Coibacins A-D, Anti-leishmanial Marine Cyanobacterial Polyketides with Intriguing Biosynthetic Origins

Marcy J. Balunas, Manuel F. Grosso, Francisco A. Villa, Niclas Engene, Kerry L. McPhail, Kevin Tidgewell, Laura M. Pineda R., Lena Gerwick, Carmenza Spadafora, Dennis E. Kyle, and William H. Gerwick*
wgerwick@ucsd.edu

Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California San Diego, La Jolla, California 92093, Smithsonian Tropical Research Institute, Ancón, Panama City, Panamá, Instituto de Investigaciones Científicas y Servicios de Alta Technología, Ciudad del Saber, Clayton, Panama City, Panamá, College of Pharmacy, Oregon State University, Corvallis, Oregon 97331, Department of Global Health, College of Public Health, University of South Florida, Tampa, Florida 33612

Contents

S1: General Experimental Procedures

Optical rotations were obtained with a Jasco P-2000 polarimeter. UV and IR spectra were recorded on Beckman Coulter DU-800 and Nicolet IR100 FT-IR spectrophotometers, respectively. Circular dichroism spectra were obtained on a Jasco J-815 CD spectrometer. NMR spectra were recorded on a JEOL Eclipse 400 MHz spectrometer and referenced to residual solvent signals (δ_H 7.26, δ_C 77.2 for $CDCl_3$). High resolution mass spectra were acquired on an Agilent ESI-TOF mass spectrometer. HPLC purifications of natural product isolates were carried out on a Merck Hitachi LaChrom HPLC system with a L-7100 pump, L-7614 degasser, and L-7455 diode array detector using a Prontosil-120 C_{18} (4.6 × 250 mm, 5μM) RP-HPLC column.

S2: Collection and Taxonomy

A cyanobacterium (Figure S3) was collected in March 2007 by hand using snorkeling methods from rocks in a bay near Uvas Island in the Coiba National Park (8^o 04.667 N, 82^o 10.995 W) in the Republic of Panama. After straining through a mesh bag to remove seawater, the sample was stored at 4 ^{o}C. Voucher specimen number PAC-03/01/07-2 is deposited at the Gerwick Lab sample repository, University of California San Diego, San Diego, California.

In February 2010, two additional coibacin-producing specimens of cyanobacteria (PAC-18-Feb-10-1 and PAC-18-Feb-10-2) were collected from similar shallow-water habitats of the rocky shore surf-zone of Uvas Island, Panama. The overall growth morphologies were distinctly different between the two specimens; PAC-18-Feb-10-1 had a thin mat-like growth morphology, while the strain PAC-18-Feb-10-2 formed a tufty thallus. Microscopically, however, the filaments of the two specimens were similar in appearance with disk-shaped cells covered by thin polysaccharide sheaths, and corresponded taxonomically with morphologically overlapping genera such as *Moorea*, *Lyngbya* or *Oscillatoria*. In addition to the "*Lyngbya*"-type filaments (PAC-18-Feb-10-1.1), the specimen PAC-18-Feb-10-1 also contained finer filaments with barrel-shaped cells (PAC-18-Feb-10-1.2). The facts that PAC-18-Feb-10-1.2 only represented a minor microbial component of the assemblage, and that both specimens produced coibacin, suggests that this associated cyanobacterium was not the metabolic producer of these compounds.

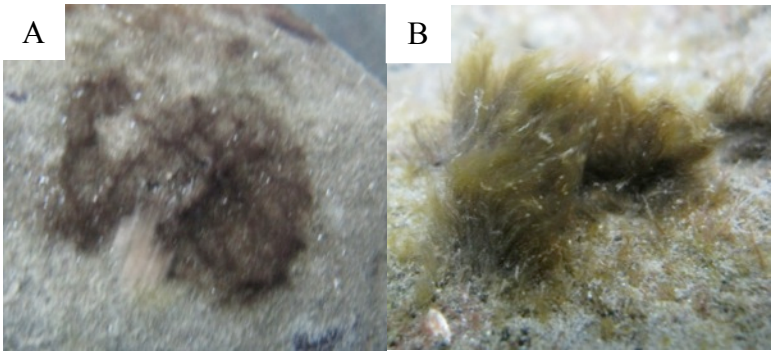


Figure S3. Photos of PAC-18-Feb-10-1 and -2 (A and B, respectively), the cf. *Oscillatoria* sp. from which the coibacins were isolated.

Phylogenetically, the two strains PAC-18-Feb-10-1.1 and PAC-18-Feb-10-2 formed a distinct, tight clade nesting between the genera *Oscillatoria* (reference strain = PCC 7515) and *Trichodesmium* (reference strain = IMS 101). This clade includes several important tropical marine NP-producing cyanobacterial strains, including the veraguamide-producing strain PAC-17-FEB-10-2 (HQ900689), the tumonoic acid-producing strain PNG05-4 (EU253968), and the microcolin-producing strain LP16 (FJ602745). Because of taxonomic ambiguity, specimens of this group have been described as *Oscillatoria*, *Blennothrix* and *Lyngbya* among others. The phylogenetic and ecological diversification of this clade with the closest related genera *Oscillatoria* and *Trichodesmium* strongly suggests that this clade represent a unique group in need of description. However, prior to taxonomic description we prefer to refer to specimens of this group as cf. *Oscillatoria*. Interestingly, the two strains PAC-18-Feb-10-1.1 and PAC-18-Feb-10-2 formed two separate sister-clades with a p-distance of 1.8% SSU rRNA gene sequence divergence. This genetic divergence suggests that these two coibacin-producing strains may represent different species, which could also explain the distinct growth morphologies between the specimens. Moreover, their evolutionary distance suggests an alternative evolution of their coibacin-producing capabilities, either by horizontal gene transfer (HGT) or distinct metabolic origin of the molecules such as associated heterotrophic bacteria.

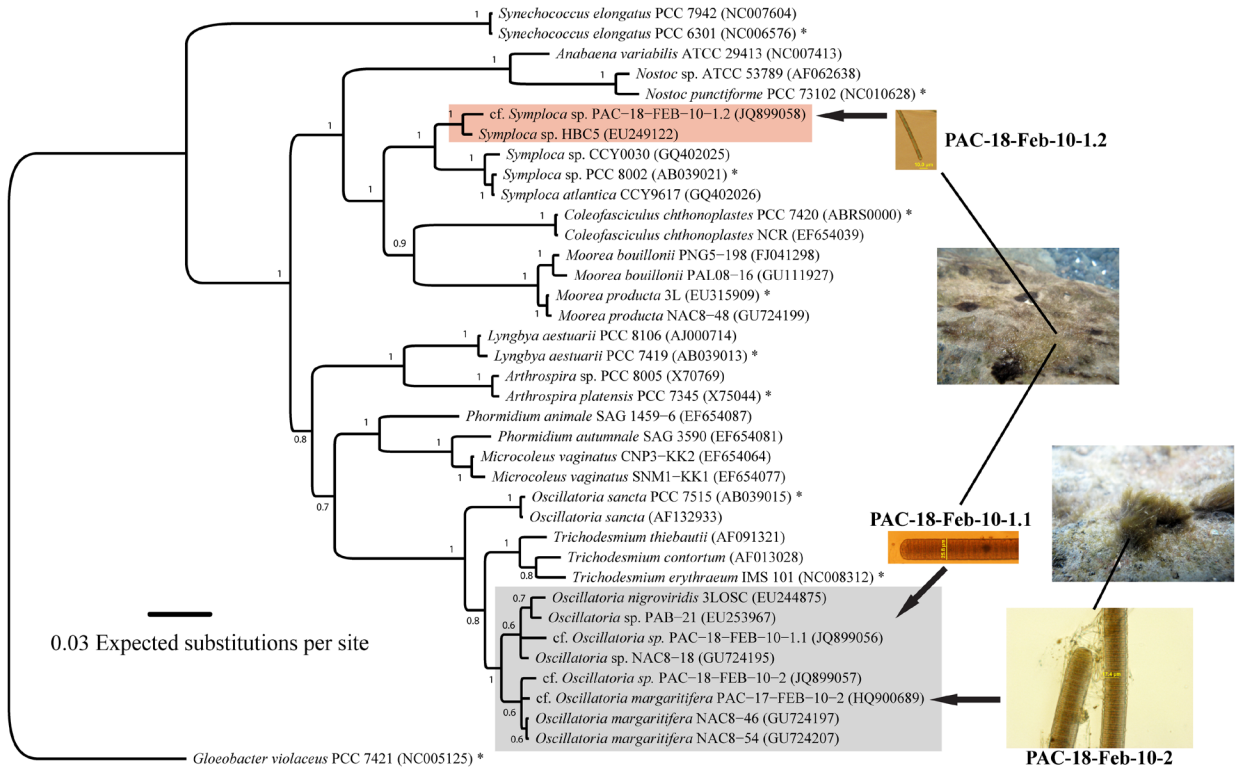


Figure S4. Evolutionary tree of the coibacin-producing cyanobacterial specimens PAC-18-Feb-10-1 and PAC-18-Feb-10-2. Note that the specimen PAC-18-Feb-10-1 is a consortium of two different types of filamentous cyanobacteria (PAC-18-Feb-10-1.1 and PAC-18-Feb-10-1.2 - highlighted by arrows) with PAC-18-Feb-10-1.1 as the predominant component. The strains PAC-18-Feb-10-1.1 and PAC-18-Feb-10-2 form a distinct clade (highlighted with a grey box) nesting between the evolutionarily closest related genera *Trichodesmium* and *Oscillatoria* (type-strains IMS 101 and PCC 7515, respectively). Moreover, PAC-18-Feb-10-1.1 and PAC-18-Feb-10-2 form two separate sister-clades (p-distance = 1.8% gene sequence divergence). The strain PAC-18-Feb-10-1.2 is phylogenetically closest to the genus *Symploca* (reference strain: PCC 8002^{T}, GenBank acc. nr. AB039021) with a p-distance of 4.9% based on the SSU rRNA gene sequence divergence (highlighted with a red box). The phylogram is based on SSU (16S) rRNA gene sequences using the bayesian (MrBayes) and maximum likelihood (PhyML) methods, and the support values are indicated as posterior probability at the nodes. The specimens are indicated as species, strain and access number in parentheses. Specimens designated with (^{T}) represent type-strains; reference-strains obtained from *Bergey's Manual* are designated with an asterisk. The scale bar is indicated at 0.03 expected nucleotide substitutions per site, corrected using the General Time Reversal (GTR) model.

S5: Morphological Characterization

Morphological characterization was performed using an Olympus IX51 epifluorescent microscope (1000X) equipped with an Olympus U-CMAD3 camera. Measurements were provided as: mean ± standard deviation (SD). The filament means were the average of three filament measurements, and cell measurements the average of ten adjacent cells of three filaments. Morphological comparison and putative taxonomic identification of the cyanobacterial specimen was performed in accordance with modern classification systems (Castenholz *et al*., 2001; Komárek & Anagnostidis, 2005).

S6: Gene Sequencing

Cyanobacterial specimens were preserved for genetic analysis both as live material and in 10 mL RNA*later* (Ambion). Algal biomass (~50 mg) was partly cleaned under an Olympus VMZ dissecting microscope. Genomic DNA was extracted using the Wizard® Genomic DNA Purification Kit (Promega) following the manufacturer's specifications. DNA concentration and purity was measured on a DU® 800 spectrophotometer (Beckman Coulter). The 16S rRNA genes were initially PCR-amplified from isolated DNA using the general primer set 106F and 1509R (Nubel et al.) while consequent PCR-reactions were performed using the modified lineage-specific primers: OT106F 5'-GGACGGGTGAGTAACGCGTGA-3' and OT1445R 5'-AGTAATGACTTCGGGCGTG-3' for the "*Oscillatoria*" specimens and 106F and CSL1445R 5'-GGTAACGACTTCGGGCGTG-3' for the "*Symploca*" specimens. The PCR reaction volumes were 25 µL containing 0.5 µL (~50 ng) of DNA, 2.5 µL of 10 × PfuUltra IV reaction buffer, 0.5 µL (25 mM) of dNTP mix, 0.5 µL of each primer (10 µM), 0.5 µL of PfuUltra IV fusion HS DNA polymerase and 20.5 µL dH_2O. The PCR reactions were performed in an Eppendorf® Mastercycler® gradient as follows: initial denaturation for 2 min at 95 °C, 25 cycles of amplification, followed by 20 sec at 95 °C, 20 sec at 55 °C and 1.5 min at 72 °C, and final elongation for 3 min at 72 °C. PCR products were analyzed on a (1%) agarose-gel in SB-buffer and visualized by EtBr-staining. The PCR products were purified using a MinElute® PCR Purification Kit (Qiagen) before subcloning using the Zero Blunt® TOPO® PCR Cloning Kit (Invitrogen) following the manufacturer's specifications. Plasmid DNA was isolated using the QIAprep® Spin Miniprep Kit (Qiagen) and sequenced with M13 primers. Correlation between gene sequences and cyanobacterial filaments in the cyanobacterial assemblage was verified by

careful dissection of single filaments under an Olympus VMZ dissecting microscope followed by multiple displacement amplification (MDA). The single filaments were washed twice in 0.5 µL sterile SWBG-11 medium and twice in 0.5 µL dH_2O before transfer into 0.2 mL PCR-tubes. DNA was amplified from the single filaments by MDA using the REPLI-g® Mini Kit (Qiagen), following the manufacturer's specifications. All MDA reactions were performed in 50 µL reaction volume for 16 h at 30 °C. The 16S rRNA gene sequences are available in the DDBJ/EMBL/GenBank databases under accession numbers JQ899056- JQ899058.

S7: Phylogenetic Inference

The 16S rRNA gene sequences of PAC-18-Feb-10-1.1, PAC-18-Feb-10-1.2 and PAC-18-Feb-10-2 were aligned with evolutionarily informative cyanobacteria using the L-INS-I algorithm in MAFFT 6.717 (Katoh & Toh, 2008) and refined using the SSU secondary structures model for *Escherichia coli* J01695 (Cannone *et al*., 2002) without data exclusion. The best-fitting nucleotide substitution model optimized by maximum likelihood (ML) was selected using corrected Akaike/Bayesian Information Criterion (AIC_C/BIC) in jModeltest 0.1.1 (Posada, 2008). The evolutionary histories of the cyanobacterial genes were inferred using ML and Bayesian inference (BI) algorithms. The ML inference was performed using GARLI 1.0 (Zwickl, 2006) for the GTR+I+G model assuming heterogeneous substitution rates and gamma substitution of variable sites (proportion of invariable sites (pINV) = 0.516, shape parameter (α) = 0.446, number of rate categories = 4) with 1,000 bootstrap-replicates. The Bayesian inference was conducted using MrBayes 3.1 (Ronquist & Huelsenbeck, 2003) with four Metropolis-coupled MCMC chains (one cold and three heated) run for 1,000,000 generations. The first 25 % were discarded as burn-in and the following data set was sampled with a frequency of every 100 generations.

S8: Extraction and Isolation

The sample (41.9 g dry weight) was thawed and repetitively extracted 7 times with 2:1 CH_2Cl_2/MeOH to afford, after solvent evaporation, 871.9 mg of crude organic extract. The extract was fractionated using flash Si gel column chromatography (Aldrich, Si gel 60, 230-400 mesh, 40 x 180 mm) using 300 mL each of 100% hexanes (A), 9:1 hexanes/EtOAc (B), 4:1

hexanes/EtOAc (C), 3:2 hexanes/EtOAc (D), 2:3 hexanes/EtOAc (E), 1:4 hexanes/EtOAc (F), 100% EtOAc (G), 3:1 EtOAc/MeOH (H), and 100% MeOH (I).

Fraction D showed strong antileishmanial activity and was subjected to further fractionation using a Burdick & Jackson C_{18} RP-SPE cartridge with a 0.45 μm nylon filter eluted with 100% MeOH. The eluent was chromatographed by RP-HPLC (68% MeCN/32% H_2O, 1.0 mL/min) to yield coibacin A (**1**, 18.0 mg, t_R 58.5 min) and two impure fractions that were fractionated further. The first of these fractions was chromatographed by RP-HPLC (65% MeOH/35% H_2O, 1.0 mL/min) to afford coibacin C (**3**, 1.3 mg, t_R 36.5 min). The second fraction was chromatographed by RP-HPLC (69% MeOH/31% H_2O, 1.0 mL/min) to afford coibacins B (**2**, 1.4 mg, t_R 34.5 min) and D (**4**, 3.6 mg, t_R 29.0 min).

