On the mass spectrometric fragmentations of the bacterial sesterterpenes sestermobaraenes A–C

Anwei Hou and Jeroen S. Dickschat*

Letter

Address: Kekulé-Institute for Organic Chemistry and Biochemistry, University of Bonn, Gerhard-Domagk-Strasse 1, 53127 Bonn, Germany

Email: Jeroen S. Dickschat - dickschat@uni-bonn.de

* Corresponding author

Keywords: isotopes; mass spectrometry; reaction mechanisms; sesterterpenes; Streptomyces mobaraensis

Abstract

A $^{13}$C-labelling was introduced into each individual carbon of the recently discovered sestermobaraenes by the enzymatic conversion of the correspondingly $^{13}$C-labelled isoprenyl diphosphate precursors with the sestermobaraene synthase from Streptomyces mobaraensis. The main compounds sestermobaraenes A, B, and C were analysed by gas chromatography–mass spectrometry (GC–MS), allowing for a deep mechanistic investigation of the electron impact mass spectrometry (EIMS) fragmentation reactions of these sesterterpene hydrocarbons.

Introduction

The sestermobaraenes A–F (1–6) and sestermobaraol (7) are a series of bacterial sesterterpenes that were recently discovered by us from the actinomycete Streptomyces mobaraensis through a genome mining approach (Figure 1) [1]. All seven compounds are produced by a canonical terpene synthase, representing the first reported sesterterpene synthase of the classical type I from bacteria, that is characterised by an aspartate-rich motif (DDXXD) and an NSE triad (NDLXSXXXE) for binding of a trinuclear Mg$^{2+}$ cluster [2,3]. The Mg$^{2+}$ cations in turn bind to the diphosphate moiety of an isoprenoid diphosphate precursor and cause substrate ionisation by a diphosphate abstraction to initiate a cationic cyclisation cascade, leading to structurally highly complex and usually polycyclic terpenes in just one enzymatic transformation. The initially formed products are non-functionalised terpene hydrocarbons or, if the terminal cationic intermediate of the cyclisation cascade is trapped by water, simple alcohols. These volatile compounds can efficiently be trapped by specialised methods including the closed-loop stripping apparatus (CLSA) [4] technique or solid-phase microextraction (SPME) [5,6], and then analysed by gas chromatography–mass spectrometry (GC–MS) [7]. Through these and related techniques the volatiles from many bacteria, fungi, and plants have been investigated [8–10], which provides rapid information about the production of volatile terpenes. This
information is particularly useful in the combination with the genome sequences of the producing organism, because it allows to identify interesting candidate genes coding for terpene synthases for further studies by genome mining. A major difficulty in the GC–MS-based identification of terpenes is associated with the high similarity of the mass spectra of structurally related terpenes. For this reason, the unambiguous identification of terpenes requires either the direct comparison to an authentic standard, or, since such a standard is not always available, a very good match of the measured mass spectrum to a library spectrum and of the measured retention index to literature data. Mass spectrometric fragmentations proceed through reactions that are classified as σ-bond cleavages, α-fragmentations, inductive cleavages, McLafferty rearrangements [11], retro-Diels–Alder fragmentations [12,13], and the recently observed unusual radical-induced retro-Cope rearrangement (herein, “retro” indicates that the mass spectrometric reaction proceeds in reverse order of a thermal reaction promoted by the thermal conditions of the gaschromatographic analysis) [14]. The fragmentation reactions of structurally simple compounds such as fatty acid methyl esters have been well investigated by isotopic labelling experiments [15,16] and the knowledge allows for structural predictions based on GC–MS data [17]. The deuterium labelling technique was also applied to other compound classes such as alkylbenzenes and ketones [18-21]. For terpenes, structural proposals can only be made based on the mass spectra for structurally less complicated cases, as was exemplified for the side products of bacterial 2-methylisoborneol synthases [22], but in general the structural complexity of terpenes does not allow for such approaches. Nevertheless, more knowledge about the MS fragmentation reactions of terpenes is desirable, but represents a challenging objective as it is difficult to get access to the isotopically labelled terpenes needed for deep and conclusive insights. The early investigations by Djerassi and co-workers have made use of semisynthetic deuterated terpenes [23-25]. While deuterium can reveal specific hydrogen migrations in the fragmentation reactions, it is comparably cheap, and can often easily be introduced, e.g., into C-H-acidic positions, a drawback of deuterium usage lies in possible kinetic isotope effects [21]. Also MS/MS-based techniques have been used to study the fragmentations of terpenes [26-28], but isotopic labelling experiments can give more detailed and conclusive insights. We have recently investigated the MS fragmentation mechanisms of several sesqui- and diterpenes in a series of studies that made use of $^{13}$C-labelled terpene precursors to systematically introduce single labellings into each individual carbon position by enzymatic synthesis [14,29-32]. Here we report on the MS fragmentation mechanisms for the bacterial compounds sestermobaraenes A, B, and C, representing the first mechanistic study of this kind for sesterterpenes.

**Results and Discussion**

**Experimental basis**

The 25 isotopomers of (1$^{3}$C)geranyl farnesyl diphosphate (GFPP) were enzymatically prepared from the correspondingly labelled geranyl diphosphate (GPP), farnesyl diphosphate (FPP), geranylgeranyl diphosphate (GGPP), and isopentenyl diphosphate (IPP) with geranylfarnesyl diphosphate synthase (GFPPS) and then converted into mixtures of the sesterterpenes 1–7 by the sestermobaraene synthase from *Streptomyces mobaraensis* (SmTS1). All 1$^{3}$C-labelled terpene precursors were made available by synthesis in our laboratory in high isotopic purity with 1$^{3}$C substitutions of nearly 100% [33-37]. The compound mixtures were subsequently analysed by GC–MS and the mass spectra of the unlabelled compounds 1–3 and their 25 singly 1$^{3}$C-labelled isotopomers are summarised in Figures S1–S3 in Supporting Information File 1. Investigations
on the mass spectrometric fragmentation mechanisms for the minor products 4–7 of SmTS1 are not included in this study, because in some cases no high quality mass spectra could be obtained. The mass spectra of the unlabelled compounds show several pronounced signals for fragment ions (m/z, mass-to-charge ratio). If a signal in a mass spectrum for a particular $^{13}$C-labeled isotopomer of a compound under investigation is in comparison to the non-labelled compound clearly increased by 1 Da, this means that the labelled carbon fully contributes to the fragment ion. Accordingly, if the signal is clearly not shifted, this means the labelled carbon is not part of the fragment ion. Also cases in between these clear situations exist, namely if a signal in the mass spectrum is a result of two or more fragment ions formed from different parts of the molecule, a labelled carbon may or may not contribute to its formation. A quick overview can be given in a position-specific mass shift analysis for a fragment ion $m/z$ (PMA$_{m/z}$), in which fully contributing carbons are marked by red dots, partially contributing carbons by green dots, and carbons that do not contribute remain without a mark (Figures 2–4, vide infra). Because usually multiple fragmentation reactions lead to the formation of the ions observed in the low molecular weight region, their formation will not be discussed (an exception is the base peak at $m/z = 120$ for all three compounds). The method also finds its limitations for fragment ions buried within a group of peaks. Such fragment ions will not be discussed in this work.

**Fragmentation mechanisms for sestermobaraene A (1)**

The position-specific mass shift analyses (Figure 2) for several prominent fragment ions observed in the mass spectrum of sestermobaraene A (1) are based on a comparison of the mass spectrum of the unlabelled compound 1 to the mass spectra of the 25 isotopomers of $(^{13}$C)$_2$-1 (Figure S1, Supporting Information File 1). As can be concluded from these analyses, the fragment ions observed at $m/z = 312$, $m/z = 206$, and the base peak at $m/z = 120$ are formed by a loss of a clearly defined portion of 1, while the fragment ions at $m/z = 325$ and $m/z = 297$ arise through various reactions with losses of different portions of the molecule that can, however, still be rationalised. For the other fragment ions in the mass spectrum of 1 the situation is less clear and their formation will not be discussed here.

The formation of the fragment ion at $m/z = 325$ requires the loss of one methyl group for which only C22, C23, C24, and C25, but not C20 and C21 show a significant participation. The most prominent loss is observed for C23 in an allylic position of the double bond in 1. After electron impact ionisation preferentially at the π-system of the olefinic double bond the radical cation $1^{++}$ is obtained from which the methyl group C23 can directly be lost by an α-cleavage leading to fragment a1$^{+}$ (Scheme 1A). However, the radical centred at the bridgehead carbon C11 is orthogonal to, or in other words, not in conjugation with the radical cation at C12–13. Therefore, an energetically more feasible process may be represented by an inductive cleavage leading to b1$^{+}$, a hydrogen rearrangement to c1$^{+}$, and an α-cleavage to d1$^{+}$ (Scheme 1B). The formation of the fragment ion at $m/z = 312$ proceeds through a highly specific loss of the C8–9 portion of 1. This is explainable from b1$^{+}$ by a sequence of two α-cleavages first to e1$^{+}$ and then to f1$^{+}$ with a neutral loss of ethylene (Scheme 1C). The fragment ion at $m/z = 297$ requires the loss of C$_2$H$_7$ which can be achieved by various reactions, as indicated by the PMA$_{297}$. This may be realised by the cleavage of an intact C$_3$H$_7$ unit originating from the isopropyl group C20–19–21 or, by involving multiple C–C bond cleavages and hydrogen rearrangements, from the C25–3–4 portion. Alternatively, a combined loss of the C8–9 moiety and one methyl group (C22, C23, C24, or C25) is possible which basically combines the fragmentations of Scheme 1A and Scheme 1B. The loss of the isopropyl group C20–19–21 can be achieved by an inductive cleavage of 1$^{++}$ to g1$^{+}$ followed by an α-cleavage to h1$^{+}$ (Scheme 1D). Starting from e1$^{+}$, two α-cleavages with the extrusion of ethylene can lead to i1$^{+}$ that upon a third α-fragmentation with loss of the methyl group C23 results in j1$^{+}$ (Scheme 1E). The fragmentation of the C25–3–4 portion can be explained starting from 1$^{+}$ by a hydrogen rearrangement to k1$^{+}$ and α-cleavage to 11$^{+}$ (Scheme 1F). Another

![Figure 2: Position-specific mass shift analyses for 1. Carbons that contribute fully to the formation of a fragment ion are indicated by red dots, partially contributing carbons are marked by green dots, and unlabelled carbons do not contribute and are thus cleaved off by the fragmentation reaction.](image-url)
Scheme 1: The EIMS fragmentation mechanisms for 1 explaining the formation of the fragment ions at \( m/z \) = 325, 312, and 297. Lost carbons are marked by purple dots.
hydrogen rearrangement combined with an α-fragmentation then leads to the allyl cation \( m_1^{+} \) which may undergo a third hydrogen rearrangement to \( u_1^{+} \) and final cleavage of a propyl group to \( o_1^{+} \).

The formation of the fragment ion at \( m/z = 206 \) proceeds with the loss of the portion represented by carbons C25–3–4–5–6–10(–9–8)–11–23 and can be proposed as shown in Scheme 2A. After the ionisation to \( 1^{+} \) a hydrogen rearrangement leads to \( p_1^{+} \) that further reacts by an inductive ring opening and α-cleavage to \( q_1^{+} \). Another α-fragmentation to \( r_1^{+} \) may be followed by a hydrogen rearrangement to \( s_1^{+} \) and two α-cleavages to \( t_1^{+} \), giving an alternative mechanistic explanation for the fragment ion at \( m/z = 312 \) by loss of C8–9.

Another the hydrogen rearrangement to \( u_1^{+} \) sets the stage for a final α-fragmentation with the neutral loss of o-xylene to \( v_1^{+} \). The base peak in the mass spectrum of 1 is formed from carbons C25–3–4–5–6–10–9–8–11–23, which can also be explained starting from \( p_1^{+} \) by three sequential α-cleavages through \( w_1^{+} \) to \( x_1^{+} \) (Scheme 2B). The inductive cleavage with hydride migration leads to \( y_1^{+} \) representing the minor fragment ion at \( m/z = 122 \) that may efficiently lose two hydrogens to give the conjugated system in \( z_1^{+} \).

**Fragmentation mechanisms for sesternobaraene B (2)**

The position-specific mass shift analyses for sesternobaraene B (2) are based on the mass spectrum of the unlabelled compound

---

**Scheme 2:** The EIIMS fragmentation mechanisms for 1 explaining the formation of fragment ions at \( m/z = 206 \) and 120. Lost carbons are marked by purple dots.
in comparison to those of its 25 $^{13}$C-labelled isotopomers (Figure S2 in Supporting Information File 1). Clear results could be obtained for the fragment ions in the high mass region at $m/z = 325, 312,$ and 297, for the base peak at $m/z = 120,$ and the prominent fragment ion at $m/z = 203.$ The results of the analyses are summarised in Figure 3.

Similarly to the observations made for 1, also for 2 the formation of the fragment ion at $m/z = 325$ by loss of one methyl group proceeds by the cleavage of C22, C23, C24, or C25, while the fragmentation of C20 or C21 does not make a significant contribution. Notably, even from the olefinic methylene group C23 a methyl group can be cleaved off, which requires hydrogen rearrangements prior to the fragmentation. A possible mechanism starts from $2^+$ by the hydrogen rearrangement to $a2^+$ and a hydride shift to $b2^+$ (Scheme 3A). This hydride migration is in reverse order compared to a similar step along the cationic cyclisation cascade during the biosynthesis of 2 (Scheme S1 in Supporting Information File 1). The subsequent inductive ring opening to $c2^+$ and $\alpha$-cleavage of C23 result in $d2^+.$ The losses of the other methyl groups can be understood more easily, e.g., two $\alpha$-fragmentations from $2^+$ explain the formation of $e2^+$ with the loss of C25 (Scheme 3B). The fragment ion at $m/z = 312$ arises by the loss of the C8–9 portion through a double $\alpha$-cleavage from $2^+$, yielding to $f2^+$ (Scheme 3C). Also for compound 2 different mechanisms for the formation of the fragment ion at $m/z = 297$ are observed, including the loss of the isopropyl group C20–19–21 or the loss of C8–9 and one methyl group. The cleavage of the isopropyl group is possible from $c2^+$ by an inductive ring opening to $g2^+$ and $\alpha$-fragmentation to $h2^+$ (Scheme 3D). Alternatively, $c2^+$ can react by two $\alpha$-cleavages leading to $i2^+$ with a neutral loss of ethylene, followed by another $\alpha$-cleavage of C23 to $j2^+$ (Scheme 3E). The fragment ion at $m/z = 297$ can also be rationalised from $f2^+$ by two $\alpha$-fragmentations with the loss of C25 to result in $k2^+$ (Scheme 3F).

The position-specific mass shift analysis for $m/z = 203$ indicates the formation of this fragment ion by two overlaid mechanisms that both involve the loss of C14–15(–22)–16–17–18(–21)–20 plus either C13 or C1. A mechanistic model for the first case with loss of C13 starts from $2^+$ by a hydrogen rearrangement to $l2^+$ and an $\alpha$-fragmentation to $m2^+$, followed by another hydrogen transfer to $n2^+$ and $\alpha$-cleavage to $o2^+$ (Scheme 4A). The second possibility with the loss of C1 is explainable from $l2^+$ by a hydrogen migration to $p2^+$ and an $\alpha$-fragmentation to $q2^+$, followed by two more $\alpha$-fragmentations to $r2^+$ (Scheme 4B). A final $\alpha$-cleavage then yields the target ion $s2^+.$ The generation of the base peak ion at $m/z = 120$ from the C25–3–4–5–6–10(–9)–11–23 moiety of 2 is more difficult to understand, as it must proceed with four C–C bond cleavages. Interestingly, for 2 the base peak is made up from the same portion of the molecule as for 1, but while 1 has a bond between C3 and C11, this bond is missing in 2 that has a bond between C2 and C12 instead. For 1 the base peak was nicely explainable by the formation of an ionised aromatic ring system. In the first instance, it seems difficult to parallel this for 2, but if for the first steps after ionisation to $2^+$ a skeletal rearrangement to $t2^+$ and a hydrogen transfer to $u2^+$ are assumed, the parallelism of the fragmentation mechanisms becomes more obvious (Scheme 4C). Subsequent steps may include an inductive ring opening to $v2^+$, another hydrogen rearrangement to $w2^+$, and two $\alpha$-cleavages to $x2^+.$ Another hydrogen rearrangement and elimination of two hydrogen atoms lead to $y2^+$ which is identical to $z1^+$ in the fragmentation mechanism for the base peak ion of 1.

Fragmentation mechanisms for sestermobaraene C (3)

For sestermobaraene C (3) the position-specific mass shift analyses based on the mass spectra of the unlabelled versus all 25 isotopomers of the singly $^{13}$C-labelled material (Figure S3 in Supporting Information File 1) also gave unambiguous results for the fragment ions at $m/z = 325, 312, 297, 206,$ and the base peak at $m/z = 120$ (Figure 4), which is similar to the correspond-
Scheme 3: The EI-MS fragmentation mechanisms for 2 explaining the formation of the fragment ions at m/z = 325, 312, and 297. Lost carbons are marked by purple dots.

ing analyses for 1 and 2 not only in the nominal masses of the fragment ions, but also in terms of the portions of the carbon skeletons these fragments arise from. Thus, it can be expected that similar fragmentation reactions as discussed for 1 and 2 above can lead to their formation. One notable difference is observed for the fragment ions at m/z = 312 and 297 that are formed with a partial loss of C11–23, which was not observed for compounds 1 and 2.
The formation of the fragment ion at \( m/z = 325 \) proceeds with cleavage of C22, C23, C24, or C25, as observed before for compounds 1 and 2. Especially noteworthy is the cleavage of the methylene carbon C25, which is explainable from \( 3^+ \) by a hydrogen rearrangement to \( a3^+ \), followed by a hydride shift to \( b3^+ \) and an \( \alpha \)-fragmentation to \( c3^+ \) (Scheme 5A). The alternative loss of C22 is possible from \( 3^+ \) by two sequential \( \alpha \)-cleavages via \( d3^+ \) to \( e3^+ \) (Scheme 5B). The fragment ion at

Scheme 4: The EI-MS fragmentation mechanisms for 2 explaining the formation of the fragment ions at \( m/z = 203 \) and 120. Lost carbons are marked by purple dots.
m/z = 312 involves the loss of either the C8–9 or the C11–23 portion. The first case can be understood starting from h3** by two inductive cleavages with the neutral loss of ethylene to f3** and then g3** (Scheme 5C), while the second case may start from a3** by an α-cleavage with hydrogen rearrangement to h3** and another subsequent α-fragmentation to i3** (Scheme 5D). Similar to the observations for compounds 1 and 2, the fragment ion at m/z = 297 of 3 is generated by the loss of C8–9 and one methyl group or of the isopropyl group C20–19–21. In addition, the combined loss of C11–23 and one methyl group also contributes to its formation. The possible mechanistic models include a simple α-fragmentation with the loss of C25 from g3** to j3** (Scheme 5E), a sequence of three α-cleavages from 3** through k3** leading to l3** (Scheme 5F), and a double α-fragmentation in i3** that explains the formation of m3** (Scheme 5G).

The fragment ion at m/z = 206 arises from the C25–3–4–5–6–10(–9–8)–11–23 moiety of 3. Its formation requires multiple bond cleavages and hydrogen transfers and is thus a multistep process (Scheme 6A). Starting from 3**, a hydride shift to n3** and skeletal rearrangement lead to o3**. A subsequent hydrogen rearrangement of this primary radical yields the tertiary radical p3** that can undergo an α-fragmentation to q3**+, followed by hydrogen rearrangement to r3**+, setting the stage for the next α-cleavage to s3**+. The same principle can explain the last bond cleavage: A hydride shift to t3**+ adjusts the reactivity for the α-fragmentation to u3**+. Notably, the intermediate q3**+ is also a good starting point to explain the formation of the base peak ion at m/z = 120 (Scheme 6B). The inductive opening produces v3**+, that, upon α-cleavage with hydrogen rearrangement, leads to w3**+ (m/z = 122). The base peak ion x3** then results by the loss of two hydrogens.

