

Supporting Information

Genomics-Driven Discovery of Benzoxazole Alkaloids from the Marine-Derived *Micromonospora* sp. SCSIO 07395

Ziqian Cheng ^{1,2,†}, Qingbo Zhang ^{1,2,3,4,†}, Jing Peng ^{1,2,†}, Xiaoyang Zhao ^{1,2}, Liang Ma ^{1,3},
Changsheng Zhang ^{1,2,3,4,*}
and Yiguang Zhu ^{1,2,3,4,*}

- ¹ Key Laboratory of Tropical Marine Bio-Resources and Ecology, Guangdong Key Laboratory of Marine Materia Medica, South China Sea Institute of Oceanology, Innovation Academy for South China Sea Ecology and Environmental Engineering, Chinese Academy of Sciences, 164 West Xingang Road, Guangzhou 510301, China
- ² University of Chinese Academy of Sciences, 19 Yuquan Road, Beijing 100049, China
- ³ Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), No.1119, Haibin Road, Nansha District, Guangzhou 511458, China
- ⁴ Sanya Institute of Ocean Eco-Environmental Engineering, Yazhou Scientific Bay, Sanya 572000, China
- * Correspondence: czhang@scsio.ac.cn (C.Z.); ygzhu@scsio.ac.cn (Y.Z.); Tel.: +86-20-8902-3038 (C.Z. and Y.Z.)
- † These authors contribute equally to this work.

Table of Contents

Table S1. Deduced functions of individual <i>orfs</i> in the <i>mich</i> gene cluster	3
Table S2. Strains and plasmids used and generated in this study	4
Table S3. Primers used in this study.....	5
Table S4. The antibacterial activities of compound 1-6 (MIC, $\mu\text{g}\cdot\text{mL}^{-1}$)	6
Figure S1 The representative biosynthetic pathway of benzoxazole	7
Figure S2. The spectroscopic data of 1	8
Figure S3. The spectroscopic data of 2	15
Figure S4. The spectroscopic data of 3	23
Figure S5. The spectroscopic data of 4	30
Figure S6. The spectroscopic data of 5	37
Figure S7. The spectroscopic data of 6	45
References.....	52

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

Gene	Size (aa)	Proposed protein function	Orthologue in A33853 cluster (% deduced protein identity)	Orthologue in caboxamycin cluster (% deduced protein identity)	Orthologue in nataxazole cluster (% deduced protein identity)
<i>orf-2</i>	381	Probable acyl-CoA dehydrogenase			
<i>orf-1</i>	188	Siroheme synthase			
<i>michA</i>	494	amidohydrolase	<i>bomN</i> (60%)	<i>cbxE</i> (58%)	<i>natAM</i> (60%)
<i>michB</i>	254	2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase	<i>bomO</i> (61%)	<i>cbxG</i> (63%)	<i>natDB</i> (61%)
<i>michC</i>	215	Isochorismatase	<i>bomP</i> (55%)	<i>cbxH</i> (63%)	<i>natIS</i> (59%)
<i>michD</i>	659	Anthranilate synthase	<i>bomQ</i> (64%)	<i>cbxI</i> (58%)	<i>natAN</i> (57%)
<i>michE</i>	427	coenzyme A ligase	<i>bomJ</i> (64%)	<i>cbxC</i> (57%)	<i>NatL2</i> (63%)
<i>michF</i>	349	O-methyltransferase			
<i>michR</i>	224	YdeI/OmpD-associated family protein			
<i>orf1</i>	140	50S ribosomal protein			
<i>orf2</i>	275	Arylamine N-acetyltransferase			
<i>orf3</i>	522	aldehyde dehydrogenase			
<i>orf4</i>	177	SRPBCC family protein			
<i>orf5</i>	341	ketoacyl-ACP synthase III family protein			
<i>michG</i>	362	DAHPh synthase			
<i>orf6</i>	776	non-ribosomal peptide synthetase			

Table S2. Strains and plasmids used and generated in this study

Strains/Plasmids	Characteristics	Sources
<i>Escherichia coli</i>		
DH10B	Host strain for BAC library construction	Giboco BRL
ET12567/pUB307	Host strain for conjugation	4 1
<i>Actinomycetes</i>		
<i>Micromonospora</i> sp. SCSIO 07395	Wild type	This study
<i>S. albus</i> J1074	Host strain for heterologous expression	3 2
<i>S. albus</i> Del14	Host strain for heterologous expression	13
<i>S. albus</i> J1074/pMSBBAC2	<i>S. albus</i> J1074 containing plasmid pMSBBAC2	This study
<i>S. albus</i> J1074/pCSG8103	<i>S. albus</i> J1074 containing plasmid pCSG8103	This study
<i>S. albus</i> Del14/pMSBBAC2	<i>S. albus</i> Del14 containing plasmid pMSBBAC2	This study
<i>S. albus</i> Del14/pCSG8103	<i>S. albus</i> Del14 containing plasmid pCSG8103	This study
Plasmids		
pMSBBAC2	Apr ^r , the vector for BAC library construction	2 4
pCSG8103	Apr ^r , BAC plasmid containing entire microechmycin (<i>mich</i>) BGC	This study

Table S3. Primers used in this study

8-F1	CGTCGTCCGGCCACCCATTGATA
8-R1	CCCCCGGCCGCATTACTGTGAA
8-F2	GCCACCGCAGTCCACCGTCATTTC
8-R2	CCAGGGCCCCGGATCAGGTTGTTCA

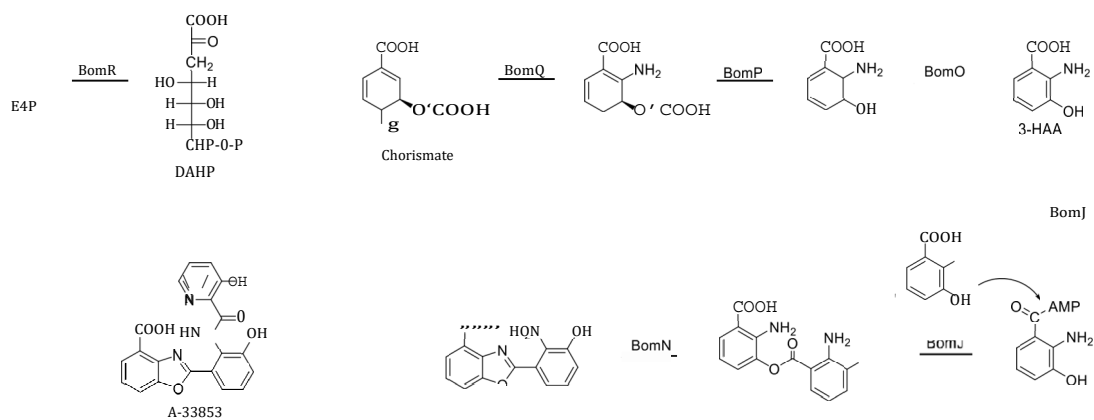
Table S4. The antibacterial activities of compound **1-6** (MIC, $\mu\text{g}\cdot\text{mL}^{-1}$)

Compounds	<i>S. aureus</i> ATCC 29213	<i>M. luteus</i> SCSIO ML01	<i>E. coli</i> ATCC 25922	<i>V. alginolyticus</i> 13214	<i>A.baumannii</i> ATCC 19606	<i>K. pneumoniae</i> ATCC 13883
1	>64	8	>64	>64	>64	>64
2	>64	>64	>64	>64	>64	>64
3	>64	>64	64	>64	>64	>64
4	>64	>64	>64	>64	>64	>64
5	>64	>64	>64	>64	>64	>64
6	>64	>64	>64	>64	>64	>64
Ciprofloxacin	0.25	1	0.0125	0.03	0.125	0.03

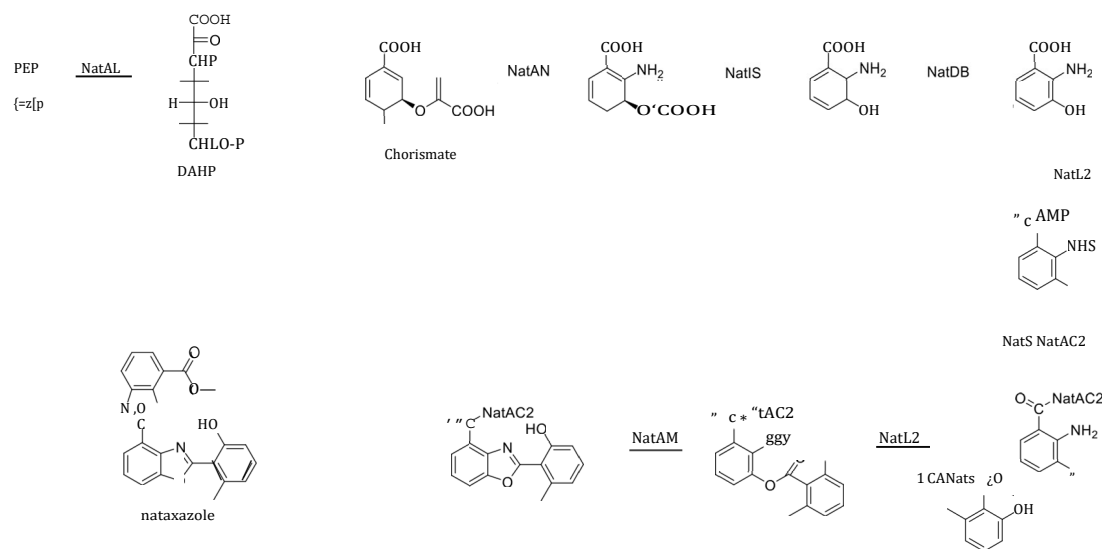
Note: Ciprofloxacin is a positive control.

Figure S1 The representative biosynthetic pathway of benzoxazole

(A) The biosynthesis of A33853



(B) The biosynthesis of nataxazole



(C) The biosynthesis of caboxamycin

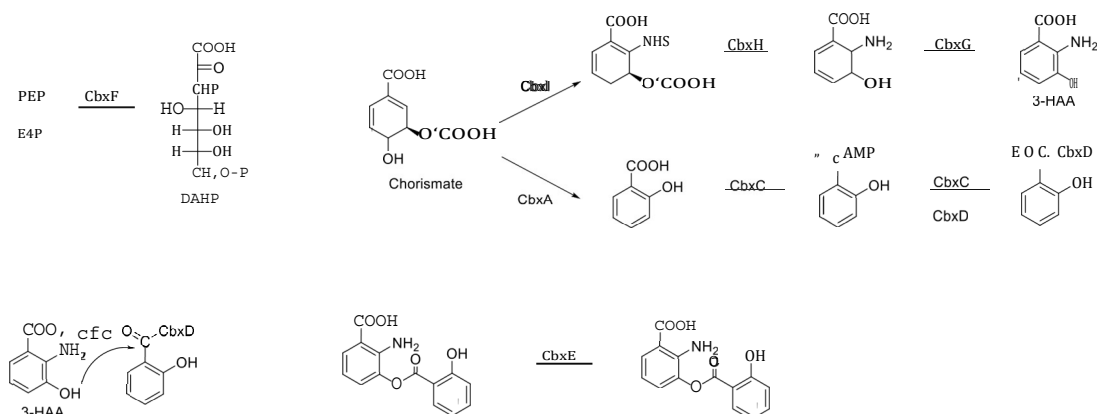
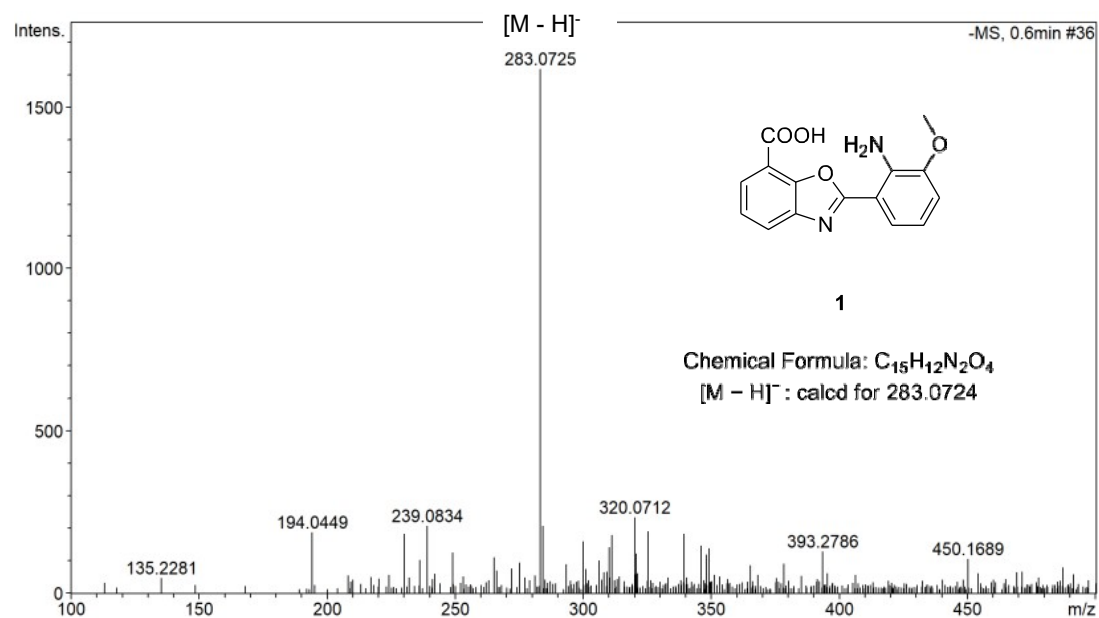


Figure S2. The spectroscopic data of **1**

(A) The HRESIMS (a) UV (b) spectrum of **1**

(a)



(b)

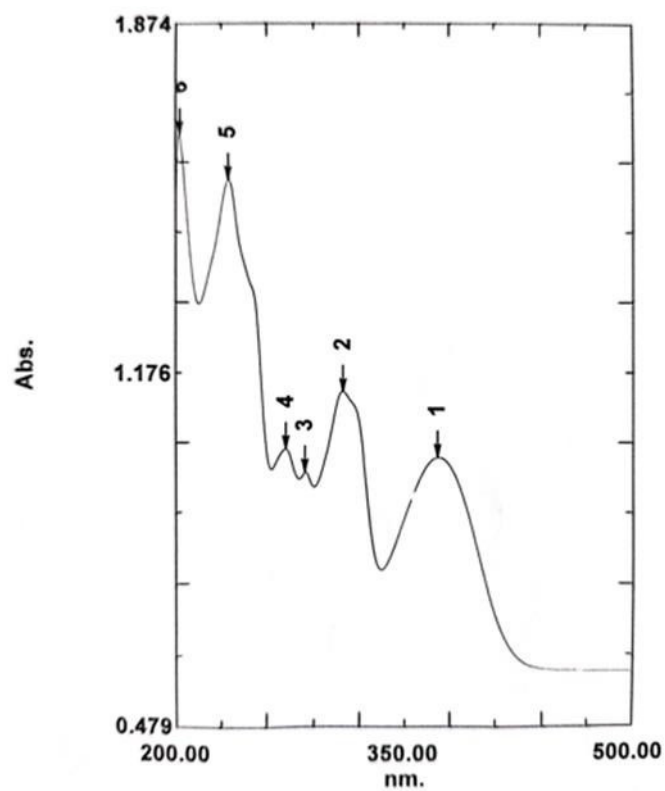


Figure S2. The spectroscopic data of **1**

(B) The ^1H -NMR spectrum of **1** (700 MHz for ^1H NMR in $\text{DMSO}-d_6$)

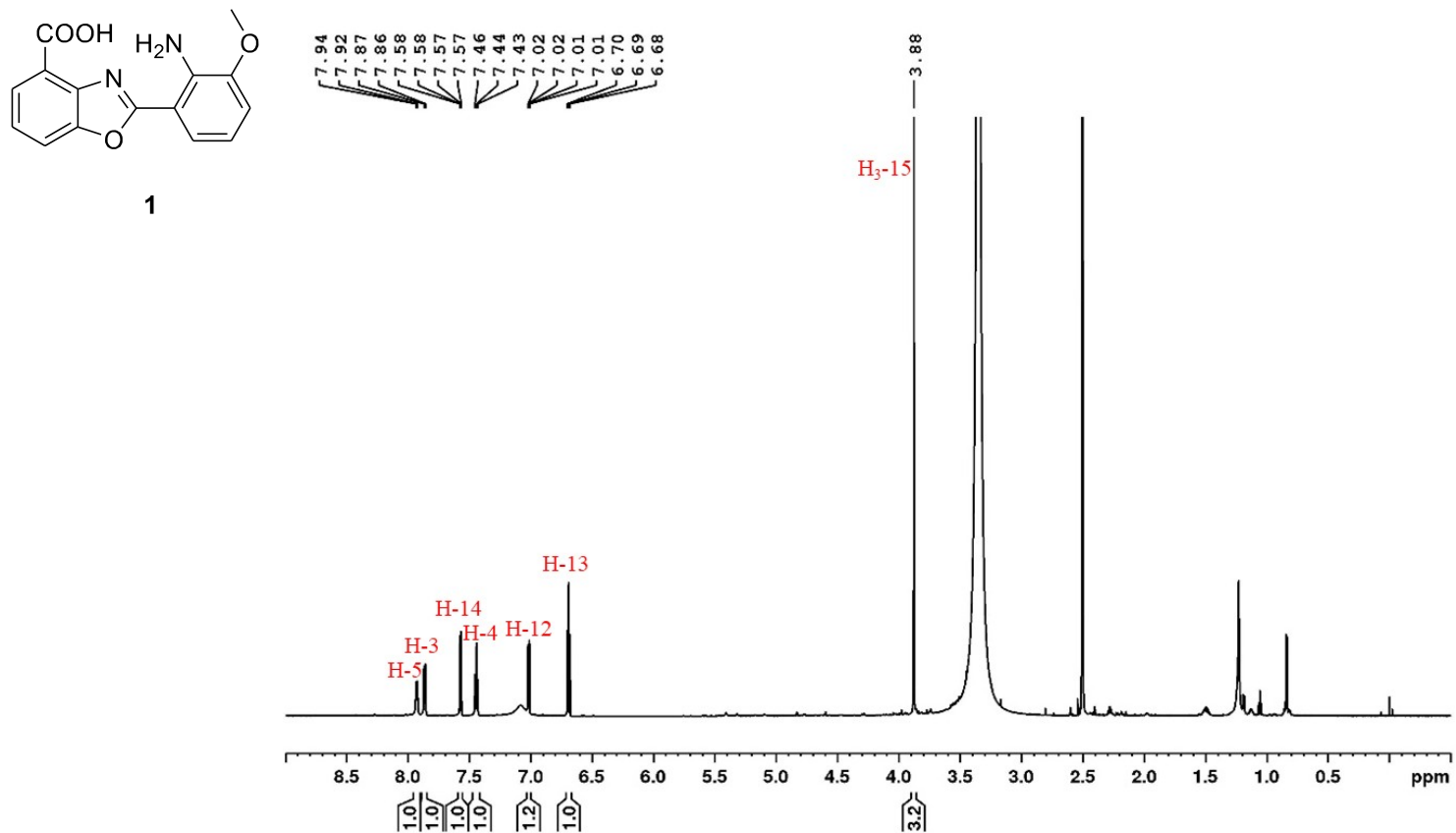


Figure S2. The spectroscopic data of **1**

(C) The ^{13}C -NMR spectrum of compound **1** in $\text{DMSO-}d_6$ (175 MHz)

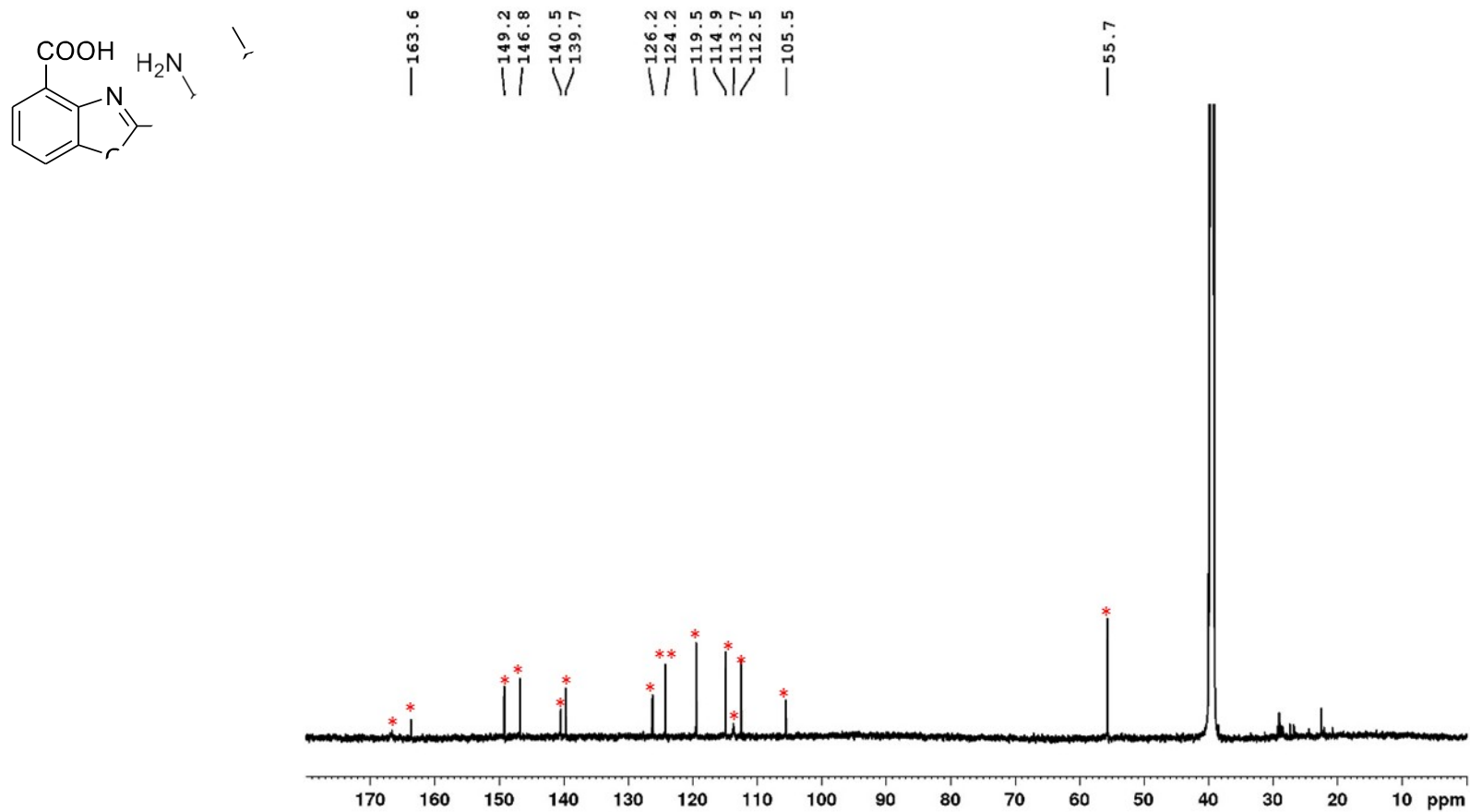


Figure S2. The spectroscopic data of **1**
(D) The DEPT135 spectrum of **1** (175MHz, DMSO-*d*₆)

