Specialized metabolites reveal evolutionary history and geographic dispersion of a multilateral symbiosis

Taise T. H. Fukuda,[1,2†] Eric J. N. Helfrich,[1,3,4†] Emily Mevers,[1,5] Weilan G. P. Melo,[2] Ethan B. Van Arnam,[1,6] David R. Andes,[7] Cameron R. Currie,[8] Monica T. Pupo,[2*] Jon Clardy[1*]

[1]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA, 02115, USA.

[2]School of Pharmaceutical Sciences of Ribeirão Preto, University of São Paulo, Ribeirão Preto, SP, 14040-903, Brazil.

[3]Institute for Molecular Bio Science, Goethe University Frankfurt, 60438 Frankfurt, Germany.
[4]LOEWE Center for Translational Biodiversity Genomics (TBG), 60325 Frankfurt, Germany.
[5]Department of Chemistry, Virginia Tech, Blacksburg, VA 24061, USA.
[6]Keck Science Department of Claremont McKenna, Pitzer, and Scripps Colleges, Claremont, CA ,91711, USA

[7]Department of Medicine, University of Wisconsin School of Medicine and Public Health, Madison, WI, 53705, USA.

[8]Department of Bacteriology, University of Wisconsin-Madison, Madison, WI, 53706, USA.

[†]These authors contributed equally to this work.

List of Contents

Table S1. *Pseudonocardia* spp. isolated from fungus-growing ant colonies collected in Brazil. SID number (Strain identification), HID number (Host identification). * genome sequenced strains.

Strain Code (ICBG)	Location	Host	Bacteria (16S)	GPS coordinates
162*	USP campus, SP	Unknown	*Pseudonocardia*	-21.16806, -47.85230
595	Itatiaia, RJ	*Acromyrmex*	*Pseudonocardia*	-22.433433, -44.615317
598	Anavilhanas, AM	*Apterostigma*	*Pseudonocardia*	-2.533528, -60.833222
599	Itatiaia, RJ	Unknown	*Pseudonocardia*	-22.433433, -44.615317
600	Itatiaia, RJ	*Acromyrmex*	*Pseudonocardia*	-22.461967, -44.594833
601*	Anavilhanas, AM	*Apterostigma*	*Pseudonocardia*	-2.532900, -60.832683
602	Anavilhanas, AM	*Apterostigma*	*Pseudonocardia*	-2.523167, -60.825528
604	Anavilhanas, AM	*Acromyrmex*	*Pseudonocardia*	-2.53290, -60.83268
618	Itatiaia, RJ	*Acromyrmex*	*Pseudonocardia*	-22.462000, -44.600000
1019	Anavilhanas, AM	*Apterostigma*	*Pseudonocardia*	-2.523167, -60.825528
1024	Anavilhanas, AM	*Acromyrmex*	*Pseudonocardia*	-2.27047, -61.01814
1025	Anavilhanas, AM	*Trachymyrmex*	*Pseudonocardia*	-2.270472, -61.018139
1030	Itatiaia, RJ	*Apterostigma*	*Pseudonocardia*	-22.449850, -44.611133
1034*	Anavilhanas, AM	*Trachymyrmex*	*Pseudonocardia*	-2.27103, -61.01902
1041	Itatiaia, RJ	Unknown	*Pseudonocardia*	-22.462000, -44.600000
1042	USP campus, SP	*Acromyrmex*	*Pseudonocardia*	-21.16750, -47.84639
1043	Itatiaia, RJ	*Acromyrmex*	*Pseudonocardia*	-22.462000, -44.600000
1050	Anavilhanas, AM	*Trachymyrmex*	*Pseudonocardia*	-2.271361, -61.018333
1052	Itatiaia, RJ	*Cyphomyrmex*	*Pseudonocardia*	-22.462700, -44.593167
1102	Anavilhanas, AM	*Trachymyrmex*	*Pseudonocardia*	-2.27047, -61.01814
1104	Itatiaia, RJ	*Acromyrmex*	*Pseudonocardia*	-22.433433, -44.615317
1111	Anavilhanas, AM	*Trachymyrmex*	*Pseudonocardia*	-2.533933, -60.835017
1122*	Ducke Reserve, AM	Unknown	*Pseudonocardia*	-2.93173, -59.97373
1123	Anavilhanas, AM	*Trachymyrmex*	*Pseudonocardia*	-2.270472, -61.018139
1124	Anavilhanas, AM	*Trachymyrmex*	*Pseudonocardia*	-2.27047, -61.01814
1125	Anavilhanas, AM	*Trachymyrmex*	*Pseudonocardia*	-2.27047, -61.01814
1126	Itatiaia, RJ	*Acromyrmex*	*Pseudonocardia*	-22.433150, -44.615467

1127*	Anavilhanas, AM	*Trachymyrmex*	*Pseudonocardia*	-2.523167, -60.825528
1137	Itatiaia, RJ	*Acromyrmex*	*Pseudonocardia*	-22.433433, -44.615317
1138	Anavilhanas, AM	*Trachymyrmex*	*Pseudonocardia*	-2.27047, -61.01814
1140	Anavilhanas, AM	*Trachymyrmex*	*Pseudonocardia*	-2.271444, -61.018694
1142*	Itatiaia, RJ	*Trachymyrmex*	*Pseudonocardia*	-22.462000, -44.600000
1143	Anavilhanas, AM	*Trachymyrmex*	*Pseudonocardia*	-2.270472, -61.018139
1144	Anavilhanas, AM	*Trachymyrmex*	*Pseudonocardia*	-2.523783, -60.825667
1145	Anavilhanas, AM	*Trachymyrmex*	*Pseudonocardia*	-2.27047, -61.01814
1146	Anavilhanas, AM	*Trachymyrmex*	*Pseudonocardia*	-2.523700, -60.825867
1244	Anavilhanas, AM	*Trachymyrmex*	*Pseudonocardia*	-2.533722, -60.833444
1247	Ducke Reserve, AM	*Trachymyrmex*	*Pseudonocardia*	-2.9317333, -59.9737333
1288*	Anavilhanas, AM	*Trachymyrmex*	*Pseudonocardia*	-2.2704722, -61.0181389
1289	Ducke Reserve, AM	*Trachymyrmex*	*Pseudonocardia*	-2.9317333, -59.9737333
1292	Anavilhanas, AM	*Trachymyrmex*	*Pseudonocardia*	-2.27047, -61.01814
1293*	Anavilhanas, AM	*Trachymyrmex*	*Pseudonocardia*	-2.523667, -60.825883

Table S2. *Pseudonocardia* spp. isolated from fungus-growing ant colonies collected in Panama. SID number (Strain identification).

Strain Code	Location	Host	Bacteria (16S)	GPS coordinates
PA2	PLR-Rio Limbo Plot	*Apterostigma*	*Pseudonocardia*	9.1608056, -79.7449444
PA4	PLR-Rio Limbo Plot	*Apterostigma*	*Pseudonocardia*	9.1608889, -79.7450556
PA5	PLR-Rio Limbo Plot	*Apterostigma*	*Pseudonocardia*	9.1608889, -79.7450556
PA6	PLR-Rio Limbo Plot	*Apterostigma*	*Pseudonocardia*	9.1611111, -79.7450833
PA9	PLR-Rio Limbo Plot	*Apterostigma*	*Pseudonocardia*	9.1622222, -79.7452500
PA11	PLR-Rio Limbo Plot	*Apterostigma*	*Pseudonocardia*	9.1622222, -79.7452500
PA13	PLR-Rio Limbo Plot	*Apterostigma*	*Pseudonocardia*	9.1630278, -79.7451111
PA15	PLR-Rio Limbo Plot	*Apterostigma*	*Pseudonocardia*	9.1630278, -79.7451111
PA17	PLR-Rio Limbo	*Apterostigma*	*Pseudonocardia*	9.1592500, -79.7504444
PA18	PLR-Rio Limbo	*Apterostigma*	*Pseudonocardia*	9.1592500, -79.7504444
PA20	PLR-Rio Limbo	*Apterostigma*	*Pseudonocardia*	9.1592500, -79.7504444
PA22	PLR-Rio Limbo	*Apterostigma*	*Pseudonocardia*	9.1600833, -79.7449444
PA24	PLR-Rio Limbo	*Apterostigma*	*Pseudonocardia*	9.1600833, -79.7449444
PA25	PLR-Rio Limbo	*Apterostigma*	*Pseudonocardia*	9.1600833, -79.7449444
PA27	PLR-Rio Limbo	*Apterostigma*	*Pseudonocardia*	9.1600833, -79.7449444
PA42	PLR-Rio Limbo	*Apterostigma*	*Pseudonocardia*	9.1596944, -79.7445000
PA43	PLR-Rio Limbo	*Apterostigma*	*Pseudonocardia*	9.1595000, -79.7437778
PA45	PLR-[end of road]	*Apterostigma*	*Pseudonocardia*	9.1562222, -79.7418889
PA47	PLR-[end of road]	*Apterostigma*	*Pseudonocardia*	9.1562222, -79.7418889
PA53	PLR-Rio Frijolito	*Apterostigma*	*Pseudonocardia*	9.1495000, -79.7315000
PA55	PLR-Rio Frijolito	*Apterostigma*	*Pseudonocardia*	9.1495000, -79.7315000
PA56	PLR-Rio Frijolito	*Apterostigma*	*Pseudonocardia*	9.1495000, -79.7315000
PA57	PLR-Rio Frijolito	*Apterostigma*	*Pseudonocardia*	9.1495000, -79.7315000
PA59	PLR-Rio Frijolito	*Apterostigma*	*Pseudonocardia*	9.1495000, -79.7315000
PA60	PLR-Juan Grande bridge	*Apterostigma*	*Pseudonocardia*	9.1347500, -79.7224722
PA61	PLR-Juan Grande bridge	*Apterostigma*	*Pseudonocardia*	9.1347500, -79.7224722
PA63	PLR-Juan Grande bridge	*Apterostigma*	*Pseudonocardia*	9.1347500, -79.7224722

PA64	PLR-Juan Grande bridge	*Apterostigma*	*Pseudonocardia*	9.1347500, -79.7224722
PA65	PLR-Juan Grande bridge	*Apterostigma*	*Pseudonocardia*	9.1347500, -79.7224722
PA67	Gamboa, Road to PLR	*Apterostigma*	*Pseudonocardia*	9.1200278, -79.7064167
PA69	Gamboa, Road to PLR	*Apterostigma*	*Pseudonocardia*	9.1200278, -79.7064167
PA71	Gamboa, Road to PLR	*Apterostigma*	*Pseudonocardia*	9.1200278, -79.7064167
PA72	Gamboa, Road to PLR	*Apterostigma*	*Pseudonocardia*	9.1200278, -79.7064167
PA76	Gamboa, 140 Jadwin Ave "Bamboo Hill"	*Apterostigma*	*Pseudonocardia*	9.1196944, -79.6962500
PA78	Gamboa, 140 Jadwin Ave "Bamboo Hill"	*Apterostigma*	*Pseudonocardia*	9.1196944, -79.6962500
PA86	Gamboa, 140 Jadwin Ave "Bamboo Hill"	*Apterostigma*	*Pseudonocardia*	9.1196944, -79.6962500
PA87	Gamboa, 176 Williamson Road	*Apterostigma*	*Pseudonocardia*	9.1152778, -79.6978333
PA89	Gamboa, 176 Williamson Road	*Apterostigma*	*Pseudonocardia*	9.1144722, -79.6978056
PA91	Gamboa, 176 Williamson Road	*Apterostigma*	*Pseudonocardia*	9.1144722, -79.6978056
PA93	Gamboa, 176 Williamson Road	*Apterostigma*	*Pseudonocardia*	9.1144722, -79.6978056
PA95	Gamboa, 176 Williamson Road	*Apterostigma*	*Pseudonocardia*	9.1144722, -79.6978056
PA97	Gamboa, 176 Williamson Road	*Apterostigma*	*Pseudonocardia*	9.1144722, -79.6978056
PA99	Gamboa, 176 Williamson Road	*Apterostigma*	*Pseudonocardia*	9.1144722, -79.6978056
PA140	Plantation Road	*Apterostigma*	*Pseudonocardia*	9.0783767, -79.6593883
PA141	Plantation Road	*Apterostigma*	*Pseudonocardia*	9.0783767, -79.6593883
PA144	Ancon Hill	*Apterostigma*	*Pseudonocardia*	8.9612500, -79.5509167
PA146	PLR-Rio Limbo	*Apterostigma*	*Pseudonocardia*	9.1596944, -79.7445000

