

**SUPPLEMENTARY INFORMATION**

**Structure elucidation and biosynthetic gene cluster analysis of caniferolides A-D, new bioactive 36-membered macrolides from the marine-derived *Streptomyces caniferus* CA-271066**

Ignacio Pérez-Victoria,\* Daniel Oves-Costales,\* Rodney Lacret, Jesús Martín, Marina Sánchez-Hidalgo, Caridad Díaz, Bastien Cautain, Francisca Vicente, Olga Genilloud and Fernando Reyes.

*Fundación MEDINA, Centro de Excelencia en Investigación de Medicamentos Innovadores en Andalucía, Parque Tecnológico de Ciencias de la Salud, Avda. del Conocimiento 34, 18016 Armilla, Granada, Spain.*

\* Corresponding authors: ignacio.perez-victoria@medinaandalucia.es  
daniel.oves@medinaandalucia.es

## *Supplementary Information Contents*

**General experimental procedures**

**Strain isolation, identification and culture conditions**

**Extraction and isolation of caniferolides A-D**

**Compounds characterization**

**Assays of antifungal activity**

**Assays of antiproliferative activity**

**Genomic DNA isolation, sequencing and bioinformatic analysis**

**Fig. S1.** Phylogenetic tree relating *Streptomyces* strains producing caniferolides A-D (our CA-271066 strain) and deplelides A and B (strain MM581-NF15).

**Table S1.** Deduced functions of the ORFs found in the 187 Kb genomic region harbouring the caniferolides biosynthetic gene cluster.

**Fig. S2.** UPGMA phylogenetic analysis of the AT domains from the type I polyketide synthases ScaP1-ScaP8 (left) and the sequence alignment of the three conserved motifs (I, II and III) within the AT domains which are diagnostic of substrate specificity.

**Fig. S3.** Sequence alignment of the loop and catalytic regions within the KR domains highlighting the aminoacid residues which are diagnostic of the stereochemical outcome of the keto-reduction.

**Fig. S4.** Sequence alignment of the HxxxGxxxxP-containing region within the DH domains.

**Fig. S5.** Sequence alignment of the ER domains.

**Fig. S6.** Proposed biosynthetic pathway leading to the alkylnaphthoquinone moiety.

**Table S2.** Predicted stereochemical outcome compared to the stereochemical assignment obtained by NMR.

**Fig. S7.** UV (DAD) spectrum of caniferolide A (**1**).

**Fig. S8.** ESI-TOF spectrum (negative mode) of caniferolide A (**1**).

**Fig. S9.** ESI-TOF spectrum (positive mode) of caniferolide A (**1**) including expansions showing key in-source fragment ions.

**Fig. S10.**  $^1\text{H}$  NMR spectrum ( $\text{CD}_3\text{OD}$ , 500 MHz) of caniferolide A (**1**).

**Fig. S11.**  $^{13}\text{C}$  NMR spectrum ( $\text{CD}_3\text{OD}$ , 125 MHz) of caniferolide A (**1**).

**Fig. S12.** JRES spectrum of caniferolide A (**1**).

**Fig. S13.** COSY spectrum of spectrum of caniferolide A (**1**).

**Fig. S14.** TOSY spectrum of spectrum of caniferolide A (**1**).

**Fig. S15.** NOESY spectrum of spectrum of caniferolide A (**1**).

**Fig. S16.** Edited HSQC spectrum of caniferolide A (**1**).

**Fig. S17.** HSQC-TOCSY spectrum of caniferolide A (**1**).

**Fig. S18.** HMBC spectrum of spectrum of caniferolide A (**1**).

**Fig. S19.** UV (DAD) spectrum of caniferolide B (**2**).

**Fig. S20.** ESI-TOF spectrum (positive mode) of caniferolide B (**2**) including expansions showing key in-source fragment ions.

**Fig. S21.**  $^1\text{H}$  NMR spectrum ( $\text{CD}_3\text{OD}$ , 500 MHz) of caniferolide B (**2**).

**Fig. S22.**  $^{13}\text{C}$  NMR spectrum ( $\text{CD}_3\text{OD}$ , 125 MHz) of caniferolide B (**2**).

**Fig. S23.** JRES spectrum of caniferolide B (**2**).

**Fig. S24.** COSY spectrum of spectrum of caniferolide B (**2**).

**Fig. S25.** TOSY spectrum of spectrum of caniferolide B (**2**).

**Fig. S26.** NOESY spectrum of spectrum of caniferolide B (**2**).

**Fig. S27.** Edited HSQC spectrum of caniferolide B (**2**).

**Fig. S28.** HSQC-TOCSY spectrum of caniferolide B (**2**).

**Fig. S29.** HMBC spectrum of spectrum of caniferolide B (**2**).

**Fig. S30.** UV (DAD) spectrum of caniferolide C (**3**).

**Fig. S31.** ESI-TOF spectrum (positive mode) of caniferolide C (**3**) including expansions showing key in-source fragment ions.

**Fig. S32.**  $^1\text{H}$  NMR spectrum ( $\text{CD}_3\text{OD}$ , 500 MHz) of caniferolide C (**3**).

**Fig. S33.**  $^{13}\text{C}$  NMR spectrum ( $\text{CD}_3\text{OD}$ , 125 MHz) of caniferolide C (**3**).

**Fig. S34.** JRES spectrum of caniferolide C (**3**).

**Fig. S35.** COSY spectrum of spectrum of caniferolide C (**3**).

**Fig. S36.** TOSY spectrum of spectrum of caniferolide C (**3**).

**Fig. S37.** NOESY spectrum of spectrum of caniferolide C (**3**).

**Fig. S38.** Edited HSQC spectrum of caniferolide C (**3**).

**Fig. S39.** HSQC-TOCSY spectrum of caniferolide C (**3**).

**Fig. S40.** HMBC spectrum of spectrum of caniferolide C (**3**).

**Fig. S41.** UV (DAD) spectrum of caniferolide D (**4**).

**Fig. S42.** ESI-TOF spectrum (positive mode) of caniferolide D (**4**) including expansions showing key in-source fragment ions.

**Fig. S43.**  $^1\text{H}$  NMR spectrum ( $\text{CD}_3\text{OD}$ , 500 MHz) of caniferolide D (**4**).

**Fig. S44.**  $^{13}\text{C}$  NMR spectrum ( $\text{CD}_3\text{OD}$ , 125 MHz) of caniferolide D (**4**).

**Fig. S45.** JRES spectrum of caniferolide D (**4**).

**Fig. S46.** COSY spectrum of spectrum of caniferolide D (**4**).

**Fig. S47.** TOSY spectrum of spectrum of caniferolide D (**4**).

**Fig. S48.** NOESY spectrum of spectrum of caniferolide D (**4**).

**Fig. S49.** Edited HSQC spectrum of caniferolide D (**4**).

**Fig. S50.** HMBC spectrum of spectrum of caniferolide D (**4**).

**Fig. S51.** Structure of caniferolide A (**1**) highlighting with different color its stereoclusters.

**Fig. S52.** Data sets employed of Kishis's universal NMR database method.

**Fig. S53.** Determination of the relative configuration of the C-4 to C-7 stereocluster.

**Fig. S54.** Determination of the relative configuration of the C-31 to C-41 stereocluster.

**Fig. S55.** Equivalent segment in caniferolide A (**1**), deplelide A and brasiliolide C key for determination of the 31S configuration in both **1** and deplelide A.

**Fig. S56.** Truncated molecular models showing the key NOEs which connect the relative configuration of the carbohydrate stereoclusters between themselves and with the macrolide aglycon.

**Fig. S57.** Determination of the relative configuration of the C-2''' to C-3''' stereocluster.

**Fig. S58.** Truncated molecular model showing the key NOEs originally reported for PM100118 which connect the relative configuration of the axenose and rhodinose stereoclusters.

**Fig. S59.** Molecular model of caniferolide A (**1**) aglycon.

**Fig. S60.** Molecular model of caniferolide A (**1**).

**Limitations on the application of Matsunaga's empirical rule for establishing the relative stereochemistry of the 1,5-diol motifs in the caniferolides.**

**Fig. S61.** Overlay of three low-energy conformers of caniferolide A (**1**) compatible with the NMR data.

**Fig. S62.** Expansion of the three low-energy conformers of caniferolide A (**1**) and implications for the application of Matsunaga's empirical rule to assign the relative stereochemistry of the 1,5-diol motifs.

**Fig. S63.** Growth inhibition curves of **1-4** and amphotericin B against *Aspergillus fumigatus* ATCC46645.

**Fig. S64.** Growth inhibition curves of **1-4** and amphotericin B against *Candida albicans* MY1055.

**Fig. S65.** Antiproliferative activity curves of **1-3** and doxorubicin against A549 cell line (lung carcinoma).

**Fig. S66.** Antiproliferative activity curves of **1-3** and doxorubicin against A2058 cell line (skin melanoma).

**Fig. S67.** Antiproliferative activity curves of **1-4** and doxorubicin against HepG2 cell line (hepatocellular carcinoma).

**Fig. S68.** Antiproliferative activity curves of **1-3** and doxorubicin against MCF-7 cell line (breast adenocarcinoma).

**Fig. S69.** Antiproliferative activity curves of **1-3** and doxorubicin against MiaPaca-2 cell line (pancreas carcinoma).

## General experimental procedures

Optical rotations were measured with a Jasco P-2000 polarimeter. IR spectra were registered with a JASCO. NMR spectra were recorded on a Bruker Avance III spectrometer (500 and 125 MHz for <sup>1</sup>H and <sup>13</sup>C NMR, respectively) equipped with a 1.7mm MicroCryoprobe, using the signal of the residual solvent as internal reference ( $\delta_{\text{H}}$  3.31 and  $\delta_{\text{C}}$  49.15 ppm for CD<sub>3</sub>OD). LC-UV-MS and LC-HRESIMS analyses were performed as previously described<sup>1,2</sup> on an Agilent 1100 single quadrupole LC-MS system and a Bruker maXis QTOF mass spectrometer coupled to an Agilent 1200 LC. Sephadex LH-20 chromatographic fractionation was carried out by gravity elution and manual collection. Semipreparative RP-HPLC was performed on a GILSON GX-281 322H2 LC with UV-vis detection. Acetone used for extraction was analytical grade. Solvents employed for isolation were HPLC grade. Molecular models were generated using Chem3D Pro 12.0. The structures were generated with a manual model-building approach to roughly satisfy the observed NOEs and three-bond coupling constants (i. e. dihedral angles) followed by energy-minimization with the MM2 force field using as gradient convergence criteria an RMS value of 0.001. Molecular modelling figures were generated with PyMol.

## Strain isolation, identification and culture conditions

Strain CA-271066 was isolated from an ascidian collected at the seaside at 2 meters depth in Bahía Ana Chaves in São Tomé (São Tomé and Príncipe). A similarity-based search with the 16S rRNA sequence (1393 nt) against the EzBioCloud database indicated that the strain is closely related to *Streptomyces caniferus* DSM 41453(T) (100% similarity).<sup>3</sup>

A one liter fermentation of the producing microorganism was obtained as follows: a seed culture of the strain was obtained by inoculating two 150 × 25 mm tubes containing 16 mL of seed-medium (soluble starch 20 g/L, glucose 10 g/L, NZ Amine Type E 5 g/L, meat extract 3 g/L, peptone 5 g/L, yeast extract 5 g/L, sea salts 30 g/L, calcium carbonate 1 g/L, pH 7) with 0.8 mL of freshly thawed inoculum stock of CA-271066. The tubes were incubated at 28 °C, 70% relative humidity and 220 rpm for about 96 hours. The fresh inoculum thus generated was mixed and employed to inoculate (2.5% v/v) twenty 250 mL conical flasks each containing 50 mL of APM9-modified medium (glucose 50 g/L, soluble starch 12 g/L, soy flour 30 g/L, CoCl<sub>2</sub>·6H<sub>2</sub>O 2 mg/L, sea salts 30 g/L, calcium carbonate 7 g/L, pH 7). The flasks were incubated at 28 °C, 70% relative humidity and 220 rpm for 6 days before harvesting.

## Extraction and isolation of caniferolides A-D

The 1 L fermentation was extracted with acetone (1 L) under continuous shaking at 220 rpm for 2h. The mycelium was separated by filtration and the supernatant (ca. 2L) was concentrated to 1L under a stream of nitrogen. The aqueous crude extract was extracted with ethyl acetate and the resulting organic extract was fractionated on Sephadex LH-20 using methanol/dichloromethane (2:1) to afford seven fractions (A-G) being fractions C and D active against *A. fumigatus* and *C. albicans*. LH-20 fraction C was subjected to purification by reversed-phase semipreparative HPLC (column Agilent Zorbax RX-C8, 9.4 × 250 mm, 7 µm; 3 ml min<sup>-1</sup>) with a linear gradient of CH<sub>3</sub>CN/H<sub>2</sub>O from 40 to 50% CH<sub>3</sub>CN over 40 min yielding **1** (10.0 mg, R<sub>t</sub> 26 min), **3** (4.0 mg, R<sub>t</sub> 32 min), **2** (3.6 mg, R<sub>t</sub> 42 min, beginning of washing step at 100% CH<sub>3</sub>CN)

and another impure subfraction (in the washing step) which was further purified by reversed-phase semipreparative HPLC (column XBridge C-18, 10 × 150 mm, 5 µm; 3 ml min<sup>-1</sup>, UV detection at 210 and 254 nm) with a linear gradient of CH<sub>3</sub>CN/H<sub>2</sub>O, from 50 to 60% CH<sub>3</sub>CN over 30 min to yield **4** (1.0 mg, R<sub>t</sub> 16 min). Further amounts of **1** (2.0 mg) and **2** (6.8 mg) could be obtained by analogous semipreparative HPLC purification of LH-20 fraction D.

## Compounds characterization

**Caniferolide A (1).** Yellow amorphous solid; [α]<sub>D</sub><sup>24</sup> +11.4° (c 0.29, CH<sub>3</sub>OH); UV (DAD) λ<sub>max</sub> 255, 270 (sh) 338 nm; (+)-HRESIMS *m/z* 1592.8057 [M+NH<sub>4</sub>]<sup>+</sup> (calcd for C<sub>77</sub>H<sub>126</sub>NO<sub>31</sub>S<sup>+</sup>, 1592.8029, Δ 1.8 ppm); <sup>1</sup>H and <sup>13</sup>C NMR data in Tables 1 and 2 (main text).

**Caniferolide B (2).** Yellow amorphous solid; [α]<sub>D</sub><sup>24</sup> +5.0° (c 0.65, CH<sub>3</sub>OH); UV (DAD) λ<sub>max</sub> 255, 270 (sh) 338 nm; (+)-HRESIMS *m/z* 1512.8419 [M+NH<sub>4</sub>]<sup>+</sup> (calcd for C<sub>77</sub>H<sub>126</sub>NO<sub>28</sub><sup>+</sup>, 1512.8461, Δ 2.8 ppm); <sup>1</sup>H and <sup>13</sup>C NMR data in Tables 1 and 2 (main text).

**Caniferolide C (3).** Yellow amorphous solid; [α]<sub>D</sub><sup>24</sup> +6.3° (c 0.075, CH<sub>3</sub>OH); UV (DAD) λ<sub>max</sub> 255, 270 (sh) 338 nm; (+)-HRESIMS *m/z* 1496.8540 [M+NH<sub>4</sub>]<sup>+</sup> (calcd for C<sub>77</sub>H<sub>126</sub>NO<sub>27</sub><sup>+</sup>, 1496.8512, Δ 1.9 ppm); <sup>1</sup>H and <sup>13</sup>C NMR data in Tables 1 and 2 (main text).

**Caniferolide D (4).** Yellow amorphous solid; [α]<sub>D</sub><sup>24</sup> +3.3° (c 0.10, CH<sub>3</sub>OH); UV (DAD) λ<sub>max</sub> 255, 270 (sh) 338 nm; (+)-HRESIMS *m/z* 1576.8106 [M+NH<sub>4</sub>]<sup>+</sup> (calcd for C<sub>77</sub>H<sub>126</sub>NO<sub>30</sub>S<sup>+</sup>, 1576.8080, Δ 1.6 ppm); <sup>1</sup>H and <sup>13</sup>C NMR data in Tables 1 and 2 (main text).

## Assays of antifungal activity

Caniferolides A-D were evaluated for their antifungal activity against the filamentous fungus *Aspergillus fumigatus* ATCC46645 following a previously described assay based on the use of resazurin as viability indicator.<sup>4</sup> The compounds were also tested against the yeast *Candida albicans* MY1055 using a broth microdilution assay as previously described.<sup>5, 6</sup> MIC is defined as the lowest concentration of antifungal compound inhibiting visible growth. Amphotericin B was employed as positive control.

## Assays of antiproliferative activity

Caniferolides A-D were evaluated for their antiproliferative activity activity against the following human tumoral cell lines: A549 (CCL-185), lung carcinoma; A2058 (CRL-11147), skin melanoma; Hep G2 (HB-8065), hepatocellular carcinoma; MCF-7 (HTB-22), breast adenocarcinoma; MiaPaca-2 (CRL-1420), pancreatic carcinoma. The human cell lines were obtained from the American Type Culture Collection (Manassas, VA) and cultured in Dulbecco's modified Eagle's medium, supplemented with 10% fetal bovine serum (Sigma) and antibiotics-antimycotics. Cell cultures were maintained in a humified incubator at 37°C with 5% CO<sub>2</sub> and passaged when confluent using trypsin/EDTA.

The cell viability in the presence of caniferolides A-D was evaluated and quantified using the MTT assay. This colorimetric assay is employed for measuring the activity of cellular enzymes that reduce the tetrazolium dye, MTT (3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide), to its insoluble formazan, giving a purple color. This assay measures mitochondrial metabolic activity via NAD(P)H-dependent cellular oxidoreductase enzymes and may, under defined conditions, reflect the number of viable cells.<sup>7</sup>

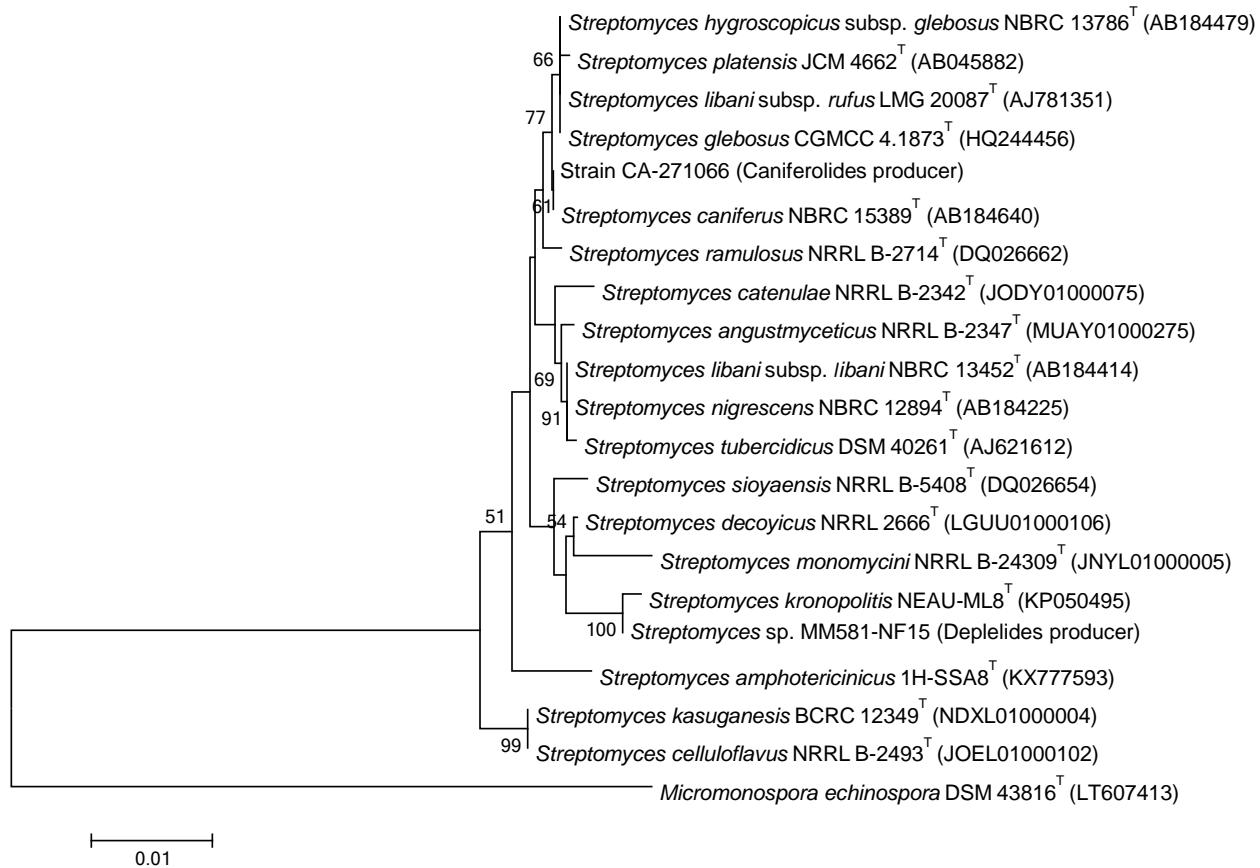
Briefly, cells were seeded at a concentration of  $1 \times 10^4$  cells/well in 200  $\mu\text{L}$  culture medium and incubated at 37°C in 5% CO<sub>2</sub> using 96 wells microplates (BD Falcon). After 24 hours, the medium was replaced with a final volume of 195  $\mu\text{L}$  and 5  $\mu\text{L}$  of each compound solution and controls were added to the plates. 8 mM methyl methanesulfonate (MMS) acts as a positive control and 0.5% DMSO as a negative control. On the last column there are four points of rotenone and doxorubicin with an initial concentration of 10 mM and dilution ½. Plates were incubated at 37°C in 5% CO<sub>2</sub> incubator for 24 hours. After this time, an MTT solution was prepared at 5 mg/mL in PBS 1x and then diluted at 0.5 mg/mL in MEM without phenol red. The sample solution in wells was flicked off and 100  $\mu\text{L}$  of MTT dye was added to each well. The plates were gently shaken and incubated for 3 hours at 37°C in 5% CO<sub>2</sub> incubator. The supernatant was removed and 100  $\mu\text{L}$  of DMSO 100% was added. The plates were gently shaken to solubilize the formed formazan. The absorbance was measured using a Victor2 (Wallac) multireader at a wavelength of 570 nm.

### Genomic DNA isolation, sequencing and bioinformatic analysis

Genomic DNA from strain CA-271066 was extracted and purified as previously described from cultures grown in ATCC-2 liquid medium [0.5% yeast extract (Difco, Franklin Lakes, NJ, USA), 0.3% beef extract (Difco), 0.5% peptone (Difco), 0.1% dextrose (Difco), 0.2% starch from potato (Panreac, Barcelona, Spain), 0.1% CaCO<sub>3</sub> (E. Merck, Darmstadt, Germany), and 0.5% NZ amine E (Sigma, St Louis, MO, USA)].<sup>8</sup>

The whole genome of *S. caniferus* CA-271066 was sequenced *de novo* using a combined strategy of paired-end Illumina HiSeq 2500 sequencing (Service XS, Leiden, the Netherlands; <http://www.servicexs.com>) and single-molecule real-time PacBio sequencing (Macrogen, Seoul, Korea, <http://www.macrogen.com>). Glimmer v3.2 was used for annotation.<sup>9</sup> The contigs obtained from Illumina and PacBio were assembled using Geneious 9.1.8 software (Biomatters, [www.geneious.com](http://www.geneious.com)).<sup>10</sup> The complete sequence of the genomic fragment spanning 186,923 nt containing the caniferolides gene cluster was deposited in GenBank under the accession number MK303577.

The identification of potential biosynthetic gene clusters was carried out with antiSMASH 4.0.0rc1.1. The annotation of the ORFs within the caniferolides biosynthetic gene cluster was carried out employing the corresponding translated sequences and the BLASTP algorithm at the National Center for Biotechnology (NCBI) website.<sup>11</sup> Protein sequence alignments were carried out under the Geneious platform, which employs ClustalW 2.1 (BLOSUM cost matrix). The UPGMA phylogenetic tree of the AT domains was carried out under the Geneious platform employing the Jukes-Cantor genetic distance model.



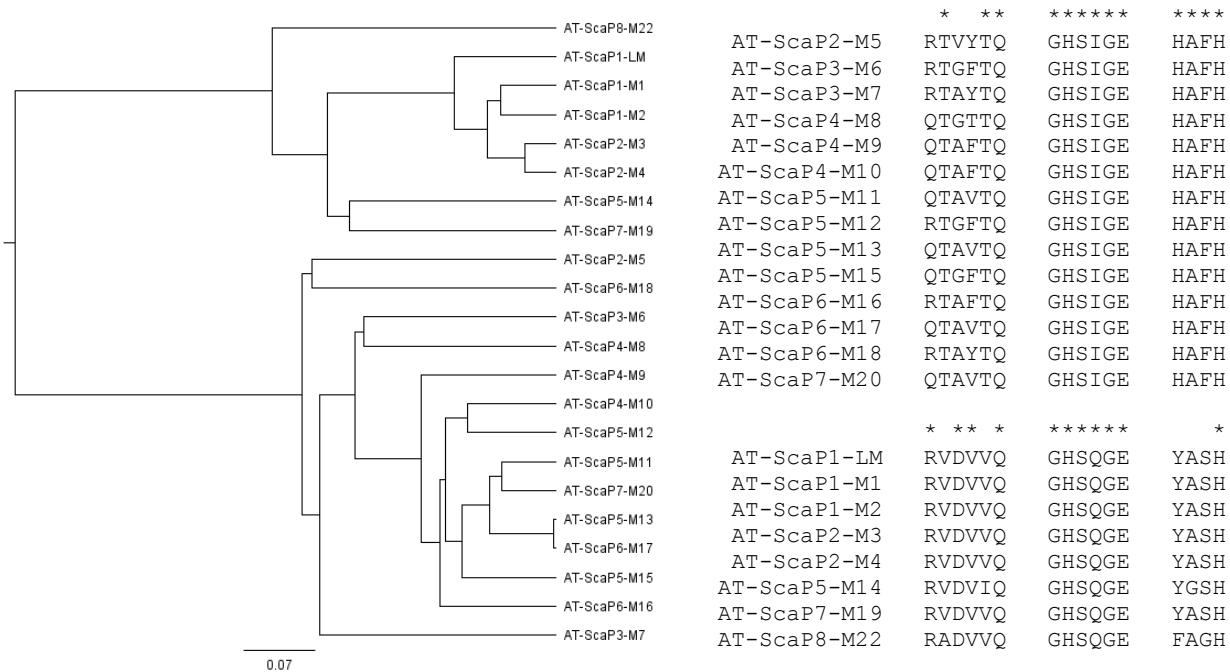
**Figure S1.** Neighbour-joining tree built with MEGA 6.06 based on nearly complete 16S rDNA gene sequences of strain CA-271066 (caniferolides producer), strain MM581-NF15 (deplelides producer) and the closest type strains of the genus *Streptomyces*. *Micromonospora echinospora* DSM 43816(T) was employed as an out-group. Bootstrap support shown in the nodes (only values higher than 50% are displayed). The scale bar indicates 0.01 substitutions per site.

**Table S1.** Deduced functions of the ORFs found in the 187 Kb genomic region harbouring the caniferolides biosynthetic gene cluster.

Gene	Aa lenght	Proposed function	Protein homologue	ID/SM	PM100117/PM100118 analogue	ID/SM
<i>scaH1</i>	238	Hypothetical protein	SEE95350.1	78/84		
<i>scaH2</i>	235	Hypothetical protein	WP_030934905.1	68/75		
<i>scaH3</i>	74	Hypothetical protein	WP_093637847.1	97/98		
<i>scaPT</i>	239	4'-phosphopantetheinyl transferase	SEE95396.1	82/88		
<i>scaTR1</i>	958	LuxR family transcriptional regulator	EJJ02893.1	89/93		
<i>scaTR2</i>	281	PAS-LuxR transcriptional regulator	WP_078568923.1	91/93		
<i>scaTR3</i>	936	LuxR family transcriptional regulator	WP_006607811.1	83/89		
<i>scaTR4</i>	258	Gntr family transcriptional regulator	SEE95458.1	89/92		
<i>scaTR5</i>	210	TetR family transcriptional regulator	WP_078568922.1	93/96		
<i>scaDH1</i>	470	NDP-hexose 2,3-dehydratase	SEE95483.1	92/95	<i>gonD1</i>	86/91
<i>scaP1</i>	4213	Type I polyketide synthase (LM, M1, M2)	WP_096211654.1	85/89	<i>gonP1</i>	82/87
<i>scaP2</i>	4949	Type I polyketide synthase (M3 to M5)	WP_093488000.1	90/93	<i>gonP2</i>	87/91
<i>scaP3</i>	3926	Type I polyketide synthase (M6 and M7)	SOE09192.1	87/90	<i>gonP3</i>	84/88
<i>scaP4</i>	5289	Type I polyketide synthase (M8 to M10)	WP_093488014.1	87/91	<i>gonP4</i>	82/88
<i>scaP5</i>	8444	Type I polyketide synthase (M11 to M15)	CUW01175.1	91/94	<i>gonP5</i>	91/94
<i>scaP6</i>	5240	Type I polyketide synthase (M16 to M18)	CUW01176.1	92/94	<i>gonP6</i>	91/94
<i>scaP7</i>	3642	Type I polyketide synthase (M19 and M20)	WP_096211649.1	90/93	<i>gonP7</i>	84/89
<i>scaF1</i>	67	Ferredoxin	WP_093488020.1	89/92	<i>gonF1</i>	71/77
<i>scaCP1</i>	408	Cytochrome P450	SEF10144.1	91/94		
<i>scaGT1</i>	418	Glycosyltransferase	WP_093488024.1	93/97	<i>gonG1</i>	84/91

Gene	Aa lenght	Proposed function	Protein homologue	ID/SM	PM100117/PM100118 analogue	ID/SM
<i>scaGS</i>	358	Glucose-1-phosphate thymidyltransferase	SCK45567.1	96/98	<i>gonGS</i>	93/96
<i>scaDH2</i>	326	NDP-glucose-4,6-dehydratase	WP_093488029.1	95/97	<i>gonD2</i>	91/93
<i>scaR1</i>	319	NAD(P)H-dependent ketoreductase (SDR fold)	WP_086716628.1	88/91	<i>gonR1</i>	62/70
<i>scaGT2</i>	410	Glycosyltransferase	SEE95784.1	91/94	<i>gonG2</i>	69/78
<i>scaDH3</i>	433	PMP-dependent NDP-4-keto-hexose 3-dehydratase	WP_093488035.1	97/98	<i>gonD3</i>	90/95
<i>scaP8</i>	2390	Type I polyketide synthase (LM and M21)	WP_093489589.1	88/92	<i>gonP8</i>	57/68
<i>scaS</i>	343	3-oxoacyl-ACP synthase	WP_086716618.1	96/98	<i>gonS1</i>	80/89
<i>scaA</i>	145	Cold-shock domain-containing protein	WP_040902049.1	92/97	<i>gonA</i>	70/83
<i>scaT1</i>	283	ABC transporter permease component	SEE96078.1	93/95	<i>gonT1</i>	81/88
<i>scaT2</i>	260	ABC transporter permease component	WP_006608193.1	96/98	<i>gonT2</i>	88/94
<i>scaR2</i>	326	NAD(P)H-dependent ketoreductase (SDR fold)	SOE09154.1	89/93	<i>gonR2</i>	75/82
<i>scaE</i>	206	dTDP-4-keto-2,6-dideoxy-3,5-epimerase	WP_096211634.1	91/93	<i>gonE</i>	80/88
<i>scaCM</i>	410	NDP-hexose-3-C-methyltransferase	WP_093649296.1	96/97	<i>gonCM</i>	90/94
<i>scaH4</i>	243	Hypothetical protein	WP_093488114.1	93/95	<i>gonH1</i>	79/87
<i>scaST</i>	429	Sulfotransferase	WP_006608188.1	91/94		
<i>scaSK</i>	201	Adenylyl-sulfate kinase	WP_006608187.1	82/86		
<i>scaO1</i>	56	Truncated flavin-dependent oxidoreductase (inactive?)	WP_030728166.1	84/96		
<i>scaO2</i>	264	Flavin-dependent oxidoreductase	WP_030985862.1	88/93		
<i>scaB</i>	70	Truncated IS481 family transposase	WP_030317787.1	84/90		
<i>scaC</i>	392	Efflux permease, MFS family	WP_067136124.1	93/95		
<i>scaD</i>	183	Acetyltransferase (GNAT)	WP_097257079.1	79/85		

Gene	Aa lenght	Proposed function	Protein homologue	ID/SM	PM100117/PM100118 analogue	ID/SM
<i>scaF</i>	348	LD-carboxypeptidase	WP_014671118.1	84/90		
<i>scaG</i>	1015	NRPS (adenylation and PCP domains)	WP_063828046.1	89/92		
<i>scaI</i>	3364	IPT/TIG domain-containing protein	WP_093488126.1	77/84		
<i>scaH5</i>	327	Hypothetical protein	WP_006608182.1	86/89		
<i>scaH6</i>	133	Hypothetical protein	WP_093488130.1	83/89		
<i>scaT</i>	256	Thioesterase	WP_096211626.1	95/96	<i>gonT</i>	88/94
<i>scaM1</i>	301	Menaquinone biosynthesis (chorismate dehydratase)	WP_093488134.1	92/95	<i>gonM1</i>	88/94
<i>scaM2</i>	422	Menaquinone biosynthesis (dehypoxanthine fthalosine cyclase)	SEE96290.1	97/98	<i>gonM2</i>	92/93
<i>scaF2</i>	101	Ferredoxin	WP_093488136.1	100/100	<i>gonF2</i>	95/96
<i>scaM3</i>	389	Menaquinone biosynthesis (aminofthalosine synthase)	WP_006608176.1	98/99	<i>gonM3</i>	95/98
<i>scaJ</i>	360	Putative arylcarboxylate reductase	WP_096211625.1	92/95		
<i>scaM4</i>	287	Menaquinone biosynthesis (1,4-dihydroxy-6-naphthoate synthase)	WP_096211624.1	89/93	<i>gonM4</i>	87/90
<i>scaTR6</i>	191	TetR family transcriptional regulator	WP_086715957.1	90/94		
<i>scaH7</i>	148	Hypothetical protein	WP_077192303.1	84/92		
<i>scaR3</i>	359	NDP-hexose-3-ketoreductase	SEE96378.1	87/91	<i>gonR3</i>	84/91
<i>scaCP2</i>	405	Cytochrome P450	WP_093488146.1	98/98	<i>gonCP</i>	75/79
<i>scaK</i>	529	Carboxyl-transferase	WP_096215236.1	95/96		
<i>scaL</i>	441	ATP-dependent acyl-CoA synthetase	WP_093488150.1	92/94		
<i>scaN</i>	302	Putative carboxylesterase	SEE96439.1	82/87		
<i>scaH8</i>	61	Hypothetical protein	XP_021453993.1	33/53		
<i>scaH9</i>	71	Hypothetical protein	WP_040902041.1	58/67		



**Figure S2.** UPGMA phylogenetic analysis of the AT domains from the type I polyketide synthases ScaP1-ScaP8 (left) and the sequence alignment of the three conserved motifs (I, II and III) within the AT domains which are diagnostic of substrate specificity. Asterisks indicate conserved residues.

KR-ScaP1-M2	HTAGVVGEARPL	SSGAGVWNGGQGPYAAANA	Type A1	2R, 3S
KR-ScaP4-M10	HAAGAV-DSVSL	SSIAAAANGSGGQAAYAAGNA	Type A1	2R, 3S
KR-ScaP5-M12	HAAGVV-DSVPL	SSIAGVWGSGGQAAYAAGNA	Type A1	2R, 3S
KR-ScaP6-M16	HAAGVV-DSVPL	SSIAGVWGSGGQAAYAAGNA	Type A1	2R, 3S
KR-ScaP1-M1	HAAGVSALGSL	SSISGTWGVVAE[GAYAAANA	Type A2	2S, 3S
KR-ScaP2-M3	HTAGVGLLVPL	SSVAGTWGSGD[GAYAASNS	Type A2	2S, 3S
KR-ScaP2-M4	HTAGVLDDGVI	SSMAGTLGGPGQGSYAAANA	Type B1	2R, 3R
KR-ScaP2-M5	HTAGVLDDGVL	SSLAGTFGGVGQGNYAAANA	Type B1	2R, 3R
KR-ScaP3-M6	HAAGVLDDGVI	SSAAATFGGPQGNYAAAGNA	Type B1	2R, 3R
KR-ScaP3-M7	HTAGVLDDGML	SSTAGVFGAPGQANYAAANA	Type B1	2R, 3R
KR-ScaP4-M8	HTAGVLDDGTL	SSIAGTLGAAGQANYAAANA	Type B1	2R, 3R
KR-ScaP4-M9	HTAGVLDDGML	SSVAGTFFGSAGQANYAAANA	Type B1	2R, 3R
KR-ScaP5-M11	HAAAVLDDGVL	SSFAGS1GAAGQANYAAANA	Type B1	2R, 3R
KR-ScaP5-M15	HTAGVLDDGIL	SSVAGTFFGSAGQANYAAANA	Type B1	2R, 3R
KR-ScaP6-M17	HTAGVLDDGVL	SSVAGTFGGSGQGNYAAANA	Type B1	2R, 3R
KR-ScaP6-M18	HAAGVLDDGML	SSFAGTAGATGQANYAAANA	Type B1	2R, 3R
KR-ScaP7-M20	HLAGVVLDDGVL	SSAAGTFFGSAGQANYAAANA	Type B1	2R, 3R
KR-ScaP5-M14	HVAGVLDDGVI	SSTSGTISGPGLNYAPGNA	Type B2	2S, 3R
KR-ScaP7-M19	HTAAVLDDGTL	SSTAGTIGAPGHANYAPGNA	Type B2	2S, 3R
KR-ScaP8-M22	HTAVVTDDGIV	SSIAGVLGGAGQGNYAAANV	Type B1	2R, 3R

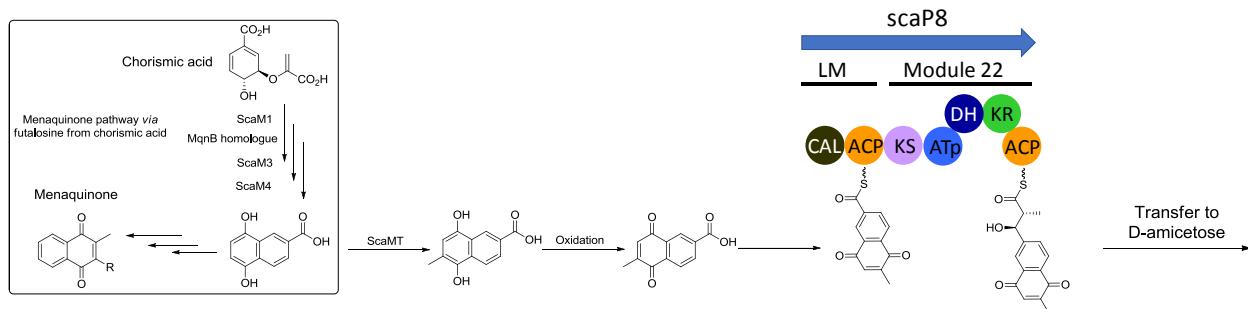
**Figure S3.** Sequence alignment of the loop and catalytic regions within the KR domains highlighting the aminoacid residues which are diagnostic of the stereochemical outcome of the keto-reduction.

	* * *
DH-ScaP2-M5	PWLADHALDGTLFPGTGFLELALQAR
DH-ScaP3-M6	PWLADRAVDGRVPLPTSAFLDLALHLG
DH-ScaP3-M7	PWLADHALAGTAVVPGTAFVELAVQAG
DH-ScaP4-M9	PWLADHVVDGSLLVPNTAFLLELALRAA
DH-ScaP5-M13	PWLAEHRIAGAIIVPSTALLELAVRAG
DH-ScaP5-M15	PWLGDHAVTGTVLFPGTGFLELAVRAA
DH-ScaP6-M17	PWLAEHCVSGSVLLPGTAFLELAVRAG
DH-ScaP7-M20	AWLADHAVLGRVILPATAYLDSLAVSAG
DH-ScaP8-M22	DWLTDHRIAGSAVLPGPAPFVELALRAG

**Figure S4.** Sequence alignment of the HxxxGxxxxP-containing region within the DH domains. The asterisks indicate the position of the HxxxGxxxxP motif. The mutation in this motif found in the DH domain from module M6 is highlighted in yellow.

	*	-----
ER-ScaP3-M7	FRDVNLVLMYPGDAGAFGL..69aa..LGGVQAGESVLVHAGAGGVGMAAVQLARHLG	
ER-ScaP4-M9	FRDVNLVLTYPGDAGELGL..69aa..LGGVQAGESVLVHAGAGGVGMAAVQLARHLG	
ER-ScaP5-M15	FRDVNLNALGMYPGEAGALGS..69aa..LGEVQAGESVLVHSAAGGVGMAAVQLARHLG	
ER-ScaP6-M17	FRDVNLNALGMYPGDAVALGI..69aa..LGELQAGESVLVHAAAGGVGMAAVQLAWHLG	

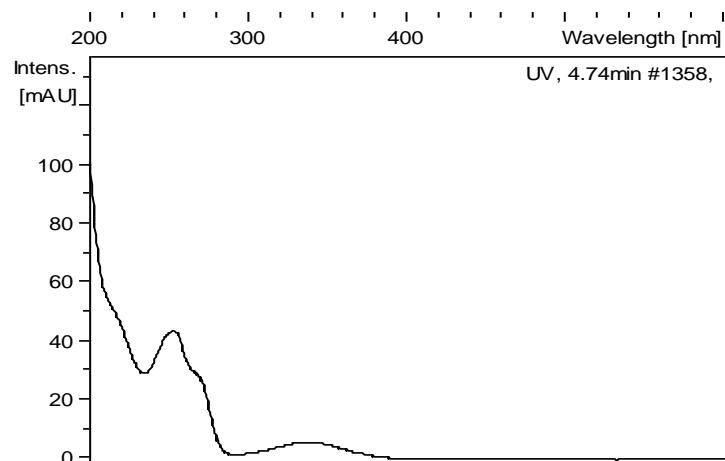
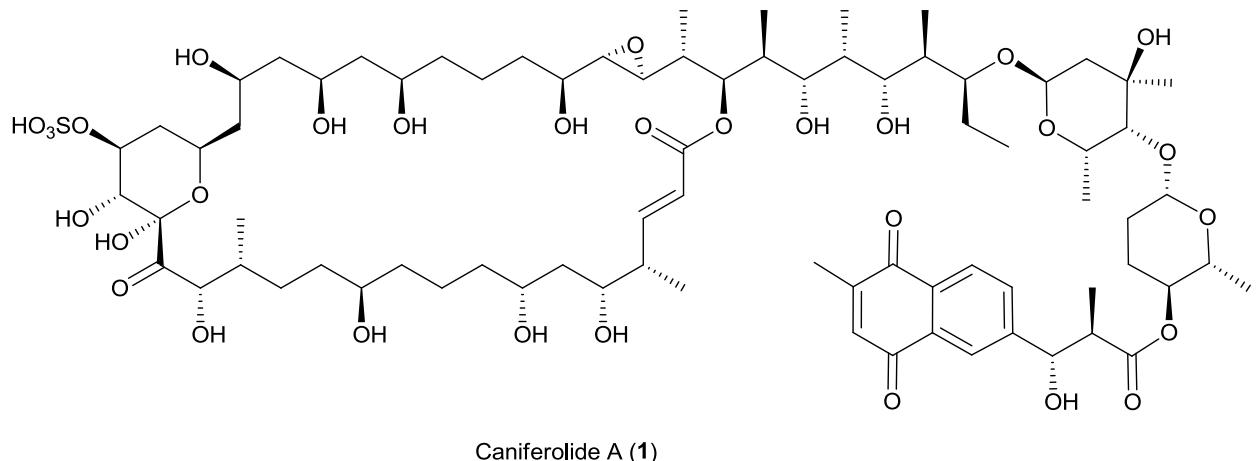
**Figure S5.** Sequence alignment of the ER domain. The asterisk indicates the position of the conserved Y residue conferring (2*S*) stereochemistry. The hyphens indicate the position of the NADPH biding site.



