

The EPA CompTox Chemicals Dashboard to support structure identification and chemical forensics using mass spectrometry

***Antony Williams¹, Alex Chao², Tom Transue³, Tommy Cathey³,
Elin Ulrich⁴ and Jon Sobus⁴***

1) *National Center for Computational Toxicology, U.S. Environmental Protection Agency, RTP, NC*

2) *Oak Ridge Institute of Science and Education (ORISE) Research Participant, RTP, NC*

3) *General Dynamics Information Technology, RTP, NC*

4) *National Exposure Research Laboratory, U.S. Environmental Protection Agency, RTP, NC*

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


*Fall 2019
ACS Fall Meeting, San Diego*

- The CompTox Chemicals Dashboard - web-based database of 875k substances
- Associated data including:
 - Experimental and predicted physicochemical data
 - *In vivo* hazard data
 - *In vitro* bioactivity screening data
 - Link farm to tens of public resources
- Integrated modules – read-across, lit search
- Data mappings and searches supporting Mass Spectrometry & structure identification

CompTox Chemicals Dashboard

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





United States
Environmental Protection
Agency

HomeAdvanced SearchBatch SearchLists▼PredictionsDownloads

Share▼



ChemicalsProduct/Use CategoriesAssay/Gene

☐ Identifier substring search

875 Thousand Chemicals

See what people are saying, read the dashboard [comments!](#)
Cite the Dashboard Publication [click here](#)

Latest News

[Read more news](#)


Journal of Cheminformatics article regarding "MS-Ready structures"


March 9th, 2019 at 1:09:45 PM

A recent article describes "MS-Ready structures", what they are, how they are generated and details regarding the benefits of these structures in navigating structure relationships across the dashboard. The article is published in the Journal of Cheminformatics [here](#).


BASIC Search

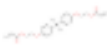
Chemicals **Product/Use Categories** **Assay/Gene**

 Bisphenol

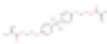


Bisphenol A
DTXSID7020182

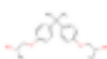




Bisphenol A bis(2-hydroxyethyl ether) diacrylate
DTXSID6066991

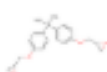


Bisphenol A bis(2-hydroxyethyl ether) dimethacrylate
DTXSID1066992

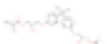


Bisphenol A bis(2-hydroxypropyl) ether
DTXSID8051592

Bisphenol A carbonate polymer
DTXSID6027840



Bisphenol A diglycidyl ether
DTXSID6024624



Bisphenol A glycidyl methacrylate
DTXSID7044841

Detailed Chemical Pages

DETAILS

EXECUTIVE SUMMARY

PROPERTIES

ENV. FATE/TRANSPORT

HAZARD

▶ ADME

▶ EXPOSURE

▶ BIOACTIVITY

SIMILAR COMPOUNDS

GENRA (BETA)

RELATED SUBSTANCES


SYNONYMS

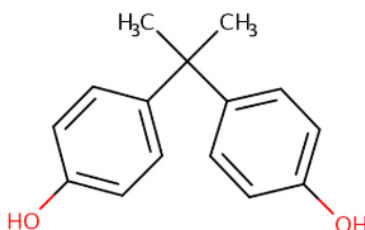
▶ LITERATURE

LINKS

COMMENTS

[Home](#) [Advanced Search](#) [Batch Search](#) [Lists](#) [Predictions](#) [Downloads](#) [Copy](#) [Share](#) [Submit Comment](#)

 **Bisphenol A**
80-05-7 | DTXSID7020182
Searched by DSSTox Substance Id.




Cc1ccc(O)cc1C(C)(C)c2ccc(O)cc2



Wikipedia


Bisphenol A (BPA) is an organic synthetic compound with the chemical formula $(\text{CH}_3)_2\text{C}(\text{C}_6\text{H}_4\text{OH})_2$ belonging to the group of diphenylmethane derivatives and bisphenols, with two hydroxyphenyl groups. It is a colorless solid that is soluble in organic solvents, but poorly soluble in water (0.344 wt % at 83 °C). BPA is a starting material for the synthesis of plastics, primarily certain polycarbonates

[Read more](#)

Intrinsic Properties

 **Molecular Formula:** $\text{C}_{15}\text{H}_{16}\text{O}_2$  [Mol File](#)  [Find All Chemicals](#)

 **Average Mass:** 228.291 g/mol  [Isotope Mass Distribution](#)

 **Monoisotopic Mass:** 228.11503 g/mol

Structural Identifiers


Linked Substances

Presence in Lists

Record Information

Quality Control Notes

Sources of Exposure to Chemicals

 United States
Environmental Protection
Agency

HomeAdvanced SearchBatch SearchLists ▼PredictionsDownloads

Copy ▼Share ▼Submit Comment

Search all data

Bisphenol A

80-05-7 | DTXSID7020182

Searched by DSSTox Substance Id.

DETAILS

EXECUTIVE SUMMARY

PROPERTIES

ENV. FATE/TRANSPORT

HAZARD

▶ ADME

Download ▼

Columns ▼10 ▼

Search query

Product and Use Categories (PUCs)

Product or Use Categorization	Categorization type	Number of Unique Products
manufacturing, metals	CPCat Cassette	17
adhesive	CPCat Cassette	17
	CPCat Cassette	16
	CPCat Cassette	12
	CPCat Cassette	11
	CPCat Cassette	8
	CPCat Cassette	8
	CPCat Cassette	8
	CPCat Cassette	7
	CPCat Cassette	6

First<<<12345678910>>>Last

EXPOSURE

PRODUCT & USE CATEGORIES

CHEMICAL WEIGHT FRACTION

CHEMICAL FUNCTIONAL USE

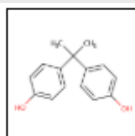
TOXICS RELEASE INVENTORY

MONITORING DATA

EXPOSURE PREDICTIONS

PRODUCTION VOLUME

Identifiers to Support Searches



Bisphenol A

80-05-7 | DTXSID7020182

Searched by Approved Name.

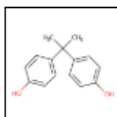
Synonyms

 Download

25

Search query

Synonym	Quality
Bisphenol A	Valid
4,4'-(Propane-2,2-diyl)diphenol	Valid
Phenol, 4,4'-(1-methylethylidene)bis-	Valid
80-05-7 Active CAS RN	Valid
BPA	Valid
4,4'-Propane-2,2-diylidiphenol	Valid
Phenol, 4,4'-(1-methylethylidene)bis-	Valid
4-06-00-06717 Beilstein Registry Number	Beilstein
(4,4'-Dihydroxydiphenyl)dimethylmethane	Good
2,2-Bis(4'-hydroxyphenyl) propane	Good
2,2'-Bis(4-hydroxyphenyl)propane	Good
2,2-BIS-(4-HYDROXY-PHENYL)-PROPANE	Good
2,2-Bis(4'-hydroxyphenyl)propane	Good
2,2-Bis(p-hydroxyphenyl)propane	Good
2,2-Di(4-Hydroxyphenyl) Propane	Good



Bisphenol A

80-05-7 | DTXSID7020182

Searched by Approved Name.

DETAILS

EXECUTIVE SUMMARY

PROPERTIES

ENV. FATE/TRANSPORT

HAZARD

▶ ADME

▶ EXPOSURE

▶ BIOACTIVITY

SIMILAR COMPOUNDS

GENRA (BETA)

RELATED SUBSTANCES

SYNONYMS

▶ LITERATURE

LINKS

COMMENTS

General

EPA Substance Registry Service

Household Products Database

Chemical Entities of Biological Interest (ChEBI)