Coibacin A: $[\alpha]_D$ = +46 (*c* 1 mg/mL, $CHCl_3$); UV (CH_3CN) λ_{max} (log ε) 204 (5.74), 231 (5.95), 239 (5.97) nm; IR (film) ν_{max} 3408, 2923, 2853, 2363, 1722, 1457, 1381, 1244, 1093, 986, 817 cm^{-1}; ^1H and ^{13}C NMR data, see Table 1; HRESIMS *m/z* $[M+Na]^+$ 307.1670 (calcd for $C_{19}H_{24}O_2Na$, 307.1669).

Coibacin B: $[\alpha]_D$ = +59 (*c* 1 mg/mL, $CHCl_3$); UV (CH_3CN) λ_{max} (log ε) 205 (6.19), 227 (6.26) nm; IR (film) ν_{max} 3422, 2923, 2363, 1721, 1449, 1381, 1244, 1053, 990, 961, 817 cm^{-1}; ^1H and ^{13}C NMR data, see Table 1; HRESIMS *m/z* $[M+Na]^+$ 281.1513 (calcd for $C_{17}H_{22}O_2Na$, 281.1512).

Coibacin C: $[\alpha]_D$ = +18 (*c* 1 mg/mL, $CHCl_3$); UV (CH_3CN) λ_{max} (log ε) 204 (5.62), 226 (5.61), 269 (4.60) nm; IR (film) ν_{max} 3447, 2923, 2854, 2362, 1718, 1453, 1384, 1251, 1091, 770 cm^{-1}; ^1H and ^{13}C NMR data, see Table 1; HRESIMS *m/z* $[M+Na]^+$ 289.0968 (calcd for $C_{15}H_{19}ClO_2Na$, 289.0966).

Coibacin D: $[\alpha]_D$ = +6 (*c* 1 mg/mL, $CHCl_3$); UV (CH_3CN) λ_{max} (log ε) 203 (6.36), 226 (5.36), 269 (4.26) nm; IR (film) ν_{max} 3448, 2922, 2852, 2362, 1722, 1384, 1247, 1037, 969, 816 cm^{-1}; ^1H and ^{13}C NMR data, see Table 1; HRESIMS *m/z* $[M+Na]^+$ 291.1124 (calcd for $C_{15}H_{21}ClO_2Na$, 291.1122).

S9: Bioassay Methods

Chemicals: Defined Fetal Bovine Serum (FBS) and penicillin-streptomycin were purchased from HyClone (Logan, UT). Dulbecco's modified Eagle medium (DMEM) was purchased from Gibco (Auckland, New Zealand). Lipopolysaccharide (LPS) (*Escherichia coli*,

serotype 026:B6), 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT), and sulfanilamide were purchased from Sigma-Aldrich (St. Louis, MO). N-(1-napthyl) ethylenediamine dihydrochloride (NED) was purchased from Ricca Chemical Company (Arlington, TX).

Cell culture: Cells from the mouse macrophage cell line RAW264.7 (ATCC; Manassas, VA) were cultured in DMEM with 4 mM L-glutamine and 4.5 g/l glucose supplemented with 10% FBS, penicillin, and streptomycin. Unless otherwise stated RAW264.7 cells were seeded in 96-well plates (5 x 10^4 cells/well), and after 1 day were stimulated with 3 μg/mL LPS in the absence or presence of pure compound (1 to 10 μg/mL) for 24 h in triplicate wells at 37 °C with 5% CO_2.

Nitrite assay: The generation of NO was assessed in the supernatant of cell cultures by quantification of nitrite using the Griess reaction (Green et al., 1982). In brief, 50 μl of each supernatant were added to 96-well plates together with 50 μl 1% sulfanilamide in 5% phosphoric acid plus 50 μl 0.1% NED in water, and the optical density was measured at 570 nm. The IC_{50} value, the sample concentration resulting in 50% inhibition of NO production, was determined using non-linear regression analysis (% nitrite versus concentration).

MTT assay: Cell viability was determined by mitochondrial-dependent reduction of MTT to formazan quantified at 570 and 630 nm. Cells were incubated with 1 mg/mL MTT at 37 °C for 25 min, the medium was aspirated, and resuspended in 100 μl DMSO for solubilization of the formazan dye. The percent survival was determined by comparison with the control group.

ELISA assay: RAW264.7 cells were seeded in 6-well plates (8 x 10^5 cells/well), and after 1 day each well was stimulated with LPS (3 μg/mL) in the absence or presence of coibacins A (10 μg/mL), B, C, and D (3 μg/mL) in triplicate at 37 °C with 5% CO_2. The secretion of mouse TNF-α and IL-6 were measured by ELISA using the Ready-Set-Go! kit obtained from eBioscience (San Diego, CA). Media from three separate wells of RAW264.7 cells after treatment with coibacin A and LPS for 6 h were serially diluted and assayed, and the optical density was measured at 450 nm.

Statistics: All experiments were repeated at least three times. Data are presented as the mean ± standard deviation for the indicated number of independently performed experiments. Student's t-test was used for the determination of statistical significance with $P < 0.05$ being considered significant.

Figure S10. ^{1}H NMR spectrum of Coibacin A (**1**) (in $CDCl_3$ at 400 MHz)

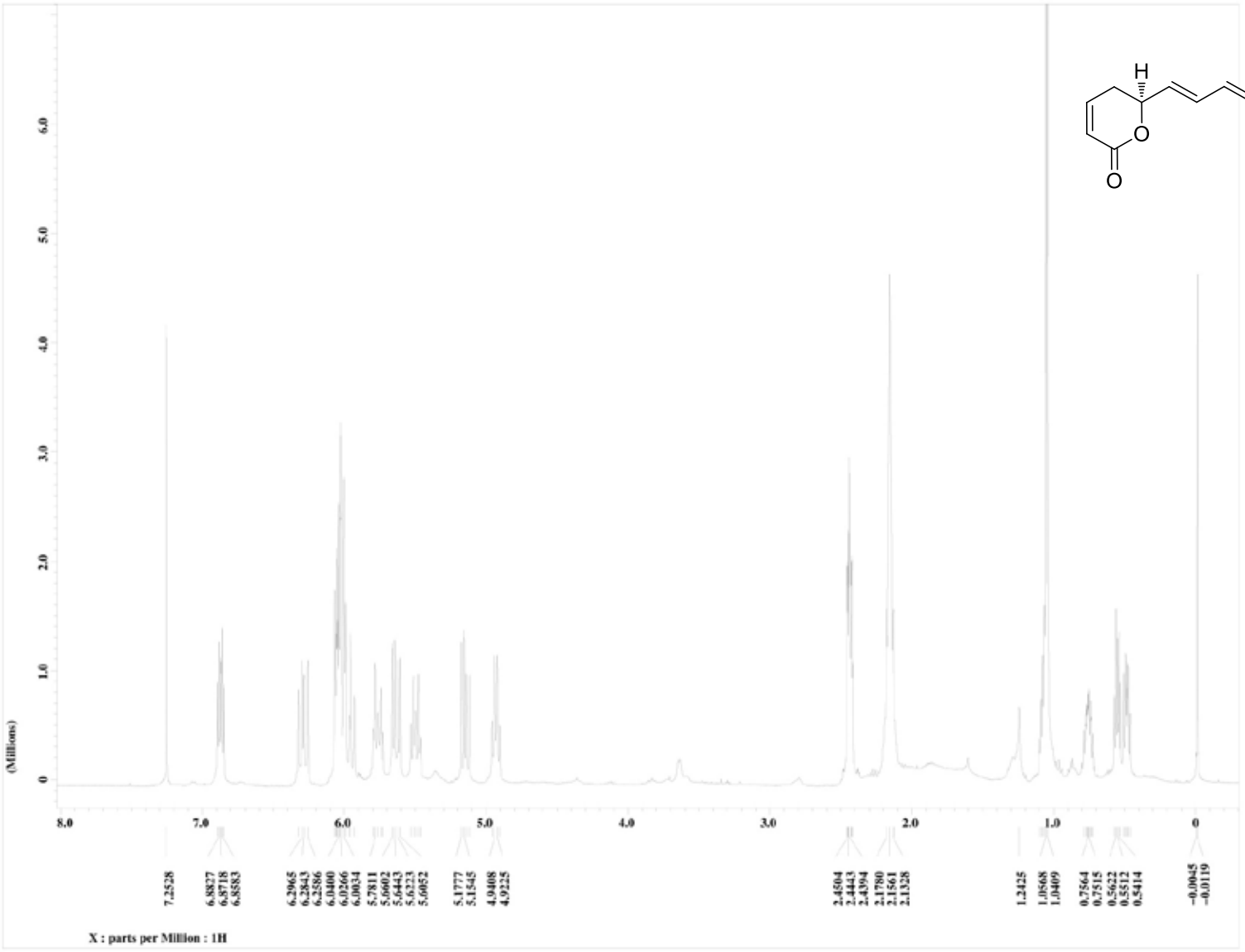


Figure S11. ^{13}C NMR spectrum of Coibacin A (**1**) (in $CDCl_3$ at 100 MHz)

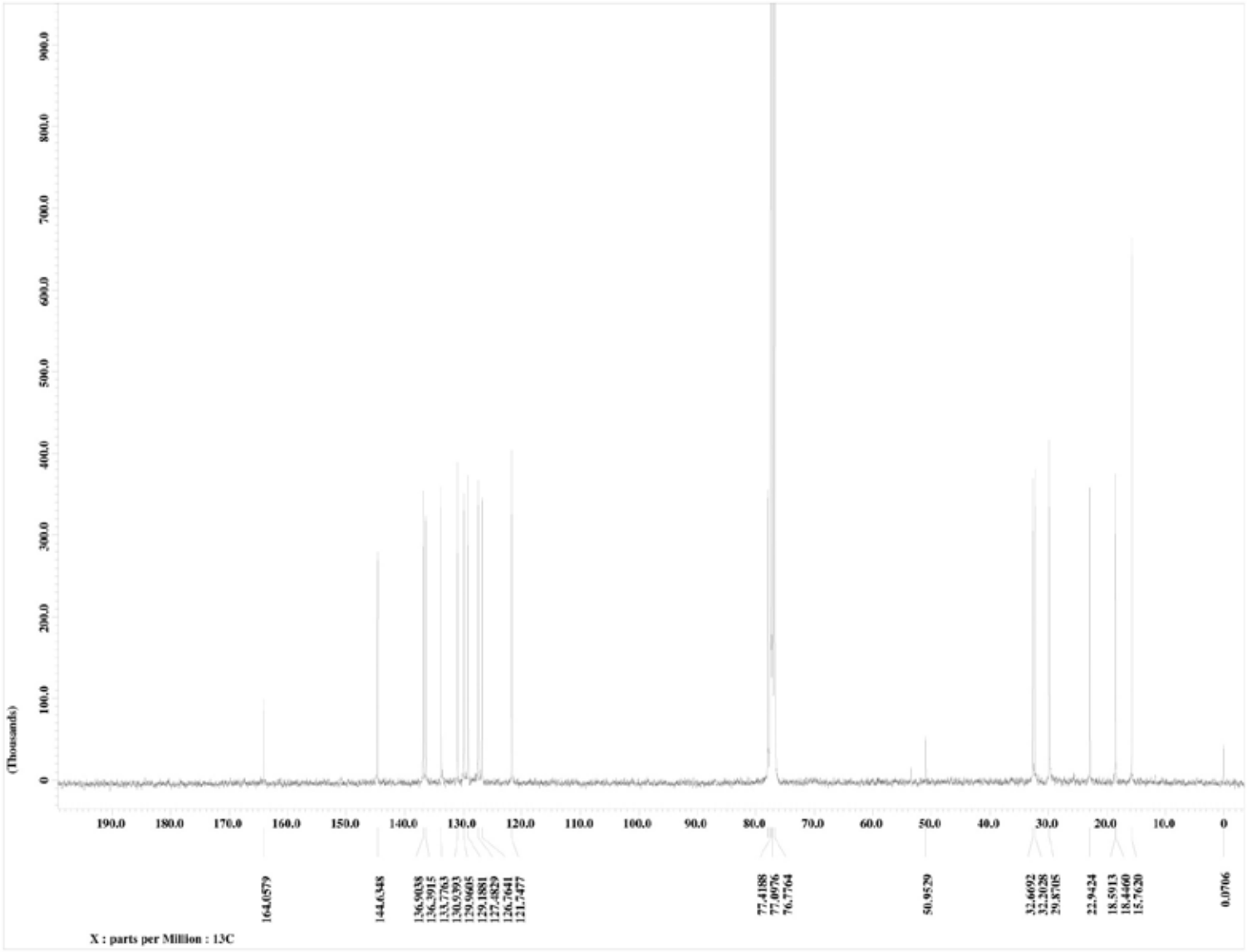


Figure S12. COSY spectrum of Coibacin A (**1**) (in $CDCl_3$ at 400 MHz)

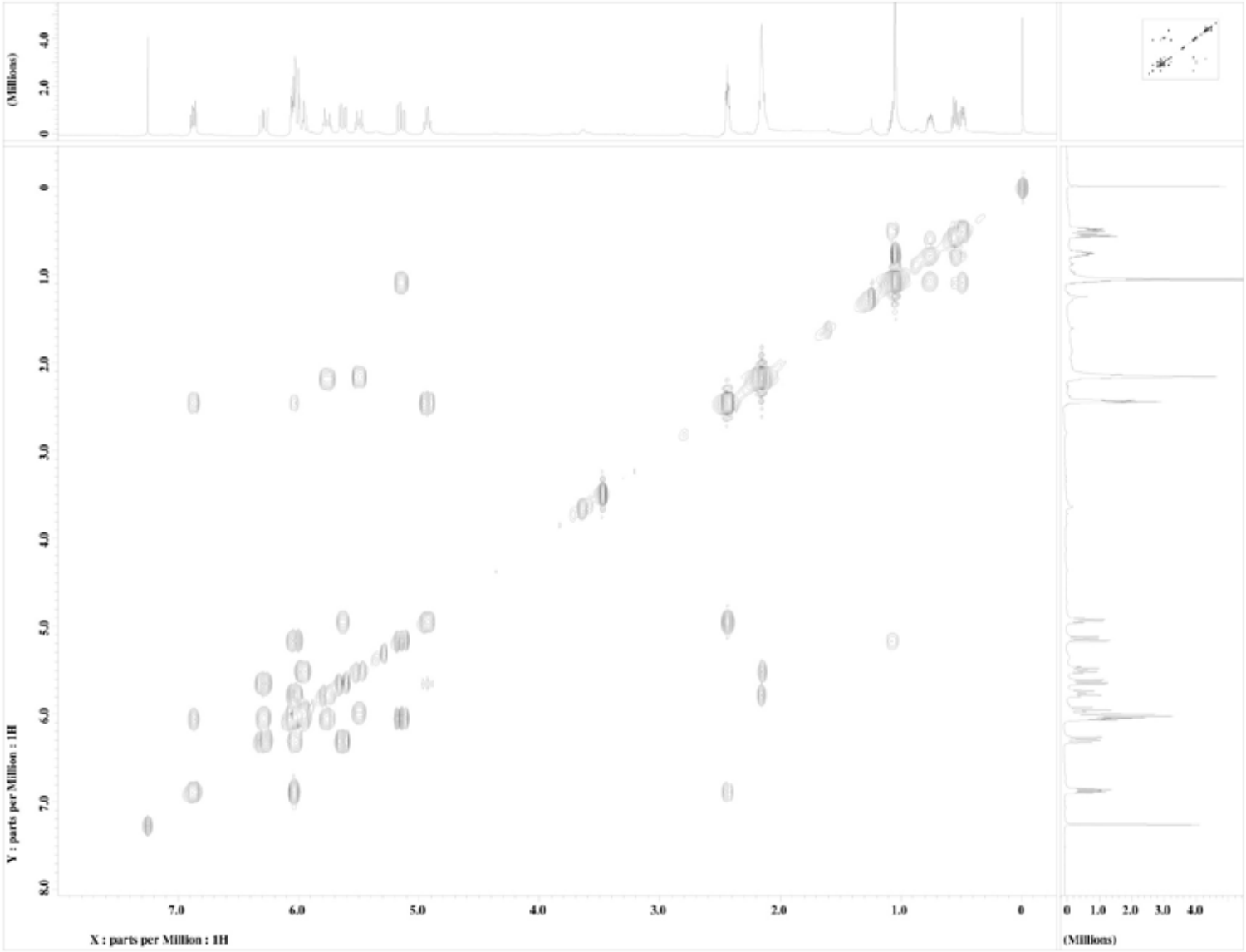


Figure S13. HSQC spectrum of Coibacin A (**1**) (in $CDCl_3$ at 400 MHz)

