# Table of contents

**Experimental procedures:**

**Strains and media**

Fungus *Calcarisporium arbuscula* NRRL 3705 (ATCC® 46034™) is the wild type strain. The *ΔhdaA* and *ΔarbA* mutants were constructed by replacement of *hdaA* and *arbA* by *hph* (Hygromycin-resistance gene), respectively. Spores of *C. arbuscula* were collected after 6 days on potato dextrose agar (PDA, BD). PDA was routinely used for production of compounds from *C. arbuscula*. Other media MEPA (3 % malt extract, 0.3 % soy flour, 2 % agar), Czapek yeast agar (CYA) (3 % sucrose, 0.5 % yeast extract, 0.1 % K2HPO4, 0.01 % (V/V) Czapek concentrate, 2 % agar), YMEG (0.4 % glucose, 0.4 % yeast extract, 1 % malt extract, 0.2 % $CaCO_3$, 2 % agar), YG (0.5 % yeast extract, 2 % glucose, 0.04 % (V/V) trace element) were also tested for compound productivity.

**Gene knock-out in *Calcarisporium arbuscula***:

Primers for gene knock-out were listed in Table S6.

*arbA* and *hdaA* genes were deleted in *C. arbuscula* according to the hygromycin split-marker strategy.[1] Hygromycin-resistance gene *hph* upstream and downstream fragments were amplified from plasmid pAN-7 (Addgene) with primers hph-up F, hph-up R and hph-dn F, hph-dn R, respectively, and digested with *Not*I/*Sac*II and *Sac*I/*Not*I to insert into the self-ligated T-vector pTA2 (Toyobo) to create plasmid pXM01 and pXM02, respectively. The primers for upstream and downstream homologous fragments are: arbA-up F, arbA-up R and arbA-dn F, arbA-dn R for *arbA* gene, hdaA-up F, hdaA-up R and hdaA-dn F, hdaA-dn R for *hdaA* gene. The amplified upstream and downstream fragments were ligated into pXM01 or pXM02, respectively, after enzyme digestion: *Kpn*I/*Not*I for arbA-up, *Bam*HI/*Not*I for hdaA-up, *Not*I/*Kpn*I for arbA-down and *Not*I/*Xba*I for hdaA-down.

Split-marker DNA was introduced into *C. arbuscula* by protoplast transformation. All the deletion cassettes were amplified with universal primers M13-F and M13-R from the above plasmids. All PCR products were precipitated with ethanol and dissolved in STC buffer (1.2 M sorbitol, 10 mM $CaC1_2$, 10 mM Tris-HCl, pH 7.5). *C. arbuscula* spores were collected on PDA (potato dextrose agar, Fluka) for 6 days at 25°C, and induced to young germ tubes in PDB (potato dextrose broth, Fluka) at 25°C for 7 hours with 180 rpm agitation. Cells were collected, washed twice with osmotic medium (1.2 M $MgCl_2$, 10 mM sodium phosphate, pH 5.8 ) and resuspended in the enzyme cocktail solution (3 mg/ml Lysing Enzymes, 3 mg/ml Yatalase in osmotic medium) at 30°C for 4 hours. After twice wash with STC buffer, 50 μl of protoplasts were gently mixed with 50 μl of DNA and incubated for 1 hour on ice. 300 μl of PEG 4000 solution (60% PEG 4000, 50 mM $CaC1_2$, 50 mM Tris-HCl, pH 7.5) was added for 100 μl protoplast mixture, incubated at room temperature for 30 min and plated on the regeneration selection medium (PDA, 1.2 M sorbitol, 250 μg/ml hygromycin B). After incubation at room temperature for about 2 weeks, the transformants were inoculated in PDB medium with stationary incubation for about 1 week.

Mycelia were collected, lyophilized and grounded for cell disruption. Cell lysate was solubilized in LETS buffer (10 mM Tris-HCl, pH 8.0, 20 mM EDTA, 0.5% SDS, 0.1 M LiCl) and extracted twice with phenol/chloroform. Genomic DNA was precipitated with ethanol, and resuspended in TE buffer. The genotypes of all mutants were verified by PCR.

**RNA preparation and reverse transcription PCR (RT-PCR)**:

*C. arbuscula* wild type and the *ΔhdaA* mutant were grown on solid PDA media (BD) at room temperature for 4 weeks. Mycelia were collected and RNA was prepared by the hot acid phenol method with minor modifications.[2] Briefly, mycelia were resuspended in 500 μl of TES buffer (50 mM Tris-HCl, pH 7.2, 10 mM EDTA, 1% SDS) with 100 mM β-mercaptoethanol and 500 μl of acid phenol (phenol/chloroform/isopentanol, pH 6.2, Ambion). The extraction mixture was incubated at 65°C for 1 hour with vortex every 15 min,

and centrifuged at 15000 rpm for 15 min to save the aqueous phase. The supernatant was extracted twice again with the acid phenol. RNA was precipitated with ethanol and resuspended in RNase-free water. Genomic DNA was further removed by digestion with RNase-free DNase I (Ambion). RNA was purified by acid phenol extraction and ethanol precipitation. RNA integrity was confirmed by electrophoresis on TBE agarose gel, and concentration was determined by Nanodrop (Thermo Scientific).

cDNA was prepared from 500 ng of total RNA by SuperScript® II Reverse Transcriptase (Invitrogen) with random primers as described by the manufacture. PCR was performed with GoTaq® Green Master Mix (Promega) with 2 ng of reverse transcribed RNA. Primers and PCR cycles were listed in Table S7-S10.

**Chemical analysis and characterization of compounds:**

LC-MS was conducted with a Shimadzu 2010 EV liquid chromatography mass spectrometer by using both positive and negative electrospray ionization and a Phenomenex Luna 5 μm, 2.0 mm × 100 mm $C_{18}$ reverse-phase column. Samples were dissolved in methanol before separated on a linear gradient of 5% to 99% $CH_3CN$ (v/v) in $H_2O$ at a flow rate of 0.1 mL/min. $^1H$, $^{13}C$ and 2D NMR spectra were obtained using $CDCl_3$ or DMSO-$d_6$ as solvent on a Bruker AV500 spectrometer with a 5 mm dual cryoprobe at the UCLA Molecular Instrumentation Center.

**Purification of aurovertin B (1) and D (2) from *C. arbuscula* wild type:**

*C. arbuscula* was cultivated in PDB (6 L) at 25 °C for 10 d. The culture was filtered through cheesecloth. The mycelia were extracted with acetone for three times. The combined extracts were evaporated to dryness under reduced pressure to afford the residue (15.6 g).The residue was dissolved in $H_2O$ (1 L) to form a suspension, followed by extraction with $CHCl_3$ (4 × 500 mL). The combined organics were dried over MgSO4. The crude products (6.0 g) were separated by a RP-18 column (Redi*Sep* Rf Gold C18 Column, 20–40 μ, 86 g, MeCN-$H_2O$ 2:8→5:5) to afford four fractions (Fr.1–4). Fr.1 was purified by flash chromatography (MeCN-$H_2O$ 1:9→3:7) to provide aurovertins D (**2**) (33 mg). Aurovertin B (**1**) (110 mg) was obtained from Fr. 3 by recrystallization with MeCN-$H_{2O}$ (9:1).

**Purification of compounds 3-12 from the *ΔhdaA* mutant**:

*C. arbuscula ΔhdaA* mutant were grown on 6 liter of solid PDA media (BD) divided into 60 large 150×15 $mm^2$ Petri dishes at room temperature for 30 days. The cultures were extracted with 1.0 liter of acetyl acetate and evaporated to dryness (1.5 g). The dried extract was partitioned between hexane and methanol. The MeOH fraction was evaporated to dryness (1.0 g), and purified by ISCO-CombiFlash® Rf 200 (Teledyne Isco, Inc) with a gradient of $H_2O$ and $CH_3CN$ (linear gradient of 20% to 80% $CH_3CN$ over 50 min at 20 mL/min on 100 g reversed phase silica gel ($C_{18}$) to give 17 fractions (Fr.1~ Fr.17). Fraction Fr.10 was purified by Sephadex LH-20 eluted with MeOH to give compound **8** (4.0 mg). Fraction Fr.11 was purified by Sephadex LH-20 eluted with MeOH to give 14 subfractions (Fr.11-1~ Fr.11-14). Fr.11-3 was further purified by HPLC eluted with 35% $CH_3CN$ in $H_2O$ at a flow rate of 2.5 mL/min to give compound **9** (4.0 mg, $t_R$= 30.5 min) and **4** (1.0 mg, $t_R$= 30.5 min). Fr.11-6 was further purified by HPLC with a gradient of $H_2O$ and $CH_3CN$ (linear gradient of 30% to 50% $CH_3CN$ over 30 min at 2.5 mL/min) to give compound **12** (1.2 mg, $t_R$= 25.5 min). Fr.11-12 was purified by Sephadex LH-20 eluted with MeOH to give compound **7** (1.1 mg). Fr.14 was further purified by HPLC with a gradient of $H_2O$ and $CH_3CN$ (linear gradient of 50% to 65% $CH_3CN$ over 30 min at 2.5 mL/min) to give **5** (1.5 mg, $t_R$= 14.2 min), **3** (8 mg, $t_R$ = 16.2 min), and **10** (1.4 mg, $t_R$= 17.2 min). Pure **3** was crystallized as white needles from MeOH/$CHCl_3$ (3:1) with an initial concentration of 4 mg/ml. Fraction Fr.15 was purified by Sephadex LH-20 eluted with MeOH to give compound **6** (3.0 mg) and **11** (10 mg). **11** was crystallized as white needles from

$CH_3CN$ / $H_2O$ (6:4) with an initial concentration of 4 mg/ml.

***arbumycin*** (**3**), white needles. $^1$H NMR (500 MHz, $CDCl_3$) and $^{13}$C NMR (125 MHz, $CDCl_3$) see Table S11 and attached spectra. ESI-MS *m/z* 514.2 $[M + H]^+$ and 536.2 $[M + Na]^+$. HRESIMS *m/z* 514.31077 $[M + H]^+$ (calcd for 514.31284, $C_{25}H_{44}N_3O_8$); 512.29621[M - H]$^-$ (calcd for 512.29719, $C_{25}H_{42}N_3O_8$).

***arbumelin*** (**5**), white powder; $^1$H NMR (500 MHz) and $^{13}$C NMR (125 MHz) in $CD_3OD$ or DMSO-$d_6$ as well as 2D NMR spectra see Table S12-S13 and attached spectra. ESI-MS *m/z* 910.2 $[M + H]^+$ and 922.2 $[M + Na]^+$. HRESIMS *m/z* 910.5062 $[M + H]^+$ (calcd for 910.4674, $C_{44}H_{64}N_9O_{12}$); 932.4591[M + $H^+$ (calcd for 932.4494, $C_{44}H_{64}N_9O_{12}Na$).

***arbusculic acid A*** **(11)** , white needles. $^1$H NMR (500 MHz, DMSO-$d_6$) and $^{13}$C NMR (125 MHz, DMSO-$d_6$) see Table S14 and attached spectra. ESI-MS *m/z* 351.1 $[M + H]^+$ and 349.1 [M - H]$^-$.

***arbusculic acid B*** **(12),** yellow powder; $^1$H NMR (500 MHz) and $^{13}$C NMR (125 MHz) as well as 2D NMR spectra in $CD_3OD$ or DMSO-$d_6$ see Table S15 and attached spectra. ESI-MS *m/z* 707.2[M + H]$^+$ and 705.2 [M - H]$^-$. HRESIMS *m/z* 705.2545[M - H]$^-$ (calcd for 705.25526, $C_{38}H_{41}O_{13}$).

**Structure characterization of 3-4 and absolute configuration of amino acid moieties in arbumelin (4)**:

To elucidate the structures new peptides (**3**-**4**), a combination of COSY, HSQC, HMBC and HSQC-TOCSY were firstly used to identify amino acids or hydroxyl acids moieties. Sequential assignment was mainly based on the correlations between the proton and HN of amino acid i with the corresponding resonances of amino acids i+1 and i–1, respectively, in the NOESY spectrum, and HMBC correlations from the proton and the NH of amino acid i to the carbonyl of amino acid i and i+1.[3]

0.4 mg of **4** was subjected to acid hydrolysis at 110 °C for 18 h with 6 N HCl (600 μL), and then the hydrolysates were dried under a steam of $N_2$ gas and redissolved in $H_2O$ (200 μL). To one portion (50 μL) was added 20 μL of 1M $NaHCO_3$ and 100 μL of 1% *N*-$\alpha$-(2,4-Dinitro-5-fluorophenyl)-L-alaninamide (FDAA) in acetone, after which the mixtures were heated at 80 °C for 5 min. The reaction mixtures were cooled, neutralized with 1N HCl (20 μL) and diluted with 300 μL $CH_3CN$. About 150 μL of each solution of FDAA derivatives was analyzed by LC-MS using a $C_{18}$ reverse-phase column (Phenomenex Luna, 2.0 × 100 mm, 5μ). Aqueous $CH_3CN$ containing 0.1% formic acid was used as mobile phase with linear gradient elution (30-55% for 40 min) at a flow rate of 0.1 ml/min. FDAA-derivatived amino acids were detected by absorption at 340 nm.

The retention times (min) of the FDAA-derivatized standards were 33.9 min for D-Leu (*m/z*: 384.1 $[M+H]^+$), 29.5 min for L-Leu (*m/z*: 384.1 $[M+H]^+$), 23.1 min for L-Tyr (*m/z*: 434.1 $[M+H]^+$), 24.5 min for D-Tyr (*m/z*: 434.1 $[M+H]^+$), 29.5 min for D-Val (*m/z*: 370.1 $[M+H]^+$), 25.1 min for L-Val (*m/z*: 370.1 $[M+H]^+$), 22.0 min for L-Ser (*m/z*: 358.1 $[M+H]^+$), 22.5 min for D-Ser (*m/z*: 358.1 $[M+H]^+$), and 16.1 min for Gly (*m/z*: 328.1$[M+H]^+$).

The retention times of the FDAA-derivatized hydrolysate of **4** were 29.5 min (L-Leu, *m/z*: 384.1 $[M+H]^+$), 23.1 min(L-Tyr, *m/z*: 434.1 $[M+H]^+$), 29.5 min (D-Val, *m/z*: 370.1 $[M+H]^+$), 25.1 min (L-Val *m/z*: 370.1 $[M+H]^+$), 22.0 min (L-Ser, *m/z*: 358.1 $[M+H]^+$), 16.1 min (Gly, *m/z*: 328.1$[M+H]^+$).

Since there is one each of D and L-valine in the peptide, we assigned the position of D-valine in the peptide at position 5 based on the presence of the E domain in module 5. Similarly, since no O-Me-D-Tyr standard was available, we similarly assigned the stereochemistry of this position as D based on the presence of an epimerase domain.

**Table S1**. Polyketide synthase (PKS)

| No | Contig | Module organization | PKS type | Homolog accession number | Species with the homolog |
|---|---|---|---|---|---|
| 1 | 5 | KS-AT-ACP-TE | NR-PKS | XP_001225797.1 | *Chaetomium globosum* CBS 148.51 |
| 2 | 157 | KS-AT-DH-MT-ER-KR-ACP | HR-PKS | EJP62832.1 | *Beauveria bassiana* ARSEF 2860 |
| 3 | 517 | KS-AT-DH-ER-KR-ACP | HR-PKS | EGR46571.1 | *Trichoderma reesei* QM6a |
| 4 | 661 | KS-AT-DH-MT-ER-KR-ACP | HR-PKS | EMR65646.1 | *Eutypa lata* UCREL1 |
| 5 | 692 | KS-AT-DH-MT-ER-KR-ACP | HR-PKS | XP_003235732.1 | *Trichophyton rubrum* CBS 118892 |
| 6 | 840 | KS-AT-DH-ER-KR-ACP | HR-PKS | XP_390724.1 | *Fusarium graminearum* PH-1 |
| 7 | 921 | KS-AT-ACP-MT-TR | NR-PKS | BAD44749.1 | *Monascus purpureus* |
| 8 | 959 | KS-AT-DH-MT-ER-KR-ACP | HR-PKS | ABQ42548.1 | *Solorina crocea* |
| 9 | 1048 | KS-AT-ACP-MT-TR | NR-PKS | EFY84249.1 | *Metarhizium acridum* CQMa 102 |
| 10 | 1080 | KS-AT-DH-MT-ER-KR-ACP | HR-PKS | EMR61679.1 | *Eutypa lata* UCREL1 |
| 11 | 1333 | KS-AT-ACP-MT-TR | NR-PKS | EFY89365.1 | *Metarhizium acridum* CQMa 102 |
| 12 | 1452 | KS-AT-DH-MT-KR-ACP | HR-PKS | EFY96172.1 | *Metarhizium anisopliae* ARSEF 23 |
| 13 | 1462 | KS-AT-DH-MT-KR-ACP | HR-PKS | ACM42412.1 | *Chaetomium chiversii* |
| 14 | 1530 | KS-AT-ACP | NR-PKS | EFY97792.1 | *Metarhizium anisopliae* ARSEF 23 |
| 15 | 1538 | KS-AT-DH-ER-KR-ACP-TE | HR-PKS | EMT73989.1 | *Fusarium oxysporum* f. sp. cubense race 4 |
| 16 | 1693 | AT-KS-AT-ACP-ACP-TE | NR-PKS | ACH72912.1 | *Aspergillus ochraceoroseus* |
| 17 | 1737 | KS-AT-DH-MT-ER-KR-ACP | HR-PKS | EFY96790.1 | *Metarhizium anisopliae* ARSEF 23 |
| 18 | 1781 | KS-AT-DH-MT-ER-KR-ACP | HR-PKS | EFY94872.1 | *Metarhizium anisopliae* ARSEF 23 |
| 19 | 1936 | KS-AT-DH-MT-ER-KR-ACP | HR-PKS | CCT74300.1 | *Fusarium fujikuroi* IMI 58289 |
| 20 | 1978 | KS-AT-DH-MT-ER-KR-ACP | HR-PKS | AFN68295.1 | *Alternaria alternata* |
| 21 | 1998 | KS-AT-DH-MT-ER-KR-ACP | HR-PKS | EJP68806.1 | *Beauveria bassiana* ARSEF 2860 |
| 22 | 2016 | KS-AT-DH-MT-KR-ACP-P450 | HR-PKS | EKJ68630.1 | *Fusarium pseudograminearum* CS3096 |
| 23 | 2032 | KS-AT-DH-MT-ER-KR-ACP-mevDPdecarb | HR-PKS | XP_002485885.1 | *Talaromyces stipitatus* ATCC 10500 |
| 24 | 2051 | KS-AT-DH-MT-ER-KR-ACP | HR-PKS | XP_002486631.1 | *Talaromyces stipitatus* ATCC 10500 |
| 25 | 2134 | KS-AT-DH-MT-KR-T-C-A-T-E | PKS-NRPS | EME40768.1 | *Dothistroma septosporum* NZE10 |
| 26 | 2142 | KS-AT-DH | NR-PKS | XP_002152245.1 | *Talaromyces marneffei* ATCC 18224 |
| 27 | 2163 | KS-AT-DH-MT-ER-KR-ACP | HR-PKS | KFH43187.1 | *Acremonium chrysogenum* ATCC 11550 |
| 28 | 2174 | KS-AT-DH-MT-ER-KR-ACP | HR-PKS | EHK20022.1 | *Trichoderma virens* Gv29-8 |
| 29 | 2197 | KS-AT-ACP | NR-PKS | ADI24953.1 | *Penicillium aethiopicum* |
| 30 | 2221 | A-T-KS-AT-KR-ACP-TE | NRPS-PKS | EMR62731.1 | *Eutypa lata* UCREL1 |
| 31 | 2230 | KS-AT-DH-ACP-TE | HR-PKS | KEY71921.1 | *Stachybotrys chartarum* IBT 7711 |
| 32 | 2290 | KS-AT-DH-MT-KR-ACP | HR-PKS | GAA83278.1 | *Aspergillus kawachii* IFO 4308 |
| 33 | 2305 | KS-AT-ACP-TE | NR-PKS | EHK23001.1 | *Trichoderma virens* Gv29-8 |
| 34 | 2333 | KS-AT-DH-ER-KR-ACP | HR-PKS | EGX94103.1 | *Cordyceps militaris* CM01 |
| 35 | 2361 | KS-AT-DH-ER-KR-ACP | HR-PKS | XP_007786831.1 | *Endocarpon pusillum* Z07020 |
| 36 | 2384 | KS-AT-DH-MT-ER-KR-ACP | HR-PKS | XP_002152245.1 | *Talaromyces marneffei* ATCC 18224 |
| 37 | 2385 | KS-AT-DH-MT-KR-C-A-T-TE | PKS-NRPS | EMR62487.1 | *Eutypa lata* UCREL1 |
| 38 | 2437 | KS-AT-DH-KR-ACP | HR-PKS | EJP63776.1 | *Beauveria bassiana* ARSEF 2860 |
| 39 | 2521 | KS-AT-ACP-MT-TE | NR-PKS | XP_007596699.1 | *Colletotrichum fioriniae* PJ7 |
| 40 | 2686 | KS-AT-DH-MT-KR-ACP | HR-PKS | EFZ04272.1 | *Metarhizium anisopliae* ARSEF 23 |
| 41 | 2720 | KS-AT-DH-ER-KR-ACP-A-T-TR | PKS-NRPS | EFY98483.1 | *Metarhizium anisopliae* ARSEF 23 |

**Table S2**. Non-ribosomal peptide synthase (NRPS) (excluding PKS-NRPS hybrids)

| No | Contig | Module organization | Homolog accession number | Species with the homolog |
|---|---|---|---|---|
| 1 | 5 | ATC-ATEC-ATC-ATC-A(MT)TC-A(MT)TC | KFG77713.1 | *Metarhizium anisopliae* ARSEF 23 |
| 2 | 453 | ATC-ATC-ATC-ATEC-ATEC-ATC-ATC-ATC-ATC | EHK50755.1 | *Trichoderma atroviride* IMI 206040 |
| 3 | 567 | ATC-ATC | EGR45389.1 | *Trichoderma reesei* QM6a |
| 4 | 840 | ATC-ATEC-ATEC-ATEC-ATC-ATC | XP_390720.1 | *Fusarium graminearum* PH-1 |
| 5 | 921 | TC-ATC-ATR | EHK45804.1 | *Trichoderma atroviride* IMI 206040 |
| 6 | 1462 | TC-ATEC-ATEC-ATC-ATC-ATEC-ATC | BAO73252.1 | *Lecanicillium* sp. HF627 |
| 7 | 1538 | C-ATC-TE | EIT81855.1 | *Aspergillus oryzae* 3.042 |
| 8 | 1747 | ACT-ATEC-ATC-ATC-ATC | EFY94500.1 | *Metarhizium anisopliae* ARSEF 23 |
| 9 | 1811 | ATC-ATC-ATC-ATC-ATC | EHK21622.1 | *Trichoderma virens* Gv29-8 |
| 10 | 1981 | TC-ATEC-ATC | XP_007826599.1 | *Metarhizium anisopliae* ARSEF 23 |
| 11 | 2021 | C-ATEC-ATC-ATC-ATC-ATEC | EHK21757.1 | *Trichoderma virens* Gv29-8 |
| 12 | 2034 | ATC-ATC-TC-ATC-TC-TC | EHK26839.1 | *Trichoderma virens* Gv29-8 |
| 13 | 2353 | ATC-ATEC-ATC-ATEC-ATC-ATEC | EHK20429.1 | *Trichoderma virens* Gv29-8 |
| 14 | 2362 | ATEC-ATC-ATR | XP_008596990.1 | *Beauveria bassiana* ARSEF 2860 |
| 15 | 2384 | TC-ATC-ATC-ATC-ATC-ATEC-ATC-ATC-ATEC-ATC-TE | KFH42120.1 | *Acremonium chrysogenum* ATCC 11550 |
| 16 | 2410 | ATC-ATC-ATC-ATC-A(MT)TC-ATEC | XP_003044554.1 | *Nectria haematococca* mpVI 77-13-4 |
| 17 | 2474 | ATC-TC | KEY64460.1 | *Stachybotrys chartarum* IBT 7711 |
| 18 | 2507 | ATC-ATC-ATC-ATC-ATC | EHK21757.1 | *Trichoderma virens* Gv29-8 |

**Table S3**. Genes involving terpene synthesis

| No | Contig | Putative functions | Homolog accession number | Species with the homolog |
|---|---|---|---|---|
| 1 | 693 | Isoprenoid Biosynthesis enzymes, class I | CAX94841.1 | *Trichoderma brevicompactum* |
| 2 | 797 | Trans-Isoprenyl Diphosphate Synthase | XP_002846999.1 | *Arthroderma otae* CBS 113480 |
| 3 | 1299 | Trichodiene synthase (TRI5) | AGZ87953.1 | *Trichoderma brevicompactum* |
| 4 | 1737 | ent-copalyl diphosphate synthase | KFH47392.1 | *Acremonium chrysogenum* ATCC 11550 |
| 5 | 1789 | Trans-Isoprenyl Diphosphate Synthases | XP_390273.1 | *Fusarium graminearum* PH-1 |
| 6 | 1930 | ent-kaurene synthase | XP_002849529.1 | *Arthroderma otae* CBS 113480 |
| 7 | 1936 | Copalyl diphosphate synthase | KFH47392.1 | *Acremonium chrysogenum* ATCC 11550 |
| 8 | 2327 | pentalenene synthase | ETS03221.1 | *Trichoderma reesei* RUT C-30 |
| 9 | 2560 | Isoprenoid Biosynthesis enzymes, Class 1 | EAA31071.2 | *Neurospora tetrasperma* FGSC 2509 |

**Table S4.** A domain signatures in contig 2507 NRPS.

| | Residues in A domain signature | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| A1 domain | D | G | L | F | I | G | I | P | I | K |
| A2 domain | A | G | Q | F | C | G | T | T | Y | K |
| A3 domain | D | V | Y | T | V | L | A | V | L | K |
| A4 domain | D | M | H | I | L | M | G | C | F | K |

**Table S5**. A domain signatures in contig 453 NRPS.

| | Residues in A domain signature | | | | | | | | | | Proposed substrate |
|---|---|---|---|---|---|---|---|---|---|---|---|
| A1 domain | D | L | Y | P | I | A | I | P | V | K | L-Leu |
| A2 domain | D | A | F | P | A | G | C | P | V | K | L-Tyr |
| A3 domain | D | M | F | Q | L | I | T | P | V | K | Gly |
| A4 domain | D | M | S | G | I | A | G | P | C | K | D-*O*-methyl-Tyr |
| A5 domain | D | I | L | Y | I | G | A | P | A | K | D-Val |
| A6 domain | D | V | M | L | L | G | V | P | C | K | L-Ser |
| A7 domain | D | M | F | Q | L | I | A | P | C | K | Gly |
| A8 domain | D | M | F | Q | L | I | S | P | C | K | Gly |
| A9 domain | D | M | F | N | L | G | A | P | F | K | L-Val |

**Table S6**. Primers for gene knock-out in *C. arbuscula*.

| Primer name | Primer sequence |
|---|---|
| hph-up F | ACCTGGCGGCCGCTACAACGACCATCAAAGTC |
| hph-up R | ACTGACCGCGTCTGCTGCTCCATACAA |
| hph-dn F | CAGGGAGCTCTCGGAGGGCGAAGAATC |
| hph-dn R | GAATGCGGCCGCAGGATTACCTCTAAACAAGTG |
| arbA-up F | CCTGGTACCGTTGGGATGTG |
| arbA-up R | AAGCGGCCGCTTGGAATTCACA |
| arbA-dn F | AAAGCGGCCGCAACGTCTACGC |
| arbA-dn R | AAGGTACCCTAGATCATCAGCCT |
| hdaA-up F | CGCAAGAAAGTCCAAGCC |
| hdaA-up R | ACCTGGCGGCCGCTCCATCCTAGCGTAGTGAC |
| hdaA-dn F | CCTGGCGGCCGCGGACACAACTGAGGAAGATG |
| hdaA-dn R | ACCTGTCTAGACTGGGGGAATCGTTGGAC |
| M13-F | GTAAAACGACGGCCAGTG |
| M13-R | CCTTTGTCGATACTGGTAC |

