CHEMBIOCHEM

DOI: 10.1002/cbic.201000510

Linking Chemical and Microbial Diversity in Marine Sponges: Possible Role for Poribacteria as Producers of Methyl-Branched Fatty Acids

Thomas Hochmuth,[a] Holger Niederkrüger,[a] Christine Gernert,[b] Alexander Siegl,[b] Stefan Taudien,[c] Matthias Platzer,[c] Phillip Crews,[d] Ute Hentschel,[b] and Jörn Piel*[a]

Many marine sponges contain massive numbers of largely uncultivated, phylogenetically diverse bacteria that seem to be important contributors to the chemistry of these animals. Insights into the diversity, origin, distribution, and function of their metabolic gene communities are crucial to dissect the chemical ecology and biotechnological potential of sponge symbionts. This study reveals a sharp dichotomy between high and low microbial abundance sponges with respect to polyketide synthase (PKS) gene content, the presence of methyl-branched fatty acids, and the presence of members of the symbiotic candidate phylum "Poribacteria". For the symbiont-rich sponge *Cacospongia mycofijiensis*, a source of the tubulin-inhibiting fijianolides (=laulimalides), near-exhaustive large-scale sequencing of PKS gene-derived PCR amplicons was conducted. Although these amplicons exhibit high diversity at the sequence level, almost all of them belong to a single, architecturally unique group of PKSs present in "Poribacteria" and are proposed to synthesize methyl-branched fatty acids. Three components of this PKS were studied in vitro, providing initial insight into its biochemistry.

Introduction

Sponges are the most ancient multicellular animals, dating back to the Precambrian period, more than 635 million years ago.[1] Their natural product diversity is among the highest found in nature. There is growing evidence that bacterial symbionts play a crucial role as producers of sponge-derived metabolites.[2] In many sponge species, termed "high microbial abundance" (HMA) sponges, up to half the biomass can consist of microbial symbionts, whereas "low microbial abundance" (LMA) species, which live in the same habitats, contain much fewer symbionts.[3] Often, distantly related HMA sponges from different oceans share remarkably similar microbial communities.[3a] In general, attempts to access sponge-derived natural products of biomedical significance through mariculture or culture of sponge symbionts have remained unsuccessful.[4] However, by using a cultivation-independent approach, in two cases we were able to assign biosynthetic genes for known "sponge" polyketides to bacterial producers,[5] indicating that symbionts could be the true source of many other polyketides as well. Indeed, a high diversity of bacterial-type polyketide synthase (PKS) genes can be detected in many sponges.[5b,6] It is an important task to understand the function and taxonomic distribution of these genes, as such insight will provide new avenues for natural product discovery and greatly aid biosynthetic gene cloning to develop sustainable production sources.

The vast majority of PKS genes detected by PCR in sponges belong to an unusual type of small PKS (Figure 1). These genes are widely distributed in, and at the same time highly specific for, sponges and have therefore been termed *sup*, from "sponge symbiont ubiquitous PKS.[6b] Genetic analysis of FACS-sorted individual symbiont cells revealed that *sup* genes are present in members of the candidate phylum "Poribacteria",[7] a highly unusual bacterial group that is nearly exclusively found in sponge symbioses.[8] It is unknown whether these as-yet unculturable bacteria are the only *sup* source and what the metabolic products of these genes are. Although the Sup enzymes clearly classify as PKSs from a phylogenetic standpoint, their architecture is remarkably similar to that of animal FASs (Figure 1).[6b] A major difference to these enzymes is the presence of an apparently functional methyltransferase (MT) domain, suggesting that the PKS synthesizes methyl-branched fatty acids. Sponges are indeed one of the richest known sources for mid-chain-branched FAs (MBFAs), and these lipids have often been proposed to be of bacterial origin.[9] The Sup system therefore appears to be a prime candidate for their biosynthesis. To obtain more insight into its distribution and func-

[a] *T. Hochmuth, H. Niederkrüger, Prof. Dr. J. Piel*
Kekulé Institute of Organic Chemistry and Biochemistry
University of Bonn, Gerhard-Domagk-Strasse 1, 53121 Bonn (Germany)
Fax: (+49) 228-739712
E-mail: joern.piel@uni-bonn.de

[b] *C. Gernert, A. Siegl, Prof. Dr. U. Hentschel*
Department of Botany II, Julius von Sachs Institute for Biological Sciences
University of Würzburg
Julius-von-Sachs Platz 3, 97082 Würzburg (Germany)

[c] *Dr. S. Taudien, Dr. M. Platzer*
Genome Analysis, Leibniz Institute for Age Research, Fritz Lipmann Institute
Beutenbergstrasse 11, Beutenberg Campus, 07745 Jena (Germany)

[d] *Prof. Dr. P. Crews*
Department of Chemistry and Biochemistry, University of California Santa Cruz
1156 High Street, Santa Cruz, CA 95064 (USA)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.201000510.

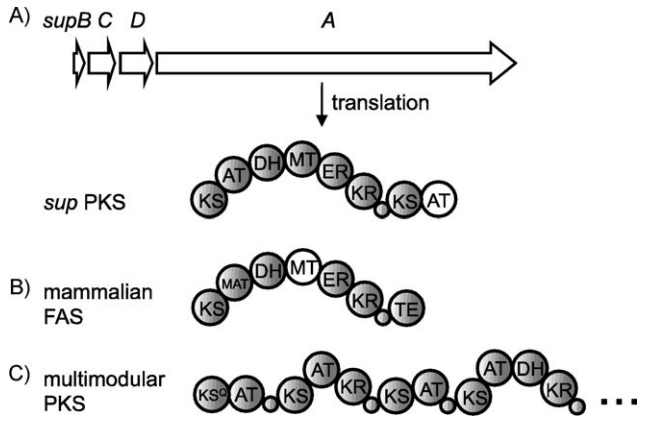

Figure 1. Comparison of selected PKS-like enzyme architectures. A) Architecture of the *sup* gene cluster isolated from the *T. swinhoei* metagenome and translated PKS. SupB is an ACP, SupC is a 4′-phosphopantetheinyl transferase (PPTase), and SupD is a putative hydrolase. B) FASs involved in animal fatty acid biosynthesis. C) Bacterial multimodular PKSs involved in complex polyketide biosynthesis. Domains shown in white lack active site motifs. Abbreviations: KS, ketosynthase; KS^Q, ketosynthase involved in chain initiation; AT, acyltransferase; DH, dehydratase; MT, methyltransferase; ER, enoyl reductase; KR, keto reductase; MAT, malonyl-/acetyltransferase; TE, thioesterase. Small unlabeled circles represent acyl carrier proteins (ACPs).

