GNPS-guided discovery of xylacremolide C and D, evaluation of their putative biosynthetic origin and bioactivity studies of xylacremolide A and B†

Felix Schalk,‡a Janis Fricke,‡a Janis Fricke,‡a,‡b Soohyun Um,a Benjamin H. Conlon,‡b Hannah Maus,‡c Nils Jäger,d Thorsten Heinzel,d Tanja Schirmeister,‡c Michael Poulsen‡b and Christine Beemelmanns*a

Targeted HRMS2-GNPS-based metabolomic analysis of Pseudoxylaria sp. X187, a fungal antagonist of the fungus-growing termite symbiosis, resulted in the identification of two lipopeptidic congeners of xylacremolides, named xylacremolide C and D, which are built from D-phenylalanine, L-proline and an acetyl-CoA starter unit elongated by four malonyl-CoA-derived ketide units. The putative xyla gene cluster was identified from a draft genome generated by Illumina and PacBio sequencing and RNAseq studies. Biological activities of xylacremolide A and B were evaluated and revealed weak histone deacetylase inhibitor (HDACi) and antifungal activities, as well as moderate protease inhibition activity across a panel of nine human, viral and bacterial proteases.

Introduction

While most fungal species belonging to the genus Xylaria (Ascomycota: Xylariaceae) are of saprotrophic nature and globally abundant,1,2 members of the subgenus Pseudoxylaria thrive predominantly on deteriorating comb material of fungus-cultivating termites that live in an obligate symbiosis with specialized fungal cultivars in the genus Termimycyes (Basidiomycota).3,4 Our recent studies focused on the relationship between the termites’ fungal cultivar and the competitive and/or antagonistic behaviour of Pseudoxylaria to better understand their coevolutionary relationship.5,6 Based on the metabolomic analysis of fungus-fungus co-cultures, we found that isolate Pseudoxylaria sp. X802 produces several biologically active small molecules including the antibacterial tetracyclic peptides pseudoxylallamycins and epoxy-cytochalasins.7,8 Following up on this, we performed a high-resolution tandem mass spectrometry (HRMS3)-based analysis of guttation droplets that appear on aging aerial hyphae of the isolate X187 (Fig. 1a).9 Processing of the MS2-data using Global Natural Products Social Molecular Networking Analysis (GNPS)10 revealed the production of four linear peptides called pseudoxyllaramides A–D and two PKS-NRPS-based hybrids named xylacremolides A and B.11 The identified xylacremolides share similar structural features with acremolides12 and FR235222,13 a known histone deacetylase inhibitor, isolated from Acro- monium spp. and saroclides from the mangrove-derived fungus Sarocladium kiliei HDN11-112.14 Intrigued by the high abundance of xylacremolide A and B in guttation droplets of Pseudoxylaria sp. X187 (from now on named X187), and inspired by the findings that fungal PKS-based megaenzymes are often inherently promiscuous, we intensified our metabolomic and genomic analysis of X187. Here, we report for the first time on the characterization of two novel xylacremolide congeners named xylacremolide C and D, and on detailed bioactivity studies of xylacremolide A and B. In silico analysis of genomic and RNA sequencing data of X187 allowed us to propose a putative biosynthetic PKS-NRPS-based gene cluster of xylacremolides in Pseudoxylaria for the first time.

Results and discussion

Metabolomic analysis and isolation

HRMS3-based metabolomic analysis of guttation droplet of X187 and subsequent GNPS-based visualization of the acquired MS3-data was performed according to an established procedure (Fig. 1b and S1†).11 GNPS-networks were screened for molecular ion peaks belonging to xylacremolide A and B and putative structural congeners that varied in chain length and oxidation
of the PKS moiety. The identified m/z subnetwork related to xylacremolide A and B (m/z 403.223 (3), m/z 417.238 (4)) contained a total of 50 connected nodes with molecular ion peaks that were partially assigned to congeners carrying additional methylene units (Δm/z 14.016, Δm/z 28.032) and derivatives with varying C₃H₄OH moieties (Δm/z 45.037). The subcluster containing molecular ion peaks of 3 and 4 was connected by a single node (m/z 443.252) to two additional subclusters containing the molecular ion peaks m/z 461.263 and m/z 459.248, respectively. Subsequent dereplication of HRMS values using Antibase and Scifinder indicated that X187 might produce yet unreported xylacremolide congeners.†