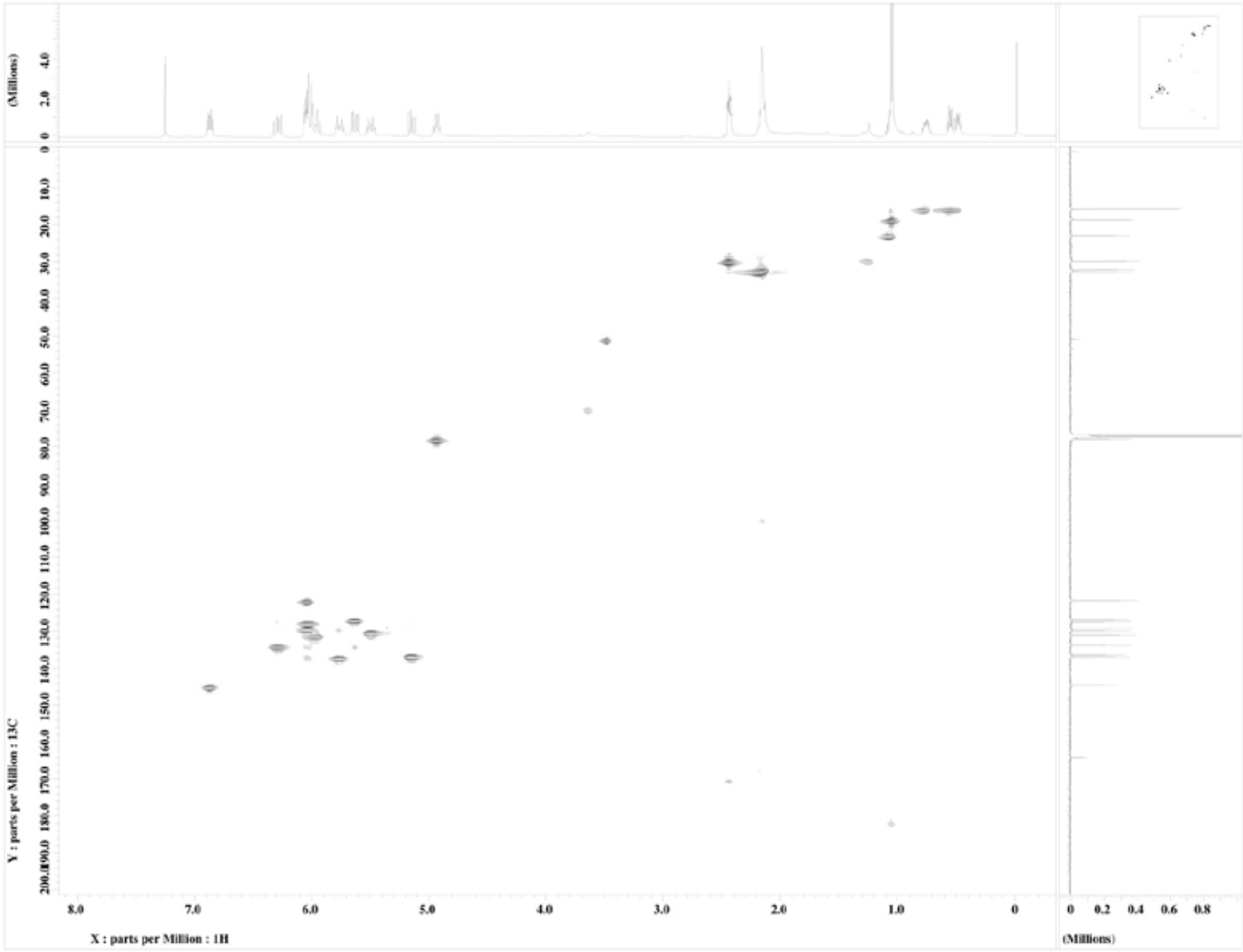


Figure S14. HMBC spectrum of Coibacin A (**1**) (in $CDCl_3$ at 400 MHz)

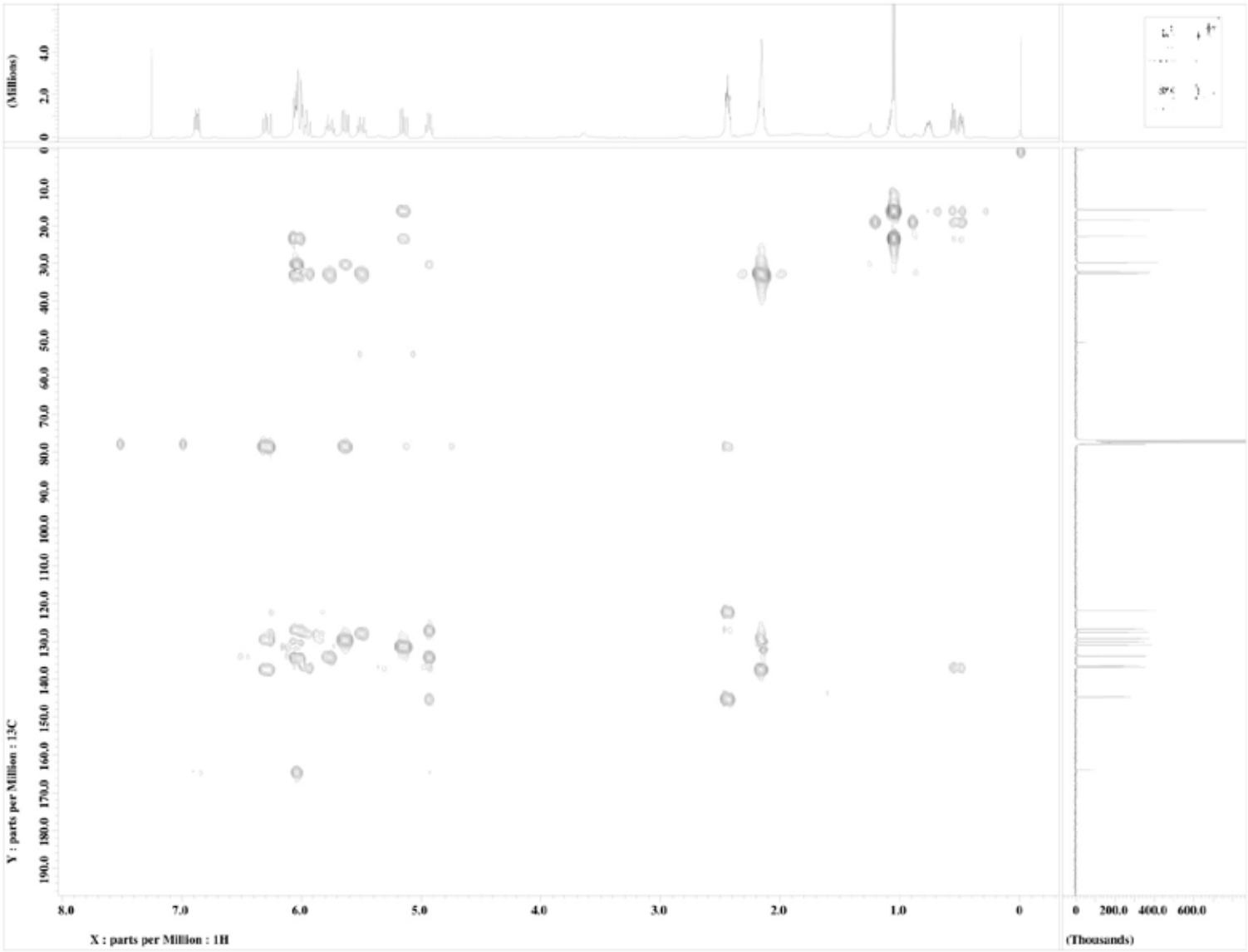


Figure S15. NOESY spectrum of Coibacin A (**1**) (in $CDCl_3$ at 400 MHz)

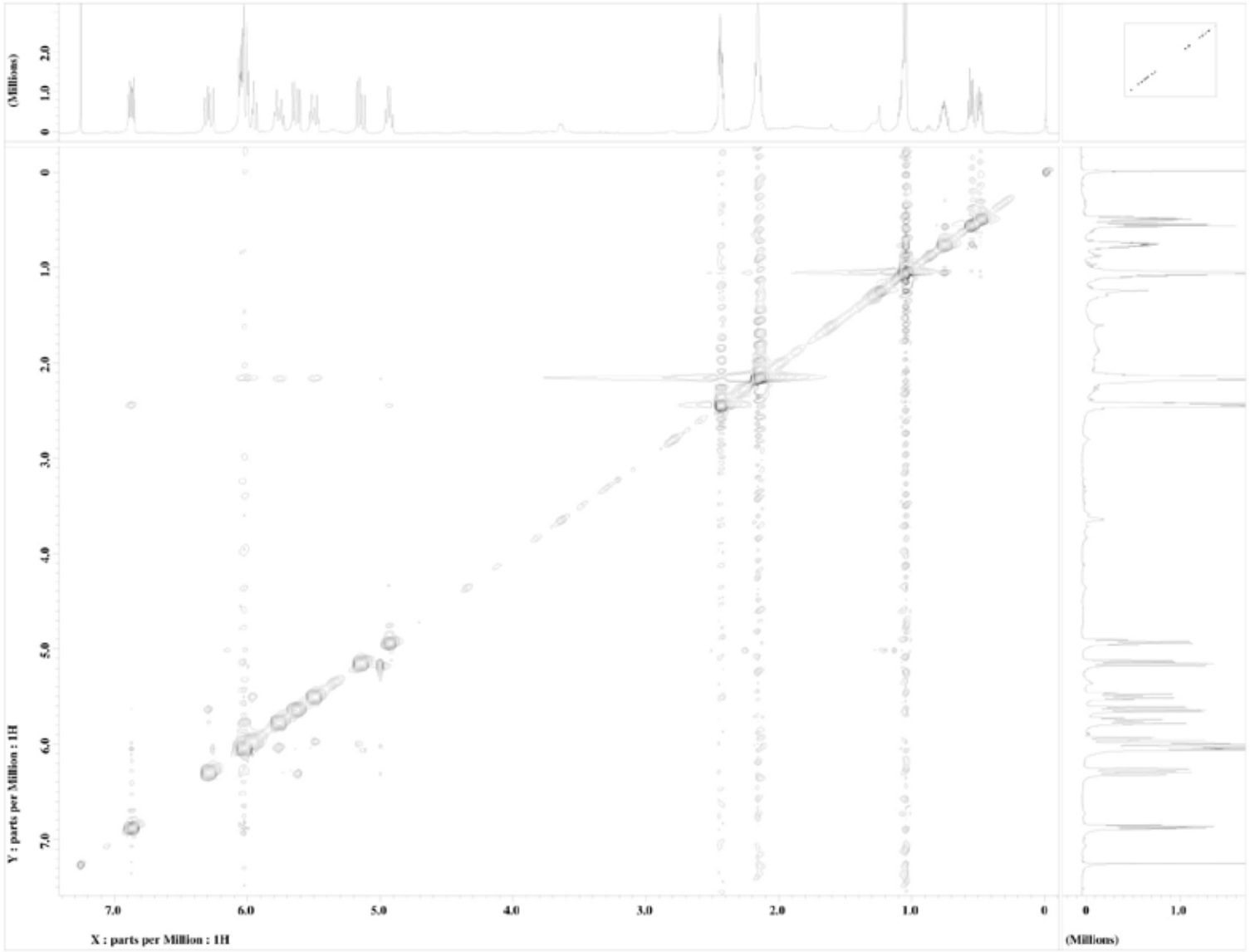


Figure S16. ^{1}H spectrum of Coibacin B (**2**) (in $CDCl_3$ at 400 MHz)

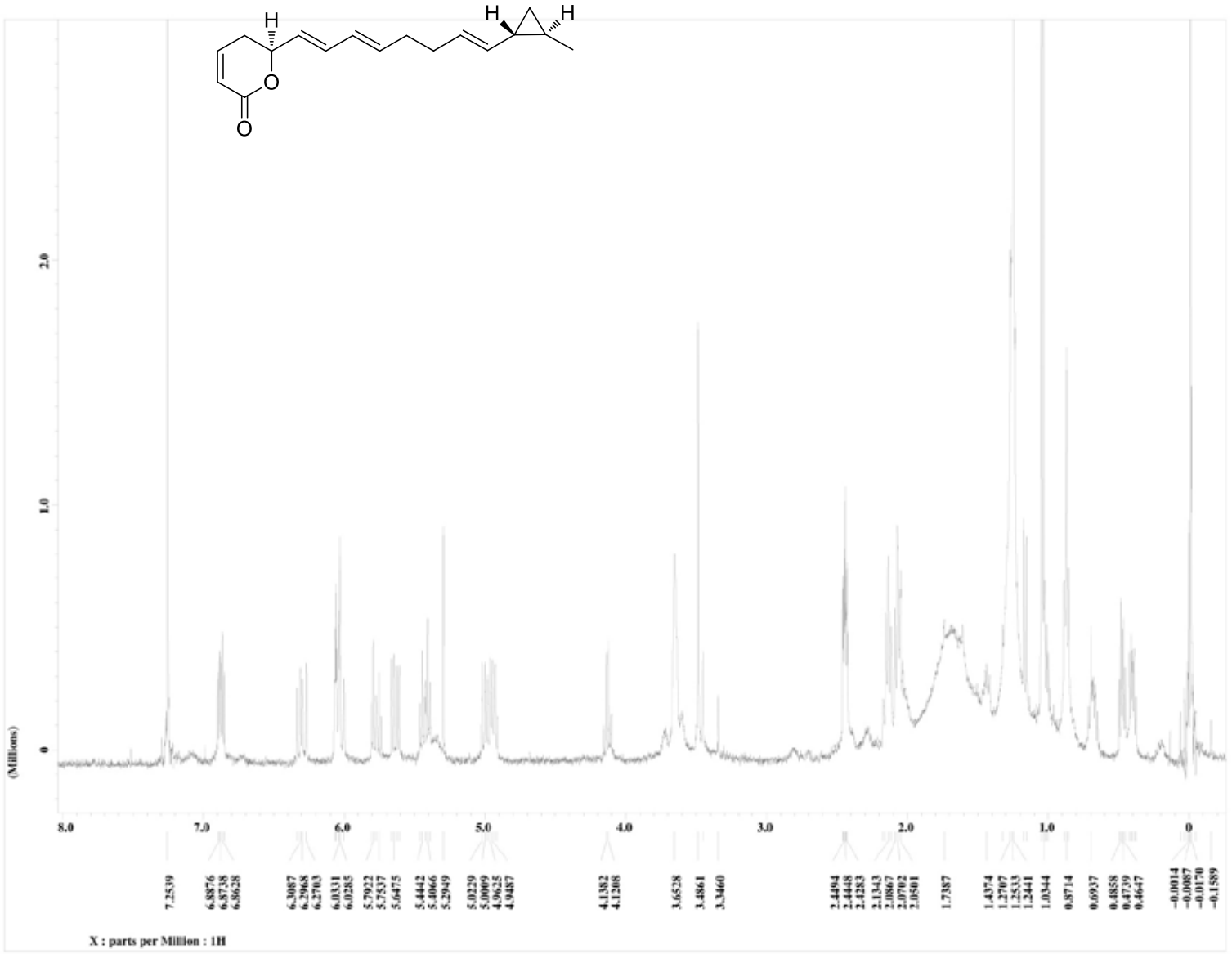


Figure S17. ^{13}C spectrum of Coibacin B (**2**) (in $CDCl_3$ at 100 MHz)

