Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry.

Genome mining of cryptic tetronate natural products from a PKS-NRPS encoding gene cluster in *Trichoderma harzianum* t-22

Yiguang Zhu,‡[a,b] Junfeng Wang,‡[a,b] Pengyun Mou,[b] Yan Yan,[a,b] Mengbin Chen,[a] Yi Tang*[a]

[a]Department of Chemical and Biomolecular Engineering, University of California Los Angeles, CA 90095, USA. E-mail: yitang@ucla.edu

[b]Key Laboratory of Tropical Marine Bio-resources and Ecology, Guangdong Key Laboratory of Marine Materia Medica, South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou 510301, China.

‡These authors contribute equally to this work.

Table of Contents

Supplemental Materials and Methods

General DNA manipulation techniques

All DNA manipulations in this study were conducted according to manufacturers' protocols. DNA restriction enzymes were used as recommended by the manufacturer (New England Biolabs, NEB). PCR was performed according to recommended protocol using Q5 High-Fidelity DNA polymerase (NEB). Genomic DNA extraction was carried out following the instructions of Zymo Quick-DNA Fungal/Bacterial Miniprep Kit. Plasmids were confirmed by restriction enzyme digestion analysis and sequencing.

The RNA extractions were performed using RiboPureTM Yeast RNA Isolation Kit (Ambion) following the manufacturer's instructions. Residual genomic DNA in the extracts was digested by DNase I (2 U/μL) (Invitrogen) at 37°C for 4 hours. SuperScript III First-Strand Synthesis System (Invitrogen) was used for cDNA synthesis with Oligo-dT primers following instructions from the user manual.

Sequence analysis

The *orfs* were determined using the 2ndfind program(http://biosyn.nih.go.jp/2ndfind/). Protein sequences were compared with BLAST programs (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

A. nidulans protoplast preparation and transformation

A. nidulans A1145 ΔSTΔEM was grown on CD agar plates containing supplements (10 mM uridine, 5mM uracil, 0.5 μg/mL pyridoxine HCl and 2.5 μg/mL riboflavin) at 30°C for 5 days. Fresh spores were inoculated into 50 mL liquid CD media containing supplements in 250 mL flask and germinated at 30°C, 250 rpm for approximately 16 h. Mycelia were harvested by centrifugation at 3,500 rpm for 10 min, and washed with 10 mL osmotic buffer (1.2 M $MgSO_4$, 10 mM sodium phosphate, pH 5.8). Then the mycelia were transferred into 30 mL of osmotic buffer containing 100 mg lysing enzymes from *Trichoderma* and 60 mg Yatalase in a 250 mL flask and were shaken at 80 rpm for 4 h at 30°C. Cells were collected

in a 30 mL Corex tube and overlaid gently with 10 mL of trapping buffer (0.6 M sorbitol, 0.1 M Tris-HCI, pH 7.0). After centrifugation at 3,500 rpm for 15 min at 4°C, protoplasts were collected from the interface of the two buffers. The protoplasts were then transferred to a sterile 15 mL falcon tube and washed by 10 mL STC buffer (1.2 M sorbitol, 10 mM$CaCl_2$, 10 mM Tris-HCl, pH 7.5). The protoplasts were then resuspended in 1 mL STC buffer and 60 µL aliquots of the protoplasts were stored in -80°C for transformation.

The plasmids extracted from *E. coli* were added to 60 µL *A. nidulans* A1145 protoplast suspension and the mixture was incubated on ice for 60 min. After that, 600 µL of PEG solution (60% PEG, 50 mM$CaCl_2$ and 50 mM Tris-HCl, pH 7.5) was added to the protoplast mixture, followed by additional incubation at room temperature for 20 min. The mixture was spread on the regeneration medium (CD solid medium with 1.2 M sorbitol and appropriate supplements including 10 mM uridine, 5 mM uracil and/or 0.5 µg/mL pyridoxine HCl and/or 2.5 µg/mL riboflavin depending on the plasmids being transformed) and incubated at 30°C for 2-3 days until single colony appear.

Table S1. Deduced functions of individual *orfs* in the *thn* gene cluster

Gene name	Size (aa)	Proposed protein function	Homolog (identity)	Homolog Organism
thaA	4044	PKS-NRPS synthetase	CcsA (40%)	*Aspergillus clavatus* NRRL 1
thaB	563	Transporter	GsfJ (49%)	*Penicillium aethiopicum*
thaC	321	2-oxoglutarate-dependent dioxygenase	TropC (38%)	*Talaromyces stipitatus* ATCC 10500
thaD	406	FAD-dependent monooxygenase	CctM (27%)	*Trichophyton benhamiae* CBS 112371
thaE	369	Trans-enoyl reductase	FSL5 (44%)	*Fusarium graminearum* PH-1

Table S2. ^{1}H and ^{13}C NMR Data for **1**-**3** (500, 125 MHz, CD_3OD, TMS, δ ppm).

c	**1**		**2**		**3**	
	δ_C	δ_H (J in Hz)	δ_C	δ_H (J in Hz)	δ_C	δ_H (J in Hz)
1	n.d.[a]		n.d.[a]		n.d.[a]	
2	98.9[b], C		99.1[b], C		95.9[b], C	
3	193.6, C		193.5, C		196.1, C	
4	78.2, CH	4.97, m	78.1, CH	5.00, m	78.3, CH	4.67, m
5	36.9, CH_2	2.99, dd (17.0, 3.7); 2.74, dd (17.0, 7.1)	36.7, CH_2	3.00, dd (17.1, 3.5); 2.77, dd (17.1, 6.8)	38.5, CH_2	2.90, dd (16.5, 3.3); 2.46, dd (16.5, 9.0)
6	172.7, C		172.6, C		174.5, C	
1’	191.6, C		191.1, C		196.0, C	
2’	134.1, C		133.7, C		137.0, C	
3’	147.3, CH	6.84, t (6.8)	148.0, CH	6.91, t (6.8)	138.6, CH	6.12, t (6.6)
4’	30.2, CH_2	2.39, m	30.2, CH_2	2.40, m	29.5, CH_2	2.24, m; 2.21, m
5’	32.3, CH_2	2.26, m	32.3, CH_2	2.30, m; 2.25, m	32.6, CH_2	2.20, m; 2.03, m
6’	131.6, CH	5.62, m	131.9, CH	5.64, m	133.3, CH	5.65, m
7’	132.9, CH	6.06, dd (14.9, 10.3)	132.8, CH	6.04, m	132.1, CH	6.01, m
8’	130.1, CH	5.97, dd (14.9, 10.3)	130.3, CH	6.00, m	130.5, CH	6.04, m
9’	139.5, CH	5.44, dd (15.0, 7.9)	139.1, CH	5.45, dd (14.8, 8.1)	137.0, CH	5.51, m
10’	39.8, CH	2.03, m	34.6, CH	2.27, m	34.8, CH	2.64, m
11'	30.9, CH_2	1.32, m	40.7, CH_2	1.52, m	42.7, CH_2	2.26, m
12'	12.2, CH_3	0.86, t (7.4)	61.1, CH_2	3.54, m	176.5, C	
13'	12.2, CH_3	1.88, s	12.1, CH_3	1.89, s	12.9, CH_3	1.80, s
14'	20.7, CH_3	0.98, d (6.8)	21.1, CH_3	1.01, d (6.8)	20.5, CH_3	1.05, d (6.8)

[a] Signals not detected from HMBC correlations; [b]Signals acquired from ^{13}C NMR spectrum.

Table S3. ^{1}H and ^{13}C NMR Data for **4**-**6** (500, 125 MHz, TMS, δ ppm).

No.	**4**[a]		**5**[b]		**6**[c]	
	δ_C	δ_H (J in Hz)	δ_C	δ_H (J in Hz)	δ_C	δ_H (J in Hz)
1	n.d.[d]		n.d.[d]		n.d.[d]	
2	93.6 [e], C		n.d.[d]		96.9 [e], C	
3	n.d.[d]		n.d.[d]		182.1, C	
4	154.5, C		148.1, C		155.6, C	
5	85.3, CH_2	4.70, s;	96.5, CH_2	5.44, d (2.5);	89.0, CH_2	5.04, d (1.6);
		4.46, s		5.26, s		7.71, d, (1.5)
1’	191.7, C		192.7, C		195.7, C	
2’	136.8, C		150.2, C		138.0, C	
3’	134.7, CH	5.90, m	132.8, CH	overlap	139.3, CH	6.15, t (6.0)
4’	28.0, CH_2	2.14, m	29.7, CH_2	2.43, m	29.7, CH_2	2.28, m
5’	31.4, CH_2	2.14, m; 1.23, m	31.2, CH_2	2.30, m	32.7, CH_2	2.21, m
6’	131.6, CH	5.63, m	130.9, CH	5.62, m	133.3, CH	5.67, m
7’	130.7, CH	6.02, m	131.5, CH	6.04, m	132.1, CH	6.03, m
8’	128.7, CH	5.95, m	129.0, CH	6.01, m	130.6, CH	6.03, m
9’	138.0, CH	5.45, dd (15.1, 7.4)	138.5, CH	5.49, dd (14.2, 8.1)	137.0, CH	5.51, m
10’	37.7, CH	2.01, m	34.0, CH	2.31, m	34.8, CH	2.64, m
11'	29.3, CH_2	1.23, m	39.9, CH_2	1.57, m	42.6, CH_2	2.25, m
12'	11.7, CH_3	0.80, t (6.3)	61.4, CH_2	3.65, m	176.4, C	
13'	12.9, CH_3	1.69, s	12.2, CH_3	1.92, s	12.8, CH_3	1.81, s
14'	20.0, CH_3	0.94, d (6.5)	20.9, CH_3	1.02, d (6.7)	20.5, CH_3	1.06, d (6.8)

[a]Measured in DMSO-d_6; [b]Measured in $CDCl_3$; [c]Measured in CD_3OD; [d]Signals not detected from HMBC correlations; [e]Signals acquired from ^{13}C NMR spectrum.

Table S4. Strains and plasmids used and generated in this study

Strains/Plasmids	**Characteristics**	**Sources**
E. coli		
TOP10	Host strain for cloning	Invitrogen
BL21(DE3)	Host strain for protein expression	Novagen
Fungi		
S. cerevisiae BJ5464-NpgA	Host for *in vivo* homologous recombination to construct *A. nidulans* plasmids	[1]
T. harzianum t-22	Wild type	[2]
A. nidulans A1145 ΔSTΔEM	Host strain for heterologous expression	[3]
ZYG001	*A. nidulans* A1145 containing plasmid pYTU, pYTP and pYTR	This study
ZYG002	*A. nidulans* A1145 containing plasmid pZYG001	This study
ZYG003	*A. nidulans* A1145 containing plasmid pZYG001 and pZYG002	This study
ZYG004	*A. nidulans* A1145 containing plasmid pZYG001 and pZYG003	This study
ZYG005	*A. nidulans* A1145 containing plasmid pZYG001, pZYG002 and pZYG003	This study
ZYG006	*A. nidulans* A1145 containing plasmid pZYG001, pZYG002 and pZYG004	This study
ZYG007	*A. nidulans* A1145 containing plasmid pZYG005	This study
Plasmids		

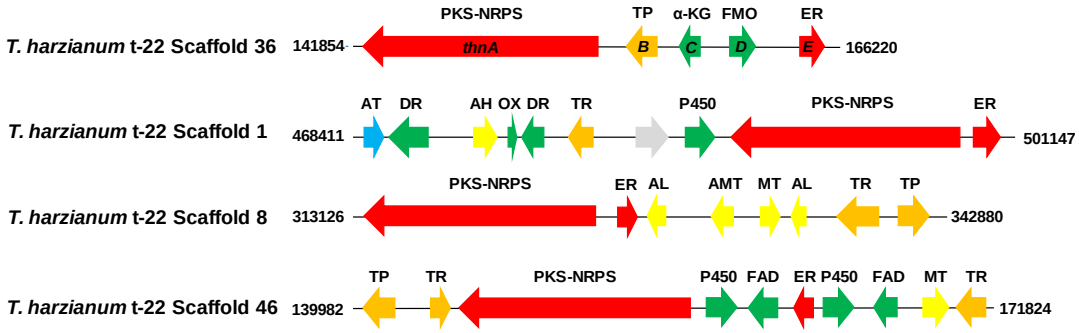
pET28b	Km^R, expression vector	Novagen
pYTU	Heterologous expression vector in *A. nidulans* A1145	[1]
pYTP	Heterologous expression vector in *A. nidulans* A1145	[1]
pYTR	Heterologous expression vector in *A. nidulans* A1145	[1]
pZYG001	pYTP containing glaA-*thnA* and gpdA-*thnE*	This study
pZYG002	pYTU containing glaA-*thnC*	This study
pZYG003	pYTR containing gpdA-*thnD*	This study
pZYG004	PYTR containing gpdA-*thnD* and glaA-*thnB*	This study
pZYG005	pYTP containing glaA-*thnA*	This study
pZYG105	pET28b containing 960bp *thnC* fragment	This study

Table S5. Primers used in this study

Primers	Plasmids	Sequences
TH22-22AE-1	pZYG001	CCTCGCGGGTGTTCTTGACGATGGCATCCTCCTGATCTTCCGAACTGGTCGTACCTGGCG
TH22-22AE-2		TGCGATCGGCTCCAGGTTTTGGGATCCCATTGCTGAGGTGTAATGATGCTGGGGATGAAG
TH22-22AE-3		CTTCATCCCCAGCATCATTACACCTCAGCAATGGGATCCCAAAACCTGGAGCCGATCGCA
TH22-22AE-4		GGCCGCATAGTTGGCCTGGCCGAGGTTACCAACAACGCACGCGGCTGACGAGAACAAGAC
TH22-22AE-5		GTCTTGTTCTCGTCAGCCGCGTGCGTTGTTGGTAACCTCGGCCAGGCCAACTATGCGGCC
TH22-22AE-6		CTCCCGTCACCCAAATCAATTCACCGGAGTCCGGTCTTGCGACAGTGTGACCAATCACAC
TH22-22AE-7		GTGTGATTGGTCACACTGTCGCAAGACCGGACTCCGGTGAATTGATTTGGGTGACGGGAG
TH22-22AE-8		AACTGCCCCATCGTAGTAGCGGACGACATTGTTTAGATGTGTCTATGTGGCGGGGTAATG
TH22-22AE-9		CATTACCCCGCCACATAGACAATCTAAACAATGTCGTCCGCTACTACGATGGGGCAGTT
TH22-22AE-10		AAGCTTGATATCGAATTCCTGCAGCCCGGGCGTTAAAAGGTGTCCAAGTAGAGCAACGGC
TH22-22C-1	pZYG002	CTTCATCCCCAGCATCATTACACCTCAGCAATGTCGGAAGAAGATATCTCGTTGCCCATC
TH22-22C-2		AGTGGAGGACATACCCGTAATTTTCTGGGCCGCAACTTGCGTGCCTAGGCATGTGATTGG
TH22-22D-1	pZYG003	ATTACCCCGCCACATAGACACATCTAAACAATGCACGTTCTTATTGCAGGAGCCGGGCT
TH22-22D-2		TAAAGGGTATCATCGAAAGGGAGTCATCCAGCTATGGGCCTAGGGCAATATTCTACAGTC
TH22-22BD-1	pZYG004	ATTACCCCGCCACATAGACACATCTAAACAATGCACGTTCTTATTGCAGGAGCCGGGCT
TH22-22BD-2		CGCCAGGTACGACCAGTTCGGAAGATCAGGGCTATGGGCCTAGGGCAATATTCTACAGTC
TH22-22BD-3		GACTGTAGAATATTGCCCTAGGCCCATAGCCCTGATCTTCCGAACTGGTCGTACCTGGCG
TH22-22BD-4		TCTCGTGGCACTGTAGTCTCCGCTCATTGCTGAGGTGTAATGATGCTGGGGATGAAG
TH22-22BD-5		CTTCATCCCCAGCATCATTACACCTCAGCAATGAGCGGAGACTACAGTGCCACGAGA
TH22-22BD-6		TAAAGGGTATCATCGAAAGGGAGTCATCCAGAAGCAATATCACGTAGCATGCATCTTC
TH22-22A-1	pZYG005	CCTCGCGGGTGTTCTTGACGATGGCATCCTCCTGATCTTCCGAACTGGTCGTACCTGGCG
TH22-22A-2		TGCGATCGGCTCCAGGTTTTGGGATCCCATTGCTGAGGTGTAATGATGCTGGGGATGAAG

TH22-22A-3		CTTCATCCCCAGCATCATTACACCTCAGCAATGGGATCCCAAAACCTGGAGCCGATCGCA
TH22-22A-4		GGCCGCATAGTTGGCCTGGCCGAGGTTACCAACAACGCACGCGGCTGACGAGAACAAGAC
TH22-22A-5		GTCTTGTTCTCGTCAGCCGCGTGCGTTGTTGGTAACCTCGGCCAGGCCAACTATGCGGCC
TH22-22A-6		GATGAGACCCAACAACCATGATACCAGGGGCCGGTCTTGCGACAGTGTGACCAATCAC
ThnC-EF	pZYG105	TTTTGTTTAACTTTAAGAAGGAGATATACCATGTCGGAAGAAGATATCTCGTTGCCCATC
ThnC-ER		GATCTCAGTGGTGGTGGTGGTGGTGGAGCTCTCCAACCTGAGCAAG

Figure S1. Selected PKS-NRPS gene clusters in *T. harzianum* t-22



PKS-NRPS: polyketide synthase and nonribosomal peptide synthetase; ER: trans enoyl reductase; TP: transporter; TR: transcriptional regulator; α-KG: α-ketoglutarate-dependent dioxygenase; FMO: flavin-dependent monooxygenase; DR: dehydrogenase, OX: oxygenase; AT: acetyltransferase; AH: alpha/beta hydrolase; AL: aldolase; AMT: aminotransferase; MT: methyltransferase.

Figure S2. *thn* cluster is highly conserved in several *Trichoderma* spp.