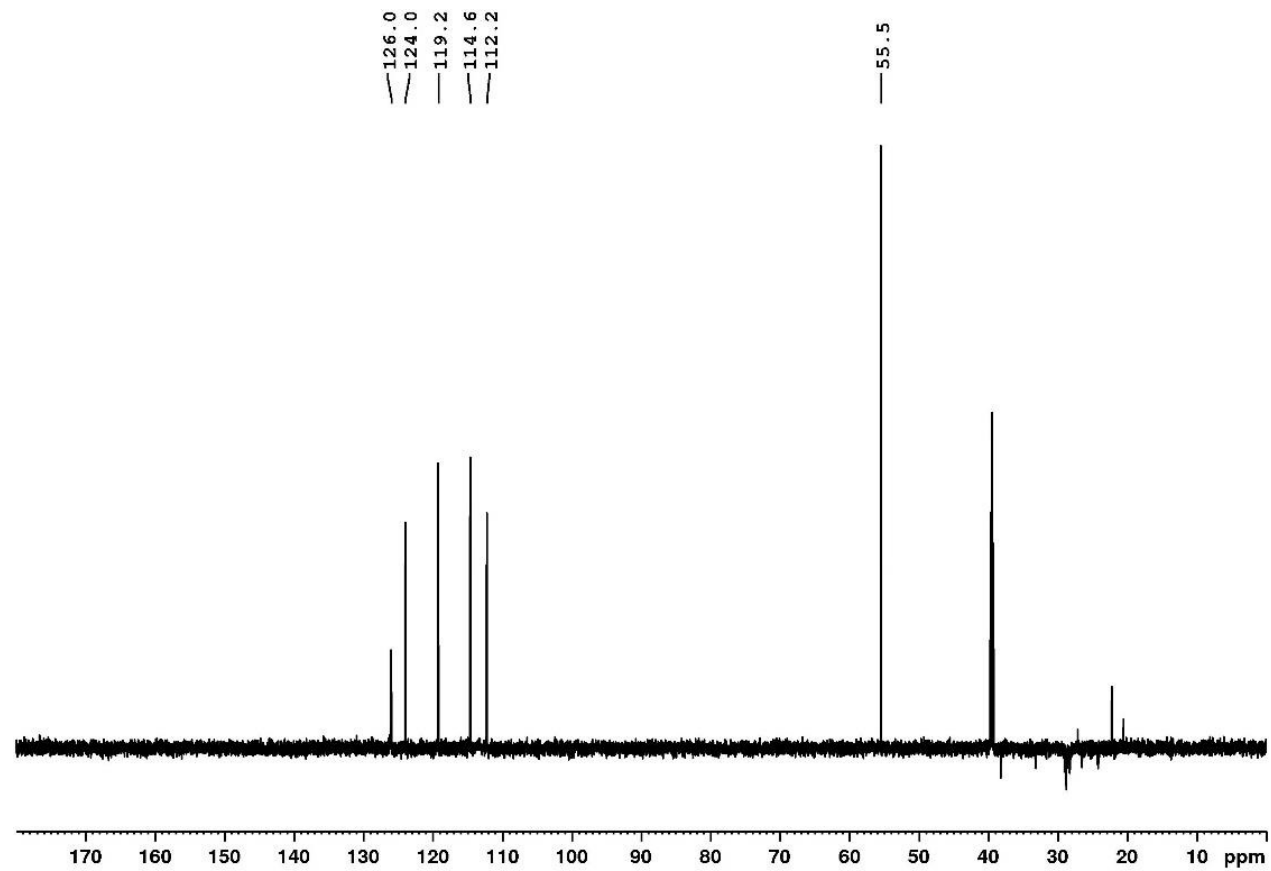
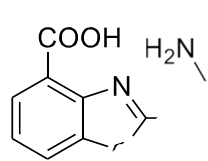
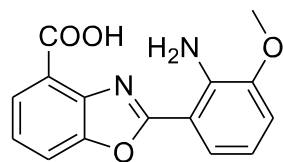


Figure S2. The spectroscopic data of **1**
(E) The HSQC spectrum of compound **1**



1

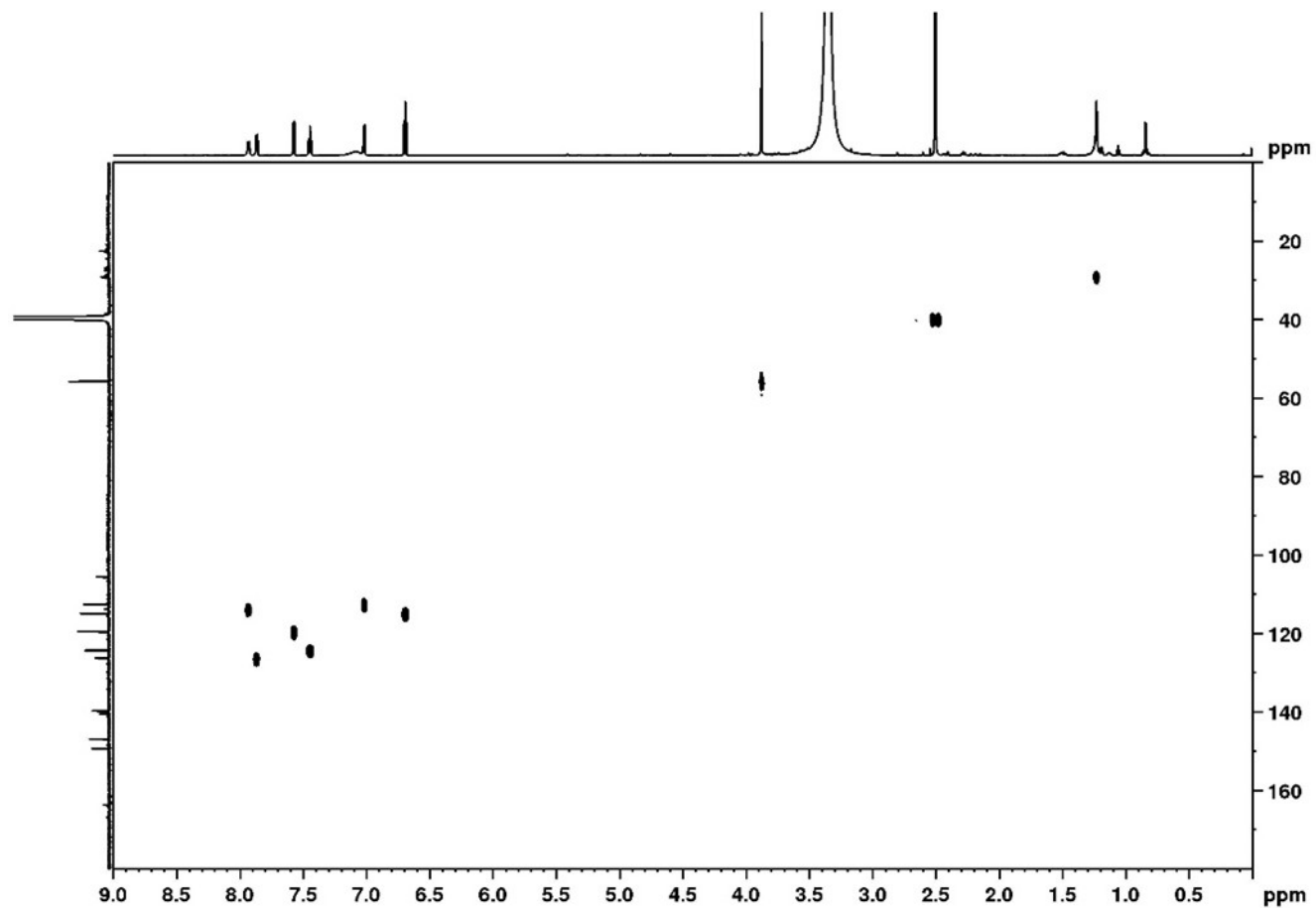


Figure S2. The spectroscopic data of **1**
(F) The HMBC spectrum of **1**

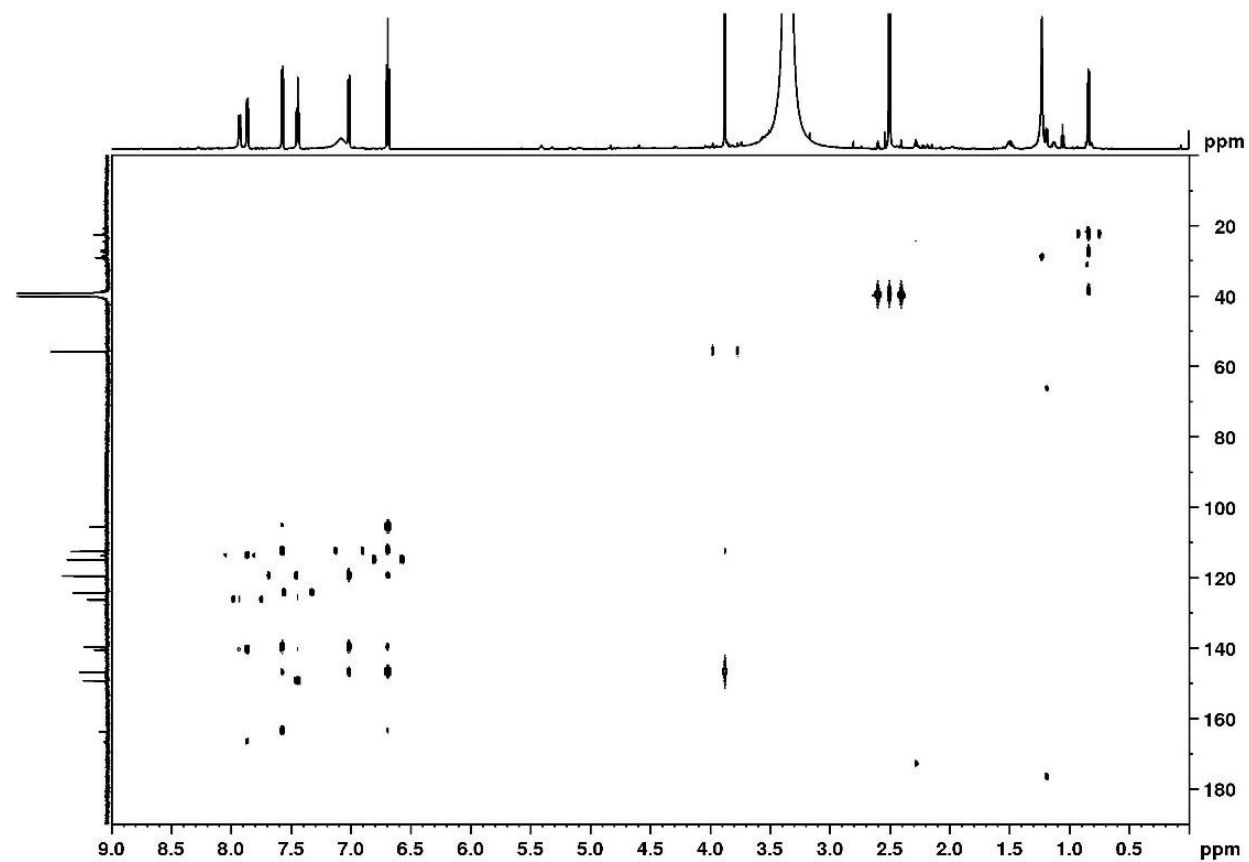
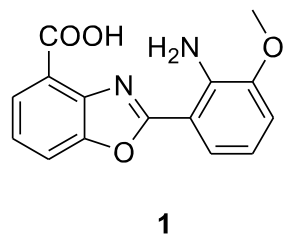
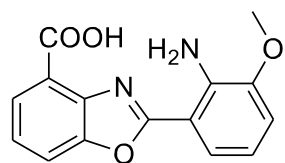


Figure S2. The spectroscopic data of **1**
(G) The ^1H - ^1H COSY spectrum of **1**



1

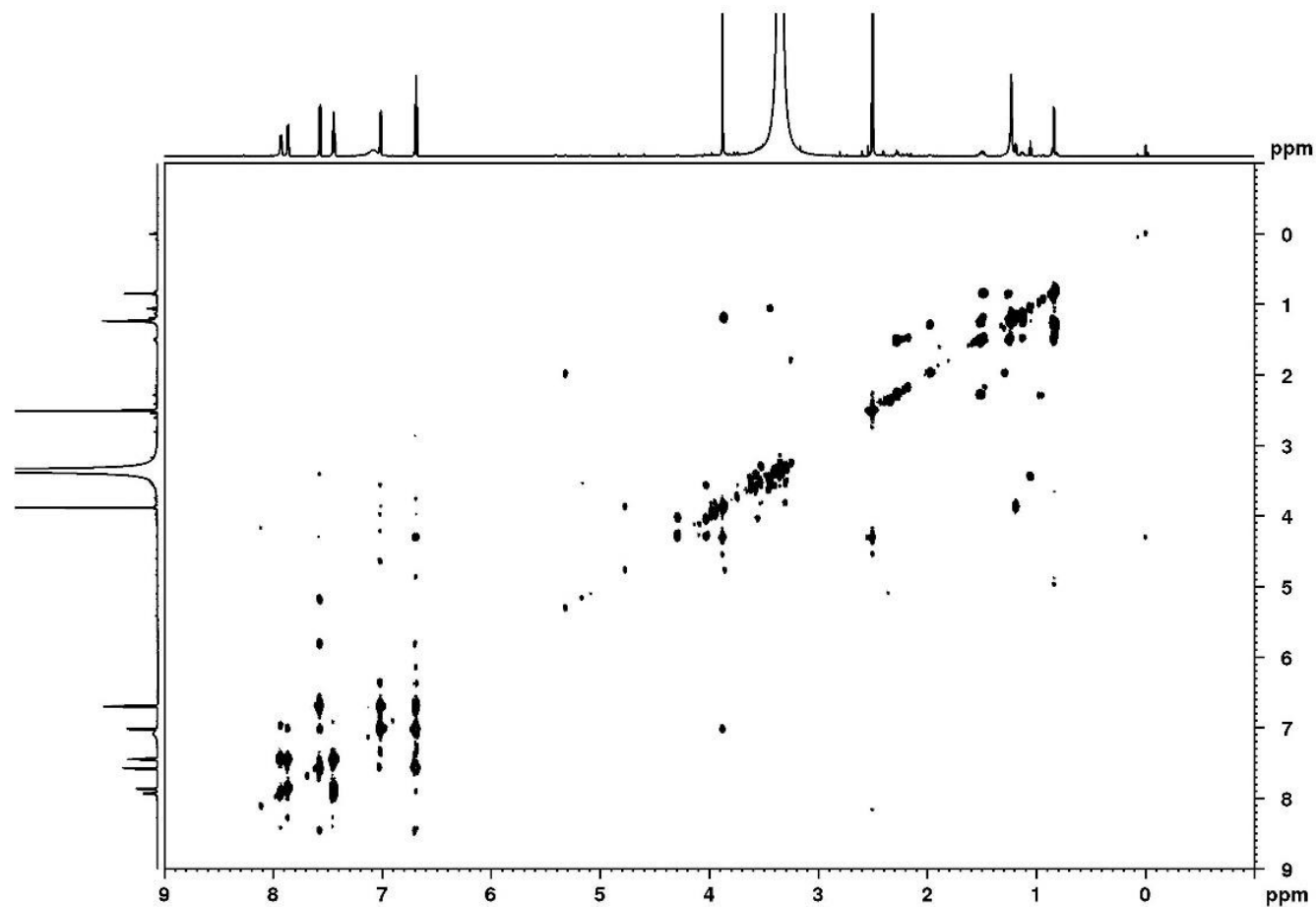
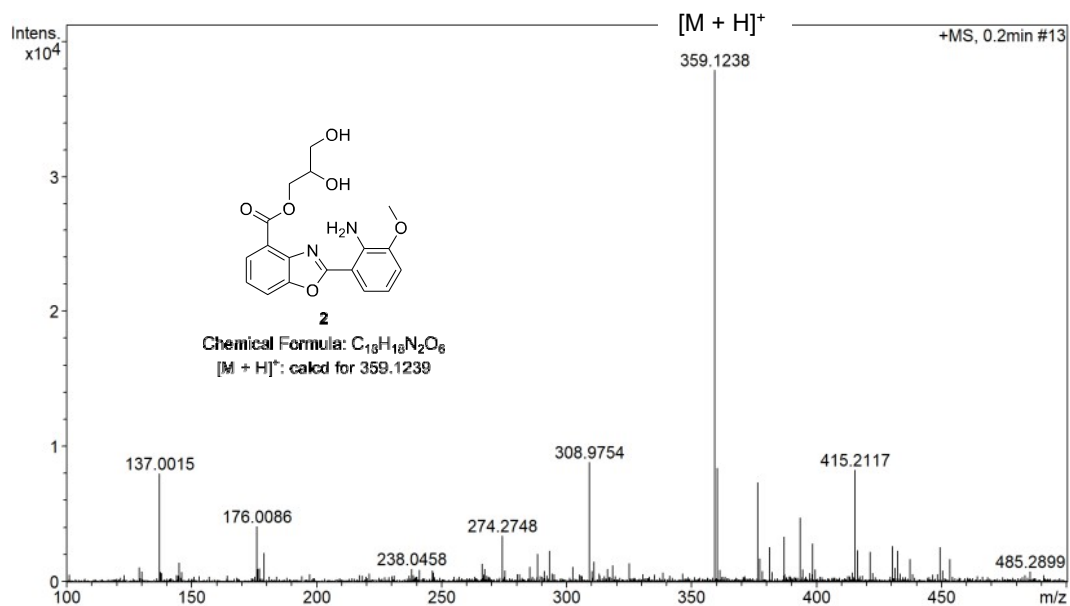


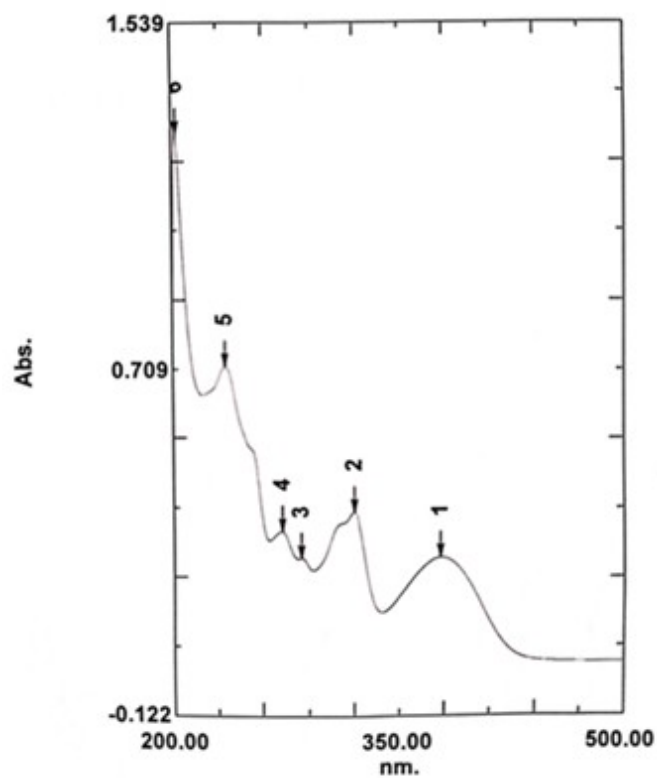
Figure S3. The spectroscopic data of **2**

(A) The HRESIMS (a), UV (b), CD (c) spectrum and chiral HPLC analysis (d) of **2**

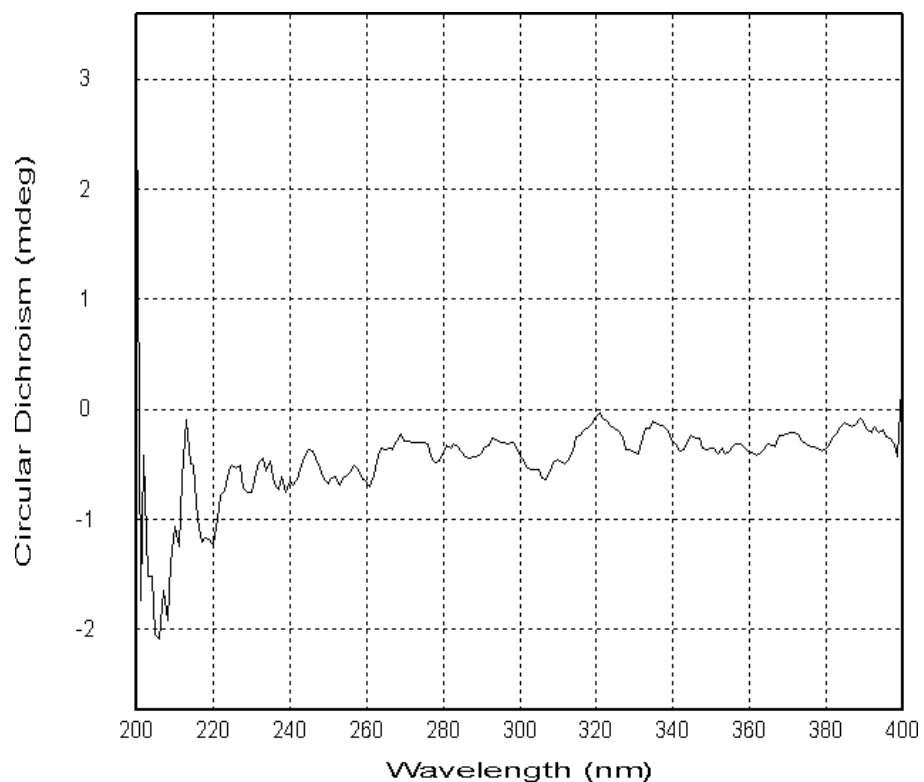
(a)



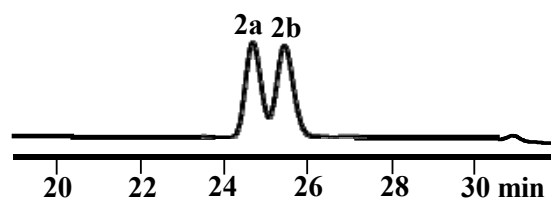
(b)



(c)



(d) Chiral HPLC analysis of compound 2



Chiral HPLC analysis of compound **2** was conducted by using a chiral column (Lux Cellulose-3, 5 μm , 250 \times 4.6 mm, phenomenex) with UV detection at 265 nm under the following program: solvent system (solvent A, 10% CH_3CN in water; solvent B, 90% CH_3CN in water); 38% B (0–35 min), flow rate at 1 mL min^{-1} .

