

Supporting Information.

A new molecular graph-based deep learning algorithm facilitates an imaging-based strategy for rapid discovery of small molecules modulating biomolecular condensates

Authors:

Peng Gao¹, Qi Zhang¹, Devin Keely², Don W. Cleveland³, Yihong Ye⁴, Wei Zheng¹, Min Shen^{1*}, and Haiyang Yu^{2*#}

Affiliations

¹*The National Center for Advancing Translational Sciences (NCATS), National Institutes of Health (NIH), MD 20850, USA*

²*Center for Alzheimer's and Neurodegenerative Diseases, Department of Molecular Biology, Peter O'Donnell Jr. Brain Institute, UT Southwestern Medical Center, TX, 75287, USA*

³*Department of Cellular and Molecular Medicine, UC San Diego, CA, 92093, USA*

⁴*National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institutes of Health (NIH), MD 20850, USA;*

*Correspondence:

haiyang.yu@utsouthwestern.edu
shenmin@mail.nih.gov

#Current affiliation: *Department of Neuroscience, Genentech, Inc., South San Francisco, CA 94080, USA. Email: yu.haiyang@gene.com*

Supporting Information Content:

1. Clustering results of the qHTS screened compounds
2. The verified effect of compound A1-A5 on cellular LLPS compartments
3. QC reports of compound A1-A5

Clustering of the screened compounds

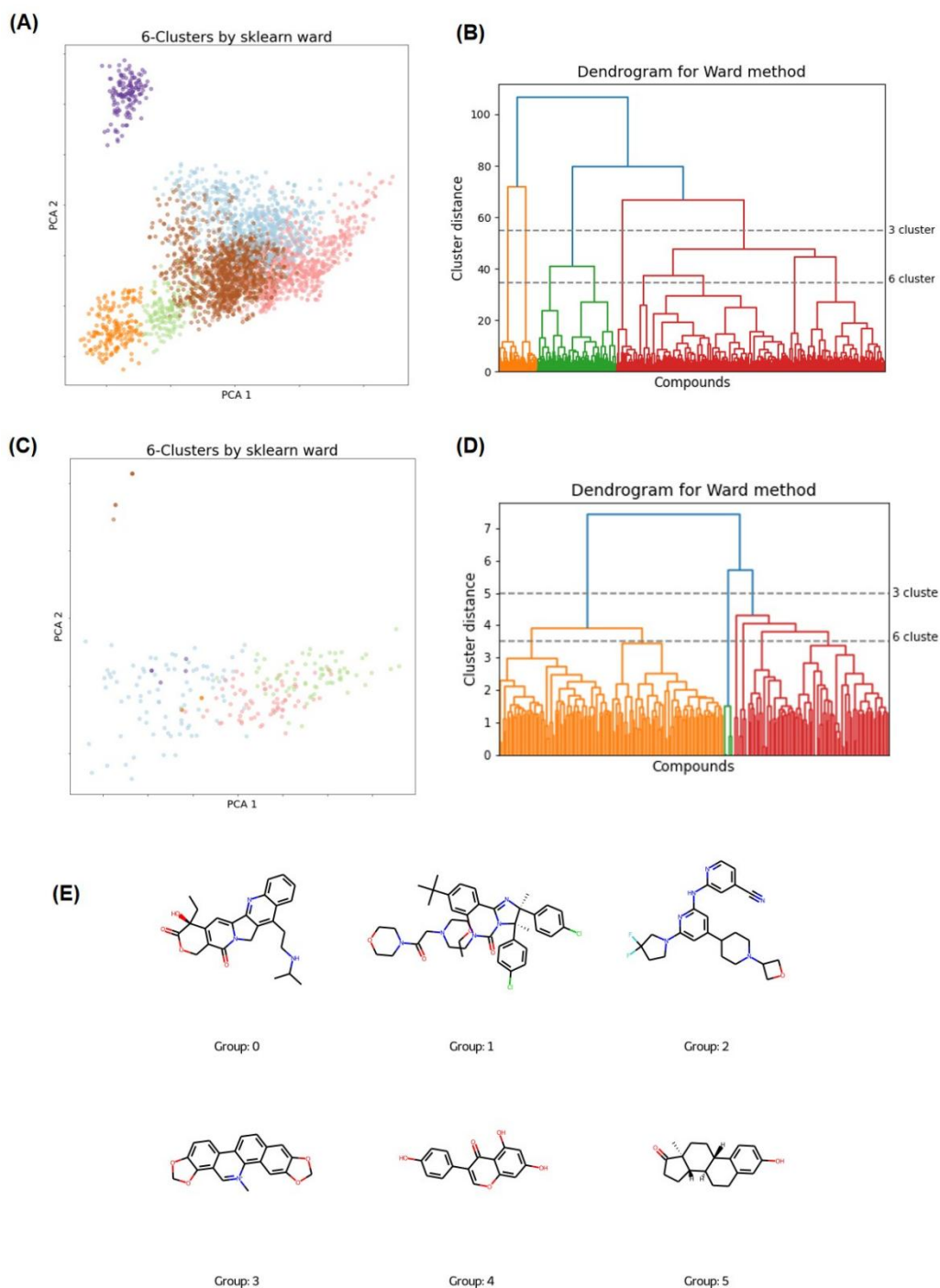


Figure S1. The Principle Component Analysis (PCA) & clustering of the qHTS screened compounds by Ward method. (A-B) The clustering result of all the screened compounds by qHTS, the compounds are divided into six clusters (represented by six colors) for statistics. **(C-D)** The clustering result of the identified active compounds, the compounds are divided into six clusters (represented by six colors) for statistics. **(E)** Some of the screened compounds with different scaffolds.

DLD1 cells → Compound (10uM) for 24h → NaAsO₂ (250uM) for 1 hour → Fix, Immunostaining and imaging