We then evaluated the relative abundance of identified molecular ion peaks in culture extracts obtained from different growth conditions. Methanolic extracts were obtained from cultures grown on potato-dextrose-agar (PDA; samples taken after 18, 28 and 42 days) and ISP2 medium (samples taken after 18 and 42 days). Subsequent comparative HRMS-analysis of culture extracts revealed a very stable xylacremolide composition within all extracts with xylacremolide B as the main product, followed by non-methylated derivative xylacremolide A (Fig. 1c, S2 and Table S1†). Although the abundance of other molecular ion peaks belonging to putative new congeners was found to be about one order of magnitude lower compared to xylacremolide B, we pursued a MS-guided isolation of those molecular ions peaks (Fig. 1).

For this purpose X187 was grown on 30 plates (PDA, 26 °C) and mycelium-covered agar plates were extracted with methanol after 28 days. Culture extracts were subjected to MS-guided semi-preparative liquid chromatography yielding two new xylacremolide congeners, xylacremolide named C (1) and D (2) in analytical purity (Fig. 2 and S3†). The first isolated molecule (HR-ESI-MS m/z [M + H]+ calcd for C₂₅H₃₇O₆N₂ 461.2646, found 461.2644, Δppm = −0.434) was structurally analyzed by comparative 1D (1H and 13C) and 2D (HSQC, COSY, and HMBC) NMR spectroscopy (Tables 1, S2 and Fig. S6–S11†). On the basis of the molecular formula, nine degrees of unsaturation were conjectured. The 1H NMR spectrum of 1 showed features of a peptidic compound bearing two α-proton signals [δH 4.85 and 4.25] and one NH signal [δH 8.46]. Also, three carbonyl carbon signals [δC 174.9, 171.9, and 169.8] and two α-carbon signals [δC 59.7 and 51.3] based on the HSQC and HMBC spectra supported that compound 1 is a peptide-derived compound (Table S1†). Comparative analysis of 1H and 2D NMR data of 1 and revealed two amino acid residues including a proline (Pro) and a phenylalanine (Phe) and 3,5,7-trihydroxy-2-methyldecanoic acid (TMDA). The α-proton [H-2, δH 4.85] of Pro showed homonuclear correlations to H-3 [δH 2.02 and 1.79]. Additional 1H–1H correlations of H-3/H-4 [δH 1.80 and 1.45], H-4/H-5 [δH 3.79 and 3.37] and respective HMBC correlations from H-2 to C-3 [δC 29.5], H-3 to C-2 and C-4 [δC 59.7 and 24.9], and H-5 [δC 3.79, 3.72] to C-3 and C-4 supported the assumption of a ring structure (Fig. 2b).

Similarly, the α-proton [H-7, δH 4.85] of Phe revealed 1H–1H-COSY correlations to H-8 [δH 2.92 and 2.86] and the amide proton [δH 8.46]. The connectivities were supported by HMBC correlations from the α-proton H-7 to C-8 [δC 36.7], from the H-8
Table 1  Chemical shift values of compound 1 and 2 (600/125 MHz, DMSO-d6)

