JOURNAL OF NATURAL PRODUCTS

pubs.acs.org/jnp

Article

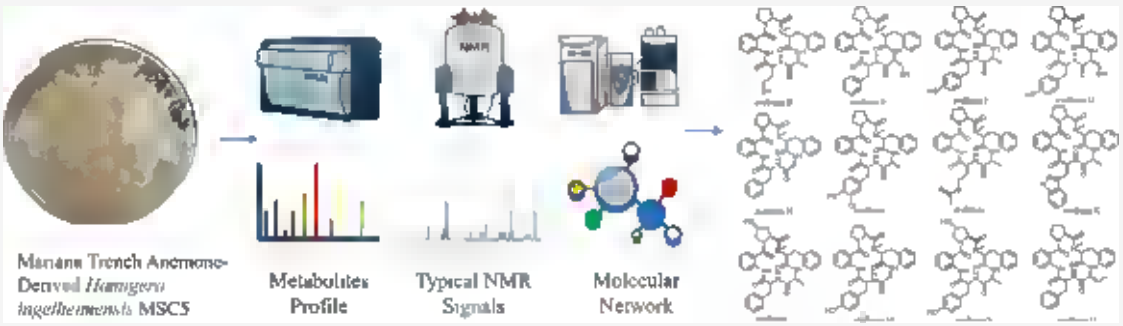
Cyclopeptide Avellanins D−O with Antimalarial Activity from the Mariana Trench Anemone-Derived *Hamigera ingelheimensis* MSC5

Hao Li, Yuling Chen, Bingqing Tang, Zhengjie Liu, Bo Peng, Jiajun Li, Han Gao, Sibao Wang, and Zhiyong Li*

Cite This: *J. Nat. Prod.* 2024, 87, 2695−2708

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



ABSTRACT: Marine microorganisms are a treasure trove of natural products, especially those in extreme marine environments, which may produce novel natural products. Herein, biosynthetic gene cluster analysis combined with an integrated metabolomic strategy incorporating matrix-assisted laser desorption ionization-time-of-flight mass spectrometry (MALDI-TOF MS), nuclear magnetic resonance (NMR), and liquid chromatography−mass spectrometry/mass spectrometry (LC-MS/MS) based Global Natural Products Social Molecular Networking (GNPS) was used to discover new compounds from the Mariana trench anemone-derived fungus *Hamigera ingelheimensis* MSC5. As a result, 12 new cyclic pentapeptides, avellanins D−O (**1**−**12**), were isolated, together with a known cyclic pentapeptide avellanin C (**13**). All the structures and absolute configurations were elucidated using NMR, mass spectrometry, X-ray diffraction analysis, and Marfey's method. A plausible biosynthetic pathway for the avellanins was proposed based on the gene cluster analysis of *H. ingelheimensis* MSC5. Bioassay revealed that compound **6** exhibited potent antimalarial activity with an IC_{50} value of 0.19 ± 0.09 μM.

Natural products (NPs) have long been a critical source of leading drugs, with almost half of the small-molecule drugs approved by the Food and Drug Administration (FDA) coming from NPs or their derivatives.[1] However, the low discovery rate of novel bioactive compounds has become a bottleneck.[2,3] To date, metabolomic strategies have been successfully used to mitigate this problem and accelerate the discovery of new NPs.[4−9] Meanwhile, with advances in sequencing technologies, genomic data have emerged as a valuable reservoir of unique biosynthetic gene clusters (BGCs), facilitating the discovery of novel NPs.[10] The integration of genomics and metabolomics has proven effective in uncovering novel NPs, including pseudonochelin,[11] ecteinamines,[12] and xylomyrocins.[13]

Deep-sea organisms are generally considered to be a potential reservoir of new NPs due to their unique habitat conditions.[14,15] Notably, over 70% of deep-sea NPs show bioactivity,[16,17] such as the deep-sea-derived analgesic agent ziconotide[18] and the orphan drug marizomib used for the treatment of multiple myeloma.[19] Despite these successes, the exploration of NPs from deep-sea extreme environments is still in its early stages,[20] and only 2% of the 35,000 known marine NPs[21,22] were isolated from the deep sea.[15,17]

Many species within the order Eurotiales are well-known for producing diverse bioactive secondary metabolites, such as those in the genera *Penicillium*, *Aspergillus*, and *Talaromyces*. However, the lesser-known genus *Hamigera* has only around 30−40 reported secondary metabolites, including anthraquinone pigments, cyclic peptides, and macrolides. In previous studies, most strains of *Hamigera* were isolated from soil, with only one strain derived from a sponge and another from beetles.[23−27] Thus far, no *Hamigera* species from deep-sea sources have been reported. Moreover, the extremely limited genomic data of *Hamigera* species, with only one report to

Received: June 26, 2024
Revised: November 14, 2024
Accepted: November 19, 2024
Published: November 27, 2024

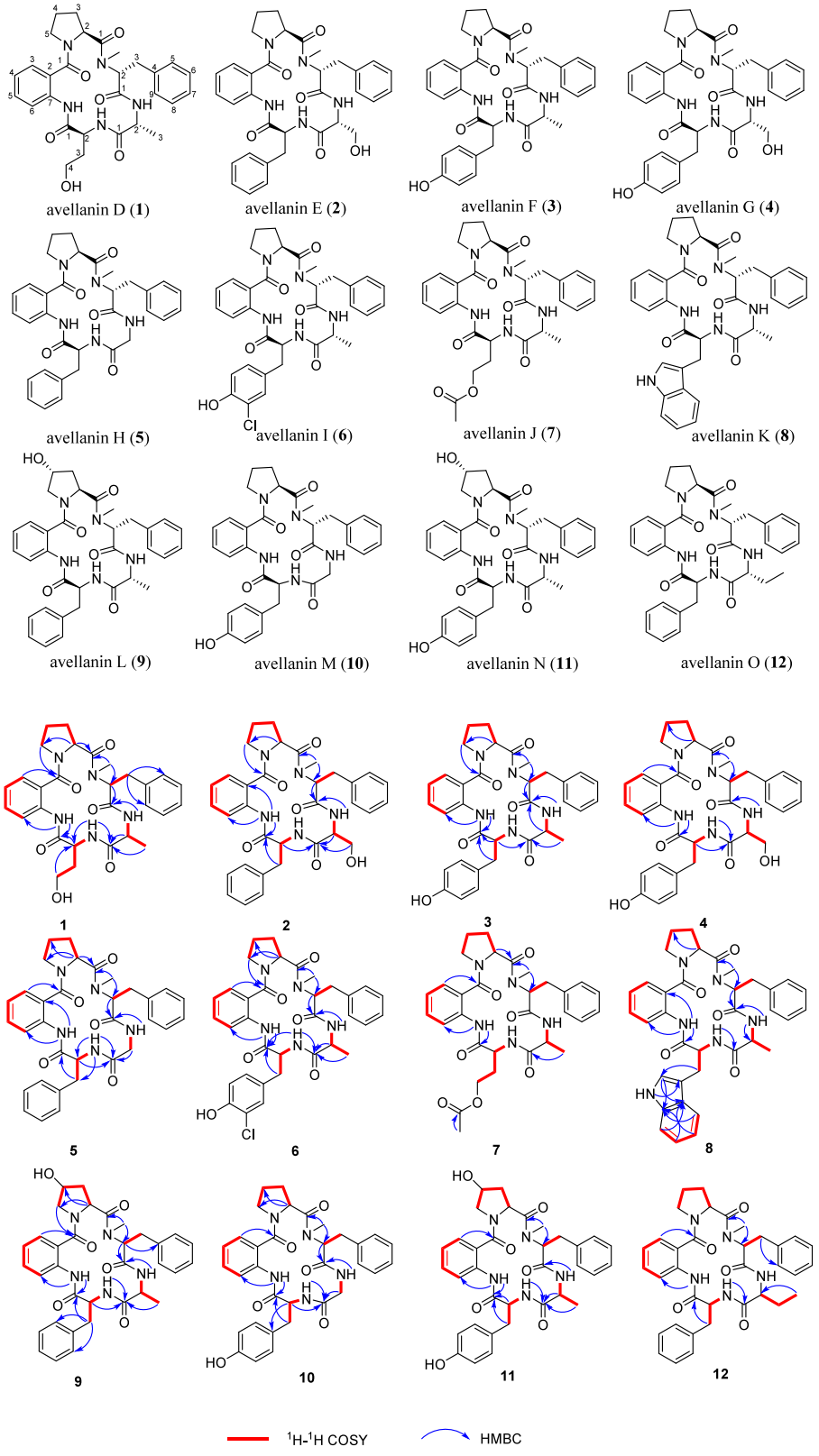


Figure 1. Key $^{1}H-^{1}H$ COSY and HMBC correlations of compounds **1**−**12**.

date,[28] has further hindered our understanding of their biosynthetic gene clusters and metabolic potential.

In recent years, various strategies including bioactivity-guided screening,[29] metabolic interaction,[30,31] precursor-directed biosynthesis,[32] and heterologous expression[33] have been used for microbial NP production in our lab. In this study, we comprehensively employed the genomic BGC analysis OSMAC (one strain many compounds) strategy and an integrated metabolomics platform incorporating MALDI-TOF-MS, NMR, and LC-MS/MS based GNPS molecular

https://doi.org/10.1021/acs.jnatprod.4c00740
J. Nat. Prod. 2024, 87, 2695−2708

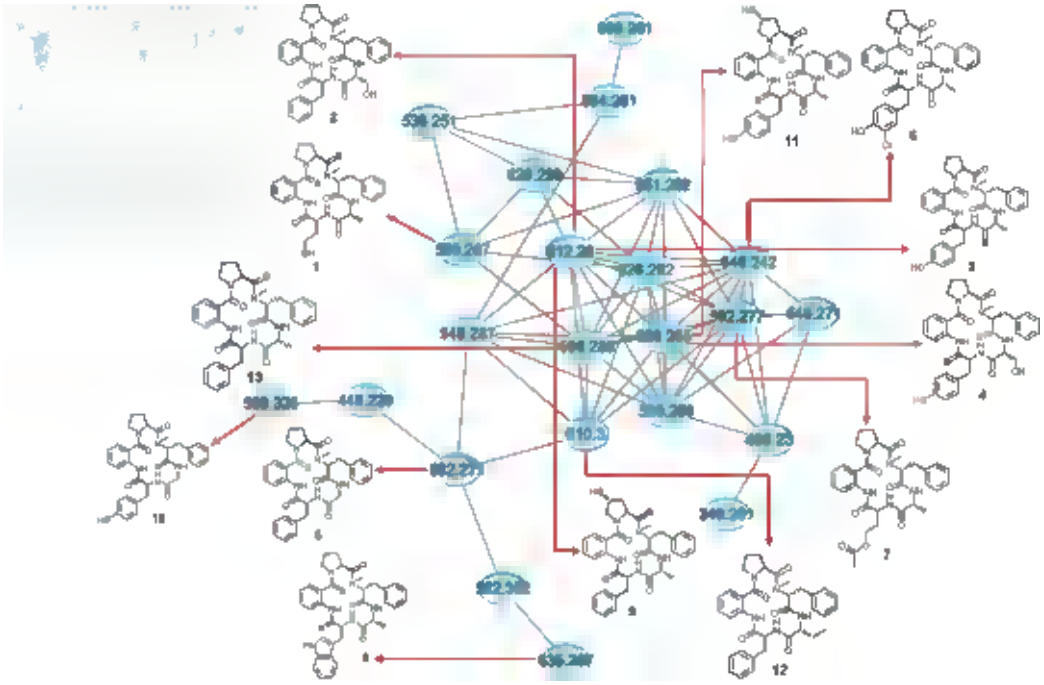


Figure 2. GNPS molecular network of the EtOAc extract from *Hamigera ingelheimensis* MSC5

networking, for the rapid discovery of novel NPs from the Mariana Trench anemone derived fungus *H. ingelheimensis* MSC5. As a result, 12 new cyclic pentapeptides, avellanins D−O (**1**−**12**), were isolated and characterized. Additionally, some of these compounds were evaluated for their antiproliferative and antimalarial activities.

RESULTS AND DISCUSSION

Biosynthetic Gene Cluster Analysis and OSMAC Strategy for the Fungus *H. ingelheimensis* MSC5. The whole genome of *H. ingelheimensis* MSC5 was sequenced on an Illumina NovaSeq combined with a PacBio Sequel platform, resulting in a 149× coverage, 40 scaffolds, and a genome size of 28.7 Mb. Twenty-seven putative BGCs were revealed with antiSMASH (Table S1),[34] including 10 nonribosomal peptide synthetase (NRPS) BGCs, which may be responsible for the production of peptide NPs. Next, an OSMAC strategy was adopted to explore optimal culture conditions for *H. ingelheimensis* MSC5 (Table S3). The result showed that rice medium was suitable for cultivation, considering the yield and chemical diversity of the NPs (Figure S1).

