Title

A Family of Nonribosomal Peptides Modulate Collective Behavior in Pseudovibrio Bacteria Isolated from Marine Sponges**

Journal

Angewandte Chemie International Edition, 60(29)

ISSN

1433-7851

Authors

Ióca, Laura P
Dai, Yitao
Kunakom, Sylvia
et al.

Publication Date

2021-07-12

DOI

10.1002/anie.202017320

Peer reviewed

HHS Public Access

Author manuscript

Angew Chem Int Ed Engl. Author manuscript; available in PMC 2022 July 12.

Published in final edited form as:

Angew Chem Int Ed Engl. 2021 July 12; 60(29): 15891–15898. doi:10.1002/anie.202017320.

A novel family of nonribosomal peptides modulate collective behavior in *Pseudovibrio* bacteria isolated from marine sponges

Dr. Laura P. Ióca[1,2,4], **Yitao Dai**[1,2], **Dr. Sylvia Kunakom**[1,2], **Jennifer Diaz-Espinosa**[1,2], **Dr. Aleksej Krunic**[1], **Dr. Camila M. Crnkovic**[1,2], **Prof. Dr. Jimmy Orjala**[1,2], **Prof. Dr. Laura M. Sanchez**[1], **Antonio G. Ferreira**[3], **Roberto G. S. Berlinck**[4,*], **Prof. Dr. Alessandra S. Eustáquio**[1,2,*]

[1]Department of Pharmaceutical Sciences, College of Pharmacy, University of Illinois at Chicago, Chicago, IL 60607, USA

[2]Center for Biomolecular Sciences, College of Pharmacy, University of Illinois at Chicago, Chicago, IL 60607, USA

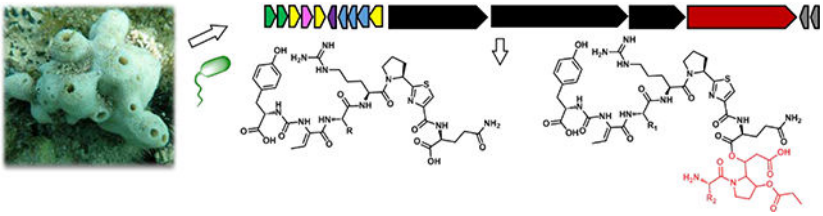
[3]Departamento de Química, Universidade Federal de São Carlos, São Carlos, SP 13565-905, Brazil

[4]Instituto de Química de São Carlos, Universidade de São Paulo, São Carlos, SP 13560-970, Brazil

Abstract

Although swarming motility and biofilms are opposed collective behaviors, both contribute to bacterial survival and host colonization. *Pseudovibrio* bacteria have attracted attention because they are part of the microbiome of healthy marine sponges. Two-thirds of *Pseudovibrio* genomes contain a member of a nonribosomal peptide synthetase-polyketide synthase gene cluster family, which is also found sporadically in *Pseudomonas* pathogens of insects and plants. After developing reverse genetics for *Pseudovibrio*, we isolated heptapeptides with an ureido linkage and related nonadepsipeptides we termed pseudovibriamides A and B, respectively. A combination of genetics and imaging mass spectrometry experiments showed heptapetides were excreted, promoting motility and reducing biofilm formation. In contrast to lipopeptides widely known to affect motility/biofilms, pseudovibriamides are not surfactants. Our results expand current knowledge on metabolites mediating bacterial collective behavior.

Graphical Abstract



[*]**To whom correspondence should be addressed**: rgsberlinck@iqsc.usp.br, and ase@uic.edu.

Pseudovibriamides. *Pseudovibrio* are α-Proteobacteria proposed to contribute to marine sponge health. Lack of reverse genetics have hindered exploration of the ecological and biotechnological potential of *Pseudovibrio*. Here we describe reverse genetics methods that enabled identification of 12 pseudovibriamides, which we show affect motility and biofilm behavior of *Pseudovibrio*.

Keywords

natural products; peptides; bacterial metabolites; swarming motility; biofilm formation; *Pseudovibrio*; *Pseudomonas*

Introduction

Collective behavior is a common feature of biological systems. Bacterial swarming is the collective motion of bacterial cells on a solid surface, which is powered by rotating flagella.[1,2] In contrast, biofilms are nonmotile, self-organized cellular aggregates bounded by extracellular polymers.[3] Although bacterial biofilms represent a concept opposite to motility, both contribute to population survival and host colonization.[4,5] Identifying which specialized metabolites mediate these processes will bring us closer to understanding collective, coordinated behavior in bacteria.

Herein we report a link between swarming and biofilms, and a nonribosomal peptide synthetase (NRPS) polyketide synthase (PKS) gene cluster family that is found in both *Pseudovibrio* and *Pseudomonas* bacteria known to colonize animals and plants (Fig. 1). We termed this biosynthetic gene cluster (BGC) *ppp* for *P*seudovibrio and *P*seudomonas nonribosomal peptide (Fig. 2).

Pseudomonas is a genus of γ-Proteobacteria discovered in the 19th century. *Pseudomonas* species can be free-living or associated with eukaryotic hosts including fungi, plants, and animals, with some species being pathogenic to their hosts.[6] In contrast, the genus *Pseudovibrio* of marine α-Proteobacteria was only recently described in 2004.[7] *Pseudovibrio* bacteria have been predominantly isolated from healthy marine sponges.[8] The current hypothesis is that *Pseudovibrio* spp. form a mutualistic relationship with marine sponges, although details of the interaction remain to be elucidated.[8,9] The identification of specialized metabolites produced by *Pseudovibrio* is an important step towards understanding *Pseudovibrio*'s ecological roles. For instance, *Pseudovibrio* spp. produce the antibiotic tropodithietic acid[10] which is active against sponge pathogenic, marine *Vibrio* spp..[11] Additional metabolites reported from *Pseudovibrio* include the cytotoxic 2-methylthio-1,4-naphthoquinone and indole alkaloid tetra(indol-3-yl)ethanone, bromotyrosine-derived alkaloids, the red pigment heptyl prodiginin, and anti-fouling diindol-3-ylmethanes.[8,12–15]

In this investigation we report the complete genome of *Pseudovibrio brasiliensis* strain Ab134, the establishment of reverse genetics methods for this marine sponge-derived isolate, the discovery and structures of a novel family of nonribosomal peptides which we named pseudovibriamides (**1–12**, Fig. 2), and pseudovibriamides' role in affecting flagellar motility and biofilm formation. Using *Pseudovibrio* as a model, our results reveal novel specialized

metabolites employed by bacteria to modulate motility and biofilms, ultimately expanding current knowledge of bacteria collective behavior.

Results

The *ppp* BGC from *Pseudovibrio* is located in a plasmid.