| Nr | Chemical Shifts | Chemical Shifts |
|----|----------------|----------------|
|    | 1              | 2              |
|    | 13C           | 1H (J in Hz)   | 13C            | 1H (J in Hz)   |
| 1  | 171.9, s      | —              | 174.6, s       | —              |
| 2  | 59.7, d       | 4.26, dd (8.5, 2.5) | 59.4, d       | 4.32, m        |
| 3  | 29.5, t       | 2.02, m        | 26.8, t        | 2.02, m        |
| 4  | 24.9, t       | 1.79, m        | 1.79, m        | 1.79, m        |
| 5  | 47.1, t       | 3.79, m        | 47.0, t        | 3.78, m        |
| 6  | 169.8, s      | —              | 167.1, s       | —              |
| 7  | 51.3, d       | 4.85, ddd (9.0, 7.5, 6.0) | 51.4, d       | 4.94, ddd (9.0, 7.5, 6.0) |
| 8  | 36.7, t       | 2.92, dd (13.5, 6.0) | 36.0, t       | 2.92, dd (13.5, 6.0) |
| 9  | 138.2, s      | —              | 138.1, s       | —              |
| 10, 14 | 129.7, d  | 7.25, m        | 129.7, d       | 7.26, m        |
| 11, 13 | 128.4, d  | 7.25, m        | 128.4, d       | 7.26, m        |
| 12 | 126.7, d      | 7.17, m        | 126.7, d       | 7.18, m        |
| NH | 8.46, d       | (7.5)          | 8.96, d        | (7.5)          |
| 15  | 174.7, s      | —              | 170.8, s       | —              |
| 16  | 42.1, d       | 2.61, m        | 51.5, d        | 3.58, q (6.5)  |
| H25 | 13.5, q       | 0.91, d (7.0)  | 13.6, q        | 0.91, d (6.5)  |
| 17  | 71.3, d       | 3.74, m        | 205.5, s       | —              |
| 18  | 43.4, t       | 1.46, m        | 50.6, t        | 2.70, m        |
| 19  | 64.4, d       | 3.77, m        | 63.2, d        | 4.28, m        |
| 20  | 41.7, t       | 1.83, m        | 33.9, t        | 1.85, m        |
| 21  | 74.3, d       | 1.53, m        | 1.53, m        | —              |
| 22  | 33.8, t       | 4.64, m        | 74.3, d        | 4.68, m        |
| 23  | 19.0, t       | 1.60, m        | 39.2, t        | 1.63, m        |
| 24  | 13.5, q       | 0.89, dd (7.5, 7.5) | 13.4, q       | 0.87, dd (7.5, 7.5) |

The trihydroxy-methylenecanoic acid (TMDA) was deduced by 1H–1H correlations from the methyl protons [H5–25, δH 0.89] to the methine proton [H–16, δH 2.61] bearing three hydroxyl carbons [C–17, C–19, and C–21; δC 71.3, 64.4, and 74.3] and one methyl group [C–25; δC 13.3; δH 0.91] at C-16 [δC 42.1] supported by HMBC correlations. The connectivity of the two amino acid residues was confirmed with the NOESY correlation between H-5 of Pro and H-7 of Phe. The amide proton [δH 7.17] at C-16 [δC 129.7] to C-10/C-14 [δC 126.7] of Phe. The benzene ring in the phenylalanine was determined by COSY and HMBC correlations between H-10/H-14 [δH 7.25] and C-12 [δC 126.7] and C-9, as well as from H-12 [δH 7.17] to C-10/C-14 [δC 129.7].

The trihydroxy-methylenecanoic acid (TMDA) was deduced by 1H–1H correlations from the methyl protons [H5–25, δH 0.89] to the methine proton [H–16, δH 2.61] bearing three hydroxyl carbons [C–17, C–19, and C–21; δC 71.3, 64.4, and 74.3] and one methyl group [C–25; δC 13.3; δH 0.91] at C-16 [δC 42.1] supported by HMBC correlations. The connectivity of the two amino acid residues was confirmed with the NOESY correlation between H-5 of Pro and H-7 of Phe. The amide proton [δH 7.17] at C-16 [δC 129.7] to C-10/C-14 [δC 126.7] of Phe. The benzene ring in the phenylalanine was determined by COSY and HMBC correlations between H-10/H-14 [δH 7.25] and C-12 [δC 126.7] and C-9, as well as from H-12 [δH 7.17] to C-10/C-14 [δC 129.7].

Based on the assumption that xylacremolides share the same biosynthetic origin and the similar NMR chemical shift pattern, we deduced the identical stereocchemical assignment for compound 1 and 2 as shown for xylacremolide B (4) by X-ray crystallography, while the configuration at position C-21 remained unsolved.

Biosynthesis

To gain better insights into the enzymatic machinery responsible for the production of the xylacremolides, we thus sequenced X187 using Illumina MiSeq and PacBio sequencing technology that yielded 499 contigs in a 39 Mbp draft genome assembly. We then set out to identify the putative biosynthetic pathway that was deduced to be of NRPS and PKS-based origin. 10–17 Mining of the obtained draft genome sequence using FungiSmash™ revealed only one biosynthetic gene cluster (BGC) region, designated as the xya cluster (Fig. 3; GenBank accession number MW579544), that encodes for a highly-
reducing PKS (HRPKS) xyaA (KS-AT-DH-CM-ER-KR), a bimodular nonribosomal peptide synthase (NRPS) xyaB (C_1\_A_2\_T_1\_E_1\_C_2\_A_2\_T_2\_C_2), a major-facilitator superfamily (MFS) transporter xyaC and two short-chain dehydrogenases/reductases (SDR) xyaD/xyaE. Subsequent RNA-Seq analysis showed that all xya genes were transcribed under producing conditions. In silico analysis of the putative substrate preference of the A domains of XyaB using NRPSPredictor2 revealed a preference for D-Phe for A_1 and a hydrophobic amino acid with L-Pro being suggested by the closest neighbour prediction for A_2, which matches the structure of the xylacremolides. Additionally, the encoded proteins XyaA and XyaB showed high similarity (70% and 65% similar amino acid sequence) to the characterized ValA and ValB from Aspergillus terreus ATCC 20542, responsible for the