PubChem

ChempSpider

CPCat

DrugBank

HMDB

Wikipedia

MSDS Lookup

ChEMBL

Chemical Vendors

CalEPA Office of Environmental Health Hazard Assessment

NIOSH Chemical Safety Cards

ToxPlanet

ACS Reagent Chemicals

Wikidata

ChemHat: Hazards and Alternatives Toolbox

Wolfram Alpha

ScrubChem

ECHA Brief Profile

Toxicology

ACToR

DrugPortal

CCRIS

ChemView

CTD

eChemPortal

Gene-Tox

HSDB

ToxCast Dashboard 2

LactMed

International Toxicity Estimates for Risk

ATSDR Toxic Substances Portal

Superfund Chemical Data matrix

NIOSH IDLH Values

ACToR PDF Report

Toxics Release Inventory

CREST

National Air Toxics Assessment

Publications

Toxline

Environmental Health Perspectives

NIEHS

National Toxicology Program

Google Books

Google Scholar

Google Patents

PPRTVWEB

PubMed

IRIS Assessments

EPA HERO

NIOSH Skin Notation Profiles

NIOSH Pocket Guide

RSC Publications

BioCaddie DataMed

Springer Materials

Federal Register

Regulations.gov

Bielefeld Academic Search Engine

CORE Literature Search

Analytical

FOR-IDENT

NEMI: National Environmental Methods Index

RSC Analytical Abstracts

Tox21 Analytical Data

MONA: MassBank North America

mzCloud

NIST NIST IR Spectrum

NIST NIST MS Spectrum

Prediction

2D NMR HSQC/HMBC Prediction

Carbon-13 NMR Prediction

Proton NMR Prediction


ChemRTP Predictor


LSERD


NIST WebBook

<https://webbook.nist.gov/chemistry/>


Analytical

 FOR-IDENT

 NEMI: National Environmental
Methods Index

 RSC Analytical Abstracts

 Tox21 Analytical Data

 MONA: MassBank North
America

 mzCloud

 NIST IR Spectrum

 NIST MS Spectrum

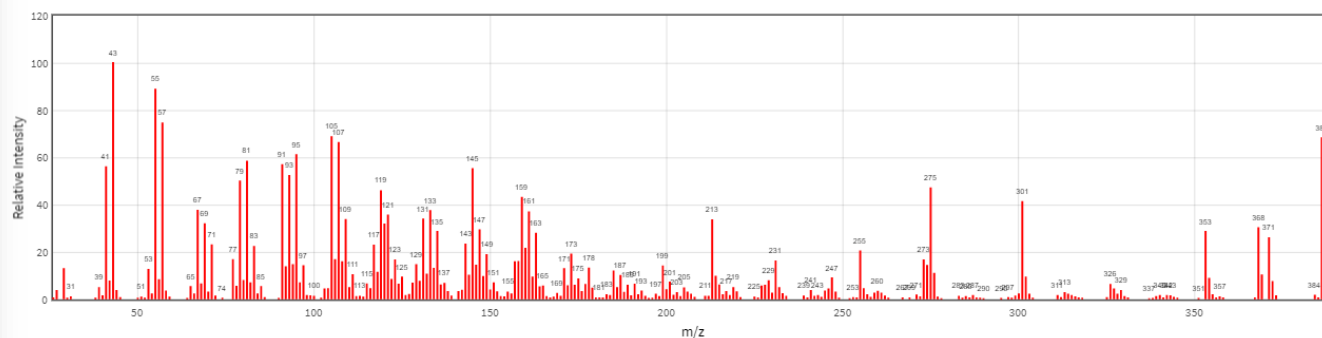
Spectrum

Plot

Help / Software credits

Cholesterol


Mass Spectrum





MassBank of North America

<https://mona.fiehnlab.ucdavis.edu>


Analytical


 FOR-IDENT

 NEMI: National Environmental
Methods Index

 RSC Analytical Abstracts





 Tox21 Analytical Data



 MONA: MassBank North
America

 mzCloud





 NIST NIST IR Spectrum

 NIST NIST MS Spectrum

MoNA - MassBank of North America  Spectra  Downloads  Upload  Help

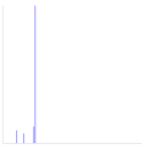
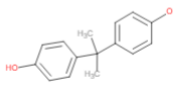
Search...  

Display Generated Query

    9




10 records/page ▾

Bisphenol A Score: ★★★★★

Q instrument	LTQ Orbitrap XL Thermo Sc...
Q instrument type	LC-ESI-ITFT
Q ms level	MS2
Q ionization	ESI
Q collision energy	30 % (nominal)
Q retention time	14.0 min
Q precursor m/z	229.1223
Q precursor type	[M+H] ⁺
Q ionization mode	positive
Q accession	EA016309

Originally submitted to the MassBank High Quality Mass Spectral Database

Batch Searching

Aggregate data for a list of chemicals



Trends in Environmental Analytical Chemistry

Volume 20, October 2018, e00059



Opioid occurrence in environmental water samples—A review

Marina Celia Campos-Mañas ^a, Imma Ferrer ^b  , E. Michael Thurman ^b, Ana Agüera ^a

 [Show more](#)

<https://doi.org/10.1016/j.teac.2018.e00059>

[Get rights and content](#)

Batch Search Names

Buprenorphine
 Codeine
 Dextromethorphan
 Dihydrocodeine
 Dihydromorphine
 Ethylmorphine
 Fentanyl
 Heroin
 Hydrocodone
 Hydromorphone
 Ketamine
 Meperidine
 Methadone
 Morphine
 Morphinone
 Naloxone
 Naltriben
 Oxycodone
 Oxymorphone
 Propoxyphene
 Sufentanil
 Tramadol

Step 1 Step 2 Step 3 Step 4 Step 5 Step 6

Step Five: Choose Data Fields to Download

Please enter one identifier per line

Select Input Type(s)

- ☒ Identifiers
 - ☒ Chemical Name ⓘ
 - ☐ CASRN ⓘ
 - ☐ InChIKey ⓘ
 - ☐ DSSTox Substance ID ⓘ
 - ☐ DSSTox Compound ID ⓘ
 - ☐ InChIKey Skeleton ⓘ
 - ☐ MS-Ready Formula(e) ⓘ
 - ☐ Exact Formula(e) ⓘ
 - ☐ Monoisotopic Mass ⓘ

Enter Identifiers to Search (searches should be limited to <5000 identifiers)







Buprenorphine
 Codeine
 Dextromethorphan
 Dihydrocodeine
 Dihydromorphine
 Ethylmorphine
 Fentanyl
 Heroin
 Hydrocodone
 Hydromorphone

Excel
Download






INPUT	FOUND_BY	DTXSID
Buprenorphine	Approved Name	DTXSID2022705
Codeine	Approved Name	DTXSID2020341
Dextromethorphan	Approved Name	DTXSID3022908
Dihydrocodeine	Approved Name	DTXSID5022936
Dihydromorphine	Approved Name	DTXSID7048908
Ethylmorphine	Approved Name	DTXSID1046760
Fentanyl	Approved Name	DTXSID9023049
Heroin	Synonym	DTXSID6046761
Hydrocodone	Approved Name	DTXSID8023131
Hydromorphone	Approved Name	DTXSID8023133
Ketamine	Approved Name	DTXSID8023187
Meperidine	Approved Name	DTXSID9023253
Methadone	Approved Name	DTXSID7023273
Morphine	Approved Name	DTXSID9023336

Add Other Data of Interest






Chemical Identifiers

- ☒ DTXSID 
- ☒ 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 

INPUT	DTXSID	CASRN	MOLECULAR_FORMULA	MONOISOTOPIC	MS_READY_SMILES
Buprenorphine	DTXSID202	52485-79-7	C29H41NO4	467.3035588	[H]C12CC3=C4C
Codeine	DTXSID202	76-57-3	C18H21NO3	299.1521435	[H]C12CC3=C4C
Dextromethamphetamine	DTXSID302	125-71-3	C18H25NO	271.1936144	[H]C12CC3=C(C=
Dihydrocodone	DTXSID502	125-28-0	C18H23NO3	301.1677936	[H]C12CC3=C4C
Dihydromorphine	DTXSID704	509-60-4	C17H21NO3	287.1521435	[H]C12CC3=C4C
Ethylmorphine	DTXSID104	76-58-4	C19H23NO3	313.1677936	[H]C12CC3=C4C
Fentanyl	DTXSID902	437-38-7	C22H28N2O	336.2201635	CCC(=O)N(C1CC
Heroin	DTXSID604	561-27-3	C21H23NO5	369.1576228	[H]C12CC3=C4C
Hydrocodone	DTXSID802	125-29-1	C18H21NO3	299.1521435	[H]C12CC3=C4C
Hydromorphone	DTXSID802	466-99-9	C17H19NO3	285.1364935	[H]C12CC3=C4C
Ketamine	DTXSID802	6740-88-1	C13H16ClNO	237.0920418	CNC1(CCCCC1=
Meperidine	DTXSID902	57-42-1	C15H21NO2	247.1572289	CCOC(=O)C1(CC
Methadone	DTXSID702	76-99-3	C21H27NO	309.2092645	CCC(=O)C(CC(C)
Morphine	DTXSID902	57-27-2	C17H19NO3	285.1364935	[H]C12CC3=C4C
Morphinone	DTXSID501	467-02-7	C17H17NO3	283.1208434	[H]C12CC3=C4C
Naloxone	DTXSID802	465-65-6	C19H21NO4	327.1470582	[H]C12CC3=C4C
Naltrexone	-	-	-	-	-
Oxycodone	DTXSID502	76-42-6	C18H21NO4	315.1470582	[H]C12CC3=C4C
Oxymorphone	DTXSID502	76-41-5	C17H19NO4	301.1314081	[H]C12CC3=C4C
Propoxyphene	DTXSID102	469-62-5	C22H29NO2	339.2198292	CCC(=O)OC(CC1
Sufentanil	DTXSID602	56030-54-7	C22H30N2O2S	386.2027994	CCC(=O)N(C1=C
Tramadol	DTXSID908	27203-92-5	C16H25NO2	263.188529	COC1=CC=CC(=

Chemical Lists of Interest...