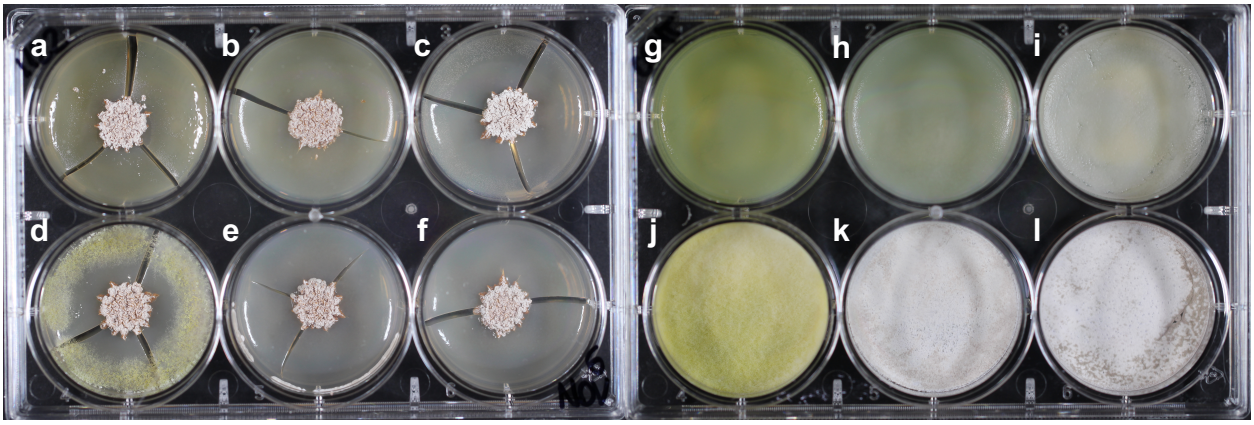

Figure. S1. Bipartite antagonism assay between *Pseudonocardia* sp. ICBG1122 (center of wells) and *Escherichia coli* **(a)**, *Pseudomonas aeruginosa* **(b)**, *Candida albicans* **(c)**, *Trichoderma reesei* **(d)**, *Streptomyces* sp. I8 **(e)** and *Streptomyces* sp. I17 **(f)**. Negative controls are shown (g-l). *Pseudonocardia* sp. ICBG1122 shows inhibitory activity against *C. albicans*, *T. reesei*, *Streptomyces* sp. I8 and *Streptomyces* sp. I17.

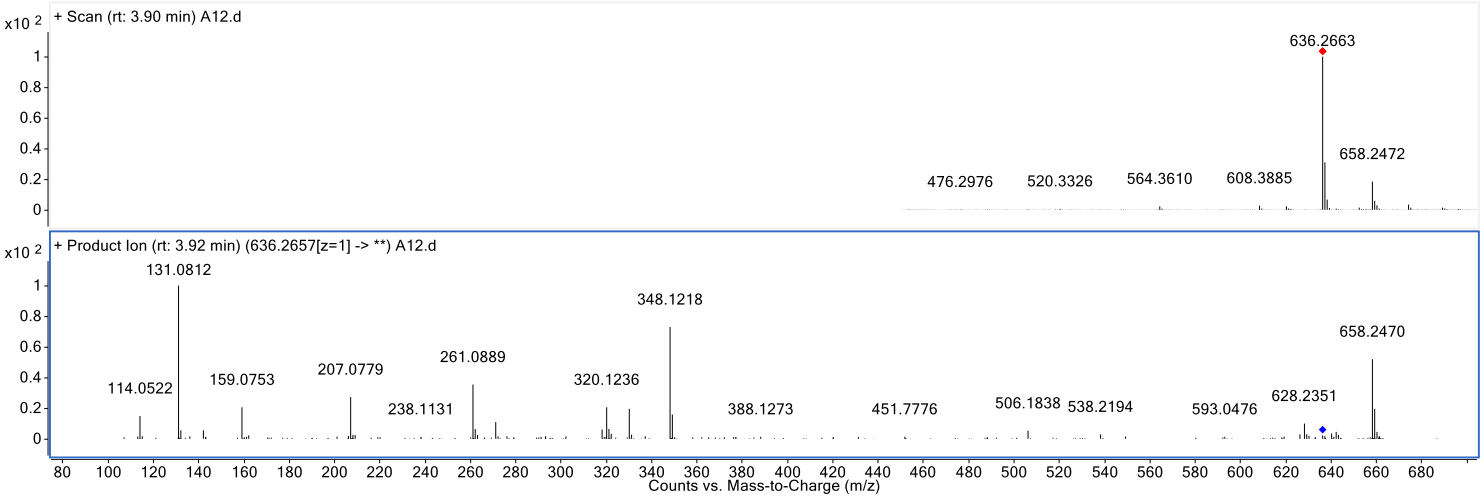

Figure S2. Parent mass and characteristic MS2 fragmentation pattern of attinimicin. HR-ESI-MS $[M+H]^+$ of attinimicin (top); HR-ESI-MS/MS $[M+H]^+$ of attinimicin (bottom).

Table S3. Attinimicin detection (molecular formula: $C_{27}H_{37}N_7O_{11}$; theoretical *m/z* 636.2624 $[M+H]^+$)

Strain Code (ICBG)	Genome sequenced	Attinimicin detected				
		BGC by PCR	LCMS			
			Detected	Retention time (min)	Mass	Error (ppm)
162	yes	yes	no			
598	no	yes	yes	3.82	636.2617	-1.1
599	no	yes	no			
600	no	yes	yes	3.96	636.2646	3.5
601	yes	yes	yes	3.97	636.2653	4.6
602	no	yes	yes	3.99	636.2650	4.1
604	no	yes	yes	3.99	636.2639	2.4
1019	no	yes	yes	3.98	636.2662	6.0
1024	no	no	no			
1025	no	no	no			
1030	no	yes	no			
1034	yes	yes	yes	3.97	636.2659	5.5
1042	no	yes	no			
1050	no	yes	no			
1052	no	yes	no			
1102	no	yes	no			
1111	no	yes	no			
1122	yes	yes	no			
1123	no	yes	no			
1124	no	yes	yes	3.84	636.2666	6.6
1125	no	yes	no			
1126	no	yes	no			
1127	yes	yes	no			
1138	no	no	no			
1140	no	yes	no			
1142	yes	yes	yes	3.96	636.2639	2.4
1143	no	yes	no			
1144	no	no	no			
1145	no	yes	yes	3.94	636.2651	4.2
1146	no	yes	no			
1244	no	no	no			
1247	no	yes	no			
1288	yes	yes	no			
1289	no	yes	no			
1292	no	yes	no			
1293	yes	yes	no			

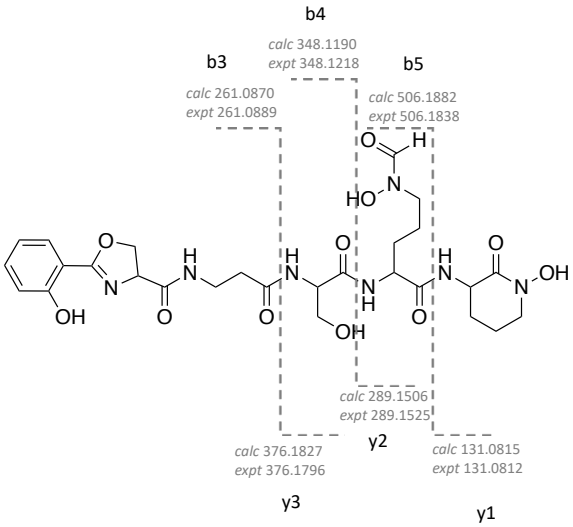

Figure S3. MS/MS fragmentation of attinimicin. Calculated and experimental *m/z* values for observed fragments.

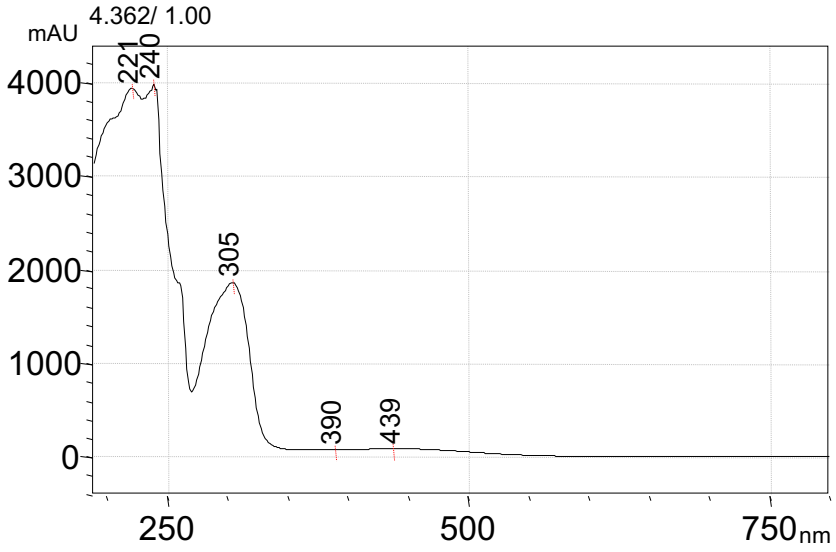

Figure S4. UV spectrum of attinimicin in acetonitrile:water 1:2.

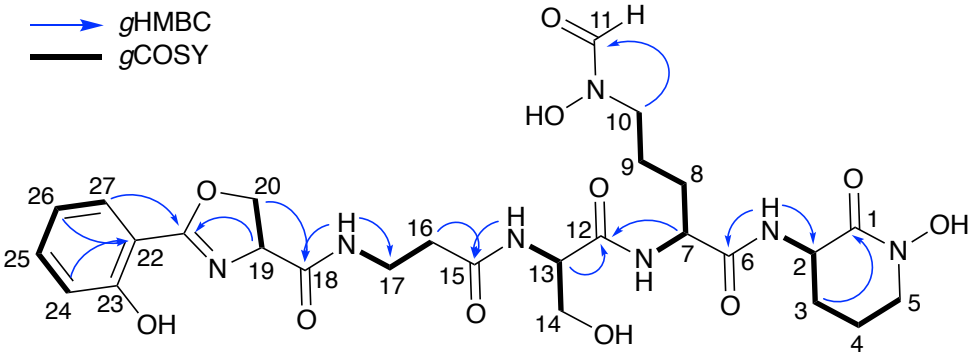

Figure S5. Key NMR correlations of attinimicin.

Table S4. NMR spectroscopic data of attinimicin in dimethylsulfoxide-d_6.

	δ_C	δ_H	Multiplicity (Hz)	COSY	HMBC
1	174.0				
2	49.5	4.36	m	2-NH, 3	1, 3, 4
2-NH		8.21	d (8.4)	2	2, 6
3a	29.9	1.48	m	2, 4	4
3b		1.65	m	2	1, 4
4	23.3	1.59	m	3a, 4	3
5	46.3	3.48	m	4	3
6	171.1				
7	51.1	4.30	m	7-NH, 8	6, 8, 12
7-NH		8.23	d (8.7)	7	7
8a	27.5	1.65	m	7, 9	7
8b		1.88	m	7	7, 9
9	20.3	1.89	m	8a, 10	7, 8
10	51.0	3.45	m	9	8, 9, 11
11	164.6	8.00	s		10
12	168.4				
13	50.8	4.96	m	13-NH, 14	12, 14
13-NH		7.64	dd (7.9, 1.5)	13	15
14a	61.1	3.52	m	13	12
14b		3.63	dd (10.7, 3.7)	13	12
15	169.9				
16	35.0	2.37	t (6.9)	17	15, 17
17	35.7	3.26	m	16, 17-NH	15, 16
17-NH		8.33	t (5.6)	17	17, 18
18	169.6				
19	67.4	4.92	dd (10.0, 7.9)	20a, 20b	18, 20, 21
20a	69.4	4.52	t (8.1)	19	18, 19, 21
20b		4.61	dd (8.3, 10.0)	19	18, 19, 21
21	165.8				
22	110.0				
23	159.1				
24	116.6	7.00	d (7.5)	25	22, 26
25	134.0	7.46	dt (7.5, 1.6)	24, 26	27
26	119.1	6.94	t (7.6)	25	22, 24, 25, 27
27	128.1	7.63	dd (7.6, 1.6)	26	25, 21

.

