

AMOS: the EPA database of analytical methods and open mass spectral database supporting non-targeted analysis

Gregory Janesch¹, Erik Carr¹, Vicente Samano², James McCord³, Jacqueline Bangma³, Jon Sobus⁴ and **Antony Williams**⁴

1. ORAU Student Services Contractor 2. Senior Environmental Employment Program 3. Center for Environmental Measurement and Modeling and 4. Center for Computational Toxicology & Exposure,

ALL at the U.S. Environmental Protection Agency

Background



- A huge number of openly available sources exist for spectra, documentation of analytical procedures, etc.
- Search engines can easily find high-traffic sources, but maybe not niche-but-high-quality ones. Most are not "structurally-enabled"
- Useful to have complementary types of data alongside each other, especially with consistent substance identifiers
- Non-targeted analysis can benefit from a broad, high-quality experimental database as a reference

About AMOS - General



 AMOS is a cheminformatics application integrating spectra and analytical methods

 Developed in ~16 months as a "proof-of-concept" application and not yet available publicly

 Provides mappings between substance(s), method documents and experimental spectra

About AMOS - Data



Three broad categories of records:

```
- Spectra (>195,000)
```

– Methods (>3500)

Fact Sheets (>800)

- Most data are open access, some are just external links
- Data is being continually updated (new datasets & updates)
- Many chemicals of interest to EPA PFAS, pesticides, etc.

Underpinning Chemistry: DSSTox Database





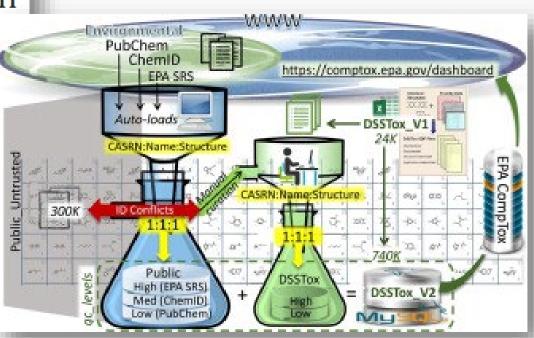
Computational Toxicology

Volume 12, November 2019, 100096



EPA's DSSTox database: History of development of a curated chemistry resource supporting computational toxicology research

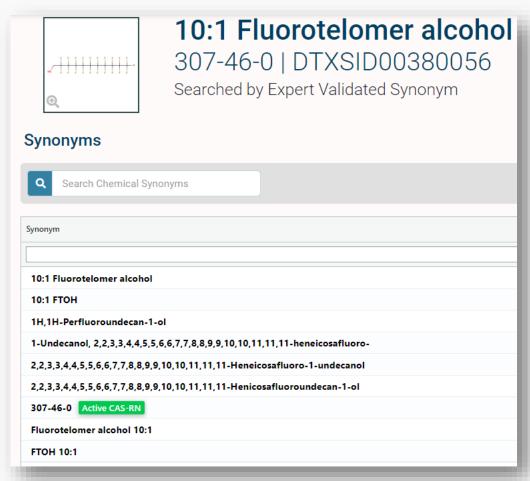
<u>Christopher M. Grulke</u>^a, <u>Antony J. Williams</u>^a, <u>Inthirany Thillanadarajah</u>^b, <u>Ann M. Richard</u>^a ♀ ⊠



About AMOS - Curation



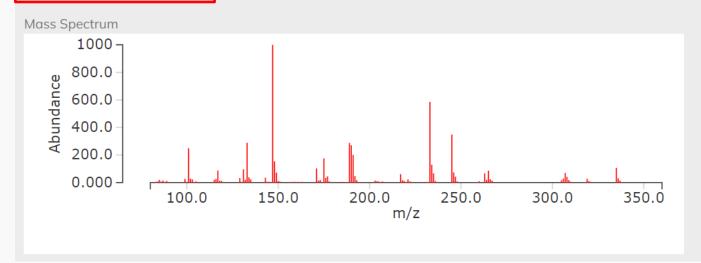
- Identifiers vary between different sources so we must curate
- A single chemical can have dozens of names
 - FTOH 10:1
 - 10:1 FTOH
 - 10:1 Fluorotelomer alcohol
 - 2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,10,10,11,11,11-Henicosafluoroundecan-1-ol
 - 1-Undecanol, 2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,10,10,11,11,11-heneicosafluoro-
- ...and it rarely stops here...

