Curating and Sharing Structures and Spectra for the Environmental Community

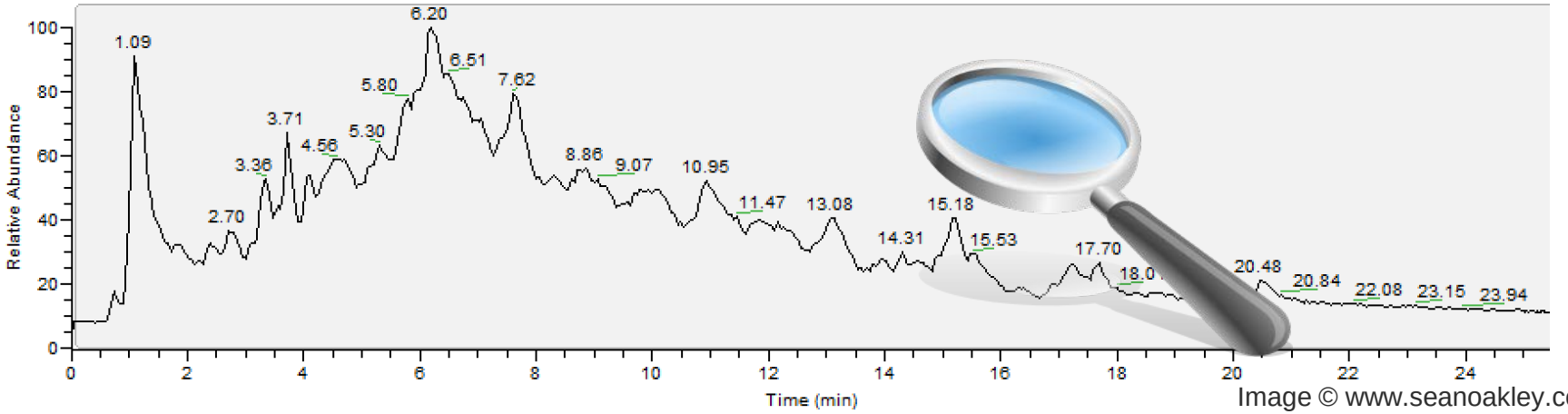


Image © www.seanoakley.com/

Emma Schymanski
Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg.
Email: emma.schymanski@uni.lu

Antony J. Williams (NCCT, US EPA, Research Triangle Park, NC, USA)

The Goal:

- To identify as many substances in environmental samples with high resolution mass spectrometry as possible

www.norman-data.eu

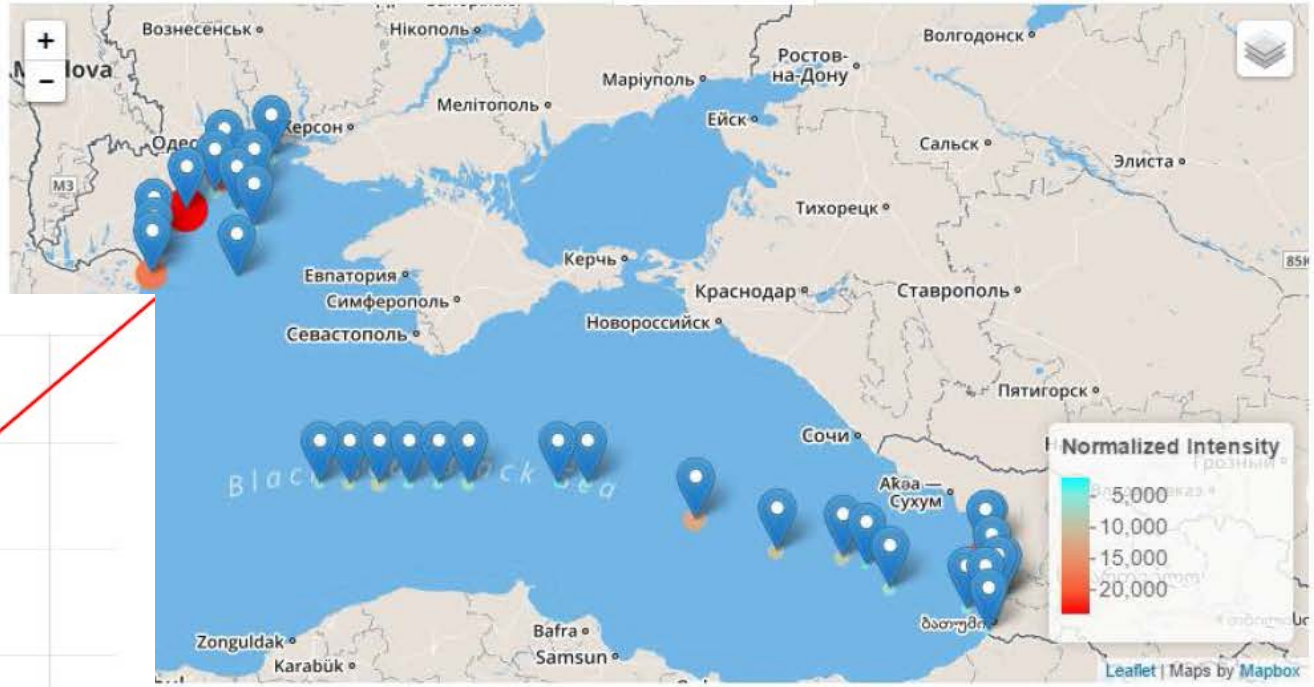
NORMAN Digital Sample Freezing Platform | Main Page | Batch mode

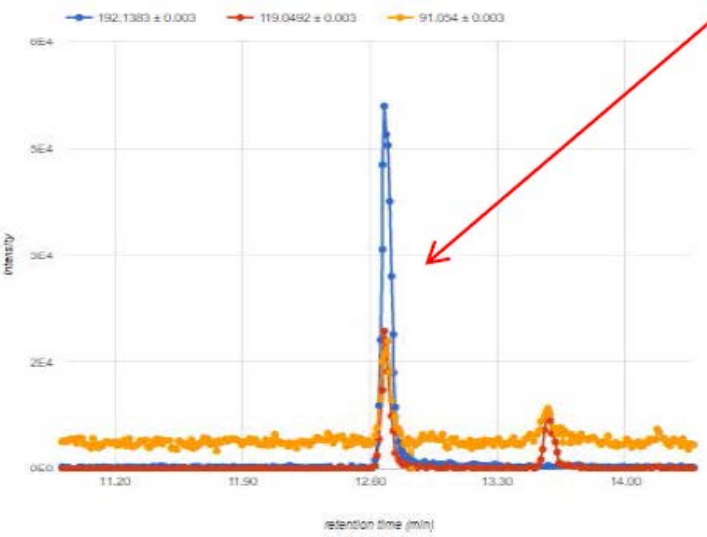
Choose Emerging Substance or input mass of interest and experimental RTI

Substance name or CAS or StdInChIKey

DEET [134-62-3] [MMOXZBCLCQITDF-UHFFFAOYSA-N]

Contributed Samples | Results | Chromatograms | Interactive Map | Help





NORMAN Digital Sample Freezing Platform
- Retrospective screening across Europe

Aligizakis *et al* in prep.

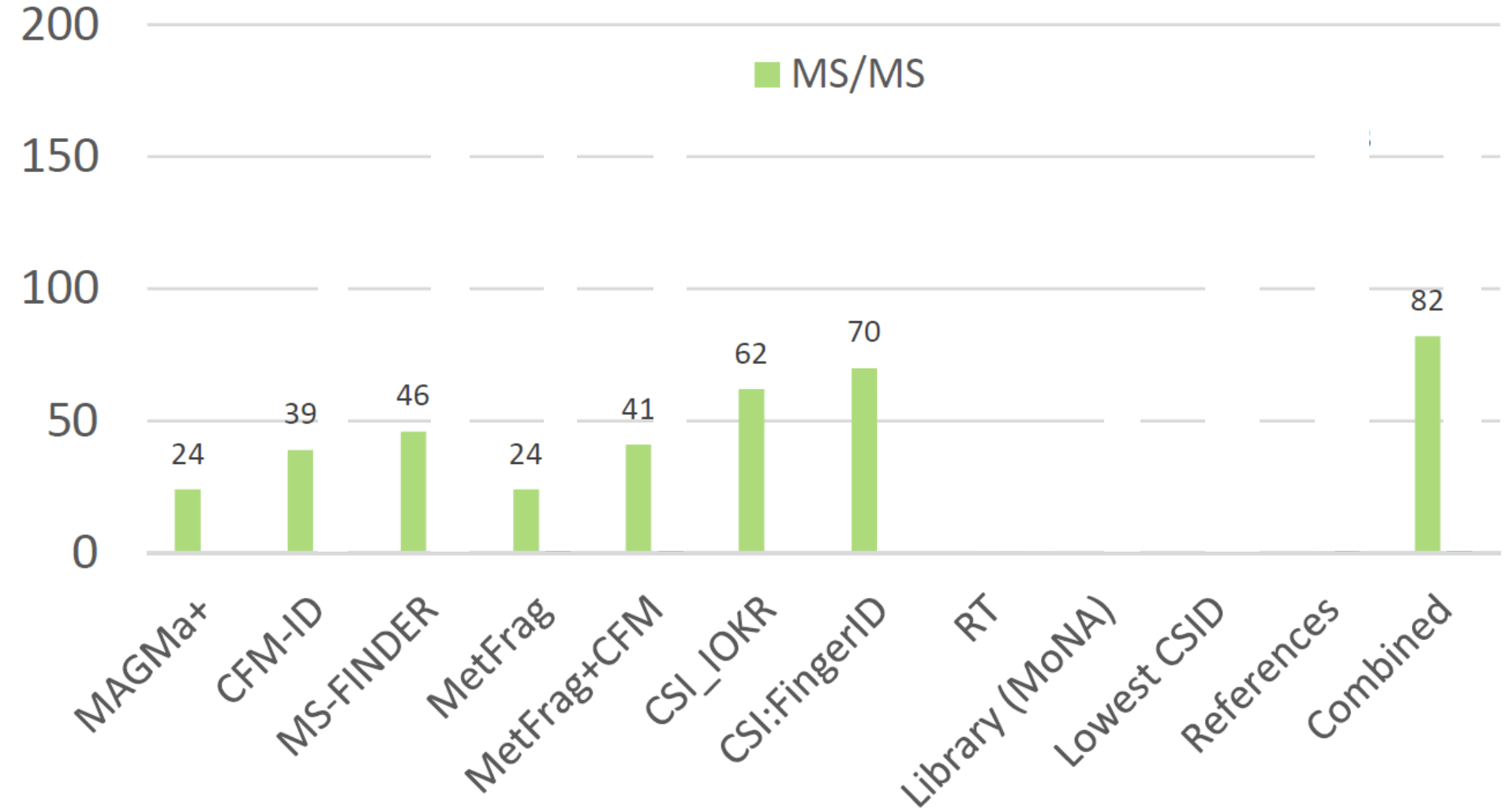
The Goal:

- To identify as many substances in environmental samples with high resolution mass spectrometry as possible

- To do this we need:
  - Mass Spectra in reference libraries
  - Metadata

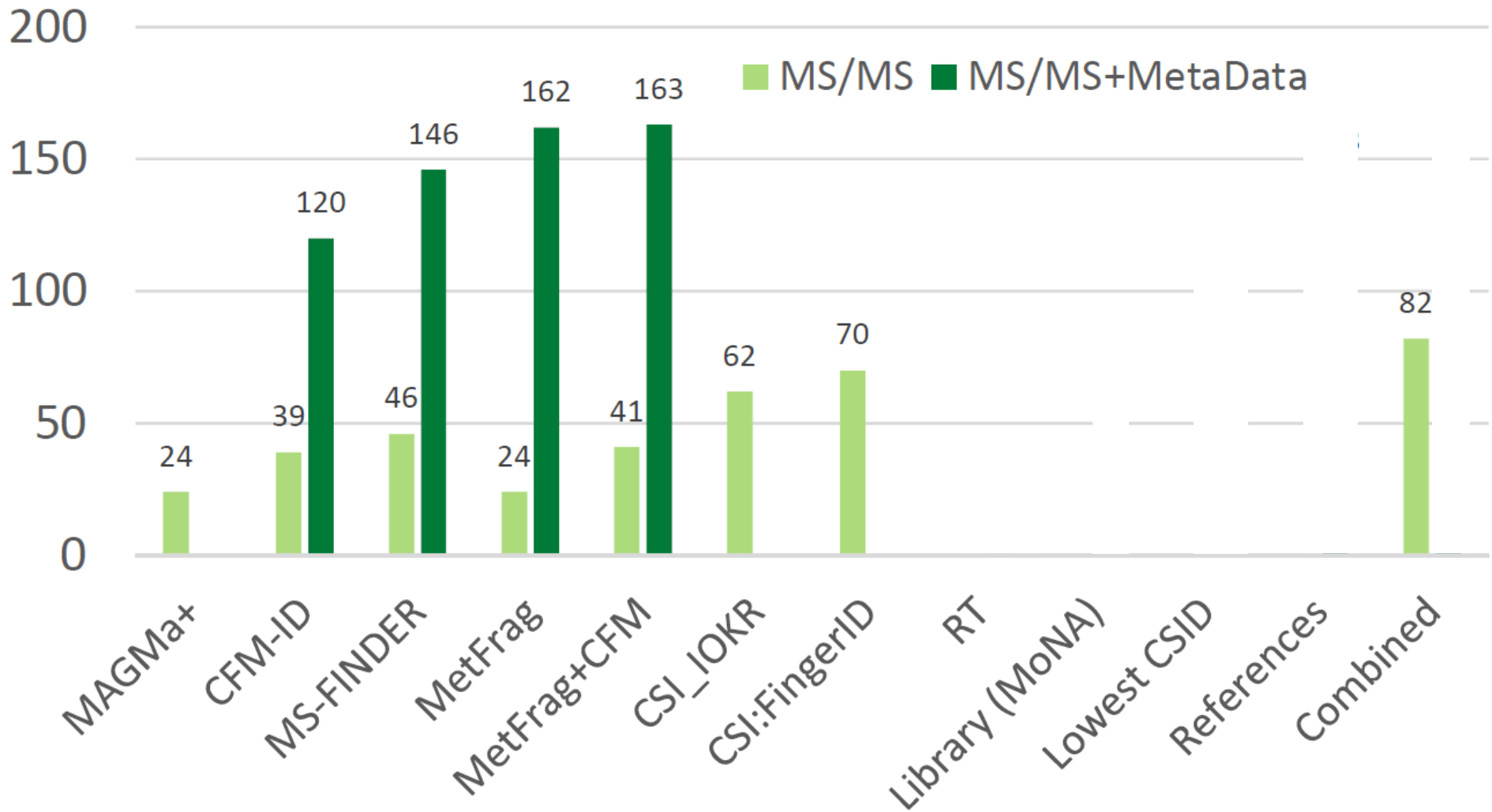
(and fancy computational methods … but that's another story…)

