin, monactin, dinactin and macrotretolide G, respectively) with manual annotation of the neutral loss and proposed structure of the fragment ion at  $m/z$  213.1482 of parent ion at  $m/z$  779.4933 indicating the presence of macrotretolide G instead of trinactin. (E) Structures of all four dereplicated macrotretolides. Me: methyl; Et: ethyl; iPr: isopropyl.

**Fig. S10.** (A) MS2-network of “active” extract of FHG110508 against *Staphylococcus aureus* ATCC 25923 with focus on the cluster representing all five detected macrotretolide derivatives and their adduct ions (dots of parent ions found as hits in our internal database or AntiBase

430 M. Oberpaul *et al.*

are marked in gold). (B) Overlaid Base peak Chromatogram (grey) and extracted ion chromatograms of macrotetrolide **14–18** (**14**  $m/z$  737.4462  $[M+H]^+$ ,  $C_{40}H_{65}O_{12}^+$  (red); **15**  $m/z$  751.4618  $[M+H]^+$ ,  $C_{41}H_{67}O_{12}^+$  (green); **16**  $m/z$  765.4779  $[M+H]^+$ ,  $C_{42}H_{69}O_{12}^+$  (blue); **17**  $m/z$  779.4933  $[M+H]^+$ ,  $C_{43}H_{71}O_{12}^+$  (yellow); **18**  $m/z$  793.5104  $[M+H]^+$ ,  $C_{44}H_{73}O_{12}^+$  (black)) of the 50 fold concentrated extract (in MeOH) with 5  $\mu$ l injection volume. (C) MS/MS fragmentation of the precursor ions **14–18** dereplicated as nonactin, monactin, dinactin, macrotetrolide G, and macrotetrolide D, respectively, with manual annotation of the neutral loss and proposed structure of the fragment ion at  $m/z$  213.1482 of parent ion

at  $m/z$  779.4933 indicating the presents of macrotetrolide G instead of trinactin and fragment ion at  $m/z$  213.1487 of parent ion at  $m/z$  793.5104 indicating the presents of macrotetrolide D instead of tetranactin.

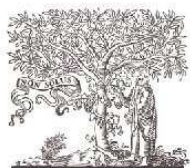
**Table S1.** Overview of the cultured genera in mISEM<sup>2</sup> pH 7.2, pH 5.5 and VL55-xyl pH 5.5. 16S rRNA gene sequencing was applied using the primer 1492R on mechanically disrupted culture broth grown for 7 days.

**Table S2.** Cosine Similarity table – Data for Fig. affiliation.

**Video S1.** Video showing the production of droplets with the  $\mu$ Encapsulator from Dolomite.

## Publication II

MethodsX 8 (2021) 101565



ELSEVIER

Contents lists available at ScienceDirect

MethodsX

journal homepage: [www.elsevier.com/locate/mex](http://www.elsevier.com/locate/mex)

Method Article

### Two-step generation of monodisperse agarose-solidified double emulsions (w/w/o) excluding an inner oil barrier



Stephan Brinkmann<sup>a</sup>, Markus Oberpaul<sup>a</sup>, Jens Glaeser<sup>a,b,\*</sup>,  
Till F. Schäberle<sup>a,c,\*</sup>

<sup>a</sup> Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Branch for Bioresources, 35392 Giessen, Germany

<sup>b</sup> Evotec International GmbH, Göttingen 37079, Germany

<sup>c</sup> Institute for Insect Biotechnology, Justus-Liebig-University of Giessen, Giessen 35392, Germany

**Copyright:** This article is Open Access (CC BY 4.0).

#### **Contribution:**

Stephan Brinkmann designed and developed the technical method to generate agarose-solidified double emulsions (w/w/o) excluding an inner oil barrier. Therefore, he planned and performed all the experiments. He established and validated this novel microfluidic application method in various experiments. He collected and analyzed the data and prepared the figures for the manuscript. He also wrote the draft of the manuscript as well as the following revisions.



Contents lists available at ScienceDirect

MethodsX

journal homepage: [www.elsevier.com/locate/mex](http://www.elsevier.com/locate/mex)

## Method Article

## Two-step generation of monodisperse agarose-solidified double emulsions (w/w/o) excluding an inner oil barrier



Stephan Brinkmann<sup>a</sup>, Markus Oberpaul<sup>a</sup>, Jens Glaeser<sup>a,b,\*</sup>,  
Till F. Schäberle<sup>a,c,\*</sup>

<sup>a</sup> Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Branch for Bioresources, 35392 Giessen, Germany

<sup>b</sup> Evotec International GmbH, Göttingen 37079, Germany

<sup>c</sup> Institute for Insect Biotechnology, Justus-Liebig-University of Giessen, Giessen 35392, Germany

## A B S T R A C T

Miniaturization of biomedical and chemical research areas is performed using microfluidic techniques. Droplet-based microfluidic applications are of high interest for various applications, e.g., high-throughput screening assays. Many of them are based on simple water-in-oil (w/o) or oil-in-water (o/w) emulsions that are easily to produce. More complex assays based on separate compartments require the use of multiple emulsions, such as water-in-oil-in-water (w/o/w) or oil-in-water-in-oil (o/w/o) emulsions. In this study an easy, fast to establish method to generate agarose-solidified (w/w/o) double emulsions with  $\sim 55 \mu\text{m}$  in diameter, in which both agarose-phases are not separated by a surfactant stabilized oil is described. An off-chip emulsion-breaking and washing step of the inner agarose droplets based on density gradient centrifugation was designed, offering new possibilities for high-throughput assays on picoliter scale. In brief, this paper reports:

- the protocol to generate agarose-solidified (w/w/o) double emulsions non-separated by surfactant stabilized oil;
- an off-chip washing protocol of agarose-solidified emulsions based on density gradient centrifugation.

© 2021 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

## A R T I C L E I N F O

**Method name:** Two-step generation of monodisperse agarose-solidified double emulsions (w/w/o) excluding an inner oil barrier  
**Keywords:** Droplet-based microfluidics, Water-in-water-in-oil emulsion, Microreactor, Microencapsulation, Microdroplet, High-throughput assay

**Article history:** Received 15 August 2021; Accepted 28 October 2021; Available online 2 November 2021

\* Corresponding authors at: Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Branch for Bioresources, 35392 Giessen, Germany.

E-mail addresses: [jens.glaeser@evotec.com](mailto:jens.glaeser@evotec.com) (J. Glaeser), [till.f.schaerberle@agr.uni-giessen.de](mailto:till.f.schaerberle@agr.uni-giessen.de) (T.F. Schäberle).

<https://doi.org/10.1016/j.mex.2021.101565>

2215-0161/© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

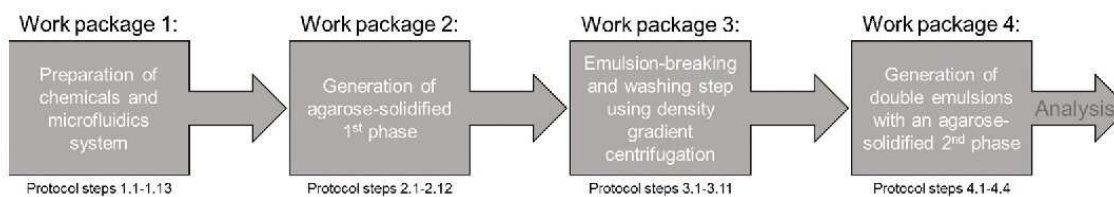
## Specifications table

Subject Area:	Materials Science
More specific subject area:	Microfluidic double emulsion
Method name:	Two-step generation of monodisperse agarose-solidified double emulsions (w/w/o) excluding an inner oil barrier
Name and reference of original method:	Combination of high-throughput Microfluidics and FACS technologies to leverage the numbers game in natural product discovery <a href="http://doi.org/10.1111/1751-7915.13872">http://doi.org/10.1111/1751-7915.13872</a>
Resource availability:	Microfluidic hardware/software: <a href="https://www.dolomite-microfluidics.com/microfluidic-systems/%C2%B5encapsulator/">https://www.dolomite-microfluidics.com/microfluidic-systems/%C2%B5encapsulator/</a>

## Background

The field of microfluidics set new standards in miniaturization of biomedical and chemical research areas. Droplet-based microfluidics is used as a tool for small scale single-cell or cell culture analysis, chemical synthesis as well as high-throughput screening [1]. Achieved by dispersions of stabilized liquids within continuous immiscible fluids, thousands of emulsions are generated within minutes, which subsequently are used as micro compartments suitable for such experiments [2]. Beside simple water-in-oil (w/o) or oil-in-water (o/w) emulsions, more complex ones, carrying single or multiple emulsions that represent more compartments, enable the build-up of systems with higher complexity. However, the latter are more difficult to generate. Widely used are water-in-oil-in-water (w/o/w) and oil-in-water-in-oil (o/w/o) emulsions, which can be employed for applications such as drug delivery vehicles, cell carriers, barcoding of droplets, microscale sensors, and more [3]. Other combinations of phases such as w/w and w/w/w are challenging, however applicable using aqueous phases of different properties, e.g., density, viscosity, and refractive index [4,5]. With millions of emulsions generated within minutes, powerful high-throughput analysis and sorting tools are required to retrieve events of interest. Optofluidic setups [6,7], fluorescence-activated droplet sorter (FADS) [8,9] or commercially available multichannel fluorescence-activated cell sorter (FACS) [10,11] are such tools. They differ in the analyzable emulsion-types, sample throughput and sorting capabilities, assay readout, and biocompatibility.

In this protocol (Fig. 1), we demonstrate the generation of agarose-solidified (w/w/o) double emulsions (diameter:  $\sim 55 \mu\text{m}$ , volume:  $\sim 87 \text{ pL}$ ) building on agarose-solidified (w/o) emulsions (first phase; diameter:  $\sim 40 \mu\text{m}$ , volume:  $\sim 33.5 \text{ pL}$ ) that are produced beforehand with the same microfluidic setup. The workflow includes an emulsion-breaking step of the first phase that results in the release of all compartments of the first phase into an aqueous phase. This is followed by a density-gradient washing step thereof to wash the first phase droplets and while working with e.g., cell cultures or bacteria to separate them from motile cells that did not get stuck within the agarose droplets. Subsequently, a second phase of liquid agarose is added. This procedure enables that both agarose phases are not compartmentalized by the fluorocarbon oil (Novec<sup>TM</sup> 7500), which is stabilized with a surfactant (Pico-Surf<sup>TM</sup> 1), necessary for emulsion stability. The emulsion-breaking step is essential, since the oil and surfactant are known to function as a barrier, which however is not completely understood until today, for several molecules with slow diffusion properties [12,13]. Moreover, with no diffusion barrier and the second phase to be added any time after incubation of the first phase, this method allows various applications for e.g., two-layer high-throughput cell culture



**Fig. 1.** Overview of the complete workflow to generate agarose-solidified (w/w/o) double emulsions.



and screening assays that rely on an intermediate washing step. This is supported by the use of a low-melting agarose that allows operations at low temperatures feasible for cell culture experiments. Applicable for e.g., fluorescence-activated cell sorting (FACS)-based technologies, a variable high-throughput readout of an envisaged assay, based on such double emulsions, is given. To the best of our knowledge, this is the first report of successfully generated agarose-solidified double emulsions not separated by an inner oil barrier. Designated to run on the commercially available microfluidic 'μEncapsulator System' (Dolomite Microfluidics, Royston, UK), trained microfluidic users can easily implement this method in any lab.

### Overview of the method

Fig. 1 depicts the general procedure to generate agarose-solidified (w/w/o) double emulsion using the commercially available microfluidic 'μEncapsulator System' (Dolomite Microfluidics), including solution and microfluidics system preparation (work package 1), generation of agarose-solidified (w/o) emulsion (first phase, work package 2), an emulsion-breaking and washing step of the first phase (work package 3), and the generation of the final agarose-solidified (w/w/o) double emulsions (work package 4). Work package 3 is critical for this application, especially while working with cell cultures. The emulsion of the solidified agarose and the fluorocarbon oil Novec™ 7500 stabilized with Pico-Surf™ 1 (first phase) needs to be de-emulsified, since otherwise in the end both agarose phases would be separated by a diffusion barrier not fully understood [12,13]. The emulsion-breaking results in release of the first phase content into the aqueous phase that did not get stuck within the agarose, e.g., still motile cells. As a solution to separate agarose-solidified droplets of the first phase from other compartments, we designed a washing step based on density gradient centrifugation using Nycodenz®.

### Materials

#### Chemicals

- Pico-Surf™ 1 (5% (w/w) in Novec™ 7500) (Sphere Fluidics, Cambridge, UK, prod. no. C022)
- Novec™ 7500 (Iolitec Ionic Liquids Technologies GmbH, Heilbronn, GER, prod. no. FL-0004-HP)
- Pico-Break™ 1 – Emulsion Breaking Solution (Sphere Fluidics, Cambridge, UK, prod. no. C081)
- SeaPrep agarose (Lonza, Basel, Switzerland, prod. no. 50,302)
- Nycodenz® (Axis-Shield Poc AS, Oslo, Norway, prod. no. 1,002,424)
- 2-Propanol (Honeywell International Inc., Morristown, New Jersey, US, prod. no. 34,965)

#### Equipment

- Temperature control unit, Meros TCU-100 (TCU, Dolomite Microfluidics, Royston, UK, prod. no. 3,200,428)
- μEncapsulator Top Interface (Dolomite Microfluidics, Royston, UK, prod. no. 3,200,569)
- Linear Connector (Dolomite Microfluidics, Royston, UK, prod. no. 3,000,024)
- μEncapsulator 1 Sample Reservoir Chip (Dolomite Microfluidics, Royston, UK, prod. no. 3,200,444)
- μEncapsulator 1 - 2 Reagent Droplet Chip (50 μm etch depth), fluorophilic, (Dolomite Microfluidics, Royston, UK, prod. no. 3,200,445)
- Pressure pumps, Mitos P-Pump (Dolomite Microfluidics, Royston, UK, prod. no. 3,200,176)
  - Mitos Flow Rate Sensor (Dolomite Microfluidics, Royston, UK, prod. no. 3,200,098, 3,200,099)
  - Mitos Sensor Interfaces (Dolomite Microfluidics, Royston, UK, prod. no. 3,200,200)
- Four-way connector (Dolomite Microfluidics, Royston, UK, prod. no. 3,200,454)
- Air compressor Fengda AS-189
- High-speed CMOS camera PL-D721CU (Navitar Inc., Rochester, NY, USA)
- Stereomicroscope Stemi SV 11 (Carl Zeiss, Oberkochen, Germany)

- Halogen light source KL 2500 LCD (Schott AG, Mainz, Germany)
- Fluorescence microscope DM2000 LED equipped with a DFC450 C camera (Leica Microsystems, Wetzlar, Germany)
- Centrifuge 5810 R (Eppendorf AG, Hamburg, Germany)
- Centrifuge 5424 R (Eppendorf AG, Hamburg, Germany)
- Fridge
- Neubauer chambers (0.1 mm depth, Paul Marienfeld GmbH & CO. KG, Lauda-Königshofen, Germany)
- Adjustable 10, 100 and 1000  $\mu$ L pipettes

#### Consumables

- 1 mL NORM-JECT syringes (Henke-Sass, Wolf GmbH, Tuttlingen, Germany prod. no. 4010.200V0)
- 30 mL syringes (Becton, Dickinson and Company, Franklin Lakes, New Jersey, US, prod. no. 309,650)
- Needles – Sterican® Gr. 14, G 23  $\times$  1 1/4" /  $\varnothing$  0,60  $\times$  30 mm, blue (B. Braun SE, Melsungen, Germany, prod. no. 4,657,640)
- 15 mL centrifuge tubes (Greiner Bio-One International GmbH, Frickenhausen, Germany, prod. no. 188,261)
- 0.2  $\mu$ m CA syringe filter (Corning Inc., Corning, NY, USA, prod. no. 431,224)
- 1.5 mL reaction tubes (SARSTEDT AG & Co. KG, Nümbrecht, Germany)
- Sterile pipette tips
- Microorganisms used as showcases in this study:
  - Environmental microorganisms isolated in our previous study [13]
  - *E. coli* DH5 $\alpha$  pFU95 (expressing *gfp<sub>mut3.1</sub>*)
  - *E. coli* DH5 $\alpha$  pFU96 (expressing *dsRed2*)

#### Software

- Dolomite Flow Control Center (Dolomite UK, prod. no. 3,200,475)
- Pixelink Capture Software (Navitar Inc., Rochester, NY, USA)
- Leica Application Suite v4.8.0 (Leica Microsystems CMS GmbH, Heerbrugg, Switzerland)

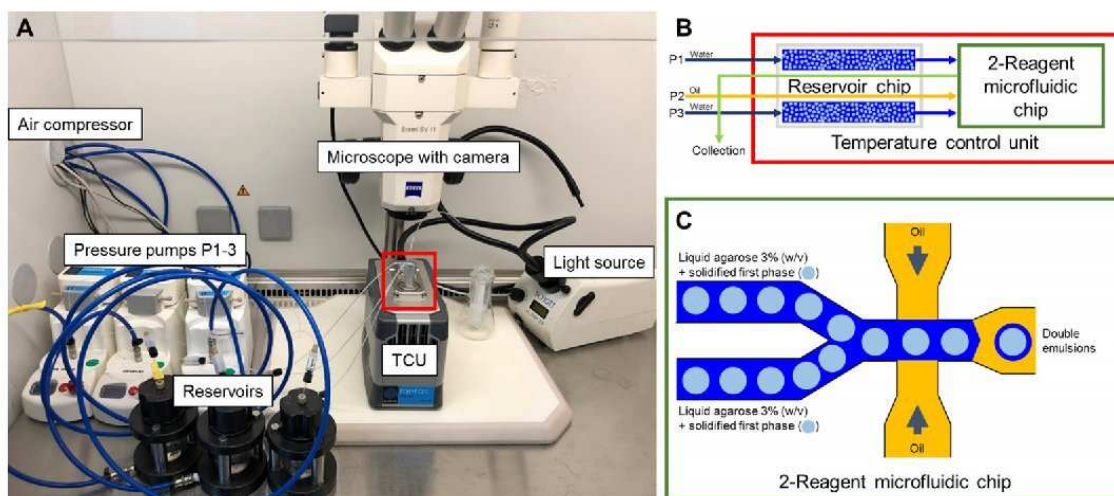
### Experimental protocol

In this section each work package and therein the individual tasks (protocol steps) are described step-wise and in detail.

#### 1. Work package – Preparation of chemicals and microfluidic system

The commercially available microfluidic ‘ $\mu$ Encapsulator System’ (Dolomite Microfluidics) is very user-friendly (Fig. 2). The assembly of all parts, the cleaning and sterilization of chips as well as the troubleshooting of blockages is described in the ‘ $\mu$ Encapsulator System User Manual’ (<https://www.dolomite-microfluidics.com/support/downloads/>). Prior to each experiment, prepare and perform the following chemicals and steps:

- 1.1 Prepare 20 mL Novec™ 7500 containing 1% (w/w) Pico-Surf™ 1 (mix 4 mL Pico-Surf™ 1 5% (w/w) in Novec™ 7500 with 16 mL Novec™ 7500).
- 1.2 Prepare 40 mL water.
- 1.3 Prepare 3% (w/v) SeaPrep agarose in water (3 g SeaPrep in 100 mL water) and autoclave. **NOTE** Add a magnetic stirrer because after autoclaving the agarose is not dissolved well. Stir it at 60 °C until a homogeneous solution is obtained.
- 1.4 Prepare 20% and 30% (w/v) Nycodenz® solution in water (10 g and 15 g Nycodenz® in 50 mL water each).
- 1.5 Prepare 70% isopropanol (HPLC grade).



**Fig. 2.** Overview of the hardware setup (a), the temperature control unit (b) and the 2-Reagent microfluidic chip (c) used for the generation of the agarose-solidified 1st and 2nd phase.

**Important** Filter all solvents through 0.2  $\mu\text{m}$  CA syringe filters to prevent blockage of microfluidic channels.

Microfluidic system (Fig. 2a):

- 1.6 Fill reservoirs with water (pump 1 and 3) and Novec™ 7500 with 1% (w/w) Pico-Surf™ 1 (pump 2).
- 1.7 Turn on the air compressor for pressure supply and operate it at 500–700 kPa.
- 1.8 Turn on all three pressure pumps, the TCU (temperature control unit), light source, and computer.
- 1.9 Open the Flow Control Center (FCC) Software (Dolomite Microfluidics) and Pixelink Capture Software (Navitar).

**IMPORTANT** The FCC software needs to be opened after the hardware is switched on to ensure connection.

- 1.10 Set the temperature of the TCU to 30 °C to prevent solidification of the low-melting agarose during encapsulation.
- 1.11 Priming the system (Fig. 2b) is done with the Linear Connector connected to the reservoir chip. Apply flow rates of 5  $\mu\text{L}/\text{min}$  on each channel for 5 min. This is necessary to prime all channels with their respective fluids and to push e.g., air bubbles out of the system.
- 1.12 Set the flow rates to 0  $\mu\text{L}/\text{min}$  and disconnect the Linear Connector. **NOTE** This ensures no backflow of fluids into their reservoirs without using valves.
- 1.13 Empty both sample reservoirs on the sample reservoir chip.

→ The system is now ready to be used for the generation of the first phase or for the generation of double-emulsions in the second step.

## 2. Work package – Generation of agarose-solidified 1st phase

We previously used the following protocol steps for encapsulation and cultivation of environmental microorganisms [13]. To obtain the here described agarose-solidified (w/w/o) double emulsions, it was necessary to use SeaPrep agarose instead of the previously used SeaPlaque one. Agarose-solidified (w/w/o) double emulsions were not possible to be obtained from a first phase produced with SeaPlaque agarose, probably due to different material properties.

2.1. Prepare the loading suspension to obtain a first phase of 1.5% (w/v) agarose droplets in a 1.5 mL reaction tube as following:

- 200  $\mu$ L 3% (w/v) SeaPrep in water
- 200  $\mu$ L water

**NOTE** Working with cell cultures, the water can be exchanged with the cells in their specific medium and other components. It is also possible to prepare the agarose with medium in advance to prevent dilution after mixing the liquid agarose and the water phase (cells + medium).

2.2 Vortex the solution for 10 s.

2.3 Load  $\sim$ 90  $\mu$ L into each reservoir of the reservoir chip.

2.4 Pipet 5  $\mu$ L of Novec<sup>TM</sup> 7500 post each loaded sample. **NOTE** This keeps the loaded sample separated from the water pushing it towards the 2-Reagent Droplet Chip (Fig. 2c).

2.5 Close the lid and connect the Linear Connector.

2.6 Set the flow rates for both pumps 1 and 3 supplying the loaded sample towards the 2-Reagent microfluidic chip to 2  $\mu$ L/min and the flow rate for the oil (pump 2) to 40  $\mu$ L/min. **NOTE** The pressure of the oil pump will be in the range of 350–550 mbar and of the water pumps between 550 and 800 mbar. A single pump should not exceed 1200 mbar, otherwise the channel of this pump is blocked. For the troubleshooting of blockages see the ' $\mu$ Encapsulator System User Manual' (<https://www.dolomite-microfluidics.com/support/downloads/>).

2.7 Wait 2–3 min for pressures to stabilize and then start to collect the generated droplets in a 1.5 mL reaction tube for about 25 min. **NOTE** A pressure drop indicates an emptied reservoir and therefore the end of the encapsulation.

2.8 Set the flow rate of all three pumps to 0  $\mu$ L/min.

2.9 Incubate the first phase at 8 °C for 20 min to solidify the liquid agarose-in-oil emulsions.

2.10 Disconnect the Linear Connector and empty both reservoirs of the reservoir chip.

2.11 Clean and sterilize the reservoir chip and the 2-Reagent microfluidic chip as following:

- To solve remaining residues pipet water up and down in each reservoir a few times followed by 70% isopropanol.
- Load both reservoirs with 70% isopropanol.
- Close the lid and connect the Linear Connector.
- Set all pumps to 7  $\mu$ L/min and run the system for 10 min.
- Load both reservoirs with water.
- Close the lid and connect the Linear Connector.
- Set all pumps to 7  $\mu$ L/min and run the system for 10 min.

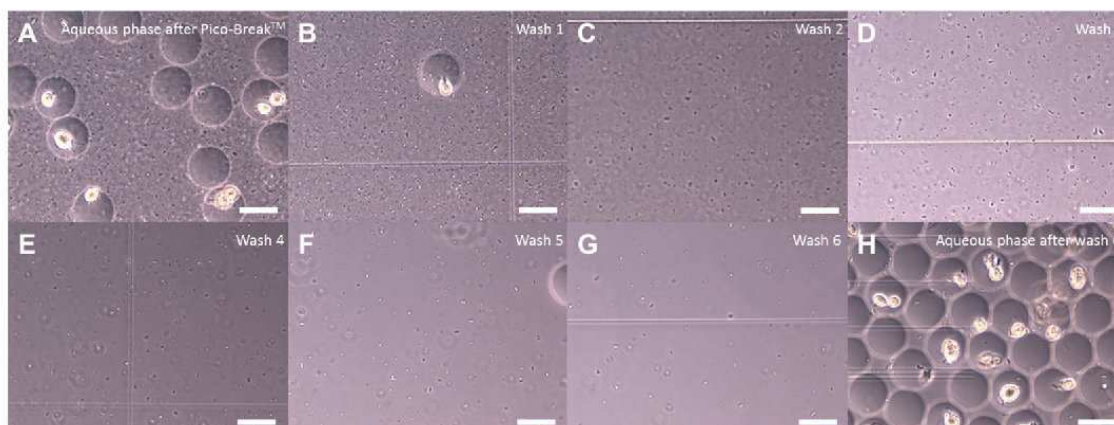
2.12 After cleaning and sterilization, stop the flow of all three pumps. **NOTE** For long term storage all channels are dried/emptied.

→ At this point the microfluidic system is on hold. A shut down can be done by disconnecting the power of all individual parts, releasing the pressure from the air compressor and shutting down the computer.

### 3. Work package – Emulsion-breaking and washing step using density gradient centrifugation

The generation of agarose-solidified double emulsions with both agarose phases not separated by the surfactant (Pico-Surf<sup>TM</sup> 1) stabilized fluorocarbon oil (Novec<sup>TM</sup> 7500) requires an emulsion-breaking step of the first phase. Otherwise, both agarose phases are oil-separated – a barrier probably unwanted for some high-throughput assays, since diffusion properties for several molecules are not completely understood today [12,13]. The emulsion-breaking step is performed using Pico-Break<sup>TM</sup> 1, according to the manufacturer's instruction with minor adaptations ([https://spherefluidics.com/wp-content/uploads/2019/03/Pico-Break\\_User-Guide-March-2019.pdf](https://spherefluidics.com/wp-content/uploads/2019/03/Pico-Break_User-Guide-March-2019.pdf)).

This results in non-oil/surfactant-separated agarose-solidified droplets in an aqueous phase. Separated environments are still preserved, e.g., for cell cultures, or as in our case bacteria, that have been encapsulated and got stuck in the solidified agarose of the first phase. However, while working



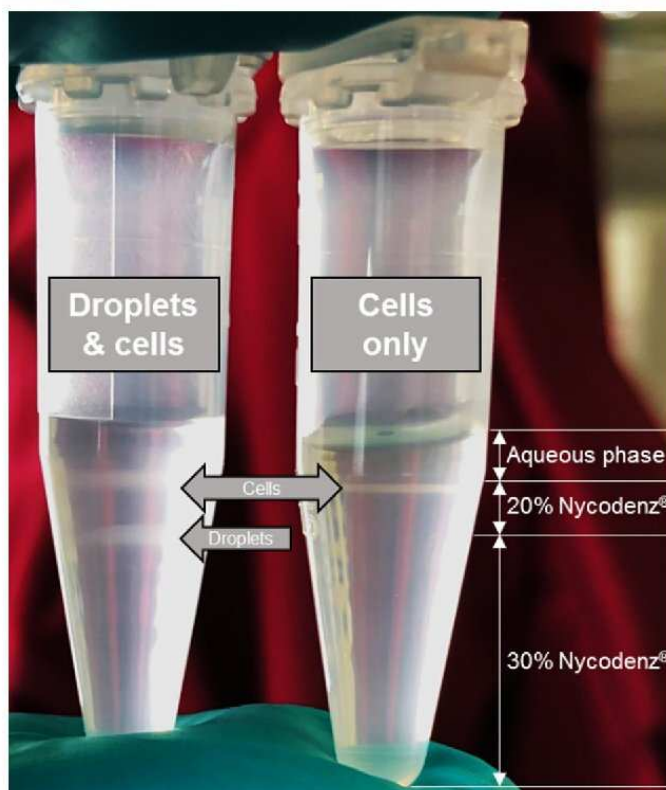
**Fig. 3.** Aqueous phase before (a), in-between (b-g) and after six washing steps consisting of centrifugation and supernatant exchange. Scale bar: 40  $\mu\text{m}$ .

with bacteria, we observed a huge problem of free swimming non-agarose-attached cells in the aqueous phase after emulsion-breaking (Fig. 3a). Applying this mixed aqueous phase to the second encapsulation to add the outer agarose layer, resulted in cross contamination of all solidified droplet-compartments by free swimming cells. A simple washing step by centrifugation and supernatant exchange resulted in a reduction of free swimming cells (Fig. 3b-g); however, even after six repetitions a few of them were observed in the aqueous phase (Fig. 3h). Therefore, a density gradient centrifugation protocol based on Nycodenz® (Figs. 4, 5a-i) was established to completely separate the free swimming cells (Fig. 5h) from the agarose-solidified droplets (Fig. 5i) containing non-motile cells with the following protocol steps:

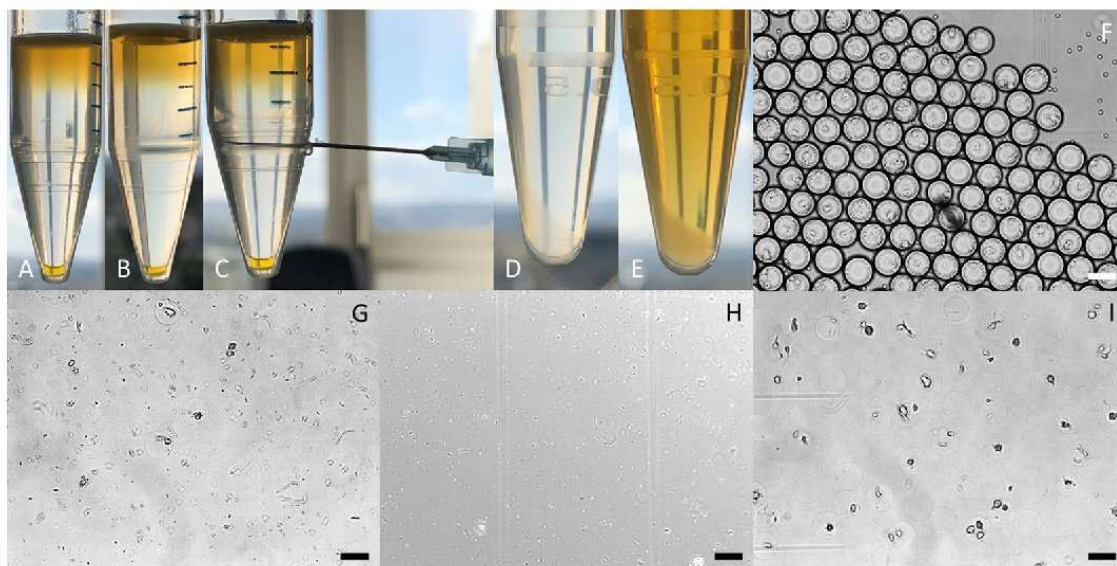
- 3.1 Take the 1.5 mL reaction tube containing the first phase (agarose-solidified droplets and oil, Fig. 5f) and remove as much of the Pico-Surf™ 1 oil (bottom layer) as possible using a standard pipet. **NOTE** This reduces the amount of Pico-Break™ 1 necessary to break the w/o-emulsion.
- 3.2 The droplet layer has a volume of approximately 100  $\mu\text{L}$ . Add 250  $\mu\text{L}$  Pico-Break™ 1 and gently agitate the mixture by inverting. **NOTE** The solution turns orange and starts to disperse.
- 3.3 Centrifuge the sample at  $1000 \times g$  for 1 min to disperse the aqueous (clear) and flurous (orange) phase.
- 3.4 Tilt 1.5 mL reaction tube to an angle of  $45^\circ$  and remove most of the orange flurous phase using a standard pipet.

Add 500  $\mu\text{L}$  water to the remaining solution in the 1.5 mL reaction tube and transfer it to a 15 mL centrifuge tube. **NOTE** While working with cell cultures, the water can be exchanged with specific medium and other components.

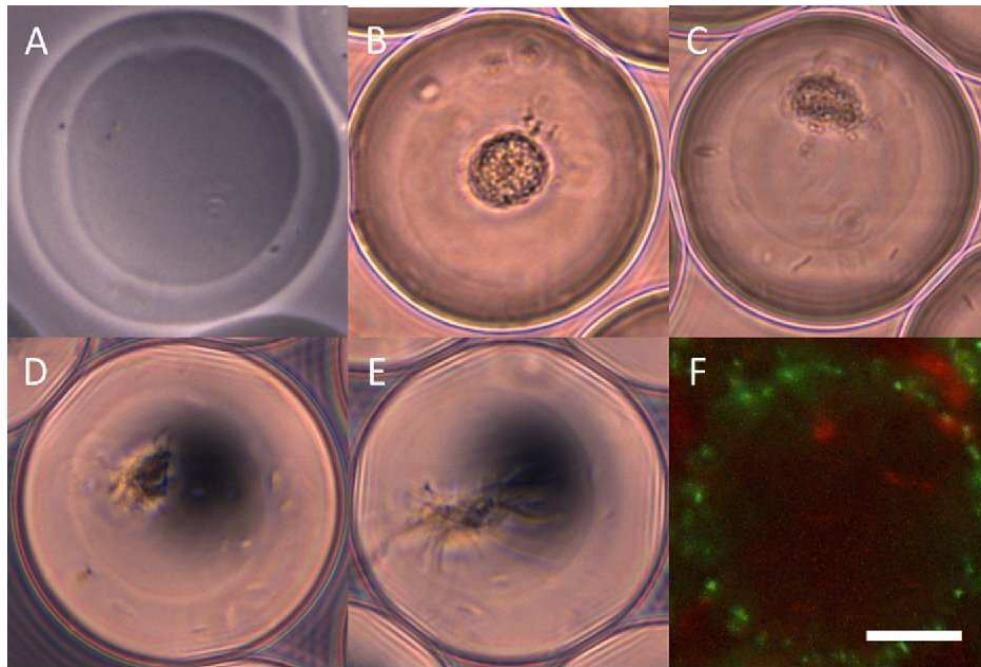
- 3.5 Carefully pipet 2 mL 20% Nycodenz® solution underneath this aqueous phase.
- 3.6 Carefully pipet 1 mL 30% Nycodenz® solution underneath the 20% Nycodenz® solution. Fig. 5a depicts how the solution should look like at this step.
- 3.7 Centrifuge at  $2000 \times g$  for 1 min. **IMPORTANT** Do not centrifuge at higher speed or elongate this step, otherwise the cell-layer will not be separated from the droplet-layer (Fig. 5b).
- 3.8 To extract the droplet-layer, place a needle on top of a 1 mL syringe, pierce slowly through the centrifuge tube beneath the droplet-layer, which is situated on top of the 30% Nycodenz® solution, and pull out the droplets (Fig. 5c).
- 3.9 Transfer the clear content to a fresh 1.5 mL reaction tube.
- 3.10 Centrifuge the solution at  $1000 \times g$  for 1 min and discard the supernatant until 200  $\mu\text{L}$  remain. **NOTE** The droplet-layer is hard to spot in a clear solution (Fig. 5d).
- 3.11 Wash the droplet-layer by adding 1000  $\mu\text{L}$  water (or medium), centrifugation at  $1000 \times g$  for 1 min (Fig. 5e) and discard the supernatant until 100  $\mu\text{L}$  are left.



**Fig. 4.** Successful separation of free-swimming cells and agarose-solidified droplets of the first phase using density gradient centrifugation based on Nycodenz®.



**Fig. 5.** Droplet-layer extraction after density gradient centrifugation (before (a) and after (b) centrifugation, pierced needle on droplet-layer level (c), milky droplet pellet in water (d), milky droplet pellet in medium (e)). Microscopic picture of the w/o emulsion of the first phase before (f) and after Pico-Break™ 1 treatment (g), the cell-layer (h) and the droplet-layer (i) after density gradient centrifugation. Scale bar: 40  $\mu\text{m}$ .



**Fig. 6.** Examples of agarose-solidified double emulsions with both phases empty (a), first phase carrying a microcolony of *E. coli* DH5 $\alpha$  derived from a single cell (b–e), and one with dsRed producing *E. coli* DH5 $\alpha$  pFU96 cells in the first phase and GFP producing *E. coli* DH5 $\alpha$  pFU95 cells in the outer layer (f). Scale Bar: 20  $\mu$ m.

→ At this point, the emulsion of the first phase is broken, the droplet-layer is washed and potentially free-swimming cells are discarded. The droplet-layer can now directly be used to be encapsulated a second time to add the second agarose layer.

#### 4. Work package – Generation of double emulsions with an agarose-solidified 2nd phase

This work package is highly similar to the second one, since the same system and chips are used. The main difference for the production of double emulsions is that a different flow rate on the oil pump (pump 2) is applied.

4.1 Prepare the loading suspension to obtain double emulsions in a 1.5 mL reaction tube as following:

- 100  $\mu$ L 3% (w/v) SeaPrep in water
- ~100  $\mu$ L droplet-layer of the first phase after work package 3 (oil-free and washed)

4.2 Perform the previously described protocol Steps 2.2 to 2.5.

4.3 Set the flow rates for both pumps 1 and 3 supplying the loaded sample towards the 2-Reagent microfluidic chip to 2  $\mu$ L/min and the flow rate for the oil (pump 2) to 15  $\mu$ L/min. **NOTE** The pressure of the oil pump will be in the range of 450–650 mbar and of the water pumps between 800 and 1100 mbar. A single pump should not exceed 1400 mbar, otherwise the channel of this pump is blocked. For the troubleshooting of blockages see the ‘ $\mu$ Encapsulator System User Manual’ (<https://www.dolomite-microfluidics.com/support/downloads/>).

4.4 Perform the previously described protocol Steps 2.7 to 2.12.

→ At this point, agarose-solidified double emulsions (w/w/o) without an inner oil barrier are generated and can be analyzed, e.g., using FACS, as previously described for agarose-solidified droplets of the first phase [13].

## Method validation

Some agarose-solidified double emulsion examples are shown in Fig. 6. They are ~55 µm in diameter with a volume of ~87 pL and they can be empty (Fig. 6a) or e.g., carry bacteria. In our case we did experiments with first phases carrying single *E. coli* DH5α cells, incubated for a few hours and engulfed with a second agarose layer (Fig. 6b–e). An encapsulation of dsRed producing *E. coli* DH5α pFU96 cells in the first phase and GFP producing *E. coli* DH5α pFU95 cells in the outer layer (Fig. 6f) shows how cells are distributed in both compartments. Usually the microfluidics setup and all chemicals are prepared within half an hour. A single encapsulation generating the first or second phase takes with preparation about 45 min. To save time between experiments, the emulsion-breaking and washing step on the next sample can be applied while the microfluidics system runs with another sample. Therefore, several encapsulations are possible on one day.

## Acknowledgments

This work was financially supported by the Hessen State Ministry of Higher Education, Research and the Arts (HMWK) via the state initiative for the development of scientific and economic excellence for the LOEWE Center for Insect Biotechnology and Bioresources. Sanofi-Aventis Deutschland GmbH and Evotec International GmbH contributed in the framework of the Sanofi-Fraunhofer Natural Products Center of Excellence/Fraunhofer-Evotec Natural Products Center of Excellence.

## Declaration of Competing Interest

J.G. is employed at Evotec International GmbH. All authors declare no competing interests.

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

- [1] L. Shang, Y. Cheng, Y. Zhao, Emerging droplet microfluidics, *Chem. Rev.* 117 (2017) 7964–8040, doi:[10.1021/acs.chemrev.6b00848](https://doi.org/10.1021/acs.chemrev.6b00848).
- [2] N. Convery, N. Gadegaard, 30 years of microfluidics, *Micro Nano Eng.* 2 (2019) 76–91, doi:[10.1016/j.mne.2019.01.003](https://doi.org/10.1016/j.mne.2019.01.003).
- [3] T.Y. Lee, T.M. Choi, T.S. Shim, R.A.M. Frijns, S.H. Kim, Microfluidic production of multiple emulsions and functional microcapsules, *Lab Chip* 16 (2016) 3415–3440, doi:[10.1039/c6lc00809g](https://doi.org/10.1039/c6lc00809g).
- [4] H. Cheung Shum, J. Varnell, D.A. Weitz, Microfluidic fabrication of water-in-water (w/w) jets and emulsions, *Biomicrofluidics* 6 (2012) 12808–128089, doi:[10.1063/1.3670365](https://doi.org/10.1063/1.3670365).
- [5] A. Sauret, Forced generation of simple and double emulsions in all-aqueous systems, *Appl. Phys. Lett.* 100 (2012) 154106, doi:[10.1063/1.3702434](https://doi.org/10.1063/1.3702434).
- [6] M. Tovar, S. Hengoju, T. Weber, L. Mahler, M. Choudhary, T. Becker, M. Roth, One sensor for multiple colors: fluorescence analysis of microdroplets in microbiological screenings by frequency-division multiplexing, *Anal. Chem.* 91 (2019) 3055–3061, doi:[10.1021/acs.analchem.8b05451](https://doi.org/10.1021/acs.analchem.8b05451).
- [7] S. Hengoju, S. Wohlfeil, A.S. Munser, S. Boehme, E. Beckert, O. Shvydkiv, M. Tovar, M. Roth, M.A. Rosenbaum, Optofluidic detection setup for multi-parametric analysis of microbiological samples in droplets, *Biomicrofluidics* 14 (2020) 24109, doi:[10.1063/1.5139603](https://doi.org/10.1063/1.5139603).
- [8] D. Vallejo, A. Nikoomezar, B.M. Paegel, J.C. Chaput, Fluorescence-activated droplet sorting for single-cell directed evolution, *ACS Synth. Biol.* 8 (2019) 1430–1440, doi:[10.1021/acssynbio.9b00103](https://doi.org/10.1021/acssynbio.9b00103).
- [9] L. Mahler, K. Wink, R.J. Beulig, K. Scherlach, M. Tovar, E. Zang, K. Martin, C. Hertweck, D. Belder, M. Roth, Detection of antibiotics synthesized in microfluidic picolitre-droplets by various actinobacteria, *Sci. Rep.* 8 (2018) 13087, doi:[10.1038/s41598-018-31263-2](https://doi.org/10.1038/s41598-018-31263-2).
- [10] A. Zinchenko, S.R.A. Devenish, B. Kintses, P.Y. Colin, M. Fischlechner, F. Hollfelder, One in a million: flow cytometric sorting of single cell-lysate assays in monodisperse picolitre double emulsion droplets for directed evolution, *Anal. Chem.* 86 (2014) 2526–2533, doi:[10.1021/ac403585p](https://doi.org/10.1021/ac403585p).
- [11] J.C. Baret, Surfactants in droplet-based microfluidics, *Lab Chip* 12 (2012) 422–433, doi:[10.1039/c1lc20582j](https://doi.org/10.1039/c1lc20582j).
- [12] P. Gruner, B. Riechers, L.A. Chacón Orellana, Q. Brosseau, F. Maes, T. Beneyton, D. Pekin, J.C. Baret, Stabilisers for water-in-fluorinated-oil dispersions: key properties for microfluidic applications, *Curr. Opin. Colloid Interface Sci.* 20 (2015) 183–191, doi:[10.1016/j.cocis.2015.07.005](https://doi.org/10.1016/j.cocis.2015.07.005).
- [13] M. Oberpaul, S. Brinkmann, M. Marner, S. Mihajlovic, B. Leis, M.A. Patras, C. Hartwig, A. Vilcinskas, P.E. Hammann, T.F. Schäberle, M. Spohn, J. Glaeser, Combination of high-throughput microfluidics and FACS technologies to leverage the numbers game in natural product discovery, *Microb. Biotechnol.* (2021), doi:[10.1111/1751-7915.13872](https://doi.org/10.1111/1751-7915.13872).



## CHAPTER 2 – BACTEROIDETES

---

Many of the nowadays known anti-infective compound classes were found while screening microorganisms belonging to the Actinobacteria, Firmicutes, Proteobacteria and Myxobacteria. While many of them are easily accessible under laboratory conditions, NP research faces constant rediscovery of known anti-infective molecules.

In the scope of this work, the Bacteroidetes phylum, a taxonomic branch not yet intensively explored in relation to the production of NPs, should be investigated for its potential to biosynthesize valuable bioactive NPs. Within the publication – *“Genomic and Chemical Decryption of the Bacteroidetes Phylum for Its Potential to Biosynthesize Natural Products”* – genomics and metabolomics technologies have been applied to uncover and highlight multiple Bacteroidetes genera for extended NP production capabilities. Based on this large *in silico* and *in vitro* generated datasets, the following two publications – *“Novel Glycerophospholipid, Lipid and N-acyl Amino Acids from Bacteroidetes: Isolation, Structure Elucidation and Bioactivity”* and *“Identification, Characterization and Synthesis of Natural Parasitic Cysteine Protease Inhibitors – Pentacidins are More Potent Falcitidin Analogs”* – depict the large untapped chemical space of the Bacteroidetes genus *Chitinophaga*. In total, the identification and characterization of several novel bioactive NPs has been archived.

## Publication III



RESEARCH ARTICLE



### Genomic and Chemical Decryption of the Bacteroidetes Phylum for Its Potential to Biosynthesize Natural Products

Stephan Brinkmann,<sup>a</sup> Michael Kurz,<sup>b</sup> Maria A. Patras,<sup>a</sup> Christoph Hartwig,<sup>a</sup> Michael Marner,<sup>a</sup> Benedikt Leis,<sup>a</sup> André Billion,<sup>a</sup> Yolanda Kleiner,<sup>a</sup> Armin Bauer,<sup>b</sup> Luigi Toti,<sup>b</sup> Christoph Pöverlein,<sup>b</sup> Peter E. Hammann,<sup>c</sup> Andreas Vilcinskas,<sup>a,d</sup> Jens Glaeser,<sup>a,c</sup> Marius Spohn,<sup>a</sup> Till F. Schäberle<sup>a,d</sup>

<sup>a</sup>Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Branch for Bioresources, Giessen, Germany

<sup>b</sup>Sanofi-Aventis Deutschland GmbH, Frankfurt am Main, Germany

<sup>c</sup>Evotec International GmbH, Göttingen, Germany

<sup>d</sup>Institute for Insect Biotechnology, Justus-Liebig-University Giessen, Giessen, Germany

**Copyright:** This article is Open Access (CC BY 4.0).

**Supporting Information** (not included in this dissertation file):

<https://journals.asm.org/doi/10.1128/spectrum.02479-21#supplementary-materials>

#### **Contribution:**

Stephan Brinkmann conceived and designed most of the study, including (i) the bioinformatics study (antiSMASH study together with Marius S. Spohn), (ii) the metabolomics study, (iii) the isolation and characterization of the chitinopeptins A–D, as well as (iv) the biosynthetic gene cluster identification and analysis. He coordinated all and performed most of the experiments (BiG-SCAPE analysis together with André Billion; NMR studies were done by Michael Kurz; Marfey's analysis together with Maria A. Patras; activity testing together with Michael Marner; synthesis of  $\beta$ -hydroxy amino acids by Yolanda Kleiner and Christoph Pöverlein. He analyzed the data and prepared all figures for the manuscript. Stephan Brinkmann wrote the draft of the manuscript with revisions of Marius S. Spohn and Till F. Schäberle as well as the following revisions.



# Genomic and Chemical Decryption of the Bacteroidetes Phylum for Its Potential to Biosynthesize Natural Products

Stephan Brinkmann,<sup>a</sup> Michael Kurz,<sup>b</sup> Maria A. Patras,<sup>a</sup> Christoph Hartwig,<sup>a</sup> Michael Marner,<sup>a</sup> Benedikt Leis,<sup>a</sup> André Billion,<sup>a</sup> Yolanda Kleiner,<sup>a</sup> Armin Bauer,<sup>b</sup> Luigi Toti,<sup>b</sup> Christoph Pöverlein,<sup>b</sup> Peter E. Hammann,<sup>c</sup> Andreas Vilcinskas,<sup>a,d</sup> Jens Glaeser,<sup>a,c</sup> Marius Spohn,<sup>a</sup> Till F. Schäberle<sup>a,d</sup>

<sup>a</sup>Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Branch for Bioresources, Giessen, Germany

<sup>b</sup>Sanofi-Aventis Deutschland GmbH, Frankfurt am Main, Germany

<sup>c</sup>Evotec International GmbH, Göttingen, Germany

<sup>d</sup>Institute for Insect Biotechnology, Justus-Liebig-University Giessen, Giessen, Germany

**ABSTRACT** With progress in genome sequencing and data sharing, 1,000s of bacterial genomes are publicly available. Genome mining—using bioinformatics tools in terms of biosynthetic gene cluster (BGC) identification, analysis, and rating—has become a key technology to explore the capabilities for natural product (NP) biosynthesis. Comprehensively, analyzing the genetic potential of the phylum Bacteroidetes revealed *Chitinophaga* as the most talented genus in terms of BGC abundance and diversity. Guided by the computational predictions, we conducted a metabolomics and bioactivity driven NP discovery program on 25 *Chitinophaga* strains. High numbers of strain-specific metabolite buckets confirmed the upfront predicted biosynthetic potential and revealed a tremendous uncharted chemical space. Mining this data set, we isolated the new iron chelating nonribosomally synthesized cyclic tetradeca- and pentadecalipodepsipeptide antibiotics chitinopeptins with activity against *Candida*, produced by *C. eiseniae* DSM 22224 and *C. flava* KCTC 62435, respectively.

**IMPORTANCE** The development of pipelines for anti-infectives to be applied in plant, animal, and human health management are dried up. However, the resistance development against compounds in use calls for new lead structures. To fill this gap and to enhance the probability of success for the discovery of new bioactive natural products, microbial taxa currently underinvestigated must be mined. This study investigates the potential within the bacterial phylum Bacteroidetes. A combination of omics-technologies revealed taxonomical hot spots for specialized metabolites. Genome- and metabolome-based analyses showed that the phylum covers a new chemical space compared with classic natural product producers. Members of the Bacteroidetes may thus present a promising bioresource for future screening and isolation campaigns.