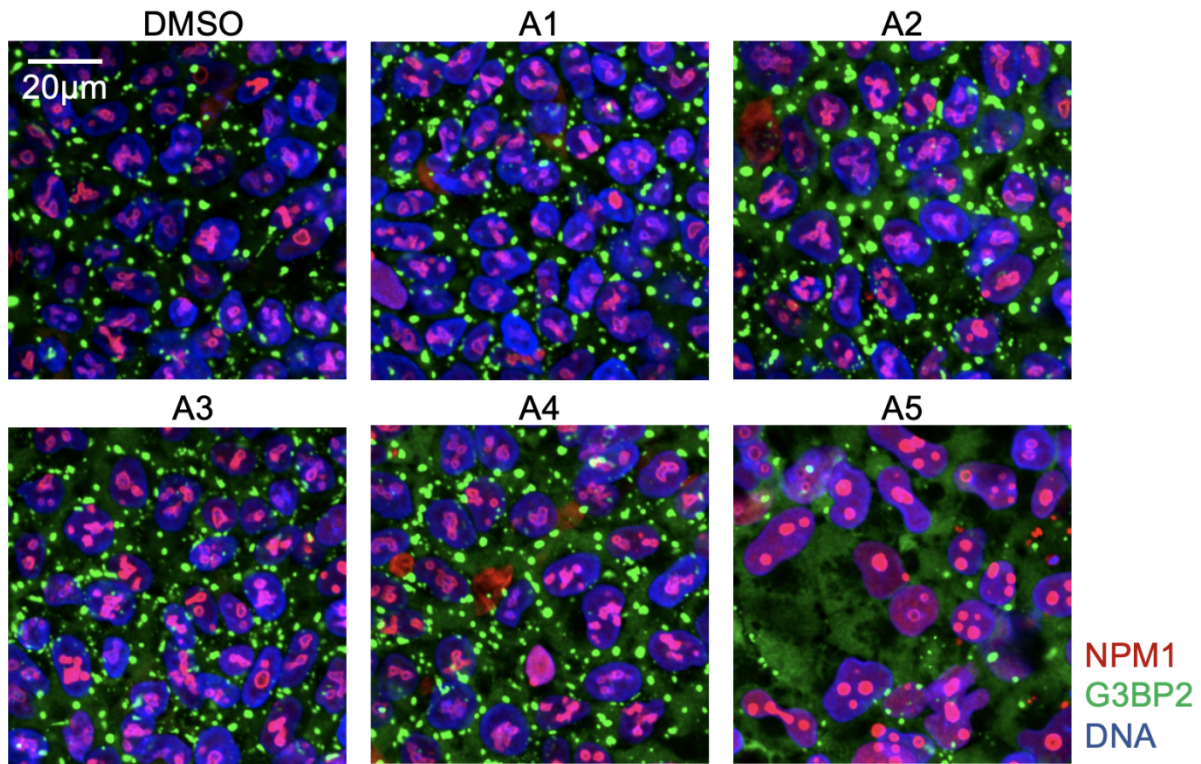
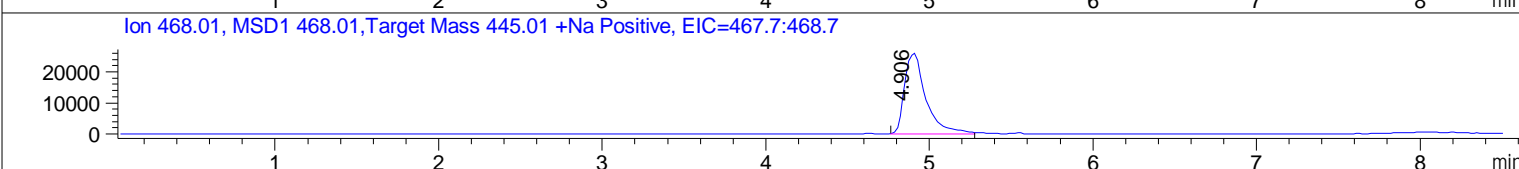
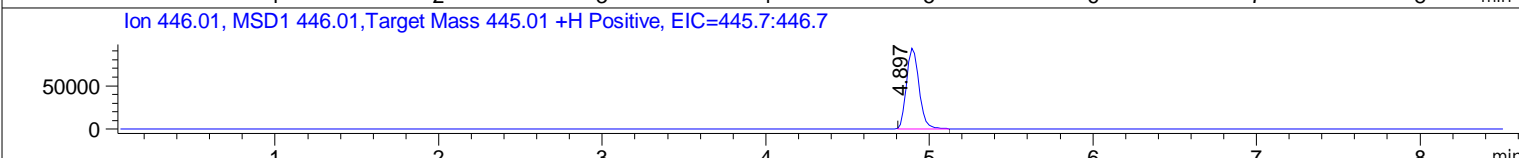
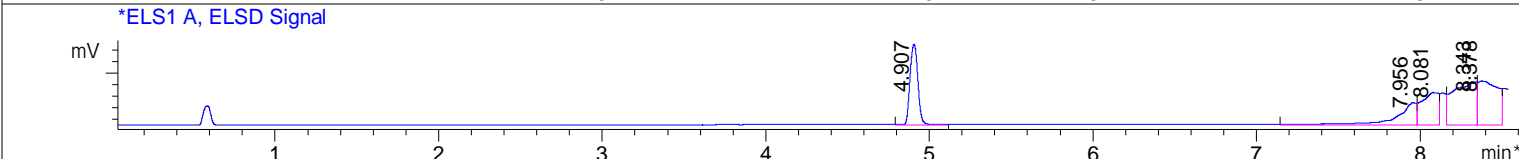
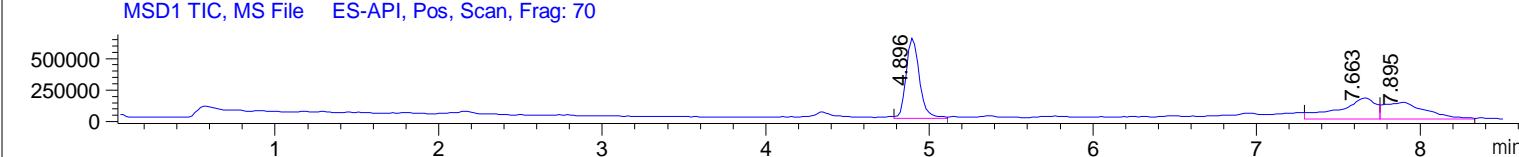
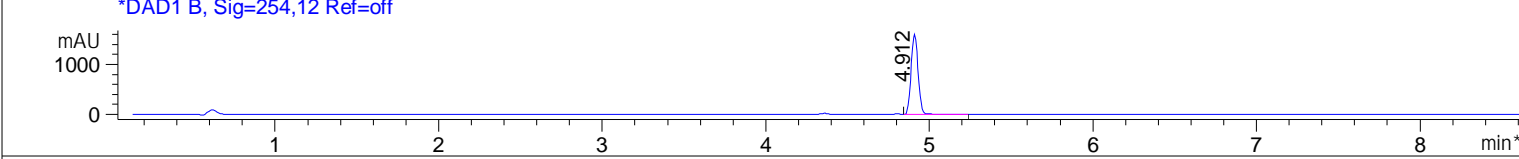
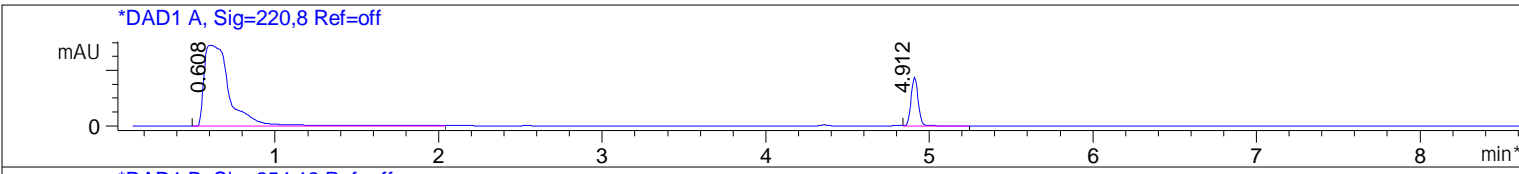


Figure S2. The effect of compound A1-A5 on other cellular LLPS compartments. To determine whether compounds A1-A5 modulates other LLPS phenomenon in the cells, Stress granules were induced and stress granule marker G3BP1 and nucleolus marker NPM1 were stained. DLD1 cells were treated with compounds A1-A5 and DMSO was used as control. After 24-hour compound treatment, cells were treated with 250µM sodium arsenite (flow chart on the top). Cells were fixed and stained with G3BP1 in green and NPM1 in red.