The Power of the Metadata (Top 1 ranks)



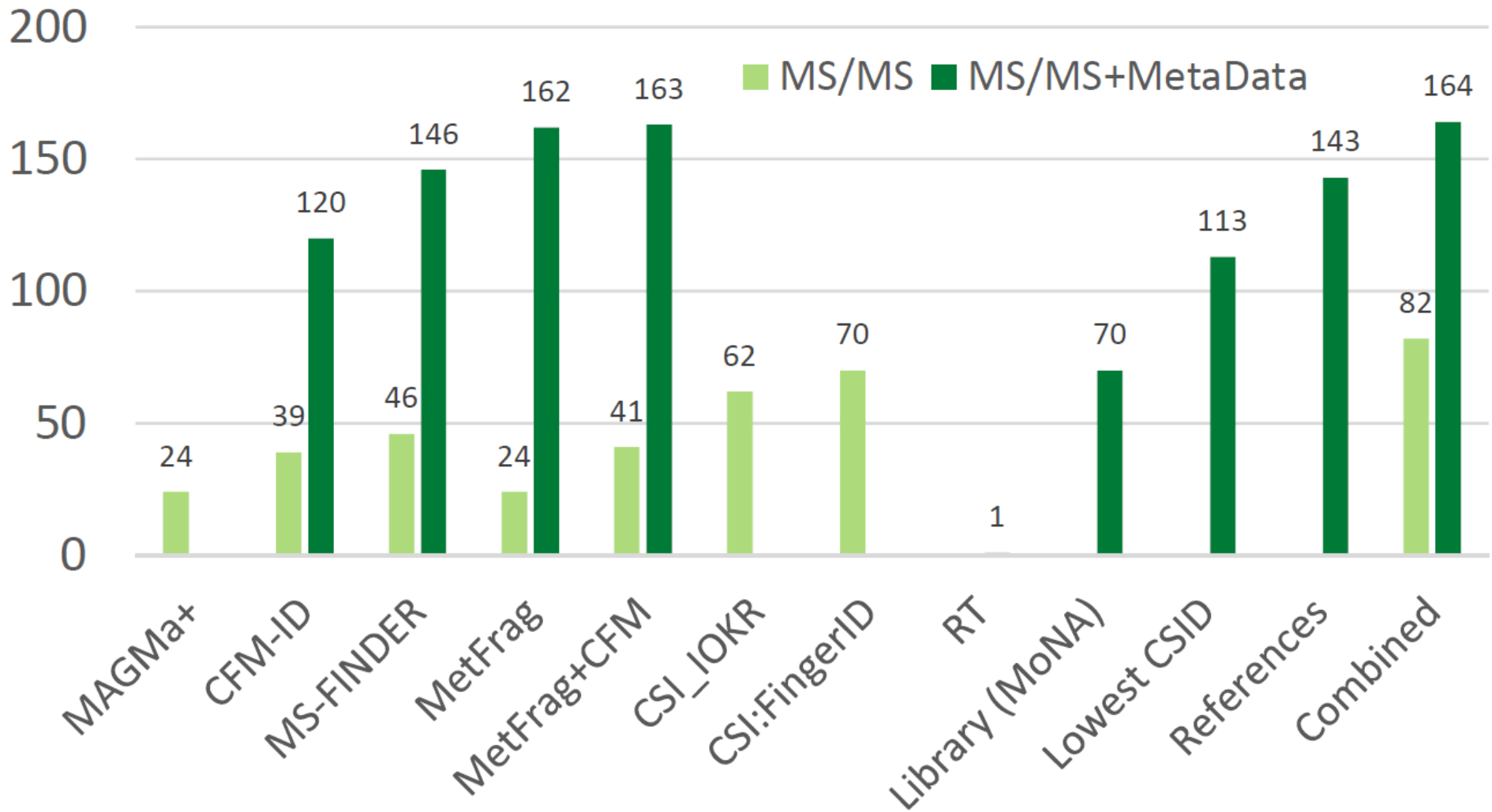
Schymanski *et al*, 2017, J Cheminf., DOI: 10.1186/s13321-017-0207-1

The Power of the Metadata (Top 1 ranks)



Schymanski *et al*, 2017, J Cheminf., DOI: 10.1186/s13321-017-0207-1

The Power of the Metadata (Top 1 ranks)



Schymanski *et al*, 2017, J Cheminf., DOI: 10.1186/s13321-017-0207-1

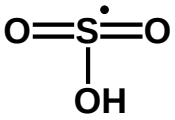
MetFrag2.3: Non-target Identification


eawag aquatic research solutions

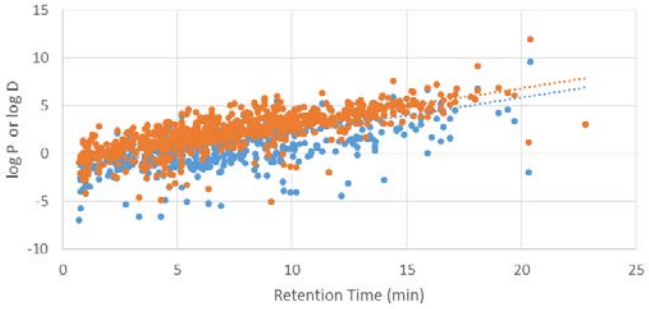
Status: 2010 => 2016

mz [M-H]⁻
213.9637
± 5 ppm

Elements: C,N,S
5 ppm
0.001 Da



RT: 4.54 min
355 InChI/RTs



ChemSpider
Search and share chemistry

or

PubChem OPEN CHEMISTRY DATABASE

References
External Refs
Data Sources
RSC Count
PubMed Count

Literature Patents

MetFrag2.3

Suspect Lists

TOFF IDENT
Chemistry Dashboard

MoNA
MassBank of North America
MassBank.eu

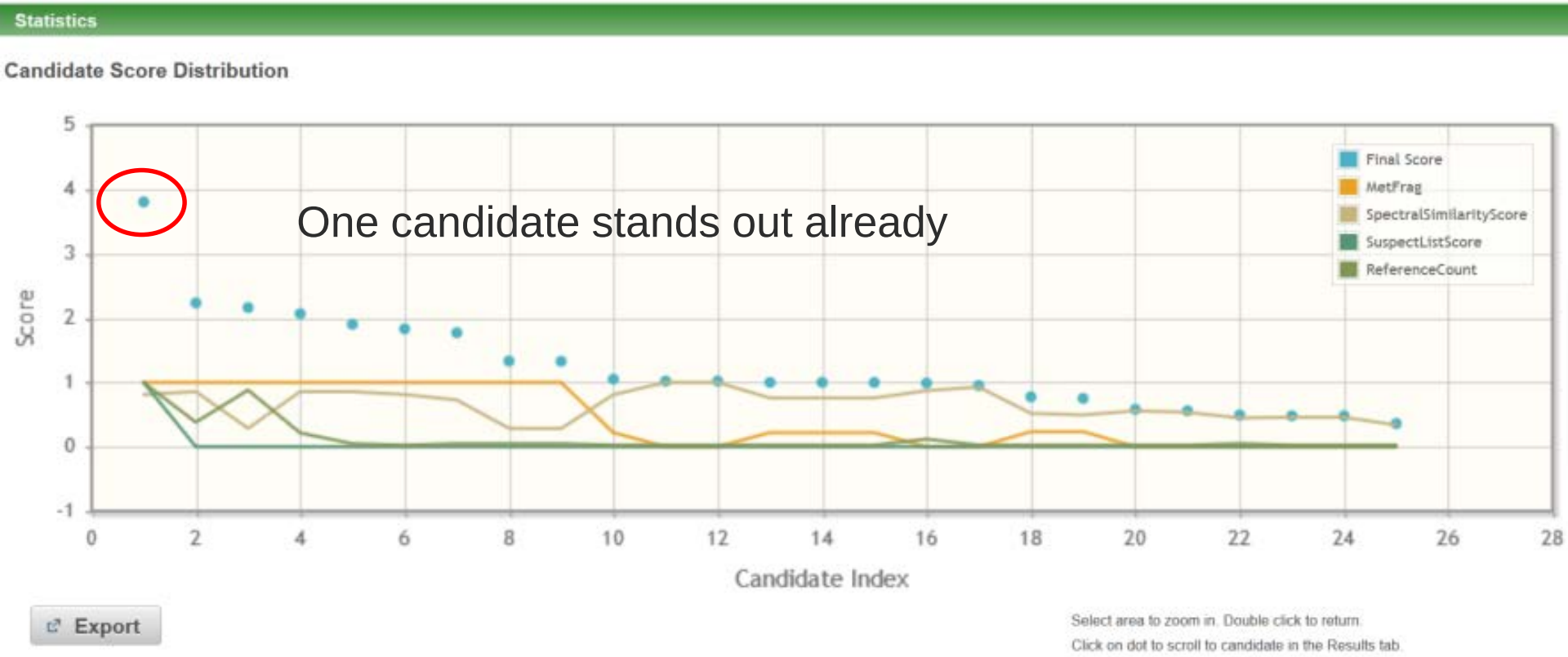
MS/MS

134.0054	339689
150.0001	77271
213.9607	632466

Ruttkies, Schymanski, Wolf, Hollender, Neumann (2016) *J. Cheminf.*, 2016, DOI: 10.1186/s13321-016-0115-9

Non-target Identification and Metadata

- Helps prioritize interesting candidates rapidly …
- …assuming candidates are in databases …
- https://msbi.ipb-halle.de/MetFragBeta/



Ruttkies, Schymanski, Wolf, Hollender, Neumann (2016) *J. Cheminf.*, 2016, DOI: 10.1186/s13321-016-0115-9

Example from Schymanski *et al* 2014, ES&T, 48: 1811-1818. DOI: 10.1021/es4044374

Suspect Screening Allows Efficient Data Exploration

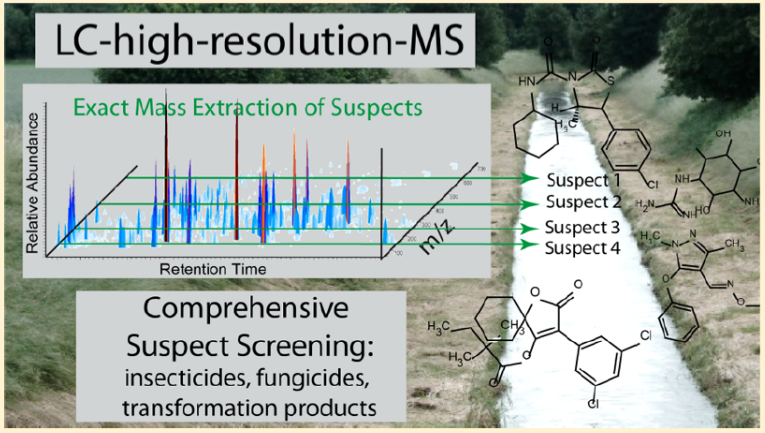
Alleviating the Reference Standard Dilemma Using a Systematic Exact Mass Suspect Screening Approach with Liquid Chromatography-High Resolution Mass Spectrometry

Christoph Moschet,[§,∥] Alessandro Piazzoli,[§,∥] Heinz Singer,[*,§] and Juliane Hollender[§,∥]

[§]Eawag, Swiss Federal Institute of Aquatic Science and Technology, Überlandstrasse 133, 8600 Dübendorf, Switzerland
[∥]Institute of Biogeochemistry and Pollutant Dynamics, ETH Zürich, 8092 Zürich, Switzerland

