Supplementary Information

Genome-driven discovery of new serrawettin W2 analogues from *Serratia fonticola* DSM 4576

Haolin Qiu [a], Yang Xiao [a], Ling Shen [a], Tao Han [a], Qiang He [c], Aiying Li [d], Peng Zhang [a] and Xiaofeng Cai [a,b]

[a]*School of Pharmacy, Tongji Medical College of Huazhong University of Science and Technology, Hubei Key Laboratory of Natural Medicinal Chemistry and Resource Evaluation, Wuhan 430030, People's Republic of China.*

[b]*State Key Laboratory of Dao-di Herbs, Beijing 100700, P. R. China.*

[c]*Xianning Public Inspection Center of Hubei Province, Xianning 437000, P. R. China.*

[d]*Helmholtz International Lab for Anti-Infectives, Shandong University-Helmholtz Institute of Biotechnology, State Key Laboratory of Microbial Technology, Shandong University, Qingdao 266237, P.R. China.*

[*]Corresponding authors.

E-mail: caixiaofeng@hust.edu.cn

Contents

Table S1 A domain specificity prediction of *sefA*.

A domain	Stachelhaus sequence	most likely amino acid predicted	amino acid detected
A1	DASTVAAVCK	Tyr	Tyr
A2	DAYFLGVTYK	Val	Ile
A3	DALFIGCVFK	Leu	Leu
A4	DALFVGGVWK	Val	Phe
A5	DVWHFSLVDK	Ser	Ser

Table S2 HRESIMS data of natural and synthetic compounds.

compound	sum formula	calcd. $[M + H]^+$	found $[M + H]^+$	Δppm
sefopeptide A (**1**)	$C_{47}H_{71}N_5O_9$	850.5325	850.5326	0.17
sefopeptide B (**2**)	$C_{45}H_{67}N_5O_9$	822.5012	822.5022	1.27
sefopeptide C (**3**)	$C_{48}H_{71}N_5O_9$	862.5325	862.5335	1.21
sefopeptide D (**4**)	$C_{49}H_{73}N_5O_9$	876.5481	876.5470	1.26

Table S3 Bacterial strains used in this study.

strain	genotype
E.coli DH10B	F-*mcr*A, Δ(*mrr-hsd*RMS-*mcr*BC), Φ80*lac*ZΔM15, Δ*lac*X74, *end*A1, *rec*A1, *deo*R, Δ(*ara leu*)7697, *ara*D139, *gal*U, *gal*K, *rps*L, *nup*G, λ–
E.coli BAP1	BL21(DE3) *Δprp*RBCD::T7prom-*sfp*, T7prom-*prpE*[1]
Serratia fonticola DSM 4576	wild type

Table S4 Primers used in this study.

primer	sequence (5'-3')	targeting DNA fragment	plasmid
Duet-F	TGCTTAAGTCGAACAGAAA	pCOLA-Duet vector backbone (3629 bp)	pHL1
Duet-R	GGTATATCTCCTTATTAAAG		
sef1-F	TTTTGTTTAACTTTAATAAGGAGAT ATACCATGGCGAATAATATGGAAA GAAT	fragment Ⅰ of *sefA* from *Serratia fonticola* DSM4576 (9678 bp)	
sef1-R	GGCGACAGAATGATGCTCAA		
sef2-F	CTAAAGGGGTGATGATTGAGC	fragment Ⅱ of *sefA* from *Serratia fonticola* DSM4576 (9782 bp)	
sef2-R	AATACGATTACTTTCTGTTCGACTTA AGCATTAGCGGATGTGGCTAATCA		

Table S5 Plasmids used in this study.

plasmid	genotype/description
pCOLA-Duet	3,719 bp, contains T7 promoter, Km^R
pHL1	22,994bp, *sefA* from *Serratia fonticola* DSM 4576 genomic DNA assembled into pCOLA-Duet, Km^R

Table S6 ^{1}H (400 MHz) and ^{13}C (100 MHz) NMR data for **1** in DMSO-d_6 (δ in ppm).

Subunit	Position	δ_C, type	δ_H (J [Hz])	Subunit	Position	δ_C, type	δ_H (J [Hz])
(R)-3-	1	168.68, C			28	35.28, CH	1.52, m
Hydroxy-	2a	39.41, CH_2	2.22, dd (14.1, 3.0)		29	24.69, CH_2	0.89, m
tetradecanoic	2b		2.44, overlap		30	10.32, CH_3	0.70, overlap
Acid	3	71.64, CH	4.94, m		31	14.83, CH_3	0.70, overlap
	4	33.07, CH_2	1.46, m		32	172.87, C	
	5	24.14, CH_2	1.23, m	Leu	33		8.68, d (5.7)
	6	29.13, CH_2	1.23, m		34	52.42, CH	3.88, overlap
	7	29.04, CH_2	1.23, m		35	39.08, CH_2	1.26, m
	8	29.02, CH_2	1.23, m		36	23.71, CH	1.23, m
	9	28.93, CH_2	1.23, m		37	21.45, CH_3	0.70, overlap
	10	28.83, CH_2	1.23, m		38	22.66, CH_3	0.78, overlap
	11	28.73, CH_2	1.23, m		39	171.65, C	
	12	31.29, CH_2	1.20, m	Phe	40		8.11, d (5.4)
	13	22.08, CH_2	1.22, m		41	54.10, CH	4.32, m
	14	13.94, CH_3	0.82, overlap		42a	35.35, CH_2	2.84, t (12.7)
Tyr	15		7.96, overlap		42b		3.22, dd (13.9, 3,9)
	16	53.23, CH	4.62, q (7.5)		43	138.64, C	
	17a	37.55, CH_2	2.64, dd (13.5, 7.3)		44	129.18, CH	7.17, m
	17b		2.75, dd (13.5, 7.1)		45	127.94, CH	7.24, m
	18	127.44, C			46	126.07, CH	7.18, m
	19	130.03, CH	6.92, dd (8.7, 2.6)		47	127.94, CH	7.24, m
	20	114.74, CH	6.58, dd (8.7, 2.6)		48	129.18, CH	7.17, m
	21	155.73, C			49	170.97, C	
	22	114.74, CH	6.58, dd (8.7, 2.6)	Ser	50		7.44, s
	23	130.03, CH	6.92, dd (8.7, 2.6)		51	56.27, CH	4.06, q (6.0)
	24		9.12, s		52a	61.18, CH_2	3.62, m
	25	171.03, C			52b		3.67, m
Ile	26		7.99, overlap		53		4.82, t (6.7)
	27	56.94, CH	3.89, overlap		54	169.56, C	

Table S7 ^{1}H (400 MHz) and ^{13}C (100 MHz) NMR data for **2** in DMSO-d_6 (δ in ppm).

Subunit	Position	δ_C, type	δ_H (J [Hz])	Subunit	Position	δ_C, type	δ_H (J [Hz])
(*R*)-3-Hydroxy-dodecanoic acid	1	168.68, C			27	24.69, CH_2	0.89, m
	2a	39.41, CH_2	2.22, dd (14.1, 3.0)		28	10.32, CH_3	0.70, overlap
	2b		2.44, overlap		29	14.83, CH_3	0.70, overlap
	3	71.64, CH	4.94, m		30	172.87, C	
	4	33.07, CH_2	1.46, m	Leu	31		8.68, d (5.7)
	5	24.14, CH_2	1.23, m		32	52.42, CH	3.88, overlap
	6	29.05, CH_2	1.23, m		33	39.08, CH_2	1.26, m
	7	28.90, CH_2	1.23, m		34	23.71, CH	1.23, m
	8	28.83, CH_2	1.23, m		35	21.45, CH_3	0.70, overlap
	9	28.78, CH_2	1.23, m		36	22.66, CH_3	0.78, overlap
	10	31.29, CH_2	1.20, m		37	171.65, C	
	11	22.08, CH_2	1.22, m	Phe	38		8.11, d (5.4)
	12	13.94, CH_3	0.82, overlap		39	54.10, CH	4.32, m
Tyr	13		7.96, overlap		40a	35.35, CH_2	2.84, t (12.7)
	14	53.23, CH	4.62, q (7.5)		40b		3.22, dd (13.9, 3,9)
	15a	37.55, CH_2	2.64, dd (13.5, 7.3)		41	138.64, C	
	15b		2.75, dd (13.5, 7.1)		42	129.18, CH	7.17, m
	16	127.44, C			43	127.94, CH	7.24, m
	17	130.03, CH	6.92, dd (8.7, 2.6)		44	126.07, CH	7.18, m
	18	114.74, CH	6.58, dd (8.7, 2.6)		45	127.94, CH	7.24, m
	19	155.73, C			46	129.18, CH	7.17, m
	20	114.74, CH	6.58, dd (8.7, 2.6)		47	170.97, C	
	21	130.03, CH	6.92, dd (8.7, 2.6)	Ser	48		7.44, s
	22		9.12, s		49	56.27, CH	4.06, q (6.0)
	23	171.03, C			50a	61.18, CH_2	3.62, m
Ile	24		7.99, overlap		50b		3.67, m
	25	56.94, CH	3.89, overlap		51		4.82, t (6.7)
	26	35.28, CH	1.52, m		52	169.56, C	

Table S8 ^{1}H (400 MHz) and ^{13}C (100 MHz) NMR data for **3** in DMSO-d_6 (δ in ppm).

Subunit	Position	δ_C, type	δ_H (J [Hz])	Subunit	Position	δ_C, type	δ_H (J [Hz])
(3*R*,8*Z*)-3-Hydroxy-8-pentadecenoic acid	1	168.68, C			29	35.28, CH	1.52, m
	2a	39.41, CH_2	2.22, dd (14.1, 3.0)		30	24.69, CH_2	0.89, m
	2b		2.44, overlap		31	10.32, CH_3	0.70, overlap
	3	71.63, CH	4.94, m		32	14.83, CH_3	0.70, overlap
	4	33.08, CH_2	1.46, m		33	172.87, C	
	5	24.07, CH_2	1.24, m	Leu	34		8.68, d (5.7)
	6	29.17, CH_2	1.24, m		35	52.42, CH	3.88, overlap
	7	26.57, CH_2	1.99, p (6.0)		36	39.08, CH_2	1.26, m
	8	129.65, CH	5,34, overlap		37	23.71, CH	1.23, m
	9	129.63, CH	5,34, overlap		38	21.45, CH_3	0.70, overlap
	10	26.57, CH_2	1.99, p (6.0)		39	22.66, CH_3	0.78, overlap
	11	28.83, CH_2	1.30, m		40	171.65, C	
	12	28.49, CH_2	1.24, m	Phe	41		8.11, d (5.4)
	13	30.86, CH_2	1.23, m		42	54.10, CH	4.32, m
	14	21.98, CH_2	1.24, m		43a	35.35, CH_2	2.84, t (12.7)
	15	13.93, CH_3	0.82, overlap		43b		3.22, dd (13.9, 3,9)
Tyr	16		7.96, overlap		44	138.64, C	
	17	53.23, CH	4.62, q (7.5)		45	129.18, CH	7.17, m
	18a	37.55, CH_2	2.64, dd (13.5, 7.3)		46	127.94, CH	7.24, m
	18b		2.75, dd (13.5, 7.1)		47	126.07, CH	7.18, m
	19	127.44, C			48	127.94, CH	7.24, m
	20	130.03, CH	6.92, dd (8.7, 2.6)		49	129.18, CH	7.17, m
	21	114.74, CH	6.58, dd (8.7, 2.6)		50	170.97, C	
	22	155.73, C		Ser	51		7.44, s
	23	114.74, CH	6.58, dd (8.7, 2.6)		52	56.27, CH	4.06, q (6.0)
	24	130.03, CH	6.92, dd (8.7, 2.6)		53a	61.18, CH_2	3.62, m
	25		9.12, s		53b		3.67, m
	26	171.03, C			54		4.82, t (6.7)
Ile	27		7.99, overlap		55	169.56, C	
	28	56.94, CH	3.89, overlap				

Table S9 ^{1}H (400 MHz) and ^{13}C (100 MHz) NMR data for **4** in DMSO-d_6 (δ in ppm).

Subunit	Position	δ_C, type	δ_H (J [Hz])	Subunit	Position	δ_C, type	δ_H (J [Hz])
(3*R*,9*Z*)-3-Hydroxy-9-hexadecenoic acid	1	168.68, C			29	56.94, CH	3.89, overlap
	2a	39.41, CH_2	2.22, dd (14.1, 3.0)		30	35.28, CH	1.52, m
	2b		2.44, overlap		31	24.69, CH_2	0.89, m
	3	71.64, CH	4.94, m		32	10.32, CH_3	0.70, overlap
	4	33.09, CH_2	1.46, m		33	14.83, CH_3	0.70, overlap
	5	24.07, CH_2	1.23, m		34	172.87, C	
	6	28.28, CH_2	1.23, m	Leu	35		8.68, d (5.7)
	7	29.12, CH_2	1.28, m		36	52.42, CH	3.88, overlap
	8	26.58, CH_2	1.99, p (6.0)		37	39.08, CH_2	1.26, m
	9	129.63, CH	5,33, overlap		38	23.71, CH	1.23, m
	10	129.63, CH	5,33, overlap		39	21.45, CH_3	0.70, overlap
	11	26.61, CH_2	1.99, p (6.0)		40	22.66, CH_3	0.78, overlap
	12	29.17, CH_2	1.28, m		41	171.65, C	
	13	28.50, CH_2	1.23, m	Phe	42		8.11, d (5.4)
	14	31.14, CH_2	1.21, m		43	54.10, CH	4.32, m
	15	22.07, CH_2	1.23, m		44a	35.35, CH_2	2.84, t (12.7)
	16	13.93, CH_3	0.82, overlap		44b		3.22, dd (13.9, 3,9)
Tyr	17		7.96, overlap		45	138.64, C	
	18	53.23, CH	4.62, q (7.5)		46	129.18, CH	7.17, m
	19a	37.55, CH_2	2.64, dd (13.5, 7.3)		47	127.94, CH	7.24, m
	19b		2.75, dd (13.5, 7.1)		48	126.07, CH	7.18, m
	20	127.44, C			49	127.94, CH	7.24, m
	21	130.03, CH	6.92, dd (8.7, 2.6)		50	129.18, CH	7.17, m
	22	114.74, CH	6.58, dd (8.7, 2.6)		51	170.97, C	
	23	155.73, C		Ser	52		7.44, s
	24	114.74, CH	6.58, dd (8.7, 2.6)		53	56.27, CH	4.06, q (6.0)
	25	130.03, CH	6.92, dd (8.7, 2.6)		54a	61.18, CH_2	3.62, m
	26		9.12, s		54b		3.67, m
	27	171.03, C			55		4.82, t (6.7)
Ile	28		7.99, overlap		56	169.56, C	

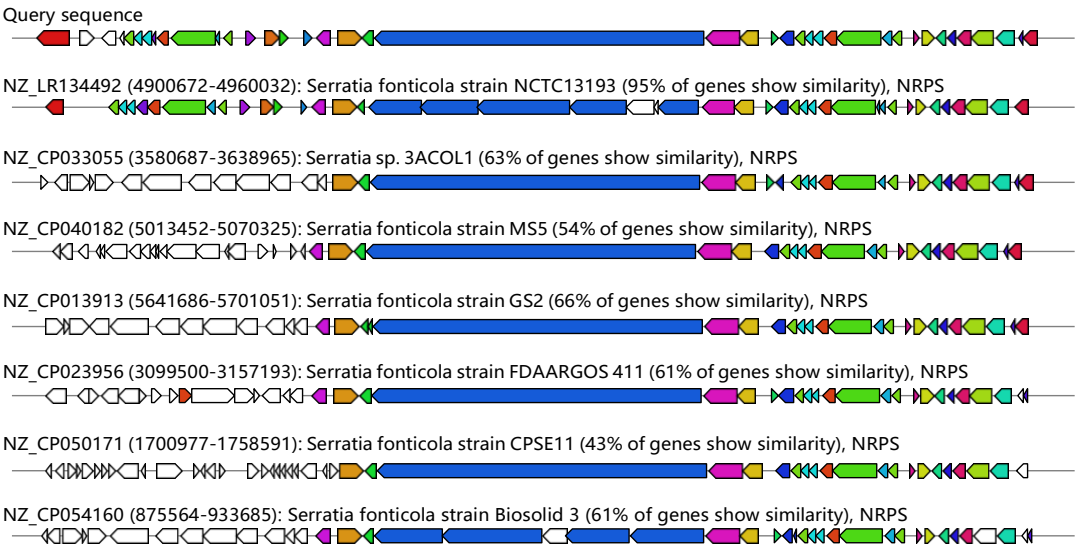


Fig. S1 Alignment of *sefA* and other similar BGCs. Similarity analysis was ccomplished using antiSMASH[2].

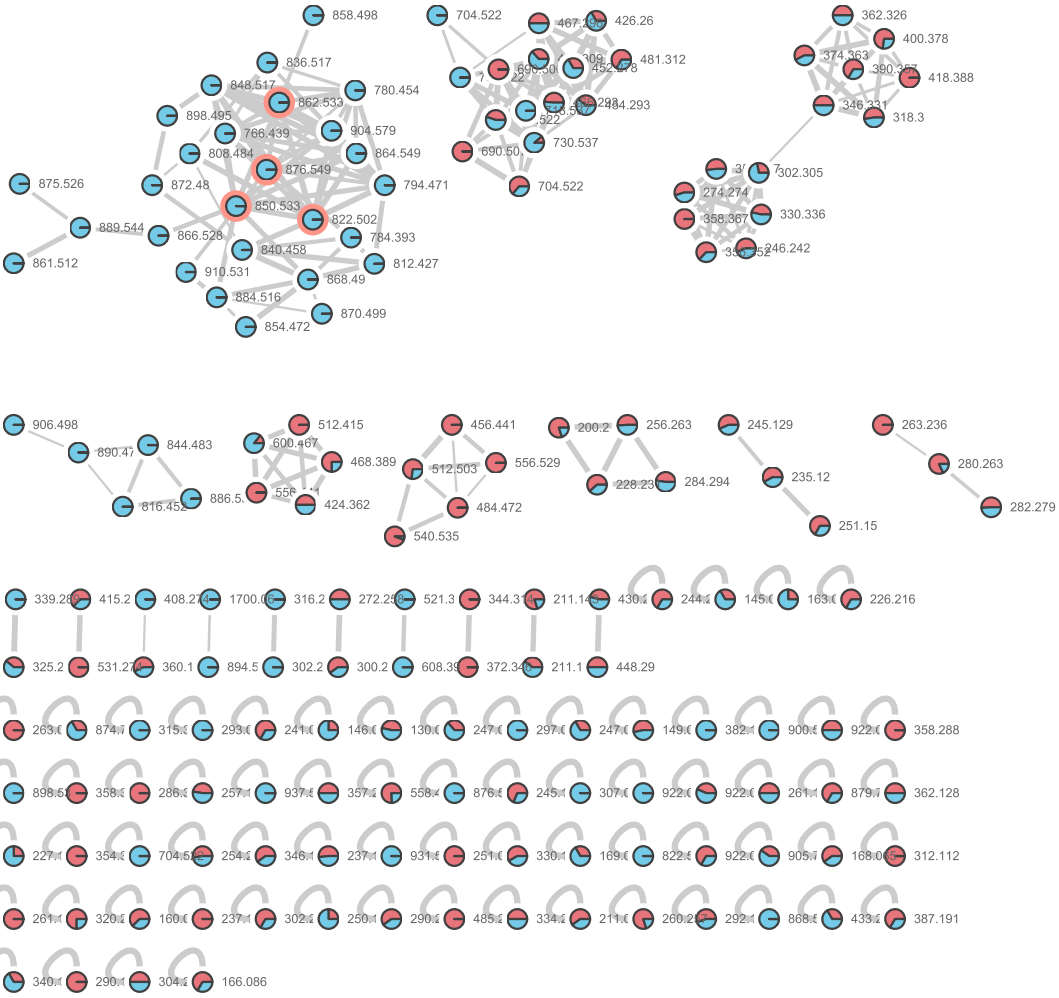


Fig. S2 Molecular network of the secondary metabolites produced by *E. coli* BAP1 pHL1 and pCOLA-Duet generated using the HRLC-MS/MS data. Nodes are labelled with the corresponding m/z values (detected in the positive mode). The light blue nodes correspond to molecules present in *E. coli* BAP1 pHL1, while red nodes represent those found in *E. coli* BAP1 pCOLA-Duet. A blue-only node signifies its exclusive presence in *E. coli* BAP1 pHL1. The thickness of the lines indicates the strength of the association. sefopeptides A-D (**1**-**4**) are highlighted with red borders around the nodes.

