Supporting Information - Discovery of Small-Molecule Antagonists of the PWWP Domain of NSD2 - Supporting Information

Renato Ferreira de Freitas[‡,|,#], Yanli Liu[‡,||,#], Magdalena M. Szewczyk[‡], Naimee Mehta[†], Fengling Li[‡], David McLeod[††], Carlos Zepeda-Velázquez[††], David Dilworth[‡], Ronan P. Hanley[†], Elisa Gibson[‡], Peter J. Brown[‡], Rima Al-Awar[††], Lindsey I. James[†], Cheryl H. Arrowsmith[‡,&], Dalia Barsyte-Lovejoy[‡,!], Jinrong Min[‡,%], Masoud Vedadi[‡,!], Matthieu Schapira[‡,!,*], Abdellah Allali-Hassani[‡,|||,#,*]

[‡] Structural Genomics Consortium, University of Toronto, Toronto, Ontario M5G 1L7, Canada

[†] Center for Integrative Chemical Biology and Drug Discovery, Division of Chemical Biology and Medicinal Chemistry, UNC Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, United States

[††] Drug Discovery Program, Ontario Institute for Cancer Research, Toronto, ON M5G 0A3, Canada

[&] Princess Margaret Cancer Centre and Department of Medical Biophysics, University of Toronto, Toronto, ON, M5G 2M9, Canada

[%] Department of Physiology, University of Toronto, Toronto, ON M5S 1A8, Canada

[!] Department of Pharmacology & Toxicology, University of Toronto, Toronto, Ontario M5S 1A8, Canada

[|] present address: Centro de Ciências Naturais e Humanas, Universidade Federal do ABC, Rua Arcturus 3, São Bernardo do Campo, SP 09606-070, Brazil.

[||] present address: College of Pharmaceutical Sciences, Soochow University, Suzhou, Jiangsu 215123, China

[|||] present address: Incyte, 1801 Augustine Cut-Off, Wilmington, DE 19803, USA

[#] equal contributors

[*] To whom correspondence should be addressed

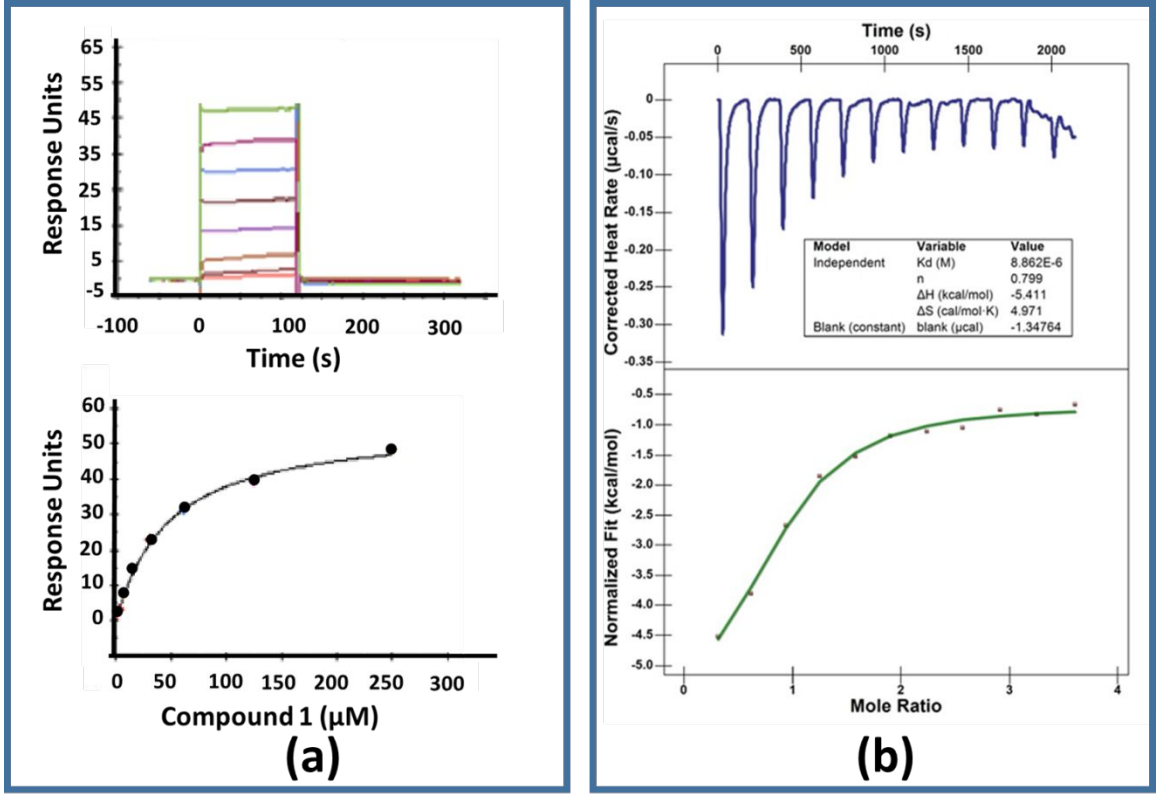


Figure S1. (a) Compound **1** binds NSD2-PWWP1 with a K_d of 41 ± 8 μM in SPR experiments conducted in triplicate. SPR-sensorgram (upper pane) for the interaction of **1** with NSD2-PWWP1 domain. Multi-cycle kinetics was used with concentrations from 2 μM to 250 μM (dilution factor of 0.5 was used to yield 8 concentrations). 90 s contact time and 120 s dissociation time at 100 μL/min were used. The steady state values were determined and plotted as a function of the concentration (lower pane). A single binding site model was fitted to the data to calculate K_d. (b) ITC results showing raw data after integration baseline correction (upper pane) and integrated data and regression (lower pane). n is the number of molecules per binding site, K_d is the association constant, ΔH is the change in enthalpy, and ΔS is the change in entropy.

Table S2. Thermal shift in differential static light scattering (DSLS) induced by **1** against seven PWWP domains.

Protein	**ΔT_{agg} at 400 μM (°C)**
BRPF1-PWWP	-0.5
DNMT3b-PWWP	-1.2
MSH6-PWWP	-1.4
NSD2-PWWP1	3.9
NSD3-PWWP1	-0.3
ZMYND11-PWWP	-1.2
ZCWPW1-PWWP	-1.4

Characterization of Compound 1

The original sample was obtained from Enamine (Catalog Number Z1483746373) and assumed to be a racemic mixture. Attempted separation of all 4 diastereoisomers by chiral SFC resulted in only two

peaks in the SFC chromatogram. Peaks showed identical NMR spectra but exhibited equal and opposite optical rotations. These must be enantiomers and thus the original sample must have been made by a diasteroselective route (Figure S2).

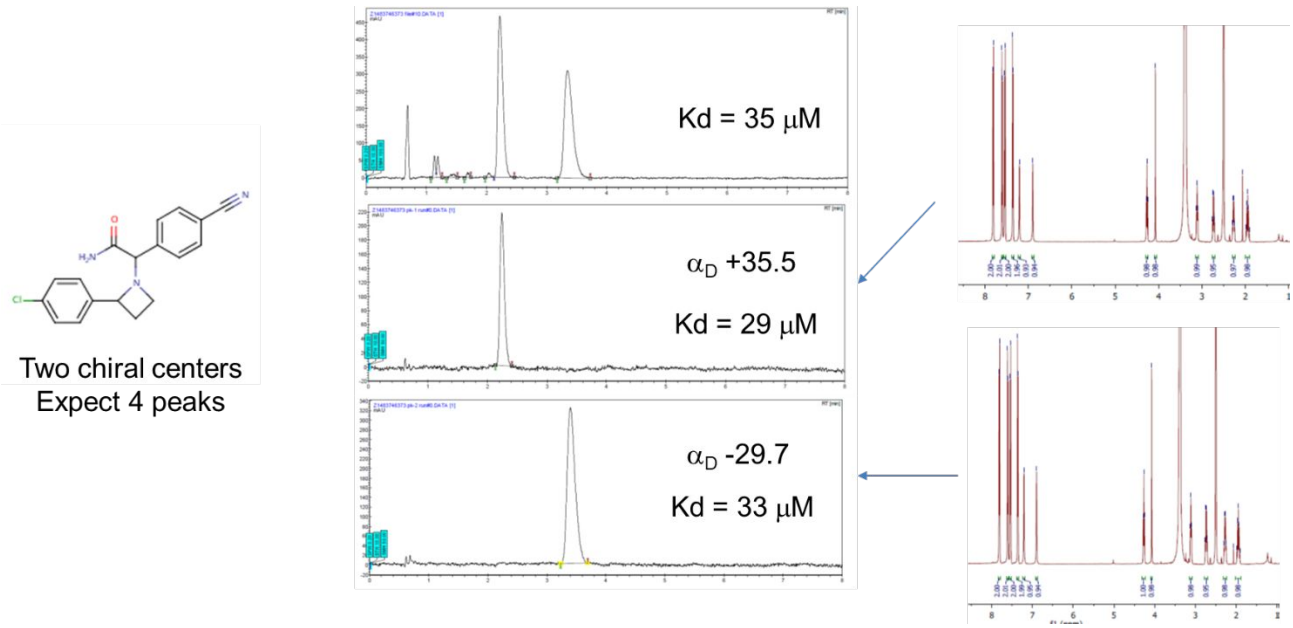


Figure S2. Separation of enantiomers from original sample by chiral SFC.

A search of the literature found a paper describing this diastereoselective synthesis from a substituted boronic acid and cyclic amine (Nanda K. K. and Trotter B. W. **Tet. Lett.** (2005) 2025-28). In an effort to determine the activity of the other diasteroisomers, a non-diastereoselective route was developed involving azetidine displacement of an a-bromoester (Figure S3).

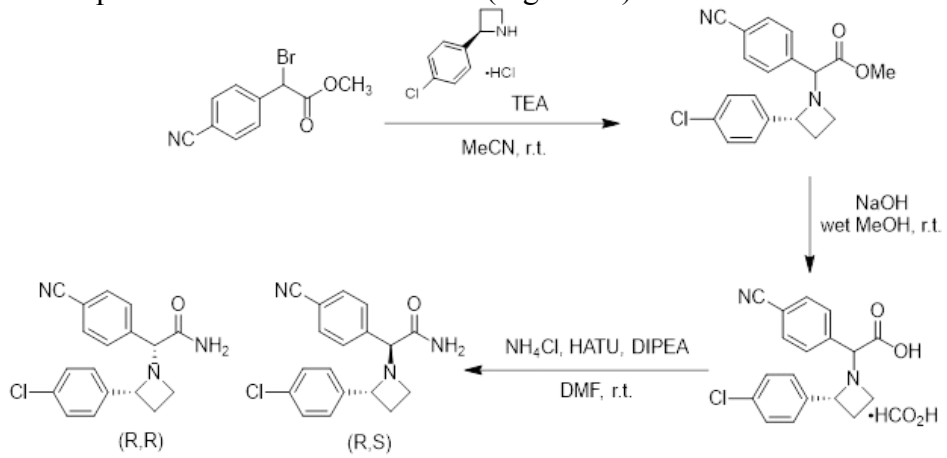


Figure S3. Synthesis from (*R*)-azetidine gave two diastereoisomers. Similar results were obtained starting from (*S*)-azetidine to give the *(S,S)* and (*S,R*) diastereoisomers.

In this case, all four isomers were detected by chiral SFC, however, the diastereoisomers not found in the original material were inactive. The assignment of stereochemistry was based on the diastereoselective product in the original paper which was confirmed by X-ray crystallography (Figure S4).

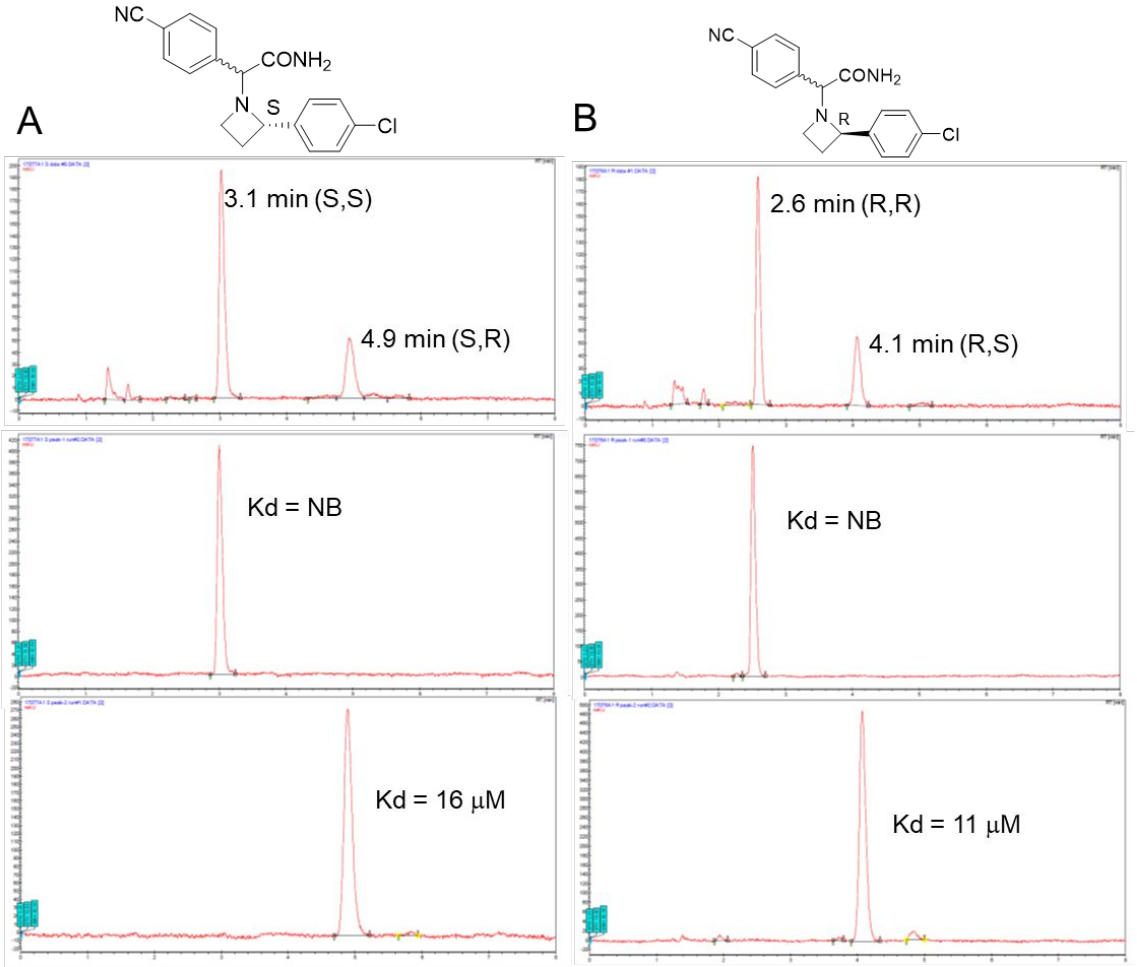


Figure S4. Chiral preparative SFC separation of diastereoisomers. A) Reaction products from (*S*)-azetidine gave peaks at 3.1 and 4.9 mins. B) Products from (*R*)-azetidine gave peaks at 2.6 and 4.1 mins. The original enantiomers correspond to the peaks at 4.1 and 4.9 mins. Lower panels indicate the analytical chiral SFC analysis of collected peaks.

.

Experimental Procedures and Characterization data

Separation of commercial material

The commercially available material (30 mg; Enamine Z1483746373) was separated by chiral SFC using an IA column (2 x 25 cm) using 30% MeOH/CO_2 as eluant at a flowrate of 60 mL.min^{-1} and 100 bar pressure. Detection by UV at 220 nm. Injection volume was 1.5 mL of 3 mg.mL^{-1} in MeOH. Analytical chiral SFC (Fig. 1) used IA column (0.46 x 25 cm) with 40% MeOH/CO_2 as eluant at a flowrate of 3 mL.min^{-1} and 100 bar pressure. Detection by UV at 220 nm.

2R-(2-(4-chlorophenyl)azetidin-1-yl)-2S-(4-cyanophenyl)acetamide (13 mg). $[\alpha]_d^{23}$ +35.5° (c=0.205; MeOH). ^{1}H NMR (500 MHz, DMSO) δ 7.81 (d, *J* = 8.2 Hz, 2H), 7.60 (d, *J* = 8.2 Hz, 2H), 7.54 (d, *J* = 8.4 Hz, 2H), 7.36 (d, *J* = 8.4 Hz, 2H), 7.21 (s, 1H), 6.90 (s, 1H), 4.27 (t, *J* = 8.1 Hz, 1H), 4.07 (s, 1H), 3.11 (t, *J* = 6.5 Hz, 1H), 2.73 (dd, *J* = 16.1, 8.0 Hz, 1H), 2.27 (q, *J* = 8.0 Hz, 1H), 1.95 (p, *J* = 9.0 Hz, 1H).