P. brasiliensis strain Ab134 was isolated from the sponge *Arenosclera brasiliensis*.[16] A draft genome had previously been reported.[17] Here, we re-sequenced *P. brasiliensis* strain Ab134 using short-read Illumina and long-read Nanopore sequencing. Hybrid assembly followed by confirmation of plasmids via PCR (Fig. S1 and Tables S1–S2) led to a genome of 6.03 Mb divided into one chromosome and five plasmids (Fig. 3 and Table S3). Seven biosynthetic gene clusters were identified using antiSMASH (Table S4),[18] including the NRPS-PKS *ppp* gene cluster located on plasmid 2 (Fig. 3). Using BLAST analyses we further identified a *tda* gene cluster putatively encoding the biosynthesis of the known antibiotic tropodithietic acid[19] on plasmid 1 (Fig. 3).

The *ppp* BGC is found in eukaryote-interacting bacteria.

Bioinformatics analyses using multigene blast[20] revealed the *ppp* BGC to be conserved in various species of *Pseudovibrio* and also found in *Pseudomonas*, although less frequently (Fig. 1c, Fig. S2, and Table S5). All members of the *ppp* BGC family include one NRPS-PKS gene (*pppD*) and either three or two NRPS genes (*pppABC* or *pppBC*, respectively, Fig. 1b). Other conserved genes encode a transporter (*pppG*), three proteins of unknown function (*pppHIJ*), and an iron/α-ketoglutarate-dependent dioxygenase (*pppK*). Additional transporter genes (e.g. *pppL* and *pppN*) and regulatory genes (e.g. *pppP* and *pppO* encoding a two-component system sensor histidine kinase and response regulator, respectively) are also found flanking the clusters (Fig. 1b). Moreover, the BGC present in *Pseudovibrio* also encodes a phosphopantheteinyl transferase (*pppE*), and a type II thioesterase (*pppF*) not present in *Pseudomonas* BGCs, and a *N*-acetyltransferase family protein (*pppM*) that is present only in the short version of the *Pseudomonas* BGC (Fig. 1b).

The larger version of the *ppp* BGC is found in 16 out of 24 *Pseudovibrio* genomes available in GenBank, 13 of which have been isolated from marine sponges, two from seawater, and one from a marine tunicate. In addition, this larger version of the *ppp* BGC was also found in at least two *Pseudomonas* species known to be plant pathogens, *Pseudomonas asplenii* ATCC23835 and *Pseudomonas fuscovaginae* LMG2158. The shorter version of the *ppp* BGC was more common in *Pseudomonas* than the longer one and was found in about 4% of *Pseudomonas* genomes (25 out of 602 complete genomes analyzed) including plant pathogen *P. syringae* pv. *lapsa* ATCC10859, and insect pathogen *P. entomophila* L48. Based on the NRPS-PKS domain organization (Fig. 2), we expected the full-length product to contain nine amino acids and one acetate unit. The amino acid prediction for all *ppp* BGCs is highly conserved (Fig. S3). The domain organization found in the *ppp* assembly line is reminiscent of the rimosamide NRPS-PKS BGC of *Streptomyces rimosus* (Fig. S2, BGC 19–21), which likewise contains two thioesterase domains and a terminal, AT-less PKS module.[21]

Establishment of reverse genetics for *P. brasiliensis*.

In order to aid with compound identification, and to probe the function of the *ppp* BGC, we developed reverse genetics for *P. brasiliensis* Ab134. No methods for genetic engineering of *Pseudovibrio* had been previously reported. After establishing procedures based on protocols available for other members of the Rhodobacteraceae family,[22,23] mutants in which the *pppA* gene was replaced with a kanamycin-resistance marker via homologous recombination were obtained (Fig. S4a–b). Comparative metabolite analysis of the wild-type and three independent mutants was then employed to identify the encoded products.

Identification of pseudovibriamides.

Comparative metabolite analysis of the wild-type and *pppA* mutant strains using mass spectrometry led to the identification of five signals that were detected only in the wild-type Ab134 strain (Fig. 4a) suggesting they were the products of the *ppp* gene cluster. The two larger compounds of *m/z* 1156 and 1170 differ only by a methyl group (Δ14 Da). Tandem mass spectrometry (LC-MS/MS) analyses (Fig. S5) showed *m/z* 844 as a fragment of the two larger compounds, indicating their relatedness. The other two smaller compounds, *m/z* 826 and *m/z* 858 differ from *m/z* 844 only by the removal of water (−Δ18 Da) and the addition of a methyl group (+Δ14 Da), respectively.

To interrogate metabolites with *m/z* ~800, we next generated *pppD* deletion mutants. As shown in Fig. 2, we hypothesized metabolites with *m/z* ~800 to be the products of modules 1–7 encoded in *pppABC*. If true, deletion of *pppD* should only abolish production of the full-length products. *pppD* mutants were generated by replacement with a kanamycin-resistance marker as done for *pppA* (Fig. S4c). Mass spectrometry confirmed that *pppD* mutants do not produce the full-length products and accumulate only the metabolites with *m/z* ~800 (Fig. 4a) which, as we will describe later, were confirmed to be heptapeptides. We named the heptapeptides pseudovibriamides A and the full-length products pseudovibriamides B.

Pseudovibriamide A is excreted and pseudovibriamide B is largely cell-associated.

Investigation of the spatial distribution of pseudovibriamides in microbial colonies by MALDI-TOF imaging mass spectrometry (IMS) analyses (Fig. 4b) indicated that the full-length products (*m/z* 1156 and 1170) appeared to be mainly colony-associated whereas the heptapeptide with *m/z* 844 were excreted. This observation was corroborated by our inability to identify the full-length products in liquid culture supernatants whereas the heptapeptides are readily identified.

Isolation and structure elucidation of pseudovibriamides A.

After testing four different growth conditions, growth on solid media provided superior pseudovibriamide relative yields compared to liquid cultures (Fig. S6). Three growth experiments and extraction procedures were carried out to obtain enough material for isolation and structure elucidation. We first extracted 400 plates (8-L, extract F). After several rounds of fractionation and HPLC purification, we obtained 5.5 mg of **1** and 2.7 mg of **2**. The HRESIMS (Fig. S7) analysis of **1** displayed a $[M + H]^+$ at *m/z* 844.3389,