Figure S3. The spectroscopic data of **2**

(B) The ^1H -NMR spectrum of **2** (700 MHz for ^1H NMR in $\text{DMSO}-d_6$)

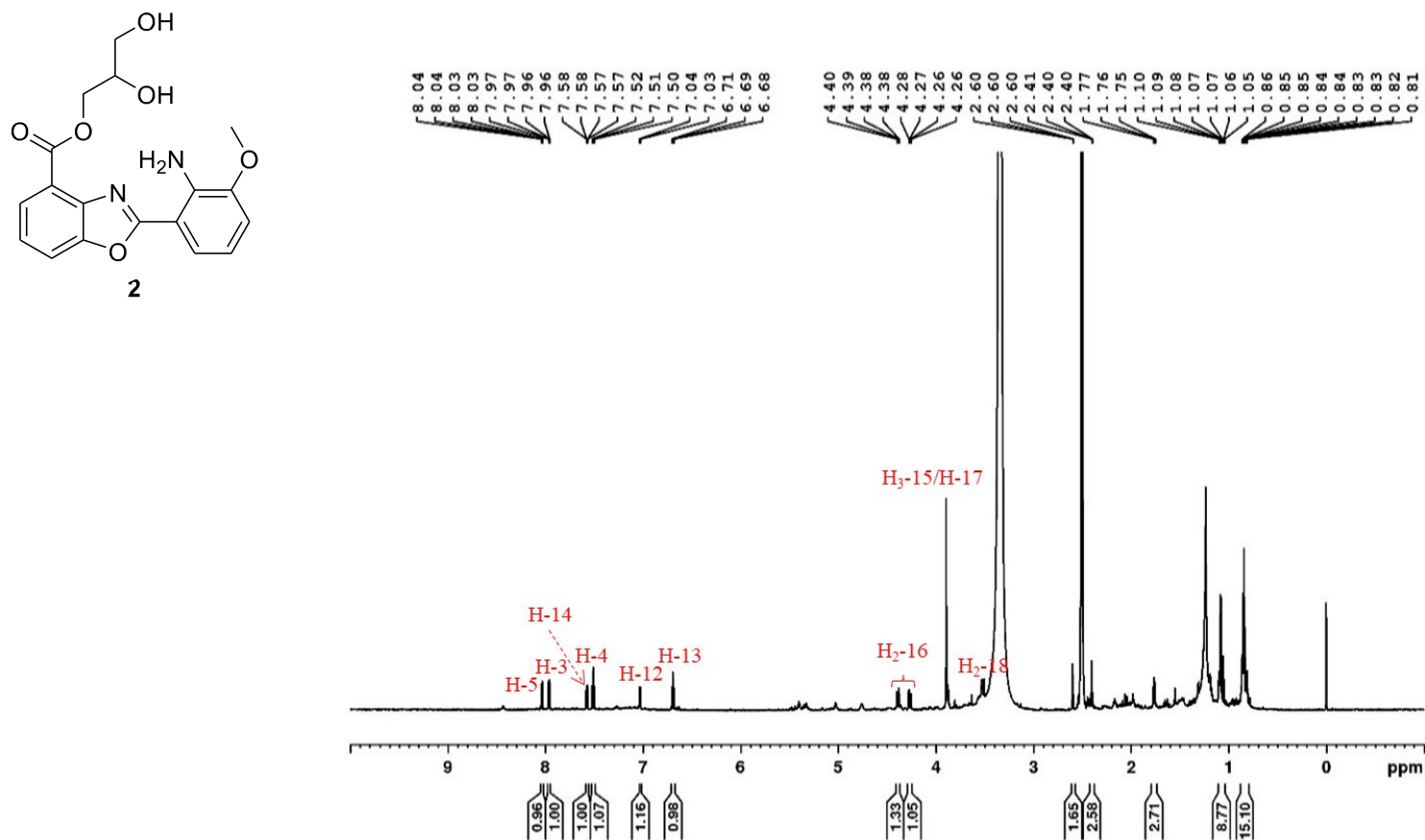


Figure S3. The spectroscopic data of **2**

(C) The ^{13}C NMR spectrum of compound **2** in $\text{DMSO-}d_6$ (175 MHz)

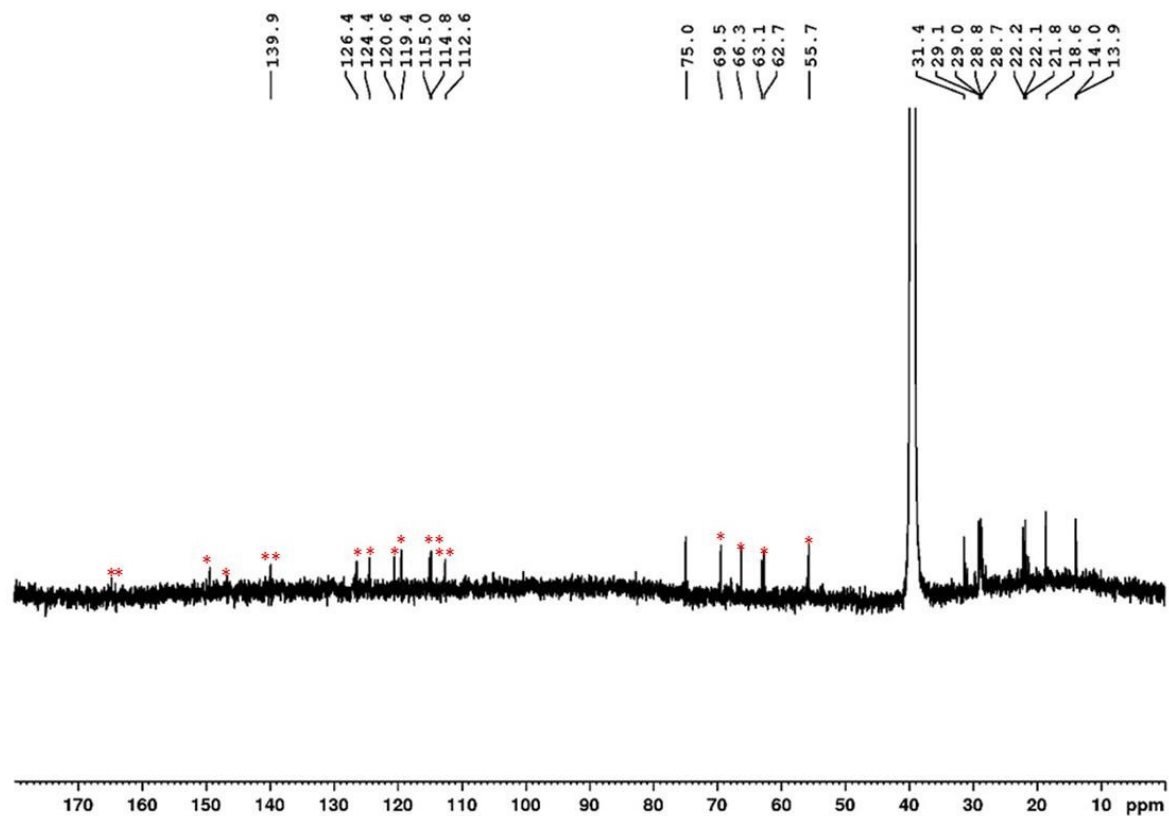
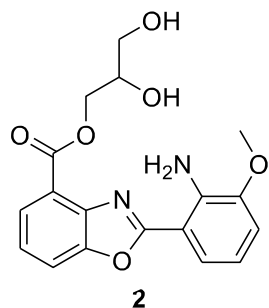


Figure S3. The spectroscopic data of **2**
(D) The DEPT135 spectrum of **2**

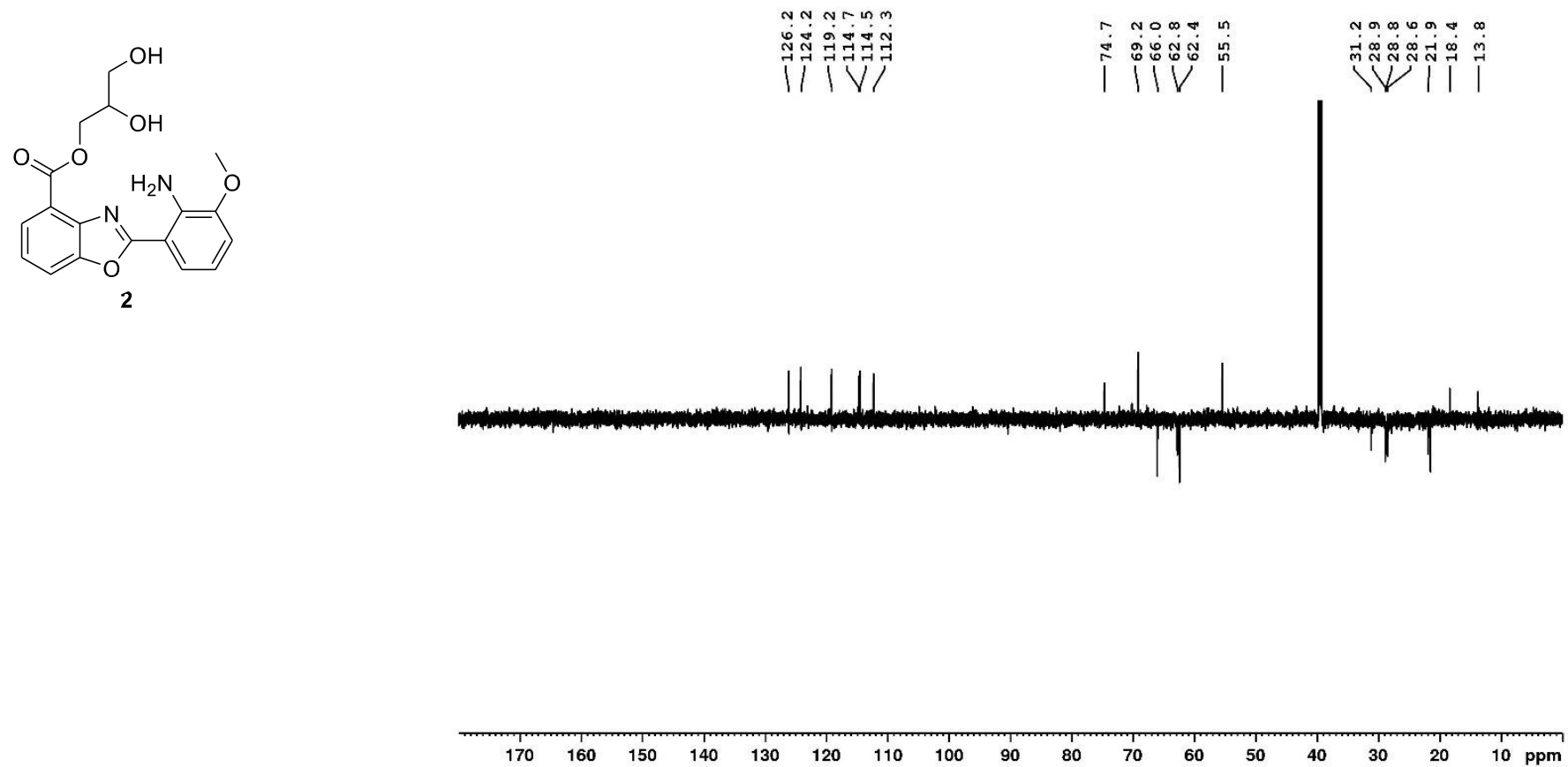


Figure S3. The spectroscopic data of **2**
(E) The HSQC spectrum of compound **2**

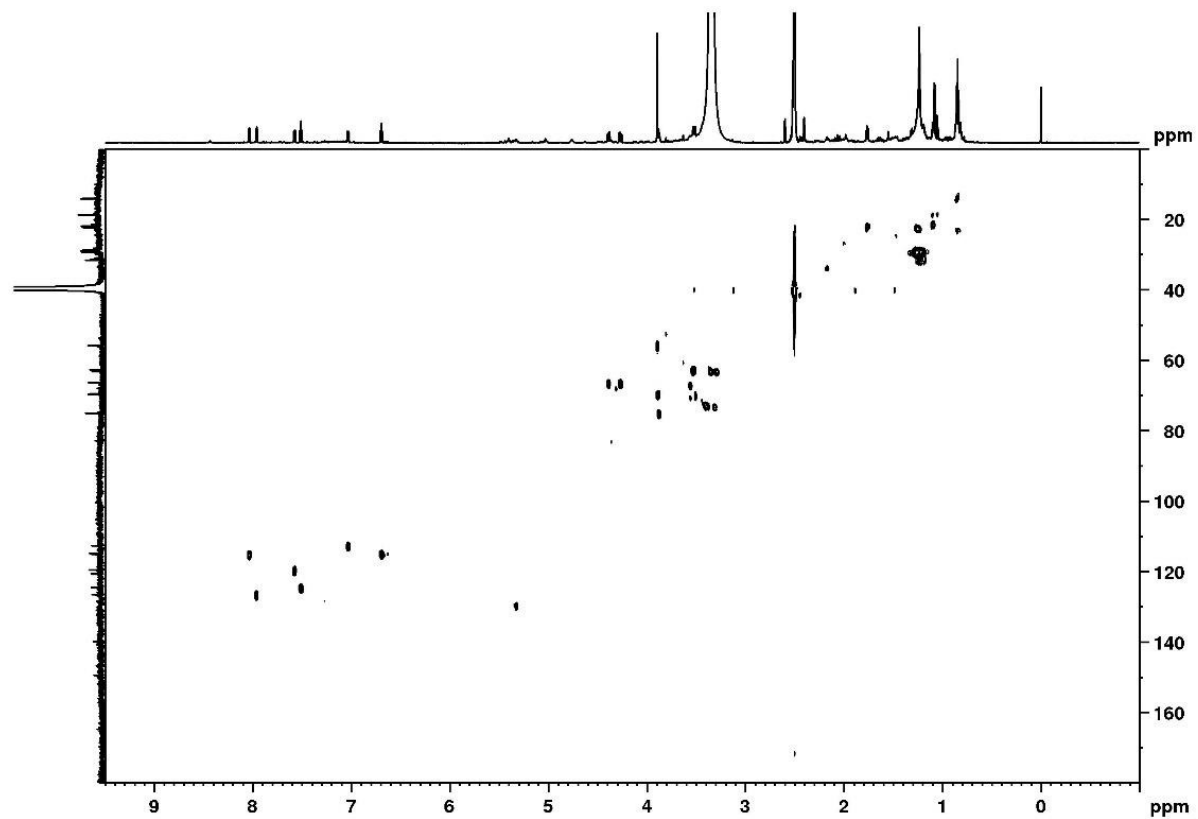
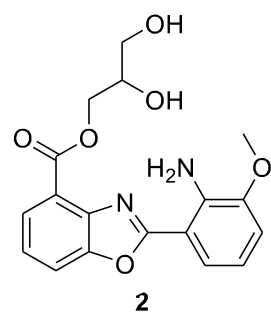


Figure S3. The spectroscopic data of **2**
(F) The HMBC spectrum of **2**

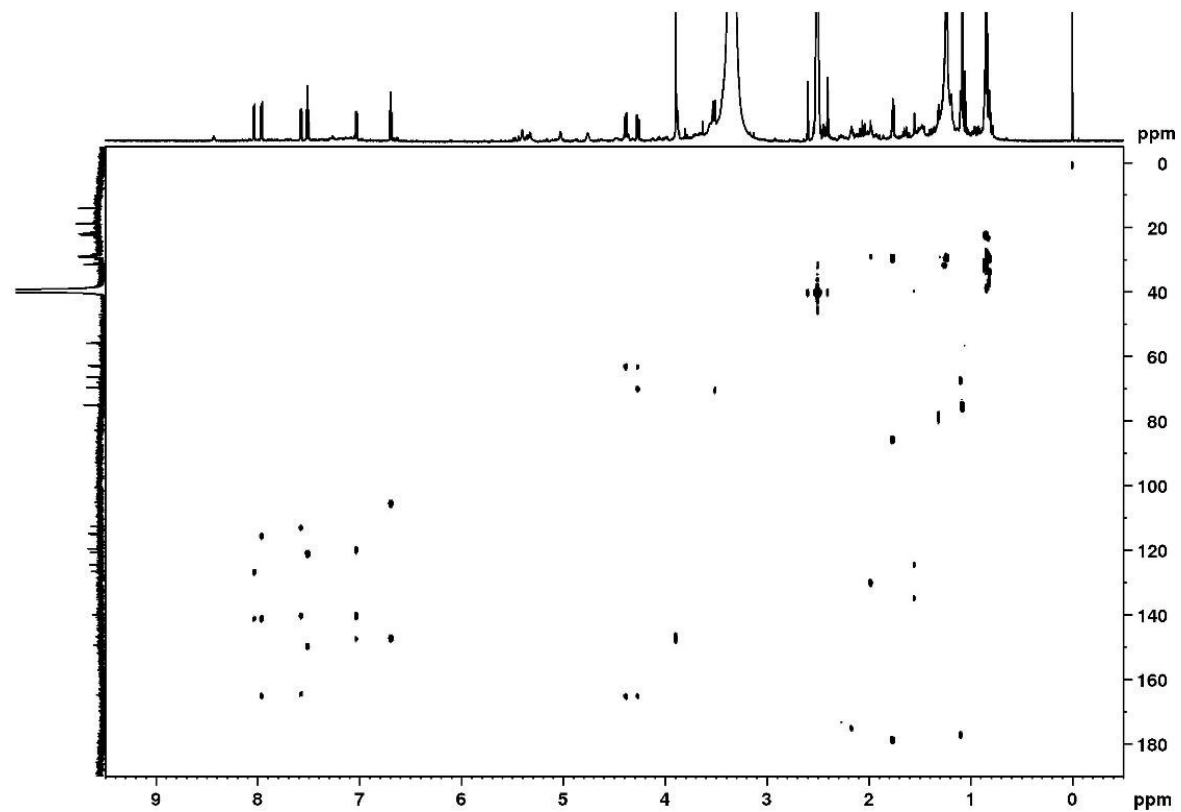
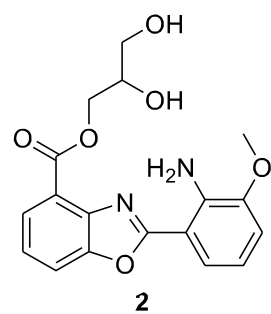


Figure S3. The spectroscopic data of **2**
(G) The ^1H - ^1H COSY spectrum of **2**

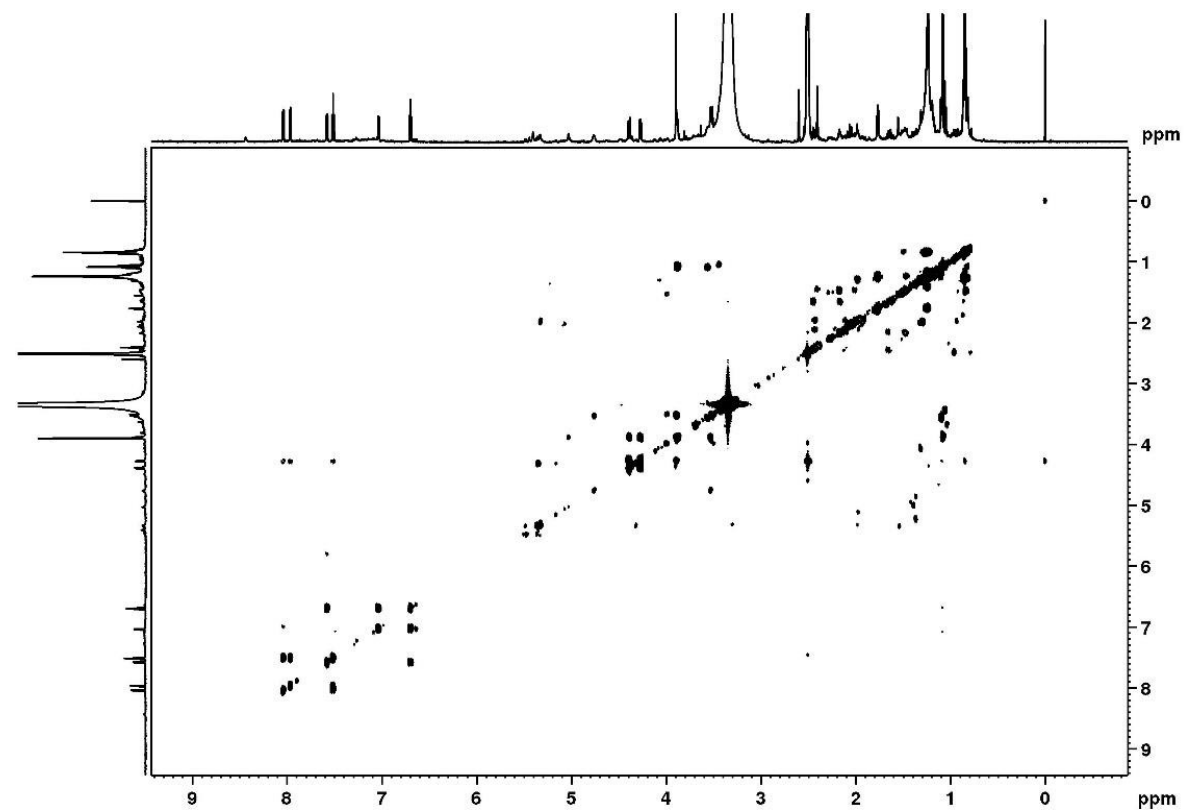
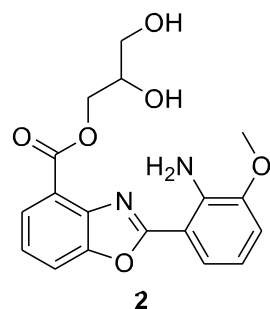
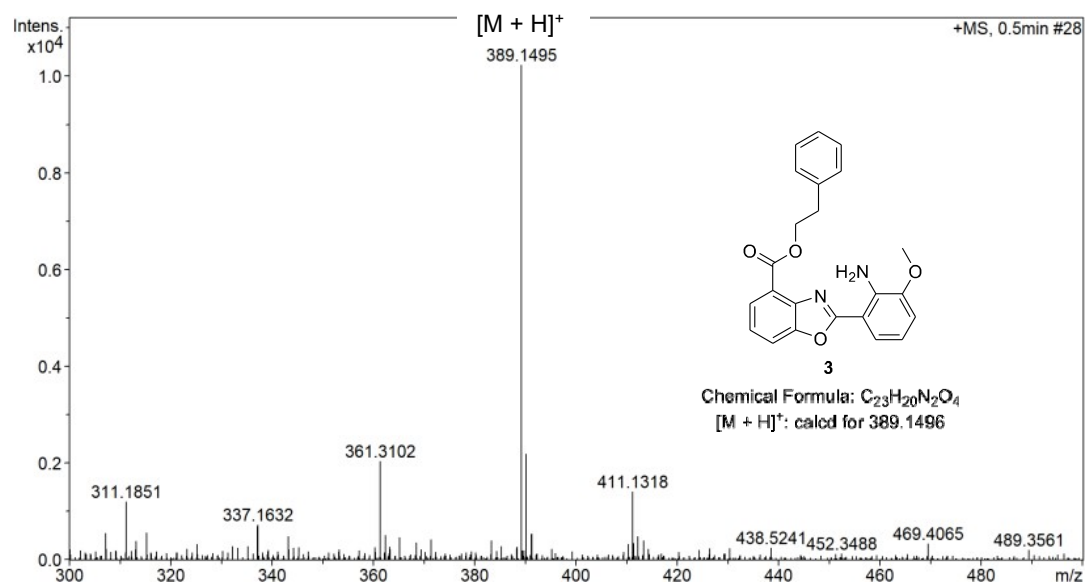