2S-(2-(4-chlorophenyl)azetidin-1-yl)-2R-(4-cyanophenyl)acetamide (12 mg). $[\alpha]_d^{23}$ -29.7° (c=0.337; MeOH). ^{1}H NMR (500 MHz, DMSO) δ 7.81 (d, *J* = 8.2 Hz, 2H), 7.60 (d, *J* = 8.2 Hz, 2H), 7.54 (d, *J* = 8.3 Hz, 2H), 7.36 (d, *J* = 8.3 Hz, 2H), 7.21 (s, 1H), 6.90 (s, 1H), 4.27 (t, *J* = 8.1 Hz, 1H), 4.07 (s, 1H), 3.11 (t, *J* = 6.5 Hz, 1H), 2.73 (dd, *J* = 16.1, 8.0 Hz, 1H), 2.27 (q, *J* = 8.0 Hz, 1H), 1.95 (p, *J* = 9.0 Hz, 1H).

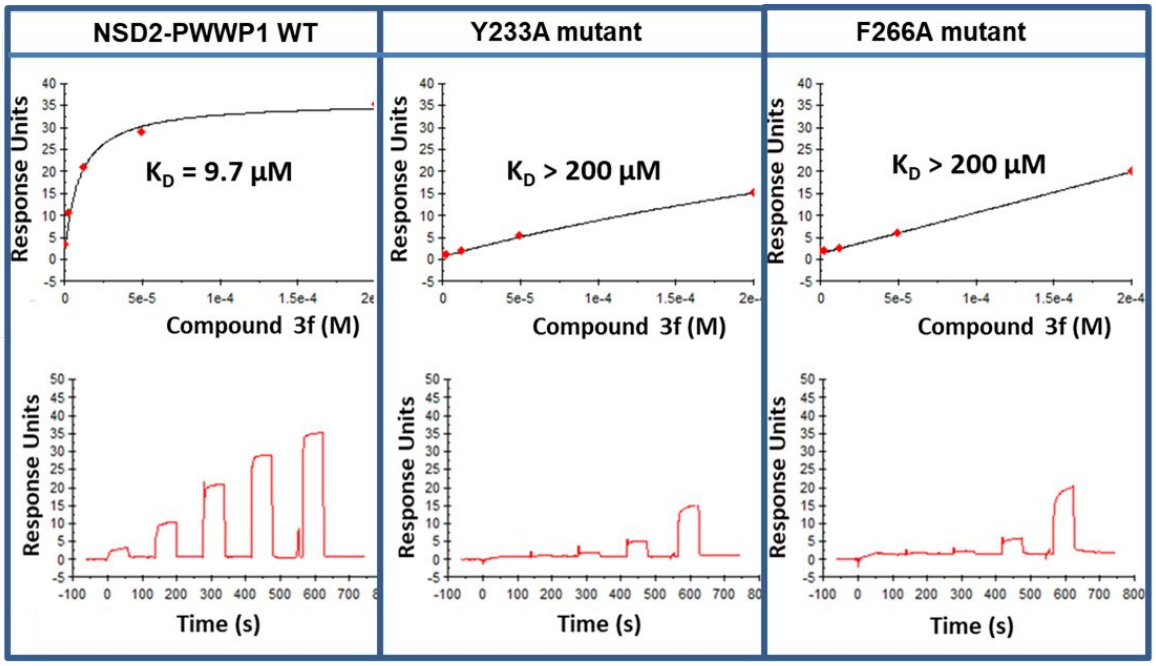


Figure S5: compound **3f** binds wild-type NSD2-PWWP1 but not Y233A or F266A mutants in SPR experiments

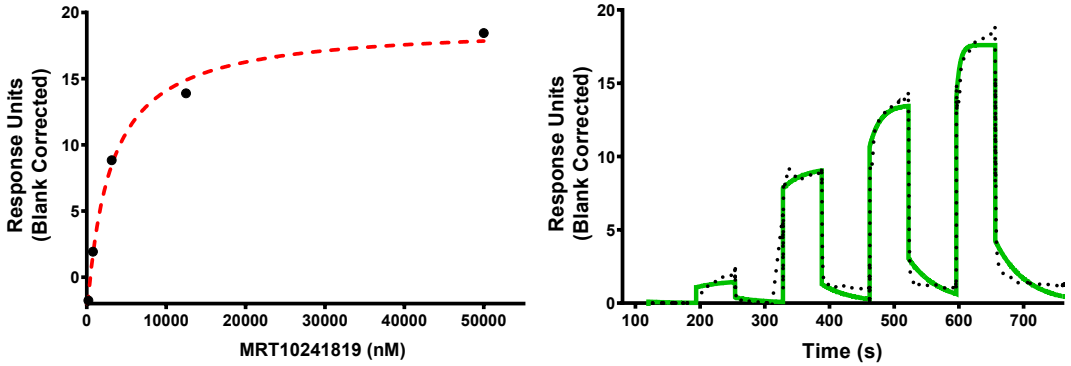


Figure S6. Compound 3f binds NSD2-PWWP1 with a K_d of 3.42 ± 0.45 µM in a SPR experiments conducted in triplicate with 0.5% DMSO concentration. Biotinylated NSD2-PWWP1 domain (208-368) was immobilized on the flow cell of an SA sensor chip in 1x HBS-EP buffer, yielding 5000 RU. The biotinylated RBBP5 (2-538) was immobilized on another flow cell of SA chip, yielding 4300 RU as a negative control. Using the same buffer with 0.5% DMSO and single cycle kinetics with 60 s contact time and a dissociation time of 120s at a flow rate of 75 µL/min. The compound was tested at 50 µM as the highest concentration and dilution factor of 0.25 was used to yield 5 concentrations.

Table S3. Crystallography data and refinement statistics

	NSD2 + **3f**
PDB Code	6UE6
Data collection	
Space group	$P2_12_12_1$

Cell dimensions	
a, *b*, *c* (Å)	69.2, 70.3, 228.8
α, β, γ (°)	90.0,90.0,90.0
Resolution (Å) (highest resolution shell)	49.33-2.40(2.49-2.40)
Measured reflections	287140
Unique reflections	44525
R_{merge}	8.4(0.997)
$I/\sigma I$	14.0(1.9)
Completeness(%)	99.7(97.5)
Redundancy	6.4(5.2)
Refinement	
Resolution (Å)	45.3-2.40
No. reflections (test set)	44447(2275)
R_{work}/ R_{free} (%)	23.6/25.2
No. atoms	
Protein	7136
Compound	160
B-factors (Å^2)	
Protein	61.4
Compound	47.5
RMSD	
Bond lengths (Å)	0.010
Bond angles (º)	0.99
Ramachandran plot % residues	
Favored	99.5
Additional allowed	0.5
Generously allowed	0
Disallowed	0

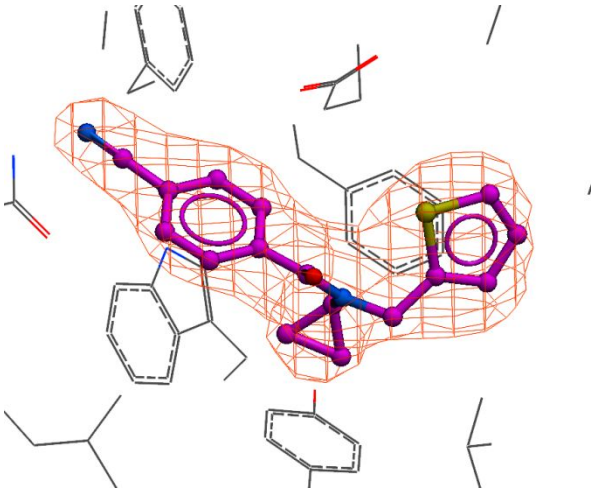
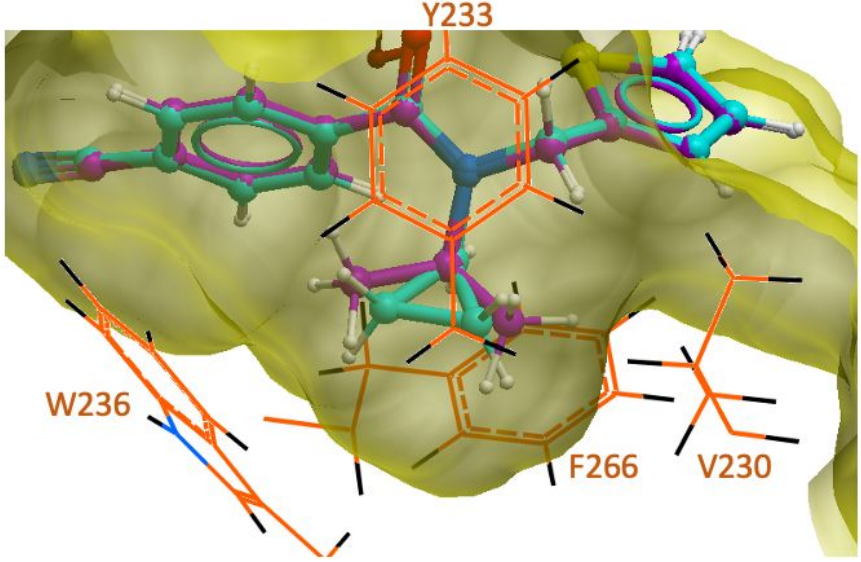
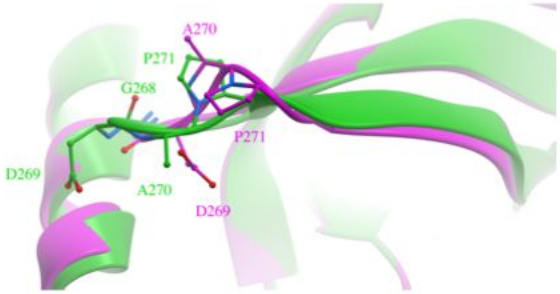


Figure S7: One sigma electron density unambiguously defines the ligand binding pose

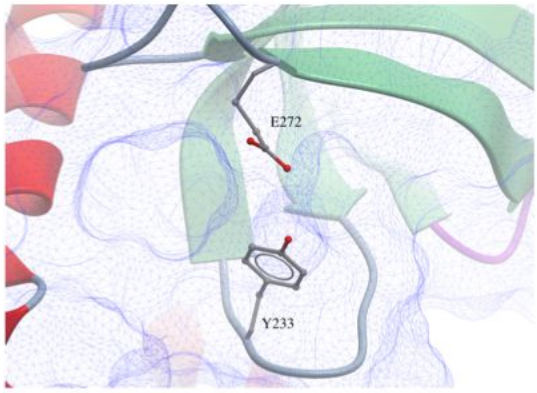


Ligand interaction VdW Energy
compound 3f (cyclopropyl) : -27.5 kcal/mol
Compound 3c (isopropyl): 1.1 kcal/mol

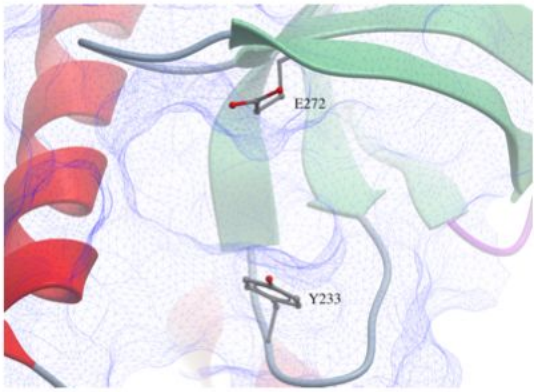
Figure S8. A computational model indicates steric clashes with the isopropyl of compound **3c**. The cyclopropyl of **3f** was replaced with an isopropyl in ICM (Molsoft, San Diego). The isopropyl methyl groups are more distant resulting in increased bulk (the isopropyl geometry is absolutely similar to the one found in ligand bk1, PDB code 3I7B). The energy of the modified flexible ligand was locally minimized in the internal coordinate space, while conserved atoms were tethered to their original position (with a tether weight tzWeight=200). The Van der Waals energy of the bound ligand was calculated in ICM (force-field: protein: ECEPP/3, ligand: mmff).



(a)



(b)



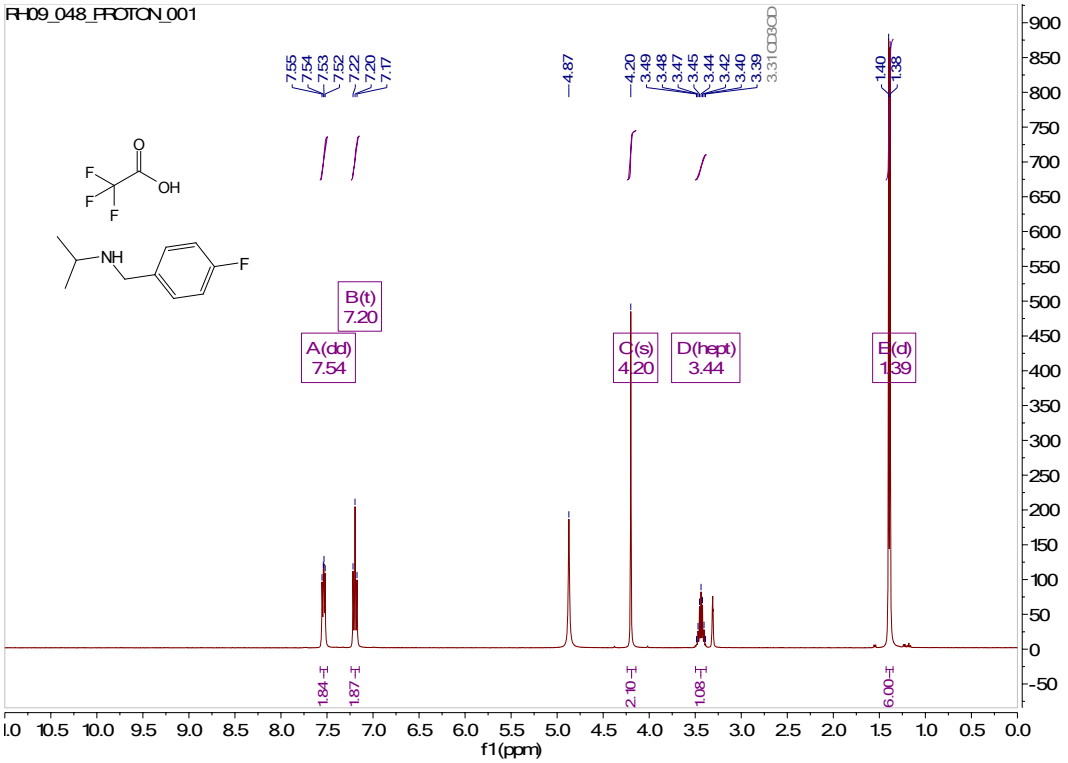
(c)

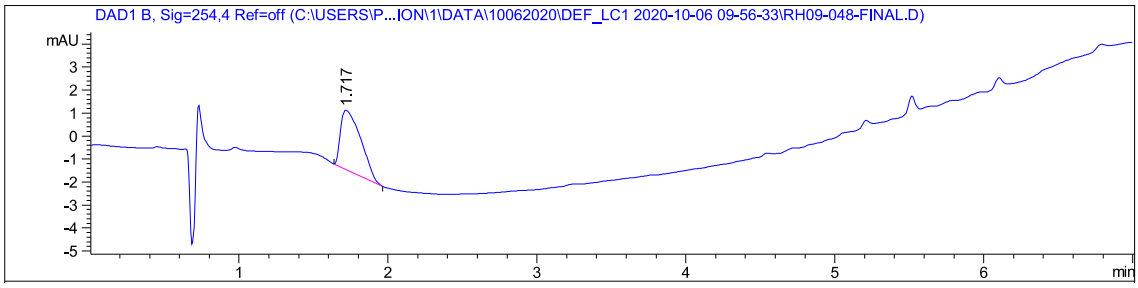
Figure S9. (a) Comparison between the apo (PBD ID: 5VC8, magenta) and bound (green) conformation of the NSD2-PWWP1. The loop residues G268, D269, A270, and P271 connecting the β3 and β4 strands

display different conformations in the two structures. (b) In the apo structure the residues E272 and Y233 are closing the pocket. (c) When the ligand binds these residues move away opening the pocket.

