Supporting Information

A Rebeccamycin Analog Provides Plasmid-Encoded Niche Defense

Ethan B. Van Arnam,[†] Antonio C. Ruzzini,[†] Clarissa S. Sit,[†] Cameron R. Currie,[‡] and Jon Clardy[†]*

[†]*Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, Massachusetts, 02115*

[‡]*Department of Bacteriology, University of Wisconsin-Madison, 6145 Microbial Science Building, 1550 Linden Drive, Madison, Wisconsin, 53706*

*Correspondence should be addressed to J.C. (jon_clardy@hms.harvard.edu)

Supporting Information Contents

Supplementary methods

General chemical analysis procedures: UV-visible absorbance spectra were collected on an Amersham Biosciences Ultrospec 5300 Pro spectrophotometer. HPLC purifications were carried out on an Agilent 1100 Series analytical HPLC system equipped with a photo diode array detector. High resolution mass spectrometry analysis was performed on a Bruker Maxis Impact LC-q-TOF mass spectrometer. ^{1}H, gCOSY, and HSQC NMR experiments were performed on a Varian VNMRS 600 MHz spectrometer equipped with a triple resonance HCN inverse probe. ^{13}C experiments were performed on a Varian 400 MHz spectrometer equipped with a Varian OneNMR probe. gHMBC experiments were performed on a Varian INOVA 500 MHz spectrometer equipped with a triple resonance HCN coldprobe. Chemical shifts were referenced to the residual solvent peaks in DMSO-d_6 and THF-d_8. Optical rotation was measured on a Jasco P-2000 polarimeter fitted with a microcell (10 mm path length).

9-Methoxyrebeccamycin production and purification: Spores of *Pseudonocardia* BCI2 were diluted into sterile double distilled water (ddH_2O) and spread onto ISP2 agar plates (BD Difco™ ISP2; 60 mL agar per 150 x 15 mm Petri dish), which were incubated at 30 °C for 14 days. Agar was then cut into squares and soaked in ethyl acetate overnight to extract organic components from the solid media. This extract was decanted and the agar was soaked in an additional volume of ethyl acetate for 2 hrs. The combined ethyl acetate extracts were concentrated *in vacuo* and adsorbed onto celite for dry packing onto a 10 g C_{18} SepPak column (Waters) that had been conditioned and pre-equilibrated with 20% isopropanol in water. Fractions were eluted with a step gradient of 20%, 40%, 60%, 80%, and 100% isopropanol and concentrated to dryness. Consecutive fractions from elution at 40% and 60% were most active in inhibition of *Pseudonocardia* PLR1. Semipure material from these fractions was purified by reversed-phase HPLC (Phenomenex Luna 5u C18 semipreparative column, 250 × 10.00 mm, 3 mL/min) with an isocratic solvent mixture of 55% acetonitrile in water. 9-Methoxyrebeccamycin eluted at 11.3 minutes. Final purification was performed by silica flash chromatography, eluting with a mixture of 2:1 tetrahydrofuran/hexanes. The overall yield of pure 9-methoxyrebeccamycin (isolated as an amorphous yellow solid) was 6 mg/L of agar.

9-Methoxyrebeccamycin (2): $[\alpha]_D^{26}$ +107° (THF); UV (MeOH) λ_{max} (log ε) 207 (4.4), 244 (4.6), 296 (4.4), 317 (4.5), 401 (3.6) nm; NMR spectral data, see **Supplemental Tables S4 and S5**; HR-ESI-TOFMS *m/z* 600.0920 $[M+H]^+$ (calcd for $C_{28}H_{23}Cl_2N_3O_8$: 600.0935)

Intruder assay: An intruder assay was performed using a protocol adapted from Poulsen (2007).[1] A concentrated ddH_2O suspension of spores of the "resident" *Pseudonocardia* strain (5 µL) was spotted in the center of an agar plate (ISP2 1% agar; 8 mL agar per 60 x 15 mm Petri dish). The plates were incubated at 30 °C for 13 days. A ddH_2O suspension of spores of the "intruder" *Pseudonocardia* (250 uL; approximately 10^7 spores) was spread around the "resident" colony using a ddH_2O-wetted sterile cotton swab, covering all remaining surface area of the plate. The plates were incubated at 30 °C for an additional 6 days and the diameter of the inhibitory zone surrounding the intruder colony was measured. Two replicates were performed for each pairing.

Spot-on-lawn inhibitory assay: Anti-*Pseudonocardia* activity of crude extracts, pure 9-methoxyrebeccamycin, and rebeccamycin was assessed using a spot-on-lawn assay (**Figure S2**). A soft agar lawn of each *Pseudonocardia* strain in **Table S1** was prepared by mixing spores (approximately 10^{11}, estimated by hemocytometer counting) into 10 mL molten ISP2 soft agar (55 °C, 0.75% agar). This mixture was poured on top of a plate of ISP2 agar (60 mL in a 150 x 15 mm Petri dish) and allowed to solidify. 5 μL of each of the following DMSO solutions was spotted directly onto each prepared *Pseudonocardia* lawn:

I. DMSO only
II. Extracts of strains BCI1 and BCI2: Each strain was grown on ISP2 agar and ethyl acetate extracts were prepared as described for production and purification of 9-methoxyrebeccamycin. Extracts were evaporated to dryness and redissolved in DMSO (1 mL DMSO per 10 mL agar of the *Pseudonocardia* culture)
III. 9-Methoxyrebeccamycin and rebeccamycin: 9-Methoxyrebeccamycin and rebeccamycin (Santa Cruz Biotech) were dissolved in DMSO at concentrations of 20 μg/mL.

The plates were incubated at 30 °C for 8 days, at which point the resulting inhibitory zones were photographed

Determination of minimum inhibitory concentration: DMSO solutions of rebeccamycin (Santa Cruz Biotech) and 9-methoxyrebeccamycin were prepared as serial 2-fold dilutions starting at 30 mM and mixed into molten (55 °C) ISP2 agar at 1% DMSO by volume. These agar solutions were dispensed into 96-well plates (100 μL per well) and allowed to solidify. A concentrated suspension of *Pseudonocardia* spores in ddH_2O was spotted on top of the agar in each well (1 μL per spot, approximately 10^9 spores per spot). The plates were incubated at 30 °C in a humidified chamber for 3 days and the minimum concentration at which there was no visible *Pseudonocardia* growth was noted.

Genome data deposition: Complete genomes have been deposited in the GenBank database and raw sequence data have been deposited in the Sequence Read Archive (SRA). BCI1 (*Pseudonocardia* sp. EC080619-01) can be accessed using Genbank accession nos. CP012184-86 and SRA accession no. SRP061657. BCI2 (*Pseudonocardia* sp. EC080610-09) can be accessed using Genbank accession nos. CP012181-3 and SRA accession no. SRP061656.

Sequence comparison and analysis: Conserved replicons in the BCI genomes were compared using two *in silico* DNA-DNA hybridization methods: (i) the Genome-to-Genome Distance calculator (GGDC)[2] and (ii) average nucleotide identity (ANI) calculator.[3] The 9-methoxyrebeccamycin gene cluster annotations were performed using antiSMASH2,[4] blastp (refseq db, updated 08/09/2015), and pairwise sequence alignments using previously reported clusters.[5] The Geneious aligner[6] was used for pairwise alignment and construction of a phylogenetic tree using the chromopyrrolic acid synthases (EspD from AB339 was used at the outgroup). The accession numbers for the gene clusters used in this analysis are as follows: *reb* (AJ414559.1); AB339 (KF551865.1); AB857 (KF551866.1); AB1533 (KF551870.1); AR1455 (KF551872.1) Atm (DQ297453.1); NM747 (KF551862.1).

Supplementary data

Table S1. *Pseudonocardia* isolates used in this study

Abbreviation	Strain ID	Collection Location	Ant Species
BCI1	EC080619-01	Barro Colorado Island, Panama Canal Zone, Panama	*Apterostigma dentigerum*
BCI2	EC080610-09	Barro Colorado Island, Panama Canal Zone, Panama	*Apterostigma dentigerum*
FRP1	EC080625-04	Frijoles Peninsula, Panama Canal Zone, Panama	*Apterostigma dentigerum*
LS1	HH130629-09	La Selva Biological Station, Costa Rica	*Apterostigma sp.*
LS4	EC060123-09	La Selva Biological Station, Costa Rica	*Apterostigma dentigerum*
PE1	AL041005-10	Peru	*Trachymyrmex cornetzi*
PLR1	EC080529-05	Pipeline Road, Panama Canal Zone, Panama	*Apterostigma dentigerum*
PLR2	EC080529-01	Pipeline Road, Panama Canal Zone, Panama	*Apterostigma dentigerum*

Table S2. Intruder assay inhibitory zones (cm)

Resident	**Intruder *Pseudonocardia***							
Pseudonocardia	**BCI1**	**BCI2**	**FRP1**	**LS1**	**LS4**	**PE1**	**PLR1**	**PLR2**
BCI1	0	0	0	0	0	0	0	0
BCI2	3.0	0	2.0	1.4	3.3	2.2	2.6	3.2

Table S3. MIC values for 9-methoxyrebeccamycin and rebeccamycin against three ant-associated *Pseudonocardia*

	BCI1	**BCI2**	**PLR1**
9-Methoxyrebeccamycin	0.59 μM	>300 μM	0.15 μM
Rebeccamycin	1.2 μM	>300 μM	0.29 μM

Figure S1. Colony images for *Pseudonocardia* BCI1 and BCI2 grown from spores on ISP2 agar for 7 days

BCI1

BCI2

Figure S2. Anti-*Pseudonocardia* spot-on-lawn activity of extracts, 9-methoxyrebeccamycin, and rebeccamycin

