http://www.orcid.org/0000-0002-2668-4821

Using Cheminformatics Approaches to Develop a Structure Searchable Database of Analytical Methods

Antony Williams, *Greg Janesch, Sakuntala Sivasupramaniam, Tyler Carr and Nancy Baker*

Center for Computational Toxicology & Exposure, U.S. Environmental Protection Agency

August 2023: ACS Fall Meeting, San Francisco, CA

Disclaimer

- The views expressed in this presentation are those of the author and do not necessarily reflect the views or policies of the U.S. EPA

- This presentation is on a proof-of-concept tool in development – NOT yet publicly available

Building a Methods Database

- Simple Vision: I want to find the best method(s) associated with a ***chemical*** and/or **class of chemicals**

- Answer the question “I cannot find a method for my chemical” - HELP

- The Approach:
  - Aggregate MS method documents (and adjust the definition of “what is a useful method”)
  - Extract chemistry (mostly CASRN and Names)
  - Map CASRN and Names to structures
  - Deliver a proof-of-concept application to search a database by names, CASRNs, InChIKeys and ultimately structure

Most people start with Google

- People have many places to search for methods, and there is no one integration hub, except for search engines

- Search engines can return so many hits – then you filter by analytes, matrix, analytical methodology, so many synonyms and abbreviations for so many chemicals

Synonyms, Abbreviations and Chemicals

Open Access | Published: 12 August 2015

Fast analysis of 29 polycyclic aromatic hydrocarbons (PAHs) and nitro-PAHs with ultra-high performance liquid chromatography-atmospheric pressure photoionization-tandem mass spectrometry

	0.5% anisole in toluene (dopant A)		
			Accuracy a
Compound	Linear Range (ng/mL)	R^2	10 ng/mL
ACPY	50–500	0.9978	98.8[a] (7.6)
ACP	5–200	0.9998	98.0(3.6)
FLU	20–500	0.9954	103[b] (2.5)
PHEN	2–500	0.997	101(6.0)
ANTHR	2–500	0.9976	101(5.1)
FL	2–200	0.9978	107(4.1)
PYR	2–200	0.9994	105(3.8)
BNT	2–500	0.9982	98.2(7.4)
CPP	2–500	0.9996	98.6(6.5)
BAA	2–500	0.9996	99.9(7.6)
CHRY	2–500	0.9994	98.9(5.3)

ELSEVIER

Food Control

Volume 62, April 2016, Pages 322-329

Simultaneous analysis of twenty-six mycotoxins in durum wheat grain from Italy

The certified standards of AFB1, AFB2, AFG1, AFG2, OTA, STG, ZEN, NIV, DON, 3-AcDON, 15-AcDON, DAS, NEO, T-2 and HT-2 toxin, FB1, FB2, FB3, BEA, ENNs (A, A1, B, B1), AOH, AME, and TEN were purchased from Sigma Aldrich (Madrid, Spain).

The individual stock solutions of AFB1, AFB2, AFG1, AFG2, OTA, STG, ENs (A, A1, B, B1), BEA, AOH, AME, and TEN at 500μg/mL were prepared in acetonitrile, whereas ZEN, FB1, FB2, FB3, NIV, DAS, NEO, DON, 3-AcDON, 15-AcDON, T-2 and HT-2 toxin were prepared at 1000μg/mL in methanol. A working mixed standard solution at 5 and

- CAS Numbers, Names and Abbreviations can limit what's possible…

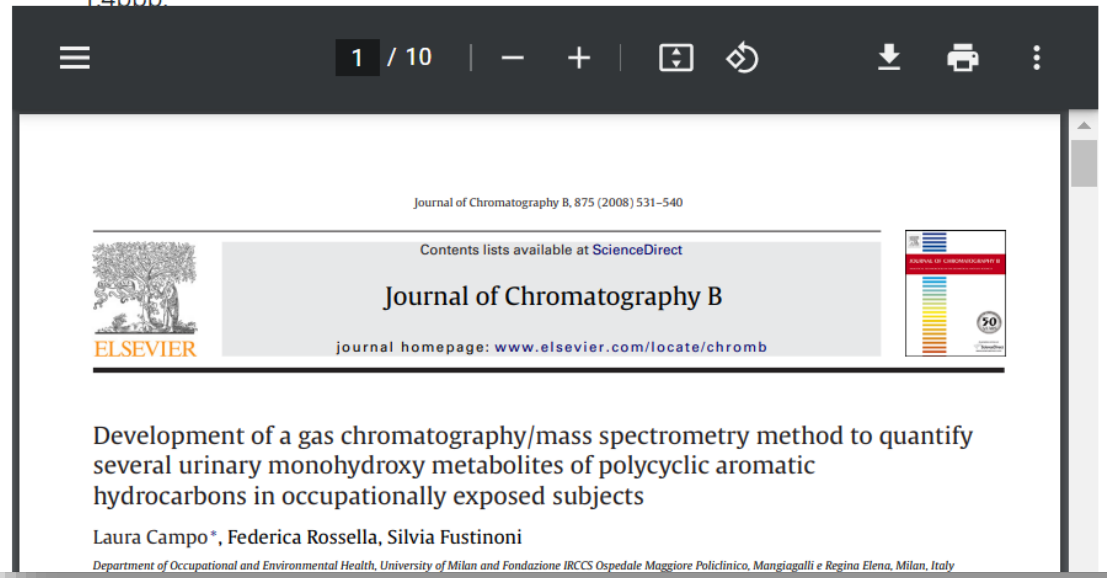
Might this be a better view?

Development of a gas chromatography/mass spectrometry method to quantify several urinary monohydroxy metabolites of polycyclic aromatic hydrocarbons in occupationally exposed subjects

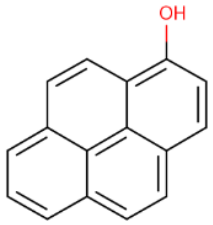
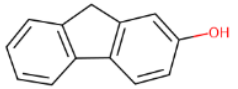
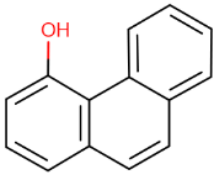
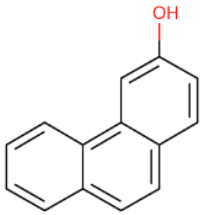
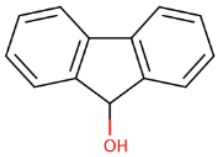
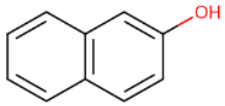
Author: Laura Campo, Federica Rossella, Silvia Fustinoni

Focus/Analyte: PAHs and their degradates

Synopsis: Determination of pAHs and their degradates in by GC/MS with a LOQ of 0.1-1.4ppb.



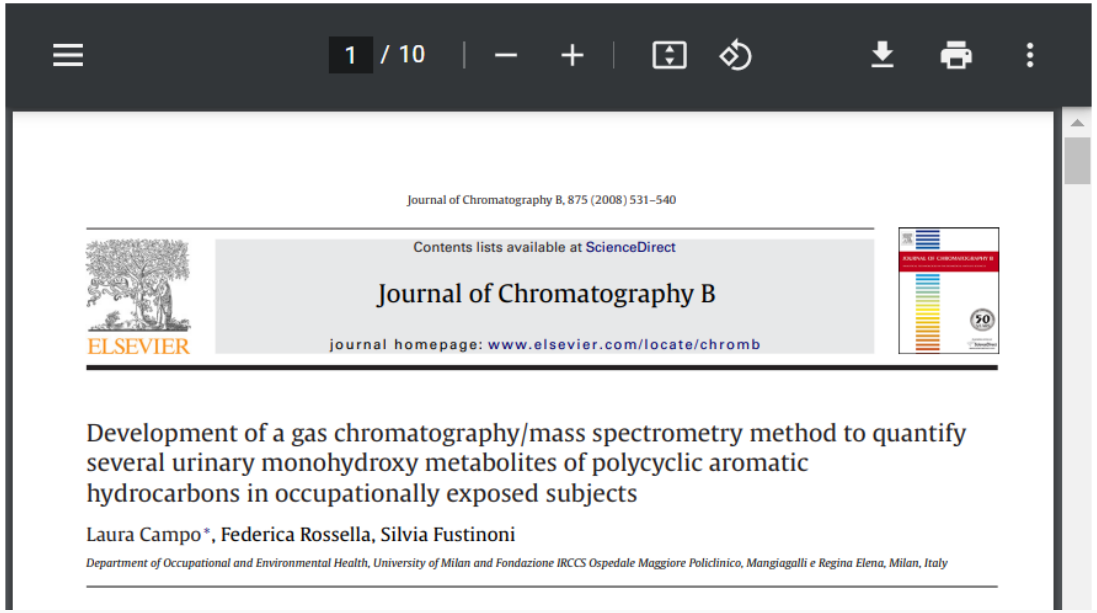
Compounds (12) (grid) | Compounds (12) (table)

1-Hydroxypyrene  DTXSID1038298 ↗	2-Hydroxyfluorene  DTXSID2047569 ↗	4-Hydroxyphenanthrene  DTXSID30227324 ↗
3-Hydroxyphenanthrene  DTXSID3052723 ↗	Fluoren-9-ol  DTXSID4052683 ↗	2-Naphthalenol  DTXSID5027061 ↗

Might this be a better view?

Development of a gas chromatography/mass spectrometry method to quantify several urinary monohydroxy metabolites of polycyclic aromatic hydrocarbons in occupationally exposed subjects

Author: Laura Campo, Federica Rossella, Silvia Fustinoni
Focus/Analyte: PAHs and their degradates
Synopsis: Determination of pAHs and their degradates in by GC/MS with a LOQ of 0.1-1.4ppb.