File ..mschem\03-23\270323-NCGC00137390-01-08038.D Tgt Mass (EZX): 445.01
 Injection Date : 27-Mar-23, 11:10:46 Seq. Line : 0
 Sample Name : NCGC00137390-01 Location : D2F-A1
 Acq. Operator : Hsiu-Ling Lin Inj : 1
 Spec. Reported : UV Integration Inj Volume : -3 ul
 Acq. Method : C:\Users\Public\Documents\ChemStation\1\Methods\FINAL_GRAD_NO_PRINT.M
 Analysis Method : C:\Users\Public\Documents\ChemStation\1\Methods\FINAL_GRAD_NO_PRINT.M
 Sample Info : 0050688158 WalkUp method: 'FINAL_GRD_NO PRINT' Mol Wt: 445.01
 Method Info : FINAL GRD BUT NO PRINT



Integration Results for DAD1 A, Sig=220,8 Ref=off

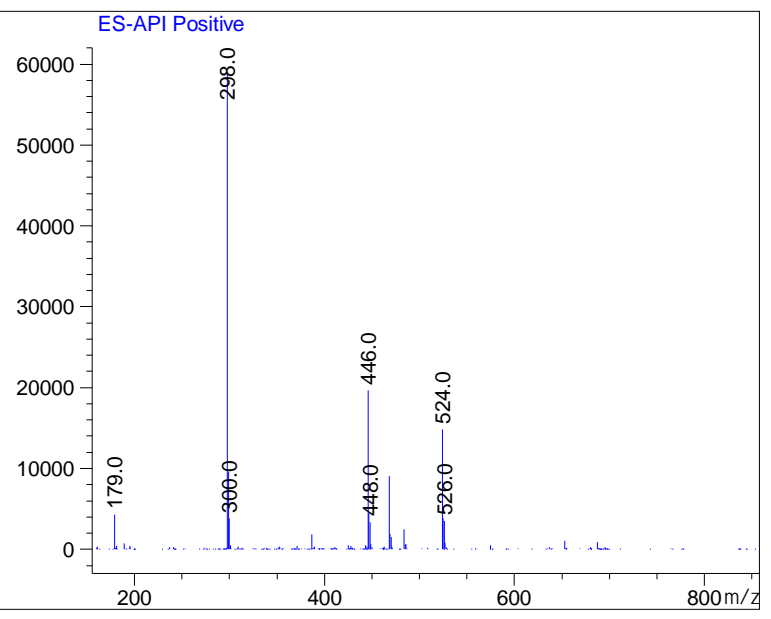
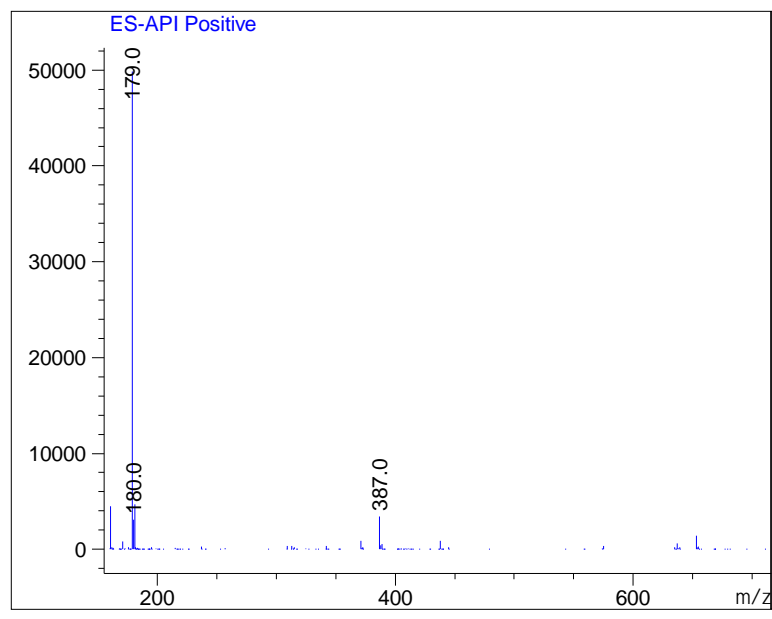
RetTim	Width	Area	Height	Area%	MS(+)
0.61	0.15	31303.33	2902.91	85.26	179
4.91	0.05	5413.07	1749.94	14.74	298