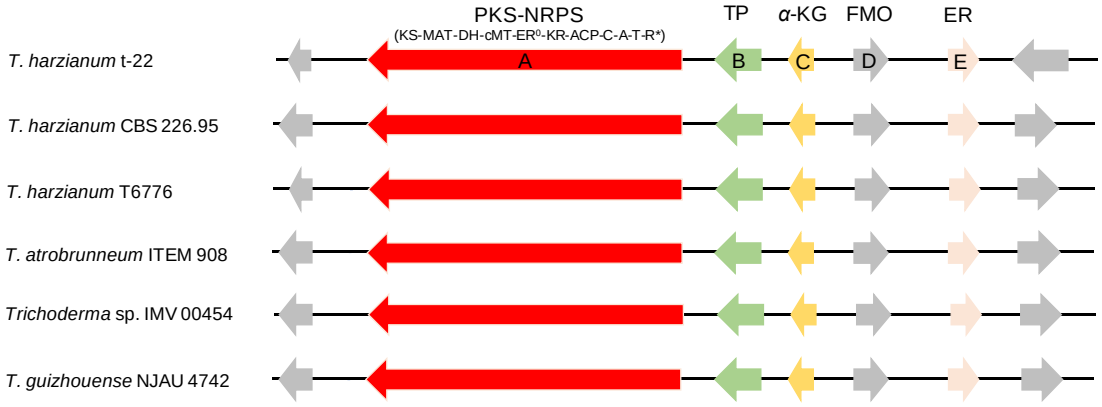
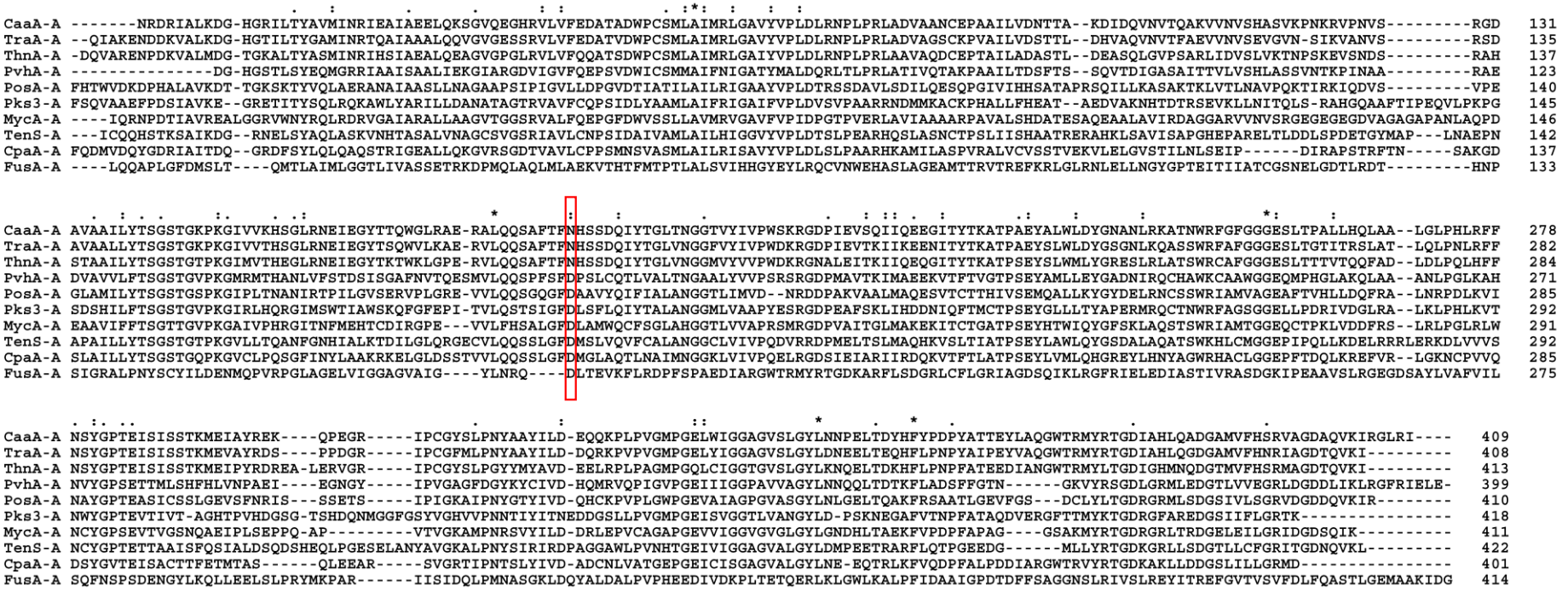
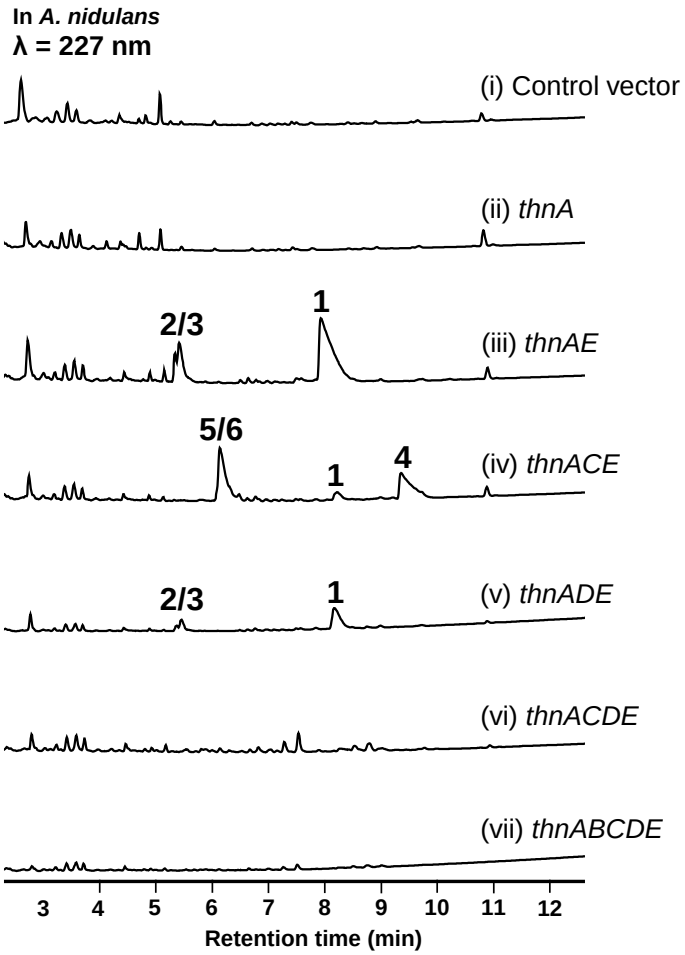


Figure S3. Amino acid alignment of ThnA-A with A domains of reported PKS-NRPS genes



Conserved amino acid residues for A domain recognition active sites were labeled in red box. CaaA (EHA18001), TraA (QBK15049), PvhA (AZZ09613), PsoA (ABS87601), Pks3 (AAS46233), MycA (G2Q9A5), TenS (EJP63694), CpaA (BAG82673), FusA (AFP73394).

Figure S4. HPLC analysis of extracts from *A. nidulans* transformant containing different combinations of *thn* genes.



In the presence of the FAD-dependent monooxygenase *thnD with thnACE*, no detectable metabolite is produced. Each transformant was cultured on CDST agar at 28°C for 3 days before extraction of metabolites.

Figure S5. The spectroscopic data of **1**

(A) The HRESIMS spectrum of **1**

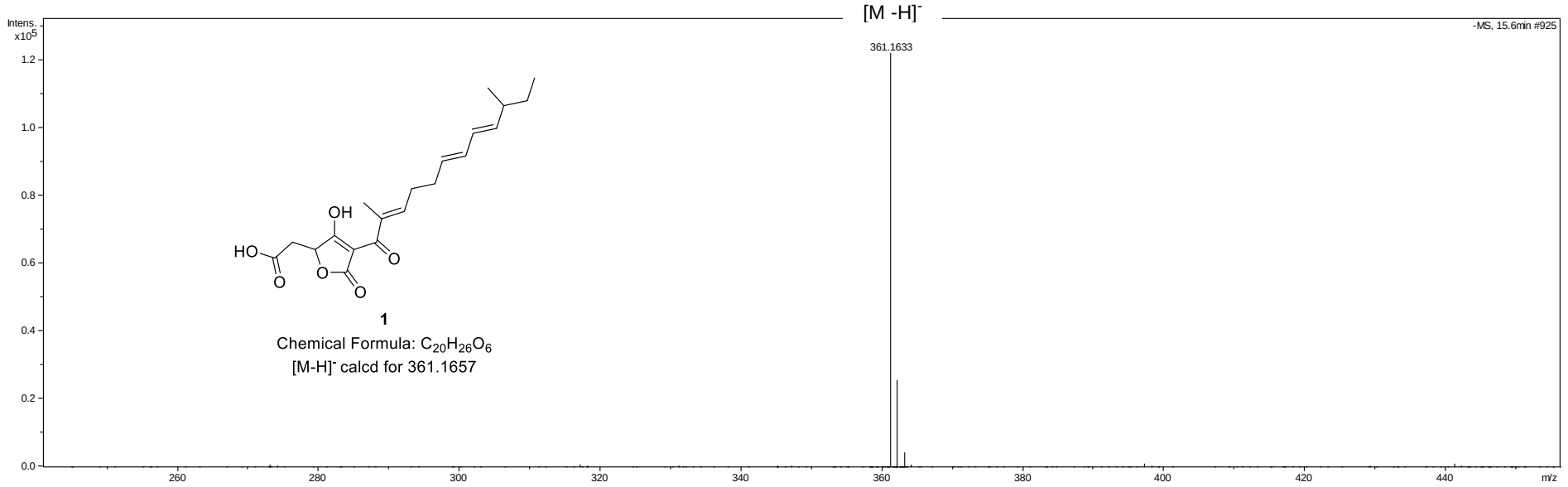


Figure S5. The spectroscopic data of **1**

(B) The ^{1}H-NMR spectrum of **1** (500 MHz for ^{1}H NMR in CD_3OD)

Figure S5. The spectroscopic data of **1**

(C) The ^{13}C NMR spectrum of compound **1** in CD_3OD (125 MHz)

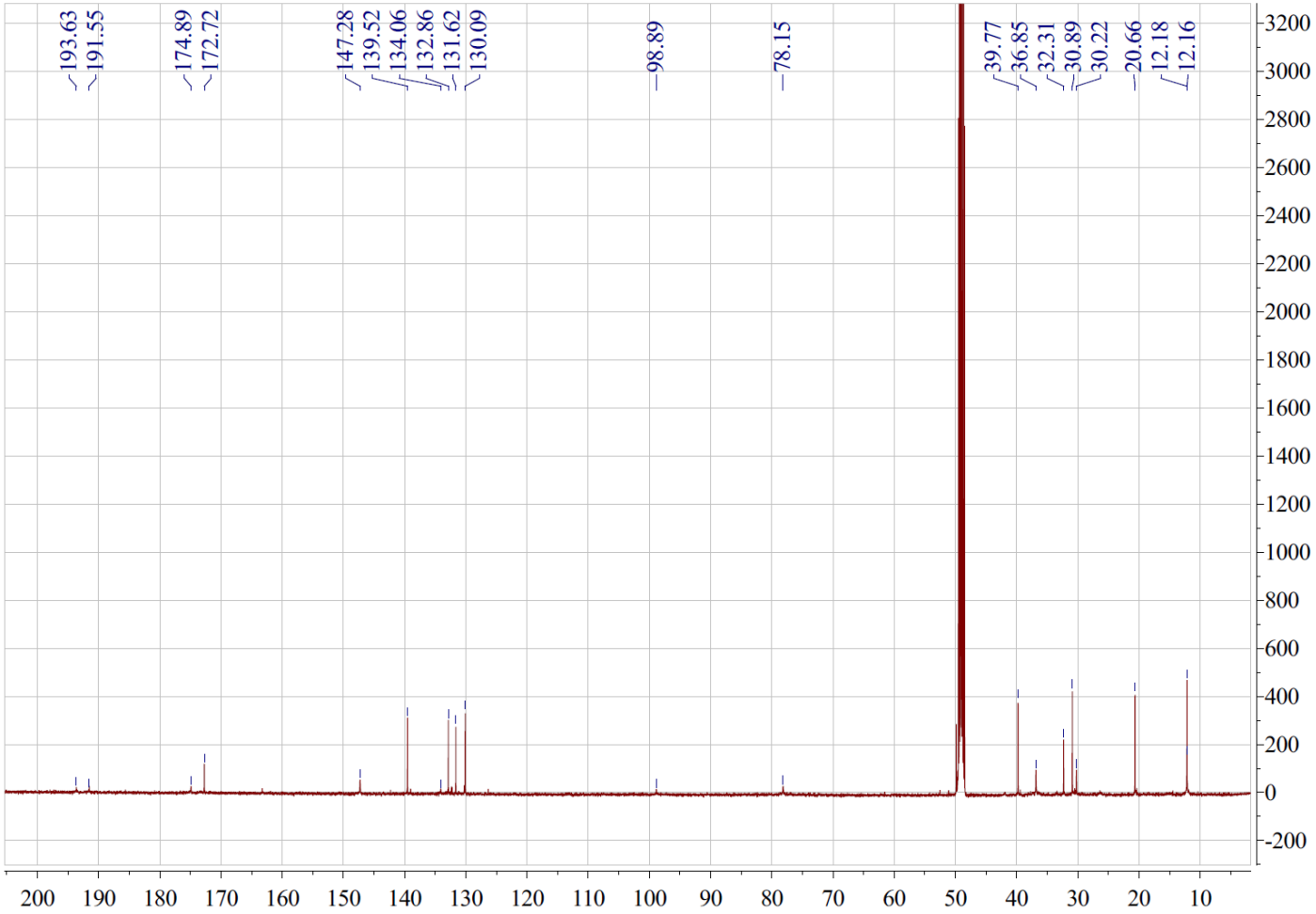


Figure S5. The spectroscopic data of **1**

(D) The HSQC spectrum of compound **1** in CD_3OD (125 MHz)

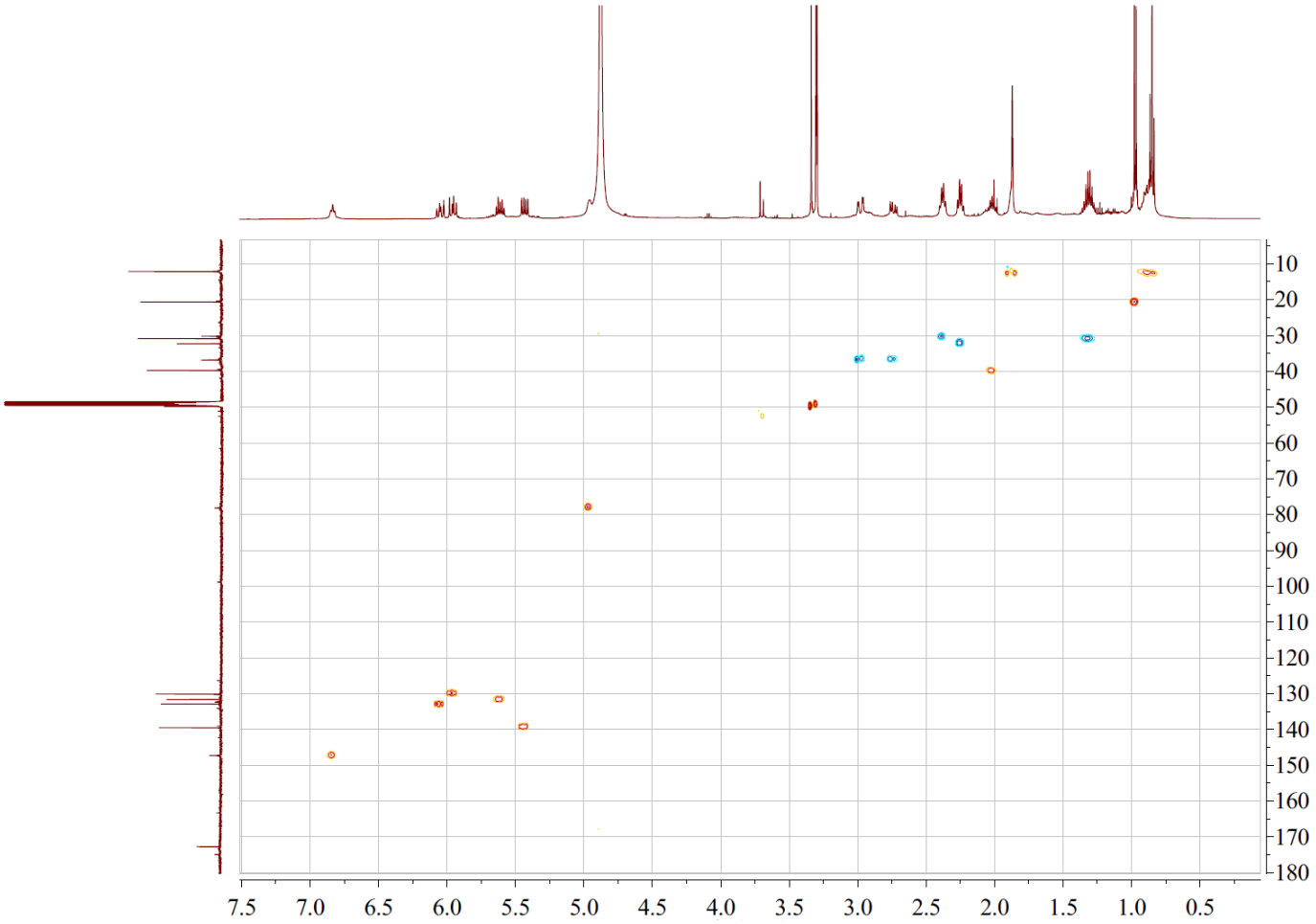


Figure S5. The spectroscopic data of **1**

(E) The HMBC spectrum of **1** (500 MHz for ^{1}H NMR in CD_3OD)

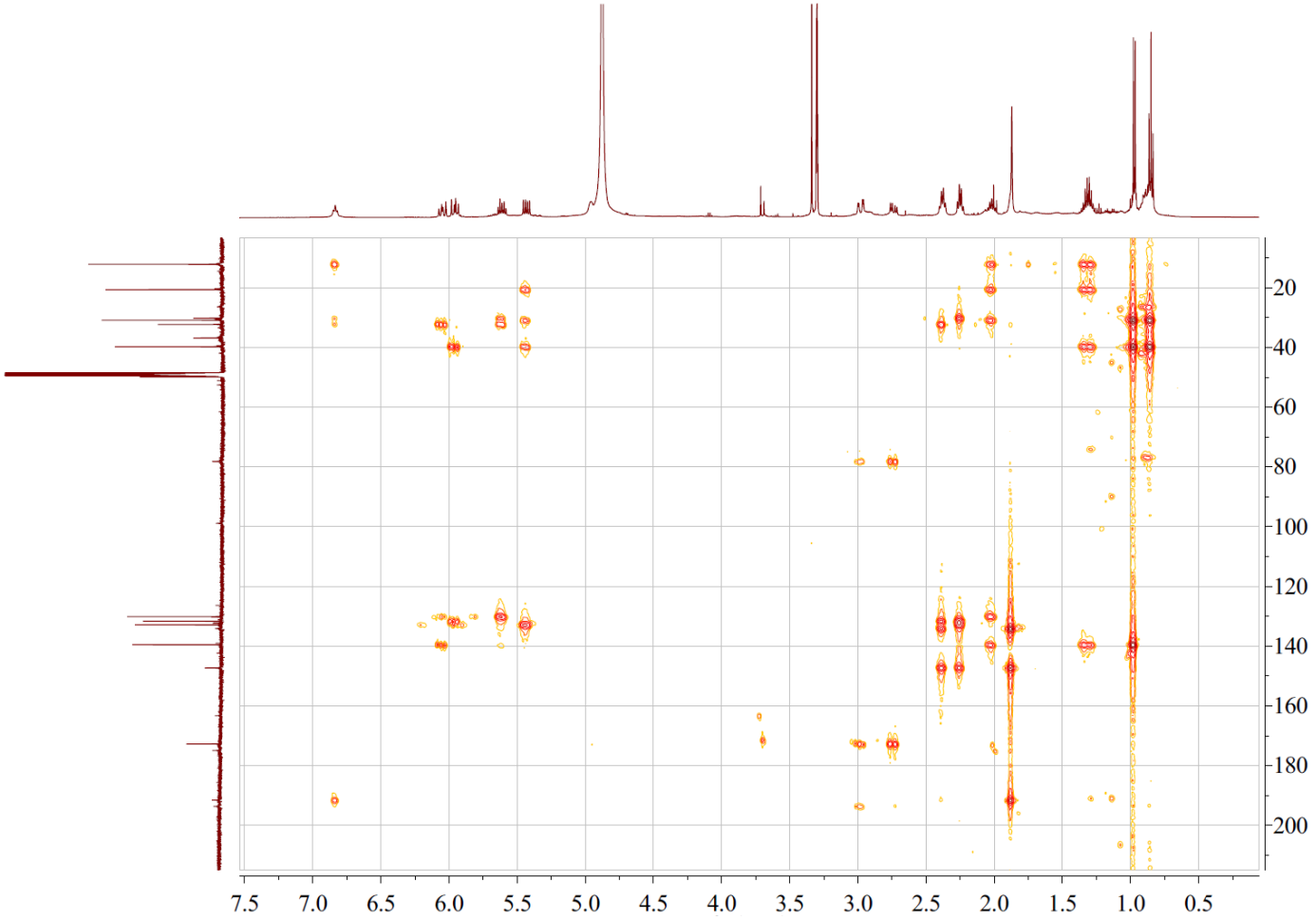


Figure S5. The spectroscopic data of **1**

(F) The ^{1}H-^{1}H COSY spectrum of **1** (500 MHz for ^{1}H NMR in CD_3OD)