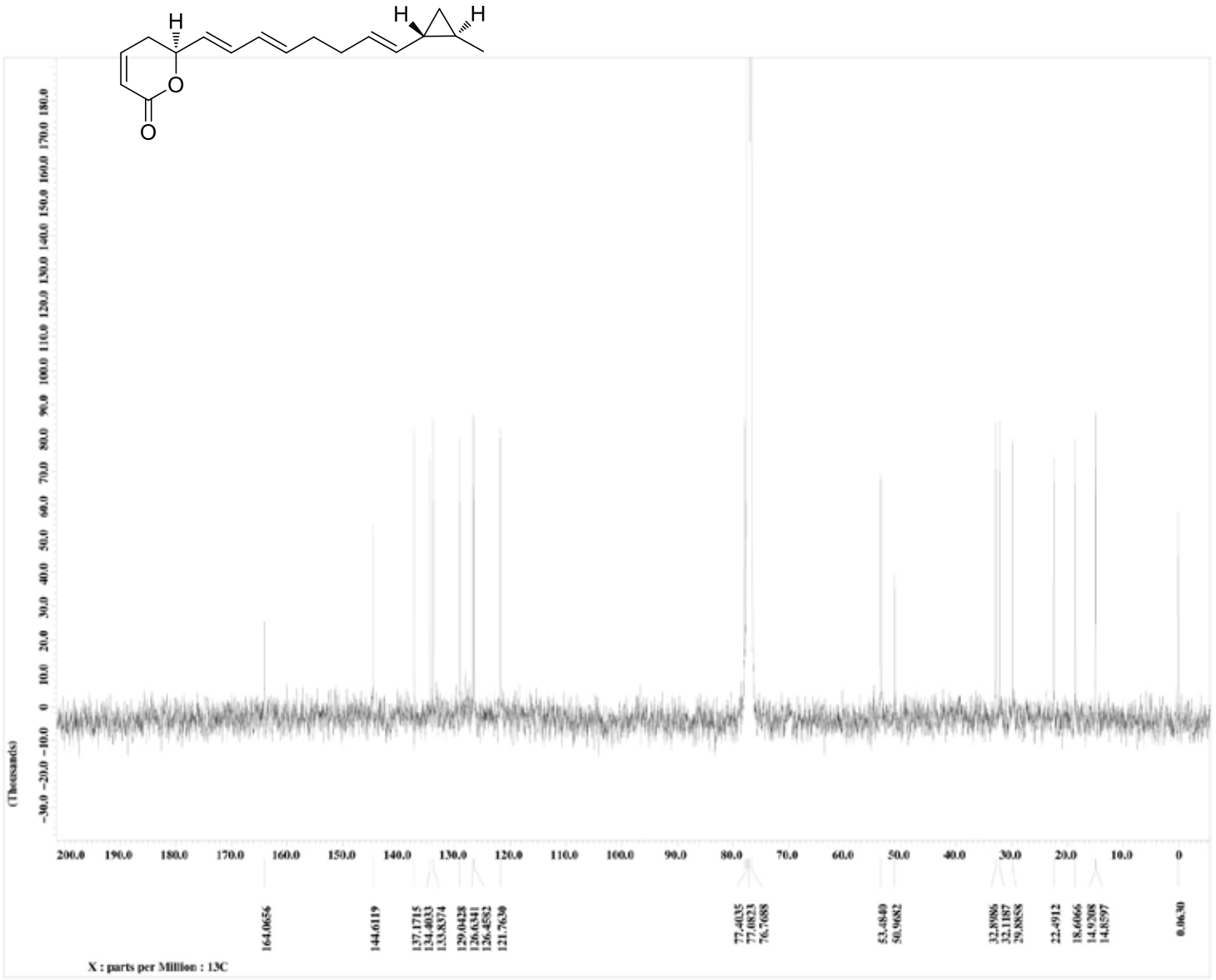


Figure S18. COSY spectrum of Coibacin B (**2**) (in $CDCl_3$ at 400 MHz)

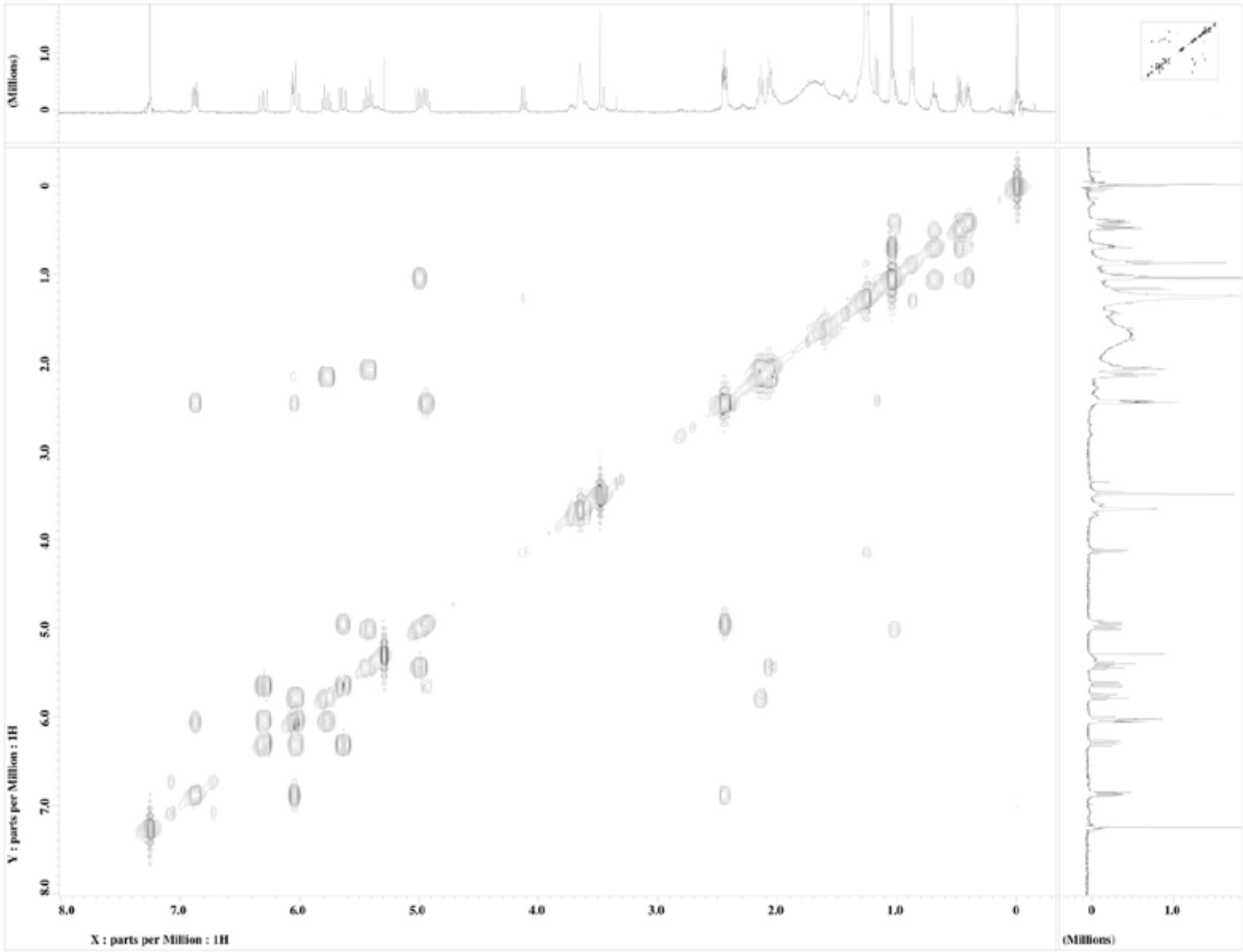


Figure S19. HSQC spectrum of Coibacin B (**2**) (in $CDCl_3$ at 400 MHz)

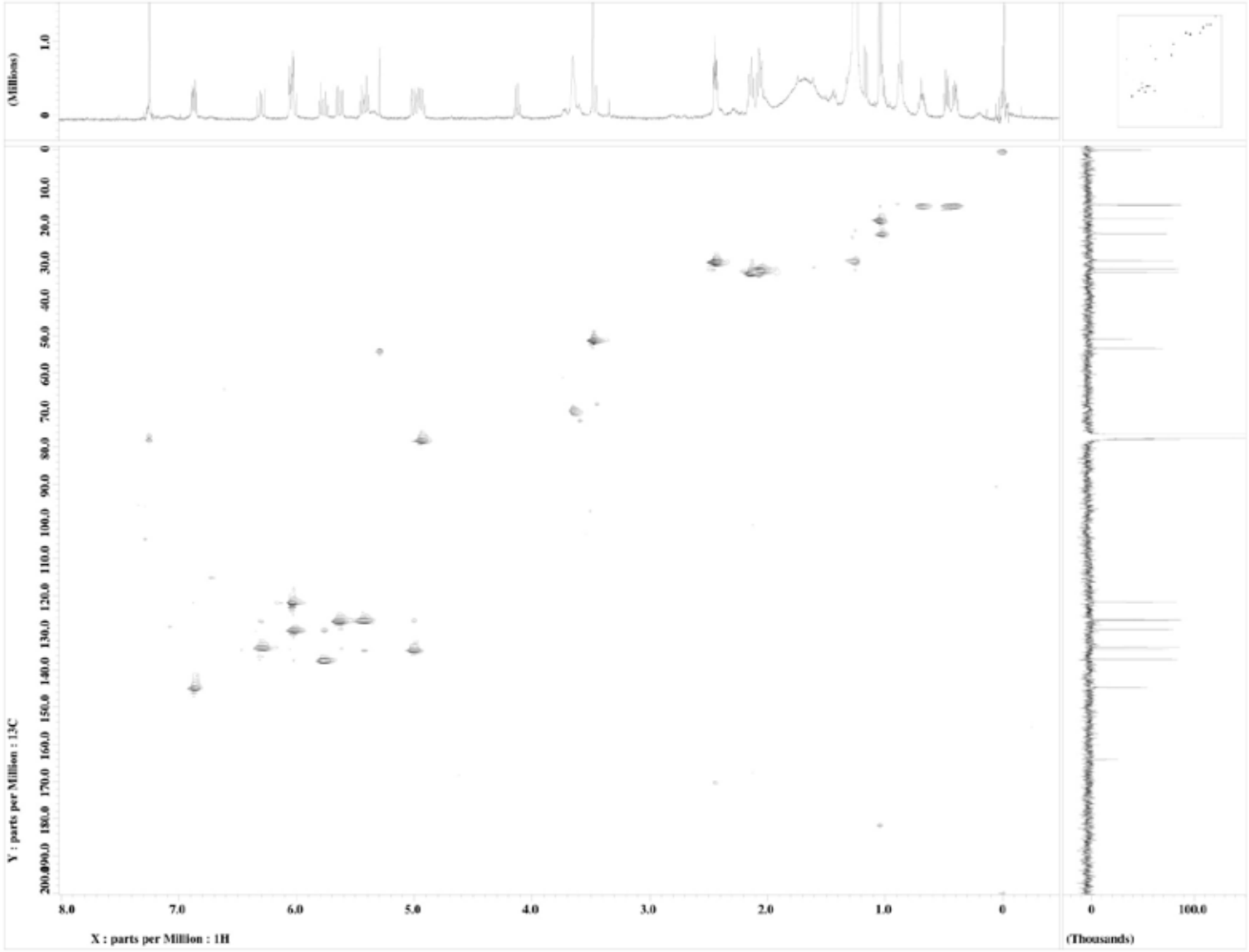


Figure S20. HMBC spectrum of Coibacin B (**2**) (in $CDCl_3$ at 400 MHz)

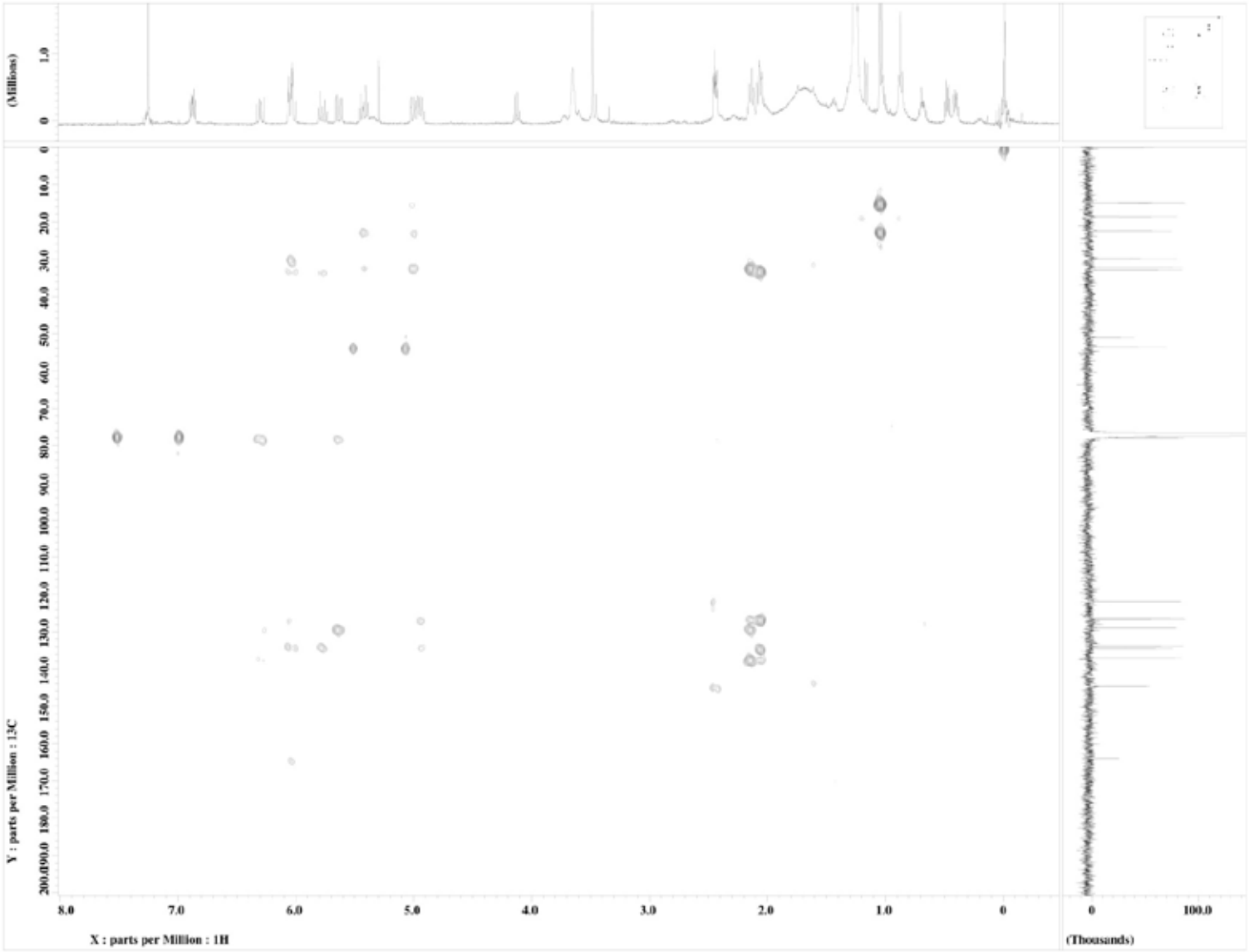


Figure S21. NOESY spectrum of Coibacin B (**2**) (in $CDCl_3$ at 400 MHz)