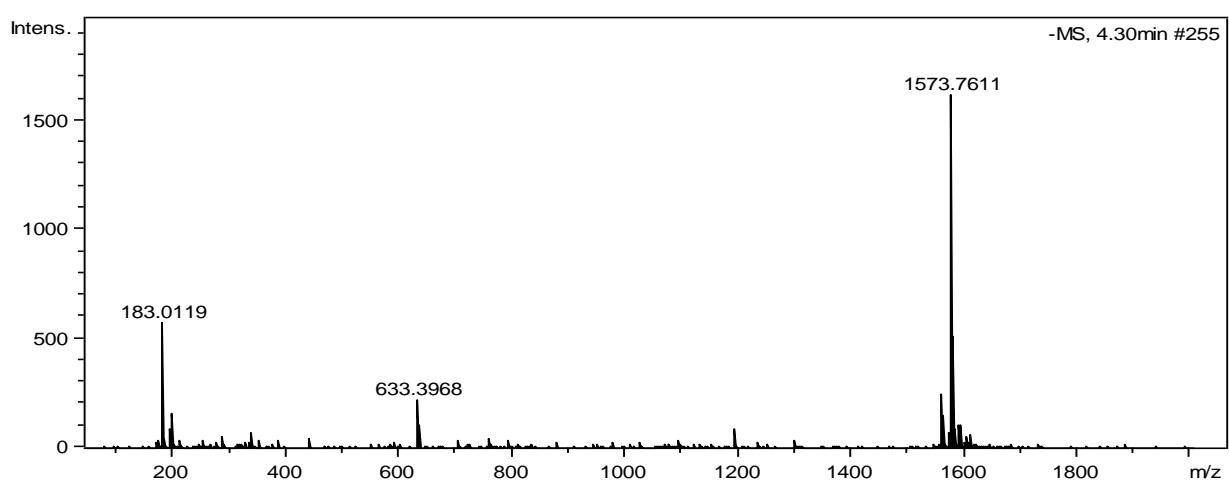
**Figure S6.** Proposed biosynthetic pathway leading to the alkynaphthoquinone moiety.

**Table S2.** Predicted stereochemical outcome (according to the type of ketoreductase in each module) compared to the stereochemical assignment obtained by NMR. To facilitate the comparison, the macrocyclization and post-PKS tayloring steps are not taken into account to avoid the changes that such modifications would produce on the Cahn-Ingold Prelog stereochemical descriptors of the last column. See also Fig. 7 in the main text.

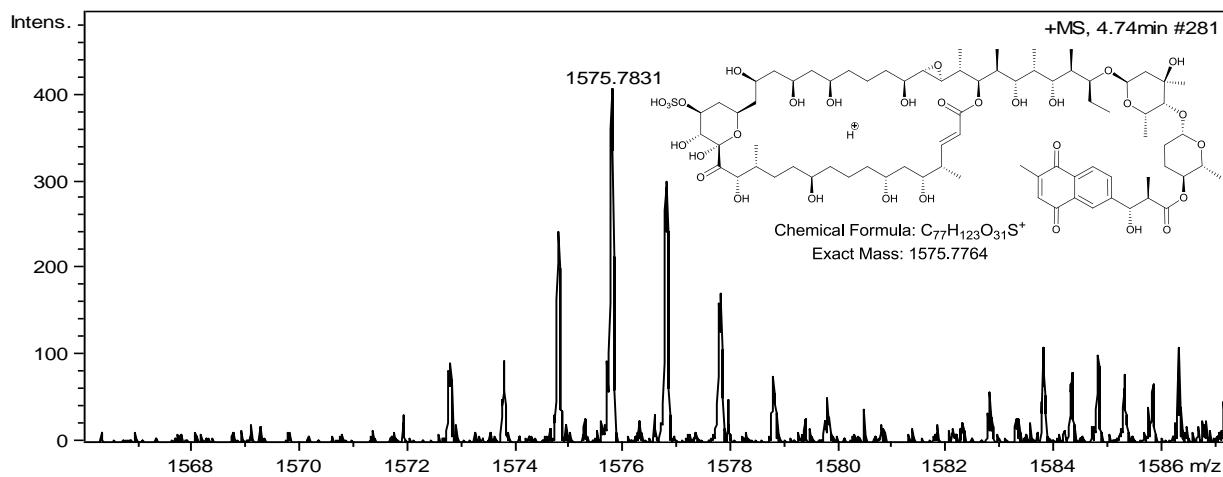
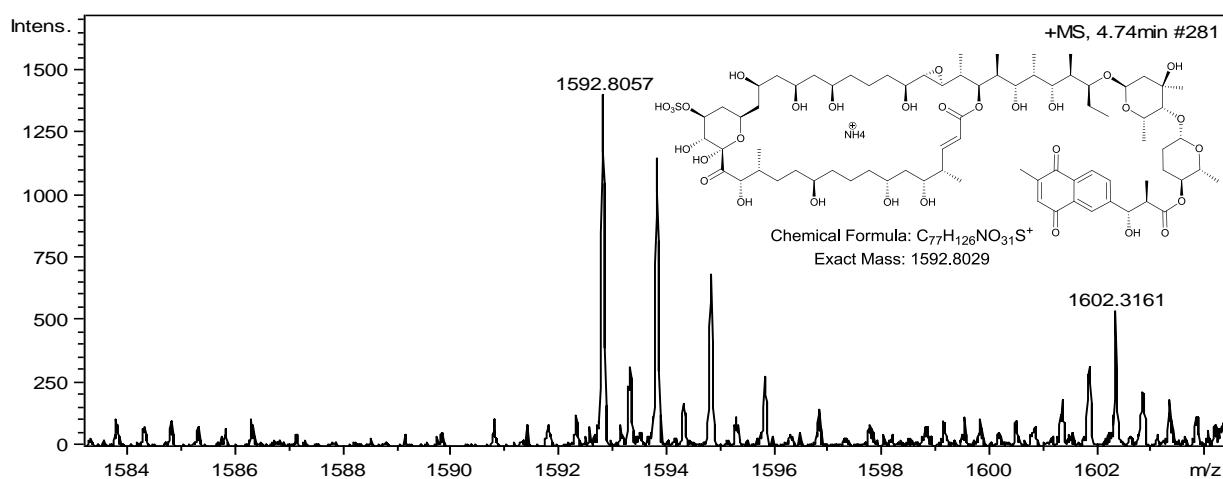
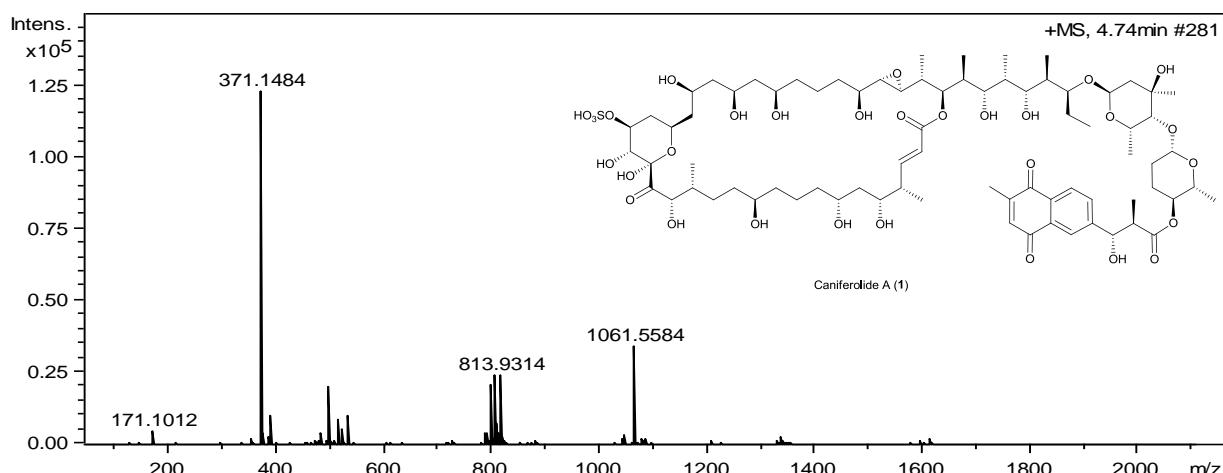
Module	Type of KR	Predicted by bioinformatic analysis		Spectroscopic data
		Keatinge-Clay "RS" system	Cahn-Ingold-Prelog RS system	Cahn-Ingold-Prelog RS system
Module 1	A2	"S,S" (C-40,41)	S,S (C-40,41)	S,S (C-40,41)
Module 2	A1	"R,S" (C-38,39)	R,S (C-38,39)	R,S (C-38,39)
Module 3	A2	"S,S" (C-36,37)	R,R (C-36,37)	R,R (C-36,37)
Module 4	B1	"R,R" (C-34,35)	S,S (C-34,35)	S,S (C-34,35)
Module 5	B1	E double bond	E double bond	trans epoxide
Module 6	B1	"R" (C-31)	S (C-31)	S (C-31)
Module 8	B1	"R" (C-27)	S (C-27)	S (C-27)
Module 10	A1	"S" (C-23)	S (C-23)	S (C-23)
Module 11	B1	"R" (C-21)	R (C-21)	R (C-21)
Module 12	A1	"S" (C-19)	S (C-19)	S (C-19)
Module 14	B2	"S,R" (C-14,15)	R,R (C-14,15)	R,R (C-14,15)
Module 16	A1	"S" (C-11)	R (C-11)	R (C-11)
Module 18	B1	"R" (C-7)	R (C-7)	R (C-7)
Module 19	B2	"S,R" (C-4,5)	R,R (C-4,5)	S,R (C-4,5)
Module 20	B1	E double bond	E double bond	E double bond
Module 22	B1	"R,R" (C-2'''',3''')	R,S (C-2'''',3''')	R,S (C-2'''',3''')



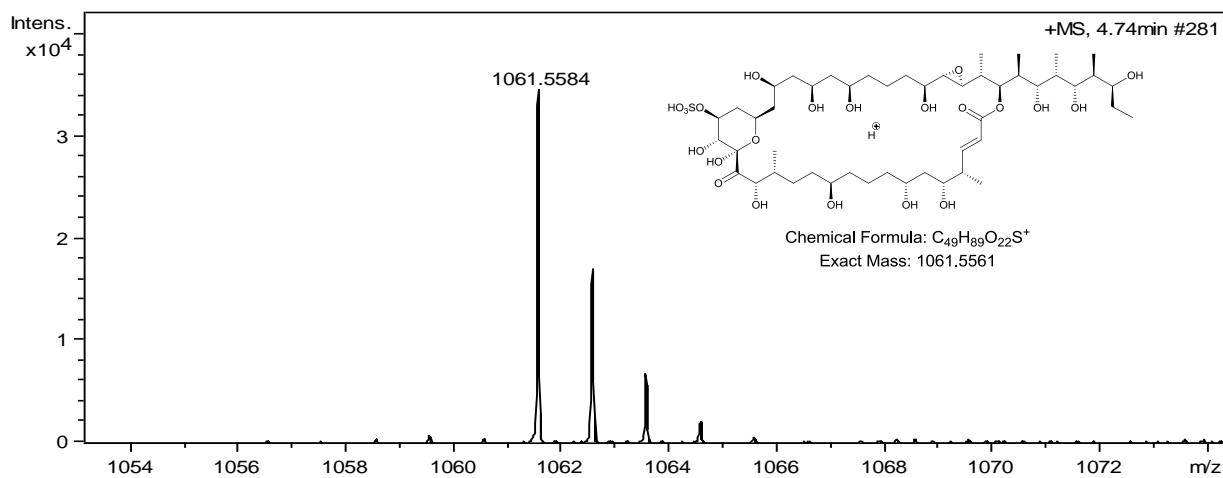
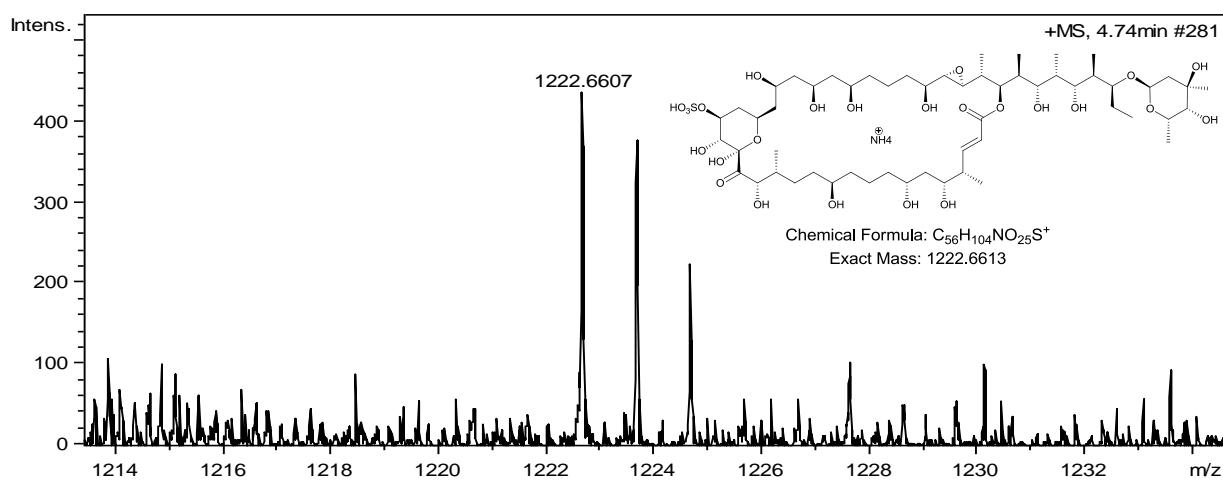
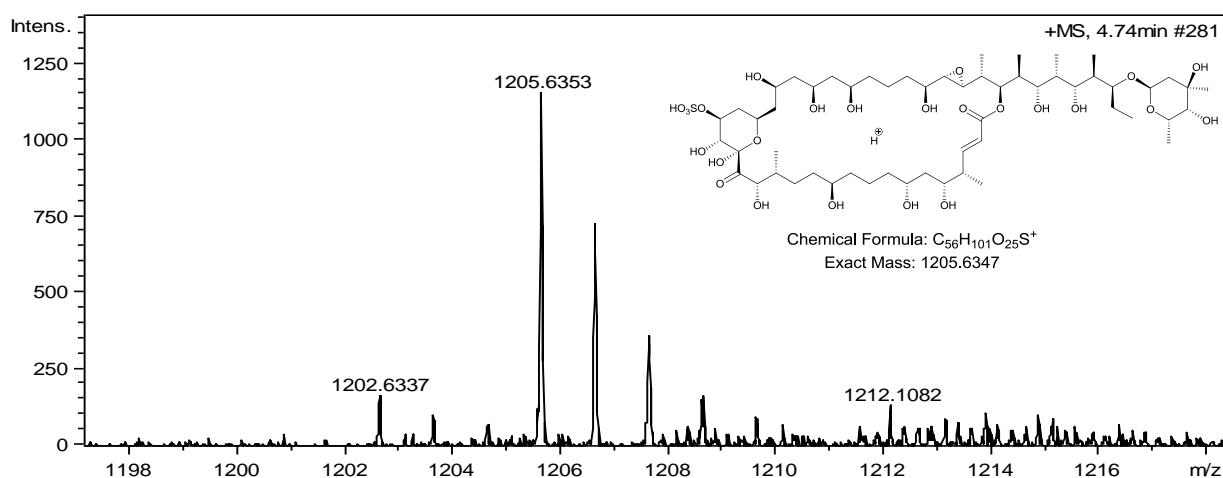
**Figure S7.** UV (DAD) spectrum of caniferolide A (1).



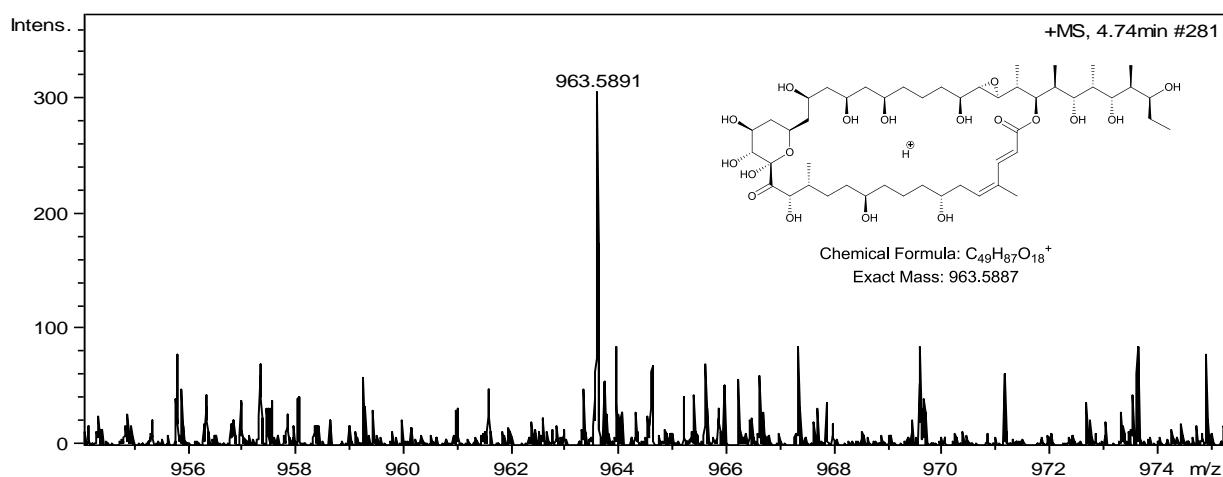
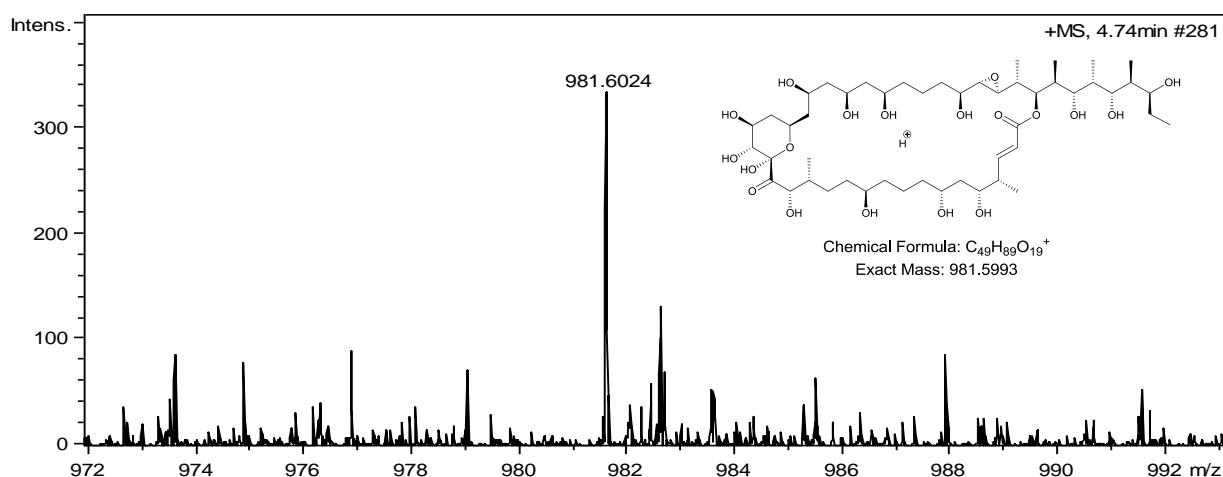
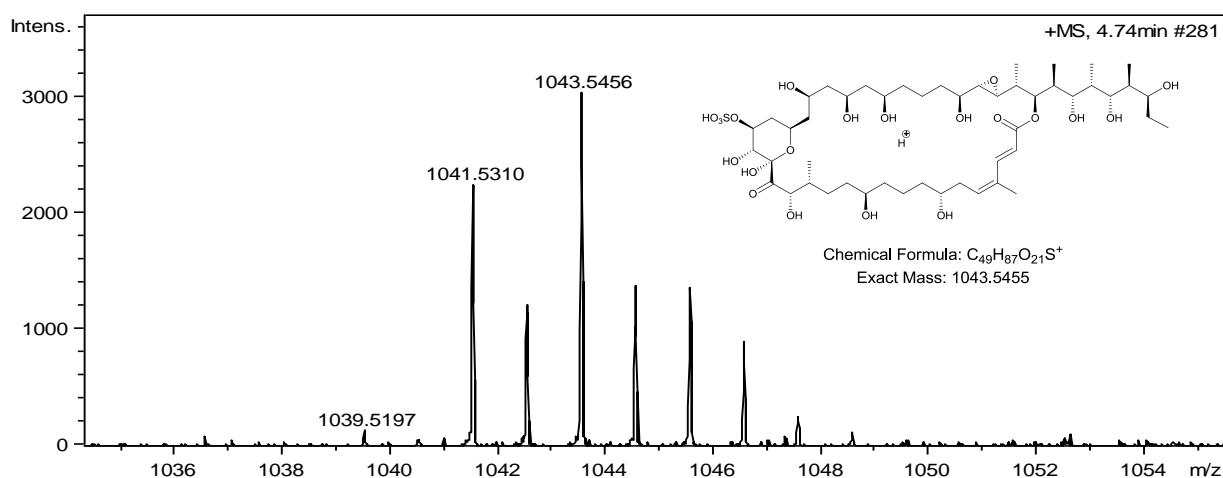
**Figure S8.** (-) ESI-TOF MS of caniferolide A (1).



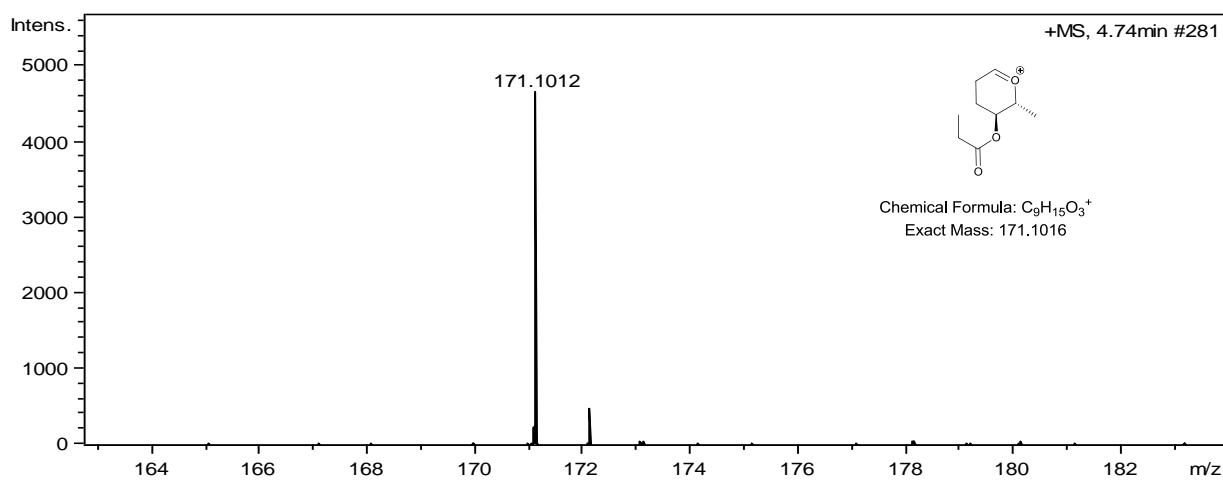
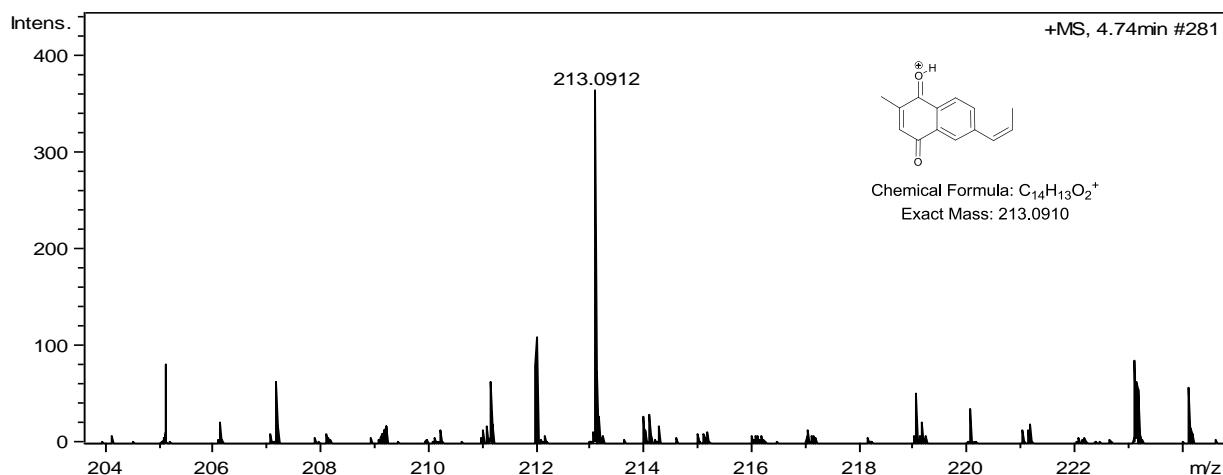
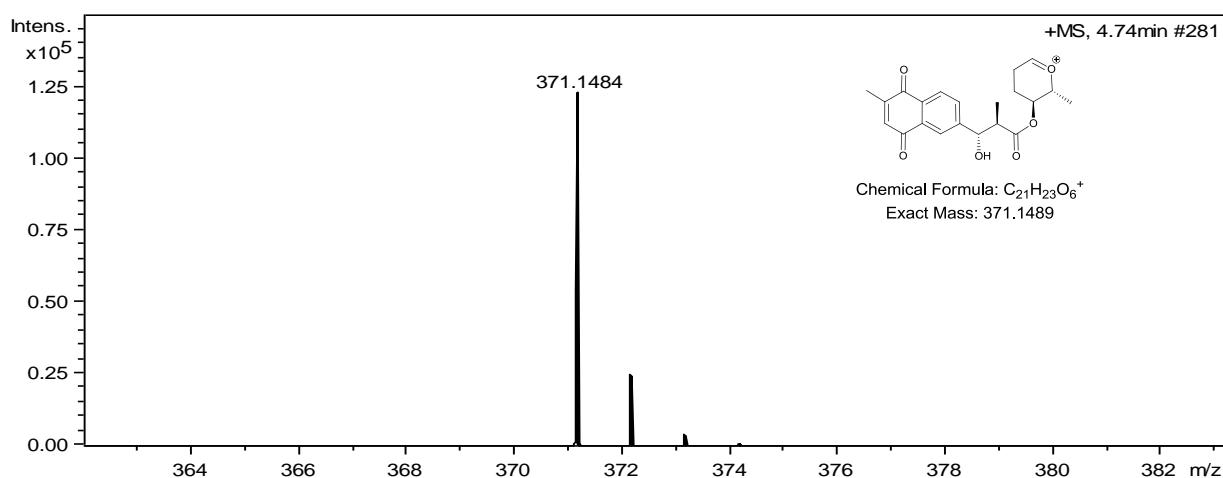
**Figure S9.** (+) ESI-TOF MS of caniferolide A (**1**) including expansions showing key in-source fragment ions.



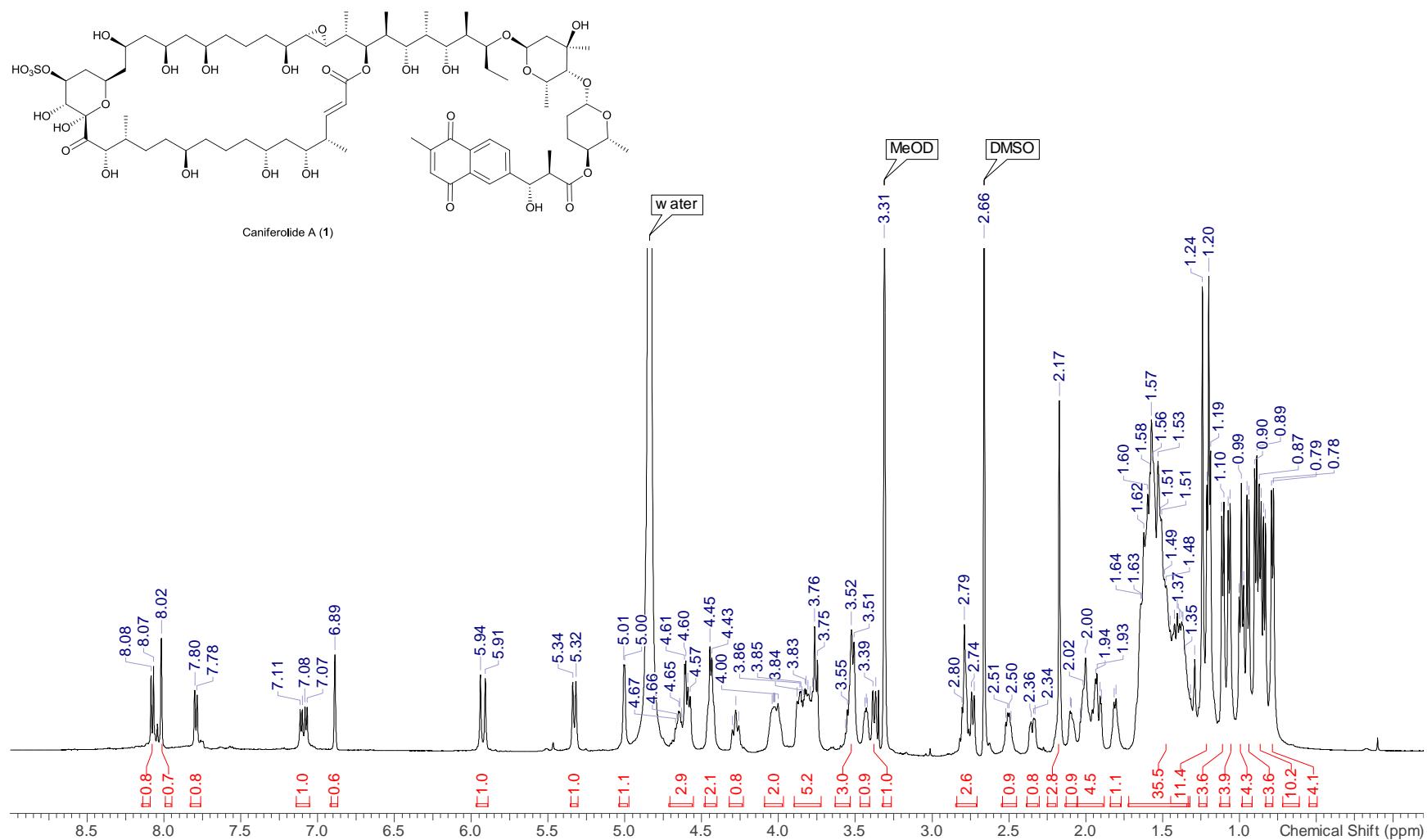
**Figure S9 cont.** (+) ESI-TOF MS of caniferolide A (**1**) including expansions showing key in-source fragment ions.



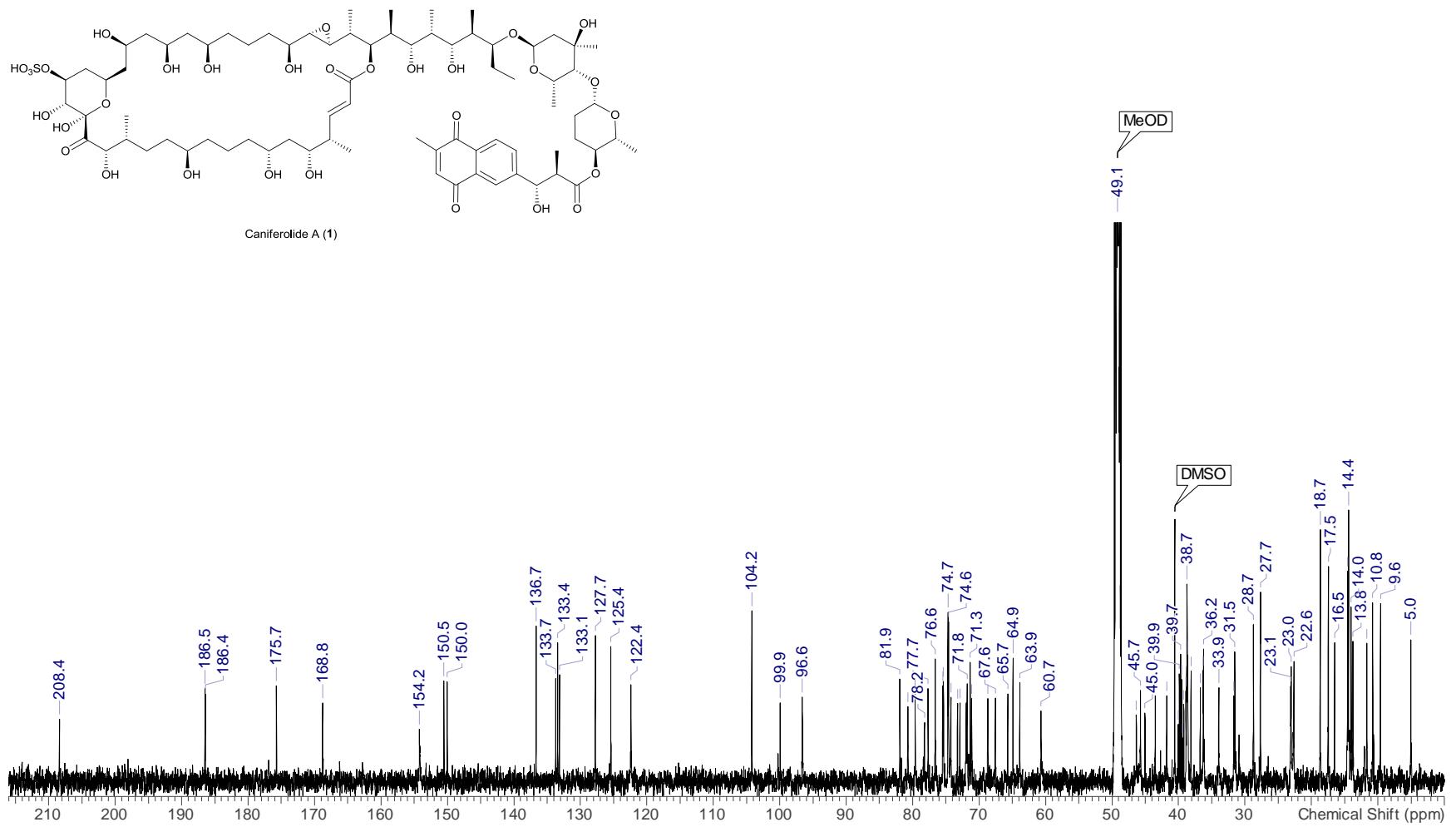
**Figure S9 cont.** (+) ESI-TOF MS of caniferolide A (**1**) including expansions showing key in-source fragment ions.



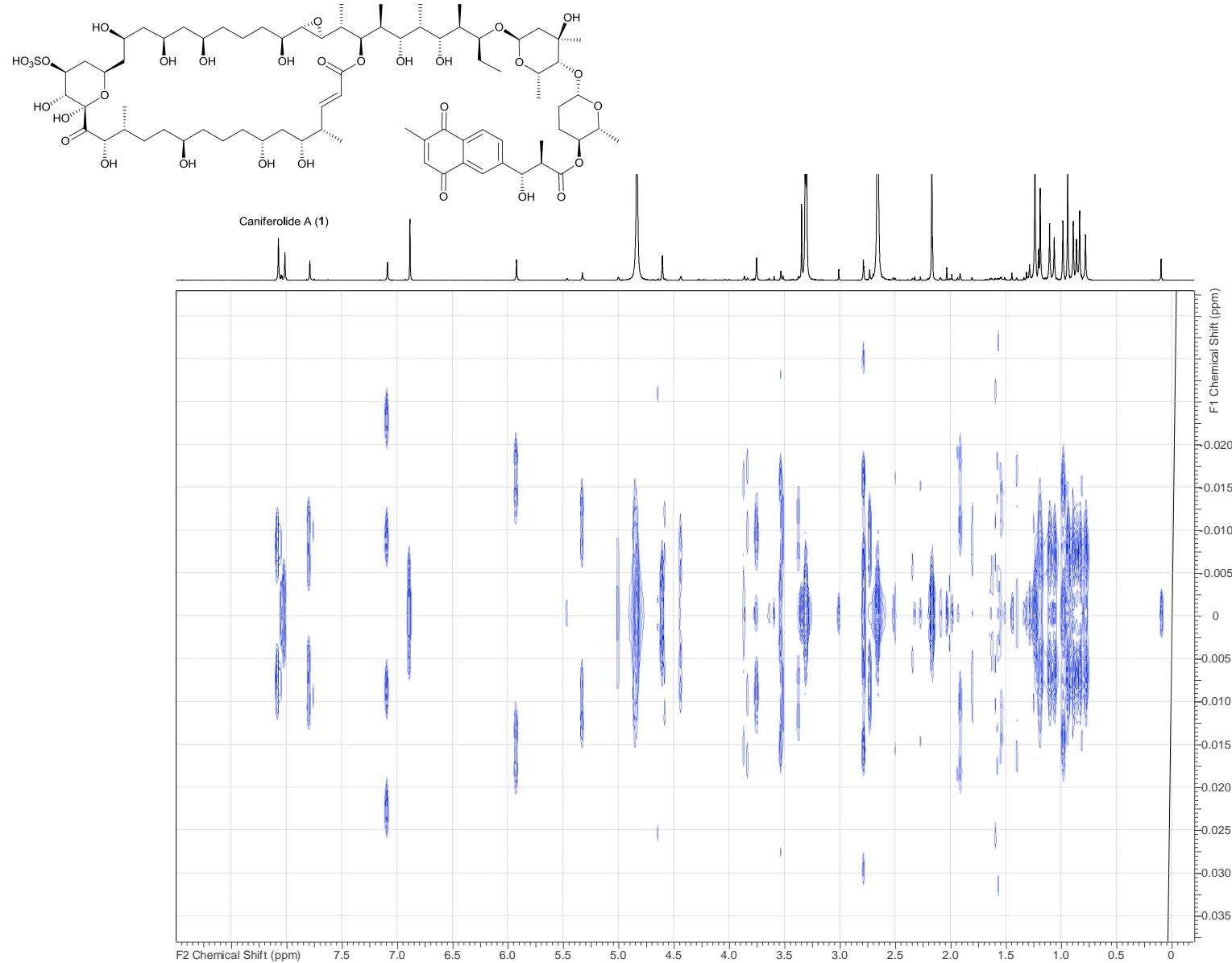
**Figure S9 cont.** (+) ESI-TOF MS of caniferolide A (**1**) including expansions showing key in-source fragment ions.



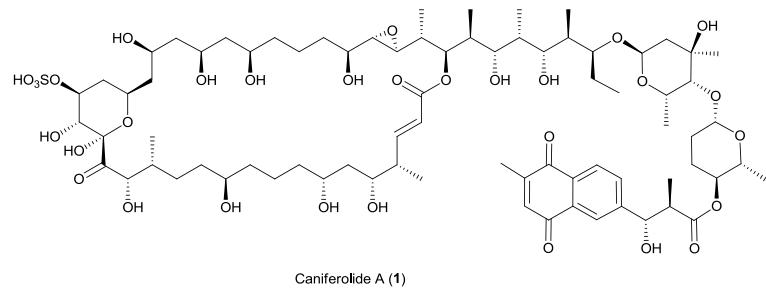
**Figure S10.**  $^1\text{H}$  NMR spectrum ( $\text{CD}_3\text{OD}$ , 500 MHz) of caniferolide A (**1**).



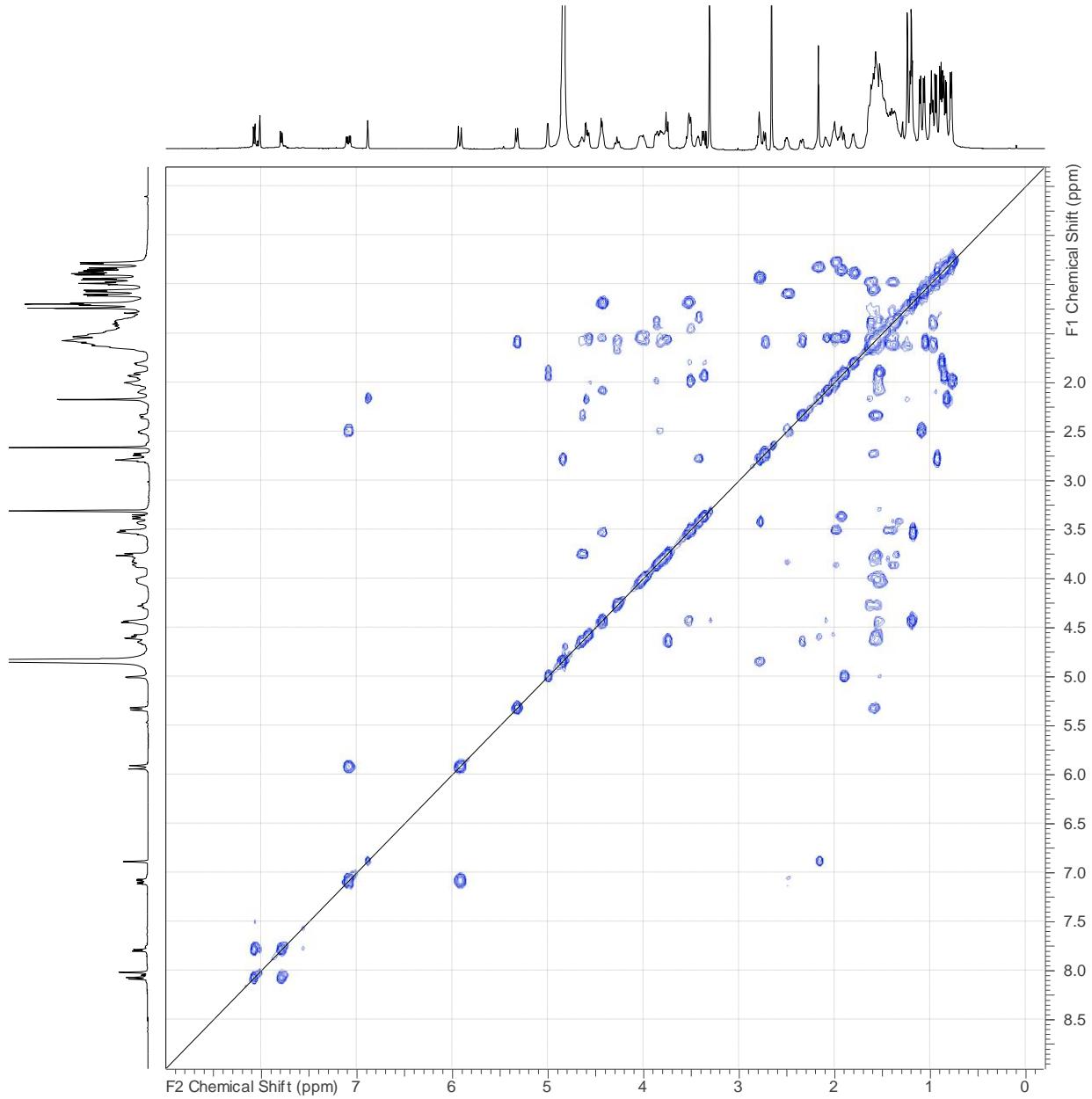
**Figure S11.** <sup>13</sup>C NMR spectrum (CD<sub>3</sub>OD, 125 MHz) of caniferolide A (**1**).