corresponding to a molecular formula $C_{36}H_{50}N_{11}O_{11}S$ (calc. 844.3412, error −2.3 ppm). Inspection of 1D and 2D NMR data for **1** (Figs. S8–S15, Table S6) allowed the identification of five amino acids as Tyr, Ala, Arg, Pro and Gln. Two additional modified residues were identified. First, a dehydrated Thr (dehydroaminobutyric acid, *Z*-Dhb), based on a COSY correlation between the methyl group at δ_H 1.59 (d) and the vinylic methine hydrogen at δ_H 5.95 (q), as well as by observing a ROESY correlation between the methyl resonance at δ_H 1.59 with NH at δ_H 7.81, pointing to a *Z*-DhB residue. Second, a thiazole group (thz), based on HMBC coupling from H_β (thz, δ_H 8.16, s) to C_α (Pro, δ_C 58.5), C_α (thz, δ_C 149.1), C_δ (thz, δ_C 173.7) and the Gln amide carbonyl group at δ_C 173.2. The connectivity between the amino acid residues was confirmed by observation of HMBC correlations (Fig. S16) and by MS^2 analysis (Fig. S7, and Table S7). Inspection of the MS^2 fragmentation pattern (Table S7) indicated the loss of a complete Tyr residue (*m/z* of 663.2673). A second ion at *m/z* 637.2866 indicated the loss of a Tyr-CO fragment. Such fragment losses would be possible due to the presence of an ureido linkage, leading to a reversal of chain polarity. HMBC correlations from the amine (NH-Tyr, δ_H 6.42) and $C_\alpha H_2$ (Tyr, δ_H 4.30) to C_{CO} (δ_C 155.2), in addition to ^{15}N-HMBC (Fig. S15) correlations from the NH_{Tyr} (δ_H 6.42) to the *N*-ureido nitrogen at δ_N 88.51, as well as from the NH_{Dhb} (δ_H 7.81) to the second N-ureido nitrogen at δ_N 92.17 (Table S6), confirmed our hypothesis. The absolute configurations of all amino acids was established as L by advanced Marfey's analysis (Figs. S17–21),[24,25] thus completing the structure of **1**, named pseudovibriamide A1 (Fig. 2).

Fractionation of extract FD (12-L, 600 plates) provided 2 mg of **2** and additional 8.8 mg of **1**. Pseudovibriamide A2 (**2**, Fig. 2) samples obtained from the two growth experiments were combined. The HRESIMS (Fig. S22) analysis of **2** displayed a $[M + H]^+$ at *m/z* 826.3311, corresponding to a molecular formula $C_{36}H_{48}N_{11}O_{10}S$ (calc. 826.3306, error 0.6 ppm), suggesting **1** undergoes dehydration and/or cyclization leading to **2**. Inspection of the MS^2 fragmentation pattern (Fig. S22, Table S8) indicated a cyclization between Tyr and *Z*-Dhb, leading the formation of a unique imidazolidinyl-dione ring. Further analysis of 1D and 2D NMR data for **2** (Figs. S23–S29, Table S9) and advanced Marfey's analysis allowed the confirmation of L-Tyr, *Z*-Dhb, L-Ala, L-Arg, L-Pro-thz and L-Gln (Figs. S30–S34).

Further fractionation of the FD extract led to the isolation of 1 mg of **3** and 0.2 mg of **4** (Fig. 2). Analysis of the MS^2 fragmentation (Tables S10–S11, Figs. S35–S36) and advanced Marfey's analysis (Figs. S37–S45) of pseudovibriamide A3 (**3**) and A4 (**4**) indicated the presence of a Gly instead of Ala (Fig. S37) as the only difference between **1** and **2**. We also obtained 0.5 mg of a mixture of **5** and **6** containing two chromatographic peaks with $[M + H]^+$ at *m/z* 858.35 (Fig. S46) and molecular formula $C_{37}H_{52}N_{11}O_{11}S$, indicating the presence of a methylated form of **1**. The MS^2 fragmentation (Figs. S47–S48, Table S12) supports the structures **5** and **6** (Fig. 2), methylated at Gln and Tyr, respectively. Although these compounds are possibly artifacts of MeOH extraction, such hypothesis requires further experimental confirmation.

Isolation and structure elucidation of pseudovibriamides B.

To isolate the full-length products we generated extract C from cells only (20-L, 1000 plates, Fig. S49). MS-guided purification provided 5.9 mg of **7**, 6.2 mg of a mixture of **8** and **9**, 2.7

mg of **10** and 1.4 mg of a mixture of **11** and **12.** Unfortunately, pseudovibriamides B degraded during the acquisition of NMR data. Degradation was also observed for pseudovibriamide A1, since the bond between NH_{Dhb} and $C_{\alpha'\text{-}Dhb}$ is easily hydrolyzed (Figs. S50–51, Table S13). The HRESIMS (Fig. S52) analysis of **7** displayed a $[M + H]^+$ at *m/z* 1156.5106, corresponding to a molecular formula $C_{51}H_{74}N_{13}O_{16}S$ (calc. 1156.5097, error 0.8 ppm). Its degradation product (Fig. S53) displayed $[M + H]^+$ at *m/z* 950.4446, corresponding to the molecular formula $C_{41}H_{64}N_{11}O_{13}S$ (calc. 950.4406, error 4.2 ppm). The MS^2 data indicated that peptide **7** shares the carbon skeleton of **1** (Fig. S7); therefore, we could identify the presence of L-Tyr, *Z*-Dhb, L-Ala, L-Arg, L-Pro-thz and L-Gln by NMR analysis (Figs. S54–S60, Table S14) and confirm the presence of L-amino acids by Marfey's analysis (Figs. S61–S66). Analysis of 1D and 2D NMR spectra of **7** and advanced Marfey's analysis indicated the presence of a L-Val residue (Fig. S66). Based on homology between the *ppp* BCG and the rimosamide BGC, we hypothesized that **7** also presented a modified hydroxyproline (Hyp) residue. An HMBC correlation between the CH of δ_H at 5.10 (δ_C 70.4) and CO (173.0) indicated a propionylated Hyp fragment. The correlation between $CH_{\alpha'\text{-}Hyp}$ at δ_H 5.33 (δ_C 69.3) and CO_{Gln} at δ_C 170.8 confirmed the final nonadepsipeptide structure of **7** (Fig. 2), named as pseudovibriamide B1. Additional confirmation of the amino acid sequence was established by MS^2 analysis (Fig. S52 and Table S15).

Comprehensive and detailed analysis of MS and MS^2 data of pseudovibriamides B2-B6 (Fig. S67, S77–S78, Table S16–S18) together with advanced Marfey's analysis (Fig. S68–S76, S79–S92) led to the proposed structures for **8–12**, respectively (Fig. 2).

Pseudovibriamides have no apparent antibacterial activity.

P. brasiliensis Ab134 wild-type has been reported to have antibacterial activity against *Vibrio* spp..[16] We observed the same level of activity against *V. fisheri* as test strain for the wild-type and the *pppA*$^-$ mutant strains, indicating that pseudovibriamides are not responsible for the antibacterial activity of *P. brasiliensis* Ab134 (Fig. S93).

Pseudovibriamides modulate motility and biofilm formation.

We observed the diameter of *pppA* mutant colonies to be significantly smaller than wild-type colonies grown on solid agar (Fig. S93), and hypothesized pseudovibriamides to be involved in surface motility. The only surface motility mode[1,2] that has been previously described for both α- and γ-Proteobacteria is swarming. Yet, *Pseudovibrio*'s ability to swarm had not been previously reported. *Pseudovibrio* spp. do contain one to several lateral or subpolar flagella[7,26] so that the minimum requirement for swarming motility is met. Indeed, we were able to show that the wild-type Ab134 is capable of swarming (Fig. S94). Moreover, we found that Eiken agar is required to allow observation of the swarming phenotype, as swarming could not be observed on more commonly used agars such as Bacto agar, most likely due to Eiken agar's lower surface tension (Fig. S94). Agars of inherently low surface tension have been shown to be required for swarming of bacteria that do not produce surfactants.[1] Based on their structures, no surfactant activity is expected for pseudovibriamides, and no surfactant activity was observed for strain Ab134 using two accepted surfactant assays[27] – drop collapse and oil spreading (Fig. S95). Finally, Ab134 swarming motility decreases with increasing agar concentrations, as has been observed for

other temperate swarmers (Fig. S96).[2,28] Having established swarming assays using 0.5% (*w/v*) Eiken agar, we were then able to show that the *pppA* mutant has an impaired swarming phenotype (Fig. 5a).