**Table S7**. RT-PCR primers for polyketide synthases (PKS)

| contig | sequence | cycle number | contig | sequence | cycle number |
|---|---|---|---|---|---|
| 5 | GCGGCAGTCAGGCTTGG | 35 | 2016 | ATGCGTATCCCGATGG | 35 |
| | CTTTGGGCTTGGTCGGAG | | | GCGAGCAGGTATTTCTGG | |
| 157 | CGCAATTACAGACAAAGAC | 35 | 2032 | CACCTCGTCTCAAAATGG | 28 |
| | GGACGTAGTGAGAGGAC | | | GGGAAGTTGTTCTCGGAC | |
| 517 | CCGGCGGCAGGAATGGTG | 28 | 2051 | GCAAACTTACTCTTGGAC | 35 |
| | GTCCGGGATACGAACTTTG | | | TGGCATCGGCTTCAATAG | |
| 661 | CGCGGGATGTGGGTTAC | 35 | 2134 | GACTCCACTCAGGCTCTC | 35 |
| | CGGCGTGTGCGTTTGTG | | | GAAATGACACCTTCGGC | |
| 692 | GCCGGCATGGGGAAAC | 35 | 2142 | GAGATTGCGCAGCCATTG | 28 |
| | TGGCGCCGGCAGCATAAG | | | GCCCGAGACTGTGACAC | |
| 840 | TTCCGGCGTCGTCTCC | 35 | 2163 | CCGCCCGGCTGGATTG | 35 |
| | GGCGTGTCACAGGTCC | | | TCCCCTACCTCATCTAAC | |
| 921 | CGCCTCGGCATACCTGTC | 28 | 2174 | TGGTACGAGGGGGTGGG | 28 |
| | AGTTCGCCCTCAGTCCAG | | | TGAGACTGTGGCGGCAAG | |
| 959 | GAGCCCGGAGGATGTTAC | 28 | 2197 | CGAGAAGAGGTGAAGAC | 28 |
| | CTGGCCCTTGCGAATGG | | | GCGGAACAATGGAGAGTC | |
| 1048 | CGACGCCGACGCCTTTG | 28 | 2221 | GGTTTAGCGACGGAATC | 35 |
| | AGAGCCCGTACTGCCGAA | | | CCTGGTGCCAAGTAATG | |
| 1080 | GCACTGGGCAAATAGAG | 35 | 2230 | AGCACAAGCCCCTACCG | 28 |
| | GAAGCGACCAAGGGGAG | | | AAGGACCAACCACCAAGG | |
| 1333 | CCTGCCTCTCCCCTTCC | 35 | 2290 | TCGCGCTGGTGCCTTGG | 35 |
| | TCCCCTTCCAGAATGCG | | | TTAGAACCCTGGACAAAAG | |
| 1452 | CTTCCTTCTCCATCCG | 35 | 2305 | CGTGGTTCTATGCTCGC | 35 |
| | GTCGCGAGTGAACAAG | | | ACGCAAATGGTTCGCAA | |
| 1462 | CCGTAAGCTCTGTTTCC | 28 | 2333 | CGCCGGTCCTCTACAAC | 35 |
| | CTCCGGCGGCGTAGTTG | | | GTCTTCTTCGGCAAACTC | |
| 1530 | CTGGTGGAAATACGAC | 35 | 2361 | GCAGTGTCCATTCCTTTC | 28 |
| | AAAACCCGAGCCCTGAG | | | CGTTGTCTGCGAGTAGTC | |
| 1538 | CCGAAGCCGTCATCTAC | 35 | 2384 | GGACCCTCAACAACGAC | 28 |
| | CTCGTTGAATCTTCTTGGG | | | TACGCCGCCGACGAAG | |
| 1693 | TTGTCATACACAACCACTG | 28 | 2385 | TTGGCAACAGTATCACG | 35 |
| | CAAGGCTAACTGAACGAC | | | GTCTTTGTGTAACCCTCC | |
| 1737 | TTGGCGCGGGTACAGG | 35 | 2437 | CTCGGGCGCAGTCTTAC | 28 |
| | CAAATGGTAAGAGCAAAGG | | | CGGATTACCAAGCACGG | |
| 1781 | GAAGGCGGGCAAGATG | 35 | 2521 | TTCCCAAGATTTGACGGC | 28 |
| | GGACAGAAAACAGAAGC | | | CTGCTGGCGACATTGCG | |
| 1936 | CTGAAGTTCAGACTATG | 35 | 2686 | CACCCTCGTATTCCTGA | 35 |
| | GCCAGTGATAAGATTCG | | | GACCAATAAGCCAGTAG | |
| 1978 | GCGGCATACTGTGTTGG | 28 | 2720 | CACCGCCAGCCCATTCAC | 28 |
| | ATTCTTCAGAAACCTGTTG | | | CGAGCGCGACATAGCAGG | |
| 1998 | CACCCTCTAAATGTCGC | 35 | | | |
| | CGTCCGAACAGTAATGC | | | | |

**Table S8**. RT-PCR primers for non-ribosomal peptide synthases (NRPS)

| contig | sequence | cycle number |
|---|---|---|
| 5 | CAGGCGCGTATGTAATG | 35 |
| | CGGGCTTGTTCAGTATC | |
| 453 | GGCTCAGCTCCCTATTC | 28 |
| | GTCCCCTAGACCCAGAG | |
| 567 | CTGGCTGCTGCTGCGAG | 28 |
| | GAGCCGTAAGTCATTGTTG | |
| 840 | GAACTCCAATGCTCTGAC | 35 |
| | ACCTCGGCCAAACCCTG | |
| 921 | CGAGTCATCCATTTGCTC | 28 |
| | AGGGCTGTGAGATTTTGATA | |
| 1462 | GAAGTTCTAAAGCAGGTC | 28 |
| | CATCACAAGGCTCTCTAC | |
| 1538 | CCATATCTAACCTACCAG | 35 |
| | ACACTGCTGTTTCCGATG | |
| 1747 | TCCAACCGCATTAGCCC | 28 |
| | ATCCTCGAGCCCGTCAA | |
| 1811 | TTAGGCGGAGAAGAAGTC | 28 |
| | CGTCGGCCGGATTGTATT | |
| 1981 | GCCCAAGACAGCCCAAG | 35 |
| | CAAGCTGCGTGCCCACA | |
| 2021 | GGTCCCGGCCTTTTCAC | 35 |
| | TGTAGGTAGTCGTCAAG | |
| 2034 | CAGCAATATCGGCAGGC | 28 |
| | CGTGATACAATAGTAGCAA | |
| 2353 | CTCCTGGGACGGCGAAC | 35 |
| | CGTGAGCAGCGATACTC | |
| 2362 | CCCTCCCTCTTGCGAC | 35 |
| | CGAGTTATCCTGACCATC | |
| 2384 | TTCGGGCATGGTCCTAC | 28 |
| | GCGCCGCAGAGAAACAC | |
| 2410 | CTGCACCGCCCTGGAGA | 28 |
| | GAGCGCAGATATGAGAAC | |
| 2474 | TTGTTCGCCCAGCCCTC | 35 |
| | GAAGTTGCTCTGTCTAAATAC | |
| 2507 | TGCCGGCTGTATTCAAGA | 28 |
| | CCAGTAGAAGATTTATCG | |

**Table S9**. RT-PCR primers of genes for terpene synthesis

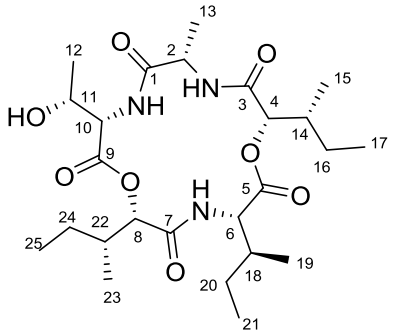
| contig | sequence | cycle number in RT-PCR |
|---|---|---|
| 693 | CCACTGCGTCGGGTCTAC | 28 |
| | CGCAATCTGGAATAACTC | |
| 797 | CAGCGAAGCCAACGATG | 28 |
| | CCAGAGCAAGGAAGTAGAC | |
| 1299 | GGTGGCGTCTGGTAAAC | 28 |
| | ATCGCGCTCGGCATTAAAC | |
| 1737 | CGGCTCTTTTACATCAACC | 35 |
| | GAAACCGACACCACGAC | |
| 1789 | GACGGGCGATAAGGAACG | 28 |
| | GCCAGAAGAGGTCCATCC | |
| 1930 | GCCCGAGAAACTTGAGCA | 28 |
| | CGAGCGGTACAGAGCAG | |
| 1936 | GGAGAAGCGTAGCAAGC | 28 |
| | AGCGGACCCAGGCGTATG | |
| 2327 | GCTCTGTGACTGGGGCAAC | 35 |
| | CGGCTCCGATAGAAACAC | |
| 2560 | AGACGCGGCCCCTCAAC | 28 |
| | CGAAGATGTACCCCACC | |

**Table S10**. RT-PCR primers for *hdaA* and house-keeping genes

| contig | sequence | cycle number in RT-PCR |
|---|---|---|
| *hdaA* | ACGCAGAATACAACTCGC | 28 |
| | ATCCTGCCGAAATAATGAC | |
| *actA* | AAGGCCGGTTTCGCTGG | 28 |
| | GGACGGCCTGGATGGAG | |
| *tubC* | ACAACTGGGCCAAGGGTC | 28 |
| | CGGGGAAACGCAGGCAGG | |
| *gpdA* | CATCGTCTTCCGCAATG | 28 |
| | CTGCCGTCGTAGGTCTTC | |

**Table S11.** NMR data of arbumycin (**3**) in $CDCl_3$

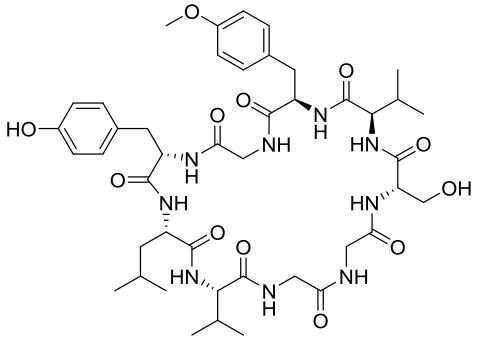
| amino acids | | No. | $_H$, mult, (*J* in Hz) | $_C$ | COSY | HMBC |
|---|---|---|---|---|---|---|
| Ala | C=O | 1 | | 172.76 C | | |
| | α | 2 | 4.21, t (7.6) | 51.58 CH | H13, NH | C1, C3, C13 |
| | β | 13 | 1.56, d (7.2) | 17.33 $CH_3$ | H2 | C1 |
| | | α–NH | 7.52 d (8.3) | | H2 | C2, C3 |
| HMVA-1 | C=O | 3 | | 171.39 C | | |
| | | 4 | 4.67, d (7.3) | 80.66 CH | H4 | C3, C14, C15 |
| | | 14 | 1.92, m | 36.32 CH | H4, H15, H16 | |
| | | 15 | 1.00, d (7.0) | 14.68 $CH_3$ | H14 | |
| | | 16 | 1.62, m | 24.89 $CH_2$ | H14, H17 | |
| | | | 1.10, m | | | |
| | | 17 | 0.91, d (7.0) | 11.33 $CH_3$ | H16 | |
| Ile | C | 5 | | 172.70 C | | |
| | | 6 | 4.62, dd (6.3, 10.0) | 57.52 CH | H18, NH | C5, C18, C19 |
| | | 18 | 2.04, m | 35.01 CH | H6, H19, H20 | |
| | | 19 | 0.98, d (7.0) | 16.18 $CH_3$ | H18 | |
| | | 20 | 1.45, m | 24.92 $CH_2$ | | |
| | | | 1.21, m | | | |
| | | 21 | 0.91, d (7.0) | 11.47 $CH_3$ | H20 | |
| | | NH | 7.77, d (10.0) | | H6 | C7 |
| HMVA-2 | C | 7 | | 170.56 C | | |
| | | 8 | H15', d (4.2) | 79.73 CH | H22 | C9, C22, C23 |
| | | 22 | 1.97, m | 36.69 CH | H8, H23, H24 | |
| | | 23 | 1.0, d (7.0) | 14.52 $CH_3$ | H22 | |
| | | 24 | 1.46, m | 25.91 $CH_2$ | | |
| | | | 1.26, m | | | |
| | | 25 | 0.91, d (7.0) | 11.64 $CH_3$ | H24 | |
| Thr-9 | C=O | 9 | | 170.29 C | | |
| | | 10 | 4.43, (7.1) | 59.33 CH | H11, NH | C12 |
| | | 11 | 4.47, m | 68.18 CH | | C9 |
| | | 12 | 1.27, d (6.5) | 20.8 $CH_3$ | | |
| | | NH | 7.44, brs | | | |




arbumycin (**3**)

**Table S12**. NMR data of arbumelin (**4**) in $CD_3OD$

| no | amino acids | | $\delta_H$, mult. (*J* in Hz) | δc | HMBC |
|---|---|---|---|---|---|
| 1 | **Leu** | **C=O** | | 174.3, C | |
| 2 | | | 4.60, dd (4.1,12.2) | 49.6, CH | C1, C3 |
| 3 | | | 1.64, ddd (4.1, 12.2, 14.4) | 38.0, $CH_2$ | |
| | | | 1.07, m | | |
| 4 | | | 1.38, m | 24.3, CH | |
| 5 | | | 0.82, d (6.6) | 19.4, $CH_3$ | |
| 6 | | | 0.79, d (6.6) | 22.6, $CH_3$ | |
| 7 | **Tyr1** | **C=O** | | 173.2, C | |
| 8 | | | 4.23, dd (5.5, 6.3) | 58.8, CH | C7, C9, C10 |
| 9 | | | 3.03, d (4.0) | 35.5, $CH_2$ | C7 |
| 10 | | Ph-1 | | 128.0, C | |
| 11 | | Ph-2, 6 | 7.26,d (8.6) | 129.8, CH × 2 | C12, C13, C11 |
| 12 | | Ph-3, 5 | 6.72, d (8.6) | 115.0, CH × 2 | C10, C12, C13 |
| 13 | | Ph-4 | | 156.0, C | |
| 14 | **Gly1** | **C=O** | | 172.4, C | |
| 15 | | | 4.44 d (17.2) | 41.7, $CH_2$ | C14 |
| | | | 3.56, d (17.2) | | C14 |
| 16 | **Tyr2** | **C=O** | | 172.4, C | |
| 17 | | | 4.38, dd (12.1, 4.0) | 56.8, CH | C16, C18, C20 |
| 18 | | | 3.22, dd (4.3, 14.4) | 35.0, $CH_2$ | C16 |
| | | | 2.67, dd (12.3, 14.6) | | C16 |
| 19 | | Ph-1 | | 129.4, C | |
| 20 | | Ph-2, 6 | 7.18, d (8.8) | 128.7, CH×2 | C19,C20, C22 |
| 21 | | Ph-3, 5 | 6.87, d (8.8) | 113.9, CH×2 | C10, C21, C22 |
| 22 | | Ph-4 | | 158.9, C | |
| 23 | | Ph-4-$OCH_3$ | 3.76, s | 54.3, $CH_3$ | C22 |
| 24 | **Val2** | C=O | | 173.7, C | |
| 25 | | | 3.74, d (6.1) | 63.2, CH | C24 |
| 26 | | | 1.89, m | 28.9, CH | C24, C25, C28 |
| 27 | | | 0.85, d (6.7) | 18.0, $CH_3$ | C25, C26 |
| 28 | | | 0.76, d (6.7) | 17.6, $CH_3$ | C25, C26 |
| 29 | **Ser** | **C=O** | | 174.7, C | |
| 30 | | | 4.74, dd (10.5, 5.9) | 53.0, CH | C29, C31, C32 |
| 31 | | | 3.99, t (10.0) | 61.7, $CH_2$ | C29 |
| | | | 3.82, dd (10.0, 6.0) | | C29 |
| 32 | **Gly2** | **C=O** | | 170.0, C | |
| 33 | | | 4.13, d (17.2) | 42.3, $CH_2$ | C32 |
| | | | 3.46, d (17.2) | | C32 |
| 34 | **Gly3** | **C=O** | | 171.1, C | |
| 35 | | | 3.91, d (17.2) | 43.6, $CH_2$ | C34,C36 |
| | | | 3.69, d (17.2) | | C36, C34 |
| 36 | **Val-1** | C=O | | 175.8, C | |
| 37 | | | 3.66, m | 62.5, CH | C36 |
| 38 | | | 2.17, m | 28.2, CH | C37, C39 |
| 39 | | | 1.07, d (6.3) | 19.4, $CH_3$ | C37, C38 |
| 40 | | | 0.94, d (6.3) | 18.2, $CH_3$ | C37, C38 |



arbumelin (**4**)

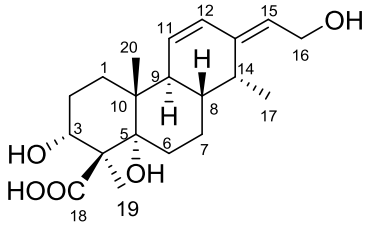
**Table S13**. NMR data of arbumelin (**4**) in DMSO-$d_6$

| no | **amino acids** | | $\delta_H$, mult. (*J* in Hz) | δc | COSY | HMBC | NOESY |
|---|---|---|---|---|---|---|---|
| 1 | **Leu** | **C=O** | | 174.3, C | | | |
| 2 | | | 4.37, ddd (13.0, 9.1, 4.2) | 49.8, CH | 2-NH, H3 , H3 | C1 | 2-NH,37-NH |
| 3 | | | 1.53, ddd (3.6, 11.1, 15.0)<br>1.08, d (15.0) | 38.2, $CH_2$ | | | |
| 4 | | | 1.33, m | 24.6, CH | | | |
| 5 | | | 0.73, d (6.4) | 20.9, $CH_3$ | | | |
| 6 | | | 0.71, d (6.8) | 23.5, $CH_3$ | | | |
| 2-NH | | | 7.80, d (8.8) | | | C7 | H4, H3β, H6 |
| 7 | **Tyr1** | **C=O** | | 171.7, C | | | |
| 8 | | | 4.05, dd (6.6, 6.6) | 58.5, CH | 8-NH, H9 | C7 | 2-NH, 8-NH, H11 |
| 9 | | | 2.86, d (6.9) | 36.1, $CH_2$ | H8 | C7 | 2-NH, 8-NH, H11 |
| 8-NH | | | 7.45, d (6.6) | | | C14 | H9, H15 |
| 10 | | Ph-1 | | 128.2, C | | | |
| 11 | | Ph-2, 6 | 7.13,d (8.5) | 130.3, CH × 2 | H12 | C12 | |
| 12 | | Ph-3, 5 | 6.63, d (8.5) | 115.6, CH × 2 | H11 | C10, C12, C13 | |
| 13 | | Ph-4 | | 156.5, C | | | |
| 14 | **Gly1** | **C=O** | | 171.5, C | | | |
| 15 | | | 4.15 dd (13.4, 9.2) | 42.4, $CH_2$ | 15-NH, H15 | C14 | 15-NH, 8-NH |
| | | | 3.38, overlap (m) | | 15-NH | C16 | |
| 15-NH | | | 7.41, m | | | | H15, H17 |
| 16 | **Tyr2** | **C=O** | | 171.2, C | | | |
| 17 | | | 4.17, m | 57.2, CH | 17-NH | C16 | 17-NH, H20, 15-NH |
| 18 | | | 3.03, dd (13.7, 3.7) | 35.3, $CH_2$ | H17 | | 17-NH |
| | | | 2.67, dd (13.7, 13.7) | | H17 | C16 | 17-NH |
| 17-NH | | | 7.846, d (6.3) | | | | H18 , H18 , H2 |
| 19 | | Ph-1 | | 129.9, C | | | |
| 20 | | Ph-2, 6 | 7.17, d (8.5) | 130.4, CH ×2 | H21 | C20, C22 | |
| 21 | | Ph-3, 5 | 6.81, d (8.5) | 114.2, CH ×2 | H20 | C19, C21, C22 | |
| 22 | | Ph-4 | | 158.5, C | | | |
| 23 | | Ph-4-$OCH_3$ | 3.69, s | 55.5, $CH_3$ | | C22 | |
| 24 | **Val2** | C=O | | 173.9, C | | | |
| 25 | | | 3.64, dd (11.5, 5.9) | 63.1, CH | 25-NH | C24 | |
| 26 | | | 2.00, m | 28.4, CH | | | |
| 27 | | | 0.93, d (6.4) | 20.4, $CH_3$ | | | |
| 28 | | | 0.84, d (6.9) | 19.5, $CH_3$ | | | |
| 25-NH | | | 9.87, brs | | H25 | | H30, H25, |
| 29 | **Ser** | **C=O** | | 174.6, C | | | |
| 30 | | | 4.72, dd (15.5, 9.2) | 53.1, CH | 30-NH | C29 | 25-NH, 30-NH |
| 31 | | | 3.68, m, overlap | 62.4, $CH_2$ | | C29 | |
| | | | 3.60, dd (11.2, 6.5) | | | | |
| 30-NH | | | 7.06, d (7.4) | | | | H30, 30-NH |
| 32 | **Gly2** | **C=O** | | 168.9, C | | | |
| 33 | | | 3.85, dd (17.1, 5.5) | 43.0, $CH_2$ | 33-NH | C32 | 33-NH, 30-NH |
| | | | 3.30, m, overlap | | | C32, C34 | 33-NH |
| 33-NH | | | 7.87, t (6.0) | | H33 | C34 | H33 , H33 |
| 34 | **Gly3** | **C=O** | | 169.6, C | | | |
| 35 | | | 3.73, dd (17.1, 5.5) | 43.6, $CH_2$ | 35-NH | C34 | 35-NH, 33-NH |
| | | | 3.57, dd (17.1, 5.5) | | | C36 | 35-NH |
| 35-NH | | | 9.22, brs | | H35 | C36 | H35 , H35 , |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | | | H33 |
| 36 | **Val-1** | C=O | | 174.8, C | | | |
| 37 | | | 3.66, m | 61.7, CH | 37-NH | | |
| 38 | | | 1.87, m | 28.9, CH | | | |
| 39 | | | 0.74 d (6.4) | 19.0, $CH_3$ | | | |
| 40 | | | 0.70, d (6.9) | 19.1, $CH_3$ | | | |
| 37-NH | | | 7.31, d (5.9) | | H37 | C1 | H2, H37 |

**Table S14**. NMR data of arbusculic acid A (**11**) in DMSO-$d_6$

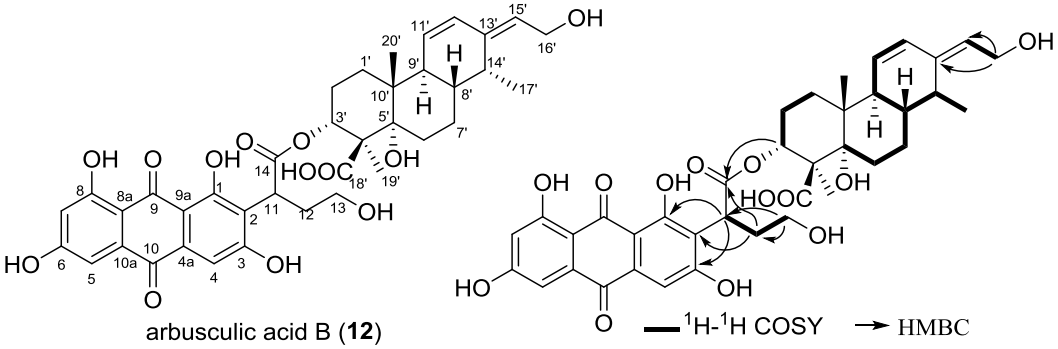
| no | $\delta_H$, mult, (*J* in Hz) | δc | COSY | HMBC | NOSEY |
|---|---|---|---|---|---|
| | 1.65, d (13.0) | 26.0, $CH_2$ | H2, H1 | C2 | |
| | 1.40, m, overlap | | H2, H1 | | |
| 2 | 2.18, dt (13.3, 3.5) | 26.5, $CH_2$ | H1 , H1 , | | H3 |
| | 1.57, d (14.0) | | H2 , H1 | | |
| 3 | 3.95, t (7.5) | 72.1, CH | H2 H2 | C5, C4, C1 | H2, H6 , H19 |
| 4 | | 49.8, C | | | |
| 5 | | 78.4, C | | | |
| 6 | 2.12, ddd (13.3, 4.0, 4.0) | 28.9, $CH_2$ | H6 | C7 | |
| 6 | 1.61, m | | | | |
| 7 | 1.72, ddd (13.3, 13.3,4.2) | 24.4, $CH_2$ | H7 | | |
| 7 | 1.08, d (12.8) | | H7 , H6 , | C8, C6 | |
| 8 | 1.45, t (11.0) | 36.6, CH | H9 | | H20 |
| 9 | 2.41, d (10.9) | 41.2, CH | H8 | C8, C11, C10, C20 | |
| 10 | | 41.9, C | | | |
| 11 | 5.61, d (10.0) | 130.0, CH | H12 | C8, C10, C13 | |
| 12 | 5.91,dd (10.0, 2.2) | 130.4, CH | H1' | C9, C13, C14 | H15 |
| 13 | | 141.7, C | | | |
| 14 | 2.45, m | 33.9, CH | H17 | C8,C12, C13, C15, C17 | H8, H16 |
| 15 | 5.19, t (6.8) | 126.5, CH | H16 | C12, C13, C14, C16 | H16 |
| 16 | 4.04, m | 57.2, $CH_2$ | H15, 16OH | C13, C15 | H14 |
| 16 | 3.96, m | | H16 | | |
| 17 | 0.78, d (7.0) | 13.9, $CH_3$ | H14 | C8, C13, C14 | H9 |
| 18 | | 177.5, C | | | |
| 19 | 1.25, s | 20.4, $CH_3$ | | C3, C4, C5,C18 | H3, H20, 5-OH |
| 20 | 0.72, s | 16.1, $CH_3$ | | C5, C1, C9 | H2, H8, H19 |
| 3-OH | 5.76, d () | / | H3 | C2, C3, C4 | H3 |
| 5-OH | 5.40, s | / | | C4, C5, C6 | H9, 19-OH |
| 16-OH | 4.51, t (5.3) | / | H16 | C15, C16 | |

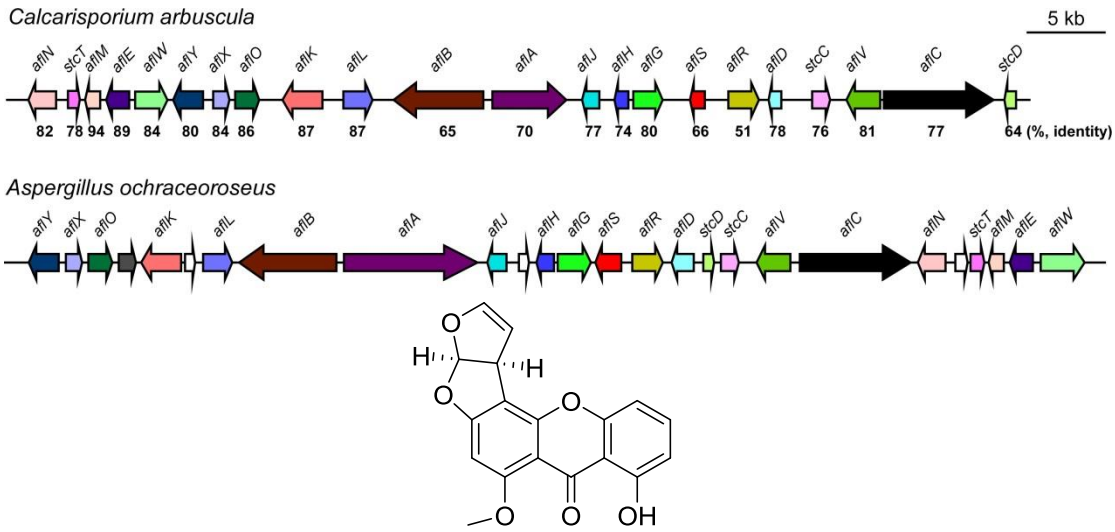



arbusculic aicd A (**11**)

**Table S15**. NMR data of arbusculic acid B (**12**) in $CD_3OD$

| no | mult, (*J* in Hz) | C | COSY | C |
|---|---|---|---|---|
| | | 162.2, C | | |
| | | 119.7, C | | |
| | | 163.1, C | | |
| | 7.28, s | 108.9, CH | | C2, C3, C10, C9a |
| | | 133.7, C | | |
| | 7.08, d (2.5) | 107.98, CH | H7 | C6, C7, C10 |
| | | 165.5, C | | |
| | 6.49, d (2.5) | 107.7, CH | H5 | C5, C6, C8 |
| | | 165.1, C | | |
| | | 108.9, C | | |
| | | 189.4, C | | |
| | | 109.1, C | | |
| | | 181.7, C | | |
| | | 135.2, C | | |
| | 4.44,dd (5.4, 9.1) | 36.5, CH | H12 , H12 | C1, C2, C12, C13, C14 |
| | 2.52, m | 31.4, $CH_2$ | H12 | C2, C11, C13, C14 |
| | 2.16,m | | H12 , H13 | |
| | 3.58, m | 59.7, $CH_2$ | H12 , H12 | C11, C12 |
| | | 172.0, C | | |
| | 1.30, m | 26.0, $CH_2$ | H2', H1' | C5' |
| | 1.01, m | | H2' | |
| 2' | 2.27, m | 22.8, $CH_2$ | H1' , H1' , H2' | |
| | 1.81, d (14.0) | | H2' , H1' | |
| 3' | 5.30, t (2.5) | 77.4, CH | H2' H2' | C14, C5', C4', C1' |
| 4' | | 50.4, C | | |
| 5' | | 77.6, C | | |
| 6' | 2.15, m | 28.3, $CH_2$ | H7' , H7' | |
| 6' | 1.59 ,m | | H7' | C5' |
| 7' | 1.53, m | 23.9, $CH_2$ | H7' | C5', C10' |
| 7' | 0.98, m | | | |
| 8' | 1.42, m | 36.3, CH | H9', H14' | |
| 9' | 1.94, | 40.7, CH | H12', H11', H8' | |
| 10' | | 41.0, C | | |
| 11' | 5.34, d (10.1) | 128.9, CH | H12', H9' | C8', C9', C13' |
| 12' | 5.73,dd (10.1, 2.7) | 129.6, CH | H11', H9' | C9', C14' |
| 13' | | 143.5, C | | |
| 14' | 2.32, m | 33.8, CH | H17' | C8' |
| 15' | 5.13, dd (6.8, 6.8) | 123.2, CH | H16' , H16' | C10', C12', C14' |
| 16' | 4.11, dd (7.4, 12.8) | 57.2, $CH_2$ | H15' | C13', C15' |
| 16' | 4.01,dd (6.8, 12.8) | | H15' | |
| 17' | 0.38, d (7.2) | 12.2, $CH_3$ | H14' | C8', C13' C14' |
| 18' | | 177.4, C | | |
| 19' | 1.17, s | 18.5, $CH_3$ | | C3', C4', C18' |
| 20' | 0.73, s | 15.5, $CH_3$ | | C5', C1', C10' |






sterigmatocystin (**6**)

**Figure S1**. Gene cluster comparison of sterigmatocystin (**6**) from *Calcarisporium arbuscula* and *Aspergillus ochraceoroseus*. Protein identities are shown on each homologous enzymes.