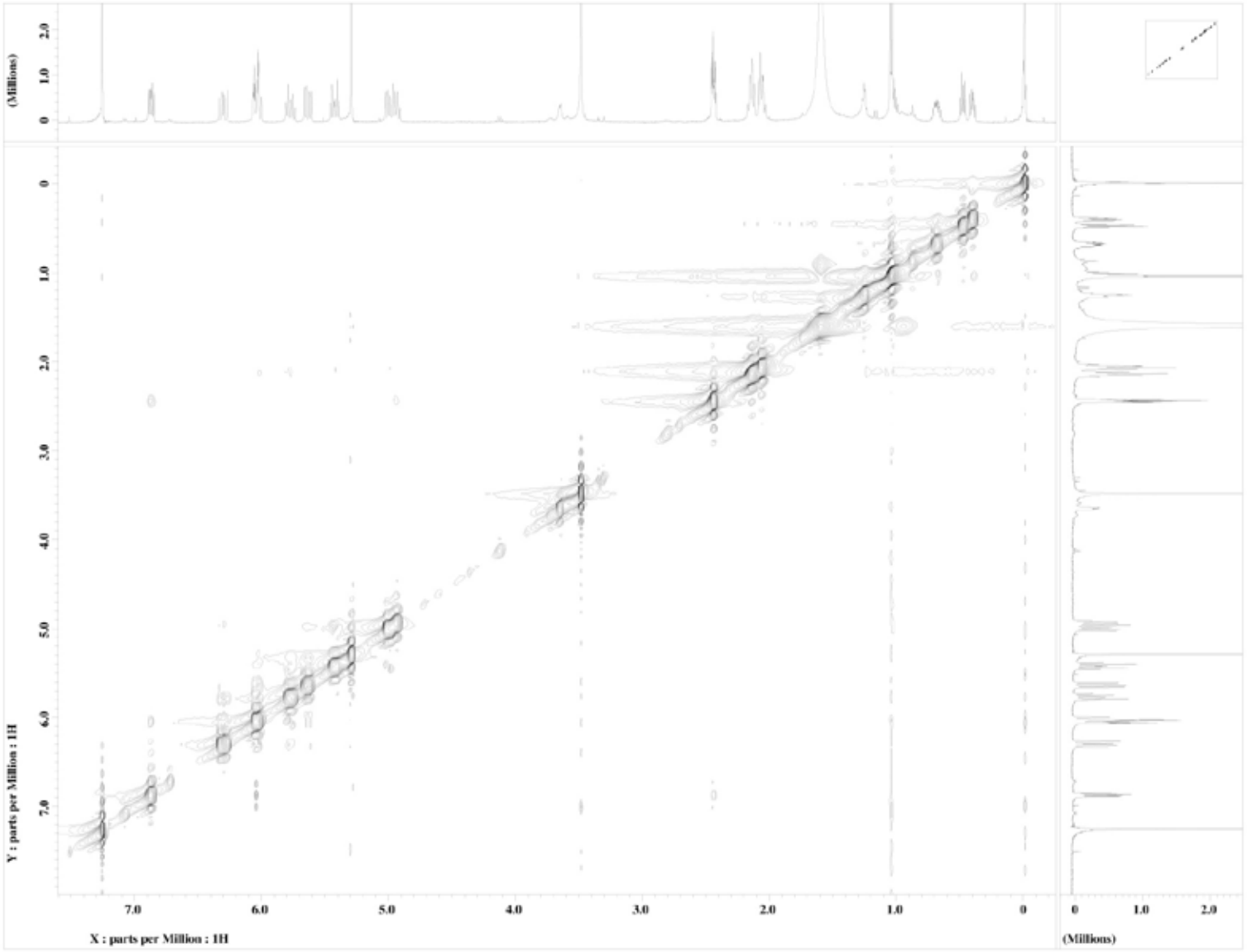


Figure S22. ^{1}H spectrum of Coibacin C (**3**) (in $CDCl_3$ at 400 MHz)

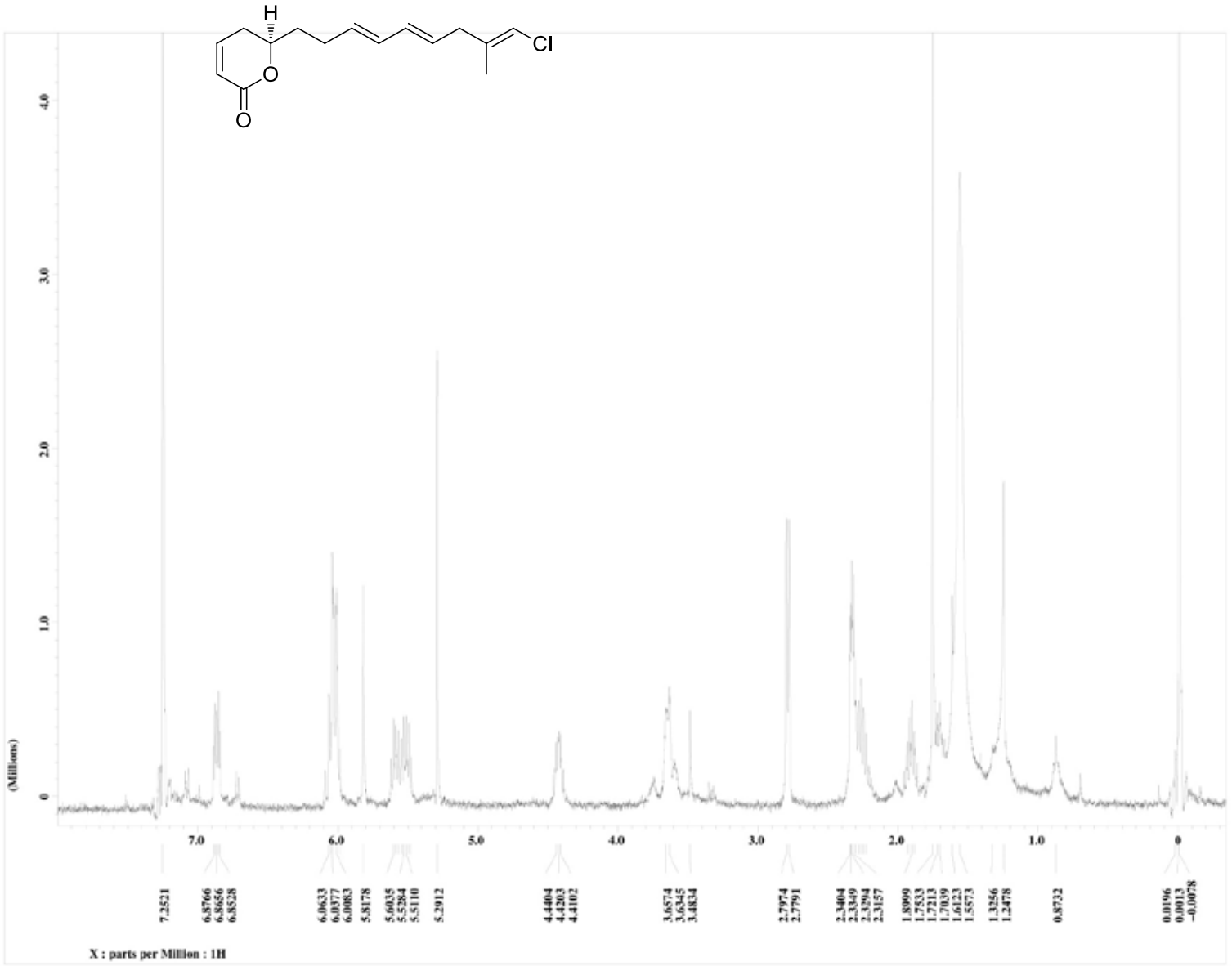


Figure S23. ^{13}C spectrum of Coibacin C (**3**) (in $CDCl_3$ at 100 MHz)

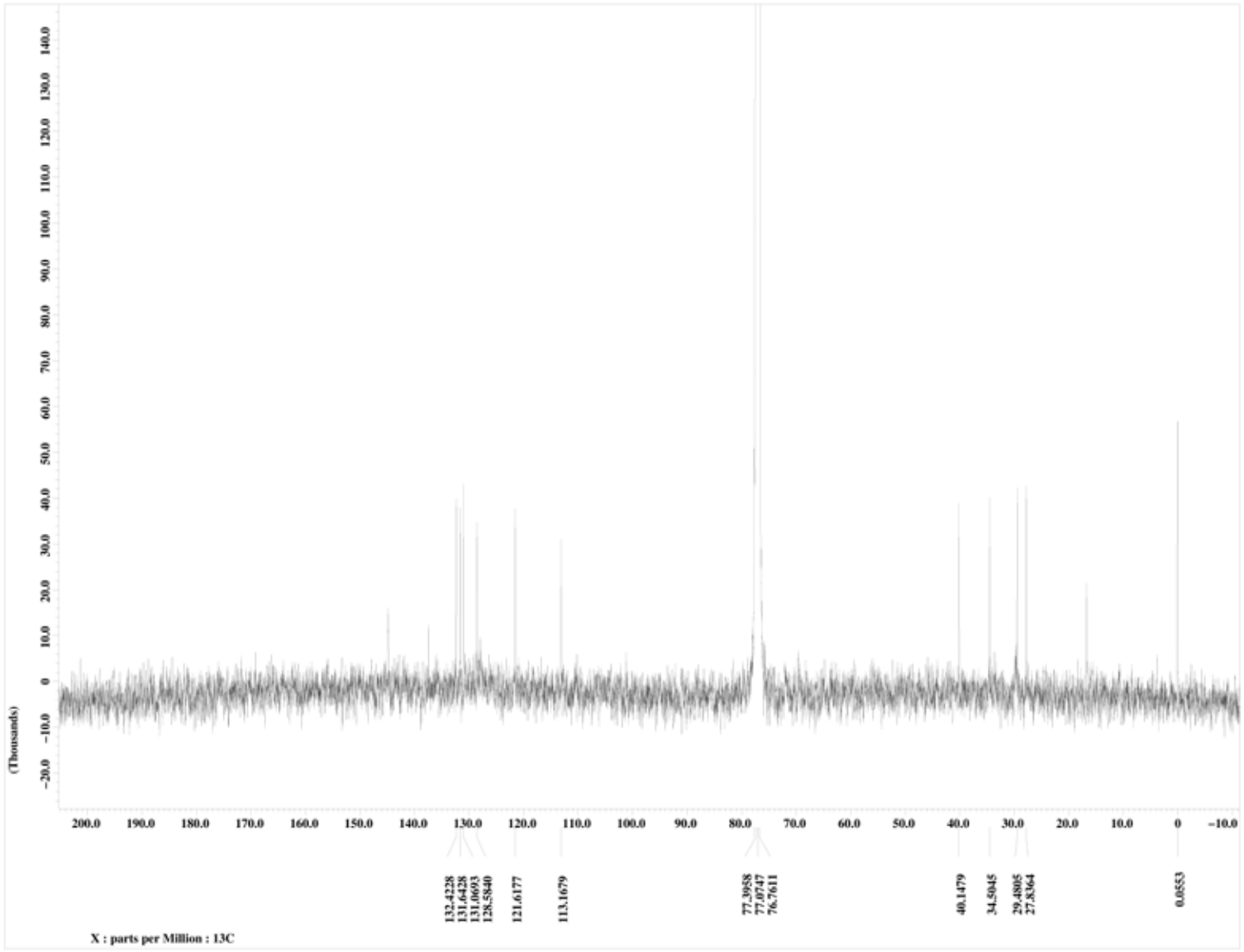


Figure S24. COSY spectrum of Coibacin C (**3**) (in $CDCl_3$ at 400 MHz)

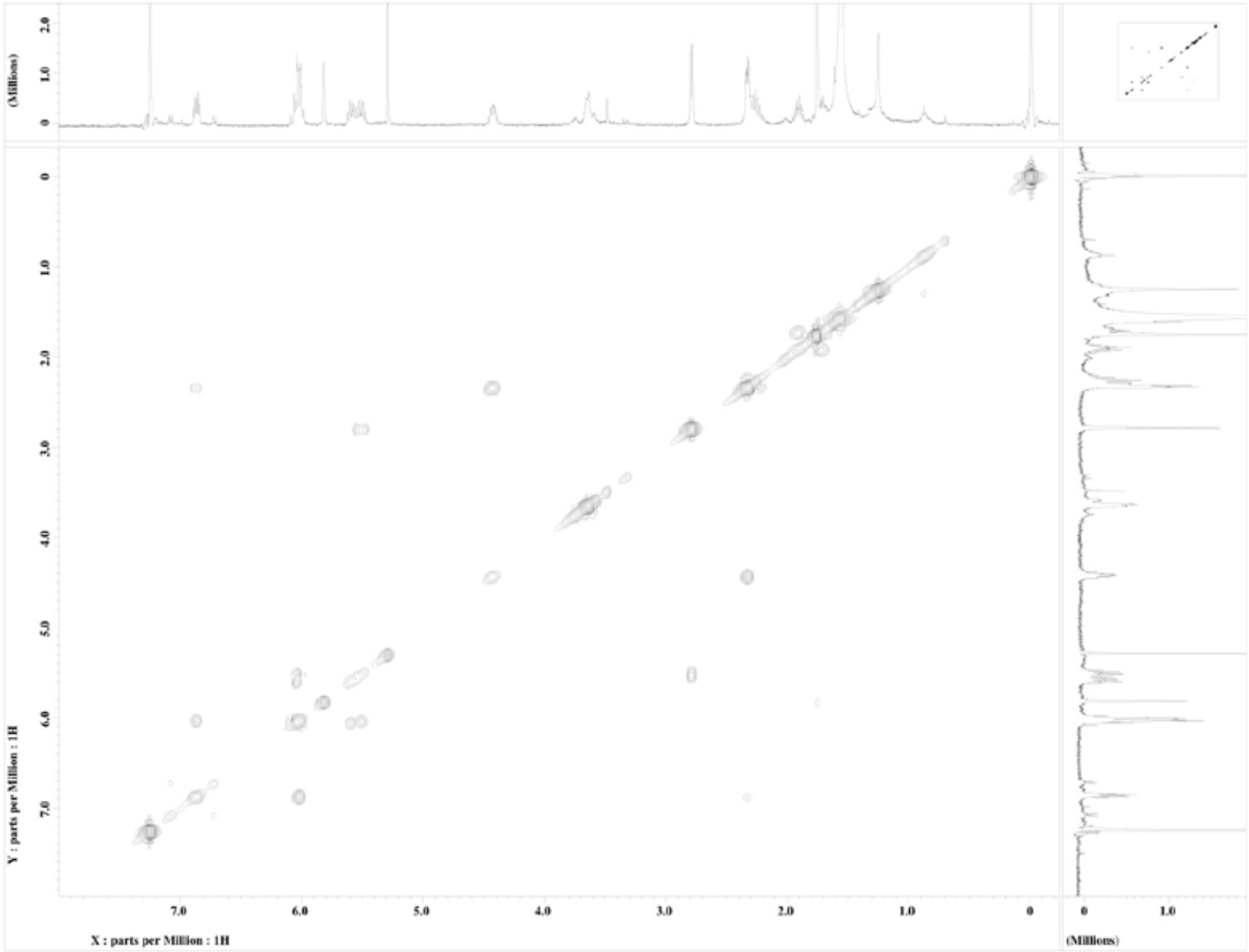


Figure S25. HSQC spectrum of Coibacin C (**3**) (in $CDCl_3$ at 400 MHz)