**Conclusion**

In this work we demonstrated that 13C-labellings can efficiently be introduced by terpene synthase catalysed reactions into each single position of a terpene, which is useful for the deep investigations on mass spectrometric fragmentation reactions. The present study provides the first example for such investigations on sesterterpene fragmentations. The applied method, once the synthetic 13C-labelled oligoprenyl diphosphates are at hand, is superior to any other approach for the introduction of labellings, also because the labelled terpene precursors can be used for studies on many different terpenes for which terpene synthases are available. In the present case it is intriguing to learn that, although the structures of the three investigated sesterterpenes are different, not only similar fragment ions are observed, but also similar reactions lead to their formation, which is most prominently observed for the common base peak ion at m/z = 120 for all three compounds. This means that the sesterterpenes have a common intrinsic reactivity that is in the first instance reflected by their joint biosynthesis, but also by their similar behaviour in the comparably high-energy chemistry of mass spectrometric fragmentation reactions. Further support for the similar reactivity of the investigated compounds during biosynthesis and mass spectrometric fragmentations is given by the notable observation of hydride shifts that occur in both of these processes. However, the three compounds show also some differences in their mass spectrometric fragmentation, e.g., for compound 2 a strong fragment ion is observed at m/z = 203, which is much less relevant for the other two compounds. It should be emphasised that the mechanistic hypotheses presented in this work are solely based on the 13C-labellings, while specific hydrogen migrations would need to be followed by deuterium labellings, but in these cases data interpretation may be hampered by kinetic isotope effects. Nevertheless, at the current stage it cannot be excluded that such experiments could demonstrate the need for a refinement of the fragmentation mechanisms for certain fragment ions presented here. We will continue our investigations on terpene fragmentations in EIMS in the future by the strategy applied in this work to learn more about the underlying reaction mechanisms.
Scheme 5: The EIMS fragmentation mechanisms for 3 explaining the formation of the fragment ions at m/z = 325, 312, and 297. Lost carbons are marked by purple dots.
Scheme 6: The EIMS fragmentation mechanisms for 3 explaining the formation of the fragment ion at \( m/z \) = 206 and the base peak ion at \( m/z \) = 120. Lost carbons are marked by purple dots.

**Experimental**

**Preparation of \(^{13}\text{C}\)-labelled compounds 1–3 and GC–MS analysis**

The 25 isotopomers of \(^{13}\text{C}\)-1, \(^{13}\text{C}\)-2, and \(^{13}\text{C}\)-3 were prepared enzymatically with SmTS1 from the correspondingly labelled oligoprenyl diphosphates as reported previously [1]. The compounds were obtained as mixtures that were directly analysed by GC–MS. The GC–MS analyses were performed using a 7890A GC connected to a 5977A mass selective detector (Agilent, Hewlett-Packard Company, Wilmington, USA). The gas chromatographic separation was done using a HP5-MS fused silica capillary column (30 m, 0.25 mm i.d., 0.25 μm film, Agilent). The GC settings were 1) inlet pressure: 77.1 kPa, He 23.3 mL min\(^{-1}\); 2) injector temperature: 250 °C; 3) injection volume: 2 μL; 4) injector operation mode: splitless (60 s valve time); 5) carrier gas: He at 1.2 mL min\(^{-1}\); 6) temperature program: 5 min at 50 °C, then increasing with a ramp of 5 °C min\(^{-1}\) to 320 °C. The MS settings were 1) transfer line: 300 °C; 2) electron energy: 70 eV.
This work was funded by the DFG (DI1536/7-2).

ORCID® iDs
Jeroen S. Dickschat - https://orcid.org/0000-0002-0102-0631

References
License and Terms

This is an Open Access article under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0). Please note that the reuse, redistribution and reproduction in particular requires that the authors and source are credited.

The license is subject to the Beilstein Journal of Organic Chemistry terms and conditions: (https://www.beilstein-journals.org/bjoc)

The definitive version of this article is the electronic one which can be found at: https://doi.org/10.3762/bjoc.16.231
Secondary metabolites of *Bacillus subtilis* impact the assembly of soil-derived semisynthetic bacterial communities

Heiko T. Kiesewalter¹, Carlos N. Lozano-Andrade¹, Mikael L. Strube² and Ákos T. Kovács*¹

Full Research Paper

Address:
1Bacterial Interactions and Evolution Group, DTU Bioengineering, Technical University of Denmark, Kgs. Lyngby, Denmark and
2Bacterial Ecophysiology and Biotechnology Group, DTU Bioengineering, Technical University of Denmark, Kgs. Lyngby, Denmark

Email:
Ákos T. Kovács* - atkovacs@dtu.dk

* Corresponding author

Keywords:
*Bacillus subtilis*; bacterial community; chemical ecology; *Lysinibacillus fusiformis*; nonribosomal peptides; surfactin

Abstract

Secondary metabolites provide *Bacillus subtilis* with increased competitiveness towards other microorganisms. In particular, nonribosomal peptides (NRPs) have an enormous antimicrobial potential by causing cell lysis, perforation of fungal membranes, enzyme inhibition, or disruption of bacterial protein synthesis. This knowledge was primarily acquired in vitro when *B. subtilis* was competing with other microbial monocultures. However, our understanding of the true ecological role of these small molecules is limited. In this study, we have established soil-derived semisynthetic mock communities containing 13 main genera and supplemented them with *B. subtilis* P5_B1 WT, the NRP-deficient strain *sfp*, or single-NRP mutants incapable of producing surfactin, plipastatin, or bacillaene. Through 16S amplicon sequencing, it was revealed that the invasion of NRP-producing *B. subtilis* strains had no major impact on the bacterial communities. Still, the abundance of the two genera *Lysinibacillus* and *Viridibacillus* was reduced. Interestingly, this effect was diminished in communities supplemented with the NRP-deficient strain. Growth profiling of *Lysinibacillus fusiformis* M5 exposed to either spent media of the *B. subtilis* strains or pure surfactin indicated the sensitivity of this strain towards the biosurfactant surfactin. Our study provides a more in-depth insight into the influence of *B. subtilis* NRPs on semisynthetic bacterial communities and helps to understand their ecological role.

Introduction

In nature, bacteria live in complex communities where they interact with various other microorganisms. Most microbial communities are influencing biochemical cycles and impact agriculture, from which the latter is primarily mediated due to plant-growth promotion [1–4]. Extensive research has been conducted in the last decade to scrutinise the occurring natural pro-
cesses and their impact on the environment, to investigate the functions and interactions of community members, such as metabolite cross-feeding interactions, and to eventually engineer them [5-7]. The soil is one of the five main habitats of bacteria and archaea [8]. Soil is very heterogeneous since it exhibits spatial variability in terms of nutrient availability and geochemical features [9]. Therefore, soil consists of microbial hotspots, indicating faster process rates than the average soil [10]. One such microbial hotspot is the rhizosphere, harbouring microbial communities where various interactions between bacteria, fungi, and plants take place [11]. The composition of microbial communities depends on multiple factors. Studies have revealed that the composition of bacterial soil communities varies at the same sampling site during different seasons [12,13]. Moreover, it has been recently demonstrated that precipitation rates have a significant impact on bacterial communities since bacterial soil communities have a higher diversity in dry than in rainy seasons [14]. Besides the seasonal factors, even different plant species with varying root exudates as well as various soil types impact the microbial community composition in the rhizosphere [15-20]. Microbial communities can consist of hundreds and thousands of diverse species, which makes investigations very challenging and hard to reproduce. One alternative approach is to establish a host-associated synthetic community, usually with members of the same kingdom, with a defined composition but fewer members [19,21]. Lebeis et al. used an artificial community of 38 bacterial strains to demonstrate that plant phytohormones sculpt the root microbiome [19]. In comparison, Niu et al. established a seven-species bacterial community based on host selection to mimic the principle root microbiome of maize [22].

Secondary metabolites (SMs) are believed to be important mediators of the interactions between microorganisms [23]. Many of them are well-studied in vitro, but the true ecological role of SMs is still the subject of investigations. Different opinions about their primary role in nature exist in the literature; some share the view that SMs are mainly microbial weapons but others instead designate them as signalling molecules [24-27]. Additionally, Pettit [28] and Wakefield et al. [29] have demonstrated in 2009 and 2017, respectively, that some bacterial or fungal biosynthetic gene clusters are silent when strains are grown in monocultures under standard laboratory conditions but are expressed in intra- or interkingdom co- or multicultures. Furthermore, they could show that some SMs had a higher production rate in multicultures, highlighting that neighbouring organisms induce and increase the SM production in the tested strains.

*Bacillus subtilis* is a well-studied soil bacterium and is used as a model organism for biofilm formation and sporulation [30]. It has been shown that several members of the *B. subtilis* species complex have exceptional plant growth promoting and plant health improving properties by suppressing plant pathogenic bacteria and fungi [31]. However, it is not completely understood how soil-administered *Bacillus* spp. affect the indigenous microbial communities. Gadhave et al. have shown that the supplementation of *B. subtilis*, *Bacillus amyloliquefaciens* (now identified as *Bacillus velezensis*), and *Bacillus cereus* to the roots of broccoli plants led to species-dependent changes in the diversity, evenness, and relative abundances of endophytic bacterial communities [32]. Like many other soil bacteria, *B. subtilis* and other *Bacillus* spp. produce various SMs [33,34]. The most prominent and bioactive SMs are nonribosomal peptides (NRPs), of which isoforms belong to the families of surfactins, fengycins, or iturins [35,36] (Figure 1). They are biosynthesised by large enzyme complexes, nonribosomal peptide synthetases (NRPSs). For the biosynthesis of *B. subtilis* NRPs, the phosphopantetheinyl transferase Sfp is needed since it has been shown to activate the peptidyl carrier protein domains, converting it from the inactive apo-form to the active holo-form [37]. *B. subtilis* has four sfp-dependent SMs, of which three are synthesised by NRPS gene clusters (surfactin, plipastatin, and bacillibactin) and one by a hybrid NRPS–PKS gene cluster (bacillaeana, Figure 1). The well-studied biosurfactant surfactin, encoded by the *srfAA-AD* gene cluster, reduces the surface tension needed for swelling and sliding motility [38,39]. The surfactin bioactivity is specifically evoked by the surfactant activity triggering cell lysis due to penetration of the bacterial lipid bilayer membranes and the formation of ion-conducting channels [40-42]. The bioactivity of surfactin was shown against *Listeria* spp. and *Legionella monocytogenes* [43,44]. It is presumed that the antifungal plipastatin, expressed from the *ppSA-E* gene cluster, acts as an inhibitor of phospholipase A2, forming pores in the fungal membrane and causing morphological changes in the fungal membrane and cell wall [45,46]. This antifungal potential was demonstrated primarily against various filamentous fungi [47-51]. The broad-spectrum antibiotic bacillaeana, synthesised by the *pksB-S* gene cluster, is mainly targeting bacterial protein synthesis [52]. Still, it was also shown that it could protect cells and spores from predation [53]. We recently demonstrated that the production of these NRPs varies among coisolated *B. subtilis* environmental strains due to missing core genes or potentially altered gene regulation, highlighting the existing natural diversity of SM production in this species [51].

In this study, we focus on soil-derived semisynthetic bacterial mock communities and describe how these are affected by a *B. subtilis* strain that was previously isolated from the same sampling site from which the bacterial mock communities originated. With an NRP-mutant-based approach, we investigated...
the impact of NRPs on the establishment and composition of the bacterial communities. We previously demonstrated that the strain P5_B1 produces the NRPs surfactin and plipastatin and has further BGC predictions for the NRPs bacillaene and bacillibactin [51]. It was revealed by 16S rRNA amplicon sequencing that the established semisynthetic mock communities contained 13 genera with a relative abundance of >0.19% in at least one mock community. Furthermore, it was demonstrated that the addition of \textit{B. subtilis} suppressed the genera \textit{Lysinibacillus} and \textit{Viridibacillus}. Additional optical density (OD)-based growth monitoring of the selected strain \textit{Lysinibacillus fusiformis} M5 confirmed the impact of \textit{B. subtilis}-produced surfactin on the growth.

**Results**

**Impact of \textit{B. subtilis} secondary metabolites on taxonomic groups in semisynthetic mock communities**

We established soil-derived semisynthetic mock communities and supplemented them with \textit{B. subtilis} WT P5_B1, the corresponding NRP-deficient strain \textit{sfp}, or the single-NRP mutants.
Figure 2: Overview of the experimental setup. A soil suspension, obtained from a soil sample, was used as an inoculum for four independent replicates and preincubated for 12 h. Enriched precultures were aliquoted and supplemented with 10% B. subtilis strains or left untreated and incubated for 48 h. DNA was extracted from the soil sample, preincubated soil suspensions, and mock communities. Parts of this figure were created using BioRender.com.

Figure 3: The taxonomic summaries are showing the relative abundance of the most abundant genera for each replicate of the soil sample (“soil”), 12 h preincubated soil suspensions (“Pre”), and untreated (“Control”) or treated mock communities with either B. subtilis wild type (“WT”), the NRP-deficient strain sfp, the surfactin mutant srfAC, the plipastatin mutant ΔppsC, or the bacillaene mutant ΔpksL, cocultivated for 48 h. Genera are classified as “other” when the relative abundance is <2% (“Soil”), <1% (“Pre”), or <0.19% (in all differently treated mock communities).
Figure 4: Diversity analyses of the soil sample (“Soil”), 12 h preincubated soil suspensions (“Pre”), and untreated (“Control”) or treated mock communities with either *B. subtilis* wild type (“WT”), the NRP-deficient strain *sfp*, the surfactin mutant *srfAC*, the plipastatin mutant *ΔppsC*, or the bacillaene mutant *ΔpksL*, cocultivated for 48 h. A) Alpha diversity (in Shannon) of the different samples. Each point represents a replicate, while the line indicates the mean of the Shannon diversity indexes. B) Beta diversity of the mock communities calculated with the Bray–Curtis dissimilarity and visualised as circles in a nMDS. The vectors, each labelled with the corresponding genus, represent the ASVs, with the highest correlating with the nMDS ordination. The vector lengths are proportional to the level of correlation.
ically assigned to the genera Lysinibacillus, Acinetobacter, and Viridibacillus, correlated with the control and two sfp-treated communities. This observation suggests that the absence of NRP-producing B. subtilis resulted in an increased abundance of these. Furthermore, two ASVs of the genus Acinetobacter correlated best with the communities supplemented with the NRP-producing B. subtilis strains, hinting a higher frequency of these in NRP-treated communities. Additionally, three ASVs, identified as Pseudomonas, Citrobacter, and Sphingobacterium, correlated with two communities treated with the surfactin mutant. A similar but smaller correlation with two bacillaene mutant-treated communities was detectable as well. These results imply a negative impact of either surfactin or bacillaene on the four ASVs. Interestingly, the vector-based analysis suggests that, depending on the ASVs, the genus Acinetobacter is both positively and negatively affected by the NRPs.

In conclusion, the alpha diversity analyses revealed that species richness and evenness were reduced in the in vitro communities compared to the soil community. Furthermore, 12 h preincubated soil suspensions showed a reduced diversity compared to the mock communities incubated for 48 h. Nevertheless, we could not detect an effect of the supplemented B. subtilis strains on the diversity. However, the beta diversity results suggested that the addition of NRP-producing B. subtilis strains influenced the composition of the mock communities. Mainly ASVs belonging to the genera Lysinibacillus, Viridibacillus, and Acinetobacter were affected by the presence or absence of B. subtilis NRPs in the bacterial mock communities.

The diversity, in particular the species evenness, increased independently of the treatment in all established mock communities, compared to the precultivated soil suspensions and contained 11–18 genera (Figure 3). The most abundant genera, having a proportion greater than 0.19% in at least one B. subtilis-treated or untreated mock community were Acinetobacter, Lysinibacillus, Pseudomonas, Chryseobacterium, Bacillus, Sphingobacterium, Stenotrophomonas, Paenibacillus, Citrobacter, Serratia, Achromobacter, Viridibacillus, and Pantoea. Note-worthy, the prevalence of the Bacillus genus was comparable in the B. subtilis-treated communities (4–9%) and the control (5–10%). In the latter, the present Bacillus sp. originated only from the soil suspension, highlighting that the additional supplementation of B. subtilis did not affect the relative abundance of the genus Bacillus after 48 h cocultivation. Interestingly, the only genera detected in both the in vitro mock communities and the soil samples were Bacillus, Lysinibacillus, and Paenibacillus. The remaining most abundant genera in the mock communities were below the detection limit.

The comparison of the abundance ratios between the control communities and B. subtilis WT-treated communities revealed that Lysinibacillus and Viridibacillus were significantly decreased 9.4-fold (P ≤ 0.001) and 8.3-fold (P ≤ 0.01), respectively, in the communities supplemented with B. subtilis WT (Figure 5A). None of the other genera was affected by the addition of this strain. In comparison, we could only detect a 1.8-fold significant reduction (P ≤ 0.05) of Lysinibacillus in the sfp-treated communities compared to the untreated communities, and thus a greatly diminished effect compared to the

**Figure 5:** Abundance ratios for each genus and replicate (points) in the control community compared to the WT-treated (A) and to the sfp-treated community (B). Red-box plots highlight the statistical significance, which is defined as P ≤ 0.05 (*), P ≤ 0.01 (**), and P ≤ 0.001 (***)
WT-treated samples was evident (Figure 5B). Also, we could not observe a significant reduction of Viridibacillus, but besides Lysinibacillus, also Stenotrophomonas was 1.7-fold ($P \leq 0.05$) significantly reduced in these communities. The direct comparison of WT- and sfp-treated communities confirmed the NRP-dependent suppression of both Lysinibacillus and Viridibacillus in the WT-treated communities and the suppression of Stenotrophomonas in the sfp-treated communities (Figure S2, Supporting Information File 1).

Concentrating on Lysinibacillus, the highest abundance of this genus was discernible in the control assays (13.9%), which was significantly different compared to all other B. subtilis-treated assays (Figure 6). However, when B. subtilis P5_B1 WT was added to the mock communities, a significant decrease ($P \leq 0.001$) of Lysinibacillus (1.2%) compared to the control communities was discovered. Furthermore, when we added the NRP-deficient strain sfp, we could notice a significantly higher abundance of Lysinibacillus (8.6%) compared to the WT-treated communities ($P \leq 0.001$) but still a significantly lower prevalence compared to the control communities ($P \leq 0.05$). Compared to the WT-treated communities, the frequency of Lysinibacillus was slightly but not significantly higher in the communities treated with the single-NRP mutants srfAC (2.0%) and ΔppsC (3.3%). The abundance of Lysinibacillus in the assays containing the ΔpksL strain (5.3%) was significantly higher ($P \leq 0.01$) than in the WT-treated assays. However, the Lysinibacillus abundance in ΔpksL-treated communities was not significantly different from the ΔppsC- or sfp-treated communities.

In summary, Lysinibacillus was affected by the addition of B. subtilis independent of the NRPs, but when B. subtilis strains capable of producing them were present, the impact on Lysinibacillus was enhanced. Furthermore, the results indicate that bacillaene had the strongest and surfactin the weakest effect on Lysinibacillus in the mock communities.

The second genus affected by the addition of B. subtilis was Viridibacillus, which had a very low abundance in the control mock communities (0.49%) compared to Lysinibacillus (Figure S3, Supporting Information File 1). However, when B. subtilis WT was added to the community, Viridibacillus indicated a significantly lower ($P \leq 0.01$) abundance (0.03%) compared to the WT and the control communities. Notably, in two of the WT-treated community replicates, Viridibacillus was below the detection level. Nevertheless, the abundance of this genus in the sfp-treated communities (0.26%) was statistically not significant in comparison to the WT and the control communities. Furthermore, the addition of the single-NRP mutants srfAC, ΔppsC, and ΔpksL resulted in communities with Viridibacillus frequencies similar to the WT-treated communities (0.08%, 0.05%, and 0.00%, respectively). Viridibacillus as well as Lysinibacillus was affected by the addition of B. subtilis to the communities. However, no particular NRP could be assigned to the reduced frequency of Viridibacillus.

**Growth properties of L. fusiformis M5 supplemented with B. subtilis spent media**

The main finding from the semisynthetic mock community experiment indicated that the genus Lysinibacillus was negatively affected by the addition of B. subtilis P5_B1 WT and that NRPs enhance the suppression. To dissect the direct impact of a particular NRP in this inhibition, we monitored the growth of L. fusiformis M5, a previously isolated Lysinibacillus species [54], over 24 h when treated with different proportions of spent media from B. subtilis WT and the corresponding NRP mutants (Figure 7). When we added 52.80% of spent medium to L. fusiformis, we observed the fastest entry into the exponential growth phase in the untreated assay. Interestingly, the addition of spent medium of either WT, ΔppsC, or ΔpksL caused a delay of entering into this growth phase of approximately 11–13 h in L. fusiformis compared to the control. Such a strong effect was not observed when the spent medium of the sfp or srfAC mutant was added. The addition of these two spent media caused only a slight delay of the exponential growth phase of L. fusiformis, although spent sfp medium had a lower effect on L. fusiformis compared to spent srfAC medium. When 23.00% of spent medium was added, no growth differences could be detected anymore between the control and the sfp-treated assays in the exponential growth phase. Furthermore, the effect of spent WT
Figure 7: Growth curves of *L. fusiformis* M5 exposed to spent media from 48 h *B. subtilis* cultures and without treatment (“control”). The spent medium concentration of 10.02% to 52.80%, acquired with a serial dilution, indicates the proportion of spent medium from the total volume. The error bars represent the standard error. *N ≥ 6*. OD<sub>600</sub> = optical density at 600 nm.

medium seems to be reduced at this concentration, but the spent media of ΔppsC and ΔpksL maintained their growth inhibition potential. The lowest concentration of a spent medium having an inhibitory effect was 10.02%. At this concentration, only the spent media of ΔppsC and ΔpksL affected the growth of *L. fusiformis*, even though it was weakened compared to using higher concentrations. Intriguingly, a higher level of aggregation was observed in the *L. fusiformis* assays supplemented with the spent medium of *sfp* compared to the other assays, which caused higher and variable OD measurements in the stationary phase of the growth curves (Figure S4, Supporting Information File 1). Finally, it was noted that the final cell density was slightly higher in the assays supplemented with the spent medium compared to the control assays.