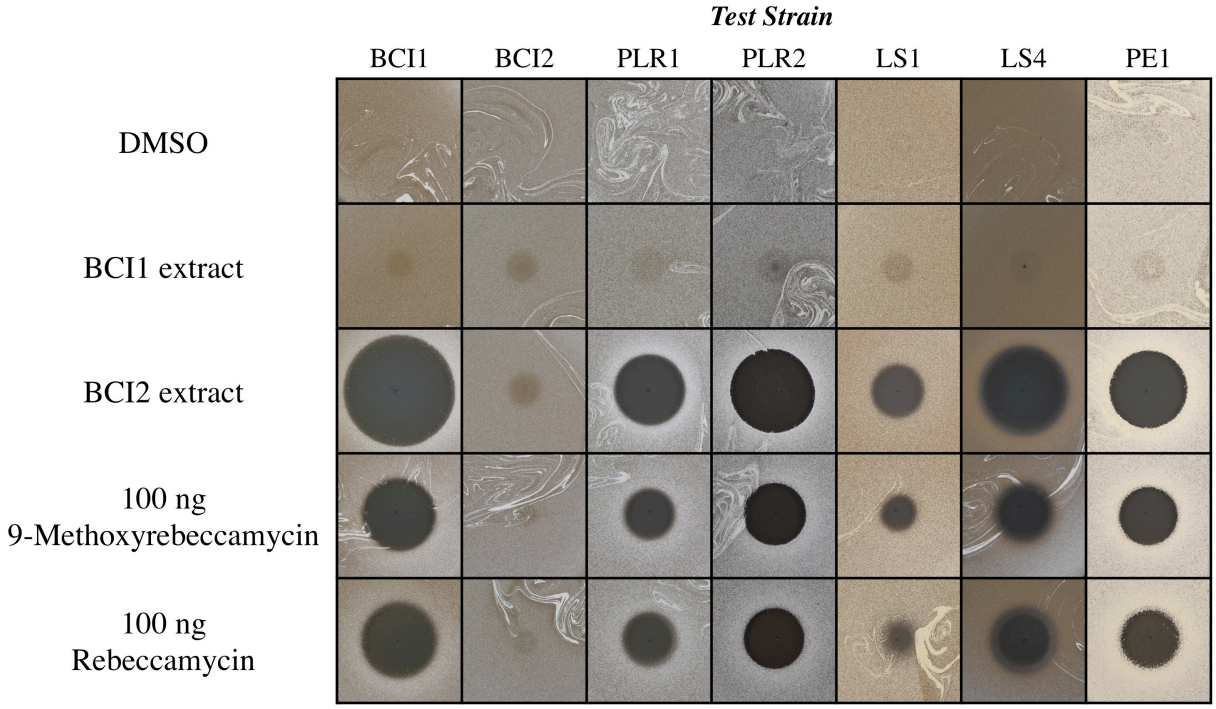
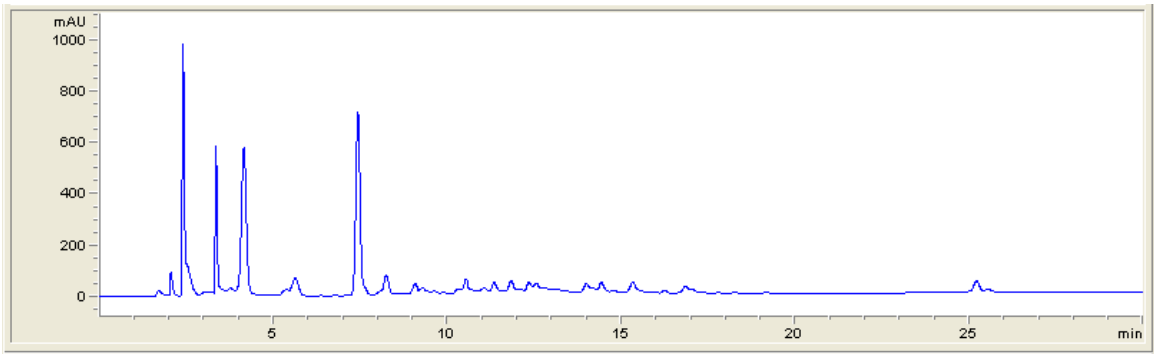


Figure S3. HPLC chromatograms (254 nm absorbance) of *Pseudonocardia* **(a)** BCI1 and **(b)** BCI2 extracts

a



b

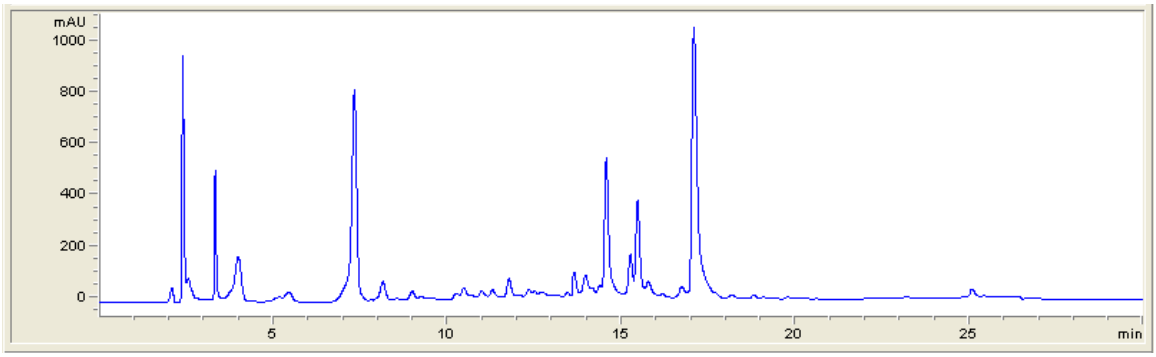
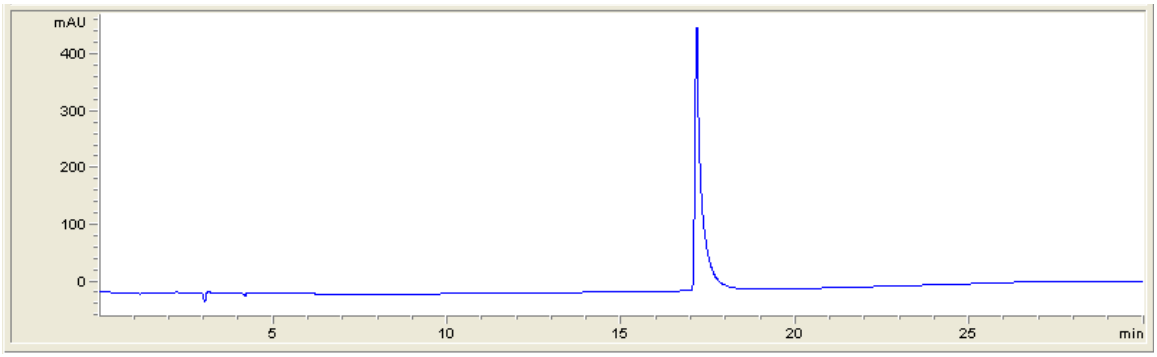


Figure S4. HPLC-MS trace of purified 9-methoxyrebeccamycin. **(a)** 254 nm absorbance; **(b)** extracted ion for 9-methoxyrebeccamycin $[M+H]^+$ (m/z = 600)

a



b

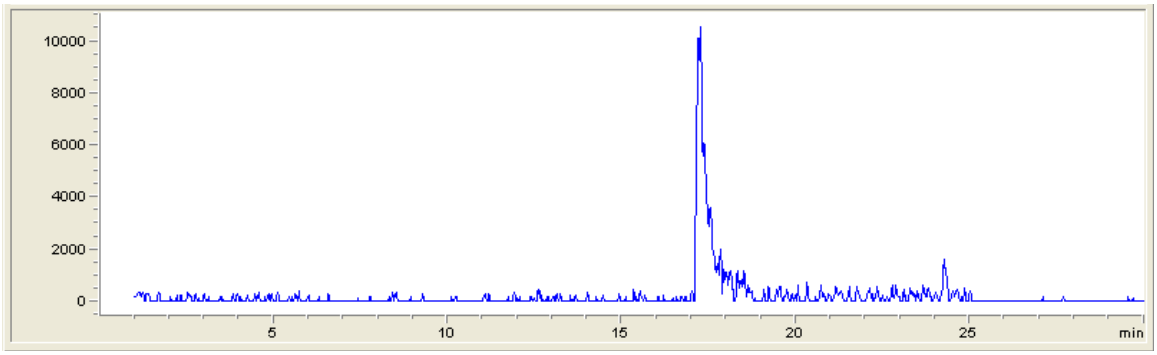


Figure S5. UV spectrum of 9-methoxyrebeccamycin

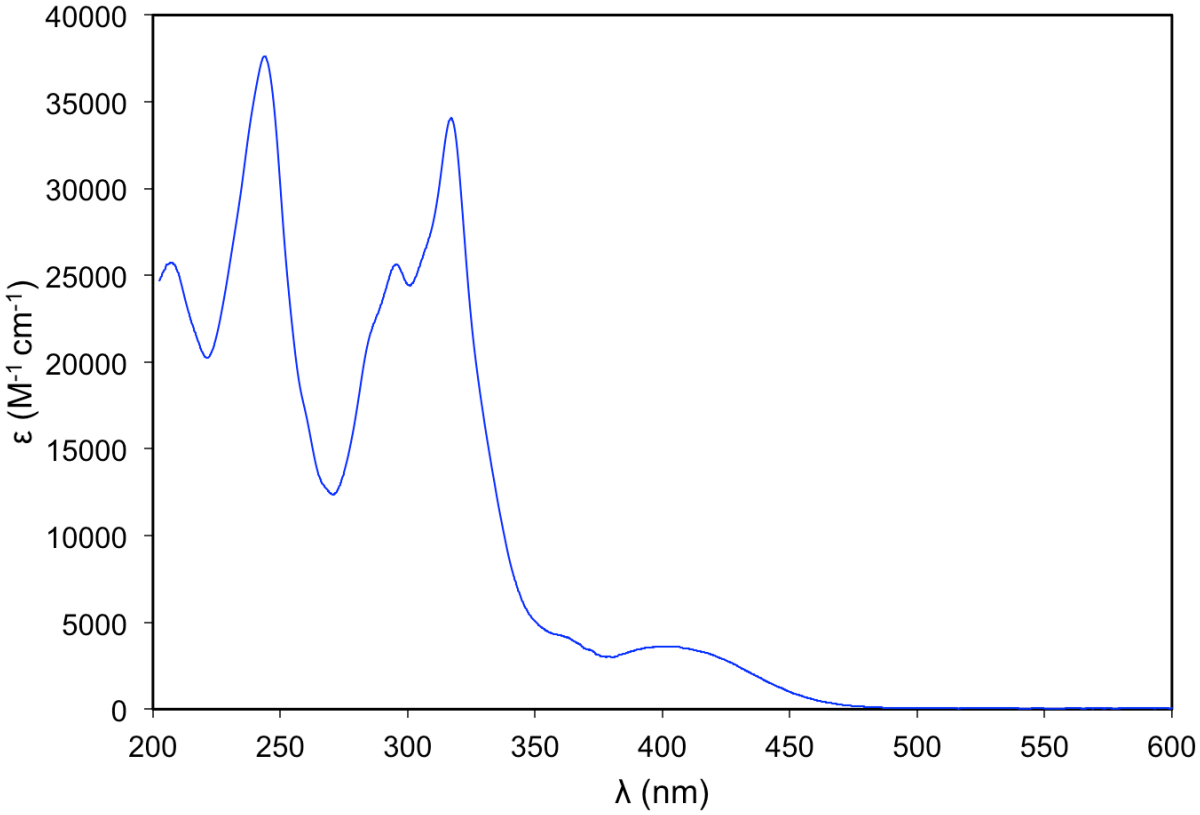


Table S4. NMR spectral data for 9-methoxyrebeccamycin (**2**) in DMSO-d_6. *J* values for sugar ^{1}H resonances obscured in the 1D ^{1}H spectrum were extracted from a *J*-resolved ^{1}H spectrum.