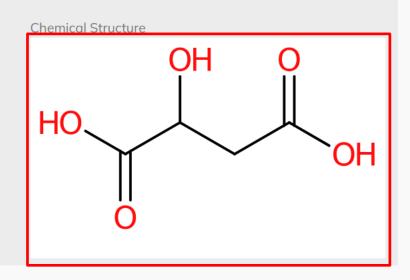


Curation of Structure-Spectra is challenging



L-(-)-Malic acid; GC-EI-TOF; MS; 3 TMS; BP:147





CH\$NAME: L-(-)-Malic acid

CH\$COMPOUND_CLASS: Natural Product

CH\$FORMULA: C₄H₆O₅

CH\$EXACT_MASS: 134.02152

CH\$SMILES: OC(=0)CC(0)C(0)= \bigcirc

CH\$IUPAC: InChI=1S/C4H6O5/c5-2(4(8)9)1-3(6)7/h2,5H,1H2,(H,6,7)(H,8,9)/t2-/m0/s1

CH\$LINK: CAS <u>636-61-3</u>

CH\$LINK: KEGG C00497

CH\$LINK: INCHIKEY BJEPYKJPYRNKOW-REOHCLBHSA-N

CH\$LINK: COMPTOX DTXSID30273987

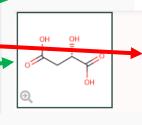
InChikey: BJEPYKJPYRNKOW-REOHCLBHSA-N

SMILES: O[C@@H](CC(O)=O)C(O)=O



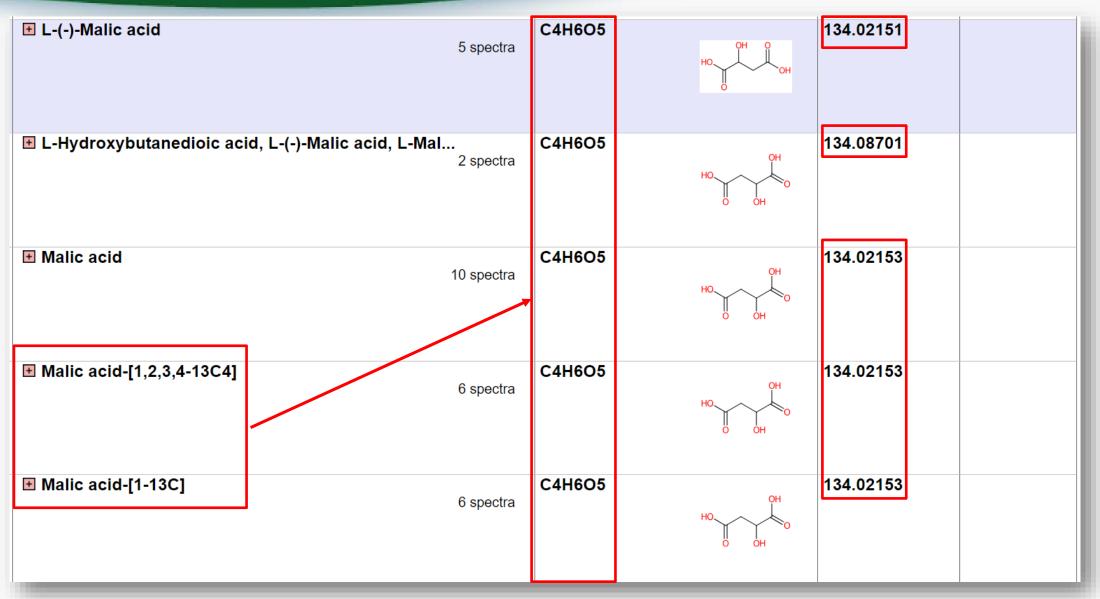
97-67-6 | DTXSID30273987

Searched by DTXSID30273987.



Curation of Structure-Spectra is challenging Multiple Malic Acids



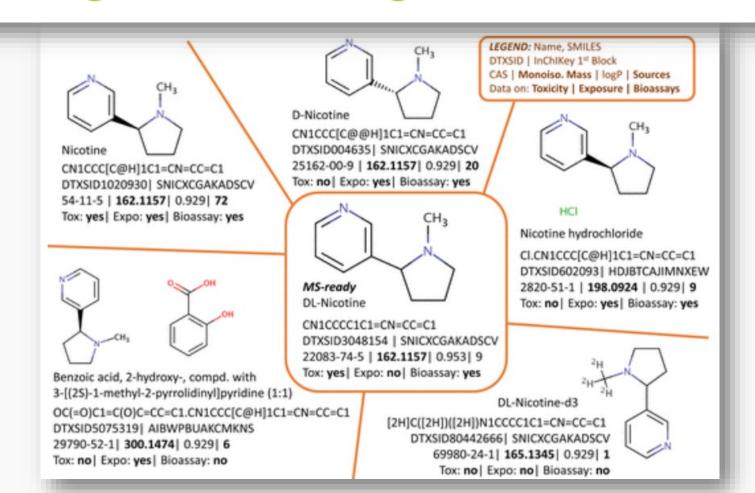


We understand spectra collapse by structure



Open Science for Identifying "Known Unknown" Chemicals

Emma L. Schymanski*† and Antony J. Williams*‡



Spectra



 About 195,000 experimental spectra covering almost 25,000 substances (not including externally-linked ones)

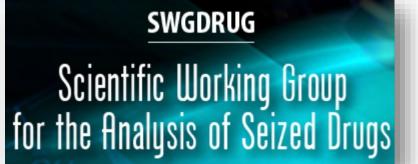
- Most are from external, open-access sources
 - About 95% between MassBank EU, MoNA, & HMDB
- EPA labs now providing spectra (especially PFAS)

Includes metadata like instrument settings (when possible)

Sources of Spectra

United States
Environmental Protection
Agency

- ► Agency for Toxic Substances and Disease Registry
- ► Agilent
- ► Analytical QC
- ► Cayman Spectral Library
- ► Center for Forensic Science Research & Education
- ► Centers for Disease Control and Prevention
- ▶ Human Metabolome Database
- ▶ Jozef Stefan Institute
- ► MassBank Europe
- MassBank of North America
- ► National Environmental Methods Index
- ► National Institute for Occupational Safety and Health
- ► National Institute of Standards and Technology
- ► National Pesticide Information Center
- ► Scientific Working Group for the Analysis of Seized Drugs
- ► SCIEX







Methods

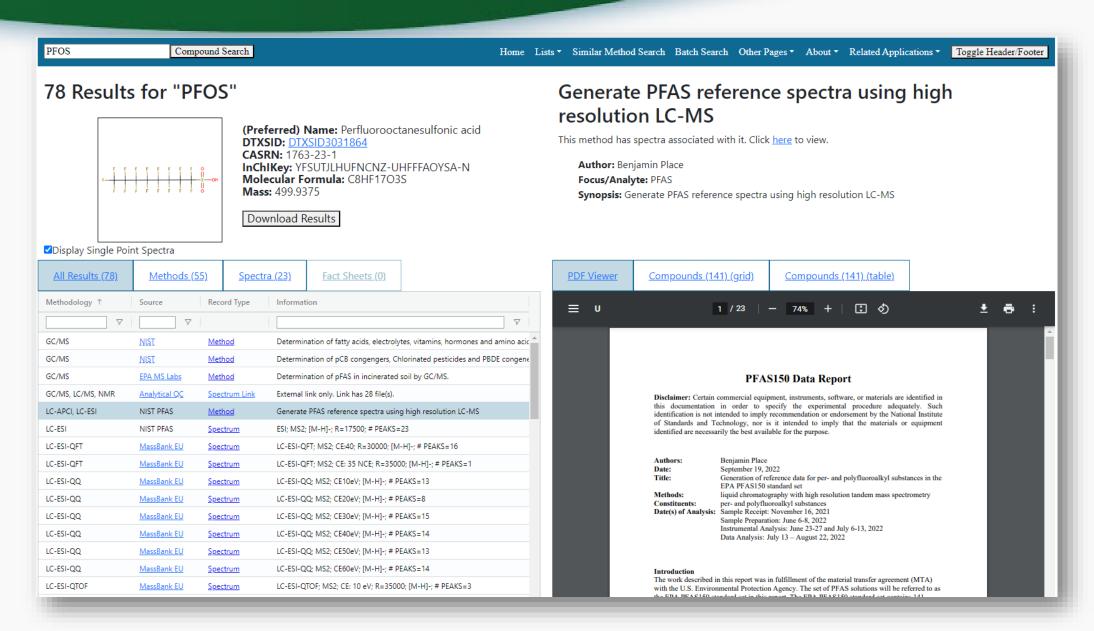


- Almost 3500 in AMOS so far from an assortment of vendors, publications, and government agencies
 - Agencies include: US-EPA, DEA, CDC, FDA, OSHA, USGS, USDA
 - Vendors include: Agilent, Shimadzu, LECO, Sciex and talking to others
- Methods can be linked to sets of spectra