These results revealed that *B. subtilis*-mediated inhibition of *L. fusiformis* is NRP-dependent since the spent medium of the NRP-deficient strain *sfp* had an only minor impact. Moreover, we hypothesise that surfactin is responsible for the direct inhibitory effect on *L. fusiformis*, as this was the only spent medium of an NRP mutant strain with lowered inhibition compared to spent media of other single NRP mutants.

Impact of surfactin on the growth of *L. fusiformis*

To confirm the inhibitory effect of surfactin on *L. fusiformis*, we exposed this strain to different concentrations of pure surfactin dissolved in methanol and monitored its growth over 24 h. The growth of *L. fusiformis* was delayed in the exponential growth phase when surfactin was supplemented in concentrations between 31.25 µg/mL and 500 µg/mL (Figure 8). At a surfactin concentration of 500 µg/mL, the cell density in the stationary phase was lower than the control. At a concentration of 250 µg/mL, the cell density reached a level similar to the untreated control. However, when surfactin was added in concentrations between 125 and 31.25 µg/mL, after an initial growth delay into the exponential phase, the cell densities in all treatments exceeded the ones of the control. The highest concentration of the solvent methanol of 5% had only a minor inhibiting effect on *L. fusiformis*, whereas lower concentrations of methanol showed no inhibition (Figure S5, Supporting Information File 1). These results suggest that surfactin has growth inhibitory effects on *L. fusiformis*, and we hypothesise that it might act as the key inhibitory *B. subtilis* NRP under the tested conditions.
Discussion

*B. subtilis* is known to produce a wide range of different SMs that target a large number of various micro- and macroorganisms [35]. Our study demonstrates that the NRPs produced by the recently isolated environmental strain of *B. subtilis* P5_B1 did not strongly impact the overall soil-derived semisynthetic mock community but reduced the abundance of the genera *Lysinibacillus* and *Viridibacillus* (Figure 9). Moreover, it reveals that the strain *L. fusiformis* M5 was directly affected by the *B. subtilis* lipopeptide surfactin in a monitored growth experiment.

We studied the bacterial community compositions by sequencing the two variable regions V3 and V4 of the 16S rRNA gene. Noteworthy, some limitations of this technique are well known. In 2014, Poretsky et al. revealed that amplicon sequencing of the 16S rRNA gene indicates a lower sequence diversity and substantial differences in the relative abundances of specific genus-assigned taxa compared to metagenomics [55]. Moreover, 16S amplicon sequencing of single variable regions rarely allows sufficient discrimination below the family or genus level, and therefore intragenus differentiation and heterogeneity cannot be addressed [55]. Furthermore, the fundamental problem is that bacteria harbour various copy numbers of the 16S rRNA gene in the genomes, which biases quantification studies [56]. Alpha diversity analyses based on the Shannon estimation revealed that diversity was strongly reduced in in vitro cultivations. Furthermore, it was disclosed that the precultured soil suspension had the lowest diversity index because mainly the genera *Bacillus* and *Acinetobacter* were enriched, which can probably be traced back to different growth rates among the present species. A substantial shift in the community compositions was observed between in vivo and in vitro communities since the majority of the genera present in the in vitro communities was below the detection limit in the soil sample. However, during the 12 h precultivation of the soil suspension, bacteria were exposed to different nutrient availabilities, changed physical conditions, such as the temperature, a liquid environment, and the loss of the spatial soil structure. These conditions were most likely selecting for generalist bacteria capable of proliferating under the given conditions and independently from other bacteria. During the following 48 h cocultivation, depletion of the primary nutrient sources and metabolic cross-feeding further shaped the community assembly. In 2018, Goldford et al. revealed that the main sources of metabolic cross-feeding are secreted metabolic by-products from the community members [57]. They further highlighted that bacterial communities stabilised after approximately eight to nine 48 h cocultivations. In our study, bacterial communities were only cocultivated once for 48 h, suggesting that the assembly of the bacterial communities has not yet reached a stable phase, which explains the differences between the precultures and cocultivated mock communities.

The Shannon index showed no differences among the established and differently treated mock communities, which primarily consisted of 13 genera. Even though *Bacillus* was the most abundant genus in the precultures, further incubation for 48 h resulted in a decreased relative abundance independently if the respective *B. subtilis* strains were seeded or the precultures were untreated. It shows that the initial dominance of *Bacillus* could not be maintained at prolonged incubation. The *B. subtilis* strains were added at a community assembly phase when *Bacillus* was the dominating genus, so that the general genera distribution was not expected to be influenced extensively. Nevertheless, after 48 h cocultivation, the final relative abundance of the *Bacillus* genus was not increased in the communities treated with *B. subtilis* when compared to the control. This observation highlights that the presence or absence of NRPs did not affect the competitiveness of *B. subtilis*. However, the 16S amplicon sequencing did not allow the detection of interactions and competitions within the *Bacillus* genus. The composition of this genus could vary among the differently treated communities. Nonetheless, the beta diversity analysis indicated a dissimilarity between the untreated and treated mock communities. Besides, two of the communities treated with the *sfp* mutant showed the highest similarity to the untreated communities,
Figure 9: Overview on the biosynthetic pathways of surfactin (A), plipastatin (B), and bacillaene (C) produced by B. subtilis. The lightning bolt indicates the proteins for which the corresponding coding genes were deleted in the mutant strains.
sugest that the supplementation of the NRP-producing *B. subtilis* strains affected the communities. The vectors of *Acinetobacter* ASVs had a direction either to NRP-treated or NRP-untreated communities, indicating that the NRPs influenced species within the same genus differently.

In microbial communities, the amount of interactions and relations increases with the number of community members. The established semisynthetic mock communities in this study contained at least 13 genera with a relative abundance >0.19%. Therefore, it can be assumed that various interactions between them occurred. Nevertheless, we could observe statistically significant reductions of the two genera, *Lysinibacillus* and *Viridibacillus*, in communities supplemented with the NRP-producing *B. subtilis* wild type strain. In contrast, in communities supplemented with the NRP-deficient mutant *sfp*, *Lysinibacillus* was more frequent than in the wild type-treated communities. This observation indicates that NRPs have a greater impact on suppressing *Lysinibacillus*. However, further factors are involved in the suppression since the *sfp* mutant maintained a reduction of *Lysinibacillus*, even though to a weaker extent. Moreover, no particular NRP could be allocated to the inhibition of the *Lysinibacillus* genus in these semisynthetic communities, but bacillaene displayed the highest impact on the suppression. An inhibition of *Viridibacillus* mediated by NRPs was also observable, but for this genus, bacillaene had the lowest impact. However, these results must be interpreted with caution and need further investigations since *Viridibacillus* was one of the lowest abundant genera in the mock communities, and abundance calculations are sensitive to the depth of sequencing. Besides the suppression of *Lysinibacillus* and *Viridibacillus, Stenotrophomonas* was uniquely suppressed in the communities supplemented with the *sfp* mutant but not when the WT strain was added. This observation might be evoked by inhibiting other species, which in turn facilitates a lower inhibition of *Stenotrophomonas*.

Previous studies revealed that the introduction of SM-producing bacteria to a bacterial community had no major impact on the entire composition. The tropodithietic acid-producing marine bacterium *Phaeobacter inhibens* did not strongly influence the microbiome diversity of the oyster *Ostrea edulis* but reduced the relative abundance of the orders *Vibionales* and *Mycoplasmatales* [58]. Similar results were achieved when *B. velezensis* FZB42 was successfully applied as a biocontrol agent to lettuce in soil [59]. The authors could not see a substantial impact on the rhizosphere bacterial community by the supplemented biocontrol strain, whereas the sampling time and additional inoculation of the fungal plant pathogen influenced the community to a greater extent. Apart from soluble SM, volatile organic compounds (VOCs) are as well capable of impacting a microbial community. In 2020, Cosetta et al. demonstrated that VOCs of cheese rind-associated fungi have both growth-stimulating and -inhibiting properties on members of the rind microbiome [60]. The authors could reveal that the VOC-mediated shift of the bacterial community was caused due to growth promotion of *Vibrio* spp. These studies and the results from the semisynthetic mock community experiment of this study highlight that the overall impact of SMs on the targeted microbial communities is low, which suggests that they are no mass destruction compounds. However, in all communities, distinct genera or species were suppressed or promoted, emphasizing the potential of SMs to shape microbial communities.

To investigate if *Lysinibacillus* is sensitive to any particular NRP of *B. subtilis*, we exposed the isolate *L. fusiformis* M5 to the spent media of the respective *B. subtilis* strains and monitored the growth. *L. fusiformis* M5 has been isolated from soil and demonstrated to impact the biofilm colony development of *B. subtilis* [54]. Interestingly, the modulation of the biofilm development was mediated by the primary metabolite hypoxanthine secreted by *L. fusiformis*. Of note, the impact of *B. subtilis* was not noticed on *L. fusiformis* in the mixed colony biofilm communities, possibly due to the use of the NRP-negative *B. subtilis* strain 168, which harbours a spontaneous frameshift mutation in the *sfp* gene [54]. Testing the impact of the natural isolate *B. subtilis* P5_B1 and the corresponding NRP mutant derivatives revealed that the spent media from both the NRP-deficient strain *sfp* and the surfactin-deficient strain *srfAC* had the lowest impact on the growth of *L. fusiformis*. In addition, the spent media of ∆ppsC and ∆pksL maintained the bioactivity at low concentrations, whereas the effect of WT was already strongly reduced at this level of the spent medium. This difference could occur, on the one hand, due to higher levels of surfactin in the two mutants compared to the wild type. On the other hand, the spent medium originated from cultures with an OD600 value of 3.0. Cultures with higher ODs were diluted before the harvesting, and WT cultures exhibited overall the highest ODs among the strains. Since the NRPs concentration is not proportional to the final OD due to, e.g., the occurrence of cell lysis, the spent media might be slightly differently diluted among the strains. Therefore, minor differences might be observable in the assays supplemented with highly diluted spent media. The observation that *L. fusiformis* displays a slightly higher cell density when the bacterial spent medium is supplemented might be due to the availability of additional nutrients. Nevertheless, the supernatant and pure compound supplementation demonstrated that surfactin is a direct suppressor of *L. fusiformis*. However, as the spent media of the *sfp* and ∆*srfAC* strains still had a growth inhibition effect, it is plausible that next to surfactin, further NRPs and even other compounds might provoke a slight growth suppression of *Lysinibacillus*.  

![Image](https://example.com/image.png)
When \textit{L. fusiformis} was exposed to surfactin concentrations between 31.25 and 125 µg/mL, higher final cell densities were detectable compared to assays treated with higher levels of surfactin or in the control. Interestingly, in 2020, Arjes et al. demonstrated that surfactin enhances the availability of oxygen to \textit{B. subtilis} by increasing the oxygen diffusivity \cite{61}, which might also positively affect the growth of \textit{L. fusiformis}.

Experiments with differently treated semisynthetic mock communities have demonstrated that \textit{Lysinibacillus} and \textit{Viridibacillus} were affected by the addition of an NRPs-producing \textit{B. subtilis} strain. \textit{Lysinibacillus} was least affected in the mock communities supplemented with the \textit{B. subtilis} \textit{ΔpksL} strain incapable of producing bacillaene, suggesting that bacillaene is the most active compound against this genus. In contrast, the growth curve experiments showed that \textit{L. fusiformis} M5 is most sensitive to surfactin. Importantly, our analysis does not reveal which \textit{Lysinibacillus} species were present in the mock communities, and therefore their sensitivity might be different from the test species \textit{L. fusiformis} used. Moreover, the spent medium was harvested from pure cultures of \textit{B. subtilis} grown in an undiluted complex medium, which might have changed the production of NRPs due to the lacking impact of the community members and the level of nutrients. Thus, lower concentrations of the NRPs in the mock communities might affect \textit{Lysinibacillus} differently compared to the monoculture growth experiments supplemented with spent media. Finally, \textit{Lysinibacillus} can also be affected indirectly by \textit{B. subtilis} NRPs in the mock communities. Bacillaene is described as a wide-spectrum antibiotic disrupting the protein synthesis in bacteria \cite{34,52}. The observations suggest that it has the most substantial impact on specific members of the mock community, and consequently an indirect effect on \textit{Lysinibacillus}. Nevertheless, the exact mechanisms at play remain to be deciphered.

Interestingly, the two genera \textit{Lysinibacillus} and \textit{Viridibacillus} of the mock communities are, besides \textit{Paenibacillus}, the closest relatives of \textit{B. subtilis}. The fact that suppression effects are only observable for these genera could presumably be caused by the higher overlap in the ecological niches, triggering competition for the same nutrients. Indeed, a higher phylogenetic and metabolic similarity between bacteria increases the probability of antagonism \cite{62}.

We could not quantify the concentrations of \textit{B. subtilis} NRPs in the mock communities since the detection of low concentrations is still under development. However, a better understanding of their impact on the mock communities could be realised by further experiments investigating the effect of supplemented pure NRP compounds, e.g., surfactin and bacillaene. The impact of antibiotics on algae-associated bacterial communities was investigated by Geng et al. in 2016, who revealed a dose-dependent influence of pure tropodithietic acid on the microbiome structure of \textit{Nannochloropsis salina} \cite{63}. Such pure NRP supplementations in various concentrations would allow exploring their effects on bacterial community assembly. Furthermore, in vivo experiments could reveal the impact of NRPs on microbial communities in complex natural systems, similar to the study from Chowdhury et al. from 2013 \cite{59}. Noteworthy, our study focused only on NRPs, but additional SMs, such as bacteriocins, are predicted for \textit{B. subtilis} \textit{P5_B1} as well \cite{51}. Future investigations should investigate the impact of both bacteriocins and NRPs on microbial communities.

### Conclusion

In summary, this study demonstrates that nonribosomal peptides of \textit{B. subtilis} \textit{P5_B1} have only a minor impact on the overall structure of soil-derived semisynthetic bacterial mock communities but suppress the genera \textit{Lysinibacillus} and \textit{Viridibacillus} significantly. Furthermore, it highlights the bioactivity of surfactin against \textit{L. fusiformis} M5.

### Experimental

#### Strains, media, and chemicals

All strains used in this study are listed in Table S1. Supporting Information File 1. For routine growth, bacterial cells were cultured in tryptic soy broth (TSB, CASO Broth, Sigma-Aldrich) containing 17 g·L\(^{-1}\) casein peptone, 3 g·L\(^{-1}\) soy peptone, 5 g·L\(^{-1}\) sodium chloride, 2.5 g·L\(^{-1}\) dipotassium hydrogen phosphate, and 2.5 g·L\(^{-1}\) glucose.

#### Semisynthetic mock community assay

Semisynthetic soil communities were obtained from the soil of sampling site P5 (55.788800, 12.558300) \cite{51,64}. 1 g soil was mixed in a 1:9 ratio with a 0.9% saline solution, vortexed on a rotary shaker for 15 min, and allowed to sediment for 2 min. Four independent communities were established by inoculating 10-times diluted TSB (0.1 × TSB) with 1% soil suspension taken from the middle part of the liquid phase, followed by incubation at 21–23 °C and 250 rpm for 12 h. Simultaneously, pregrown \textit{B. subtilis} \textit{P5_B1} WT and the corresponding NRP mutant derivatives were inoculated in 0.1 × TSB and incubated in parallel using the same conditions. After 12 h precultivation, 3 mL aliquots of the soil suspension were transferred into six glass tubes. One tube was left untreated and functioned as control, whereas the remaining five were supplemented with respective \textit{B. subtilis} strains by adding 10% of the final volume. The cultures were incubated at 21–23 °C and 250 rpm for 48 h. DNA was extracted from two replicates of the initial soil sample, the 12 h precultivated soil suspensions and the
B. subtilis-treated or untreated mock communities cocultivated for 48 h.

DNA extraction

Environmental- and semisynthetic-community genomic DNA was extracted from either 250 mg soil or 250 µL bacterial culture, respectively, by using the DNeasy PowerSoil Pro Kit (QIAGEN) and following the manufacturer’s instructions.

Amplification of 16S rRNA hypervariable regions V3-V4

The V3-V4 region of the 16S rRNA gene was PCR-amplified from the extracted DNA samples using Fw_V3V4 (5’-CCTACGGGNGGCWGCAG-3’) and Rv_V3V4 (5’-GACTACHVGGGTATCTAATCC-3’) primers that were tagged with short barcodes with a length of eight nucleotides, listed in Table S2, Supporting Information File 1. The PCR reactions contained 10.6 µL DNase-free water, 12.5 µL TEMPase Hot Start 2x Master Mix, 0.8 µL of each primer (10 µM), and 0.3 µL of 50 ng/µL DNA template. The PCR was performed using the conditions of 95 °C for 15 min, followed by 30 cycles of 95 °C for 30 s, 62 °C for 30 s, 72 °C for 30 s, and finally, 72 °C for 5 min. All V3-V4 amplicons were purified using the NucleoSpin gel and PCR cleanup kit (Macherey-Nagel) and pooled in equimolar ratios. The amplicon pool was submitted to Novogene Europe Company Limited (United Kingdom) for high-throughput sequencing on an Illumina NovaSeq 6000 platform with 2 million reads (2 × 250 bp paired-end reads). Raw sequence data is available at NCBI: PRJNA658074.

Sequencing data preprocessing

The multiplexed sequencing data was imported into the QIIME 2 pipeline (version 2020.6) [65,66]. The paired-end sequences were demultiplexed with the QIIME 2 plugin cutadapt [67]. The minimum overlap of partial matches between the read and the barcode sequence was set to 5 nucleotides to reduce random matches. The QIIME 2 implementation DADA2 was used to denoise and merge paired-end reads [68]. In total, 362,475 reads were assigned to the respective samples with an average of 12,083 reads per sample (range: 751 to 34,802; Table S3, Supporting Information File 1). The 16S rRNA reference sequences with a 99% identity criterion obtained from the SILVA database release 132 were trimmed to the V3-V4 region, bound by the primer pair used for amplification, and the product length was limited to 200–500 nucleotides [69]. The taxonomy was assigned to the sequences in the feature table generated by DADA2 by using the VSEARCH-based consensus taxonomy classifier [70]. A tree for phylogenetic diversity analyses was generated with FastTree 2 from the representative sequences [71-73].

Relative species abundance and phylogenetic diversity analyses

QIIME 2 artefacts were imported into the R software (4.0.2) with the R package qiime2R, and further analyses were conducted in the R package phyloseq [74-76]. The taxonomy summaries were achieved by merging ASVs of the same genera and calculating their relative abundance in each sample. Differences in the presence of the most abundant genera in the control communities, in the communities supplemented with B. subtilis WT as well as in the communities supplemented with B. subtilis spf, were investigated by calculating the abundance ratios of the different treated communities for each replicate. If species were not detected in some of the replicates, 0 values were replaced with the lowest detected value of the genus to avoid infinite values or 0 values in the ratio calculations. Rarefaction curves of the samples were calculated and visualised with the R package ranacapa [77]. Diversity analyses of the B. subtilis-treated and untreated samples were performed with ASV counts multiplied by factor 100,000 and transformed into integer proportions. The alpha diversity was estimated with the Shannon diversity index in the R package phyloseq [76]. The beta diversity was determined by dissimilarities among the samples with the Bray–Curtis distance and visualised in a nMDS with the R package vegan [78]. The correlation of individual ASVs on the overall bacterial community composition was calculated with the envfit function with 999 permutations from the R package vegan. The most correlating (R² > 0.6) ASVs were added to the nMDS ordination plot. All graphical visualisations were realised with ggplot2 [79].

Statistical analysis

The statistical significance was determined with the square root of the tested values. The normality and equality of the variances were tested with the Shapiro–Wilk normality test and the Levene test, respectively. If one of the tests was rejected, the nonparametric Kruskal–Wallis rank sum test was performed instead. The statistical significance of pairs was determined with the Welch two-sample t-test, and the differences among groups >2 was determined with the one-way analysis of variance (ANOVA) test and the Tukey HSD test. The statistical significance was determined with an alpha level <0.05.