Position	δ_H	mult (*J* in Hz)	δ_C	
1			116.21	C
2	7.72	dd (7.8, 1.0)	126.89	CH
3	7.44	t (7.8, 7.8)	122.06	CH
4	9.08	d (7.9)	123.41	CH
4a			123.38	C
4b			117.30	C
4c			121.98	C
5			170.36	C
6-NH	11.34	s		
7			170.62	C
7a			120.76	C
7b			118.68	C
7c			123.40	C
8	8.91	d (1.9)	116.18	CH
9			146.95	C
9-OMe	4.09	s	56.79	CH_3
10	7.30	d (1.9)	110.62	CH
11			125.66	C
11a			129.78	C
12a			129.40	C
12b			129.44	C
13-NH	10.74	s		
13a			137.02	C
1'	6.88	d (9.2)	85.66	CH
2'	3.71	td (8.9, 8.9, 5.6)	72.62	CH
2'-OH	5.05	d (5.6)		
3'	3.59	td (8.8, 8.8, 5.8)	77.30	CH
3'-OH	5.42	d (5.8)		
4'	3.67	t (9.4, 9.4)	79.29	CH
4'-OMe	3.61	s	60.15	CH_3
5'	3.84	ddd (10.0, 4.2, 2.6)	79.92	CH
6'	3.94	m	59.76	CH_2
6'-OH	5.29	t (5.4, 5.4)		

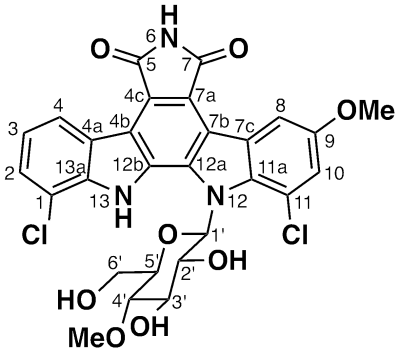


Table S5. NMR spectral data for 9-methoxyrebeccamycin (**2**) in THF-d_8. *J* values for sugar ^{1}H resonances obscured in the 1D ^{1}H spectrum were extracted from a *J*-resolved ^{1}H spectrum.

Position	δ_H	mult (*J* in Hz)	δ_C	
1			117.55	C
2	7.54	dd (7.6, 1.0)	127.45	CH
3	7.28	t (7.8, 7.8)	122.41	CH
4	9.10	dt (8.0, 0.7, 0.7)	124.88	CH
4a			125.09	C
4b			118.88	C
4c			123.35	C
5			170.80	C
6-NH	10.12	s		
7			171.03	C
7a			122.04	C
7b			120.39	C
7c			125.33	C
8	9.03	d (1.9)	118.31	CH
9			148.10	C
9-OMe	4.10	s	56.96	CH_3
10	7.16	d (2.0)	111.54	CH
11			127.40	C
11a			131.32	C
12a			130.66	C
12b			130.87	C
13-NH	10.73	s		
13a			138.63	C
1'	7.00	d (9.4)	87.31	CH
2'	3.85	ddd (9.4, 8.6, 5.2)	74.54	CH
2'-OH	4.36	d (5.2)		
3'	3.67	td (8.8, 8.8, 4.4)	79.81	CH
3'-OH	4.73	d (4.4)		
4'	3.82	m	80.47	CH
4'-OMe	3.71	s	61.08	CH_3
5'	3.82	m	82.02	CH
6'	4.07	ddd (12.1, 6.2, 2.4)	61.70	CH_2
	4.13	dt (12.2, 4.7, 4.7)		
6'-OH	4.58	dd (6.2, 5.1)		

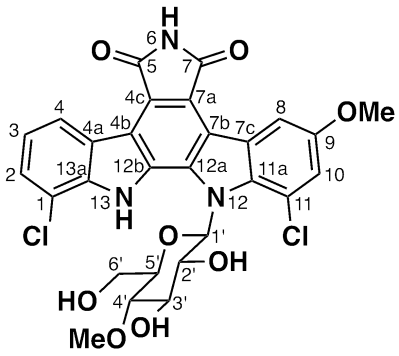
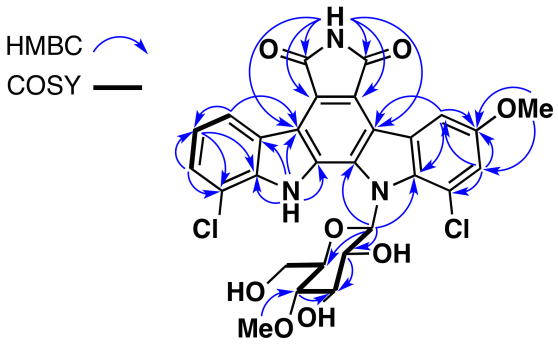


Figure S6. (a) Key 9-methoxyrebeccamycin HMBC and COSY correlations. **(b)** Sugar coupling constants for 9-methoxyrebeccamycin in DMSO-d_6. These values are consistent with the β configuration of 4-*O*-methylglucose.

a



b

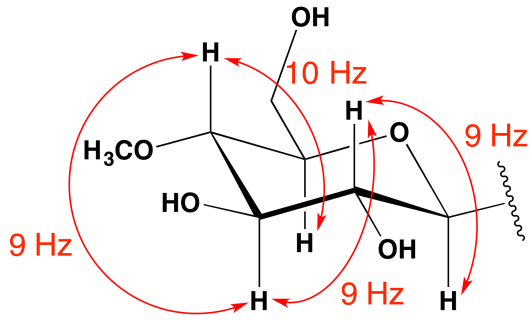
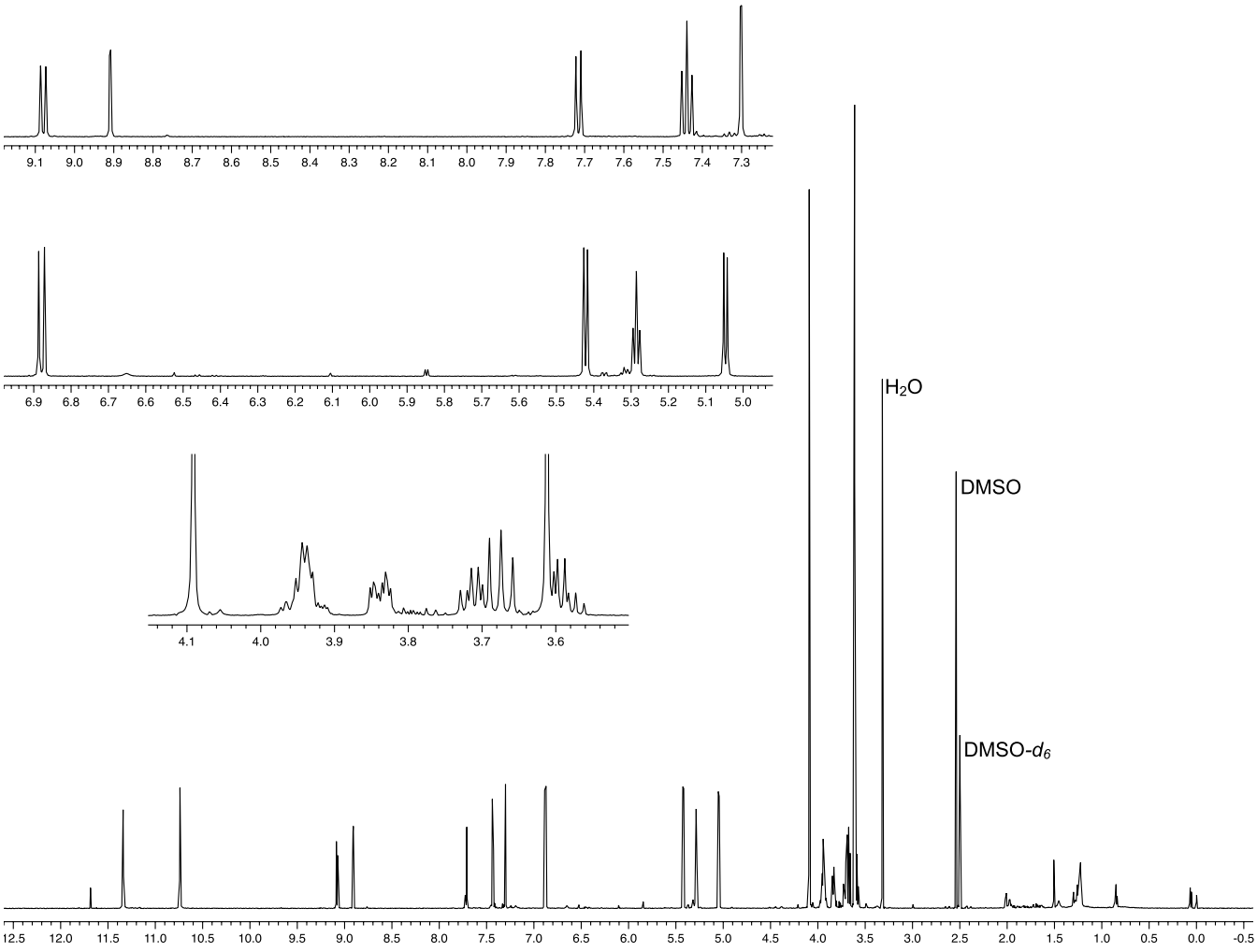
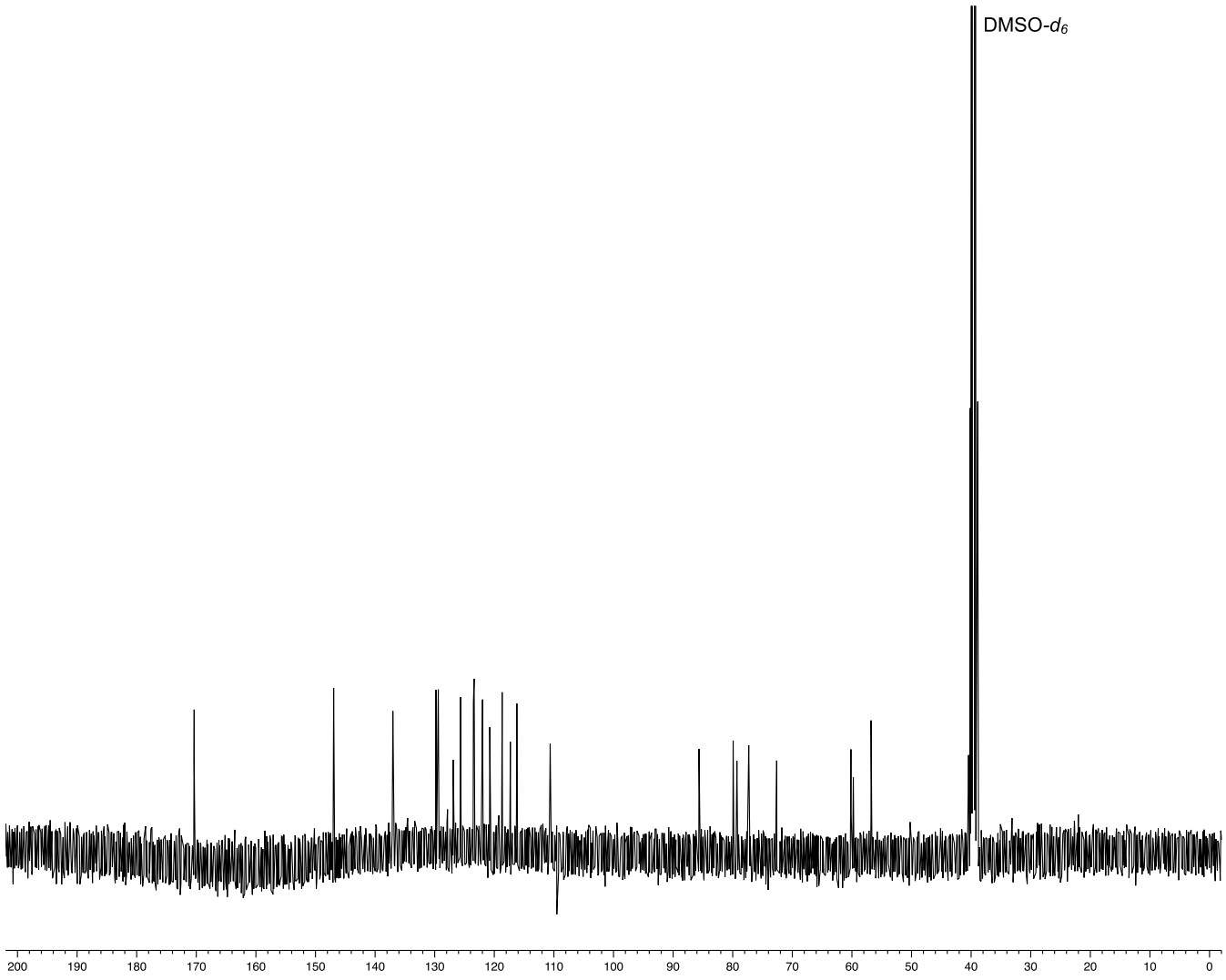