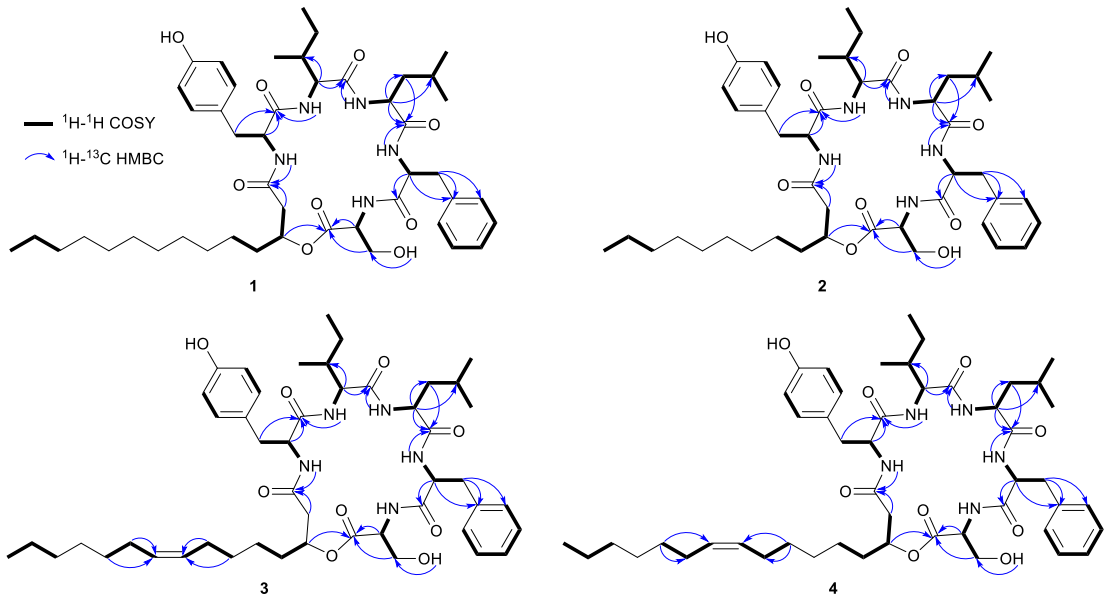


Fig. S3 ^{1}H-^{1}H COSY and key ^{1}H-^{13}C HMBC correlations of compounds **1**-**4.**

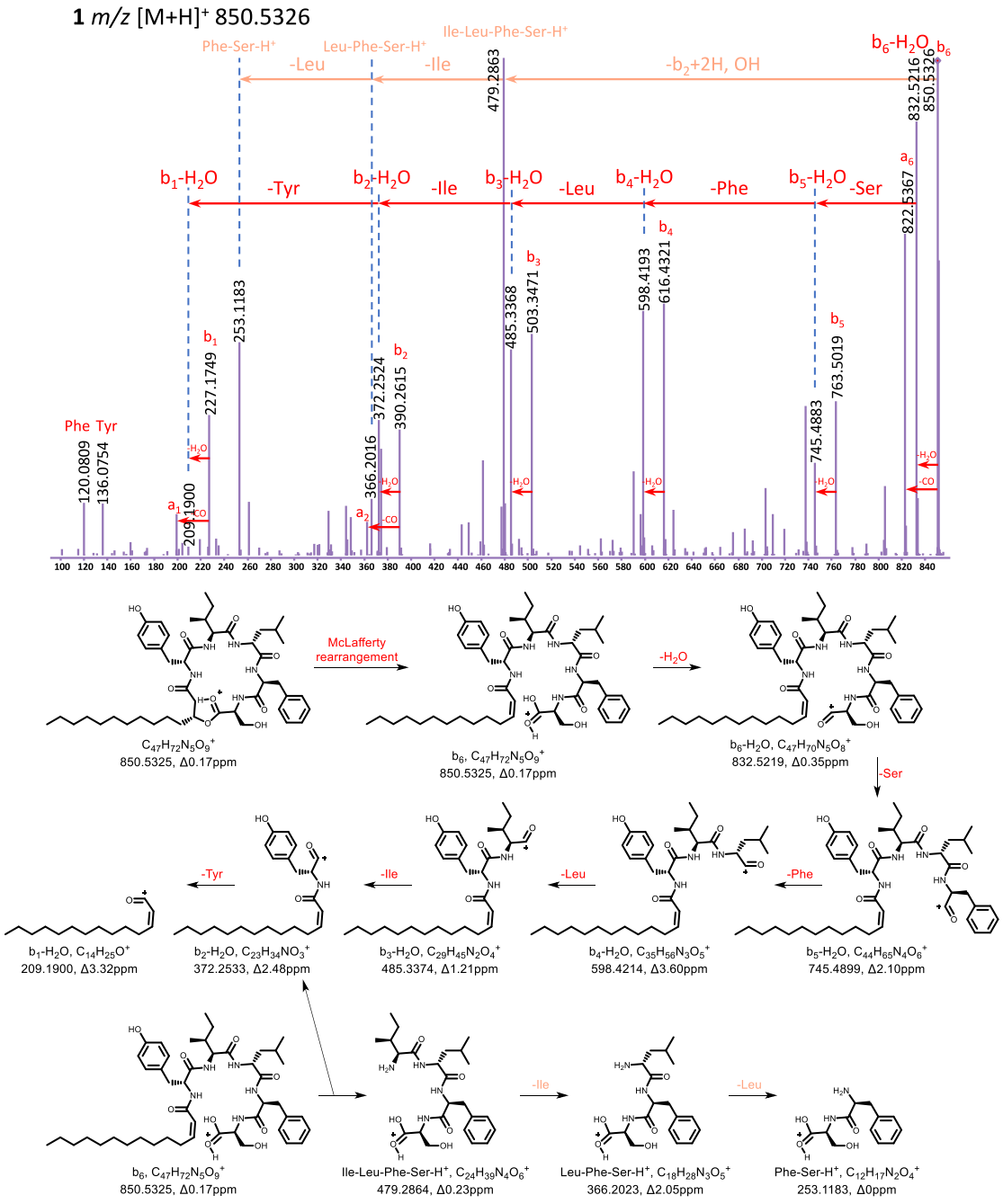


Fig. S4 MS/MS spectra and fragmentation pathways with proposed fragment structures for **1**.

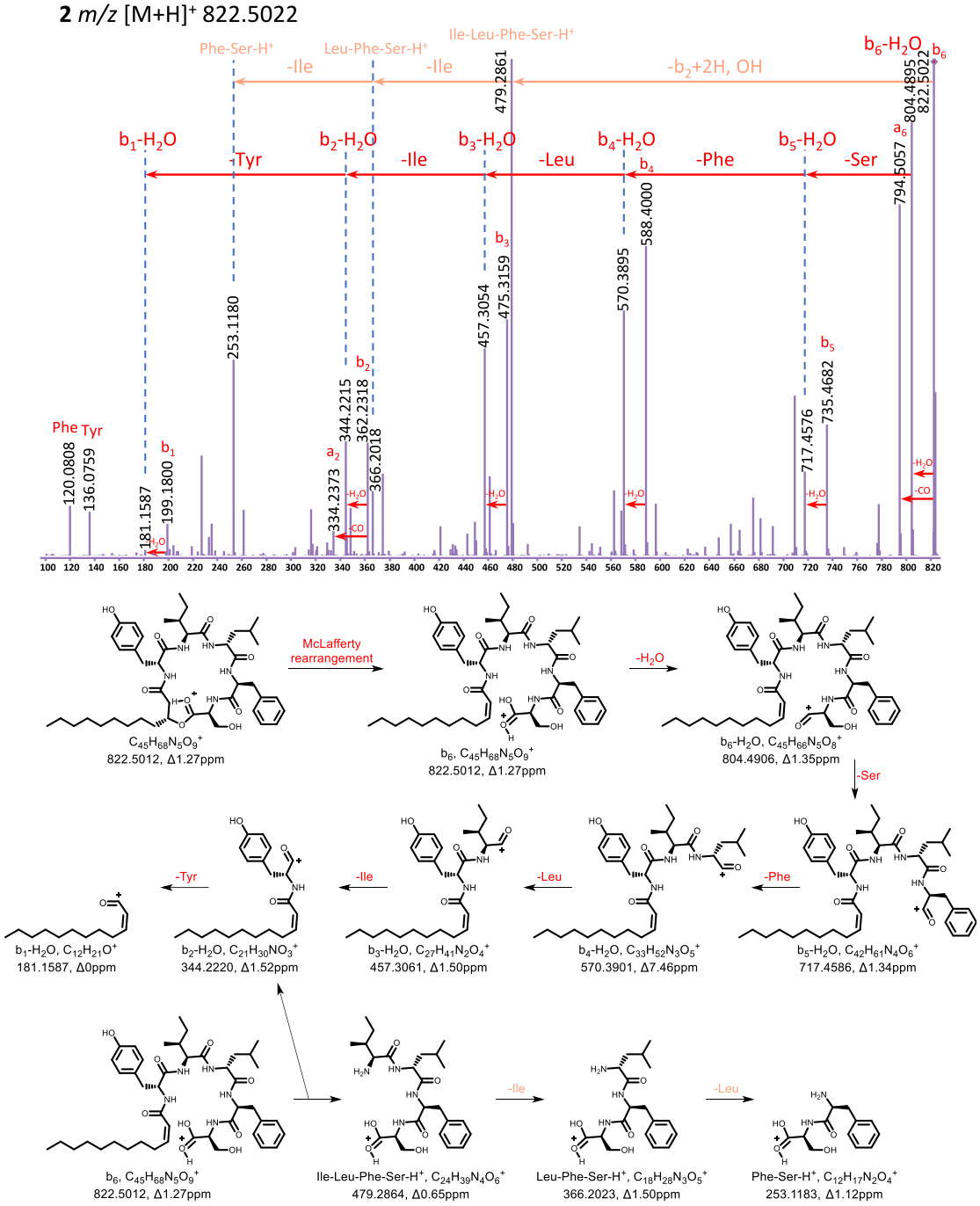


Fig. S5 MS/MS spectra and fragmentation pathways with proposed fragment structures for **2**.

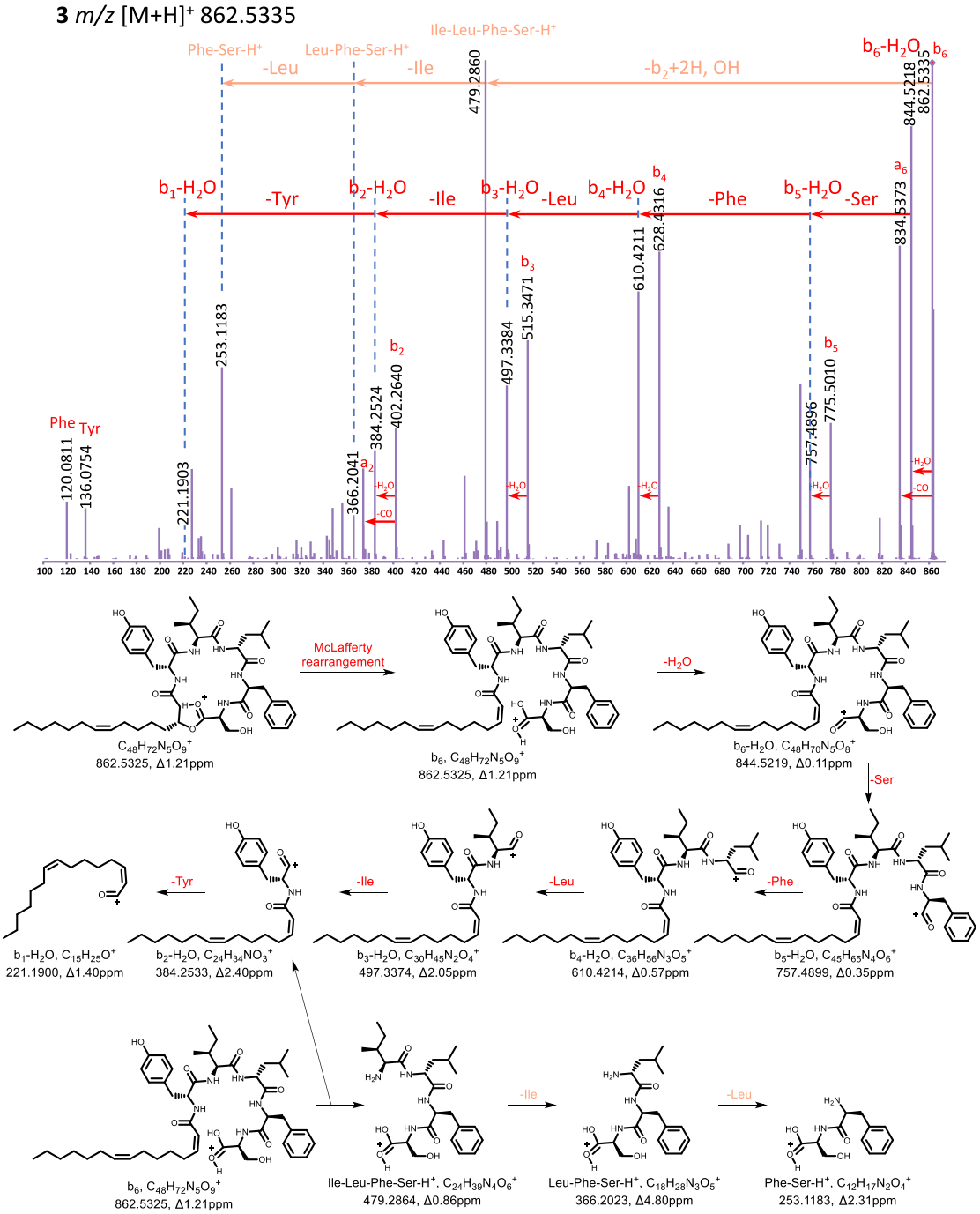


Fig. S6 MS/MS spectra and fragmentation pathways with proposed fragment structures for **3**.