Figure S4. The spectroscopic data of **3**

(A) The HRESIMS (a) UV (b) of **3**

(a)



(b)

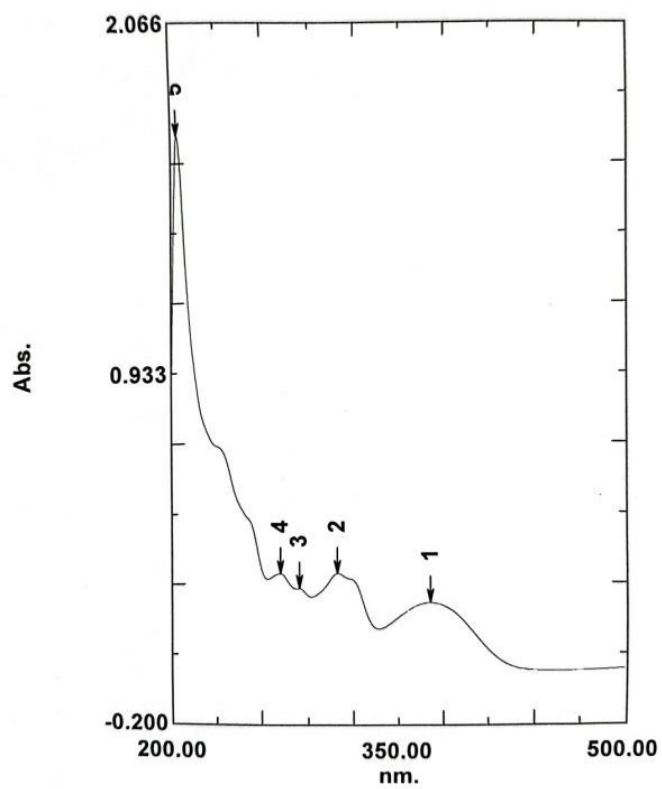


Figure S4. The spectroscopic data of **3**

(B) The ^1H -NMR spectrum of **3** (700 MHz for ^1H NMR in $\text{DMSO}-d_6$)

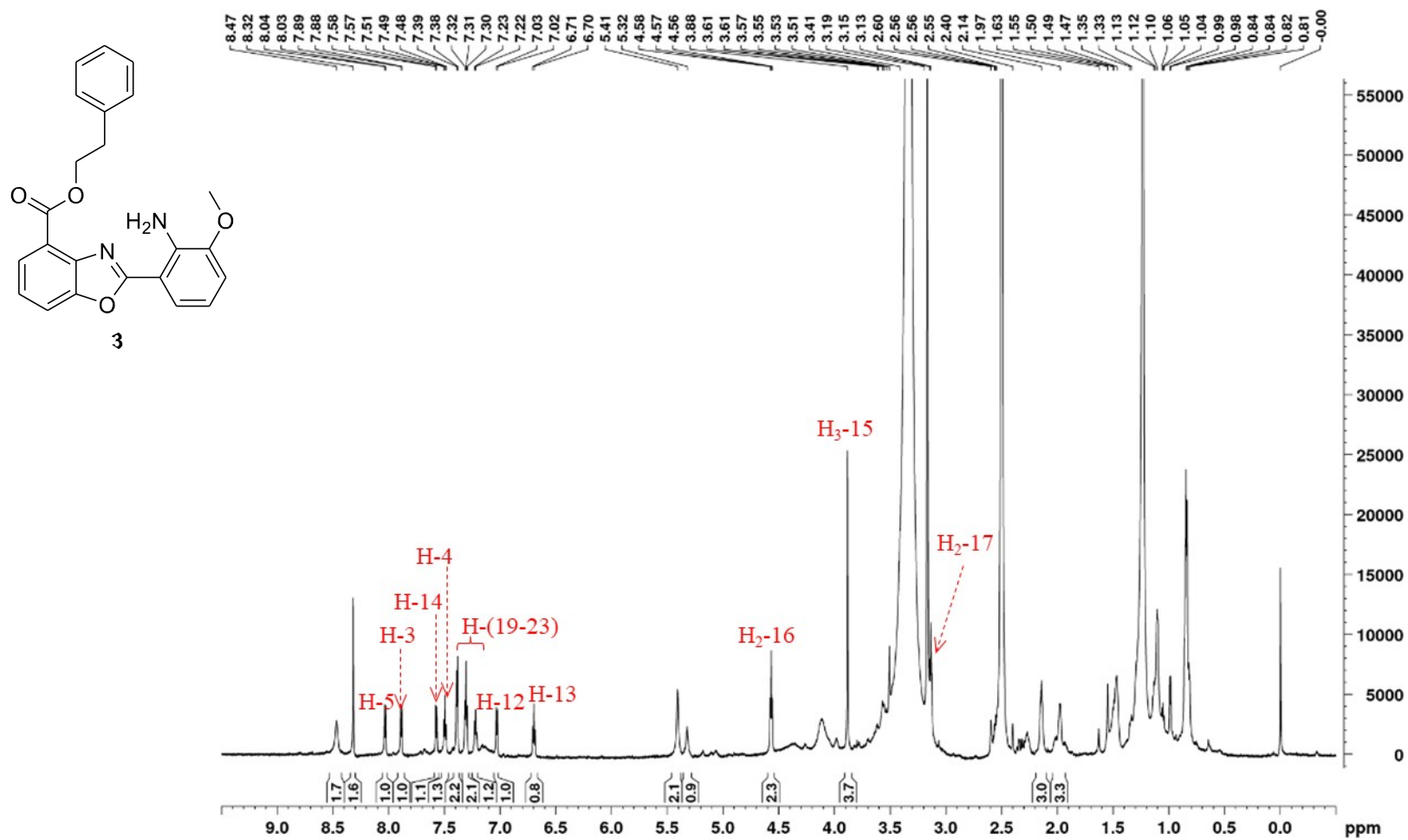


Figure S4. The spectroscopic data of **3**
 (C) The ^{13}C NMR spectrum of compound **3** in $\text{DMSO}-d_6$ (175 MHz)

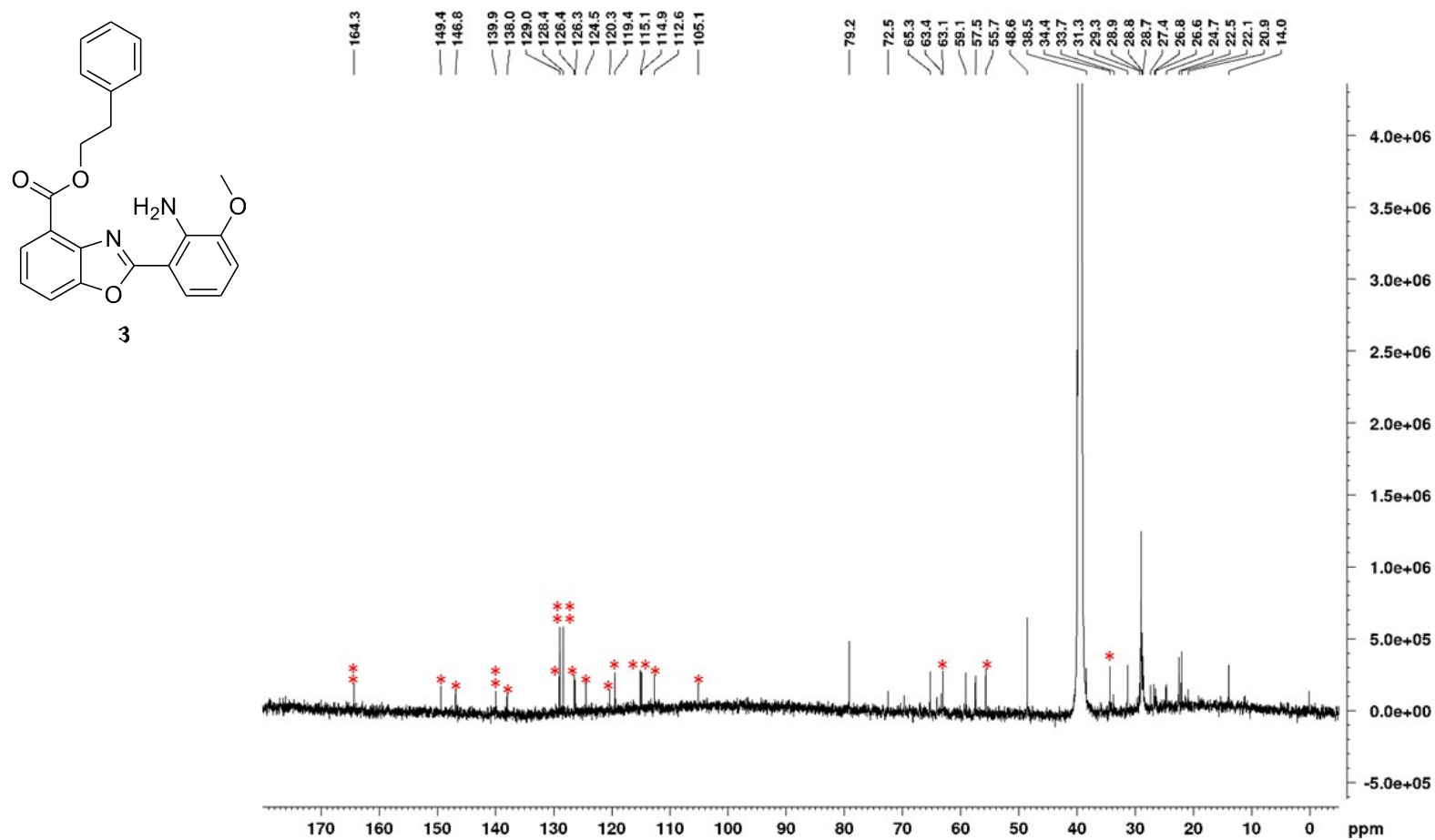


Figure S4. The spectroscopic data of **3**
(D) The DEPT135 spectrum of **3**

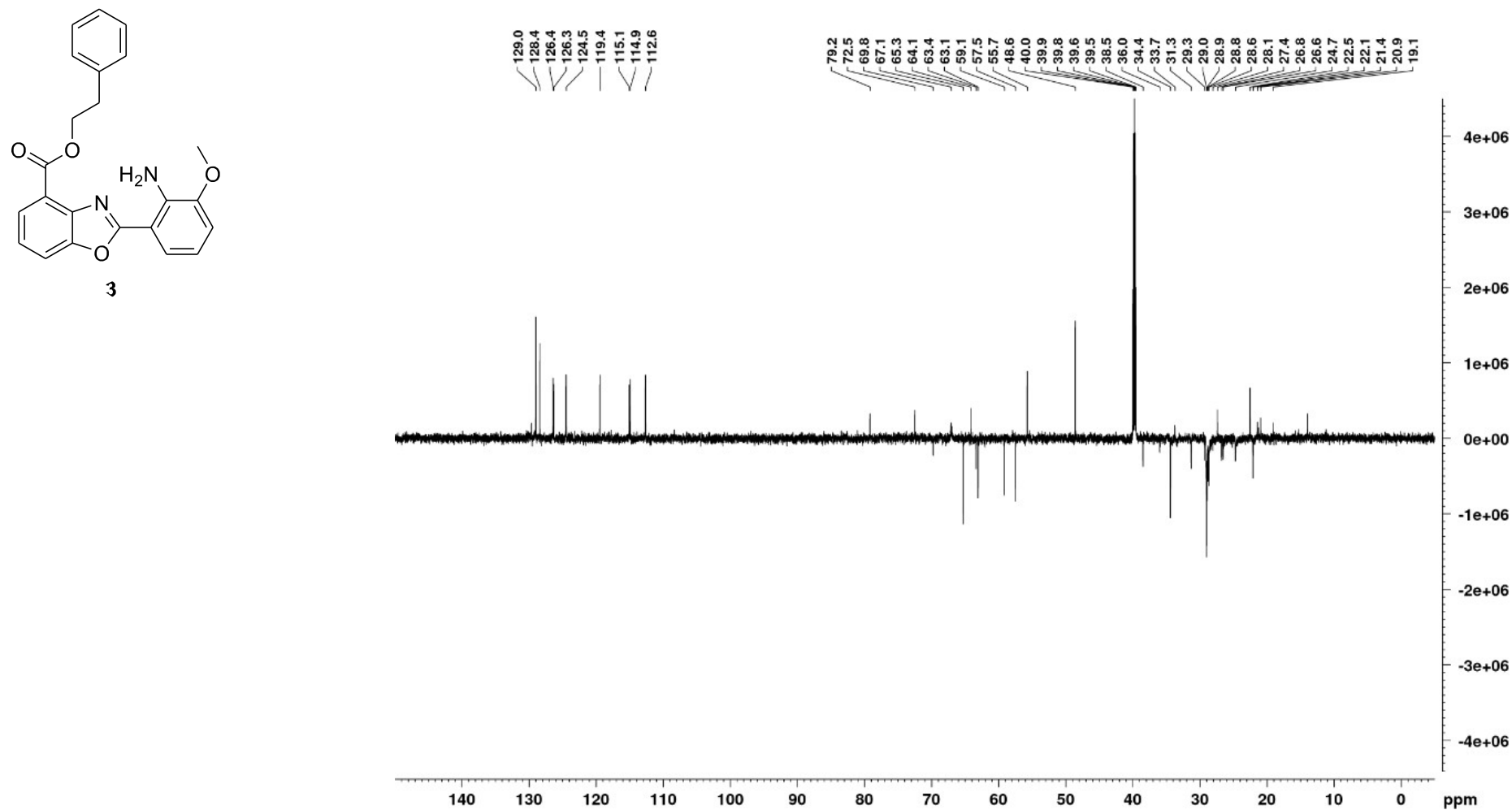


Figure S4. The spectroscopic data of **3**
(E) The HSQC spectrum of compound **3**

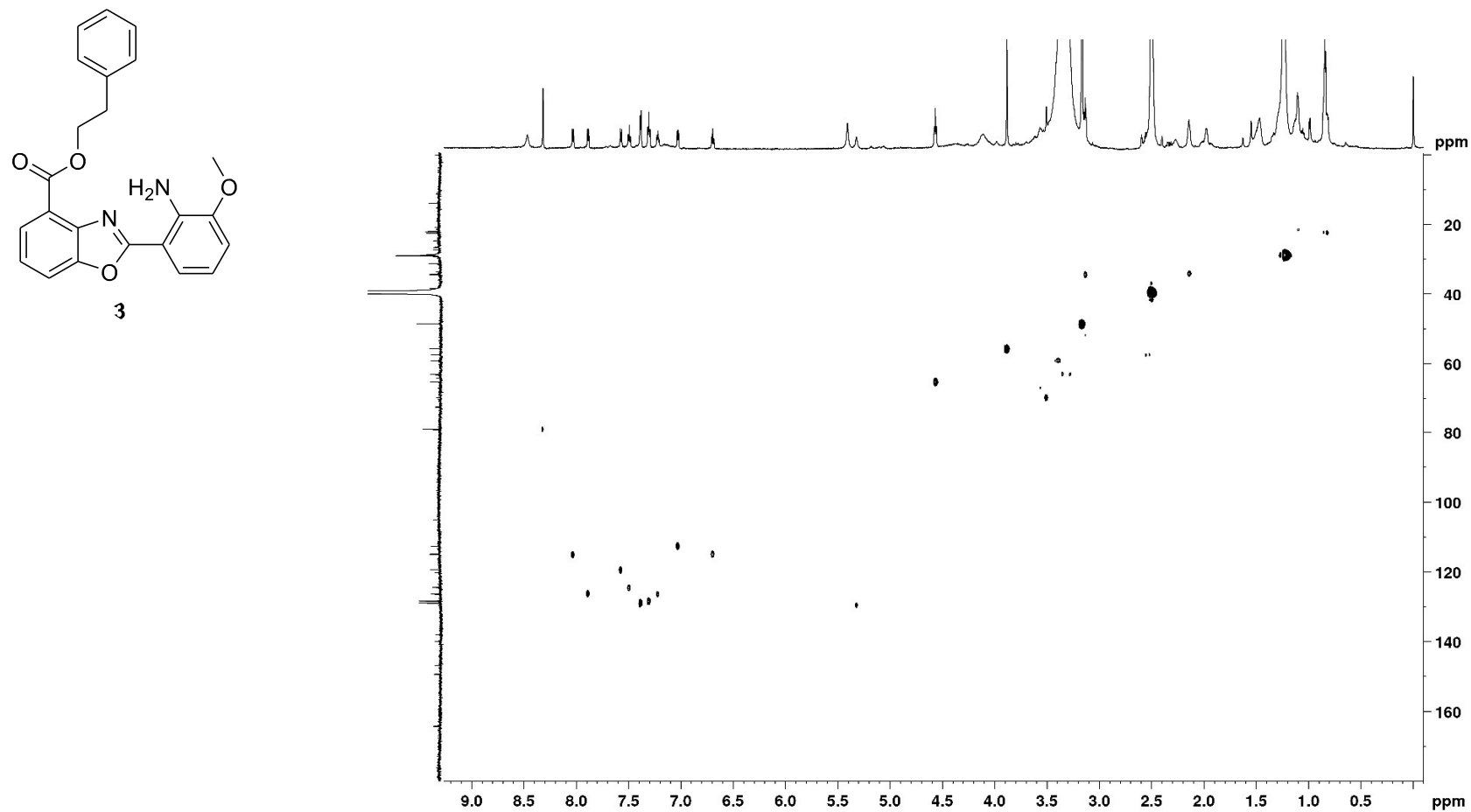


Figure S4. The spectroscopic data of **3**
(F) The HMBC spectrum of **3**

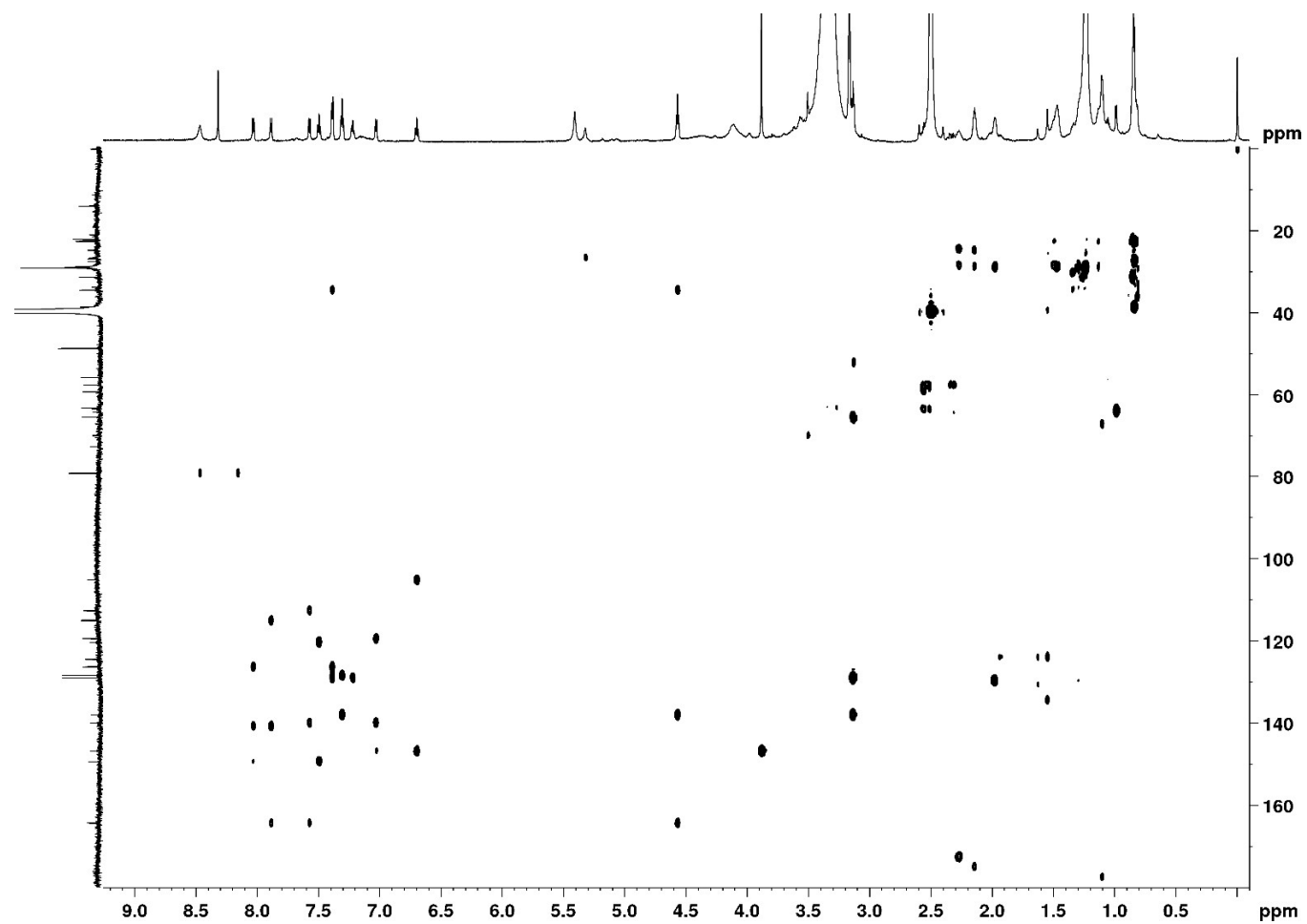
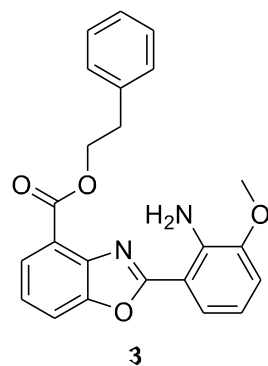