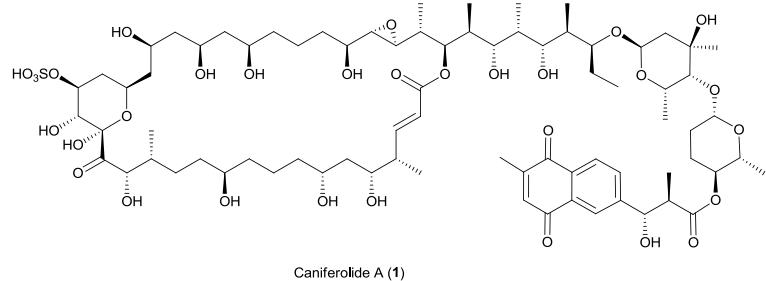
**Figure S12.** JRES spectrum of caniferolide A (**1**).



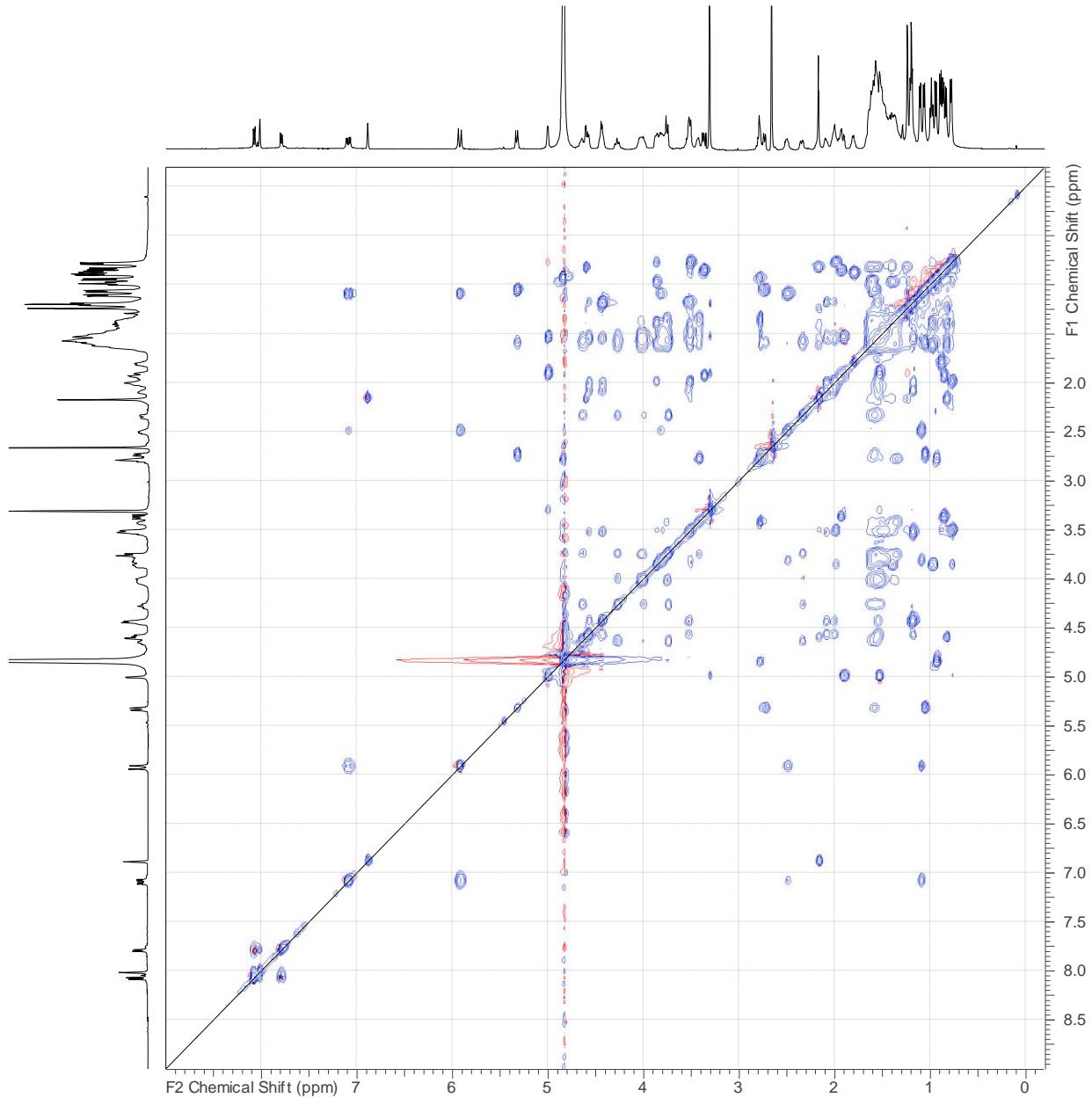
Caniferolide A (1)



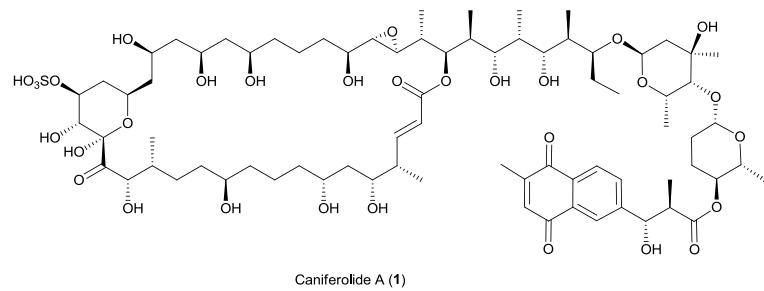
**Figure S13.** COSY spectrum of caniferolide A (1).



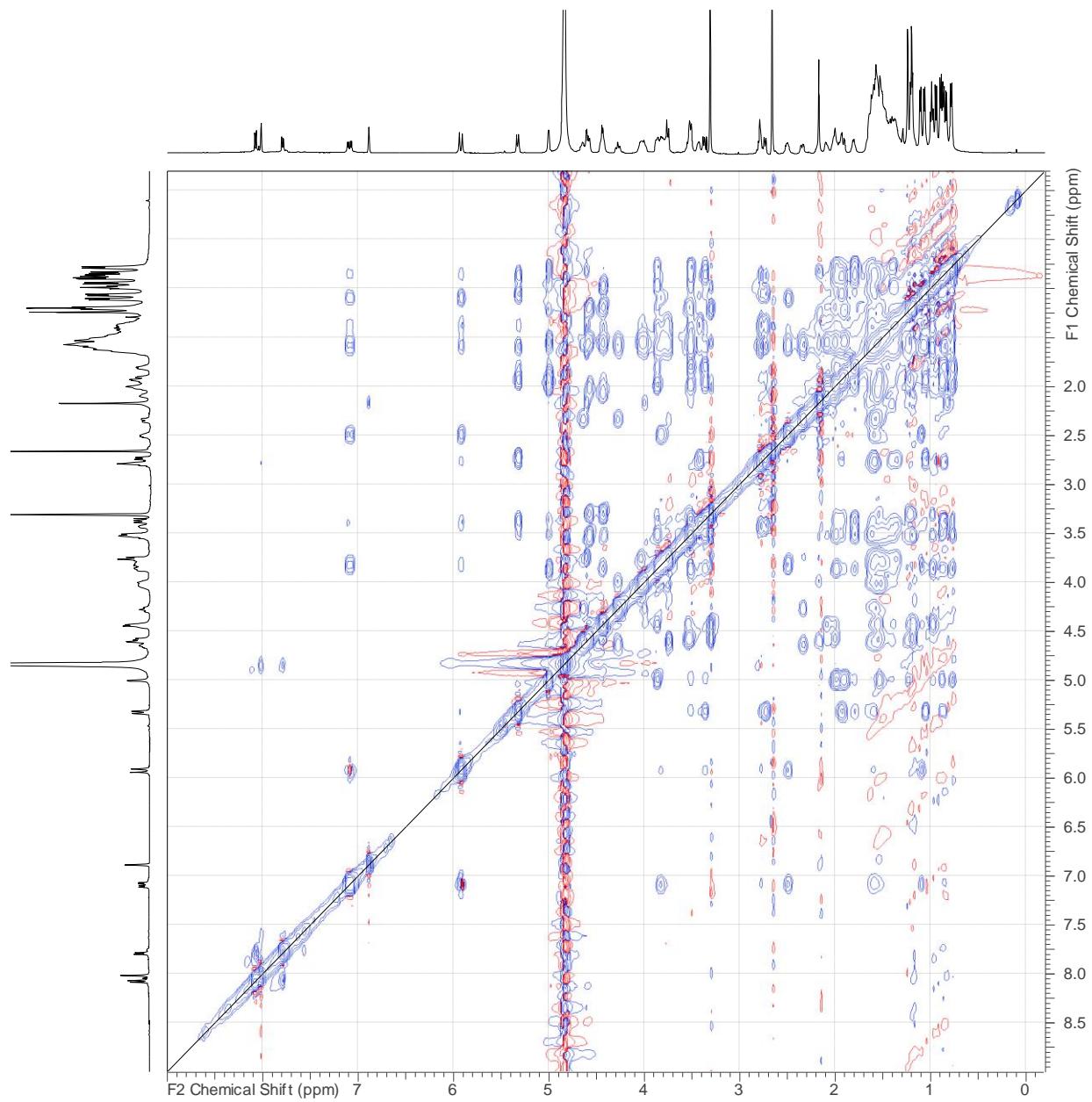
Caniferolide A (**1**)



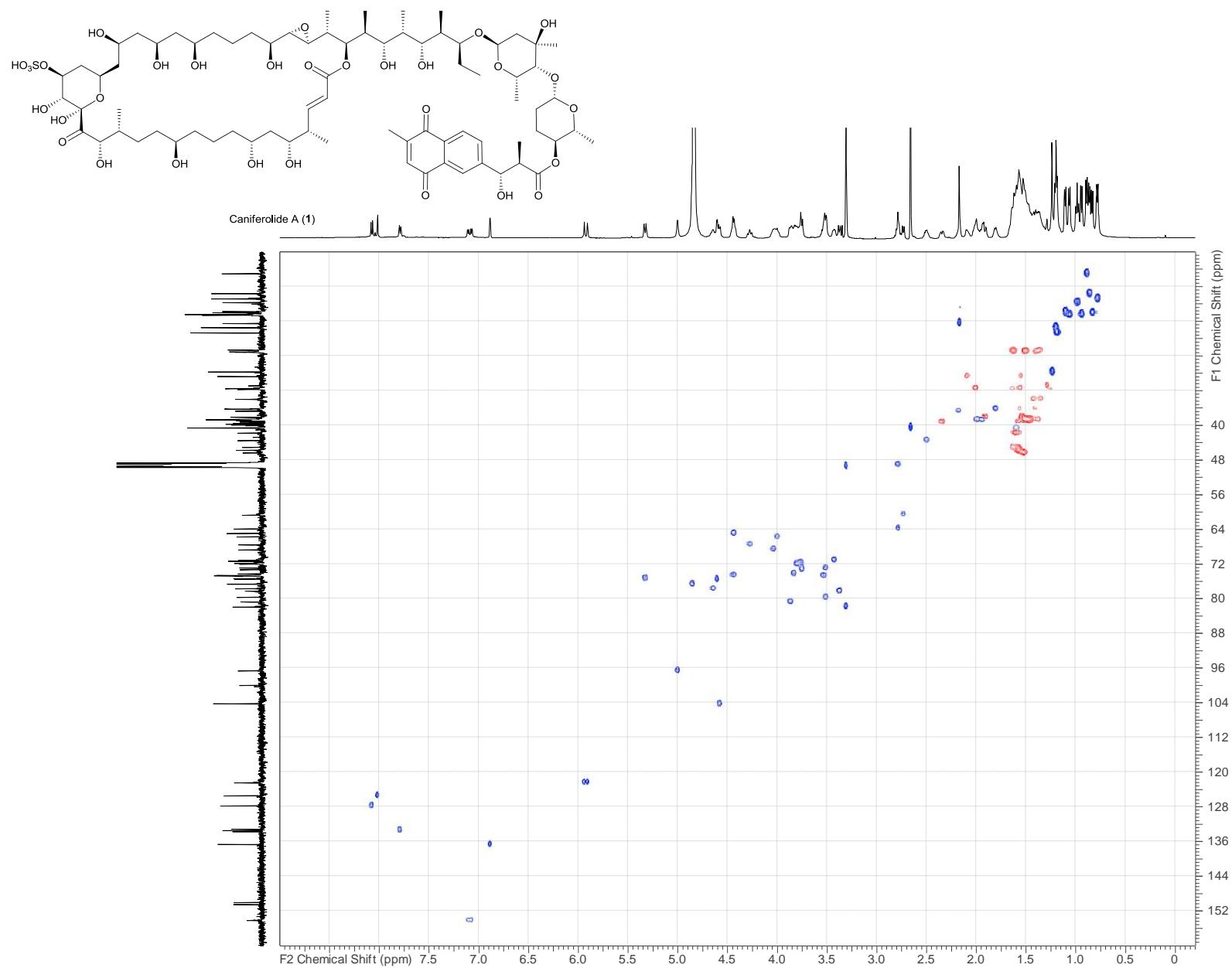
**Figure S14.** TOCSY spectrum of caniferolide A (**1**).



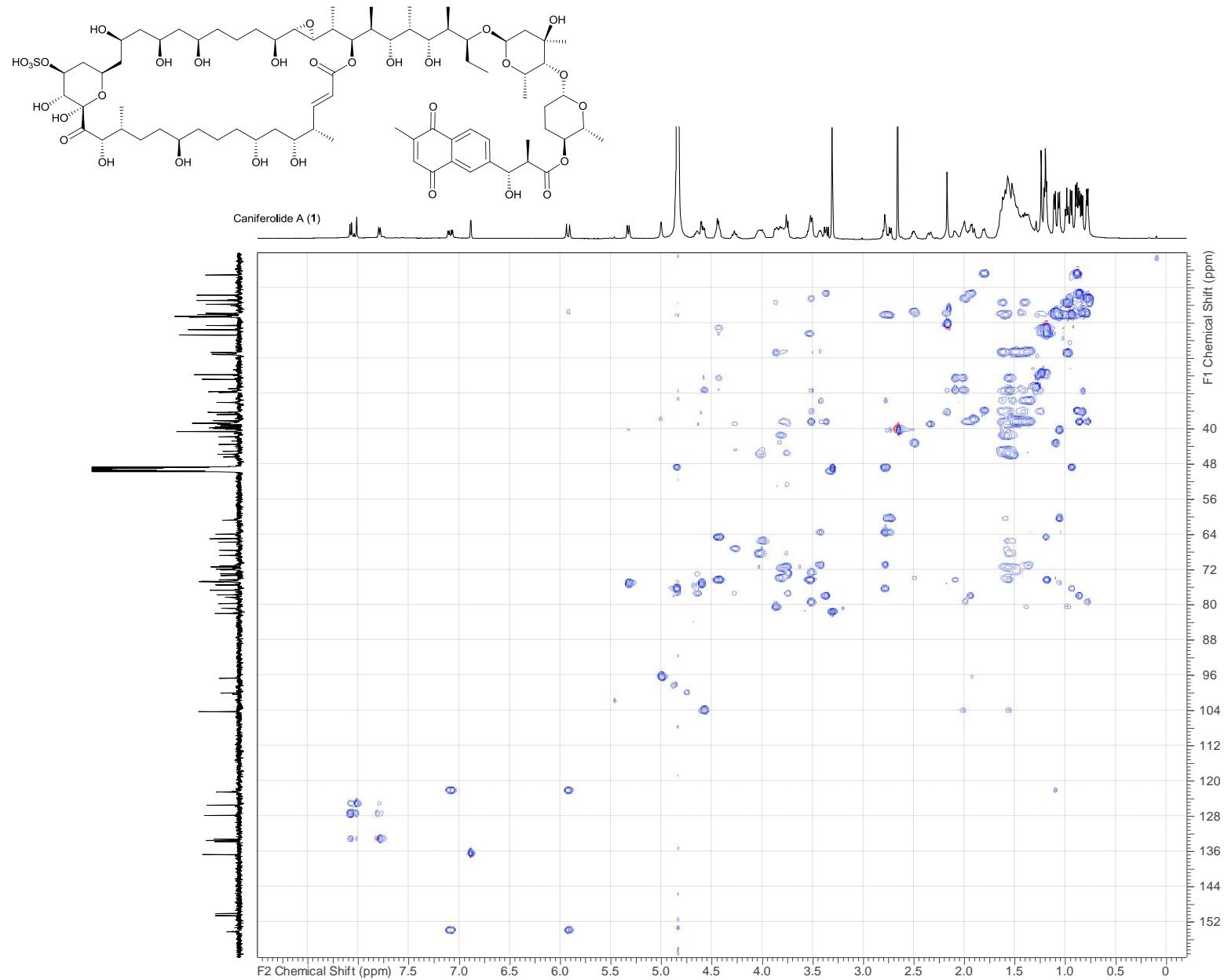
Caniferolide A (1)



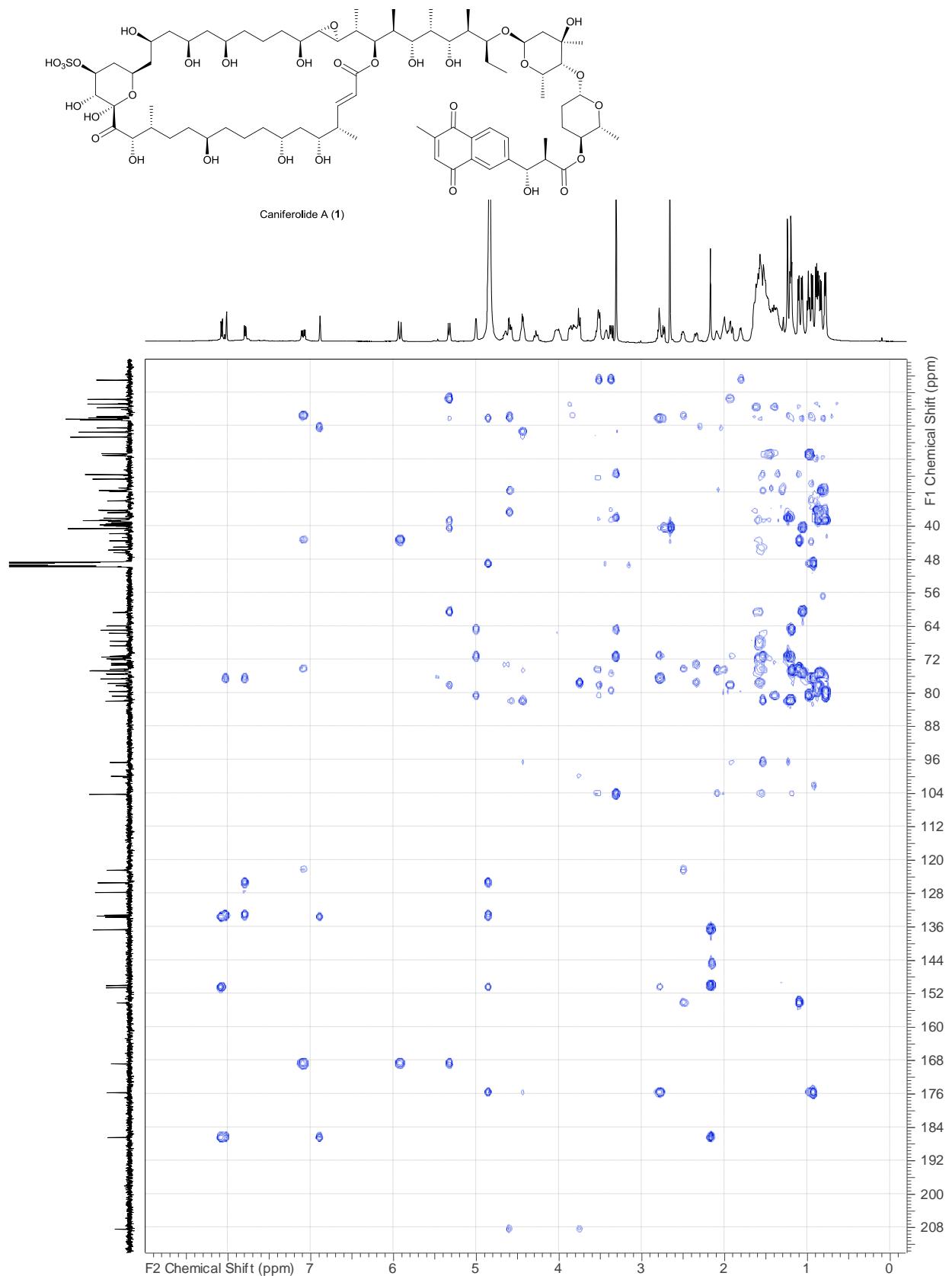
**Figure S15.** NOESY spectrum of caniferolide A (1).



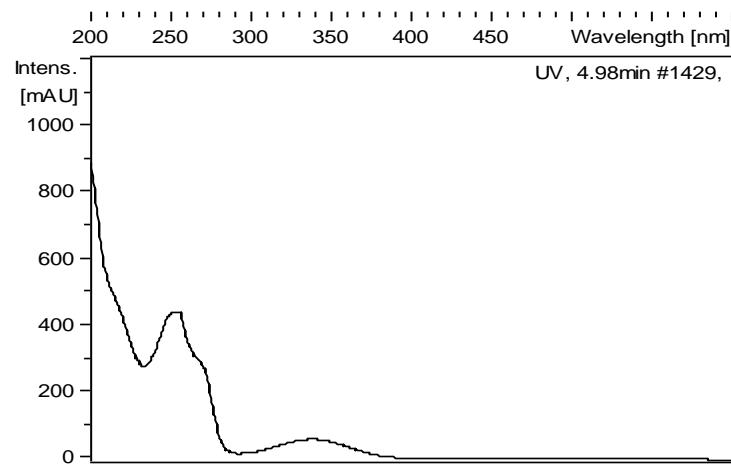
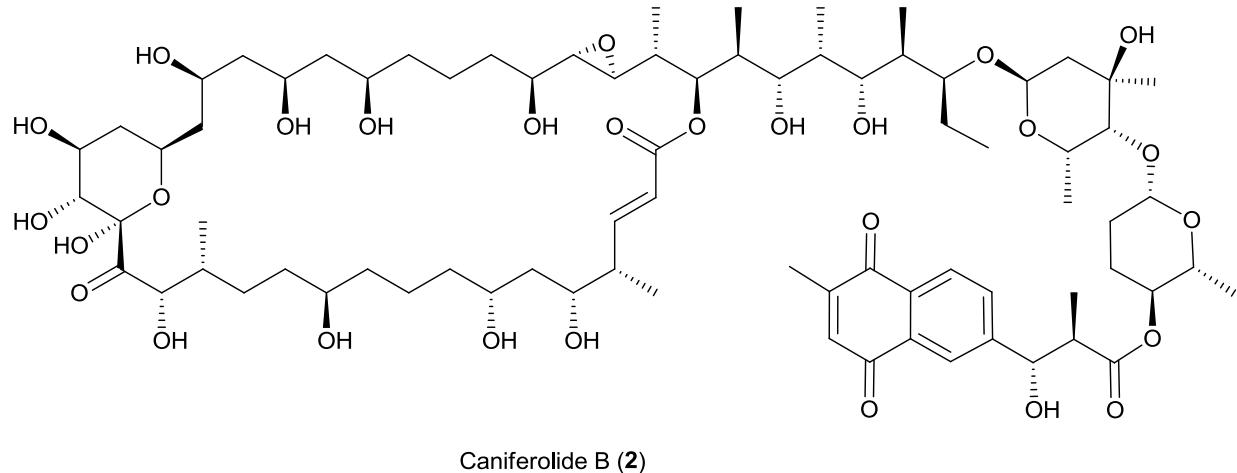
**Figure S16.** Edited HSQC spectrum of caniferolide A (**1**).



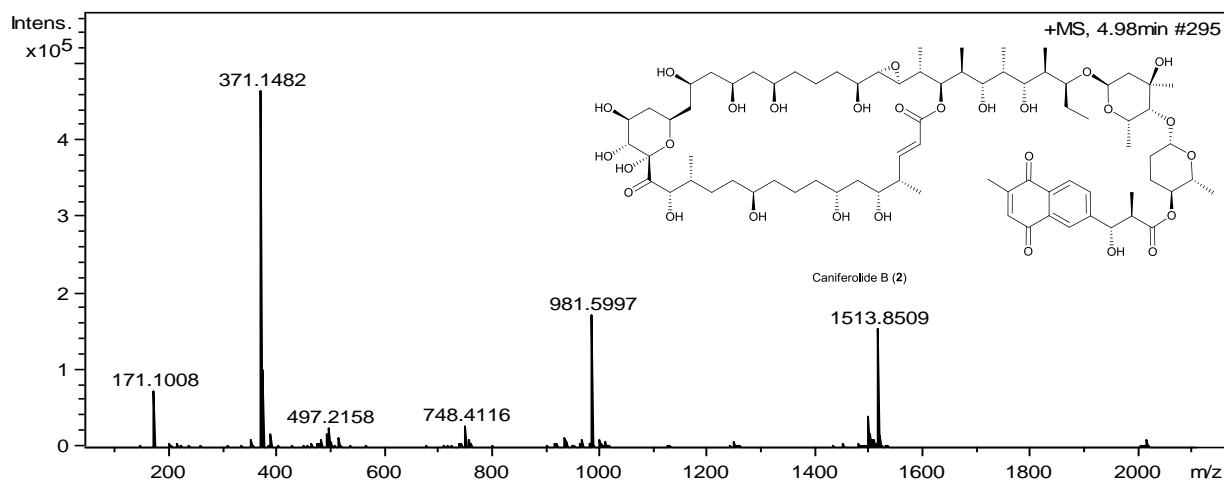
**Figure S17.** HSQC-TOCSY spectrum of caniferolide A (**1**).



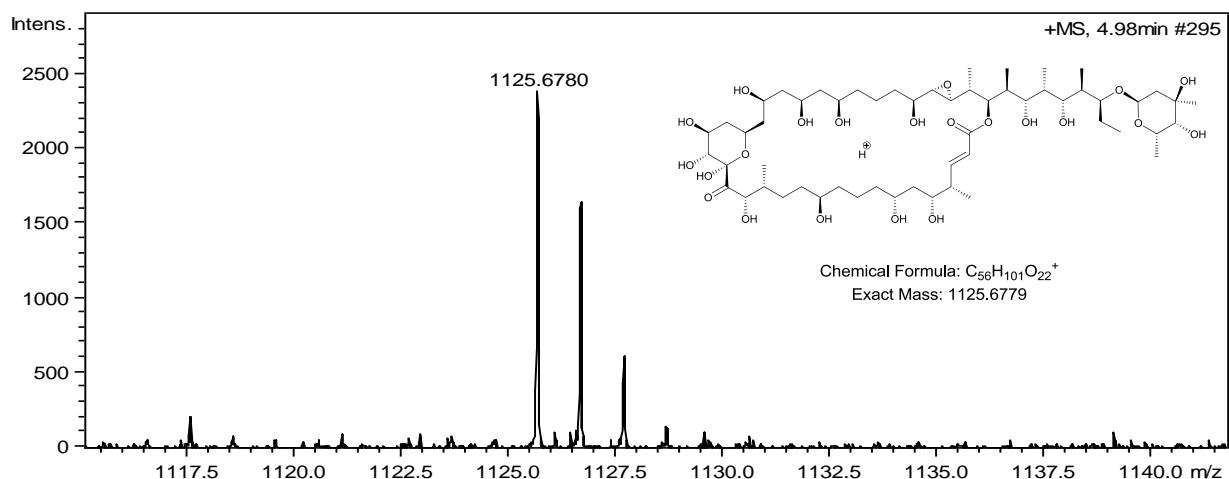
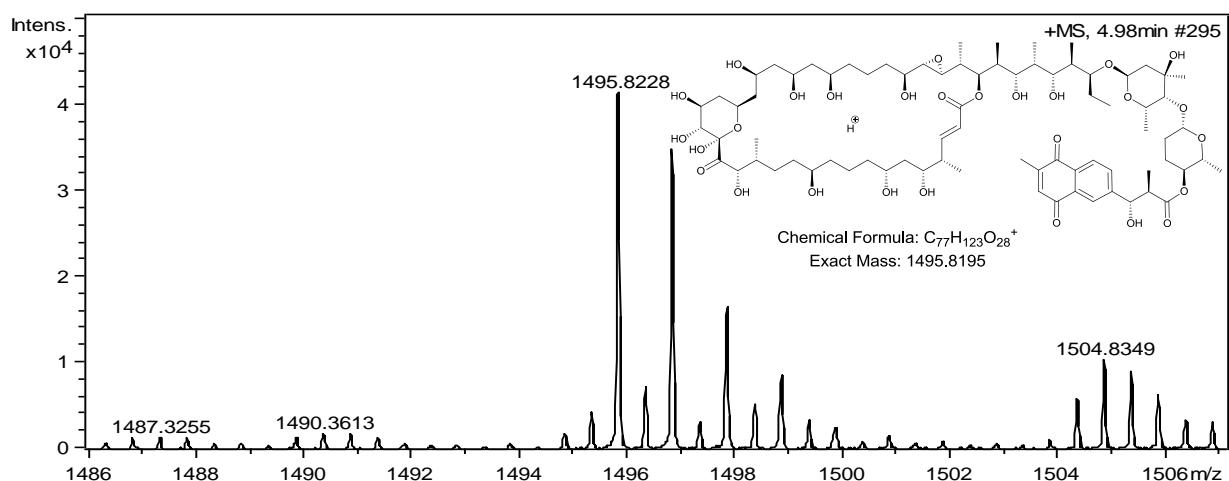
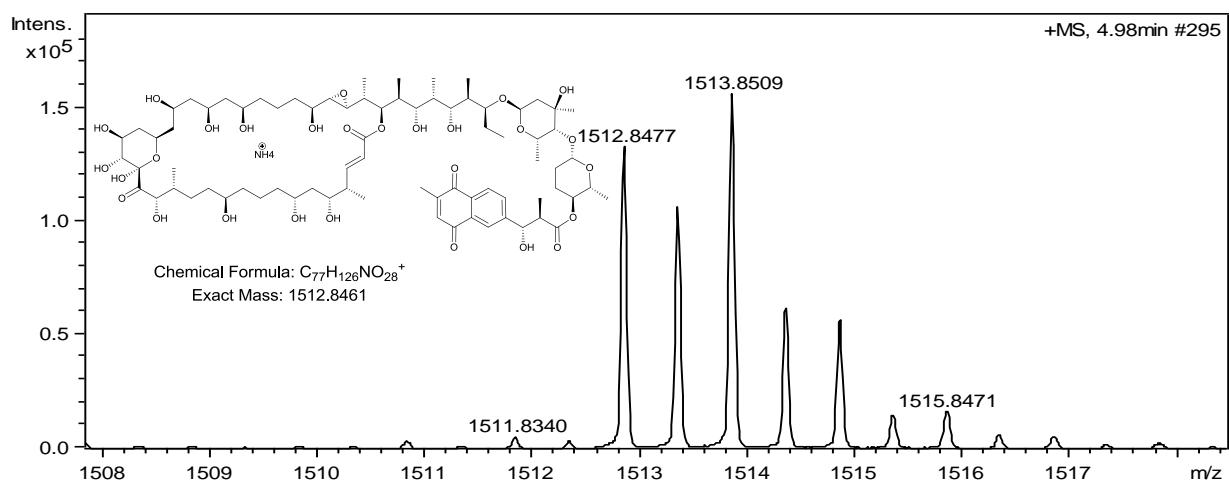
**Figure S18.** HMBC spectrum of caniferolide A (**1**).



**Figure S19.** UV (DAD) spectrum of caniferolide B (2).



**Figure S20.** (+) ESI-TOF MS of caniferolide B (2) including expansions showing key in-source fragment ions.



**Figure S20 cont. (+) ESI-TOF MS of caniferolide B (**2**) including expansions showing key in-source fragment ions.**

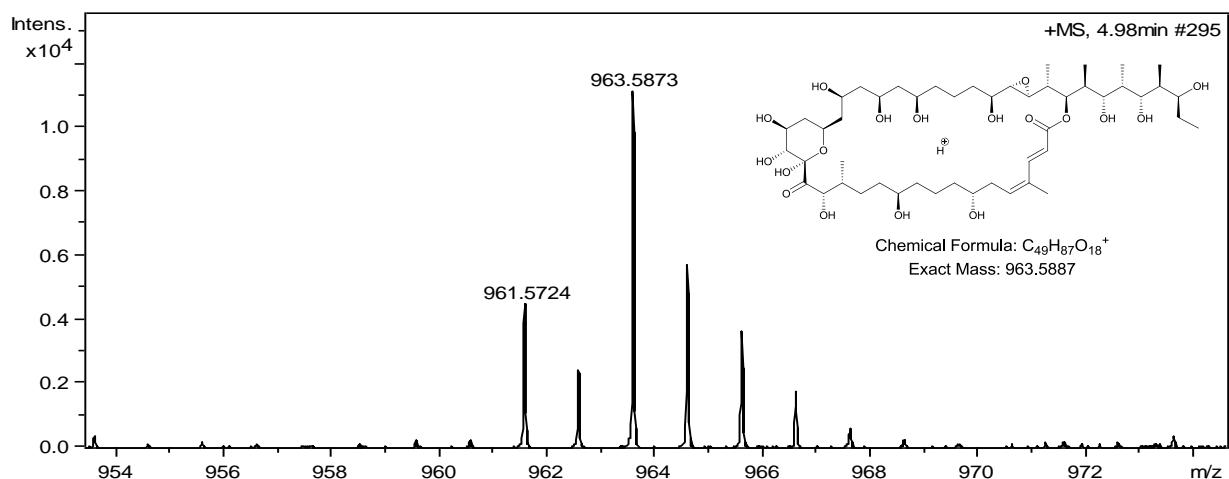
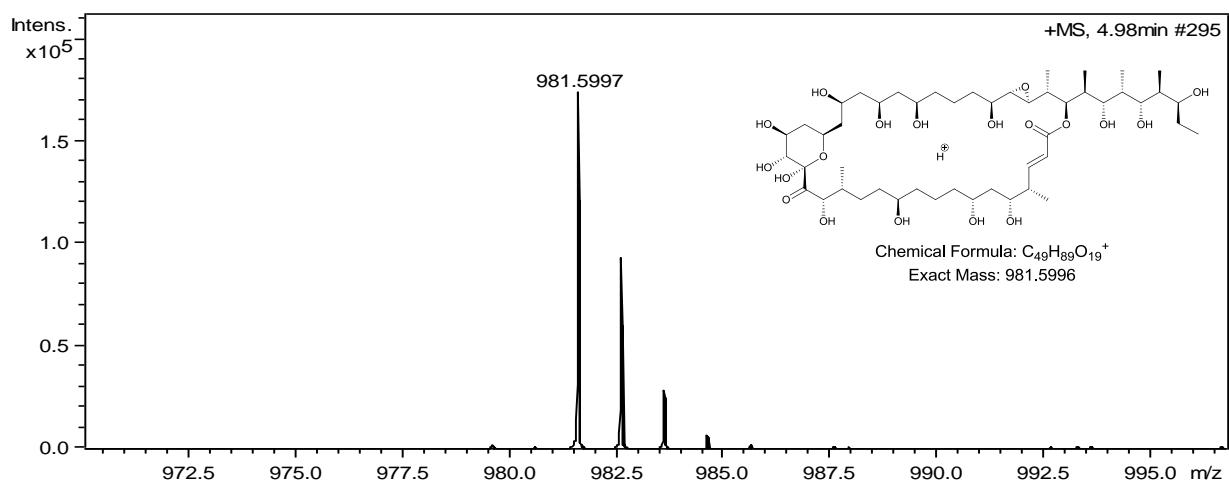
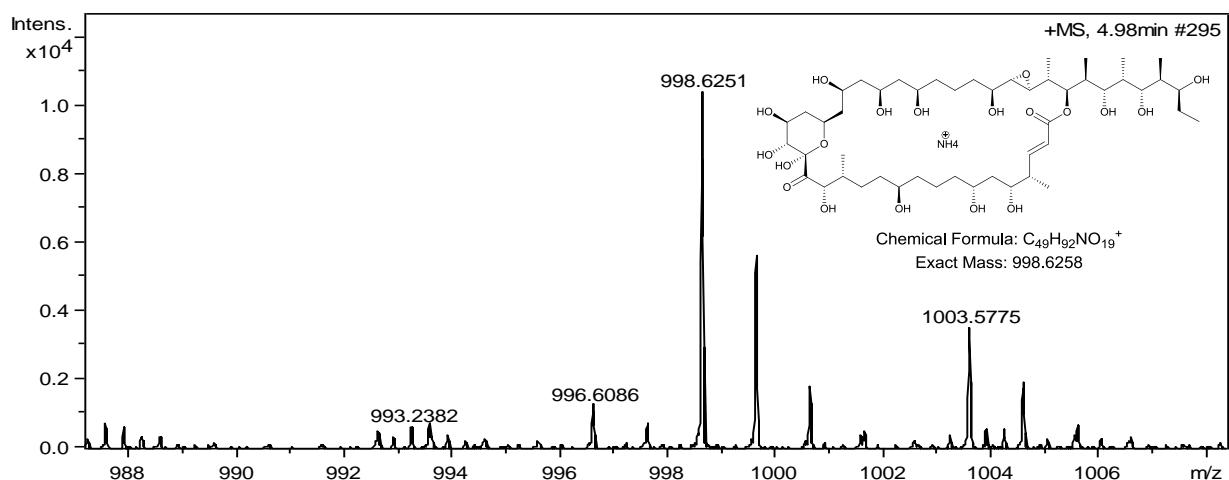
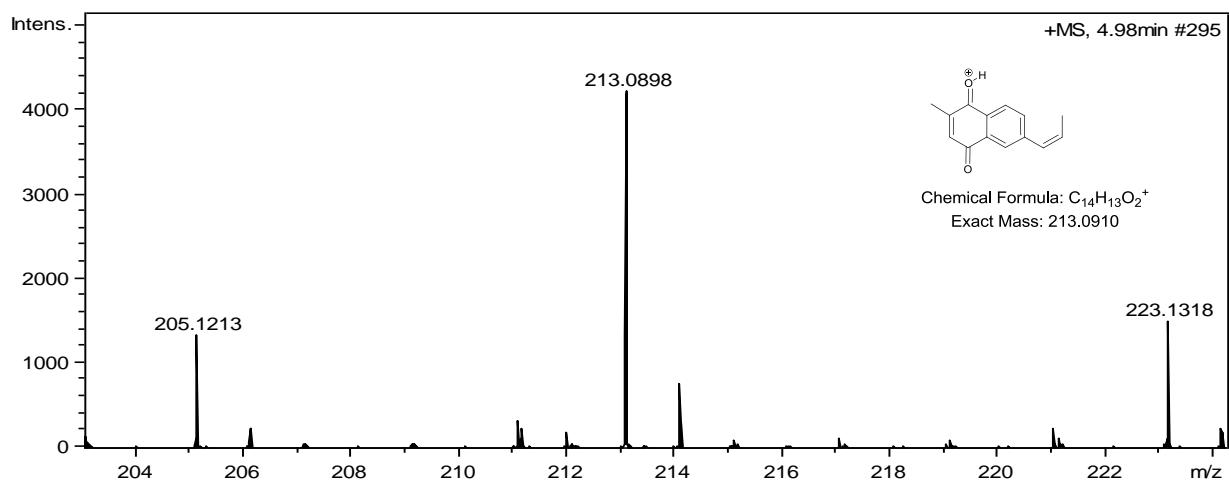
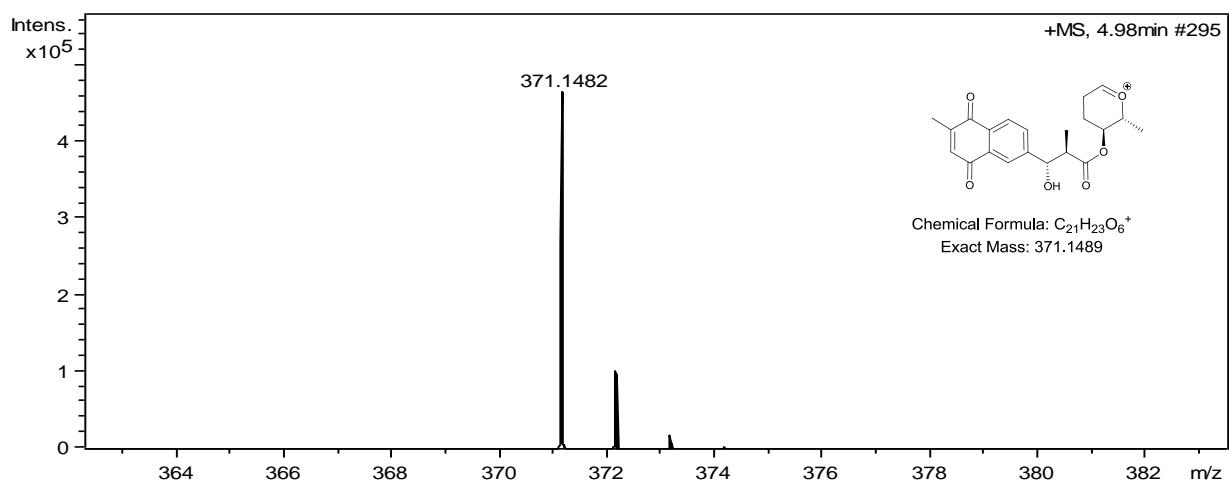
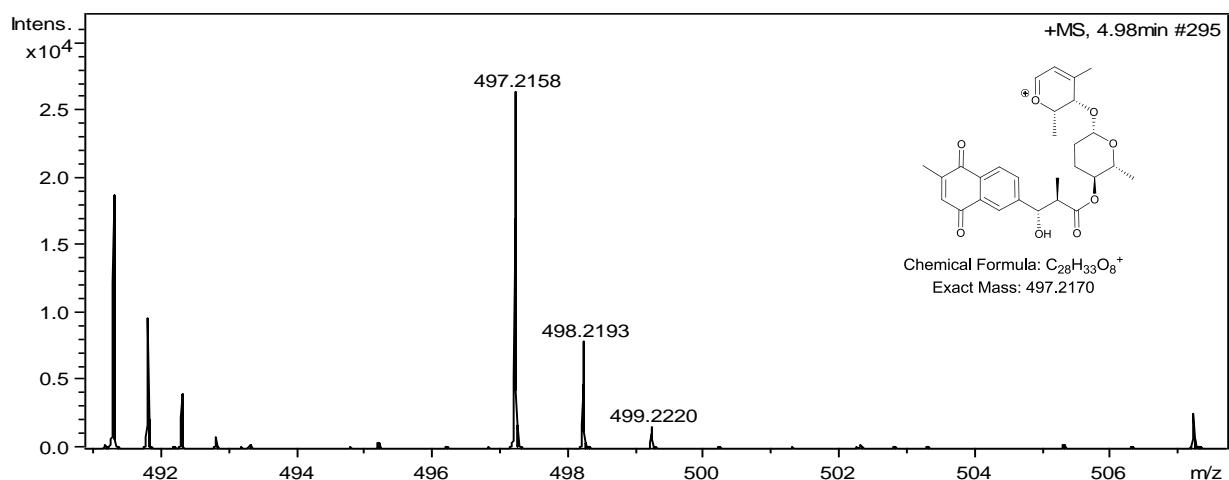
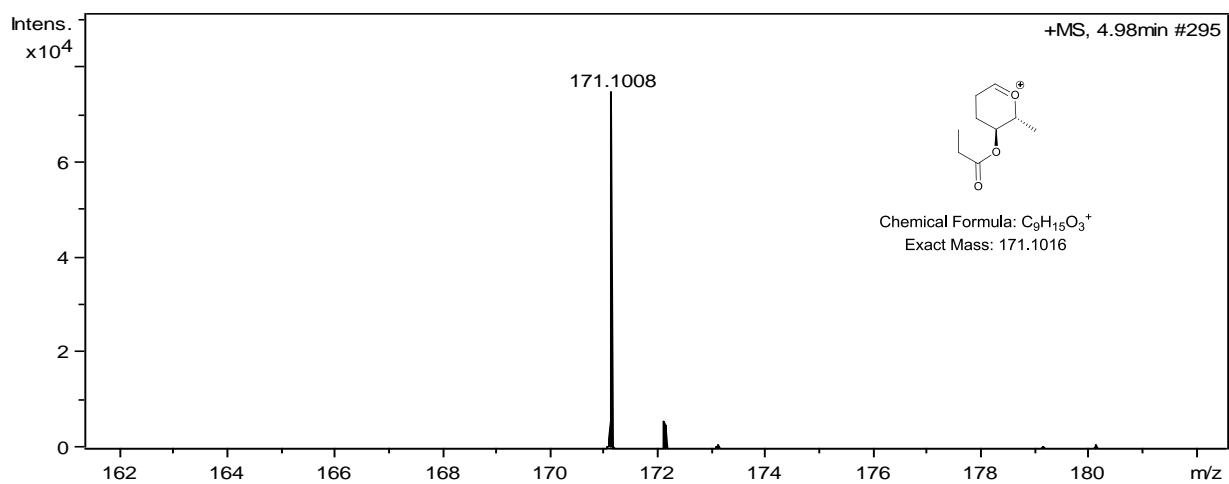


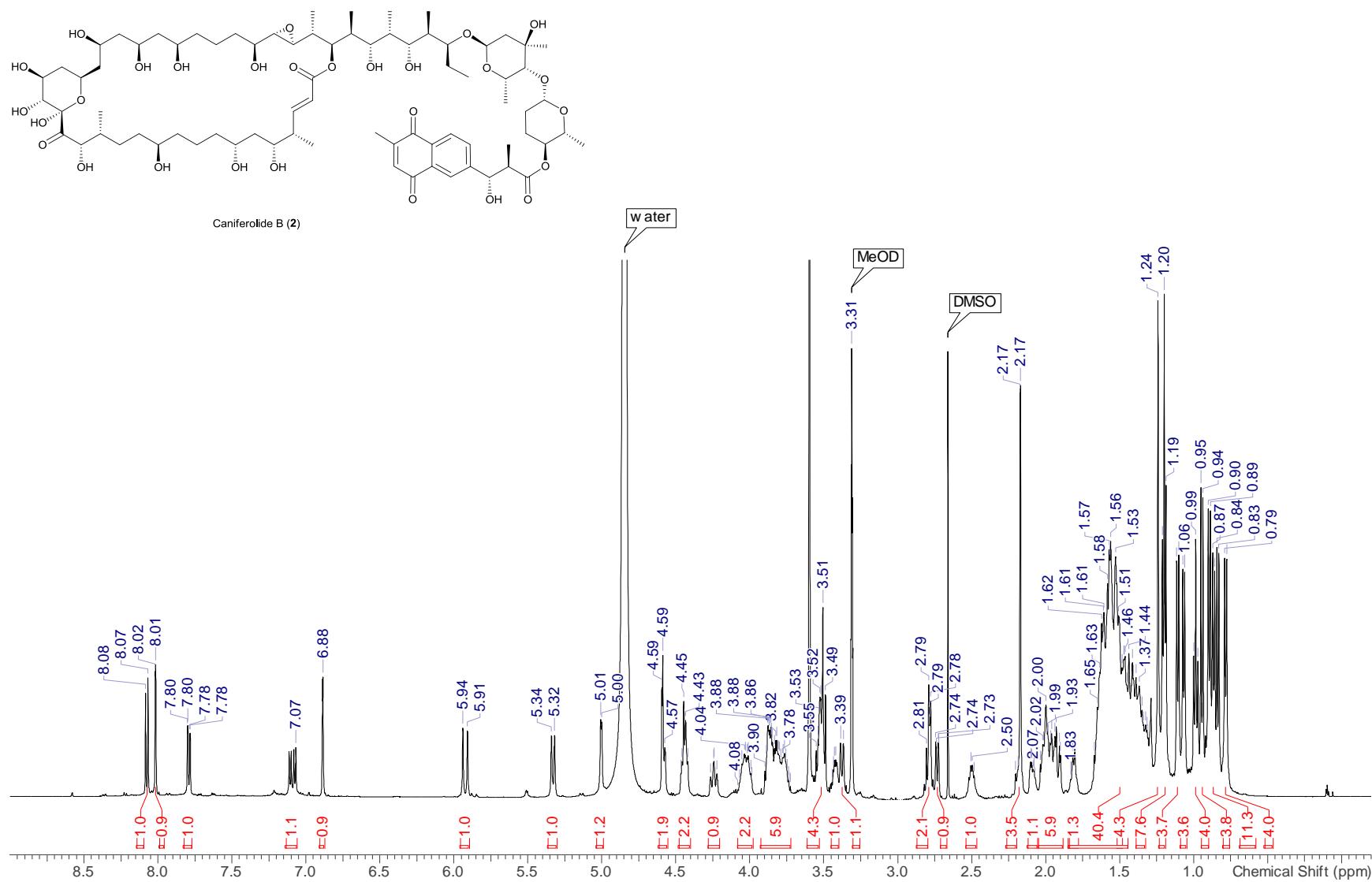
Figure S20 cont. (+) ESI-TOF MS of caniferolide B (**2**) including expansions showing key in-source fragment ions.