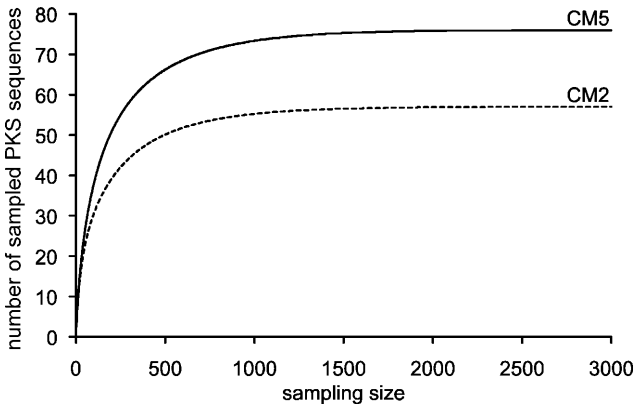

Figure 2. Rarefaction curves for PKS sequences amplified from *C. mycofijiensis* CM2 (····) and CM5 (——).

tion, we conducted an extensive analysis of *sup* genes, lipid content, and Poribacteria present in six sponge species. The data reveal a significant dichotomy in the PKS content and fatty acid metabolism of sponge–bacterial associations and lend initial chemical evidence to the hypothesis that poribacterial symbionts are an important source of sponge MBFAs.

Results and Discussion

PKS diversity in *Cacospongia mycofijiensis*

Previous studies have estimated the PKS diversity in sponges by sequencing a limited number of randomly sampled PCR amplicons.[5b,6] To obtain more detailed information about the numbers and types of PKS genes present in sponge metagenomes, we selected *Cacospongia mycofijiensis* as an HMA model sponge for 454-based deep sequencing of amplicons. Specimens collected at Vanuatu belong to the chemically most interesting sponges, as they contain members of at least three distinct groups of bioactive polyketides or hybrid polyketide–peptides: fijianolides (laulimalides), latrunculin A, and mycothiazole.[10] For PKS detection, degenerate PCR primers specific for the ketosynthase (KS) domain of PKSs, but not of FASs, were tested with total *C. mycofijiensis* DNA as template. Amplicon mixtures obtained from two different sponge specimens (*C. mycofijiensis* CM2 and CM5 collected at the same time from the same site) were then subjected to 454 sequencing. In total, 30473 sequences were obtained, which we hoped would provide an almost complete coverage of the PKS diversity present in the amplified DNA. Indeed, rarefaction analysis showed that saturation was already reached after sampling of about 2000 sequences (Figure 2). The data therefore allow the conclusion that, with duplicates and close orthologues eliminated at a 5% similarity cutoff, around 57 and 76 different PKS sequences were amplified from *C. mycofijiensis* CM2 and CM5, respectively. Only 15 sequences were found to be common between the two specimens, resulting in a total of 118 distinct PKS amplicons. Surprisingly, in spite of the large sequence space covered, only five of the sequences could be attributed to non-*sup* PKS types. Two amplicons resembled PKSs from dinoflagellates and other protists. Three further exhibited the highest similarity to either *cis*-AT PKSs, hybrid NRPS-*cis*-AT PKSs, or *cis*-AT PKS modules integrated into *trans*-AT PKSs. Studies are currently in progress to investigate the role of the latter three PKS sequences in the production of bioactive compounds from *C. mycofijiensis*. However, one complication is that there can be significant chemotype differences between colonies of this sponge growing in close proximity.[10d] The diversity of *sup* genes within individual specimens, but even more so among intraspecific chemotypes, is remarkable, considering that close homologues are virtually absent in sources other than sponges. The high abundance of *sup* amplicons also underscores that general KS-based PCR primers are not suitable for the isolation of genes encoding complex polyketide biosynthesis from sponge metagenomes. This illustrates the great importance of applying more specific targeting strategies.[5b]

Comparative analysis of poribacterial 16S rRNA and PKS genes of HMA and LMA sponges

The distribution of *sup* genes among various sponge species was next analyzed to obtain insight as to whether Poribacteria might indeed be a prominent source of the *sup* genes and of MBFAs. For this purpose, *C. mycofijiensis* was genetically and chemically compared with five other HMA sponges (*Aplysina aerophoba*, *Theonella swinhoei*, *Agelas dilatata*, *Ircinia felix*, and *Xestospongia muta*) and six LMA sponges (*Dysidea etheria*, *Dysidea avara*, *Callyspongia vaginalis*, *Niphates digitalis*, *Ptilocaulis* sp., and *Amphimedon compressa*). When we analyzed these 12 sponges with PCR primers specific for the poribacterial 16S rRNA gene, all HMA sponges were positive, whereas none of the LMA specimens yielded a PCR product (Figure 3A). This result suggests that Poribacteria are exclusively present and

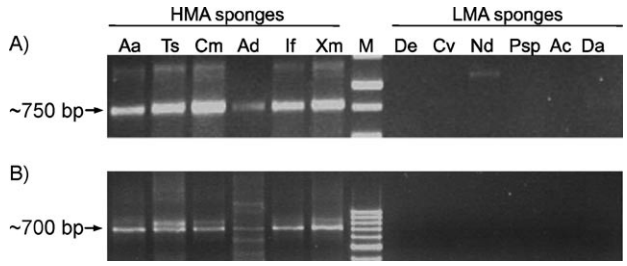

Figure 3. PCR-based detection of A) Poribacteria and of B) *sup* genes in HMA and LMA sponges. Primers for the 16S rRNA gene and the KS domain were used, respectively. Abbreviations: Aa, *Aplysina aerophoba*; Ts, *Theonella swinhoei*; Cm, *Cacospongia mycofijiensis*; Ad, *Agelas dilatata*; If, *Ircinia felix*; Xm, *Xestospongia muta*; M, DNA size marker; De, *Dysidea etheria*; Cv, *Callyspongia vaginalis*; Nd, *Niphates digitalis*; Psp, *Ptilocaulis* sp.; Ac, *Amphimedon compressa*; Da, *Dysidea avara*.

