

**This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.**

**Author(s):** Kust, Andreja; Mareš, Jan; Jokela, Jouni; Urajová, Petra; Hájek, Jan; Saurav, Kumar; Voráčková, Kateřina; Fewer, David P.; Haapaniemi, Esa; Permi, Perttu; Řeháková, Klára; Sivonen, Kaarina; Hrouzek, Pavel

**Title:** Discovery of a Pederin Family Compound in a Nonsymbiotic Bloom-Forming Cyanobacterium

**Year:** 2018

**Version:** Accepted version (Final draft)

**Copyright:** © 2018 American Chemical Society

**Rights:** In Copyright

**Rights url:** <http://rightsstatements.org/page/InC/1.0/?language=en>

**Please cite the original version:**

Kust, A., Mareš, J., Jokela, J., Urajová, P., Hájek, J., Saurav, K., Voráčková, K., Fewer, D. P., Haapaniemi, E., Permi, P., Řeháková, K., Sivonen, K., & Hrouzek, P. (2018). Discovery of a Pederin Family Compound in a Nonsymbiotic Bloom-Forming Cyanobacterium. *ACS Chemical Biology*, 13(5), 1123-1129. <https://doi.org/10.1021/acscchembio.7b01048>

## Discovery of a pederin family compound in a non-symbiotic bloom-forming cyanobacterium

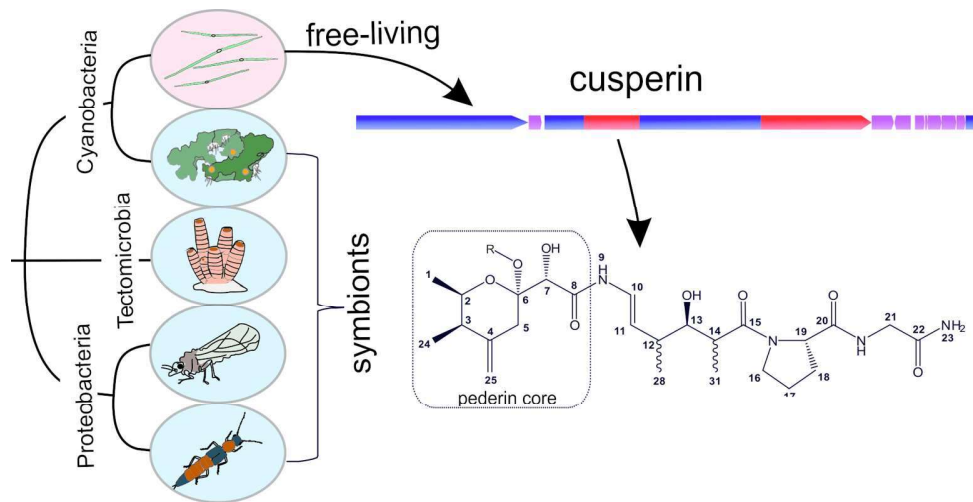
Andreja Kust, Jan Mareš, Jouni Jokela, Petra Urajová, Jan Hájek, Kumar Saurav, Kateřina Vorá#ová, David P Fewer, Esa Haapaniemi, Perttu Permi, Klára #eháková, Kaarina Sivonen, and Pavel Hrouzek

ACS Chem. Biol., **Just Accepted Manuscript** • DOI: 10.1021/acscchembio.7b01048 • Publication Date (Web): 23 Mar 2018

Downloaded from <http://pubs.acs.org> on March 26, 2018

### Just Accepted

“Just Accepted” manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides “Just Accepted” as a service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. “Just Accepted” manuscripts appear in full in PDF format accompanied by an HTML abstract. “Just Accepted” manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are citable by the Digital Object Identifier (DOI®). “Just Accepted” is an optional service offered to authors. Therefore, the “Just Accepted” Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the “Just Accepted” Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these “Just Accepted” manuscripts.



160x81mm (300 x 300 DPI)

1  
2  
3 **1 Discovery of a pederin family compound in a non-symbiotic bloom-forming**  
4 **2 cyanobacterium**  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

4 Andreja Kust#<sup>†, ‡, §</sup>, Jan Mareš#<sup>†, ‡, §</sup>, Jouni Jokela<sup>1</sup>, Petra Urajová<sup>†</sup>, Jan Hájek<sup>†, §</sup>, Kumar  
5 Saurav<sup>†</sup>, Kateřina Voráčová<sup>†</sup>, David P. Fewer<sup>1</sup>, Esa Haapaniemi<sup>⊥</sup>, Perttu Permi<sup>⊥, ¶</sup>, Klára  
6 Řeháková<sup>†</sup>, Kaarina Sivonen<sup>1</sup> and Pavel Hrouzek\*<sup>†, §</sup>

7 <sup>†</sup>*Institute of Microbiology CAS, Center Algatech, Třeboň, Czech Republic;*

8 <sup>‡</sup>*The Czech Academy of Sciences, Biology Centre, Institute of Hydrobiology, České*  
9 *Budějovice, Czech Republic;*

10 <sup>§</sup>*University of South Bohemia, Faculty of Science, České Budějovice, Czech Republic;*

11 <sup>1</sup>*Department of Microbiology, University of Helsinki, FI-00014 Helsinki, Finland;*

12 <sup>⊥</sup>*Department of Chemistry, University of Jyväskylä, FI-40010 Jyväskylä, Finland*

13 <sup>¶</sup>*Department of Biology and Environmental Science, Nanoscience Center, University of*  
14 *Jyväskylä, FI-40010 Jyväskylä, Finland*

15  
16 #These authors contributed equally to this work  
17

18 **Correspondence:** P Hrouzek, Institute of Microbiology CAS, Center Algatech, Novohradská  
19 237, Opatovický Mlýn, 379 81 Třeboň, Czech Republic, Email: [hrouzekp@gmail.com](mailto:hrouzekp@gmail.com), Tel.:  
20 +420-384-340-470, Fax: +420-384-340-415.  
21

**1 Abstract**

2 **The pederin family includes a number of bioactive compounds isolated from symbiotic**  
3 **organisms of diverse evolutionary origin. Pederin is linked to beetle-induced dermatitis**  
4 **in humans and pederin family members possess potent antitumor activity caused by**  
5 **selective inhibition of the eukaryotic ribosome. Their biosynthesis is accomplished by a**  
6 **polyketide/non-ribosomal peptide synthetase machinery employing an unusual trans-**  
7 **acyltransferase mechanism. Here we report a novel pederin type compound, cusperin,**  
8 **from the free-living cyanobacterium *Cuspidothrix issatschenkoi* (earlier**  
9 ***Aphanizomenon*). The chemical structure of cusperin is similar to that of nosperin**  
10 **recently isolated from the lichen cyanobiont *Nostoc* sharing the tetrahydropyran moiety**  
11 **and major part of the linear backbone. However, the cusperin molecule is extended by a**  
12 **glycine residue and lacks one hydroxyl substituent. Pederins were previously thought to**  
13 **be exclusive to symbiotic relationships. However, *C. issatschenkoi* is a non-symbiotic**  
14 **planktonic organism and a frequent component of toxic water blooms. Cusperin is**  
15 **devoid of the cytotoxic activity reported for other pederin family members. Hence, our**  
16 **findings raise questions about the role of pederin analogues in cyanobacteria and**  
17 **broaden the knowledge of ecological distribution of this group of polyketides.**

