

Supporting Information

Marinoterpins A–C, Rare Linear Merosesterterpenoids from Marine-Derived Actinomycete Bacteria of the Family Streptomycetaceae

Min Cheol Kim,[†] Jaclyn M. Winter,^{*,‡} Ratnakar N. Asolkar,[†] Chollaratt Boonlarpradab,[†] Reiko Cullum,[†] and William Fenical^{*,†,§}

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

[‡] Department of Medicinal Chemistry, University of Utah, Salt Lake City, Utah 84112, USA,

[§] Skaggs School of Pharmacy and Pharmaceutical Science, University of California, San Diego, La Jolla, CA, 92093, USA

Corresponding Authors

William Fenical

wfenical@ucsd.edu

Phone: 858-534-2133

Jaclyn M. Winter

Jaclyn.winter@utah.edu

Phone: 801-585-7117

Table of Contents

Table S1. 2D COSY, HMBC, and ROESY NMR data of marinoterpin A (1) (CD ₃ OD, 500 MHz).....	S3
Table S2. 2D COSY, HMBC, and NOESY NMR data of marinoterpin B (2) (CD ₃ OD, 500 MHz).....	S4
Table S3. 2D COSY, HMBC, and ROESY NMR data of marinoterpin C (3) (CD ₃ OD, 500 MHz).....	S5
Figure S1. ¹ H NMR spectrum of marinoterpin A (1) (500 MHz) in CD ₃ OD.....	S6
Figure S2. ¹³ C NMR spectrum of marinoterpin A (1) (125 MHz) in CD ₃ OD.....	S7
Figure S3. ¹ H- ¹ H COSY spectrum of marinoterpin A (1) (500 MHz) in CD ₃ OD.....	S8
Figure S4. HSQC spectrum of marinoterpin A (1) (500 MHz) in CD ₃ OD.....	S9
Figure S5. HMBC spectrum of marinoterpin A (1) (500 MHz) in CD ₃ OD.....	S10
Figure S6. ROESY spectrum of marinoterpin A (1) (500 MHz) in CD ₃ OD.....	S11
Figure S7. Homo 2D <i>J</i> -resolved ¹ H NMR spectrum of marinoterpin A (1) (500 MHz) in CD ₃ OD.....	S12
Figure S8. HR-ESI-TOFMS spectrum of marinoterpin A (1).....	S13
Figure S9. ¹ H NMR spectrum of marinoterpin B (2) (500 MHz) in CD ₃ OD.....	S14
Figure S10. ¹³ C NMR spectrum of marinoterpin B (2) (125 MHz) in CD ₃ OD.....	S15
Figure S11. ¹ H- ¹ H COSY NMR spectrum of marinoterpin B (2) (500 MHz) in CD ₃ OD.....	S16
Figure S12. HMQC NMR spectrum of marinoterpin B (2) (500 MHz) in CD ₃ OD.....	S17
Figure S13. HMBC NMR Spectrum of marinoterpin B (2) (500 MHz) in CD ₃ OD.....	S18
Figure S14. Mass spectrum of marinoterpin B (2).....	S19
Figure S15. ¹ H NMR Spectrum of marinoterpin C (3) (500 MHz) in CD ₃ OD.....	S20
Figure S16. ¹³ C NMR spectrum of marinoterpin C (3) (125 MHz) in CD ₃ OD.....	S21
Figure S17. ¹ H- ¹ H COSY NMR spectrum of marinoterpin C (3) (500 MHz) in CD ₃ OD.....	S22
Figure S18. HMQC NMR spectrum of marinoterpin C (3) (500 MHz) in CD ₃ OD.....	S23
Figure S19. HMBC NMR spectrum of marinoterpin C (3) (500 MHz) in CD ₃ OD.....	S24
Figure S20. <i>mrt</i> biosynthetic cluster from <i>Streptomyces</i> . sp. AJS-327.....	S25
Table S4. Annotation of the 16 open reading frames of the <i>mrt</i> gene cluster from <i>Streptomyces</i> sp. AJS-327.....	S25
Figure S21. Phylogenetic analysis of aryl-CoA ligases.....	S26
Figure S22. Phylogenetic analysis of polyprenyl synthases.....	S27
Figure S23. Phylogenetic analysis of prenyltransferases.....	S28
Figure S24. Phylogenetic analysis of flavin monooxygenases.....	S29
References	S29

Table S1. 2D COSY, HMBC, and ROESY NMR data for marinoterpin A (**1**) (CD₃OD, 500 MHz).

Marinoterpin A (1)				
No.	δ_H/δ_C	COSY	HMBC	ROESY
1 -N-O				
2	-/147.6			
3	-/135.5			
4	8.48/129.5		C-1', C-2, C-5, C-8a	H-5, H-2a', H-2b'
4a	-/128.9			
5	8.16/130.8	H-6	C-4, C-7, C-8a	H-4
6	7.79/130.2	H-5, H-7	C-4a, C-7, C-8	
7	7.97/134.2	H-6, H-8	C-5, C-8a	
8	8.66/120.0	H-7	C-4a, C-6, C-7, C-8a	
8a	-/142.4			
9	2.78/16.1		C-2, C-3, C-8, C-8a	
1'	-/202.9			
2'	3.12/49.8		C-1', C-3', C-4', C-21'	H-4, H-3'
	2.97		C-1', C-3', C-4', C-21'	
3'	2.17/30.3		C-1', C-2', C-4', C-5', C-21'	H-21'
4'	1.51/ 37.4		C-3', C-5', C-6', C-21'	
	1.37			
5'	1.99/27.8		C-4', C-6', C-7'	
6'	5.33/121.3		C-8', C-22'	H-8a'
7'	-/140.1			
8'	2.59/40.3	H-9'	C-6', C-7', C-22'	
	2.18	H-9'	C-7', C-9'	
9'	4.57/77.5	H-8', H-10'	C-7', C-11', C-22'	
10'	5.19/126.2	H-9', H-23'	C-8', C-9', C-23'	
11'	-/140.7			
12'	2.04/40.0		C-10', C-13', C-14', C-23'	
13'	2.02/28.6		C-11', C-12', C-14'	
14'	5.31/120.8		C-16', C-24'	H-16a'
15'	-/140.2			
16'	2.59/40.2	H-17'	C-14', C-15'	
	2.18	H-17'	C-15', C-17'	
17'	4.57/77.4	H-16', H-18'	C-15', C-19', C-24'	
18'	5.19/125.9	H-17', H-25'	C-16', C-17', C-20', C-25'	
19'	-/138.0			
20'	1.73/25.9		C-18', C-19', C-25'	
21'	1.03/20.1		C-2', C-3', C-4'	H-3'
22'	4.39/69.0		C-6', C-7', C-8', C-9'	
	4.23			
23'	1.70/16.7		C-9', C-10', C-11', C-12'	
24'	4.39/69.0		C-14', C-15', C-16', C-17'	
	4.23			
25'	1.70/18.3		C-17', C-18', C-19', C-20'	

Table S2. 2D COSY, HMBC, and NOESY NMR data for marinoterpin B (**2**) (CD₃OD, 500 MHz).

Marinoterpin B (2)				
No.	δ_H/δ_C	COSY	HMBC	NOESY
1 -N-O				
2	-/149.6			
3	-/134.4			
4	7.89/129.2		C-2, C-3, C-8a, C-1'	H-5
4a	-/129.8			
5	7.97/129.3	H-6	C-4, C-7, C-8a	H-4, H-6
6	7.69/129.5	H-5, H-7	C-4a, C-8	
7	7.82/131.7	H-6, H-8	C-5, C-8a	H-6, H-8
8	8.60/119.7	H-7	C-4a, C-6	H-7
8a	-/138.5			
9	2.73/15.0		C-2, C-3	H-1'
1'	3.61/33.1	H-2'	C-2, C-3, C-4, C-2', C-3'	H-9, H-21'
2'	5.33/122.1	H-1'	C-1', C-4', C-21'	H-4'
3'	-/139.4			
4'	2.18/40.2	H-5'	C-5'	H-2'
5'	2.11/28.7	H-4', H-6'	C-6', C-7'	H-22a', H-22b'
6'	5.32/121.0	H-5'	C-22'	H-8a', H-8b'
7'	-/140.2			
8'	2.18/40.3	H-9'	C-7'	H-6'
	2.58	H-9'	C-7', C-22'	H-6', H-9'
9'	4.55/77.5	H-8', H-10'	C-22'	H-8b', H-23'
10'	5.18/126.3	H-9'	C-8', C-23'	H-12'
11'	-/140.7			
12'	2.06/40.0	H-13'	C-10', C-11', C-13', C-23'	H-10'
13'	2.05/28.6	H-12', H-14'	C-14', C-15'	H-24a', H-24b'
14'	5.29/120.8	H-13'		H-16a', H-16b'
15'	-/140.2			
16'	2.18/40.3	H-17'	C-15'	H-14'
	2.58	H-17'	C-15', C-24'	H-17'
17'	4.55/77.4	H-16', H-18'	C-24'	H-16b', H-25'
18'	5.18/125.9	H-17'	C-16', C-20', C-25'	H-20'
19'	-/138.0			
20'	1.72/25.9		C-18', C-19', C-25'	H-18'
21'	1.79/16.4		C-2', C-3', C-4'	H-1'
22'	4.21 d (13)/69.0	H-6'	C-6', C-7'	H-5'
	4.38 t (12.5)	H-6'	C-6', C-7', C-9'	H-5'
23'	1.67/16.7		C-10', C-11', C-12'	H-9'
24'	4.21/ 69.0	H-14'	C-14', C-15'	H-13'
	4.38	H-14'	C-14', C-15', C-17'	H-13'
25'	1.69/18.3		C-18', C-19', C-20'	H-17'

Table S3. 2D COSY and HMBC NMR data for marinoterpin C (**3**) (CD₃OD, 500 MHz).

Marinoterpin C (3)			
Position	δ_H/δ_C	COSY	HMBC
1 -N			
2	-/163.1		
3	-/138.3		
4	7.67/135.5		C-2, C-3, C-5', C-8a, C-1'
4a	-/121.3		
5	7.64/128.2	H-6	C-4, C-7, C-8a
6	7.29/122.6	H-5, H-7	C-4a, C-8
7	7.60/129.8	H-6, H-8	C-5, C-8a
8	7.55/114.4	H-7	C-4a, C-6
8a	-/138.5		
9	4.55/77.1		C-2, C-3
1'	3.33/28.9	H-2'	C-2, C-3, C-4, C-2', C-3'
2'	5.41/121.2	H-1'	C-1', C-4', C-21'
3'	-/138.2		
4'	2.19/39.1	H-5'	C-5'
5'	2.13/27.7	H-4', H-6'	C-6', C-7'
6'	5.37/129.6	H-5'	C-22'
7'	-/138.8		
8'	2.18/39.1	H-9'	C-7'
	2.58	H-9'	C-7', C-22'
9'	4.55/76.3	H-8', H-10'	C-22'
10'	5.19/124.8	H-9'	C-8', C-23'
11'	-/139.6		
12'	2.06/38.5	H-13'	C-10', C-11', C-23'
13'	2.05/28.9	H-12', H-14'	C-14', C-15'
14'	5.30/120.1	H-13'	
15'	-/137.1		
16'	2.58/39.5	H-17'	C-15', C-24'
17'	4.55/76.1	H-16', H-18'	C-24'
18'	5.18/125.2	H-17'	C-16', C-20', C-25'
19'	-/139.6		
20'	1.73/24.8		C-18', C-19', C-25'
21'	1.61/16.1		C-2', C-3', C-4'
22'	4.21/67.9	H-6'	C-6', C-7'
	4.39	H-6'	C-6', C-7', C-9'
23'	1.73/15.1		C-10', C-11', C-12'
24'	4.21/67.9	H-14'	C-14', C-15'
	4.39	H-14'	C-14', C-15', C-17'
25'	1.71/15.5		C-18', C-19', C-20'

Figure S1. ^1H NMR spectrum of marinoterpin A (**1**) (500 MHz) in CD_3OD .