widely abundant in HMA sponges, but missing in the LMA category of sponges. A second PCR series with degenerate KS primers, followed by amplicon sequencing, revealed that *sup* genes are predominant in the amplified DNA of all HMA sponges (Figure 3B, Table 1). In contrast, none of the LMA sponges yielded a PCR product that belongs to the *sup* or any other PKS type. Thus, a strictly positive correlation exists between microbial abundance, the presence of Poribacteria, and the presence of *sup* genes. Importantly, because the KS primers can also amplify DNA from multimodular PKSs, the data also indicate that LMA sponges are generally not a major source of complex polyketides. In agreement, to the best of our knowledge polyketides have not been isolated from the six LMA species examined in this work, with the possible exception of amphitoxins in *A. compressa*,[11] which might be generated by an unusual PKS or FAS variant.[12] These findings may provide useful guidance for natural product discovery from sponges, as HMA sponges can be easily identified by microscopy or PCR.

Table 1. Genetic and chemical comparison of HMA and LMA sponges.

Species	Poribacterial amplicons	*sup* amplicons (total KS[a])	Distinct *sup* amplicons	Distinct MBFAs detected
Cacospongia mycofijiensis (HMA)	+	30468 (30473)	118	37
Aplysina aerophoba (HMA)	+	45 (45)	7	n.a.[b]
Theonella swinhoei (HMA)	+	46 (46)	10	32
Agelas dilatata (HMA)	+	2 (9)	2	16
Xestospongia muta (HMA)	+	44 (44)	7	n.a.[c]
Ircinia felix (HMA)	+	45 (45)	3	21
Dysidiea avara (LMA)	–	0 (0)[d]	0	1[e]
Dysidea etheria (LMA)	–	0 (0)[d]	0	0
Callyspongia vaginalis (LMA)	–	0 (0)[d]	0	n.a.
Niphates digitalis (LMA)	–	0 (0)[d]	0	n.a.
Ptilocaulis sp. (LMA)	–	0 (0)[d]	0	n.a.
Amphimedon compressa (LMA)	–	0 (0)[d]	0	n.a.[f]

[a] Number of total KS sequences based on RFLP analysis and sequencing. [b] Six MBFAs were previously reported.[32] [c] One MBFA (12Me16:0) was previously reported.[33] [d] No amplification observed. [e] Only in traces. [f] No MBFA according to a previous report.[18]

Comparative lipid analysis

A comparative FA analysis showed remarkable parallels to the genetic data. Because exhaustive lipid analyses are very time consuming, a subset of four HMA sponges (*C. mycofijiensis*, *T. swinhoei*, *A. dilatata*, and *I. felix*) and two LMA sponges (*D. etheria* and *D. avara*) were selected for these studies. In total, around 450 individual GC peaks from 50 sponge samples were assessed by GC–MS. These belonged to 186 different FAs. With 47 congeners, the diversity of MBFA was remarkably high. The largest number (43 different MBFAs) was present in *C. mycofijiensis*, and only *Agelas oroides* was previously known to contain more variants.[9b] However, while these lipids were abundant in the HMA sponges, virtually no MBFAs were found in the LMA animals. Due to this difference, the FA chromatograms of HMA sponges are highly characteristic and can be easily recognized (Figure 4). Because it is not yet known to which unit MBFAs in sponges are bound, we separated lipids into neutral (NL), glycolipid (GL), and polar lipid (PL) fractions. All fractions were found to contain MBFAs, suggesting that these moieties are present in diverse lipid types (table S1 in the Supporting Information). In the four bacteriosponges, the MBFA share of the total FAs ranged from 18% (*A. dilatata* PL fraction) to 34% (*I. felix* GL fraction). Most MBFAs were saturated and contained one mid-chain methyl branch, and a subset of these carried an additional branch of the *iso* or *anteiso* type. 11-Methyloctadecanoic (11Me18:0), 10-methylhexadecanoic acid (10Me16:0) and 9-methyltetradecanoic acid (9Me14) belonged to the most prominent compounds. In *C. mycofijiensis* (NL fraction), 11Me18:0 represented up to 12% to the total FA content. This MBFA is frequently reported as an abundant FA in sponges and might serve as a biomarker for HMA specimens.[13] All other MBFAs occurred in minor amounts, but could be assigned in almost all cases from their equivalent chain lengths (ECL)[14] and mass spectra of the methyl ester, pyrrolidide, and picolinyl derivatives.[15] These FAs exhibited

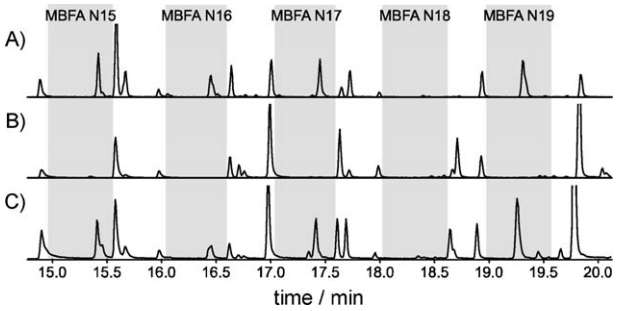

Figure 4. GC–MS analysis of the glycolipid fraction of three selected sponges. A) *T. swinhoei* (HMA); B) *D. etheria* (LMA); C) *A. dilatata* (HMA). The region in which MBFAs can be detected is shown in gray, along with the numbers of carbon atoms (N15 = 15 carbon atoms). Shown are extracted ion chromatograms at *m/z* 74, which is the mass of the characteristic McLafferty ion.

