Supporting Information

Structures and comparative characterization of biosynthetic gene clusters for cyanosporasides, enediyne-derived natural products from marine actinomycetes

Amy L. Lane,[†,‡] *Sang-Jip Nam,*[†,£] *Takashi Fukuda,*[†,⊥] *Kazuya Yamanaka,*[†] *Christopher A. Kauffman,*[†] *Paul R. Jensen,*[†] *William Fenical,*[†,§,*] *Bradley S. Moore*[†,§,*]

[†]Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California at San Diego, La Jolla, California 92093-0204

[§]The Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California at San Diego, La Jolla, California 92093

[‡]Current address: Chemistry Department, University of North Florida, Jacksonville, Florida 32224

[£]Current address: College of Pharmacy and Research Institute of Life and Pharmaceutical Sciences, Sunchon National University, Suncheon 540-950, Republic of Korea

[⊥]Current address: School of Pharmacy, Kitasato University, 5-9-1, Shirokane, Minato-ku, Tokyo 108-8641, Japan

*Email: bsmoore@ucsd.edu; wfenical@ucsd.edu

Table of Contents

Experimental Section

General Methods. Optical rotations were measured using a Rudolph Research Autopol III polarimeter or a Jasco DIP-1000 digital polarimeter with a 10-cm cell. UV spectra were recorded in a Varian Cary UV-visible spectrophotometer or a Hitachi U-2800 spectrophotometer with a path length of 1 cm, and IR spectra were recorded on a Perkin-Elmer 1600 FT-IR spectrometer or an Jasco FT/IR-460 plus spectrometer. CD spectra were collected in an AVIV model 215 CD spectrometer or J-720 Spectropolarimeter with a 0.5 cm long cell. ^{1}H and 2D NMR spectra data were recorded at 500 or 600 MHz in DMSO-d_6, or CD_3CN solution containing Me_4Si as internal standard on Varian Inova spectrometers. ^{13}C NMR spectra were acquired at 150 MHz on a Varian Inova spectrometer. High resolution ESI-TOF mass spectra were provided by The Scripps Research Institute, La Jolla, CA or by the mass spectrometry facility at the Department of Chemistry and Biochemistry at the University of California, San Diego, La Jolla, CA. Low-resolution LC/MS data were measured using an Agilent series 1100 LC/MS system. The assembly and analysis of DNA sequence data was carried out with use of Geneious 5.0 software, and protein alignments were prepared with ClustalW.

Collection and Phylogenetic Analysis of Strain CNS-143 and CNT-179. The marine-derived actinomycete strain CNS-143 was isolated from a marine sediment sample collected at a depth of 500 m off Palau in 2004. The strain appears to represent a new *Salinispora* sp. (proposed as "*Salinispora pacifica*", DQ092624) based on 16S rDNA analysis and DNA-DNA hybridization studies. Actinomycete strain CNT-179 was isolated from a sediment sample collected at Bahamas in 2007 and identified as a *Streptomyces* sp. (JX972220) based on 16S rDNA analysis.

Cultivation and Extraction. "*S. pacifica*" CNS-143 was cultured in forty 2.8 l Fernbach flasks, each containing 1 l of seawater-based A1-BFe medium (10 g starch, 4 g yeast extract, 2 g peptone, 1 g $CaCO_3$, 40 mg $Fe_2(SO_4)_3$•$4H_2O$, 100 mg KBr) and shaken at 230 rpm at 27 °C. After seven days of cultivation, sterilized XAD-7 resin (20 g/l) was added to adsorb the organic products, and the culture and resin were shaken at 215 rpm for 2 h. The resin was filtered through cheesecloth, washed with deionized water, and eluted with acetone. The acetone was removed under reduced pressure, and the resulting aqueous layer was extracted with EtOAc (3 × 500 ml). The EtOAc soluble fraction was dried *in vacuo* to yield 3 g of extract. *Streptomyces* sp. CNT-179 was cultivated and extracted analogously, to yield 5 g of extract.

Isolation of Cyanosporasides C and D (3–4). The extract of CNS-143 (3 g) was fractionated by open column chromatography on silica gel (25 g), employing a step gradient of CH_2Cl_2 and MeOH. The CH_2Cl_2/MeOH 5:1 fraction contained a mixture of both cyanosporasides, which were purified by reversed-phase HPLC (Phenomenex Ultracarb C30, 250 × 100 mm, 5 μm, 100 Å, 2.0 ml/min, UV = 210 nm) using a H_2O:MeCN gradient from 5% to 100% MeCN (0.05% TFA) over 60 min to afford cyanosporasides C (**3**, 5.6 mg) and D (**4**, 0.5 mg), as pale yellow oils.

Isolation of Cyanosporaside D–F (4–6). The crude extract of CNT-179 (5 g) was dissolved in a small volume of chloroform, applied on a silica gel column (30 g, 34 x 150 mm, 200-450 mesh), and eluted stepwise with 100% chloroform, 50:1, 25:1, 10:1, 5:1, 3:1, and 1:1 (v/v) of chloroform:MeOH, and 100% MeOH (150 ml each). Cyanospraside D (**4**) was observed in the 3:1 chloroform:MeOH fraction. Compound **4** was purified by HPLC using an ODS column (250 x 10 mm, Phenomenex C18) with a mobile phase of 43:57 MeCN:H_2O in 0.05% TFA at a flow rate of 2.5 ml/min and UV detection at 254 nm. Under this condition, **4** was eluted at 26.0 min. This peak was collected and concentrated to yield 1.7 mg of yellow oil. Cyanospraside E (**5**) was

observed in the 25:1 chloroform:MeOH fraction. Compound **5** was purified by HPLC using an ODS column (250 x 10 mm, Phenomenex C18) with a 50-minute linear H_2O:MeCN gradient from 20 to 45% MeCN in 0.05% TFA at a flow rate of 2.5 ml/min and UV detection at 254 nm. Under these conditions, **5** exhibited a retention time of 27.0 min. This peak was collected and concentrated to yield 2.7 mg of yellow oil. Cyanospraside F (**6**) was observed in the 10:1 chloroform:MeOH fraction. This fraction was purified by HPLC with an ODS column (250 x 10 mm, Phenomenex C18) using a 60-minute linear H_2O:MeCN gradient from 10 to 43% MeCN in 0.05% TFA with a flow rate of 2.5 ml/min and UV detection at 254 nm. Under these conditions, **6** was eluted at 37.0 min, yielding 1.8 mg as a yellow oil.

Cyanosporaside C (3, 5.6 mg**):** pale yellow oil; $[\alpha]_D^{21}$ +18 (*c* 0.5, MeOH); UV (MeOH) λ_{max} (log *e*) 244 (4.0), 301 (3.9), 328 (3.6) nm; IR (KBr) ν_{max} 3448, 2938, 2255, 1733, 1710, 1606, 1131, 1072 cm^{-1}; ^1H and 2D-NMR (600 MHz, CD_3CN), see Table 1; HRESIMS $[M+Na]^+$ *m/z* 482.0982 (calcd for $C_{23}H_{22}{}^{35}ClNO_7$, 482.0977).

Cyanosporaside D (**4**, 1.7 mg): pale yellow oil; $[\alpha]_D^{21}$ + 29 (*c* 0.1, MeOH); IR (KBr) ν_{max} 3424, 2924, 1731, 1611, 1456, 1374, 1238 cm^{-1}; UV (MeOH) λmax (log *e*) 241 (3.6), 302 (3.1), 328 (3.0); ^1H and ^{13}C NMR (600 MHz, DMSO-*d*$_6$), see Table 2; HRESIMS $[M+Na]^+$ *m/z* 324.0397 ($C_{16}H_{12}{}^{35}ClNO_3$, calcd for 324.0398)

Cyanosporaside E (**5**, 2.7 mg): pale yellow oil; $[\alpha]_D^{21}$ + 27 (*c* 0.1, MeOH); IR (KBr) ν_{max} 3421, 2924, 1629, 1554, 1413 cm^{-1}; UV (MeOH) λmax (log *e*) 241 (3.6), 302 (3.1), 328 (3.0); ^1H and ^{13}C NMR (600 MHz, DMSO-*d*$_6$), see Table 2; HRESIMS $[M+Na]^+$ *m/z* 282.0294 ($C_{14}H_{10}{}^{35}ClNO_2$, calcd for 282.0292)

Cyanosporaside F (**6**, 1.8 mg): pale yellow oil; $[\alpha]_D^{21}$ + 82 (*c* 0.2, MeOH); IR (KBr) ν_{max} 3414, 2921, 1596, 1453, 1378, 1205 cm^{-1}; UV (MeOH) λmax (log *e*) 242 (3.7), 299 (3.5), 328

(3.4); ^{1}H and 2D NMR (600 MHz, DMSO-d_6), see Table 3; HRESIMS $[M+Na]^+$ *m/z* 513.1219 ($C_{24}H_{27}{}^{35}ClN_2O_5S$, calcd 513.1221)

***Bis*-MTPA Esters of 3 (3a/3b).** Cyanosporaside C (**3**, 1.0 mg) was dissolved in freshly distilled dry pyridine (2 ml) and several dry crystals of dimethylaminopyridine were added. The mixture was stirred for 15 min at RT. Treatment with *R*-(–)-MTPA-Cl yielded the *bis*-*S*-MTPA ester at 60°C after 12 h. The reaction was quenched by 1 ml of MeOH. After removal of solvent under vacuum, the residue was purified by reversed-phase HPLC (Phenomenex Luna 5*u* C18 (2) 100 Å, 250 × 100 mm) with a flow rate of 2.0 ml/min and UV detection (at 210 nm) using a gradient solvent system from 5% to 100 % MeCN (0.05% TFA) over 50 min. The *bis*-*S*-MTPA ester was obtained at 40 min. The *bis*-*R*-MTPA ester was prepared with *S*-MTPA-Cl in the same manner. $\Delta\delta_{S-R}$ values for the *S*- and *R*-MTPA esters of **1** were recorded in ppm in DMSO-d_6.

***Bis*-*S*-MTPA Ester of Cyanosporaside C (3a)**; ^{1}H NMR (600 MHz, DMSO-d_6): δ 7.66-7.37 (m, 6H), 7.36 (d, 1H, *J* = 2.0 Hz), 7.35-7.12 (m, 4H), 7.11 (d, 1H, *J* = 2.0 Hz), 6.91 (d, 1H, *J* = 6.0 Hz), 6.55 (s, 1H), 6.43 (dd, 1H, *J* = 6.0, 2.5 Hz), 5.61 (d, 1H, *J* = 2.5 Hz), 5.37 (d, 1H, *J* = 8.0 Hz), 4.31 (d, 1H, *J* = 8.0 Hz), 3.76 (d, 2H, *J* = 14.0 Hz), 3.63 (s, 3H), 3.49 (q, 1H, *J* = 6.5 Hz), 2.98 (s, 3H), 2.12 (s, 3H), 1.21 (s, 3H), 1.12 (d, 3H, *J* = 6.5 Hz); ESI-MS *m/z* 914 $[M+Na]^+$.

***Bis*-*R*-MTPA Ester of Cyanosporaside C (3b)**; ^{1}H NMR (600 MHz, DMSO-d_6): δ 7.60-7.36 (m, 6H), 7.35 (d, 1H, *J* = 2.0 Hz), 7.34-7.06 (m, 4H), 7.05 (d, 1H, *J* = 2.0 Hz), 7.04 (d, 1H, *J* = 6.0 Hz), 6.59 (s, 1H), 6.65 (dd, 1H, *J* = 6.0, 2.5 Hz), 5.85 (d, 1H, *J* = 2.5 Hz), 5.34 (d, 1H, *J* = 8.0 Hz), 4.43 (d, 1H, *J* = 8.0 Hz), 3.70 (d, 2H, *J* = 14.0 Hz), 3.43 (s, 3H), 3.51 (q, 1H, *J* = 6.5 Hz), 3.19 (s, 3H), 2.11 (s, 3H), 1.20 (s, 3H), 1.13 (d, 3H, *J* = 6.5 Hz); ESI-MS *m/z* 914 $[M+Na]^+$.

Genomic DNA library preparation for "*S. pacifica*" CNS-143 and *Streptomyces* sp. CNT-179. For both "*S. pacifica*" CNS-143 and *Streptomyces* sp. CNT-179, genomic DNA (gDNA)

was isolated from 7-10 day A1+BFe fermentations following established protocols.[1] For *Streptomyces* sp., a ~1,000 member clone library of 30-50 kB genomic DNA fragments was prepared with the SuperCos vector (Agilent Technologies) following manufacturer-supplied protocols with some modifications. Briefly, SuperCos was digested with XbaI and this linearized vector dephosphorylated by treatment with calf intestinal phosphatase (CIP), and then cleaved into two fragments through digestion with BamHI. *Streptomyces* sp. gDNA was digested with Sau3A1 to yield DNA fragments of ~30-50 kb, as determined by agarose gel electrophoresis. These fragments were dephosphorylated by treatment with CIP. This gDNA was ligated into the SuperCos vector using T4 DNA Ligase (New England Biolabs). Ligations were packaged using MaxPlax Lambda Packaging Extracts (EpiCentre), and the packaging reaction titered using *E. coli* XL1-MRF Blue cells following instructions provided by EpiCentre. For "*S. pacifica*" CNS-143, a ~500 member clone library was prepared using pCC1FOS vector (Epicentre) and following manufacturer instructions.