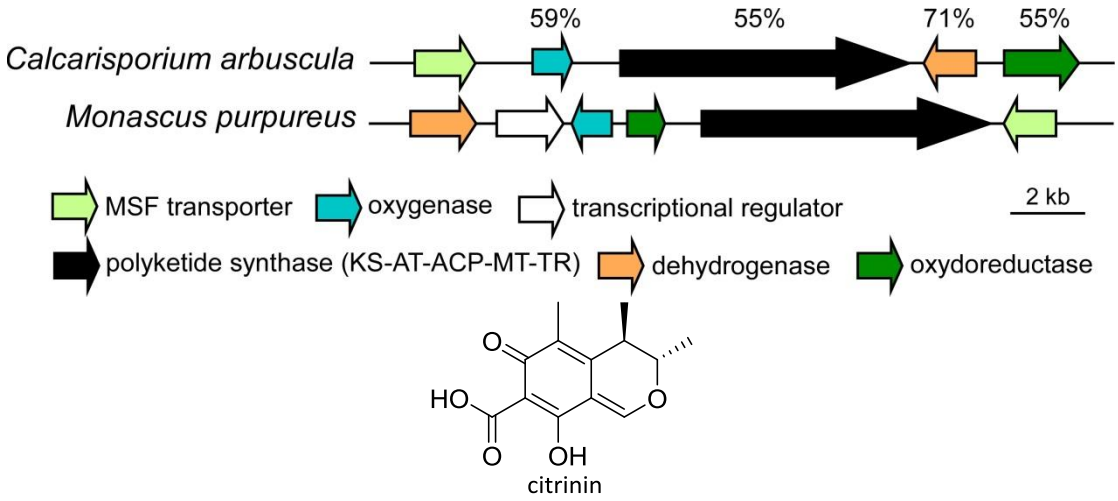



**Figure S2**. Gene cluster comparison of citrinin from *Calcarisporium arbuscula* (contig 921) and *Monascus purpureus*. Protein identities are shown on each homologous enzymes.

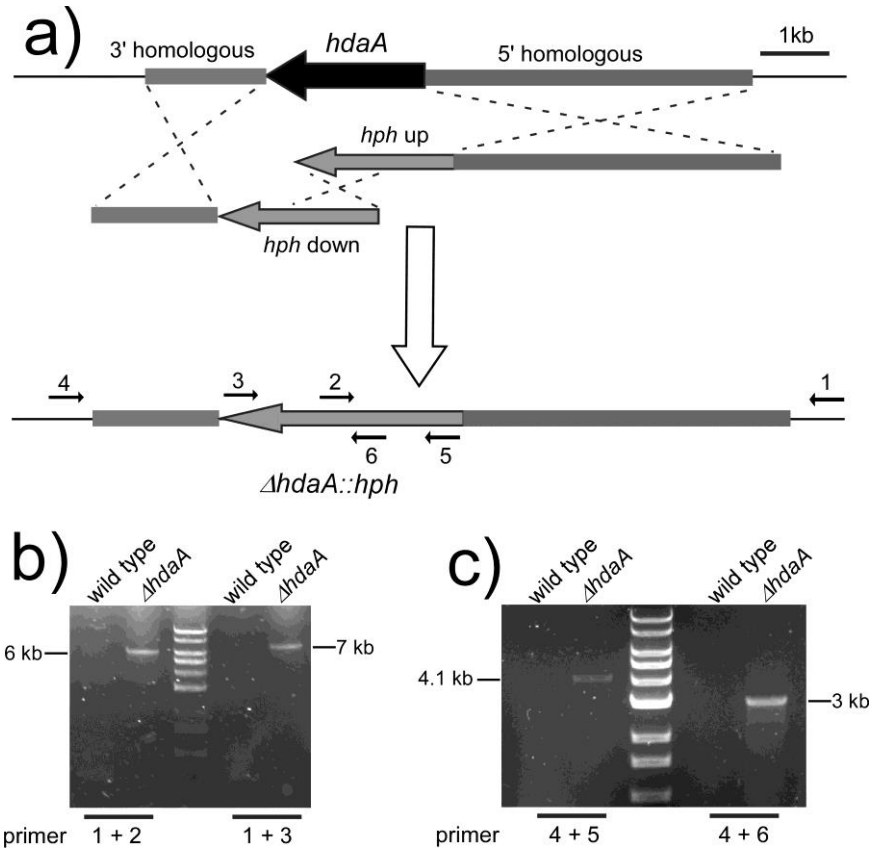



**Figure S3**. Deletion of *hdaA* gene in *C. arbuscula*.

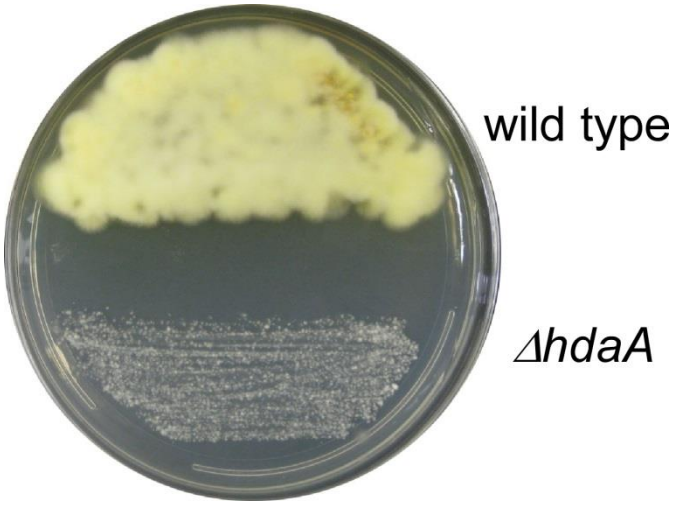



**Figure S4**. Phenotype of *C. arbuscula* wild type and the *ΔhdaA* mutant grown on PDA for 7 days.

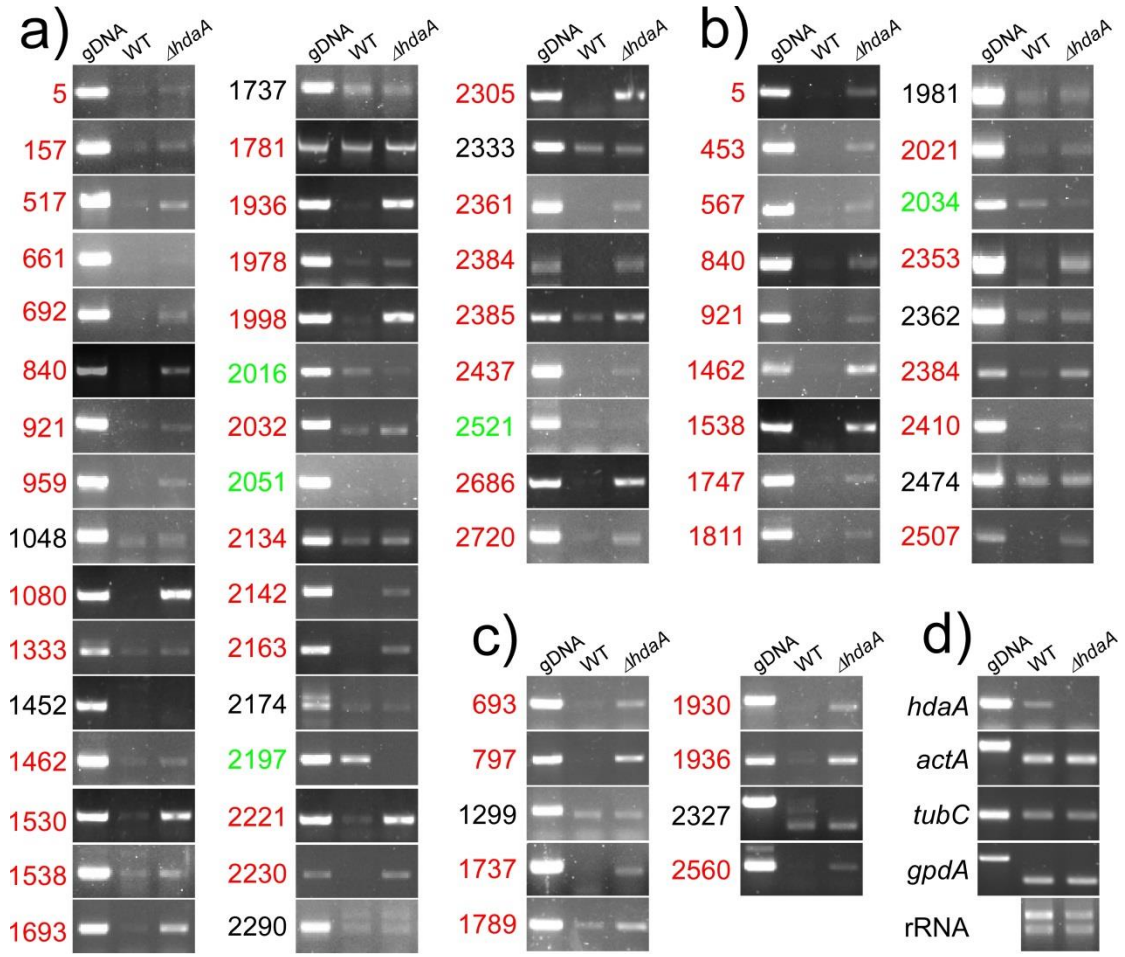



**Figure S5**. RT-PCR analysis of gene expression for biosynthetic core genes in PKS (a), NRPS (b), terpene synthesis (c) and internal controls (d) from *Calcarisporium arbuscula* wild type and the *ΔhdaA* mutant. The house-keeping genes *actA*, *tubC* and *gpdA* were the internal controls. rRNA were shown for loading control and RNA integrity. Up-regulated genes were shown in red, down-regulated genes in green and non-regulated genes in black.

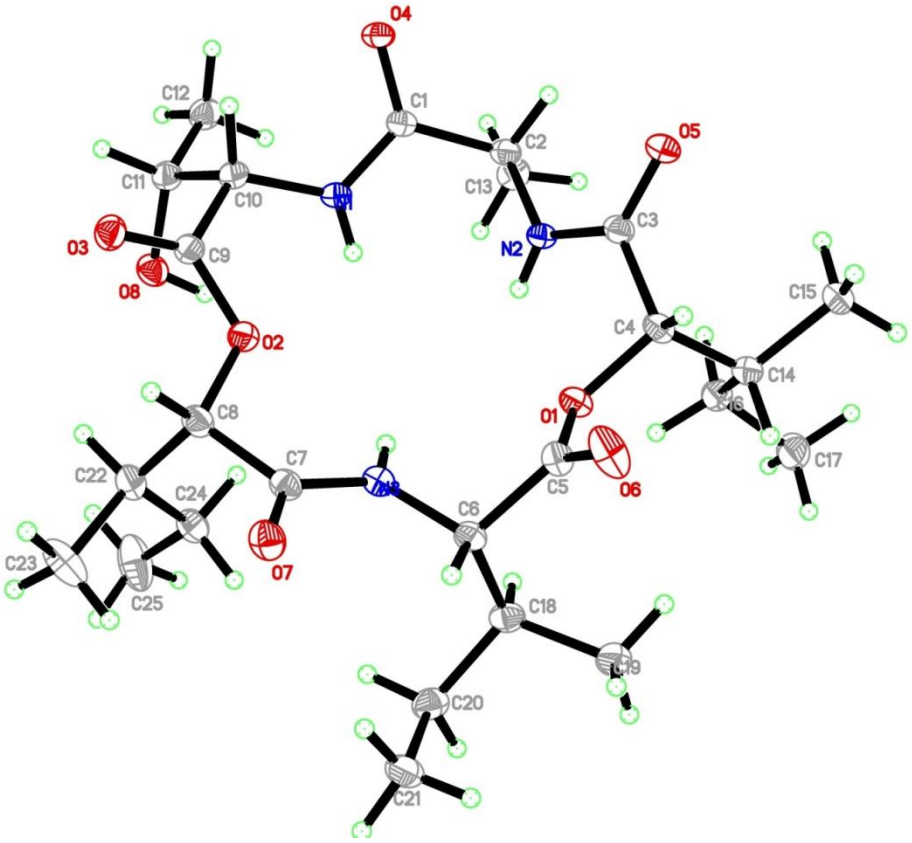



**Figure S6**. ORTEP representation of the asymmetric unit of the crystal of **3**. Ellipsoids displayed at 50% probability.

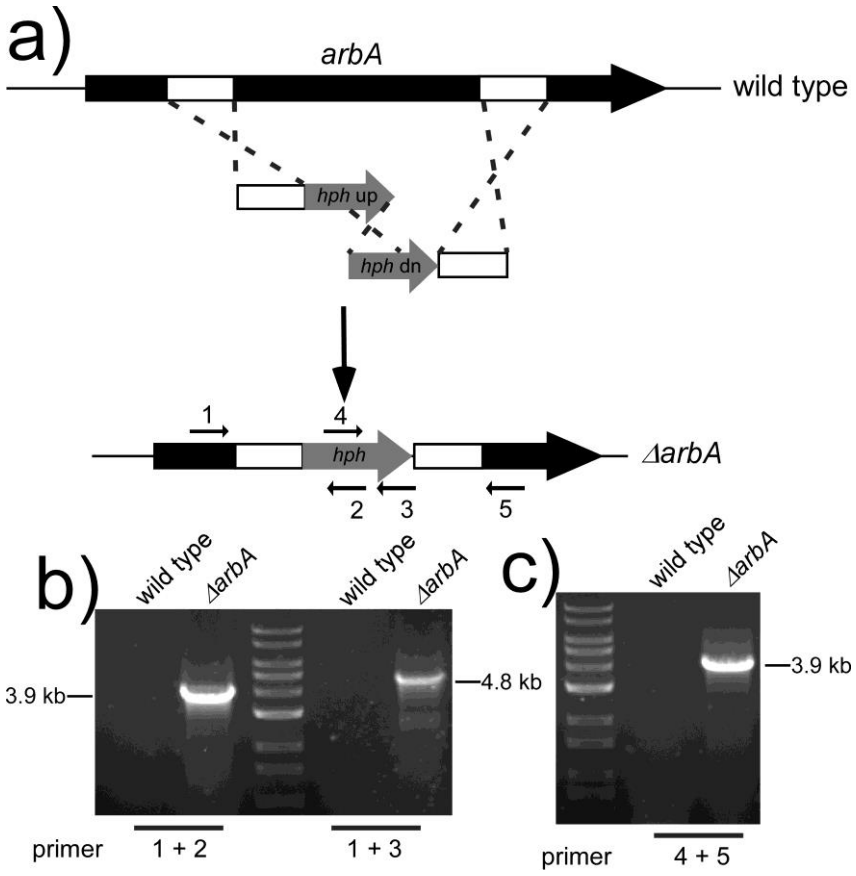



**Figure S7.** Deletion of *arbA* in *C. arbuscula*.

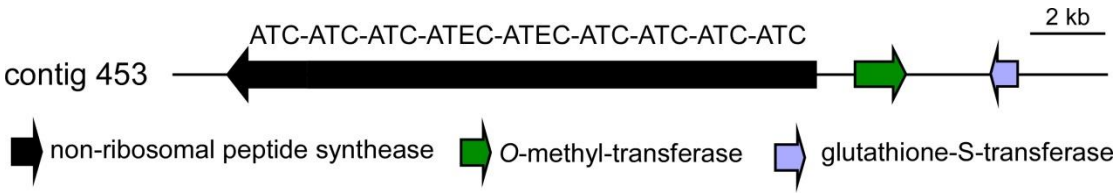



**Figure S8**. Gene cluster of contig 453 for the possible biosynthesis of arbumelin (**4**).

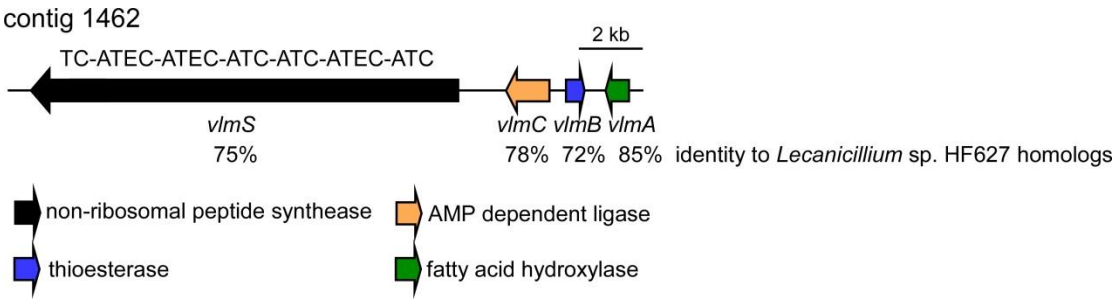



**Figure S9**. Gene cluster of contig 1462 for the biosynthesis of verlamelin (**5**) in *C. arbuscula*. The protein sequence identity to *Lecanicillium* sp. HF627 homologs responsible for verlamelin biosynthesis are also shown.

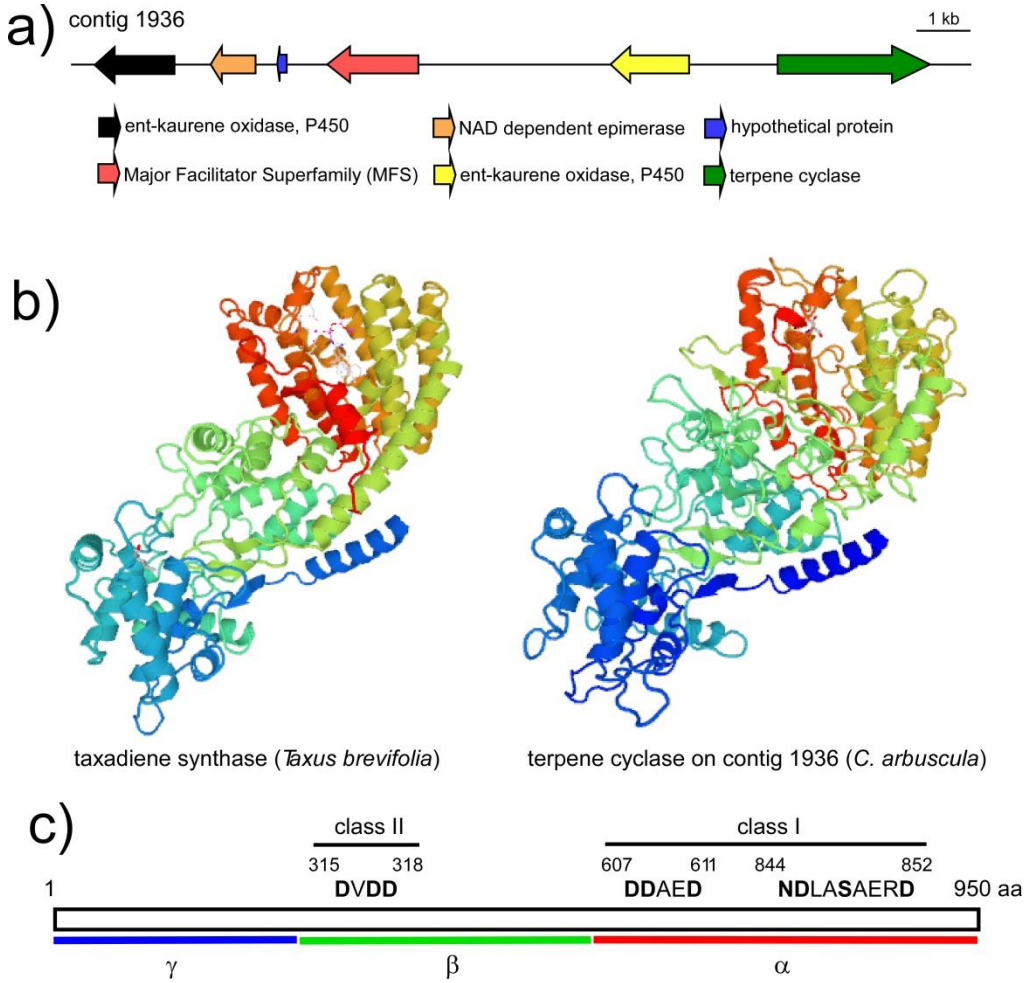



**Figure S10**. a) Gene cluster of contig 1936 for the possible biosynthesis of three diterpenes (**9**, **10**, **11**). b) Structural modeling of taxadiene synthase (*Taxus brevifolia*) and the predicted terpene cylcase (*C. arbuscula*) on SWISS-MODEL server. The structure of terpene cylcase (*C. arbuscula*) is modeled based on taxadiene synthase (*Taxus brevifolia*). c) Domain organization of the predicted terpene cylcase (*C. arbuscula*) based on the structural modeling. Three putative α, β, γ domains are shown, and the conserved motifs from class I and class II terpene cyclases are labeled together with the amino acid positions. The highly conserved residues are bold.