**KEYWORDS** natural products, antifungal, NRPS, metabolomics, genomics, Bacteroidetes

The steady utilization of antimicrobials in all areas affected by human and animal diseases, but also in agriculture, supports the distribution of resistance across all environmental niches (1–4). Consequently, mortality rates accumulate—caused by multiresistant pathogens not treatable with approved anti-infectives (5). At the same time, approximately 15% of all crop production is lost to plant pathogen diseases nowadays (6). With devastating socioeconomic consequences for human health and food supply security, the increase of antifungal drug resistances has become a global problem (7). To sustain the medical care and the food supply for a growing population, continued innovation in discovery of new lead anti-infectives with improved activities and/or novel mechanisms of action is of greatest need (8, 9). Small molecules from biological origin—natural products

**Editor** Eva C. Sonnenschein, Technical University of Denmark

**Ad Hoc Peer Reviewer** Scott Alexander Jarmusch, DTU; John Vollmers, Karlsruhe Institute of Technology

**Copyright** © 2022 Brinkmann et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Address correspondence to Marius Spohn, marius.spohn@ime.fraunhofer.de, or Till F. Schäberle, till.f.schaerberle@agr.uni-giessen.de.

The authors declare no conflict of interest.

**Received** 2 December 2021

**Accepted** 29 March 2022

(NPs)—are a rich source of new chemical entities and have always been a major inspiration for the development of anti-infectives and control agents. Between 1981 and 2019, 36.3% of new medicines based on small molecules and approved by the U.S. Food and Drug Administration (FDA) were NPs, derivatives of those or synthetic compounds that utilize an NP pharmacophore (10).

A promising NP group is represented by the cyclic lipopeptides (CLPs), sharing a common structural core composed of a lipid tail linked to a cyclized oligopeptide (11). CLP biosynthesis is phylogenetically dispersed over the bacterial kingdom. This translates into an immense structural diversity arising from differences in amino acid sequences (length, type, configuration, and modification) and the composition of the fatty acid moiety. Again, these variations result in heterogeneous biological activities of CLPs and promote their investigation in several research fields (11). For example, recently isolated isopedopeptins (12) overcome multidrug resistance of Gram-negative pathogens, while daptomycin (13) (Cubicin), an FDA-approved drug, is marketed for the treatment of complicated skin and soft tissue infections caused by Gram-positive bacteria (14). Today, further promising CLPs are already under clinical investigation and attack multiple molecular targets (11). In contrast, other CLPs like fengycin (15) exhibit an intrinsic antifungal activity (16, 17) and several CLP-producing bacteria have been registered with the U.S. Environmental Protection Agency (EPA) for their application as biocontrol agents (18). Their biosynthesis and production is mostly investigated in the genera *Pseudomonas*, *Bacillus*, and *Streptomyces* spp. (19). These taxa belong to the best-studied NP producers. Traditionally, antimicrobial NPs from bacterial origin had been isolated within large cultivation campaigns on those talented microorganisms. While those genera have delivered the vast majority of structurally different NPs over the last decades, they represent only a limited phylogenetic diversity (20). In addition to classical approaches, computational evaluation of genomic data paved the way to discover novel NPs in a target oriented manner by applying bioinformatics tools for biosynthetic gene cluster (BGC) identification, analyzation, and rating (21–26). These approaches provided evidence of a yet not fully exploited biosynthetic potential in terms of chemical novelty (27–31) certified with complementary metabolomics studies (32, 33). The predicted level of novelty increases by shifting from classical NP producing taxa toward a currently still underexplored phylogenetic and chemical space (34, 35). The validity of this dogma in NP discovery was recently shown, e.g., by the discovery of teixobactin (36) and darobactin (37), both produced by rare Proteobacteria.

The Bacteroidetes phylum represents such an underexplored phylogenetic space, too (34). Although, easy to cultivate and widely spread through all environments (38, 39), there are only a few biologically active molecules described from this taxonomic clade (e.g., isopedopeptins [12], elansolids [40], pinensins [41], formadicins [42], TAN-1057A-D [43], katanosins [44], and ariakemicins [45]). Whereas environmental studies already allowed a first glance at the genetic potential of the phylum (46), there is no systematic investigation of their actual genetic and metabolic repertoire and the overall number of isolated compounds remains low.

Motivated by this gap, we examined the NP biosynthesis potential of the Bacteroidetes phylum by computational analysis of 600 publicly available genomes for their BGC amount, type, and diversity. This revealed the accumulation of NP production capability in terms of BGC amount in certain taxonomic hot spots. The vast majority of Bacteroidetes BGC of the RiPPs, NRPS, PKS, and hybrid NRPS/PKS classes are unique compared with BGCs of any other phylum, in turn providing strong evidence of an overall high potential to discover novel scaffolds from these phylum's hot spots. Particularly, the genus *Chitinophaga* represents an outstanding group, on average encoding 15.7 BGCs per strain and enriched in NRPS and PKS BGCs. Based on this analysis, we selected 25 members of this genus for a cultivation and screening program. Our systematic chemotype-barcoding matrix pointed toward a tremendous chemical space of new NPs within this data set. Almost no known NPs were identified, strengthening the upfront-performed computational strain evaluation and selection. Eventually, this process led to the discovery of new

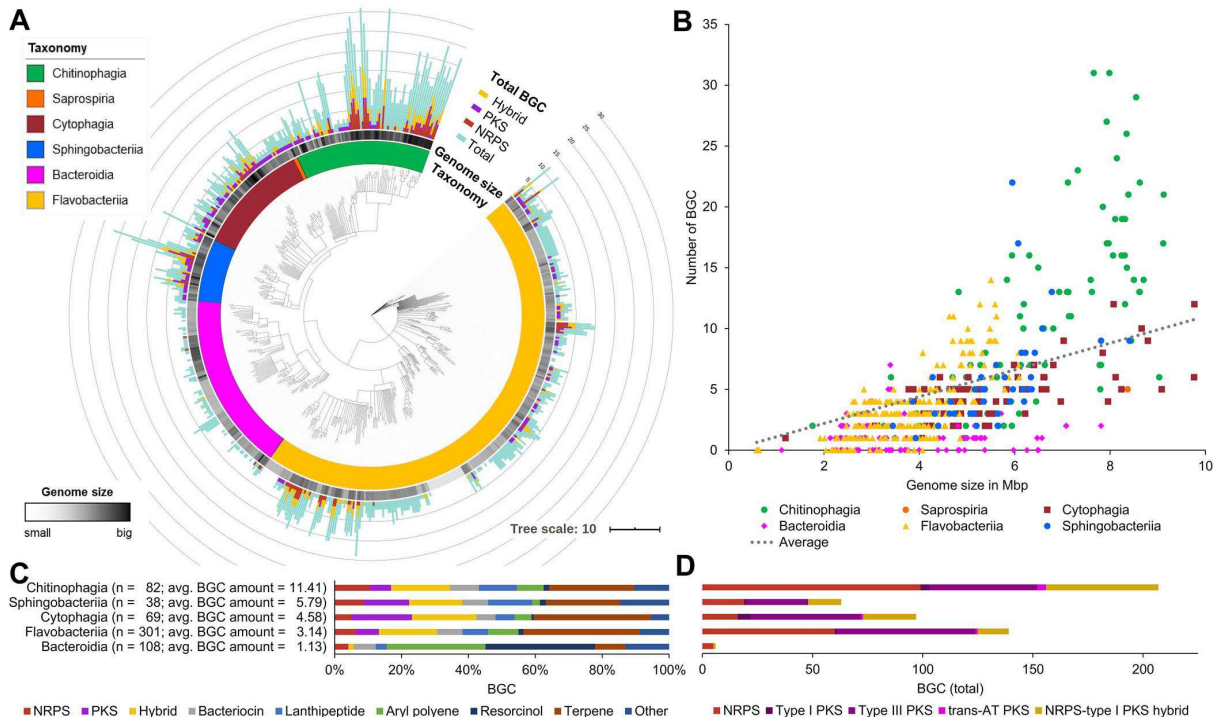
nonribosomally synthesized cyclic tetradeca- and pentadecalipodepsipeptides with iron chelating properties and candidal activity. Those CLPs named chitinopeptins A, B, C1+C2, and D1+D2 were isolated from *C. Eiseniae* DSM 22224 and *C. flava* KCTC 62435. The structures as well as the absolute stereochemistry of all amino acids were elucidated by extensive NMR studies and advanced Marfey's Analysis.

## RESULTS

**Bioinformatics analysis of the phylum Bacteroidetes.** The large and diverse phylum Bacteroidetes harbors Gram-stain-negative, chemo-organotrophic, non-spore forming rod shaped bacteria (47), graded into six so-called classes (48, 49). Members have colonized all types of habitats, including soil, ocean, freshwater, and the gastrointestinal tract of animals (38). Species from the mostly anaerobic *Bacteroidia* class are predominantly found in gastrointestinal tracts, while environmental Bacteroidetes belong primarily to the *Flavobacteriia*, *Cytophagia*, *Chitinophagia*, *Saprospiria*, and *Sphingobacteriia* classes (48, 49). Environmental studies based on amplicon diversity of adenylation and ketosynthase domains gave a first glance to the genetic potential of the phylum for the biosynthesis of NPs (46). In order to map the Bacteroidetes phylum systematically in terms of their BGC potential, we selected publicly available, closed, and annotated genomes in addition with some whole genome shotgun (WGS) projects at the time of data processing. In total, 600 genomes were analyzed using the “antibiotics and secondary metabolite analysis shell” (antiSMASH 5.0) (50). The determined total BGC amount as well as specific amount of NRPS, PKS, and hybrid BGCs was assigned to each single strain and set to the taxonomic context of the Bacteroidetes phylum, based on a phylogenetic tree calculated on complete 16S rRNA gene sequences (Fig. 1A). Assigning their BGC amount and types over the phylogenetic tree enabled comparisons between different classes and genera in terms of BGC amount and type. In most cases, a small linear positive correlation between genome size and number of secondary metabolite BGCs per genome is given, a phenomenon known from other bacteria (51) (Fig. 1A and B).

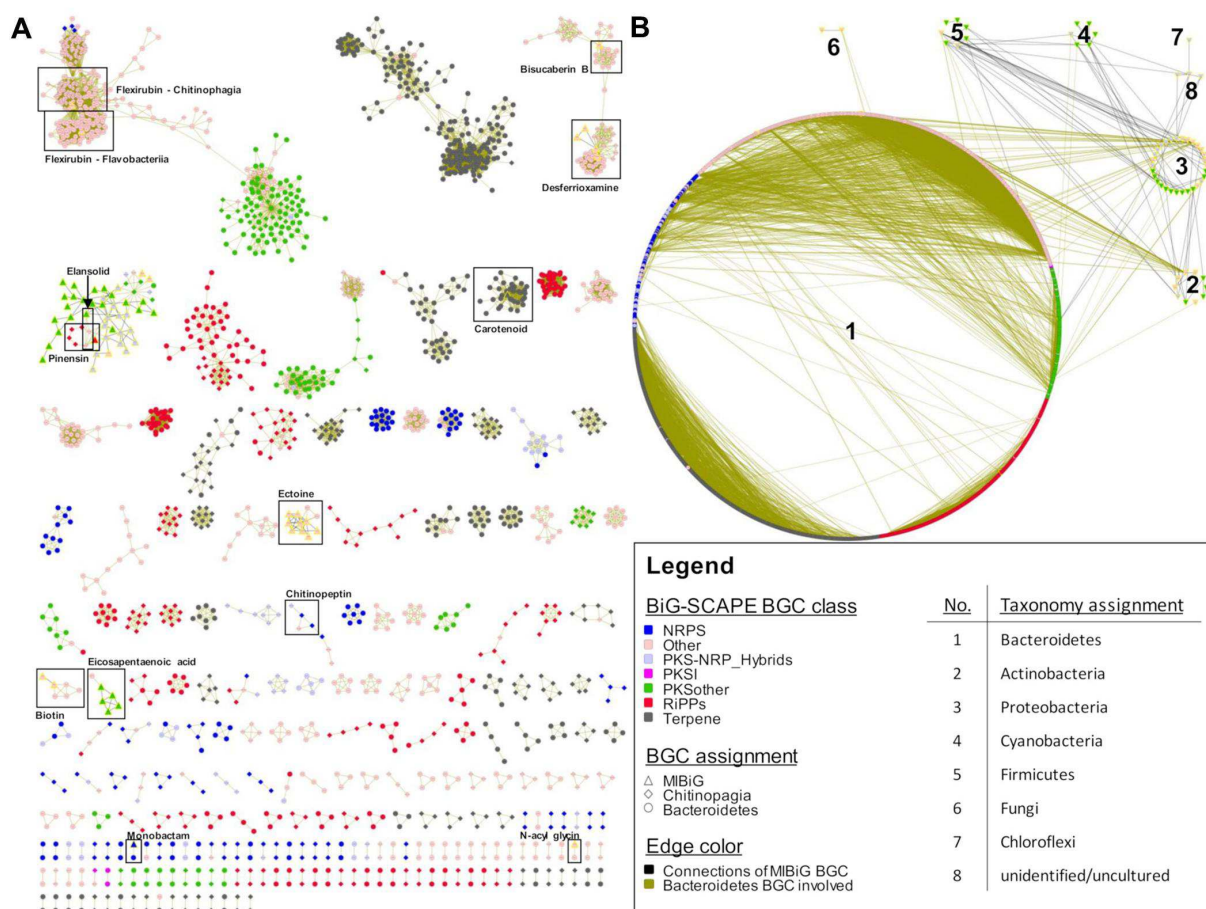
Mainly bacteria of the classes *Bacteroidia* and *Flavobacteriia*, with the smallest average genome size (3.78 and 3.51 Mb) and an average BGC amount per strain (1.15 and 3.19 BGCs), display less significance for NP discovery. Exceptions are found in the genera *Kordia* (5.33 Mb) and 10 BGCs on average, three unique genomes analyzed [ $n = 3$ ] and *Chryseobacteria* (4.43 Mb) and 6.19 BGCs on average,  $n = 48$  with up to five BGCs of the NRPS and/or PKS type. Many strains of these classes are pathogens and inhabit environments that are characterized by higher stability and lower complexity (e.g., guts) (38). NP production is an adaptive mechanism providing evolutionary fitness upon changing environmental conditions and in the presence of growth competitors (52). In accordance, the most talented bacterial NP producers, like the Actinomycetes (30) and Myxobacteria (53) are mainly found in highly competitive environments as e.g., soils. This correlation can also be seen within the Bacteroidetes phylum. In contrast to the anaerobic and pathogenic species, a higher BGC load was observed in the freely living and aerobic classes. The *Sphingobacteriia* and *Cytophagia* classes have an average genome size of 5.55 Mb and 5.57 Mb respectively, and an average BGC load of 5.79 and 4.63 per strain. An outlier is the genus *Pedobacter* with up to 22 BGCs on a single genome. Nevertheless, our analysis revealed that the *Chitinophagia* class outcompetes the other phyla members in respect to BGC amount per genome. Summarized, the class matched up with 11.4 BGCs per strain and genomes of an average size of 6.64 Mb. Within this class, the genus *Chitinophaga* ( $n = 47$ ) accumulates an enriched amount of 15.7 BGCs per strain on an average genome size of 7.51 Mb. Thirty percent of their BGCs belong to the classes of NRPS and PKS, including rare *trans*-AT PKS BGCs. The genus with the second highest BGC load within the Bacteroidetes phylum is *Taibaiella* that also belongs to the *Chitinophagia* class, with a 27% smaller genome size and on average 19% less BGCs (Fig. 1C and D).

In order to discover novel chemistry, the pure amount of BGCs is of subordinate importance in comparison to the BGCs divergence, predicted to translate into structural



**FIG 1** Bioinformatics analysis of 600 genomes of the phylum Bacteroidetes points toward high biosynthetic genetic potential of the *Chitinophagia* class. (A) Consensus tree based on maximum-likelihood method (RAxML model v8 [108], GTR GAMMA with 1,000 bootstraps) of 600 16S rRNA gene sequences color coded on class level. For each strain the genome size and biosynthetic gene cluster amount and types are depicted. Tree is annotated using iTOL v4 (110). (B) Correlation of the total gene cluster amount with the genome size of each strain. (C) Analysis of the BGC types in the individual classes. BGC types: NRPS, nonribosomal peptide; PKS, polyketide; hybrid, cluster containing more than one BGC type; and other, remaining BGC types not separately listed. (D) Detailed look onto the most essential BGC types responsible for the production of bioactive NPs. Partial BGCs on contigs <10 kb of WGS projects are not included in all graphs.

diversity within the encoded metabolites (20, 54). Thus, we expanded the computational analysis by examining the sequential and compositional similarity of the BGCs detected in the 600 genomes using the “biosynthetic gene similarity clustering and prospecting engine” (BiG-SCAPE v1.0.0) (26). BiG-SCAPE creates a distance matrix by calculating the distance between every pair of BGC in the data set. The distance matrix combines three metrics, the percentage of shared domain types (Jaccard index), the similarity between aligned domain sequences (Domain sequence similarity) and the similarity of domain pair types (Adjacency index). The comparative analysis of the Bacteroidetes BGCs with the integrated MIBiG (Minimum Information about a Biosynthetic Gene cluster, v1.4) database (55) enabled their correlation to 1,796-deposited BGCs and consequently the correlation of their synthesized metabolites. In total, in 415 of the 600 genomes analyzed, 2,594 BGCs were detected and grouped with a default similarity score cutoff of  $c = 0.6$  into a sequence similarity network with 306 gene cluster families (GCFs). Only 12 GCFs clustered with MIBiG reference BGCs of known function. Together, those 12 GCFs comprise 11.5% (298 BGCs) of all detected Bacteroidetes BGCs. Nine of them belonged to the BiG-SCAPE BGC classes of “PKSother,” “Terpenes,” or “Other.” These GCFs encode known NP classes like biotin, ectoine, *N*-acyl glycine, and eicosapentaenoic acid, as well as products of the NRPS-independent siderophore (NIS) synthetase type, precisely desferrioxamine and bisucaberin B (56), forming two connected though distinct clouds (Fig. 2). A Cytophagales specific GCF of the terpene class includes the MIBiG BGC0000650, encoding the carotenoid flexixanthin (57). The Bacteroidetes are well known producer of flexirubin-like pigments (aryl polyenes), which is reflected in a conserved biosynthesis across several genera (58, 59). The flexirubin gene cluster cloud (GCC) covers 268 BGCs from 252



**FIG 2** BiG-SCAPE (26) analysis of the phylum Bacteroidetes highlights a giant uncovered genetic potential. (A) A global network of all depicted gene cluster families with a cutoff of 0.6. Known ones are marked, named and are identified by known biosynthetic gene cluster (BGC) deposited at MiBiG (all known and deposited Bacteroidetes BGCs are found). (B) Same BGCs sorted by taxonomy. BiG-SCAPE BGC classes: NRPS, nonribosomal peptide; PKS, polyketide; and RiPPs, ribosomally synthesized and posttranslational modified peptides. Singeltons (unique BGCs without any connection) are not shown. Visualization and manipulation by Cytoscape v3.4.0 (113).

individual strains and five of six analyzed classes. In our analysis, only the newly formed *Saprospira* class (49) was an exception. However, considering that the analysis included only two *Saprospira* strains it does not yet allow any integral assessment of its capabilities to produce these yellow pigments. The flexirubin GCC can be divided into at least five distinct GCFs. Resolved on class level this revealed a specific *Flavobacteriia* family including BGC0000838 from *Flavobacterium johnsoniae* UW101 (58) as well as a specific *Chitinophaga* family including BGC0000839 from *Chitinophaga pinensis* DSM 2588 (59). The latter in turn being directly connected to a third GCF, in majority covering BGCs from its genus *Chitinophaga* while not including a reference BGC.

In addition, the reference BGCs described to encode the bioactive NPs monobactam SQ 28,332 (60, 61), elansolid (40, 62), and pinensins (41), are annotated to three distinct GCFs (Fig. 2A). The monobactam BGC (BGC0001672) was unique and only identified in its described producer strain *Flexibacter* sp. ATCC 35103. Elansolids and pinensins represent patent protected chemical entities active against Gram-positive bacteria and filamentous fungi and yeasts, respectively. The complete elansolid encoding BGC (BGC0000178), almost 80 kbp in size, was identified in the genomes of strain *Chitinophaga* sp. YR627 and *Chitinophaga pinensis* DSM 2588 (=DSM 28390 [63]) with the genetic potential to produce elansolid already proposed for *C. pinensis* (64) (Fig. S1A). Besides the original producer strain, *Chitinophaga sancti* DSM 21134 (65), these

strains provide alternative bioresources to access these polyketide-derived macrolides. In addition, both strains also harbor the pinensin BGC directly co-localized with the elansolid-type BGC. This co-localization leads to an artificial connectivity between both GCFs by using the chosen BiG-SCAPE parameters. Manual curation revealed six strains carrying a pinensin-like BGC in their genome in total (Fig. S1B). The alignment of the RiPP core peptide revealed that only the amino acid sequence from strain *Chitinophaga* sp. YR627 was identical to the described pinensin sequence. The other strains show amino acid sequence variations, pointing toward structural variance (Fig. S1C).

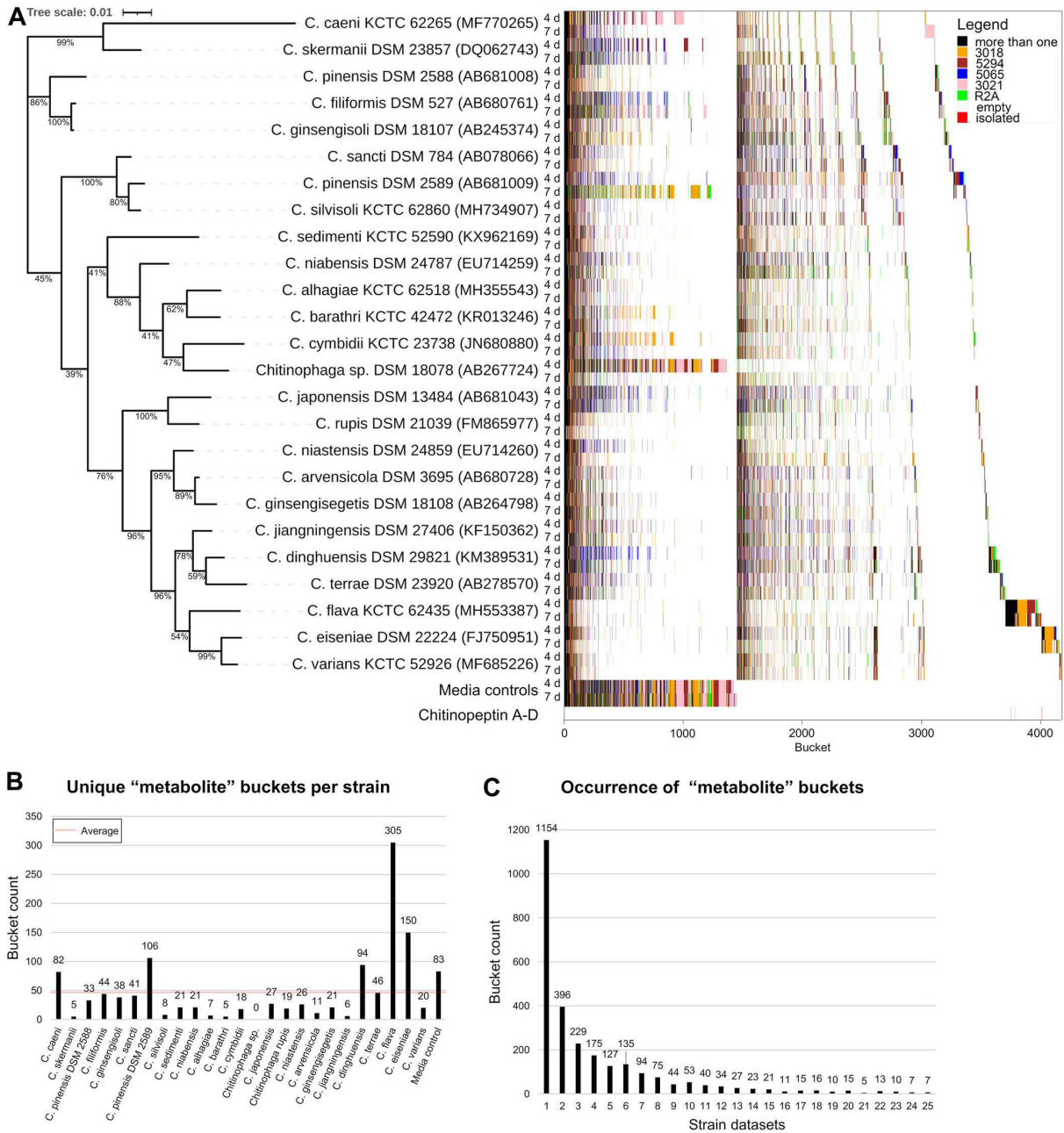
With >200 GCFs identified and only 12 of them annotated toward known BGCs and their metabolites, the sequential and compositional similarity analysis revealed a BGC diversity within the Bacteroidetes phylum, differing from the composition of known BGCs deposited in the MiBIG database.

Extension of this similarity network analysis toward taxonomic relations on phylum level showed that Bacteroidetes BGC of the RiPPs, NRPS, PKS, and hybrid NRPS/PKS classes are unique compared with BGCs of any other phylum (Fig. 2B). This in turn provides a strong evidence of a general high potential to discover novel metabolites from this phylum. The majority (~66%) of all GCFs belonging to the above-mentioned BGC classes are found in the *Chitinophagia* class, only representing 13.7% of the analyzed strains. Within this class, the genus *Chitinophaga* can be prioritized in terms of BGC amount and composition. In respect to many more complete unique and novel RiPP, NRPS, PKS, and hybrid BGCs thereof (not depicted in the network), this is a strong indication that the biosynthetic potential within this genus is far from being fully exploited. It can be considered as the most promising starting point for the discovery of novel metabolites within this phylum.

**Metabolomics of the *Chitinophaga*.** Based on the genomic data evaluation, we selected the *Chitinophaga* for performing a bioactivity guided NP discovery program. NPs are considered to be nonessential metabolites for bacterial growth and reproduction but rather providing evolutionary fitness, thus, being expressed as adaptive response to changing environmental conditions. Consequently, the discovered BGC potential is not expected to translate into the actually produced metabolite pattern under laboratory conditions (22). A common theme of strategies to approach this challenging link is the cultivation in several media variants exposing the strains to various stress conditions, e.g., nutrient depletion (66, 67). As a consequence of nutrient depletion, bacteria enter the stationary phase and reduce or even cease growth, often found to coincide with induction of secondary metabolite production (66, 68). To trigger these events, we cultivated a diversity of 25 *Chitinophaga* strains (Table S1) in five different media for 4 as well as 7 days.

The metabolites were extracted from freeze-dried culture broths with methanol and the organic extracts were subsequently analyzed by ultra-high performance liquid chromatography-high resolution mass spectrometry (UHPLC-QTOF-HR-MS). LC-MS data sets from a total of 250 extracts (and media controls) were examined allowing the definition of strain-specific molecular features. In an initial step, features (represented by  $m/z$ , retention time, isotope pattern) were calculated within all extracts. Curation of all data sets was necessary to filter background noise and confirm the authenticity of defined features. This curation step helped to avoid false uniqueness due to concentrations near the corresponding detection limit and to reduce the possibility of picking up background noise. Furthermore, the possibility of multiple mass spectrometric features for any NP contributes to the complexity of those data sets, e.g., by the formation of different ion adducts and in-source-generated fragment ions of single molecules. The final data set consisted of 93,526 features. Those were aligned into 4,188 buckets with a bucket being defined as an  $m/z$  and retention time (RT) region hosting all features with matching  $m/z$  and RT (69). We created a chemotype-barcoding matrix of this complex data set, allowing its visualization and evaluation (Fig. 3A). After normalization of the data set by buckets congruent with the media controls (in total 1,452),





**FIG 3** Taxonomical arrayed chemotype-barcoding matrix reveals an uncharted chemical space within the genus *Chitinophaga*. (A) The tree is based on a Clustal W alignment (111) of available 16S rRNA gene sequences of 25 *Chitinophaga* strains available for cultivation. The tree was calculated using MEGA v7.0.26 with the maximum-likelihood method and GTR-Gamma model (112). Percentage on the tree branches indicate values of 1,000 bootstrap replicates with a bootstrap support of more than 50%. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The bucketing process is depicted as a chemotype-barcode matrix and color coded by the condition each individual bucket was present. (B) Bar blot of the unique “metabolite” buckets of each strain. (C) Bar blot depicting the occurrence of “metabolite” buckets in the data sets.

we determined in total 2,736 buckets as representing produced metabolites of the investigated *Chitinophaga* set.

The detected buckets were analyzed for presence in all combinations considering utilized cultivation media and incubation time. They were put in order according to 16S rRNA sequence similarity on strain level. In order to facilitate data interpretation

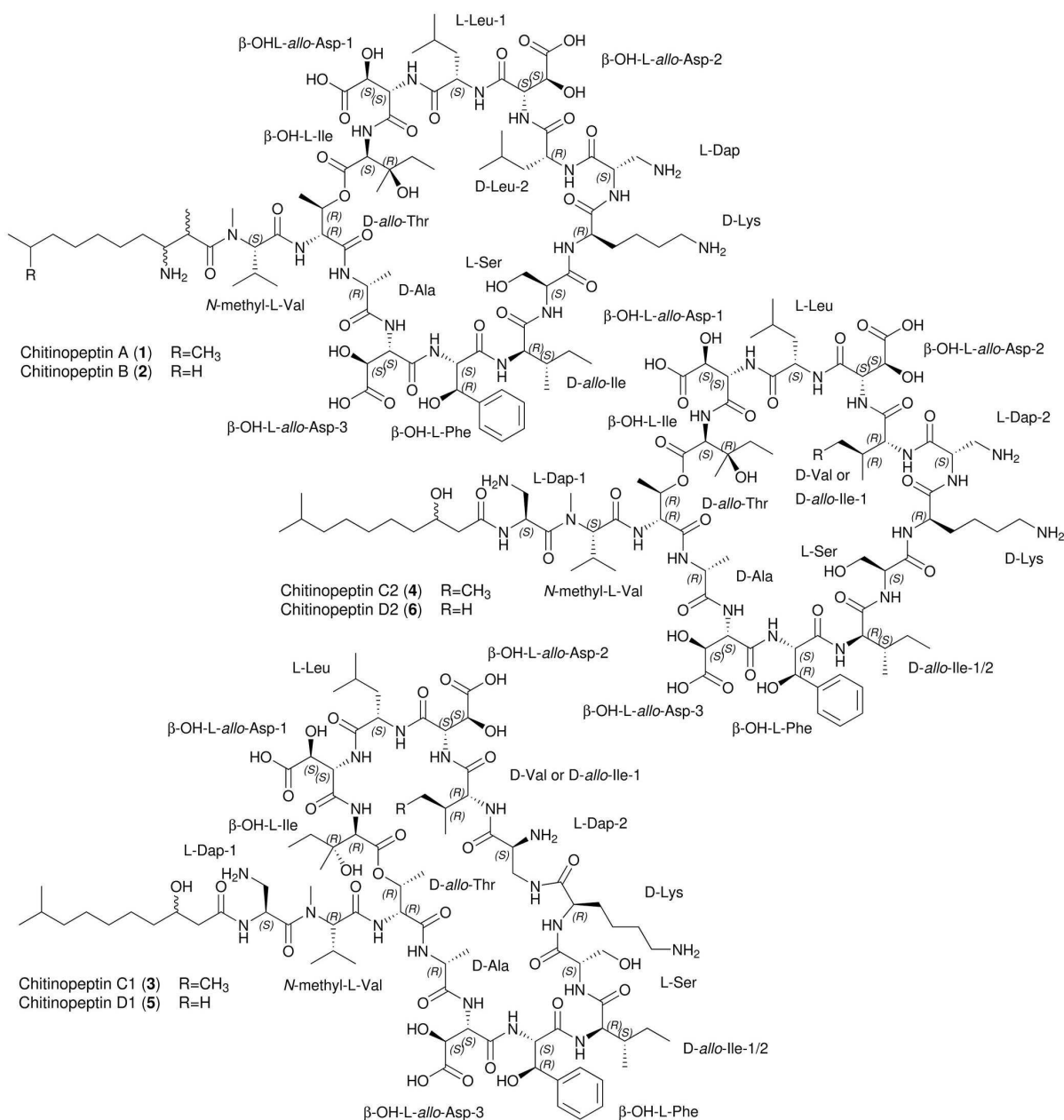
and to identify strain specific as well as conserved metabolites, all buckets were sorted according to their frequency of appearance. Comparative visualization of the short (4 days) and prolonged (7 days) incubation time revealed differences on the chemotype profile of the strains. Most strains metabolize the majority of media components within the first 4 days of cultivation, recognized by only a small number of “media buckets” still present in the extracts after this incubation period. A second population of rather slow-growing strains appeared to shift the metabolic profile only after 7 days of incubation in comparison to the respective media control (e.g., *C. caeni* KCTC 62265, *C. dinghuensis* DSM 29821, *C. niastensis* DSM 24859, *C. barathri*, and *C. cymbidii*). Especially *Chitinophaga* sp. DSM 18078 required a prolonged incubation time to metabolize the media ingredients, producing 77.4% more metabolite buckets after 7 days of incubation in comparison to the earlier sampling time. Within the whole data set, on average 26.6% (92.4) more metabolite buckets were detected after seven than after 4 days of incubation. In contrast, a fraction of bacterial metabolite buckets disappeared within the extracts of five strains after prolonged incubation, showing the necessity to vary cultivation conditions to access a possibly comprehensive metabolite profile of each investigated strain. The many media-specific buckets (colored) compared with the ones being produced in various media (black) also show the effect of variations of the bacterial nutrient supply (Fig. 3A). In combination with the varied cultivation period, this led to an average of approximately 46 unique buckets per strain. Especially *C. flava* KCTC 62435 and its close relative *C. eiseniae* DSM 22224 outcompeted the others by producing 305 and 150 unique metabolite buckets respectively (Fig. 3B), indicating their status as best-in-class producers. In total 1,154 buckets (~42%) of the entire data set were identified only in one respective strain data set, representing strain-specific metabolites (Fig. 3C). This high level of strain-specific metabolites shows the heterogeneity of the investigated strain set. In parallel, this experiment depicts a structural diversity in terms of molecular size up to 5145.051 Da and covered polarity range (Fig. S2).

Analysis of the conserved metabolites (which includes primary and secondary metabolites) revealed 357 buckets to be present in at least 10 out of the 25 total analyzed strains while only the low number of seven buckets could be detected within samples of all *Chitinophaga* strains investigated. Based on their MS<sup>2</sup>-fragmentation patterns sharing the loss of long carbon chains, we postulated a structural relationship between four of them and assigned them as hitherto unknown (amino/phospho) lipids.

Next, the complete LC-MS data set was examined for the presence of structurally characterized microbial NPs (~1,700) deposited in our in-house database on the basis of accurate *m/z*, RT, and isotope pattern. The frequency of rediscovery was zero. Considering the known bias of the database for natural products from classical NP producing taxa such as Actinobacteria, Myxobacteria, and fungi, this confirmed the low congruency toward these taxa, which was also found by our BGC categorization study. A complementary scan of LC-MS/MS data of the entire data set was compared with *in silico* fragments of >40k NPs deposited in the commercial database AntiBase (70). Congruently, no database-recorded NP was identified within our data set besides falcitidin (71), an acyltetrapeptide produced by a *Chitinophaga* strain. Although this finding shows the general applicability of this workflow, the underrepresentation of NPs isolated from the phylum Bacteroidetes also in public databases is still a severe limitation for comprehensive categorization of their omics-data today.

However, as a consequence, these data provided a high confidence level in the investigated strain portfolio and showed that the metabolite spectrum produced by these strains is largely underexplored and different from the metabolites produced by classical NP producer taxa. Even though strains of the genus *Chitinophaga* are phylogenetically closely related, they produce a heterogeneity of metabolites; thereby, showing a high number of strain-specific metabolites, associated with a likelihood for chemical novelty.

**Chitinopeptins, new CLPs from *C. eiseniae* and *C. flava*.** A correlation of those untapped buckets with antibacterial activity was performed by screening the organic crude extracts against a panel of opportunistic microbial pathogens. In particular, methanol extracts of *Chitinophaga eiseniae* DSM 22224 and *Chitinophaga flava* KCTC



**FIG 4** Chemical structures of the chitinopeptins A–D. Compounds 1 and 2 are produced by *C. eisensiae* DSM 22224, compounds 3 to 6 by *C. flava* KCTC 62435.

62435, the two strains with the highest level of metabolic uniqueness, exhibited strong activity against *Candida albicans* FH2173. Bioactivity assay and UHPLC-HR-ESI-MS guided fractionation led to the identification of six new cyclic lipodepsipeptides. *C. eisensiae* produced the two tetradecalipodepsipeptides chitinopeptins A and B (1 and 2) with molecular formulae  $C_{82}H_{137}N_{17}O_{28}$  (1,  $[M+H]^+$  1809.0052) and  $C_{81}H_{135}N_{17}O_{28}$  (2,  $[M+H]^+$  1794.9907). Whereas *C. flava* assembled the four pentadecalipodepsipeptides chitinopeptins C1+C2 and D1+D2 (3 to 6) with molecular formulae  $C_{84}H_{140}N_{18}O_{30}$  (3 and 4,  $[M+H]^+$  1882.0177) and  $C_{83}H_{138}N_{18}O_{30}$  (5 and 6,  $[M+H]^+$  1868.0059) (Fig. 4). All

compounds were present in the MS spectra as pairs of  $[M + 3H]^{3+}$  and  $[M + 2H]^{2+}$  ions (Fig. S3A). All six native peptides appeared to be highly stable, because only poor yields of fragment ions arose using electrospray ionization source (Fig. S3B), even under elevated collision energy conditions (up to 55 eV), thereby preventing MS/MS based structure prediction and structural relationship analysis using molecular networking.

The chemical structures of the six compounds were determined by extensive NMR studies using 1D- $^1H$ , 1D- $^{13}C$ , DQF-COSY, TOCSY, ROESY, multiplicity edited-HSQC, and HMBC spectra (Fig. 4). The analysis of the compounds in most “standard” NMR solvents was hampered by either extreme line broadening (DMSO, MeOH, pyridine) or poor solubility ( $H_2O$ , acetone). However, a mixture of  $H_2O$  and  $CD_3CN$  (ratio 1:1) gave rise to NMR spectra of high quality and confirmed the presence of peptides. In order to obtain a good dispersion of the amide resonances,  $^1H$ -spectra were acquired at different temperatures between 290 and 305 K. A temperature of 299 K or 300 K, respectively, was found to be the best compromise considering signal dispersion and line broadening (Fig. S4 to S31 and Table S2 and S3).

The first compound to be studied was chitinopeptin A. The analysis of the NMR spectra revealed the presence of several canonical amino acids (1 Thr, 1 Ala, 1 Ile, 1 Ser, 1 Lys, and 2 Leu), and several hydroxylated amino acids (3  $\beta$ -OH Asp, 1  $\beta$ -OH Phe, 1  $\beta$ -OH Ile). In addition, one *N*-methyl Val and one 2,3-diaminopropionic acid (Dap) moiety could be assigned. The sequence of the amino acids was established by correlations in the ROESY ( $NH_i/NH_{i+1}$ ,  $NH_{i+1}/H\alpha_i$ ) and HMBC spectrum ( $C'_i/NH_{i+1}$ ). The formation of a cyclic peptide was indicated by the  $^1H$ -chemical shift of the  $\beta$ -proton of the Thr residue in position 2 (5.23 ppm) and the correlation in the HMBC spectrum between the carbonyl carbon of the C-terminal  $\beta$ -OH Ile (position 14) and the  $\beta$ -proton of the Thr residue.

Aside from the amino acids, a modified fatty acid residue was identified which could be described as 2,9-dimethyl-3-amino decanoic acid. Correlations in the HMBC spectrum between the carboxyl carbon (C1) and the *N*-methyl group of the *N*-methyl Val proved its position at the *N*-terminus of the peptide. The structure of chitinopeptin B was almost identical to the structure of CLP 1. The only difference was the substitution of the 2,9-dimethyl-3-amino decanoic acid by 3-amino-9-methyl decanoic acid.

CLPs 3 and 4 were isolated as a 5:4 mixture of two components. One of the main differences compared with the structures above was an additional Dap residue, which was inserted between the fatty acid moiety and *N*-methyl Val. Both components contain an Ile instead of a Leu (CLPs 1 and 2) at position 10. Furthermore, the 2,9-dimethyl-3-amino decanoic acid is replaced by a 3-hydroxy-9-methyl decanoic acid. The two components 3 and 4 differ in the constitution of the Dap in position 9. In one component 4, the peptide bond between the  $\alpha$ -amino function and the carbonyl group of Lys is formed, while in the other component 3, the  $\beta$ -amino group (side chain) is connected to the carbonyl group of Lys. The same pair of structures as for CLPs 3 and 4 is obtained in the case of CLPs 5 and 6. In contrast to the previous ones, both components contain a Val in position 10 instead of Leu or Ile, respectively.

The absolute stereochemistry of the amino acids in the CLPs was determined by using advanced Marfey's Analysis (72). Comparison of the RTs with (commercially available) reference amino acids allowed identification of nine out of 14 (CLPs 1 and 2) and 10 out of 15 amino acids (CLPs 3–6), respectively. RTs of *N*-methyl-L-Val, D-allo-Thr, D-Ala, D-allo-Ile (2  $\times$  for CLPs 3 and 4), L-Ser, D-Lys, L-Dap (2 $\times$  for CLPs 3 to 6), D-Leu (only CLPs 1 and 2), L-Leu, and D-Val (only CLPs 5 and 6) matched the reference ones. Assigning L- as well as D-leucine within structures 1 and 2 to position 13 and 10, respectively, was possible because position 10 was the only variable position in all six depsipeptides. Either Ile (CLPs 3 and 4) or Val (CLPs 5 and 6) were identified at this position with all amino acids having D-configuration. Therefore, it can be assumed that D-Leu is present at position 10 in CLPs 1 and 2 (Fig. S32 and S33).

Authentic samples of the  $\beta$ -hydroxyamino acids or suitable precursors were synthesized utilizing modified literature known procedures. All four stereoisomers of  $\beta$ -hydroxyaspartic acid were obtained from (–)-dibenzyl D-tartrate or (+)-dibenzyl L-tartrate,

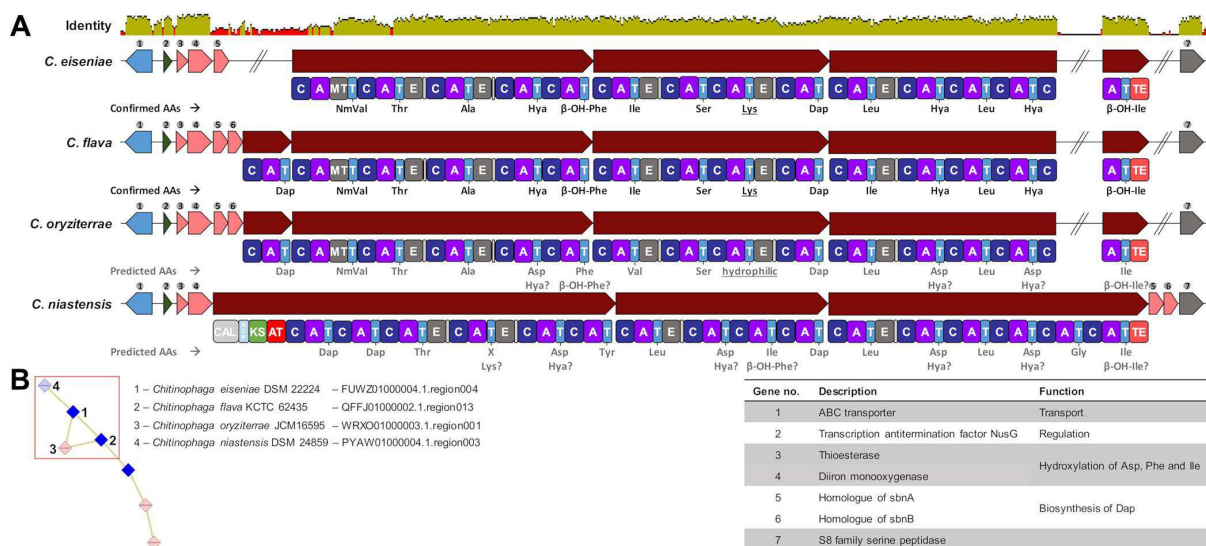
respectively, according to a procedure described by Breuning et al. (73). While the *anti*-isomers are directly accessible, the *syn*-isomers were obtained by selective base-induced epimerization of the azido-intermediates and separation of the two isomers by HPLC. Cbz-protected *L*-isomers of the  $\beta$ -hydroxyphenylalanines and  $\beta$ -hydroxyisoleucines were synthesized starting from an orthoester protected *L*-serine aldehyde, initially described by Blaskovich and Lajoie (74–76). To cover the corresponding *D*-isomers for analytical purposes, racemic samples of the amino acids were produced by racemization of the orthoester protected *L*-serine aldehyde by simple chromatography on silica (74). Advanced Marfey's Analysis determined (2*S*,3*S*)-3-hydroxyaspartic acid, (2*S*,3*R*)-3-hydroxyphenylalanine and (2*S*,3*R*)-3-hydroxyisoleucine as the absolute stereochemistry of  $\beta$ -hydroxyamino acids for all six CLPs (Fig. S34 to 36).

To the best of our knowledge, these represent the first CLPs described from the genus *Chitinophaga* and after the recently described isopedopeptins (12), the second CLP family of the entire phylum. They contain a high number of non-proteinogenic amino acids (i.e., 12 of 14 amino acids in CLPs 1 and 2, and 13 of 15 in CLPs 3 to 6). Beta-hydroxylations of Asp, Phe, and Ile are the most abundant modifications and *N*-methyl Val and Dap are incorporated into the peptide backbone. Furthermore, during LC-MS analysis, a mass shift of 52.908 Da accompanied with an emerging UV maximum at 310 nm, as well as a shift in RT was observed for compounds 1 to 6. This traced back to the coordination of the compounds to iron impurities during LC-analysis, which was confirmed by the addition of Fe(III)-citrate to the compounds prior to LC-MS analysis (Fig. S37). We postulated that the RT shift is due to a conformational change and altered polarities as a consequence of iron complexation. Iron coordination is a known feature of siderophores, produced by bacteria upon low iron stress (77). To investigate the impact of iron on the CLPs production, CLP 1C. *eiseniae* was cultured in 3018 medium (for composition see Materials and Methods section) supplemented with different iron concentrations. However, the overall production of CLPs 1 and 2 and their iron complexes was not repressed by increased iron levels, contrasting the iron-responsive productivity of classical siderophores (Fig. S38).

Chitinopeptins A to D were tested against eight Gram-negative and five Gram-positive bacteria, as well as against three filamentous fungi and *C. albicans* (Table S4). For all tested CLPs, activity was observed against *M. catarrhalis* ATCC 25238 and *B. subtilis* DSM 10, exhibiting MICs down to 2  $\mu$ g/mL. The tetradecalipodepsipeptides 1 and 2 exhibited activity at 4 to 8  $\mu$ g/mL against *C. albicans* FH2173, while the pentadecalipodepsipeptides exhibited MICs of only 16  $\mu$ g/mL. Screenings against filamentous fungi revealed MICs of 16  $\mu$ g/mL against *Z. tritici* MUCL45407, while no activity was observed against *A. flavus* ATCC 9170 and *F. oxysporum* ATCC 7601. To investigate the impact of iron-binding on the bioactivity, CLP 1 was tested as a representative also in its iron complexed form confirmed by LC-MS analysis (Fig. S37). This revealed that the bioactive potency of the iron complex is reduced in comparison to its iron free form, although not completely suppressed (Table S4).

**BGCs corresponding to chitinopeptins.** In order to identify the BGCs encoding the chitinopeptins' biosynthesis, we scanned the genomes of *C. eiseniae* (FUWZ01.1) and *C. flava* (QFFJ01.1) for NRPS-type BGCs matching the structural features of the molecules. The number and predicted substrate specificity of the A-domains, the overall composition of the NRPS assembly line, as well as precursor supply and post-assembly modifications were taken into account. We identified the BGCs in *C. eiseniae* and *C. flava* congruent to the CLP structure in each case (Fig. 5A). Furthermore, the positioning of all epimerization domains within the detected NRPS genes, encoding the conversion of *L*- to *D*-amino acids, is in agreement with the determined stereochemistry of the molecules. Thereby, the domains are classically embedded in the NRPS assembly lines. No racemase(s) encoded in *trans* or C domains catalyzing the conversion of amino acids, are observed as it is the case e.g., in the BGC of the stechlisins, CLPs produced by *Pseudomonas* sp. (78).

In our BiG-SCAPE network, these BGCs from *C. eiseniae* and *C. flava* were part of a GCF, and manual inspection confirmed two further related BGCs within the genomes of *C. oryzae* JCM16595 (WRX001.1) and *C. niastensis* DSM 24859 (PYAW01.1) (Fig. 5B). Besides the



**FIG 5** Biosynthetic gene clusters responsible for the production of chitinopeptin A to D and further derivatives. (A) BGCs of strains *C. eiseniae*, *C. flava*, *C. oryztarreae*, and *C. niastensis* with matching amino acid sequence and additional biosynthetic genes responsible for the hydroxylation of Asp, Phe, and Ile or the in cooperation of 2,3-diaminopropionic acid. Identity was calculated with a standard MAFFT alignment (107). (B) Gene cluster family of this iron chelating cyclic lipodepsipeptides. AA, amino acids; NmVal, *N*-Me-Val; Hya,  $\beta$ -hydroxyaspartic acid; C, condensation domain; A, adenylation domain; MT, nitrogen methyltransferase; T, peptidyl-carrier protein domain; E, epimerization domain; TE, thioesterase domain; CAL, co-enzyme A ligase domain; ACP, acyl-carrier protein domain; KS, ketosynthase domain; AT, acyltransferase domain.

structural NRPS genes, further genes are conserved between all four BGCs, predicted to encode an ATP-binding cassette (ABC) transporter, a transcription factor, an S8 family serine peptidase, a thioester reductase domain, and a metal  $\beta$ -lactamase fold metallo-hydrolase. The biosynthesis of these CLPs requires a  $\beta$ -hydroxylation tailoring reaction of the precursor Asp, Phe, and Ile. This structural feature is also present in chloramphenicol and its biosynthesis was shown to be catalyzed by the diiron-monooxygenase CmlA that catalyzes substrate hydroxylation by dioxygen activation. CmlA coordinates two metal ions within a His-X-His-X-Asp-His motif and possesses a thioester reductase domain (79). Both features are also present in all four detected BGCs, encoded by two separate genes annotated as thioester reductase domain and metal  $\beta$ -lactamase fold metallo-hydrolase (gene 3 and 4) (Fig. 5B and Table S5). Moreover, all four BGCs contain genes with sequence similarity to *sbnA* and *sbnB*, encoding enzymes that catalyze the synthesis of the non-proteinogenic amino acid Dap (80, 81) (Table S6). The presence of genes encoding for supply with precursor amino acid(s) is a common feature in BGCs corresponding to NPs described in the Firmicutes and Actinobacteria phyla (82–87). Interestingly, within the BGC of *C. eiseniae* a *sbnB* homologue is missing. To rule out an error during genome sequencing, assembly, and annotation, we amplified the respective section by PCR and confirmed the published genome sequence. However, not encoded in *cis*, *C. eiseniae* carries further *sbnA* and *sbnB* homologues. These are encoded in other NRPS-type BGCs (FUWZ01000006, location: 555,550 to 557,479 and FUWZ01000004, location: 273,583 to 275,522) and potentially function in *trans* to compensate the absence of the gene within the chitinopeptin A and B BGC. In general, we observed that the Dap subcluster *sbnA/B* is an abundant genetic feature of Bacteroidetes secondary metabolism, present in many BGCs. This is not restricted to *Chitinophaga*, but expanded to different genera and in consequence, a structural feature of Bacteroidetes NPs predictively found with high frequency. Indeed, Dap moieties are present in the known antibacterial Bacteroidetes compounds isopedopeptins (12) and TAN-1057 A to D (43).