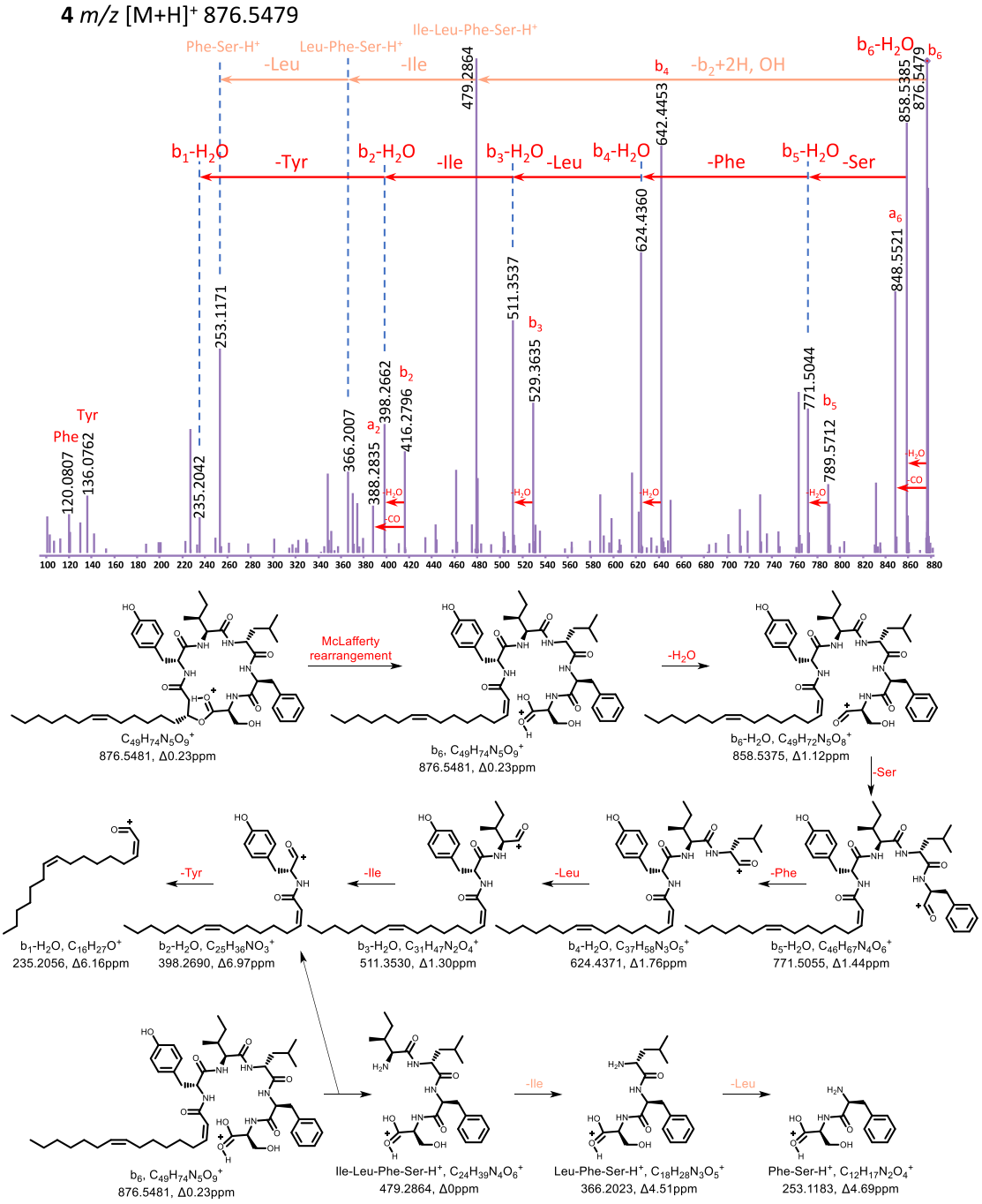
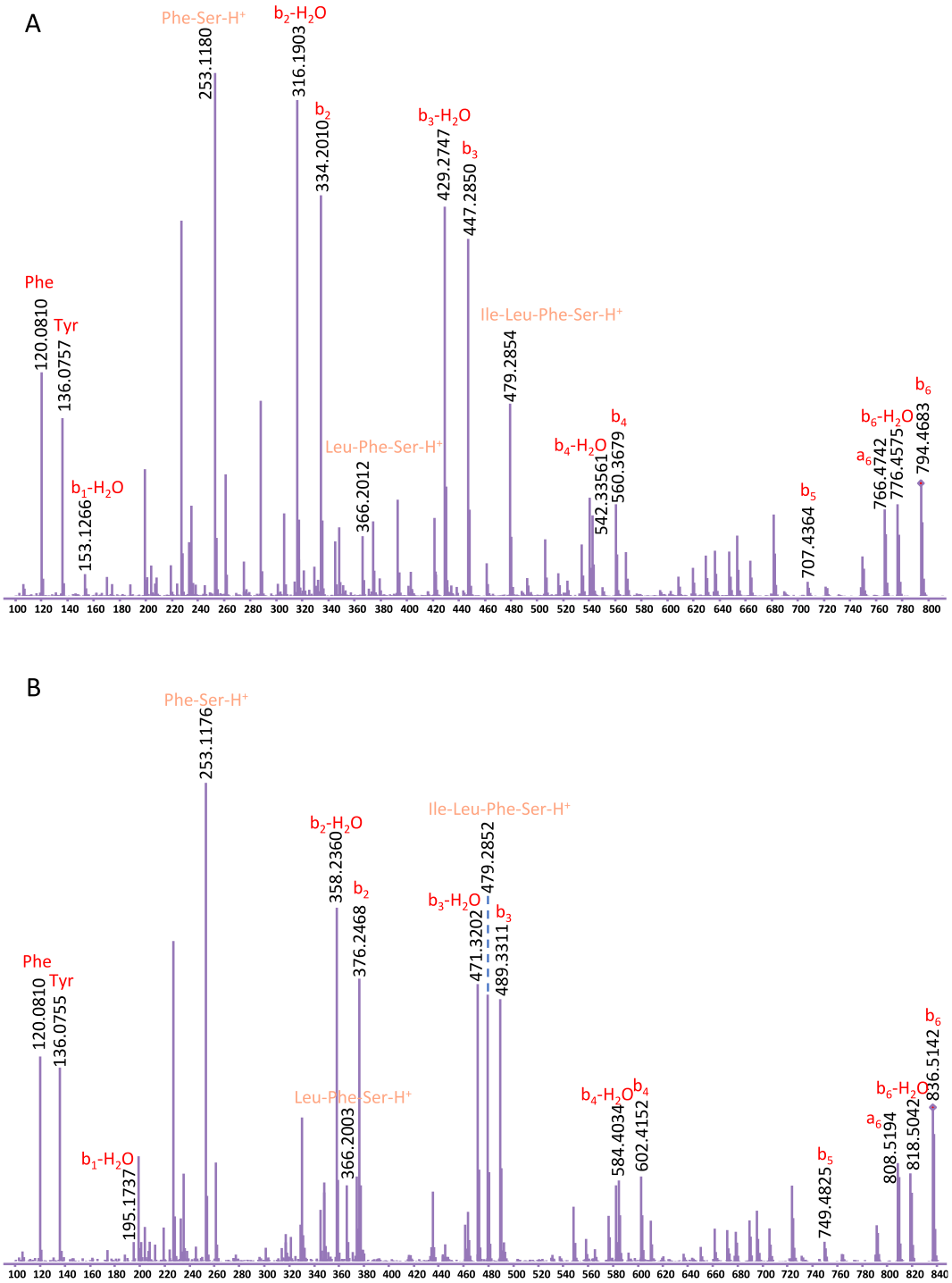


Fig. S7 MS/MS spectra and fragmentation pathways with proposed fragment structures for **4**.


A
Phe-Ser-H+
253.1180
b2-H2O
316.1903
b2
334.2010
b3-H2O
429.2747
b3
447.2850
Phe
120.0810
Tyr
136.0757
Ile-Leu-Phe-Ser-H+
479.2854
b4
b4-H2O
542.33561
560.3679
Leu-Phe-Ser-H+
366.2012
b1-H2O
153.1266
b5
707.4364
b6
b6-H2O
a6
766.4742
776.4575
794.4683
100 120 140 160 180 200 220 240 260 280 300 320 340 360 380 400 420 440 460 480 500 520 540 560 580 600 620 640 660 680 700 720 740 760 780 800
B
Phe-Ser-H+
253.1176
Ile-Leu-Phe-Ser-H+
b2-H2O
358.2360
479.2852
b2
376.2468
b3-H2O
471.3202
b3
489.3311
Phe
120.0810
Tyr
136.0755
b6
836.5142
b6-H2O
a6
808.5194
818.5042
b4
b4-H2O
584.4034
602.4152
Leu-Phe-Ser-H+
366.2003
b1-H2O
195.1737
b5
749.4825
100 120 140 160 180 200 220 240 260 280 300 320 340 360 380 400 420 440 460 480 500 520 540 560 580 600 620 640 660 680 700 720 740 760 780 800 820 840

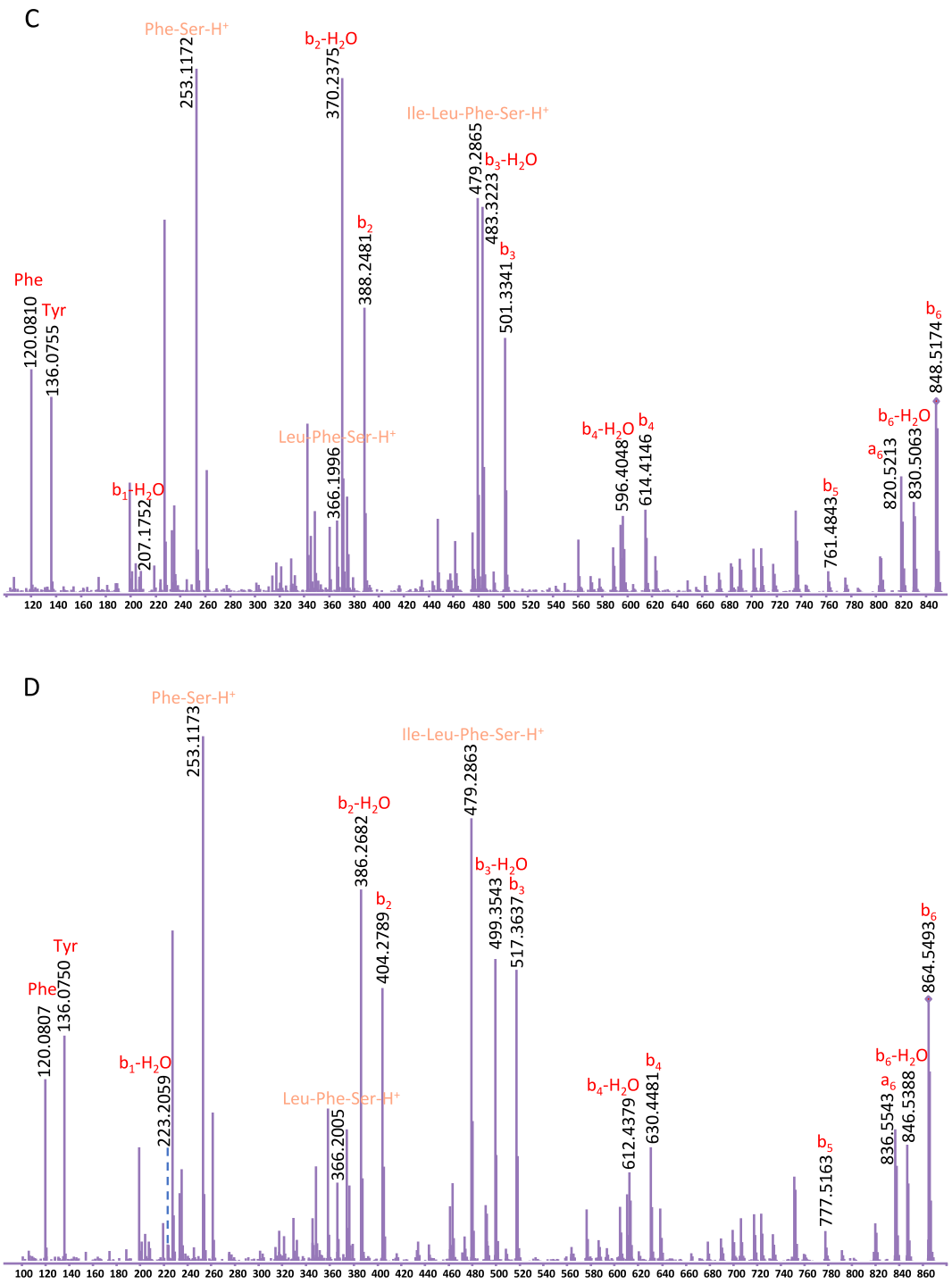
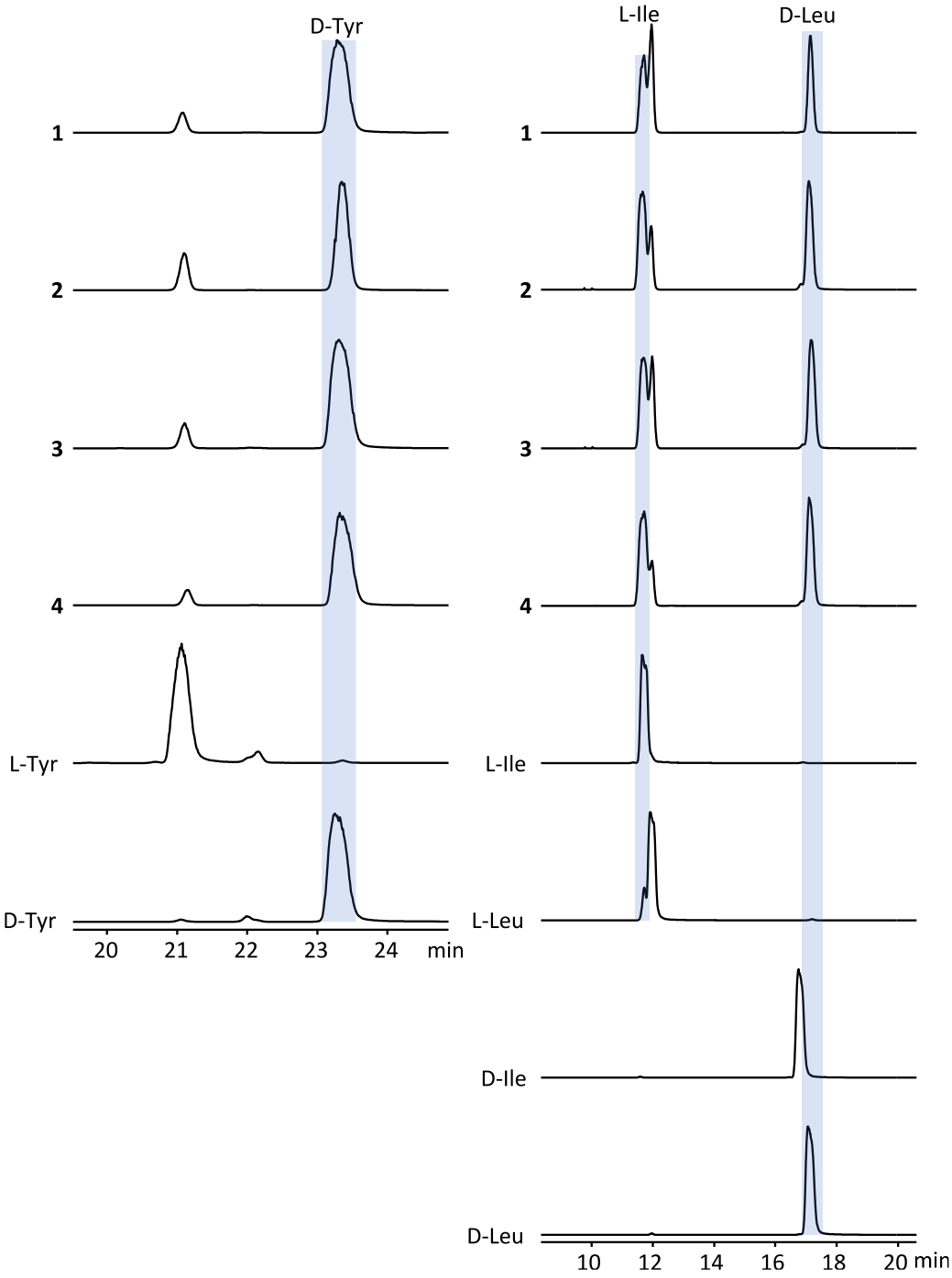


Fig S8 We selected nodes from the molecular network with molecular weights of 794.471 (A), 836.517 (B), 848.517 (C), and 864.549 (D) for the analysis of their MS/MS spectra. This analysis reveals that these analogues share the same peptide core as sefopeptides A-D but exhibit variations only in the fatty acid chain.


D-Tyr
L-Ile
D-Leu
1
2
3
4
L-Tyr
D-Tyr
L-Ile
L-Leu
D-Ile
D-Leu
20
21
22
23
24
min
10
12
14
16
18
20
min

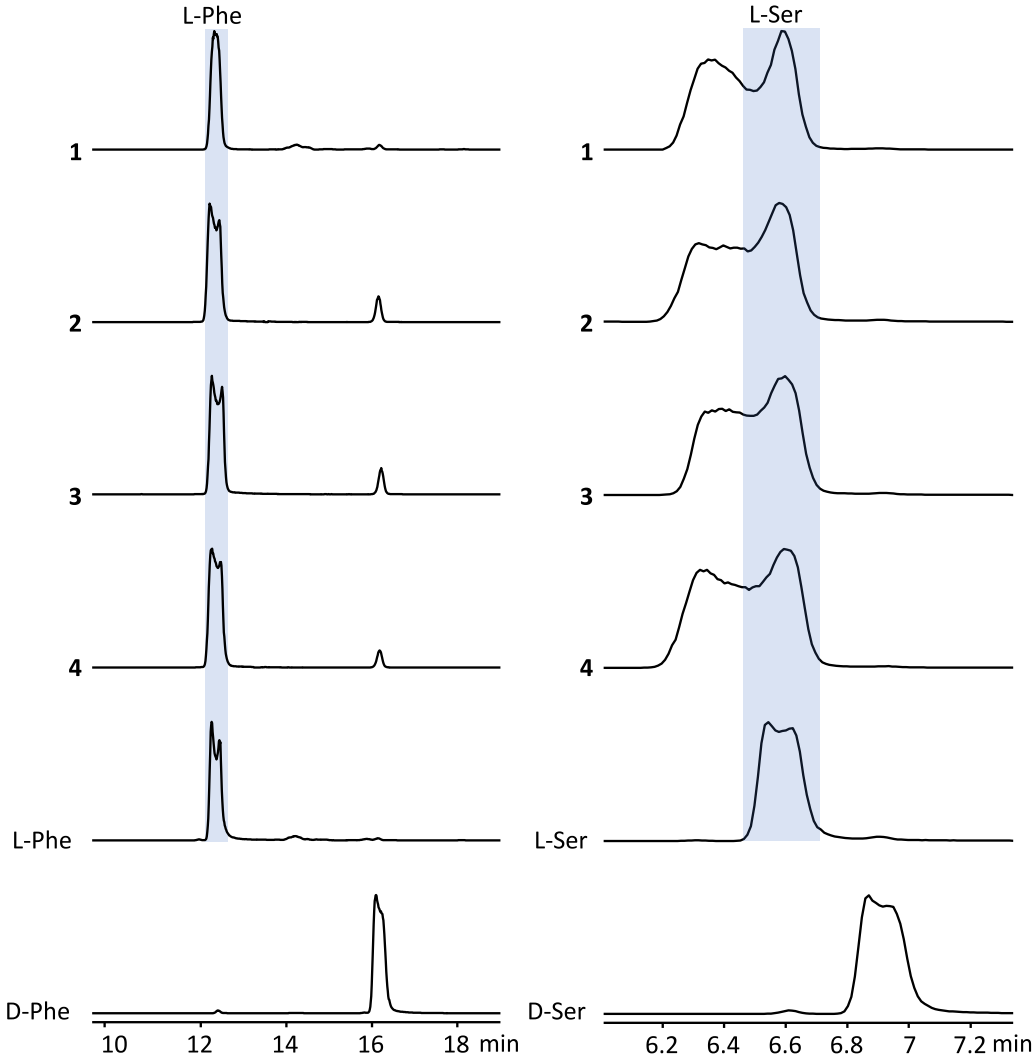


Fig. S9 Configuration determination of amino acids in **1**-**4** using the Marfey's method. HRLC-MS analysis of hydrolyzed **1**-**4** and amino acid standards derivatized with L-FDLA. Depicted are EIC traces for Tyrosine (Tyr, *m/z* 770 $[M + H]^+$, di-substituted Marfey's derivatives[3]), isoleucine (Ile, *m/z* 426 $[M + H]^+$), leucine (Leu, *m/z* 426 $[M + H]^+$), phenylalanine (Phe, *m/z* 460 $[M + H]^+$), and serine (Ser, *m/z* 400 $[M + H]^+$). A minor peak of L-Leu is likely due to racemization of D-Leu during acid hydrolysis[4].