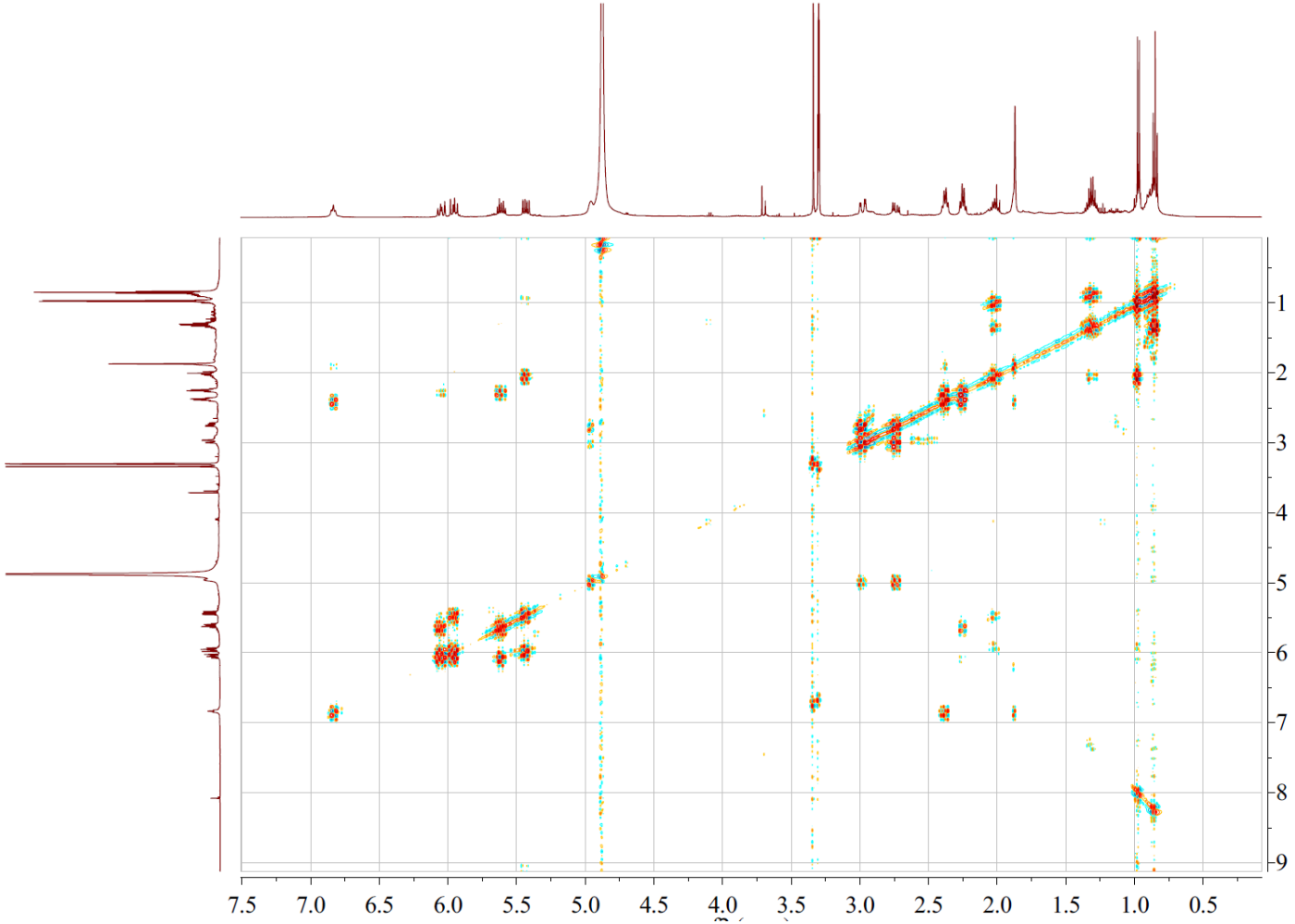


Figure S5. The spectroscopic data of **1**

(G) The NOESY spectrum of **1** (500 MHz for ^{1}H NMR in CD_3OD)

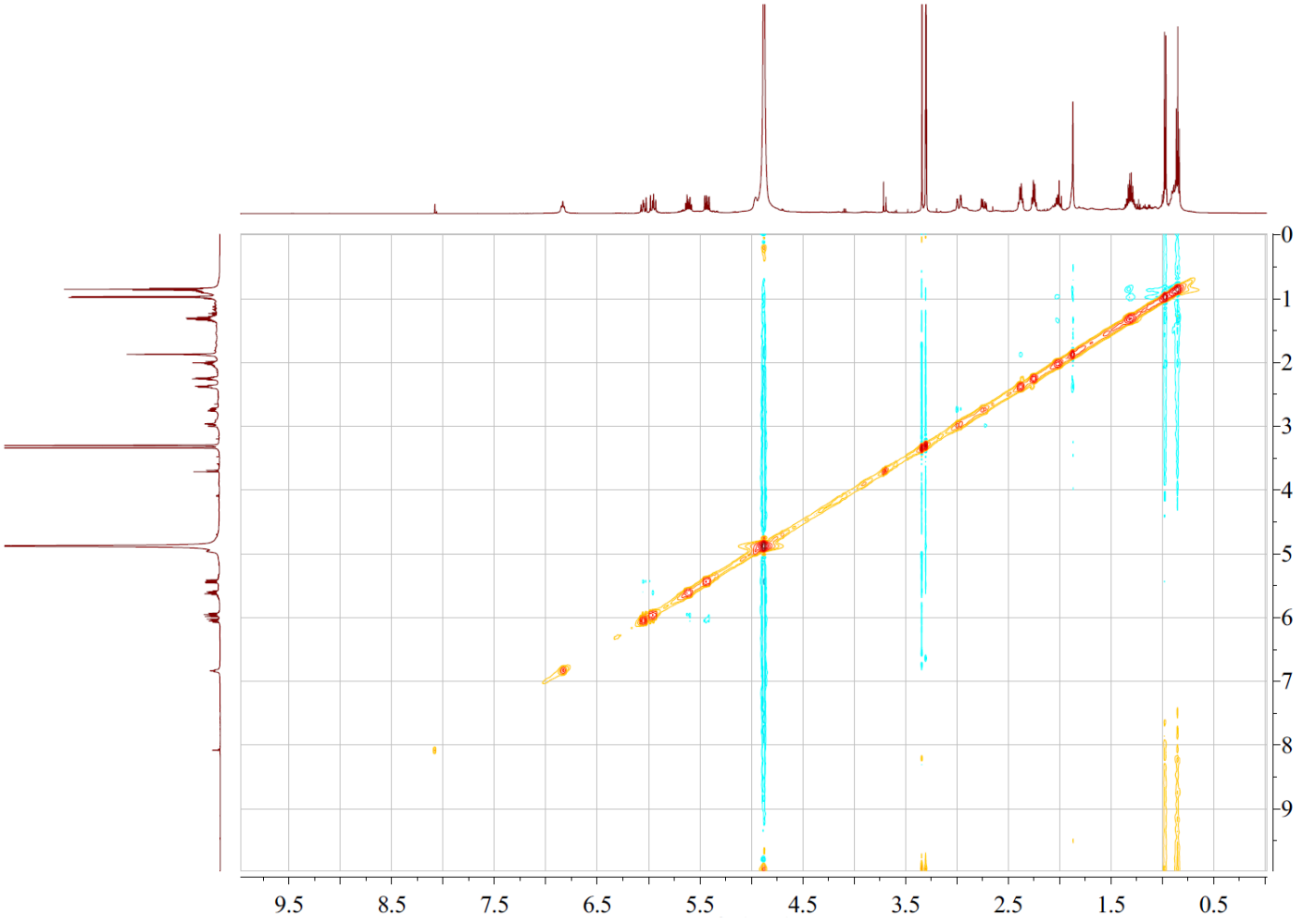


Figure S6. The spectroscopic data of **2**

(A) The HRESIMS spectrum of **2**

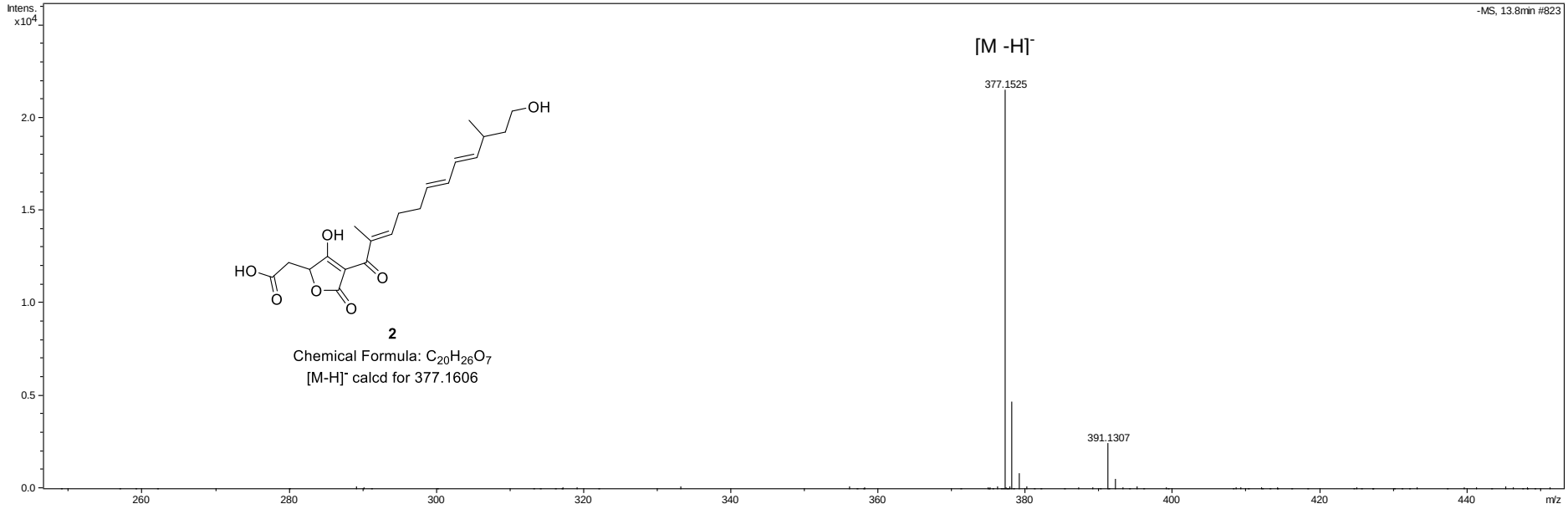


Figure S6. The spectroscopic data of **2**

(B) The ^{1}H-NMR spectrum of **2** (500 MHz for ^{1}H NMR in CD_3OD)

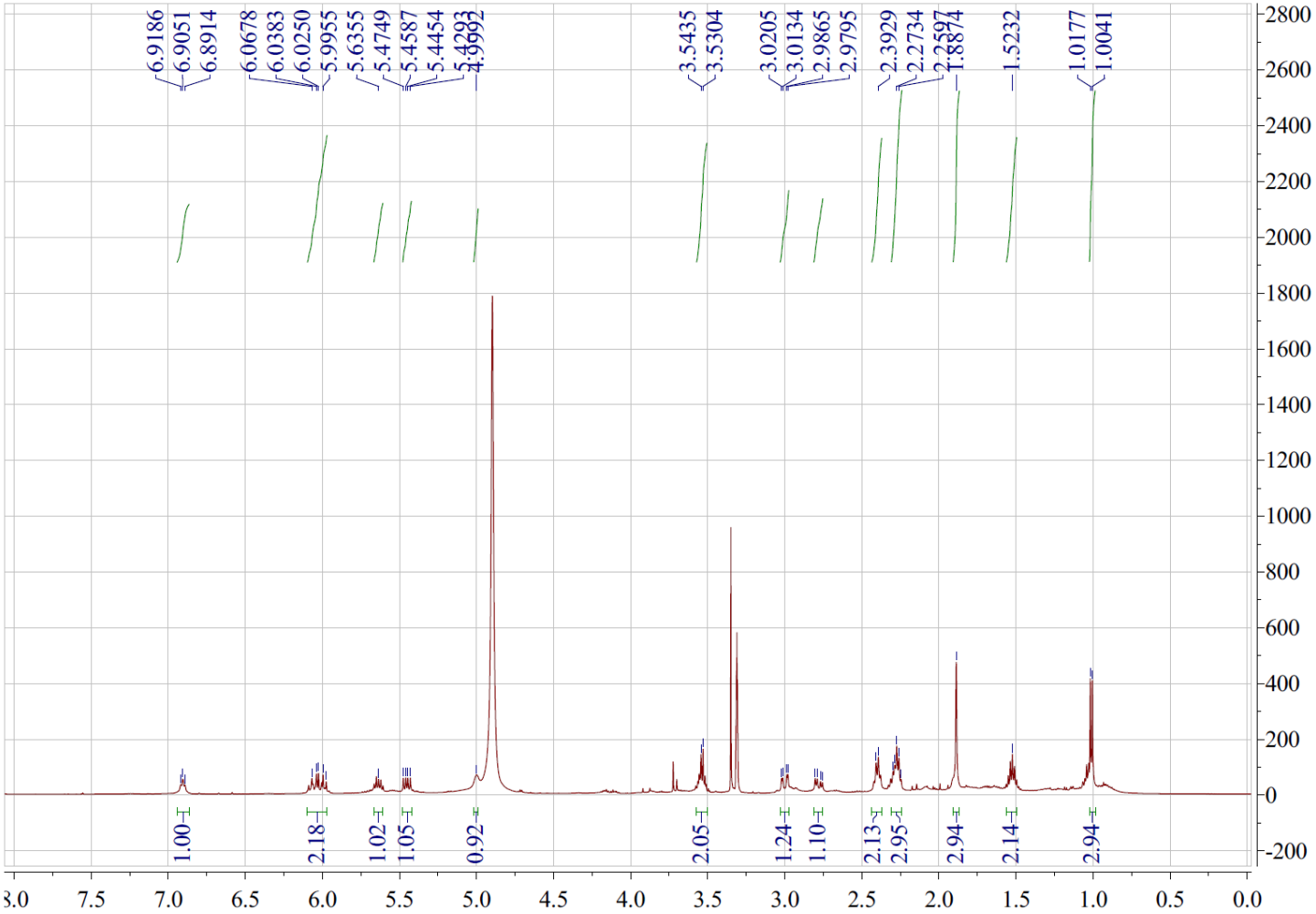


Figure S6. The spectroscopic data of **2**

(C) The ^{13}C NMR spectrum of compound **2** in CD_3OD (125 MHz)

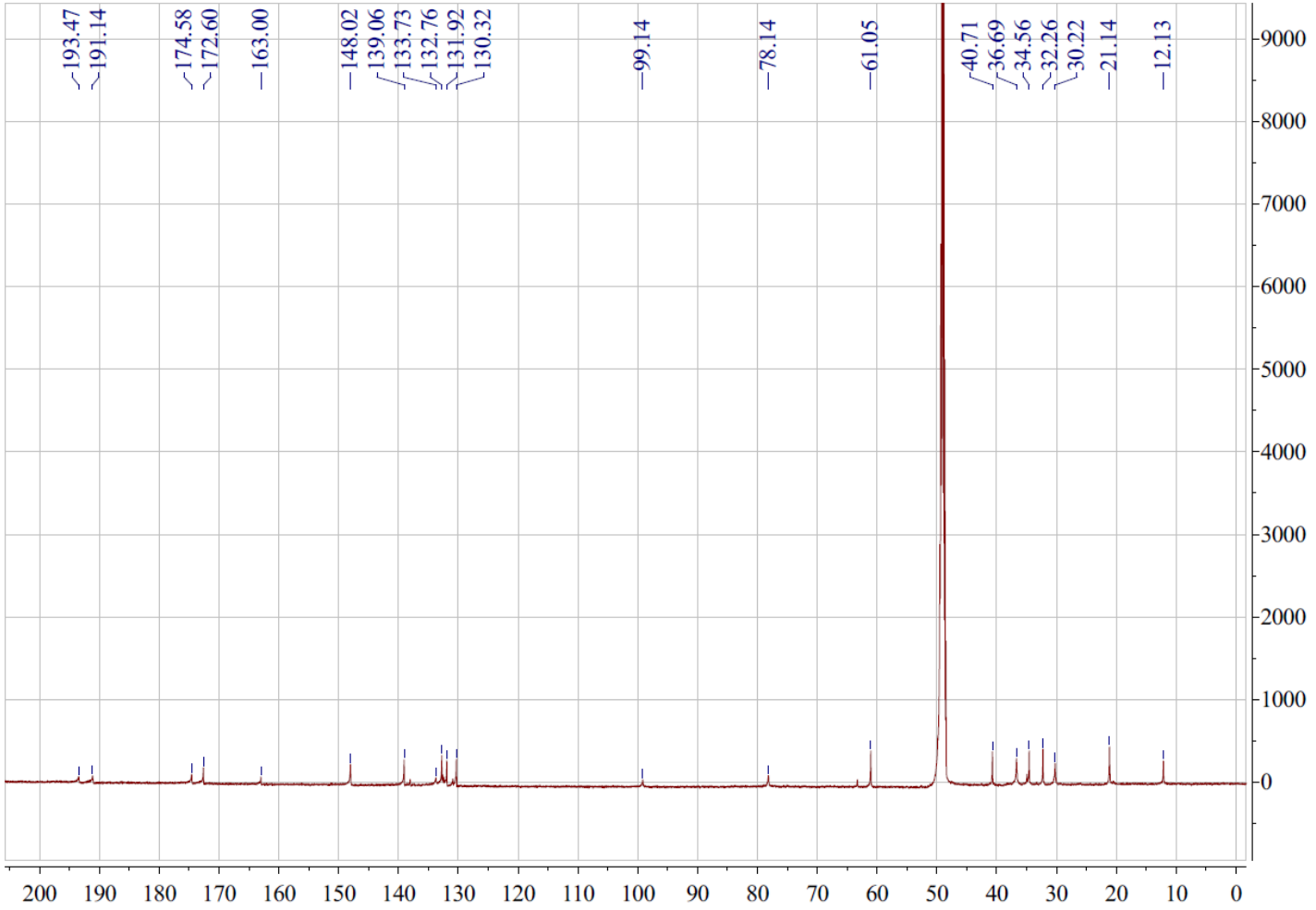


Figure S6. The spectroscopic data of **2**

(D) The HSQC spectrum of compound **2** in CD_3OD (125 MHz)

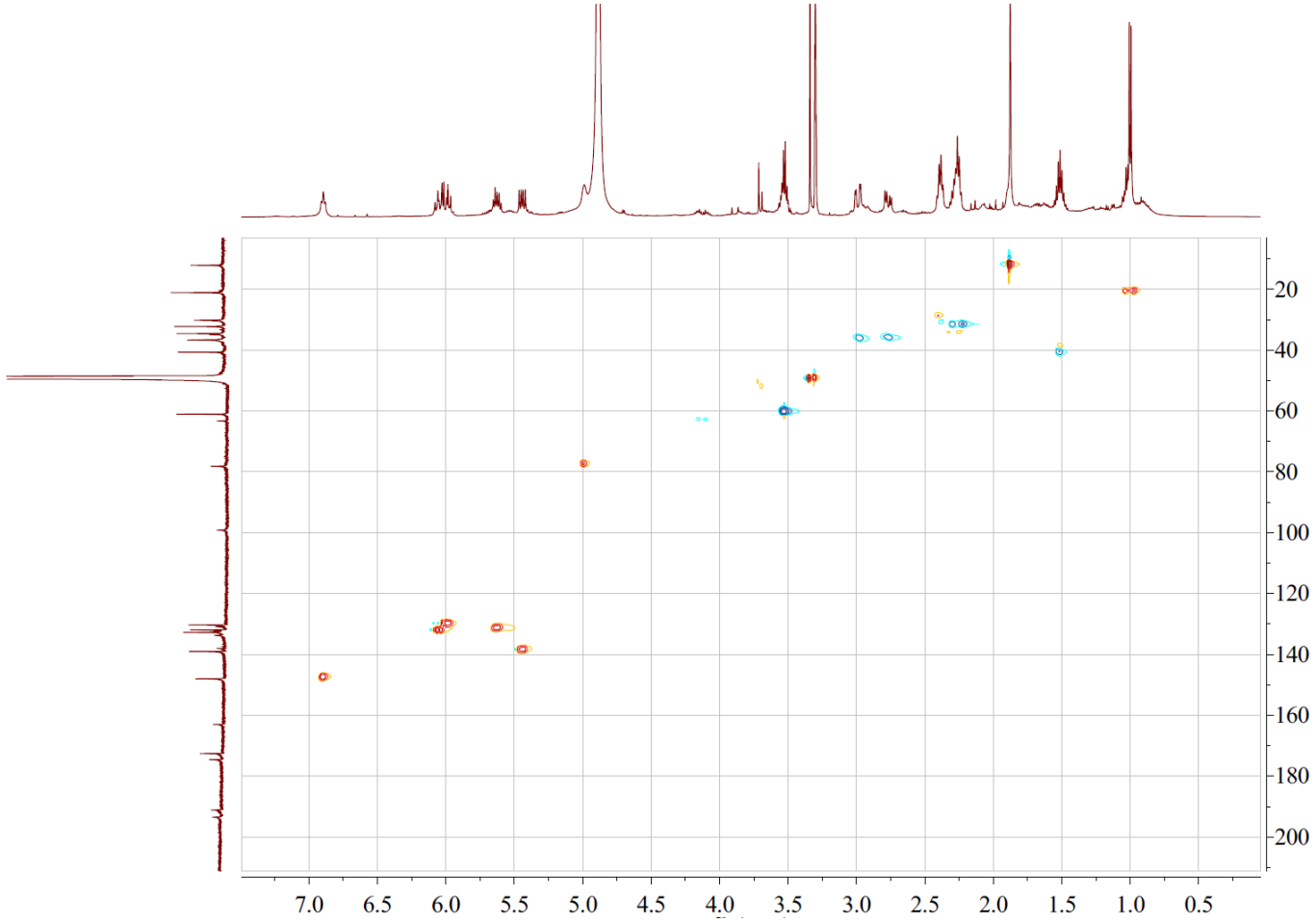


Figure S6. The spectroscopic data of **2**

(E) The HMBC spectrum of **2** (500 MHz for 1H NMR in CD_3OD)