Compounds (12) (grid)

Compounds (12) (table)

Download Compound Info

Structure	DTXSID	Links	CASRN	Compound Name
	DTXSID1038298	Search	5315-79-7	1-Hydroxypyrene
	DTXSID2047569	Search	2443-58-5	2-Hydroxyfluorene
	DTXSID30227324	Search	7651-86-7	4-Hydroxyphenanthrene
	DTXSID3052723	Search	605-87-8	3-Hydroxyphenanthrene
	DTXSID4052683	Search	1689-64-1	Fluoren-9-ol
	DTXSID5027061	Search	135-19-3	2-Naphthalenol
	DTXSID6021793	Search	90-15-3	1-Naphthol

When methods are mapped to chemistry…

- The advantages of mapping chemicals directly to methods
  - When chemicals are mapped it opens access to many other tools
  - Chemical structures allow for QSAR modeling

Transparency in Modeling through Careful Application of OECD's QSAR/QSPR Principles via a Curated Water Solubility Data Set

Charles N. Lowe*, Nathaniel Charest*, Christian Ramsland, Daniel T. Chang, Todd M. Martin, and Antony J. Williams

Cite this: *Chem. Res. Toxicol.* 2023, 36, 3, 465−478 Article Views Altmetric Citations Share Add

Predicting compound amenability with liquid chromatography-mass spectrometry to improve non-targeted analysis

Charles N. Lowe, Kristin K. Isaacs, Andrew McEachran, Christopher M. Grulke, Jon R. Sobus, Elin M. Ulrich, Ann Richard, Alex Chao, John Wambaugh & Antony J. Williams

Analytical and Bioanalytical Chemistry **413**, 7495–7508 (2021) | Cite this article

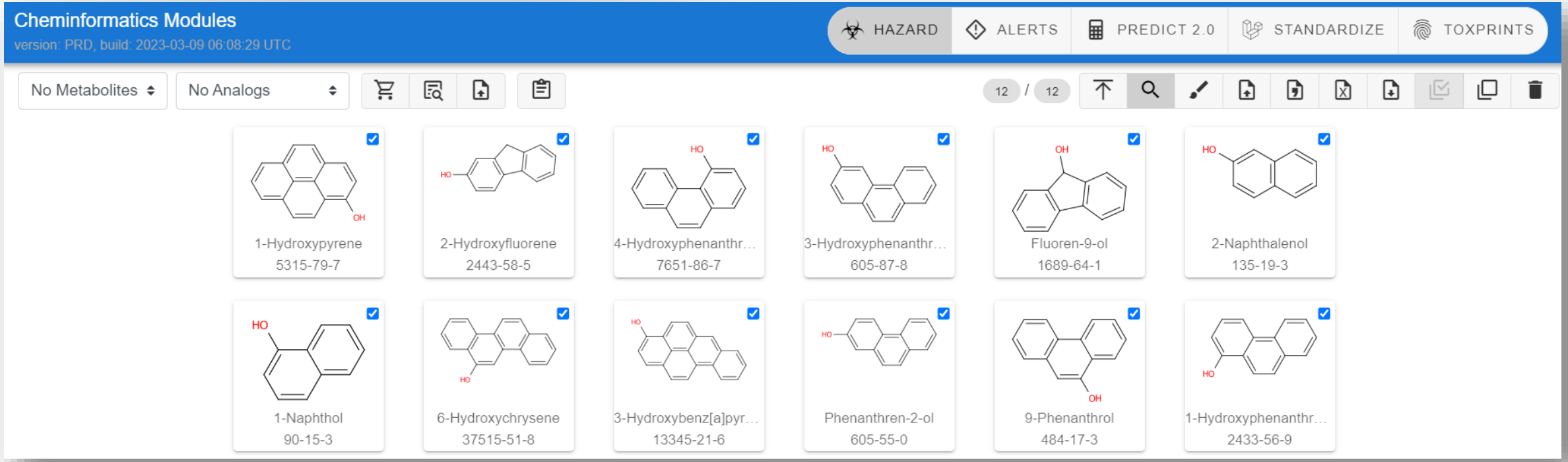
OPERA models for predicting physicochemical properties and environmental fate endpoints

Kamel Mansouri, Chris M. Grulke, Richard S. Judson & Antony J. Williams

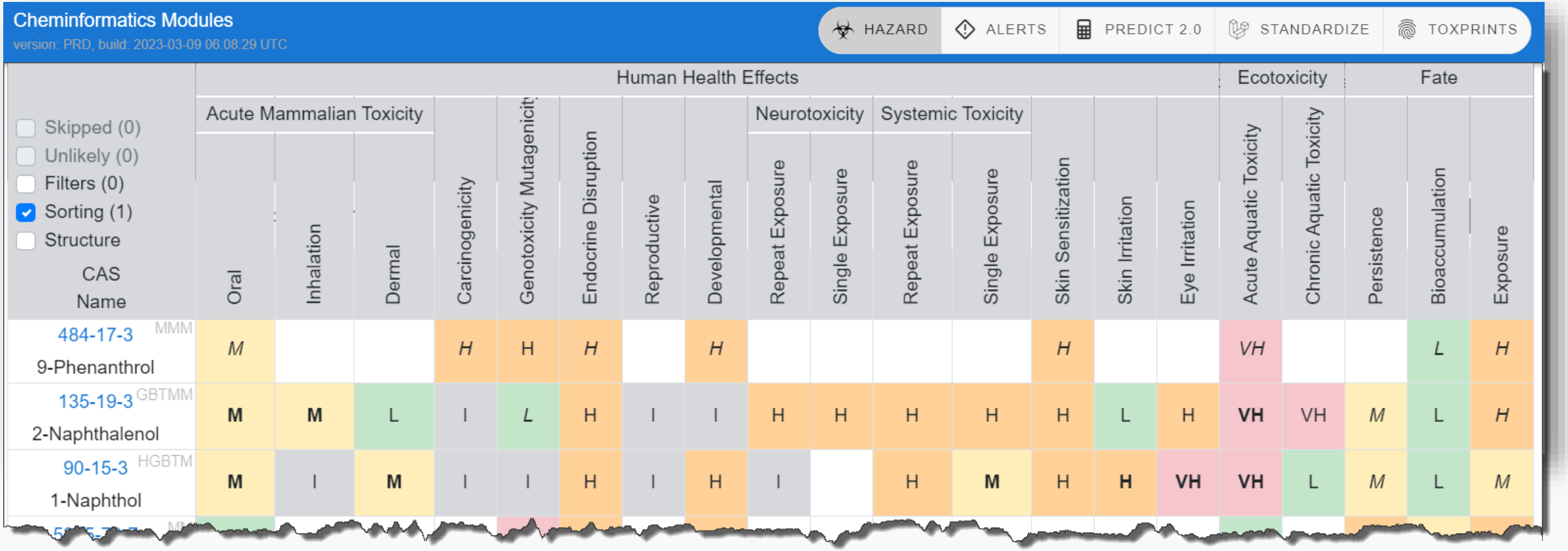
Journal of Cheminformatics **10**, Article number: 10 (2018) | Cite this article

17k Accesses | **221** Citations | **25** Altmetric | Metrics

…and what if we could then profile toxicity?



…and what if we could then profile toxicity?



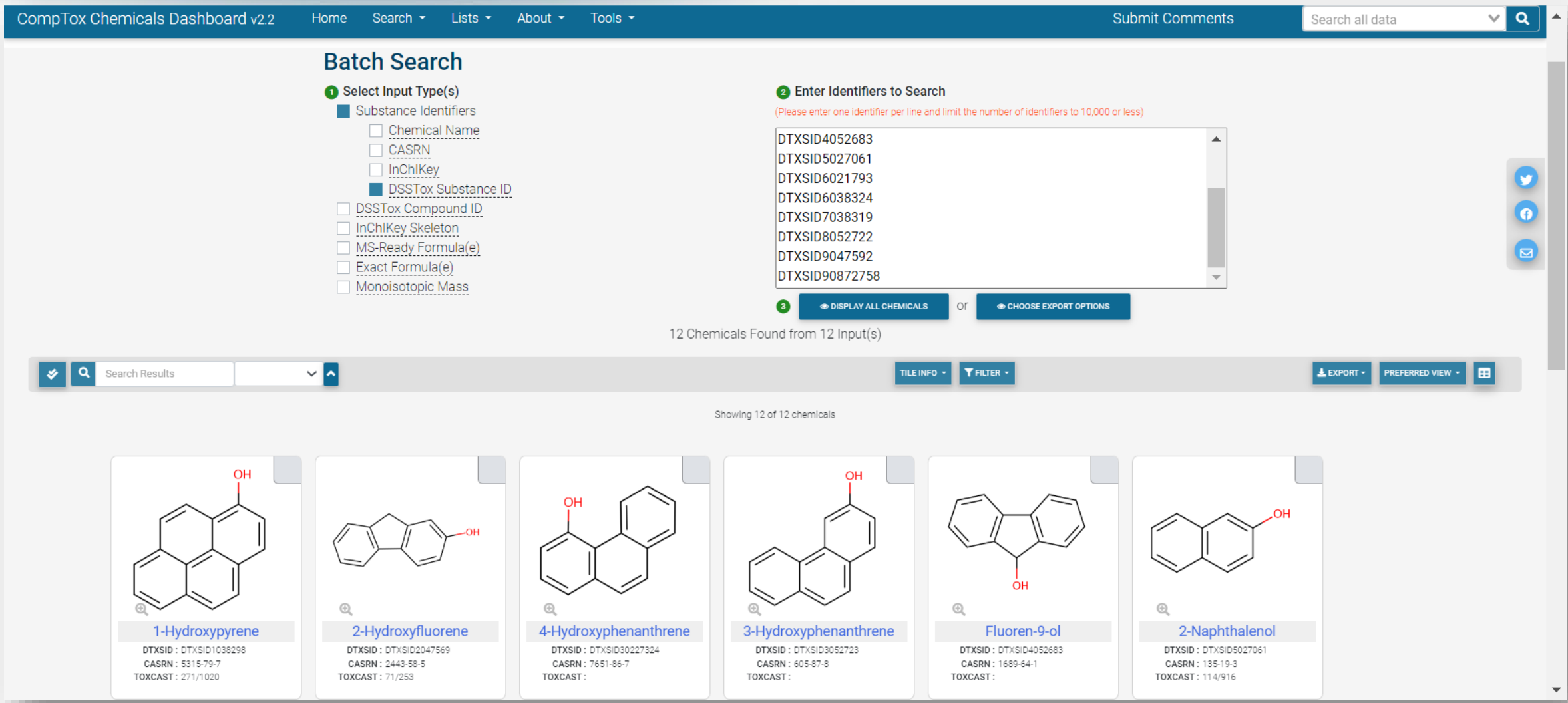
Cheminformatics Modules
version: PRD, build: 2023-03-09 06:08:29 UTC

HAZARD | ALERTS | PREDICT 2.0 | STANDARDIZE | TOXPRINTS

Skipped (0)
Unlikely (0)
Filters (0)
Sorting (1)
Structure

CAS / Name	Human Health Effects: Acute Mammalian Toxicity: Oral	Acute Mammalian Toxicity: Inhalation	Acute Mammalian Toxicity: Dermal	Carcinogenicity	Genotoxicity Mutagenicity	Endocrine Disruption	Reproductive	Developmental	Neurotoxicity: Repeat Exposure	Neurotoxicity: Single Exposure	Systemic Toxicity: Repeat Exposure	Systemic Toxicity: Single Exposure	Skin Sensitization	Skin Irritation	Eye Irritation	Ecotoxicity: Acute Aquatic Toxicity	Ecotoxicity: Chronic Aquatic Toxicity	Fate: Persistence	Fate: Bioaccumulation	Fate: Exposure
484-17-3 MMM 9-Phenanthrol	*M*			*H*	H	*H*		*H*					*H*			*VH*			*L*	*H*
135-19-3 GBTMM 2-Naphthalenol	**M**	**M**	L	I	*L*	H	I	I	H	H	H	H	H	L	H	**VH**	VH	*M*	L	*H*
90-15-3 HGBTM 1-Naphthol	**M**	I	**M**	I	I	H	I	H	I		H	**M**	H	**H**	**VH**	**VH**	L	*M*	L	*M*

…or simply harvest data from the CompTox Chemicals Dashboard



What data would you like???