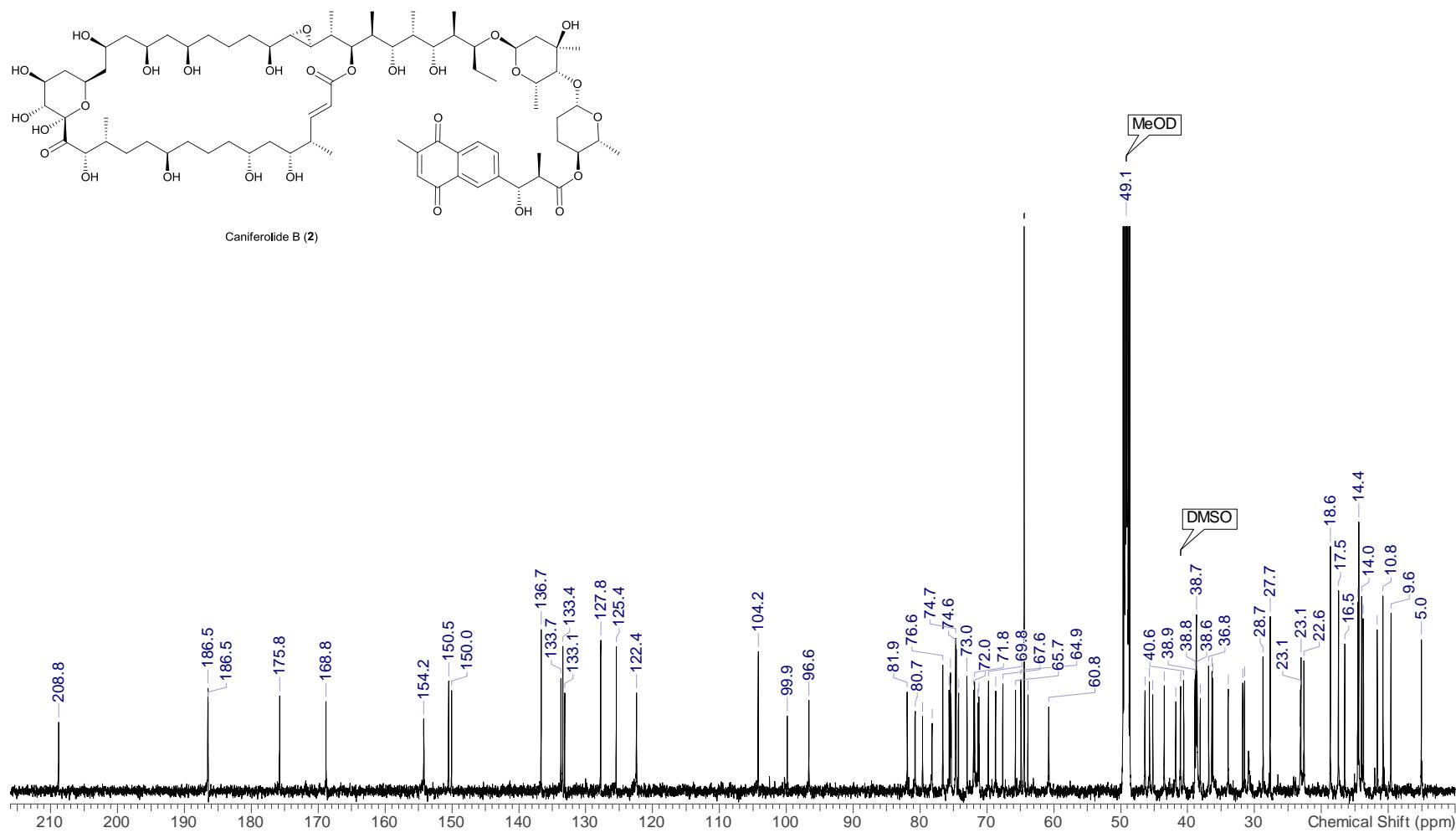
**Figure S20cont.** (+) ESI-TOF MS of caniferolide B (**2**) including expansions showing key in-source fragment ions.



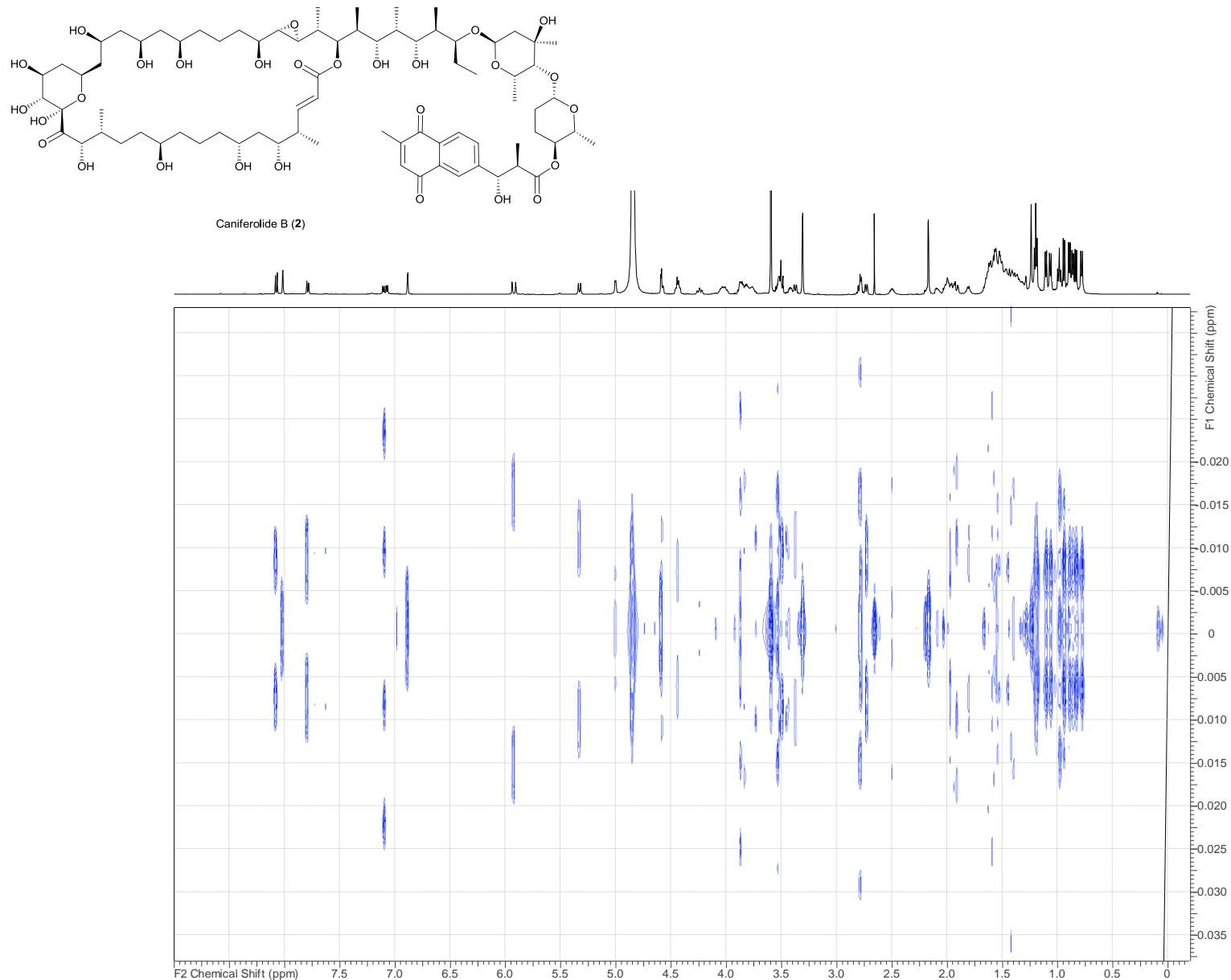
**Figure S20 cont.** (+) ESI-TOF MS of caniferolide B (**2**) including expansions showing key in-source fragment ions.



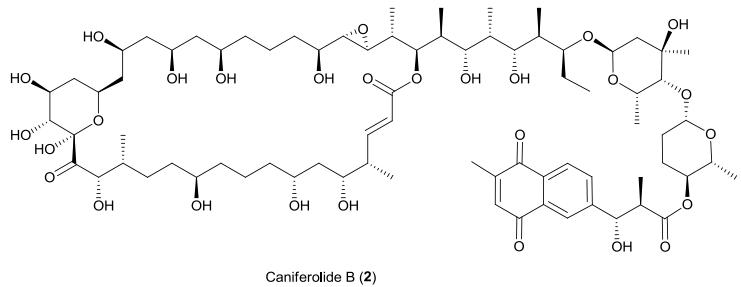
**Figure S21.** <sup>1</sup>H NMR spectrum (CD<sub>3</sub>OD, 500 MHz) of caniferolide B (**2**).



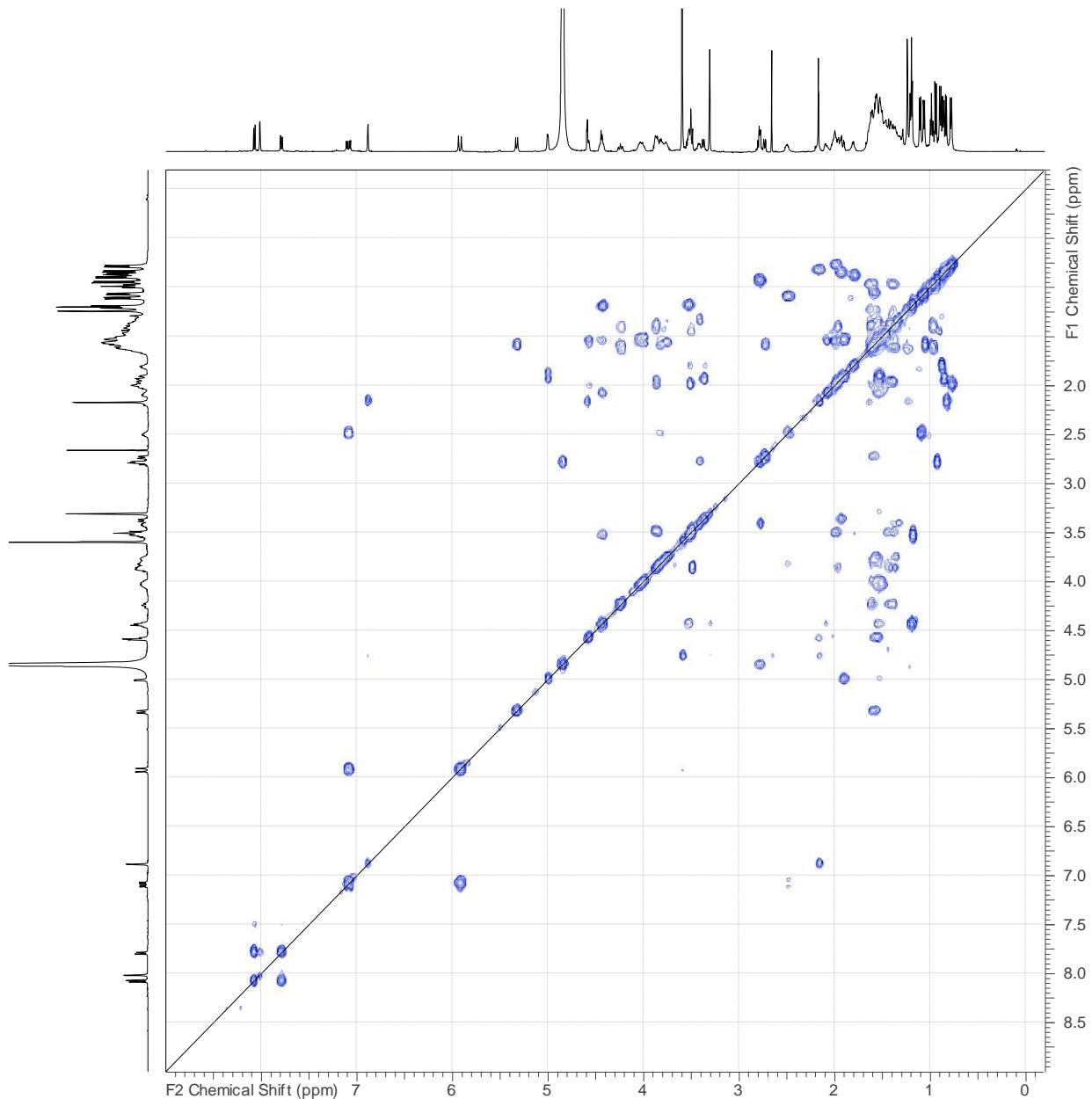
**Figure S22.** <sup>13</sup>C NMR spectrum (CD<sub>3</sub>OD, 125 MHz) of caniferolide B (**2**).



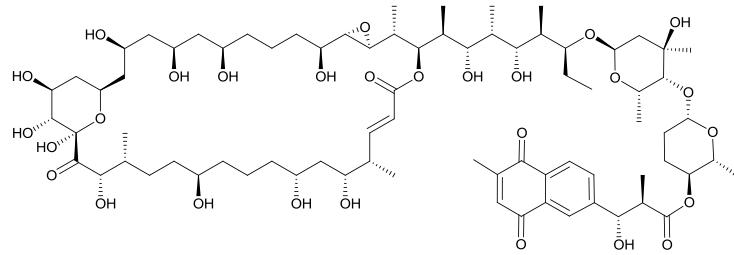
**Figure S23.** JRES spectrum of caniferolide B (2).



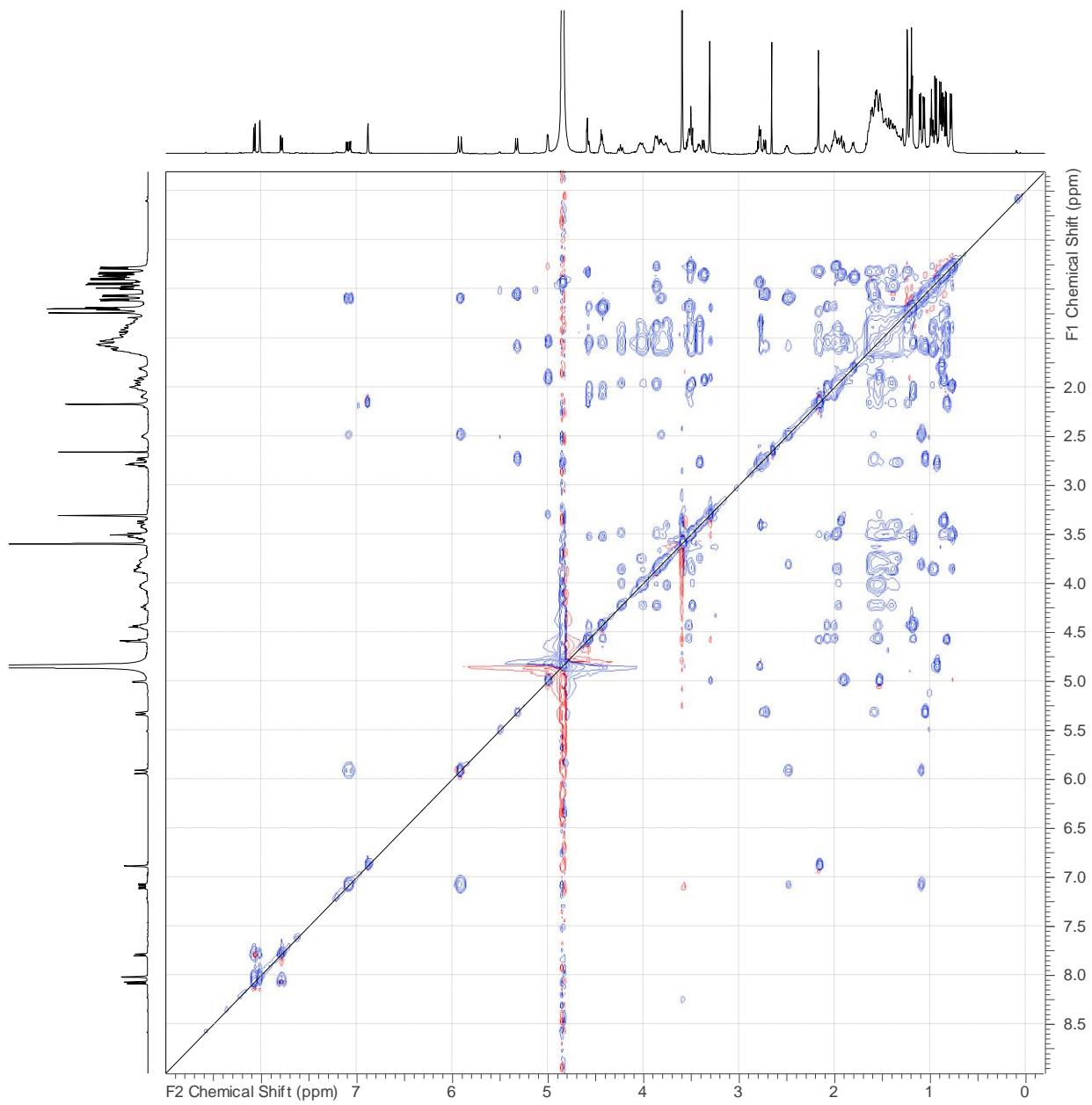
Caniferolide B (2)



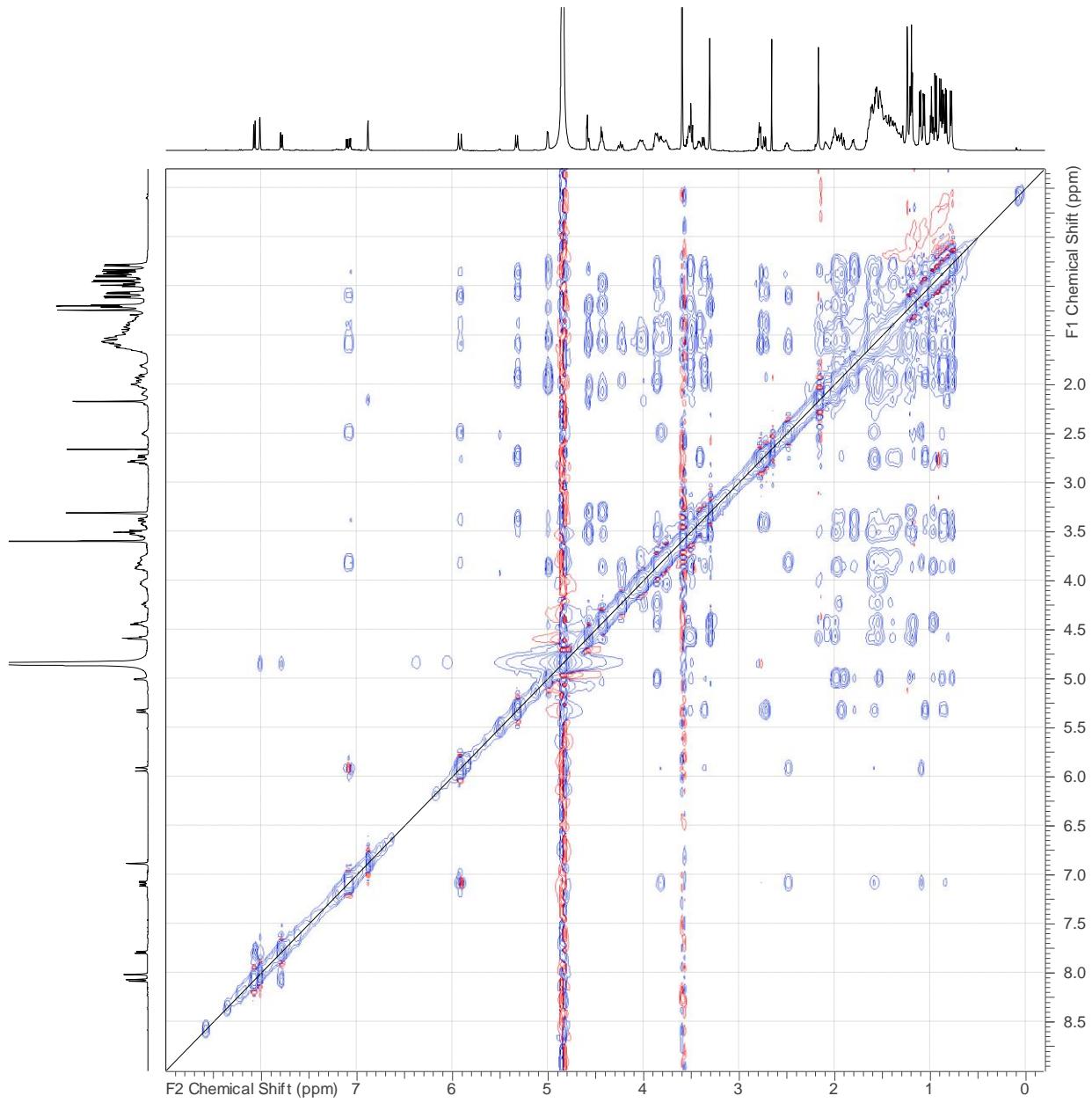
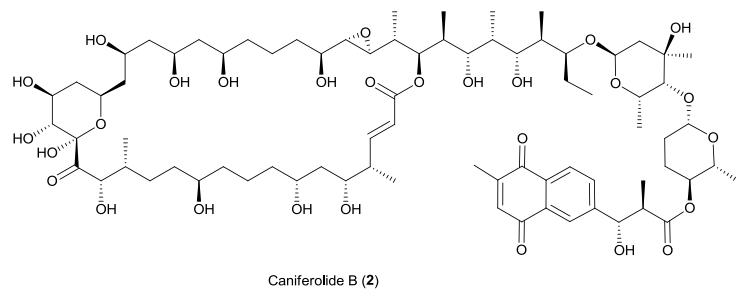
**Figure S24.** COSY spectrum of caniferolide B (2).



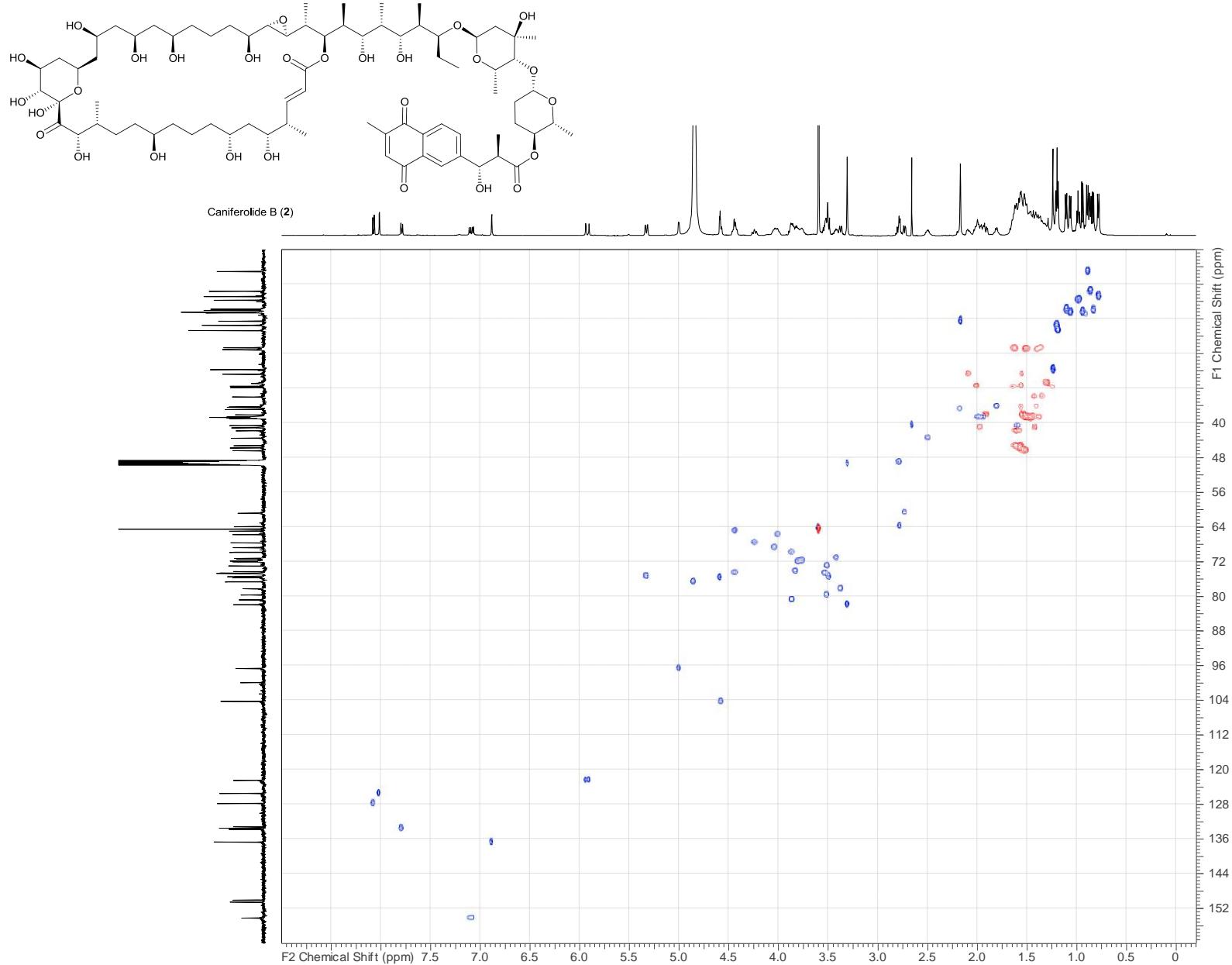
Caniferolide B (**2**)



**Figure S25.** TOCSY spectrum of caniferolide B (**2**).



**Figure S26.** NOESY spectrum of caniferolide B (2).



**Figure S27.**Edited HSQC spectrum of caniferolide B (**2**).

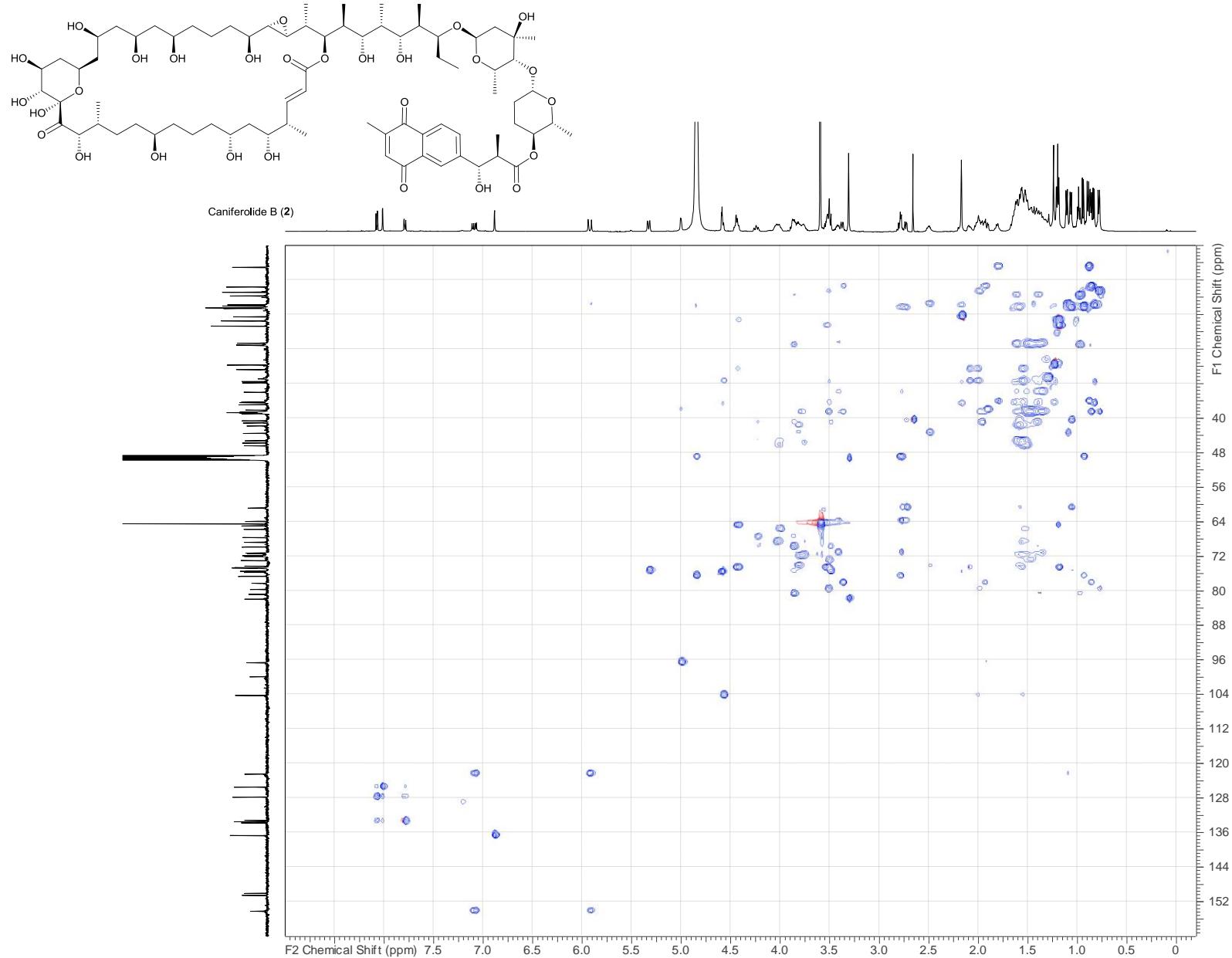
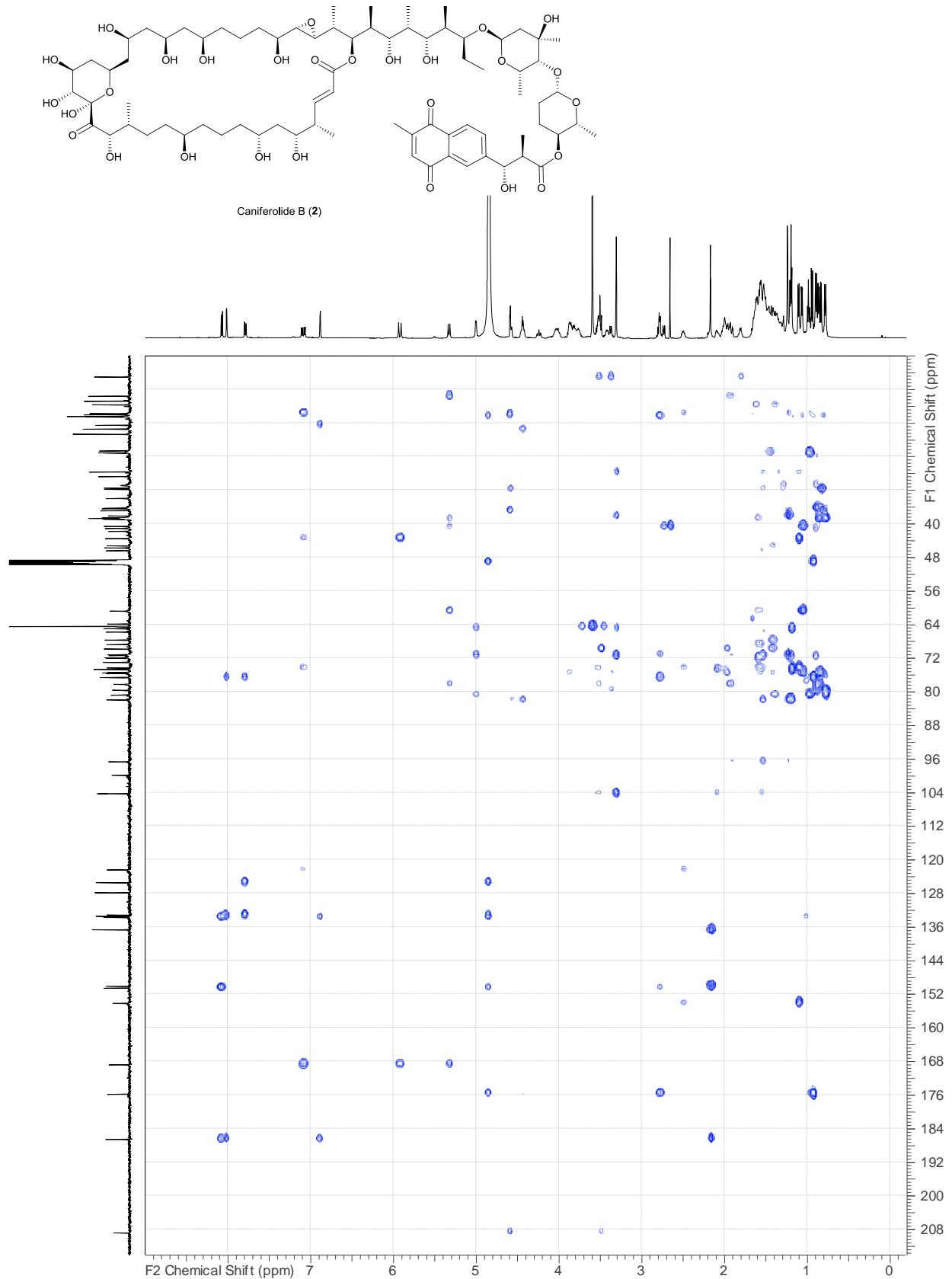
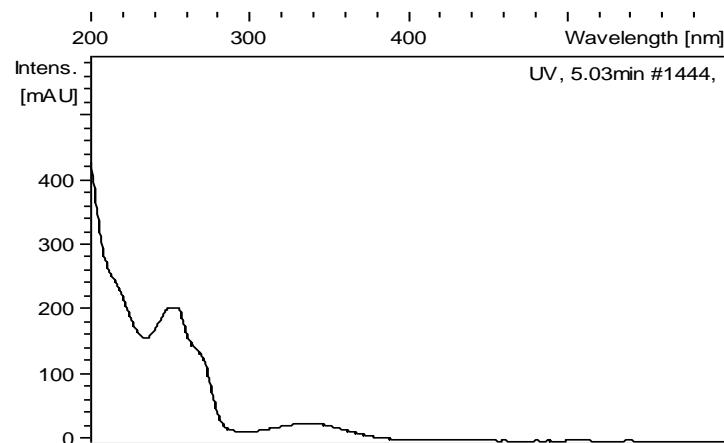
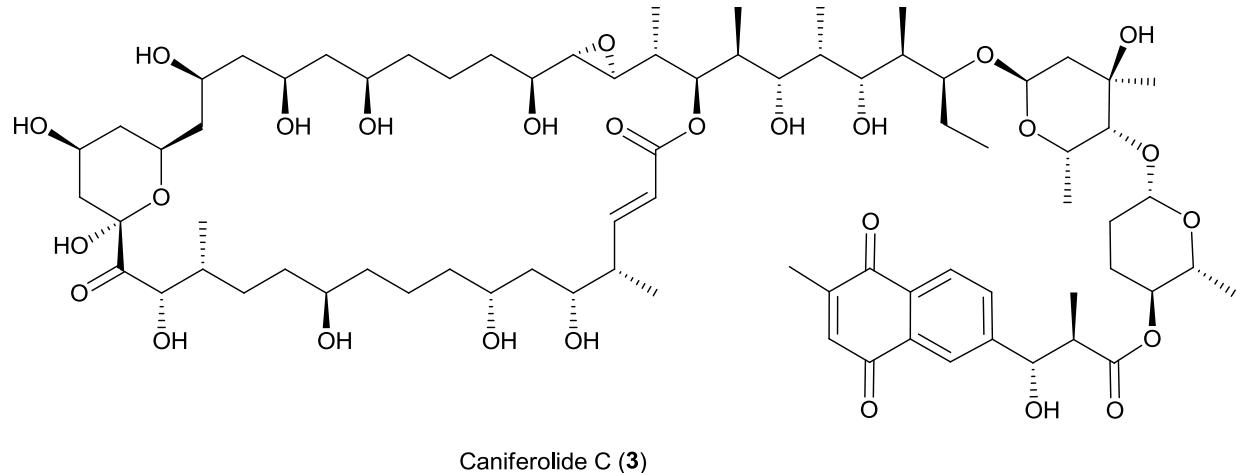


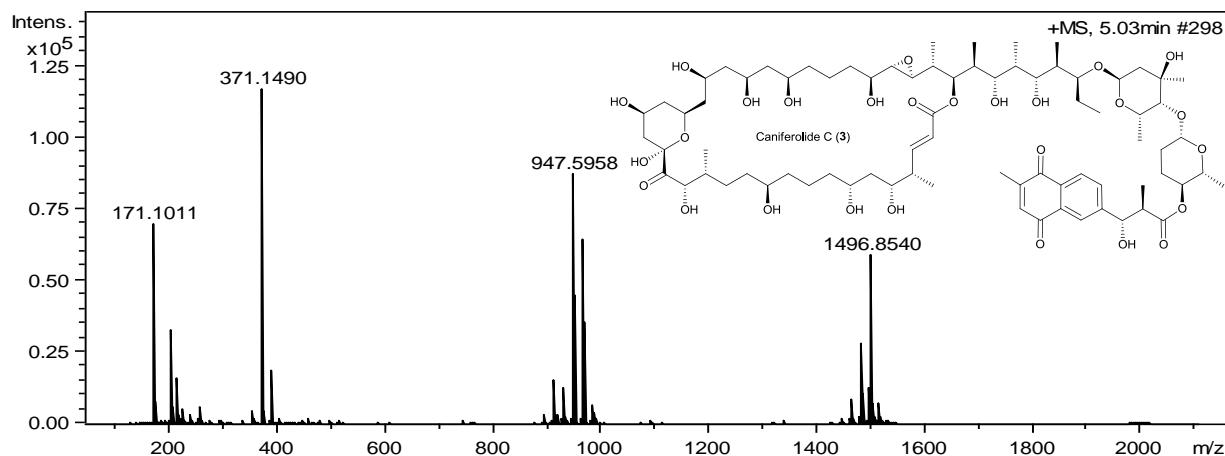
Figure S28. HSQC-TOCSY spectrum of caniferolide B (**2**).