Growth monitoring of L. fusiformis supplemented with B. subtilis spent media and pure surfactin

Spent media of B. subtilis strains were harvested from cultures grown in TSB medium at 37 °C and 250 rpm for 48 h immediately before the growth experiments. The cultures were adjusted to OD₆₀₀ 3.0 and centrifuged for 4 min at 5,000g. Subsequently,
the supernatants were passed through 0.22 µm filters and stored at 4 °C. The growth experiments were performed in 96-well microplates. The wells of the first column were filled with 30 µL 10 × TSB, 30 µL *L. fusiformis* culture adjusted to OD$_{600}$ 0.1 in 1 × TSB, and 240 µL of the appropriate spent *B. subtilis* medium or water (untreated control). 100 µL *L. fusiformis* culture adjusted to OD$_{600}$ 0.01 in 1 × TSB was added to the wells of the remaining columns. A 1.5-fold serial dilution of the spent media was performed column-by-column. A surfactin stock solution was prepared by dissolving 10 mg of surfactin (Sigma-Aldrich) in 1 mL methanol (MeOH). The wells of the first column were filled with 170 µL 1 × TSB, 20 µL *L. fusiformis* culture adjusted to OD$_{600}$ 0.1 in 1 × TSB, and 10 µL surfactin, 10 µL MeOH (solvent control), or 10 µL 1 × TSB (untreated control). To the wells of the remaining columns, 100 µL *L. fusiformis* culture was added adjusted to OD$_{600}$ 0.01 in 1 × TSB. A 2-fold serial dilution of surfactin or MeOH was performed column-by-column. In both assays, the growth of *L. fusiformis* was monitored in a microplate reader (BioTek Synergy HTX Multi-Mode Microplate Reader). The microplates were incubated at 30 °C with continuous shaking (600 rpm). OD$_{600}$ was measured in 15 min intervals over 24 h. All graphical visualisations were prepared using ggplot2 [79].

**Supporting Information**

**Supporting Information File 1**

Bacterial strains used in this study, 16S rRNA V3-V4 primer list, number of sequencing reads per sample, and supporting figures.

[https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-16-248-S1.pdf](https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-16-248-S1.pdf)

**Acknowledgements**

The authors thank the suggestions of Lone Gram and the CeMiSt centre members on the project. A part of the graphical abstract was created using BioRender.com.

**Funding**

This project was supported by the Danish National Research Foundation (DNRF137) for the Center for Microbial Secondary Metabolites (CeMiSt).

**ORCID® iDs**

Heiko T. Kieseewalter - https://orcid.org/0000-0001-5966-5894
Carlos N. Lozano-Andrade - https://orcid.org/0000-0003-2805-4505
Mikael L. Strube - https://orcid.org/0000-0003-0905-5705
Ákos T. Kovács - https://orcid.org/0000-0002-4465-1636

Preprint

A non-peer-reviewed version of this article has been previously published as a preprint: https://doi.org/10.1101/2020.08.20.259788
Identification of volatiles from six marine *Celeribacter* strains

Anuj Kumar Chhalodia\(^1\), Jan Rinkel\(^1\), Dorota Konvalinkova\(^1\), Jörn Petersen\(^2\) and Jeroen S. Dickschat \(*^1\)

**Abstract**

The volatiles emitted from six marine *Rhodobacteraceae* species of the genus *Celeribacter* were investigated by GC–MS. Besides several known compounds including dimethyl trisulfide and \(\text{S}\)-methyl methanethiosulfonate, the sulfur-containing compounds ethyl (\(E\))-3-(methylsulfanyl)acrylate and 2-(methyldisulfanyl)benzothiazole were identified and their structures were verified by synthesis. Feeding experiments with \([\text{methyl}-{\text{\(^2\)}}\text{H\(_3\)}]{\text{methionine}}, [\text{methyl}-\text{\(^{13}\)}\text{C}]{\text{methionine}}\) and \([\text{34}\text{S}]{\text{3-(dimethylsulfonio)propanoate}}}\) (DMSP) resulted in the high incorporation into dimethyl trisulfide and \(\text{S}\)-methyl methanethiosulfonate, and revealed the origin of the methylsulfanyl group of 2-(methyldisulfanyl)benzothiazole from methionine or DMSP, while the biosynthetic origin of the benzothiazol-2-ylsulfanyl portion could not be traced. The heterocyclic moiety of this compound is likely of anthropogenic origin, because 2-mercaptobenzothiazole is used in the sulfur vulcanization of rubber. Also in none of the feeding experiments incorporation into ethyl (\(E\))-3-(methylsulfanyl)acrylate could be observed, questioning its bacterial origin. Our results demonstrate that the *Celeribacter* strains are capable of methionine and DMSP degradation to widespread sulfur volatiles, but the analysis of trace compounds in natural samples must be taken with care.

**Introduction**

Bacteria from the roseobacter group belong to the most abundant microbial species in marine ecosystems [1,2]. They are present from polar to tropical regions, in marine sediments, in estuarine and open ocean environments in different pelagic zones ranging from surface waters to depths of >2,000 m [3,4]. Some species are associated with other marine organisms, e.g., *Thalassococcus halodurans* DSM 26915\(^T\) has been isolated from the marine sponge *Halichondria panicea* [5], and *Phaeobacter gallaeciensis* DSM 26640\(^T\) is an isolate from the scallop *Pecten maximus* [6]. Important interactions are also observed between bacteria from the roseobacter group and various types of marine algae, e.g., the first described organisms
Roseobacter litoralis DSM 6996\textsuperscript{T} and R. denitrificans DSM 7001\textsuperscript{T} were obtained from seaweed [7], while Dinoroseobacter shibae DSM 16493\textsuperscript{T} and Marinovum algicola DSM 10251\textsuperscript{T} are both isolates from the dinoflagellate Prorocentrum lima [8,9]. Especially in algal blooms bacteria of the roseobacter group are highly abundant [10], and here they belong to the main players involved in the enzymatic degradation of the algal sulfur metabolite 3-(dimethylsulfinio)propanoate (DMSP, Scheme 1) [11]. Its catabolism leads either through the demethylation pathway by action of the enzymes DmdABCD to methanethiol (MeSH, Scheme 1A) [12] or through lysis by DddD [13] or hydrolytic cleavage by one of the known DMSP lyases (DddW [14], DddP [15], DddQ [16], DddL [17], DddY [18] or DddK [19]) to dimethyl sulfide (DMS, Scheme 1B).

It has already been pointed out in the 1970s and 1980s that atmospheric DMS is important for the global sulfur cycle [20] and influences the climate on Earth, known as CLAW hypothesis according to the authors’ initials (Carlson, Lovelock, Andreae, Warren) [21], which underpins the relevance of this algal–bacterial interaction. Isotopic labeling experiments demonstrated that also in laboratory cultures roseobacter group bacteria efficiently degrade DMSP into sulfur volatiles [22,23], but also from other sulfur sources including 2,3-dihydroxy-propane-1-sulfonic acid (DHPS, Scheme 1C) labeling was efficiently incorporated into sulfur volatiles [24,25]. Notably, DHPS is produced in large quantities by the marine diatom Thalassiosira pseudonana [26], and diatoms from this genus live in symbiotic relationship with bacteria of the roseobacter group [27]. Another interesting aspect of sulfur metabolism in marine bacteria from the roseobacter group is the production of the sulfur-containing antibiotic tropodithietic acid (TDA) in Phaeobacter piscinae DSM 103509\textsuperscript{T} [28], a compound that is in equilibrium with its tautomer thiotropocin [29] that was first described from Pseudomonas sp. CB-104 [30]. Its biosynthesis depends on the clustered tda genes [31] and has been studied by

---

**Scheme 1:** Sulfur metabolism in bacteria from the roseobacter group. A) DMSP demethylation by DmdABCD, B) DMSP hydrolysis by DddP and lysis by DddW, DddP, DddQ, DddL, DddY or DddK, and C) structures of DHPS and sulfur-containing secondary metabolites.
feeding experiments with labeled precursors to the wildtype and gene knockout strains of \textit{P. inhibens} DSM 17395\textsuperscript{T}, demonstrating the formation of TDA from phenylalanine through phenylacetly-CoA and the phenylacetly-CoA catabolon \cite{32,33}. These experiments also led to a suggestion for the mechanism for sulfur incorporation, but further research is required for a deep understanding of TDA biosynthesis. Besides its function as an antibiotic, TDA acts as a signaling molecule, similar to \textit{N-acylhomoserine lactones}, at concentrations 100 times lower than required for a significant antibiotic activity \cite{34}. The biosynthesis of tropane \cite{35} and of the algicidal sulfur-containing roseobacticides \cite{36} are most likely connected to the TDA pathway. Interestingly, in the interaction with marine algace \textit{P. inhibens} can change its lifestyle from a symbiotic relationship during which the antibiotic TDA and growth stimulants are produced to a pathogenic interaction promoted by lignin degradation products in fading algal blooms that induce roseobacticide biosynthesis \cite{36}. All these examples demonstrate the importance of sulfur metabolism for marine bacteria from the roseobacter group. Here we report on the volatiles emitted by six \textit{Celeribacter} species with a special focus on sulfur volatiles. The results from feeding studies with labeled precursors demonstrate that the \textit{Celeribacter} strains can form sulfur volatiles from methionine and DMSP, but also showed that some of the detected sulfur compounds are not or only partly of bacterial origin.

\textbf{Results and Discussion}

\textbf{Headspace analysis}

The volatiles released by six marine \textit{Celeribacter} type strains, including \textit{C. marinus} DSM 100036\textsuperscript{T}, \textit{C. neptuni} DSM 26471\textsuperscript{T}, \textit{C. manganoides} DSM 27541\textsuperscript{T}, \textit{C. baekdonensis} DSM 27375\textsuperscript{T}, \textit{C. halophilus} DSM 26270\textsuperscript{T} and \textit{C. indicus} DSM 27257\textsuperscript{T}, were collected through a closed-loop stripping apparatus (CLSA) on charcoal \cite{37}. After extraction with dichloromethane the obtained extracts were analyzed by GC–MS (Figure 1). The compounds were identified by the comparison of the recorded EI mass spectra to library spectra and of retention indices \cite{38} to tabulated literature data (Table 1), or by a direct comparison to authentic standards. The structures of the identified compounds are shown in Figure 2.

While the headspace extracts from \textit{C. marinus}, \textit{C. neptuni} and \textit{C. manganoides} were particularly rich, the extracts from \textit{C. baekdonensis}, \textit{C. halophilus} and \textit{C. indicus} contained fewer compounds. Most of the observed volatiles are well known \cite{39,40} and were thus readily identified from their mass spectra and retention indices. Pyrazines including methylpyrazine (1), 2,5-dimethylpyrazine (2) and trimethylpyrazine (3) were present in the extracts from all six strains. Notably, also several \textit{α-hydroxyketones} that have been described as biosynthetic pre-
cursors to pyrazines \cite{40}, represented by 3-hydroxypentan-2-one (4), 2-hydroxypentan-3-one (5) and 2-hydroxyhexan-3-one (6), were observed in some of the investigated strains. A series of aldehydes ranging from hexanal (7) to tetradecanal (13) was found in strain specific patterns, with all identified compounds present in the bouquet from \textit{C. manganoides}. A similar series of \textit{γ-lactones} spanning from pentan-4-olide (14) to dodecan-4-olide (20), in addition to 3-methylbutan-4-olide (21) and 4-methylhex-5-en-4-olide (22), was detected in strain-specific patterns, with almost all of these compounds present in \textit{C. marinus}; only \textit{C. halophilus} did not emit lactones. Furans included furan-2-ylmethanol (23), furfural (24), and 2-acetylfuran (25). Cyclohexanol (26) was observed only once in \textit{C. marinus}, and aromatic compounds included benzyl alcohol (27), benzaldehyde (28) and salicylaldehyde (29), acetoephone (30) and \textit{o-aminocacetophenone} (31), 2-phenylethanol (32), and phenylacetone (33). 6-Methylhept-5-en-2-one (34) was detected in all strains, while its saturated analog 6-methylheptan-2-one (35) was only emitted by \textit{C. baekdonensis} and geranylacetone (36) only by the three productive species \textit{C. marinus}, \textit{C. neptuni}, and \textit{C. manganoides}. Compounds 34 and 36 have been described as non-enzymatic degradation products arising from the side chain in menaquinones \cite{58}. Sulfur-containing compounds included dimethyl trisulfide (37), released by all six species, \textit{S-methyl methanethiosulfonate} (38), 2-acetyltiazole (39), and benzothiazole (40), the latter also in the extracts from all six strains. In addition, the extracts from the three species \textit{C. marinus}, \textit{C. neptuni} and \textit{C. baekdonensis} contained an additional volatile (41) whose mass spectrum (Figure 3A) was not included in our libraries. Furthermore, ethyl 3-(methylsulfanyl)acrylate (42) was found in \textit{C. marinus} and \textit{C. manganoides}, but the measured retention index (\textit{I} = 1177) did not allow to distinguish between the E and the Z isomer for which retention indices of \textit{I} = 1144 (E) and \textit{I} = 1158 (Z) were reported \cite{53}. Therefore, for an unambiguous structural assignment for compounds 41 and 42 the synthesis of reference compounds was required.

\textbf{Synthesis of reference compounds}

The mass spectrum of the component 41 showed strong similarities to the library mass spectrum of 2-mercaptobenzothiazole that has a molecular weight of 167 Da. The isotope pattern of the molecular ion at \textit{m/z} = 213 indicated the presence of three sulfur atoms. The strong base peak at \textit{m/z} = 167 in the mass spectrum of 41 suggested a benzothiazol-2-ylsulfanyl moiety, while the mass difference to the molecular ion pointed to the connection to a methylsulfanyl group. Taken together, this analysis resulted in the structural proposal of 2-(methylsulfanyl)benzothiazole for 41. For the structural verification a synthesis was performed by a BF\textsubscript{3}·OEt\textsubscript{2}-catalyzed reaction of bis(benzothiazol-2-yl)disulfane with dimethyl disulfide, giving
Figure 1: Total ion chromatograms of headspace extracts from A) _C. marinus_ DSM 100038T, B) _C. neptunius_ DSM 26471T, C) _C. manganoxidans_ DSM 27541T, D) _C. baekdonensis_ DSM 27375T, E) _C. halophilus_ DSM 26270T, and F) _C. indicus_ DSM 27257T. Peaks arising from known contaminants are indicated by asterisks.
Table 1: Volatiles from Celeribacter spp.

<table>
<thead>
<tr>
<th>Compounda</th>
<th>b</th>
<th>t(lit.)b</th>
<th>Id. c</th>
<th>Occurrenced</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-hydroxypentan-2-one (4)</td>
<td>812</td>
<td>815 [39]</td>
<td>ri, ms</td>
<td>C</td>
</tr>
<tr>
<td>hexanal (7)</td>
<td>813</td>
<td>806 [39]</td>
<td>ri, ms</td>
<td>B</td>
</tr>
<tr>
<td>2-hydroxypentan-3-one (5)</td>
<td>818</td>
<td>818 [40]</td>
<td>ri, ms</td>
<td>C</td>
</tr>
<tr>
<td>methylpyrazine (1)</td>
<td>831</td>
<td>826 [41]</td>
<td>ri, ms</td>
<td>A B C D E F</td>
</tr>
<tr>
<td>furfural (24)</td>
<td>841</td>
<td>841 [42]</td>
<td>ri, ms</td>
<td>B C</td>
</tr>
<tr>
<td>furan-2-ylmethanol (23)</td>
<td>861</td>
<td>863 [43]</td>
<td>ri, ms</td>
<td>A B C D E F</td>
</tr>
<tr>
<td>cyclohexanol (26)</td>
<td>888</td>
<td>886 [44]</td>
<td>ri, ms</td>
<td>A</td>
</tr>
<tr>
<td>2-hydroxyhexan-3-one (6)</td>
<td>899</td>
<td>900 [40]</td>
<td>ri, ms</td>
<td>A</td>
</tr>
<tr>
<td>heptanal (8)</td>
<td>906</td>
<td>901 [45]</td>
<td>ri, ms</td>
<td>C</td>
</tr>
<tr>
<td>2,5-dimethylpyrazine (2)</td>
<td>912</td>
<td>908 [45]</td>
<td>ri, ms</td>
<td>A B C D E F</td>
</tr>
<tr>
<td>2-acetylfurane (25)</td>
<td>913</td>
<td>909 [45]</td>
<td>ri, ms</td>
<td>C D</td>
</tr>
<tr>
<td>pentan-4-olide (14)</td>
<td>953</td>
<td>956 [46]</td>
<td>ri, ms</td>
<td>A D F</td>
</tr>
<tr>
<td>3-methylbutan-4-olide (21)</td>
<td>957</td>
<td>958 [47]</td>
<td>ri, ms</td>
<td>A F</td>
</tr>
<tr>
<td>6-methylheptan-2-one (35)</td>
<td>959</td>
<td>962 [48]</td>
<td>ri, ms</td>
<td>D</td>
</tr>
<tr>
<td>benzaldehyde (28)</td>
<td>961</td>
<td>952 [45]</td>
<td>ri, ms</td>
<td>A B C D E F</td>
</tr>
<tr>
<td>dimethyl trisulfide (37)</td>
<td>970</td>
<td>968 [49]</td>
<td>ri, ms</td>
<td>A B C D E F</td>
</tr>
<tr>
<td>6-methylhept-5-en-2-one (34)</td>
<td>988</td>
<td>981 [45]</td>
<td>ri, ms</td>
<td>A B C D E F</td>
</tr>
<tr>
<td>trimethylpyrazine (3)</td>
<td>1000</td>
<td>1000 [45]</td>
<td>ri, ms</td>
<td>A B C D E F</td>
</tr>
<tr>
<td>2-acetylthiazole (39)</td>
<td>1017</td>
<td>1014 [45]</td>
<td>ri, ms</td>
<td>A B C D</td>
</tr>
<tr>
<td>benzyl alcohol (27)</td>
<td>1033</td>
<td>1026 [45]</td>
<td>ri, ms</td>
<td>C</td>
</tr>
<tr>
<td>4-methylhex-5-en-4-olide (22)</td>
<td>1039</td>
<td>1034 [45]</td>
<td>ri, ms</td>
<td>B C</td>
</tr>
<tr>
<td>salicylaldehyde (29)</td>
<td>1042</td>
<td>1039 [45]</td>
<td>ri, ms</td>
<td>B D</td>
</tr>
<tr>
<td>hexan-4-olide (15)</td>
<td>1052</td>
<td>1056 [50]</td>
<td>ri, ms</td>
<td>A</td>
</tr>
<tr>
<td>S-methyl methanethiosulfonate (38)</td>
<td>1061</td>
<td>1068 [51]</td>
<td>ri, ms</td>
<td>A B C D F</td>
</tr>
<tr>
<td>acetophenone (30)</td>
<td>1065</td>
<td>1059 [45]</td>
<td>ri, ms</td>
<td>A B C</td>
</tr>
<tr>
<td>nonanal (9)</td>
<td>1103</td>
<td>1100 [45]</td>
<td>ri, ms</td>
<td>A B C</td>
</tr>
<tr>
<td>2-phenylethanol (32)</td>
<td>1111</td>
<td>1106 [45]</td>
<td>ri, ms</td>
<td>B C</td>
</tr>
<tr>
<td>phenylacetone (33)</td>
<td>1127</td>
<td>1124 [52]</td>
<td>ri, ms</td>
<td>A</td>
</tr>
<tr>
<td>ethyl (E)-3-(methylsulfonyl)acrylate (42)</td>
<td>1177</td>
<td>1144 [53]</td>
<td>ms</td>
<td>A</td>
</tr>
<tr>
<td>decanal (10)</td>
<td>1203</td>
<td>1201 [45]</td>
<td>ri, ms</td>
<td>A B C F</td>
</tr>
<tr>
<td>benzothiazole (40)</td>
<td>1221</td>
<td>1222 [54]</td>
<td>ri, ms</td>
<td>A B C D E F</td>
</tr>
<tr>
<td>octan-4-olide (16)</td>
<td>1252</td>
<td>1250 [45]</td>
<td>ri, ms</td>
<td>A</td>
</tr>
<tr>
<td>o-aminoacetophenone (31)</td>
<td>1292</td>
<td>1296 [55]</td>
<td>ri, ms</td>
<td>A</td>
</tr>
<tr>
<td>undecan-11</td>
<td>1298</td>
<td>1305 [45]</td>
<td>ri, ms</td>
<td>B C</td>
</tr>
<tr>
<td>nonan-4-olide (17)</td>
<td>1354</td>
<td>1358 [45]</td>
<td>ri, ms</td>
<td>A C</td>
</tr>
<tr>
<td>dodecan-12</td>
<td>1400</td>
<td>1408 [45]</td>
<td>ri, ms</td>
<td>A B C</td>
</tr>
<tr>
<td>geranylacetone (36)</td>
<td>1445</td>
<td>1453 [45]</td>
<td>ri, ms</td>
<td>A B C</td>
</tr>
<tr>
<td>decan-4-olide (18)</td>
<td>1461</td>
<td>1465 [45]</td>
<td>ri, ms</td>
<td>A C</td>
</tr>
<tr>
<td>undecan-4-olide (19)</td>
<td>1568</td>
<td>1569 [45]</td>
<td>ri, ms</td>
<td>C</td>
</tr>
<tr>
<td>tetradecanal (13)</td>
<td>1605</td>
<td>1611 [45]</td>
<td>ri, ms</td>
<td>B C</td>
</tr>
<tr>
<td>dodecan-4-olide (20)</td>
<td>1673</td>
<td>1676 [45]</td>
<td>ri, ms</td>
<td>A C</td>
</tr>
<tr>
<td>2-(methylsulfonyl)benzothiazole (41)</td>
<td>1860</td>
<td></td>
<td>std</td>
<td>A B D</td>
</tr>
</tbody>
</table>

aIdentified by GC–MS, known typical contaminants such as plasticizers are not included and all listed compounds were not detected in blank runs with medium plates (except traces of benzaldehyde); bretention index on a HP5-MS GC column and comparison to literature data from the same or a similar type of GC column; cidentification based on ri: matching retention index (difference between measured retention index and literature data ≤10 points); ms: mass spectrum matching to a database spectrum, std: direct comparison to an authentic standard; doccurrence in A: C. marinus DSM 100036T, B: C. neptunius DSM 26471T, C: C. manganoxidans DSM 27541T, D: C. baekdonensis DSM 27375T, E: C. halophilus DSM 26270T, and F: C. indicus DSM 27257T.
Figure 2: Structures of the identified volatile compounds in the headspace extracts from six Celeribacter type strains.