Swimming is an alternative mode of bacterial motility that is likewise powered by flagella but that involves moving in liquid rather than on solid surfaces.[1] We investigated whether swimming was also affected by the *pppA* deletion. A swimming assay showed that loss of pseudovibriamides A and B also impaired swimming motility (Fig. 5a). All three independent *pppA* mutants analyzed showed the same impaired motility phenotype (Fig. S97). To make sure that no unintended mutations were introduced during mutant generation, we sequenced the genome of one of the mutants using Illumina sequencing. Mapping of Illumina reads to our reference genome showed that no further mutations were present besides the *pppA* deletion. To provide further evidence for the link between the *ppp* gene cluster, pseudovibriamides, and the motility phenotype, a genetic complementation mutant was generated in which *pppA* was knocked back in the *pppA*$^{-}$ mutant, i.e. the kanamycin resistance marker was replaced with *pppA* via homologous recombination leading to a reversal to the original wild-type genotype (Fig. S98). Pseudovibriamide production and the swarming phenotype were restored in the *pppA* knock-in mutant (Fig. S98).

Motility and biofilm development are considered opposite behaviors.[29,30] Thus, we employed an established biofilm assay[31] to evaluate the effect of the *pppA* mutation on the ability of *Pseudovibrio* to form biofilms. We observed biofilm adherence and microcolony formation to be increased in the *pppA* mutant compared to the wild-type (Fig. 5b and Table S19). In contrast, flagellar motility of *pppD* mutants that still produce heptapeptides pseudovibriamides A was comparable to the wild-type (Fig. 5a), whereas biofilm adherence and microcolony formation was reduced in *pppD* mutants (Fig. 5b and Table S19).

Discussion

Microbes play crucial roles in the function and health of plants and animals.[32,33] Marine sponges are the most ancient animals and are a dominant component of our oceans from reef systems to polar sea floors.[34] Microorganisms can comprise up to one-third of sponge biomass.[34] Ecological questions such as the importance of microbes to sponge health and the role of specialized metabolites produced by sponge microbes are especially relevant since sponge disease negatively affects the ecology of reef systems.[35] *Pseudovibrio* is a genus of α-Proteobacteria found in healthy sponges but absent from diseased ones, and it has been hypothesized that *Pseudovibrio* spp. contribute to marine sponge health.[8] Lack of reverse genetics methods have hindered full exploration of the ecological and biotechnological potential of *Pseudovibrio* bacteria.[8] The reverse genetics procedures described here not only enabled identification of pseudovibriamides but also allowed us to identify their role in motility and biofilm formation.

The widespread nature of the *ppp* BGC in *Pseudovibrio* spp. suggests that pseudovibriamides play an important ecological role in this genus. Our results show that the *ppp* gene cluster is located in a plasmid of *P. brasiliensis* Ab134 (Fig. 3). Although most *Pseudovibrio* genomes are draft genomes preventing determination of replicon count, the

ppp BGC also appears to be present in plasmids of at least five other *Pseudovibrio* spp. (Table S20), due to the presence of replication genes *repB* or *repC* in the same contig. In contrast, the *ppp* BGC is located in the chromosome of *Pseudomonas* strains (Fig. S99). Plasmid localization facilitates horizontal gene transfer and may help explain the wider distribution of the *ppp* BGC in *Pseudovibrio* compared to *Pseudomonas* (Fig. 1c). Moreover, it is interesting to note that the *ppp* BGC is located next to NRPS genes encoding lipopeptides in the two *Pseudomonas* species that contain the full-length *ppp* cluster (Fig. S99). *Pseudomonas* lipopeptides are known to influence swarming and biofilm formation at least in part due to their surfactant activity.[36] Physical clustering of the *ppp* BGC with lipopeptide-encoding NRPS BGCs may facilitate co-regulation of different metabolites that are involved in the same processes.

The sequence of amino acids observed in pseudovibriamide's structures agrees with the NRPS module organization (Fig. 2 and Fig. S3). The amino acid that could not be predicted was identified as L-Ala (module 3), for which the specificity code shows little conservation.[37] Gly can also be substituted at this position as in **3**, **4**, and **10**. Moreover, three amino acid modifications are present in pseudovibriamides, i.e. heterocyclization of cysteine followed by oxidation leading to a thiazole ring, reversal of tyrosine's chain polarity via an ureido linkage, and dehydration of Thr to yield Dhb acid.

Thiazole biosynthesis in nonribosomal peptides is well documented and involves a heterocyclization (Cy) domain and an oxidoreductase domain,[38] both of which are present in the respective module encoded in *pppB* (Fig. 2). Several peptides harbor a reversal in chain polarity through an ureido linkage (Table S21). The formation of this unusual connection has been investigated *in vitro*, indicating that the C domain catalyzes the reaction using bicarbonate/CO_2.[39,40] Phylogenetic analysis of domains encoded in *pppA* from *Pseudovibrio* and *Pseudomonas* revealed that the first C domain (at module 2) clades together with condensation domains known to lead to ureido linkages (Fig. S100), supporting the ureido functionality determined from spectroscopic data. The implications of this structural feature for the phenotypes associated with the pseudovibriamides are, as of yet, unknown.

Finally, for all full-length *ppp* BGCs, the second A domain present in PppA is predicted to load Thr (Fig. S3). Our spectroscopic analyses of pseudovibriamides supports dehydration of Thr to yield Dhb acid. None of the *ppp* standalone genes are predicted to catalyze dehydration. However, phylogenetic analysis of the corresponding C domain of module 3 shows that it clades together with C domains that likewise process dehydroamino acids in the biosynthesis of mycrocystin and nodularin from cyanobacteria (Fig. S101, Table S22). The C domain of McyA (module 2) has been hypothesized to play a role in dehydration of Ser and Thr leading to microcystins containing dehydroalanine (Dha) and Dhb, respectively.[41–44] Thus, although experimental evidence is outstanding, the *ppp* NRPS appears to contain three C domains that putatively catalyze reactions other than simple amino acid condensation, i.e. reversal of chain polarity via an ureido linkage, dehydration, and cyclization (Fig. 2).

Another intriguing and unique feature of pseudovibriamides **2**, **4**, **11**, and **12** is the presence of an imidazolidinyl-dione ring. While this heterocycle is commonly found in marine-derived alkaloids (Table S23), to the best of our knowledge, it has not been reported as a bacterial peptide residue. Additional studies are required to unveil how this cyclization proceeds.