18 Pederin is a toxin containing two tetrahydropyran cores that causes dermatitis<sup>1</sup>. Pederin was  
19 firstly reported from the beetle *Paederus fuscipes* and widely believed to be produced by  
20 eukaryotes<sup>2,3</sup>. However, the discovery of the pederin biosynthetic gene cluster demonstrated  
21 that bacterial symbionts of insects, marine invertebrates, and lichens are the true producers of  
22 these compounds<sup>4</sup>. Biosynthesis of pederin and its homologues is accomplished by hybrid  
23 polyketide/nonribosomal peptide synthases (PKS/NRPS) employing an unusual trans-  
24 acyltransferase (AT) PKS mechanism<sup>5</sup>, in which the PKS enzymatic modules utilize a  
25 common AT encoded in a separate gene. The biosynthesis of all pederin family compounds

1  
2  
3 1 seems to share the first several steps forming the typical tetrahydropyran moiety attached to  
4  
5 2 an acyl chain with a glycine residue<sup>6</sup>.

6  
7  
8 3 Metagenomic sequencing of symbiotic consortia has recently uncovered a large and highly  
9  
10 4 specific repertoire of natural products, particularly in bacterial symbionts of marine  
11  
12 5 sponges<sup>7,8</sup>. Pederins are currently also believed to be specifically associated with prokaryote-  
13  
14 6 eukaryote symbioses as they have been reported exclusively from symbiotic  
15  
16 7 associations<sup>6,7,9,10,11</sup>. However, a recent discovery of a pederin analogue in a purportedly free-  
17  
18 8 living strain of the proteobacterium *Labrenzia* sp. PHM005 isolated from marine sediment  
19  
20 9 challenged this hypothesis<sup>12</sup>.

21  
22  
23 10 The potential of free-living organisms to produce pederins is further corroborated with our  
24  
25 11 report of a novel pederin analogue, cusperin, identified for the first time in a non-symbiotic  
26  
27 12 cyanobacterium, *Cuspidothrix issatschenkoi* (Usačev) Rajaniemi et al. *Cuspidothrix* (earlier  
28  
29 13 *Aphanizomenon*) is a frequent bloom-forming species in freshwater bodies<sup>13</sup>, which belongs  
30  
31 14 to the lineage of typical freshwater planktonic heterocytous cyanobacteria, clearly distant  
32  
33 15 from all groups of the typical pederin producers (Figure 1).

### 34 35 36 37 **[Figure 1]**

38  
39  
40 17 Inspection of a draft genome of *C. issatschenkoi* strain CHARLIE-1 led to the identification  
41  
42 18 of a previously unreported PKS/NRPS gene cluster, exhibiting >70% similarity with the  
43  
44 19 nosperin biosynthetic gene cluster<sup>6</sup>. The core region of this biosynthetic gene cluster is almost  
45  
46 20 identical to that reported for pederin compounds (Figure 2a, Supplementary Table 1, 2). The  
47  
48 21 putative biosynthetic gene cluster (~54 kb) consisted of 11 protein coding ORFs: two genes  
49  
50 22 encoding multidomain PKS/NRPS proteins, eight accessory genes, and a single gene  
51  
52 23 encoding a peculiar AT in *trans* position (Figure 2a). Further evidence suggesting the  
53  
54 24 presence of a *trans*-AT PKS system in the cusperin synthetase is the absence of AT-domains  
55  
56  
57  
58  
59  
60

1  
2  
3 1 in the deduced PKS proteins (Figure 2b). Detailed analysis of the individual components of  
4  
5 2 the gene cluster are provided in Supporting Information.

6  
7  
8 **[Figure 2]**

9  
10 4 Analysis of methanol extracts of *C. issatschenkoi* CHARLIE-1 led to the identification of two  
11  
12 5 putative products of the PKS/NRPS gene cluster. Their chemical structures were elucidated  
13  
14 6 using a combination of high-performance liquid chromatography connected to high resolution  
15  
16 7 mass spectrometer with electrospray ionization (HPLC-ESI-HRMS/MS) and nuclear  
17  
18 8 magnetic resonance (NMR) spectroscopy (Figure 3, 4 Supplementary Figures 1–6,  
19  
20 9 Supplementary Tables 3, 4). A prominent chromatographic peak was obtained by HPLC-MS  
21  
22 10 analysis of *Cuspidothrix* extract containing a compound detected at  $m/z=561.2896$   $[M+Na]^+$   
23  
24 11 with calculated elemental composition of  $C_{26}H_{42}N_4O_8+Na^+$  ( $\Delta$  0.2 ppm) corresponding to  
25  
26 12 cusperin A (Figure 3). Its MS/MS exhibited attributes analogous to those reported for  
27  
28 13 nosperin, e.g. the loss of methanol (producing ion a) and the loss of a whole substituted  
29  
30 14 tetrahydropyran moiety (producing ion b) identical to those of nosperin<sup>6</sup> (Figure 3b).

31  
32  
33  
34  
35 **[Figure 3]**

36  
37  
38 16 Data from the analysis of  $^1H$ ,  $^1H$ - $^1H$  NOESY,  $^{13}C$ -HSQC,  $^{13}C$ -HMBC and  $^{15}N$ -HSQC spectra  
39  
40 17 (Supplementary Figures 2–6) of cusperin showed high similarity to nosperin NMR data<sup>6</sup>  
41  
42 18 which proved that their general structure was the same (Supplementary Table 4). Compared to  
43  
44 19 nosperin, cusperin lacked C-28 hydroxyl, C-20 hydroxyl was changed to carbonyl and next to  
45  
46 20 it there was an extra amide group (C-21-NH). Due to the lack of the hydroxyl, C-28 showed a  
47  
48 21 methyl signal ( $\delta C$  19.8,  $\delta H$  1.00) and this difference affected the nearby shift values  
49  
50 22 (Supplementary Table 4). Four carbonyl signals were found (Supplementary Figure 3), one  
51  
52 23 more than in nosperin as one hydroxyl of nosperin is replaced with carbonyl in cusperin.  
53  
54 24 Cusperin NH-9 signal (d,  $\delta H$  8.40) matched with nosperin (d,  $\delta H$  8.57). The extra C-21-NH  
55  
56  
57  
58  
59  
60

1  
2  
3 1 signal (dd,  $\delta$ H 7.11) of cusperin was broad but the signals dd structure could be seen which  
4  
5 2 was characteristic for glycine in peptide structures.  $^{15}$ N-HSQC spectrum showed the  $\delta$ H 7.11  
6  
7 3 proton connection to a nitrogen atom (Supplementary Figure 4). The most important  $^{13}$ C-  
8  
9 4 HMBC and  $^1$ H- $^1$ H NOESY correlations are shown in Supplementary Figure 3 and 5.