Chemical Identifiers

- [x] DTXSID
- [x] Chemical Name
- [ ] DTXCID
- [ ] CAS-RN
- [ ] InChIKey
- [ ] IUPAC Name

Structures

- [ ] Mol File
- [ ] SMILES
- [ ] InChI String
- [ ] MS-Ready SMILES
- [ ] QSAR-Ready SMILES

Intrinsic and Predicted Properties

- [ ] Molecular Formula
- [ ] Average Mass
- [ ] Monoisotopic Mass
- [ ] TEST Model Predictions
- [ ] OPERA Model Predictions
- [ ] ToxPrint fingerprints (ChemoTyper)
- [ ] ToxPrint single fingerprints

Metadata

- [ ] Curation Level Details
- [ ] Safety Data
- [ ] NHANES/Predicted Exposure
- [ ] Data Sources
- [ ] Include ToxVal Data Availability
- [ ] Assay Hit Count
- [ ] Number of PubMed Articles
- [ ] PubChem Data Sources
- [ ] CPDat Product Occurrence Count
- [ ] IRIS
- [ ] PPRTV
- [ ] Wikipedia Article
- [ ] QC Notes
- [ ] Include links to ACToR reports

Enhanced Data Sheets

- [ ] MetFrag Input File (Beta)
- [ ] Abstract Sifter Input File
- [ ] Synonyms and Identifiers
- [ ] Related Substance relationships
- [ ] Associated ToxCast Assays
- [ ] ToxValDB Details
- [ ] Physicochemical Property Values

Presence in Lists

Title	Description
40CFR1164	40 CFR 116.4 Designation of Hazardous Substances (Above Ground Storage Tanks)
40CFR355	40CFR355 Extremely Hazardous Substance List and Threshold Planning Quantities
ACSREAG	LIST: ACS Reagent Chemicals
AEGLVALUES	AEGLS: Acute Exposure Guideline Levels
AGCHEMWEAPONS	WEAPONS: Australia Group
ALGALTOX	LIST: Algal Toxins
ALLSURFACTANTS	CATEGORY: Surfactants
AMINOACIDS	CATEGORY: Amino acids
AMPHIBOLES	Amphibole minerals
ANITMICROB2	Antimicrobial Ingredients in Building Materials
ANTIBIOTICS	CATEGORY|PHARMACEUTICALS: Antibiotics
ANTIMICROBIALS	CATEGORY|WIKILIST|ANTIMICROBIALS: Antimicrobials from Wikipedia
AOPSTRESSORS	List of Adverse Outcome Pathway Stressors

Rows: **424**

Introducing AMOS
Analytical Methods and Spectra Database

- Three types of data in the database:
  - Methods (regulatory, lab manuals and SOPs, publications, tech notes)
  - Spectra (from public domain and our own laboratories)
  - Monographs (harvested from SWGDRUG and other sites)
- Some methods have associated spectra
- Some data are just ***externally*** linked
- Currently contains around 195,000 spectra, 700,000 external links, 800 “Fact Sheets” and ~3500 methods
- ALL data are growing in number

Where are there methods?

- Agency-based methods
  - EPA, USGS, USDA, CDC, FDA, OSHA, DEA, …
- ASTM and ISO methods
- Vendor application notes – Thermo, Waters, Agilent, Sciex, Shimadzu, LECO, ….
- Peer-reviewed articles
- Laboratory Documents – lab manuals, SOPs

A view of the methods list

Method #	Name	Year	Methodology	Source	Analyte	Chemical Class	Matrix	# Compounds
10.1155/2015/592763	Development and V...	2015	GC/ECD	Hindawi	Metamitron		Soil	1
JAMC-8838219	Development and A...	2020	LC/MS	Hindawi	Antituberculosis Dr...		Serum (human)	4
PMC-AL-2074-2088	Quantitation of Phe...	2021	LC/MS	HHS	Phenolic benzotriaz...		Plasma (rat)	9
10.1016/j.foodcont.2023.109772	Determination of 63...	2023	HPLC/MS	Food Control	Mycotoxins		Grain products	62
10.1016/j.foodchem.2016.05.151	QuEChERS-based p...	2016	HPLC/MS	Food Chemi...	Quaternary ammoni...	Surfactant	Dairy products	7
C-010.02	Determination of 16...	2021	LC/MS	FDA	PFAS		Processed food	16
C-011.01	Determination of Ch...	2019	LC/MS	FDA	Chloramphenicol an...		Cobia, croaker, and shrimp	2
T-PFAS-WI14355	Polyfluorinated alkyl...	2018	LC/MS	Eurofins	PFAS		Aqueous samples	22
10.1186/s12302-021-00556-1	Method developme...	2021	HPLC/MS	Environment...	Microcystins	Cyanotoxin	Water (surface)	8
10.1039/C9EM00554D	Comprehensive scre...	2020	LC/MS	Environment...	Quaternary ammoni...	Surfactant	Sediment, water (waste)	28
Method 545	Method 545: Deter...	2015	LC/MS	EPA-OW	Anatoxin-a, Cylindro...	Cyanotoxin	Water (finished)	2
GRM-91.10	2018-06-44045101-...	1995	Immunoassay	EPA-ECM	Spinosad		Water	15
GRM-94.21	2018-06-44045102-...	1995	Immunoassay	EPA-ECM	Spinosad		Sediment	15
GRM-94.20	2018-06-44045103-...	1995	HPLC/UV	EPA-ECM	Spinosad and its de...		Soil/sediment	5
GRM-94.12	2018-06-44045105-...	1995	HPLC/UV	EPA-ECM	Spinosad and its de...		Water	5
D9513	2018-06-44084504-...	1995	LC/MS	EPA-ECM	Quinclorac and its d...		Soil (sand, sandy loam)	3
D9513	2018-06-44086601-...	1996	LC/MS	EPA-ECM	Quinclorac and its d...		Soil (sand, sandy loam)	3
RAM-278/01	2018-06-44104807-...	1996	LC/MS	EPA-ECM	Tralkoxydim		Soil	3

Filtering the list for interests…

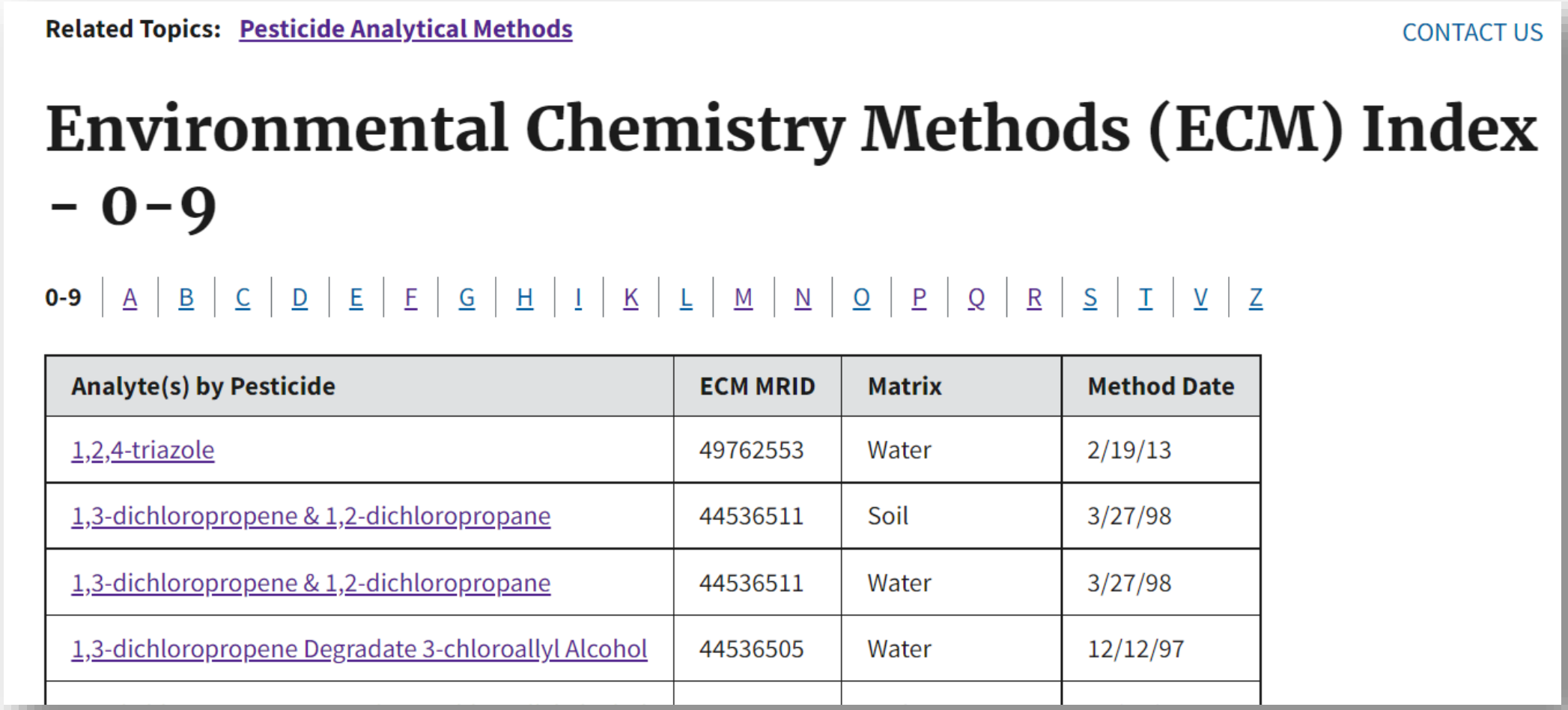
- Look for pesticides studied in water, by GC/MS, after 1990

Method #	Name	Year	Method...	Source	Analyte	Chemical Class	Matrix	# Compounds
		1990				pesticides	water	
EPA-1656a	Method 1656, Revision A: ...	2020			Organohalide pesticides	Pesticides	Wastewater, soil, sludge, sedimen...	31
EPA-1657a	Method 1657, Revision A: ...	2020			Organophosphorus Pestici...	Pesticides	Wastewater, soil, sludge, sedimen...	47
EPA-608.1	Organochlorine Pesticides i...	1993			Organochlorine pesticides	Pesticides	Water (municipal waste, industrial...	7
EPA-608.2	Certain Organochlorine Pes...	1993			Organochlorine pesticides	Pesticides	Water (municipal waste, industrial...	6
EPA-614	Organophosphorus Pestici...	1992			Organophosphorus pestici...	Pesticides	Water (municipal waste, industrial...	8
EPA-614.1	Organophosphorus Pestici...	1992	GC	USEPA	Organophosphorus pestici...	Pesticides	Water (municipal waste, industrial...	4
EPA-617	Organohalide Pesticides an...	1992	GC	USEPA	Organohalide pesticides an...	Pesticides	Water (municipal waste, industrial...	25
EPA-619	Triazine Pesticides in Munic...	1992	GC/MS	USEPA	Triazine pesticides	Pesticides	Water (municipal waste, industrial...	11
EPA-622	Organophosphorus Pestici...	1992	GC	USEPA	Organophosphorus pestici...	Pesticides	Water (municipal waste, industrial...	20
EPA-622.1	Thiophosphate Pesticides i...	1992	GC/AFD	USEPA	Thiophosphate pesticides	Pesticides	Water (municipal waste, industrial...	7
Method-507	Determination of nitrogen ...	1995	GC/NPD	USEPA	Nitrogen and phosphorus ...	Pesticides	Water (ground, finished)	46
AOAC-990.06	Organochlorine Pesticides i...	1993	GC/ECD	NEMI	Organochlorine pesticides	Pesticides	Water (finished)	28

Greater than
1990
AND OR
Equals
Filter...

