In vitro and *in silico* studies of SARS-CoV-2 main protease M^{pro} inhibitors isolated from *Helichrysum bracteatum*

Gehad Abdel Wahab,[a] Walaa S. Aboelmaaty,[a] Mohamed Farid Lahloub,[a] and Amal Sallam [*a]

[a]*Pharmacognosy Department, Faculty of Pharmacy, Mansoura University, Mansoura, 35516 Egypt.*

*Corresponding author:

Amal Sallam, Department of Pharmacognosy, Faculty of Pharmacy, Mansoura University, Mansoura, 35516 Egypt. Tel. +201092017949; fax. +2 050 2247496.

E-mail: asallam@mans.edu.eg

Other authors e-mail addresses:

Gehad Abdel Wahab: e-mail: gehadabdelwahab@mans.edu.egor algehadalhaq@yahoo.com.

Walaa S. Aboelmaaty: e-mail: walaa_safwat@mans.edu.egor walaa_m_s@yahoo.com

Mohamed Farid Lahloub: e-mail: mfilah@yahoo.com.

ORCID

Gehad Abdel Wahab: 0000-0001-6424-2445

Walaa S. Aboelmaaty: 0000-0002-7806-7884

Mohamed Farid Lahloub: 0000-0001-6769-3145

Amal Sallam: 0000-0003-3577-2046

In vitro and _in silico_ studies of SARS-CoV-2 main protease M^{pro} inhibitors isolated from _Helichrysum bracteatum_

Abstract

Discovering SARS-CoV-2 inhibitors from natural sources is still a target that capture the interest of many researchers. In this study, the methanolic extract of *Helichrysum bracteatum* leaves besides compounds **(1-18)** isolated and identified from it were evaluated *in vitro* for their inhibitory activities against SARS-CoV-2 main protease (M^{pro}) using Fluorescence Resonance Energy Transfer assay (FRET-based assay). Based on 1D and 2D spectroscopic techniques, compounds **(1-18)** were identified as 24-*β*-ethyl-cholesta-5(6),22(23),25(26)-triene-3-ol **(1)**, *α*-amyrin **(2)**, linoleic acid **(3)**, 24-*β*-ethyl-cholesta-5(6),22(23),25(26)-triene-3-O-*β*-D-glucoside **(4)**, 1,3-propanediol-2-amino-1-(3`,4`-methylenedioxyphenyl)**(5)**, (-)-(7*R*,8*R*,8`*R*)-acuminatolide **(6)**, (+)-piperitol **(7)**, 5,7,4`-trihydroxy-8,3`-dimethoxy flavanone **(8)**, 5,7,4`-trihydroxy-6-methoxy flavanone **(9)**, 4`,5-dihydroxy-3`,7,8-trimethoxyflavone **(10)**, 5,7-dihydroxy-3`,4`,5`,8-tetramethoxy flavone **(11)**, 1,3-propanediol-2-amino-1-(4`-hydroxy-3`-methoxyphenyl)**(12)**, 3`,5`,5,7-tetrahydroxy-6-methoxyflavanone **(13)**, simplexoside (piperitol-O-*β*-D-glucoside) **(14)**, pinoresinol monomethyl ether-*β*-D-glucoside **(15)**, orientin **(16)**, luteolin-3`-O-*β*-D-glucoside **(17)** and 3,5-dicaffeoylquinic acid **(18)**. Compounds **6**, **12** and **14** showed comparable inhibitory activities against SARS-COV-2 M^{pro} with IC_{50} values of 0.917±0.05, 0.476±0.02 and 0.610±0.03µM, respectively compared with the control lopinavir with an IC_{50} value of 0.225±0.01 µM. The other tested compounds showed considerable inhibitory activities. Molecular docking study for the tested compounds was carried out to correlate their binding modes and affinities for SARS-COV-2 M^{pro}enzyme with the *in vitro* results. Analyzing the results of the *in vitro* assay together with the obtained *in silico* results led to the conclusion that the phenylpropanoids, lignans and flavonoids could be considered suitable drug leads for developing anti-COVID-19 therapeutics. Moreover, the phenylpropanoid skeleton oxygenated at C3, C4 of the phenyl moiety and at C1, C3 of the propane part constitute an essential core of the SARS-COV-2 M^{pro} inhibitors, thus could be proposed as scaffold for the design of new anti-COVID-19 drugs.

Key words: *Helichrysum bracteatum*, FRET-based assay, SARS-COV-2 M^{pro} inhibitors, Molecular docking, anti-COVID-19.

Content

Tables **Page:**

Compound 1

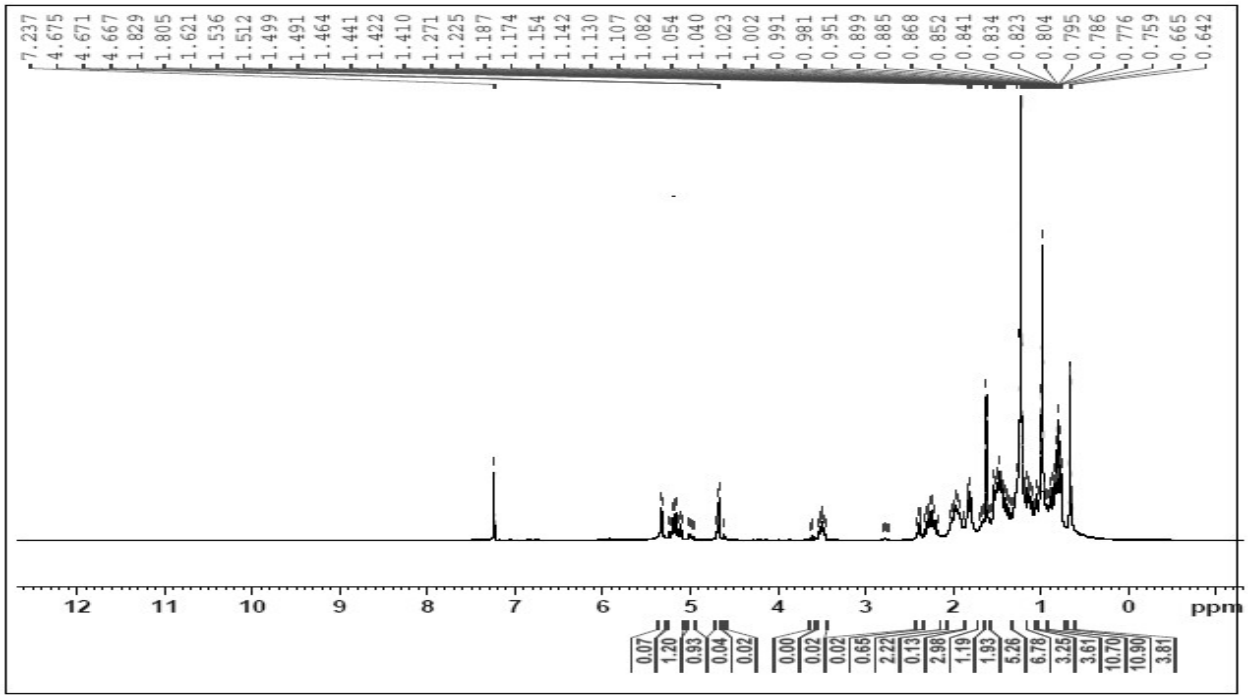


Figure S1: ^{1}H NMR spectrum of compound 1

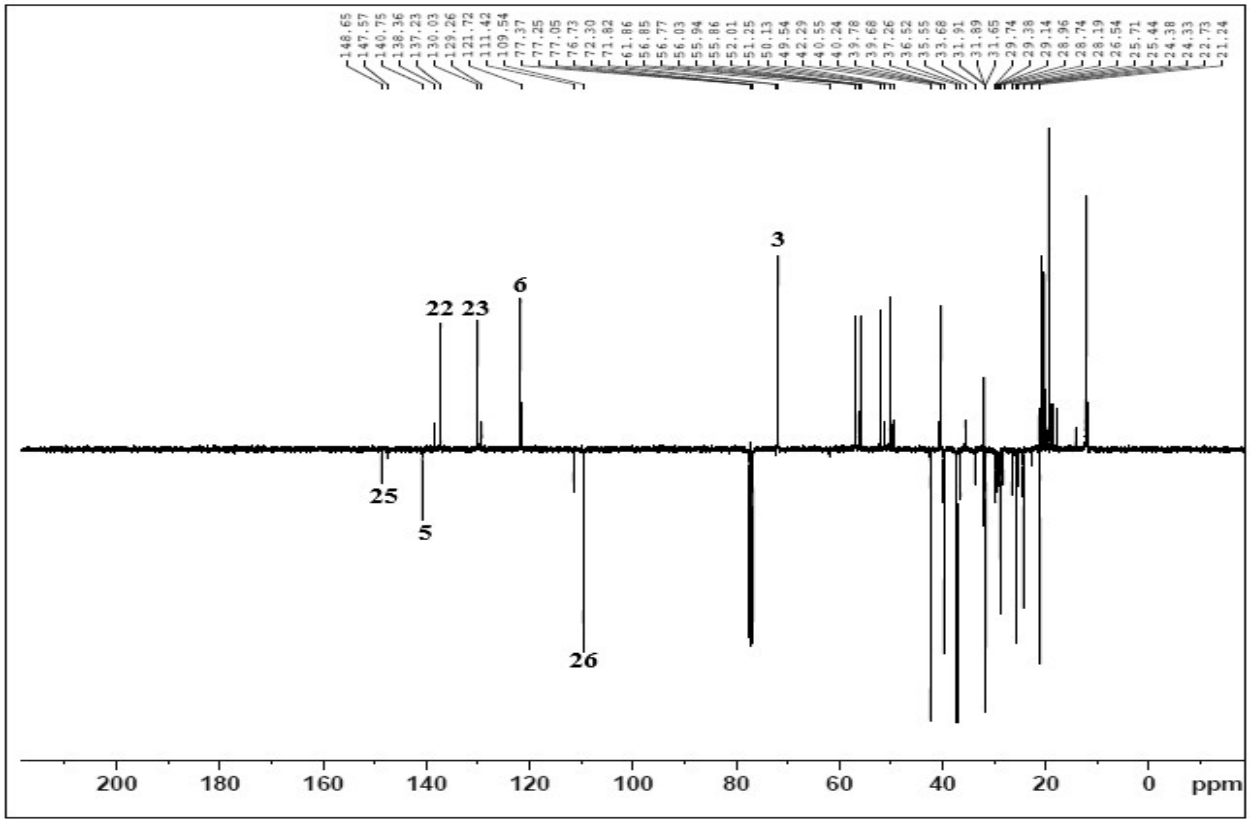


Figure S2: DEPTQ spectrum of compound 1 in $CDCl_3$

Compound 2

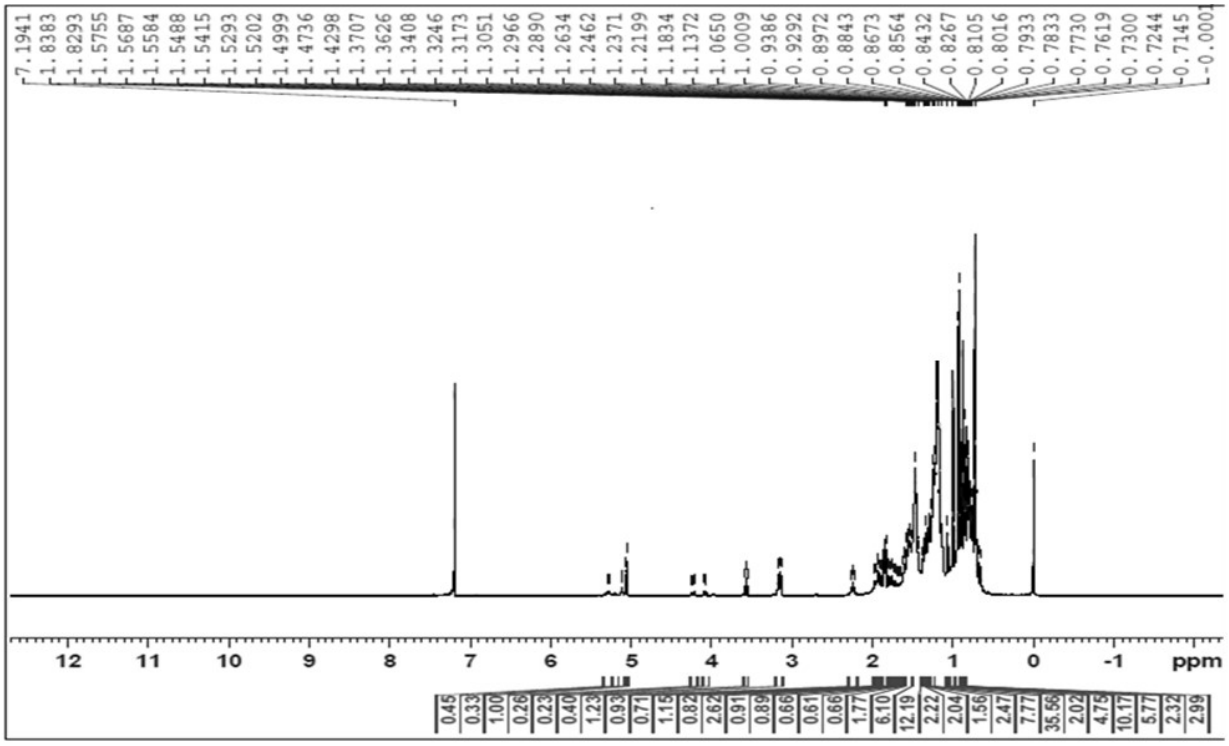


Figure S3: ^{1}H NMR spectrum of compound 2 in $CDCl_3$

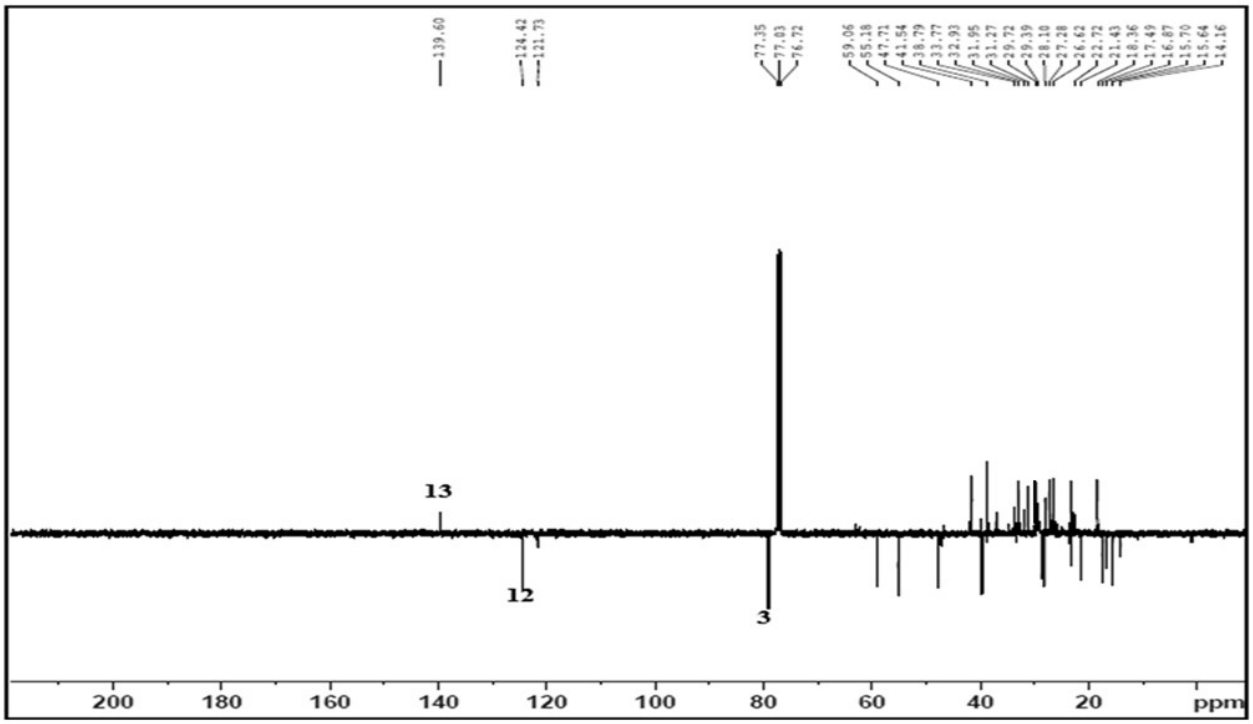


Figure S4: APT spectrum of compound 2 in $CDCl_3$

Compound 3

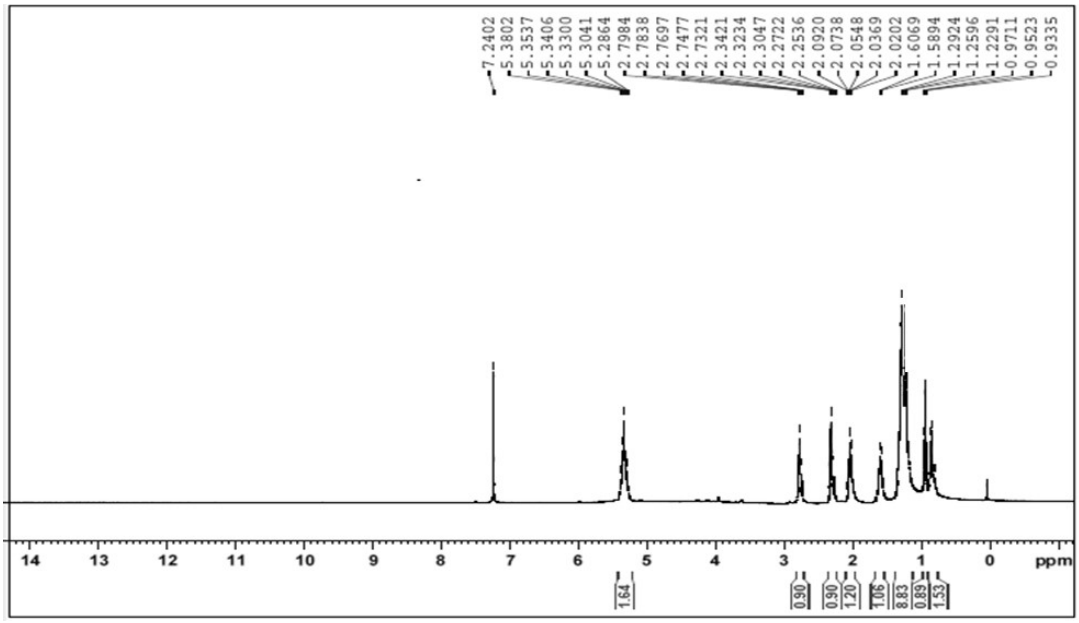


Figure 5: ^{1}HNMR spectrum of compound 3

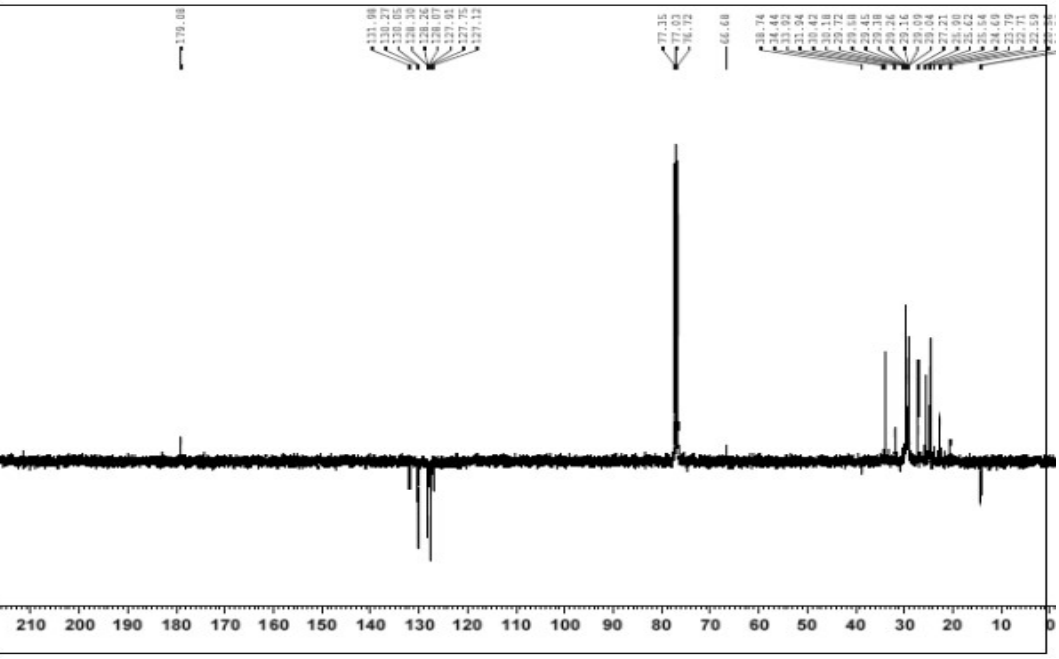


Figure S6: APT spectrum of compound 3 in $CDCl_3$

Compound 4

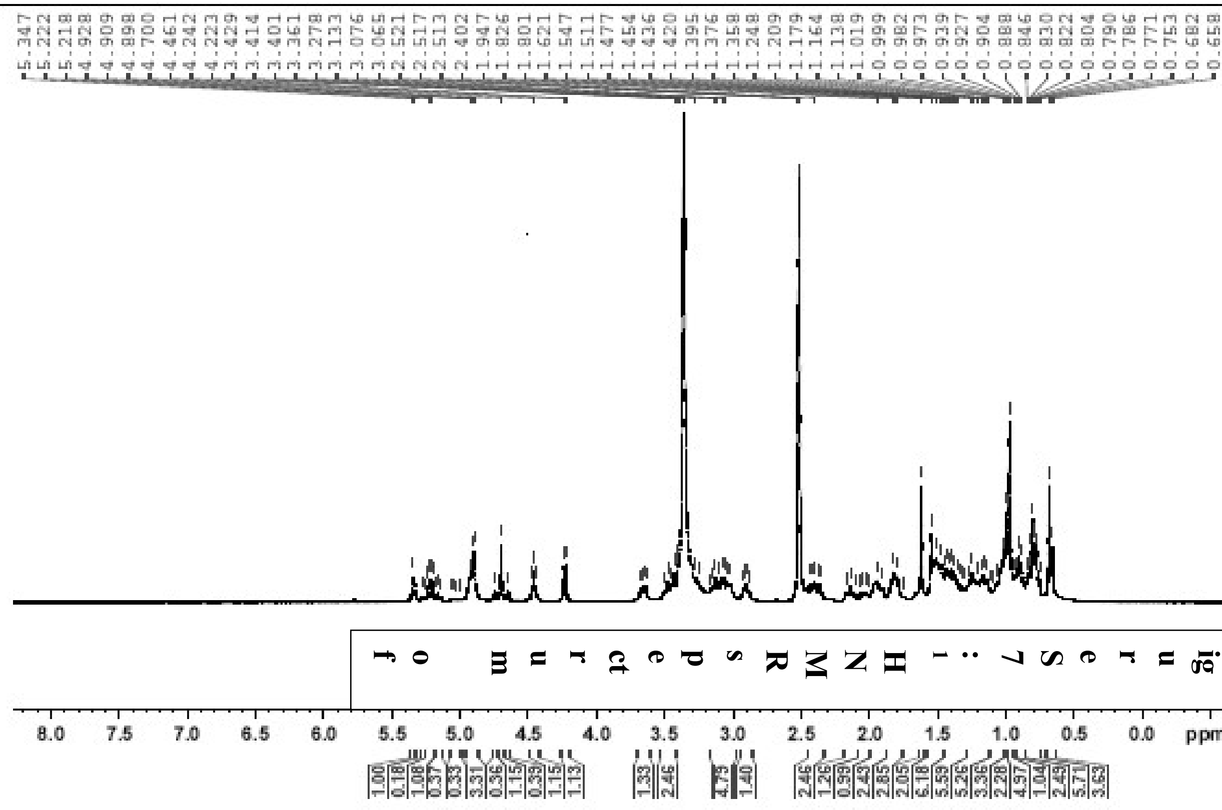
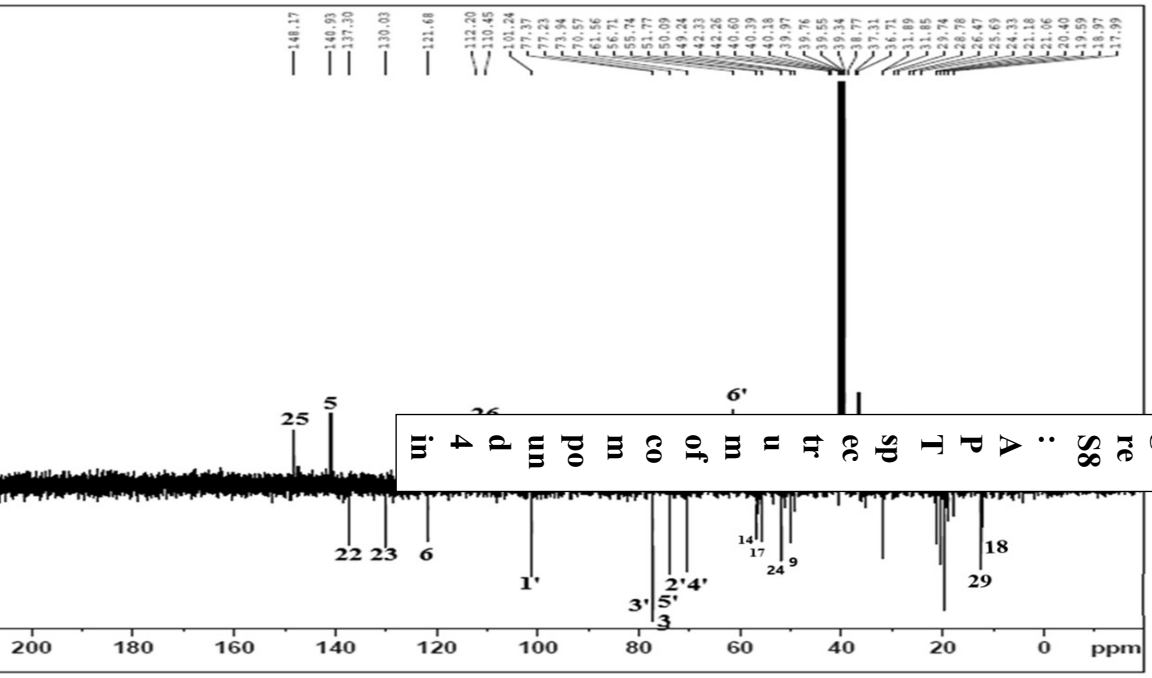