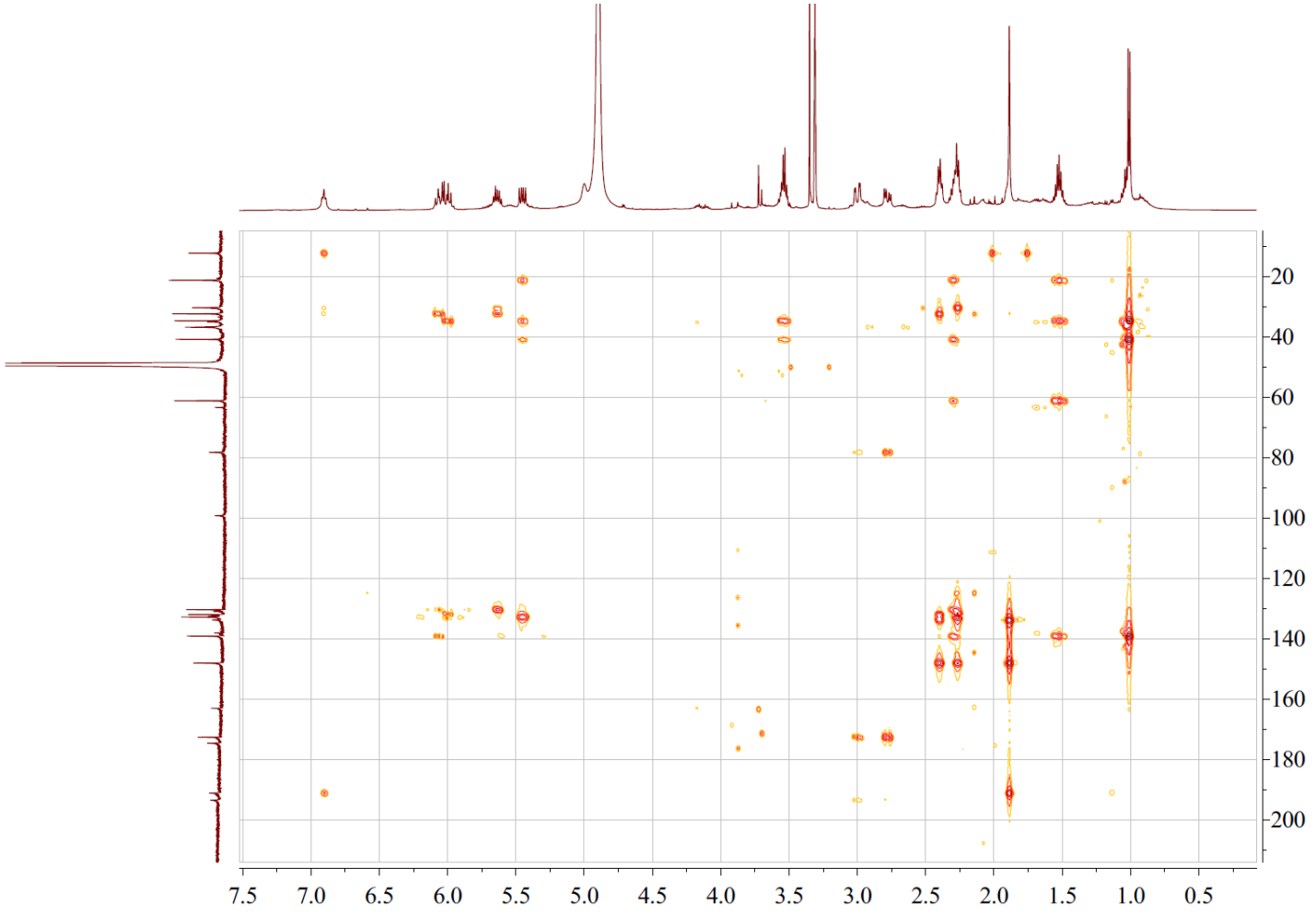


Figure S6. The spectroscopic data of **2**

(F) The ^{1}H-^{1}H COSY spectrum of **2** (500 MHz for ^{1}H NMR in CD_3OD)

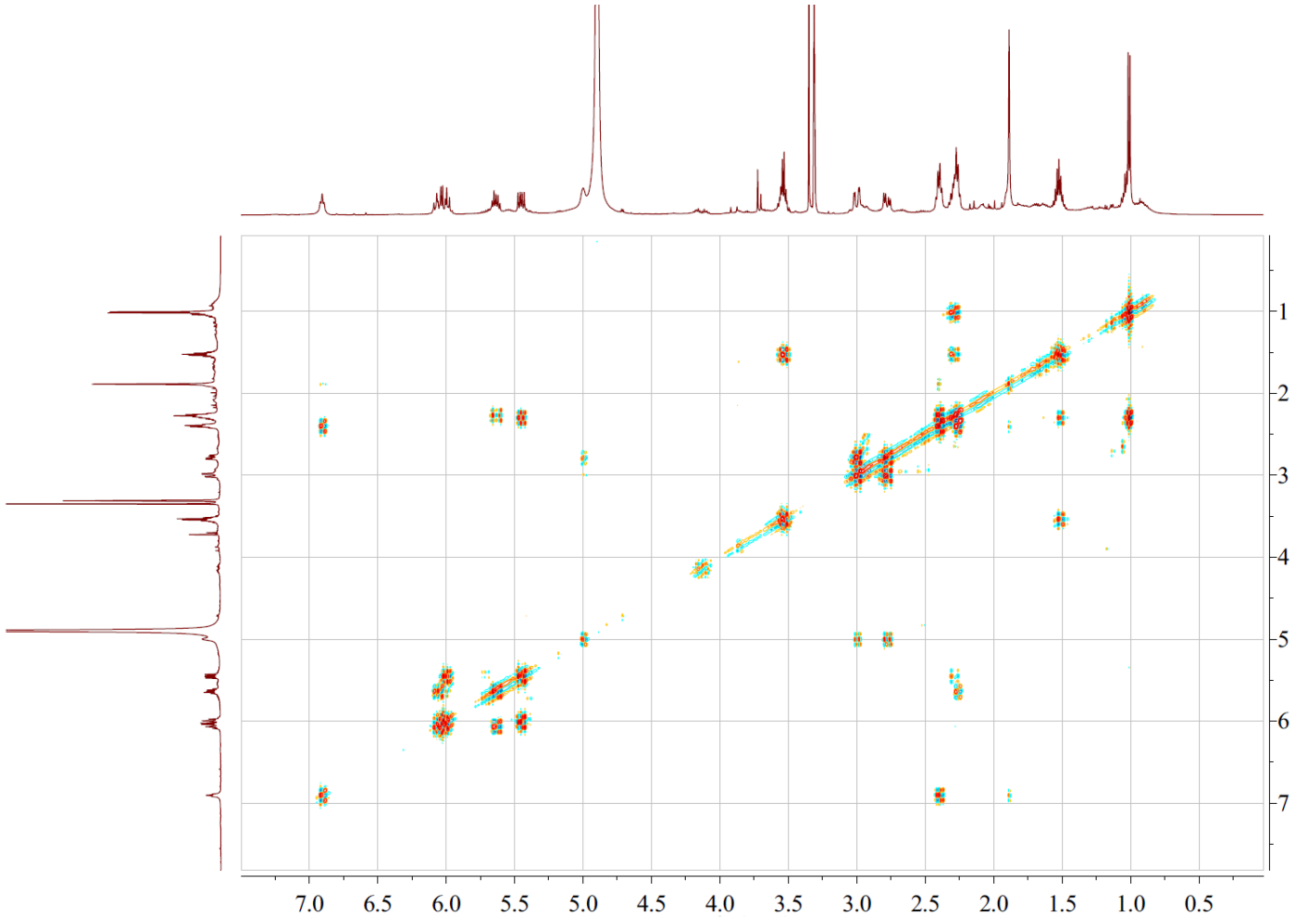


Figure S6. The spectroscopic data of **2**

(G) The NOESY spectrum of **2** (500 MHz for ^{1}H NMR in CD_3OD)

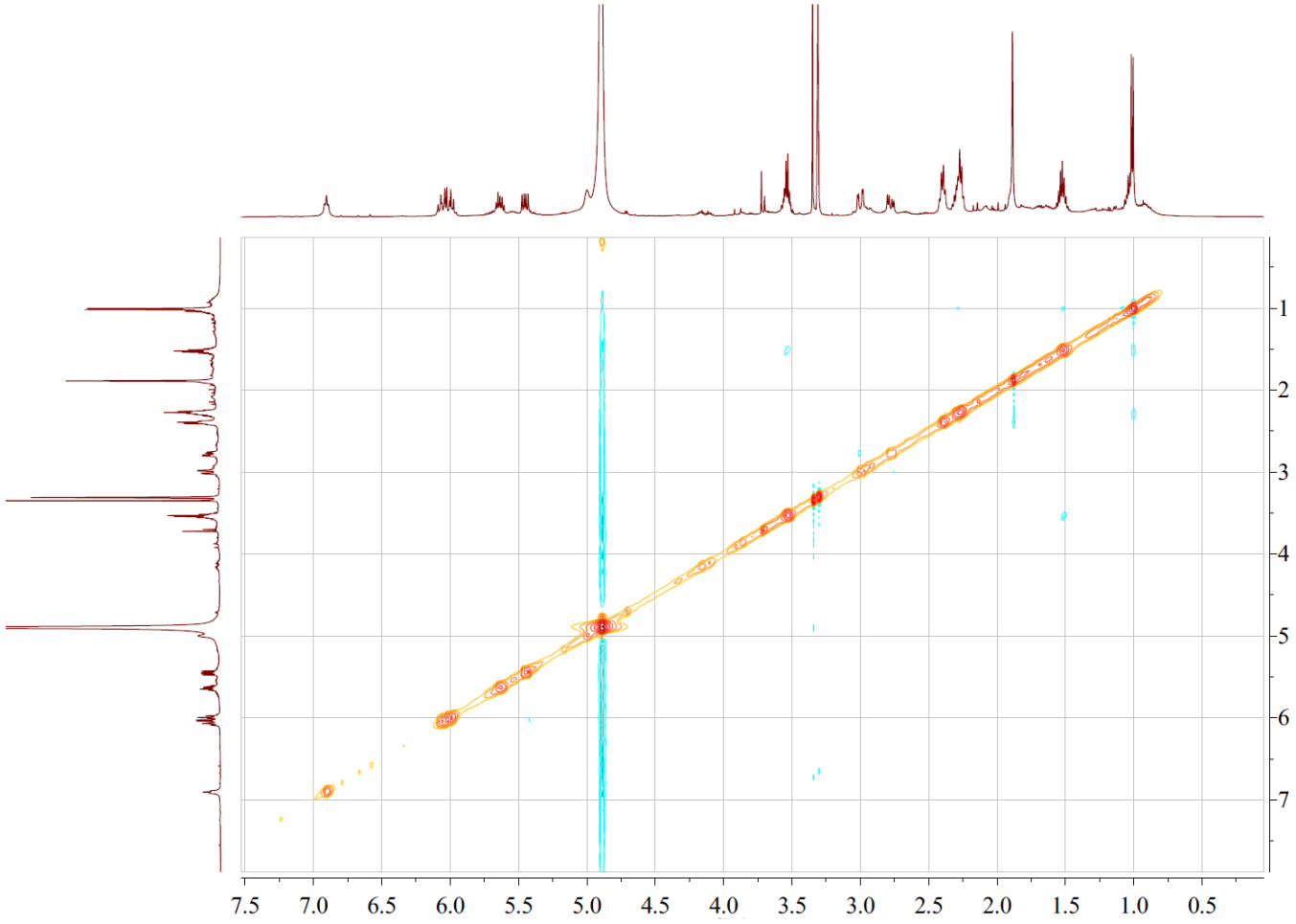


Figure S7. The spectroscopic data of **3**

(A) The HRESIMS spectrum of **3**

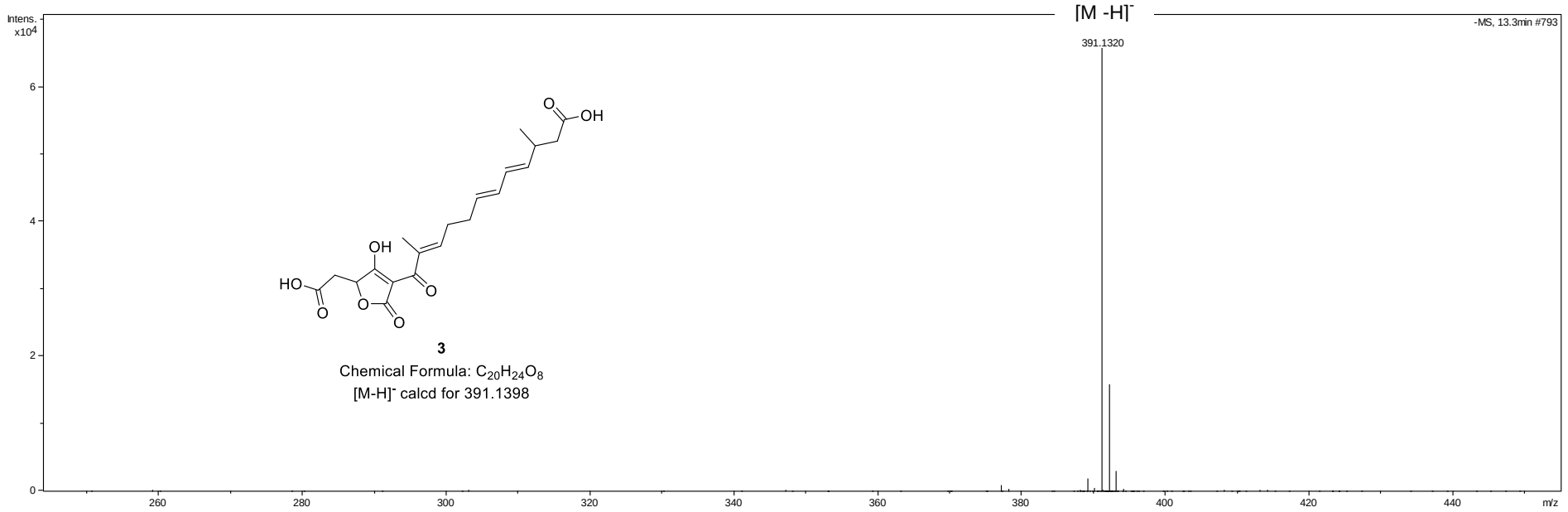


Figure S7. The spectroscopic data of **3**

(B) The ^{1}H-NMR spectrum of **3** (500 MHz for ^{1}H NMR in CD_3OD)

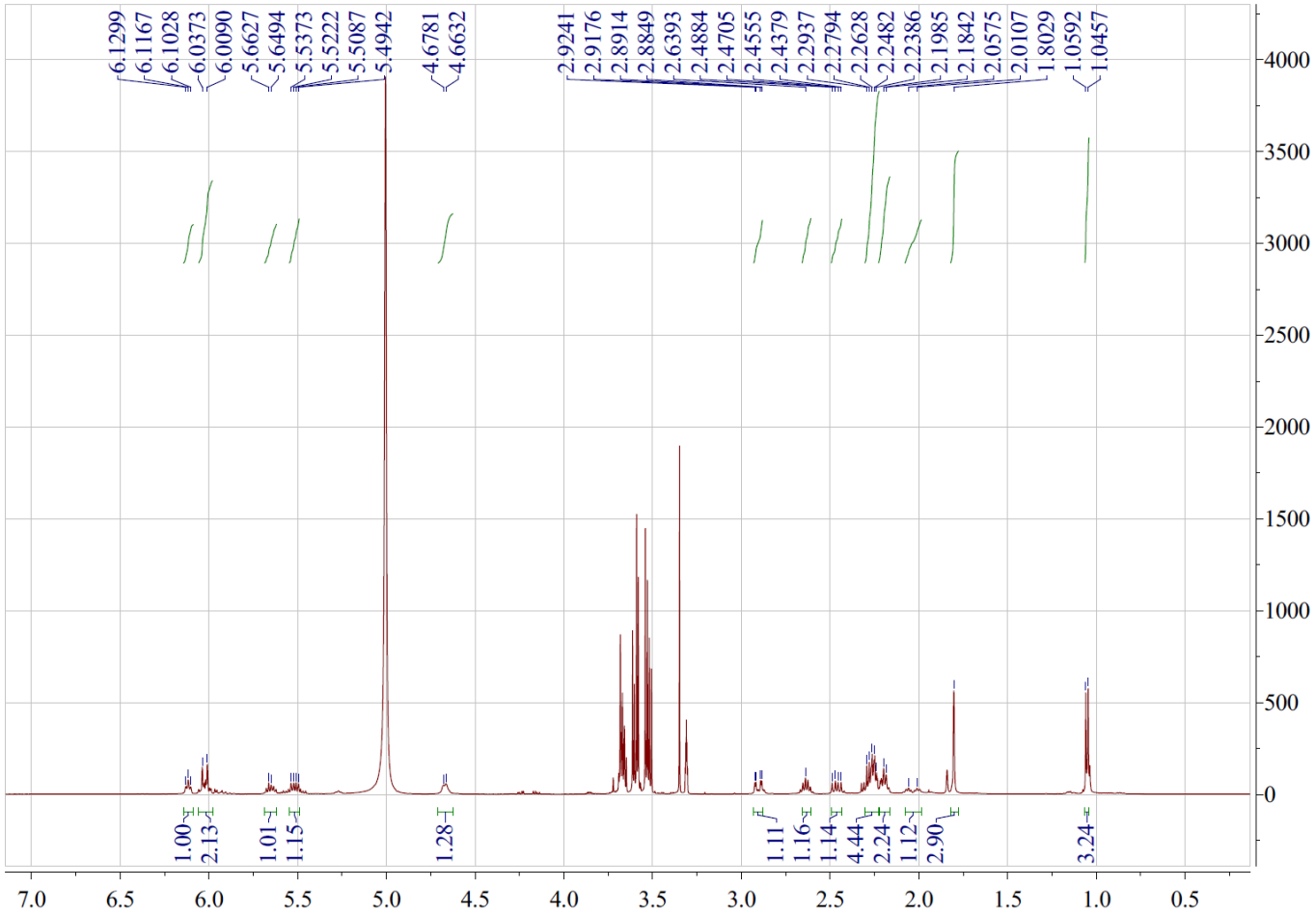


Figure S7. The spectroscopic data of **3**

(C) The ^{13}C NMR spectrum of compound **3** in CD_3OD (125 MHz)