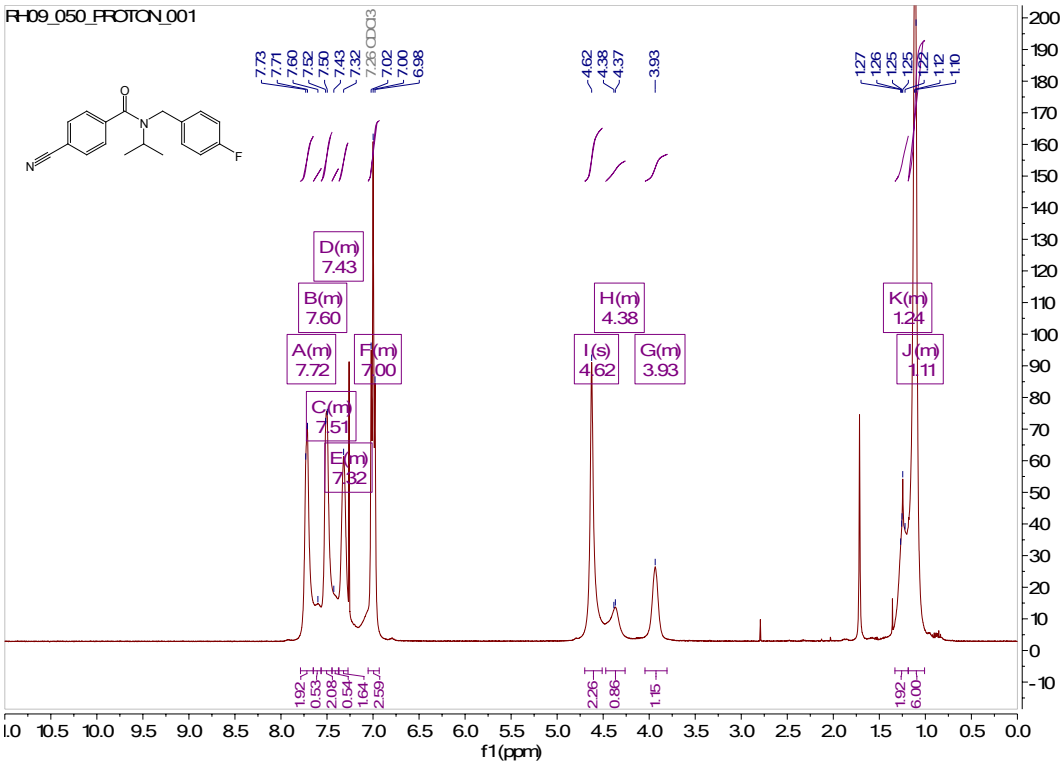
^{1}H NMR Spectra

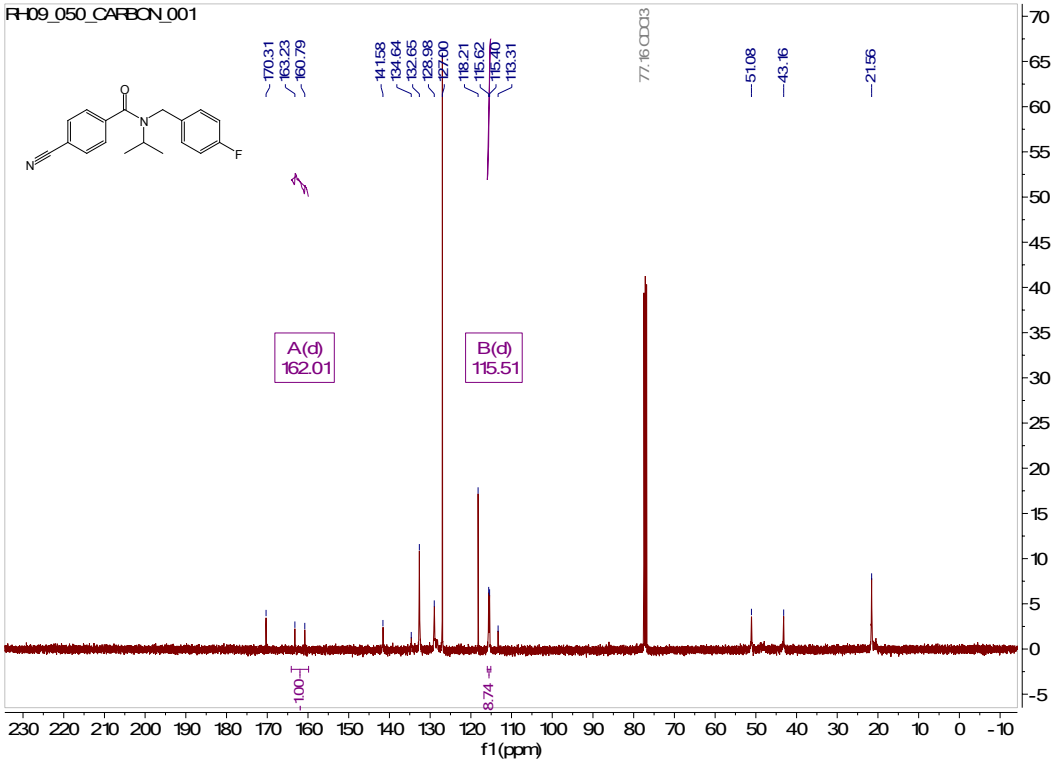
N-isopropyl-4-fluorobenzylamine

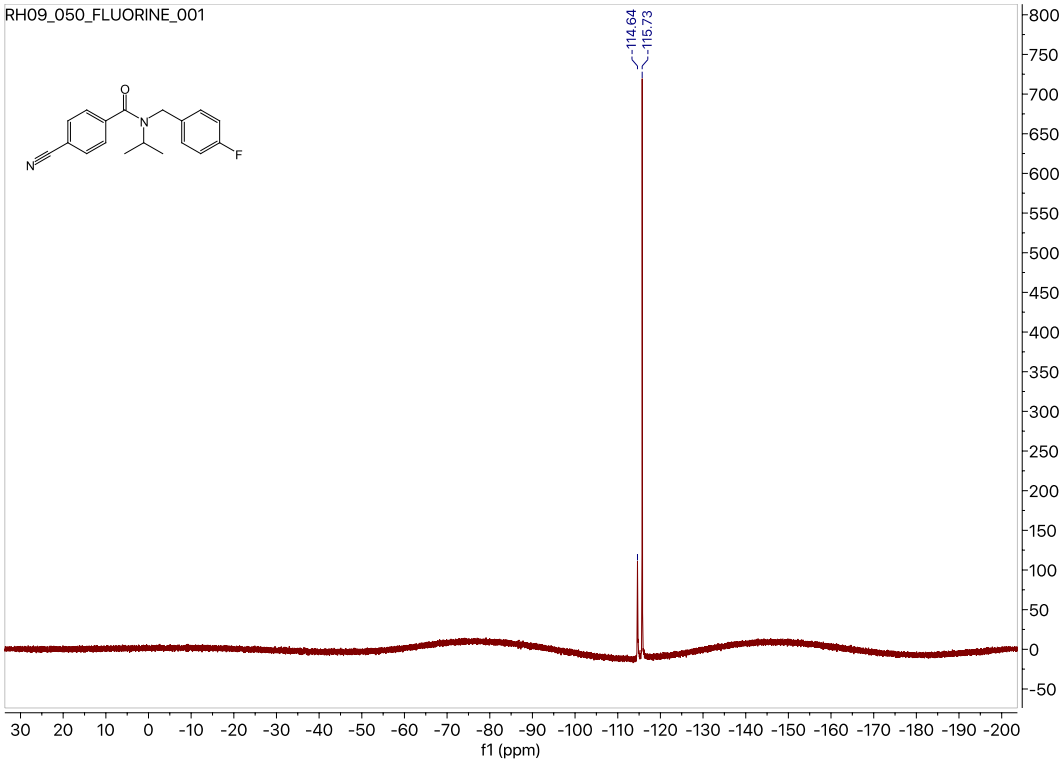


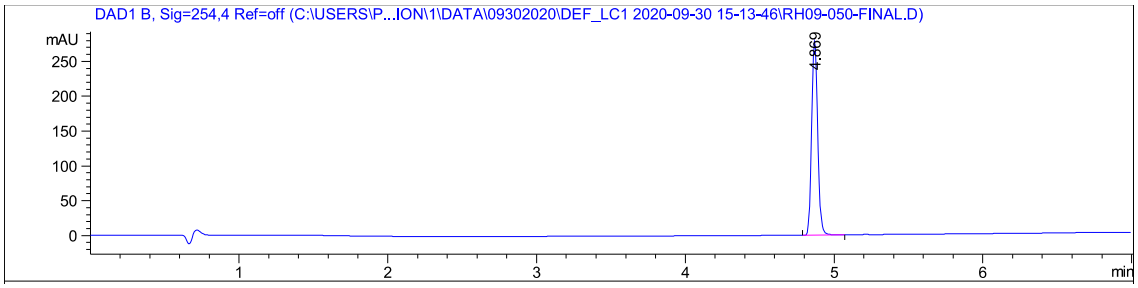


Compound 3c

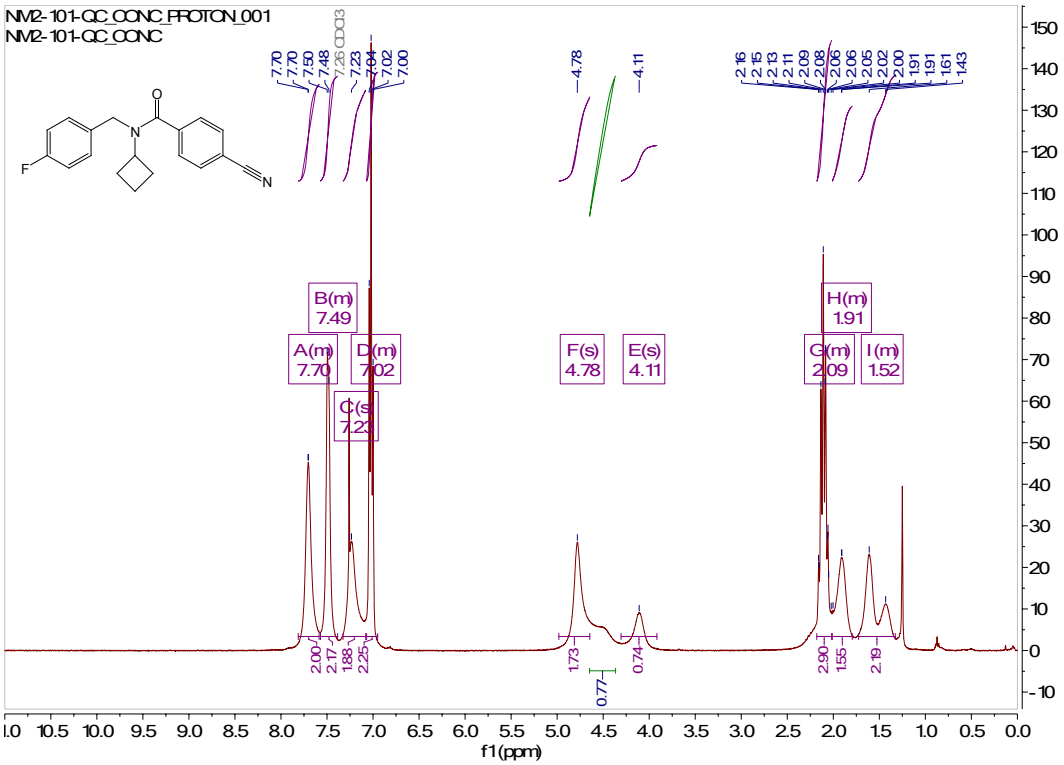

RH09_050_PROTON_001
7.73
7.71
7.60
7.52
7.50
7.43
7.32
7.26 CDCl3
7.02
7.00
6.98
4.62
4.38
4.37
3.93
1.27
1.26
1.25
1.25
1.22
1.12
1.10
D(m) 7.43
B(m) 7.60
A(m) 7.72
F(m) 7.00
C(m) 7.51
E(m) 7.32
H(m) 4.38
I(s) 4.62
G(m) 3.93
K(m) 1.24
J(m) 1.11
1.92
0.53
2.08
0.54
1.64
2.59
2.26
0.86
1.15
1.92
6.00
f1(ppm)

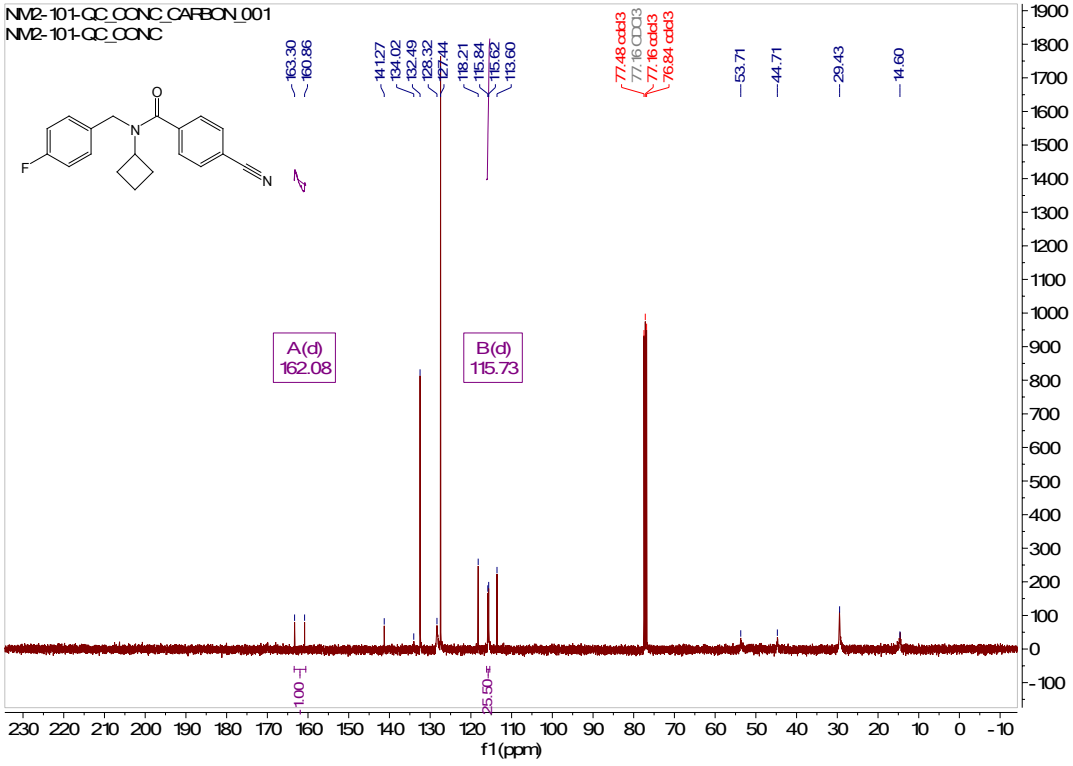

RH09_050_CARBON_001
170.31
163.23
160.79
141.58
134.64
132.65
128.98
127.00
118.21
115.62
115.40
113.31
77.16 CDCl3
51.08
43.16
21.56
A(d) 162.01
B(d) 115.51
1.00
8.74
f1(ppm)

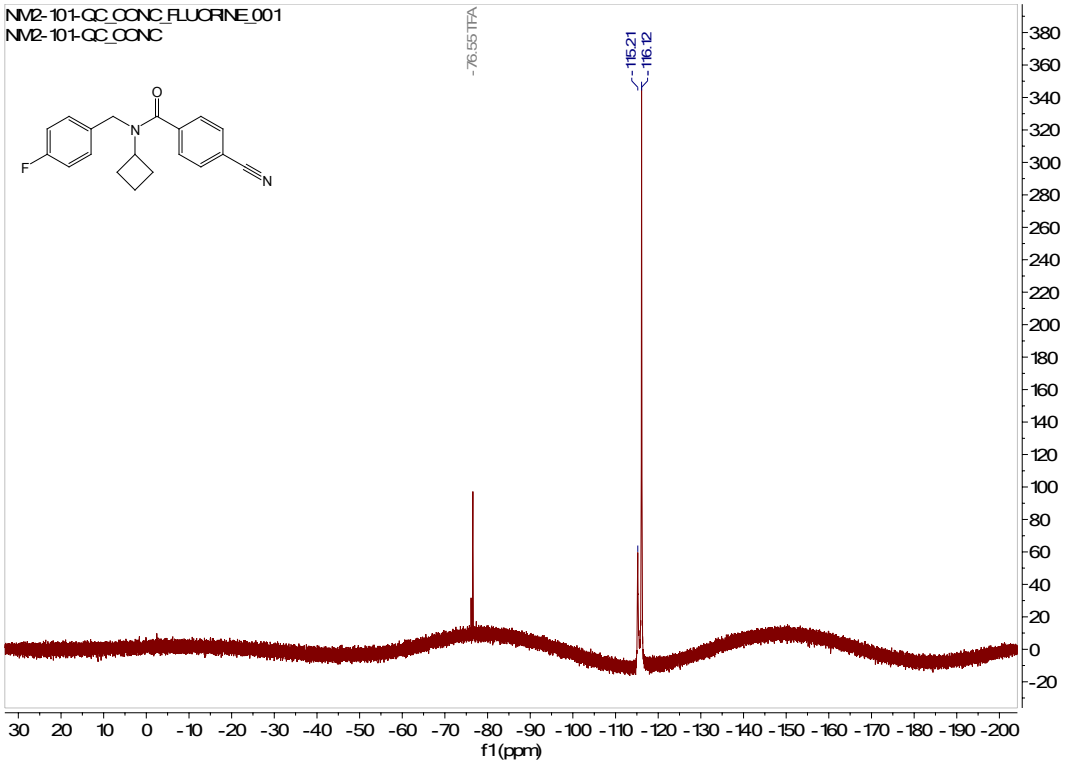

RH09_050_FLUORINE_001
-114.64
-115.73
f1 (ppm)