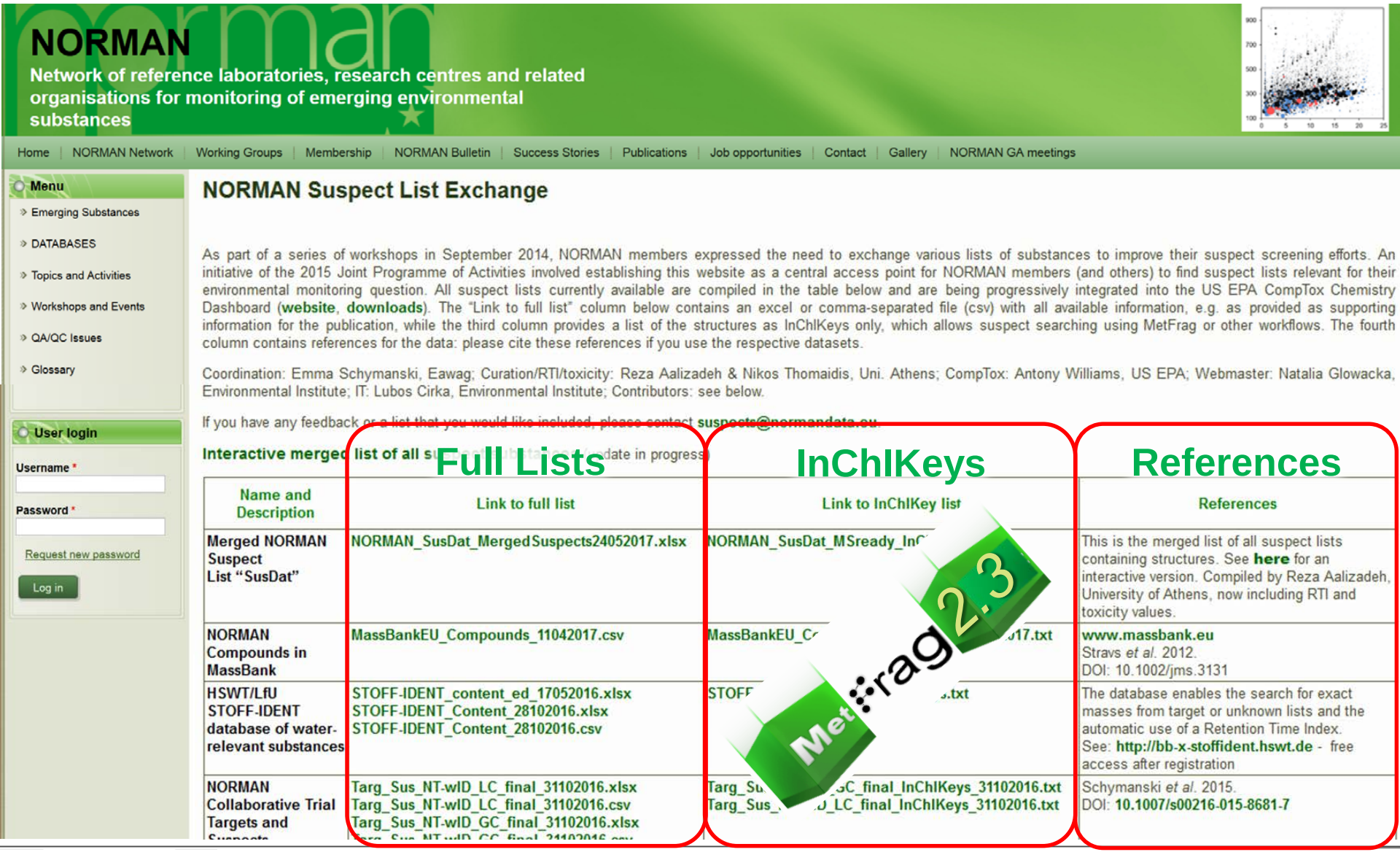
Supporting Information

ABSTRACT: In this study, the efficiency of a suspect screening strategy using liquid chromatography-high resolution mass spectrometry (LC-HRMS) without the prior purchase of reference standards was systematically optimized and evaluated for assessing the exposure of rarely investigated pesticides and their transformation products (TPs) in 76 surface water samples. Water-soluble and readily ionizable (electrospray ionization) substances, 185 in total, were selected from a list of all insecticides and fungicides registered in Switzerland and their major TPs. Initially, a solid phase extraction-LC-HRMS method was established using 45 known, persistent, and high sales volume pesticides. Seventy percent of these target substances had limit of



NORMAN Suspect Exchange & SusDat

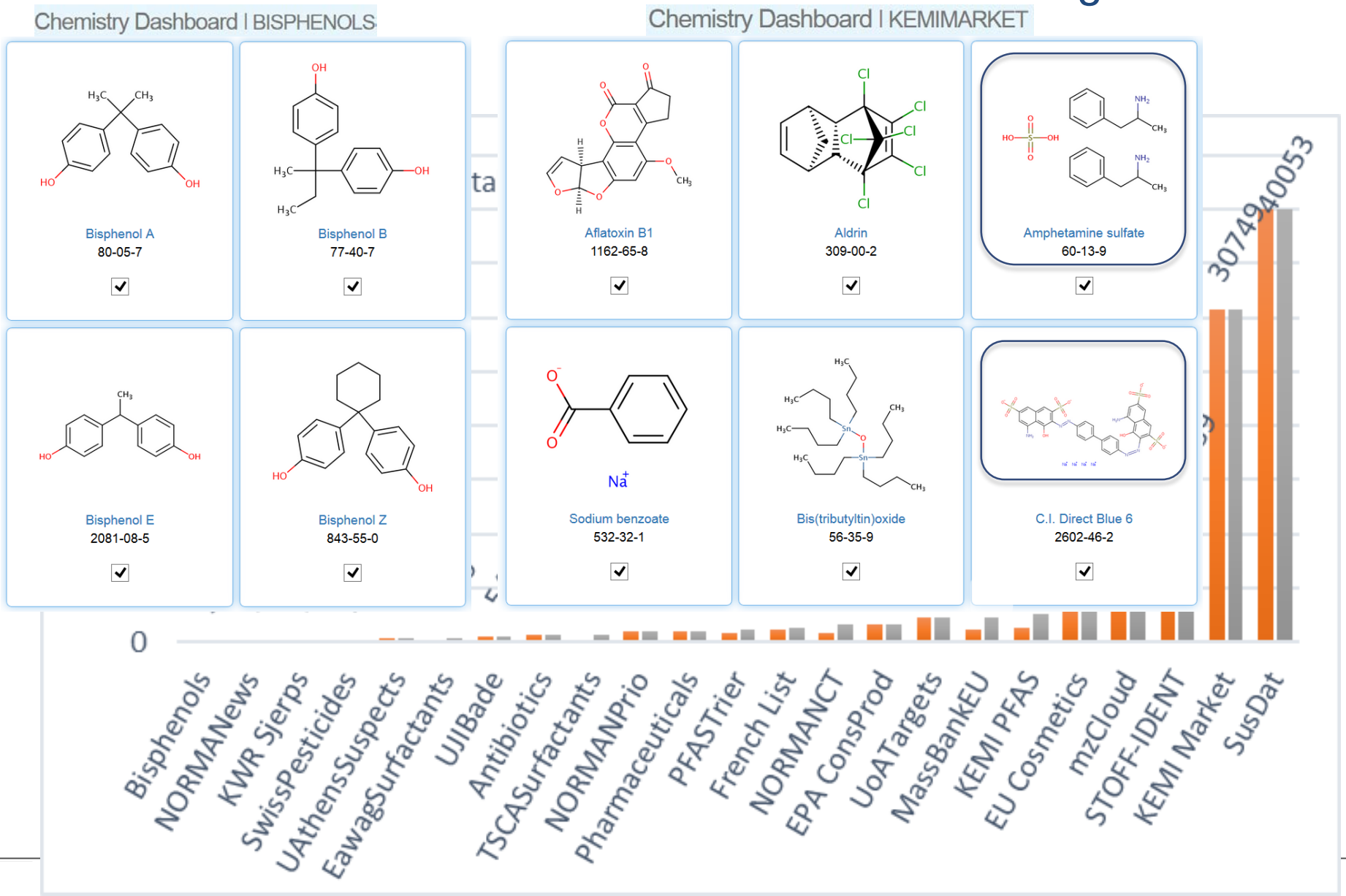
- http://www.norman-network.com/?q=node/236



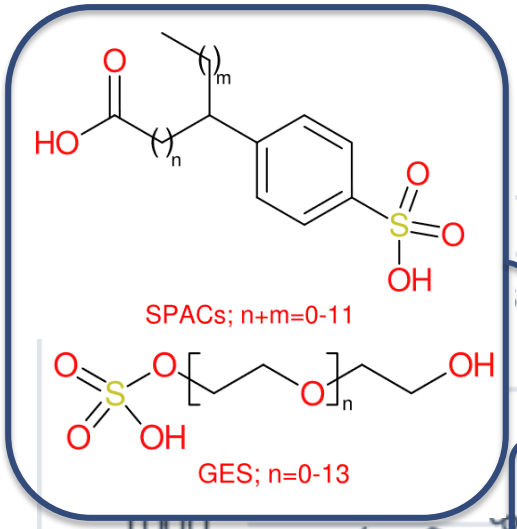
Schymanski, Aalizadeh *et al*. in prep.

Specialised Lists through to Market Lists

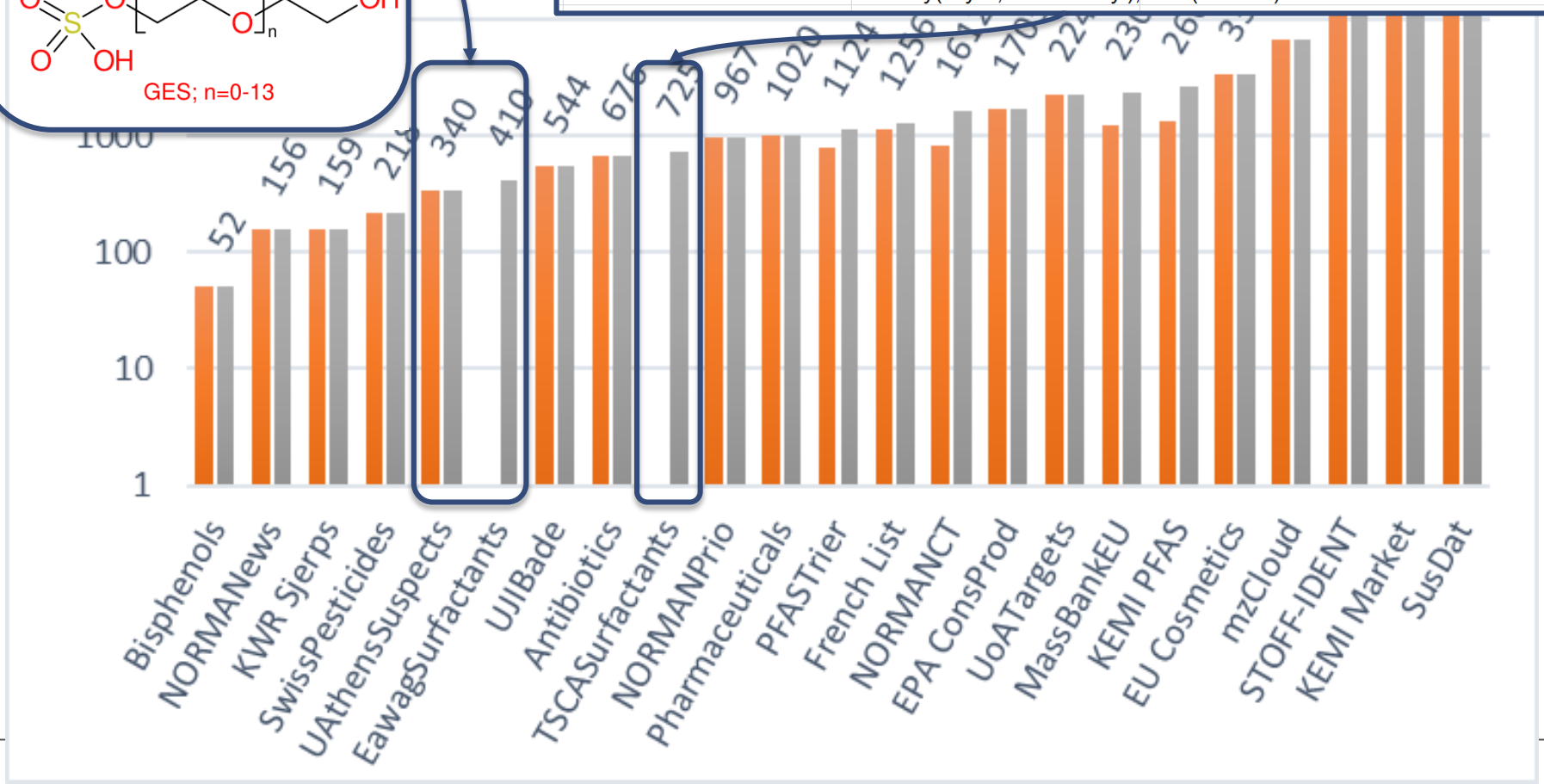
- Now 23 lists available online … from small to large!