**Figure S29.** HMBC spectrum of caniferolide B (**2**).



**Figure S30.** UV (DAD) spectrum of caniferolide C (3).



**Figure S31.** (+) ESI-TOF MS of caniferolide C (3) including expansions showing key in-source fragment ions.

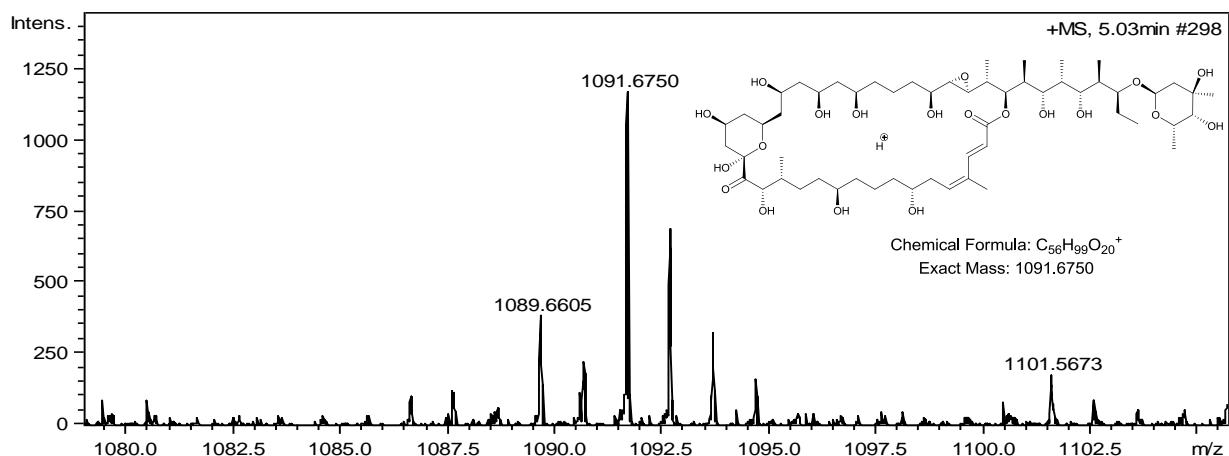
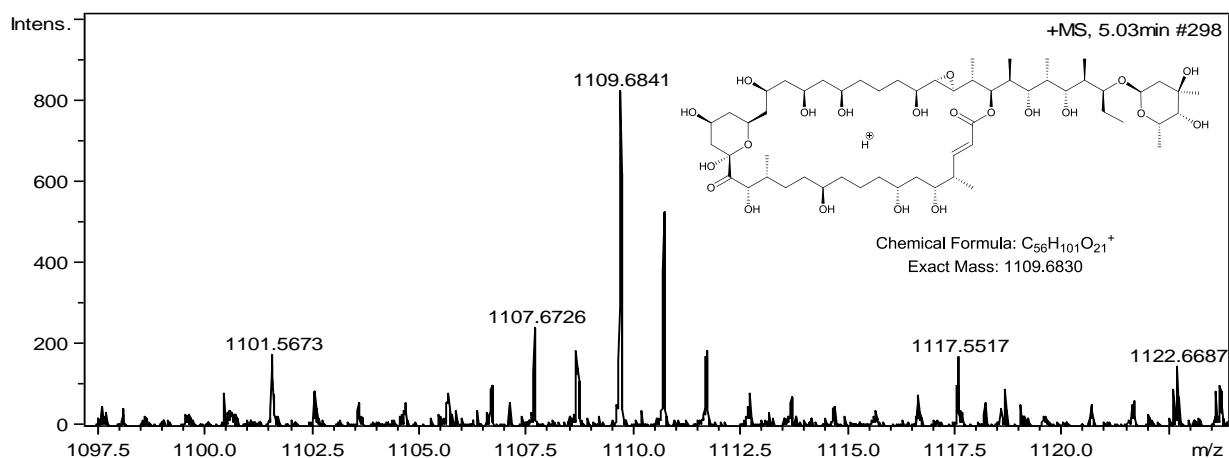
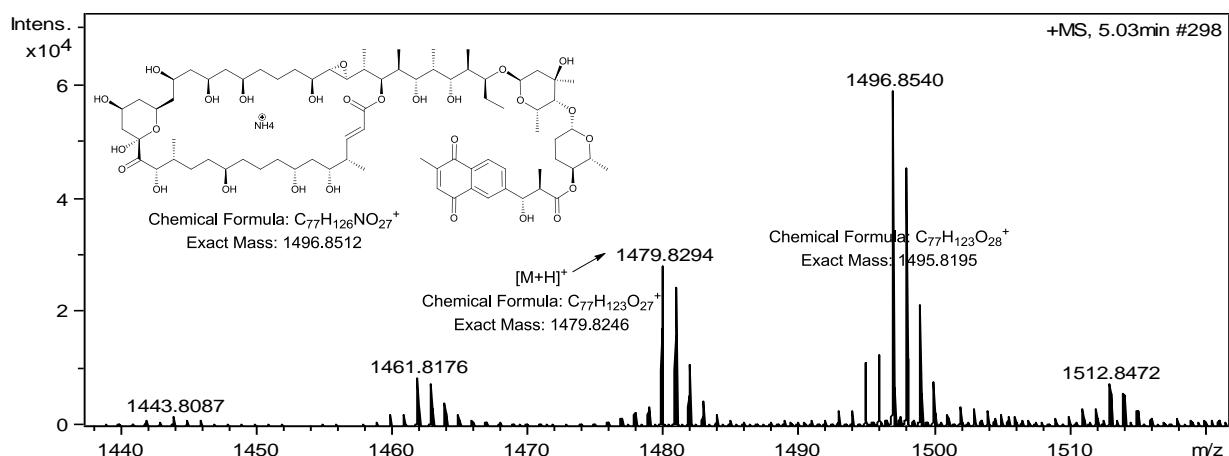
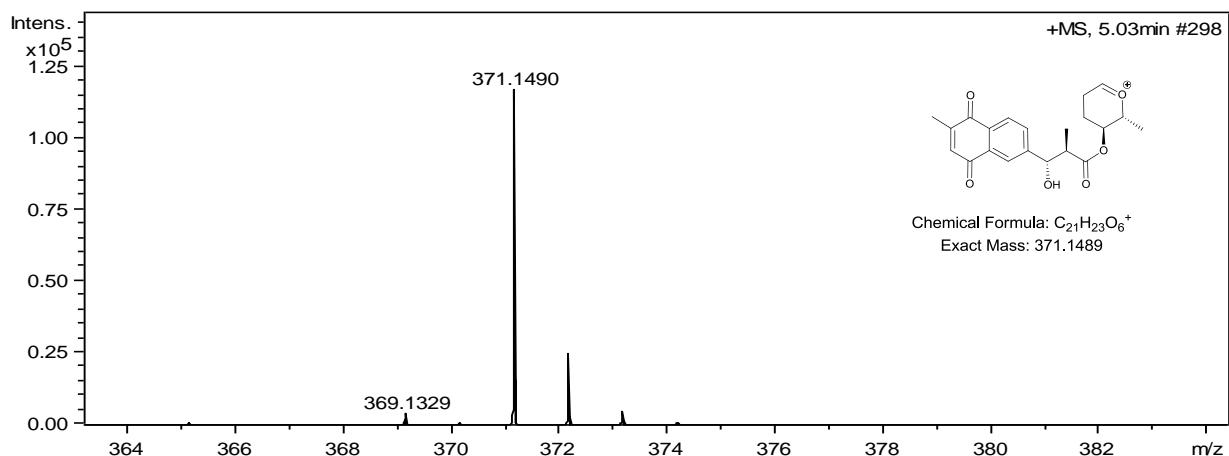
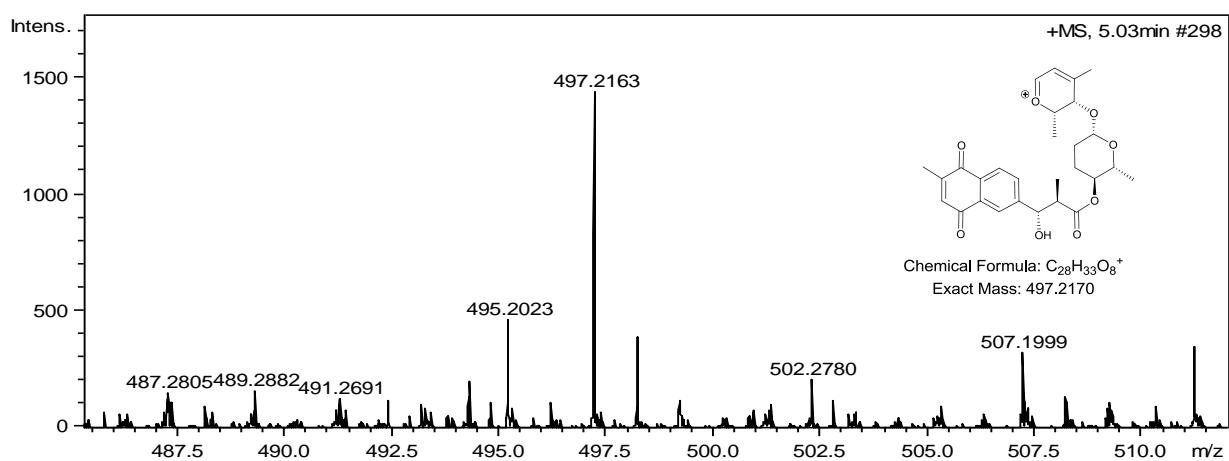
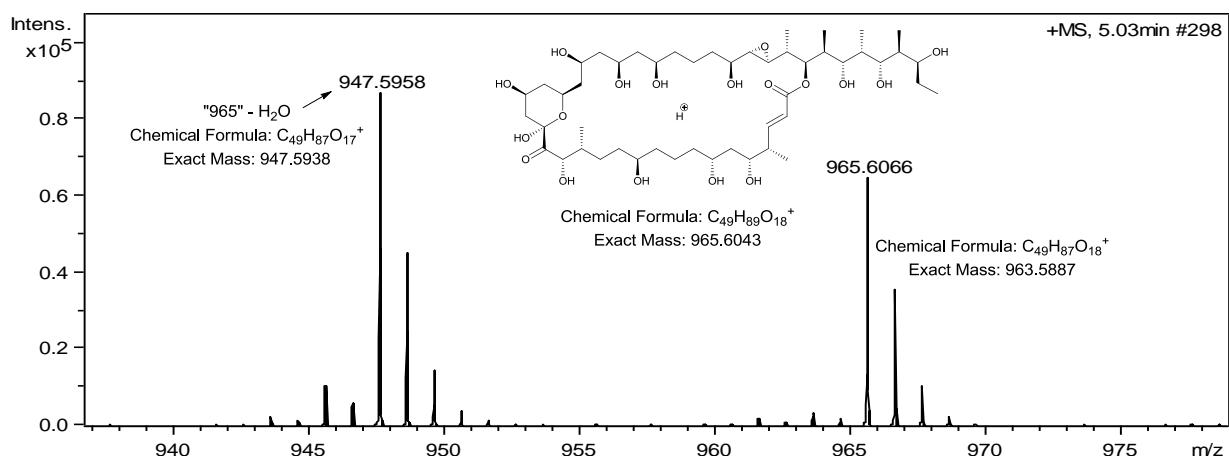
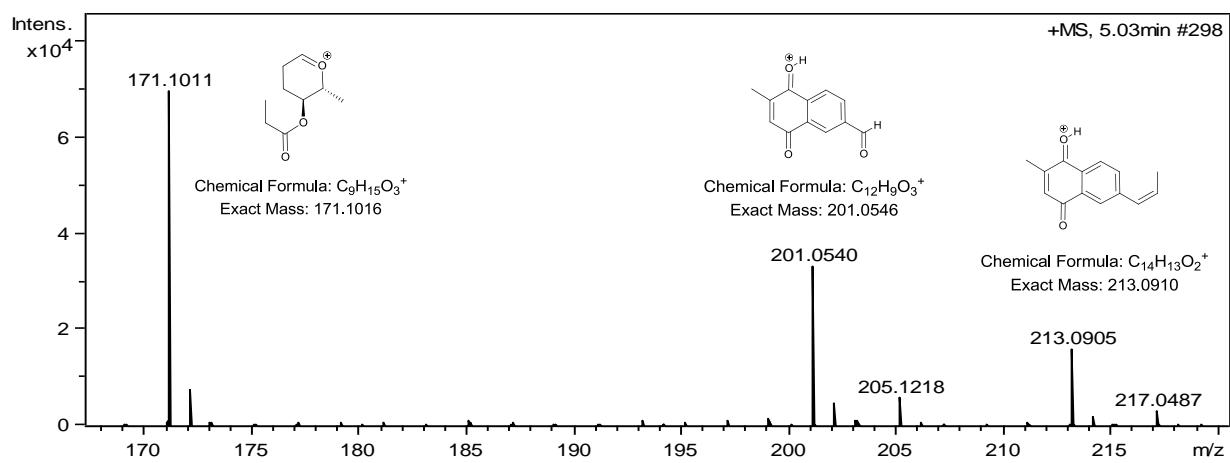


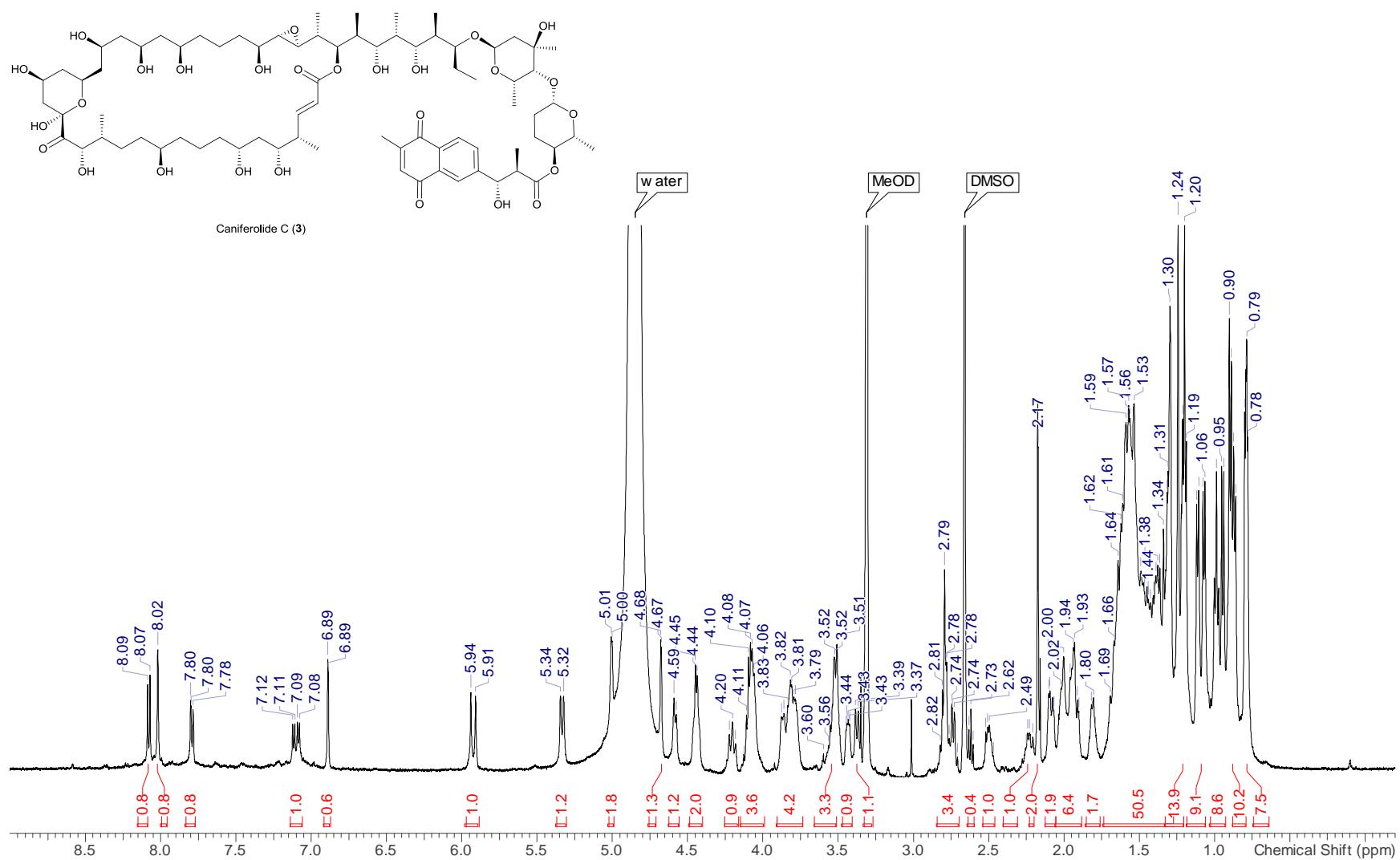
Figure S31 cont. (+) ESI-TOF MS of caniferolide C (3) including expansions showing key in-source fragment ions.

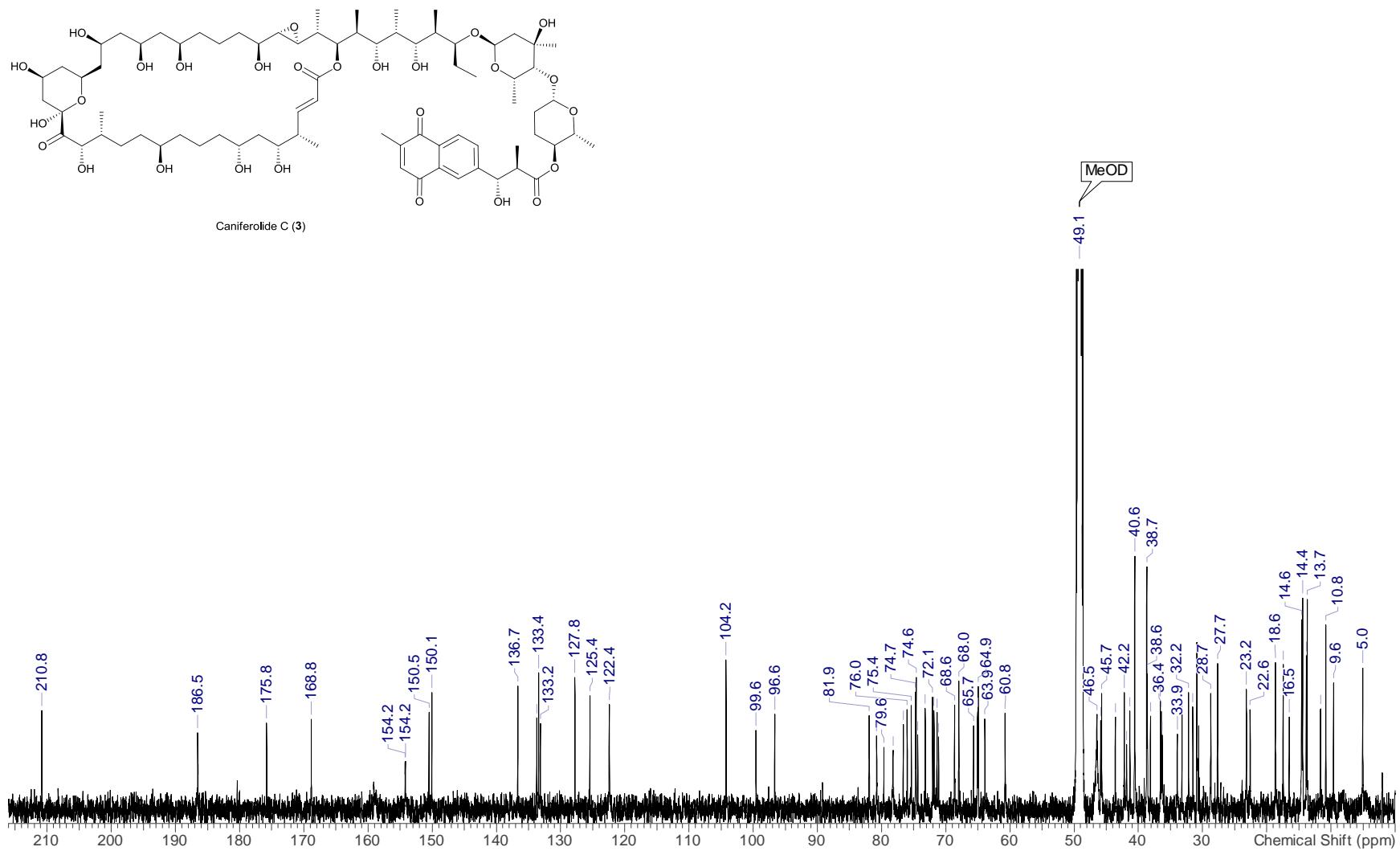


**Figure S31 cont.** (+) ESI-TOF MS of caniferolide C (**3**) including expansions showing key in-source fragment ions.

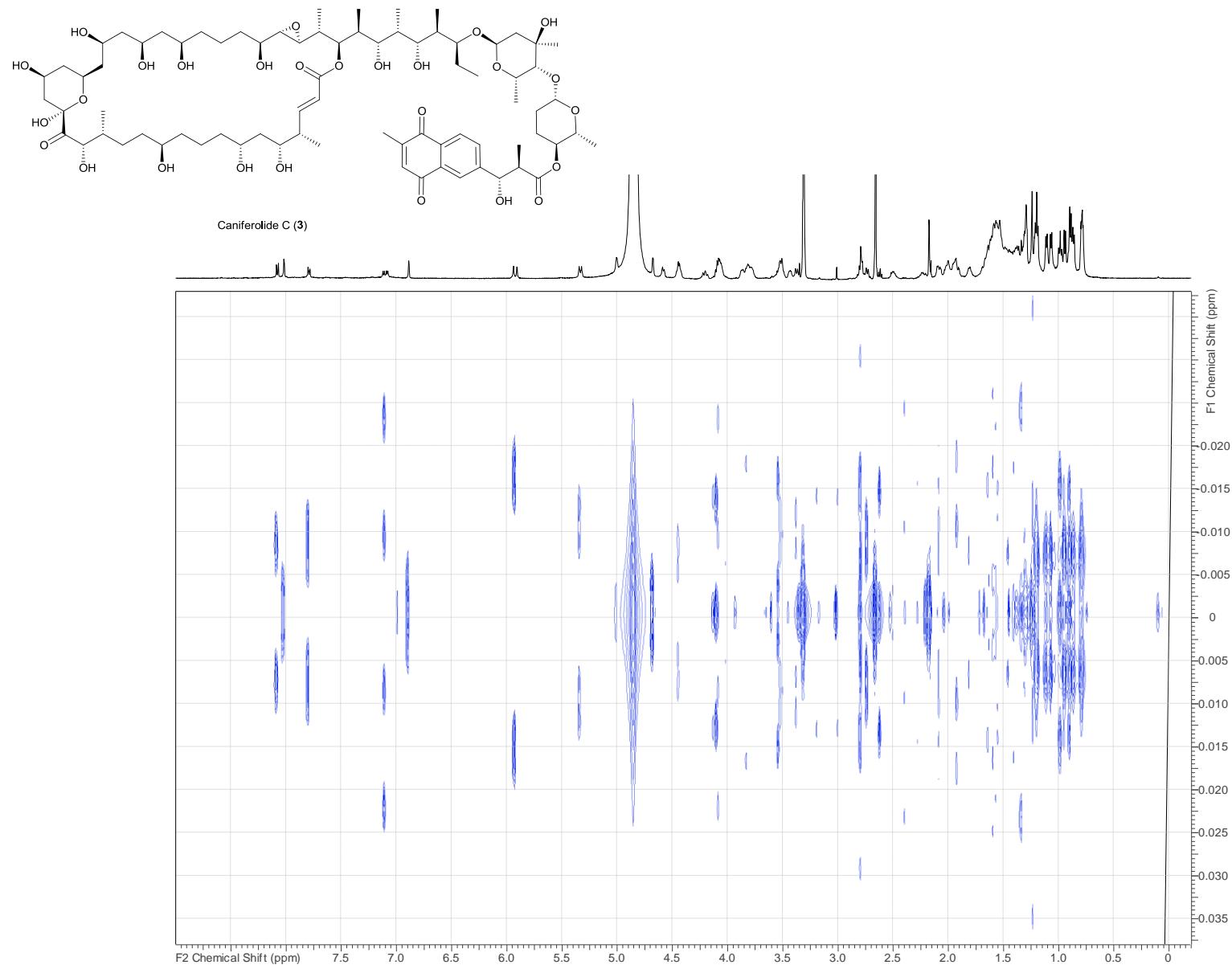


**Figure S31 cont.** (+) ESI-TOF MS of caniferolide C (**3**) including expansions showing key in-source fragment ions.

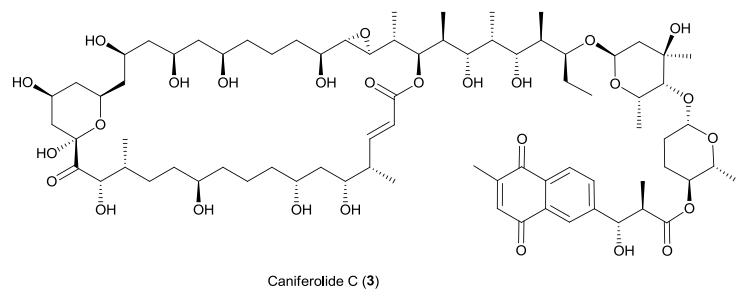




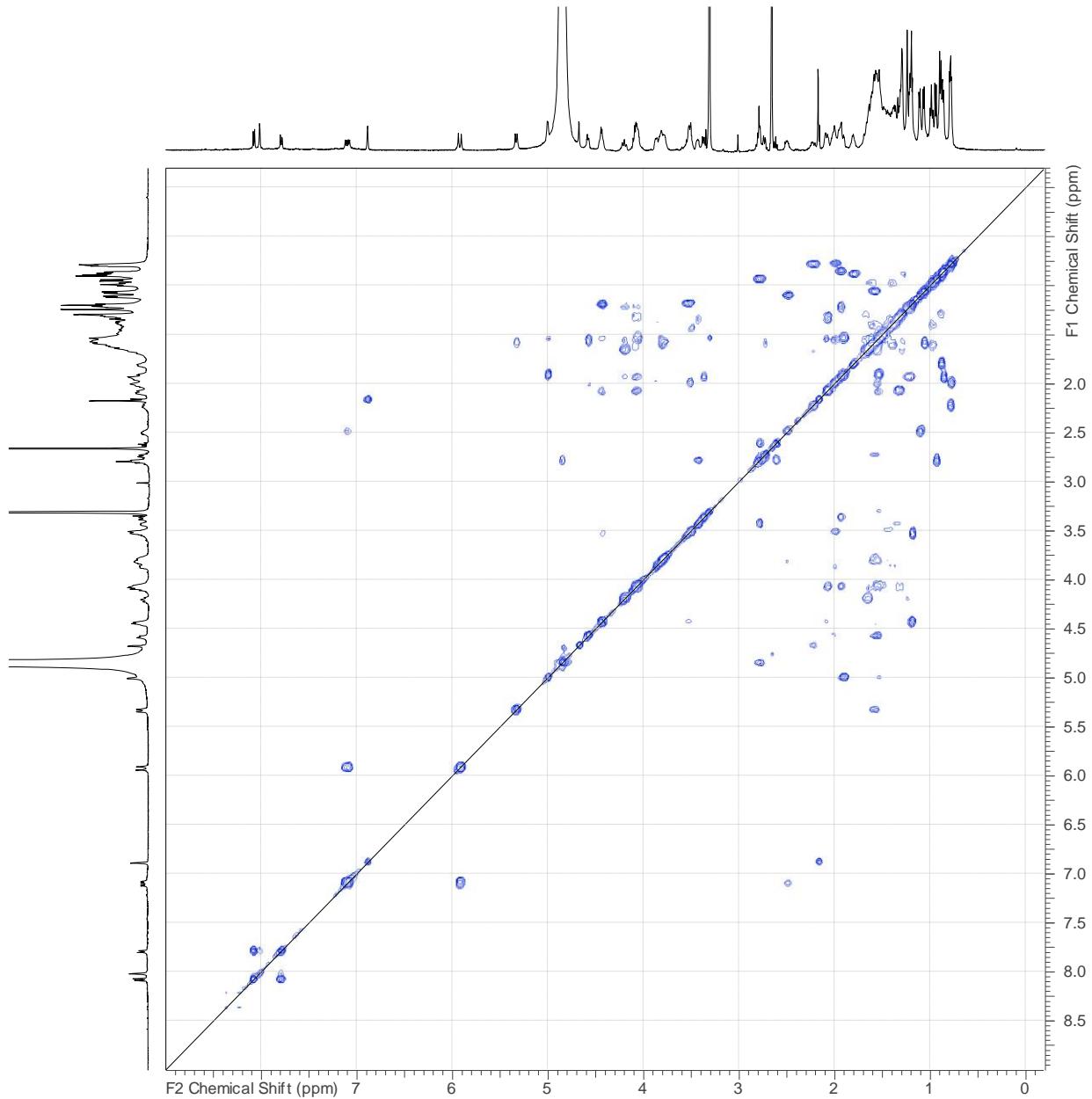
**Figure S33.**  $^{13}\text{C}$  NMR spectrum ( $\text{CD}_3\text{OD}$ , 125 MHz) of caniferolide C (3).



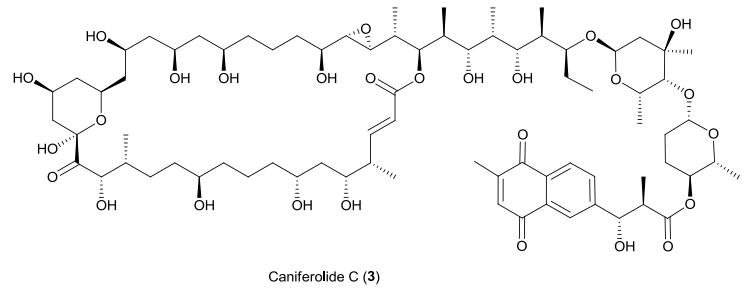
**Figure S34.** JRES spectrum of caniferolide C (3).



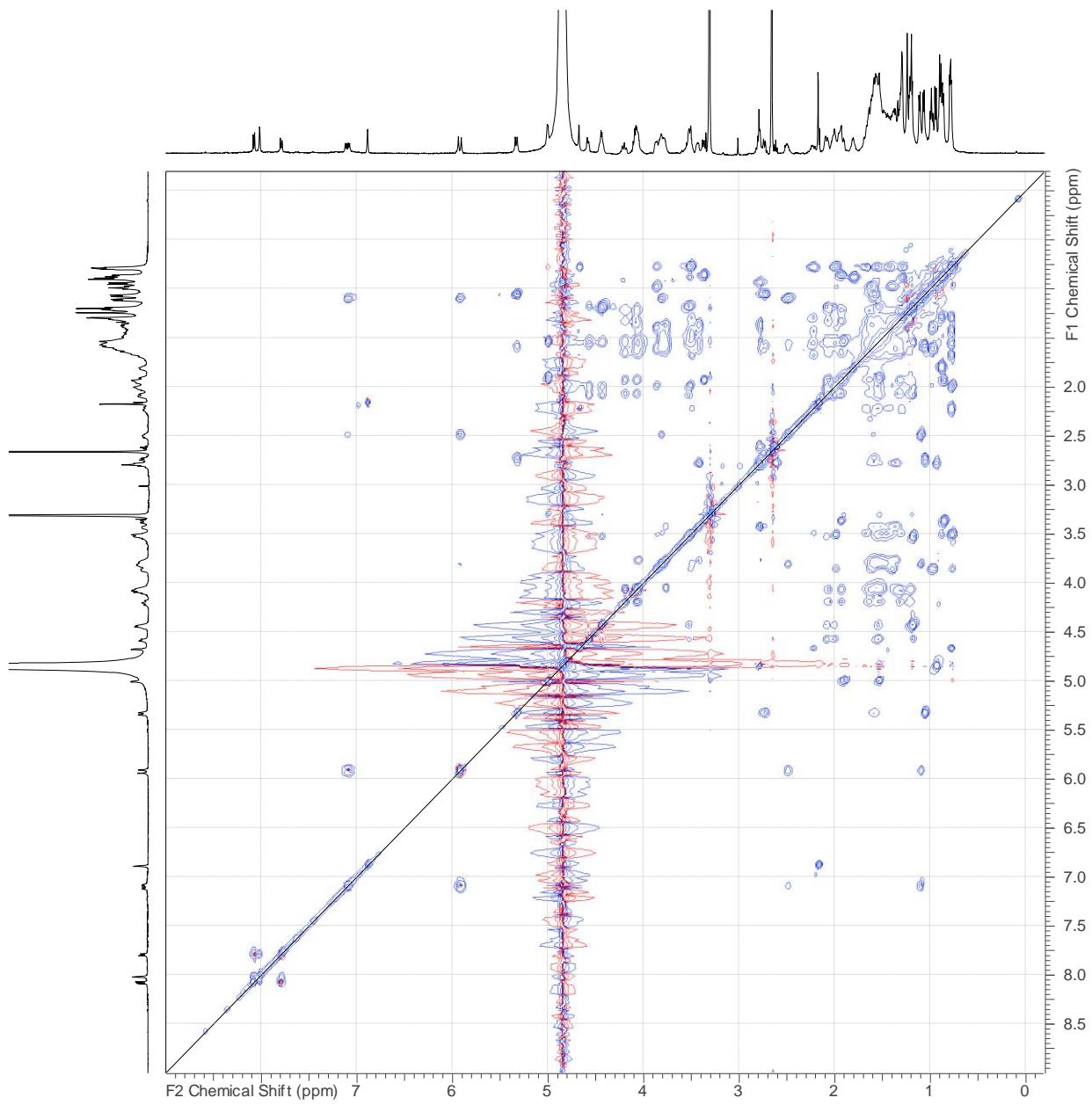
Caniferolide C (3)



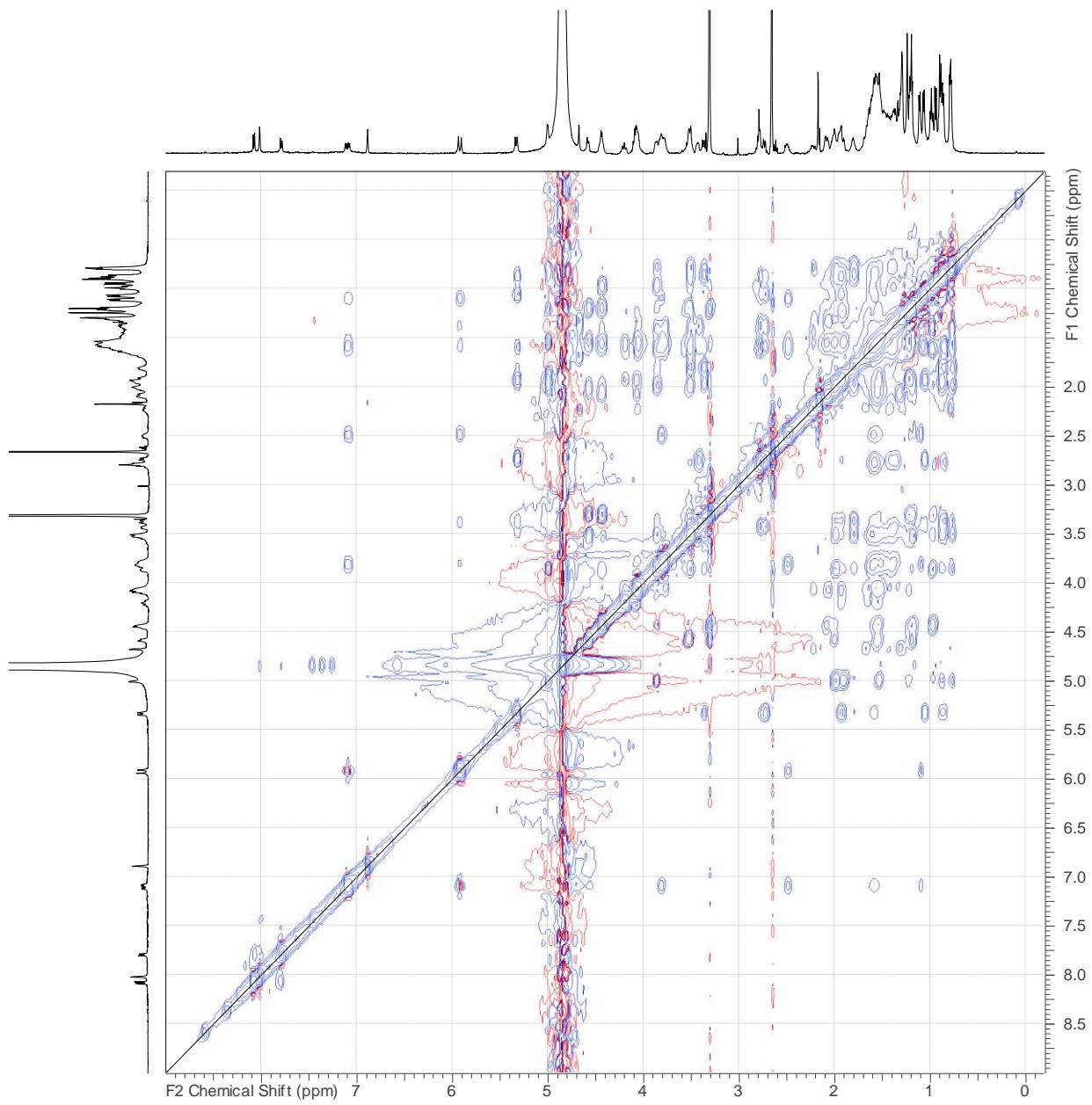
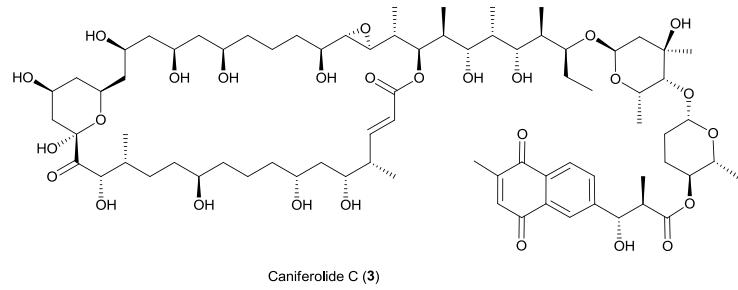
**Figure S35.** COSY spectrum of caniferolide C (3).



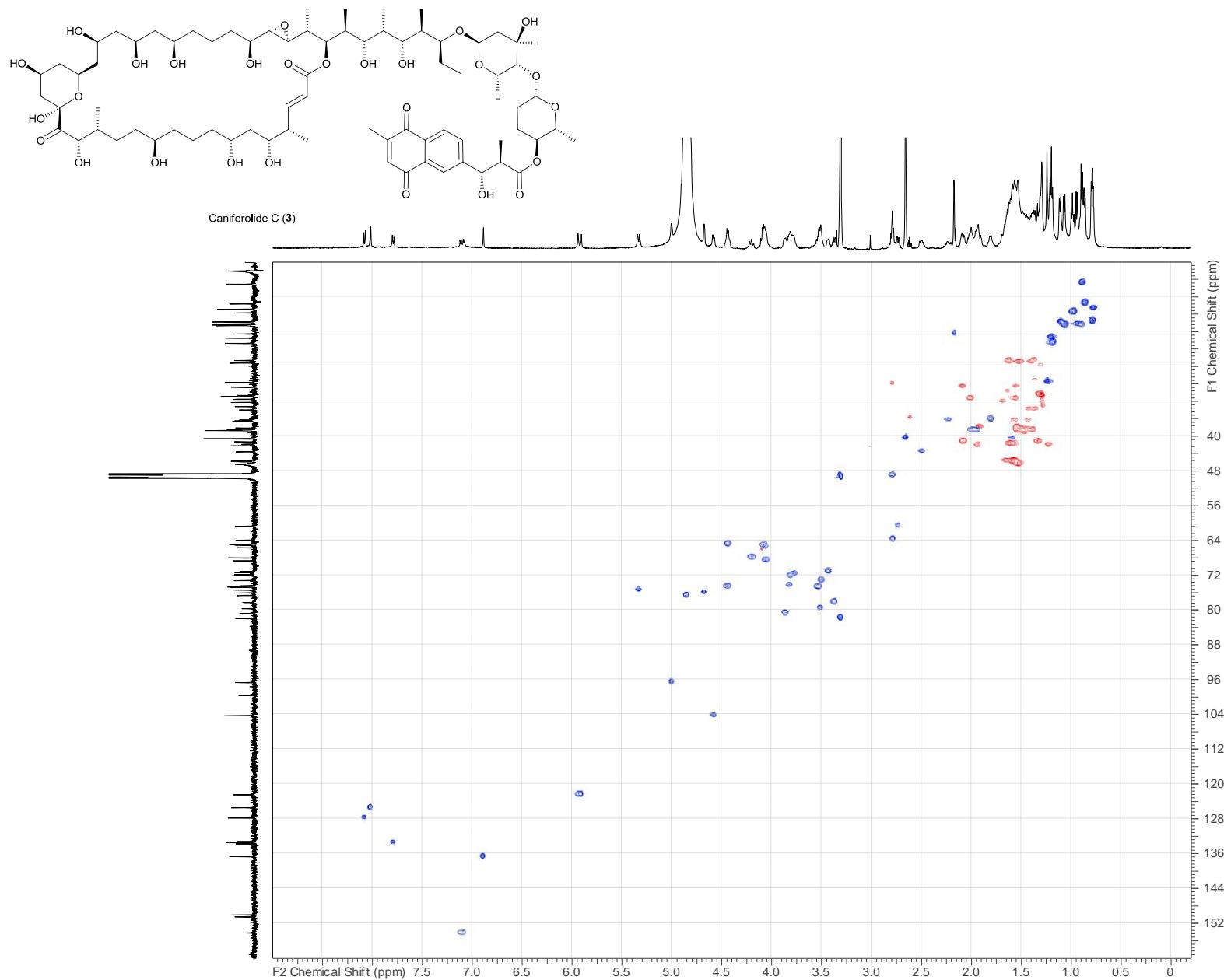
Caniferolide C (3)



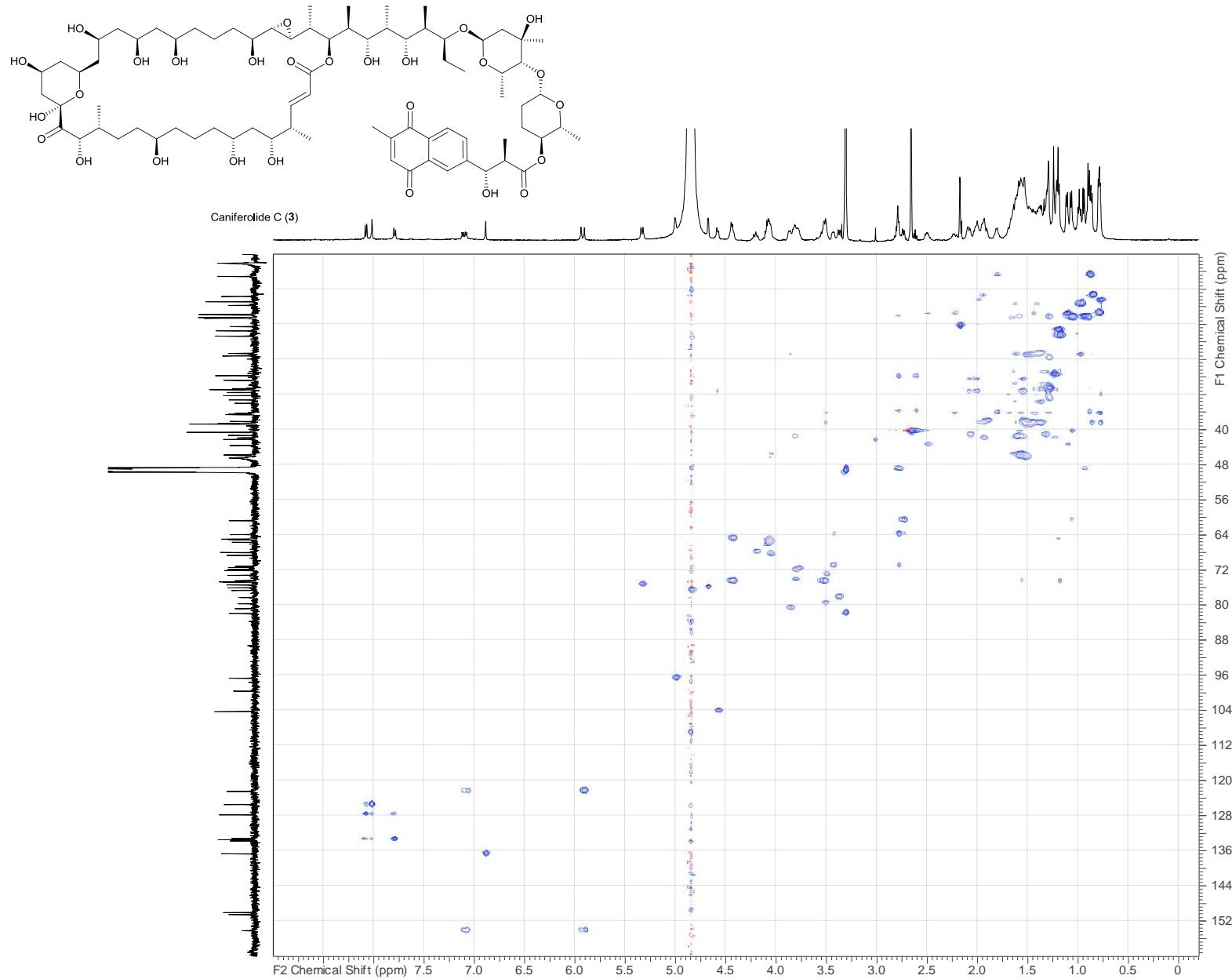
**Figure S36.** TOCSY spectrum of caniferolide C (3).



