Supporting Information

Design, Synthesis, and Mechanism Study of Benzenesulfonamide-Containing Phenylalanine Derivatives as Novel HIV-1 Capsid Inhibitors with Improved Antiviral Activities

Lin Sun[†], Alexej Dick[‡], Megan E. Meuser[‡], Tianguang Huang[†], Waleed A. Zalloum[§], Chin-Ho Chen[∥], Srinivasulu Cherukupalli[†], Shujing Xu[†], Xiao Ding[†], Ping Gao[†], Dongwei Kang[†], Erik De Clercq[#], Christophe Pannecouque[#], Simon Cocklin[‡,*], Kuo-Hsiung Lee[⊥,*], Xinyong Liu[†,*], Peng Zhan[†,*]

[†]*Department of Medicinal Chemistry, Key Laboratory of Chemical Biology (Ministry of Education), School of Pharmaceutical Sciences, Shandong University, 44 West Culture Road, Ji'nan, Shandong 250012, People's Republic of China*

[‡]*Department of Biochemistry & Molecular Biology, Drexel University College of Medicine, Philadelphia, Pennsylvania 19102, United States*

[§]*Department of Pharmacy, Faculty of Health Science, American University of Madaba, P.O Box 2882, Amman 11821, Jordan*

[∥]*Duke University Medical Center, Surgical Oncology Research Facility, Box 2926, Durham, North Carolina 27710, United States*

[⊥]*Natural Products Research Laboratories, Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, North Carolina 27599, United States*

[#]*Rega Institute for Medical Research, Laboratory of Virology and Chemotherapy, K.U. Leuven, Herestraat 49 Postbus 1043 (09.A097), B-3000, Leuven, Belgium*

**E-mail: sc349@drexel.edu.*

**E-mail: khlee@unc.edu.*

*E-mail: xinyongl@sdu.edu.cn.

*E-mail: zhanpeng1982@sdu.edu.cn.

Contents:

I. Molecular Dynamics Simulation on 8a and 6k

1. Methods

The most active compounds from Series I (**6k**) and II (**8a**) were docked to the active site using Autodock 4.2.6 with default settings.[S1] To keep the consistency of the MD simulation with the previously published HIV-1 CA monomer inhibitors, we used the same procedure for MD simulation and its analysis, refer to our previously published research for methodology.[S2]

2. Results and Discussion

The most active inhibitors of Series II (**8a**) and I (**6k)** were chosen for the MD for 1 μs to find their binding interactions to the active site. The root mean square deviation (RMSD) of amino acids during the simulation was calculated for the monomer upon binding of both compounds, see **Figure S1 A** and **C**. According to the figure, binding of both compounds produced similar RMSD deviation from the X-ray structure. Also, the figure shows that the HIV-1 CA monomer exists in different conformational forms with a highly abundant conformational form. This indicates that the inhibitor could bind with different binding modes to the protein. The RMSD of **8a** and **6k** was calculated and plotted in **Figure S1 B** and **C**, respectively. The figure shows that **8a** exists in different conformations, and **6k** exists in few conformational ensembles. Different RMSD indicates that both inhibitors bind with different binding modes to the active site.

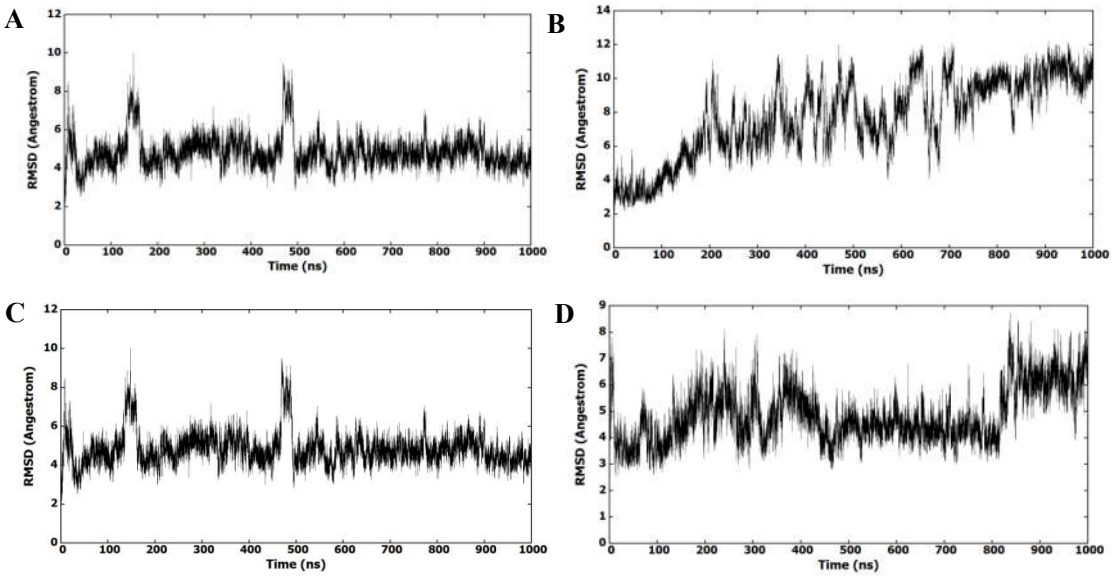


Figure S1. RMSD (heavy atoms) of amino acids of HIV-1 CA monomer in **8a** (**A**) and **6k** (**C**) in reference to the first frame of the MD simulation. RMSD (heavy atoms) of the bound inhibitors **8a** (**B**) and **6k** (**D**) regarding the docked conformer.

To further investigate the binding studies of both compounds to the active site, the entire trajectory has been clustered based on each inhibitor (no fit). Clustering resulted in eighteen clusters for **8a** and three clusters for **6k**. The clustering procedure yielded two most populated clusters of the eighteen clusters in **8a**, whereas **6k** yielded one most populated cluster of the three clusters. **Figure S2 A** and **B** shows representative structure interactions of the first (24.3%) and second (10.3%) clusters of **8a**, while **Figure S3 C** and **D** shows first (69.0%) and second (11.0%) clusters of **6k**. The figure shows that both inhibitors bind in different binding modes to the active site. The binding of both compounds showed similarity to the binding of the reference inhibitor **PF-74**, where the core scaffold is oriented to the inside of the active site and the substituent is oriented to the outside of the active site.