Where are there methods?

- 900 method documents from the EPA harvested



Related Topics: Pesticide Analytical Methods

CONTACT US

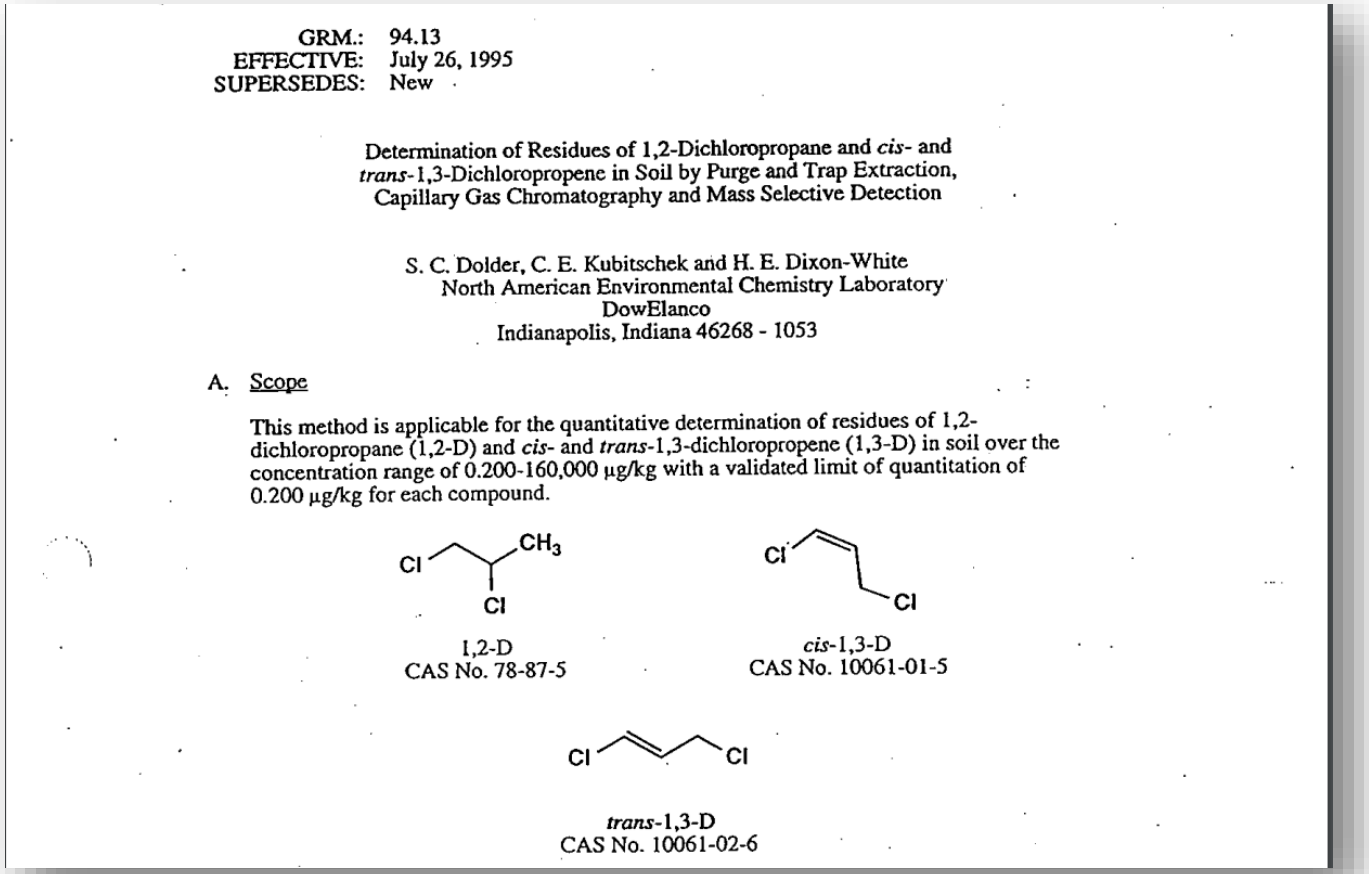
Environmental Chemistry Methods (ECM) Index – 0–9

0-9 | A | B | C | D | E | F | G | H | I | K | L | M | N | O | P | Q | R | S | T | V | Z

Analyte(s) by Pesticide	ECM MRID	Matrix	Method Date
1,2,4-triazole	49762553	Water	2/19/13
1,3-dichloropropene & 1,2-dichloropropane	44536511	Soil	3/27/98
1,3-dichloropropene & 1,2-dichloropropane	44536511	Water	3/27/98
1,3-dichloropropene Degradate 3-chloroallyl Alcohol	44536505	Water	12/12/97

Many Scanned Documents!!!

- Methods generally developed by the agrochemical companies
- Include parents *plus* degradation products
- Lots of scanned, old, documents but the historical records are still of significant use
- Electronic document forms of old documents still of benefit



GRM.: 94.13
EFFECTIVE: July 26, 1995
SUPERSEDES: New

Determination of Residues of 1,2-Dichloropropane and *cis*- and *trans*-1,3-Dichloropropene in Soil by Purge and Trap Extraction, Capillary Gas Chromatography and Mass Selective Detection

S. C. Dolder, C. E. Kubitschek and H. E. Dixon-White
North American Environmental Chemistry Laboratory
DowElanco
Indianapolis, Indiana 46268 - 1053

A. Scope

This method is applicable for the quantitative determination of residues of 1,2-dichloropropane (1,2-D) and *cis*- and *trans*-1,3-dichloropropene (1,3-D) in soil over the concentration range of 0.200-160,000 µg/kg with a validated limit of quantitation of 0.200 µg/kg for each compound.

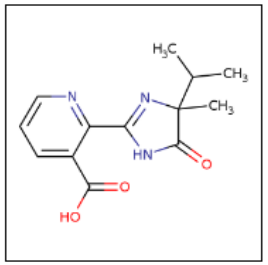
1,2-D
CAS No. 78-87-5

cis-1,3-D
CAS No. 10061-01-5

trans-1,3-D
CAS No. 10061-02-6

Embedding the old Method PDFs

Search Results for "Imazapyr"



(Preferred) Name: Imazapyr
DTXSID: DTXSID8034665
CASRN: 81334-34-1
InChIKey: CLQMBPJKHLGMQK-UHFFFAOYNA-N
Molecular Formula: C13H15N3O3
Mass: 261.281

Spectrum Type	Source	Record Type	View
	Environmental Chemistry Methods	Method	PDF
	Environmental Chemistry Methods	Method	PDF
LC-MS+	MoNA	Spectrum	Spectrum
LC-MS+	MassBank EU	Spectrum	Spectrum
LC-MS+	MassBank EU	Spectrum	Spectrum
LC-MS+	MassBank EU	Spectrum	Spectrum
LC-MS+	MassBank EU	Spectrum	Spectrum
LC-MS+	MassBank EU	Spectrum	Spectrum

Imazapyr

MRID: 41891501
Date: 10/1/89
Matrix: Water
Registrant: American Cyanamide Co
Analysis: HPLC/UV
Limit of Quantitation: 5.0 µg/L

US EPA - ECM for Imazapyr ... 1 / 8 90%

41891501

CYANAMID C3197 CONFIDENTIAL Page 5 of 22

SOP M1900
A. Khunachak/hm
09/05/89
Approved by:
J. M. Devine 09/05/89 Date

AMERICAN CYANAMID COMPANY
AGRICULTURAL RESEARCH DIVISION
CHEMICAL DEVELOPMENT
P. O. Box 400
Princeton, New Jersey 08540 USA

RECOMMENDED METHOD OF ANALYSIS

Imazapyr (CL 243,997): HPLC Method for the Determination of CL 243,997 Residues in Water

A. Principle

Residues of CL 243,997 are extracted from water by using a C18 solid phase extraction (SPE) cartridge. Additional cleanup and specificity are achieved by

Embedding New Method PDFs

EPA United States Environmental Protection Agency

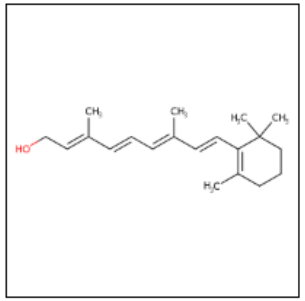
Search EPA.gov

Environmental Topics | Laws & Regulations | About EPA

Search Term [ ] Search

Home | Monograph List | About this App | Toggle Header/Footer

Search Results for "Retinol"



(Preferred) Name: Retinol
DTXSID: DTXSID3023556
CASRN: 68-26-8
InChIKey: FPIPGXGPPPQFEQ-OVSJKPMPSA-N
Molecular Formula: C20H30O
Mass: 286.459

Spectrum Type	Source	Record Type	View
GC-MS	Spectrabase	Spectrum	External Link
GC-MS	Spectrabase	Spectrum	External Link
GC-MS	Spectrabase	Spectrum	External Link
GC-MS	Spectrabase	Spectrum	External Link
GC-MS	Spectrabase	Spectrum	External Link
GC-MS	Spectrabase	Spectrum	External Link
GC-MS	Spectrabase	Spectrum	External Link
GC-MS	Spectrabase	Spectrum	External Link
LC-MS	Agilent	Method	PDF
LC-MS+	MoNA	Spectrum	Spectrum

Determination of Fat-Soluble Vitamins in Foods Using Agilent Chem Elut S Extraction with LC/DAD and LC/MS/MS Triple-Quadrupole

Author: Hui Zhao
Focus/Analyte: Reliable sample preparation and identification/quantitation in various food matrices
Synopsis: This application note describes a method for the determination of fat-soluble vitamins, including vitamin A (retinol), vitamin D3 (cholecalciferol), vitamin D2 (ergocalciferol), and vitamin E (α-tocopherol) in complex food matrices, including infant formula, egg, canned tuna, and mushroom. Samples were saponified as sample pretreatment, extracted using Agilent Chem Elut S (Supported Liquid Extraction (SLE)) 12 mL cartridges, and fat-soluble vitamins were then simultaneously identified and quantified by an Agilent 1290 Infinity II LC coupled to an Agilent diode array detector (DAD) and Agilent 6470 triple quadrupole LC/MS in series. Data were analyzed using Agilent MassHunter workstation software.