chain lengths between 14 and 24 carbon atoms and carried methyl branches at all positions between C2 and C17 with the exception of positions 4, 5, and 13. Mid-chain branches occurred preferentially at positions 10 and 11. Branches were most often encountered at even carbon numbers, but were also found at odd-numbered positions. The latter are not in accordance with methylations at acidic α-centers of nascent thioester intermediates or with the use of methylmalonyl-CoA extenders, as is likely for the *sup* and other PKSs. The presence of these odd-numbered branches suggests a different mechanism, such as double bond methylation,[16] β-branching (the conversion of β-keto groups into carbon branches),[17] or carbon chain shortening.[12] Several of the saturated MBFAs are novel natural products: 15Me20:0, 17Me22:0, and the *iso*-FAs 11,15-dimethylhexanoic acid (11Me-i16:0), 12Me-i17:0, 10Me-i18:0, 12Me-i19:0, 11Me-i20:0, and 16Me-i23:0. Five monoenoic FAs were also detected. Of these, 8-methyleicosa-7-enoic acid (8Me20:1Δ7) is novel and occurred in several HMA sponges in surprisingly large amounts (in *C. mycofijiensis* it represented ~7% of the total phospholipid FAs (table S1)). Another unprecedented FA is 11Me18:1Δ6, which is present in substantial quantities in *C. mycofijiensis* and *I. felix*. The other three unsaturated MBFAs occurred only in trace amounts and exhibited chain lengths of 16 or 18 carbon atoms; however, the position of the methyl group and/or the double bond could not be determined unequivocally. Of these, the 7Me16:1 acid is noteworthy, as it is the only MBFA found at very low concentrations in an LMA sponge (0.38% in *D. avara*; table S1). The general absence of MBFAs in LMA sponges is also corroborated by published reports for *A. compressa*[18] and *Pericharax heteroaphis*,[19] both of which contain low numbers of symbionts.[20] Taken together, with ten novel compounds detected, the examined HMA sponges represent a rich source of MBFAs; this stands in stark contrast to the examined LMA species, but parallels the distribution of Poribacteria and *sup* genes.

Functional study of *sup* genes

To obtain additional insight into the function of the *sup* PKSs, a gene cluster previously cloned from the metagenome of *T. swinhoei* (Figure 1)[6b] was transferred into several hosts, including *E. coli*, *Pseudomonas putida*, *Streptomyces albus*, and *Corynebacterium glutamicum* under the control of various promoters. However, no change in the metabolic profile was observed in any strain. Because the *sup* PKS contains no TE-like domain that normally releases the final product from the enzyme, one possible reason for the absence of new lipids could be the lack of a dedicated transferase and/or acceptor molecule in the hosts. We next attempted to express the large 3460-residue protein SupA in *E. coli* to study its function in vitro. However, only traces of soluble enzyme were detected. Individual domains and monofunctional proteins were therefore expressed and assayed. After cloning the free-standing ACP gene *supB* into pHIS8[21] and expression in *E. coli*, a soluble His_8-tagged protein was obtained. ESI-micrO TOF-Q-MS analysis revealed that the ACP is present exclusively in the *apo* form (*m/z* 11422.5 for $[M+H]^+$, 11422.93 calcd, Figure 5A) lacking

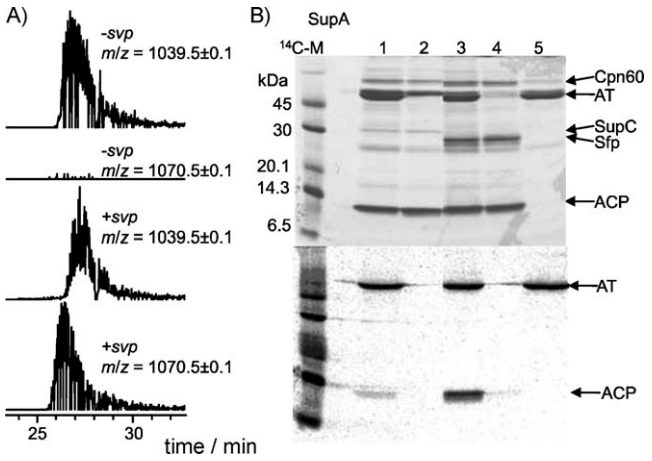

Figure 5. In vitro analysis of Sup proteins. A) LC–FTICR–MS analysis of the free-standing ACP SupB overexpressed in the absence (upper two panels) or presence (lower panels) PPTase Svp. Extracted ion chromatograms are shown. The *m/z* values correspond to the prominent ion at the charge state 11+ for the *apo*- (*m/z* 1039.4) and *holo*-ACP (*m/z* 1070.5). B) SDS gel (top) and autoradiographic analysis (bottom) of phosphopantetheinylation and acyl transfer test reactions conducted with $SupA_{ACP}$, the PPTases SupC and Sfp, the $SupA_{AT1}$. Lanes: ^{14}C-M, radiolabeled protein size marker; 1) $SupA_{ACP}$ + $SupA_{AT1}$ + SupC; 2) $SupA_{ACP}$ + SupC; 3) $SupA_{ACP}$ + $SupA_{AT1}$ + Sfp; 4) $SupA_{ACP}$ + Sfp; 5) $SupA_{AT1}$. All reactions contained [2-^{14}C]malonyl-CoA. The expected position of the proteins is shown. Cpn60 is a chaperone present in the expression strain *E. coli* ArcticExpress (DE3) RIL.