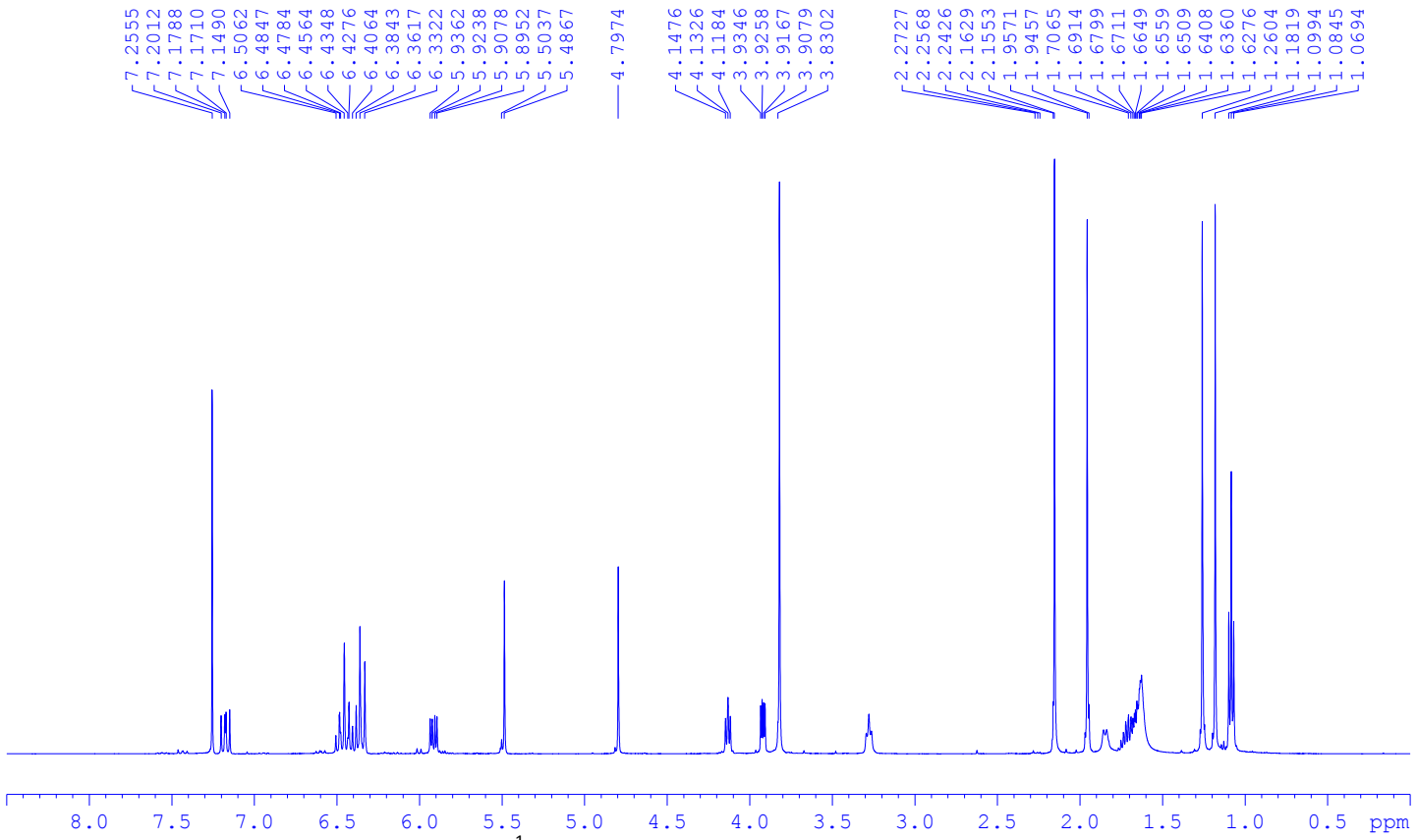



**Figure S11**: $^{1}$H NMR of aurovertin B (**1**) in $CDCl_3$ (500 MHz)

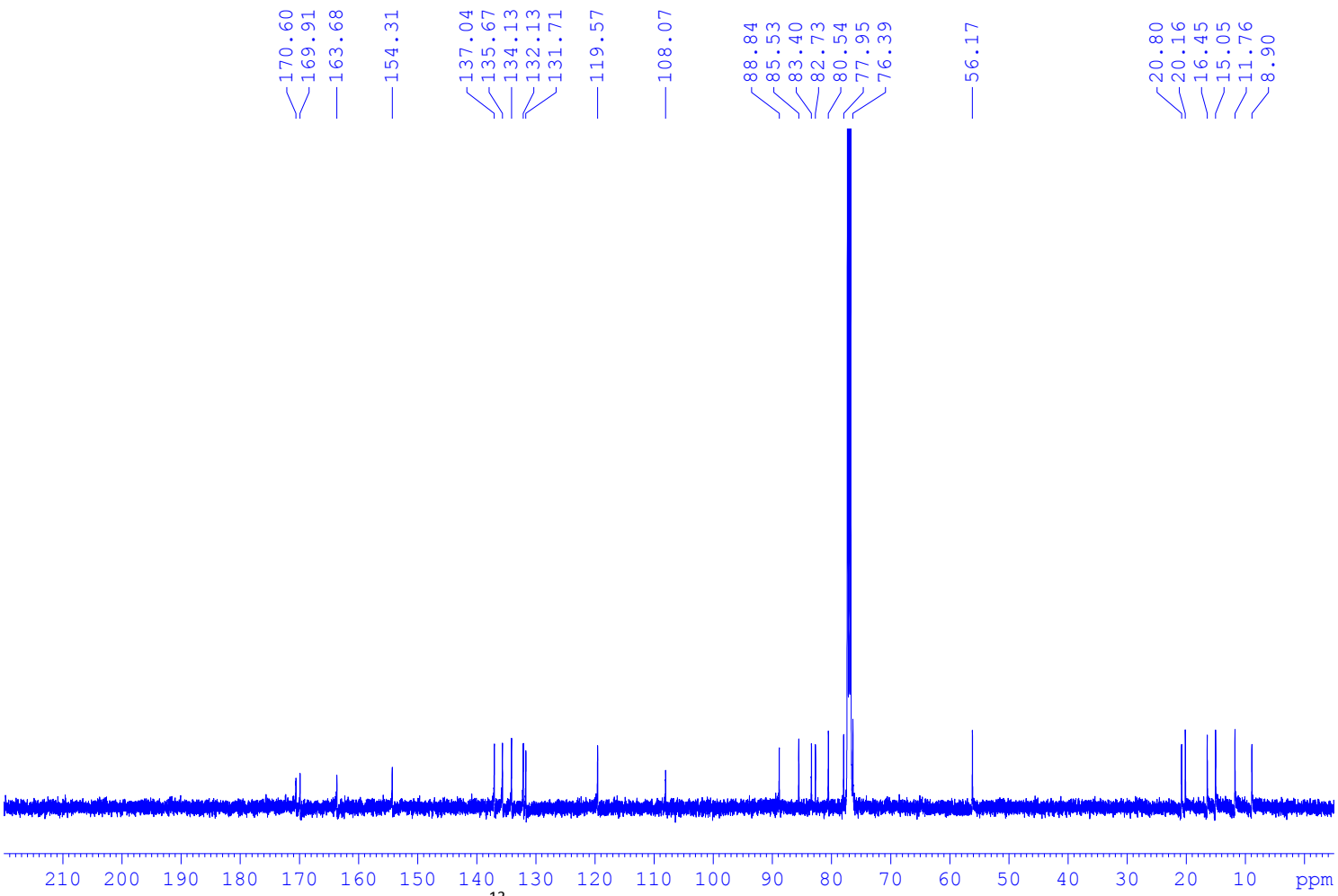



**Figure S12**: $^{13}C$ NMR of aurovertin B (**1**) in $CDCl_3$ (125 MHz)

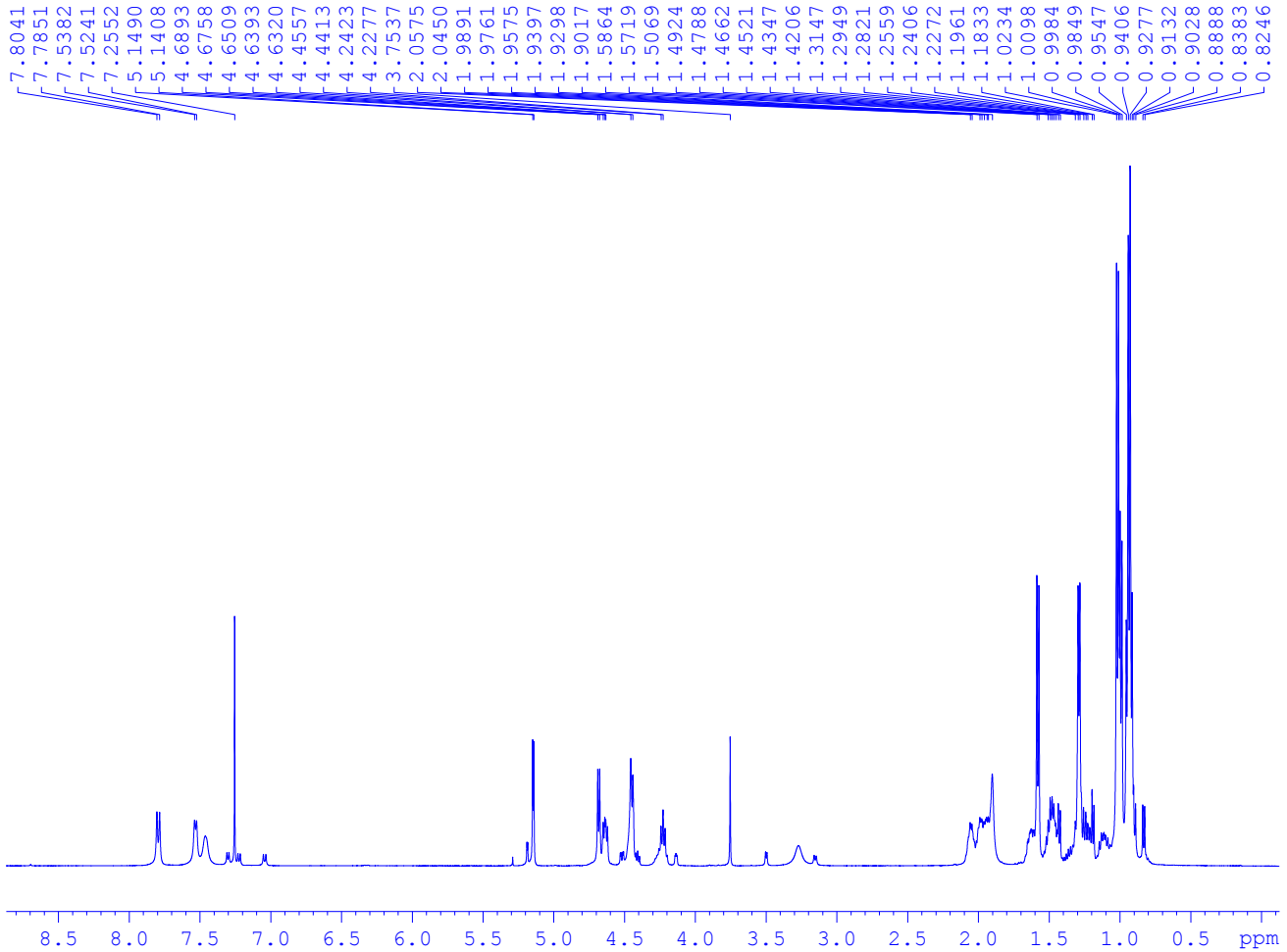



**Figure S13**: $^{1}$H NMR of arbumycin (**3**) in $CDCl_3$ (500 MHz)

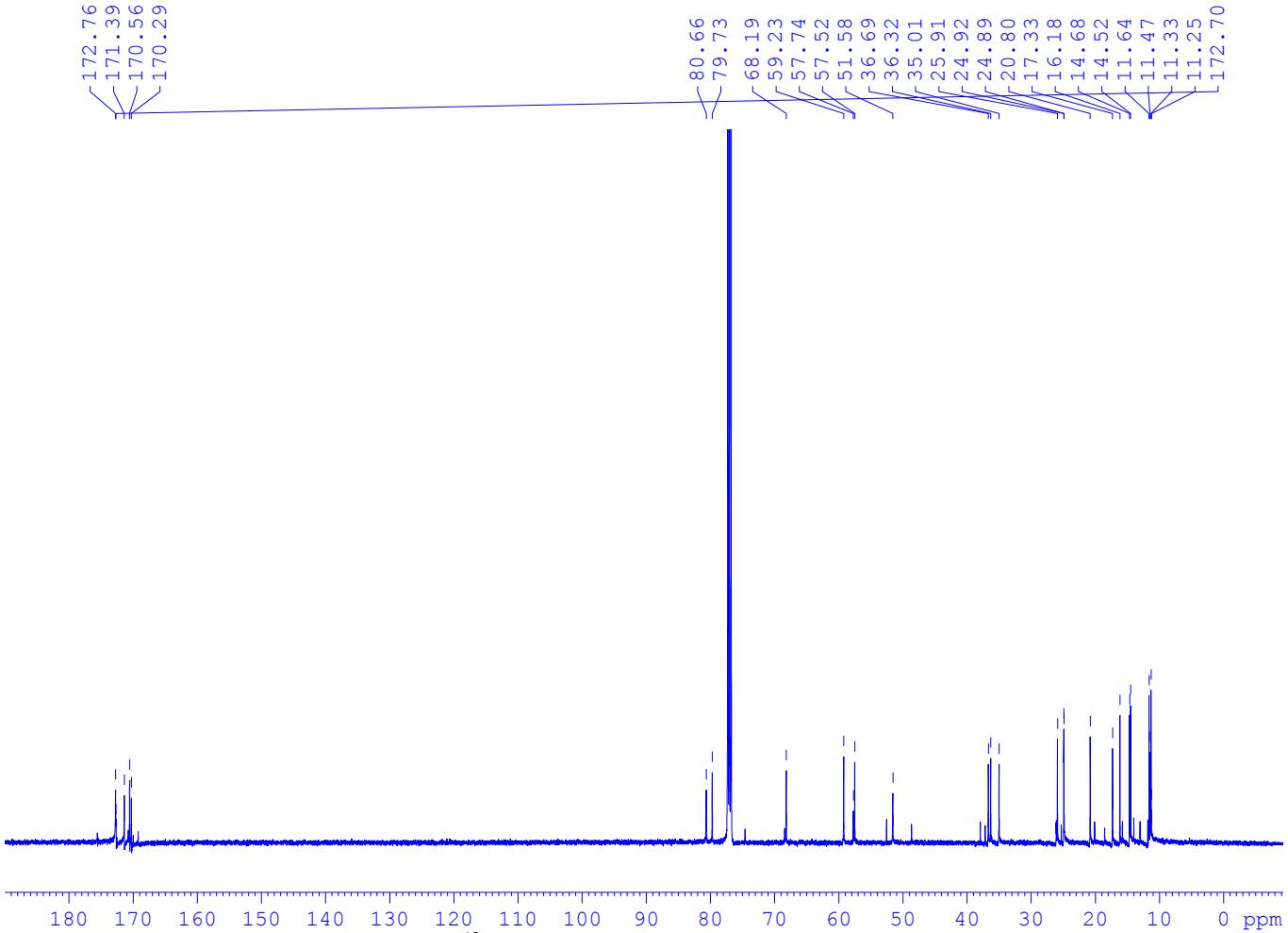



**Figure S14**: $^{13}C$ NMR of arbumycin (**3**) in $CDCl_3$ (125 MHz)

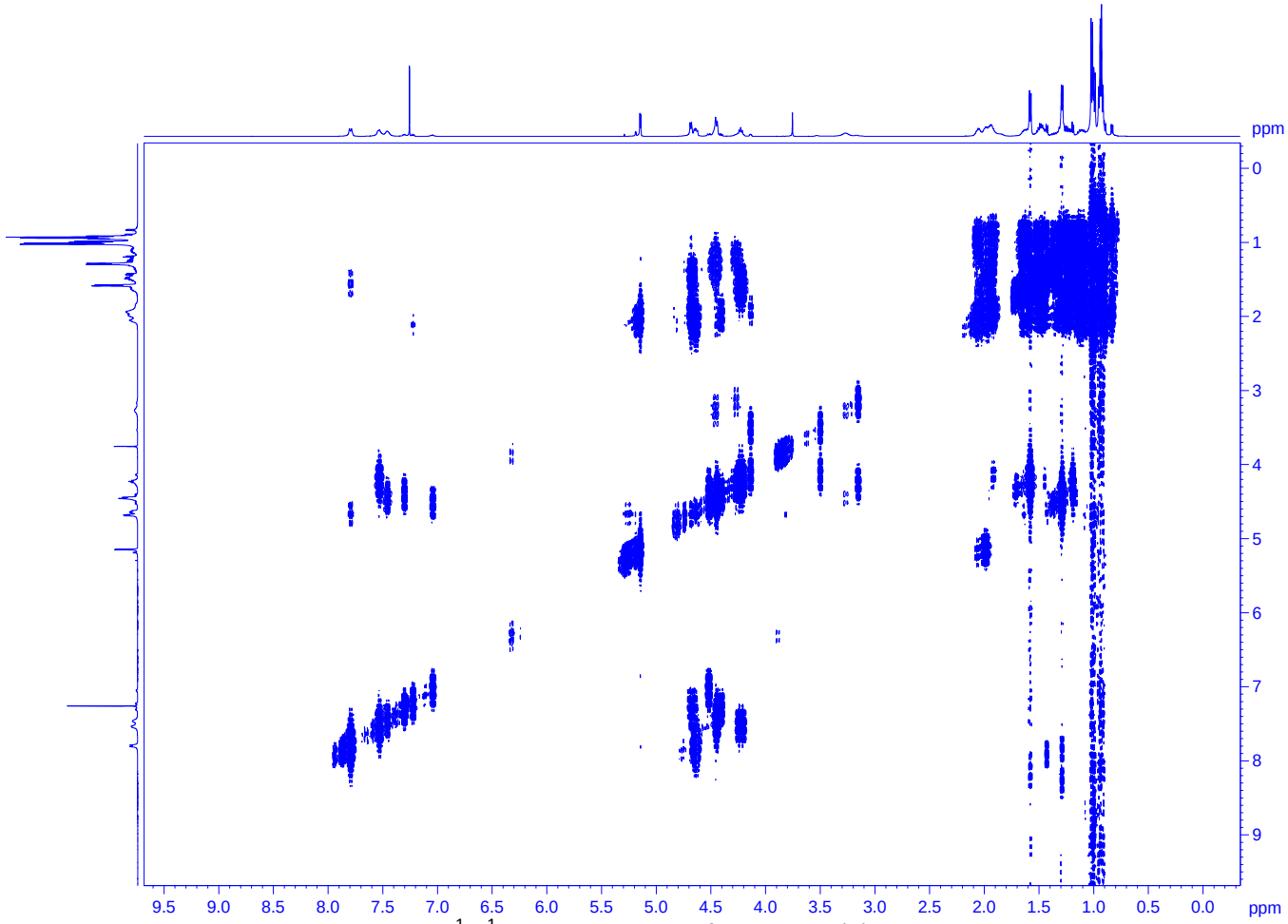



**Figure S15**: $^{1}H$ $^{1}H$-COSY spectrum of arbumycin (**3**) in $CDCl_3$

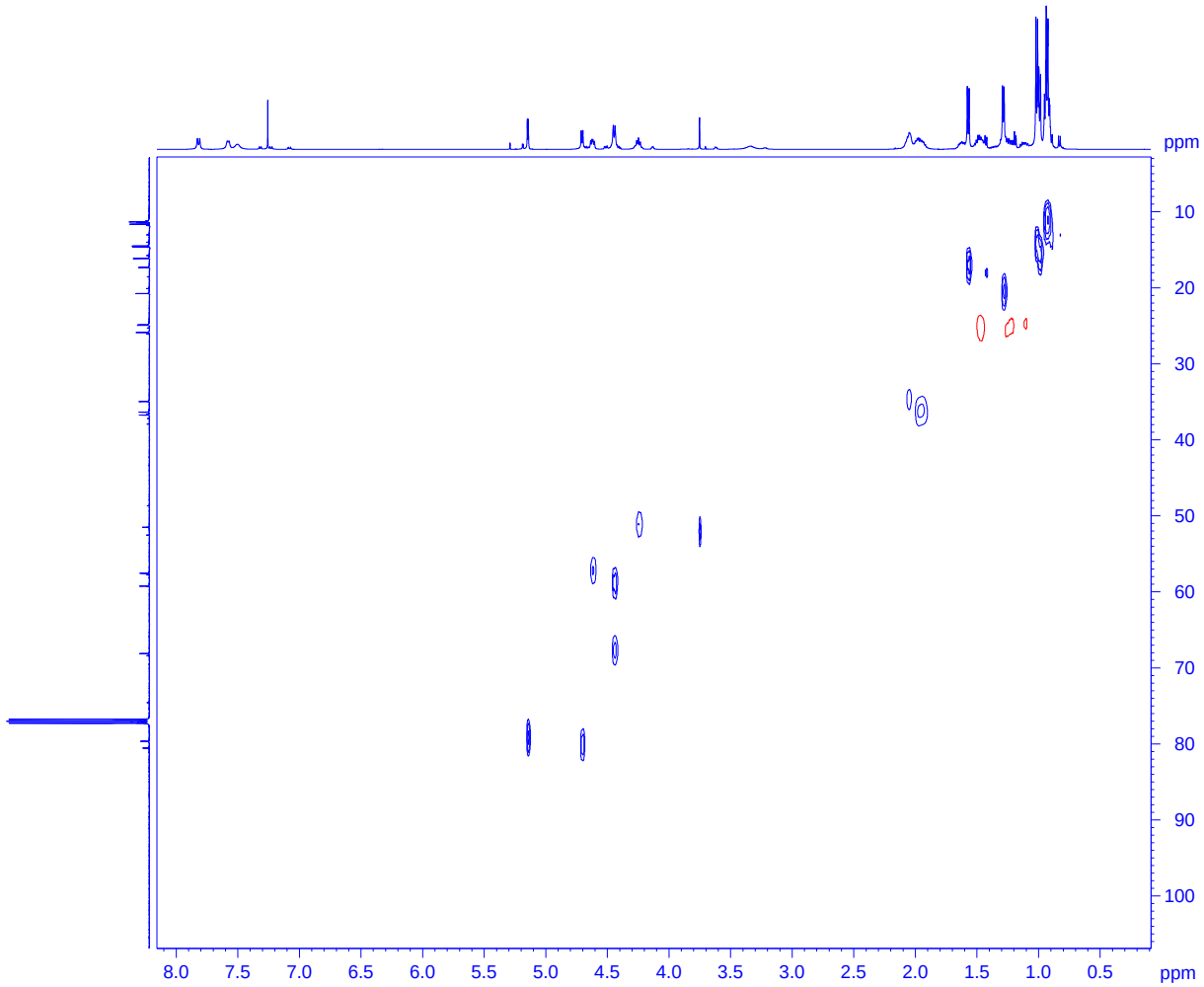



**Figure S16**: HSQC of arbumycin (**3**) in $CDCl_3$

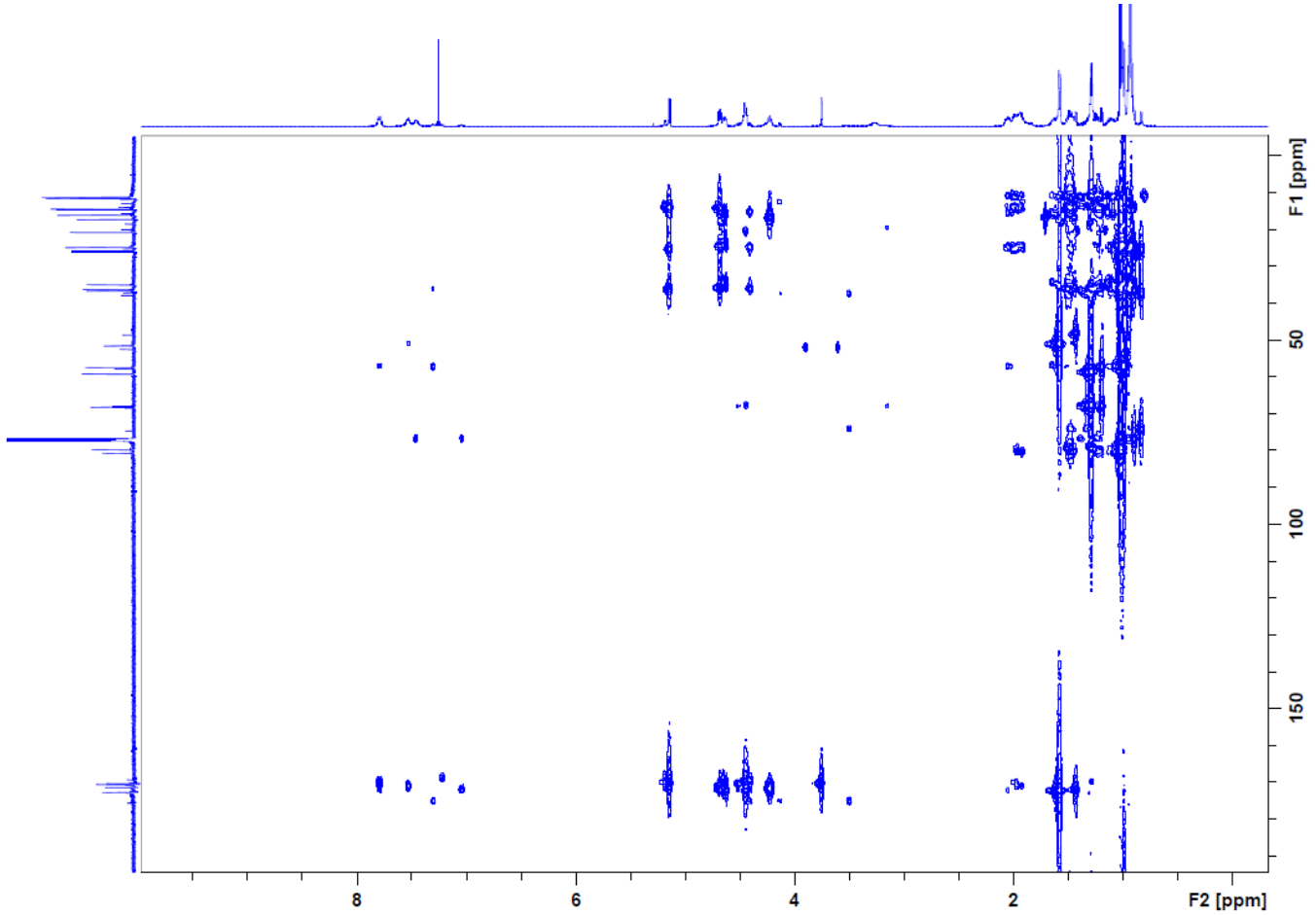



**Figure S17**: HMBC of arbumycin (**3**) in $CDCl_3$

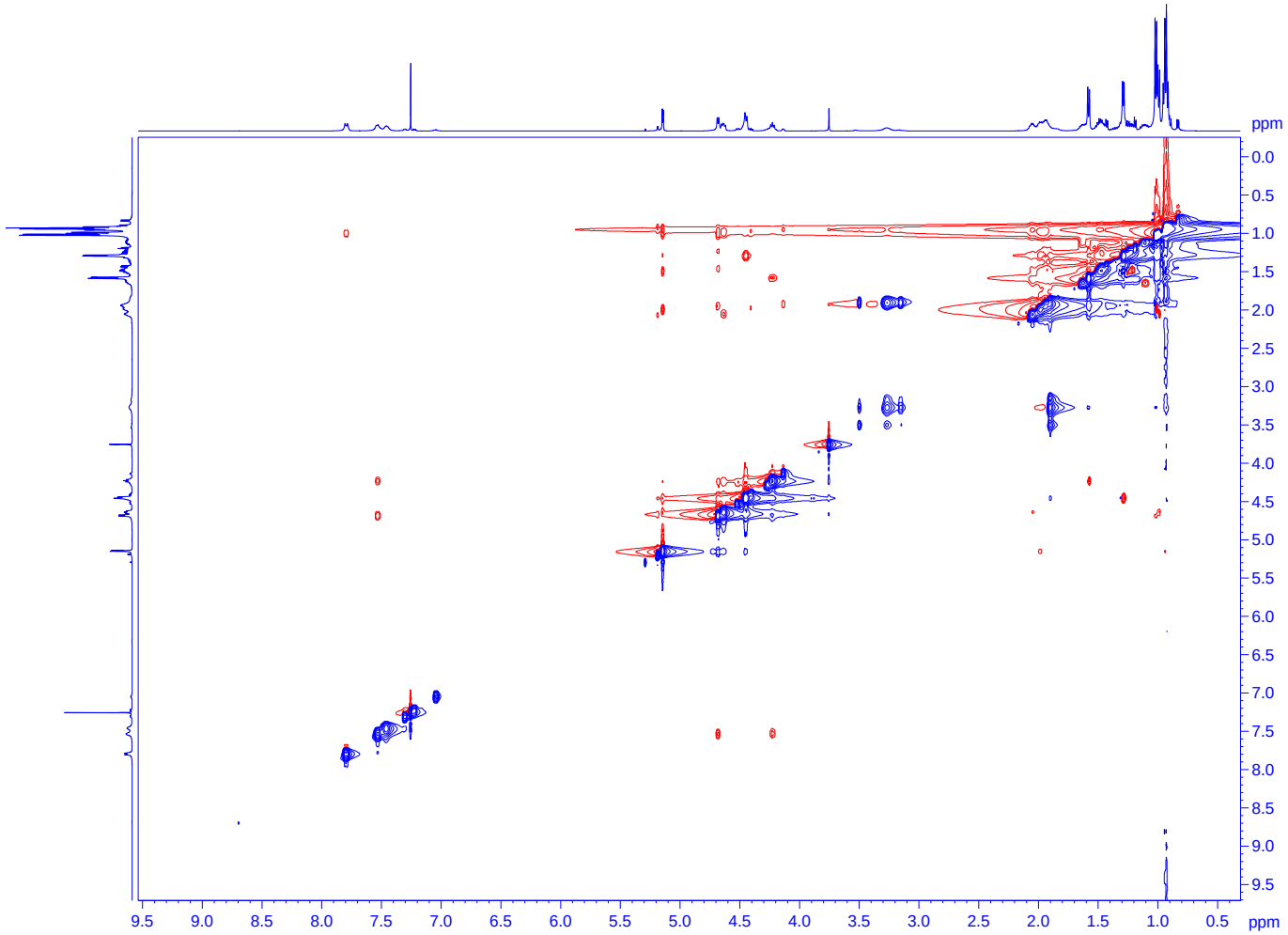



**Figure S18**: NOSEY of arbumycin (**3**) in $CDCl_3$

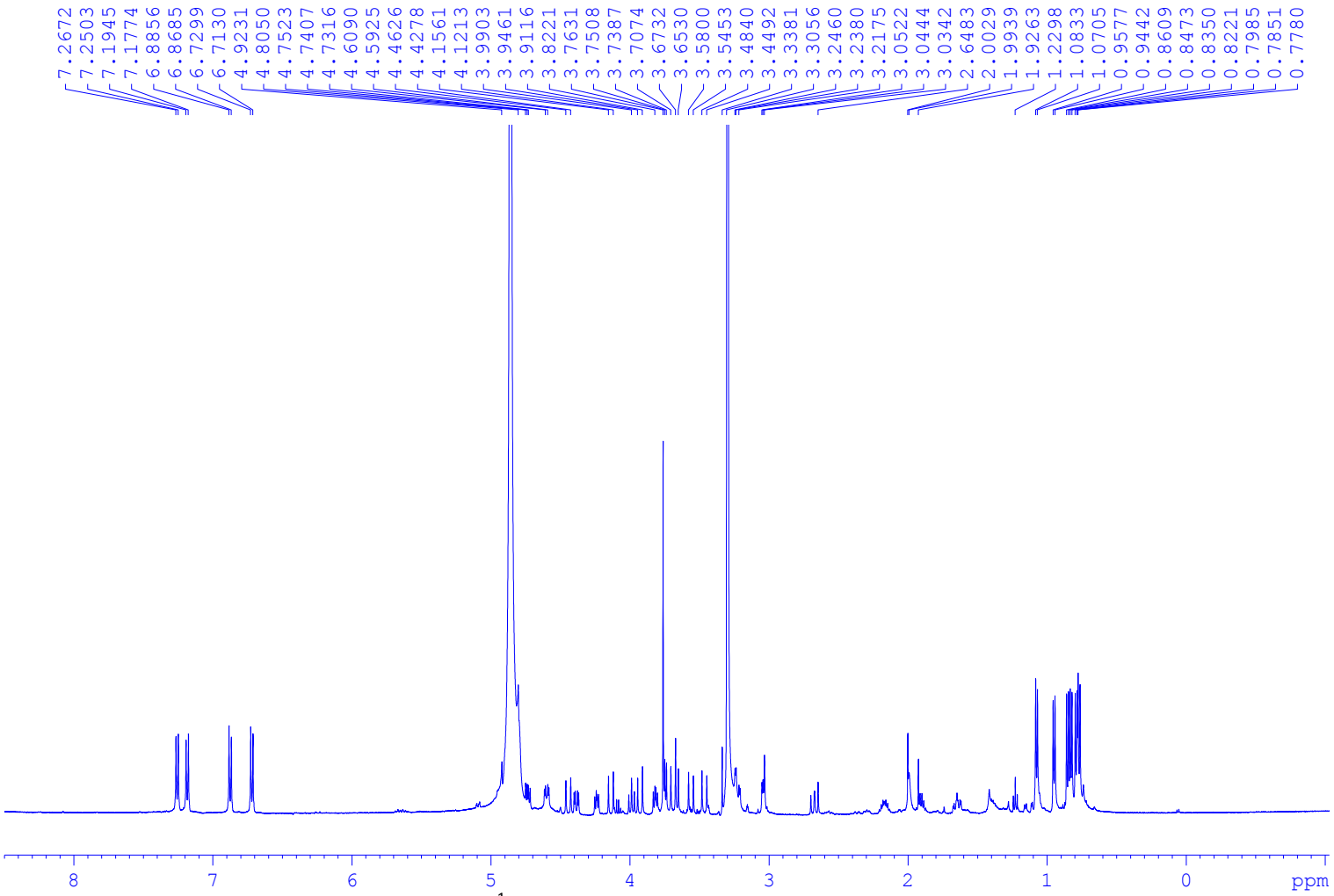



**Figure S19**: $^{1}$H NMR of arbumelin (**4**) in $CD_3OD$ (500 MHz)

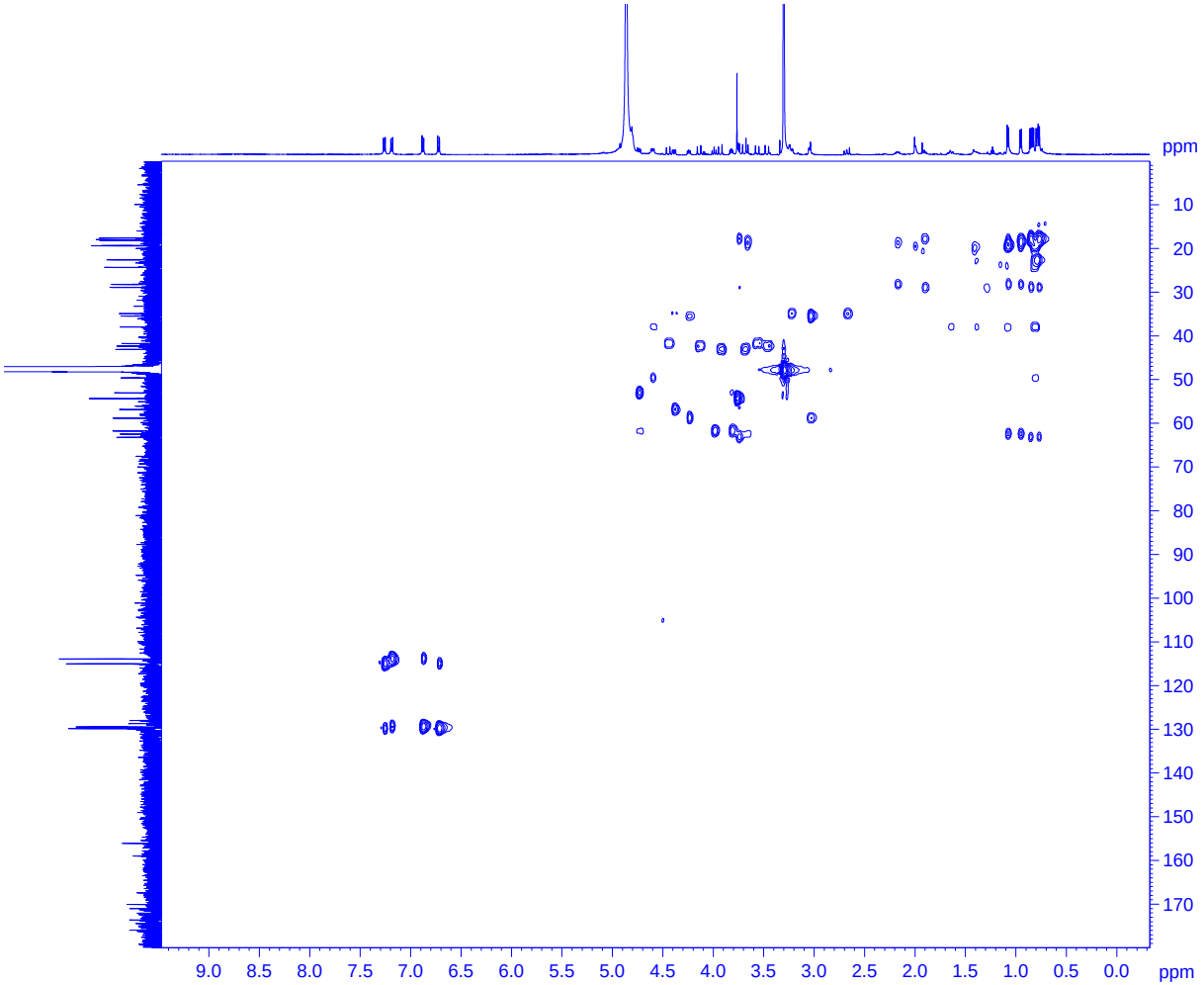