Figure 3: EI mass spectra of A) unlabeled 2-(methylsulfonyl)benzothiazole (41) and of labeled 41 after feeding of B) (methyl-2H₃)methionine, C) (methyl-13C)methionine and D) (34S)DMSP.
access to 41 with a yield of 64% (Scheme 2). The synthetic compound 41 showed an identical mass spectrum and retention index compared to the volatile in the Celeribacter extracts. The Z and E stereoisomers of 42 were obtained by the Michael addition of NaSMe to ethyl propiolate (45), yielding a mixture of stereoisomers inseparable by silica gel column chromatography (92%). The major stereoisomer was found to be (Z)-42 (dr 94:6), whose preferred formation may be a result of a chalcogen–chalcogen interaction between the sulfur and an ester oxygen. This phenomenon was first described in supramolecular structures by Gleiter [59] and later also used to explain the outcome of organocatalytic reactions [60]. The pure stereoisomers of 42 were isolated by preparative HPLC, for which the best separation was achieved using a YMC ChiralART Cellu-
lose-SC column. This yielded 70% of (Z)-42 and 6% of (E)-42, and their analysis by GC–MS showed retention indices of \( I = 1177 \) for (E)-42 and \( I = 1200 \) for (Z)-42, revealing that the compound in the headspace extracts of C. marinus DSM 100036\(^T\) and C. manganoxidans DSM 27541\(^T\) was identical to (E)-42.

Feeding experiments with isotopically labeled precursors

The biosynthesis of sulfur volatiles in C. marinus was investigated in a series of feeding experiments with isotopically labeled precursors. Feeding of (methyl-\(^3\)C)methionine resulted in the efficient incorporation of labeling into 37 (79% incorporation rate, Figure S1I in Supporting Information File 1), 38 (78%, Figure S1J in Supporting Information File 1) and the S-methyl group of 41 (84%), as indicated by a shift of the molecular ion from \( m/z = 213 \) to 216 (Figure 3B, deuterated compounds can be separated from their non-deuterated analogs by gas chromatography [61]). The base peak appears at \( m/z = 168 \), demonstrating its formation with participation of one deuterium from the S-methyl group. Analogous results were obtained by feeding of (methyl-\(^1\)\(^3\)C)methionine, showing incorporation into 37 (74%, Figure S1C in Supporting Information File 1), 38 (71%, Figure S1G in Supporting Information), and the MeS group of 41 (71%, Figure 3C; the signal at \( m/z = 213 \) represents inseparable unlabeled 41 that, in contrast to a deuterated compound, cannot be separated from \(^3\)C-labeled 41 by gas chromatography). Furthermore, feeding of [\(^3\)H]DMSP gave an incorporation into the MeS groups of 37 (50%, Figure S1D), into both sulfur atoms of 38 (47%, Figure S1H in Supporting Information File 1), but only into one sulfur atom of 41 (46%), as indicated by the molecular ion at \( m/z = 215 \), while no signals at \( m/z = 217 \) and 219 were visible that would account for the incorporation of labeling into two or three of the sulfur atoms in 41 (Figure 3D; also here the signal at \( m/z = 213 \) represents inseparable unlabeled 41). In this experiment, the base peak did not change which allowed the localization of labeling specifically in the MeS group of 41.

The fact that no incorporation was observed for the other two sulfur atoms of 41 prompted us to further investigate the biosynthetic origin of the benzo-thiazol-2-ylsulfanyl portion of 41 to establish its natural origin. Several feeding experiments with central primary metabolites including (\(^1\)\(^3\)C)\(^5\)ribose and (indole-\(^2\)H\(_2\)jtryptophan were performed, but none of these experiments resulted in a detectable incorporation of labeling. Conclusively, a non-biological origin of this part of the molecule seems likely, which may also explain why the detection of 41 in Celeribacter was not always reproducible. Notably, 2-mercaptobenzothiazole is used in the sulfur vulcan-
ization of rubber and could react spontaneously with MeSH of bacterial origin in the presence of oxygen to form 41, giving a reasonable explanation for its formation.

Also none of the feeding experiments with the various labeled precursors resulted in an incorporation of labeling into the

Scheme 2: Synthesis of sulfur-containing compounds detected in the Celeribacter headspace extracts. A) Synthesis of 2-(methyl disulfanyl)benzothiazole (41) and B) synthesis of ethyl (Z)- and (E)-3-(methylsulfanyl)acrylate (42).
sulfur volatiles 39, 40, and 42, which also questioned their natural origin. This finding is rather surprising for 42, especially regarding the feeding experiment with \(^{34}S\)DMSP, because its formation would be explainable by a DMSP degradation through the demethylation pathway, for which all relevant enzymes are encoded in the six Celeribacter strains (only a DmdA homolog is missing in C. indicus, Table S1 in Supporting Information File 1), and e.g., transesterification of the DmdC product with EtOH (Scheme 1A). Compound 42 is not a widespread sulfur volatile, but has been reported before from pineapples [53], pears [62], passion fruits [63], and apples [64].

Conclusion

Six marine Celeribacter strains were investigated for their volatiles, leading to the identification of 42 compounds from different classes, including several sulfur volatiles. However, feeding experiments with isotopically labeled precursors suggested that only the widespread compounds dimethyl trisulfide (37) and \(\alpha\)-methyl methanethiosulfonate (38) are of natural origin, while no labeling from any of the fed precursors was incorporated into 2-acetyltiazole (39), benzothiazole (40), and ethyl \((E)\)-3-(methylsulfanyl)acrylate (42), thus questioning their natural source from Celeribacter. These results demonstrate that the six Celeribacter strains are able to degrade methionine and DMSP with formation of MeSH as a source for the likely non-enzymatic oxidation in the presence of air to 37 and 38, opening possibilities for future studies on methionine and DMSP degrading enzymes and pathways in Celeribacter. Our study also shows that the results from trace compound analyses must be taken with care and contaminations from other sources must always be taken into consideration. For the unusual compound 2-(methylsulfanyl)benzothiazole (41) the incorporation of labeling was observed only into the MeS group, while the benzothiazol-2-ylsulfanyl portion is likely of anthropogenic origin from the rubber vulcanization agent 2-mercaptobenzothiazol-2-ylsulfanyl portion is likely of anthropogenic origin. This finding is rather surprising for 42, especially regarding the feeding experiment with \(^{34}S\)DMSP, because its formation would be explainable by a DMSP degradation through the demethylation pathway, for which all relevant enzymes are encoded in the six Celeribacter strains (only a DmdA homolog is missing in C. indicus, Table S1 in Supporting Information File 1), and e.g., transesterification of the DmdC product with EtOH (Scheme 1A). Compound 42 is not a widespread sulfur volatile, but has been reported before from pineapples [53], pears [62], passion fruits [63], and apples [64].

General synthetic and analytical methods

Reactions were carried out in oven-dried flasks under Ar atmosphere and using distilled and dried solvents. Chemicals were obtained from Sigma-Aldrich (St. Louis, USA). Column chromatography was performed on silica gel (0.04–0.06 nm) purchased from Acros Organics (Geel, Belgium) with distilled solvents. NMR spectroscopy was performed on a Bruker (Billerica, USA) Avance III HD Ascend (500 MHz) spectrometer. Solvent peaks were used for referencing (\(^1\)H NMR: CDCl\(_3\) δ = 7.26 ppm, \(^{13}\)C NMR: CDCl\(_3\) δ = 77.16 ppm) [65]. Multiplicities are indicated by s (singlet) and d (doublet), coupling constants \(J\) are given in Hz. IR spectra were recorded on a Bruker (Karlsruhe, Germany) equipped with a UV–vis detector MWL 2.1L (deuterium lamp, 190–700 nm) and a YMC ChiralART Cellulose-SC column (5 \(\mu\)m; 250 × 20 mm) with a guard column of the same type (30 × 20 mm). The elution was performed with hexane/propanol 60:40 (isocratic) at a flow rate of 10 mL min\(^{-1}\) (36 bar). The UV–vis absorption was monitored at 275 nm.

Experimental

Strains, culture conditions, and feeding experiments

All six Celeribacter type strains were cultivated at 28 °C on marine broth agar plates. In case of feeding experiments, the isotopically labeled compound (1 mM) was added to the agar medium before inoculation.

Collection of volatiles

The volatiles emitted by Celeribacter spp. agar plate cultures were collected on charcoal filters (Chromtech, Idstein, Germany), precision charcoal filters charged with 5 mg of charcoal) by use of a closed-loop stripping apparatus as developed by Grob and Zürcher [37]. After a collection time of 24 h the charcoal was extracted with CH\(_2\)Cl\(_2\) (50 \(\mu\)L) and the extract was analyzed by GC–MS.

GC–MS

GC–MS analyses were carried out through a 7890B GC – 5977A MD system (Agilent, Santa Clara, CA, USA). The GC was equipped with a HP5-MS fused silica capillary column (30 m, 0.25 mm i.d., 0.50 \(\mu\)m film) and operated with the settings 1) inlet pressure: 77.1 kPa, He flow: 23.3 mL min\(^{-1}\), 2) injection volume: 2 \(\mu\)L, 3) splitless injection, 4) temperature program: 5 min isothermal at 50 °C, then increasing with 5 °C min\(^{-1}\) to 320 °C, and 5) He carrier gas flow: 1.2 mL min\(^{-1}\). The parameters of the MS were 1) transfer line temperature: 250 °C, 2) ion source temperature: 230 °C, 3) quadrupole temperature: 150 °C, and 4) electron energy: 70 eV. Retention indices were calculated from retention times in comparison to those of a homologous series of \(n\)-alkanes (C\(_7\)–C\(_{40}\)).

Synthesis of 2-(methylsulfanyl)benzothiazole (41)

1,2-Bis(benzothiazol-2-yl)disulfane (43, 1.00 g, 3.00 mmol, 1 equiv) and dimethyl sulfide (44, 0.28 g, 3.00 mmol, 1 equiv) were dissolved in dry CH\(_2\)NO\(_2\) (10 mL) and dry CH\(_2\)Cl\(_2\) (10 mL). The solution was cooled to 0 °C and then treated with BF\(_3\)·Et\(_2\)O (43 mg, 0.3 mmol, 0.1 equiv). After stirring at 0 °C for 3 hours and at room temperature overnight, the reaction was


427
quenched by the addition of water (10 mL) and extracted with ethyl acetate (3 × 50 mL). The combined extracts were dried with MgSO₄ and concentrated. The residue was purified by column chromatography (cyclohexane/ethyl acetate 1:1) to give 41 as a colorless solid (0.82 g, 3.85 mmol, 64%). Rₜ 0.60 (cyclohexane/ethyl acetate 5:1; TLC visualized with UV illumination at 366 nm); GC (HP-5MS): I = 1854; IR (diamond-ATR) v: 3060 (s), 2916 (s), 1425 (w), 1310 (s), 1236 (s), 1005 (w), 756 (w), 431 (s) cm⁻¹; ¹H NMR (500 MHz, CDCl₃, 298 K) δ 7.88 (dd, J = 8.1, 1.2, 0.7 Hz, 1H, CH), 7.87 (dd, J = 7.9, 1.2, 0.6 Hz, 1H, CH), 7.43 (dd, J = 8.3, 7.3, 1.2 Hz, 1H, CH), 7.33 (dd, J = 8.2, 7.2, 1.2 Hz, 1H, CH), 2.67 (s, 3H, CH₃) ppm; ¹³C NMR (125 MHz, CDCl₃, 298 K) δ 172.50 (C), 155.17 (C), 144.44 (CH), 23.62 (CH₃), 14.67 (CH₃) ppm.

Synthesis of ethyl (Z)-3-(methylsulfanyl)acrylate ((Z)-42) and ethyl (E)-3-(methylsulfanyl)acrylate ((E)-42)

Ethyl propiolate (45, 70 mg, 0.71 mmol, 1 equiv) was dissolved in distilled water (5 mL) followed by the addition of sodium methanethiolate (50 mg, 0.71 mmol, 1 equiv). The solution was stirred for 30 minutes at room temperature. Water (5 mL) was added and the product was extracted with ethyl acetate (3 × 10 mL). The combined extracts were dried over MgSO₄ and concentrated to afford the crude product. Purification by column chromatography (cyclohexane/ethyl acetate 99:1) gave a mixture of stereoisomers (Z)-42 and (E)-42 as pale yellow oil (96 mg, 0.65 mmol, 92%, dr 94:6 by H NMR). The product mixture was separated by preparative HPLC to give pure (Z)-42 (73 mg, 0.50 mmol, 70%) and (E)-42 (6 mg, 0.04 mmol, 6%).

(Z)-42. Rₜ 0.74 (cyclohexane/ethyl acetate 1:1); GC (HP-5MS): I = 1200; IR (diamond-ATR) v: 2982 (w), 2927 (w),1695 (m), 1569 (m), 1434 (w), 1374 (w), 1300 (w), 1266 (w), 1213 (m), 1166 (s), 1095 (w), 1033 (w), 986 (w), 961 (w), 800 (w), 727 (w), 687 (w) cm⁻¹; ¹H NMR (700 MHz, CDCl₃, 298 K) δ 7.04 (d, J = 10.14 Hz, 1H, CH), 5.83 (d, J = 10.14 Hz, 1H, CH), 4.20 (q, J = 7.15 Hz, 2H, CH₂), 2.39 (s, 3H, CH₃), 1.29 (t, J = 7.17 Hz, 3H, CH₃) ppm; ¹³C NMR (175 MHz, CDCl₃, 298 K) δ 157.8 (s), 144.4 (CH), 143.4 (CH), 135.9 (C), 126.3 (CH), 124.7 (CH), 122.2 (CH), 121.2 (CH), 23.6 (CH₃) ppm.

(E)-42. Rₜ 0.76 (cyclohexane/ethyl acetate 1:1); GC (HP-5MS): I = 1177; IR (diamond-ATR) v: 2980 (w), 2925 (w), 1701 (s), 1578 (s), 1444 (w), 1366 (w), 1322 (w), 1297 (m), 1251 (s), 1161 (s), 1095 (w), 1037 (m), 945 (m), 886 (w), 832 (w), 799 (w), 702 (w) cm⁻¹; ¹H NMR (700 MHz, CDCl₃, 298 K) δ 7.76 (d, J = 14.93 Hz, 1H, CH), 5.68 (d, J = 14.90 Hz, 1H, CH), 4.21 (q, J = 7.14 Hz, 2H, CH₂), 2.35 (s, 3H, CH₃), 1.31 (t, J = 7.13 Hz, 3H, CH₃) ppm; ¹³C NMR (175 MHz, CDCl₃, 297 K) δ 165.59 (C), 147.21 (CH),113.56 (CH), 60.55 (CH₂), 27.26 (CH₃), 14.67 (CH₃) ppm.

Supporting Information
Supporting Information File 1
DMSP demethylation pathway in Celeribacter spp. and copies of spectra.

Acknowledgements
We thank Andreas Schneider (Bonn) for HPLC separation of (E)- and (Z)-42.

Funding
This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 34509606 – TRR 51 within the frame of the Transregional Collaborative Research Center “Roseobacter”.

ORCID® IDs
Jeroen S. Dickschat - https://orcid.org/0000-0002-0102-0631

References
   doi:10.1002/anie.201805383
59. Werz, D. B.; Staebe, T. H.; Benisch, C.; Rausch, B. J.; Rominger, F.; 
60. Leverett, C. A.; Purohit, V. C.; Romo, D. Angew. Chem., Int. Ed. 2010, 
   49, 9479–9483. doi:10.1002/anie.201004671
   doi:10.1039/c3np70080a
   40, 1925–1929. doi:10.1021/jf00022a040
63. Werkhoff, P.; Güntert, M.; Krammer, G.; Sommer, H.; Kaulen, J. 
   Nudelman, A.; Stoltz, B. M.; Bercaw, J. E.; Goldberg, K. I. 
   Organometallics 2010, 29, 2176–2179. doi:10.1021/om100106e

License and Terms

This is an Open Access article under the terms of the 
Creative Commons Attribution License 
(https://creativecommons.org/licenses/by/4.0). Please note 
that the reuse, redistribution and reproduction in particular 
requires that the author(s) and source are credited and that 
individual graphics may be subject to special legal 
provisions.

The license is subject to the Beilstein Journal of Organic 
Chemistry terms and conditions: 
(https://www.beilstein-journals.org/bjoc/terms)

The definitive version of this article is the electronic one 
which can be found at: 
https://doi.org/10.3762/bjoc.17.38
Breakdown of 3-(allylsulfonio)propanoates in bacteria from the *Roseobacter* group yields garlic oil constituents

Anuj Kumar Chhalodia and Jeroen S. Dickschat*

Full Research Paper

**Abstract**

Two analogues of 3-(dimethylsulfonio)propanoate (DMSP), 3-(diallylsulfonio)propanoate (DAllSP), and 3-(allylmethylsulfonio)propanoate (AllMSP), were synthesized and fed to marine bacteria from the *Roseobacter* clade. These bacteria are able to degrade DMSP into dimethyl sulfide and methanethiol. The DMSP analogues were also degraded, resulting in the release of allylated sulfur volatiles known from garlic. For unknown compounds, structural suggestions were made based on their mass spectrometric fragmentation pattern and confirmed by the synthesis of reference compounds. The results of the feeding experiments allowed to conclude on the substrate tolerance of DMSP degrading enzymes in marine bacteria.

**Introduction**

The name of the allyl group has been introduced by Wertheim in 1844 when he investigated the constituents of garlic oil and derives from the botanical name of garlic (*Allium sativum*) [1]. During that time, the structures of the garlic oil constituents and also of the allyl group remained unknown, but its formula was correctly assigned as C$_3$H$_5$. Five decades later, Semmler reported on the nature of allyl propyl disulfide (1), diallyl disulfide (2), diallyl trisulfide (3), and diallyl tetrasulfide (4) from garlic oil (Scheme 1A) [2]. The antibacterial principle in garlic was identified in 1944 by Cavallito et al. as allicin (5) [3], a formal oxidation product of disulfide 2. Not only 5, but also several other sulfur compounds from garlic are today known to exhibit diverse biological activities, including inter alia antibacterial, antifungal, antioxidant, anti-inflammatory, and anticancer effects [4]. Later on, also heterocyclic compounds including 2-vinyl-1,3-dithiine (6) and 3-vinyl-3,4-dihydro-1,2-dithiine (7) were discovered [5]. The formation of these volatile sulfur compounds starts from alliin (9) [6], a non-volatile precursor that is stored in garlic and related plants and only degraded into sulfur volatiles upon wounding by the pyridoxal phosphate (PLP) dependent alliinase (Scheme 1B) [7]. This initial enzyme-catalyzed reaction yields one equivalent of allyl-
Scheme 1: Volatile allyl sulfides. A) Compounds known from garlic oil, B) mechanism of formation from alliin (9) by the PLP-dependent allinase (PLP: pyridoxalphosphate) and subsequent spontaneous reactions.

sulfenic acid (10), pyruvic acid (11), and ammonia from 9, followed by a series of proposed spontaneous reactions [5,8]. Through these transformations, acid 10 can undergo a dimerization with elimination of water to allicin (5). The hydrolysis of 5 results in allylsulfinic acid (12) and allyl thiol (13), the latter of which can react with another molecule of 5 to yield 10 and 2. Alternatively, 5 can decompose to 10 and thioacroleine (14) by a Cope elimination, which explains the formation of the heterocycles 6 and 7 by dimerization through a [4 + 2] cycloaddition [5]. Compounds 6 and 7 were also reported to be formed from 5 during gas chromatographic (GC) analysis by an unknown mechanism [9] (7 was confused with its double bond regiosomer 3-vinyl-3,6-dihydro-1,2-dithione (8) in this study [5]). Under these conditions the formation of the heterocyclic disulfides 7 and 8 may not involve a dimerization of 14, as a [4 + 2] cycloaddition is not a preferred gas-phase reaction.