![Fig. 3](image-url)

Fig. 3 (a) Structures of acremolide (Acremonium sp.), saroclide (Sarocladium kiliense), tolyprolinol (Tolypocladium sp. FKI-7981), colletotrichamide (Colletotrichum gloeosporioides JS419), and valactamide (Aspergillus terreus). (b) Schematic representation of the xya BGC in Pseudoxylaria sp. X187 and comparison to homologues BGCs from Aspergillus terreus ATCC 20542, Tolypocladium capitatum CBS 113982, Colletotrichum gloeosporioides Lc1 and Acremonium chrysogenum ATCC 11550, which were identified using BLAST. Shown are genes coding for highly-reducing polyketide synthases (dark blue: HR-PKS), non-ribosomal peptide synthases (light blue: NRPS), short-chain dehydrogenases/reductases (red: SDR) and major facilitator superfamily transporters (dark grey: MFS). Non related genes are shown in light grey. Sequence similarities of the corresponding proteins of the homologues BGCs regarding XyaA/B are shown in percentages. (c) Proposed biosynthesis of the xylacremolides.
biosynthesis of the structurally related lipopeptides valactamide A–H20 (Fig. 3a). Furthermore, homologues BGCs have been identified in silico in other ascomycetes (Fig. 3b and Table S2†). Among them Tolypocladium intermedia, Colletotrichum gloeosporioides and Acremonium chrysogenum, produce toxypoline,21 colletorichiamides22 and acremolides,23 which are all lipopeptides comprised of two amino acids and an aliphatic part pointing towards a similar biosynthetic origin (Fig. 3a). Although these three predicted BGCs have not yet been biochemically characterized, our findings support the assignment of the xya BGC and could set the stage for further biomolecular and biotechnological investigations.

The proposed biosynthetic assembly line for xylacremolides (Fig. 3c) is consistent with both the domain architectures of XyaA and XyaB and their molecular structure. Here, we propose that the KS domain of the iterative HRPKS XyaA incorporates an acetyl starter unit, followed by subsequent elongation with either three or four malonyl-CoA derived ketide units. The first ketide is fully reduced to an alkane by the KR, DH and ER domains, whereas the following units are partially reduced to the corresponding alcohols. Notably, the CMeT domain catalyzes the α-methylation of five membered polyketide chains, but can bypass this reaction in the four-membered products, which leads to the production of 3. The C1 domain of XyaB then catalyzes the amide bond formation of the tetra- and pentaketide chains with ω-Phe, which was isomerized by the included E domain. Chain elongation is then catalyzed by the C2 domain incorporating ω-Pro, followed by macro lactonization by the C7 domain releasing the main products, which is typical for fungal NRPSes.26–28 Subsequent tailoring oxidation reaction of the secondary alcohol (1) to the keto functionality of 2 could be catalyzed by either XyaD or XyaE; however, incomplete reduction by XyaA cannot be excluded at this stage. Here, it is again noteworthy that congeners 1 and 2 are produced in a 2 : 1 ratio almost independently from the tested cultivation conditions (Table S1†).

Bioactivities of xylacremolides A and B

Similar to their closest structural counterparts (acremolides), xylacremolides revealed so far no cytotoxic or antibacterial activities.21 We then questioned which other cellular and/or ecological role xylacremolides might impose. A literature survey revealed that structurally related fungal metabolites trappoxin and FR235222 were previously investigated and found to be potent inhibitors of histone deacetylase (HDAC),23,24 and are thus being investigated in cancer therapy, as well as in parasitic and inflammatory diseases. To test for HDACi activity, we selected the two most abundant xylacremolide derivatives, xylacremolides A and B, and found weak inhibitory activity of both compared to the positive controls valproic acid and trichostatin A (tested concentration range: 0.25–1.0 μM; Fig. S4†).