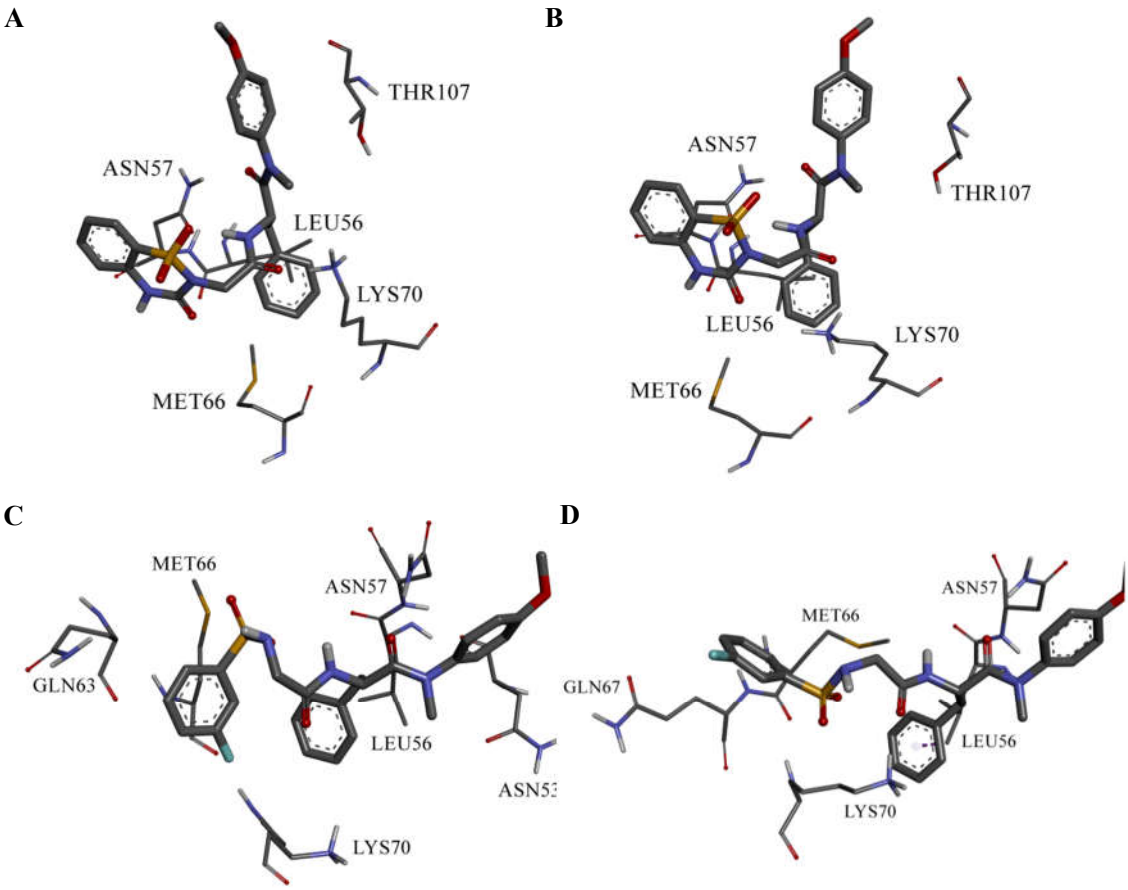


Figure S2. Binding interactions of **8a** in the first (**A**) and second (**B**) clusters. Binding interactions of **6k** in the first (**C**) and second (**D**) clusters.

The phenyl ring of the core region of both **8a** and **6k** forms hydrophobic interaction with Lys70 in the first cluster, and it does not form any interactions with Lys70 in the second cluster. Lys70 could form an ion-induced dipole with the benzene ring of the core region in the first cluster of **8a** and the second cluster of **6k**. Also, both inhibitors form aliphatic-aromatic hydrophobic interactions with Leu56 with its core benzene ring in the first cluster similar to the binding of **PF-74**. Leu56 is far from binding to the aromatic ring in the second cluster for both inhibitors. Accordingly, it is clear that, both

inhibitors bind the same way in the core region of the compound. Methoxy group of the core region does not form any interaction with the active site in both inhibitors binding. However, in the second cluster of **6k** could form a hydrogen bond with Asn57. The oxygen atom of the core region amide forms a hydrogen bond to Asn57 in all clusters of both inhibitors. The phenyl ring of methoxybenzene of the core region of **8a** forms hydrophobic interaction with Thr107 in both clusters. The amide oxygen atom of the side chain forms a hydrogen bond to Lys70, and its nitrogen atom forms a hydrogen bond to Asn57 in both clusters of **8a**. The amide oxygen atom of the side chain of **6k** forms hydrogen bond to Lys70 in the least populated cluster. The sulphonamide region of **8a** does not form any interaction with the binding site. However, the oxygen atom of the sulphonamide in **6k** is involved in a hydrogen bond with Lys70 in the least populated cluster. Benzene ring next to the sulphonamide moiety does not form any interaction to the binding site in **8a**. However, the fluorobenzene ring in **6k** forms hydrophobic interaction with Met66. To further investigate hydrogen bond formation during MD simulation, we hold hydrogen bond analysis for the whole trajectory, see **Table S1**. According to **Table S1**, it is clear that the hydrogen bond formation is preferred with two amino acids for both inhibitors, Lys70 and Asn57. Inspection of the frequency of hydrogen bond formation during the 1 μs MD simulation shows that **8a** has a higher frequency of interaction to both amino acids than **6k**. The higher frequency of hydrogen bonding in **8a** (EC_{50} = 2.11 μM) explains the higher activity of this inhibitor over that of **6k** (EC_{50} = 5.61 μM).

Table S1. Hydrogen bond analysis for **8a** and **6k** and their corresponding frequencies during the 1μs simulation.

8a			**6k**		
Residues involved		Frequency %	Residues involved		Frequency %
Asn57	MOL-O35	70.8	Asn57	MOL-O35	61.0
Asn57	MOL-H13	60.6	Asn57	MOL-H13	34.8
Lys70	MOL-O24	39.8	Lys70	MOL-O24	39.2

Refer to **Figure S3** for atom numbers

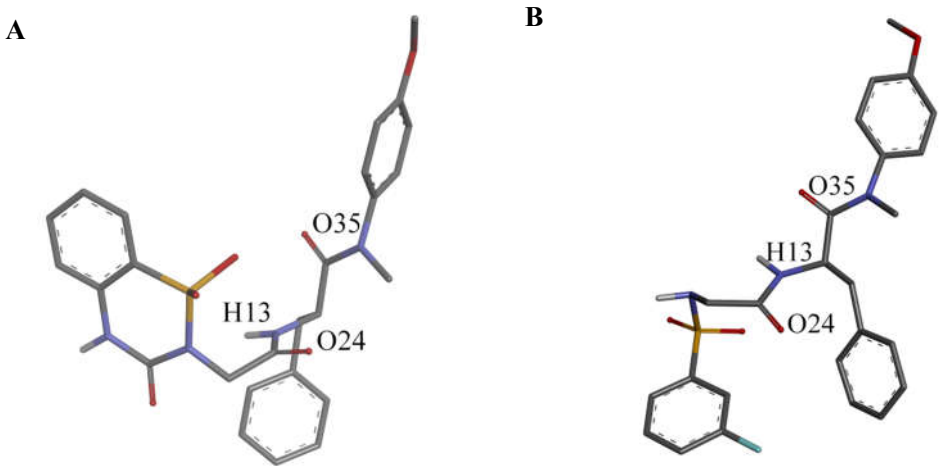


Figure S3. The numbering of atoms for hydrogen bond analysis in **Table S1** for **8a** (**A**) and **6k** (**B**).

II. Metabolic Stability in Human Liver Microsomes

The metabolic stability in human liver microsomes of compounds was determined in WuXi AppTec Co. Ltd. (Shanghai), China. The detailed procedure is as follows:

1. Test Compounds

Table S2. Compound information.