Metabolomics for the Discovery of New Cyclopeptides from the Fungus *H. ingelheimensis* MSC5. MALDI-TOF-MS enables the rapid, high-throughput screening of NPs from microbial culture extracts. ^{1}H NMR allows the rapid recognition of characteristic functional groups of the major compounds.[35] LC-MS/MS-based GNPS molecular networking enables the simultaneous analysis of numerous compounds and provides a visual representation of the chemistry present. By clustering metabolites with similar MS/MS fragmentation patterns, molecular networking facilitates the rapid identification of known compounds in databases as well as the detection of potential new compounds.[9,12] The extracts of *H. ingelheimensis* MSC5 showed some distinct metabolite peaks based on MALDI-TOF MS (Figure S2). Then the extracts were analyzed by ^{1}H NMR (Figure S3), and the presence of α-proton signals at δ_H 4.0−5.5 and amide proton signals at δ_H 6.0−10.0 indicated the main compounds were likely peptide natural products. Based on the genomic analysis of *H. ingelheimensis* MSC5 and the metabolomics analysis of the extracts, we hypothesized that *H. ingelheimensis* MSC5 may produce peptide NPs. Subsequently, LC-MS/MS-based GNPS molecular networking was performed on the EtOAc extracts of the fungus grown on rice medium to visualize all of the metabolites of this fungus. A cluster, in this molecular network, was composed of precursor ions in the range of m/z 500−700, including a node with m/z 596.285 $[M + H]^+$ (Figure 2), the MS/MS spectrum of which had a library match to the cyclic pentapeptide avellanin C (**13**).[25] Surrounding nodes with similar masses in this cluster indicated that *H. ingelheimensis* MSC5 might produce several new pentapeptide analogues.

Structure Elucidation of Avellanins D−O (1−12). The EtOAc extract of *H. ingelheimensis* MSC5 was subject to various chromatographic methods to isolate 12 new pentapeptides, avellanins D−O (**1**−**12**) (Figure 1). The structures and absolute configurations of each compound were elucidated by detailed spectroscopic methods, including 1D and 2D NMR, LC-MS/MS, X-ray single-crystal analysis, and Marfey's hydrolysis.

Compound **1** was obtained as a yellow oil. HRESIMS revealed a protonated molecule at m/z 550.2656 (calcd 550.2665), consistent with the molecular formula $C_{29}H_{35}N_5O_6$. The ^{1}H NMR data showed signals for three amide protons (δ_H 7.61, 8.81, and 10.14), one *N*-methyl group (δ_H 3.04), and four α protons (δ_H 4.68, 4.73, 4.88, and 5.64) (Table 1). The ^{13}C NMR data included five carbonyls (δ_C 168.6, 169.5, 170.0, 171.8, and 174.6), four α carbons (δ_C 48.4, 54.7, 56.3, and 57.9), one oxygenated sp^3 carbon (δ_C 60.4), and carbons associated with a monosubstituted aromatic ring (δ_C 126.9−137.2) and a disubstituted aromatic ring (δ_C 122.4−137.1) (Table 1). Further analysis of 2D NMR spectra (Figures S8−S10) led to the identification of five amino-acid components. Alanine (Ala) was confirmed by ^{1}H−^{1}H COSY correlations for

https://doi.org/10.1021/acs.jnatprod.4c00740
J. Nat. Prod. 2024, 87, 2695−2708

Table 1. ^{1}H and ^{13}C NMR Spectroscopic Data of 1−4

	1[a]		2[a]		3[a]		4[b]	
no.	δ_C, type	δ_H, mult, (J in Hz)	δ_C, type	δ_H, mult, (J in Hz)	δ_C, type	δ_H, mult, (J in Hz)	δ_C, type	δ_H, mult, (J in Hz)
	Ant		Ant		Ant		Ant	
1	169.5, C		169.9, C		169.6, C		167.8, C	
2	122.4, C		123.5, C		123.2, C		121.9, C	
3	127.7, CH	7.34, d, (7.8)	127.6, CH	7.32, m	127.5, CH	7.34, m	128.4, CH	7.62, d, (7.8)
4	123.1, CH	7.09, t, (7.6)	123.4, CH	7.10, t, (7.6)	123.2, CH	7.09, t, (7.6)	122.8, CH	7.17, t, (7.6)
5	132.2, CH	7.46, t, (6.9)	131.9, CH	7.46, t, (7.6)	132.0, CH	7.46, t, (7.8)	131.6, CH	7.50, t, (7.1)
6	121.7, CH	8.50, d, (8.3)	122.1, CH	8.42, d, (8.3)	121.7, CH	8.46, d, (8.3)	119.8, CH	8.43, d, (8.3)
7	137.1		136.7		136.8		136.9	
NH		10.14, s		9.86, s		9.85, s		10.19, s
	Pro		Pro		Pro		Pro	
1	174.6, C		174.1, C		173.8, C		173.2, C	
2	56.3, CH	4.68, m	56.5, CH	4.68, dd, (8.8, 1.4)	56.5, CH	4.67, d, (8.7)	56.3, CH	4.84, d, (8.5)
3	28.6, CH_2	1.40, m1.96, m	28.7, CH_2	1.42, m1.97, m	28.6, CH_2	1.41, m1.97, m	28.0, CH_2	1.23, m2.01, m
4	25.4, CH_2	1.88, m2.15, m	25.1, CH_2	1.86, m2.09, m	25.0, CH_2	1.85, m2.07, m	24.6, CH_2	1.82, m1.91, m
5	51.5, CH_2	3.60, m; 3.67, m	51.5, CH_2	3.56, m; 3.61, m	51.5, CH_2	3.53, m; 3.59, m	50.8, CH_2	3.37, m; 3.78, m
	*N*MePhe		*N*MePhe		*N*MePhe		*N*MePhe	
1	168.7, C		169.6, C		169.4, C		169.5, C	
2	57.9, CH	5.64, dd (12.5, 4.9)	57.7, CH	5.66, dd (12.4, 4.7)	57.6, CH	5.65, dd (12.3, 4.5)	57.3, CH	5.69, dd (12.3, 4.5)
3	33.4, CH_2	2.92, dd (15.7,12.4); 3.81, dd, (15.5, 4.8)	31.3, CH_2	2.94, dd (15.7, 12.4); 3.79, dd (15.7, 4.8)	33.0, CH_2	2.92, dd (15.3, 12.7); 3.78, dd (15.6, 4.4)	32.2, CH_2	3.00, dd (15.1, 12.3); 3.47, dd (14.8, 4.3)
4	137.2, C		137.3, C		137.3, C		138.0, C	
5,9	128.4, CH	7.19, m	128.3, CH	7.20, m	128.4, CH	7.23, m	128.5, CH	7.26, m
6,8	128.6, CH	7.28, m	128.7, CH	7.31, m	128.7, CH	7.31, m	128.1, CH	7.29, m
7	126.9, CH	7.22, m	126.9, CH	7.24, m	126.9, CH	7.25, m	126.3, CH	7.22, m
*N*Me	31.8, CH_3	3.04, s	31.3, CH_3	3.00, s	31.1, CH_3	2.96, s	31.0, CH_3	3.01, s
	Ala		Ser		Ala		Ser	
1	171.8, C		170.8, C		171.8, C		170.0, C	
2	48.4, CH	4.88, m	53.8, CH	4.71, m	48.5, CH	4.72, m	53.9, CH	4.40, m
3	18.3, CH_3	1.49, d, (8.3)	63.2, CH_2	3.75, dd (11.6, 4.8); 4.08, dd (11.6, 3.8)	18.1, CH_3	1.44, d, (7.1)	61.0, CH_2	3.55, m; 3.71, m
NH		7.61, d, (11.6)		7.49, d, (10.0)		7.25, m		7.34, d, (9.7)
OH				2.0, bs				4.69, t, (5.8)
	Hse		Phe		Tyr		Tyr	
1	170.0, C		169.2, C		169.6, C		169.5, C	
2	54.7, CH	4.73, m	54.8, CH	4.97, m	54.6, CH	4.90, m	55.0, CH	4.58, m
3	31.3, CH_2	2.19, m; 2.06, m	37.2, CH_2	3.05, dd (14.4, 9.2); 3.43, dd (14.4, 4.6)	36.2, CH_2	3.09, dd (14.4, 8.0); 3.19, dd (14.4, 4.8)	35.0, CH_2	2.83, dd (14.1, 10.6); 3.19, dd (14.3, 3.8)
4	60.4, CH_2	3.84, m; 3.76, m	136.5, CH		128.1, CH		127.8, CH	
5, 9			129.0, CH	7.20, m	130.4, CH	7.01, m	129.7, CH	6.99, d, (8.5)
6, 8			129.2, CH	7.31, m	116.1, CH	6.72, m	115.1, CH	6.64, d, (8.5)
7			127.2, CH	7.24, m	155.2, CH		155.8	
4-OH		4.20, bs						
NH		8.81, d, (7.2)		6.59, d, (8.9)		6.41, d, (9.0)		7.73, d, (8.7)
7-OH						5.65, overlapped		9.19, s

[a]In $CDCl_3$ [b]In DMSO-d_6. ^{1}H: 700 MHz and ^{13}C: 175 MHz for **2**, **3**; ^{1}H: 600 MHz and ^{13}C: 150 MHz for **1**, **4**.

NH (δ_H 7.61)/CH (δ_H 4.88)/CH_3 (δ_H 1.49) and an HMBC correlation from the methyl protons to a carbonyl carbon (δ_C 171.8) (Figure 1). An *N*Me-phenylalanine unit (*N*Me-Phe) was confirmed by the HMBC correlations of an *N*Me (δ_H 3.04) to an α carbon (δ_C 57.9), $^{1}H-^{1}H$ COSY correlations of a methine (δ_H 5.64) to a methylene (δ_H 2.92/3.81), and a phenyl ring assigned by sequential $^{1}H-^{1}H$ COSY correlations of five CH (δ_H 7.18−7.19, 7.22, 7.27−7.30) protons. Another aliphatic amino acid, proline (Pro), was deduced from sequential $^{1}H-^{1}H$ COSY (Figure 1) correlations for CH (δ_H 4.68)/CH_2 (δ_H 1.96/1.40)/CH_2 (δ_H 2.15/1.88)/CH_2 (δ_H 3.67/3.60) along with long-range correlations from the α proton (δ_H 4.68) to the methylene carbon (δ_C 51.5) and a carbonyl carbon (δ_C 174.6). In addition, a spin system consisting of four adjacent aromatic protons (δ_H 7.34, 7.09, 7.46, and 8.50) recognized in the $^{1}H-^{1}H$ COSY correlations (Figure 1) was considered to be a 1,2-disubstituted benzene ring. Correlations were also detected from the proton at δ_H 7.34 to a carbonyl carbon (δ_C 169.5) and from a NH proton (δ_H 10.14) to the aromatic carbons (δ_C 121.7), indicating the presence of the anthranilic acid unit (Ant). The last unit was assigned as homoserine (Hse) by $^{1}H-^{1}H$ COSY (Figure 1) correlations for CH (δ_H 4.73)/CH_2 (δ_H 2.06/2.19) /CH_2 (δ_H 3.76/3.84). The connectivity among the 5 amino acid residues was determined by HMBC (Figure 1) correlations from Ala-NH to *N*Me-Phe-C1, *N*Me-Phe-CH_3 to Pro-C1, Pro-H5 to Ant-C1, Ant-NH to Hse-C1, and Hse-NH to Ala-C1 (Figure 1) and supported by the MS/MS data (Figure 3). Subsequent

https://doi.org/10.1021/acs.jnatprod.4c00740
J. Nat. Prod. 2024, 87, 2695−2708

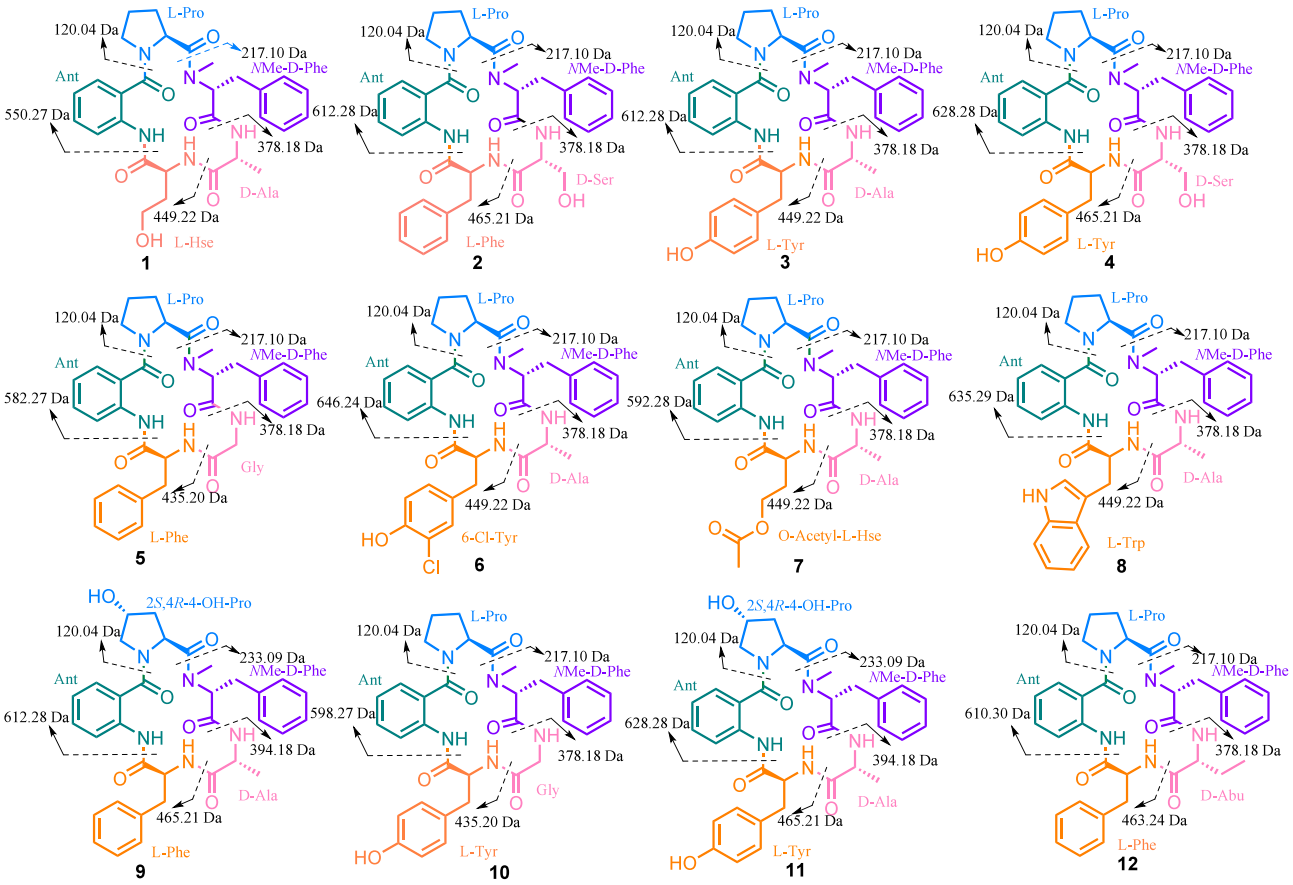


Figure 3. MS^2 fragmentation of compounds **1**−**12**.

