Mock Scenarios for Rapid Response Situations Involving Unknown Chemicals Using an NTA Approach

John T. Sloop, Alex Chao, Jennifer Gundersen, Allison L. Phillips, Jon R. Sobus, Elin M. Ulrich, Antony J. Williams, Seth R. Newton

Wednesday, September 22, 2021

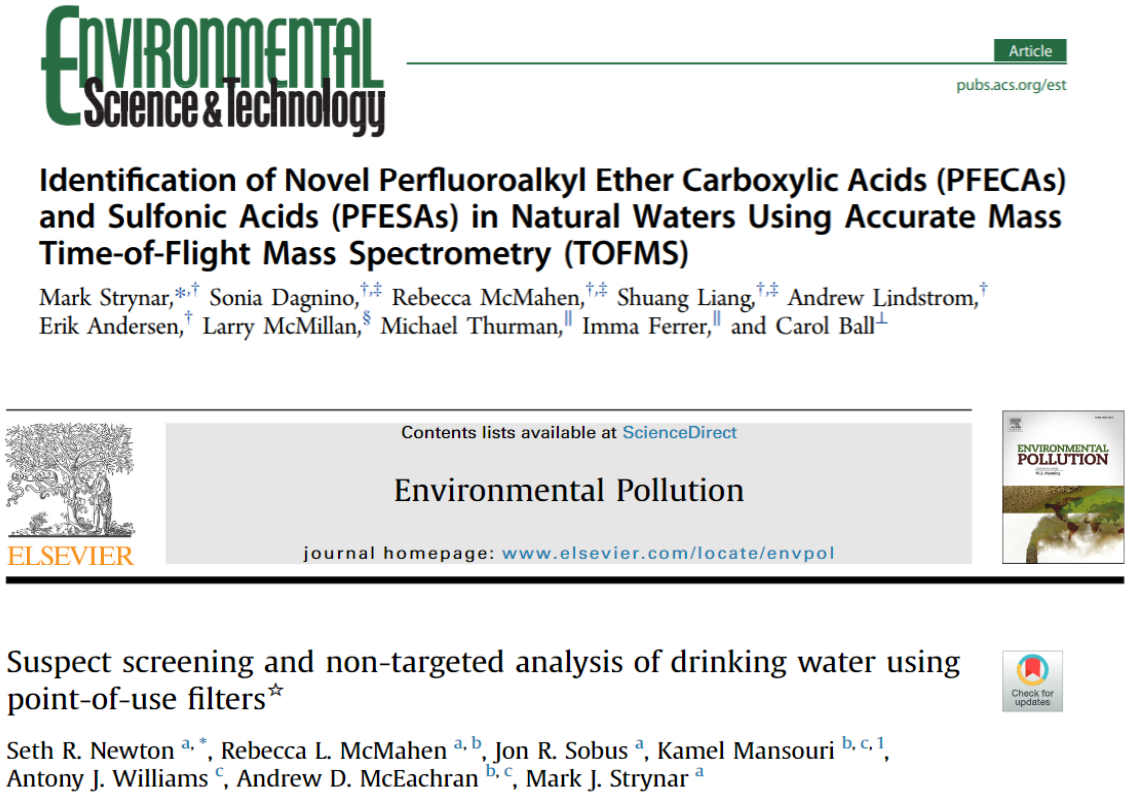
Background of Rapid Response

- National Response Center (NRC) and EPA's OLEM (Office of Land and Emergency Management)
  - In 2019 alone, over 26,000 calls
  - 37% - unknown composition
- ERLN and PHILIS
  - Targeted analytical approaches
- How we can help:
  - Non-targeted analysis (NTA) with our workflows and informatics tools

Phillips et al., ET&C 2021.

Background of Non-Targeted Analysis

- Non-targeted analysis (NTA)
- Identification of unknowns
- Processing data generated by high-resolution mass spectrometry instruments
- Utilizes both MS_1 and MS_2 data for feature identification
- Historically, this is a very slow, time-consuming workflow
  - Hope to speed this process with automation and various new tools



Environmental Science & Technology

Article

pubs.acs.org/est

Identification of Novel Perfluoroalkyl Ether Carboxylic Acids (PFECAs) and Sulfonic Acids (PFESAs) in Natural Waters Using Accurate Mass Time-of-Flight Mass Spectrometry (TOFMS)

Mark Strynar,*,† Sonia Dagnino,†,‡ Rebecca McMahen,†,‡ Shuang Liang,†,‡ Andrew Lindstrom,† Erik Andersen,† Larry McMillan,§ Michael Thurman,∥ Imma Ferrer,∥ and Carol Ball⊥

Contents lists available at ScienceDirect

Environmental Pollution

journal homepage: www.elsevier.com/locate/envpol

ELSEVIER

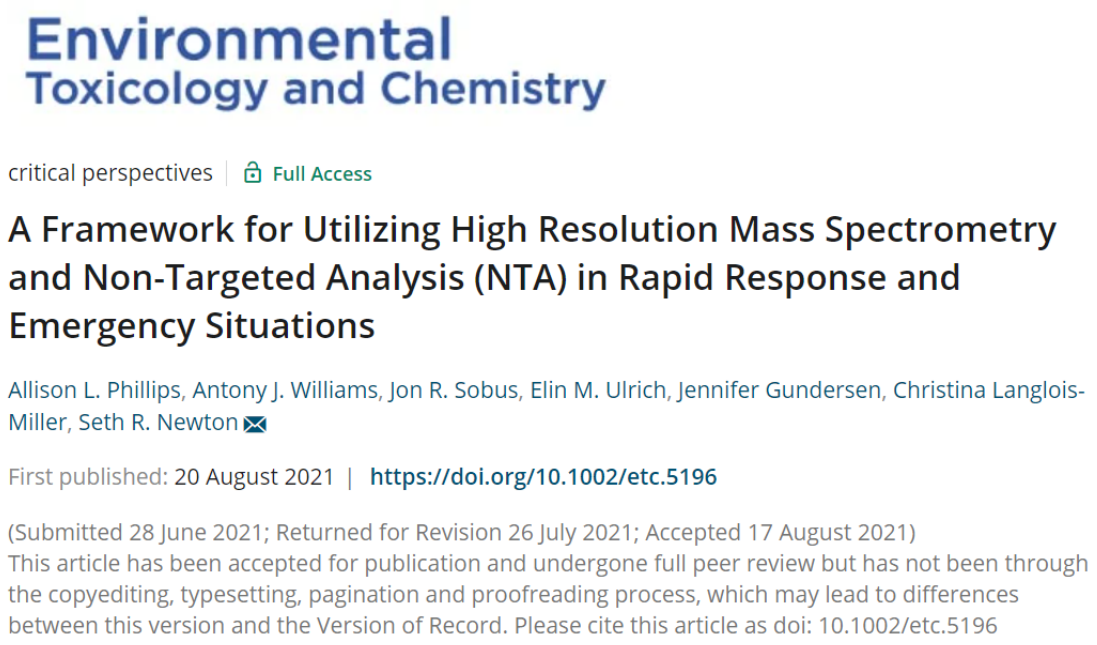
Suspect screening and non-targeted analysis of drinking water using point-of-use filters☆

Seth R. Newton [a, *], Rebecca L. McMahen [a, b], Jon R. Sobus [a], Kamel Mansouri [b, c, 1], Antony J. Williams [c], Andrew D. McEachran [b, c], Mark J. Strynar [a]

Strynar et al., ES&T 2015, 49, 11622-11630. Newton et al., Environ. Poll. 2018, 234, 297-306.