The ecology of marine bacteria in their interaction with algae is particularly interesting in which the bacteria can promote the algal growth, but can also kill their host [10,11]. For both processes, the phytohormone indole-3-acetic acid is used as a messenger molecule [10]. For the macroalga Ulva mutabilis the presence of bacteria from the Roseobacter group is even mandatory for proper algal development, and 3-(dimethylsulfonyl)propanoate (DMSP) is used as a chemotactic signal by the bacteria attracting them towards the algal host [12]. Many bacteria and fungi also release sulfur volatiles [13,14] that are especially important headspace constituents from marine bacteria of the Roseobacter group [15-17]. In these organisms, sulfur volatiles are to a large extent generated from algal (DMSP), a metabolite that is produced in massive amounts by algae [18], thus giving another example for the complex interactions between marine bacteria and algae. Known DMSP degradation pathways include its hydrolysis to dimethyl sulfide.
(DMS) and 3-hydroxypropanoic acid (15) by the enzyme DddD [19], or the lysis to DMS and acrylic acid (16) for which various enzymes including DddL [20], DddP [21], DddQ [22], DddY [23], DddW [24], and DddK [25] have been described (Scheme 2A). Furthermore, a demethylation pathway is known through which DMSP is first converted into methylmercapto-propanoic acid (17) by the tetrahydrofolate (FH4)-dependent demethylase, DmdA (Scheme 2B) [26]. Compound 17 can be transformed into the coenzyme A thioester 18 by the CoA ligase Dmdb, followed by FAD-dependent oxidation to the α,β-unsaturated compound 19 by DmdC. The attack of water to the Michael acceptor catalyzed by the enoyl-CoA hydratase DmdD yields the hemithioacetal 20 that spontaneously collapses to methanethiol (MeSH) and malonyl-CoA semialdehyde (21). This compound further degrades to acetaldehyde (22) through the thioester hydrolysis and decarboxylation [27].

Feeding of (methyl-2H₆)DMSP to Phaeobacter inhibens DSM 17395 and Rügeria pomeroyi DSM 15171 resulted in the efficient uptake of labelling into dimethyl disulfide (DMDS), the oxidative dimerization product from MeSH, showing the activity of the demethylation pathway in these bacteria. However, knockout of the dmdA gene in R. pomeroyi still gave a low incorporation of labelling into DMDS, suggesting the presence of another gene responsible for the demethylation activity [28]. Also the labelling from (34S)DMSP was efficiently incorporated into DMDS and dimethyl trisulfide (DMTS) [29]. Our previous investigations have also demonstrated that synthetic, i.e., non-natural DMSP analogues such as 3-(ethylmethyl)sulfoniopropanoate (EMSP), 3-(diethylsulfonio)propanoate (DESP), 3-(dimethylselenio)propanoate (DMSnP; this compound is also formed naturally in Spartina alterniflora in the presence of sodium selenate [30]), and even 3-(dimethyltellurio)propanoate (DMTeP) are converted by the demethylation pathway into ethanethiol, methaneselenol, and methanetellurol, respectively, that further react to various volatiles containing EtS, MeSe, and MeTe groups [31]. The in vitro incubations of these DMSP analogues with recombinant DddQ and DddW from R. pomeroyi and DddP from P. inhibens demonstrated that all substrate analogues can be degraded through the lysis pathway into the corresponding dialkyl chalcogenides; only DMTeP was not cleaved by DddQ [32]. Here we describe the synthesis of the new DMSP analogues 3-(allylmethylsulfonio)propanoate (AllMSP) and 3-(diallylsulfonio)propanoate (DAllSP) and their

Scheme 2: Degradation of DMSP by marine bacteria. A) Hydrolysis or lysis to DMS, B) demethylation pathway leading to MeSH. The color code shows which enzymes are encoded in the genomes of the strains investigated in this study.
conversion into typical garlic odor constituents by marine bacteria from the *Roseobacter* group that do not naturally occur in these organisms.

**Results and Discussion**

3-(Diallylsulfonio)propanoate (DAIIISP) and 3-(allylmethylsulfonio)propanoate (AllMSP) were synthesized by the acid-catalyzed addition of allyl methyl sulfide and diallyl sulfide, respectively, to acrylic acid (Scheme 3). The obtained DMSP analogues were fed to marine broth agar plate cultures of three strains from the *Roseobacter* group with fully sequenced genomes, including *P. inhibens* DSM 17395, *Dinoroseobacter shibae* DSM 16493, and *Oceanibulbus indolifex* DSM 14862. In all cases the bacterial cultures released a strong garlic-like odor, presumptively due to a degradation of the DMSP derivatives to sulfur-containing volatiles, similar to the compounds known from garlic, through one of the pathways shown in Scheme 2. The emitted volatiles were captured on charcoal filter traps using a closed-loop stripping apparatus (CLSA) [33], followed by the extraction of the filters with CH$_2$Cl$_2$ and analysis by gas chromatography–mass spectrometry (GC–MS) of the resulting extracts. Most of the compounds were readily identified by the comparison of their mass spectra and retention indices to published data. Every experiment was performed in triplicate to check for the reproducibility of the results. For comparison, the volatiles from all three strains grown on marine broth medium without the addition of DMSP or its analogues have been reported before [31].

Feeding of DAIIISP to *P. inhibens* resulted in the production of sulfur volatiles including several allyl derivatives (Figure 1, Figure 2A, Table 1, and Figure S1 in Supporting Information File 1). Besides the methylated sulfur compounds dimethyl trisulfide (31), dimethyl tetrasulfide (33), and S-methyl methanethiosulfonate (28) that were reported previously from *P. inhibens* [31], large amounts of diallyl sulfide (29) were observed, pointing to an efficient degradation of DAIIISP through the lysis pathway, for which the DMSP lyase DddP can account in this organism (Scheme 2). Furthermore, the compounds allyl methyl disulfide (30), diallyl disulfide (2), allyl methyl trisulfide (32), and traces of diallyl trisulfide (3) and allyl methyl tetrasulfide (34) were observed. The formation of these compounds is explainable by the deallylation of DAIIISP to 3-(allylsulfanyl)propanoic acid (37) and further degradation to allyl thiol (13) through the enzymes of the demethylation pathway that is fully established in *P. inhibens* by genes coding for DmdA–D (Scheme 4A). In the presence of air thiol 13 can then undergo an oxidative dimerization, or react analogously with MeSH to form allyl methyl disulfide (30, Scheme 4B). Similar oxidations requiring one additional unit of hydrogen sulfide can lead to the trisulfides 3 and 32 (Scheme 4C), while higher polysulfides such as 34 can arise through a metathesis reaction of two trisulfides (Scheme 4D). Also traces of methyl 3-(allylsulfanyl)propanoate (24), methyl 3-(methyldisulfanyl)propanoate (25), and methyl 3-(allyldisulfanyl)propanoate (26) were observed. While the presence of 24 can be explained by the O-methylation of the DmdA product 37 with S-adenosylmethionine (SAM, Scheme 4E), compounds 25 and 26 require a second deallylation of 37 to 3-mercapto propanoic acid (38) possibly by DmdA, the reaction with a corresponding thiol MeSH or 13, and O-methylation (Scheme 4F).
Figure 2: Total ion chromatograms of CLSA extracts obtained from feeding experiments with DAiISP fed to A) *P. inhibens*, B) *D. shibae*, and C) *O. indolifex*. Numbers at peaks refer to compounds in Figure 1. Peaks without numbers are unidentified.

Table 1: Volatiles from agar plate cultures fed with DAiISP.

<table>
<thead>
<tr>
<th>Compounda</th>
<th>( I )</th>
<th>( I_L^{b} )</th>
<th>*P. in.*c</th>
<th>*D. sh.*c</th>
<th>*O. in.*c</th>
</tr>
</thead>
<tbody>
<tr>
<td>diallyl sulfide (29)*</td>
<td>849</td>
<td>848 [34]</td>
<td>⬢ ⬢ ⬢</td>
<td>⬢ ⬢ ⬢</td>
<td>⬢ ⬢ ⬢</td>
</tr>
<tr>
<td>allyl methyl disulfide (30)</td>
<td>910</td>
<td>912 [34]</td>
<td>⬢ ⬢ ⬢</td>
<td>⬢ ⬢ ⬢</td>
<td>⬢ ⬢ ⬢</td>
</tr>
<tr>
<td>dimethyl trisulfide (31)*</td>
<td>967</td>
<td>970 [35]</td>
<td>⬢ ⬢ ⬢</td>
<td>⬢ ⬢ ⬢</td>
<td>⬢ ⬢ ⬢</td>
</tr>
<tr>
<td>S-methyl methanethiosulfonate (28)*</td>
<td>1063</td>
<td>1068 [35]</td>
<td>⬢ ⬢ ⬢ ⬢</td>
<td>⬢ ⬢ ⬢ ⬢</td>
<td>⬢ ⬢ ⬢ ⬢</td>
</tr>
<tr>
<td>diallyl disulfide (2)*</td>
<td>1074</td>
<td>1075 [34]</td>
<td>⬢ ⬢ ⬢ ⬢</td>
<td>⬢ ⬢ ⬢ ⬢</td>
<td>⬢ ⬢ ⬢ ⬢</td>
</tr>
<tr>
<td>allyl methyl trisulfide (32)</td>
<td>1136</td>
<td>1133 [36]</td>
<td>⬢ ⬢ ⬢ ⬢</td>
<td>⬢ ⬢ ⬢ ⬢</td>
<td>⬢ ⬢ ⬢ ⬢</td>
</tr>
<tr>
<td>methyl 3-(allylsulfanyl)-propanoate (24)</td>
<td>1177</td>
<td>–</td>
<td>⬢ ⬢ ⬢ ⬢</td>
<td>⬢ ⬢ ⬢ ⬢</td>
<td>⬢ ⬢ ⬢ ⬢</td>
</tr>
<tr>
<td>dimethyl tetrasulfide (33)</td>
<td>1216</td>
<td>1215 [37]</td>
<td>⬢ ⬢ ⬢ ⬢</td>
<td>⬢ ⬢ ⬢ ⬢</td>
<td>⬢ ⬢ ⬢ ⬢</td>
</tr>
<tr>
<td>methyl 3-(methyldisulfanyl)-propanoate (25)*</td>
<td>1236</td>
<td>–</td>
<td>⬢ ⬢ ⬢ ⬢</td>
<td>⬢ ⬢ ⬢ ⬢</td>
<td>⬢ ⬢ ⬢ ⬢</td>
</tr>
<tr>
<td>diallyl trisulfide (3)</td>
<td>1300</td>
<td>1300 [38]</td>
<td>⬢ ⬢ ⬢ ⬢</td>
<td>⬢ ⬢ ⬢ ⬢</td>
<td>⬢ ⬢ ⬢ ⬢</td>
</tr>
<tr>
<td>allyl methyl tetrasulfide (34)</td>
<td>1382</td>
<td>1371 [39]</td>
<td>⬢ ⬢ ⬢ ⬢</td>
<td>⬢ ⬢ ⬢ ⬢</td>
<td>⬢ ⬢ ⬢ ⬢</td>
</tr>
<tr>
<td>methyl 3-(allyldisulfanyl)-propanoate (26)*</td>
<td>1397</td>
<td>–</td>
<td>⬢ ⬢ ⬢ ⬢</td>
<td>⬢ ⬢ ⬢ ⬢</td>
<td>⬢ ⬢ ⬢ ⬢</td>
</tr>
<tr>
<td>diallyl tetrasulfide (35)</td>
<td>1551</td>
<td>1540 [38]</td>
<td>⬢ ⬢ ⬢ ⬢</td>
<td>⬢ ⬢ ⬢ ⬢</td>
<td>⬢ ⬢ ⬢ ⬢</td>
</tr>
</tbody>
</table>

*aAsterisks indicate the identity to a commercially available or synthetic reference standard. bRetention index literature data for a HP5-MS or a similar GC column. cAbbreviations are *P. in.* = *Phaeobacter inhibens*, *D. sh.* = *Dinoroseobacter shibae*, and *O. in.* = *Oceanibulbus indolifex*. Filled circles indicate the presence, non-filled circles indicate the absence of a compound in the headspace extract. The colors of the circles refer to the chromatograms in Figure 2 and Figure S1–S3 in Supporting Information File 1 with the same color.
Very similar patterns of volatiles were obtained in the feeding experiments of DAISP with \textit{D. shibae} and \textit{O. indolifex} (Figure 2B,C, Table 1 and Figures S2 and S3 in Supporting Information File 1). An additionally observed compound in one analysis of \textit{O. indolifex} was diallyl tetrasulfide (35). Both organisms also encode the DMSP demethylation pathway in their genomes, but with missing \textit{dmdD} genes in both cases. A possible explanation is, that another enoyl-CoA hydratase, e.g., from fatty acid degradation, may functionally substitute for DmdD. \textit{Dinoroseobacter shibae} additionally encodes genes for the DMSP hydrolase DddD and the DMSP lyase DddL, explaining the formation of 29, while no DMSP hydrolase or
lyase is found in O. indolifex. Still, compound 29 is observed within this organism, but in lower quantities than in P. inhibens or D. shibae, and may point to the presence of another, yet unidentified type of DMSP lyase in this organism, because control experiments with medium plates with DAllSP added did not show a spontaneous degradation to 29 that could explain its observation.

The compound identification was based on a comparison to an authentic standard or of mass spectra to data base spectra in our MS libraries and confirmed for most cases by comparison of the retention indices to literature data, only for the mass spectrum of 26 no data base hit was returned. Therefore, a structural suggestion for this compound was based on the observed fragmentation pattern of the mass spectrum (Figure 3A). The molecular ion together with its isotope pattern pointed to two sulfur atoms, while the fragment ion at m/z = 64 ([S₂]+) pointed to a disulfide. The fragment ions at m/z = 59 ([C₂O₂H₃]+) and 161 ([M − OMe]+) indicated a methyl ester, and the series of m/z = 105 ([C₃H₅S₂]+), 73 ([C₃H₃S]+), and 41 ([C₃H₅]+) suggested an allyl disulfide. Taken together, the structure of methyl 3-(allyldisulfanyl)propanoate was delineated for compound 26 that was further supported by additional fragmentations as shown in Figure 3A. In addition, compound 26 was synthesized by a method reported previously for the related compound 25 [40], through dimerization of methyl 3-mercaptopropanoate (39) to dimethyl 3,3’-disulfanediylpropanoate (40), followed by the BF₃·OEt₂-mediated metathesis with 2 (Scheme 5A). The synthetic compound 26 was identical by mass spectrum and retention index to the unknown volatile.

The feeding of AllMSP to P. inhibens resulted in the formation of large amounts of methyl 3-(methylsulfanyl)propanoate (23) in addition to smaller quantities of methyl 3-(allylsulfanyl)propanoate (24, Figure 4A, Table 2 and Figure S4 in Supporting Information File 1). While compound 23 can arise from AllMSP by deallylation to 3-(methylsulfanyl)propanoic acid (36), potentially through DmdA, and O-methylation, the deriva-
Scheme 5: Synthesis of A) methyl 3-(allylsulfanyl)propanoate (26) and B) methyl 3-(methylsulfonyl)propanoate (27).

Figure 4: Total ion chromatograms of CLSA extracts obtained from the feeding experiments with AllMSP fed to A) P. inhibens, B) D. shibae, and C) O. indolifex. Numbers at peaks refer to compounds in Figure 1. Peaks without numbers are unidentified.
tive 24 may be formed analogously through intermediate 37 (Scheme 4A and E). The higher production of 23 in comparison to 24 suggests that the deallylation of AllMSP is more efficient than its demethylation, which is surprising, because naturally DmdA catalyzes a methyl-group transfer. This finding may reflect the high reactivity of the allyl group towards nucleophiles. Other compounds originating from AllMSP included the di- and trisulfides 2, 26, 30, and 32 that pointed to a breakdown of AllMSP to 13 through the DMSP demethylation pathway and subsequent oxidative polysulfide formation (Scheme 4A–C), but their formation was lower than from DAllSP, likely because of the discussed efficient deallylation of AllMSP. Small amounts of diallyl sulfide (29) were also detected, which is the formal lysis product of DAllSP, but not of AllMSP. In first instance, its formation from AllMSP was surprising, but it is explainable by a degradation of AllMSP to 13, followed by a nucleophilic attack at the allyl group of another AllMSP molecule (Scheme 4G). For D. shibae and O. indolifex the same pattern of compounds was found (Figure 4B,C and Figures S5 and S6 in Supporting Information File 1), only the production of the deallylated compound 23 was lower, while in turn the production of the di- and trisulfides from 13 and of 29 was increased. This suggests that the deallylation of AllMSP by the DmdA variants in these organisms may be less efficient than was observed for P. inhibens. Besides these sulfur compounds, only O. indolifex, but not the other two strains, released another compound, 27, whose mass spectrum was not included in our databases. The analysis of the fragmentation pattern (Figure 3B) suggested that 27 could be methyl 3-(methylsulfonyl)propanoate, an oxidation product of 23. This hypothesis was confirmed by the chemical oxidation of 23, yielding methyl 3-(methylsulfonyl)propanoate with an identical mass spectrum and retention index to the volatile 27 (Scheme 5B). This compound may arise from 23 by the action of an oxygenase that is restricted to O. indolifex and not encoded in the genomes of the other two species. Its spontaneous formation from 23 in the presence of air can be excluded, because other cultures forming 23 did not show the release of 27.

Conclusion

Bacteria from the Roseobacter group can degrade DMSP analogues with 5-allyl groups including AllMSP and DAllSP, likely with the participation of the enzymes for DMSP (hydro)lysis and from the DMSP demethylation pathway. Because MeSH can also originate from other sources, the DMSP derivatives used in this study can lead to products that can indicate which metabolic pathways are used for their conversion. Interestingly, the volatiles formed from AllMSP and DAllSP closely resemble flavoring compounds from garlic. The demethylation pathway with all four enzymes DmdABCD and not MeSH can indicate which metabolic pathways are used for their conversion. Interestingly, the volatiles formed from AllMSP and DAllSP closely resemble flavoring compounds from garlic.

### Table 2: Volatiles from agar plate cultures fed with AllMSP.

<table>
<thead>
<tr>
<th>Compound</th>
<th>I</th>
<th>I&lt;sub&gt;lit&lt;/sub&gt;&lt;sup&gt;b&lt;/sup&gt;</th>
<th>P. in.&lt;sup&gt;c&lt;/sup&gt;</th>
<th>D. sh.&lt;sup&gt;c&lt;/sup&gt;</th>
<th>O. in.&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>diallyl sulfide (29)*</td>
<td>849</td>
<td>848 [34]</td>
<td>●●●●</td>
<td>●●●●</td>
<td>●●●●</td>
</tr>
<tr>
<td>allyl methyl disulfide (30)</td>
<td>910</td>
<td>912 [34]</td>
<td>●●●●●●●●</td>
<td>●●●●●●●●</td>
<td>●●●●●●●●</td>
</tr>
<tr>
<td>dimethyl trisulfide (31)*</td>
<td>967</td>
<td>970 [35]</td>
<td>○○○○</td>
<td>○○○○</td>
<td>○○○○</td>
</tr>
<tr>
<td>methyl 3-(methylsulfonyl)-propanoate (23)*</td>
<td>1020</td>
<td>1023 [41]</td>
<td>●●●●●●●●</td>
<td>●●●●●●●●</td>
<td>●●●●●●●●</td>
</tr>
<tr>
<td>S-methyl methanethiosulfonate (28)*</td>
<td>1063</td>
<td>1068 [35]</td>
<td>●●●●●●●●</td>
<td>○○○○</td>
<td>○○○○</td>
</tr>
<tr>
<td>diallyl disulfide (2)*</td>
<td>1074</td>
<td>1075 [34]</td>
<td>●●●●●●●●</td>
<td>○○○○</td>
<td>○○○○</td>
</tr>
<tr>
<td>allyl methyl trisulfide (32)</td>
<td>1136</td>
<td>1133 [36]</td>
<td>●●●●●●●●</td>
<td>●●●●●●●●</td>
<td>●●●●●●●●</td>
</tr>
<tr>
<td>methyl 3-(allylsulfanyl)propanoate (24)</td>
<td>1177</td>
<td>–</td>
<td>●●●●●●●●</td>
<td>●●●●●●●●</td>
<td>●●●●●●●●</td>
</tr>
<tr>
<td>dimethyl tetrasulfide (33)</td>
<td>1216</td>
<td>1215 [37]</td>
<td>○○○○</td>
<td>○○○○</td>
<td>○○○○</td>
</tr>
<tr>
<td>methyl 3-(methylidisulfanyl)-propanoate (25)*</td>
<td>1236</td>
<td>–</td>
<td>●●●●●●●●</td>
<td>○○○○</td>
<td>○○○○</td>
</tr>
<tr>
<td>diallyl trisulfide (3)</td>
<td>1300</td>
<td>1300 [38]</td>
<td>●●●●●●●●</td>
<td>●●●●●●●●</td>
<td>●●●●●●●●</td>
</tr>
<tr>
<td>methyl 3-(methylsulfonyl)propanoate (27)*</td>
<td>1353</td>
<td>–</td>
<td>○○○○</td>
<td>○○○○</td>
<td>○○○○</td>
</tr>
<tr>
<td>methyl 3-(allylsulfanyl)propanoate (26)*</td>
<td>1397</td>
<td>–</td>
<td>●●●●●●●●</td>
<td>●●●●●●●●</td>
<td>●●●●●●●●</td>
</tr>
</tbody>
</table>

*a Asterisks indicate the identity to a commercially available or synthetic reference standard. b Retention index literature data for a HP5-MS or a similar GC column. c Abbreviations are P. in. = Phaeobacter inhibens, D. sh. = Dinoroseobacter shibae, and O. in. = Oceanibulbus indolifex. Filled circles indicate the presence, non-filled circles indicate the absence of a compound in the headspace extract. The colors of the circles refer to the chromatograms in Figure 4 and Figures S4–S6 in Supporting Information File 1 with the same color.
DAIISP into diallyl sulfide, while the reason for its formation in *O. indolifex* is currently unclear and may point to an unknown type of DMSP lyase in this organism. Since the observed patterns of allylated sulfur volatiles in the three investigated strains are different, it seems possible that the DMSP (hydro)lases and the enzymes from the DMSP demethylation pathway have different activities towards AllMSP and DAIISP. In vitro studies with recombinant purified enzymes and mutational work will be needed for more detailed insights to support our hypotheses regarding the involved enzymes in AllMSP and DAIISP breakdown and will be performed in our laboratories in the future.