We then tested both compounds for antifungal activity against Aspergillus nidulans RMS011 using a spore germination assay with carboxine and amphotericin B as positive controls (Figs. 4 and S5†).25 Here, a moderate inhibitory effect of both compounds on spore germination was observed when spores were incubated for less than 20 h, while longer incubation times caused a significant decline in the inhibitory activity. Thus, it can be concluded that xylacremolides delays spore germination to moderate extent, but once A. nidulans is growing, it either tolerates or even detoxifies xylacremolides.

In light of the ecological origin of Pseudoxylaria sp. X187 (co-evolved antagonistic fungus of Termiformes), it could be speculated at this point that the inhibitory activity of xylacremolides, while moderate, allows in combination with other metabolites an overall growth advantages and supports the successful outcompetition of the fungal cultivar when colony homeostasis is disrupted.

In line with our long-term mission to identify pharmacologically relevant protease inhibitors, we tested the inhibitory activity of the two most abundant derivatives, xylacremolides A and B, on various pharmacologically important proteases using an established fluorescence-based assay (Table 2).26–29 Both rhodensin (Rd), and cruzain, cysteine proteases belonging to the papain-family, play central roles in the life cycle of the parasitic protists Trypanosoma brucei (Human African Trypanosomiasis (HAT))30 and Trypanosoma cruzi (Chagas-disease),31 respectively. Three proteases were selected for their important functions in cancer cells (Cath L, Cath B, urokinase (uPA)).22 Sortase A (SrtA)

![Fig. 4 Antifungal activity assay of compound 3 and 4 and enriched SPE fraction of Pseudoxylaria sp. X187 crude extract (SPE) against Aspergillus nidulans RMS011. Carboxine (Cbx) and amphotericin B (Amp) served as positive controls. All substrates were tested at 1 μg mL⁻¹. The percentage of 0% inhibition is considered normal growth; while 100% are considered complete growth inhibition.](image)

Table 2 Protease inhibitor activity screening of 3 and 4

| Protease     | Compound 3 [%] | Compound 4 [%] |
|--------------|---------------|---------------|
| Cathepsin L  | 7.3 ± 3.3     | 11 ± 5.0      |
| Cathepsin B  | 4.9 ± 1.4     | n.i.          |
| uPA          | n.i.          | n.i.          |
| SrtA         | 3.8 ± 6.4     | 2.5 ± 4.2     |
| Rhodesain    | 5.7 ± 5.2     | 15 ± 4.6      |
| Cruzain      | 3.6 ± 2.3     | 3.8 ± 5.4     |
| Mpro (SARS-CoV) | n.i.         | 11 ± 9        |
| PLpro (SARS-CoV) | n.i.         | 11 ± 14       |
| DENV2        | n.i.          | 4.8 ± 4.8     |
| ZIKV         | 0.7 ± 1.2     | 17 ± 3.4      |

a Inhibition at a compound concentration of 200 μM in DMSO as a percentage mean value ± standard deviation with n = 3. Inhibition rates below 0.5% are given as n.i. (no inhibition).
was selected as promising antibacterial target due to its protease and transpeptidase activity related to the formation of the peptidoglycan layer of *Staphylococcus aureus.* For antiviral activity testing, we selected four proteases (ZIKV NS2B-NS3 protease, DENV2-NS2B-NS3, Mpro (also known as 3C like protease and which is the main protease of the novel coronavirus SARS-CoV2), and PLpro). Initial inhibitor screening was performed at compound concentrations of 200 μM in DMSO, which revealed mild inhibition of cruzain, cathepsin L, DENV2 virus SARS-CoV2), protease and transpeptidase activity related to the formation of a network of biochemical interactions. The data was activated and equilibrated SPE C18ec column (1 g, Macherey-Nagel). The data was then clustered with MS-Cluster with a parent mass tolerance of 0.02 Da and a MS/MS fragment ion tolerance of 0.02 Da to create consensus spectra. Further, edges between two nodes were kept in the network if they shared at least 90% of their consensus spectra that contained less than 2 spectra were discarded. A network was then created where edges were weighted by the number of shared peaks. Finally, the maximum size of a molecular family was determined by removing all MS/MS peaks within a 17 Da of the consensus spectra. The data was filtered by removing all MS/MS peaks within ±17 Da of the precursor m/z. MS/MS spectra were window filtered by choosing only the top 6 peaks in the ±50 Da window throughout the spectrum. The data was then clustered with MS-Cluster with a parent mass tolerance of 0.02 Da and a MS/MS fragment ion tolerance of 0.02 Da to create consensus spectra. Further, consensus spectra that contained less than 2 spectra were discarded. A network was then created where edges were filtered to have a cosine score above 0.7 and more than 6 matched peaks. Further, edges between two nodes were kept in the network if each of the nodes appeared in each other’s respective top 5 most similar nodes. Finally, the maximum size of a molecular family was set to 50, and the lowest scoring edges were removed from molecular families until the molecular family size was below this threshold.