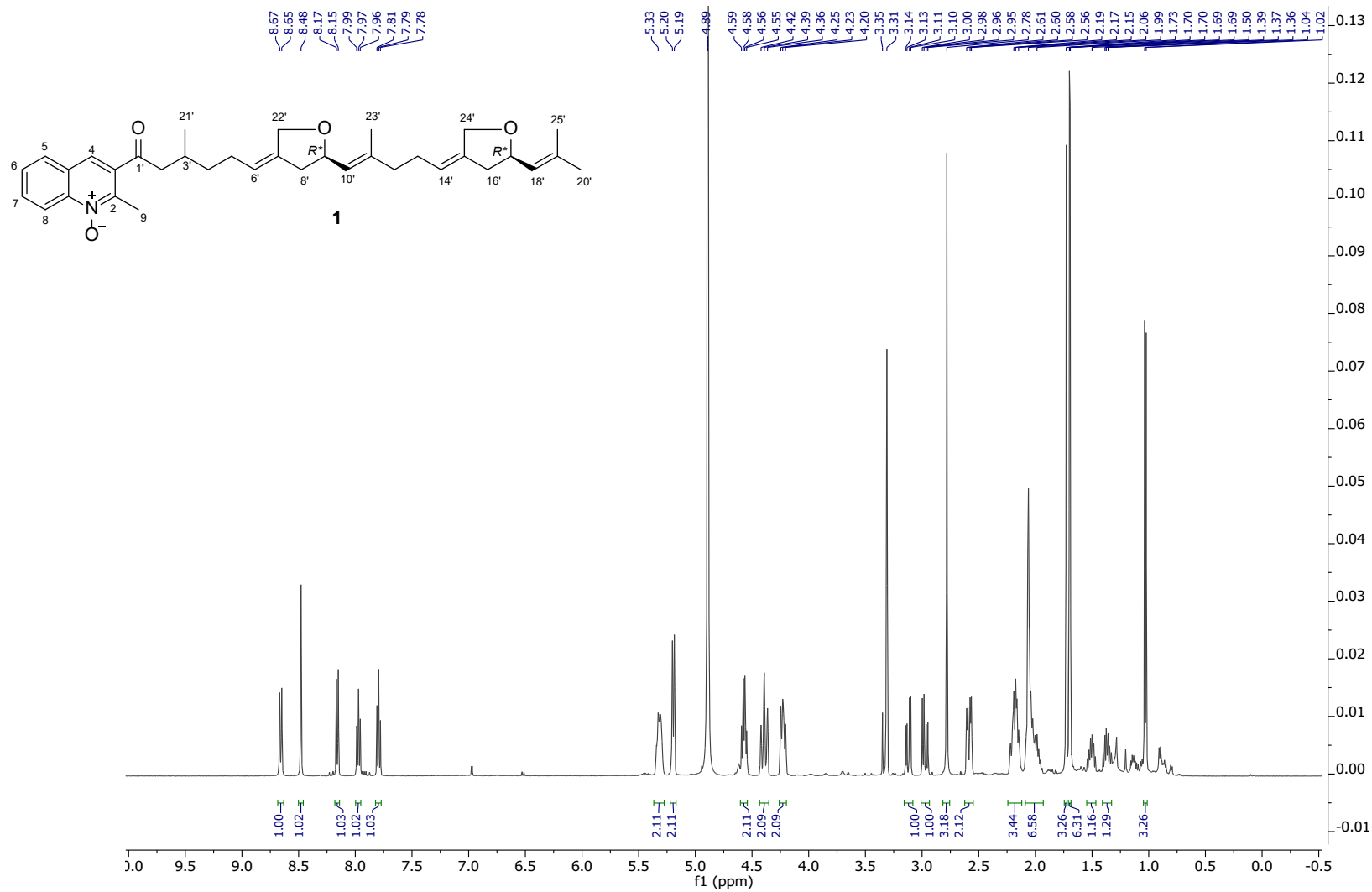


Figure S2. ^{13}C NMR spectrum of marinoterpin A (**1**) (125 MHz) in CD_3OD .

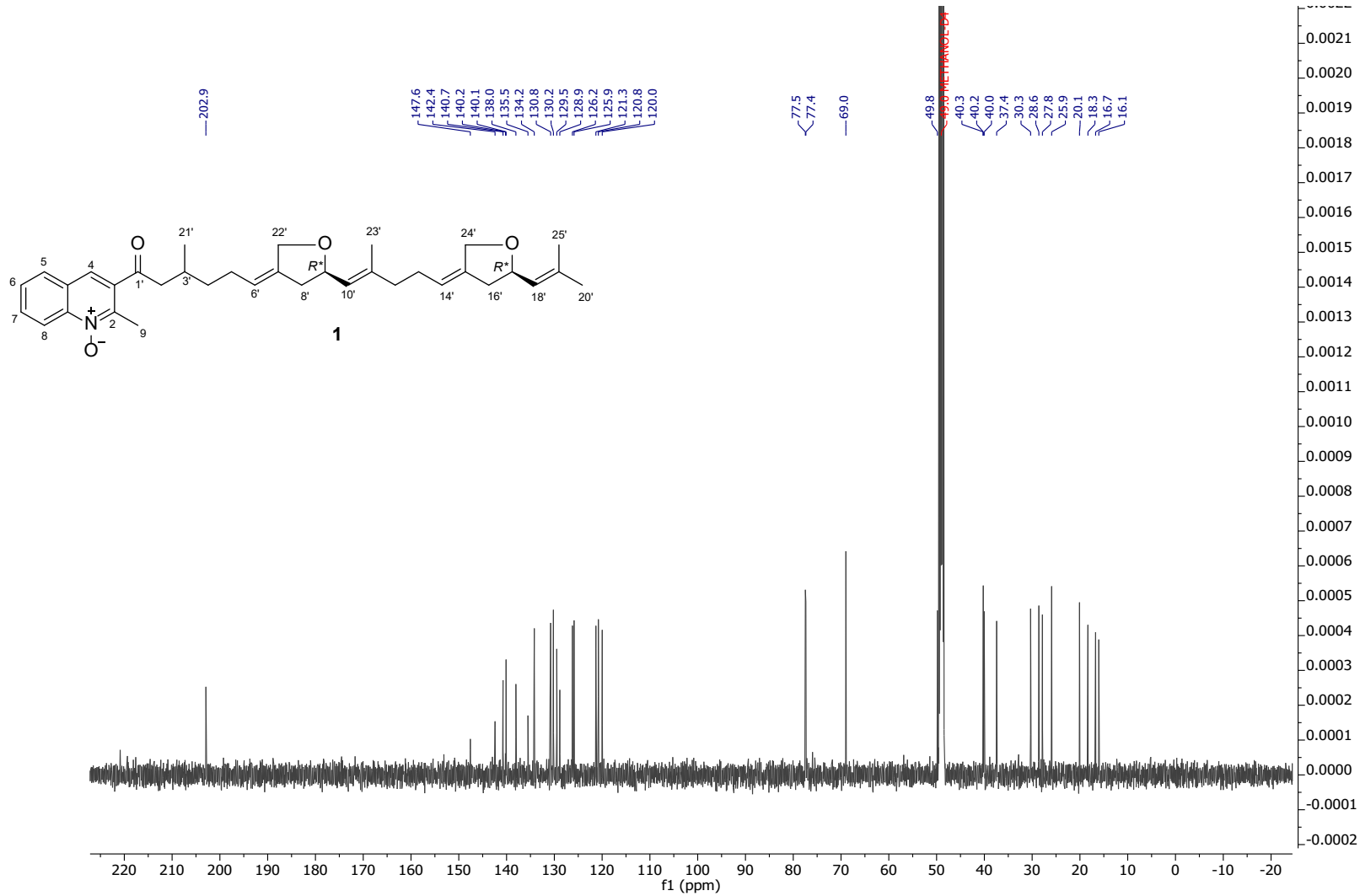


Figure S3. ^1H - ^1H COSY spectrum of marinoterpin A (**1**) (500 MHz) in CD_3OD .

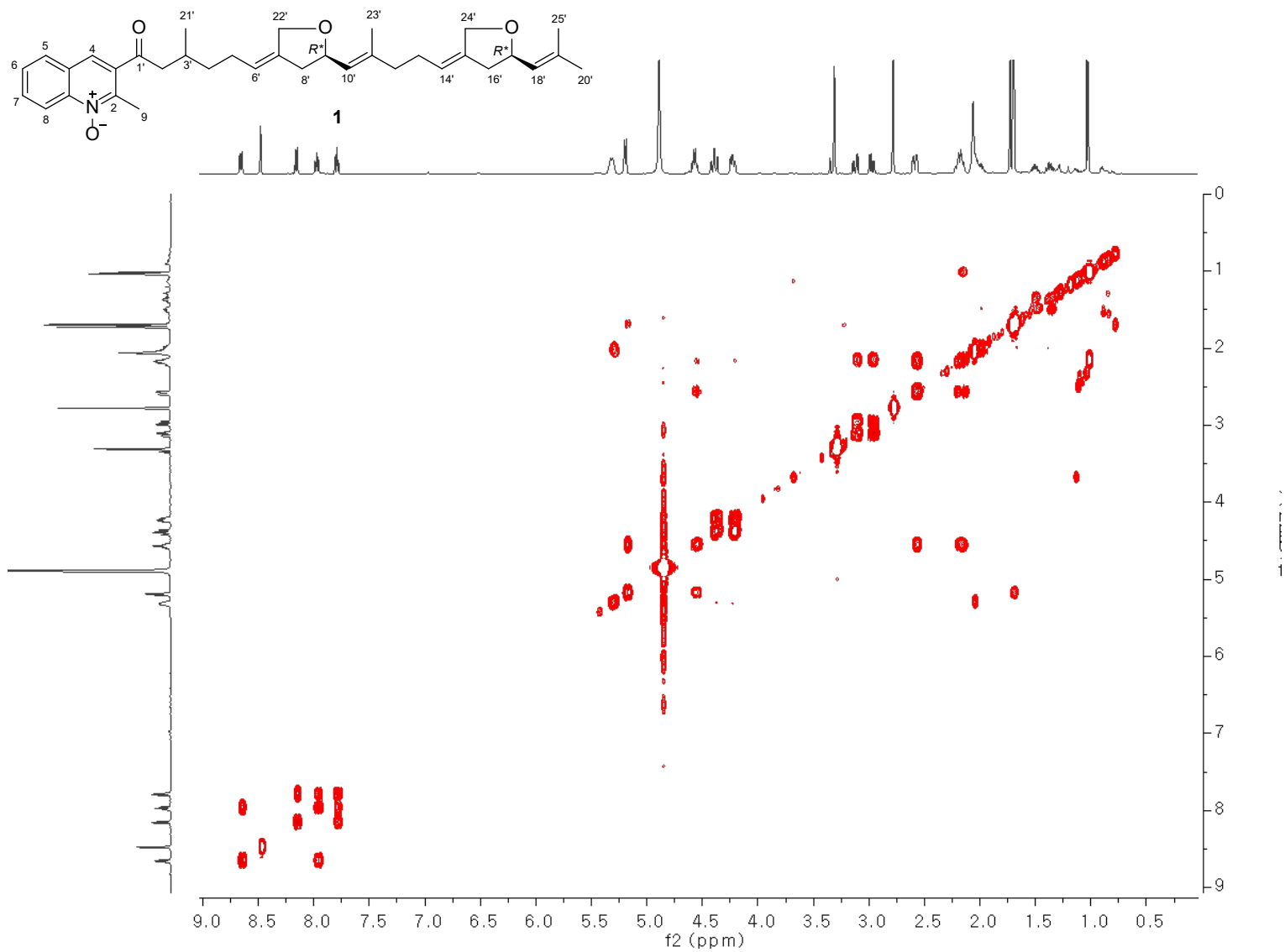


Figure S4. HSQC spectrum of marinoterpin A (**1**) (500 MHz) in CD₃OD.

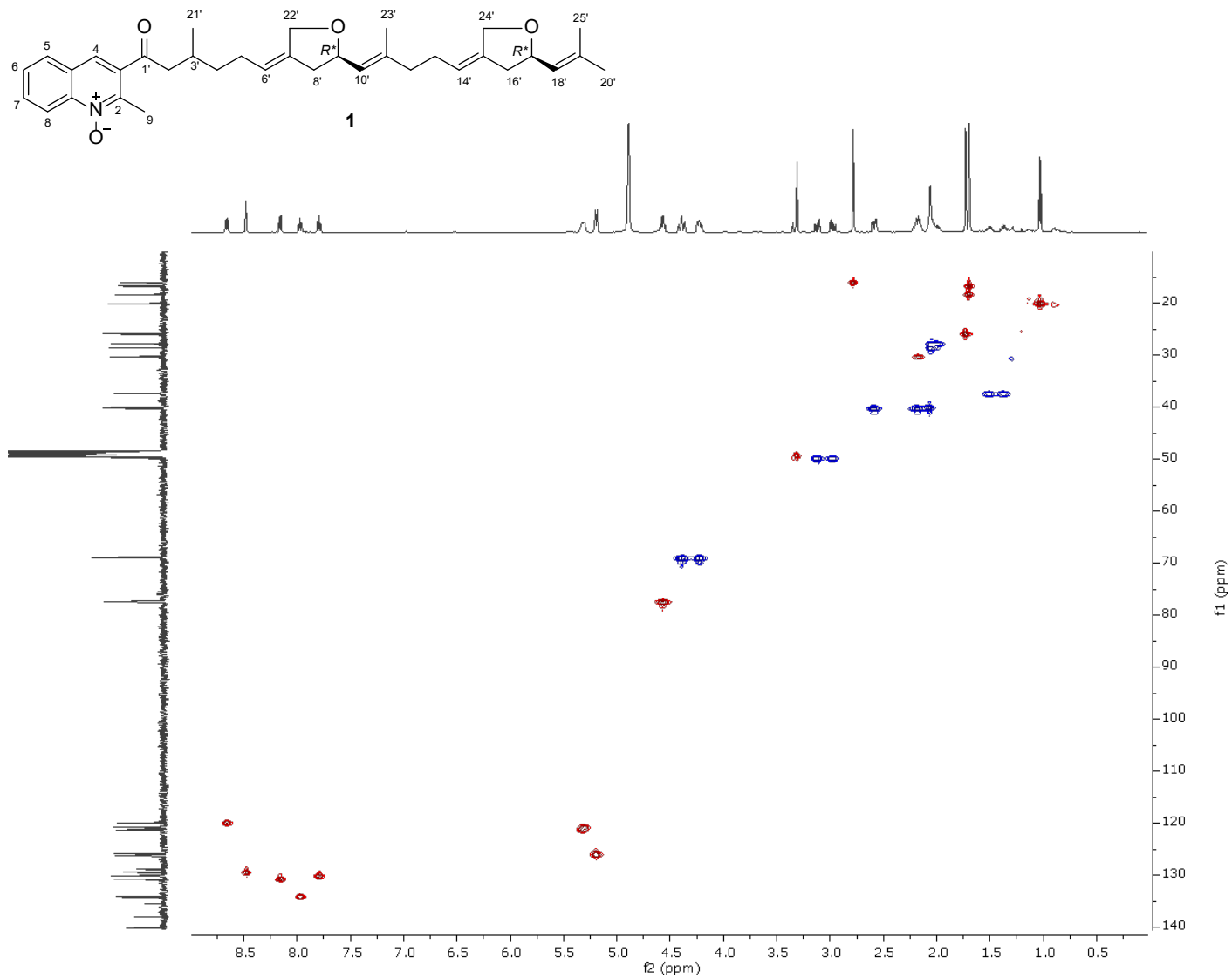


Figure S5. HMBC spectrum of marinoterpin A (**1**) (500 MHz) in CD₃OD.