…but not all are what they seem…



CAS No.	Name	MF
RN: 9002-92-0	IN: Poly(oxy-1,2-ethanediyl),	MF: (C2H4O)nC12H26O
RN: 60828-78-6	IN: Poly(oxy-1,2-ethanediyl),	MF: (C2H4O)nC12H26O
RN: 61702-78-1	IN: Poly(oxy-1,2-ethanediyl),	MF: (C2H4O)nC12H26O
RN: 65733-67-7	IN: Poly(oxy-1,2-ethanediyl),	MF: (C2H4O)nC12H26O.xH4O7P2
RN: 109075-72-1	IN: Poly(oxy-1,2-ethanediyl),	MF: (C2H4O)nC12H26O
RN: 77680-31-0	IN: Poly(oxy-1,2-ethanediyl),	MF: (C2H4O)n(C2H4O)nC19H37NO3.K
RN: 41928-09-0	IN: Poly(oxy-1,2-ethanediyl),	MF: (C2H4O)n(C2H4O)nC29H44O2
RN: 61723-87-3	IN: Poly(oxy-1,2-ethanediyl),	MF: (C2H4O)nC19H32O
RN: 51192-09-7	IN: Poly(oxy-1,2-ethanediyl),	MF: (C2H4O)n(C2H4O)nC21H40O4
RN: 68311-23-9	IN: Poly(oxy-1,2-ethanediyl),	MF: (C2H4O)nC17H34O3.Na



Example: Eawag Surfactant List

https://comptox.epa.gov/dashboard/chemical_lists/eawagsurf

Eawag Surfactant Suspect List (formulas only)	**Surfactant_Suspects_Schymanski_etal_2014.xlsx** **Surfactant_Suspects_Schymanski_etal_2014.csv**		Schymanski *et al.* 2014. DOI: **10.1021/es4044374**

SuspectID	Name	Name_ref	Formula	Monoisoto	Adduct_Sta	M+H+	M-H-	Reference	Reference_DOI	Source_ref	Source_DOI
C10-LAS	C10-LAS	C10-LAS_G	C16H26S1C	298.1603	None	299.1675	297.153	Schymansk	dx.doi.org/10.1021/es	Gonzalez_e	dx.doi.org/10.1002/rcm.3527
C11-LAS	C11-LAS	C11-LAS_G	C17H28S1C	312.1759	None	313.1832	311.1686	Schymansk	dx.doi.org/10.1021/es	Gonzalez_e	dx.doi.org/10.1002/rcm.3527
C12-LAS	C12-LAS	C12-LAS_G	C18H30S1C	326.1916	None	327.1988	325.1843	Schymansk	dx.doi.org/10.1021/es	Gonzalez_e	dx.doi.org/10.1002/rcm.3527
C13-LAS	C13-LAS	C13-LAS_G	C19H32S1C	340.2072	None	341.2145	339.1999	Schymansk	dx.doi.org/10.1021/es	Gonzalez_e	dx.doi.org/10.1002/rcm.3527
C14-LAS	C14-LAS	C14-LAS_G	C20H34S1C	354.2229	None	355.2301	353.2156	Schymansk	dx.doi.org/10.1021/es	Gonzalez_e	dx.doi.org/10.1002/rcm.3527
C3-SPC	C3-SPC	C3-SPC_Co	C9H10O5S1	230.0249	None	231.0322	229.0176	Schymansk	dx.doi.org/10.1021/es	Corada-Fer	dx.doi.org/10.1039/c1em10150a
C4-SPC	C4-SPC	C4-SPC_Co	C10H12O5S	244.0405	None	245.0478	243.0333	Schymansk	dx.doi.org/10.1021/es	Corada-Fer	dx.doi.org/10.1039/c1em10150a
C5-SPC	C5-SPC	C5-SPC_Co	C11H14O5S	258.0562	None	259.0635	257.0489	Schymansk	dx.doi.org/10.1021/es	Corada-Fer	dx.doi.org/10.1039/c1em10150a
C6-SPC	C6-SPC	C6-SPC_Co	C12H16O5S	272.0718	None	273.0791	271.0646	Schymansk	dx.doi.org/10.1021/es	Corada-Fer	dx.doi.org/10.1039/c1em10150a
C7-SPC	C7-SPC	C7-SPC_Co	C13H18O5S	286.0875	None	287.0948	285.0802	Schymansk	dx.doi.org/10.1021/es	Corada-Fer	dx.doi.org/10.1039/c1em10150a
C8-SPC	C8-SPC	C8-SPC_Co	C14H20O5S	300.1031	None	301.1104	299.0959	Schymansk	dx.doi.org/10.1021/es	Corada-Fer	dx.doi.org/10.1039/c1em10150a
C9-SPC	C9-SPC	C9-SPC_Co	C15H22O5S	314.1188	None	315.1261	313.1115	Schymansk	dx.doi.org/10.1021/es	Corada-Fer	dx.doi.org/10.1039/c1em10150a
C10-SPC	C10-SPC	C10-SPC_C	C16H24O5S	328.1344	None	329.1417	327.1272	Schymansk	dx.doi.org/10.1021/es	Corada-Fer	dx.doi.org/10.1039/c1em10150a
C11-SPC	C11-SPC	C11-SPC_C	C17H26O5S	342.1501	None	343.1574	341.1428	Schymansk	dx.doi.org/10.1021/es	Corada-Fer	dx.doi.org/10.1039/c1em10150a
C12-SPC	C12-SPC	C12-SPC_C	C18H28O5S	356.1657	None	357.173	355.1585	Schymansk	dx.doi.org/10.1021/es	Corada-Fer	dx.doi.org/10.1039/c1em10150a
C13-SPC	C13-SPC	C13-SPC_C	C19H30O5S	370.1814	None	371.1887	369.1741	Schymansk	dx.doi.org/10.1021/es	Corada-Fer	dx.doi.org/10.1039/c1em10150a
C14-SPC	C14-SPC	C14-SPC_C	C20H32O5S	384.197	None	385.2043	383.1898	Schymansk	dx.doi.org/10.1021/es	Corada-Fer	dx.doi.org/10.1039/c1em10150a
C15-SPC	C15-SPC	C15-SPC_C	C21H34O5S	398.2127	None	399.22	397.2054	Schymansk	dx.doi.org/10.1021/es	Corada-Fer	dx.doi.org/10.1039/c1em10150a
SPA-1DC	SPA-1DC	SPA-1DC_D	C9H8O7S1	259.9991	None	261.0063	258.9918	Schymansk	dx.doi.org/10.1021/es	DiCorcia_et	dx.doi.org/10.1021/es990596u
SPA-2DC	SPA-2DC	SPA-2DC_D	C10H10O7S	274.0147	None	275.022	273.0074	Schymansk	dx.doi.org/10.1021/es	DiCorcia_et	dx.doi.org/10.1021/es990596u
SPA-3DC	SPA-3DC	SPA-3DC_D	C11H12O7S	288.0304	None	289.0376	287.0231	Schymansk	dx.doi.org/10.1021/es	DiCorcia_et	dx.doi.org/10.1021/es990596u
SPA-4DC	SPA-4DC	SPA-4DC_D	C12H14O7S	302.046	None	303.0533	301.0387	Schymansk	dx.doi.org/10.1021/es	DiCorcia_et	dx.doi.org/10.1021/es990596u

Schymanski *et al* 2014, ES&T, 48: 1811-1818. DOI: 10.1021/es4044374

Eawag Surfactant List (after many late nights…)

https://comptox.epa.gov/dashboard/chemical_lists/eawagsurf

Eawag Surfactant Suspect List (formulas only)	Surfactant_Suspects_Schymanski_etal_2014.xlsx Surfactant_Suspects_Schymanski_etal_2014.csv		Schymanski *et al.* 2014. DOI: **10.1021/es4044374**

SuspectID	Name	Name_ref	Formula	Monoisoto	Adduct_Sta	M+H+	M-H-	Reference	Reference_DOI	Source_ref	Source_DOI
C10-LAS	C10-LAS	C10-LAS_G	C16H26S1C	298.1603	None	299.1675	297.153	Schymansk	dx.doi.org/10.1021/es	Gonzalez_e	dx.doi.org/10.1002/rcm.3527

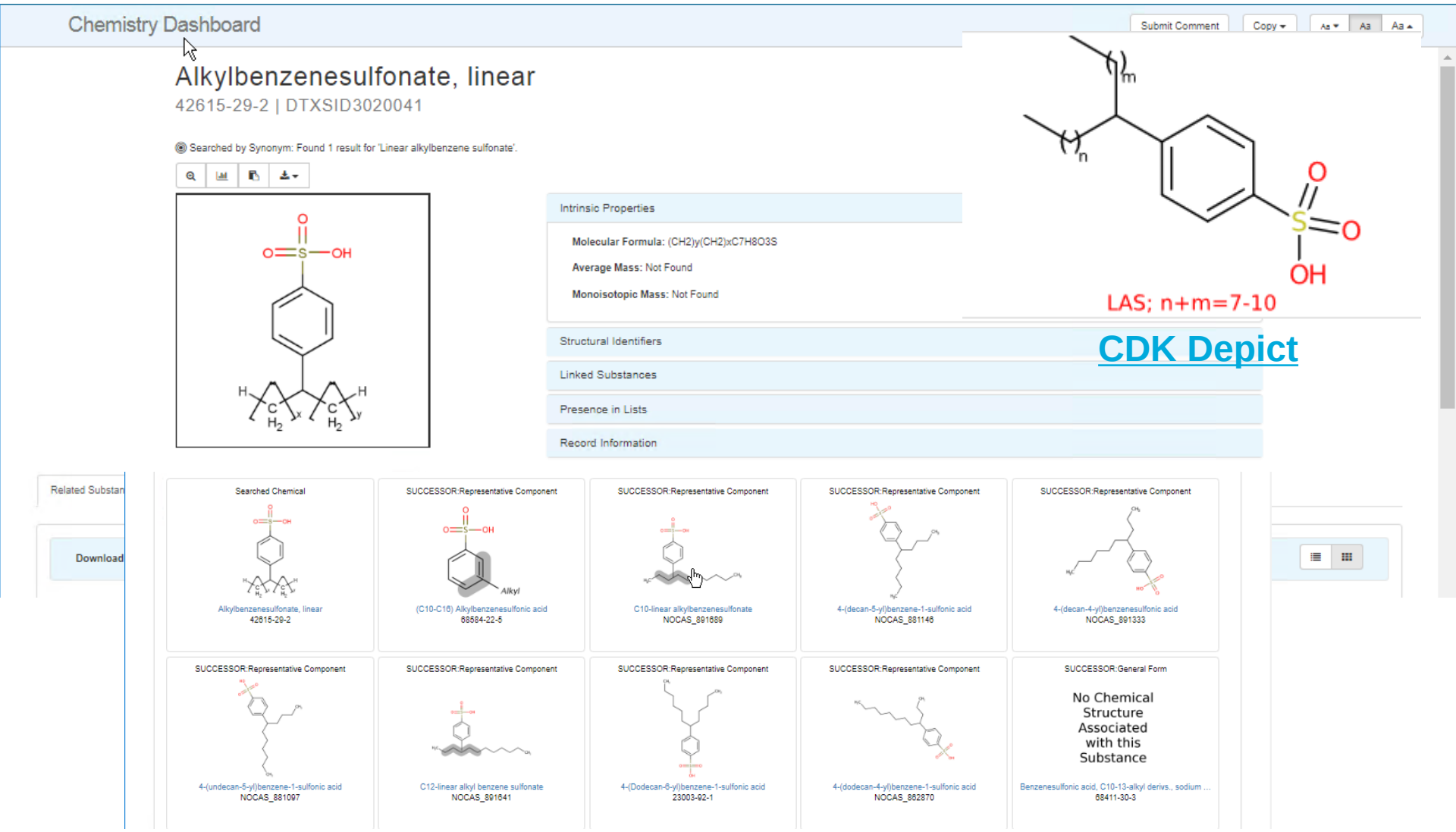
ACCESSION	NAME	NAME2	Trade Names (approx matching!)	CLASS DTXSID	REPRESENTATIVE DTXSID	Represer	CAS on Scifinder
LIT00001	C10-DATS	C10-Dialkyl tetralin sulfonate	Sirene-AlCl3	DTXSID40891636	DTXSID80891337	YES	NO
LIT00002					DTXSID00891637	YES	NO
LIT00003					DTXSID30862870	YES	NO
LIT00004					DTXSID9058600	YES	NO
LIT00005					DTXSID4059091	YES	
LIT00006					DTXSID90880419	YES	
LIT00007							
LIT00008							
LIT00009							
LIT00010							
LIT00011							
LIT00012							
LIT00013							
LIT00014							
LIT00015							

Order	SuspectID	DTXSID	PREFERRED NAME
1	Cx-LAS_class	DTXSID3020041	
2	C10toC16-LAS_c	DTXSID2028723	
5			
6			
7			
8			
9			
10			
11			
14			
19			
20			
21			
22			
23			
24			

CODE	Class_DTXSID	IDs
Cx-LAS_class	DTXSID3020041	DTXSID202 DTXSID708 DTXSID708 DTXSID408 DTXSID408 DTXSID908 DTXSID308 DTXSID30862870
C10toC16-LAS_class	DTXSID2028723	DTXSID708 DTXSID708 DTXSID408 DTXSID408 DTXSID908 DTXSID308 DTXSID30862870
C10-LAS_class	DTXSID70891689	DTXSID708 DTXSID40891333
C12_LAS_class	DTXSID90891641	DTXSID308 DTXSID30862870
Cx-SPC_class	DTXSID90891722	DTXSID808 DTXSID508 DTXSID308 DTXSID608 DTXSID008 DTXSID008 DTXSID608 DTXSID308 DTXSID608 DTXSID108 DTXSID8C
C7-SPC_class	DTXSID80891690	DTXSID50891662
C8-SPC_class		
SPA-DC_class		
Cx-DATS_cla		
C10-DATS_cl		
STAC_class		
STADCs_class		
AS_class		
AS_AES_mix		
AES_class		
SAS_class		
C12-14AEO_class	DTXSID0041931	DTXSID508 DTXSID108 DTXSID605 DTXSID501 DTXSID308 DTXSID30575670
AEO_class	DTXSID40891732	DTXSID004 DTXSID708 DTXSID808 DTXSID508 DTXSID108 DTXSID605 DTXSID501 DTXSID208 DTXSID608 DTXSID205 DTXSID3C

CODE	Class_DTXSID_Parent	Class_DTXSIDs_TPs	
Cx-LAS_class	DTXSID3020041	DTXSID90891722	DTXSID108 DTXSID708 DTXSID308 DTXSID90891727
Cx-SPC_class	DTXSID90891722	DTXSID10891724	
Cx-DATS_class	DTXSID70891725	DTXSID30891726	DTXSID90891727
STAC_class	DTXSID30891726	DTXSID90891727	
NPEO_class	DTXSID1027718	DTXSID40891691	
OPEO_class	DTXSID2042309	DTXSID60891734	