**Figure S20**: HSQC-TOCSY of arbumelin (**4**) in $CD_3OD$ (500 MHz)

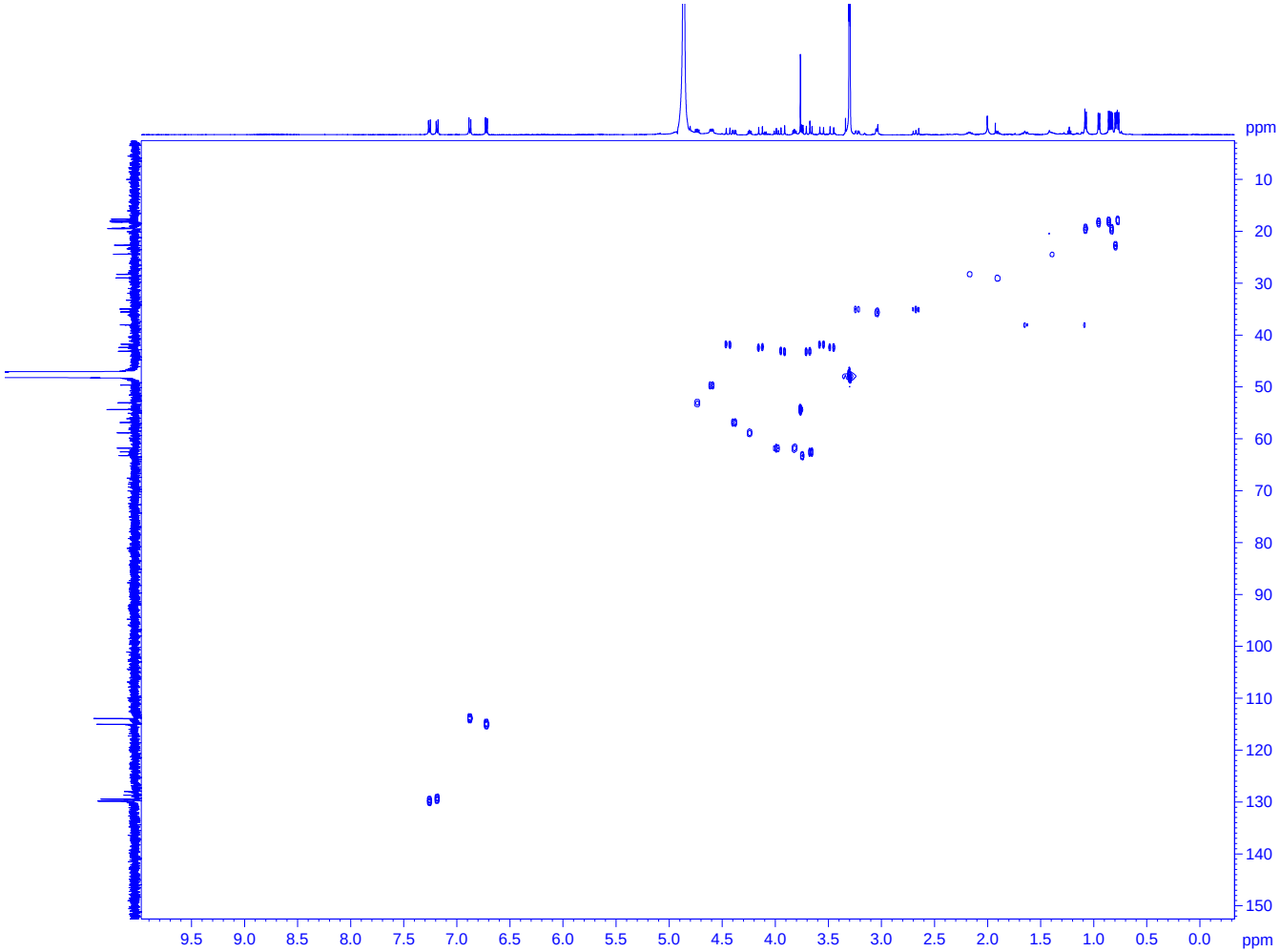



**Figure S21**: HSQC of arbumelin (**4**) in $CD_3OD$

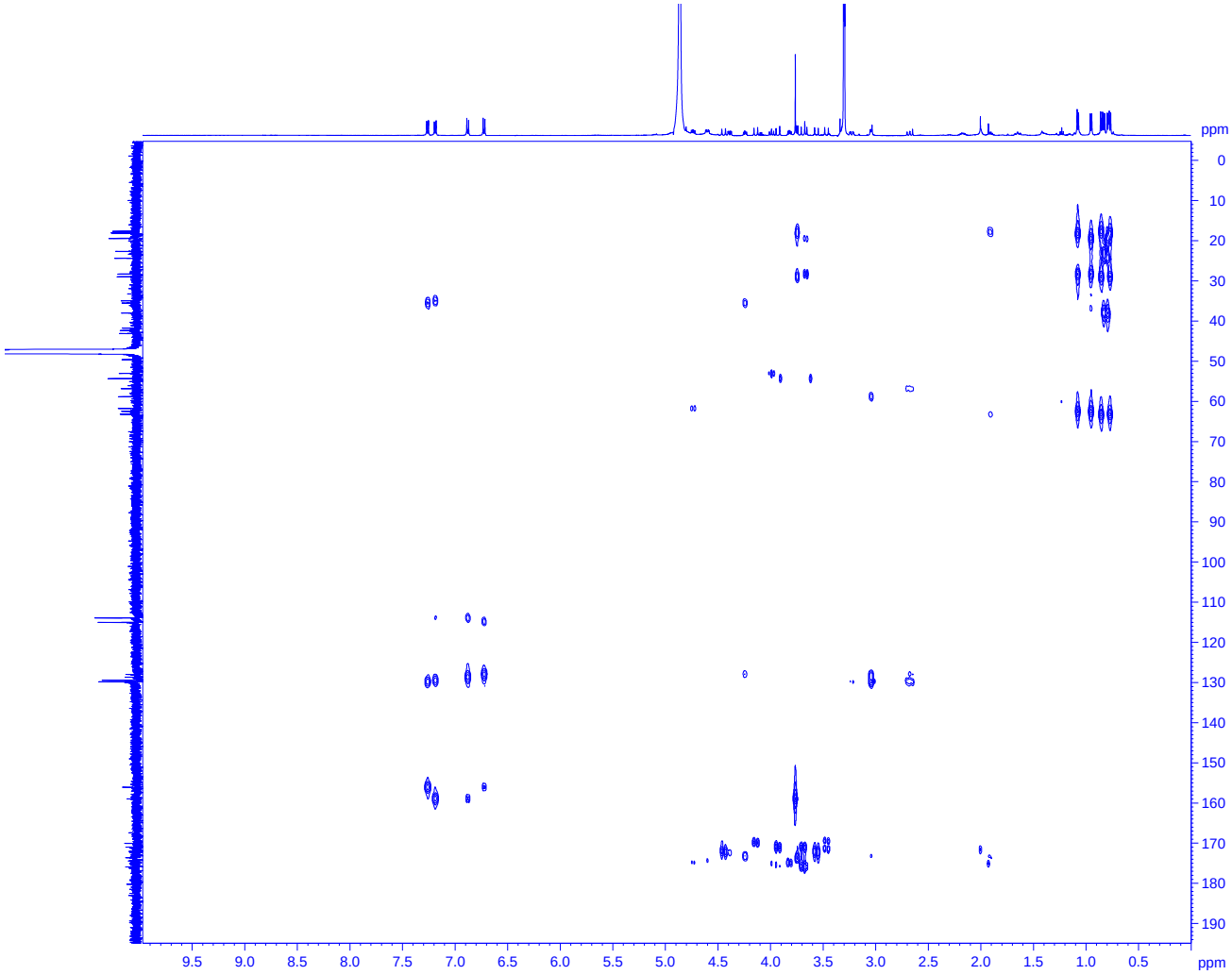



**Figure S22**: HMBC of arbumelin (**4**) in $CD_3OD$

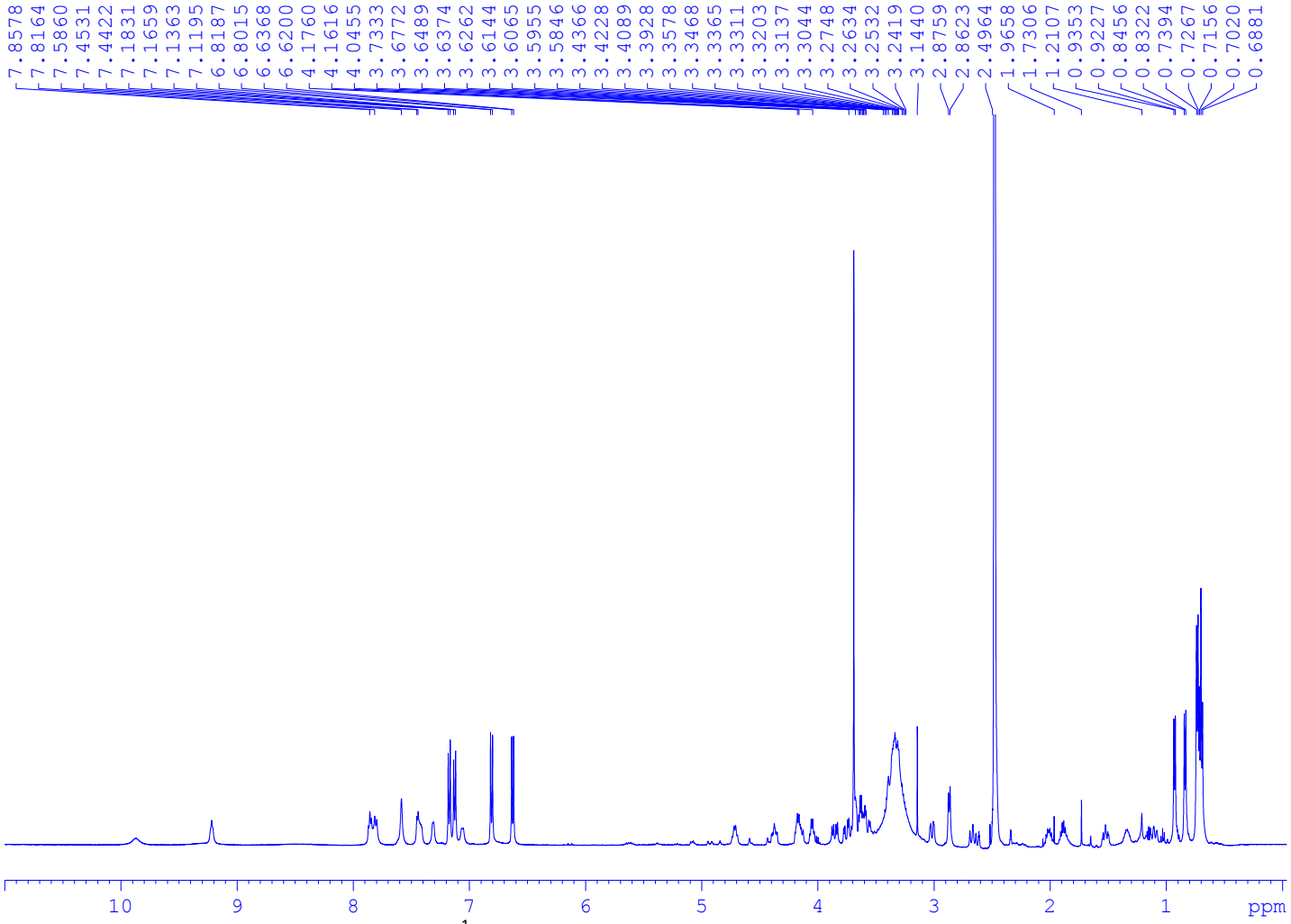



**Figure S23**: $^{1}$H NMR of arbumelin (**4**) in DMSO-$d_6$ (500 MHz)

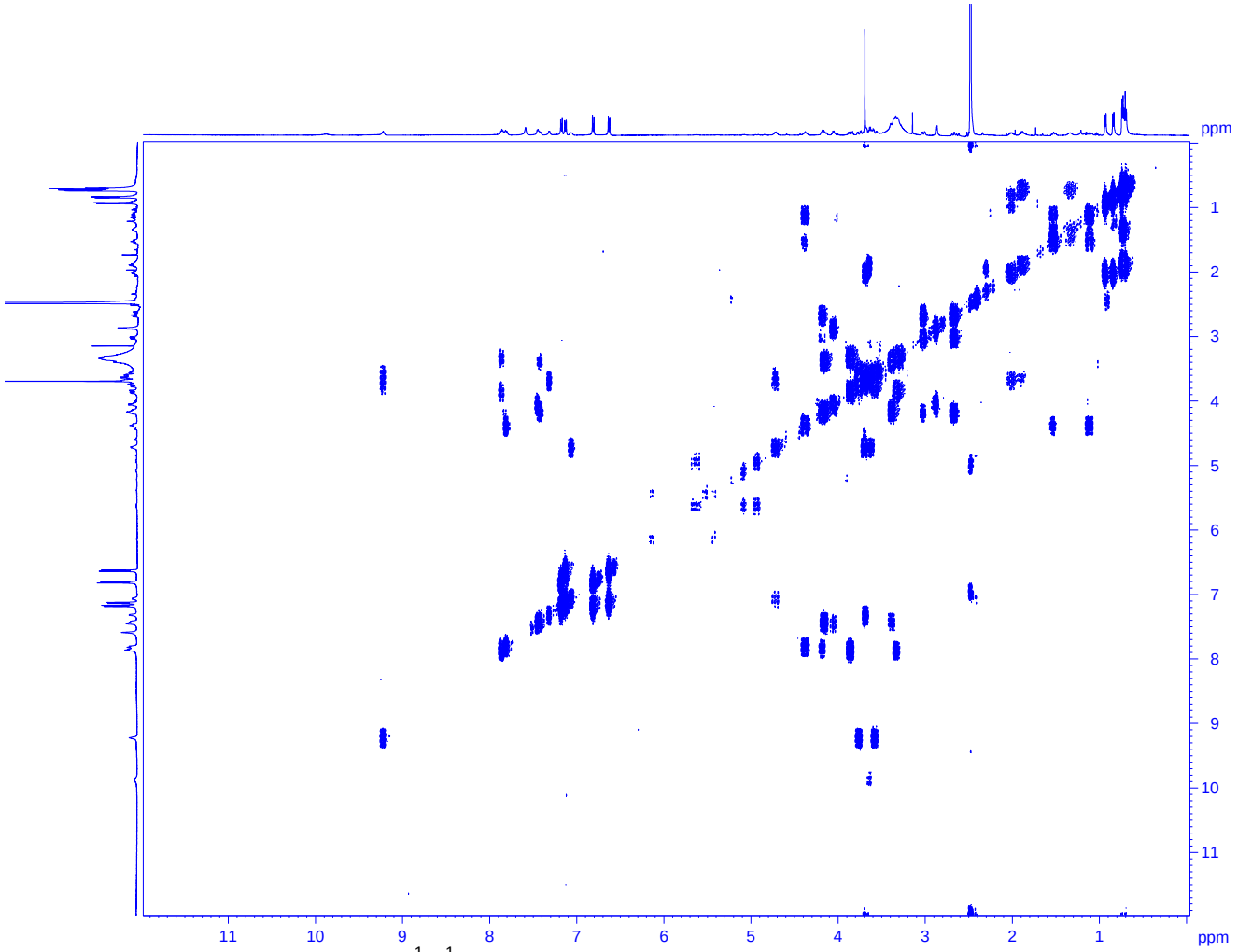



**Figure S24**: $^{1}$H $^{1}$H-COSY spectrum of arbumelin (**4**) in DMSO-$d_6$

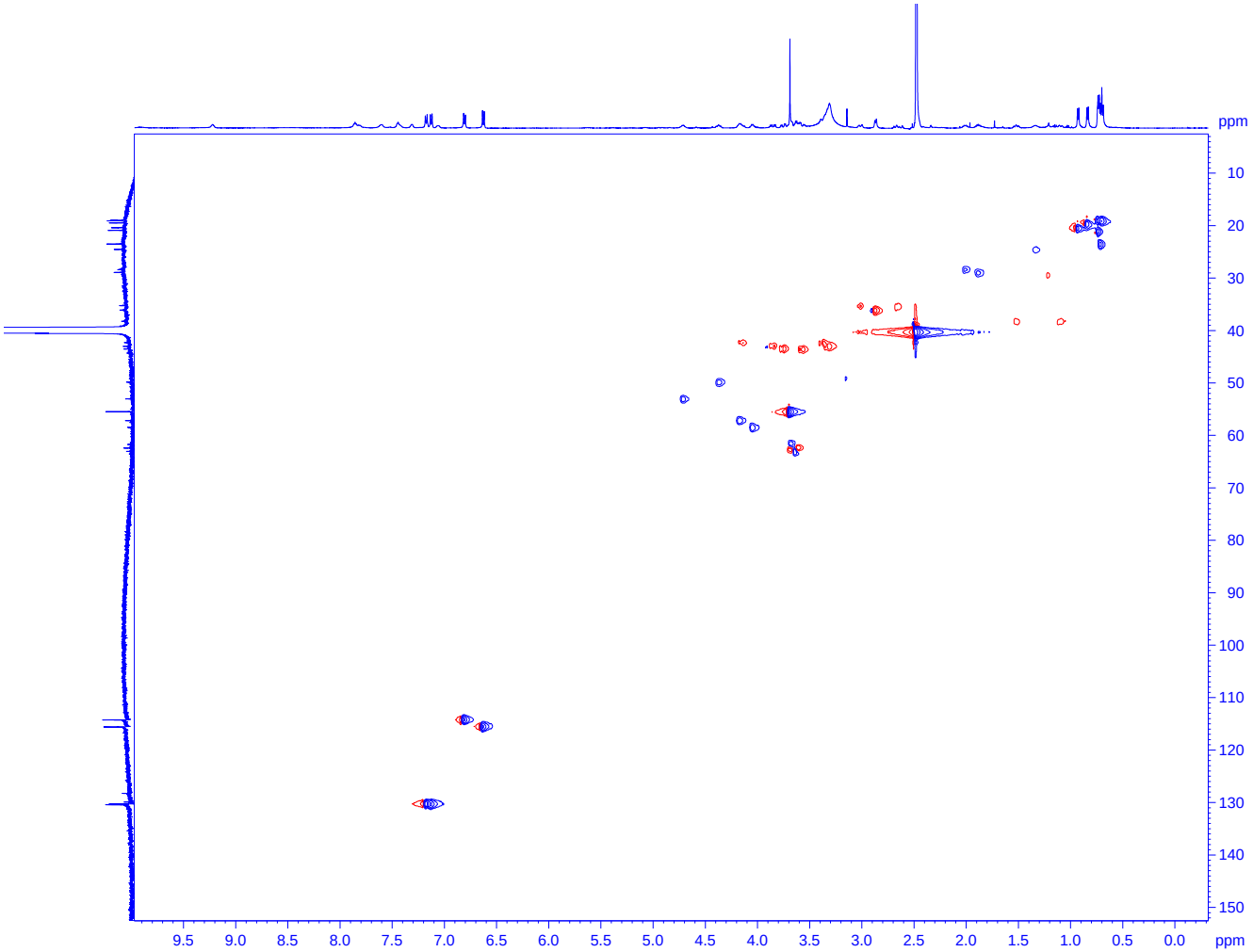



**Figure S25**: HSQC of arbumelin (**4**) in DMSO-$d_6$

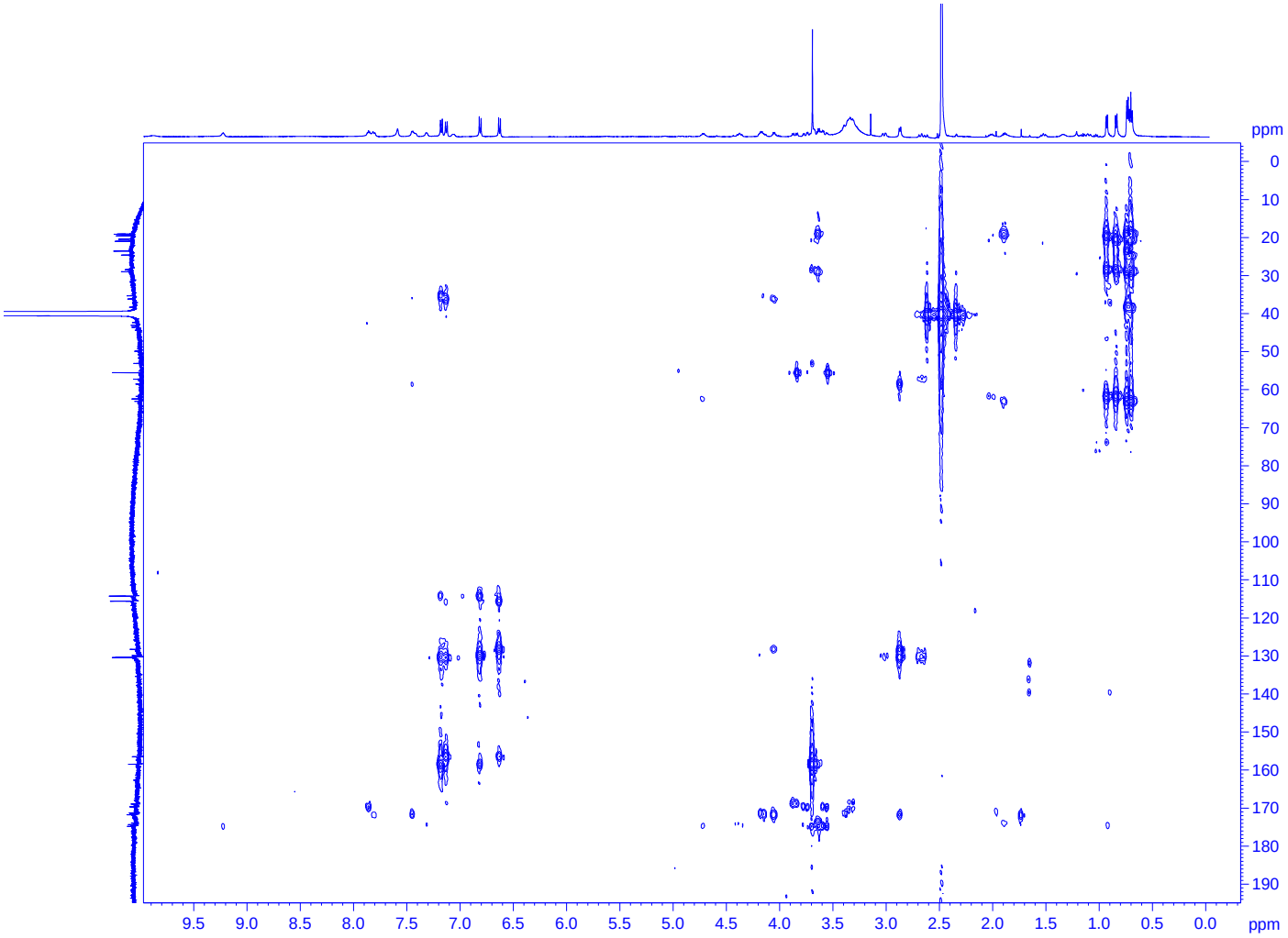



**Figure S26**: HMBC of arbumelin (**4**) in DMSO-$d_6$

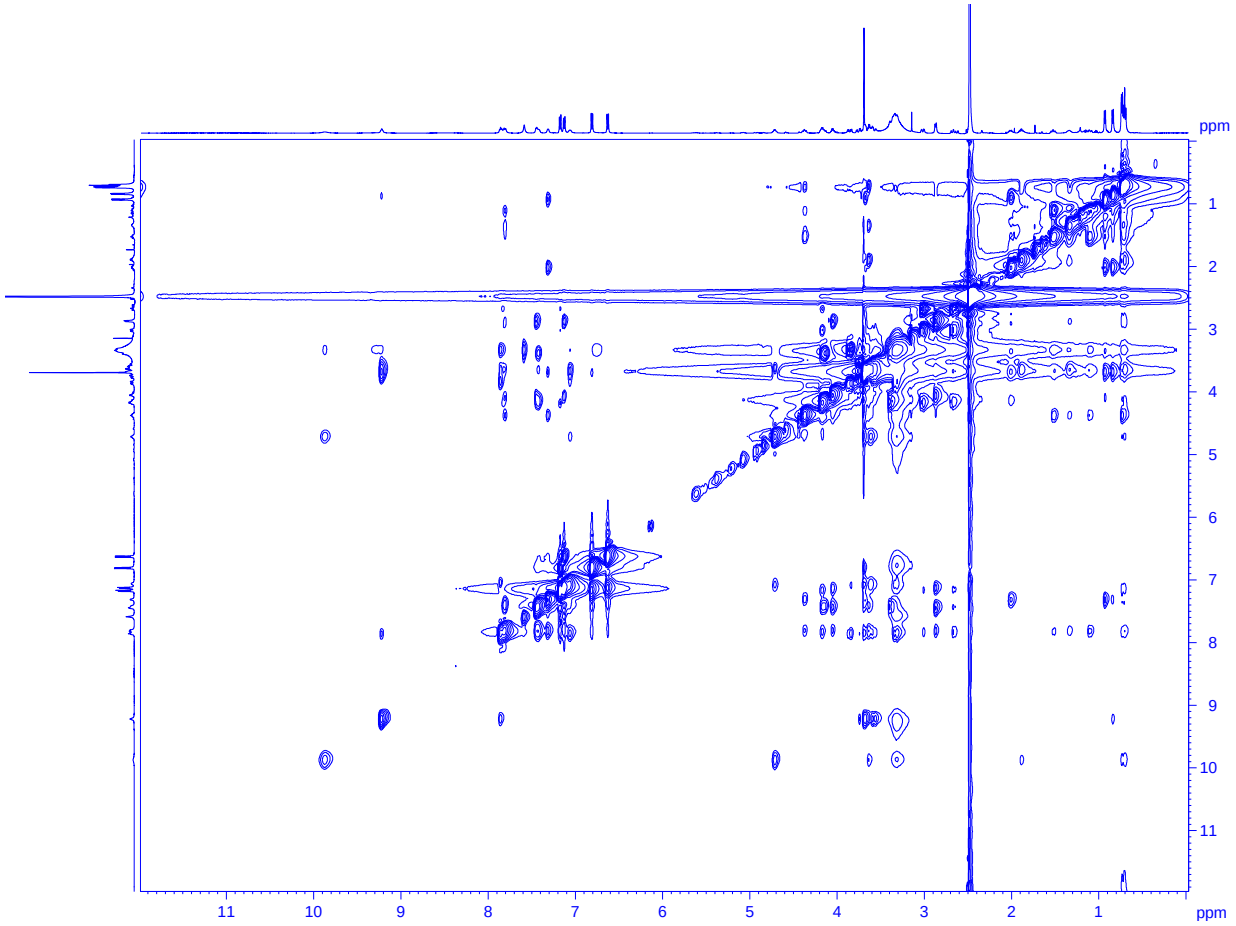



**Figure S27**: NOSEY of arbumelin (**4**) in DMSO-$d_6$

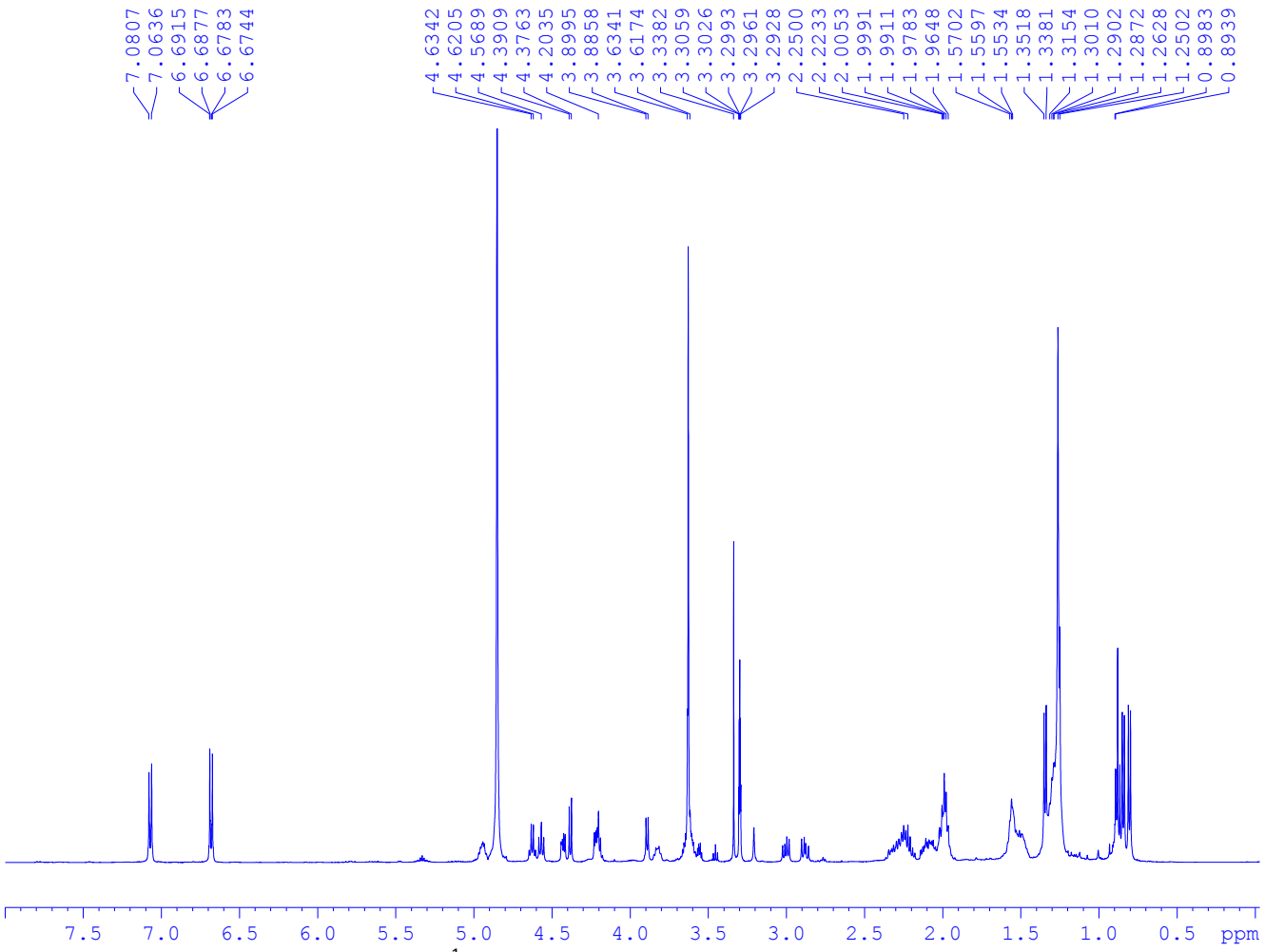



**Figure S28**: $^{1}H$ NMR of verlamelin (**5**) in $CD_3OD$ (500MHz)

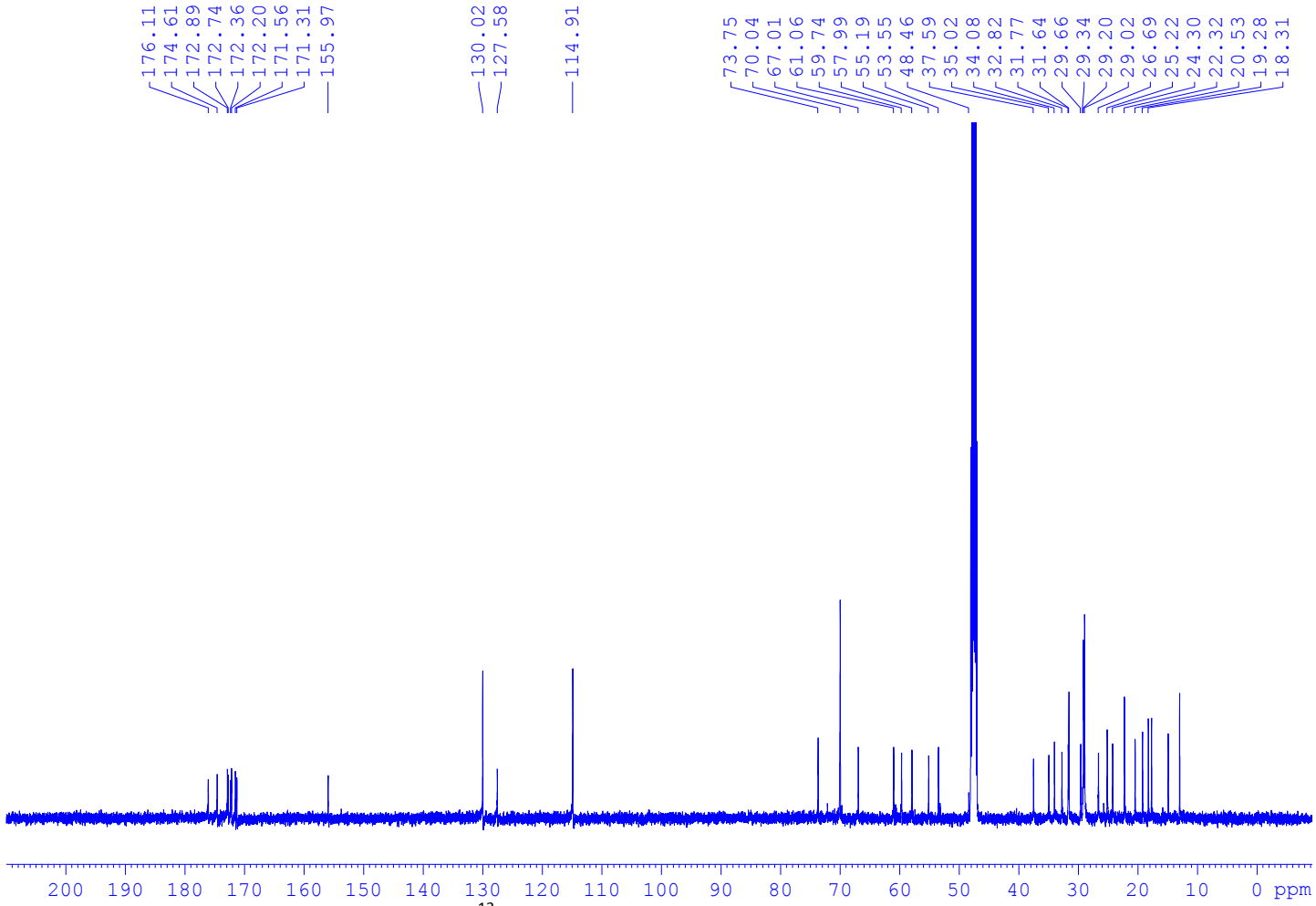