Considering the composition of the structural NRPS genes in terms of A domain number as well as predicted A domain substrate specificity (88), we propose that *C. oryztarreae* and *C. niastensis* carry the potential to provide additional structural variety

to the chitinopeptins (Table S7). The NRPS gene of *C. niastensis* encodes 16 A-domain-containing modules, predictably producing hexadecapeptides. Furthermore, an initial PKS1 module replaces the C-starter domains present in the three other BGCs of this GCF. This points toward to an attachment of carboxylic acid residues to the peptide scaffold, resulting in further compound diversification. While *C. oryzae* was not available for cultivation, we inspected the extracts of *C. niastensis* for putative further CLPs. Indeed, although only detected in traces, possible products could be identified based on comparable RTs and isotope patterns with  $m/z$  of 680.9817 [M + 3H]<sup>3+</sup> and 685.6545 [M + 3H]<sup>3+</sup> (Fig. S39).

## DISCUSSION

Mining microbes for their secondary/specialized metabolites continues to be an essential and valuable part of drug discovery pipelines. With the constant rediscovery of already known NPs, the past decades are representative for the limitations of classical approaches focusing on the most talented producer taxa that only represent a limited phylogenetic space of the bacterial kingdom (34). However, the microbial diversity is growing with each new phylum added to the tree of life (20). In parallel, rapid improvements of resolution and accuracy allowed us to apply technologically more complex metabolomics- and genomics-guided methods to mine nature (24, 25, 31, 33). Genomics-guided methods predict the strains phenotype and their theoretical ability to produce metabolites, based on identified BGCs and their consecutive prioritization (23). Complementary omic-technologies, especially metabolomics enable the sophisticated characterization of the strains' chemotypes, the actually detected metabolites. Standalone or applied sequentially, both technologies already made a strong impact on the field of NPs. (Semi)automatized combinations of genomics and metabolomics are under development and aim to establish a direct link between the biosynthetic potential of microbial strains and their actually expressed metabolites (29, 33, 89). The basis for this was built over the last years and gave evidence that those tools are as powerful as their underlying training data sets (90).

In the future, these approaches will likely support the field by promoting reduction of rediscovery rates and in turn allowing efficient sample prioritization with respect to the target molecules of the respective studies (91). Until today, these approaches were primarily applied on big data sets from well-known NP producing taxa such as Actinobacteria or Myxobacteria (29, 30). These studies enabled technology development, while directly expanding the overall knowledge on the still available chemical space of these taxa as well as the identification of novel secondary metabolites or derivatives of known ones. Here, we translated this approach of application of complementary omics-technologies for the first time to the currently still underexplored Bacteroidetes phylum. The comprehensive computational analysis of publicly available Bacteroidetes genomes revealed the *Chitinophaga* as the most promising genus in terms of NP production capabilities. The genome-wide evaluation of 47 *Chitinophaga* strains showed the presence of 15.7 BGCs on average, with a minimum of six and a maximum of 31 BGC. Although the number of detected BGCs is strongly dependent on the utilized bioinformatics tools and chosen settings, the most talented *Chitinophaga* competed with the BGC loads of some Actinobacteria (92). Moreover, the variation of BGC amount within a genus is a known phenomenon and also observed e.g., in the *Amycolatopsis* genus (93).

The Dictionary of Natural Products (DNP) (94) (status 01/2021) contains only 14 entries ascribed to the genus *Chitinophaga*. This does not reflect their predicted genetic potential, because e.g., for the genus *Streptomyces*, 8,969 chemical entities are deposited. The extrapolation of the approximately doubled BGC load of *Streptomyces* (25–70 BGCs) (51) results in ~650 times more chemical entries, a theoretical gap that justifies efforts to investigate the metabolic repertoire of the genus *Chitinophaga*. Our metabolomics analysis of 25 *Chitinophaga* strains confirmed the genomic and database-based prediction by empirical data. The identified chemical space is represented by more than 1,000 unique and unknown candidates that could not be associated to any microbial NP known today. Specialized metabolites do not

have a pivotal role in the survival of the microorganism. They are considered not essential for vegetative growth, as well as reproduction, but to give a fitness advantage in specialized growth conditions. As a consequence of that, these metabolites are not constantly expressed, but rather as a specific response to defined environmental conditions. Their production occurs generally in a growth-phase dependent manner and often coinciding with reduction in bacterial growth rate or even growth cessation mainly based on nutrient depletion (68). The strategy to challenge the microorganisms by varying cultivation conditions and incubation times successfully translated also in the genus *Chitinophaga* toward an increased metabolite diversity. This shows the need to adjust the cultivation time to the specific growth pace of the investigated strains. The uniqueness and value of this data set compares with data sets gathered by metabolomics studies of well-known NP producer taxa. Although comparability of such extensive approaches is challenging due to different data set size, composition as well as technical devices and settings, the detection of bucket/metabolite numbers per strain in our *Chitinophaga* strain set shows a similar range compared with the ones described for different Myxobacteria genera (32). One striking difference however, is the lack of dereplicated known metabolites, which can be explained by their underrepresentation in all reference databases utilized in this study for data categorization. This is likely due to the fact that investigations of the genus *Chitinophaga* and the Bacteroidetes phylum itself have just started.

In conclusion, our analysis reveals that the Bacteroidetes phylum has a distinct metabolic repertoire compared with classical NP producer taxa, indicated by the small overlapping taxonomic relationship of their GCFs. Eventually, our combinatorial approach in utilizing genomics and metabolomics facilitated the genus and strain prioritization and paved the way for the discovery of the chitinopeptins A to D. These NRPS-assembled CLPs exhibit primarily activity against *Candida albicans* and were found to coordinate iron. Binding of metal ions is a feature also described for other CLPs such as pseudofactin II, which displays an increased antimicrobial activity upon metal-coordination due to disruption of the cytoplasmic membrane in its chelated state (95). In general, it is known that individual CLPs possess multiple functions such as iron chelation, antimicrobial activity, and interaction in bacterial motility (19).

Incidences of fungal infections, including candidiasis caused by *Candida* species, are increasing (96). Currently, clinically used antifungals are limited to four classes: azoles, polyenes, echinocandins, and pyrimidine analogs. Associated with occurrence and spreading of resistances, this leads to alarmingly decreased treatment success and increased possibilities of fatal outcomes (97). This development clearly demands counteraction. One possible way to use is the constant discovery of NPs with novel antifungal activity. The chitinopeptins share properties with the NP family of CLPs that exhibit an intrinsic antifungal activity. Mode of action studies for these compounds suggest a membrane interaction and leakage effects up to pore-forming properties (98–101). Although certain pharmacological properties, such as metabolic stability and cell permeability, represent major challenges for drug development (102), their application as e.g., biocontrol agents and food preservatives is under investigation or already reached commercialization (16, 103). The chitinopeptins thereby exhibit similar activities against *Candida* as CLPs of the surfactin or iturin (104) and fengycin (15) group. Further studies are necessary to investigate their specific activity profile.

## MATERIALS AND METHODS

**Genomic data processing.** Genomic data were collected until January 2021. To that date, almost all publicly, complete, and annotated genomes in addition with some WGS projects (<100 scaffolds) of the phylum Bacteroidetes (600 genomes in total, Table S8) were analyzed using antiSMASH 5.0 (50) and BiG-SCAPE v1.0.0 (26). All Bacteroidetes genome sequences used in this study were downloaded from the National Center for Biotechnology Information (NCBI) database or the Department of Energy (DOE) Joint Genome Institute–Integrated Microbial Genomes & Microbiomes (JGI IMG) database (105, 106). The data set consists of all six phylogenetic classes with the following distribution: *Bacteroidia* 18% ( $n = 108$ ), *Chitinophagia* 13.7% ( $n = 82$ ), *Cytophagia* 11.5% ( $n = 69$ ), *Flavobacteriia* 50.2% ( $n = 301$ ), *Saprosipria* 0.3% ( $n = 2$ ), *Sphingobacteriia* 6.3% ( $n = 37$ ).

**Phylogenetic tree reconstruction.** For the phylogenetic tree of the complete Bacteroidetes phylum, we aligned the extracted 16S rRNA genes from each genome assembly using MAFFT (107). A maximum likelihood phylogeny was built based on the 16S rRNA genes using the program RAxML v.8.2.11 (108) with



a general time reversible (GTR) nucleotide substitution model (109) and 1,000 bootstrap replicates. We used the Interactive Tree of Life (iTol v4) (110) to visualize the phylogenetic tree. The labels of each branch are color-coded by class level. A strip data set represents the genome size of each strain, from light gray for small sizes to black for large genomes. A multivalued bar chart placed around the tree displays the total BGC amount of each strain, in parallel highlighting the specific amount of NRPS, PKS, and the hybrid BGCs of both types. Partial BGCs on contigs <10 kb of WGS projects are not included in this analysis.

The phylogenetic tree of the genus *Chitinophaga* is based on a clustal W alignment (111) of available 16S rRNA gene sequences of 25 *Chitinophaga* strains available for cultivation. The tree was calculated using MEGA v7.0.26 with the maximum-likelihood method and GTR-Gamma model (112). Percentage on the tree branches indicate values of 1,000 bootstrap replicates with a bootstrap support of more than 50%. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site.

**BiG-SCAPE-CORASON analysis.** Individual .gb files were processed using antiSMASH 5.0, including the ClusterFinder border prediction algorithm to automatically trim BGCs where gene cluster borders were possible to predict (23, 50). The “hybrids” mode of BiG-SCAPE v1.0.0 (26), which allows BGCs with mixed annotations to be analyzed together, was enabled. Several cutoffs of 0.1 to 0.9 were tested. After manual inspection, the network with a cutoff of 0.6 was visualized and annotated using Cytoscape v3.6.0 (113).

**Mass spectrometric analysis.** A quadrupole time-of-flight spectrometer (LC-QTOF maxis II, Bruker Daltonik) equipped with an electrospray ionization source in line with an Agilent 1290 infinity LC system (Agilent) was used for all UHPLC-QTOF-HR-MS and MS/MS measurements. C18 RP-UHPLC (ACQUITY UPLC BEH C18 column [130 Å, 1.7 μm, 2.1 × 100 mm]) was performed at 45°C with the following linear gradient: 0 min: 95% A; 0.30 min: 95% A; 18.00 min: 4.75% A; 18.10 min: 0% A; 22.50 min: 0% A; 22.60 min: 95% A; 25.00 min: 95% A (A: H<sub>2</sub>O, 0.1% HCOOH; B: CH<sub>3</sub>CN, 0.1% HCOOH; flow rate: 0.6 mL/min). Mass spectral data were acquired using a 50 to 2,000 *m/z* scan range at 1 Hz scan rate. MS/MS experiments were performed with 6 Hz and the top five most intense ions in each full MS spectrum were targeted for fragmentation by higher-energy collisional dissociation at 25 eV or 55 eV using N<sub>2</sub> at 10<sup>-2</sup> mbar. Precursors were excluded after two spectra, released after 0.5 min and reconsidered if the intensity of an excluded precursor increased by factor 1.5 or more.

**Chemotype-barcoding matrix. (i) Raw data processing.** Raw data processing was performed with DataAnalysis 4.4 (Bruker) using recalibration with sodium formate. RecalculateLinespectra with two thresholds (10,000 and 20,000) and subsequent FindMolecularFeatures (0.5 to 25 min, S/N = 0, minimal compound length = 8 spectra, smoothing width = 2, correlation coefficient threshold = 0.7) was performed. Bucketing was performed using ProfileAnalysis 2.3 (Bruker; 30 to 1,080 s, 100 to 6,000 *m/z*, Advanced Bucketing with 24 s ppm, no transformation, Bucketing basis = H<sup>+</sup>).

**(ii) Data management.** The complete data set was copied and processed with both line spectra thresholds (10,000 and 20,000). Bucketing was performed on both sets at the same time, resulting in one table containing all buckets deemed identical for every sample (under both thresholds). This table was subsequently curated: Buckets were only deemed present, if they were detected in the 20,000 samples (4,707 buckets did not meet this criteria and were deleted). To avoid “uniqueness” due to compounds being just above the detection threshold, entries from the 10,000 set were used for these buckets, the 20,000 entries were deleted. Eight buckets were not filled in this table and were subsequently deleted, resulting in a final table of 278 samples with 4,188 buckets.

For preparation of the barcode, for each strain (and combined control) and fermentation length (4 and 7 days) how often each bucket was present was analyzed. Because media controls were repeated several times (instead of once per strain/duration combination), they were combined first. Each bucket present in one of the control/time/media combinations was deemed present.

Classifications were “more than one” for buckets present in two or more media; “media X” for buckets present in only one media and “none” for all not present at that strain/duration combination. This table is the input table for the barcode matrix. For each strain, it was calculated if a bucket was present in at least one of its media/duration combinations. Subsequently, how many strains (excluding media controls) each bucket was present was calculated.

To achieve a readable representation, the buckets in the input table were sorted (after transformation/rotation) (i) (alphabetically) for each sample (for defragmentation of media), (ii) (decreasing) for the number of strains in which the bucket is present, and (iii) presence in any of the control conditions. Visualization was performed with the R-script deposited in the SI. All MS and MS/MS-raw data and used data sets are deposited at <http://dx.doi.org/10.24406/fordatis/188>.

**Cultivation and screening conditions. (i) Bacterial strains and culture conditions.** All 25 *Chitinophaga* strains used for this study (Table S1) were purchased from Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ) and Korean Collection for Type Cultures (KCTC). Strains were inoculated in 50 mL R2A (114) (HiMedia Laboratories, LLC.) in 300-mL flasks and incubated for 3 days, then 2% (vol/vol) culture volume was used to inoculate the main cultures (50 mL media/300 mL flask), incubated at 28°C with agitation at 180 rpm for 4 or 7 days. R2A, 3018 (1 g/L yeast extract, 5 g/L Casitone, pH 7.0, 24 mM *N*-acetylglucosamine added after autoclaving), 5294 (10 g/L soluble starch, 10 g/L glucose, 10 g/L glycerol 99%, 2.5 g/L liquid corn steep, 5 g/L peptone, 2 g/L yeast extract, 1 g/L NaCl, 3 g/L CaCO<sub>3</sub>, pH 7.2), 3021 (10 g/L glucose, 10 g/L chitin, 5 g/L soy flour, 5 g/L casein peptone, pH 7.0), and 5065 (15 g/L soluble starch, 10 g/L glucose, 10 g/L soy flour, 1 g/L yeast extract, 0.1 g/L K<sub>2</sub>HPO<sub>4</sub>, 3 g/L NaCl, pH 7.4) supplemented with 1 mL/L SL-10 trace element solution (115) and 3 mL/L vitamin solution (added after autoclaving, VL-55 medium, DSMZ) were used as main culture media.

**(ii) Screening conditions.** Freeze-dried cultures were extracted with 40-mL MeOH, dried and resuspended in 1-mL MeOH to generate 50x concentrated extracts, partially used for extract analysis by UHPLC-HR-MS (details in MS analysis part). An aliquot of 150 μL of each extract was dried again and

resuspended in 75  $\mu$ L DMSO, resulting in 100x concentrated extracts. Four concentrations (1x, 0.5x, and 0.25x twice) of each extract were screened in 384-well plate format (20  $\mu$ L assay volume) supported by liquid handling robots (Analytik Jena CyBio Well 96/384 CYBIO SW-CYBI, Thermo Fisher Scientific Matrix Wellmate and Multidrop). Assay plates were inoculated with 20,000 CFU/mL from overnight precultures (100-mL flasks filled with 30-mL cation adjusted Mueller-Hinton II Broth [MHIB, BD Difco], 37°C) diluted in MHIB of *Escherichia coli* ATCC 35218, *E. coli* ATCC 25922  $\Delta$ TolC, *Pseudomonas aeruginosa* ATCC 27853, *Klebsiella pneumoniae* ATCC 13883, or *Staphylococcus aureus* ATCC 25923 and incubated at 37°C overnight. The protocol was adapted to an inoculum of 100,000 CFU/mL for *Mycobacterium smegmatis* ATCC 607, *Candida albicans* FH2173, and 50,000 spores/mL from a spore solution for *Aspergillus flavus* ATCC 9170. Brain heart infusion (BHI) broth supplemented with 1% (vol/vol) Tween 80 for *M. smegmatis* and MHIB for *A. flavus* and *C. albicans* were used. Pre-cultures and main-cultures were incubated at 37°C for 2 days, except the pre-culture of *C. albicans* was incubated at 28°C. All assay plates and pre-cultures were incubated with agitation at 180 rpm and controlled humidity (80% RH) to prevent evaporation. Active extracts were defined as >80% growth inhibition compared with the controls. Optical density (600 nm) or for *A. flavus*, *C. albicans*, and *M. smegmatis* a quantitative ATP assay (BacTiter-Glo, Promega) based on the resulting relative light units, both measured with a LUMIstar OPTIMA Microplate Luminometer (BMG LABTECH), was used according to the manufacturer's protocol to detect growth inhibition. UHPLC-HR-ESI-MS guided fractionation helped to identify and dereplicate possible active molecules within crude extracts.

**Strain fermentation and purification of chitinopeptins.** *C. Eiseniae* DSM 22224 and *C. flava* KCTC 62435 were inoculated from plate (R2A) in 300-mL Erlenmeyer flasks filled with 100 mL R2A and incubated at 28°C with agitation at 180 rpm for 3 days. Followed by 20 L fermentations (separated in 500 mL culture volume per 2 L flasks) in 3018-medium inoculated with 2% (vol/vol) pre-culture, incubated under the same conditions for 4 days, and subsequently freeze-dried. Dried cultures were extracted with one-time culture volume MeOH. The extracts were evaporated to dryness using rotary evaporation under reduced pressure, resuspended in 3 L of 10% MeOH/H<sub>2</sub>O, and loaded onto a XAD16N column (1 L bed volume). Step-wise elution with 10%, 40%, 60%, 80%, and 100% MeOH (two-times bed volume each) was performed. The 80% and 100% fractions containing chitinopeptins were subjected to preparative followed by semi-preparative reverse-phase high-performance liquid chromatography (RP-HPLC) on C18 columns (preparative HPLC: Synergi 4  $\mu$ m Fusion-RP 80 Å [250  $\times$  21.2 mm]; semi-preparative-HPLC: Synergi 4  $\mu$ m Fusion-RP 80 Å [250  $\times$  10 mm], mobile phase: CH<sub>3</sub>CN/H<sub>2</sub>O + 0.1% HCOOH) using linear gradients of 5% to 95% organic in 40 min. Final purification was achieved using UHPLC on a ACQUITY UPLC BEH C18 column (130 Å, 1.7  $\mu$ m, 100  $\times$  2.1 mm) with the same mobile phase and a linear gradient of 30% to 60% organic in 18 min. After each step, fractions containing chitinopeptins were evaporated to dryness using a high performance evaporator (Genevac HT-12).

Chitinopeptin A (1). White, amorphous powder;  $[\alpha]_D^{20} = -27.6^\circ$  ( $c = 0.11$ , CH<sub>3</sub>OH); LC-UV (CH<sub>3</sub>CN in H<sub>2</sub>O + 0.1% HCOOH)  $\lambda_{max}$  220 nm; <sup>1</sup>H-NMR and <sup>13</sup>C-NMR data, see Table S2 and S3; HR-MS (ESI-TOF)  $m/z$  (M+H)<sup>+</sup> calcd for C<sub>82</sub>H<sub>138</sub>N<sub>17</sub>O<sub>28</sub> 1809.0037, found 1809.0052.

Chitinopeptin B (2). White, amorphous powder;  $[\alpha]_D^{20} = -21.0^\circ$  ( $c = 0.16$ , CH<sub>3</sub>OH); LC-UV [(CH<sub>3</sub>CN in H<sub>2</sub>O + 0.1% HCOOH)]  $\lambda_{max}$  222 nm; <sup>1</sup>H-NMR and <sup>13</sup>C-NMR data, see Table S2 and S3; HR-MS (ESI-TOF)  $m/z$  [M+H]<sup>+</sup> calcd for C<sub>81</sub>H<sub>136</sub>N<sub>17</sub>O<sub>28</sub> 1794.9880, found 1794.9907.

Chitinopeptin C1+C2 (3 and 4). Slightly yellowish, amorphous powder; LC-UV [(CH<sub>3</sub>CN in H<sub>2</sub>O + 0.1% HCOOH)]  $\lambda_{max}$  219 nm; <sup>1</sup>H-NMR and <sup>13</sup>C-NMR data, see Table S2 and S3; HR-MS (ESI-TOF)  $m/z$  [M+H]<sup>+</sup> calcd for C<sub>80</sub>H<sub>141</sub>N<sub>18</sub>O<sub>30</sub> 1882.0061, found 1882.0177.

Chitinopeptin D1+D2 (5 and 6). Slightly yellowish, amorphous powder; LC-UV [(CH<sub>3</sub>CN in H<sub>2</sub>O + 0.1% HCOOH)]  $\lambda_{max}$  221 nm; <sup>1</sup>H-NMR and <sup>13</sup>C-NMR data, see Table S2 and S3; HR-MS (ESI-TOF)  $m/z$  [M+H]<sup>+</sup> calcd for C<sub>83</sub>H<sub>139</sub>N<sub>18</sub>O<sub>30</sub> 1867.9905, found 1868.0059.

**Structure elucidation. (i) NMR studies.** NMR spectra of chitinopeptin A (1) were acquired on a Bruker AVANCE 700 spectrometer operating at a proton frequency of 700.13 MHz and a <sup>13</sup>C-carbon frequency of 176.05 MHz. NMR spectra of the remaining compounds were recorded on a Bruker AVANCE 500 spectrometer operating at a proton frequency of 500.30 MHz and a <sup>13</sup>C-carbon frequency of 125.82 MHz. Both instruments were equipped with a 5 mm TCI cryo probe. For structure elucidation and assignment of proton and carbon resonances 1D-<sup>1</sup>H, 1D-<sup>13</sup>C, DQF-COSY, TOCSY (mixing time 80 ms), ROESY (mixing time 150 ms), multiplicity edited-HSQC, and HMBC spectra were acquired. The raw data is deposited at <http://dx.doi.org/10.24406/fordatis/188>.

Homonuclear experiments (1D-<sup>1</sup>H, DQF-COSY, TOCSY, ROESY) were acquired in a mixture of H<sub>2</sub>O and CD<sub>3</sub>CN in a ratio of 1:1. 1D-<sup>13</sup>C, HSQC, and HMBC spectra were acquired in a mixture of D<sub>2</sub>O and CD<sub>3</sub>CN in a ratio of 1:1. In the case of CLP 1 the HMBC spectrum has also been acquired in H<sub>2</sub>O/CD<sub>3</sub>CN. <sup>1</sup>H-chemical shifts were referenced to sodium-3-(Trimethylsilyl)propionate-2,2,3,3-d<sub>4</sub>. <sup>13</sup>C-chemical shifts were referenced to the solvent signal (CD<sub>3</sub>CN, <sup>13</sup>C: 1.30 ppm).

**(ii) Advanced Marfey's analysis.** The absolute configuration of all amino acids was determined by derivatization using Marfey's reagent (72). Stock solutions of amino acid standards (50 mM in H<sub>2</sub>O), NaHCO<sub>3</sub> (1 M in H<sub>2</sub>O), and N<sub>α</sub>-(2,4-dinitro-5-fluorophenyl)-L-valinamide (L-FDVA, 70 mM in acetone) were prepared. Commercially available and synthesized standards were derivatized using molar ratios of amino acid to FDVA and NaHCO<sub>3</sub> (1/1.4/8). After stirring at 40°C for 3 h, 1 M HCl was added to obtain concentration of 170 mM to end the reaction. Samples were subsequently evaporated to dryness and dissolved in DMSO (final concentration 50 mM). L- and D-amino acids were analyzed separately using C18 RP-UHPLC-MS (A: H<sub>2</sub>O, 0.1% HCOOH; B: CH<sub>3</sub>CN, 0.1% HCOOH; flow rate: 0.6 mL/min). A linear gradient of 15% to 75% B in 35 min was applied to separate all commercially available amino acid standards except of D- and D-allo-Ile. Separation of D- and D-allo-Ile was archived using chiral HPLC (CHIRALPAK IC, 1 mL/min, 75% n-hexane, 25% isopropyl alcohol, 0.2% HCOOH). Prepared Marfey's adducts of synthesized

$\beta$ -hydroxyaspartic acids were analyzed using C18 RP-UHPLC-MS with a linear gradient of 5% to 20% B in 18 min. For the determination of stereochemistry of  $\beta$ -hydroxyphenylalanines and  $\beta$ -hydroxyisoleucines, Cbz-protected intermediates were directly subjected to acidic hydrolysis (6 M HCl, 2 h, 120°C) and subsequent treatment with Marfey's reagent on analytical scale as described before after drying. Marfey's adducts were analyzed using C18 RP-UHPLC-MS with 5% to 40% B in 18 min.

Total hydrolysis of the chitinopeptins A to D was carried out by dissolving 250  $\mu$ g of each peptide in 6 M DCl in D<sub>2</sub>O and stirring for 7 h at 160°C. The samples were subsequently evaporated to dryness. Samples were dissolved in 100  $\mu$ L H<sub>2</sub>O, derivatized with I-FDVA and analyzed using the same parameters as described before.

**(iii) Synthesis of  $\beta$ -hydroxyaspartic acids,  $\beta$ -hydroxyphenylalanines, and  $\beta$ -hydroxyisoleucines.**

The syntheses of the  $\beta$ -hydroxyamino acids were achieved by modified literature known procedures. The enantiomeric excess of all synthesized amino acids were determined by chiral HPLC on selected intermediates. The synthesized compounds were fully characterized and/or compared with literature known references. For details of syntheses and characterization see supplemental materials.

**Optical rotation.** Specific rotation was determined on a digital polarimeter (P3000, A. Krüss Optronic GmbH, Germany). Standard wavelength was the sodium d-line with 589 nm. Temperature, concentration, (g/100 mL) and solvent are reported with the determined value.

**MIC.** The MIC was determined by broth microdilution method in 96-well plates with 100  $\mu$ L assay volume per well following EUCAST instructions (116, 117) for *E. coli* ATCC 35218, *E. coli* ATCC 25922  $\Delta$ ToC, *E. coli* MG1655, *Pseudomonas aeruginosa* ATCC 27853, *P. aeruginosa* PAO750, *Klebsiella pneumoniae* ATCC 13883, *Moraxella catarrhalis* ATCC 25238, *Acinetobacter baumannii* ATCC 19606, *Bacillus subtilis* DSM 10, *Staphylococcus aureus* ATCC 25923, and *Micrococcus luteus* DSM 20030. *Listeria monocytogenes* DSM 20600 (1 day of incubation at 37°C) and *Mycobacterium smegmatis* ATCC 607 (2 days of incubation at 37°C) were grown in BHI broth supplemented with 1% (vol/vol) Tween 80. *Candida albicans* FH2173 was grown in MHIIB over 2 days of incubation at 37°C (pre-culture 2 days at 28°C). Approximate inoculation cell density for all strains was 500,000 CFU/mL, except 1,000,000 CFU/mL were used for *M. smegmatis* and *C. albicans*. For all fungal screenings 100,000 spores/mL were used and incubated at 37°C for *Aspergillus flavus* ATCC 9170 (MHIIB) and 25°C for *Z. tritici* MUCL45407 (4 g/L yeast extract, 4 g/L malt extract, 4 g/L sucrose) and *Fusarium oxysporum* ATCC 7601 (potato dextrose broth, Sigma). Growth inhibition was detected by optical density (600 nm) or for *M. luteus*, *L. monocytogenes*, *C. albicans*, *M. smegmatis*, and all fungi by quantification of free ATP using BacTiter-Glo after 24 to 48 h (72 h for *Z. tritici*). All MIC assays were performed at least in triplicate. FE (III)-citrate was added at molar ratios of 1:1 to samples in DMSO to test compounds in their iron bound conformation.

**Data availability.** MS, MS/MS, and NMR data are deposited at Fordatis - Research Data Repository of Fraunhofer-Gesellschaft (<http://dx.doi.org/10.24406/fordatis/188>).

## SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

**SUPPLEMENTAL FILE 1**, PDF file, 7.7 MB.

**SUPPLEMENTAL FILE 2**, XLSX file, 0.1 MB.

## ACKNOWLEDGMENTS

We intend to show a deep sense of gratitude to Sanja Mihajlovic as caretaker of the biobank. We also thank Mona-Katharina Bill, Sandra Semmler, Sören M. M. Schuler, Frank Förster, Christine Wehr, Jennifer Kuhn, Nadine Zucchetto, Sina Serife Abdo, Regina Zweigert, Mona Abdullahi, and Kirsten-Susann Bommersheim for their support and valuable discussions as well as Heiko Heese and Joachim Kluge for HPLC purifications of synthetic intermediates and Karin Rahn-Hotze for ee-determinations. We thank Stefan Bernhardt (Institute for Organic Chemistry, Justus-Liebig-University Giessen) for chiral HPLC analysis. This work was financially supported by the Hessen State Ministry of Higher Education, Research and the Arts (HMWK) via the state initiative for the development of scientific and economic excellence for the LOEWE Center for Insect Biotechnology and Bioresources. Sanofi-Aventis Deutschland GmbH and Evotec International GmbH funded this work in the framework of the Sanofi-Fraunhofer Natural Products Center and its follow up, the Fraunhofer-Evotec Natural Products Center of Excellence.

S.B. and M.S. conceived and designed the experiments. S.B., M.S., M.A.P., B.L., M.M., C.H., A.Bi., Y.K., C.P., and M.K. performed the experiments. S.B., M.S., M.A.P., C.P., Y.K., A.Ba., L.T., M.K., and T.F.S. analyzed the data. A.Ba., L.T., J.G., and P.E.H. initiated the project idea. J.G., M.S., and T.F.S. supervised the project. A.V. and P.E.H. initiated the public-private partnership between Fraunhofer and Sanofi (later Evotec). S.B., M.S., and T.F.S. drafted the first manuscript. S.B., M.S., and T.F.S. revised the manuscript. All authors accepted the final version of the manuscript.

## REFERENCES

- Forsberg KJ, Patel S, Gibson MK, Lauber CL, Knight R, Fierer N, Dantas G. 2014. Bacterial phylogeny structures soil resistomes across habitats. *Nature* 509:612–616. <https://doi.org/10.1038/nature13377>.
- Nesme J, Simonet P. 2015. The soil resistome: a critical review on antibiotic resistance origins, ecology and dissemination potential in telluric bacteria. *Environ Microbiol* 17:913–930. <https://doi.org/10.1111/1462-2920.12631>.
- Forsberg KJ, Reyes A, Wang B, Selleck EM, Sommer MOA, Dantas G. 2012. The shared antibiotic resistome of soil bacteria and human pathogens. *Science* 337:1107–1111. <https://doi.org/10.1126/science.1220761>.
- Finley RL, Collignon P, Larsson DGJ, McEwen SA, Li X-Z, Gaze WH, Reid-Smith R, Timinouni M, Graham DW, Topp E. 2013. The scourge of antibiotic resistance: the important role of the environment. *Clin Infect Dis* 57:704–710. <https://doi.org/10.1093/cid/cit355>.
- UN Interagency Coordination Group (IACG) on Antimicrobial Resistance. 2019. No time to wait: Securing the future from drug-resistant infections. <https://www.who.int/antimicrobial-resistance/interagency-coordination-group/final-report/en/>. Accessed June 4, 2020.
- Oerke E-C. 2006. Crop losses to pests. *J Agric Sci* 144:31–43. <https://doi.org/10.1017/S0021859605005708>.
- Kainz K, Bauer MA, Madeo F, Carmona-Gutierrez D. 2020. Fungal infections in humans: the silent crisis. *Microb Cell* 7:143–145. <https://doi.org/10.15698/mic2020.06.718>.
- Lewis K. 2013. Platforms for antibiotic discovery. *Nat Rev Drug Discov* 12:371–387. <https://doi.org/10.1038/nrd3975>.
- Schäberle TF, Hack IM. 2014. Overcoming the current deadlock in antibiotic research. *Trends Microbiol* 22:165–167. <https://doi.org/10.1016/j.tim.2013.12.007>.
- Newman DJ, Cragg GM. 2020. Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *J Nat Prod* 83:770–803. <https://doi.org/10.1021/acs.jnatprod.9b01285>.
- Schneider T, Müller A, Miess H, Gross H. 2014. Cyclic lipopeptides as antibacterial agents - potent antibiotic activity mediated by intriguing mode of action. *Int J Med Microbiol* 304:37–43. <https://doi.org/10.1016/j.ijmm.2013.08.009>.
- Nord C, Bjerketorp J, Levenfors JJ, Cao S, Strömstedt AA, Guss B, Larsson R, Hughes D, Öberg B, Broberg A. 2020. Isopropylpeptides A-H: cationic cyclic lipopeptides from *Pedobacter cryocoonitus* UP508 targeting WHO top-priority carbapenem-resistant bacteria. *ACS Chem Biol* 15:2937–2944. <https://doi.org/10.1021/acscchembio.0c00568>.
- Debono M, Barnhart M, Carrell CB, Hoffmann JA, Occolowitz JL, Abbott BJ, Fukuda DS, Hamill RL, Biemann K, Herlihy WC. 1987. A21978C, a complex of new acidic peptide antibiotics: isolation, chemistry, and mass spectral structure elucidation. *J Antibiot (Tokyo)* 40:761–777. <https://doi.org/10.7164/antibiotics.40.761>.
- Arbeit RD, Maki D, Tally FP, Campanaro E, Eisenstein BI, Daptomycin 98–01 and 99–01 Investigators. 2004. The safety and efficacy of daptomycin for the treatment of complicated skin and skin-structure infections. *Clin Infect Dis* 38:1673–1681. <https://doi.org/10.1086/420818>.
- Vanittanakom N, Loeffler W, Koch U, Jung G. 1986. Fengycin—a novel antifungal lipopeptide antibiotic produced by *Bacillus subtilis* F-29–3. *J Antibiot (Tokyo)* 39:888–901. <https://doi.org/10.7164/antibiotics.39.888>.
- Geudens N, Martins JC. 2018. Cyclic lipopeptides from *Pseudomonas* spp. - biological Swiss-Army Knives. *Front Microbiol* 9:1867. <https://doi.org/10.3389/fmicb.2018.01867>.
- Li W, Rokni-Zadeh H, de Vleeschouwer M, Ghequire MGK, Sinnaeve D, Xie G-L, Rozenski J, Madder A, Martins JC, de Mot R. 2013. The antimicrobial compound xantholysin defines a new group of *Pseudomonas* cyclic lipopeptides. *PLoS One* 8:e62946. <https://doi.org/10.1371/journal.pone.0062946>.
- Olorunleke FE, Kieu NP, Höfte M. 2015. Recent advances in pseudomonas biocontrol, p 167–198. *In* Murillo J, Vinatzer AB, Jackson WR, Arnold LW (ed), *Bacteria-plant interactions: advanced research and future trends*. Caister Academic Press UK. <https://doi.org/10.21775/9781908230584.0>.
- Raaijmakers JM, de Bruijn J, Nybroe O, Ongena M. 2010. Natural functions of lipopeptides from *Bacillus* and *Pseudomonas*: more than surfactants and antibiotics. *FEMS Microbiol Rev* 34:1037–1062. <https://doi.org/10.1111/j.1574-6976.2010.00221.x>.
- Monciardini P, Iorio M, Maffioli S, Sosio M, Donadio S. 2014. Discovering new bioactive molecules from microbial sources. *Microb Biotechnol* 7:209–220. <https://doi.org/10.1111/1751-7915.12123>.
- Adamek M, Spohn M, Stegmann E, Ziemert N. 2017. Mining bacterial genomes for secondary metabolite gene clusters. *Methods Mol Biol* 1520:23–47. [https://doi.org/10.1007/978-1-4939-6634-9\\_2](https://doi.org/10.1007/978-1-4939-6634-9_2).
- Ziemert N, Alanjary M, Weber T. 2016. The evolution of genome mining in microbes - a review. *Nat Prod Rep* 33:988–1005. <https://doi.org/10.1039/c6np00025h>.
- Cimermancic P, Medema MH, Claesen J, Kurita K, Wieland Brown LC, Mavrommatis K, Pati A, Godfrey PA, Koehrsen M, Clardy J, Birren BW, Takano E, Sali A, Lington RG, Fischbach MA. 2014. Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. *Cell* 158:412–421. <https://doi.org/10.1016/j.cell.2014.06.034>.
- Kurita KL, Lington RG. 2015. Connecting phenotype and chemotype: high-content discovery strategies for natural products research. *J Nat Prod* 78:587–596. <https://doi.org/10.1021/acs.jnatprod.5b00017>.
- Medema MH, Fischbach MA. 2015. Computational approaches to natural product discovery. *Nat Chem Biol* 11:639–648. <https://doi.org/10.1038/nchembio.1884>.
- Navarro-Muñoz JC, Selem-Mojica N, Muldowney MW, Kautsar SA, Tryon JH, Parkinson EI, de Los Santos ELC, Yeong M, Cruz-Morales P, Abubucker S, Roeters A, Lokhorst W, Fernandez-Guerra A, Cappellini LTD, Goering AW, Thomson RJ, Metcalf WW, Kelleher NL, Barona-Gomez F, Medema MH. 2020. A computational framework to explore large-scale biosynthetic diversity. *Nat Chem Biol* 16:60–68. <https://doi.org/10.1038/s41589-019-0400-9>.
- Walsh CT, Fischbach MA. 2010. Natural products version 2.0: connecting genes to molecules. *J Am Chem Soc* 132:2469–2493. <https://doi.org/10.1021/ja909118a>.
- Winter JM, Behnken S, Hertweck C. 2011. Genomics-inspired discovery of natural products. *Curr Opin Chem Biol* 15:22–31. <https://doi.org/10.1016/j.cbpa.2010.10.020>.
- van der Hoof JJJ, Mohimani H, Bauermeister A, Dorrestein PC, Duncan KR, Medema MH. 2020. Linking genomics and metabolomics to chart specialized metabolic diversity. *Chem Soc Rev* 49:3297–3314. <https://doi.org/10.1039/d0cs00162g>.
- Doroghaji JR, Albright JC, Goering AW, Ju K-S, Haines RR, Tchalukov KA, Labeda DP, Kelleher NL, Metcalf WW. 2014. A roadmap for natural product discovery based on large-scale genomics and metabolomics. *Nat Chem Biol* 10:963–968. <https://doi.org/10.1038/nchembio.1659>.
- Skinneider MA, Johnston CW, Gunabalasingam M, Merwin NJ, Kieliszek AM, MacLellan RJ, Li H, Ranieri MRM, Webster ALH, Cao MPT, Pfeifle A, Spencer N, To QH, Wallace DP, Dejong CA, Magarvey NA. 2020. Comprehensive prediction of secondary metabolite structure and biological activity from microbial genome sequences. *Nat Commun* 11:6058. <https://doi.org/10.1038/s41467-020-19986-1>.
- Hoffmann T, Krug D, Bozkurt N, Duddela S, Jansen R, Garcia R, Gerth K, Steinmetz H, Müller R. 2018. Correlating chemical diversity with taxonomic distance for discovery of natural products in myxobacteria. *Nat Commun* 9:803. <https://doi.org/10.1038/s41467-018-03184-1>.
- Behszaz B, Bode E, Gurevich A, Shi Y-N, Grundmann F, Acharya D, Caraballo-Rodríguez AM, Bouslimani A, Panitchpakdi M, Linck A, Guan C, Oh J, Dorrestein PC, Bode HB, Pevzner PA, Mohimani H. 2021. Integrating genomics and metabolomics for scalable non-ribosomal peptide discovery. *Nat Commun* 12:3225. <https://doi.org/10.1038/s41467-021-23502-4>.
- Challinor VL, Bode HB. 2015. Bioactive natural products from novel microbial sources. *Ann N Y Acad Sci* 1354:82–97. <https://doi.org/10.1111/nyas.12954>.
- Crits-Christoph A, Diamond S, Butterfield CN, Thomas BC, Banfield JF. 2018. Novel soil bacteria possess diverse genes for secondary metabolite biosynthesis. *Nature* 558:440–444. <https://doi.org/10.1038/s41586-018-0207-y>.
- Ling LL, Schneider T, Peoples AJ, Spoering AL, Engels I, Conlon BP, Mueller A, Schäberle TF, Hughes DE, Epstein S, Jones M, Lazarides L, Steadman VA, Cohen DR, Felix CR, Fetterman KA, Millett WP, Nitti AG, Zullo AM, Chen C, Lewis K. 2015. A new antibiotic kills pathogens without detectable resistance. *Nature* 517:455–459. <https://doi.org/10.1038/nature14098>.
- Imai Y, Meyer KJ, Iinishi A, Favre-Godal Q, Green R, Manuse S, Caboni M, Mori M, Niles S, Ghiglieri M, Honrao C, Ma X, Guo JJ, Makriyannis A, Linares-Otaya L, Böhringer N, Wuisan ZG, Kaur H, Wu R, Mateus A, Typas A, Savitski MM, Espinoza JL, O'Rourke A, Nelson KE, Hiller S, Noinaj N, Schäberle TF, D'Onofrio A, Lewis K. 2019. A new antibiotic selectively kills Gram-negative pathogens. *Nature* 576:459–464. <https://doi.org/10.1038/s41586-019-1791-1>.
- Thomas F, Hehemann J-H, Rebuffet E, Czjzek M, Michel G. 2011. Environmental and gut bacteroidetes: the food connection. *Front Microbiol* 2:93. <https://doi.org/10.3389/fmicb.2011.00093>.
- Mendes R, Garbeva P, Raaijmakers JM. 2013. The rhizosphere microbiome: significance of plant beneficial, plant pathogenic, and human pathogenic microorganisms. *FEMS Microbiol Rev* 37:634–663. <https://doi.org/10.1111/1574-6976.12028>.

40. Steinmetz H, Gerth K, Jansen R, Schläger N, Dehn R, Reinecke S, Kirschning A, Müller R. 2011. Elansolid A, a unique macrolide antibiotic from *Chitinophaga sancti* isolated as two stable atropisomers. *Angew Chem Int Ed Engl* 50:532–536. <https://doi.org/10.1002/anie.201005226>.
41. Mohr KI, Volz C, Jansen R, Wray V, Hoffmann J, Bernecker S, Wink J, Gerth K, Stadler M, Müller R. 2015. Pinensins: the first antifungal lantibiotics. *Angew Chem Int Ed Engl* 54:11254–11258. <https://doi.org/10.1002/anie.201500927>.
42. Katayama N, Nozaki Y, Okonogi K, Ono H, Harada S, Okazaki H. 1985. Formadimins, new monocyclic beta-lactam antibiotics of bacterial origin. I. Taxonomy, fermentation and biological activities. *J Antibiot (Tokyo)* 38: 1117–1127. <https://doi.org/10.7164/antibiotics.38.1117>.
43. Katayama N, Fukusumi S, Funabashi Y, Iwahi T, Ono H. 1993. TAN-1057 A-D, new antibiotics with potent antibacterial activity against methicillin-resistant *Staphylococcus aureus*. Taxonomy, fermentation and biological activity. *J Antibiot (Tokyo)* 46:606–613. <https://doi.org/10.7164/antibiotics.46.606>.
44. Shoji J, Hinoo H, Matsumoto K, Hattori T, Yoshida T, Matsuura S, Kondo E. 1988. Isolation and characterization of katanosins A and B. *J Antibiot (Tokyo)* 41:713–718. <https://doi.org/10.7164/antibiotics.41.713>.
45. Oku N, Adachi K, Matsuda S, Kasai H, Takatsuki A, Shizuri Y. 2008. Ariake-micins A and B, novel polyketide-peptide antibiotics from a marine gliding bacterium of the genus *Rapidithrix*. *Org Lett* 10:2481–2484. <https://doi.org/10.1021/ol8007292>.
46. Borsetto C, Amos GCA, da Rocha UN, Mitchell AL, Finn RD, Laidi RF, Vallin C, Pearce DA, Newsham KK, Wellington EMH. 2019. Microbial community drivers of PK/NRP gene diversity in selected global soils. *Microbiome* 7: 78. <https://doi.org/10.1186/s40168-019-0692-8>.
47. Paster BJ, Dewhurst FE, Olsen I, Fraser GJ. 1994. Phylogeny of bacteroides, prevotella, and porphyromonas spp. and related bacteria. *J Bacteriol* 176:725–732. <https://doi.org/10.1128/jb.176.3.725-732.1994>.
48. Hahnke RL, Meier-Kolthoff JP, García-López M, Mukherjee S, Huntemann M, Ivanova NN, Woyke T, Kyripides NC, Klenk H-P, Göker M. 2016. Genome-based taxonomic classification of bacteroidetes. *Front Microbiol* 7:2003. <https://doi.org/10.3389/fmicb.2016.02003>.
49. Hahnke RL, Meier-Kolthoff JP, García-López M, Mukherjee S, Huntemann M, Ivanova NN, Woyke T, Kyripides NC, Klenk H-P, Göker M. 2018. Corrigendum: genome-based taxonomic classification of bacteroidetes. *Front Microbiol* 9:304. <https://doi.org/10.3389/fmicb.2018.00304>.
50. Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, Medema MH, Weber T. 2019. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res* 47:W81–W87. <https://doi.org/10.1093/nar/gkz310>.
51. Belknap KC, Park CJ, Barth BM, Andam CP. 2020. Genome mining of biosynthetic and chemotherapeutic gene clusters in *Streptomyces* bacteria. *Sci Rep* 10:2003. <https://doi.org/10.1038/s41598-020-58904-9>.
52. Fajardo A, Martínez JL. 2008. Antibiotics as signals that trigger specific bacterial responses. *Curr Opin Microbiol* 11:161–167. <https://doi.org/10.1016/j.mib.2008.02.006>.
53. Wenzel SC, Müller R. 2009. The impact of genomics on the exploitation of the myxobacterial secondary metabolome. *Nat Prod Rep* 26:1385–1407. <https://doi.org/10.1039/b817073h>.
54. Medema MH, Cimemancic P, Sali A, Takano E, Fischbach MA. 2014. A systematic computational analysis of biosynthetic gene cluster evolution: lessons for engineering biosynthesis. *PLoS Comput Biol* 10:e1004016. <https://doi.org/10.1371/journal.pcbi.1004016>.
55. Medema MH, Kottmann R, Yilmaz P, Cummings M, Biggins JB, Blin K, de Bruijn I, Chooi YH, Claesen J, Coates RC, Cruz-Morales P, Duddela S, Düsterhaus S, Edwards DJ, Fewer DP, Garg N, Geiger C, Gomez-Escribano JP, Greule A, Hadjithomas M, Haines AS, Helfrich EJN, Hillwig ML, Ishida K, Jones AC, Jones CS, Jungmann K, Kegler C, Kim HU, Kötter P, Krug D, Masschelein J, Melnik AV, Mantovani SM, Monroe EA, Moore M, Moss N, Nützmann H-W, Pan G, Pati A, Petras D, Reen FJ, Rosconi F, Rui Z, Tian Z, Tobias NJ, Tsunematsu Y, Wiemann P, Wyckoff E, Yan X, Yim G, Yu F, Xie Y, Aigle B, et al. 2015. Minimum information about a biosynthetic gene cluster. *Nat Chem Biol* 11:625–631. <https://doi.org/10.1038/nchembio.1890>.
56. Fujita MJ, Nakano K, Sakai R. 2013. Bisucaberin B, a linear hydroxamate class siderophore from the marine bacterium *Tenacibaculum mesophilum*. *Molecules* 18:3917–3926. <https://doi.org/10.3390/molecules18043917>.
57. Tao L, Yao H, Kasai H, Misawa N, Cheng Q. 2006. A carotenoid synthesis gene cluster from *Algoriphagus* sp. KK10202C with a novel fusion-type lycopene beta-cyclase gene. *Mol Genet Genomics* 276:79–86. <https://doi.org/10.1007/s00438-006-0121-0>.
58. McBride NJ, Xie G, Martens EC, Lapidus A, Henrissat B, Rhodes RG, Goltsman E, Wang W, Xu J, Hunnicutt DW, Staroscik AM, Hoover TR, Cheng Y-Q, Stein JL. 2009. Novel features of the polysaccharide-digesting gliding bacterium *Flavobacterium johnsoniae* as revealed by genome sequence analysis. *Appl Environ Microbiol* 75:6864–6875. <https://doi.org/10.1128/AEM.01495-09>.
59. Schöner TA, Fuchs SW, Schöna C, Bode HB. 2014. Initiation of the flexirubin biosynthesis in *Chitinophaga pinensis*. *Microb Biotechnol* 7: 232–241. <https://doi.org/10.1111/1751-7915.12110>.
60. Singh PD, Johnson JH, Ward PC, Wells JS, Trejo WH, Sykes RB. 1983. SQ 28,332, a new monobactam produced by a *Flexibacter* sp. Taxonomy, fermentation, isolation, structure determination and biological properties. *J Antibiot (Tokyo)* 36:1245–1251. <https://doi.org/10.7164/antibiotics.36.1245>.
61. Webster ALH, Walker C, Li H, Skinnider M, Magarvey NA. [*Flexibacter*] sp. ATCC 35208 monobactam biosynthesis gene cluster. Accession no. KY502195.1, GenBank. <https://www.ncbi.nlm.nih.gov/nuccore/KY502195.1>.
62. Dehn R, Katsuyama Y, Weber A, Gerth K, Jansen R, Steinmetz H, Höfle G, Müller R, Kirschning A. 2011. Molecular basis of elansolid biosynthesis: evidence for an unprecedented quinone methide initiated intramolecular Diels-Alder cycloaddition/macrolactonization. *Angew Chem Int Ed Engl* 50:3882–3887. <https://doi.org/10.1002/anie.201006880>.
63. Gerth K, Jansen R, Mohr KI, Müller R, Volz C, Wray V. 2016. NOVEL LANTIBIOTICS. WO2016045783A1. <https://patentimages.storage.googleapis.com/e0/52/f4/2dfce4ab11326/WO2016045783A1.pdf>.
64. Teta R, Gurgui M, Helfrich EJN, Künne S, Schneider A, van Echten-Deckert G, Mangoni A, Piel J. 2010. Genome mining reveals trans-AT polyketide synthase directed antibiotic biosynthesis in the bacterial phylum bacteroidetes. *Chembiochem* 11:2506–2512. <https://doi.org/10.1002/cbic.201000542>.
65. Gerth K, Steinmetz H, Höfle G. 2009. Elansolids, novel natural metabolites of flexibacter and antibioticly active derivatives thereof. *US20110034528*.
66. Bode HB, Bethe B, Höfs R, Zeck A. 2002. Big effects from small changes: possible ways to explore nature's chemical diversity. *Chembiochem* 3:619. [https://doi.org/10.1002/1439-7633\(20020703\)3:7%3C619::AID-CBIC619%3E3.0.CO;2-9](https://doi.org/10.1002/1439-7633(20020703)3:7%3C619::AID-CBIC619%3E3.0.CO;2-9).
67. Urem M, Świątek-Połatyńska MA, Rigali S, van Wezel GP. 2016. Intertwining nutrient-sensory networks and the control of antibiotic production in *Streptomyces*. *Mol Microbiol* 102:183–195. <https://doi.org/10.1111/mmi.13464>.
68. Wohlleben W, Bera A, Mast Y, Stegmann E. 2017. Regulation of secondary metabolites of actinobacteria, p 181–232. In Wink J, Mohammadpanah F, Hamed J (ed), *Biology and biotechnology of actinobacteria*. Springer International Publishing, Cham, Switzerland.
69. Stolt R, Torgrip RJO, Lindberg J, Csenki L, Kolmert J, Schuppe-Koistinen I, Jacobsson SP. 2006. Second-order peak detection for multicomponent high-resolution LC/MS data. *Anal Chem* 78:975–983. <https://doi.org/10.1021/ac050980b>.
70. Laatsch H. 2017. *AntiBase: the natural compound identifier*, 5. Auflage. Wiley-VCH, Weinheim, Germany.
71. Somanadhan B, Kotturi SR, Yan Leong C, Glover RP, Huang Y, Flotow H, Buss AD, Lear MJ, Butler MS. 2013. Isolation and synthesis of falcitidin, a novel myxobacterial-derived acyltetrapeptide with activity against the malaria target falcipain-2. *J Antibiot* 66:259–264. <https://doi.org/10.1038/ja.2012.123>.
72. Bhushan R, Brückner H. 2004. Marfey's reagent for chiral amino acid analysis: a review. *Amino Acids* 27:231–247. <https://doi.org/10.1007/s00726-004-0118-0>.
73. Breuning A, Vicik R, Schirmeister T. 2003. An improved synthesis of aziridine-2,3-dicarboxylates via azido alcohols—epimerization studies. *Tetrahedron: Asymmetry* 14:3301–3312. <https://doi.org/10.1016/j.tetasy.2003.09.015>.
74. Blaskovich MA, Lajoie GA. 1993. Synthesis of a chiral serine aldehyde equivalent and its conversion to chiral.alpha.-amino acid derivatives. *J Am Chem Soc* 115:5021–5030. <https://doi.org/10.1021/ja00065a010>.
75. Blaskovich MA, Evindar G, Rose NGW, Wilkinson S, Luo Y, Lajoie GA. 1998. Stereoselective synthesis of Threo and Erythro  $\beta$ -Hydroxy and  $\beta$ -disubstituted- $\beta$ -hydroxy  $\alpha$ -amino acids. *J Org Chem* 63:3631–3646. <https://doi.org/10.1021/jo972294l>.
76. Hansen DB, Joullié MM. 2005. A stereoselective synthesis of (2S,3R)- $\beta$ -methoxyphenylalanine: a component of cyclomarlin A. *Tetrahedron: Asymmetry* 16: 3963–3969. <https://doi.org/10.1016/j.tetasy.2005.10.029>.
77. Hider RC, Kong X. 2010. Chemistry and biology of siderophores. *Nat Prod Rep* 27:637–657. <https://doi.org/10.1039/b906679a>.
78. Marner M, Patras MA, Kurz M, Zubeil F, Förster F, Schuler S, Bauer A, Hammann P, Vilcinskas A, Schäberle TF, Glaeser J. 2020. Molecular networking-guided discovery and characterization of stechlisins, a group of cyclic lipopeptides from a *Pseudomonas* sp. *J Nat Prod* 83:2607–2617. <https://doi.org/10.1021/acs.jnatprod.0c00263>.
79. Makris TM, Chakrabarti M, Münck E, Lipscomb JD. 2010. A family of diiron monooxygenases catalyzing amino acid beta-hydroxylation in antibiotic biosynthesis. *Proc Natl Acad Sci U S A* 107:15391–15396. <https://doi.org/10.1073/pnas.1007953107>.