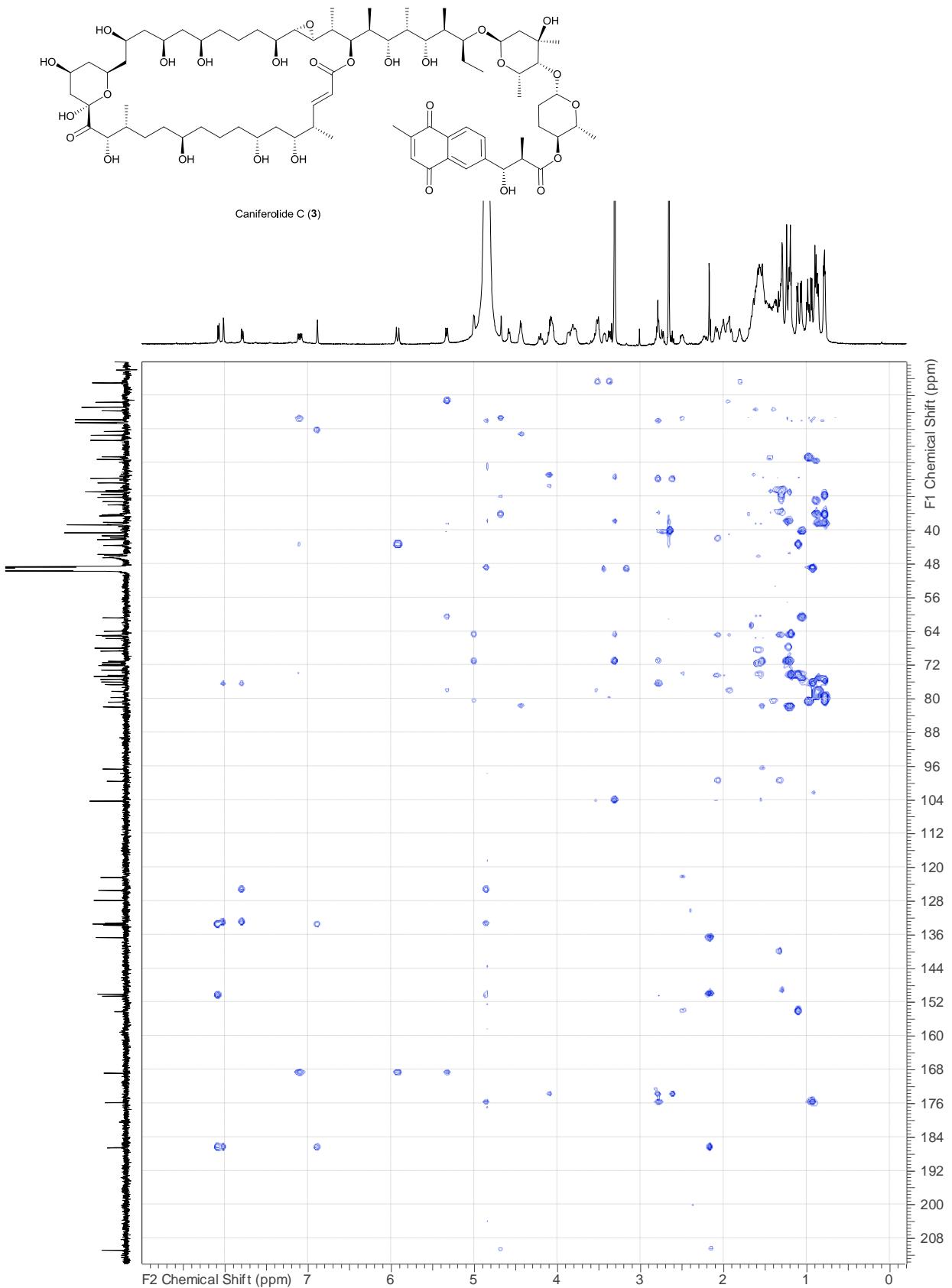
**Figure S37.** NOESY spectrum of caniferolide C (3).



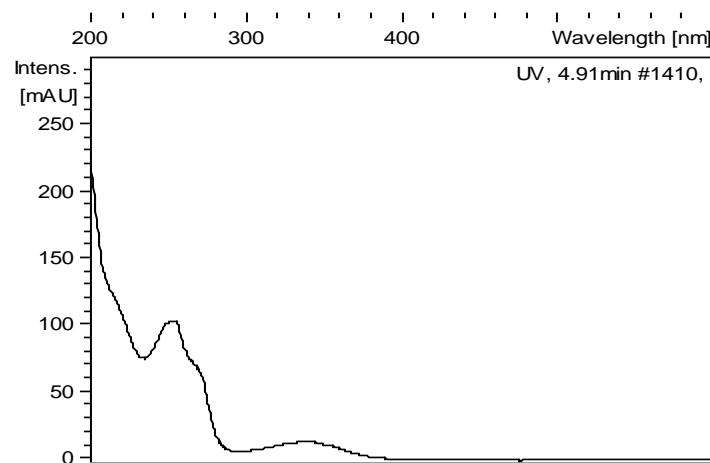
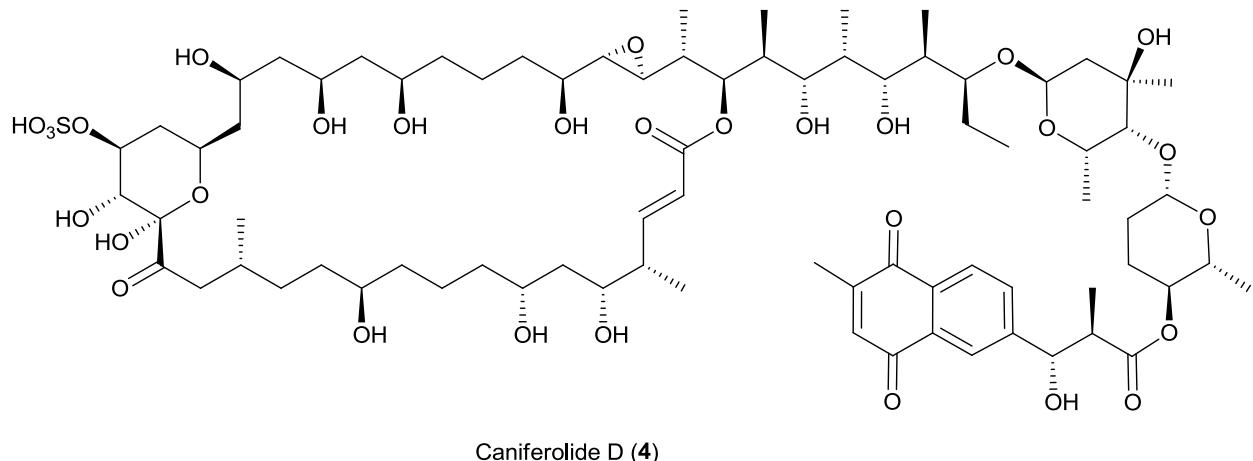
**Figure S38.** Edited HSQC spectrum of caniferolide C (3).



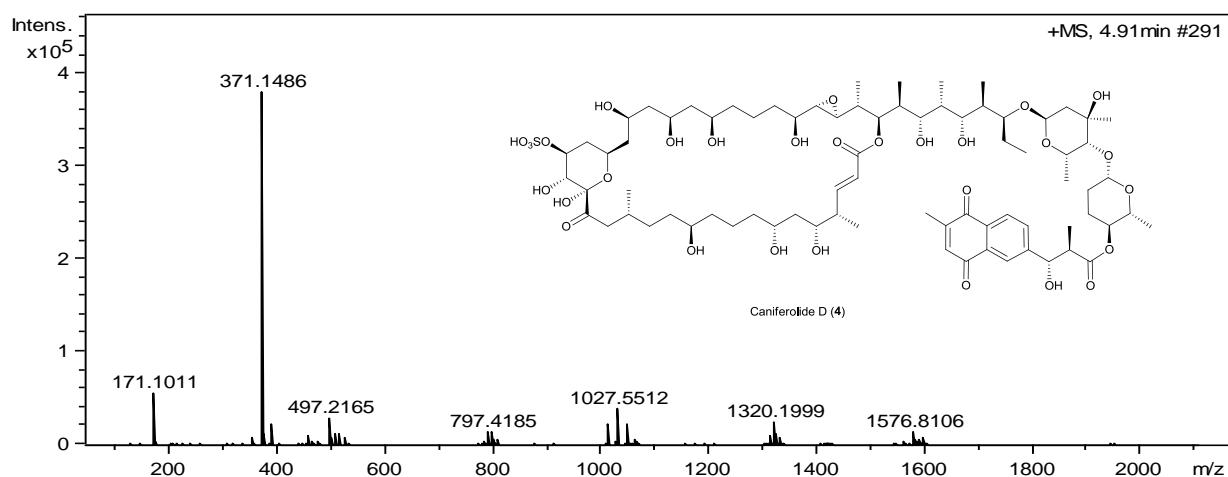
**Figure S39.** HSQC-TOCSY spectrum of caniferolide C (3).



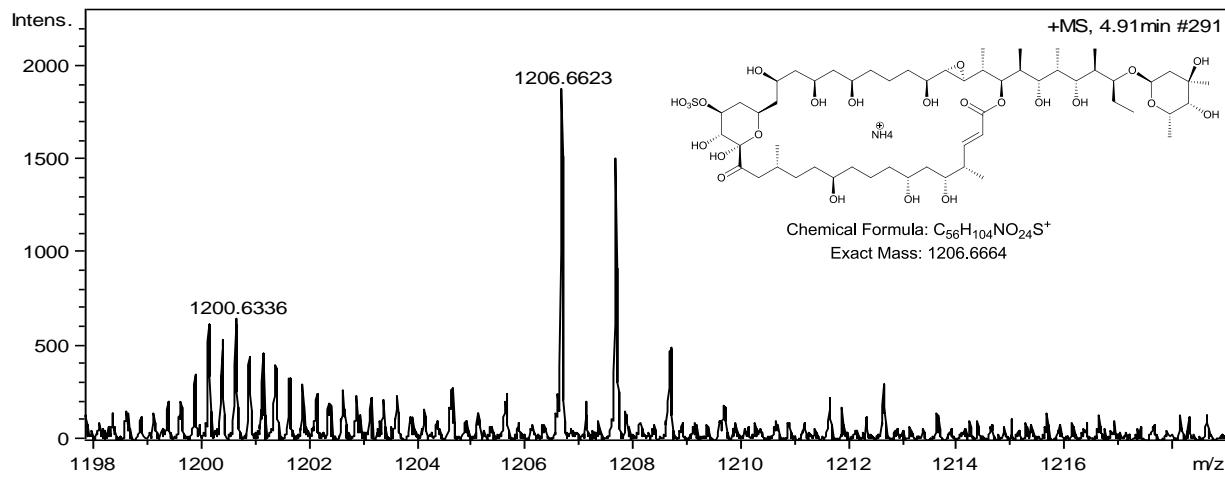
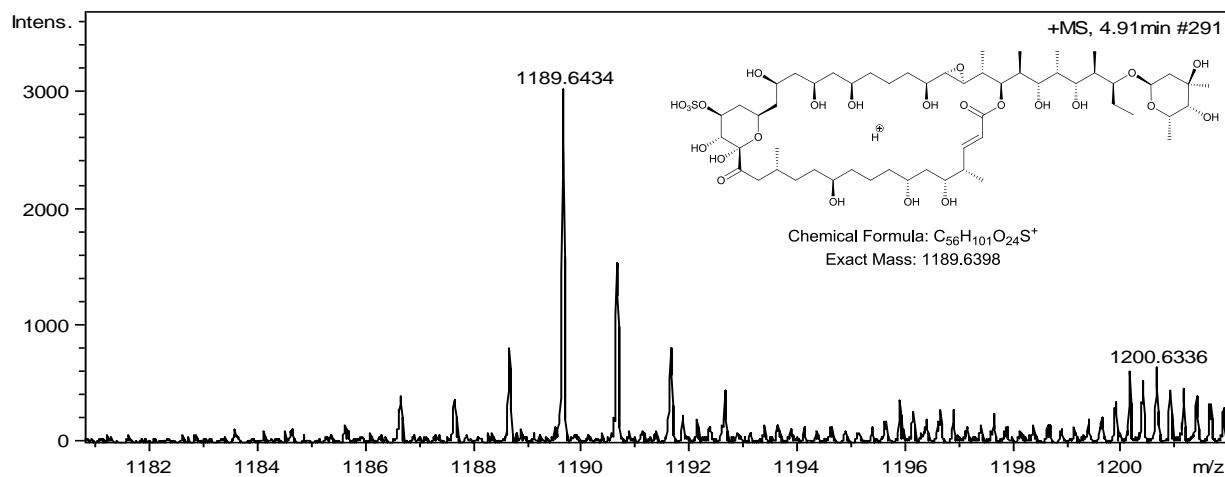
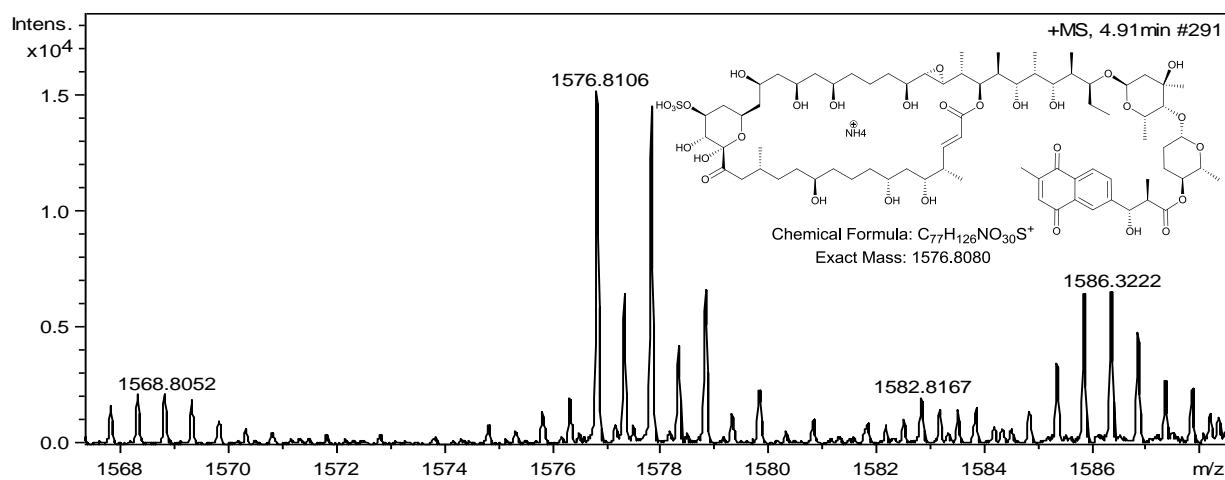
**Figure S40.** HMBC spectrum of caniferolide C (3).



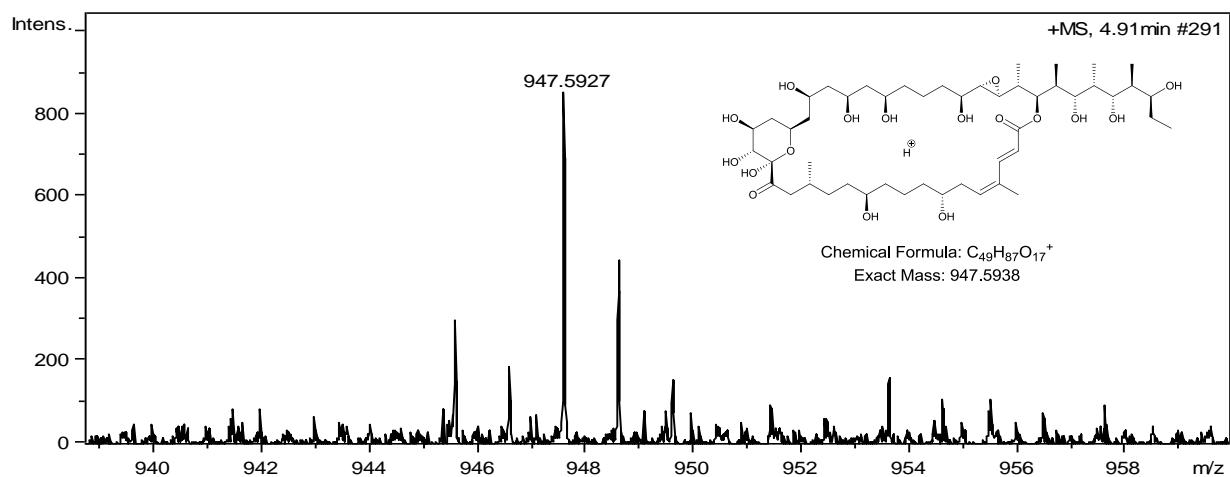
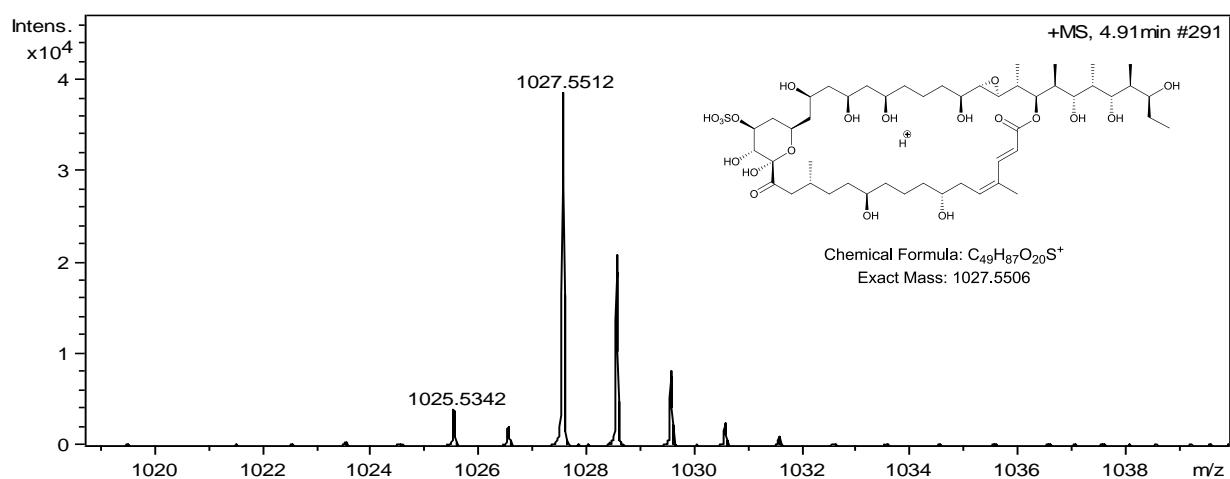
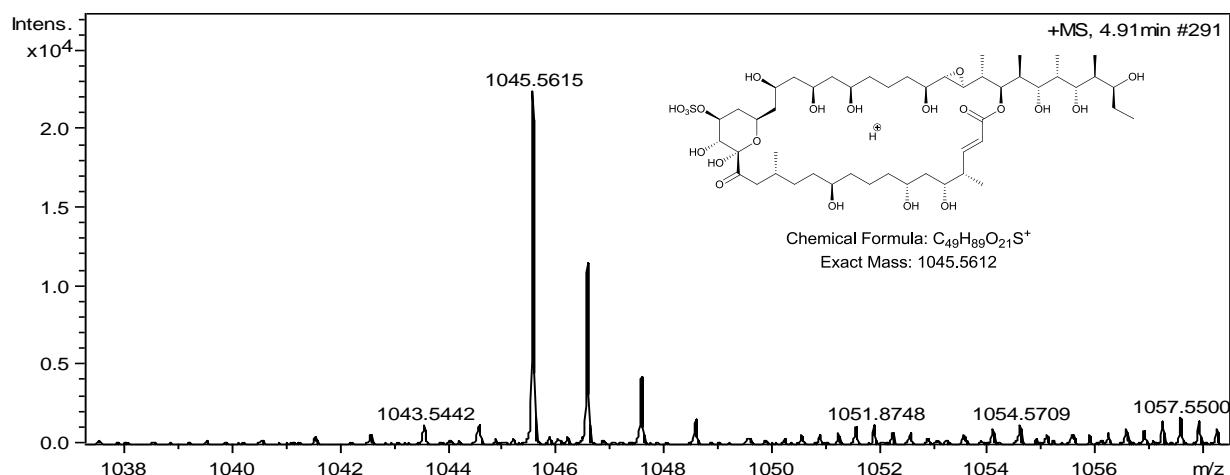
**Figure 41.** UV (DAD) spectrum of caniferolide D (4).



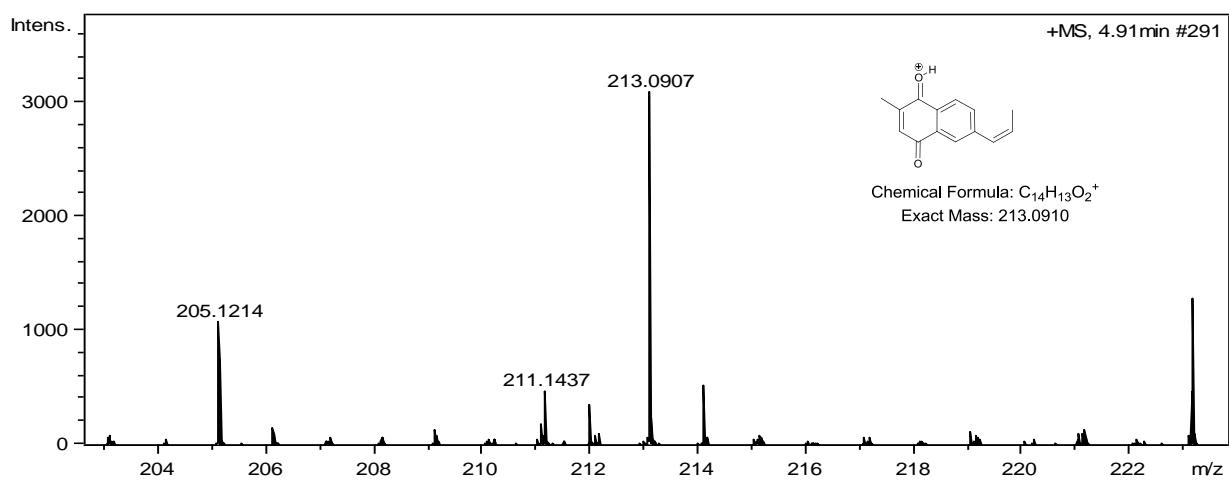
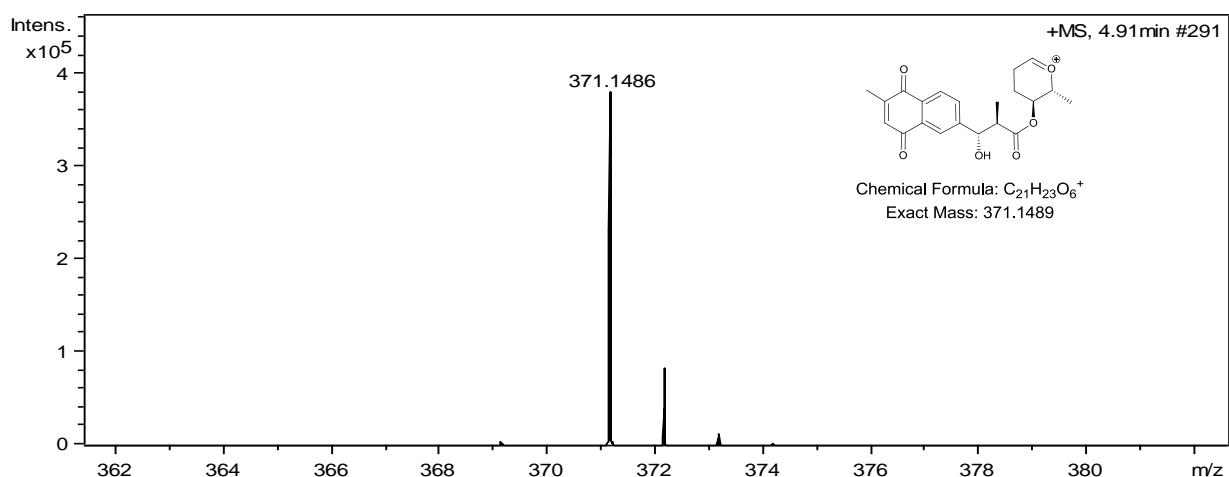
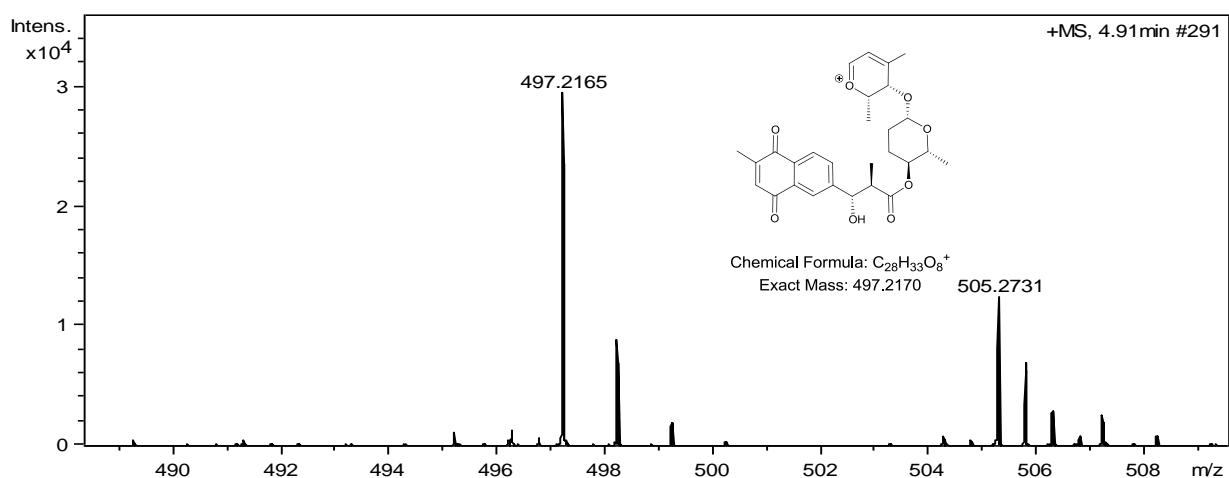
**Figure S42.** (+) ESI-TOF MS of caniferolide D (4) including expansions showing key in-source fragment ions.



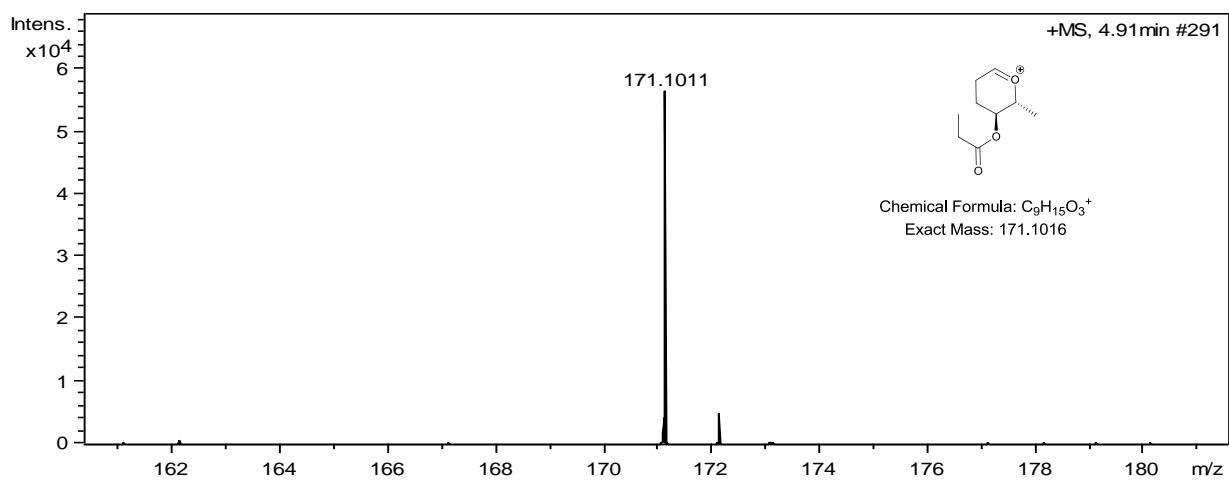
**Figure S42 cont.** (+) ESI-TOF MS of caniferolide D (**4**) including expansions showing key in-source fragment ions.



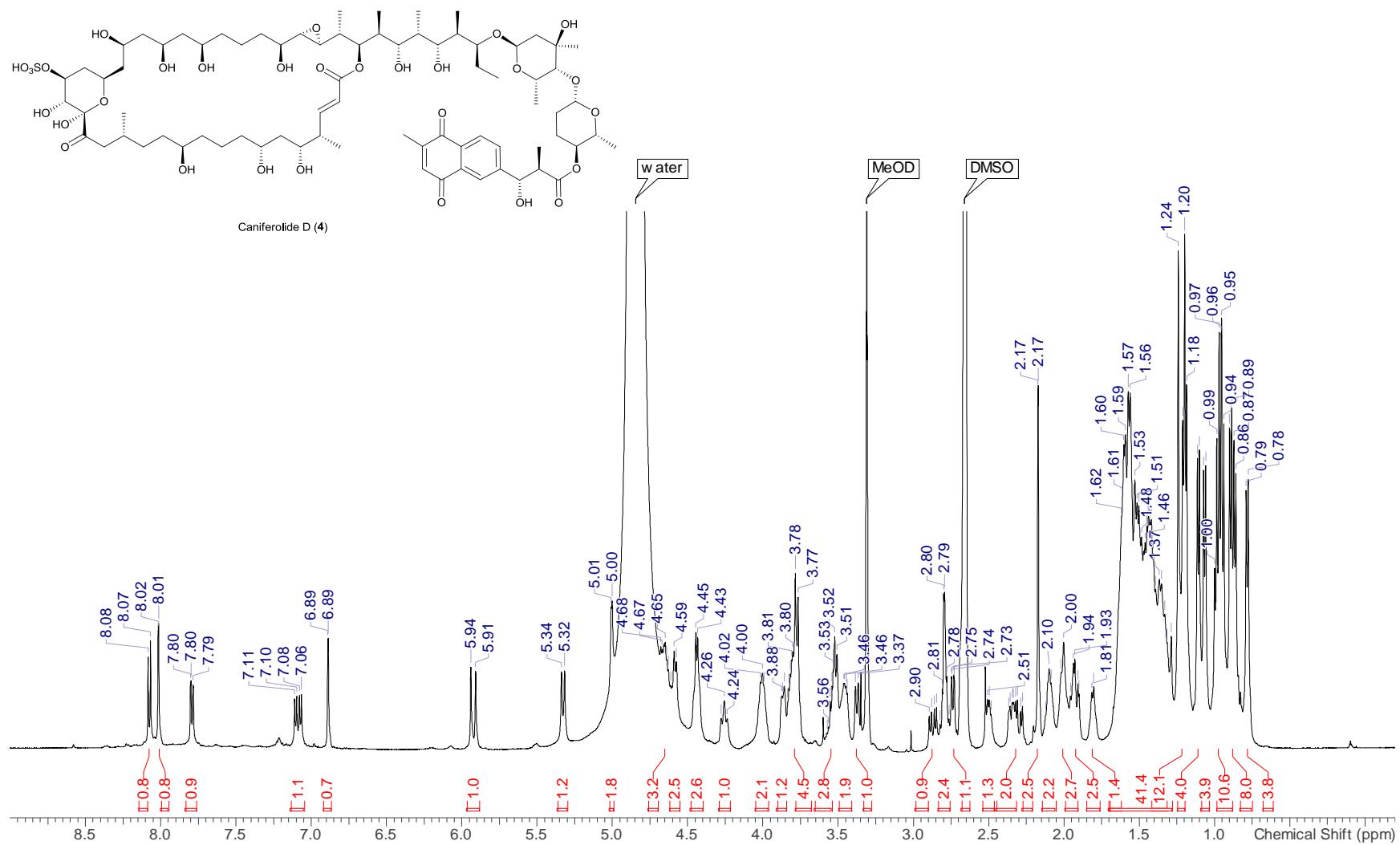
**Figure S42 cont.** (+) ESI-TOF MS of caniferolide D (**4**) including expansions showing key in-source fragment ions.



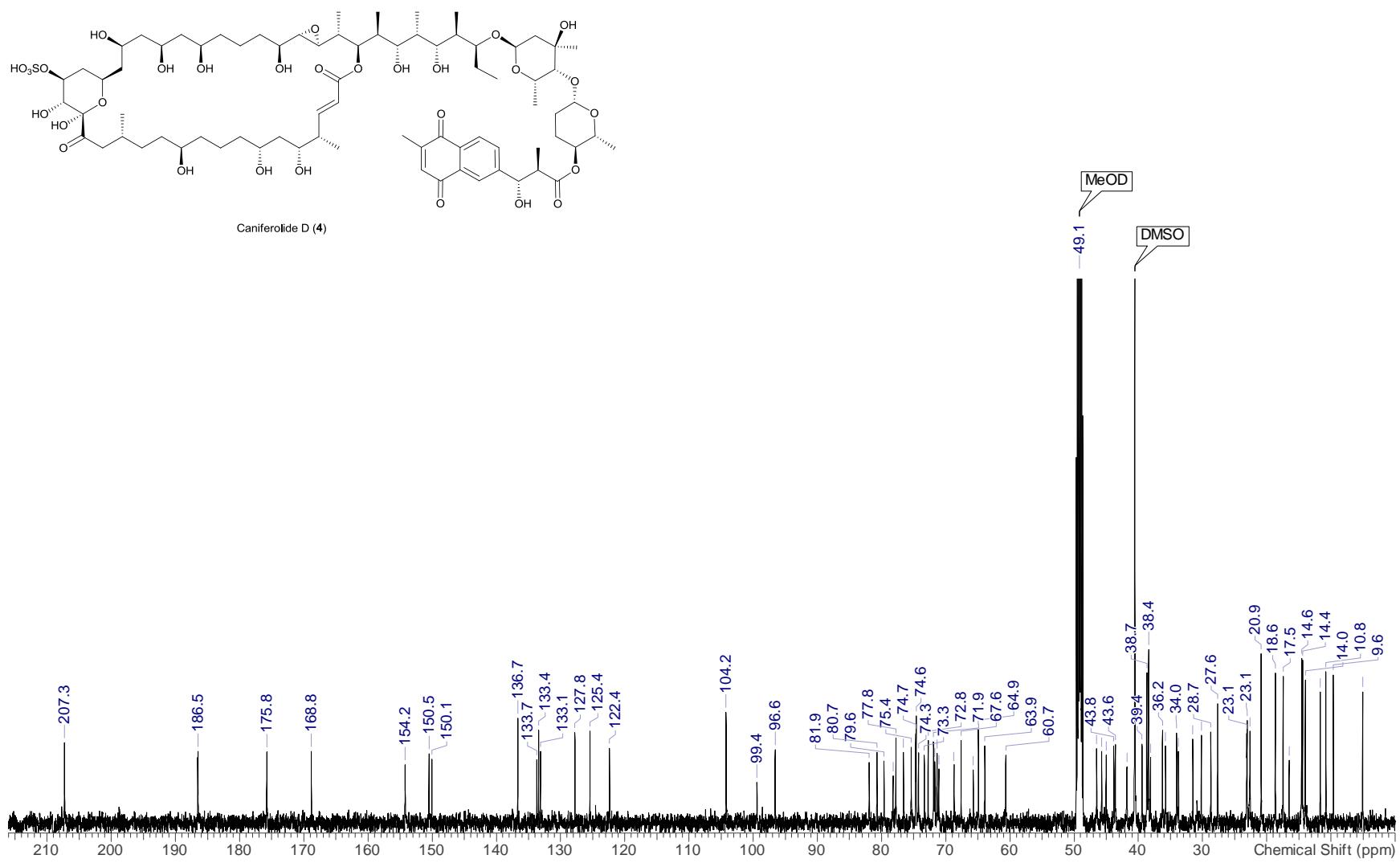
**Figure S42 cont.** (+) ESI-TOF MS of caniferolide D (**4**) including expansions showing key in-source fragment ions.



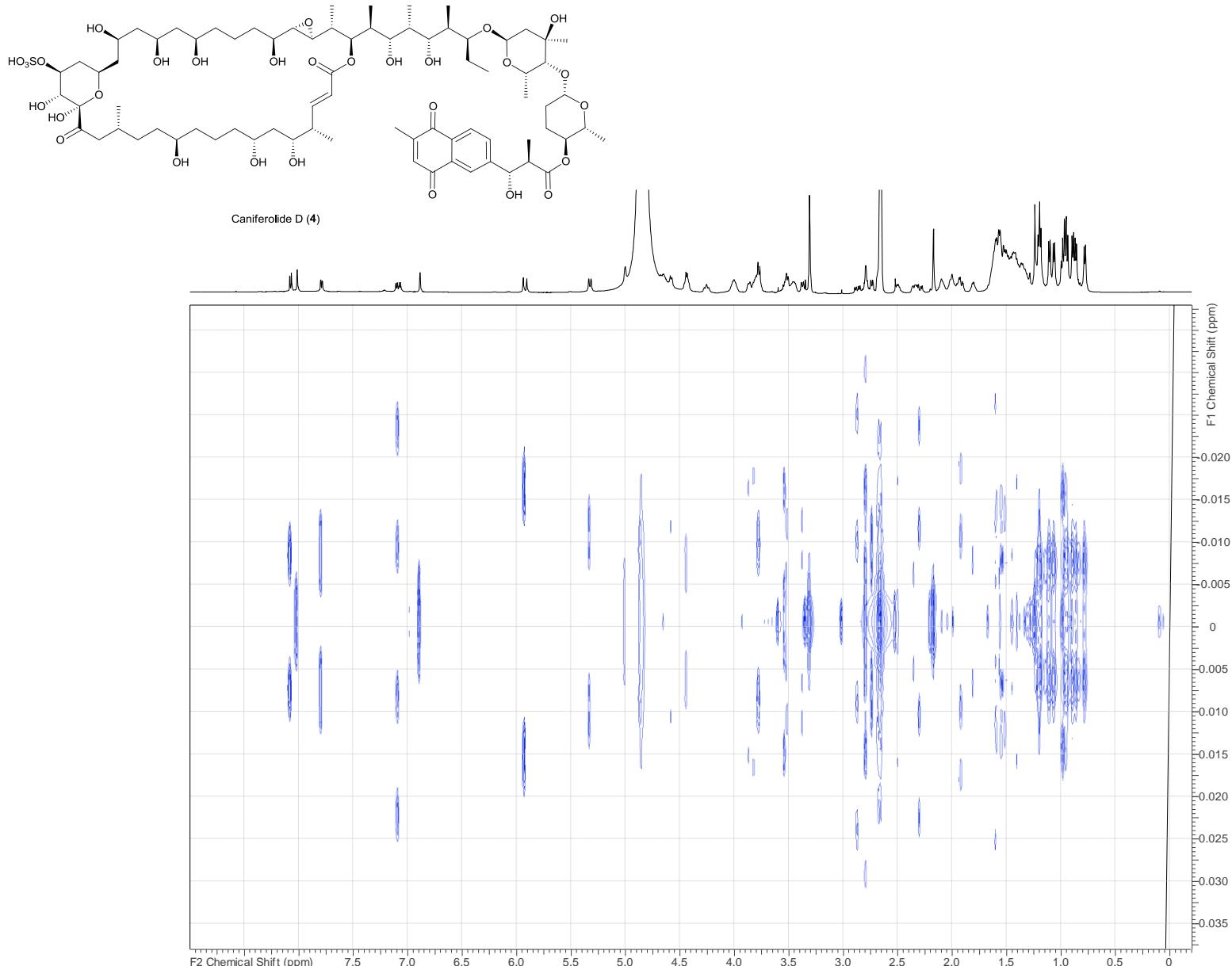
**Figure S42 cont.** (+) ESI-TOF MS of caniferolide D (**4**) including expansions showing key in-source fragment ions.



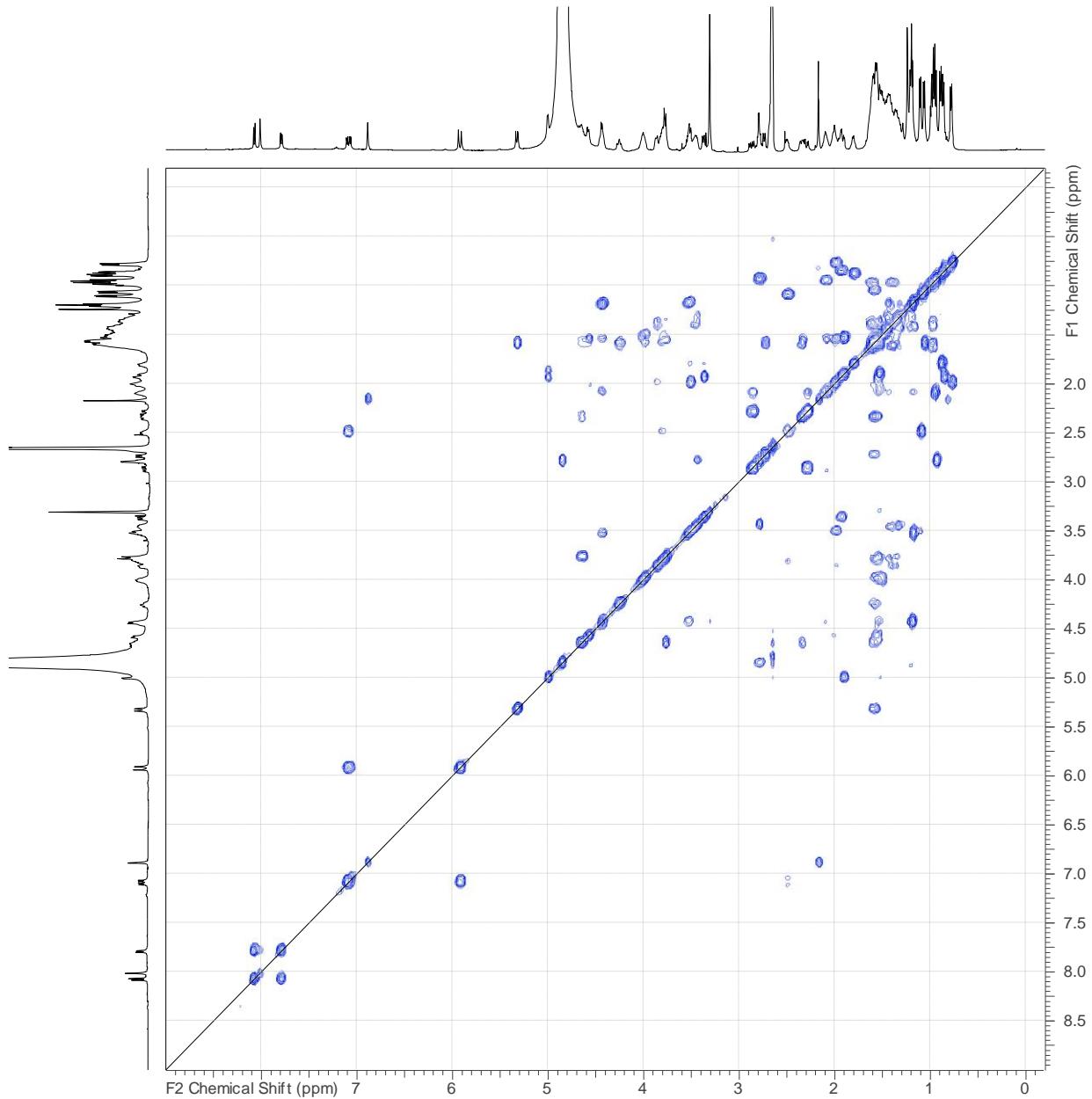
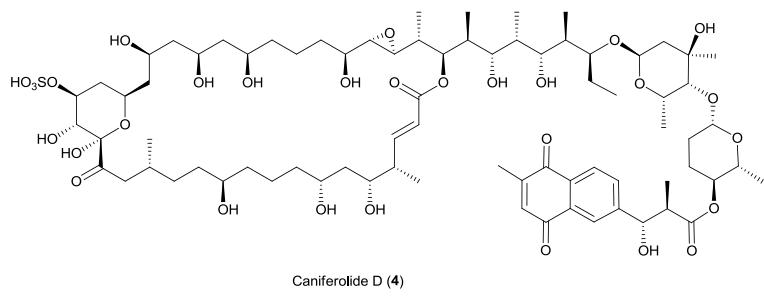
**Figure S43.**  $^1\text{H}$  NMR spectrum ( $\text{CD}_3\text{OD}$ , 500 MHz) of caniferolide D (**4**).