PDF Viewer | Compounds (grid) | Compounds (table)

Determination of Fat-Solubl... 1 / 13 100%

Application Note
Food Testing & Agriculture

Agilent
Trusted Answers

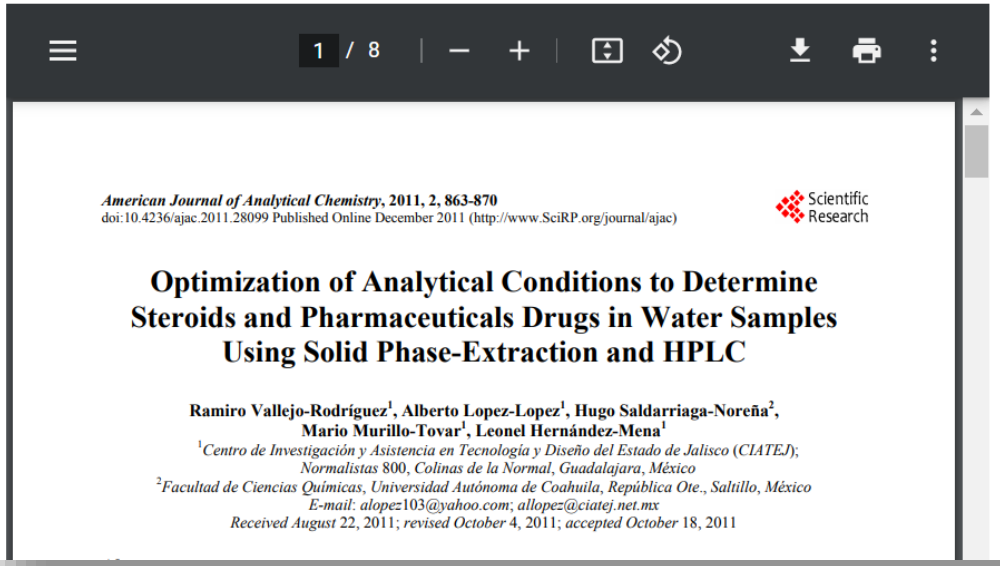
When Methods are OPEN Access

Optimization of Analytical Conditions to Determine Steroids and Pharmaceuticals Drugs in Water Samples Using Solid Phase-Extraction and HPLC

Author: Ramiro Vallejo-Rodríguez, Alberto Lopez-Lopez, Hugo Saldarriaga-Noreña, Mario Murillo-Tovar, Leonel Hernández-Mena

Focus/Analyte: 17beta-Estradiol, 17-epi-Ethynylestradiol, Naproxen, Ibuprofen

Synopsis: Optimization of Analytical Conditions to Determine Steroids and Pharmaceuticals Drugs in Water Samples Using Solid Phase-Extraction and HPLC



American Journal of Analytical Chemistry, 2011, 2, 863-870
doi:10.4236/ajac.2011.28099 Published Online December 2011 (http://www.SciRP.org/journal/ajac)

Scientific Research

Optimization of Analytical Conditions to Determine Steroids and Pharmaceuticals Drugs in Water Samples Using Solid Phase-Extraction and HPLC

Ramiro Vallejo-Rodríguez[1], Alberto Lopez-Lopez[1], Hugo Saldarriaga-Noreña[2], Mario Murillo-Tovar[1], Leonel Hernández-Mena[1]

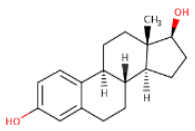
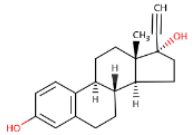
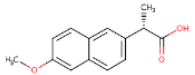
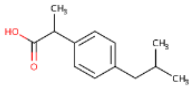
[1]*Centro de Investigación y Asistencia en Tecnología y Diseño del Estado de Jalisco (CIATEJ); Normalistas 800, Colinas de la Normal, Guadalajara, México*

[2]*Facultad de Ciencias Químicas, Universidad Autónoma de Coahuila, República Ote., Saltillo, México*

E-mail: alopez103@yahoo.com; allopez@ciatej.net.mx

Received August 22, 2011; revised October 4, 2011; accepted October 18, 2011

Compounds (4) (grid) | Compounds (4) (table)

17beta-Estradiol  DTXSID0020573 ↗	17-epi-Ethynylestradiol  DTXSID00873493 ↗	Naproxen  DTXSID4040686 ↗
Ibuprofen  DTXSID5020732 ↗		

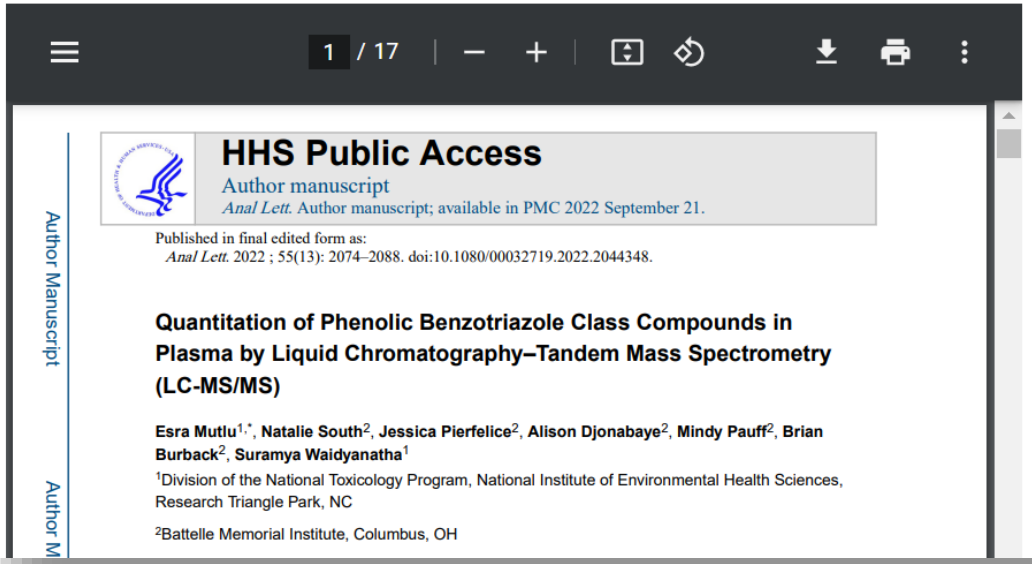
When Methods are PubMed OPEN Access

Quantitation of Phenolic Benzotriazole Class Compounds in Plasma by Liquid Chromatography–Tandem Mass Spectrometry

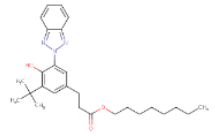
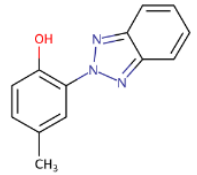
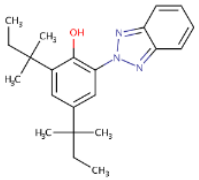
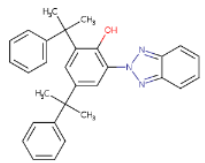
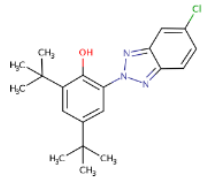
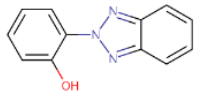
Author: Ersa Mutlu, Natalie South, Jessica Pierfelice, Alison Djonabaye, Mindy Pauff, Brian Burback, Suramya Waidyanatha

Focus/Analyte: Phenolic benzotriazole compounds

Synopsis: Determination of phenolic benzotriazole compounds in plasma (rat) by LC/MS with a LOQ of 5.0-10.0ppb.



Compounds (9) (grid) | Compounds (9) (table)

Octyl 3-[3-(2H-benzotriaz...  DTXSID001016713 ↗	2-(2H-Benzotriazol-2-yl)-4...  DTXSID1027479 ↗	2-(2-Hydroxy-3,5-di-tert-p...  DTXSID2027886 ↗
2-(2H-Benzotriazol-2-yl)-4,...  DTXSID2028985 ↗	Phenol, 2-(5-chloro-2H-be...  DTXSID4038893 ↗	2-(2H-Benzotriazole-2-yl)p...  DTXSID60877078 ↗
Phenol, 2-(2H-benzotriazol...	Octrizole	Bumetrizole

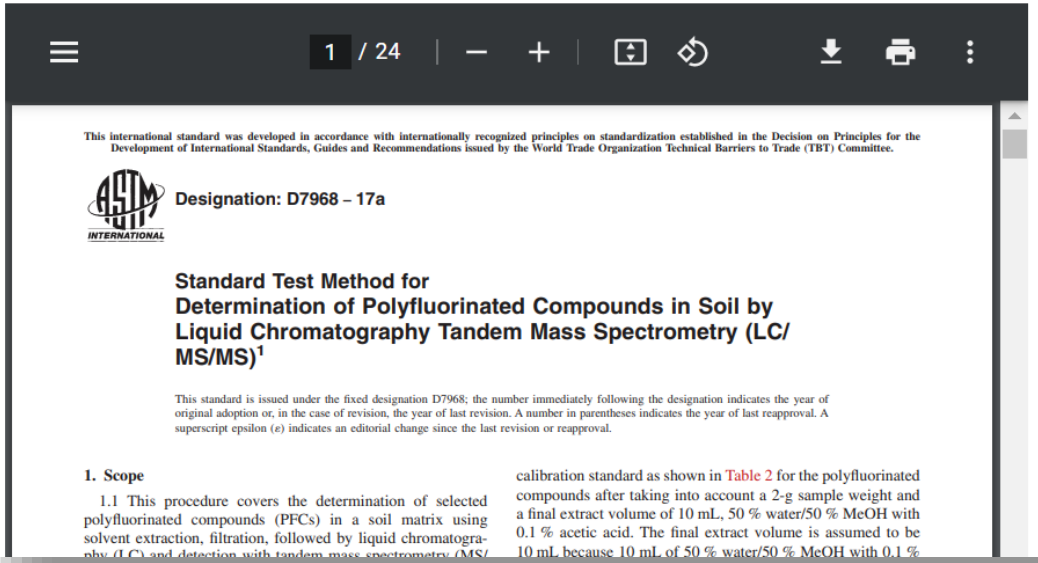
Proprietary Methods for INTERNAL Access

Standard Test Method for Determination of Polyfluorinated Compounds in Soil by Liquid Chromatography Tandem Mass Spectrometry (LC/MS/MS)

Author: ASTM International

Focus/Analyte: PFAS

Synopsis: Determination of PFAS in soil (clay, sand, silt) by LC/MS.



1 / 24

This international standard was developed in accordance with internationally recognized principles on standardization established in the Decision on Principles for the Development of International Standards, Guides and Recommendations issued by the World Trade Organization Technical Barriers to Trade (TBT) Committee.

Designation: D7968 – 17a

Standard Test Method for Determination of Polyfluorinated Compounds in Soil by Liquid Chromatography Tandem Mass Spectrometry (LC/MS/MS)[1]

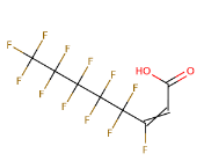
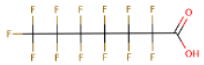
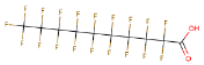
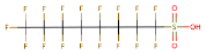
This standard is issued under the fixed designation D7968; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ε) indicates an editorial change since the last revision or reapproval.