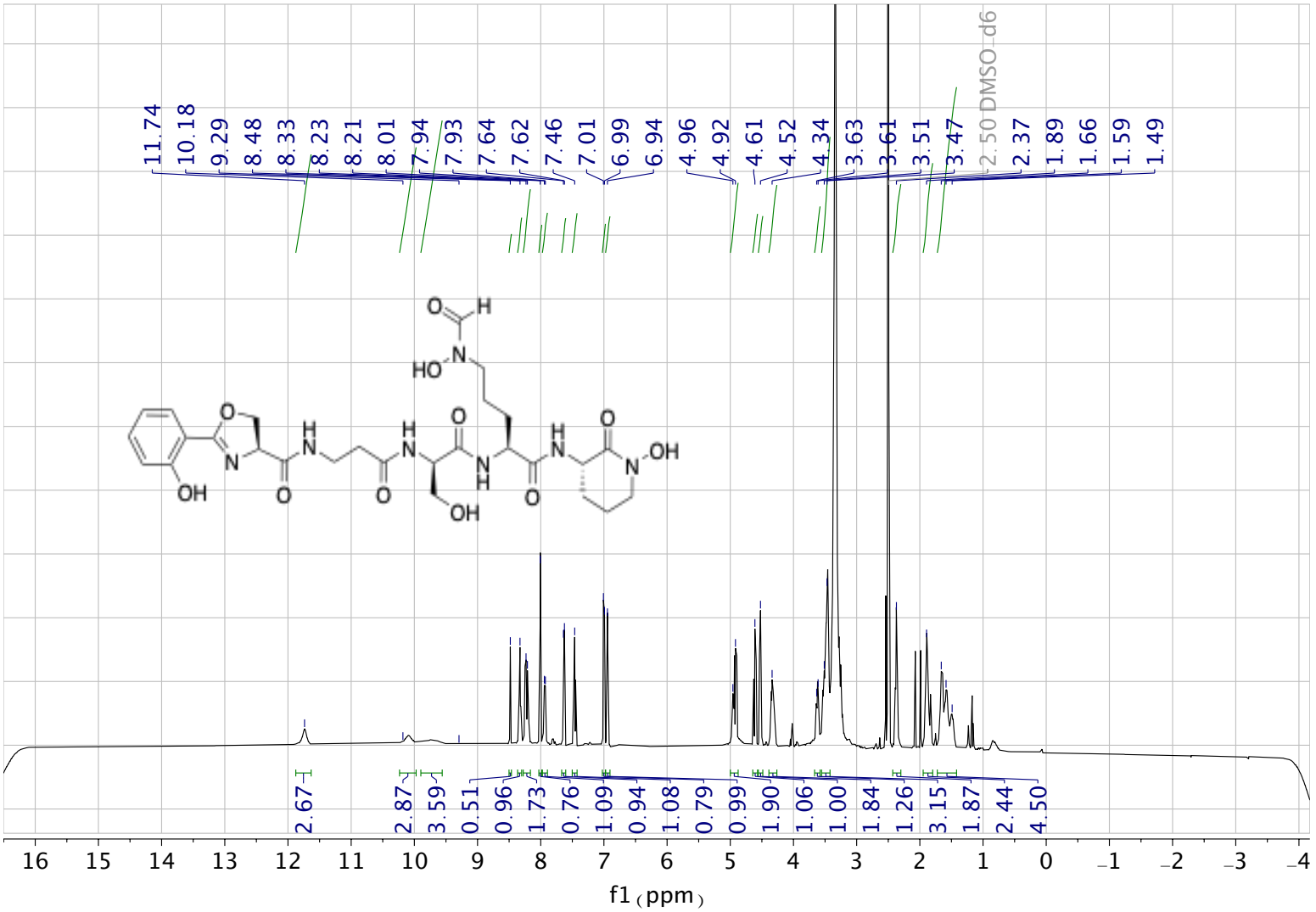

Figure S6. ^{1}H NMR spectrum of attinimicin (**3**) in dimethylsulfoxide-*d*$_6$, 600 MHz.

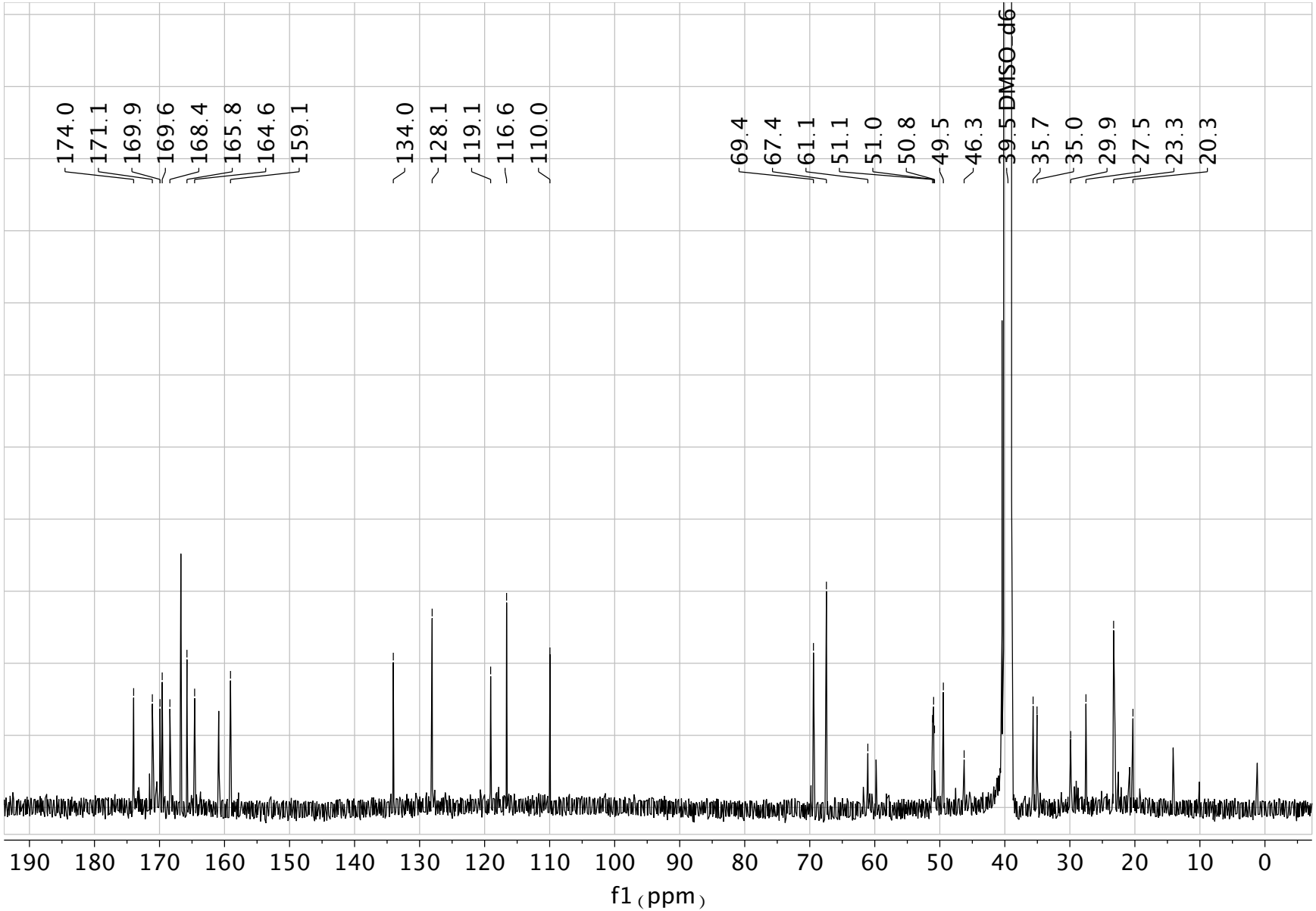

Figure S7. ^{13}C NMR spectrum of attinimicin (**3**) in dimethylsulfoxide-*d*$_6$, 150 MHz.

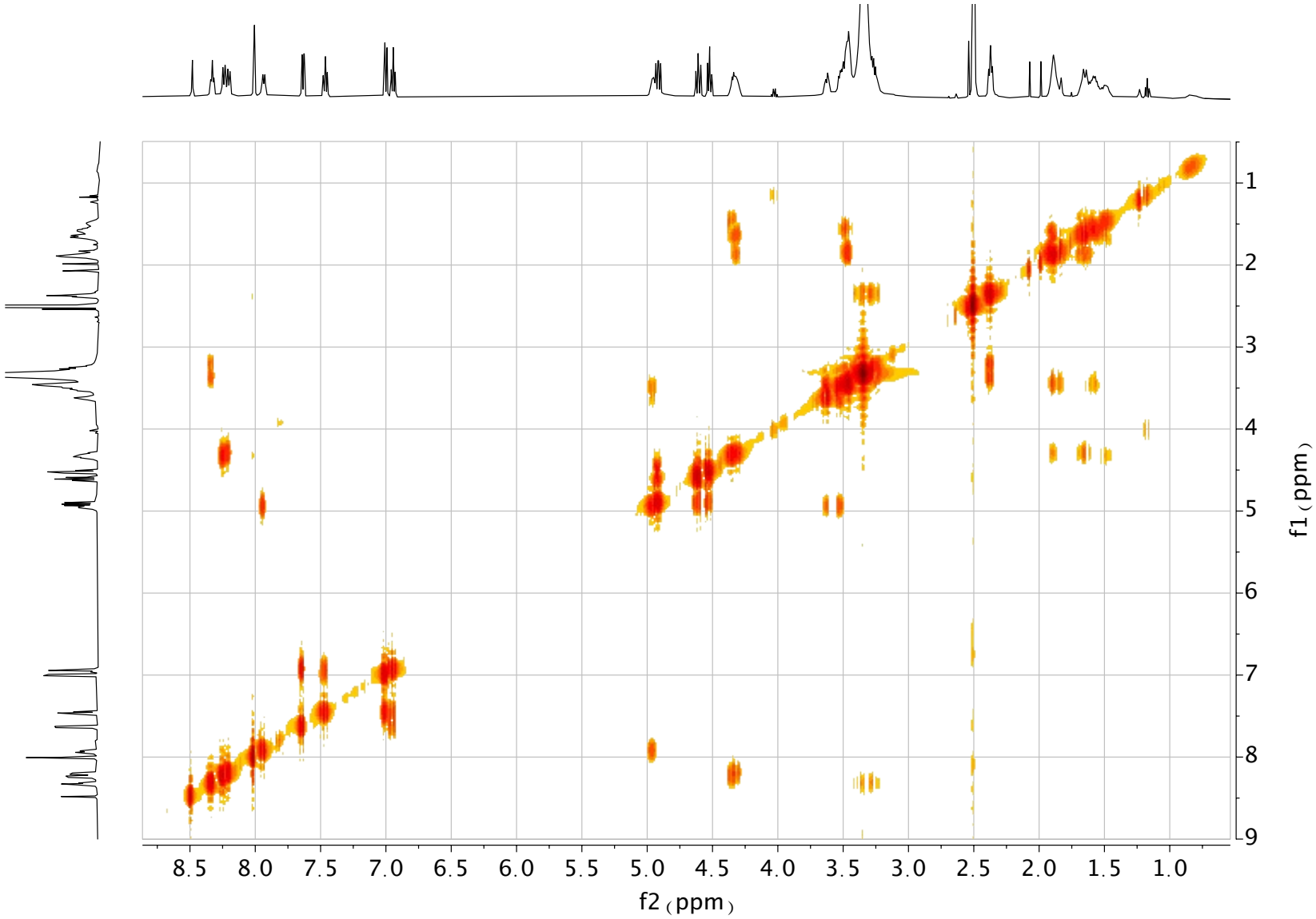

Figure S8. ^{1}H,^{1}H-COSY spectrum of attinimicin (**3**) in dimethylsulfoxide-*d*$_6$.