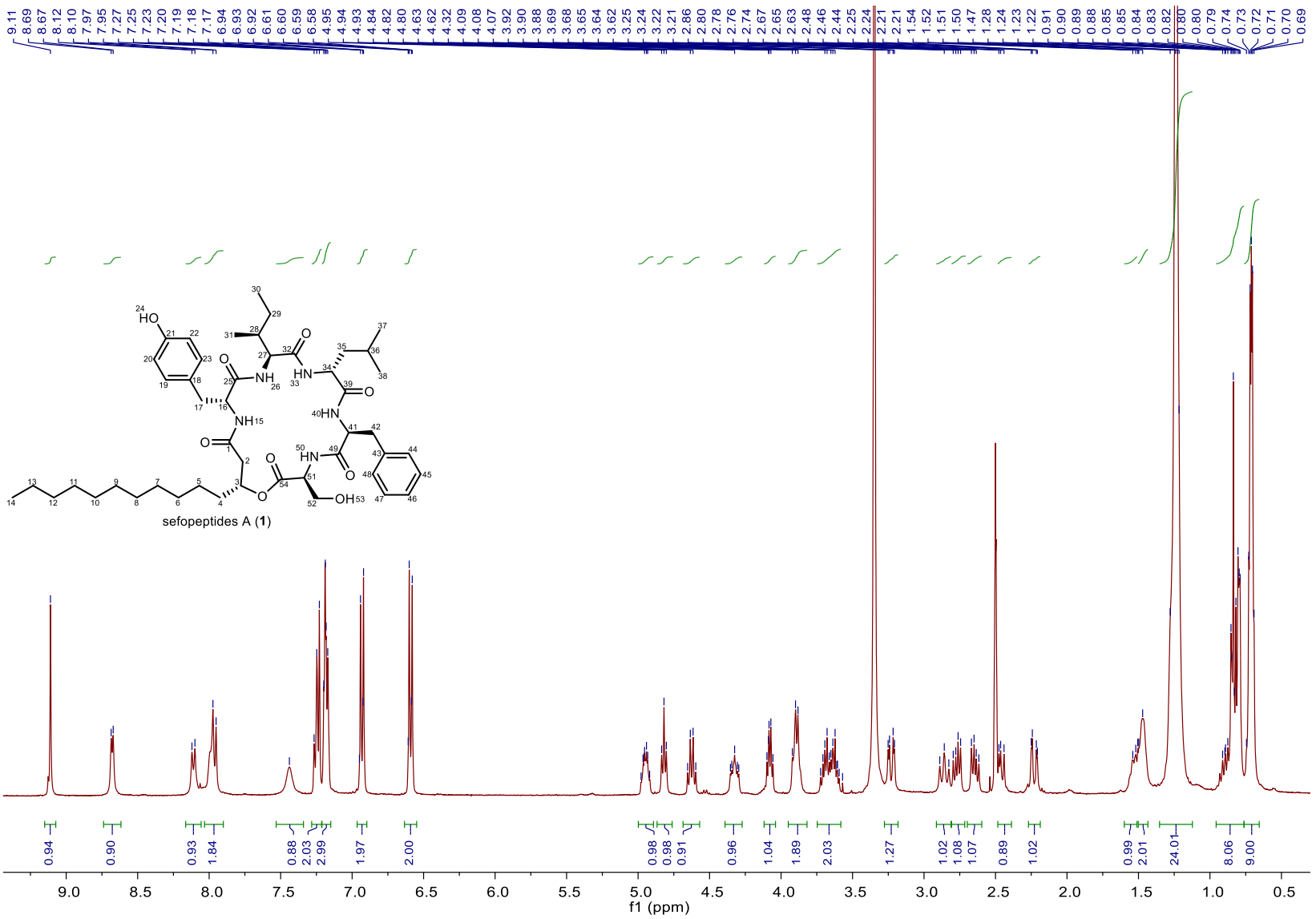


Fig. S10 ^{1}H NMR (400 MHz, DMSO-d_6) spectrum of sefopeptides A (**1**).

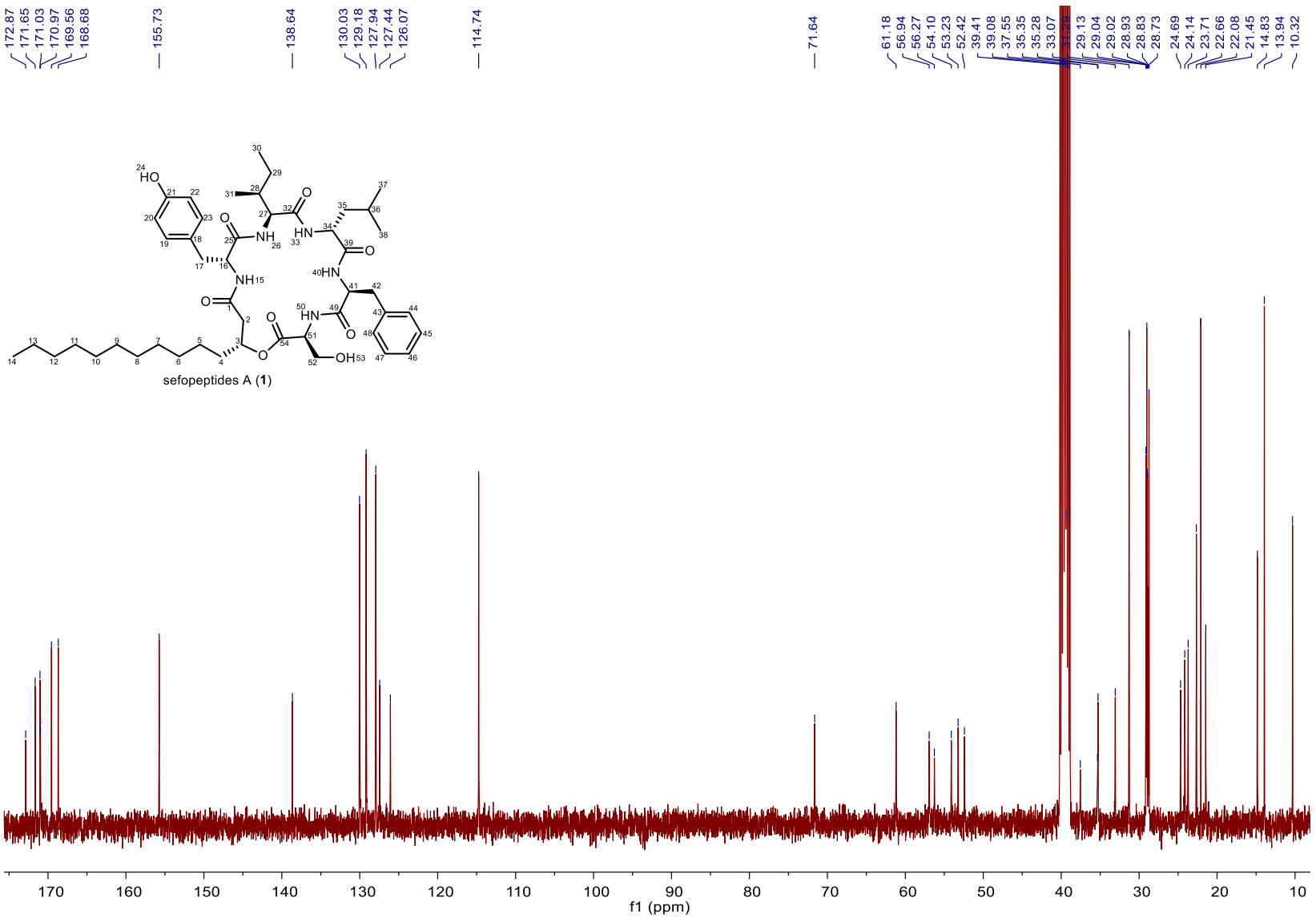


Fig. S11 ^{13}C NMR (100 MHz, DMSO-d_6) spectrum of sefopeptides A (**1**).

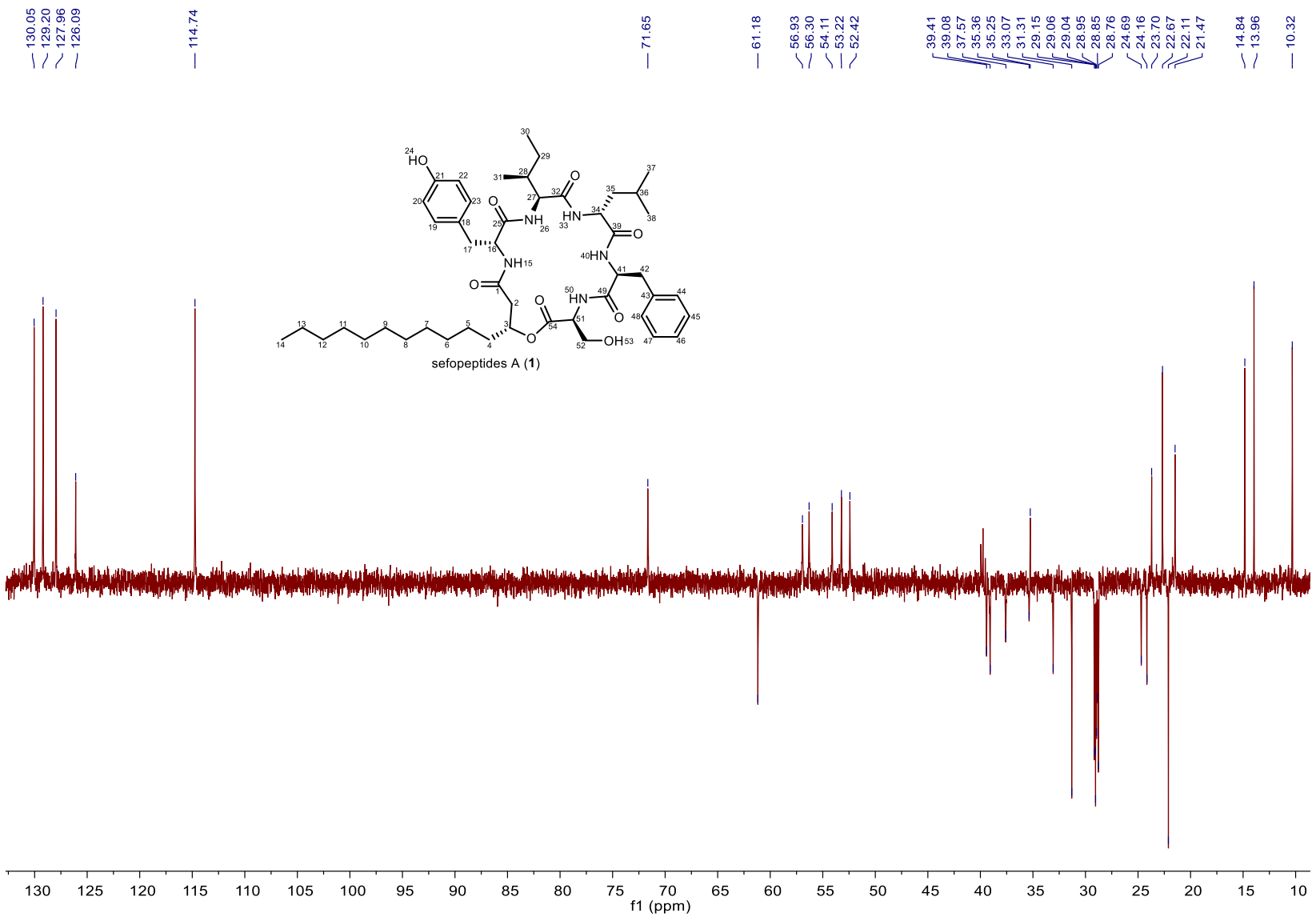


Fig. S12 DEPT-135 (100 MHz, DMSO-d_6) spectrum of sefopeptides A (**1**).

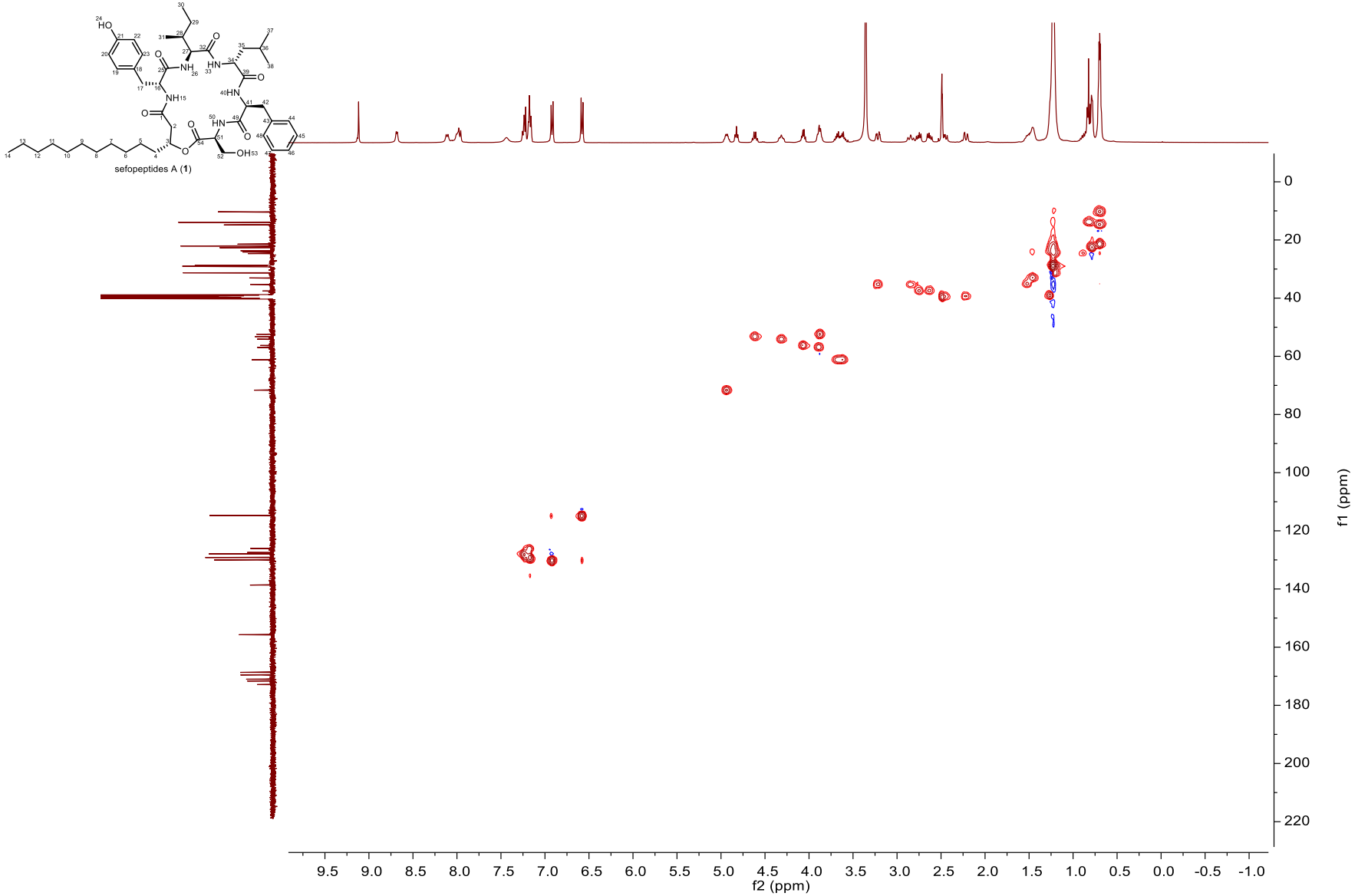


Fig. S13 HSQC (DMSO-d_6) spectrum of sefopeptides A (**1**).

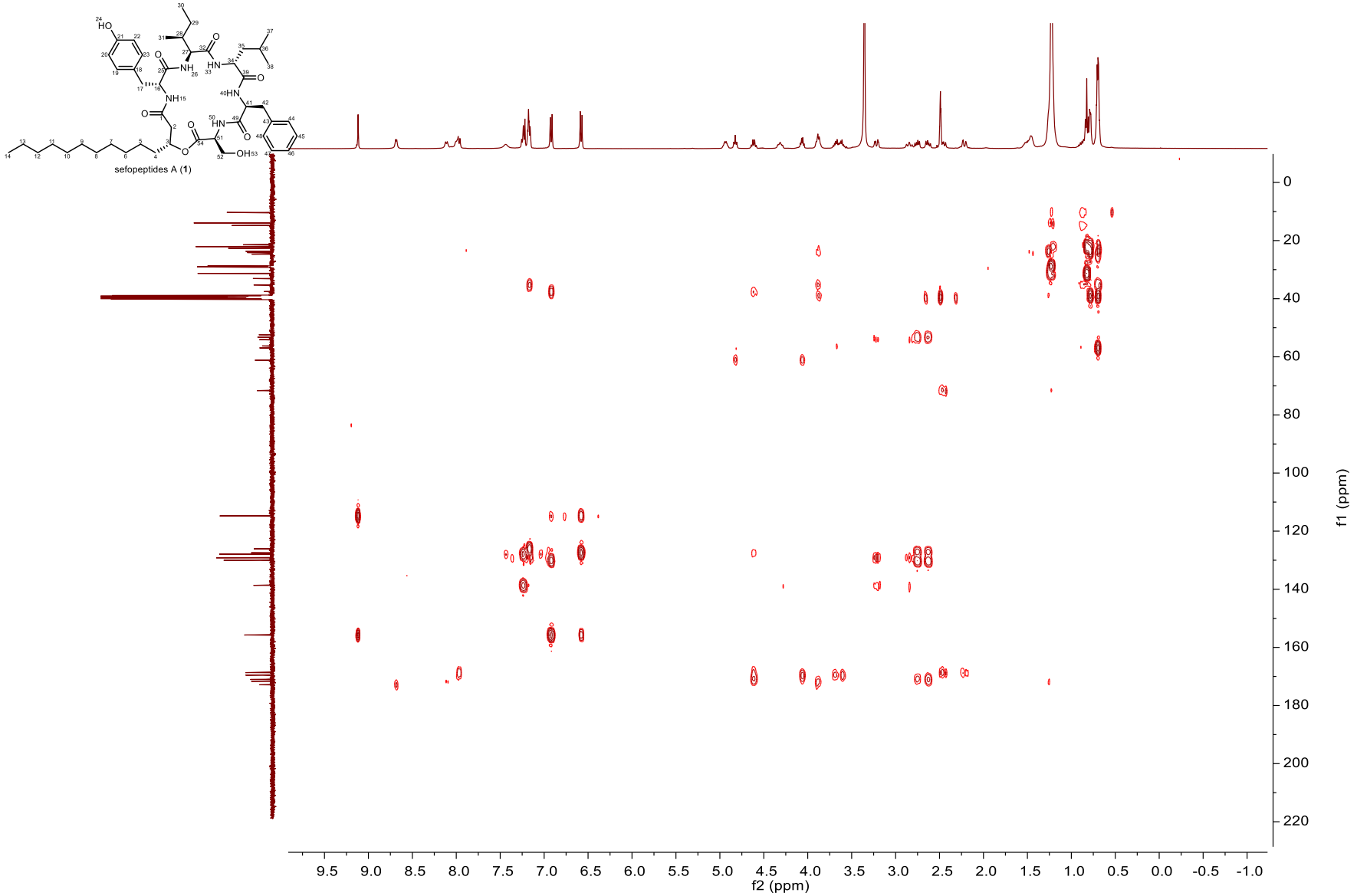


Fig. S14 ^{1}H-^{13}C HMBC (DMSO-d_6) spectrum of sefopeptides A (**1**).