Figure S7: ^1H NMR spectrum of



Fi gu re S8 : A P T sp ec tr u m of co m po un d 4 in

Compound 5

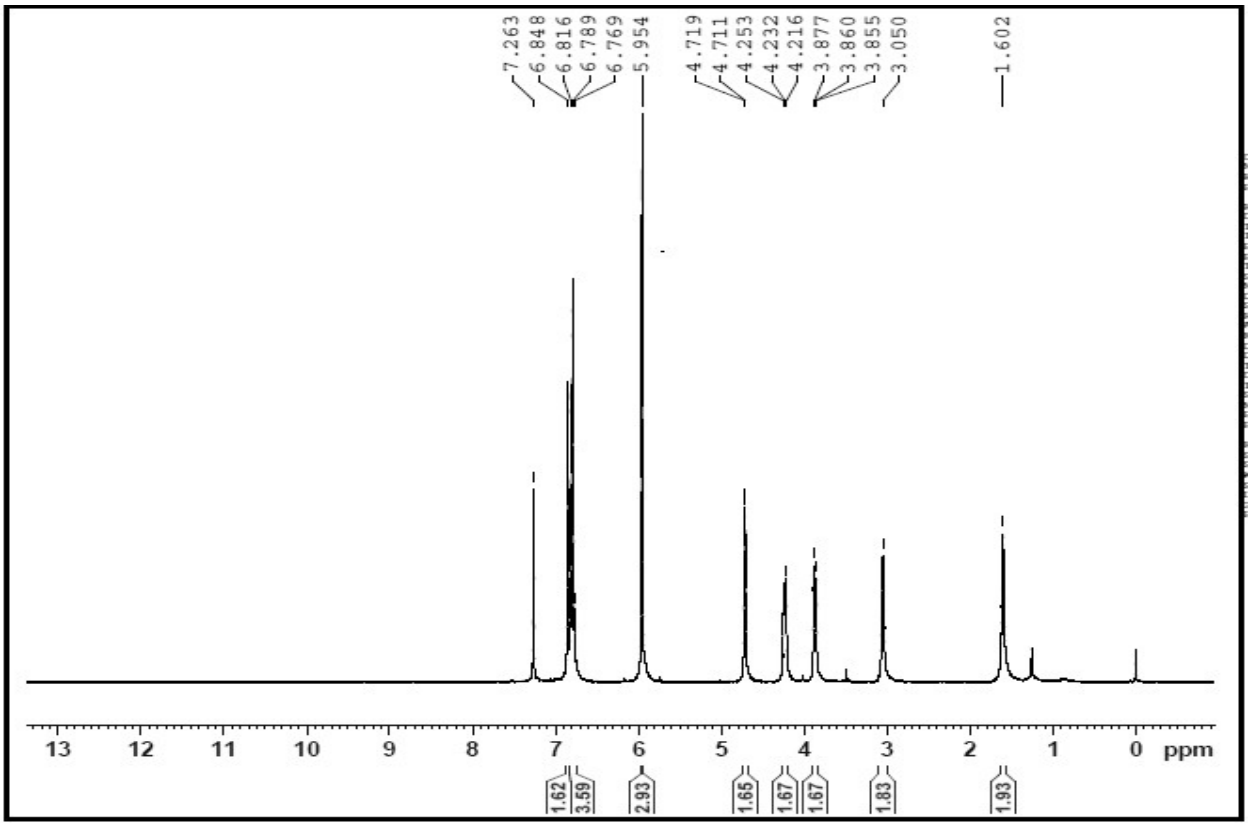


Figure S19: Unreported 1H NMR

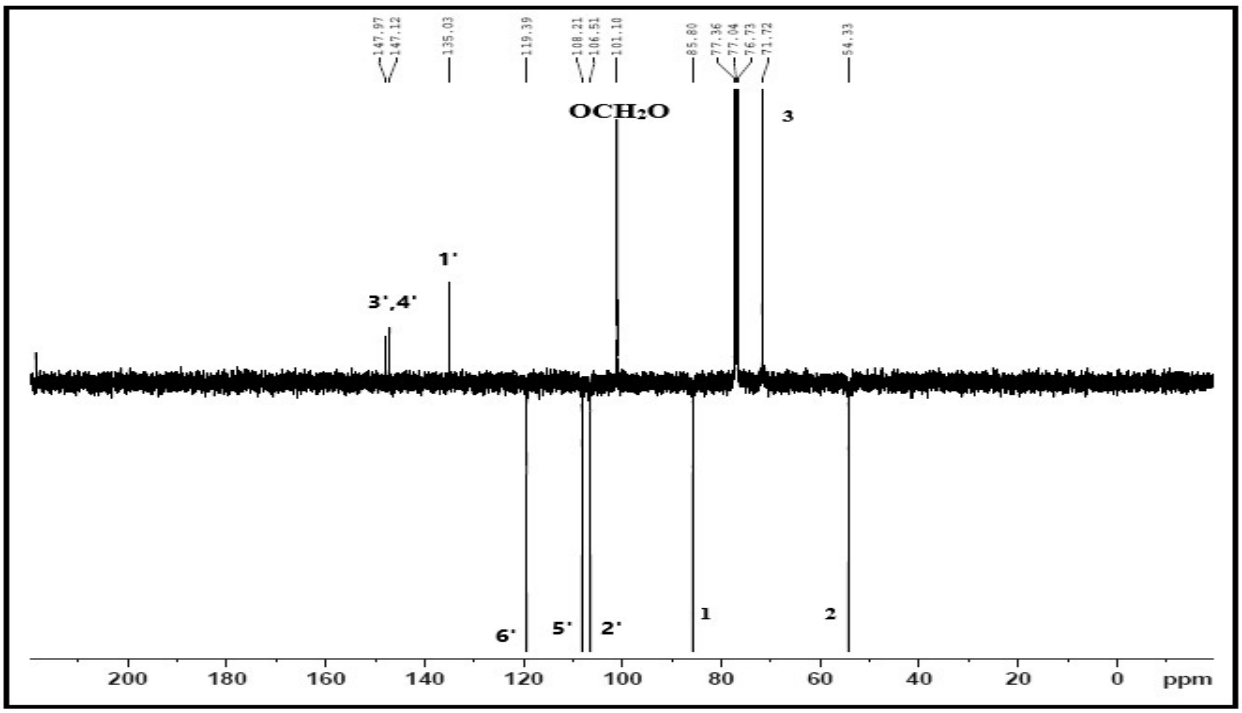


Figure S20: APT spectrum of compound 5in $CDCl_3$

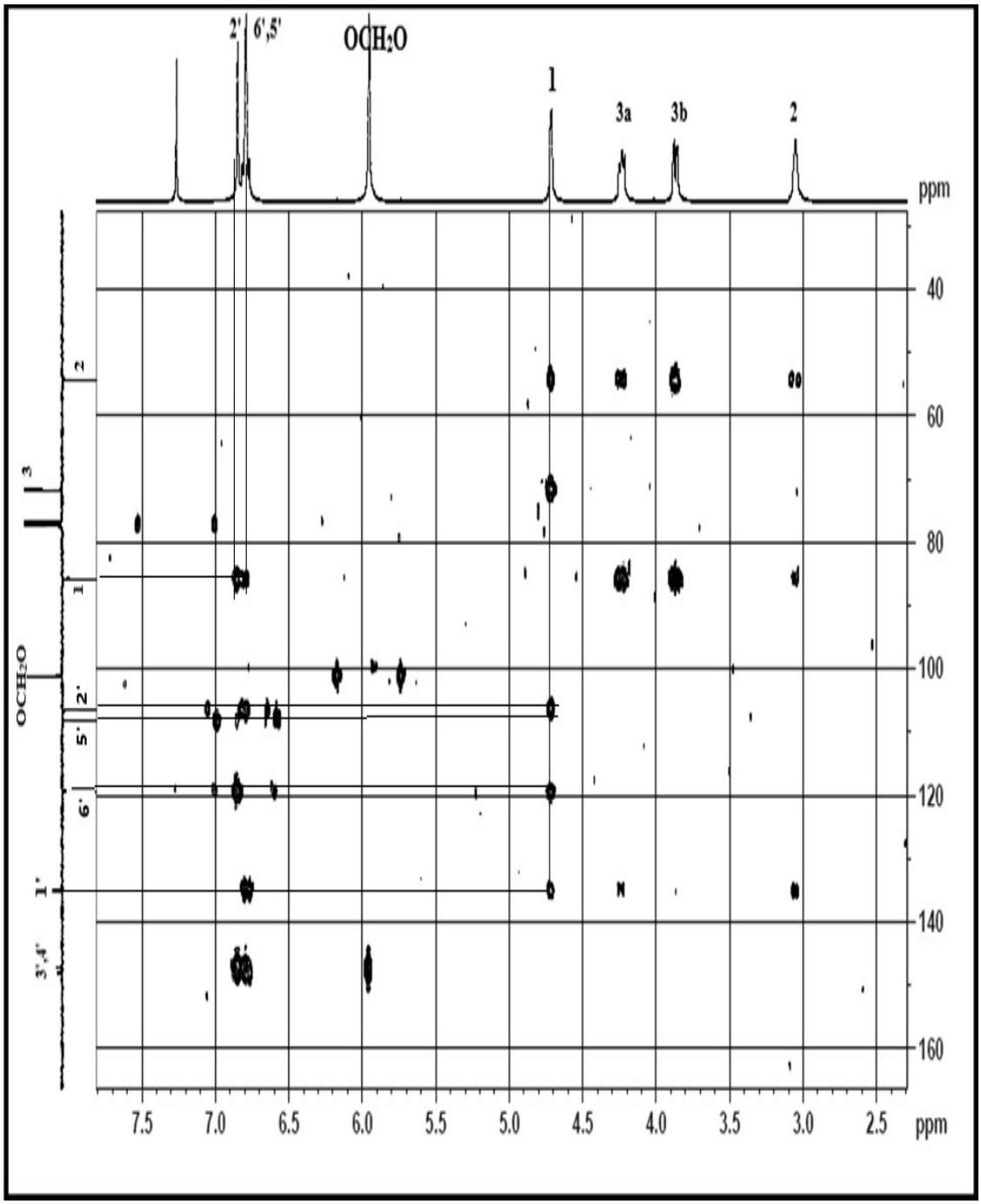


Figure S11: HMBC correlations of compound 5 in $CDCl_3$

Compound 6

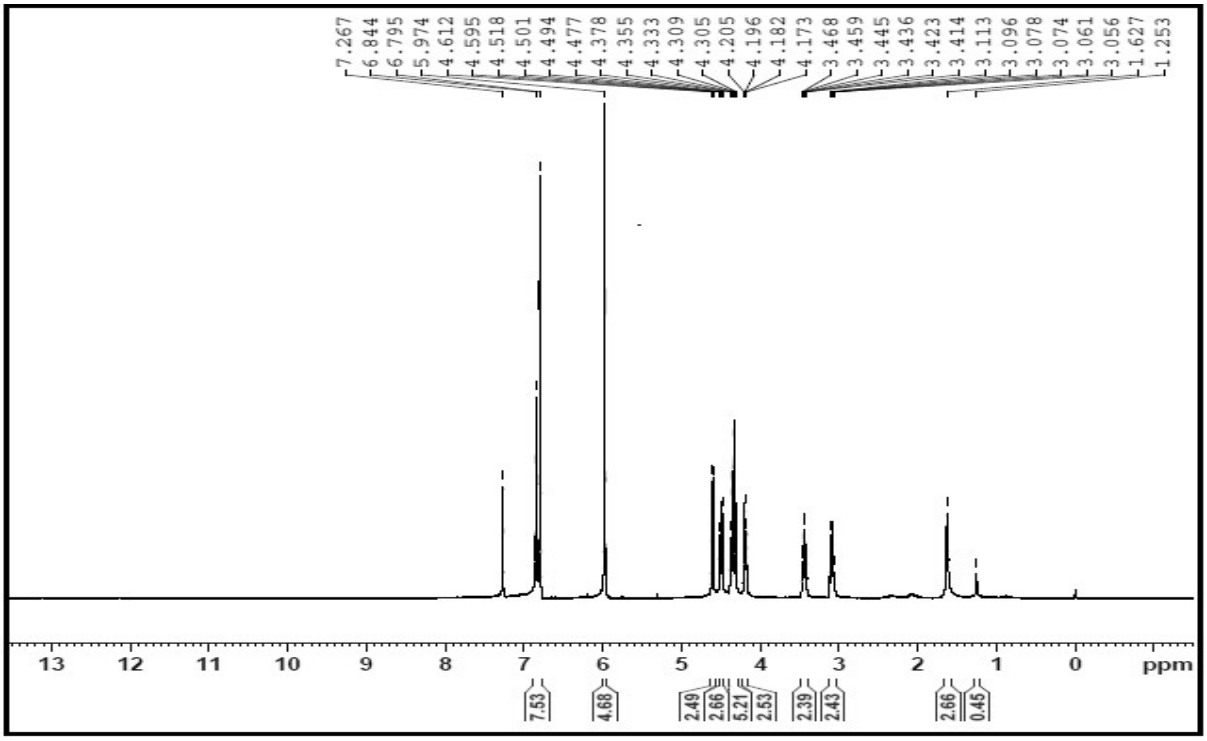


Figure S11: ^{1}H NMR spectrum of compound 6in CDCl$_3$

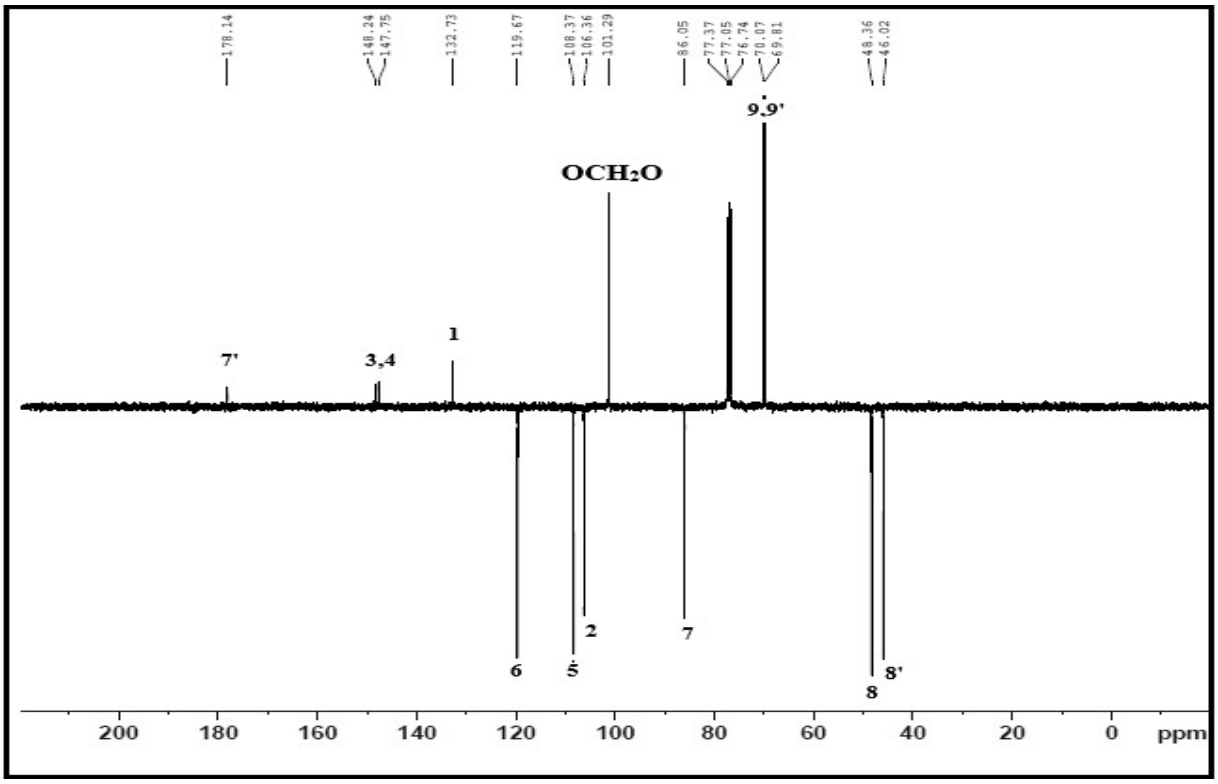


Figure S13: APT spectrum of compound 6in CDCl$_3$

Compound 7

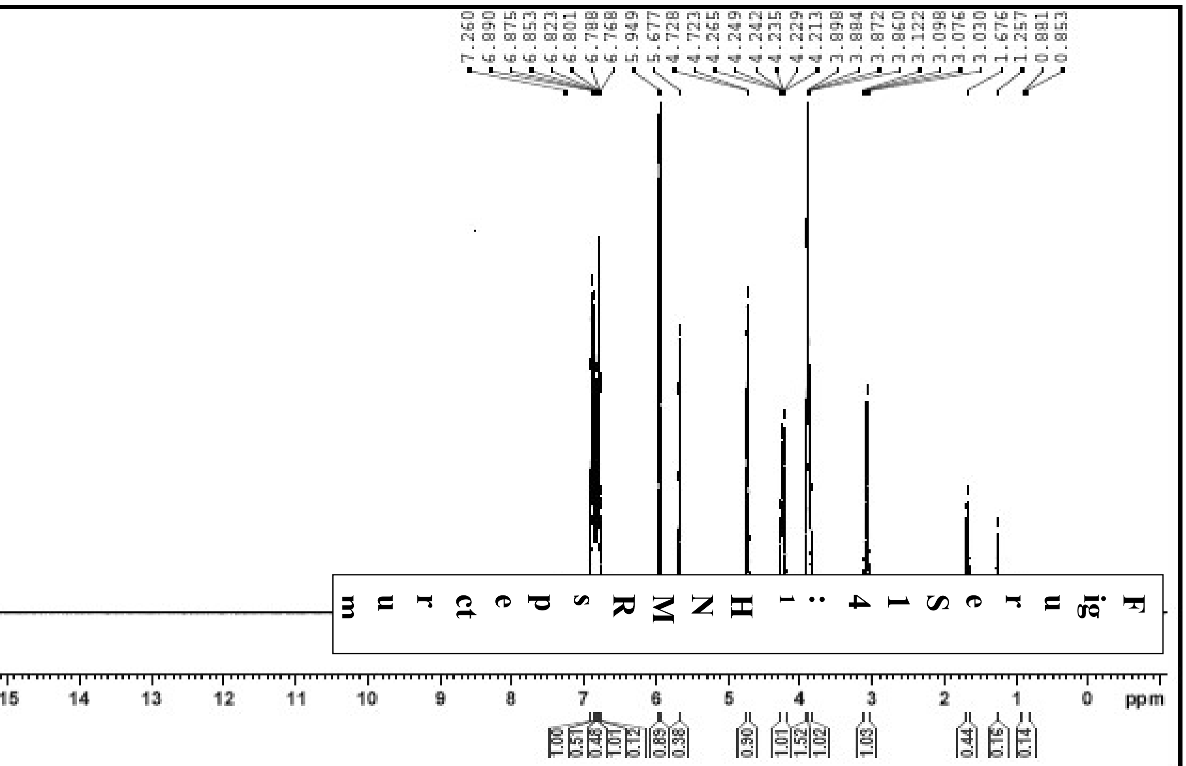
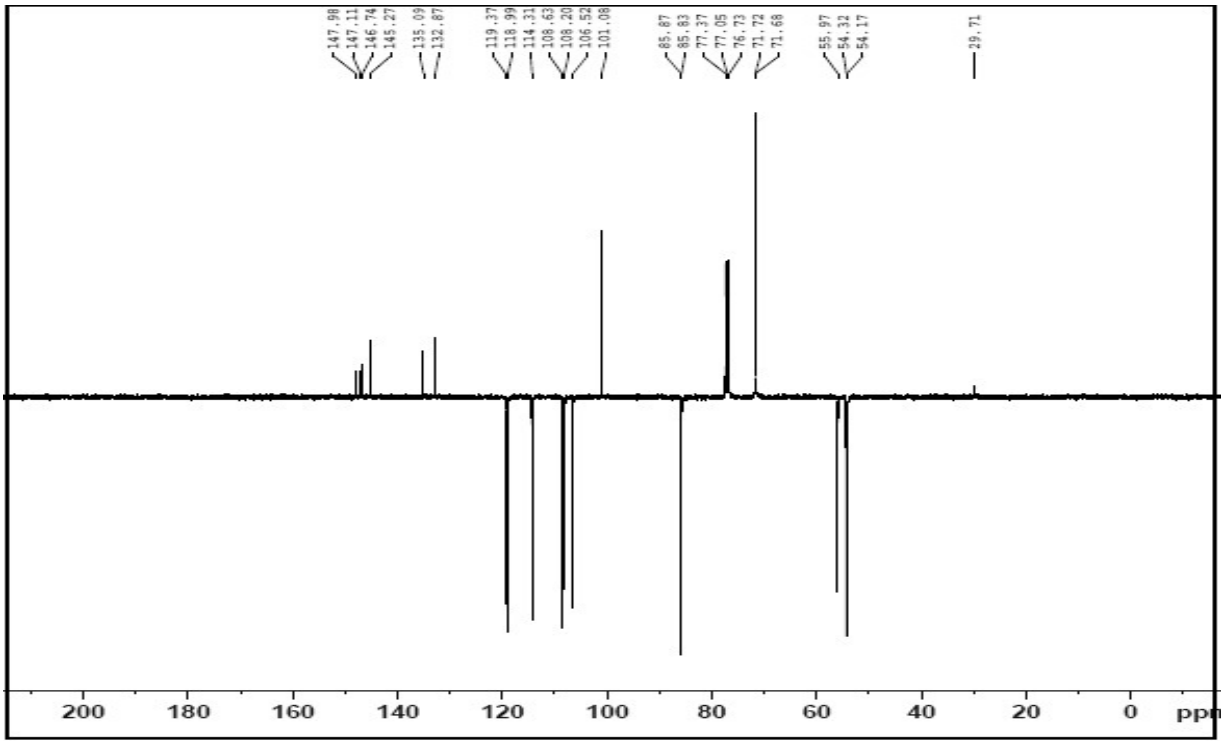