11  
12 5 The MS/MS analysis further detected the presence of a second cusperin variant with  $m/z$  at  
13  
14 6 547.2740  $[M+Na]^+$  and calculated elemental composition of  $C_{25}H_{40}N_4O_8+Na^+$  ( $\Delta -0.3$  ppm).  
15  
16 7 While cusperin A provided ions corresponding to  $[M-CH_2OH+H]^+$  in both MS and MS/MS  
17  
18 8 ( $m/z$  507. 2812,  $\Delta$  0.3ppm), cusperin B provided only  $[M-H_2O+H]^+$  fragment (Supplementary  
19  
20 9 Figure 6). This corresponds to the losses of a methoxy- and a hydroxy- group from the  
21  
22 10 molecular ion of cusperin A and B, respectively. Based on the combination of MS and NMR  
23  
24 11 data we concluded that cusperin A presents the methylacetal variant (bearing methoxy group  
25  
26 12 on C-6) and cusperin B presents the hemiacetal variant (bearing hydroxy group on C-6).  
27  
28 13 Purification (see Supporting Information and Supplementary Table 5) yielded 0.3 mg of  
29  
30 14 cusperin A and B in molar ratio 70/30 (Supplementary Figure 6).  
31  
32  
33  
34

35 15 The origin of the structural differences between cusperin and nosperin as well as the  
36  
37 16 occurrence of cusperin A and B variants were explained by bioinformatic prediction of  
38  
39 17 cusperin biosynthesis (Figure 2b). First, the presence of a terminal glycine in cusperin instead  
40  
41 18 of a reduced biketide was consistent with an additional NRPS module encoded in *cusC*  
42  
43 19 containing an A-domain with a predicted substrate specificity for glycine (Supplementary  
44  
45 20 Table 2). The second difference observed was the lack of the C-12 hydroxyl in cusperin,  
46  
47 21 explained by the lack of cytochrome-450 like protein responsible for the hydroxylation step in  
48  
49 22 nosperin<sup>6</sup>. Four nitrogen atoms were predicted in the cusperin scaffold based on the presence  
50  
51 23 of three NRPS modules each incorporating a single amino acid and an asparagine synthetase  
52  
53 24 responsible for the final transamination step (Figure 2b). The number of nitrogen atoms was  
54  
55  
56  
57  
58  
59  
60



1 further confirmed by  $^{15}\text{N}$  isotopic substitution ( $m/z$  at 565.2685,  $\Delta -0.9$  ppm). The predicted  
2 product of *cusF* was an *O*-methyltransferase, likely involved in facultative *O*-methylation of  
3 the hydroxy group at C6, which results in formation of the observed cusperin  
4 methylacetal/hemiacetal variants (cusperin A and B). The individual predicted biosynthetic  
5 steps are explained in detail in Supporting Information.

#### 6 [Figure 4]

7 The cusperin gene cluster exhibited high similarity in the arrangement of the first five  
8 PKS/NRPS modules to those observed in several pederin members (pederin, theopedierins,  
9 mycalamides, onnamides) which allowed us to infer the absolute configuration at chiral  
10 carbons C2, C3, C6, and C7 (Figure 4b). Besides, the arrangement of the first 10 modules was  
11 consistent with the nosperin gene cluster, further suggesting the configuration of C10, C11  
12 and C13 are identical to nosperin. Comparison of the NMR coupling constants and shift  
13 values (Supplementary Table 4) further indicates that the stereochemistry is likely identical  
14 between these two compounds (Figure 4b). Bioinformatics analysis of the specific sequence  
15 motifs of the KR domains have again confirmed D-OH configuration at C2 and C13 based on  
16 the presence of a conserved aspartate residue recently demonstrated to play a key role in  
17 stereospecificity of the KR domain<sup>14,15,16</sup> (Supplementary Table 6). Finally, the lack of the  
18 epimerization domain in the NRPS module 10 supported the incorporation of an L-Pro residue  
19 as per bioinformatics analysis of pathway A-domains. Equivalent to nosperin, C12 and C14  
20 remained without stereochemical assignments.

21 The strong cytotoxic activity of pederin family members was found due to their specific  
22 binding to the large ribosomal subunit<sup>17</sup>. Recent structure-activity studies defined the key  
23 features of pederin analogues involved in their action<sup>18</sup>. Firstly, the presence of alkoxy- or  
24 hydroxy- substitution at C-10 facilitates the molecule rotation and subsequent binding, while

1  
2  
3 1 the lack of any substituent at this position leads to diminished activity<sup>18</sup>. Additionally, the  
4  
5 2 presence of methoxy- group at C-13 is favoured over the hydroxyl group. In comparison to  
6  
7 3 other natural pederin analogues, cusperin A and B exhibited only a weak cytotoxic activity  
8  
9 4 (Supplementary Table 7). The cusperin molecule features both traits, which do not favour the  
10  
11 5 interaction with ribosome as C-10 is not functionalized and C-13 bears the less efficient -OH  
12  
13  
14 6 group.

15  
16 7 The bioactivity of cusperin demonstrates that not all natural pederin analogues exert high  
17  
18 8 cytotoxic activity and thus they do not act generally and exclusively as ribosome poisons.  
19  
20 9 Interestingly, high production levels of cusperin A and B were recorded in the *Cuspidothrix*  
21  
22 10 culture (up to 2.5 mg/g of dry weight). With regard to the high content and easy isolation  
23  
24 11 protocol, discovery of cusperin opens up an opportunity to address the ecological role of this  
25  
26 12 type of compounds.