Figure S4. The spectroscopic data of **3**
(G) The ^1H - ^1H COSY spectrum of **3**

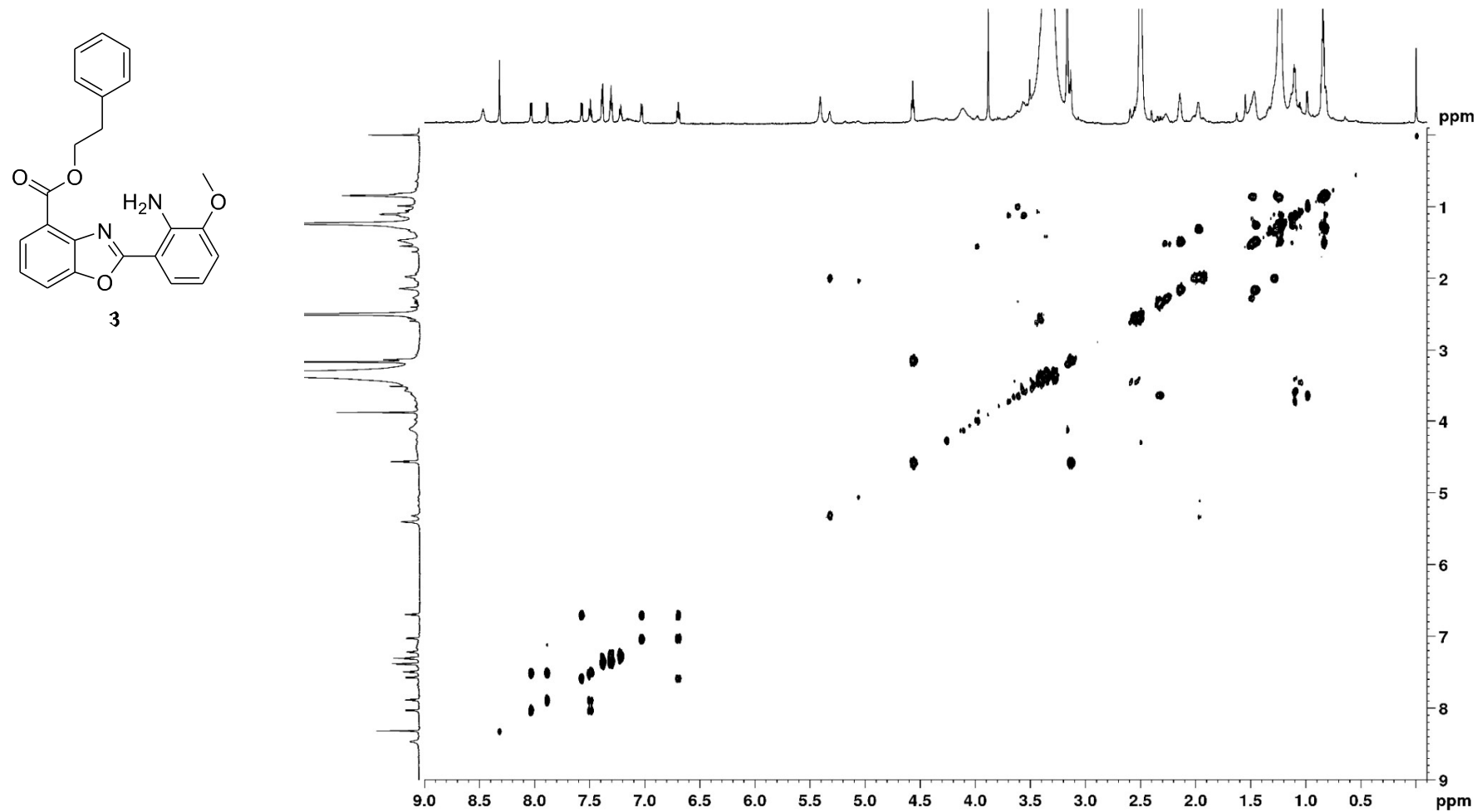
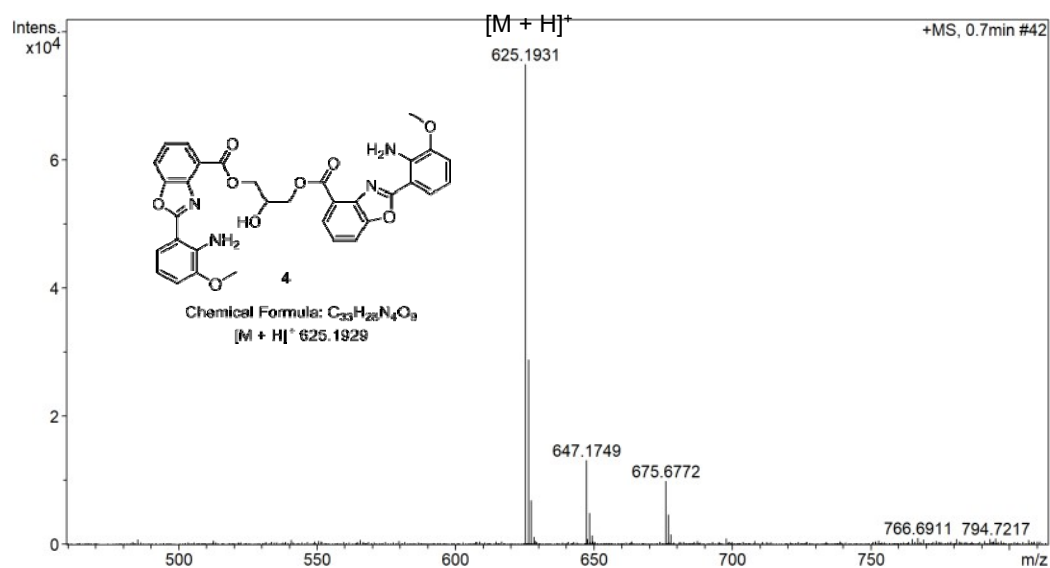


Figure S5. The spectroscopic data of **4**

(A) The HRESIMS (a) and UV (b) spectrum of **4**

(a)



(b)

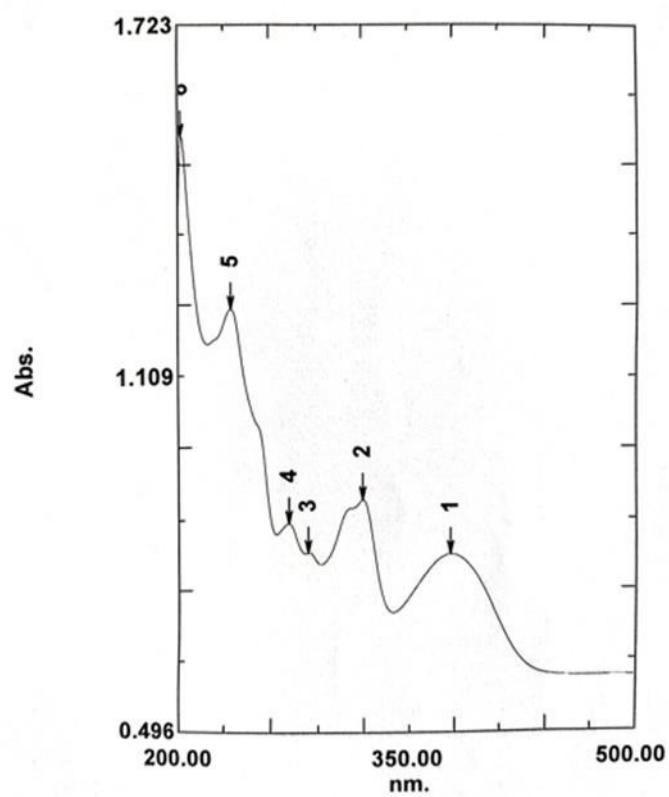


Figure S5. The spectroscopic data of **4**

(B) The ^1H -NMR spectrum of **4** (700 MHz for ^1H NMR in CDCl_3)

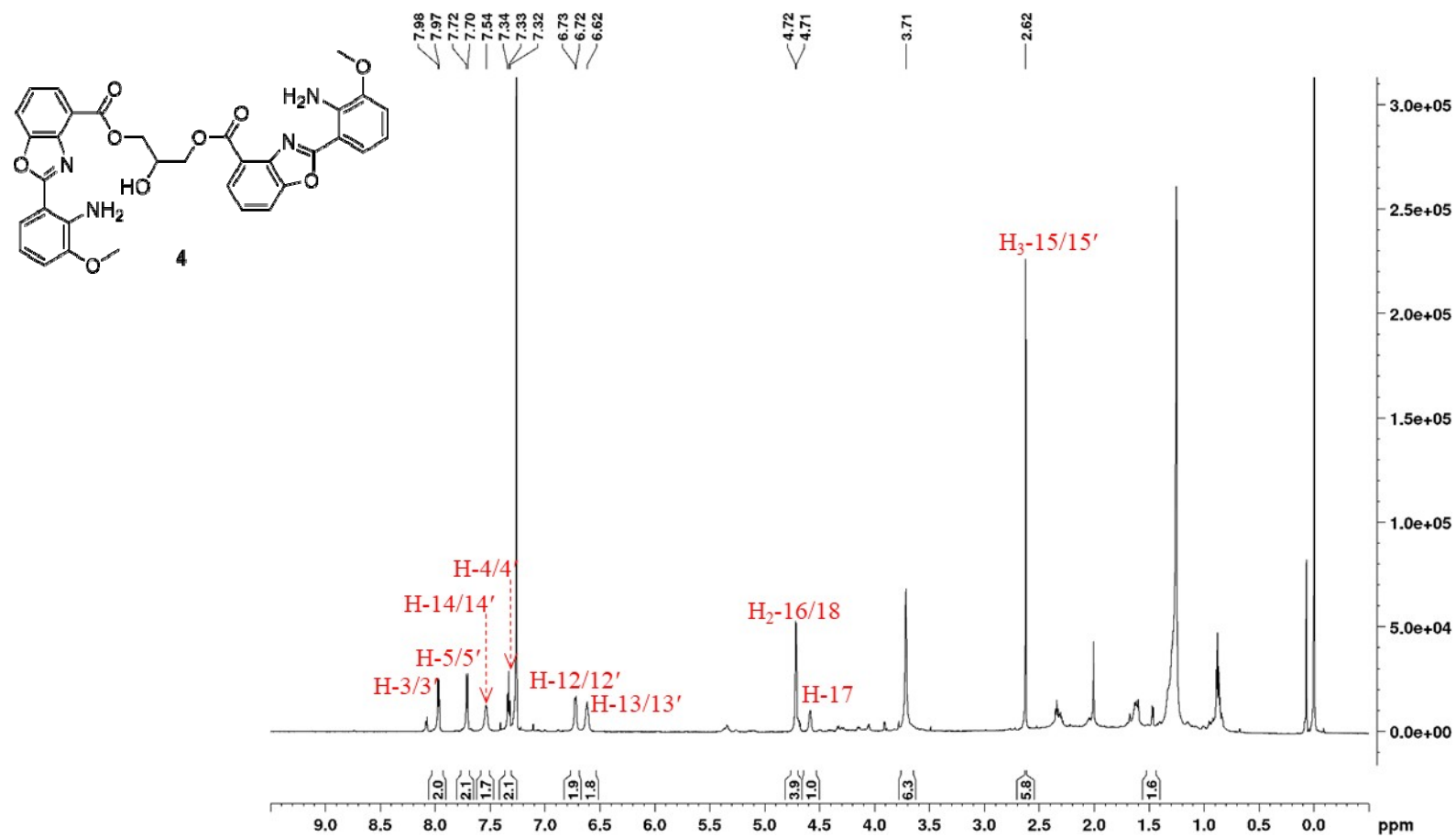


Figure S5. The spectroscopic data of **4**
 (C) The ^{13}C NMR spectrum of compound **4** in CDCl_3 (175 MHz)

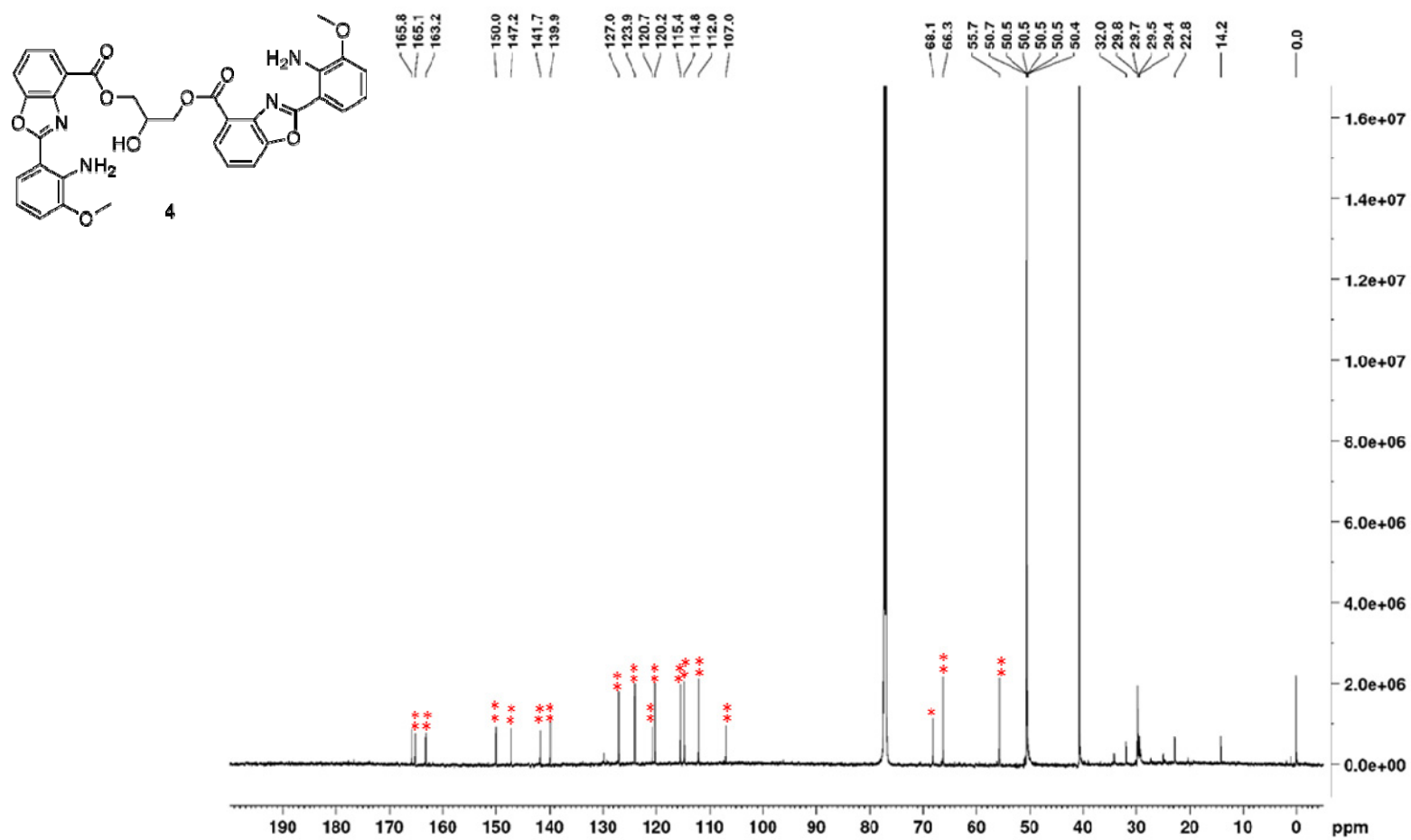
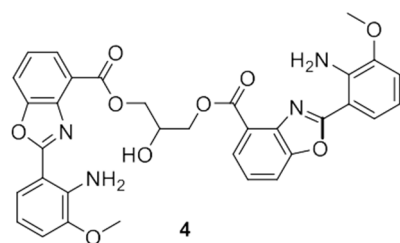


Figure S5. The spectroscopic data of **4**
(D) The DEPT135 spectrum of **4**



4

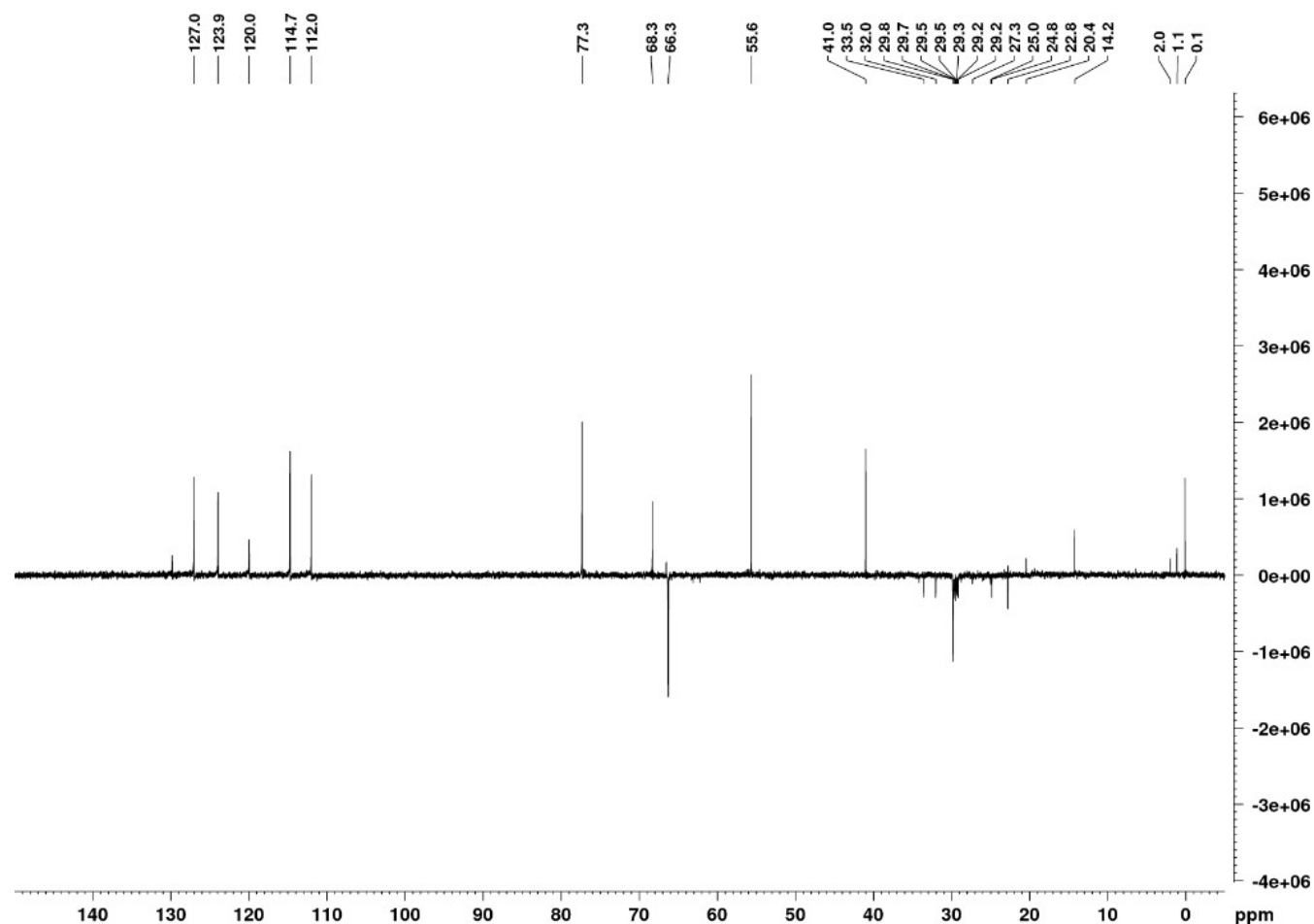


Figure S5. The spectroscopic data of **4**
(E) The HSQC spectrum of compound **4**