the 4′-phosphopantetheinyl residue that serves as covalent attachment site for the growing chain. This shows that the *E. coli* 4′-phosphopantetheinyl transferase (PPTase) AcpS does not accept the type II ACP of the *sup* pathway, although it is specific for ACPs of type II FASs.[22] Co-expression of SupB with Svp, a PPTase from the bleomycin PKS–NRPS pathway,[23] resulted in the production of a mixture of *apo*- and *holo*-ACP (*m/z* 11762.5 for $[M+H]^+$, 11762.13 calcd, for the *holo* form). In an analogous way, the expressed ACP domain of SupA ($SupA_{ACP}$) was only obtained as a *holo*-protein when expressed in the presence of Svp. To study phosphopantetheinylation and acyl transfer in vitro, the ACP and the first AT region of SupA as well as SupC, the PPTase of the *sup* cluster, were individually expressed as His_8-tagged proteins. As an alternative PPTase, Sfp with broad substrate specificity from *Bacillus subtilis*,[22] was also prepared. Incubation of the AT with [2-^{14}C]malonyl-CoA and analysis of the reaction mixture by radio SDS-PAGE (Figure 5B, lane 5) showed an attachment of the radiolabel to the protein, suggesting that the AT is functional and accepts malonyl-CoA as substrate. In the absence of the AT, no or only a weak signal was detected for the ACP (Figure 5B, lanes 2 and 4). In contrast, incubation of the ACP with a PPTase and subsequent treatment with the AT in the presence of the acyl donor resulted in clear labeling of both the ACP and the AT (lanes 1 and 3). These data thus suggest that SupA uses malonyl building blocks to assemble its product.

Conclusions

Symbiotic bacteria emerge as an important source of natural products in sponges, but remain poorly studied. Herein we

provide further support for the role of symbionts as producers of polyketides and methyl-branched fatty acids by revealing a positive correlation between bacterial abundance, the presence of PKS genes, and of MBFAs. Although bacteria can be detected at low levels in LMA sponges, the candidate phylum "Poribacteria" as a likely source of MBFAs was noticeably absent in our samples, indicating that a general, distinct dichotomy exists not only in terms of bacterial numbers but also of metabolic content and taxonomic composition. Of particular interest to natural product discovery could be our observation that PKS genes in general were not detectable in the examined LMA sponges, although the employed primers are specific for all major groups of PKSs involved in complex polyketide biosynthesis. If this correlation is confirmed by wider-scale analyses, simple microscopic examination of a collected sponge could streamline sponge-based drug discovery by allowing for the pre-selection of polyketide-containing candidates. Finally, deep sequencing of PKS types detected in *C. mycofijiensis* provided the first exhaustive analysis of PKS types detectable in a symbiont-rich sponge and revealed the first gene candidates for pharmacologically relevant compounds present in this species. Isolation of the corresponding gene clusters is currently in progress in our laboratories.

Experimental Section

General: *E. coli* XL1 Blue MRF′ (Stratagene) was used for cloning experiments and was grown in liquid or on solid Luria–Bertani (LB) medium (1.5% agar) at 37 °C. Sequencing was performed on ABI 377XL (Applied Biosystems) or GS FLX automated sequencers (Roche). Masses of *sup* proteins were determined with an ESI-micrO TOF-Q-MS instrument (Bruker Daltonik, Bremen, Germany). Fatty acid methyl esters (FAMEs) were analyzed by GC–MS (Shimadzu QP2010) on an HP-1 column (30 m × 0.25 mm, film thickness 0.25 μm). Helium was used as a carrier gas, injection was in split mode. The program was as follows: start with 50 °C, heat to 290 °C at a rate of 10 °C min^{-1}, hold for 15 min at 290 °C. Pyrrolidides and picolinyl esters were measured the same way as FAMEs, with the exception that the temperature was held at 290 °C for 25 instead of 15 min.

Sponge collection: HMA sponges (*A. dilatata*, *C. vaginalis*, *I. felix*, and *X. muta*) and LMA sponges (*A. compressa*, *D. etheria*, *Ptilocaulis* sp., and *N. digitalis*) were collected offshore Little San Salvador Island, Caribbean Sea, Bahamas (24°34.39′N; 75°58.00′W). Additional harvests were collected: *T. swinhoei* (HMA) at Hachijo Island, Japan (33°06′N, 139°47′E), *C. mycofijiensis* (HMA), collection numbers 02600 for CM2 and 02601 for CM5, offshore Vanuatu, Southern Pacific Ocean (17°34′S, 168°19′E), and *A. aerophoba* (HMA) and *D. avara* (LMA) offshore Rovinj, Croatia (45°05′N, 13°38′E) by scuba diving at depths of 5–15 m. Individuals were placed separately into plastic bags and brought to the surface. Immediately after collection, sponge tissues were cut into pieces and stored at −80 °C until further use.

DNA isolation and PCR screening: Genomic DNA was isolated from sponge samples with the Fast DNA spin kit for soil (Q-Biogene, Heidelberg, Germany) as described previously.[8] PCR screening was performed with the poribacterial 16S rRNA gene specific primers 389f and 1130r[8] and the KS gene fragment specific primers KSDPQQF and KSHGTGR.[8] The specificity of the poribacterial primers was established previously.[8,24] As an additional control, the PCR product from *A. dilatata* DNA was cloned and sequenced to confirm that it represents poribacterial 16S rRNA genes. PCR products were then cloned into pGEM-T easy (Promega) and transformed into electrocompetent *E. coli* Novablue cells. Colony PCRs on at least 48 clones per PCR product were performed at an annealing temperature of 45 °C using the vector primers SP6 and T7. The resulting PCR product was digested with HaeIII and MspI, and the plasmid insert was sequenced from at least one representative per RFLP pattern,[25] resulting in 3–10 different KS sequences per sponge. KS sequences were attributed to PKS groups by BlastX searches and phylogenetic analysis using BioEdit and ClustalX2.