Compound No.	Compound ID	Batch No.	Exact Mass	Stock Concentration (mM)
1	11l	NA	579.22	10
2	PF-74	NA	425.21	10
Control	Testosterone		288.42	10
Control	Diclofenac		295.14	10
Control	Propafenone		341.44	10

2. Experimental Procedure

2.1. Test Compound and Control Working Solution Preparation:

2.1.1. Intermediate solution: 5 μL of compound stock solution (10 mM in dimethyl sulfoxide (DMSO)) were diluted with 495 μL of methanol (MeOH) (intermediate solution concentration: 100 μM, 99% MeOH)

2.1.2. Working solution: 50 μL of compound intermediate solution (100 μM) was diluted with 450 μL of 100 mM potassium phosphate buffer (working solution concentration: 10 μM, 9.9% MeOH)

2.2. NADPH Cofactor Preparation:

2.2.1. Materials:

NADPH powder: β-Nicotinamide adenine dinucleotide phosphate reduced form, tetrasodium salt; NADPH·4Na (Vendor: Chem-Impex International, Cat. No. 00616)

2.2.2. Preparation Procedure:

The appropriate amount of NADPH powder was weighed and diluted into a 10 mM $MgCl_2$ solution (working solution concentration: 10 unit/mL; final concentration in reaction system: 1 unit/mL)

2.3. Liver Microsomes Preparation:

2.3.1. Materials:

Table S3. Liver microsomes information.

Species	Product Information	Vendor	Abbrevation
Human	Cat No. 452117 Lot No. 38292	Corning	HLM

2.3.2. Preparation Procedure:

The appropriate concentrations of microsome working solutions were prepared in 100 mM potassium phosphate buffer

2.4. Stop Solution Preparation:

Cold (4°C) acetonitrile (ACN) containing 100 ng/mL tolbutamide and 100 ng/mL labetalol as internal standards (IS) was used as the stop solution

2.5. Assay Procedure:

2.5.1. Using an Apricot automation workstation, 10 μL/well of compound working solution were added to all 96-well reaction plates except the blank (T0, T5, T10, T20, T30, T60, and NCF60)

2.5.2. An Apricot automation workstation was used to add 80 μL/well of microsome solution to all reaction plates (Blank, T0, T5, T10, T20, T30, T60, and NCF60)

2.5.3. All reaction plates containing mixtures of compound and microsomes were pre-incubated at 37°C for 10 minutes

2.5.4. An Apricot automation workstation was used to add 10 μL/well of 100 mM potassium phosphate buffer to reaction plate NCF60

2.5.5. Reaction plate NCF60 was incubated at 37°C, and timer 1 was started

Table S4. NCF60 incubation.

Time Point	Start Time	End Time
NCF60	1:00:00	0:00:00

2.5.6. After pre-incubation, an Apricot automation workstation was used to add 10

μL/well of NADPH regenerating system to every reaction plate except NCF60 (Blank, T0, T5, T10, T20, T30, and T60) to start the reaction

Table 2.3: Final Concentration of Each Component in Incubation Medium

Component	Concentration
Microsome	0.5 mg protein/mL
Test Compound	1 μM
Control Compound	1 μM
MeOH	0.99%
DMSO	0.01%

2.5.7. The reaction plates were incubated at 37°C, and timer 2 was started

Table 2.4: Reaction Plates Incubation

Time Point	Start Time	End Time
Blank	1:00:00	0:00:00
T60	1:00:00	0:00:00
T30	0:30:00	0:00:00
T20	0:20:00	0:00:00
T10	0:10:00	0:00:00
T5	0:05:00	0:00:00
T0	Stop solution was added prior to microsome and NADPH solutions	

2.5.8. An Apricot automation workstation was used to add 300 μL/well of stop solution to each reaction plate at its appropriate end time point to terminate the reaction

2.5.9. Each plate was sealed and shaken for 10 minutes

2.5.10. After shaking, each plate was centrifuged at 4000 rpm and 4°C for 20 minutes

2.5.11. During centrifugation, an Apricot automation workstation was used to add 300 μL/well of HPLC grade water to eight new 96-well plates

2.5.12. After centrifugation, an Apricot automation workstation was used to transfer 100 μL of supernatant from each reaction plate to its corresponding bioanalysis plate

2.5.13. Each bioanalysis plate was sealed and shaken for 10 minutes before LC-MS/MS analysis

3. Data Analysis

3.1. The equation of first-order kinetics was used to calculate T1/2 and CLint(mic) (μL/min/mg):

Equation of first-order kinetics:

$$C_t = C_0 \bullet e^{-k_e \bullet t}$$

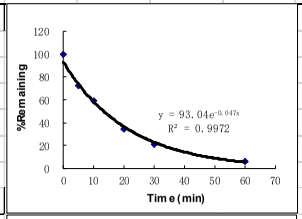
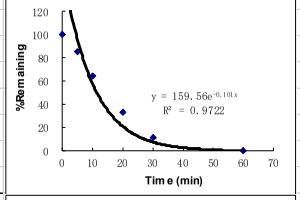
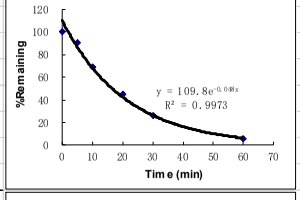
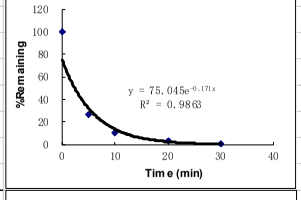
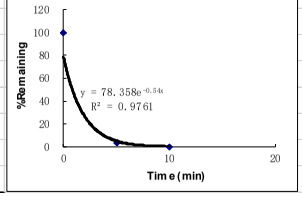
$$\text{when } C_t = \frac{1}{2} C_0,$$

$$T_{1/2} = \frac{Ln2}{k_e} = \frac{0.693}{k_e}$$

$$CL_{int(mic)} = \frac{0.693}{\text{In vitro } T_{1/2}} \bullet \frac{1}{\text{mg / mL microsomal protein in reaction system}}$$

$$CL_{int(liver)} = CL_{int(mic)} \bullet \frac{\text{mg microsomes}}{\text{g liver}} \bullet \frac{\text{g liver}}{\text{kg body weight}}$$