Cyanosporaside gene cluster identification and sequencing for *Streptomyces* sp. CNT-179. Genomic DNA from *Streptomyces* sp. CNT-179 was subjected to PCR amplification of enediyne-associated genes using degenerate primers and reaction conditions described by Liu et al.[2] PCR products were cloned into pGEM-T (Promega) and sequenced. This sequence information was used to design specific primers for screening the *Streptomyces* sp. clone library for enediyne gene clusters. The library was PCR screened by systematically pooling cosmid clones from individual rows and columns of plates. PCRs were completed with a forward primer sequence of CGTCCCCGACGCCACCCT and reverse sequence of AGGACCCGCTCCAGCGTC. Each reaction mixture consisted of 0.5 μl of Taq DNA Polymerase (New England Biolabs), 2.5 μl 10X ThermoPol Buffer, 250 nM of each primer, 250

μM dNTPs, 2 μl of cosmid DNA, 1.25 μl of DMSO, and sufficient sterile water to yield a final reaction volume of 25 μl. PCRs were cycled at 95°C for 2 min, followed by 30 cycles of 95°C for 30 s, 66°C for 1 min, and 72°C for 1.5 min, and a final cycle of 72°C for 10 min followed by a 4°C hold. In total, three clones (pLC12, pDG11, pGD12) from the *Streptomyces* sp. library were further pursued to yield the complete cyanosporaside gene cluster. The *Streptomyces* sp. cyanosporaside gene cluster sequence was established by sequencing of the complete, overlapping pGD12 and pLC12 cosmids, and further extended by primer walking of the pDG11 cosmid. DNA sequencing and assembly was conducted by Macrogen (Seoul, South Korea) and Beckman Coulter Genomics (Beverley, MA, USA). DNA sequence data were analyzed using FramePlot[3] to identify putative orfs and BLAST searches to assign functions of corresponding proteins.[4] NapDos was employed to further analyze individual PKSE domains.[5]

Cyanosporaside gene cluster identification and sequencing for "*S. pacifica*" CNS-143. For "*S. pacifica*" CNS-143, the full cyanosporaside gene cluster sequence was determined through evaluation of draft genome sequencing data and primer walking. Genome sequence data for this strain was obtained by 454 pyrosequencing at ~8X coverage along with Sanger sequencing at ~0.5X coverage. Genome sequence data were assembled with the Celera Assembler using the Best Overlap graph and referenced with the completed genome sequence of *S. arenicola* CNS-205 (CP000850). Analysis of these data via BLAST searches revealed a partially sequenced enediyne gene cluster, corresponding to cosmids pO19, pI19, and pA20. To close gaps within the gene cluster, primer walking was carried out on these cosmids as necessary.

Inactivation of cyanosporaside biosynthetic genes in "*S. pacifica*" CNS-143. Gene inactivation experiments were performed using λ-RED recombination technology as described by Gust and coworkers[6] and previously adapted for gene knockouts in *Salinispora tropica*.[7]

Briefly, an apramycin resistance cassette (*oriT-aac(3)IV* surrounded by FRT sites) was amplified by PCR from plasmid pIJ773 with primers listed in Table S1. Each ~1.4 kb PCR-amplified cassette was introduced by electroporation into *E. coli* BW25113 carrying recombination plasmid pKD20 and the targeted fosmid. Transformants were selected by antibiotic resistance, and gene replacement on the fosmid confirmed by colony PCR and DNA sequencing. Mutant fosmids were introduced into *E. coli* S17-1 cells by electroporation and then transferred into "*S. pacifica*" CNS-143 by intergeneric conjugation. Mutants, in which the apramycin resistance cassette was integrated into the "*S. pacifica*" chromosome, were selected on A1 agar by overlaying plates with 200 μg/ml apramycin and 100 μg/ml nalidixic acid. Following incubation at 28°C for 5-9 days, exconjugants were randomly selected and those exhibiting an Apr^r/Chl^s phenotype were evaluated by PCR and DNA sequencing for double crossover mutation.

Comparative metabolite profiling between wild type "*S. pacifica*" and gene elimination mutants. Wild type "*S. pacifica*" CNS-143 and gene elimination mutants prepared above were grown in 50 ml liquid A1+BFe media for 7 days at 28°C with shaking at 200 rpm. Amberlite XAD-7 resin (0.5 g) was added to each culture and cultures shaken for 3-4 h. The cells and resin were poured over cheesecloth and washed with deionized water. The cheesecloth was then soaked overnight in acetone to extract cyanosporasides from the resin. Resulting acetone extracts were filtered to remove resin and concentrated to dryness *in vacuo*. Prior to LC/MS analysis, extracts were pre-fractionated using C18 HyperSep SPE columns (Thermo Scientific) eluted with 25:75 MeOH:H_2O, 50:50 MeOH:H_2O, 75:25 MeOH:H_2O, and MeOH. Using a Phenomenex Luna C18 column (4.6 mm × 100 mm, 5 μm), LC/MS analyses were conducted with a flow rate of 0.7 ml/min and mobile phase gradient from 10–100% MeCN in H_2O over 26 min with positive mode ESI.

Biochemical induction assay for DNA-interfering compounds. The above-prepared "*S. pacifica*" CNS-143 and *Streptomyces* sp. CNT-179 chemical extracts were evaluated for activity in the biochemical induction assay as previously described,[8] with all assays conducted using 10 μl of a 10 mg/ml extract stock solubilized in MeOH. Assays were also completed using 10 μl aliquots of "*S. pacifica*" and *Streptomyces* sp. cultures at fermentation time points from 3-7 days.

Table S1. PCR primers for amplification of extended antibiotic resistance cassette during preparation of "*S. pacifica"* CNS-143 gene elimination mutants.

primer	sequence	gene elimination target (predicted protein function)
GTmut-F	GTCTGCGTCGTCTCCCACGCCCCCATCGCGGAGTCGACCattccggggatccgtcgacc	
GTmut-R	CACAGCGGTCGCCGAGATCTCGTTGCCGACGAAATCGAGtgtaggctggagctgcttc	*cyaGT* (glycosyltransferase)
MTmut-F	GGCCACTGGATCTCGGGTTCTGCCCACAGTGCAGCCCTGattccggggatccgtcgacc	
MTmut-R	GCCGCCGTGAGCATCGACGAGATCGATATCAAAGACTTCtgtaggctggagctgcttc	*cyaA3* (C-methyltransferase)
ATmut-F	GCCCTGTTGGTCTTCCTTTTCCATACTGCTATTTTCGTGattccggggatccgtcgacc	
ATmut-R	CCACCCGCCAAAAAGGGATGCGGCGAAAAGGCCGAGCAGtgtaggctggagctgcttc	*cyaA4* (acyltransferase)
EPmut-F	CGGGCAAATGATGGATCGGGAGCGACGCCACCGAGGATGattccggggatccgtcgacc	
EPmut-R	AGGCAGGTTTGGGCGGGTCGAATGGGCGGTTGACTTTCAtgtaggctggagctgcttc	*orf24* (epimerase)
ENEmut-F	ATGAAACCTACCCGGATCGATATTCCTGAGGCCCAGTTGattccggggatccgtcgacc	
ENEmut-R	TCAGGAGCGTAGGTCGCTGGCGAAGGCACGGAGGTCCGCtgtaggctggagctgcttc	*cyaF* (enediyne epoxide hydrolase)
AMmut-F	GTGACCATCAACCAGTTGTTTACCACATTGTCCAAGGCCattccggggatccgtcgacc	
AMmut-R	TCAGCGCACGGCGACCCAACCGTCGGGGGTGGTGCGCGAtgtaggctggagctgcttc	*cyaN1* (aminotransferase)
OXmut-F	GTGATGTATGACTACGTAATTGTCGGTGCCGGCTCTGCGattccggggatccgtcgacc	
OXmut-R	TCACCCAGCCACGGGAACCGACTCCGAGACCAGCGCCGCtgtaggctggagctgcttc	*cyaN2* (oxidoreductase)
P450mut-F	CTCGAATATCGACACGCTGTGTTCAACGCGCTGCGGGCCattccggggatccgtcgacc	
P450mut-R	GATGTCCGGAATTCGAGTCAGCAGTTCCCGAAACATCACtgtaggctggagctgcttc	*orf9* (cytochrome P450)

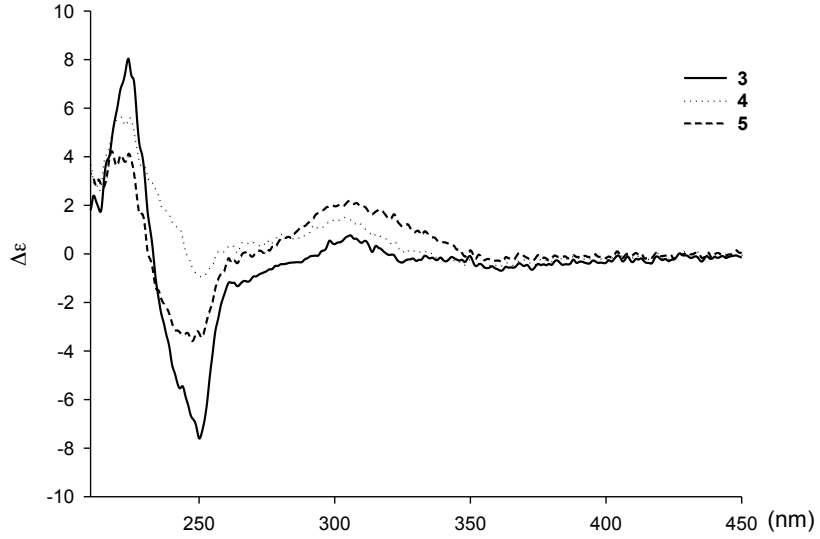
Table S2. NMR spectroscopic data for cyanosporasides C-F (**3-6**)

No.	**3**[a]		**4**[b]		**5**[b]		**6**[c]	
	δ_C, mult.[d]	δ_H (*J* in Hz)	δ_C, mult.[e]	δ_H (*J* in Hz)	δ_C, mult.[e]	δ_H (*J* in Hz)	δ_C, mult.[d]	δ_H (*J* in Hz)
1	131.1, CH	6.76, d (6.0)	132.2, CH	6.92, d (5.7)	127.7, CH	6.67, d (5.7)	128.2, CH	6.66, d (5.5)
2	146.6, CH	6.66, dd (6.0, 2.5)	142.1, CH	6.61, dd (5.7, 2.5)	147.9, CH	6.64, dd (5.7, 2.2)	146.5, CH	6.58, dd (5.5, 2.0)
3	76.5, CH	4.68, d (2.5)	77.0, CH	5.52, d (2.5)	75.5, CH	4.72, d (2.2)	82.2, CH	4.58, d (2.0)
3a	97.5, C		90.1, C		90.8, C		89.3, C	
3b	144.7, C		147.8, C		149.6, C		143.4, C	
4	125.5, CH	7.38, d (2.0)	123.2, CH	7.21, s	123.1, CH	7.31, s	130.9, C	
5	130.9, C		130.7, C		130.0, C		125.5, C	
6	130.5, CH	7.31, d (2.0)	128.7, CH	7.30, s	127.9, CH	7.26, s	130.5, CH	7.35, d (7.5)
7	127.3, C		126.1, C		125.5, C		120.6, CH	7.16, d (7.5)
7a	147.5, C		144.7, C		145.2, C		149.6, C	
8	130.9, CH	6.54, s	117.3, CH	6.52, s	116.1, CH	6.41, s	118.2, CH	6.38, s
8a	159.5, C		160.7, C		161.9, C		159.9, C	
9	23.5, CH_2	3.83, s	21.9, CH_2	4.04, s	21.7, CH_2	4.05, s	20.8, CH_2	3.97, s
10	119.6, C		119.2, C		119.2, C		118.0, C	
1′	101.5, CH	3.56, d (8.0)	170.2, C				106.4, CH	4.49, d (8.5)
2′	76.8, CH	4.09, d (8.0)	22.4, CH_3	1.70, s			42.7, CH	2.27, ddd (13.0, 8.5, 4.0)
3′	203.8, C						38.2, CH_2	1.08, m (3′ax) 1.93, m (3′eq)
4′	84.6, C						70.7, CH_2	2.97, ddd (10.0, 9.0, 4.0)
5′	74.8, CH	3.03, q (6.5)					75.5, CH	3.20, dq (9.0, 6.0)
6′	14.4, CH_3	1.22, q (6.5)					18.0, CH_3	1.15, q (6.0)
7′	15.9, CH_3	1.24, s						
1″	171.8, C						30.2, CH_2	1.83, 1.60, m
2″	21.3, CH_3	2.05, s					38.1, CH_2	2.69, 2.63, m
2″-NH								7.55, t (6.0)
1‴							168.8, C	
2‴							22.5, CH_3	1.70, s

[a]500 MHz for ^{1}H NMR and 75 MHz for ^{13}C NMR in CD_3CN. [b]600 MHz for ^{1}H NMR in DMSO-d_6. [c]600 MHz for ^{1}H NMR and 125 MHz for ^{13}C NMR in DMSO-d_6. [d] Numbers of attached protons were determined by analysis of 2D spectra. [e] The chemical shifts of carbons and numbers of attached protons were determined by analysis of 2D spectra.

Figure S1. (a) CD spectra of cyanosporasides C (**3**), D (**4**) and E (**5**) in methanol ($1.0e^{-4}$ M) CD spectrum of cyanosporaside F (**6**) in methanol ($9.7e^{-4}$ M).

a.



b.