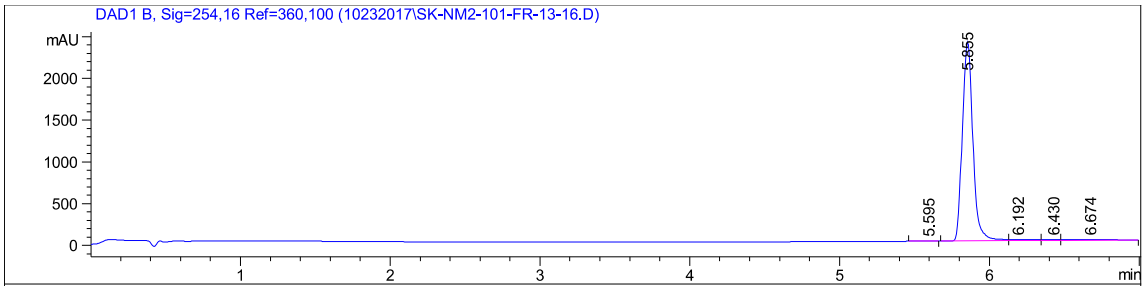

DAD1 B, Sig=254,4 Ref=off (C:\USERS\P...ION\1\DATA\09302020\DEF_LC1 2020-09-30 15-13-46\RH09-050-FINAL.D)
mAU
4.869
min

Compound 3d

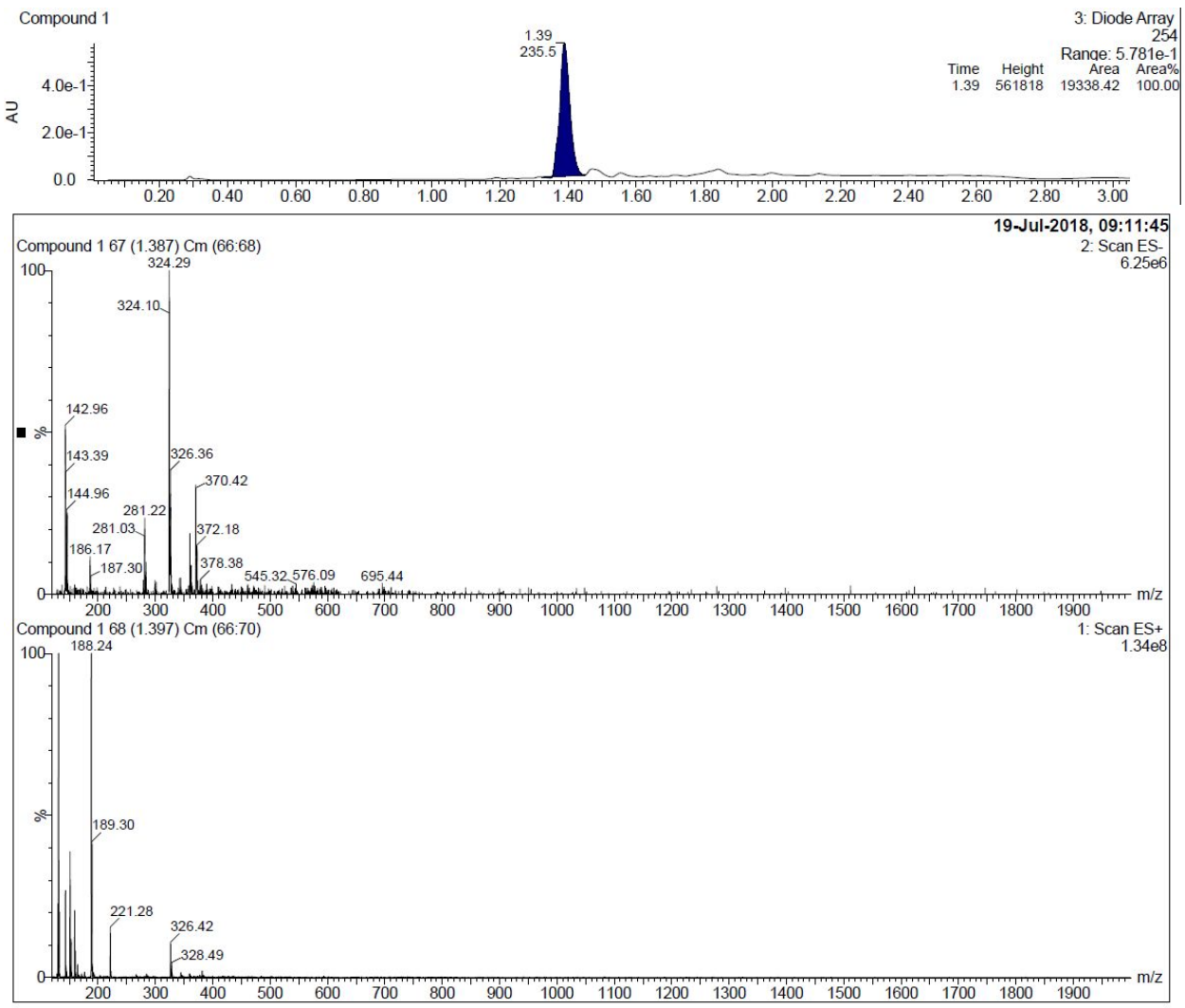

NM2-101-QC_CONC_PROTON_001
NM2-101-QC_CONC
B(m) 7.49
A(m) 7.70
D(m) 7.02
C(s) 7.23
F(s) 4.78
E(s) 4.11
H(m) 1.91
G(m) 2.09
I(m) 1.52
f1(ppm)

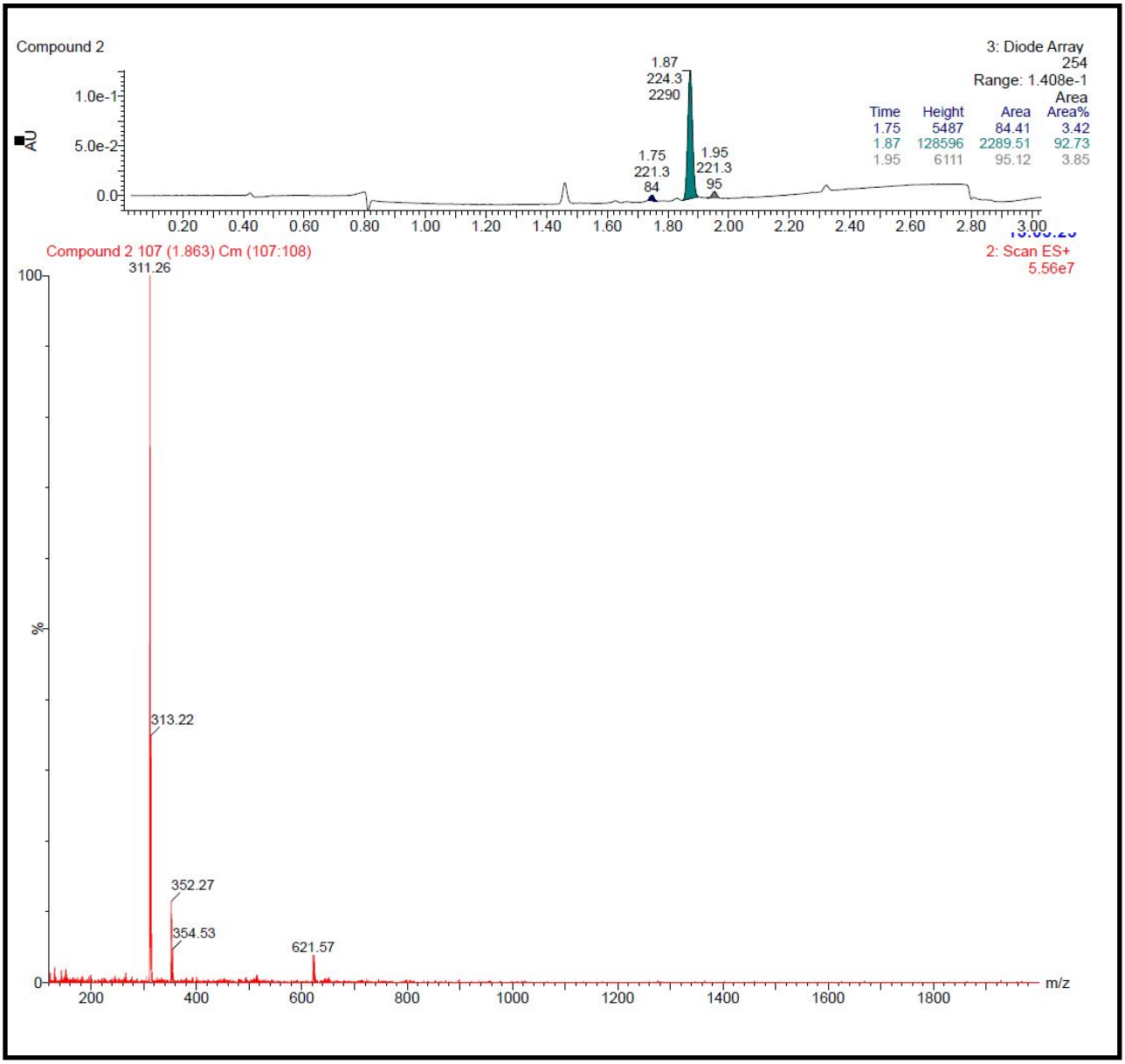

NM2-101-QC_CONC_CARBON_001
NM2-101-QC_CONC
A(d) 162.08
B(d) 115.73
f1(ppm)

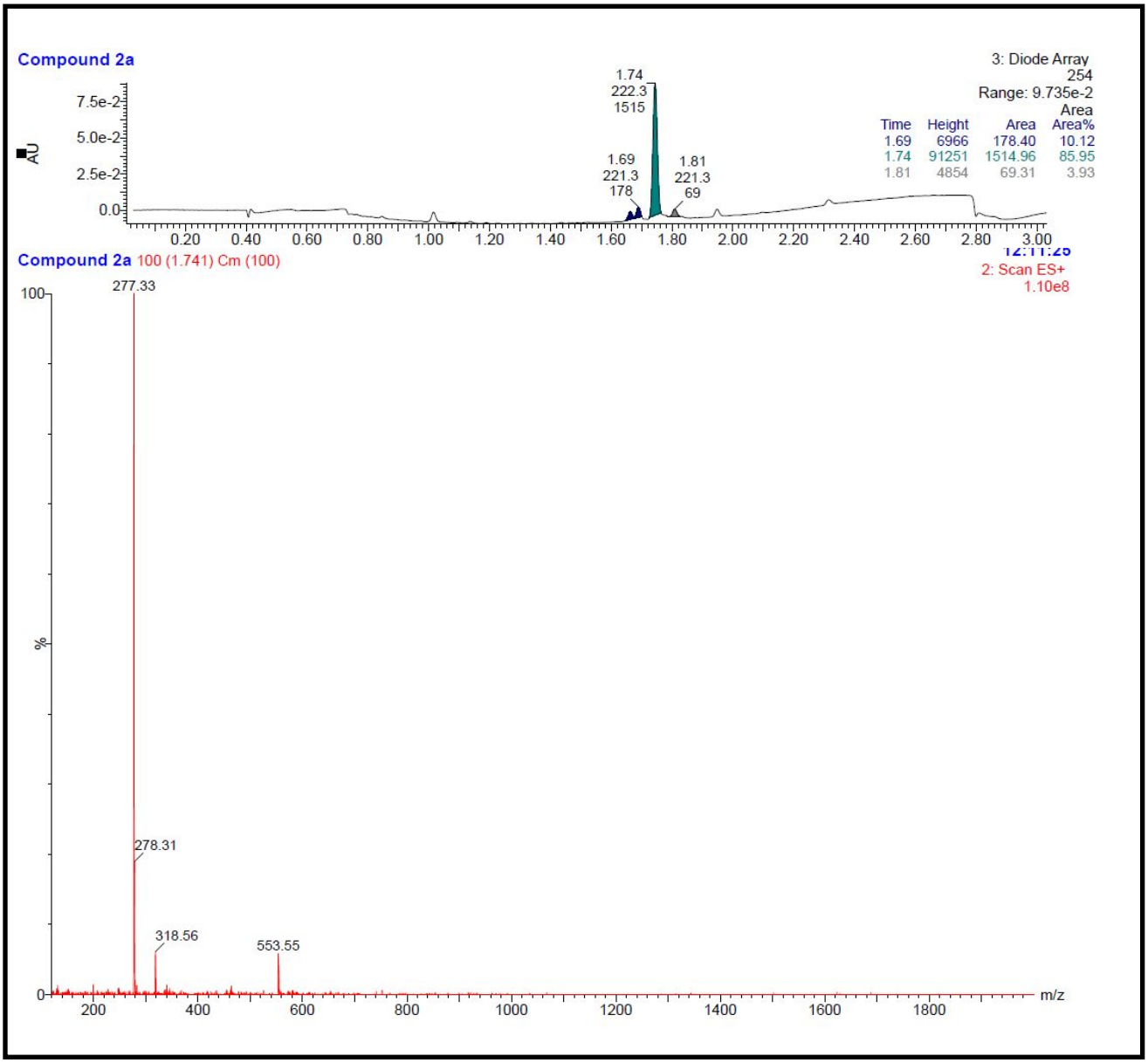

NM2-101-QC_CONC_FLUORINE_001
NM2-101-QC_CONC
-76.55 TFA
-115.21
-116.12
f1 (ppm)

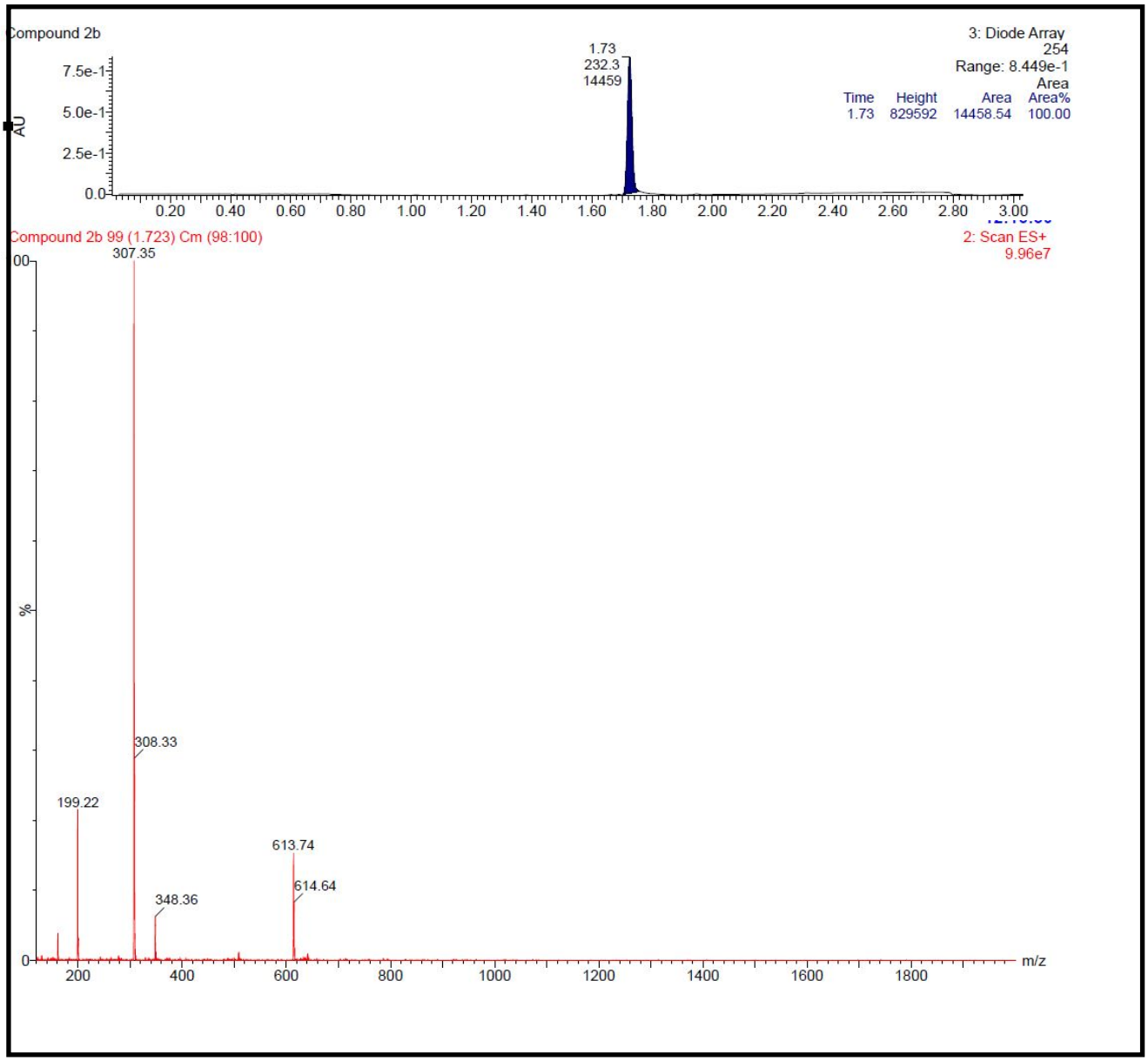

DAD1 B, Sig=254,16 Ref=360,100 (10232017\SK-NM2-101-FR-13-16.D)
mAU
5.595
5.855
6.192
6.430
6.674
min