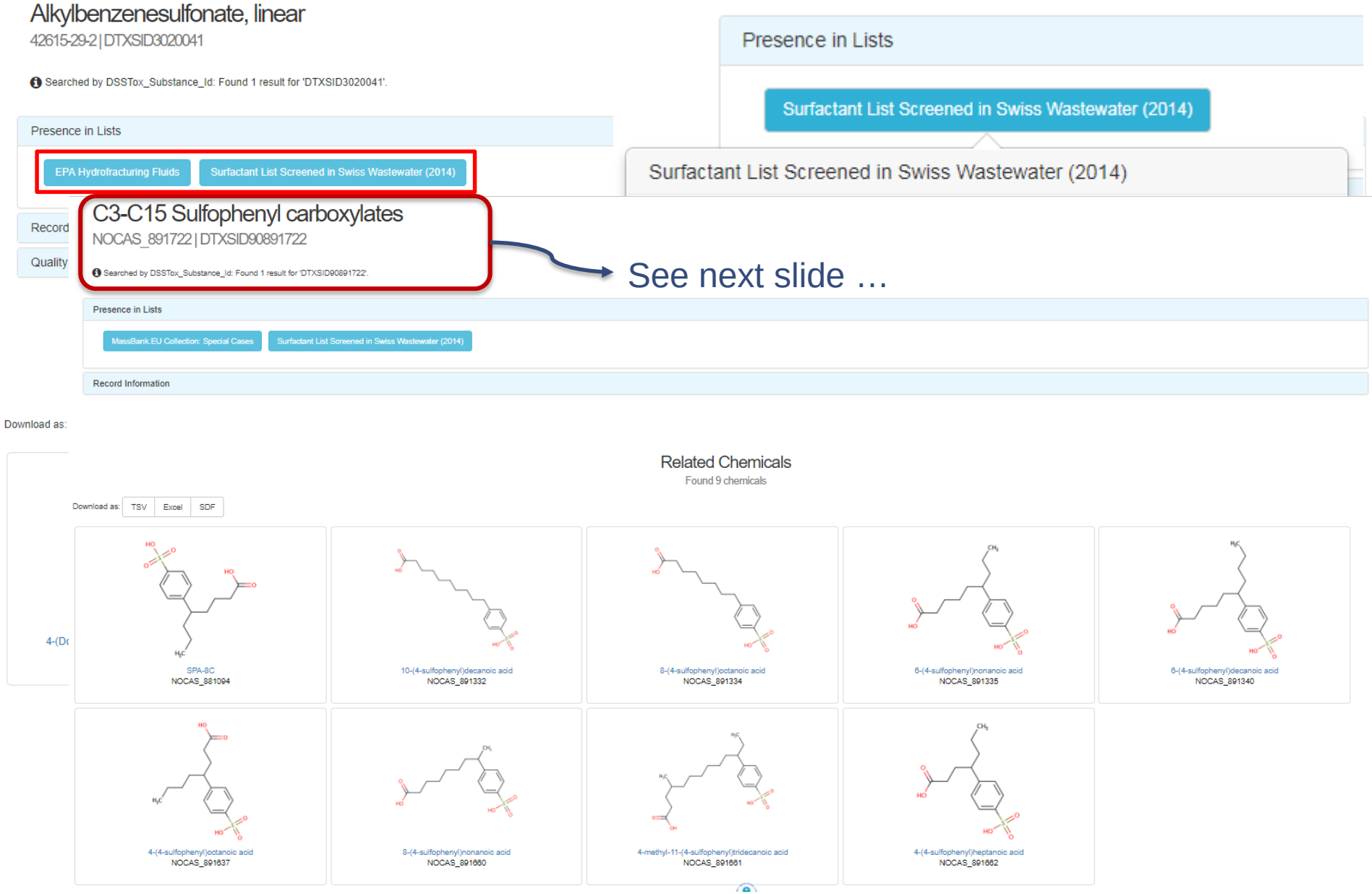
Eawag Surfactant List in CompTox Dashboard



https://comptox.epa.gov/dashboard/chemical_lists/eawagsurf

Schymanski, Grulke, Williams et al, in prep. & Williams *et al.* 2017 *J. Cheminformatics* 9:61 DOI: 10.1186/s13321-017-0247-6

Cross-Linking with Lists in CompTox Dashboard



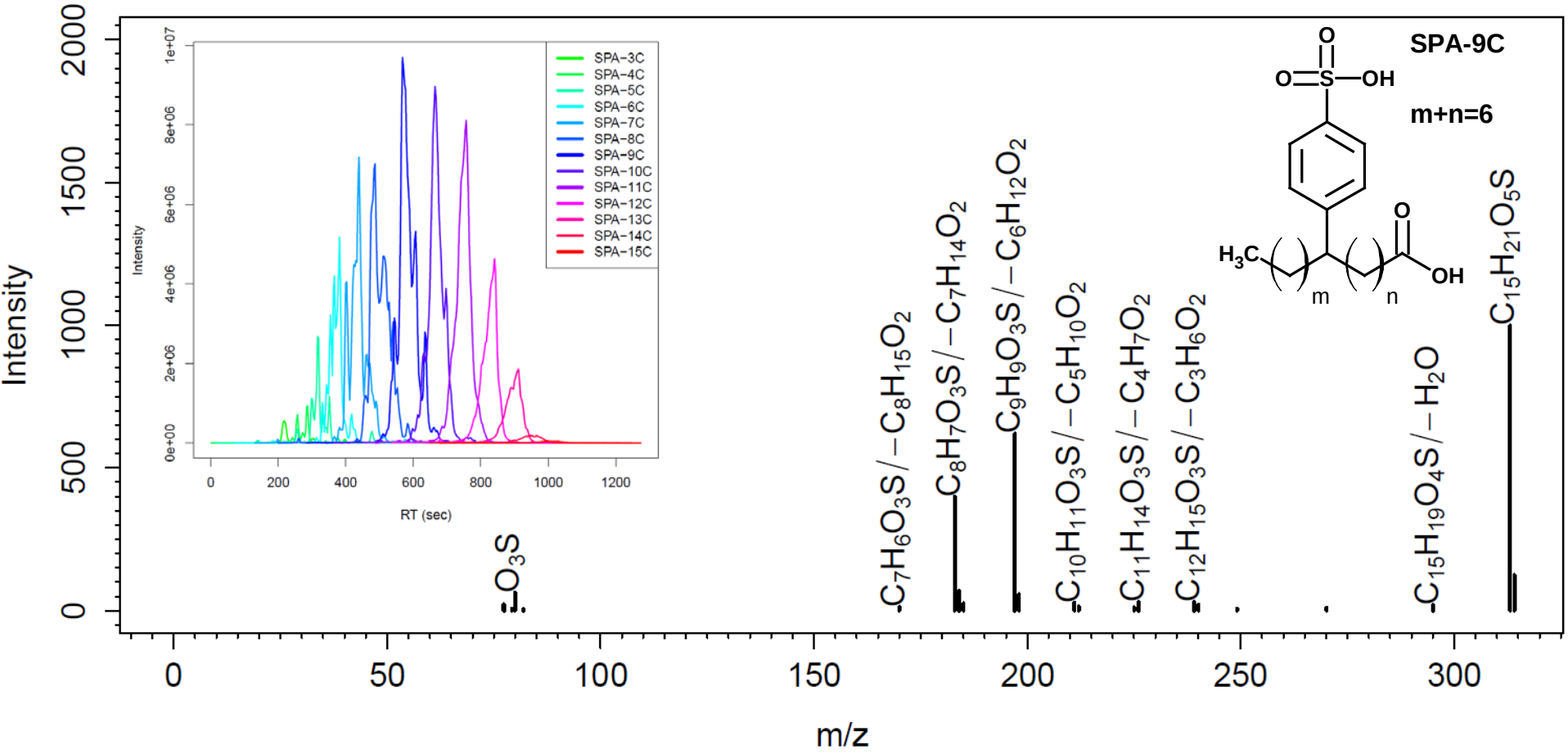
Supporting Evidence for Homologues

eawag aquatic research

Chromatography and MS/MS Annotation

RMassBank

https://github.com/MassBank/RMassBank/



Formulas: http://sourceforge.net/projects/genform/
Meringer *et al,* 2011, *MATCH* 65, 259-290
Data: Schymanski *et al*. 2014, ES&T, 48: 1811-1818. DOI: 10.1021/es4044374

MassBank.eu

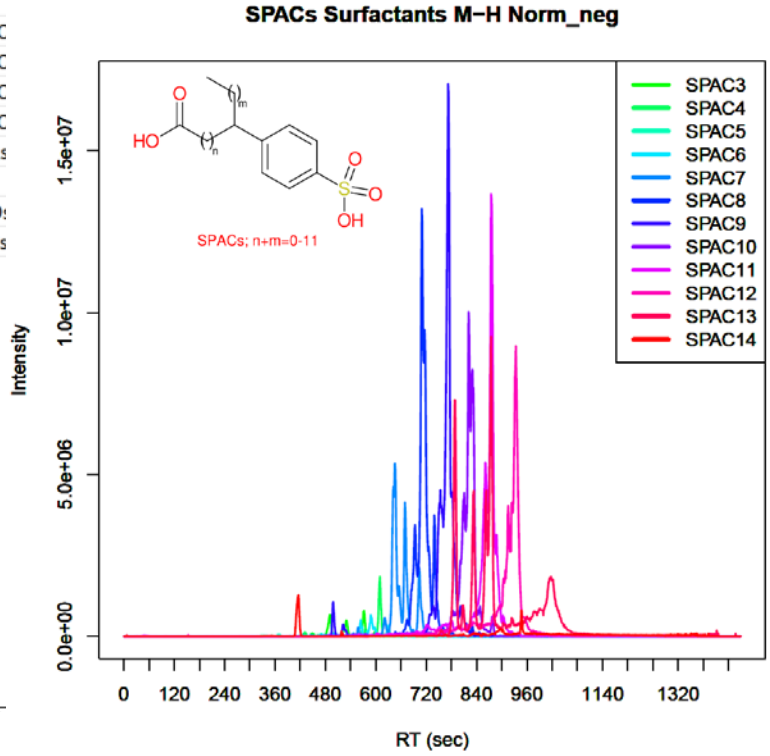
Literature: LIT00034,35
Sample: ETS00002
Standard: ETS00016,17,19,20

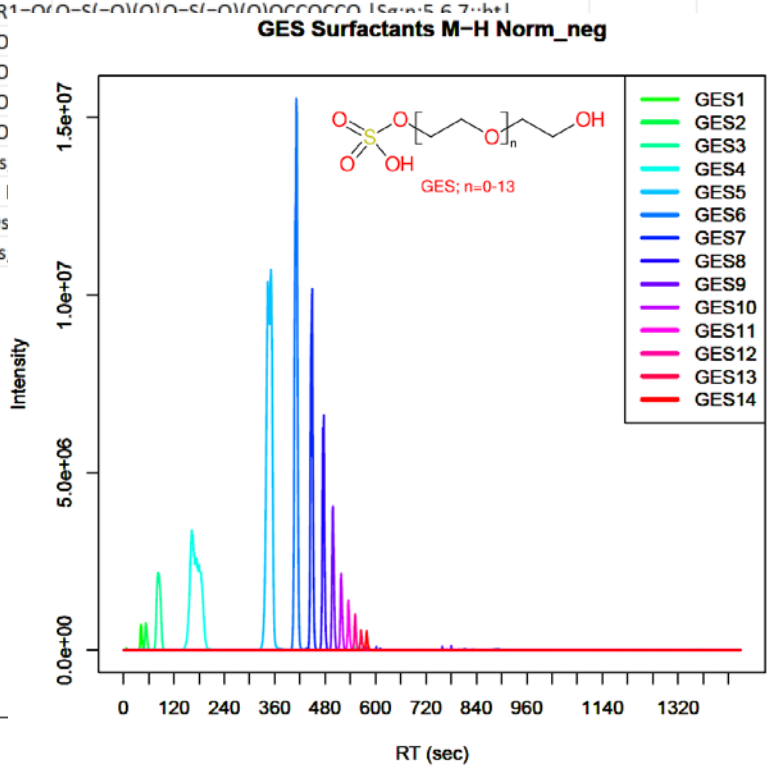
Stravs *et al*. *(*2013), *J. Mass Spectrom,* 48(1):89-99. DOI: 10.1002/jms.3131