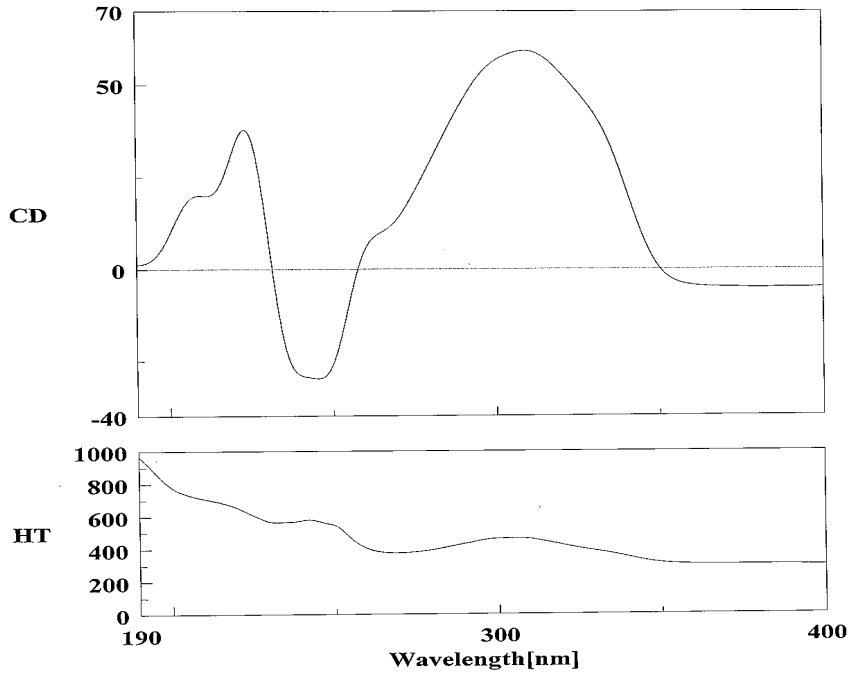
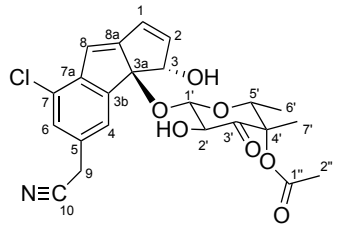


Figure S2. ^{1}H NMR spectrum (500 MHz) of cyanosporaside C (**3**) in CD_3CN.



cyanosporaside C (**3**)

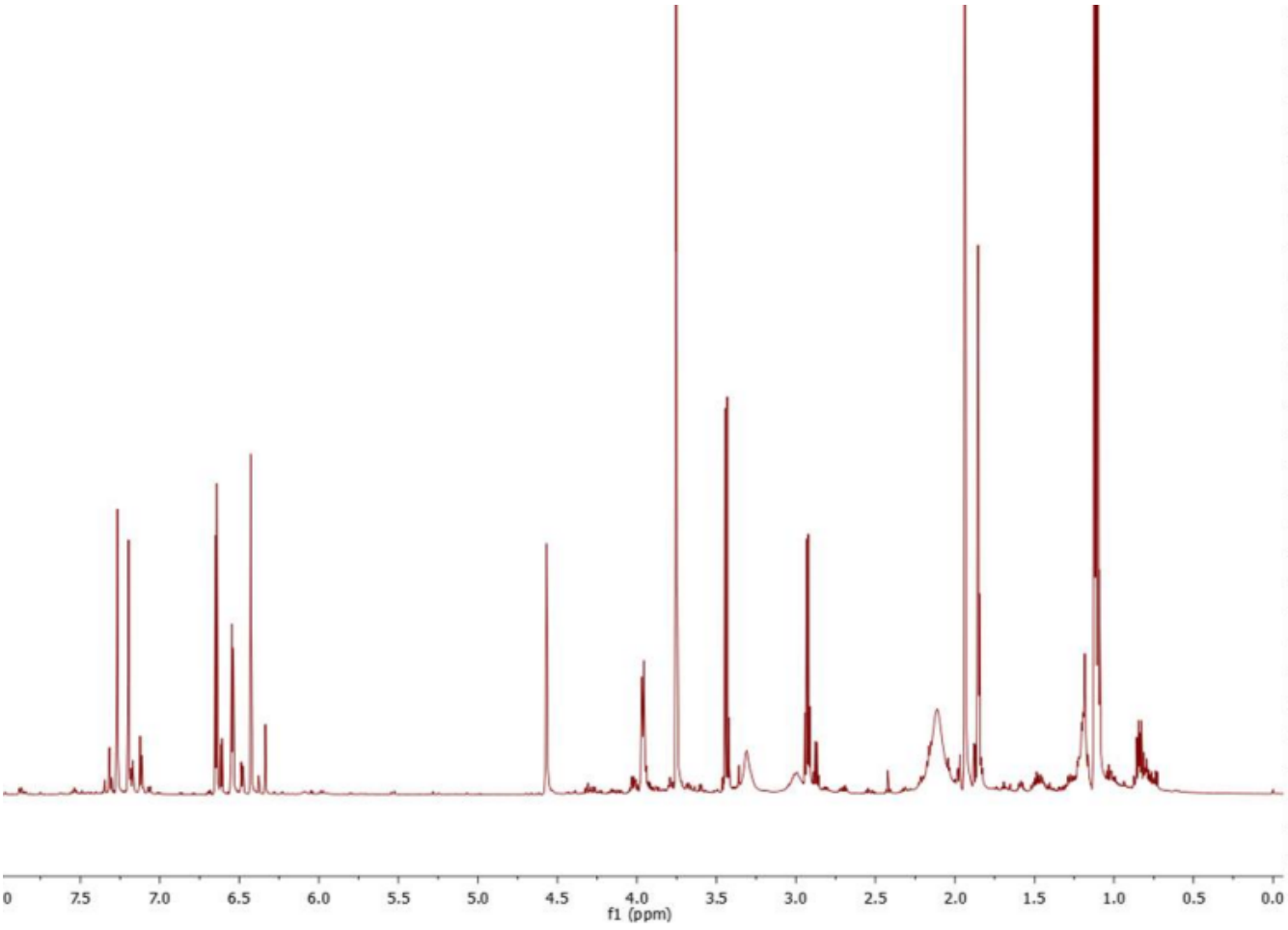


Figure S3. ^{13}C NMR spectrum (75 MHz) of cyanosporaside C (**3**) in CD_3CN.

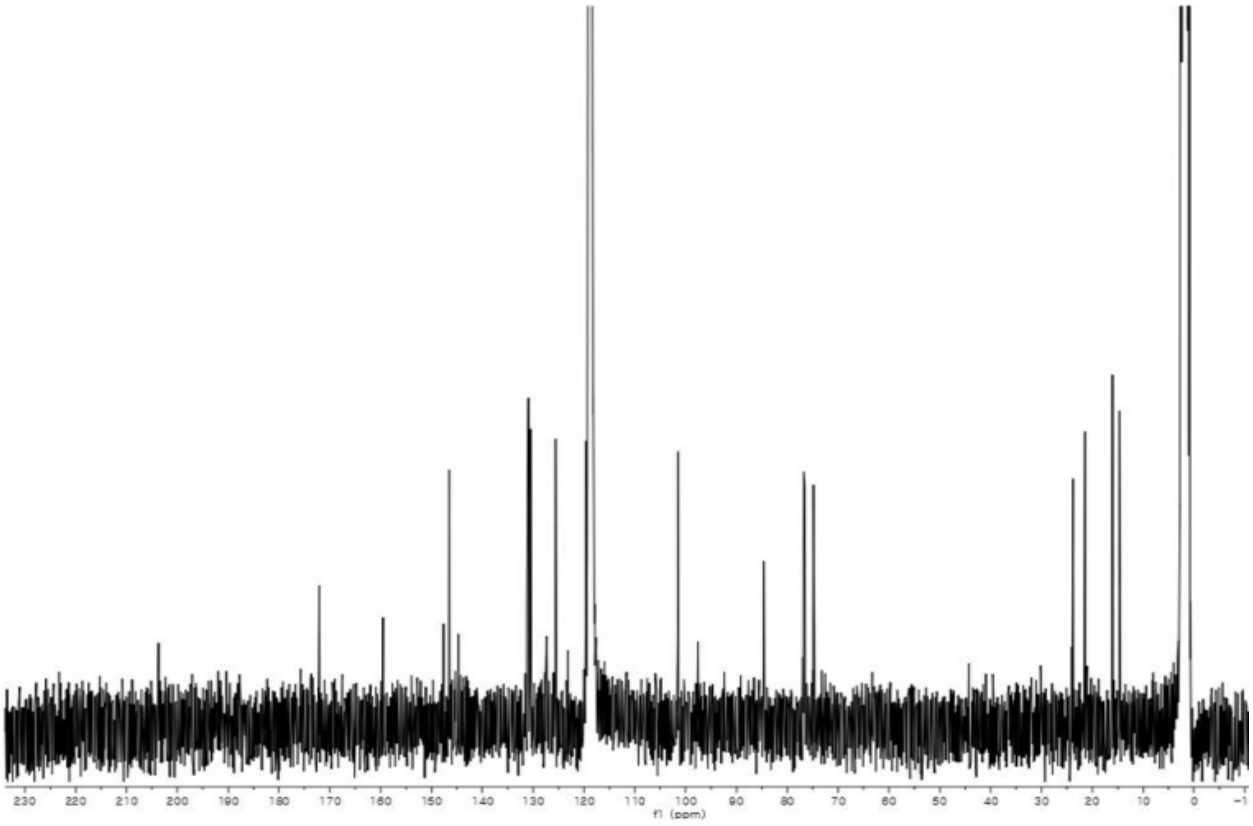


Figure S4. gCOSY spectrum (500 MHz) of cyanosporaside C (**3**) in CD_3CN.

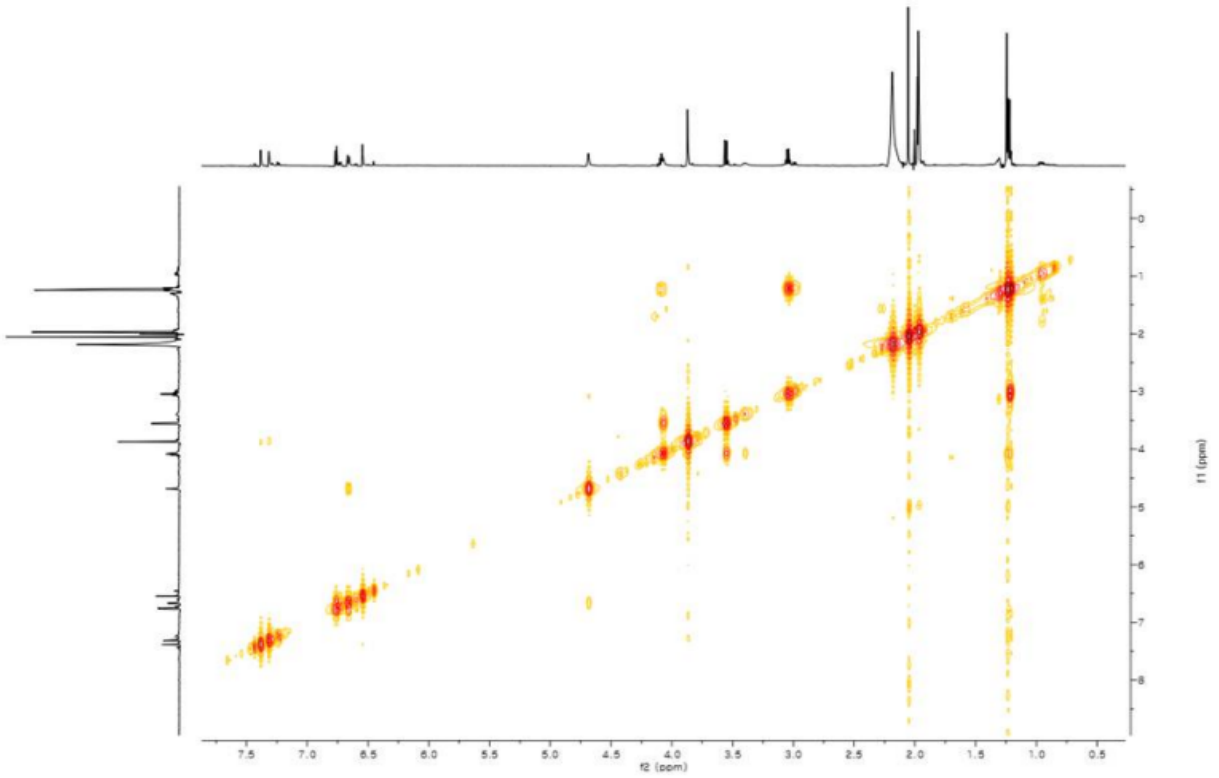


Figure S5. gHSQC spectrum (500 MHz) of cyanosporaside C (**3**) in CD_3CN.