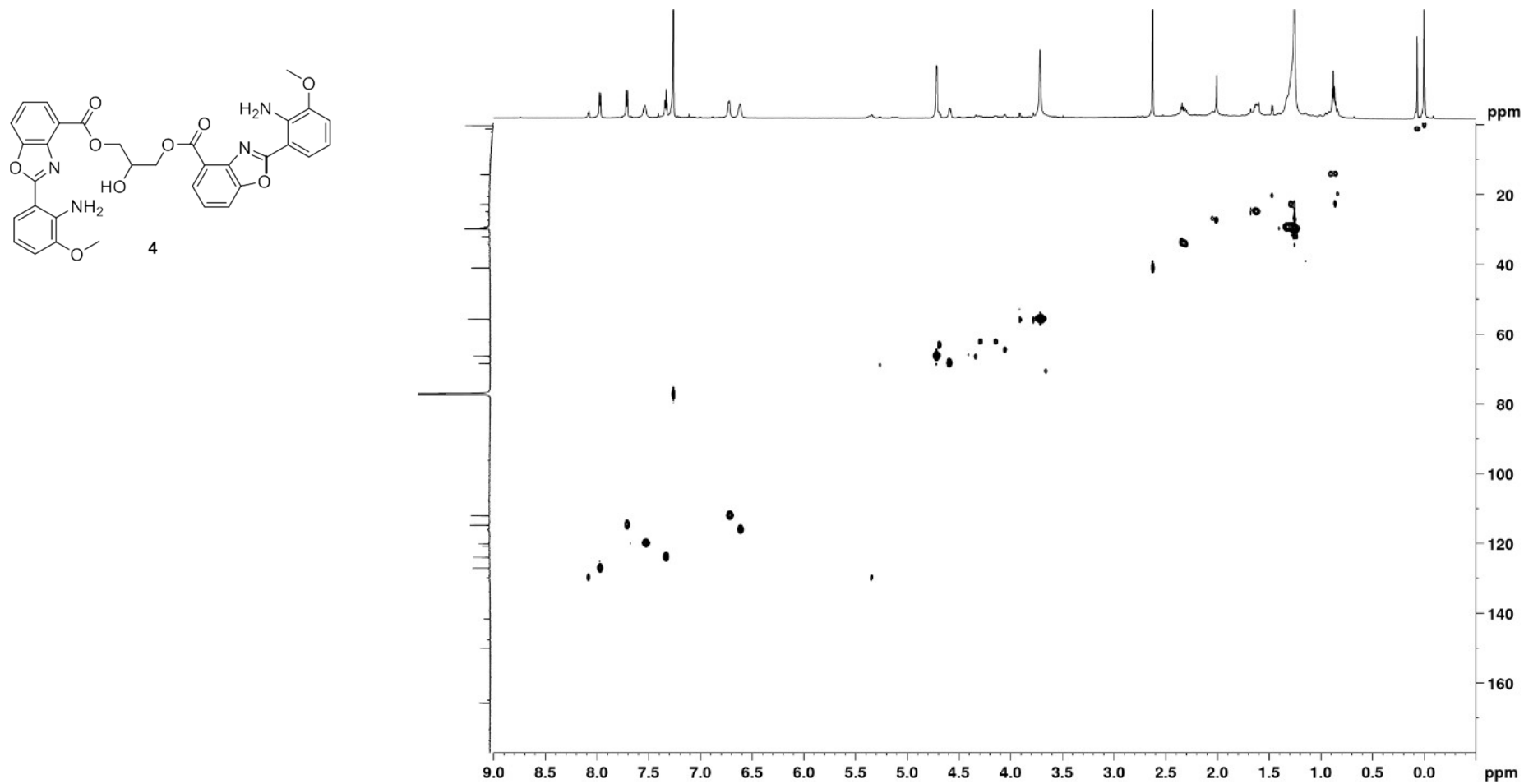


Figure S5. The spectroscopic data of **4**
(F) The HMBC spectrum of **4**

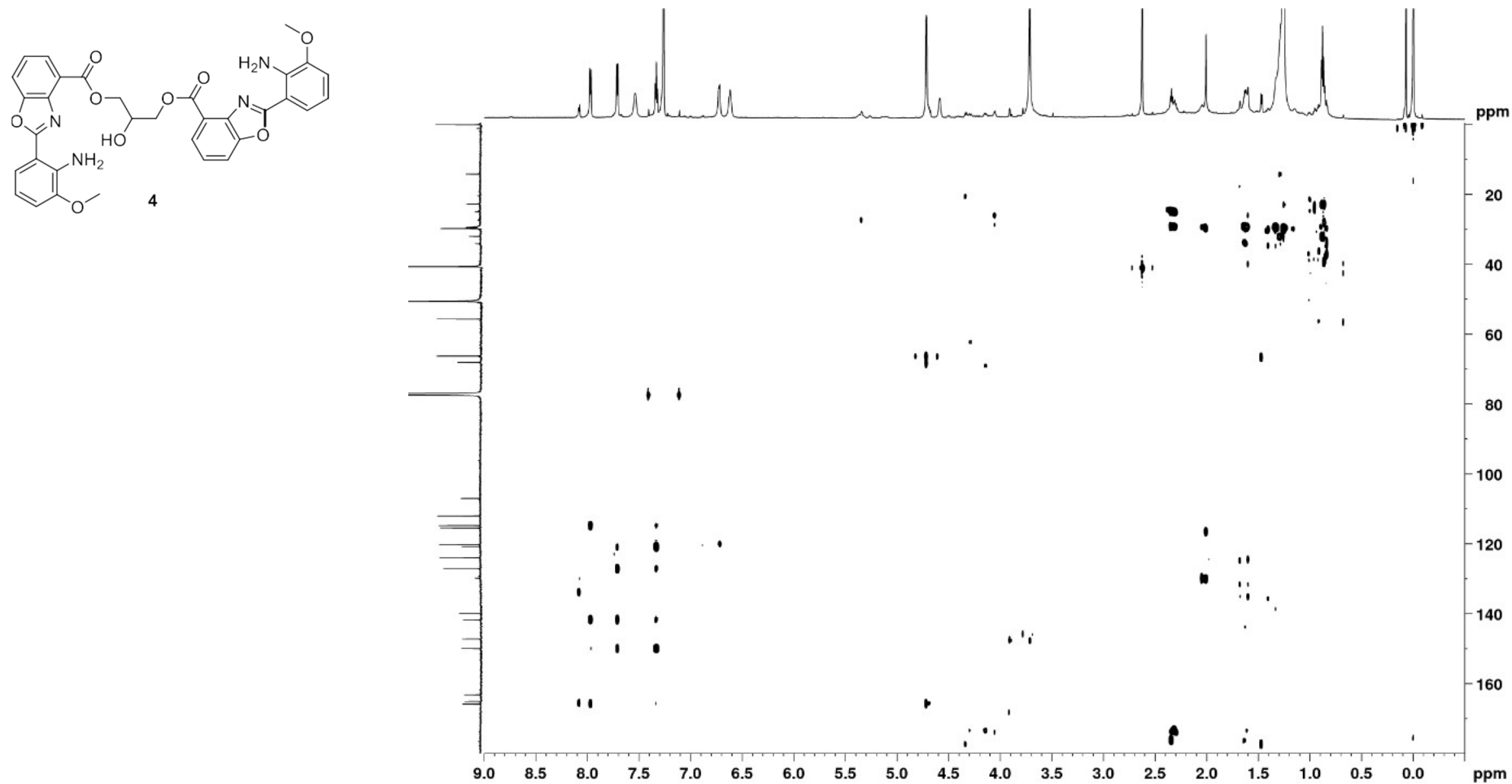


Figure S5. The spectroscopic data of **4**
(G) The ^1H - ^1H COSY spectrum of **4**

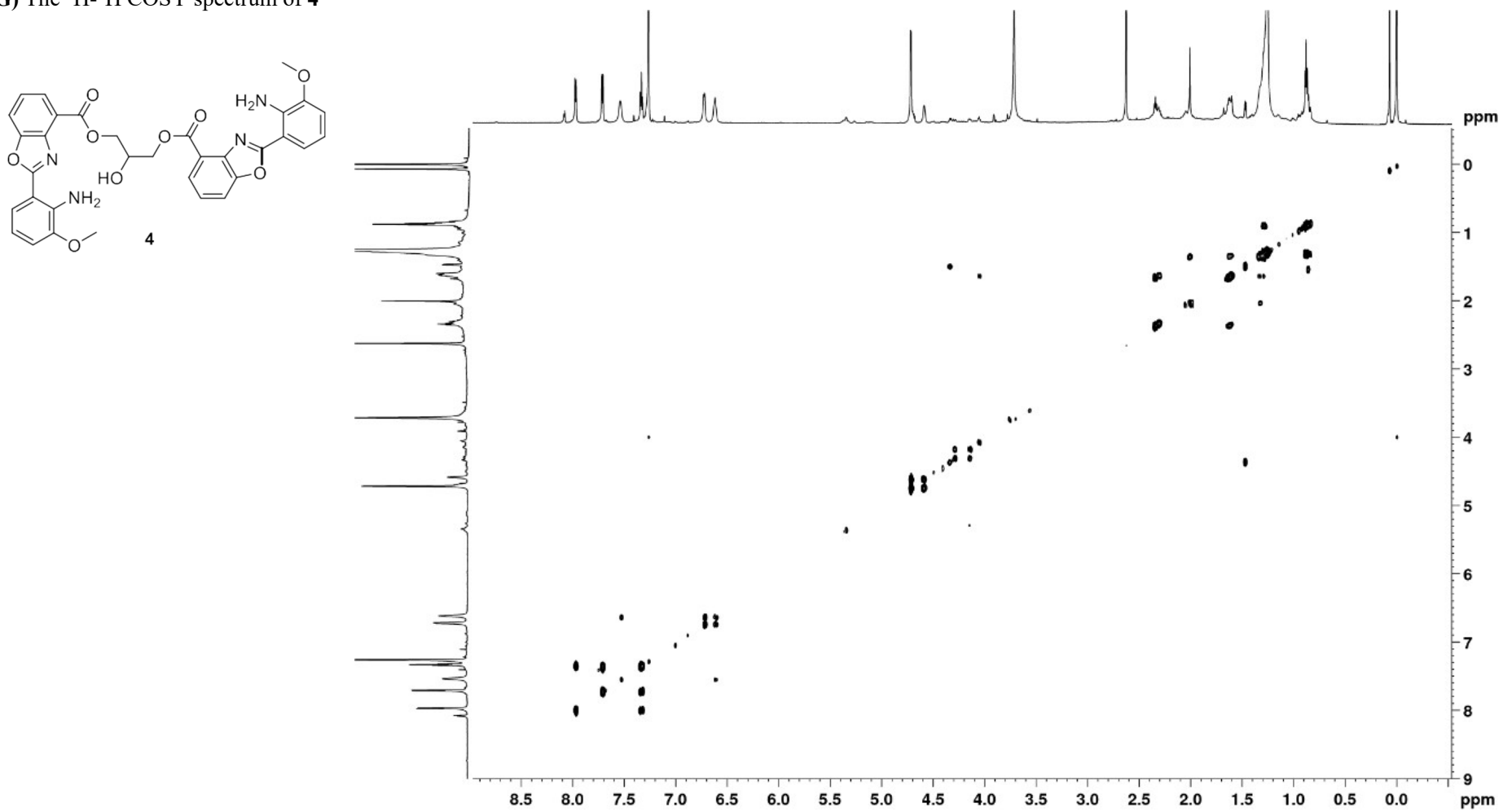
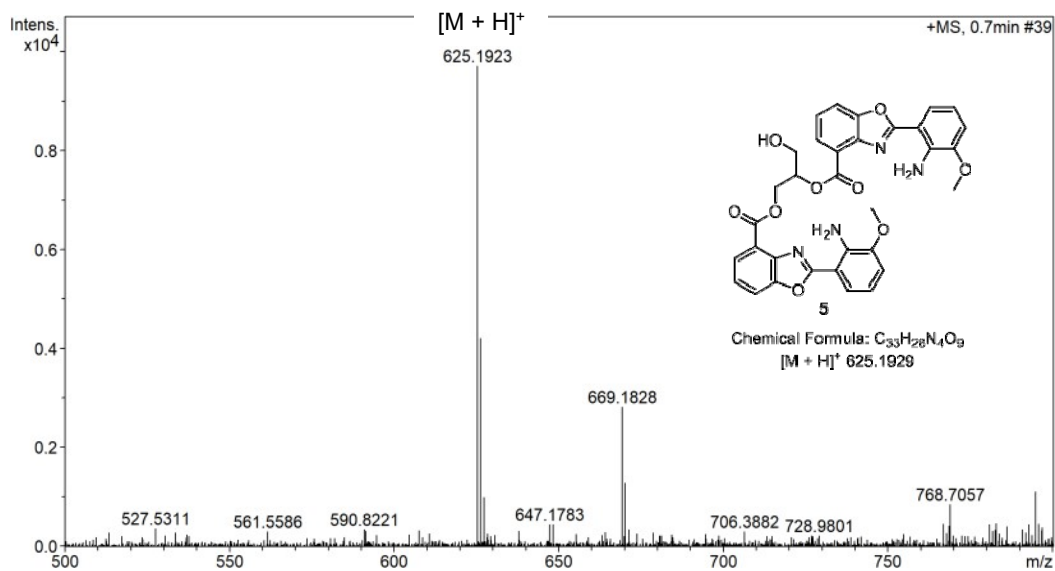


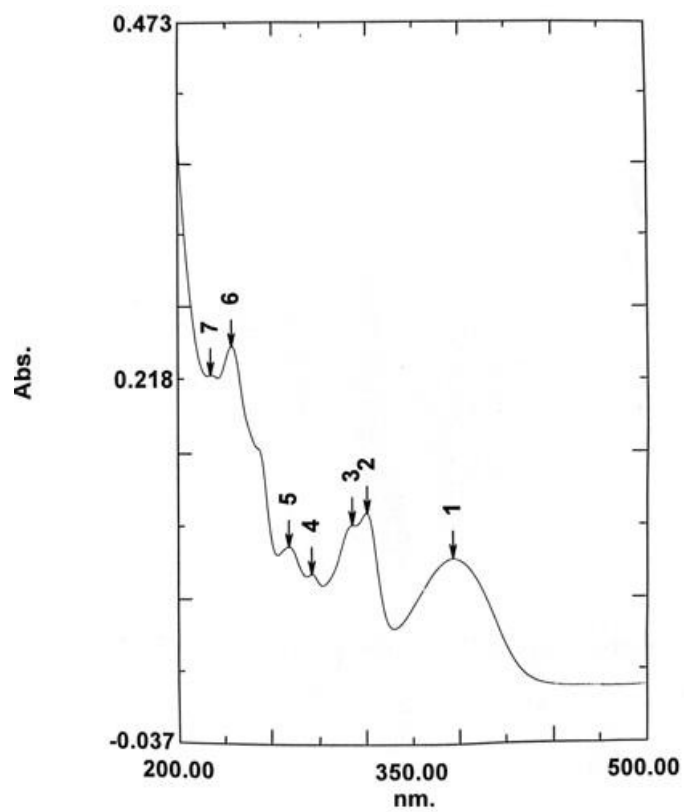
Figure S6. The spectroscopic data of **5**

(A) The HRESIMS (a), UV (b), CD (c) spectrum and chiral HPLC analysis (d) of **5**

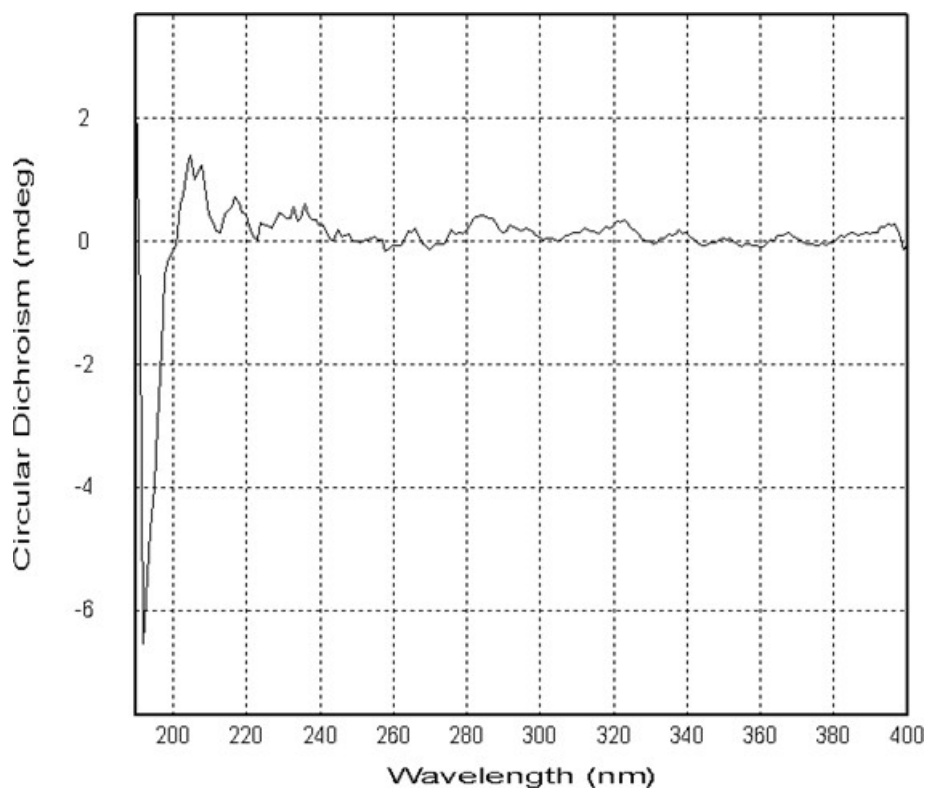
(a)



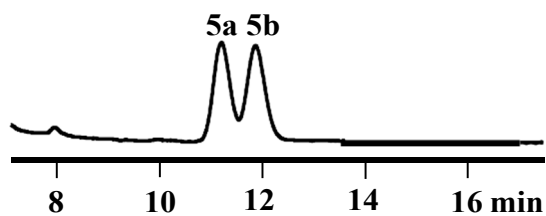
(b)



(c)



(d) Chiral HPLC analysis of **5**



Chiral HPLC analysis of compound **5** was conducted by using a chiral column (Lux Cellulose-4, 5 μ m, 250 \times 4.6 mm, phenomenex) using the following program: solvent system (solvent A, 10% CH₃CN in water; solvent B, 90% CH₃CN in water); 5% B to 100% B (0–20 min), 100% B (21–25 min), 100% B to 5%B (25–26 min), 5% B (26–30 min), flow rate at 1 mL min⁻¹.

Figure S6. The spectroscopic data of **5**

(B) The ^1H -NMR spectrum of **5** (700 MHz for ^1H NMR in CDCl_3)

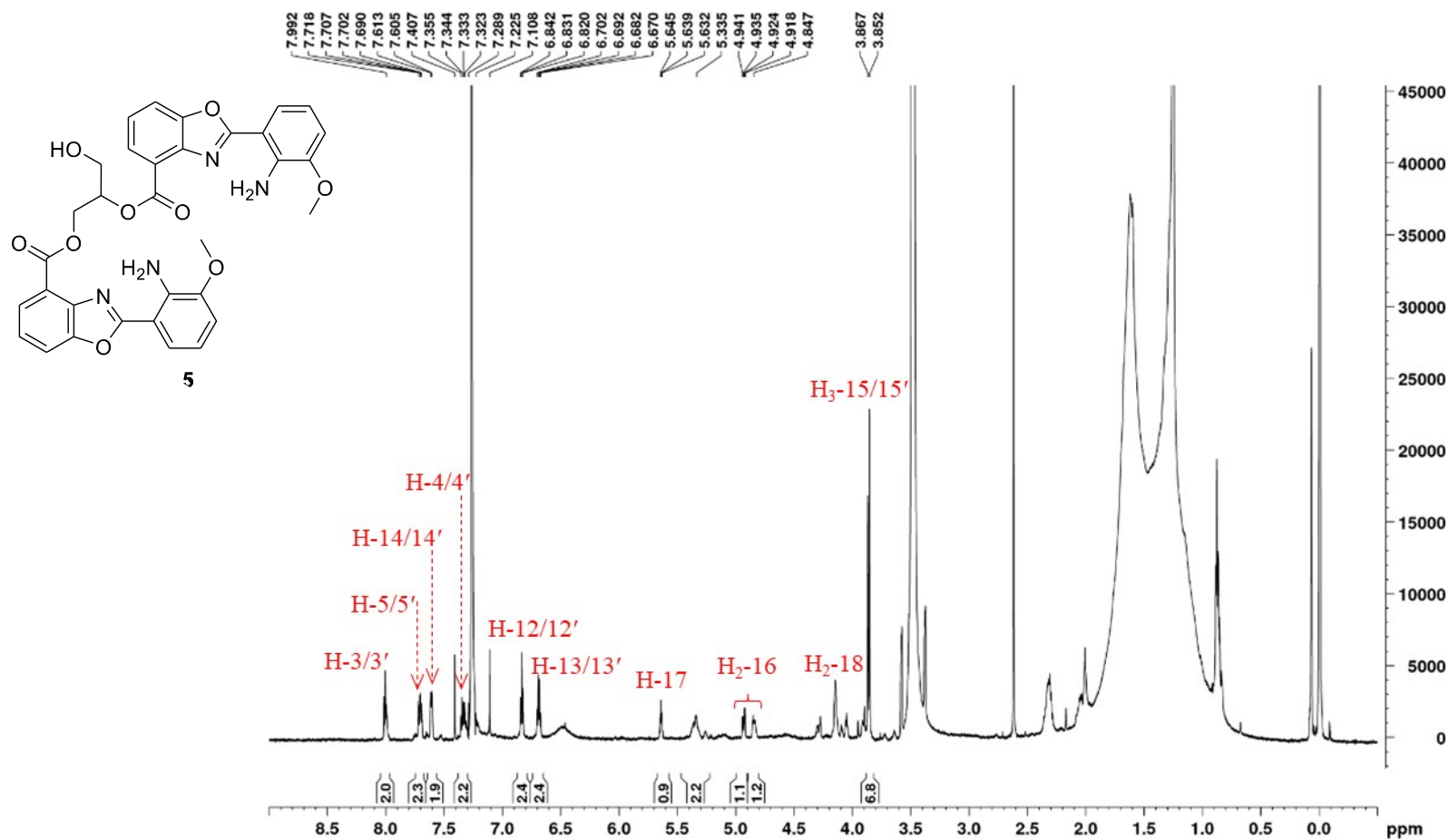


Figure S6. The spectroscopic data of **5**
 (C) The ^{13}C NMR spectrum of compound **5** in CDCl_3 (175 MHz)