80. Cheung J, Beasley FC, Liu S, Lajoie GA, Heinrichs DE. 2009. Molecular characterization of staphyloferrin B biosynthesis in *Staphylococcus aureus*. *Mol Microbiol* 74:594–608. <https://doi.org/10.1111/j.1365-2958.2009.06880.x>.
81. Beasley FC, Cheung J, Heinrichs DE. 2011. Mutation of L-2,3-diaminopropionic acid synthase genes blocks staphyloferrin B synthesis in *Staphylococcus aureus*. *BMC Microbiol* 11:199. <https://doi.org/10.1186/1471-2180-11-199>.
82. Carter JH, Du Bus RH, Dyer JR, Floyd JC, Rice KC, Shaw PD. 1974. Biosynthesis of viomycin. I. Origin of alpha, beta-diaminopropionic acid and serine. *Biochemistry* 13:1221–1227. <https://doi.org/10.1021/bi00703a026>.
83. Wang M, Gould SJ. 1993. Biosynthesis of capreomycin. 2. Incorporation of L-serine, L-alanine, and L-2,3-diaminopropionic acid. *J Org Chem* 58:5176–5180. <https://doi.org/10.1021/jo00071a029>.
84. Felnagle EA, Rondon MR, Berti AD, Crosby HA, Thomas MG. 2007. Identification of the biosynthetic gene cluster and an additional gene for resistance to the antituberculosis drug capreomycin. *Appl Environ Microbiol* 73:4162–4170. <https://doi.org/10.1128/AEM.00485-07>.
85. Zhao C, Song C, Luo Y, Yu Z, Sun M. 2008. L-2,3-diaminopropionate: one of the building blocks for the biosynthesis of Zwittermicin A in *Bacillus thuringiensis* subsp. kurstaki strain YBT-1520. *FEBS Lett* 582:3125–3131. <https://doi.org/10.1016/j.febslet.2008.07.054>.
86. Spohn M, Wohlleben W, Stegmann E. 2016. Elucidation of the zinc-dependent regulation in *Amycolatopsis japonicum* enabled the identification of the ethylenediamine-disuccinate (S,S-EDDS) genes. *Environ Microbiol* 18:1249–1263. <https://doi.org/10.1111/1462-2920.13159>.
87. Spohn M, Edenhart S, Alanjary M, Ziemert N, Wibberg D, Kalinowski J, Niedermeyer THJ, Stegmann E, Wohlleben W. 2018. Identification of a novel aminopolycarboxylic acid siderophore gene cluster encoding the biosynthesis of ethylenediaminesuccinic acid hydroxyarginine (EDHA). *Metallomics* 10:722–734. <https://doi.org/10.1039/c8mt00009c>.
88. Stachelhaus T, Mootz HD, Marahiel MA. 1999. The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem Biol* 6:493–505. [https://doi.org/10.1016/S1074-5521\(99\)80082-9](https://doi.org/10.1016/S1074-5521(99)80082-9).
89. Goering AW, McClure RA, Doroghazi JR, Albright JC, Haverland NA, Zhang Y, Ju K-S, Thomson RJ, Metcalf WW, Kelleher NL. 2016. Metabologomics: correlation of microbial gene clusters with metabolites drives discovery of a nonribosomal peptide with an unusual amino acid monomer. *ACS Cent Sci* 2:99–108. <https://doi.org/10.1021/acscentsci.5b00331>.
90. van Santen JA, Kautsar SA, Medema MH, Linington RG. 2021. Microbial natural product databases: moving forward in the multi-omics era. *Nat Prod Rep* 38:264–278. <https://doi.org/10.1039/d0np00053a>.
91. Jarmusch SA, van der Hooft JJ, Dorrestein PC, Jarmusch AK. 2021. Advances in capturing and mining mass spectrometry data are transforming natural products research. *Nat Prod Rep* 38:2066–2082. <https://doi.org/10.1039/d1np00040c>.
92. Baltz RH. 2017. Gifted microbes for genome mining and natural product discovery. *J Ind Microbiol Biotechnol* 44:573–588. <https://doi.org/10.1007/s10295-016-1815-x>.
93. Adamek M, Alanjary M, Sales-Ortells H, Goodfellow M, Bull AT, Winkler A, Wibberg D, Kalinowski J, Ziemert N. 2018. Comparative genomics reveals phylogenetic distribution patterns of secondary metabolites in *Amycolatopsis* species. *BMC Genomics* 19:426. <https://doi.org/10.1186/s12864-018-4809-4>.
94. Buckingham J. 2021. Dictionary of natural products 29.1 Chemical Search. <http://dnp.chemnetbase.com/faces/chemical/ChemicalSearch.xhtml>. Accessed January 19, 2021.
95. Janek T, Rodrigues LR, Gudiña EJ, Czyżnikowska Ż. 2016. Structure and mode of action of cyclic lipopeptide pseudofactin II with divalent metal ions. *Colloids Surf B Biointerfaces* 146:498–506. <https://doi.org/10.1016/j.colsurfb.2016.06.055>.
96. Kszepolska E, Gabaldón T. 2018. Evolutionary emergence of drug resistance in candida opportunistic pathogens. *Genes (Basel)* 9:461. <https://doi.org/10.3390/genes9090461>.
97. Odds FC, Brown AJP, Gow NAR. 2003. Antifungal agents: mechanisms of action. *Trends Microbiol* 11:272–279. [https://doi.org/10.1016/S0966-842X\(03\)00117-3](https://doi.org/10.1016/S0966-842X(03)00117-3).
98. Maget-Dana R, Ptak M, Peypoux F, Michel G. 1985. Pore-forming properties of iturin A, a lipopeptide antibiotic. *Biochimica et Biophysica Acta (BBA) - Biomembranes* 815:405–409. [https://doi.org/10.1016/0005-2736\(85\)90367-0](https://doi.org/10.1016/0005-2736(85)90367-0).
99. Maget-Dana R, Harnois I, Ptak M. 1989. Interactions of the lipopeptide antifungal iturin A with lipids in mixed monolayers. *Biochimica et Biophysica Acta (BBA) - Biomembranes* 981:309–314. [https://doi.org/10.1016/0005-2736\(89\)90042-4](https://doi.org/10.1016/0005-2736(89)90042-4).
100. Deleu M, Paquot M, Nylander T. 2005. Fengycin interaction with lipid monolayers at the air-aqueous interface-implications for the effect of fengycin on biological membranes. *J Colloid Interface Sci* 283:358–365. <https://doi.org/10.1016/j.jcis.2004.09.036>.
101. Deleu M, Paquot M, Nylander T. 2008. Effect of fengycin, a lipopeptide produced by *Bacillus subtilis*, on model biomembranes. *Biophys J* 94:2667–2679. <https://doi.org/10.1529/biophysj.107.114090>.
102. Vinogradov AA, Yin Y, Suga H. 2019. Macrocyclic peptides as drug candidates: recent progress and remaining challenges. *J Am Chem Soc* 141:4167–4181. <https://doi.org/10.1021/jacs.8b13178>.
103. Meena KR, Kanwar SS. 2015. Lipopeptides as the antifungal and antibacterial agents: applications in food safety and therapeutics. *Biomed Res Int* 2015:473050. <https://doi.org/10.1155/2015/473050>.
104. Mhammedi A, Peypoux F, Besson F, Michel G. 1982. Bacillomycin F, a new antibiotic of iturin group: isolation and characterization. *J Antibiot (Tokyo)* 35:306–311. <https://doi.org/10.7164/antibiotics.35.306>.
105. NCBI Resource Coordinators. 2018. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* 46:D8–D13. <https://doi.org/10.1093/nar/gkx1095>.
106. Markowitz VM, Chen I-MA, Palaniappan K, Chu K, Szeto E, Grechkin Y, Ratner A, Jacob B, Huang J, Williams P, Huntemann M, Anderson I, Mavromatis K, Ivanova NN, Kyrpides NC. 2012. IMG: the integrated microbial genomes database and comparative analysis system. *Nucleic Acids Res* 40:D115–122. <https://doi.org/10.1093/nar/gkr1044>.
107. Katoh K, Misawa K, Kuma K-i, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 30:3059–3066. <https://doi.org/10.1093/nar/gkf436>.
108. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>.
109. Tavare S. 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. American Mathematical Society: Lectures on Mathematics in the Life Sciences 17:57–86.
110. Letunic I, Bork P. 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* 47:W256–W259. <https://doi.org/10.1093/nar/gkz239>.
111. Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680. <https://doi.org/10.1093/nar/22.22.4673>.
112. Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 33:1870–1874. <https://doi.org/10.1093/molbev/msw054>.
113. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498–2504. <https://doi.org/10.1101/gr.1239303>.
114. Reasoner DJ, Geldreich EE. 1985. A new medium for the enumeration and subculture of bacteria from potable water. *Appl Environ Microbiol* 49:1–7. <https://doi.org/10.1128/aem.49.1.1-7.1985>.
115. Widdel F, Kohring G-W, Mayer F. 1983. Studies on dissimilatory sulfate-reducing bacteria that decompose fatty acids. *Arch Microbiol* 134:286–294. <https://doi.org/10.1007/BF00407804>.
116. EUCAST. 2003. Determination of minimum inhibitory concentrations (MICs) of antibacterial agents by broth dilution. *Clin Microbiol Infect* 9:ix–xv. <https://doi.org/10.1046/j.1469-0691.2003.00790.x>.
117. Tudela-Rodríguez LJ, Arendrup CM, Barchiesi F, Bille J, Chrystanthou E, Estrella-Cuenca M, Dannaoui E, Denning DW, Donnelly JP, Dromer F, Fegeler W, Lass-Flörl C, Moore C, Richardson M, Sandven P, Velegriaki A, Verweij P. 2008. Antifungal Susceptibility Testing (AFST) of the ESCMID European Committee for Antimicrobial Susceptibility Testing (EUCAST): method for the determination of broth dilution MICs of antifungal agents for fermentative yeasts. *Clin Microbiol Infect* 14:398–405. <https://doi.org/10.1111/j.1469-0691.2007.01935.x>.

## Publication IV



Article

### **Novel Glycerophospholipid, Lipo- and *N*-acyl Amino Acids from Bacteroidetes: Isolation, Structure Elucidation and Bioactivity**

Mona-Katharina Bill <sup>1,†</sup>, Stephan Brinkmann <sup>1,†</sup>, Markus Oberpaul <sup>1</sup>, Maria A. Patras <sup>1</sup>, Benedikt Leis <sup>1</sup>, Michael Marner <sup>1</sup>, Marc-Philippe Maitre <sup>2</sup>, Peter E. Hammann <sup>3,4</sup>, Andreas Vilcinskas <sup>1,5</sup>, Sören M. M. Schuler <sup>4,\*</sup> and Till F. Schäberle <sup>1,5,\*</sup>

- <sup>1</sup> Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Branch for Bioresources, 35392 Giessen, Germany; mona.bill@ime.fraunhofer.de (M.-K.B.); stephan.brinkmann@ime.fraunhofer.de (S.B.); markus.oberpaul@ime.fraunhofer.de (M.O.); maria.patras@ime.fraunhofer.de (M.A.P.); benedikt.leis@gmx.net (B.L.); michael.marner@ime.fraunhofer.de (M.M.); andreas.vilcinskas@ime.fraunhofer.de (A.V.)  
<sup>2</sup> Sanofi Pasteur, R&D, 69280 Marcy L'Etoile, France; MarcPhilippe.Maitre@sanofi.com  
<sup>3</sup> Sanofi-Aventis Deutschland GmbH, R&D, 65926 Frankfurt am Main, Germany; peter.hammann@npconsult.me  
<sup>4</sup> Evotec International GmbH, 37079 Göttingen, Germany  
<sup>5</sup> Institute for Insect Biotechnology, Justus-Liebig-University of Giessen, 35392 Giessen, Germany  
\* Correspondence: soeren.schuler@evotec.com (S.M.M.S.); till.f.schaerberle@agr.uni-giessen.de (T.F.S.)  
† These authors contributed equally to this work.

**Copyright:** This article is Open Access (CC BY 4.0).

**Supporting Information** (not included in this dissertation file):

<https://www.mdpi.com/1420-3049/26/17/5195/s1>

#### **Contribution:**

Mona-Katharina Bill and Stephan Brinkmann conceived and designed the study. Stephan Brinkmann identified and characterized the lipoamino acids via mass spectrometry and isolated some of these. Both of them analyzed the data and prepared the figures for the manuscript. Mona-Katharina Bill and Stephan Brinkmann wrote the draft of the manuscript as well as the following revisions.

## Article

# Novel Glycerophospholipid, Lipo- and N-acyl Amino Acids from Bacteroidetes: Isolation, Structure Elucidation and Bioactivity

Mona-Katharina Bill <sup>1,†</sup>, Stephan Brinkmann <sup>1,†</sup>, Markus Oberpaul <sup>1</sup>, Maria A. Patras <sup>1</sup>, Benedikt Leis <sup>1</sup>, Michael Marner <sup>1</sup>, Marc-Philippe Maitre <sup>2</sup>, Peter E. Hammann <sup>3,4</sup>, Andreas Vilcinskas <sup>1,5</sup>, Sören M. M. Schuler <sup>4,\*</sup> and Till F. Schäberle <sup>1,5,\*</sup>

<sup>1</sup> Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Branch for Bioresources, 35392 Giessen, Germany; mona.bill@ime.fraunhofer.de (M.-K.B.); stephan.brinkmann@ime.fraunhofer.de (S.B.); markus.oberpaul@ime.fraunhofer.de (M.O.); maria.patras@ime.fraunhofer.de (M.A.P.); benedikt.leis@gmx.net (B.L.); michael.marner@ime.fraunhofer.de (M.M.); andreas.vilcinskas@ime.fraunhofer.de (A.V.)

<sup>2</sup> Sanofi Pasteur, R&D, 69280 Marcy L'Etoile, France; MarcPhilippe.Maitre@sanofi.com

<sup>3</sup> Sanofi-Aventis Deutschland GmbH, R&D, 65926 Frankfurt am Main, Germany; peter.hammann@npconsult.me

<sup>4</sup> Evotec International GmbH, 37079 Göttingen, Germany

<sup>5</sup> Institute for Insect Biotechnology, Justus-Liebig-University of Giessen, 35392 Giessen, Germany

\* Correspondence: soeren.schuler@evotec.com (S.M.M.S.); till.f.schaerberle@agr.uni-giessen.de (T.F.S.)

† These authors contributed equally to this work.



**Citation:** Bill, M.-K.; Brinkmann, S.; Oberpaul, M.; Patras, M.A.; Leis, B.; Marner, M.; Maitre, M.-P.; Hammann, P.E.; Vilcinskas, A.; Schuler, S.M.M.; et al. Novel Glycerophospholipid, Lipo- and N-acyl Amino Acids from Bacteroidetes: Isolation, Structure Elucidation and Bioactivity. *Molecules* **2021**, *26*, 5195. <https://doi.org/10.3390/molecules26175195>

Academic Editor: Natalizia Miceli

Received: 31 July 2021

Accepted: 25 August 2021

Published: 27 August 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** The ‘core’ metabolome of the Bacteroidetes genus *Chitinophaga* was recently discovered to consist of only seven metabolites. A structural relationship in terms of shared lipid moieties among four of them was postulated. Here, structure elucidation and characterization via ultra-high resolution mass spectrometry (UHR-MS) and nuclear magnetic resonance (NMR) spectroscopy of those four lipids (two lipoamino acids (LAAs), two lysophosphatidylethanolamines (LPEs)), as well as several other undescribed LAAs and N-acyl amino acids (NAAAs), identified during isolation were carried out. The LAAs represent closely related analogs of the literature-known LAAs, such as the glycine-serine dipeptide lipids 430 (2) and 654. Most of the here characterized LAAs (1, 5–11) are members of a so far undescribed glycine-serine-ornithine tripeptide lipid family. Moreover, this study reports three novel NAAAs (N-(5-methyl)hexanoyl tyrosine (14) and N-(7-methyl)octanoyl tyrosine (15) or phenylalanine (16)) from *Olivibacter* sp. FHG000416, another Bacteroidetes strain initially selected as best in-house producer for isolation of lipid 430. Antimicrobial profiling revealed most isolated LAAs (1–3) and the two LPE ‘core’ metabolites (12, 13) active against the Gram-negative pathogen *M. catarrhalis* ATCC 25238 and the Gram-positive bacterium *M. luteus* DSM 20030. For LAA 1, additional growth inhibition activity against *B. subtilis* DSM 10 was observed.

**Keywords:** linear lipoamino acid; lipid 430; lipid 654; N-acyl amino acid; lysophosphatidylethanolamine; bacteroidetes; *Chitinophaga*; *Olivibacter*; LC-MS/MS; antimicrobial lipids

## 1. Introduction

Lipids are a diverse group of natural biomolecules. Thousands of distinct lipids, such as glycerolipids, sterol lipids, sphingolipids, lipoamino acids (LAAs), and phospholipids, are ubiquitous in all organisms. Each of them is chemically unique, and they exhibit different biological functions. Given the diversity in both the chemical and physical properties of lipids and the fact that each lipid type is involved at various stages of cellular processes, the definition of lipid function besides their primary biological role, i.e., the formation of cell membrane matrixes, is challenging. Described functions in cellular signaling, energy storage, or an implication as substrate for metabolite or protein lipidation are only examples [1].



Bacterial membrane composition differs among bacterial species and depends on the exposed environmental conditions [2]. In most cases and conditions studied, amphiphilic lipids such as glycerophospholipids are composed of two fatty acids, a glycerol moiety, a phosphate group, and variable head groups. Common examples are phosphatidylethanolamine (PE), phosphatidylglycerol (PG), cardiolipin (CL), lysyl-phosphatidylglycerol (LPG), phosphatidylinositol (PI), phosphatidic acid (PA), and phosphatidylserine (PS). Additionally, a small fraction of the bacterial membrane consists of lysophospholipids (LPLs). They are metabolic intermediates of bacterial phospholipid synthesis or they derive from membrane degradation. The most abundant member of this class is lysophosphatidylethanolamine (LPE). LPLs result from partial hydrolysis of phospholipids mediated by phospholipase A [3]. Other phosphorus-free lipids, such as sulfolipids, LAAs [2,4–14], or the growing lipid class of the *N*-acyl amino acids (NAAAs), are also reported [15]. The latter are found in all biological systems, but their functions remain unclear. It is hypothesized that these lipids are putative signaling molecules with a wide range of biological activities [16–23].

In general, a variety of physicochemical properties and bioactivities have been associated with all kinds of bacterial lipids, including hemagglutination [9], macrophage activation [10], bacterial virulence factors [24], involvement in the development of multiple sclerosis [25], and antimicrobial activities [26,27]. Moreover, simple representatives of the LAAs such as the glycine-serine dipeptide lipid 654 (also referred to as flavolipin [28], topostin D654 [29], and WB-3559 D [6,7]) and glycine-serine dipeptide lipid 430 were isolated from pathogenic Bacteroidetes strains associated with chronic periodontitis [30]. They account for osteoclast formation from RAW cells [31] and TLR2-dependent inhibition of osteoblast differentiation and function [32], and they are implicated in dendritic cell release of IL-6 mediated through engagement of TLR2 [33,34]. Furthermore, antimicrobial activity was reported for lipid 430 [35], the de-esterified enzymatic hydrolysis product of lipid 654 hydrolyzed by phospholipase A2 (PLA2) [36].

For enabling doubtless structure elucidation, various analytical methods such as MS and NMR are well established and intensively used in the field of natural product research in general. For lipids as one important sub-class, first, gas or liquid chromatography coupled to diverse mass spectrometry methods represent two key analytical techniques [37–39]. Second, NMR spectroscopy is not only widely utilized for structure elucidations of single lipids, successful applications regarding qualitative and quantitative analysis of lipids in complex mixtures are also reported in this context [40–45].

In a previous study, we identified the ‘core’ metabolome of the Bacteroidetes genus *Chitinophaga*, consisting of only seven metabolites [46]. We postulated a structural relationship among four of them based on their MS/MS fragmentation pattern, which suggested them to be unknown LAAs and LPEs. Here, we report the identification in total of 16 diverse lipids (11 LAAs, 2 LPEs, 3 NAAAs) produced by *Chitinophaga* spp. and *Olivibacter* sp. FHG000416, the ‘core’ ones included. Isolation and structure elucidation of nine of those lipids isolated from *Chitinophaga eiseniae* DSM 22224 [47] and FHG000416 was successfully achieved. Based on the four *Chitinophaga* core lipids, the known lipid 430, several novel LAA analogs thereof, and three undescribed NAAAs were characterized. In this context, we identified a novel glycine-serine-ornithine LAA family that is closely related to the previously described glycine and glycine-serine LAA families that lipid 430 and 654 belong to. Investigating them for their antimicrobial activity revealed a common growth inhibition effect against the Gram-negative pathogen *M. catarrhalis* ATCC 25238 and the Gram-positive bacterium *M. luteus* DSM 20030, with LAA 1 as the most potent one, also showing growth-inhibiting activity against *B. subtilis* DSM 10.

## 2. Results

### 2.1. Lipoamino Acids

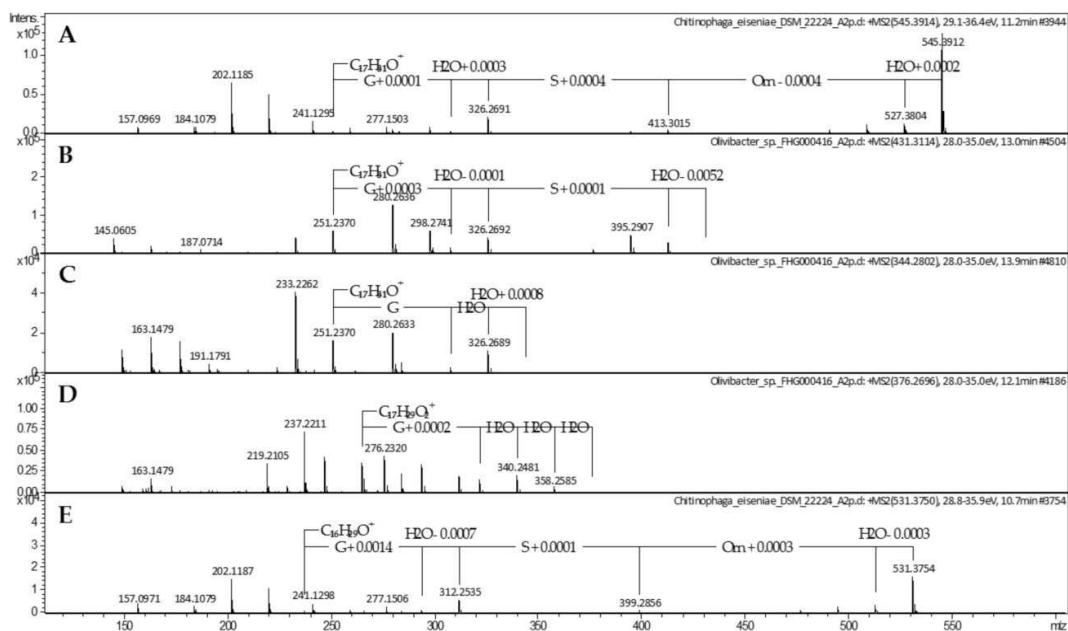
A previous study revealed a small core metabolome of the Bacteroidetes genus *Chitinophaga*, identified as a genus of underexplored talented producers of natural products. The analyzed 25 strains commonly share only 7 of the 2736 identified metabolite buckets. Four of those seven 'core' metabolites correspond to the UHR-ESI-MS ion peaks at  $m/z$  545.3908 [M + H]<sup>+</sup>, 452.2769 [M + H]<sup>+</sup>, 440.2767 [M + H]<sup>+</sup>, and 767.5896 [M + H]<sup>+</sup>, with retention times (RTs) of 11.2, 11.7, 12.1, and 17.4 min, respectively (Table S1). Based on their MS/MS fragmentation pattern, similar lipid moieties were postulated, suggesting a structural relationship [46].

For the first core metabolite, the molecular formula C<sub>27</sub>H<sub>52</sub>N<sub>4</sub>O<sub>7</sub> (**1**) was assigned according to the UHR-ESI-MS ion peak at  $m/z$  545.3908 [M + H]<sup>+</sup>. The MS/MS fragmentation of **1** revealed neutral losses of H<sub>2</sub>O, followed by the three amino acids ornithine, serine, and glycine, resulting in the fragment ion of  $m/z$  251.2366 [M + H]<sup>+</sup>. The corresponding ion formula of C<sub>17</sub>H<sub>31</sub>O<sup>+</sup> indicated a fatty acyl group based on the molecular composition and apparent carbon-to-hydrogen ratio (Figure 1A). A missing hit in a database search indicated the potential novelty of LAA **1**, containing a glycine-serine-ornithine tripeptide. However, very close relatives, such as lipid 430 (**2**) and lipid 654, belong to the glycine-serine dipeptide LAA family, known to be biosynthesized by several bacteria, including representatives of the phylum Bacteroidetes [14,30,34]. Therefore, the metabolomics data generated in our previous study was examined for the presence of both dipeptide lipids. A corresponding ion peak for lipid 430 (**2**) at  $m/z$  431.3114 [M + H]<sup>+</sup> with a RT of 13.0 min was identified in 23 of the analyzed 25 *Chitinophaga* metabolomes (Table S1). Similar serial neutral losses of H<sub>2</sub>O, serine, and glycine, together with the same remaining fragment ion of  $m/z$  251.2367 [M + H]<sup>+</sup>, strongly suggested **1** being a close derivative of **2**, expanded by an additional ornithine residue (Figure 1B). To provide a sufficient amount of the compounds for structure confirmation via NMR, our in-house Bacteroidetes-based extract library was examined for enhanced production of **1** and **2**. Lipid 430 (**2**) was enriched in extracts of strain FHG000416, which was assigned by 16S rRNA sequence analysis to the Bacteroidetes genus *Olivibacter* due to a sequence identity of ~94.5% towards *Olivibacter domesticus* DSM 18733 [48,49]. Extracts from both strains, *C. eiseniae* DSM 22224 and FHG000416, were considered as starting points for compound isolation. NMR analysis confirmed compound **2** to be lipid 430 since the data were in good agreement with the literature (Table 1) [35]. Compared thereto, <sup>1</sup>H and 2D spectra of **1** were highly similar. Accordingly, the fatty acyl motif linked to glycine and serine was identified as *iso*-heptadecanoic acid (*iso*-C<sub>17:0</sub>). An additional amide proton 2'''-NH ( $\delta_{\text{H}}$  7.66–7.59 ppm) was observed showing COSY correlation to methine proton H-2''' ( $\delta_{\text{H}}$  3.82–3.75 ppm,  $\delta_{\text{C}}$  53.6 ppm). Further correlations to methylene protons H-3''', H-4''', and H-5''' confirmed the presence of a C-terminal ornithine residue, as postulated based on the MS/MS fragmentation, identifying lipid **1** as 5-amino-2-(3-hydroxy-2-(2-(3-hydroxy-15-methylhexadecanamido)acetamido)propanamido)pentanoic acid (Figure 2; Table 1).

The stereochemistry of the chiral amino acids incorporated in LAA **1** and lipid 430 (**2**) was determined by advanced Marfey's analysis, using *N*<sub>α</sub>-(2,4-dinitro-5-fluorophenyl)-L-valinamide (L-FDVA) as Marfey's reagent [50]. D- and L-enantiomers of serine and ornithine served as RT references. In agreement with published data, the serine residue of lipid 430 (**2**) was identified as L-serine (Figure S1) [34]. This was also true for lipid **1**, which furthermore contains L-ornithine (Figure S1; Figure S2).

During processing, the extracts of FHG000416, two further derivatives of **1** and **2**, were recognized and identified based on their similar MS/MS fragmentation patterns. The two ions of  $m/z$  344.2802 [M + H]<sup>+</sup> and  $m/z$  376.2696 [M + H]<sup>+</sup> were assigned to putative lipids with molecular formulae C<sub>19</sub>H<sub>37</sub>NO<sub>4</sub> (**3**) and C<sub>19</sub>H<sub>37</sub>NO<sub>6</sub> (**4**). Compared with **1** and **2**, the MS/MS fragmentation of **3** showed the identical resulting fragment ion at  $m/z$  251.2370 [M + H]<sup>+</sup> after neutral losses of two water and one glycine molecule (Figure 1C).

Therefore, we assumed that **1**, **2**, and **3** share the same acyl chain and glycine residue. In contrast, the MS/MS fragmentation pattern of **4** also showed admittedly the neutral loss of a glycine but also the loss of several water molecules. The resulting fragment ion at  $m/z$  265.2162  $[M + H]^+$  corresponded to the formula of  $C_{17}H_{29}O_2^+$ , indicating modifications in the acyl residue (Figure 1D). Finally, 1D and 2D-NMR spectra confirmed an *iso*- $C_{17:0}$  aliphatic acyl group linked to glycine via a peptide bond (Table 1). In the case of lipid **3**, the  $\alpha$ -methylene protons of the acyl group H-2 ( $\delta_H$  2.19 ppm) showed COSY correlation with the single methine proton H-3 ( $\delta_H$  3.79–3.74 ppm), suggesting substitution. The chemical shift of the HSQC-correlating carbon atom C-3 ( $\delta_C$  67.4 ppm) furthermore supported the hypothesis of a hydroxyl group being attached in  $\beta$ -position (Table 1). Thus, **3** was identified as (3-hydroxy-15-methylhexadecanoyl)glycine (Figure 2). For derivative **4**, NMR spectra undergirded the expected similarities, except for the methine proton H-3 ( $\delta_H$  3.44 ppm,  $\delta_C$  74.5 ppm) showing COSY correlations with methine instead of methylene protons in  $\alpha$ - (H-2,  $\delta_H$  4.19 ppm) and  $\gamma$ -position (H-4,  $\delta_H$  3.40–3.35 ppm) of the acyl chain. In agreement with the chemical shift of the HSQC-correlating carbon atoms C-2 and C-4 ( $\delta_C$ -2 70.9 ppm,  $\delta_C$ -4 69.5 ppm), hydroxylation was determined at both sites, thereby identifying **4** as (2,3,4-trihydroxy-15-methylhexadecanoyl)glycine (Figure 2; Table S2).



**Figure 1.** MS/MS spectra of lipoamino acids **1–4** (A–D) and **7** (E) produced by *Chitinophaga eiseniae* DSM 22224 (**1**, **7**) and *Olivibacter* sp. FHG000416 (**2–4**).

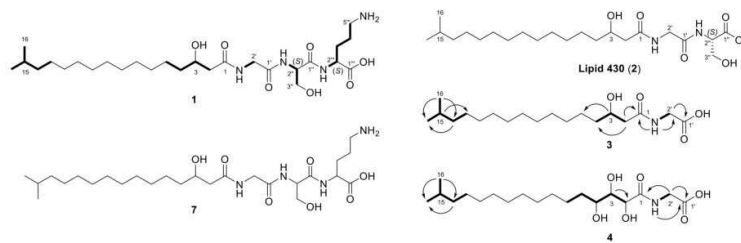
In contrast with its de-esterified hydrolysis product lipid 430 (**2**) [36], lipid 654 ( $m/z$  655.525  $[M + H]^+$ ) carrying an additional ester-linked *iso*-pentadecenoic acid (*iso*- $C_{15:0}$ ) was not observed in our previously generated *Chitinophaga* extracts (Table S1) [46]. Nevertheless, we searched for corresponding lipid 654 analogs because LAA **1** represents a close derivative of lipid 430 (**2**) containing an additional ornithine amino acid attached to the glycine-serine dipeptide. In 8 of 25 *Chitinophaga* metabolomes, the corresponding parent ion at  $m/z$  769.6051  $[M + H]^+$  with molecular formula  $C_{42}H_{80}N_4O_8$  (**5**) was found (Table S1). Neutral losses of ornithine-serine-glycine and the resulting fragment ion at  $m/z$  251.2372  $[M + H]^+$  were identical to the MS spectrum of compound **1** (Figure 3A). Moreover, the MS/MS fragmentation pattern containing the characteristic fragment ion at  $m/z$  527.3814  $[M + H]^+$  (loss of 242.2251 Da ( $C_{15}H_{30}O_2$ )) proved the

presence of the expected ester-linked fatty acyl moiety (*iso*-C<sub>15:0</sub>) at C-3. Therefore, compound **5** (2-(3-aminopropyl)-5-(hydroxymethyl)-26-methyl-12-(12-methyltridecyl)-4,7,10,14-tetraoxo-13-oxa-3,6,9-triazaheptacosanoic acid) was clearly identified as the ornithine analog of the literature-known lipid 654 [34] (Figure 4).

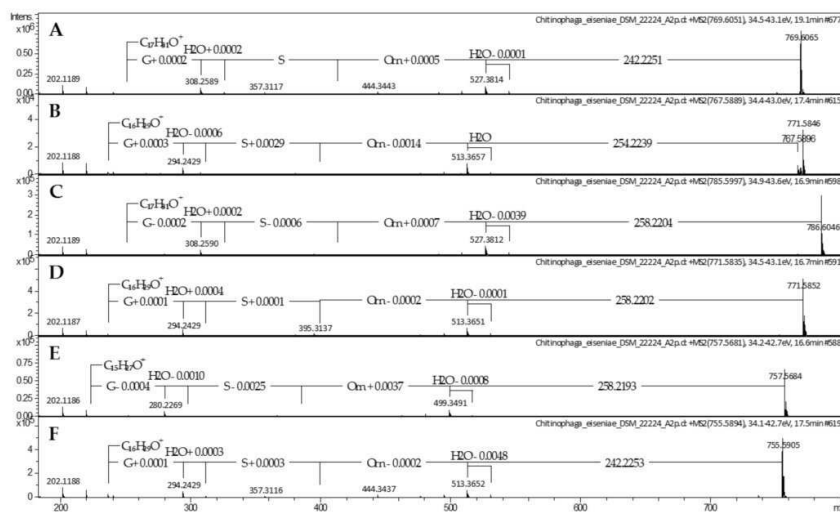
**Table 1.** <sup>1</sup>H and <sup>13</sup>C data of compounds **1–3** (**1**: 500 MHz/126 MHz, DMSO-*d*<sub>6</sub>; **2**: 400 MHz/101 MHz, MeOD-*d*<sub>4</sub>; **3**: 600 MHz/101 MHz, DMSO-*d*<sub>6</sub>).

Position	1		2		3	
	$\delta_{\text{H}}$ , (J in Hz)	$\delta_{\text{C}}$ , Type <sup>a</sup>	$\delta_{\text{H}}$ , (J in Hz)	$\delta_{\text{C}}$ , Type	$\delta_{\text{H}}$ , (J in Hz)	$\delta_{\text{C}}$ , Type
1		171.1, C		175.0, C		171.1, C
2	2.20, d (6.1)	43.3, CH <sub>2</sub>	2.41, dd (14.0, 4.3), 2.34, dd (14.0, 8.6)	44.8, CH <sub>2</sub>	2.19, d (6.4)	43.6, CH <sub>2</sub>
3	3.82–3.75, m	67.2, CH	4.02–3.94, m	70.0, CH	3.79–3.74, m	67.4, CH
4	1.39–1.31, m	36.7, CH <sub>2</sub>	1.53–1.46, m	38.4, CH <sub>2</sub>	1.40–1.34, m, 1.33–1.27, m	36.8, CH <sub>2</sub>
5	1.39–1.31, m <sup>b</sup> , 1.31–1.26, m <sup>b</sup>	24.8, CH <sub>2</sub>	1.50–1.43, m, 1.38–1.32, m	26.7, CH <sub>2</sub>	1.40–1.35, m, 1.31–1.25, m	25.1, CH <sub>2</sub>
6	1.26–1.22, m <sup>b</sup>	28.0–28.9 <sup>c</sup> , CH <sub>2</sub>	1.36–1.25, m <sup>e</sup>	30.7–31.0 <sup>f</sup> , CH <sub>2</sub>	1.26–1.21, m <sup>g</sup>	29.0–29.3 <sup>h</sup> , CH <sub>2</sub>
7	1.26–1.22, m <sup>b</sup>	28.0–28.9 <sup>c</sup> , CH <sub>2</sub>	1.36–1.25, m <sup>e</sup>	30.7–31.0 <sup>f</sup> , CH <sub>2</sub>	1.26–1.21, m <sup>g</sup>	29.0–29.3 <sup>h</sup> , CH <sub>2</sub>
8	1.26–1.22, m <sup>b</sup>	28.0–28.9 <sup>c</sup> , CH <sub>2</sub>	1.36–1.25, m <sup>e</sup>	30.7–31.0 <sup>f</sup> , CH <sub>2</sub>	1.26–1.21, m <sup>g</sup>	29.0–29.3 <sup>h</sup> , CH <sub>2</sub>
9	1.26–1.22, m <sup>b</sup>	28.0–28.9 <sup>c</sup> , CH <sub>2</sub>	1.36–1.25, m <sup>e</sup>	30.7–31.0 <sup>f</sup> , CH <sub>2</sub>	1.26–1.21, m <sup>g</sup>	29.0–29.3 <sup>h</sup> , CH <sub>2</sub>
10	1.26–1.22, m <sup>b</sup>	28.0–28.9 <sup>c</sup> , CH <sub>2</sub>	1.36–1.25, m <sup>e</sup>	30.7–31.0 <sup>f</sup> , CH <sub>2</sub>	1.26–1.21, m <sup>g</sup>	29.0–29.3 <sup>h</sup> , CH <sub>2</sub>
11	1.26–1.22, m <sup>b</sup>	28.0–28.9 <sup>c</sup> , CH <sub>2</sub>	1.36–1.25, m <sup>e</sup>	30.7–31.0 <sup>f</sup> , CH <sub>2</sub>	1.26–1.21, m <sup>g</sup>	29.0–29.3 <sup>h</sup> , CH <sub>2</sub>
12	1.26–1.22, m <sup>b</sup>	28.0–28.9 <sup>c</sup> , CH <sub>2</sub>	1.36–1.25, m <sup>e</sup>	30.7–31.0 <sup>f</sup> , CH <sub>2</sub>	1.26–1.21, m <sup>g</sup>	29.0–29.3 <sup>h</sup> , CH <sub>2</sub>
13	1.26–1.22, m <sup>b</sup>	26.4, CH <sub>2</sub>	1.36–1.25, m <sup>e</sup>	28.5, CH <sub>2</sub>	1.26–1.21, m <sup>g</sup>	26.8, CH <sub>2</sub>
14	1.16–1.11, m	38.1, CH <sub>2</sub>	1.17, q (6.7)	40.2, CH <sub>2</sub>	1.15–1.11, m	38.5, CH <sub>2</sub>
15	1.50, non (6.6)	27.0, CH	1.50, sept (6.6)	29.2, CH	1.49, non (6.6)	27.4, CH
16	0.85, d (6.6)	22.2, CH <sub>3</sub>	0.88, d (6.7)	23.0, CH <sub>3</sub>	0.84, d (6.6)	22.5, CH <sub>3</sub>
1'		168.6, C		171.6, C		n.o. <sup>d</sup>
2'	3.73, d (4.6)	41.5, CH <sub>2</sub>	3.98, d (16.8), 3.90, d (16.7)	43.5, CH <sub>2</sub>	3.70, dd (17.4, 5.7), 3.65, dd (17.4, 5.7)	41.0, CH <sub>2</sub>
2-NH	8.10–8.04, m		8.30, t (5.6)		8.01, t (5.7)	
1'''		n.o. <sup>d</sup>		173.4, C		
2''	4.25, dd (12.6, 6.2)	54.7, CH	4.50, t (4.3)	56.2, CH		
2''-NH	7.91–7.83, m		8.04, d (7.7)			
3''	3.56, m, 3.46, m	62.0, CH <sub>2</sub>	3.91, dd (11.5, 4.5), 3.83, dd (11.3, 3.8)	63.0, CH <sub>2</sub>		
1'''		n.o. <sup>d</sup>				
2'''	3.82–3.75, m	53.6, CH				
2'''-NH	7.66–7.59, m					
3'''	1.80–1.71, m, 1.69–1.61, m	n.o. <sup>d</sup>				
4'''	1.66–1.56, m	24.0, CH <sub>2</sub>				
5'''	2.80–2.73, m	38.4, CH <sub>2</sub>				

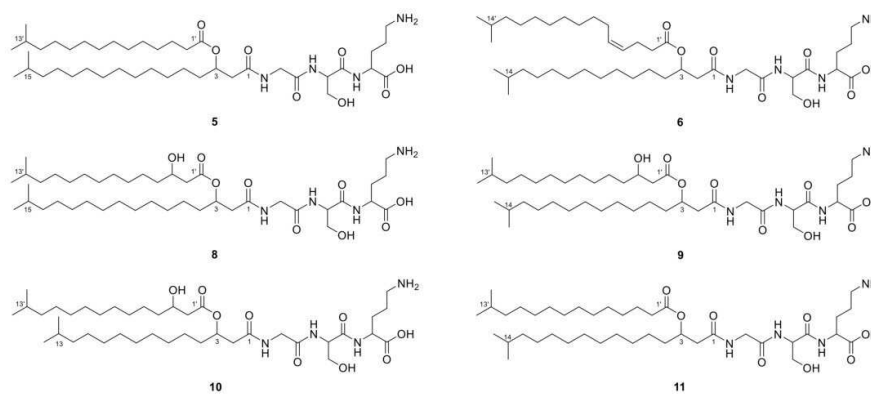
<sup>a</sup> Extracted from HSQC spectra; <sup>b,c,e-h</sup> Signals are overlapping; <sup>d</sup> Not observed due to extreme line broadening.



**Figure 2.** Chemical structures of lipoamino acids **1–4** and **7** produced by *C. eiseniae* DSM 22224 (**1**, **7**) and *Olivibacter* sp. FHG000416 (**2–4**). The COSY (in bold) and key H→C HMBC (arrows) correlations observed for lipids **1**, **3**, and **4** are indicated.



**Figure 3.** MS/MS spectra of lipoamino acids **5** (**A**), **6** (**B**), and **8–11** (**C–F**) produced by *Chitinophaga eiseniae* DSM 22224.



**Figure 4.** Postulated chemical structures based on UHR-ESI-MS/MS spectra of lipoamino acids **5**, **6**, and **8–11**, produced by *Chitinophaga eiseniae* DSM 22224.

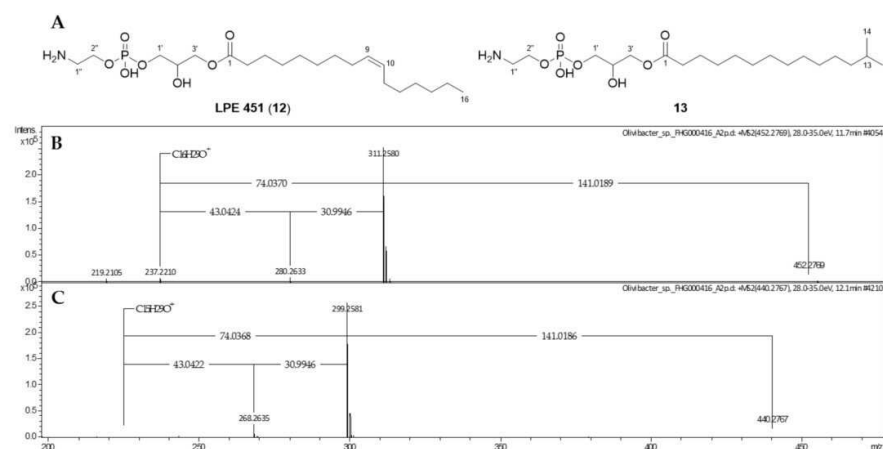
Based on this finding, the second *Chitinophaga* core metabolite with a UHR-ESI-MS ion peak at  $m/z$  767.5896  $[M + H]^+$  was assigned the molecular formula  $C_{42}H_{78}N_4O_8$  (6). Compared with compound 5, a difference of 2.0155 Da indicated a double bond in one of the fatty acid residues. The MS/MS fragmentation pattern revealed the amino acid sequence ornithine-serine-glycine. Furthermore, compared with lipid 5, a fragment ion at  $m/z$  513.3657  $[M + H]^+$  resulted from the loss of an ester-linked fatty acid residue (254.2239 Da) with additional 11.9988 Da ( $1 \times C$ ) at C-3. Together with the remaining fragment ion at  $m/z$  237.2212  $[M + H]^+$ ,  $C_{16}H_{29}O^+$ , 14.0160 Da less compared with 1–3 and 5), we postulate lipid 6 to carry two *iso*-palmitic acid (*iso*- $C_{16}$ ) residues, a saturated amide-linked *iso*- $C_{16:0}$ , and an ester-linked mono-unsaturated *iso*- $C_{16:1}$  (Figure 3B). In line with reported glycine and glyceryl-serine analogs [13], we assume the double bond to be in  $\Delta^4$  position and with *Z* configuration for lipid 6, identifying it as (*Z*)-2-(3-aminopropyl)-5-(hydroxymethyl)-27-methyl-12-(11-methyldodecyl)-4,7,10,14-tetraoxo-13-oxa-3,6,9-triazaoctacos-17-enoic acid (Figure 4). However, structural confirmation by NMR was not possible due to a low production titer in large-scale fermentation of *C. eiseniae* DSM 22224.

With lipid 6 carrying a different amide-linked fatty acid residue than 5, we assumed to find a correlating close derivative of lipids 1–3 that carried an amide-linked *iso*- $C_{16:0}$  fatty acid residue instead of an *iso*- $C_{17:0}$ . In total, 18 *Chitinophaga* strain data sets contained the matching positive parent ion at  $m/z$  531.3754 according to molecular formula  $C_{26}H_{50}N_4O_7$  (7) (Table S1). A missing ester-linked fatty acid residue and an otherwise identical UHR-ESI-MS/MS spectrum of compound 7 compared with lipid 6 confirmed the presence of an *iso*- $C_{16:0}$  fatty acid residue amide linked to a glycine-serine-ornithine amino acid motif (Figure 1E). Therefore, lipid 7 was identified as 5-amino-2-(3-hydroxy-2-(2-(3-hydroxy-14-methylpentadecanamido)acetamido)propanamido)pentanoic acid (Figure 2).