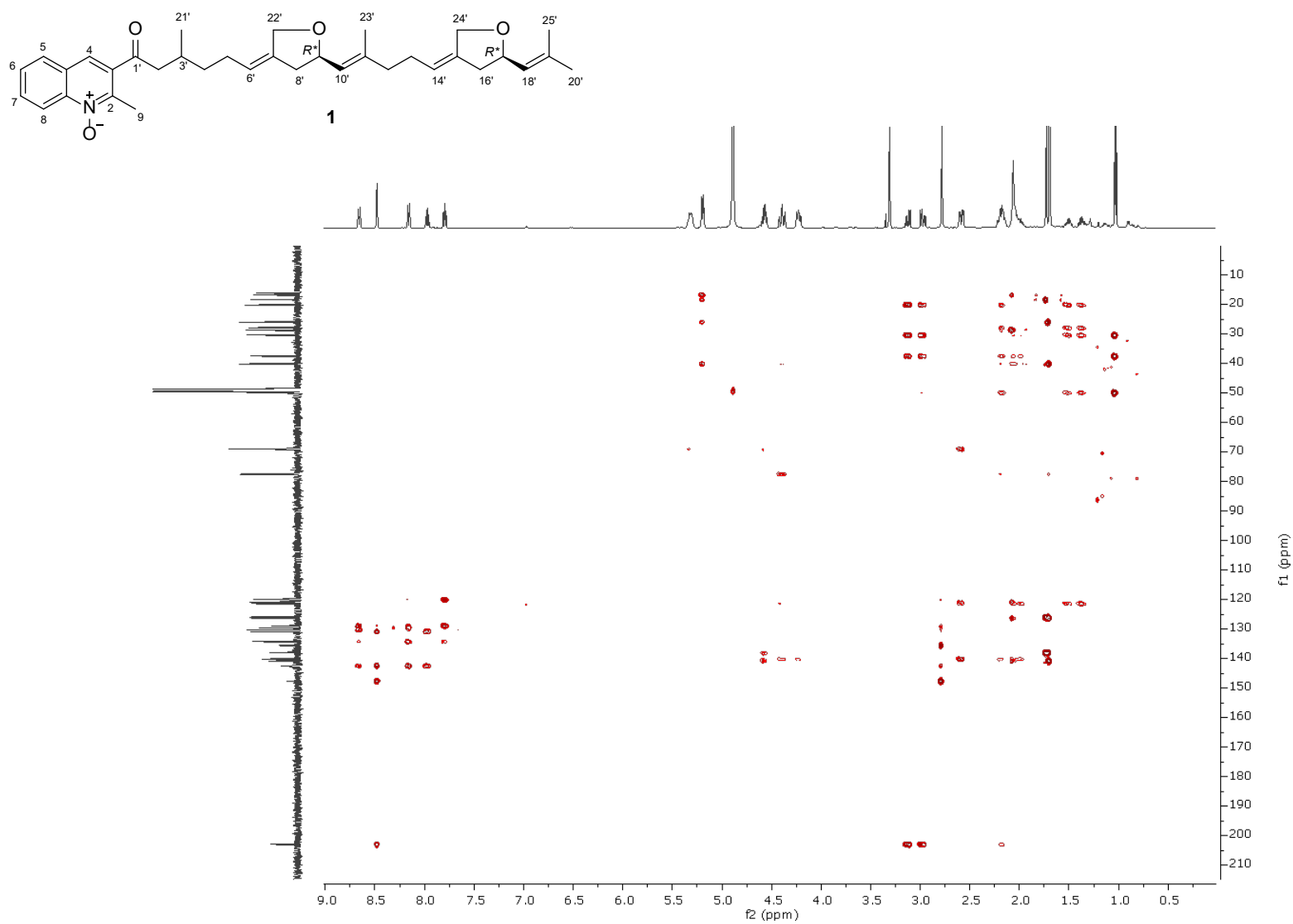


Figure S6. ROESY spectrum of marinoterpin A (**1**) (500 MHz) in CD₃OD.

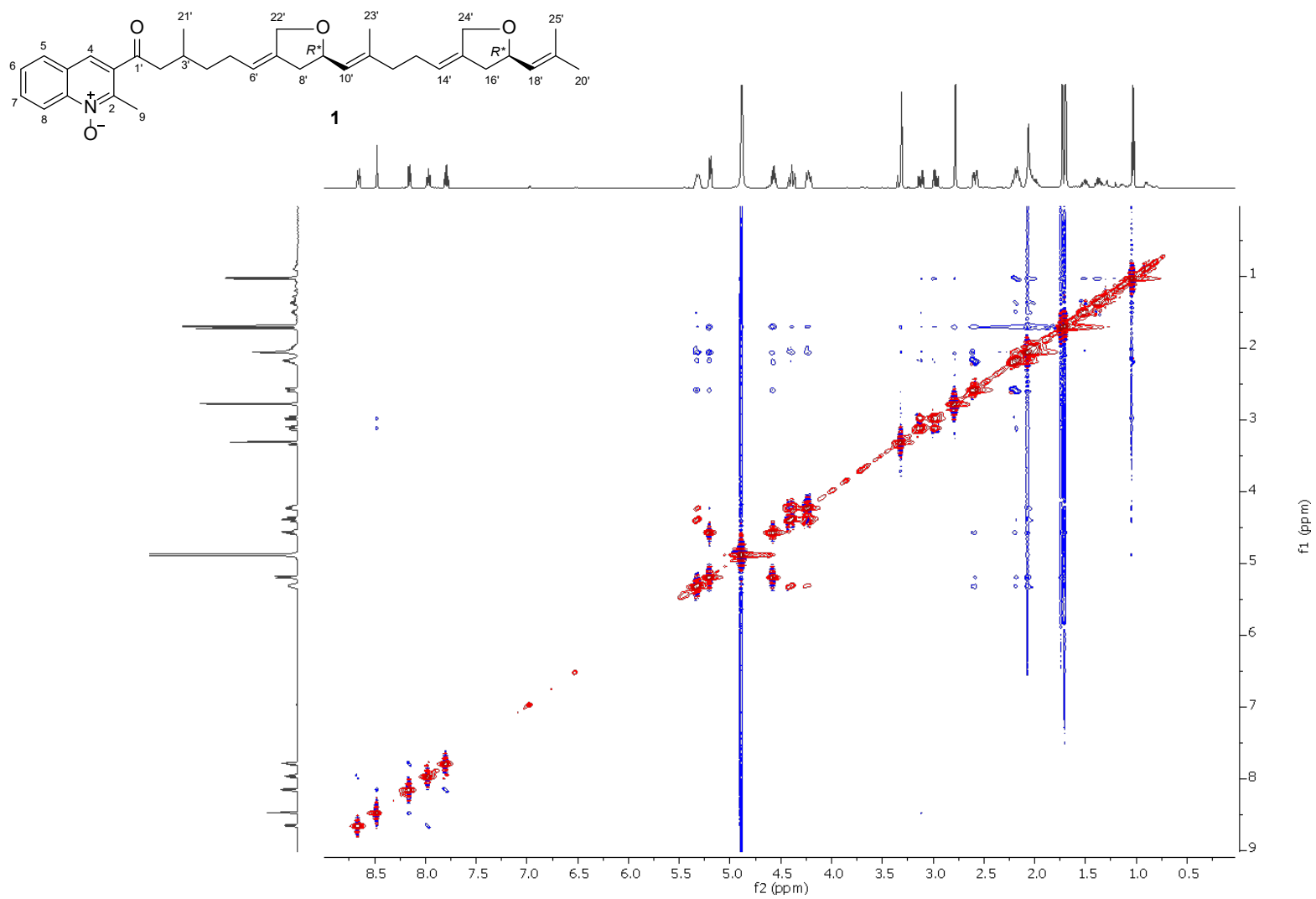


Figure S7. Homo 2D *J*-resolved ^1H NMR spectrum of marinoterpin A (**1**) (500 MHz) in CD_3OD .

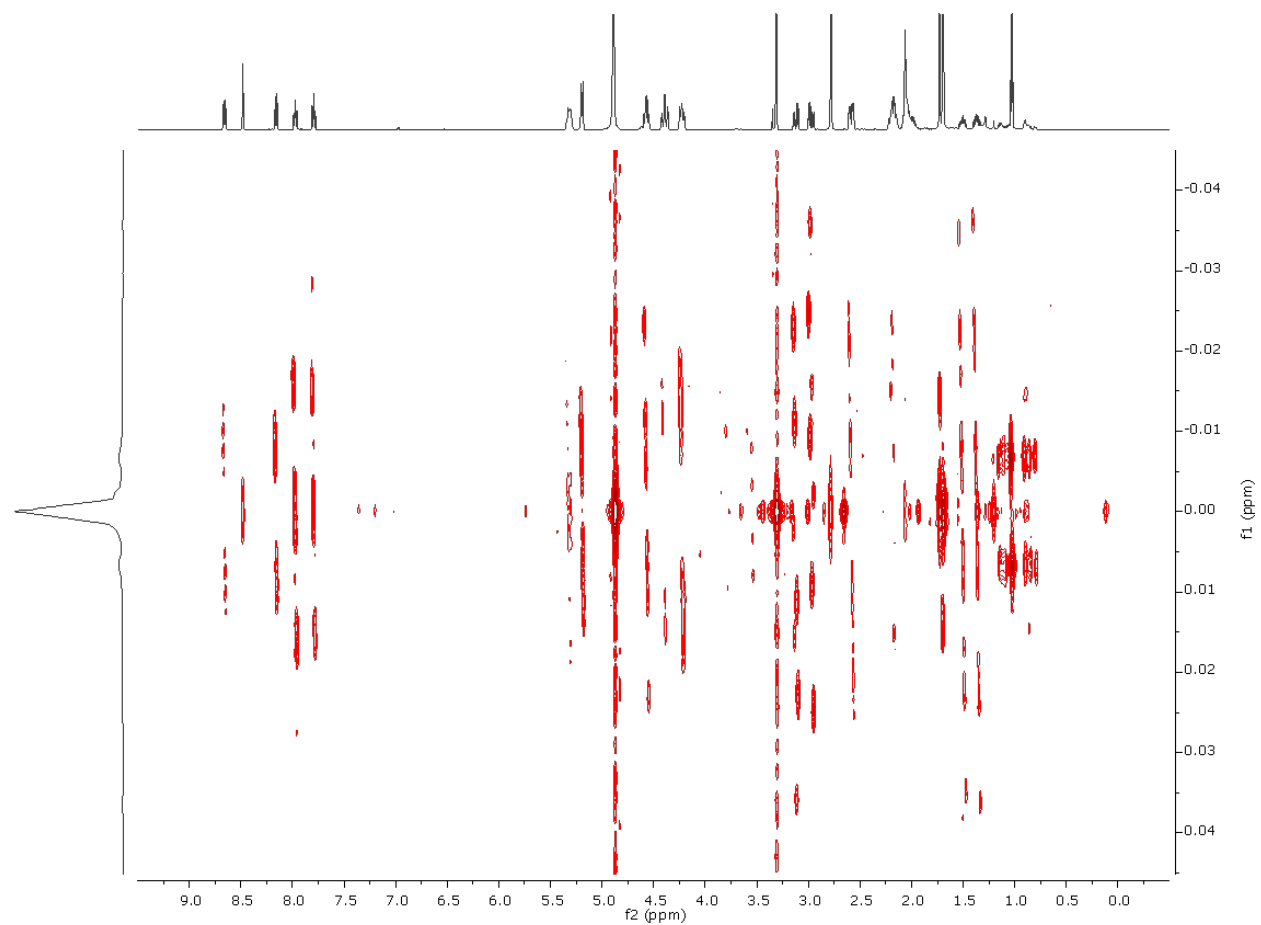
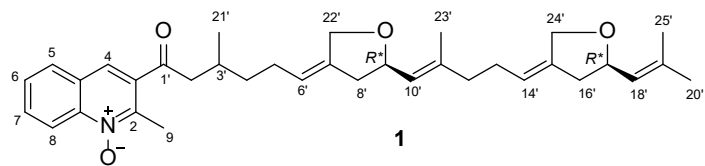


Figure S8. HR-ESI-TOFMS spectrum of marinoterpin A (**1**).