454 sequencing of *C. mycofijiensis* PKS amplicons: For the two *C. mycofijiensis* specimens CM2 and CM5, the KS amplicon mixture was subjected to direct 454 sequencing. For this purpose, two GS FLX shotgun libraries were prepared and subjected to emulsion PCR following the instructions of the GS FLX Sample Preparation and emPCR Manuals (Roche Diagnostics, December 2007). Sequencing was done on two lanes of a 16-lane gasket per library on a 70 × 75 FLX picotiter plate of a Roche GS FLX sequencer, resulting in 15064 reads (3.1 Mb/average read length 208 bp) and 15409 reads (3.5 Mb/average read length 230 bp) for CM2 and CM5, respectively. Sequences were assembled independently for each DNA using the Roche gsAssembler with a minimal initial match value of 0.95 leading to 271 and 291 contigs for CM2 and CM5, respectively. After exclusion of all contigs consisting of $\leq$10 sequences and of low quality, 67 contigs for CM2 (mean length 218 bp, 11 to 803 sequences per contig, mean 100) and 84 contigs for CM5 (mean length 208 bp, 11 to 557 sequences per contig, mean 78) were used for further analyses.

Lipid analysis: Frozen (−80 °C) sponge specimens were ground in a mortar under liquid nitrogen and immediately extracted according to Bligh and Dyer.[26] Subsequently, the total lipids were separated over a silica gel column (Unisil, 100–200 mesh, Clarkson Chromatography Products, South Williamsport, PA, USA), prepared in a large Pasteur pipet, as described previously.[27] Briefly, neutral lipids (NL) were eluted with chloroform, glycolipids/sulfolipids (GL) with acetone, and polar lipds (PL) were eluted with methanol. The obtained fractions were methylated under alkaline conditions as described previously.[27b] Briefly, ~10 mg lipid fraction was diluted in 2 mL 0.2 M KOH (in dry MeOH) and incubated at 60 °C for 30 min. After cooling and adding 200 μL MeOH/acetic acid (9:1), FAMEs were extracted with hexane (2 × 3 mL). Pyrrolidides were made directly from FAMEs as described previously[15,28] by adding 1–2 mL pyrrolidine and heating for 1–2 h at 100 °C in a closed vial. Picolinyl esters were made by transesterification of the lipids as described previously.[29] Briefly, lipids were dissolved in 1 mL CH_2Cl_2 and added to 1 mL of a 2:1 mixture of 3-pyridinylmethanol (Sigma–Aldrich) and KO*t*Bu (1.0 M in THF, Sigma–Aldrich). Incubation was performed for 40 min at 50 °C.

Cloning of *sup* genes for in vivo studies: The genes were amplified by PCR using Expand High FidelityPLUS PCR system (Roche). The following primers were used: supB-forward (5′-AAA GAG CTC ACA TCA ACC GAA GAT CGC ATC-3′) and reverse (5′-AAA AAG CTT TCA CCC AGC GTG TGT ATC GAG GT-3′) for the ACP SupB, supA-ACP forward (5′- AAA GAG CTC CGG CTG CAG GAG ACG CCC GCC GCG GAG-3′) and reverse (5′-AAA AAG CTT TCA TAG TTC CTC CGC CAG ATG GCG GGC GAG-3′) for the ACP of SupA, supC forward (5′-AAA GAG CTC TTC GGC AAC GAT CGG CTC ATG GC-3′) and reverse (5′-AAA AAG CTT TCA TCG ATA CAC TCG CCC TGC AT-3′) for the PPTase SupC, and supA-AT1 forward (5′-AAA GAG CTC CGA CAG CGC GAG GCC CGT CTA CTG CCT-3′) and reverse (5′-AAA AAG CTT

TCA TAG CAA AGG GTG AAC GGC CCT TGG GCG-3′) for the AT_1 of SupA. All forward primers have a 5′-end extension with a restriction site for SacI, and the reverse primers have a HindIII extension followed by a stop codon (sites underlined). The ORFs were cloned without their native start codons. The PCR products were subcloned using T/A cloning[30] into pBluescriptII SK(+) (Stratagene) and sequenced at GATC (Konstanz, Germany) to confirm the correct sequence. The subclones were cut with SacI and HindIII and ligated into the expression vector pHIS8,[21] or pHIS8 Svp, respectively, to give plasmids pTH97 (SupA-ACP *holo* form), pHN80 (SupA-ACP *apo* form), pTH89 (SupB), pTH90 (SupC), and pTH86 (SupA-AT_1). pHIS8 Svp is derived from pHIS8 and carries the PPTase gene *svp* inserted into the *Not*I site of the vector.[31]

Protein expression: Proteins were as expressed as N-terminal His_8-tag fusion proteins. The ACPs SupB and SupA-ACP were expressed in *E. coli* BL21(DE3). For this purpose, 15 mL overnight cultures in LB medium supplied with 50 μg mL^{-1} kanamycin (LB-Km^{50}) were used to inoculate 600 mL of LB-Km^{50} and grown at 30–37 °C until an OD_{600} of ~0.6–1.0 was reached. The flasks were cooled to 16 °C, induced with 0.5 mM IPTG, and incubated further at 16 °C for ~16–20 h. The AT and KS-AT1 di-domain of SupA and the PPTase SupC were expressed in *E. coli* ArcticExpress™ (DE3) RIL (Stratagene). Overnight cultures (10 mL) in LB containing 20 μg mL^{-1} gentamycin, 50 μg mL^{-1} streptomycin, and 50 μg mL^{-1} kanamycin were used to inoculate 400 mL LB without antibiotics and grown at 30–37 °C until an OD_{600} of ~0.6–1.0 was reached. The flasks were cooled to 11–12 °C, induced with 0.5 mM IPTG, and incubated further at this temperature for additional 16–24 h. After expression, all cultures were harvested by centrifugation at 4 °C and 5000 rpm (9700 *g*). To obtain cell lysates, 2–6 mL lysis buffer (50 mM Tris-HCl, 500 mM NaCl, 10 mM $MgCl_2$, pH 8.0) were added to the cell pellets, and the mixture was sonicated and subsequently centrifuged for 20 min at 14000 rpm (7000 *g*) and 4 °C. The supernatants containing the soluble proteins were incubated with 200 μL Ni-NTA (Qiagen) for 1 h on ice with shaking, and then loaded onto Poly-Prep™ chromatography columns (Bio-Rad). Proteins were eluted from the column with lysis buffer at increasing imidazole concentrations. Aliquots (20 μL) were mixed with same volumes of 2×SDS loading buffer (according to the Qiagen protocol), incubated for 5 min at 99 °C, and monitored by SDS-PAGE.