Figure S14: 1H NMR Spectrum



Compound 8

Figure S15: APT spectrum of compound 7 in $CDCl_3$

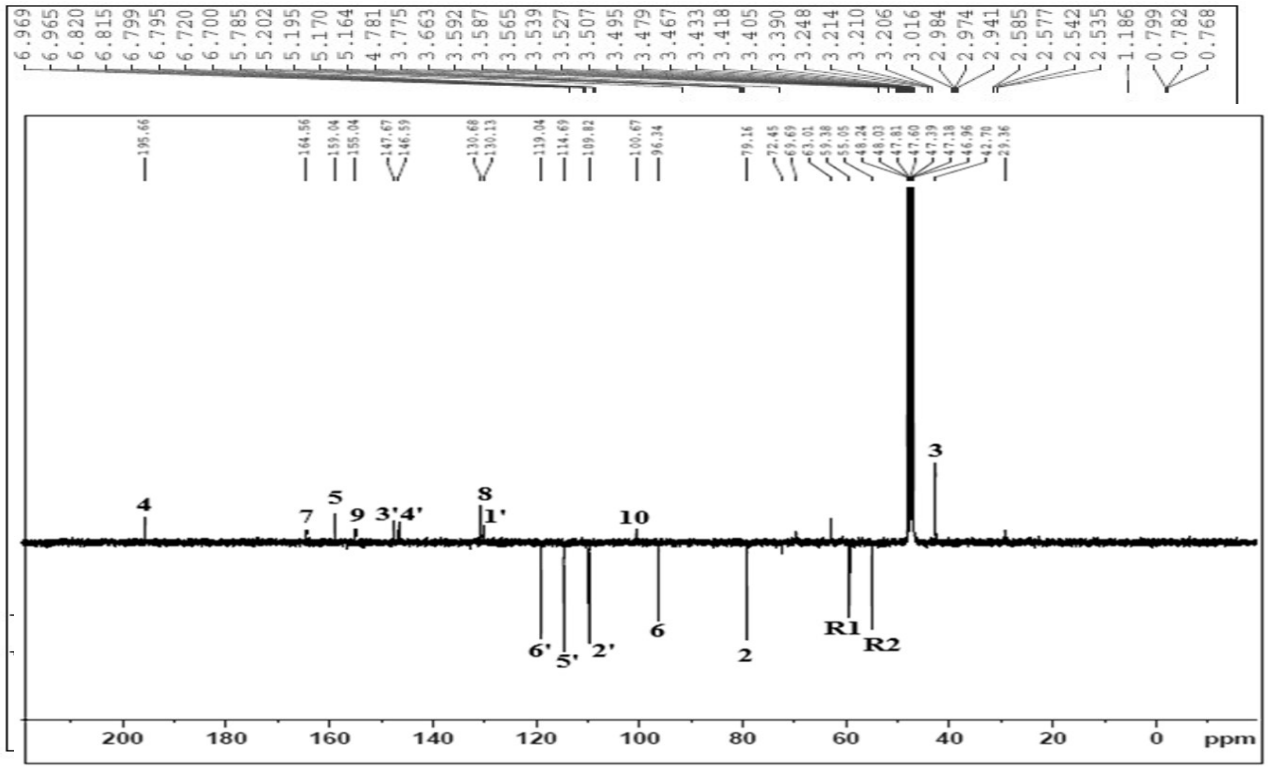


Figure S16: ^{1}H NMR spectrum of compound 8

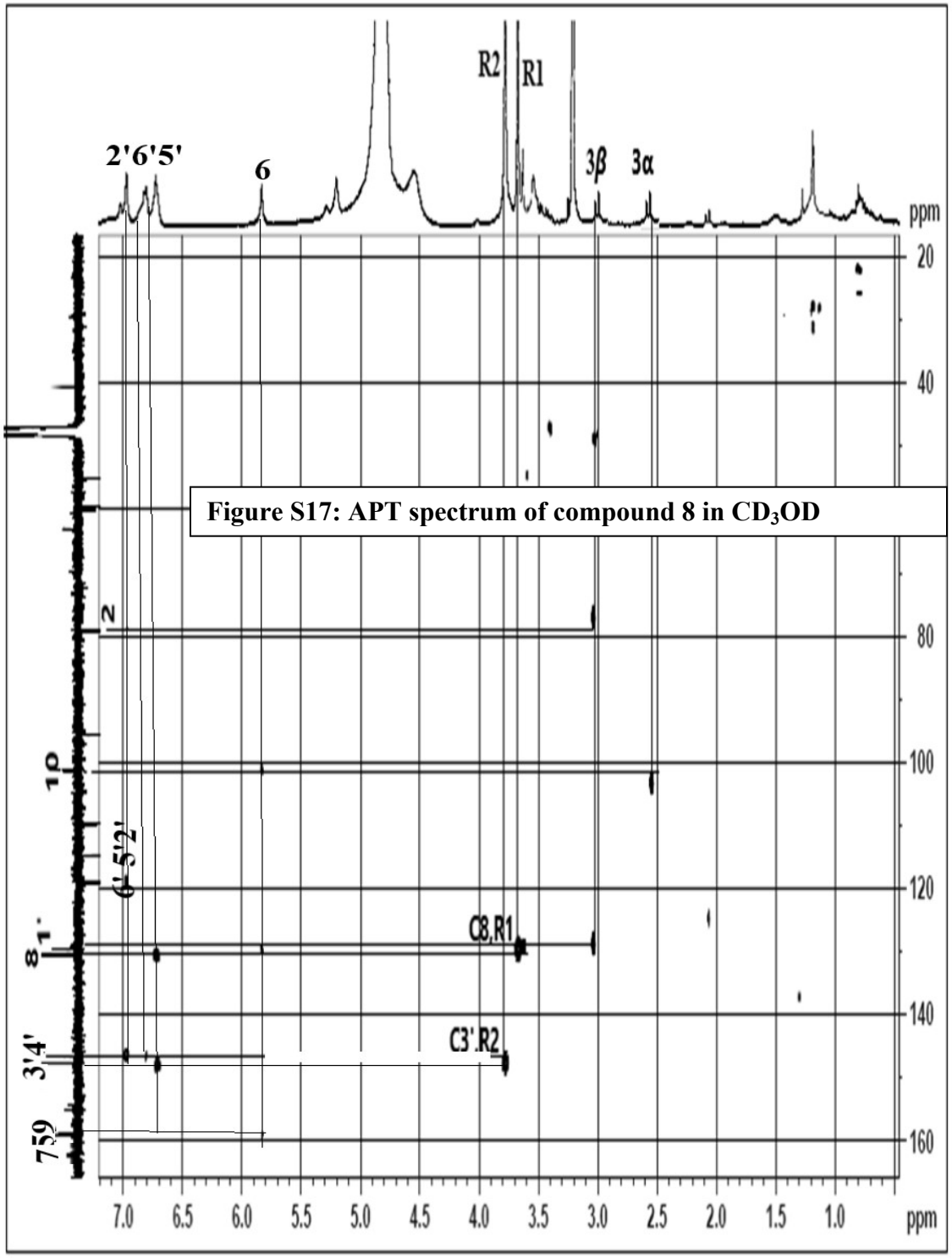


Figure S17: APT spectrum of compound 8 in CD_3OD

Figure S18: HMBC spectrum of compound HM 8

Compound 9

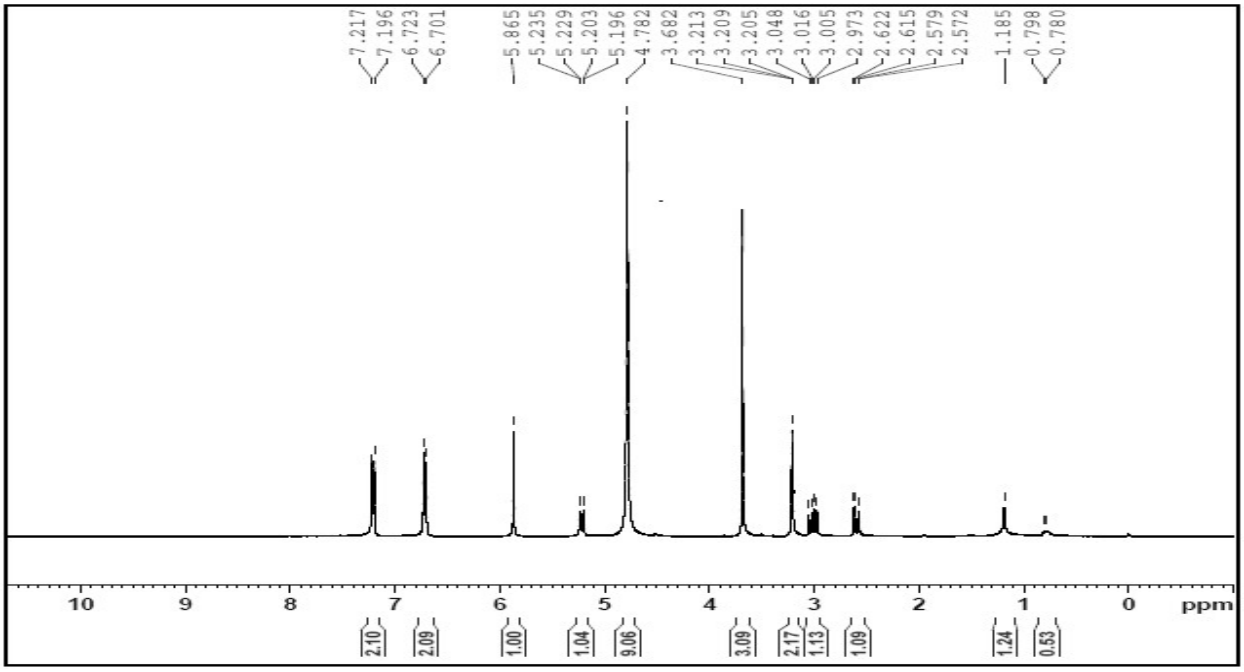


Fig. 11: Serum HNMR spectrum

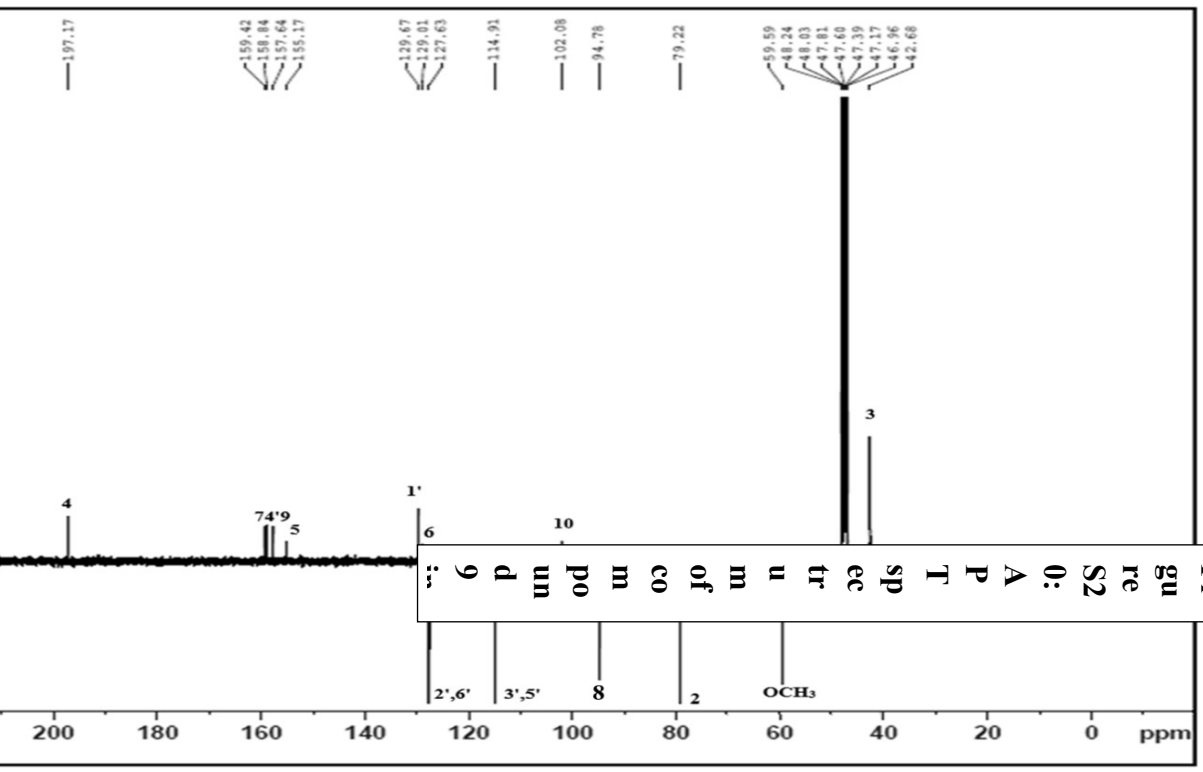


Figure S20: APT spectrum of compound 10

Compound 10

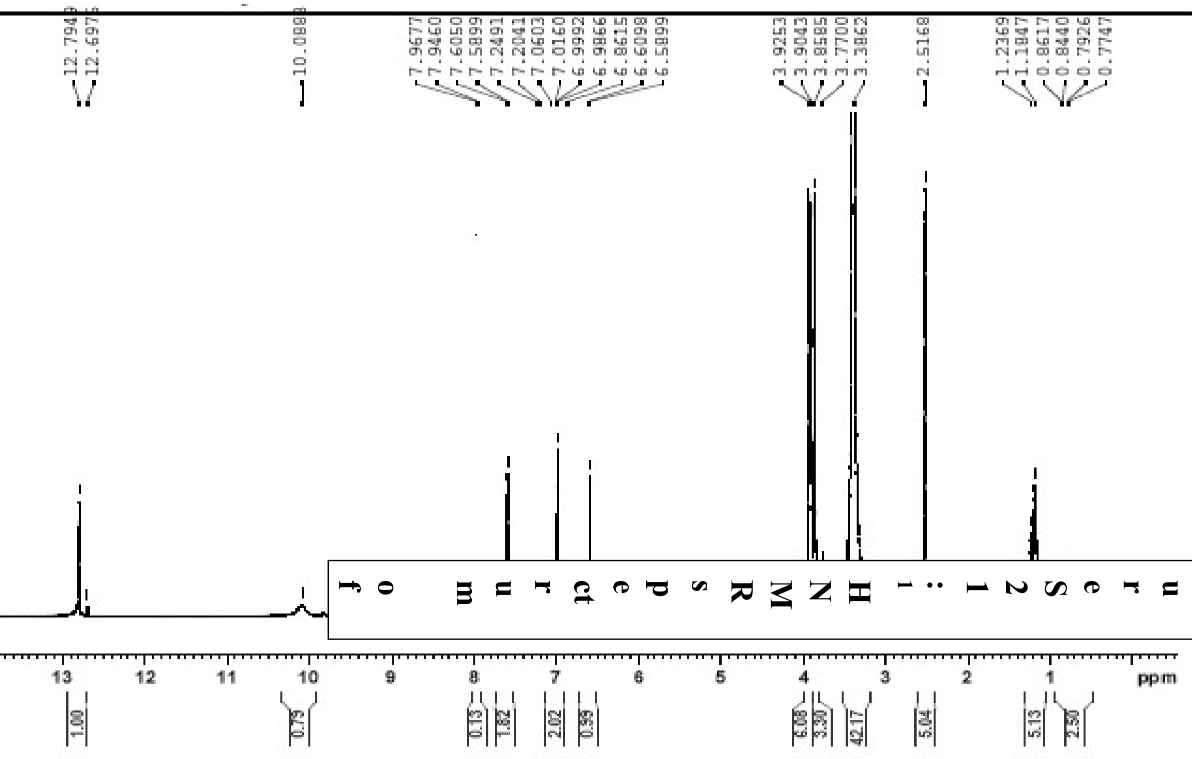


Figure S21: ^1H NMR spectrum of

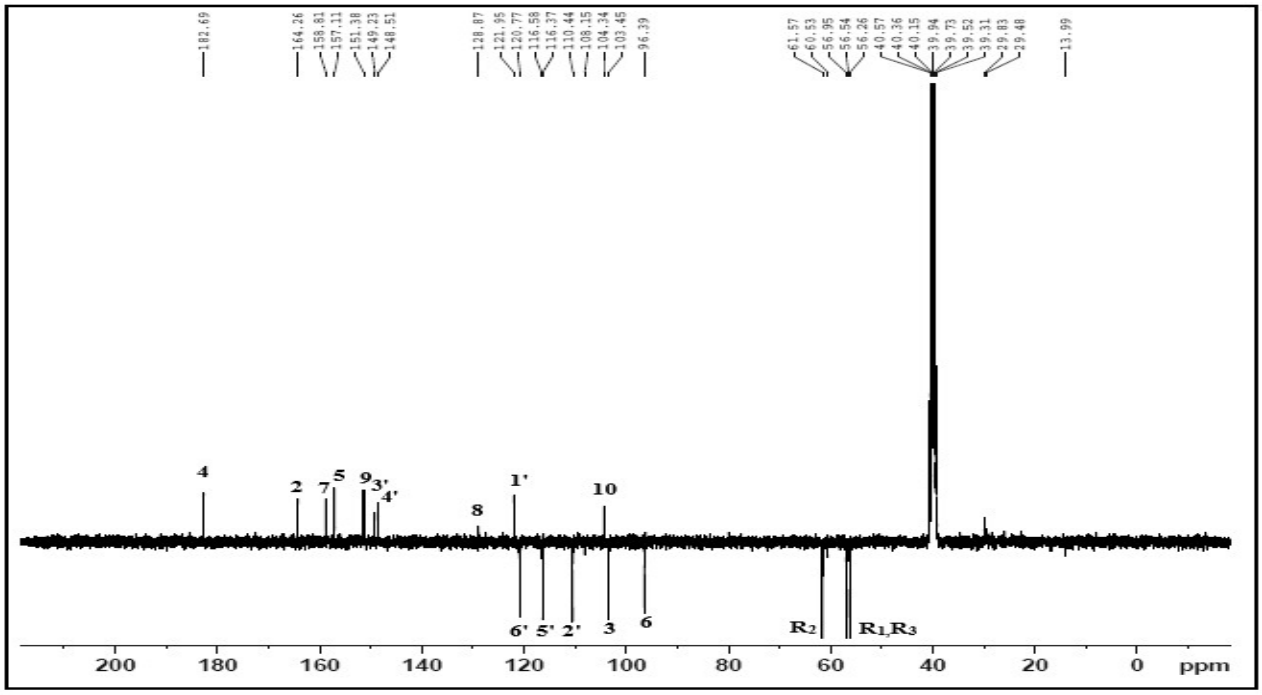


Figure S22: DEPTQ spectrum of compound 10 in DMSO-d_6

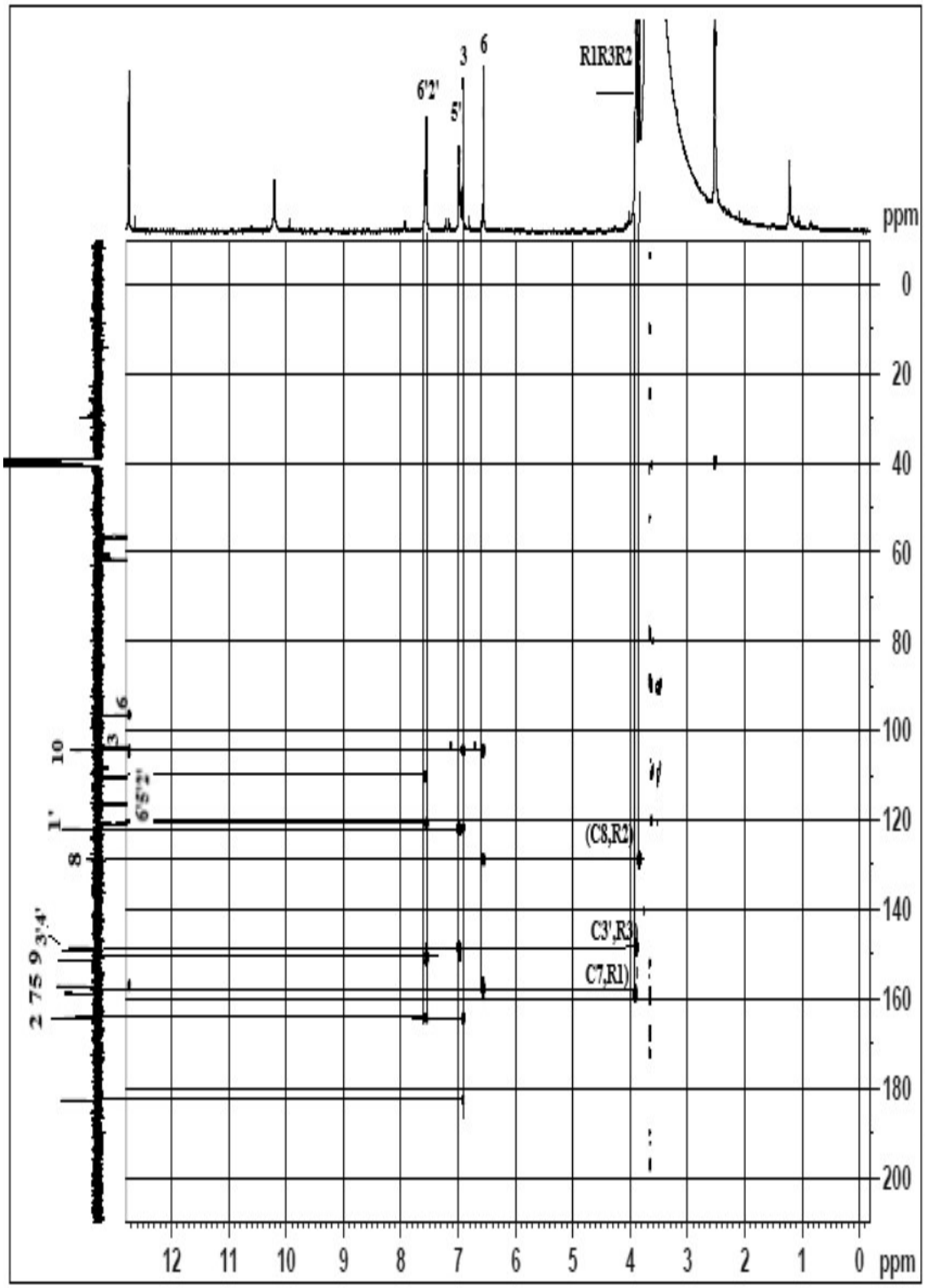


Figure S23: HMBC spectrum of compound 10

Compound 11

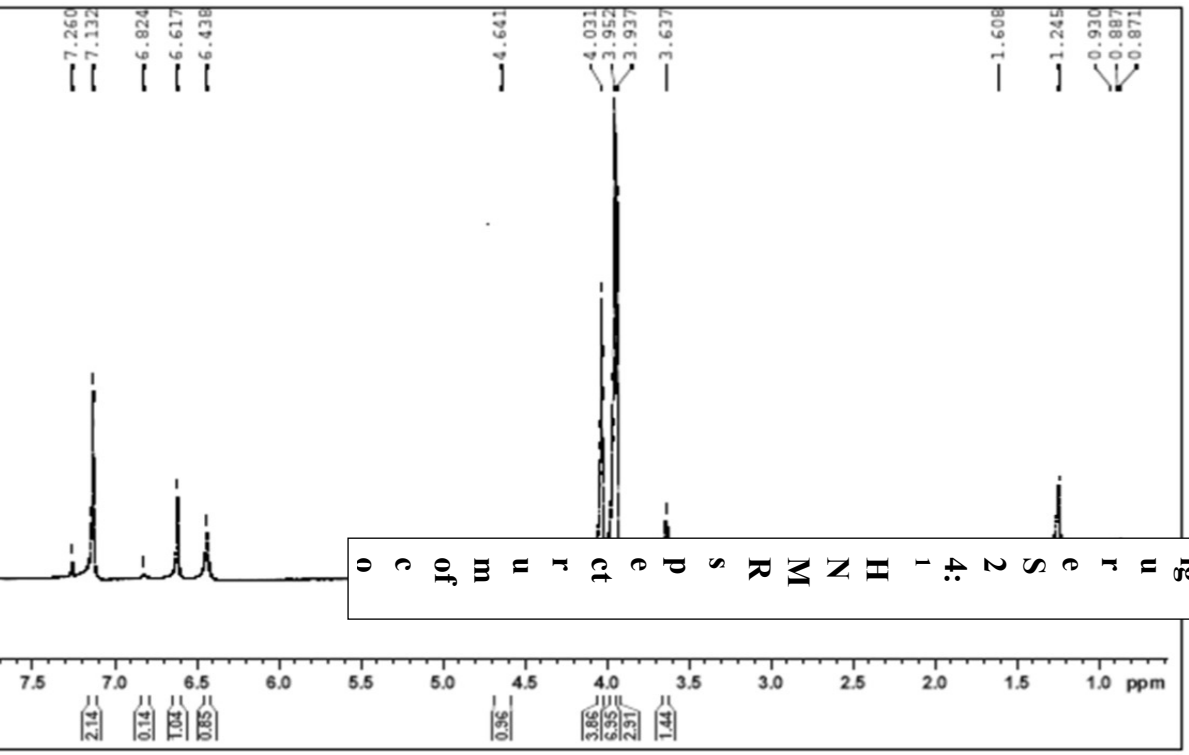
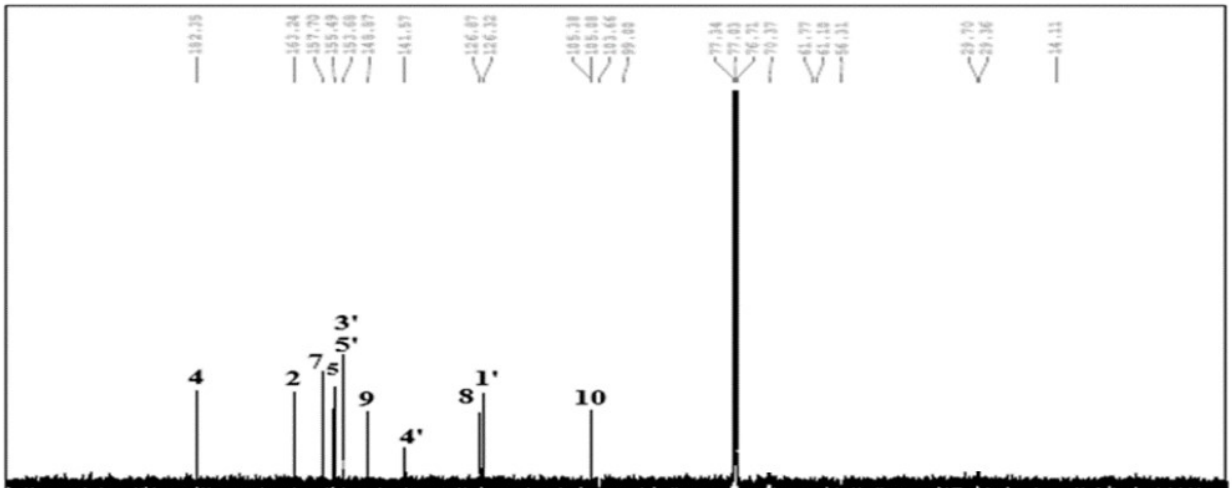


Figure 4: 1HNMR spectrum of compound 11



Compound 12

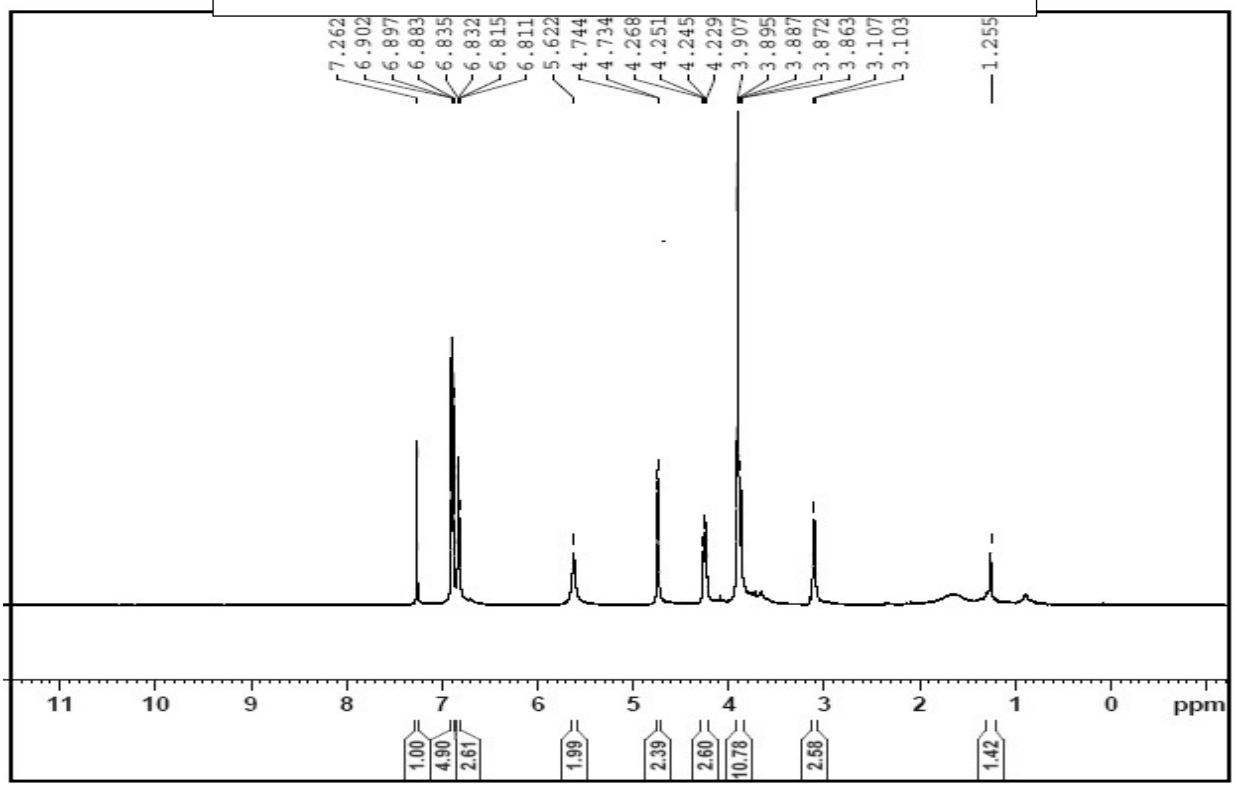


Figure S2: ^{1}H NMR spectrum of compound 12 recorded in CDCl3 at 400 MHz

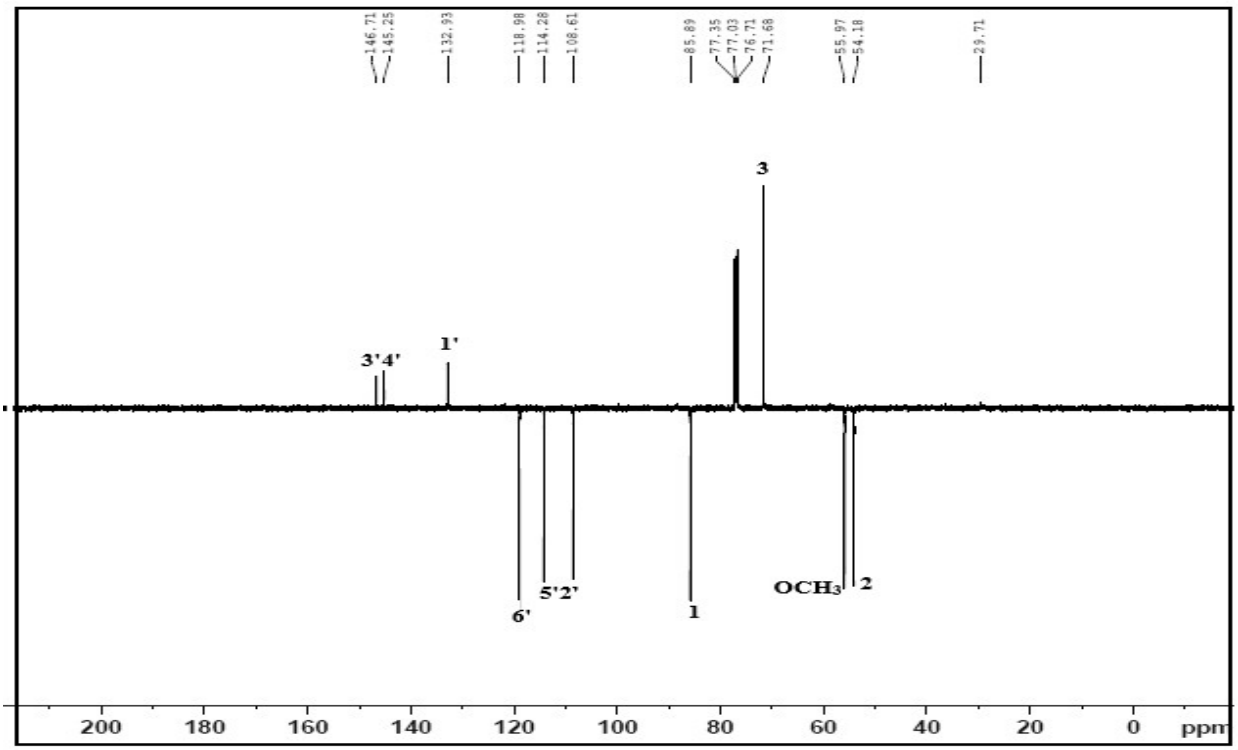


Figure S27: APT spectrum of compound 12 in CD_3OD

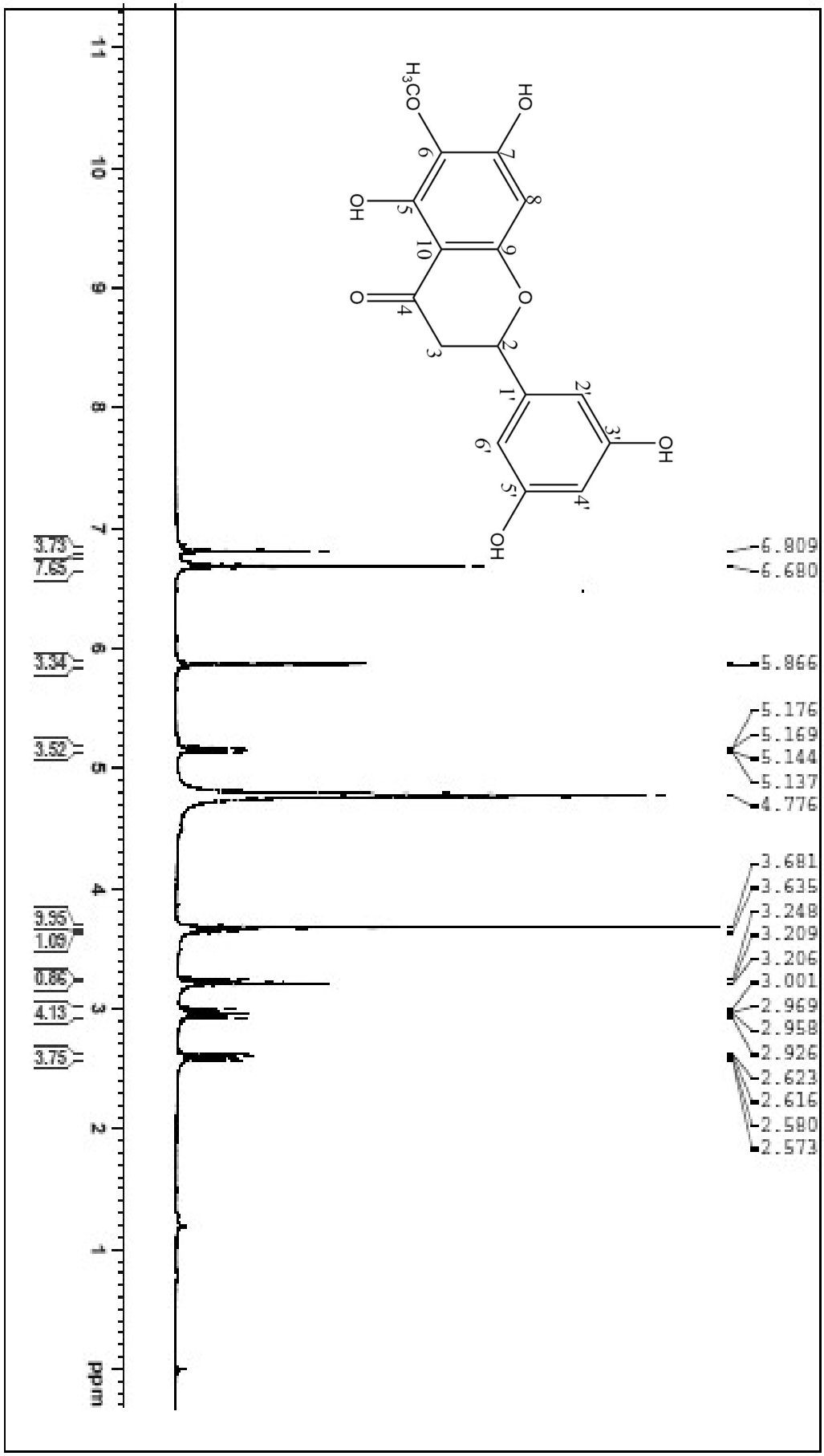


Figure S28: ^{1}H NMR spectrum of compound 13 in CD_3OD

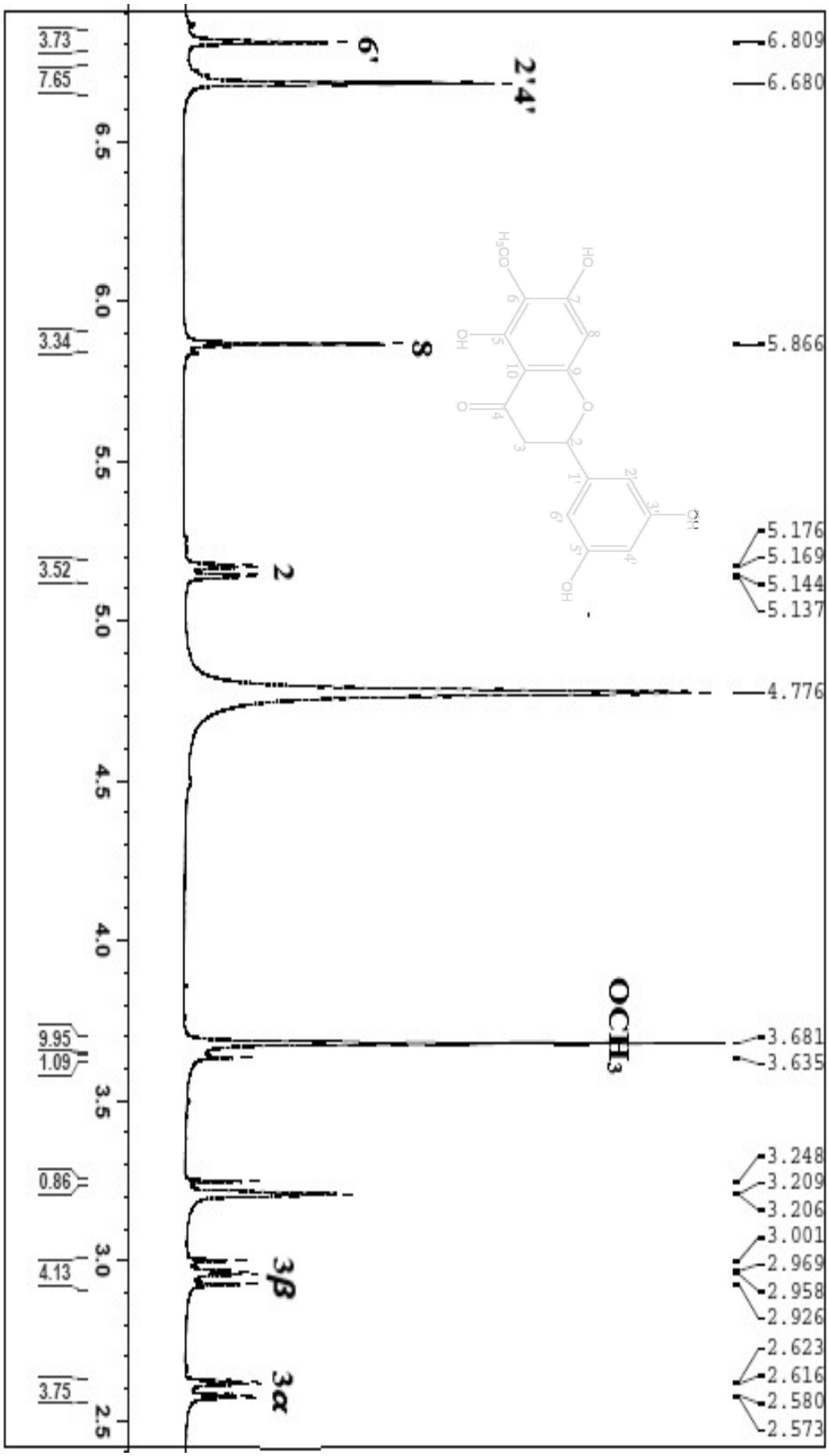


Figure S29: ^{1}H-NMR spectrum of compound 13 (expansion at δ(2.5-6.8 ppm)

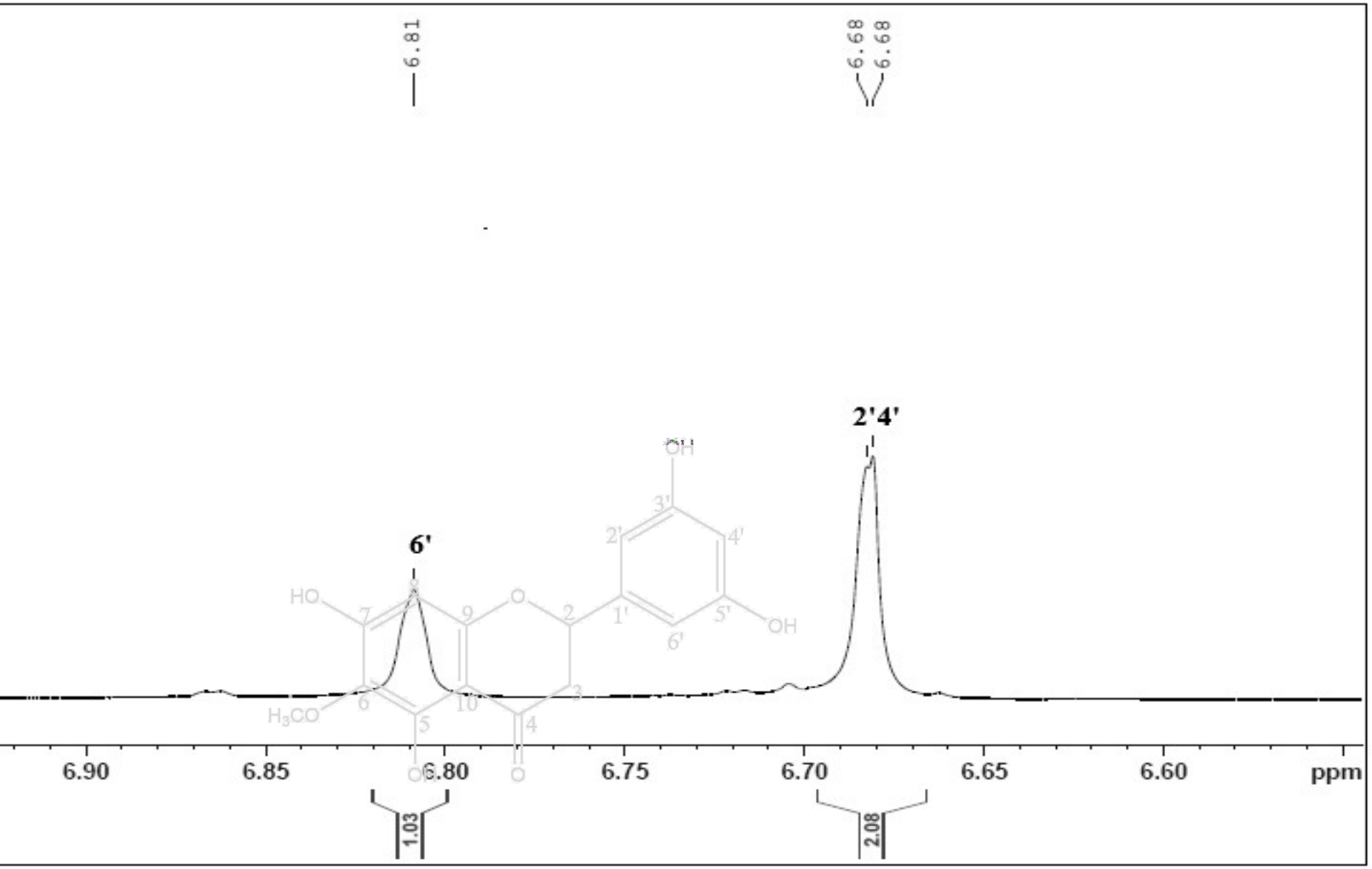


Figure S30: ^{1}H-NMR spectrum of compound 13 (expansion at δ(6.60-6.95 ppm)