27  
28  
29  
30 13 *Cuspidothrix*, a bloom-forming cyanobacterium which has the potential to produce  
31  
32 14 compounds that are a threat to human health<sup>19</sup>, has been characterized as the first free-living  
33  
34 15 cyanobacterium producing a pederin-type compound, cusperin. Our findings contradict the  
35  
36 16 recently suggested restricted occurrence of pederin family members in symbiotic bacteria,  
37  
38 17 and dispute their general role directly associated with specific symbioses between eukaryotes  
39  
40 18 and bacteria<sup>6</sup>. Cusperin production may however play a role in other microbe-microbe  
41  
42 19 relationships occurring in freshwater blooms that are currently not well characterized. The  
43  
44 20 selective advantage for the production of a pederin family compound by a non-symbiotic  
45  
46 21 organism remains unknown and further studies are needed to decipher its possible ecological  
47  
48 22 significance; especially as cytotoxic activity is weak compared to other pederins.  
49  
50

51 23 METHODS  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 1 **Cultivation of the strain.** *C. issatschenkoi* CHARLIE-1 was isolated from a water bloom  
4 sample collected from fishpond Papež near Dobříš, Czech Republic. A monoclonal culture  
5 was established from a single filament and maintained in the culture collection of the Biology  
6 Centre of CAS, Institute of Hydrobiology in 250 ml Erlenmeyer flasks with WC medium<sup>20</sup> at  
7 21°C and a light intensity of 50  $\mu\text{mol m}^{-2} \text{s}^{-1}$  (16:8 L: D cycle). For isolation procedure the  
8 batch culture of the strain was grown in 10 l Erlenmeyer flasks using BG-11 medium bubbled  
9 with 1.5% CO<sub>2</sub> enriched air at 25°C with an illumination of 100  $\mu\text{mol m}^{-2} \text{s}^{-1}$ . The biomass  
10 for isolation of <sup>15</sup>N isotopically substituted cusperin was cultivated under the same cultivation  
11 conditions in BG-11 medium supplemented with Na<sup>15</sup>NO<sub>3</sub> (Sigma Aldrich, Cat. No. 364606)  
12 and the culture was re-inoculated twice each time after 20 days of cultivation period. Strain  
13 identification was performed based on morphological characters<sup>21</sup> and position in a  
14 phylogenetic tree based on the 16S rRNA gene.

15  
16  
17 13 **Phylogenetic analysis.** The 16S rRNA gene sequences of pederin family compound  
18 producers available through NCBI and the cusperin producer *C. issatschenkoi* CHARLIE-1  
19 were aligned with a set of closest BLAST hits, representatives of the corresponding bacterial  
20 phyla, and a representative set of cyanobacterial sequences. A phylogenetic tree was  
21 constructed using the Bayesian Inference. See Supporting Information for details.

22  
23  
24 18 **Genomic and bioinformatic analysis.** Single filaments of *C. issatschenkoi* CHARLIE-1  
25 were isolated using the glass capillary technique and prepared for whole genome sequencing  
26 as described previously<sup>22</sup>. The filaments were utilized as a template for whole-genome  
27 amplification by multiple-displacement amplification (MDA) using Repli-G Mini Kit  
28 (Qiagen). Sixteen MDA products that passed the quality check were pooled and sent for  
29 commercial *de novo* genome sequencing (EMBL Genomics Core Facility, Heidelberg,  
30 Germany) using Illumina MiSeq Pair-End library with 250 bp reads, 350 bp average insert

1 length, and 1.4 Gbp data yield. The data from Illumina were assembled using CLC Bio  
2 Genomics Workbench v. 10.5 (Qiagen). The genomic assembly is available under NCBI  
3 accession number PGEM00000000.1 and the cusperin biosynthetic gene cluster is available  
4 under accession number MG518226.

5 Genomic scaffolds were investigated using BLASTp in Geneious Pro R10 (Biomatters)  
6 software package to identify putative PKS/NRPS gene clusters, employing cyanobacterial A-  
7 domains and KS-domains as queries. In target genomic regions, open reading frames were  
8 predicted (Glimmer 3)<sup>23</sup>, the deduced proteins were functionally annotated (BLASTp and  
9 CDD searches), and in the putative PKS/NRPS proteins individual enzymatic domains were  
10 identified (antiSMASH 4.0)<sup>24</sup>. The substrate specificities of KS domains in the *trans*-AT PKS  
11 modules were predicted based on similarity to the closest functionally characterized hits in the  
12 BLASTp analysis (Supplementary Table 2), which were unequivocally represented by KS  
13 domains from the nosperin pathway (analysed by phylogenetics in a previous study<sup>6</sup>). The  
14 predicted specificity of NRPS adenylation domains was inferred using a combination of tools  
15 implemented in antiSMASH 4.0 (Stachelhaus code, NRPS Predictor3, pHMM search). A  
16 biosynthetic scheme was constructed based on the functional annotation of the deduced  
17 PKS/NRPS domains and accessory enzymes, and by comparing to known biosynthesis  
18 pathways of pederin-family compounds. Stereochemical assignment of cusperin was assessed  
19 based on NMR data, specific KR sequence motifs<sup>14,15,16</sup> and overall arrangement of  
20 NRPS/PKS modules<sup>14</sup>.

21 **HPLC-ESI-HRMS.** Crude extract of *C. issatschenkoi* CHARLIE-1 was analyzed using high-  
22 performance liquid chromatograph (Dionex UltiMate 3000 UHPLC+ (Thermo Scientific,  
23 Sunnyvale, CA, USA) connected to Bruker Impact HD high resolution mass spectrometer  
24 (Bruker, Billerica, Massachusetts, USA) with electrospray ionization in positive mode.

1  
2  
3 1 Separation was performed on RP-C18 column (Phenomenex Kinetex, 2.6  $\mu\text{m}$ ; 4.6  $\times$  150,  
4  
5 2 Torrance, CA, USA) using H<sub>2</sub>O (A) and acetonitrile (B) (both containing 0.1% HCOOH) as a  
6  
7 3 mobile phase at a flow rate of 0.5 ml min<sup>-1</sup>. The linear gradient was as follows: A/B 85/15 (0  
8  
9 4 min), 85/15 (in 1 min), 0/100 (in 19 min), 0/100 (in 25 min), and 85/15 (in 30 min). The  
10  
11 5 following settings of mass spectrometer were used: drying temperature, 200 °C; drying gas  
12  
13 6 flow, 12 l min<sup>-1</sup>; nebulizer gas pressure, 3 bar; capillary voltage, 3.8 kV; end plate offset, 500  
14  
15 7 V. Spectra were collected in the range  $m/z$  20–2000 with a spectra rate of 2 Hz. The  
16  
17 8 automated fragmentation was used in the method with the precursor ion selection 20–2000  
18  
19 9  $m/z$  for with the isolation window width of 1Da. The ramping of the collision energy was  
20  
21 10 based on the molecular mass: 20 eV at 400  $m/z$  and 70 eV at 1000  $m/z$  for analysis. The mass  
22  
23 11 spectrometer was calibrated with sodium formate clusters at the beginning of each analysis.  
24  
25 12 Further the lock mass calibration was applied during the analysis using hexakis (1*H*,1*H*,2*H*-  
26  
27 13 perfluoroethoxy) phosphazene (ES-TOF  $m/z$  622.02896, 97%) purchased from ABCR GmbH  
28  
29 14 & Co., KG, Germany. Quantification of cusperin A and B in the *C. issatschenkoi* CHARLIE-  
30  
31 15 1 extract was performed using method described above with purified compounds as external  
32  
33 16 standard (concentrations: 2.5, 5, 10, 25, 50, 100  $\mu\text{g}\cdot\text{ml}^{-1}$  were used to establish the calibration  
34  
35 17 curve). See Supporting Information for details.  
36  
37  
38  
39  
40