**Experimental**

**Compound identification based on GNPS analysis**

*Pseudoxylaria* sp. X187-2 was grown on 30 potato-dextrose (PD) agar plates (0.8 L of PDA, 2% agar, standard plates 16 mm × 92 mm, 20 mL per plate) for four weeks at room temperature. Mycelium-covered agar plates were cut into small pieces (0.5 cm × 0.5 cm), extracted twice with 300 mL MeOH, and methanolic extracts dried under vacuum. Crude extracts were subjected to UPLC-ESI-HRMS measurements (Dionex Ultimate 3000 system combined with a Q-Exactive Plus mass spectrometer equipped with an electrospray ion (ESI) source (Thermo Scientific)). The resulting HRMS² was subjected to MS² fragmentation based network analysis carried out on the global natural products social molecular networking platform (GNPS). The data was filtered by removing all MS/MS peaks within ±17 Da of the precursor m/z. MS/MS spectra were window filtered by choosing only the top 6 peaks in the ±50 Da window throughout the spectrum. The data was then clustered with MS-Cluster with a parent mass tolerance of 0.02 Da and a MS/MS fragment ion tolerance of 0.02 Da to create consensus spectra. Further, consensus spectra that contained less than 2 spectra were discarded. A network was then created where edges were filtered to have a cosine score above 0.7 and more than 6 matched peaks. Further, edges between two nodes were kept in the network if each of the nodes appeared in each other’s respective top 5 most similar nodes. Finally, the maximum size of a molecular family was set to 50, and the lowest scoring edges were removed from molecular families until the molecular family size was below this threshold.

**Analysis of production levels**

*Pseudoxylaria* sp. X187-2 was cultivated on 30 PDA 16 × 92 mm, 20 mL per plate, 2% agar) and plates incubated for four weeks at room temperature. Mycelium-covered agar plates were cut into small pieces and extracted twice with 600 mL MeOH. Extracts were filtered through filter paper, concentrated under reduced pressure, and enriched using C18 SPE cartridges as described before. Xylacremolides A-D eluted predominantly as a mixture within the 60% MeOH fraction with some residual elution in the 80% MeOH fraction from the SPE column. Fractions were concentrated under reduced pressure to yield a white-brown solid material which was subjected to further purification by semi-preparative HPLC until analytical compounds were obtained.

**Large-scale cultivation**

*Pseudoxylaria* sp. X187-2 was cultivated on 30 PDA 16 × 92 mm, 20 mL per plate, 2% agar) and plates incubated for four weeks at room temperature. Mycelium-covered agar plates were cut into small pieces and extracted twice with 600 mL MeOH. Extracts were filtered through filter paper, concentrated under reduced pressure, and enriched using C18 SPE cartridges as described before. Xylacremolides A-D eluted predominantly as a mixture within the 60% MeOH fraction with some residual elution in the 80% MeOH fraction from the SPE column. Fractions were concentrated under reduced pressure to yield a white-brown solid material which was subjected to further purification by semi-preparative HPLC until analytical compounds were obtained.

**Structure elucidation**

Xylacremolide C (1). White solid; [α]D²⁵ −0.18 (c 0.1 w/v%, MeOH); UV (acetonitrile), λmax 201, 219 nm; IR (ATR) rmax 3743, 3621, 3311, 2702, 2322, 1616, 1562 cm⁻¹; NMR spectral data, see Table S2† HRMS (ESI) m/z [M + H⁺] calecd for C25H34O6N2 461.2646, found 461.2644, Δppm = 0.434.