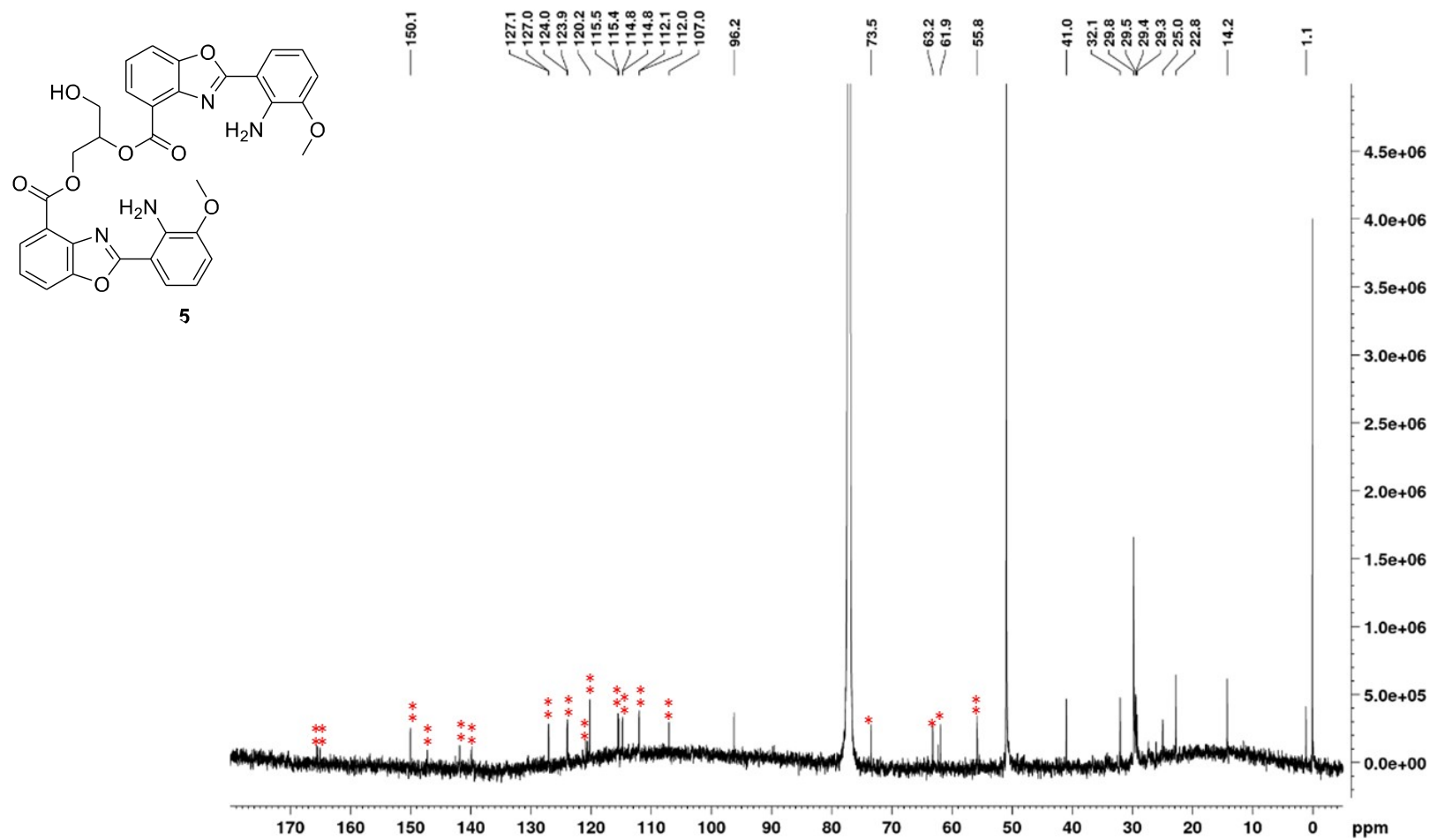


Figure S6. The spectroscopic data of **5**
(D) The DEPT135 spectrum of **5**

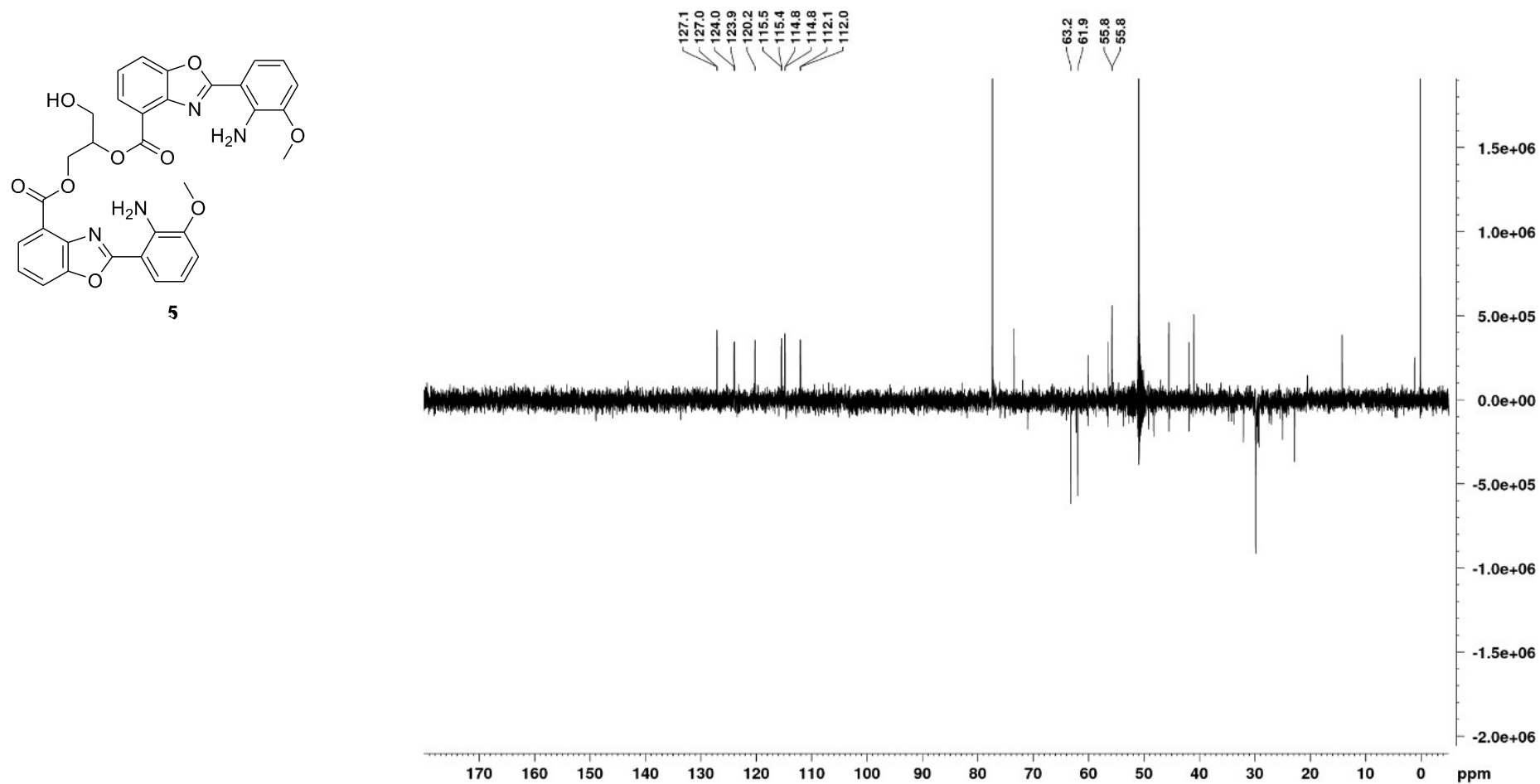


Figure S6. The spectroscopic data of **5**
(E) The HSQC spectrum of compound **5**

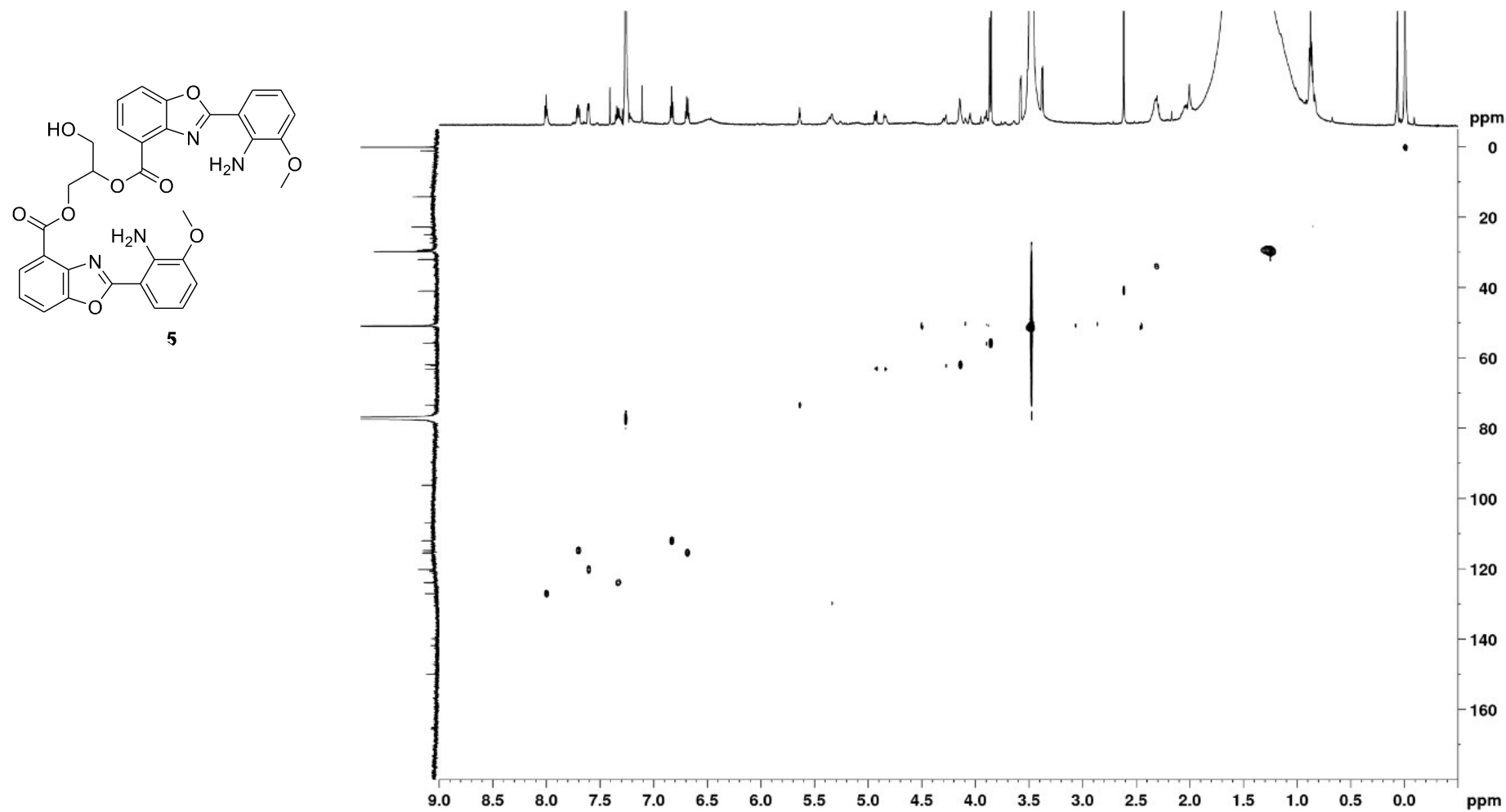


Figure S6. The spectroscopic data of **5**
(F) The HMBC spectrum of **5**

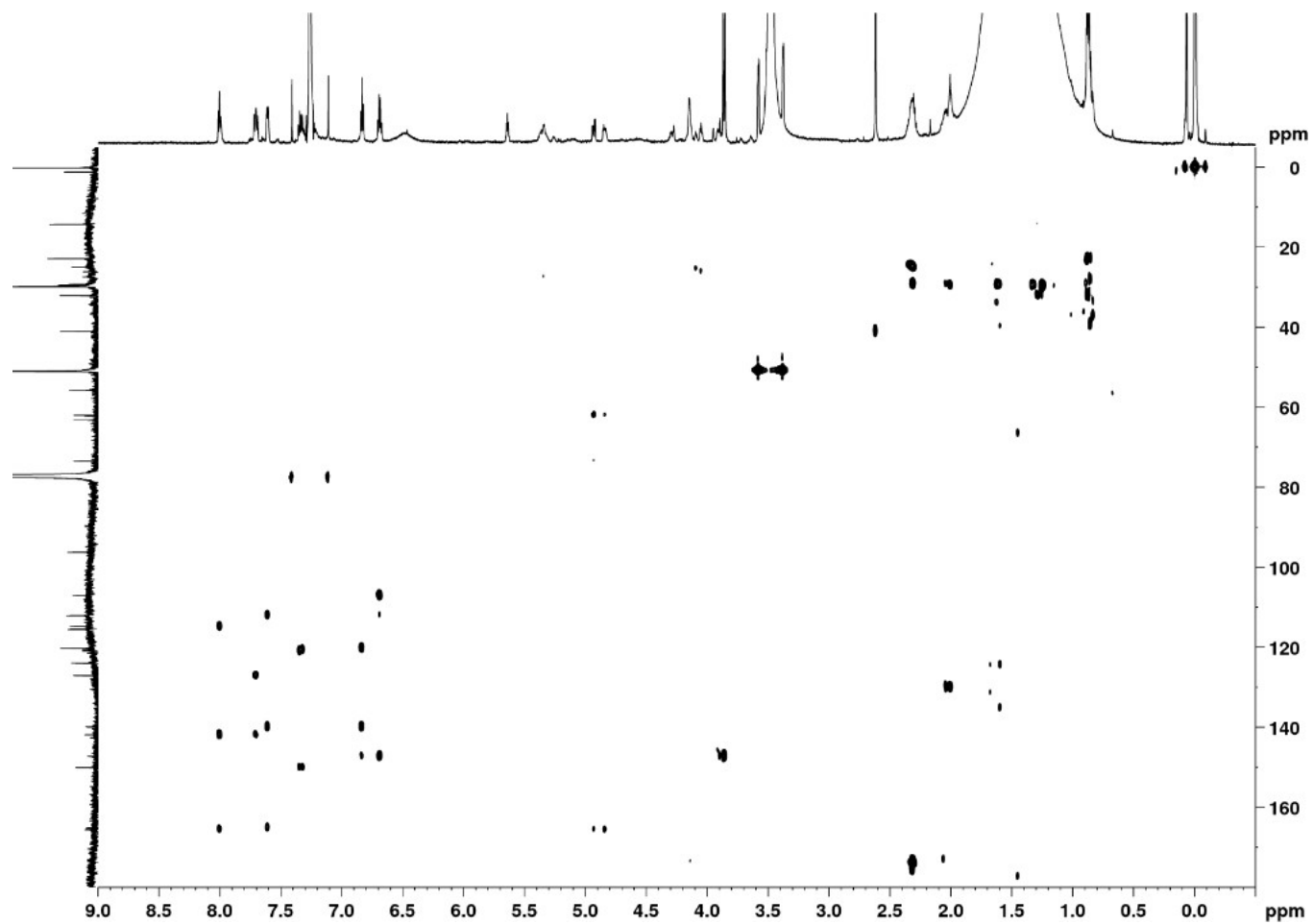
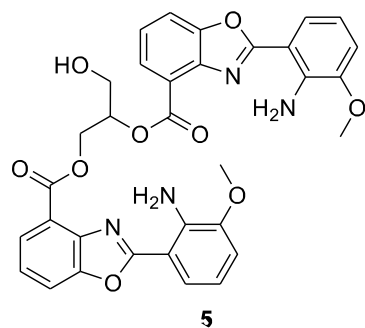


Figure S6. The spectroscopic data of **5**
(G) The ^1H - ^1H COSY spectrum of **5**

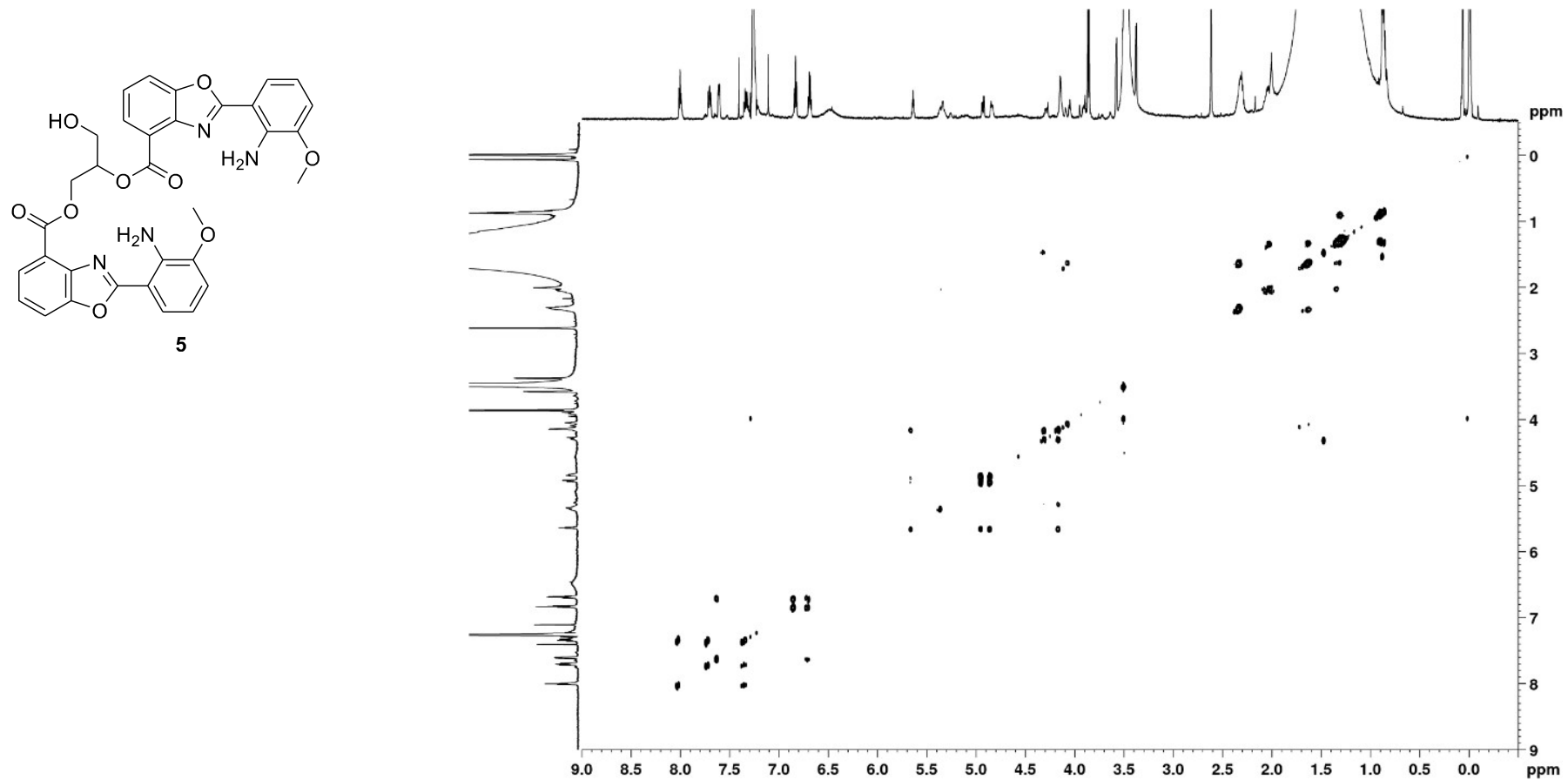
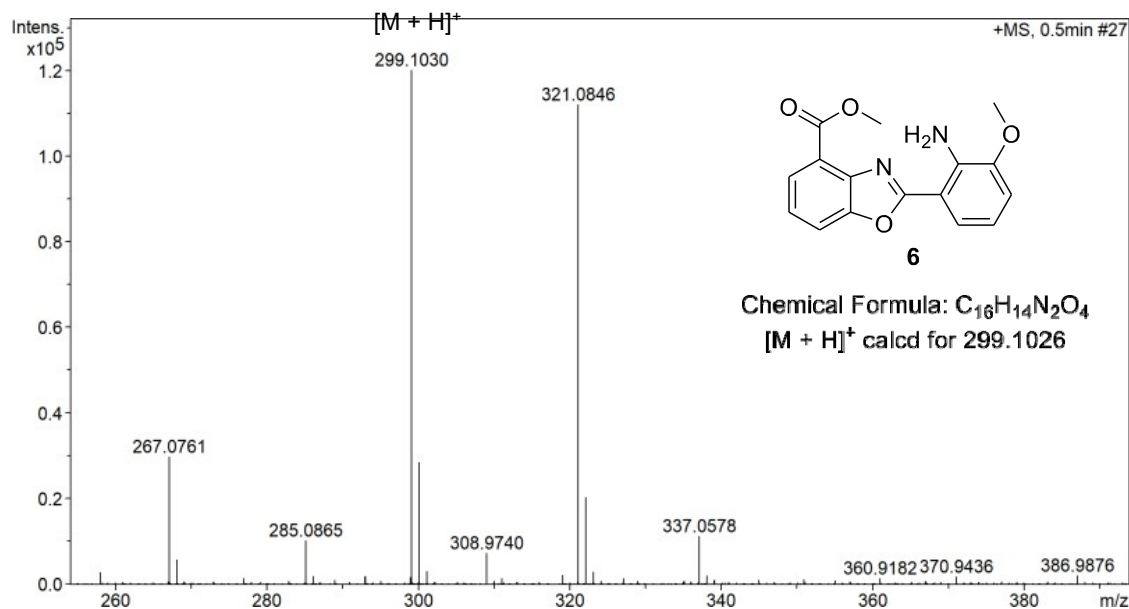


Figure S7. The spectroscopic data of **6**

(A) The HRESIMS (a) UV (b) spectrum of **6**

(a)



(b)

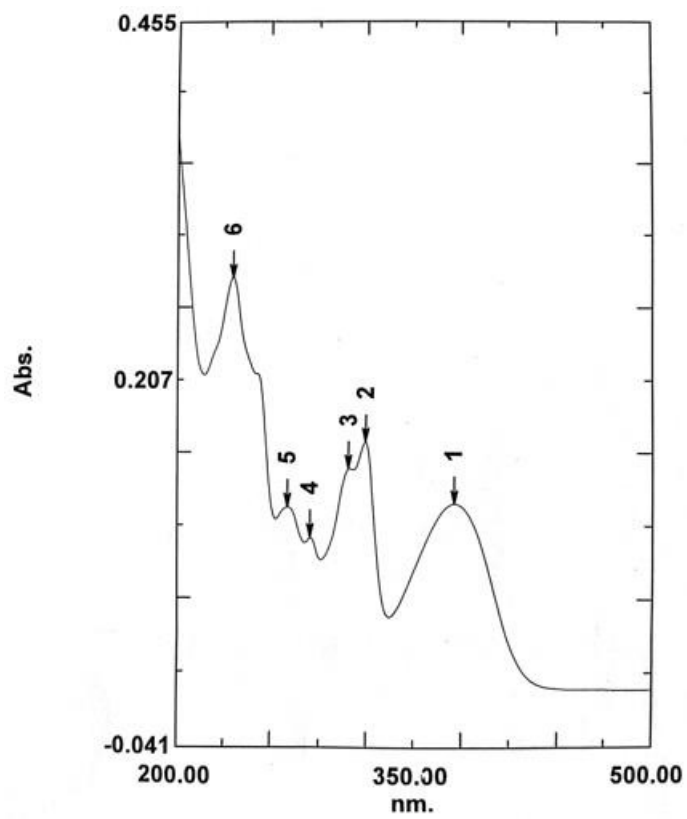


Figure S7. The spectroscopic data of **6**

(B) The ^1H -NMR spectrum of **6** (700 MHz for ^1H NMR in CDCl_3)