41 18 **Extraction, purification and quantification.** Freeze-dried biomass of *C. issatschenkoi*  
42  
43 19 CHARLIE-1 was extracted twice with 400 ml methanol/water (40/60, v/v) using bath  
44  
45 20 sonicator, resulting supernatants were combined and partially evaporated on rotary evaporator  
46  
47 21 and remaining solvent was diluted by distilled water so that the final MeOH concentration did  
48  
49 22 not exceed 5%. Further the sample was loaded on DCS 18 SPE cartridge and retained  
50  
51 23 compounds were eluted with MeOH. Eluent was evaporated till dryness and reconstituted in  
52  
53 24 10 ml of MeOH prior the high-performance liquid chromatography (HPLC) purification. The  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 1 compounds were purified using three consecutive preparative HPLC separation steps (for  
4  
5 2 detailed information on the purification procedure see the Supporting Information).

6  
7  
8 3 **Nuclear magnetic resonance spectroscopy.** All NMR spectra were collected using a Bruker  
9  
10 4 Avance III HD 800 MHz NMR spectrometer, equipped with cryogenically cooled TCI  $^1\text{H}$ ,  
11  
12 5  $^{13}\text{C}$ ,  $^{15}\text{N}$  triple resonance probehead. Data were collected at 25 °C. For the assignment of  
13  
14 6 cusperin  $^1\text{H}$ ,  $^{13}\text{C}$  and  $^{15}\text{N}$  resonances,  $^1\text{H}$  preset experiment along with two-dimensional total  
15  
16 7 correlation spectroscopy (TOCSY) and nuclear Overhauser effect spectroscopy (NOESY)  
17  
18 8 experiments as well as heteronuclear single quantum coherence ( $^{13}\text{C}$  HSQC and  $^{15}\text{N}$  HSQC)  
19  
20 9 and heteronuclear multiple bond correlation ( $^{13}\text{C}$  HMBC) experiments were employed. For  
21  
22 10 details of the NMR measurement please see Supporting Information.

23  
24  
25  
26  
27 11 **Cytotoxicity testing.** In order to test the antiproliferative and cytotoxic activity usually  
28  
29 12 reported for pederin family members, cusperin bioactivity was tested on human cervical  
30  
31 13 cancer cells (HeLa). For cell maintenance and plating before the experiment see Supporting  
32  
33 14 Information. Serial dilutions (in the range from 0.5 – 20  $\mu\text{M}$ ) of cusperin A and B variants  
34  
35 15 were added as technical triplicates so the concentration of the vehicle (MeOH) did not exceed  
36  
37 16 1%. The viability was determined after 72 h of exposure. To assess cell viability we used  
38  
39 17 three end-point methods: the ATP content measured using the CellTiter-Glo® Luminescent  
40  
41 18 Viability Assay (Promega, G7570), MTT assay, and determination of cell counts (384-well  
42  
43 19 plates) of treated and control cells at 72 h by optical microscopy. The output data represents  
44  
45 20 results of three independent experiments. For detailed method description see Supporting  
46  
47 21 Information.

48  
49  
50  
51 22 **Accession Codes.** PGEM00000000.1 and MG518226.

52  
53  
54  
55 23 **Conflict of interest**

1  
2  
3 1 The authors declare no conflict of interest.  
4

5  
6 2 **Acknowledgements**  
7

8  
9 3 This work was supported by the Czech Science Foundation grant number 14–18067S; the  
10 4 grant of the Faculty of Science, University of South Bohemia GAJU 158/2016/P; Ministry of  
11 5 Education, Youth and Sports of the Czech Republic - National Programme of Sustainability I  
12 6 grant number LO1416 and ALGAMIC project grant number CZ.1.05/2.1.00/19.0392. Access  
13 7 to instruments and other facilities was supported by the Czech research infrastructure for  
14 8 systems biology C4SYS (grant number LM2015055). This work was also supported by the  
15 9 Academy of Finland grant 1273798 to KS. We would like to thank E. Kozlíková-Zapomělová  
16 10 and K. Čapková for isolating and maintaining the strain, T. Galica for help with formatting of  
17 11 Figure 2, and M. Mehrshad for help with sequence annotation.  
18

19  
20  
21  
22  
23  
24  
25  
26  
27  
28 12 **Supporting Information**  
29

30  
31  
32 13 The Supporting Information is available free of charge on the ACS Publications website. The  
33 14 Supporting Information file contains additional details on methods and results concerning the  
34 15 phylogenetic analysis, genomic and bioinformatic analysis, HPLC-MS/MS, and cytotoxic  
35 16 tests, Supplementary Figures 1, 2, 3, 4, 5 and 6; Supplementary Tables 1, 2, 3, 4, 5, 6 and 7,  
36 17 and their captions.  
37  
38  
39  
40  
41  
42

43  
44 18 **References**  
45

- 46  
47 19 [1] Frank, J.-H., and Kanamitsu, K. (1987) *Paederus*, sensu lato (Coleoptera: Staphylinidae):  
48 20 natural history and medical importance, *J. Med. Entomol.* 24, 155–191.  
49  
50  
51 21 [2] Ueta A. (1949) *Kyushu Igaku Zasshi, J. Kurume. Med. College., Kyushu* 249.  
52  
53  
54 22 [3] Pavan, M., and Bo, G. (1953) Pederin, toxic principles obtained in the crystalline state  
55 23 from the beetle *Paederus fuscipes* Curt., *Physiol. Comp. Oecol.* 3, 307–312.  
56  
57  
58  
59  
60