Rapid Response Framework Paper

- Laid the framework of how NTA could be applied in the field of Rapid Response (RR)
- HRMS vs. traditional, low-resolution instrumentation
- NTA has been proven as a tool for identifying unknowns
  - Even in rapid response scenarios



Environmental
Toxicology and Chemistry

critical perspectives | Full Access

A Framework for Utilizing High Resolution Mass Spectrometry and Non-Targeted Analysis (NTA) in Rapid Response and Emergency Situations

Allison L. Phillips, Antony J. Williams, Jon R. Sobus, Elin M. Ulrich, Jennifer Gundersen, Christina Langlois-Miller, Seth R. Newton

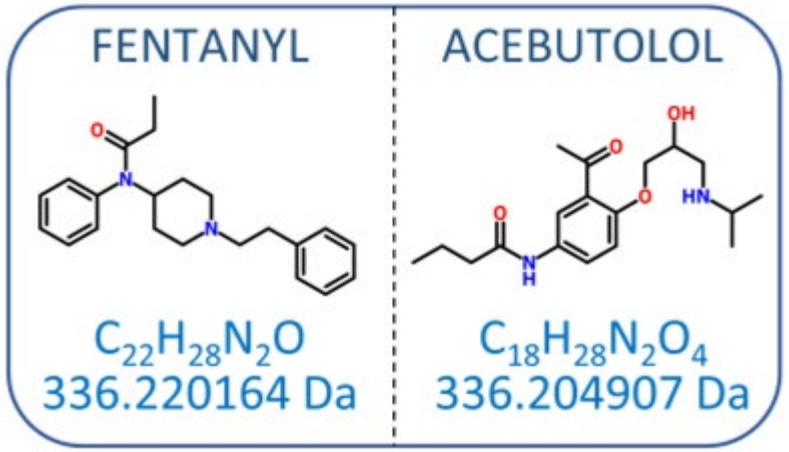
First published: 20 August 2021 | https://doi.org/10.1002/etc.5196

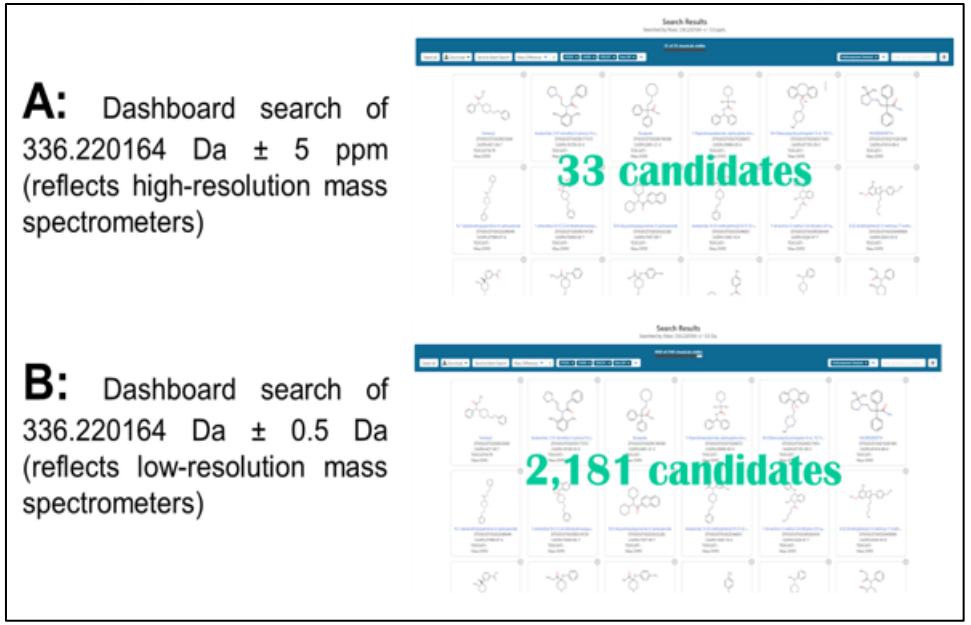
(Submitted 28 June 2021; Returned for Revision 26 July 2021; Accepted 17 August 2021)
This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/etc.5196

Phillips et al., ET&C 2021. Strynar et al., ES&T 2015, 49, 11622-11630. Newton et al., Environ. Poll. 2018, 234, 297-306. Hollender et al., 2017. Bader et al., 2016.

NTA as a tool for RR

- Rapid screening of common suspects
- Identification of unknowns
- Informatics tools at EPA

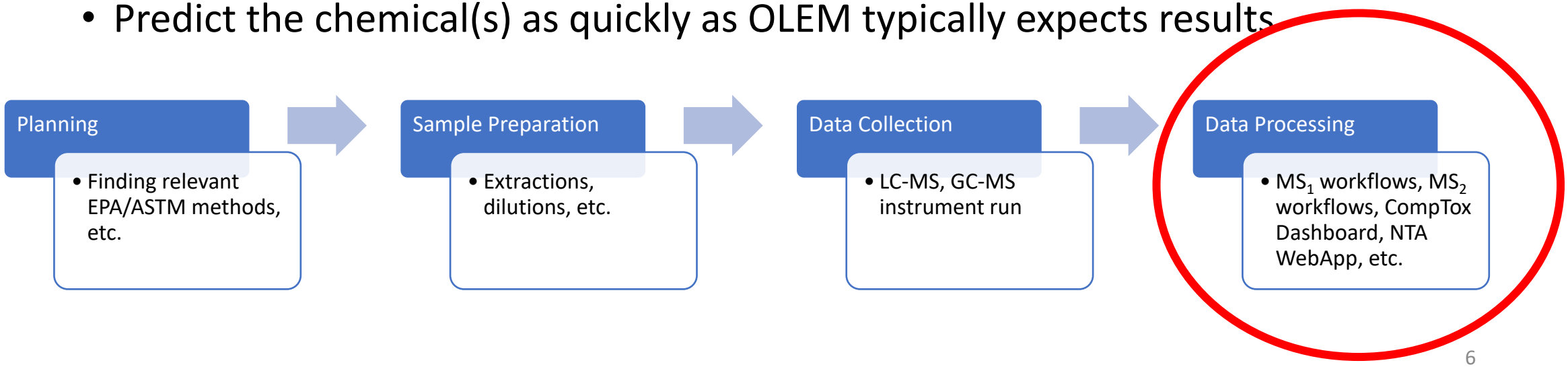




Phillips et al., ET&C 2021.

Goals of the RR mock scenarios

- Create multiple scenarios in which some type of “chemical spill” is mimicked
- Using our instrumental analyses and data processing workflows, show we can:
  - Correctly identify compound(s) of interest
  - Predict the chemical(s) as quickly as OLEM typically expects results



List of currently performed mock scenarios

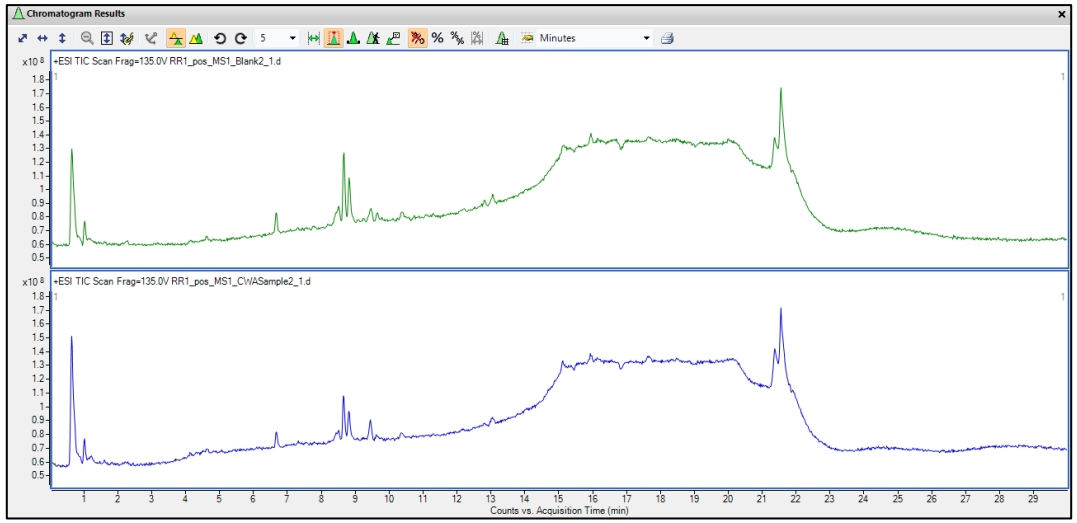
- *Method development mock scenarios:*
  1. Unidentified chemical warfare agent (CWA) released inside a building
  2. *Different* unidentified CWA released inside a building
- *“Real” mock scenarios:*
  3. Novichok nerve agent used to poison foreign operative via spiking into alcoholic beverage
  4. AFFF spill in river water