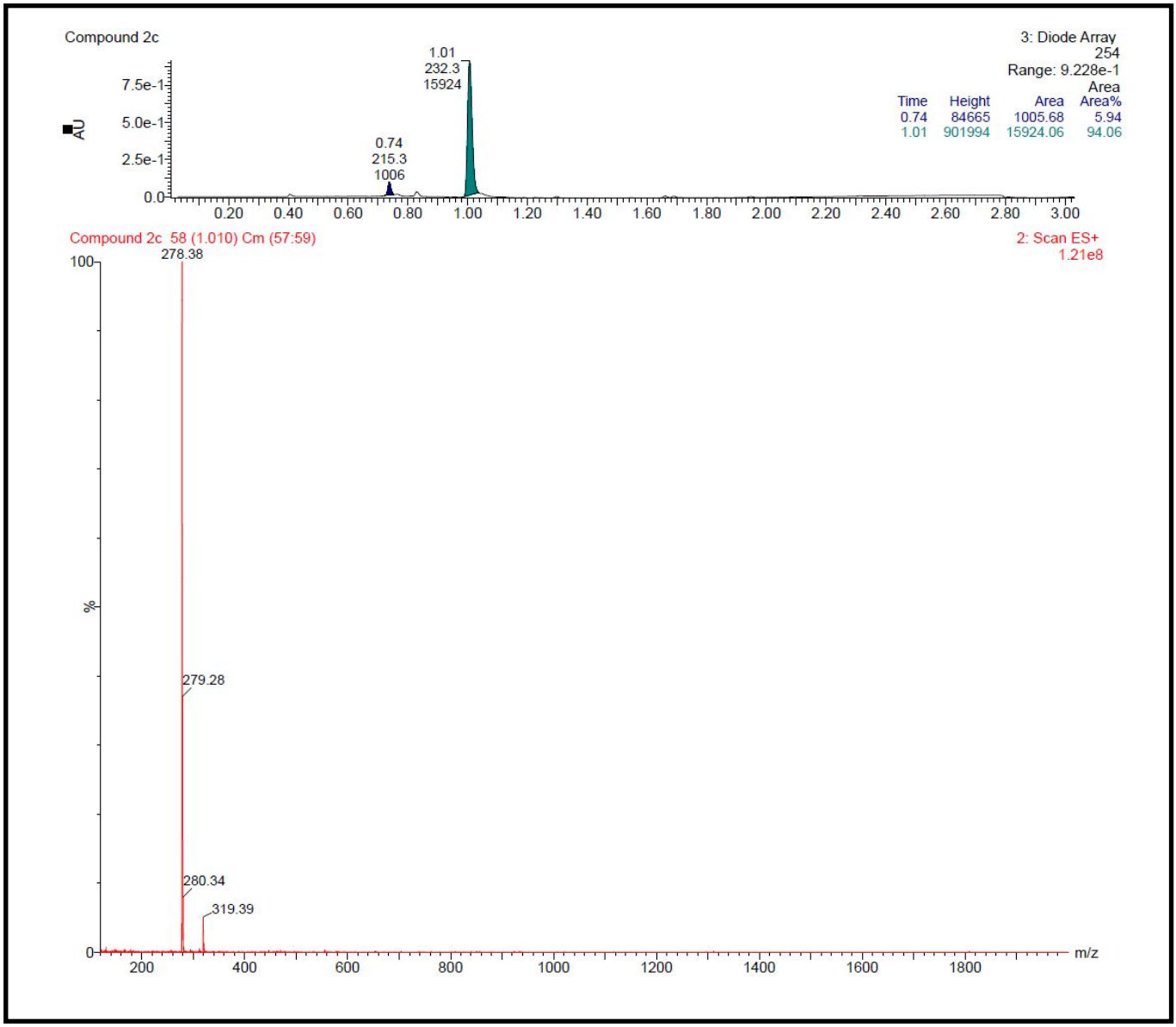
Compound purity: analytical spectra

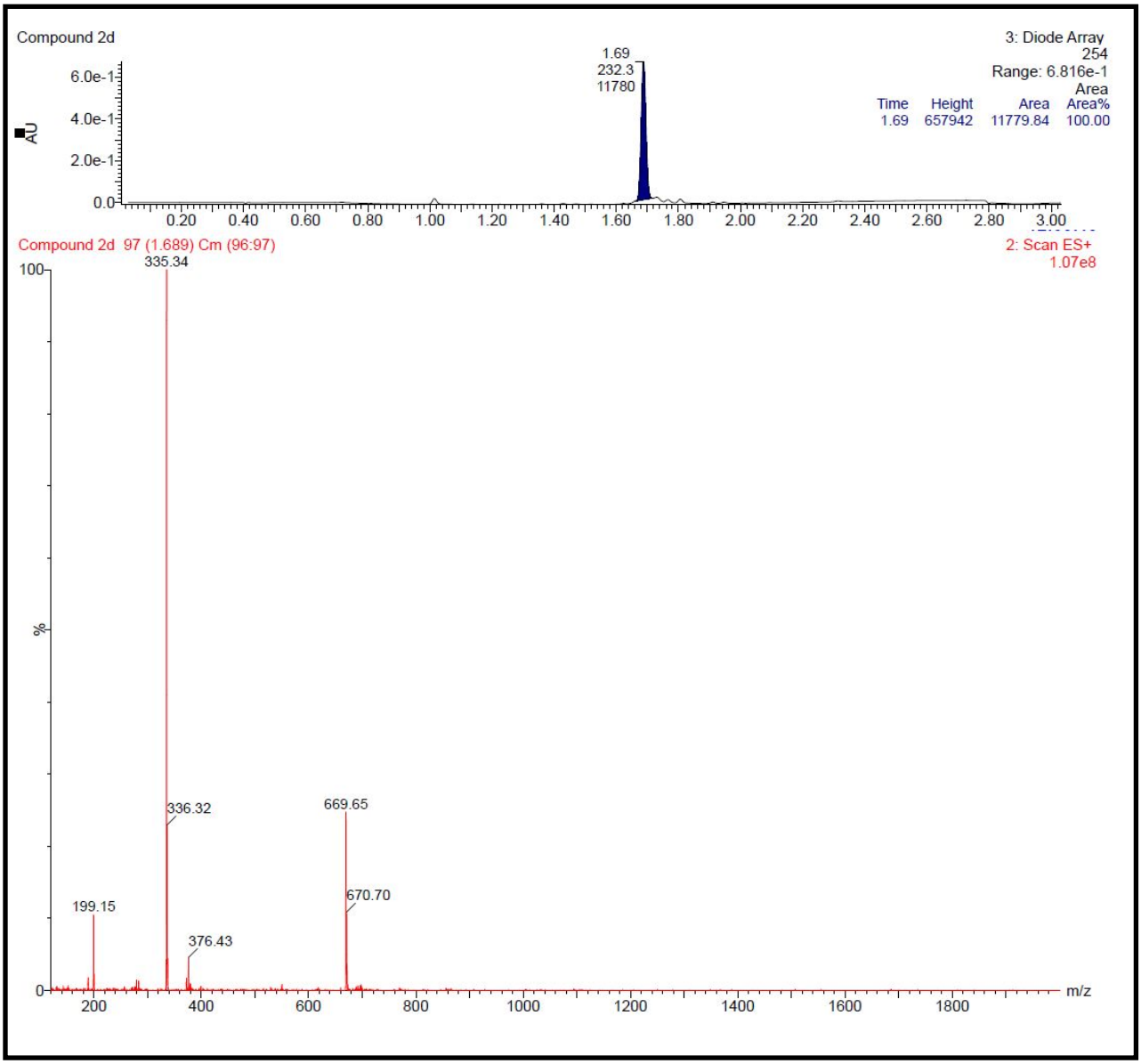


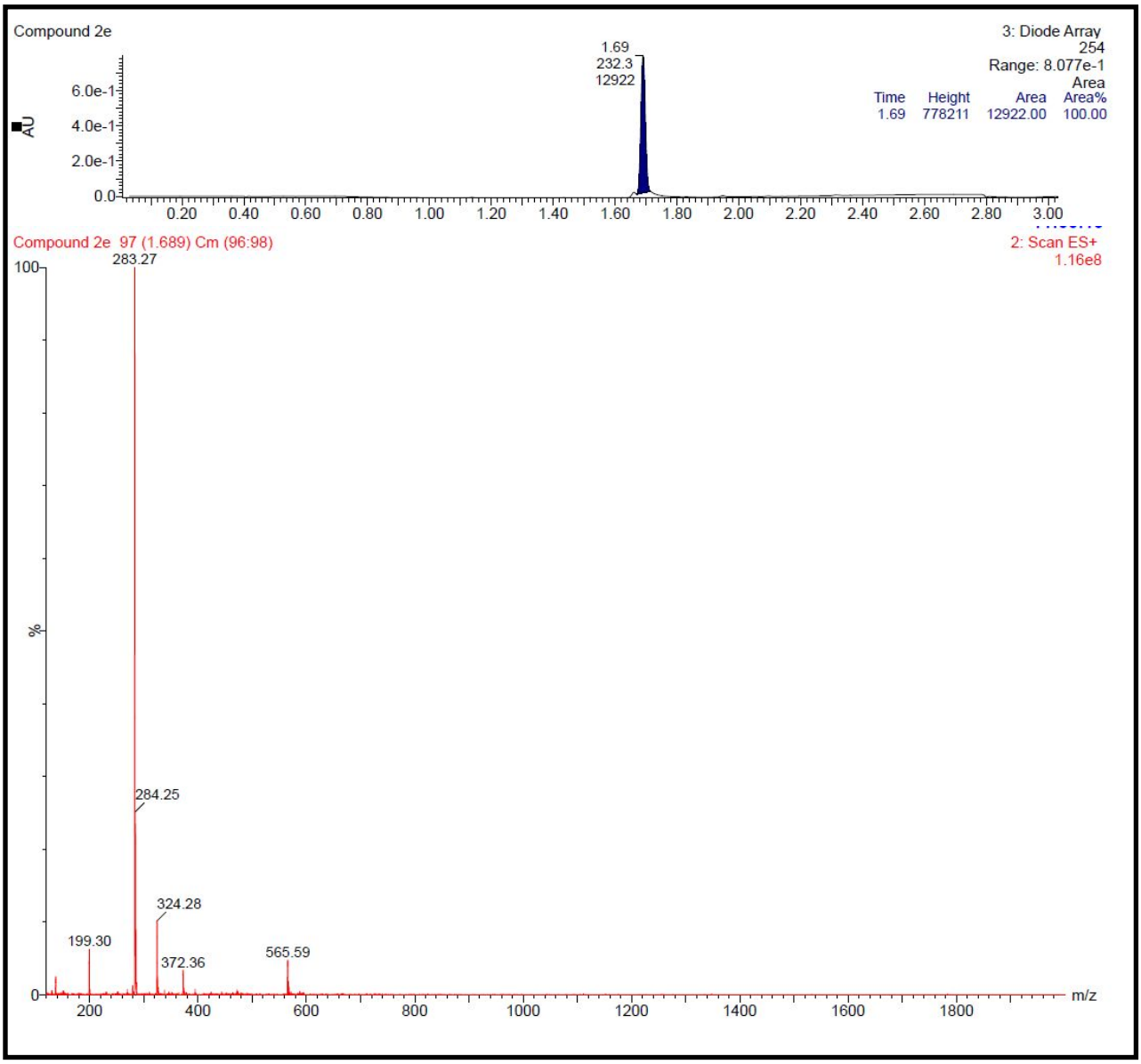

Compound 2
3: Diode Array
254
Range: 1.408e-1
Area
Time Height Area Area%
1.75 5487 84.41 3.42
1.87 128596 2289.51 92.73
1.95 6111 95.12 3.85
1.87
224.3
2290
1.75
221.3
84
1.95
221.3
95
AU
1.0e-1
5.0e-2
0.0
0.20 0.40 0.60 0.80 1.00 1.20 1.40 1.60 1.80 2.00 2.20 2.40 2.60 2.80 3.00
Compound 2 107 (1.863) Cm (107:108)
2: Scan ES+
5.56e7
100
%
0
311.26
313.22
352.27
354.53
621.57
200 400 600 800 1000 1200 1400 1600 1800
m/z

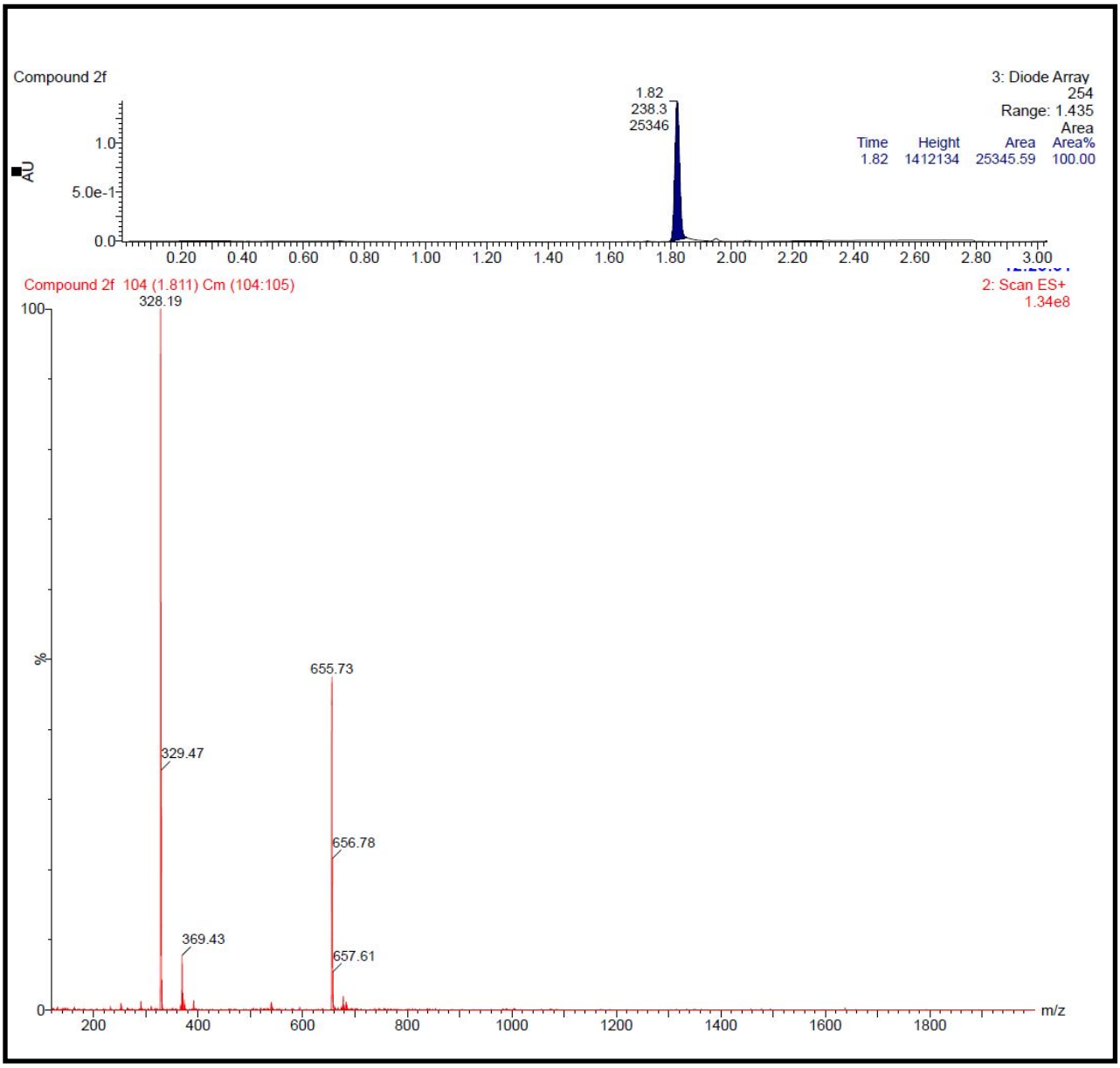

Compound 2a
3: Diode Array
254
Range: 9.735e-2
Area
Time Height Area Area%
1.69 6966 178.40 10.12
1.74 91251 1514.96 85.95
1.81 4854 69.31 3.93
AU
7.5e-2
5.0e-2
2.5e-2
0.0
1.74
222.3
1515
1.69
221.3
178
1.81
221.3
69
0.20 0.40 0.60 0.80 1.00 1.20 1.40 1.60 1.80 2.00 2.20 2.40 2.60 2.80 3.00
Compound 2a 100 (1.741) Cm (100)
2: Scan ES+
1.10e8
100
%
0
277.33
278.31
318.56
553.55
200 400 600 800 1000 1200 1400 1600 1800
m/z