- 1  
2  
3 1 [4] Piel, J. (2002) A polyketide synthase-peptide synthetase gene cluster from an uncultured  
4 bacterial symbiont of *Paederus* beetles, *Proc. Natl. Acad. Sci. U. S. A.* 99, 14002–14007.  
5  
6  
7  
8 3 [5] Helfrich, E.-J.-N., and Piel, J. (2016) Biosynthesis of polyketides by trans-AT polyketide  
9 synthases, *Nat. Prod. Rep.* 33, 231–316.  
10  
11  
12  
13 5 [6] Kampa, A., Gagunashvili, A.-N., Gulder, T.-A.-M., Morinaka, B. I., Daolio, C.,  
14 Godejohann, M., Miao, V.-P.-W., Piel, J., and Andresson, O.-S. (2013) Metagenomic natural  
15 product discovery in lichen provides evidence for a family of biosynthetic pathways in  
16 diverse symbioses, *Proc. Natl. Acad. Sci. U. S. A.* 110, E3129–E3137.  
17  
18  
19  
20  
21  
22 9 [7] Wilson, M.-C., Mori, T., Ruckert, C., Uria, A.-R., Helf, M.-J., Takada, K., Gernert, C.,  
23 Steffens, U.-A.-E., Heycke, N., Schmitt, S., Rinke, C., Helfrich, E.-J.-N., Brachmann, A.- O.,  
24 Gurgui, C., Wakimoto, T., Kracht, M., Crusemann, M., Hentschel, U., Abe, I., Matsunaga, S.,  
25  
26  
27  
28 12 Kalinowski, J., Takeyama, H., and Piel, J. (2014) An environmental bacterial taxon with a  
29 large and distinct metabolic repertoire, *Nature* 506, 58–62.  
30  
31  
32  
33 14 [8] Freeman, M.-F., Vagstad, A.-L., and Piel, J. (2016) Polytheonamide biosynthesis  
34 showcasing the metabolic potential of sponge-associated uncultivated '*Entotheonella*'  
35 bacteria, *Curr. Opin. Chem. Biol.* 31, 8–14.  
36  
37  
38  
39  
40  
41 17 [9] Piel, J., Hui, D.-Q., Wen, G.-P., Butzke, D., Platzer, M., Fusetani, N., and Matsunaga, S.  
42 (2004) Antitumor polyketide biosynthesis by an uncultivated bacterial symbiont of the  
43 marine sponge *Theonella swinhoei*, *Proc. Natl. Acad. Sci. U. S. A.* 101, 16222–16227.  
44  
45  
46  
47  
48 20 [10] Fisch, K.-M., Gurgui, C., Heycke, N., van der Sar, S.-A., Anderson, S.-A., Webb, V.-L.,  
49 Taudien, S., Platzer, M., Rubio, B.-K., Robinson, S.-J., Crews, P., and Piel, J. (2009)  
50 Polyketide assembly lines of uncultivated sponge symbionts from structure-based gene  
51 targeting, *Nat. Chem. Biol.* 5, 494–501.  
52  
53  
54  
55  
56  
57  
58  
59  
60