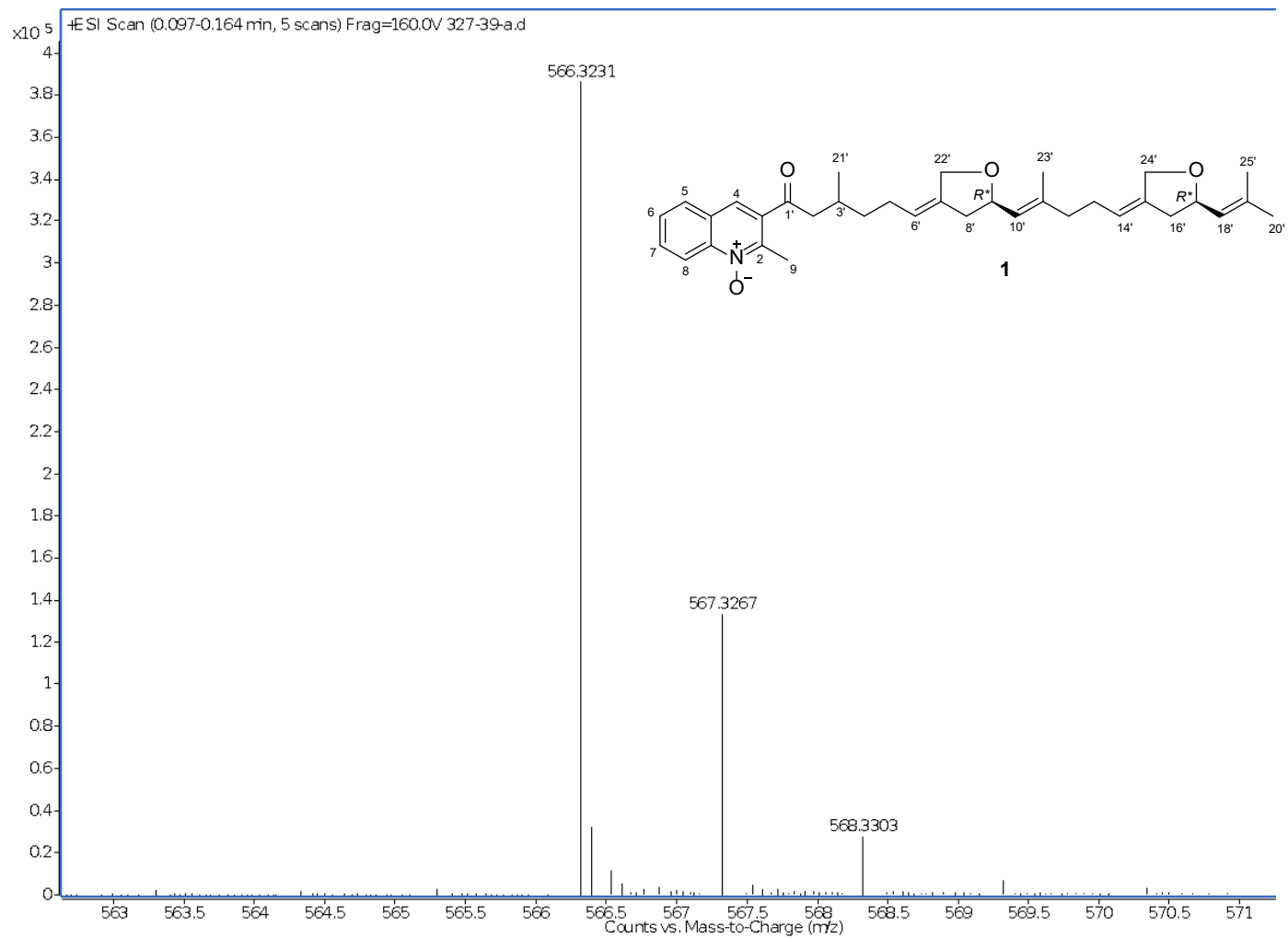


Figure S9. ^1H NMR spectrum of marinoterpin B (**2**) at (500 MHz) in CD_3OD .

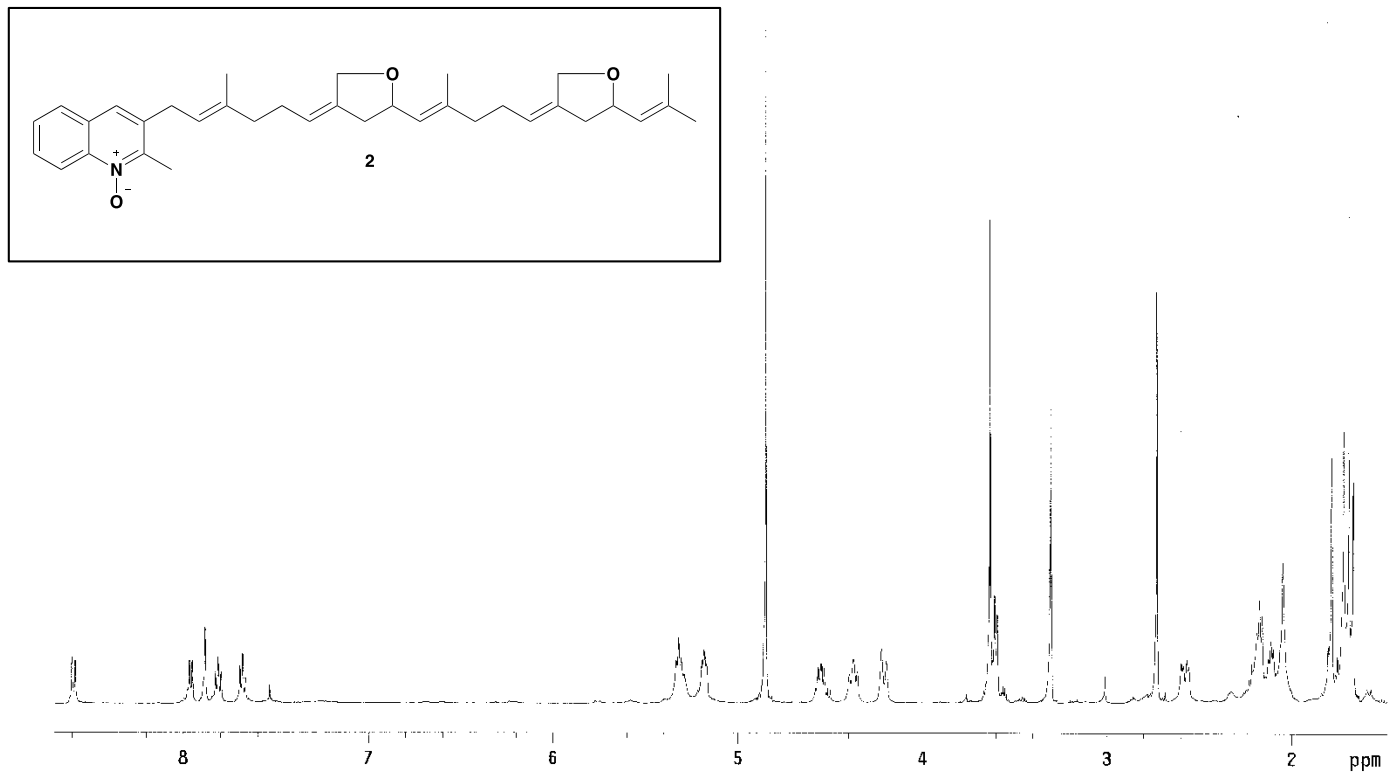


Figure S10. ^{13}C NMR spectrum of marinoterpin B (**2**) (125 MHz) in CD_3OD

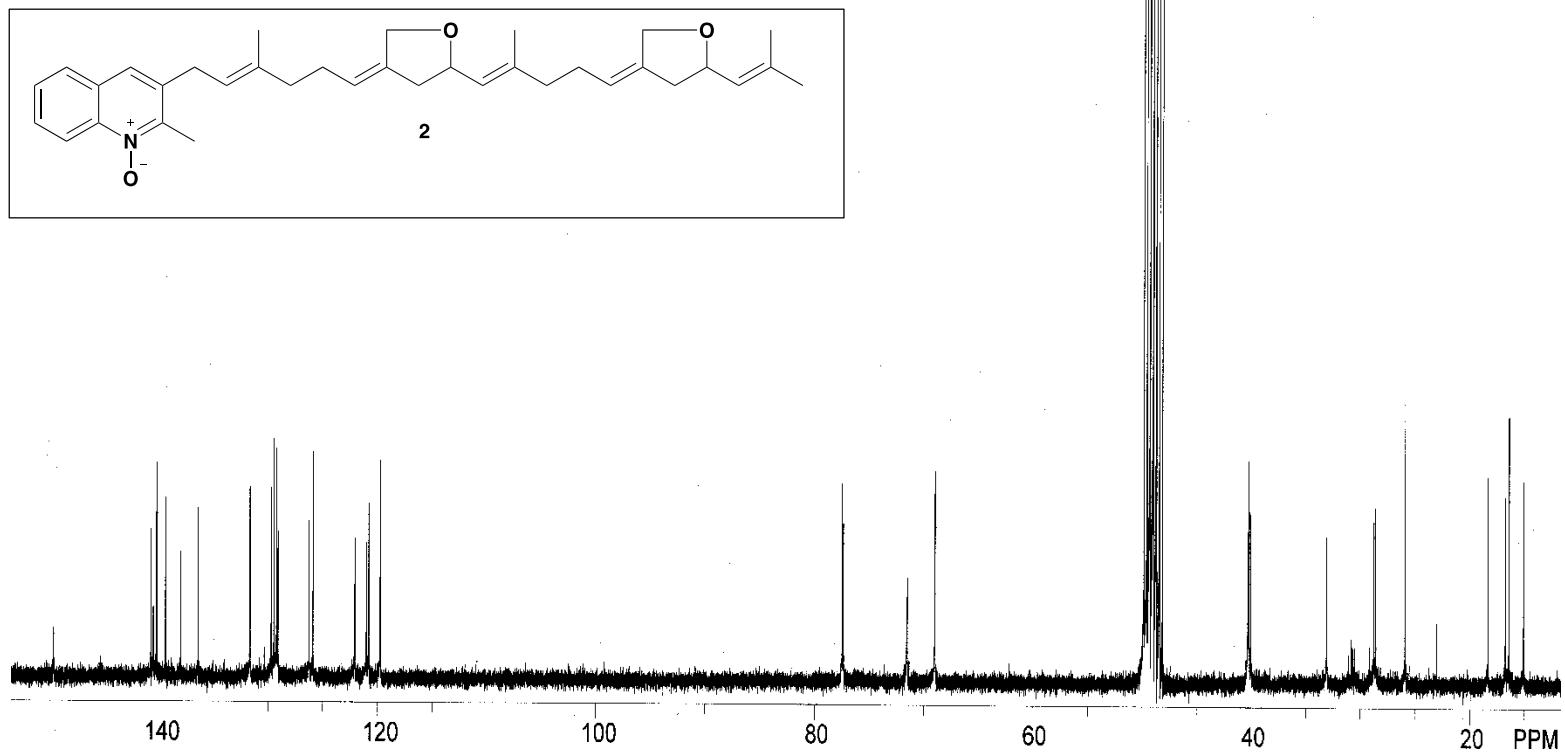


Figure S12. HMQC NMR spectrum of marinoterpin B (**2**) (500 MHz) in CD₃OD.