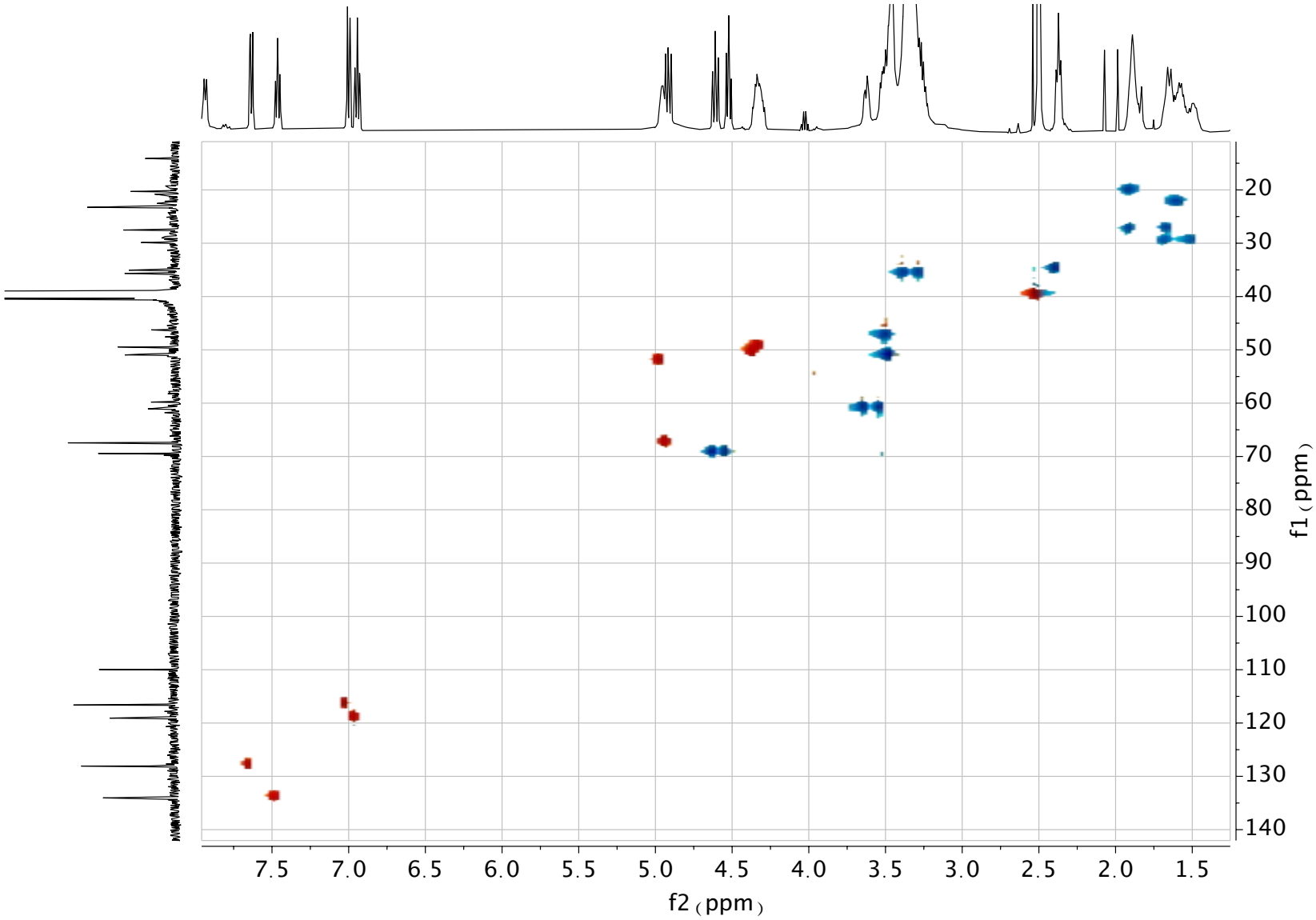

Figure S9. PS-HSQC spectrum of attinimicin (**3**) in dimethylsulfoxide-d_6, 600 MHz. Methine and methyl resonances are shown in red; methylenes are shown in blue.

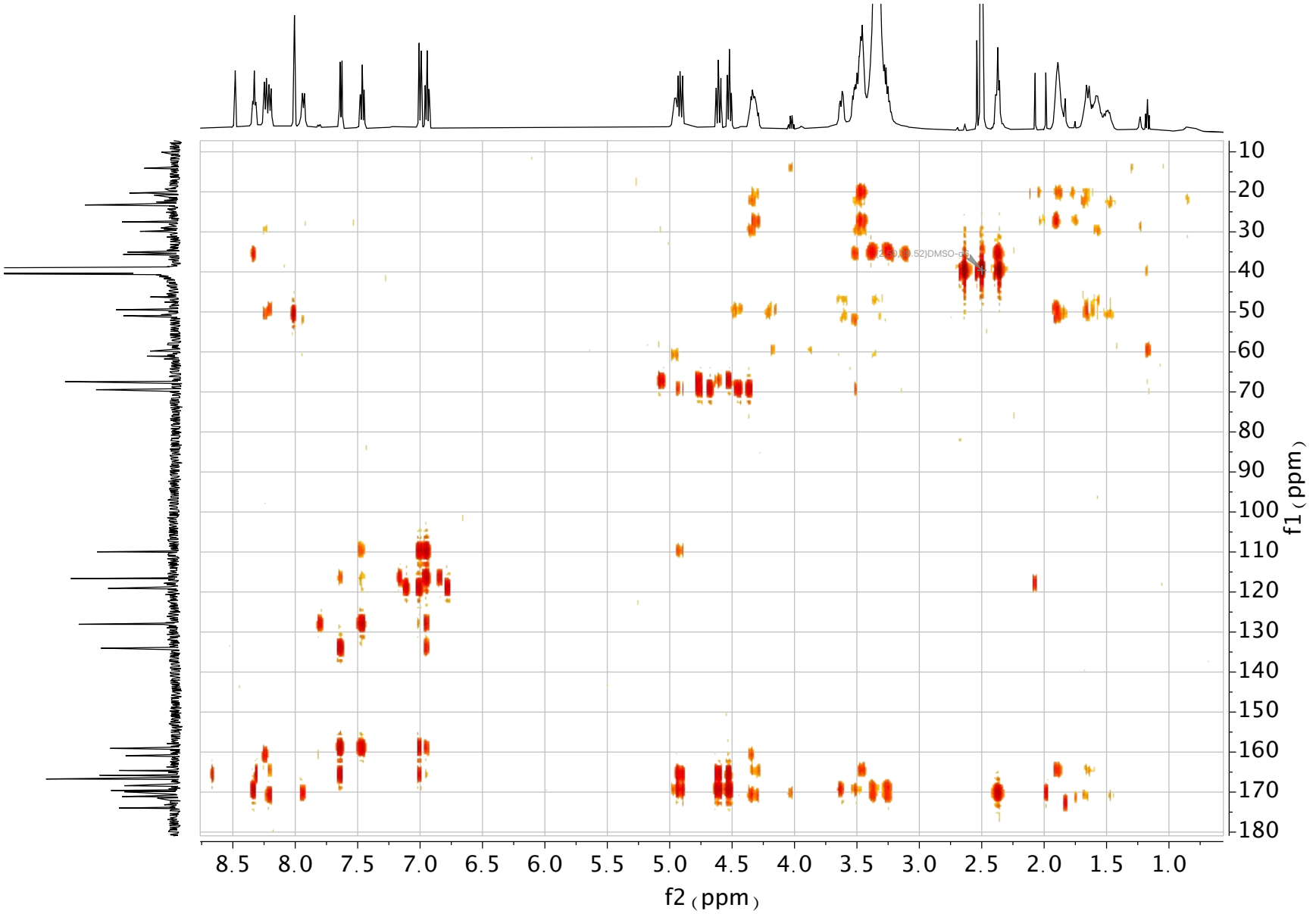

Figure S10. HMBC spectrum of attinimicin (**3**) in dimethylsulfoxide-d_6, 600 MHz.

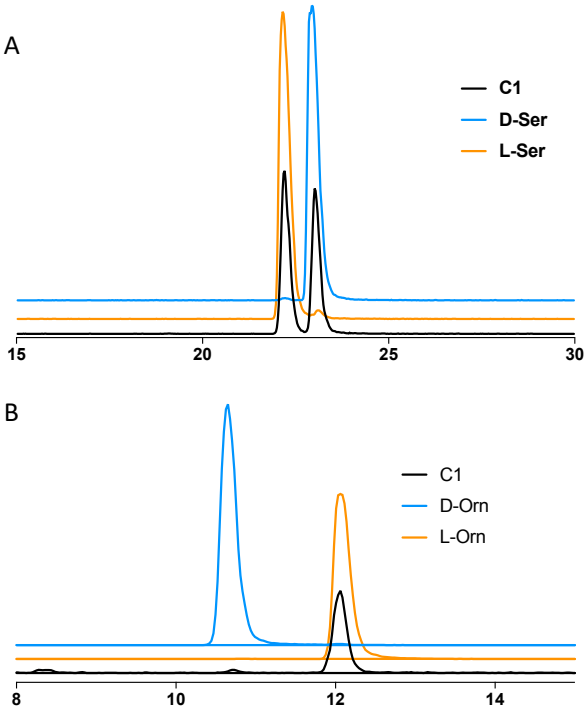

Figure S11. HPLC chromatograms of Marfey's analysis of amino acids after hydrolysis of **3**. (**A**) EIC of the ion *m/z* 358.10 ($[M+H]^+$), positive mode ESI, corresponding to serine residues. (**B**) EIC of the ion *m/z* 385.15 ($[M+H]^+$), positive mode ESI, corresponding to ornithine residues.

Table S5. Marfey's analysis of attinimicin.

Amino acid	*m/z* $[M+H]^+$	Retention time (min): Marfey product of L-amino acid standard	Retention time (min): Marfey product of D-amino acid standard	Retention time (min): Marfey product of amino acid from attinimicin hydrolysis
Serine	358.10	22.16	22.95	22.20/23.02*
Ornithine	385.15	12.06	10.65	12.06

* indicates the presence of both D and L serine in the structure of attinimicin, as seen in Supplementary Figure 10.

Table S6. Quality assessment of genome assemblies

	ICBG1293	ICBG 601	ICBG 1142	ICBG 1122	ICBG 1034	ICBG 162	ICBG 1127	ICBG 1288
# contigs (>= 0 bp)	86	280	54	83	32	32	85	14
# contigs (>= 1000 bp)	86	280	54	83	32	32	85	14
# contigs (>= 5000 bp)	79	220	48	49	32	32	78	14
# contigs (>= 10000 bp)	64	173	42	43	25	32	72	13
# contigs (>= 25000 bp)	29	99	25	26	10	13	60	12
# contigs (>= 50000 bp)	26	45	13	21	8	9	37	11
Total length (>= 0 bp)	7959327	13311475	7434153	6899411	6680590	7108174	7491467	7163992
Total length (>= 1000 bp)	7959327	13311475	7434153	6899411	6680590	7108174	7491467	7163992
Total length (>= 5000 bp)	7937130	13130099	7420899	6807696	6680590	7108174	7470679	7163992
Total length (>= 10000 bp)	7817963	12796468	7372246	6765367	6615032	7108174	7419395	7155839
Total length (>= 25000 bp)	7318660	11575456	7110298	6476681	6416272	6817729	7209024	7145828
Total length (>= 50000 bp)	7209581	9697141	6688353	6320833	6360557	6683277	6353194	7113635
# contigs	86	280	54	83	32	32	85	14
Largest contig	939100	3455861	2163479	1193048	2850860	2274010	947933	3135515
Total length	7959327	13311475	7434153	6899411	6680590	7108174	7491467	7163992
GC (%)	73.17	73.52	72.59	74.07	74.27	74.42	73.22	73.71
N50	418257	143282	833121	413978	1138390	1743544	198270	1055934
N75	272220	43919	412174	197656	1094868	516861	74417	625011
L50	7	8	3	5	2	2	11	2
L75	13	52	6	10	3	4	26	4
# N's per 100 kbp	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

# contigs:	total number of contigs in the assembly.
Largest contig:	length of the longest contig.
Total length:	total number of bases.
GC (%):	total number of G and C nucleotides, divided by the total number of bases.
N50:	length for which the collection of all contigs of that length or longer covers at least half an assembly.
N75:	defined similarly to N50 but with 75% instead of 50%.
L50 and L75:	number of contigs equal to or longer than N50 and N75, respectively.
# N's per 100 kbp:	average number of uncalled bases (N's) per 100000 bases.

Table S7. Predicted proteins of the attinimicin BGC

Protein	Size (aa)	Proposed function	Sequence similarity, Organism	Identity/ Similarity (%)	GenBank accession no.
AttT1	347	iron compound ABC transporter, periplasmic	iron-siderophore ABC transporter substrate-binding protein, *Pseudonocardia* sp. Ae505_Ps2	92%/96%	WP_084786367.1
AttA	88	ornithine N-monooxygenase	L-ornithine N (5)-oxygenase family protein, *Pseudonocardia ammonioxydans*	49%/63%	WP_093336896.1
AttB	69	MbtH family protein	MbtH, *Pseudonocardia* sp. EV170527-09	91%/94%	WP_149665922.1
AttC	145	Aspartate decarboxylase	Aspartate 1-decarboxylase, *Pseudonocardia* sp. Ae505_Ps2	97%/99%	OLM14204.1
AttD	1151	NRPS: PCP, C, A, ACP	NRPS, *Pseudonocardia* sp. HH130629-09	94%/96%	WP_060712263.1
AttE	4742	NRPS: C, A, PCP, C, A, PCP, C, A, PCP, C, A, PCP, E	NRPS, *Pseudonocardia* sp. Ae505_Ps2	91%/93%	WP_075553901.1
AttT2	452	major facilitator transporter	MFS transporter, *Pseudonocardia* sp. EV170527-09	91%/91%	WP_149665670.1
AttF	227	methionyl-tRNA formyltransferase	methionyl-tRNA formyltransferase, *Pseudonocardia sediminis*	54%/66%	RZT86011.1
AttG	452	salicylate synthase	salicylate synthase, *Pseudonocardia* sp. EV170527-09	92%/93%	WP_149665668.1
AttH	546	NRPS: A	AMP-binding protein, *Pseudonocardia* sp. HH130629-09	95%/96%	WP_060712259.1
AttI	270	hydrolase	alpha/beta fold hydrolase, *Pseudonocardia* sp. AL041005-10	95%/97%	WP_062394477.1
AttT3	384	iron compound ABC transporter, periplasmic	ABC transporter substrate-binding protein, *Pseudonocardia* sp. EV170527-09	88%/92%	WP_149665665.1