### Experimental

**Strains and culture condition**

*Phaeobacter inhibens* DSM 14862, *Dinoroseobacter shibae* DSM 16493, *Oceanibulbus indolifex* DSM 14862 were precultured in full strength marine broth medium (MB 2216, Roth) at 28 °C with shaking at 180 rpm until the OD value reached about 1.0.

**Feeding experiments and sampling of volatiles**

Headspace samplings for each strain were done in triplicates. For the feeding experiments, DAIISP or AllMSP (1 mM) were added to the full strength marine broth agar medium (MB2216) after autoclaving. The medium was then transferred into glass Petri dishes. The agar plates were inoculated with the precultures (400 μL), incubated for two days at 28 °C and then subjected for headspace extraction to a CLSA [33] for 24 h. The released volatiles were collected on charcoal filters (Chromtech, Idstein, Germany), followed by the extraction of the filters with dichloromethane (50 mL), and analysis of the extracts by GC–MS.

**GC–MS**

The GC–MS analyses were carried out on a HP7890A GC system connected to a HP5975C mass selective detector fitted with a HP-SMS fused silica capillary column (30 m × 0.22 mm i.d., 0.25 μm film, Hewlett-Packard). The conditions were: inlet pressure: 67 kPa, He 23.3 mL min⁻¹; injection volume: 1 μL; injector: 250 °C; transfer line: 300 °C; electron energy: 70 eV. The GC was programmed as follows: 50 °C (5 min isothermal), increasing at 5 °C min⁻¹ to 320 °C and operated in the splitless mode (60 s valve time); carrier gas (He): 1.2 mL min⁻¹. The retention indices were determined from *n*-alkane standards (C₈–C₃₂) [42].

**General synthetic methods**

All chemicals were purchased from TCI (Deutschland) or Sigma-Aldrich Chemie (Germany), and used without purification. Solvents were distilled and dried by standard methods. NMR spectra were recorded on a Bruker (Billerica, USA) Avance III HD Prodigy (500 MHz) or on an Avance III HD Cryo (700 MHz) NMR spectrometer. The spectra were referenced against solvent signals (¹H NMR, residual proton signal: D₂O δ = 4.79 ppm, CDC₁₃ δ = 7.26 ppm, d₆-DMSO δ = 2.50 ppm; ¹³C NMR: CDC₁₃ δ = 77.16 ppm, d₆-DMSO δ = 39.52 ppm). The coupling constants are given in Hz. IR spectra were recorded on a Bruker α spectrometer equipped with a diamond-ATR probe. The relative intensities of signals are indicated by w (weak), m (medium), and s (strong).

**Synthesis of allyl DMSP derivatives**

A mixture of acrylic acid (0.72 g, 10 mmol) and diallyl sulfide or allylmethyl sulfide (10 mmol) was treated with 2 N HCl at 80 °C for 4 h. The reaction mixture was concentrated in vacuo and the residue was purified by silica gel column chromatography (CH₂Cl₂/Methanol 5:1), followed by recrystallization from methanol/diethyl ether 1:1 to yield the pure compounds.

**DAIISP·HCl.** Yield: 960 mg (4.32 mmol, 43%). ¹H NMR (D₂O, 700 MHz) δ 5.98 (ddt, δ = 17.5, 10.2, 7.4, 2H), 5.73 (d, δ = 10.2, 2H), 5.72 (d, δ = 17.0, 2H), 4.08 (d, δ = 7.4, 4H), 3.43 (t, J = 6.9, 2H), 2.78 (t, J = 6.9, 2H); ¹³C NMR (D₂O, 175 MHz) δ 177.05 (C), 127.65 (2 × CH), 123.54 (2 × CH₂), 41.53 (2 × CH₂), 35.08 (CH₂), 31.68 (CH₂); HRMS–EI (m/z): calcd for [C₈H₁₅S₂]⁺, 187.0787; found, 187.0790.

**AllMSP·HCl.** Yield: 1.06 g (5.41 mmol, 54%). ¹H NMR (D₂O, 700 MHz) δ 5.96 (ddt, δ = 17.5, 10.2, 7.5, 1H), 5.74 (d, δ = 10.2, 1H), 5.72 (J = 17.2, 1H), 5.71 (d, δ = 17.2, 1H), 4.13 (dd, δ = 13.4, 7.4, 1H), 4.09 (dd, δ = 13.4, 7.5, 1H), 3.58 (dt, δ = 13.7, 6.9, 1H), 3.47 (dt, δ = 13.5, 6.7, 1H), 3.04 (t, δ = 6.8, 2H), 2.91 (s, 3H); ¹³C NMR (D₂O, 175 MHz) δ 173.77 (C), 128.19 (2 × CH), 122.74 (2 × CH₂), 43.82 (CH₂), 35.84 (CH₂), 28.75 (CH₂), 21.72 (CH₃); HRMS–EI (m/z): calcd for [C₈H₁₅O₂S]⁺, 161.0631; found, 161.0630.

**Synthesis of dimethyl 3,3’-disulfanediyldioproanoate (40)**

A solution of methyl 3-mercaptopropanoate (6.00 g, 50.0 mmol, 1.0 equiv) and triethylamine (5.05 g, 50.0 mmol) in DMF (10 mL) was treated for 24 h at 40 °C. The reaction was quenched by the addition of water and the aqueous phase extracted with ethyl acetate. The extract was dried with MgSO₄ and then concentrated in vacuo. The residue was purified by silica column chromatography (cyclohexane/EtOAc 5:1) to give compound 40 (1.80 g, 7.56 mmol, 30%) as pale yellow oil. TLC
Rf 0.44 (cyclohexane/EtOAc 10:3); IR (diamond-ATR) ν: 2998 (w), 2952 (w), 2845 (w), 2256 (w), 1730 (m), 1436 (w), 1354 (w), 1240 (w), 1215(w), 1195 (w), 1171 (w), 1139 (w), 1046 (w), 1017 (w), 979 (w), 907 (w), 822 (w), 726 (m), 648 (w), 435 (w) cm⁻¹; ¹H NMR (CDCl₃, 500 MHz) δ 3.64 (6H, 3H); 2.87 (t, J = 7.2, 4H), 2.68 (t, J = 7.2, 4H) ppm; ¹³C NMR (CDCl₃, 125 MHz) δ 72.11 (2 × C), 51.90 (2 × CH₃), 33.93 (2 × CH₂), 33.16 (2 × CH₂) ppm.

Synthesis of methyl 3-(allyldisulfanyl)propanoate (26)

To a solution of dimethyl 3,3’-disulfoanidlypropanoate (40, 0.50 g, 2.10 mmol, 1.0 equiv) and diallyl disulfide (2, 0.31 g, 2.10 mmol, 1.0 equiv) in dry DCM (10 mL) and CH₃NO₂ (10 mL) at 0°C BF₃·OEt₂ (30 mg, 0.21 mmol, 0.1 equiv) was added. The reaction mixture was stirred at 0°C for 3 h and at room temperature overnight. The reaction was quenched by the addition of water and extracted with ethyl acetate. The extracts were dried with MgSO₄ and concentrated in vacuo. The obtained residue was purified by silica gel column chromatography (cyclohexane/EtOAc 5:1) to give compound 26 (0.23 g, 1.20 mmol, 57%). TLC Rf = 0.72 (cyclohexane/EtOAc = 1:1); IR (diamond-ATR) ν: 3082 (w), 2950 (w), 2845 (w), 1736 (s), 1634 (w), 1435 (w), 1354 (w), 1277 (w), 1240 (w), 1126 (w), 1117 (w), 1144 (w), 1017 (w), 985 (w), 922 (w), 859 (w), 820 (w), 756 (w), 722 (w), 669 (w), 582 (w), 478 (w), 435 (w) cm⁻¹; ¹H NMR (CDCl₃, 500 MHz) δ 5.83 (ddt, J = 17.1, 9.9, 7.3, 1H), 5.19 (ddt, J = 16.9, 1.3, 1.3, 1H), 5.14 (dm, J = 10.0, 1H), 3.69 (s, 3H), 3.32 (dm, J = 7.3, 2H), 2.91 (t, J = 7.2, 2H), 2.72 (t, J = 7.2, 2H) ppm; ¹³C NMR (CDCl₃, 125 MHz) δ 172.14 (C), 132.71 (CH), 119.40 (CH₂), 52.04 (CH₂), 41.60 (CH₂), 33.87 (CH₂), 33.40 (CH₂) ppm; HRMS–El (m/z): calc for [C₇H₁₂O₅S₂]⁺: 566.0142; found: 566.0132.

Synthesis of methyl 3-(methylsulfonyl)propanoate (27)

To a stirred solution of [(n-C₄H₉)₄N]₄(MoO₄)₂ (5 mg, 2.5 μmol, 0.001 equiv) [43] in methanol (4 mL), methyl 3-methylthiopropanoate (335 mg, 2.50 mmol, 1.0 equiv) was added at 40°C. After the reaction mixture was stirred for 5 minutes, 30% hydrogen peroxide solution (0.52 mL, 0.57 g, 5.0 mmol, 2.0 equiv) was added dropwise. The color of the reaction mixture changed from colorless to yellow. The reaction mixture was filtered and concentrated in vacuo to give pure 27 (0.34 g, 2.05 mmol, 82%) as colorless solid. TLC Rf 0.17 (cyclohexane/EtOAc 1:1); IR (diamond-ATR) ν: 3014 (w), 2948 (w), 2932 (w), 1762 (m), 1687 (w), 1442 (w), 1433 (w), 1418 (w), 1375 (w), 1331 (w), 1306 (m), 1373 (m), 1259 (m), 1203 (w), 1180 (w), 1131 (m), 1056 (w), 1004 (w), 988 (w), 971 (w), 956 (w), 989 (w), 786 (w), 774 (w), 749 (w), 601 (w), 514 (w), 505 (w), 441 (w) cm⁻¹; ¹H NMR (d6-DMSO, 500 MHz) δ 3.63 (3H, 3H), 3.37 (t, J = 7.5, 2H), 3.01 (s, 3H), 2.78 (t, J = 7.5, 2H) ppm; ¹³C NMR (d6-DMSO, 125 MHz) δ 170.79 (C), 51.88 (CH₃), 49.14 (CH₂), 40.21 (CH₃), 26.89 (CH₂) ppm.

References

A new glance at the chemosphere of macroalgal–bacterial interactions: In situ profiling of metabolites in symbiosis by mass spectrometry

Marine Vallet¹, Filip Kaftan², Veit Grabe³, Fatemeh Ghaderiardakani⁴, Simona Fenizia⁴,⁵, Aleš Svatоš², Georg Pohnert¹,⁴,⁶ and Thomas Wichard*⁴

Abstract

Symbiosis is a dominant form of life that has been observed numerous times in marine ecosystems. For example, macroalgae coexist with bacteria that produce factors that promote algal growth and morphogenesis. The green macroalga *Ulva mutabilis* (Chlorophyta) develops into a callus-like phenotype in the absence of its essential bacterial symbionts *Roseovarius* sp. MS2 and *Maribacter* sp. MS6. Spatially resolved studies are required to understand symbiont interactions at the microscale level. Therefore, we used mass spectrometry profiling and imaging techniques with high spatial resolution and sensitivity to gain a new perspective on the mutualistic interactions between bacteria and macroalgae. Using atmospheric pressure scanning microprobe matrix-assisted laser desorption/ionisation high-resolution mass spectrometry (AP-SMALDI-HRMS), low-molecular-weight polar compounds were identified by comparative metabolomics in the chemosphere of *Ulva*. Choline (2-hydroxy-N,N,N-trimethylethanol-1-aminium) was only determined in the alga grown under axenic conditions, whereas ectoine (1,4,5,6-tetrahydro-2-methyl-4-pyrimidinecarboxylic acid) was found in bacterial presence. Ectoine was used as a metabolic marker for localisation studies of *Roseovarius* sp. Within the tripartite community because it was produced exclusively by these bacteria. By combining confocal laser scanning microscopy (cLSM) and AP-SMALDI-HRMS, we proved that *Roseovarius* sp. MS2 settled mainly in the rhizoidal zone (holdfast) of *U. mutabilis*. Our findings provide the fundament to decipher bacterial symbioses with multicellular hosts in aquatic ecosystems.
in an ecologically relevant context. As a versatile tool for microbiome research, the combined AP-SMALDI and cLSM imaging analysis with a resolution to level of a single bacterial cell can be easily applied to other microbial consortia and their hosts. The novelty of this contribution is the use of an in situ setup designed to avoid all types of external contamination and interferences while resolving spatial distributions of metabolites and identifying specific symbiotic bacteria.

Introduction

In intertidal zones with high temporal and spatial ecosystem variations, bacteria and macroalgae establish close mutualistic relationships, in which both gain reciprocal benefits forming an ecological unit (holobiont) [1-3]. Chemical exchange and physical proximity are the basis of this algae–bacterial mutualism [4], but little is known about the spatial distribution of the bacteria on the algal host and the locally released and exchanged compounds within the algal chemosphere [3]. Bacterial biofilms on macroalgae can be crucial for developing algae and their interactions with other marine organisms. The exchange of resources in this spatially limited region is of high interest for understanding the macroalgal–bacterial interactions. The chemosphere was proposed as a region that supports chemical mediator-based cross-kingdom interactions [3]. High-throughput sequencing analysis provides the abundance and composition of the bacterial community on macroalgal surfaces [5,6]. It does not reveal any information on the metabolically active bacteria and the spatial distribution of substances exchanged. While the study of bacterial symbiosis is often limited to either chemistry or microscopy work, recent functional and metabolomics methods are available to enable chemical imaging of specialised metabolites involved in host–bacteria interactions.

In our study, comparative metabolomics using atmospheric pressure scanning microprobe matrix-assisted laser desorption/ionisation high-resolution mass spectrometry (AP-SMALDI-HRMS) enables the identification of specialised metabolites of the marine macroalga Ulva mutabilis (Chlorophyta) and its associated essential bacteria, a model system for cross-kingdom interactions [7]. The method provides a tool to formulate hypotheses about metabolic processes in the phycosphere while preserving spatial structure. This novel depth of insight into a multicellular host and bacteria interactions can characterise natural products in symbiotic interactions.

Algal growth and morphogenesis-promoting factors (AGMPFs) are required for the development of the model organism Ulva mutabilis [7]. They are provided by a combination of two essential bacteria, Maribacter sp. MS6 and Roseovarius sp. MS2 forming a tripartite community [3,7,8] (see also Figure 1a and the Graphical Abstract). In turn, Roseovarius sp. benefits from the released photosynthetic glycerol as a carbon source [9]. Axenic Ulva germ cells (i.e. gametes) develop into a callus-like phenotype composed of undifferentiated cells with malformed cell walls [8,10]. Up to now, the bacterial sesquiterpenoid thallusin, released by Maribacter sp. [11,12], is the only known AGMPF that induces morphogenesis such as rhizoid and cell-wall formation in Ulva spp. [11,12] or thallus development in Monostroma spp. [13]. The Roseovarius-factor that promotes cell division is still unknown [3,8]. Algal substances are released into the surrounding environment to attract epiphytic bacteria and initiate the cross-kingdom interaction [14,15]. Ulva attracts Roseovarius sp. MS2 through the sulphur-containing zwiterion dimethylsulphonopropionate (DMSP), resulting in biofilm formation on the algal surrounding [9]. The bacterium subsequently uses the provided glycerol for growth and transforms DMSP into methanethiol and dimethyl sulphide [9].

The metabolic activities of marine bacteria and algae can be surveyed using mass spectrometry-based methods. For example, stable sulphur isotope ($^{34}$S) labelled DMSP was used to track DMSP uptake and degradation by marine bacteria, and secondary ion mass spectrometry was applied to visualise it at the single-cell level [16]. The interaction between epibiotic bacteria on algal surfaces and their metabolic activities can be monitored in situ or using an imprinting method by desorption electrospray ionisation mass spectrometry [17,18]. In U. mutabilis gametophytes, matrix-assisted laser desorption ionisation mass spectrometry imaging (MALDI-MSI) was used to identify cell differentiation markers [19]. However, there has yet to be a thorough investigation of associated-mutualistic bacteria. MALDI-MSI has been shown to have high sensitivity and spatial resolution at the microscale in plant tissues, plankton, and other microbes [20,21].

The application of a MALDI matrix to a sample is an important part of the MALDI-MSI experiment. MALDI-MS can be used to identify proteins and metabolic signatures [22-24] from bacteria and macroalgae, as well as biofilms [25]. The primary function of the applied matrix is to improve the quality of the MS spectra, particularly the signal intensities of the compounds of interest. In some cases, the matrix might also work in opposition to this premise, suppressing desired ions. Then, matrix-free approaches such as LDI-HRMS can overcome this limiting phenomenon and have been applied for species-level microalgae identification based on metabolic profile fingerprint matching [26-28].
Our research combines cutting-edge laser scanning microscopy and high-resolution mass spectrometry to uncover Ulva/bacteria interactions and specialised metabolites at the microscale level. In this study, we demonstrate that the chemosphere of *U. mutabilis* changes depending on the presence or absence of the bacterial symbionts (*Roseovarius* sp. MS2 and *Maribacter* sp. MS6). As a result, specific metabolic markers can be used to identify bacteria in the vicinity of *U. mutabilis*. We used an untargeted comparative metabolomics approach that also provides micrometre-resolved MS imaging data through AP-SMALDI-HRMS. Two sample preparations, matrix-free LDI and MALDI, were performed to increase the range of metabolites recovered with this type of ionisation. We identified significant metabolites that define the host–bacteria interactions based on spectral similarity with standards. Using combined imaging mass spectrometry and confocal laser scanning microscopy, we then linked the chemical and microscopic observations that characterise the symbiotic association (cLSM).