The structure of pseudovibriamide B is reminiscent of depsipeptide rimosamides from *Streptomyces rimosus*.[21] The rimosamide BGC likewise contains a NRPS-PKS system with two TE domains in which RmoI encodes a trimodular NRPS-PKS that loads Val/Ile, Pro and malonate, respectively, analogously to PppD (Fig. 2). One of the TE domains may be responsible for transesterification leading to pseudovibriamides B as recently reported for the unrelated depsipeptide FR900359.[45] Both the *ppp* and *rmo* clusters also encode putative iron/α-ketoglutarate-dependent dioxygenases (*pppK* and *rmoL*) which may be responsible for hydroxylation of Pro. The enzyme responsible for alkylation of Hyp has not been identified in either case. The *ppp* gene cluster of *P. brasiliensis* Ab134 and the short version of the *Pseudomonas ppp* BGC both contain a gene with sequence similarity to GNAT family proteins as candidate to catalyze propyl transfer.

Our results show that heptapeptides such as **1** (*m/z* 844 in Fig. 4b) are excreted to promote flagellar motility while reducing biofilm formation (Fig. 5), whereas nonadepsipeptides such as **7** and **8**/**9** (*m/z* 1156 and 1170 in Fig. 4b) are cell-associated. It is unclear whether the excreted heptapeptides are direct products of module 7 or hydrolysis products of the nonadepsipeptides (Fig. 2).

In conclusion, collective behaviors such as swarming motility and biofilm formation are beneficial for bacterial survival and for the colonization of eukaryotic hosts[2,5]. The majority of nonribosomal peptides shown to be implicated in swarming motility and biofilms are lipopeptides having surfactant activity.[46,47] In contrast, pseudovibriamides have distinct structures and are not surfactants. Identification of pseudovibriamides and their link to motility and biofilm formation opens the door to future studies on how they mediate these processes, ultimately expanding current knowledge of bacterial collective behavior.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Dr. F. Thompson (Federal University of Rio de Janeiro) for the *Pseudovibrio brasiliensis* strain Ab134 and for sharing genome data prior to publication, Dr. D. Schneider (Grenoble Alpes University) for the pDS132 vector, UIC's Sequencing Core for genome sequencing, Dr. G. Pauli (UIC) for access to QToF-MS used in preliminary studies, and Dr. T. Molinski (University of California San Diego) for the FDTA Marfey's reagent. Genome sequencing and assembly was performed by the UIC Research Resources Center, supported in part by NCATS through Grant UL1TR002003. Financial support for this work was provided by the Division of Integrative Organismal Systems of the National Science Foundation under grant IOS-1917492 (to ASE), by the São Paulo Research Foundation FAPESP grants 2013/50228-8 and 2015/01017-0 (to RGSB), by FAPESP scholarships 2016/05133-7 and 2018/10742-8 (to LPI), National Institutes of Health National Institute of General Medical Sciences R01GM125943 (to LMS) and by startup funds from the Department of Pharmaceutical Sciences (to ASE and LMS) and the Center for Biomolecular Sciences, University of Illinois at Chicago (to ASE).

References

[1]. Kearns DB, Nat. Rev. Microbiol 2010, 8, 634–644. [PubMed: 20694026]

[2]. Mattingly AE, Weaver AA, Dimkovikj A, Shrout JD, J. Bacteriol 2018, 200, e00014–18. [PubMed: 29555698]

[3]. Flemming HC, Wuertz S, Nat. Rev. Microbiol 2019, 17, 247–260. [PubMed: 30760902]

[4]. Verstraeten N, Braeken K, Debkumari B, Fauvart M, Fransaer J, Vermant J, Michiels J, Trends Microbiol. 2008, 16, 496–506. [PubMed: 18775660]

[5]. Gloag ES, Marshall CW, Snyder D, Lewin GR, Harris JS, Santos-Lopez A, Chaney SB, Whiteley M, Cooper VS, Wozniaka DJ, MBio 2019, 10, e01698–19. [PubMed: 31409682]

[6]. Peix A, Ramírez-Bahena MH, Velázquez E, Infect. Genet. Evol 2018, 57, 106–116. [PubMed: 29104095]

[7]. Shieh WY, Te Lin Y, Jean WD, Int. J. Syst. Evol. Microbiol 2004, 54, 2307–2312. [PubMed: 15545476]

[8]. Romano S, Appl. Environ. Microbiol 2018, 84, e02516–17. [PubMed: 29453252]

[9]. Alex A, Antunes A, PLoS One 2018, 13, e0194368. [PubMed: 29775460]

[10]. Harrington C, Reen FJ, Mooij MJ, Stewart FA, Chabot JB, Guerra AF, Glöckner FO, Nielsen KF, Gram L, Dobson ADW, O'Gara F, Mar. Drugs 2014, 12, 5960–5978. [PubMed: 25513851]

[11]. Duan Y, Petzold M, Saleem-Batcha R, Teufel R, ChemBioChem 2020, 23, 2384.

[12]. Sertan-De Guzman AA, Predicala RZ, Bernardo EB, Neilan BA, Elardo SP, Mangalindan GC, Tasdemir D, Ireland CM, Barraquio WL, Concepcion GP, FEMS Microbiol. Lett 2007, 277, 188–196. [PubMed: 18031339]

[13]. Wang KL, Xu Y, Lu L, Li Y, Han Z, Zhang J, Shao CL, Wang CY, Qian PY, Mar. Biotechnol 2015, 17, 624–632.

[14]. Rodrigues AMS, Rohée C, Fabre T, Batailler N, Sautel F, Carletti I, Nogues S, Suzuki MT, Stien D, Tetrahedron Lett. 2017, 58, 3172–3173.

[15]. Nicacio KJ, Ióca LP, Fróes AM, Leomil L, Appolinario LR, Thompson CC, Thompson FL, Ferreira AG, Williams DE, Andersen RJ, Eustaquio AS, Berlinck RGS, J. Nat. Prod 2017, 80, 235–240. [PubMed: 28191971]

[16]. Rua CPJ, Trindade-Silva AE, Appolinario LR, Venas TM, Garcia GD, Carvalho LS, Lima A, Kruger R, Pereira RC, Berlinck RGS, Valle RAB, Thompson CC, Thompson F, PeerJ 2014, 2014, e419.

[17]. Fróes AM, Freitas TC, Vidal L, Appolinario LR, Leomil L, Venas T, Campeão ME, Silva CJF, Moreira APB, Berlinck RGS, Thompson FL, Thompson CC, Front. Mar. Sci 2018, 5, 81.

[18]. Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, Medema MH, Weber T, Nucleic Acids Res. 2019, 47, 81–87.