4. Raw Data

Compound ID	Compound & Species	Time (min)	Analyte Peak Area	IS Peak Area	Analyte/IS	%Remaining	Time (min)	%Remaining	Ln(%Remaining)		R^2	k_e (min^{-1})	$T_{1/2}$ (min)	$CL_{int(mic)}$ (μL/min/mg)	Remaining (T=60min)	Remaining (NCF=60min)
Diclofenac	Diclofenac HLM 0.5	Blank	0	129,714	0.000	0.0					0.9972	0.0468	14.8	93.6	5.9%	94.8%
Diclofenac	Diclofenac HLM 0.5	60	4,073	122,629	0.033	5.9	60	5.9	1.8							
Diclofenac	Diclofenac HLM 0.5	30	15,369	125,952	0.122	21.6	30	21.6	3.1							
Diclofenac	Diclofenac HLM 0.5	20	25,182	129,971	0.194	34.3	20	34.3	3.5							
Diclofenac	Diclofenac HLM 0.5	10	42,907	127,947	0.335	59.4	10	59.4	4.1							
Diclofenac	Diclofenac HLM 0.5	5	56,358	138,244	0.408	72.2	5	72.2	4.3							
Diclofenac	Diclofenac HLM 0.5	0	75,507	133,638	0.565	100.0	0	100.0	4.6							
Diclofenac	Diclofenac HLM 0.5	NCF60	69,123	129,039	0.536	94.8										
Propafenone	Propafenone HLM 0.5	Blank	0	130,771	0.000	0.0					0.9722	0.1007	6.9	201.4	0.3%	100.0%
Propafenone	Propafenone HLM 0.5	60	318	126,610	0.003	0.3	60	0.3	-1.3							
Propafenone	Propafenone HLM 0.5	30	13,353	126,122	0.106	11.3	30	11.3	2.4							
Propafenone	Propafenone HLM 0.5	20	41,620	134,055	0.310	33.2	20	33.2	3.5							
Propafenone	Propafenone HLM 0.5	10	82,344	135,453	0.608	65.0	10	65.0	4.2							
Propafenone	Propafenone HLM 0.5	5	105,613	131,783	0.801	85.7	5	85.7	4.5							
Propafenone	Propafenone HLM 0.5	0	127,228	136,116	0.935	100.0	0	100.0	4.6							
Propafenone	Propafenone HLM 0.5	NCF60	114,075	122,013	0.935	100.0										
Testosterone	Testosterone HLM 0.5	Blank	0	158,059	0.000	0.0					0.9973	0.0476	14.6	95.2	6.1%	84.0%
Testosterone	Testosterone HLM 0.5	60	5,315	150,780	0.035	6.1	60	6.1	1.8							
Testosterone	Testosterone HLM 0.5	30	22,469	148,162	0.152	26.4	30	26.4	3.3							
Testosterone	Testosterone HLM 0.5	20	39,844	154,619	0.258	44.9	20	44.9	3.8							
Testosterone	Testosterone HLM 0.5	10	61,993	156,039	0.397	69.3	10	69.3	4.2							
Testosterone	Testosterone HLM 0.5	5	82,633	159,304	0.519	90.4	5	90.4	4.5							
Testosterone	Testosterone HLM 0.5	0	93,108	162,322	0.574	100.0	0	100.0	4.6							
Testosterone	Testosterone HLM 0.5	NCF60	74,835	155,295	0.482	84.0										
11l	I-3LHLM 0.5	Blank	0	93,948	0.000	0.0					0.9863	0.1710	4.1	342.0	0.0%	93.4%
11l	I-3LHLM 0.5	60	0	90,801	0.000	0.0										
11l	I-3LHLM 0.5	30	1,690	105,535	0.016	0.4	30	0.4	-0.8							
11l	I-3LHLM 0.5	20	10,473	98,253	0.107	3.0	20	3.0	1.1							
11l	I-3LHLM 0.5	10	36,647	100,105	0.366	10.2	10	10.2	2.3							
11l	I-3LHLM 0.5	5	97,803	102,693	0.952	26.4	5	26.4	3.3							
11l	I-3LHLM 0.5	0	357,315	99,199	3.602	100.0	0	100.0	4.6							
11l	I-3LHLM 0.5	NCF60	317,067	94,229	3.365	93.4										
PF-74	PF-74HLM 0.5	Blank	0	89,811	0.000	0.0					0.9761	0.5401	1.3	1080.3	0.3%	93.1%
PF-74	PF-74HLM 0.5	60	2,765	83,354	0.033	0.3										
PF-74	PF-74HLM 0.5	30	3,492	94,141	0.037	0.4										
PF-74	PF-74HLM 0.5	20	7,455	98,523	0.076	0.7										
PF-74	PF-74HLM 0.5	10	4,358	92,356	0.047	0.5	10	0.5	-0.8							
PF-74	PF-74HLM 0.5	5	34,911	103,258	0.338	3.2	5	3.2	1.2							
PF-74	PF-74HLM 0.5	0	1,003,927	95,951	10.463	100.0	0	100.0	4.6							
PF-74	PF-74HLM 0.5	NCF60	924,996	94,927	9.744	93.1										

III. Analytical Method for Pharmacokinetics Assay

Analytical Method

Instrument LC-MS/MS (Agilent 6460)

Matrix Rat Plasma

Analyte(s) **11l**

Internal standard(s) TBTM

MS conditions Positive Ion, ESI

MRM Detection

Compounds	Q1	Q3	DP (v)	CE (v)
11l	580.4	296.5	110	15
TBTM	271.4	171.9	90	16

HPLC conditions

Mobile Phase

Mobile Phase A: H_2O

(0.1% Formic Acid)

Mobile Phase B: ACN

Time (min)	Mobile phase B (%)
0.00	5

1.00	95
2.00	95
2.01	5
4.00	5

Column: Agilent Eclipse Plus C18(2.1×50mm，5μm)

Flow rate: 0.3 mL·min^{-1}

MS parameters

Gas flow (L/min)	11
Nebulizer (psi)	45
Gas temp (°C)	300
Capillary (V)	4000

Sample preparation	To each 40μL plasma, 160 μL ACN which contains 20 ng·mL^{-1}. TBTM were added. After a thorough vortex for 1 min, the mixture was centrifuged at 13000rpm for 5 min at 8°C, and a10 μL aliquot of the supernatant was injected into the LC-MS/MS for analysis.