Xylacremolide D (2). White solid; [α]D²⁵ −1.63 (c 0.1 w/v%, MeOH); UV (acetonitrile), λmax 201, 219 nm; IR (ATR) rmax 3412, 3290, 2821, 2704, 2325, 2023, 1660 cm⁻¹; NMR spectral data, see Table S2† HRMS (ESI) m/z [M + H⁺] calecd for C25H34O6N2 459.2490, found 459.2485, Δppm −1.088.

**Genome sequencing and in silico analysis**

DNA extractions for whole-genome sequencing were performed using a scaled-up CTAB extraction. Whole-genome sequencing was performed using a 150 bp paired-end shotgun (BGIseq) and long-read (Pacific sequencer) sequencing by BGI (Hongkong). Sequencing results were checked for quality using FastQC version 0.11.8 (ref. 38) and MultiQC version 1.7 (ref. 39). Kmer depth was calculated using Jellyfish version 2.2.10 (ref. 40) and Kmer-based estimates of genome size, heterozygosity and repeat content generated using GenomeScope.

A hybrid *de novo* genome assembly, combining BGIseq and PacBio data, was performed using SPAdes version 3.13.0 (ref. 41). Gen models were predicted with Augustus (ref. 42). Homologues gene clusters were identified using the protein–protein BLAST (blastp) function from the NCBI. RNA-Seq data was mapped against the identified
Protease inhibition assays.26-29 Compounds were dissolved in DMSO (20 mM). Fluorometric assays were performed at a concentration of 200 μM. The inhibitory effect of compounds was determined for the following enzymes: rhodamine, cruzain, cathepsin L, cathepsin B, SARS-CoV2 MPro, SARS-CoV PLpro, DENV2 NS2B-NS3, ZIKV NS2B-NS3, urokinase (uPA), sortase A (Sta). Assay conditions were: 45 μL buffer, 1.25 μL enzyme in buffer, 2.5 μL compound in DMSO, 1.25 μL substrate in DMSO (total volume 50 μL). The increase of the fluorescence was recorded over a period of 10 min in intervals of 30 s (exception: Sta 30 min). Inhibition was calculated from the slope of fluorescence increase in the presence of the compound in relation to the slope in the presence of DMSO.

Buffers and substrates. Rhodamine (50 mM Na-acetate pH 5.5, 5 mM EDTA, 200 mM NaCl, 5 mM DTT, 10 μM Z-Phe-Arg-AMC); cruzain (50 mM Na-acetate pH 5.5, 5 mM EDTA, 200 mM NaCl, 5 mM DTT, 5 μM Z-Phe-Arg-AMC); cathepsin L (50 mM Tris pH 6.5, 5 mM EDTA, 200 mM NaCl, 2 mM DTT, 6.25 μM Z-Phe-Arg-AMC); cathepsin B (50 mM Tris pH 6.5, 5 mM EDTA, 200 mM NaCl, 1 mM DTT, 100 μM Z-Phe-Arg-AMC); SARS-CoV2 MPro (20 mM Tris pH 7.5, 0.1 mM EDTA, 200 mM NaCl, 1 mM DTT, 50 μM Dabcyl-KTSAVLQS GFRKM-Edans); SARS-CoV PLpro (20 mM Tris pH 7.5, 0.1 mM EDTA, 200 mM NaCl, 1 mM DTT, 50 μM Z-Arg-Leu-Arg-Gly-Gly-AMC); DENV2 and ZIKV NS2B-NS3 (50 mM Tris pH 9.0, 20% glycerol, 1 mM Chaps, 100 μM Boc-Gly-Arg-AMC); urokinase (50 mM Tris pH 7.4, 50 mM NaCl, 0.5 mM EDTA, 240 μM Z-Gly-Gly-Arg-AMC), sortase A (5 mM Tris pH 7.5, 15 mM NaCl, 0.5 mM CaCl₂, 0.5 mM Gly₄, 25 μM Abz-LPETG-Dap(dnp)-OH).