[19]. Geng H, Bruhn JB, Nielsen KF, Gram L, Belas R, Appl. Environ. Microbiol 2008, 74, 1535–1545. [PubMed: 18192410]

[20]. Medema MH, Takano E, Breitling R, Mol. Biol. Evol 2013, 30, 1218–1223. [PubMed: 23412913]

[21]. McClure RA, Goering AW, Ju KS, Baccile JA, Schroeder FC, Metcalf WW, Thomson RJ, Kelleher NL, ACS Chem. Biol 2016, 11, 3452–3460. [PubMed: 27809474]

[22]. Dziewit L, Czarnecki J, Prochwicz E, Wibberg D, Schlüter A, Pühler A, Bartosik D, Front. Microbiol 2015, 6, 852. [PubMed: 26347732]

[23]. Philippe N, Alcaraz J-P, Coursange E, Geiselmann J, Schneider D, Plasmid 2004, 51, 246–255. [PubMed: 15109831]

[24]. Salib MN, Molinski TF, J. Org. Chem 2017, 82, 10181–10187. [PubMed: 28846849]

[25]. Freire VF, Slivinski J, Quintana-Bulla JI, Moraes FC, Paradas WC, Salgado LT, Pereira RC, Moura RL, Amado-Filho GM, Ferreira AG, Berlinck RGS, Rev. Bras. Farmacogn 2019, 29, 715–719.

[26]. Fukunaga Y, Kurahashi M, Tanaka K, Yanagi K, Yokota A, Harayama S, Int. J. Syst. Evol. Microbiol 2006, 56, 343–347. [PubMed: 16449437]

[27]. Youssef NH, Duncan KE, Nagle DP, Savage KN, Knapp RM, McInerney MJ, J. Microbiol. Methods 2004, 56, 339–347. [PubMed: 14967225]

[28]. Partridge JD, Harshey RM, J. Bacteriol 2013, 195, 909–918. [PubMed: 23264580]

[29]. Belas R, Trends Microbiol. 2014, 22, 517–527. [PubMed: 24894628]

[30]. van Ditmarsch D, Boyle KE, Sakhtah H, Oyler JE, Nadell CD, Déziel É, Dietrich LEP, Xavier JB, Cell Rep. 2013, 4, 697–708. [PubMed: 23954787]

[31]. Merritt JH, Kadouri DE, O'Toole GA, Curr. Protoc. Microbiol 2005, Chapter 1, Unit 1B.1.

[32]. Tkacz A, Poole P, J. Exp. Bot 2015, 66, 2167–2175. [PubMed: 25908654]

[33]. McFall-Ngai M, Hadfield MG, Bosch TCG, Carey HV, Domazet-Lošo T, Douglas AE, Dubilier N, Eberl G, Fukami T, Gilbert SF, Hentschel U, King N, Kjelleberg S, Knoll AH, Kremer N, Mazmanian SK, Metcalf JL, Nealson K, Pierce NE, Rawls JF, Reid A, Ruby EG, Rumpho M, Sanders JG, Tautz D, Wernegreen JJ, Proc. Natl. Acad. Sci. U. S. A 2013, 110, 3229–3236. [PubMed: 23391737]

[34]. Hentschel U, Piel J, Degnan SM, Taylor MW, Nat. Rev. Microbiol 2012, 10, 641–654. [PubMed: 22842661]

[35]. Webster NS, Blackall LL, ISME J 2009, 3, 1–3. [PubMed: 18971962]

[36]. Geudens N, Martins JC, Front. Microbiol 2018, 9, 1867. [PubMed: 30158910]

[37]. Challis GL, Ravel J, Townsend CA, Chem. Biol 2000, 7, 211–224. [PubMed: 10712928]

[38]. Walsh CT, Nat. Prod. Rep. 2016, 33, 127–135. [PubMed: 26175103]

[39]. Ramel C, Tobler M, Meyer M, Bigler L, Ebert MO, Schellenberg B, Dudler R, BMC Biochem. 2009, 10, 26. [PubMed: 19863801]

[40]. Imker HJ, Walsh CT, Wuest WM, J. Am. Chem. Soc 2009, 131, 18263–18265. [PubMed: 19968303]

[41]. Tillett D, Dittmann E, Erhard M, Von Döhren H, Börner T, Neilan BA, Chem. Biol 2000, 7, 753–764. [PubMed: 11033079]

[42]. Christiansen G, Fastner J, Erhard M, Börner T, Dittmann E, J. Bacteriol 2003, 185, 564–572. [PubMed: 12511503]

[43]. Christiansen G, Yoshida WY, Blom JF, Portmann C, Gademann K, Hemscheidt T, Kurmayer R, J. Nat. Prod 2008, 71, 1881–1886. [PubMed: 18939865]

[44]. Ziemert N, Podell S, Penn K, Badger JH, Allen E, Jensen PR, PLoS One 2012, 7, e34064. [PubMed: 22479523]

[45]. Hermes C, Richarz R, Wirtz DA, Patt J, Hanke W, Kehraus S, Voß JH, Küppers J, Ohbayashi T, Namasivayam V, Alenfelder J, Inoue A, Mergaert P, Gütschow M, Müller CE, Kostenis E, König GM, Crüsemann M, Nat. Commun 2021, 12, 144. [PubMed: 33420046]

[46]. Aleti G, Lehner S, Bacher M, Compant S, Nikolic B, Plesko M, Schuhmacher R, Sessitsch A, Brader G, Environ. Microbiol 2016, 18, 2634–2645. [PubMed: 27306252]

[47]. Berti AD, Greve NJ, Christensen QH, Thomas MG, J. Bacteriol 2007, 189, 6312–6323. [PubMed: 17601782]

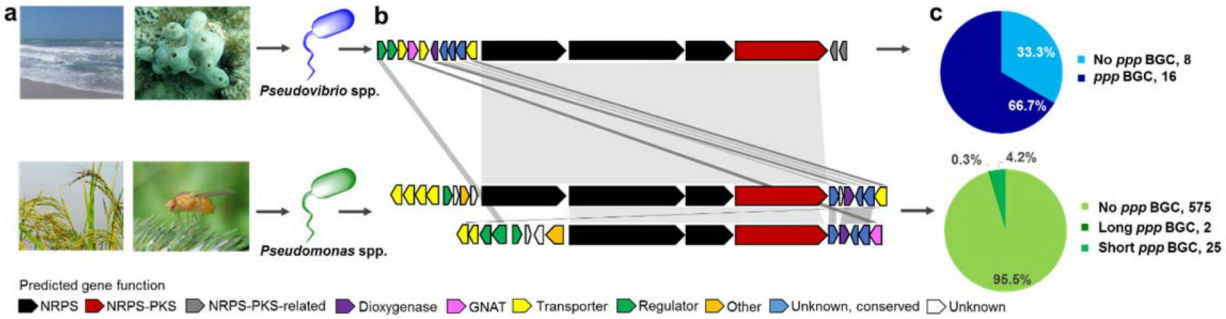


Fig. 1. The *ppp* gene cluster family conserved in *Pseudovibrio* and *Pseudomonas* bacteria known to establish associations with eukaryotic hosts.