Results and Discussion

Comparative metabolomics using AP-SMALDI-HRMS identifies metabolites in axenic algae and those present during macroalgal–bacterial symbiosis

Axenic gametes of *U. mutabilis* (phenotype slender) were allowed to settle onto glass plates in Petri dishes filled with growth medium. In the absence of the symbionts, the axenic gametes developed into undifferentiated cells known as the callus-like form [8,29]. In the second set of samples, algae were inoculated with the two marine bacteria, *Roseovarius* sp. MS2 and *Maribacter* sp. MS6, developing into a phenotype composed of bilayer cells and organised tissues, as previously reported [8]. The algal germlings incubated with the marine bacteria showed a rhizoidal zone that serves for substrate attachment and a thallus zone. From apex to rhizoid, *Ulva* germlings had an average length of 50 to 150 µm after three weeks of growth. The samples were recovered, dried on tissue, and for MALDI, immediately covered with 2,5-dihydroxybenzoic acid (DHB) applied by spraying. We targeted either specialised tissues (rhizoidal zone versus thallus) or the whole alga germlings (axenic callus versus alga in symbiosis) using a mounted AP-SMALDI camera (Figure 1a). The metabolic profiles of tissue and whole algae were obtained from callus or alga in symbiosis using AP-SMALDI-HRMS with the two sample preparations, either with matrix deposition (MALDI-HRMS) or matrix-free analysis (LDI-HRMS) (Figure 1a and b). The data matrix was generated by processing the raw spectra, and the data tables produced were from 1534 to 4986 features (m/z) (Figure 1b and Table S1 in Supporting Information File 1). The principal component analysis (PCA) visualised differences between metabolic profiles of axenic algae, algae in symbiosis, and specialised tissues (thallus, rhizoidal zone), analysed either with LDI or MALDI-HRMS. The metabolic profiles of axenic algae and algae in symbiosis were significantly different, while tissue-specialised metabolomes were less differentiated in the PCA score plots (Figure 1b). Significant features in the loading plots were listed in a heatmap to compare their relative abundance of intensities averaged per sample class (Figure 1c). Among the statistically significant features in all datasets (Table S1, Supporting Information File 1), six metabolites were identified, which were annotated using spectral similarity with analytical standards. For example, the features m/z 104.1064 and m/z 143.0815 were selected among the significant markers of the MALDI-HRMS profiling of axenic algae and the rhizoid tissue (whole alga profiling) of *U. mutabilis* grown with the marine symbiotic bacteria, respectively (Figure 1c). The heatmap shows the complementarities of both methods, LDI or MALDI-HRMS, as the significant features m/z 104.1064 and m/z 143.0815 have only been detected by one of the two methods.

Identification of metabolites in *Ulva–bacteria* symbiosis

To identify the selected markers found by the comparative metabolomics study, we searched several mass spectra libraries, including METLIN, and determined the chemical formula based on exact mass. We also used spectral similarity matching of data acquired from analytical standards. Choline was identified from the molecular peak m/z 104.1064 for [M]+ (calculated m/z as 104.1069 ± 4.8 ppm for C₇H₁₄N+O) in the profiles of axenic *U. mutabilis* (Figure 2a). This small polar metabolite was linked to the metabolic homeostasis of *Ulva lactuca* during tidal cycles [30]. Choline is the precursor of the membrane constituent phosphatidylcholine [31]. We inferred that the accumulation of choline in axenic *U. mutabilis* germlings might correlate with the absence of the key bacterial morphogen thallusin, which induces cell wall and rhizoid formation. The accompanying formation of cell wall protrusions might disrupt the cell membrane arrangement indicated by choline accumulation. Screening the tripartite community *Ulva–Roseovarius–Maribacter* identified ectoine as a metabolic marker of the rhizoidal zone (Figure 2b). The molecular formula C₆H₁₀N₂O₂ was deduced from the molecular peak at m/z 143.0817 for [M + H]+ (± 1.4 ppm) and m/z 165.0636 for [M + Na]+ (± 1.2 ppm) detected in the AP-SMALDI-HRMS profiles of the standard and rhizoid tissue of *U. mutabilis* in symbiosis with the marine bacteria. To separate algal and bacterial metabolism, single colonies of *Roseovarius* sp. MS2 and *Maribacter* sp. MS6 were deposited onto glass slides and analysed with AP-SMALDI-HRMS/MS. Using spectral similarity matching based on the fragmentation pattern
obtained from AP-SMALDI-HRMS experiments, we proved that the bacterial symbiont *Roseovarius* sp. MS2 produces ectoine (Figure 2c). This observation supports earlier assumptions that the rhizoidal zone is mainly colonised by *Roseovarius* sp. MS2 [8,29].

Ectoine is a known osmoprotectant produced by marine bacteria and phytoplankton with high concentrations during saline stress conditions [32]. It has not yet been described in the *Ulva*-bacteria symbiosis. Not all essential genes for ectoine biosynthesis reported by [33] were found in the *U. mutabilis* genome [34], providing further support for the bacterial origin of ectoine. Homologs of EctA (UM017_0070.1, E value 0.34), EctB (UM084_0040.1, E value < 0.0001) that provide the central intermediate N-acetyl-2,4-diaminobutyrate and EctD (UM025_0127.1, E value 0.094) an ectoine hydroxylase could be identified. However, a homolog gene for EctC (ectoine synthase) is missing in the *U. mutabilis* genome. In addition, despite the low E value of EctB, the reciprocal NCBI-blast search against the anoxygenic photosynthetic halophile and ectoine-producing bacterium *Halorhodospira halochloris* [35] did not confirm the presence of the sequence in the algal genome. Therefore, it is unlikely that the alga produces ectoine. In summary, ectoine is indicative of *Roseovarius* sp. MS2 in the tripartite community and can serve for localisation studies.

**Figure 1**: Untargeted comparative metabolomics using AP-SMALDI-HRMS highlighted metabolites involved in *Ulva*-bacteria symbiosis. a) The study looked at axenic algae with cell wall protrusions, the whole algae, and specific tissues with bacterial symbionts. b) The profiles of axenic alga (“axenic”) were contrasted with alga with bacterial symbionts (“symbiosis”) in the PCA score plots for LDI and MALDI-HRMS. The ellipses represent the 95% confidence region. c) The significant features (m/z) characterising axenic algae, algae in symbiosis, or differentiated tissues (blades/thallus, rhizoids) are represented in a heatmap with their relative abundance. The colour scale represents the averaged TIC normalised intensities per sample class (red colour for high intensity, blue for low intensity).

**Localisation of bacterial symbionts of *Ulva mutabilis* using fluorescence microscopy and imaging mass spectrometry**

Based on the above results, we combined LDI-MS imaging mass spectrometry and cLSM using a non-specific fluorescence labelling probe to visualise the bacterial cells living in symbiosis with *U. mutabilis*. Following a one-month incubation in clean cuvette slides placed in Petri dishes filled with medium, axenic and bacteria-inoculated *U. mutabilis* germlings were stained with SYBR Gold, a sensitive probe forming a complex with DNA with high fluorescence quantum yield [36]. In the axenic callus-like form, the nuclei of algal cells and the bacterial cells accumulated around the rhizoidal tissue and exhibited the specific fluorescence after SYBR Gold staining (Figure 3a) as previously described [8,37]. These findings indicated that bacteria are associated with their algal host during symbiosis.

In parallel, we visualised the metabolites produced by the biofilm formed around *U. mutabilis* by imaging analysis with AP-SMALDI-HRMS. Three replicates of each the axenic algae, algae in symbiosis, germlings, and bacterial cells in monocultures were imaged after matrix deposition by AP-SMALDI-HRMS over a centimetre-scaled area (Figure 3b). The algal pigment chlorophyll was localised with the algal tissues (Figure 3b and Figure S1 in Supporting Information File 1). Even though...
Figure 2: Identification of significant features associated with axenic or bacterial symbiont-associated alga Ulva mutabilis (phenotype slender). a) The structural determination was achieved by spectral matching with the analytical standards using AP-SMALDI-HRMS. b) Relative amounts of ectoine ([M + H]+) were determined by AP-SMALDI-HRMS measurements to compare different tissues: axenic and algae in symbiosis. One-way ANOVA with a Fisher HSD post hoc test found choline to be significant in profiles of axenic algae (F = 42, P-value < 0.0001) and ectoine in profiles of rhizoidal zones of algae in symbiosis (F = 4, P-value < 0.005) (colour code with reference to Figure 1a). c) Ectoine ([M + H]+, precursor ion) was identified in a single colony of the bacterial symbiont Roseovarius sp. MS2 using AP-SMALDI-HRMS/MS analysis.

most of the seawater media was removed from the Ulva samples during sample preparation, crystallisation of seawater salts on the sample surface occurred. The size of the crystals and their distribution within an imaged area were examined using a digital microscope and found to be homogeneous and consistent across the samples and experiments. As a result, the ion
a) Cytochemical staining with confocal laser scanning microscopy

b) AP-SMALDI-HRMS

Symbiont-free alga

Roseovarius sp. MS2

Alga in symbiosis with bacteria

Figure 3: Visualisation of algae Ulva mutabilis grown under axenic conditions or with bacterial symbionts Roseovarius sp. MS2 and Maribacter sp. MS6. a) Images acquired after nucleic acid staining with SYBR gold and with confocal laser scanning microscopy. The protrusion of alga grown without bacterial symbiont is highlighted (red arrow). DIC: differential interference microscopy. b) The images show ectoine spatial localisation and thus the presence of Roseovarius sp. (m/z 143.0814 for [M + H]^+; m/z 165.0633 for [M + Na]^+; shown in green) as well as chlorophyll (m/z 892.5360 for [M + H]^+; m/z 614.2375 fragment shown in white). These metabolite traces are visible in axenic algae, symbiotic algae, and cell cultures of bacteria Roseovarius sp. MS2. White arrows indicate the rhizoidal zones.
suppression effect caused by the presence of seawater crystals on the *Ulva* samples and surroundings was consistent across all measurements (Supporting Information File 1, Figures S1 and S2).

Ectoine was detected in both profiling and imaging MS spectra as the [M + Na]$^+$ adduct at m/z 165.0633. Ectoine was mainly found around the rhizoid in elevated amounts. Thus, *Roseovarius* sp. MS2 became visible in the rhizoidal zone and on the thallus due to the exclusive production of ectoine within the tripartite community (Figure 3b). AP-SMALDI-HRMS studies extended to the entire clade of motile Rhodobacteraceae will reveal other characteristic metabolites of the *Ulva*–bacteria interactions. Those species attracted by *U. mutabilis* (e.g., through DMSP) that use the provided photosynthates [9], will preferentially succeed the previously described competitive colonisations of *Ulva* spp. [38,39]. Also, related species of *Roseovarius* sp. MS2 can often release unknown AGMPFs [29,40], which further foster the bacterial–algal interactions. As more species-specific metabolic markers become available, AP-SMALDI imaging will be a powerful tool to track these dynamic microbial colonisation processes using the *U. mutabilis* model system with a designed microbiome.

**Conclusion**

Metabolic profiling of whole alga and specialised tissues conducted with AP-SMALDI-HRMS enabled identifying specific metabolites in host–bacteria symbiosis. We report the first identification of choline and ectoine as markers of symbiont-free *U. mutabilis* and rhizoid tissue of algae in symbiosis with bacteria. We visualised the rhizoidal zone formed by the bacterial symbionts using chemical staining, confocal laser scanning microscopy, and imaging mass spectrometry. Notably, ectoine was used as a metabolic marker to identify bacteria in the biofilm associated with *U. mutabilis* and the algal surface. Visualising the spatial distribution of epiphytic bacteria in the phycosphere will contribute to the general understanding of the chemically mediated cross-kingdom interactions. The combined AP-SMALDI and cLSM imaging with resolution down to the level of a single bacterial cell introduced here can be applied to other microbial consortia and their hosts and will be instrumental for microbiome research.

**Experimental**

**Biological experiments and imaging microscopy**

The laboratory strains of *U. mutabilis* (sl-G[mt+]) are direct descendants of the original isolates collected by B. Føyn in Portugal (Ria Formosa) in 1958 [8]. This strain is used as a model organism in cross-kingdom interactions [7,34,37] and cultivated under standardised conditions [41,42]. *Ulva* strains are available from the corresponding author (Thomas Wichard, Friedrich-Schiller-Universität Jena, Germany).

Gametogenesis of *U. mutabilis* was induced by chopping harvested algal tissue, and released gametes were purified from accompanying bacteria according to the protocol of Wichard and Oertel (2010) [41]. The strains *Roseovarius* sp. MS2 (Genbank EU359909) and *Maribacter* sp. MS6 (Genbank EU359911) were originally isolated from *U. mutabilis* [8] and were cultivated in Marine Broth medium (Roth, Germany) at 20 °C. *Ulva* gametes were either grown axenically or inoculated with the bacteria (final optical density OD$_{620}$ = 0.001). All algae were cultivated in *Ulva* culture medium (UCM) [43] at 18 °C with the illumination of about 60 μmol photons m$^{-2}$ s$^{-1}$ under a 17:7 light/dark regime. Axenic *Ulva* gametes deposited on cleaned glass slides and inoculated with bacteria MS2/MS6 were prepared following the procedure for in situ MS imaging described by Kessler et al. [19]. Briefly, algal gametes were inoculated to 10 mL medium in 9 cm diameter sterile Petri dishes with a clean and autoclaved glass slide (25 mm × 75 mm) with cavities (Paul Marienfeld, Germany) on the bottom; samples were incubated for one month at 18 °C in static conditions. An inverted microscope was used to monitor the algal growth. Transmitted light microscopy pictures were obtained using a Keyence BHX-500 digital microscope. Samples were recovered with pliers and fixed with glutaraldehyde 1% (Merck), stained with SYBR Gold (1% in DMSO, Invitrogen, Thermo Fisher); a cover slide was added, followed by incubation in the dark at 15 °C for 15 min. Cavity slides were spotted with 100 μL of SYBR Gold or unstained bacterial monoculture (*Roseovarius* sp MS2 or *Maribacter* sp. MS6) to use them as controls. Fluorescence images (1024 × 1024) were acquired using a Zeiss cLSM 880 (Carl Zeiss AG, Oberkochen, Germany) with a Plan-Apochromat 20 × 0.8 and 488 nm Argon-laser excitation (5% transmission). Emission wavelengths for SYBR Gold (490–650 nm) and chlorophyll A (653–735 nm) were separated via the spectral detection unit. Transmitted light was detected by the transmitted light-PMT. The effect of an additional quick washing step was tested by gently adding 100 µL of sterile MQ water for two seconds. The controls consisted of bacteria grown for one week in monoculture in 40 mL of marine broth medium and the axenic medium with fixative and stain. All the experiments with glass slides were performed in biological triplicates.

**Genome analysis**

To identify the putative biosynthetic gene cluster (ect gene cluster) in *U. mutabilis* [34], the algal genome was searched for the gene ectoine hydroxylase (ectD) and also for a specialised aspartokinase (ask_ect). Aspartokinase (Ask), along with
t-aspartate-p-semialdehyde-dehydrogenase (Asd), provides the precursor t-ASA for ectoine biosynthesis [33,44,45]. Homologs of the enzymes of the ectoine pathway from *Halorhodospira halochloris* were identified by BLAST searches of the *U. mutabilis* genome at ORCAE using default parameters (https://bioinformatics.psb.ugent.be/orcae/overview/Ulvmu).

**AP-SMALDI-HRMS metabolic profiling and imaging**

All standards and *Ulva* samples were analysed via AP-SMALDI (AP-SMALDI10, TransMit, Germany) ion source equipped with a UV (337 nm) nitrogen laser (LTB MNL-106, LTB, Germany) coupled to a high-resolution mass spectrometer Q-Exactive Plus (Thermo Fisher Scientific, Bremen, Germany). Glass slides with one month-grown algal gametophytes were gently recovered from a Petri dish filled with sterile water using a sterile tweezer and dipped for one second in sterile ultrapure water to remove the excess salts before metabolic profiling. When algae were investigated directly on a glass slide before in situ MS imaging, blotting paper was used to remove sea water (see also Supporting Information File 1). The desired area of a glass slide covered with algal individuals was first marked, photographed, and finally fixed on the AP-SMALDI metal target.

AP-SMALDI profiling and imaging experiments unless otherwise stated were enhanced by a 2,5-dihydroxybenzoic acid (DHB) MALDI matrix. A methanolic solution of the DHB matrix at a concentration of 4 mg mL\(^{-1}\) was applied onto a sample via SunChrom MALDI spotter (SunChrom GmbH, Germany). The spraying method was optimised using the following parameters: line distance 2 mm, spraying speed 800 mm min\(^{-1}\) with 5 seconds drying time, and matrix solution flow rate in 4 cycles from 10 µL up to 30 µL min\(^{-1}\). Solvents used in this study were all LCMS analytical grade. 2,5-Dihydroxybenzoic acid with a purity of above 98% and high purity MS-grade methanol were purchased from Sigma-Aldrich (Germany).

All *Ulva* samples were imaged in the positive ion mode using a step size of 5 µm and with the number of laser shots per spot set to 30 (approximately 1.2 µJ shot\(^{-1}\)) within the laser frequency of 60 Hz. MS spectra were acquired in a mass range from m/z 100 to m/z 1000 with a resolving power of 280000. Pseudo ion intensity maps of selected m/z values were generated using the Mirion V3 software package with an m/z width of 0.01 u.

In the profiling mode, the single *Ulva* individuals were targeted visually and ablated with a laser spot size of approximately 10 µm in positive and negative polarity in a mass range from m/z 100 to m/z 1000. The other parameters stayed like for the MSI mode. In profiling, the same area of the rhizoid and the tip of the thallus of different individuals were analysed by laser ablation over one-minute time acquisition. Axenic and alga in symbiosis germlings were profiled with a UV laser along a longitudinal axis to investigate the effect of bacteria on metabolism changes in *U. mutabilis*.

The size of the sample groups analysed by AP-SMALDI-HRMS in profiling mode was n = 10 for thallus tissue, n = 9 for rhizoid tissue, n = 8 for axenic callus, and n = 10 for alga in symbiosis. Matrix-free experiments (LDI-HRMS) were performed in profiling mode under the same experimental conditions as the AP-SMALDI-HRMS. The size of the sample groups was defined as follows: n = 6 for rhizoid, n = 7 for thallus and whole alga profiling, n = 10 for alga in symbiosis, and n = 15 for axenic alga.

The metabolic profiles of nutrient media were obtained by analysing 30 µL deposited onto cleaned glass slides and following the same protocol used for the *Ulva* samples. In the late exponential stage, bacterial monoculture was recovered from agar plates with a 10 µL loop and diluted in 100 µL of sterile water. Five microliters of the solution were spotted onto a glass slide and analysed in AP-SMALDI-HRMS mode in positive and negative polarity.

The data acquired in MSI mode were collected with Xcalibur software version 2.8 SP1 Build 2806 (Thermo Fisher Scientific, Germany) while the acquisition of spatial scans, pre-defined in the x- and y-direction as rectangular sample regions, was controlled by the MCP (Master Control Program, TransMIT GmbH, Giessen, Germany). The raw data acquired in profiling mode were visualised in Thermo Xcalibur™ version 3.0.63 (Thermo Fisher Scientific, Germany) and then converted to netCDF format using the Thermo File Converter tool. Data preprocessing was performed to extract the intensities in each profile, excluding the features of the nutrient medium using a script adapted from the MALDIquant package [46]. Spectra were de-noised with a signal-to-noise ratio of 5. Normalisation was done based on total ion current (TIC) recommended for MALDI-MS analysis [47]. All spectra, images, R data, scripts, and results from the statistical analysis were uploaded and are freely accessible in the Max Planck repository Edmond (https://dx.doi.org/10.17617/3.4v).

**Significant features analysis and metabolite identification**

Data analysis was conducted in MetaboAnalyst 4.0 [48] to perform univariate and multivariate statistical tests and find significant differences in intensities and the presence or absence of metabolites in the samples. Pareto scaling and cube root trans-
formation were conducted to normalise the datasets before the multivariate statistics. PCA highlighted the metabolic differences between axenic and alga in symbiosis and between thallus and rhizoid tissues. Significant features were searched in the PCA loading plots and also in the pattern hunter plots obtained from a correlation analysis based on the Pearson correlation coefficient R. A one-way ANOVA with Fisher’s LSD post hoc test (P-value < 0.05) was performed, and the relative amounts of the significant features were displayed as a boxplot. The selected significant features were further searched in the raw HRMS profiles to identify those with the reliable isotopic pattern assigned to a metabolite. The m/z values were searched in the METLIN database, using a mass deviation equal to or lower than five ppm, which suggested several known natural products such as ectoine [49].

To confirm the identity of the significant features, mass spectral information was compared with analytical standards analysed with the AP-SMALDI-HRMS (DSMP, chlorophyll-a, ectoine, choline). MS/MS experiments were performed with AP-SMALDI-HRMS to match the fragmentation pattern between the standard ectoine and bacteria monoculture profile. fragmentation spectra of ectoine were acquired from the bacterial isolate *R. varius* sp. MS2 and an ectoine standard. To perform a measurement, 4 μL of ectoine at concentration 50 μM was pipetted onto a clean glass slide (washed with dH2O, acetone) and overlaid with 2 μL of a methanolic solution of the DHB matrix at a concentration of 4 mg mL−1. All-ion fragmentation (AIF) mode was set as follows: molecular ion of ectoine at m/z 143.1; isolation window m/z ± 0.2; 45 NCE. The peak resolution was set at 28000, and the mass range was set from m/z 50 to m/z 300.

Supporting Information

Supporting Information File 1

Details on sample preparation and additional figures. [https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-17-91-S1.pdf](https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-17-91-S1.pdf)

Funding

SF was funded by the International Max Planck Research School Exploration of Ecological Interactions with Molecular Techniques. This work was supported by an MPG Fellowship awarded to GP and by the Deutsche Forschungsgemeinschaft through Grant No. SFB 1127/2 ChemBioSys-239748522 (GP, TW) and within the framework of the priority program (SPP 1158) “Antarctic Research with comparative investigations in Arctic ice areas” under the project number #424256657 (FG, TW).

References