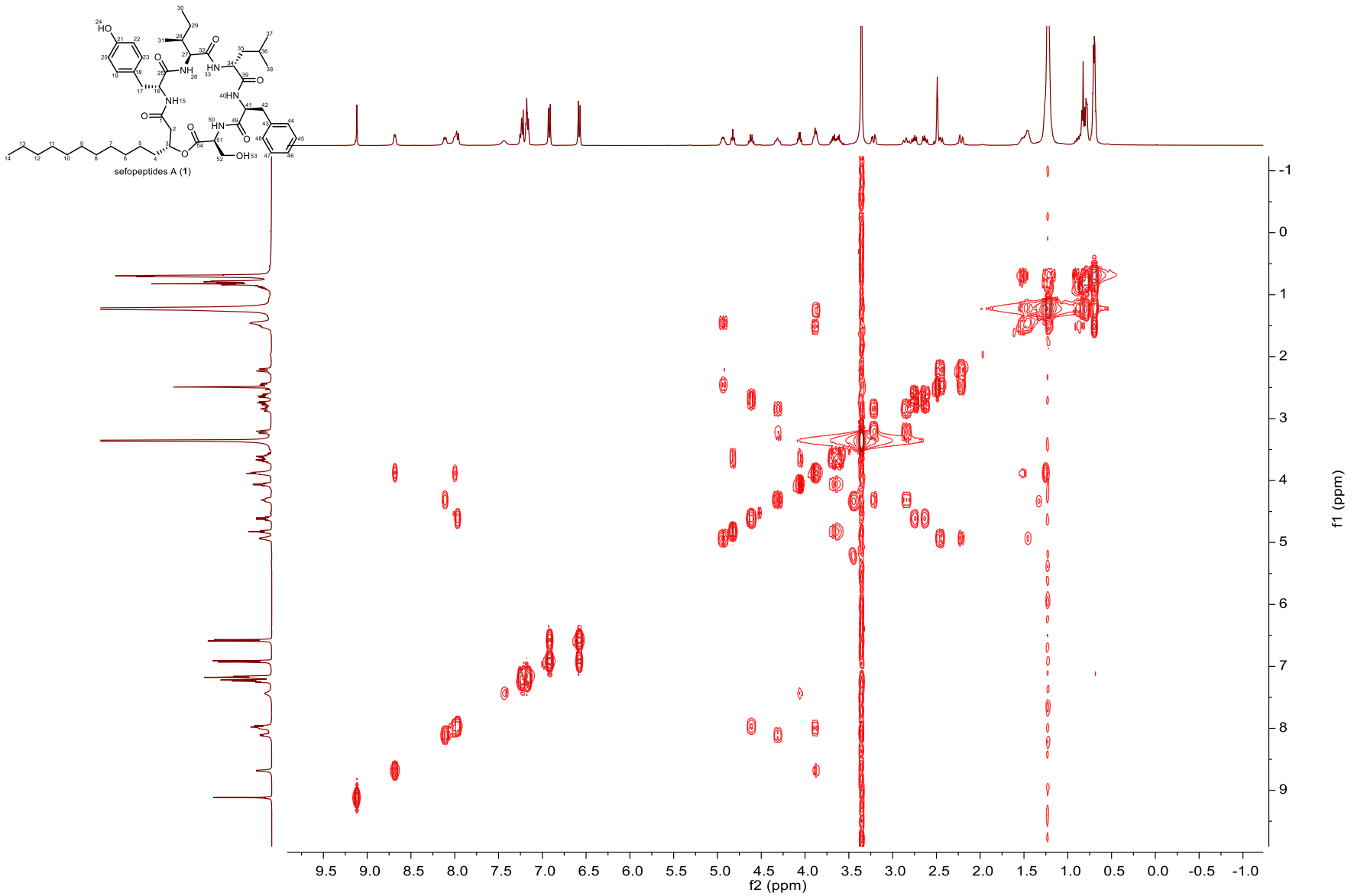


Fig. S15 ^{1}H-^{1}H COSY (DMSO-d_6) spectrum of sefopeptides A (**1**).

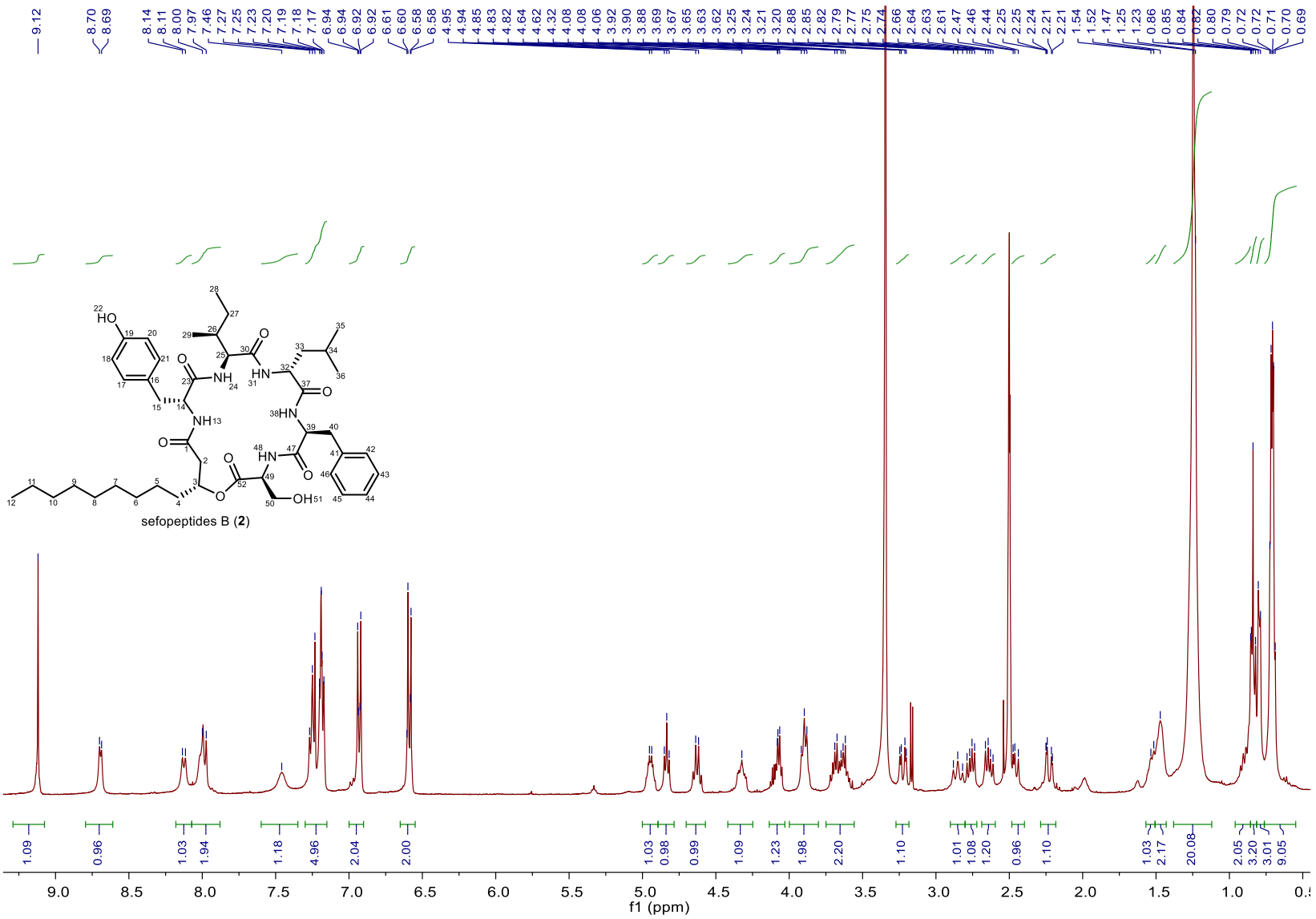


Fig. S16 ^{1}H NMR (400 MHz, DMSO-d_6) spectrum of sefopeptides B (**2**).

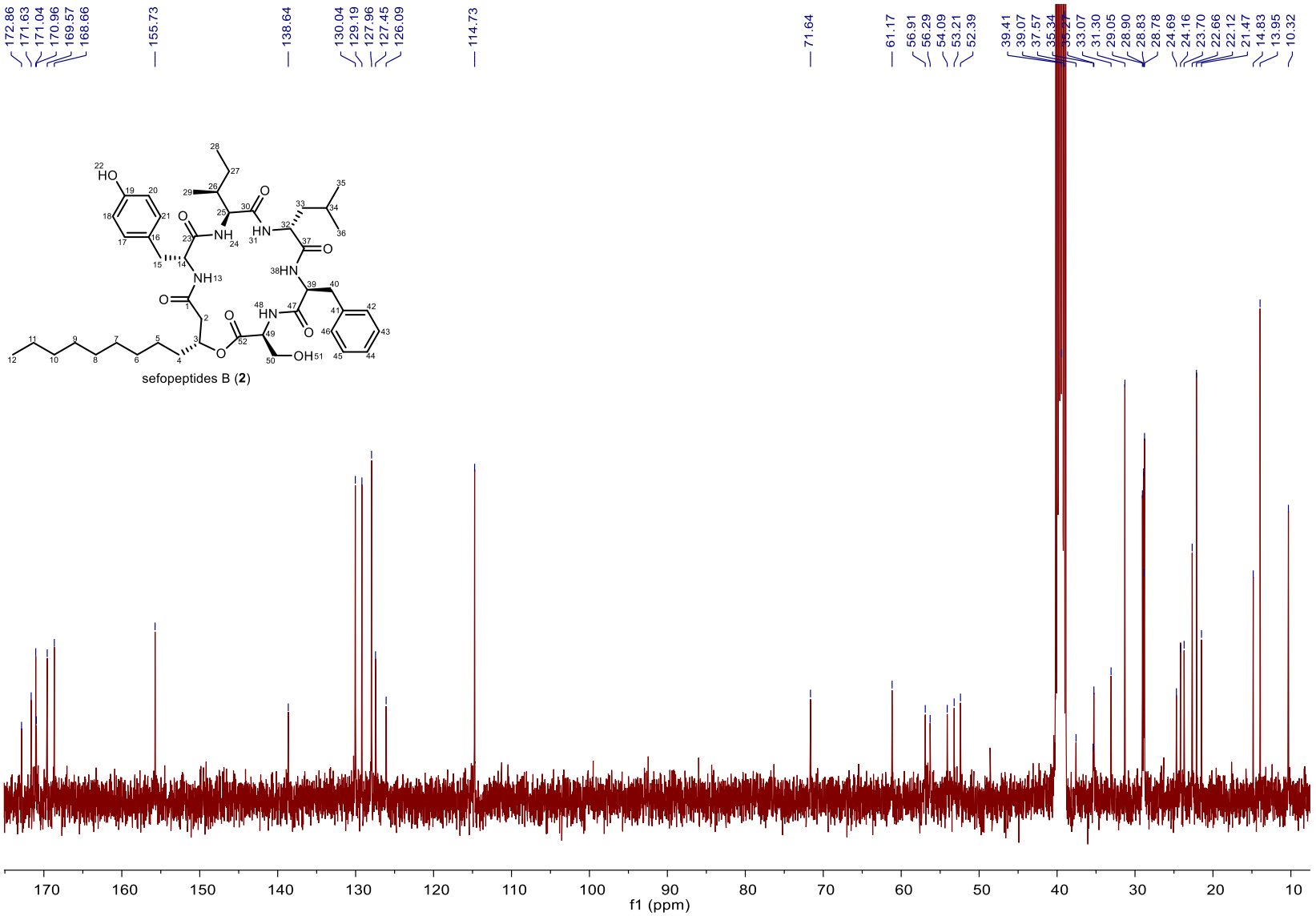


Fig. S17 ^{13}C NMR (100 MHz, DMSO-d_6) spectrum of sefopeptides B (**2**).

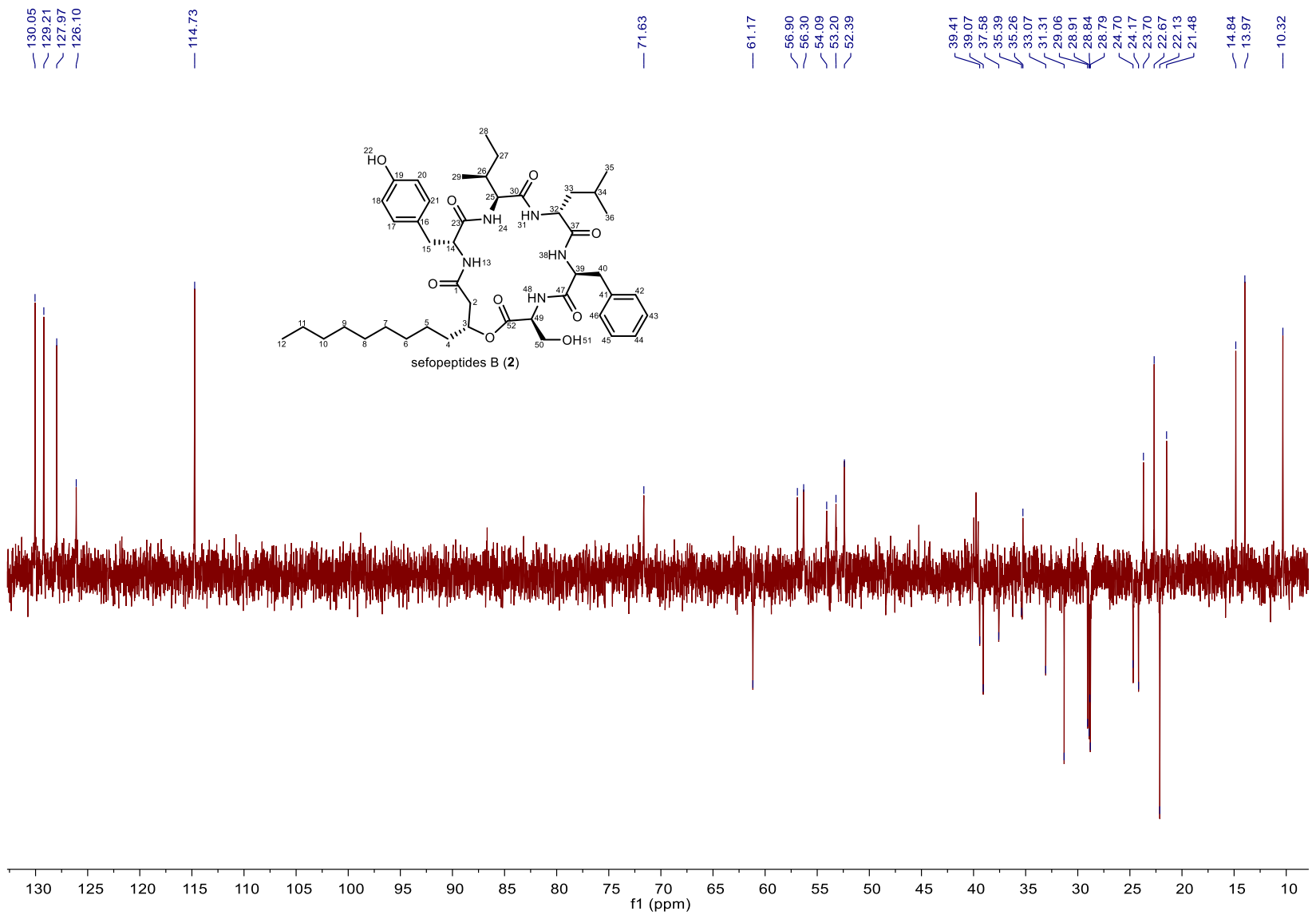


Fig. S18 DEPT-135 (100 MHz, DMSO-d_6) spectrum of sefopeptides B (**2**).

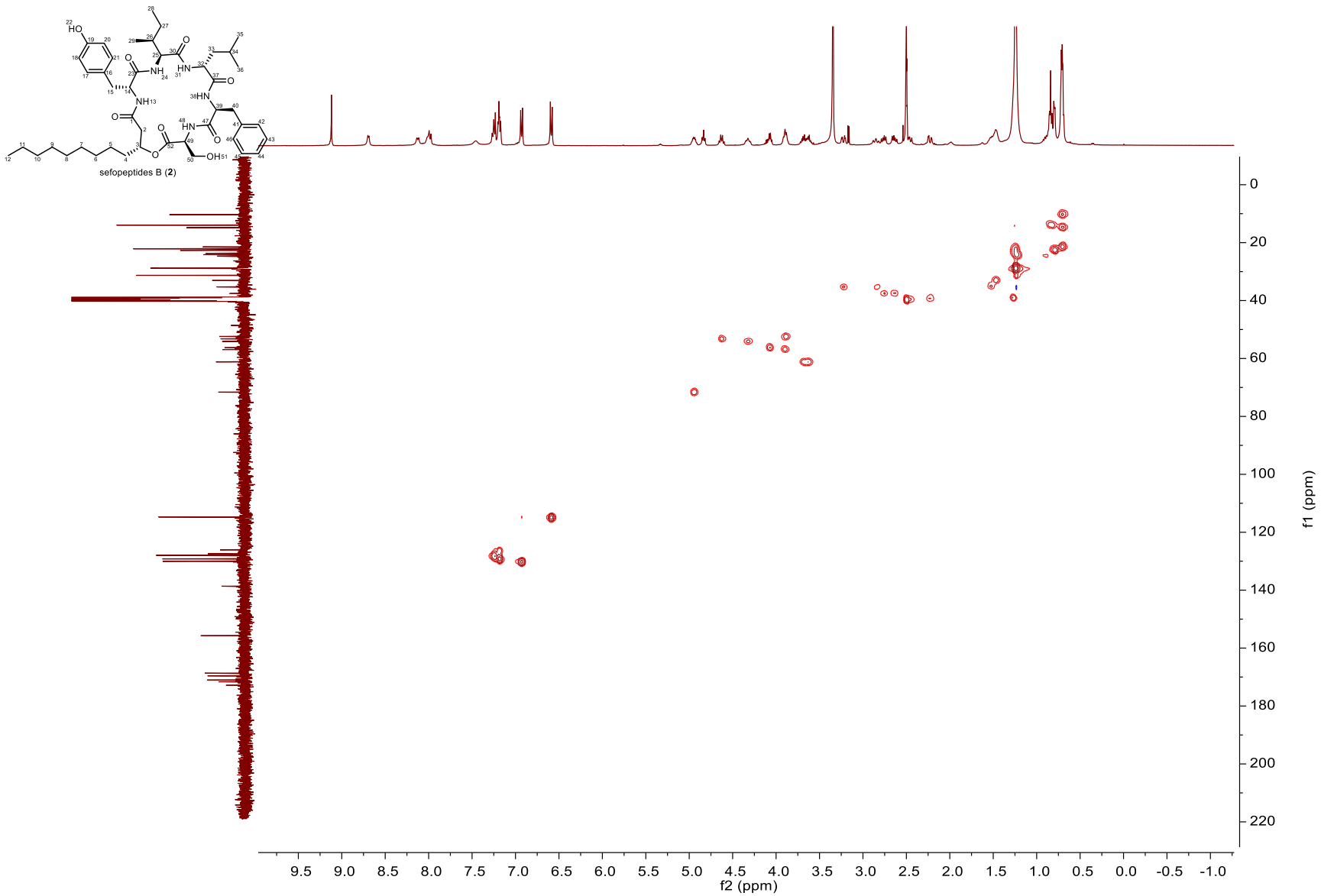


Fig. S19 HSQC (DMSO-d_6) spectrum of sefopeptides B (**2**).