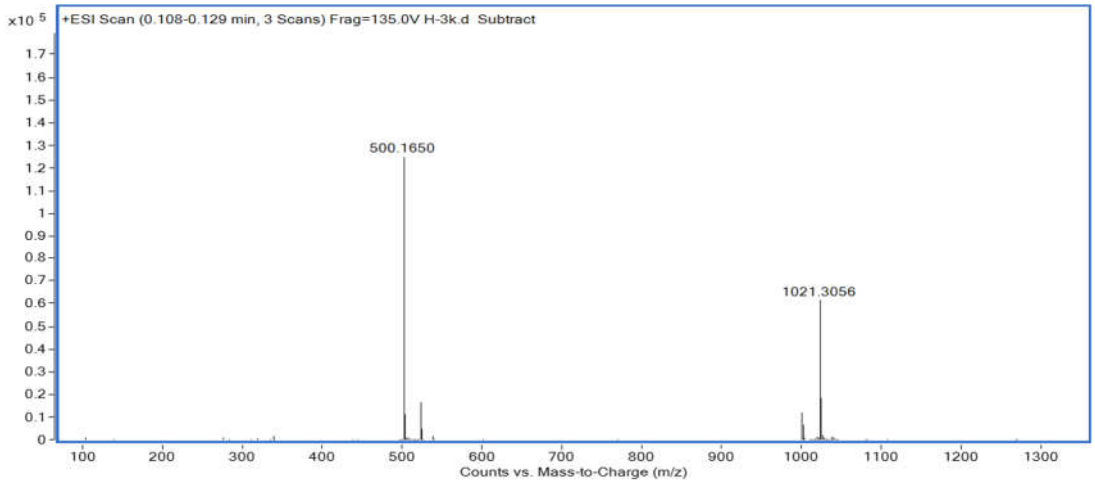
IV. References

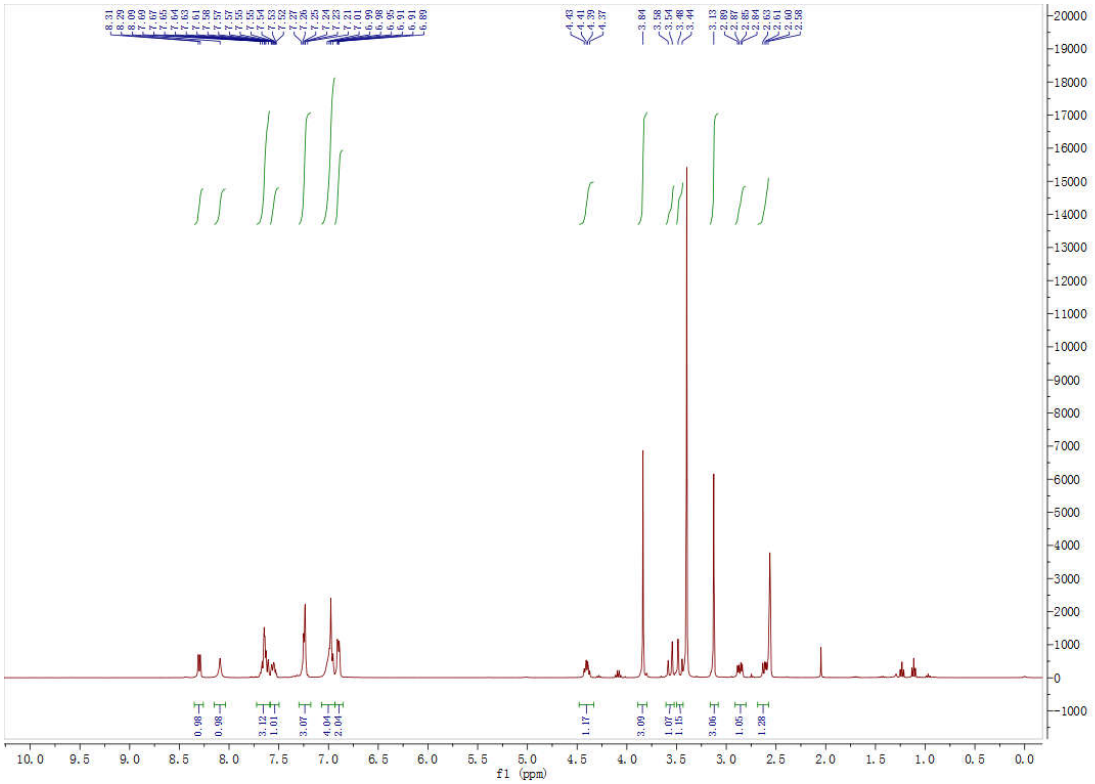
S1. Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. Autodock4 and AutoDockTools4: automated docking with selective receptor flexiblity. *J Comput Chem.* **2009,** 16: 2785-2791.

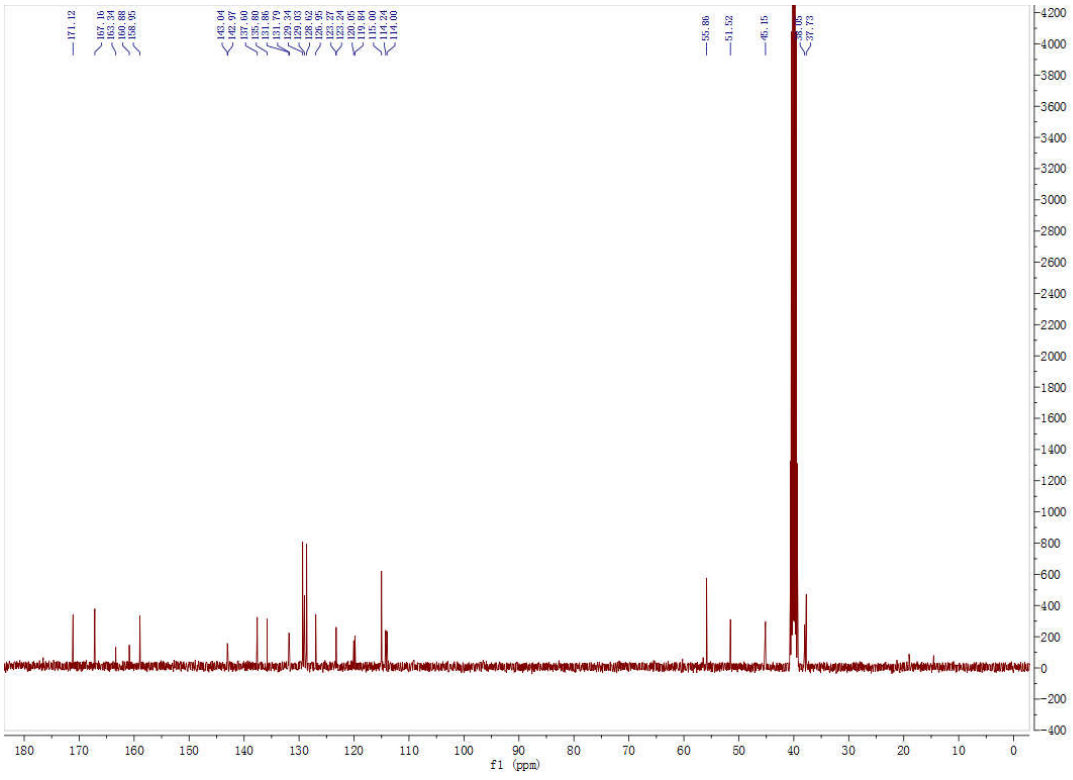
S2. Wu, G.; Zalloum, W. A.; Meuser, M. E.; Jing, L.; Kang, D.; Chen, C. H.; Tian, Y.; Zhang, F.; Cocklin, S.; Lee, K. H.; Liu, X.; Zhan, P. Discovery of phenylalanine derivatives as potent HIV-1 capsid inhibitors from click chemistry-based compound library. *Eur J Med Chem.* **2018**, 158, 478-492.