Marfey's analysis confirmed the absolute configurations of each amino acid (Figure S101). Thus, compound **1** was determined as *cyclo* (Ant[1]-L-Pro[2]-*N*Me-D-Phe[3]-D-Ala[4]-L-Hse[5]) and named avellanin D.

Compound **2** was obtained as a colorless crystal. HRESIMS revealed a protonated molecule at *m*/*z* 612.2811 (calcd 612.2822), consistent with the molecular formula $C_{34}H_{37}N_5O_6$, with an index of hydrogen deficiency (IHD) of 19. The ^{1}H NMR data displayed three amide NH protons (δ_H 6.59, 7.49, 9.86), one *N*-methyl group (δ_H 3.04), and four α protons (δ_H 4.68, 4.71, 4.97, 5.66) (Table 1). The ^{13}C NMR data displayed 34 carbon signals, five deshielded signals for the carbonyl carbons at δ_C 169.2, 169.6, 169.9, 170.8, and 174.1, and four α carbons at δ_C 53.8, 54.8, 56.5, and 57.7 (Table 1). These indicated that **2** should be a pentapeptide and an analogue of **1**. A careful comparison of the NMR spectra between **2** (Figures S13−S18) and **1** (Figures S5−S10) showed one oxygenated sp^3 carbon (δ_C 63.2) in place of the methyl group (δ_C 18.3), indicating that the fourth amino acid residue should be serine (Ser). This was confirmed by MS/MS data analysis (Figure 3). Furthermore, a monosubstituted phenyl ring (δ_C 127.2−136.5) replaced an oxygenated carbon (δ_C 60.4), suggesting that the fifth amino acid residue should be a phenylalanine (Phe). The NMR signals of the other three amino acid residues of compound **2** were consistent with those of compound **1**. Thus, the planar structure of **2** was defined as *cyclo*-(Ant[1]-Pro[2]-*N*Me-Phe[3]-Ser[4]-Phe[5]) and supported by the HMBC and MS/MS data (Figures 1, 3). The structure, including the absolute configuration, was further confirmed by X-ray diffraction with Cu Kα radiation (Figure 4). Therefore, compound **2** was determined unambiguously as *cyclo*-(Ant[1]-L-Pro[2]-*N*Me-D-Phe[3]-D-Ser[4]-L-Phe[5]) and named avellanin E.

Compound **3** was obtained as a white crystal. HRESIMS revealed a protonated molecule at *m*/*z* 612.2809 (calcd 612.2822), consistent with the molecular formula $C_{34}H_{37}N_5O_6$, making **3** an isomer of **2** with different retention times and different MS/MS fragmentation. Comparison of the 1D NMR data (Table 1) of compounds **3** and **2** revealed that the hydroxymethine group (δ_C 63.2) of Ser[4] in **2** was substituted by a methyl group (δ_C 18.2/δ_H 1.44) in **3**, suggesting the presence of alanine (Ala). Additionally, Phe[5] in **2** was substituted by a tyrosine (Tyr) in **3** indicated by the deshielded carbon (δ_C = 155.2). The Ala[4] and Tyr[5] residues in **3** were further established by ^{1}H−^{1}H COSY and HMBC correlations (Figure 1). The planar structure of **3** was assigned as *cyclo* (Ant[1]-Pro[2]-*N*Me-Phe[3]-Ala[4]-Tyr[5]). The absolute configuration was further confirmed by X-ray diffraction with Cu Kα radiation (Figure 4). Thus, compound **3** was determined as *cyclo*-(Ant[1]-L-Pro[2]-*N*Me-D-Phe[3]-D-Ala[4]-L-Tyr[5]) and named avellanin F.

Compound **4** was obtained as a white powder. HRESIMS revealed a protonated molecule at *m*/*z* 628.2764 (calcd 628.2771), consistent with the molecular formula $C_{34}H_{37}N_5O_7$. Compounds **4** and **3** have very similar structures, as revealed by a comparison of 1D NMR data (Table 1). The main difference is that the methyl carbon (δ_C 18.2) in **3** was substituted by a hydroxymethine carbon (δ_C 61.0) in **4**, corresponding to an additional 16 amu (O) compared to **3** and indicating the presence of Ser[4] in **4**. Further detailed comparisons of the 2D NMR spectra of **4** and **3** (Figures

https://doi.org/10.1021/acs.jnatprod.4c00740
J. Nat. Prod. 2024, 87, 2695−2708

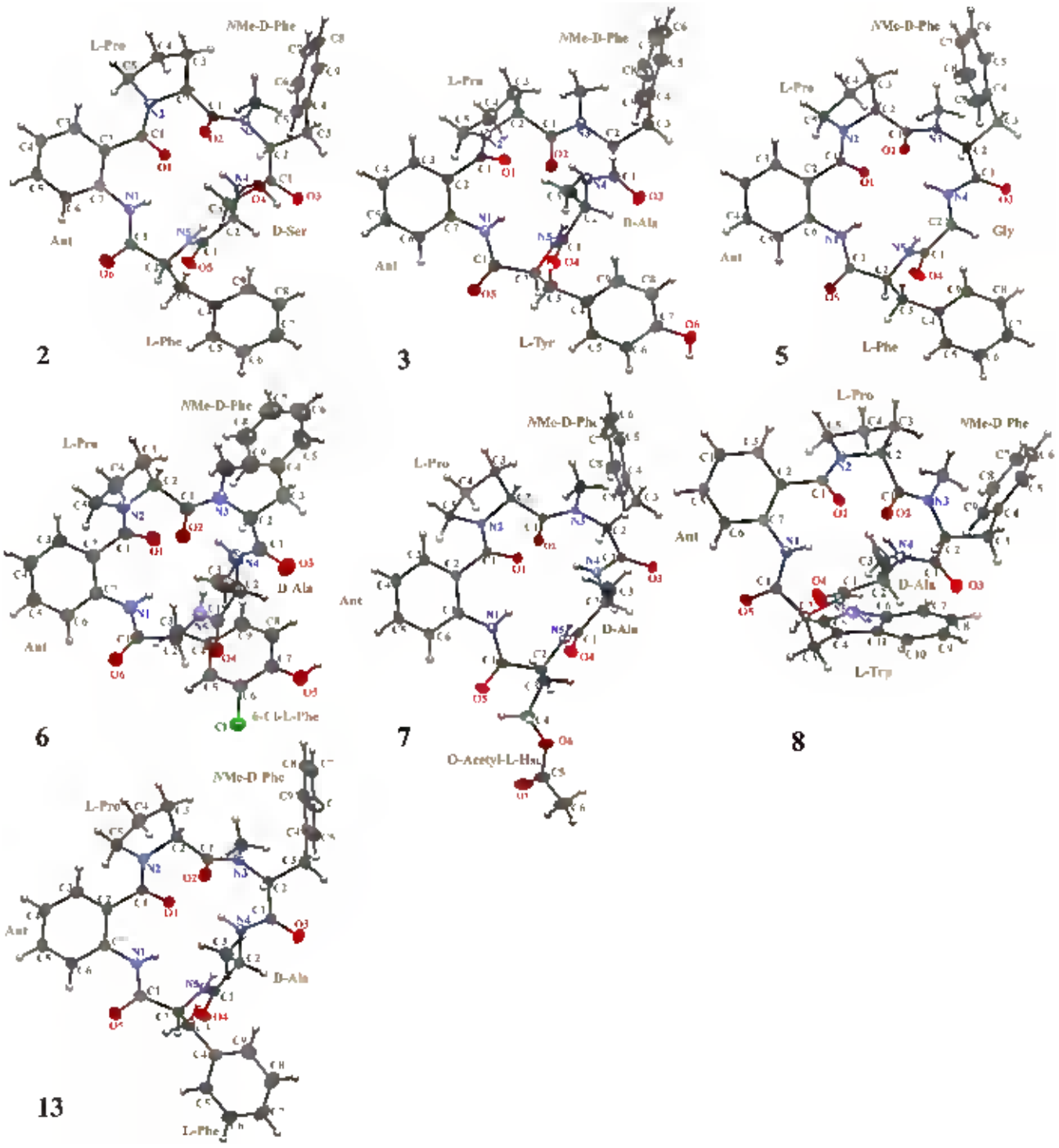


Figure 4. X-ray crystal structures of compounds **2**, **3**, **5**, **6**, **7**, **8**, and **13** illustrating their absolute configurations.

S32−S34, Figures S24−S26) revealed the planar structure of **4** as *cyclo*-(Ant[1]-Pro[2]-*N*Me-Phe[3]-Ser[4]-Tyr[5]), supported by the HMBC and MS/MS data (Figures 1, 3). Subsequent Marfey's analysis established the absolute configurations of each amino acid (Figure S102). Thus, compound **4** was determined as *cyclo*-(Ant[1]-L-Pro[2]-*N*Me-D-Phe[3]-D-Ser[4]-L-Tyr[5]) and named avellanin G.

Compound **5** was obtained as colorless needle crystals. HRESIMS revealed a protonated molecule at *m*/*z* 582.2704 (calcd 582.2716), consistent with the molecular formula $C_{33}H_{35}N_5O_5$. Compounds **5** and **2** exhibited very similar ^{13}C and DEPT-135 NMR spectroscopic data (Figures S14, S15, Figures S38, S39), except that the methine α carbon (δ_C 53.8) in **2** was substituted by a methylene carbon (δ_C 42.6) in **5**, and one oxygenated sp^3 carbon (δ_C 63.2) in **2** disappeared in **5**. These findings indicated the presence of Gly in compound **5**. The peptide sequence was defined as *cyclo*-(Ant[1]-Pro[2]-*N*Me-Phe[3]-Gly[4]-Phe[5]) by HMBC correlations and MS/MS data (Figures 1, 3). The absolute configuration was further confirmed by X-ray diffraction with Cu Kα radiation (Figure