1st Mock Scenario:
CWA released inside building

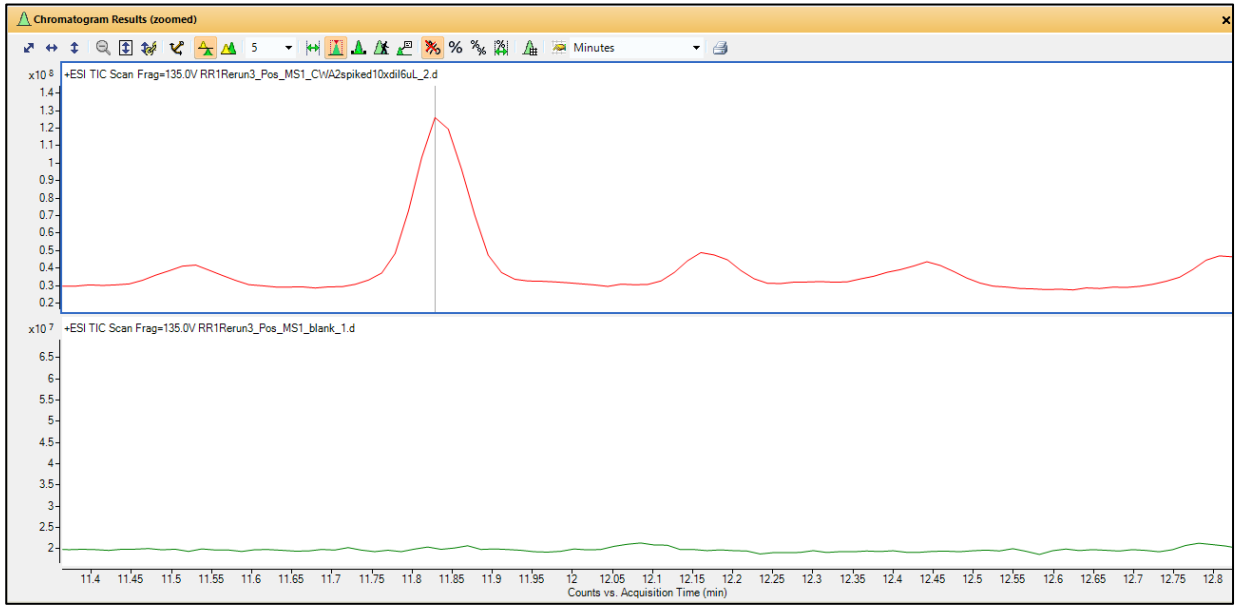
METHOD 3572

EXTRACTION OF CHEMICAL AGENTS FROM WIPE SAMPLES USING MICROEXTRACTION

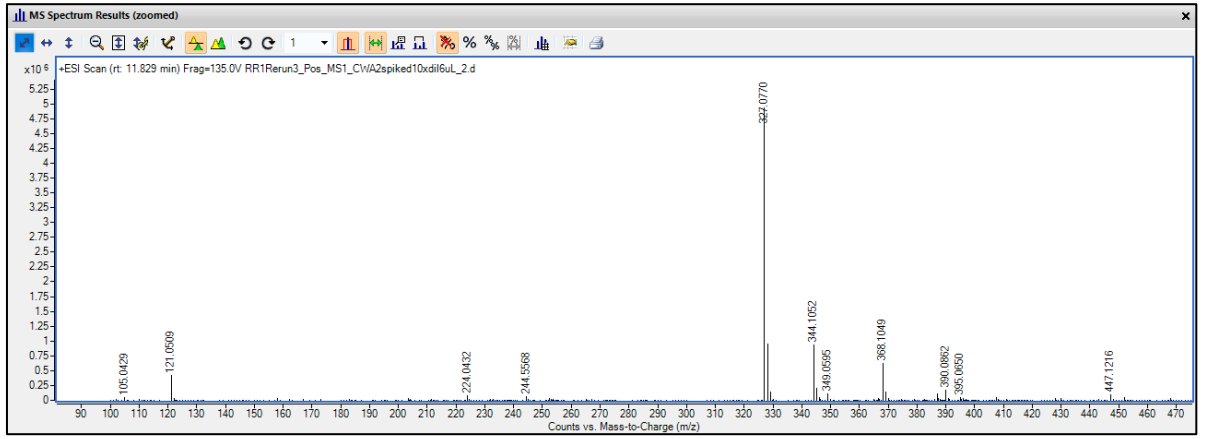
SW-846 is not intended to be an analytical training manual. Therefore, method procedures are written based on the assumption that they will be performed by analysts who are formally trained in at least the basic principles of chemical analysis and in the use of the subject technology.