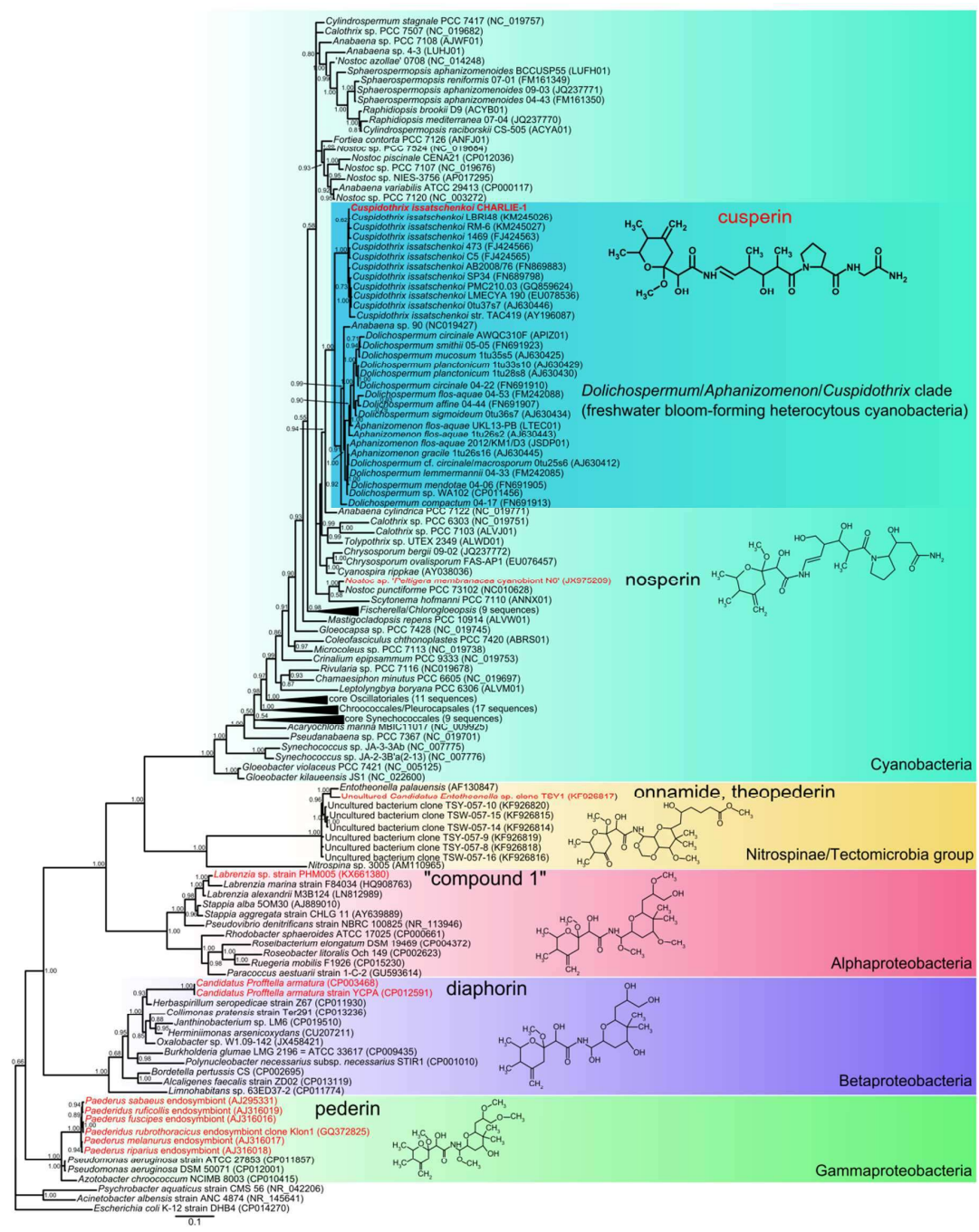


- 1  
2  
3 1 [11] Nakabachi, A., Ueoka, R., Oshima, K., Teta, R., Mangoni, A., Gurgui, M., Oldham, N. -  
4  
5 2 J., van Echten-Deckert, G., Okamura, K., Yamamoto, K., Inoue, H., Ohkuma, M., Hongoh,  
6  
7 3 Y., Miyagishima, S., Hattori, M., Piel, J., and Fukatsu, T. (2013) Defensive bacteriome  
8  
9 4 symbiont with a drastically reduced genome, *Curr. Biol.* *23*, 1478–1484.  
10  
11  
12 5 [12] Schleissner, C., Canedo, L. -M., Rodriguez, P., Crespo, C., Zuniga, P., Penalver, A., de  
13  
14 6 la Calle, F., and Cuevas, C. (2017) Bacterial production of a pederin analogue by a free-living  
15  
16 7 marine Alphaproteobacterium, *J. Nat. Prod.* *80*, 2170–2173.  
17  
18  
19 8 [13] Kaštovský, J., Hauer, T., Mareš, J., Krautová, M., Bešta, T., Komárek, J., Desortová, B.,  
20  
21 9 Heteša, J., Hindáková, A., Houk, V., Janeček, E., Kopp, R., Marvan, P., Pumann, P.,  
22  
23 10 Skácelová, O., and Zapomělová, E. (2010) A review of the alien and expansive species of  
24  
25 11 freshwater cyanobacteria and algae in the Czech Republic, *Biol. Invasions* *12*, 3599–3625.  
26  
27  
28 12 [14] Ueoka, R., Uria, A.-R., Reiter, S., Mori, T., Karbaum, P., Peters, E.-E., Helfrich, E.-J.-  
29  
30 13 N., Morinaka, B.-I., Gugger, M., Takeyama, H., Matsunaga, S., and Piel, J. (2015) Metabolic  
31  
32 14 and evolutionary origin of actin-binding polyketides from diverse organisms, *Nat. Chem.*  
33  
34 15 *Biol.* *11*, 705–712.  
35  
36  
37 16 [15] Caffrey, P. (2003) Conserved amino acid residues correlating with ketoreductase  
38  
39 17 stereospecificity in modular polyketide synthases, *ChemBioChem* *4*, 654–657.  
40  
41  
42 18 [16] Reid, R., Piagentini, M., Rodriguez, E., Ashley, G., Viswanathan, N., Carney, J., Santi,  
43  
44 19 D. - V., Hutchinson, and C. - R. and McDaniel R. (2003) A model of structure and catalysis  
45  
46 20 for ketoreductase domains in modular polyketide synthases, *Biochemistry* *42*, 72–79.  
47  
48  
49 21 [17] Mosey, R.-A., and Floreancig, P.-E. (2012) Isolation, biological activity, synthesis, and  
50  
51 22 medicinal chemistry of the pederin/mycalamide family of natural products, *Nat. Prod. Rep.*  
52  
53 23 *29*, 980–995.  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 [18] Wan, S.-Y., Wu, F.-H., Rech, J.-C., Green, M. -E., Balachandran, R., Horne, W. -S.,  
4 Day, B.-W., and Floreancig, P.-E. (2011) Total synthesis and biological evaluation of  
5 pederin, psymberin, and highly potent analogs, *J. Am. Chem. Soc.* *133*, 16668–16679.  
6  
7  
8  
9  
10 [19] Selwood, A.-I., Holland, P.-T., Wood, S.-A., Smith, K.-F., and McNabb, P. -S. (2007)  
11 Production of anatoxin-a and a novel biosynthetic precursor by the cyanobacterium  
12 *Aphanizomenon issatschenkoi*, *Environ. Sci. Technol.* *41*, 506–510.  
13  
14  
15  
16  
17 [20] Guillard, R.-R., and Lorenzen, C. J. (1972) Yellow-green algae with chlorophyllide c, *J.*  
18 *Phycol.* *8*, 10–14.  
19  
20  
21  
22  
23 [21] Komárek, J. (2013) Cyanoprokaryota. 3. Heterocytous Genera, in *Süswasserflora von*  
24 *Mitteleuropa/Freshwater flora of Central Europe* (Büdel, B., Gärtner, G., Krienitz, L., and  
25 Schagerl, M., Eds), pp 1130, Springer Spektrum.  
26  
27  
28  
29  
30 [22] Mareš, J., Hájek, J., Urajová, P., Kopecký, J., and Hrouzek, P. (2014) A hybrid non-  
31 ribosomal peptide/polyketide synthetase containing fatty-acyl ligase (FAAL) synthesizes the  
32 beta-amino fatty acid lipopeptides puwainaphycins in the cyanobacterium *Cylindrospermum*  
33 *alatosporum*, *PLoS One* *9*, e111904.  
34  
35  
36  
37  
38  
39 [23] Delcher, A.-L., Bratke, K.-A., Powers, E.-C., and Salzberg, S.-L. (2007) Identifying  
40 bacterial genes and endosymbiont DNA with Glimmer, *Bioinformatics* *23*, 673–679.  
41  
42  
43  
44 [24] Blin, K., Wolf, T., Chevrette, M. G., Lu, X.-W., Schwalen, C.-J., Kautsar, S.-A., Duran,  
45 H.-G.-S., Santos, E., Kim, H.-U., Nave, M., Dickschat, J.-S., Mitchell, D.-A., Shelest, E.,  
46 Breitling, R., Takano, E., Lee, S.-Y., Weber, T., and Medema, M.-H. (2017) antiSMASH 4.0-  
47 improvements in chemistry prediction and gene cluster boundary identification, *Nucleic*  
48 *Acids Res.* *45*, W36–W41.  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

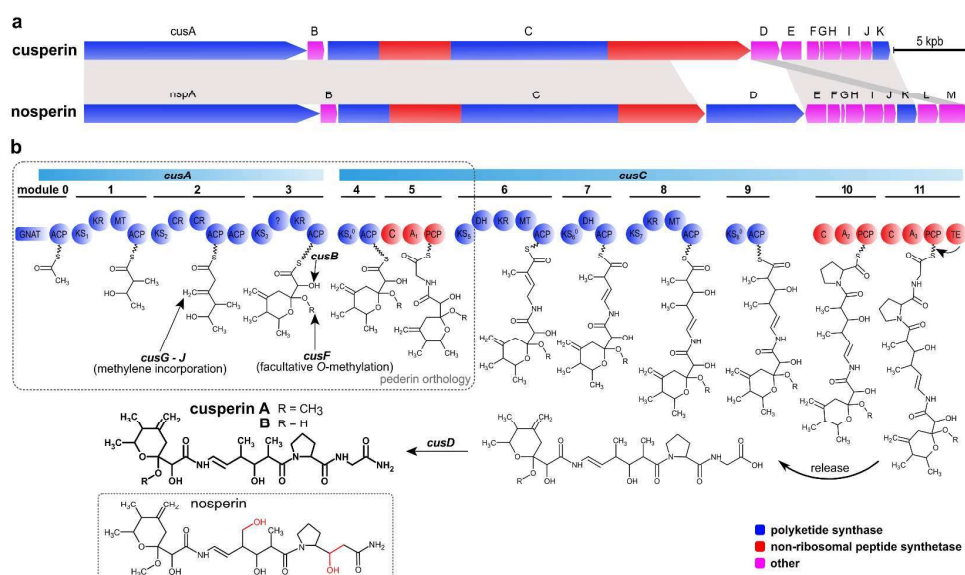
1  
2  
3 1 **Figure 1.** Phylogenetic position of *Cuspidothrix issatschenkoi* CHARLIE-1 among the  
4  
5 2 producers of pederin family compounds. The unrooted tree was inferred from 16S rRNA  
6  
7 3 gene data using Bayesian Inference (BI). BI posterior probabilities are indicated near nodes.  
8  
9 4 The major bacterial lineages containing producers of pederin analogues are color-shaded. The  
10  
11 5 strains producing pederin family compounds are printed in red, representative structures of  
12  
13 6 the products are shown for each group.  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