(**a**) *Pseudovibrio* and *Pseudomonas* species isolated from marine invertebrates, and from terrestrial plants and animals, respectively, contain a member of a (**b**) hybrid nonribosomal peptide synthetase (NRPS) polyketide synthase (PKS) biosynthetic gene cluster (BGC) family we named *ppp*. Genes are color-coded based on predicted function as indicated. GNAT, GCN5-related *N*-acetyltransferase. **c**) Prevalence of the *ppp* BGC in *Pseudovibrio* (24) and *Pseudomonas* (602) genomes.

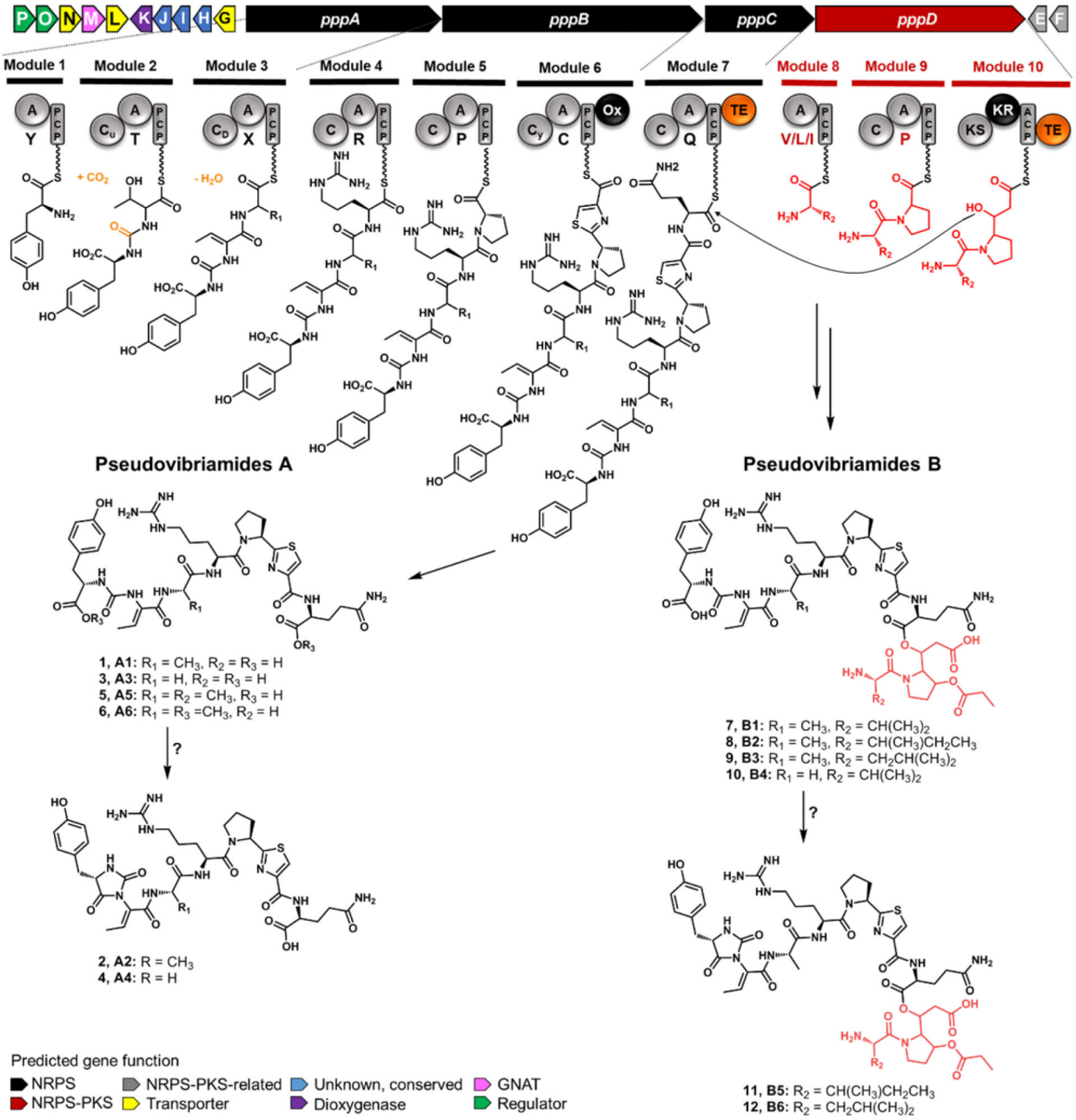


Fig. 2. Organization of the *ppp* BGC from *Pseudovibrio brasiliensis* Ab134, and proposed biosynthesis of isolated pseudovibriamides.

Marfey's analyses were performed for all pseudovibriamides, except **5** and **6**. The single letter amino acid code is indicated below each A domain. If the amino acid could not be predicted, it is indicated with "X". Genes are colored based on predicted function and as indicated. NRPS, nonribosomal peptide synthetase. PKS, polyketide synthase. NRPS-PKS-related, 4'-phosphopantetheinyl transferase (*pppE*) and type II thioesterase (*pppF*), respectively. GNAT, GCN5-related *N*-acetyltransferase. Domain key: A, adenylation; ACP/PCP, acyl/peptidyl carrier protein; C, condensation; C_u, ureido-linkage formation

condensation domain; C_D, dehydration condensation domain; C_y, condensation and heterocyclization; KR, ketoreductase; KS, ketosynthase; Ox, oxidase; TE, thioesterase.

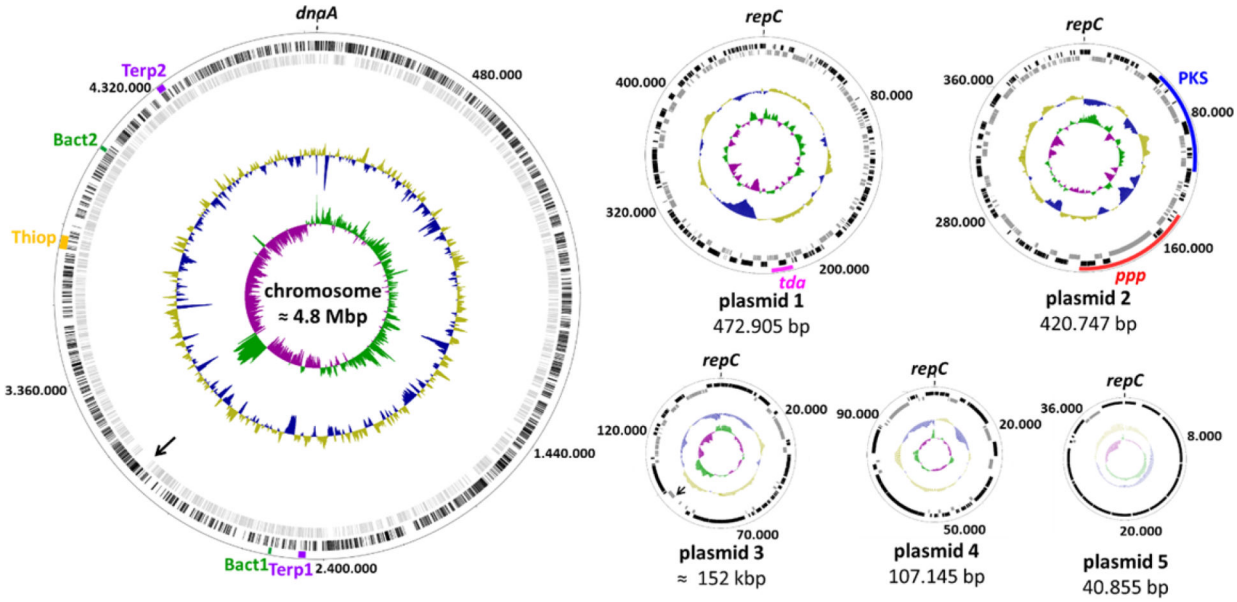


Fig. 3. The genome of *Pseudovibrio brasiliensis* Ab134.