**Figure S29**: $^{13}C$ NMR of verlamelin (**5**) in $CD_3OD$ (125 MHz)

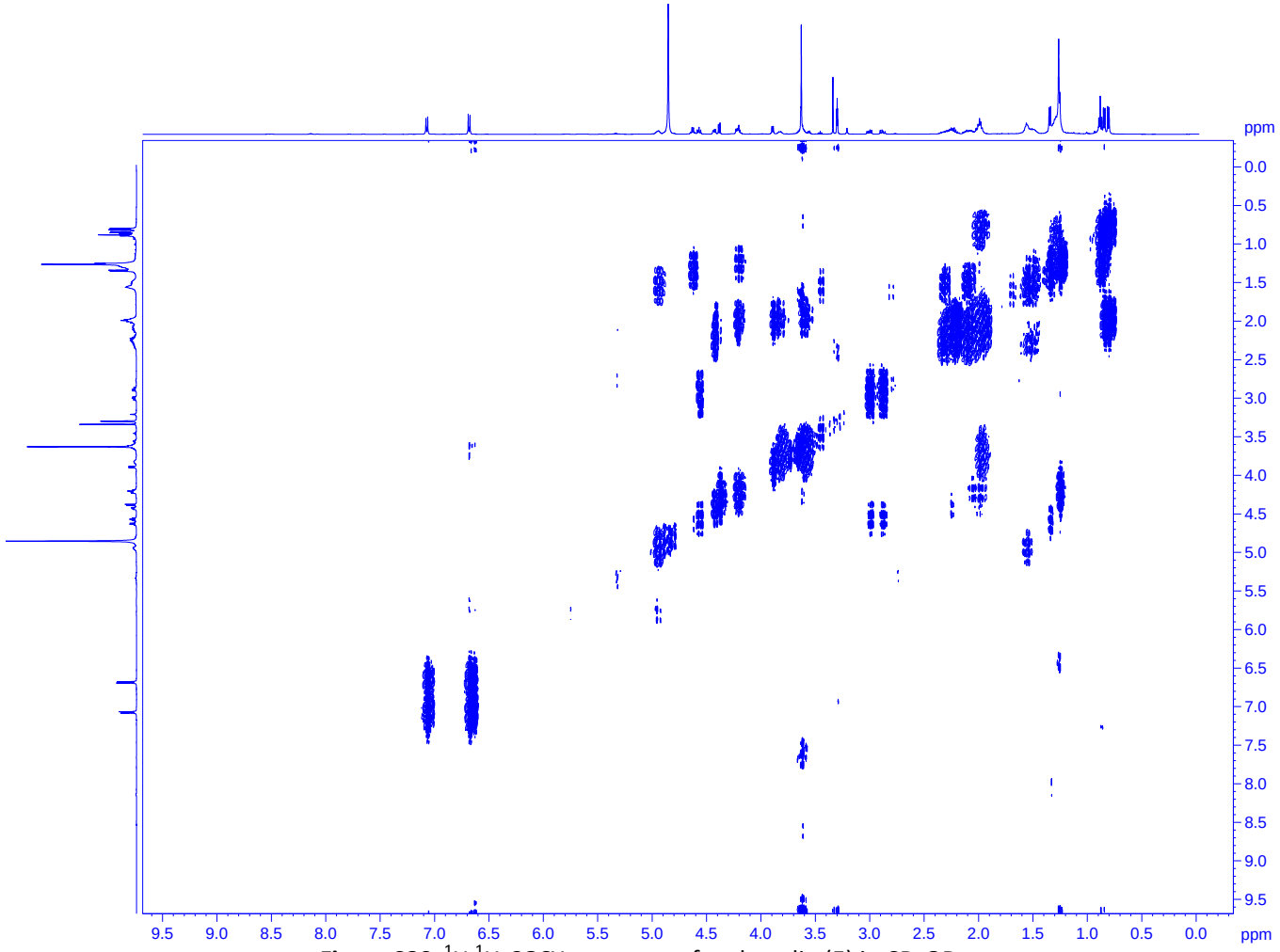



**Figure S30**: $^{1}$H $^{1}$H-COSY spectrum of verlamelin (**5**) in $CD_3OD$

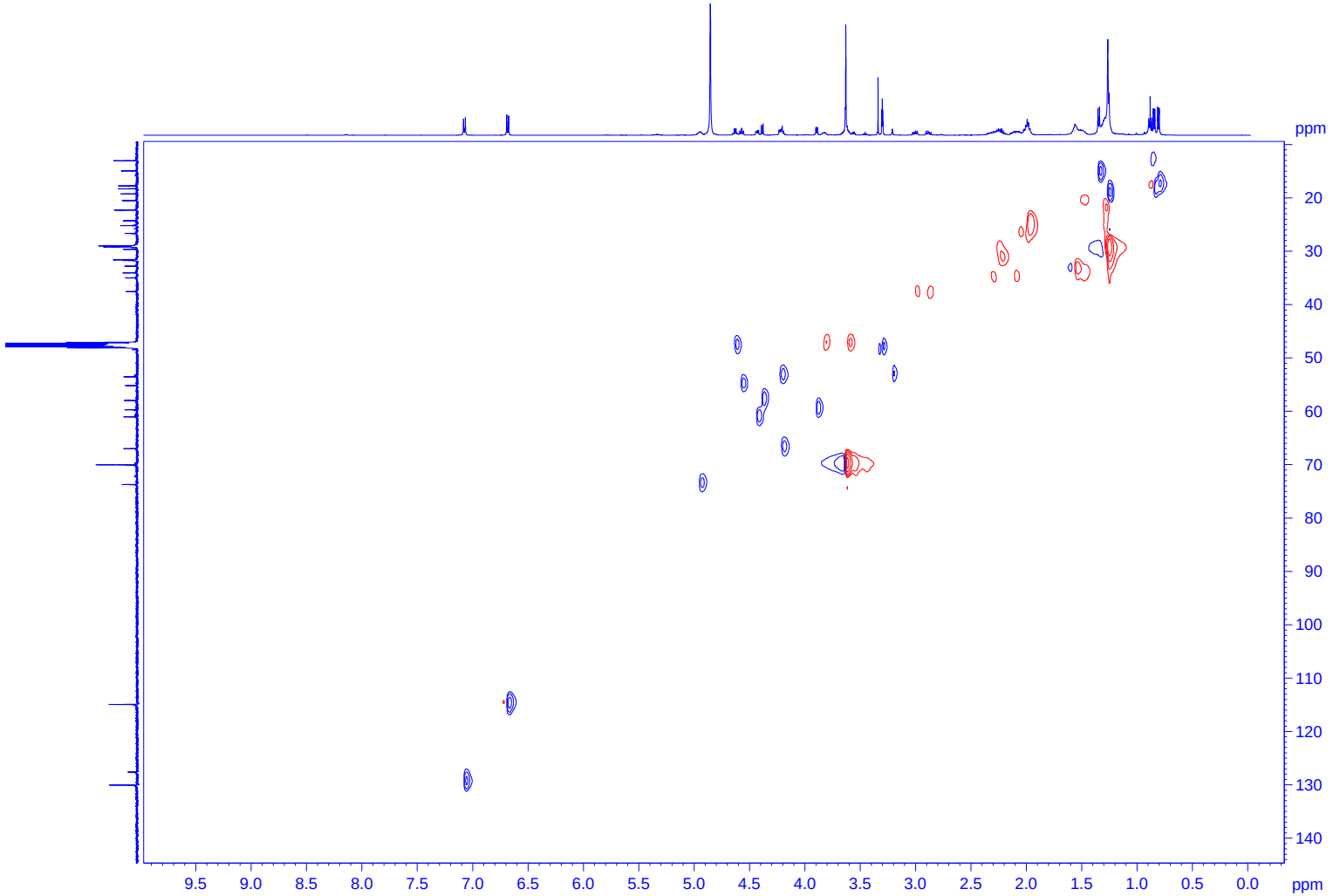



**Figure S31**: HSQC of verlamelin (**5**) in $CD_3OD$

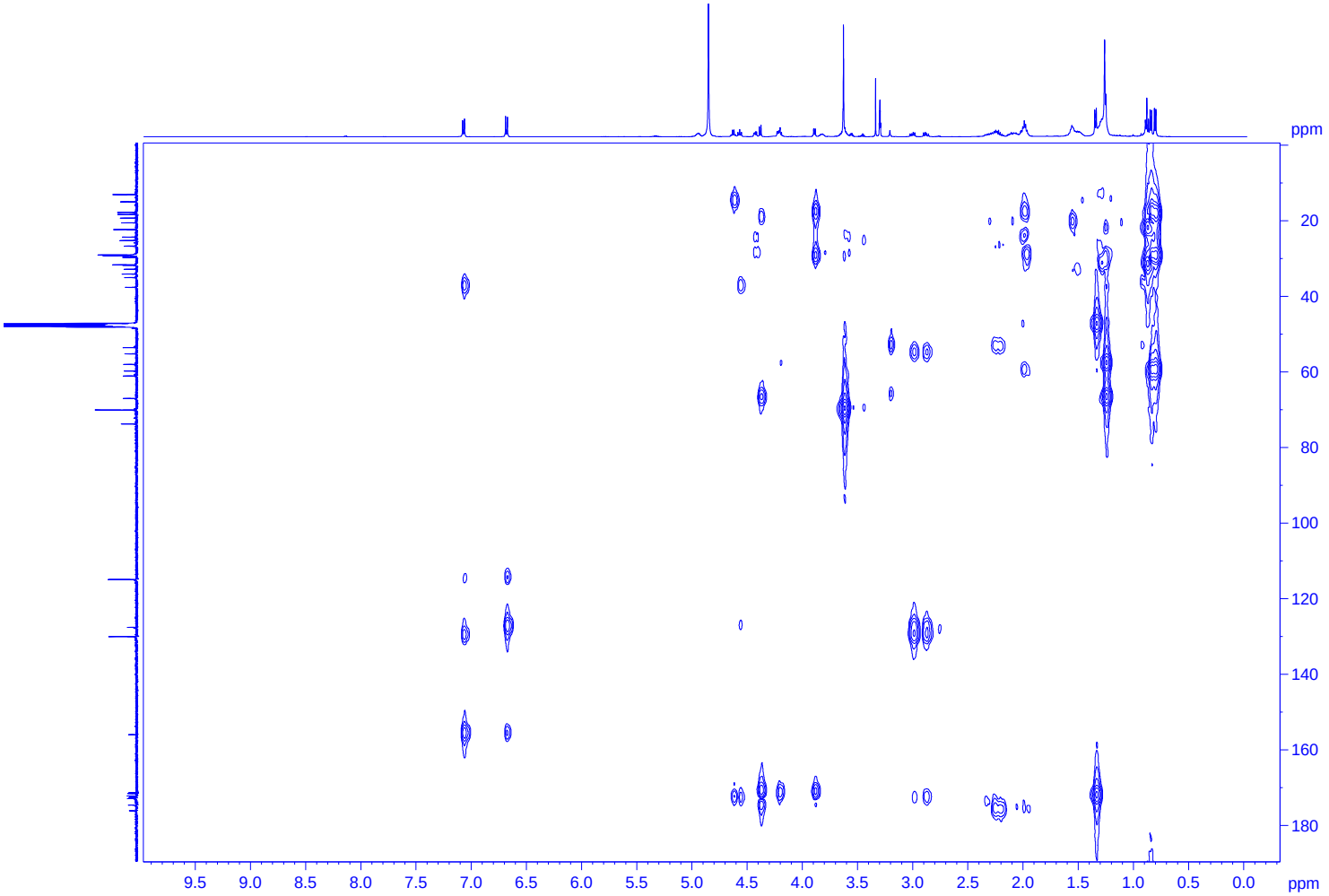



**Figure S32**: HMBC of verlamelin (**5**) in $CD_3OD$

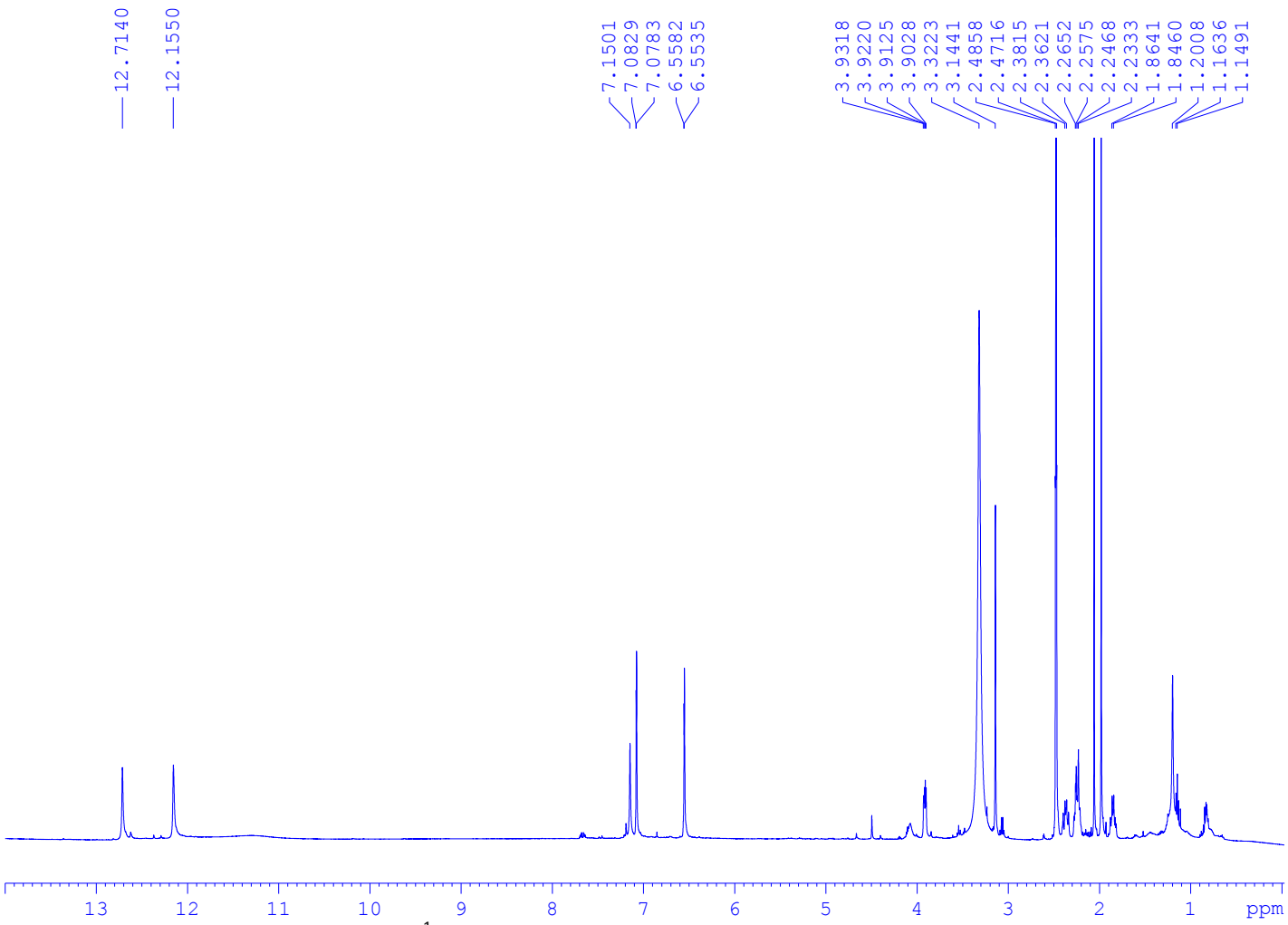



**Figure S33**: $^{1}$H NMR of paeciloquinone B (**8**) in DMSO-$d_6$ (500MHz)

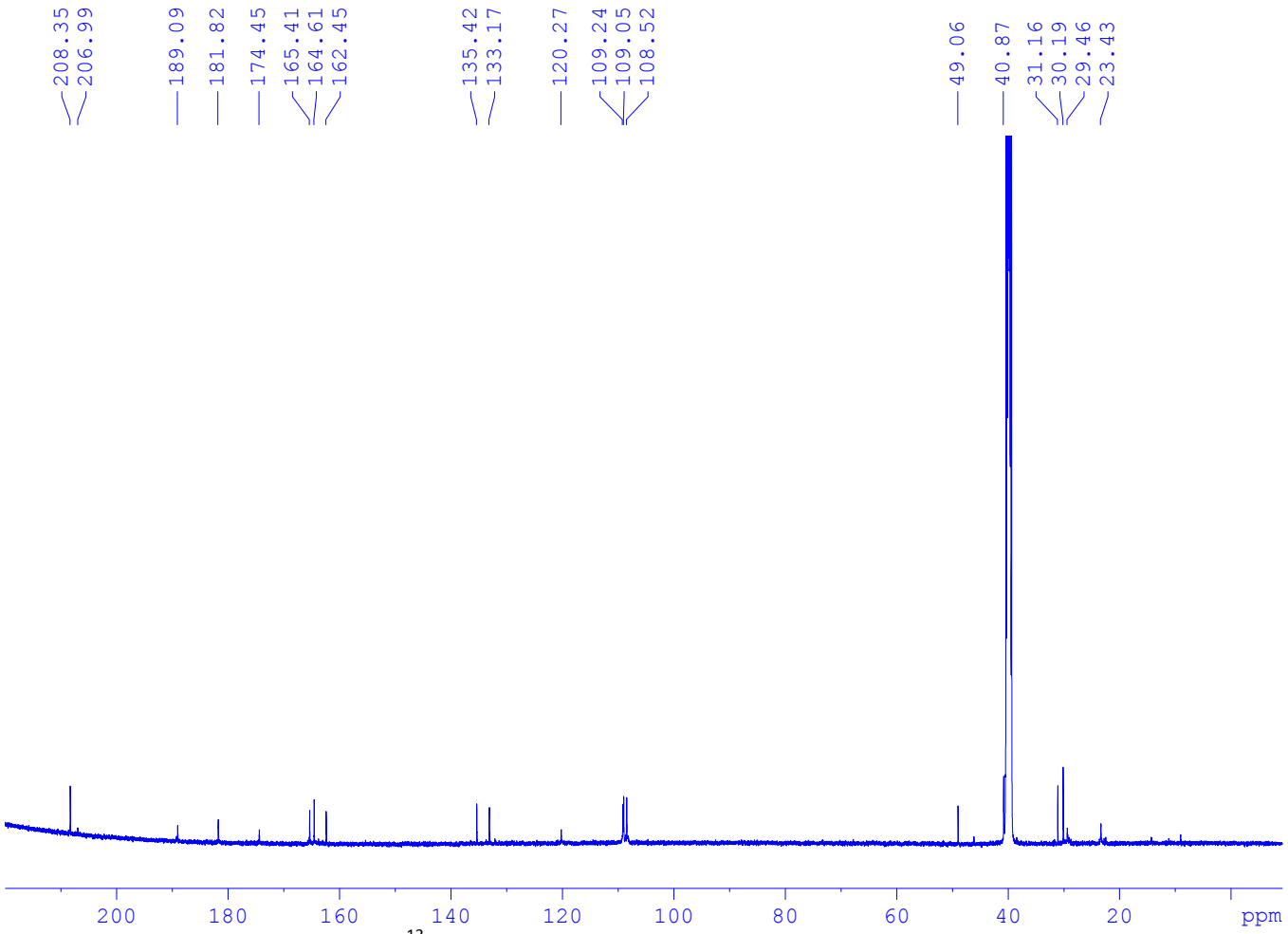



**Figure S34:** $^{13}$C NMR of paeciloquinone B (**8**) in DMSO-$d_6$ (125 MHz)

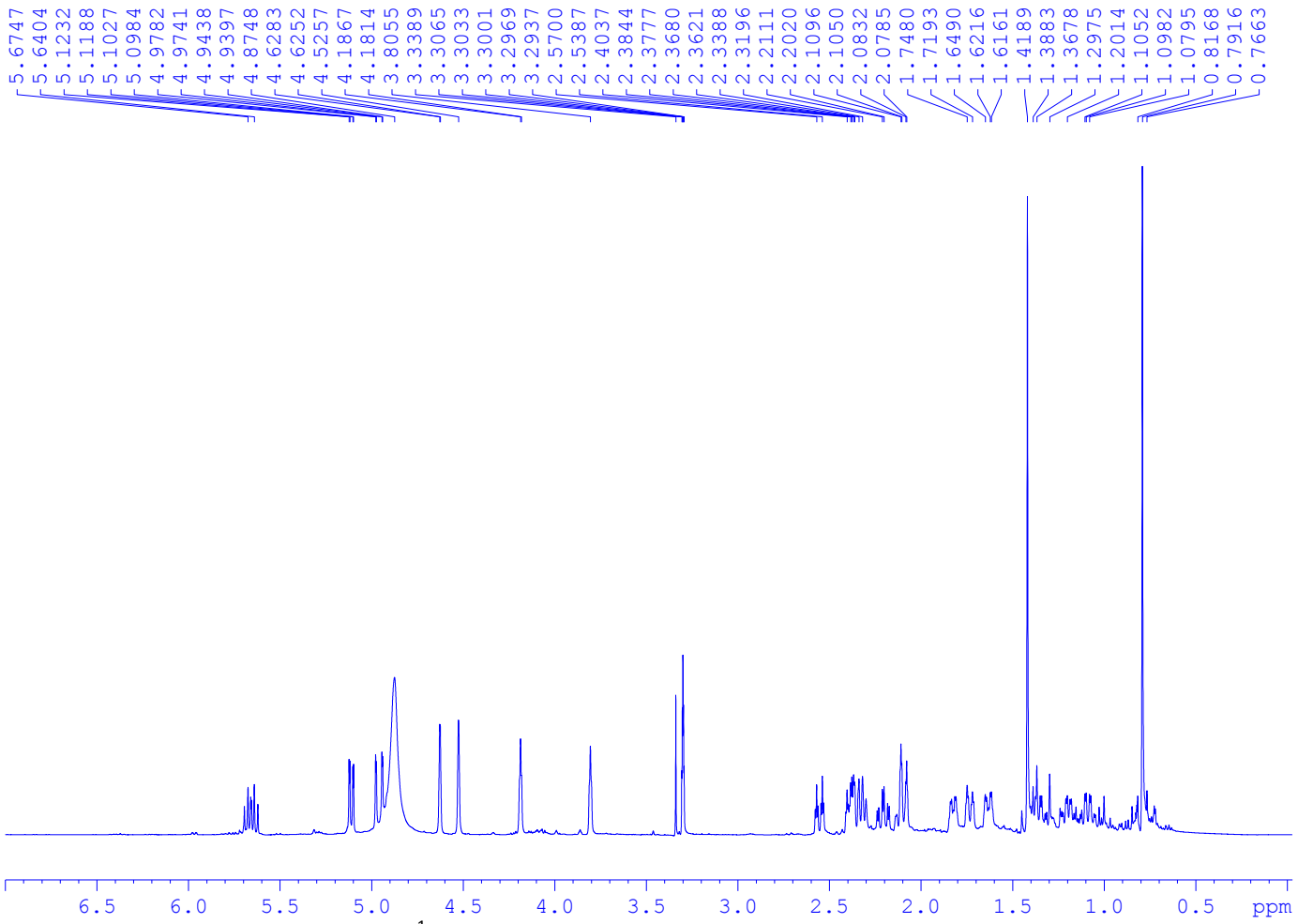



**Figure S35**: $^{1}$H NMR of zythiostromic acid A (**9**) in $CD_3OD$ (500MHz)

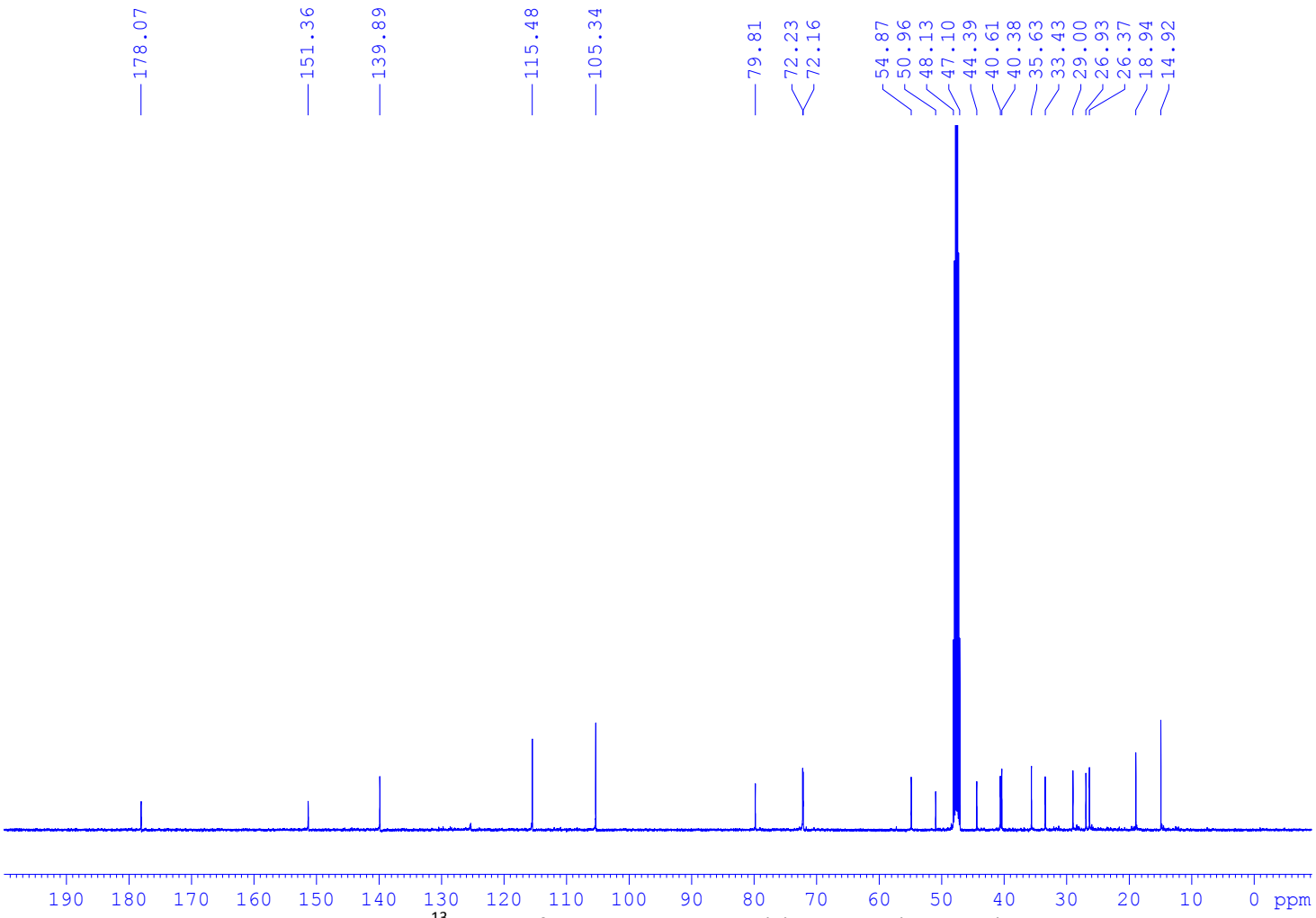



**Figure S36:** $^{13}C$ NMR of zythiostromic acid A (**9**) in $CD_3OD$ (125 MHz)

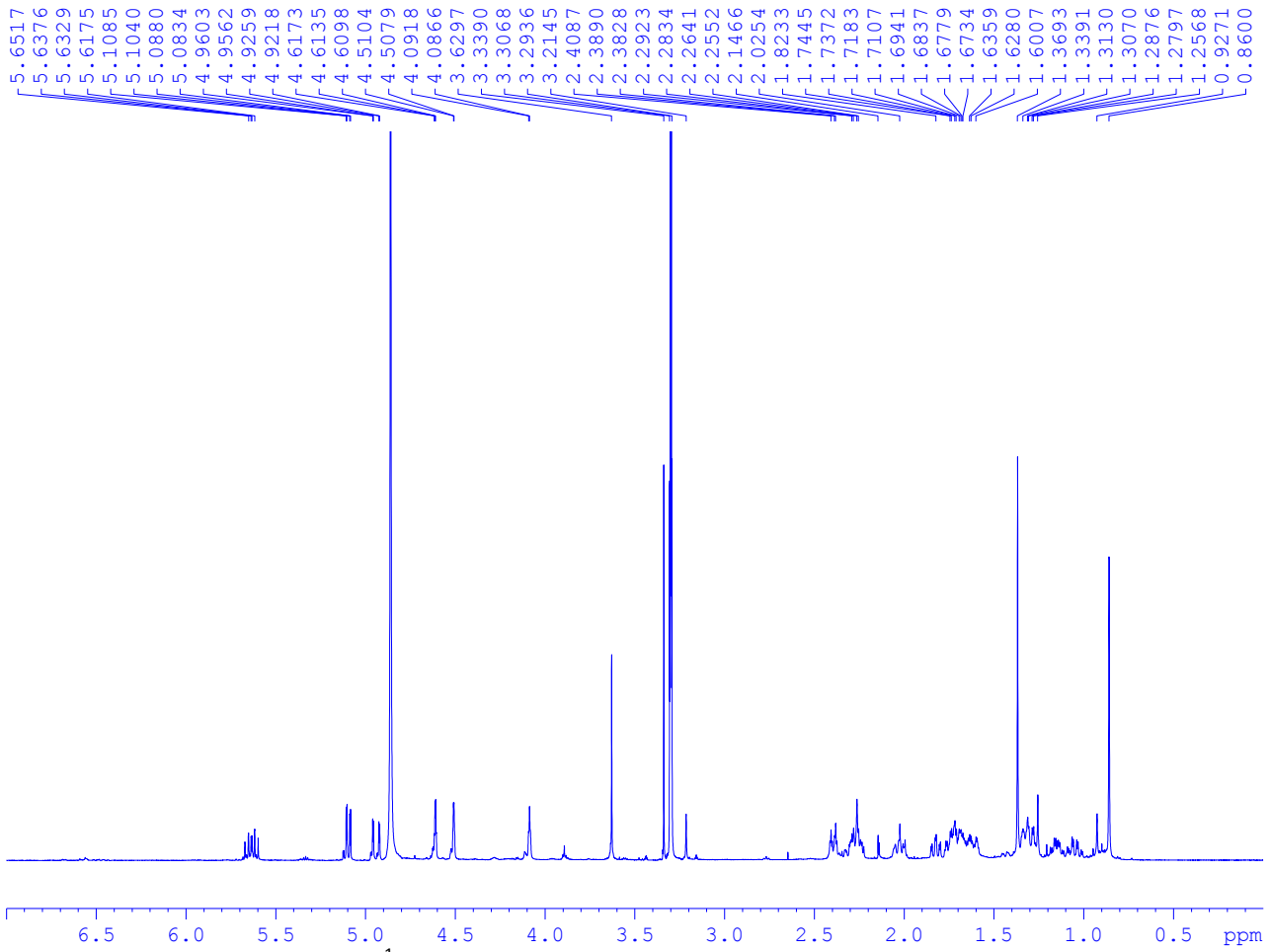



**Figure S37**: $^{1}$H NMR of zythiostromic acid B (**10**) in $CD_3OD$ (500MHz)

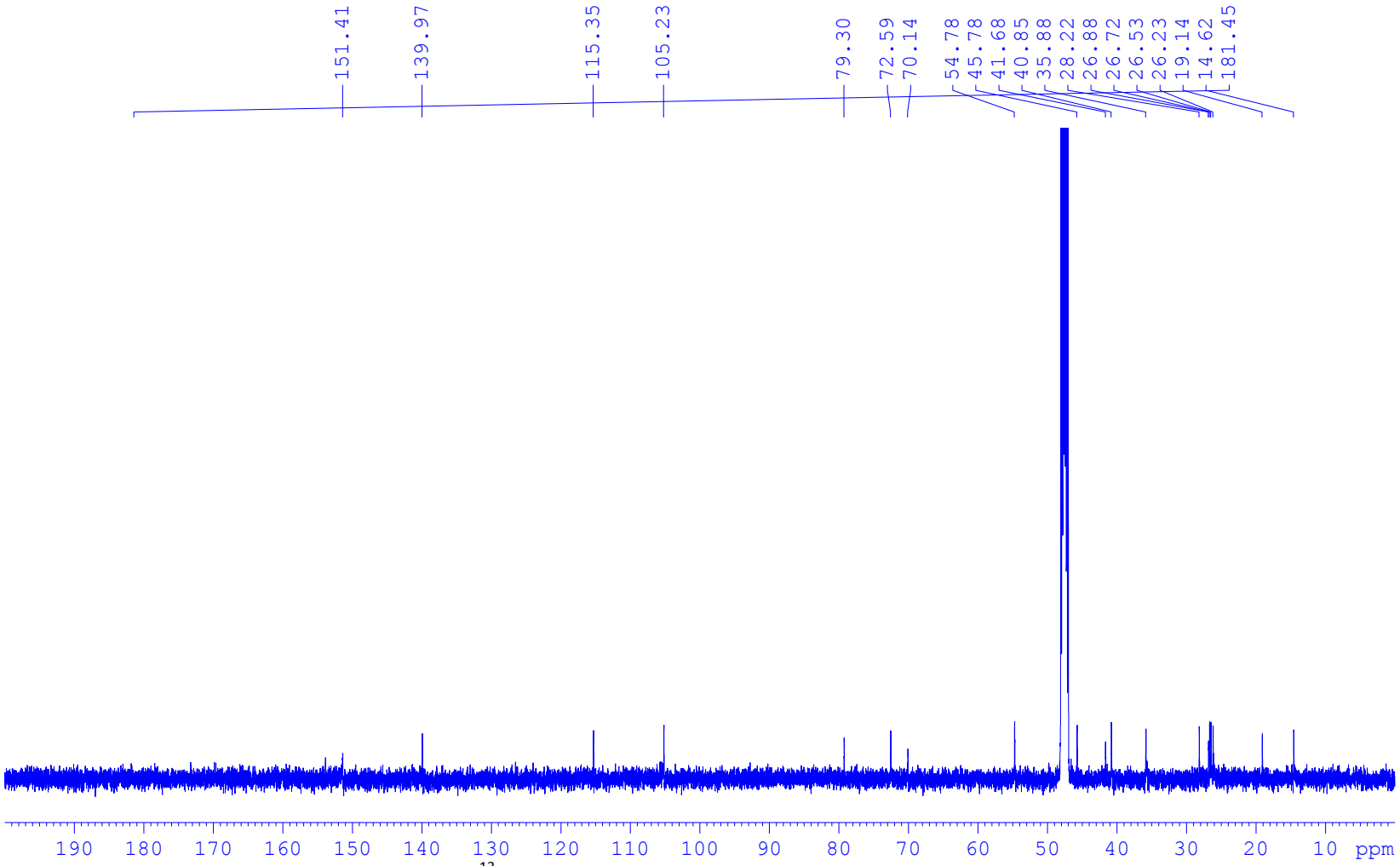