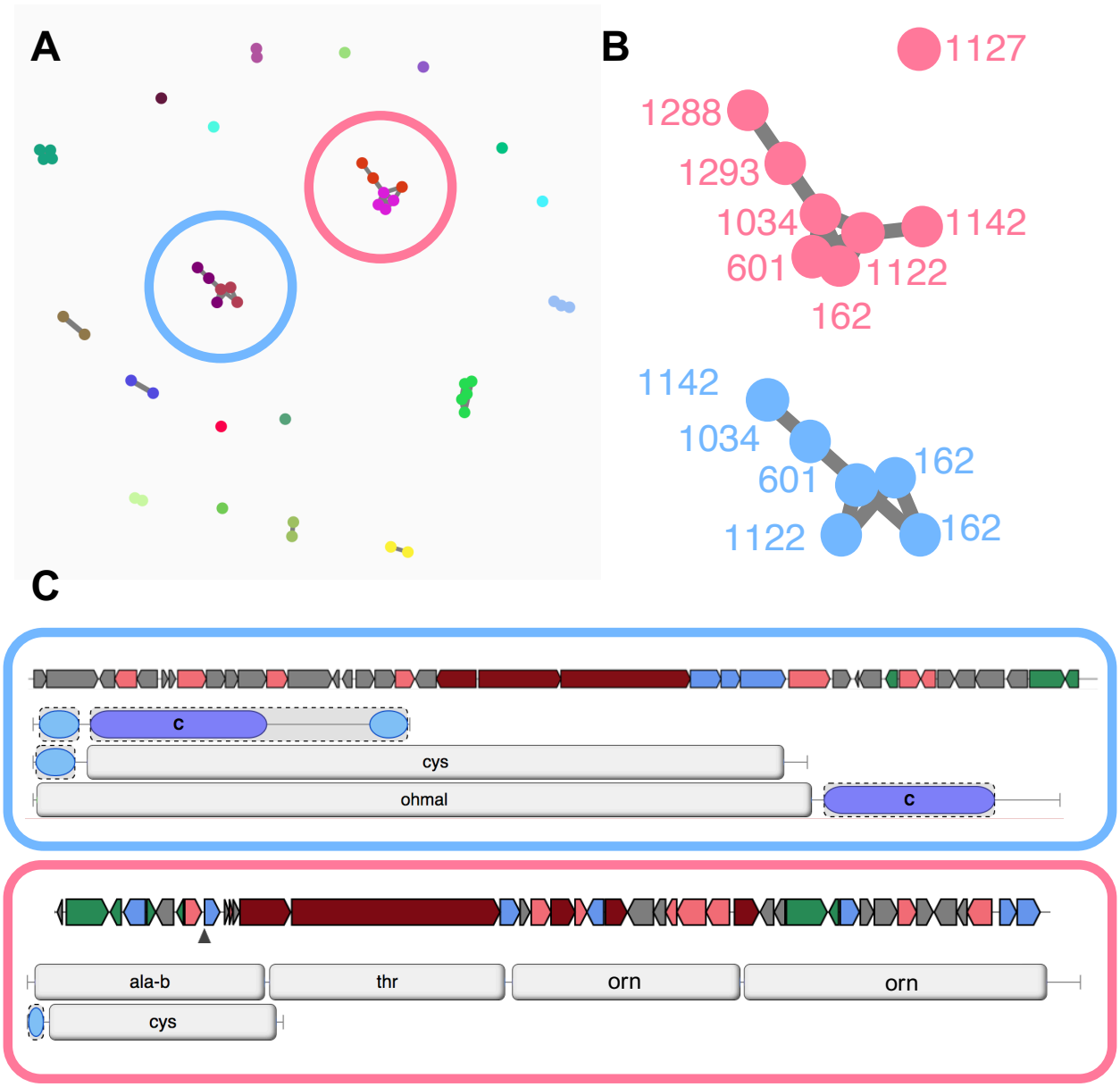

Figure S12. BiG-SCAPE analysis of sequenced attinimicin producers. (**A**) BiG-SCAPE network containing BGCs identified as NRPS, RiPP or hybrid. Two BGCs were identified to be conserved in several producers, the putative *att* BGC and a NRPS-PKS hybrid. (**B**) BiG-SCAPE analysis of attinimicin producers and the MIBiG databsase. Displayed are BGC families identified in attinimicin producers that contain more than five nodes (NRPS, RiPPs or hybrid BGCs). Numbers correspond to *Pseudonocardia* spp. ICBG stains. Node colors refer to clusters in A. Note that the NRPS-PKS hybrid is duplicated in *Pseudonocardia* spp. ICBG162 but absent in other attinimicin producers. The *att* BGC did not cluster with any NRPS responsible for the biosynthesis of attinimicin-like peptides in the MIBiG database as a likely result of the overall architecture of the biosynthetic island in which the *att* BGC is located (even at the lowest cut-off). (**C**) BGC architecture of clusters present in more than five *Pseudonocardia* spp. strains. Box colors represent clusters in A and B.

Table S8. *In-silico* adenylation domain substrate specificity predictions

Domain	SANDPUMA prediction	Phylogenetic prediction	Incorporated substrate
AttD-A	Cys	Cys	Ser
AttE-A_1	β-Ala	β-Ala	β-Ala
AttE-A_2	Thr	Thr	Ser
AttE-A_3	Not predicted	hOrn	hOrn
AttE-A_4	Not predicted	hOrn	hOrn

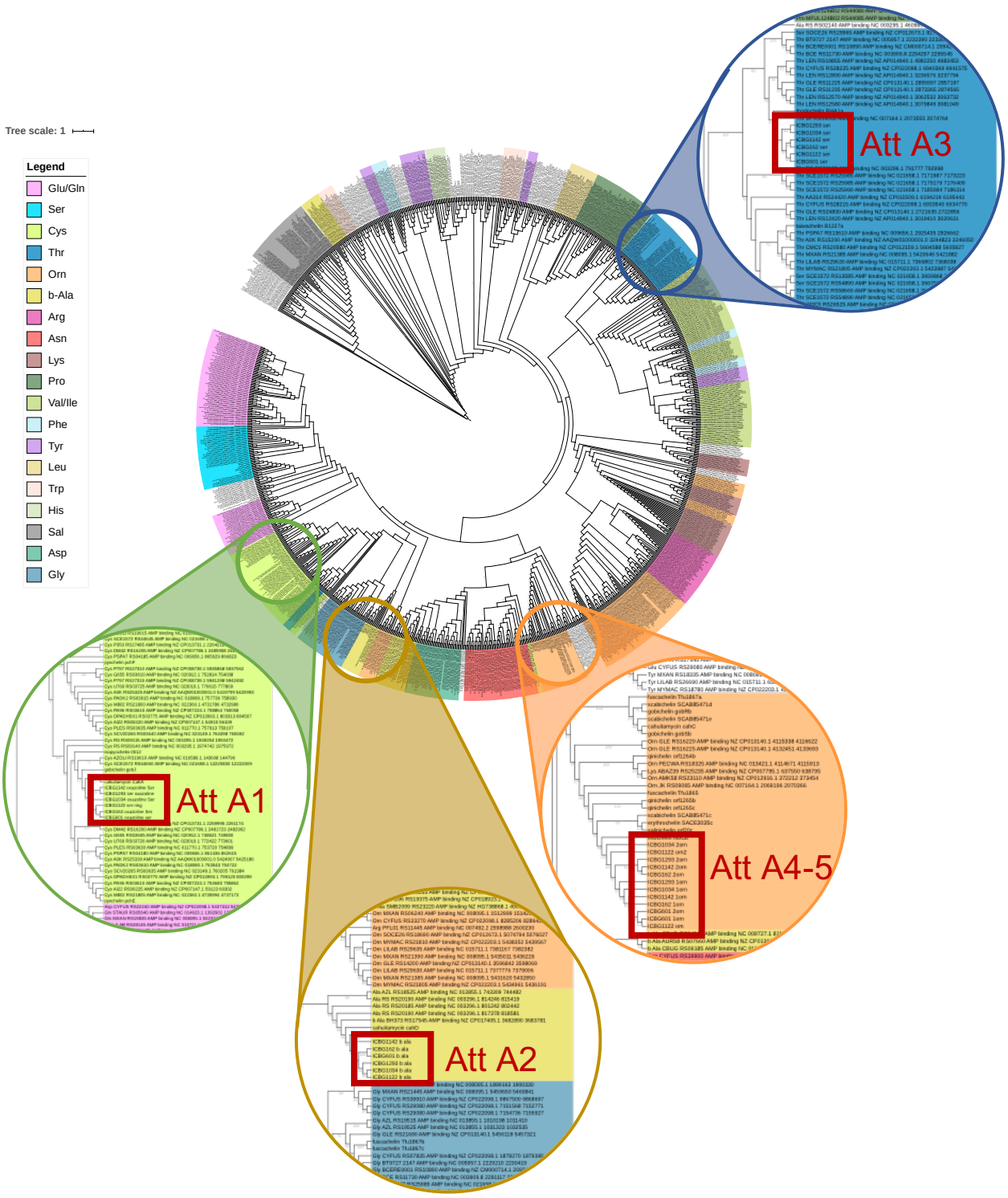

Figure S13. Phylogenetic tree of A domains. Different colors represent different substrate specificities. The phylogeny was reconstructed using IQ-TREE[1], employing the LG+F+R10 model of amino acid substitution. Support values are based on 1000 bootstrap replicates. Rectangular views of branches containing attinimicin A domains are shown. Att-derived A domains are highlighted in red. It is suggested that a duplication event resulted in the A domains of modules 4 and 5.

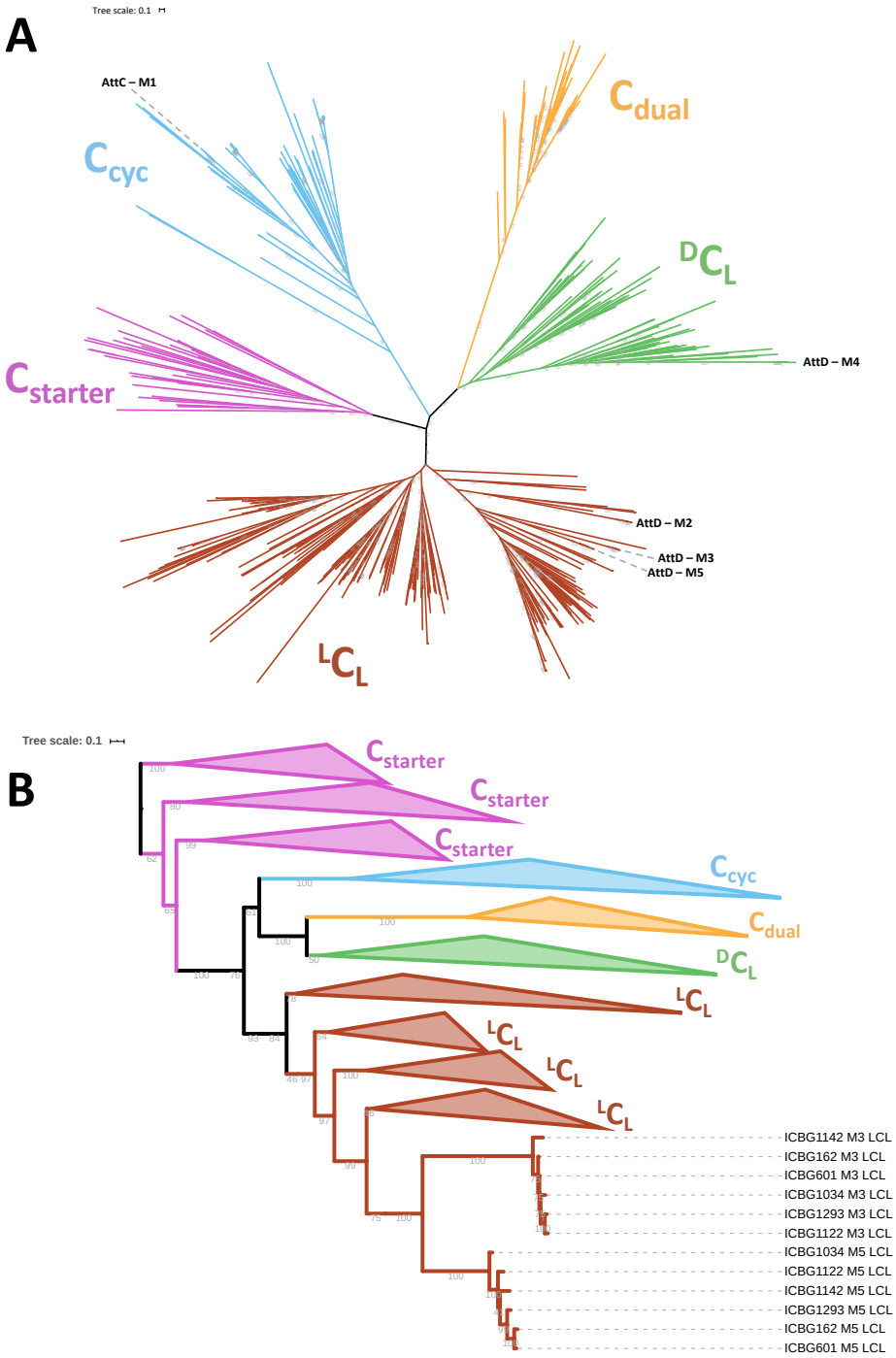

Figure S14. Phylogeny of C domains. $C_{starter}$, condensation domain in the first module of peptide biosynthesis; Cyc, cyclization domain; $^{D}C_{L}$, condensation domain that catalyzes the formation of the peptide bond between the terminal D-amino acid of the growing peptide chain and an L-amino acid to be incorporated; $^{L}C_{L}$, condensation domains condense two L-amino acids. The phylogeny was reconstructed using IQ-TREE[1], employing the LG+F+R8 model of amino acid substitution. Support values are based on 1000 bootstrap replicates; **(A)** Unrooted phylogenetic tree of C domains. Attinimicin C domains are shown; **(B)** Rectangular view of the tree. It is suggested that a duplication event resulted in the C domains of modules 3 and 5.