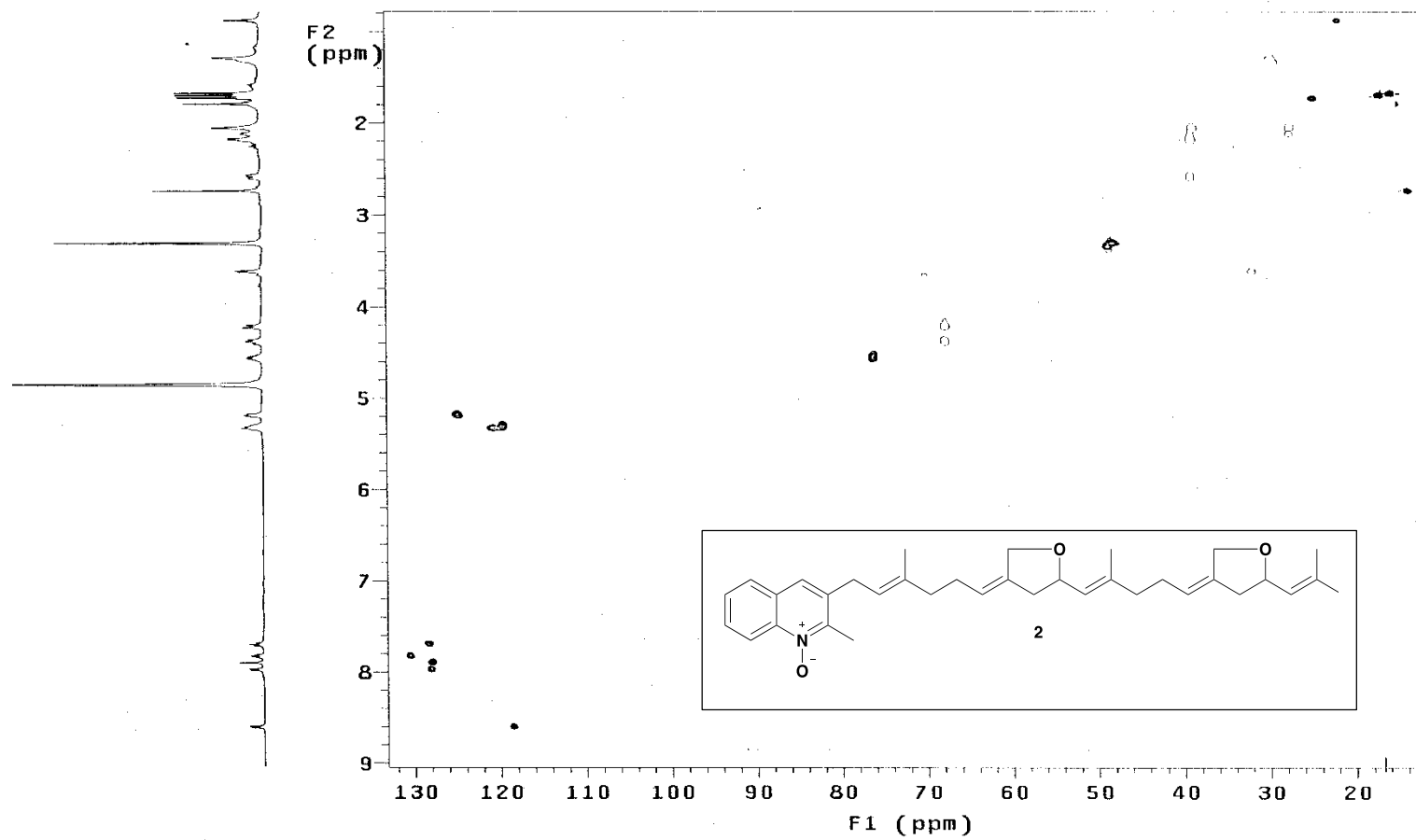


Figure S13. HMBC NMR spectrum of marinoterpin B (**2**) (500 MHz) in CD₃OD.

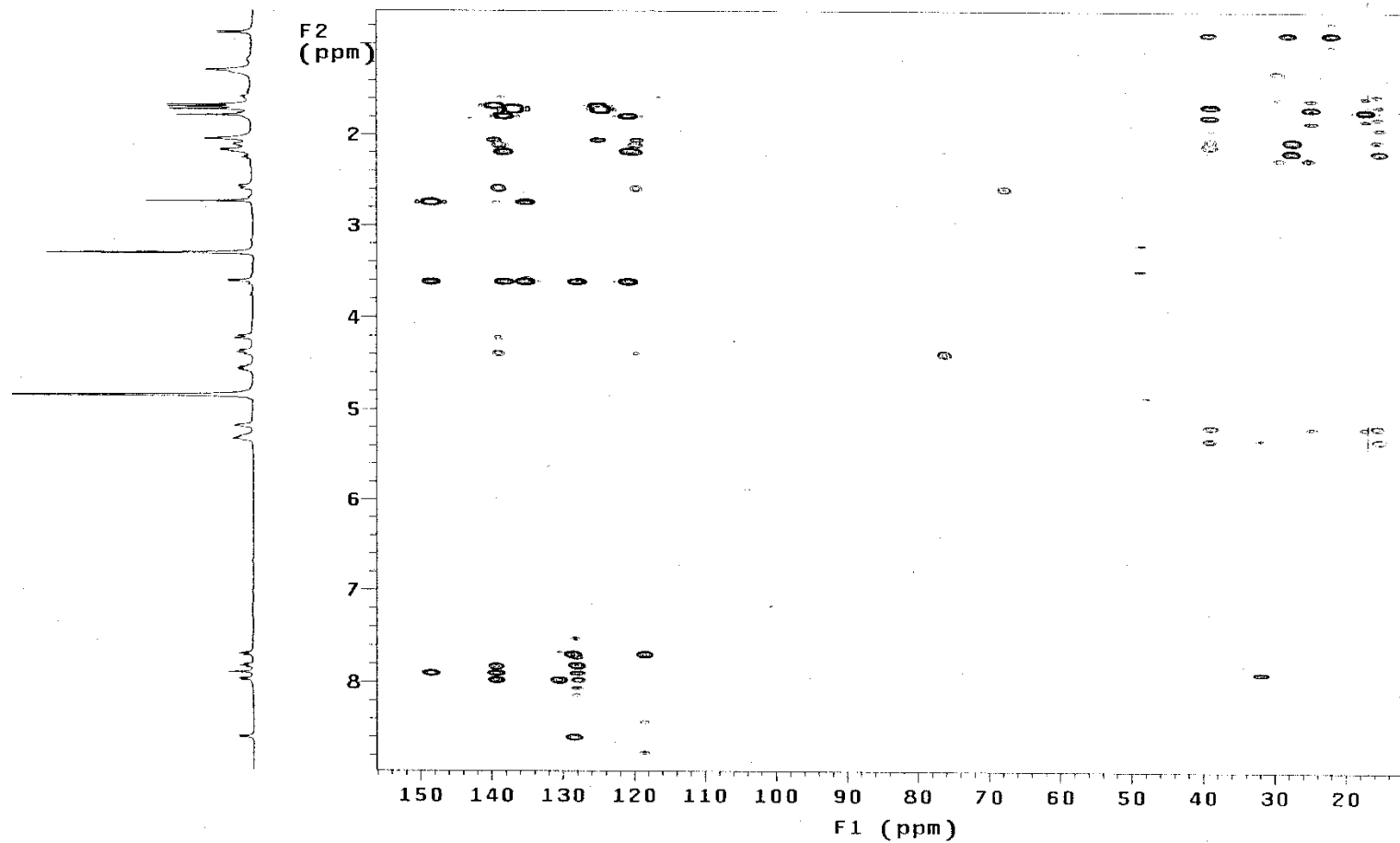
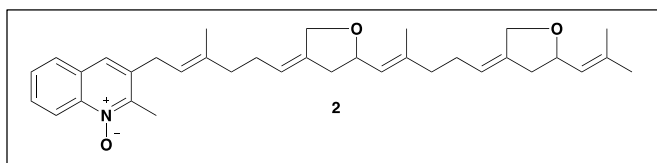


Figure S14. ESI mass spectrum of marinoterpin B (2).

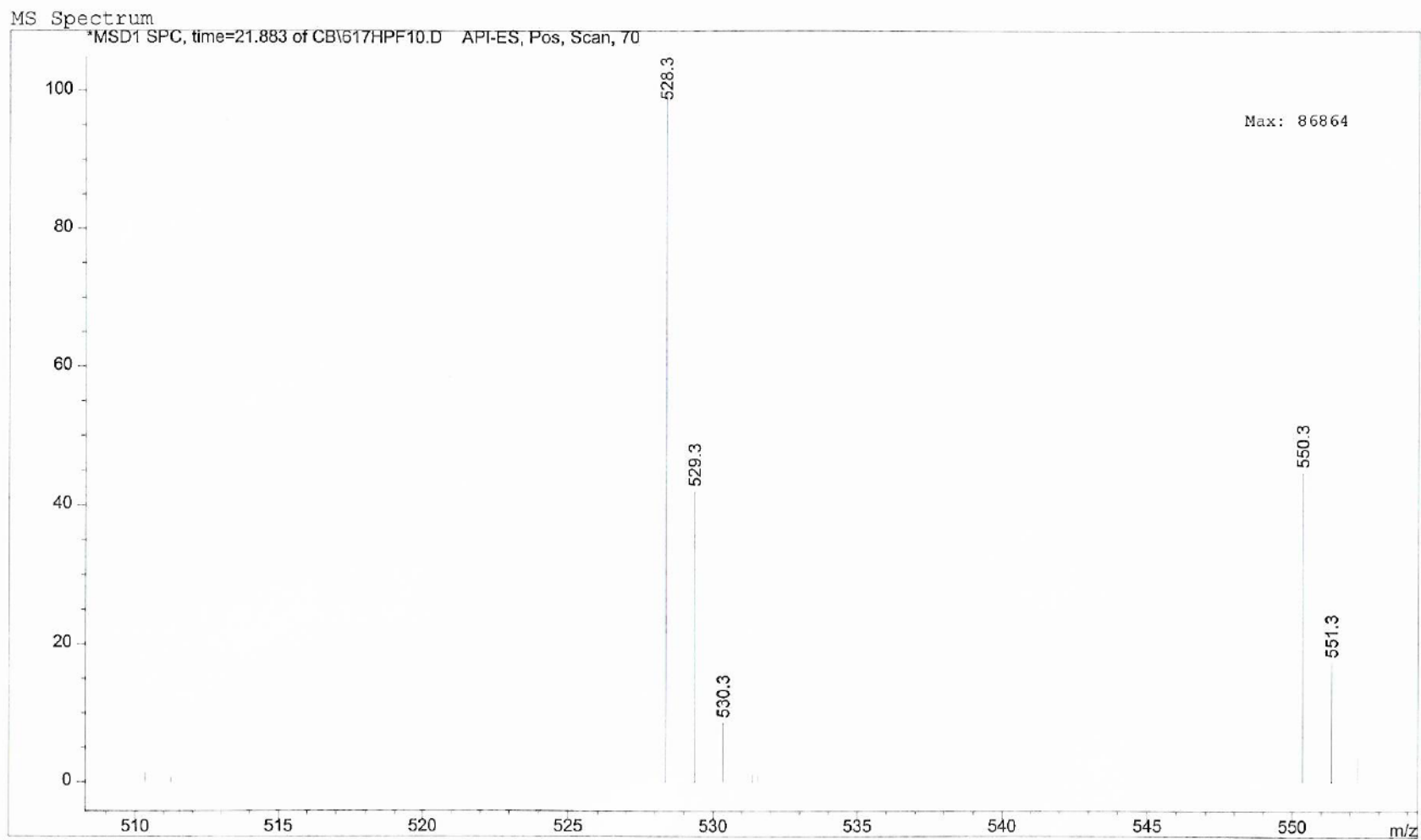


Figure S15. ^1H NMR spectrum of marinoterpin C (**3**) (500 MHz) in CD_3OD .

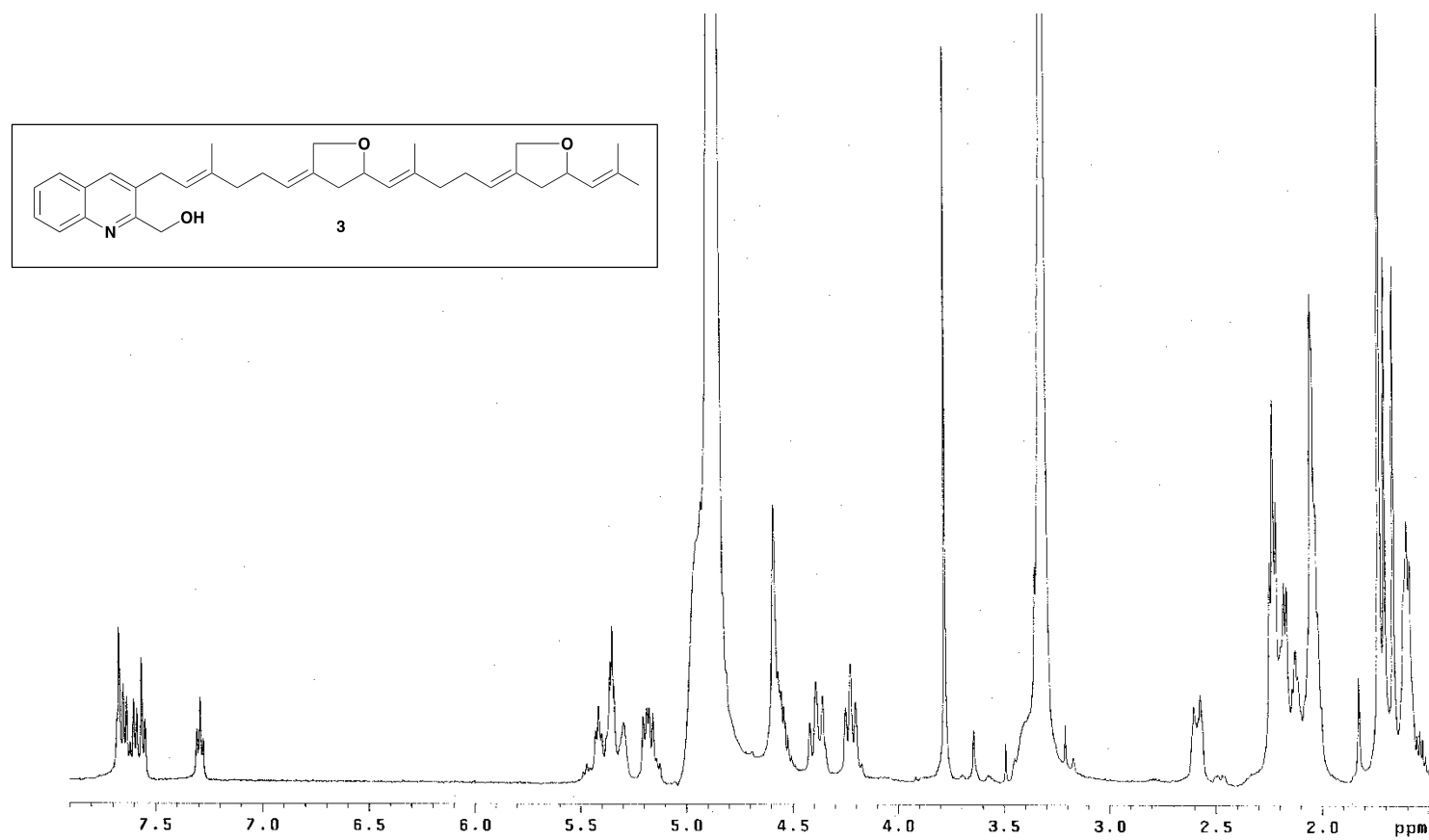


Figure S16. ^{13}C NMR spectrum of marinoterpin C (**3**) (125 MHz) in CD_3OD .