https://doi.org/10.1021/acs.jnatprod.4c00740
J. Nat. Prod. 2024, 87, 2695−2708

Table 2. ^{1}H and ^{13}C NMR Spectroscopic Data of 5−8

	5[b]		6[a]		7[a]		8[b]	
no.	δ_C, type	δ_H, mult, (J in Hz)	δ_C, type	δ_H, mult, (J in Hz)	δ_C, type	δ_H, mult, (J in Hz)	δ_C, type	δ_H, mult, (J in Hz)
	Ant		Ant		Ant		Ant	
1	167.7, C		169.6, C		169.5, C		167.9, C	
2	122.3, C		123.4, C		123.3, C		121.4, C	
3	128.4, CH	7.63, d, (7.7)	127.3, CH	7.29, m	127.5, CH	7.34, d, (7.8)	128.6, CH	7.62, d, (7.8)
4	123.1, CH	7.18, t, (7.6)	123.8, CH	7.10, t, (7.5)	123.3, CH	7.11, t, (7.6)	122.9, CH	7.17, t, (7.6)
5	131.6, CH	7.50, t, (7.0)	131.8, CH	7.46, t, (7.4)	132.2, CH	7.46, t, (6.9)	131.8, CH	7.51, overlapped
6	120.0, CH	8.42, d, (8.3)	121.9, CH	8.40, d, (8.3)	121.9, CH	8.40, d, (8.5)	119.4, CH	8.56, d, (8.4)
7	136.8, C		136.3, C		137.3, C		137.0, C	
NH		10.15, s		9.69, s		9.93, s		10.11, s
	Pro		Pro		Pro		Pro	
1	173.1, C		173.9, C		174.2, C		172.9, C	
2	56.5, CH	4.84, dd, (8.5, 2.4)	56.5, CH	4.65, dd, (9.0, 2.4)	56.4, CH	4.71, dd, (8.5, 1.1)	56.1, CH	4.81, dd, (8.8, 2.1)
3	28.1, CH_2	1.31, m; 2.05, m	28.7, CH_2	1.41, m; 1.96, m	28.8, CH_2	1.45, m; 1.99, m	28.0, CH_2	1.23, m; 2.01, m
4	24.7, CH_2	1.83, m; 1.93, m	24.9, CH_2	1.85, m; 2.07, m	25.2, CH_2	1.87, m; 2.13, m	24.7, CH_2	1.82, m; 1.90, m
5	50.7, CH_2	3.40, m; 3.76, m	51.4	3.49, m; 3.57, m	51.6	3.61, m; 3.63, m	51.0	3.78, m; 3.36, m
	*N*MePhe		*N*MePhe		*N*MePhe		*N*MePhe	
1	169.3, C		169.3, C		169.5, C		168.9, C	
2	57.5, CH	5.58, dd (11.9, 4.5)	57.4, CH	5.56, dd (12.6, 4.8)	57.7, CH	5.80, dd (12.2, 4.8)	56.8, CH	5.60, dd (12.3, 4.7)
3	32.3, CH_2	3.01, dd (15.0, 12.1); 3.42, m	33.1, CH_2	2.95, m; 3.83, dd (15.9, 4.9)	33.2, CH_2	2.96, m; 3.80, dd (15.5, 4.8)	32.0, CH_2	2.99, dd (15.3, 12.0); 3.41, overlapped
4	137.9, C		137.3, C		136.7, C		137.9, C	
5, 9	128.5, CH	7.26, m	128.3, CH	7.22, m	128.3, CH	7.19, m	128.1, CH	7.29, m
6, 8	128.3, CH	7.29, m	128.7, CH	7.32, m	128.7, CH	7.29, m	128.5, CH	7.22, m
7	126.4, CH	7.23, m	126.9, CH	7.24, m	126.9, CH	7.23, m	126.3, CH	7.21, m
*N*Me	31.1, CH_3	2.98, s	31.0, CH_3	2.93, s	31.3, CH_3	3.02, s	30.8, CH_3	2.94, s
	Gly		Ala		Ala		Ala	
1	169.3, C		171.9, C		172.2, C		172.0, C	
2	42.6, CH_2	3.50, dd (16.6, 2.5); 4.18, dd (16.6, 8.8)	48.5, CH	4.74, m	48.6, CH	4.89, m	47.3, CH	4.47, m
3			18.2, CH_3	1.45, d, (7.2)	18.3, CH_3	1.52, d, (7.2)	17.7, CH_3	1.21, d, (7.0)
NH		7.53, m		7.15, m		7.42, m		7.84, d, (7.6)
	Phe		6-Cl-Phe		O-Acetyl-Hse		Trp	
1	169.3, C		169.3, C		169.7, C		169.8, C	
2	54.7, CH	4.71, m	54.2, CH	4.93, m	51.8, CH	4.86, m	54.9, CH	4.65, m
3	35.9, CH_2	2.94, dd (14.2, 10.7); 3.35, m	35.9, CH_2	3.08, dd (14.3, 5.2); 3.19, dd (14.3, 7.1)	30.6, CH_2	2.10, m; 2.41, m	26.1, CH_2	3.13, dd (12.6, 5.1); 3.41, m
4	137.8, C		129.4, C		61.5, CH_2	4.15, m; 4.23, m	110.1, C	
5	128.8, CH	7.23, m	129.4, CH	6.97, d, (8.3)	171.1, C		123.0, CH	7.09, d, (2.5)
6	128.2, CH	7.28, m	120.6, C		21.1, CH_3	2.06, s	136.1, C	
7	126.4	7.19, m	150.9, C				111.4, CH	7.33, d, (8.0)
8	128.2, CH	7.28, m	116.7, CH	6.92, d, (8.3)			121.0, CH	7.06, t, (7.0)
9	128.8, CH	7.23, m	129.8, CH	7.17, overlapped			118.4, CH	6.96, t, (6.9)
10							117.9, CH	7.52, d, (7.5)
11							127.1, C	
NH		7.95, d, (8.5)		6.50, d, (9.2)		6.85, d, (9.0)		10.83, d, (1.5)
7-OH				5.65, bs				

[a]In $CDCl_3$ [b]In DMSO-d_6. ^{1}H: 700 MHz and ^{13}C: 175 MHz for **6**, **7**, and **8**; ^{1}H: 600 MHz and ^{13}C: 150 MHz for **5**.

4). Thus, compound **5** was determined as *cyclo*-(Ant[1]-L-Pro[2]-*N*Me-D-Phe[3]-Gly[4]-L-Phe[5]) and named avellanin H.

Compound **6** was obtained as a colorless crystal. HRESIMS revealed a protonated molecule at m/z 646.2423 (calcd 646.2432), consistent with the molecular formula $C_{34}H_{36}N_5O_6Cl$. Compounds **6** and **3** have similar 1D NMR spectra (Figures S21−S23, Figures S45−S47). The position of chlorination was determined at Tyr-C6 through HMBC, $^{1}H-^{1}H$ COSY, and HSQC correlations to a non-hydrogenated, deshielded carbon at δ_C 120.6 (Figure 1, Figures S48−S50). The planar structure was defined as *cyclo*-(Ant[1]-Pro[2]-*N*Me-Phe[3]-Ala[4]-6-Cl-Tyr[5]). The absolute configuration was further confirmed by X-ray diffraction with Cu Kα radiation (Figure 4). Thus, the structure of compound **6** was determined as *cyclo*-(Ant[1]-L-Pro[2]-*N*Me-D-Phe[3]-D-Ala[4]-6-Cl-L-Tyr[5]) and named avellanin I.

Compound **7** was obtained as a colorless crystal. HRESIMS revealed a protonated molecule at m/z 592.2765 (calcd

https://doi.org/10.1021/acs.jnatprod.4c00740
J. Nat. Prod. 2024, 87, 2695−2708

Table 3. ^{1}H and ^{13}C NMR Spectroscopic Data of 9−12

	9[a]		10[a]		11[a]		12[a]	
no.	δ_C, type	δ_H, mult, (J in Hz)	δ_C, type	δ_H, mult, (J in Hz)	δ_C, type	δ_H, mult, (J in Hz)	δ_C, type	δ_H, mult, (J in Hz)
		Ant		Ant		Ant		Ant
1	169.7, C		169.6, C		169.8, C		169.6, C	
2	123.4, C		123.4, C		122.9, C		123.3, C	
3	127.5, CH	7.32, m	127.4, CH	7.33, m	127.7, CH	7.32, m	127.3, CH	7.32, m
4	123.6, CH	7.09, t, (7.6)	123.3, CH	7.10, t, (7.6)	123.5, CH	7.11, t, (7.4)	123.3, CH	7.10, m
5	131.9, CH	7.45, t, (7.6)	131.9, CH	7.46, t, (8.5)	132.1, CH	7.47, t, (6.9)	131.8, CH	7.46, m
6	122.0, CH	8.37, d (8.5)	121.9, CH	8.42, d (8.3)	121.8, CH	8.43, d (8.4)	122.0, CH	8.40, d (8.3)
7	136.6, C		136.7, C		136.7, C		136.5, C	
NH		9.64, s		9.79, s		9.66, s		9.87, s
	4-OH-Pro		Pro		4-OH-Pro		Pro	
1	174.2, C		173.9, C		174.0, C		173.9, C	
2	55.4, CH	4.80, m	56.6, CH	4.66, m	55.4, CH	4.78, dd, (8.7, 4.6)	56.5, CH	4.68, d, (9.3)
3	36.7, CH_2	1.71, m; 1.94, m	28.6, CH_2	1.43, m; 1.97, m	36.6, CH_2	1.70, m; 1.94, m	28.6, CH_2	1.40, m; 1.97, m
4	69.8, CH	4.54, m	25.0, CH_2	1.86, m; 2.06, m	69.7, CH	4.56, m	25.0, CH_2	1.85, m; 2.11, m
4-OH		2.20, bs				not detected		
5	57.7, CH_2	3.59, m; 3.61, m	51.4, CH_2	3.53, m; 3.59, m	57.6, CH_2	3.61, m; 3.62, m	51.4, CH_2	3.58, m; 3.60, m
	*N*MePhe		*N*MePhe		*N*MePhe		*N*MePhe	
1	169.2, C		169.9, C		169.3, C		169.6, C	
2	58.1, CH	5.59, dd (12.5, 4.2)	57.8, CH	5.61, dd (12.5, 4.6)	58.0, CH	5.57, dd (12.3, 4.7)	57.6, CH	5.65, dd (12.5, 4.9)
3	33.2, CH_2	2.92, dd (15.7, 12.5); 3.81, dd (15.7, 4.7)	33.2, CH_2	2.93, dd (15.9, 12.3); 3.81, dd (15.6, 4.7)	33.2, CH_2	2.92, dd (15.8, 12.4); 3.80, dd (15.2, 4.0)	33.2, CH_2	2.95, m; 3.84, dd (15.9, 4.1)
4	137.3, C		137.3, C		137.2, C		137.4, C	
5, 9	128.3, CH	7.20, m	128.3, CH	7.22, m	128.4, CH	7.23, m	128.3, CH	7.19, m
6, 8	128.9, CH	7.30, m	128.7, CH	7.31, m	128.7, CH	7.32, m	128.9, CH	7.31, m
7	127.0, CH	7.24, m	126.9, CH	7.24, m	127.0, CH	7.25, m	127.2, CH	7.22, m
*N*Me	31.4, CH_3	3.00, s	31.3, CH_3	2.97, s	31.4, CH_3	3.00, s	31.2, CH_3	3.04, s
	Ala		Gly		Ala		2-Abu	
1	172.0, C		169.1, C		171.8, C		169.3, C	
2	48.4, CH	4.68, m	42.8, CH	3.55, m4.68, m	48.5, CH	4.73, m	53.9, CH	4.55, m
3	18.1, CH_3	1.43, d, (7.2)			18.1, CH_3	1.44, d, (7.4)	24.9, CH_2	1.70, m2.07, m
4							10.5, CH_3	0.93, t, (7.4)
NH		7.27, m		7.34, dd (10.2, 2.5)		7.28, m		7.41, d, (1.6)
	Phe		Tyr		Tyr		Phe	
1	169.6, C		169.4, C		169.6, C		169.7, C	
2	54.8, CH	4.92, m	54.3, CH	4.96, m	54.6, CH	4.87, m	54.5, CH	4.97, m
3	37.1, CH_2	3.03, dd (14.2, 9.1); 3.38, dd (14.4, 4.7)	36.2, CH_2	3.09, dd (14.2, 7.8); 3.19, dd (14.4, 4.8)	36.1, CH_2	3.11, dd (14.3, 7.5); 3.18, dd (14.4, 4.8)	37.2, CH_2	3.05, m; 3.38, dd (14.3, 6.4)
4	136.4, C		128.2, C		128.2, C		136.5, C	
5, 9	129.2, CH	7.20, m	130.5, CH	7.02, m	130.4, CH	7.00, m	129.3, CH	7.21, m
6, 8	128.8, CH	7.31, m	116.1, CH	6.73, m	116.3, CH	6.72, m	128.2, CH	7.29, m
7	127.0, CH	7.25, m	155.2		155.1		126.9, CH	7.21, m
NH		6.54, d, (8.9)		6.45, d, (9.3)		6.45, d, (9.0)		6.48, d, (8.5)
7-OH				5.65, overlapped		7.10, overlapped		

[a]In $CDCl_3$. ^{1}H: 700 MHz and ^{13}C: 175 MHz for **9**, **10**, **11**, and **12**.

592.2771), consistent with the molecular formula $C_{31}H_{37}N_5O_7$. Compound **7** displayed 1D NMR spectra very similar to those of **1** (Figures S5−S7, S53−S55). An additional methyl carbon (δ_C = 14.3) and a deshielded carbonyl (δ_C = 171.1) indicated the presence of an acetyl group. Subsequent HMBC and MS/MS data revealed the planar structure of **7** as *cyclo*-(Ant[1]-Pro[2]-*N*Me-Phe[3]-Ala[4]-O-Acetyl-Hse[5]) (Figures 1, 3). The absolute configuration was further confirmed by X-ray diffraction with Cu Kα radiation (Figure 4). Thus, compound **7** was determined as *cyclo*-(Ant[1]-L-Pro[2]-*N*Me-D-Phe[3]-D-Ala[4]-O-Acetyl-L-Hse[5]) and named avellanin J.

Compound **8** was obtained as a colorless crystal. HRESIMS revealed a protonated molecule at m/z 635.2968 (calcd 635.2981), consistent with the molecular formula $C_{36}H_{38}N_6O_5$. Comparison of the 1D NMR spectra of **8** (Figures S61−S63) with those of compound **1** suggested the replacement of an Hse residue in **1** by a Trp residue in **8**. The sequential $^{1}H-^{1}H$ COSY correlations of H7, H8, H9, and H10 and the HMBC correlations of H7 to C11and H8 to C6 indicated a disubstituted aromatic ring (Figure 1). Furthermore, the HMBC correlations of H3 to C5, H5 to C6/C11, and the −NH to C2 together confirmed the presence of tryptophan, consistent with the MS/MS fragmentation data (Figures 1, 3). Thus, compound **8** was assigned as *cyclo*-(Ant[1]-Pro[2]-*N*Me-Phe[3]-Ala[4]-Trp[5]). The absolute configuration was further confirmed by X-ray diffraction with Cu Kα radiation (Figure