Conclusions

Reinvestigation and intensified HRMS²-GNPS-based metabolomic analysis of Pseudoxylaria sp. X187 resulted in the identification of two additional lipopeptide congeners of xylacremolides, which we named xylacremolide C and D. While xylacremolide A and B are built from three malonyl-CoA derived ketide units, the newly identified congeners are built from four malonyl-CoA derived ketide units, which indicates towards a promiscuous PKS. Evaluation of possible biological activities revealed moderate to weak antifungal and HDACi activity, as well as moderate protease inhibition activities across a panel of nine human, viral and bacterial proteases. Although their single bioactivities alone are likely not responsible for the antagonistic behaviour of Pseudoxylaria sp. X187, the combined actions might allow for a considerable growth advantage.

Conflict of interest

There are no conflicts to declare.

Acknowledgements

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy – EXC 2051 with Project-ID 390713860 and CRC 1127 with Project-ID 239748522 to CB and TH. Funded by the European Research Council (ERC-CoG – 771349) and The Danish Council for Independent Research (DFF – 701400178) to MP. We thank the Humboldt Foundation for a postdoctoral fellowship to SU, and Mrs Heike Heinecke (Hans Knöll Institute) for recording NMR spectra.

Notes and references

1 F. Song, S.-H. Wu, Y.-Z. Zhai, Q.-C. Xuan and T. Wang, Chem. Biodiversity, 2014, 11, 673–694.
2 S. E. Helaly, B. Thongbai and M. Stadler, Nat. Prod. Rep., 2018, 35, 992–1014.
3 M. Poulsen, H. Hu, C. Li, Z. Chen, L. Xu, S. Otani, S. Nygaard, T. Nobre, S. Klaubauf, P. M. Schindler, F. Hauser, H. Pan, Z. Yang, A. S. M. Sonnenberg, Z. Wilhelm De Beer, Y. Zhang, M. J. Wingfield, C. J. P. Grimmelikhuijzen, R. P. De Vries, J. Korb, D. K. Aanen, J. Wang, J. J. Boomsma and G. Zhang, Proc. Natl. Acad. Sci. U. S. A., 2014, 111, 14500–14505.
4 S. Otani, V. L. Challinor, N. B. Kreuzenbeck, S. Kildgaard, S. Krath Christensen, L. L. M. Larsen, D. K. Aanen, S. A. Rasmussen, C. Beemelmanns and M. Poulsen, Sci. Rep., 2019, 9, 1–10.
5 A. A. Visser, V. I. D. Ros, Z. W. De Beer, A. J. M. Debets, E. Hartog, T. W. Kuyper, T. Lessoe, B. Slippers and D. K. Aanen, Mol. Ecol., 2009, 18, 553–567.
6 A. A. Visser, P. W. Kooij, A. J. M. Debets, T. W. Kuyper and D. K. Aanen, Fungal Ecol., 2011, 4, 322–332.
7 H. Guo, N. B. Kreuzenbeck, S. Otani, M. Garcia-Altares, H. M. Dahse, C. Weigel, D. K. Aanen, C. Hertweck, M. Poulsen and C. Beemelmanns, Org. Lett., 2016, 18, 3338–3341.
8 H. Guo, A. Schmidt, P. Stephan, L. Raguz, D. Braga, M. Kaiser, H. M. Dahse, C. Weigel, G. Lackner and C. Beemelmanns, ChemBioChem, 2018, 19, 2307–2311.
9 M. Gareis and C. Gottschalk, Mycotoxin Res., 2014, 30, 151–159.
10 M. Wang, J. J. Carver, V. V. Phelan, L. M. Sanchez, N. Garg, Y. Peng, D. D. Nguyen, J. Watrous, C. A. Kapono, T. Luzzatto-Knaan, C. Porto, A. Bouslimani, A. V. Melnik,
41 G. W. Vurture, F. J. Sedlazeck, M. Nattestad, C. J. Underwood, H. Fang, J. Gurtowski and M. C. Schatz, *Bioinformatics*, 2017, 33, 2202–2204.
42 A. Bankevich, S. Nurk, D. Antipov, A. A. Gurevich, M. Dvorkin, A. S. Kulikov, V. M. Lesin, S. I. Nikolenko, S. Pham, A. D. Prjibelski, A. V. Pyshkin, A. V. Sirotkin, N. Vyahhi, G. Tesler, M. A. Alekseyev and P. A. Pevzner, *J. Comput. Biol.*, 2012, 19, 455–477.
43 M. Stanke, R. Steinkamp, S. Waack and B. Morgenstern, *Nucleic Acids Res.*, 2004, 32, 309–312.