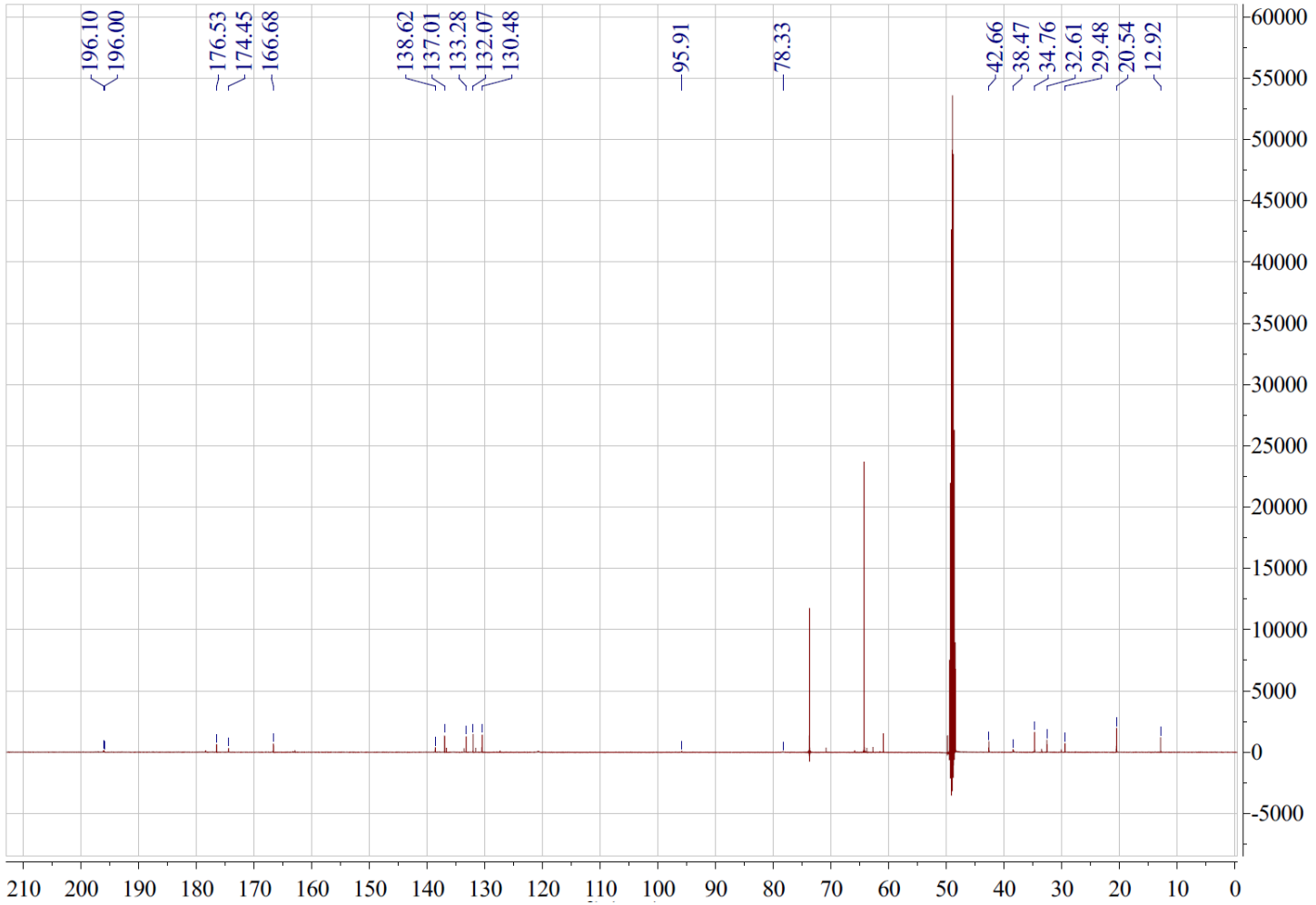


Figure S7. The spectroscopic data of **3**

(D) The HSQC spectrum of compound **3** in CD_3OD (125 MHz)

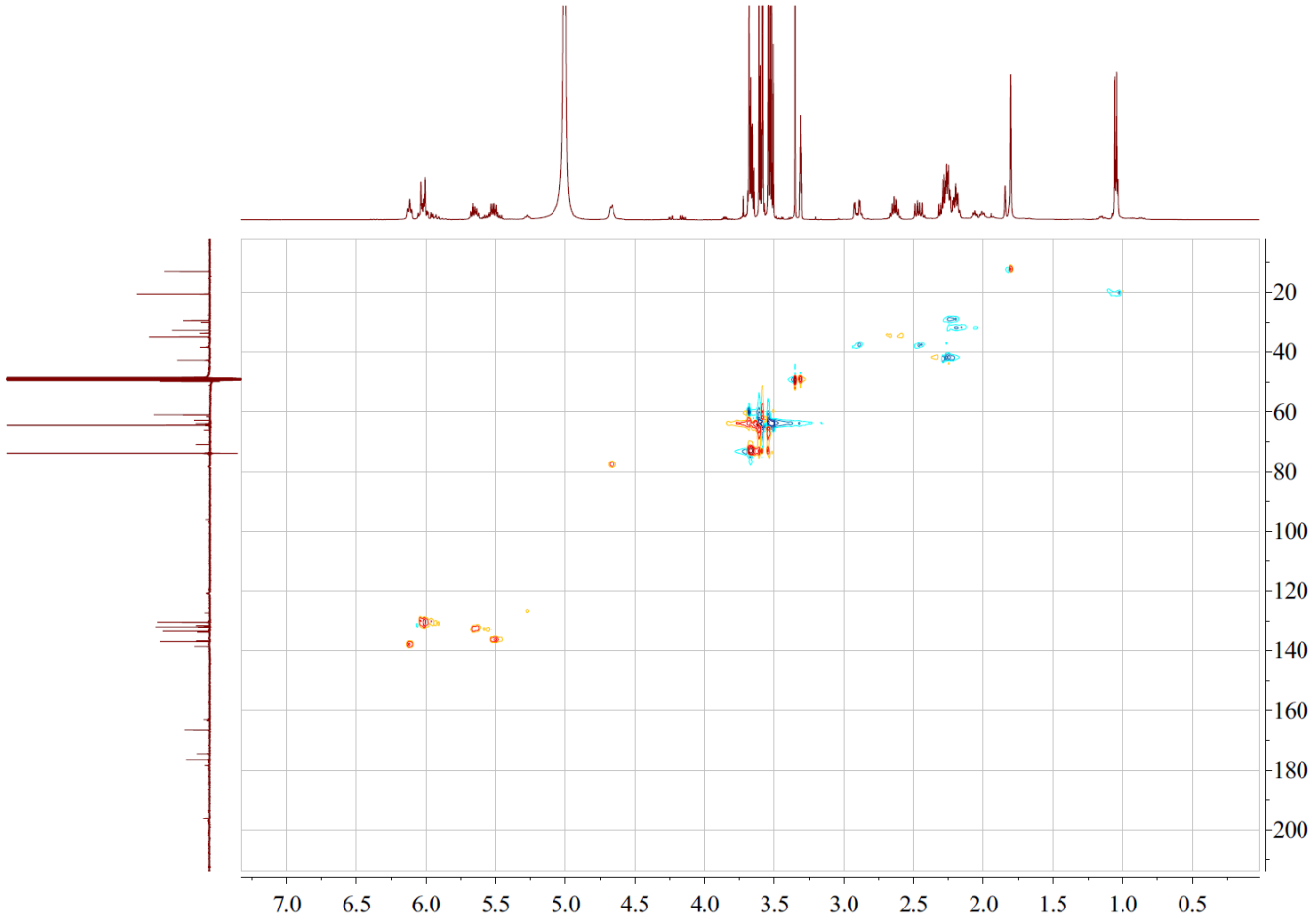


Figure S7. The spectroscopic data of **3**

(E) The HMBC spectrum of **3** (500 MHz for ^{1}H NMR in CD_3OD)

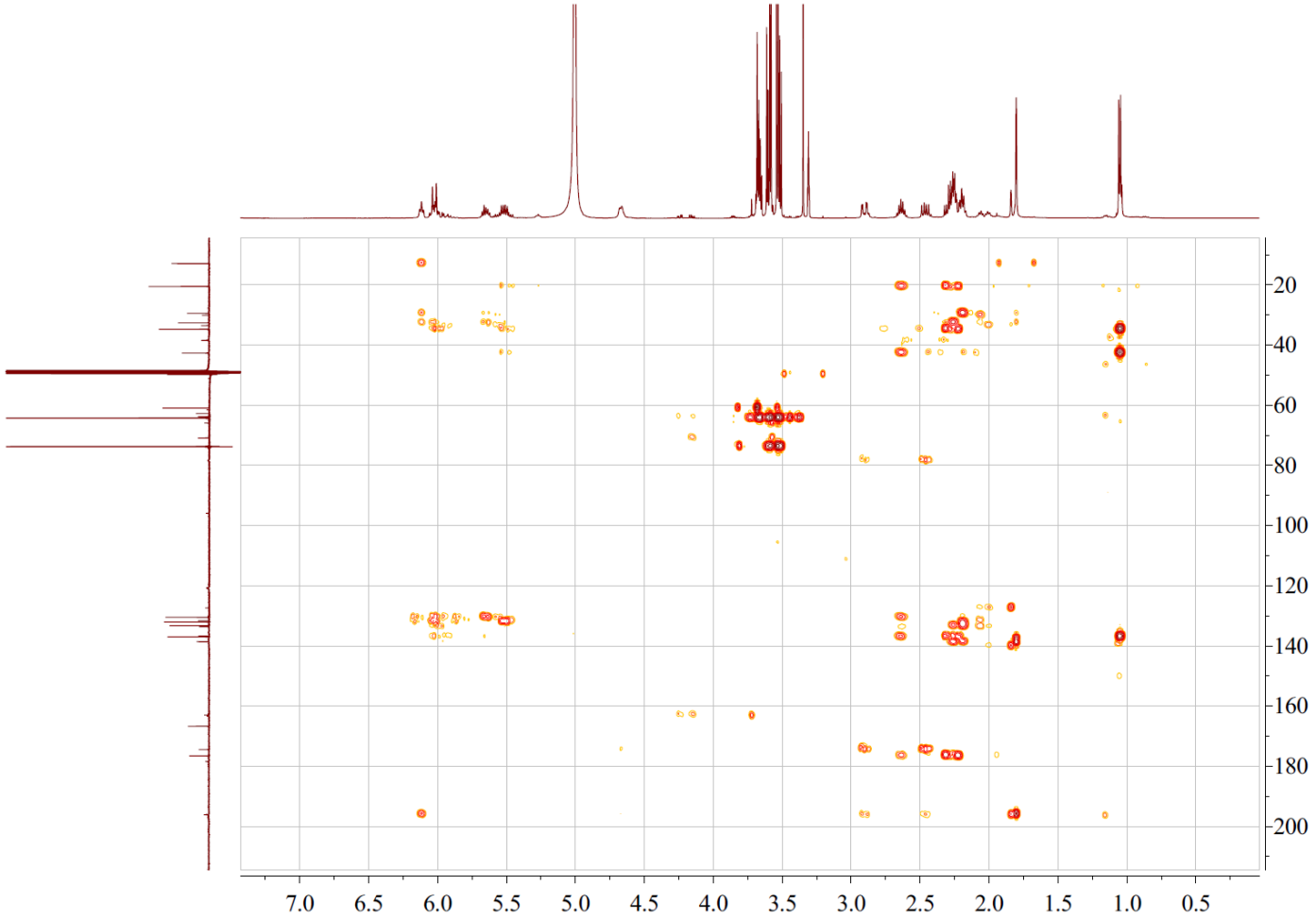


Figure S7. The spectroscopic data of **3**

(F) The ^{1}H-^{1}H COSY spectrum of **3** (500 MHz for ^{1}H NMR in CD_3OD)

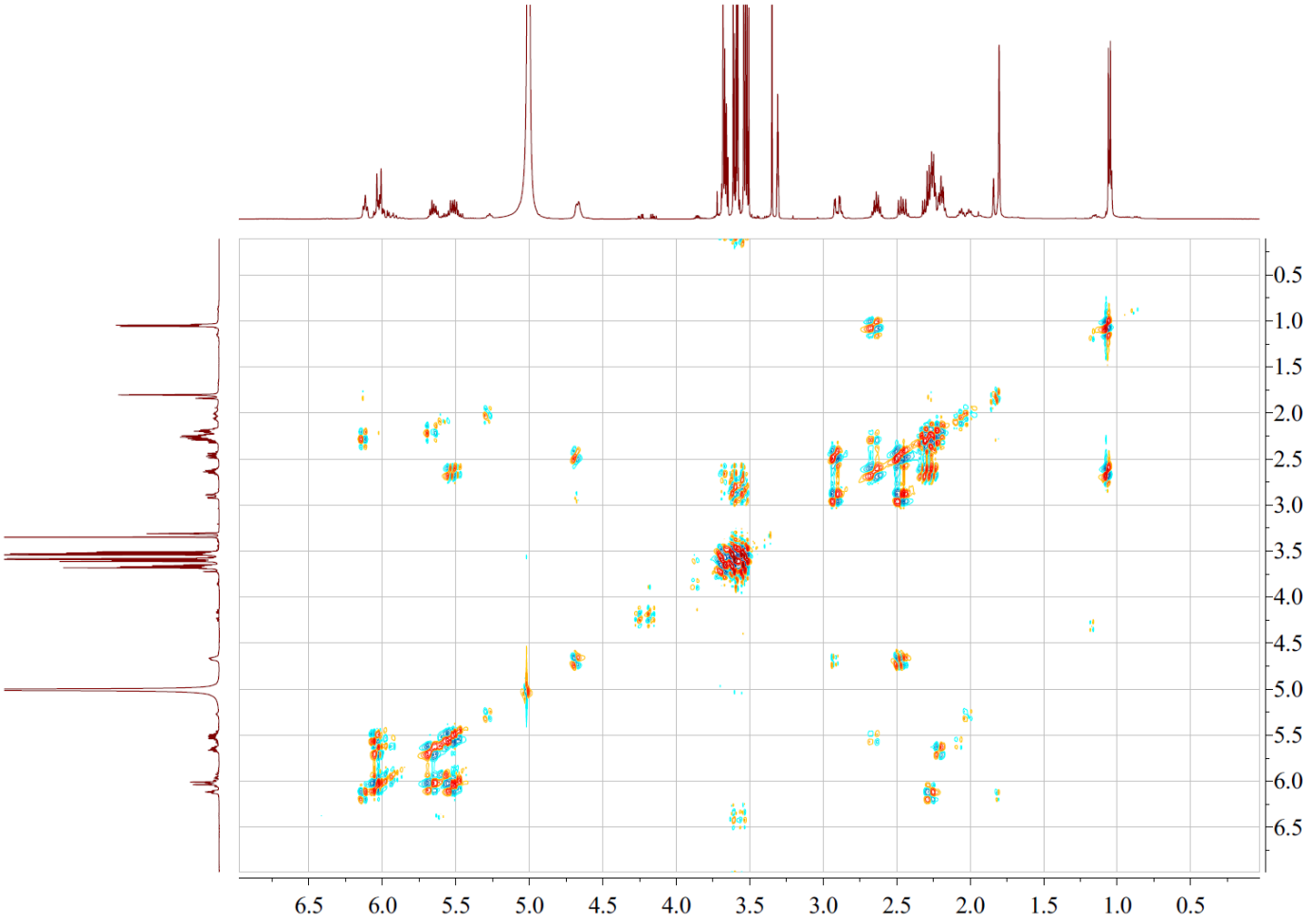


Figure S8. The spectroscopic data of **4**

(A) The HRESIMS spectrum of **4**

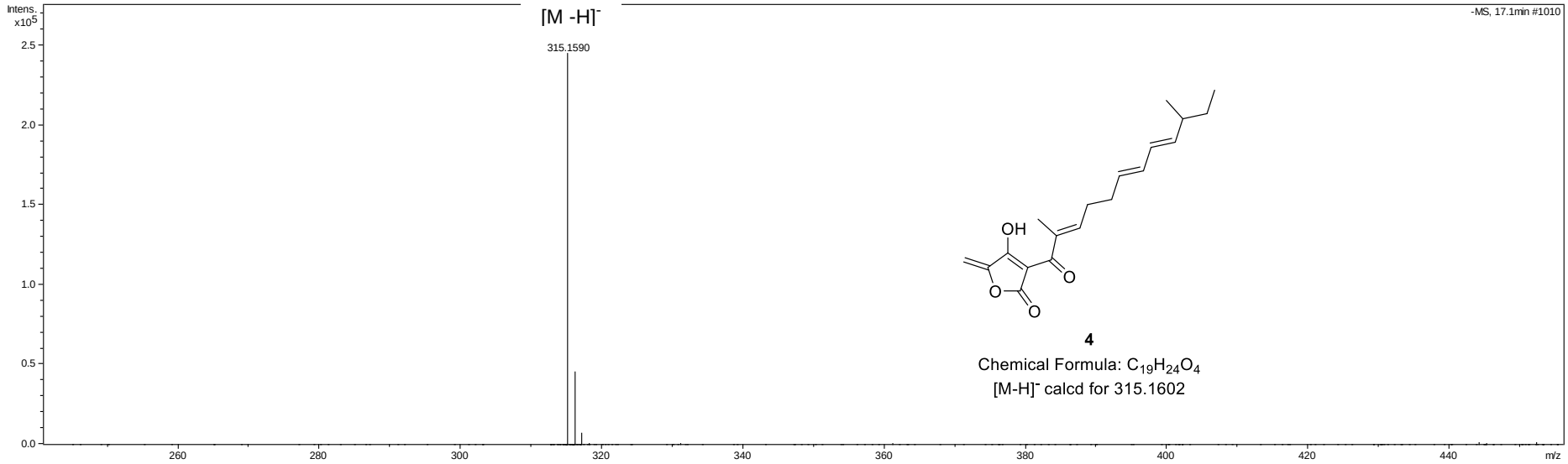


Figure S8. The spectroscopic data of **4**

(B) The ^{1}H-NMR spectrum of **4** (500 MHz for ^{1}H NMR in DMSO-d_6)

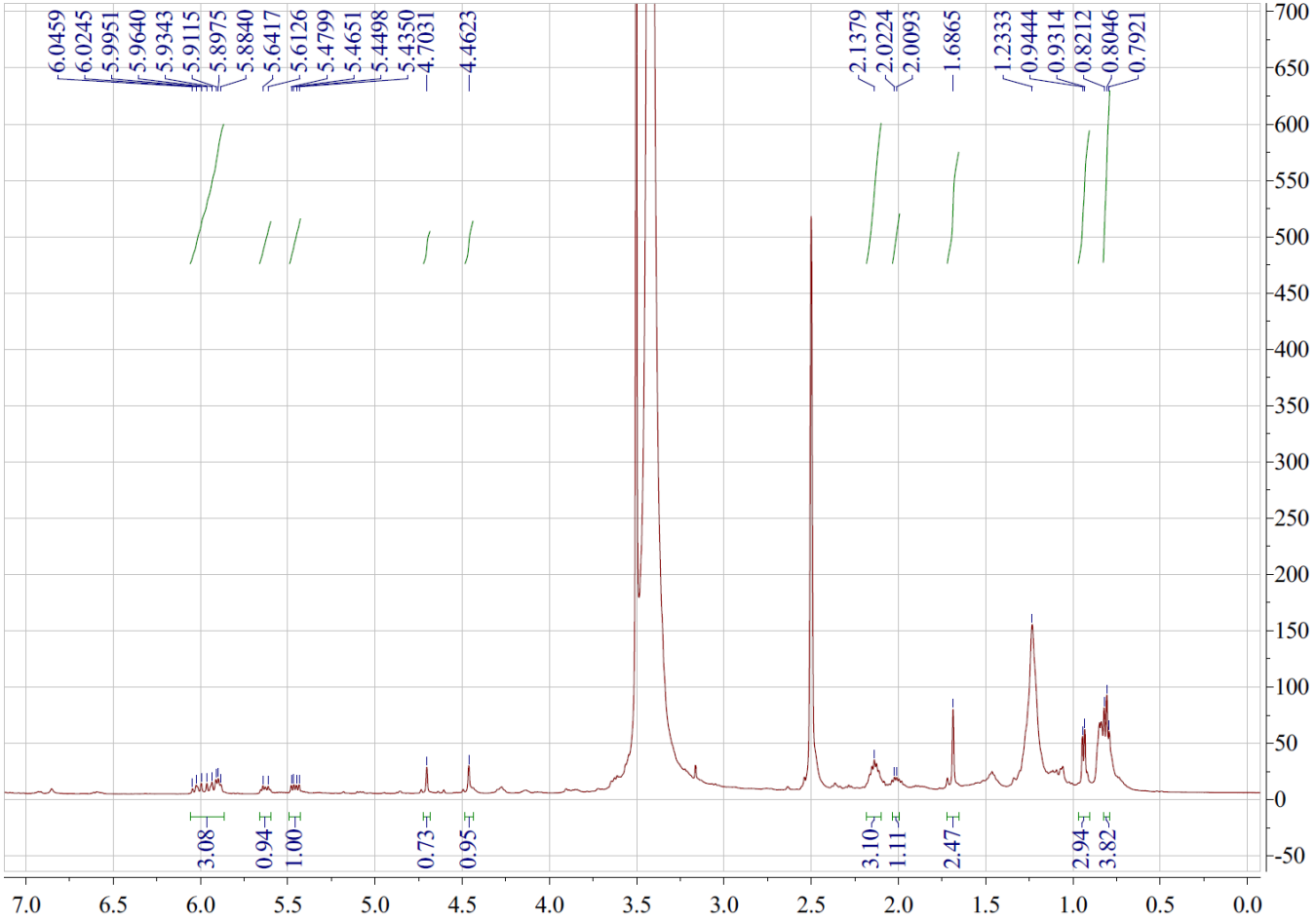


Figure S8. The spectroscopic data of **4**

(C) The ^{13}C NMR spectrum of compound **4** in DMSO-d_6 (125 MHz)