Table S9. Samples collected in Itatiaia National Park for *in situ* detection of attinimicin. Sample and nest codes, locations and GPS coordinates are included on this table. Attinimicin was detected in samples marked with an asterisk.

Code	Sample	Host	Nest code	GPS coordinates
TF01	Ant	*Acromyrmex*	GPC02-3	-22.4516111, -44.6077500
TF02*	Ant	*Acromyrmex*	TF03-6	-22.4468056, -44.6106111
TF03*	Ant	*Acromyrmex*	TF03-3	-22.4458056, -44.6111389
TF04*	Queen	*Acromyrmex*	TF03-6	-22.446805, -44.6106111
TF05*	Ant	*Acromyrmex*	CFP02-1	-22.4508889, -44.6103056
TF06	Garden	*Acromyrmex*	CFP02-1	-22.4508889, -44.6103056
TF07	Garden	*Acromyrmex*	WM03-1	-22.4409722, -44.6095833
TF08*	Garden	*Acromyrmex*	GPC02-1	-22.4502222, -44.6111667
TF09	Garden	*Acromyrmex*	GPC02-3	-22.4516111, -44.6077500
TF10	Garden	*Acromyrmex*	TF03-3	-22.4458056, -44.6111389
TF11	Garden	*Acromyrmex*	GPC02-2	-22.4508889, -44.6103056
TF12*	Garden	*Acromyrmex*	WM02-2	-22.4508889, -44.6103056
TF13	Garden	*Acromyrmex*	TF03-1	-22.4409722, -44.6095833
TF14*	Garden	*Acromyrmex*	TF03-2	-22.4409722, -44.6095833
TF15	Garden	*Acromyrmex*	CFP03-3	-22.4409722, -44.6095833
TF16	Garden	*Acromyrmex*	WM03-6	-22.4468056, -44.6106111
TF17	Garden	*Acromyrmex*	TF02-2	-22.4514444, -44.6125556
TF18	Garden	*Acromyrmex*	WM03-2	-22.4409722, -44.6095833
TF19	Garden	*Acromyrmex*	MTP03-1	-22.4458056, -44.6111389
TF20	Ant	*Acromyrmex*	TF02-2	-22.4514444, -44.6125556
TF21*	Ant	*Acromyrmex*	CFP03-3	-22.4409722, -44.6095833
TF22*	Ant	*Acromyrmex*	WM03-6	-22.4468056, -44.6106111
TF23*	Ant	*Acromyrmex*	TF03-2	-22.4409722, -44.6095833
TF24	Ant	*Acromyrmex*	GPC02-1	-22.4502222, -44.6111667
TF25	Ant	*Acromyrmex*	WM03-2	-22.4409722, -44.6095833
TF26	Ant	*Acromyrmex*	MTP03-1	-22.4458056, -44.6111389
TF27*	Ant	*Acromyrmex*	WM03-1	-22.4409722, -44.6095833
TF28*	Ant	*Acromyrmex*	WM02-2	-22.4508889, -44.6103056
TF29	Ant	*Acromyrmex*	GPC02-2	-22.4508889, -44.6103056
TF30	Ant	*Acromyrmex*	TF01-3	-22.4516111, -44.6077500
TF31	Ant	*Acromyrmex*	TF03-1	-22.4409722, -44.6095833

Table S10. pFe comparison of attinimicin and the siderophores used in medicine to treat acute iron poisoning[2]

Compound name	pFe
Attinimicin	18.57 ± 0.81
Desferrioxamine	26.5
Deferiprone	20.70
Deferasirox	23.18

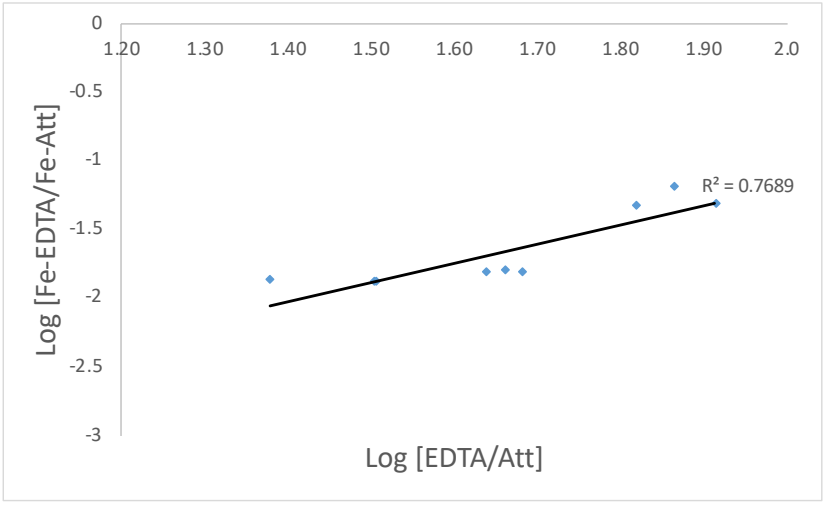

Figure S15. Plots of log [EDTA/Att] against log [(Fe-EDTA)/(Fe-Att)]. Experiments were performed in triplicates.

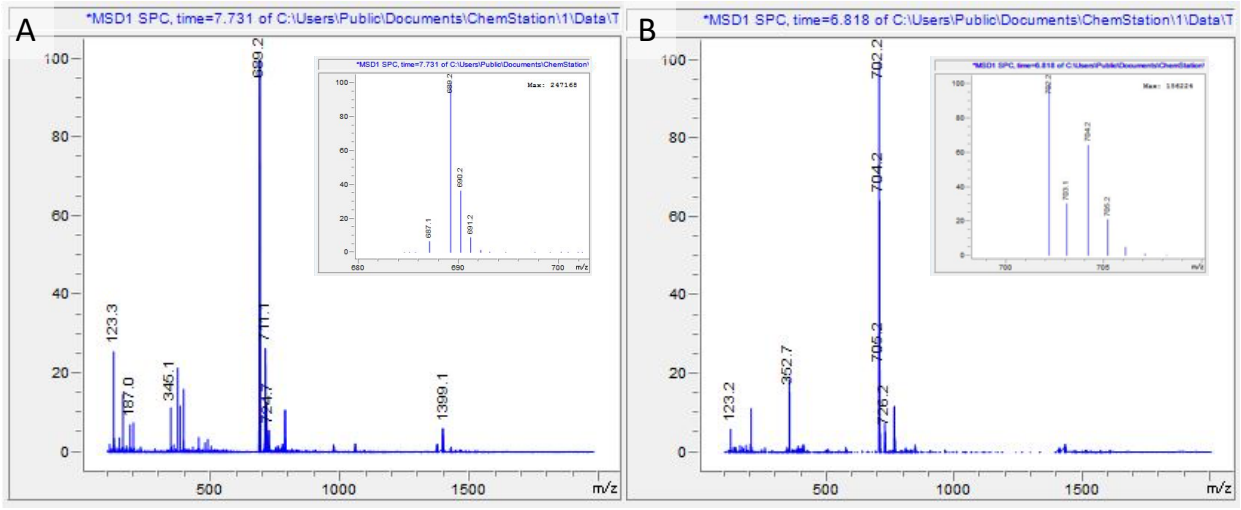

Figure S16. MS spectra of attinimicin complexed with metals. **(A)** Low-resolution MS spectrum of Fe-attinimicin. **(B)** Low-resolution MS spectrum of Ga-attinimicin.

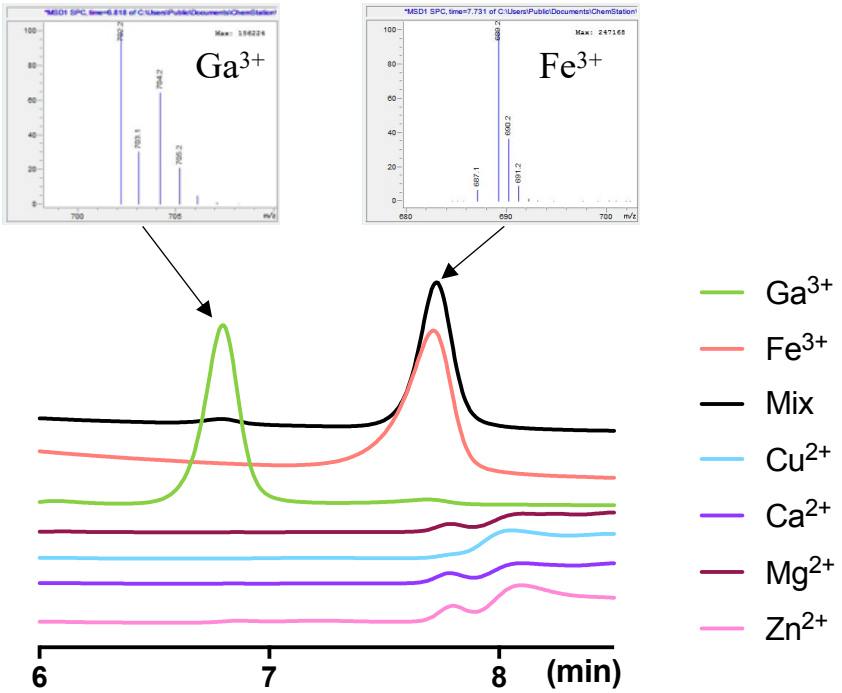

Figure S17. HPLC chromatogram (λ = 254 nm) of attinimicin when incubated with different metals**:** Ga^{3+}, Fe^{3+}, Ca^{2+}, Zn^{2+}, Cu^{2+}, Mg^{2+} and a mixture containing all metals in equimolar ratio. The chromatogram of the mixture indicates that attinimicin exhibits selectivity toward Fe^{3+}. Non-complexed attinimicin is not shown above.

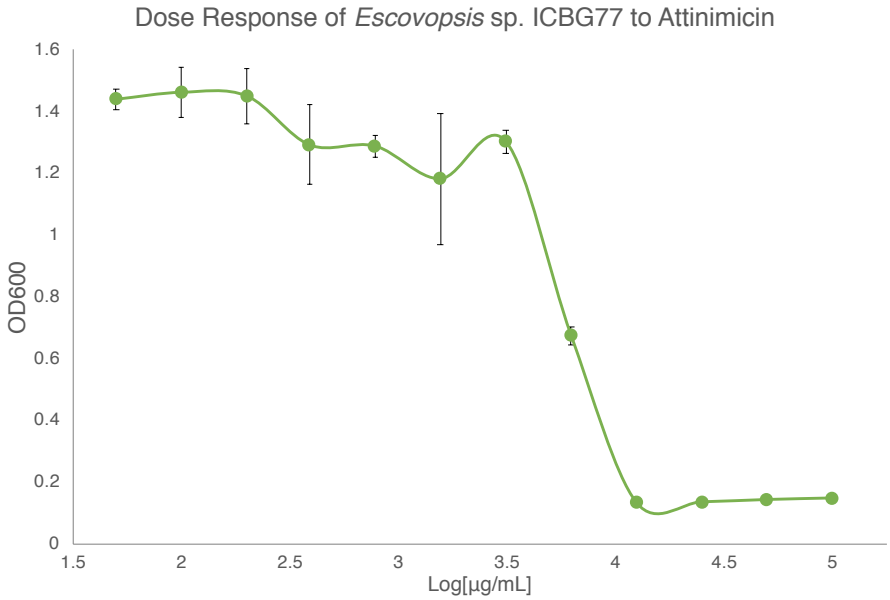

Figure S18. Dose response of *Escovopsis* sp. ICBG77 to attinimicin.