https://doi.org/10.1021/acs.jnatprod.4c00740
J. Nat. Prod. 2024, 87, 2695−2708

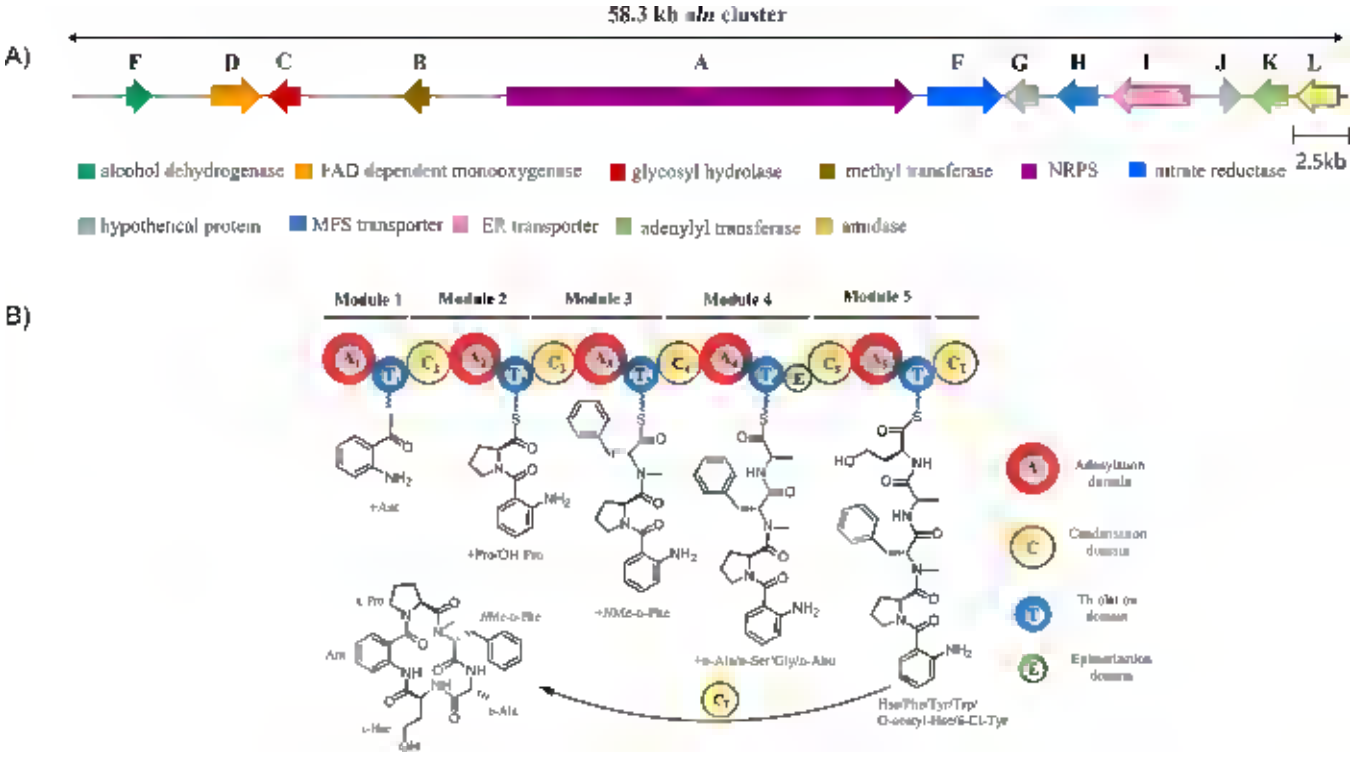


Figure 5. Modular architecture of the avellanin assembly line and the proposed biosynthesis pathway. (A) Organization of the putative avellanin biosynthetic gene cluster (*aln* BGC) in *H. ingelheimensis* MSC5. (B) Proposed biosynthetic pathway for compounds **1**−**13**.

4). Thus, compound **8** was determined as *cyclo*-(Ant[1]-L-Pro[2]-*N*Me-D-Phe[3]-D-Ala[4]-L-Trp[5]) and named avellanin K.

Compound **9** was obtained as a white powder. HRESIMS revealed a protonated molecule at *m*/*z* 612.2820 (calcd 612.2822), consistent with the molecular formula $C_{34}H_{37}N_5O_6$. Compound **9** has a very similar structure to the known compound avellanin C (**13**),[25] as revealed by a comparison of their 1D NMR spectra (Figures S69−S71). The slight difference was that Pro-C4 (δ_C 14.3) in **13** was substituted by an oxygenated carbon (δ_C 69.8) in **9**, indicating the presence of a hydroxy group on Pro in **9**, which was consistent with the 16 mass unit difference. HMBC and MS/MS data revealed the planar structure as *cyclo*-(Ant[1]-4-OH-Pro[2]-*N*Me-Phe[3]-Ala[4]-Phe[5]) (Figures 1, 3). Subsequent Marfey's analysis confirmed the absolute configurations of each amino acid (Figures S103−S104). Thus, compound **9** was determined as *cyclo*-(Ant[1]-2*S*, 4*R*-4-OH-L-Pro[2]-*N*Me-D-Phe[3]-D-Ala[4]-L-Phe[5]) and named avellanin L.

Compound **10** was obtained as a white powder. HRESIMS revealed a protonated molecule at *m*/*z* 598.2658 (calcd 598.2665), consistent with the molecular formula $C_{33}H_{35}N_5O_6$. Compound **10** has a very similar structure to **3** as revealed by the comparison of their 1D ^{13}C NMR spectra (Figures S22, S79). A methyl carbon (δ_C 18.1) disappeared, and a methine α carbon (δ_C 48.5) in **3** was substituted by a methylene carbon (δ_C 42.8) in **10**, indicating that the Ala[4] in **3** was replaced by a glycine in **10**, corresponding to a difference of 14 amu. HMBC and MS/MS data revealed the planar structure of **10** as *cyclo*-(Ant[1]-Pro[2]-*N*Me-Phe[3]-Gly[4]-Tyr[5]) (Figures 1, 3). Subsequent Marfey's analysis confirmed the absolute configurations of each amino acid (Figure S105). Thus, compound **10** was determined as *cyclo*-(Ant[1]-L-Pro[2]-*N*Me-D-Phe[3]-Gly[4]-L-Tyr[5]) and named avellanin M.

Compound **11** was obtained as a white powder. HRESIMS revealed a protonated molecule at *m*/*z* 628.2764 (calcd 628.2771), consistent with the molecular formula $C_{34}H_{37}N_5O_7$. Compound **11** has very similar 1D NMR spectra to those of **9** (Figures S69−S71, S86−S88), except that Phe[5] in **9** was replaced by Tyr[5] in **11**, corresponding to the 16 amu difference. HMBC and MS/MS data revealed the planar structure of **11** as *cyclo*-(Ant[1]-4-OH-Pro[2]-*N*Me-Phe[3]-Ala[4]-Tyr[5]) (Figures 1, 3). Subsequently Marfey's analysis confirmed the absolute configurations of each amino acid (Figures S104 and S106). Thus, compound **11** was determined as *cyclo*-(Ant[1]-2*S*, 4*R*-4-OH-Pro[2]-*N*Me-D-Phe[3]-D-Ala[4]-L-Tyr[5]) and named avellanin N.

Compound **12** was obtained as a white powder. HRESIMS revealed a protonated molecule at *m*/*z* 610.3023 (calcd 610.3029), consistent with the molecular formula $C_{35}H_{39}N_5O_5$. Compound **12** has a similar structure to **5**, as revealed through the comparison of their 1D NMR spectra (Figures S37−S39, S94−S96). A methyl (δ_C 10.5/δ_H 0.93) group, a methylene carbon (δ_C 24.9/δ_H 2.07, 1.69), and the ^{1}H-^{1}H COSY (Figure 1) correlations collectively demonstrated that the Gly[4] in **5** was replaced by a 2-aminobutyric acid[4] (Abu) in **12**. HMBC and MS/MS data revealed the planar structure of **12** as *cyclo*-(Ant[1]-Pro[2]-*N*Me-Phe[3]-Abu[4]-Phe[5]) (Figures 1, 3). Subsequent Marfey's analysis confirmed the absolute configurations of each amino acid (Figure S107). Thus, compound **12** was determined as *cyclo*-(Ant[1]-L-Pro[2]-*N*Me-D-Phe[3]-D-Abu[4]-L-Phe[5]) and named avellanin O.

Proline Conformational Analysis of Avellanins D−O (1−12). Since the *cis*/*trans* conformation of the proline residues can have a large impact on the overall peptide conformation, we sought to determine the conformation of each proline in **1**−**12**. First, for the compounds with XRD data (**2**, **5**, **6**, **7**, **8**, and **13**) (Figure 4), the crystal structures clearly

https://doi.org/10.1021/acs.jnatprod.4c00740
J. Nat. Prod. 2024, 87, 2695−2708

show a torsion angle of approximately 180° ($\omega \approx 180°$) between Ant[1]-Pro[2], confirming that the peptide bond formed by proline in these compounds adopts a *trans* conformation. Second, this conformation of Ant[1]-Pro[2] was also confirmed by the small chemical shift difference (3.2−3.8 ppm, Tables 1−3) between the C3 and C4 carbons of Pro[2] observed in the ^{13}C NMR spectrum of **1**−**8**, **10**, and **12** ($\Delta\delta_{C3\text{-}C4}$: *cis* > 8 ppm, *trans* < 6 ppm), consistent with the presence of a *trans* proline.[36,37] However, this rule does not apply to compounds **9** and **11**, as their proline residues feature a hydroxy group at C4. For the conformation of these, we turned to the NOESY spectrum of compound **11** (Figure S75). The detectable NOE correlations of H3-Ant[1]/H5-4-OH-Pro[2] rather than H3-Ant[1]/H2-4-OH-Pro[2] suggest that the amide bond of Ant[1]-4-OH-Pro[2] adopts a *trans* conformation (Figure S75), and considering the biosynthetic origin, we propose that the proline conformations of **9** and **11** are the same.

Proposed Biosynthetic Pathway. antiSMASH revealed 10 potential NRPS BGCs of 27 total predicted BGCs in the genome of this strain (Table S1). Among them, the *aln* BGC contains a typical NRPS gene *aln*A, which encodes an expected pentamodular NRPS containing 16 domains (A_1-T_1, C_2-A_2-T_2, C_3-A_3-T_3, C_4-A_4-T_4-E, C_5-A_5-T_5, and C_T) (Figure 5, Table S2). The products encoded by *aln* BGC are highly consistent with the backbone of avellanins. However, this expected NRPS BGC lacked an integral *N*-methyltransferase (*N*-MeT) domain, which is typically found in BGCs encoding *N*-methyl-cyclopeptides, such as that for the isaridins.[36] Bioinformatics analysis instead revealed an adjacent methyltransferase *alnB*, which may perform this activity *in trans*, as in the biosynthesis of the cycloaspeptides.[38] Another intriguing aspect is the third module, which is involved in the formation of the *N*Me-D-Phe residue. However, no E or C/E domains are present in the third module. Therefore, we suggest that the A domain of this third NRPS module directly accepts a D-amino acid. Based on the architecture of this NRPS and the predicted substrate specificity of the adenylation (A) domain, a biosynthetic pathway is proposed for compounds **1**−**13** (Figure 5). Additionally, avellanin D (**1**) and avellanin J (**7**) contain a peptide backbone featuring rare homoserine and *O*-acetyl-homoserine residues. To the best of our knowledge, only 5 NPs containing a homoserine residue have been reported,[39−42] and we could not find any other *O*-acetyl-homoserine-containing cyclopeptides in the literature. This may be because both homoserine and *O*-acetyl-homoserine are the intermediate products in microbial metabolism,[43] and thus are rarely used as common amino acid building blocks for microbial secondary metabolites.

Antiproliferative Activity and Antimalarial Activity. Compounds **1**−**6** were evaluated for cytotoxicity, but no activity was observed. Compounds **1**, **2**, **3**, **5**, **6**, and **13** were also tested for antimalarial activity, and most of them exhibited activity against the human protozoal parasite *Plasmodium falciparum* 3D7, with IC_{50} values ranging from 0.19 to 16 μM. Notably, avellanin I (**6**) displayed potent inhibitory activity with an IC_{50} value of 0.19 ± 0.09 μM (Table 4). Previous reports showed that avellanins have a pressor effect and inhibitory effect on quorum sensing.[24,25] This study is the first to demonstrate that avellanins exhibit antimalarial activity. Moreover, avellanin I (**6**) is the most potent antimalarial cyclic pentapeptide from fungi reported to data.[44,45]

In summary, 12 new cyclopeptides, avellanins D−O (**1**−**12**), were isolated from the Mariana Trench anemone derived fungus *H. ingelheimensis* MSC5, using a combination of genomic BGC analysis, OSMAC, and metabolomics strategies. Notably, avellanin D (**1**) and avellanin J (**7**) contain peptide backbones with rare homoserine and *O*-acetyl-homoserine residues. Bioassay results indicated that the halogenated cyclopeptide avellanin I (**6**) exhibited potent antimalarial activity with an IC_{50} value of 0.19 ± 0.09 μM. These results not only report 12 new cyclopeptides with antimalarial activity but also highlight an efficient metabolomics platform for the discovery of novel NPs from deep-sea microbes.

Table 4. Antimalarial Activity of Avellanins

Compound	Antimalarial activity (IC_{50}, μM) against *P. falciparum* 3D7
avellanin D (**1**)	>50
avellanin E (**2**)	16 ± 3
avellanin F (**3**)	10 ± 1
avellanin H (**5**)	3 ± 2
avellanin I (**6**)	0.19 ± 0.09
avellanin C (**13**)	9.4 ± 0.7
artemisinin[a]	0.010 ± 0.002
chloroquine[a]	0.032 ± 0.005

[a]Artemisinin and chloroquine were used as positive controls.

EXPERIMENTAL SECTION

General Experimental Procedures. The melting points were measured by using a Buchi M-560 melting point instrument. Optical rotations were measured in MeOH by using a JASCO P-2000 polarimeter at 300 K. UV spectra were obtained with a ThermoFisher Evolution 300 UV−visible spectrophotometer. Micro-IR spectra were acquired on a Nicolet iN10 Fourier transform infrared spectrometer with KBr disks. 1D and 2D NMR data were collected on Bruker AVANCE III HD 600/700 MHz NMR spectrometers. HRESIMS experiments were conducted on a Thermo Vanquish UHPLC/Q Exactive Plus LC-MS/MS system with a 2.1 mm × 100 mm, 1.7 μm ACQUITY UPLC BEH C18 column. HPLC analysis was performed on an Agilent 1260 system, coupled to a UV detector G1314B, and equipped with a 4.6 × 150 mm, 5 μm Agilent C18 column. Semipreparative HPLC was carried out using a Waters 1525 pump equipped with a Waters 2489 detector and a 9.4 × 250 mm, 5 μm Eclipse XDB-C18 column. Column chromatography (CC) was performed using silica gel (200−300 mesh, Marine Chemical Factory of Qingdao). Fractions were monitored by thin-layer chromatography and a Bruker ultrafleXtreme MALDI TOF/TOF mass spectrometer. All solvents and reagents were purchased from Tianjin Kemiou Chemical Reagent Co. Ltd., Bide Pharmatech Co. Ltd., Shanghai yuanye Bio-Technology Co. Ltd., J & K Scientific, and Aladdin.