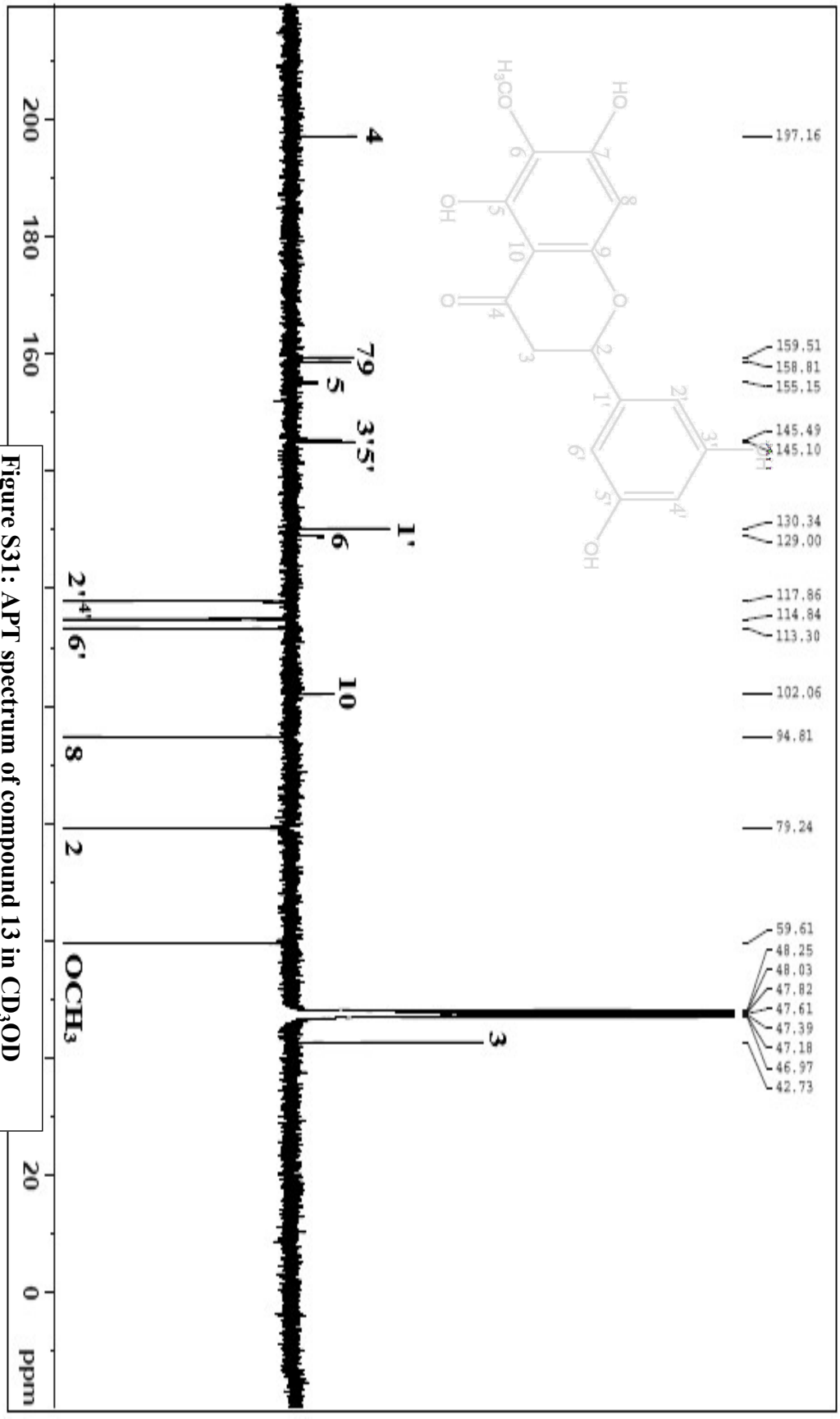


Figure S31: APT spectrum of compound 13 in CD_3OD

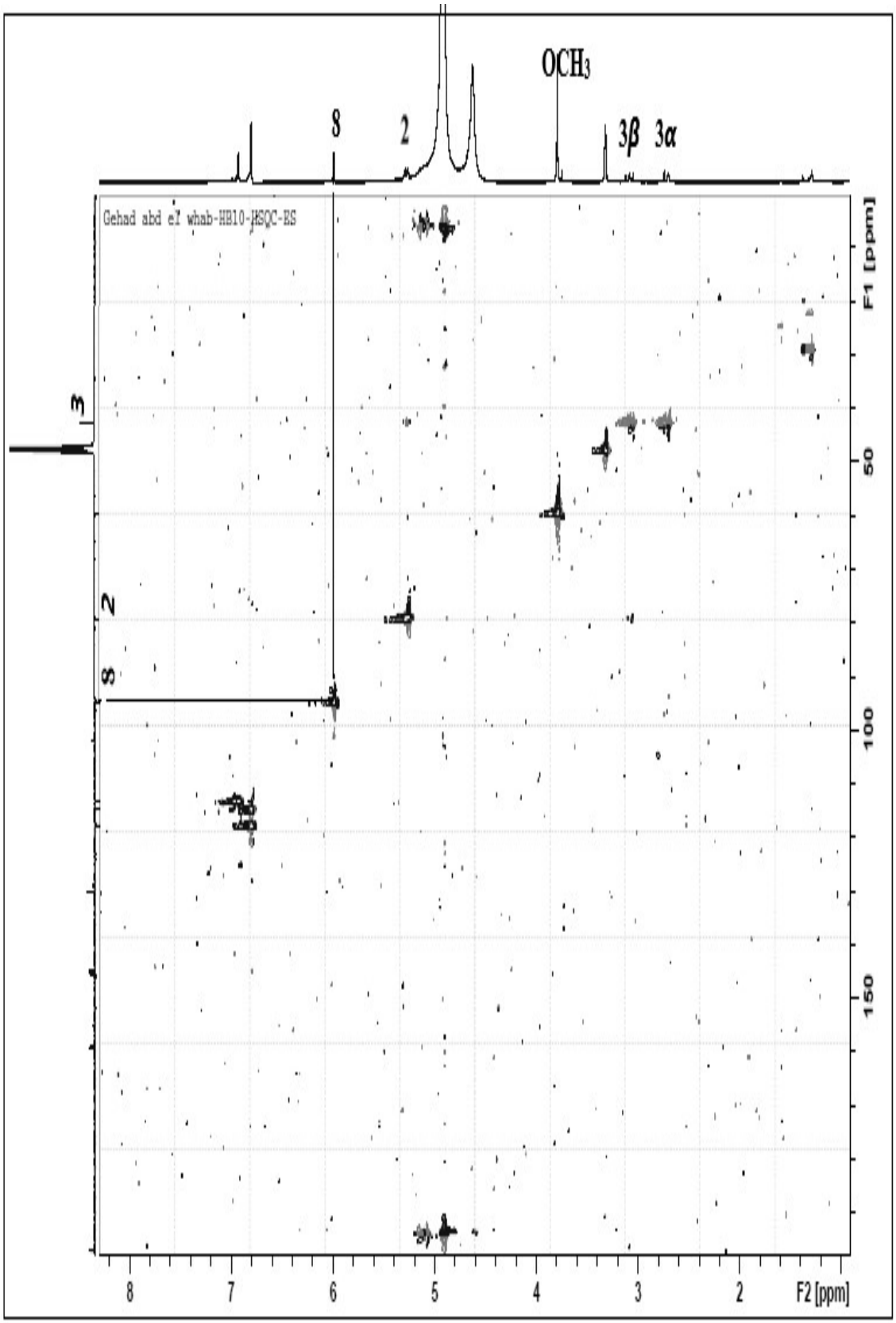


Figure S32: HSQC spectrum of compound

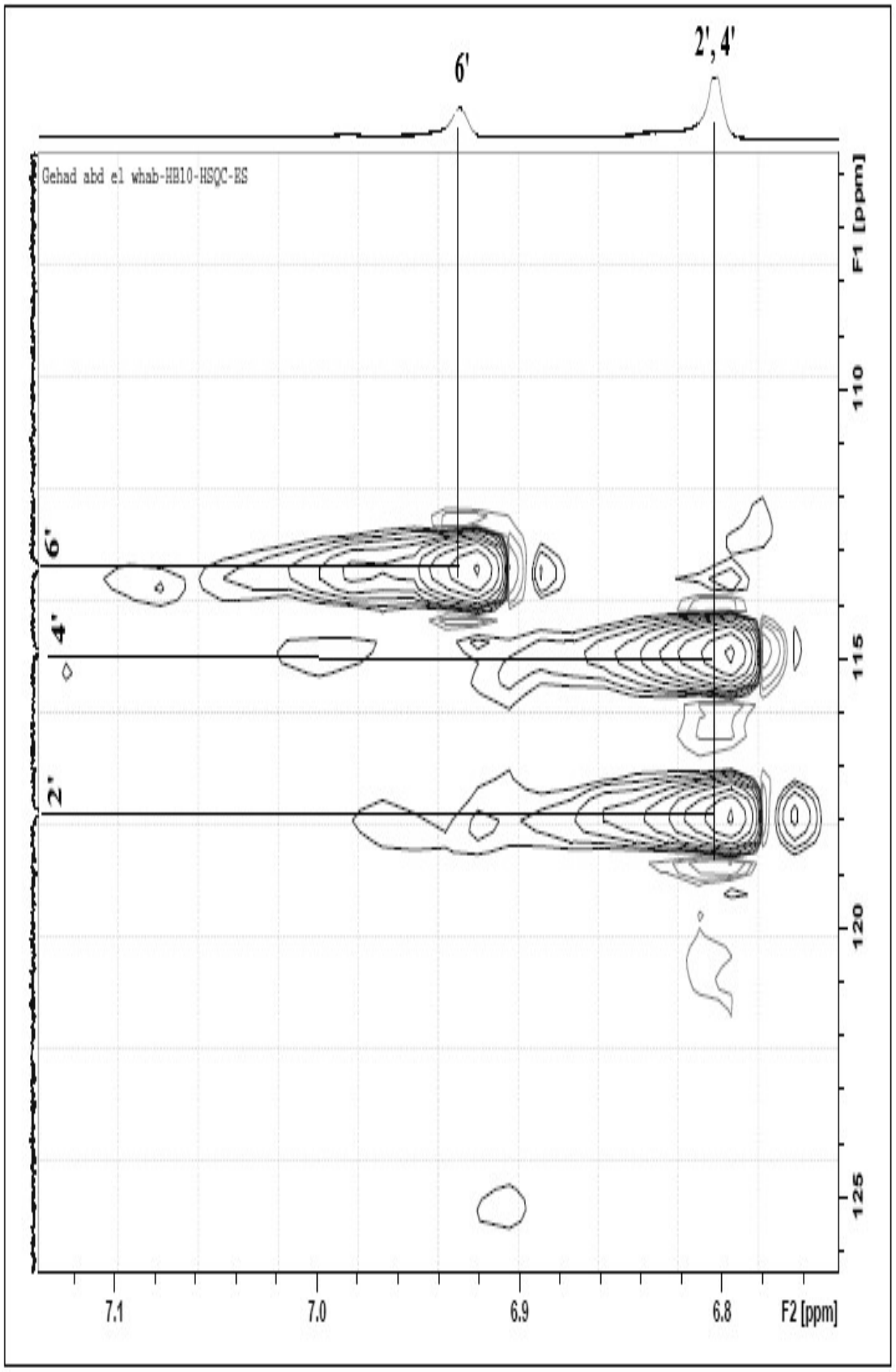


Figure S3: HSQC experiment of compound

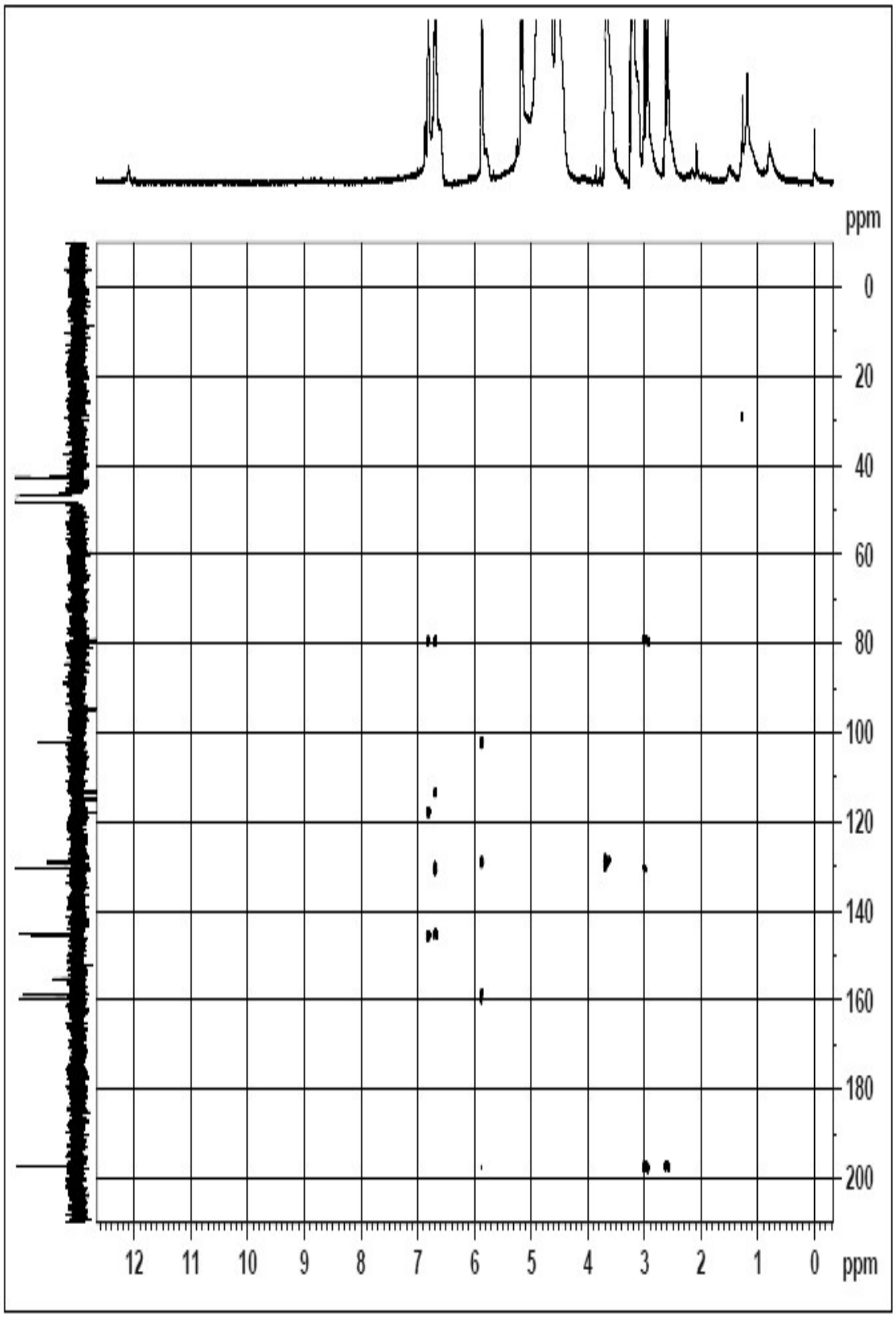


Figure S34: HMBC spectrum of compound

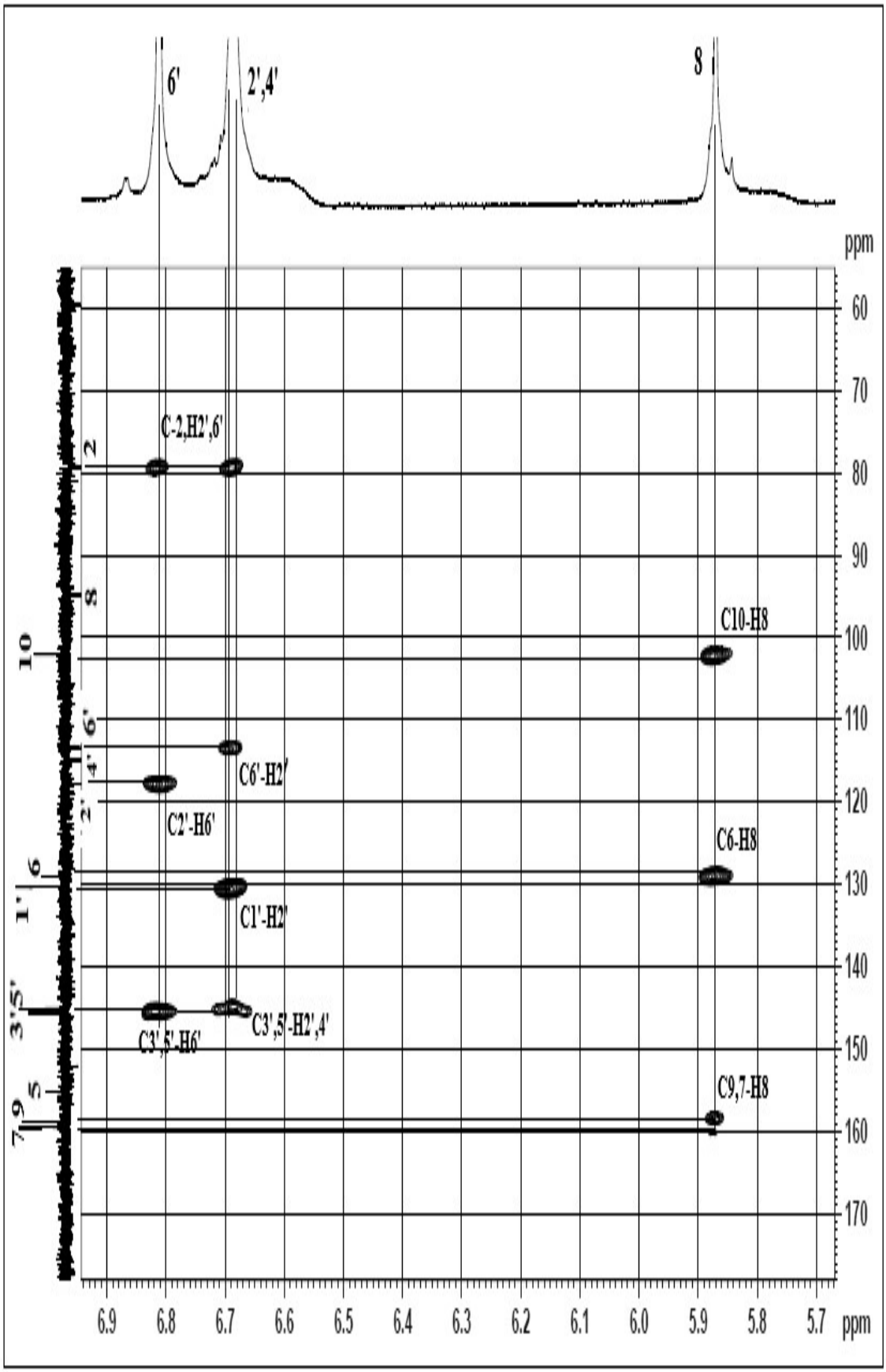


Figure S35: HMBC experiment of compound

[
M
+

Figure S36: LC-MS (+ ESI) spectrum of compound 13

Figure (37A): Proposed fragmentation pattern for compound 13 according to LC-ESI+-MS positive mode spectrum

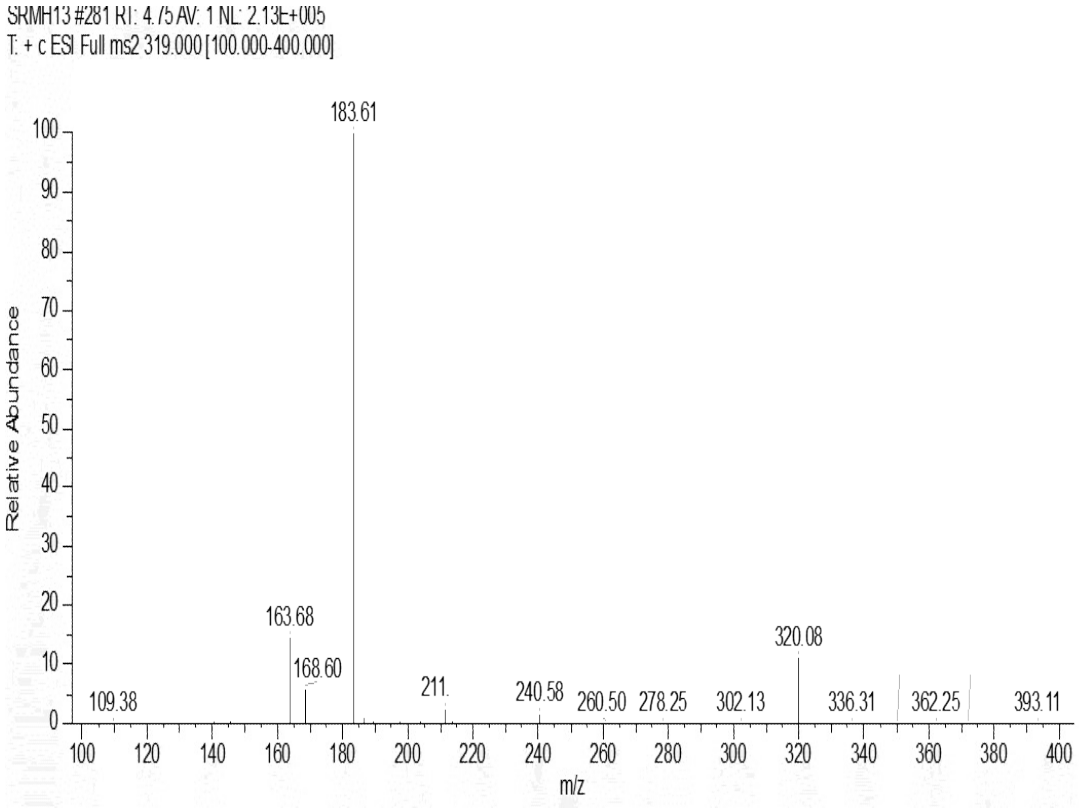
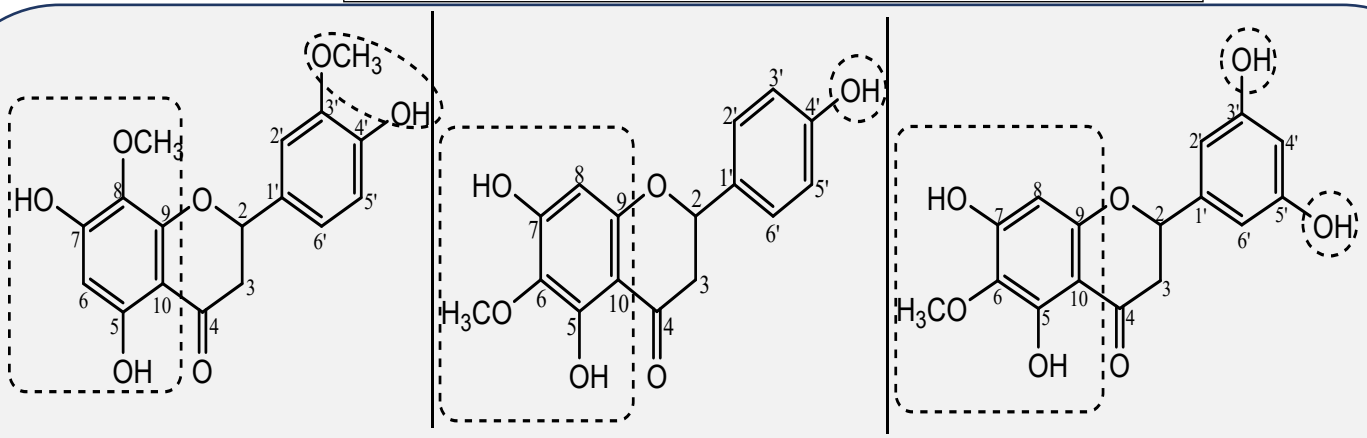


Figure S37B: LC-MS (+ ESI) fragmentation of component



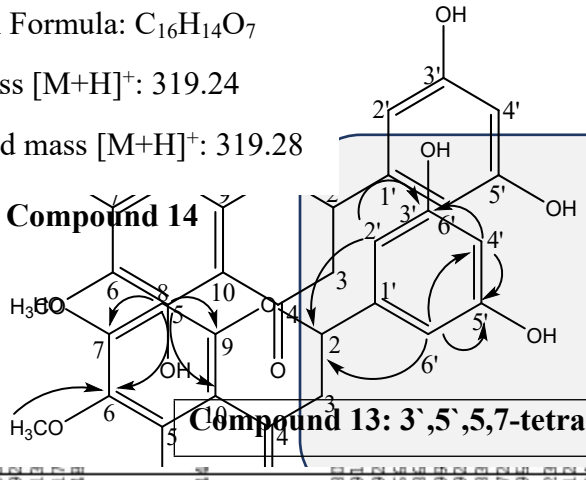
Compound HM 8	Compound HM9	Compound HM 13
5,7,4`- trihydroxy- 8,3` dimethoxy flavanone	5,7,4`-trihydroxy-6-methoxy flavanone	3`,5`,5,7-tetrahydroxy-6-methoxyflavanone

Figure S38: Structural differences between compounds 8,9& 13

Chemical Formula: $C_{16}H_{14}O_7$

Exact mass $[M+H]^+$: 319.24

Calculated mass $[M+H]^+$: 319.28



Compound 14

Compound 13: 3`,5`,5,7-tetrahydroxy-6-methoxyflavanone

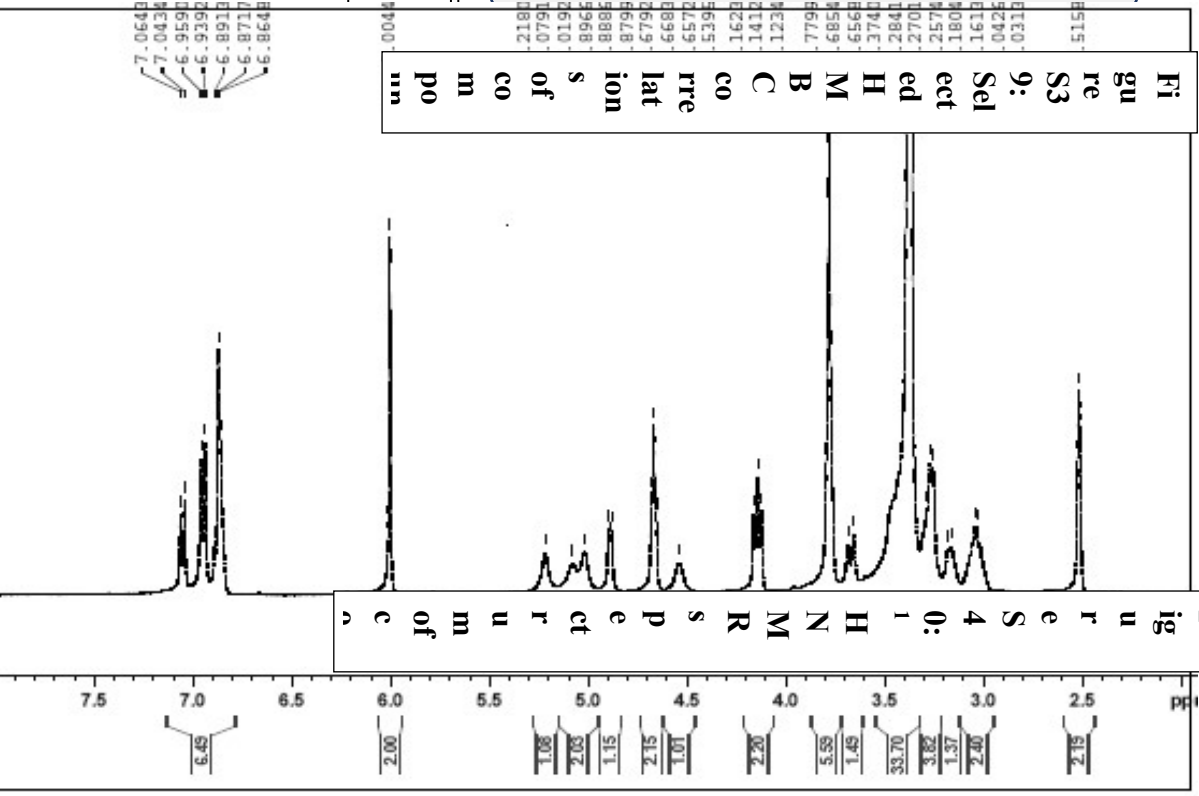
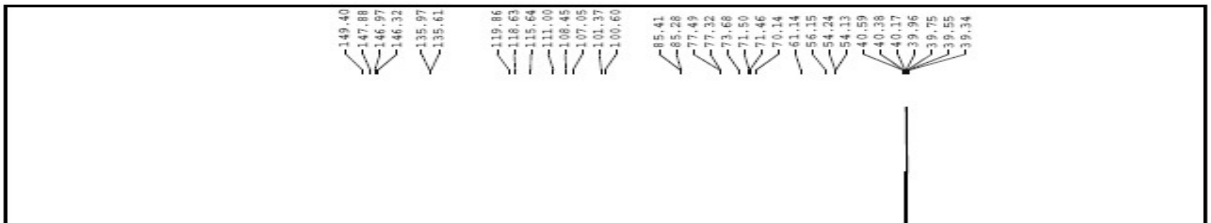


Figure S39: Selected 1H NMR spectrum of compound 13

Figure S40: Selected HMBC correlations of compound 13



Compound 15

Figure S41: APT spectrum of compound 14 in DMSO-d_6

Compound 15

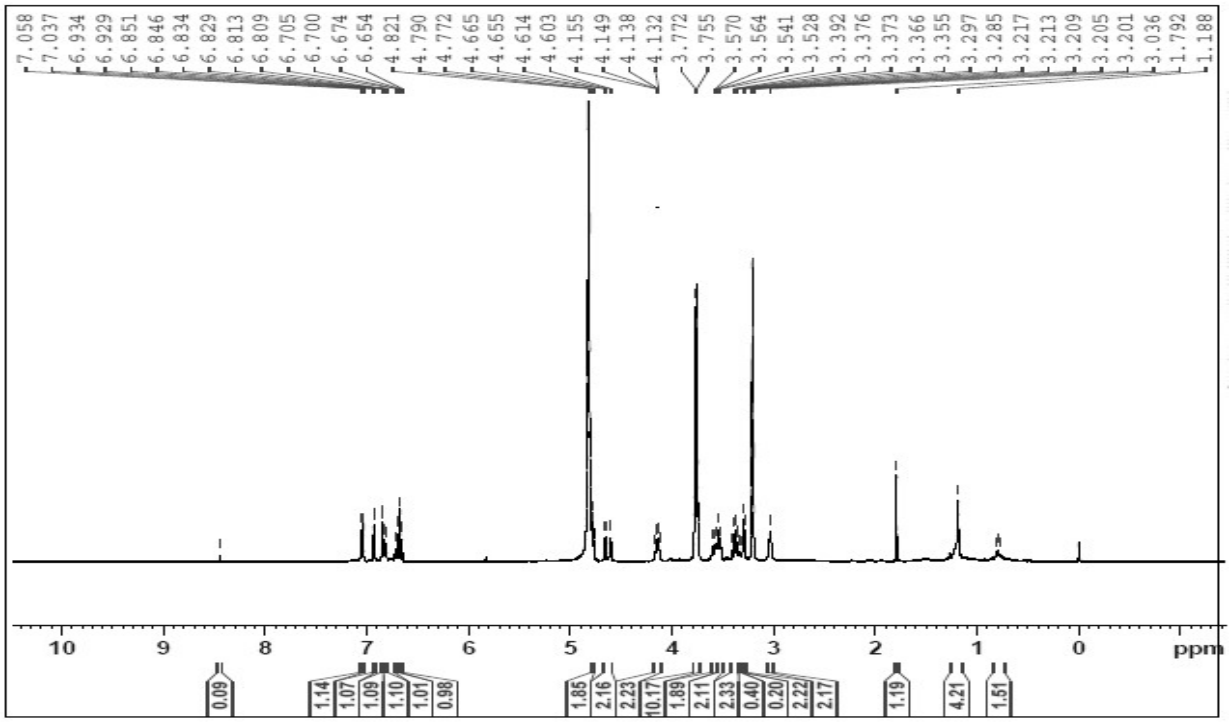


Figure S42: ^{1}H NMR spectrum of compound 14 in DMSO-d_6

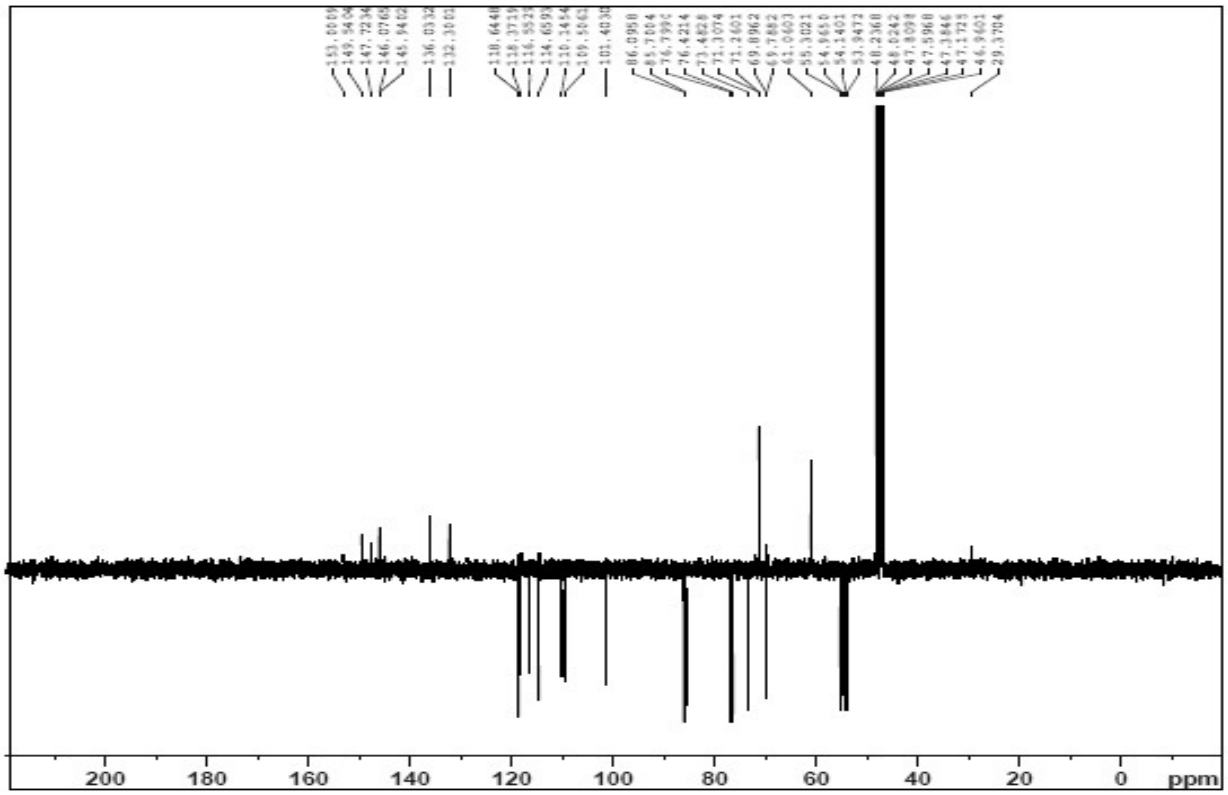


Figure S43: APT spectrum of compound 15 in CD_3OD

Compound 16

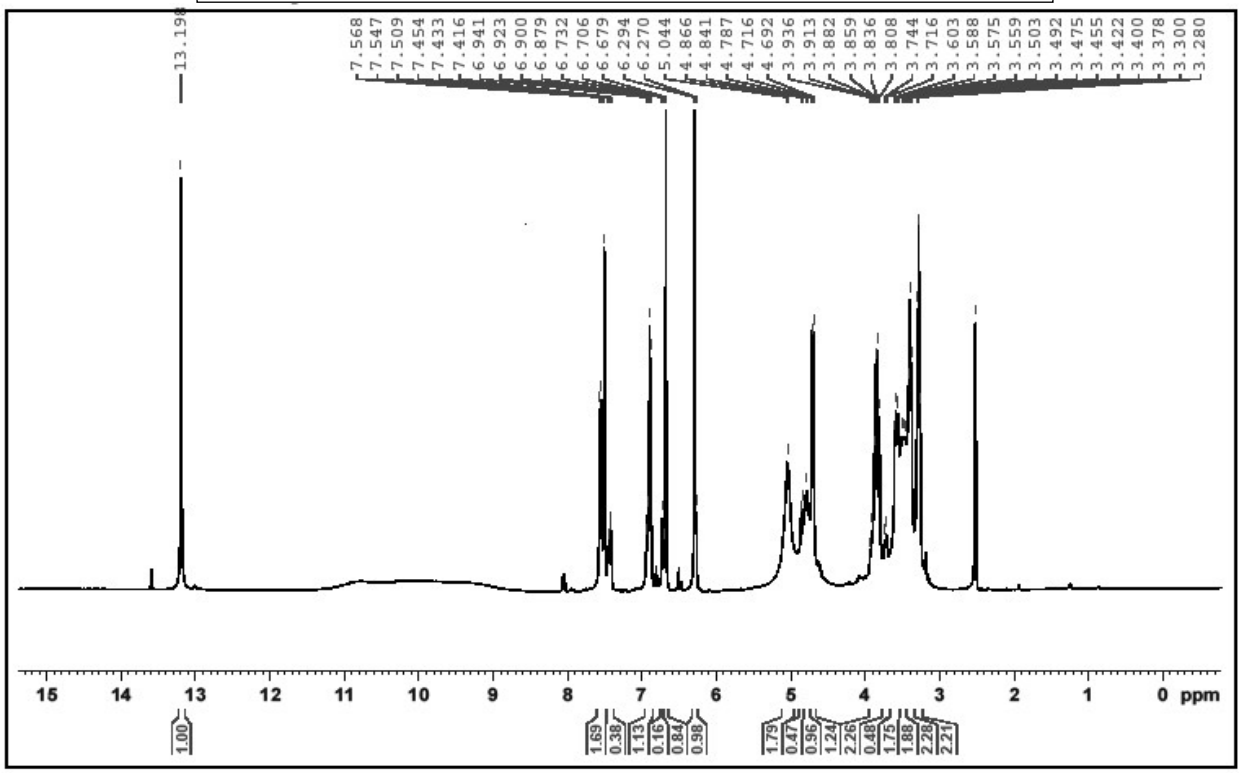
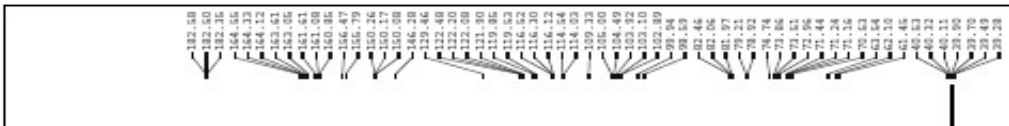