The outer circle depicts the size in base pairs. Lane 1 (from the outside in) shows the eight predicted BGCs, color-coded based on biosynthetic class: bact, bacteriocin (green); terp, terpene (purple); thiop, thiopeptide (yellow); *tda*, tropodithietic acid (pink); PKS, polyketide synthase (blue). The *ppp* BGC (red) investigated in this study is located on plasmid 2. Lanes 2 and 3 show predicted open reading frames (ORFs) on the leading (black) and lagging (gray) strands, respectively. Lanes 4 and 5 depict normalized plot of GC content (yellow/blue) and normalized plot of GC skew (purple/green), respectively. The chromosome is oriented to *dnaA*, plasmids are oriented to *repC*. See Table S3 for further details. Replicons are represented not strictly to scale. Two remaining gaps (one in the chromosome and one in plasmid 3) are indicated with arrows. GenBank genome accession codes: CP074126 to CP074131 for chromosome to plasmid 5, respectively.

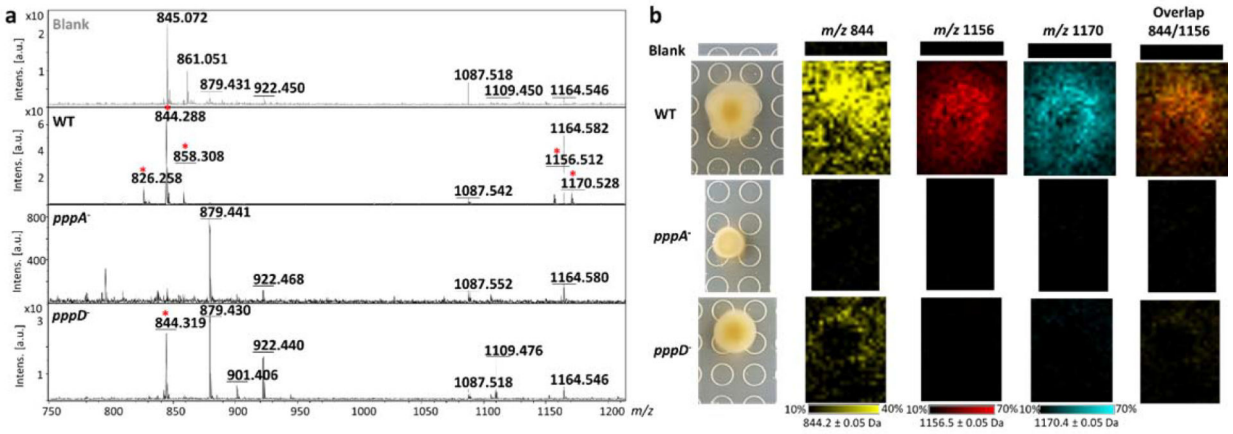


Fig. 4. Comparative metabolite analysis of wild-type, *pppA* and *pppD* mutant strains of *P. brasiliensis* Ab134 using MALDI-TOF mass spectrometry.

(**a**) dried-droplet analysis of culture extracts. Peaks of interest are highlighted with red asterisks. (**b**) Imaging mass spectrometry analysis showing the spatial distribution of selected compounds on and around microbial colonies. *P. brasiliensis* wild-type and mutant strains were grown on marine agar for 18 h before matrix application and analysis.

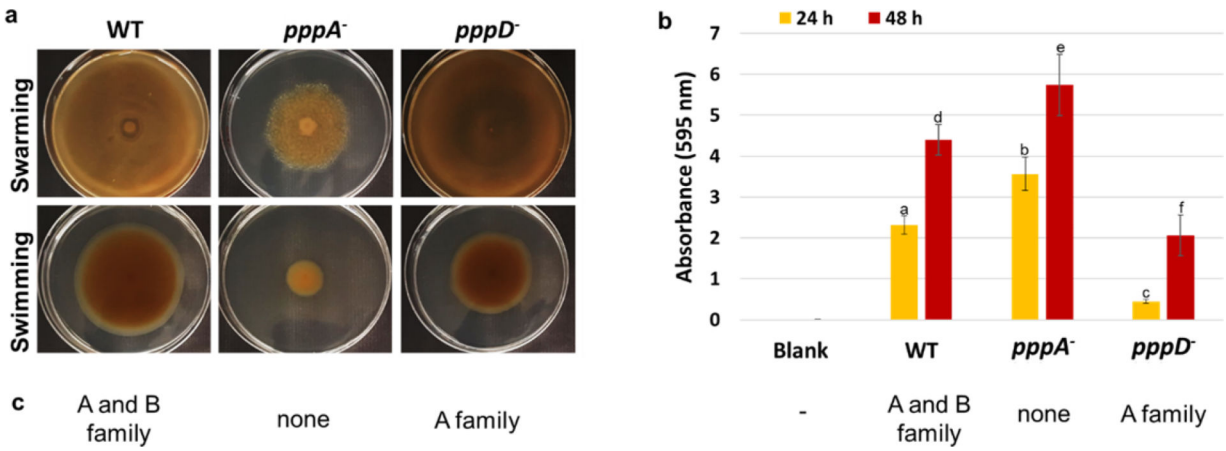


Fig. 5. Effect of *pppA* and *pppD* deletion on flagellar motility and biofilm formation.

(**a**) Swarming assay (top) and swimming assay (bottom) performed on marine agar containing either 0.5% or 0.3% Eiken agar, respectively. Plates were inoculated on top of the agar (swarming) or inside the agar (swimming) with 5 μL of cryo-preserved cultures of Ab134 wild-type (WT) or *pppA*$^{-}$/*pppD*$^{-}$ mutant strains normalized to OD_{600} of 1.0. Pictures shown were taken at 48 hours after inoculation. The assay was performed multiple times, each time in triplicates, and similar results were obtained each time. Representative results are shown. (**b**) Biofilm assay using 0.1% crystal-violet solution to quantify biofilm adherence and microcolony formation. Bars are average of six replicates (Table S19), and error bars represent standard deviation. Lowercase letters above bar indicate statistically different values according to the Tukey test 95% confidence level (Table S19). (**c**) Pseudovibriamides produced by each bacterial strain.