Fungal Strain and Genome Sequencing. *H. ingelheimensis* MSC5 was isolated from a sea anemone sample from the Mariana Trench (11.722° N, 138.723° E) in the Pacific Ocean and identified based on ITS sequencing. Genome sequencing was performed at Shanghai Personalbio Technology Co., Ltd. The resulting contigs were assembled into 40 scaffolds using CANU 1.7.1.[46] Gene prediction was conducted using Augustus (version 3.03), GlimmerHMM (version 3.0.1), and GeneMark-ES (version 4.35), and the gene functions were annotated using seven databases (NR, eggNOG, KEGG, SwissProt, GO, P450, and TCDB). The BGCs of *H. ingelheimensis* MSC5 were analyzed using antiSMASH 6.1.1 fungal version (https://fungismash.secondarymetabolites.org/). The genome sequence of this strain can be accessed via BioProject accession ID PRJNA1105766.

MALDI-TOF MS Data Acquisition. α-Cyano-4-hydroxycinnamic acid (HCCA) and 2,5-dihydroxybenzoic acid (DHB) were used as the MALDI matrix. A 2 μL amount of microbial crude extract was spotted directly on a stainless steel 384-well target plate and overlaid with 2 μL of the matrix. The MALDI-TOF spectra were recorded in

https://doi.org/10.1021/acs.jnatprod.4c00740
J. Nat. Prod. 2024, 87, 2695−2708

reflector/positive ion mode, and each sample was ablated 1000 times in 1 s with the laser energy set to 60%. The m/z range for detecting secondary metabolites was selected as 300−2000. All mass data were analyzed using FlexAnalysis 3.3 software.

OSMAC Strategy. Twelve different media (Table S3) were tested to determine the optimal cultural conditions for *H. ingelheimensis* MSC5. The liquid shake flask cultures were incubated at 28 °C and 180 rpm for 8 days. For stationary cultures, the rice medium was incubated at 28 °C for 30 days. EtOAc extracts of those different cultural media of *H. ingelheimensis* MSC5 were dissolved in methanol to a concentration of 1 mg/mL. HPLC analysis was performed by using a UHPLC system. The mobile phase consisted of 0.05% formic acid in water (A) and MeCN (B). The optimized elution gradient was as follows: linear gradient from 5% to 95% solvent B from 0.5 to 7.5 min, followed by maintenance at 95% B from 7.5 to 11.5 min, and re-equilibration at 5% B from 11.5 to 13.5 min, at a constant flow rate of 0.4 mL/min, injection volume: 5 μL.

LC-MS/MS-Based Molecular Network Analysis. EtOAc extracts of the rice culture of *H. ingelheimensis* MSC5 were dissolved in methanol to a concentration of 0.1 mg/mL. LC-MS/MS analysis was performed on a UHPLC system coupled to a Q Exactive Plus mass spectrometer equipped with an ESI source. The mobile phase consisted of 0.05% formic acid in water (A) and MeCN (B). The optimized elution gradient was as follows: linear gradient from 5% to 95% solvent B from 0.5 to 7.5 min, followed by maintenance at 95% B from 7.5 to 11.5 min, and re-equilibration at 5% B from 11.5 to 13.5 min, at a constant flow rate of 0.4 mL/min, injection volume: 1 μL. ESI source parameters were set as follows: capillary temperature: 320 °C; spray voltage: 3.2 kV; s-lens RF level: 50 V. Data were acquired using Xcalibur 3.0 (Thermo Fisher). Raw data were converted to mzXML format files using MSCovertGUI 3.0 software. Molecular networking was performed using GNPS (https://gnps.ucsd.edu/).[9] The generated molecular network was analyzed and visualized in Cytoscape 3.7.2.[47] The GNPS analysis results can be viewed at https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=64bd8e682e99460f9c0c57c98b3b0201.

Large-Scale Cultivation and Extraction of *H. ingelheimensis* MSC5. The strain *H. ingelheimensis* MSC5 was incubated in a 250 mL Erlenmeyer flask containing 100 mL of PDB liquid medium (soluble starch 10 g L^{-1}, peptone 2 g L^{-1}, yeast extract powder 4 g L^{-1}, NaCl 5 g L^{-1}) and shaken at 180 rpm and 28 °C for 3 days. After that, the seed culture was transferred into a 500 mL flask containing 80 g of rice and 120 mL of seawater with an inoculation rate of 5%. The culture was incubated at room temperature for 30 days. The whole culture (50 L) of *H. ingelheimensis* MSC5 was extracted with an equal volume of EtOAc three times and dried in vacuo to yield 39.0 g of extract.

Purification. The extract was separated into 13 fractions (Fr. 1−Fr. 13) by silica gel column chromatography using gradient elution with ethyl acetate−petroleum ether (0−100%) and MeOH−CH_2Cl_2 (0−50%), respectively. Fr. 9 (4.2 g) and Fr. 10 (3.1 g) were subjected to an RP-silica gel column eluting with a stepwise gradient of MeOH−H_2O (10−100%) to yield 11 fractions (Fr. 9.1−Fr. 9.5 and Fr. 10.1−Fr. 10.6). Fr. 9.4 (224.9 mg) was further purified by semipreparative HPLC (Eclipse XDB-C18 5 μm, 9.4 mm × 250 mm, 4.0 mL/min gradient elution over 45 min from 90% H_2O−MeCN to 20% H_2O−MeCN with a constant 0.05% TFA modifier) to yield compounds **4** (0.8 mg, t_R = 15.4 min), **3** (4.1 mg, t_R = 16.5 min), **9** (2.1 mg, t_R = 18.1 min), **1** (6.9 mg, t_R = 18.6 min), and **2** (4.7 mg, t_R = 19.3 min). Fr. 10.4 (154.4 mg) was purified by semipreparative HPLC with the same elution method to yield compounds **7** (1.7 mg, t_R = 8.4 min), **8** (0.9 mg, t_R = 16.5 min), **10** (2.1 mg, t_R = 16.5 min), and **11** (0.7 mg, t_R = 16.5 min). Fr. 10.5 (38.4 mg) was purified by semipreparative HPLC eluting with 38% MeCN to afford compounds **5** (9.4 mg, t_R = 27.5 min) and **6** (6.3 mg, t_R = 30.1 min). Fr. 10.6 (22.1 mg) was purified by semipreparative HPLC eluting with 60% MeCN to afford compound **12** (1.1 mg, t_R = 10.2 min). A known compound, avellanin C (**13**), was obtained from Fr. 8 as the major component.

*Avellanin D (**1**).* Yellow oil; $[\alpha]^{20}_D$ +107.8 (c 0.018, MeOH); UV (MeOH) λ_{max} (log ε) 212 (0.9), 219 (0.9) nm; IR (KBr) ν_{max} 3326, 2933, 1678, 1502, 1301, 761 cm^{-1}; ^1H and ^{13}C NMR, see Table 1; HRESIMS m/z 550.2656 $[M + H]^+$ (calcd for $C_{29}H_{36}N_5O_6^+$, 550.2660). (IR, NMR, and HRESIMS spectra, see Figures S4−S11.)

*Avellanin E (**2**).* Colorless crystal; mp 161−163 °C; $[\alpha]^{20}_D$ +50.5 (c 0.019, MeOH); UV (MeOH) λ_{max} (log ε) 207 (2.9) nm; IR (KBr) ν_{max} 3334, 2928, 1678, 1415, 1108, 760 cm^{-1}; ^1H and ^{13}C NMR, see Table 1; HRESIMS m/z 612.2811 $[M + H]^+$ (calcd for $C_{34}H_{38}N_5O_6^+$, 612.2817). (IR, NMR, and HRESIMS spectra, see Figures S12−S19.)

*Avellanin F (**3**).* White crystal; mp 229−232 °C; $[\alpha]^{20}_D$ +45.3 (c 0.019, MeOH); UV (MeOH) λ_{max} (log ε) 212(1.2), 250 (7.1) nm; IR (KBr) ν_{max} 3427, 3027, 2933, 1667, 1616, 762 cm^{-1}; ^1H and ^{13}C NMR, see Table 1; HRESIMS m/z 612.2809 $[M + H]^+$ (calcd for $C_{34}H_{38}N_5O_6^+$, 612.2817). (IR, NMR, and HRESIMS spectra, see Figures S20−S27.)

*Avellanin G (**4**).* White powder; $[\alpha]^{20}_D$ +118.0 (c 0.018, MeOH); UV (MeOH) λ_{max} (log ε) 217 (2.1), 263 (0.5) nm; IR (KBr) ν_{max} 3333, 2932, 1516, 1418, 757 cm^{-1}; ^1H and ^{13}C NMR, see Table 1; HRESIMS m/z 628.2764 $[M + H]^+$ (calcd for $C_{34}H_{38}N_5O_7^+$, 628.2766). (IR, NMR, and HRESIMS spectra, see Figures S28−S35.)

*Avellanin H (**5**).* Colorless needle crystal; mp 156−161 °C; $[\alpha]^{20}_D$ +38.4 (c 0.020, MeOH); UV (MeOH) λ_{max} (log ε) 282 (1.4) nm; IR (KBr) ν_{max} 3356, 2930, 1680,1499, 1414, 760 cm^{-1}; ^1H and ^{13}C NMR, see Table 2; HRESIMS m/z 582.2704 $[M + H]^+$ (calcd for $C_{33}H_{36}N_5O_5^+$, 582.2711). (IR, NMR and HRESIMS spectra, see Figures S36−S43.)

*Avellanin I (**6**).* Colorless crystal; mp 174−178 °C; $[\alpha]^{20}_D$ +48.0 (c 0.020, MeOH); UV (MeOH) λ_{max} (log ε) 206 (4.1) nm; IR (KBr) ν_{max} 3328, 2933, 1663, 1293, 905, 758 cm^{-1}; ^1H and ^{13}C NMR, see Table 2; HRESIMS m/z 646.2423 $[M + H]^+$ (calcd for $C_{34}H_{37}N_5O_6Cl^+$, 646.2427). (IR, NMR, and HRESIMS spectra, see Figures S44−S51.)

*Avellanin J (**7**).* Colorless crystal; mp 180−183 °C; $[\alpha]^{20}_D$ +89.6 (c 0.023, MeOH); UV (MeOH) λ_{max} (log ε) 206 (2.9) nm; IR (KBr) ν_{max} 3330, 2880, 1588, 1501, 1417, 759 cm^{-1}; ^1H and ^{13}C NMR, see Table 2; HRESIMS m/z 592.2765 $[M + H]^+$ (calcd for $C_{31}H_{38}N_5O_7^+$, 592.2766). (IR, NMR, and HRESIMS spectra, see Figures S52−S59.)

*Avellanin K (**8**).* Colorless crystal; mp 138−140 °C; $[\alpha]^{20}_D$ +196.0 (c 0.025, MeOH); UV (MeOH) λ_{max} (log ε) 224 (6.8) nm; IR (KBr) ν_{max} 3331, 2928, 1667, 1518, 1414, 1001, 740 cm^{-1}; ^1H and ^{13}C NMR, see Table 2; HRESIMS m/z 635.2968 $[M + H]^+$ (calcd for $C_{36}H_{39}N_6O_5^+$, 635.2976). (IR, NMR, and HRESIMS spectra, see Figures S60−S67.)

*Avellanin L (**9**).* White powder; $[\alpha]^{20}_D$ +98.3 (c 0.160, MeOH); UV (MeOH) λ_{max} (log ε) 204 (1.5) nm; IR (KBr) ν_{max} 3335, 3062, 3028, 2930, 2480, 1665 cm^{-1}; ^1H and ^{13}C NMR, see Table 3; HRESIMS m/z 612.2820 $[M + H]^+$ (calcd for $C_{34}H_{38}N_5O_6^+$, 612.2817). (IR, NMR, and HRESIMS spectra, see Figures S68−S76.)

*Avellanin M (**10**).* White powder; $[\alpha]^{20}_D$ +65.1 (c 0.070, MeOH); UV (MeOH) λ_{max} (log ε) 203 (2.3), 285 (8.4) nm; IR (KBr) ν_{max} 3336, 3027, 2934, 1676, 1617, 1518 cm^{-1}; ^1H and ^{13}C NMR, see Table 3; HRESIMS m/z 598.2658 $[M + H]^+$ (calcd for $C_{33}H_{36}N_5O_6^+$, 598.2664). (IR, NMR, and HRESIMS spectra, see Figures S77−S84.)

*Avellanin N (**11**).* White powder; $[\alpha]^{20}_D$ +46.4 (c 0.047, MeOH); UV (MeOH) λ_{max} (log ε) 206 (6.3), 241(2.6), 251(2.5) nm; IR (KBr) ν_{max} 3333, 3064, 3028, 2930, 1668, 1589 cm^{-1}; ^1H and ^{13}C NMR, see Table 3; HRESIMS m/z 628.2764 $[M + H]^+$ (calcd for $C_{34}H_{38}N_5O_7^+$, 628.2766). (IR, NMR, and HRESIMS spectra, see Figures S85−S92.)