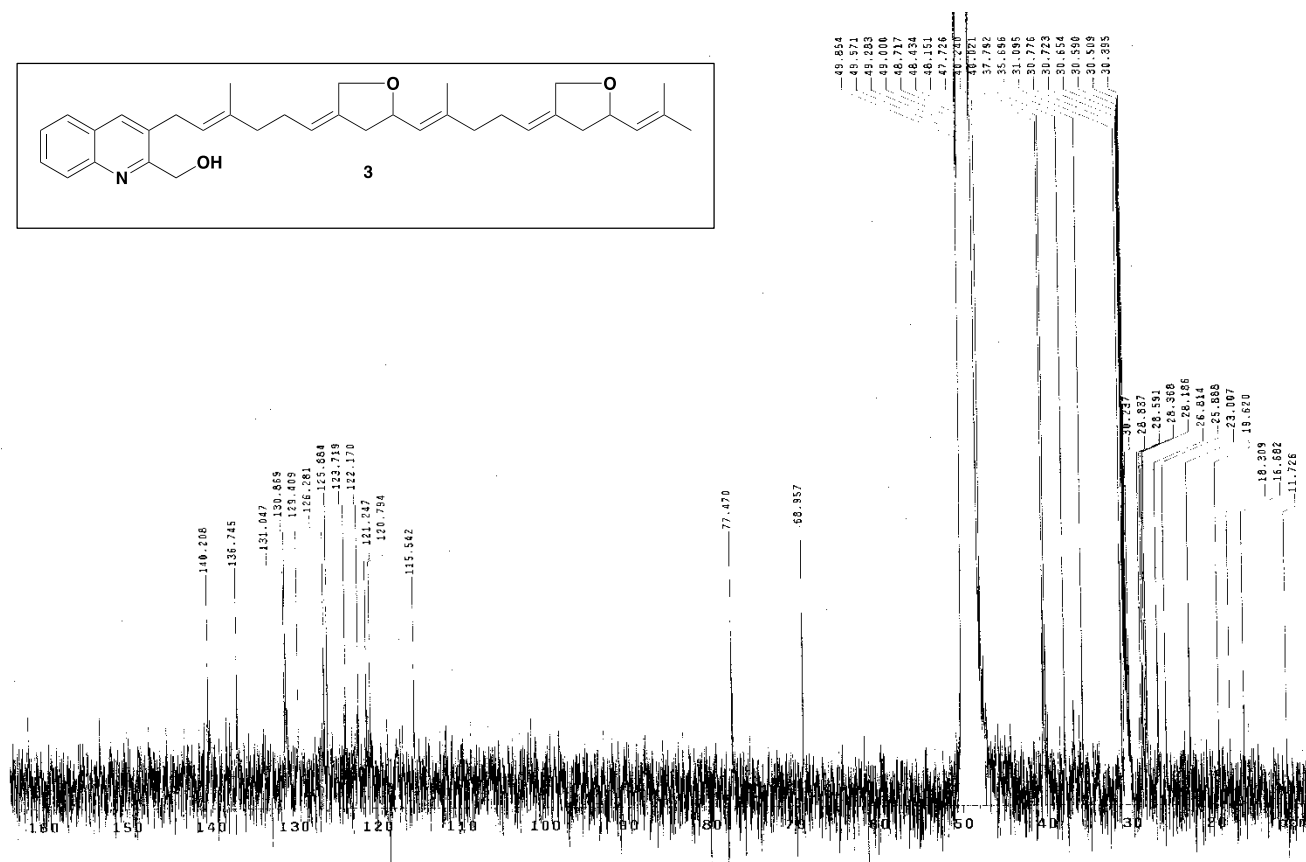


Figure S17. ^1H - ^1H COSY NMR spectrum of marinoterpin C (**3**) (500 MHz) in CD_3OD .

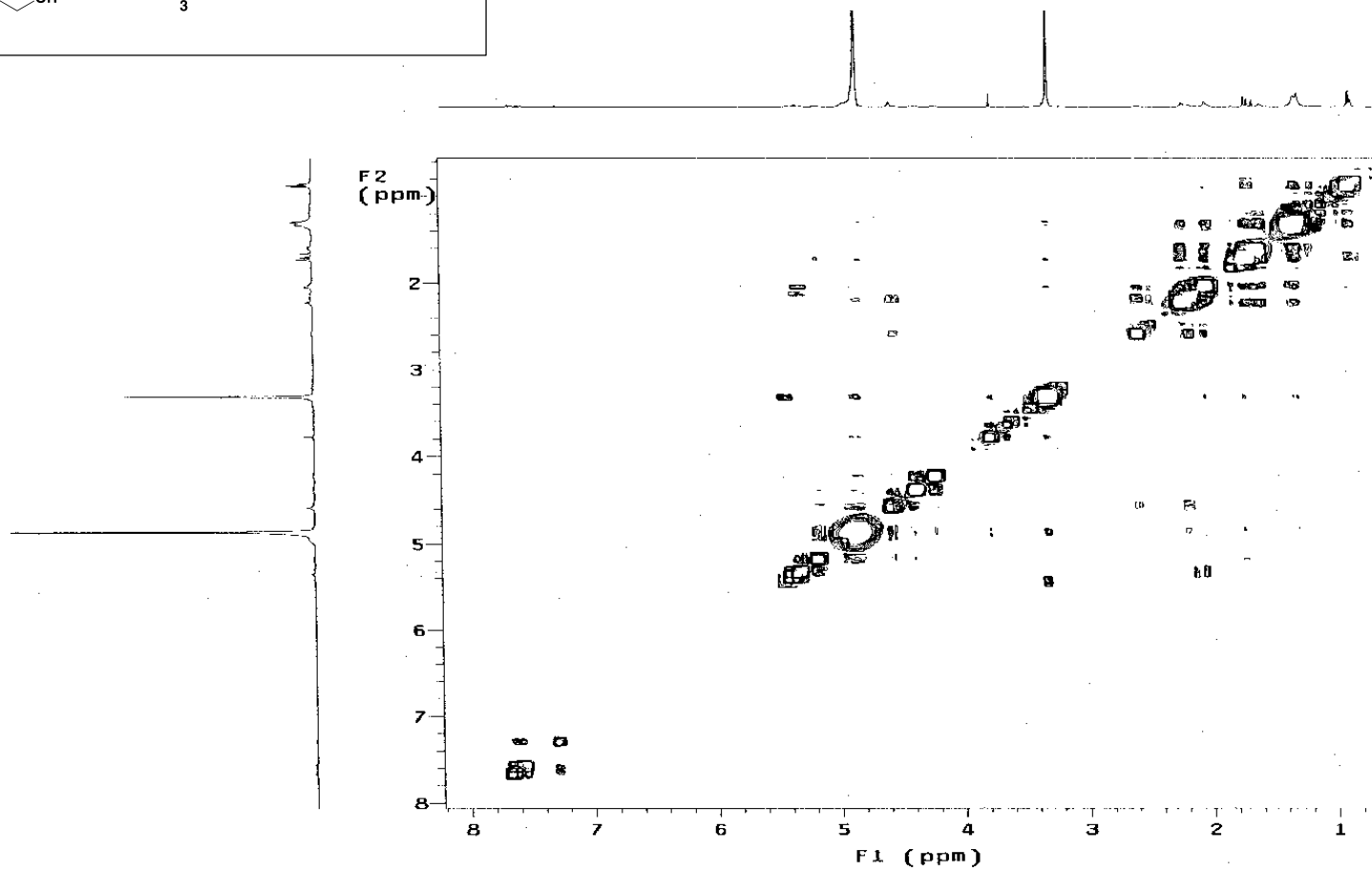
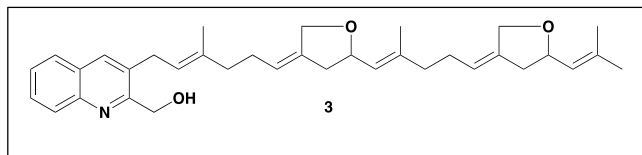


Figure S18. HMQC NMR spectrum of marinoterpin C (**3**) (500 MHz) in CD₃OD.

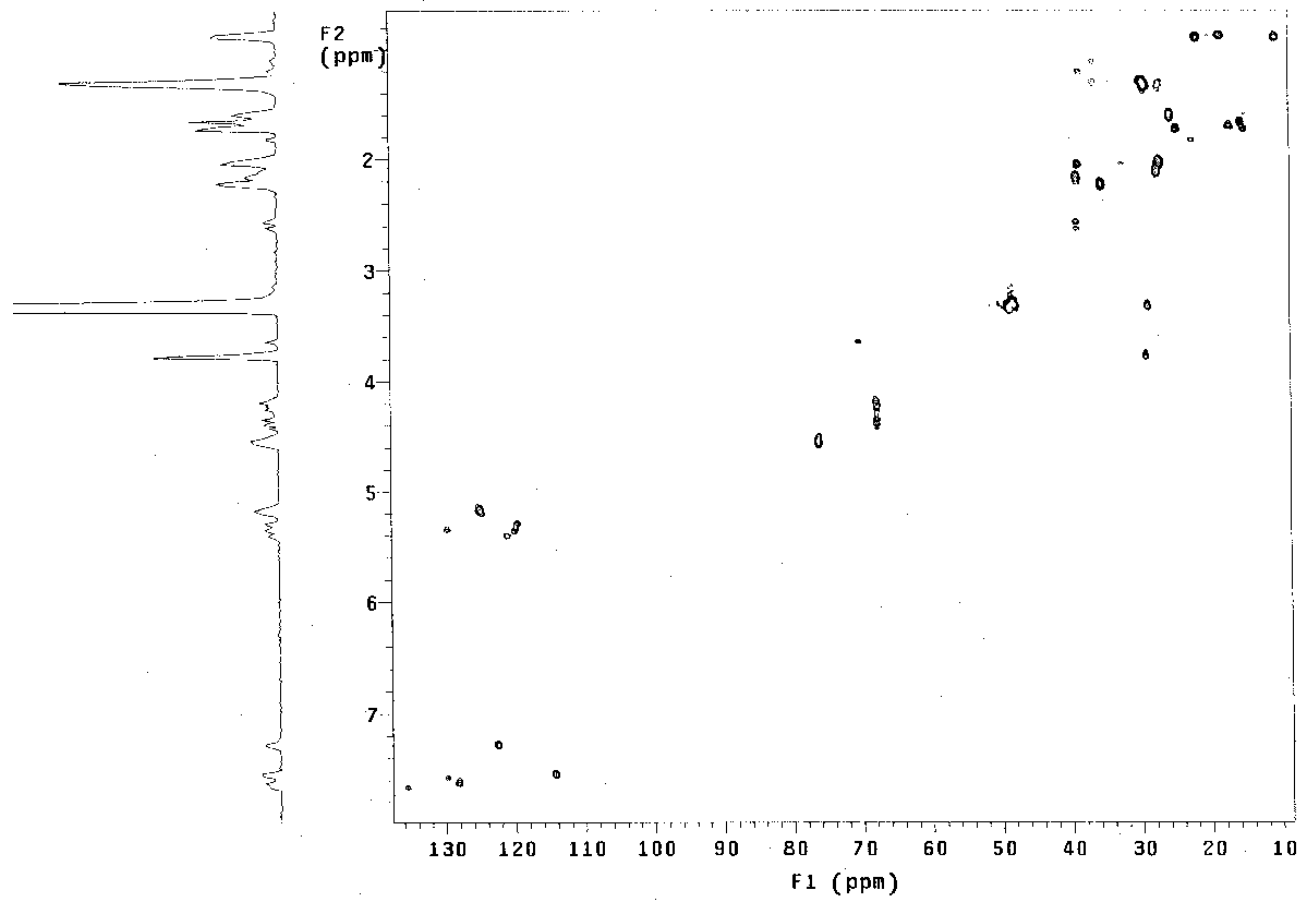
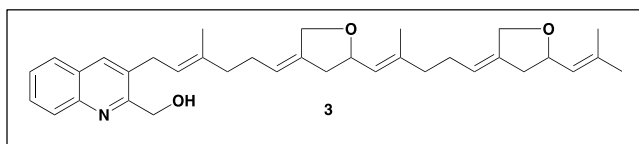


Figure S19. HMBC NMR spectrum of marinoterpin C (**3**) (500 MHz) in CD₃OD.