Using Generic Structures for Screening/Linking

- https://github.com/schymane/RChemMass/

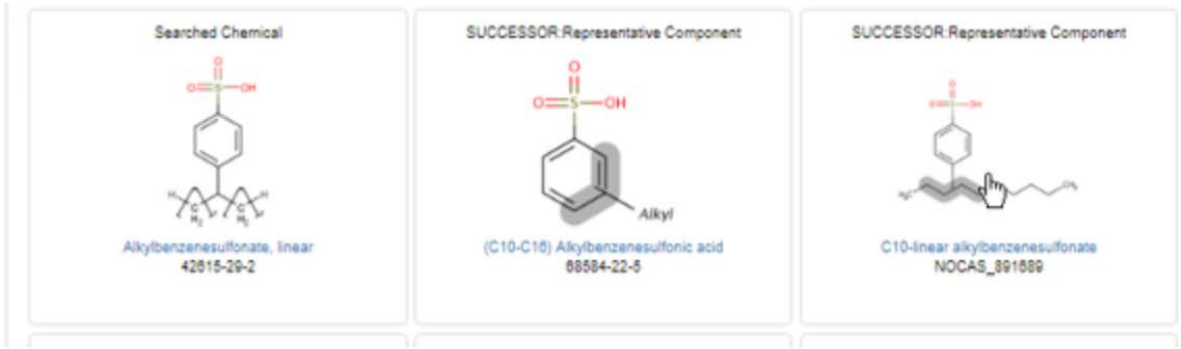
Series_ID	GenSMILES	R1toN	nR1toN	ExtraAtom	RDB_R1toN	SeriesName	Source	BabelSmile	Label1	CxSMILES_	CxSMILES_noTitle
LAS	OS(=O)(=O)	{C,C}	{9-13}			LAS	http://pubs	OS(=O)(=O)	LAS; R1=C,I	OS(=O)(=O)	OS(=O)(=O)c1ccc(cc1)C(CC)CC |Sg:n:13:m:ht,Sg:n:11:n:ht|
SPACs	OS(=O)(=O)	{C,C}	{1-13}			SPACs	http://pubs	OS(=O)(=O)	SPACs; R1=	OS(=O)(=O)	OS(=O)(=O)c1ccc(cc1)C(CC(=O)O)CC |Sg:n:15:m:ht,Sg:n:11:n:ht|
SPADCs	OS(=O)(=O)	{C,C}	{0-13}			SPADCs	http://pubs	OS(=O)(=O)	SPADCs; R1	OS(=O)(=O)	OS(=O)(=O)c1ccc(cc1)C(CC(=O)O)CC(=O)O |Sg:n:15:m:ht,Sg:n:11:n:ht|
DATS	OS(=O)(=O)	{C,C}	{0-15}			DATS	http://pubs	OS(=O)(=O)	DATS; R1=(	OS(=O)(=O)	OS(=O)(=O)c1ccc2c(c1)C(CCC2CC)CC |Sg:n:14:m:ht,Sg:n:16:n:ht|
STACs	OS(=O)(=O)	{C,C}	{1-13}			STACs	http://pubs	OS(=O)(=O)	STACs; R1=	OS(=O)(=O)	OS(=O)(=O)c1ccc2c(c1)C(CCC2CC(=O)O)CC |Sg:n:14:m:ht,Sg:n:18:n:ht|
STADCs	OS(=O)(=O)	{C,C}	{0-13}			STADCs	http://pubs	OS(=O)(=O)	STADCs; R1	OS(=O)(=O)	OS(=O)(=O)c1ccc2c(c1)C(CCC2CC(=O)O)CC(=O)O |Sg:n:14:m:ht,Sg:n:18:n:ht|
AS	O=S(=O)(O)	{C}	{11-15}			AS	http://pubs	O=S(=O)(O)	AS; R1=C(1	O=S(=O)(O)	O=S(=O)(O)OCC |Sg:n:5:n:ht|
C12AES	O=S(O)(=O)	{CCO,C}	{2-11,11-11}			C12AES	http://pubs	O=S(O)(=O)	C12AES; R1	O=S(O)(=O)	O=S(O)(=O)OCCOCCC |Sg:n:5,6,7::ht,Sg:n:9:m:ht|
C13AES	O=S(O)(=O)	{CCO,C}	{2-11,12-12}			C13AES	http://pubs	O=S(O)(=O)	C13AES; R1	O=S(O)(=O)	O=S(O)(=O)OCCOCCC |Sg:n:5,6,7::ht,Sg:n:9:m:ht|
SAS	O=S(O)(=O)	{C,C}	{9-20}			SAS	http://pubs	O=S(O)(=O)	SAS; R1=C,I	O=S(O)(=O)	O=S(O)(=O)C(CC)CC |Sg:n:5:m:ht,Sg:n:7:n:ht|
NPEOSO4s	O=S(=O)(O)	{OCC,C}	{0-2,9-9}			NPEOSO4s	http://pubs	O=S(=O)(O)	NPEOSO4s;	O=S(=O)(O)	O=S(=O)(O)OCCOc1ccc(cc1)CC |Sg:n:5,6,7::ht,Sg:n:14:m:ht|
GES											
CxAEC											
CxAEC											
CxAEC											
CxAEC											
CDEAs											
PEGs											
NPEOs											
NPECs											

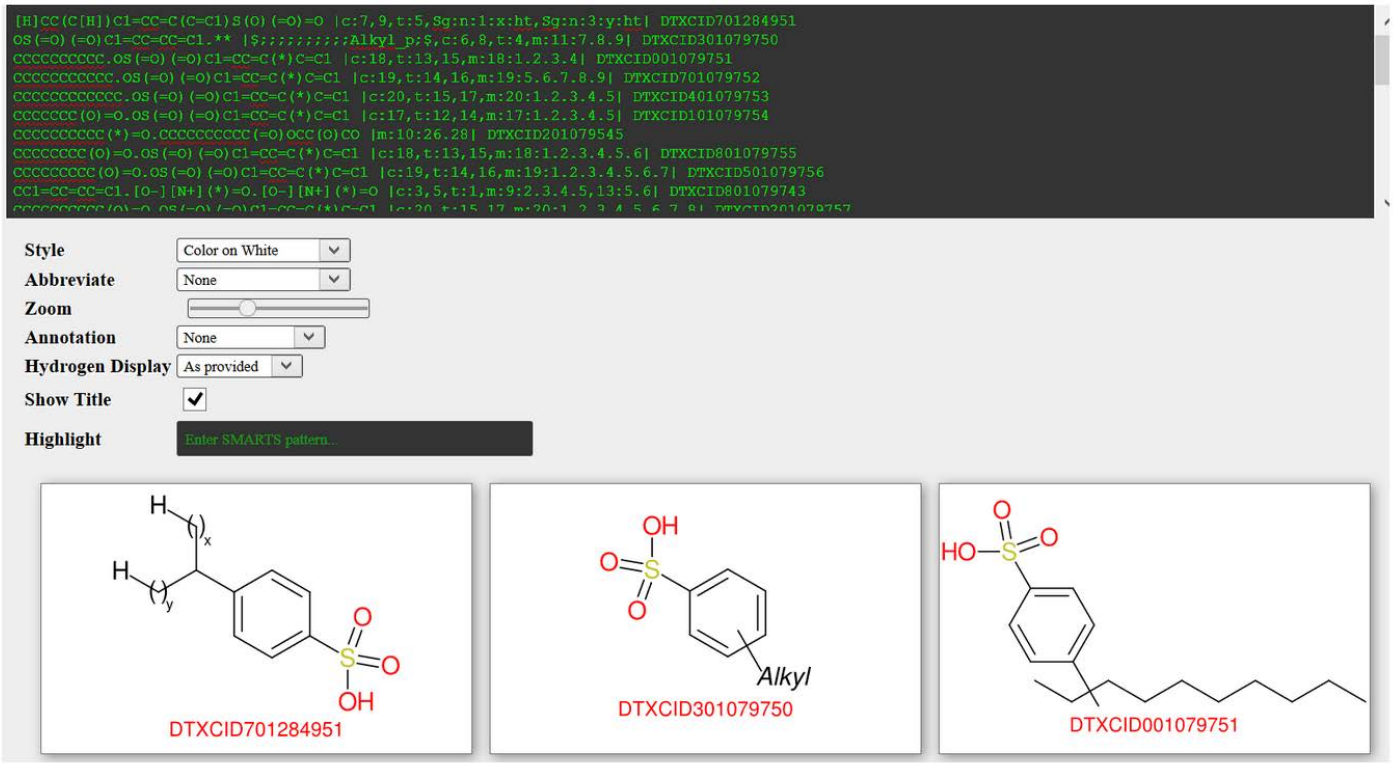




Toolkit compatibility will be vital for us …

- https://github.com/cdk/depict/issues/7

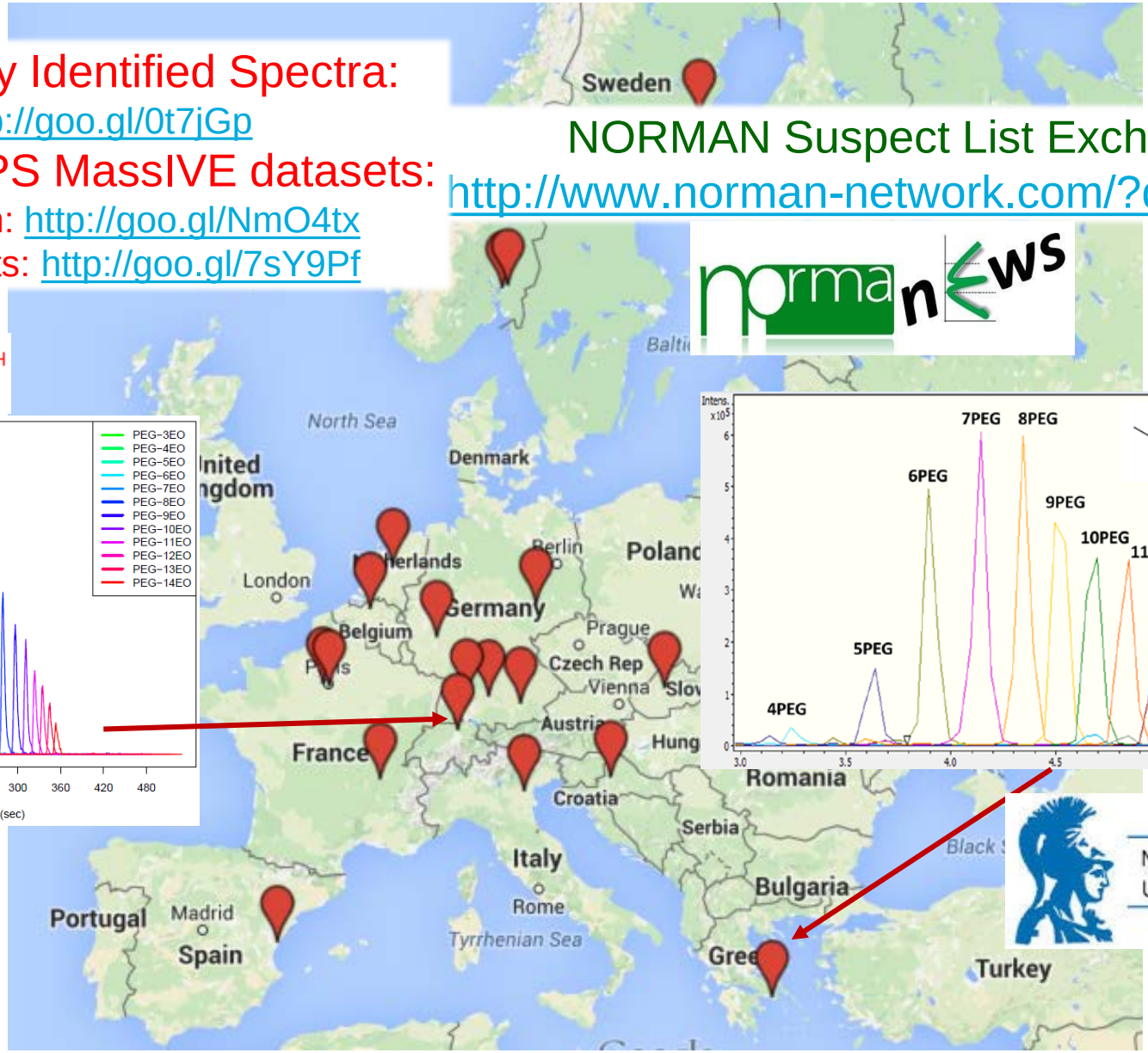


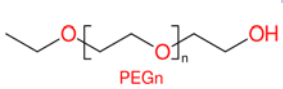


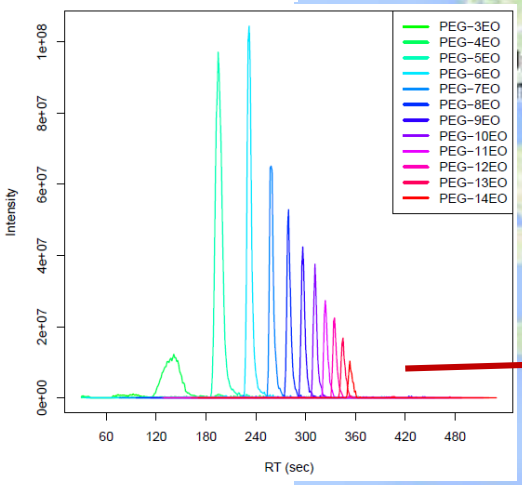
… to enable world-wide exchange of suspects

Tentatively Identified Spectra:
http://goo.gl/0t7jGp
Hits in GNPS MassIVE datasets:
TPs in skin: http://goo.gl/NmO4tx
Surfactants: http://goo.gl/7sY9Pf

NORMAN Suspect List Exchange:
http://www.norman-network.com/?q=node/236







eawag aquatic research

National and Kapodistrian UNIVERSITY OF ATHENS

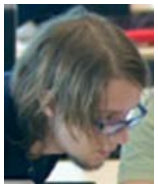
Schymanski et al. 2015, *ABC*, DOI: 10.1007/s00216-015-8681-7