1. Scope

1.1 This procedure covers the determination of selected polyfluorinated compounds (PFCs) in a soil matrix using solvent extraction, filtration, followed by liquid chromatography (LC) and detection with tandem mass spectrometry (MS/

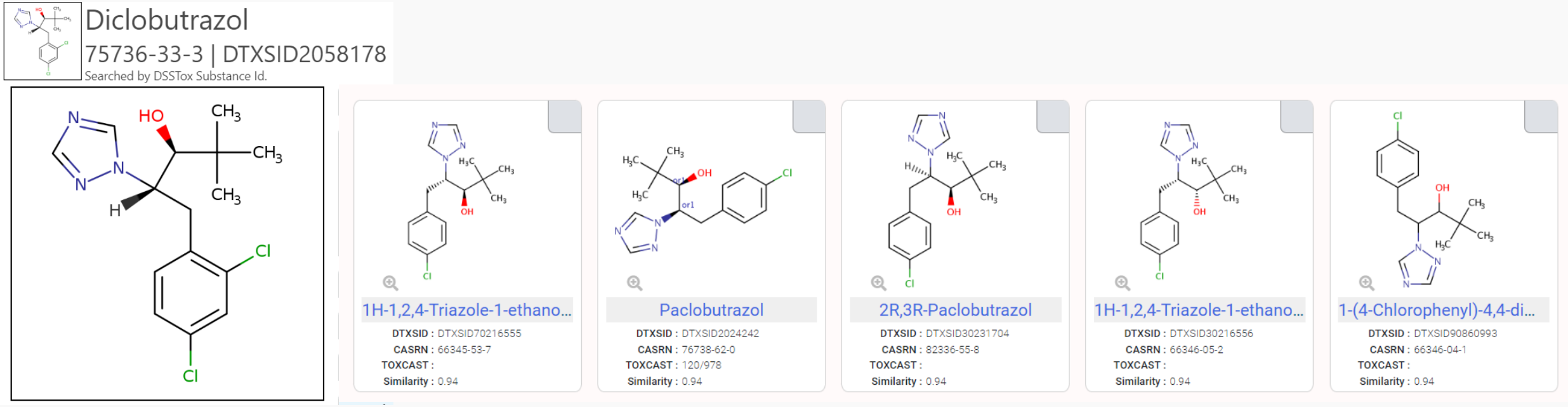
calibration standard as shown in Table 2 for the polyfluorinated compounds after taking into account a 2-g sample weight and a final extract volume of 10 mL, 50 % water/50 % MeOH with 0.1 % acetic acid. The final extract volume is assumed to be 10 mL because 10 mL of 50 % water/50 % MeOH with 0.1 %

Compounds (21) (grid) | Compounds (21) (table)

3-(Perfluoropentyl)-3-fluor...  DTXSID00892326 ↗	Perfluoropentanoate  DTXSID00892487 ↗	Perfluoroheptanoic acid  DTXSID1037303 ↗
Perfluorohexanoate  DTXSID20892484 ↗	Perfluorodecanoic acid  DTXSID3031860 ↗	Perfluorooctanesulfonic acid  DTXSID3031864 ↗
Perfluorotetradecanoic acid	Perfluorobutanoic acid	2-(Perfluorodecyl)ethanoic...

If there is no method for your chemical

- Use “Chemical Similarity Searching” so that you can find chemicals that are similar in structure space
- Use the “Tanimoto Similarity Search”



Searching for a chemical – CASRN, Name
Direct structure searching coming

Compound Identifier: Diclobutrazol | Method Search

The table below lists methods for compounds that are similar to "DTXSID2058178".

Select a row in the table to view the method on the right half of the screen. Bolded rows refer to methods which contain the chemical being searched.

Hover over a method name to see the full text of it. The number in parentheses at the end is the number of similar compounds found in the method (not necessarily the number of compounds present in the method).

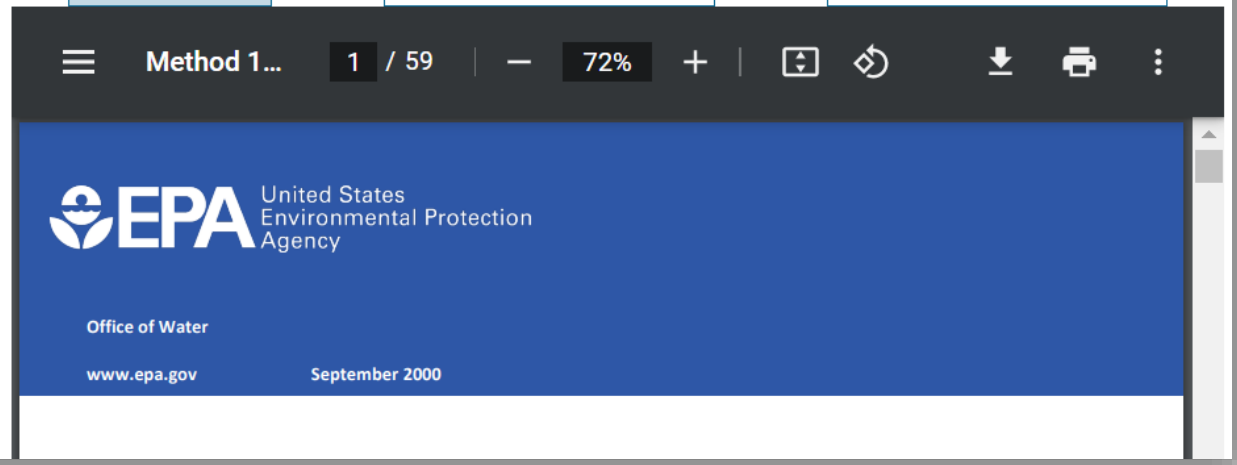
Columns can be hidden by clicking on the menu icon seen when hovering over a column name -- this brings up a menu where column visibility can be toggled.

Method Name (# compounds)	Source	Methodology	Year	Simi...	Similar DTXSID	Compou
Method 1656, Revision A: Organ	USEPA	GC/HSD	2020	0.94		
Methods of analysis-Determinat	USGS	GC/MS	2012	0.85		
Analysis of Endocrine Disrupting				0.85		
Determination of pesticides and	USGS	LC/MS	2015	0.85		
A method for the analysis of 121				0.85		
Analysis of Pesticides in Food Sa				0.85		
2021-09-der-tebuconazole-soil-	EPA-ECM		2018	0.85		
2021-09-der-tebuconazole-wate	EPA-ECM		2018	0.85		
2021-09-ecm-tebuconazole-soil	EPA-ECM		2018	0.85		
2021-09-ecm-tebuconazole-wat	EPA-ECM		2018	0.85		
2021-09-ilv-tebuconazole-soil-n	EPA-ECM		2018	0.85		

Method 1656, Revision A: Organo-Halide Pesticides in Wastewater, Soil, Sludge, Sediment, and Tissue by GC/HSD

Author: EPA-OW
Focus/Analyte: Organohalide pesticides
Synopsis: Determination of organohalide pesticides in wastewater, soil, sludge, sediment, and fish tissue by GC/HSD with a MDL of 300-2000000ppb.

PDF Viewer | Compounds (31) (grid) | Compounds (31) (table)



When Methods are Not Enough

- EPA is highly active in the field of non-targeted analysis
- We have been applying lots of cheminformatics approaches

"MS-Ready" structures for non-targeted high-resolution mass spectrometry screening studies

Andrew D. McEachran, Kamel Mansouri, Chris Grulke, Emma L. Schymanski, Christoph Ruttkies & Antony J. Williams

Journal of Cheminformatics **10**, Article number: 45 (2018) | Cite this article

6215 Accesses | **45** Citations | **14** Altmetric | Metrics

Linking *in silico* MS/MS spectra with chemistry data to improve identification of unknowns

Andrew D. McEachran, Ilya Balabin, Tommy Cathey, Thomas R. Transue, Hussein Al-Ghoul, Chris Grulke, Jon R. Sobus & Antony J. Williams

Scientific Data **6**, Article number: 141 (2019) | Cite this article

5422 Accesses | **23** Citations | **11** Altmetric | Metrics

metabolites

MDPI

Article

Revisiting Five Years of CASMI Contests with EPA Identification Tools

Andrew D. McEachran [1,*], Alex Chao [1], Hussein Al-Ghoul [1], Charles Lowe [2], Christopher Grulke [2], Jon R. Sobus [2] and Antony J. Williams [2,*]

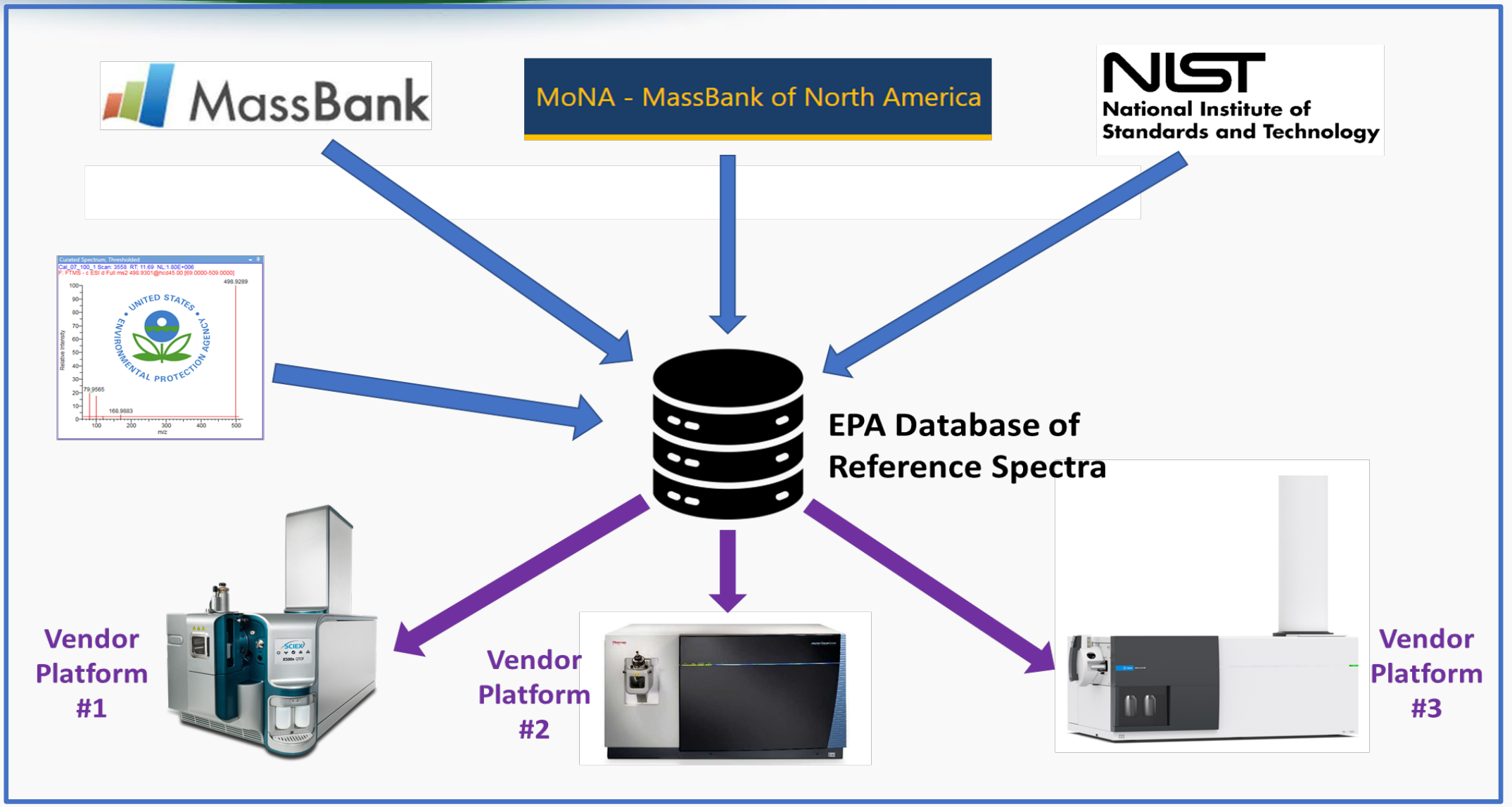
Identifying known unknowns using the US EPA's CompTox Chemistry Dashboard