Chemical Lists

[Home](#)[Advanced Search](#)[Batch Search](#)[Lists ▾](#)[Predictions](#)[Downloads](#)[Lists of Chemicals](#)[List of Assays](#) Download ▾

Columns ▾

 Copy Filtered Lists URL

List Acronym ▾	List Name ▾	Last Updated ▾	Number of Chemicals ▾	List Description ▾
HDXEXCH	MASSPECDB: Hydrogen Deuterium Exchange Standard Set - Under HDX Conditions	2018-11-07	592	Observed species (deuterated and undeuterated) from the HDXNOEX list under hydrogen deuterium exchange conditions (Ruttkies, Schymanski et al. in prep.)
HDXNOEX	MASSPECDB: Hydrogen Deuterium Exchange Standard Set - No Exchange	2018-11-07	765	Environmental standard set used to investigate hydrogen deuterium exchange in small molecule high resolution mass spectrometry (Ruttkies, Schymanski et al. in prep.)
MASSBANKEUSP	MASSPECDB: MassBank.EU Collection: Special Cases	2017-07-16	263	The MassBank.EU list contains curated chemicals (Schymanski/Williams) associated with the literature/tentative/unknown/SI spectra available on MassBank.EU that are not available as part of the full MassBank collection of reference standard spectra.
MASSBANKREF	MASSPECDB: MassBank Reference Spectra Collection	2017-07-13	1267	This MassBank list contains chemicals associated with the full MassBank collection of reference standard spectra available on MassBank.EU, MassBank.JP and MassBank of North America as well as the Open Data collection, curated by Williams/Schymanski.
MYCOTOXINS	MASSPECDB: Mycotoxins from MassBank.EU	2017-08-02	88	This is a set of mycotoxins, initiated by the contribution of spectra of 90 mycotoxins to MassBank.EU by Justin Renaud and colleagues from Agriculture and Agri-Food Canada, Government of Canada

PFAS lists of Chemicals

Select List

Download

Columns

PFAS

Copy Filtered Lists URL

List Acronym	List Name	Last Updated	Number of Chemicals	List Description
EPAPFAS75S1	PFAS EPA: List of 75 Test Samples (Set 1)	2018-06-29	74	PFAS list corresponds to 75 samples (Set 1) submitted for initial testing screens conducted by EPA researchers in collaboration with researchers at the National Toxicology Program.
EPAPFAS75S2	PFAS EPA: List of 75 Test Samples (Set 2)	2019-02-21	75	PFAS list corresponds to a second set of 75 samples (Set 2) submitted for testing screens conducted by EPA researchers in collaboration with researchers at the National Toxicology Program.
EPAPFASCAT	PFAS EPA Structure-based Categories	2018-06-29	64	List of registered DSSTox "category substances" representing PFAS categories created using ChemAxon's Markush structure-based query representations.
EPAPFASINSOL	PFAS EPA: Chemical Inventory Insoluble in DMSO	2018-06-29	43	PFAS chemicals included in EPA's expanded ToxCast chemical inventory found to be insoluble in DMSO above 5mM.
EPAPFASINV	PFAS EPA: ToxCast Chemical Inventory	2018-06-29	430	PFAS chemicals included in EPA's expanded ToxCast chemical inventory and available for testing.
EPAPFASRL	PFAS EPA: Cross-Agency Research List	2017-11-16	199	EPAPFASRL is a manually curated listing of mainly straight-chain and branched PFAS (Per- & Poly-fluorinated alkyl substances) compiled from various internal, literature and public sources by EPA researchers and program office representatives.
PFASKEMI	PFAS: List from the Swedish Chemicals Agency (KEMI) Report	2017-02-09	2416	Perfluorinated substances from a Swedish Chemicals Agency (KEMI) Report on the occurrence and use of highly fluorinated substances.
PFASMASTER	PFAS Master List of PFAS Substances	2018-07-26	5061	PFASMASTER is a consolidated list of PFAS substances spanning and bounded by the below lists of current interest to researchers and regulators worldwide.
PFASOECD	PFAS: Listed in OECD Global Database	2018-05-16	4729	OECD released a New Comprehensive Global Database of Per- and Polyfluoroalkyl Substances, (PFASs) listing more than 4700 new PFAS
PFASTRIER	PFAS Community-Compiled List (Trier et al, 2015)	2017-07-16	597	PFASTRIER community-compiled public listing of PFAS (Trier et al, 2015)

Batch Search in specific lists

<input type="checkbox"/>	INPUT	DTXSID	MASSBANKREF	NEMILIST	WRTMSD	NORMANPRI	SUSDAT
<input type="checkbox"/>	Buprenorph	DTXSID202	-	-	Y	-	Y
<input type="checkbox"/>	Codeine	DTXSID202	Y	Y	Y	Y	Y
<input type="checkbox"/>	Dextrometh	DTXSID302	Y	Y	Y	-	Y
<input type="checkbox"/>	Dihydrocod	DTXSID502	Y	-	Y	Y	Y
<input type="checkbox"/>	Dihydromor	DTXSID704	-	-	-	-	Y
<input type="checkbox"/>	Ethylmorph	DTXSID104	-	-	Y	-	Y
<input type="checkbox"/>	Fentanyl	DTXSID902	Y	-	Y	-	Y
<input checked="" type="checkbox"/>	Heroin	DTXSID604	Y	-	Y	Y	Y
<input checked="" type="checkbox"/>	Hydrocodor	DTXSID802	Y	Y	Y	Y	Y
<input type="checkbox"/>	Hydromorph	DTXSID802	-	-	Y	-	Y
<input type="checkbox"/>	Ketamine	DTXSID802	Y	-	Y	-	Y
<input checked="" type="checkbox"/>	Meperidine	DTXSID902	Y	-	Y	-	Y
<input type="checkbox"/>	Methadone	DTXSID702	Y	Y	Y	-	Y
<input checked="" type="checkbox"/>	Morphine	DTXSID902	Y	Y	Y	Y	Y
<input type="checkbox"/>	Morphinone	DTXSID501	-	-	-	-	Y
<input checked="" type="checkbox"/>	Naloxone	DTXSID802	-	-	Y	-	Y
<input type="checkbox"/>	Naltriben	-	-	-	-	-	-
<input type="checkbox"/>	Oxycodone	DTXSID502	Y	Y	Y	Y	Y
<input type="checkbox"/>	Oxymorpho	DTXSID502	-	-	Y	-	Y
<input type="checkbox"/>	Propoxyph	DTXSID102	Y	Y	Y	-	Y
<input type="checkbox"/>	Sufentanil	DTXSID602	-	-	Y	-	Y
<input type="checkbox"/>	Tramadol	DTXSID908	Y	Y	Y	Y	Y

“MS-ready” Structures

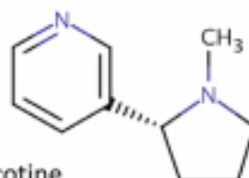
Open Science for Identifying “Known Unknown” Chemicals

Emma L. Schymanski^{*,†} and Antony J. Williams^{*,‡}



Nicotine

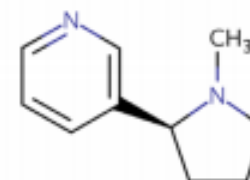
CN1CCC[C@H]1C1=CN=CC=C1
DTXSID1020930 | SNICXCGAKADSCV
54-11-5 | **162.1157** | 0.929 | **72**
Tox: **yes** | Expo: **yes** | Bioassay: **yes**



D-Nicotine

CN1CCC[C@H]1C1=CN=CC=C1
DTXSID004635 | SNICXCGAKADSCV
25162-00-9 | **162.1157** | 0.929 | **20**
Tox: **no** | Expo: **yes** | Bioassay: **yes**

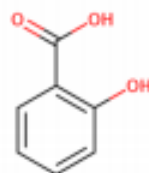
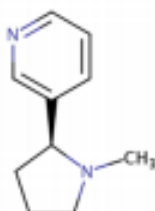
LEGEND: Name, SMILES
DTXSID | InChIKey 1st Block
CAS | Monoiso. Mass | logP | Sources
Data on: Toxicity | Exposure | Bioassays



HCl

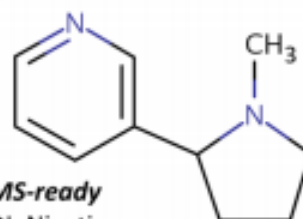
Nicotine hydrochloride

Cl.CN1CCC[C@H]1C1=CN=CC=C1
DTXSID602093 | HDJBTCJIMNXEW
2820-51-1 | **198.0924** | 0.929 | **9**
Tox: **no** | Expo: **yes** | Bioassay: **yes**



Benzoic acid, 2-hydroxy-, compd. with
3-[(2S)-1-methyl-2-pyrrolidinyl]pyridine (1:1)