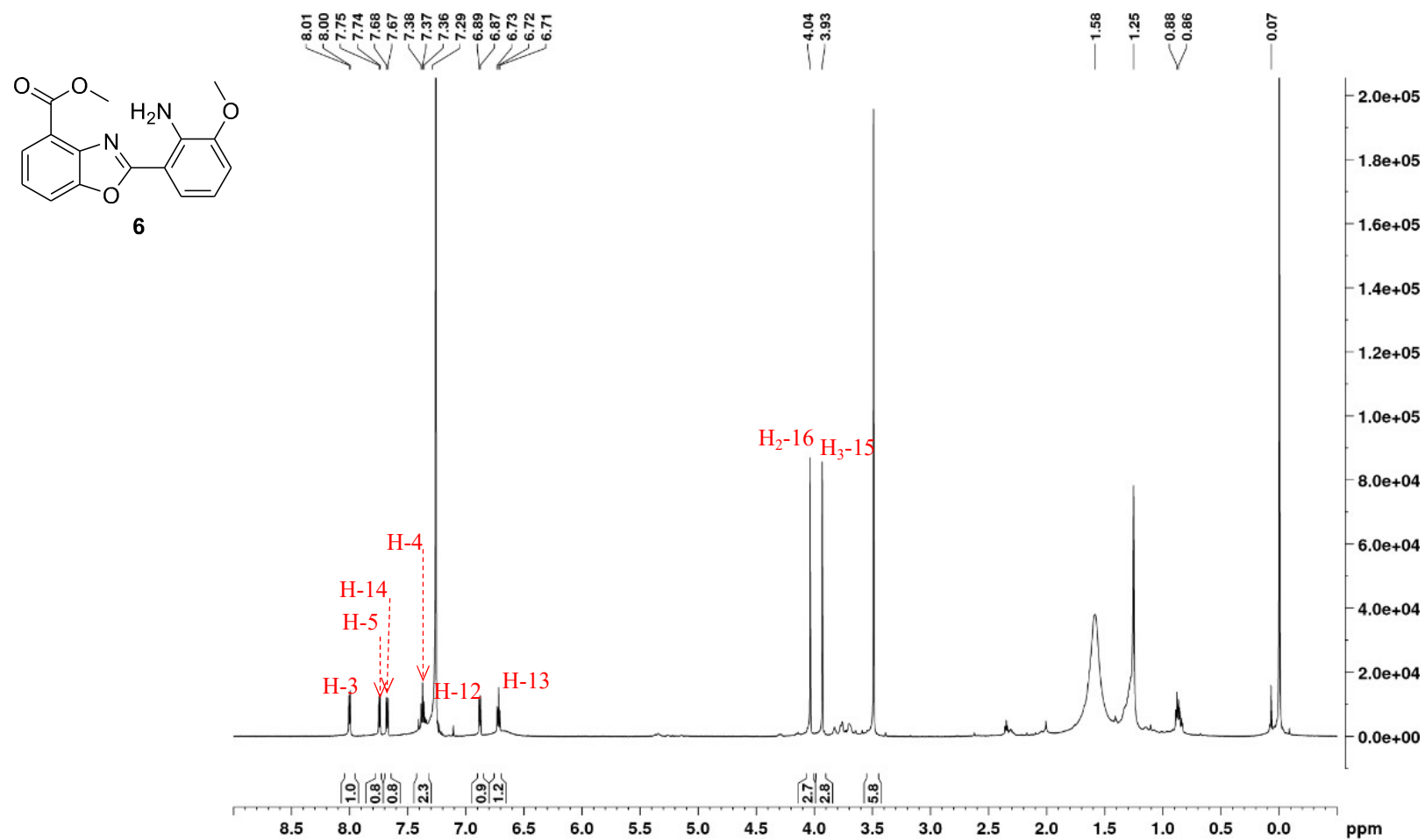


Figure S7. The spectroscopic data of **6**
(C) The ^{13}C -NMR spectrum of compound **6** in CDCl_3 (175 MHz)

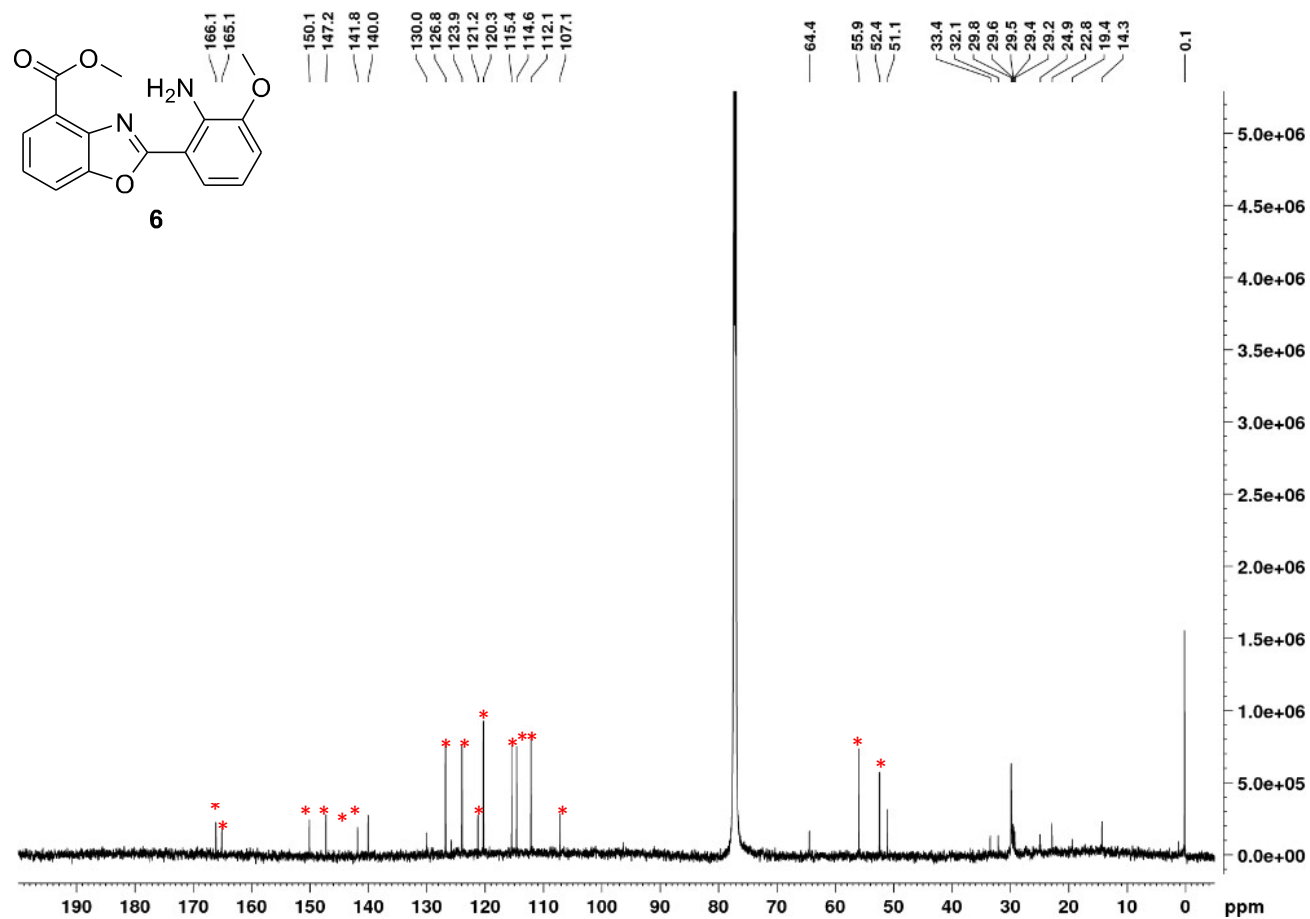


Figure S7. The spectroscopic data of **6**
(D) The DEPT135 spectrum of **6**

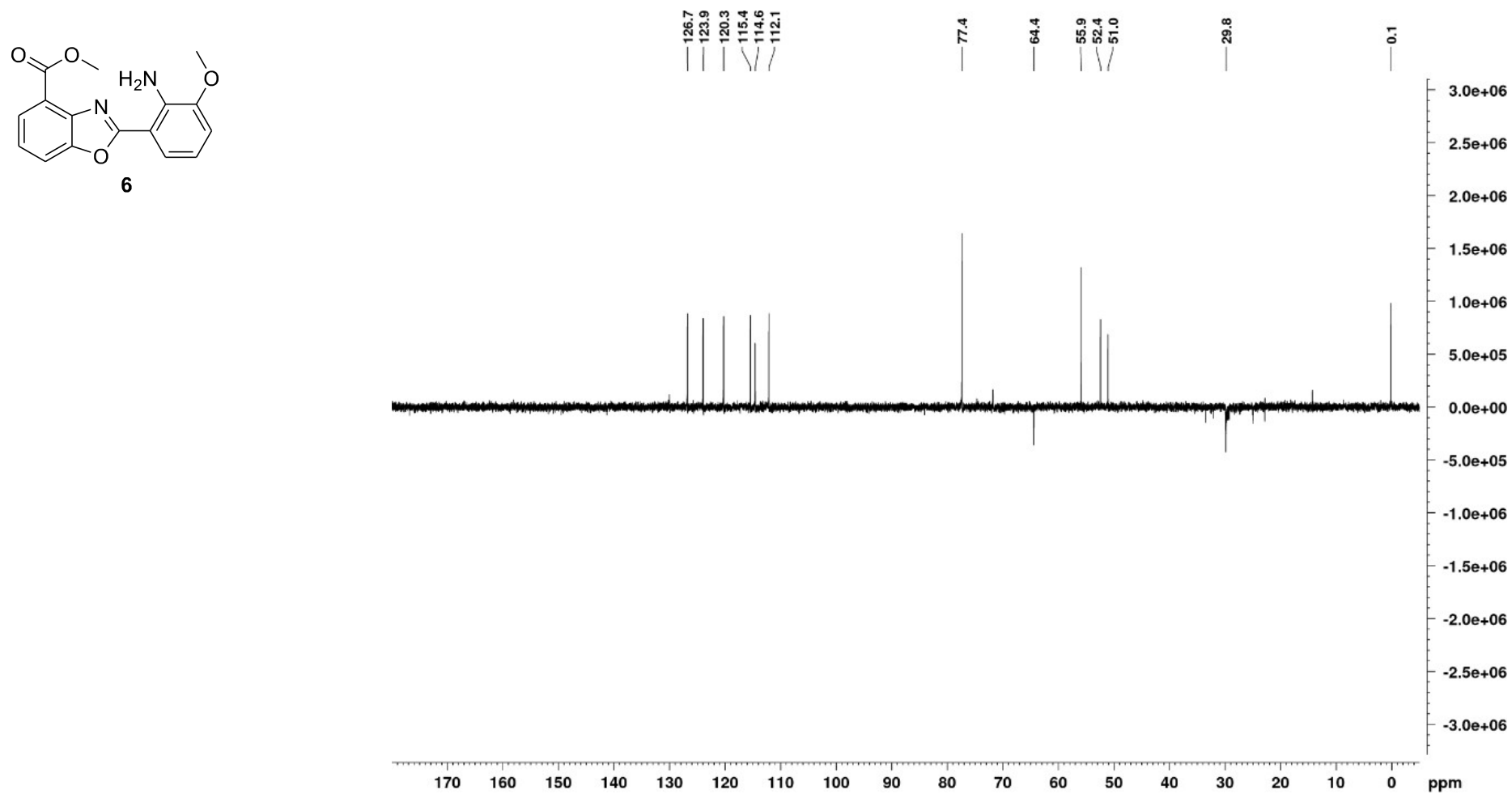


Figure S7. The spectroscopic data of **6**
(E) The HSQC spectrum of compound **6**

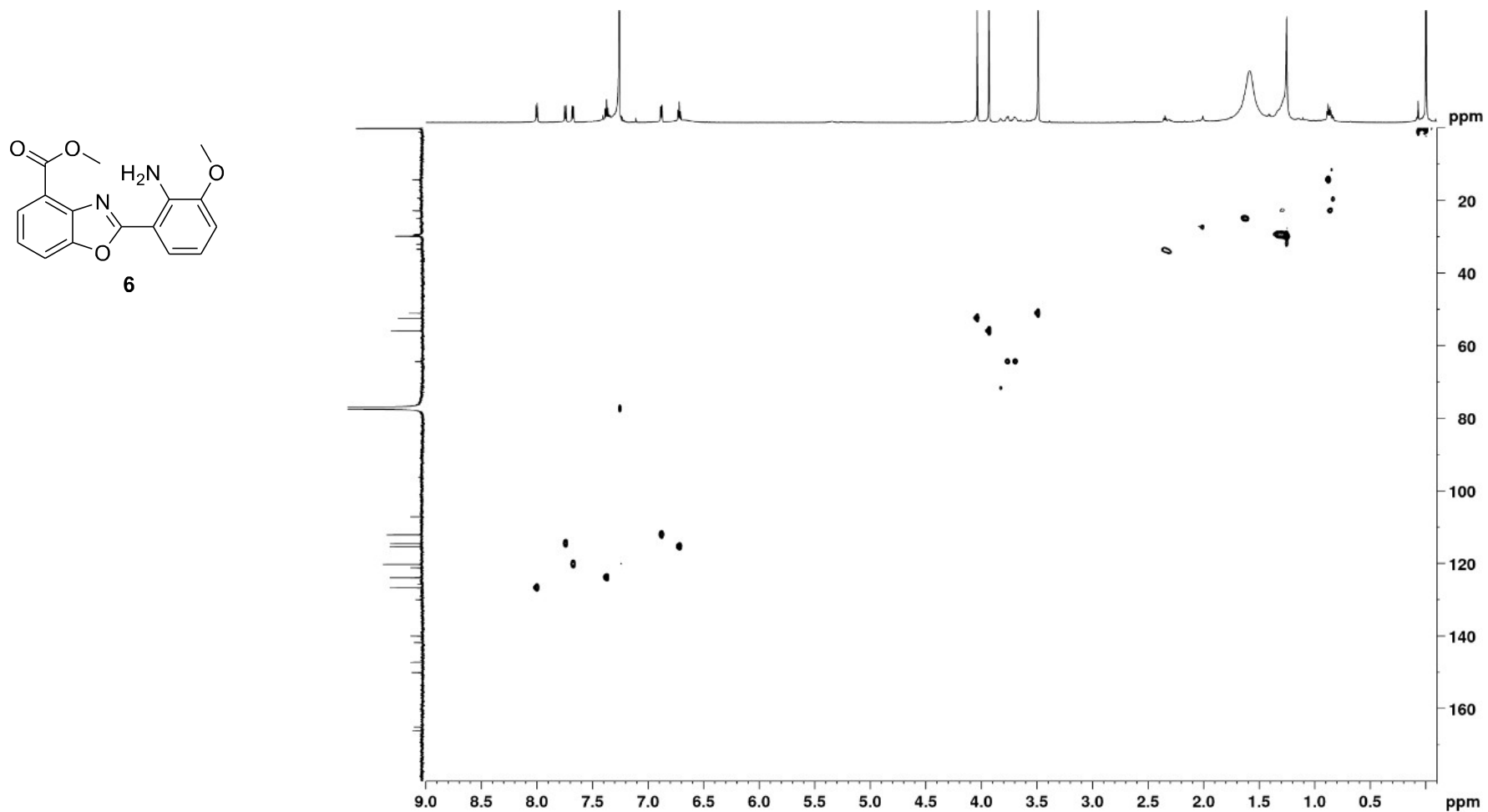


Figure S7. The spectroscopic data of **6**

(F) The HMBC spectrum of **6**

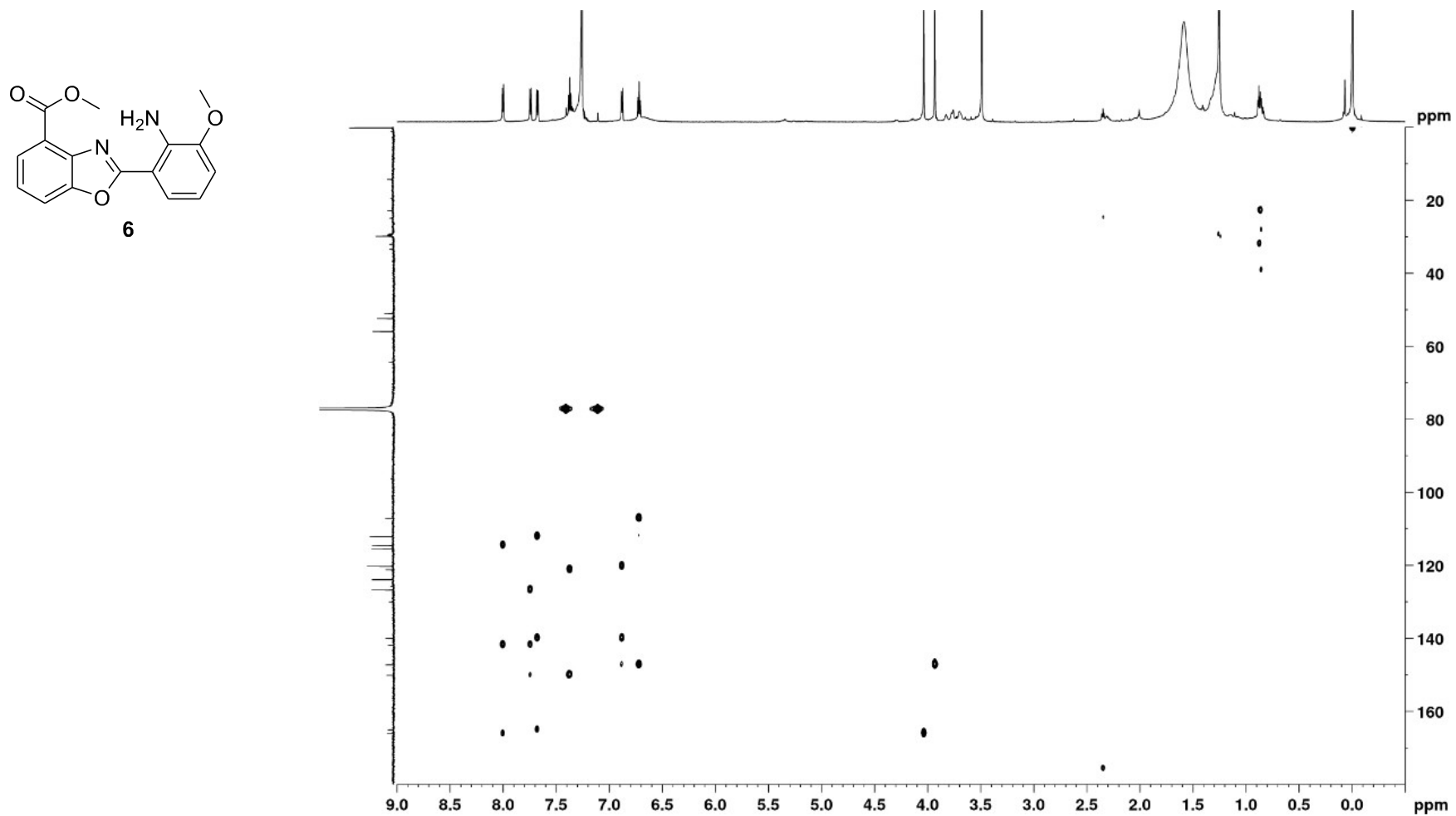
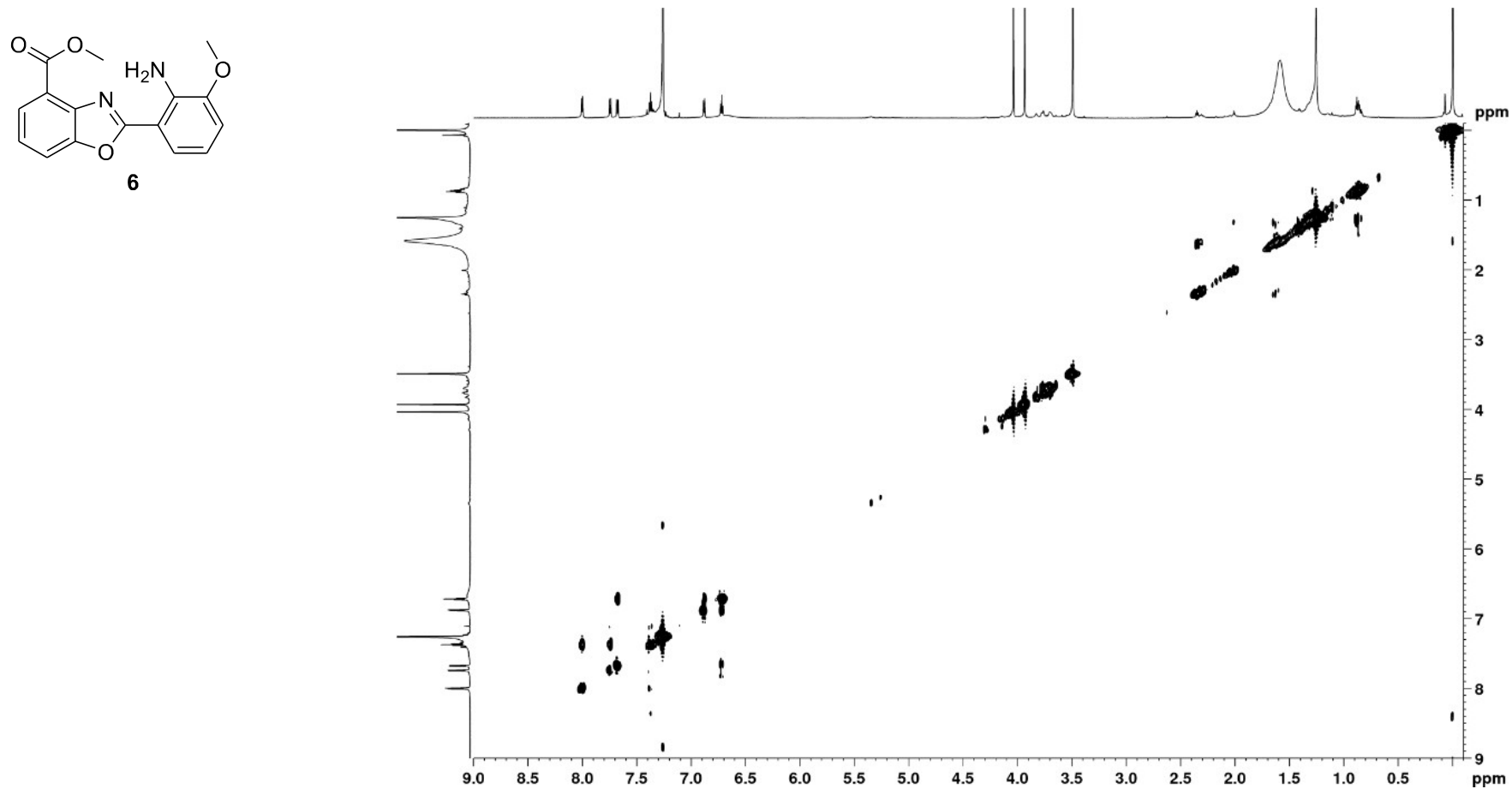


Figure S7. The spectroscopic data of **6**
(G) The ^1H - ^1H COSY spectrum of **6**



References

1. Flett, F.; Mersinias, V.; Smith, C. P., High efficiency intergeneric conjugal transfer of plasmid DNA from *Escherichia coli* to methyl DNA-restricting *streptomycetes*. *FEMS Microbiol Lett* **1997**, 155 (2), 223-229.
2. Zaburannyi, N.; Rabyk, M.; Ostash, B.; Fedorenko, V.; Luzhetskyy, A., Insights into naturally minimised *Streptomyces albus* J1074 genome. *Bmc Genomics* **2014**, 15:97.
3. Myronovskyi, M.; Rosenkranzer, B.; Nadmid, S.; Pujic, P.; Normand, P.; Luzhetskyy, A., Generation of a cluster-free *Streptomyces albus* chassis strains for improved heterologous expression of secondary metabolite clusters. *Metab Eng* **2018**, 49, 316-324.
4. Huang, S.; Li, N.; Zhou, J.; He, J., Construction of a new bacterial artificial chromosome (BAC) vector for cloning of large DNA fragments and heterologous expression in *Streptomyces*. *Wei Sheng Wu Xue Bao* **2012**, 52 (1), 30-37.