V. HRMS, ^{1}H-NMR and ^{13}C-NMR spectra for representative compounds

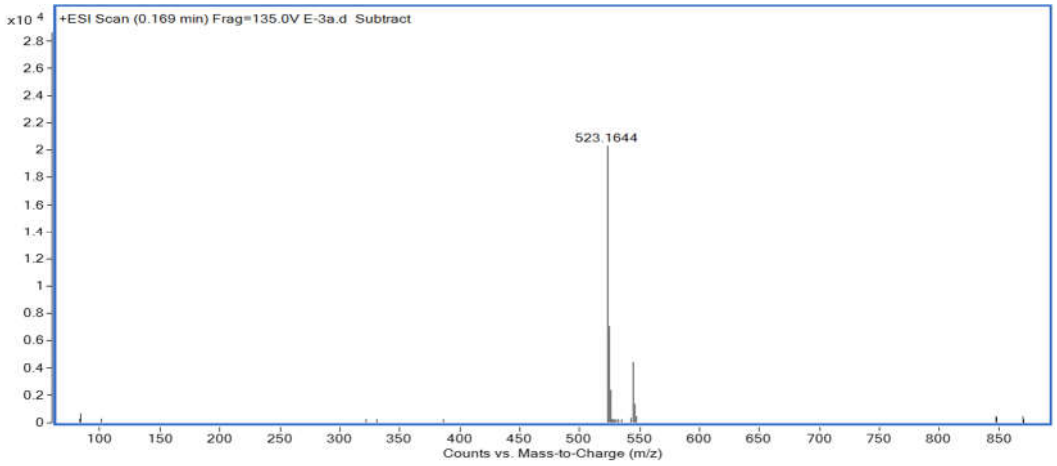
1. HRMS, ^{1}H-NMR and ^{13}C-NMR spectra for 6k

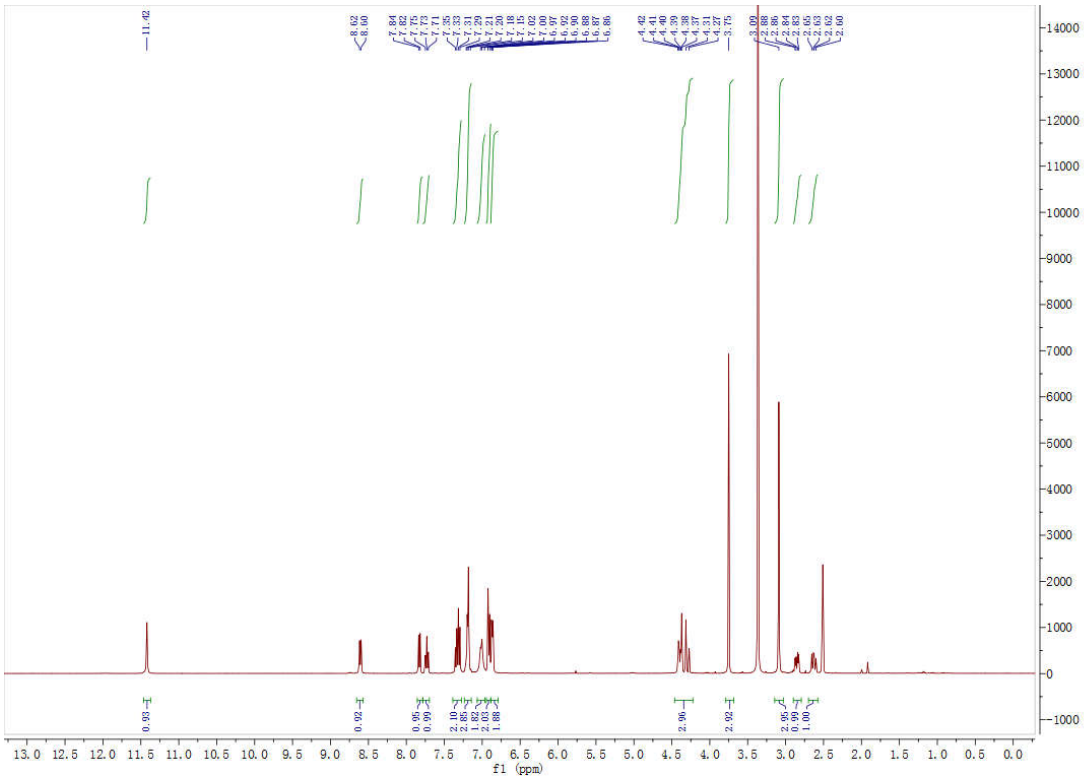


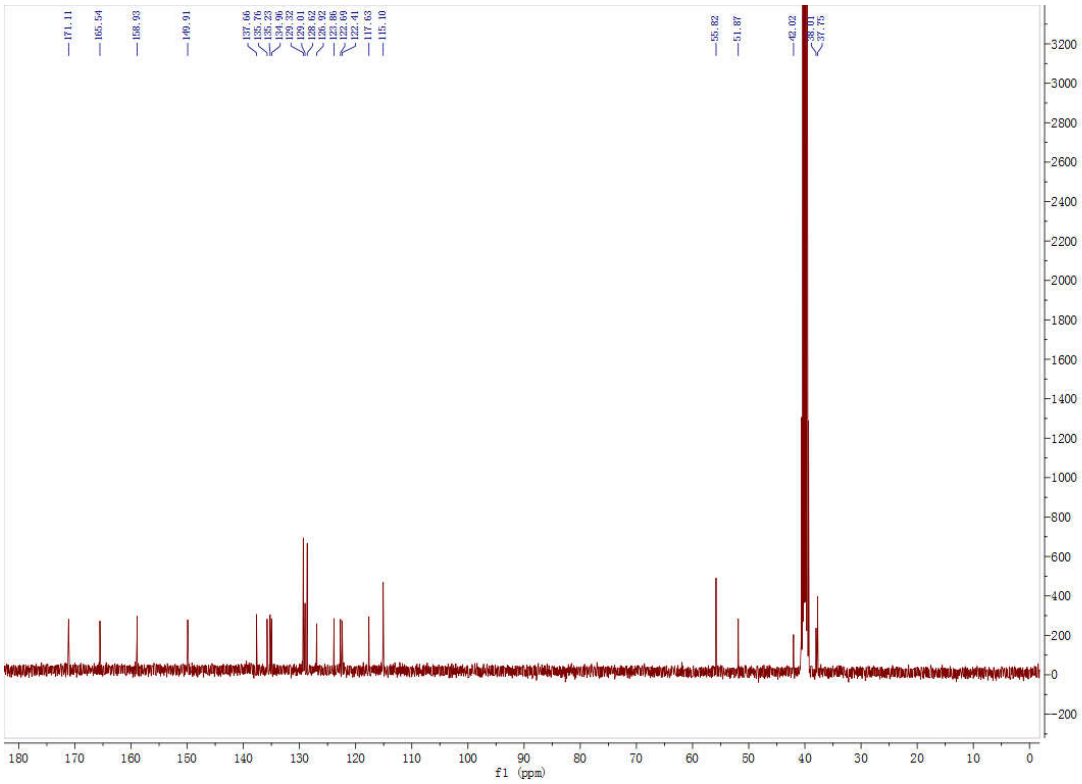




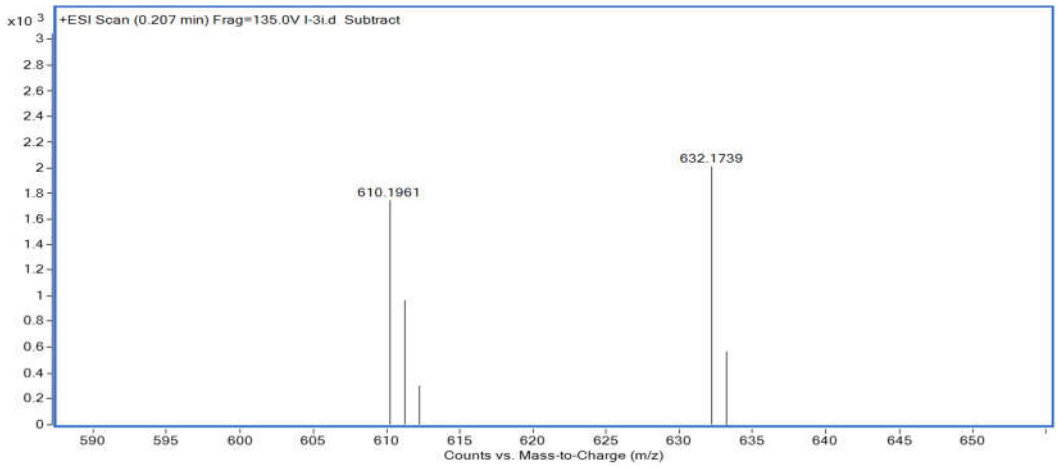
2. HRMS, ^{1}H-NMR and ^{13}C-NMR spectra for 8a