Acknowledgements







RMassBank

nontarget

enviPat Web

enviPick

enviMass

MetFrag 2.3


MassBank.eu

Stellan Fischer KEMI


ETH zürich
Department of Biology
Institute of Molecular Systems Biology

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY

emma.schymanski@uni.lu

Further Information:

www.massbank.eu

http://www.norman-network.com/?q=node/236

https://github.com/MassBank/RMassBank/

https://comptox.epa.gov/dashboard/

https://wwwen.uni.lu/lcsb/

norman

solutions

SEVENTH FRAMEWORK PROGRAMME

EU Grant 603437

LCSB

uni.lu UNIVERSITÉ DU LUXEMBOURG

2015: European Non-target Screening Trial

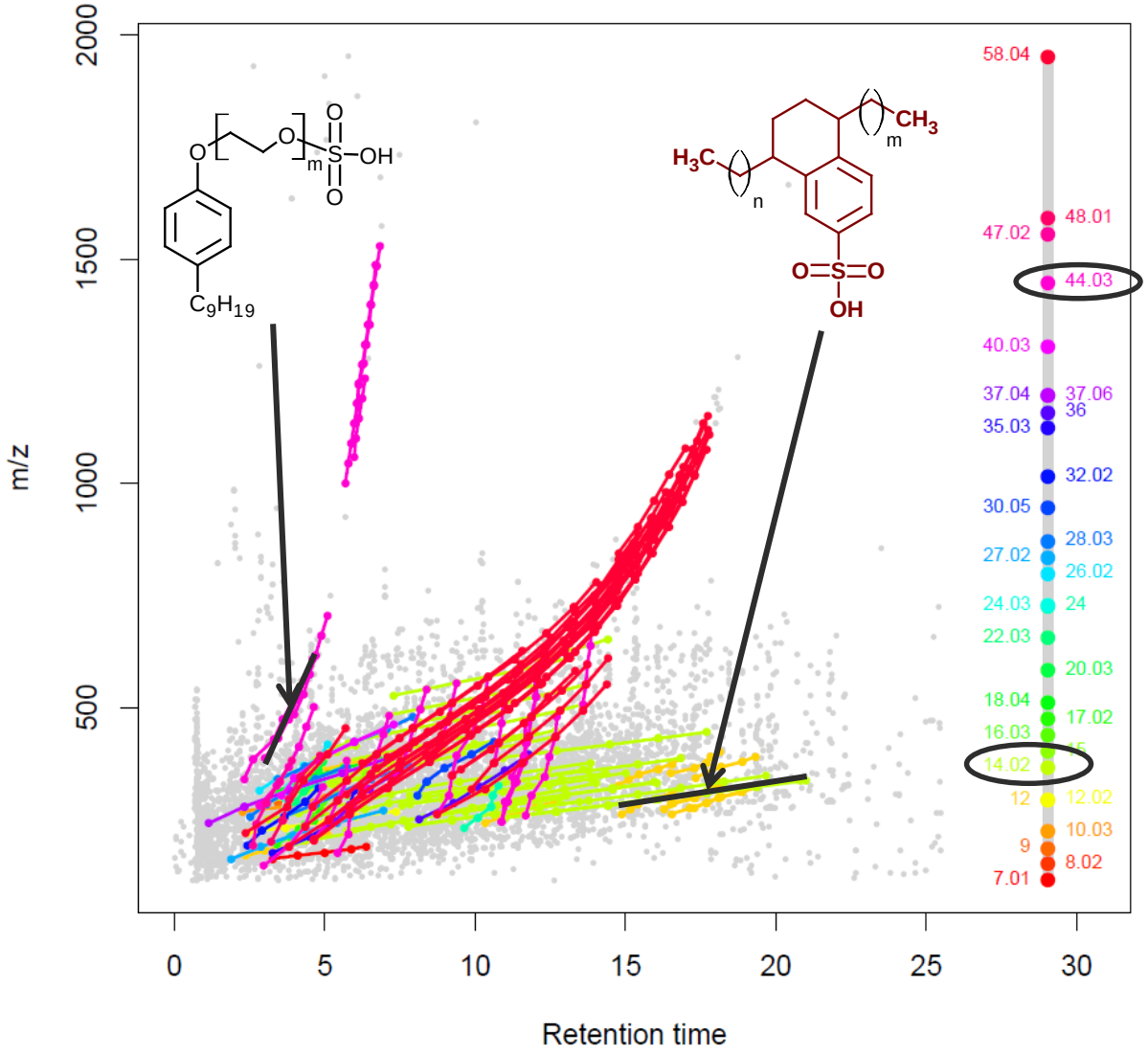
Schymanski et al. 2015, ABC, DOI: 10.1007/s00216-015-8681-7

Homologous Series Detection

eawag aquatic research

nontarget

Search for discrete mass differences



http://www.envihomolog.eawag.ch/

LCSB

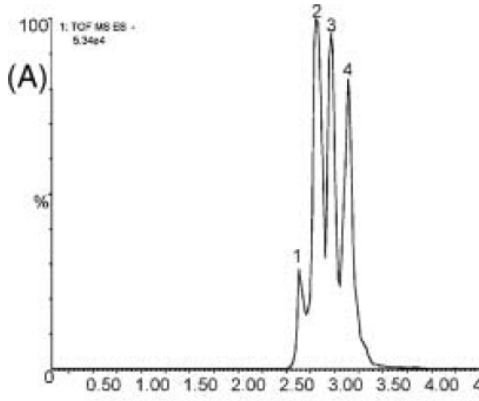
UNIVERSITÉ DU LUXEMBOURG

M. Loos & H Singer, 2017. J. Cheminf. DOI: 10.1186/s13321-017-0197-z & Schymanski et al. 2014, ES&T DOI: 10.1021/es4044374

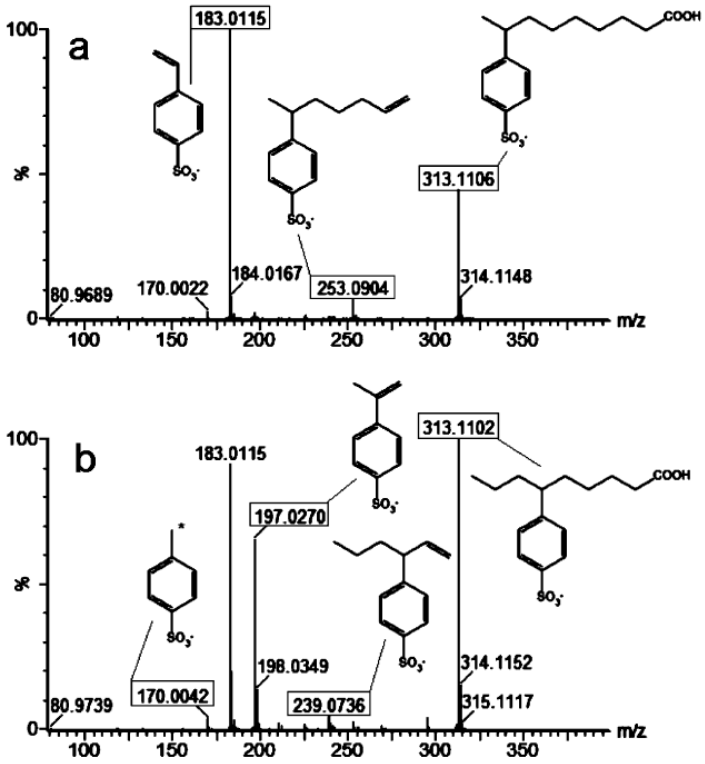
Surfactant Screening From Literature

Literature sources

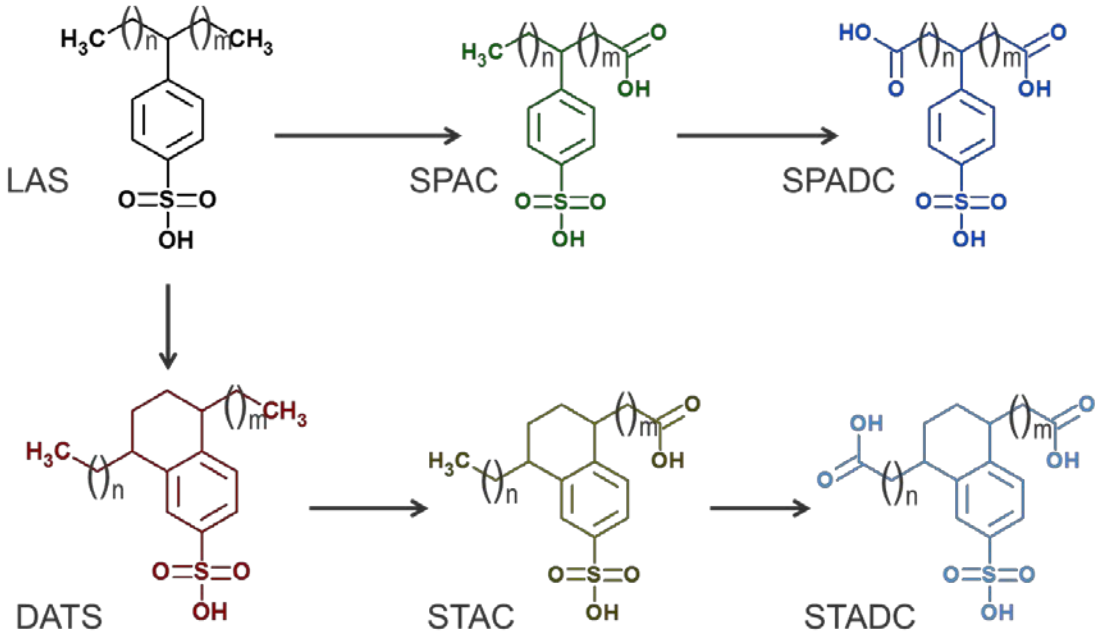
- Formulas, masses (ions), retention times and intensities
- Spectra of selected compounds (different instruments)