Comparison of Sample vs. Blank chromatogram using Qualitative Navigator



Comparison of Sample vs. Blank chromatogram using Qualitative Navigator

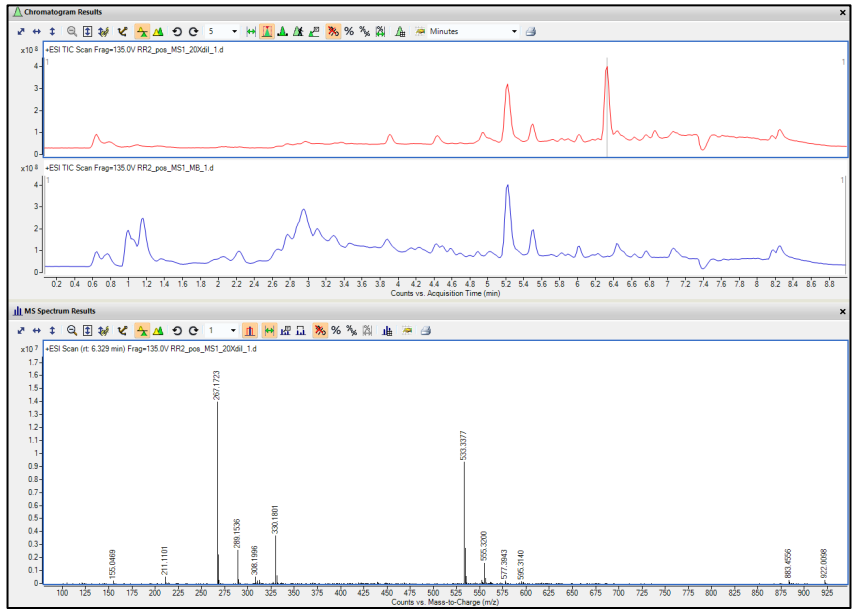


Extracted MS spectrum of sample chromatogram at RT 11.829 min

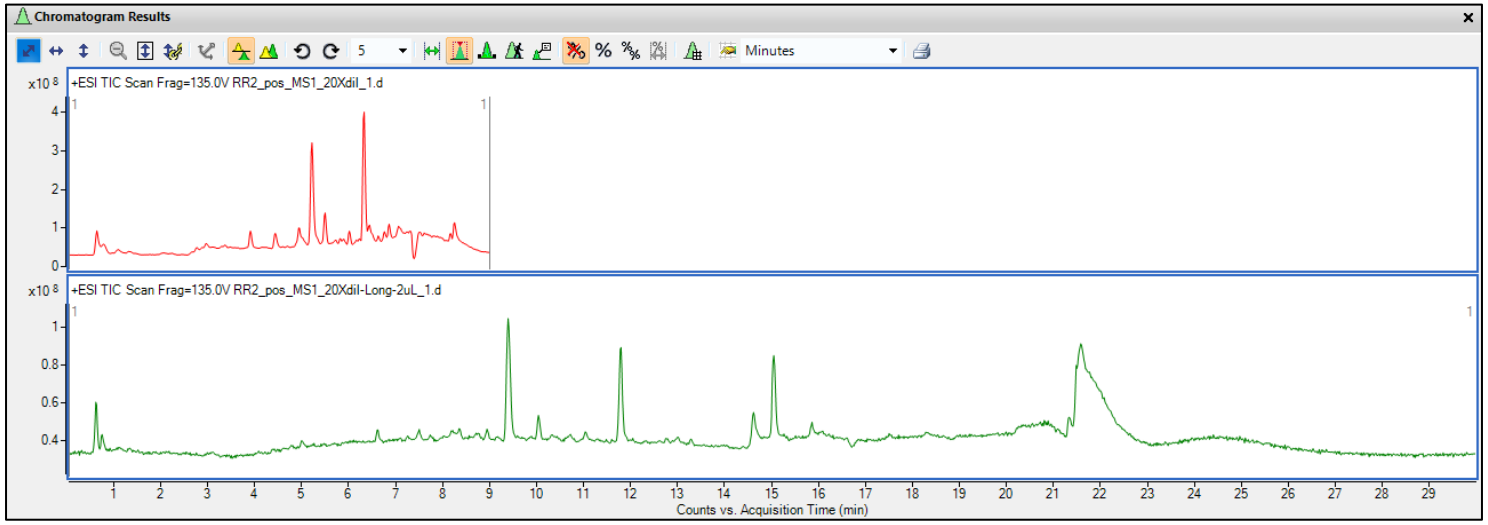
Important takeaway: Need better method for range finding!

2nd Mock Scenario

- Goals:
  - Test a rapid range-finding method
  - Start to determine appropriate workflows for MS_1 and MS_2 data
- Range finding method
  - Shorter, 8-minute LC-MS run for MS_1 analysis
  - Used to:
    - Determine appropriate concentration/dilution factor
    - Determine appropriate polarity (ESI+ or ESI-)
  - It works!
- **Tributyl Phosphate**



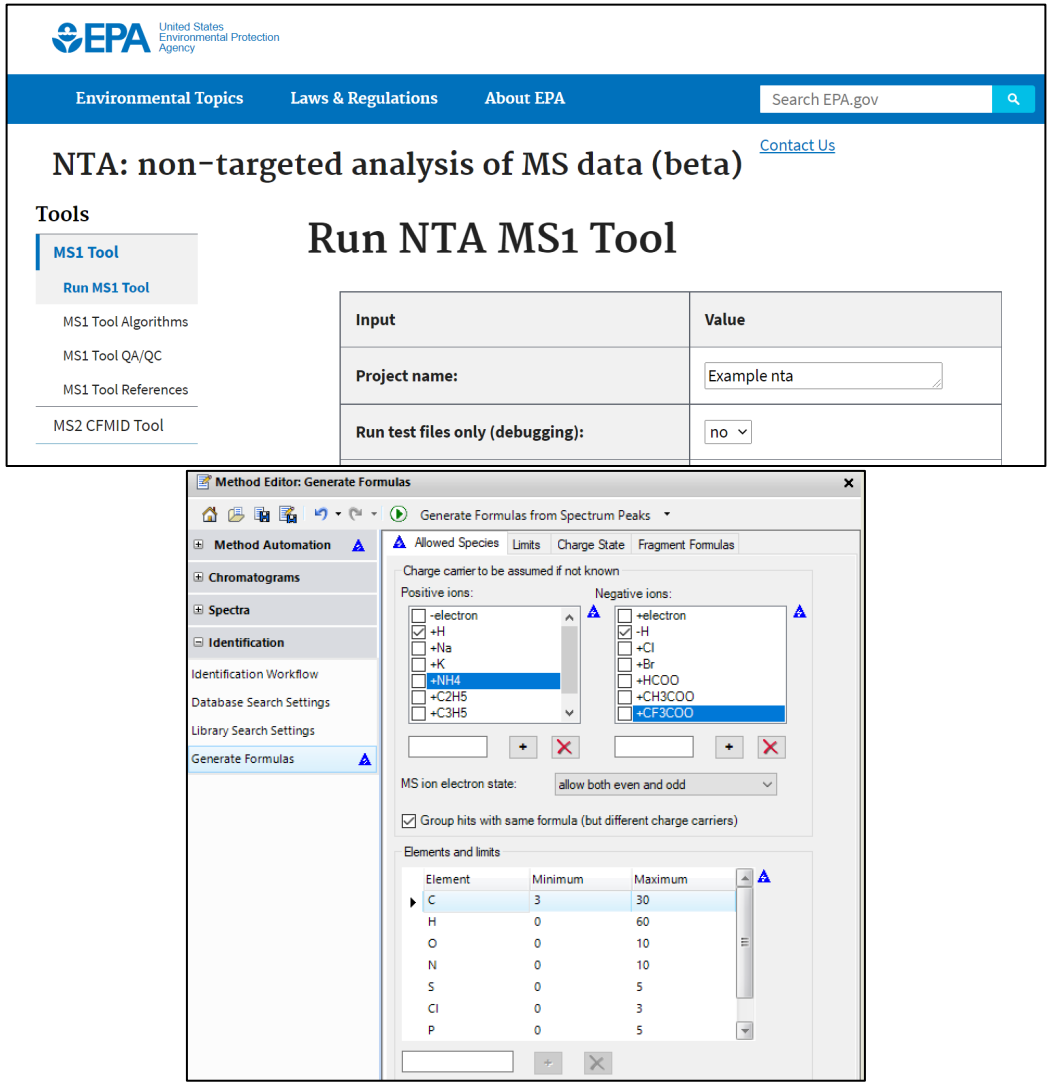
Comparison of Sample vs. Blank chromatogram using Qualitative Navigator, with extracted MS spectrum of sample at RT 6.329 min (using faster, range finding method)



Short vs. Long LC-MS method sample chromatograms

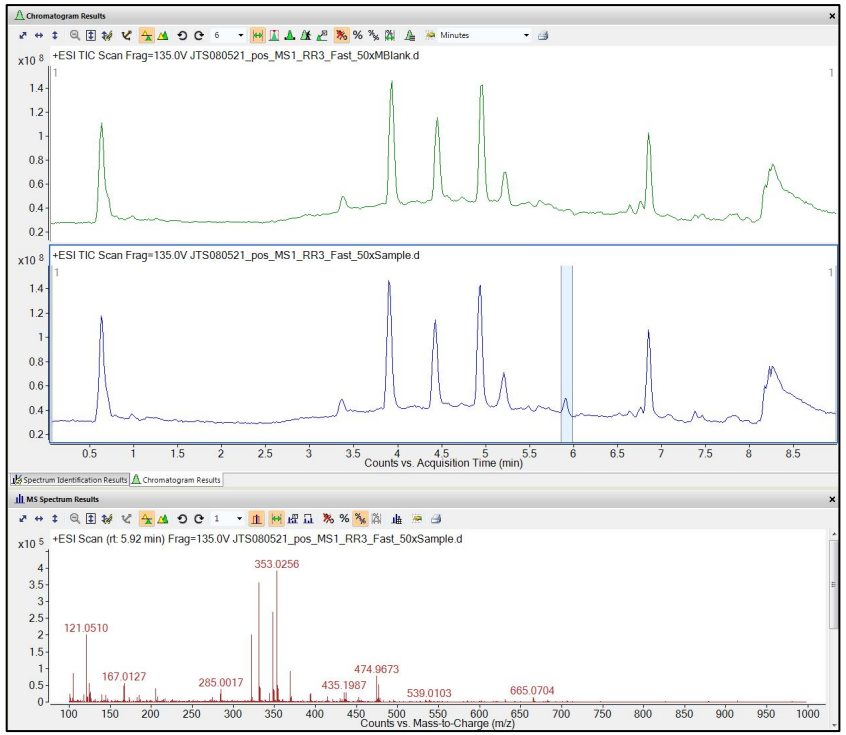
2nd Mock Scenario

- Data processing workflows:
  - MS_1:
    - Formula matching to MS Ready Formula (Dashboard)
    - WebApp Search by Mass
    - Molecular formula prediction
  - MS_2:
    - Matching to spectral libraries
    - WebApp match to CFM-ID in-silico database