Integration Results for DAD1 B, Sig=254,12 Ref=off

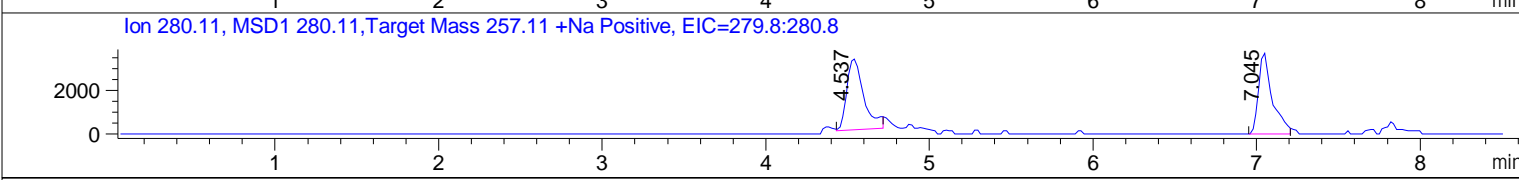
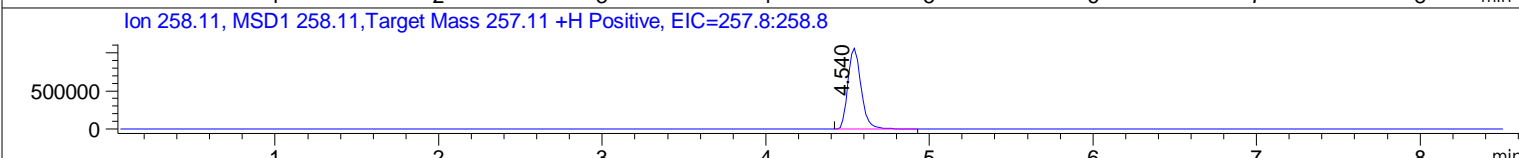
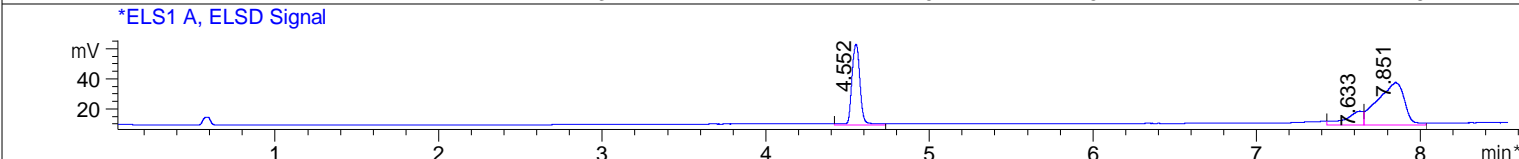
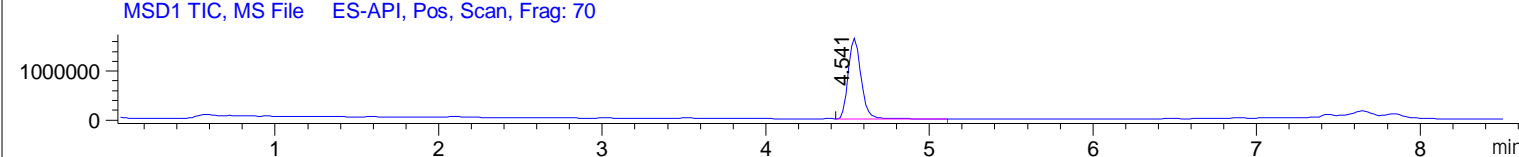
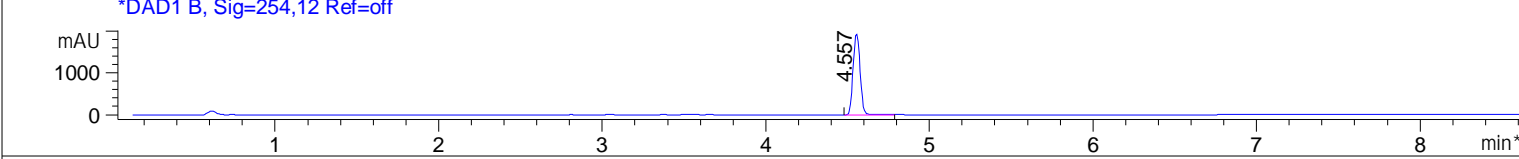
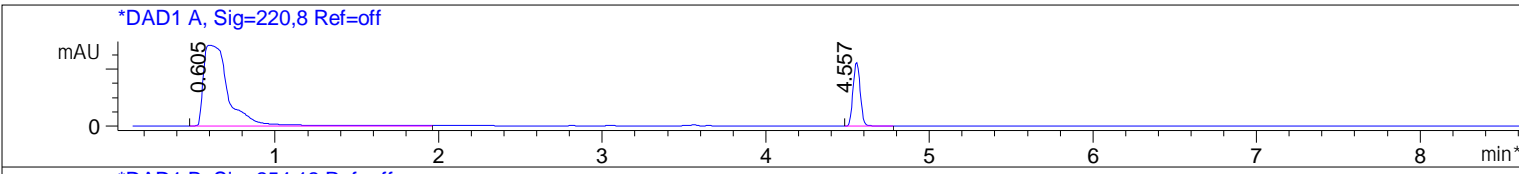
RetTim	Width	Area	Height	Area%	MS(+)
4.91	0.05	4843.56	1606.18	100.00	298

Ret. Time: 0.61 <<<< POSITIVE SPECTRA >>>>

Ret. Time: 4.91



File ..mschem\03-23\270323-NCGC00017285-04-08039.D Tgt Mass (EZX): 257.11
 Injection Date : 27-Mar-23, 11:20:57 Seq. Line : 0
 Sample Name : NCGC00017285-04 Location : D2F-B1
 Acq. Operator : Hsiu-Ling Lin Inj : 1
 Spec. Reported : UV Integration Inj Volume : -3 ul
 Acq. Method : C:\Users\Public\Documents\ChemStation\1\Methods\FINAL_GRAD_NO_PRINT.M
 Analysis Method : C:\Users\Public\Documents\ChemStation\1\Methods\FINAL_GRAD_NO_PRINT.M
 Sample Info : 0068079948 WalkUp method: 'FINAL_GRD_NO PRINT' Mol Wt: 257.11
 Method Info : FINAL GRD BUT NO PRINT



Integration Results for DAD1 A, Sig=220,8 Ref=off

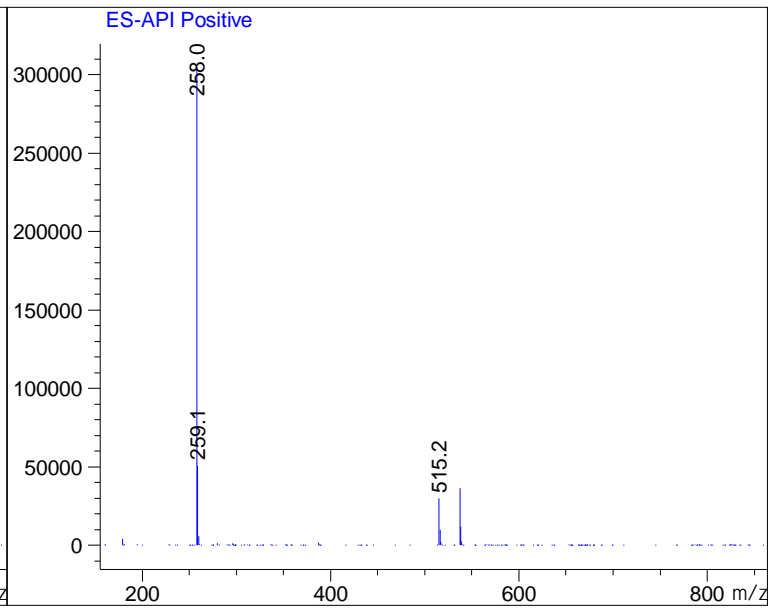
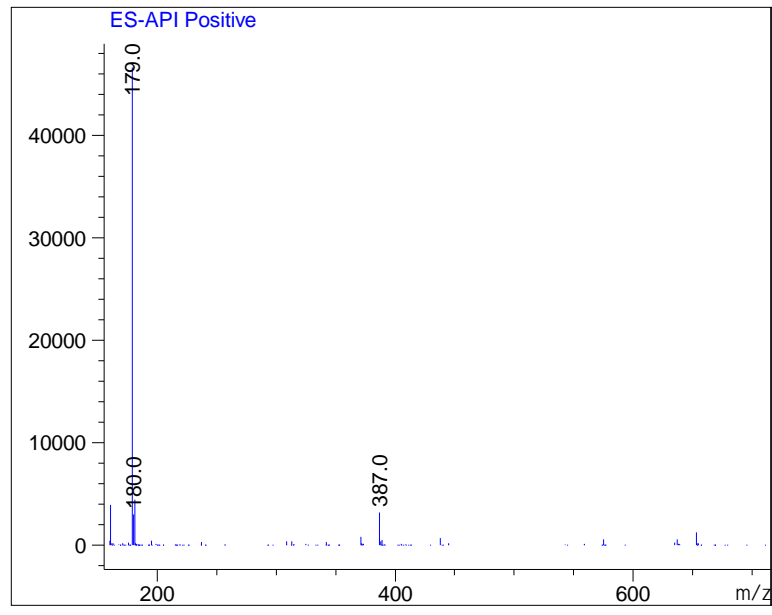
RetTim	Width	Area	Height	Area%	MS(+)
0.60	0.16	30177.99	2855.86	81.38	179
4.56	0.05	6905.01	2255.73	18.62	258