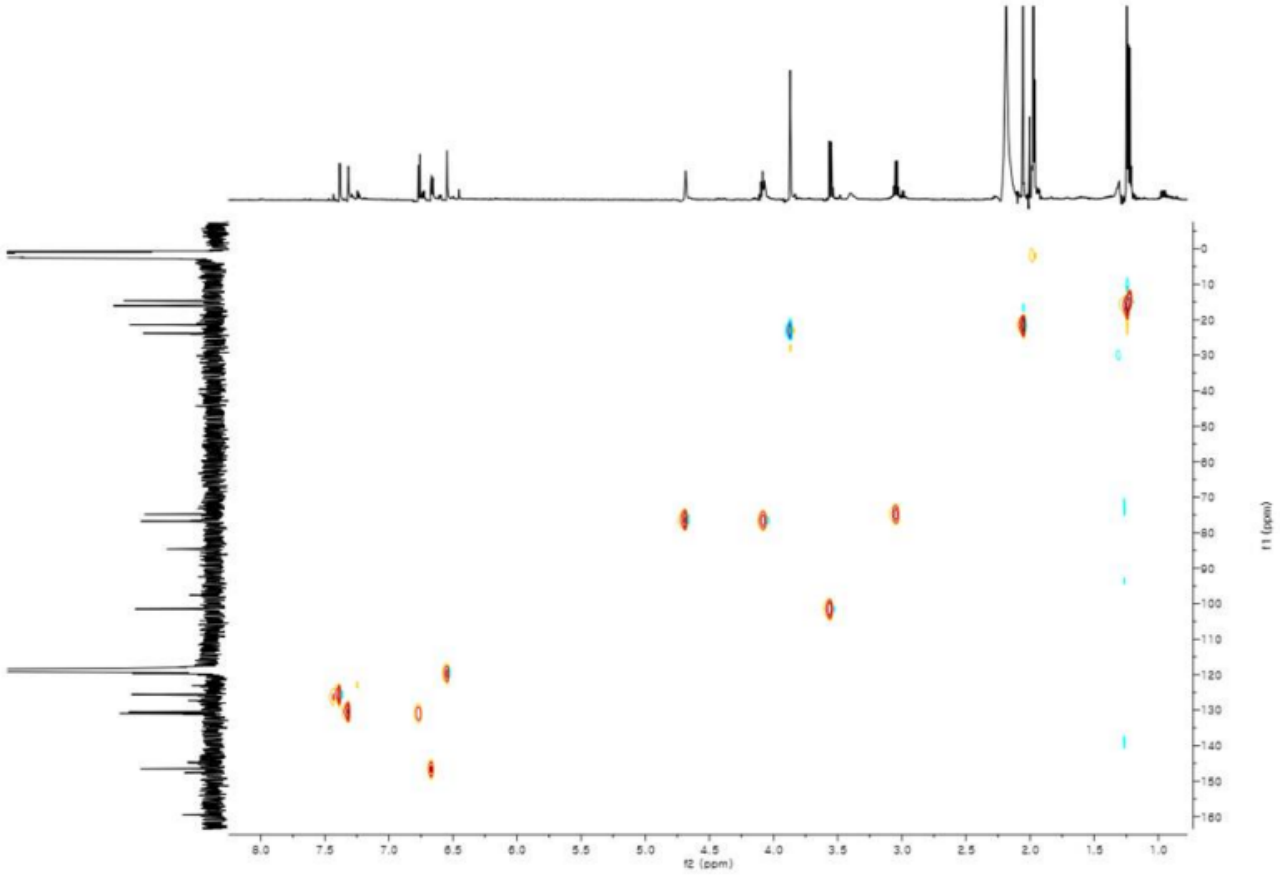


Figure S6. gHMBC spectrum (500 MHz) of cyanosporaside C (**3**) in CD_3CN.

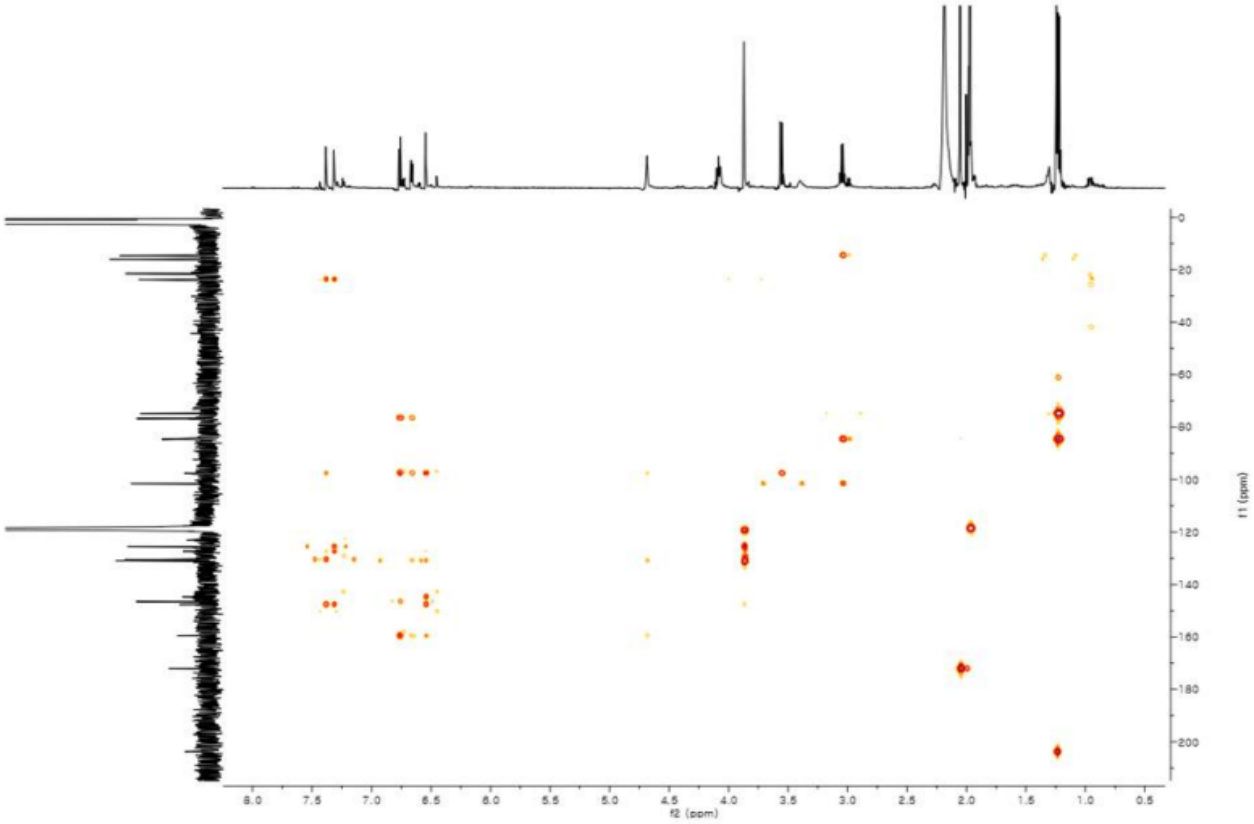


Figure S7. ROESY spectrum (500 MHz) of cyanosporaside C (**3**) in CD_3CN.

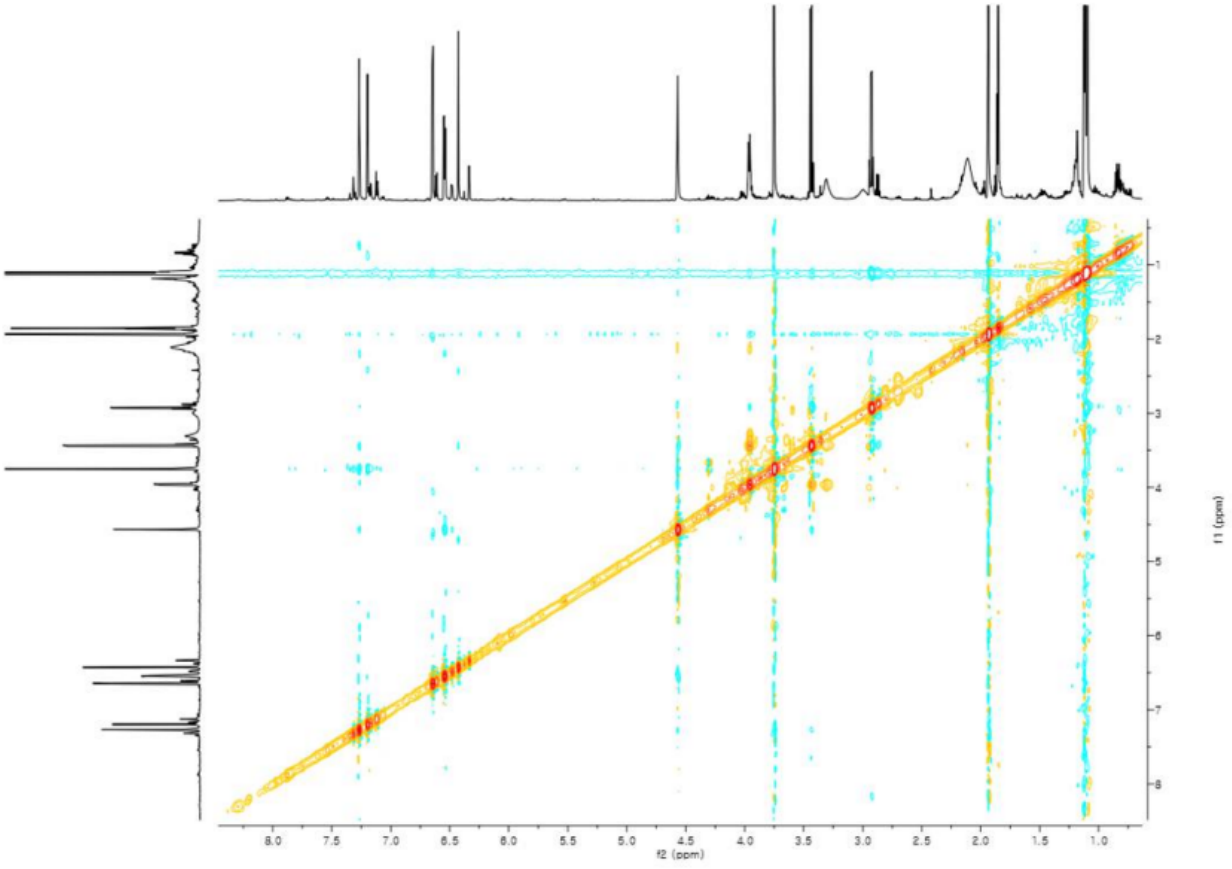


Figure S8. ^{1}H NMR spectrum (600 MHz) for *bis*-*S*-MTPA ester (**3a**) of cyanosporaside C in DMSO-d_6.

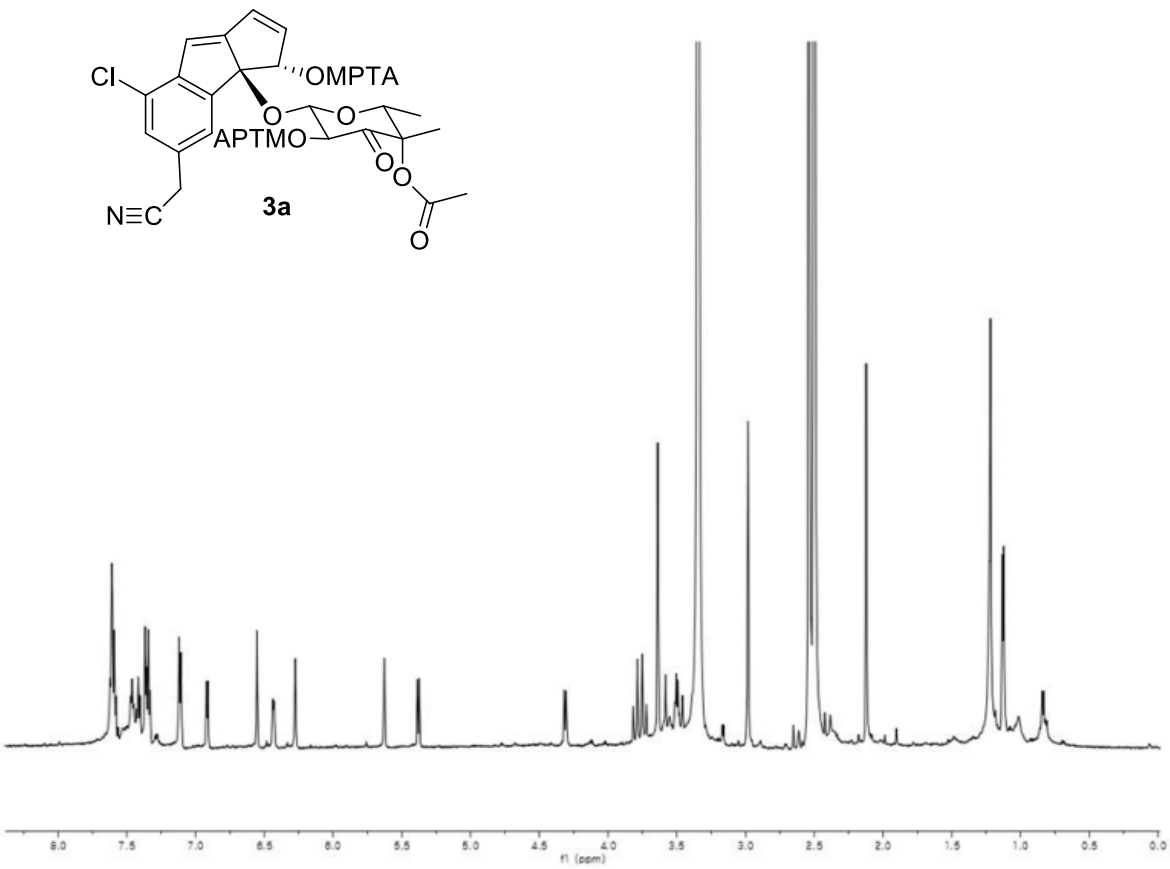


Figure S9. ^{1}H NMR spectrum (600 MHz) for *bis*-*R*-MTPA ester (**3b**) of cyanosporaside C in DMSO-d_6.