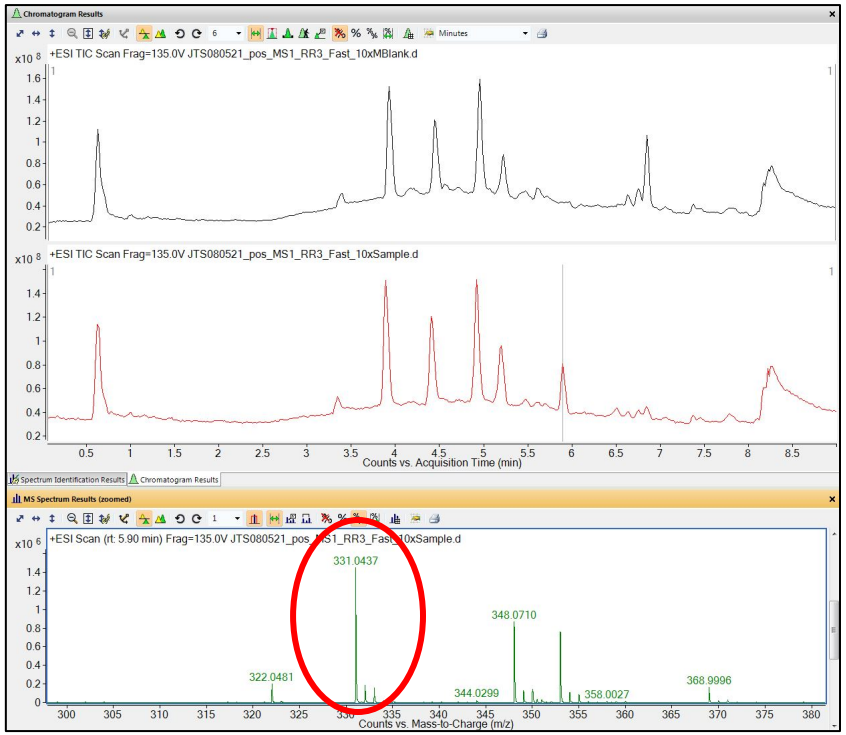


3rd Mock Scenario: Foreign operative poisoned with nerve agent

- Nerve agent, similar to Novichok nerve agents used in attacks in 2019, used to poison alcoholic beverage of foreign operative
  - 1000x – 10x dilution



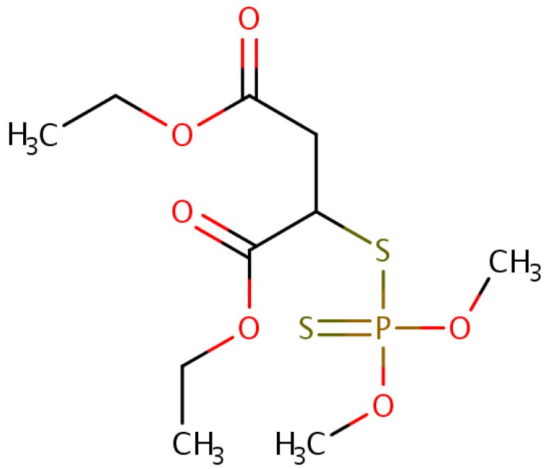
50x dilution (353.0256)



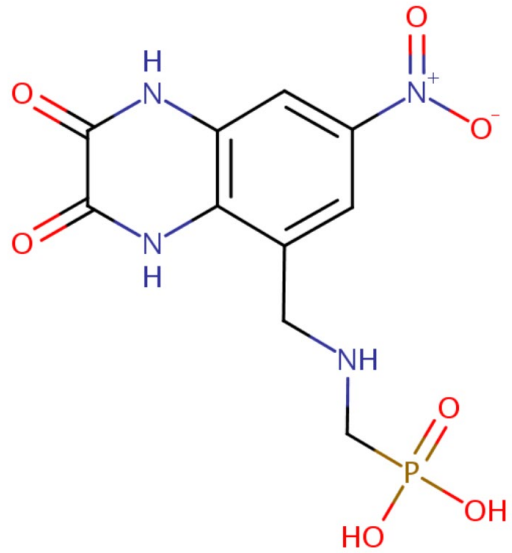
10x dilution (331.0437)

3rd Mock Scenario

- MS_1 Results
- **Formula matching to MS-Ready formula**
  - $C_{10}H_{19}O_6PS_2$, scored 89.2
- **NTA WebApp by Mass**
  - 49 hits for mass 330.0365
    - Top hits on Data Source hits:
      - N=250, Malathion ($C_{10}H_{19}O_6PS_2$)
      - N=33, Isomalathion ($C_{10}H_{19}O_6PS_2$)
      - N=17, Becampanel ($C_{10}H_{11}N_4O_7P$)
- **Molecular formula prediction (MFG)**
  - Top hit = $C_{10}H_{19}O_6PS_2$, score of 99.11



Malathion ($C_{10}H_{19}O_6PS_2$)



Becampanel ($C_{10}H_{11}N_4O_7P$)

Sobus et al., Anal. Bioanal. Chem. 2019, 411, 835-851. McEachran et al., Anal. Bioanal. Chem. 2017, 409, 1729-1735.

3rd Mock Scenario

- MS_2 Results
- **Matching to spectral libraries (MS_2)**
  - Two hits, very low scores for both
    - 25.48 and 27.32
- **Match to CFM-ID in-silico database (using WebApp)**
  - Compound did not fragment well during experimental MS_2 instrument run

	A	B	C	D	E	F	G	H	I	J
1		MASS	DTXCID	energy0	energy1	energy2	energy_sum	SCORE	MATCHES	ASS_in_MGF
35	60	330.0386	DTXCID101468832	0.087187	0.025464	4.54E-05	0.112696472	0.38967	2	331.0437
36	66	330.0353	DTXCID30120528	0.090655	0.018483	0.002036	0.111173803	0.384405	1	331.0437
37	54	330.0376	DTXCID9039286	0.081561	0.025079	0.001431	0.108070872	0.373676	11	331.0437
38	45	330.0368	DTXCID00253039	0.077258	0.024637	0.000909	0.10280428	0.355465	12	331.0437
39	63	330.0362	DTXCID20923065	0.087711	0.009323	0.000359	0.09739334	0.336756	2	331.0437
40	36	330.0338	DTXCID10944941	0.05986	0.031783	0.004535	0.096178137	0.332554	2	331.0437
41	51	330.0365	DTXCID0094732	0.079164	0.014115	3.86E-05	0.093317551	0.322663	3	331.0437
42	48	330.0342	DTXCID30292072	0.078537	0.007852	7.22E-05	0.086461002	0.298955	5	331.0437
43	18	330.0376	DTXCID40305006	0.056618	0.02121	0.001098	0.07892695	0.272905	11	331.0437
44	21	330.0376	DTXCID60936755	0.056658	0.021175	0.001046	0.078878876	0.272739	11	331.0437
45	30	330.0342	DTXCID40736013	0.059408	0.013258	0.000622	0.07328803	0.253407	5	331.0437
46	33	330.0376	DTXCID70366544	0.059526	0.012034	0.000349	0.071909439	0.248641	11	331.0437
47	27	330.0376	DTXCID90422015	0.058559	0.011936	0.000292	0.070787533	0.244761	11	331.0437
48	42	330.0365	DTXCID20950093	0.064721	0.00454	0.00025	0.069511138	0.240348	3	331.0437
49	24	330.0376	DTXCID20437450	0.058348	0.010627	0.000231	0.069206914	0.239296	11	331.0437
50	39	330.0344	DTXCID201381768	0.063693	0.002811	6.56E-05	0.066569781	0.230178	1	331.0437
51	12	330.0394	DTXCID90735511	0.046751	0.012217	4.06E-05	0.059008426	0.204033	2	331.0437
52	15	330.0368	DTXCID30320041	0.052302	0.005887	1.52E-05	0.05820453	0.201253	12	331.0437
53	6	330.0361	DTXCID8032099	0.034243	0.006696	0.000448	0.041386966	0.143103	2	331.0437
54	9	330.0365	DTXCID301307347	0.036081	0.002364	7.46E-05	0.03851932	0.133188	3	331.0437
55	3	330.0376	DTXCID40552159	0.027685	0.007708	0.000324	0.035717031	0.123498	11	331.0437
56	0	330.0361	DTXCID80791	0.014177	0.009685	0.000145	0.024006415	0.083007	2	331.0437