Integration Results for DAD1 B, Sig=254,12 Ref=off

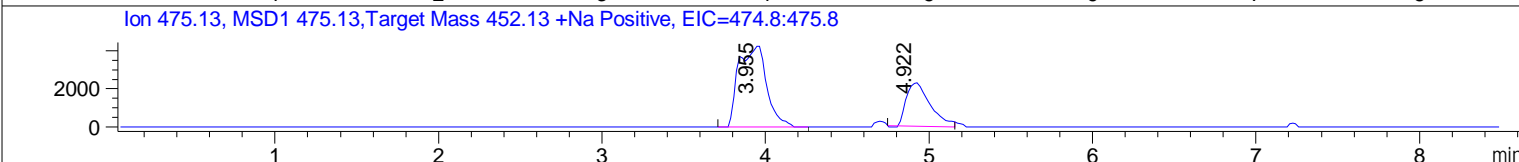
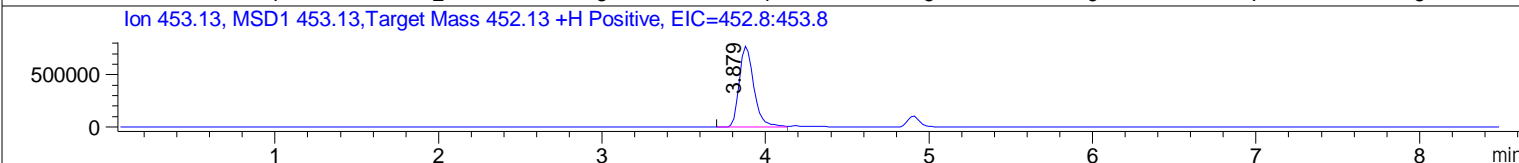
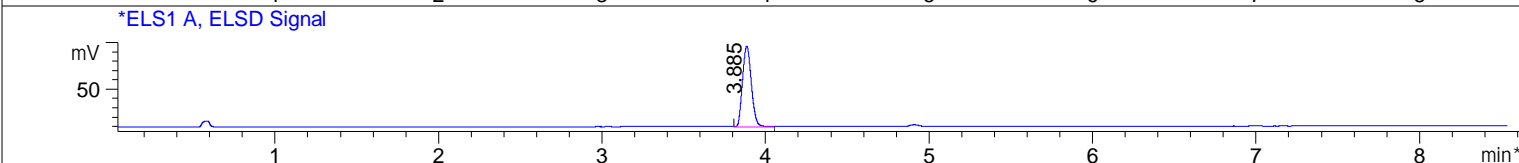
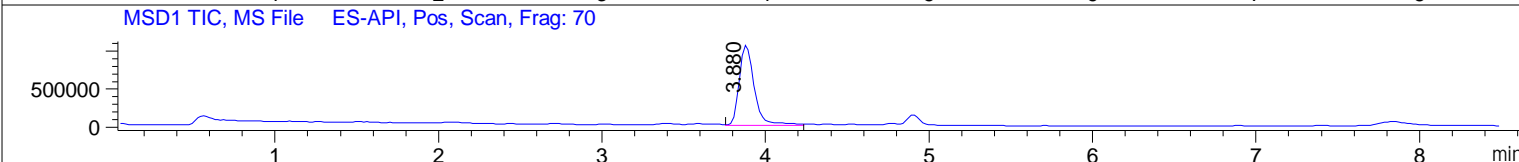
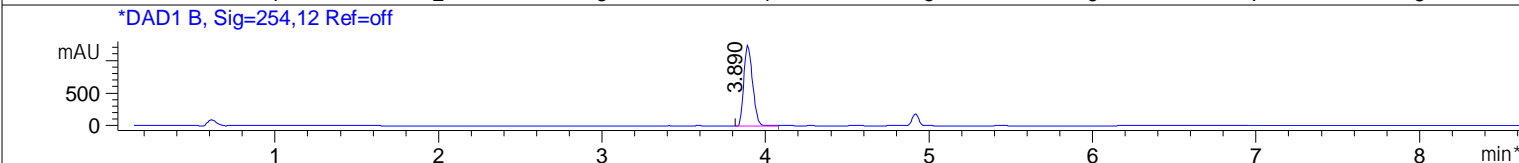
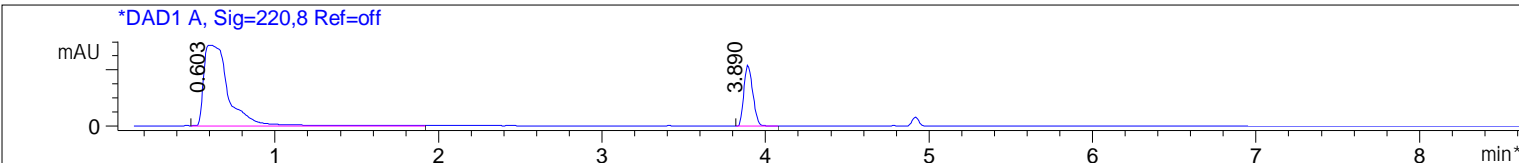
RetTim	Width	Area	Height	Area%	MS(+)
4.56	0.05	5763.82	1938.35	100.00	258