**Figure S38:** $^{13}C$ NMR of zythiostromic acid B (**10**) in $CD_3OD$ (125 MHz)

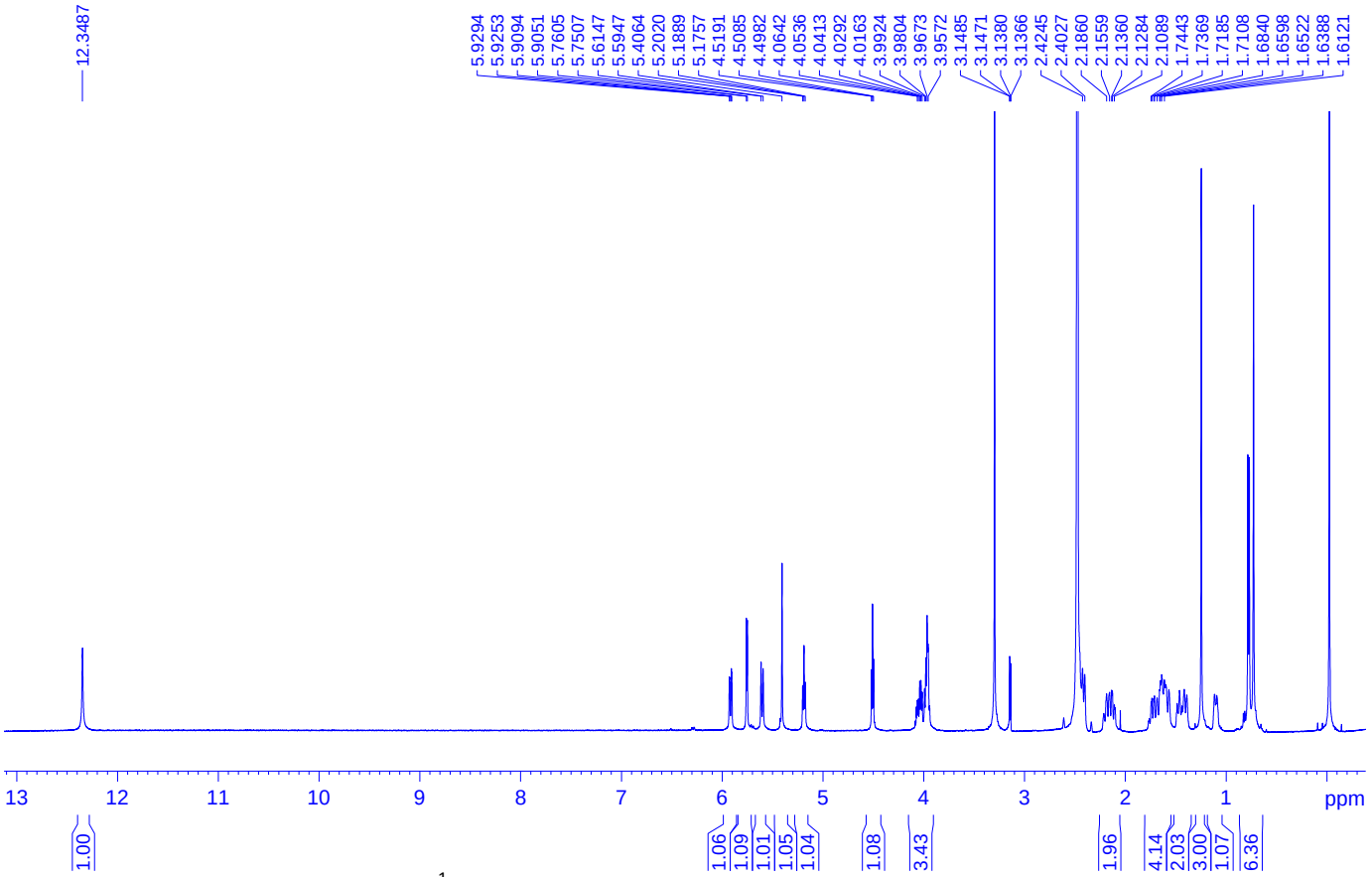



**Figure S39**: $^{1}$H NMR of arbusculic acid A (**11**) in DMSO-$d_6$ (500 MHz)

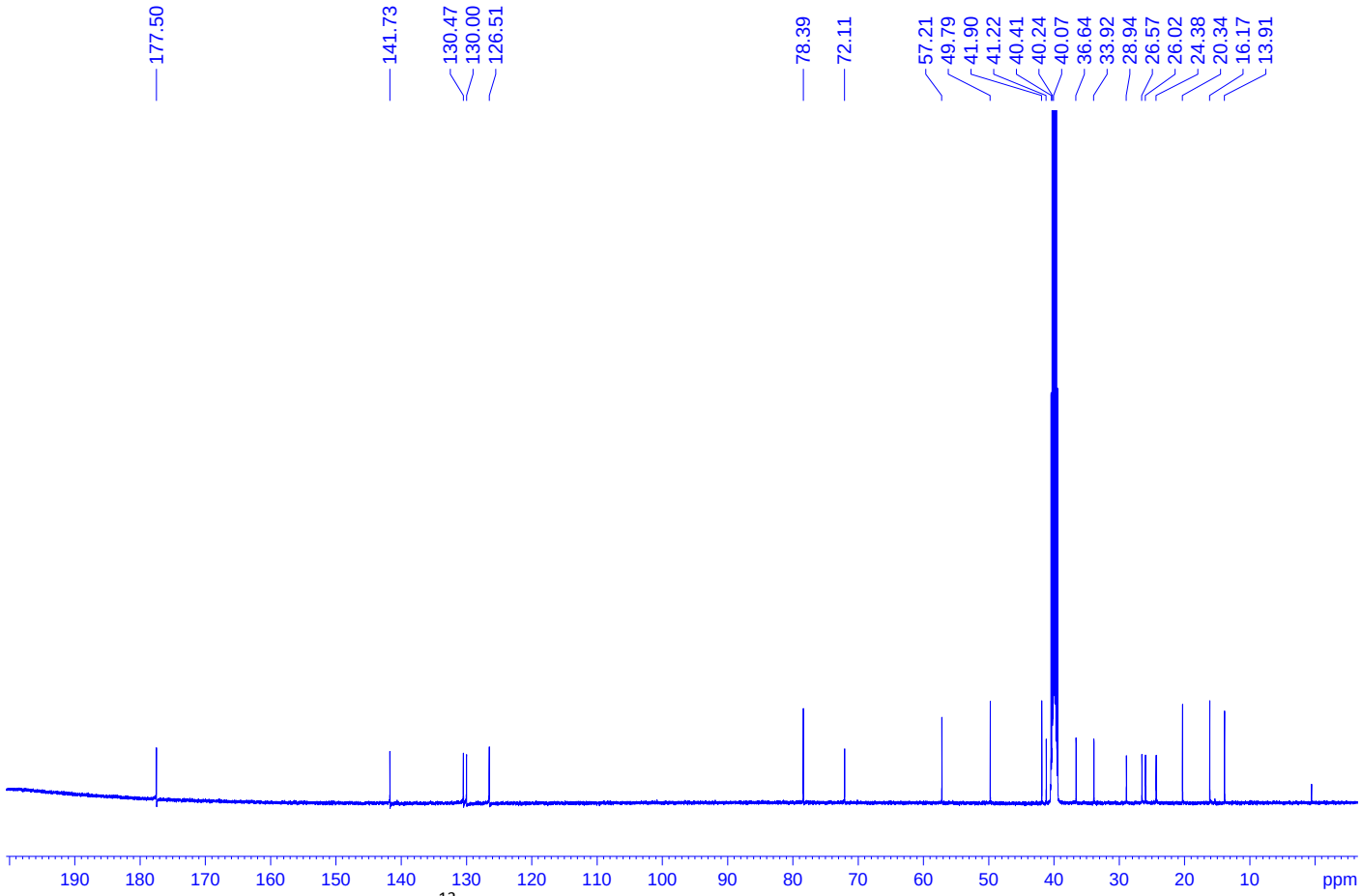



**Figure S40**: $^{13}$C NMR of arbusculic acid A (**11**) in DMSO-$d_6$ (125 MHz)

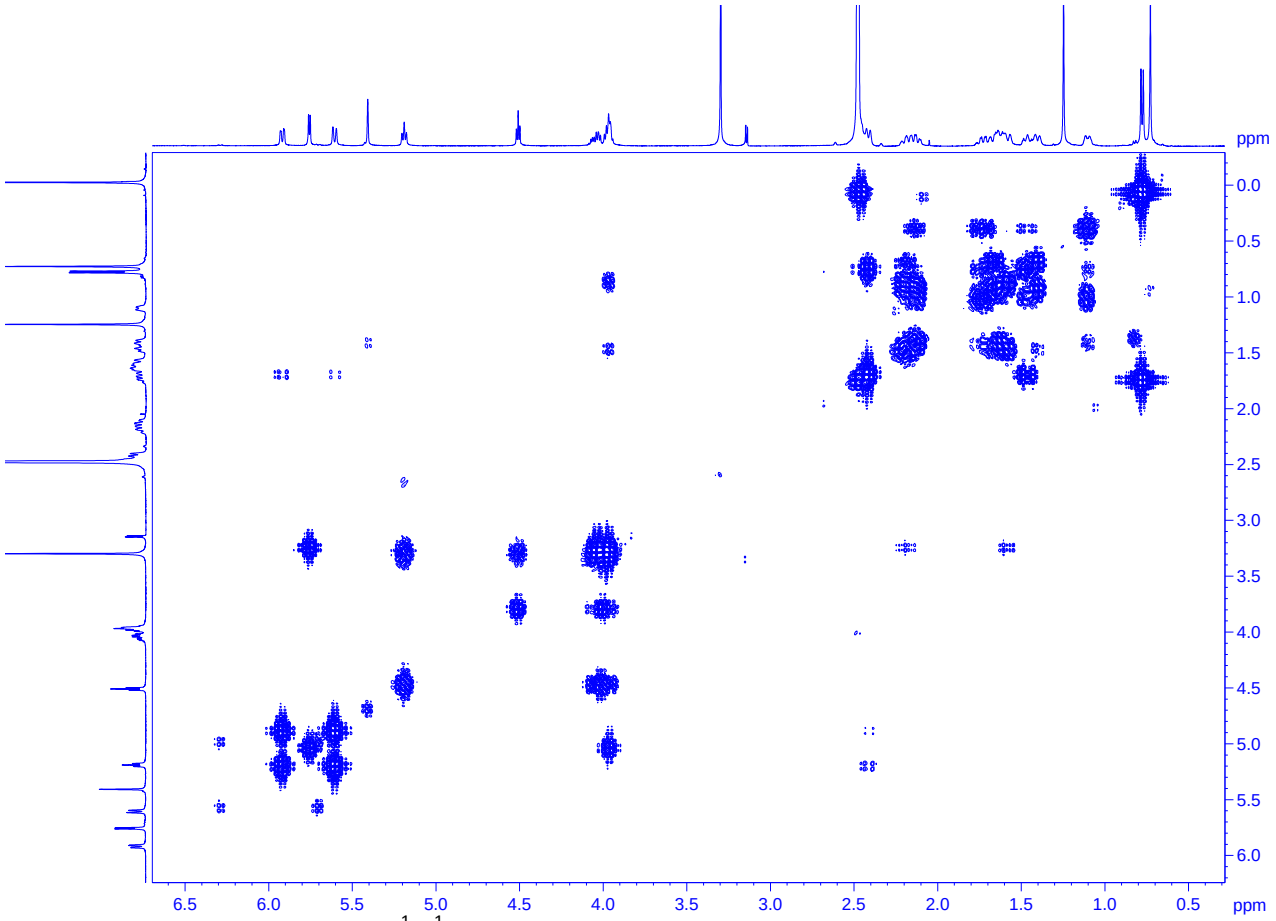



**Figure S41**: $^{1}$H $^{1}$H-COSY NMR of arbusculic acid A (**11**) in DMSO-$d_6$

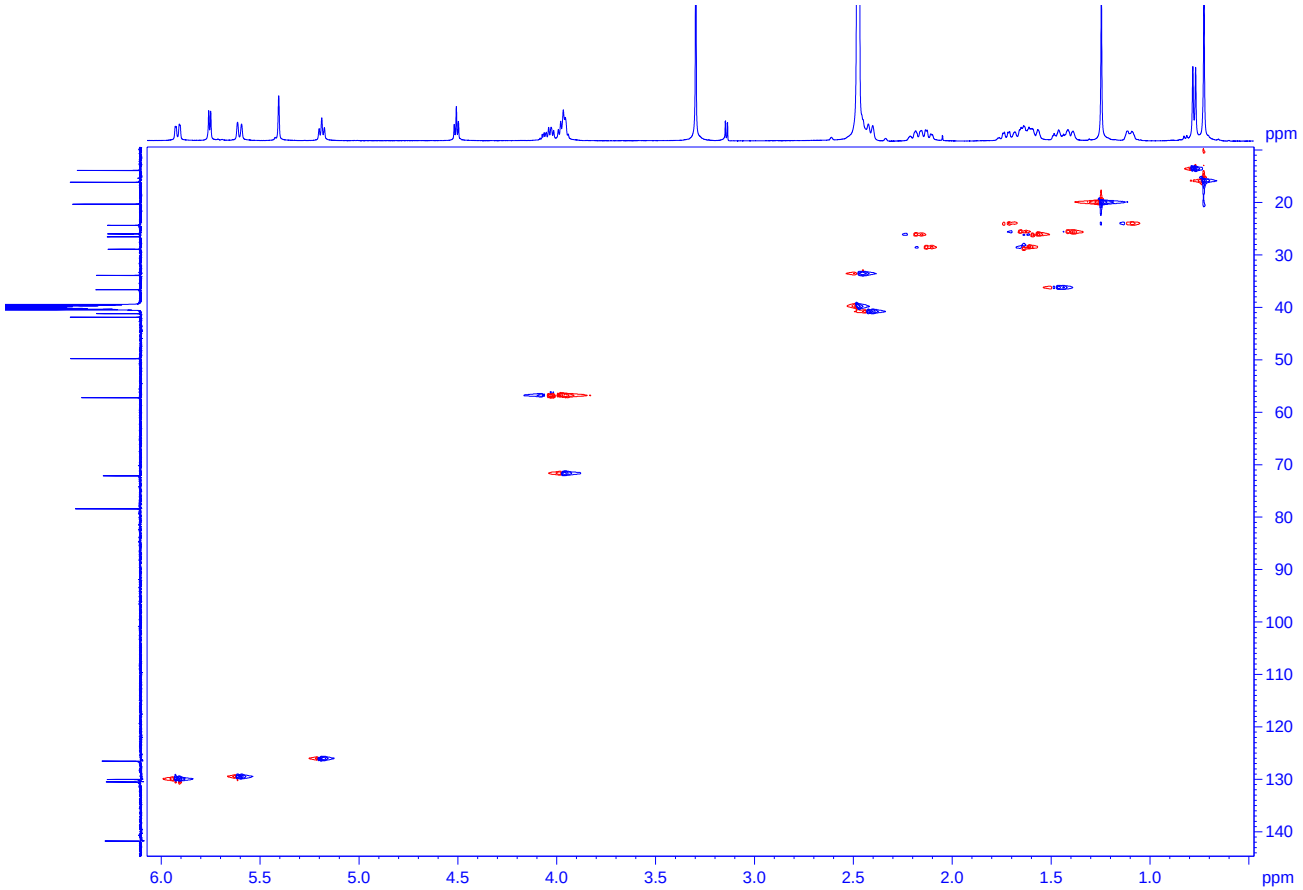



**Figure S42**: HSQC of arbusculic acid A (**11**) in DMSO-$d_6$

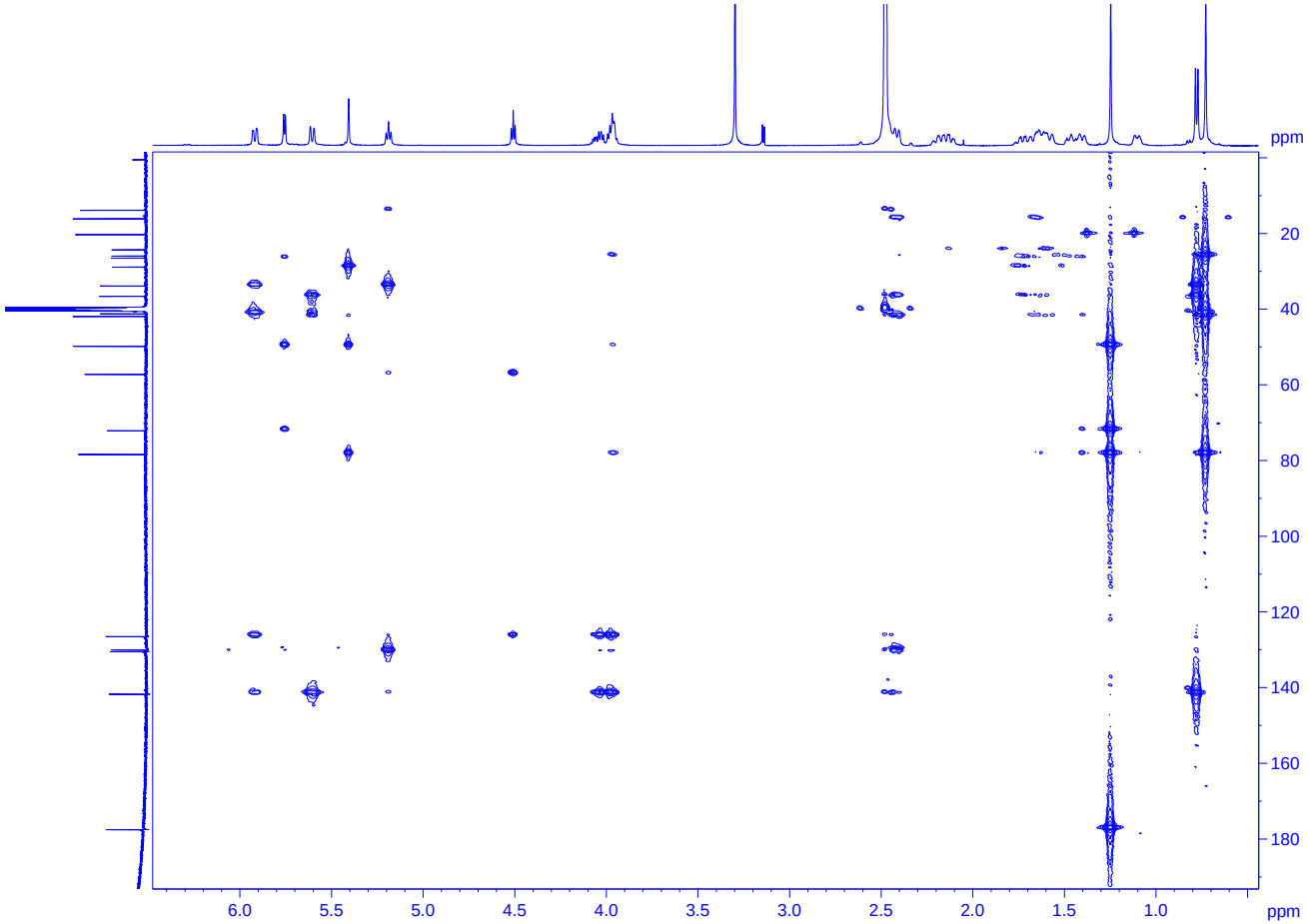



**Figure S43**: HMBC of arbusculic acid A (**11**) in DMSO-$d_6$

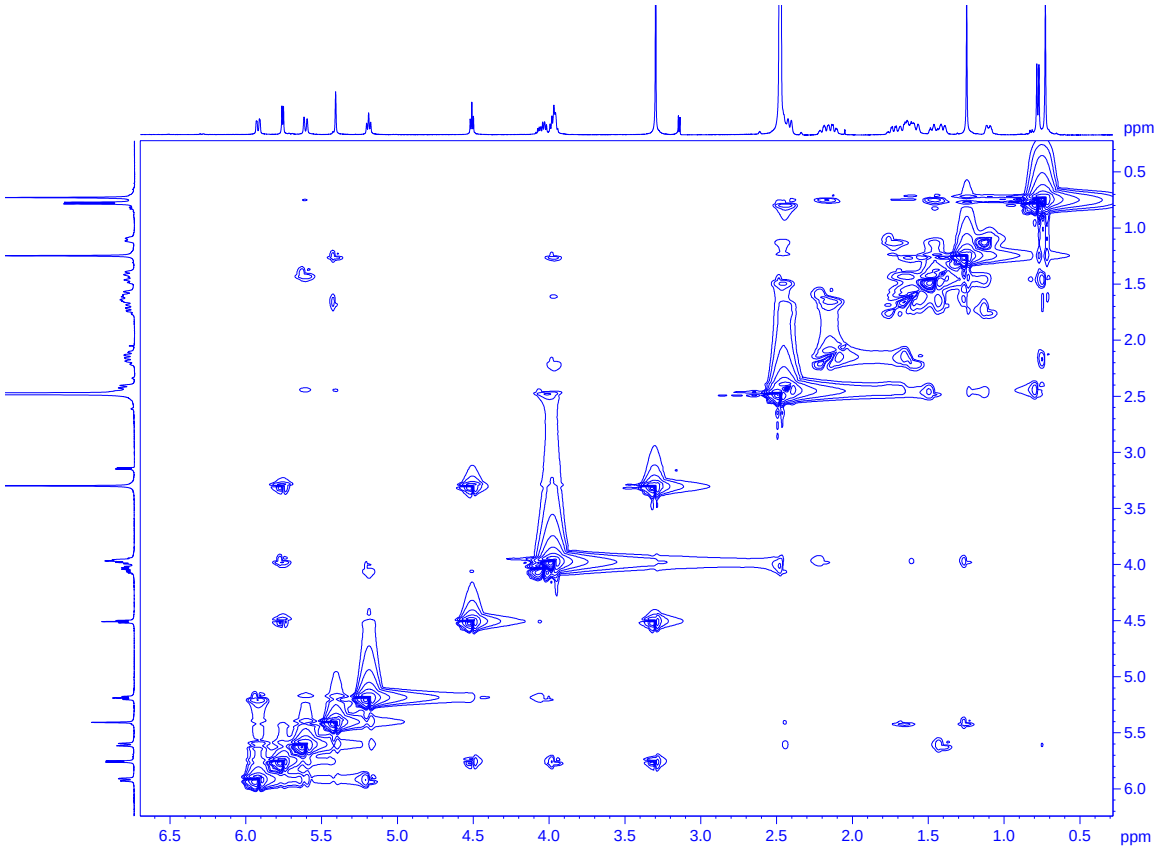



**Figure S44**: NOESY of arbusculic acid A (**11**) in DMSO-$d_6$

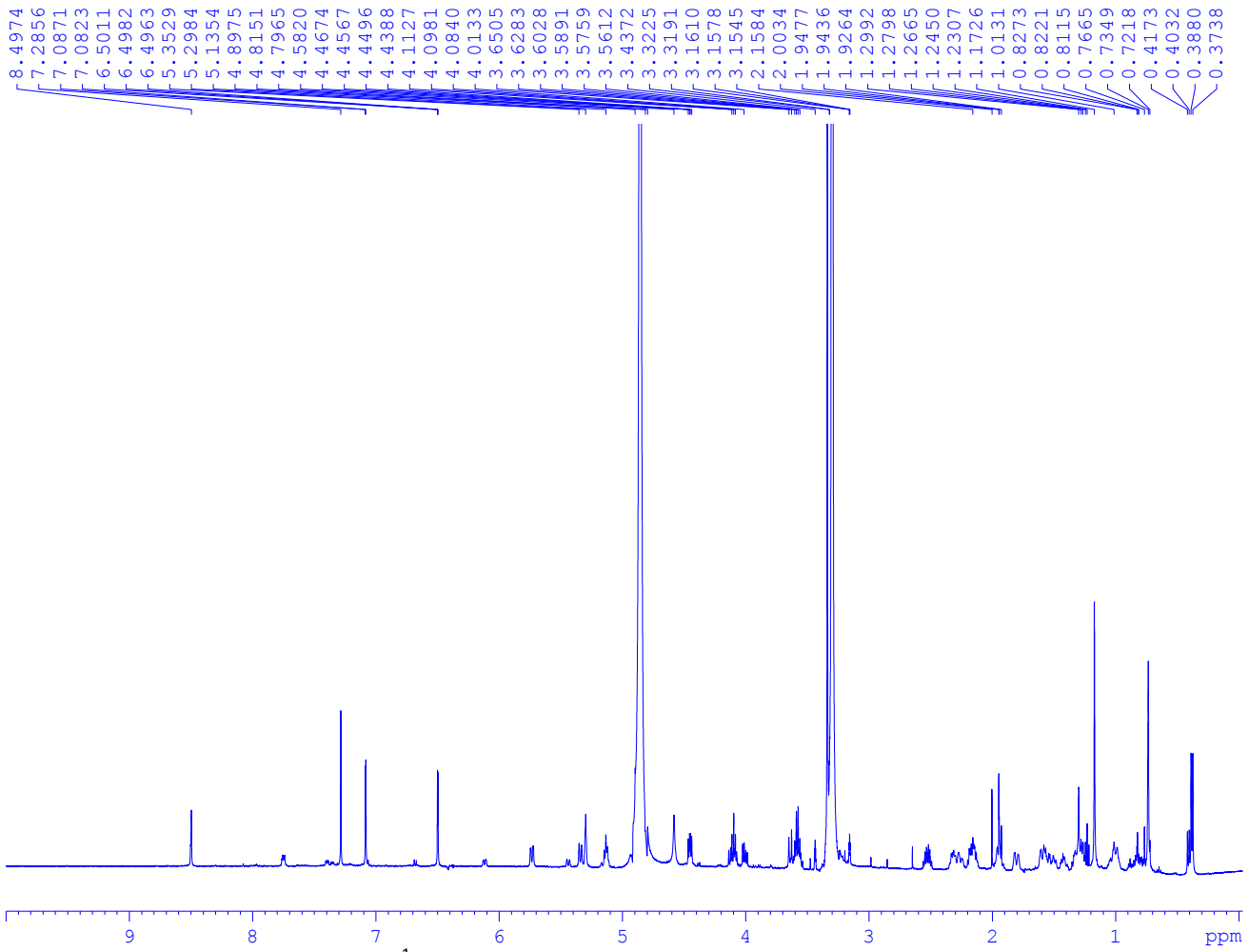



**Figure S45**: $^{1}$H NMR of arbusculic acid B (**12**) in $CD_3OD$ (500MHz)

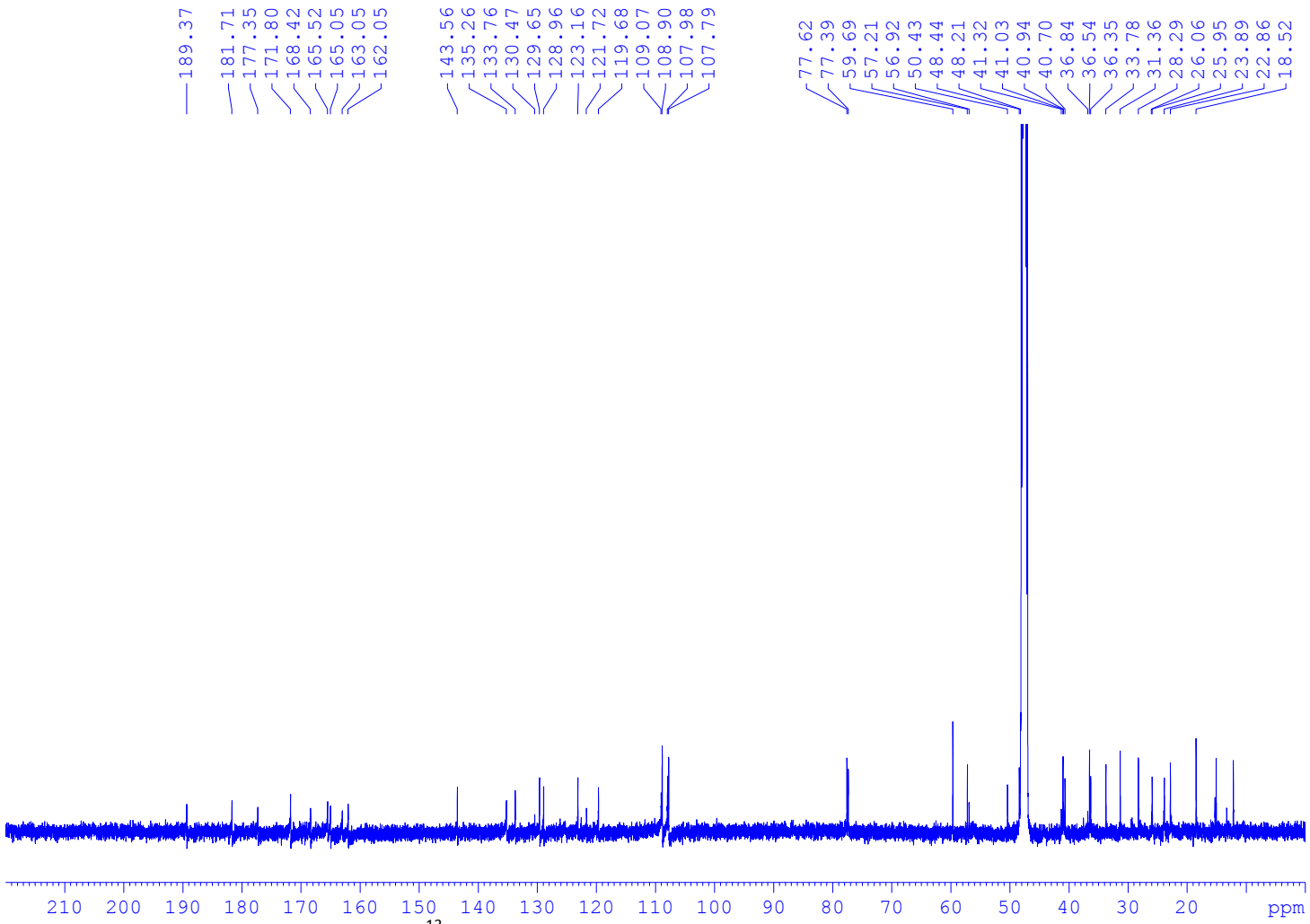