In addition to lipids 5 and 6, we observed four closely related metabolites with similar masses and RTs: 8 ( $m/z$  785.6016  $[M + H]^+$ ,  $C_{42}H_{80}N_4O_9$ , 16.9 min), 9 ( $m/z$  771.5852  $[M + H]^+$ ,  $C_{40}H_{78}N_4O_9$ , 16.7 min), 10 ( $m/z$  757.5681  $[M + H]^+$ ,  $C_{40}H_{76}N_4O_9$ , 16.6 min), and 11 ( $m/z$  755.5894  $[M + H]^+$ ,  $C_{41}H_{78}N_4O_8$ , 17.5 min) in the extract of *C. eiseniae* DSM 22224. Sharing the same tripeptide moiety of glycine-serine-ornithine, these lipid 654 analogs vary in the length and hydroxylation of their acyl chains. The analysis of the UHR-ESI-MS/MS spectra of LAAs 8–10 revealed a neutral loss of 258.2204 Da ( $C_{15}H_{28}O_3$ ). Compared with the neutral loss of 242.2251 Da ( $C_{15}H_{28}O_2$ ) of LAA 5, the additional 15.9953 Da depicted the presence of an additional oxygen atom in the ester-linked *iso*- $C_{15:0}$  fatty acid residue (Figure 3C–E). In accordance with the isolated LAAs 2 and 3, the corresponding hydroxyl group was assigned to C-3' position. With all other fragment ions identical to LAA 5, the structure of LAA 8 was postulated as 2-(3-aminopropyl)-16-hydroxy-5-(hydroxymethyl)-26-methyl-12-(12-methyltridecyl)-4,7,10,14-tetraoxo-13-oxa-3,6,9-triazaheptacosanoic acid (Figure 4). LAA 9 shared the characteristic fragment ion at  $m/z$  237.2213  $[M + H]^+$ , identified as an unsaturated amide-linked *iso*- $C_{16:0}$  fatty acid residue with LAA 6 (Figure 3D). Therefore, 9 was identified as 2-(3-aminopropyl)-16-hydroxy-5-(hydroxymethyl)-26-methyl-12-(11-methyldodecyl)-4,7,10,14-tetraoxo-13-oxa-3,6,9-triazaheptacosanoic acid (Figure 4). For LAA 10, a characteristic fragment ion was detected at  $m/z$  223.2059  $([M + H]^+, C_{15}H_{27}O^+)$  with 14.0154 Da less compared with the one of 9 (Figure 3E). This indicated lipid 10 to be a LAA with an ornithine-serine-glycine tripeptide amide linked to an *iso*- $C_{15:0}$  fatty acid residue that carries another C-3 ester-linked hydroxylated *iso*- $C_{15:0}$  fatty acid residue. The chemical structure of the new compound 10 is 2-(3-aminopropyl)-16-hydroxy-5-(hydroxymethyl)-26-methyl-12-(10-methylundecyl)-4,7,10,14-tetraoxo-13-oxa-3,6,9-triazaheptacosanoic acid (Figure 4). The analysis of the UHR-ESI-MS/MS spectrum of the last derivative of this lipid family revealed the ester-linked fatty acid to have an *iso*- $C_{15:0}$  and the amide-linked one an *iso*- $C_{16:0}$  moiety without further modifications (Figure 3F). Thus, the new LAA 11 was identified as 2-(3-aminopropyl)-5-(hydroxymethyl)-26-methyl-12-(11-methyldodecyl)-4,7,10,14-tetraoxo-13-oxa-3,6,9-triazaheptacosanoic acid (Figure 4).

## 2.2. Phospholipids

In addition to the aforementioned LAAs, dereplication of the last two *Chitinophaga* core buckets with  $m/z$  452.2769  $[M + H]^+$  and 440.2767  $[M + H]^+$  suggested the closely related molecular formulae  $C_{21}H_{42}NO_7P$  (**12**) and  $C_{20}H_{42}NO_7P$  (**13**), respectively. A database query for **12** provided LPE 451 as a structural hypothesis, based on similar MS and MS/MS spectra in accordance with the literature [51]. Neutral losses of the phosphatidylethanolamine group (141.018 Da) followed by the  $CH_2OH$  group (30.994 Da) or the glycerol moiety (74.037 Da), which are reported for phospholipids such as the identified LPE [52,53], resulted in a key fragment ion at  $m/z$  237.2210  $[M + H]^+$  ( $C_{16}H_{29}O^+$ , DBE = 3). Highly similar MS/MS fragmentation of **13** led to the fragment ion at  $m/z$  225.2212  $[M + H]^+$  ( $C_{15}H_{29}O^+$ , DBE = 2), indicating the structural variance to be located in the acyl motif. For NMR analysis, the isolation of these compounds was again carried out from extracts of *Olivibacter* sp. FHG000416, as higher production titers were observed compared with *C. eiseniae* DSM 22224. For **12**, NMR data confirmed the occurrence of lysophosphatidylethanolamine ( $C_{16:1}$ ), as postulated by dereplication. According to published data for LPE 451, the double bond ( $\delta_{H-15/16}$  5.24–5.12 ppm) of the mono-unsaturated palmitoyl motif was assigned to  $\Delta^9$  position with *Z* configuration, with the acyl chain being attached to the glycerol moiety at *sn*-1 position (Table S3) [51]. Thus, **12** was identified as 1-(9*Z*-palmitoyl)-2-hydroxy-*sn*-glycerol-3-phosphoethanol-amine (Figure 5). In comparison, 1D and 2D-NMR data acquired for **13** revealed the double bond of the acyl chain to be missing. Instead, an isopropyl moiety ( $\delta_{H-19}$  1.29 ppm,  $\delta_{C-20}$  27.7 ppm/ $\delta_{H-20}$  0.63 ppm,  $\delta_{C-20}$  22.2 ppm) was observed, identifying the acyl chain as *iso*- $C_{15:0}$  attached to the otherwise identical molecule (Table S3). Therefore, **13** was identified as 1-isopentadecanoyl-2-hydroxy-*sn*-glycerol-3-phosphoethanolamine (Figure 5).



**Figure 5.** Chemical structures (A) and MS/MS spectra of LPE 451 (**12**, B) and its new derivative (**13**, C), isolated from *Olivibacter* sp. FHG000416.

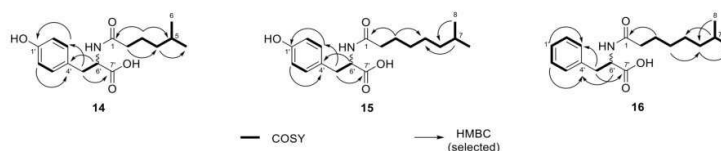
## 2.3. N-acyl Amino Acids from *Olivibacter* sp. FHG000416

During processing extracts from *Olivibacter* sp. FHG000416, three additional metabolites with lipid-like properties (e.g., surface activity) were observed. For compound **14**, the molecular formula  $C_{16}H_{23}NO_4$  ( $m/z$  294.1696  $[M + H]^+$ ) was determined by UHR-ESI-MS. The MS/MS fragmentation pattern showed a fragment at  $m/z$  182.0811, matching the protonated form of tyrosine ( $[M + H]^+$ ,  $C_9H_{12}NO_3^+$ ). This hypothesis was further supported by neutral losses of  $NH_3$  and  $H_2O$ , commonly known for the fragmentation of amino acids (Figure S3A) [15]. Furthermore, the neutral loss of 112.0885 Da ( $C_7H_{12}O$ ) implied the loss of a saturated acyl group. The 1D and 2D NMR experiments confirmed this structural proposal (Table 2), showing an isopropyl group at the end of the *iso*- $C_{7:0}$  acyl

moiety, which was attached to tyrosine via a peptide bond. Therefore, **14** was identified as *N*-(5-methyl)hexanoyl tyrosine (Figure 6).

**Table 2.**  $^1\text{H}$  and  $^{13}\text{C}$  data of compounds **14–16** ( $^1\text{H}$ : 400 MHz,  $^{13}\text{C}$ : 101 MHz, MeOD- $d_4$ ).

Position	<b>14</b>		<b>15</b>		<b>16</b>	
	$\delta_{\text{H}}$ , (J in Hz)	$\delta_{\text{C}}$ , Type	$\delta_{\text{H}}$ , (J in Hz)	$\delta_{\text{C}}$ , Type	$\delta_{\text{H}}$ , (J in Hz)	$\delta_{\text{C}}$ , Type
1'		157.3, C		157.2, C	7.23–7.17, m	127.8, CH
2'	6.69, dd (6.6, 1.9)	116.1, CH	6.68, dd (6.5, 2.1)	116.1, CH	7.30–7.22, m	129.4, CH
3'	7.04, dd (6.6, 1.9)	131.2, CH	7.03, dd (6.6, 1.5)	131.3, CH	7.26–7.21, m	130.3, CH
4'		129.2, C		129.4, C		138.7, C
5'	3.11, dd (14.0, 4.8), 2.83, dd (14.0, 9.3)	37.7, CH <sub>2</sub>	3.11, dd (14.0, 4.9), 2.84, dd (13.9, 8.9)	37.9, CH <sub>2</sub>	3.22, dd (13.9, 4.8), 2.93, dd (13.9, 9.5)	38.5, CH <sub>2</sub>
6'	4.59, dd (9.3, 4.6)	55.3, CH	4.57, dd (8.9, 4.9)	55.7, CH	4.66, dd (9.4, 4.8)	55.1, CH
7'		175.3, C		175.9, C		175.2, C
1		176.1, C		175.9, C		176.1, C
2	2.13, t (7.5)	37.1, CH <sub>2</sub>	2.15, t (7.4)	37.0, CH <sub>2</sub>	2.14, t (7.4)	36.9, CH <sub>2</sub>
3	1.55–1.47, m	24.8, CH <sub>2</sub>	1.56–1.48, m	27.0, CH <sub>2</sub>	1.55–1.46, m	26.9, CH <sub>2</sub>
4	1.14–1.05, m	39.4, CH <sub>2</sub>	1.32–1.24, m	28.2, CH <sub>2</sub>	1.32–1.23, m	28.2, CH <sub>2</sub>
5	1.52–1.46, m	29.0, CH	1.24–1.18, m	30.4, CH <sub>2</sub>	1.23–1.14, m	30.4, CH <sub>2</sub>
6	0.86, d (6.6)	22.8/22.9, CH <sub>3</sub>	1.20–1.12, m	40.0, CH <sub>2</sub>	1.22–1.10, m	40.0, CH <sub>2</sub>
7			1.57–1.49, m	29.1, CH	1.55–1.49, m	29.1, CH
8			0.88, d (6.6)	23.0, CH <sub>3</sub>	0.87, d (6.6)	23.1/23.0, CH <sub>3</sub>



**Figure 6.** Chemical structures of *N*-acyl amino acids (**14–16**) isolated from *Olivibacter* sp. FHG000416, with observed COSY (in bold) and key H→C HMBC (arrows) correlations.

Based on UHR-ESI-MS analysis, compounds **15** ( $m/z$  322.2013  $[\text{M} + \text{H}]^+$ ,  $\text{C}_{18}\text{H}_{27}\text{NO}_4$ ) and **16** ( $m/z$  306.2063  $[\text{M} + \text{H}]^+$ ,  $\text{C}_{18}\text{H}_{27}\text{NO}_3$ ) shared the same saturated acyl group indicated by the neutral loss of 140.1201 Da ( $\text{C}_9\text{H}_{16}\text{O}$ ). The difference of 28.0317 Da compared with **14** is equivalent to two additional methylene groups matching an *iso*- $\text{C}_{9,0}$  acyl group. Identical to NAAA **14**, the acyl group of **15** was attached to tyrosine (Figure S3B). In contrast, the fragment ion of NAAA **16** at  $m/z$  166.0862 ( $[\text{M} + \text{H}]^+$ ,  $\text{C}_9\text{H}_{12}\text{NO}_2^+$ ) corresponded to one oxygen atom less and was thereby assumed to be phenylalanine (Figure S3C). NMR analysis confirmed both structural proposals based on the MS data. NAAA **15** was determined as *N*-(7-methyl)octanoyl tyrosine and **16** as *N*-(7-methyl)octanoyl phenylalanine (Figure 6, Table 2). Finally, advanced Marfey's analysis of NAAAs (**14–16**) revealed that all three compounds had been isolated as an enantiomeric mixture. The D/L-ratio was determined by UV signal integration as 1:16 for **14**, 1:1.7 for **15**, and 1.6:1 for **16** (Figure S4–S6).



#### 2.4. Antimicrobial Activity of Lipids Isolated from Bacteroidetes

The antimicrobial activity of compounds **1–4** and **12–16** isolated either from *C. eiseniae* DSM 22224 or *Olivibacter* sp. FHG000416 was determined by microbroth dilution assay against a panel of 12 indicator strains up to a test concentration of 64 µg/mL. No growth inhibition effect was observed for LAA **4** and the NAAAs **15** and **16**. *N*-(5-methyl)hexanoyl tyrosine (**14**) exhibited very low effect against Gram-negative *M. catarrhalis* ATCC 25238 and Gram-positive *M. luteus* DSM 20030 at the highest concentration of 64 µg/mL. In contrast, LAAs **1–3** and both LPEs (**12** and **13**) showed growth inhibiting activity in a range of 4–16 µg/mL and 16–64 µg/mL against *M. catarrhalis* ATCC 25238 and *M. luteus* DSM 20030, respectively. Moreover, LAA **1** inhibited *B. subtilis* DSM 10 up to 8 µg/mL and *C. albicans* FH2173 as well as *E. coli* ATCC 35218 up to 64 µg/mL—the latter only when tested in bicarbonate-supplemented screening medium (MHC) (Table 3).

**Table 3.** MIC values (µg/mL) of compounds **1–4** and **12–16**. MHC = cation-adjusted Mueller–Hinton II medium supplemented with 3.7 g/L bicarbonate, n.d. = not determined.

	Compounds								
	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>12</b>	<b>13</b>	<b>14</b>	<b>15</b>	<b>16</b>
<i>E. coli</i> ATCC 35218 (MH-II)	>64	>64	>64	>64	>64	>64	>64	>64	>64
<i>E. coli</i> ATCC 35218 (MHC)	64	>64	>64	>64	>64	>64	>64	>64	>64
<i>E. coli</i> ATCC 25922 ΔToIC	>64	>64	>64	>64	>64	>64	>64	>64	>64
<i>P. aeruginosa</i> ATCC 27853	>64	>64	>64	>64	>64	>64	>64	>64	>64
<i>K. pneumoniae</i> ATCC 13883	>64	>64	>64	>64	>64	>64	>64	>64	>64
<i>M. catarrhalis</i> ATCC 25238	4–8	16–32	8	>64	4–8	16	64	>64	>64
<i>A. baumannii</i> ATCC 19606	n.d.	n.d.	>64	>64	n.d.	n.d.	>64	>64	>64
<i>B. subtilis</i> DSM 10	8	>64	>64	>64	>64	>64	>64	>64	>64
<i>S. aureus</i> ATCC 25923	>64	>64	>64	>64	>64	>64	>64	>64	>64
<i>M. luteus</i> DSM 20030	64	32–64	>64	>64	16	16	64	>64	>64
<i>L. monocytogenes</i> DSM 20600	n.d.	n.d.	>64	>64	>64	n.d.	>64	>64	>64
<i>M. smegmatis</i> ATCC 607	>64	n.d.	>64	n.d.	>64	n.d.	>64	>64	>64
<i>C. albicans</i> FH2173	64	>64	>64	>64	>64	>64	>64	>64	>64

### 3. Discussion

Bacterial small molecules are of great importance for medicinal, industrial, and agricultural applications [54]. Within this class of compounds, lipids represent a structurally diverse class of metabolites with a variety of biological functions [2,26,30]. Advances in LC-MS emerged the field of lipidomics, allowing high-throughput detection and sophisticated analysis of complex lipid samples. Identification of abundant known lipid classes and structure predictions of lipids new to science are possible. However, compared with NMR techniques, LC-MS measurements that require lower amounts of sample will not provide enough data to deduce the planar structure unambiguously [39].

In the current study, we described the characterization and structure elucidation of the unknown previously determined core lipids of the Bacteroidetes genus *Chitinophaga* [46]. Interestingly, we observed the highest production titers of three of the four lipids in extracts of *Olivibacter* sp. FHG000416, a strain belonging to another Bacteroidetes genus. Therefore, compound isolation was carried out from these two strains: *Olivibacter* sp. FHG000416 and *Chitinophaga eiseniae* DSM 22,224. Two of the four lipids are identified as LPEs, with **13** being an undescribed derivative of the literature-known LPE 451 (**12**). The remaining ones (**1** and **6**), together with several further derivatives thereof, are novel lipids classified as LAAs. The literature frequently described the production of LAAs by various bacteria, with some of them being isolated and chemically fully characterized [5]. These are often mono- or dipeptide lipids containing glycine, serine, ornithine, or glycine-serine as amino acid residues amide-linked to an *iso*-fatty acid ester at C-3 with different degrees of unsaturation [6–12]. Based on MS/MS and NMR experiments, the lipids **3** and **4** were identified as new LAA glycine derivatives. The *Chitinophaga* core lipids **1** and **6**, as well

as lipids **5** and **8–11** form a LAA family with an undescribed tripeptide moiety of glycine-serine-ornithine. This novel tripeptide LAA family is most closely related to the intensively studied glycine-serine dipeptide lipids 654 and 430 (**2**). Interestingly, with lipid 430 (**2**), only the de-esterified hydrolysis product of lipid 654 [34] was detected in the methanolic extracts of 23 of 25 *Chitinophaga* strains generated in our previous study [46]. Recent intensive studies of human pathogenic Bacteroidetes such as *P. gingivalis* revealed both dipeptide lipids to engage TLR2 [34]. They are involved in the development of two chronic inflammatory diseases, including periodontitis and atherosclerosis [30]. Furthermore, studies implicated lipid 654 to be involved in the development of multiple sclerosis [25]. Therefore, further studies are necessary to investigate the effects of the here described novel tripeptide LAAs in terms of immune response and whether they might also be part of the cell membrane of pathogens such as *P. gingivalis*.

The same applies to the new *N*-acyl tyrosine and phenylalanine analogs. They belong to a growing family of microbial secondary metabolites isolated from bacteria [22,23], fungi [21], or from environmental DNA expressed in heterologous hosts such as *E. coli* [16–20]. It is hypothesized that these lipids are possible signaling molecules with a wide range of biological activities from anti-cancer therapy targets to antimicrobial lipids [16–23]. With no antimicrobial activity observed, further studies are necessary to elucidate their biological function.

In addition to the structure-associated as well as the target organism-oriented antimicrobial activities of lipids, a membrane destabilization mechanism has been investigated for several decades [26,55]. In this context, a ‘carpet’ mechanism is suggested, which results in a detergent-like membrane permeation and/or disintegration [55,56]. Based on the reported findings, we selected a panel of 12 microorganisms to cover a wide range of possible targets for the herein described NPs. The antimicrobial profiling of the herein tested lipids revealed a frequent growth inhibitory effect towards *M. catarrhalis*, with lipid **1** additionally showing growth inhibiting activity against *B. subtilis*. The more hydrophobic cell surface of *M. catarrhalis*, compared with the surfaces of, e.g., *E. coli* and *P. aeruginosa*, is believed to be the reason for high accessibility of hydrophobic agents to the cell surface [57–59]. Therefore, we assume *M. catarrhalis* to be the most susceptible pathogen towards a suggested membranolytic mode of action mediated by LAAs carrying a hydrophobic lipid moiety and a short hydrophilic amino acid moiety. The mode of action needs to be confirmed in further studies.

#### 4. Materials and Methods

##### 4.1. Isolation of *Olivibacter* sp. FHG000416

*Olivibacter* sp. FHG000416 is incorporated into the Fraunhofer strain collection. This strain was isolated in 2016 from termite carton nest material of *Coptotermes gestroi* kindly provided by Prof. Dr. Rudy Plarre (Bundesanstalt für Materialforschung und –prüfung, Berlin). In brief, living cells were retrieved using nycodenz density gradient method, as described elsewhere [60]. Diluted cell suspensions ( $10^{-2}$ – $10^{-6}$ ) were plated on R2A Agar (DMSZ medium 830) and incubated for seven days at 28 °C. In order to isolate single strains, single colonies were propagated four times, then affiliated by using 16S rRNA gene sequencing with the primer pair E8F (5′-GAGTTTGATCCTGGCTCAG-3′) and 1492R (5′-AGAGTTTGATCCTGGCTCAG-3′) [61]. As judged on nearly full-length 16S rRNA sequence (MZ073637) comparison, FHG000416 is phylogenetically most closely related to *Olivibacter domesticus* DSM 18733 [48,49], with only ~94.5% identity.

##### 4.2. Mass Spectrometric Analysis

For all UHPLC-QTOF-UHR-MS and MS/MS measurements, a quadrupole time-of-flight spectrometer (LC-QTOF maXis II, Bruker Daltonics, Bremen, Germany) equipped with an electrospray ionization source in line with an Agilent 1290 infinity LC system (Agilent Technologies, Santa Clara, CA, USA) was used. C18 RP-UHPLC (ACQUITY UPLC BEH C18 column (130 Å, 1.7 µm, 2.1 × 100 mm)) was performed at 45 °C with the following

linear gradient (A: H<sub>2</sub>O, 0.1% HCOOH; B: CH<sub>3</sub>CN, 0.1% HCOOH; flow rate: 0.6 mL/min): 0 min: 95% A; 0.30 min: 95% A; 18.00 min: 4.75% A; 18.10 min: 0% A; 22.50 min: 0% A; 22.60 min: 95% A; 25.00 min: 95% A. A 50 to 2000 *m/z* scan range at 1 Hz scan rate was used to acquire mass spectral data. The injection volume was set to 5  $\mu$ L. MS/MS experiments were performed at 6 Hz, and the top five most intense ions in each full MS spectrum were targeted for fragmentation by higher-energy collisional dissociation at 25 eV using N<sub>2</sub> at 10<sup>-2</sup> mbar. Precursors were excluded after 2 spectra, released after 0.5 min, and reconsidered if the intensity of an excluded precursor increased by a factor of 1.5 or more. Data were analyzed using the Bruker Data Analysis 4.0 software package.

#### 4.3. Isolation of Lipids

*C. eiseniae* DSM 22224 and *Olivibacter* sp. FHG000416 were inoculated from plate (R2A) in 300 mL Erlenmeyer flasks filled with 100 mL R2A and incubated at 28 °C with agitation at 180 rpm for 3 d. A 20 L fermentation (separated in 500 mL culture volume per 2 L flasks) of *C. eiseniae* in medium 3018 (1 g/L yeast extract, 5 g/L casitone, pH 7.0) was inoculated with 2% (*v/v*) pre-culture and incubated under the same conditions for 4 d. Using the same conditions, 7 and 20 L fermentations of *Olivibacter* sp. were performed in medium 5065 (15 g/L soluble starch, 10 g/L glucose, 10 g/L soy flour, 1 g/L yeast extract, 0.1 g/L K<sub>2</sub>HPO<sub>4</sub>, 3 g/L NaCl, pH 7.4) and 5294 (10 g/L soluble starch, 10 g/L glucose, 10 g/L glycerol 99%, 2.5 g/L liquid corn steep, 5 g/L peptone, 2 g/L yeast extract, 1 g/L NaCl, 3 g/L CaCO<sub>3</sub>, pH 7.2), respectively. An additional 40 L cultivation of *Olivibacter* sp. in medium 5294 was carried out using the same conditions as before. Cultures were subsequently lyophilized using a delta 2–24 LSCplus (Martin Christ Gefriertrocknungsanlagen GmbH, Osterode am Harz, Germany).

The dried culture of *C. eiseniae* and the 7 L culture of *Olivibacter* sp. were extracted with one-time culture volume MeOH for the isolation of lipid 1, lipid 430 (2), and 3, respectively. The extracts were evaporated to dryness using rotary evaporation under reduced pressure, resuspended in 3 L of 10% MeOH/H<sub>2</sub>O, and separately loaded onto a XAD16N column (1 L bed volume). Step-wise elution with 10%, 40%, 60%, 80%, and 100% MeOH (2-times bed volume each) was performed. The 80% and 100% fractions containing the lipids were further fractionated by preparative (Synergi<sup>TM</sup> Fusion-RP 80 Å, 10  $\mu$ m, 250  $\times$  21.2 mm) and/or semi-preparative HPLC (Nucleodur<sup>®</sup> C18 Gravity-SB, 3  $\mu$ m, 250  $\times$  10 mm) using gradients of 60–95% and 50–95% CH<sub>3</sub>CN (0.1% HCOOH) in water (0.1% HCOOH), respectively. Final purification was achieved by analytical HPLC (Synergi<sup>TM</sup> Fusion-RP 80 Å, 4  $\mu$ m, 250  $\times$  4.6 mm) or UHPLC fractionation (Acquity UPLC<sup>®</sup> BEH C18, 1.7  $\mu$ m, 100  $\times$  2.1 mm) using a custom-made fraction collector (Zinsser-Analytik, Eschborn, Germany).

The 20 L fermentation of *Olivibacter* sp. was the starting point for the isolation of lipid 4. Due to the enlarged volume, LLE was performed as an additional purification step after MeOH extraction using ethyl acetate and water. In addition to that, the isolation procedure was highly identical for lipid 3. For semi-preparative HPLC, an adapted gradient of 55–95% CH<sub>3</sub>CN in water was used.

Extraction of the dried 40 L culture of *Olivibacter* sp. with MTBE/MeOH [62] was performed to isolate LPE 451 (12), 13, and NAAAs 14–16. Combined organic layers were subsequently fractionated by preparative and semi-preparative HPLC using gradients of 40–95% and 60–95% CH<sub>3</sub>CN (0.1% HCOOH) in water (0.1% HCOOH), respectively. Again, final purification was achieved by UHPLC fractionation. After each step, fractions containing compounds of interest were evaporated to dryness using a high performance evaporator (Genevac HT-12).

(2S)-5-Amino-2-((2S)-3-hydroxy-2-(2-(3-hydroxy-15 methylhexadecanamido)acetamido)propanamido)pentanoic acid (1). Colorless solid;  $[\alpha]_D^{20.6} +21.1$  (*c* 0.19, MeOH); LC-UV (CH<sub>3</sub>CN/H<sub>2</sub>O)  $\lambda_{max}$  223 nm; <sup>1</sup>H-NMR (500 MHz, DMSO-*d*<sub>6</sub>) data and 2D spectra, see Table 1, Figure S7–S10; UHRMS (ESI-TOF) *m/z* [M + H]<sup>+</sup> calcd for C<sub>27</sub>H<sub>52</sub>N<sub>4</sub>O<sub>7</sub><sup>+</sup> 545.3909, found 545.3914.

Lipid 430 (**2**). Colorless solid;  $[\alpha]_D^{25.9} +16.3$  (c 0.37, MeOH); LC-UV (CH<sub>3</sub>CN/H<sub>2</sub>O)  $\lambda_{\max}$  220 nm; <sup>1</sup>H-NMR (400 MHz, MeOD-*d*<sub>4</sub>) and <sup>13</sup>C-NMR (101 MHz, MeOD-*d*<sub>4</sub>) data, see Table 1; UHRMS (ESI-TOF) *m/z* [M + H]<sup>+</sup> calcd for C<sub>22</sub>H<sub>42</sub>N<sub>2</sub>O<sub>6</sub><sup>+</sup> 431.3116, found 431.3115.

(3-Hydroxy-15-methylhexadecanoyl)glycine (**3**). Colorless solid;  $[\alpha]_D^{21.7} +66.7$  (c 0.02, MeOH); LC-UV (CH<sub>3</sub>CN/H<sub>2</sub>O)  $\lambda_{\max}$  224 nm; <sup>1</sup>H (600 MHz, DMSO-*d*<sub>6</sub>) and <sup>13</sup>C-NMR (151 MHz, DMSO-*d*<sub>6</sub>) data, see Table 1, Figure S11–S15; UHRMS (ESI-TOF) *m/z* [M + H]<sup>+</sup> calcd for C<sub>19</sub>H<sub>37</sub>NO<sub>4</sub><sup>+</sup> 344.2795, found 344.2802.

(2,3,4-Trihydroxy-15-methylhexadecanoyl)glycine (**4**). Colorless solid;  $[\alpha]_D^{21.7} -76.9$  (c 0.03, MeOH); LC-UV (CH<sub>3</sub>CN/H<sub>2</sub>O)  $\lambda_{\max}$  224 nm; <sup>1</sup>H (600 MHz, DMSO-*d*<sub>6</sub>) and <sup>13</sup>C-NMR (151 MHz, DMSO-*d*<sub>6</sub>) data, see Table S2, Figure S16–S20; UHRMS (ESI-TOF) *m/z* [M + H]<sup>+</sup> calcd for C<sub>19</sub>H<sub>37</sub>NO<sub>6</sub><sup>+</sup> 376.2694, found 376.2696.

1-(9Z-Palmitoyl)-2-hydroxy-*sn*-glycerol-3-phosphoethanol-amine (**12**). Colorless solid; LC-UV (CH<sub>3</sub>CN/H<sub>2</sub>O)  $\lambda_{\max}$  223 nm; <sup>1</sup>H (600 MHz, CDCl<sub>3</sub>/MeOD-*d*<sub>4</sub> 2:1) and <sup>13</sup>C-NMR (151 MHz, CDCl<sub>3</sub>/MeOD-*d*<sub>4</sub> 2:1) data, see Table S3, Figure S21–S23; UHRMS (ESI-TOF) *m/z* [M + H]<sup>+</sup> calcd for C<sub>21</sub>H<sub>42</sub>NO<sub>7</sub>P<sup>+</sup> 452.2772, found 452.2769.

1-Isopentadecanoyl-2-hydroxy-*sn*-glycerol-3-phosphoethanolamine (**13**). Colorless solid; LC-UV (CH<sub>3</sub>CN/H<sub>2</sub>O)  $\lambda_{\max}$  224 nm; <sup>1</sup>H (600 MHz, CDCl<sub>3</sub>/MeOD-*d*<sub>4</sub> 2:1) and <sup>13</sup>C-NMR (151 MHz, CDCl<sub>3</sub>/MeOD-*d*<sub>4</sub> 2:1) data, see Table S3, Figure S24–S29; UHRMS (ESI-TOF) *m/z* [M + H]<sup>+</sup> calcd for C<sub>20</sub>H<sub>42</sub>NO<sub>7</sub>P<sup>+</sup> 440.2772, found 440.2767.

*N*-(5-Methyl)hexanoyl tyrosine (**14**). Yellowish solid; LC-UV (CH<sub>3</sub>CN/H<sub>2</sub>O)  $\lambda_{\max}$  224, 277 nm; <sup>1</sup>H and <sup>13</sup>C-NMR data, see Table 2, Figure S30–S34; UHRMS (ESI-TOF) *m/z* [M + H]<sup>+</sup> calcd for C<sub>16</sub>H<sub>23</sub>NO<sub>4</sub><sup>+</sup> 294.1700, found 294.1698.

*N*-(7-Methyl)octanoyl tyrosine (**15**). Yellowish solid; LC-UV (CH<sub>3</sub>CN/H<sub>2</sub>O)  $\lambda_{\max}$  225, 277 nm; <sup>1</sup>H and <sup>13</sup>C-NMR data, see Table 2, Figure S35–S39; UHRMS (ESI-TOF) *m/z* [M + H]<sup>+</sup> calcd for C<sub>18</sub>H<sub>27</sub>NO<sub>4</sub><sup>+</sup> 322.2013, found 322.2009.

*N*-(7-Methyl)octanoyl phenylalanine (**16**). Off-white solid; LC-UV (CH<sub>3</sub>CN/H<sub>2</sub>O)  $\lambda_{\max}$  217 nm; <sup>1</sup>H and <sup>13</sup>C-NMR data, see Table 2, Figure S40–S44; UHRMS (ESI-TOF) *m/z* [M + H]<sup>+</sup> calcd for C<sub>18</sub>H<sub>27</sub>NO<sub>3</sub><sup>+</sup> 306.2064, found 306.2062.

#### NMR Studies

NMR spectra of LAA **1** were recorded on a Bruker AVANCE III 500 spectrometer (<sup>1</sup>H: 500 MHz, <sup>13</sup>C: 125 MHz) equipped with a 10 mm MNP cryo probe. For all remaining compounds, NMR spectra were acquired on a Bruker AVANCE II/III HD 400 spectrometer (<sup>1</sup>H: 400 MHz, <sup>13</sup>C: 101 MHz) or an AVANCE III HD 600 spectrometer (<sup>1</sup>H: 600 MHz, <sup>13</sup>C: 151 MHz). Chemical shifts ( $\delta$ ) given in parts per million (ppm) are referenced to the residual solvent signals of DMSO-*d*<sub>6</sub> ( $\delta_{\text{H}}$  2.50 and  $\delta_{\text{C}}$  39.5), CD<sub>3</sub>OD ( $\delta_{\text{H}}$  3.31 and  $\delta_{\text{C}}$  49.0), and CDCl<sub>3</sub> ( $\delta_{\text{H}}$  7.26 and  $\delta_{\text{C}}$  77.2, in a 2:1 mixture with CD<sub>3</sub>OD).

#### 4.4. Optical Rotation

Specific rotation was determined on a digital polarimeter (P3,000, A. Krüss Optronic GmbH). Standard wavelength was the sodium D-line with 589 nm. Temperature, concentration (g/100 mL), and solvents are reported with the determined value.

#### 4.5. Advanced Marfey's Analysis

The absolute configuration of all amino acids was determined by derivatization using Marfey's reagent [50]. Stock solutions of amino acid standards (50 mM in H<sub>2</sub>O), NaHCO<sub>3</sub> (1 M in H<sub>2</sub>O), and L-FDVA (70 mM in acetone) were prepared. Commercially available standards were derivatized using molar ratios of amino acid to FDVA and NaHCO<sub>3</sub> (1/1.4/8). After stirring at 40 °C for 3 h, 1 M HCl was added to obtain a final concentration of 170 mM to end the reaction. Samples were subsequently evaporated to dryness and dissolved in DMSO (final concentration 50 mM). L- and D-amino acids were analyzed

separately using C18 RP-UHPLC-MS with the standard gradient (for details see Section 4.2) at a flow rate of 0.6 mL/min.

Total hydrolysis of compounds **1**, lipid 430 (**2**), and **14–16** was carried out by dissolving 250 µg of each compound in 6 M deuterio-hydrochloric acid (DCl in D<sub>2</sub>O) and stirring for 7 h at 160 °C. The sample was subsequently evaporated to dryness. Samples were dissolved in 100 µL H<sub>2</sub>O, derivatized, and analyzed using the same parameters as described before.

#### 4.6. Minimal Inhibitory Concentration (MIC)

Microbroth dilution assays were performed in 96-well plates to determine the minimum inhibitory concentrations (MIC) of purified compounds dissolved in DMSO and were tested in triplicate following EUCAST instructions with minor adaptations [63,64]. A cell concentration of  $5 \times 10^5$  cells/mL was adjusted for all bacteria from an overnight culture (37 °C, 180 rpm) in cation-adjusted Mueller–Hinton II medium (BD). All tested organisms are summarized in Table 3. Dilution series of rifampicin, tetracycline, and gentamicin were used as control antibiotics (64–0.03 µg/mL) to ensure that concentrations achieved a range of effects from none to complete growth inhibition of the test strain. Negative controls were cell suspensions without test sample or antibiotic control. The turbidity was measured with a microplate spectrophotometer at 600 nm (LUMIstar Omega BMG Labtech) to assess cell growth after overnight incubation (18 h, 37 °C, 180 rpm, 85% rH).

*Mycobacterium smegmatis* ATCC 607 was grown in brain–heart infusion broth supplemented with Tween 80 (1.0% v/v) at 37 °C and 180 rpm for 48 h before the cell density was adjusted in cation-adjusted Mueller–Hinton II medium. The gentamicin control was replaced with isoniazid. Assay read out was done by cell viability assessment after 48 h (37 °C, 180 rpm, 85% rH) by ATP quantification (BacTiter-Glo, Promega), according to the manufacturer's instructions.

*Candida albicans* FH2173 was incubated at 28 °C for 48 h before diluting to  $1 \times 10^6$  cells/mL in cation-adjusted Mueller–Hinton II medium. Assays were incubated at 37 °C for 48 h with nystatin as positive control and were evaluated by ATP quantification (BacTiter-Glo, Promega).

Pre- and main cultures of *Micrococcus luteus* DSM 20030 and *Listeria monocytogenes* DSM 20600 were incubated for two days, and the assay readout was done by ATP quantification, as described before.

**Supplementary Materials:** The following are available online, Figure S1: L-FDVA adducts of serine, Figure S2: L-FDVA adducts of ornithine, Figure S3: MS/MS spectra of NAAAs **14–16** (A–C) including postulated MS/MS fragmentation pathway, Figure S4: Double L-FDVA adducts of tyrosine, Figure S5: Integrated UV signals corresponding to double L-FDVA adducts of tyrosine for hydrolyzed NAAA **14** (top) and hydrolyzed NAAA **15** (bottom), Figure S6: L-FDVA adducts of phenylalanine, Figure S7: <sup>1</sup>H-NMR (500 MHz, DMSO-*d*<sub>6</sub>) of lipid **1**, Figure S8: <sup>1</sup>H-<sup>1</sup>H COSY (500 MHz, DMSO-*d*<sub>6</sub>) spectrum of lipid **1**, Figure S9: <sup>1</sup>H-<sup>13</sup>C HSQC (500 MHz, DMSO-*d*<sub>6</sub>) spectrum of lipid **1**, Figure S10: <sup>1</sup>H-<sup>13</sup>C HMBC (500 MHz, DMSO-*d*<sub>6</sub>) spectrum of lipid **1**, Figure S11: <sup>1</sup>H-NMR (600 MHz, DMSO-*d*<sub>6</sub>) of lipid **3**, Figure S12: <sup>1</sup>H-<sup>1</sup>H COSY (600 MHz, DMSO-*d*<sub>6</sub>) spectrum of lipid **3**, Figure S13: <sup>1</sup>H-<sup>13</sup>C HSQC (600 MHz, DMSO-*d*<sub>6</sub>) spectrum of lipid **3**, Figure S14: <sup>1</sup>H-<sup>13</sup>C HMBC (600 MHz, DMSO-*d*<sub>6</sub>) spectrum of lipid **3**, Figure S15: <sup>13</sup>C-NMR (151 MHz, DMSO-*d*<sub>6</sub>) of lipid **3**, Figure S16: <sup>1</sup>H-NMR (600 MHz, DMSO-*d*<sub>6</sub>) of lipid **4**, Figure S17: <sup>1</sup>H-<sup>1</sup>H COSY (600 MHz, DMSO-*d*<sub>6</sub>) spectrum of lipid **4**, Figure S18: <sup>1</sup>H-<sup>13</sup>C HSQC (600 MHz, DMSO-*d*<sub>6</sub>) spectrum of lipid **4**, Figure S19: <sup>1</sup>H-<sup>13</sup>C HMBC (600 MHz, DMSO-*d*<sub>6</sub>) spectrum of lipid **4**, Figure S20: <sup>13</sup>C-NMR (151 MHz, DMSO-*d*<sub>6</sub>) of lipid **4**, Figure S21: <sup>1</sup>H-NMR (600 MHz, CDCl<sub>3</sub>/MeOD-*d*<sub>4</sub> 2:1) of LPE 451 (**12**), Figure S22: <sup>13</sup>C-NMR (151 MHz, CDCl<sub>3</sub>/MeOD-*d*<sub>4</sub> 2:1) of LPE 451 (**12**), Figure S23: <sup>31</sup>P-NMR (243 MHz, CDCl<sub>3</sub>/MeOD-*d*<sub>4</sub> 2:1) of LPE 451 (**12**), Figure S24: <sup>1</sup>H-NMR (600 MHz, CDCl<sub>3</sub>/MeOD-*d*<sub>4</sub> 2:1) of LPE **13**, Figure S25: <sup>1</sup>H-<sup>1</sup>H COSY (600 MHz, CDCl<sub>3</sub>/MeOD-*d*<sub>4</sub> 2:1) of LPE **13**, Figure S26: <sup>1</sup>H-<sup>13</sup>C HSQC (600 MHz, CDCl<sub>3</sub>/MeOD-*d*<sub>4</sub> 2:1) of LPE **13**, Figure S27: <sup>1</sup>H-<sup>13</sup>C HMBC (600 MHz, CDCl<sub>3</sub>/MeOD-*d*<sub>4</sub> 2:1) of LPE **13**, Figure S28: <sup>13</sup>C-NMR (151 MHz, CDCl<sub>3</sub>/MeOD-*d*<sub>4</sub> 2:1) of LPE **13**, Figure S29: <sup>31</sup>P-NMR (243 MHz, CDCl<sub>3</sub>/MeOD-*d*<sub>4</sub> 2:1) of LPE **13**, Figure S30: <sup>1</sup>H-NMR (400 MHz, MeOD-*d*<sub>4</sub>) of NAAA **14**, Figure S31: <sup>1</sup>H-<sup>1</sup>H COSY (400 MHz, MeOD-*d*<sub>4</sub>) of NAAA **14**, Figure S32: <sup>1</sup>H-<sup>13</sup>C HSQC (400 MHz, MeOD-*d*<sub>4</sub>) of NAAA **14**, Figure S33: <sup>1</sup>H-<sup>13</sup>C HMBC (400 MHz, MeOD-*d*<sub>4</sub>) of NAAA **14**, Figure S34:

<sup>13</sup>C-NMR (101 MHz, MeOD-*d*<sub>4</sub>) of NAAA 14, Figure S35: <sup>1</sup>H-NMR (400 MHz, MeOD-*d*<sub>4</sub>) of NAAA 15, Figure S36: <sup>1</sup>H-<sup>1</sup>H COSY (400 MHz, MeOD-*d*<sub>4</sub>) of NAAA 15, Figure S37: <sup>1</sup>H-<sup>13</sup>C HSQC (400 MHz, MeOD-*d*<sub>4</sub>) of NAAA 15, Figure S38: <sup>1</sup>H-<sup>13</sup>C HMBC (400 MHz, MeOD-*d*<sub>4</sub>) of NAAA 15, Figure S39: <sup>13</sup>C-NMR (101 MHz, MeOD-*d*<sub>4</sub>) of NAAA 15, Figure S40: <sup>1</sup>H-NMR (400 MHz, MeOD-*d*<sub>4</sub>) of NAAA 16, Figure S41: <sup>1</sup>H-<sup>1</sup>H COSY (400 MHz, MeOD-*d*<sub>4</sub>) of NAAA 16, Figure S42: <sup>1</sup>H-<sup>13</sup>C HSQC (400 MHz, MeOD-*d*<sub>4</sub>) of NAAA 16, Figure S43: <sup>1</sup>H-<sup>13</sup>C HMBC (400 MHz, MeOD-*d*<sub>4</sub>) of NAAA 16, Figure S44: <sup>13</sup>C-NMR (101 MHz, MeOD-*d*<sub>4</sub>) of NAAA 16, Table S1: Overview of all lipids with their molecular formula, predicted and found masses within the metabolomics data of our previous study, and in how many of the investigated 25 *Chitinophaga* metabolomes each lipid is present, Table S2: <sup>1</sup>H and <sup>13</sup>C data of compound 4 (<sup>1</sup>H: 600 MHz, <sup>13</sup>C: 151 MHz, DMSO-*d*<sub>6</sub>), Table S3: <sup>1</sup>H and <sup>13</sup>C data of compounds 12 and 13 (<sup>1</sup>H: 600 MHz, <sup>13</sup>C: 151 MHz, CDCl<sub>3</sub>/MeOD-*d*<sub>4</sub> 2:1).

**Author Contributions:** Conceptualization, M.-K.B., S.B., P.E.H., S.M.M.S. and T.F.S.; investigation, M.-K.B., S.B., M.O., M.A.P., B.L., M.M. and M.-P.M.; data curation, M.-K.B. and S.B.; writing—original draft preparation, M.-K.B. and S.B.; writing—review and editing, all authors; visualization, M.-K.B. and S.B.; supervision, S.M.M.S. and T.F.S.; project administration, P.E.H.; funding acquisition, P.E.H. and A.V. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was financially supported by the Hessen State Ministry of Higher Education, Research and the Arts (HMWK) via the state initiative for the development of scientific and economic excellence for the LOEWE Center for Insect Biotechnology and Bioresources. Sanofi-Aventis Deutschland GmbH and Evotec International GmbH contributed to the framework of the Sanofi-Fraunhofer Natural Products Center of Excellence/Fraunhofer-Evotec Natural Products Center of Excellence.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The 16S rRNA gene sequence of *Olivibacter* sp. FHG000416 is available at reference number MZ073637 (GenBank).

**Acknowledgments:** The authors thank Luigi Toti, Marius Spohn, Sanja Mihajlovic, Jennifer Kuhn, and Christine Wehr for assistance in fermentation, Mona Abdullahi for assisting the isolation, Christoph Hartwig for technical support, as well as Kirsten-Susann Bommersheim for determination of MICs. We thank Jens Glaeser and Heike Hausmann for valuable discussions and the NMR department of the Justus-Liebig-University Giessen for technical assistance.

**Conflicts of Interest:** S.M.M.S. and P.E.H. are or have been employed by Evotec International GmbH and Sanofi-Aventis Deutschland GmbH, respectively.