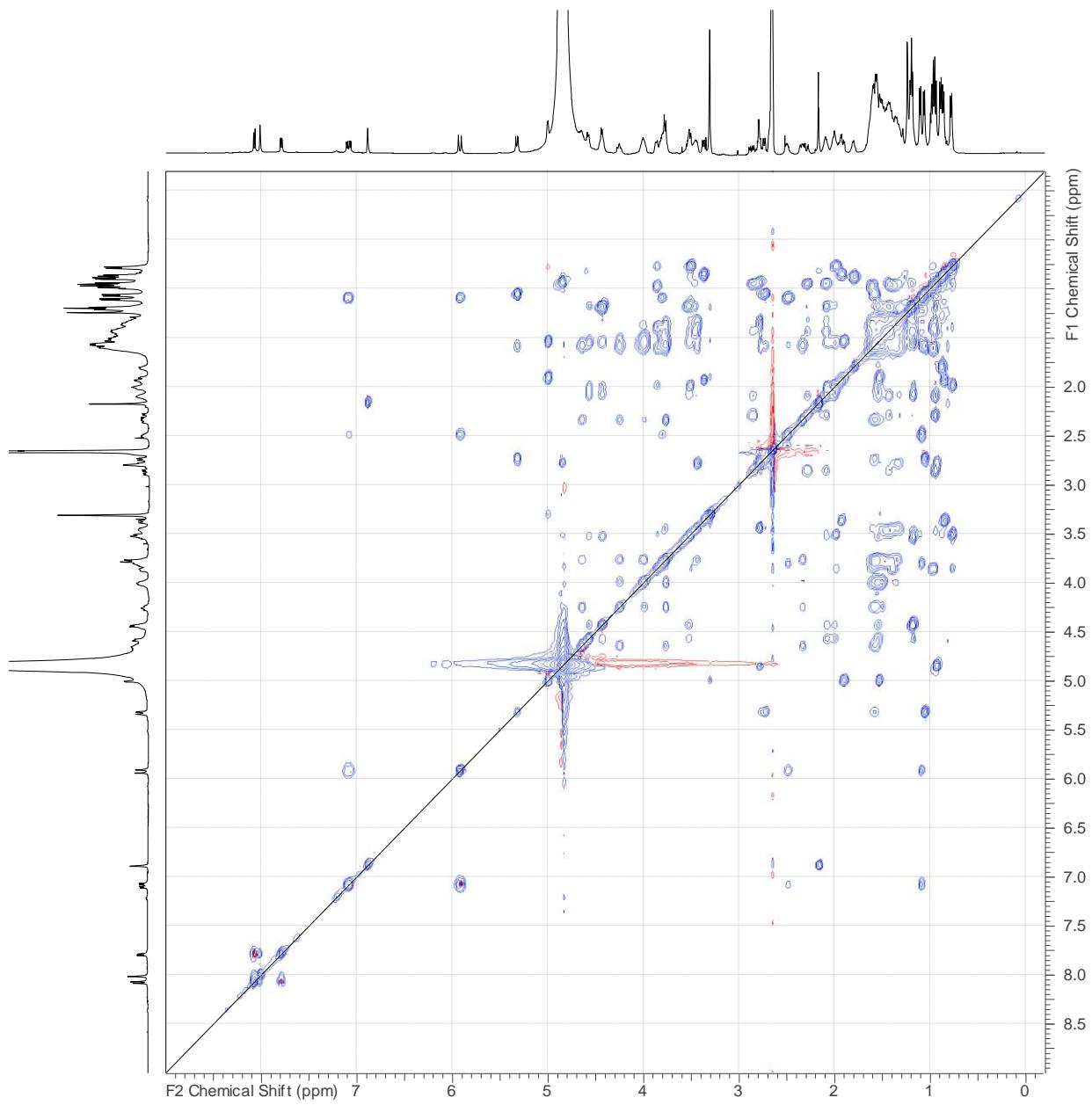
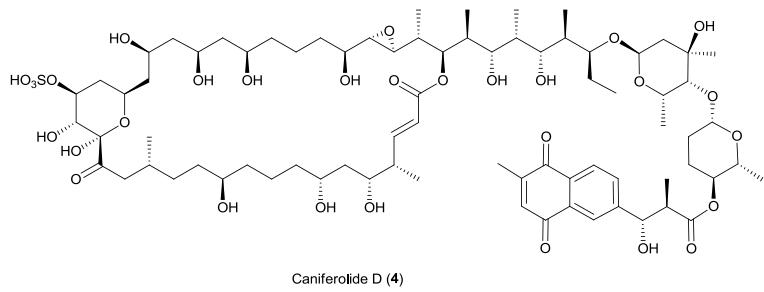
**Figure S44.** <sup>13</sup>C NMR spectrum (CD<sub>3</sub>OD, 125 MHz) of caniferolide D (**4**).



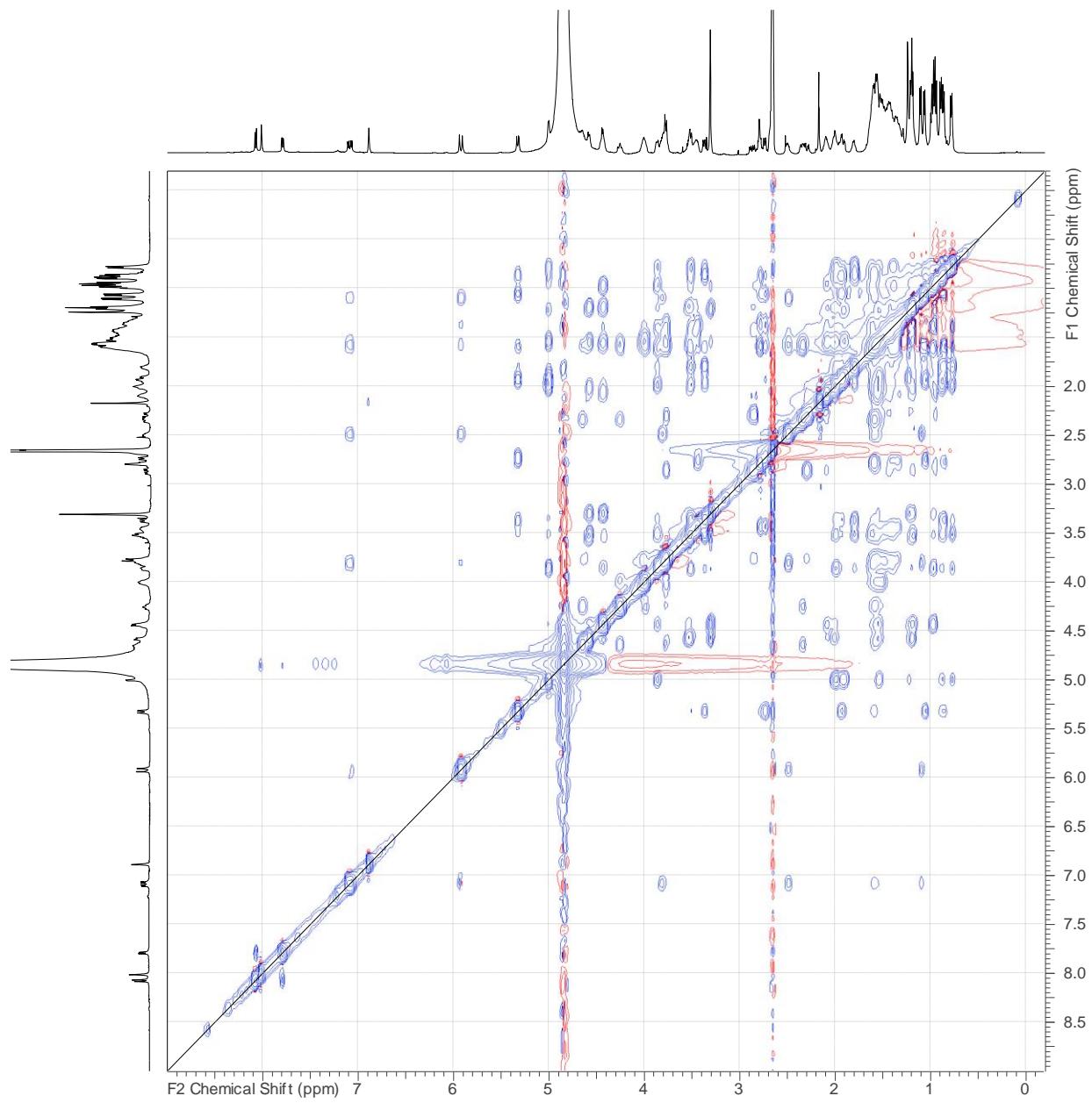
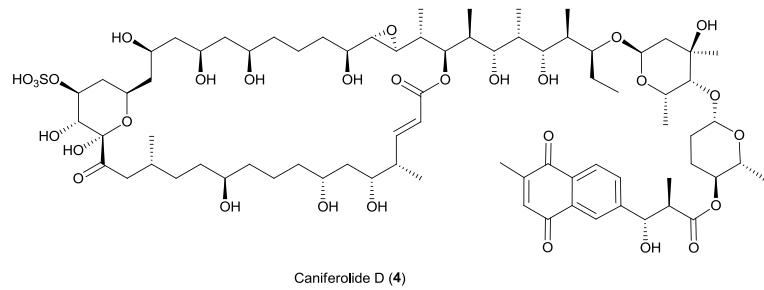
**Figure S45.** JRES spectrum of caniferolide D (**4**).



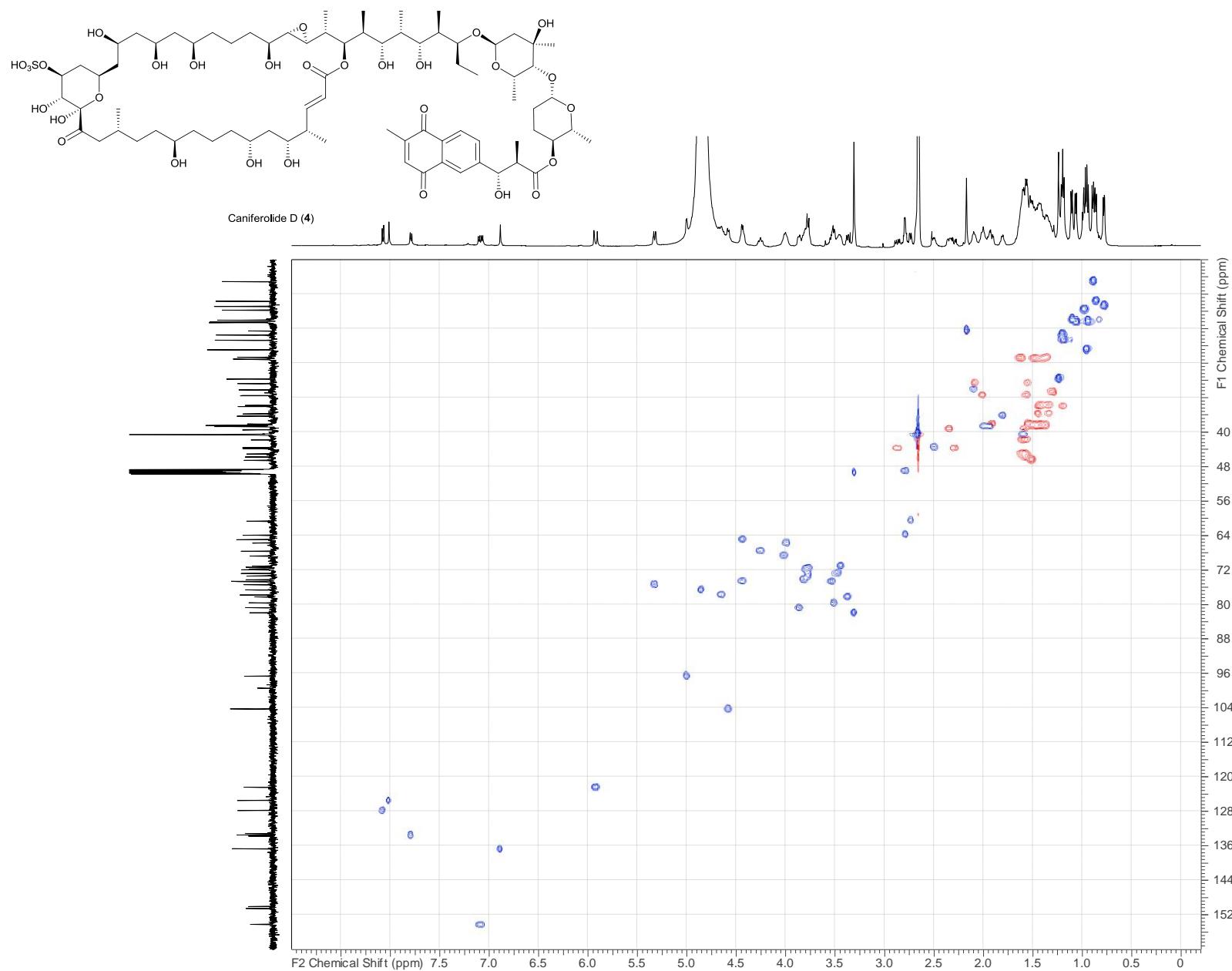
**Figure S46.** COSY spectrum of caniferolide D (**4**).



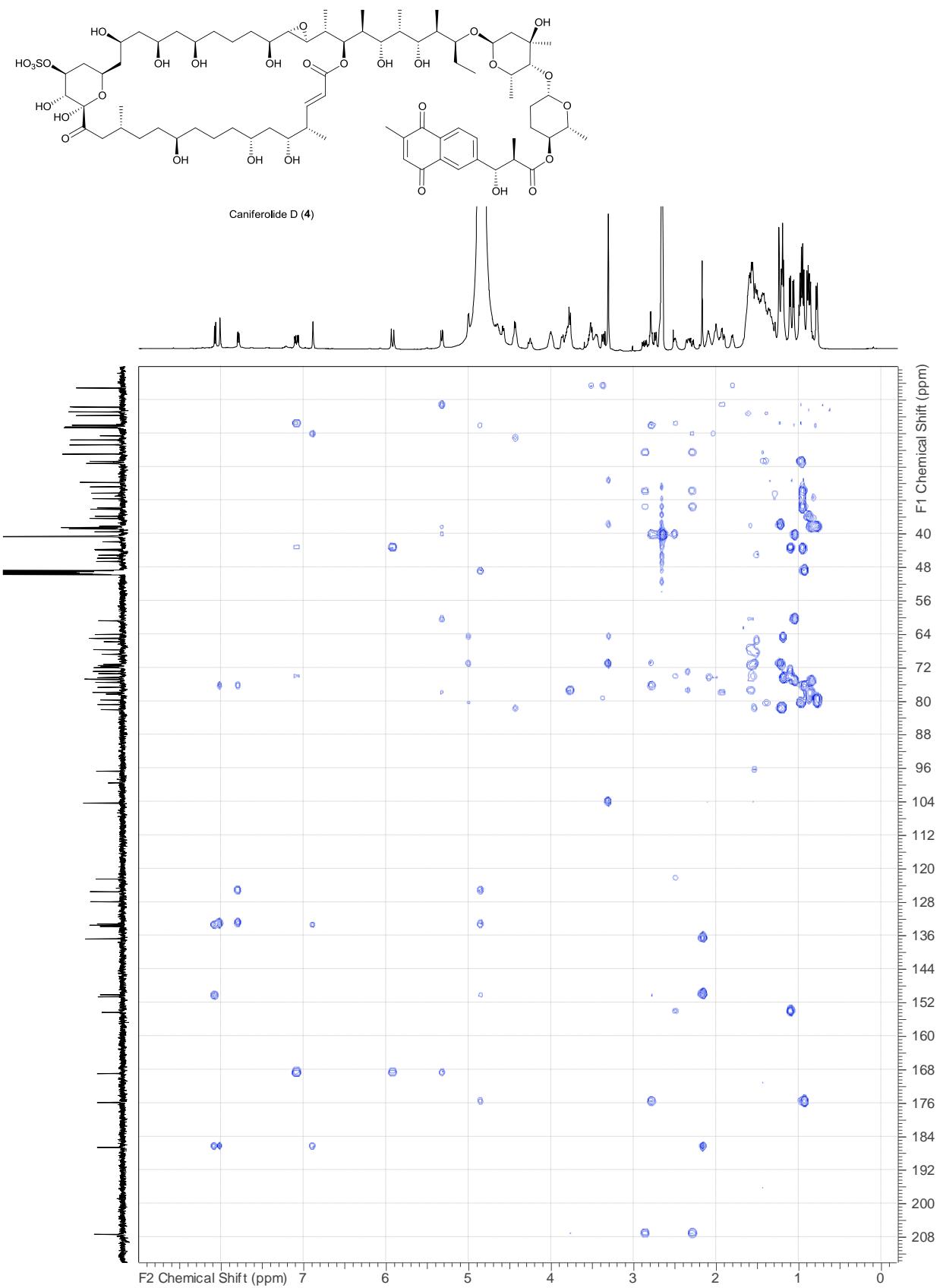
**Figure S47.** TOCSY spectrum of caniferolide D (**4**).



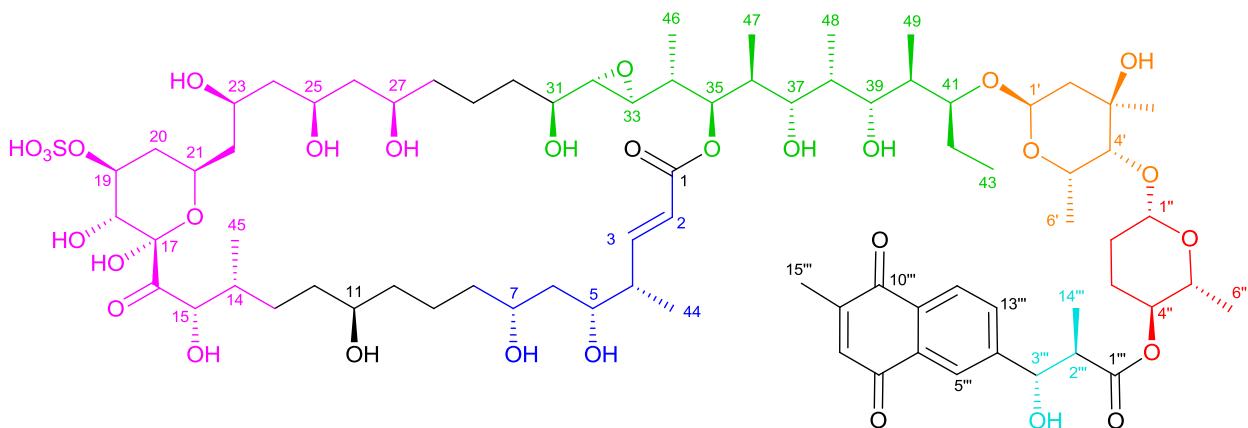
**Figure 48.** NOESY spectrum of caniferolide D (4).



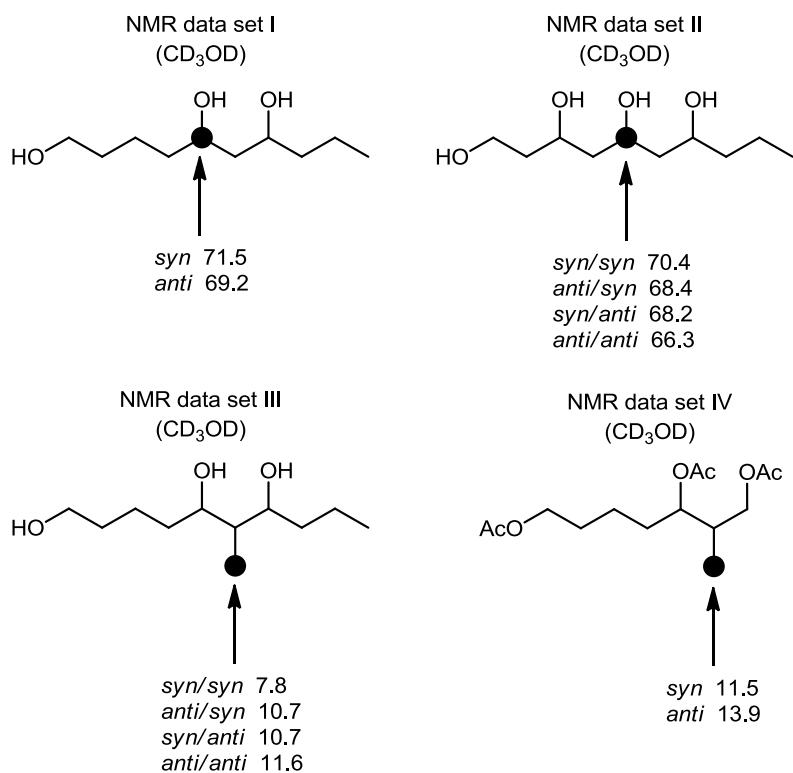
**Figure S49.** Edited HSQC spectrum of caniferolide D (4).



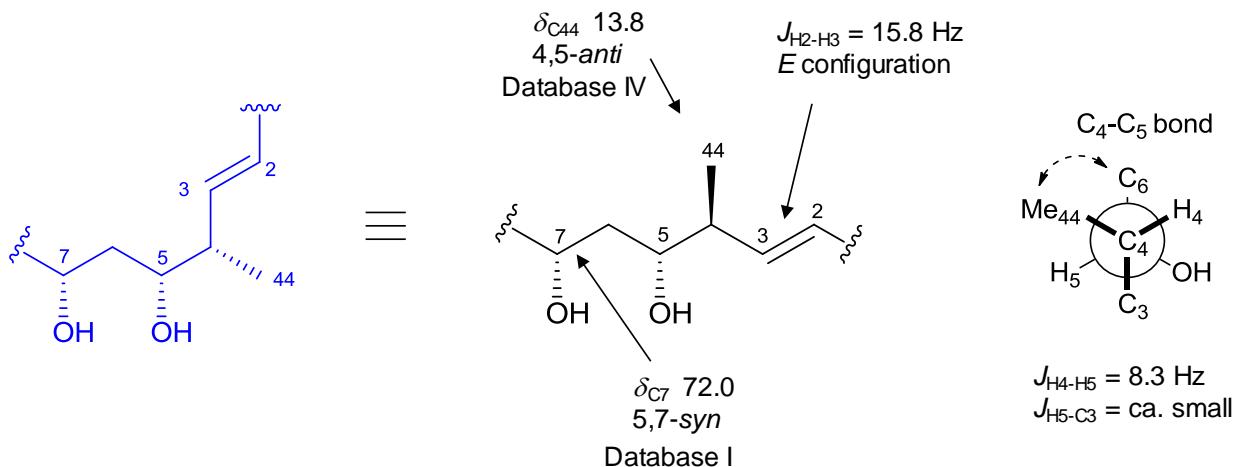
**Figure S50.** HMBC spectrum of caniferolide D (**4**).



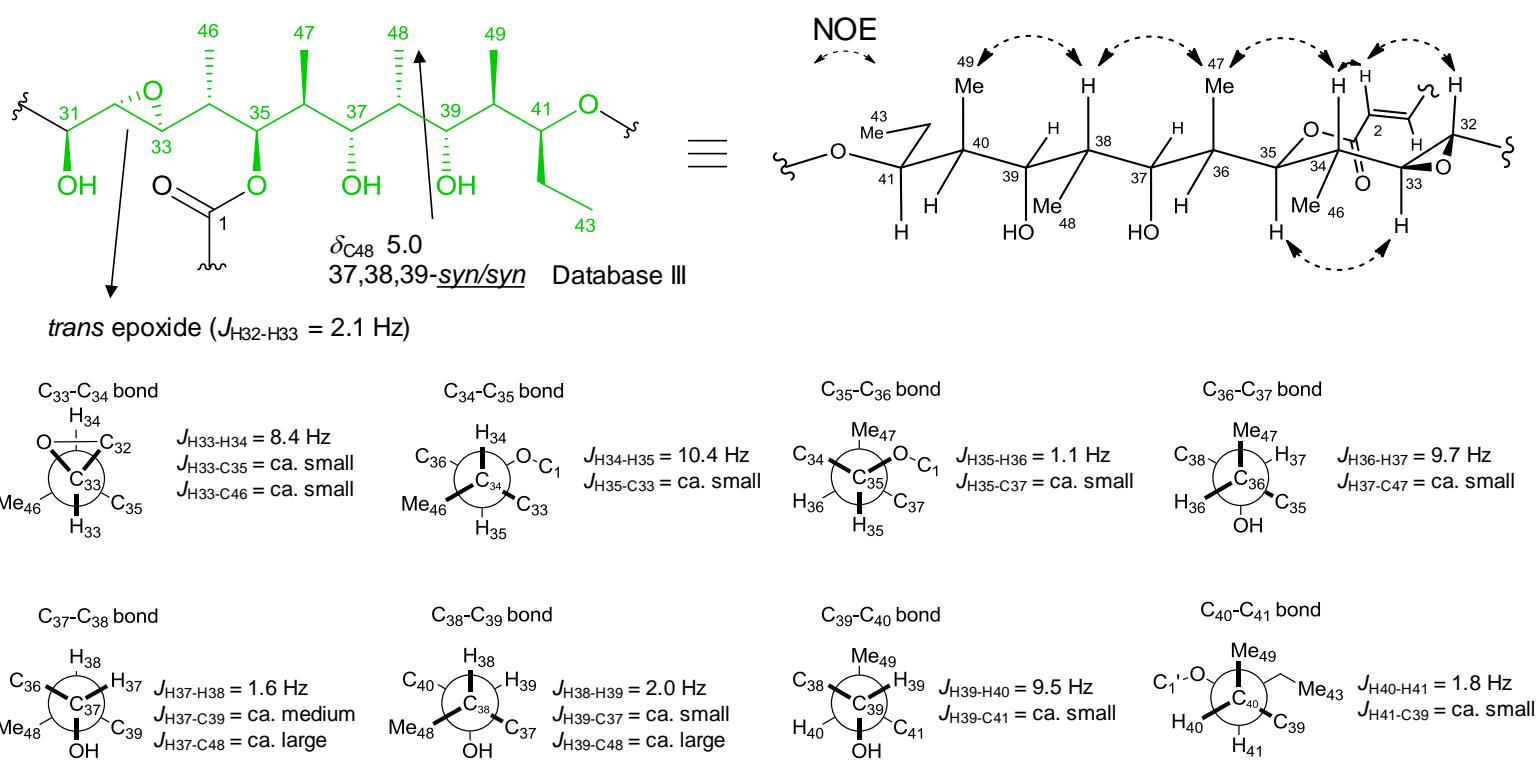
**Figure S51.** Structure of caniferolide A (**1**) highlighting with different color its stereoclusters. It was possible by NMR to relate the relative configuration of the carbohydrate stereoclusters between themselves and with the macrolide aglycon.



**Figure S52.** Data sets of Kishis's universal NMR database method for determining the relative configuration of 1,3-diol (I), 1,3,5-triol (II), 2-methyl-1,3-diol (III) and 1-methyl-2-alcohol (IV).

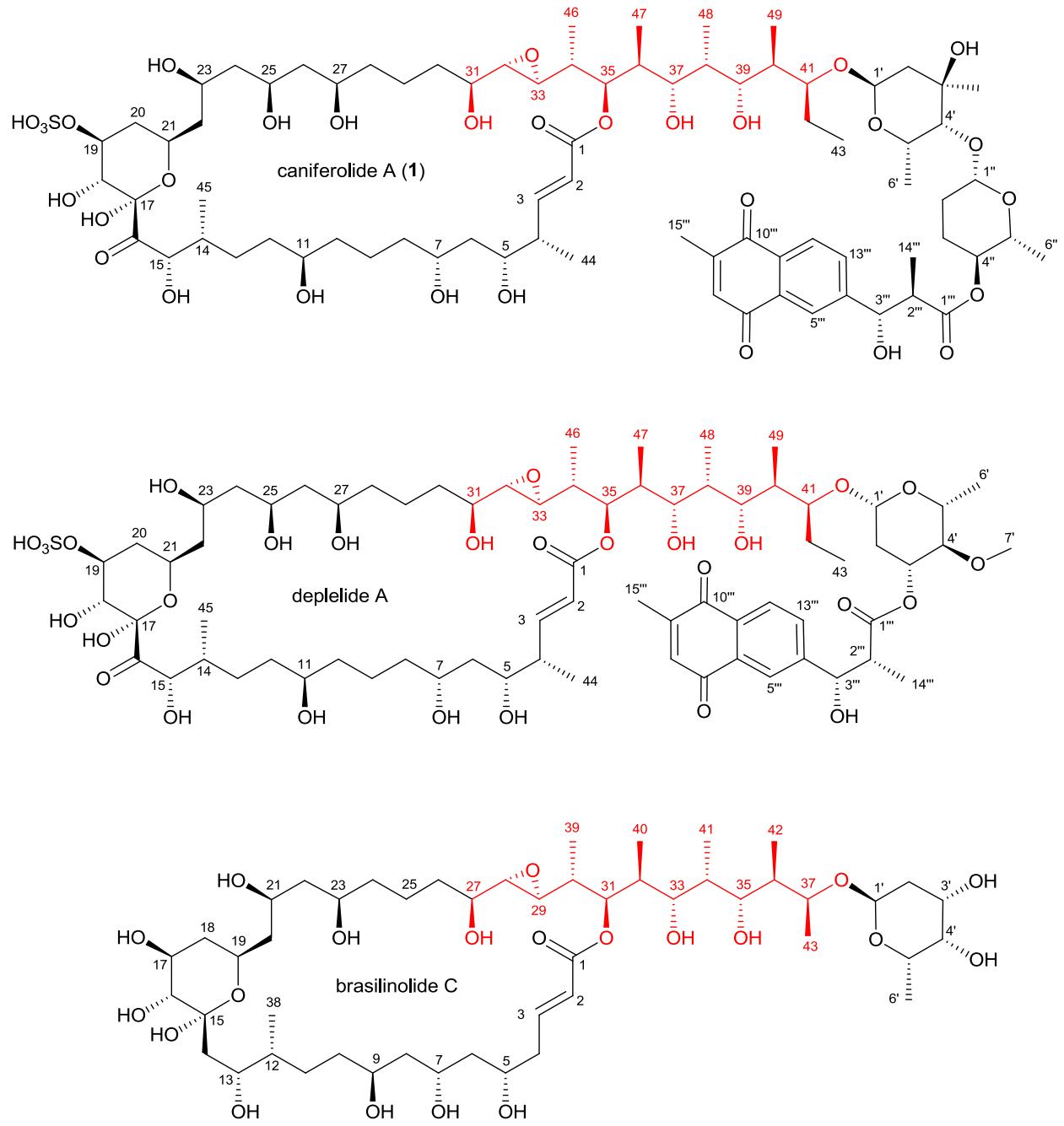


**Figure S53.** Determination of the relative configuration of the C-4 to C-7 stereocluster

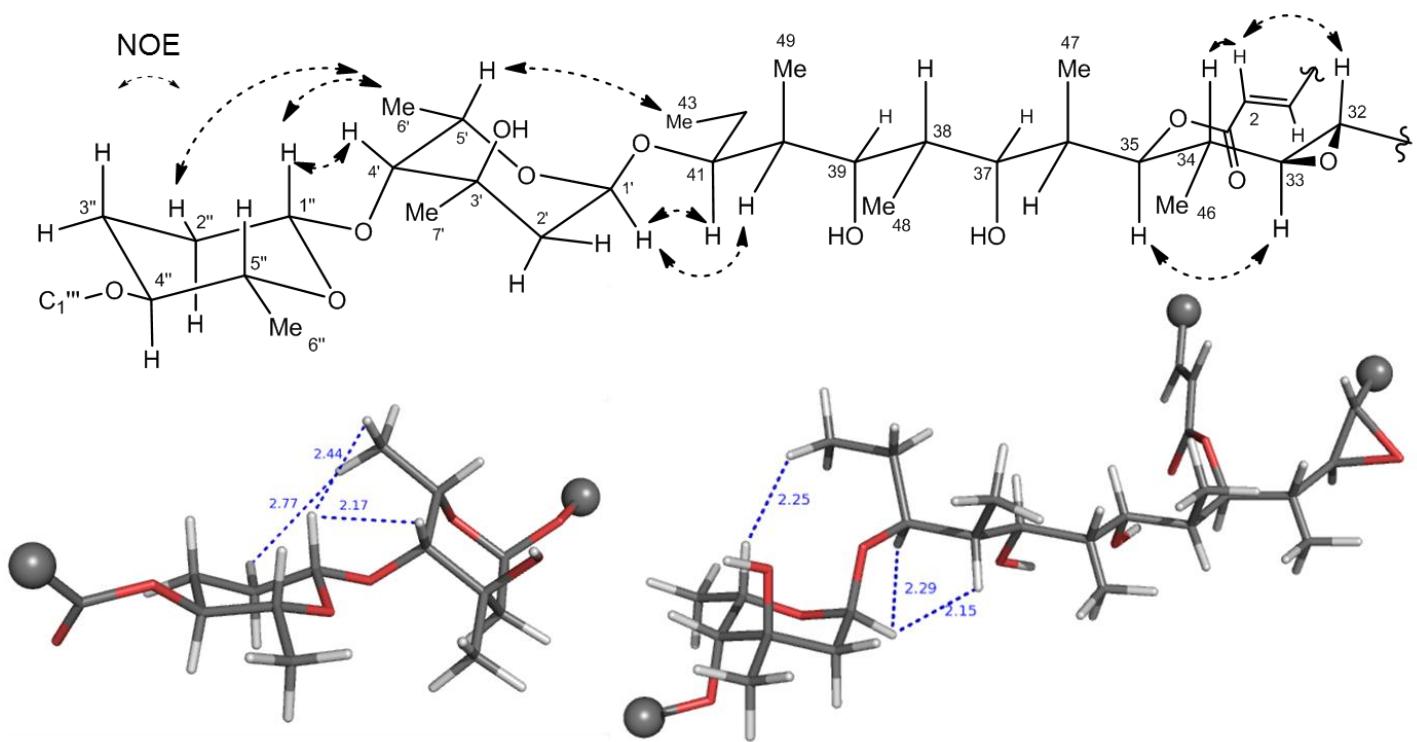


Relative configuration from C<sub>31</sub> to C<sub>41</sub> was corroborated by comparisons with NMR data of the C<sub>27</sub> to C<sub>37</sub> segment of brasiliolide C

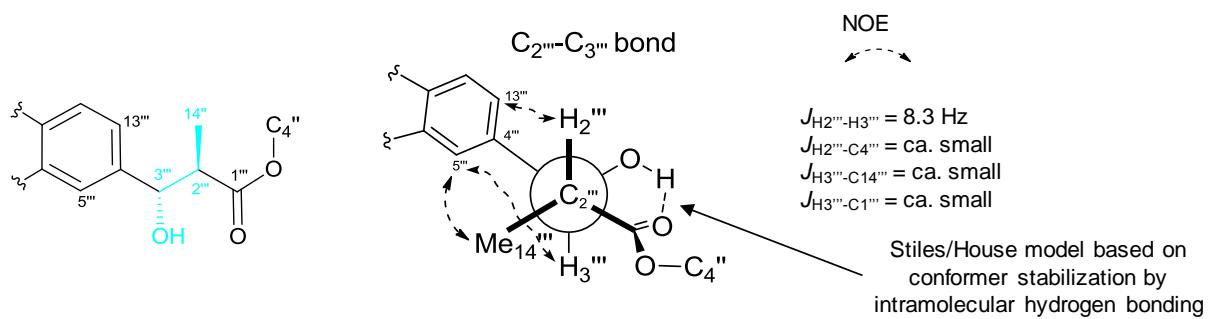
**Figure S54.** Determination of the relative configuration of the C-31 to C-41 stereocluster.



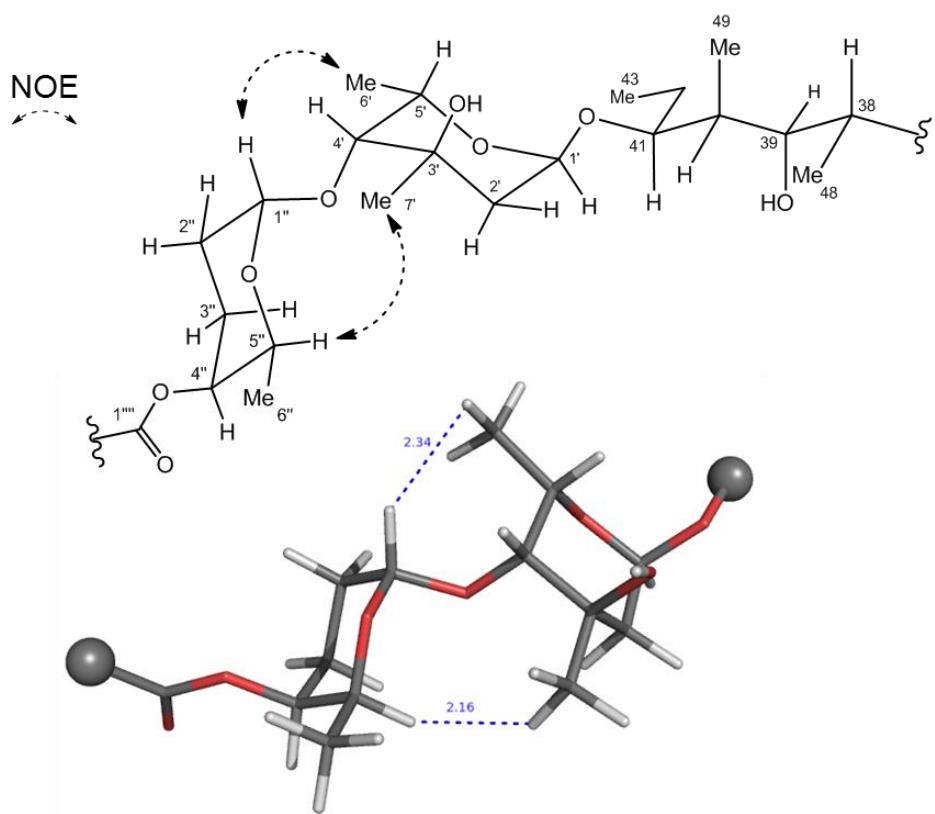
**Figure S55.** Structure of caniferolide A (**1**), deplelide A and brasiliolide C. The equivalent segment of brasiliolide C used for determination of the 31S configuration in both **1** and deplelide A is highlighted in red color.



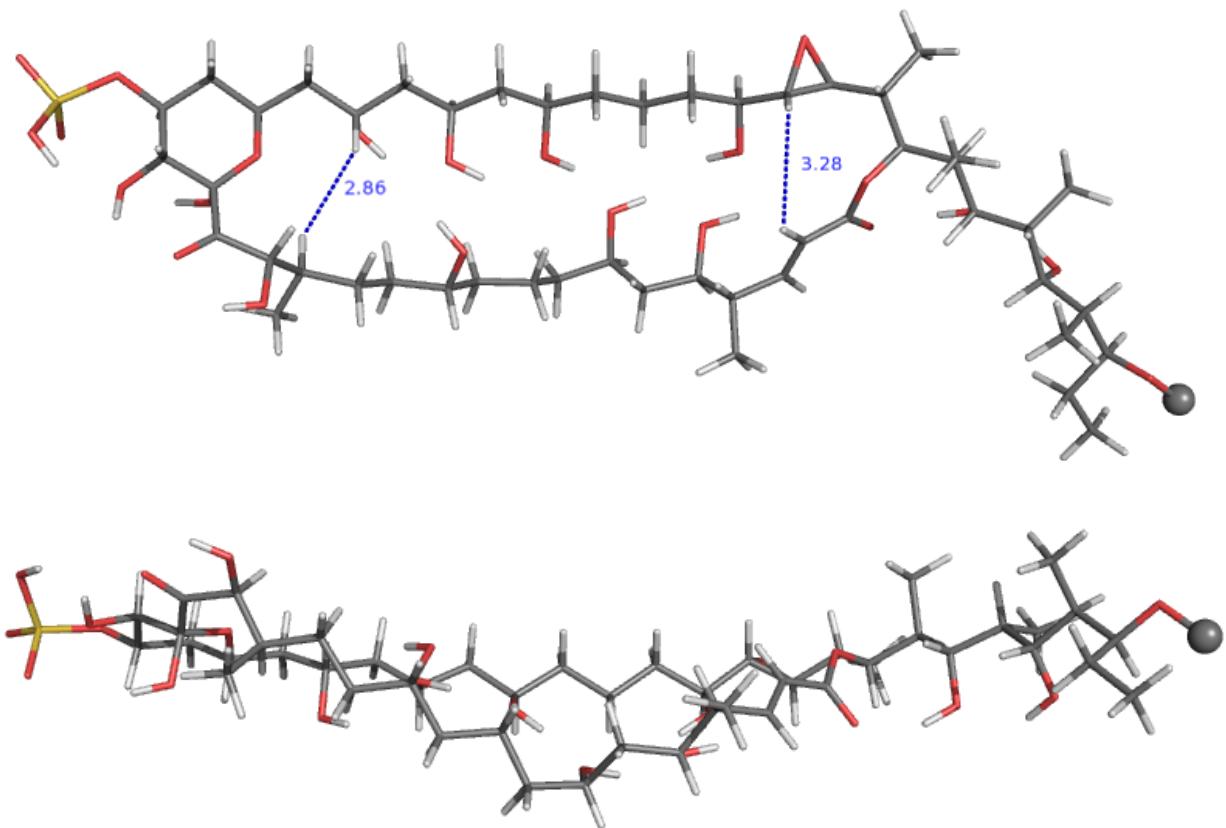
**Figure S56.** Truncated energy-minimized molecular models showing the distances associated with the key NOEs which connect the relative configuration of the carbohydrate stereoclusters between themselves and with the macrolide aglycon.