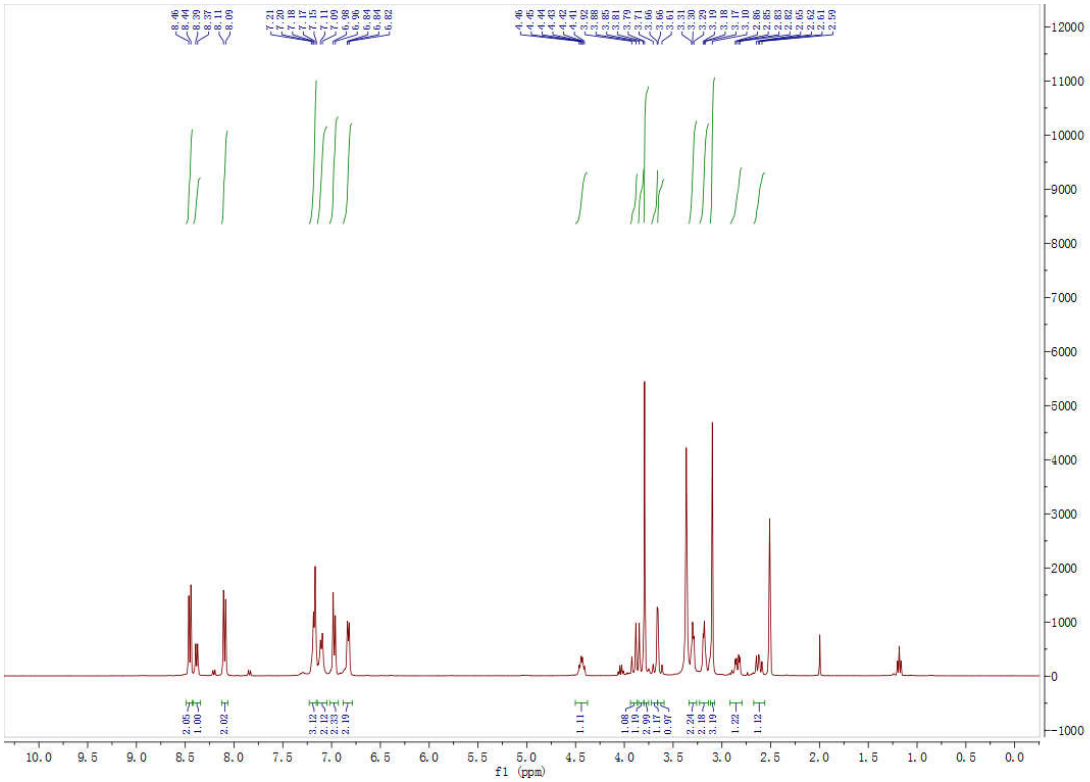

x10 4
+ESI Scan (0.169 min) Frag=135.0V E-3a.d Subtract
523.1644
Counts vs. Mass-to-Charge (m/z)

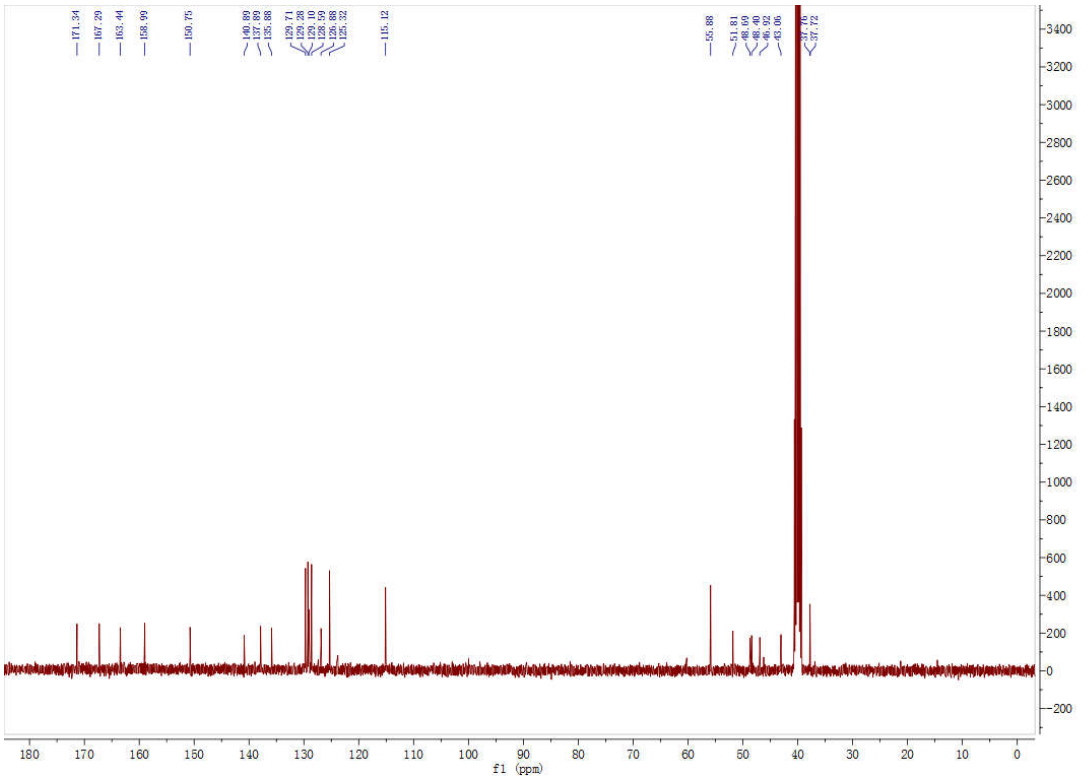

f1 (ppm)