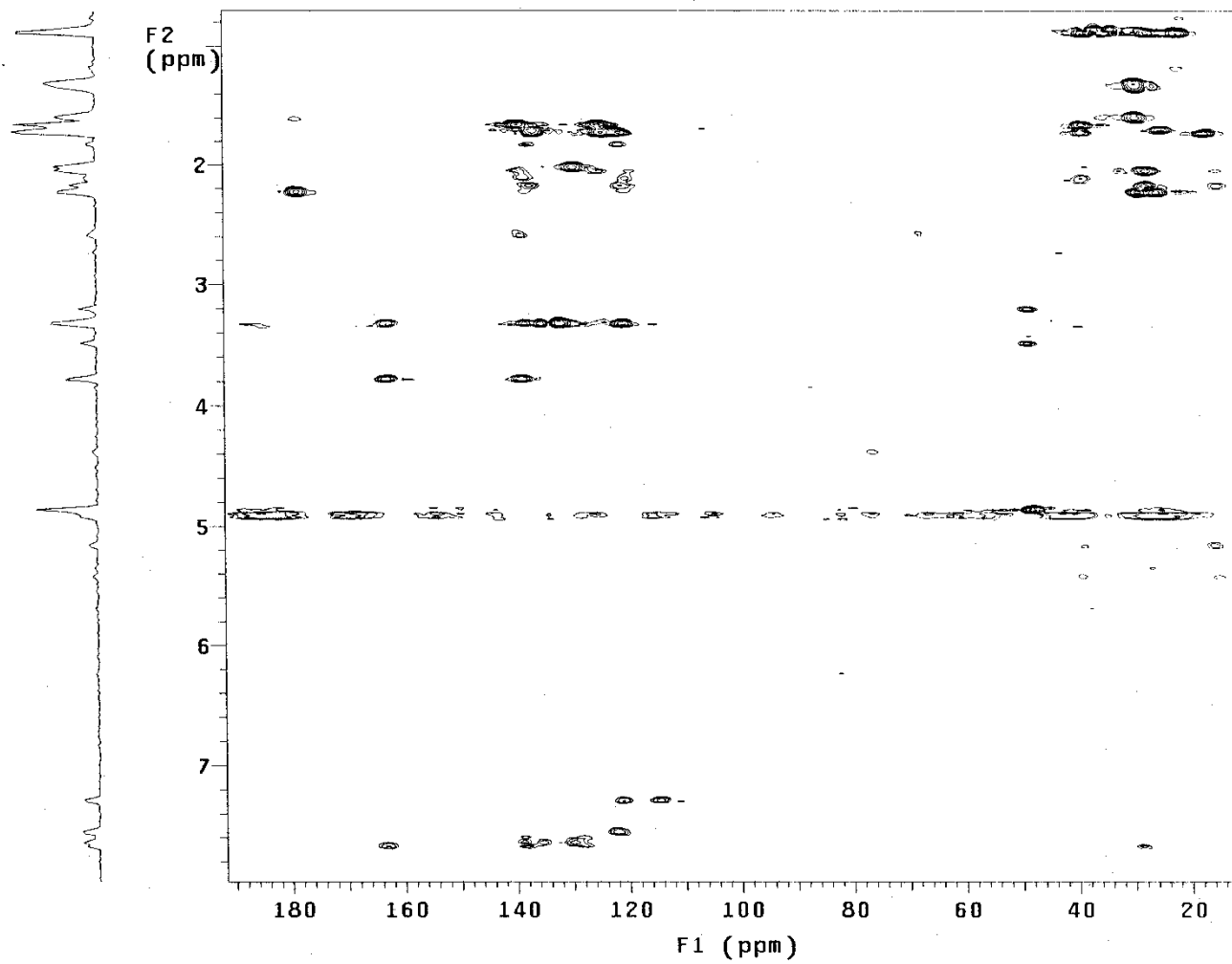
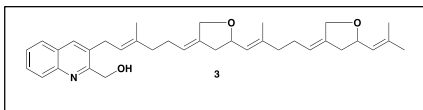


Figure S20. Organization of the *mrt* biosynthetic cluster in *Streptomyces* sp. AJS-327. The arrows represent open reading frames and point in the direction of transcription.

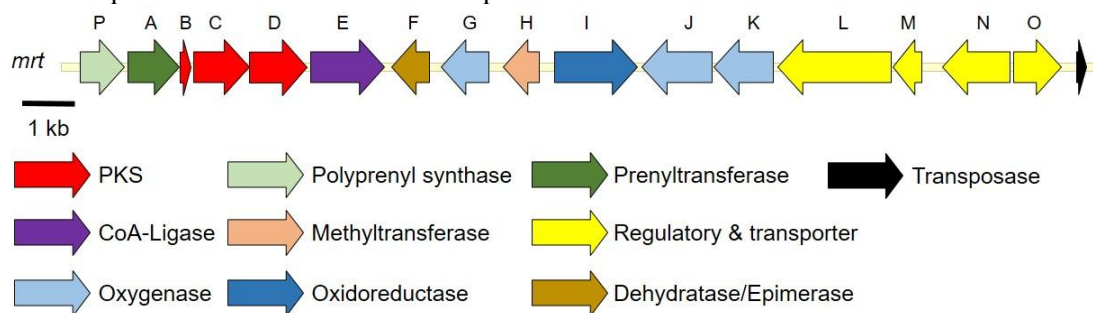
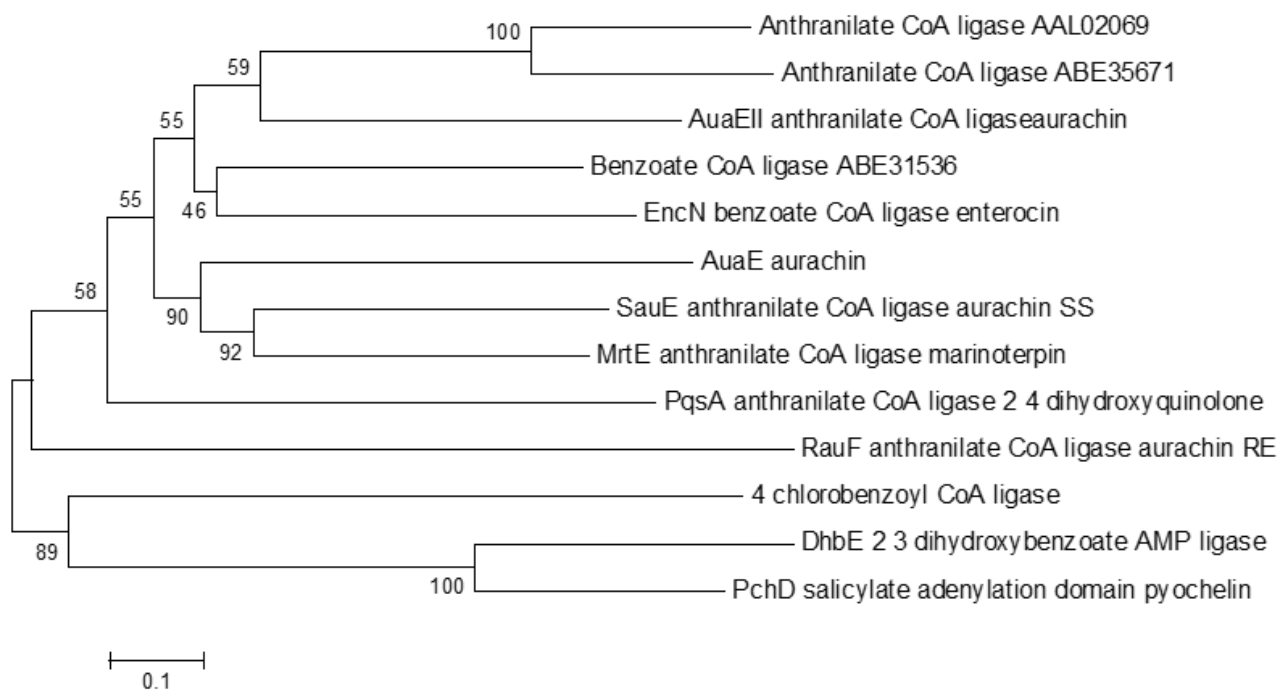


Table S4. Annotated functions of the open reading frames in the *mrt* biosynthetic cluster

Gene Product	NCBI Reference Sequence	Proposed function	Sequence similarity (origin)	Similarity / identity (%)	Accession number
MrtP	WP_180926943	Polyprenyl synthase	<i>Streptomyces</i> sp. CNZ298	68/78	MBB6534444
MrtA	WP_180926944	UbiA family prenyltransferase	<i>Streptomyces</i> sp. CNZ279	72/79	WP_099885763
MrtB	WP_180926945	Acyl carrier protein	Unclassified <i>Streptomyces</i> sp.	71/86	WP_099885764
MrtC	WP_180926946	Beta-ketoacyl-[acyl-carrier-protein] synthase family protein	Unclassified <i>Streptomyces</i> sp.	67/74	WP_099885765
MrtD	WP_180926947	Beta-ketoacyl-CLF synthase II	<i>Streptomyces vitaminophilus</i>	59/69	WP_058032795
MrtE	WP_180926948	Benzoate-CoA ligase	<i>Streptomyces</i> sp. CNZ279	72/79	WP_099885767
MrtF	WP_180926949	NAD-dependent epimerase / dehydratase	<i>Streptomyces vitaminophilus</i>	79/85	WP_078500755
MrtG	WP_180926950	FAD-dependent mono-oxygenase	<i>Streptomyces vitaminophilus</i>	77/87	WP_018385507
MrtH	WP_180926951	Methyltransferase	<i>Streptomyces armeniacus</i>	71/83	AXK34857
MrtI	WP_180926952	FAD-dependent oxidoreductase	Unclassified <i>Streptomyces</i> sp.	75/84	WP_099885772
MrtJ	WP_180928508	Cytochrome P450	<i>Streptomyces vitaminophilus</i>	76/86	WP_018385504
MrtK	WP_180926953	Cytochrome P450	<i>Streptomyces</i> sp. CNZ298	81/87	MBB6534452
MrtL	WP_180926954	FtsX-like permease family protein	<i>Streptomyces</i> sp. CNZ298	79/80	WP_185018974
MrtM	WP_180926955	ABC transporter ATP-binding protein	Unclassified <i>Streptomyces</i> sp.	82/91	WP_099885768
MrtN	WP_180926956	MFS transporter	<i>Streptomyces</i> sp. Tu 3180	71/78	WP_159533361
MrtO	WP_180926957	LysR family transcriptional regulator	<i>Streptomyces radiopugnans</i>	69/77	WP_093654096

Figure S21. Phylogenetic relatedness of select adenylate forming enzymes. MUSCLE¹ was used to create multiple sequence alignments. A) The neighbor-joining tree was constructed using aryl-CoA ligases involved in enterocin biosynthesis, EncN (accession number AAF81733); aurachin SS biosynthesis, SauE; aurachin RE biosynthesis, RauF; marinoterpin biosynthesis, mrtE; anthranilate CoA ligases (accession numbers ABE35671, AAL0206, AAL02069); 2,3-dihydroxybenzoate AMP ligase, DhbE (accession number AFV18074); 4-chlorobenzoyl-CoA ligase (accession number AAN10109); benzoate-CoA ligase (accession number ABE31536); aurachin biosynthesis, AuaE and AuaEII; 2,4-dihydroxyquinolone biosynthesis, PqsA (accession number WP_003140754); pyochelin biosynthesis, PchD adenylation domain (accession number AAD55799). The scale bar indicates 0.1 changes per amino acid. B) Alignment of the adenylate-forming A10 core region of adenylate-forming enzymes.² The lysine residue essential for forming the adenylate intermediate is highlighted in red.