Ret. Time: 0.60 <<<< POSITIVE SPECTRA >>>>

Ret. Time: 4.56



File ..mschem\03-23\270323-NCGC00123946-01-08040.D Tgt Mass (EZX): 452.13
 Injection Date : 27-Mar-23, 11:30:56 Seq. Line : 0
 Sample Name : NCGC00123946-01 Location : D2F-C1
 Acq. Operator : Hsiu-Ling Lin Inj : 1
 Spec. Reported : UV Integration Inj Volume : -3 ul
 Acq. Method : C:\Users\Public\Documents\ChemStation\1\Methods\FINAL_GRAD_NO_PRINT.M
 Analysis Method : C:\Users\Public\Documents\ChemStation\1\Methods\FINAL_GRAD_NO_PRINT.M
 Sample Info : 0050708857 WalkUp method: 'FINAL_GRD_NO PRINT' Mol Wt: 452.13
 Method Info : FINAL GRD BUT NO PRINT



Integration Results for DAD1 A, Sig=220,8 Ref=off

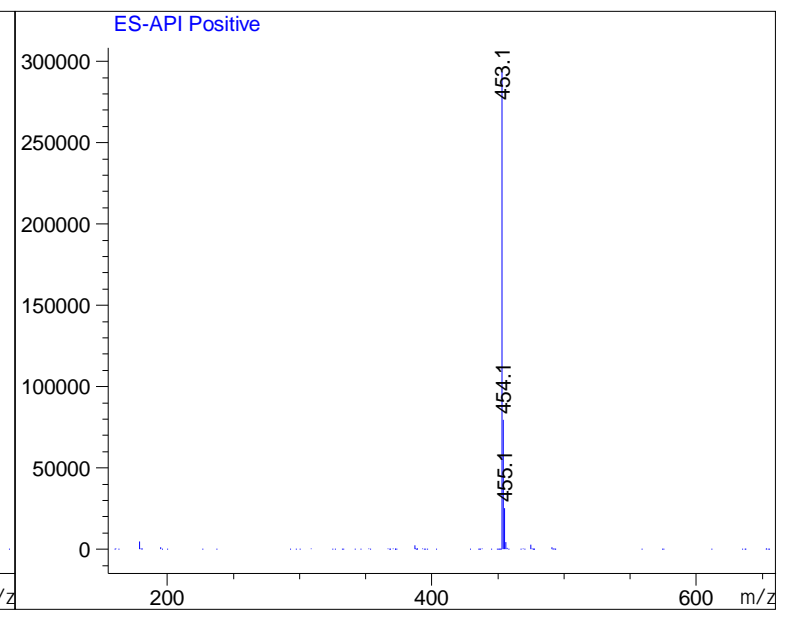
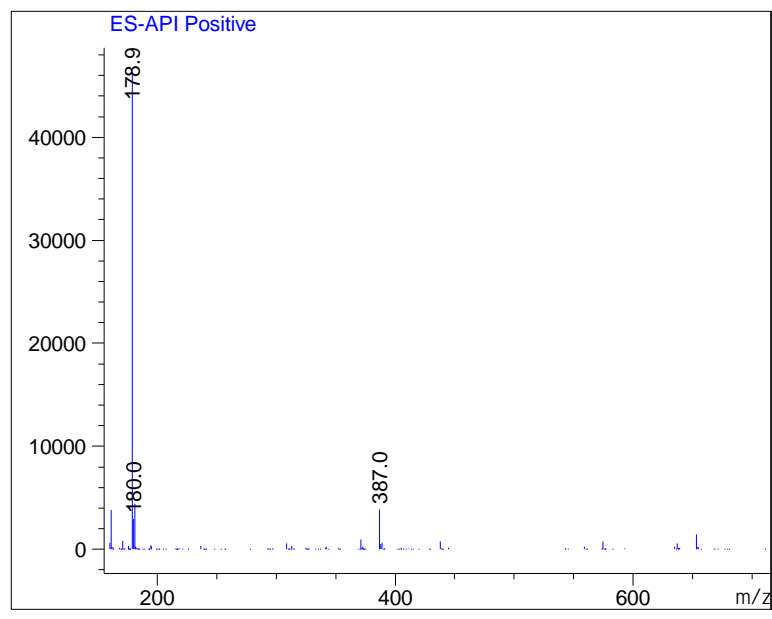
RetTim	Width	Area	Height	Area%	MS(+)
0.60	0.16	30328.42	2879.10	78.74	179
3.89	0.06	8190.48	2165.96	21.26	453

Integration Results for DAD1 B, Sig=254,12 Ref=off

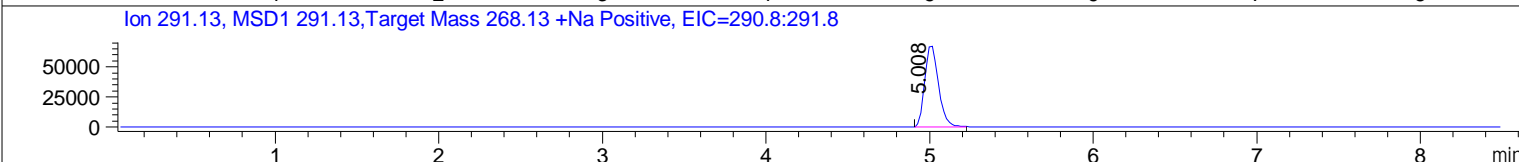
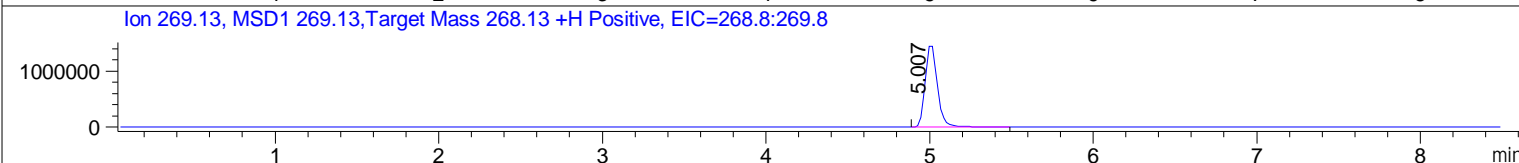
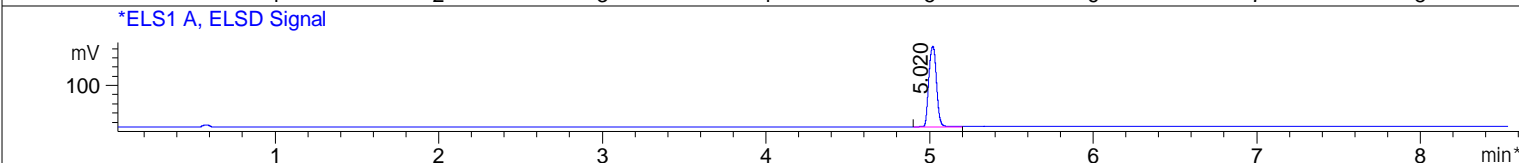
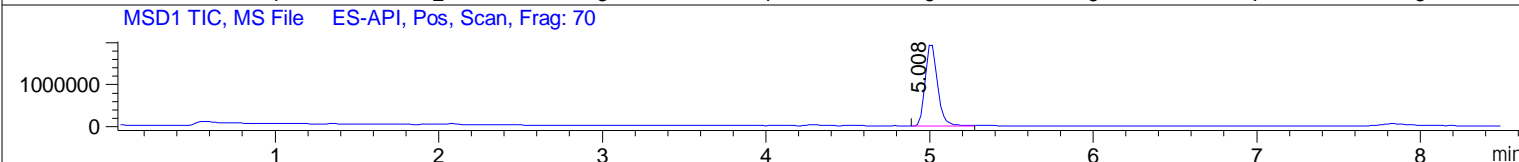
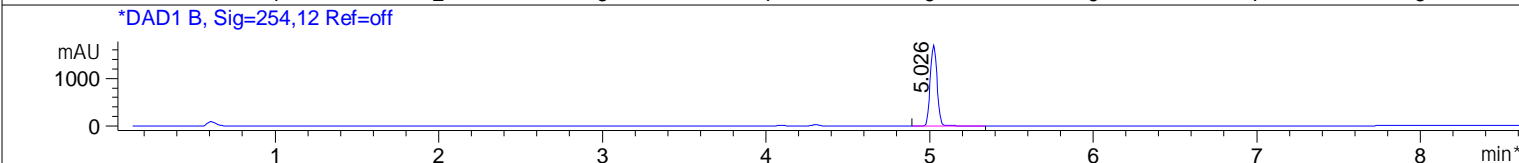
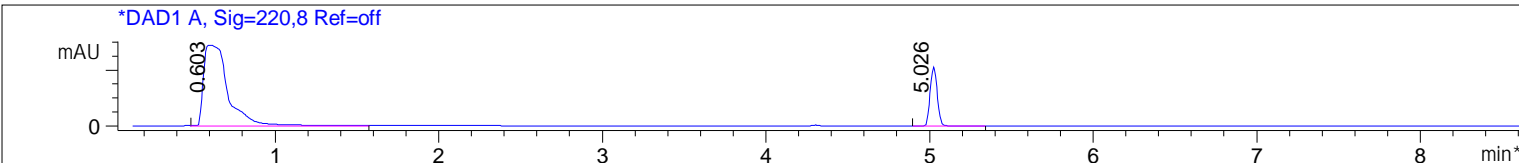
RetTim	Width	Area	Height	Area%	MS(+)
3.89	0.06	4565.96	1235.76	100.00	453