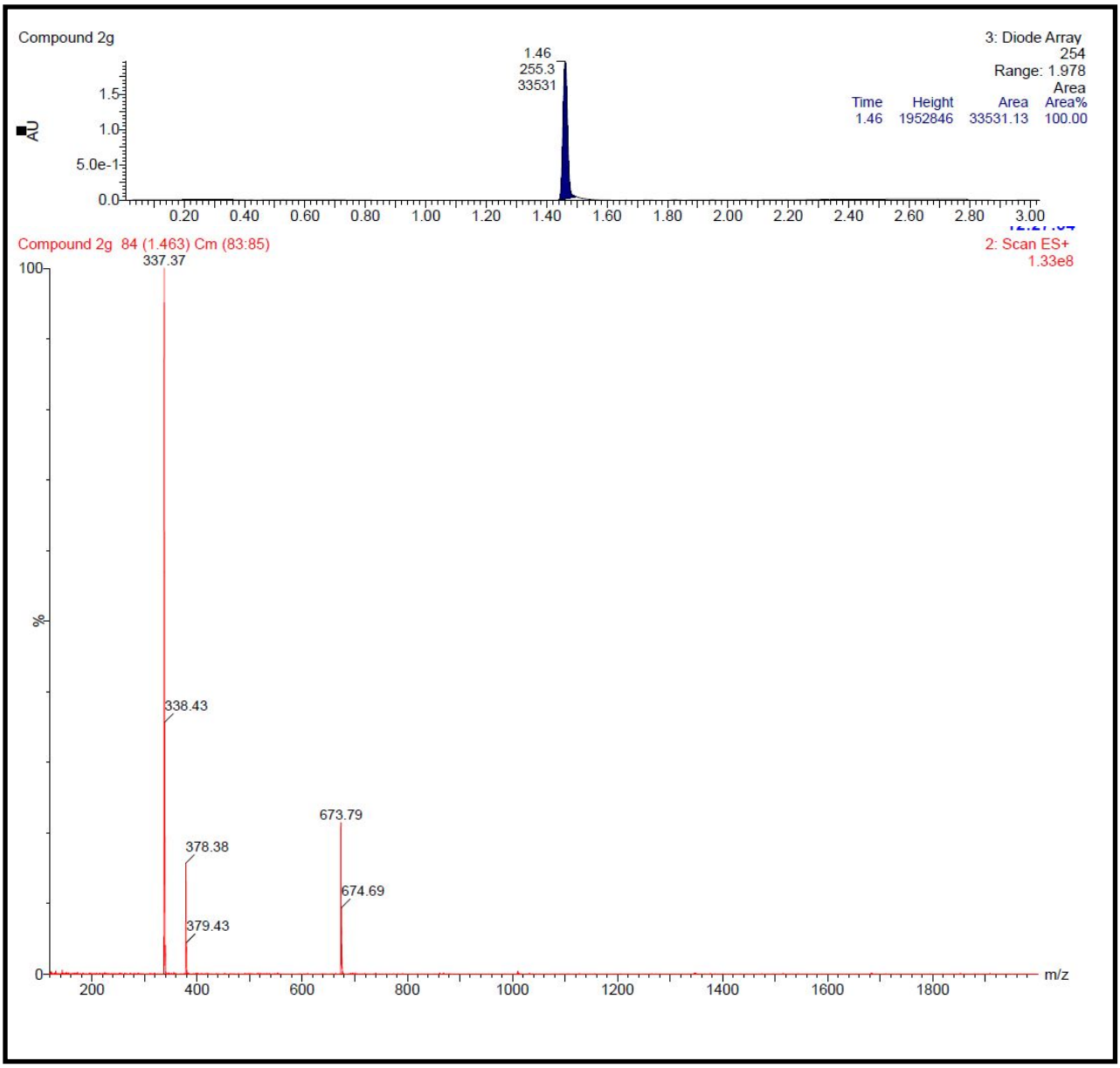

Compound 2b
3: Diode Array
254
Range: 8.449e-1
Area
Time Height Area Area%
1.73 829592 14458.54 100.00
1.73
232.3
14459
AU
7.5e-1
5.0e-1
2.5e-1
0.0
0.20 0.40 0.60 0.80 1.00 1.20 1.40 1.60 1.80 2.00 2.20 2.40 2.60 2.80 3.00
Compound 2b 99 (1.723) Cm (98:100)
2: Scan ES+
9.96e7
307.35
308.33
199.22
613.74
614.64
348.36
%
0
200 400 600 800 1000 1200 1400 1600 1800
m/z

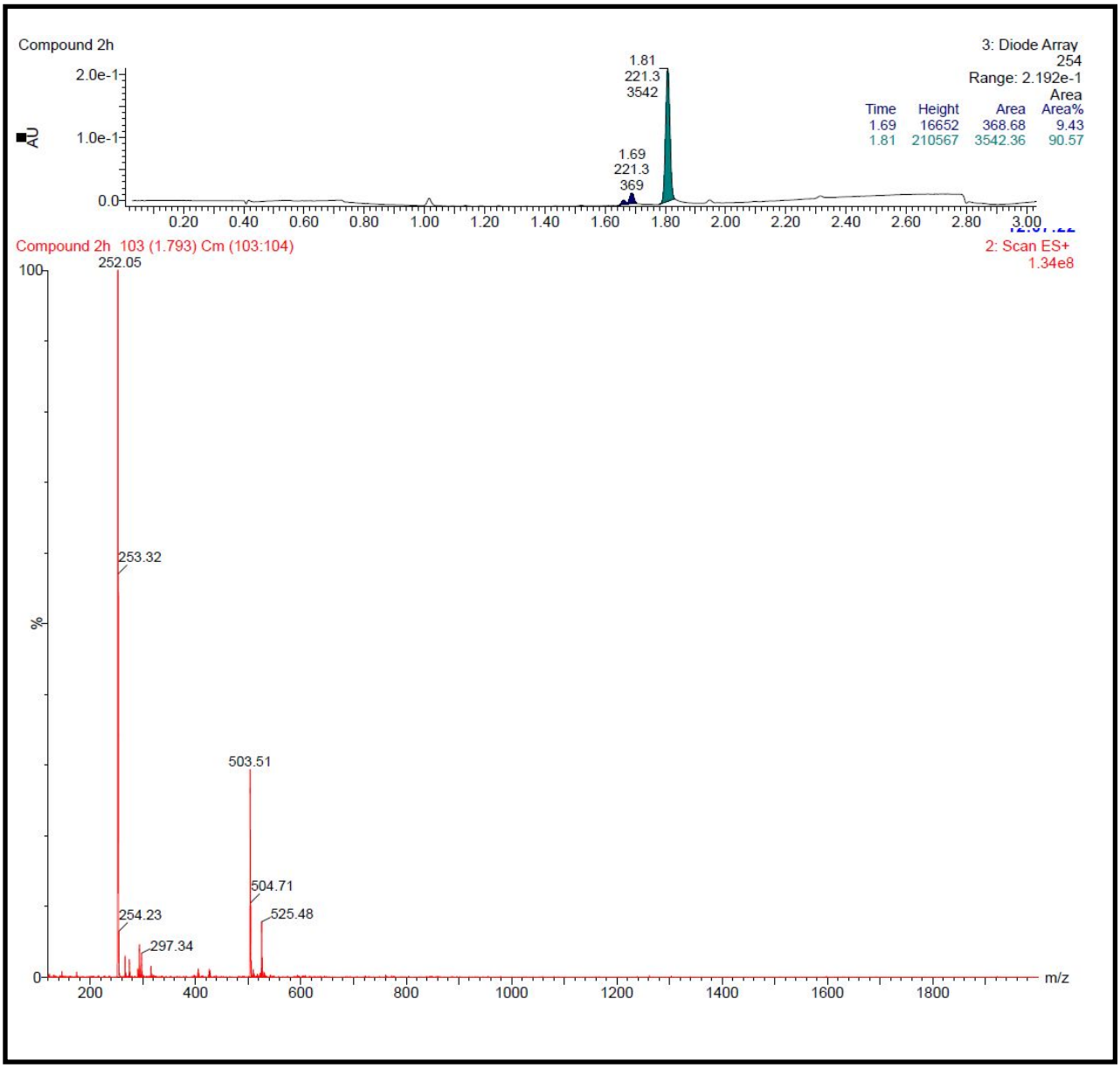

Compound 2c
3: Diode Array
254
Range: 9.228e-1
Area
Time Height Area Area%
0.74 84665 1005.68 5.94
1.01 901994 15924.06 94.06
1.01
232.3
15924
0.74
215.3
1006
AU
7.5e-1
5.0e-1
2.5e-1
0.0
0.20 0.40 0.60 0.80 1.00 1.20 1.40 1.60 1.80 2.00 2.20 2.40 2.60 2.80 3.00
Compound 2c 58 (1.010) Cm (57:59)
2: Scan ES+
1.21e8
100
%
278.38
279.28
280.34
319.39
0
200 400 600 800 1000 1200 1400 1600 1800
m/z

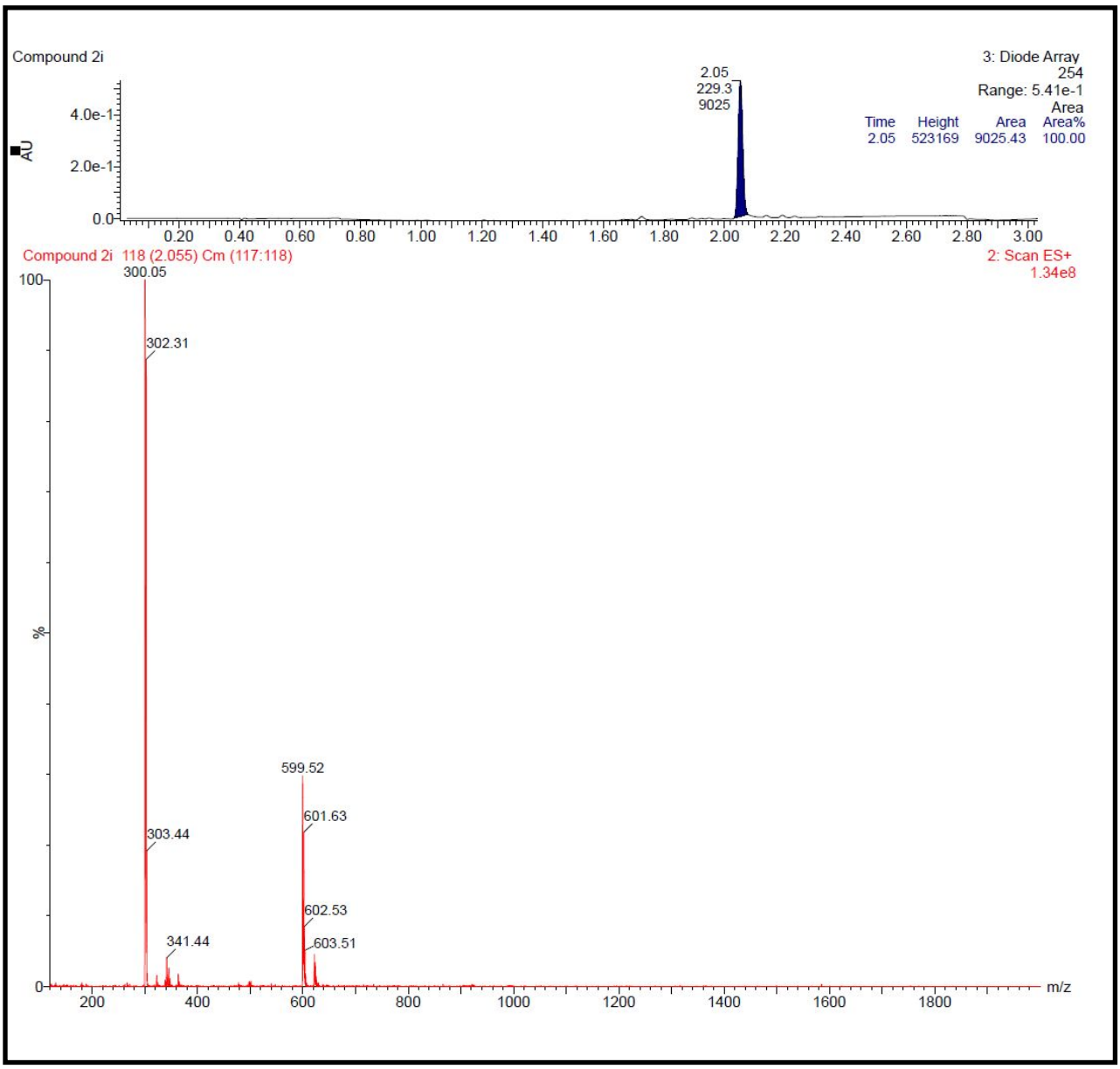

Compound 2d
3: Diode Array
254
Range: 6.816e-1
Area
Time Height Area Area%
1.69 657942 11779.84 100.00
1.69
232.3
11780
AU
6.0e-1
4.0e-1
2.0e-1
0.0
0.20 0.40 0.60 0.80 1.00 1.20 1.40 1.60 1.80 2.00 2.20 2.40 2.60 2.80 3.00
Compound 2d 97 (1.689) Cm (96:97)
2: Scan ES+
1.07e8
100
%
0
335.34
336.32
669.65
670.70
199.15
376.43
200 400 600 800 1000 1200 1400 1600 1800
m/z

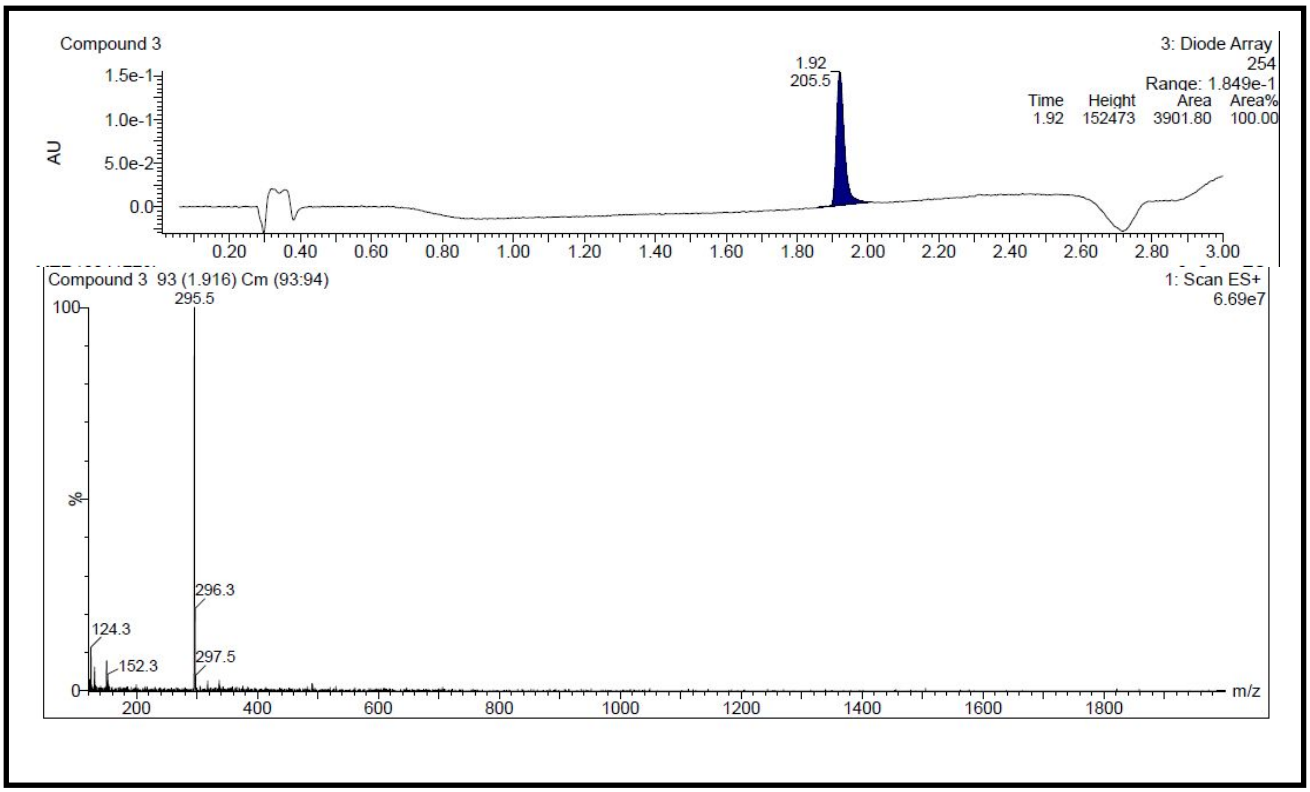

Compound 2e
3: Diode Array
254
Range: 8.077e-1
Area
Time Height Area Area%
1.69 778211 12922.00 100.00
1.69
232.3
12922
AU
6.0e-1
4.0e-1
2.0e-1
0.0
0.20 0.40 0.60 0.80 1.00 1.20 1.40 1.60 1.80 2.00 2.20 2.40 2.60 2.80 3.00
Compound 2e 97 (1.689) Cm (96:98)
2: Scan ES+
1.16e8
100
%
283.27
284.25
324.28
199.30
372.36
565.59
0
200 400 600 800 1000 1200 1400 1600 1800
m/z