 Searchable can be filtered by analytes, matrix, analytical methodology, source

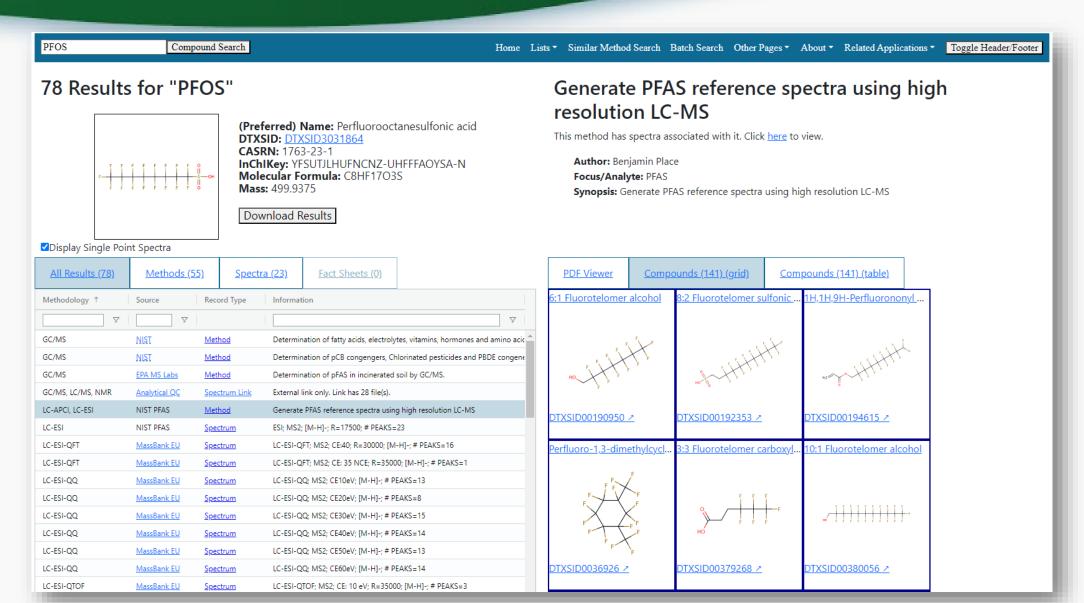
General Search





General Search





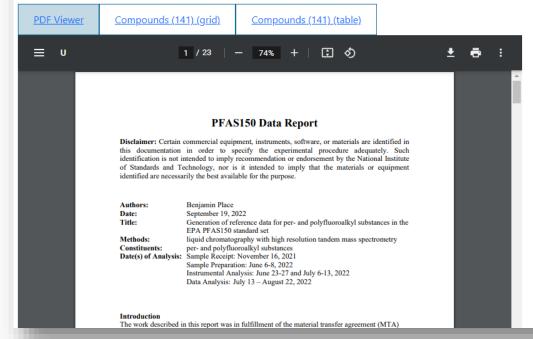
Methods can have associated spectra





Author: Benjamin Place **Focus/Analyte:** PFAS

Synopsis: Generate PFAS reference spectra using high resolution LC-MS

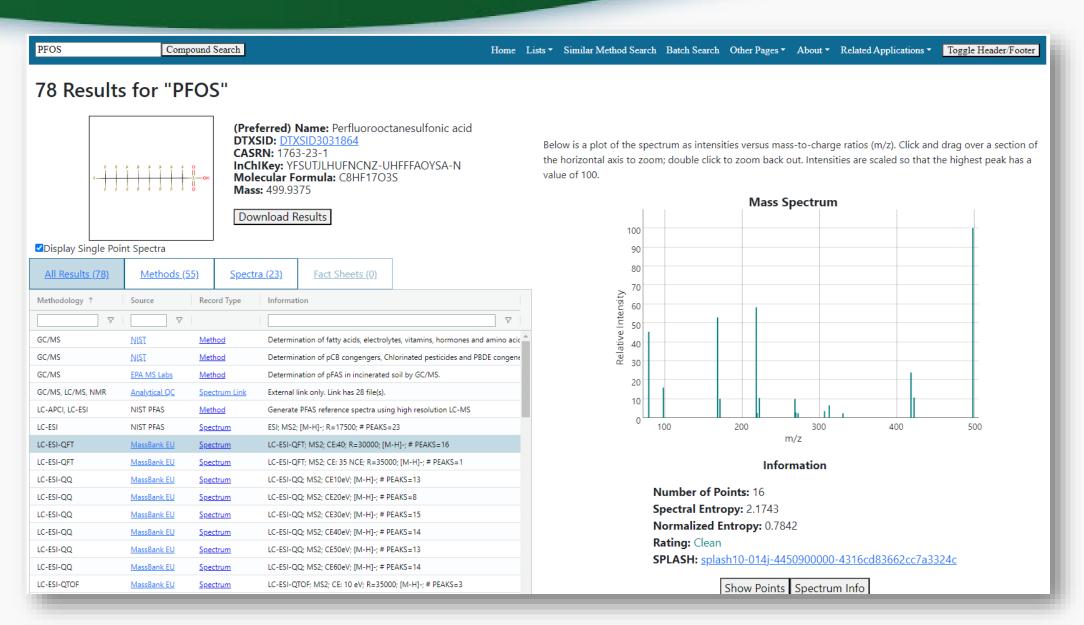


This is a list of the 68 spectra associated with this method, organized by compound identifier. Double-click on a row to show the spectrum in a modal window.