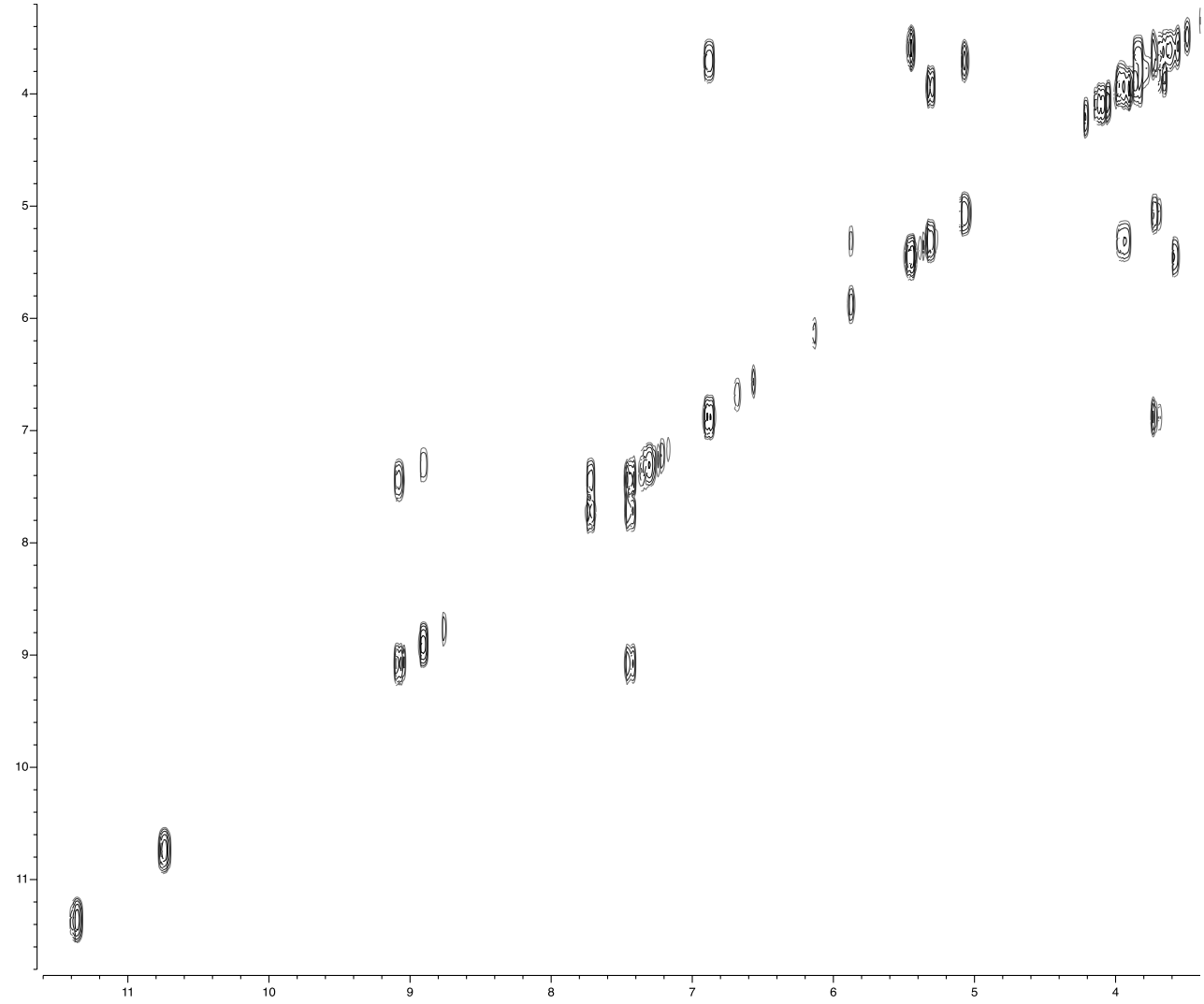
Figure S7. 9-Methoxyrebeccamycin NMR spectra in DMSO-d_6



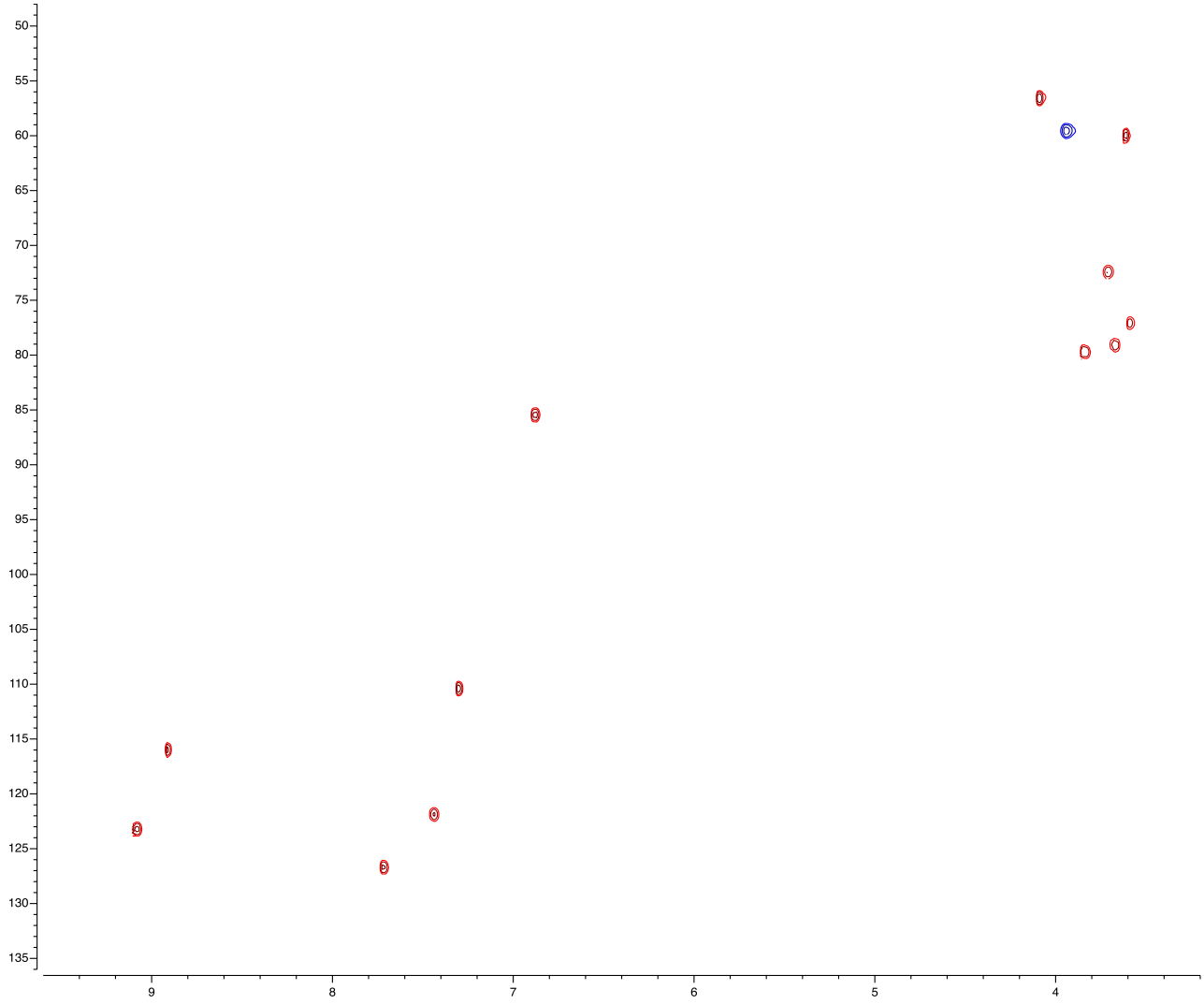
(a) ^{1}H NMR spectrum of 9-methoxyrebeccamycin in DMSO-d_6