**Sample Availability:** Samples of the compounds are not available from the authors.

## References

1. Harayama, T.; Riezman, H. Understanding the diversity of membrane lipid composition. *Nat. Rev. Mol. Cell Biol.* **2018**, *19*, 281–296. [CrossRef]
2. Sohlenkamp, C.; Geiger, O. Bacterial membrane lipids: Diversity in structures and pathways. *FEMS Microbiol. Rev.* **2016**, *40*, 133–159. [CrossRef]
3. Snijder, H.J.; Dijkstra, B.W. Bacterial phospholipase A: Structure and function of an integral membrane phospholipase. *Biochim. Biophys. Acta (BBA) Mol. Cell Biol. Lipids* **2000**, *1488*, 91–101. [CrossRef]
4. Asselineau, J. Bacterial Lipids Containing Amino Acids or Peptides Linked by Amide Bonds. In *Fortschritte der Chemie Organischer Naturstoffe/Progress in the Chemistry of Organic Natural Products*; Asselineau, J., Kagan, J., Eds.; Springer: Vienna, Austria, 1991; pp. 1–85. ISBN 978-3-7091-9084-5.
5. Batrakov, S.G.; Nikitin, D.I.; Mosezhnyi, A.E.; Ruzhitsky, A.O. A glycine-containing phosphorus-free lipoaminoacid from the gram-negative marine bacterium *Cyclobacterium marinum* WH. *Chem. Phys. Lipids* **1999**, *99*, 139–143. [CrossRef]
6. Yoshida, K.; Iwami, M.; Umehara, Y.; Nishikawa, M.; Uchida, I.; Kohsaka, M.; Aoki, H.; Imanaka, H. Studies on WB-3559 A, B, C and D, new potent fibrinolytic agents. I. Discovery, identification, isolation and characterization. *J. Antibiot.* **1985**, *38*, 1469–1475. [CrossRef]
7. Uchida, I.; Yoshida, K.; Kawai, Y.; Takase, S.; Itoh, Y.; Tanaka, H.; Kohsaka, M.; Imanaka, H. Studies on WB-3559 A, B, C and D, new potent fibrinolytic agents. II. Structure elucidation and synthesis. *J. Antibiot.* **1985**, *38*, 1476–1486. [CrossRef]
8. Tahara, Y.; Kameda, M.; Yamada, Y.; Kondo, K. A new lipid the ornithine and taurine-containing “Cerilipin”. *Agric. Biol. Chem.* **1976**, *40*, 243–244. [CrossRef]

9. Kawai, Y.; Yano, I.; Kaneda, K. Various kinds of lipoamino acids including a novel serine-containing lipid in an opportunistic pathogen *Flavobacterium*. Their structures and biological activities on erythrocytes. *Eur. J. Biochem.* **1988**, *171*, 73–80. [[CrossRef](#)] [[PubMed](#)]
10. Kawai, Y.; Akagawa, K. Macrophage activation by an ornithine-containing lipid or a serine-containing lipid. *Infect. Immun.* **1989**, *57*, 2086–2091. [[CrossRef](#)] [[PubMed](#)]
11. Kawazoe, R.; Okuyama, H.; Reichardt, W.; Sasaki, S. Phospholipids and a novel glycine-containing lipoamino acid in *Cytophaga johnsonae* Stanier strain C21. *J. Bacteriol.* **1991**, *173*, 5470–5475. [[CrossRef](#)] [[PubMed](#)]
12. Morishita, T.; Sato, A.; Hisamoto, M.; Oda, T.; Matsuda, K.; Ishii, A.; Kodama, K. N-type calcium channel blockers from a marine bacterium, *Cytophaga* sp. SANK 71996. *J. Antibiot.* **1997**, *50*, 457–468. [[CrossRef](#)]
13. Chianese, G.; Esposito, F.P.; Parrot, D.; Ingham, C.; de Pascale, D.; Tasdemir, D. Linear Aminolipids with Moderate Antimicrobial Activity from the Antarctic Gram-Negative Bacterium *Aequorivita* sp. *Mar. Drugs* **2018**, *16*, 187. [[CrossRef](#)]
14. Gomi, K.; Kawasaki, K.; Kawai, Y.; Shiozaki, M.; Nishijima, M. Toll-like receptor 4-MD-2 complex mediates the signal transduction induced by flavolipin, an amino acid-containing lipid unique to *Flavobacterium meningosepticum*. *J. Immunol.* **2002**, *168*, 2939–2943. [[CrossRef](#)] [[PubMed](#)]
15. Tan, B.; O'Dell, D.K.; Yu, Y.W.; Monn, M.F.; Hughes, H.V.; Burstein, S.; Walker, J.M. Identification of endogenous acyl amino acids based on a targeted lipidomics approach. *J. Lipid Res.* **2010**, *51*, 112–119. [[CrossRef](#)] [[PubMed](#)]
16. Brady, S.F.; Clardy, J. Long-Chain N-Acyl Amino Acid Antibiotics Isolated from Heterologously Expressed Environmental DNA. *J. Am. Chem. Soc.* **2000**, *122*, 12903–12904. [[CrossRef](#)]
17. Brady, S.F.; Chao, C.J.; Clardy, J. New natural product families from an environmental DNA (eDNA) gene cluster. *J. Am. Chem. Soc.* **2002**, *124*, 9968–9969. [[CrossRef](#)] [[PubMed](#)]
18. Brady, S.F.; Clardy, J. Palmitoylputrescine, an antibiotic isolated from the heterologous expression of DNA extracted from bromeliad tank water. *J. Nat. Prod.* **2004**, *67*, 1283–1286. [[CrossRef](#)]
19. Brady, S.F.; Clardy, J. N-acyl derivatives of arginine and tryptophan isolated from environmental DNA expressed in *Escherichia coli*. *Org. Lett.* **2005**, *7*, 3613–3616. [[CrossRef](#)]
20. Clardy, J.; Brady, S.F. Cyclic AMP directly activates NasP, an N-acyl amino acid antibiotic biosynthetic enzyme cloned from an uncultured beta-proteobacterium. *J. Bacteriol.* **2007**, *189*, 6487–6489. [[CrossRef](#)] [[PubMed](#)]
21. Phoon, C.W.; Somanadhan, B.; Heng, S.C.H.; Ngo, A.; Ng, S.B.; Butler, M.S.; Buss, A.D.; Sim, M.M. Isolation and total synthesis of gymnastatin N, a POLO-like kinase 1 active constituent from the fungus *Arachniotus punctatus*. *Tetrahedron* **2004**, *60*, 11619–11628. [[CrossRef](#)]
22. Touré, S.; Desrat, S.; Pellissier, L.; Allard, P.M.; Wolfender, J.L.; Dusfour, I.; Stien, D.; Eparvier, V. Characterization, Diversity, and Structure-Activity Relationship Study of Lipoamino Acids from *Pantoea* sp. and Synthetic Analogues. *Int. J. Mol. Sci.* **2019**, *20*, 1083. [[CrossRef](#)]
23. Vitale, G.A.; Sciarretta, M.; Cassiano, C.; Buonocore, C.; Festa, C.; Mazzella, V.; Núñez Pons, L.; D'Auria, M.V.; de Pascale, D. Molecular Network and Culture Media Variation Reveal a Complex Metabolic Profile in *Pantoea* cf. *eucrina* D2 Associated with an Acidified Marine Sponge. *Int. J. Mol. Sci.* **2020**, *21*, 6307. [[CrossRef](#)] [[PubMed](#)]
24. Teng, O.; Ang, C.K.E.; Guan, X.L. Macrophage-Bacteria Interactions-A Lipid-Centric Relationship. *Front. Immunol.* **2017**, *8*, 1836. [[CrossRef](#)] [[PubMed](#)]
25. Farrokhi, V.; Nemati, R.; Nichols, F.C.; Yao, X.; Anstadt, E.; Fujiwara, M.; Grady, J.; Wakefield, D.; Castro, W.; Donaldson, J.; et al. Bacterial lipodipeptide, Lipid 654, is a microbiome-associated biomarker for multiple sclerosis. *Clin. Transl. Immunol.* **2013**, *2*, e8. [[CrossRef](#)]
26. Yoon, B.K.; Jackman, J.A.; Valle-González, E.R.; Cho, N.J. Antibacterial Free Fatty Acids and Monoglycerides: Biological Activities, Experimental Testing, and Therapeutic Applications. *Int. J. Mol. Sci.* **2018**, *19*, 1114. [[CrossRef](#)]
27. Alves, E.; Dias, M.; Lopes, D.; Almeida, A.; Domingues, M.D.R.; Rey, F. Antimicrobial Lipids from Plants and Marine Organisms: An Overview of the Current State-of-the-Art and Future Prospects. *Antibiotics* **2020**, *9*, 441. [[CrossRef](#)]
28. Shiozaki, M.; Deguchi, N.; Mochizuki, T.; Wakabayashi, T.; Ishikawa, T.; Haruyama, H.; Kawai, Y.; Nishijima, M. Revised structure and synthesis of flavolipin. *Tetrahedron* **1998**, *54*, 11861–11876. [[CrossRef](#)]
29. Nemoto, T.; Ojika, M.; Takahata, Y.; Andoh, T.; Sakagami, Y. Structures of topostins, DNA topoisomerase I inhibitors of bacterial origin. *Tetrahedron* **1998**, *54*, 2683–2690. [[CrossRef](#)]
30. Olsen, I.; Nichols, F.C. Are Sphingolipids and Serine Dipeptide Lipids Underestimated Virulence Factors of *Porphyromonas gingivalis*? *Infect. Immun.* **2018**, *86*, e00035-18. [[CrossRef](#)] [[PubMed](#)]
31. Mirucki, C.S.; Abedi, M.; Jiang, J.; Zhu, Q.; Wang, Y.-H.; Safavi, K.E.; Clark, R.B.; Nichols, F.C. Biologic Activity of *Porphyromonas endodontalis* complex lipids. *J. Endod.* **2014**, *40*, 1342–1348. [[CrossRef](#)] [[PubMed](#)]
32. Wang, Y.-H.; Nemati, R.; Anstadt, E.; Liu, Y.; Son, Y.; Zhu, Q.; Yao, X.; Clark, R.B.; Rowe, D.W.; Nichols, F.C. Serine dipeptide lipids of *Porphyromonas gingivalis* inhibit osteoblast differentiation: Relationship to Toll-like receptor 2. *Bone* **2015**, *81*, 654–661. [[CrossRef](#)]
33. Nichols, F.C.; Housley, W.J.; O'Connor, C.A.; Manning, T.; Wu, S.; Clark, R.B. Unique lipids from a common human bacterium represent a new class of Toll-like receptor 2 ligands capable of enhancing autoimmunity. *Am. J. Pathol.* **2009**, *175*, 2430–2438. [[CrossRef](#)] [[PubMed](#)]

34. Clark, R.B.; Cervantes, J.L.; Maciejewski, M.W.; Farrokhi, V.; Nemati, R.; Yao, X.; Anstadt, E.; Fujiwara, M.; Wright, K.T.; Riddle, C.; et al. Serine lipids of *Porphyromonas gingivalis* are human and mouse Toll-like receptor 2 ligands. *Infect. Immun.* **2013**, *81*, 3479–3489. [[CrossRef](#)]
35. Schneider, Y.K.-H.; Hansen, K.Ø.; Isaksson, J.; Ullsten, S.; H Hansen, E.; Hammer Andersen, J. Anti-Bacterial Effect and Cytotoxicity Assessment of Lipid 430 Isolated from *Algibacter* sp. *Molecules* **2019**, *24*, 3991. [[CrossRef](#)]
36. Nemati, R.; Dietz, C.; Anstadt, E.J.; Cervantes, J.; Liu, Y.; Dewhirst, F.E.; Clark, R.B.; Finegold, S.; Gallagher, J.J.; Smith, M.B.; et al. Deposition and hydrolysis of serine dipeptide lipids of Bacteroidetes bacteria in human arteries: Relationship to atherosclerosis. *J. Lipid Res.* **2017**, *58*, 1999–2007. [[CrossRef](#)]
37. Blanksby, S.J.; Mitchell, T.W. Advances in mass spectrometry for lipidomics. *Annu. Rev. Anal. Chem.* **2010**, *3*, 433–465. [[CrossRef](#)]
38. Hancock, S.E.; Poad, B.L.J.; Batarseh, A.; Abbott, S.K.; Mitchell, T.W. Advances and unresolved challenges in the structural characterization of isomeric lipids. *Anal. Biochem.* **2017**, *524*, 45–55. [[CrossRef](#)] [[PubMed](#)]
39. Cajka, T.; Fiehn, O. Comprehensive analysis of lipids in biological systems by liquid chromatography-mass spectrometry. *Trends Analyt. Chem.* **2014**, *61*, 192–206. [[CrossRef](#)] [[PubMed](#)]
40. Swiezewska, E.; Wójcik, J. NMR of carbohydrates, lipids and membranes. In *Nuclear Magnetic Resonance: A Review of the Literature Published between January 2009 and May 2010*; Kamińska-Trela, K., Aliev, A.E., Eds.; RSC Publications: Cambridge, UK, 2011; pp. 344–390. ISBN 978-1-84973-147-8.
41. Swiezewska, E.; Wójcik, J. NMR of lipids and membranes. In *Nuclear Magnetic Resonance*; Kamińska-Trela, K., Wojcik, J., Eds.; Royal Society of Chemistry: Cambridge, UK, 2012; pp. 320–347; ISBN 978-1-84973-373-1.
42. Pikula, S.; Bandorowicz-Pikula, J.; Groves, P. NMR of lipids. In *Nuclear Magnetic Resonance*; Wójcik, J., Kamińska-Trela, K., Eds.; The Royal Society of Chemistry: Cambridge, UK, 2013; pp. 362–382; ISBN 978-1-84973-577-3.
43. Pikula, S.; Bandorowicz-Pikula, J.; Groves, P. NMR of lipids. In *Nuclear Magnetic Resonance*; Kamińska-Trela, K., Aliev, A.E., Bandorowicz-Pikula, J., Buda, S., D'Errico, G., de Dios, A.C., Eds.; Royal Society of Chemistry: Cambridge, UK, 2015; pp. 385–406; ISBN 978-1-78262-052-5.
44. Pikula, S.; Bandorowicz-Pikula, J.; Groves, P. Recent Advances in NMR Studies of Lipids. In *Annual Reports on NMR Spectroscopy*; Webb, G.A., Ed.; Elsevier Science: Burlington, NJ, USA, 2015; pp. 195–246; ISBN 9780128030905.
45. Alexandri, E.; Ahmed, R.; Siddiqui, H.; Choudhary, M.I.; Tsiafoulis, C.G.; Gerathanassis, I.P. High Resolution NMR Spectroscopy as a Structural and Analytical Tool for Unsaturated Lipids in Solution. *Molecules* **2017**, *22*, 1663. [[CrossRef](#)] [[PubMed](#)]
46. Brinkmann, S.; Kurz, M.; Patras, M.A.; Hartwig, C.; Marner, M.; Leis, B.; Billion, A.; Kleiner, Y.; Bauer, A.; Toti, L.; et al. Genomic and chemical decryption of the Bacteroidetes phylum for its potential to biosynthesize natural products. *BioRxiv* **2021**. [[CrossRef](#)]
47. Yasir, M.; Chung, E.J.; Song, G.C.; Bibi, F.; Jeon, C.O.; Chung, Y.R. *Chitinophaga eiseniae* sp. nov., isolated from vermicompost. *Int. J. Syst. Evol. Microbiol.* **2011**, *61*, 2373–2378. [[CrossRef](#)]
48. Vaz-Moreira, I.; Nobre, M.F.; Nunes, O.C.; Manaia, C.M. *Pseudosphingobacterium domesticum* gen. nov., sp. nov., isolated from home-made compost. *Int. J. Syst. Evol. Microbiol.* **2007**, *57*, 1535–1538. [[CrossRef](#)]
49. Siddiqi, M.Z.; Liu, Q.; Lee, S.Y.; Choi, K.D.; Im, W.-T. *Olivibacter ginsenosidimutans* sp nov., with ginsenoside converting activity isolated from compost, and reclassification of *Pseudosphingobacterium domesticum* as *Olivibacter domesticus* comb. nov. *Int. J. Syst. Evol. Microbiol.* **2018**, *68*, 2509–2514. [[CrossRef](#)] [[PubMed](#)]
50. Bhushan, R.; Brückner, H. Marfey's reagent for chiral amino acid analysis: A review. *Amino Acids* **2004**, *27*, 231–247. [[CrossRef](#)] [[PubMed](#)]
51. Woznica, A.; Cantley, A.M.; Beemelmans, C.; Freinkman, E.; Clardy, J.; King, N. Bacterial lipids activate, synergize, and inhibit a developmental switch in choanoflagellates. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 7894–7899. [[CrossRef](#)]
52. Kayganich, K.A.; Murphy, R.C. Fast atom bombardment tandem mass spectrometric identification of diacyl, alkylacyl, and alk-1-enylacyl molecular species of glycerophosphoethanolamine in human polymorphonuclear leukocytes. *Anal. Chem.* **1992**, *64*, 2965–2971. [[CrossRef](#)] [[PubMed](#)]
53. Murphy, R.C.; Harrison, K.A. Fast atom bombardment mass spectrometry of phospholipids. *Mass Spectrom. Rev.* **1994**, *13*, 57–75. [[CrossRef](#)]
54. Bérdy, J. Bioactive microbial metabolites. *J. Antibiot.* **2005**, *58*, 1–26. [[CrossRef](#)]
55. Makovitzki, A.; Avrahami, D.; Shai, Y. Ultrashort antibacterial and antifungal lipopeptides. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 15997–16002. [[CrossRef](#)]
56. Shai, Y. Mode of action of membrane active antimicrobial peptides. *Biopolymers* **2002**, *66*, 236–248. [[CrossRef](#)]
57. Gotoh, N.; Tanaka, S.; Nishino, T. Supersusceptibility to hydrophobic antimicrobial agents and cell surface hydrophobicity in *Branhamella catarrhalis*. *FEMS Microbiol. Lett.* **1989**, *59*, 211–213. [[CrossRef](#)]
58. Fomsgaard, J.S.; Fomsgaard, A.; Høiby, N.; Bruun, B.; Galanos, C. Comparative immunochemistry of lipopolysaccharides from *Branhamella catarrhalis* strains. *Infect. Immun.* **1991**, *59*, 3346–3349. [[CrossRef](#)] [[PubMed](#)]
59. Tsujimoto, H.; Gotoh, N.; Nishino, T. Diffusion of macrolide antibiotics through the outer membrane of *Moraxella catarrhalis*. *J. Infect. Chemother.* **1999**, *5*, 196–200. [[CrossRef](#)]
60. Oberpaul, M.; Zumkeller, C.M.; Culver, T.; Spohn, M.; Mihajlovic, S.; Leis, B.; Glaeser, S.P.; Plarre, R.; McMahon, D.P.; Hammann, P.; et al. High-Throughput Cultivation for the Selective Isolation of Acidobacteria From Termite Nests. *Front. Microbiol.* **2020**, *11*, 597628. [[CrossRef](#)]



61. Kämpfer, P.; Dott, W.; Martin, K.; Glaeser, S.P. *Rhodococcus defluvii* sp. nov., isolated from wastewater of a bioreactor and formal proposal to reclassify *Corynebacterium hoagii* and *Rhodococcus equi* as *Rhodococcus hoagii* comb. nov. *Int. J. Syst. Evol. Microbiol.* **2014**, *64*, 755–761. [[CrossRef](#)]
62. Matyash, V.; Liebisch, G.; Kurzchalia, T.V.; Shevchenko, A.; Schwudke, D. Lipid extraction by methyl-tert-butyl ether for high-throughput lipidomics. *J. Lipid Res.* **2008**, *49*, 1137–1146. [[CrossRef](#)]
63. Determination of minimum inhibitory concentrations (MICs) of antibacterial agents by broth dilution. *Clin. Microbiol. Infect.* **2003**, *9*, ix. [[CrossRef](#)]
64. Arendrup, M.C.; Guinea, J.; Cuenca-Estrella, M.; Meletiadou, J.; Mouton, J.W.; Lagrou, K.; Howard, S.J.; The Subcommittee on Antifungal Susceptibility Testing (AFST) of the ESCMID European Committee for Antimicrobial Susceptibility Testing (EUCAST). Eucast definitive document EDef 7.1: Method for the determination of broth dilution MICs of antifungal agents for fermentative yeasts. *Clin. Microbiol. Infect.* **2008**, *14*, 398–405. [[CrossRef](#)]

## Publication V



pubs.acs.org/acschemicalbiology

Articles

### Identification, Characterization, and Synthesis of Natural Parasitic Cysteine Protease Inhibitors: Pentacitidins Are More Potent Falcitidin Analogues

Stephan Brinkmann,<sup>∇</sup> Sandra Semmler,<sup>∇</sup> Christian Kersten, Maria A. Patras, Michael Kurz, Natalie Fuchs, Stefan J. Hammerschmidt, Jenny Legac, Peter E. Hammann, Andreas Vilcinskis, Philip J. Rosenthal, Tanja Schirmeister, Armin Bauer,\* and Till F. Schäberle\*



Cite This: *ACS Chem. Biol.* 2022, 17, 576–589



Read Online

**Copyright:** Copyright © 2022, American Chemical Society (permission for reuse: page - 136 -)

**Supporting Information** (not included in this dissertation file):

<https://pubs.acs.org/doi/10.1021/acscchembio.1c00861>

#### **Contribution:**

Stephan Brinkmann conceived and designed the study. Stephan Brinkmann identified and characterized the natural pentacitidins via molecular networking and isolated them. Stephan Brinkmann and Sandra Semmler characterized all >30 natural falcitidin analogs by analyzing the individual MS<sup>2</sup>-spectra of each compound. Synthesis has been performed by Sandra Semmler, while purification of the synthetic analogs has been done by Stephan Brinkmann. For activity and docking studies, the collaborations with the Johannes Gutenberg University Mainz (Prof. Dr. Tanja Schirmeister) and the University of California, San Francisco (Prof. Dr. Philip. J. Rosenthal) were initiated by Stephan Brinkmann and Till F. Schäberle. The biosynthetic gene clusters have been identified and analyzed by Stephan Brinkmann. Apart from the synthesis data and the docking study, he analyzed all data and prepared all figures for the manuscript. Stephan Brinkmann and Sandra Semmler wrote the draft of the manuscript as well as the following revisions.

# Identification, Characterization, and Synthesis of Natural Parasitic Cysteine Protease Inhibitors: Pentacitidins Are More Potent Falcitidin Analogues

Stephan Brinkmann,<sup>▽</sup> Sandra Semmler,<sup>▽</sup> Christian Kersten, Maria A. Patras, Michael Kurz, Natalie Fuchs, Stefan J. Hammerschmidt, Jenny Legac, Peter E. Hammann, Andreas Vilcinskas, Philip J. Rosenthal, Tanja Schirmeister, Armin Bauer,\* and Till F. Schäberle\*

Cite This: *ACS Chem. Biol.* 2022, 17, 576–589

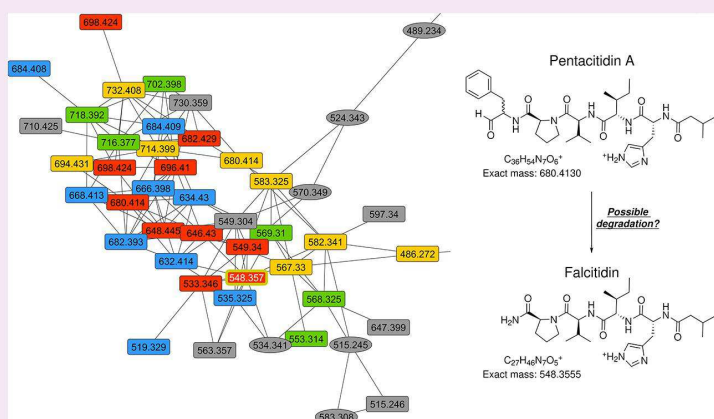
Read Online

ACCESS |

Metrics & More

Article Recommendations

Supporting Information



**ABSTRACT:** Protease inhibitors represent a promising therapeutic option for the treatment of parasitic diseases such as malaria and human African trypanosomiasis. Falcitidin was the first member of a new class of inhibitors of falcipain-2, a cysteine protease of the malaria parasite *Plasmodium falciparum*. Using a metabolomics dataset of 25 *Chitinophaga* strains for molecular networking enabled identification of over 30 natural analogues of falcitidin. Based on MS/MS spectra, they vary in their amino acid chain length, sequence, acyl residue, and C-terminal functionalization; therefore, they were grouped into the four falcitidin peptide families A–D. The isolation, characterization, and absolute structure elucidation of two falcitidin-related pentapeptide aldehyde analogues by extensive MS/MS spectrometry and NMR spectroscopy in combination with advanced Marfey's analysis was in agreement with the in silico analysis of the corresponding biosynthetic gene cluster. Total synthesis of chosen pentapeptide analogues followed by in vitro testing against a panel of proteases revealed selective parasitic cysteine protease inhibition and, additionally, low-micromolar inhibition of  $\alpha$ -chymotrypsin. The pentapeptides investigated here showed superior inhibitory activity compared to falcitidin.

## INTRODUCTION

Parasitic diseases such as malaria and sleeping sickness (human African trypanosomiasis, HAT) are poverty-associated diseases that have a massive impact on human life, especially in tropical areas.<sup>1,2</sup> Protozoan parasites of the *Trypanosoma* genus, transmitted by tsetse flies, are responsible for HAT.<sup>2</sup> With fewer than 3000 cases in 2015, the number of reported cases of HAT has reached a historically low level in recent years.<sup>2</sup> On the contrary, 229 million malaria cases and 409,000 deaths were estimated in 2019.<sup>3</sup> *Plasmodium falciparum* is the most virulent malaria parasite, causing nearly all deaths<sup>3</sup> as well as most drug-resistant infections.<sup>4,5</sup> Resistance to most available antimalarials, including artemisinin-based combination regi-

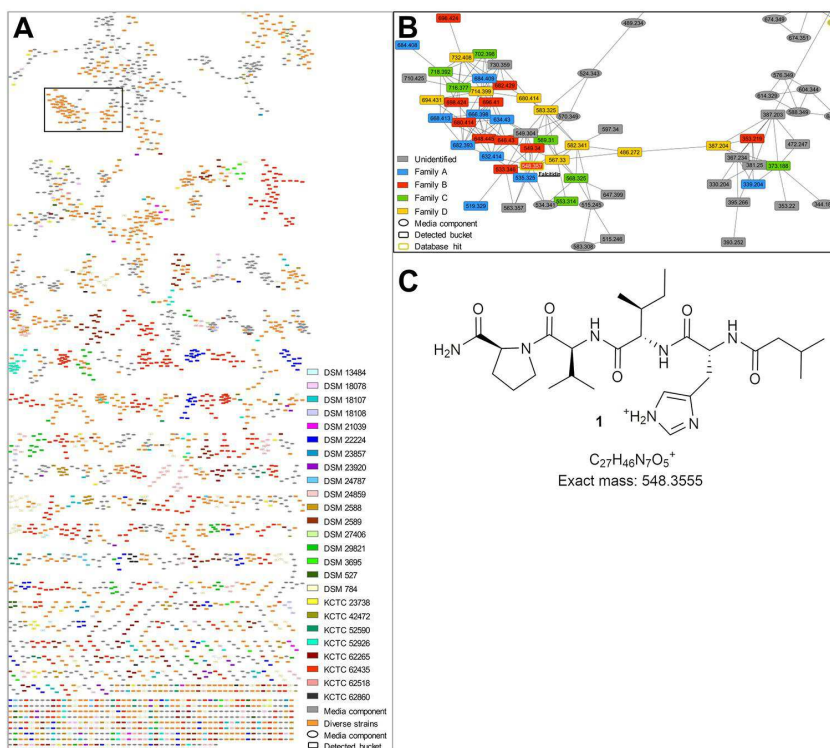
mens that are the standard treatment for falciparum malaria, is of great concern. Hence, there is a great need to discover new active compounds to serve as lead structures for the development of new antimalarial drugs.<sup>5,4</sup>

**Received:** October 30, 2021

**Accepted:** February 25, 2022

**Published:** March 9, 2022





**Figure 1.** MS<sup>2</sup>-networking analysis revealed more than 30 natural falcitidin analogues. (a) Complete MS<sup>2</sup>-network based on extracts of 25 *Chitinophaga* strains. (b) Close view of the falcitidin network color-coded according to the four molecule families A–D. (c) Molecular structure, formula, and calculated mass of falcitidin (1), which was putatively annotated for the parent ion of  $m/z$  548.357 [ $M + H$ ]<sup>+</sup>.

Papain-family cysteine proteases represent interesting therapeutic targets against several infectious diseases, including malaria (falcipains) and HAT (rhodesain).<sup>6</sup> Synthetic or natural product-based small molecule inhibitors were investigated and proved falcipains and rhodesain to be promising targets.<sup>7,8</sup> However, developing compounds that target these cysteine proteases into drugs has proved challenging, in part due to the difficulty of achieving selectivity.<sup>9</sup> The acyltetrapeptide falcitidin and its synthetic analogues were described as first members of a new class of cysteine protease inhibitors. In a falcipain-2 assay, the natural compound showed an IC<sub>50</sub> value in the low micromolar range, and the optimized synthetic analogues showed sub-micromolar IC<sub>50</sub> activity against *Plasmodium falciparum* in a standard blood-cell assay. Instead of reactive groups that covalent-reversibly or covalent-irreversibly bind to active-site cysteines, these compounds share a C-terminal amidated proline.<sup>10,11</sup>

High structural diversity in combination with diverse biological activities make natural products a valuable source in the search for new lead candidates for drug development. Technological improvements in resolution and accuracy of spectrometry methods followed by complex and large dataset analysis allow the application of metabolomics techniques based on, for example, ultrahigh-performance liquid chromatography in line with high-resolution tandem mass spectrometry (UHPLC-HRMS/MS) for the discovery and characterization of these metabolites present in a given sample.<sup>12</sup> Data visualization and interpretation using tandem mass spectrometry (MS/MS) networks (molecular networking) allow the

automatic annotation of MS/MS spectra against library compounds as well as the identification of signals of interest and their analogues.<sup>13</sup>

In this study, molecular networking was applied to analyze a previously generated dataset based on bacterial organic extracts originated from 25 *Chitinophaga* strains (phylum Bacteroidetes) in which falcitidin was putatively annotated.<sup>14</sup> We aimed to investigate if additional natural analogues were produced by these strains and discovered more than 30 natural analogues. A gene encoding a multimodular nonribosomal peptide synthetase was identified and linked in silico to the production of the pentapeptide aldehydes. After isolation and structure elucidation of two natural aldehyde analogues, total synthesis of all four pentapeptide aldehydes and their carboxylic acid and alcohol derivatives was successfully achieved. In activity testing against a panel of proteases, they performed better than the reference falcitidin and revealed low-micromolar IC<sub>50</sub> activity against  $\alpha$ -chymotrypsin, falcipain-2, and falcipain-3. Three of them also showed activity against rhodesain.

## RESULTS AND DISCUSSION

**Discovery of Falcitidin Analogues.** Exploration of the genomic and metabolomic potential of the Bacteroidetes genus *Chitinophaga* revealed a high potential to find chemical novelty.<sup>14</sup> In the framework of this previous metabolomic study of *Chitinophaga*, an inhibitor of the antimalarial target falcipain-2, falcitidin (1),<sup>10</sup> was detected.<sup>14</sup> Originally wrongly described as myxobacterium-derived tetrapeptides produced

Table 1. Overview of Molecular Features of all Identified Natural Falcitidin Analogues Aligned into Four Protein Families

family	fatty acid residue	amino acid chain					C-terminal group	molecular formula	molecular mass			ion	cpd no.
		1	2	3	4	5			calcd	MS <sup>2</sup> network	found		
A	iVal	H	V				COOH	C <sub>16</sub> H <sub>27</sub> N <sub>4</sub> O <sub>4</sub>	339.2027	339.204	339.2030	[M+H] <sup>+</sup>	
	iVal	H	V	V	P		CHO	C <sub>26</sub> H <sub>43</sub> N <sub>6</sub> O <sub>5</sub>	519.3289	519.329	519.3297	[M+H] <sup>+</sup>	
	iVal	H	V	V	P		CONH <sub>2</sub>	C <sub>26</sub> H <sub>44</sub> N <sub>7</sub> O <sub>5</sub>	534.3398	534.341	534.3402	[M+H] <sup>+</sup>	
	iVal	H	V	V	P		COOH	C <sub>26</sub> H <sub>43</sub> N <sub>6</sub> O <sub>6</sub>	535.3239	535.325	535.3244	[M+H] <sup>+</sup>	
	iVal	H	V	V	P	L	CHO	C <sub>32</sub> H <sub>54</sub> N <sub>7</sub> O <sub>6</sub>	632.4130	632.414	632.4148	[M+H] <sup>+</sup>	
	iVal	H	V	V	P	L	CH <sub>2</sub> OH	C <sub>32</sub> H <sub>56</sub> N <sub>7</sub> O <sub>6</sub>	634.4287	634.430	634.4298	[M+H] <sup>+</sup>	
	iVal	H	V	V	P	F	CHO	C <sub>33</sub> H <sub>52</sub> N <sub>7</sub> O <sub>6</sub>	666.3974	666.398	666.3982	[M+H] <sup>+</sup>	15
	iVal	H	V	V	P	F	CH <sub>2</sub> OH	C <sub>33</sub> H <sub>54</sub> N <sub>7</sub> O <sub>6</sub>	668.4130	668.413	668.4117	[M+H] <sup>+</sup>	12
	iVal	H	V	V	P	F	COOH	C <sub>33</sub> H <sub>52</sub> N <sub>7</sub> O <sub>7</sub>	682.3923	682.393	682.3918	[M+H] <sup>+</sup>	8
B	iVal	H	V	V	P	F	CHO	C <sub>33</sub> H <sub>54</sub> N <sub>7</sub> O <sub>7</sub>	684.4079	684.408	684.4085	[M+H <sub>2</sub> O+H] <sup>+</sup>	15 <sup>a</sup>
	iVal	H	I				COOH	C <sub>17</sub> H <sub>29</sub> N <sub>4</sub> O <sub>4</sub>	353.2183	353.219	353.2185	[M+H] <sup>+</sup>	
	iVal	H	I	V	P		CHO	C <sub>27</sub> H <sub>45</sub> N <sub>6</sub> O <sub>5</sub>	533.3446	533.346	533.3442	[M+H] <sup>+</sup>	
	iVal	H	I	V	P		CONH <sub>2</sub>	C <sub>27</sub> H <sub>46</sub> N <sub>7</sub> O <sub>5</sub>	548.3555	548.357	548.3563	[M+H] <sup>+</sup>	1
	iVal	H	I	V	P		COOH	C <sub>27</sub> H <sub>45</sub> N <sub>6</sub> O <sub>6</sub>	549.3395	549.340	549.3397	[M+H] <sup>+</sup>	17
	iVal	H	I	V	P	L	CHO	C <sub>33</sub> H <sub>56</sub> N <sub>7</sub> O <sub>6</sub>	646.4287	646.430	646.4305	[M+H] <sup>+</sup>	
	iVal	H	I	V	P	L	CH <sub>2</sub> OH	C <sub>33</sub> H <sub>58</sub> N <sub>7</sub> O <sub>6</sub>	648.4443	648.445	648.4445	[M+H] <sup>+</sup>	
	iVal	H	I	V	P	F	CHO	C <sub>36</sub> H <sub>54</sub> N <sub>7</sub> O <sub>6</sub>	680.4130	680.414	680.4131	[M+H] <sup>+</sup>	2
	iVal	H	I	V	P	F	CH <sub>2</sub> OH	C <sub>36</sub> H <sub>56</sub> N <sub>7</sub> O <sub>6</sub>	682.4287	682.429	682.4270	[M+H] <sup>+</sup>	14
C	iVal	H	I	V	P	F	COOH	C <sub>36</sub> H <sub>54</sub> N <sub>7</sub> O <sub>7</sub>	696.4079	696.410	696.4072	[M+H] <sup>+</sup>	10
	iVal	H	I	V	P	F	CHO	C <sub>36</sub> H <sub>56</sub> N <sub>7</sub> O <sub>7</sub>	698.4236	698.424	698.4238	[M+H <sub>2</sub> O+H] <sup>+</sup>	2 <sup>a</sup>
	PA	H	V				COOH	C <sub>19</sub> H <sub>25</sub> N <sub>4</sub> O <sub>4</sub>	373.1870	373.188	373.1874	[M+H] <sup>+</sup>	
	PA	H	V	V			COOH	C <sub>24</sub> H <sub>34</sub> N <sub>5</sub> O <sub>5</sub>	472.2554	not present	472.2560	[M+H] <sup>+</sup>	
	PA	H	V	V	P		CHO	C <sub>29</sub> H <sub>41</sub> N <sub>6</sub> O <sub>5</sub>	553.3133	553.314	553.3135	[M+H] <sup>+</sup>	
	PA	H	V	V	P		CONH <sub>2</sub>	C <sub>29</sub> H <sub>42</sub> N <sub>7</sub> O <sub>5</sub>	568.3242	568.325	568.3248	[M+H] <sup>+</sup>	
	PA	H	V	V	P		COOH	C <sub>29</sub> H <sub>41</sub> N <sub>6</sub> O <sub>6</sub>	569.3082	569.310	569.3101	[M+H] <sup>+</sup>	
	PA	H	V	V	P	F	CHO	C <sub>38</sub> H <sub>50</sub> N <sub>7</sub> O <sub>6</sub>	700.3817	not present	700.3806	[M+H] <sup>+</sup>	16
	PA	H	V	V	P	F	CH <sub>2</sub> OH	C <sub>38</sub> H <sub>52</sub> N <sub>7</sub> O <sub>6</sub>	702.3974	702.398	702.3956	[M+H] <sup>+</sup>	13
D	PA	H	V	V	P	F	COOH	C <sub>38</sub> H <sub>50</sub> N <sub>7</sub> O <sub>7</sub>	716.3766	716.377	716.3765	[M+H] <sup>+</sup>	9
	PA	H	V	V	P	F	CHO	C <sub>38</sub> H <sub>52</sub> N <sub>7</sub> O <sub>7</sub>	718.3923	718.392	718.3923	[M+H <sub>2</sub> O+H] <sup>+</sup>	16 <sup>a</sup>
	PA	H	I				COOH	C <sub>30</sub> H <sub>27</sub> N <sub>4</sub> O <sub>4</sub>	387.2027	387.203	387.2033	[M+H] <sup>+</sup>	
	PA	H	I	V			COOH	C <sub>25</sub> H <sub>36</sub> N <sub>5</sub> O <sub>5</sub>	486.2711	486.272	486.2715	[M+H] <sup>+</sup>	
	PA	H	I	V	P		CHO	C <sub>30</sub> H <sub>43</sub> N <sub>6</sub> O <sub>5</sub>	567.3289	567.330	567.3300	[M+H] <sup>+</sup>	
	PA	H	I	V	P		CONH <sub>2</sub>	C <sub>30</sub> H <sub>44</sub> N <sub>7</sub> O <sub>5</sub>	582.3398	582.341	582.3408	[M+H] <sup>+</sup>	
	PA	H	I	V	P		COOH	C <sub>30</sub> H <sub>43</sub> N <sub>6</sub> O <sub>6</sub>	583.3239	583.325	583.3252	[M+H] <sup>+</sup>	
	PA	H	I	V	P	L	CHO	C <sub>36</sub> H <sub>54</sub> N <sub>7</sub> O <sub>6</sub>	680.4130	680.414	680.4115	[M+H] <sup>+</sup>	
	PA	H	I	V	P	L	CH <sub>2</sub> OH	C <sub>36</sub> H <sub>56</sub> N <sub>7</sub> O <sub>6</sub>	682.4287	not present	682.4288	[M+H] <sup>+</sup>	
D	PA	H	I	V	P	F	CHO	C <sub>39</sub> H <sub>52</sub> N <sub>7</sub> O <sub>6</sub>	714.3974	714.399	714.3975	[M+H] <sup>+</sup>	3
	PA	H	I	V	P	F	CH <sub>2</sub> OH	C <sub>39</sub> H <sub>54</sub> N <sub>7</sub> O <sub>6</sub>	716.4136	not present	716.4132	[M+H] <sup>+</sup>	4
	PA	H	I	V	P	F	COOH	C <sub>39</sub> H <sub>52</sub> N <sub>7</sub> O <sub>7</sub>	730.3923	not present	730.3939	[M+H] <sup>+</sup>	11
	PA	H	I	V	P	F	CHO	C <sub>39</sub> H <sub>54</sub> N <sub>7</sub> O <sub>7</sub>	732.4079	732.408	732.4082	[M+H <sub>2</sub> O+H] <sup>+</sup>	3 <sup>a</sup>

<sup>a</sup>Nodes found in MS<sup>2</sup> network, no MS/MS spectra found due to water ion adduct.

by *Chitinophaga* sp. Y23, falcitidin and its synthetic analogues are first members of a new class of cysteine protease inhibitors without a reactive group that covalent-reversibly or covalent-irreversibly binds to the active-site cysteine residue.<sup>10,11</sup> Based on extracts of all 25 previously cultured *Chitinophaga* strains,<sup>14</sup> here, a complementary molecular networking study was carried out to analyze the chemical profiles (Figure 1a). This facilitated the discovery of more than 30 natural analogues of falcitidin (1,  $m/z$  548.3563 [M+H]<sup>+</sup>, molecular formula C<sub>27</sub>H<sub>46</sub>N<sub>7</sub>O<sub>5</sub>) (Figure 1b,c and Table 1) biosynthesized by *Chitinophaga eiseniae* DSM 22224, *Chitinophaga dinghuensis* DSM 29821, and *Chitinophaga varians* KCTC 52926. Analysis of the corresponding MS/MS spectra led to the assignment of four peptide families that differed in the amino acid chain at position 2 and the acyl residue of the peptide. Members of families A and B shared a fragment ion of  $m/z$  222.1237 [M

+H]<sup>+</sup> corresponding to the molecular formula C<sub>11</sub>H<sub>16</sub>N<sub>3</sub>O<sub>2</sub><sup>+</sup>. The fragment can be assigned to the described isovaleroyl (iVal) residue attached to the first amino acid histidine.<sup>10</sup> In contrast, a fragment ion of  $m/z$  256.1084 [M+H]<sup>+</sup>, corresponding to the molecular formula C<sub>14</sub>H<sub>14</sub>N<sub>3</sub>O<sub>2</sub><sup>+</sup>, was detected in peptides of families C and D (Figures S1–S4). A difference within the acyl moiety was assumed due to the identical number of three nitrogen atoms, also indicating a histidine at position one. Moreover, the MS/MS fragmentation pattern of family members A and B as well as C and D varied in position 2 of the amino acid chain, either carrying valine or isoleucine. Within each peptide family, members differed in chain length from two to five amino acids and their C-terminal functional group (Table 1). Interestingly, a corresponding falcitidin analogue with the unusual C-terminal amidated (CONH<sub>2</sub>-group) proline, indicated by a neutral loss of

**Table 2. NMR Data Comparison for Pentacitidin A (2) in DMSO-*d*<sub>6</sub> for the Natural Isolated Compound (700 MHz, 176 MHz) and the Synthetic Compound (600 MHz, 151 MHz)<sup>a</sup>**

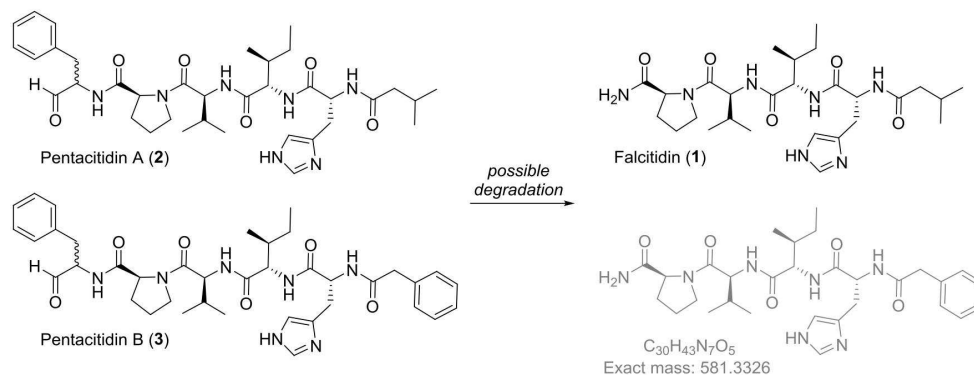
position	$\delta^{13}\text{C}$ (ppm)		$\delta^1\text{H}$ (ppm) (m, f, J)		position	$\delta^{13}\text{C}$ (ppm)		$\delta^1\text{H}$ (ppm) (m, f, J)	
	natural isolated	synthetic	natural isolated	synthetic		natural isolated	synthetic	natural isolated	synthetic
phenylalaninal									
CHO	200.3	200.3	9.47, 9.41	9.47 (s, 0.2H), 9.41 (s, 0.2H)		18.5	18.45/18.41	0.86	
NH			8.38	8.36 (d, 0.2H, J = 7.7 Hz), 8.34 (d, 0.2H, J = 7.0 Hz)				7.82	7.69–7.63 (m, 1H)
$\alpha$ -CH	59.5	59.7/59.5	4.30	4.35–4.24 (m, 2H, overlay), 4.24–4.16 (m, 2H, overlay)	$\alpha$ -CH	56.3 <sup>b</sup>	56.32/56.29	4.25	4.24–4.16 (m, 2H)
$\beta$ -CH <sub>2</sub>	33.5	33.6/33.3	3.13, 2.74	3.15–3.07 (m, 1H), 2.91–2.83 (m, 2H, overlay), 2.78–2.69 (m, 2H, overlay)	$\beta$ -CH	36.9 <sup>b</sup>	36.9/36.8	1.65	1.70–1.60 (m, 2H)
$\gamma$ -C <sub>quart</sub>	137.6	137.6			$\gamma$ -CH <sub>2</sub>	24.0	24.0	1.25	1.30–1.22 (m, 1H)
CH <sub>arom</sub>	129.3, 128.1, 126.3, 126.2	129.3, 129.2, 129.1, 128.9, 128.2, 128.1, 126.24, 126.19	7.23, 7.27, n.a.	7.29–7.12 (m, 5H)	$\gamma$ -CH <sub>3</sub>	15.2	15.2	0.71	0.76–0.67 (m, 6H, overlay)
proline					$\delta$ -CH <sub>3</sub>	11.0	11.0	0.74	0.76–0.67 (m, 6H, overlay)
CO	172.0 <sup>b</sup>	172.1/172.0			histidine				
$\alpha$ -CH	59.1 <sup>b</sup>	59.1/59.0	4.29	4.35–4.24 (m, 2H, overlay)	CO	170.0	170.9		
$\beta$ -CH <sub>2</sub>	29.3	29.31/29.28	1.92, 1.54	2.02–1.83 (m, 5H, overlay), 1.83–1.70 (m, 2H, overlay), 1.60–1.51 (m, 1H)	NH			8.16	8.00–7.94 (m, 2H, overlay)
$\gamma$ -CH <sub>2</sub>	24.4 <sup>b</sup>	24.44/24.38	1.79, 1.74	2.02–1.83 (m, 5H, overlay), 1.83–1.70 (m, 2H, overlay), 3.47 (m, 1H)	$\alpha$ -CH	51.7 <sup>b</sup>	52.7	4.71	4.61–4.55 (m, 1H)
$\delta$ -CH <sub>2</sub> N	47.0 <sup>b</sup>	47.00/46.96	3.75, 3.53	3.78–3.68 (m, 1H), 3.57–3.47 (m, 1H)	$\beta$ -CH <sub>2</sub>	27.9 <sup>b</sup>	29.9*	3.03	2.91–2.83 (m, 2H, overlay)
valine					$\gamma$ -C <sub>quart</sub>	n.a.	134.2		
CO	169.6	169.58/169.55			$\delta$ -CH <sub>arom</sub>	116.7 <sup>b</sup>	n.o.	~7.2	6.76 (s, 1H)
NH			8.05	8.00–7.94 (m, 2H, overlay)	$\epsilon$ -CH <sub>arom</sub>	133.8	134.5	~8.7	7.50 (s, 1H)
$\alpha$ -CH	55.8	55.7	4.25	4.35–4.24 (m, 2H, overlay)	$\epsilon$ -NH			n.o.	n.o.
$\beta$ -CH	29.7	29.7	1.98	2.02–1.83 (m, 5H, overlay)	isovaleroyl				
$\gamma$ -CH <sub>3</sub>	19.0	19.03/19.00	0.88	0.90–0.84 (m, 6H, overlay)	CO	171.6	171.4		

<sup>a</sup>Due to the presence of two diastereomers (see above), two sets of signals are obtained. Therefore, for some signals, two chemical shift values are given. n.o., not observed; n.a., not assigned due to line broadening; asterisks denote data obtained from the HMBC or HSQC spectrum; overlay means that the signal overlaps with another signal. <sup>b</sup>Broad signal.

114.0794 Da, was found in each family. Falcitidin (iVal-H-I-V-P-CONH<sub>2</sub>) itself belongs to family B. Its analogues with the same neutral loss were assigned to the UHR-ESI-MS ion peaks at *m/z* 534.3402 [M+H]<sup>+</sup> (C<sub>26</sub>H<sub>44</sub>N<sub>7</sub>O<sub>5</sub>, family A), *m/z* 568.3248 [M+H]<sup>+</sup> (C<sub>29</sub>H<sub>42</sub>N<sub>7</sub>O<sub>5</sub>, family C), and *m/z* 582.3408 [M+H]<sup>+</sup> (C<sub>30</sub>H<sub>44</sub>N<sub>7</sub>O<sub>5</sub>, family D). Additionally, within all families, tetrapeptide analogues carrying a C-terminal-unmodified (COOH-group) proline and analogues with an aldehyde moiety (CHO-group) indicated by a neutral loss of 99.0686 Da (C<sub>3</sub>H<sub>9</sub>NO) were found. Besides these tetrapeptide analogues of falcitidin, truncated di- or tripeptides with C-terminal COOH-groups and larger pentapeptide analogues were predicted based on MS/MS spectra. The spectra of the pentapeptides, which were found in all families, revealed either

phenylalanine or leucine/isoleucine in the C-terminal position. Besides an unmodified C-terminal phenylalanine (COOH-group), additional analogues carrying a C-terminal phenylalaninal (CHO-group) and phenylalaninol (CH<sub>2</sub>OH-group) were predicted. The last two analogues were proposed based on the neutral losses of 149.0841 Da (C<sub>9</sub>H<sub>11</sub>NO) and 151.1006 Da (C<sub>9</sub>H<sub>13</sub>NO), respectively. For analogues with a leucine/isoleucine at position 5, only analogues with a CHO- and CH<sub>2</sub>OH-group were detected, with neutral losses of 115.1031 Da (C<sub>6</sub>H<sub>13</sub>NO) and 117.1164 Da (C<sub>6</sub>H<sub>15</sub>NO), respectively (Table 1 and Figures S1–S4).

**Natural Compound Isolation and Spectroscopic Analysis.** In order to confirm the proposed structures of the pentapeptide falcitidin analogues, their functional C-terminal



**Figure 2.** Chemical structures of isolated novel pentacitidin A (2) and B (3) as well as their possible degradation analogues, falcitidin (1) and the corresponding tetrapeptide to pentacitidin B. Black represents synthesized structures; gray represents MS<sup>2</sup> spectra-confirmed structures.

groups and the unknown acyl chain of peptide families C and D, isolation and spectroscopic analysis of natural pentapeptides with  $m/z$  680.4131 [M+H]<sup>+</sup> (2, C<sub>36</sub>H<sub>54</sub>N<sub>7</sub>O<sub>6</sub>),  $m/z$  714.3975 [M+H]<sup>+</sup> (3, C<sub>39</sub>H<sub>52</sub>N<sub>7</sub>O<sub>6</sub>), and  $m/z$  716.4132 [M+H]<sup>+</sup> (4, C<sub>39</sub>H<sub>54</sub>N<sub>7</sub>O<sub>6</sub>) were achieved. Isolation was performed by the adsorption of the metabolites on XAD16N resin and sequential C18-RP-HPLC followed by C18-RP-UHPLC fractionations starting from a 20 L fermentation of *C. eiseniae* DSM 22224. Based on the analysis of several two-dimensional NMR spectra, including DQF-COSY, TOCSY, ROESY, multiplicity-edited HSQC, and HMBC experiments, the structure of the isolated compound 2 was determined to be the phenylalaninal extended version of falcitidin (iVal-H-I-V-P-F-CHO) determined from the additional aromatic signals  $\delta_{\text{H}} = 7.27$  and 7.23 ppm as well as the aldehyde protons at  $\delta_{\text{H}} = 9.47$  and 9.41 ppm and  $\delta_{\text{C}} = 200.3$  ppm (Table 2 and Figure S5). We proposed the name pentacitidin A. The presence of two signals for several protons and carbon atoms indicated an epimerization of the stereogenic center of the C-terminal phenylalaninal (Figure 2). The analysis of the NMR data of the isolated compound 3 revealed the same peptide sequence. However, instead of the isovaleroyl moiety, 3 contains a phenyl acetyl (PA) residue at the N terminus, leading to the final sequence PA-H-I-V-P-F-CHO (Table 3 and Figure S6). Again, the protons of the terminal aldehyde moiety gave rise to a set of two signals (e.g.,  $\delta_{\text{H}}$  of aldehyde proton = 9.46 and 9.41 ppm), indicating the epimerization of its stereogenic center (Figure 2). Accordingly, we proposed the name pentacitidin B for compound 3. The NMR sample of compound 4 contained a mixture of 4 and aldehyde 3, not allowing for a complete assignment of the NMR signals. The phenylalaninol moiety, as indicated by MS/MS results, was confirmed by the presence of a second methylene group adjacent to C <sub>$\alpha$</sub>  ( $\delta_{\text{H}} = 3.27$  ppm,  $\delta_{13\text{C}} = 61.8$  ppm), which in the COSY spectrum is correlated with an exchangeable proton at 4.7 ppm (OH). The reduction of the aldehyde function also leads to a significant high-field shift of the H <sub>$\alpha$</sub>  and the amide proton (Figure S7), verifying the structure of compound 4 to be PA-H-I-V-P-F-CH<sub>2</sub>OH (Scheme 1).

Finally, advanced Marfey's analysis was conducted to determine the absolute configuration of the amino acids.<sup>15</sup> Total hydrolysis of a sample containing 3 and 4 followed by chemical derivatization with *N* <sub>$\alpha$</sub> -(2,4-dinitro-5-fluorophenyl)-L-valinamide (Marfey's reagent) and UHPLC-MS comparison to reference substrates confirmed the literature-known D-His-L-

Val/Ile-L-Val-L-Pro configuration<sup>10</sup> and the incorporation of L-phenylalanine at position 5 by verifying the presence of L-phenylalaninol (Figure S9). All stereogenic centers were further confirmed by total synthesis and comparison of NMR data to natural isolated compounds (Tables 2 and 3).