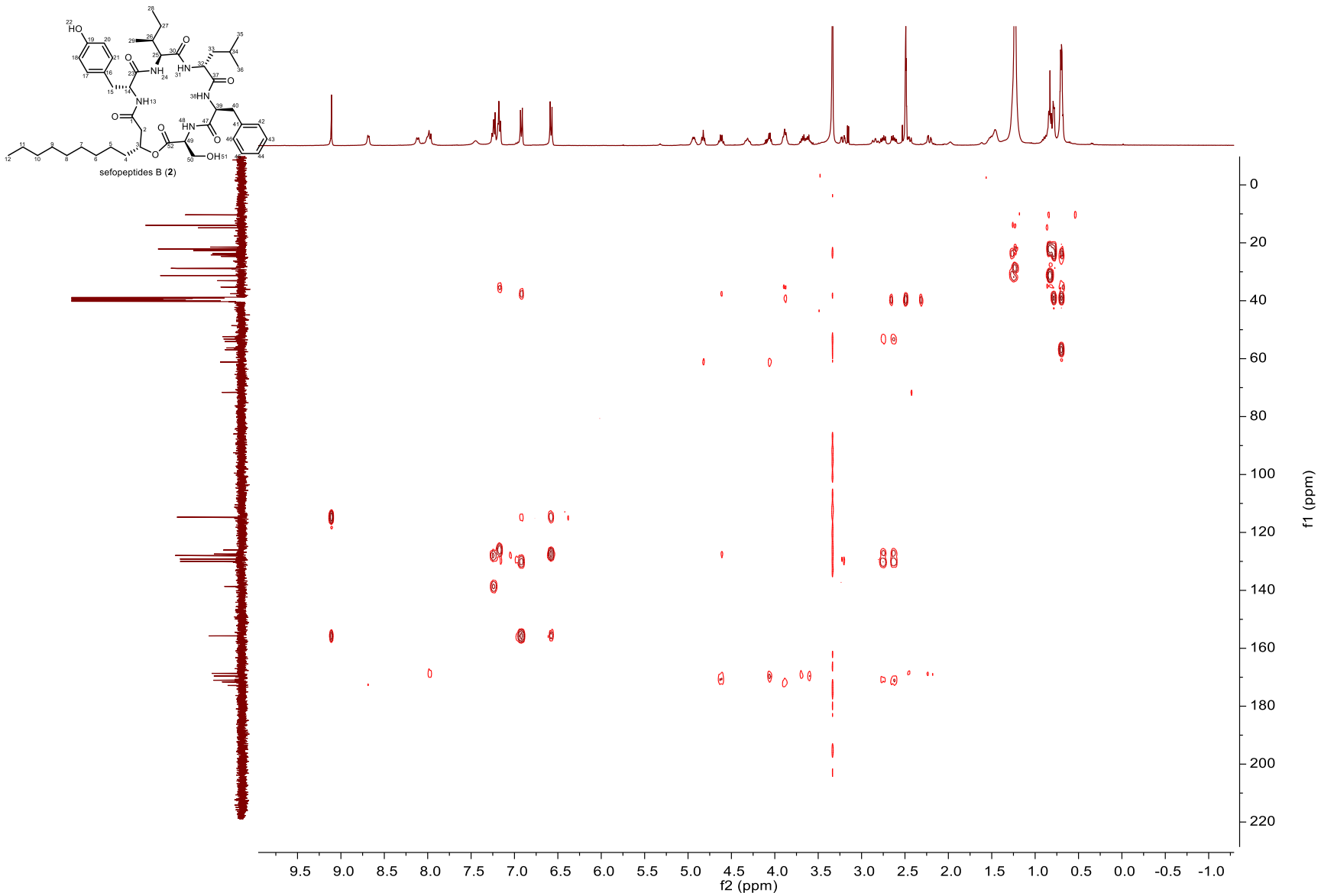


Fig. S20 ^{1}H-^{13}C HMBC (DMSO-d_6) spectrum of sefopeptides B (**2**).

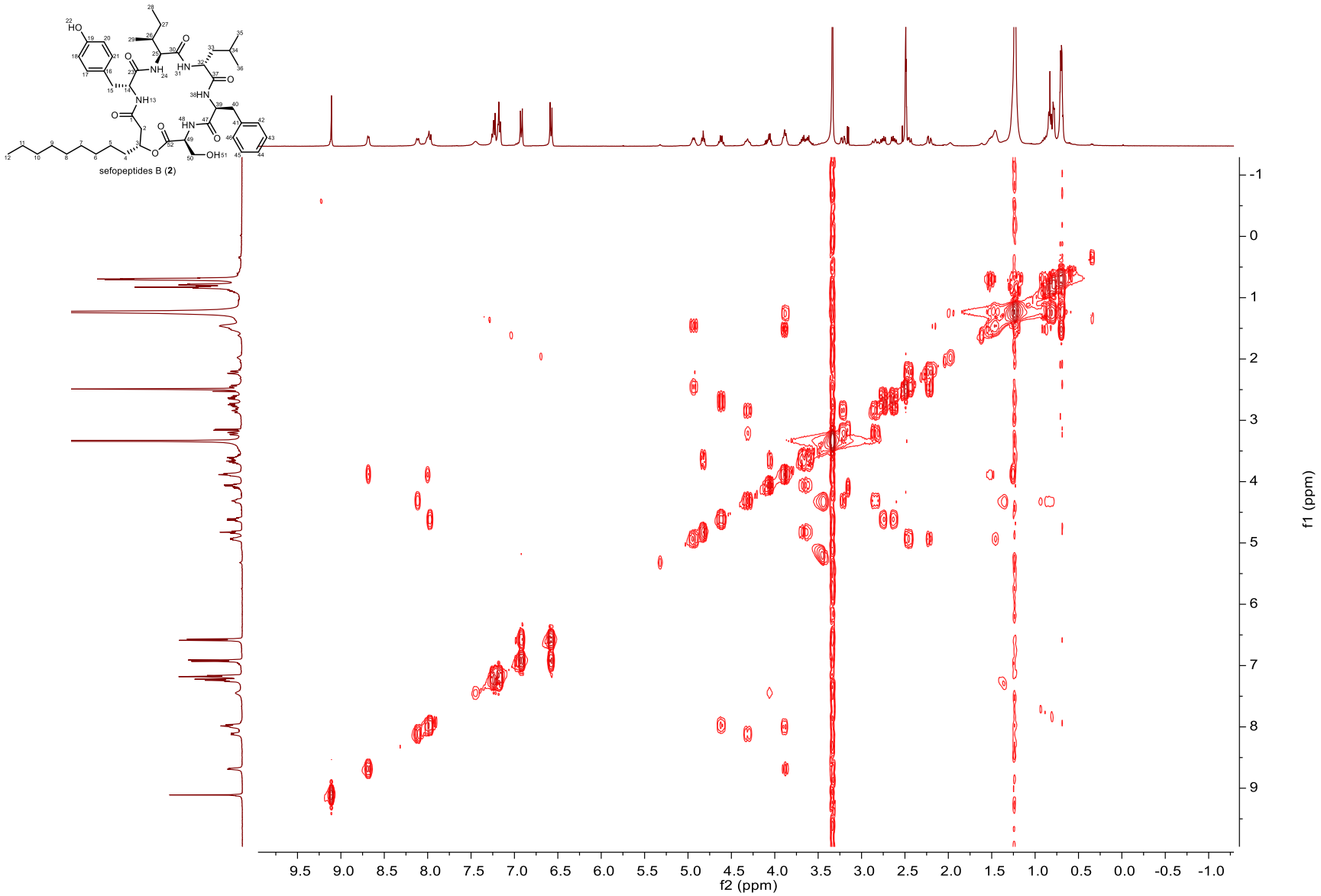


Fig. S21 ^{1}H-^{1}H COSY (DMSO-d_6) spectrum of sefopeptides B (**2**).

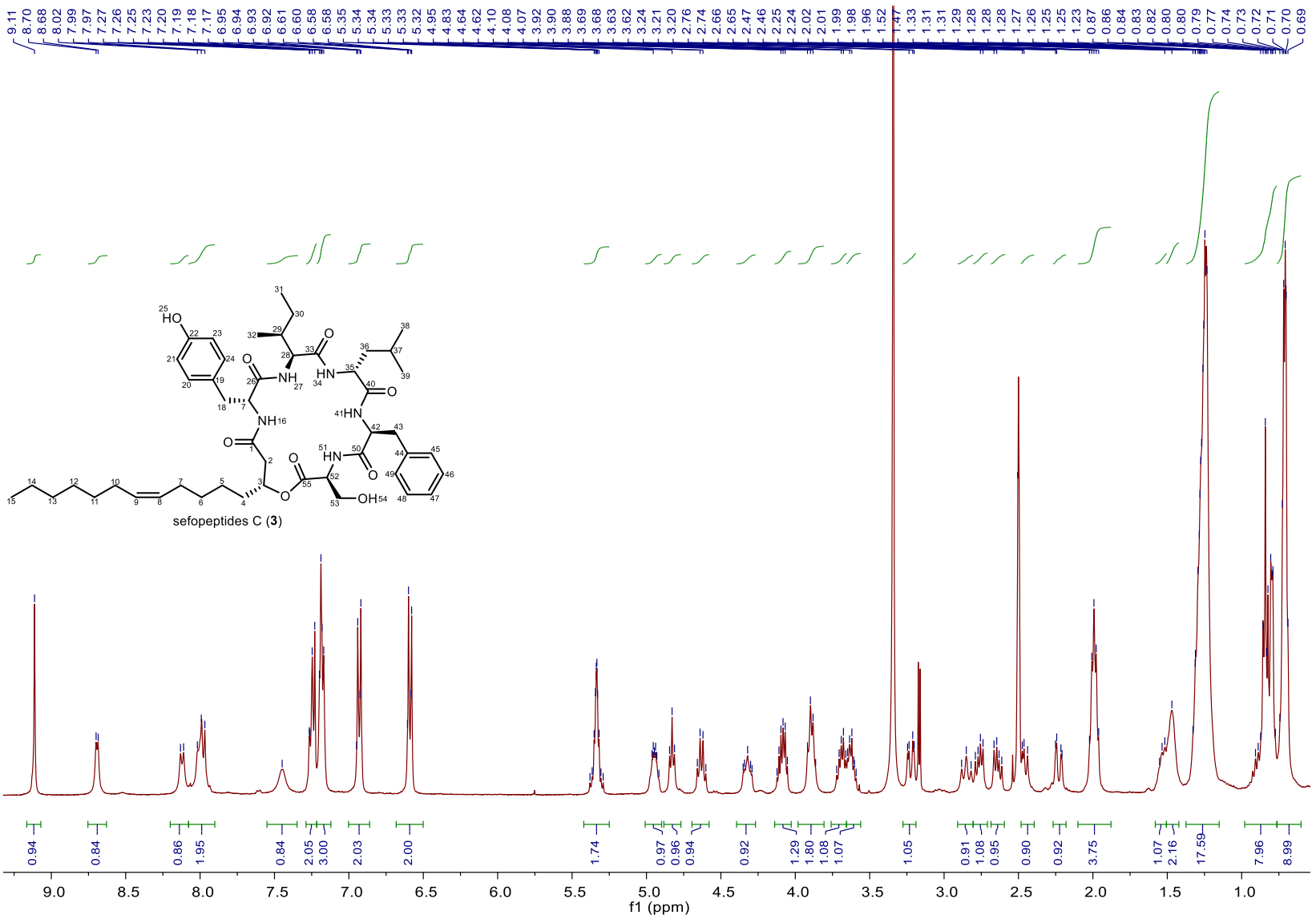


Fig. S22 ^{1}H NMR (400 MHz, DMSO-d_6) spectrum of sefopeptides C (**3**).

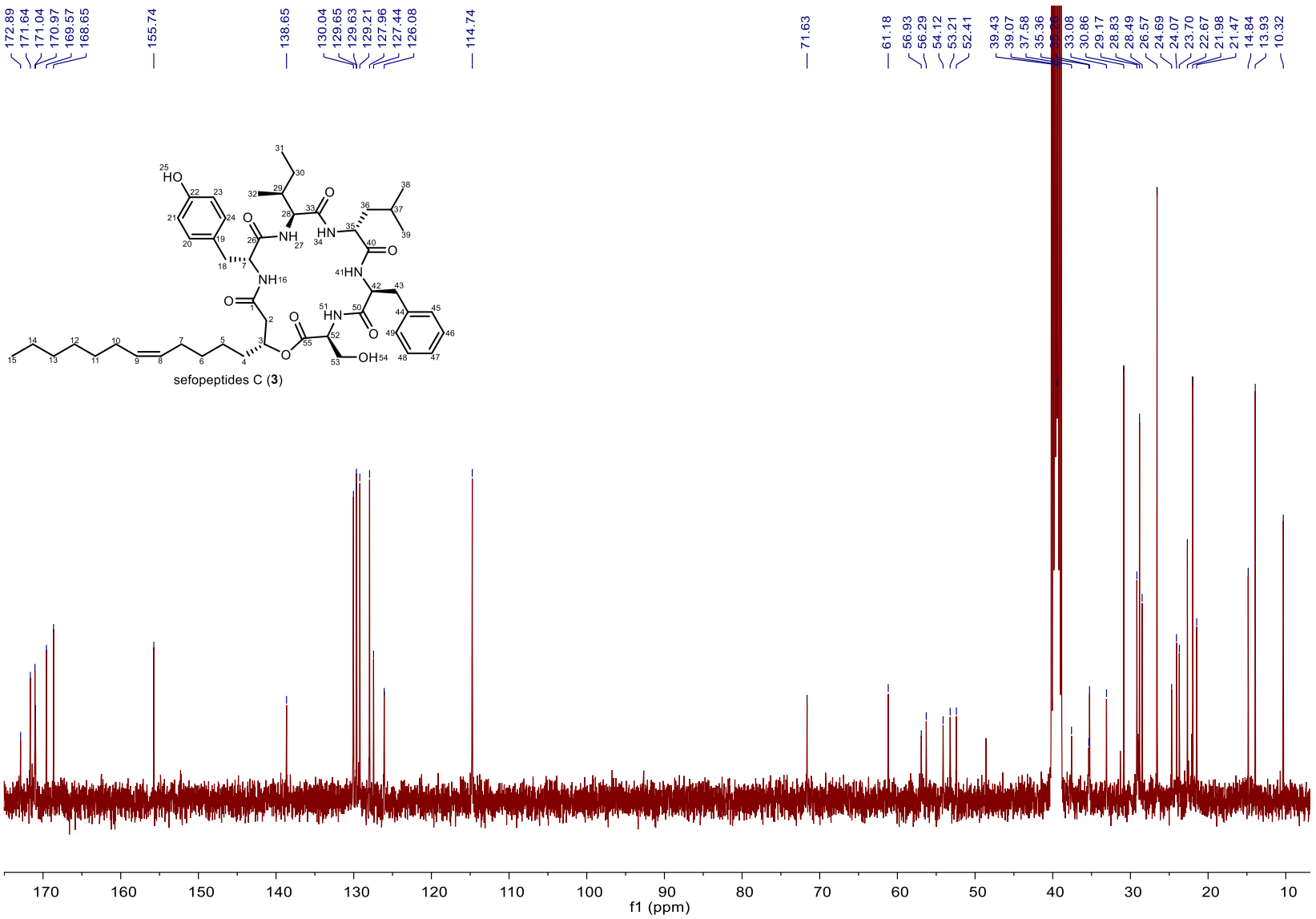


Fig. S23 ^{13}C NMR (100 MHz, DMSO-d_6) spectrum of sefopeptides C (**3**).

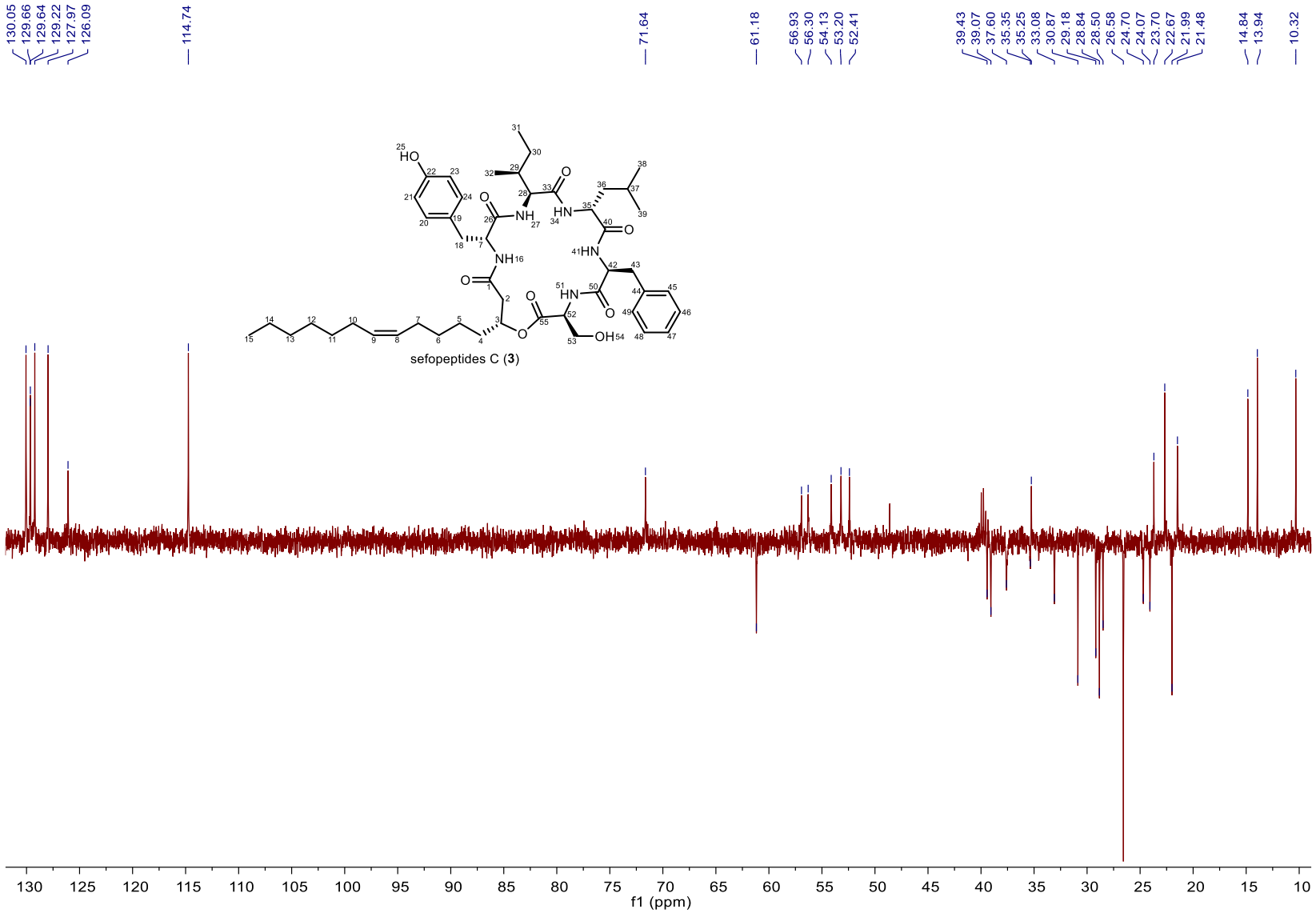


Fig. S24 DEPT-135 (100 MHz, DMSO-d_6) spectrum of sefopeptides C (**3**).