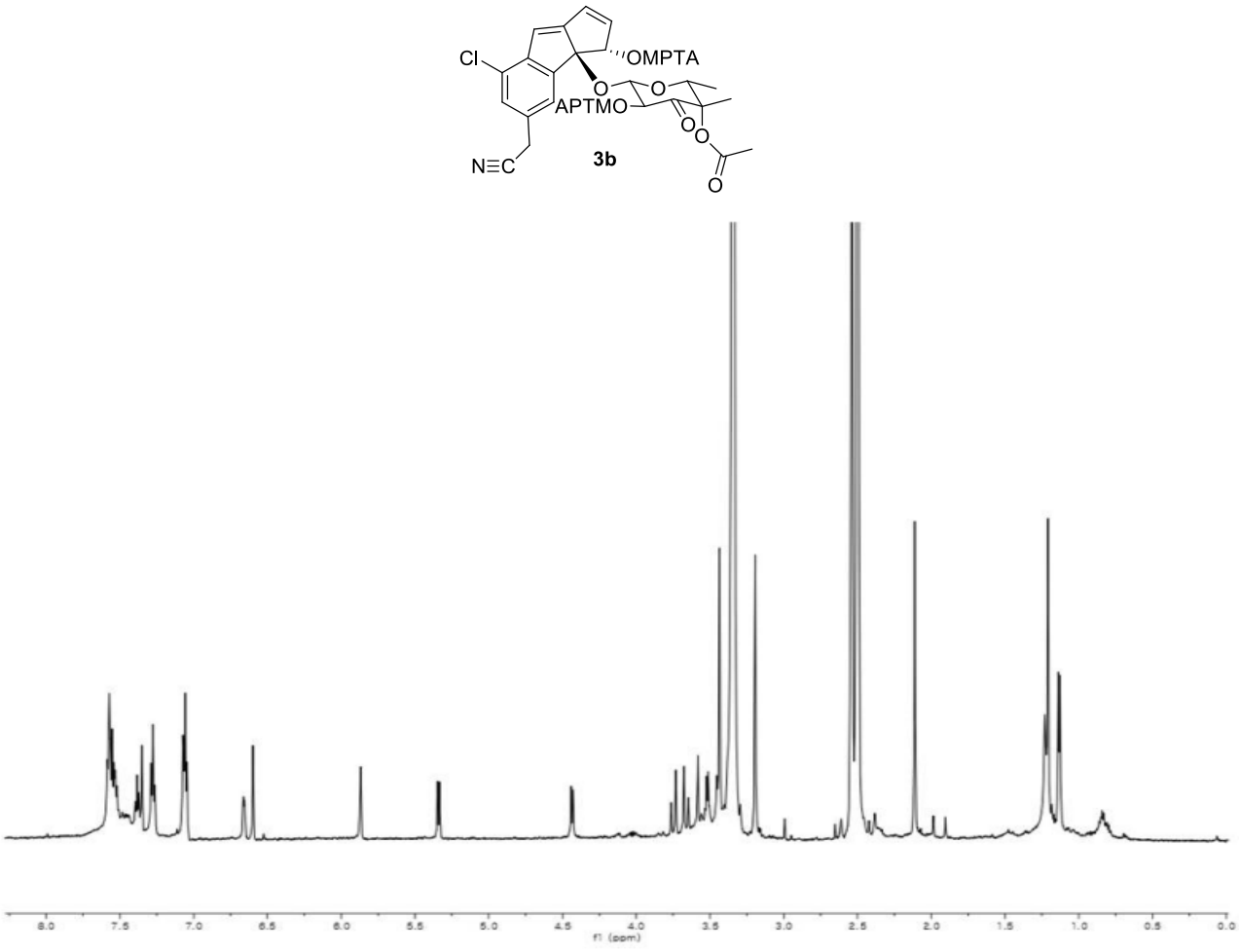
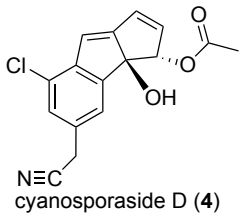


Figure S10. ^{1}H NMR spectrum (600 MHz) of cyanosporaside D (**4**) in DMSO-d_6



cyanosporaside D (**4**)

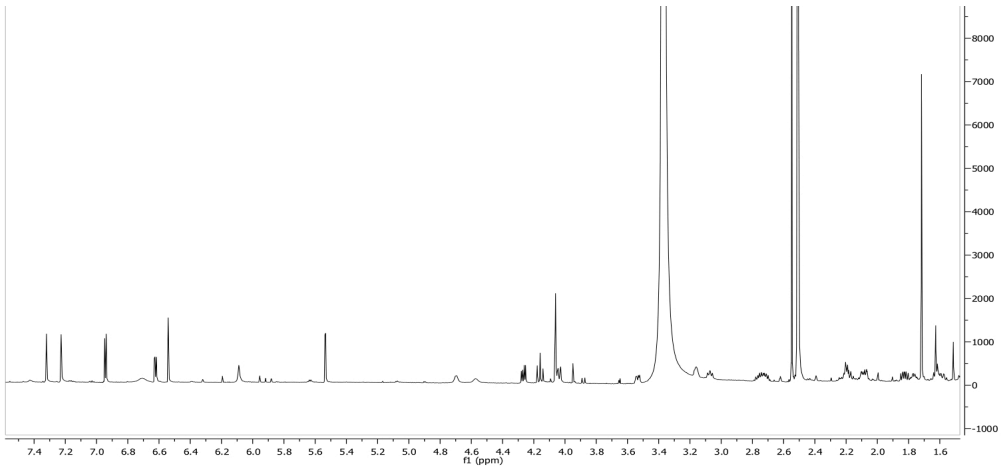
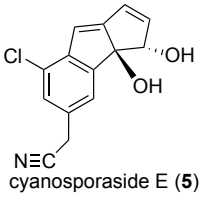


Figure S11. ^{1}H NMR spectrum (600 MHz) of cyanosporaside E (**5**) in DMSO-d_6.



cyanosporaside E (**5**)

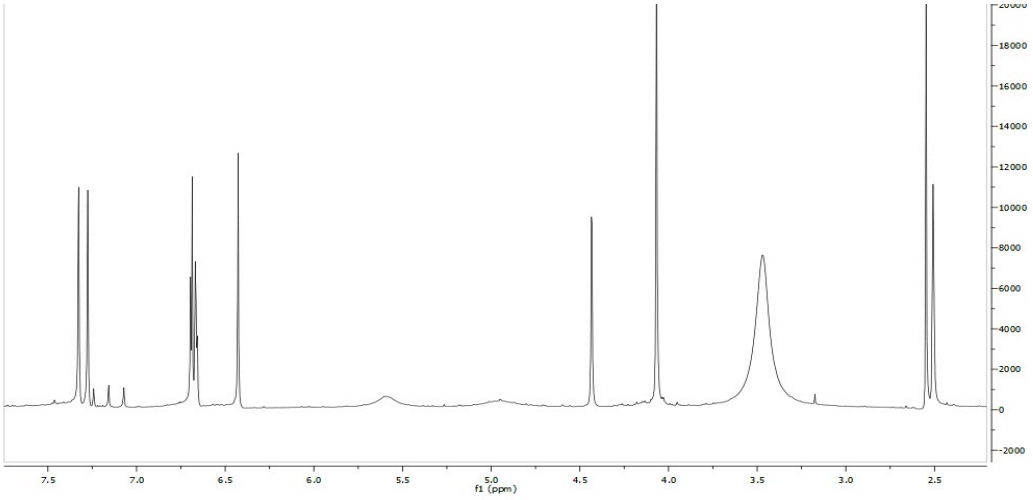
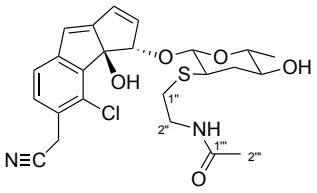


Figure S12. ^{1}H NMR spectrum (600 MHz) of cyanosporaside F (**6**) in DMSO-d_6.



cyanosporaside F (**6**)

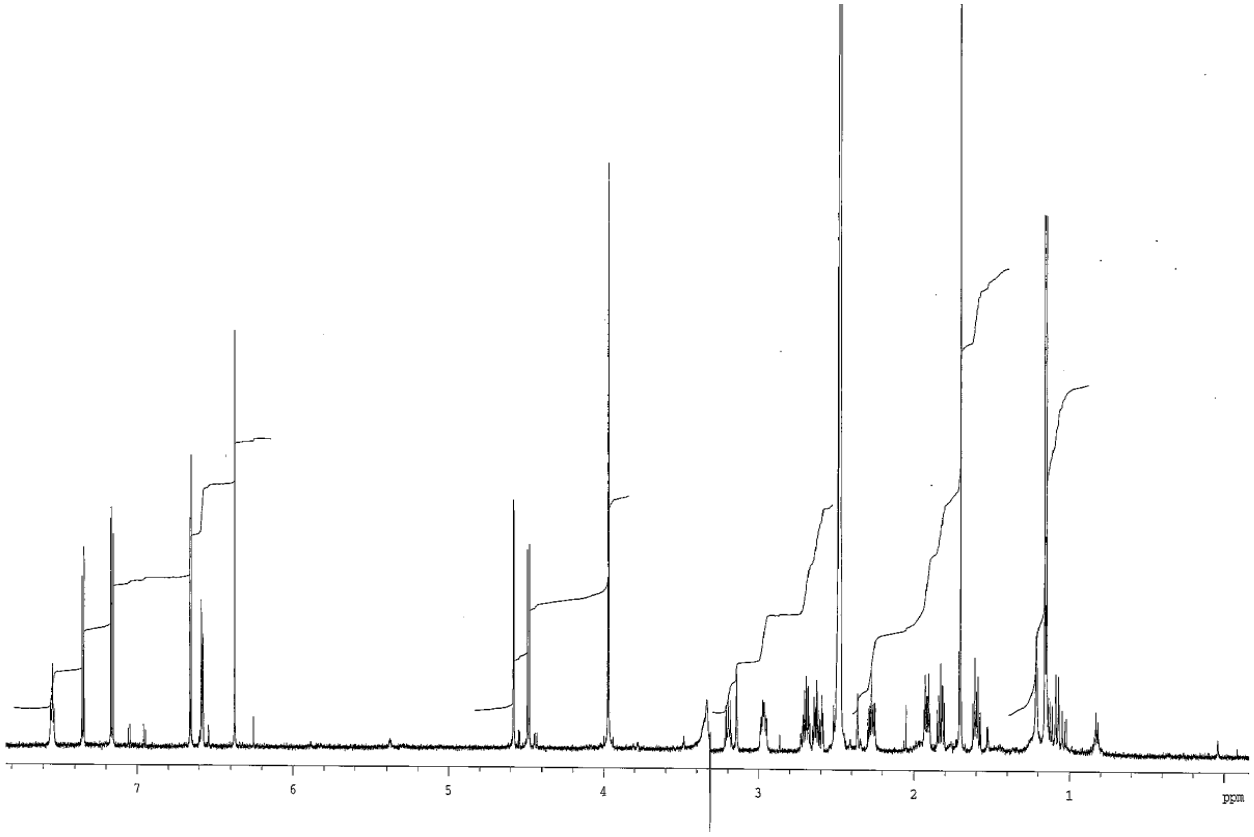


Figure S13. ^{13}C NMR spectrum (125 MHz) of cyanosporaside F (**6**) in DMSO-d_6.

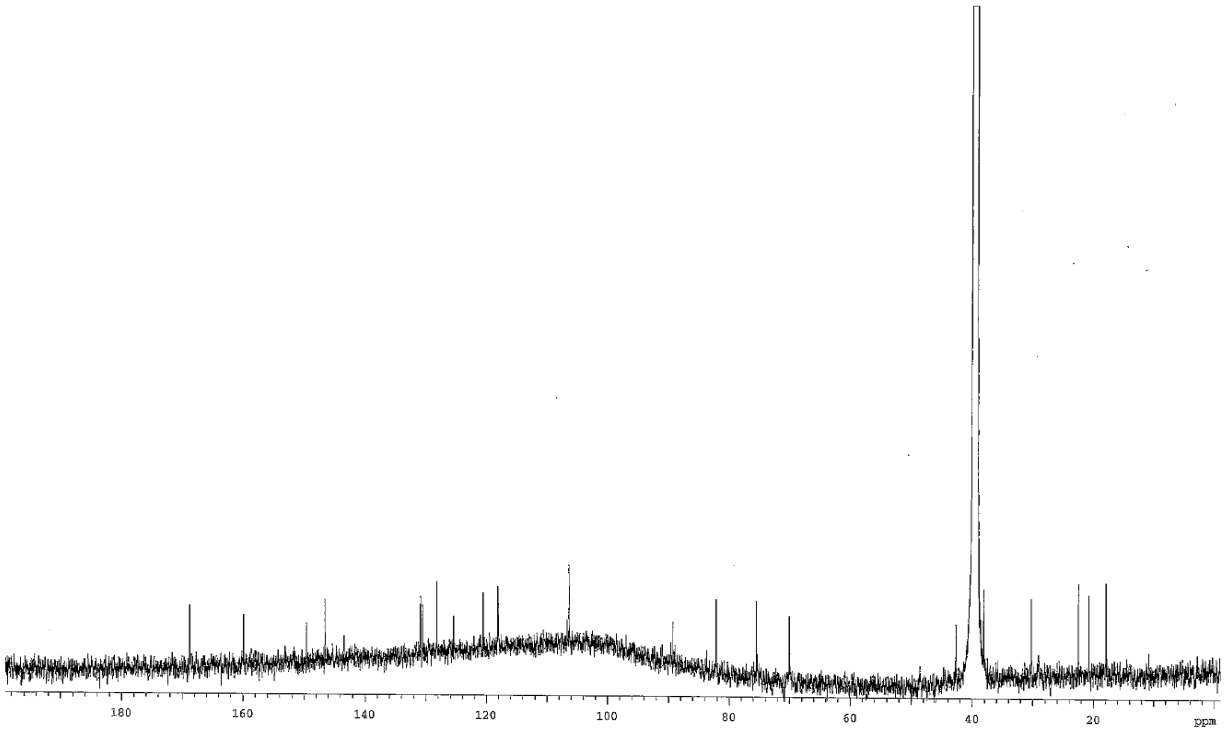


Figure S14. gCOSY NMR spectrum (600 MHz) of cyanosporaside F (**6**) in DMSO-d_6.