Figure S44: ^1H NMR spectrum of compound 16 in DMSO



Compound 17

Figure S45: APT spectrum of compound 16 in DMSO

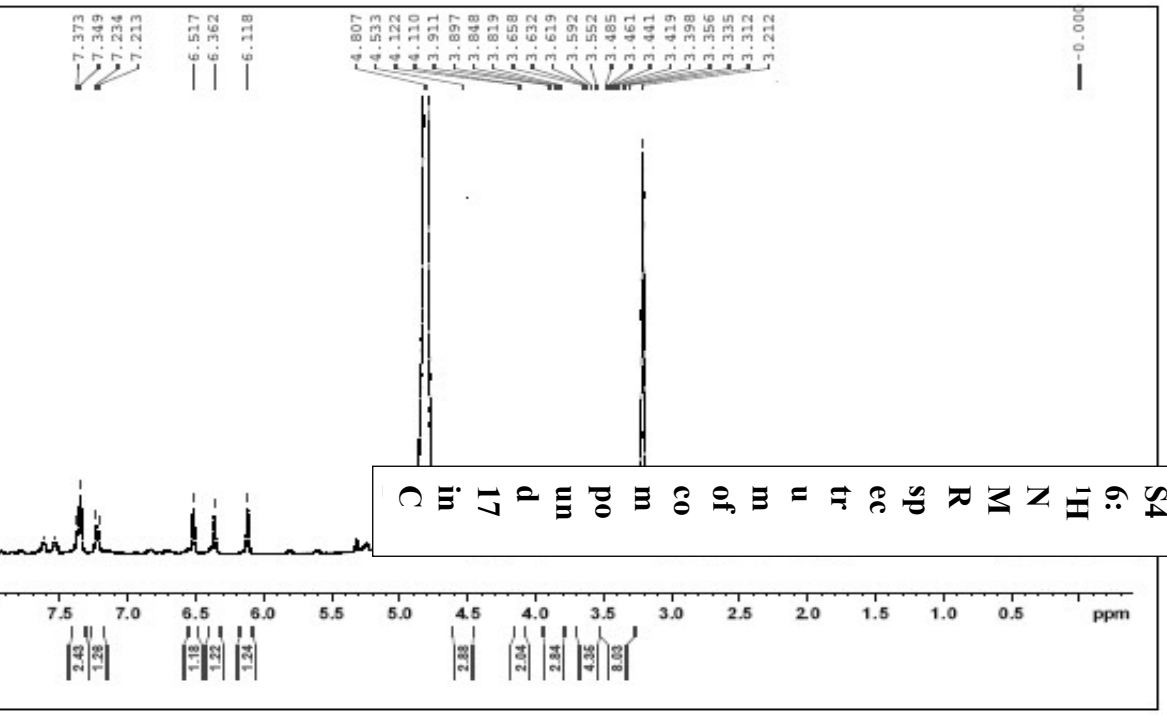


Figure S46: ^1H NMR spectrum of compound 17 in $CDCl_3$

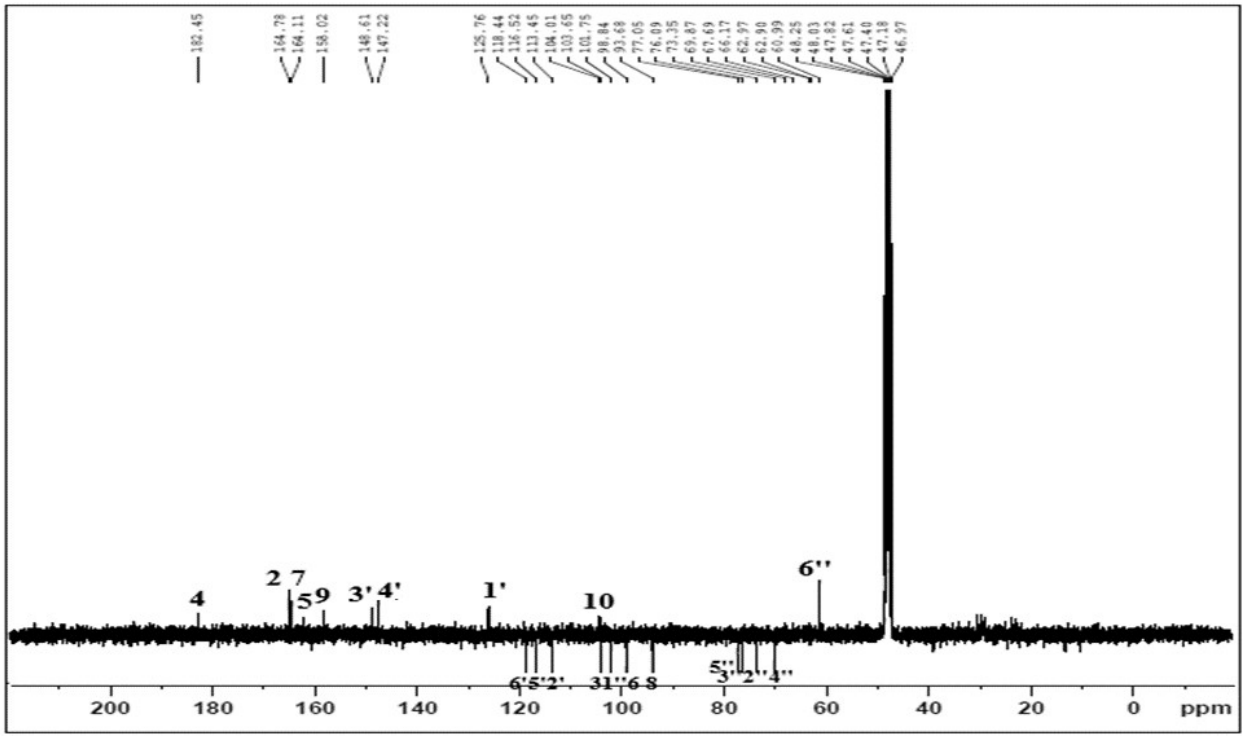


Figure S47: APT spectrum of compound 17 in D_3COOD

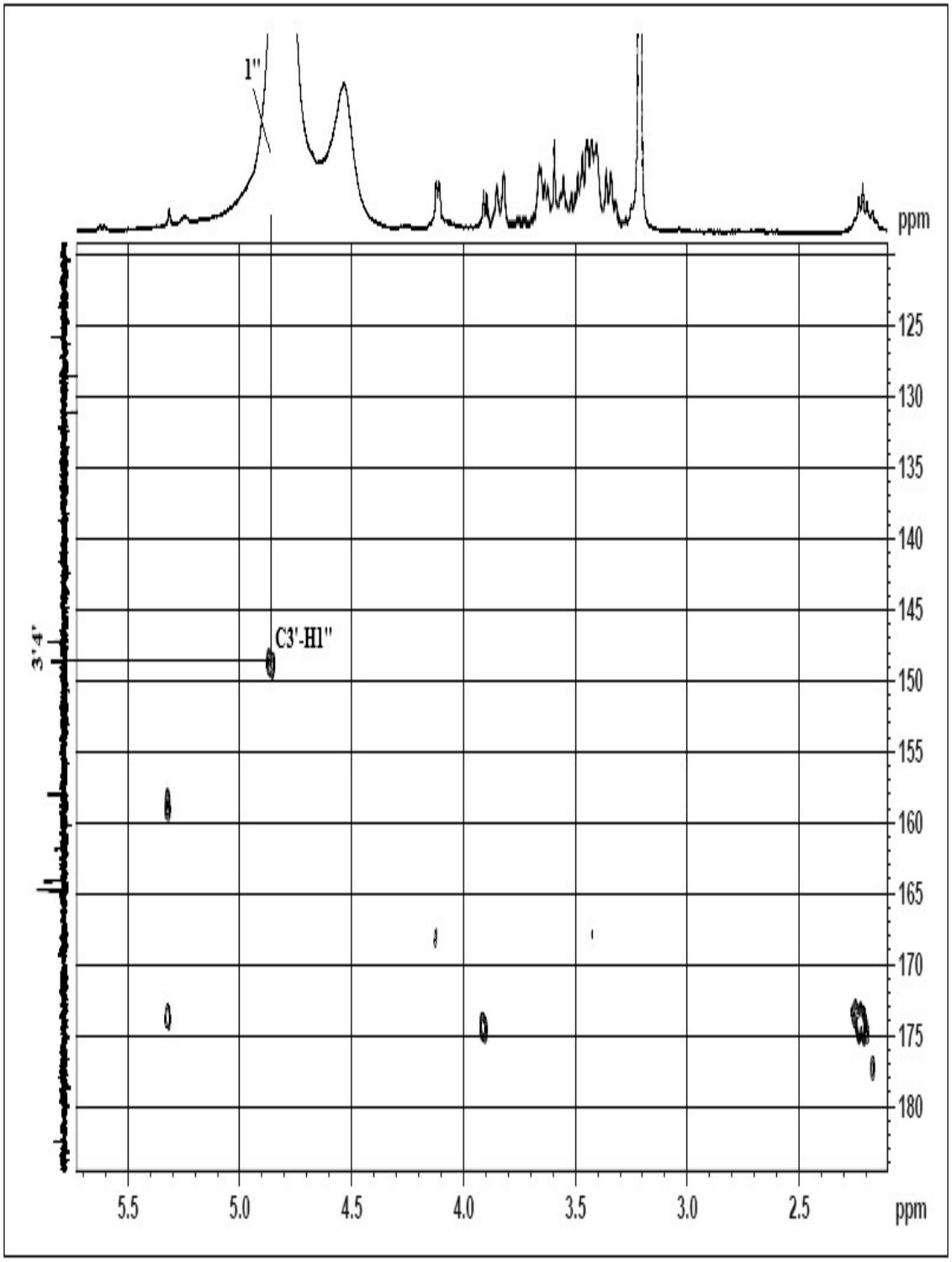


Figure S48: HMBC spectrum of compound

Compound 18

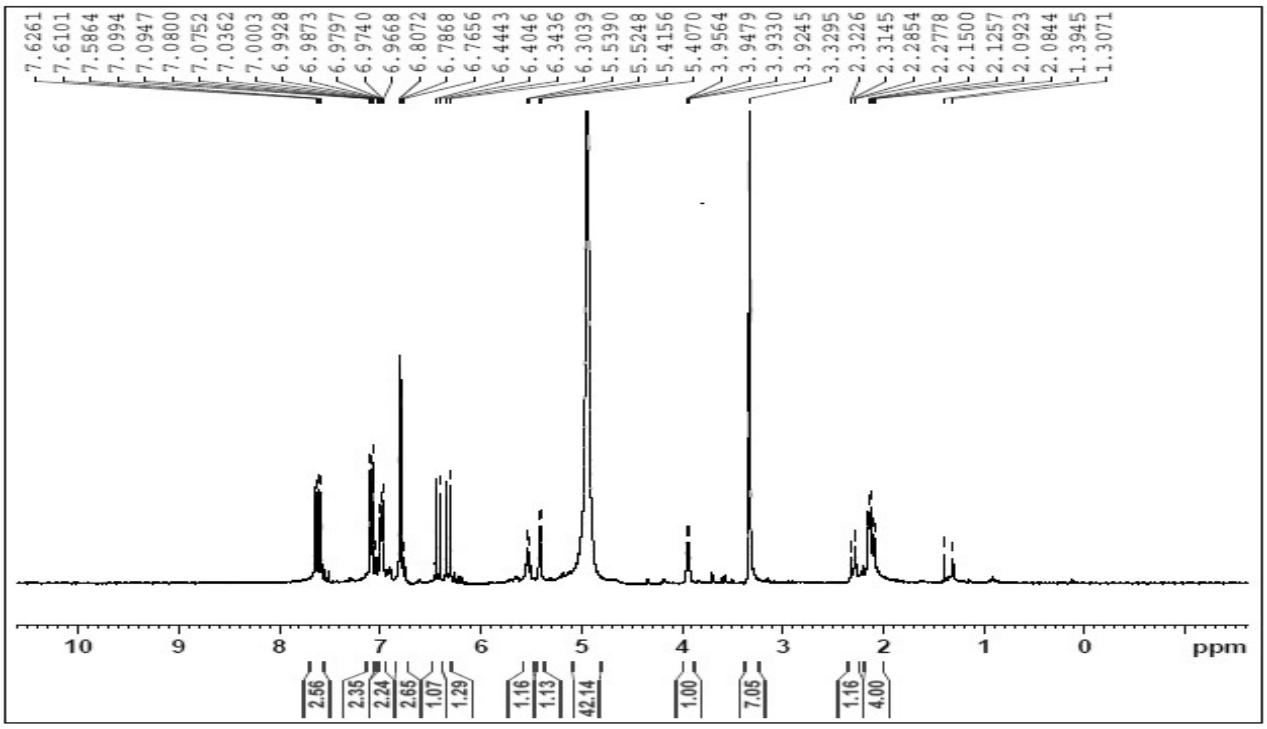


Fig. 9: ^{1}H NMR Spectrum of compound 18

Figure S50: APT spectrum of compound 18i in CD_3OD

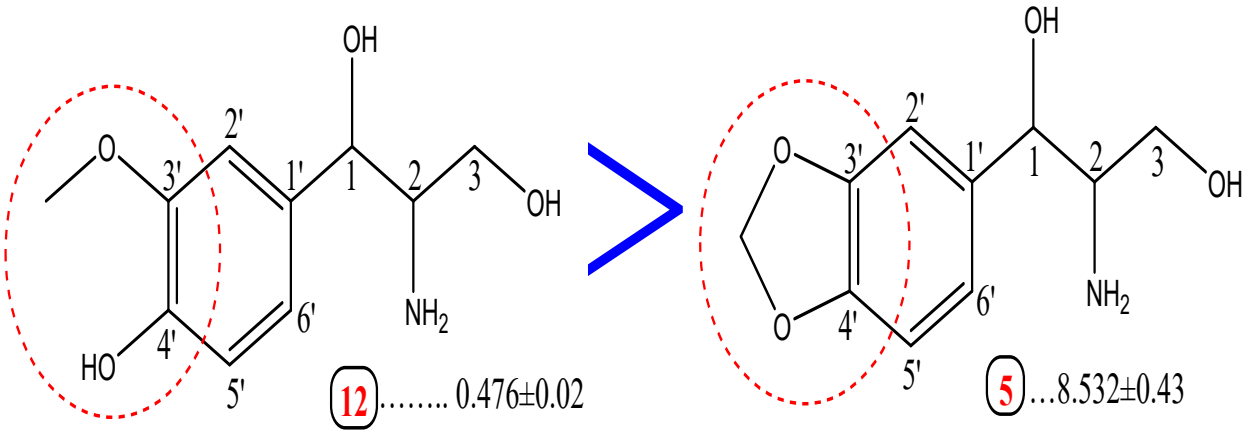


Figure S5 1: SAR of compounds 12 and 5

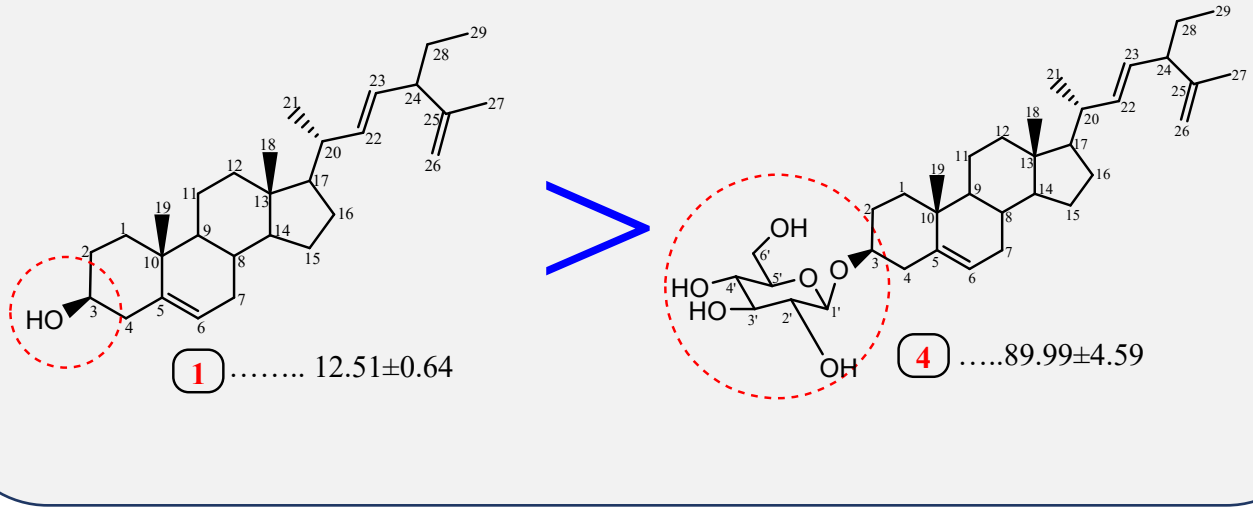


Figure S5 2: SAR of compounds 1 and 4

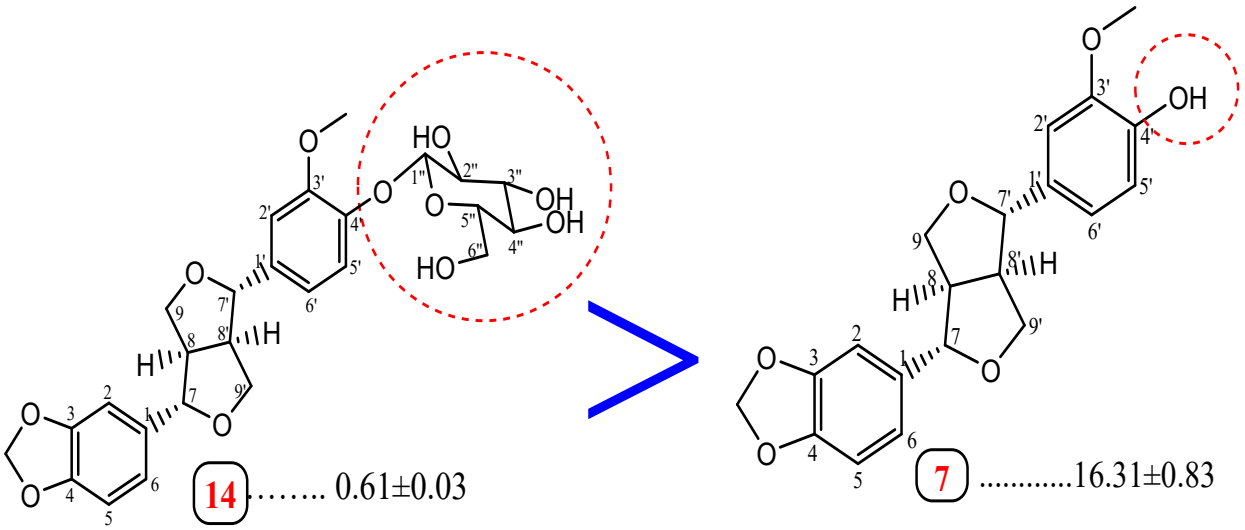


Figure 3: SAR of compounds 14 and 7

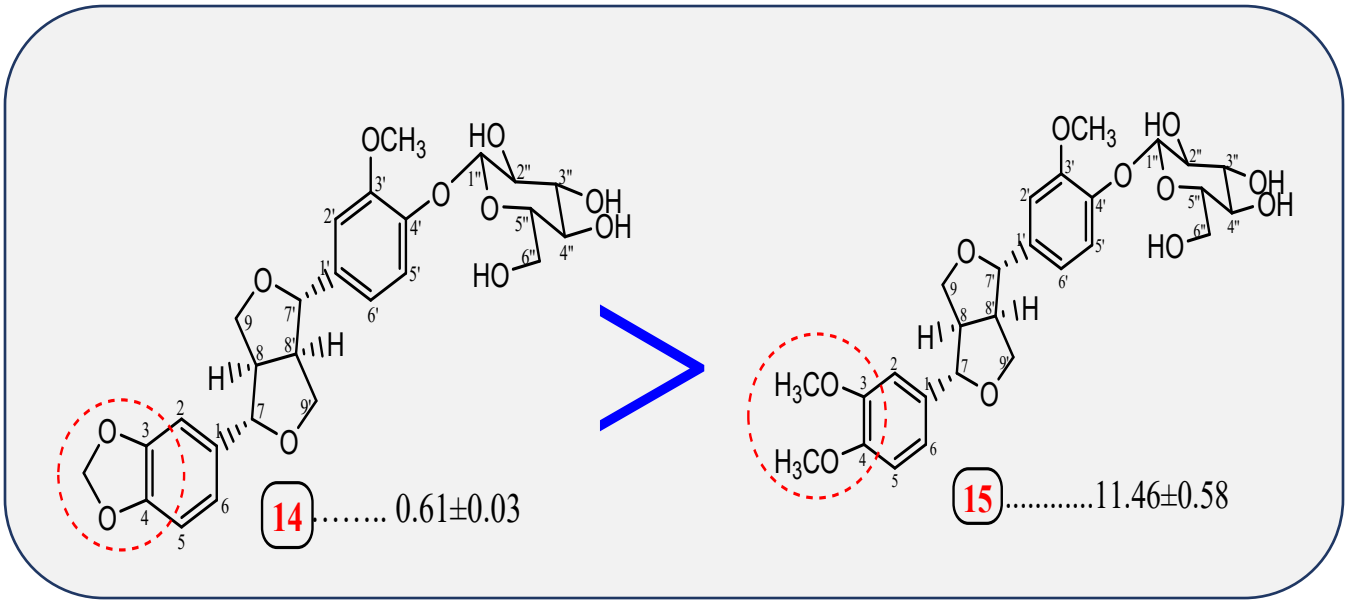


Figure 4: SAR of compounds 14 and 15

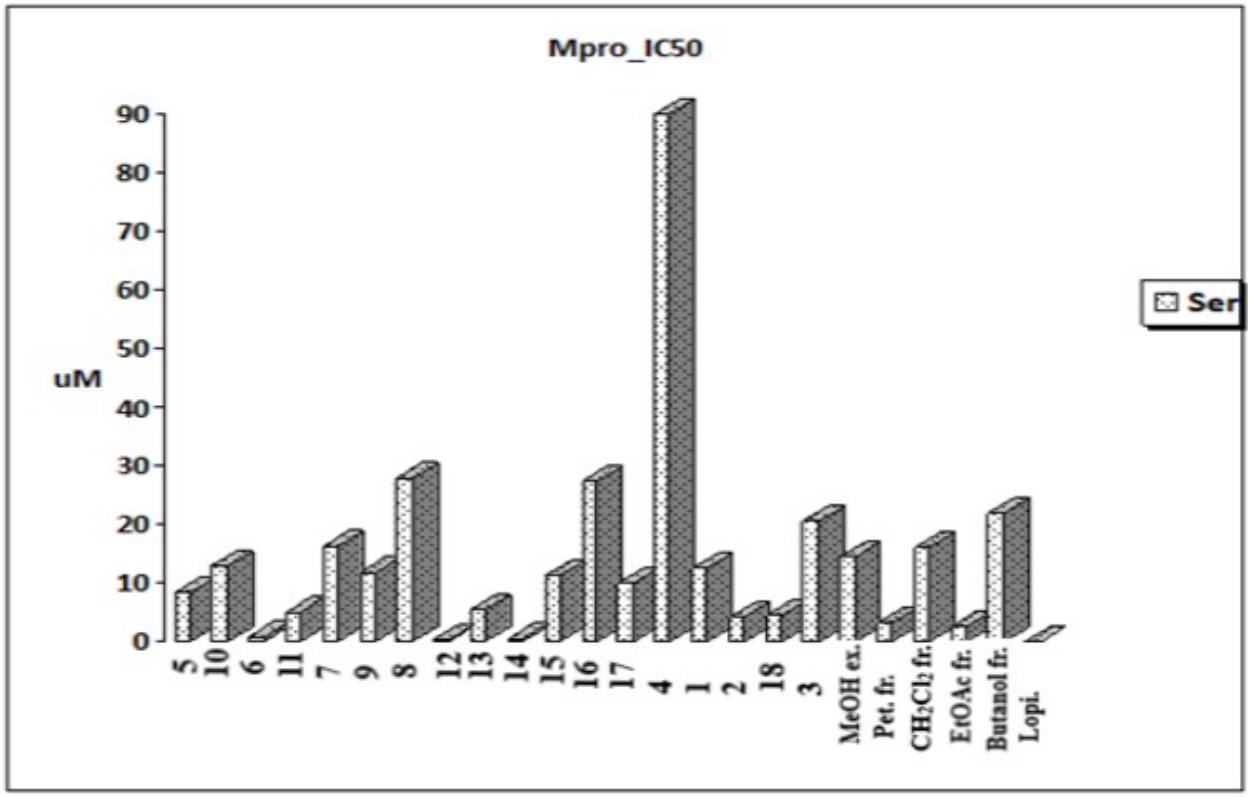


Figure (S55): TheSARS-COV-2M^{pro}inhibition (IC_{50} µg/ml) of methanolic extract, fractions and the isolated compounds against the standard lopinavir

Table S1: UV λ_{max} (nm) of compounds 8, 9, 13, 10, 11, 16 & 17 in methanol and in different shift reagents

Compounds	Band	MeOH	$NaOCH_3$	$AlCl_3$	$AlCl_3$/ HCl	NaOAc	NaOAc/ H_3BO_3
Flavanones							
8	**I**	236	245	235	235	235	238
	II	289	**327**	310	**309**	**329**	293
9	**I**	232	243	231	231	234	238
	II	291	**326**	307	**311**	**331**	294
13	**I**	234	238	236	234	237	237
	II	289	**291**	293	**308**	**320**	295
Flavones							
10	**I**	345	**410**	**357**	**351**	**409**	**328**
	II	276	270	**307**	**306**	**272**	276
11	**I**	323	**306**	**345**	**345**	**314**	**313**
	II	278	287	**310**	**310**	**294**	299
Glucosidated flavones							
16	**I**	346	**406**	**424**	**382**	**392**	**380**
	II	271	277	**346**	**361**	**277**	268
17	**I**	331	**387**	**363**	**364**	**397**	**336**
	II	271	279	**342**	**339**	**310**	290

Table S2: ^{1}H-NMR and APT values of compound 8, 9 and 13in CD_3OD*

C/H Position	Compound HM 8 [5,7,4`- trihydroxy-8,3`- dimethoxy flavanone]		Compound HM 9 [5,7,4`-trihydroxy-6-methoxy flavanone]		Compound HM 13 [3`,5`,5,7-tetrahydroxy-6-methoxyflavanone]		
	^{1}H-NMR	**APT**	**^{1}H-NMR**	**APT**	**^{1}H-NMR**	**APT**	
2	5.18 (1H, dd, J_1= 2.8, J_2= 12.6)	79.2	5.22 (1H, dd, J_1= 2.8, J_2= 13)	79.2	5.16 (1H, dd, J= 2.8, 12.8)	79.2	CH
3	***3α***= 2.56 (1H, dd, J_1= 2.8, J_2= 17.2) ***3β***= 2.97 (1H, dd, J_1= 12.8, J_2= 17.2)	42.7	***3α***= 2.60 (1H, dd, J_1= 2.8, J_2= 17) ***3β***= 3.01 (1H, dd, J= 12.8, 17.2)	42.7	***3α***= 2.60 (1H, dd, J= 2.8, 17.2) ***3β***= 2.96 (1H, dd, J= 12.8, 17.2)	42.7	CH_2
4	----------	195.7	----------	197.2	----------	197.2	Q
5	----------	159.0	----------	155.2	----------	155.2	Q
6	5.79 (1H, s)	96.3	----------	129.0	----------	129.0	Q
7	----------	164.6	----------	159.4	----------	159.5	Q
8	----------	130.7	5.87 (1H, s)	94.8	5.87 (1H, s)	94.8	CH
9	----------	155.0	---------	157.6	---------	158.8	Q
10	----------	100.7	---------	102.1	---------	102.1	Q
OCH_3	3.66 (3H, s)	59.4	3.67 (3H, s)	59.6	3.68 (3H, s)	59.6	CH_3
OCH_3	3.78 (3H, s)	55.1	---------	-----	---------	------	-----
1`	----------	130.1		129.7		130.3	Q
2`	6.97 (1H, d, J= 1.6)	109.8	7.21 (2H, d, J= 8.4)	127.6	6.68 (2H, s)	117.9	CH
3`	----------	147.7	6.71 (2H, d, J= 8.8)	114.9	----------	145.1	Q
4`	----------	146.6	----------	158.8	6.68 (2H, s)	114.8	CH
5`	6.71 (1H, d, J= 8)	114.7	6.71 (2H, d, J= 8.8)	114.9	----------	145.5	Q
6`	6.81 (1H, dd, J_1= 8.2, J_2= 2)	119.0	7.21 (2H, d, J= 8.4)	127.6	6.81 (1H, s)	113.3	CH

* δ values of compounds **8, 9& 13** are expressed in ppm and coupling constants (*J*) in Hz. ^{1}H-NMR and APT were measured in CD_3ODat 400 and 100 MHz respectively.