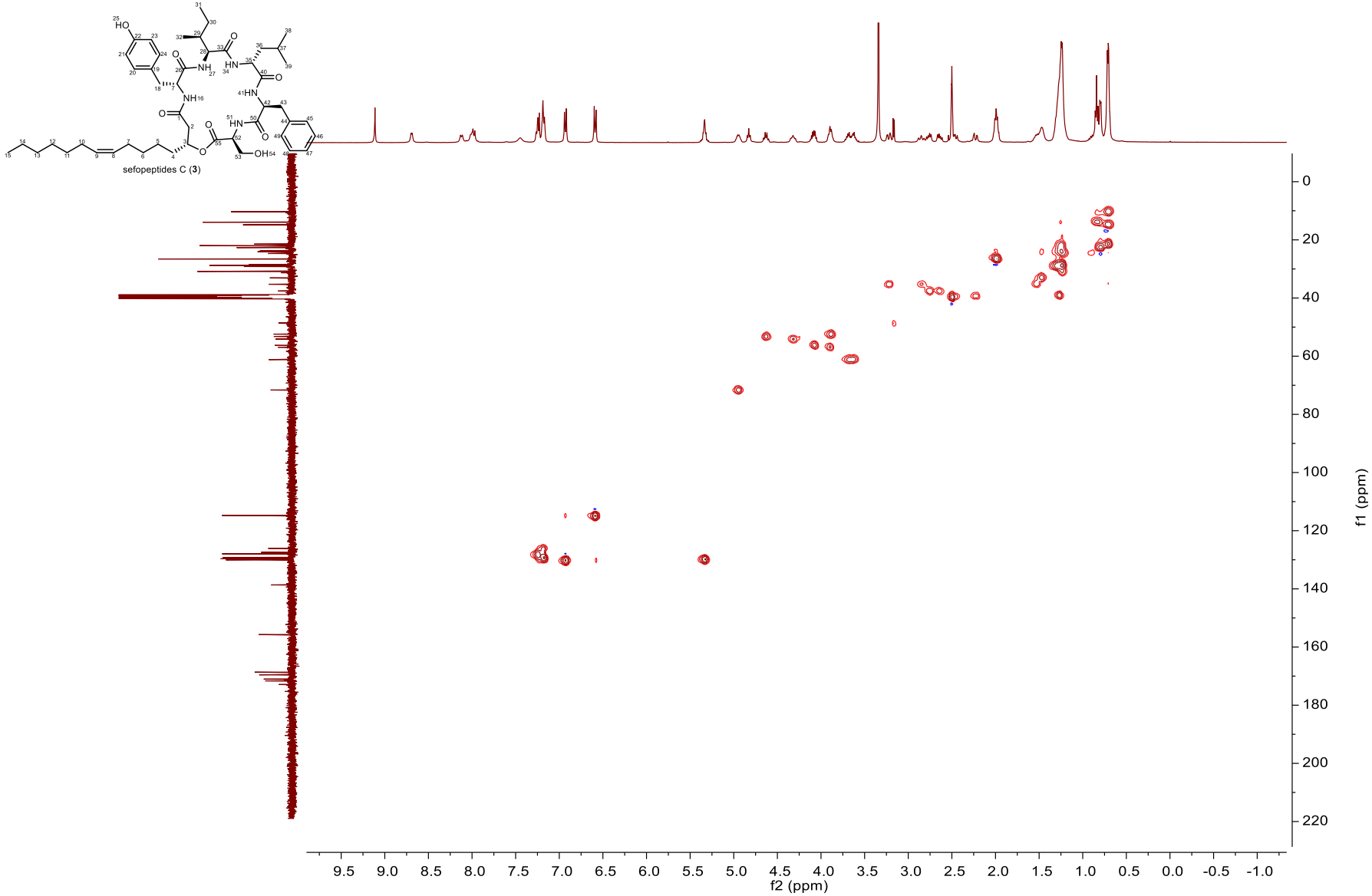


Fig. S25 HSQC (DMSO-d_6) spectrum of sefopeptides C (**3**).

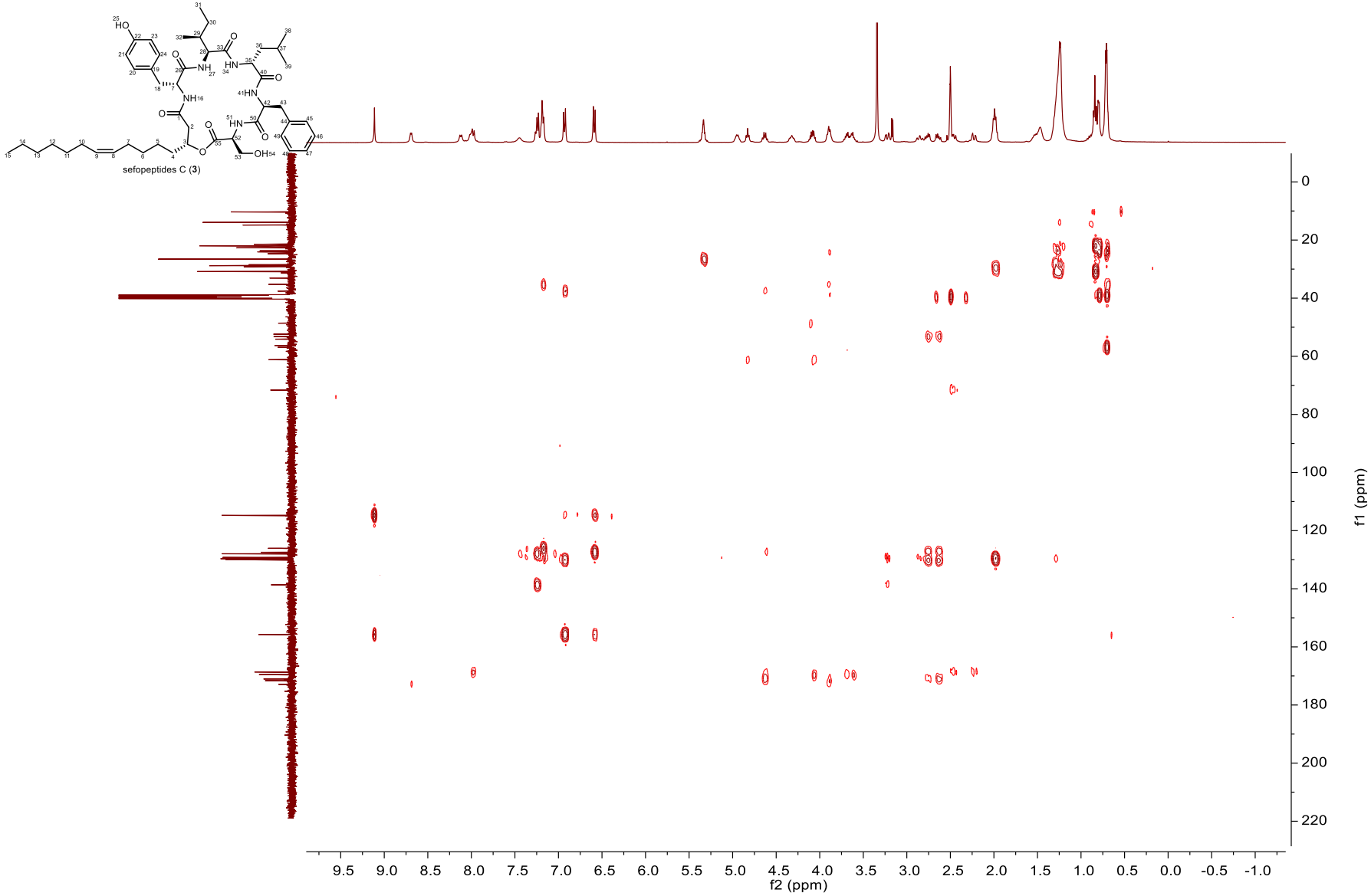


Fig. S26 ^{1}H-^{13}C HMBC (DMSO-d_6) spectrum of sefopeptides C (**3**).

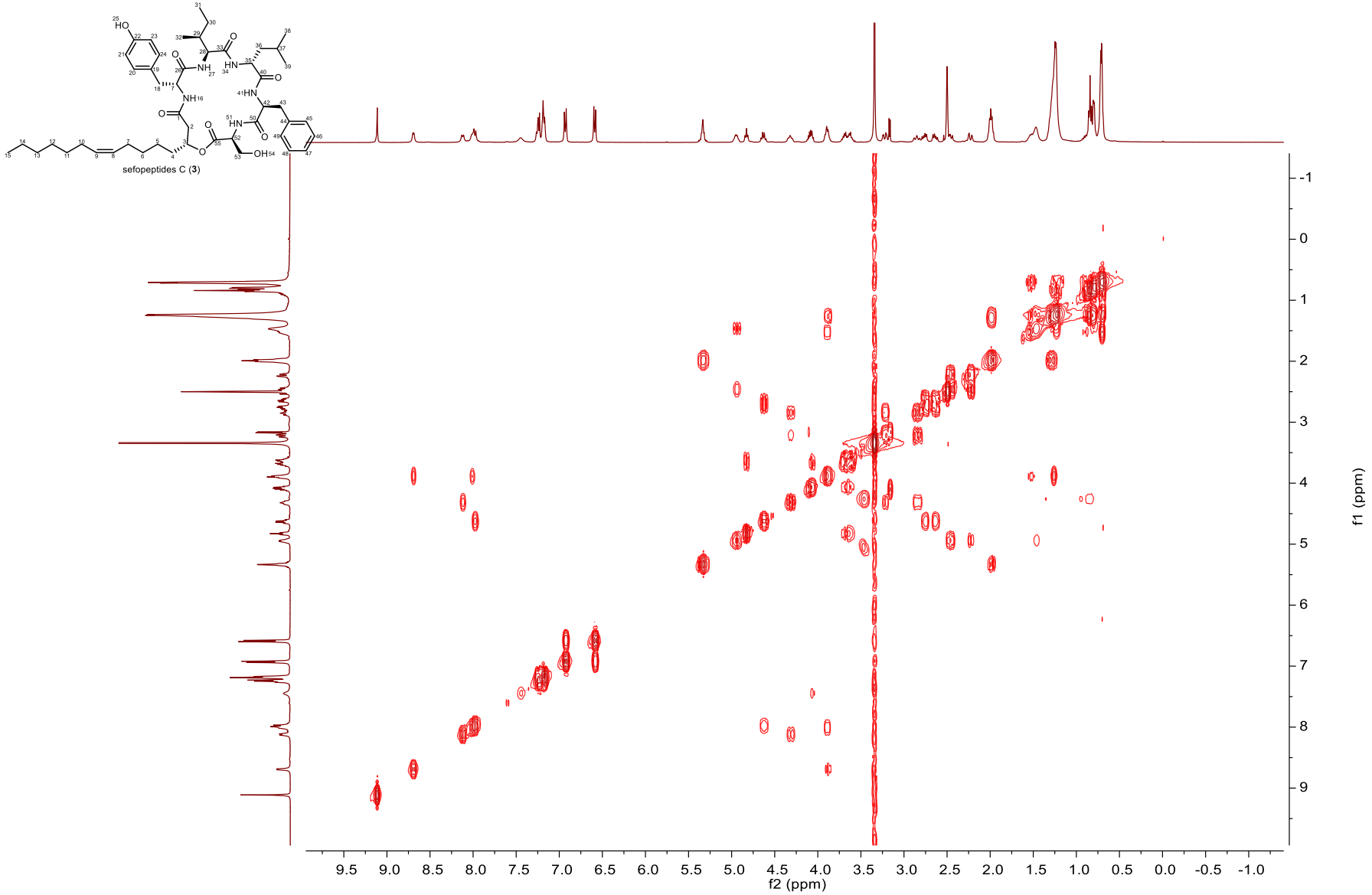


Fig. S27 ^{1}H-^{1}H COSY (DMSO-d_6) spectrum of sefopeptides C (**3**).

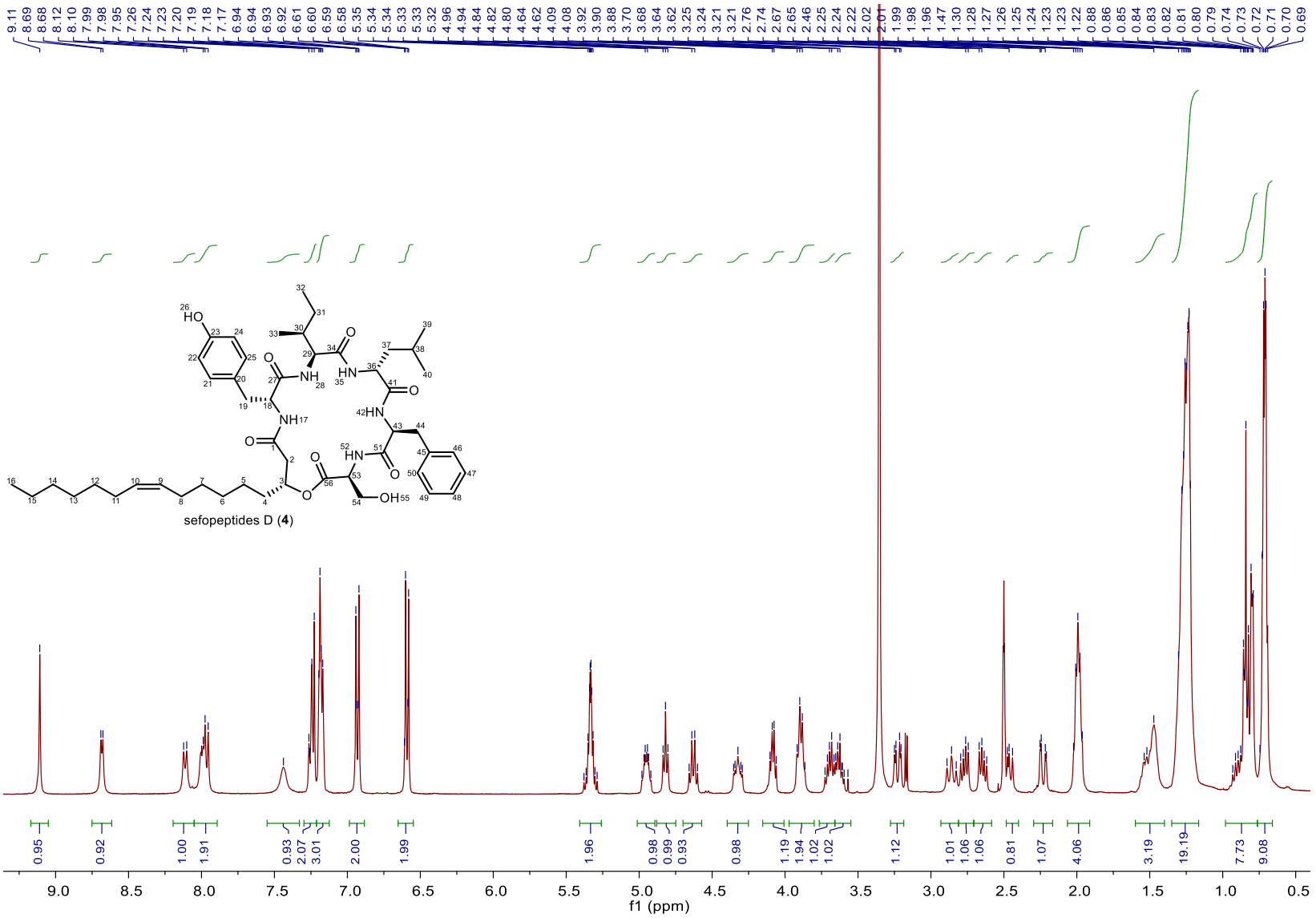


Fig. S28 ^{1}H NMR (400 MHz, DMSO-d_6) spectrum of sefopeptides D (**4**).