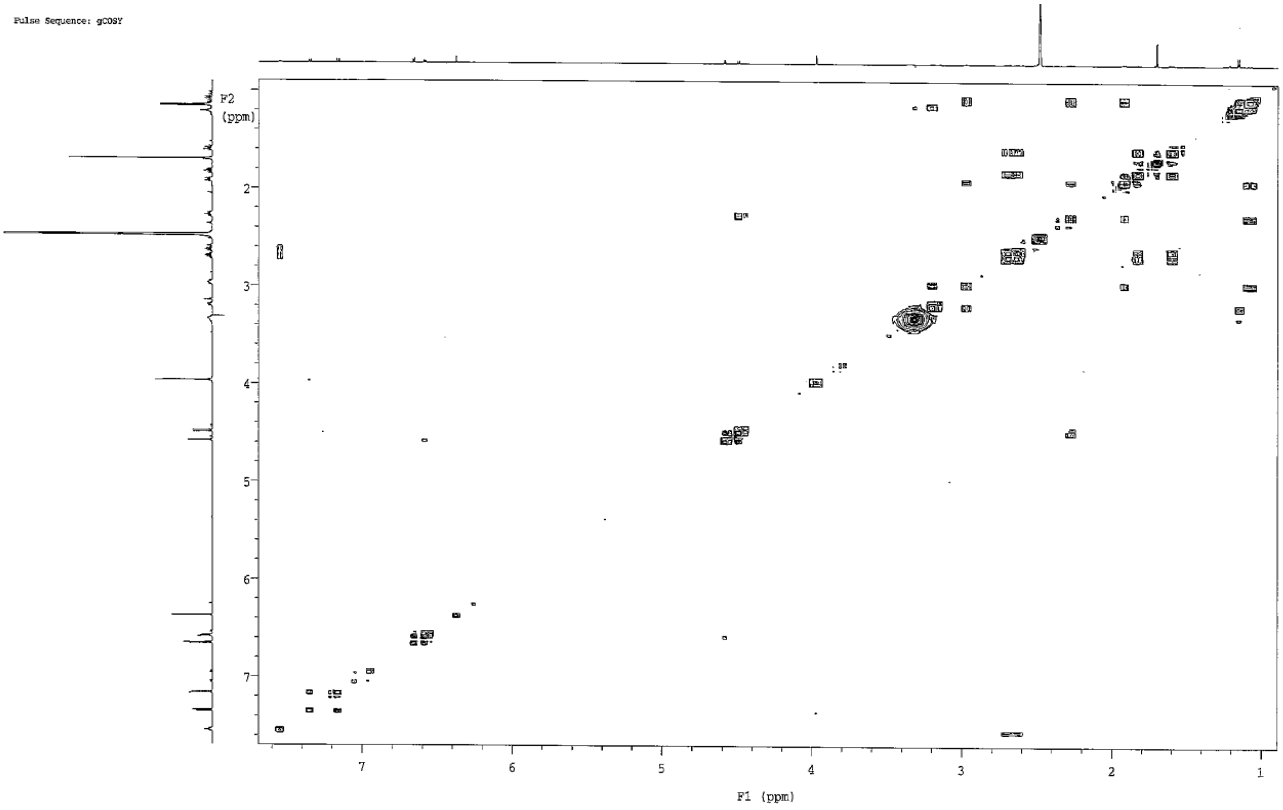


Figure S15. gHSQC NMR spectrum (600 MHz) of cyanosporaside F (**6**) in DMSO-d_6

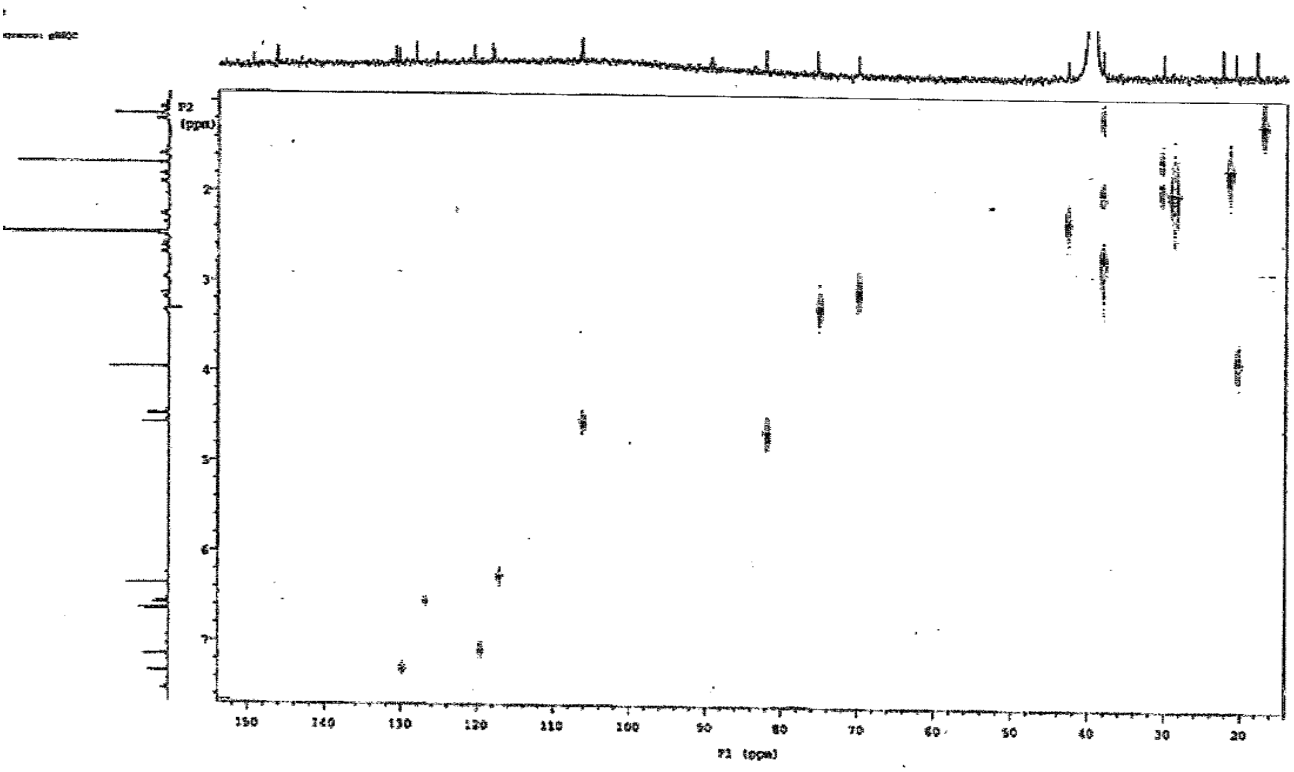


Figure S16. gHMBC NMR spectrum (600 MHz) of cyanosporaside F (**6**) in DMSO-d_6.

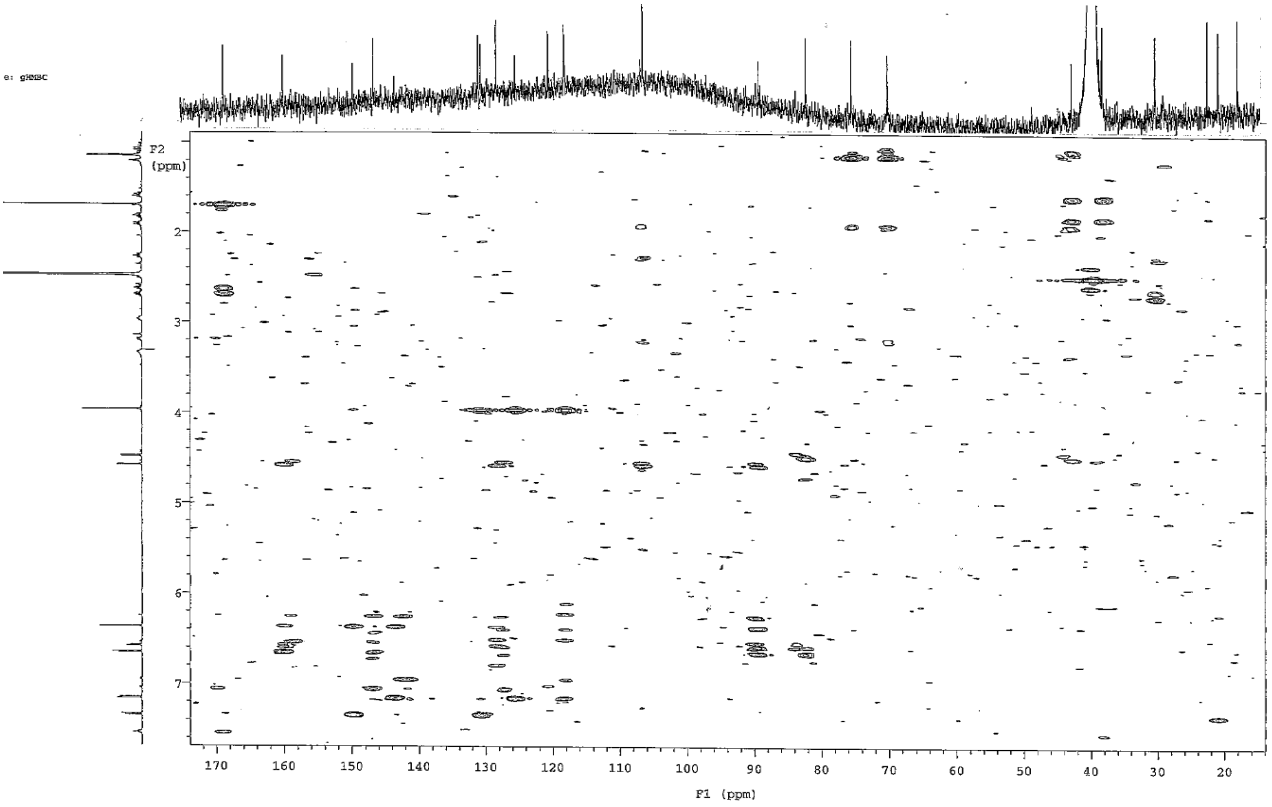


Figure S17. NOESY NMR spectrum (600 MHz) of cyanosporaside F (**6**) in DMSO-d_6.