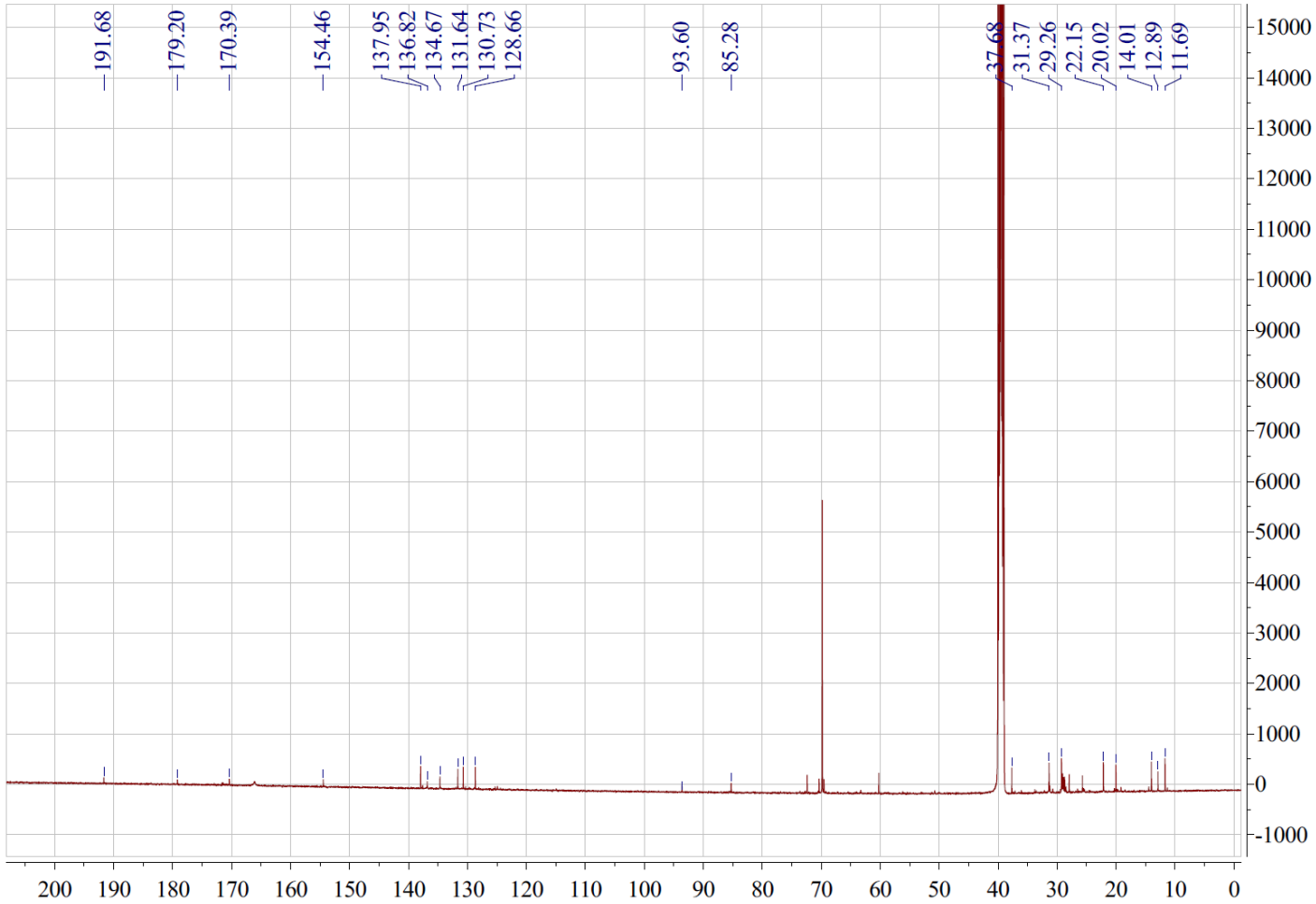


Figure S8. The spectroscopic data of **4**

(D) The HSQC spectrum of compound **4** in DMSO-d_6 (125 MHz)

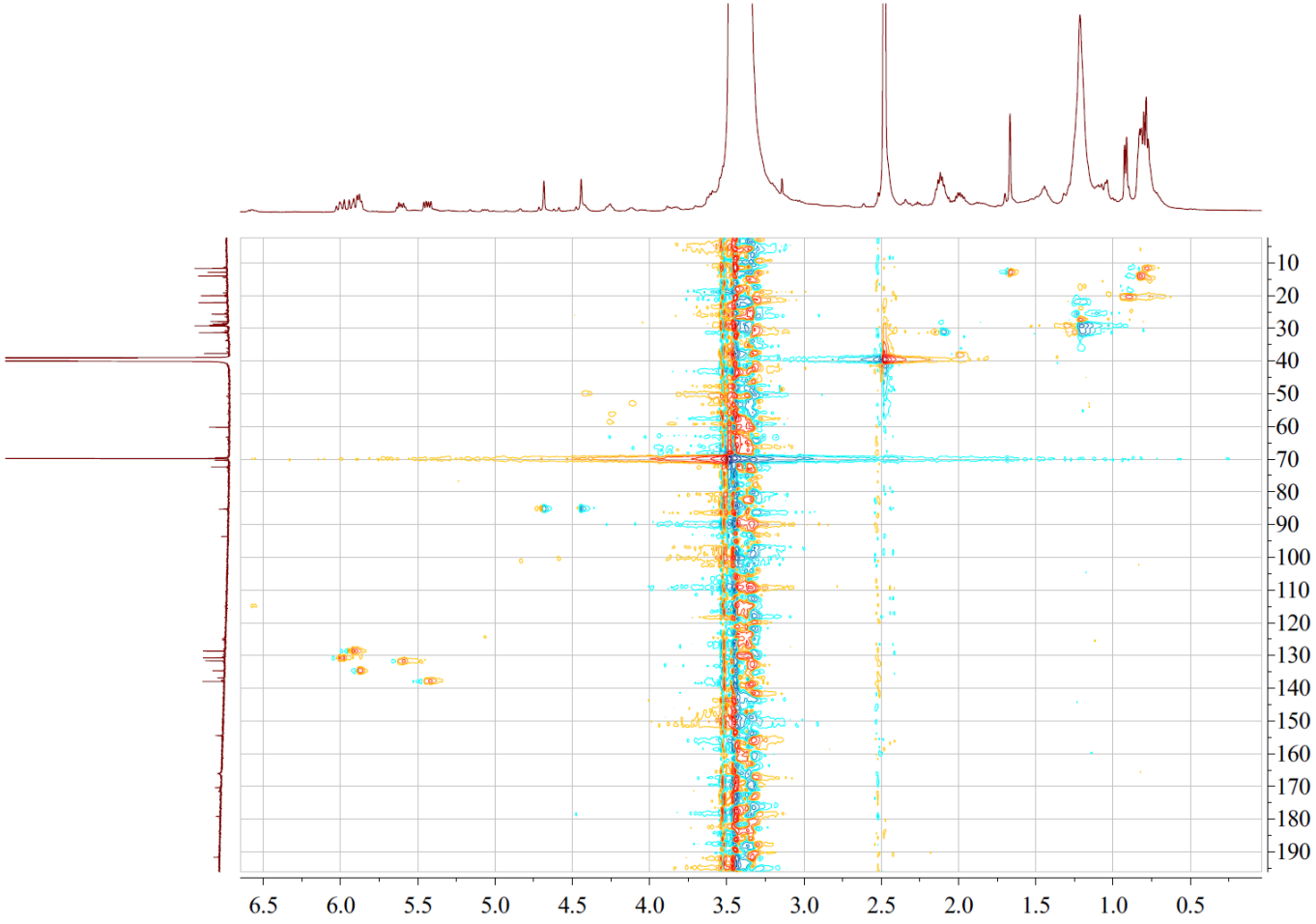


Figure S8. The spectroscopic data of **4**

(E) The HMBC spectrum of **4** (500 MHz for ^{1}H NMR in DMSO-d_6)

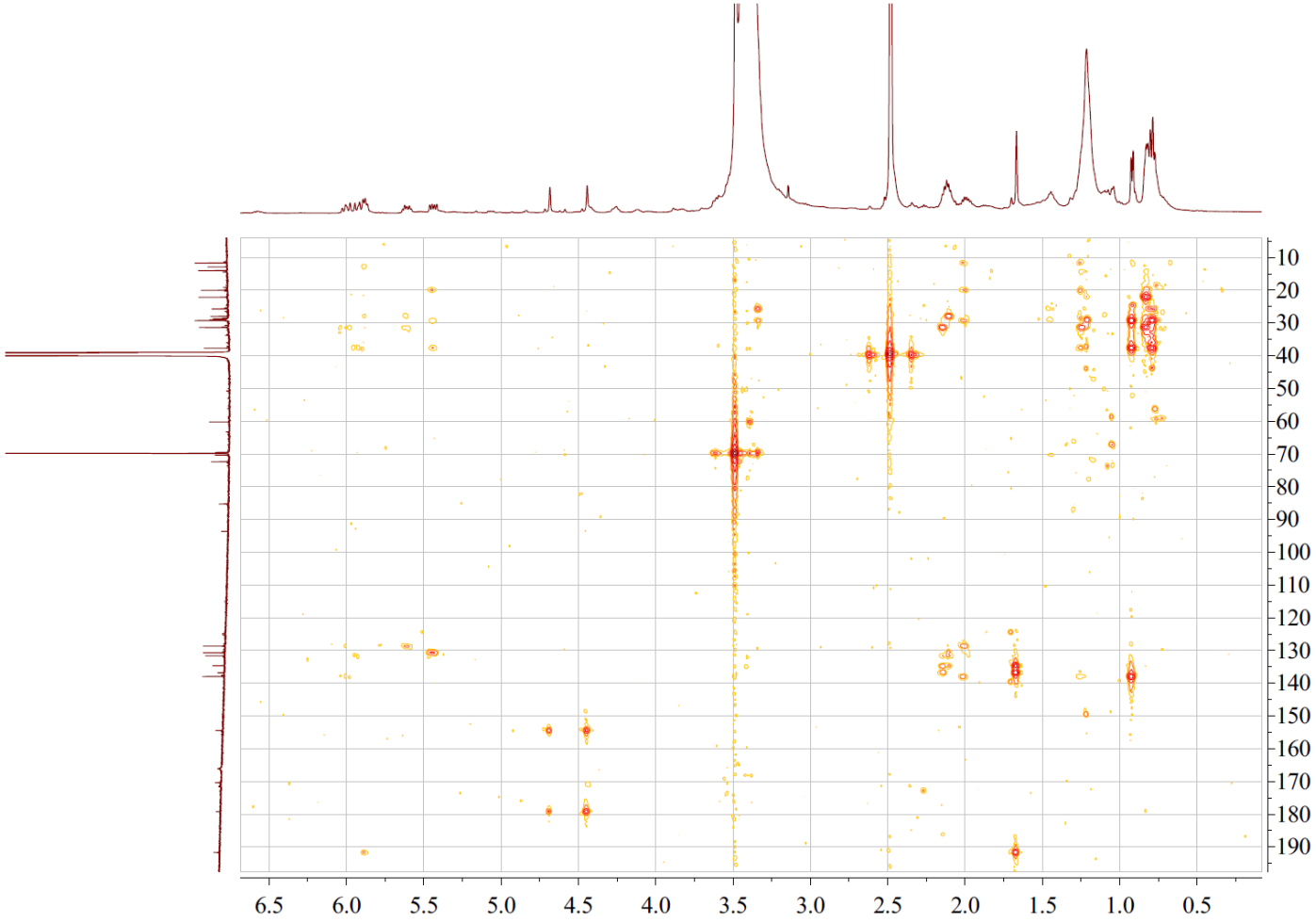


Figure S8. The spectroscopic data of **4**

(F) The ^{1}H-^{1}H COSY spectrum of 4 (500 MHz for ^{1}H NMR in DMSO-d_6)

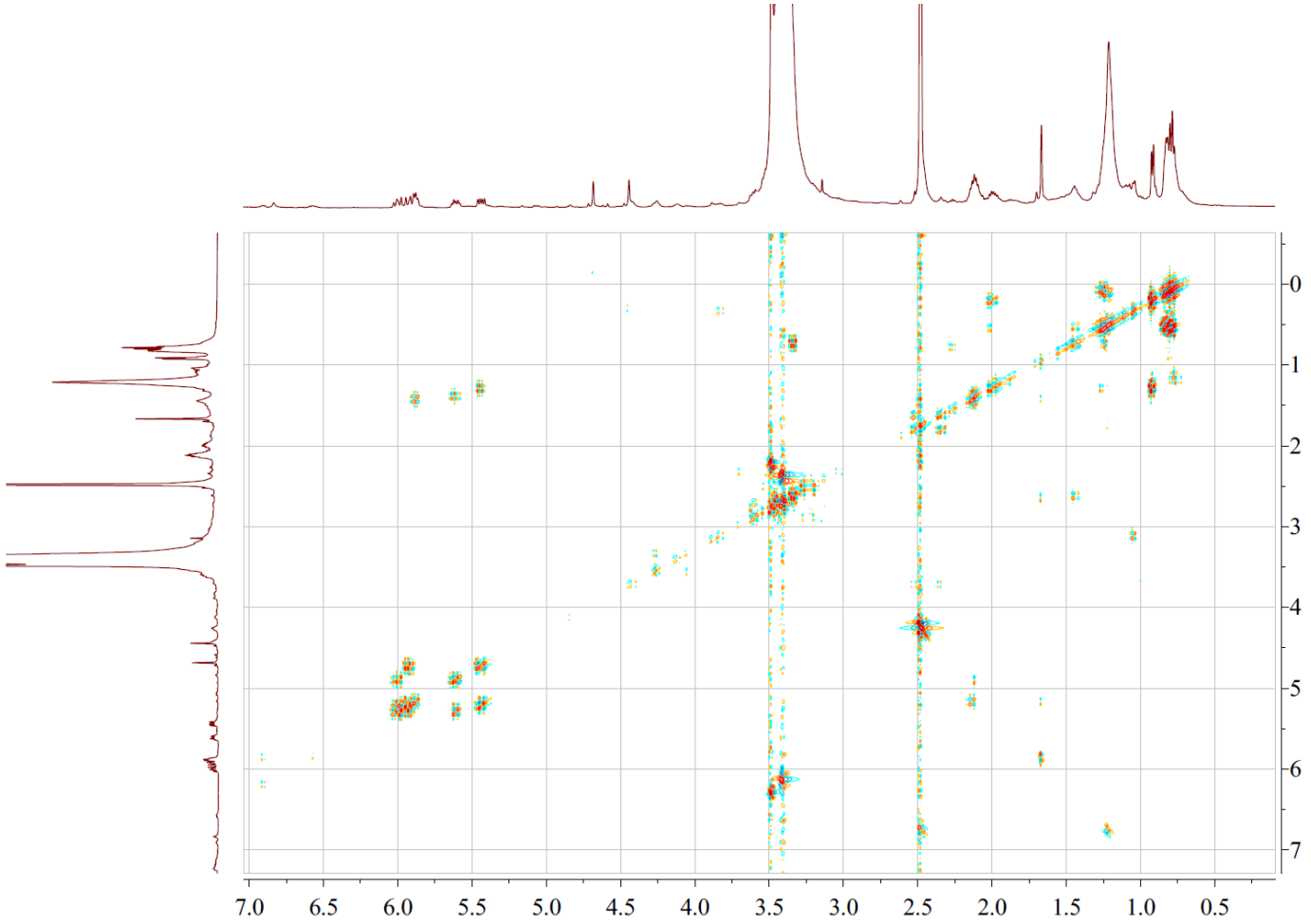


Figure S9. The spectroscopic data of **5**

(A) The HRESIMS spectrum of **5**

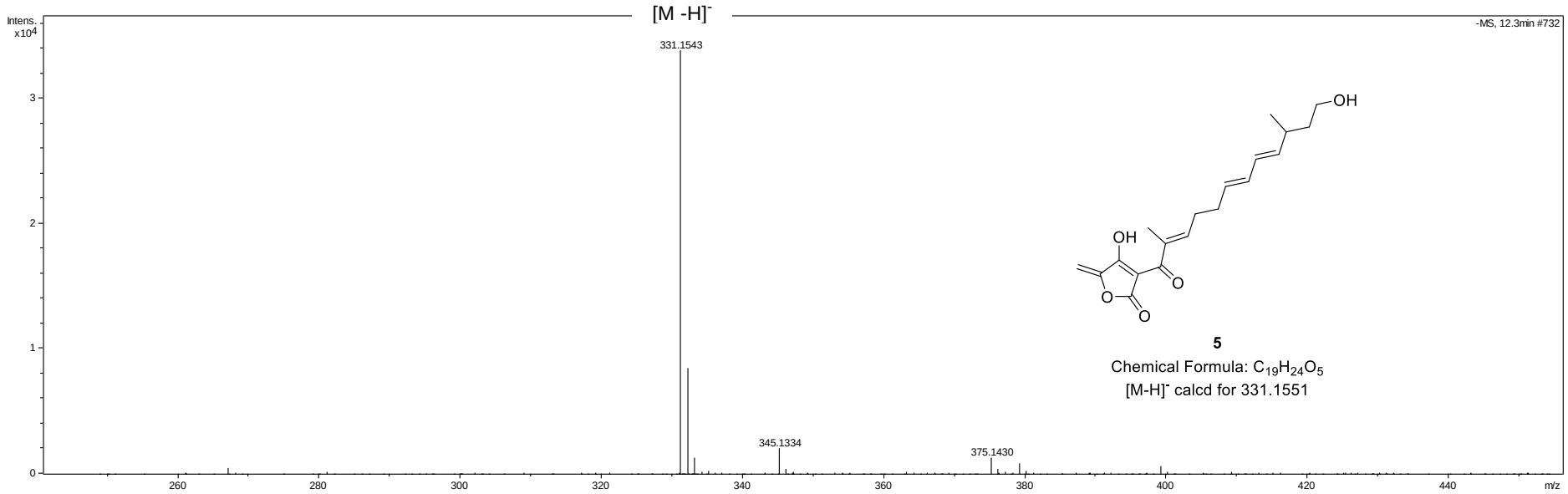


Figure S9. The spectroscopic data of **5**

(B) The ^{1}H-NMR spectrum of **5** (500 MHz for ^{1}H NMR in $CDCl_3$)