DTXSID	Compound Name
DTXSID00190950	6:1 Fluorotelomer alcohol
DTXSID00194615	1H,1H,9H-Perfluorononyl acrylate
DTXSID0059879	1H,1H,5H-Perfluoropentanol
DTXSID10382147	3-(Perfluoro-2-butyl)propane-1,2-diol
DTXSID20337446	2-(Perfluorooctyl)ethanthiol
DTXSID3066215	(Heptafluorobutanoyl)pivaloylmethane
DTXSID40380797	1H,1H-Perfluoro-3,6,9-trioxadecan-1-ol
DTXSID70381090	1H,1H,8H,8H-Perfluoro-3,6-dioxaoctane-1,8-diol
DTXSID70381151	Perfluorooctanamidine
DTXSID7060332	(Perfluorobutyryl)-2-thenoylmethane
DTXSID80310730	Octafluoroadipamide
DTXSID80375107	11:1 Fluorotelomer alcohol
DTXSID80379721	1H,1H,6H,6H-Perfluorohexane-1,6-diol diacrylate
DTXSID9059832	1H,1H,7H-Dodecafluoro-1-heptanol
DTXSID00192353	8:2 Fluorotelomer sulfonic acid
DTXSID00379268	3:3 Fluorotelomer carboxylic acid
DTXSID1032646	N-Ethylperfluorooctane sulfonamide
DTXSID1037303	Perfluoroheptanoic acid
DTXSID1067629	N-Methylperfluorooctanesulfonamide
DTXSID20179883	2-(Perfluorohexyl)ethylphosphonic acid

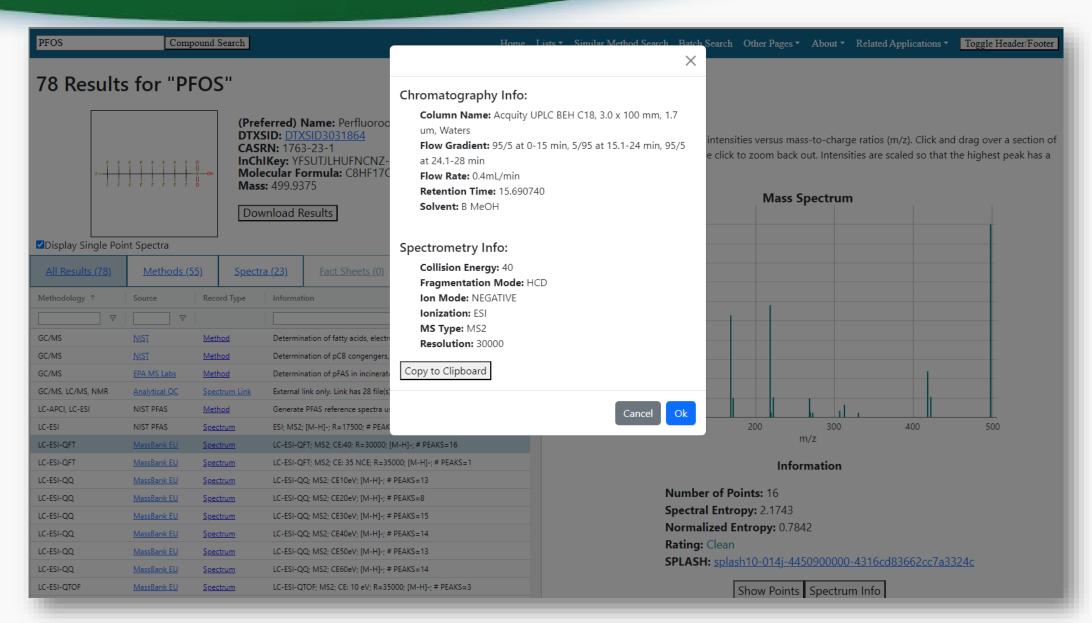
General Search finds spectra too





Chromatographic and Spectral Parameters





Curation is separating spectra-structures Spectral Search based on InChl Key



■ L-(-)-Malic acid 5 spectra	C4H6O5	134.02151
■ L-Hydroxybutanedioic acid, L-(-)-Malic acid, L-Mal 2 spectra	C4H6O5	134.08701
Malic acid 10 spectra	C4H6O5	134.02153
Malic acid-[1,2,3,4-13C4] 6 spectra	C4H6O5	134.02153
■ Malic acid-[1-13C] 6 spectra	C4H6O5	134.02153

Home Page

The search will return records of all three types -- spectra, fact sheets, and methods. The init

BJEPYKJPYRNKOW-REOHCLBHSA-N

Compound Search

○Methods

OFact Sheets

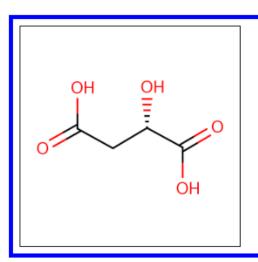
○Spectra

Exact InChlKey Search: 34 spectra/1 method



An exact match for the InChIKey "BJEPYKJPYRNKOW-REOHCLBHSA-N" was found. However, several other InChIKeys with the same first block were discovered, if you were looking for one of those. Click on the InChIKey to be directed to search results for that substance.

EXACT MATCH:



(Preferred) Name: (2S)-2-Hydroxybutanedioic acid

DTXSID: DTXSID30273987

CASRN: 97-67-6

InChlKey: BJEPYKJPYRNKOW-REOHCLBHSA-N

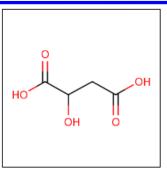
Molecular Formula: C4H6O5

Mass: 134.02153

Records available: 34 spectra, 1 methods

Other spectra based on first block InChlKey We map spectra to specific structures





(Preferred) Name: Malic acid DTXSID: DTXSID0027640

CASRN: 6915-15-7
InChlKey: BJEPYKJPYRNKOW-UHFFFAOYSA-N

Molecular Formula: C4H6O5

Mass: 134.02153

Records available: 66 spectra, 3 methods



(Preferred) Name: 2-Hydroxy(2-¹³C)butanedioic acid

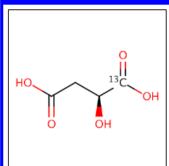
DTXSID: DTXSID90745771 **CASRN:** 143435-96-5

InChlKey: BJEPYKJPYRNKOW-VQEHIDDOSA-N

Molecular Formula: C3[13C]H6O5

Mass: 135.02487

Records available: 5 spectra



(Preferred) Name: (2S)-2-Hydroxy(1-13C)butanedioic acid

DTXSID: DTXSID60745925 **CASRN:** 180991-05-3

InChlKey: BJEPYKJPYRNKOW-GZPBOPPUSA-N

Molecular Formula: C3[13C]H6O5

Mass: 135.02487

Records available: 3 spectra



(Preferred) Name: (+)-Malic acid

DTXSID: DTXSID90892496

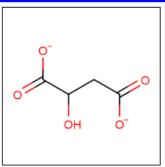
CASRN: 636-61-3

InChlKey: BJEPYKJPYRNKOW-UWTATZPHSA-N

Molecular Formula: C4H6O5

Mass: 134.02153

Records available: 23 spectra



(Preferred) Name: 2-Hydroxybutanedioate

DTXSID: DTXSID00933521

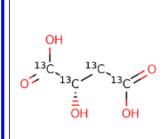
CASRN: 149-61-1

InChlKey: BJEPYKJPYRNKOW-UHFFFAOYSA-L

Molecular Formula: C4H4O5

Mass: 132.00697

Records available: 1 spectra



(Preferred) Name: Malic acid-[1,2,3,4-13C4]