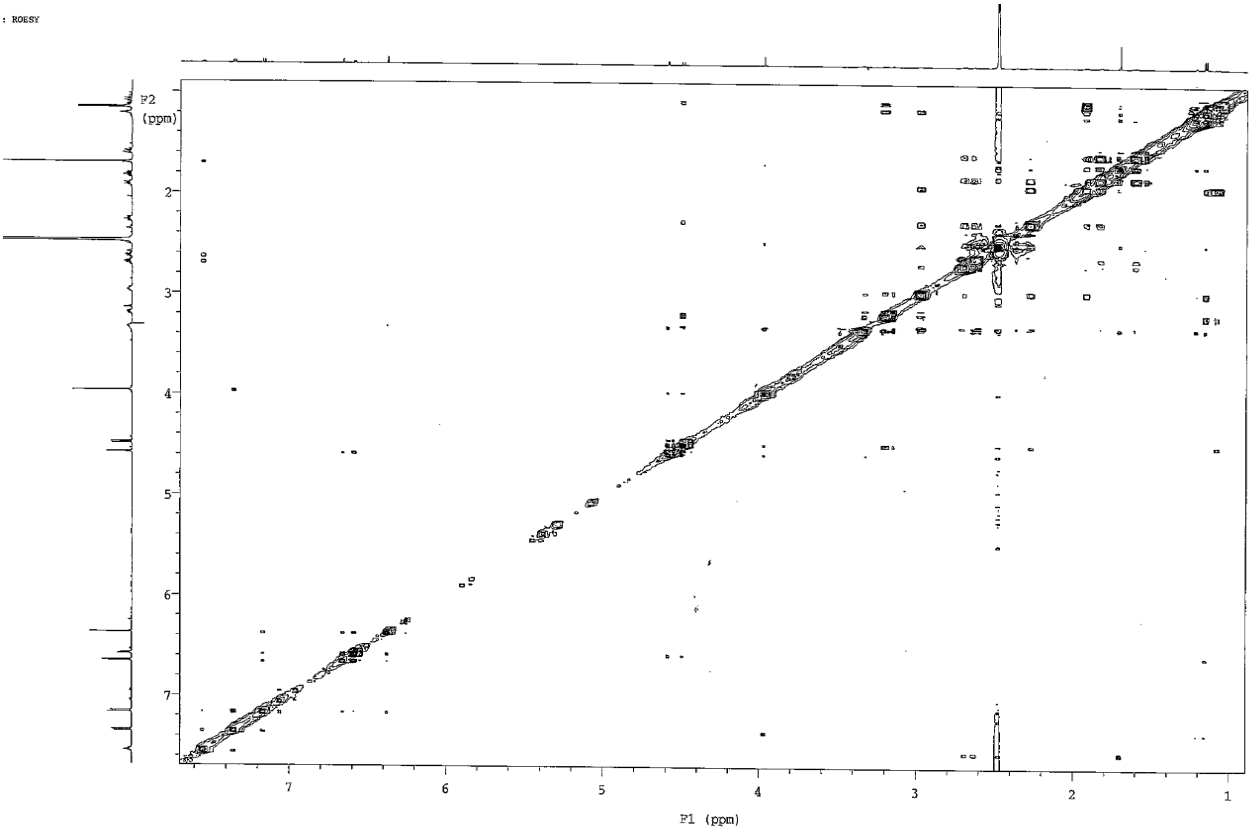
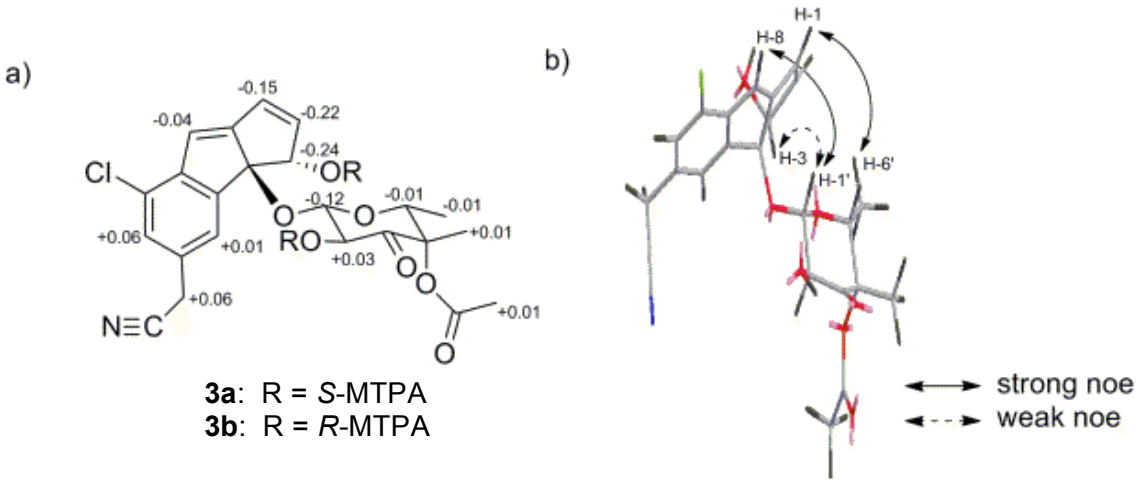


Figure S18. (a) $\Delta\delta_{S-R}$ of ^{1}H for *S*- and *R*-*bis*-MTPA esters of **3**, and (b) selected key ROESY correlations for determination of the configuration at C-3a for **3**.



3a: R = *S*-MTPA
3b: R = *R*-MTPA

Table S3. Proposed functions of proteins from analysis of the "*S. pacifica"* biosynthetic gene cluster (*cya*).

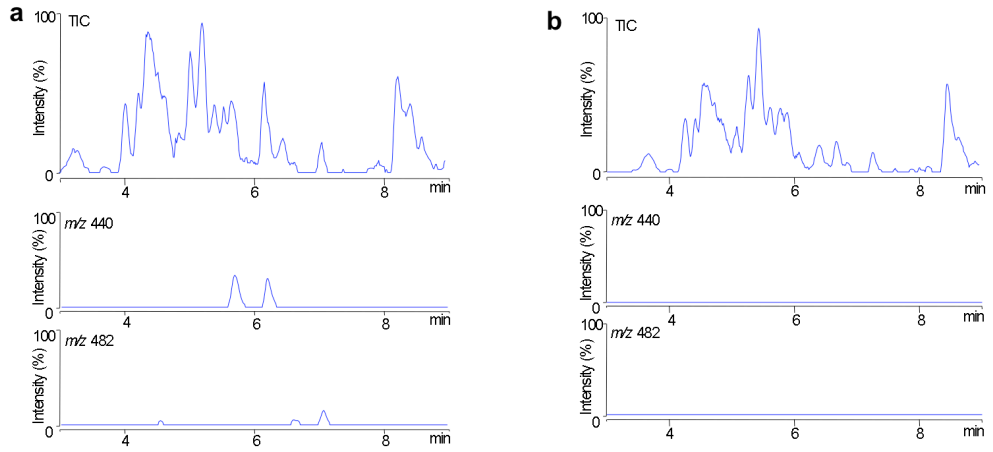
Gene	translated protein size (#aa)	annotation	homolog (NCBI accession number)	identity/ similarity (%)
orf(-6)	277	ABC transporter	Amir_0511 (YP_003098324)	54/70
orf(-5)	309	ABC transporter	AMED_8454 (ADJ50150)	67/79
orf(-4)	224	hypothetical protein	MCAG_00864 (ZP_04604607)	83/91
orf(-3)	327	transcription regulator	MCAG_00863 (ZP_04604606)	84/88
orf(-2)	205	hypothetical protein	Strop_2746 (YP_001159567)	98/98
orf(-1)	303	nucleotide kinase	Micau_0535 (YP_003833678)	60/72
		Predicted upstream *cya* cluster boundary		
cyaR1	146	transcription regulator	Strop_1067 (YP_001157916)	91/96
orf2	223	hypothetical protein	Strop_2742 (YP_001159564)	62/66
cyaR2	149	antibiotic resistance	ROP_05350 (YP_002777727)	67/79
cyaA2	335	TDP-hexose 4,6-dehydratase	LipDig2 (ABB05098)	66/78
cyaGT	392	glycosyltransferase	MdpA6 (ABY66016)	42/63
cyaA3	410	C-methyltransferase	NovU (AF170880_21)	36/52
orf7	198	dihydrofolate reductase	Strop_1085 (YP_001157934)	87/91
cyaA4	393	O-acyltransferase	MppN (ZP_06594431)	38/55
orf9	429	cytochrome P450 monooxygenase	NigD (ABC84463)	45/64
cyaR3	541	transport protein	Sros_0831 (YP_003336588)	53/70
cyaJ	159	conserved hypothetical protein	MdpJ (ABY66022)	47/66
orf12	176	unknown	BBR47_40930 (YP_002773574)	63/75
cyaJ2	96	conserved hypothetical protein	MdpJ (ABY66022)	54/77
cyaJ3	151	conserved hypothetical protein	MdpJ (ABY66022)	45/66
cyaN1	461	aminotransferase	MdpC7 (ABY66011)	58/73
cyaN2	519	oxidoreductase	AMED_6837 (ADJ48560)	72/82
cyaF	389	expoxide hydrolase	SgcF (AAL06662)	55/69
cyaD2	392	oxidoreductase	MdpD2 (ABY66021)	58/74
cyaA1	354	glucose-1-phosphate thymidylyltransferase	Lct49 (ABX71132)	63/76
cyaE11	272	conserved hypothetical protein	MdpE11 (ABY66001)	55/69
cyaM	360	conserved hypothetical protein	MdpM (ABY66000)	44/55
cyaE9	544	oxidoreductase	MdpE9 (ABY65999)	79/87
cyaE8	189	conserved hypothetical protein	MdpE8 (ABY65998)	54/69
orf24	347	sugar epimerase	Strop_2702 (YP_001159524)	62/76
cyaE2	308	conserved hypothetical protein	MdpE2 (ABY65995)	65/78
cyaE3	273	conserved hypothetical protein	MdpE3 (AAQ17107)	63/71
cyaE	1936	enediyne polyketide synthase	MdpE (AAQ17110)	63/73
cyaE10	147	thioesterase	MdpE10 (AAQ17111)	79/91
cyaE6	148	oxidoreductase	MdpE6 (AAQ17112)	64/76
cyaE7	288	cytochrome P450 hydroxylase	MdpE7 (AAQ17113)	65/75
cyaR4	292	AraC family transcriptional regulator	MdpR2 (ABY65996)	57/68
cyaR5	232	MerR transcription regulator	Sare_0557 (YP_001535477)	49/62
cyaE4	683	conserved hypothetical protein	MdpE4 (AAQ17108)	50/64
cyaE5	359	conserved hypothetical protein	MdpE5 (AAQ17109)	54/69
		Predicted downstream *cya* cluster boundary		
orf35	150	hypothetical protein	Sare_2908 (YP_001537725)	87/93
orf36	390	radical SAM-domain containing protein	Strop_2740 (YP_001159562)	90/93
orf37	314	hypothetical protein	Strop_2739 (YP_001159561)	86/93
orf38	252	peptidase	Strop_2738 (YP_001159560)	93/96

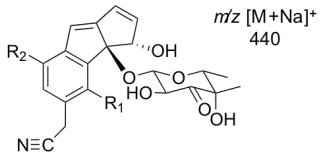
Table S4. Proposed functions of proteins from analysis of the *Streptomyces* sp. CNT-179 biosynthetic gene cluster (*cyn*) and comparison with proteins encoded by the "*S. pacifica"* gene cluster (*cya*).

CNT-179 gene	translated protein size (#aa)	annotation	homolog (NCBI accession number)	identity/ similarity (%)	*"S. pacifica"* homolog	identity/ similarity (%)
orf(-8)	313	helicase	Ndas_0664 (YP_003678615)	75/82	-	-
orf(-8)	356	recombinase A	Ndas_0666 (YP_003678617)	93/98	-	-
orf(-7)	192	recombination regulator	Ndas_0667 (YP_003678618)	75/85	-	-
orf(-6)	683	peptidase	Ndas_0668 (YP_003678619)	70/81	-	-
orf(-5)	328	TRAP transporter	Bsel_1504 (YP_003699582)	42/62	-	-
orf(-4)	161	hypothetical protein	HMP7215_0817 (ZP_06265972)	34/50	-	-
orf(-3)	688	TRAP transporter	SELS_00598 (ZP_05898028)	37/55	-	-
orf(-2)	496	RNA modification enzyme	Ndas_0673 (YP_003678624)	93/96	-	-
orf(-1)	423	cytochrome P450	Ema4 (AAT45263)	51/64	-	-
		Predicted upstream *cyn* cluster boundary				
cynR1	249	TetR family transcription regulator	Ndas_0675 (YP_003678626)	44/58	-	-
orf2	335	tRNA isopentenyltransferase	Ndas_0676 (YP_003678627)	82/89	-	-
cynE	1944	enediyne PKS	MdpE (AAQ17110)	55/67	CyaE	53/65
cynE10	158	thioesterase	MdpE10 (AAQ17111)	73/84	CyaE10	81/88
cynE6	172	oxidoreductase	MdpE6 (AAQ17112)	52/64	CyaE6	54/64
orf6	541	methylmalonyl-CoA decarboxylase	Mdp orf(-2) (ABY66032)	70/78	-	-
cynR2	223	TetR family transcription regulator	Mdp orf(-3) (ABY66033)	65/75	-	-
cynR3	274	AraC family transcription regulator	MdpR2 (ABY65996)	36/56	CyaR4	37/52
cynE2	324	conserved hypothetical protein	MdpE2 (ABY65995)	64/70	CyaE2	63/74
cynE3	344	conserved hypothetical protein	MdpE3 (AAQ17107)	55/67	CyaE3	59/65
cynE7	448	cytochrome P450 hydroxylase	MdpE7 (AAQ17113)	66/76	CyaE7	58/70
cynE8	189	conserved hypothetical protein	MdpE8 (ABY65998)	55/72	CyaE8	57/68
cynE9	550	oxidoreductase	MdpE9 (ABY65999)	78/87	CyaE9	74/84
cynM	343	conserved hypothetical protein	MdpM (ABY66000)	45/54	CyaM	38/50
cynE11	270	conserved hypothetical protein	MdpE11 (ABY66001)	63/75	CyaE11	58/70
cynA	416	*C*-methyltransferase	SgcA3 (AAL06661)	36/53	CyaA3	81/89
cynGT	446	glycosyltransferase	MdpA6 (ABY66016)	37/58	CyaGT	81/90
cynJ	154	hypothetical protein	MdpJ (ABY66022)	43/61	CyaJ	56/68
cynF	387	epoxide hydrolase	SgcF (AAL06662)	60/73	CyaF	64/76
cynD2	397	oxidoreductase	MdpD2 (ABY66021)	53/70	CyaD2	72/82
orf21	136	hypothetical protein	Cwoe_0858 (YP_003392666)	66/80	-	-
cynA3	463	glucose 3,4-dehydratase	UrdQ (AF269227_2)	73/83	-	-
cynGT2	394	glycosyltransferase	CalG4 (AF505622_37)	34/50	CyaGT	21/34
cynN2	536	oxidoreductase	AMED_6837 (YP_003768962)	63/76	CyaN2	65/78
cynN1	506	aminotransferase	MdpC7 (ABY66011)	56/73	CyaN1	70/82
cynA4	413	acyltransferase	MppN (ZP_06594431)	36/54	CyaA4	34/49
orf27	247	hypothetical protein	OB9013 (NP_691834)	61/76	-	-
cynR4	404	transport protein	Sare_3494 (YP_001538287)	46/60	-	-
orf29	431	cytochrome P450	NigD (ABC84463)	47/63	Orf9	68/80
cynA5	412	acyltransferase	MegY (AF263245_5)	39/57	CyaA4	63/74
orfR5	531	transport protein	MdpR3 (ABY66009)	49/67	CyaR3	60/77
orf32	147	heat shock protein	RER_51090 (YP_002768556)	50/65	-	-
cynR6	122	AsrR family transcription regulator	SACE_7306 (YP_001109389)	58/68	-	-