Gonzalez et al. Rapid Comm. Mass Spec. 2008, 22: 1445-54

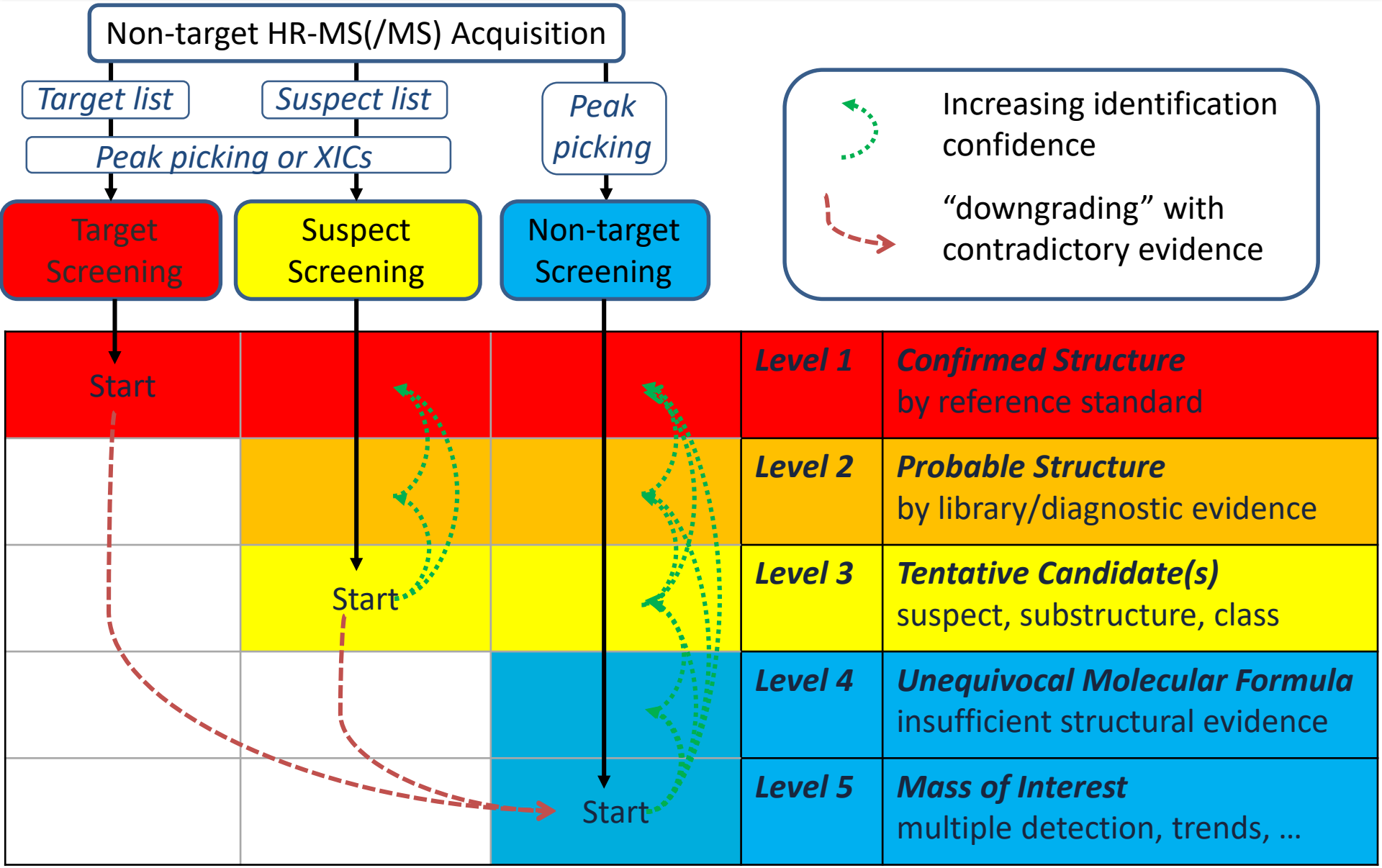


Lara-Martin et al. EST. 2010, 44: 1670-1676



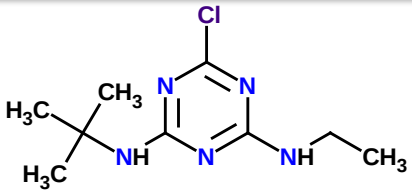
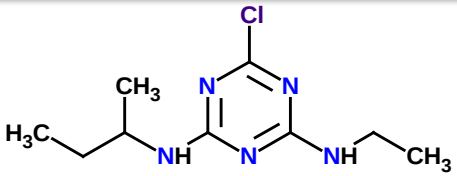
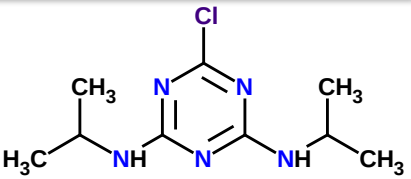
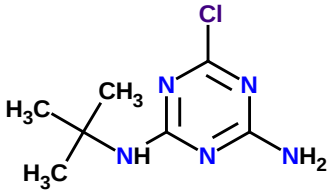
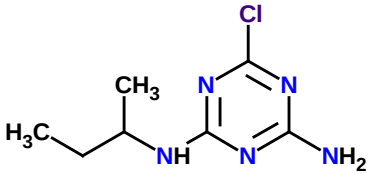
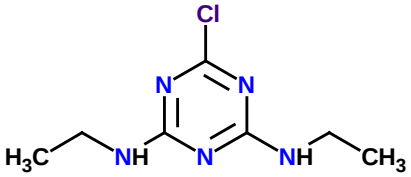
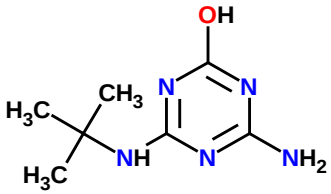
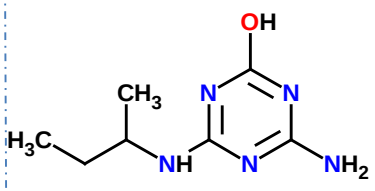
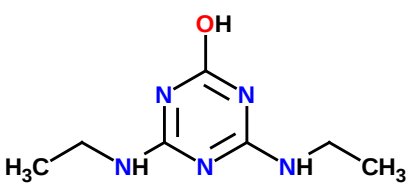
Schymanski *et al*. *(*2014), *ES&T,* 48: 1811-1818. DOI: 10.1021/es4044374

2015: European Non-target Screening Trial



Schymanski *et al*, 2014, ES&T. DOI: 10.1021/es5002105 & Schymanski *et al*. 2015, ABC, DOI: 10.1007/s00216-015-8681-7

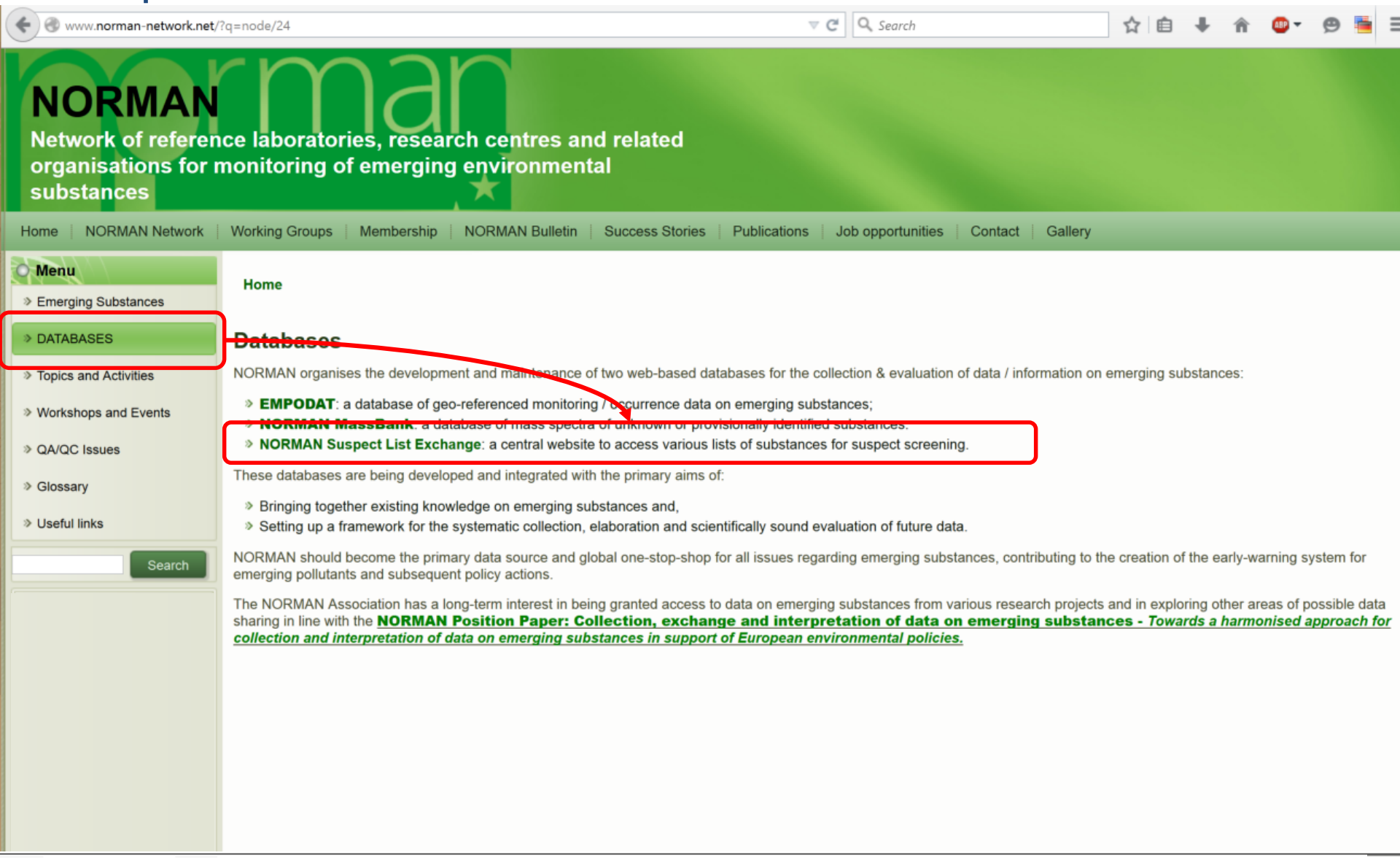
Reported Identification Depended on Source List

$C_9H_{16}ClN_5$ *m/z 229.1094 Da*	 Terbutylazine Detects: 12; # Refs: 220	 Sebutylazine Detects: 3; # Refs: 51	(no related compound at this mass)	 Propazine Detects: 3; # Refs: 201
$C_7H_{12}ClN_5$ *m/z 201.0781 Da*	 Terbutylazine-desethyl Detects: 9; # Refs: 92	 Sebutylazine-desethyl Detects: 1; # Refs: 14	 Simazine Detects: 4; # Refs: 518	(no related compound at this mass)
$C_7H_{13}N_5O$ *m/z 183.1120 Da*	 Terbutylazine-desethyl-2-hydroxy Detects: 2; # Refs: 57	 Sebutylazine-desethyl-2-hydroxy Detects: 0; # Refs: 3	 Simazine-2-hydroxy Detects: 2; # Refs: 66	(no related compound at this mass)

Schymanski et al. 2015, ABC, DOI: 10.1007/s00216-015-8681-7

Enter: NORMAN Suspect Exchange

- …part of the NORMAN Databases Collection



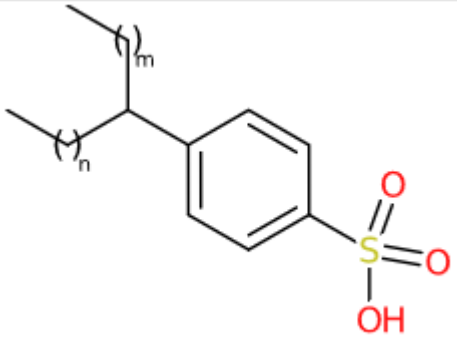
Schymanski et al. 2015, ABC, DOI: 10.1007/s00216-015-8681-7

Eawag Surfactant List in CompTox Dashboard

Alkylbenzenesulfonate, linear

42615-29-2 | DTXSID3020041

Searched by Synonym: Found 1 result for 'Linear alkylbenzene sulfonate'.



LAS; n+m=7-10

CDK Depict

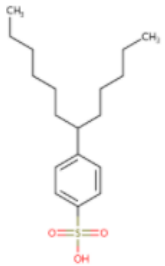
Presence in Lists

Surfactant List Screened in Swiss Wastewater (2014)

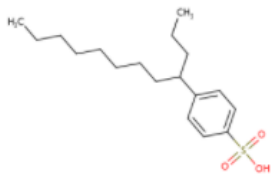
Surfactant List Screened in Swiss Wastewater (2014)

EAWAGSURF is a list of surfactants screened in Swiss wastewater effluents as part of a 2014 study. Structures/mixtures are being progressively curated and linked (Schymanski/Williams). Further details in Schymanski et al 2014, DOI: 10.1021/es4044374

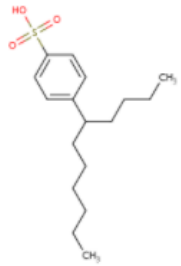
cals



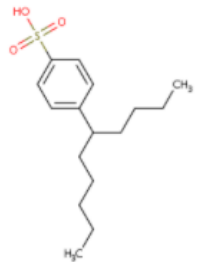
4-(Dodecan-6-yl)benzene-1-sulfon...
23003-92-1



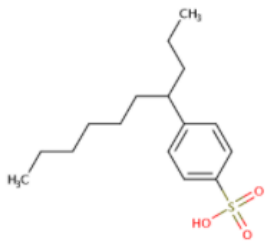
4-(dodecan-4-yl)benzene-1-sulfoni...
NOCAS_862870



C11-LAS
NOCAS_881097



4-(decan-5-yl)benzene-1-sulfonic ...
NOCAS_881146



4-(decan-4-yl)benzenesulfonic acid
NOCAS_891333

https://www.slideshare.net/AntonyWilliams/markush-enumeration-to-manage-mesh-and-manipulate-substances-of-unknown-or-variable-composition

Curation never stops …

Undefined mixtures (UVCBs)

Cleaning up lists to remove errors

Mol_ID	Name	EDITED NAMES FOR INPUT INTO SEARCH	CAS_RN	Merged DTXSIDs	DTXSID Based on Name	Preferred Name
SA8750	By-Product	By-Product	NA	-	-	NO_MATCH
stpQQR1546	C10-DATS|C10-Dialkyl tetr	C10-DATS|C10-Dialkyl tetralin sulfonate|8	NA	-	-	NO_MATCH
SA2074	C10-LAS	C10-LAS	NA	-		NO_MATCH
stpQQR1582	C10LAS|C10-linear alkylbe	C10LAS|C10-linear alkylbenzyl sulfonate|4	NA	-	-	NO_MATCH
SA14931	C10-phosphonic	C10-phosphonic	NA	-	-	NO_MATCH
StpBB815	C12-15 ALKYL BENZOATE	C12-15 ALKYL BENZOATE	68411-27-8	-	-	NO_MATCH
SA13282	C12-AE5S	C12-AE5S	NA	-	-	NO_MATCH
stpQQR1548	C12-LAS|C12-linear alkyl b	C12-LAS|C12-linear alkyl benzene sulfonat	NA	-	-	NO_MATCH
stpQQR690	C14-SAS (TENTATIVE)|tetr	C14-SAS (TENTATIVE)|tetradecane-7-sulfo	NA	-	-	NO_MATCH
stpQQR1557	C16EOx|C16EO2|C16-alcc	C16EOx|C16EO2|C16-alcohol polyethoxyl	NA	-	-	NO_MATCH
stpQQR1556	C18EOx|C18EO2|C18-alcc	C18EOx|C18EO2|C18-alcohol polyethoxyl	4439-32-1	-	-	NO_MATCH
SA14932	C4-phosphonic	C4-phosphonic	NA	-	-	NO_MATCH
SA14929	C6-phosphonic	C6-phosphonic	NA	-	-	NO_MATCH
stpQQR1583	C7SPC|C7-sulfophenyl car	C7SPC|C7-sulfophenyl carboxylates|4-(de	NA	-	-	NO_MATCH
SA14930	C8-phosphonic	C8-phosphonic	NA	-	-	NO_MATCH
stpQQR1547	C8-SPC|C8-Sulfophenyl ca	C8-SPC|C8-Sulfophenyl carboxylic acid|4-(	NA	-	-	NO_MATCH
stpQQR1576	CA5PE2C|7-{4-[2-(carboxy	CA5PE2C|7-{4-[2-(carboxymethoxy)ethoxy	NA	-	-	NO_MATCH
stpQQR1578	CA6PE2	CA6PE2	NA	-	-	NO_MATCH
stpQQR1577	CA6PE2C	CA6PE2C	NA	-	-	NO_MATCH
stpQQR1575	CA8PE2C	CA8PE2C	NA	-	-	NO_MATCH
SA9863	cacotheline	cacotheline	561-20-6	-	-	NO_MATCH
SAn15715	Caerulomycin A	Caerulomycin A	21802-37-9	-	-	NO_MATCH
SA5151	cafedrine	cafedrine	58166-83-9	-	-	NO_MATCH

(many) more registrations…

Target, Suspect and Non-Target Screening

KNOWNS → **TARGET ANALYSIS** → Targets found

SUSPECTS → **SPECTRUM SEARCH** → Spectral match

SUSPECTS → **SUSPECT SCREENING** → Suspects found

No Prior Knowledge → **NON-TARGET SCREENING** → Masses of interest → (Molecular formula) → **DATABASE SEARCH** / **STRUCTURE GENERATION**

Candidate selection (retention time, MS/MS, calculated properties)

Time, Effort & Number of Compounds….