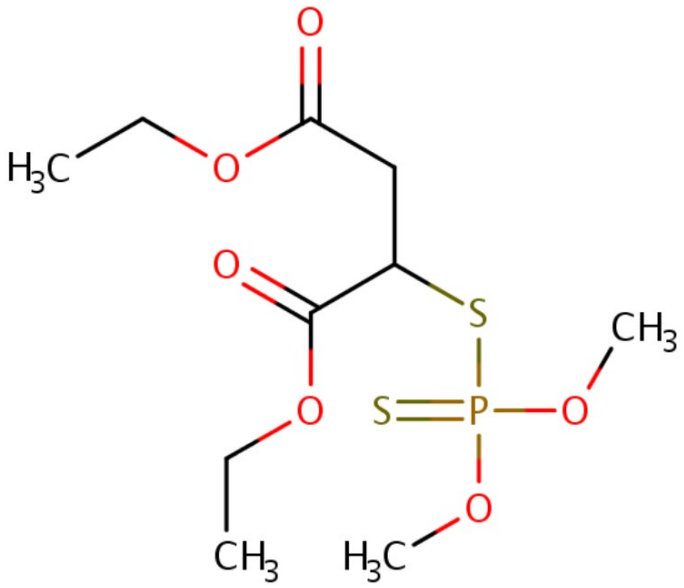
WebApp CFM-ID matches

	A	B	C	D	E	F	G	H	I	
1	INPUT	FOUND_BY	DTXSID	PREFERRED_NAME	DTXCID	IUPAC_NAI	MOLECULA	MONOISOT	DATA_SOURCES	ATM
2	DTXCID80791	DTXCID	DTXSID4020	Malathion	DTXCID80	Diethyl 2-[(d	C10H19O6F	330.03607	209	7.7
3	DTXCID8032099	DTXCID	DTXSID9058	Isomalathion	DTXCID80	Diethyl 2-[(n	C10H19O6F	330.03607	28	7.8
4	DTXCID9039286	DTXCID	DTXSID0068	Benzoic acid, 2,2'-(dioxydica	DTXCID90	-	C16H10O8	330.03757	16	8.0
5	DTXCID0094732	DTXCID	DTXSID5017	Becampanel	DTXCID00	([(7-Nitro-2,	C10H11N4C	330.03654	16	1.0
6	DTXCID4089361	DTXCID	DTXSID8016	NS 1608	DTXCID40	N-(5-Chloro-	C14H10ClF3	330.03829	13	1.
7	DTXCID4099409	DTXCID	DTXSID2017	3,3'-Di-O-methylellagic acid	DTXCID40	2,7-Dihydro>	C16H10O8	330.03757	10	2.2
8	DTXCID30120528	DTXCID	DTXSID0019	Benzylidene bis(dimethyldith	DTXCID30	Phenylmeth	C13H18N2S	330.03528	9	2.3
9	DTXCID00256060	DTXCID	DTXSID5030	Diselenide, dihexyl	DTXCID00	Dihexyldisel	C12H26Se2	330.0365	8	2.8
10	DTXCID8035683	DTXCID	DTXSID2066	Cyclotetrasiloxane, (chlorom	DTXCID80	2-(Chlorome	C8H23ClO4	330.03619	7	2.3
11	DTXCID90422015	DTXCID	DTXSID4047	3,5,6,8-Tetrahydroxy-1-meth	DTXCID90	3,5,6,8-Tetra	C16H10O8	330.03757	7	1.
12	DTXCID10299049	DTXCID	DTXSID5034	Cyclodisilazane, 1,3-bis(chlc	DTXCID10	1,3-Bis[chlo	C8H24Cl2N2	330.03936	5	2.1
13	DTXCID30292072	DTXCID	DTXSID8034	1-(4-chlorophenyl)-3-(5-phen	DTXCID30	N-(4-Chloro	C15H11ClN4	330.03421	5	1
14	DTXCID80128599	DTXCID	DTXSID2020	3,4'-Di-O-methylellagic acid	DTXCID80	2,8-Dihydro>	C16H10O8	330.03757	4	2
15	DTXCID401337543	DTXCID	DTXSID7090	7-Chloro-2-methyl-5-phenyl-5	DTXCID40	7-Chloro-2-n	C15H11ClN4	330.03421	3	1.3
16	DTXCID70266898	DTXCID	DTXSID3031	Bis(2-ethylbutyl)diselane	DTXCID70	Bis(2-ethylb	C12H26Se2	330.0365	3	3.3
17	DTXCID501367644	DTXCID	DTXSID3093	{5-[(3,5-Dichlorophenyl)sulfar	DTXCID50	{5-[(3,5-Dich	C14H16Cl2N	330.03604	3	1.7
18	DTXCID301367648	DTXCID	DTXSID2093	{4-[(3,5-Dichlorophenyl)sulfar	DTXCID30	{4-[(3,5-Dich	C14H16Cl2N	330.03604	3	1.7
19	DTXCID30276115	DTXCID	DTXSID8032	[1,1'-Biphenyl]-2,2',6,6'-tetrac	DTXCID30	[1,1'-Biphen	C16H10O8	330.03757	3	8.0
20	DTXCID30529557	DTXCID	DTXSID6057	[1,1'-Biphenyl]-3,3',5,5'-tetrac	DTXCID30	[1,1'-Biphen	C16H10O8	330.03757	3	8.0

Dashboard search results of all hits from CFM-ID search

3rd Mock Scenario

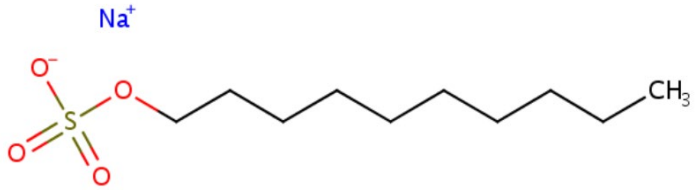
- Final results and prediction
- Based on very low scores of all MS_2 results, decided to rely on MS_1 results for prediction, which all agreed with $C_{10}H_{19}O_6PS_2$ as compound of interest
- Malathion was top hit based on Data Sources, by a landslide
  - Correct prediction!
- Start to finish, total of 14 working hours



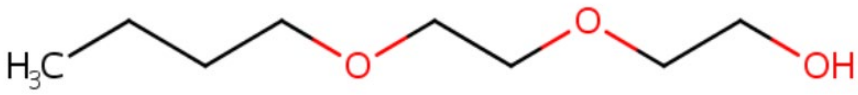
Malathion ($C_{10}H_{19}O_6PS_2$)

4th Mock Scenario: AFFF in River Water

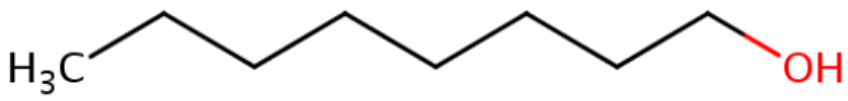
- AFFF are water-based, typically contain hydrocarbon-based and fluoro-surfactants
- Examples of common chemicals:



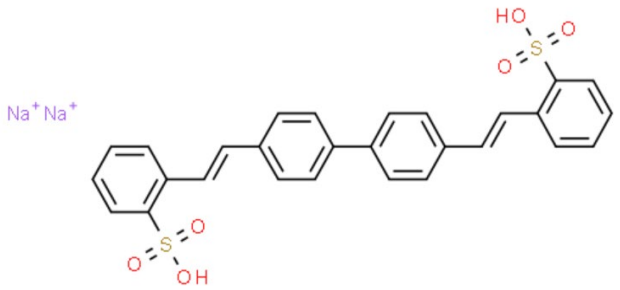
Sodium Decyl Sulfate



2-(2-Butoxyethoxy)ethanol

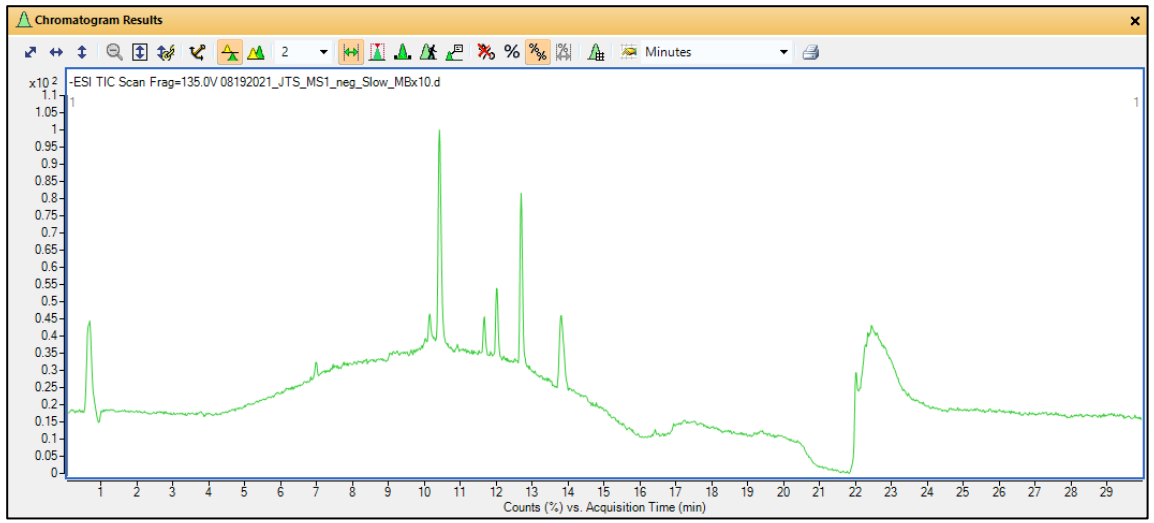


1-Octanol

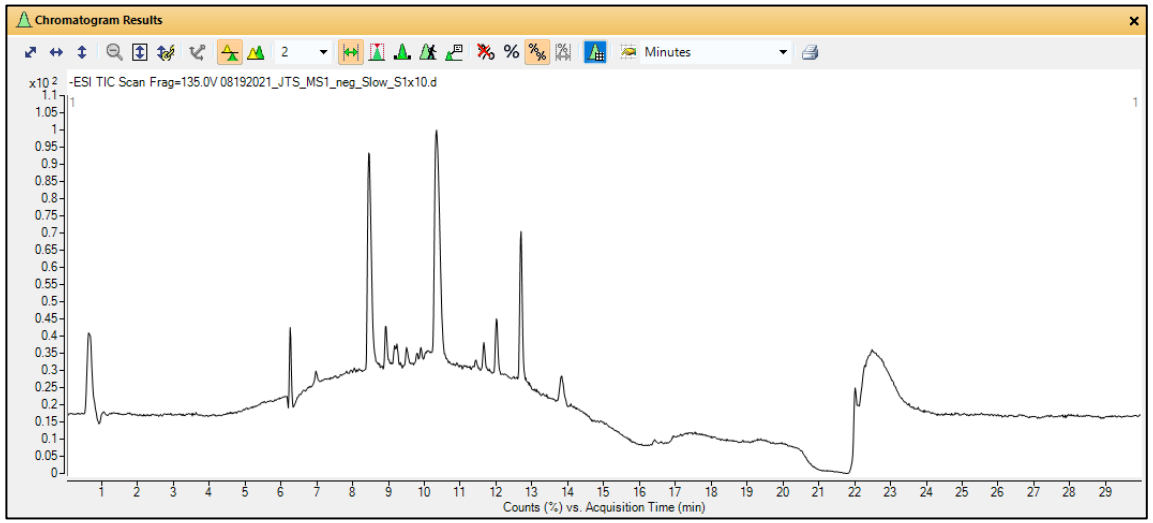


4,4’-bis-(sulfostyryl)-biphenyl disodium salt

4th Mock Scenario



Matrix blank chromatogram



Sample chromatogram

Compound	Blankx10	S1x10	Annotations	Compound Name	Formula	Frequency	Ionization	Mass	MS1	RT	Sample:Blank	Ratio >10?
C10H15N2OP		191000000	C10H15N2OP [	C10H15N2OP	C10 H15 N2 O P	2	Esi-	210.0924	(419.1786, 3137.86	8.447	191000000	YES
C10H15N2OP 8.447		191000000	C10H15N2OP [	C10H15N2OP 8.447	C10 H15 N2 O P	2	Esi-	210.0928	(209.0855, 3807232	8.447	191000000	YES
C13H17FNO2	261	139000000	C13H17FNO2 [	C13H17FNO2	C13 H17 F N O2	2	Esi-	238.1263	(475.2412, 3566.87	10.234	5320.166877	YES
C15H13Cl2NO4S	689000	58800000	C15H13Cl2NO4	C15H13Cl2NO4S	C15 H13 Cl2 N O4 S	2	Esi-	372.9944	(371.9872, 3168988	10.393	0.85341074	NO
C16H15Cl4O5P	507000	48000000	C16H15Cl4O5P	C16H15Cl4O5P	C16 H15 Cl4 O5 P	2	Esi-	457.9414	(914.8747, 17757.1	12.693	0.946745562	NO
C11H11IN6O4	488000	43200000	C11H11IN6O4 [	C11H11IN6O4	C11 H11 I N6 O4	2	Esi-	417.9874	(834.967, 61789.5)(	10.393	0.885245902	NO
503.9514@13.822	880000	38600000		503.9514@13.822		2	Esi-	503.9514	(502.9441, 2734899	13.822	0.438636364	NO
C20H14Cl4O6	1990000	17300000	C20H14Cl4O6 [	C20H14Cl4O6	C20 H14 Cl4 O6	2	Esi-	489.9541	(488.9468, 1586030	12.007	0.869346734	NO
C19H16OS3	1130000	15000000	C19H16OS3 [C	C19H16OS3	C19 H16 O S3	2	Esi-	356.0364	(355.0295, 314841.5	10.201	1.327433628	NO
C3HF5O3	1300000	13600000	C3HF5O3 [C3 H	C3HF5O3	C3 H F5 O3	2	Esi-	179.9852	(178.9779, 1831067	0.617	1.046153846	NO
C6H2F4S	565339	9605196	C6H2F4S [C6 H	C6H2F4S	C6 H2 F4 S	2	Esi-	181.9809	(180.9734, 849558.(	22.252	1.699631893	NO
C11H16O2S	1	9197612	C11H16O2S [C	C11H16O2S	C11 H16 O2 S	2	Esi-	212.0884	(211.0811, 780841.3	8.464	9197612	YES
C10H19O6PS2	814552	8937208	C10H19O6PS2 [	C10H19O6PS2	C10 H19 O6 P S2	2	Esi-	330.0367	(329.0294, 179909.3	10.218	1.097202218	NO
515.8854@10.385	357884	7635923		515.8854@10.385		2	Esi-	515.8854	(514.8783, 276921.8	10.385	2.134081536	NO
555.8394@12.693	542945	7071045		555.8394@12.693		2	Esi-	555.8394	(554.832, 257630.4	12.693	1.30242331	NO
591.816@12.693	576296	6595675		591.816@12.693		2	Esi-	591.816	(590.8087, 139818.(	12.693	1.144526893	NO
C16H11ClN2O9S2	796326	6499723	C16H11ClN2O9	C16H11ClN2O9S2	C16 H11 Cl N2 O9 S2	2	Esi-	473.9593	(472.952, 636688.1	11.656	0.816212197	NO
C9H20O7		6290139	C9H20O7 [C9 H	C9H20O7	C9 H20 O7	2	Esi-	240.1203	(239.1125, 389785.8	10.32	6290139	YES
C5H8N2OS3	483879	6264446	C5H8N2OS3 [C	C5H8N2OS3	C5 H8 N2 O S3	2	Esi-	207.98	(206.9727, 748900.4	0.617	1.294629345	NO
C6H10N2O3	6740721	6148373	C6H10N2O3 [C	C6H10N2O3	C6 H10 N2 O3	2	Esi-	158.0675	(157.0603, 890131.9	6.973	0.91212394	NO