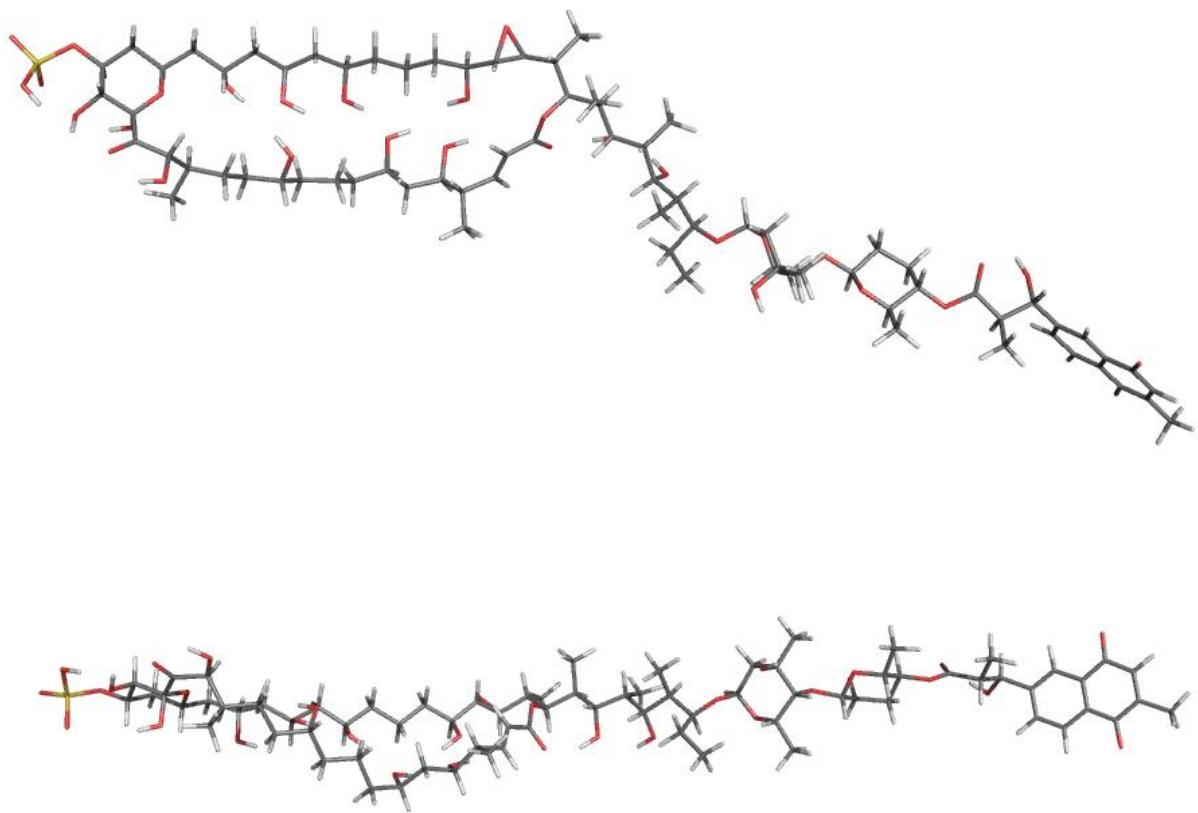
**Figure S57.** Determination of the relative configuration of the C-2''' to C-3''' stereocluster based on the Stiles-House rule.



**Figure S58.** Truncated energy-minimized molecular model showing the distances associated with the key NOEs originally reported for PM100118<sup>12</sup> which connect the relative configuration of the axenose and rhodinose stereoclusters demonstrating unambiguously the L absolute configuration of the rhodinose residue.



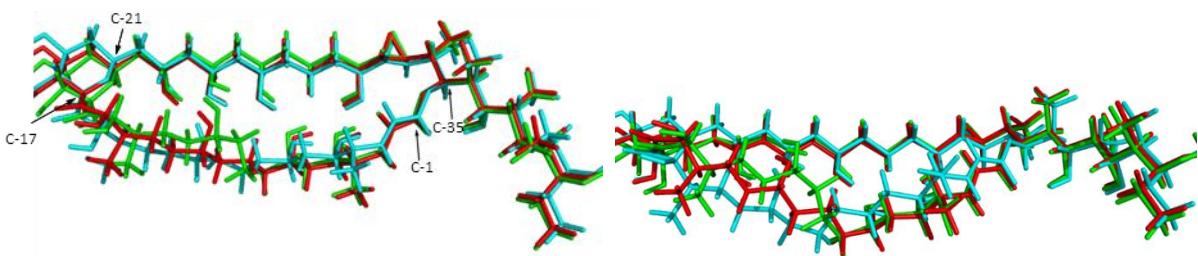
**Figure S59.** Energy-minimized molecular model (top and side views) of caniferolide A (**1**) aglycon satisfying the NOE correlations,  $^3J_{HH}$  and  $^3J_{CH}$  employed for determining the relative configuration. The model displays the final absolute configuration corroborated after bioinformatic gene cluster analysis. Two weak key NOEs connecting the two main linear segments of the macrocycle are highlighted.



**Figure S60.** Energy-minimized molecular model (top and side views) of caniferolide A (**1**) satyisfying the NOE correlations,  $^3J_{HH}$  and  $^3J_{CH}$  employed for determining the relative configuration.

## Limitations on the application of Matsunaga's empirical rule for establishing the relative stereochemistry of the 1,5-diol motifs in the caniferolides

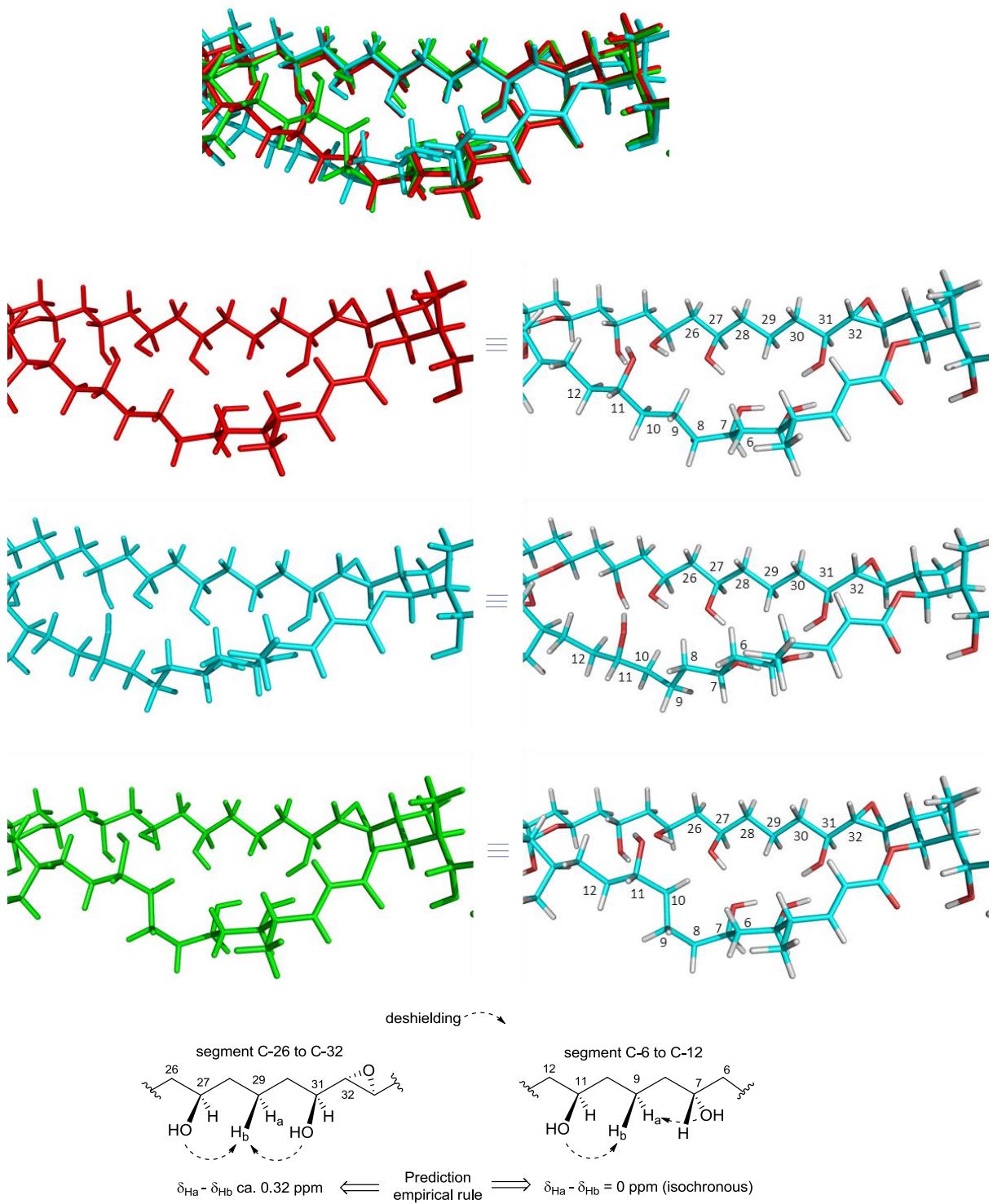
Matsunaga et al. have proposed an empirical rule to assign the relative stereochemistry of linear 1,5-diols.<sup>13</sup> However, its application when the 1,5-diol motif is embedded within a macrocycle must be taken cautiously and not surprisingly has not been employed to date. A reliable application of this empirical rule for such cases requires that the 1,5-diol motif is contained within a linear and fully extended segment of the macrocycle. In the case of the caniferolides, there are two main (pseudo)linear regions emanating from the embedded tetrahydropyran ring within the macrocycle. The first one comprises 15 carbon atoms, from C-21 to C-35 while the second one comprises 17 carbon atoms, from C-1 to C-17. This difference in length (two carbon atoms) between these two pseudolinear regions translates into a different global conformation for each segment when modelling the 3D structure of the macrocyle. The mentioned difference in global conformation can be easily observed just comparing the 3D structure (Fig. S61) of three low-energy conformers compatible with the NMR data (coupling constants and NOESY correlations).



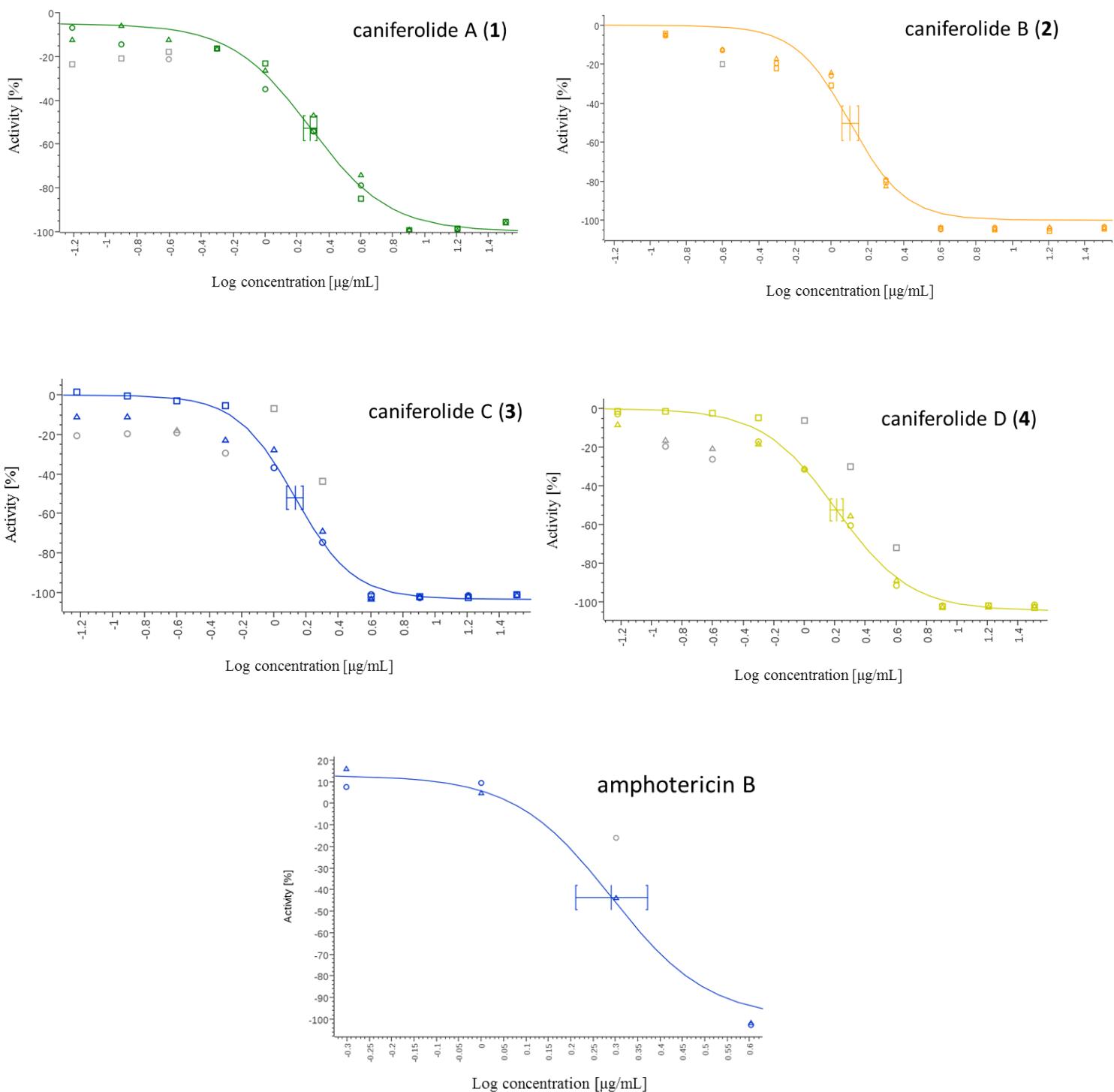
**Figure S61.** Overlay (top and side views) of three low-energy conformers (represented in red, cyan and green color) of caniferolide A (**1**) compatible with the NMR data. Sulfate and sugars are not included for the sake of clarity.

In Fig. S61 it can be clearly seen that the conformation of the region from C-21 to C-35 is homogeneous and essentially identical among the different low-energy conformers, being basically a linear extended carbon chain. However, the conformation of the region from C-1 to C-17 is different in each low-energy conformer.

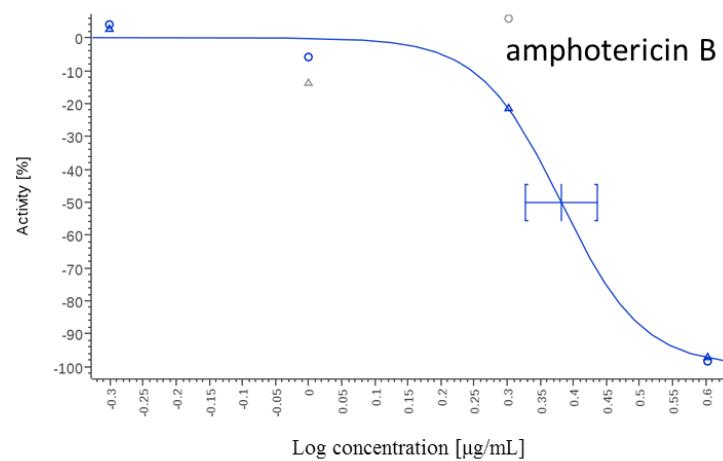
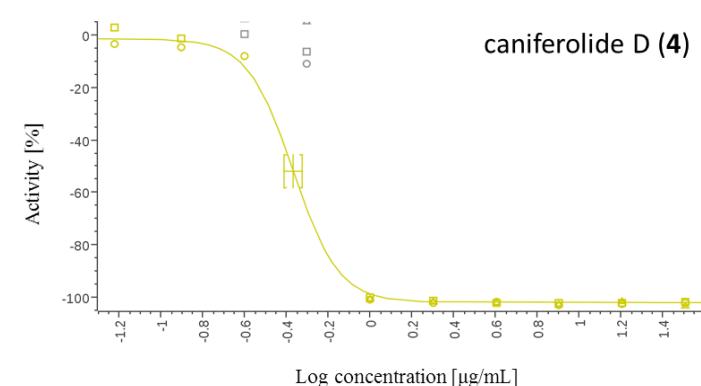
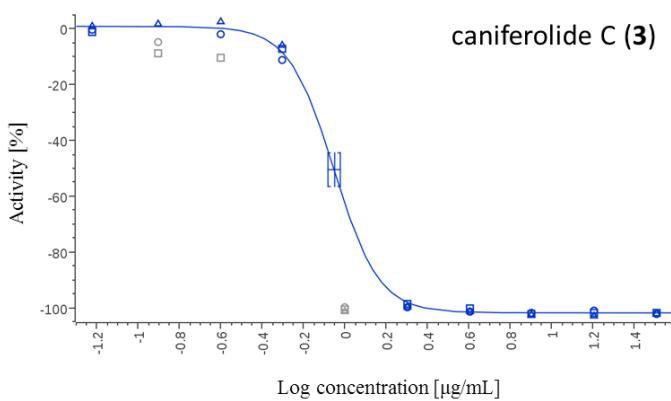
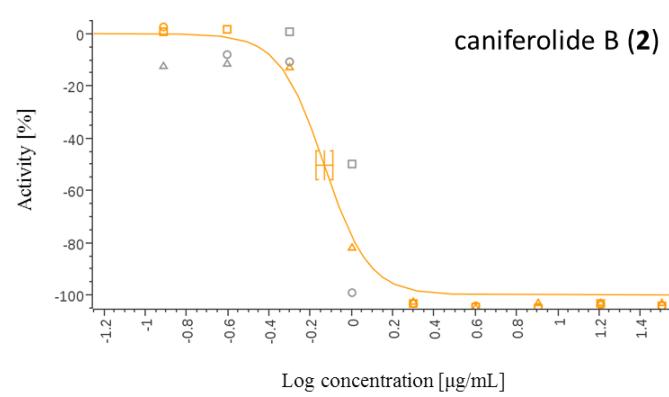
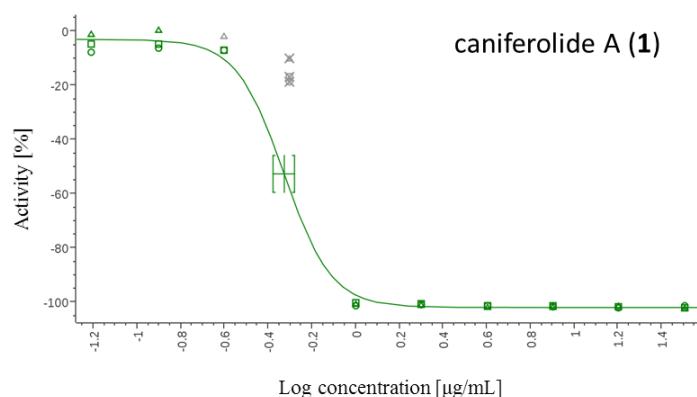
Fig. S62 show how this translates for the two 1,5 diol motifs present in the caniferolides (C-7/C-11 and C-27/C-31). For the diol contained in the C-26 to C-32 segment the requirement of a full extended linear conformation for applying the empirical rule is met in each low-energy conformer. Consequently, methylene protons H-29<sub>a</sub> and H-29<sub>b</sub> resonate at 1.37 and 1.63 ppm, having a difference of 0.26 ppm in agreement with the prediction of the empirical rule for *syn* 1,5-diols. However, for the diol contained in the C-6 to C-12 segment the rule cannot be applied since in any of the conformers there is a turn which avoids a fully extended linear conformation. For this reason, protons H-9<sub>a</sub> and H-9<sub>b</sub> are not isochronous, as the rule of Matsunaga predicts, but resonate at 1.40 and 1.51 ppm, having a difference of 0.11 ppm which cannot be employed alone to unambiguously establish the relative configuration (*syn* or *anti*) of this 1,5-diol (at positions C-7 and C-11).



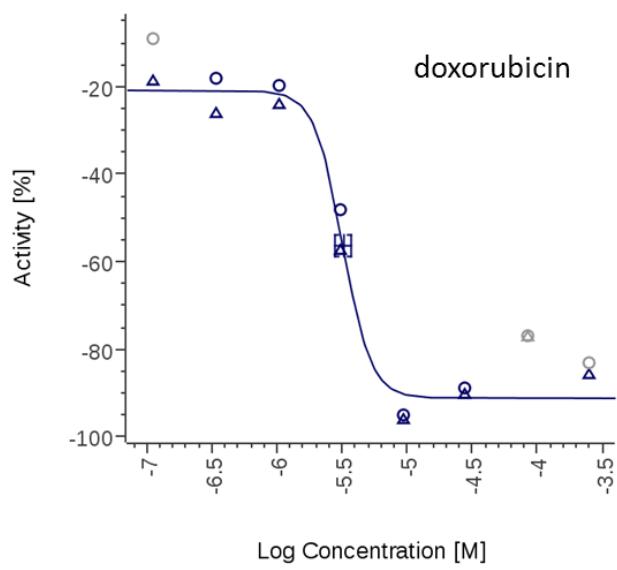
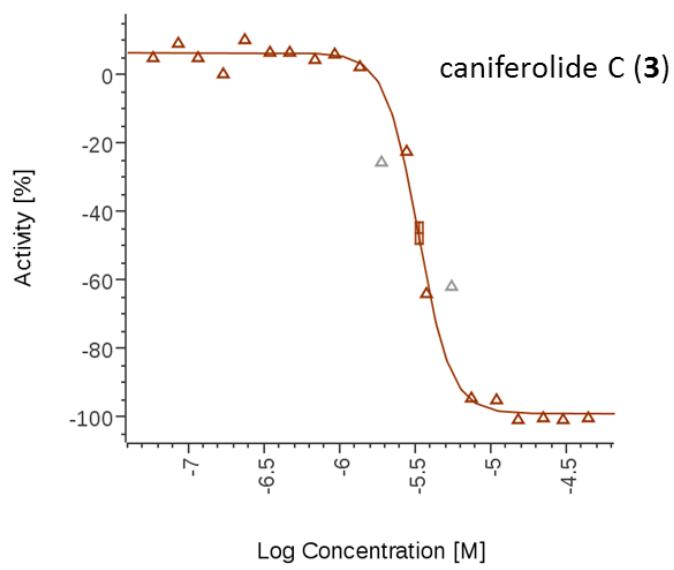
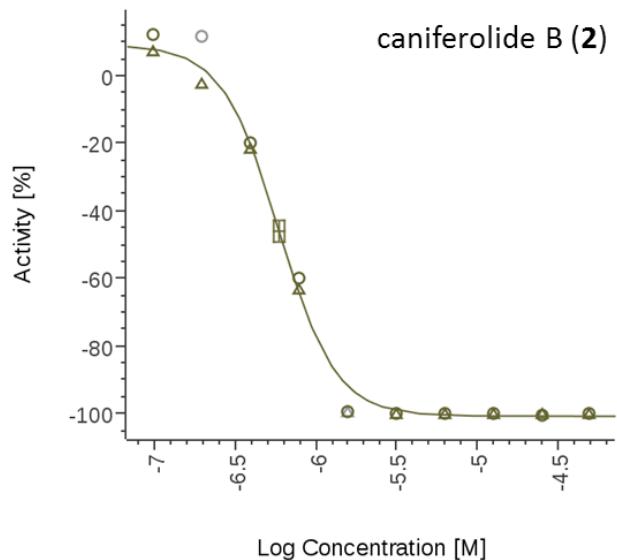
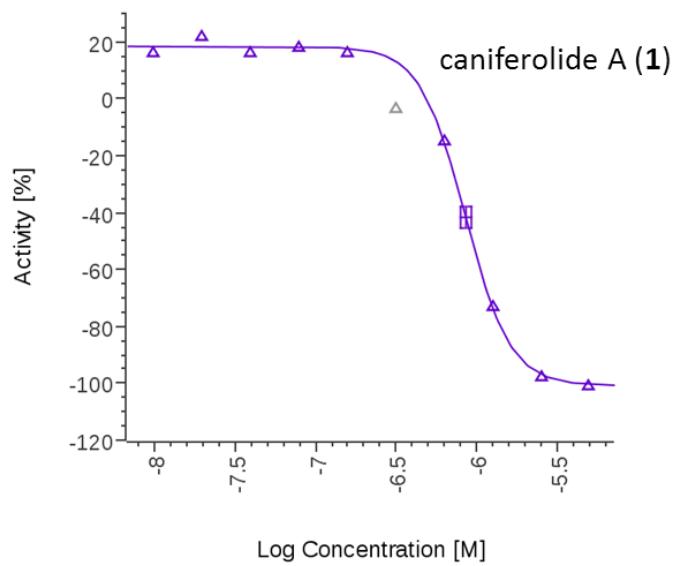
**Figure S62.** Expansion of the overlay of three low-energy conformers of caniferolide A and numbering of the key carbons corresponding to the two segments (C-6 to C-12 and C-26 to C-32) containing the 1,5-diol motifs in each conformer. For these segments, the ideal extended linear conformation that would allow an application of Matsunaga's empirical rule to assign the relative stereochemistry of these 1,5-diol motifs is sketched in the lower part of the figure. In all conformers, the C-26 to C-32 segment meets the requirements for applying the empirical rule with confidence, however none of the conformers show an extended conformation for the C-6 to C-12 segment which, in all cases, contains a turn which makes this segment incompatible with the empirical rule.



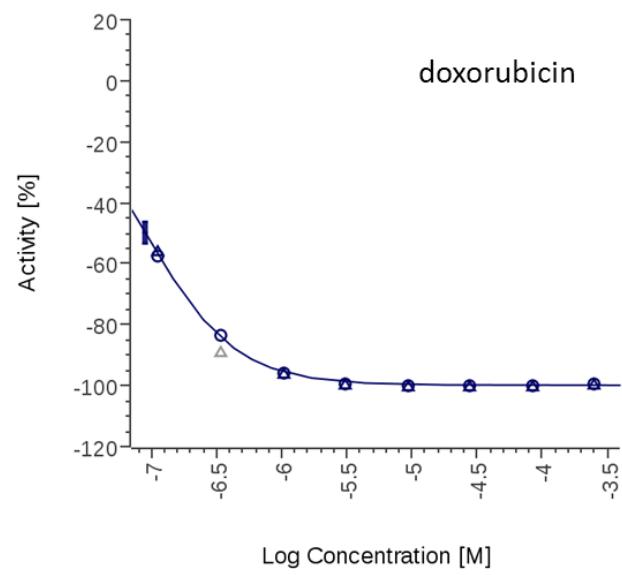
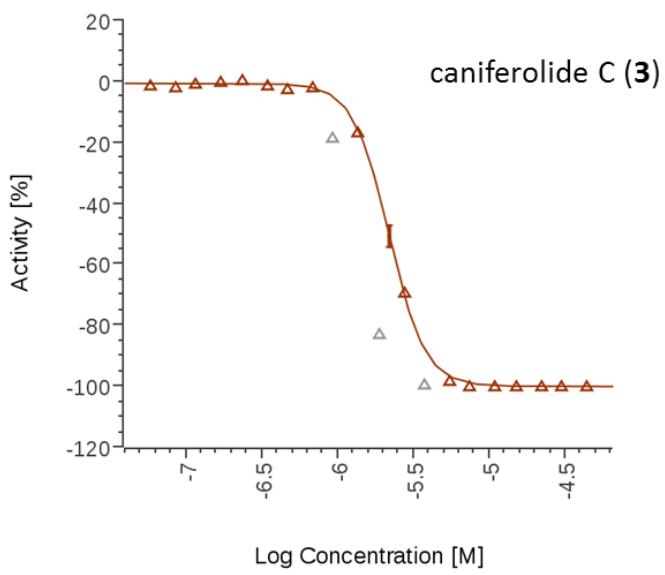
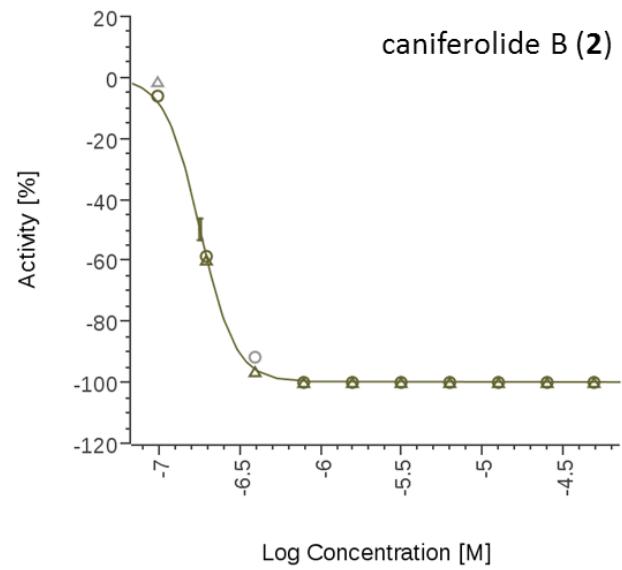
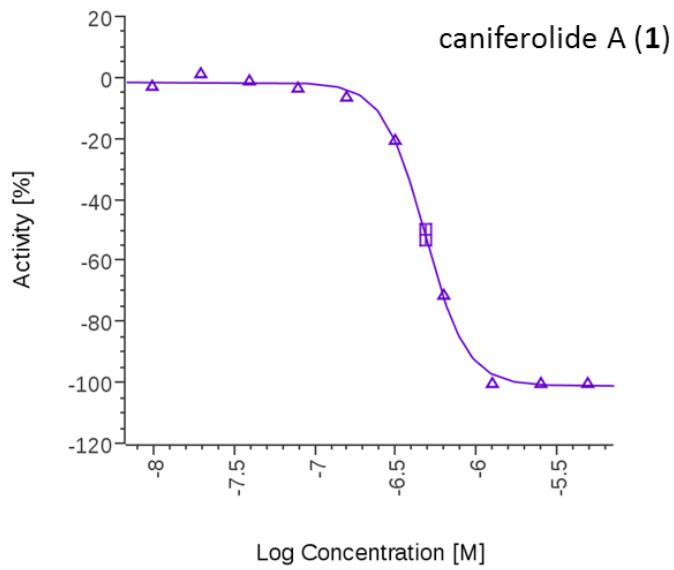
**Figure S63.** Growth inhibition curves of **1-4** and amphotericin B against *Aspergillus fumigatus* ATCC46645.



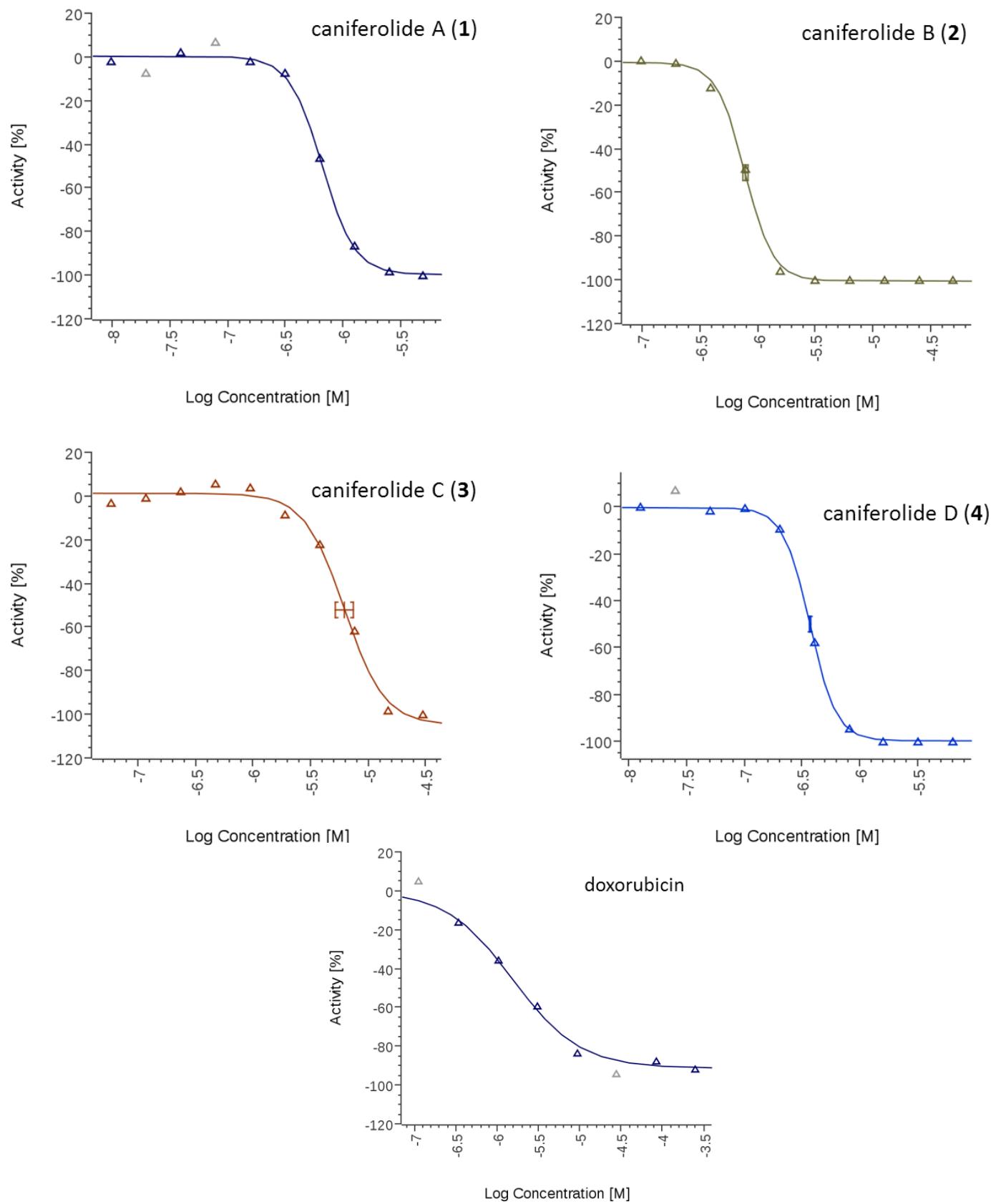
**Figure S64.** Growth inhibition curves of **1-4** and amphotericin B against *Candida albicans* MY1055.



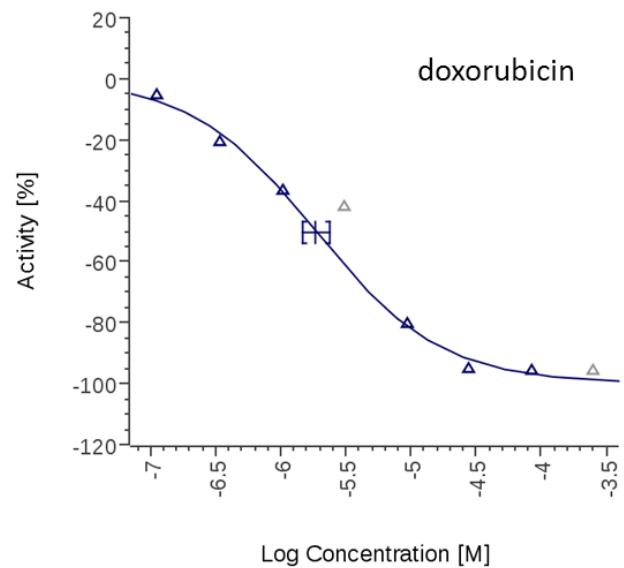
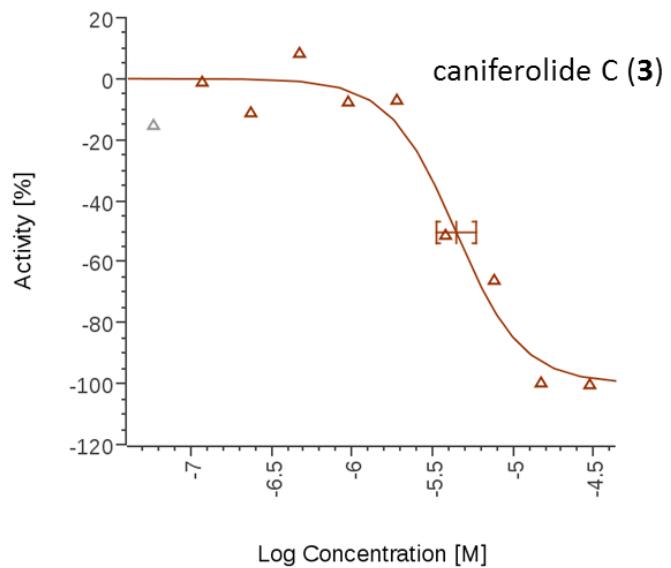
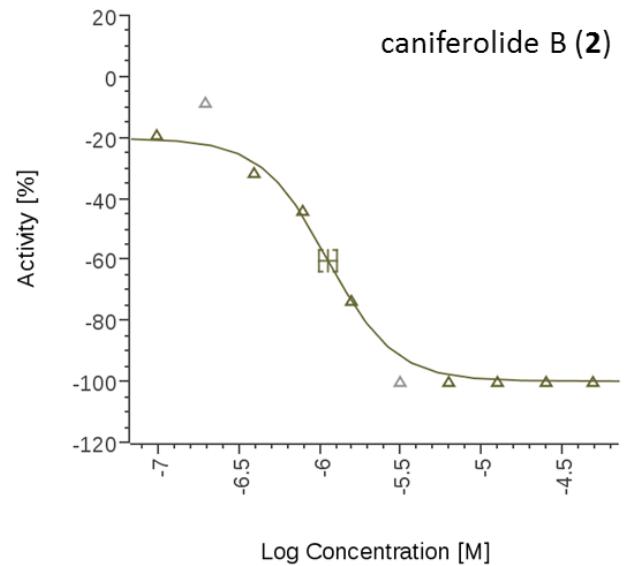
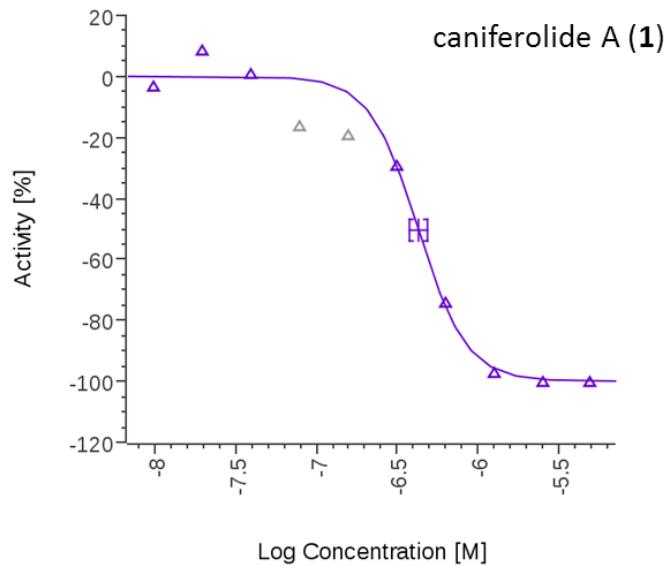
**Figure S65.** Antiproliferative activity curves of **1-3** and doxorubicin against A549 cell line (lung carcinoma).



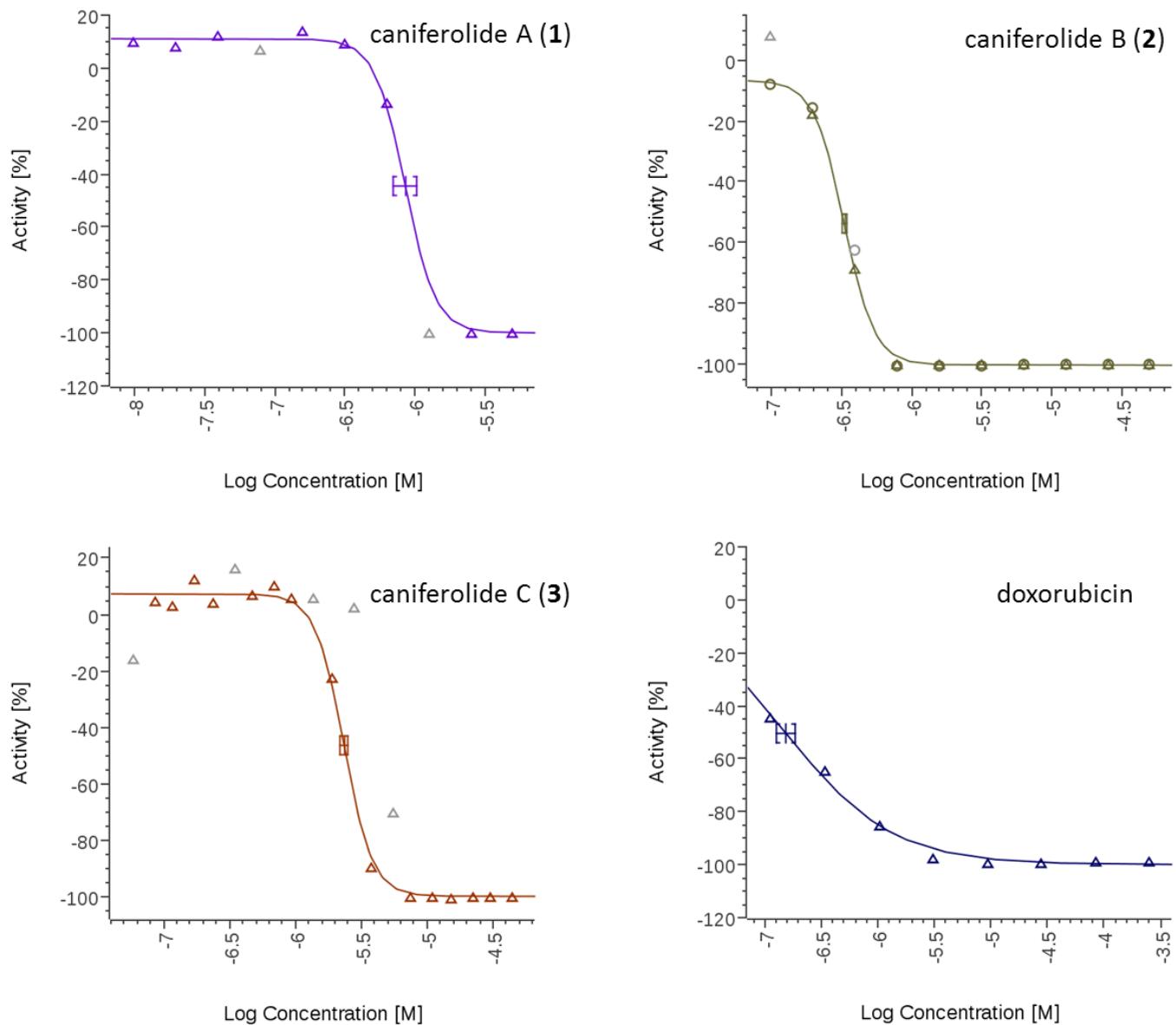
**Figure S66.** Antiproliferative activity curves of **1-3** and doxorubicin against A2058 cell line (skin melanoma).



**Figure S67.** Antiproliferative activity curves of **1-4** and doxorubicin against Hep G2 cell line (hepatocellular carcinoma).



**Figure S68.** Antiproliferative activity curves of **1-3** and doxorubicin against MCF-7 (breast adenocarcinoma).



**Figure S69.** Antiproliferative activity curves of **1-3** and doxorubicin against MiaPaca-2 (pancreatic carcinoma).

## References

1. J. Martín, G. Crespo, V. González-Menéndez, G. Pérez-Moreno, P. Sánchez-Carrasco, I. Pérez-Victoria, L. M. Ruiz-Pérez, D. González-Pacanowska, F. Vicente, O. Genilloud, G. F. Bills and F. Reyes, *J. Nat. Prod.*, 2014, **77**, 2118-2123.
2. I. Pérez-Victoria, J. Martín and F. Reyes, *Planta Med.*, 2016, **82**, 857-871.
3. S.-H. Yoon, S.-M. Ha, S. Kwon, J. Lim, Y. Kim, H. Seo and J. Chun, *Int. J. Syst. Evol. Microbiol.*, 2017, **67**, 1613-1617.
4. M. C. Monteiro, M. De La Cruz, J. Cantizani, C. Moreno, J. R. Tormo, E. Mellado, J. R. De Lucas, F. Asensio, V. Valiante, A. A. Brakhage, J. P. Latgé, O. Genilloud and F. Vicente, *J. Biomol. Screen.*, 2012, **17**, 542-549.
5. F. J. Ortíz-López, M. C. Monteiro, V. González-Menéndez, J. R. Tormo, O. Genilloud, G. F. Bills, F. Vicente, C. Zhang, T. Roemer, S. B. Singh and F. Reyes, *J. Nat. Prod.*, 2015, **78**, 468-475.
6. K. Bartizal, T. Scott, G. K. Abruzzo, C. J. Gill, C. Pacholok, L. Lynch and H. Kropp, *Antimicrob. Agents Chemother.*, 1995, **39**, 1070-1076.
7. T. Mosmann, *J Immunol Methods*, 1983, **65**, 55-63.
8. T. Kieser, J. I. Foundation, M. J. Bibb, M. J. Buttner, K. F. Chater and D. A. Hopwood, *Practical Streptomyces Genetics*, John Innes Foundation, 2000.
9. G. Aggarwal and R. Ramaswamy, *J. Biosci.*, 2002, **27**, 7-14.
10. M. Kearse, R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Meintjes and A. Drummond, *Bioinformatics*, 2012, **28**, 1647-1649.
11. <https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins>.
12. M. Pérez, C. Schleissner, R. Fernández, P. Rodríguez, F. Reyes, P. Zuñiga, F. de la Calle and C. Cuevas, *J. Antibiot.*, 2015, **69**, 388.
13. Y. Miyata and S. Matsunaga, *Tetrahedron Lett.*, 2008, **49**, 6334-6336.