DTXSID: DTXSID301353776 **CASRN:** 1258274-51-9

InChlKey: BJEPYKJPYRNKOW-JCDJMFQYNA-N

Molecular Formula: [13C]4H6O5

Mass: 138.03494

Records available: 6 spectra

Methods List Filtering



Compound Search Home Lists ▼ Similar Method Search Batch Search Other Pages ▼ About ▼ Related Applications ▼ Toggle Header/Footer Below is a list of methods currently in the database. Double-click on a row to display the method and its compounds in another tab. Cells in the table with a dashed underline have hovertext, usually for expanding out abbreviations in the cell content (it will take a second or two for the hovertext to appear). 2714 methods in total are present in the database; 2714 are currently displayed. Full Table Filter 0 Copy filters to clipboard Download Compounds Reset Filters Method # Year Methodology Chemical Class Matrix Name Source Analyte # Compounds EPA-1656a Method 1656, Revision A: Organo-Halide I 2020 GC/HSD <u>USEPA</u> Organohalide pesticides Pesticides Water (waste), soil, sludge, sediment, and fish tissue 31 EPA-1657a Method 1657, Revision A: Organo-Phosph 2020 GC/FPD **USEPA** Organophosphorus Pesticides Pesticides Water (waste), soil, sludge, sediment, and fish tissue 47 EPA-601 Method 601: Purgeable Halocarbons 1994 GC/MS USEPA Halocarbons Halocarbons Municipal and industrial wastewater 29 CLG-MRM3.03 Screening and Confirmation of Animal Dru 2022 <u>USDA</u> LC/MS Drugs Drugs Kidney tissue (bovine, avian, porcine, caprine, ovine 108 CLG-PFAS2.03 Screening, Determination and Confirmatic 2021 LC/MS USDA **PFAS** PFAS Kidney tissue (bovine, avian, porcine, caprine, ovine 16 CLG-MRM1.08 Screening and Confirmation of Animal Dru 2018 LC/MS USDA Drugs Drugs Kidney tissue (bovine, avian, porcine, caprine, ovine 97 CLG-AMG2.08 Screening and Confirmation for Aminogly 2021 LC/MS USDA Aminoglycosides Kidney tissue (bovine, porcine, avian, ovine, caprine 9 Drugs CLG-AMG4.03 Screening for Aminoglycosides by LC-MS- 2020 LC/MS USDA Aminoglycosides Drugs Kidney tissue (bovine, porcine, avian, ovine, caprine 12 CLG-PST5.09 Screening for Pesticides by LC/MS and GC 2022 GC/MS;LC/MS USDA Pesticides Pesticides Muscle tissue (bovine, caprine, equine, ovine, porcii 117 CLG-AGON1.10 Screening Determination and Confirmatio 2018 USDA Beta-agonists Drugs Liver tissue (bovine, porcine), and muscle tissue (bc 5 LC/MS CLG-AVR.04 Determination of Ivermectin, Doramectin, 2011 LC/FLD USDA Ivermectin, Doramectin and Moxidectin Drugs Liver tissue (bovine, ovine, porcine, caprine, equine) 7 CLG-AVR1.03 Liquid Chromatography/Atmospheric Pres 2011 LC/MS USDA Ivermectin, Doramectin and Moxidectin Drugs Liver tissue (bovine, ovine, porcine, caprine, equine) 3 CLG-CAM.07 Determination and Confirmation of Chlora 2018 GC/ECD;GC/MS USDA Chloramphenicol Drugs Muscle tissue (bovine, avian, fish, porcine, equine) 1 CLG-CBX4.02 Screen of Carbadox Metabolite QCA by HI 2016 Carbadox degradate QCA

Vet drug

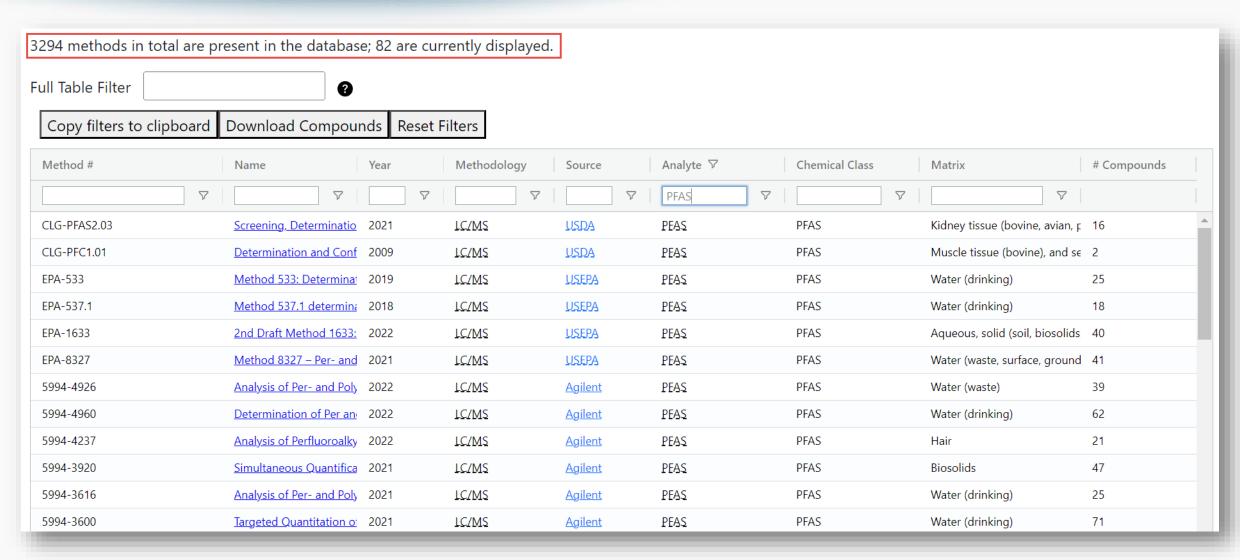
Liver tissue (porcine)

LC/MS

USDA

Methods List: 82 Methods regarding PFAS





Similar Method Search



Compound Identifier 1-Propionyl-lysergic acid Method Search

The table below lists methods for compounds that are similar to "1-Propionyl-lysergic acid diethylamide".