**Identification of the Biosynthetic Gene Cluster.** The publicly available genomes of all three producers, *C. eiseniae* DSM 22224 (FUWZ01000000), *C. dinghuensis* DSM 29821 (QLMA01000000), and *C. varians* KCTC 52926 (JACVFB01000000), were scanned with antiSMASH<sup>16</sup> for NRPS-type biosynthetic gene clusters (BGCs) matching the structural features of the molecules. The number and predicted substrate specificity of the A-domains (Table S1) as well as the precursor supply and post-assembly modifications were taken into account. BGCs congruent to the pentapeptide structure were identified in each case. Furthermore, an epimerization domain to catalyze the conversion of L- to D-amino acids is positioned in agreement with the determined stereochemistry of the molecules. Interestingly, manual search of other publicly available *Chitinophaga* genomes revealed the presence of similar BGCs in genomes of *C. varians* Ae27 (JA-BAIA01000000) and *C. niastensis* DSM 24859 (PYAW00000000) (Figure 3a), the latter also being included in the molecular networking analysis. For *C. niastensis*, however, no production of pentacitidins was observed, which could correlate with the small differences in the A-domain specificity for amino acids 3 (His instead of Val) and 5 (beta-hydroxy-tyrosine (Bht) instead of Phe) (Table S1). The MAFFT alignment<sup>17</sup> of all five BGCs allowed clear cluster boarder prediction and revealed minor variations in the upstream region. In conclusion, pentacitidin biosynthesis is putatively encoded by a single NRPS core gene of 18.5 kbp in length. No additional genes encoding proteins involved in further modifications or transportation were conserved between all five BGCs (Figure 3a). The NRPS starts with a starter condensation (C-starter) domain responsible for the addition of iVal or PA to the peptide core. A terminal reductase domain (TD) at the C-terminal end in four BGCs should be responsible for the reductive release process (Figure 3b).<sup>18</sup> The presence of alcohol and aldehyde pentapeptides suggests a similar reduction as shown for example in the biosynthesis of the siderophore myxochelin.<sup>19–21</sup> The (peptidyl)acyl thioester attached to the carrier protein could be reduced first to an aldehyde and then to an alcohol via a four-electron reduction during the product release.<sup>22,23</sup> The

Table 3. NMR Data Comparison for Pentacitidin B (3) in DMSO-*d*<sub>6</sub> for the Natural Isolated Compound (500 MHz, 126 MHz) and the Synthetic Compound (600 MHz, 151 MHz)<sup>a</sup>

position	$\delta^{13}\text{C}$ (ppm)		$\delta^1\text{H}$ (ppm) (m, <i>f</i> , <i>J</i> )	
	natural isolated (major)	synthetic	natural isolated (major)	synthetic
phenylalaninal				
CHO	200.3, 200.5	200.3	9.46, 9.41	9.47 (s, 0.2H) 9.41 (s, 0.2H)
NH			8.38, 8.36	8.36 (d, 0.25H, <i>J</i> = 7.7 Hz) 8.34 (d, 0.25H, <i>J</i> = 7.0 Hz)
$\alpha$ -CH	59.5, 59.8	59.1 / 59.0	4.29, 4.17	4.24–4.16 (m, 2H, overlay)
$\beta$ -CH <sub>2</sub>	33.5, 33.3	33.5 / 33.3	3.12/2.74, 3.09/2.85	3.15–3.06 (m, 1H, overlay) 2.93–2.83 (m, 2H, overlay) 2.78–2.71 (m, 1H, overlay)
$\gamma$ -C <sub>quart</sub>	137.7	137.6		
CH <sub>arom</sub>	129.3, 129.2128.1, 128.2126.2, 126.3	129.3, 129.2, 129.14, 129.09, 129.0, 128.2, 128.1, 128.0, 126.25, 126.20, 126.16	7.21, 7.24, 7.18	7.29–7.13 (m, 10H, overlay)
proline				
CO	172.0	172.1 / 172.0		
$\alpha$ -CH	59.0, 59.1	59.5	4.29, 4.30	4.33–4.24 (m, 2H, overlay)
$\beta$ -CH <sub>2</sub>	29.3, 29.4	29.32 / 29.29	1.91/1.54, 1.97/1.74	2.02–1.84 (m, 2H, overlay) 1.84–1.69 (m, 2H, overlay) 1.59–1.51 (m, 1H)
$\gamma$ -CH <sub>2</sub>	24.4, 24.5	24.5 / 24.4	1.78/1.74, 1.91/1.80	2.02–1.84 (m, 2H, overlay) 1.84–1.69 (m, 2H, overlay)
$\delta$ -CH <sub>2</sub> N	46.98, 47.02	47.01 / 46.97	3.75/3.50, 3.75/3.54	3.78–3.68 (m, 1H) 3.57–3.47 (m, 1H)
valine				
CO	169.58	169.61 / 169.58		
NH			7.98, 7.96	7.95 (t, 1H, <i>J</i> = 9.0 Hz)
$\alpha$ -CH	55.8	55.8	4.26	4.33–4.24 (m, 2H, overlay)
$\beta$ -CH	29.7	29.7	1.97	2.02–1.84 (m, 2H, overlay)
$\gamma$ -CH <sub>3</sub>	19.0 18.5	19.03 / 19.00 18.5 / 18.4	0.87, 0.85	0.94–0.82 (m, 7H, overlay)
isoleucine				
CO	170.75	170.7		
NH			7.76	7.77–7.71 (m, 1H)
$\alpha$ -CH	56.39	56.40 / 56.37	4.21	4.24–4.16 (m, 2H, overlay)
$\beta$ -CH	36.79	36.8 / 36.7	1.64	1.69–1.59 (m, 1H)
$\gamma$ -CH <sub>2</sub>	23.9	23.9	1.22, 0.88	1.27–1.17 (m, 1H) 0.94–0.82 (m, 7H, overlay)
$\gamma$ -CH <sub>3</sub>	15.3	15.3	0.69	0.75–0.66 (m, 6H, overlay)
$\delta$ -CH <sub>3</sub>	11.0	11.0	0.72	0.75–0.66 (m, 6H, overlay)
histidine				
CO	170.8	170.8		
NH			8.29	8.26 (d, 1H, <i>J</i> = 8.1 Hz)
$\alpha$ -CH	52.9	52.9	4.58	4.62–4.55 (m, 1H)
$\beta$ -CH <sub>2</sub>	n.a.	30.1*	2.89, 2.75	2.93–2.83 (m, 2H, overlay) 2.78–2.71 (m, 1H, overlay)



Table 3. continued

position	$\delta^{13}\text{C}$ (ppm)		$\delta^1\text{H}$ (ppm) (m, f, j)	
	natural isolated (major)	synthetic	natural isolated (major)	synthetic
$\gamma\text{-C}_{\text{quart}}$	n.a.	n.o.		
$\delta\text{-CH}_{\text{arom}}$	n.a.	n.o.	6.75	6.76 (s, 1H)
$\epsilon\text{-CH}_{\text{arom}}$	134.5	134.5	7.49	7.51 (s, 1H)
$\epsilon\text{-NH}$				n.o.
phenylacetyl				
CO	169.9	169.9		
$\text{CH}_2$	42.0	42.0	3.43	3.45 (d, 1H, $J = 14.1$ Hz) 3.42 (d, 1H, $J = 14.1$ Hz)
$\text{C}_{\text{quart}}$	136.2	136.2		
$\text{CH}_{\text{arom}}$	128.97, 128.05, 126.18	129.3, 129.2, 129.14, 129.09, 129.0, 128.2, 128.1, 128.0, 126.25, 126.20, 126.16	7.17, 7.22, 7.19	7.29–7.13 (m, 10H, overlay)

<sup>a</sup>n.o., not observed; n.a., not assigned due to line broadening; asterisks denote data obtained from the HMBC or HSQC spectrum; overlay means that the signal overlaps with another signal.

BGC in the genome of *C. varians* Ae27 carries a NAD\_binding\_4 domain instead of the TD. This domain family is sequence-related to the C-terminal region of the male sterility protein in the arabidopsis species<sup>24</sup> and a jojoba acyl-CoA reductase.<sup>25</sup> The latter is known to catalyze a similar reduction reaction, with the formation of a fatty alcohol from a fatty acyl substrate.<sup>25</sup>

**Degradation Hypotheses.** A BGC congruent to the pentapeptide structure and the presence of the unusual tetrapeptide analogue after NMR study (Figure S5) of its corresponding pentapeptide aldehyde points toward falcitidin and its identified analogues from families A, C, and D with their unusual C-terminal amidated proline to be degradation products (Figure 2). Biochemically, a cleavage by a carboxypeptidase in the presence of ammonia could result in the C-terminal amide group of the proline, which can explain the tetrapeptides in the extract. Chemically, the decomposition of the pentapeptide to the unexpected amide version of the proline could be based on two hypothetical mechanisms: (i) base catalysis by the imidazole moiety of the histidine<sup>26</sup> or (ii) an oxidative decomposition via an *N,O*-acetal-like intermediate.<sup>27</sup> These hypotheses and their mechanisms will need to be evaluated in further studies. Interestingly, UHPLC control measurements of the pure pentacitidin samples after NMR spectroscopy always revealed the presence of mixtures of the pentapeptide and its corresponding tetrapeptide analogue with the C-terminal amidated proline. The NMR samples were dissolved in DMSO, dried in vacuo, resolved in MeOH, and measured. For example, within the control measurement of pentacitidin B (3), the falcitidin analogue of family D ( $m/z$  582.3408  $[\text{M}+\text{H}]^+$ ,  $\text{C}_{30}\text{H}_{44}\text{N}_7\text{O}_5$ ) was detected (Figure 2 and Figure S5). However, at the present stage, it also cannot be excluded that falcitidin and pentacitidin analogues are biosynthesized by the same NRPS assembly line and module skipping takes place, which will result in the shorter tetrapeptides.

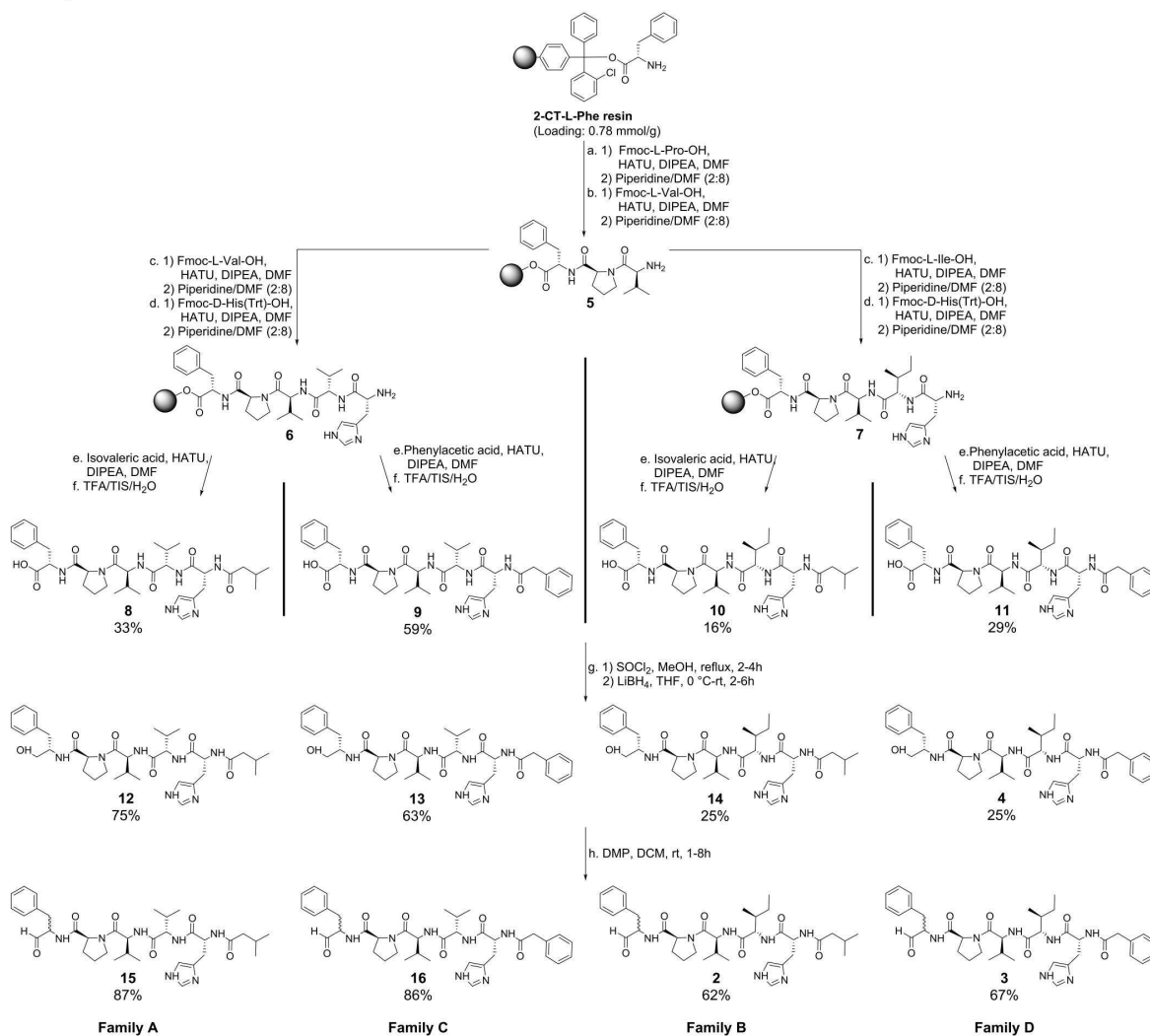
**Total Synthesis of Pentapeptide Falcitidin Analogues.** The scarcity of natural material and difficulties in isolation due to the co-elution of aldehydes and alcohol analogues made it necessary to synthesize the most promising pentapeptide phenylalanine-aldehydes for biological testing. The application of a split-approach solid-phase peptide synthesis (SPPS)<sup>28</sup> followed by functional group interconver-

sion gave access to a variety of falcitidin pentapeptide analogues of families A, B, C, and D (Scheme 1).

Splitting the synthesis after the third amino acid and for the attachment of the two different fatty acids allows the synthesis of all four chosen acid analogues (2, 3, 15, and 16) based on 2-chlorotrytil-L-Phe resin (2-CT-L-Phe). The functionalization of the peptide acid yields the corresponding methyl ester, the alcohol, and the aldehyde. Reduction to the alcohol was done via the methyl ester after direct reduction of the acid could not be achieved. The methyl ester was gained by conversion with thionyl chloride in methanol followed by the reduction with  $\text{LiBH}_4$  in THF. For the oxidation to the aldehyde, Dess–Martin periodinane (DMP) was chosen as a very mild reagent. We found that small traces of methanol as a stabilizer in DCM as the solvent made the reaction take longer, up to 8 h instead of 1 h, and was responsible for incomplete conversion. Epimerization of the stereogenic center of the aldehydes could not be avoided, nor was it necessary to be avoided based on the data for the natural isolated compounds. They were obtained as diastereomers for the phenylalanine moiety of roughly equal proportions based on NMR data (Figures S20–S23 and Tables S12–S15). Overall yields of SPPS for the four different chains varied between 16 and 59%. The reduction yielded the corresponding alcohols over two steps with 25–75% yield, and oxidation to the aldehydes achieved 62–87% yield (Scheme 1).

For reference purposes, falcitidin (1) was synthesized using the same SPPS approach instead of the literature-known liquid phase method.<sup>11</sup> This afforded acid analogue 17 with a yield of 78% and, after amidation, falcitidin (1) with a 41% overall yield (Scheme 2). Additionally, the acid analogue 17 and falcitidin (1) themselves present great functional groups for further derivatization and the introduction of different war heads, like a nitrile or azide group, to increase the potency as a potential inhibitor.<sup>29,30</sup>

**Bioactivity.** Falcitidin (1) was previously reported to display an  $\text{IC}_{50}$  of 6  $\mu\text{M}$  against falcipain-2.<sup>10</sup> Therefore, the falcitidin pentapeptide aldehyde analogues 2, 3, 15, and 16 were tested in a similar in vitro assay against falcipain-2 together with 1 as control. All four aldehydes were active with an  $\text{IC}_{50}$  of 41.5  $\mu\text{M}$  (15) to 23.7  $\mu\text{M}$  (2). However, the originally described activity of 1 could not be reproduced, with an  $\text{IC}_{50}$  of >50  $\mu\text{M}$ . This indicated differences in assay

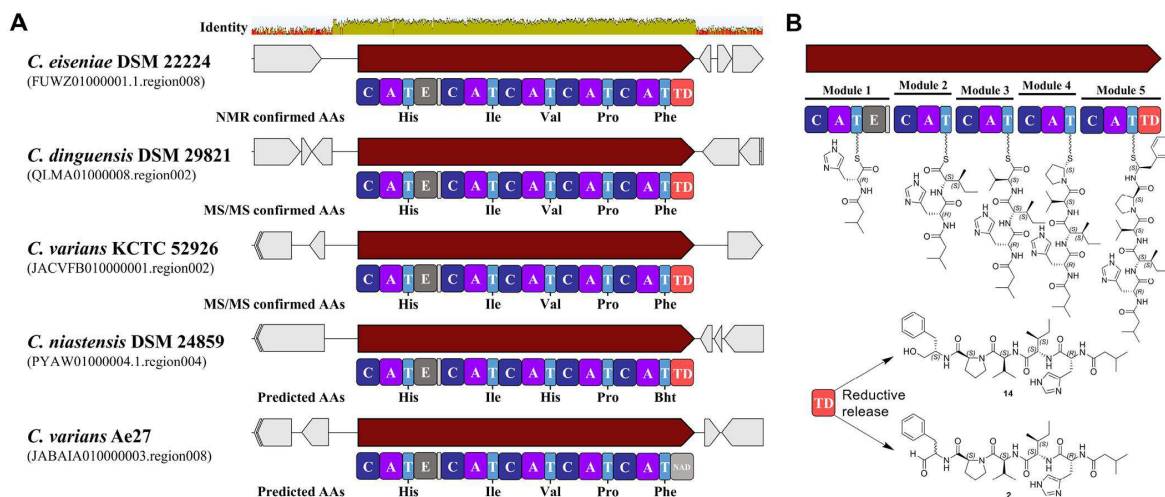
Scheme 1. SPPS Split Approach (a–f) and Functional Group Interconversions (g–h) to Chosen Falcitidin Pentapeptide Analogues of Families A, B, C, and D<sup>a</sup>

<sup>a</sup>Standard SPPS conditions apply, and reactions were carried out at room temperature if not noted otherwise. Reduction of the acid to the alcohol was achieved via the corresponding methyl ester. The aldehyde was obtained by Dess–Martin oxidation; the stereogenic center of phenylalanine could not be retained. Detailed conditions can be found in the [Supporting Information](#).

sensitivity, which made it difficult to put the aldehyde activities in line with the literature. Aldehydes **3** and **15** also displayed activities against the related *P. falciparum* cysteine protease falcipain-3 (66% sequence identity with falcipain-2), with IC<sub>50</sub> values of 45.4 and 42.5 μM, respectively. Inhibition of both falcipains is of importance, since falcipain-3 is able to compensate for the knockout of falcipain-2.<sup>31</sup> A counter-screen against human cysteine proteases cathepsins B and L and sortase A of *Staphylococcus aureus* as a surrogate for unrelated cysteine proteases revealed no activity for pentapeptide aldehydes **2**, **3**, **15**, and **16** as well as the C-terminal acid (**11**) and alcohol (**4**) of **3**, thus demonstrating selectivity over these off-targets. Activities with IC<sub>50</sub> values of 57.7 μM (**2**) to 17.1 μM (**15**) were also observed for the falcipain-homologue cysteine protease of *Trypanosoma brucei rhodesiense* rhodesain.

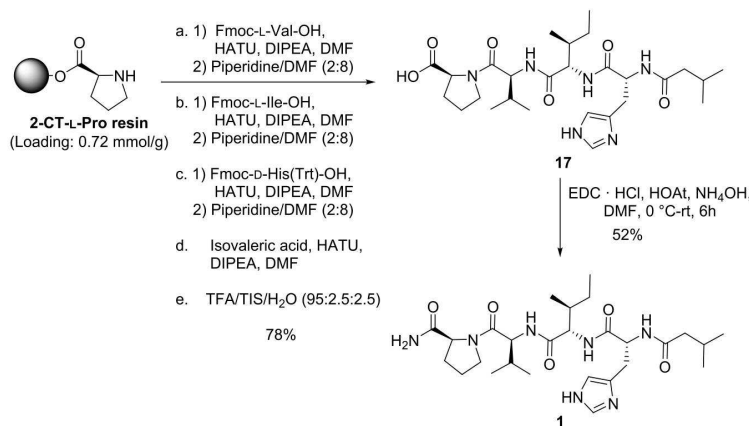
Most interestingly, aldehydes **3**, **15**, and **16** displayed higher activities, with IC<sub>50</sub>s of 3.7 μM (**15**) and 1.5 μM (**16**), against α-chymotrypsin, which was tested as a prototype of human serine proteases, while the serine membrane protease matriptase-2 (TMPRSS6) was not inhibited (Table 4). To further elucidate these observations, molecular docking studies were performed.

**Docking Studies.** Due to the high number of rotatable bonds and the associated degrees of conformational freedom in the molecules under elucidation, docking of the full-length peptides is challenging.<sup>32–34</sup> Hence, truncated tripeptides with an N-terminal acetyl (ace)-cap to avoid a nonpresent charge of full-length inhibitors (**2**, **3**, **4**, **11**, **15**, and **16**) were used for docking studies. First, a conventional noncovalent docking was performed. However, aldehydes are known electrophilic



**Figure 3.** Putative biosynthetic gene clusters for pentacitidin peptides found in various *Chitinophaga* strains. (a) Nucleotide alignment of all five BGCs using MAFFT alignment<sup>67</sup> allowed clear cluster boarder prediction and revealed minor variations in the upstream region (identity). The single NRPS core gene that is responsible for pentacitidin biosynthesis is shown in dark red, and neighboring genes that do not belong to the pentacitidin BGC are shown in gray. (b) Biosynthetic hypothesis for pentacitidins. AA, amino acids; C, condensation domain; A, adenylation domain; T, peptidyl-carrier protein domain; E, epimerization domain; Bht, beta-hydroxy-tyrosine; TD, terminal reductase domain; NAD, NAD\_binding\_4 domain.

#### Scheme 2. Synthesis of Falcitidin (1) via Acid Analogue 17



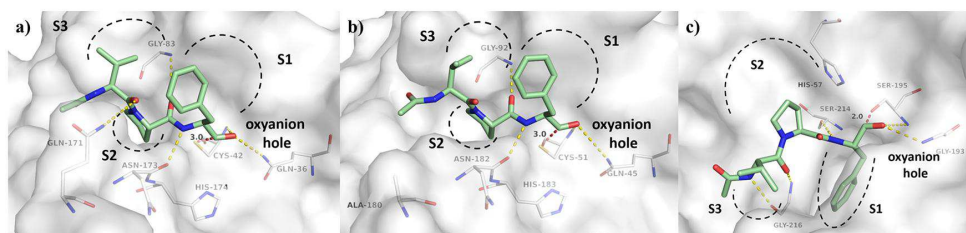
**Table 4.** Overview of Protease Activity Data<sup>42</sup>

no.	peptide sequence							IC <sub>50</sub> (μM)							
								Rhod	falcipain-2	falcipain-3	CatB	CatL	α-CT	SrtA	TMPRSS6
1	iVal	H	I	V	P	-	CONH <sub>2</sub>	-	>50	>50	-	-	-	-	-
17	iVal	H	I	V	P	-	COOH	-	n.d.	n.d.	-	-	-	-	-
11	PA	H	I	V	P	F	COOH	-	n.d.	n.d.	-	-	-	-	-
4	PA	H	I	V	P	F	CH <sub>2</sub> OH	-	n.d.	n.d.	-	-	-	-	-
2	iVal	H	I	V	P	F	CHO	57.7	33.8	45.4	-	-	-	-	-
15	iVal	H	V	V	P	F	CHO	17.1	42.5	-	-	-	2.2	-	-
3	PA	H	I	V	P	F	CHO	25.4	23.7	42.5	-	-	3.7	-	-
16	PA	H	V	V	P	F	CHO	-	38.2	-	-	-	1.5	-	-

<sup>a</sup>n.d., not determined; hyphens denote that the compounds were not active.

warheads and able to form covalent-reversible hemithioacetal adducts with catalytic cysteine residues.<sup>35</sup> Therefore, the predicted poses with special attention to the distance between

the nucleophilic sulfur of the catalytic cysteine residue and the electrophilic carbon of the aldehyde moiety were elucidated, and an additional covalent docking was performed. The



**Figure 4.** Noncovalent docking predicted binding modes of truncated peptidic protease inhibitor ace-Val-Pro-Phe-aldehyde (green carbon atoms) in complex with falcipain-2 (PDB-ID 3BPF) (a), falcipain-3 (PDB-ID 3BPM) (b), and chymotrypsin (PDB-ID 1AFQ) (c). Proteases are depicted as white transparent surfaces; for clarity, only residues forming polar interactions (yellow dashed lines) and catalytic Cys/Ser and His residues are labeled and depicted as lines. Substrate binding sites S1–S3 are schematically indicated. The distance between nucleophilic sulfur (Cys-42 in falcipain-2 and Cys-51 in falcipain-3) or oxygen (Ser-195 in chymotrypsin) of the catalytic center to the electrophilic carbon atom of the aldehyde is depicted as a dashed red line and labeled with its distance in Å.

tripeptides mimic the orientation of the protease substrates by addressing S3–S1 (Figure 4a,b). Based on these results (Table S16), the low inhibitory potency of the C-terminal alcohol (4) and acid (11) moieties as well as that for falcitidin cannot simply be explained by noncovalent interactions as the docking scores are within a similar range or even higher when compared to the aldehydes. However, based on the predicted binding poses, the covalent reaction between aldehydes and catalytic cysteine residues, which might contribute to higher affinity, seems likely for molecules with proteinogenic L-Phe as P1 residue, with a distance of 3.0 Å between the electrophilic carbon atom and the nucleophilic sulfur of Cys for both falcipain-2 and falcipain-3 as well as the aldehyde oxygen coordinated by hydrogen bonds in the oxyanion hole, but rather unlikely for D-Phe (distances of 7.3 and 5.0 Å for falcipain-2 and falcipain-3, respectively). Additionally, only small differences between covalent and noncovalent binding modes were observed, indicating that no larger conformational changes need to take place during or after reaction. While testing against  $\alpha$ -chymotrypsin was conducted to demonstrate selectivity over human off-target serine proteases, the high affinity is reasonable, as inhibitors with aromatic moieties deeply buried in the S1 pocket of the protease were reported previously as well as cleavage preference after Phe.<sup>36–38</sup> Aldehydes are able to form covalent-reversible hemiacetal adducts with serine;<sup>39</sup> moreover, an increased potency is indicated by proximity of the aldehyde to the catalytic Ser-195 residue in the predicted binding modes for both L- and D-Phe at P1 (Figure 4c and Table S16). In conclusion, the higher docking scores for the covalent binding compared to the ones for the noncovalent ones indicate that the literature-known covalent-reversible binding mode in form of the hemithioacetal/hemiacetal adducts takes place for the cysteine/serine residue and the aldehyde moiety of the pentacitidins. Binding of L-Phe seems to be more likely based on the docking results (distance) than D-Phe.

Falcitidin (1) was found to display poor whole-cell activity against chloroquine-sensitive *P. falciparum* strain 3D7 ( $IC_{50} > 10 \mu M$ ), while a structure–activity relationship (SAR) study identified a synthetic trifluoromethyl analogue displaying sub-micromolar  $IC_{50}$  activity.<sup>11</sup> Together with our results, this allows further development of the SAR, with the aim to increase both potency and selectivity with an improved peptidic recognition sequence. However, no general selectivity issue against serine proteases can be expected, as matrilysin-2 was not inhibited by the compounds. Additionally, the docking

studies revealed a rather unlikely distance for the covalent binding of D-Phe-analogues. With all compounds having been tested as mixtures of D- and L-Phe due to the aldehyde's natural epimerization of its stereogenic center,<sup>40,41</sup> further analogues with other warheads preventing epimerization such as a nitrile group,<sup>29,30</sup> for example, should ideally be synthesized and tested in L-configurations only.

## CONCLUSIONS

In this study, the metabolic networking analysis of *Chitinophaga* strains led to the discovery of over 30 N-acyl oligopeptides structurally related to falcitidin, an inhibitor of the antimalarial cysteine protease falcipain-2.<sup>10</sup> They were classified into four peptide families (A–D) based on variations in their amino acid chain length, sequence, acyl residue, and C-terminal functionalization. The analysis of MS/MS spectra revealed each family to contain truncated di- or tripeptides and larger pentapeptides, including molecules with classical C-terminal aldehyde moieties, which are supposed to covalently react with the active-site cysteine and serine residues of proteases.<sup>9</sup> Isolation and structure elucidation of two novel natural pentapeptide aldehydes validated the MS/MS fragmentation pattern analysis. A BGC congruent to the pentapeptide structure was identified in silico. Together with the pentacitidin NMR studies, this indicated falcitidin and its tetrapeptide analogues carrying the C-terminally amidated proline to be degradation products from the described pentacitidins. Total synthesis of all four pentapeptide aldehydes and their carboxylic acid and alcohol derivatives was successfully achieved using a solid-phase peptide synthesis (SPPS) split approach followed by functional C-terminal group interconversion and gave access to the most promising aldehyde analogues, allowing their biological profiling. Selective in vitro activity against parasitic cysteine proteases rhodesain, falcipain-2, and falcipain-3, together with a low-micromolar  $IC_{50}$  inhibition of the serine protease  $\alpha$ -chymotrypsin, was observed. This forms the basis for future studies to develop optimized derivatives with increased potency and selectivity against targeted proteases as well as to elucidate the occurrence of falcitidin analogues carrying the unusual C-terminal amidated proline.

## EXPERIMENTAL SECTION

**General Experimental Procedures.** For all UHPLC-QTOF-UHR-MS and MS/MS measurements, a quadrupole time-of-flight spectrometer (LC-QTOF maXis II, Bruker Daltonics, Bremen, Germany) equipped with an electrospray ionization source in line

with an Agilent 1290 infinity II LC system (Agilent Technologies, CA, United States) was used. C18 RP-UHPLC (ACQUITY UPLC BEH C18 column (130 Å, 1.7  $\mu\text{m}$ ,  $2.1 \times 100$  mm)) was performed at 45 °C with the following linear gradient (A:  $\text{H}_2\text{O}$ , 0.1% HCOOH; B:  $\text{CH}_3\text{CN}$ , 0.1% HCOOH; flow rate: 0.6  $\text{mL min}^{-1}$ ): 0 min: 95% A; 0.30 min: 95% A; 18.00 min: 4.75% A; 18.10 min: 0% A; 22.50 min: 0% A; 22.60 min: 95% A; and 25.00 min: 95% A. A 50–2000  $m/z$  scan range at a 1 Hz scan rate was used to acquire mass spectral data. The injection volume was set to 5  $\mu\text{L}$ . MS/MS experiments were performed at 6 Hz, and the top five most intense ions in each full MS spectrum were targeted for fragmentation by higher-energy collisional dissociation at 25 eV using  $\text{N}_2$  at  $10^{-2}$  mbar. Precursors were excluded after two spectra, released after 0.5 min, and reconsidered if the intensity of an excluded precursor increased by a factor of 1.5 or more. Data were analyzed using the Bruker DataAnalysis 4.0 software package. Specific rotation was determined on a digital polarimeter (P3000, A. Krüss Optronic GmbH, Germany). The standard wavelength was the sodium D-line with 589 nm. Temperature, concentration (g  $100 \text{ mL}^{-1}$ ), and solvent are reported with the determined value.

**NMR Spectroscopy.** The NMR spectra of natural isolated pentacitidins were acquired on a Bruker AVANCE 700 spectrometer (700 MHz for  $^1\text{H}$  and 176 MHz for  $^{13}\text{C}$ ) and a Bruker AVANCE 500 spectrometer (500 MHz for  $^1\text{H}$  and 126 MHz for  $^{13}\text{C}$ ). Both instruments were equipped with a 5 mm TCI cryoprobe. For structure elucidation and the assignment of proton and carbon resonances, 1D- $^1\text{H}$ , 1D- $^{13}\text{C}$ , DQF-COSY, TOCSY (mixing time 80 ms), ROESY (mixing time 150 ms), multiplicity-edited HSQC, and HMBC spectra were acquired. The NMR spectra of synthesized molecules were recorded on an AVANCE III HD 600 spectrometer (600 MHz for  $^1\text{H}$ , 151 MHz for  $^{13}\text{C}$ ) from Bruker Biospin (Bruker Biospin GmbH, Rheinstetten, Germany).  $^1\text{H}$  and  $^{13}\text{C}$  chemical shifts were reported in ppm and referenced to the corresponding residual solvent signal (DMSO- $d_6$ :  $\delta_{\text{C}} = 39.52$  ppm,  $\delta_{\text{H}} = 2.50$  ppm).  $\delta_{\text{C}}$  shifts marked with asterisks were not observed in the  $^{13}\text{C}$  NMR spectrum but were obtained either from HMBC or HSQC data.

**MS/MS Networking.** Molecular networking was performed following established protocols.<sup>13,42</sup> In brief, parent ions are represented by a list of fragment mass/intensity value pairs within the raw data (\*.d files) converted with MSConvert (Proteowizard package32) into plain text files (\*.mgf). These ions are included in the final network once they share at least six fragments (tolerance  $\Delta\text{ppm}$  0.05) with at least one partner ion.<sup>43</sup> Deposited compounds from an in silico fragmented<sup>44</sup> commercial database (Antibase 2017<sup>45</sup>) as well as our in-house reference compound MS/MS database were included in the final network to highlight known NPs. A visualization of the network was constructed in Cytoscape v3.6.0.<sup>46</sup> Edges were drawn between scan nodes with a cosine similarity of  $>0.7$ .<sup>47</sup>

**Strain Fermentation and Purification of Falcitidin Analogues.** A preculture (R2A, 100 mL in a 300 mL Erlenmeyer flask) of *C. eiseniae* DSM 22224 was inoculated from a plate (R2A) and incubated at 28 °C with agitation at 180 rpm for 3 days. A 20 L fermentation in medium 3018 (1 g  $\text{L}^{-1}$  yeast extract and 5 g  $\text{L}^{-1}$  casitone, pH 7.0) inoculated with 2% (v/v) preculture was carried out in separate 2 L flasks filled with 500 mL of culture volume at 28 °C with agitation at 180 rpm for 4 days. The culture broth was subsequently freeze-dried using a Delta 2-24 LSCplus (Martin Christ Gefriertrocknungsanlagen GmbH, Osterode am Harz, Germany). The sample was extracted with a one-time culture volume of  $\text{CH}_3\text{OH}$ , evaporated to dryness using rotary evaporation under reduced pressure, and resuspended in 3 L of 10%  $\text{CH}_3\text{OH}/\text{H}_2\text{O}$ . The extract was loaded onto a XAD16N column (1 L bed volume) and eluted step-wise with 10, 40, 60, 80, and 100%  $\text{CH}_3\text{OH}$  (2 times bed volume each). The 80 and 100% fractions containing falcitidin analogues were pooled, and the sample was adjusted to 200  $\text{mg mL}^{-1}$  in methanol to achieve further separation using preparative C18-RP-HPLC (Synergi 4  $\mu\text{m}$  Fusion-RP 80 Å (250  $\times$  21.2 mm), Phenomenex Inc.) by eluting in a linear gradient increasing from 25 to 75%  $\text{CH}_3\text{CN}$  (+0.1% HCOOH) in 22 min. Fractions of interest were concentrated to 100  $\text{mg mL}^{-1}$  for semipreparative C18-RP-HPLC (Synergi 4  $\mu\text{m}$  Fusion-

RP 80 Å (250  $\times$  10 mm), Phenomenex Inc.) using a linear gradient from 15 to 50%  $\text{CH}_3\text{CN}$  (+0.1% HCOOH) in 29 min. Final purification of the samples of interest (100  $\text{mg mL}^{-1}$ ) was achieved using UHPLC on a ACQUITY UPLC BEH C18 column (130 Å, 1.7  $\mu\text{m}$ , 100  $\times$  2.1 mm, Waters Corporation), eluting in an isocratic gradient of 27.50%  $\text{CH}_3\text{CN}$  (+0.1% HCOOH) in 18 min. In total, isolation yielded 1.5 mg of 2, 1 mg of 3, and 2 mg of a mixture of 3 and 4.

**Pentacitidin A (2).** Amorphous, white powder; see Table 2 for  $^1\text{H}$  and  $^{13}\text{C}$  NMR data; positive HR-ESIMS  $m/z$  680.4131 [ $\text{M}+\text{H}$ ] $^+$ , calculated mass for  $\text{C}_{36}\text{H}_{54}\text{N}_7\text{O}_6$ : 680.4130;  $\Delta = 0.15$  ppm.

**Pentacitidin B (3).** Amorphous, white powder; see Table 3 for  $^1\text{H}$  and  $^{13}\text{C}$  NMR data; positive HR-ESIMS  $m/z$  714.3975 [ $\text{M}+\text{H}$ ] $^+$ , calculated mass for  $\text{C}_{39}\text{H}_{52}\text{N}_7\text{O}_6$ : 714.3974;  $\Delta = 0.14$  ppm.

**Advanced Marfey's Analysis.** The absolute configurations of all amino acids were determined by derivatization using Marfey's reagent.<sup>15</sup> Stock solutions of amino acid standards (50 mM in  $\text{H}_2\text{O}$ ),  $\text{NaHCO}_3$  (1 M in  $\text{H}_2\text{O}$ ), and  $N_{\alpha}$ -(2,4-dinitro-5-fluorophenyl)-L-valinamide (L-FDVA, 70 mM in acetone; Sigma-Aldrich, St. Louis, MO, United States) were prepared. Commercially available and synthesized standards were derivatized using molar ratios of amino acid to L-FDVA and  $\text{NaHCO}_3$  (1/1.4/8). After stirring at 40 °C for 3 h, 1 M HCl was added to obtain a concentration of 170 mM to end the reaction. Samples were subsequently evaporated to dryness and dissolved in DMSO (final concentration 50 mM). L- and D-amino acids were analyzed separately using C18 RP-UHPLC-MS (A:  $\text{H}_2\text{O}$ , 0.1% HCOOH; B:  $\text{CH}_3\text{CN}$ , 0.1% HCOOH; flow rate: 0.6  $\text{mL min}^{-1}$ ). A linear gradient of 15–75% B in 35 min was applied to separate all amino acid standards. Total hydrolysis of the peptide sample containing 3 and 4 was carried out by dissolving 250  $\mu\text{g}$  in 6 M DCl in  $\text{D}_2\text{O}$  and stirring for 7 h at 160 °C. The sample was subsequently evaporated to dryness. Samples were dissolved in 100  $\mu\text{L}$   $\text{H}_2\text{O}$ , derivatized with L-FDVA, and analyzed using the same parameters as described before.

**Fluorometric Assays.** Rhodesain (Rhod),<sup>48–50</sup> *Staphylococcus aureus* sortase A (SrtA),<sup>51</sup> and human matriptase-2 (TMPRSS6)<sup>52</sup> were expressed and purified as published previously, cathepsins B and L (CatB and CatL; human liver, Calbiochem) and  $\alpha$ -chymotrypsin ( $\alpha$ -CT; Sigma-Aldrich, St. Louis, MO, United States) were purchased. For these proteases except SrtA, fluorescence increase upon cleavage of the fluorogenic substrates was monitored without incubation with a TECAN Infinite F200 Pro fluorimeter (excitation  $\lambda = 365$  nm; emission  $\lambda = 460$  nm) in white, flat-bottom 96-well microtiter plates (Greiner Bio-One, Kremsmünster, Austria) with a total volume of 200  $\mu\text{L}$ . Inhibitors and substrates were prepared as stock solutions in DMSO to a final DMSO content of 0.5%. Inhibitors were screened at final concentrations of 20  $\mu\text{M}$  and eventually at 1  $\mu\text{M}$ . IC<sub>50</sub> values were determined for compounds for which an inhibition of  $>50\%$  at a concentration of 20  $\mu\text{M}$  was observed. All assays were performed in technical triplicates and normalized to the activity of DMSO instead of the inhibitors by measuring the increase of fluorescence signal over 10 min. The data were analyzed using GraFit V 5.0.13<sup>53</sup> (Erithacus Software, Horley, UK; <http://erithacus.com/grafit/>).

Cbz-Phe-Arg-AMC (Bachem, Bubendorf BL, Switzerland) was used as a substrate for Rhod, CatB, and CatL. The enzymes were incubated at room temperature in enzyme incubation buffer (Rhod: 50 mM sodium acetate (pH 5.5), 5 mM EDTA, 200 mM NaCl, and 2 mM DTT; CatB/L: 50 mM Tris-HCl (pH 6.5), 5 mM EDTA, 200 mM NaCl, and 2 mM DTT) for 30 min. One hundred and eighty microliters of assay buffer (Rhod: 50 mM sodium acetate (pH 5.5), 5 mM EDTA, 200 mM NaCl, and 0.005% Brij35; CatB/L: 50 mM Tris-HCl (pH 6.5), 5 mM EDTA, 200 mM NaCl, and 0.005% Brij35) were added to the 96-well plates; afterward, the respective enzyme in enzyme incubation buffer (5  $\mu\text{L}$ ; to yield final concentrations for Rhod, 0.01  $\mu\text{M}$ ; CatB, 0.1  $\mu\text{M}$ ; and CatL, 0.2  $\mu\text{M}$ ) was added followed by 10  $\mu\text{L}$  of DMSO (control) or inhibitor solution in DMSO and, finally, the substrate (5  $\mu\text{L}$ ; final concentrations: Rhod, 10  $\mu\text{M}$ ; CatB, 100  $\mu\text{M}$ ; and CatL, 6.5  $\mu\text{M}$ ) was added.<sup>54</sup>

Transpeptidation efficacy of SrtA was examined in vitro as described previously.<sup>51</sup> Briefly, SrtA was diluted in assay buffer (50 mM Tris–HCl (pH 7.50) and 150 mM NaCl) to a final concentration of 1  $\mu$ M. The FRET-pair substrate Abz-LPETG-Dap(Dnp)-OH (Genscript, Piscataway, NJ, United States) and the tetraglycine (Sigma-Aldrich, St. Louis, MO, United States) were added at 25  $\mu$ M and 0.5 mM, respectively.

$\alpha$ -Chymotrypsin (final concentration, 0.4  $\mu$ M) was dissolved in assay buffer containing 50 mM Tris–HCl (pH 8.0), 100 mM NaCl, and 5 mM EDTA. Suc-Leu-Leu-Val-Tyr-AMC (Bachem, Bubendorf BL, Switzerland) was used as a substrate at a final concentration of 52.5  $\mu$ M.<sup>55</sup>

Proteolytic activity of matriptase-2 was measured in a final concentration of 2.5 nM enzyme in 180  $\mu$ L reaction buffer (50 mM Tris–HCl (pH 8.0), 150 mM NaCl, 5 mM CaCl<sub>2</sub>, and 0.01% (v/v) T<sub>x-100</sub>). After addition of the inhibitors, the reaction was initiated without further incubation by adding the substrate Boc-Leu-Arg-Arg-AMC (Bachem, Bubendorf BL, Switzerland,  $K_M = 36.1 \pm 5.8 \mu$ M) to a final concentration of 100  $\mu$ M.

The recombinant enzymes falcipain-2 and falcipain-3 were expressed and purified as previously described.<sup>56</sup> Stock solutions of the compounds, the substrate, and the positive control E-64 (Sigma-Aldrich, St. Louis, MO, United States) were prepared at 10 mM in DMSO. The compounds were incubated in 96-well white flat-bottom plates with 30 nM of recombinant falcipain-2 or falcipain-3 at room temperature in assay buffer (100 mM sodium acetate (pH 5.5) with 5 mM DTT) for 10 min. After incubation, the fluorogenic substrate Z-Leu-Arg-AMC (R&D Systems, Minneapolis, MN, United States) was added at a concentration of 25  $\mu$ M in a final assay volume of 200  $\mu$ L. Fluorescence was monitored with a Varioskan Flash (Thermo Fisher Scientific Inc., Waltham, MA, United States) with excitation at 355 nm and emission at 460 nm. The IC<sub>50</sub> was calculated using GraphPad Prism (GraphPad Software, San Diego, CA, United States) based on a sigmoidal dose–response curve.

**Molecular Docking.** Molecular docking was performed against falcipain-2 (complex with E-64, PDB-ID 3BPF),<sup>57</sup> falcipain-3 (complex with leupeptin, PDB-ID 3BPM),<sup>57</sup> and chymotrypsin (complex with D-leucyl-L-phenylalanyl-p-fluorobenzylamide, PDB-ID 1AFQ).<sup>56</sup> Conventional noncovalent template-based docking was performed with HYBRID v3.3.0.3 (OpenEye Scientific Software, Santa Fe, NM, US; <http://eyesopen.com>).<sup>58,59</sup> The receptor was prepared using the make\_receptor tool version 3.3.0.3 under default settings for potential field generation around the reference ligand for 3BPM and 1AFQ. As the complexed ligand E-64 of 3BPF does not reach toward S1 of falcipain-2, the potential field was generated using leupeptin from the aligned structure of 3BPM (falcipain-3) for which a similar binding behavior for falcipain-2 and falcipain-3 can be expected.<sup>60</sup> Eight hundred ligand conformers per molecule for docking were generated using Omega Pose (OMEGA v3.1.0.3, OpenEye Scientific Software, Santa Fe, NM, US; <http://eyesopen.com>).<sup>61</sup> Covalent docking was performed using MOE 2020.09.<sup>62</sup> Using the covalent reaction of the acetalization from aldehyde and Cys/Ser (for E-64 redocking from epoxide to beta-hydroxythioether), a rigid docking with GB/VI scoring was applied for the initial placement of 50 poses, from which the 10 best-scoring ones were refined using the ASE scoring function. Docking setups were validated by the redocking of crystallographic reference ligands by pose inspection and RMSD calculation (Table S16).

## ■ ASSOCIATED CONTENT

### SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acschembio.1c00861>.

MS/MS-based assignment of peptide sequences; Marfey's analysis; detailed description of all syntheses; <sup>1</sup>H and <sup>13</sup>C spectra and data of all synthesized compounds and the natural isolated ones; docking results (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Authors

Till F. Schäberle – *Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Branch for Bioresources, Giessen 35392, Germany; Institute for Insect Biotechnology, Justus-Liebig-University of Giessen, Giessen 35392, Germany;* [orcid.org/0000-0001-9947-8079](https://orcid.org/0000-0001-9947-8079); Email: [Armin.Bauer@sanofi.com](mailto:Armin.Bauer@sanofi.com)

Armin Bauer – *Sanofi-Aventis Deutschland GmbH, R&D, Frankfurt am Main 65926, Germany;* [orcid.org/0000-0002-2171-1333](https://orcid.org/0000-0002-2171-1333); Phone: +49-641-97-219140; Email: [Till.Schaerberle@ime.fraunhofer.de](mailto:Till.Schaerberle@ime.fraunhofer.de)

### Authors

Stephan Brinkmann – *Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Branch for Bioresources, Giessen 35392, Germany;* [orcid.org/0000-0003-4299-1079](https://orcid.org/0000-0003-4299-1079)

Sandra Semmler – *Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Branch for Bioresources, Giessen 35392, Germany*

Christian Kersten – *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, Mainz 55128, Germany;* [orcid.org/0000-0001-9976-7639](https://orcid.org/0000-0001-9976-7639)

Maria A. Patras – *Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Branch for Bioresources, Giessen 35392, Germany*

Michael Kurz – *Sanofi-Aventis Deutschland GmbH, R&D, Frankfurt am Main 65926, Germany*

Natalie Fuchs – *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, Mainz 55128, Germany;* [orcid.org/0000-0001-6404-676X](https://orcid.org/0000-0001-6404-676X)

Stefan J. Hammerschmidt – *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, Mainz 55128, Germany*

Jenny Legac – *Department of Medicine, University of California, San Francisco, California 94143, United States*

Peter E. Hammann – *Evotec International GmbH, Göttingen 37079, Germany*

Andreas Vilcinskis – *Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Branch for Bioresources, Giessen 35392, Germany; Institute for Insect Biotechnology, Justus-Liebig-University of Giessen, Giessen 35392, Germany*

Philip J. Rosenthal – *Department of Medicine, University of California, San Francisco, California 94143, United States*

Tanja Schirmeister – *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, Mainz 55128, Germany*

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acschembio.1c00861>

### Author Contributions

<sup>v</sup>S.B. and S.S. contributed equally to this work.

### Funding

This work was financially supported by the Hessen State Ministry of Higher Education, Research, and the Arts (HMWK) via the state initiative for the development of scientific and economic excellence for the LOEWE Center for Insect Biotechnology and Bioresources. Sanofi-Aventis Deutschland GmbH and Evotec International GmbH contributed in the framework of the Sanofi-Fraunhofer Natural

Products Center of Excellence/Fraunhofer-Evotec Natural Products Center of Excellence.

### Notes

The authors declare no competing financial interest.

### ACKNOWLEDGMENTS

The authors would like to thank J. Glaeser, S. Schuler, Y. Kleiner, and C. Pöverlein for valuable discussions. We thank C. Hartwig, V. Öhler, and the NMR department of the Justus-Liebig-University Giessen for technical assistance. We further thank OpenEye Scientific for free academic software licenses and T. Steinmetzer and coworkers from the Philipps University of Marburg for the kind gift of matriptase-2 coding plasmid for recombinant protein expression.