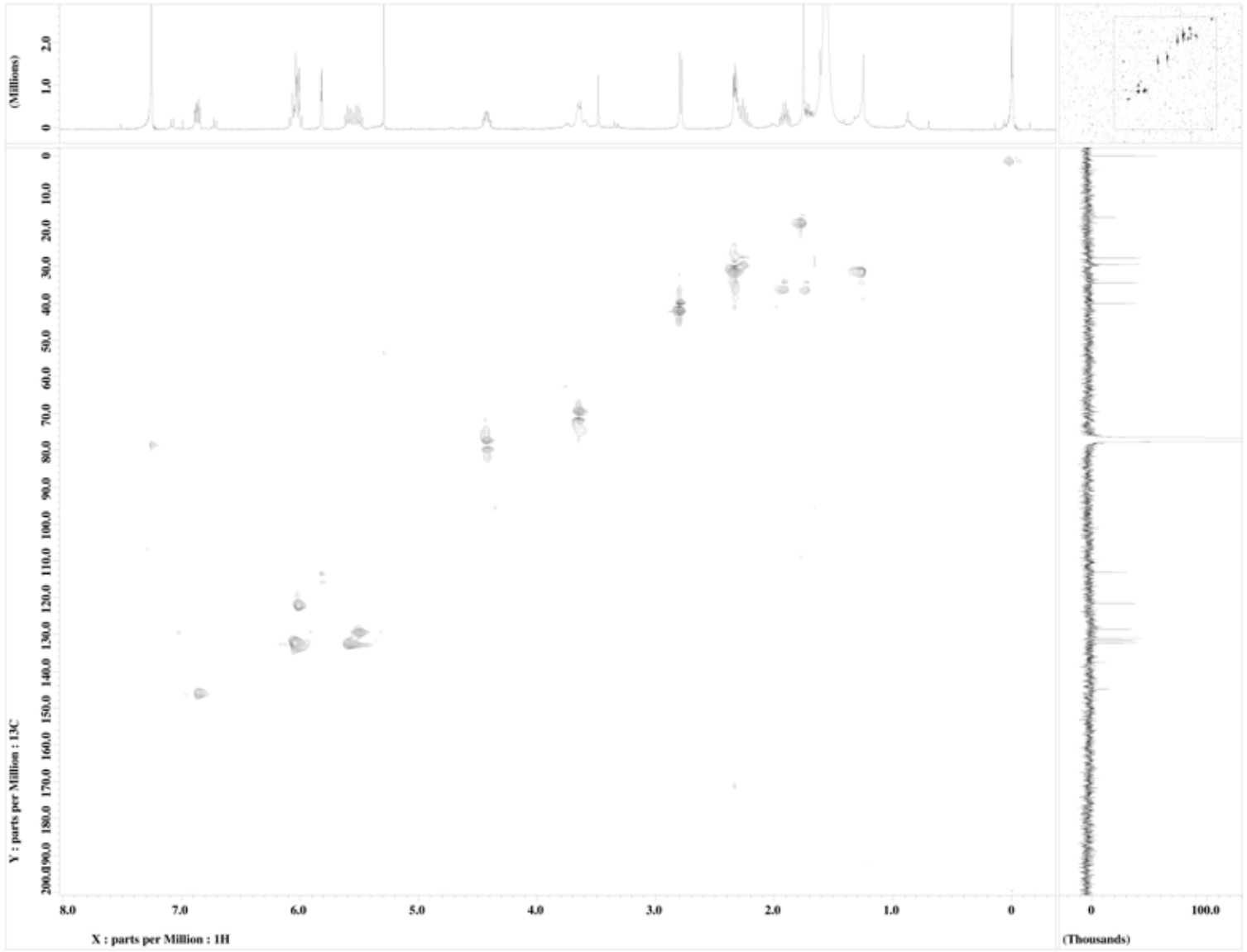


Figure S26. HMBC spectrum of Coibacin C (**3**) (in $CDCl_3$ at 400 MHz)

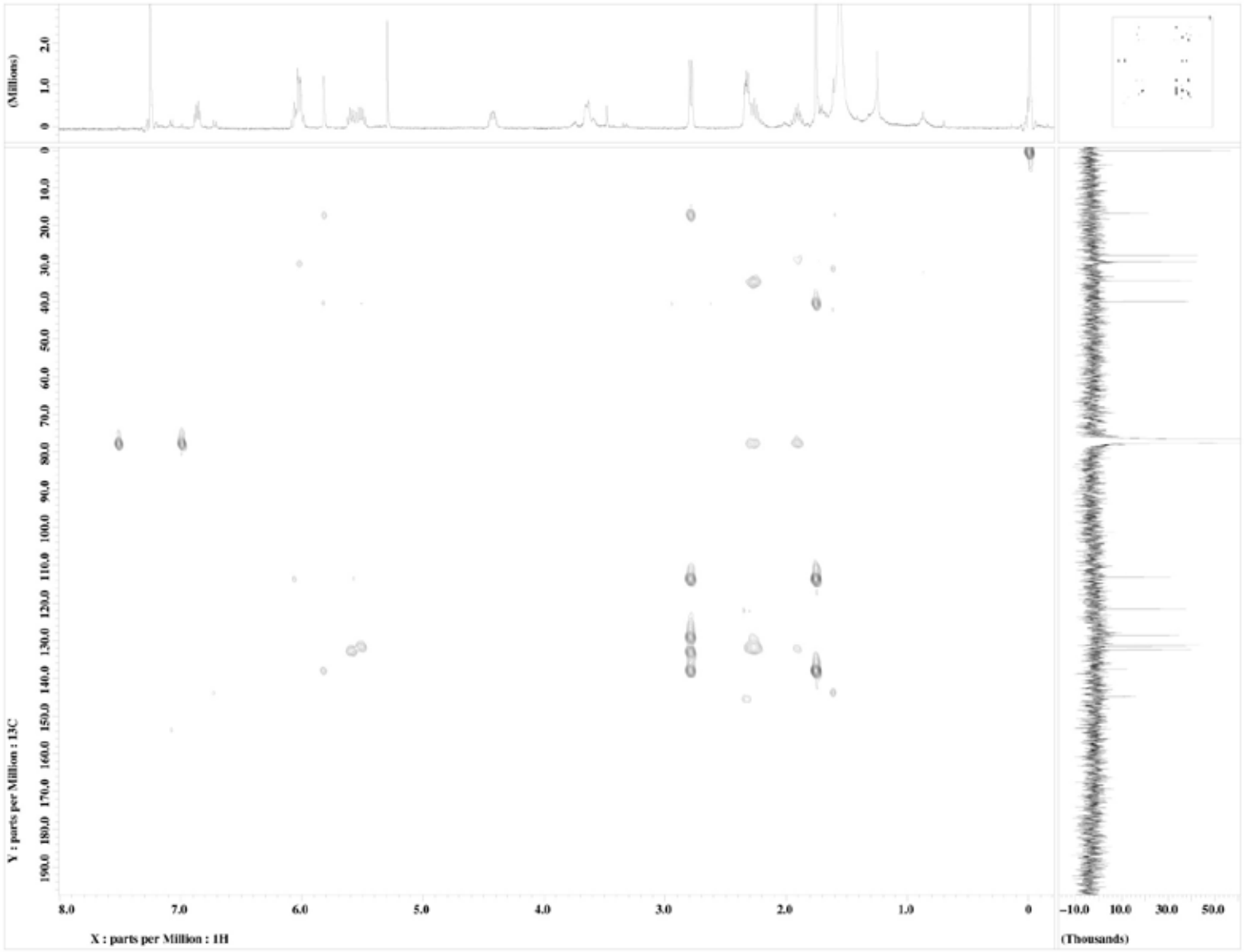


Figure S27. NOESY spectrum of Coibacin C (**3**) (in $CDCl_3$ at 400 MHz)

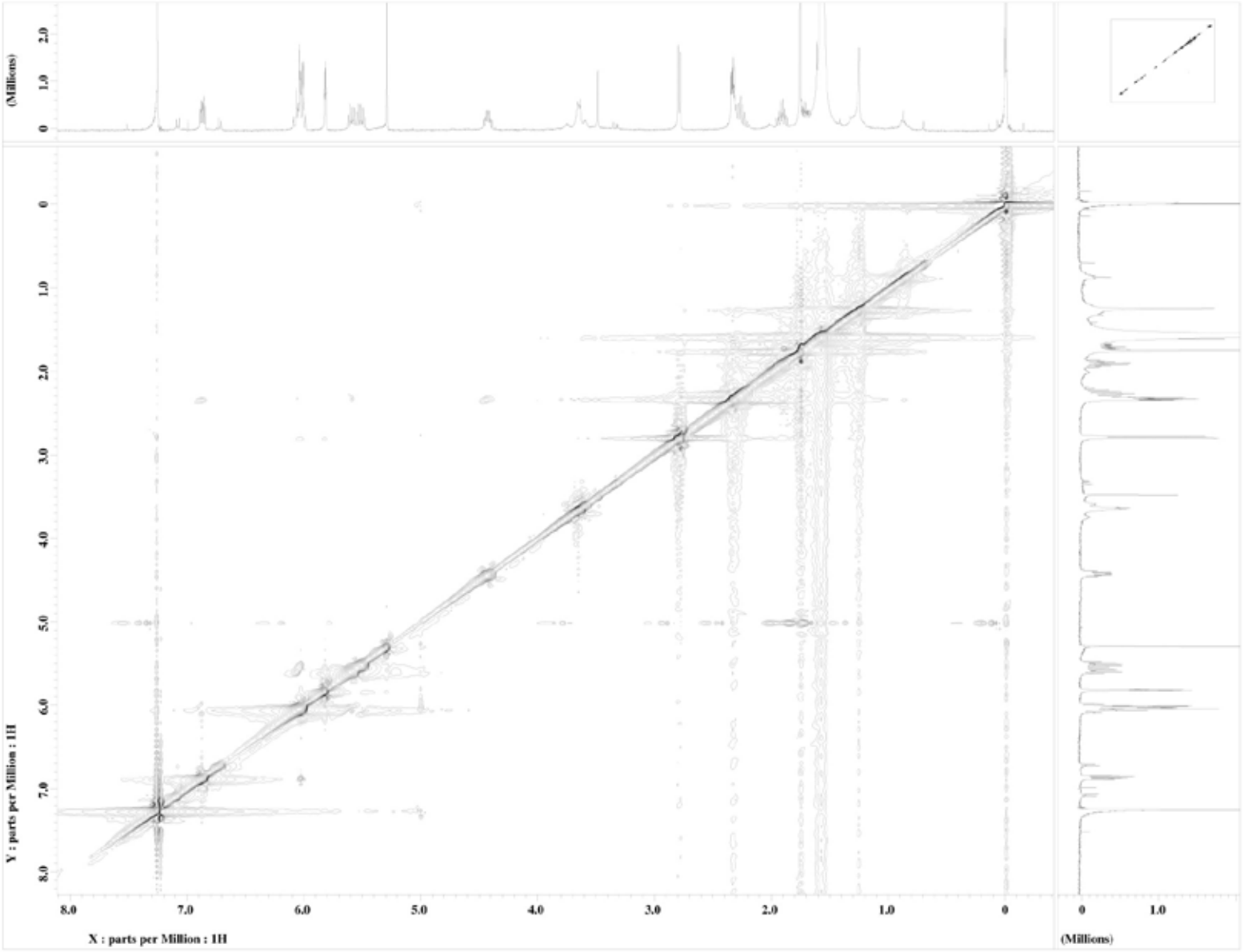


Figure S28. ^{1}H spectrum of Coibacin D (**4**) (in $CDCl_3$ at 400 MHz)

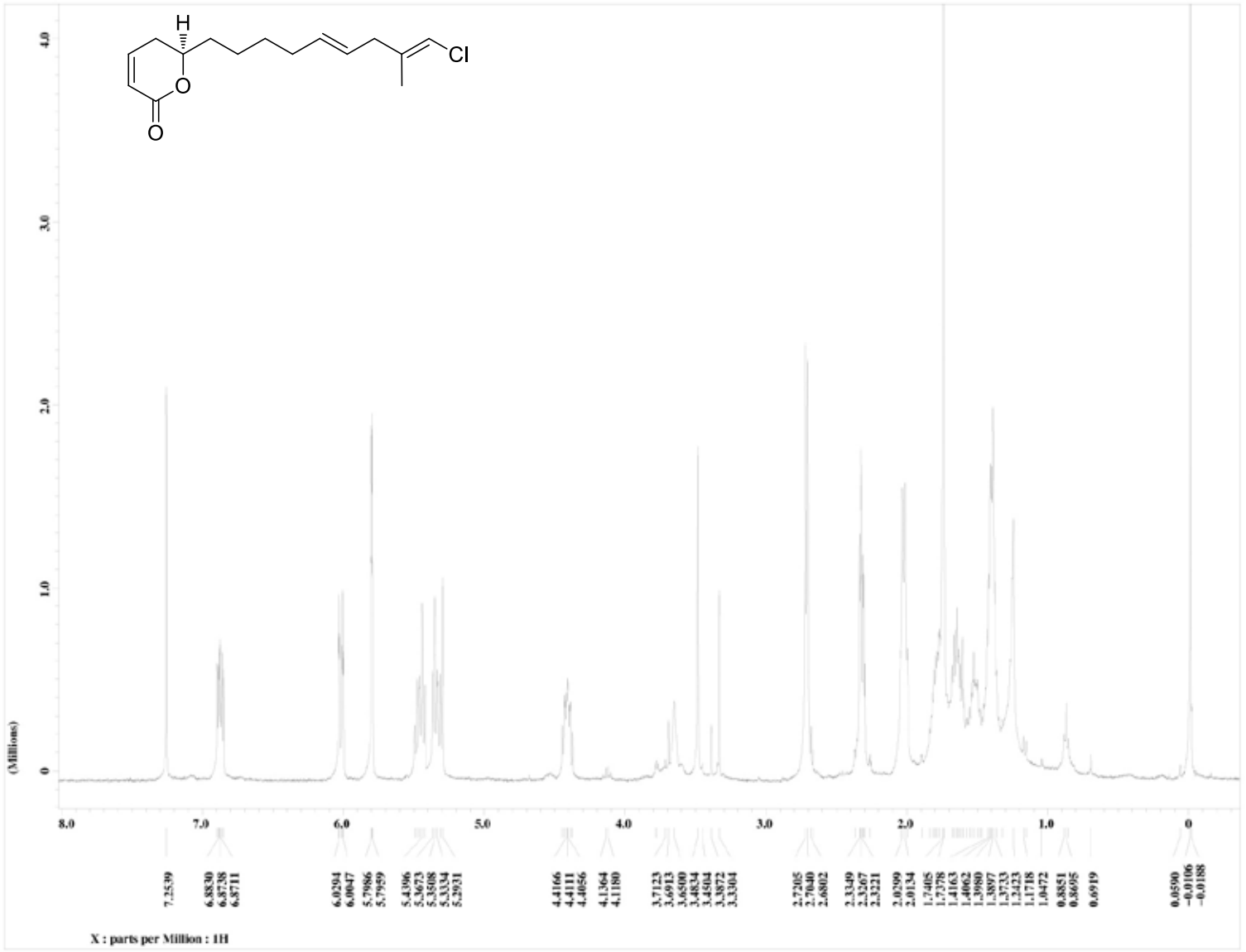


Figure S29. ^{13}C spectrum of Coibacin D (**4**) (in $CDCl_3$ at 100 MHz)