Table S3: HMBC correlations of compound 13 deduced from HMBC (Figures S34, S35&S39)

Proton	Proton (Values in ppm)	Correlated Carbon (s) (Values in ppm)
OCH_3	3.68	129.0 (C-6)
H-8	5.87	129.0 (C-6), 159.5 (C-7), 158.8 (C-9), 102.1 (C-10)
H- 2`	6.68	130.3 (C-1`), 114.8 (C-4`), 145.1 (C-3`), 145.5 (C-5`), 113.3 (C-6`), 79.2 (C-2)
H-4`	6.68	145.1 (C-3`), 145.5 (C-5`)
H-6`	6.81	117.9 (C-2`), 145.1 (C-3`), 145.5 (C-5`), 79.2 (C-2)

Table S4: Results of the SARS-COV-2M ^{pro} inhibitory activity of the methanolic extract of the leaves of *H. bracteatum* and its fractions compared with the standard lopinavir

Extract& Fractions	*In vitro*SARS-COV-2M ^{pro} IC_{50} IC_{50} (µg/ml)
Lopinavir (Standard)	**0.141**±0.01
Methanolic extract	14.47±0.74
Pet. ether fraction	3.466±0.18
Methylene fraction	16.05±0.82
Ethyl fraction	2.589±0.13
Butanol fraction	21.9±1.12

Table S5: The SARS-COV-2M ^{pro}inhibition (IC_{50} µM), docking scores[a] and type of binding interactions of the isolated compounds (1-18) and the standard compound (lopinavir)

Comp-Ound (code)	Compounds (name)	*In vitro*SARS-COV-2 M^{pro} IC_{50} uM IC_{50} (µmole)	Binding energy (kcal/mol) (docking score)	Type of binding interactions
Stand-ard	**Lopinavir**	**0.225±0.01**	-9.61	H-bond with Glu 166 & Gln 189
1	24-*β*-ethyl-cholesta-5(6),22(23),25(26)-triene-3-ol	12.51±0.64	-9.99	H-bond with Glu 166 & Phe 140
2	*α*-amyrin	4.185±0.21	-10.29	H-bond with Glu 166
3	Linoleic acid	20.67±1.05	-10. 39	H-bond with Thr 190 & Arg 188
4	24-*β*-ethyl-cholesta-5(6),22(23),25(26)-triene-3-O-*β*-D-glucoside	89.99±4.59	-11.92	Three H-bonds with Gln 189
5	1,3-propanediol-2-amino-1-(3`,4`-methylene dioxyphenyl)	8.532±0.43	-8.97	-Two H-bonds with Gln 189 - H-bond with Glu 166 & Gln 192
6	(-)-(7*R*,8*R*,8`*R*)-acuminatolide	0.917±0.05	-9.39	H-bond with Glu 166 & Ser 144
7	(+)-piperitol	16.31±0.83	-12.34	H-bond with Glu 166, Ser 46, Thr 25 & Thr 45
8	5,7,4`-trihydroxy-8,3`-	27.86±1.42	-12.69	H-bond with Glu 166, Gly

	dimethoxy flavanone			143 & Leu 141
9	5,7,4`-trihydroxy-6-methoxy flavanone	11.83±0.6	-11.49	-H-bond with Glu 166, Gly 143& Ser 144 -Two H-bonds with His 163
10	4`,5-dihydroxy-3`,7,8-trimethoxyflavone	12.83±0.65	-13.45	-H-bond with Glu 166 & Leu 141 - Two H-bonds with Gly 143
11	5,7-dihydroxy- 3`,4`,5`,8-tetramethoxy flavone	5.069±0.26	-12.48	H-bond with Glu 166, Cys 145, Gly 143 & Ser 144
12	1,3-propanediol-2-amino-1-(4`-hydroxy-3`-methoxyphenyl)	0.476±0.02	-10.79	-Two H-bonds with Glu 166 - H-bond with Gln 189, Thr 190 & Arg 188
13	3`,5`,5,7-tetrahydroxy-6-methoxyflavanone	5.565±0.28	-12.81	H-bond with Glu 166, His 163 & Leu 167

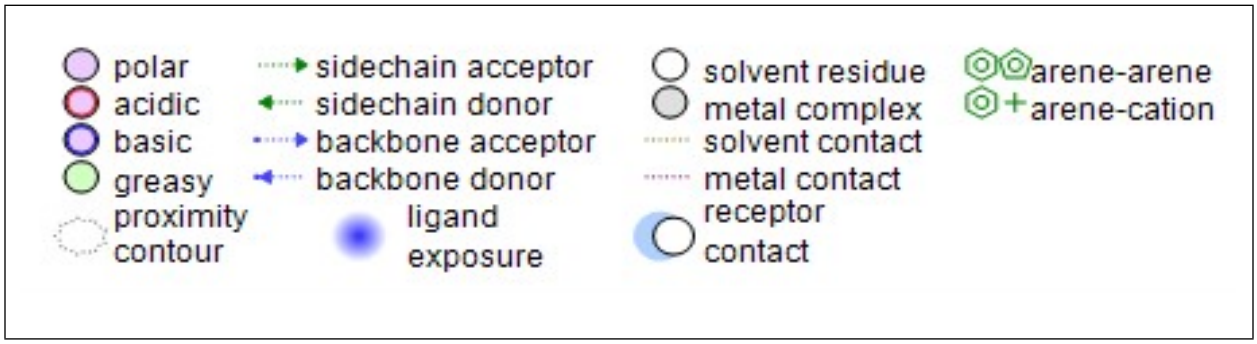
14	Simplexoside (piperitol-O-*β*-D-glucoside)	0.61±0.03	-12.96	-H-bond with Gln 189, Glu 166, Thr 26, Thr 24 & Ser 46 -Two H-bonds with Gly 143
15	Pinoresinol monomethyl ether-*β*-D-glucoside	11.46±0.58	-11.69	-Three H- bonds with Glu 166 -Two H-bonds with Thr 190 - H-bond with Gln 192, Thr 26 & Arg 188
16	Orientin	27.5±1.4	-14.34	-Two H-bonds with Glu 166 - H-bond with His 163 & Arg 188
17	Luteolin 3`-O-*β*-D-glucoside	10.12±0.52	-15.61	-Two H-bonds with Glu 166 - H-bond with His 163, Phe 140, Thr 24 & Thr 25
18	3,5-dicaffeoylquinic acid	4.74±0.24	-16.24	-H-bond with Glu 166, Gln 189, Leu 141 & Thr 25

[a] Docking was performed using MOE 2009.10 towards the active site of M ^{pro} (code: 6LU7)

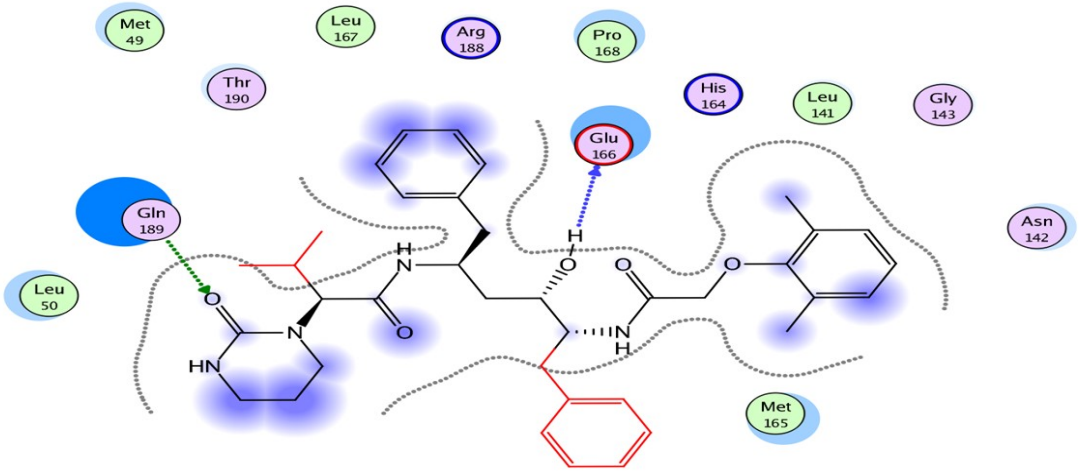
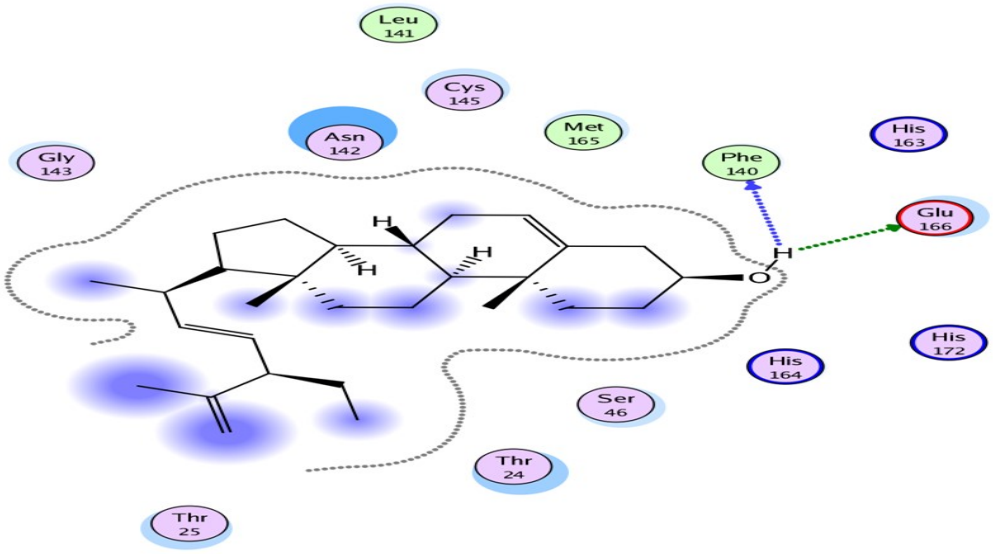
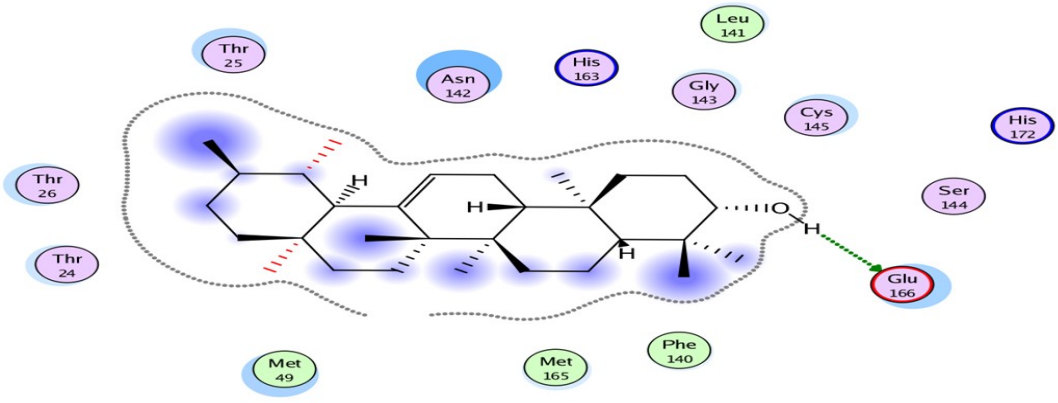
[b] All data are presented as mean value ± SD for three independent experiments.

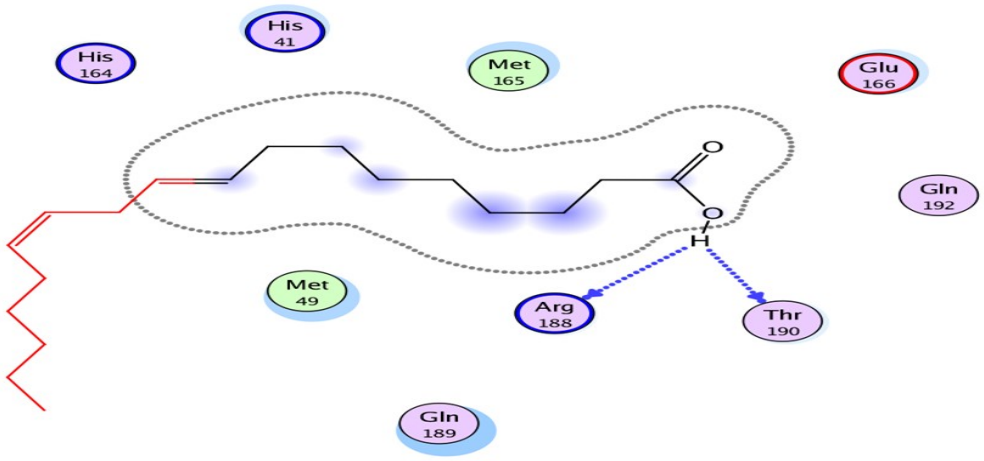
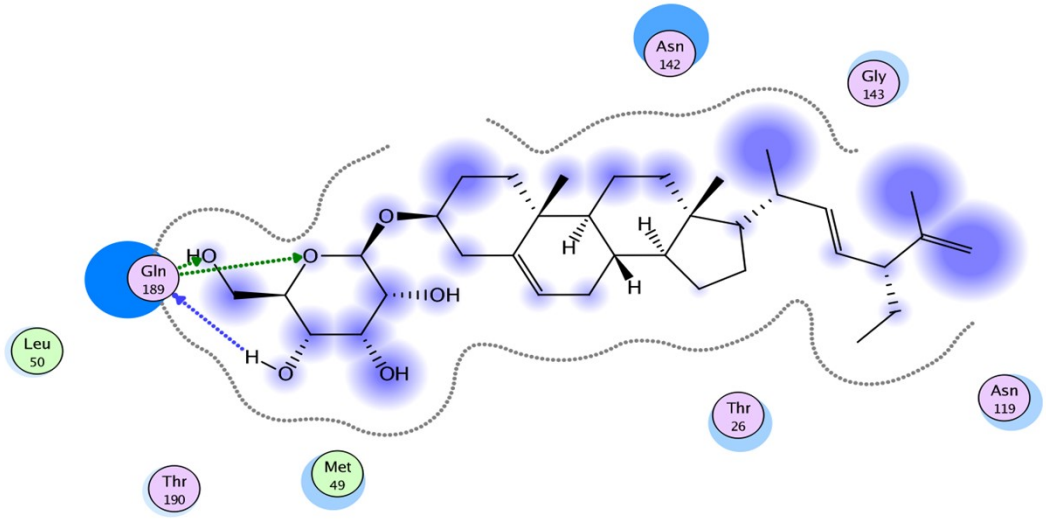
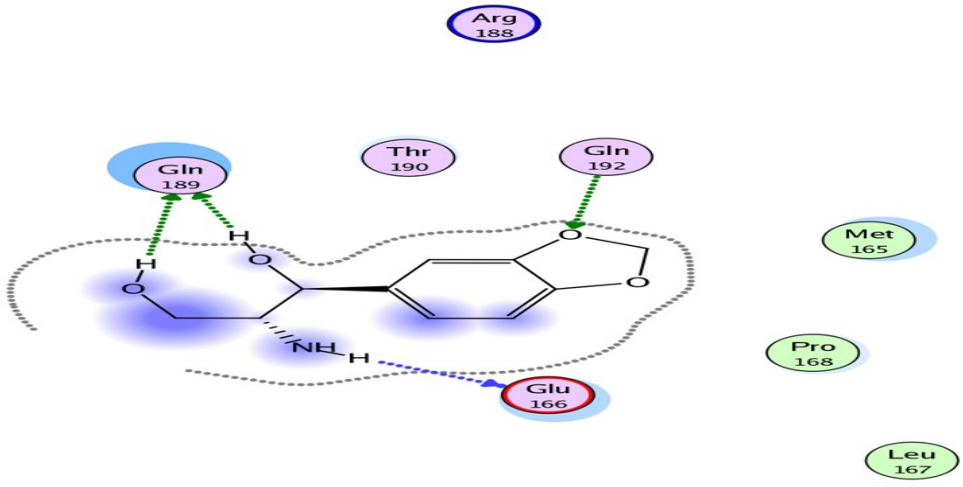
[c] Lopinavir was used as a positive control.

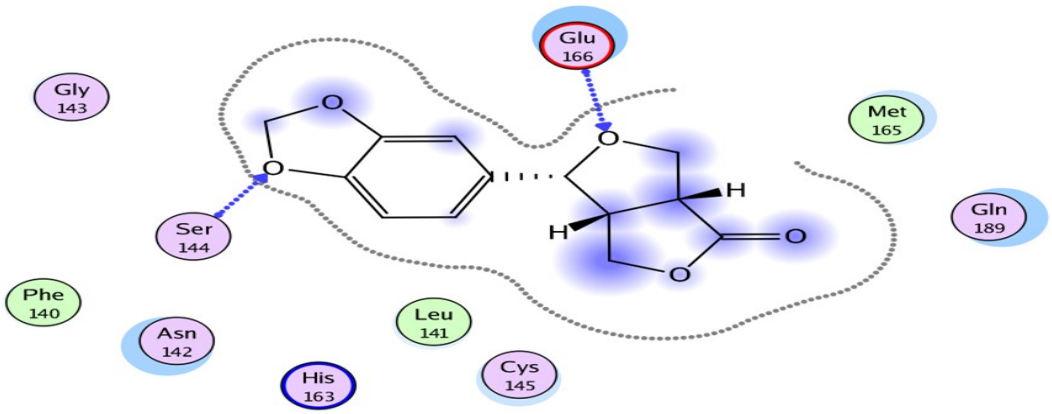
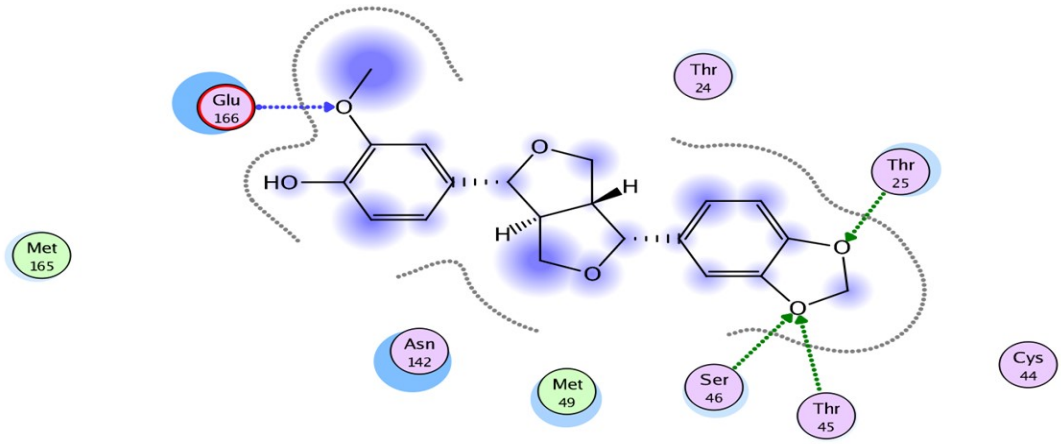
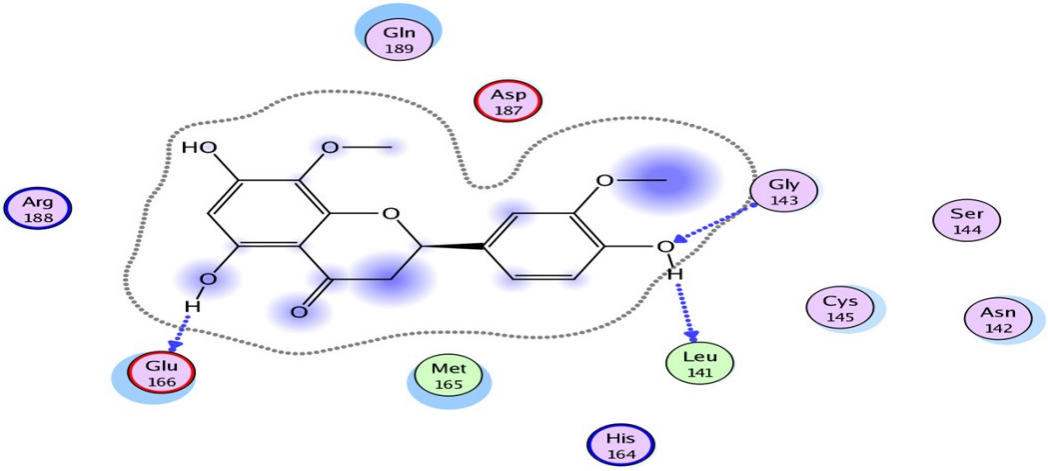
Table S6: 2D binding mode and residues involved in the recognition of the standard lopinavir and the isolated compounds docked and minimized in the SARS-COV-2M^{pro}binding pocket

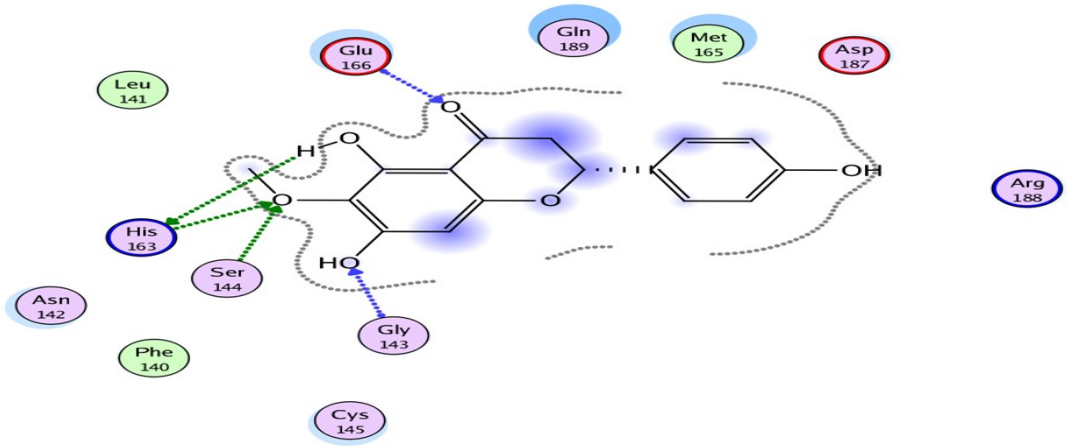
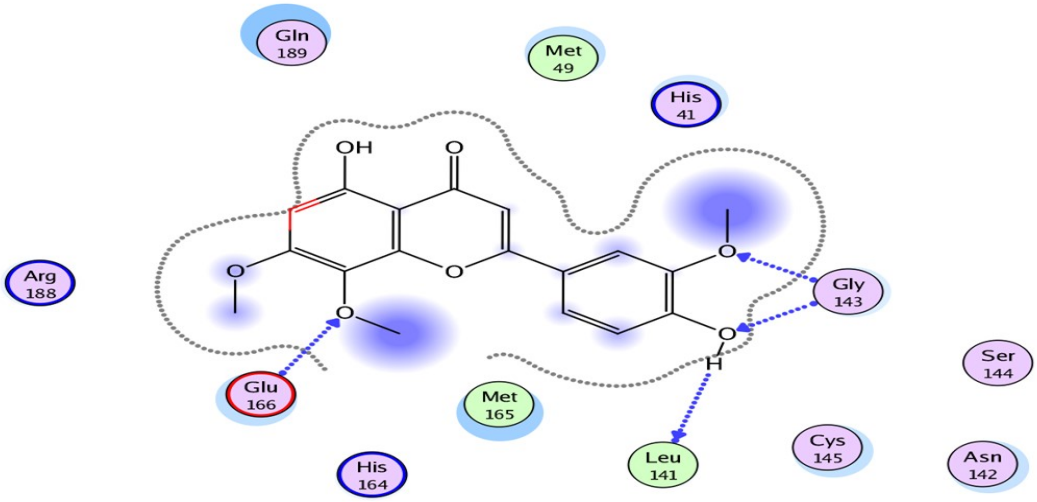
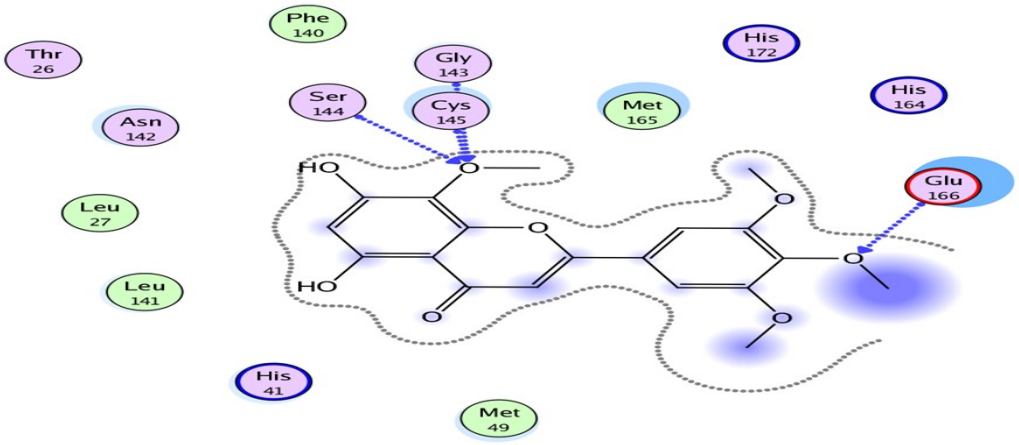


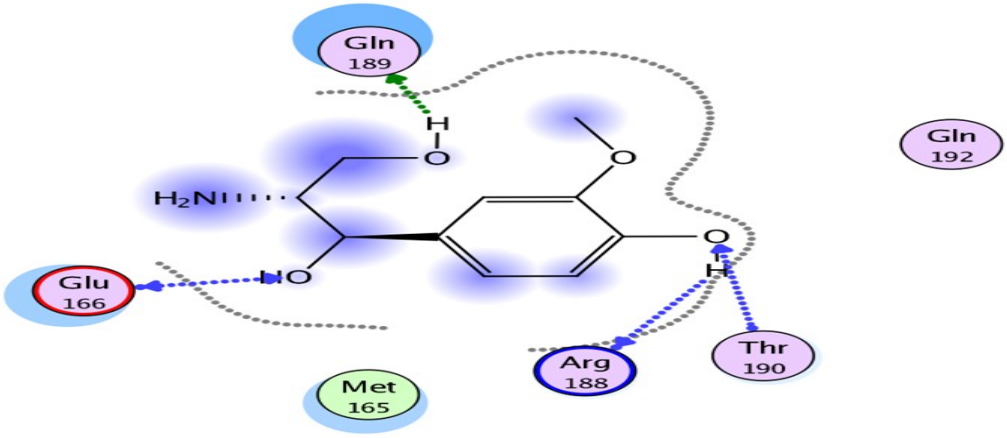
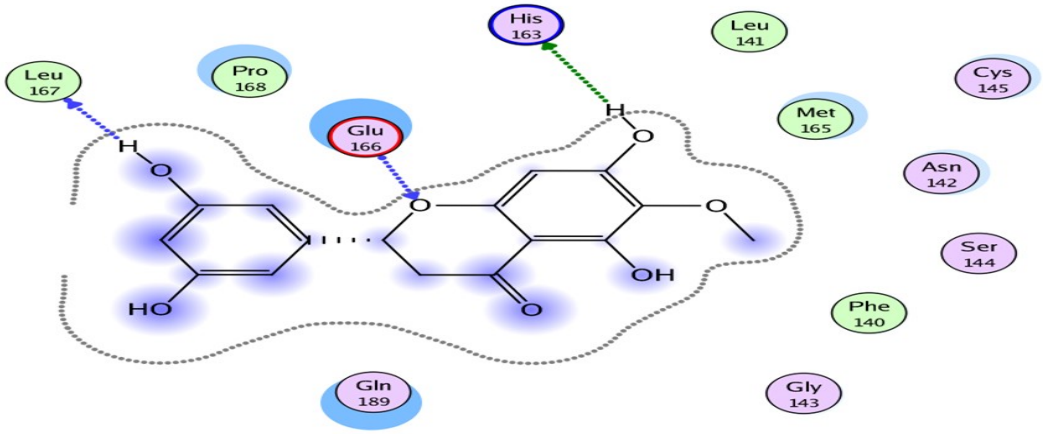
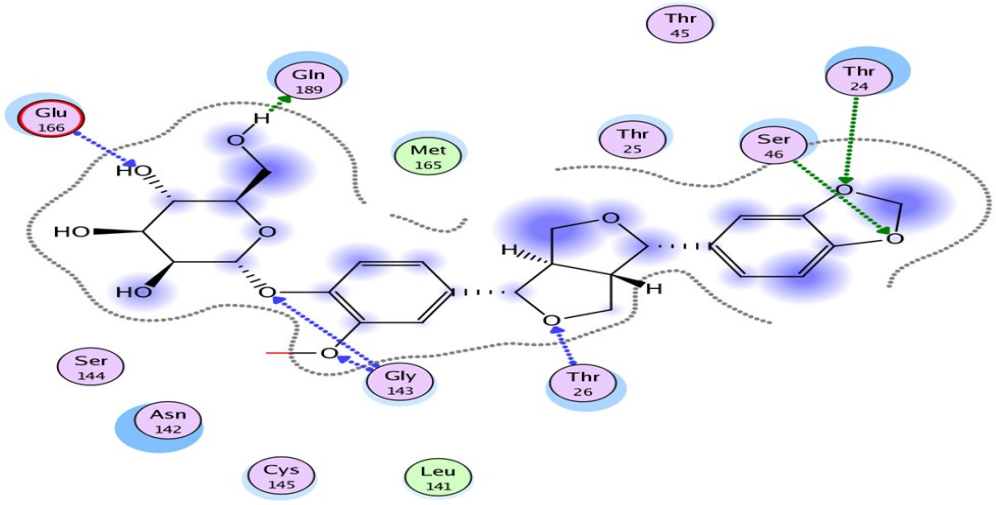
Code	2D binding mode and residues