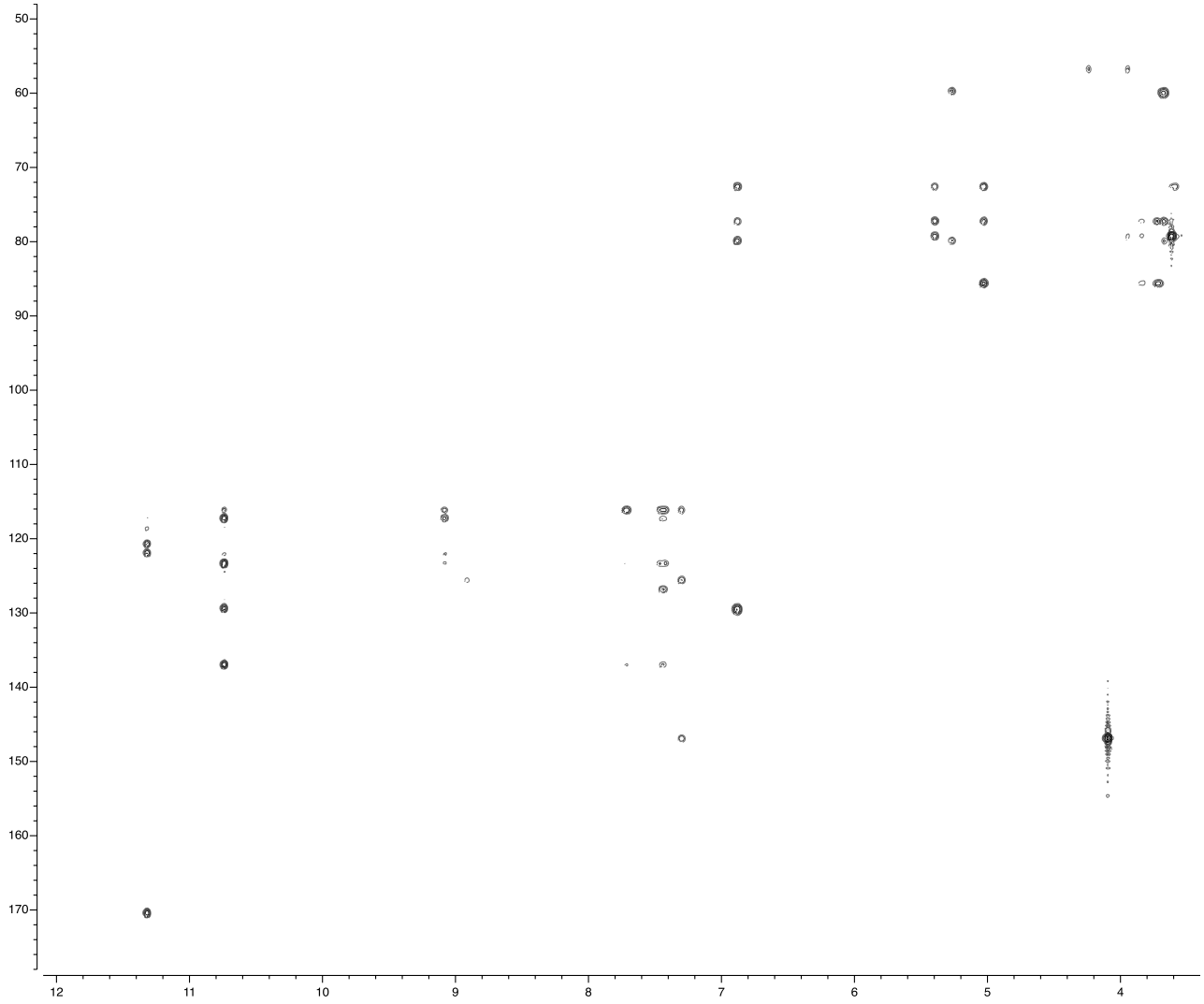
(b) ^{13}C NMR spectrum of 9-methoxyrebeccamycin in DMSO-d_6



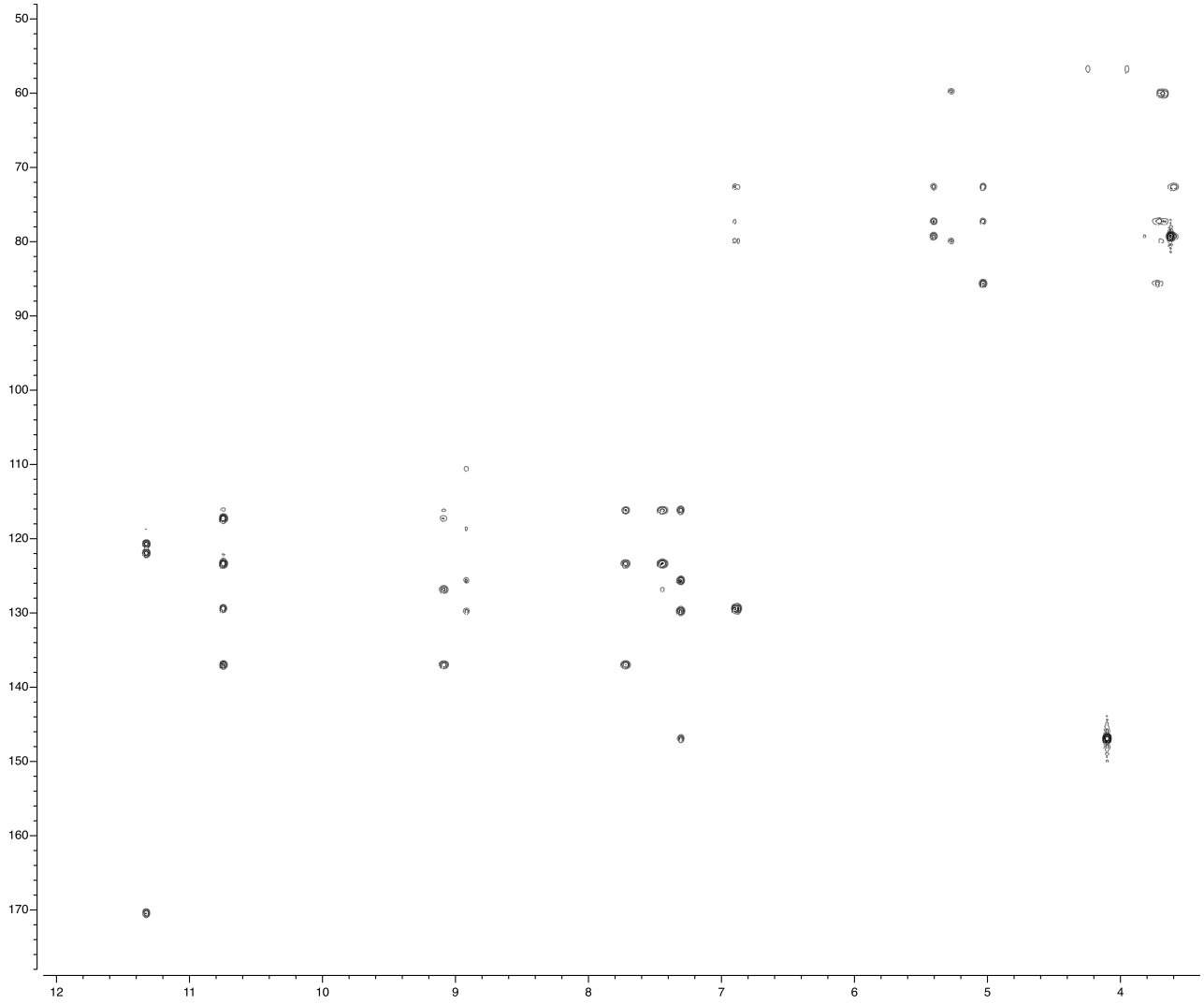
(c) gCOSY spectrum of 9-methoxyrebeccamycin in DMSO-d_6



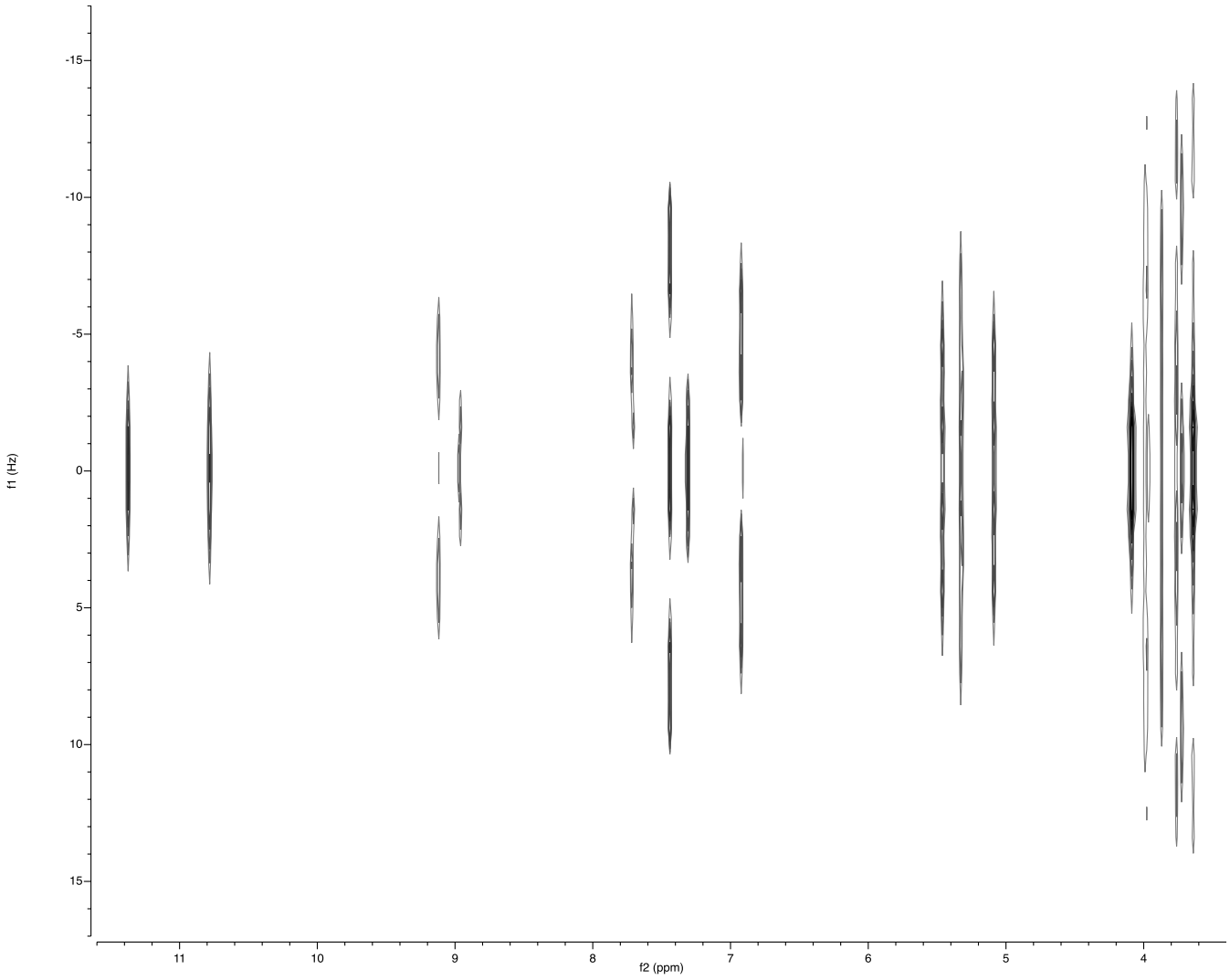
(d) Multiplicity-edited HSQC NMR spectrum of 9-methoxyrebeccamycin in DMSO-d_6. CH and CH_3 group correlations are shown in red and the CH_2 group correlation is shown in blue.