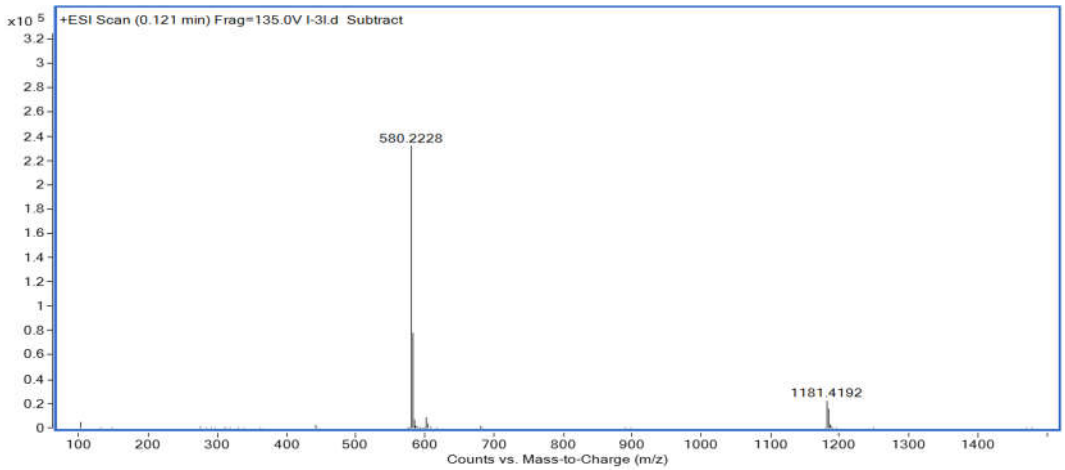
3. HRMS, ^{1}H-NMR and ^{13}C-NMR spectra for 11i

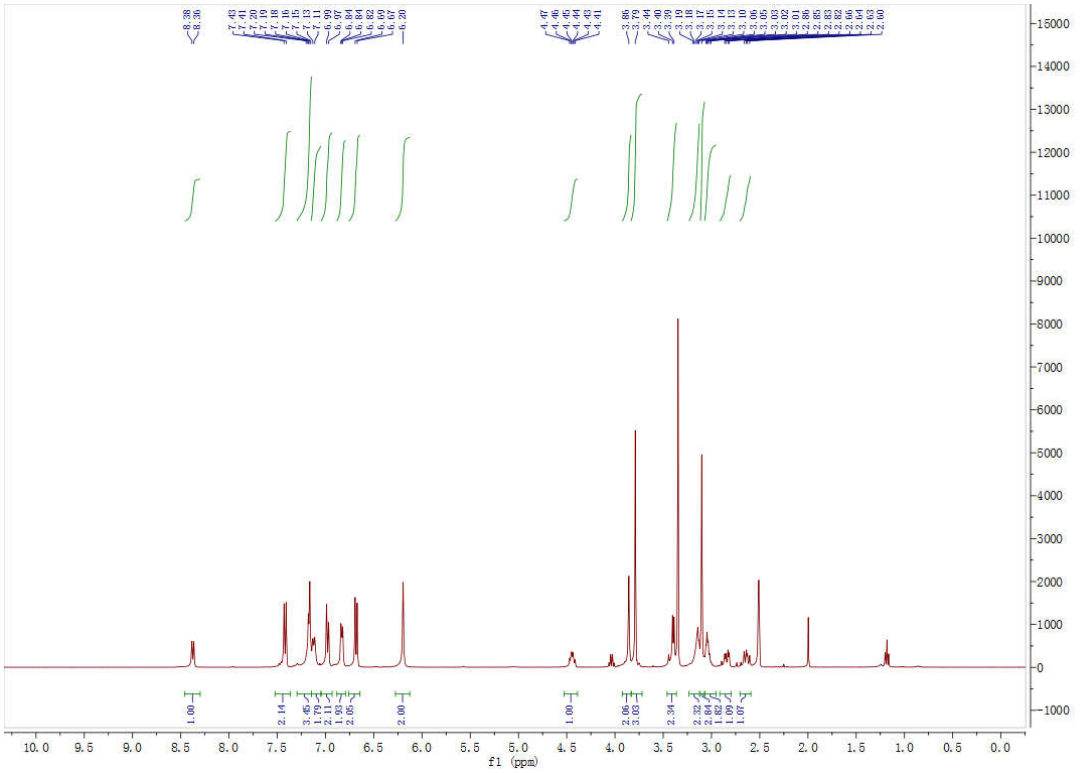


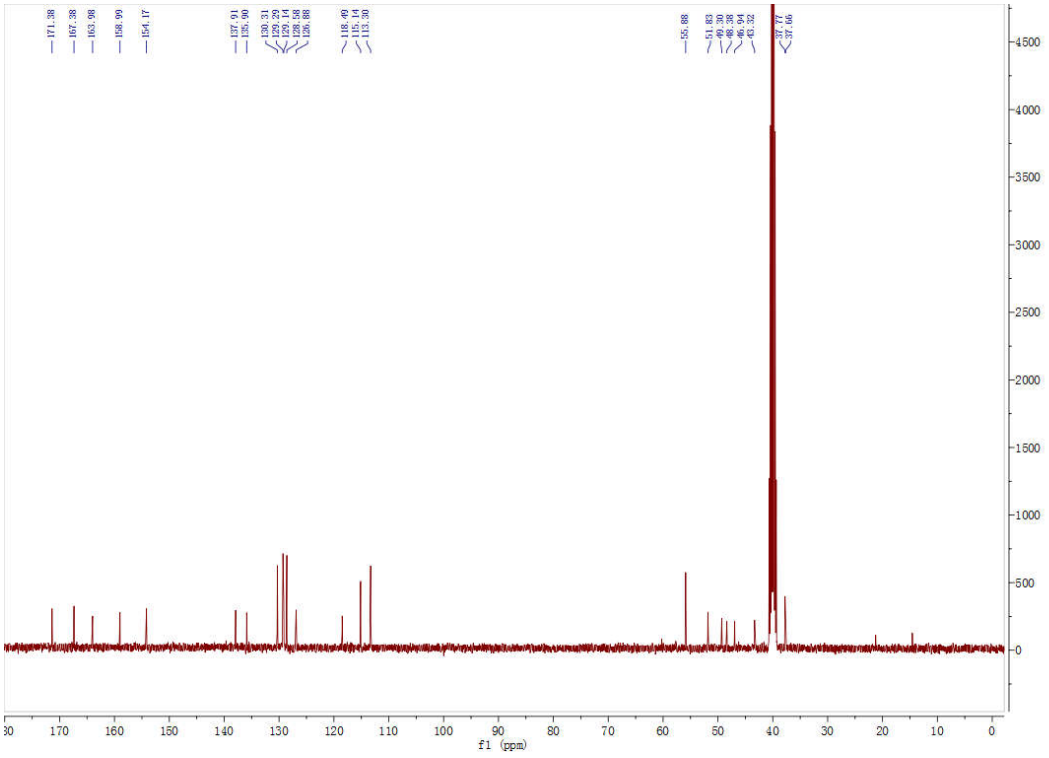




4. HRMS, ^{1}H-NMR and ^{13}C-NMR spectra for 11l


x10 5
+ESI Scan (0.121 min) Frag=135.0V I-3l.d Subtract
580.2228
1181.4192
Counts vs. Mass-to-Charge (m/z)


f1 (ppm)


171.38
167.38
163.98
158.99
154.17
137.91
135.90
130.31
129.29
129.14
128.94
126.88
118.49
115.14
113.30
55.88
51.83
49.30
48.98
46.94
43.32
37.77
37.66
4500
4000
3500
3000
2500
2000
1500
1000
500
0
30
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)