MPP output file, organized by feature abundance in sample

4th Mock Scenario

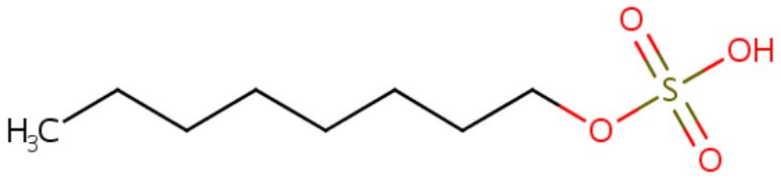
- Picked list of “top” features for further investigation
  - 11 features from ESI-
  - 6 features from ESI+
- Chosen based on being the top 5 most abundant features in the sample that had a Sample:Blank ratio >10
- Next set of features chosen based on the same requirement of Sample:Blank ratio >10, but these features also had an obviously negative mass defect

MS_1:
- Formula matching to MS Ready Formula (Dashboard)
- WebApp Search by Mass
- Molecular formula prediction

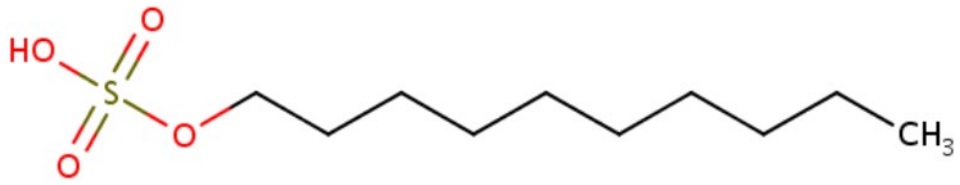
MS_2:
- Matching to spectral libraries
- WebApp match to CFM-ID in-silico database

4th Mock Scenario

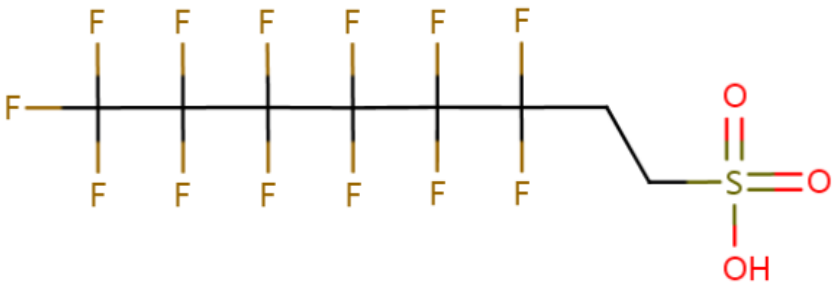
- Confident structure identification (*Level 2 Schymanski*)



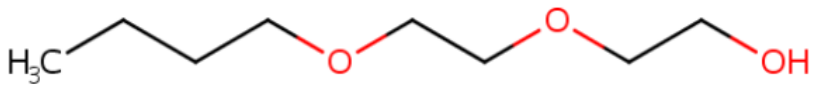
ESI- Feature 1 – **Octyl Hydrogen Sulfate**



ESI- Feature 2 – **Decyl Hydrogen Sulfate**



ESI- Feature 9 – **6:2 Fluorotelomer Sulfonic Acid**

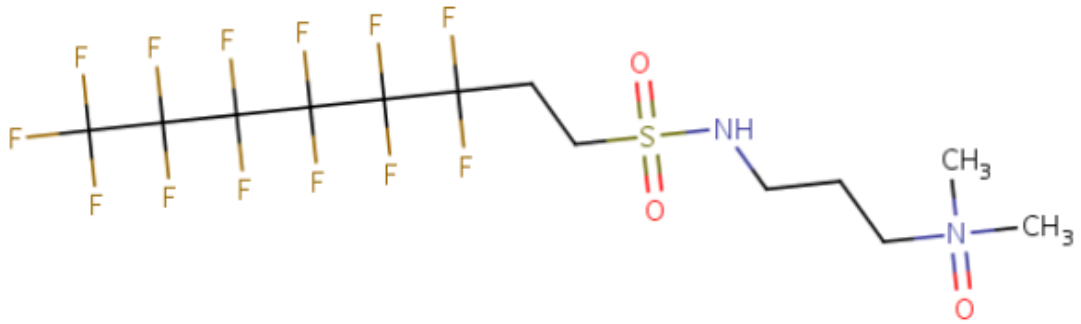


ESI+ Feature 3 – **2-(2-Butoxyethoxy)ethanol**

Schymanski et al., Environ. Sci. Technol. 2014, 48 (4) 2097-2098.

4th Mock Scenario

- Probable structure identification
  - (*Level 3 Schymanski*)

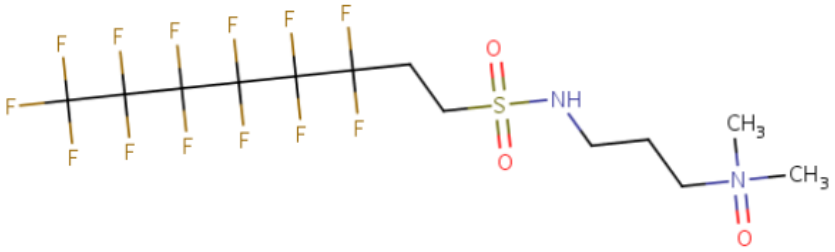


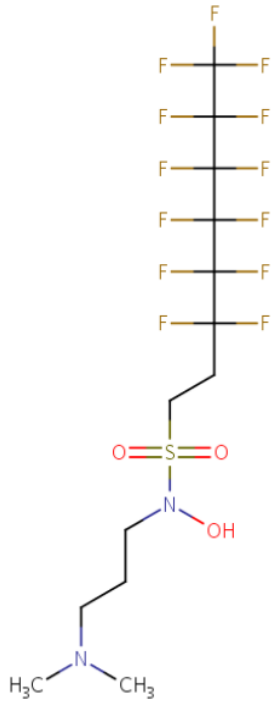
ESI- Feature 5 – either **N,N-Dimethyl-3-((perfluorohexyl)ethylsulfonyl)aminopropanamine N-oxide**

$C_{13}H_{17}F_{13}N_2O_3S$

4th Mock Scenario

- Probable structure identification
  - (*Level 3 Schymanski*)





ESI- Feature 5 – either **N,N-Dimethyl-3-((perfluorohexyl)ethylsulfonyl)aminopropanamine N-oxide** *or* **N-[3-(Dimethylamino)propyl]-3,3,4,4,5,5,6,6,7,7,8,8,8-tridecafluoro-N-hydroxyoctane-1-sulfonamide**

$C_{13}H_{17}F_{13}N_2O_3S$

4th Mock Scenario

- Formula assignment (*Level 4 Schymanski*):
  - ESI+ Feature 11 and 13: Best predictions are formula $C_5H_5Cl_2N_3S$ and $C_4H_3ClN_2O_3$
- Masses of interest (*Level 5 Schymanski*):
  - ESI- Features 11, 12, 17, and 18: *m/z* 306.9832, 256.9545, 134.9874, and 334.9557
  - ESI+ Feature 12: *m/z* 100.9915
- Incorrect assignments:
  - ESI- Features 3 and 4: isotopologues of other features
  - ESI- Feature 15: false positive
- 50% identified at Tier 1 confidence, and 50% identified at Tier 2

Conclusions/Next Steps

- Things we've learned:
  - Range finding is really important!
  - We utilize MS_1 data much more than expected
  - The WebApp is a useful tool, and is being customized further for these applications
- Things we can improve upon
- Next steps for future mock scenarios

Acknowledgements

EPA ORD
Alex Chao
Jennifer Gundersen
Seth R. Newton
Allison L. Phillips
Jon R. Sobus
Elin M. Ulrich
Antony J. Williams
Agilent
Jarod N. Grossman

Sloop.John@epa.gov