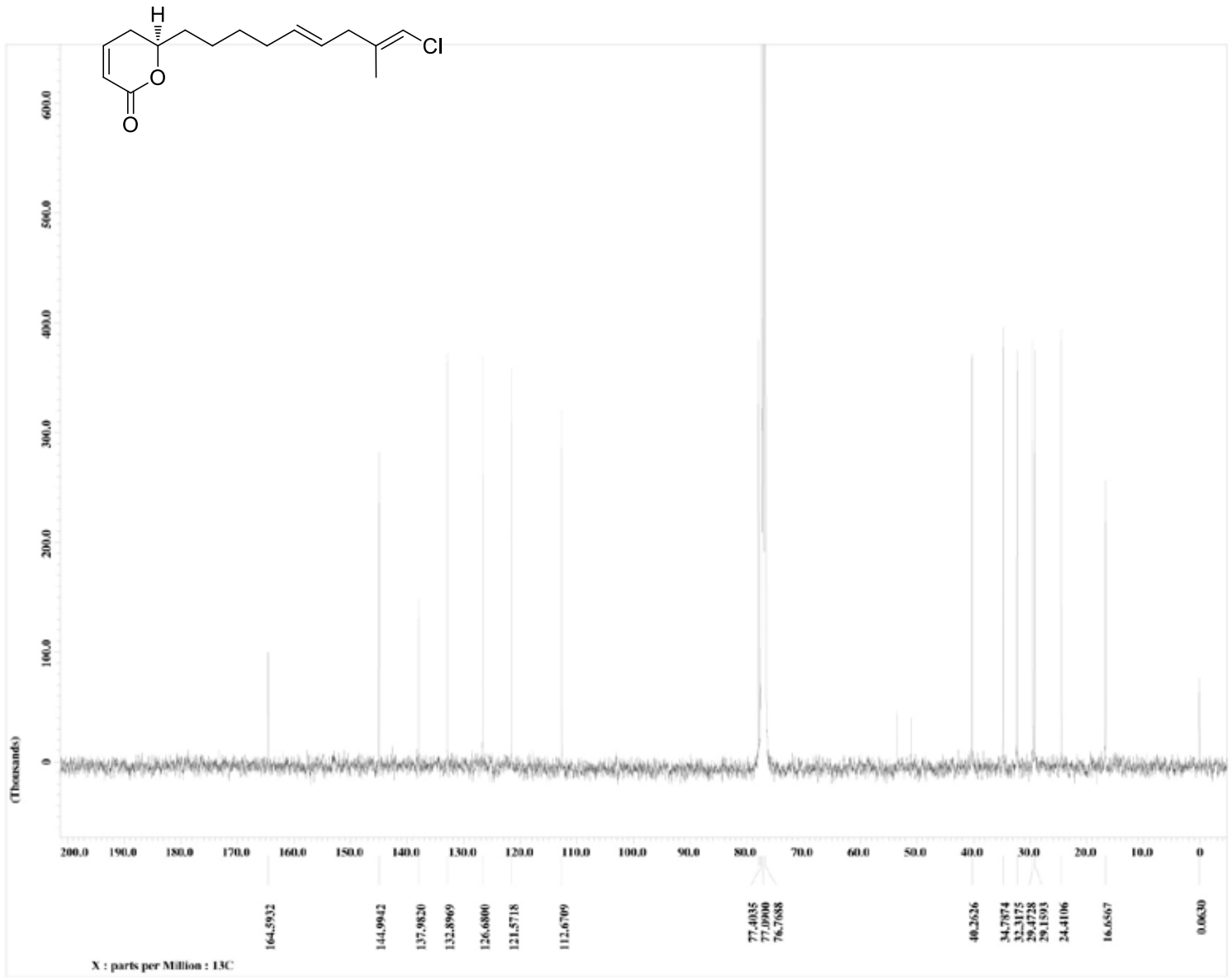


Figure S30. COSY spectrum of Coibacin D (**4**) (in $CDCl_3$ at 400 MHz)

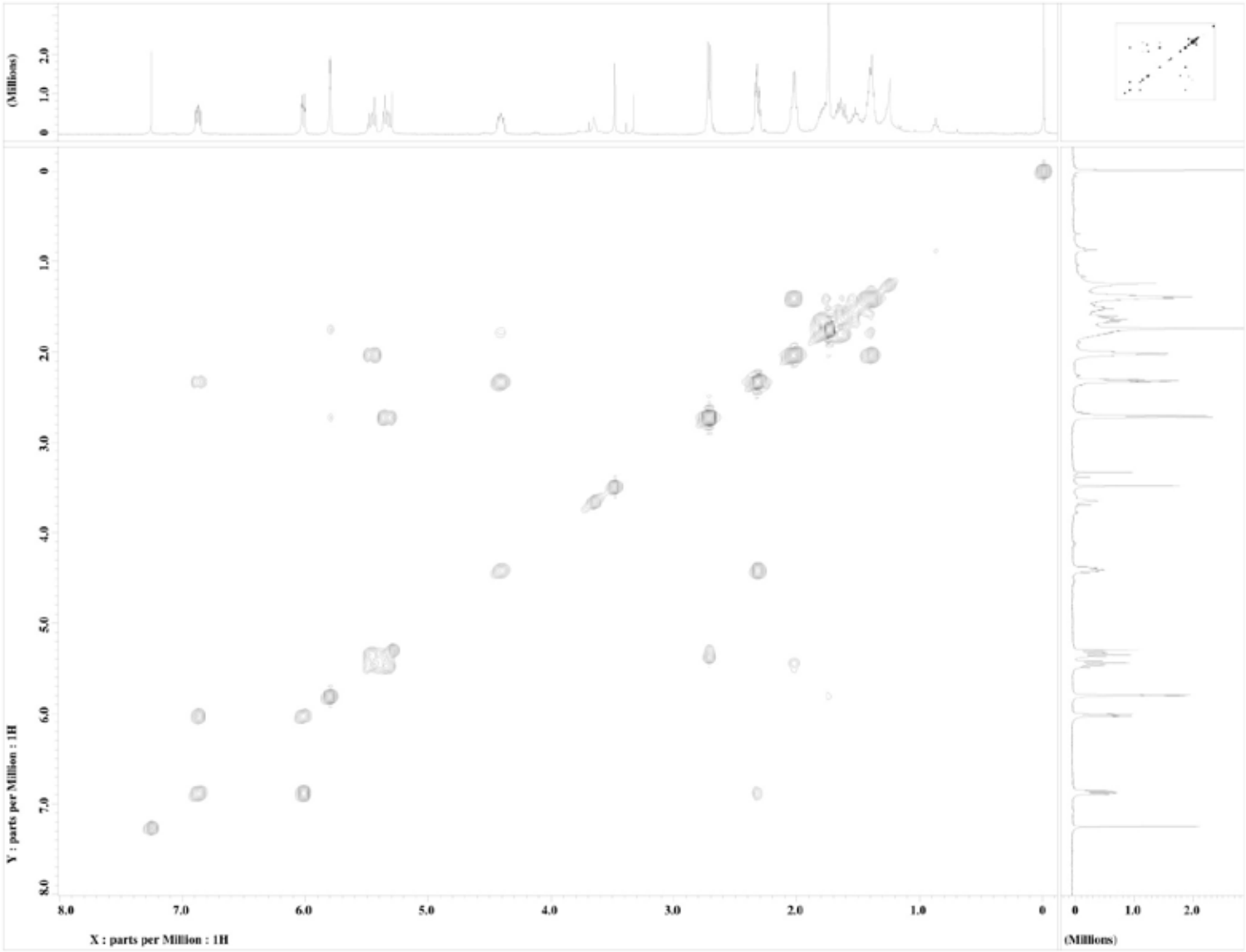


Figure S31. HSQC spectrum of Coibacin D (**4**) (in $CDCl_3$ at 400 MHz)

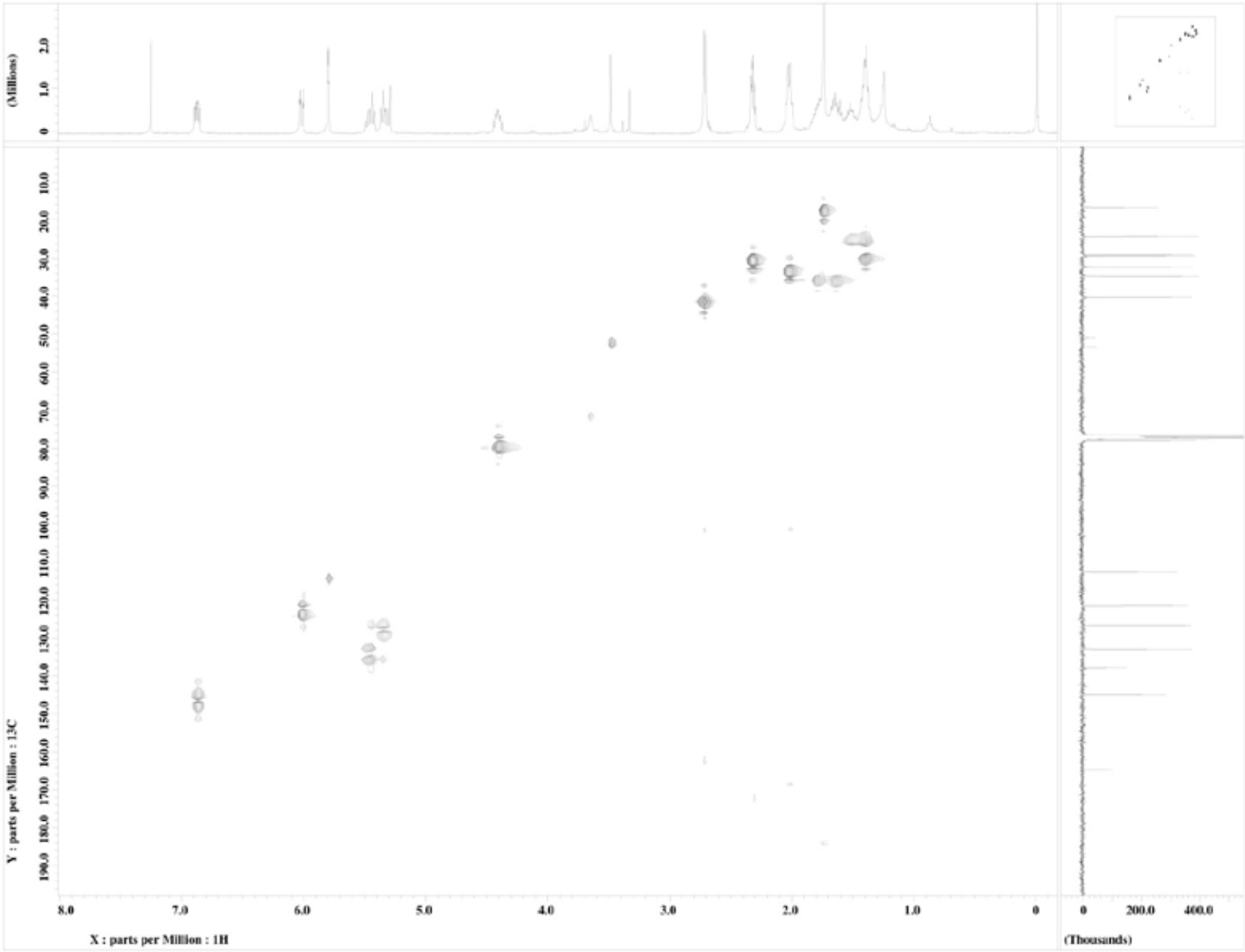


Figure S32. HMBC spectrum of Coibacin D (**4**) (in $CDCl_3$ at 400 MHz)

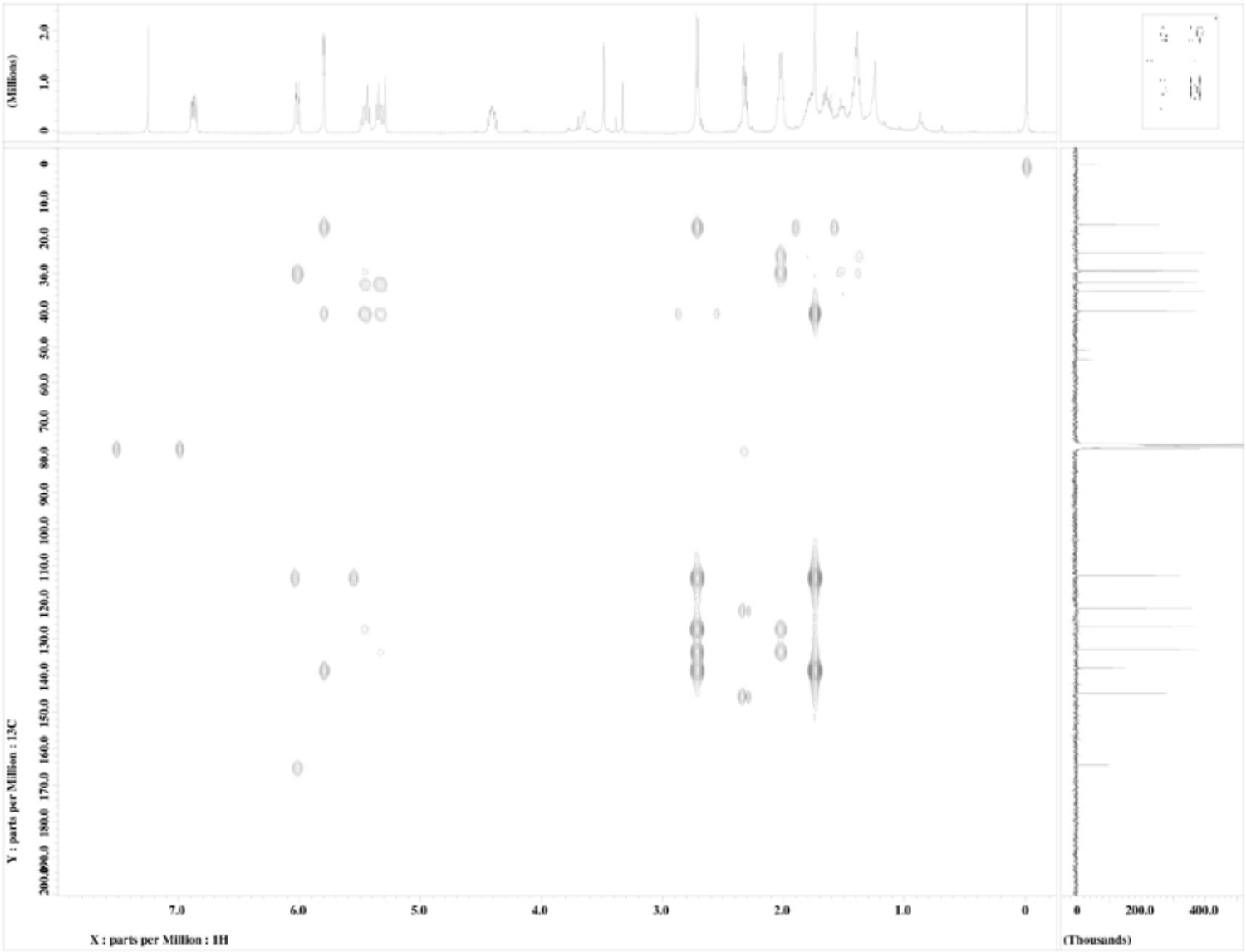
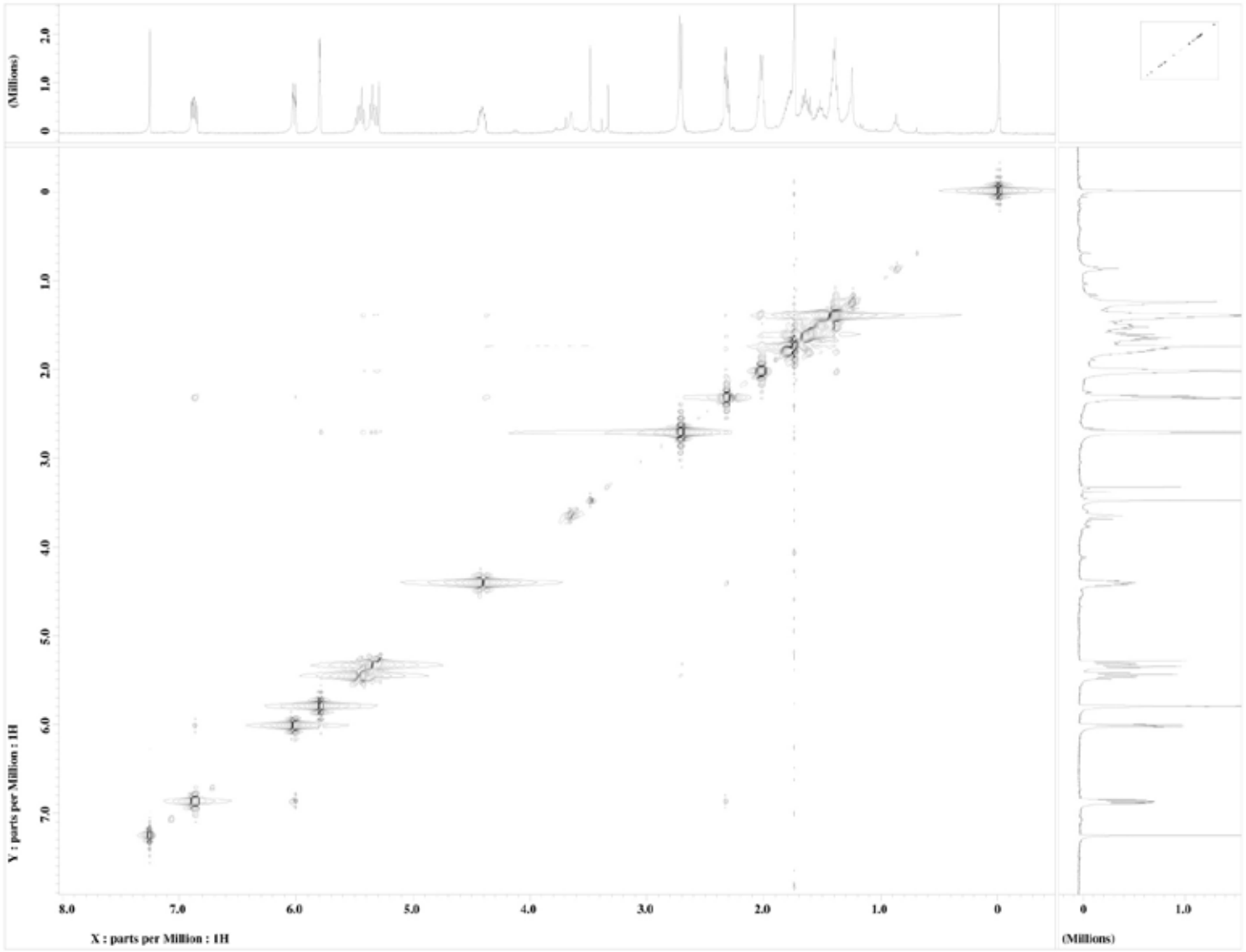