OC(=O)C1=CC(=O)C=CC=C1.CN1CCC[C@H]1C1=CN=CC=C1
DTXSID5075319 | AIBWPBUAKCMKNS
29790-52-1 | **300.1474** | 0.929 | **6**
Tox: **no** | Expo: **yes** | Bioassay: **no**



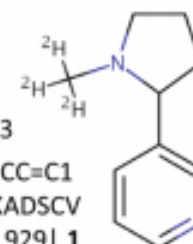
MS-ready

DL-Nicotine

CN1CCCC1C1=CN=CC=C1
DTXSID3048154 | SNICXCGAKADSCV
22083-74-5 | **162.1157** | 0.953 | **9**
Tox: **yes** | Expo: **no** | Bioassay: **yes**

DL-Nicotine-d3

[2H]C([2H])([2H])N1CCCC1C1=CN=CC=C1
DTXSID80442666 | SNICXCGAKADSCV
69980-24-1 | **165.1345** | 0.929 | **1**
Tox: **no** | Expo: **no** | Bioassay: **no**



“MS-Ready Structures”

<https://doi.org/10.1186/s13321-018-0299-2>

McEachran et al. *J Cheminform* (2018) 10:45
<https://doi.org/10.1186/s13321-018-0299-2>

Journal of Cheminformatics

METHODOLOGY


Open Access



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

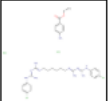
Andrew D. McEachran^{1,2*}, Kamel Mansouri^{1,2,3}, Chris Grulke², Emma L. Schymanski⁴, Christoph Ruttkies⁵ and Antony J. Williams^{2*}

MS-Ready Mappings

 **EPA** United States
Environmental Protection
Agency

Home Advanced Search Batch Search Lists Predictions Downloads

Copy Share Submit Comment Search all data



Progaron

108532-15-6 | DTXSID20148579

Searched by DSSTox Substance Id.

DETAILS

EXECUTIVE SUMMARY

PROPERTIES

ENV. FATE/TRANSPORT

HAZARD

ADME

EXPOSURE

BIOACTIVITY

SIMILAR COMPOUNDS

GENRA (BETA)

RELATED SUBSTANCES

SYNONYMS

LITERATURE

LINKS

COMMENTS

Intrinsic Properties

Molecular Formula: $C_{31}H_{43}Cl_4N_{11}O_2$ Mol File Find All Chemicals

Average Mass: 743.56 g/mol Isotope Mass Distribution

Structural Identification

Linked Substances

Same Connectivity: 1 record

Mixtures, Components: 0 records

MS-Ready Mappings: DTXCID301804: 12 records; DTXCID0013314: 11 records;

Similar Compounds: 0 records

Quality Control Notes

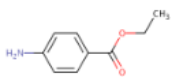
MS-Ready Mappings Set

MS-Ready Mappings of Benzocaine (Isotopes pre-filtered)

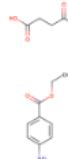
Select all Download Send to Batch Search Default DTXSID PubChem CPDAT

10 of 12 chemicals visible

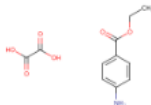
Isotopes Filter by Name or CASRN



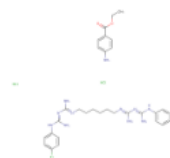
Benzocaine
DTXSID: DTXSID8021804
PubChem: 184
CPDAT: 42



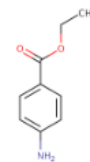
Anesthesine succinate
DTXSID: DTXSID60148336
PubChem: 10
CPDAT: 0



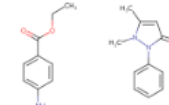
Anesthesine oxalate
DTXSID: DTXSID20148337
PubChem: 6
CPDAT: 0



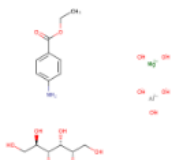
Progaron
DTXSID: DTXSID20148579
PubChem: 5
CPDAT: 0



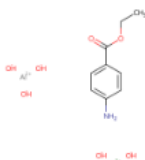
Benzocaine hydrochloride
DTXSID: DTXSID50177812
PubChem: 33
CPDAT: 0



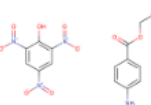
Antipyrine mixture with benzocaine
DTXSID: DTXSID80212866
PubChem: 9
CPDAT: 0



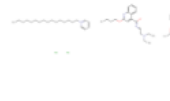
Almagel A-neo
DTXSID: DTXSID60227559
PubChem: 9
CPDAT: 0



Almagel
DTXSID: DTXSID70227560
PubChem: 9
CPDAT: 0



Ethyl 4-aminobenzoate-2,4,6-trinitroph...
DTXSID: DTXSID70787033
PubChem: 5
CPDAT: 0



1-Hexadecylpyridin-1-ium 2-butoxy-N-...
DTXSID: DTXSID50997335
PubChem: 0
CPDAT: 0

Mass and Formula Searching

Advanced Searches

Mass Search

Mass Search

± Min/Max

Adduct

Neutral



All Adducts



Choose adduct from dropdown

191.131

Da

±

5

Da

ppm

Search 

Advanced Searches

Mass Search

Search Results

Searched by Mass: 191.131 +/- 5.0 ppm.

329 of 329 chemicals visible

Select all

Download

Send to Batch Search

Mass Difference

DTXSID

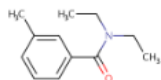
CASRN

TOXCAST

Mass Diff

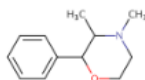
Multicomponent Chemicals

Filter by Name or CASRN



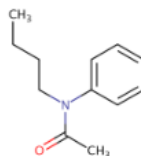
DEET

DTXSID: DTXSID2021995
CASRN: 134-62-3
TOXCAST: 12/768
Mass Diff: 0.000014



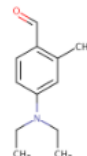
Phendimetrazine

DTXSID: DTXSID1023447
CASRN: 634-03-7
TOXCAST: -
Mass Diff: 0.000014



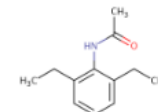
N-Butylacetanilide

DTXSID: DTXSID2042197
CASRN: 91-49-6
TOXCAST: -
Mass Diff: 0.000014



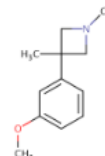
Benzaldehyde, 4-(diethylamino)-2-methyl-

DTXSID: DTXSID4059041
CASRN: 92-14-8
TOXCAST: -
Mass Diff: 0.000014



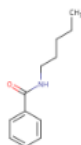
Acetanilide, 2',6'-diethyl-

DTXSID: DTXSID90168148
CASRN: 16665-89-7
TOXCAST: -
Mass Diff: 0.000014



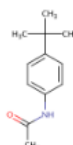
Azetidine, 1,3-dimethyl-3-(m-methoxyphenyl)-

DTXSID: DTXSID40173560
CASRN: 19832-26-9
TOXCAST: -
Mass Diff: 0.000014



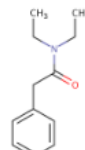
Benzamide, N-pentyl-

DTXSID: DTXSID20174196
CASRN: 20308-43-4
TOXCAST: -
Mass Diff: 0.000014



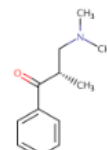
p-t-Butylacetanilide

DTXSID: DTXSID00174238
CASRN: 20330-45-4
TOXCAST: -
Mass Diff: 0.000014



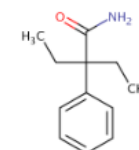
N,N-Diethylphenylacetamide

DTXSID: DTXSID00179048
CASRN: 2431-96-1
TOXCAST: -
Mass Diff: 0.000014



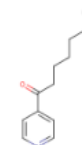
3-(Dimethylamino)-2-methylpropionophenone

DTXSID: DTXSID60180796
CASRN: 26171-50-6
TOXCAST: -
Mass Diff: 0.000014



Butyramide, 2-ethyl-2-phenyl-

DTXSID: DTXSID60184653
CASRN: 30568-39-9
TOXCAST: -
Mass Diff: 0.000014



1-Heptanone, 1-(4-pyridyl)-

DTXSID: DTXSID40186594
CASRN: 32941-30-3
TOXCAST: -
Mass Diff: 0.000014

MS-Ready Structures for Formula Search

Molecular Formula Search

☒ MS Ready Formula  ☐ Exact Formula 










Formula


Please use the format of the following example: C₆H₈O₂ or C₆H(8-10)O(0-2)

Search 

MS-Ready Structures Batch Searches

Select Input Type(s) +/- 5 ppm

- ☐ Identifiers
 - ☐ Chemical Name 
 - ☐ CASRN 
 - ☐ InChIKey 
 - ☐ DSSTox Substance ID 
- ☐ DSSTox Compound ID 
- ☐ InChIKey Skeleton 
- ☐ MS-Ready Formula(e) 
- ☐ Exact Formula(e) 
- ☒ Monoisotopic Mass 

 Display All Chemicals

Enter Identifiers to Search [\(search\)](#)

41.0265
56.02621
53.0265
58.0418|
93.0578
113.9639
151.8754
69.9377
77.9872

This search is based on what we refer to as "Mass Spec (MS) Ready" structures. All chemicals within the database are treated in a manner such that all are desalted, mixtures are separated, and stereochemistry is removed as Mass Spectrometry detects the major components of a salt or mixture and is insensitive to stereochemistry. As an example, a search for the monoisotopic mass of phenol will return phenol, sodium phenolate and calcium phenoxide. See the publication for more details: <https://doi.org/10.1186/s13321-018-0299-2>.