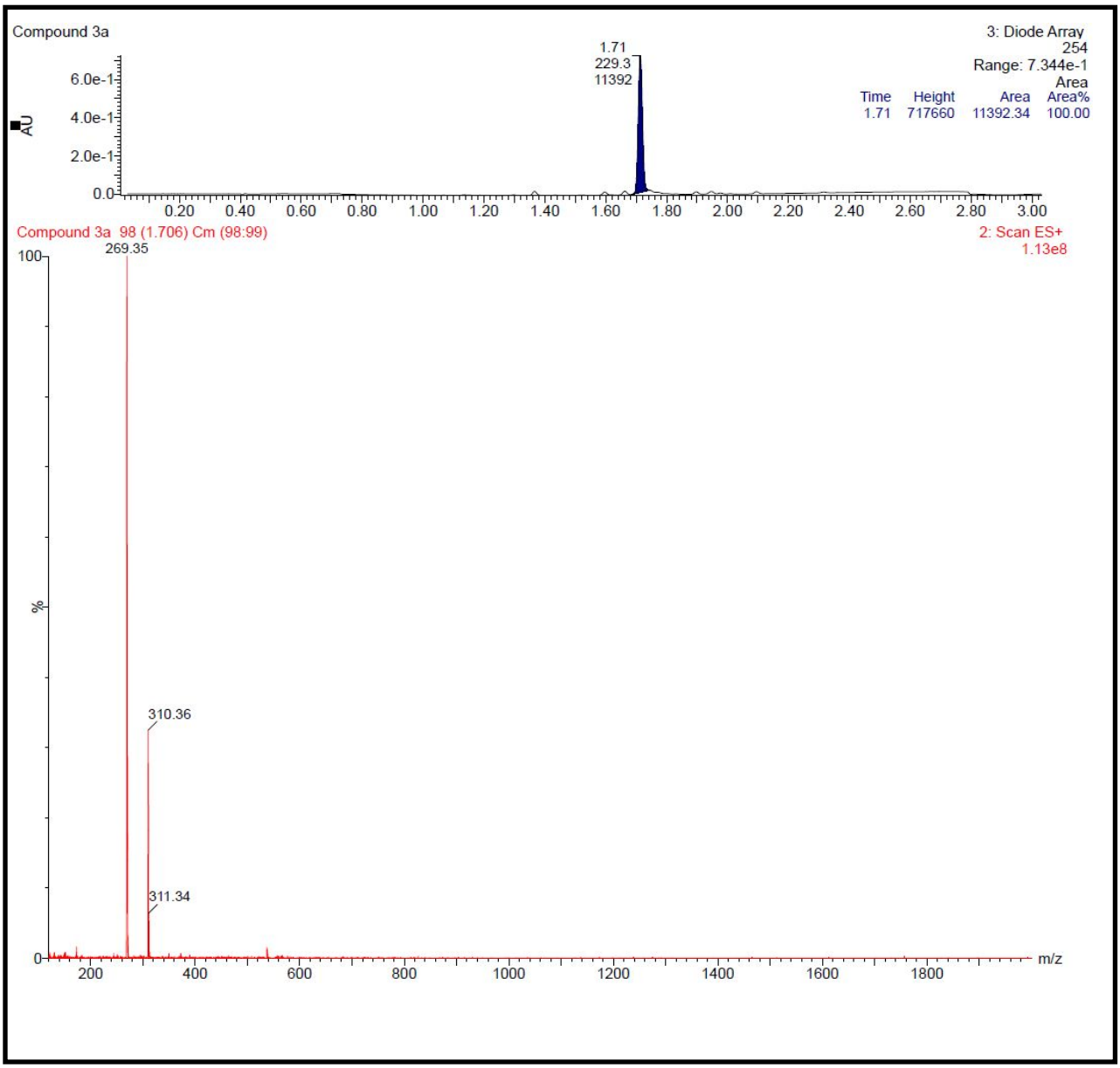

Compound 2f
3: Diode Array
254
Range: 1.435
Area
Time Height Area Area%
1.82 1412134 25345.59 100.00
1.82
238.3
25346
AU
1.0
5.0e-1
0.0
0.20 0.40 0.60 0.80 1.00 1.20 1.40 1.60 1.80 2.00 2.20 2.40 2.60 2.80 3.00
Compound 2f 104 (1.811) Cm (104:105)
2: Scan ES+
1.34e8
100
%
0
328.19
655.73
329.47
656.78
369.43
657.61
200 400 600 800 1000 1200 1400 1600 1800
m/z

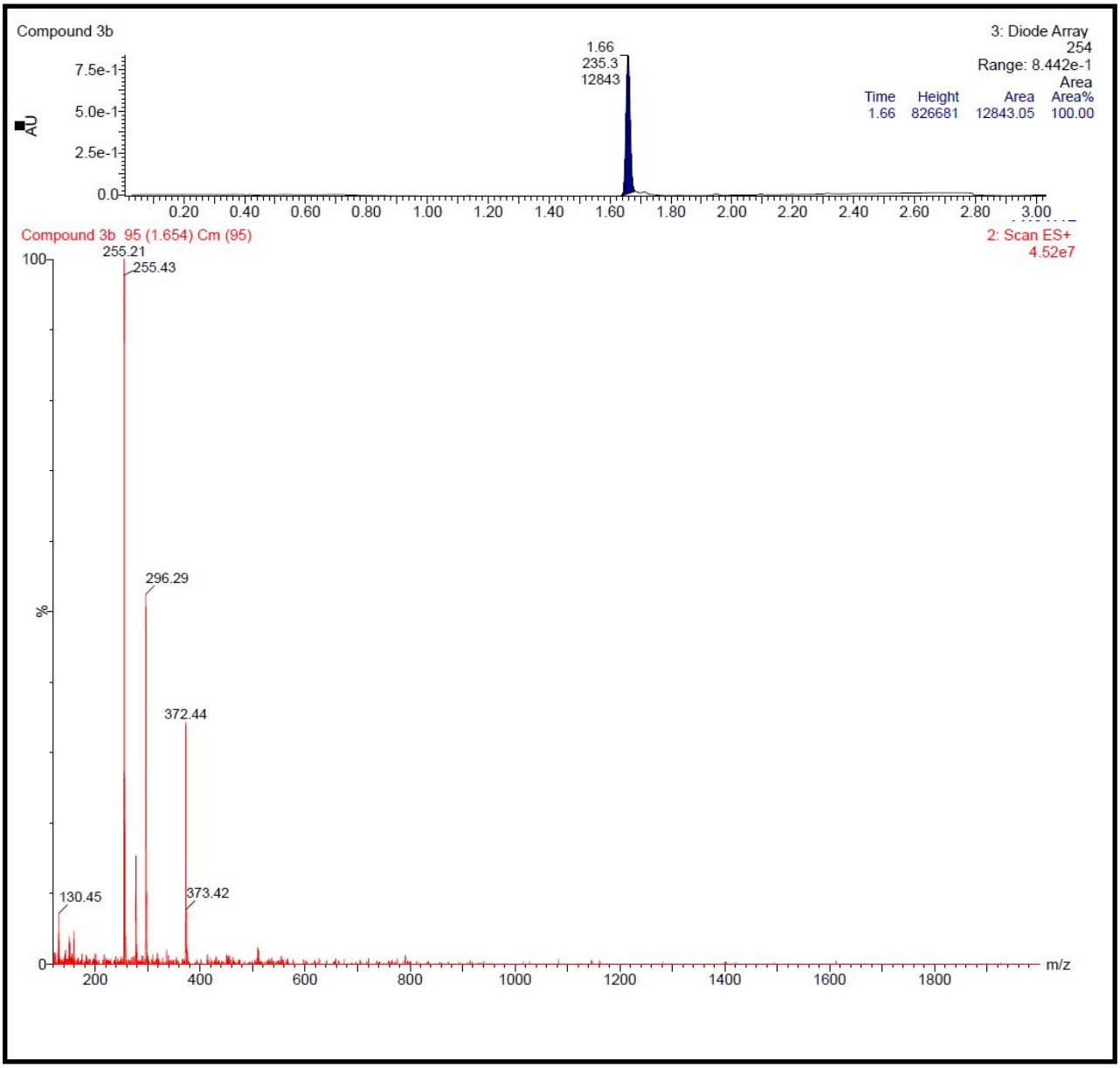

Compound 2g
3: Diode Array
254
Range: 1.978
Area
Time Height Area Area%
1.46 1952846 33531.13 100.00
1.46
255.3
33531
AU
1.5
1.0
5.0e-1
0.0
0.20 0.40 0.60 0.80 1.00 1.20 1.40 1.60 1.80 2.00 2.20 2.40 2.60 2.80 3.00
Compound 2g 84 (1.463) Cm (83:85)
2: Scan ES+
1.33e8
100
%
0
337.37
338.43
673.79
378.38
674.69
379.43
200 400 600 800 1000 1200 1400 1600 1800
m/z


Compound 2h
3: Diode Array
254
Range: 2.192e-1
Area
Time Height Area Area%
1.69 16652 368.68 9.43
1.81 210567 3542.36 90.57
1.81
221.3
3542
1.69
221.3
369
AU
2.0e-1
1.0e-1
0.0
0.20 0.40 0.60 0.80 1.00 1.20 1.40 1.60 1.80 2.00 2.20 2.40 2.60 2.80 3.00
Compound 2h 103 (1.793) Cm (103:104)
2: Scan ES+
1.34e8
100
%
0
252.05
253.32
503.51
504.71
525.48
254.23
297.34
200 400 600 800 1000 1200 1400 1600 1800
m/z

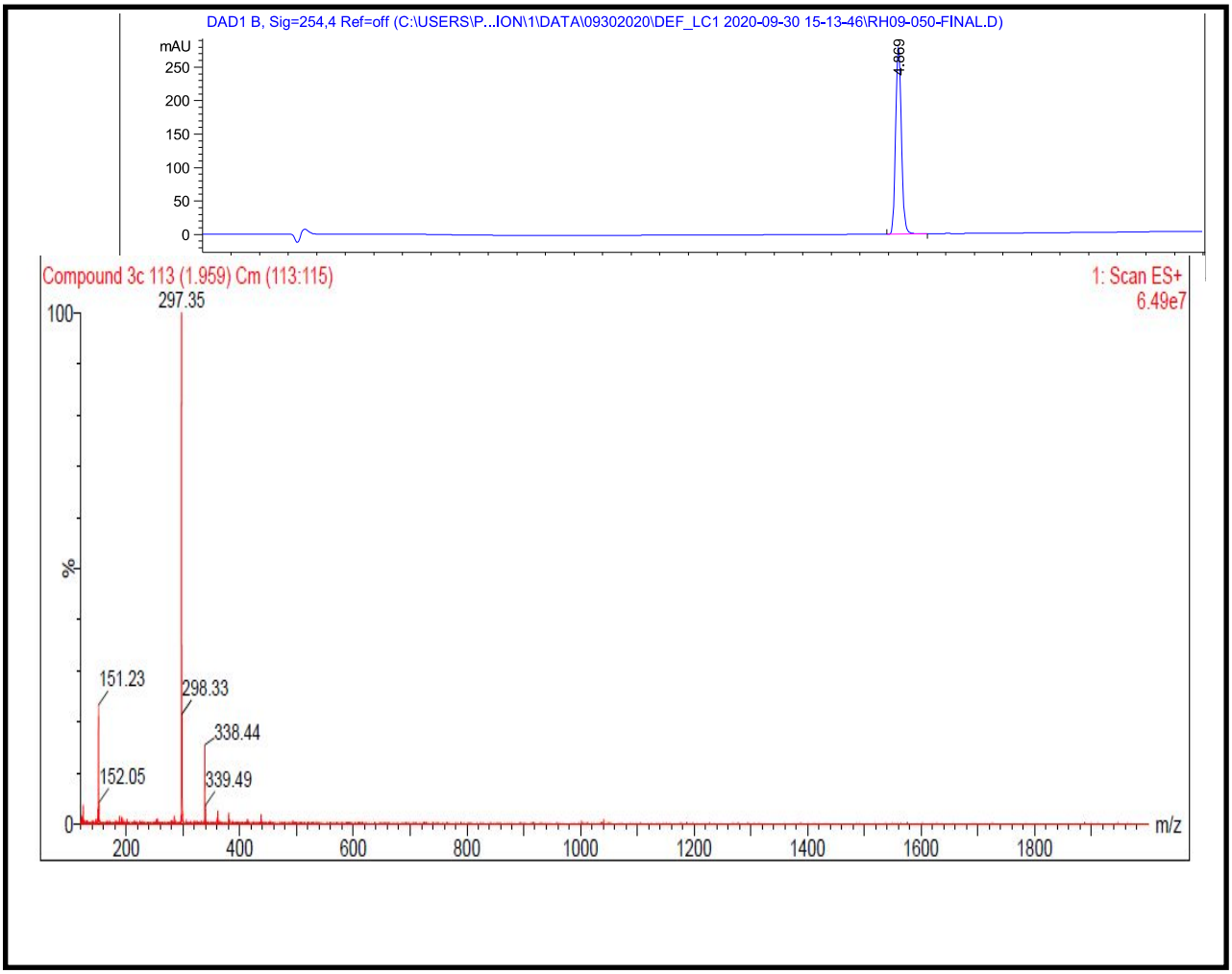

Compound 2i
3: Diode Array
254
Range: 5.41e-1
Area
Time Height Area Area%
2.05 523169 9025.43 100.00
2.05
229.3
9025
AU
4.0e-1
2.0e-1
0.0
0.20 0.40 0.60 0.80 1.00 1.20 1.40 1.60 1.80 2.00 2.20 2.40 2.60 2.80 3.00
Compound 2i 118 (2.055) Cm (117:118)
2: Scan ES+
1.34e8
100
%
0
300.05
302.31
303.44
341.44
599.52
601.63
602.53
603.51
200 400 600 800 1000 1200 1400 1600 1800
m/z