(e) gHMBC spectrum of 9-methoxyrebeccamycin in DMSO-d_6 acquired with the multiple bond C-H coupling parameter (jnxh) set to 4 Hz

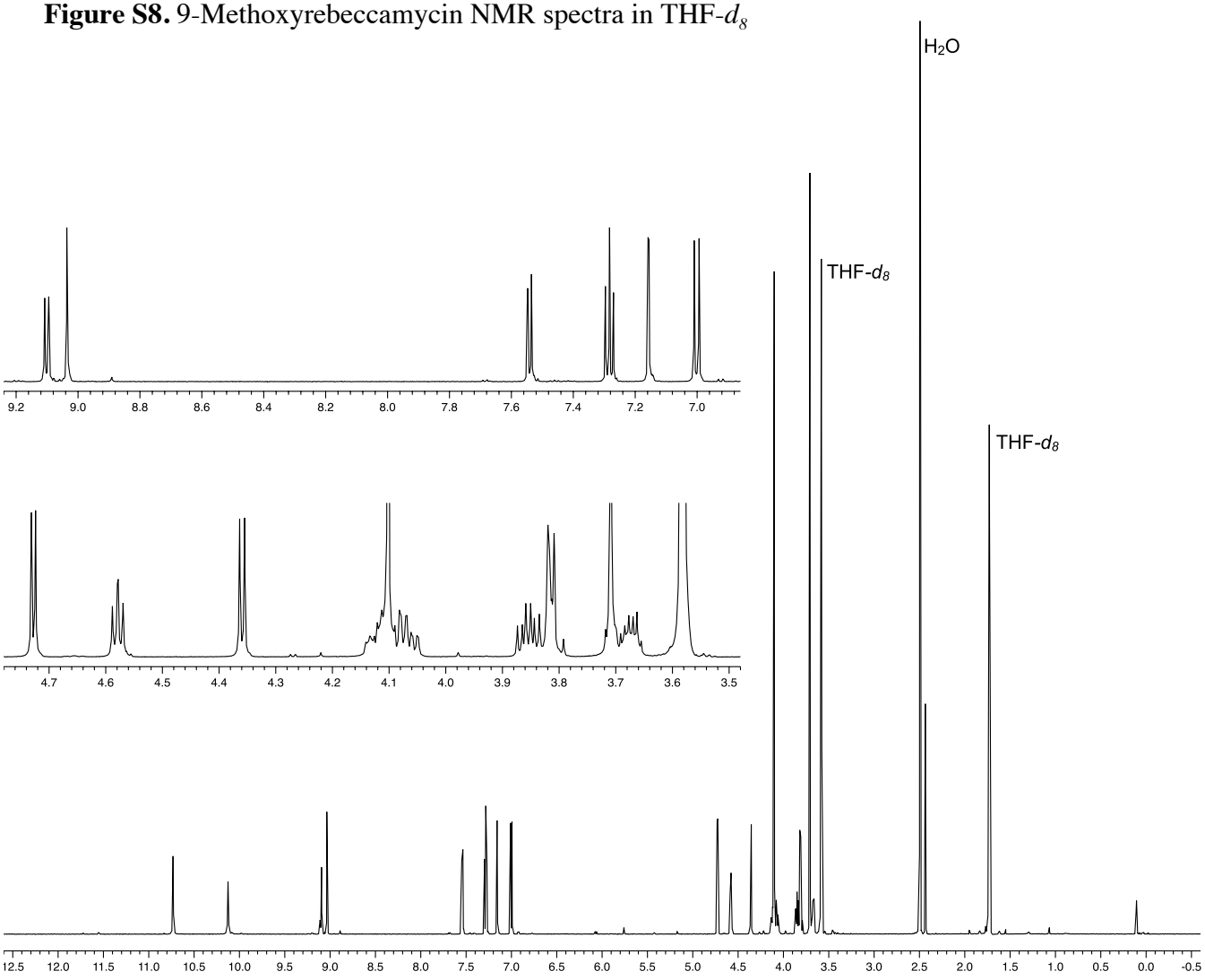


(f) gHMBC spectrum of 9-methoxyrebeccamycin in DMSO-d_6 acquired with the multiple bond C-H coupling parameter (jnxh) set to 8 Hz

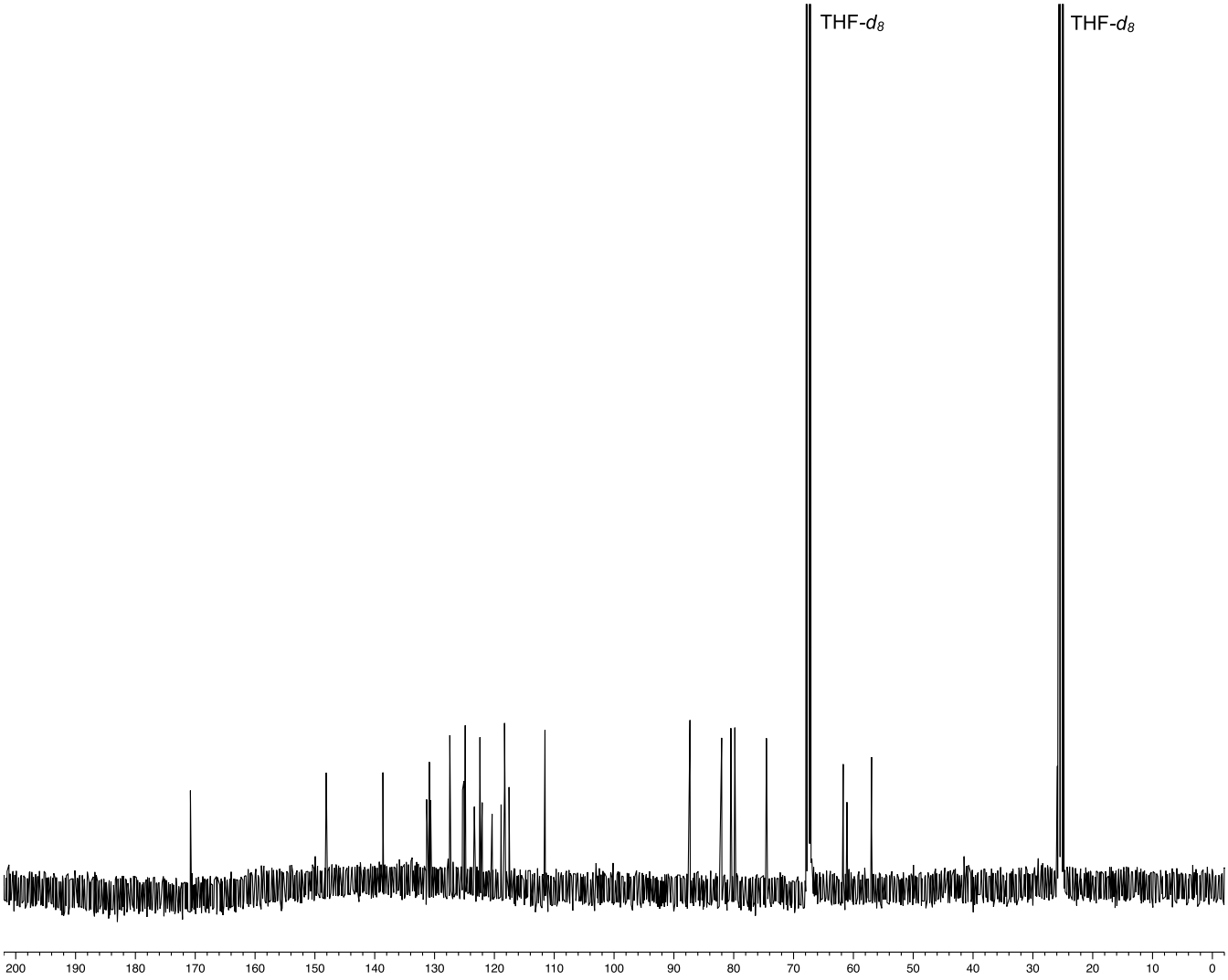


(g) *J*-resolved ^{1}H spectrum of 9-methoxyrebeccamycin in DMSO-d_6

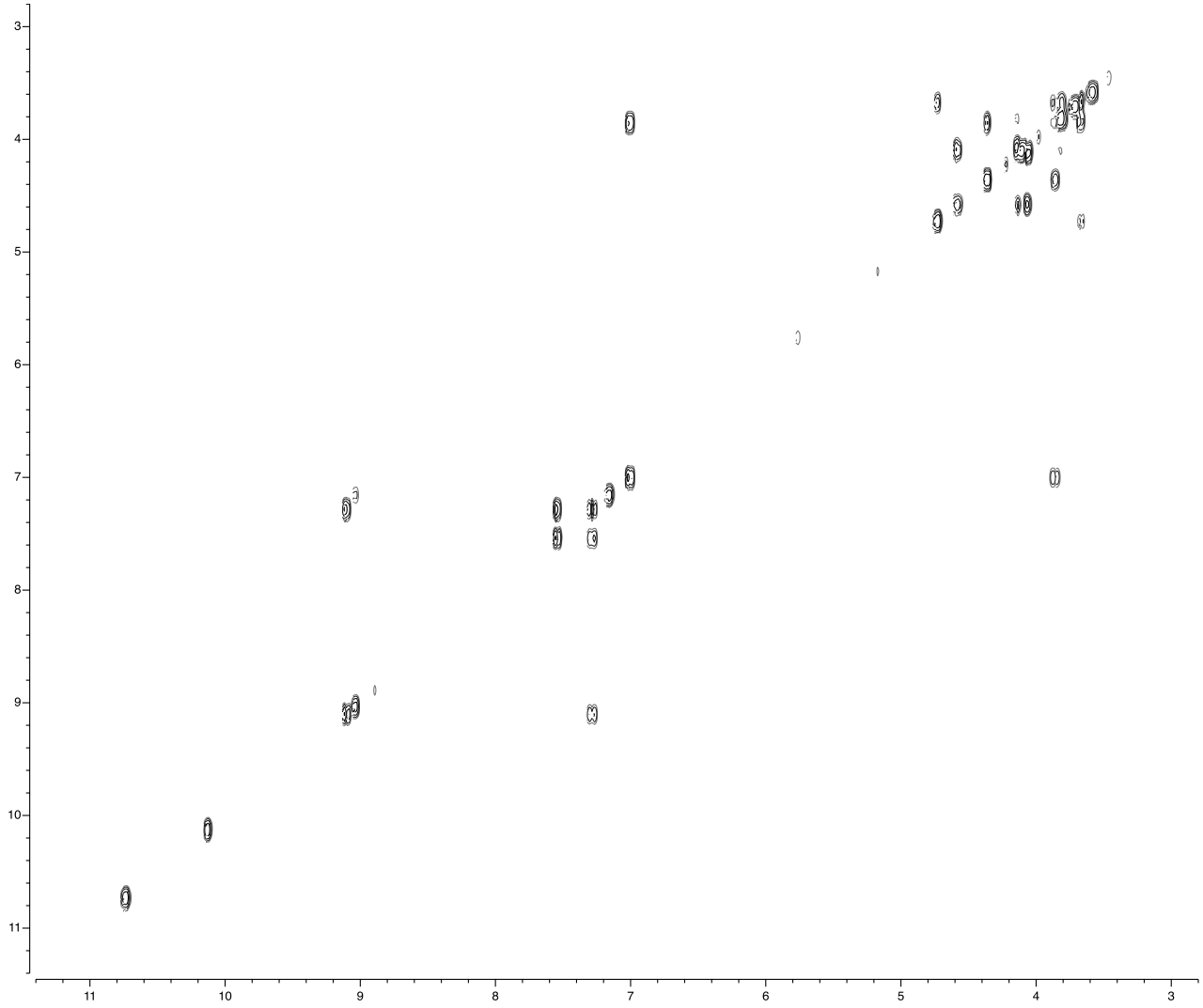
Figure S8. 9-Methoxyrebeccamycin NMR spectra in THF-d_8