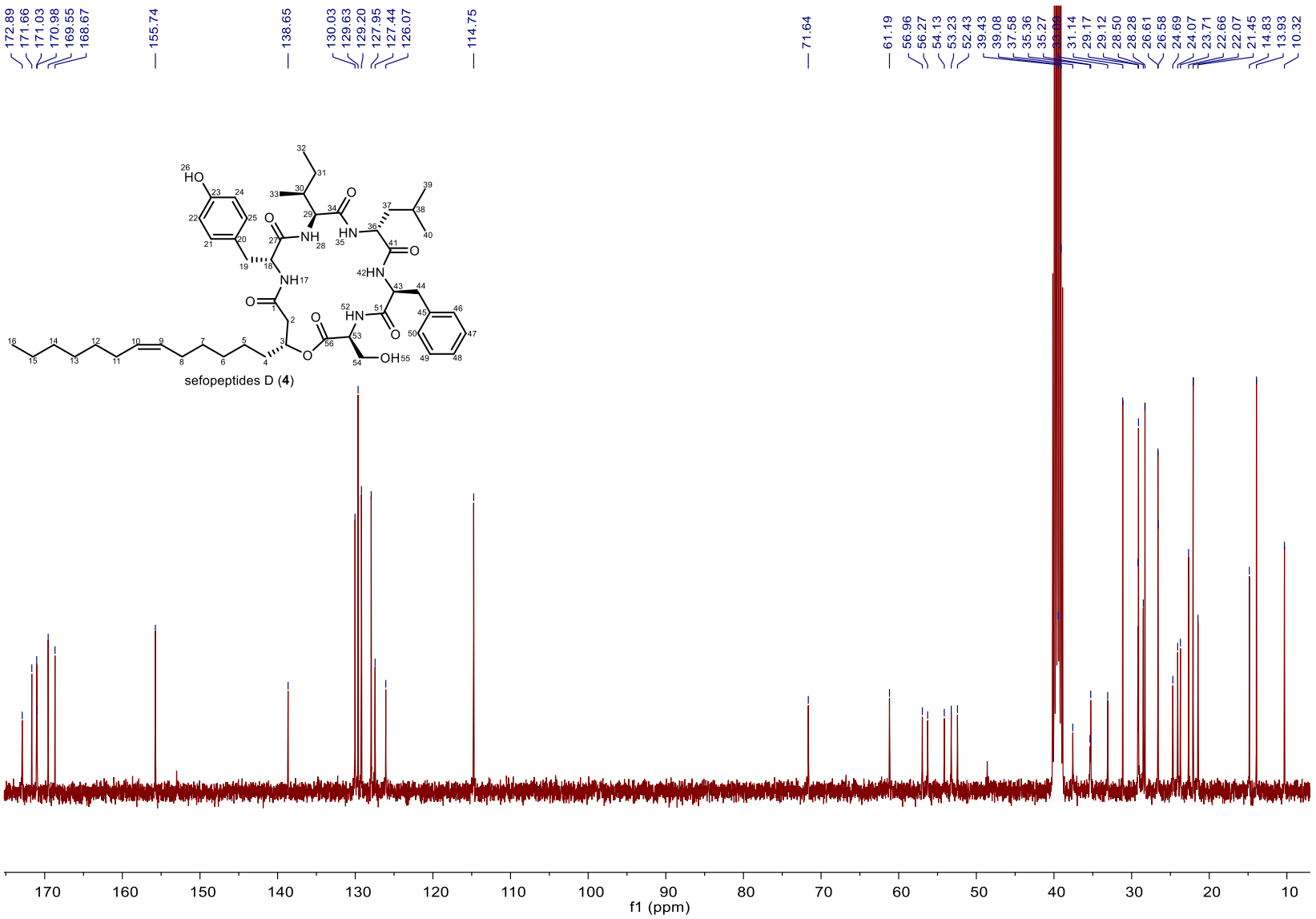


Fig. S29 ^{13}C NMR (100 MHz, DMSO-d_6) spectrum of sefopeptides D (**4**).

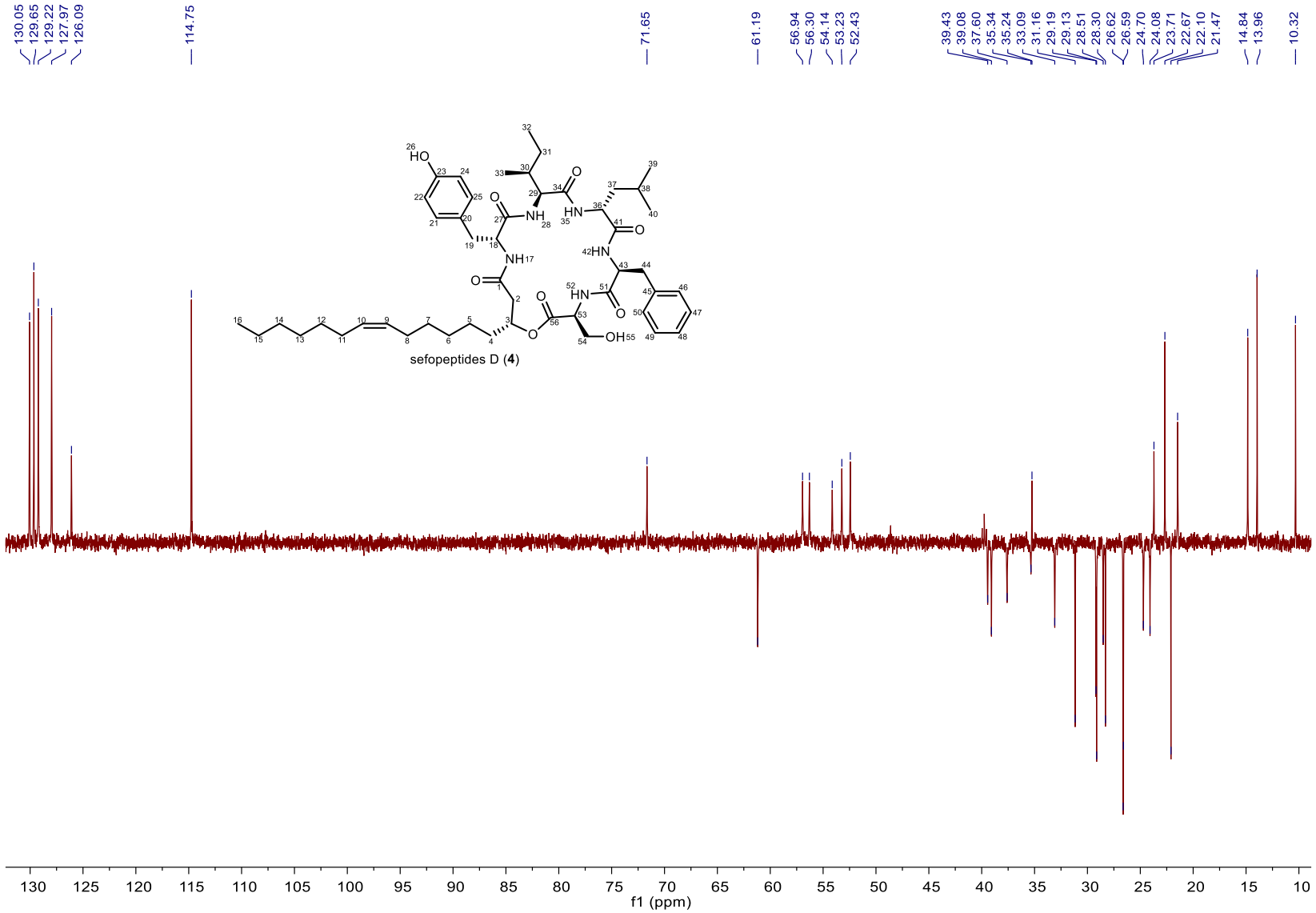


Fig. S30 DEPT-135 (100 MHz, DMSO-d_6) spectrum of sefopeptides D (**4**).

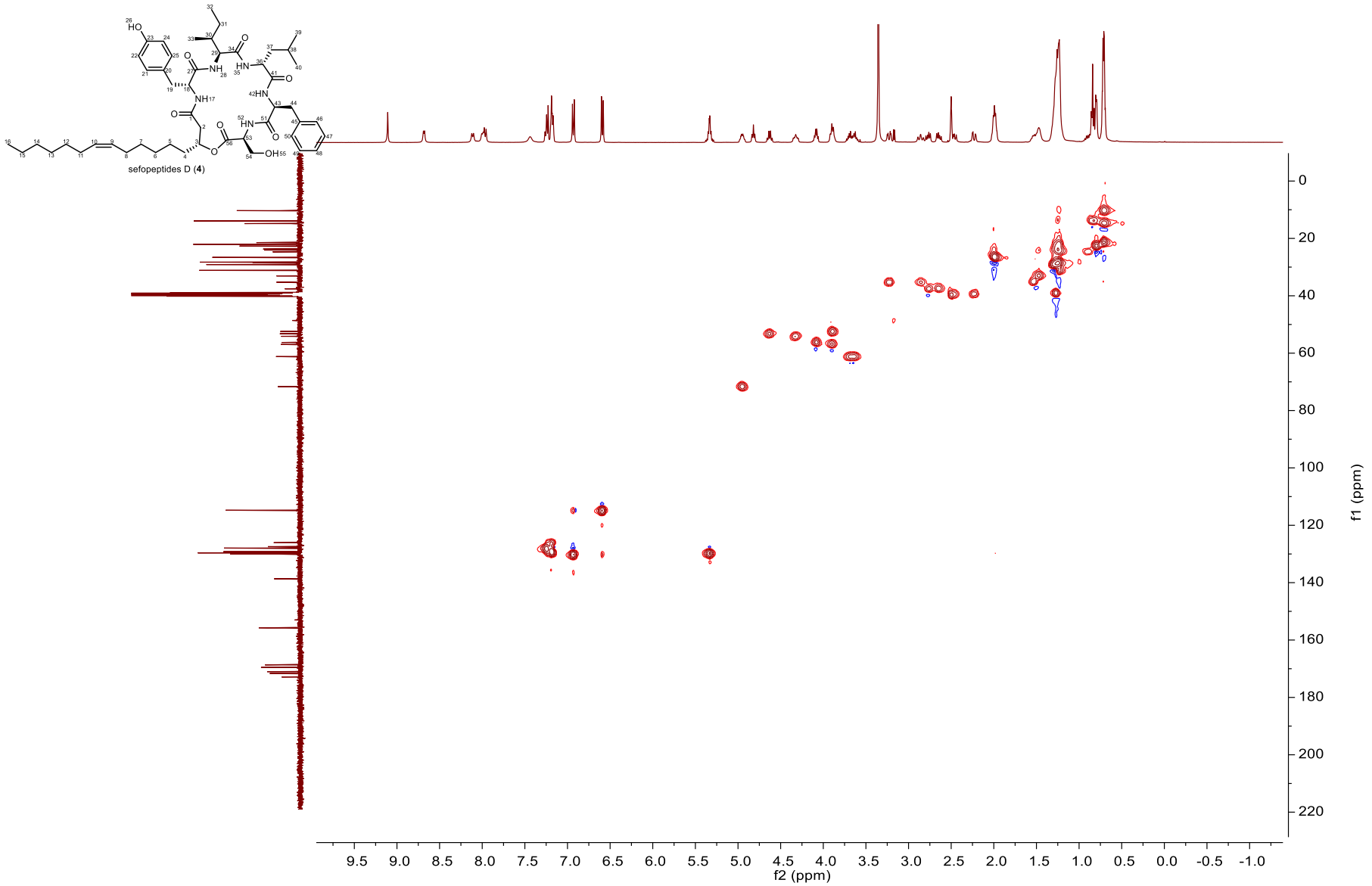


Fig. S31 HSQC (DMSO-d_6) spectrum of sefopeptides D (**4**).

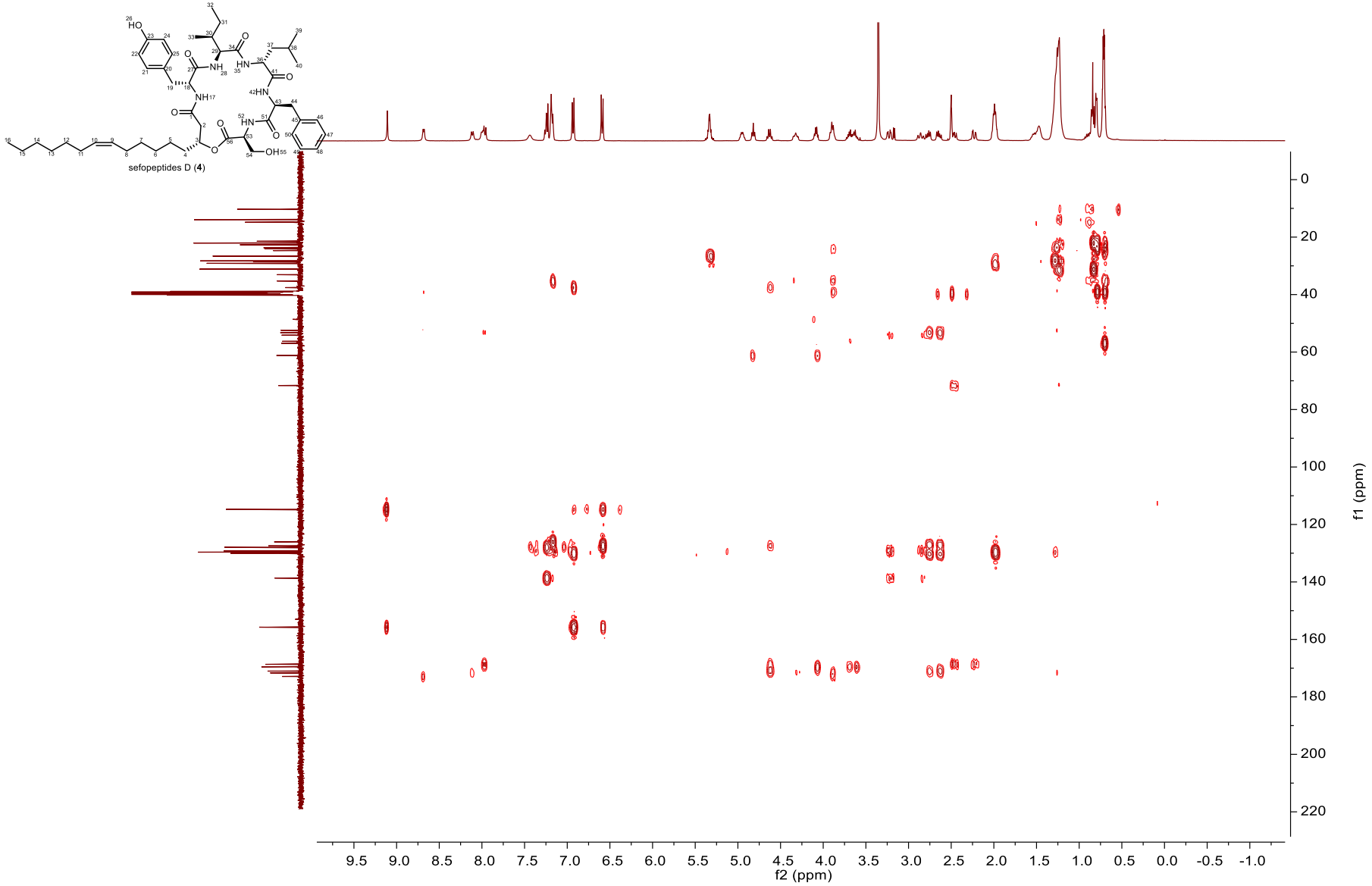


Fig. S32 ^{1}H-^{13}C HMBC (DMSO-d_6) spectrum of sefopeptides D (**4**).

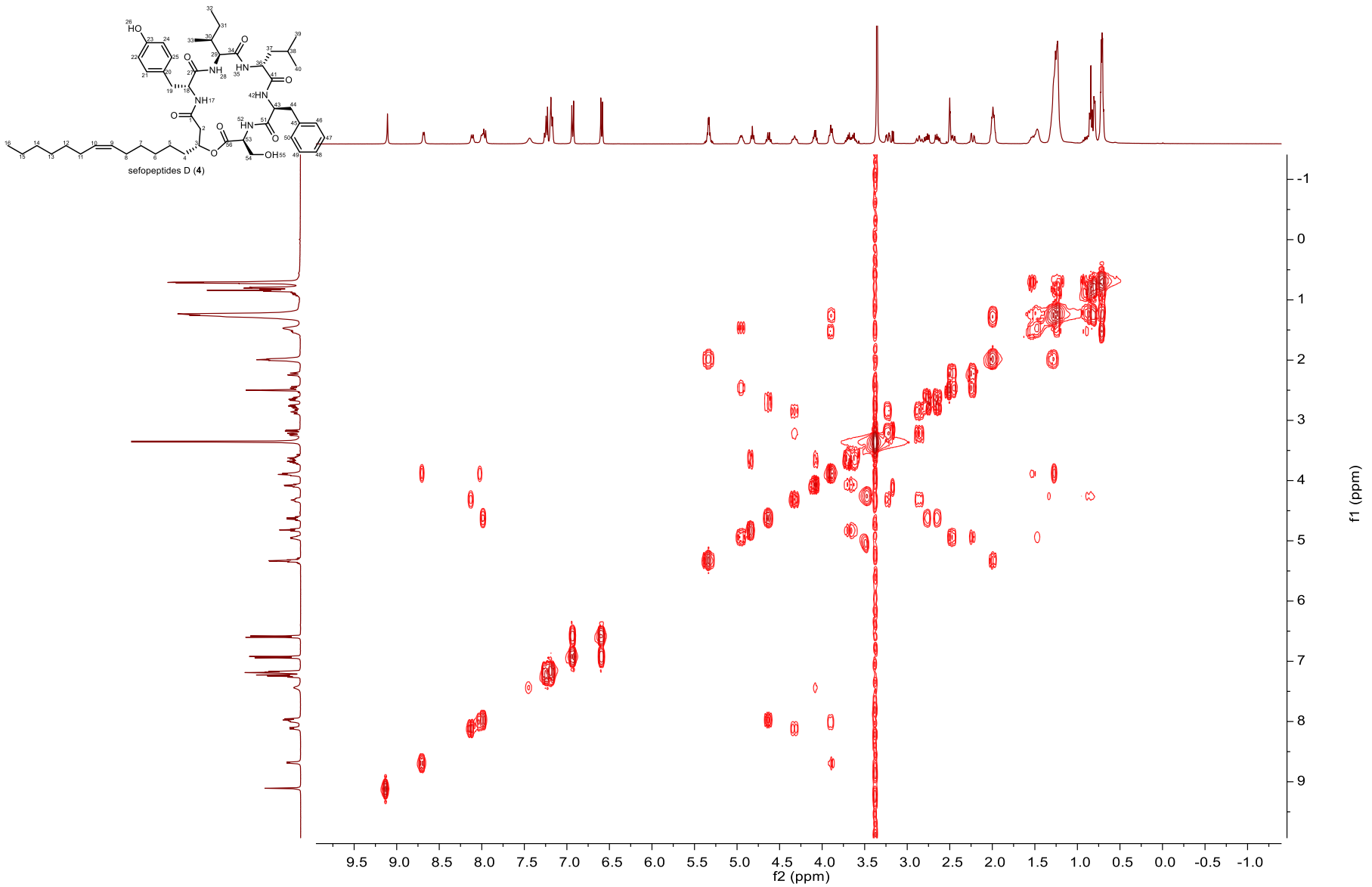


Fig. S33 ^{1}H-^{1}H COSY (DMSO-d_6) spectrum of sefopeptides D (**4**).

1. B. A. Pfeifer, S. J. Admiraal, H. Gramajo, D. E. Cane and C. Khosla, *Science*, 2001, **291**, 1790-1792.
2. K. Blin, S. Shaw, H. E. Augustijn, Z. L. Reitz, F. Biermann, M. Alanjary, A. Fetter, B. R. Terlouw, W. W. Metcalf, E. J. N. Helfrich, G. P. van Wezel, M. H. Medema and T. Weber, *Nucleic Acids Res.*, 2023, **51**, W46-W50.
3. R. Bhushan and H. Bruckner, *Amino Acids*, 2004, **27**, 231-247.
4. J. M. Manning, *J. Am. Chem. Soc.*, 1970, **92**, 7449-7454.