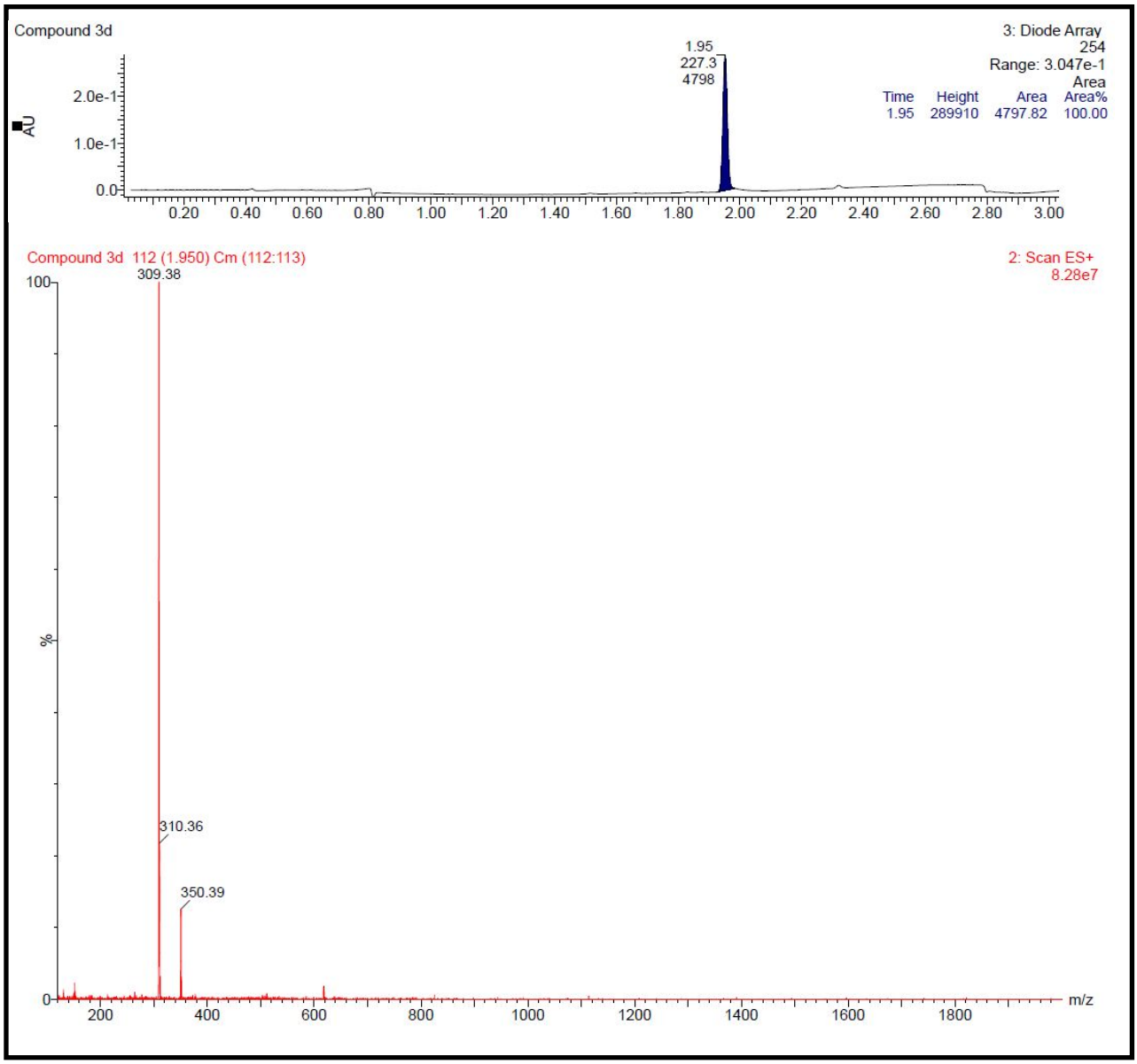

Compound 3
3: Diode Array
254
Range: 1.849e-1
1.92
205.5
Time Height Area Area%
1.92 152473 3901.80 100.00
AU
1.5e-1
1.0e-1
5.0e-2
0.0
0.20 0.40 0.60 0.80 1.00 1.20 1.40 1.60 1.80 2.00 2.20 2.40 2.60 2.80 3.00
Compound 3 93 (1.916) Cm (93:94)
1: Scan ES+
6.69e7
295.5
100
%
296.3
124.3
297.5
152.3
0
200 400 600 800 1000 1200 1400 1600 1800
m/z

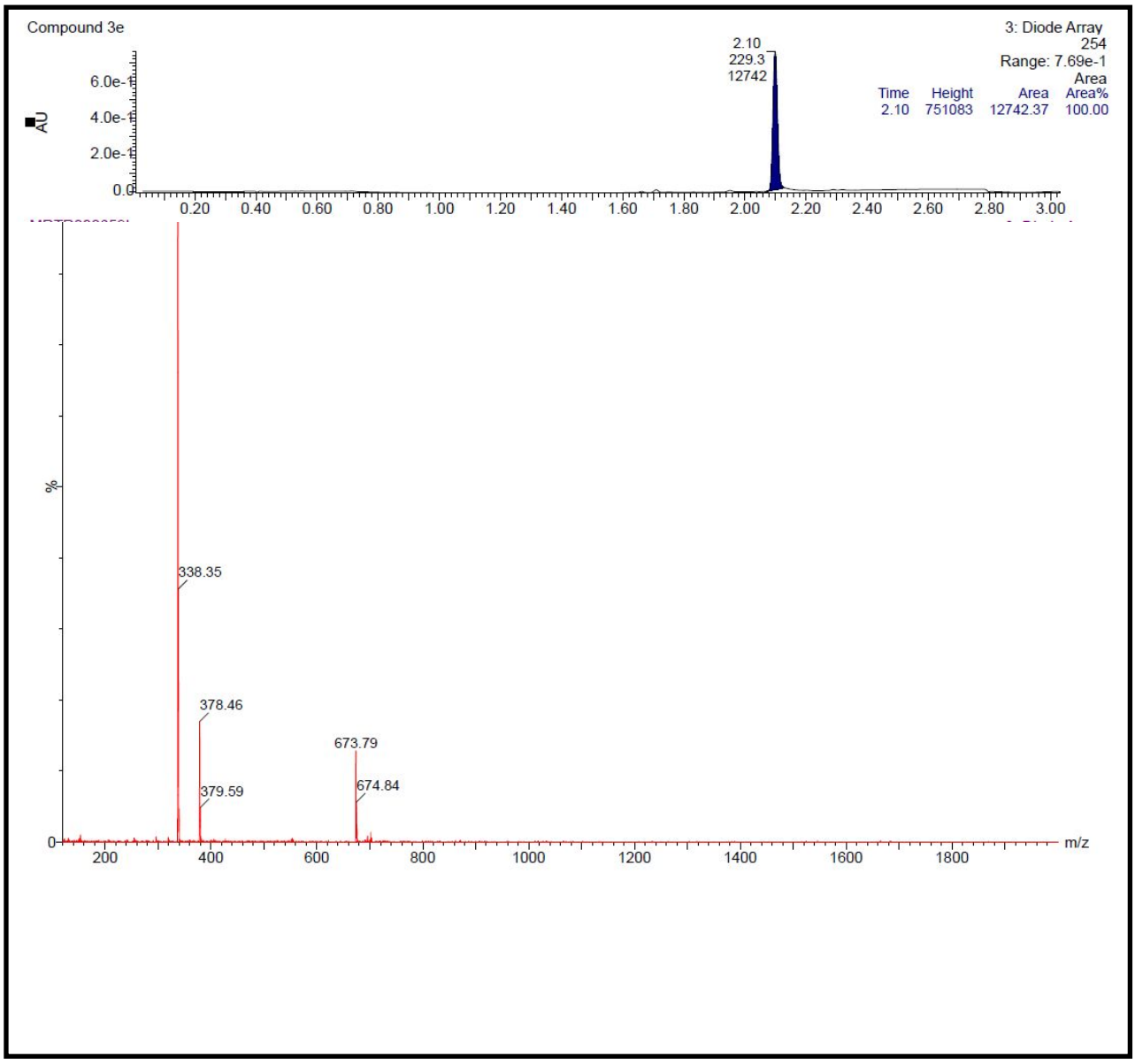

Compound 3a
3: Diode Array
254
Range: 7.344e-1
1.71
229.3
11392
Area
Time Height Area Area%
1.71 717660 11392.34 100.00
AU
6.0e-1
4.0e-1
2.0e-1
0.0
0.20 0.40 0.60 0.80 1.00 1.20 1.40 1.60 1.80 2.00 2.20 2.40 2.60 2.80 3.00
Compound 3a 98 (1.706) Cm (98:99)
2: Scan ES+
1.13e8
100
269.35
%
310.36
311.34
0
200 400 600 800 1000 1200 1400 1600 1800
m/z

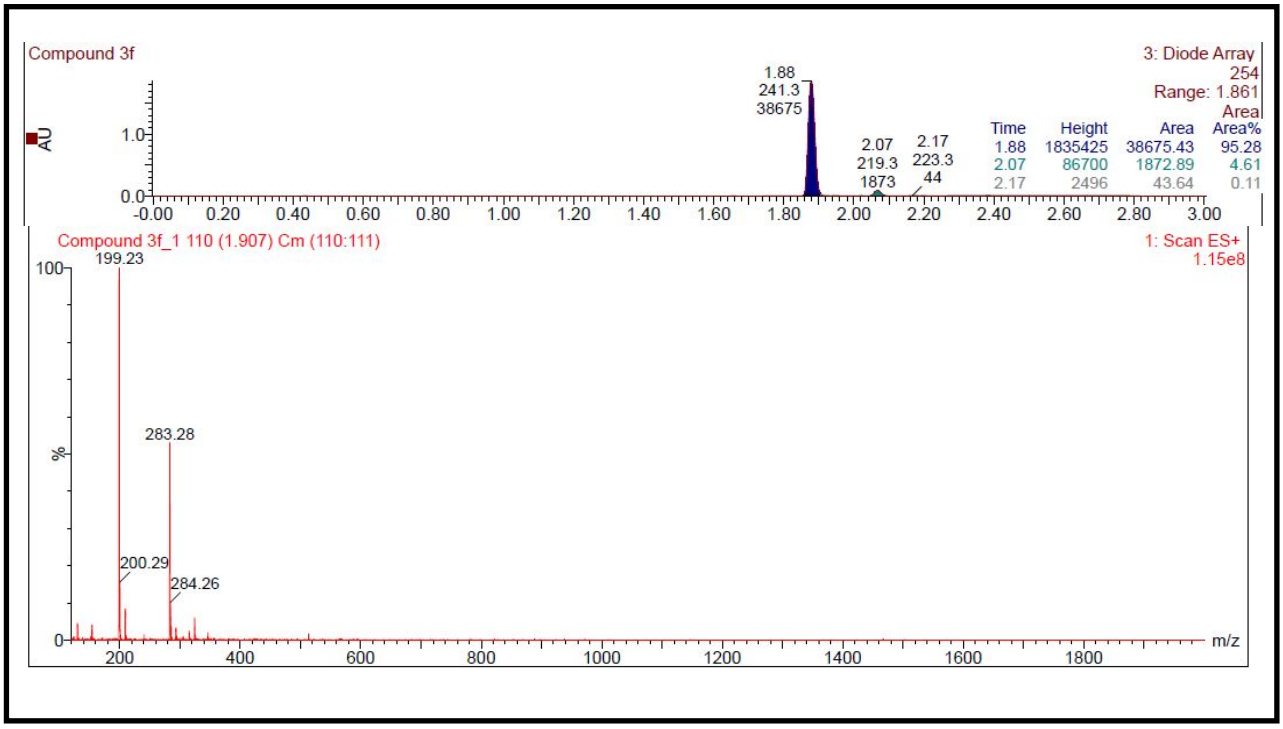

Compound 3b
3: Diode Array
254
Range: 8.442e-1
Area
Time Height Area Area%
1.66 826681 12843.05 100.00
1.66
235.3
12843
AU
7.5e-1
5.0e-1
2.5e-1
0.0
0.20 0.40 0.60 0.80 1.00 1.20 1.40 1.60 1.80 2.00 2.20 2.40 2.60 2.80 3.00
Compound 3b 95 (1.654) Cm (95)
2: Scan ES+
4.52e7
100
%
255.21
255.43
296.29
372.44
130.45
373.42
0
200 400 600 800 1000 1200 1400 1600 1800
m/z

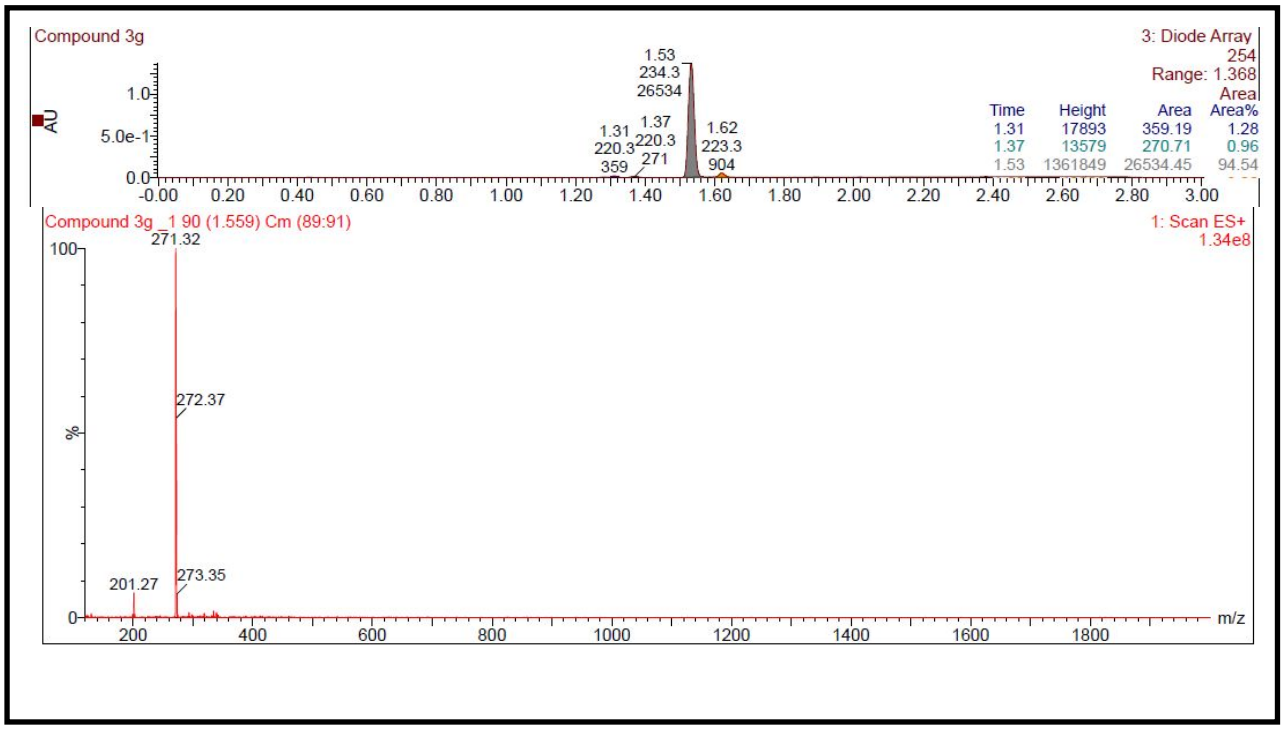
Compound 3c


DAD1 B, Sig=254,4 Ref=off (C:\USERS\P...ION\1\DATA\09302020\DEF_LC1 2020-09-30 15-13-46\RH09-050-FINAL.D)
mAU
250
200
150
100
50
0
4.869
Compound 3c 113 (1.959) Cm (113:115)
1: Scan ES+
6.49e7
297.35
100
%
151.23
298.33
338.44
152.05
339.49
0
200
400
600
800
1000
1200
1400
1600
1800
m/z


Compound 3d
3: Diode Array
254
Range: 3.047e-1
Area
Time Height Area Area%
1.95 289910 4797.82 100.00
1.95
227.3
4798
AU
2.0e-1
1.0e-1
0.0
0.20 0.40 0.60 0.80 1.00 1.20 1.40 1.60 1.80 2.00 2.20 2.40 2.60 2.80 3.00
Compound 3d 112 (1.950) Cm (112:113)
2: Scan ES+
8.28e7
309.38
100
%
310.36
350.39
0
200 400 600 800 1000 1200 1400 1600 1800
m/z


Compound 3e
3: Diode Array
254
Range: 7.69e-1
2.10
229.3
12742
Area
Time Height Area Area%
2.10 751083 12742.37 100.00
AU
6.0e-1
4.0e-1
2.0e-1
0.0
0.20 0.40 0.60 0.80 1.00 1.20 1.40 1.60 1.80 2.00 2.20 2.40 2.60 2.80 3.00
%
338.35
378.46
673.79
379.59
674.84
0
200 400 600 800 1000 1200 1400 1600 1800
m/z


Compound 3f
3: Diode Array
254
Range: 1.861
AU
1.88
241.3
38675
2.07
219.3
1873
2.17
223.3
44
Time Height Area Area%
1.88 1835425 38675.43 95.28
2.07 86700 1872.89 4.61
2.17 2496 43.64 0.11
Compound 3f_1 110 (1.907) Cm (110:111)
1: Scan ES+
1.15e8
199.23
283.28
200.29
284.26
%
m/z


Compound 3g
3: Diode Array
254
Range: 1.368
AU
1.0
5.0e-1
0.0
1.53
234.3
26534
1.31
220.3
359
1.37
220.3
271
1.62
223.3
904
Time Height Area Area%
1.31 17893 359.19 1.28
1.37 13579 270.71 0.96
1.53 1361849 26534.45 94.54
-0.00 0.20 0.40 0.60 0.80 1.00 1.20 1.40 1.60 1.80 2.00 2.20 2.40 2.60 2.80 3.00
Compound 3g_1 90 (1.559) Cm (89:91)
1: Scan ES+
1.34e8
100
%
0
271.32
272.37
201.27
273.35
200 400 600 800 1000 1200 1400 1600 1800
m/z