A)



B)

Anthranilate-CoA ligase_AAL02069	PRAIEFRDSLPRTEETG K LQRFRLREGKP-----
Anthranilate-CoA ligase_ABE35671	PREIVFVDLTPRTETG K LKRFELRTIA-----
SauE	PQTVRVVEELPVTSTG K TARHLIRREMEKQS---
RauF	PDICVSVPRIPLTSTG K IDRGACHRLLARRVGSRL
DhbE	PDRIEFIESFPQTGVG K VSKKELRKVIAEKLITVK
4-chlorobenzoyl-CoA ligase	PKRYFILDQLPKNALN K VLRRLQLVQQVSS-----
Benzoate-CoA ligase_ABE31536	PRDIVFVDDLPKTATG K IQRFKLREQS-----
AuaE	PSKIEVLSGSAS-----
AuaEII	PRWIEFVTELPKTATG K IQRFKLRSAA-----
EncN	PHLVEFAADLPKTP K IQRFALRSQETGKADAPA
PqsA	PSQLHVLPALPRNDNG K LARAELRHLADTLYHDNL
PchD	PDRIELVPAFPQTGIG K ISKKDLRERLRRELEARA
MrtE	PRVLRVDDLPRATATG K IARHTLRAEAVAGAAAPT

Figure S22. Phylogenetic relatedness of select polyprenyl synthases. MUSCLE¹ was used to create multiple sequence alignments. The neighbor-joining tree was constructed using the farnesyl diphosphate synthase involved in aurachin RE biosynthesis, RaulI (accession number WP_052416887); geranyl/farnesyl diphosphate synthase involved in aurachin SS biosynthesis, SauI (accession number ARM20258); geranyl diphosphate synthase involved in napyradiomycin biosynthesis, NapT7 (accession number ABS50482); geranyl diphosphate synthase involved in furaquinocin biosynthesis (accession number); geranyl diphosphate synthase involved in furaquinocin biosynthesis, Fur19 (accession number BAE78987); geranyl diphosphate synthase involved in furanonaphthoquinone biosynthesis, Fnq23 (accession number CAL34101); Mcl2 and Mcl22 involved in merochlorin biosynthesis (accession numbers AGH68887 and WP_024888525); sesterterpene diphosphate synthase involved in marinoterpin biosynthesis, MrtP (accession number WP_180926943); sesterterpene diphosphate synthase involved in atolypene biosynthesis, AtoC (Accession number WP_063745200); sesterterpene diphosphate synthase involved in sestermobaraene biosynthesis(accession number WP_004941318); StsA involved in somaliensene A and B biosynthesis (accession number WP_078571316); StsB involved in somaliensene A and B biosynthesis (accession number WP_010468026);farnesyl diphosphate synthase involved in BE-40644 biosynthesis, FdpS (accession number BAD07374); solanesyl diphosphate synthase (accession number XP_002879508); undecaprenyl diphosphate synthases (accession numbers WP_003640735 and O82827); geranylgeranyl diphosphate synthases (accession numbers NP_001032354, XP_003308874, XP_003824547, XP_004028673, XP_009249955, CAD6651761, and XP_003267333); farnesyl diphosphate synthases (accession numbers NP_001129293, WP_013069157, WP_000183378, and XP_001160391); and decaprenyl diphosphate synthase (accession number WP_013068881). The scale bar indicates 0.2 changes/amino acid.

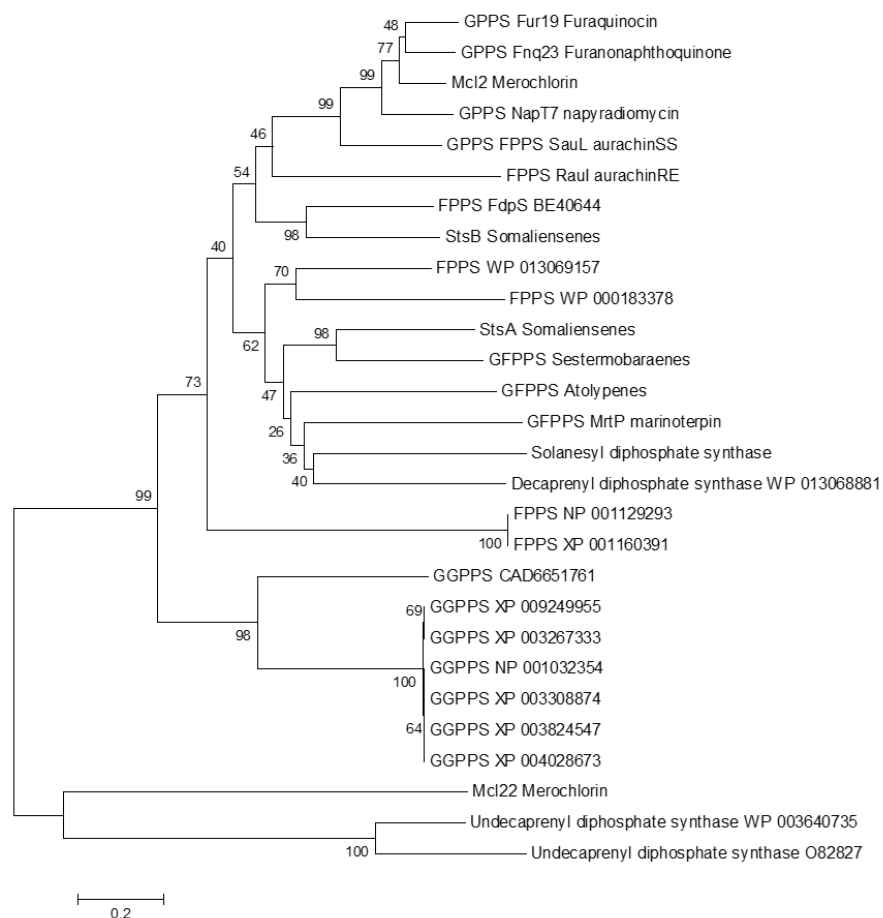


Figure S23. Phylogenetic relatedness of selected prenyltransferases. MUSCLE¹ was used to create multiple sequence alignments. The neighbor-joining tree was constructed using prenyltransferases involved in aurachin RE biosynthesis, RauB (accession number WP_080726500); aurachin biosynthesis, AuaA (accession number CAL48953); aurachin SS biosynthesis, SauA (accession number WP_176101155); marinoterpin biosynthesis, MrtA (accession number WP_180926944); napyradiomycin biosynthesis, NapT8 and NapT9 (accession numbers ABS50489 and ABS50490); 4-hydroxybenzoate octaprenyltransferase, UbiA (accession number WP_160799794); BE-40644 biosynthesis, (accession number BAD07390); furaquinocin biosynthesis, Fur7 (accession number BAE78988); furanonaphthoquinone biosynthesis, Fnq26 and Fnq (accession numbers A2AXG5 and CAL34106); merochlorin biosynthesis, Mcl23 (accession number AGH68908); and novobiocin biosynthesis, NovQ (accession number WP_079127921). The scale bar indicates 0.2 changes per amino acid.

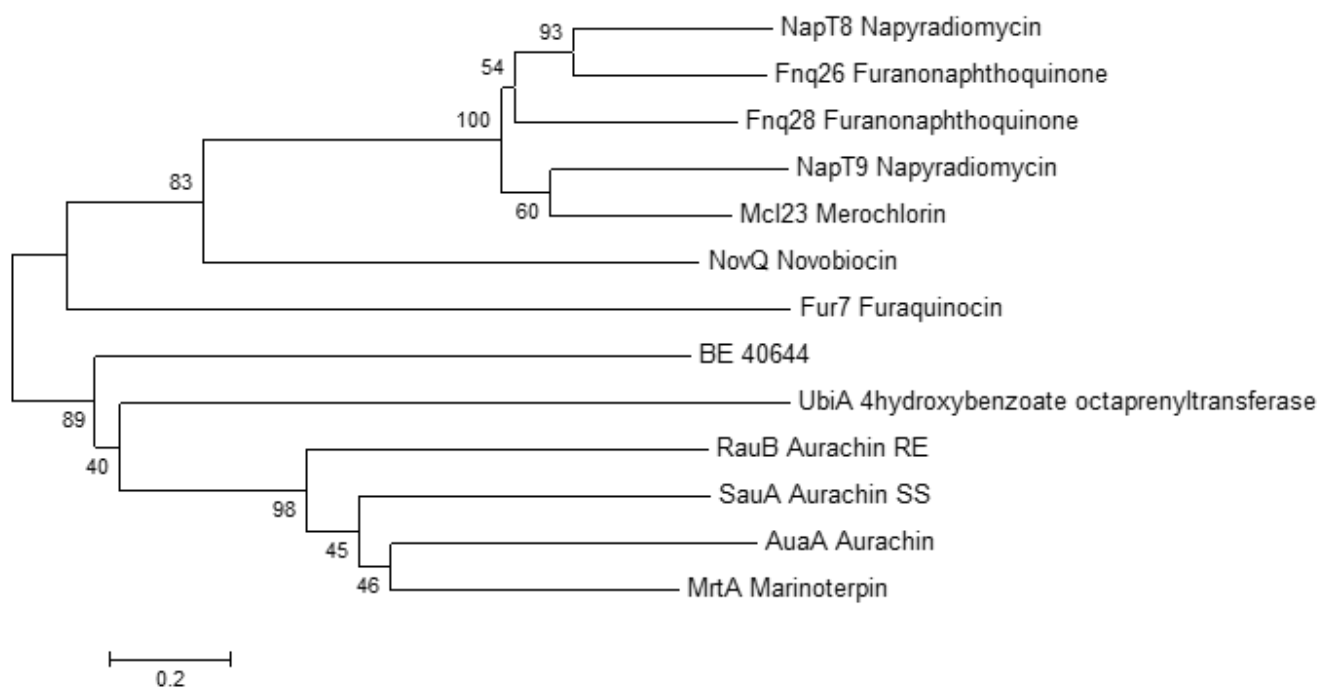
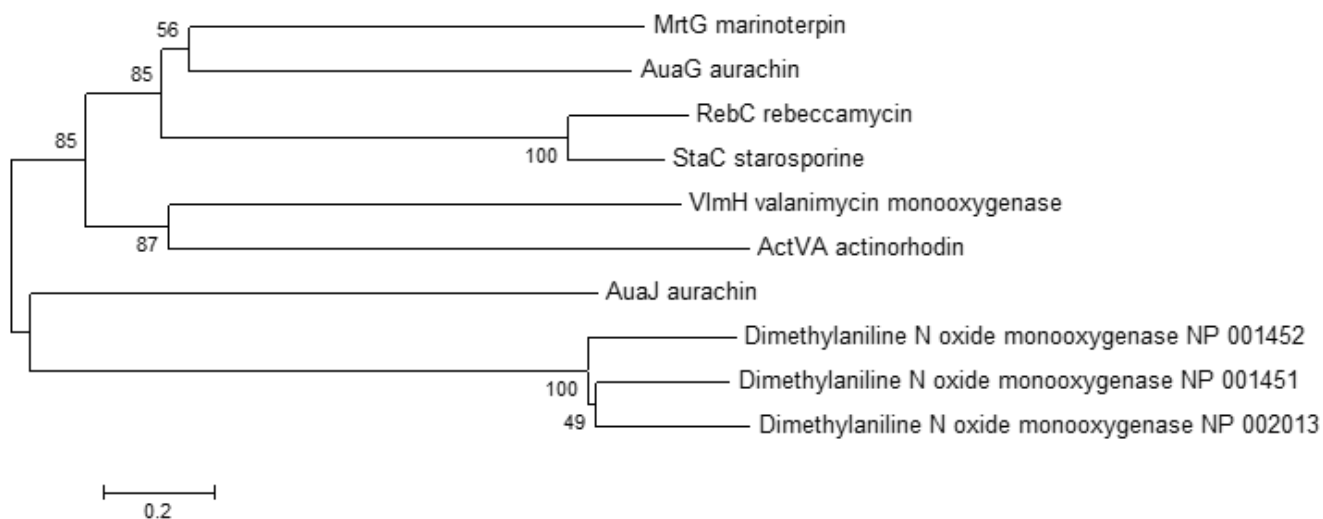


Figure S24. Phylogenetic relatedness of select flavin monooxygenases. MUSCLE¹ was used to create multiple sequence alignments. The neighbor-joining tree was constructed using flavin monooxygenases involved in valanimycin biosynthesis, VImH (accession number P96072); aurachin biosynthesis, AuaJ and AuaG (accession numbers CCD27752 and H1ZZA4, respectively); marinoterpin biosynthesis, MrtG (accession number WP_0180926950); actinorhodin biosynthesis, ActVA (accession number WP_011030043); rebeccamycin biosynthesis, RebC (PDB 2R0C_A); starosporine biosynthesis, StaC (accession number ABI94390); and dimethylaniline *N*-oxide monooxygenases (accession numbers NP_001452, NP_001451, and NP_002013). The scale bar indicates 0.2 changes per amino acid.



References

1. Edgar, R. C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC bioinformatics*. 2004, 5, 113.
2. Pistorius, D., Li, Y., Mann, S., Müller, R. Unprecedented anthranilate priming involving two enzymes of the acyl adenylating superfamily in aurachin biosynthesis. *J Am Chem Soc*. 2011, 133, 12362.