- **EXACT Formula:** C₁₀H₁₆N₂O₈: 3 Hits

☐ MS Ready Formula  ☒ Exact Formula 


Formula

C₁₀H₁₆N₂O₈


Select all


Download 


Send to Batch Search

Default 



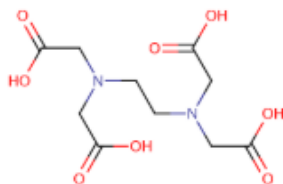
DTXSID 

PubChem 

CPDAT 



3 of 3 chemi

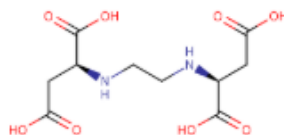


Ethylenediaminetetraacetic acid

DTXSID: DTXSID6022977

PubChem: 158

CPDAT: 387

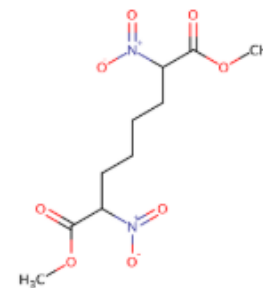


N,N'-Ethylenedi-L-aspartic acid

DTXSID: DTXSID1051852

PubChem: 25

CPDAT: 8



Dimethyl 2,7-dinitrooctanedioate

DTXSID: DTXSID20498864

PubChem: 5

CPDAT: 0

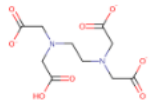
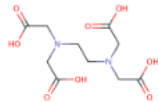
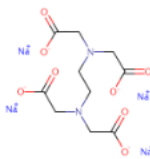
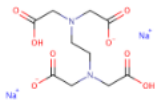
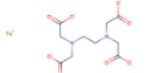
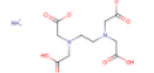
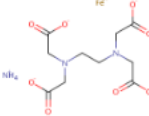
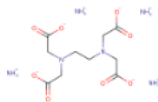
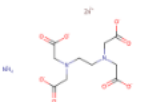
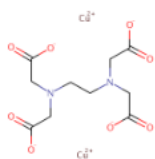
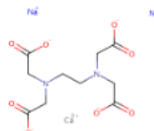
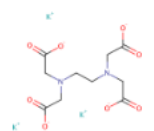
MS-Ready Mappings

- **Same Input Formula: C₁₀H₁₆N₂O₈**
- **MS Ready Formula Search: 125 Chemicals**

125 chemicals

Select all Download Send to Batch Search Default DTXSID PubChem CPDAT

Hide chemicals that are: Filter by Name or CASR

 <p>Trisodium ethylenediaminetetraacetate DTXSID: DTXSID7020556 PubChem: 33 CPDAT: 82</p>	 <p>Ethylenediaminetetraacetic acid DTXSID: DTXSID6022977 PubChem: 158 CPDAT: 387</p>	 <p>Ethylenediaminetetraacetic acid tetrasod... DTXSID: DTXSID3026350 PubChem: 57 CPDAT: 1227</p>	 <p>Ethylenediaminetetraacetic acid, disodiu... DTXSID: DTXSID9027073 PubChem: 56 CPDAT: 1359</p>	 <p>Ethylenediaminetetraacetic acid ferric so... DTXSID: DTXSID5027774 PubChem: 53 CPDAT: 62</p>	 <p>Diammonium dihydrogen ethylenediami... DTXSID: DTXSID9027813 PubChem: 12 CPDAT: 17</p>
 <p>Ferrate(1-), [[N,N'-1,2-ethanediy]bis[N-]]... DTXSID: DTXSID9027815 PubChem: 24 CPDAT: 20</p>	 <p>Tetraammonium ethylenediaminetetraac... DTXSID: DTXSID8027820 PubChem: 11 CPDAT: 12</p>	 <p>Zincate(2-), [[N,N'-1,2-ethanediy]bis[N-]]... DTXSID: DTXSID8028343 PubChem: 5 CPDAT: 10</p>	 <p>EDTA, copper salt DTXSID: DTXSID0034564 PubChem: 8 CPDAT: 10</p>	 <p>Calcium disodium ethylenediaminetetra... DTXSID: DTXSID2036409 PubChem: 42 CPDAT: 29</p>	 <p>Tetrapotassium ethylenediaminetetra... DTXSID: DTXSID3036442 PubChem: 25 CPDAT: 36</p>

- 125 chemicals returned in total
 - 8 of the 125 are **single component** chemicals
 - 3 of the 8 are **isotope-labeled**
 - 3 are **neutral compounds** and 2 are **charged**

- Singleton searches are useful but we work with **thousands** of masses and formulae!
- Typical questions
 - What is the list of chemicals for the formula $C_xH_yO_z$
 - What is the list of chemicals for a mass +/- error
 - Can I get chemical lists in Excel files? In SDF files?
 - Can I include properties in the download file?

Batch Searching Formula/Mass

Batch Search


Step 1 Step 2 Step 3 Step 4 Step 5 Step 6


Step Five: Choose Data Fields to Download


Please enter one identifier per line ✕


Select Input Type(s)


☐ Identifiers


☐ Chemical Name 


☐ CASRN 


☐ InChIKey 


☐ DSSTox Substance ID 


☐ DSSTox Compound ID 

☐ InChIKey Skeleton 

☐ MS-Ready Formula(e) 

☐ Exact Formula(e) 

☒ Monoisotopic Mass 

 Display All Chem

+/- ppm

Enter Identifiers to Search (searches should be limited to <5000 identifiers)

41.0265

56.02621

53.0265

58.0418

93.0578

113.9639

151.8754

69.9377

77.9872

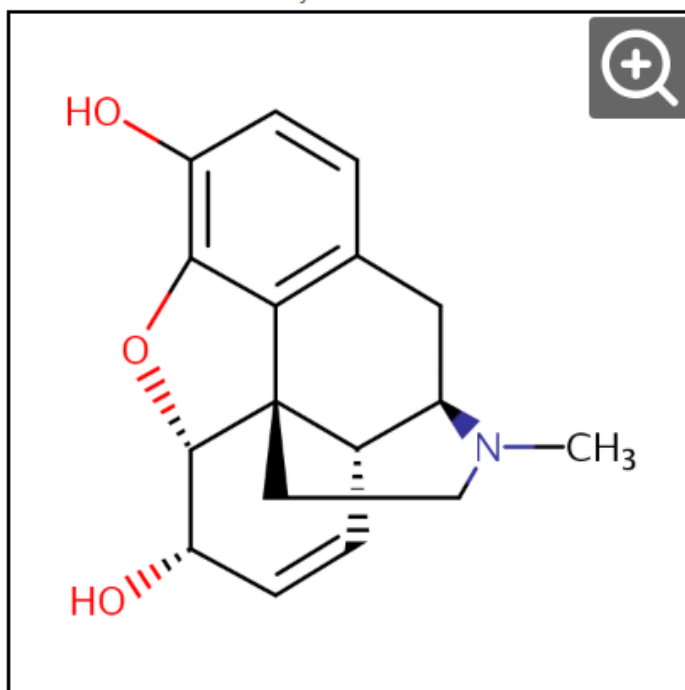
This search is based on what we refer to as "Mass Spec (MS) Ready" structures. All chemicals within the database are treated in a manner such that all are desalted, mixtures are separated, and stereochemistry is removed as Mass Spectrometry detects the major components of a salt or mixture and is insensitive to stereochemistry. As an example, a search for the monoisotopic mass of phenol will return phenol, sodium phenolate and calcium phenoxide. See the publication for more details: <https://doi.org/10.1186/s13321-018-0299-2>.