Protein assays: For the in vitro phosphopantetheinylation assays, 160 μL ACP (~0.3 mg mL^{-1}), 160 μL PPTase (~0.1 mg mL^{-1}), and 4 μL CoA (10 μM final concentration) were incubated at 30 °C for 2 h in the buffer of the protein elution fractions (50 mM Tris-HCl, 500 mM NaCl, 10 mM $MgCl_2$, pH 8.0, 150–250 mM imidazole). For the AT-ACP assays, 4 μL DTT or TCEP (2 mM final concentration), 5 μL AT (~0.2 mg mL^{-1}), 9 μL ACP (~0.3 mg mL^{-1}), and 2 μL [2-^{14}C]malonyl-CoA (0.01 mCi mL^{-1}) or 1 μL [1-^{14}C]acetyl-CoA (0.02 mCi mL^{-1}) were incubated for 5 min at 20 °C or 30 s on ice in the buffer of the elution fractions. All reactions were stopped by adding the same volume of 2×SDS loading buffer and incubating for 5 min at 99 °C. All enzymes were dissolved in lysis buffer (see above), containing 150–250 mM imidazole. After separation by SDS-PAGE, the radioactivity was detected and visualized by overlaying the dried gel on a phosphorimager screen (Fuji), incubating it for 24–48 h in a cassette, and exposing the screen on the Phosphorimager Fujix BAS1000 (Raytest, Straubenhardt, Germany).

Accession number: The amplicon sequences have been deposited in GenBank with accession number SRA024372.

Acknowledgements

This work was supported by grants of the DFG (PI 430 1-3 and SFB 624 to J.P., and SFB 567 TP C3 to U.H.). We thank S. Matsunaga and N. Fusetani (University of Tokyo) for obtaining T. swinhoei*, P. Wenzel (University of California Santa Cruz) for the* C. mycofijiensis *samples, S. Zea for sponge identification, B. S. Moore for the plasmid pHIS8 Svp, J. J. Reddick and J. K. Williams for the* sfp *expression plasmid, S. Waldvogel for access to GC–MS, M. Engeser and C. Sondag for assistance with mass spectrometry, and L. Eggeling for assistance with expression in* C. glutamicum.

Keywords: fatty acids · metagenomics · polyketides · sponges · symbiosis

[1] G. D. Love, E. Grosjean, C. Stalvies, D. A. Fike, J. P. Grotzinger, A. S. Bradley, A. E. Kelly, M. Bhatia, W. Meredith, C. E. Snape, S. A. Bowring, D. J. Condon, R. E. Summons, *Nature* **2009**, *457*, 718.

[2] J. Piel, *Nat. Prod. Rep.* **2009**, *26*, 338.

[3] a) U. Hentschel, J. Hopke, M. Horn, A. B. Friedrich, M. Wagner, J. Hacker, B. S. Moore, *Appl. Environ. Microbiol.* **2002**, *68*, 4431; b) M.-L. Schläppy, S. Schöttner, G. Lavik, M. Kuypers, D. de Beer, F. Hoffmann, *Mar. Biol.* **2010**, *157*, 593; c) M. W. Taylor, R. Radax, D. Steger, M. Wagner, *Microbiol. Mol. Biol. Rev.* **2007**, *71*, 295.

[4] D. Sipkema, R. Osinga, W. Schatton, D. Mendola, J. Tramper, R. H. Wijffels, *Biotechnol. Bioeng.* **2005**, *90*, 201.

[5] a) J. Piel, D. Q. Hui, G. P. Wen, D. Butzke, M. Platzer, N. Fusetani, S. Matsunaga, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 16222; b) K. M. Fisch, C. Gurgui, N. Heycke, S. A. van der Sar, S. A. Anderson, V. L. Webb, S. Taudien, M. Platzer, B. K. Rubio, S. J. Robinson, P. Crews, J. Piel, *Nat. Chem. Biol.* **2009**, *5*, 494.

[6] a) J. Piel, D. Q. Hui, N. Fusetani, S. Matsunaga, *Environ. Microbiol.* **2004**, *6*, 921; b) L. Fieseler, U. Hentschel, L. Grozdanov, A. Schirmer, G. P. Wen, M. Platzer, S. Hrvatin, D. Butzke, K. Zimmermann, J. Piel, *Appl. Environ. Microbiol.* **2007**, *73*, 2144; c) A. Schirmer, R. Gadkari, C. D. Reeves, F. Ibrahim, E. F. Delong, C. R. Hutchinson, *Appl. Environ. Microbiol.* **2005**, *71*, 4840; d) A. Dobson, J. Kennedy, J. Marchesi, *J. Biotechnol.* **2007**, *131*, S215; e) J. Kennedy, C. E. Codling, B. V. Jones, A. D. W. Dobson, J. R. Marchesi, *Environ. Microbiol.* **2008**, *10*, 1888; f) T. K. Kim, J. A. Fuerst, *Environ. Microbiol.* **2006**, *8*, 1460.

[7] a) A. Siegl, U. Hentschel, *Environ. Microbiol. Reports* **2010**, *2*, 507; b) A. Siegl, J. Kamke, T. Hochmuth, J. Piel, M. Richter, C. G. Liang, T. Dandekar, U. Hentschel, *ISME J.* **2010**, doi:10.1038/ismej.2010.95.

[8] L. Fieseler, M. Horn, M. Wagner, U. Hentschel, *Appl. Environ. Microbiol.* **2004**, *70*, 3724.

[9] a) F. T. Gillan, I. L. Stoilov, J. E. Thompson, R. W. Hogg, C. R. Wilkinson, C. Djerassi, *Lipids* **1988**, *23*, 1139; b) V. Thiel, A. Jenisch, G. Wörheide, A. Löwenberg, J. Reitner, W. Michaelis, *Org. Geochem.* **1999**, *30*, 1; c) V. Thiel, M. Blumenberg, J. Hefter, T. Pape, S. Pomponi, J. Reed, J. Reitner, G. Wörheide, W. Michaelis, *Naturwissenschaften* **2002**, *89*, 60.