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.

Met	thod Name (# compounds)	Source	Methodology	Year	Simi	i ↓	Similar DTXSID	Compound
~	Highly sensitive determination c	Analytical and	LC/MS	2014	~	0.88		
		Analytical and	LC/MS	2014		0.88	DTXSID1023231	Lysergide
		Analytical and	LC/MS	2014		0.54	DTXSID00436244	2-Oxo-3-hy
~	A fully automated approach for	Journal of Chrc	LC/MS	2018	~	0.88		
		Journal of Chrc	LC/MS	2018		0.88	DTXSID1023231	Lysergide
>	Performance Evaluation of LC-M	Journal of AOA	LC/MS	2019	>	0.84		
>	Development of a Novel LC-MS,	Molecules	LC/MS	2023	>	0.84		
>	Screening and Verifying Mycoto	Agilent	LC/MS	2016	>	0.64		
>	Determination of 63 mycotoxins	Food Control	LC/MS	2023	>	0.64		
>	Quantitative Screening of Multir	Agilent	LC/MS	2021	>	0.63		
>	An End-To-End Workflow for Qu	Agilent	LC/MS	2020	>	0.63		
>	Quantitative Screening of Multir	Agilent	LC/MS	2021	>	0.63		
>	Fast, rugged and sensitive ultra I	Journal of Chrc	LC/MS	2015	>	0.59		
>	Method development and valida	Environmental	LC/MS	2021	>	0.59		
>	Quantitation of Microcystins and	SCIEX	LC/MS	2017	>	0.59		
>	Identification of Unknown Micro	Agilent	LC/MS	2014	>	0.59		
>	LC-MS/MS Validation and Quant	Toxins	LC/MS	2022	>	0.59		
>	On-line solid-phase extraction c	Toxicon	LC/MS	2015	>	0.59		
>	Development and Application o	Toxins	LC/MS	2022	>	0.59		
_	Multi manal data atian af dansa a	CCIEV	LC/MS	2020		0.50		

Highly sensitive determination of 68 psychoactive pharmaceuticals, illicit drugs, and related human metabolites in wastewater by liquid chromatographytandem mass spectrometry

Author: Viola L. Borova, Niki C. Maragou, Pablo Gago-Derrero, Constantinos Pistos, Nikolaos S. Thomaidis

Focus/Analyte: Psychoactive substances

Synopsis: Determination of psychoactive substances in water (waste) by LC/MS.

PDF Viewer	Compo	ounds (68) (grid)	Comp	ounds (68) (table)	
<u>Chlorpromazine</u>		<u>Ephedrine</u>		<u>Mirtazapine</u>	
DTXSID0022808 2		CH ₃	сн₃	H ₃ C N N N	
DTXSID0022808 >		DTXSID0022985 /		DTXSID0023325 🗷	
N-Desmethylcloza	<u>pine</u>	2-Oxo-3-hydroxy-l	<u>SD</u>	Ecgonine methyl est	er er
CI NIH		CH ₅ CH ₅		H _C NOH	.CH₃
DTXSID0042616 >		DTXSID00436244 /	_	DTXSID00891435 ∠	
<u>Oxazepam</u>		<u>Lysergide</u>		<u>Thiopental</u>	

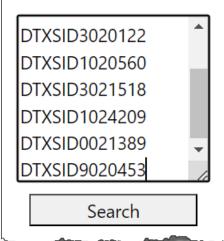
Batch Search



This page allows for batch searching of compounds by DTX here to retrieve them.

The results returned from this search are essentially the sar retrieved from the database.

Note: if two or more searched DTXSIDs are present in the s in it.



Searching a set of DTXSIDs downloads sets of spectra and methods and links to original data

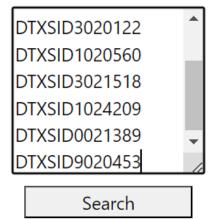
Batch Search



This page allows for batch searching of compounds by DTX here to retrieve them.

The results returned from this search are essentially the sar retrieved from the database.

Note: if two or more searched DTXSIDs are present in the s in it. Searching a set of DTXSIDs downloads sets of spectra and methods and links to original data



dtxsid 🔻	casrn 💌	preferr	methodologies *	source 🕶	link	▼ re	ecord_ 🔻	description
DTXSID0020446	330-54-1	Diuron	['LC/MS']	Analytical	http://v	/26 Sp	ectrum	External link only. Link has 8 file(s).
DTXSID0020446	330-54-1	Diuron	['LC/MS']	MoNA	https://	m(Sp	oectrum	MS2; Positive; # PEAKS=9
DTXSID0020446	330-54-1	Diuron	['LC/MS']	MoNA	https://	m(Sp	oectrum	MS2; Negative; # PEAKS=53
DTXSID0020446	330-54-1	Diuron	['LC/MS']	MoNA	https://	m(Sp	oectrum	MS2; Positive; # PEAKS=25
DTXSID0020446	330-54-1	Diuron	['LC/MS']	MoNA	https://	m(Sp	oectrum	MS2; Positive; # PEAKS=54
DTXSID0020446	330-54-1	Diuron	['GC/MS', 'LC/MS'	USDA	https://	w، M	lethod	Determination of pesticides in muscle tissue (bovine, caprine, equine, ovine, p
DTXSID0020446	330-54-1	Diuron	['LC/MS']	EPA-ECM	https://	w، M	lethod	Determination of linuron, diuron and their degradates in water (drinking, grounds)
DTXSID0020446	330-54-1	Diuron	['LC/MS']	EPA-ECM	https://	w، M	lethod	DER of method for determination of diuron, Linuron, Desmethoxy linuron, No
DTXSID0020446	330-54-1	Diuron	['LC/MS']	EPA-ECM	https://	w، M	lethod	Determination of linuron, Diuron and their degradates in soil (loam, silty clay
DTXSID0020446	330-54-1	Diuron	['LC/MS']	EPA-ECM	https://	w، M	lethod	ILV of method for determination of linuron, Diuron and their degradates in so
DTXSID0020446	330-54-1	Diuron	['LC/MS']	EPA-ECM	https://	w، M	lethod	DER of method for determination of diuron and its degradates in water (drink
DTXSID0020446	330-54-1	Diuron	['LC/MS']	EPA-ECM	https://	w، M	lethod	Determination of diuron, Linuron and their degradates in water by LC/MS with
DTXSID0020446	330-54-1	Diuron	['LC/MS']	EPA-EMC	https://	w، M	lethod	ILV of method for determination of diuron degradates PDMU and cPMU in wa
DTXSID0020446	330-54-1	Diuron	['LC/MS']	EPA-ECM	https://	w، M	lethod	ILV of method for determination of diuron and its degradates DCPMU and mC
DTXSID0020446	330-54-1	Diuron	['LC/MS']	USEPA	https://	w\ M	lethod	This method covers the use of high-performance liquid chromatography (HPL