Searching batches using MS-Ready Formula (or mass) searching

	A	B	C	D	E	F	G
1	INPUT	DTXSID	CASRN	PREFERRED NAME	MOL FORMULA	MONOISOTOPIC MASS	DATA SOURCES
2	C14H22N2O3	DTXSID2022628	29122-68-7	Atenolol	C14H22N2O3	266.163042576	46
3	C14H22N2O3	DTXSID0021179	6673-35-4	Practolol	C14H22N2O3	266.163042576	32
4	C14H22N2O3	DTXSID4048854	841-73-6	Bucolome	C14H22N2O3	266.163042576	20
5	C14H22N2O3	DTXSID1045407	13171-25-0	Trimetazidine dihydrochloride	C14H24Cl2N2O3	338.116398	19
6	C14H22N2O3	DTXSID0045753	56715-13-0	R-(+)-Atenolol	C14H22N2O3	266.163042576	19
7	C14H22N2O3	DTXSID2048531	5011-34-7	Trimetazidine	C14H22N2O3	266.163042576	14
8	C14H22N2O3	DTXSID10239405	93379-54-5	Esatenolol	C14H22N2O3	266.163042576	12
9	C14H22N2O3	DTXSID50200634	52662-27-8	N-(2-Diethylaminoethyl)-2-(4-hydroxyphenoxy)acetamide	C14H22N2O3	266.163042576	7
10	C14H22N2O3	DTXSID4020111	51706-40-2	dl-Atenolol hydrochloride	C14H23ClN2O3	302.1397203	6
11	C14H22N2O3	DTXSID1068693	51963-82-7	Benzenamine, 2,5-diethoxy-4-(4-morpholinyl)-	C14H22N2O3	266.163042576	5
12	C18H34N2O6S	DTXSID3023215	154-21-2	Lincomycin	C18H34N2O6S	406.213757997	35
13	C18H34N2O6S	DTXSID7047803	859-18-7	Lincomycin hydrochloride	C18H35ClN2O6S	442.1904357	22
14	C18H34N2O6S	DTXSID20849438	1398534-62-7	PUBCHEM_71432748	C18H35ClN2O6S	442.1904357	1
15	C10H12N2O	DTXSID1047576	486-56-6	Cotinine	C10H12N2O	176.094963014	40
16	C10H12N2O	DTXSID8075330	50-67-9	Serotonin	C10H12N2O	176.094963014	22
17	C10H12N2O	DTXSID8044412	2654-57-1	4-Methyl-1-phenylpyrazolidin-3-one	C10H12N2O	176.094963014	18
18	C10H12N2O	DTXSID80165186	153-98-0	Serotonin hydrochloride	C10H13ClN2O	212.0716407	11
19	C10H12N2O	DTXSID2048870	29493-77-4	(4R,5S)-4-methyl-5-phenyl-4,5-dihydro-1,3-oxazol-2-amine	C10H12N2O	176.094963014	10
20	C10H12N2O	DTXSID10196105	443-31-2	6-Hydroxytryptamine	C10H12N2O	176.094963014	9
21	C10H12N2O	DTXSID90185693	31822-84-1	1,4,5,6-Tetrahydro-5-phenoxy pyrimidine	C10H12N2O	176.094963014	7
22	C10H12N2O	DTXSID40178777	2403-66-9	2-Benzimidazolepropanol	C10H12N2O	176.094963014	7
23	C10H12N2O	DTXSID80157026	13140-86-8	N-Cyclopropyl-N'-phenylurea	C10H12N2O	176.094963014	6
24	C10H12N2O	DTXSID30205607	570-14-9	4-Hydroxytryptamine	C10H12N2O	176.094963014	6
25	C14H18N4O3	DTXSID5023900	17804-35-2	Benomyl	C14H18N4O3	290.137890456	68
26	C14H18N4O3	DTXSID3023712	738-70-5	Trimethoprim	C14H18N4O3	290.137890456	51
27	C14H18N4O3	DTXSID40209671	60834-30-2	Trimethoprim hydrochloride	C14H19ClN4O3	326.1145682	8
28	C14H18N4O3	DTXSID70204210	55687-49-5	Benzenemethanol, 4-((2,4-diamino-5-pyrimidinyl)methyl)-2-	C14H18N4O3	290.137890456	5
29	C14H18N4O3	DTXSID20152671	120075-57-2	6-Methoxy-4-(3-(N,N-dimethylamino)propylamino)-5,8-quin	C14H18N4O3	290.137890456	4
30	C14H18N4O3	DTXSID30213742	63931-79-3	1H-1,2,4-Benzotriazepine-3-carboxylic acid, 4,5-dihydro-4-	C14H18N4O3	290.137890456	3
31	C14H18N4O3	DTXSID30219608	69449-07-6	2,4-Pyrimidinediamine, 5-((3,4,5-trimethoxyphenyl)methyl)-	C14H20N4O4	308.14845514	3
32	C14H18N4O3	DTXSID20241155	94232-27-6	L-Aspartic acid, compound with 5-((3,4,5-trimethoxyphenyl	C18H25N5O7	423.175398165	3
33	C14H18N4O3	DTXSID80241156	94232-28-7	L-Glutamic acid, compound with 5-((3,4,5-trimethoxyphenyl	C19H27N5O7	437.191048229	3
34	C14H18N4O3	DTXSID20143781	101204-93-7	1H-Pyrido(2,3-e)-1,4-diazepine-2,3,5-trione, 4-(2-(diethylam	C14H18N4O3	290.137890456	3
35	C12H11N7	DTXSID6021373	396-01-0	Triamterene	C12H11N7	253.107593382	52
36	C12H11N7	DTXSID00204465	5587-93-9	Ampyrimine	C12H11N7	253.107593382	7
37	C12H11N7	DTXSID5064621	7300-26-7	Benzenamine, 4-azido-N-(4-azidophenyl)-	C12H9N7	251.091943318	4
38	C12H11N7	DTXSID00848025	90293-82-6	Sulfuric acid-6-phenylpteridine-2,4,7-triamine (1/1)	C12H13N7O4S	351.074973101	1
39	C12H11N7	DTXSID50575293	92310-83-3	(1E)-N-Phenyl-1,2-bis(1H-1,2,4-triazol-1-yl)ethan-1-imine	C12H11N7	253.107593382	1
40	C8H9NO2	DTXSID2020006	103-90-2	Acetaminophen	C8H9NO2	151.063328534	75
41	C8H9NO2	DTXSID6025567	134-20-3	Methyl 2-aminobenzoate	C8H9NO2	151.063328534	50

Additional Mass-Spec Functionality

Formula-Based Search



Wikipedia

Morphine is a pain medication of the opiate family which is found naturally in a number of plants and animals. It acts directly on the central nervous system (CNS) to decrease the feeling of pain. It can be taken for both acute pain and chronic pain. It is frequently used for pain from myocardial infarction and during labor. It can be given by mouth, by injection into a muscle, by injection under the skin, intravenously, injection into the space around the

...

[Read more](#)

Intrinsic Properties



Molecular Formula: $C_{17}H_{19}NO_3$

Mol File

Find All Chemicals



Average Mass: 285.343 g/mol

Isotope Mass Distribution



Monoisotopic Mass: 285.136493 g/mol

Structural Identifiers

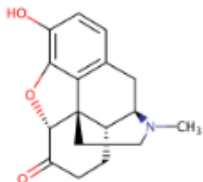
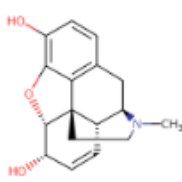
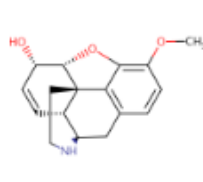
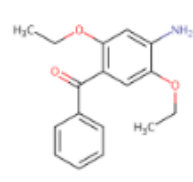
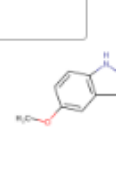
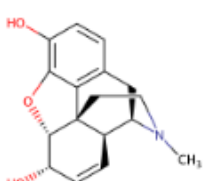
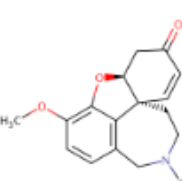
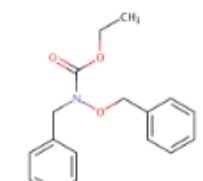
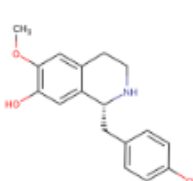
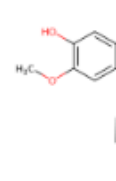
Select Chemicals of Interest

Searched by Exact Molecular Formula: C17H19NO3

6 of 120 chemicals selected

Search Default

Hide chemicals that are:

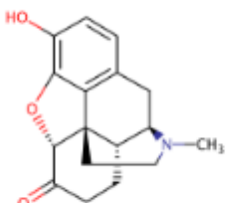
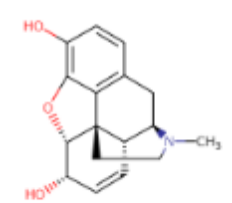
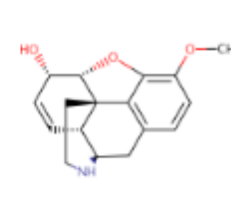
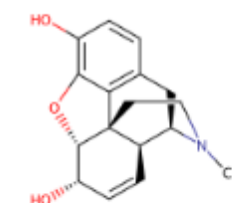
 <p>Hydromorphone DTXSID: DTXSID8023133 PubChem: 39 CPDAT: 21</p>	 <p>Morphine DTXSID: DTXSID9023336 PubChem: 52 CPDAT: 37</p>	 <p>Norcodeine DTXSID: DTXSID8046327 PubChem: 20 CPDAT: 2</p>	 <p>Methanone, (4-amino-2,5-diethoxyphenyl)- DTXSID: DTXSID3071696 PubChem: 16 CPDAT: 0</p>	 <p>Carbazone DTXSID: DTXSID... PubChem: 1 CPDAT: 0</p>
 <p>Morphinan-3,6-diol, 7,8-didehydro-4,5-epoxy- DTXSID: DTXSID40167308 PubChem: 8 CPDAT: 0</p>	 <p>Narwedine, (+/-)- DTXSID: DTXSID60168190 PubChem: 43 CPDAT: 0</p>	 <p>Carbamic acid, (phenylmethoxy)(phenyl)- DTXSID: DTXSID50171931 PubChem: 11 CPDAT: 0</p>	 <p>7-Isoquinolinol, 1,2,3,4-tetrahydro-1-((4-hydroxyphenyl)methyl)- DTXSID: DTXSID70176367 PubChem: 21 CPDAT: 0</p>	 <p>Cherylin DTXSID: DTXSID... PubChem: 1 CPDAT: 0</p>

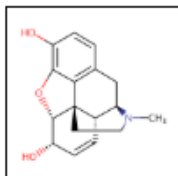
Prune to list of interest

6 of 120 chemicals selected

Deselect all Download Send to Batch Search Default

DTXSID PubChem CPDAT

 <p>Hydromorphone DTXSID: DTXSID8023133 PubChem: 39 CPDAT: 21</p>	 <p>Morphine DTXSID: DTXSID9023336 PubChem: 52 CPDAT: 37</p>	 <p>Norcodeine DTXSID: DTXSID8046327 PubChem: 20 CPDAT: 2</p>	 <p>Morphinan-3,6-diol, 7,8-didehydro-4,5-... DTXSID: DTXSID40167308 PubChem: 8 CPDAT: 0</p>
--	---	--	---




Morphine

57-27-2 | DTXSID9023336

Searched by Approved Name.

Abstract Sifter

1) Select PubMed starting point query then 2) click on Retrieve. 

Select a Query Term



Retrieve Articles

Select a Query Term

- Hazard
- Fate and Transport
- Metabolism/PK/PD
- Chemical Properties
- Exposure
- Mixtures
- Male Reproduction
- Androgen Disruption
- Female Reproduction
- GeneTox
- Cancer
- Clinical Trials
- Embryo and embryonic development
- Child (infant through adolescent)
- Dust and Exposure
- Food and Exposure
- Water and Exposure
- Algae
- Disaster / Emergency

Optionally, edit the query before retrieving.

"57-27-2" OR "Morphine"

Literature Searching

Child (infant through adolescent)

Dust and Exposure


Food and Exposure

Water and Exposure

Algae

Disaster / Emergency



 Optionally, edit the query before retrieving.

("57-27-2" OR "Morphine") AND ((water OR groundwater OR drinking water) AND Environmental Exposure)



Work in Progress

- Registering and curating numerous lists
 - NIST library of chemicals
 - United States Geological Survey chemicals in water
 - Scientific Working Group for the Analysis of Seized Drugs
 - Synthetic Cannabinoids
 - Mycotoxins

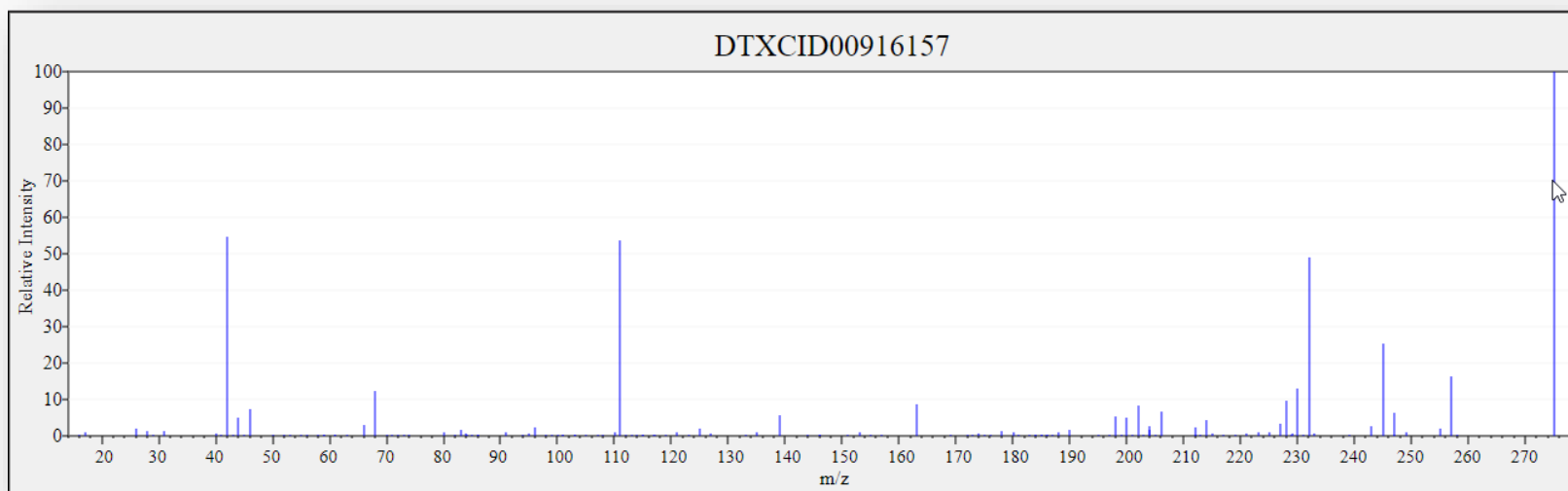
- CFM-ID
 - Viewing and Downloading pre-predicted spectra
 - Search spectra against the database
- Structure/substructure/similarity search
- Access to API and web services

Predicted Mass Spectra

<http://cfmid.wishartlab.com/>



- MS/MS spectra prediction for ESI+, ESI-, and EI
- Predictions generated and stored for >800,000 structures, to be accessible via Dashboard



Search Expt. vs. Predicted Spectra



Non Target Analysis Prototype

Mass Search

±

Min/Max

321.136493476

Da

±

0.0000002

Da

ppm

Molecular Formula Search

Molecular Formula

Mass or Formula must be entered before searching spectrum

Ionization Type

ESI+ ▼

ESI+

ESI-

EI

Spectra Input

Single Energy

Multiple

304.1332052 11.6199475
198.0913404 7.308439899
123.0440559 6.538348292
196.0756904 5.269463115
216.1019051 4.700461978
200.1080005 4.800144384

Peak Match Window:


0.02

Da

ppm

Search

Search Expt. vs. Predicted Spectra



United States
Environmental Protection
Agency

Home Advanced Search Batch Search Lists Predictions Downloads

Share Search all data

Spectra Input

Single Energy

304.1332052 11.61

198.0913404 7.30

123.0440559 6.53

196.0756904 5.28

216.1019051 4.70

200.1080005 4.80

Peak Match

Search

TSV CSV Excel

Chemical Structure ID

DTXCID101048191

DTXCID101181567

DTXCID50879086

DTXCID60686349

DTXCID00830900

DTXCID10971176

DTXCID60301242

DTXCID40703048

DTXCID60349982

DTXCID10316649

Showing 1 to 10 of 38 entries

Chemical Structure ID	Score (10eV)
DTXCID101048191	0.22
DTXCID101181567	0.19
DTXCID50879086	0.17
DTXCID60686349	0.14
DTXCID00830900	0.13
DTXCID10971176	0.12
DTXCID60301242	0.12
DTXCID40703048	0.11
DTXCID60349982	0.11
DTXCID10316649	0.09