**Figure 2.** The predicted biosynthetic pathway of cusperin. a) Gene arrangement of the proposed cusperin biosynthetic gene cluster and comparison to nosperin biosynthetic gene cluster. Homologous regions are indicated by grey shading. b) Proposed biosynthesis of

cusperin A and B is co-linear with the arrangement of modules in CusA and CusC. Part of the pathway orthologous to pederin, onnamides, theopederins, and diaphorin is indicated by a dashed box. The structure of nosperin is shown for comparison, differences from cusperin A are highlighted in red colour. A, adenylation domain; ACP, acyl carrier protein; C, condensation domain; CR, crotonase (enoyl-CoA hydratase); DH, dehydratase; GNAT, GCN5-related N-acetyltransferase; MT, methyltransferase; NRPS, non-ribosomal peptide synthetase; KR, ketoreductase; KS, ketosynthase; PCP, peptidyl carrier protein; PKS, polyketide synthase; TE, thioesterase.

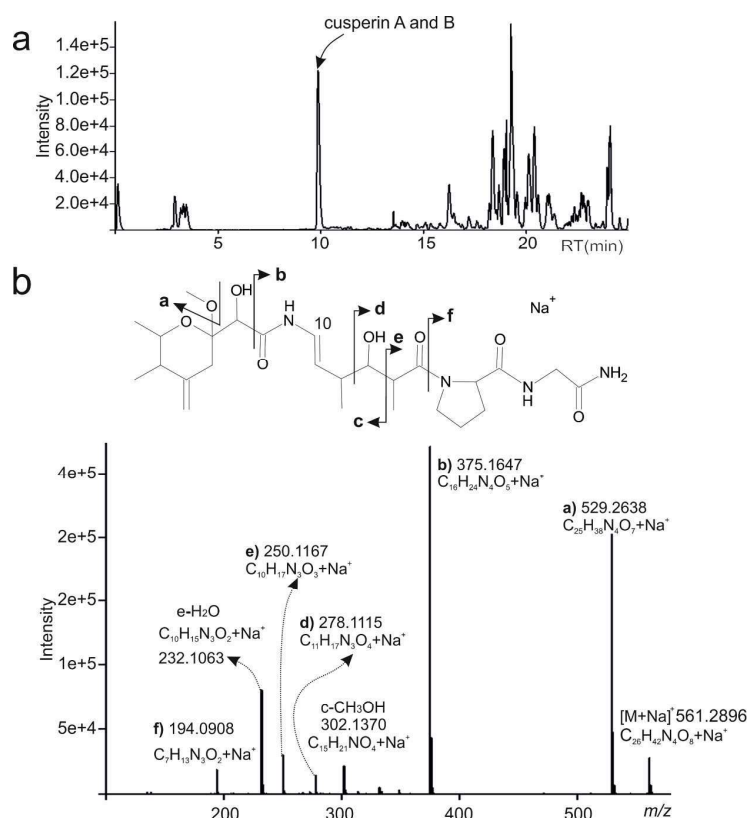


9

10

11 **Figure 3.** Structural analysis of cusperin A. a) HPLC-HRMS/MS base peak chromatogram of

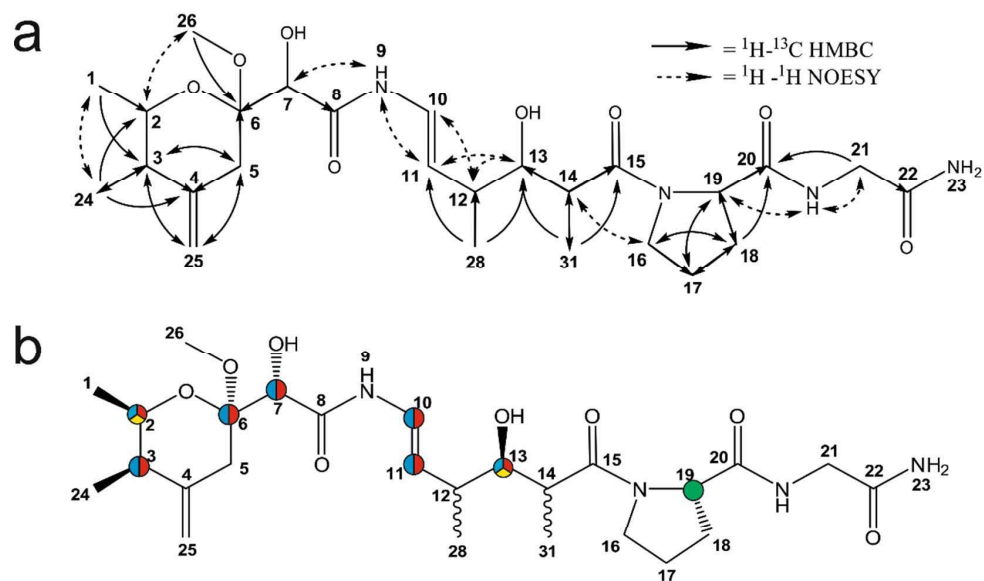
12 *Cuspidothrix* crude extract. b) Fragmentation pattern and HRMS/MS spectrum of cusperin A.



1

2 **Figure 4.** NMR analysis and stereochemical assignment of cusperin A. a) <sup>13</sup>C HMBC (solid  
 3 arrows) and <sup>1</sup>H-<sup>1</sup>H NOESY (dashed arrows) correlations of cusperin. b) Stereochemical  
 4 arrangement of cusperin chiral centres based on NMR and bioinformatic analysis. Blue,  
 5 domain organization identical to nosperin, supporting a similar arrangement of chiral centers;  
 6 red, NMR coupling constants and shifts; yellow, ketoreductases (KRs) predicted specificity  
 7 based on a specific sequence motif (the presence of a conserved Asp residue); green,  
 8 nonribosomal peptide synthase (NRPS) domain prediction (absence of an epimerase domain).

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



1

2

23 *Supporting Information Available:* This material is available free of charge via the Internet.

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60