Ret. Time: 0.60 <<<< POSITIVE SPECTRA >>>>

Ret. Time: 3.89



File ..mschem\03-23\270323-NCGC00126599-01-08041.D Tgt Mass (EZX): 268.13
 Injection Date : 27-Mar-23, 11:40:54 Seq. Line : 0
 Sample Name : NCGC00126599-01 Location : D2F-D1
 Acq. Operator : Hsiu-Ling Lin Inj : 1
 Spec. Reported : UV Integration Inj Volume : -3 ul
 Acq. Method : C:\Users\Public\Documents\ChemStation\1\Methods\FINAL_GRAD_NO_PRINT.M
 Analysis Method : C:\Users\Public\Documents\ChemStation\1\Methods\FINAL_GRAD_NO_PRINT.M
 Sample Info : 0050684671 WalkUp method: 'FINAL_GRD_NO PRINT' Mol Wt: 268.13
 Method Info : FINAL GRD BUT NO PRINT



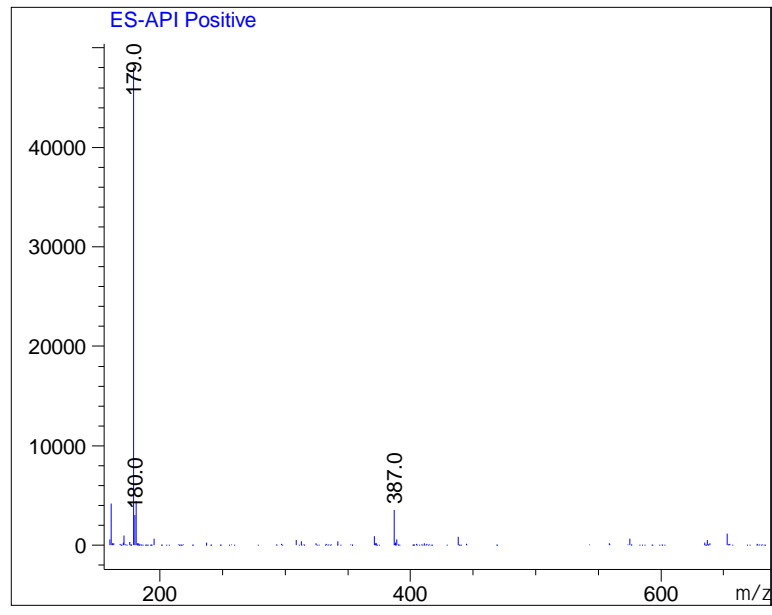
Integration Results for DAD1 A, Sig=220,8 Ref=off

RetTim	Width	Area	Height	Area%	MS(+)
0.60	0.16	30140.36	2897.92	82.48	179
5.03	0.05	6403.17	2118.90	17.52	269

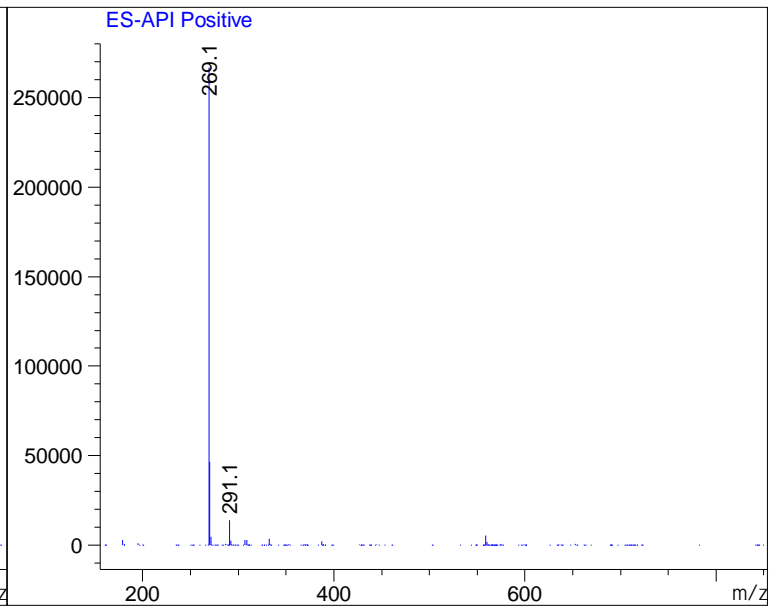
Integration Results for DAD1 B, Sig=254,12 Ref=off