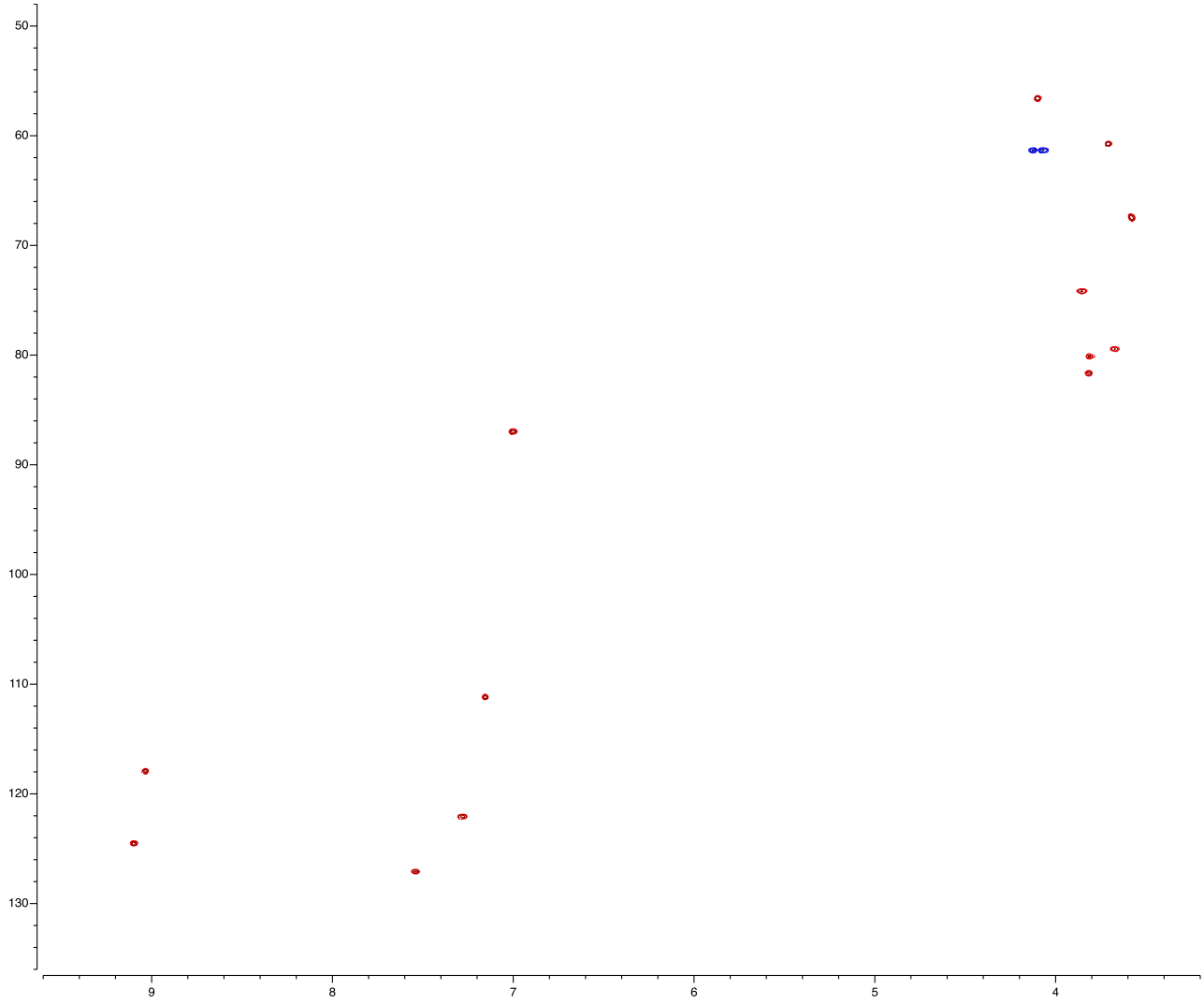
(a) ^{1}H NMR spectrum of 9-methoxyrebeccamycin in THF-d_8



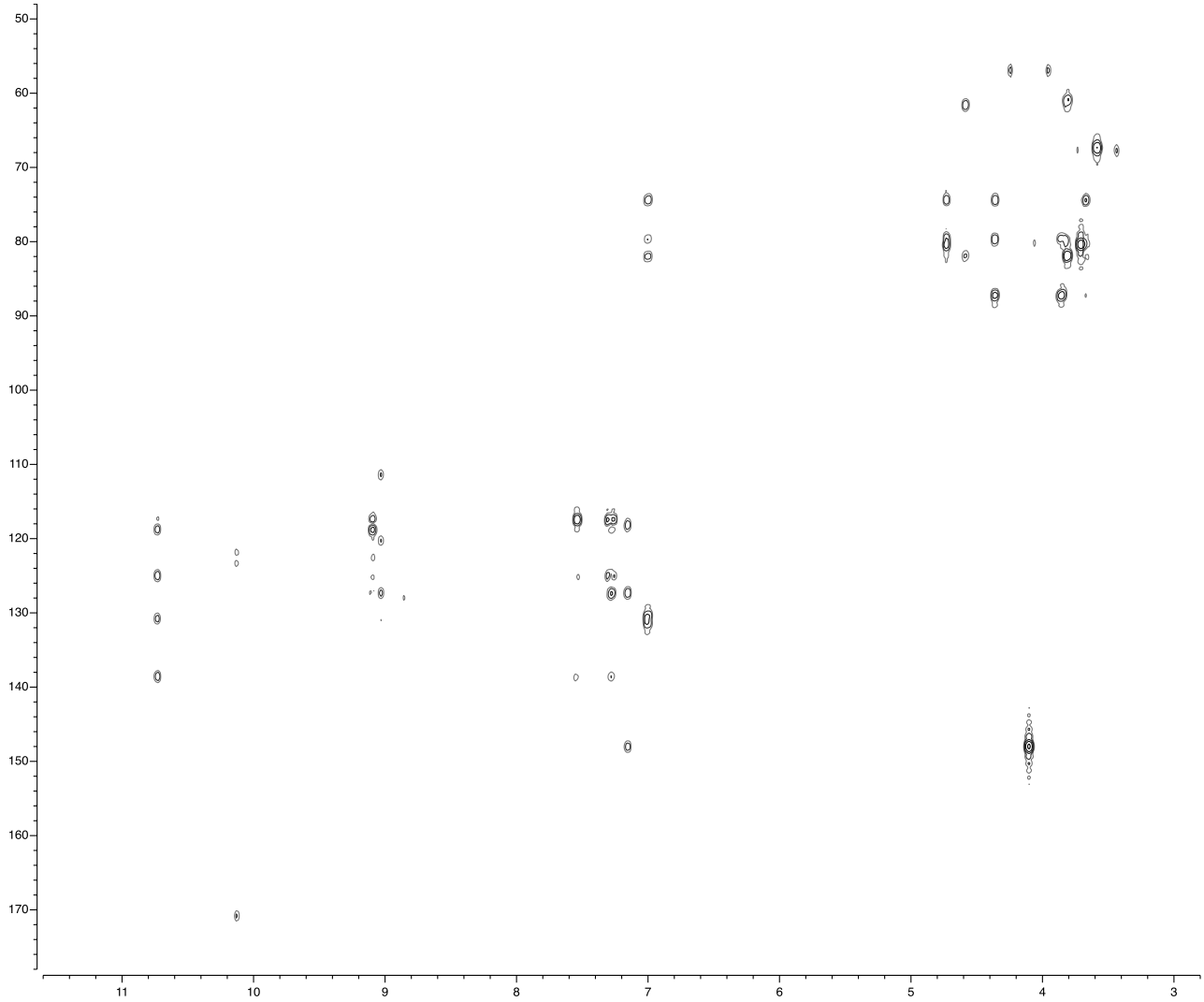
(b) ^{13}C NMR spectrum of 9-methoxyrebeccamycin in THF-d_8



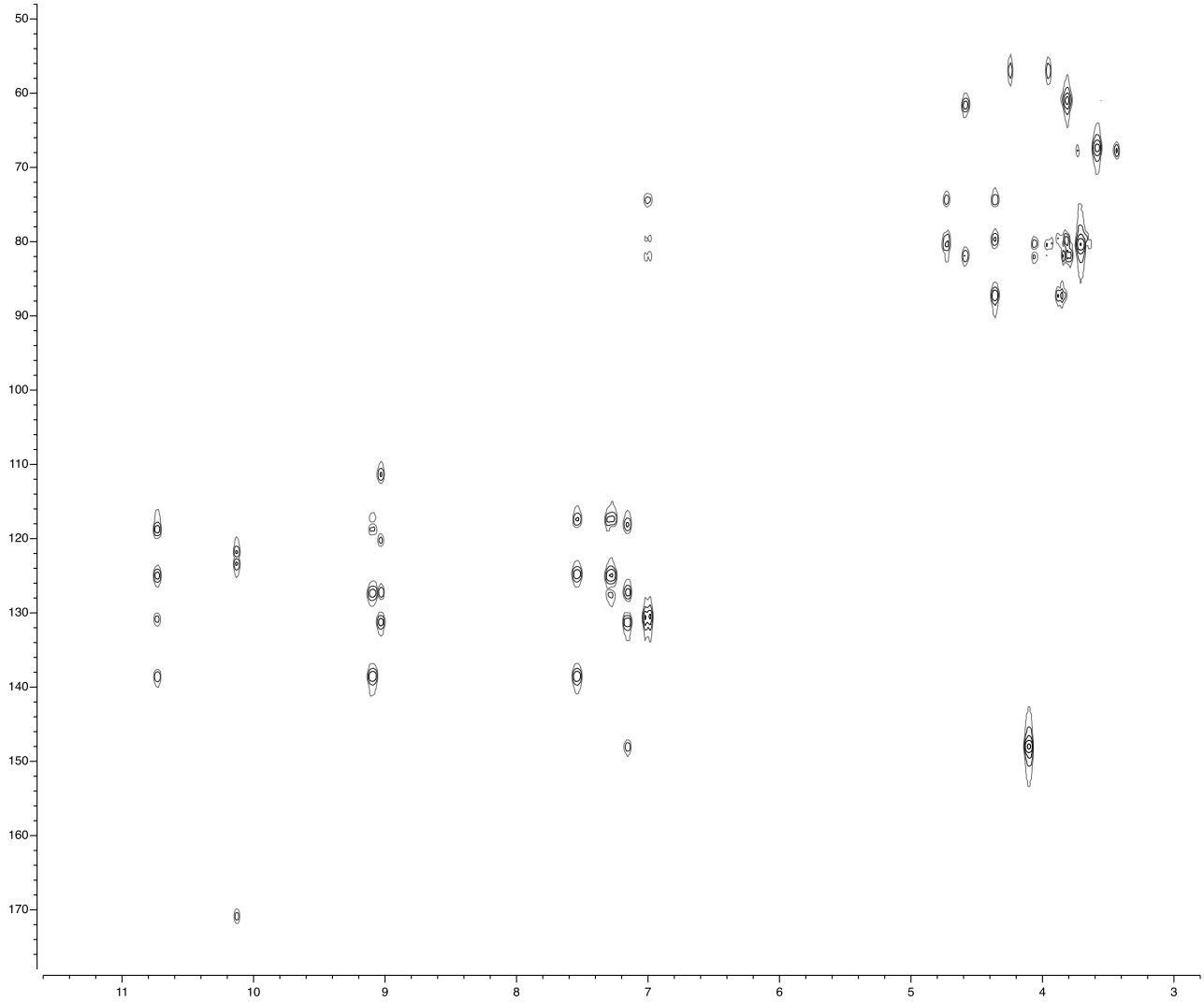
(c) gCOSY spectrum of 9-methoxyrebeccamycin in THF-d_8