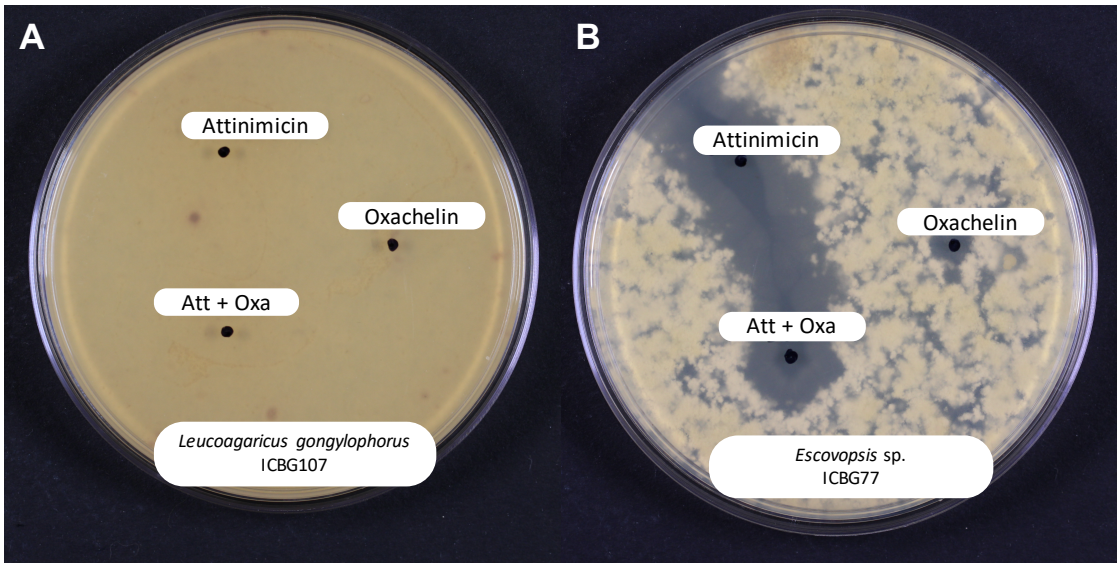

Figure S19. Inhibitory spot-on-lawn assay of attinimicin and/or oxachelin. (**A**) Attinimicin and/or oxachelin (in our laboratory's pure compound library resulting from an unpublished project) showed no activity against the cultivar *Leucoagaricus gongylophorus* ICBG107 at concentrations up to 100 µg; (**B**) Inhibition of *Escovopsis* sp. ICBG77 by attinimicin and/or oxachelin (100 µg). Attinimicin showed more potent activity against the fungal pathogen than oxachelin.

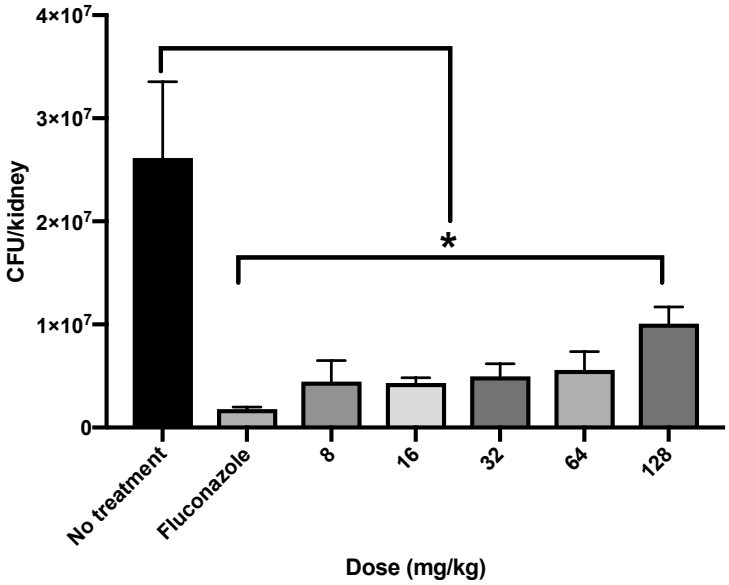

Figure S20. *In vivo* activity of attinimicin against *Candida albicans*. Fluconazole (4 mg/kg, non-concurrent) was used as a positive control. Asterisk indicates significance.

Table S11. Primers used in this study

Primer	Sequences 5' to 3'	Description
attH2F	5'-TCGTCGSCGAAGSCGAKCACCAG-3'	Forward primer for first amplification of *attH*
attH2R	5'-CCGTSGTCCTSGTGATGGGCMMS-3'	Reverse primer for first amplification of *attH*
attH3F	5'-CASYTSSAGCTGGGYCCG-3'	Forward primer for second amplification of *attH*
attH3R	5'-GSRCGGKTTCAGSRTCGACGA-3'	Reverse primer for second amplification of *attH*
att1b_F	5'-TGTTCGACCTGTTCGCGGTGCT-3'	Forward primer for first amplification of *attD*
att1b_R	5'-TGASCCCRTCGATGAAGGCGAT-3'	Reverse primer for first amplification of *attD*
att2b_F	5'-ATGACCAGTCTCGACGAGCGCAA-3'	Forward primer for second amplification of *attD*
att2b_R	5'-YGTAGCAGGCGAGSACCACCGT-3'	Reverse primer for second amplification of *attD*
FC27	5'-AGAGTTTGATCCTGGCTCAG-3'	Forward primer for amplification of 16S rDNA
RC1492	5'-TACGGCTACCTTGTTACGACTT-3'	Reverse primer for amplification of 16S rDNA

Materials and Methods

General experimental procedures

NMR spectra were recorded using a Bruker Avance III HD 600 MHz spectrometer equipped with a helium High Sensitivity Prodigy Cryoprobe (600 and 150 MHz for ^{1}H and ^{13}C NMR, respectively). High-performance liquid chromatography- low resolution electrospray-ionization mass spectrometry (HPLC-LR-ESI-MS) data were obtained using an Agilent 1200 series HPLC system equipped with a photo-diode array detector and a 6130 ESI quadrupole mass spectrometer. HPLC-HR-ESI-MS was carried out using an Agilent 6530 QTOF mass spectrometer equipped with a 1290 Infinity Binary HPLC system. HPLC purifications were carried out using an Agilent 1100 or 1200 series HPLC systems (Agilent Technologies) equipped with a photodiode array detector. UV spectra were acquired on an Amersham Biosciences Ultrospec 5300-pro UV/Visible spectrophotometer. All solvents were of HPLC quality. DNA samples were quantified using NanoDrop 1000 Spectrophotometer (Thermo Scientific). No unexpected or unusually high safety hazards were encountered with any of the experiments outlined below.

Strains isolation and bioactivity screening

Strains were isolated from samples collected in the Brazilian Amazon (Anavilhanas National Park and Ducke Reserve), and in Southeastern Brazil (Itatiaia National Park and University of São Paulo, Ribeirão Preto campus) (authorizations/permits: SISBIO #46555-1 to #46555-7 and CNPq #010936/2014-9). Bacteria were isolated as previously described,[3] and cultured at 30 ºC, using chitin medium supplemented with the antifungals nystatin and cycloheximide (per liter: 4 g chitin, 0.7 g K_2HPO_4, 0.3 g KH_2PO_4, 0.5 g $MgSO_4 \cdot 5H_2O$, 0.01 g $FeSO_4 \cdot 7H_2O$, 0.01 g $ZnSO_4 \cdot 7H_2O$, 0.01 g $MnCl_2 \cdot 4H_2O$, 20 g of agar, 0.04 g/L nystatin, and 0.05 g/L cycloheximide). Bacteria were subcultured on International Streptomyces Project Medium 2 (ISP2) agar plates (per liter: 4 g yeast extract, 10 g malt extract, 4 g dextrose, 20 g agar) with antifungals (0.04 g/L nystatin, and 0.05 g/L cycloheximide). *Pseudonocardia* strains were identified by 16S rRNA sequencing.

Binary challenge assays were performed between all bacterial isolates from ant nests and six reporter strains: *Escherichia coli*, *Pseudomonas aeruginosa*, *Candida albicans*, *Trichoderma reesei*, *Streptomyces* sp. I8 and *Streptomyces* sp. I17. Bacterial isolates were inoculated in the center of wells of 6-well plates containing ISP2 agar and incubated for 10 days at 30 ºC, after which the reporter strain was inoculated side-by-side in a distance of 1 cm.

HPLC-MS

Brazilian *Pseudonocardia* were grown for 10 days on liquid ISP2 containing Diaion® HP-20 resin. From each culture, the resin was filtered, washed with deionized water and extracted with acetone. Extracts from Brazilian *Pseudonocardia* isolates were analyzed by HPLC-HR-ESI-MS using a Phenomenex Kinetex 2.6 mm EVO C18 100Å (100 x 2.1 mm) column using the following conditions: 10% ACN + 0.1% FA/H_2O + 0.1% FA for 1 min, 10% to 100% ACN + 0.1% FA over 9 min, 100% 100% ACN + 0.1% FA for 3.5 min at a flow rate of 0.3 mL/min. Panamanian *Pseudonocardia* were grown for 14 days on ISP2 agar. From each solid culture, 2 cm² was excised and extracted in 1.5 mL isopropanol for 6 hours. These extracts were dried under reduced pressure, redissolved in 1.1 mL 10% methanol in water, and loaded onto a pre-equilibrated 96-well Strata-X reversed phase SPE (Phenomenex). The SPE was rinsed with 1 mL 10% methanol in water, and the sample was eluted with 1 mL methanol and subjected to LCMS analysis. The extracts of Panamanian strains were analyzed using the following HPLC method: hold 10% ACN + 0.1% FA/H_2O + 0.1% FA for 0.5 min then gradient to 100% ACN + 0.1% FA over 7.5 min then hold 100% ACN + 0.1% FA for 1 min, returning to 10% ACN + 0.1% FA/H_2O + 0.1% FA, with a flow rate of 0.5 mL/min. MS spectra were acquired in positive ion mode and a mass range of 50-1,700 *m/z*. For MS and MS/MS measurements the electrospray ionization (ESI) parameters were set to 10 L/min sheath gas flow, 3.0 L/min auxiliary gas flow, and 325°C gas temperature. The spray voltage was set to 3500 V, the inlet capillary temperature was set to 300°C. The most abundant 5 ions per MS scan were selected for fragmentation and then subsequently excluded from refragmentation for 10 secs. The collision energies were set to 10, 20, and 30 eV. Compounds were de-novo annotated manually using the Agilent MassHunter Workstation software and compared to the Antibase database.

Isolation of attinimicin

Spore stocks of *Pseudonocardia* sp. ICBG162 were reactivated in ISP2 medium and 2 mL were transferred to 50 mL of fresh ISP2 medium. One liter of liquid ISP2 medium was inoculated with 50 mL of the starter culture of the attinimicin producers. The strains were cultured in 4 L flasks containing 1 L of liquid ISP2 medium containing Diaion® HP-20 (30 g/L). A total of 10 flasks were cultured in shakers for 10 d, at 30 ºC, and 200 rpm shaking. The HP-20 resin was filtered, washed with deionized water and extracted twice with acetone (300 mL). The organic extracts were dried under reduced pressure. Attinimicin was further purified by semi-preparative HPLC, using a Phenomenex Synergi™ 10 µm Hydro-RP 80 Å column (250 x 10 mm), using the following conditions: 10% ACN for 5 min, then a linear gradient to 28% ACN over 35 min, linear gradient to 100% over 2 min then hold 100% ACN for 2 min, with a flow rate of 3 mL/min. We obtained yields of attinimicins between 4-6 mg/L.

Determination of the absolute configuration of the amino acids in attinimicin

The absolute configuration of the amino acids present in attinimicin was determined using Marfey's method.[4] Attinimicin (1 mg) was hydrolyzed at 110 ºC in 500 µL of 6 N HCl (aq) for 18 h. HCl was removed under reduced pressure and the dry material was resuspended in 500 µL of DI H_2O and dried three times to remove residual acid. The hydrolysate was dissolved in 100 µL of 1 N $NaHCO_3$ (aq) and incubated with 50 µL of 10 mg/mL L-FDAA, in acetone, at 80 ºC for 3 min. The reaction was quenched by adding 50 µL of 2 N HCl (aq). Aqueous acetonitrile (1:1 – 100 µL) was added to dissolve the mixture. The same derivatization process was performed with D and L amino acid standards. A 10 µL aliquot of the hydrolysate derivative was analyzed by HRLC-ESI-MS using an analytical C18 column (Phenomenex Kinetex 2.6 µm EVO C18, 100 Å, 100 x 2.1 mm), with a gradient solvent system (20% to 60% ACN with 0.1% FA over 40 min). The absolute configuration of individual amino acids was determined by comparing retention times of the hydrolysate derived amino acids conjugates (Ser: 22.20 and 23.02 min; Orn: 12.06 min) with amino acid standards (L-Ser: 22.16 min; D-Ser: 22.95 min; L-Orn: 12.06 min; D-Orn: 10.65 min) subjected to Marfey's derivatization and identified by MS.