[10] a) P. Crews, Y. Kakou, E. Quinoa, *J. Am. Chem. Soc.* **1988**, *110*, 4365; b) Y. Kakou, P. Crews, G. J. Bakus, *J. Nat. Prod.* **1987**, *50*, 482; c) E. Quinoa, Y. Kakou, P. Crews, *J. Org. Chem.* **1988**, *53*, 3642; c) T. Amagata, T. A. Johnson, R. H. Cichewicz, K. Tenney, S. L. Mooberry, J. Media, M. Edelstein, F. A. Valeriote, P. Crews, *J. Med. Chem.* **2008**, *51*, 7234; d) T. A. Johnson, T. Amagata, K. V. Sashidhara, A. G. Oliver, K. Tenney, T. Matainaho, K. K.-H. Ang, J. H. McKerrow, P. Crews, *Org. Lett.* **2009**, *11*, 1975.

[11] S. Albrizio, P. Ciminiello, E. Fattorusso, S. Magno, *J. Nat. Prod.* **1995**, *58*, 647.

[12] a) A. Cutignano, A. Tramice, S. De Caro, G. Villani, G. Cimino, A. Fontana, *Angew. Chem.* **2003**, *115*, 2737; *Angew. Chem. Int. Ed.* **2003**, *42*, 2633; A. Cutignano, G. Cimino, A. Giordano, G. d'Ippolito, A. Fontana, *Tetrahedron Lett.* **2004**, *45*, 2627.

[13] a) J. Nechev, W. W. Christie, R. Robaina, F. M. de Diego, A. Ivanova, S. Popov, K. Stefanov, *Hydrobiologia* **2002**, *489*, 91; b) J. Nechev, W. W.

Christie, R. Robaina, F. d. Diego, S. Popov, K. Stefanov, *Eur. J. Lipid Sci. Technol.* **2002**, *104*, 800; c) N. M. Carballeira, N. Montano, J. Vicente, A. D. Rodriguez, *Lipids* **2007**, *42*, 519; d) V. Thiel, A. Jenisch, G. Wörheide, A. Löwenberg, J. Reitner, W. Michaelis, *Org. Geochem.* **1999**, *30*, 1; e) N. M. Carballeira, M. Pagan, A. D. Rodriguez, *J. Nat. Prod.* **1998**, *61*, 1049; f) A. Martínez, C. Duque, Y. Fujimoto, *Lipids* **1997**, *32*, 565.
[14] D. Hedrick, A. Peacock, P. Long, D. White, *Lipids* **2008**, *43*, 843.
[15] W. W. Christie, *The AOCS Lipid Library*, **2009**: http://lipidlibrary.aocs.org/index.html (accessed October 27, 2010).
[16] P. H. Buist, *Nat. Prod. Rep.* **2007**, *24*, 1110.
[17] a) C. T. Calderone, W. E. Kowtoniuk, N. L. Kelleher, C. T. Walsh, P. C. Dorrestein, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 8977; b) L. C. Gu, J. Y. Jia, H. C. Liu, K. Hakansson, W. H. Gerwick, D. H. Sherman, *J. Am. Chem. Soc.* **2006**, *128*, 9014.
[18] N. M. Carballeira, V. Negrón, E. D. Reyes, *J. Nat. Prod.* **1992**, *55*, 333.
[19] A. Schreiber, G. Worheide, V. Thiel, *Chem. Phys. Lipids* **2006**, *143*, 29.
[20] U. Hentschel, L. Fieseler, M. Wehrl, C. Gernert, M. Steinert, J. Hacker, M. Horn in *Molecular Marine Biology of Sponges* (Ed.: W. E. G. Müller), Springer, Heidelberg, **2003**, p. 60.
[21] J. M. Jez, J. L. Ferrer, M. E. Bowman, R. A. Dixon, J. P. Noel, *Biochemistry* **2000**, *39*, 890.
[22] R. H. Lambalot, A. M. Gehring, R. S. Flugel, P. Zuber, M. LaCelle, M. A. Marahiel, R. Reid, C. Khosla, C. T. Walsh, *Chem. Biol.* **1996**, *3*, 923.
[23] C. Sanchez, L. C. Du, D. J. Edwards, M. D. Toney, B. Shen, *Chem. Biol.* **2001**, *8*, 725.
[24] F. F. Lafi, J. A. Fuerst, L. Fieseler, C. Engels, W. W. L. Goh, U. Hentschel, *Appl. Environ. Microbiol.* **2009**, *75*, 5695.
[25] J. Sambrook, D. Russel, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, New York, **2001**.
[26] E. G. Bligh, W. J. Dyer, *Can. J. Biochem. Physiol.* **1959**, *37*, 911.
[27] a) H. L. Fredrickson, T. E. Cappenberg, J. W. Leeuw, *FEMS Microbiol. Lett.* **1986**, *38*, 381; b) D. C. White, D. B. Ringelberg in *Techniques in Microbial Ecology* (Eds.: R. S. Burlage, R. Atlas, D. Stahl, G. Geesey, G. Sayler), Oxford University Press, New York, **1998**, 225.
[28] W. W. Christie, *Lipid Analysis*, 3rd ed., P. J. Barnes & Associates, Bridgewater, **2003**.
[29] F. Destaillats, P. Angers, *J. Am. Oil Chem. Soc.* **2002**, *79*, 253.
[30] D. Marchuk, M. Drumm, A. Saulino, F. S. Collins, *Nucleic Acids Res.* **1991**, *19*, 1154.
[31] W. Izumikawa, Q. Cheng, B. S. Moore, *J. Am. Chem. Soc.* **2006**, *128*, 1428.
[32] J. Nechev, W. W. Christie, R. Robaina, F. de Diego, S. Popov, K. Stefanov, *Eur. J. Lipid Sci. Technol.* **2002**, *104*, 800.
[33] N. M. Carballeira, L. Maldonado, *Lipids* **1988**, *23*, 682.

Received: August 30, 2010
Published online on November 12, 2010