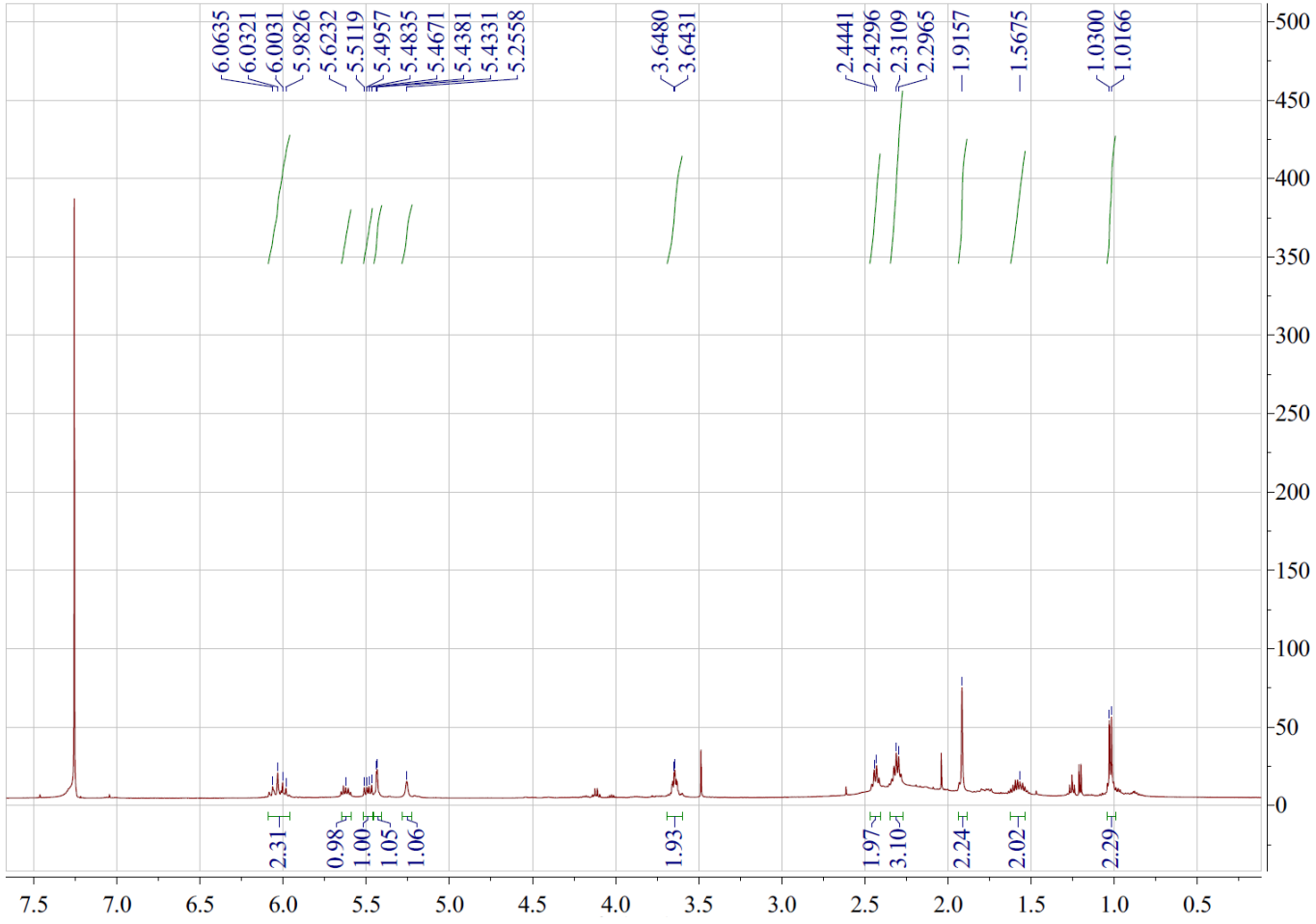


Figure S9. The spectroscopic data of **5**

(C) The ^{13}C NMR spectrum of compound **5** in $CDCl_3$ (125 MHz)

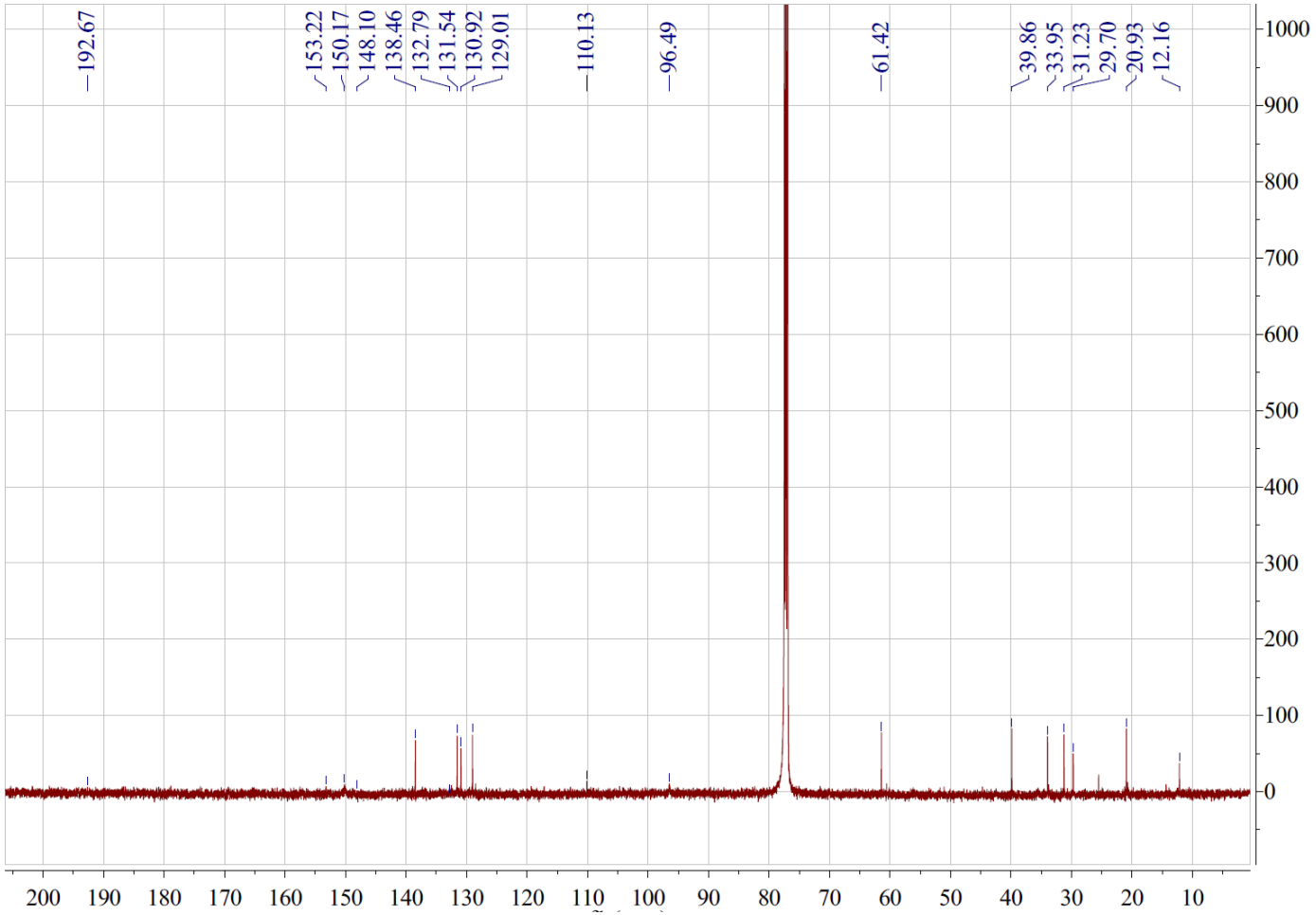


Figure S9. The spectroscopic data of **5**

(D) The HSQC spectrum of compound **5** in $CDCl_3$ (125 MHz)

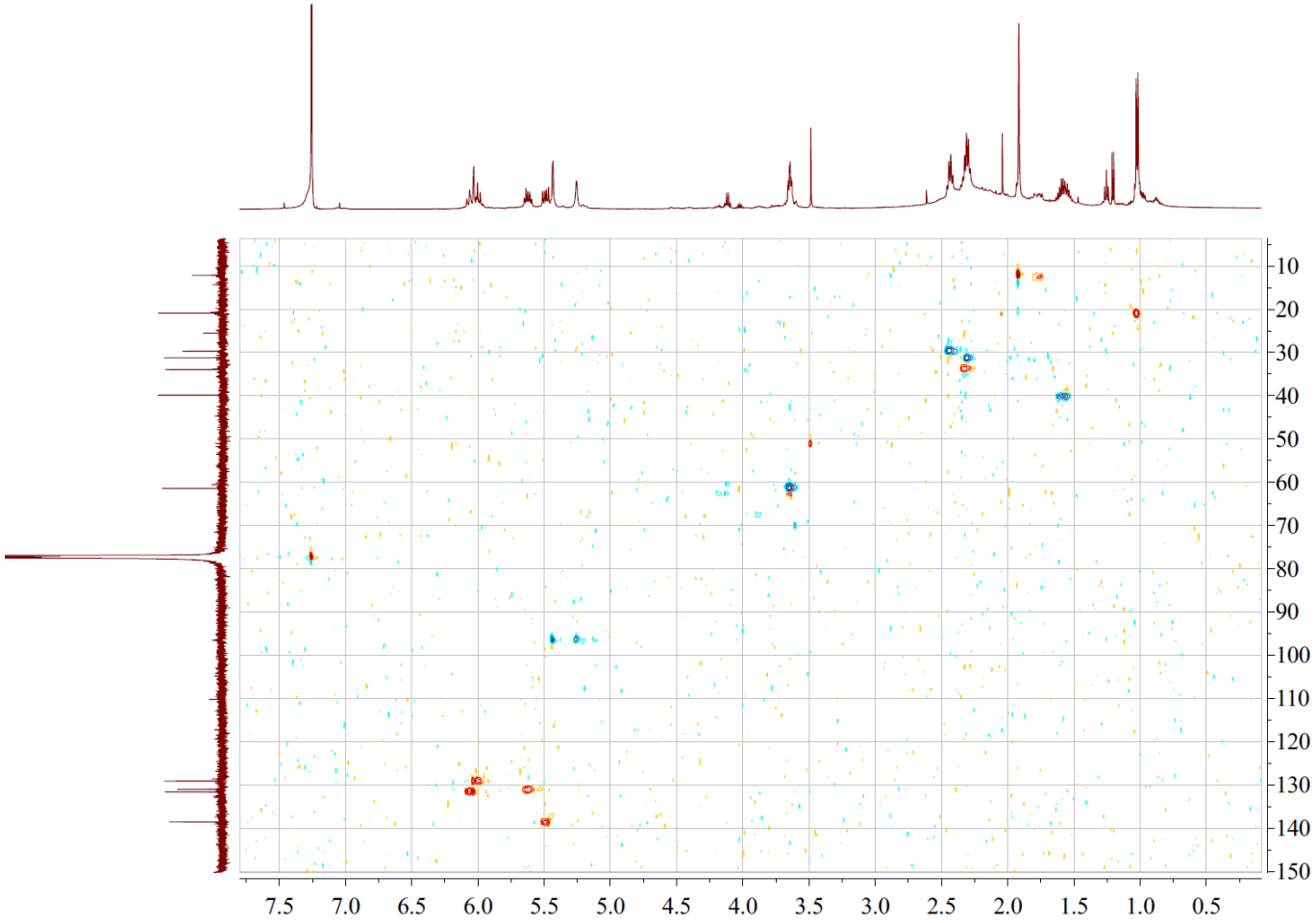


Figure S9. The spectroscopic data of **5**

(E) The HMBC spectrum of **5** (500 MHz for ^{1}H NMR in $CDCl_3$)

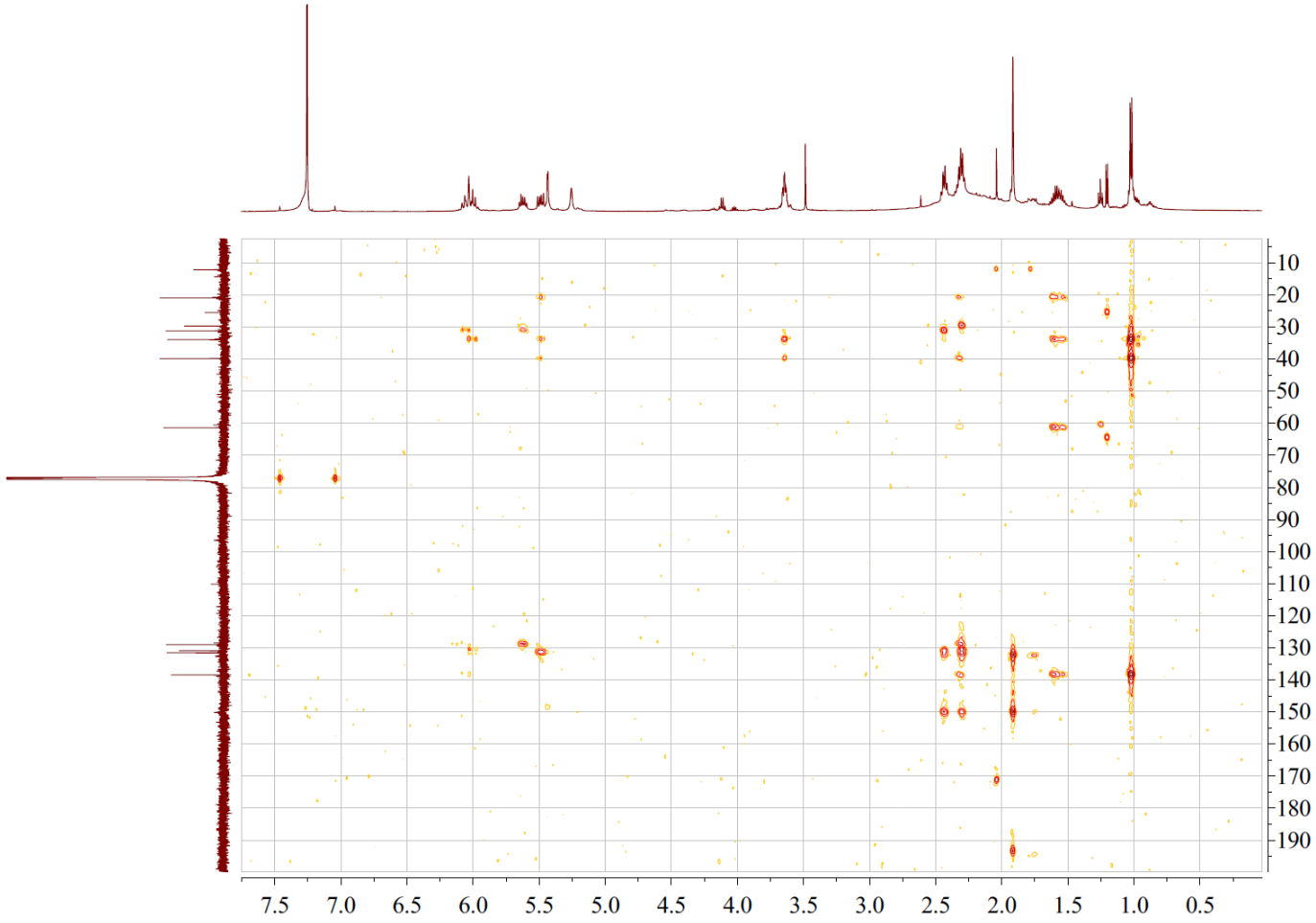


Figure S9. The spectroscopic data of **5**

(F) The ^{1}H-^{1}H COSY spectrum of **5** (500 MHz for ^{1}H NMR in $CDCl_3$)

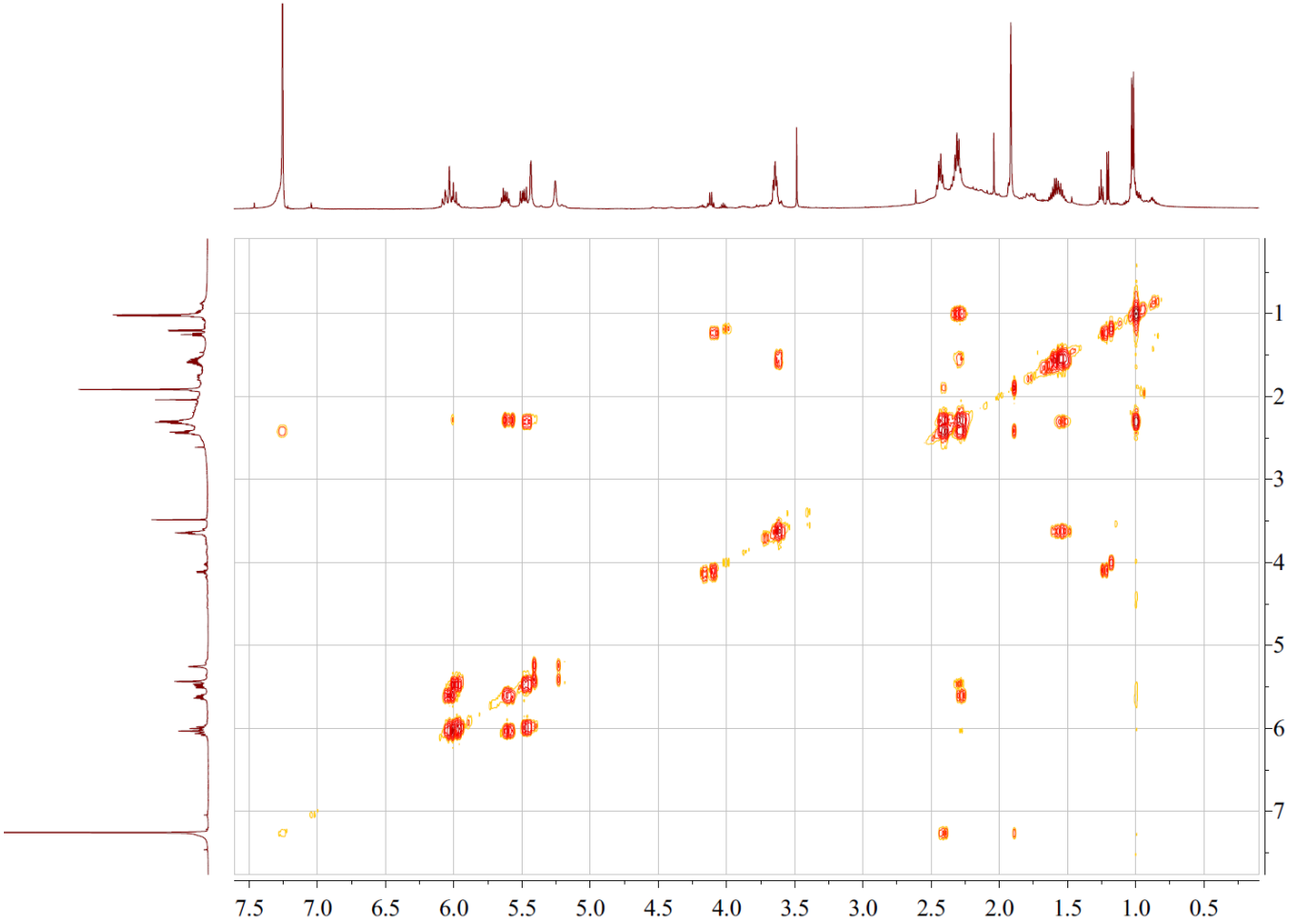


Figure S10. The spectroscopic data of **6**

(A) The HRESIMS spectrum of **6**

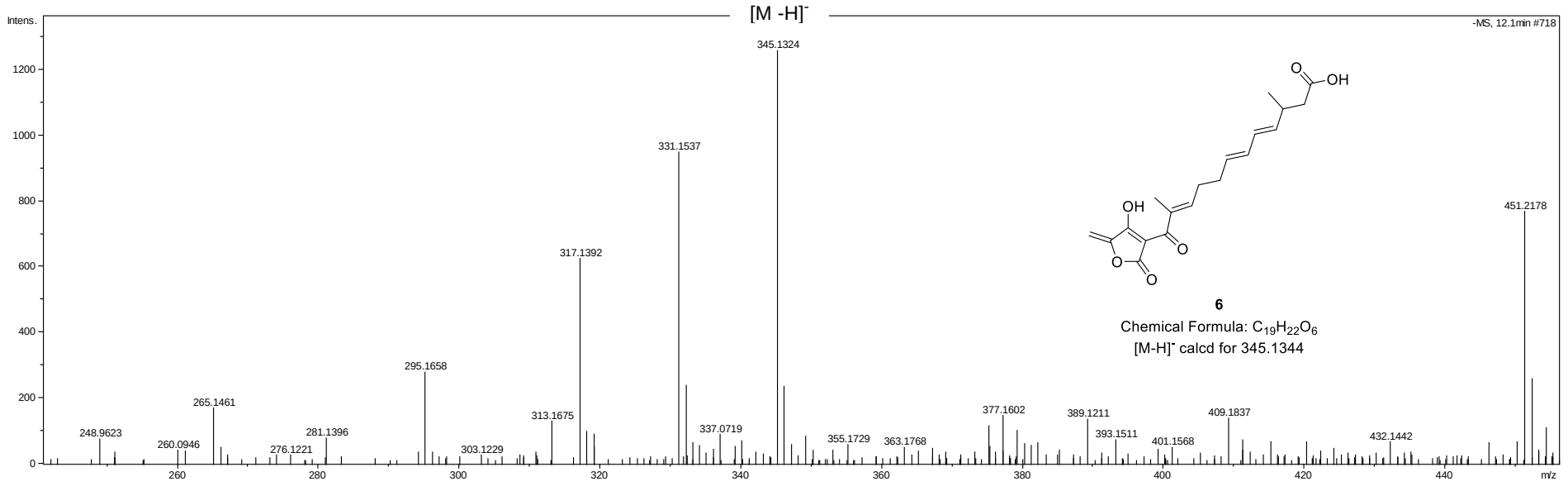


Figure S10. The spectroscopic data of **6**

(B) The ^{1}H-NMR spectrum of **6** (500 MHz for 1H NMR in CD_3OD)