Andrew D. McEachran, Jon R. Sobus & Antony J. Williams

Analytical and Bioanalytical Chemistry **409**, 1729–1735 (2017) | Cite this article

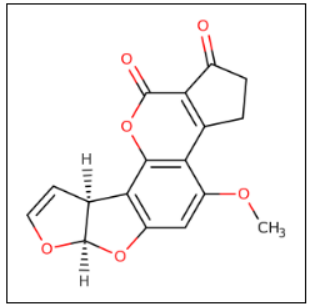
2748 Accesses | **76** Citations | **31** Altmetric | Metrics

Building a spectrum library to search against



Linking to actual spectra

33 Results for "aflatoxin B1"



(Preferred) Name: Aflatoxin B1
DTXSID: DTXSID9020035
CASRN: 1162-65-8
InChIKey: OQIQSTLJSLGHID-WNWIJWBNSA-N
Molecular Formula: C17H12O6
Mass: 312.0634

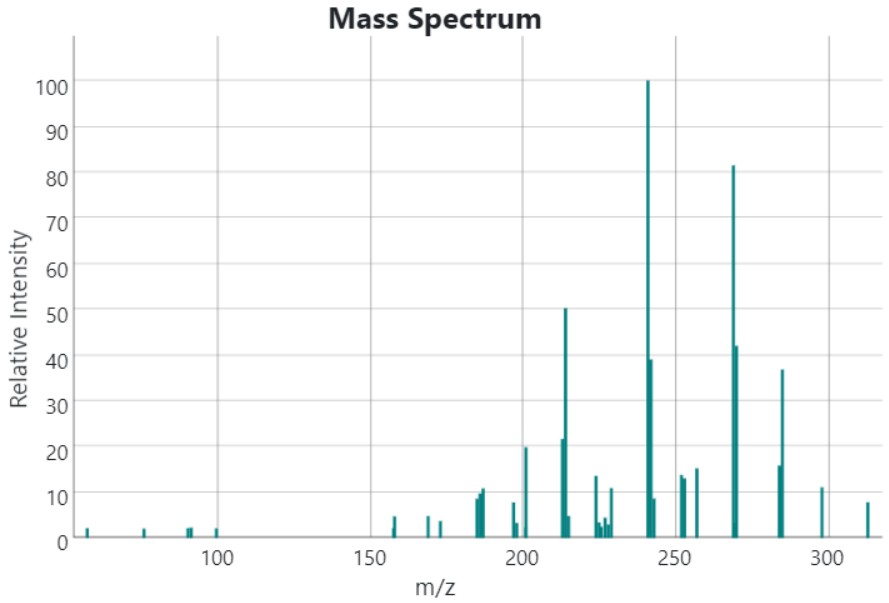
Download Results

☐ Display Single Point Spectra

All Results (33) | Spectra (23) | Monographs (0) | Methods (10)

Methodology	Source	Record Type	Information
GC/MS	MoNA	Spectrum	EI-B; Positive; # PEAKS=29
LC-ESI-ITFT	MassBank EU	Spectrum	LC-ESI-ITFT; MS2; CE: 30; R=17500; [M+H
LC-ESI-ITFT	MassBank EU	Spectrum	LC-ESI-ITFT; MS2; CE: 35; R=17500; [M+H
LC-ESI-ITFT	MassBank EU	Spectrum	LC-ESI-ITFT; MS2; CE: 50; R=17500; [M+H
LC-ESI-ITFT	MassBank EU	Spectrum	LC-ESI-ITFT; MS2; CE: 35; R=17500; [M+N
LC-ESI-ITFT	MassBank EU	Spectrum	LC-ESI-ITFT; MS2; CE: 50; R=17500; [M+N
LC-ESI-QTOF	MassBank EU	Spectrum	LC-ESI-QTOF; MS2; CE: 20 eV; R=35000; I
LC-ESI-QTOF	MassBank EU	Spectrum	LC-ESI-QTOF; MS2; CE: 30 eV; R=35000; I
LC-ESI-QTOF	MassBank EU	Spectrum	LC-ESI-QTOF; MS2; CE: 40 eV; R=35000; I
LC-ESI-QTOF	MassBank EU	Spectrum	LC-ESI-QTOF; MS2; CE: 50 eV; R=35000; I
LC/MS	MoNA	Spectrum	MS2; Positive; # PEAKS=160
LC/MS	MoNA	Spectrum	MS2; Positive; # PEAKS=133

Below is a plot of the spectrum as intensities versus mass-to-charge ratios (m/z). Click and drag over a section of the horizontal axis to zoom; double click to zoom back out. Intensities are scaled so that the highest peak has a value of 100.

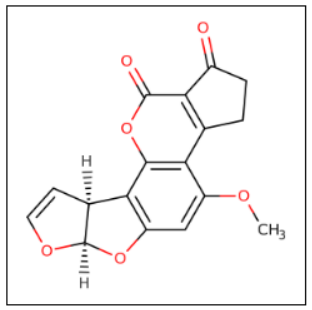


Information

Number of Points: 38
Spectral Entropy: 2.9499

Linking to actual spectra

33 Results for "aflatoxin B1"



(Preferred) Name: Afla
DTXSID: DTXSID902003
CASRN: 1162-65-8
InChIKey: OQIQSTLJSL
Molecular Formula: C1
Mass: 312.0634

Download Results

Display Single Point Spectra

All Results (33) | Spectra (23) | Monograph

Methodology	Source	Record Type	Information
GC/MS	MoNA	Spectrum	EI-B; Positive; # PEAKS
LC-ESI-ITFT	MassBank EU	Spectrum	LC-ESI-ITFT; MS2; CE:
LC-ESI-ITFT	MassBank EU	Spectrum	LC-ESI-ITFT; MS2; CE:
LC-ESI-ITFT	MassBank EU	Spectrum	LC-ESI-ITFT; MS2; CE:
LC-ESI-ITFT	MassBank EU	Spectrum	LC-ESI-ITFT; MS2; CE:
LC-ESI-ITFT	MassBank EU	Spectrum	LC-ESI-ITFT; MS2; CE:
LC-ESI-QTOF	MassBank EU	Spectrum	LC-ESI-QTOF; MS2; CE
LC-ESI-QTOF	MassBank EU	Spectrum	LC-ESI-QTOF; MS2; CE
LC-ESI-QTOF	MassBank EU	Spectrum	LC-ESI-QTOF; MS2; CE
LC-ESI-QTOF	MassBank EU	Spectrum	LC-ESI-QTOF; MS2; CE
LC/MS	MoNA	Spectrum	MS2; Positive; # PEAKS
LC/MS	MoNA	Spectrum	MS2; Positive; # PEAKS=133

Chromatography Info:

COLUMN_NAME: Agilent RRHD Eclipse 50 x 2 mm, 1.8 uM
FLOW_GRADIENT: 100:0 at 0 min, 100:0 at 0.5 min, 0:100 at 3.5 min, 0:100 at 5.5 min, 100:0 at 7 min
FLOW_RATE: 0.3 mL min-1
RETENTION_TIME: 3.1
SOLVENT: ["H2O 0.1% FA", "ACN 0.1% FA"]

Spectrometry Info:

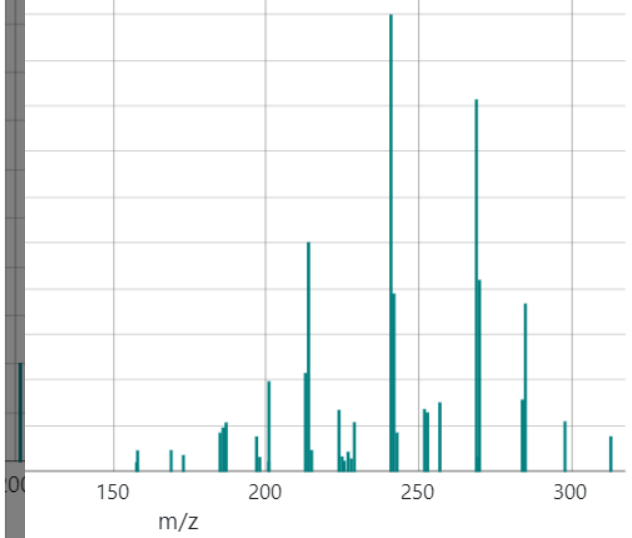
CAPILLARY_VOLTAGE: 3.9 kV
COLLISION_ENERGY: 50(NCE)
FRAGMENTATION_MODE: HCD
IONIZATION: ESI
ION_MODE: POSITIVE
MS_TYPE: MS2
RESOLUTION: 17500

Copy to Clipboard

Cancel | Ok

sities versus mass-to-charge ratios (m/z). Click and drag over a
double click to zoom back out. Intensities are scaled so that the

Mass Spectrum



Information

ints: 38
Spectral Entropy: 2.9499

There are errors EVERYWHERE: **110-75-8**

METHOD 601—PURGEABLE HALOCARBONS

1. **Scope and Application**

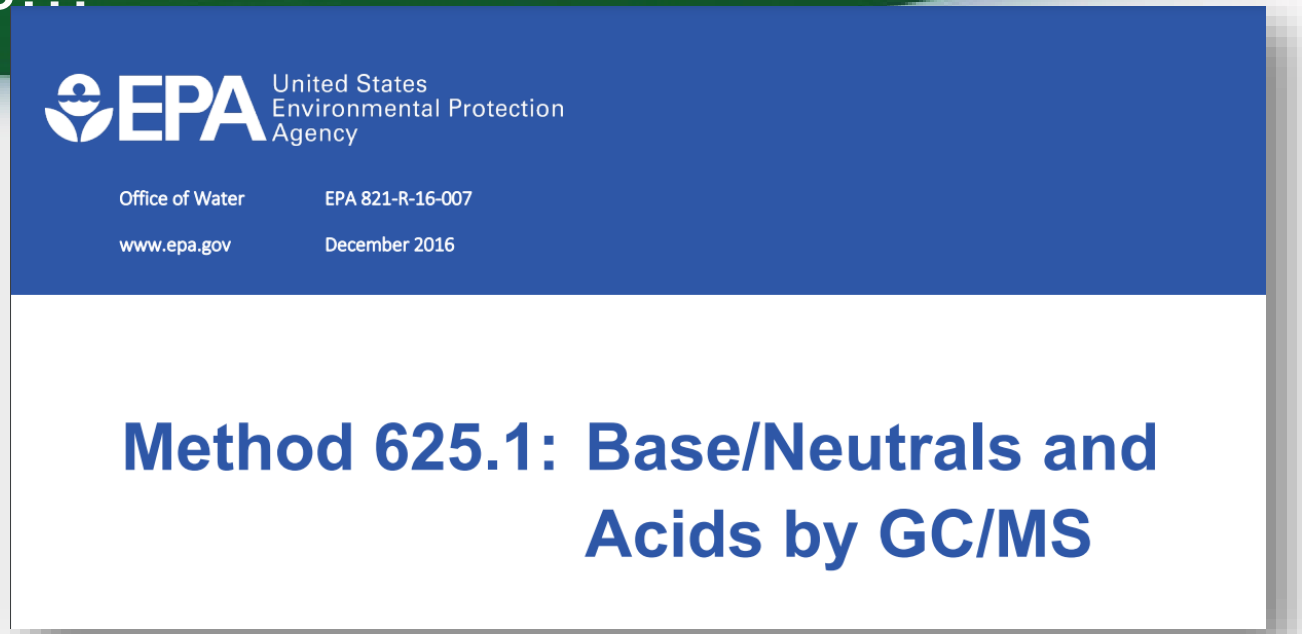
1.1 This method covers the determination of 29 purgeable halocarbons.