Spectrum Search



This page is intended for searching the database for spectra. Currently, this is just being done by mass (with a margin of error) and the methodology of the spectrum.

Mass Information 194 Da ± 0.1 Da Da Dppm WARNING: Wide mass intervals can severly slow down the app due to retrieving large spectra; there are currently no limits on what mass values can be collected, so be careful. Methodology LC/MS ~

User Spectrum

53.84601 0.040218 59.7077 0.039517 69.04531 0.269281 71.112045 0.036707 83.06077 0.233345 110.071434 2.046293

Search

Search found 292 spectra covering 27 substances.

> DTXSID0020232 (Caffeine) (41) > 1.0000 > DTXSID5040673 (Ferulic acid) (34) > 0.7781 > DTXSID1049431 (Anthrone) (1) > 0.7629 > DTXSID20282301 (5,6-DIMETHOXY-3H-ISOBENZC > 0.7511 V DTXSID601029528 (Quebrachitol) (56) V 0.5797 V DTXSID601029528 (Quebrachitol) (56) V 0.5794 MS2; CE=13V; # PEAKS=8 0.5498 MS2; CE=13V; # PEAKS=4	
> DTXSID1049431 (Anthrone) (1) > 0.7629 > DTXSID20282301 (5,6-DIMETHOXY-3H-ISOBENZC) > 0.7511 V DTXSID601029528 (Quebrachitol) (56) V 0.5797 0.5797 Orbitrap; MS2; CE=2V; # PEAKS=8 0.5794 MS2; CE=13V; # PEAKS=8	
> DTXSID20282301 (5,6-DIMETHOXY-3H-ISOBENZC > 0.7511 V DTXSID601029528 (Quebrachitol) (56) V 0.5797 0.5797 Orbitrap; MS2; CE=2V; # PEAKS=8 0.5794 MS2; CE=13V; # PEAKS=8	
V DTXSID601029528 (Quebrachitol) (56) V 0.5797 0.5797 Orbitrap; MS2; CE=2V; # PEAKS=8 0.5794 MS2; CE=13V; # PEAKS=8	
0.5797 Orbitrap; MS2; CE=2V; # PEAKS=8 0.5794 MS2; CE=13V; # PEAKS=8	
0.5794 MS2; CE=13V; # PEAKS=8	
0.5498 MS2; CE=13V; # PEAKS=4	
0.4944 Orbitrap; MS2; CE=1V; # PEAKS=5	
0.4799 Orbitrap; MS2; CE=4V; # PEAKS=7	
0.3116 Orbitrap; MS2; CE=2V; # PEAKS=12	
0.3113 Orbitrap; MS2; CE=0V; # PEAKS=2	
0.2911 Orbitrap; MS2; CE=3V; # PEAKS=13	-

Combined Mass Spectra

Spectrum Search



This page is intended for searching the database for spectra. Currently, this is just being done by mass (with a margin of error) and the methodology of the spectrum.

Mass Information

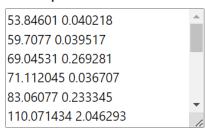


WARNING: Wide mass intervals can severly slow down the app due to retrieving large spectra; there are currently no limits on what mass values can be collected, so be careful.

Methodology

LC/MS ~

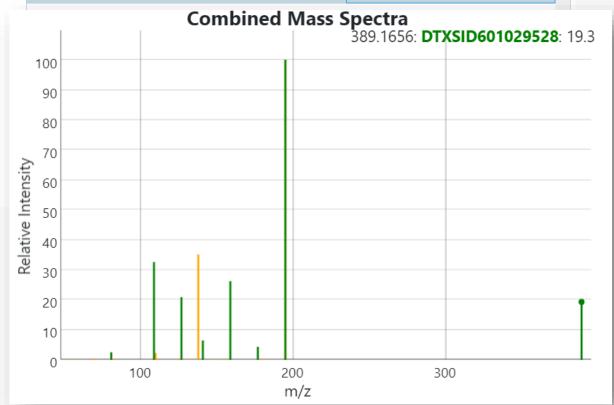
User Spectrum



Search

Search found 292 spectra covering 27 substances.

DTXSID (Name)	Sim	ilarity	Description	
> DTXSID0020232 (Caffeine) (41)	>	1.0000		^
> DTXSID5040673 (Ferulic acid) (34)	>	0.7781		
> DTXSID1049431 (Anthrone) (1)	>	0.7629		
> DTXSID20282301 (5,6-DIMETHOXY-3H-ISOBENZC	>	0.7511		
∨ DTXSID601029528 (Quebrachitol) (56)	~	0.5797		
		0.5797	Orbitrap; MS2; CE=2V; # PEAKS=8	



How can these data be used?