Figure S10. The spectroscopic data of **6**

(C) The ^{13}C NMR spectrum of compound **6** in CD_3OD (125 MHz)

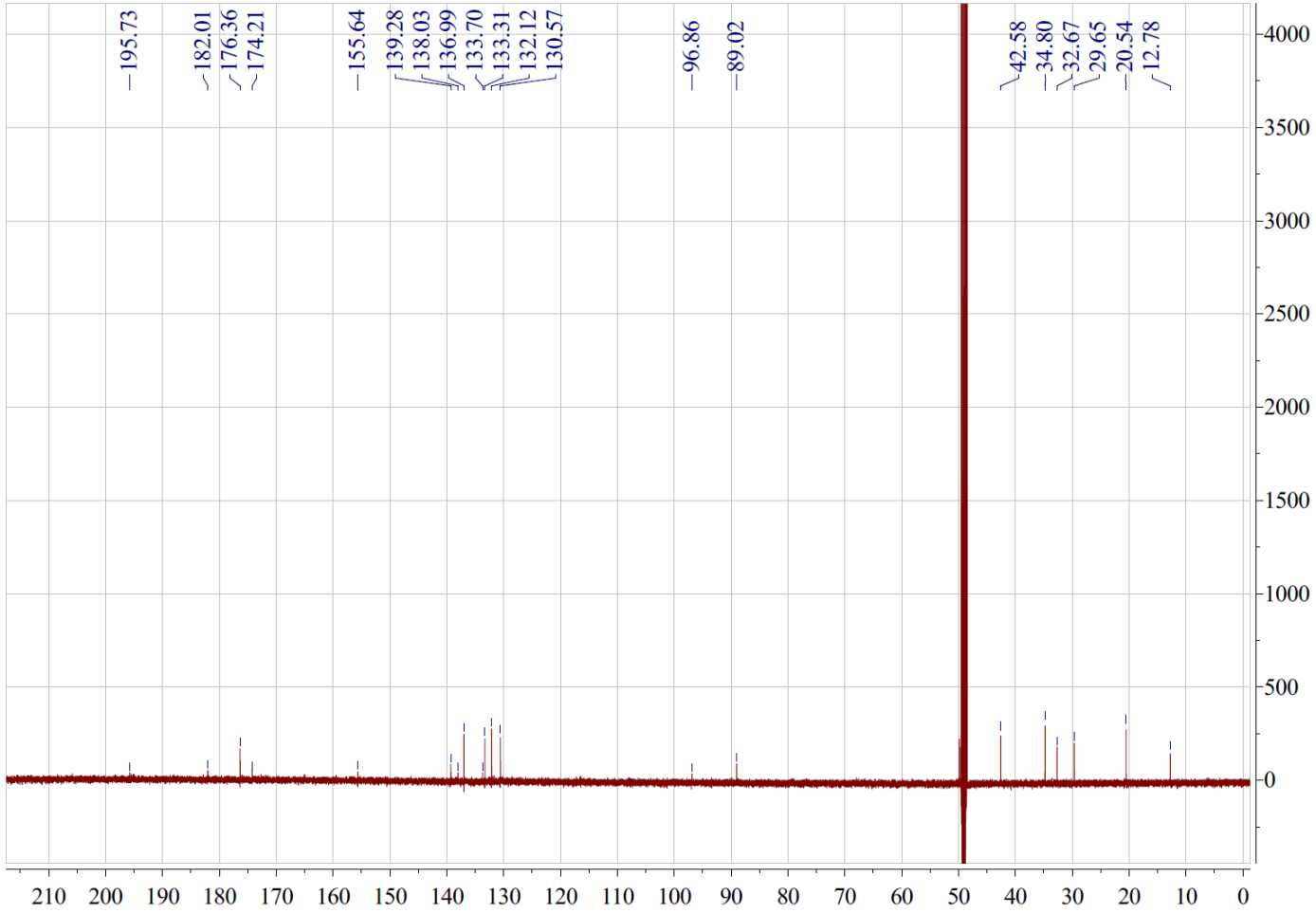


Figure S10. The spectroscopic data of **6**

(D) The HSQC spectrum of compound **6** in CD_3OD (125 MHz)

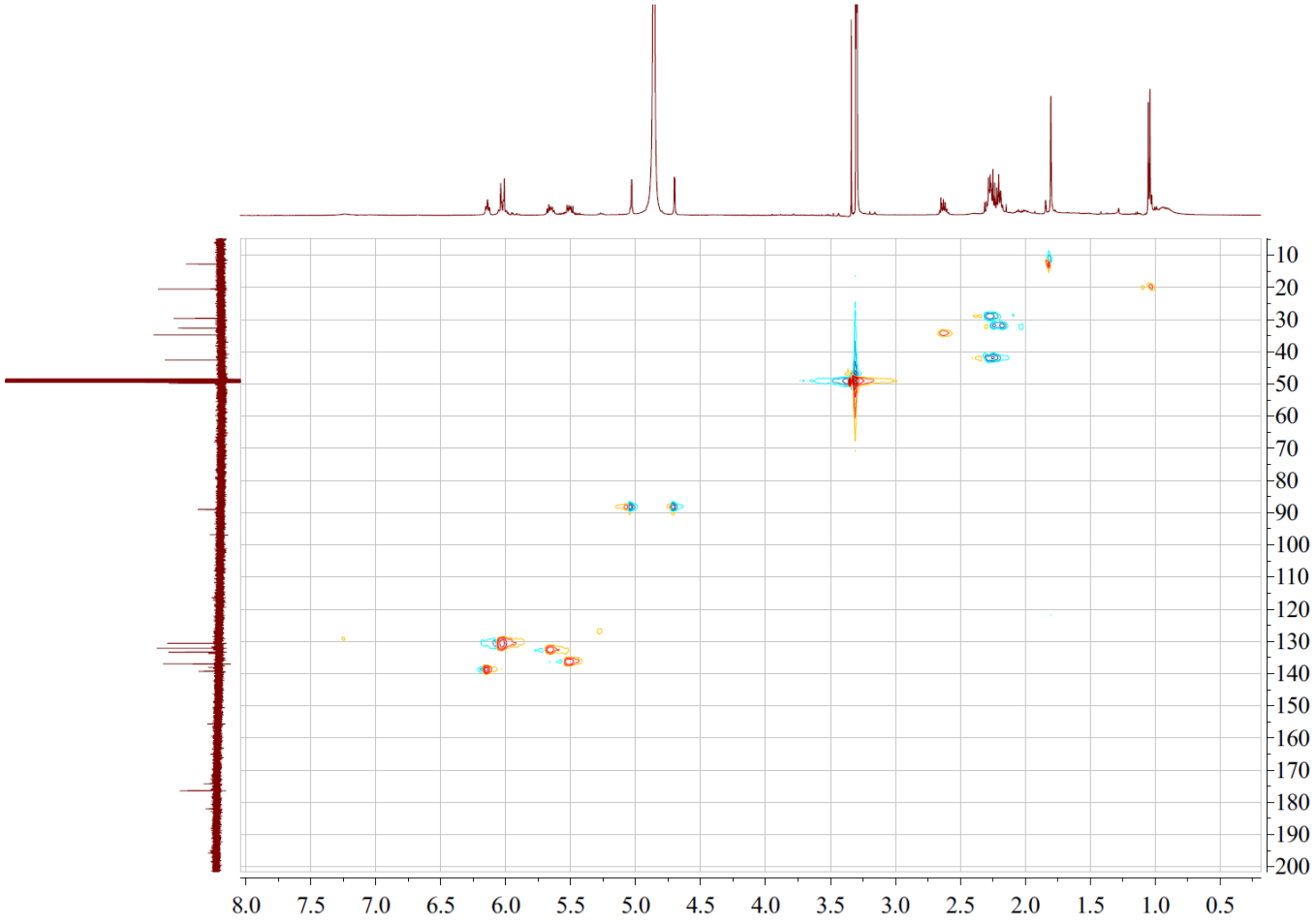


Figure S10. The spectroscopic data of **6**

(E) The HMBC spectrum of **6** (500 MHz for ^{1}H NMR in CD_3OD)

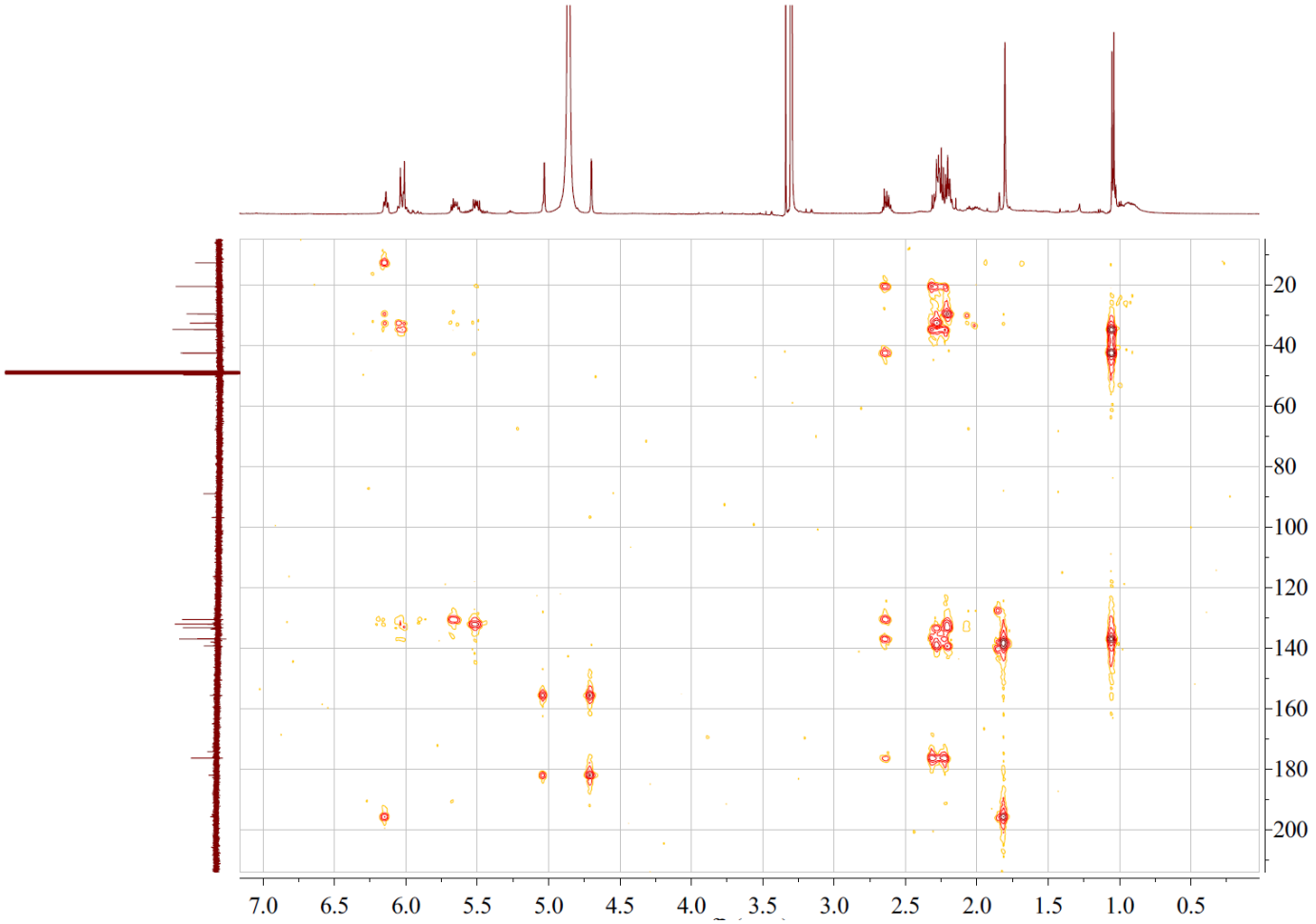


Figure S10. The spectroscopic data of **6**

(F) The ^{1}H-^{1}H COSY spectrum of **6** (500 MHz for ^{1}H NMR in CD_3OD)

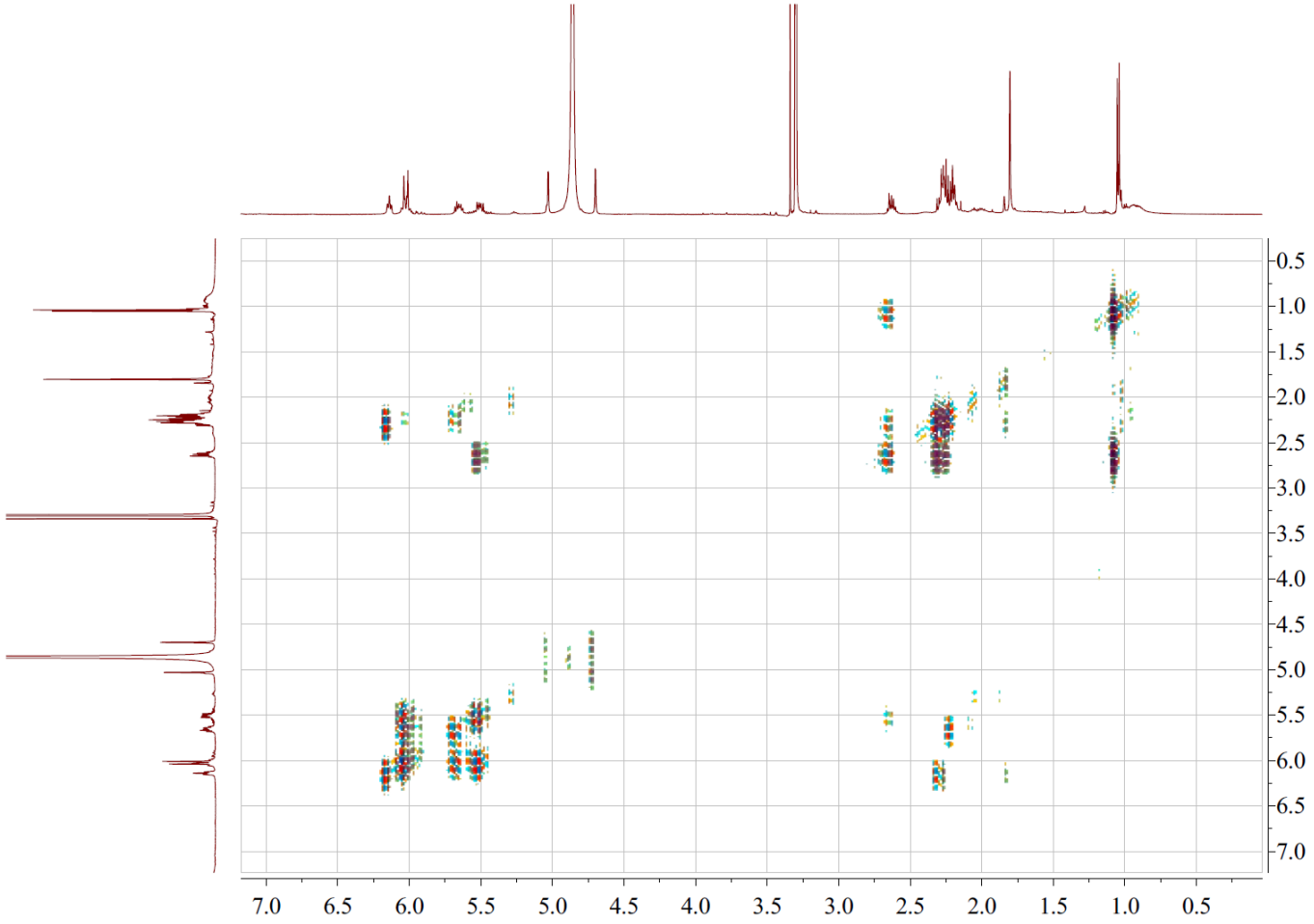
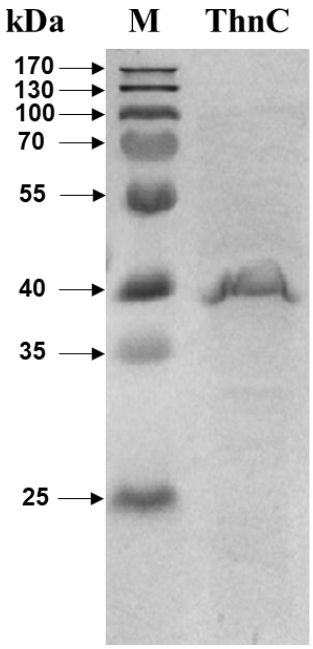


Figure S11. Bioinformatic analysis of ThnC amino acid sequences.

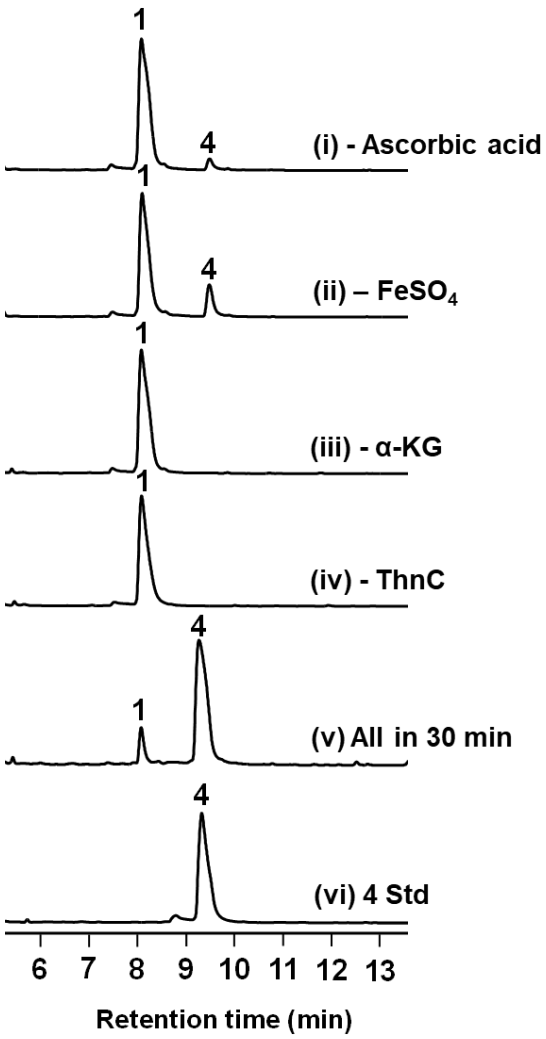
TraH from terrestric acid gene cluster (*Penicillium crustosum*); TropC from tropolone gene cluster (*Talaromyces stipitatus* ATCC 10500); AsL3 from Xenovulene A gene cluster (*Sarocladium* sp. schorii); CitB from Citrinin gene cluster (*Monascus ruber*); VidW from Validamycin B gene cluster (*Streptomyces hygroscopicus* subsp. *limoneus*). The conserved His^1-X-Asp/Glu-X_n-His^2 iron-binding motifs were labeled in red box

Figure S12. SDS-PAGE analysis of purified ThnC.



Purification of C-$(His)_6$-tagged ThnC from *E. coli* BL21(DE3)/pZYG105.The acrylamide percentage of the SDS-PAGE gel is 12 %.

Figure S13. HPLC analysis of in vitro assays of ThnC with substrate **1**



(i) minus ascorbic acid, (ii) minus $FeSO_4$, (iii) minus α-ketoglutarate (α-KG), (iv) minus ThnC, (v) a complete ThnC assay, (vi) **4** std. The ThnC enzyme assay was performed in Tris-HCl buffer (50 mm, pH 8.0) at 28°C containing 200 μM **1**, 2 μM ThnC, 5 mM L-ascorbic acid, 5 mM α-ketoglutarate, 1 mM $FeSO_4$.

References

1. N. Liu, Y. S. Hung, S. S. Gao, L. Hang, Y. Zou, Y. H. Chooi and Y. Tang, *Org Lett*, 2017, **19**, 3560-3563.
2. M. Chen, Q. Liu, S. S. Gao, A. E. Young, S. E. Jacobsen and Y. Tang, *Proc Natl Acad Sci U S A,* 2019, **116**, 5499-5504.
3. T. Nayak, E. Szewczyk, C. E. Oakley, A. Osmani, L. Ukil, S. L. Murray, M. J. Hynes, S. A. Osmani and B. R. Oakley, *Genetics*, 2006, **172**, 1557-1566.