RetTim	Width	Area	Height	Area%	MS(+)
5.03	0.05	4935.17	1709.14	100.00	269

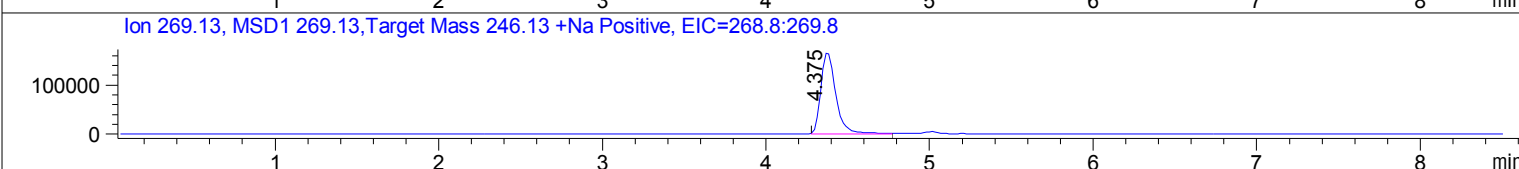
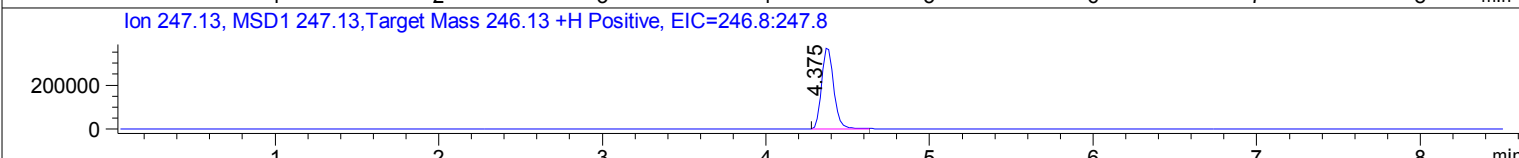
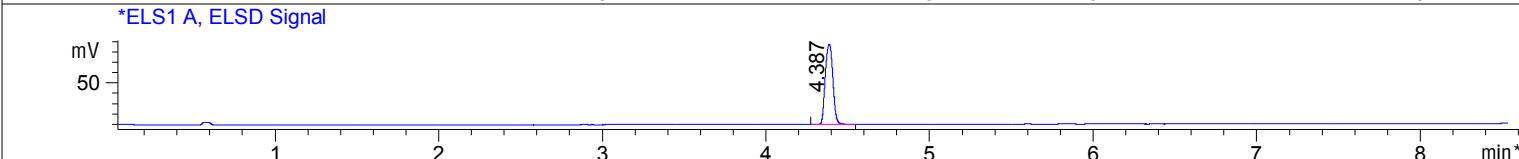
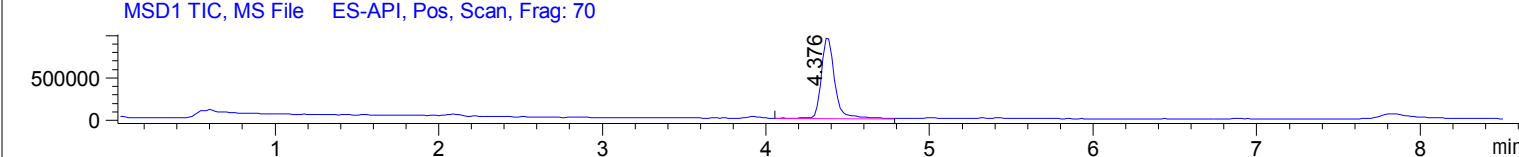
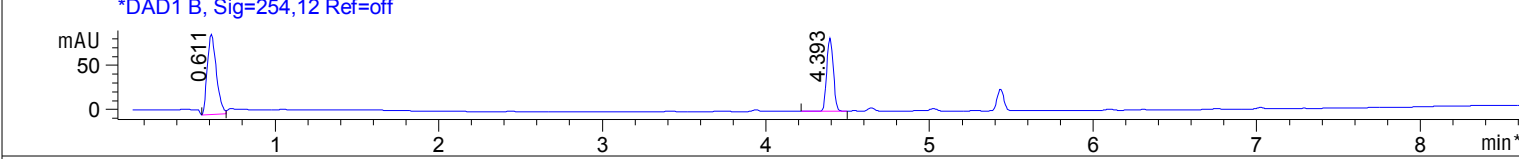
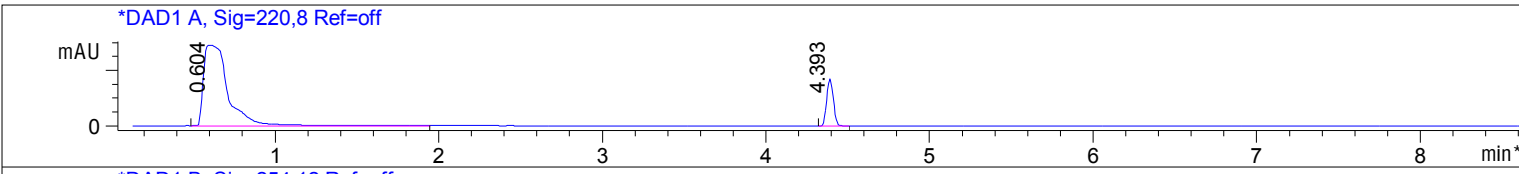
Ret. Time: 0.60 <<<< POSITIVE SPECTRA >>>>



Ret. Time: 5.03



File ..mschem\03-23\270323-NCGC00168931-02-08042.D Tgt Mass (EZX): 246.13
 Injection Date : 27-Mar-23, 11:50:54 Seq. Line : 0
 Sample Name : NCGC00168931-02 Location : D2F-E1
 Acq. Operator : Hsiu-Ling Lin Inj : 1
 Spec. Reported : UV Integration Inj Volume : -3 ul
 Acq. Method : C:\Users\Public\Documents\ChemStation\1\Methods\FINAL_GRAD_NO_PRINT.M
 Analysis Method : C:\Users\Public\Documents\ChemStation\1\Methods\FINAL_GRAD_NO_PRINT.M
 Sample Info : 0190189536 WalkUp method: 'FINAL_GRD_NO PRINT' Mol Wt: 246.13
 Method Info : FINAL GRD BUT NO PRINT



Integration Results for DAD1 A, Sig=220,8 Ref=off

RetTim	Width	Area	Height	Area%	MS (+)
0.60	0.16	30402.55	2907.19	86.30	179
4.39	0.05	4826.55	1679.70	13.70	247

Integration Results for DAD1 B, Sig=254,12 Ref=off

RetTim	Width	Area	Height	Area%	MS (+)
0.61	0.06	352.27	91.17	59.88	179
4.39	0.05	235.98	83.17	40.12	247

Ret. Time: 0.60 <<<< POSITIVE SPECTRA >>>>

Ret. Time: 4.39