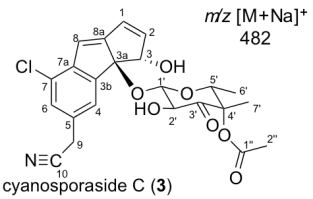
cynA1	288	glucose-1-phosphate thymidylyltransferase	LipDig1 (ABB05097)	71/85	CyaA1	36/55
cynA2	325	glucose 4,6-dehydratase	LipDig2 (ABB05098)	69/80	CyaA2	81/87
orf36	171	hypothetical protein	AMED_6842 (YP_003768967)	66/70	-	-
orf37	151	hypothetical protein	AMED_6843 (YP_003768968)	54/66	-	-
cynE4	644	conserved hypothetical protein	MdpE4 (AAQ17108)	38/51	CyaE4	40/52
cynE5	332	conserved hypothetical protein	MdpE5 (AAQ17109)	47/65	CyaE5	51/66
		Predicted downstream *cyn* cluster boundary				
orf40	270	diaminopimelate epimerase	Ndas_0677 (YP_003678628)	85/91	-	-
orf41	429	oxidase	Ndas_0681 (YP_003678632)	59/66	-	-
orf42	509	D-lactate oxidase/dehydrogenase	Ndas_0682 (YP_003678633)	77/86	-	-
orf43	567	malate synthase	Ndas_0683 (YP_003678634)	89/95	-	-
orf44	508	GTP-binding protein Hf lx	Ndas_0684 (YP_003678635)	82/87	-	-
orf45	409	imidazolonepropionase	Ndas_0685 (YP_003678636)	91/95	-	-
orf46	337	DNA polymerase III, Delta subunit	Ndas_0686 (YP_003678637)	90/94	-	-

Figure S19. LC/MS profiles of chemical extracts from wild-type "*S. pacifica*" CNS-143 and gene replacement mutants lacking hypothesized genes required for nitrile functionalization. (a) Wild type "*S. pacifica*," (b) *cyaN2::apr*^{R} oxidoreductase mutant; (c) *orf9::apr*^{R} cytochrome P450 mutant; (d) *cyaN1::apr*^{R} aminotransferase mutant. For all, the total ion count (TIC) is shown in the top panel, the middle panel shows the selected ion recording for $[M+Na]^+$ *m/z* 441 of cyanosporasides A-B (**1-2**), and the bottom panel depicts $[M+Na]^+$ *m/z* 482, representing $[M+Na]^+$ for cyanosporaside C (**3**) as well as an uncharacterized regioisomer of **3** presumed to bear the chlorine substituent at C-4 analogous to cyanosporaside A (**1**).





cyanosporaside A (**1**) R_1 = Cl, R_2 = H
cyanosporaside B (**2**) R_1 = H, R_2 = Cl



cyanosporaside C (**3**)

Figure S20. LC/MS profiles of chemical extracts from wild-type "*S. pacifica*" CNS-143 and glycosylation gene knockout mutants. (a) Wild type "*S. pacifica*," (b) *cyaGT::apr*^{R} glycosyltransferase mutant; (c) *cyaA4::apr*^{R} acyltransferase mutant; (d) *orf24::apr*^{R} epimerase mutant; (e) *cyaA3::apr*^{R} C-methyltransferase mutant. For all, the total ion count (TIC) is shown in the top panel, the middle panel shows the selected ion recording for $[M+Na]^+$ *m/z* 441 of cyanosporasides A-B (**1-2**), and the bottom panel depicts $[M+Na]^+$ *m/z* 482, representing $[M+Na]^+$ for cyanosporaside C (**3**) as well as an uncharacterized regioisomer of **3** presumed to bear the chlorine substituent at C-4 analogous to cyanosporaside A (**1**).

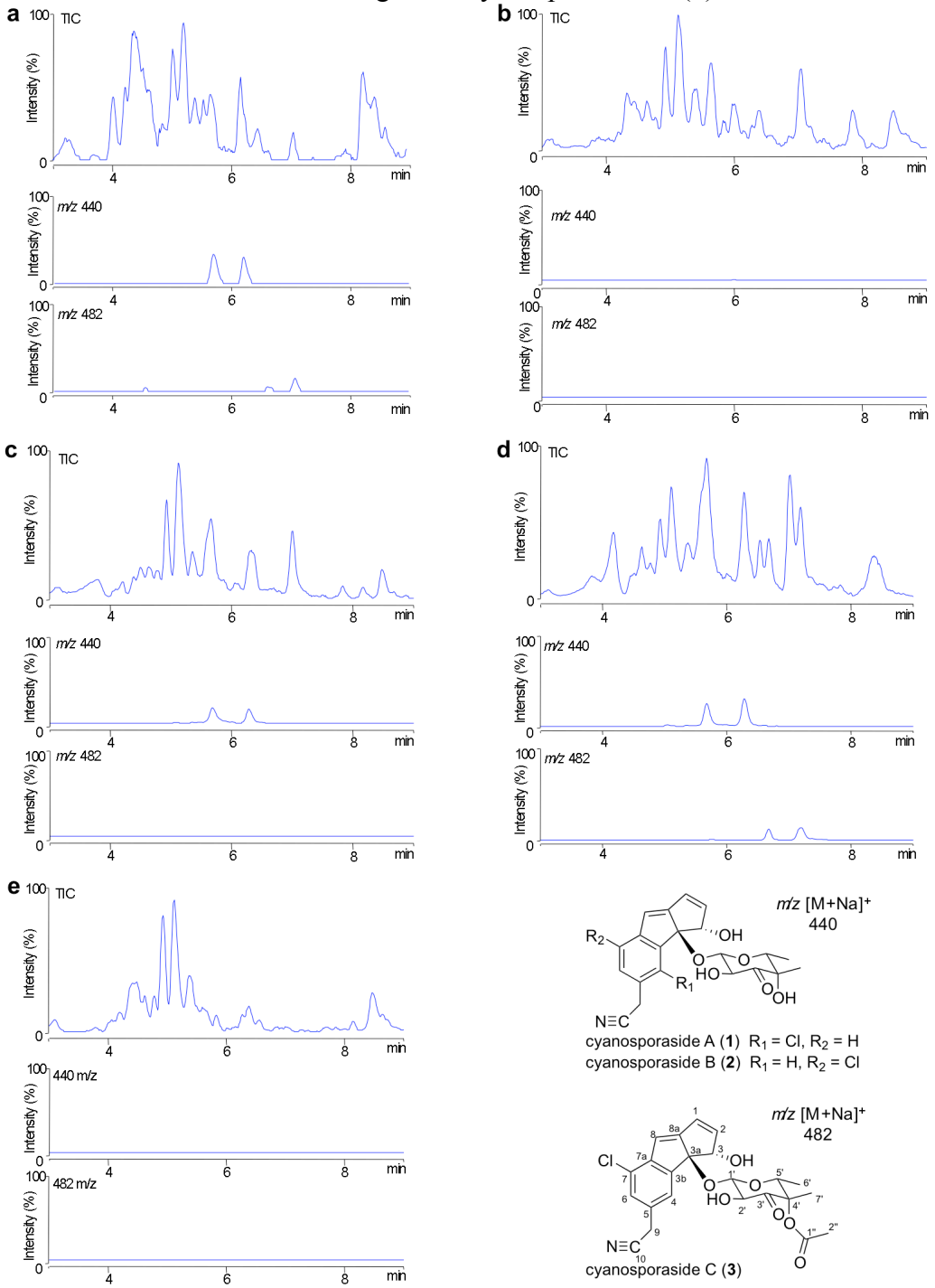
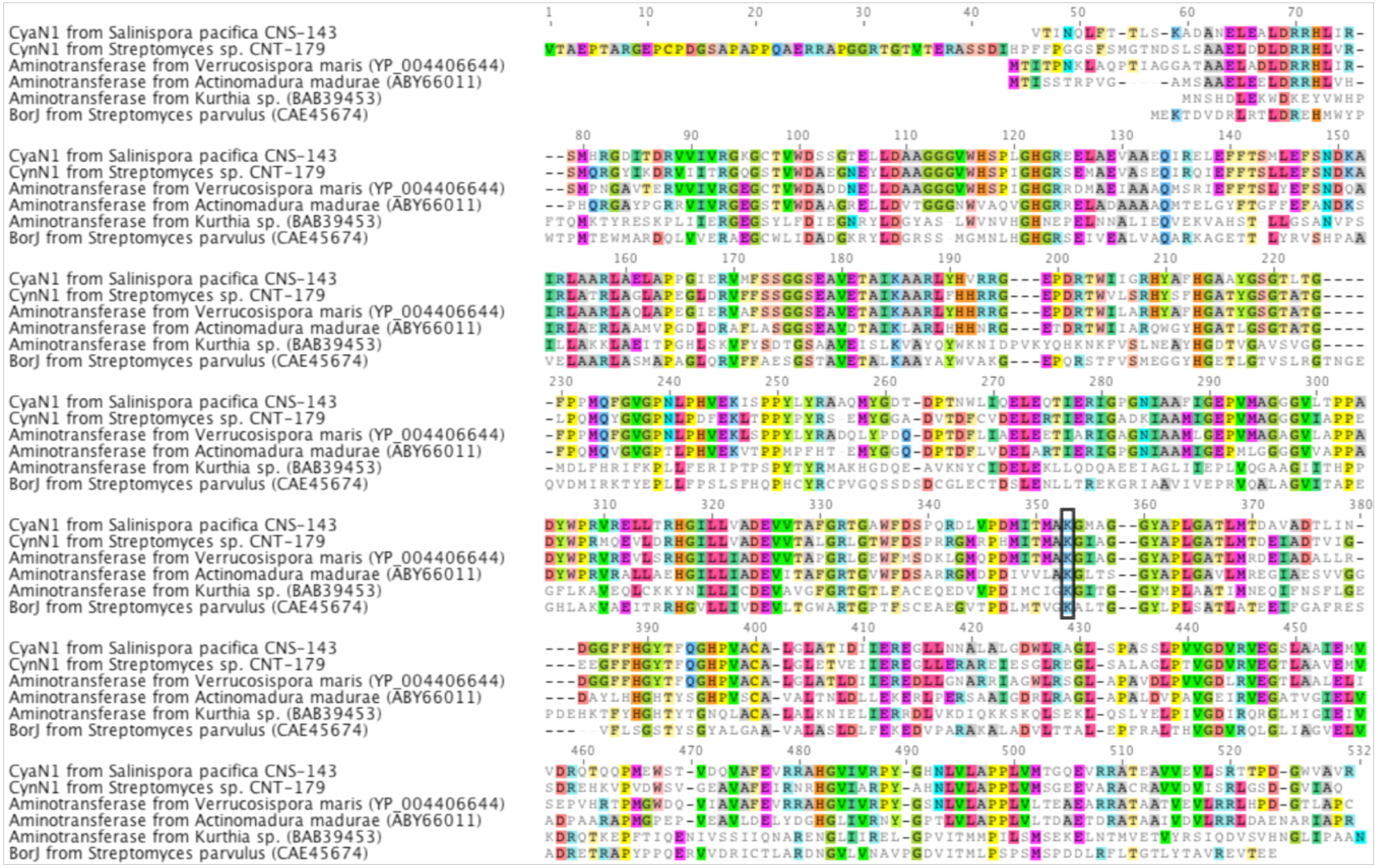


Figure S21. Alignment of aminotransferases CyaN1 and CynN1 with selected homologous group III aminotransferases including BorJ, required for nitrile functionalization in borrelidin biosynthesis. The conserved PLP-binding lysine residue is highlighted with a box. Alignments were prepared with ClustalW.



References

(1) Kieser, T.; Bibb, M. J.; Buttner, M. J.; Chater, K. F.; Hopwood, D. A. *Practical Streptomyces Genetics*; The John Innes Foundation: Norwich, 2000.

(2) Liu, W.; Ahlert, J.; Gao, Q. J.; Wendt-Pienkowski, E.; Shen, B.; Thorson, J. S. *Proc. Natl. Acad. Sci. U.S.A.* 2003, *100*, 11959-11963.

(3) Ishikawa, J.; Hotta, K. *FEMS Microbiol. Letters* 1999, *174*, 251-253.

(4) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. *J. Mol. Biol.* 1990, *215*, 403-410.

(5) Ziemert, N.; Podell, S.; Penn, K.; Badger, J.; Allen, E.; Jensen, P. R. *PloS One* **2012**, *7*, e34064.

(6) Gust, B.; Challis, G. L.; Fowler, K.; Kieser, T.; Chater, K. F. *Proc. Natl. Acad. Sci. U.S.A.* 2003, *100*, 1541-1546.

(7) Eustáquio, A. S.; Pojer, F.; Noel, J. P.; Moore, B. S. *Nat. Chem. Biol.* **2008**, *4*, 69-74.

(8) Elespuru, R. K.; White, R. J. *Cancer Research* 1983, *43*, 2819-2830.