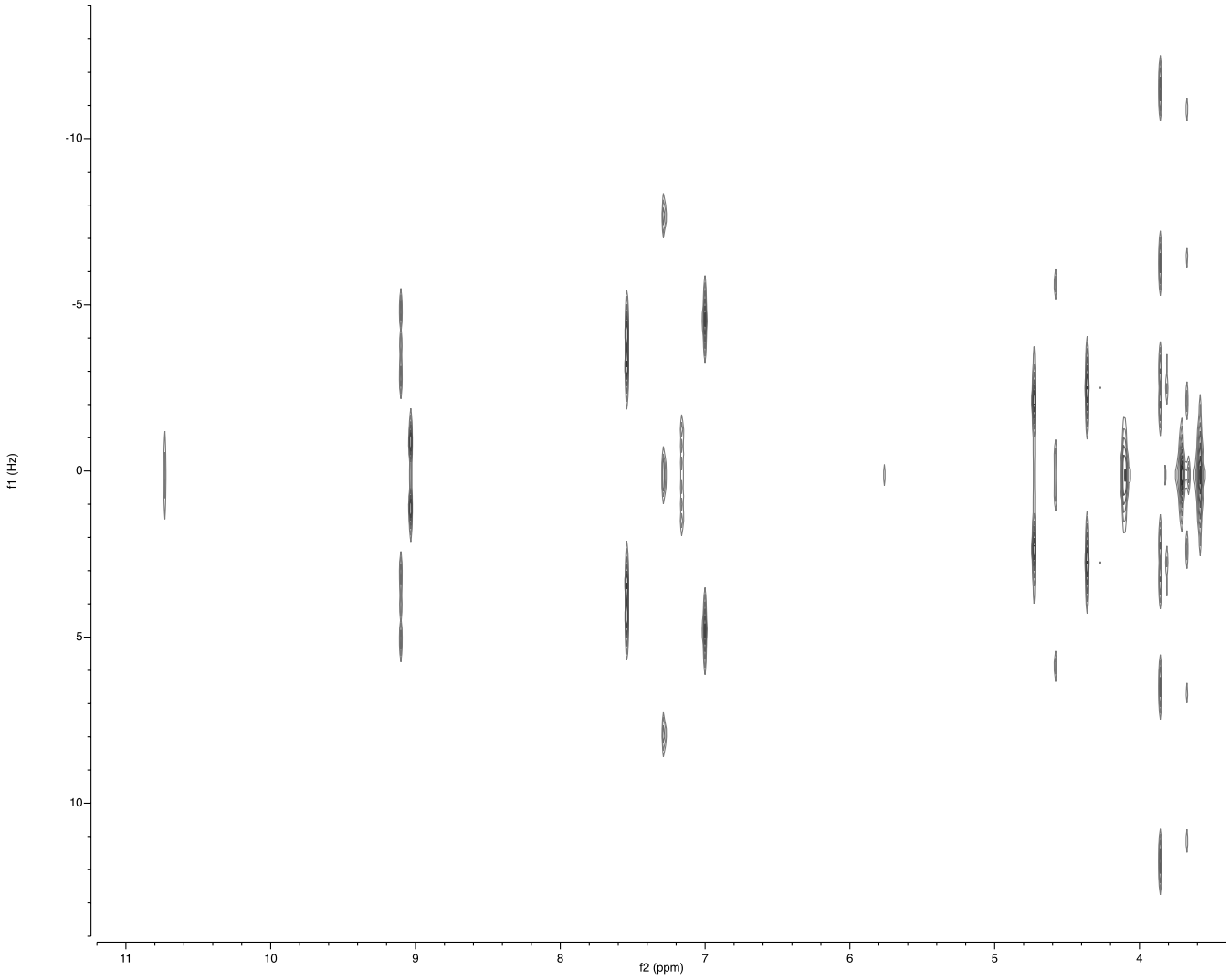
(d) Multiplicity-edited HSQC NMR spectrum of 9-methoxyrebeccamycin in THF-d_8. CH and CH_3 group correlations are shown in red and the CH_2 group correlation is shown in blue.



(e) gHMBC spectrum of 9-methoxyrebeccamycin in THF-d_8 acquired with the multiple bond C-H coupling parameter (jnxh) set to 4 Hz



(f) gHMBC spectrum of 9-methoxyrebeccamycin in THF-d_8 acquired with the multiple bond C-H coupling parameter (jnxh) set to 8 Hz



(g) *J*-resolved ^{1}H spectrum of 9-methoxyrebeccamycin in THF-d_8

Table S6. Pairwise comparison of conserved BCI *Pseudonocardia* replicons

	GGDC Probability DDH ≥70%			
replicon	**Formula 1**	**Formula 2**	**Formula 3**	**Two-way ANI**
chromosome	99.61%	98.26%	99.99%	99.98%
megaplasmid	99.58%	96.26%	99.98%	99.96%

Figure S9. Comparison of conserved BCI *Pseudonocardia* replicon architecture. Conserved sequence blocks are colored blue and red for BCI1 and BCI2, respectively, within the inner or outer rings based on the direction of the sequence. The circular plasmids are virtually identical, the circular megaplasmids are essentially identical although with large scale rearrangements.

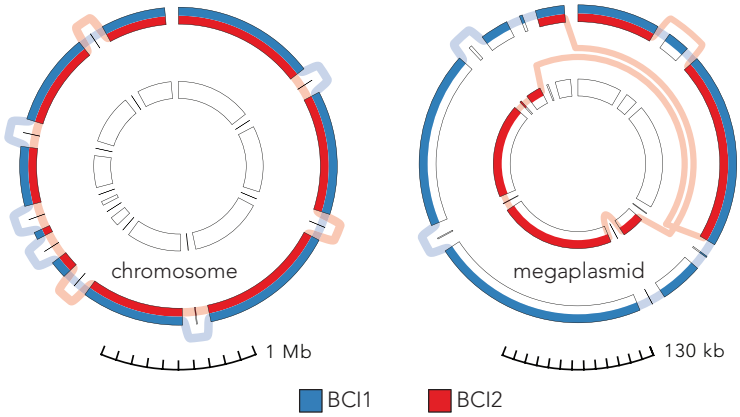


Figure S10. Relatedness of RebD homologs from rebeccamycin-like biosynthetic gene clusters. Red colored nodes indicate the presence of putative enzymes required to install the 9-methoxy group. The producing organisms or eDNA sample is named along with known natural products that have been isolated.

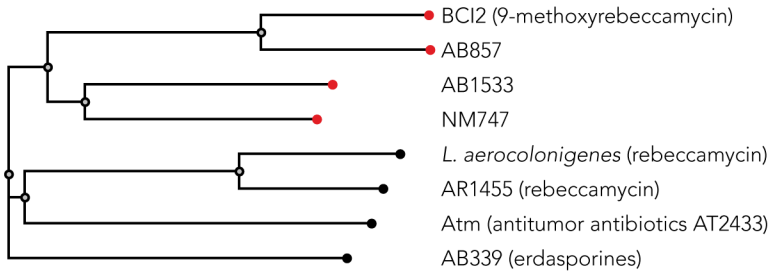


Table S7. Protein sequence comparison of 9-methoxyrebeccamycin biosynthetic gene cluster and two previously reported biosynthetic gene clusters*

Figure X annotation	blastp *v.* refseq	Pairwise identity with other reported Reb enzymes KF551866.1 AB857 cluster	AJ414559 *reb* cluster
RebM	methyltransferase type 11, WP_029021590.1, *Salinispora arenicola*	RebM (70%)	RebM (50%)
efflux	hypothetical, WP_048545302.1, *Tetrasphaera jenkinsii*	putative transmembrance efflux protein (75%)	RebT (28%)
	hypothetical, WP_048545304.1, *Tetrasphaera jenkinsii*	MarR family protein (71%)	ND
9-OMe	hypothetical, WP_044851327.1, Amycolatopsis orientalis	O-methyltransferase (67%)	ND
	serine/threonine protein kinase, WP_035923158.1, *Frankia* sp. QA3	putative helicase (54%)	ND
	hypothetical, WP_024794391.1, *Tomitella biformata*	putative SWIM Zn-finger (61%)	ND
	hypothetical, WP_048910367.1, *Streptomyces* sp. NRRL WC-3744	hypothetical, AHE14659.1 (68%)	ND
9-OH	tryptophan halogenase, WP_030431039.1, *Allokutzneria albata*	tryptophan 5-halogenase (84%)	RebH (39%)
RebP	cytochrome P450, WP_018742590.1, *Salinispora pacifica*	RebP (75%)	RebP (50%)
RebC	FAD-binding monooxygenase, WP_045932897.1, *Streptomyces* sp. NRRL B-1568	RebC (76%)	RebC (58%)
RebD	polyketide synthase, WP_025359114.1, *Kutzneria albida*	RebD (78%)	RebD (53%)
RebO	amine oxidase, WP_042222626.1, *Kutzneria albida*	RebO (78%)	RebO (67%)
RebG	glycosyl transferase, WP_031128186.1, *Streptomyces* sp. NRRL WC-3719	RebG (76%)	RebG (60%)
RebU	sodium:proton exchanger, WP_030926171.1, *Streptosporangium amethystogenes*	sodium (67%)	RebU (32%)
RebH	tryptophan halogenase, WP_040684713.1, *Nocardiopsis halotolerans*, 0, 67%/95%	RebH (84%)	RebH (72%)
RebF	oxidase, WP_016335659.1, *Amycolatopsis orientalis*	RebF (63%)	RebF (56%)
RebR	hypothetical, WP_017570355.1, *Nocardiopsis halotolerans*	RebR (56%)	RebR (37%)
	aminoglycoside phophostransferase, WP_043632787.1, *Nonomuraea candida*	ND	ND
	glycosyl transferase, WP_030063184.1, *Streptomyces natalensis*	ND	RebG (26%)
	hypothetical, WP_005202261.1, *Gordonia sputi*	ND	ND
	hypothetical, WP_015288214.1, *Mycobacterium cannettii*	ND	ND
	integrase, WP_037217458.1, *Rhodococcus* sp. R04	ND	ND

*predicted genes/pseudogenes and the gene products derived from sequences < 200 bp are omitted from the table; ND: not detected

Supplementary references

1. Poulsen M, Erhardt DP, Molinaro DJ, Lin T-L, Currie CR (2007) *PLoS ONE*. 2:e960.
2. Meier-Kolthoff JP, Auch AF, Klenk HP, & Goker M (2013) Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 14:60.
3. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. (2007) DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* 57:81-91.
4. Blin K, Medema MH, Kazempour D, Fischbach MA, Breitling R, Takano E, Weber T. (2013) antiSMASH 2.0 - a versatile platform for genome mining of secondary metabolite producers. *Nucleic Acids Res* 41:W204-212.
5. Chang FY, Ternei MA, Calle PY, & Brady SF (2013) Discovery and synthetic refactoring of tryptophan dimer gene clusters from the environment. *J Am Chem Soc* 135:17906-17912.
6. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28:1647-1649.