*Avellanin O (**12**).* White powder; $[\alpha]^{20}_D$ +21.5 (c 0.040, MeOH); UV (MeOH) λ_{max} (log ε) 206 (2.2), 211 (2.2), 294 (1.5) nm; IR (KBr) ν_{max} 3337, 2921, 2852, 1819, 1739, 1679 cm^{-1}; ^1H and ^{13}C NMR, see Table 3; HRESIMS m/z 610.3023 $[M + H]^+$ (calcd for $C_{35}H_{40}N_5O_5^+$, 610.3029). (IR, NMR, and HRESIMS spectra, see Figures S93−S100.)

X-ray Crystallographic Analysis. Colorless crystals of compounds **2**, **3**, **5**, **6**, **7**, **8**, and **13** were obtained from MeOH aqueous solution evaporated at 4 °C. Single-crystal X-ray diffraction data were collected by a Bruker SC-XRD D8 VENTURE CMOS Photon II

https://doi.org/10.1021/acs.jnatprod.4c00740
J. Nat. Prod. 2024, 87, 2695−2708

diffractometer using Helios MX multilayer monochromator Cu Kα radiation (λ = 1.54178 Å) adopting the direct-drive rotating-anode technique. All structures were solved by Intrinsic Phasing with ShelXT and refined against all F^2 values using the ShelXL and the OLEX2 suite of programs. The parameters in Crystallographic Information File (CIF) format for **2**, **3**, **5**, **6**, **7**, **8**, and **13** have been deposited with the Cambridge Crystallographic Data Center (CCDC) under the deposition numbers CCDC 2345875 (**2**), CCDC 2345869 (**3**), CCDC 2345877 (**5**), CCDC 2345880 (**6**), CCDC 2345881 (**7**), CCDC 2345882 (**8**), and CCDC 2347221 (**13**) (Figure 4).

Crystallographic data of compound **2**: $C_{34}H_{43}N_5O_9$, M = 665.73, orthorhombic, space group $P2_12_12_1$ (no. 19), a = 11.1279(3) Å, b = 13.0415(4) Å, c = 24.2057(6) Å, α = 90°, β = 90°, γ = 90°, V = 3512.84(17) Å^3, Z = 4, T = 173.00 K, μ(Cu Kα) = 0.761 mm^{-1}, D_{calc} = 1.259 g/cm^3, 34 465 reflections measured (7.7° ≤ 2θ ≤ 136.6°), 6432 unique (R_{int} = 0.0465, R_{sigma} = 0.0302), which were used in all calculations. The final R_1 was 0.0355 (I > 2σ(I)), and wR_2 was 0.0941; Flack parameter = − 0.09(7); Hooft parameter = − 0.07(7).

Crystallographic data for compound **3**: $C_{35}H_{41}N_5O_7$, M = 643.73, orthorhombic, space group $P2_12_12_1$ (no. 19), a = 10.7335(3) Å, b = 14.8195(3) Å, c = 21.1875(5) Å, α = 90°, β = 90°, γ = 90°, V = 3370.19(14) Å^3, Z = 4, T = 173.00 K, μ(Cu Kα) = 0.732 mm^{-1}, D_{calc} = 1.269 g/cm^3, 58 984 reflections measured (7.28° ≤ 2θ ≤ 137.018°), 6177 unique (R_{int} = 0.0544, R_{sigma} = 0.0230), which were used in all calculations. The final R_1 was 0.0465 (I > 2σ(I)), and wR_2 was 0.1391; Flack parameter = − 0.08(7); Hooft parameter = − 0.14(7).

Crystallographic data of compound **5**: $C_{35}H_{43}N_5O_7$, M = 645.74, orthorhombic, space group $P2_12_12_1$ (no. 19), a = 13.7582(3) Å, b = 14.4762(3) Å, c = 16.8559(4) Å, α = 90°, β = 90°, γ = 90°, V = 3357.13(13) Å^3, Z = 4, T = 173.00 K, μ(Cu Kα) = 0.735 mm^{-1}, D_{calc} = 1.278 g/cm^3, 40 210 reflections measured (8.05° ≤ 2θ ≤ 137.684°), 6207 unique (R_{int} = 0.0642, R_{sigma} = 0.0339), which were used in all calculations. The final R_1 was 0.0361 (I > 2σ(I)), and wR_2 was 0.0955; Flack parameter = 0.01 (10); Hooft parameter = 0.02(9).

Crystallographic data for compound **6**: $C_{68}H_{74}Cl_2N_{10}O_{13}$, M = 1310.27, triclinic, space group $P1$ (no. 1), a = 8.6366(2) Å, b = 10.7030(2) Å, c = 18.0421(3) Å, α = 97.563(2)°, β = 97.702(2)°, γ = 91.910(2)°, V = 1636.14(6) Å^3, Z = 1, T = 178(6) K, μ(Cu Kα) = 1.487 mm^{-1}, D_{calc} = 1.330 g/cm^3, 48 775 reflections measured (4.99° ≤ 2θ ≤ 133.196°), 11 221 unique (R_{int} = 0.0791, R_{sigma} = 0.0483), which were used in all calculations. The final R_1 was 0.0548 (I > 2σ(I)), and wR_2 was 0.1285; Flack parameter = 0.004(7); Hooft parameter = 0.01(6).

Crystallographic data of compound **7**: $C_{31}H_{40}N_5O_{8.5}$, M = 618.68, monoclinic, space group C2 (no. 5), a = 21.1753(4) Å, b = 15.1434(3) Å, c = 11.3530(2) Å, α = 90°, β = 120.2610(10)°, γ = 90°, V = 3144.46(11) Å^3, Z = 4, T = 173.00 K, μ(Cu Kα) = 0.797 mm^{-1}, D_{calc} = 1.307 g/cm^3, 25 073 reflections measured (7.578° ≤ 2θ ≤ 136.778°), 5759 unique (R_{int} = 0.0446, R_{sigma} = 0.0335), which were used in all calculations. The final R_1 was 0.0331 (I > 2σ(I)), and wR_2 was 0.0820; Flack parameter = − 0.11(8); Hooft parameter = − 0.08(8).

Crystallographic data for compound **8**: $C_{36}H_{40}N_6O_6$, M = 652.74, orthorhombic, space group $P2_12_12_1$ (no. 18), a = 14.6367(3) Å, b = 21.0611(4) Å, c = 10.8230(2) Å, α = 90°, β = 90°, γ = 90(2)°, V = 3336.35(11) Å^3, Z = 4, T = 173.00 K, μ(Cu Kα) = 0.734 mm^{-1}, D_{calc} = 1.299 g/cm^3, 31 177 reflections measured (7.354° ≤ 2Θ ≤ 136.978°), 6141 unique (R_{int} = 0.0741, R_{sigma} = 0.0434), which were used in all calculations. The final R_1 was 0.0394 (I > 2σ(I)), and wR_2 was 0.0986; Flack parameter = 0.01(12); Hooft parameter = 0.07(13).

Crystallographic data of compound **13**: $C_{71}H_{86}N_{10}O_3$, M = 1287.534, monoclinic, space group $P2_1$ (no. 4), a = 13.9438(3) Å, b = 14.6929(4) Å, c = 33.7760(8) Å, β = 97.760(1)°, V = 6856.5(3) Å^3, Z = 4, T = 173.00 K, μ(Cu Kα) = 0.707 mm^{-1}, D_{calc} = 1.247 g/cm^3, 64 328 reflections measured (5.28° ≤ 2θ ≤ 136.82°), 24 599 unique (R_{int} = 0.0566, R_{sigma} = 0.0562), which were used in all calculations. The final R_1 was 0.0517 (I > = 2σ(I)), and wR_2 was 0.1391; Flack parameter = − 0.12(8); Hooft parameter = − 0.12(8).

Marfey's Analysis. Avellanins D (**1**), G (**4**), L (**9**), M (**10**), N (**11**), and O (**12**) were hydrolyzed (0.4 mg of each compound), respectively, in 6 M HCl at 110 °C for 10 h, and the reaction mixture was evaporated to dryness. After that, the acidic hydrolysate was then dissolved in 100 μL of 1 M $NaHCO_3$, followed by the addition of 100 μL of Marfey's reagent (N^{α}-(5-fluoro-2,4-dinitrophenyl)-L-leucinamide) in acetone at a concentration of 1 mg/mL. Standard amino acids were treated in the same way. Samples and standard amino acids were each sealed and incubated at 50 °C for 1.5 h, then quenched with 2 M HCl (20 μL) and dried in vacuo. The reaction products were dissolved in 500 μL of MeOH, and 10 μL of the solution was used for UPLC-MS analysis with a reversed-phase C_{18} column (1.7 μm, 2.1 × 100 mm) under a gradient condition (flow rate: 0.4 mL/min, mobile phase: CH_3CN−H_2O, 0.1% FA in H_2O; 0−0.5 min: 25% CH_3CN; 0.5 min−1.0 min: 25−35% CH_3CN; 1.0−6.0 min: 35−50% CH_3CN; 6.0−8.5 min: 50−85% CH_3CN; 8.5−9.0 min: 85−100% CH_3CN; 9.0−11.0 min: 100% CH_3CN; 11.0−11.5 min: 100−25% CH_3CN; 11.5−13.0 min: 25% CH_3CN). The retention times for the FDLA-amino acid derivatives were as follows: L-Ala (3.4 min), D-Ala (4.4 min), L-Pro (3.5 min), D-Pro (4.1 min), *N*Me-L-Phe (5.7 min), *N*Me-D-Phe (6.1 min), L-Ser (2.6 min), D-Ser (2.8 min), L-Abu (3.8 min), D-Abu (5.3 min), L-Tyr (3.7 min), D-Tyr (4.3 min), L-Phe (5.1 min), D-Phe (7.0 min), L-Hse (2.6 min), D-Hse (2.9 min), 2*S*, 4*R*-4-OH-Pro (2.3 min), 2*S*, 4*S*-4-OH-Pro (2.5 min), 2*R*, 4*S*-4-OH-Pro (2.3 min), 2*R*, 4*R*-4-OH-Pro (2.6 min). The L-FDLA derivatives of 2*S*, 4*R*-4-OH-Pro and 2*R*, 4*S*-4-OH-Pro had the same retention time and were further separated on a chiral column (DAICEL GROUP CHIRALPAK ZWIX (-), 3 μm, 3 × 150 mm) with an isocratic elution condition (flow rate: 0.3 mL/min; mobile phase: CH_3OH−H_2O = 9:1 with 50 mM FA and 25 mM DEA). The retention time of L-FDLA derivatives of 2*S*,4*R*-4-OH-Pro was 5.83 and 5.23 min for 2*R*, 4*S*-4-OH-Pro, respectively.

Cytotoxicity Activity. The *in vitro* cytotoxicity of compounds **1**−**6** against 18 human cancer cell lines (human lung cancer cell line A549, human gastric carcinoma cell line MKN-45, human colorectal carcinoma cell line HCT-116, human cervical cancer cell line HeLa, human chronic myeloid leukemia cell line K562, human kidney clear cell carcinoma cells 786-O, human esophageal cancer cell line TE-1, human bladder cancer cell line 5637, human gallbladder cancer cells GBC-SD, human breast cancer cell line MCF-7, human hepatocellular carcinoma cell line HepG2, human brain glioma cell line SF126, human prostate cancer cell line DU145, human thyroid cancer cell line GAL-62, human pancreatic cancer cell line PATU8988T, human osteosarcoma cell line HOS, human melanoma cell line A-375, human rhabdomyosarcoma cell line A-673) and two human cell lines (fetal hepatocyte line L-02 and human embryonic kidney 293T) was determined using a cell counting kit-8 (CCK-8) with doxorubicin as the positive control (cancer cell lines were purchased from Qingdao AC Biotechnology Co., Ltd.).[48] IC_{50} values were calculated from dose−response curves using Graph Prism 8 (version 8.0.2, from GraphPad Software Inc.). In brief, cells in exponential growth at a density of 3 × 10^3 cells per well were seeded into 96-well flat-microtiter plates for 24 h. Subsequently, the culture medium was replaced with fresh medium containing eight different concentrations of the test compounds, and 10 μL of CCK-8 solution was added to each well. Then the cells were further incubated for 48 h. After a 4 h incubation at 37 °C with 5% CO_2, the absorbance was measured at 450 nm. Cell viability was calculated as a percentage relative to the control, which were considered as having 100% viable cells. Each experiment was performed in triplicate.

Antimalarial Assay. *P. falciparum* 3D7 was cultivated in RPMI 1640 complete medium with additional 25 mM HEPES, 50 mg/L hypoxanthine, 2.1 g/L $NaHCO_3$, and 0.5% bovine serum albumin, pH 7.2. Synchronous ring-stage *P. falciparum* 3D7 was generated through treatment with 5% sorbitol and subsequently diluted with human erythrocytes and complete medium to attain the initial composition of 2% hematocrit and 0.5% parasitemia. A 90 μL amount of the cellular suspension was added into individual test wells, with nonparasitized erythrocytes at 4% hematocrit used for reference control. The test compounds (in DMSO) were serially diluted with RPMI 1640

https://doi.org/10.1021/acs.jnatprod.4c00740
J. Nat. Prod. 2024, 87, 2695−2708

complete medium and transferred into test wells with a final volume of 100 μL. Each concentration was tested in triplicate with the final concentration ranging from 0 to 10^{-4} M. After a 48 h incubation at 37 °C of this tested 96-well plate, the antimalarial activity of avellanins against *P. falciparum* was evaluated using a fluorescence-based method.[49,50] In brief, 100 μL of supernatant and 100 μL of lysis buffer (0.08% (v/v) Triton X-100, 0.008% (w/v) saponin, 5 mM EDTA, 20 mM Tris, pH 7.5) with SYBR green-I solution were transferred into a 96-well plate. Then the plate was placed for 2 h in darkness at room temperature. Fluorescence was measured with a multifunctional microplate reader (Multiskan MK3, Thermo) with the excitation and emission wavelengths of 485 and 535 nm, respectively. Background fluorescence was subtracted from the empty wells, and the counts were then plotted and analyzed by linear regression using Graph Prism 8 (version 8.0.2, from GraphPad Software Inc.).