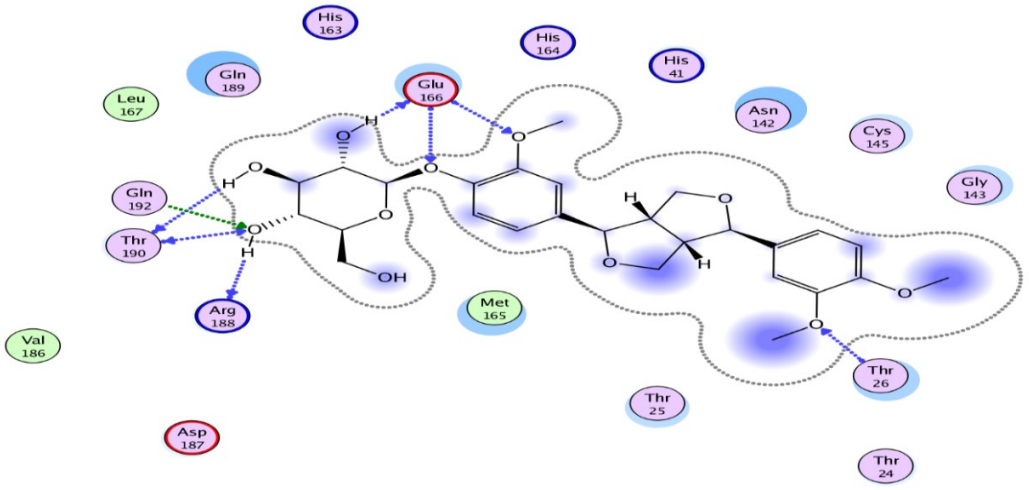
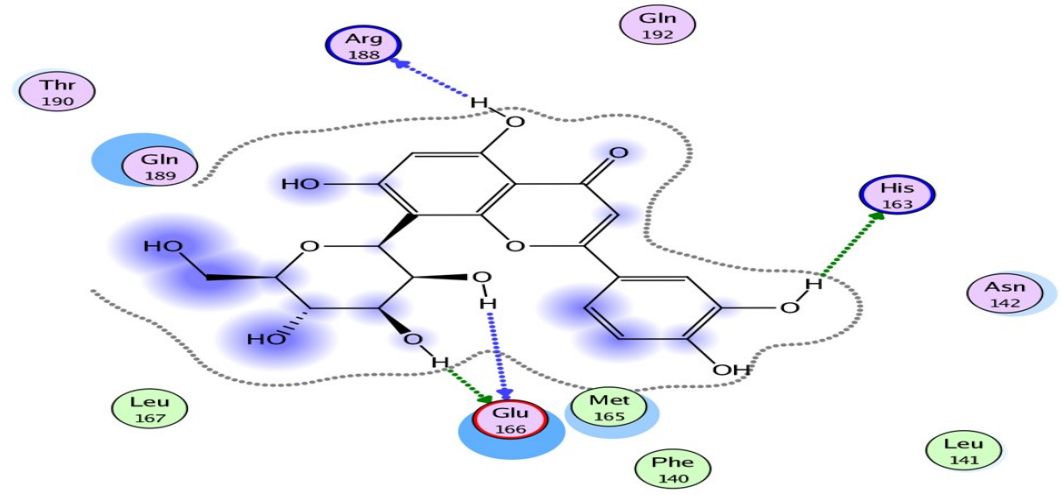
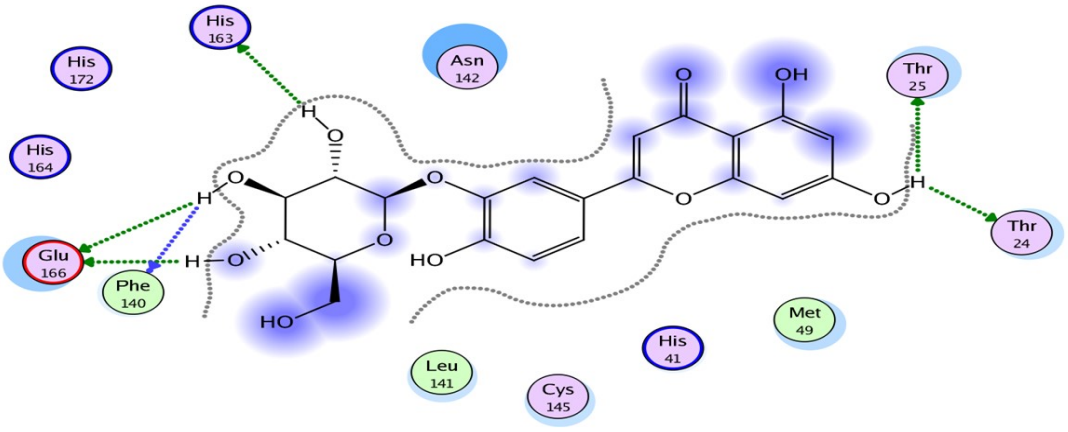
Stan-dard	**Lopinavir** 
1	**24-*β*-ethyl-cholesta-5(6),22(23),25(26)-triene-3-ol** 
2	***α*-amyrin** 

3	**Linoleic acid** 
4	**24-*β*-ethyl-cholesta-5(6),22(23),25(26)-triene-3-O-*β*-D-glucoside** 
5	**1,3-propanediol-2-amino-1-(3`,4`-methylenedioxyphenyl)** 

6	**(-)-(7*R*,8*R*,8`*R*)-acuminatolide** 
7	**(+)-piperitol** 
8	**5,7,4`-trihydroxy- 8,3`-dimethoxy flavanone** 

9	**5,7,4`-trihydroxy-6-methoxy flavanone** 
10	**4`,5-dihydroxy-3`,7,8-trimethoxyflavone** 
11	**5,7-dihydroxy- 3`,4`,5`,8-tetramethoxy flavone** 

12	**1,3-propanediol-2-amino-1-(4`-hydroxy-3`-methoxyphenyl)** 
13	**3`,5`,5,7-tetrahydroxy-6-methoxyflavanone** 
14	**Simplexoside (piperitol-O-*β*-D-glucoside)** 

15	**Pinoresinol monomethyl ether-*β*-D-glucoside** 
16	**Orientin** 
17	**Luteolin-3`-O-*β*-D-glucoside** 

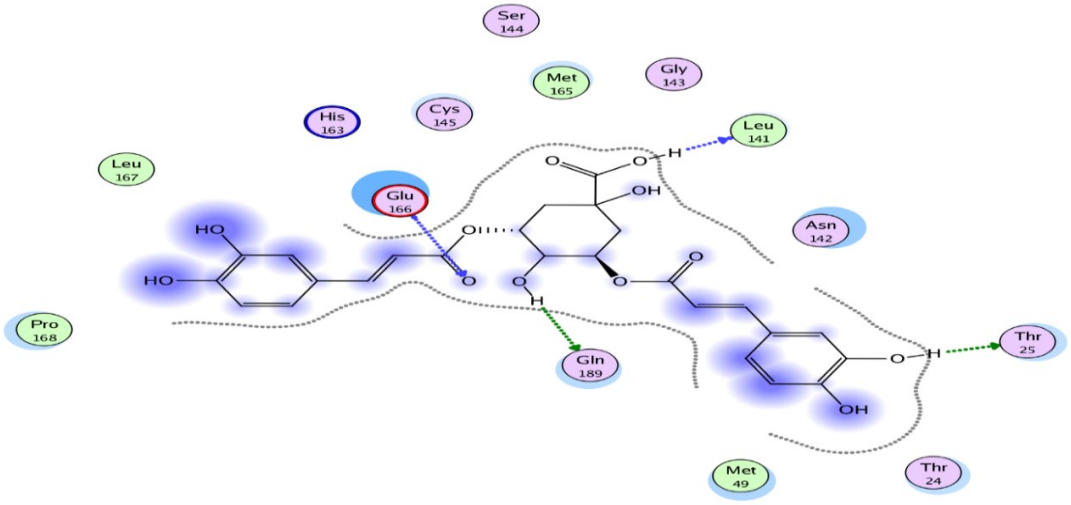
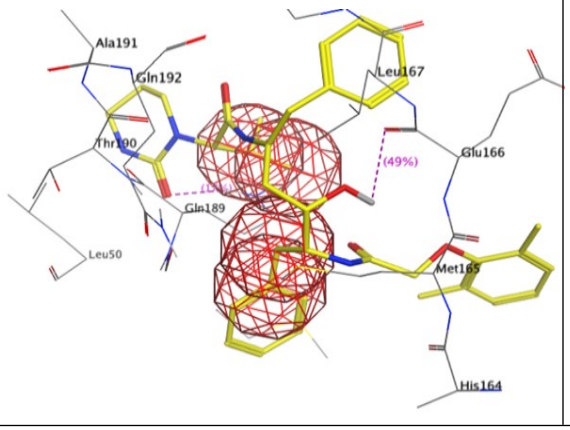
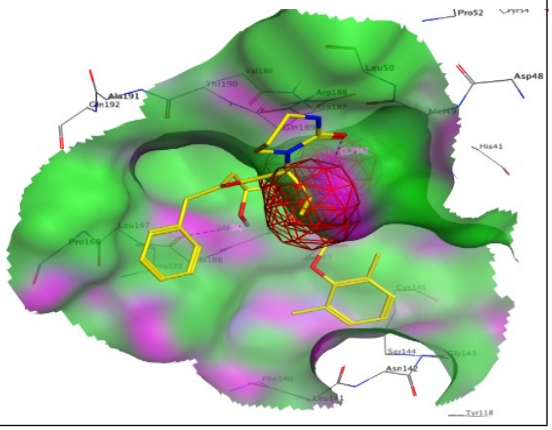
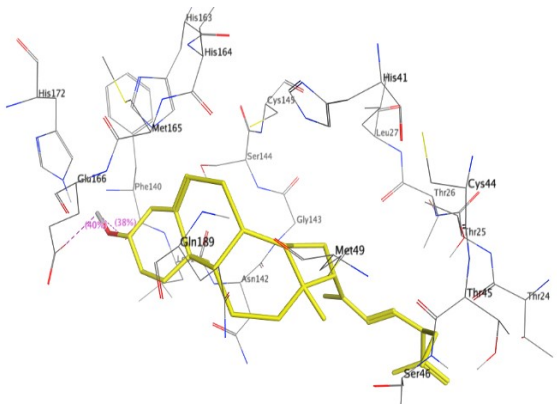
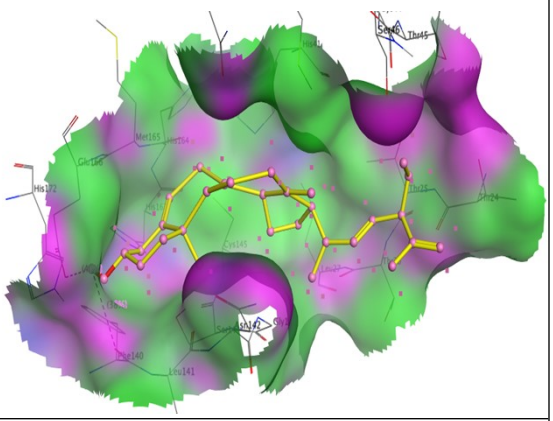
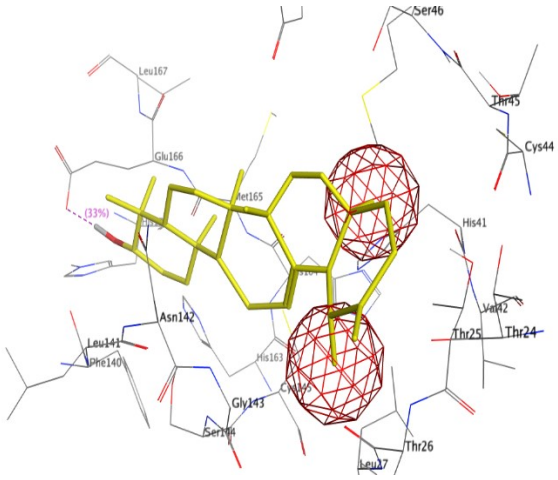
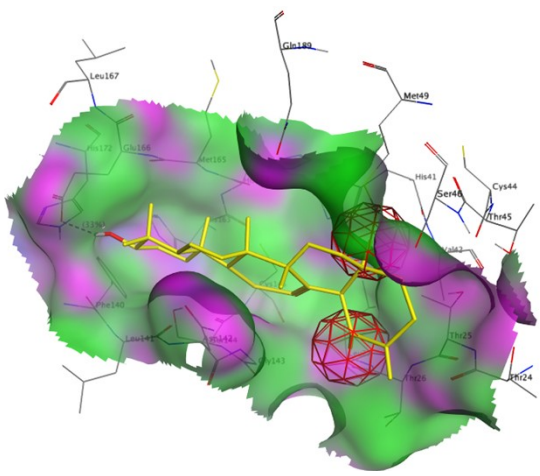
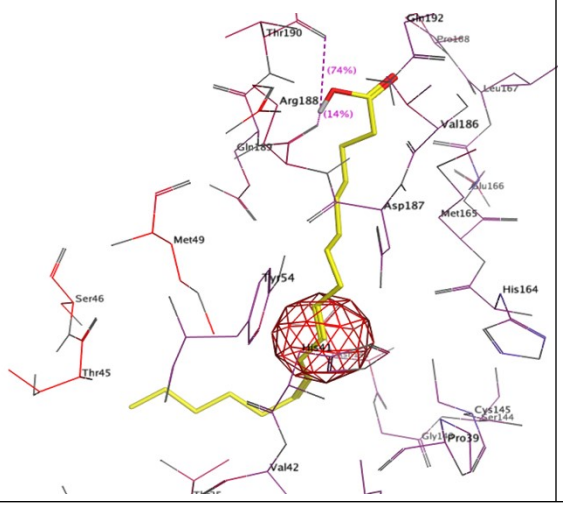
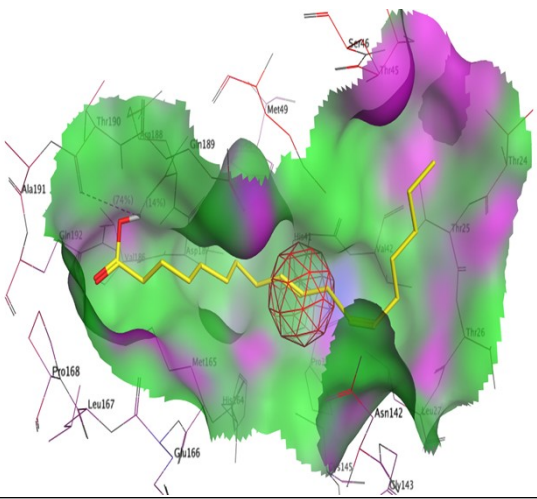
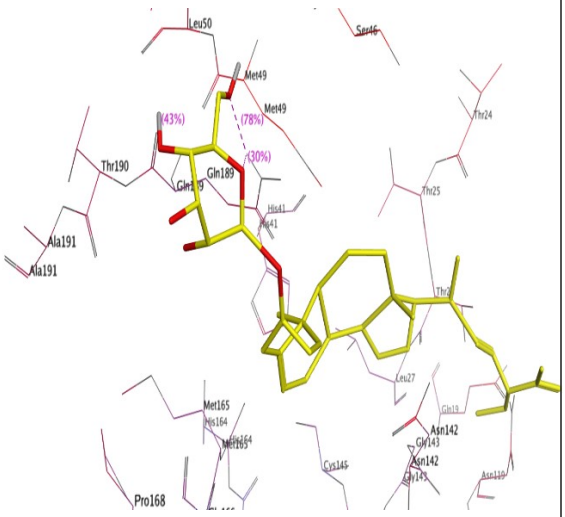
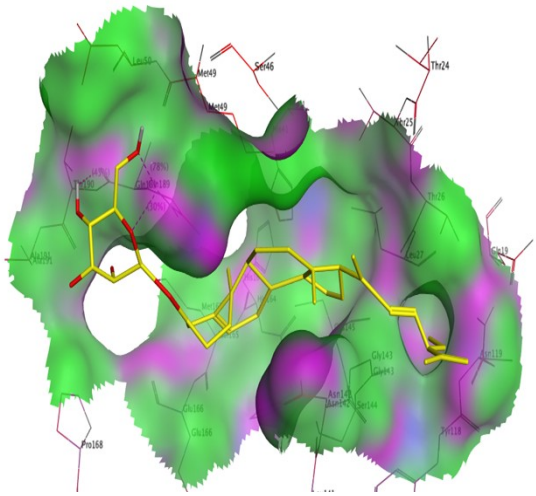
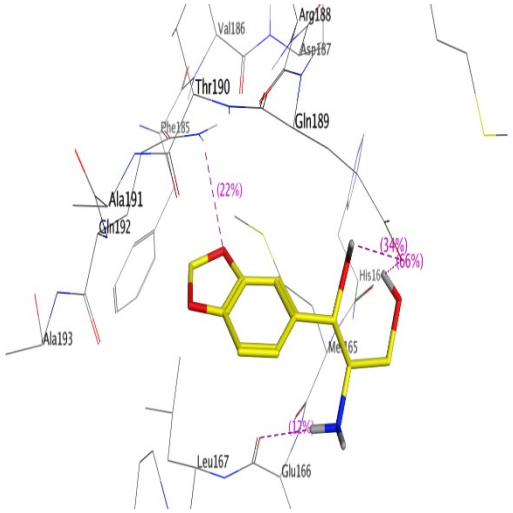
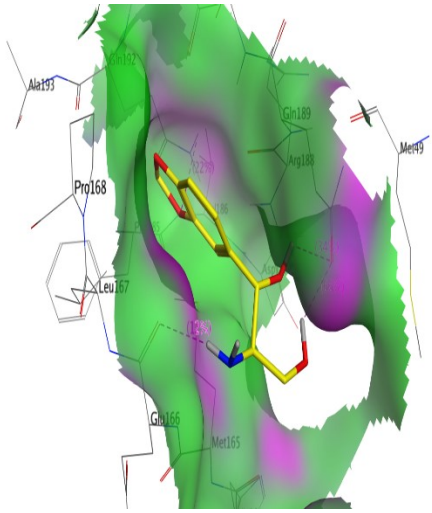
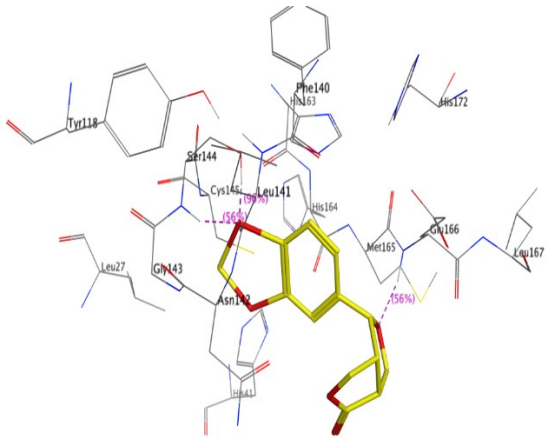
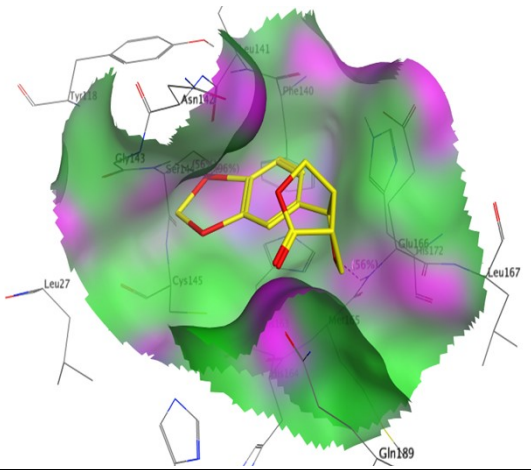
18	**3,5-dicaffeoylquinic acid** 

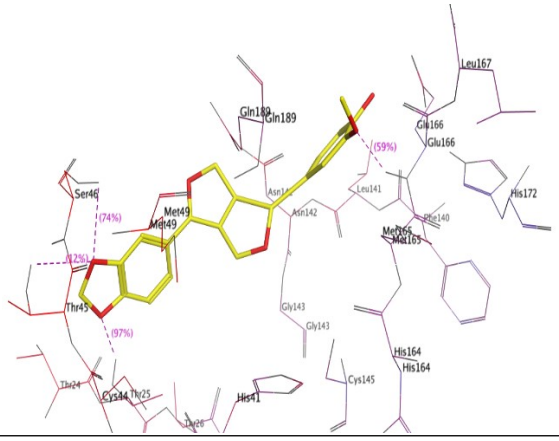
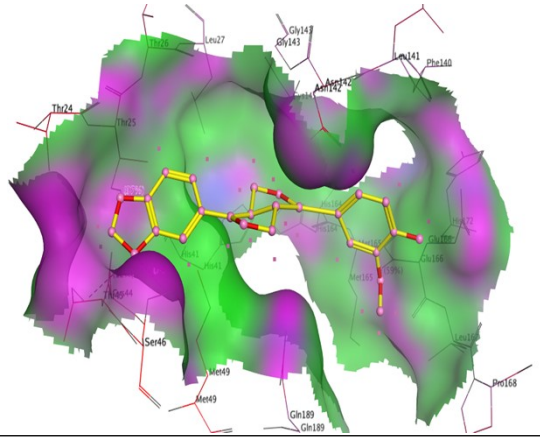
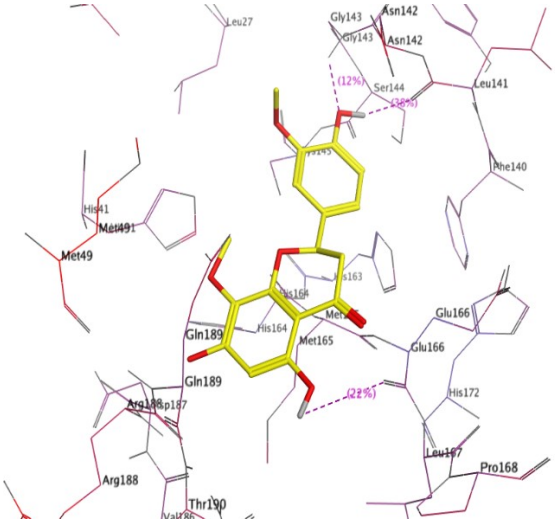
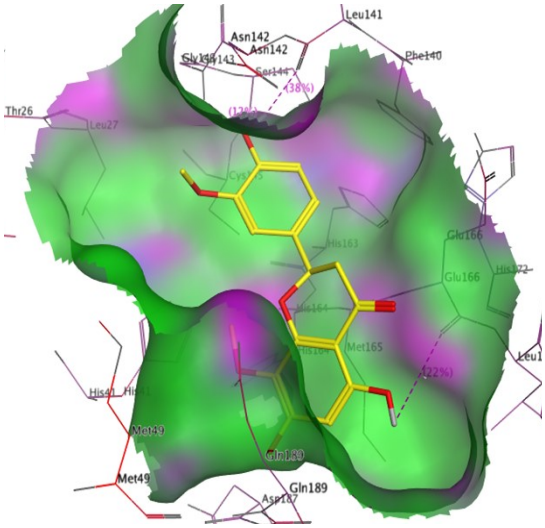
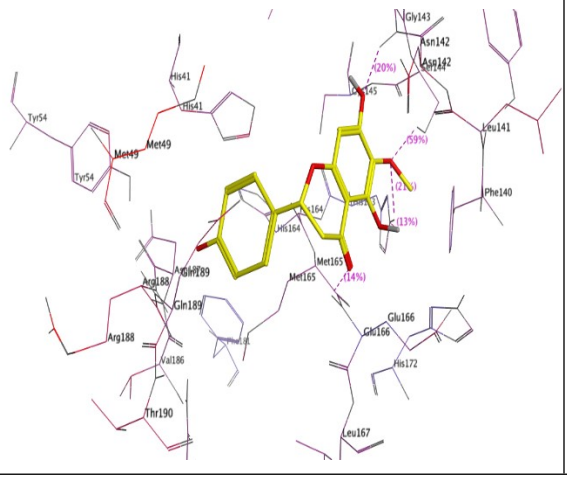
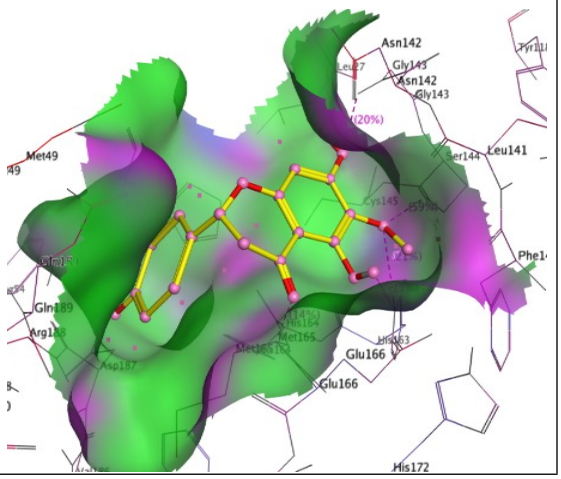
Table (S7): A) 3D binding mode and residues involved in the recognition the standard lopinavir and the isolated compounds docked and minimized in the SARS-COV-2M^{pro}binding pocket

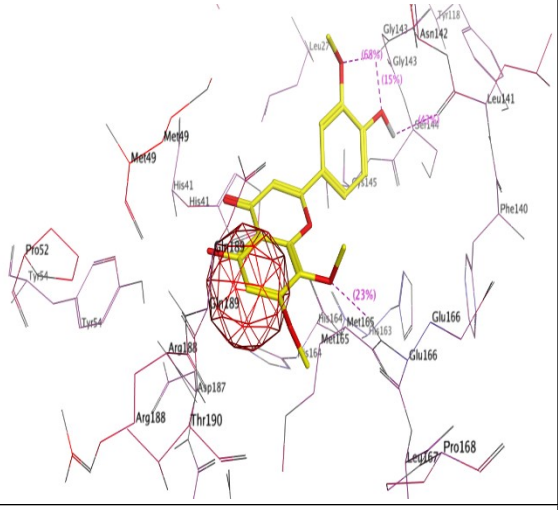
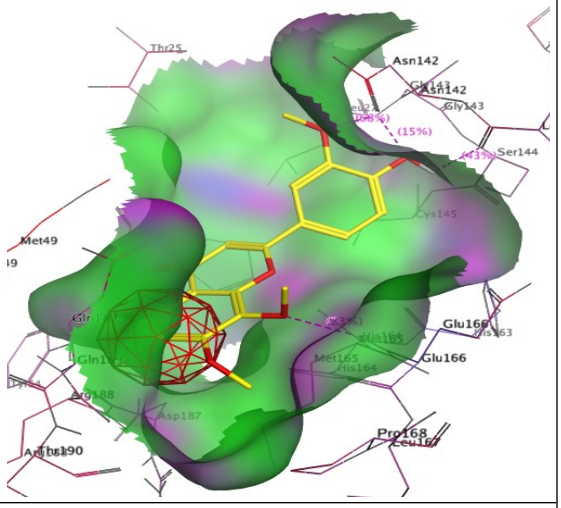
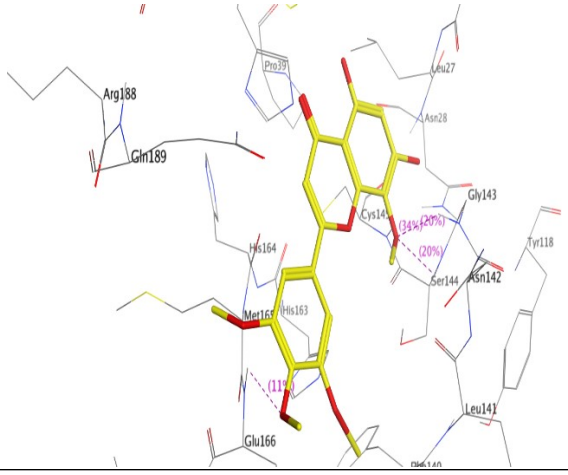
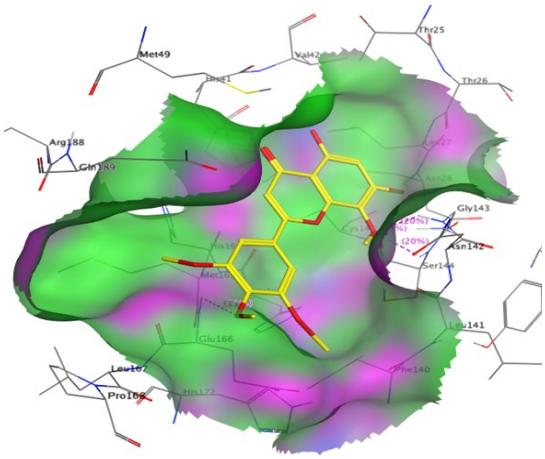
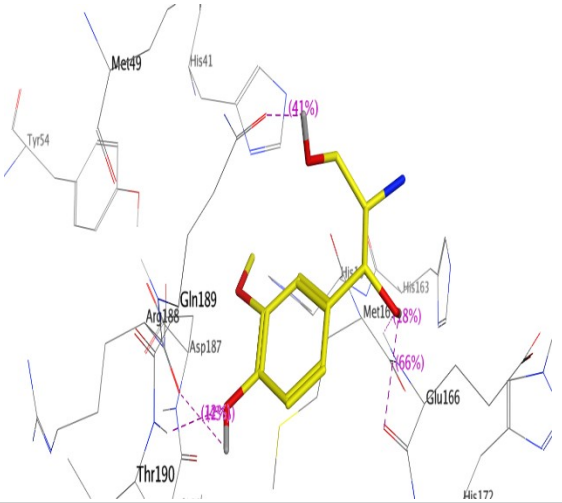
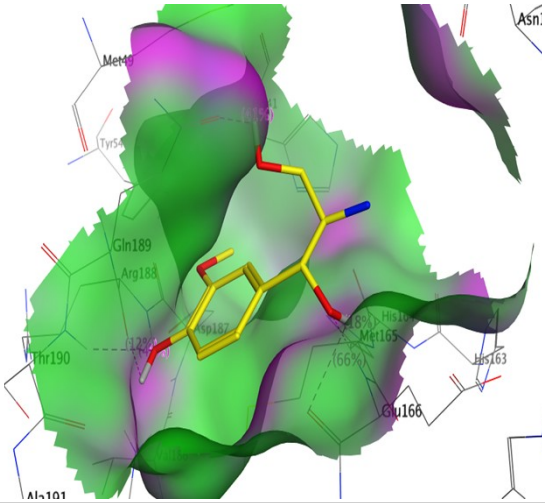
B) Surface and maps of the isolated compound compared to the standard lopinavir