DNA isolation, sequencing and analysis

Pseudonocardia spp. were cultured in 5 mL of ISP2 medium, 30 ºC, 200 rpm. After four days, cell pellets were obtained by centrifuging the cultures and discarding the supernatant. *Pseudonocardia* spp. cell pellets were incubated with a lysis solution (Tris-HCl 25 mM pH 8.5, sucrose 0.3 M, EDTA 25 mM), RNAse A (100 mg/mL) and lysozyme (50 mg/mL) at 37 ºC for 30 min, in a water bath. Then 2% SDS and proteinase K (20 mg/mL) solutions were added and the mixture was kept at 55 ºC for 30 min. Nucleic acids were extracted using 5 M NaCl (aq) and phenol:chloroform: isoamyl alcohol (25:24:1). DNA was precipitated from the aqueous phase using isopropanol and washed with 70% ethanol (aq). The genomes of *Pseudonocardia* sp. ICBG162, ICBG601, ICBG1034, ICBG1122, ICBG1127, ICBG1142, ICBG1288 and ICBG1293 were sequenced using PacBio single-molecule, real-time (SMRT) cell technology[5] at the Duke University Center for Genomic and Computational Biology (GCB), Genome Sequencing Shared Resource.

Bioinformatic analysis

Genomes were assembled using the Canu[6] pipeline (Harvard Chan Bioinformatics Core, Harvard T.H. Chan School of Public Health, Boston, MA). Quality of genome assemblies was assessed using QUAST 5.0.2[7]. The putative attinimicin gene cluster was annotated using antiSMASH 5.0,[8] BLASTp and manual analysis using Geneious 11.1.4.

Domain specificities were determined by phylogenetic analysis. For the construction of phylogenetic trees, sequences were aligned using MUSCLE,[9] trimmed manually and the alignments were refined using MUSCLE. Maximum-likelihood phylogenetic trees were reconstructed using IQ-TREE 1.6.[1] ModelFinder[10] was used for each alignment to choose the best model for protein evolution. Branch support was assessed by ultrafast bootstrap UFBoot2[11] (1000 bootstrap replicates). Phylogenetic trees were visualized with iTol.[12] Sequence similarity networks of attinimicin

BGCs were obtained using BiG-SCAPE[13] v1.0.1 docker image (image ID: c056fa35c8c6), using the internal MIBiG 1.4 dataset and a cutoff of 0.3.

In situ detection of attinimicin

Ants and samples of the fungal garden were collected at Itatiaia National Park (RJ, Brazil) in December 2018. Samples were extracted with 1 mL MeOH and dried under reduced pressure. Extracts were resuspended in 100 µL of aqueous acetonitrile (1:1) and 10 µL were analyzed by HR-ESI-MS and data-dependent MS/MS.

Determination of pFe^{III} for attinimicin and metal complexing analysis.

pFe^{III} for attinimicin was determined as reported by Abergel et al.[14] with modifications. Briefly, purified *apo*-attinimicin, obtained as described above, was dissolved in DI water and incubated with 10-fold excess of Fe^{3+} for 2 h. In order to remove the excess of free iron, the solution was applied onto a SPE C-18 column and washed with water. Fe-attinimicin was eluted with methanol and dried under reduced pressure. Fe-attinimicin was then dissolved in HEPES buffer (10 mM HEPES and 0.1 M KCl, pH 7.4) and three different concentrations of ETDA were added from EDTA stock solutions in HEPES buffer, ranging from 0.05 to 1 mM. Each reaction consisted of a total volume of 100 µL and a defined final concentration of 0.1 mM Fe-attinimicin. The reaction mixture was equilibrated at rt for 24 h, and the solutions were analyzed by HPLC-MS. Standard curves of *apo*-attinimicin and Fe-attinimicin were generated, and both attinimicin species were quantified. Concentrations of *apo*-EDTA and Fe-EDTA were calculated from total initial EDTA and [*apo*-attinimicin]. The log [EDTA]/[attinimicin] was plotted against log [Fe-EDTA]/ [Fe-attinimicin] and the data were fit to the following equation[14]:

$$\log\left(\frac{[Fe-EDTA]}{[Fe-att]}\right) = log\left(\frac{[EDTA]}{[att]}\right) + \Delta pFe$$

To test metal ion complexing selectivity, attinimicin was dissolved in water and solutions containing Fe^{3+}, Ga^{3+}, Ca^{2+}, Zn^{2+}, Cu^{2+}, Mg^{2+} in an equimolar ratio were added. The mixtures were equilibrated at rt overnight and analyzed by HPLC-MS.

PCR amplification and sequencing

For the detection of the biosynthetic gene clusters nested PCRs were performed. PCRs were conducted using Phusion® High-Fidelity DNA Polymerase (ThermoFisher, USA) and the following conditions. First amplification: 98 °C, 2 min; denaturation: 98 °C, 20 s; annealing: 64 °C, 15 s; extension: 72 °C, 1.5 min (for 35 cycles); final extension 72 °C, 4 min. Second amplification: denaturation 98 °C, 2 min; denaturation: 98 °C, 20 s; annealing: 60 °C, 15 s; extension 72 °C, 30 s (for 35 cycles); final extension: 72 °C, 4 min. For primer sets used see Supplementary Table S10. Bacterial samples were heated to 100 °C for 10 min with 50 µL of PrepMan® Ultra Sample Preparation Reagent (ThermoFisher, USA) and centrifuged. 1 µL of this solution was used as a template for the first amplification. Subsequently, PCR fragments were separated by agarose gel electrophoresis with 10 µL of SYBR Safe DNA gel stain per 100 mL of agarose in 1x TAE and visualized by ChemiDoc XRS equipment, using the Image Lab software for image acquisition and analysis. Bands with the expected product size (~600 bp) were cut out, extracted from the gel using the QIAquick PCR Purification kit (QIAGEN, USA), and the purified fragments were used for a second round of amplification as described above. The final product was verified by Sanger sequencing followed by bioinformatic analysis using Geneious Prime 11.1.4 software package.

In vitro antifungal activity of attinimicin and oxachelin

Disc-diffusion assay: Attinimicin and oxachelin (present in our laboratory's pure compound library as a result of an unpublished project, isolated under the same conditions and using the same method used for attinimicin) were each

screened for activity against *Escopovsis* sp. ICBG77 and *Leucoagaricus gongylophorus* ICBG107 using a modified version of the disc-diffusion assay. 1 mL of a *Escovopsis* sp. ICBG77 spore suspension ($1x10^5$ spores/mL) or a homogenate of *Leucoagaricus gongylophorus* ICBG107 were inoculated on PDA plates (Potato Dextrose Agar, BD Difco, USA). Attinimicin and oxachelin were dissolved in DMSO and 10 μL of these solutions were spotted on the plates. Cultures were incubated at 30 ºC and monitored for 14 days.

Minimum inhibitory concentration assays: The MIC of attinimicin was determined using a broth microdilution method. The assay was performed in duplicate. The MIC value of attinimicin was defined as the lowest concentration in which a 50% reduction in growth is observed (determined spectrophotometrically, at 600 nm). Attinimicin and Fe-attinimicin were dissolved in DMSO and 20 μL of a suspension of *Escovopsis* sp. ICBG77 spores ($1x10^5$ spores/mL) were inoculated in each well of 96-well plates, containing MH broth (Mueller Hinton Broth, BD Difco, USA). The plates were incubated at 30 ºC, 50 rpm and monitored for 7 days.

In vivo antifungal activity of attinimicin

Animals were maintained in accordance with the criteria of the Association for Assessment and Accreditation of Laboratory Animal Care. Approval for animal studies were obtained from the Animal Research Committee of the William S. Middleton Memorial Veterans Hospital. Ethical approval for experiments and protocols were obtained from the University of Wisconsin Institutional Animal Care and Use Committee. Six-week-old, specific-pathogen-free, female ICR/Swiss mice weighing 23–27 g were used for *in vivo* studies (Harlan Sprague-Dawley, Indianapolis, IN). Cyclophosphamide (Mead Johnson Pharmaceuticals, Evansville, IN) was used to induce neutropenia (100 neutrophils per mm^3) in mice. Subcutaneous injections were administered 4 days (150 mg kg^{-1}) and 1 day (100 mg kg^{-1}) before infection and 2 days after infection (100 mg kg^{-1}). *Candida albicans* was cultured on SDA 24 h before infection. To prepare the initial inoculum, three to five colonies were transferred to a sterile pyrogen-free 0.9% saline solution warmed to 35 °C. The suspension was further adjusted to a 0.6 transmittance at 530 nm and 0.1 mL of inoculum was injected via lateral tail vein 2 h prior to the start of drug administration. Animals were sacrificed by CO_2 asphyxiation, kidneys were removed and homogenized in sterile 0.9% saline at 4 °C. Serial dilutions were performed with the homogenates and aliquots were plated on SDA for viable fungal colony counts after incubation for 24 h at 35 °C. The limit of detection was 100 CFU per mL. Results were expressed as the mean number of CFU per kidney for three mice.

REFERENCES

(1) Nguyen, L. T.; Schmidt, H. A.; Von Haeseler, A.; Minh, B. Q. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol. Biol. Evol.* **2015**, *32* (1), 268–274.

(2) Nurchi, V. M.; Crisponi, G.; Lachowicz, J. I.; Medici, S.; Peana, M.; Zoroddu, M. A. Chemical Features of in Use and in Progress Chelators for Iron Overload. *J. Trace Elem. Med. Biol.* **2016**, *38*, 10–18.

(3) Cafaro, M. J.; Currie, C. R. Phylogenetic Analysis of Mutualistic Filamentous Bacteria Associated with Fungus-Growing Ants. *Can. J. Microbiol.* **2005**, *51* (6), 441–446.

(4) Marfey, P. Determination of D-Amino Acids. II. Use of a Bifunctional Reagent, 1,5-Difluoro-2,4-Dinitrobenzene. *Carlsberg Res. Commun.* **1984**, *49* (6), 591–596.

(5) Ardui, S.; Ameur, A.; Vermeesch, J. R.; Hestand, M. S. Single Molecule Real-Time (SMRT) Sequencing Comes of Age: Applications and Utilities for Medical Diagnostics. *Nucleic Acids Res.* **2018**, *46* (5), 2159–2168.

(6) Koren, S.; Walenz, B. P.; Berlin, K.; Miller, J. R.; Bergman, N. H.; Phillippy, A. M. Canu: Scalable and Accurate Long-Read Assembly via Adaptive κ-Mer Weighting and Repeat Separation. *Genome Res.* **2017**, *27* (5), 722–736.

(7) Gurevich, A.; Saveliev, V.; Vyahhi, N.; Tesler, G. QUAST: Quality Assessment Tool for Genome Assemblies. *Bioinformatics* **2013**, *29* (8), 1072–1075.

(8) Blin, K.; Shaw, S.; Steinke, K.; Villebro, R.; Ziemert, N.; Lee, S. Y.; Medema, M. H.; Weber, T. AntiSMASH 5.0: Updates to the Secondary Metabolite Genome Mining Pipeline. *Nucleic Acids Res.* **2019**, *47* (W1), W81–W87.

(9) Edgar, R. C. MUSCLE: Multiple Sequence Alignment with High Accuracy and High Throughput. *Nucleic Acids Res.* **2004**, *32* (5), 1792–1797.

(10) Kalyaanamoorthy, S.; Minh, B. Q.; Wong, T. K. F.; Von Haeseler, A.; Jermiin, L. S. ModelFinder: Fast Model Selection for Accurate Phylogenetic Estimates. *Nat. Methods* **2017**, *14* (6), 587–589.

(11) Hoang, D. T.; Chernomor, O.; von Haeseler, A.; Minh, B. Q.; Vinh, L. S. UFBoot2: Improving the Ultrafast Bootstrap Approximation. Molecular Biology and Evolution. *Mol. Biol. Evol.* **2018**, *35* (2), 518–522.

(12) Letunic, I.; Bork, P. Interactive Tree Of Life (ITOL) v4: Recent Updates and New Developments. *Nucleic Acids Res.* **2019**, *47* (W1), W256–W259.

(13) Navarro-Muñoz JC, Selem-Mojica N, Mullowney MW, Kautsar SA, Tryon JH, Parkinson EI, De Los Santos EL, Yeong M, Cruz-Morales P, Abubucker S, Roeters A. A computational framework to explore large-scale biosynthetic diversity. *Nature chemical biology*. **2020**, *16* (1), 60-68.

(14) Abergel, R. J.; Zawadzka, A. M.; Raymond, K. N. Petrobactin-Mediated Iron Transport in Pathogenic Bacteria: Coordination Chemistry of an Unusual 3,4-Catecholate/Citrate Siderophore. *J. Am. Chem. Soc.* **2008**, *130* (7), 2124–2125.