■ ASSOCIATED CONTENT

Data Availability Statement

The NMR data for compounds **1**−**12** have been deposited in The Harvard Dataverse (10.7910/DVN/HOU5DG).

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jnatprod.4c00740.

(BGC summary of *H. ingelheimensis* MSC5; media composition; OSMAC screening for 12 different culture conditions of *H. ingelheimensis* MSC5; MALDI-TOF-MS spectra of the culture supernatants of *H. ingelheimensis* MSC5; ^{1}H NMR spectra of EtOAc extracts of *H. ingelheimensis* MSC5; IR, NMR, and HRESIMS spectra of **1**−**12**; Marfey's data of **1**−**12**; MS^2 fragmentation raw data of **1**−**12** PDF)

X-ray crystallographic data for **1**−**12** (ZIP)

■ AUTHOR INFORMATION

Corresponding Author

Zhiyong Li − *State Key Laboratory of Microbial Metabolism, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, P.R. China; Yazhou Bay Institute of Deepsea Sci-Tech, Shanghai Jiao Tong University, Sanya 572025, P.R. China*; Email: zyli@sjtu.edu.cn

Authors

Hao Li − *State Key Laboratory of Microbial Metabolism, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, P.R. China; Yazhou Bay Institute of Deepsea Sci-Tech, Shanghai Jiao Tong University, Sanya 572025, P.R. China*; orcid.org/0000-0002-8594-1007

Yuling Chen − *State Key Laboratory of Microbial Metabolism, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, P.R. China*

Bingqing Tang − *State Key Laboratory of Microbial Metabolism, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, P.R. China; Yazhou Bay Institute of Deepsea Sci-Tech, Shanghai Jiao Tong University, Sanya 572025, P.R. China*

Zhengjie Liu − *State Key Laboratory of Microbial Metabolism, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, P.R. China*; orcid.org/0009-0007-7688-6636

Bo Peng − *State Key Laboratory of Microbial Metabolism, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, P.R. China*

Jiajun Li − *State Key Laboratory of Microbial Metabolism, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, P.R. China*

Han Gao − *CAS Key Laboratory of Insect Developmental and Evolutionary Biology, CAS Center for Excellence in Molecular Plant Sciences, Shanghai Institute of Plant Physiology and Ecology, Chinese Academy of Science, Shanghai 200032, P.R. China*

Sibao Wang − *CAS Key Laboratory of Insect Developmental and Evolutionary Biology, CAS Center for Excellence in Molecular Plant Sciences, Shanghai Institute of Plant Physiology and Ecology, Chinese Academy of Science, Shanghai 200032, P.R. China*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jnatprod.4c00740

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was financially supported by the Shanghai Municipal Science and Technology Major Project, National Key Research and Development Program of China (2022YFC280410), and Sanya Science and Technology Innovation Special Project (2022KJCX63). We thank Dr. Liang Dong (School of Oceanography, Shanghai Jiao Tong University) for providing the sample collected from Mariana Trench. We thank Prof. Minjuan Xu (Key Laboratory of Systems Biomedicine (Ministry of Education), Shanghai Centre for Systems Biomedicine, Shanghai Jiao Tong University) for very useful discussions. We express our gratitude to Jieli Wu, Dr. Lei Feng, and Dr. Ningjin Zhang (Instrumental Analysis Center, Shanghai Jiao Tong University, Shanghai, China) for their valuable suggestions on the spectroscopic techniques.

■ REFERENCES

(1) Newman, D. J.; Cragg, G. M. *J. Nat. Prod.* **2020**, *83*, 770−803.
(2) Genilloud, O. *Curr. Opin. Microbiol.* **2019**, *51*, 81−87.
(3) Skinnider, M. A.; Magarvey, N. A. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, E6271−E6272.
(4) Hoffmann, T.; Krug, D.; Huttel, S.; Muller, R. *Anal. Chem.* **2014**, *86*, 10780−10788.
(5) Chanana, S.; Thomas, C. S.; Zhang, F.; Rajski, S. R.; Bugni, T. S. *Metabolites* **2020**, *10*, 297.
(6) Chen, L.; Lu, W.; Wang, L.; Xing, X.; Chen, Z.; Teng, X.; Zeng, X.; Muscarella, A. D.; Shen, Y.; Cowan, A.; et al. *Nat. Methods* **2021**, *18*, 1377−1385.
(7) Wolfender, J.-L.; Litaudon, M.; Touboul, D.; Queiroz, E. F. *Nat. Prod. Rep.* **2019**, *36*, 855−868.
(8) Schorn, M. A.; Verhoeven, S.; Ridder, L.; Huber, F.; Acharya, D. D.; Aksenov, A. A.; Aleti, G.; Moghaddam, J. A.; Aron, A. T.; Aziz, S.; et al. *Nat. Chem. Biol.* **2021**, *17*, 363−368.
(9) Wang, M.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kapono, C. A.; Luzzatto-Knaan, T.; et al. *Nat. Biotechnol.* **2016**, *34*, 828−837.
(10) Scherlach, K.; Hertweck, C. *Nat. Commun.* **2021**, *12*, 3864.
(11) Zhang, F.; Ramos Alvarenga, R. F.; Throckmorton, K.; Chanana, S.; Braun, D. R.; Fossen, J.; Zhao, M.; McCrone, S.; Harper, M. K.; Rajski, S. R.; et al. *Org. Lett.* **2022**, *24*, 3998−4002.
(12) Wu, Q.; Bell, B. A.; Yan, J. X.; Chevrette, M. G.; Brittin, N. J.; Zhu, Y.; Chanana, S.; Maity, M.; Braun, D. R.; Wheaton, A. M.; et al. *J. Am. Chem. Soc.* **2023**, *145*, 58−69.
(13) Wang, C.; Xiao, D.; Dun, B.; Yin, M.; Tsega, A. S.; Xie, L.; Li, W.; Yue, Q.; Wang, S.; Gao, H.; et al. *Proc. Natl. Acad. Sci. U. S. A.* **2022**, *119*, e2123379119.

(14) Dello Russo, P.; Del Bufalo, A.; Fini, M. *EXCLI J.* **2015**, *14*, 228−236.
(15) Skropeta, D.; Wei, L. *Nat. Prod. Rep.* **2014**, *31*, 999−1025.
(16) Pilkington. *Molecules* **2019**, *24*, 3942.
(17) Siro, G.; Donald, L.; Pipite, A. *Diversity* **2023**, *15*, 30.
(18) McGivern, J. G. *Neuropsychiatr. Dis. Treat.* **2007**, *3*, 69−85.
(19) Ma, L.; Diao, A. *ANTI-CANCER. AGENT. ME. (Formerly Current Medicinal Chemistry-Anti-Cancer Agents)* **2015**, *15*, 298−306.
(20) Qin, L.; Yong, K.; Lian, X.-Y.; Zhang, Z. *Phytochem. Lett.* **2022**, *49*, 79−82.
(21) Carroll, A. R.; Copp, B. R.; Davis, R. A.; Keyzers, R. A.; Prinsep, M. R. *Nat. Prod. Rep.* **2021**, *38*, 362−413.
(22) Huo, Z.-Q.; Zhu, F.; Zhang, X.-W.; Zhang, X.; Liang, H.-B.; Yao, J.-C.; Liu, Z.; Zhang, G.-M.; Yao, Q.-Q.; Qin, G.-F. *Mar. Drugs* **2022**, *20*, 333.
(23) Igarashi, Y.; Hanafusa, T.; Gohda, F.; Peterson, S.; Bills, G. *Mycology* **2014**, *5*, 102−109.
(24) Yamazaki, M.; Horie, Y.; Bae, K.; Maebayashi, Y.; Jisai, Y.; Fujimoto, H. *Chem. Pharm. Bull. (Tokyo)* **1987**, *35*, 2122−2124.
(25) Igarashi, Y.; Gohda, F.; Kadoshima, T.; Fukuda, T.; Hanafusa, T.; Shojima, A.; Nakayama, J.; Bills, G. F.; Peterson, S. *J. Antibiot (Tokyo)* **2015**, *68*, 707−710.
(26) Almeida, C.; Perez-Victoria, I.; Gonzalez-Menendez, V.; de Pedro, N.; Martin, J.; Crespo, G.; Mackenzie, T.; Cautain, B.; Reyes, F.; Vicente, F.; et al. *J. Nat. Prod.* **2018**, *81*, 1488−1492.
(27) Klaram, R.; Dethoup, T.; Machado, F. P.; Gales, L.; Kumla, D.; Hafez Ghoran, S.; Sousa, E.; Mistry, S.; Silva, A. M. S.; Kijjoa, A. *Mar. Drugs* **2023**, *21*, 344.
(28) Caesar, L. K.; Butun, F. A.; Robey, M. T.; Ayon, N. J.; Gupta, R.; Dainko, D.; Bok, J. W.; Nickles, G.; Stankey, R. J.; Johnson, D.; et al. *Nat. Chem. Biol.* **2023**, *19*, 846−854.
(29) Li, Y.; Li, Z. *Mar. Drugs* **2021**, *19*, 537.
(30) Yu, M.; Li, Y.; Banakar, S. P.; Liu, L.; Shao, C.; Li, Z.; Wang, C. *Front. Microbiol.* **2019**, *10*, 915.
(31) Li, H.; Li, Z. *Bioengineering* **2022**, *9*, 707.
(32) Li, Y.; Zhang, F.; Banakar, S.; Li, Z. *Appl. Microbiol. Biotechnol.* **2018**, *102*, 7865−7875.
(33) Peng, Q.; Gao, G.; Lu, J.; Long, Q.; Chen, X.; Zhang, F.; Xu, M.; Liu, K.; Wang, Y.; Deng, Z.; et al. *Front. Microbiol.* **2018**, *9*, 3042.
(34) Blin, K.; Shaw, S.; Kloosterman, A. M.; Charlop-Powers, Z.; van Wezel, G. P.; Medema, M. H.; Weber, T. *Nucleic Acids Res.* **2021**, *49*, W29−W35.
(35) Shin, Y. H.; Im, J. H.; Kang, I.; Kim, E.; Jang, S. C.; Cho, E.; Shin, D.; Hwang, S.; Du, Y. E.; Huynh, T. H.; et al. *J. Am. Chem. Soc.* **2023**, *145*, 1886−1896.
(36) Jiang, M.; Chen, S.; Lu, X.; Guo, H.; Chen, S.; Yin, X.; Li, H.; Dai, G.; Liu, L. *J. Agric. Food. Chem.* **2023**, *71*, 9782−9795.
(37) Siemion, I. Z.; Wieland, T.; Pook, K. H. *Angew. Chem., Int. Ed. Engl.* **1975**, *14*, 702−703.
(38) de Mattos-Shipley, K. M. J.; Greco, C.; Heard, D. M.; Hough, G.; Mulholland, N. P.; Vincent, J. L.; Micklefield, J.; Simpson, T. J.; Willis, C. L.; Cox, R. J.; et al. *Chem. Sci.* **2018**, *9*, 4109−4117.
(39) Hendra, R.; Salib, M. N.; Molinski, T. F. *J.Nat. Prod.* **2022**, *85*, 2207−2216.
(40) Heinilä, L. M. P.; Fewer, D. P.; Jokela, J. K.; Wahlsten, M.; Ouyang, X.; Permi, P.; Jortikka, A.; Sivonen, K. *Org. Biomol. Chem.* **2021**, *19*, 5577−5588.
(41) Baldeweg, F.; Kage, H.; Schieferdecker, S.; Allen, C.; Hoffmeister, D.; Nett, M. *Org. Lett.* **2017**, *19*, 4868−4871.
(42) BORMANN, C.; HUHN, W.; ZÄHNER, H. *J. Antibiot.* **1985**, *38*, 9−16.
(43) Park, J. H.; Lee, S. Y. *Biotechnol. J.* **2010**, *5*, 560−577.
(44) Shepherd, R. A.; Earp, C. E.; Cank, K. B.; Raja, H. A.; Burdette, J.; Maher, S. P.; Marin, A. A.; Ruberto, A. A.; Mai, S. L.; Darveaux, B. A.; et al. *J. Antibiot.* **2023**, *76*, 642−649.
(45) Wang, X.; Lin, M.; Xu, D.; Lai, D.; Zhou, L. *Molecules* **2017**, *22*, 2069.
(46) Koren, S.; Walenz, B. P.; Berlin, K.; Miller, J. R.; Bergman, N. H.; Phillippy, A. M. *Genome Res.* **2017**, *27*, 722−736.
(47) Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. *Genome Res.* **2003**, *13*, 2498−2504.
(48) Tominaga, H.; Ishiyama, M.; Ohseto, F.; Sasamoto, K.; Hamamoto, T.; Suzuki, K.; Watanabe, M. *Anal. Commun.* **1999**, *36*, 47−50.
(49) Smilkstein, M.; Sriwilaijaroen, N.; Kelly, J. X.; Wilairat, P.; Riscoe, M. *Antimicrob. Agents Chemother.* **2004**, *48*, 1803−1806.
(50) Malmquist, N. A.; Moss, T. A.; Mecheri, S.; Scherf, A.; Fuchter, M. J. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 16708−16713.

https://doi.org/10.1021/acs.jnatprod.4c00740
J. Nat. Prod. 2024, 87, 2695−2708