**Figure S46**: $^{13}C$ NMR of of arbusculic acid B (**12**) in $CD_3OD$ (125 MHz)

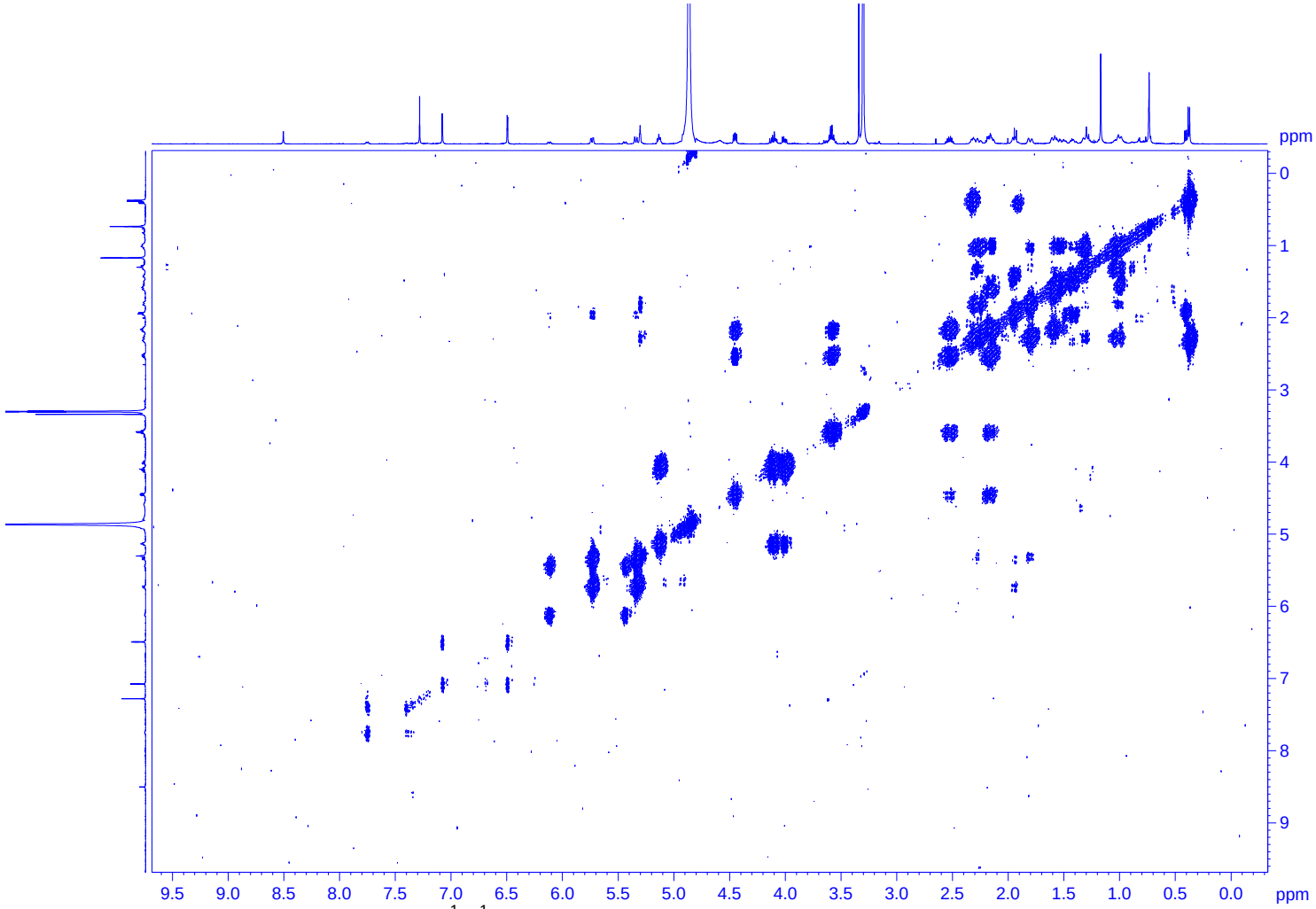



**Figure S47**: $^{1}H$ $^{1}H$-COSY spectrum arbusculic acid B (**12**) in $CD_3OD$

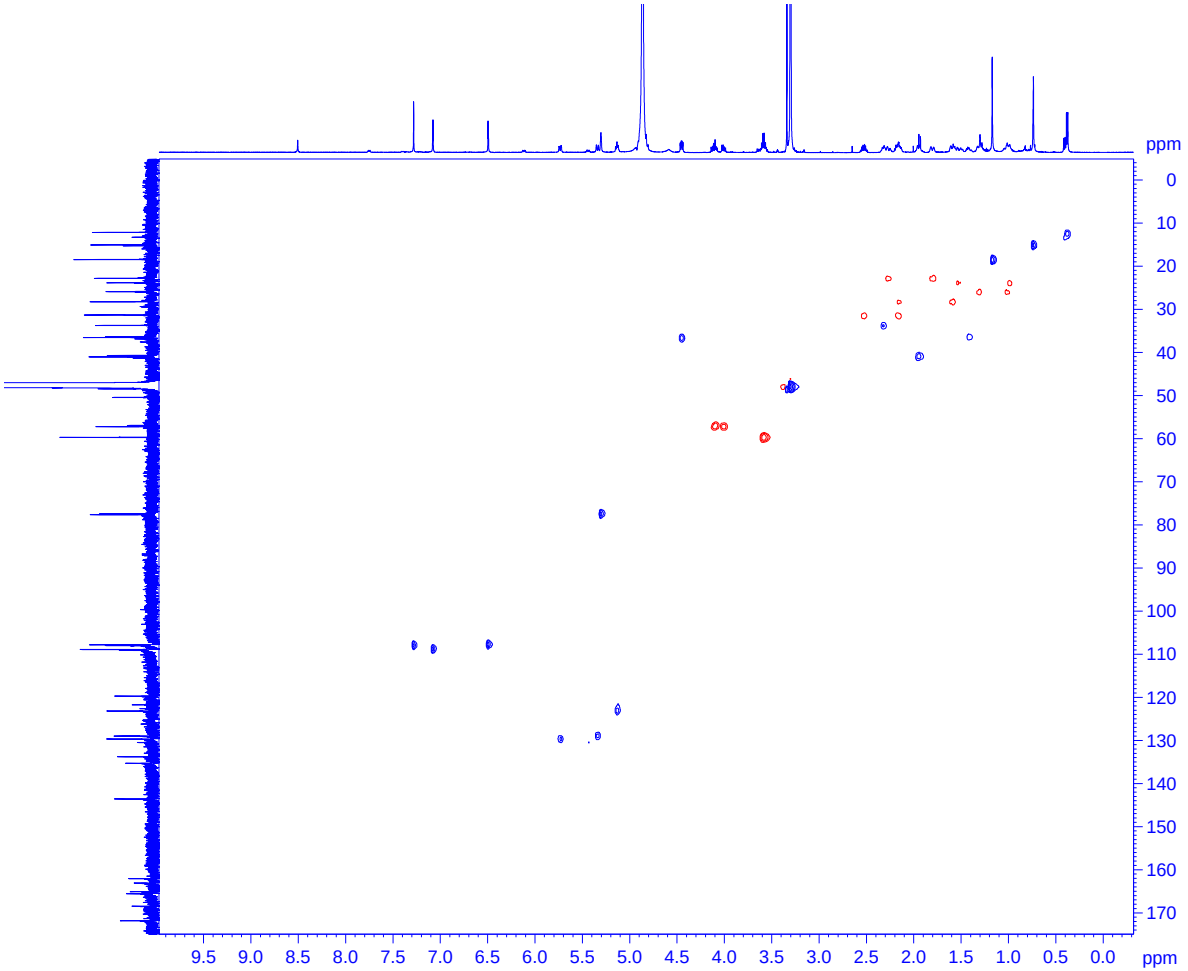



**Figure S48**: HSQC of arbusculic acid B (**12**) in $CD_3OD$

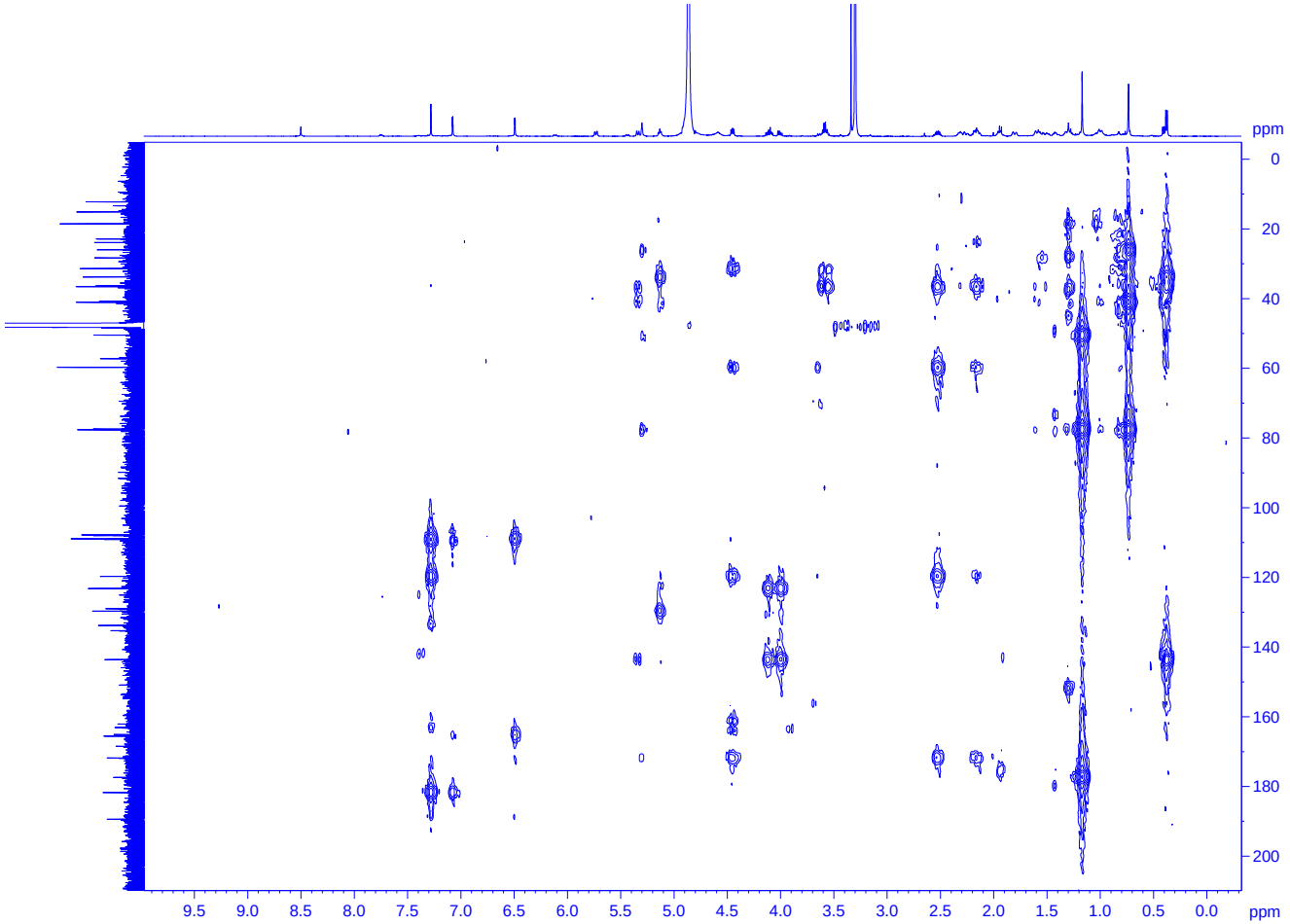
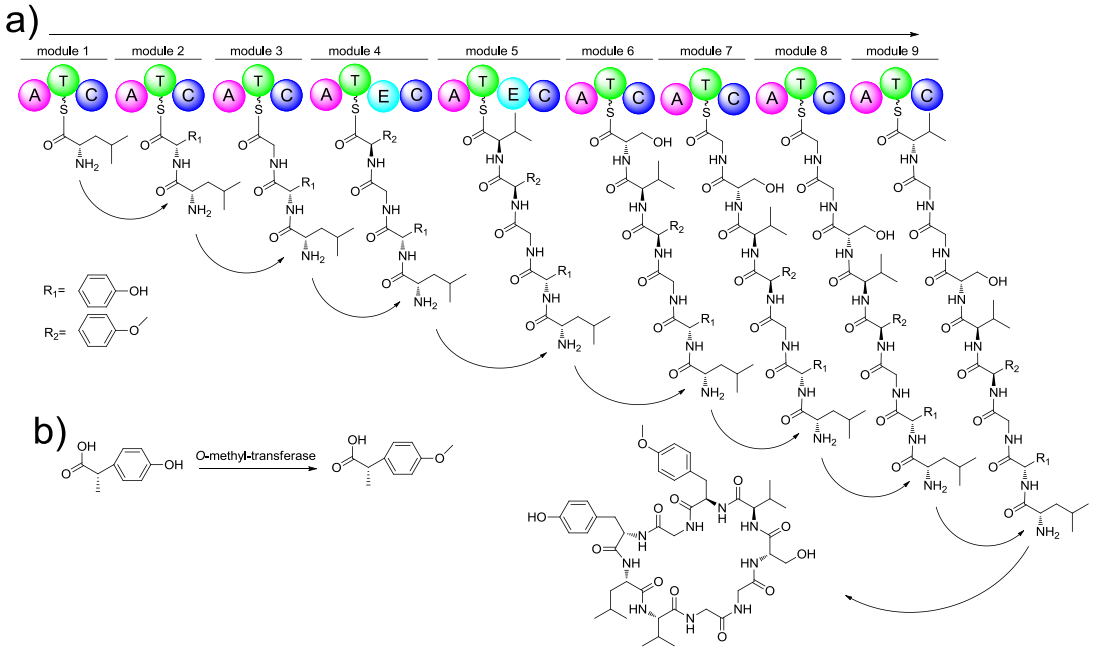
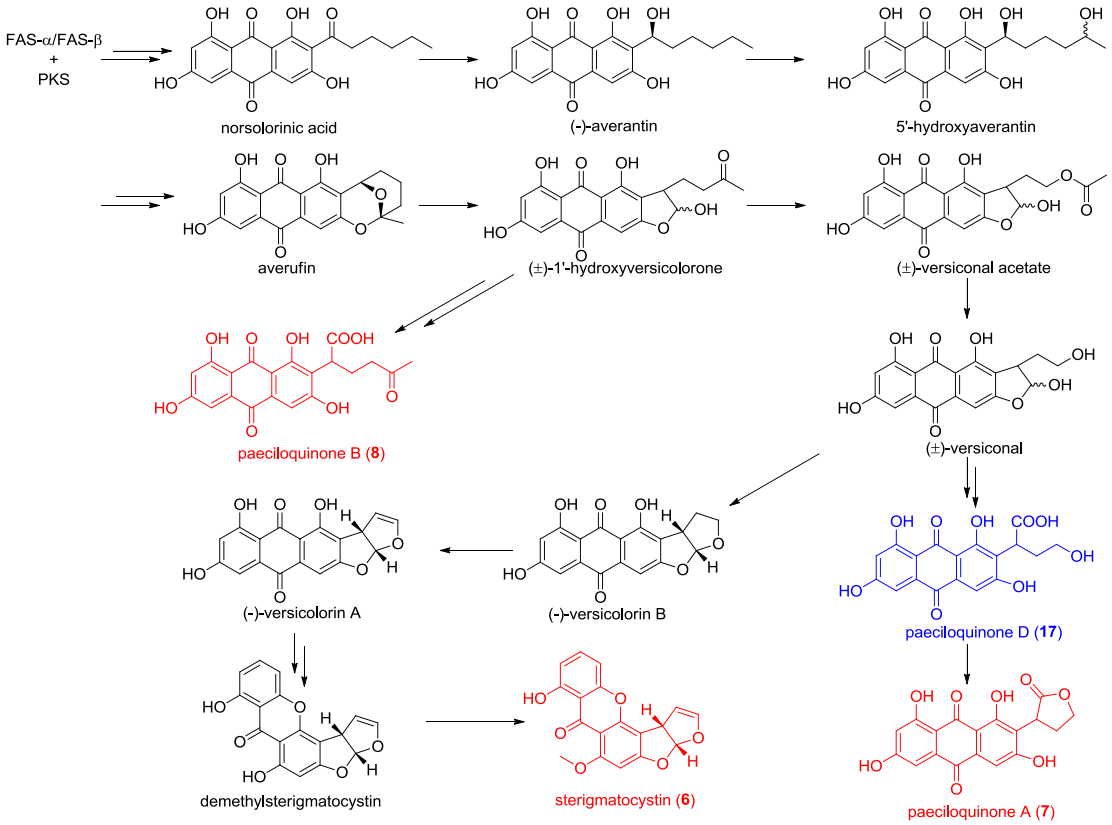



**Figure S49**: HMBC of arbusculic acid B (**12**) in $CD_3OD$




**Scheme S1**. A proposed pathway for the biosynthesis of arbumelin (**4**).




**Scheme S2**. Biosynthesis of sterigmatocystin (**6**), paeciloquinone A (**7**) and B (**8**).

**Supplementary References:**


[1] F. N. Gravelat, D. S. Askew, D. C. Sheppard, *Methods Mol. Biol.* **2012**, *845*, 119-130.

[2] M. A. Collart, S. Oliviero, *Curr. Protoc. Mol. Biol.* **2001**, *Chapter 13*, Unit13 12.

[3] T. Prasch, T. Naumann, R. L. Markert, M. Sattler, W. Schubert, S. Schaal, M. Bauch, H. Kogler, C. Griesinger, *Eur. J. Biochem.* **1997**, *244*, 501-512.