The following parameters may be determined by this method:

Parameter	STORET No.	CAS No.
Bromodichloromethane	32101	75-27-4
Bromoform	32104	75-25-2
Bromomethane	34413	74-83-9
Carbon tetrachloride	32102	56-23-5
Chlorobenzene	34301	108-90-7
Chloroethane	34311	75-00-3
2-Chloroethylvinyl ether	34576	100-75-8
Chloroform	32106	67-66-3
Chloromethane	34418	74-87-3
Dibromochloromethane	32105	124-48-1

	INPUT	FOUND_BY	DTXSID	PREFERRED_NAME
2	75-27-4	CAS-RN	DTXSID102	Bromodichloromethane
3	75-25-2	CAS-RN	DTXSID102	Bromoform
4	74-83-9	CAS-RN	DTXSID802	Methyl bromide
5	56-23-5	CAS-RN	DTXSID802	Carbon tetrachloride
6	108-90-7	CAS-RN	DTXSID402	Chlorobenzene
7	75-00-3	CAS-RN	DTXSID102	Chloroethane
8	100-75-8	Checksum F	-	-
9	67-66-3	CAS-RN	DTXSID102	Chloroform
10	74-87-3	CAS-RN	DTXSID002	Chloromethane
11	124-48-1	CAS-RN	DTXSID102	Chlorodibromomethane
12	95-50-1	CAS-RN	DTXSID602	1,2-Dichlorobenzene
13	541-73-1	CAS-RN	DTXSID602	1,3-Dichlorobenzene

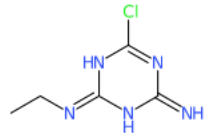
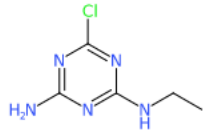
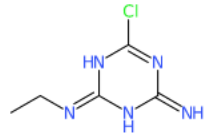
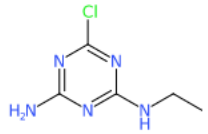
It can be challenging 9/365 chemicals…



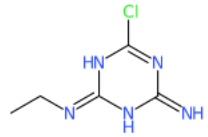
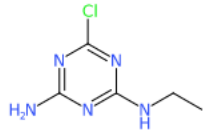
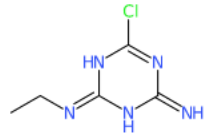
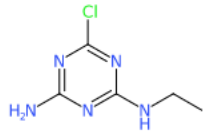
Butylate	2008-41-5
n-C10 (*n*-decane)	124-18-5
n-C12 (*n*-undecane)	112-40-2
n-C14 (*n*-tetradecane)	629-59-4
n-C16 (*n*-hexadecane)	544-76-3
n-C18 (*n*-octadecane)	593-45-3
n-C20 (*n*-eicosane)	112-95-8
n-C22 (*n*-docosane)	629-97-0
n-C24 (*n*-tetracosane)	646-31-1
n-C26 (*n*-hexacosane)	630-01-3
n-C28 (*n*-octacosane)	630-02-4
n-C30 (*n*-triacontane)	638-68-6

	CASRN	INPUT	FOUND_BY	DTXSID	PREFERRED_NAME
100-75-5	100-75-5	100-75-5	CAS number fails checksum.		
108-39-2	108-39-2	108-39-2	CAS number fails checksum.		
11098-82-	11098-82-	11098-82-5	CAS number fails checksum.		
112-40-2	112-40-2	112-40-2	CAS number fails checksum.		
2310-18-0	2310-18-0	2310-18-0	CAS number fails checksum.		
291-29-4	291-29-4	291-29-4	CAS number fails checksum.		
5218-45-2	5218-45-2	5218-45-2	CAS number fails checksum.		
58-89-8	58-89-8	58-89-8	CAS number fails checksum.		
65-50-0	65-50-0	65-50-0	CAS number fails checksum.		
6/1/2425	2425-06-1	2425-06-1	CASRN	DTXSID4020242	Captafol
6/5/2497	2497-06-5	2497-06-5	CASRN	DTXSID8041901	Disulfoton sulfone
7/6/2497	2497-07-6	2497-07-6	CASRN	DTXSID4037536	Oxydisulfoton
100-01-6	100-01-6	100-01-6	CASRN	DTXSID8020961	4-Nitrobenzenamine
100-02-7	100-02-7	100-02-7	CASRN	DTXSID0021834	4-Nitrophenol
100-25-4	100-25-4	100-25-4	CASRN	DTXSID0021836	1,4-Dinitrobenzene

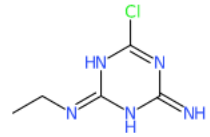
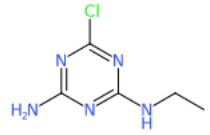
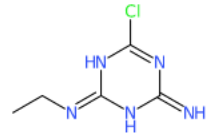
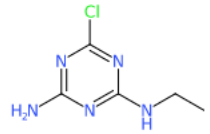
There are MANY errors in public spectral databases

⊞ **Atrazine-Desisopropyl**	4 spectra	**C5H8ClN5**		**173.04683**
⊞ **Atrazine-desisopropyl**	16 spectra	**C5H8ClN5**		**173.04680**
⊞ **Atrazine-Desisopropyl**	4 spectra	**C5H8ClN5**		**173.04683**
⊞ **Atrazine-desisopropyl**	16 spectra	**C5H8ClN5**		**173.04680**

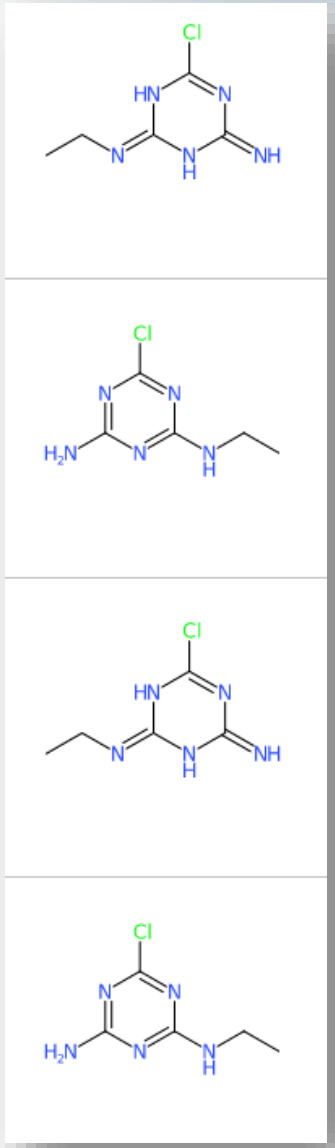
There are MANY errors in public spectral databases

⊞ **Atrazine-Desisopropyl**	4 spectra	**C5H8ClN5**		**173.04683**
⊞ **Atrazine-desisopropyl**	16 spectra	**C5H8ClN5**		**173.04680**
⊞ **Atrazine-Desisopropyl**	4 spectra	**C5H8ClN5**		**173.04683**
⊞ **Atrazine-desisopropyl**	16 spectra	**C5H8ClN5**		**173.04680**

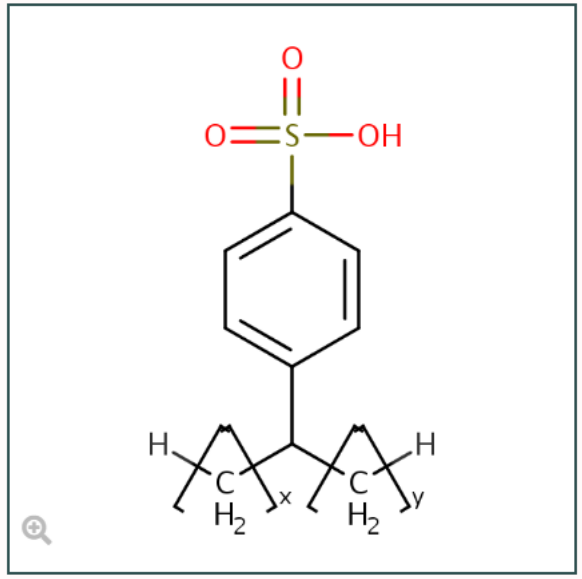
There are MANY errors in public spectral databases

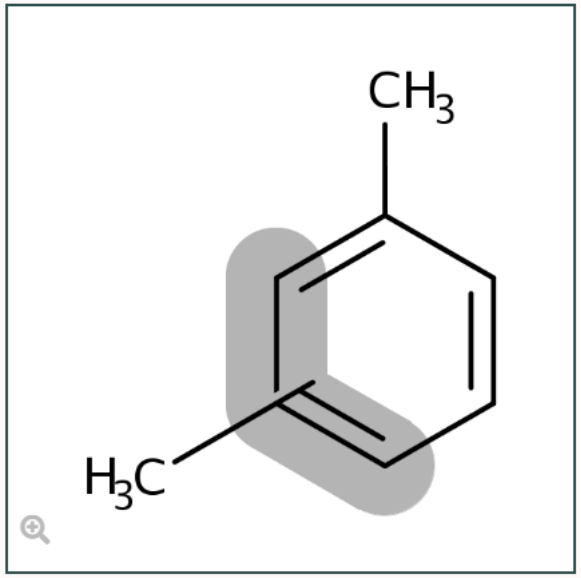
⊞ **Atrazine-Desisopropyl**	4 spectra	**C5H8ClN5**		**173.04683**
⊞ **Atrazine-desisopropyl**	16 spectra	**C5H8ClN5**		**173.04680**
⊞ **Atrazine-Desisopropyl**	4 spectra	**C5H8ClN5**		**173.04683**
⊞ **Atrazine-desisopropyl**	16 spectra	**C5H8ClN5**		**173.04680**

There are MANY errors in public spectral databases



- Chemical structure representations would ideally be standardized…consider tautomeric forms
- Not all substances are explicit and can be ambiguous representations





Conclusions

- We have built (I think) the first chemical structure indexed **open database** of “methods”
- Methods are not just “approved methods” but also standard operating procedures, application notes, lab manuals, regulatory methods, etc.
- Integrating methods to experimental spectral data will serve our non-targeted analysis efforts
- Work is underway to make it public

If you want to help…

- Send information regarding analytical methods and method articles to williams.antony@epa.gov
- If you want to add spectral data please contact me