Figure S33. NOESY spectrum of Coibacin D (**4**) (in $CDCl_3$ at 400 MHz)



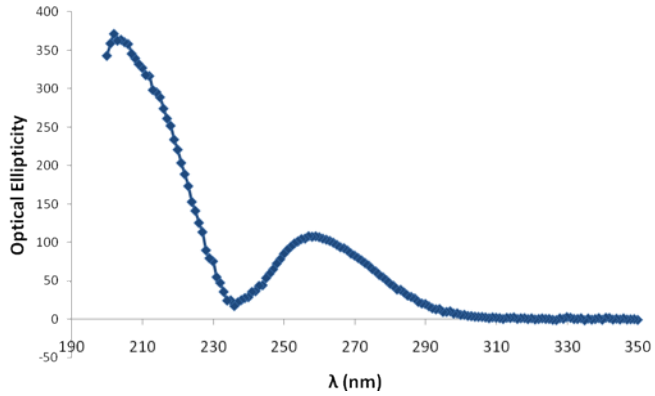


Figure S34. Circular dichroism of Coibacin A (**1**).

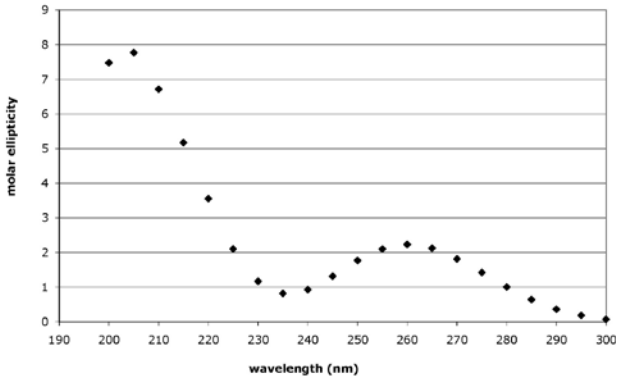


Figure S35. Circular dichroism of Coibacin D (**4**).

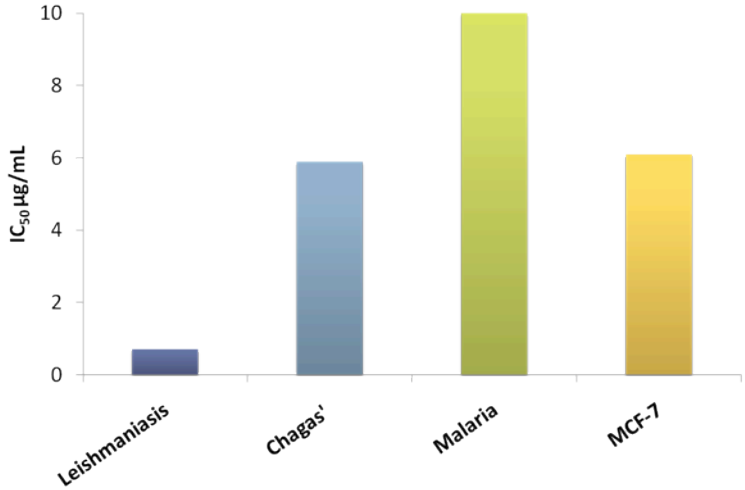


Figure S36. Activity of Coibacins A (**1**) against *Leishmania donavoni* (leishmaniasis), *Trypanosoma cruzi* (Chagas' disease), *Plasmodium falciparum* (malaria), and MCF-7 human breast cancer cells.

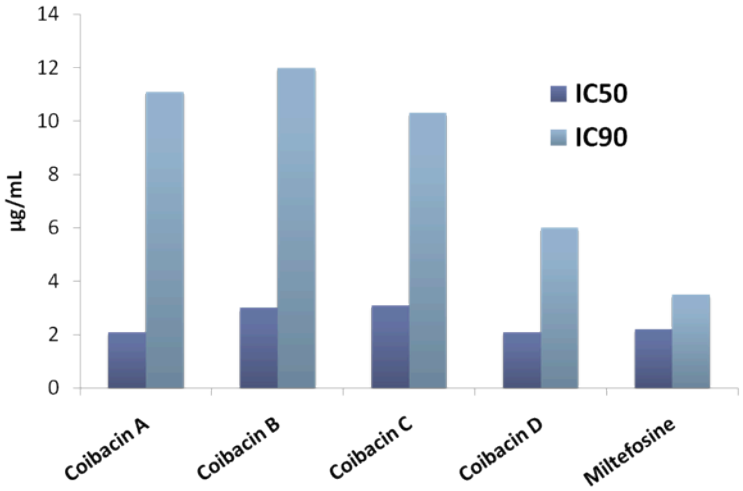


Figure S37. Activity of Coibacins A-D (**1-4**) against *Leishmania mexicana* axenic amastigotes.

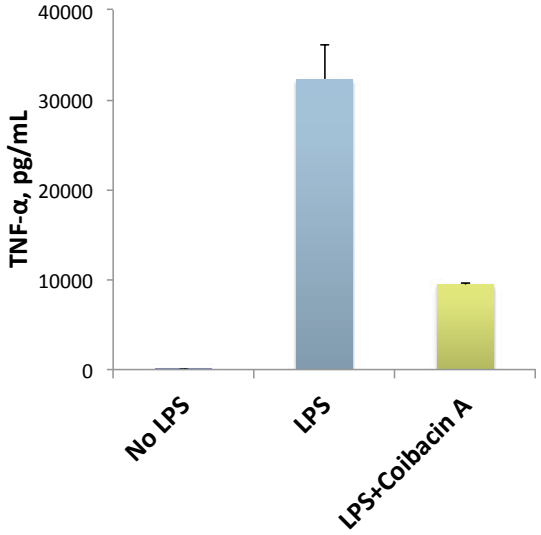


Figure S38. Changes in TNF-α protein expression induced by Coibacin A (**1**).

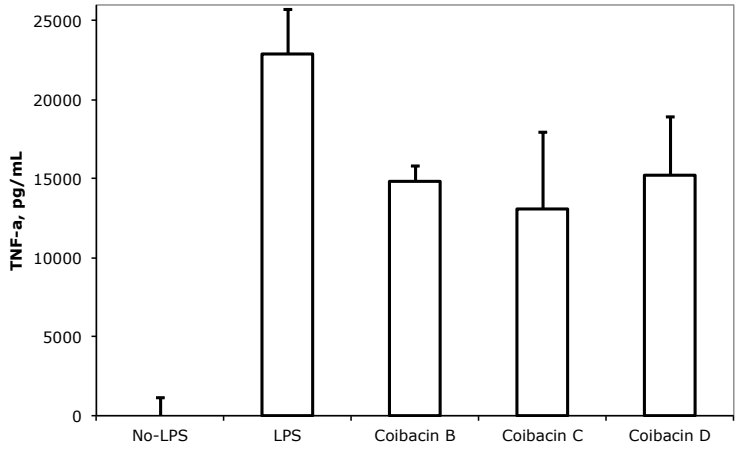


Figure S39. Changes in TNF-α protein expression induced by Coibacins B-D (**2-4**).

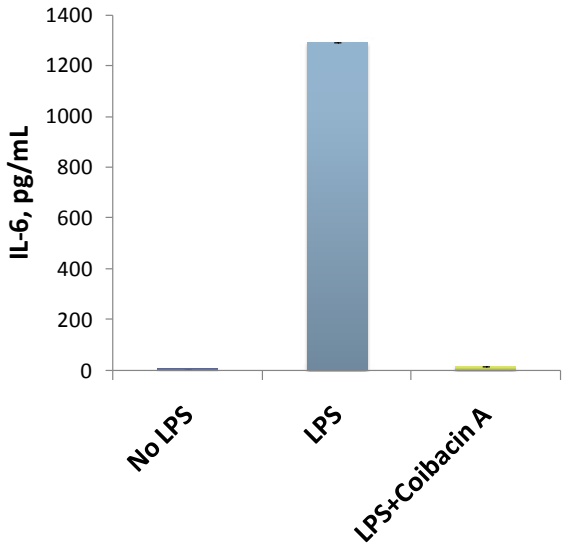


Figure S40. Changes in IL-6 protein expression induced by Coibacin A (**1**).

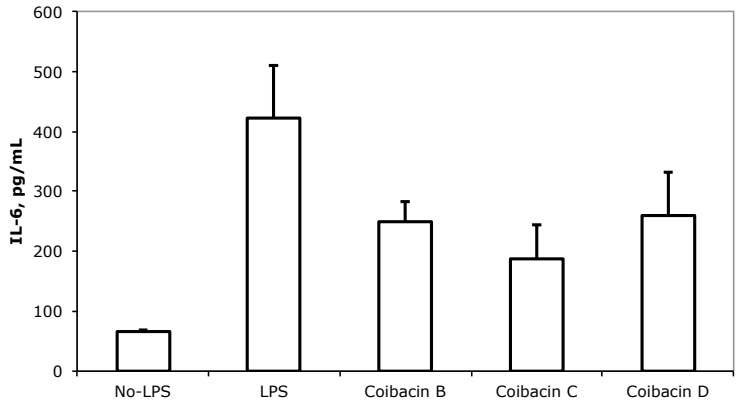


Figure S41. Changes in IL-6 protein expression induced by Coibacins B-D (**2**-**4**).

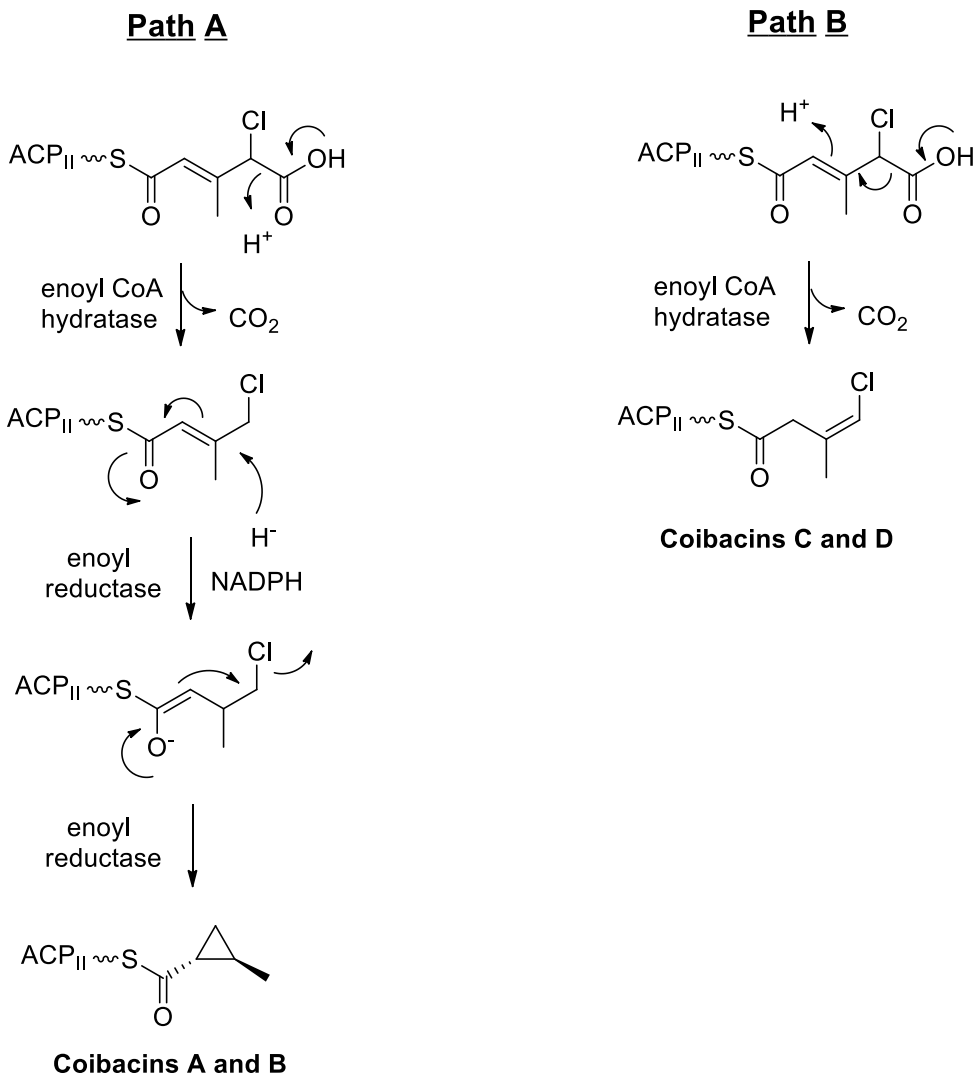


Figure S42. Variable regiochemistry of proton addition resulting from enoyl CoA hydratase catalyzed decarboxylation may explain the co-occurrence of cyclopropyl-containing metabolites (proton addition to the γ-carbon to yield coibacins A and B) and vinyl chloride-containing metabolites (proton addition to the α-carbon to yield coibacins C and D). In Path A, subsequent enoyl reductase catalyzed hydride delivery occurs to the β-carbon, and following electron shifts, results in loss of chloride to produce the cyclopropyl ring [ACPII = acyl carrier protein II].

S43: References

Cannone, J. J.; Subramanin, S.; Schnare, M. N.; Collett, J. R.; D'Souza, L. M.; Du, Y.; Feng, B.; Lin, N.; Madabusi, L. V.; Muller, K. M.; Pnde, N.; Schang, Z.; Yu, N.; Gutell R. R. *BMC Bioinformatics* **2002**, *3*, 1471–2105.

Castenholz, R. W.; Rippka, R.; Herdman, M. In *Bergey's Manual of Systematic Bacteriology*; Boone, D. R. and Castenholz, R. W., Eds.; Springer: New York, 2001; Vol. 1, pp 473-599.

Green, L. C.; Wagner, D. A.; Glogowski, J.; Skipper, P. L.; Wishnok, J. S.; Tannenbaum, S. R.; *Anal. Biochem.* **1982**, *126*, 131-138.

Katoh, K.; Toh, H. *Brief Bioinform.* **2008**, *9*, 286-298.

Komárek, J.; Anagnostidis, K. In *Süßwasserflora von Mitteleuropa*; Büdel, B.; Gärtner, G.; Krienitz, L.; Schagerl, M., Eds. Gustav Fischer: Jena, 2005, 19/2.

Nylander, J. A. A.; Wilgenbusch, J. C.; Warren, D. L.; Swofford, D. L. *Bioinformatics (Oxf.)* **2008**, *15*, 581-583.

Posada, D. *Mol. Biol. Evol.* **2008**, *25*, 1253-1256.

Ronquist, F.; Huelsenbeck, J. P. *Bioinformatics (Oxf)*. **2003**, *12*, 1572-1574.

Zwickl, D. J. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Ph.D. Thesis, The University of Texas at Austin, 2006.