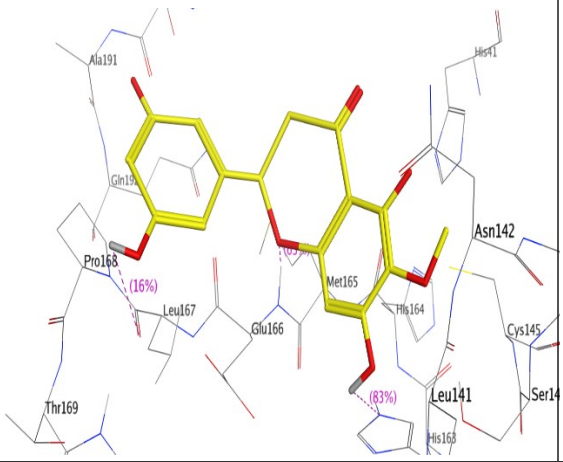
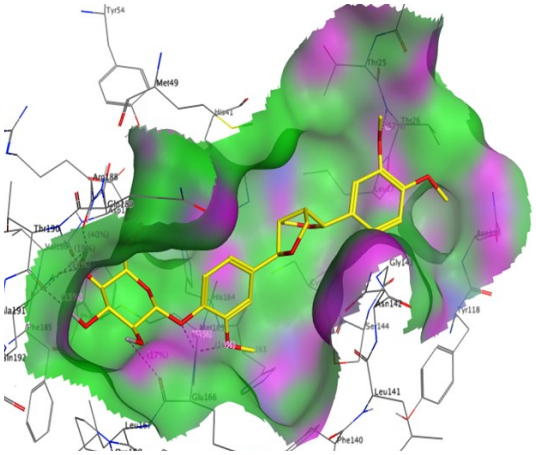
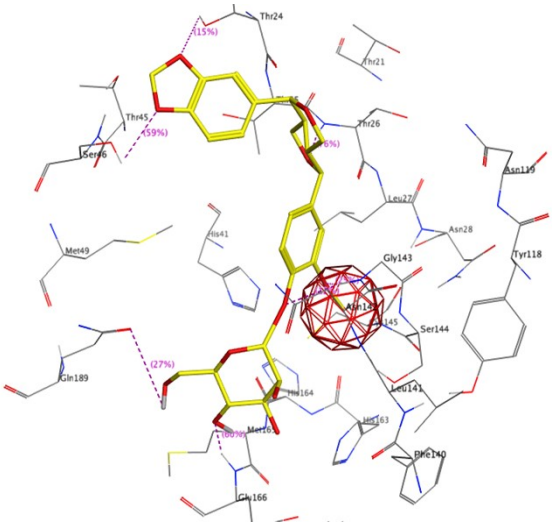
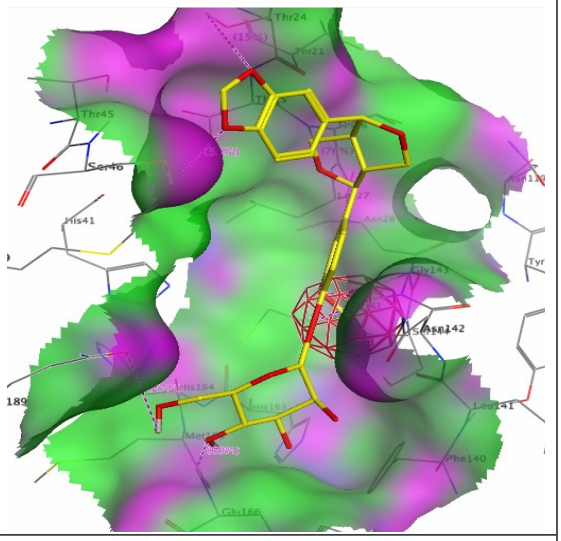
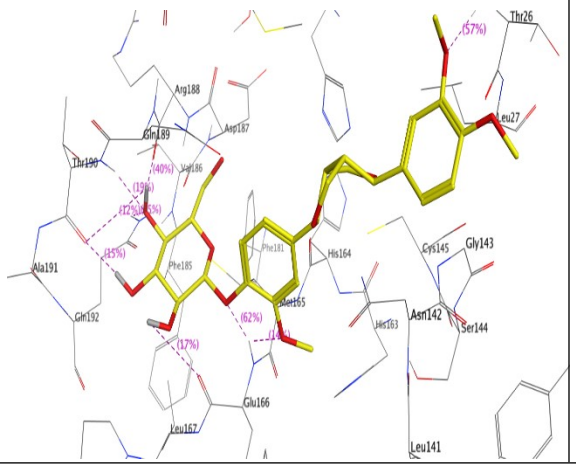
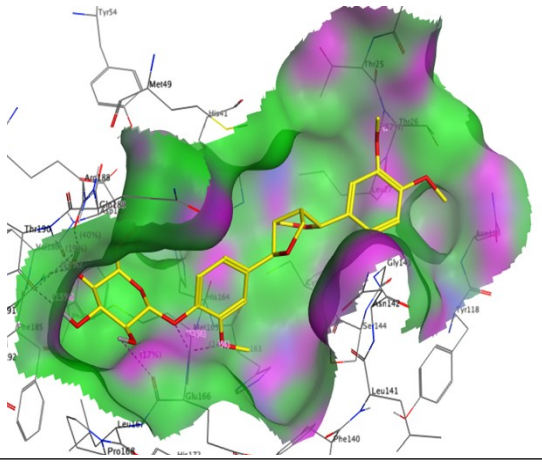
No.	Name of compounds	A	B
Stan-dard	**Lopinavir**		

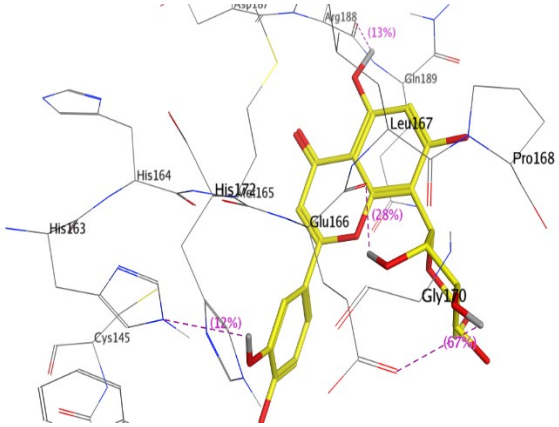
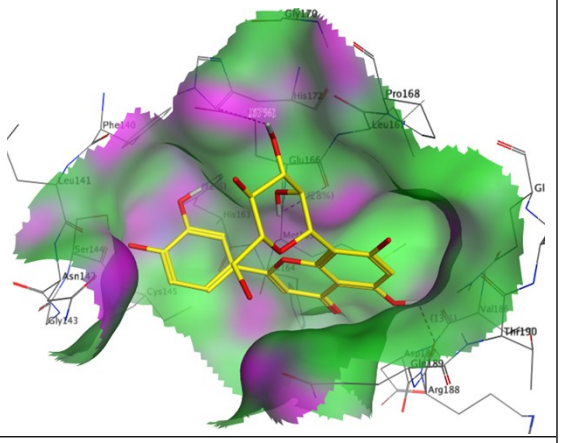
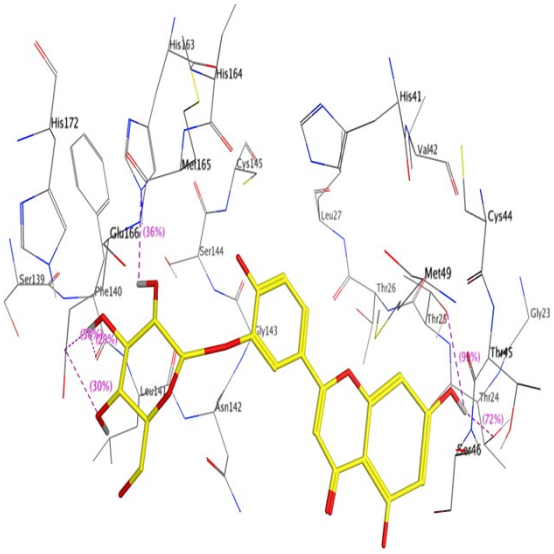
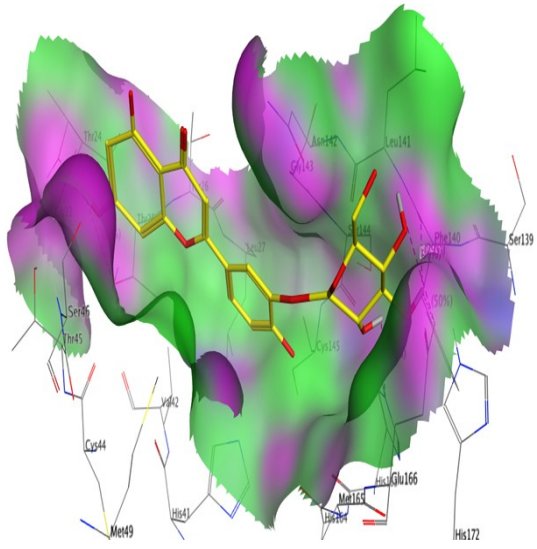
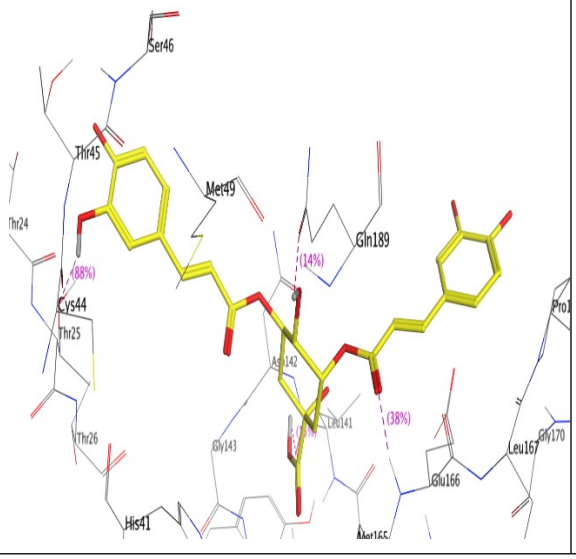
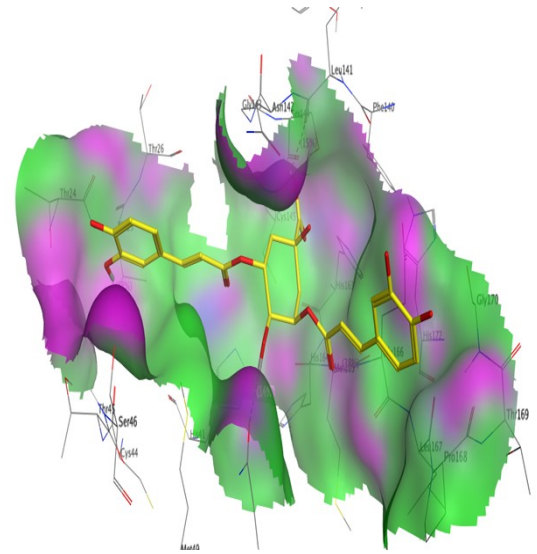
1	**24-*β*-ethyl-cholesta-5(6),22(23),25(26)-triene-3-ol**		
2	***α*-amyrin**		
3	**Linoleic acid**		

4	**24-*β*-ethyl-cholesta-5(6),22(23),25(26)-triene-3-O-*β*-D-glucoside**		
5	**1,3-propanediol-2-amino-1-(3`,4`-methylenedioxyphenyl)**		
6	**(-)-(7*R*,8*R*,8`*R*)-acuminatolide**		

7	**(+)-piperitol**		
8	**5,7,4`-trihydroxy-8,3`-dimethoxy flavanone**		
9	**5,7,4`-trihydroxy-6-methoxy flavanone**		

10	**4`,5-dihydroxy-3`,7,8-trimethoxyflavone**		
11	**5,7-dihydroxy-3`,4`,5`,8-tetramethoxy flavone**		
12	**1,3-propanediol-2-amino-1-(4`-hydroxy-3`-methoxyphenyl)**		

13	**3`,5`,5,7-tetrahydroxy-6-methoxyflavanone**		
14	**Simplexoside (piperitol-O-*β*-D-glucoside)**		
15	**Pinoresinol monomethyl ether-*β*-D-glucoside**		

16	**Orientin**		
17	**Luteolin-3`-O-*β*-D-glucoside**		
18	**3,5-dicaffeoylquinic acid**		

Data S1

Compound 1 (24-*β*-ethyl-cholesta-5(6),22(23),25(26)-triene-3-ol) was obtained as white powder. ^{1}H NMR ($CDCl_3$, 400 MHz): 3.49 (m, H-3), 5.32 (1H, d, *J*= 5.2 Hz, H-6), 0.67 (s, H-18), 0.99 (br s, H-19), 0.98 (br s, H-21), 5.22 (1H, dd, *J*= 15.2 & 8 Hz, H-22), 5.13 (1H, dd, *J*= 15.2 & 8 Hz, H-23), 4.67-4.69 (2H, m, H-26), 1.62 (s, H-27), 0.82 (1H, d, *J*= 7.6 Hz, H-29). DEPT Q ($CDCl_3$, 100 MHz): 37.3 (C-1), 31.7 (C-2), 71.8 (C-3), 39.8 (C-4), 140.8 (C-5), 121.7 (C-6), 31.9 (C-7), 31.9 (C-8), 50.1 (C-9), 36.5 (C-10), 21.1 (C-11), 39.7 (C-12), 42.3 (C-13), 56.9 (C-14), 24.3 (C-15), 28.7 (C-16), 55.9 (C-17), 12.1 (C-18), 19.4 (C-19), 40.2 (C-20), 20.8 (C-21), 137.2 (C-22), 130.0 (C-23), 52.0 (C-24), 148.7 (C-25), 109.5 (C-26), 20.3 (C-27), 25.7 (C-28), 12.2 (C-29).

Compound 2 (*α*-amyrin) was obtained as oily substance. ^{1}H NMR ($CDCl_3$, 400 MHz): 3.15 (m, H-3), 5.11 (t, H-12), 0.94 (s, 3H, H-23), 0.85 (s, 3H, H-24), 0.79 (m, H-25), 0.89 (s, 3H, H-26), 1.00 (s, 3H, H-27), 0.93 (s, 3H, H-28), 0.89 (s, 3H, H-29), 0.73 (s, 3H, H-30). APT ($CDCl_3$, 100 MHz): 38.8 (C-1), 27.3 (C-2), 77.4 (C-3), 38.5 (C-4), 55.2 (C-5), 18.4 (C-6), 32.9 (C-7), 40.0 (C-

8), 47.7 (C-9), 36.8 (C-10), 23.4 (C-11), 124.4 (C-12), 139.6 (C-13), 42.0 (C-14), 29.4 (C-15), 26.6 (C-16), 33.8 (C-17), 59.1 (C-18), 39.7 (C-19), 39.6 (C-20), 31.3 (C-21), 41.5 (C-22), 28.8 (C-23), 15.7 (C-24), 15.6 (C-25), 16.9 (C-26), 23.4 (C-27), 28.2 (C-28), 17.5 (C-29), 21.4 (C-30).

Compound 3 (linoleic acid) was obtained as white powder. ^{1}H NMR ($CDCl_3$, 400 MHz): 2.32 (H-2), 1.61 (H-3), 1.29 (H-4), 1.29 (H-7), 2.07 (H-8), 5.34 (H-9,10), 2.78 (H-11), 5.34 (H-12,13), 1.29 (H-15&17), 0.97 (H-18). APT ($CDCl_3$, 100 MHz): 179.1 (C-1), 33.9 (C-2), 24.7 (C-3), 29.0 (C-4), 29.3 (C-5), 29.6 (C-6), 29.7 (C-7), 27.2 (C-8), 130.3 (C-9), 128.3 (C-10), 25.5 (C-11), 130.1 (C-12), 127.9 (C-13), 25.6 (C-14), 29.4 (C-15), 31.9 (C-16), 22.7 (C-17), 14.3 (C-18).

Compound 4 (24-*β*-ethyl-cholesta-5(6),22(23),25(26)-triene-3-O-*β*-D-glucoside) was obtained as white powder. ^{1}H NMR (DMSO- d_6, 400 MHz): 3.66 (m, H-3), 5.34 (m, H-6), 0.68 (s, H-18), 0.99 (s, H-19), 0.97 (s, H-21), 5.28-5.14 (m, H-22,23), 4.91 (m, H-26), 1.62 (s, H-27), 0.80 (s, H-29), 5.03 (d, *J*= 8.8 Hz, H-1`), 4.24 (d, *J*= 7.6 Hz, H-2`), 4.70 (m, H-3`, 4`), 3.66 (m, H-5`), 4.46 (t, *J*= 11.2 Hz, 5.6 Hz, H-6`a), 4.70 (m, H-6`b). APT (DMSO- d_6, 100 MHz): 37.3 (C-1), 29.7 (C-2), 77.2 (C-3), 42.3 (C-4), 140.9 (C-5), 121.7 (C-6), 31.9 (C-7&8), 50.1 (C-9), 36.7 (C-10), 21.1 (C-11), 38.7 (C-12), 42.3 (C-13), 56.7 (C-14), 24.3 (C-15), 28.8 (C-16), 55.7 (C-17), 12.3 (C-18), 19.6 (C-19), 40.6 (C-20), 21.2 (C-21), 137.3 (C-22), 130.0 (C-23), 51.7 (C-24), 148.2 (C-25), 110.5 (C-26), 20.4 (C-27), 25.7 (C-28), 12.5 (C-29), 101.2 (C-1`), 73.9 (C-2`), 77.4 (C-3`), 70.6 (C-4`), 77.2 (C-5`), 61.6 (C-6`).

Compound 5 (1,3-propanediol-2-amino-1-(3`,4`-methylenedioxyphenyl) or (1`,3`-propanediol-2`-amino-1`-(1,3-benzodioxol-5-yl)) was obtained as white powder. ^{1}H NMR ($CDCl_3$, 400 MHz): 4.72 (d, 1H, *J* =3.2 Hz, H-1), 3.05 (m, 1H, H-2), 4.23 (dd, 1H, *J* = 6.4 Hz, 8.4 Hz, H-3a), 3.87 (dd, 1H, *J* = 6.8 Hz, 2 Hz, H-3b), 6.85 (br s, 1H, H-2`), 6,78 (br d, 1H, *J*= 8, H-5`), 6.80 (br d, 1H, *J*= 10.8, H-6`), 5.95 (s, 2H, $OCH_2O$). APT ($CDCl_3$, 100 MHz): 85.8 (C-1), 54.3 (C-2), 71.7 (C-3), 135.0 (C-1`), 106.5 (C-2`), 148.0 (C-3`), 147.1 (C-4`), 108.2 (C-5`), 119.4 (C-6`), 101.1 (OCH_2O).

Compound 6 ((-)-(7*R*,8*R*,8`*R*)-acuminatolide) was obtained as white powder. $^1$H NMR ($CDCl_3$, 400 MHz): 6.84& 6.79 (br. s, 3H aromatic), 5.97 (s, 2H, $OCH_2O$), 4.60 (d, 1H, $J_{7,8}$= 6.8 Hz, H-7), 3.06-3.11 (m, 1H, H-8), 3.44 (ddd, 1H, $J_{7,8`}$= 3.6 Hz, $J_{8`,9`eq}$ = 3.6 Hz, $J_{8`,8}$ = 3.6 Hz, H-8`), 4.49 (dd, 1H, $J_{9eq,9ax=}$ 6.8, $J_{8,9eq}$ = 6.8, H-9eq), 4.38-4.31 (m, 2H, H-9ax & H-9`ax) , 4.19 (dd, 1H, $J_{9`eq,9`ax=}$ 3.6, $J_{8`,9`eq}$= 3.6, H-9`eq). APT ($CDCl_3$, 100 MHz): 132.7 (C-1), 106.4 (C-2), 148.2 (C-3), 147.8 (C-4), 108.4 (C-5), 119.7 (C-6), 101.3 (OCH_2O), 86.1 (C-7), 178.1 (C-7`), 48.4 (C-8), 46.0 (C-8`), 70.1 (C-9), 69.8 (C-9`).

Compound 7 ((+)-piperitol) was obtained as white powder. ^{1}H NMR ($CDCl_3$, 400 MHz): 6.77-6.89 (6H, m, Ar-H), 5.95 (s, 2H, OCH_2O), 3.89 (s, 3H, OCH_3) 5.68 (s, 1H, 4`- OH), 4.73 (2H, d, *J*= 2, H-7&7`), 3.03-3.12 (2H, m, H-8&8`), 4.21- 4.27 (2H, m, H-9a&9`a), 3.86-3.89 (2H, m, H-9b&9`b). APT ($CDCl_3$, 100 MHz): 135.1 (C-1), 106.5 (C-2), 147.9 (C-3), 147.1 (C-4), 108.6 (C-5),

119.4 (C-6),101.1 (OCH_2O), 132.9 (C-1`), 108.2 (C-2`), 146.7 (C-3`), 145.3 (C-4`), 114.3 (C-5`), 118.9 (C-6`), 55.9 (OCH_3),85.9 (C-7), 85.8 (C-7`), 54.3 (C-8), 54.2 (C-8`), 71.7 (C-9&9`).

Compound 8 (5,7,4`-trihydroxy-8,3`-dimethoxyflavanone) was obtained as yellowish white powder. ^{1}H NMR (CD_3OD, 400 MHz): 5.18 (1H, dd, *J*= 2.8 Hz, 12.6 Hz, H-2), 2.56 (1H, dd, *J*= 2.8 Hz, 17.2 Hz, H-3α), 2.97 (1H, dd, *J*= 12.8 Hz, 17.2 Hz, H-3*β*), 5.79 (s, 1H, H-6), 3.66 (s, 3H, R_4), 3.78 (s, 3H, R_5), 6.97 (1H, d, *J*= 1.6 Hz, H-2`), 6.71 (1H, d, *J*= 8 Hz, H-5`), 6.81 (1H, dd, *J*= 8.2 Hz, 2 Hz, H-6`). APT ($CD_3OD$, 100 MHz): 79.2 (C-2), 42.7 (C-3), 195.7 (C-4), 159.0 (C-5), 96.3 (C-6), 164.6 (C-7), 130.7 (C-8), 155.0 (C-9), 100.7 (C-10), 59.4 ($R_4$), 55.1 ($R_5$), 130.1 (C-1`), 109.8 (C-2`), 147.7 (C-3`), 146.6 (C-4`), 114.7 (C-5`), 119.0 (C-6`).

Compound 9 (5,7,4`-trihydroxy-6-methoxy flavanone) was obtained as yellowish white powder. ^{1}H NMR (CD_3OD, 400 MHz): 5.22 (1H, dd, *J*= 2.8 Hz, 13 Hz, H-2), 2.60 (1H, dd, *J*= 2.8 Hz, 17 Hz, H-3α), 3.01 (1H, dd, *J*= 12.8 Hz, 17.2 Hz, H-3*β*), 5.87 (s, 1H, H-8), 3.67 (s, 3H, OCH_3), 7.21 (2H, d, *J*= 8.4 Hz, H-2`,6`), 6.71 (2H, d, *J*= 8 Hz, H-3`,5`). APT (CD_3OD, 100 MHz): 79.2 (C-2), 42.7 (C-3), 197.2 (C-4), 155.2 (C-5), 129.0 (C-6), 159.4 (C-7), 94.8 (C-8), 157.6 (C-9), 102.1 (C-10), 59.6 (OCH_3), 129.7 (C-1`), 127.6 (C-2`,6`), 114.9 (C-3`,5`), 158.8 (C-4`).

Compound 10 (4`,5-dihydroxy-3`,7,8-trimethoxyflavone) was obtained as yellow powder. ^{1}H NMR (DMSO- d_6, 400 MHz): 6.99 (1H, s, H-3), 6.59 (1H, s, H-6), 3.92 (3H, s, R_3), 3.86 (3H, s, R_4), 7.59 (1H s, H-2`), 7.00 (1H, br d, *J*= 6.7 Hz, H-5`), 7.60 (1H, d, *J*= 6.0 Hz, H-6`), 3.90 (3H, s, $R_5$), 12.97 (5-OH), 10.08 (4`-OH). DEPT Q (DMSO- d_6, 100 MHz): 164.3 (C-2), 103.5 (C-3), 182.7 (C-4), 157.1 (C-5), 96.4 (C-6), 158.8 (C-7), 128.9 (C-8), 151.4 (C-9), 104.3 (C-10), 56.9 (R_3), 61.6 (R_4), 121.9 (C-1`), 110.4 (C-2`), 149.2 (C-3`), 148.5 (C-4`), 116.6 (C-5`), 120.8 (C-6`), 56.5 (R_5).

Compound 11 (5,7-dihydroxy-3`,4`,5`,8-tetramethoxy flavone) was obtained as yellow substance. $^1$H NMR ($CDCl_3$, 400 MHz): 6.62 (1H, s, H-3), 6.43 (1H, s, H-6), 4.00 (3H, s, $R_4$), 7.13 (2H, s, H-2`,6`), 3.95 (6H, s, $R_5$, $R_7$), 3.94 (3H, s, $R_6$). APT ($CDCl_3$, 100 MHz): 163.2 (C-2), 105.4 (C-3), 182.4 (C-4), 155.5 (C-5), 99.0 (C-6), 157.7 (C-7), 126.9 (C-8), 148.9 (C-9), 105.1 (C-10), 61.8 ($R_4$), 126.3 (C-1`), 103.7 (C-2`,6`), 153.7(C-3`,5`), 141.6 (C-4`), 56.3 (R_5, R_7), 61.1 (R_6).

Compound 12 (1,3-propanediol-2-amino-1-(4`-hydroxy-3`-methoxyphenyl) was obtained as colorless needles. ^{1}H NMR ($CDCl_3$, 400 MHz): 4.74 (d, 1H, *J*= 4 Hz, H-1), 3.10 (m, 1H, H-2), 4.25 (dd, 1H, *J* =9.2 & 6.4 Hz, H-3a), 3.88 (dd, 1H, *J* =9.2 Hz, 3.6 Hz, H-3b), 6.90 (d, 1H, *J*= 2 Hz, H-2`), 6.89 (d, 1H, *J*= 7.6 Hz, H-5`), 6.82 (dd, 1H, *J*= 8.2 Hz, 1.6 Hz, H-6`), 3.91 (s, 3H, $OCH_3$). APT ($CDCl_3$, 100 MHz): 85.9 (C-1), 54.2 (C-2), 71.7 (C-3), 132.9 (C-1`), 108.6 (C-2`), 146.7 (C-3`), 145.3 (C-4`), 114.3 (C-5`), 118.9 (C-6`), 55.9 (OCH_3).

Compound 13 (3`,5`,5,7-tetrahydroxy-6-methoxyflavanone) was obtained as yellow powder. ^{1}H NMR (CD_3OD, 400 MHz): 5.16 (1H, dd, *J*= 2.8 Hz, 12.8 Hz, H-2), 2.60 (1H, dd, *J*= 2.8 Hz, 17.2 Hz, H-3α), 2.96 (1H, dd, *J*= 12.8 Hz, 17.2 Hz, H-3*β*), 5.87 (1H, s, H-8), 3.68 (s, 3H, OCH_3), 6.68 (2H, s, H-2`,4`), 6.81 (1H, s, H-6`). APT ($CD_3OD$, 100 MHz): 79.2 (C-2), 42.7 (C-3), 197.2 (C-4), 155.2 (C-5), 129.0 (C-6), 159.5 (C-7), 94.8 (C-8), 158.8 (C-9), 102.1 (C-10), 59.6 ($OCH_3$), 130.3 (C-1`), 117.9 (C-2`), 145.1 (C-3`), 114.8 (C-4`), 145.5 (C-5`),113.3 (C-6`).

Compound 14 (simplexoside (piperitol-O-*β*-D-glucoside)) was obtained as white powder. ^{1}H NMR (DMSO- d_6, 400 MHz): 6.86- 7.06 (6H, m, aromatic protons), 3.78 (s, 3H, OCH_3), 6.00 (s, 2H, OCH_2O), 4.67 (s, 2H, H-7&7`), 3.04 (2H, m, H- 8&8`), 4.14 (2H, t, *J*= 15.6, 7.1, H-9a,9`a), 3.66-3.69 (2H, d, *J*= 11.4, H-9b, 9`b), 4.88 (1H, s, *J*= 6.7 Hz, H-1``), 3.37 (m, protons of sugar). APT (DMSO- d_6, 100 MHz): 135.9 (C-1), 107.1 (C-2), 149.4 (C-3), 147.9 (C-4), 111.0 (C-5), 119.9 (C-6), 101.4 (OCH_2O), 135.6 (C-1`), 108.5 (C-2`), 146.9 (C-3`), 146.3 (C-4`), 115.6 (C-5`), 118.6 (C-6`),56.2 (OCH_3), 85.4 (C-7), 85.3 (C-7`), 54.2 (C-8), 54.1 (C-8`), 71. 6 (C-9), 71.5 (C-9`), 100.6 (C-1``), 73.7 (C-2``), 77.3 (C-3``), 70.1 (C-4``), 77.4 (C-5``), 61.1 (C-6``).

Compound 15 (pinoresinol monomethyl ether-*β*-D-glucoside) was obtained as white powder. ^{1}H NMR (CD_3OD, 400 MHz): 6.85 (d, *J*= 2 Hz, 1H, H-2), 6.66 (d, *J*= 8 Hz, 1H, H-5), 6.71 (dd, *J*= 8.2 Hz, 2 Hz, 1H, H-6), 3.76 (s, 6H, $R_{1,2}$), 6.93 (d, *J*= 2 Hz, 1H, H-2`), 7.05 (d, *J*= 8.4 Hz, 1H, H-5`), 6.82 (dd, *J*= 8 Hz, 2 Hz, 1H, H-6`), 3.77 (s, 3H, $R_3$), 4.61 (d, *J*= 4 Hz, 1H, H-7), 4.66 (d, *J*= 4 Hz, 1H, H-7`), 3.04 (m, 2H, H-8&8`), 3.53-3.61 (m, 2H, H-9a&9`a), 4.12-4.17 (m, 2H, H-9b& 9`b), 4.78 (d, *J*=7.2 Hz, 1H, H-1``), 3.29-3.41 (m, 4H, H- 2``,3``,4``,5``), 3.53-3.61 (m, 2H, H-6``a, 6``b). APT ($CD_3OD$, 100 MHz): 132.3 (C-1), 109.5 (C-2), 147.7 (C-3), 146.1 (C-4), 114.7 (C-5), 118.6 (C-6), 54.9 ($R_{1\&}R_2$), 136.0 (C-1`), 110.1 (C-2`), 149.5 (C-3`), 145.9 (C-4`), 116.6 (C-5`), 118.4 (C-6`), 55.3 ($R_3$), 86.1 (C-7), 85.7 (C-7`),54.1 (C-8), 53.9 (C-8`), 71.3 (C-9&9`), 101.4 (C-1``), 73.5 (C-2``), 76.4 (C-3``), 69.9 (C-4``), 76.8 (C-5``), 61.1 (C-6``).

Compound 16 (orientin) was obtained as yellow powder. ^{1}H NMR (DMSO- d_6, 400 MHz):6.68 (s, 1H, H-3), 6.29 (s, 1H, H-6), 7.51 (br s, 1H, H-2`), 6.89 (d, *J*= 8.4 Hz, 1H, H-5`), 7.56 (br d, *J*= 8.4 Hz, 1H, H-6`), 4.70 (d, *J*= 9.6 Hz, 1H, H-1``), 3.26- 3.94 (m, 6H, H-2``,3``,4``,5``,6``a,6``b), 13.20 (s, 5-OH).APT (DMSO- $d_6$, 100 MHz): 164.6 (C-2), 102.9 (C-3), 182.5 (C-4), 160.9 (C-5), 98.7 (C-6), 163.1 (C-7), 105.0 (C-8), 156.5 (C-9), 104.5 (C-10), 122.5 (C-1`), 114.5 (C-2`), 146.3 (C-3`), 150.2 (C-4`), 116.1 (C-5`), 119.9 (C-6`), 73.9 (C-1``), 71.2 (C-2``), 79.2 (C-3``), 71.2 (C-4``), 82.5 (C-5``), 62.1(C-6``).

Compound 17 (luteolin-3`-O-*β*-D-glucoside) was obtained as yellow powder. $^1$H NMR ($CD_3OD$, 400 MHz): 6.52 (s, H-3), 6.12 (s, H-6), 6.36 (s, H-8), 7.35 (br s, H-2`), 7.22 (d, *J*= 8.4 Hz, H-5`), 7.36 (br d, *J*= 9.6 Hz, H-6`), 4.8 (H-1``, masked), 3.3-3.9 (m, H-2``,3``,4``,5``,6``). APT (CD_3OD, 100 MHz): 164.8 (C-2), 103.7 (C-3), 182.5 (C-4), 161.8 (C-5), 98.8 (C-6), 164.1 (C-7), 93.7 (C-8), 158.0 (C-9), 104.1 (C-10), 125.8 (C-1`), 113.5 (C-2`), 148.6 (C-3`), 147.2 (C-4`), 116.5 (C-5`), 118.4 (C-6`), 101.8 (C-1``), 73.4 (C-2``), 76.1 (C-3``), 69.9 (C-4``), 77.1 (C-5``), 60.1 (C-6``).

Compound 18 (3,5-dicaffeoylquinic acid (isochlorogenic acid)) was obtained as yellow powder. ^{1}H NMR (CD_3OD, 400 MHz): 2.31 (m, H-2eq), 2.08-2.15 (m, H-2ax, 6eq, 6ax), 5.41 (m, H-3), 3.95 (dd, *J*= 9.36, 3.4, H-4), 5.53 (m, H-5), 6.79 (s, H-2`), 6.81 (s, H-2``), 7.09 (d, *J*= 7.8 Hz, H-5`, 5``), 6.99 (dd, *J*= 7.5, 2.2 Hz, H-6`, 6``),7.60 (d, *J*=15.9 Hz, H-7`), 7.63 (d, *J*= 15.9 Hz, H-7``), 6.32 (d, *J*= 15.9 Hz, H-8`), 6.42 (d, *J*= 15.9 Hz, H-8``). APT ($CD_3OD$, 100 MHz): 74.6 (C-1), 35.9 (C-2), 72.7 (C-3), 71.3 (C-4), 70.9 (C-5), 38.9 (C-6), 170.0 (C-7), 126.6 (C-1`), 126.4 (C-1``), 113.6 (C-2`), 113.7 (C-2``), 145.4 (C-3`, 3``), 148.1 (C-4`), 148.0 (C-4``), 115.1 (C-5`), 115.0 (C-5``), 121.6 (C-6`), 121.5 (C-6``), 145.5 (C-7`), 145.4 (C-7``), 114.6 (C-8`), 114.1 (C-8``), 167.4 (C-9`), 167.9 (C-9``).

End of Supplementary material file