 Assembling spectral and methods data provides additional data for building training algorithms

Analytical and Bioanalytical Chemistry (2021) 413:7495–7508 https://doi.org/10.1007/s00216-021-03713-w

RESEARCH PAPER



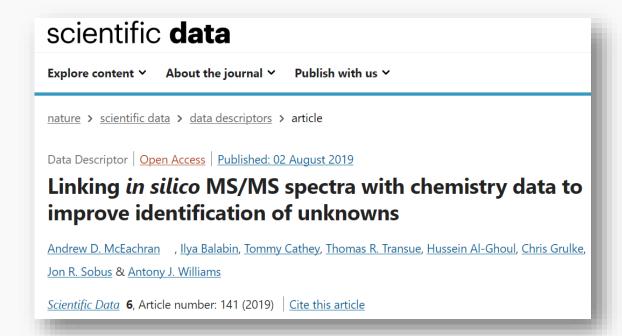
Predicting compound amenability with liquid chromatography-mass spectrometry to improve non-targeted analysis

Charles N. Lowe¹ · Kristin K. Isaacs¹ · Andrew McEachran² · Christopher M. Grulke¹ · Jon R. Sobus¹ · Elin M. Ulrich¹ · Ann Richard¹ · Alex Chao¹ · John Wambaugh¹ · Antony J. Williams¹

Spectral Searching for the NTA WebApp



NTA is supported by searching against predicted spectra

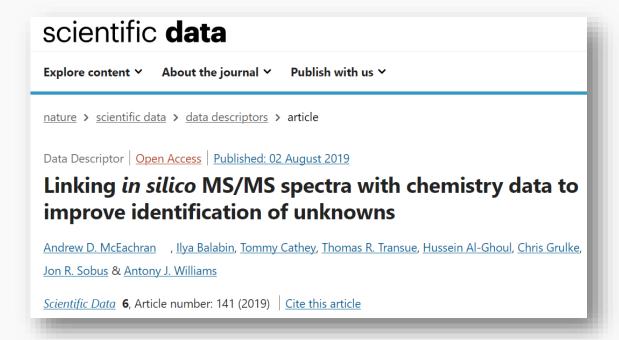


 Now experimental spectral data are being used

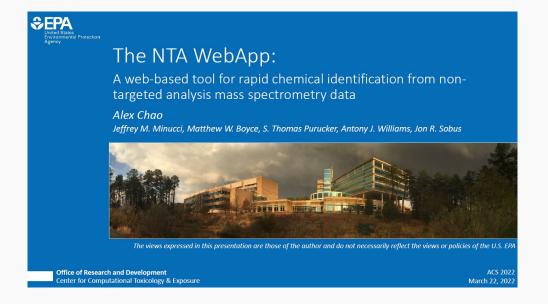
Spectral Searching for the NTA WebApp



 NTA is supported by searching against predicted spectra – NTA WebApp



 Now experimental spectral data are being used



Future Work

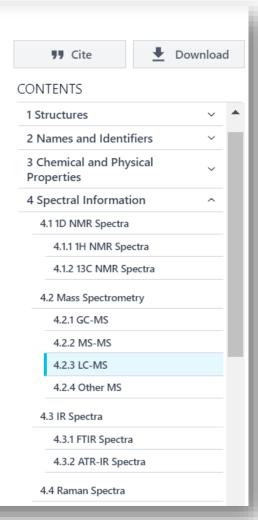


- Add more data assembled from EPA labs (standards)
- Improvements to spectral searching in testing
- Expand spectral and chromatographic metadata
- Structure, substructure and similarity searching
- Integration with other EPA applications
 - Chemical transformations database
 - ToxCast/Tox21 Analytical QC dataset for >8,000 chemicals
- Hoping to release to the public in 2024

Other sources of Methods and Spectra PubChem



4.2.3 LC-MS		@ Ľ
1 of 14		View All ☑
Accession ID	MSBNK-Keio_Univ-KO001305	
Authors	Kakazu Y, Horai H, Institute for Advanced Biosciences, Keio Univ.	
Instrument	API3000, Applied Biosystems	
Instrument Type	LC-ESI-QQ	
MS Level	MS2	
Ionization Mode	NEGATIVE	
Collision Energy	10 V	
Precursor m/z	133	
Precursor Adduct	[M-H]-	
Top 5 Peaks	132.9 999	
	114.9 107	
	89 4	
	70.9 3	
	87.3 3	
SPLASH	splash10-001i-0900000000-1decf12f117200e0f28d	



Other sources of Methods and Spectra PubChem



4.2.3 LC-MS

1 of 14	
Accession ID	MSBNK-Keio_Univ-KO001305
Authors	Kakazu Y, Horai H, Institute for Ac
Instrument	API3000, Applied Biosystems
Instrument Type	LC-ESI-QQ
MS Level	MS2
Ionization Mode	NEGATIVE
Collision Energy	10 V
Precursor m/z	133
Precursor Adduct	[M-H]-
Top 5 Peaks	132.9 999
	114.9 107
	89 4
	70.9 3
	87.3 3
SPLASH	splash10-001i-0900000000-1decf

44				C •		
11		\triangle r	1 t r	tic	'atı	on
	ıu		IUI		au	\mathbf{O}

@ 4

11.1 Analytic Laboratory Methods



9.123, 22.073. CORDIALS & LIQUEURS /OF/ ... TOTAL MALIC ACID (LEVO & INACTIVE) /DETERMINED/ BY CHROMATOGRAPHIC METHOD. 9.124, 22.081. CORDIALS & LIQUEURS /OF/ LEVO-MALIC ACID /DETERMINED BY/ POLARIZATION. /L-MALIC ACID/

Association of Official Analytical Chemists. Official Methods of Analysis. 10th ed. and supplements. Washington, DC: Association of Official Analytical Chemists, 1965. New editions through 13th ed. plus supplements, 1982., p. 13/160

▶ Hazardous Substances Data Bank (HSDB)

12.020, 22.073. BEVERAGES: NONALCOHOLIC & CONCENTRATES /OF/ TOTAL MALIC ACID /DETERMINED BY/ CHROMATOGRAPHIC METHODS. 12.021, 22.081. BEVERAGES: NONALCOHOLIC & CONCENTRATES /OF/ LEVO MALIC ACID /DETERMINED BY/ POLARIZATION. /L-MALIC ACID/

Association of Official Analytical Chemists. Official Methods of Analysis. 10th ed. and supplements. Washington, DC: Association of Official Analytical Chemists, 1965. New editions through 13th ed. plus supplements, 1982., p. 13/194

▶ Hazardous Substances Data Bank (HSDB)

11.047, 22.088. BEVERAGES: WINES /OF/ ... CITRIC & MALIC ACIDS /DETERMINED BY/ POLARIZATION.

Association of Official Analytical Chemists. Official Methods of Analysis. 10th ed. and supplements. Washington, DC: Association of Official Analytical Chemists, 1965. New editions through 13th ed. plus supplements, 1982., p. 13/160

Conclusions



 The AMOS database combines multiple kinds of data associated with analytical chemistry – primarily MS-related

DSSTox database provides the underpinning chemistry

 Data can be interrogated via a web-based interface and enabled by cheminformatics

Methods database is growing by 10-20 methods per day

If you want to help....



 Send information regarding analytical methods and method articles to <u>williams.antony@epa.gov</u>