### REFERENCES

- (1) Ashley, E. A.; Pyae Phyo, A.; Woodrow, C. J. *Malaria. Lancet* **2018**, *391*, 1608–1621.
- (2) Büscher, P.; Cecchi, G.; Jamonneau, V.; Priotto, G. Human African trypanosomiasis. *Lancet* **2017**, *390*, 2397–2409.
- (3) World Health Organization *World malaria report 2020: 20 years of global progress and challenges*; World Health Organization 2020.
- (4) Haldar, K.; Bhattacharjee, S.; Safeukui, I. Drug resistance in Plasmodium. *Nat. Rev. Microbiol.* **2018**, *16*, 156–170.
- (5) Conrad, M. D.; Rosenthal, P. J. Antimalarial drug resistance in Africa: the calm before the storm? *Lancet Infect. Dis.* **2019**, *19*, e338–e351.
- (6) Otto, H.-H.; Schirmeister, T. Cysteine Proteases and Their Inhibitors. *Chem. Rev.* **1997**, *97*, 133–172.
- (7) Ettari, R.; Previti, S.; Tamborini, L.; Cullia, G.; Grasso, S.; Zappalà, M. The Inhibition of Cysteine Proteases Rhodesain and TbCatB: A Valuable Approach to Treat Human African Trypanosomiasis. *Mini-Rev. Med. Chem.* **2016**, *16*, 1374–1391.
- (8) Rosenthal, P. J. Falcipain cysteine proteases of malaria parasites: An update. *BBA-Proteins Proteom.* **2020**, *1868*, 140362.
- (9) Drag, M.; Salvesen, G. S. Emerging principles in protease-based drug discovery. *Nat. Rev. Drug Discov.* **2010**, *9*, 690–701.
- (10) Somanadhan, B.; Kotturi, S. R.; Yan Leong, C.; Glover, R. P.; Huang, Y.; Flotow, H.; Buss, A. D.; Lear, M. J.; Butler, M. S. Isolation and synthesis of falcitidin, a novel myxobacterial-derived acyltetrapeptide with activity against the malaria target falcipain-2. *J. Antibiot.* **2013**, *66*, 259–264.
- (11) Kotturi, S. R.; Somanadhan, B.; Ch'ng, J.-H.; Tan, K. S.-W.; Butler, M. S.; Lear, M. J. Diverted total synthesis of falcitidin acyl tetrapeptides as new antimalarial leads. *Tetrahedron Lett.* **2014**, *55*, 1949–1951.
- (12) Krug, D.; Müller, R. Secondary metabolomics: the impact of mass spectrometry-based approaches on the discovery and characterization of microbial natural products. *Nat. Prod. Rep.* **2014**, *31*, 768–783.
- (13) Yang, J. Y.; Sanchez, L. M.; Rath, C. M.; Liu, X.; Boudreau, P. D.; Bruns, N.; Glukhov, E.; Wodtke, A.; de Felicio, R.; Fenner, A.; et al. Molecular networking as a dereplication strategy. *J. Nat. Prod.* **2013**, *76*, 1686–1699.
- (14) Brinkmann, S.; Kurz, M.; Patras, M. A.; Hartwig, C.; Marner, M.; Leis, B.; Billion, A.; Kleiner, Y.; Bauer, A.; Toti, L.; Pöverlein, C.; Hammann, P. E.; Vilcinskas, A.; Glaeser, J.; Spohn, M. S.; Schäberle, T. F. Genomic and chemical decryption of the Bacteroidetes phylum for its potential to biosynthesize natural products. *bioRxiv* **2021**, DOI: 10.1101/2021.07.30.454449.
- (15) Bhushan, R.; Brückner, H. Marfey's reagent for chiral amino acid analysis: a review. *Amino Acids* **2004**, *27*, 231–247.
- (16) Blin, K.; Shaw, S.; Steinke, K.; Villebro, R.; Ziemert, N.; Lee, S. Y.; Medema, M. H.; Weber, T. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* **2019**, *47*, W81–W87.
- (17) Katoh, K.; Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780.
- (18) Du, L.; Lou, L. PKS and NRPS release mechanisms. *Nat. Prod. Rep.* **2010**, *27*, 255–278.
- (19) Gaitatzis, N.; Kunze, B.; Müller, R. In vitro reconstitution of the myxochelin biosynthetic machinery of *Stigmatella aurantiaca* Sg a15: Biochemical characterization of a reductive release mechanism from nonribosomal peptide synthetases. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 11136–11141.
- (20) Silakowski, B.; Nordsiek, G.; Kunze, B.; Blöcker, H.; Müller, R. Novel features in a combined polyketide synthase/non-ribosomal peptide synthetase: the myxalimid biosynthetic gene cluster of the myxobacterium *Stigmatella aurantiaca* Sga1511. *Chem. Biol.* **2001**, *8*, 59–69.
- (21) Li, Y.; Weissman, K. J.; Müller, R. Myxochelin Biosynthesis: Direct Evidence for Two- and Four-Electron Reduction of a Carrier Protein-Bound Thioester. *J. Am. Chem. Soc.* **2008**, *130*, 7554–7555.
- (22) Chhabra, A.; Haque, A. S.; Pal, R. K.; Goyal, A.; Rai, R.; Joshi, S.; Panjikar, S.; Pasha, S.; Sankaranarayanan, R.; Gokhale, R. S. Nonprocessive 2 + 2e<sup>-</sup> off-loading reductase domains from mycobacterial nonribosomal peptide synthetases. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 5681–5686.
- (23) Barajas, J. F.; Phelan, R. M.; Schaub, A. J.; Kliever, J. T.; Kelly, P. J.; Jackson, D. R.; Luo, R.; Keasling, J. D.; Tsai, S.-C. Comprehensive Structural and Biochemical Analysis of the Terminal Myxalimid Reductase Domain for the Engineered Production of Primary Alcohols. *Chem. Biol.* **2015**, *22*, 1018–1029.
- (24) Aarts, M. G. M.; Hodge, R.; Kalantidis, K.; Florack, D.; Wilson, Z. A.; Mulligan, B. J.; Stiekema, W. J.; Scott, R.; Pereira, A. The Arabidopsis MALE STERILITY 2 protein shares similarity with reductases in elongation/condensation complexes. *Plant. J.* **1997**, *12*, 615–623.
- (25) Metz, J. I. M.; Lassner, M. Reprogramming of Oil Synthesis in Rapeseed: Industrial Applications. *Ann. N. Y. Acad. Sci.* **1996**, *792*, 82–90.
- (26) Tan, S.; Li, F.; Park, S.; Kim, S. Transition metal-free aerobic oxidative cleavage of the C–N bonds of  $\alpha$ -amino esters. *Org. Chem. Front.* **2019**, *6*, 3854–3858.
- (27) Doering, W. E.; Chanley, J. D. The autoxidation of quinone. *J. Am. Chem. Soc.* **1946**, *68*, 586–588.
- (28) Merrifield, R. B. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *J. Am. Chem. Soc.* **1963**, *85*, 2149–2154.
- (29) Serim, S.; Haedke, U.; Verhelst, S. H. L. Activity-based probes for the study of proteases: recent advances and developments. *ChemMedChem* **2012**, *7*, 1146–1159.
- (30) Gehringer, M.; Laufer, S. A. Emerging and Re-Emerging Warheads for Targeted Covalent Inhibitors: Applications in Medicinal Chemistry and Chemical Biology. *J. Med. Chem.* **2019**, *62*, 5673–5724.
- (31) Sijwali, P. S.; Rosenthal, P. J. Gene disruption confirms a critical role for the cysteine protease falcipain-2 in hemoglobin hydrolysis by *Plasmodium falciparum*. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 4384–4389.
- (32) Andrusier, N.; Mashiach, E.; Nussinov, R.; Wolfson, H. J. Principles of flexible protein-protein docking. *Proteins* **2008**, *73*, 271–289.
- (33) Pons, C.; Grosdidier, S.; Solernou, A.; Pérez-Cano, L.; Fernández-Recio, J. Present and future challenges and limitations in protein-protein docking. *Proteins* **2010**, *78*, 95–108.
- (34) Knaff, P. M.; Kersten, C.; Willbold, R.; Champanhac, C.; Crespy, D.; Wittig, R.; Landfester, K.; Mailänder, V. From In Silico to Experimental Validation: Tailoring Peptide Substrates for a Serine Protease. *Biomacromolecules* **2020**, *21*, 1636–1643.
- (35) Silva, D. G.; Ribeiro, J. F. R.; de Vita, D.; Cianni, L.; Franco, C. H.; Freitas-Junior, L. H.; Moraes, C. B.; Rocha, J. R.; Burtoloso, A. C. B.; Kenny, P. W.; et al. A comparative study of warheads for design of cysteine protease inhibitors. *Bioorg. Med. Chem. Lett.* **2017**, *27*, 5031–5035.

- (36) Kashima, A.; Inoue, Y.; Sugio, S.; Maeda, I.; Nose, T.; Shimohigashi, Y. X-ray crystal structure of a dipeptide-chymotrypsin complex in an inhibitory interaction. *Eur. J. Biochem.* **1998**, *255*, 12–23.
- (37) Stoll, V. S.; Eger, B. T.; Hynes, R. C.; Martichonok, V.; Jones, J. B.; Pai, E. F. Differences in binding modes of enantiomers of 1-acetamido boronic acid based protease inhibitors: crystal structures of gamma-chymotrypsin and subtilisin Carlsberg complexes. *Biochemistry* **1998**, *37*, 451–462.
- (38) Rawlings, N. D.; Waller, M.; Barrett, A. J.; Bateman, A. MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* **2014**, *42*, D503–D509.
- (39) Cho, S.; Dickey, S. W.; Urban, S. Crystal Structures and Inhibition Kinetics Reveal a Two-Stage Catalytic Mechanism with Drug Design Implications for Rhomboid Proteolysis. *Mol. Cell* **2016**, *61*, 329–340.
- (40) Lyons, B.; Kwan, A. H.; Jamie, J.; Truscott, R. J. W. Age-dependent modification of proteins: N-terminal racemization. *FEBS J.* **2013**, *280*, 1980–1990.
- (41) Kajita, R.; Goto, T.; Lee, S. H.; Oe, T. Aldehyde stress-mediated novel modification of proteins: epimerization of the N-terminal amino acid. *Chem. Res. Toxicol.* **2013**, *26*, 1926–1936.
- (42) Allard, P.-M.; Péresse, T.; Bisson, J.; Gindro, K.; Marcourt, L.; van Pham, C.; Roussi, F.; Litaudon, M.; Wolfender, J.-L. Integration of Molecular Networking and In-Silico MS/MS Fragmentation for Natural Products Dereplication. *Anal. Chem.* **2016**, *88*, 3317–3323.
- (43) Riyanti; Marner, M.; Hartwig, C.; Patras, M.; Wodi, S.; Rieuwpassa, F.; Ijong, F.; Balansa, W.; Schäberle, T. Sustainable Low-Volume Analysis of Environmental Samples by Semi-Automated Prioritization of Extracts for Natural Product Research (SeaPEPR). *Mar. Drugs* **2020**, *18*, 649.
- (44) Allen, F.; Greiner, R.; Wishart, D. Competitive fragmentation modeling of ESI-MS/MS spectra for putative metabolite identification. *Metabolomics* **2015**, *11*, 98–110.
- (45) Laatsch, H. *AntiBase: The Natural Compound Identifier*; 5. Edition; Wiley-VCH, 2017.
- (46) Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, *13*, 2498–2504.
- (47) Marner, M.; Patras, M. A.; Kurz, M.; Zubeil, F.; Förster, F.; Schuler, S.; Bauer, A.; Hammann, P.; Vilcinskas, A.; Schäberle, T. F.; Glaeser, J. Molecular Networking-Guided Discovery and Characterization of Stechlisins, a Group of Cyclic Lipopeptides from a *Pseudomonas* sp. *J. Nat. Prod.* **2020**, *83*, 2607–2617.
- (48) Previti, S.; Ettari, R.; Cosconati, S.; Amendola, G.; Chouchene, K.; Wagner, A.; Hellmich, U. A.; Ulrich, K.; Krauth-Siegel, R. L.; Wich, P. R.; et al. Development of Novel Peptide-Based Michael Acceptors Targeting Rhodesain and Falcipain-2 for the Treatment of Neglected Tropical Diseases (NTDs). *J. Med. Chem.* **2017**, *60*, 6911–6923.
- (49) Johé, P.; Jaenicke, E.; Neuweiler, H.; Schirmeister, T.; Kersten, C.; Hellmich, U. A. Structure, interdomain dynamics, and pH-dependent autoactivation of pro-rhodesain, the main lysosomal cysteine protease from African trypanosomes. *J. Biol. Chem.* **2021**, *296*, 100565.
- (50) Jung, S.; Fuchs, N.; Johe, P.; Wagner, A.; Diehl, E.; Yuliani, T.; Zimmer, C.; Barthels, F.; Zimmermann, R. A.; Klein, P.; et al. Fluorovinylsulfones and -Sulfonates as Potent Covalent Reversible Inhibitors of the Trypanosomal Cysteine Protease Rhodesain: Structure-Activity Relationship, Inhibition Mechanism, Metabolism, and In Vivo Studies. *J. Med. Chem.* **2021**, *64*, 12322–12358.
- (51) Barthels, F.; Marincola, G.; Marciniak, T.; Konhäuser, M.; Hammerschmidt, S.; Biermeier, J.; Distler, U.; Wich, P. R.; Tenzer, S.; Schwarzzer, D.; Ziebuhr, W.; Schirmeister, T. Asymmetric Disulfanylbenzamides as Irreversible and Selective Inhibitors of *Staphylococcus aureus* Sortase A. *ChemMedChem* **2020**, *15*, 839–850.
- (52) Steinmetzer, T.; Schweinitz, A.; Stürzebecher, A.; Dönnecke, D.; Uhland, K.; Schuster, O.; Steinmetzer, P.; Müller, F.; Friedrich, R.; Than, M. E.; Bode, W.; Stürzebecher, J. Secondary amides of sulfonlated 3-amidinophenylalanine. New potent and selective inhibitors of matriptase. *J. Med. Chem.* **2006**, *49*, 4116–4126.
- (53) Leatherbarrow, R. J. *GraFit Version 5*; Erithacus Software Ltd.: Horley, U.K., 2006.
- (54) Schirmeister, T.; Kesselring, J.; Jung, S.; Schneider, T. H.; Weickert, A.; Becker, J.; Lee, W.; Bamberger, D.; Wich, P. R.; Distler, U.; et al. Quantum Chemical-Based Protocol for the Rational Design of Covalent Inhibitors. *J. Am. Chem. Soc.* **2016**, *138*, 8332–8335.
- (55) Ludewig, S.; Kossner, M.; Schiller, M.; Baumann, K.; Schirmeister, T. Enzyme kinetics and hit validation in fluorimetric protease assays. *Curr. Top. Med. Chem.* **2010**, *10*, 368–382.
- (56) Sijwali, P. S.; Brinen, L. S.; Rosenthal, P. J. Systematic optimization of expression and refolding of the Plasmodium falciparum cysteine protease falcipain-2. *Protein Expres. Purif.* **2001**, *22*, 128–134.
- (57) Kerr, I. D.; Lee, J. H.; Pandey, K. C.; Harrison, A.; Sajid, M.; Rosenthal, P. J.; Brinen, L. S. Structures of falcipain-2 and falcipain-3 bound to small molecule inhibitors: implications for substrate specificity. *J. Med. Chem.* **2009**, *52*, 852–857.
- (58) McGann, M. FRED pose prediction and virtual screening accuracy. *J. Chem. Inf. Model.* **2011**, *51*, 578–596.
- (59) McGann, M. FRED and HYBRID docking performance on standardized datasets. *J. Comput. Aid Mol. Des.* **2012**, *26*, 897–906.
- (60) Lisk, G.; Pain, M.; Gluzman, I. Y.; Kambhampati, S.; Furuya, T.; Su, X.-Z.; Fay, M. P.; Goldberg, D. E.; Desai, S. A. Changes in the plasmodial surface anion channel reduce leupeptin uptake and can confer drug resistance in Plasmodium falciparum-infected erythrocytes. *Antimicrob. Agents Chemother.* **2008**, *52*, 2346–2354.
- (61) Hawkins, P. C. D.; Skillman, A. G.; Warren, G. L.; Ellingson, B. A.; Stahl, M. T. Conformer generation with OMEGA: algorithm and validation using high quality structures from the Protein Databank and Cambridge Structural Database. *J. Chem. Inf. Model.* **2010**, *50*, 572–584.
- (62) *Molecular Operating Environment (MOE)*; 2020.09; Chemical Computing Group ULC: 1010 Sherbrooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2020. <https://www.chemcomp.com/index.htm>.



## DISCUSSION

---

### **Droplet-based microfluidics approaches for anti-infective NP discovery**

Among various other examples, the increasing scientific challenges of discovering new anti-infective NPs in conjunction with the emerging AMRs (Murray *et al.*, 2022) continuously decrease the already declined therapeutic armamentarium. While common anti-infective compound classes were found applying classical bioactivity-based whole-cell phenotypic screening approaches during the golden age of antibiotics, microbial cultivation-based approaches still hold notable potential in discovering new anti-infective NPs. The lower chances of finding novel scaffolds for drug development must be contrasted with increased throughput. In consideration of optimal growth conditions, including natural occurring growth factors and nutrient supply, microbial dormancy, a bet-hedging strategy used by a variety of organisms to overcome unfavorable environmental conditions (Lennon and Jones, 2011) and the reason uncultivable species exist, could be overcome. Hence, as transition from dormancy to activity is a stochastic process (Epstein, 2013), the chance to cultivate previous inaccessible microbial diversity rises.

In recent years, droplet-based microfluidics has emerged as a method of choice for high-throughput processes. It enables the generation, manipulation and monitoring of individual compartments (droplets) on a picoliter-scale. Realistically, it can deliver million-fold improvements in throughput compared to conventional methods, while consuming just a few microliters of sample. With its key features of compartmentalization, miniaturization, and

parallelization, it has rapidly evolved as a prominent technology in various research fields including microbiology and anti-infective NP discovery (Kaminski *et al.*, 2016; Price and Paegel, 2016).

The results of the droplet-based microfluidic cultivation approach developed and established within this work, which is based on agarose-solidified droplets and presented in Publication I (Chapter I), show the enormous potential of microfluidic-based cultivation approaches. In total, 1071 strains of five different phyla affiliated to 57 genera were brought into culture, including underrepresented proteobacterial genera and a fastidious representative of the Acidobacteria. These results are in alliance with studies using similar droplet-based cultivation approaches (Zengler *et al.*, 2002; Akselband *et al.*, 2006; Ben-Dov *et al.*, 2009; Cao *et al.*, 2017; Terekhov *et al.*, 2017; Mahler *et al.*, 2018; Villa *et al.*, 2020; Watterson *et al.*, 2020; Mahler *et al.*, 2021). In particular, the microfluidics method was established using only commercially available microfluidic components ( $\mu$ Encapsulator 1 system, Dolomite Bio). In contrast to many published studies that have been based on bespoke systems and chips (Cao *et al.*, 2017; Terekhov *et al.*, 2017; Mahler *et al.*, 2018; Watterson *et al.*, 2020; Mahler *et al.*, 2021), this reduces complexity while allowing a maximum of scalability and reproducibility. Though using such a system, the implementation, maintenance and daily operation as well as a rapid method transfer to other laboratories is ensured.

This is also consistent with the analysis of the droplets that has to compete in terms of throughput with droplet generation at a few kilohertz. The applied flow cytometry technique in Chapter I is routinely been used for the analysis of such droplets and standardized in many laboratories (Manome *et al.*, 2001; Zengler *et al.*, 2002; Akselband *et al.*, 2006; Terekhov *et al.*, 2017). It allows to analyze and process droplets at kilohertz based on physical and chemical characteristics. However, as droplets can be of various physical properties and research questions need to be addressed with any readout possible, more variable, flexible and modular systems evolved. Many detection, analysis and sorting systems are custom made and include one or more on-chip optical (Zang *et al.*, 2013; Cao *et al.*, 2017; Mahler *et al.*, 2018; Watterson *et al.*, 2020; Mahler *et al.*, 2021), optofluidics (Sciambi and Abate, 2015; Mahler *et al.*, 2018; Tovar

*et al.*, 2019; Hengoju *et al.*, 2020b; Mahler *et al.*, 2021) and mass spectrometry-based solutions (Mahler *et al.*, 2018; Wink *et al.*, 2018). Whole fluorescence-activated droplet sorting devices were also constructed (Baret *et al.*, 2009). Nevertheless, on-chip-based solutions are technically more complex and custom made. Special equipment and skilled operators are required. Therefore, their establishment is more expensive and time-consuming compared to the implemented and standardized approach presented in Chapter I, Publication I.

On the other hand, these technical developments indicate that up- and downstream processes in anti-infective research, *i.e.*, cultivation, phylogenetic and chemical characterization of individual strains, and activity screening will in future be miniaturized as standard down to microfluidic level. This results in more complex droplet-applications and manipulation as well as technical requirements. A few examples are already published. Mahler *et al.* created a modular workflow to cultivate and screen environmental microorganisms in droplets for bioactivity using several custom made microfluidic chips. First, droplets were generated with single environmental cells, followed by incubation and reinjected into a microfluidic chip allowing picoinjection. This was used to inject fluorescent-tagged reporter strains into these droplets for screening purposes. Finally, droplets were sorted via dielectrophoresis using an on-chip droplet sorting structure based on fluorescence readout (Mahler *et al.*, 2018; Mahler *et al.*, 2021). Another more complex droplet-application allows to perform droplet-based, more destructive anti-infective research processes such as phylogenetic characterization based on sequencing. Droplet splitting could be used to create two daughter droplets, one for a destructive assay and the other to recover the droplet content (Link *et al.*, 2006; Ng *et al.*, 2015).

In order to accompany the implications outlined above, further complex and multilayer microfluidic approaches need to be developed in future. The microfluidic method published as Publication II (Chapter I) is an example of such a novel multi compartment assay. The assay builds on the same agarose-solidified agarose droplets and is performed at the same common microfluidic setup used in Publication I. Therefore, it is able to be standardized and implemented easily in other laboratories. The agarose droplets are washed applying a Nykoden<sup>®</sup> density gravity gradient. Followed by the addition of a second agarose layer to the

washed droplets using the same microfluidics system. This results in agarose-solidified double emulsions not separated by a diffusion barrier (surfactant stabilized oil). This assay is the first of its kind with a washing step in between.

First application examples are shown in Publication II with one of them been a screening assay with fluorescent-tagged indicator strains. First, microorganisms could be cultivated within the droplets and an indicator strain could be added for any kind of readout in the second step. However, for some applications, such as an antimicrobial screening assay, the washing step might be a disadvantage. Any soluble molecules will be washed out during the washing step and need to be produced again after adding the second agarose layer, including any possibly produced compound with bioactivity. Therefore, with a fast readout time, the detection of slowly produced or anti-infective compounds with week potency could be very difficult. However, facing the general miniaturization down to microfluidic-based applications, a washing step is essential for various assays (e.g., phage display). For such a purpose, the method developed in this work could be easily applied and opens up new possibilities for researchers trying to miniaturize their assays. The multilayer microfluidic approach published as Publication II could also be applied in the field of microbial co-cultivation. In most cases, microorganisms co-exist within complex microbial communities while experiencing competition and/or antagonism. Co-cultivations can simulate such interactions, leading to favorable microbial growth under laboratory conditions, and thus could likely help to cultivate species that have not been cultivated before (Kaeberlein *et al.*, 2002; Schink, 2002; Park *et al.*, 2011; Marmann *et al.*, 2014). Additionally, the microfluidic method could be adapted by adding a liquid layer rather than a second agarose layer on the washed droplets. It could be used to encapsulate environmental cells in a solid as well as liquid phase. During FACS sorting, this results in the loss of the liquid layer containing environmental microorganisms, and thus recovery of the cells grown in the solidified agarose droplet. This would allow high-throughput co-cultivation in droplets while retrieving axenic cultures. Other applications need to be tested next.

In this perspective, the work that was conducted in Chapter I addresses the recent rise of droplet-based microfluidic methods and applications in the field of anti-infective research. Given is the change to cultivate and retrieve new as well as fastidious species due to the adaptable cultivation conditions in droplets and the single cell confinement, together with a high sampling depth. The enormous potential of microfluidic-based cultivation approaches have been proven and will help to standardize such assays in future. Furthermore, the developed novel double-emulsion method has the potential to help miniaturize further assays and research fields.

### **Bacteroidetes – A yet underexplored source of natural product research**

Microorganisms are a valuable source of NPs (Newman and Cragg, 2020). Over 300.000 bacterial genomes are publically available (as of Nov. 2020) and provide a rich genomic diversity (Mukherjee *et al.*, 2021). As the sequencing capacity increases, the cost decreases – a unique chance for the research community to tap into the genomic era and revolutionize NP discovery (Mardis, 2017; Kalkreuter *et al.*, 2020). Even though the genetic potential for the biosynthesis of further promising NPs is still high for rather classical producers such as Actinobacteria, Firmicutes, and Myxobacteria, a constant rediscovery of known molecules challenges classical NP research (Ziemert *et al.*, 2016). The exploration of yet underexplored phylogenetic spaces of the bacterial kingdom can be an alternative. Here, the large potential that comes with the public genomic data can support the search for such alternatives as it helps to predict their potential to produce valuable secondary/specialized metabolites (Borsetto *et al.*, 2019).

In this context, the genomic study performed in Chapter II, Publication III explored and illuminated the Bacteroidetes phylum as such an alternative for the discovery of novel NPs. Earlier on, it has been described that the Bacteroidetes phylum harbors distinct adenylation and ketosynthase domains. This was correlated with genes coding for non-ribosomal peptide synthetases (NRPSs) and polyketide synthases (PKSs). In turn, detecting these genes in the genomes of Bacteroidetes strains, the assumption of a high biosynthetic potential to produce polyketides and non-ribosomal peptides was raised (Borsetto *et al.*, 2019). The within this work

presented study represents the first systematic and comprehensive survey rating the biosynthetic potential of all six different Bacteroidetes classes using the bioinformatics tools antiSMASH (Blin *et al.*, 2021) and BiG-SCAPE (Navarro-Muñoz *et al.*, 2020). It revealed the accumulation of NP production capabilities in only a few taxonomic hotspots, rather than a general ability of the entire phylum-members, to produce secondary/specialized metabolites. Additionally, evidence was made that the encoded chemical space of the phylum differs from that of other bacteria. In particular, the here performed genomic study revealed the *Chitinophaga* as the most promising genus in terms of secondary/specialized metabolite production capabilities. The most talented *Chitinophaga* species even competed with the BGC loads of classical producers such as Actinobacteria (Baltz, 2017).

However, as genomic diversity does not always correlates into chemical diversity (Kalkreuter *et al.*, 2020), the translation of the biosynthetic potential into chemical novelty needs to be verified. In light of this, a representative metabolomics study of the *Chitinophaga* was part of the research in Publication III. The identified chemical space comprises more than 1,000 unknown candidates that could not be associated to microbial NPs known today. Although comparability of such extensive approaches is challenging due to different dataset sizes, composition as well as technical devices and settings, this dataset competed with its uniqueness and value with other metabolomics studies of well-known NP producer taxa (Hoffmann *et al.*, 2018). One striking difference however, is the lack of dereplicated known metabolites. It can be explained by their underrepresentation in all reference databases utilized in this study for data categorization. This is representative of the fact that investigations of the genus *Chitinophaga* and the Bacteroidetes phylum itself have only just begun.

From this perspective, and among with the rather low number of literature known secondary/specialized metabolites not reflecting the phylum's potential and reviewed as part of this work (Introduction), the isolated and characterized compounds reported in Chapter II, Publication III, IV and V represent individual examples of bioactive metabolites from the untapped chemical space of the Bacteroidetes phylum. The datasets generated with the omics-technologies used in this work, stand alone or applied in sequel, were proven to make a strong

impact on each isolation project. Especially the here presented results reveal the diverse repertoire of bioactive NPs produced by members of the genus *Chitinophaga*. The chitinopeptins A-D are the first cyclic lipodepsipeptides isolated from the genus *Chitinophaga* and able to coordinate iron (Publication III). The analysis of the genomics data generated within this work revealed putative NRPS-BGCs congruent with its structures and even two more close related BGCs were found in genomes of other members of this genus. Furthermore, based on the metabolomics data, the novel lipoamino acids described in Publication IV were identified as 'core' metabolites of such strains. In parallel, the same data revealed the presence of the tetrapeptide falcitidin, an inhibitor of the antimalarial target falcipain-2 (Somanadhan *et al.*, 2013), in several *Chitinophaga* extracts. Published in Publication V, an intensive follow up study predicts this tetrapeptide to be a degradation product of the pentacitidins, novel linear pentapeptides carrying aldehyde warheads. The pentacitidins act as even better inhibitors of parasitic cysteine proteases. The NRPS-BGC of the pentacitins was also putatively annotated in at least five *Chitinophaga* genomes.

Based on these insights, the results of Chapter II comprehensively claim the NP potential of the entire Bacteroidetes phylum, especially proven for the *Chitinophaga* genus. This work has been proven to be the foundation for all upcoming NP research on this phylum. Applied on this underexplored NP source, omics-technologies revealed a large untapped chemical space with the potential to discover further promising anti-infective secondary/specialized metabolites produced by members of this phylum.

# 4

## CONCLUSION and OUTLOOK

---

Today, anti-infective therapies based on small molecules are applied to fight many diseases and are likely to become even more substantial in future. Novel methods and technical approaches have significantly advanced the field of NP discovery. The field is witnessing a renaissance relying on increasing multidisciplinary collaborations between microbiologists, data-driven computational as well as synthetic biologists, chemists, and physicists to successfully address today's challenges in the discovery and development of novel bioactive small molecules. In this perspective, the here presented work made use of such synergistic effects and expertise. The work that was conducted in Chapter I has added technical-driven, microfluidic-based high-throughput approaches to an existing NP discovery pipeline. Next steps are further improvements of the methods developed herein, especially the double emulsion one. Its potential use as a cultivation and screening assay is impeded by the washout of any bioactive molecules that might be produced at the cultivation step. After adding the second agarose layer to the agarose-solidified droplets, such molecules need to be reproduced while the indicator strain also grows within the double emulsion droplet. Therefore, preliminary experiments in this direction have already shown that only a combination of highly potent molecules and very weak or growth delayed reporter strains leads to a detectable inhibition and thus to a usable readout. A solution that needs to be tested next, might be the use of autotrophic or inducible reporter strains.

The results presented in Chapter II are based on large datasets processed with computational tools. They are representative of the great progress in the genomic era of NP discovery by



comprehensively breaking down the NP potential of the entire Bacteroidetes phylum demonstrated specifically for one of the phylum's determined hotspots in NP production capabilities, the *Chitinophaga* genus. The next steps would be to scale the workflow used here to further clades of the Bacteroidetes phylum such as the genus *Pedobacter* or *Cytophaga*. In combination with novel metabologenomic approaches that directly correlate the detected metabolites with their putative BGCs, this would help shed light on the entire treasure trove of NPs from the still underexplored Bacteroidetes phylum. Furthermore, it would prove that the workflow is applicable to other underexplored phyla and even beyond.

In the next decades, technical and scientific improvements will guide scientists to discover further classes of anti-infectives from the almost boundless chemical diversity of NPs hidden in the manifold branches of the tree of life. The future of NP research will continue to rely on classical synthetic biology and cultivation approaches, prioritized and guided by harvesting large and diverse datasets generated with high-throughput approaches.

## REFERENCES

---

- Achtman, M., and Wagner, M. (2008) Microbial diversity and the genetic nature of microbial species. *Nat Rev Microbiol* **6** (6): 431–440.
- Adamek, M., Alanjary, M., and Ziemert, N. (2019) Applied evolution: phylogeny-based approaches in natural products research. *Nat. Prod. Rep.* **36** (9): 1295–1312.
- Akselband, Y., Cabral, C., Castor, T.P., Chikarmane, H.M., and McGrath, P. (2006) Enrichment of slow-growing marine microorganisms from mixed cultures using gel microdrop (GMD) growth assay and fluorescence-activated cell sorting. *Journal of Experimental Marine Biology and Ecology* **329** (2): 196–205.
- Atanasov, A.G., Zotchev, S.B., Dirsch, V.M., and Supuran, C.T. (2021) Natural products in drug discovery: advances and opportunities. *Nat Rev Drug Discov* **20** (3): 200–216.
- Baltz, R.H. (2017) Gifted microbes for genome mining and natural product discovery. *J Ind Microbiol Biotechnol* **44** (4-5): 573–588.
- Baret, J.-C., Miller, O.J., Taly, V., Ryckelynck, M., El-Harrak, A., Frenz, L., et al. (2009) Fluorescence-activated droplet sorting (FADS): efficient microfluidic cell sorting based on enzymatic activity. *Lab Chip* **9** (13): 1850–1858.
- Belknap, K.C., Park, C.J., Barth, B.M., and Andam, C.P. (2020) Genome mining of biosynthetic and chemotherapeutic gene clusters in *Streptomyces* bacteria. *Sci Rep* **10** (1): 2003.
- Ben-Dov, E., Kramarsky-Winter, E., and Kushmaro, A. (2009) An in situ method for cultivating microorganisms using a double encapsulation technique. *FEMS Microbiology Ecology* **68** (3): 363–371.
- Bérdy, J. (2012) Thoughts and facts about antibiotics: where we are now and where we are heading. *J Antibiot (Tokyo)* **65** (8): 385–395.
- Blin, K., Shaw, S., Kloosterman, A.M., Charlop-Powers, Z., van Wezel, G.P., Medema, M.H., and Weber, T. (2021) antiSMASH 6.0: improving cluster detection and comparison capabilities. *Nucleic Acids Res* **49** (W1): W29–W35.
- Boedicker, J.Q., Vincent, M.E., and Ismagilov, R.F. (2009) Microfluidic confinement of single cells of bacteria in small volumes initiates high-density behavior of quorum sensing and growth and reveals its variability. *Angew Chem Int Ed Engl* **48** (32): 5908–5911.
- Boitard, L., Cottinet, D., Bremond, N., Baudry, J., and Bibette, J. (2015) Growing microbes in millifluidic droplets. *Eng. Life Sci.* **15** (3): 318–326.
- Bollmann, A., Lewis, K., and Epstein, S.S. (2007) Incubation of environmental samples in a diffusion chamber increases the diversity of recovered isolates. *Appl Environ Microbiol* **73** (20): 6386–6390.
- Borsetto, C., Amos, G.C.A., da Rocha, U.N., Mitchell, A.L., Finn, R.D., Laidi, R.F., et al. (2019) Microbial community drivers of PK/NRP gene diversity in selected global soils. *Microbiome* **7** (1): 78.
- Brady, S.F., Simmons, L., Kim, J.H., and Schmidt, E.W. (2009) Metagenomic approaches to natural products from free-living and symbiotic organisms. *Nat. Prod. Rep.* **26** (11): 1488–1503.
- Cao, J., Hafermann, L., and Köhler, J.M. (2017) Stochastically reduced communities-Microfluidic compartments as model and investigation tool for soil microorganism growth in structured spaces. *Eng. Life Sci.* **17** (7): 792–800.
- Chain, E., Florey, H.W., Gardner, A.D., Heatley, N.G., Jennings, M.A., Orr-Ewing, J., and Sanders, A.G. (1940) PENICILLIN AS A CHEMOTHERAPEUTIC AGENT. *The Lancet* **236** (6104): 226–228.
- Challinor, V.L., and Bode, H.B. (2015) Bioactive natural products from novel microbial sources. *Ann. N.Y. Acad. Sci.* **1354** (1): 82–97.

- Cimermancic, P., Medema, M.H., Claesen, J., Kurita, K., Wieland Brown, L.C., Mavrommatis, K., *et al.* (2014) Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. *Cell* **158** (2): 412–421.
- Davies, J. (2013) Specialized microbial metabolites: functions and origins. *J Antibiot (Tokyo)* **66** (7): 361–364.
- Dejong, C.A., Chen, G.M., Li, H., Johnston, C.W., Edwards, M.R., Rees, P.N., *et al.* (2016) Polyketide and nonribosomal peptide retro-biosynthesis and global gene cluster matching. *Nat Chem Biol* **12** (12): 1007–1014.
- Demain, A.L., and Fang, A. (2000) The natural functions of secondary metabolites. *Adv Biochem Eng Biotechnol* **69**: 1–39.
- Domagk, G. (1935) Chemotherapie der bakteriellen Infektionen. *Angew. Chem.* **48** (42): 657–667.
- Epstein, S.S. (2013) The phenomenon of microbial uncultivability. *Curr Opin Microbiol* **16** (5): 636–642.
- Fair, R.J., and Tor, Y. (2014) Antibiotics and bacterial resistance in the 21st century. *Perspect Medicin Chem* **6**: 25–64.
- Firn, R.D., and Jones, C.G. (2003) Natural products—a simple model to explain chemical diversity. *Nat. Prod. Rep.* **20** (4): 382–391.
- Harvey, A.L., Edrada-Ebel, R., and Quinn, R.J. (2015) The re-emergence of natural products for drug discovery in the genomics era. *Nat Rev Drug Discov* **14** (2): 111–129.
- Hengoju, S., Tovar, M., Man, D.K.W., Buchheim, S., and Rosenbaum, M.A. (2020a) Droplet Microfluidics for Microbial Biotechnology. *Adv Biochem Eng Biotechnol.*
- Hengoju, S., Wohlfeil, S., Munser, A.S., Boehme, S., Beckert, E., Shvydkiv, O., *et al.* (2020b) Optofluidic detection setup for multi-parametric analysis of microbiological samples in droplets. *Biomicrofluidics* **14** (2): 24109.
- Hoffmann, T., Krug, D., Bozkurt, N., Duddela, S., Jansen, R., Garcia, R., *et al.* (2018) Correlating chemical diversity with taxonomic distance for discovery of natural products in myxobacteria. *Nat Commun* **9** (1): 803.
- Hong, J. (2011) Role of natural product diversity in chemical biology. *Current Opinion in Chemical Biology* **15** (3): 350–354.
- Ingham, C.J., Sprenkels, A., Bomer, J., Molenaar, D., van den Berg, A., van Hylckama Vlieg, J.E.T., and Vos, W.M. de (2007) The micro-Petri dish, a million-well growth chip for the culture and high-throughput screening of microorganisms. *Proc Natl Acad Sci USA* **104** (46): 18217–18222.
- Ishii, S., Tago, K., and Senoo, K. (2010) Single-cell analysis and isolation for microbiology and biotechnology: methods and applications. *Appl Microbiol Biotechnol* **86** (5): 1281–1292.
- Jiang, C.-Y., Dong, L., Zhao, J.-K., Hu, X., Shen, C., Qiao, Y., *et al.* (2016) High-Throughput Single-Cell Cultivation on Microfluidic Streak Plates. *Appl Environ Microbiol* **82** (7): 2210–2218.
- Kaeberlein, T., Lewis, K., and Epstein, S.S. (2002) Isolating "uncultivable" microorganisms in pure culture in a simulated natural environment. *Science* **296** (5570): 1127–1129.
- Kalkreuter, E., Pan, G., Cepeda, A.J., and Shen, B. (2020) Targeting Bacterial Genomes for Natural Product Discovery. *Trends Pharmacol Sci* **41** (1): 13–26.
- Kaminski, T.S., Scheler, O., and Garstecki, P. (2016) Droplet microfluidics for microbiology: techniques, applications and challenges. *Lab Chip* **16** (12): 2168–2187.
- Keller, M., and Zengler, K. (2004) Tapping into microbial diversity. *Nat Rev Microbiol* **2** (2): 141–150.
- Lennon, J.T., and Jones, S.E. (2011) Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nat Rev Microbiol* **9** (2): 119–130.
- Lewis, K. (2020) The Science of Antibiotic Discovery. *Cell* **181** (1): 29–45.
- Li, J.W.-H., and Vederas, J.C. (2009) Drug discovery and natural products: end of an era or an endless frontier? *Science* **325** (5937): 161–165.

- Ling, L.L., Schneider, T., Peoples, A.J., Spoering, A.L., Engels, I., Conlon, B.P., *et al.* (2015) A new antibiotic kills pathogens without detectable resistance. *Nature* **517** (7535): 455–459.
- Link, D.R., Grasland-Mongrain, E., Duri, A., Sarrazin, F., Cheng, Z., Cristobal, G., *et al.* (2006) Electric control of droplets in microfluidic devices. *Angew Chem Int Ed Engl* **45** (16): 2556–2560.
- Ma, L., Datta, S.S., Karymov, M.A., Pan, Q., Begolo, S., and Ismagilov, R.F. (2014a) Individually addressable arrays of replica microbial cultures enabled by splitting SlipChips. *Integr Biol (Camb)* **6** (8): 796–805.
- Ma, L., Kim, J., Hatzenpichler, R., Karymov, M.A., Hubert, N., Hanan, I.M., *et al.* (2014b) Gene-targeted microfluidic cultivation validated by isolation of a gut bacterium listed in Human Microbiome Project's Most Wanted taxa. *Proc Natl Acad Sci USA* **111** (27): 9768–9773.
- Mahler, L., Niehs, S.P., Martin, K., Weber, T., Scherlach, K., Hertweck, C., *et al.* (2021) Highly parallelized droplet cultivation and prioritization of antibiotic producers from natural microbial communities. *Elife* **10**.
- Mahler, L., Wink, K., Beulig, R.J., Scherlach, K., Tovar, M., Zang, E., *et al.* (2018) Detection of antibiotics synthesized in microfluidic picolitre-droplets by various actinobacteria. *Sci Rep* **8** (1): 13087.
- Manome, A., Zhang, H., Tani, Y., Katsuragi, T., Kurane, R., and Tsuchida, T. (2001) Application of gel microdroplet and flow cytometry techniques to selective enrichment of non-growing bacterial cells. *FEMS Microbiol Lett* **197** (1): 29–33.
- Mardis, E.R. (2017) DNA sequencing technologies: 2006–2016. *Nat Protoc* **12** (2): 213–218.
- Marmann, A., Aly, A.H., Lin, W., Wang, B., and Proksch, P. (2014) Co-cultivation—a powerful emerging tool for enhancing the chemical diversity of microorganisms. *Mar Drugs* **12** (2): 1043–1065.
- Martin, K., Henkel, T., Baier, V., Grodrian, A., Schön, T., Roth, M., *et al.* (2003) Generation of larger numbers of separated microbial populations by cultivation in segmented-flow microdevices. *Lab Chip* **3** (3): 202–207.
- Medema, M.H., Cimerancic, P., Sali, A., Takano, E., and Fischbach, M.A. (2014) A systematic computational analysis of biosynthetic gene cluster evolution: lessons for engineering biosynthesis. *PLoS Comput Biol* **10** (12): e1004016.
- Mendes, R., Garbeva, P., and Raaijmakers, J.M. (2013) The rhizosphere microbiome: significance of plant beneficial, plant pathogenic, and human pathogenic microorganisms. *FEMS Microbiology Reviews* **37** (5): 634–663.
- Monciardini, P., Iorio, M., Maffioli, S., Sosio, M., and Donadio, S. (2014) Discovering new bioactive molecules from microbial sources. *Microb. Biotechnol.* **7** (3): 209–220.
- Mukherjee, S., Stamatis, D., Bertsch, J., Ovchinnikova, G., Sundaramurthi, J.C., Lee, J., *et al.* (2021) Genomes OnLine Database (GOLD) v.8: overview and updates. *Nucleic Acids Res* **49** (D1): D723–D733.
- Murray, C.J.L., Ikuta, K.S., Sharara, F., Swetschinski, L., Robles Aguilar, G., Gray, A., *et al.* (2022) Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet*.
- Navarro-Muñoz, J.C., Selem-Mojica, N., Mullowney, M.W., Kautsar, S.A., Tryon, J.H., Parkinson, E.I., *et al.* (2020) A computational framework to explore large-scale biosynthetic diversity. *Nat Chem Biol* **16** (1): 60–68.
- Newman, D.J., and Cragg, G.M. (2020) Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. *J. Nat. Prod.* **83** (3): 770–803.
- Ng, A.H.C., Li, B.B., Chamberlain, M.D., and Wheeler, A.R. (2015) Digital Microfluidic Cell Culture. *Annu Rev Biomed Eng* **17**: 91–112.
- Nge, P.N., Rogers, C.I., and Woolley, A.T. (2013) Advances in microfluidic materials, functions, integration, and applications. *Chem. Rev.* **113** (4): 2550–2583.

- Nichols, D., Cahoon, N., Trakhtenberg, E.M., Pham, L., Mehta, A., Belanger, A., *et al.* (2010) Use of icip for high-throughput in situ cultivation of "uncultivable" microbial species. *Appl Environ Microbiol* **76** (8): 2445–2450.
- Park, J., Kerner, A., Burns, M.A., and Lin, X.N. (2011) Microdroplet-enabled highly parallel co-cultivation of microbial communities. *PLoS One* **6** (2): e17019.
- Piel, J. (2011) Approaches to capturing and designing biologically active small molecules produced by uncultured microbes. *Annu. Rev. Microbiol.* **65** (1): 431–453.
- Price, A.K., and Paegel, B.M. (2016) Discovery in Droplets. *Anal Chem* **88** (1): 339–353.
- Schink, B. (2002) Synergistic interactions in the microbial world. *Antonie Van Leeuwenhoek* **81** (1-4): 257–261.
- Sciambi, A., and Abate, A.R. (2015) Accurate microfluidic sorting of droplets at 30 kHz. *Lab Chip* **15** (1): 47–51.
- Seemann, R., Brinkmann, M., Pfohl, T., and Herminghaus, S. (2012) Droplet based microfluidics. *Rep. Prog. Phys.* **75** (1): 16601.
- Somanadhan, B., Kotturi, S.R., Yan Leong, C., Glover, R.P., Huang, Y., Flotow, H., *et al.* (2013) Isolation and synthesis of falcitidin, a novel myxobacterial-derived acyltetrapeptide with activity against the malaria target falcipain-2. *J Antibiot (Tokyo)* **66** (5): 259–264.
- Stewart, E.J. (2012) Growing unculturable bacteria. *J Bacteriol* **194** (16): 4151–4160.
- Talebi Bezmin Abadi, A., Rizvanov, A.A., Haertlé, T., and Blatt, N.L. (2019) World Health Organization Report: Current Crisis of Antibiotic Resistance. *BioNanoSci.* **9** (4): 778–788.
- Tandogan, N., Abadian, P.N., Epstein, S., Aoi, Y., and Goluch, E.D. (2014) Isolation of microorganisms using sub-micrometer constrictions. *PLoS One* **9** (6): e101429.
- Teh, S.-Y., Lin, R., Hung, L.-H., and Lee, A.P. (2008) Droplet microfluidics. *Lab Chip* **8** (2): 198–220.
- Terekhov, S.S., Smirnov, I.V., Stepanova, A.V., Bobik, T.V., Mokrushina, Y.A., Ponomarenko, N.A., *et al.* (2017) Microfluidic droplet platform for ultrahigh-throughput single-cell screening of biodiversity. *Proc Natl Acad Sci USA* **114** (10): 2550–2555.
- Theberge, A.B., Courtois, F., Schaerli, Y., Fischlechner, M., Abell, C., Hollfelder, F., and Huck, W.T.S. (2010) Microdroplets in microfluidics: an evolving platform for discoveries in chemistry and biology. *Angew Chem Int Ed Engl* **49** (34): 5846–5868.
- Thomas, F., Hehemann, J.-H., Rebuffet, E., Czjzek, M., and Michel, G. (2011) Environmental and gut bacteroidetes: the food connection. *Front Microbiol* **2**: 93.
- Tovar, M., Hengoju, S., Weber, T., Mahler, L., Choudhary, M., Becker, T., and Roth, M. (2019) One Sensor for Multiple Colors: Fluorescence Analysis of Microdroplets in Microbiological Screenings by Frequency-Division Multiplexing. *Anal Chem* **91** (4): 3055–3061.
- Tracanna, V., Jong, A. de, Medema, M.H., and Kuipers, O.P. (2017) Mining prokaryotes for antimicrobial compounds: from diversity to function. *FEMS Microbiology Reviews* **41** (3): 417–429.
- Villa, M.M., Bloom, R.J., Silverman, J.D., Durand, H.K., Jiang, S., Wu, A., *et al.* (2020) Interindividual Variation in Dietary Carbohydrate Metabolism by Gut Bacteria Revealed with Droplet Microfluidic Culture. *mSystems* **5** (3).
- Vincent, M.E., Liu, W., Haney, E.B., and Ismagilov, R.F. (2010) Microfluidic stochastic confinement enhances analysis of rare cells by isolating cells and creating high density environments for control of diffusible signals. *Chem Soc Rev* **39** (3): 974–984.
- Volpatti, L.R., and Yetisen, A.K. (2014) Commercialization of microfluidic devices. *Trends Biotechnol* **32** (7): 347–350.
- Watterson, W.J., Tanyeri, M., Watson, A.R., Cham, C.M., Shan, Y., Chang, E.B., *et al.* (2020) Droplet-based high-throughput cultivation for accurate screening of antibiotic resistant gut microbes. *Elife* **9**.

- Wilson, M.C., Mori, T., Rückert, C., Uria, A.R., Helf, M.J., Takada, K., *et al.* (2014) An environmental bacterial taxon with a large and distinct metabolic repertoire. *Nature* **506** (7486): 58–62.
- Wink, K., Mahler, L., Beulig, J.R., Piendl, S.K., Roth, M., and Belder, D. (2018) An integrated chip-mass spectrometry and epifluorescence approach for online monitoring of bioactive metabolites from incubated Actinobacteria in picoliter droplets. *Anal Bioanal Chem* **410** (29): 7679–7687.
- Wohlleben, W., Mast, Y., Stegmann, E., and Ziemert, N. (2016) Antibiotic drug discovery. *Microb. Biotechnol.* **9** (5): 541–548.
- Wright, G.D. (2019) Unlocking the potential of natural products in drug discovery. *Microb. Biotechnol.* **12** (1): 55–57.
- Zang, E., Brandes, S., Tovar, M., Martin, K., Mech, F., Horbert, P., *et al.* (2013) Real-time image processing for label-free enrichment of Actinobacteria cultivated in picolitre droplets. *Lab Chip* **13** (18): 3707–3713.
- Zengler, K., Toledo, G., Rappe, M., Elkins, J., Mathur, E.J., Short, J.M., and Keller, M. (2002) Cultivating the uncultured. *Proc Natl Acad Sci USA* **99** (24): 15681–15686.
- Zhao, X., and Kuipers, O.P. (2016) Identification and classification of known and putative antimicrobial compounds produced by a wide variety of Bacillales species. *BMC Genomics* **17** (1): 882.
- Ziemert, N., Alanjary, M., and Weber, T. (2016) The evolution of genome mining in microbes - a review. *Nat. Prod. Rep.* **33** (8): 988–1005.

# ACKNOWLEDGMENTS

---

I would like to take this opportunity to express my thankfulness and appreciation to a large number of colleagues and friends without whose help all of these achievements would have never been possible.

First of all I would like to thank Prof. Dr. Andreas Vilcinskas for founding and building up the Fraunhofer IME-BR in Giessen and creating a special place to work. I would like to thank Dr. Jens Glaeser and Prof. Dr. Peter E. Hammann for accepting me as a PhD student to contribute in the framework of the Sanofi-Fraunhofer Natural Products Center of Excellence (later Fraunhofer-Evotec Natural Products Center of Excellence). Moreover, I am particularly grateful for my supervisor Prof. Dr. Till F. Schäberle for his guidance, beneficial suggestions and support over the years. I hugely appreciate the time and effort he put into this work, his financial support, and thus the opportunity to pursue my doctoral degree. Furthermore, I would like to thank Prof. Till F. Schäberle and Prof. Dr. Markus Nett for their efforts as first and second reviewer, respectively.

I would like to express a very special thank you to Dr. Marius Spohn, Dr. Benedikt Leis, Dr. Maria A. Patras, Dr. Sören M. Schuler, Dr. Sanja Mihajlovic, Dr. Christoph Pöverlein, Dr. Luigi Toti, Dr. Armin Bauer and Dr. Michael Kurz. Without your help, the scientific input, the constructive feedback and your research enthusiasm, most of our publications would not have been possible.

Furthermore, this PhD thesis includes joint publications written with the following co-authors: Markus, Mona and Sandra. It was a pleasure working with you guys. I will miss our weekend sessions in the lab and all the time we spent together at work and afterwards.

Additionally, I would like to acknowledge all my colleagues at the Fraunhofer IME-BR for making my doctoral studies a memorable time of my life.

Last but surely not the least, I would like to thank my family, friends, and especially my girlfriend for their endless love, support and patience throughout this journey.

# COPYRIGHT PERMISSION

Rightslink® by Copyright Clearance Center

https://s100.copyright.com/AppDispatchServlet



Home Help Email Support Sign in Create Account

## Identification, Characterization, and Synthesis of Natural Parasitic Cysteine Protease Inhibitors: Pentacitidins Are More Potent Falcitidin Analogues



**Author:** Stephan Brinkmann, Sandra Semmler, Christian Kersten, et al

**Publication:** ACS Chemical Biology

**Publisher:** American Chemical Society

**Date:** Mar 1, 2022

Copyright © 2022, American Chemical Society

### PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms and Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from {COMPLETE REFERENCE CITATION}. Copyright {YEAR} American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your RightsLink request. No additional uses are granted (such as derivative works or other editions). For any uses, please submit a new request.

If credit is given to another source for the material you requested from RightsLink, permission must be obtained from that source.

[BACK](#)

[CLOSE WINDOW](#)

© 2022 Copyright - All Rights Reserved | [Copyright Clearance Center, Inc.](#) | [Privacy statement](#) | [Terms and Conditions](#)  
Comments? We would like to hear from you. E-mail us at [customer@copyright.com](mailto:customer@copyright.com)