1

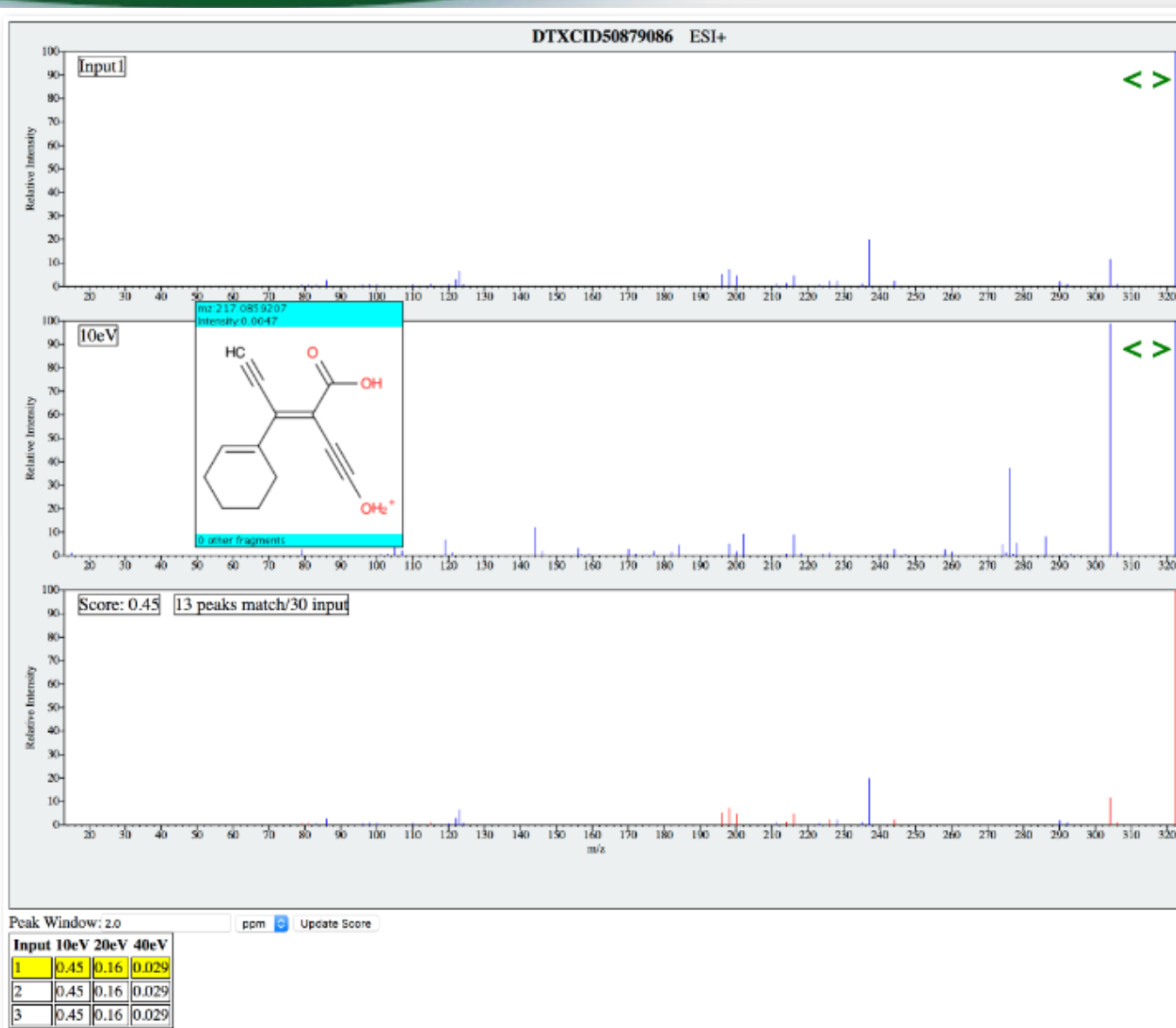
2

3

4

Next

Spectral Viewer Comparison



Non-Targeted Screening of Wastewater for Water Reuse using Mass Spectrometry

Prototype development using the US-EPA CompTox Chemicals Dashboard data and CFM-ID Fragment Prediction algorithms

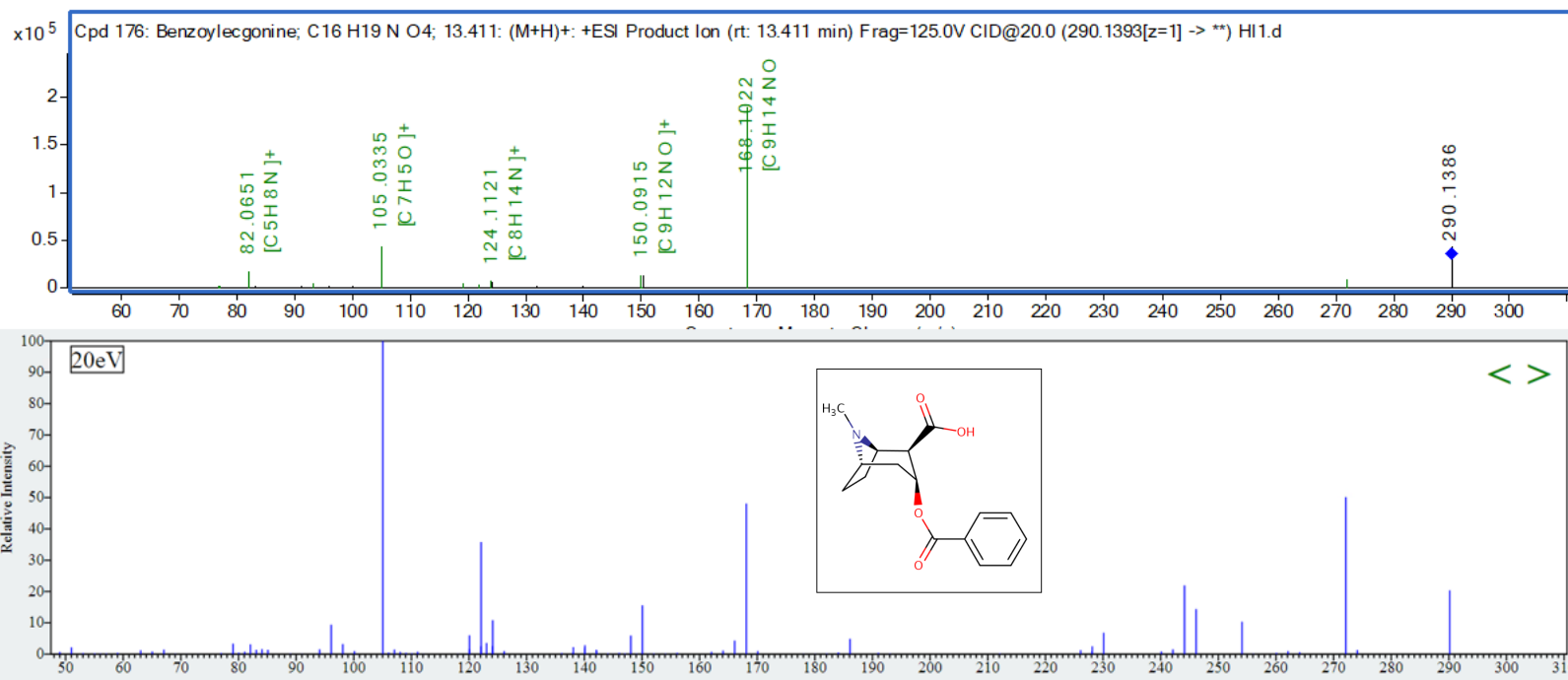
Jerry Zweigenbaum, Agilent, Wilmington DE, **Andrew McEachran**, Agilent, Santa Clara CA, **Alex Chao**, Oak Ridge Associated Universities (ORAU), US EPA, National Exposure Research Laboratory, and **Antony J. Williams**, National Center for Computational Toxicology, US EPA Research Triangle Park NC

ACS Fall 2019 National Meeting & Exposition in San Diego



Example match



MS/MS of benzoylecgonine




Predicted Data Already Public *Publication and Data Files*

Data Descriptor | OPEN | Published: 02 August 2019

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) | [Download Citation](#) 

CFM-ID Paper Data

[Dataset](#) posted on 01.03.2019, 08:38 by EPA's National Center for Computational Toxicology

88
views

17
downloads

0
citations

This upload is a zip containing the following files:

Predicted EI-MS Spectra of CompTox Chemicals Dashboard Structures:

Predicted EI-MS spectra of ~700,000 chemical structures from the CompTox Chemicals Dashboard were generated using the CFM-ID model developed by Allen, et al. (<https://doi.org/10.1021/acs.analchem.6b01622>). These data are provided in .dat ASCII format.

Predicted MS/MS Spectra in ESI-positive mode of CompTox Chemicals Dashboard Structures:

Predicted MS/MS spectra of ~700,000 chemical structures from the CompTox Chemicals Dashboard were generated using the CFM-ID model developed by Allen, et al. (<https://doi.org/10.1007/s11306-014-0676-4>) in ESI-positive mode. These data are provided in .dat ASCII format.

Predicted MS/MS Spectra in ESI-negative mode of CompTox Chemicals Dashboard Structures:

Predicted MS/MS spectra of ~700,000 chemical structures from the CompTox Chemicals Dashboard were generated using the CFM-ID model developed by Allen, et al. (<https://doi.org/10.1007/s11306-014-0676-4>) in ESI-negative mode. These data are provided in .dat ASCII format.



CATEGORIES

• Toxicology

KEYWORD(S)

Computational Toxicology

DSSTox Chemical Database

Chemicals Dashboard

Non-targeted analysis

CFM-ID

LICENCE



EXPORT

RefWorks

BibTeX

Ref. manager

Endnote

https://epa.figshare.com/articles/CFM-ID_Paper_Data/7776212/1

Prototype Development

AADashboard

atrazine

Search



Select properties to predict

H

T.E.S.T. 18

OPERA

Search

C

Exact

N

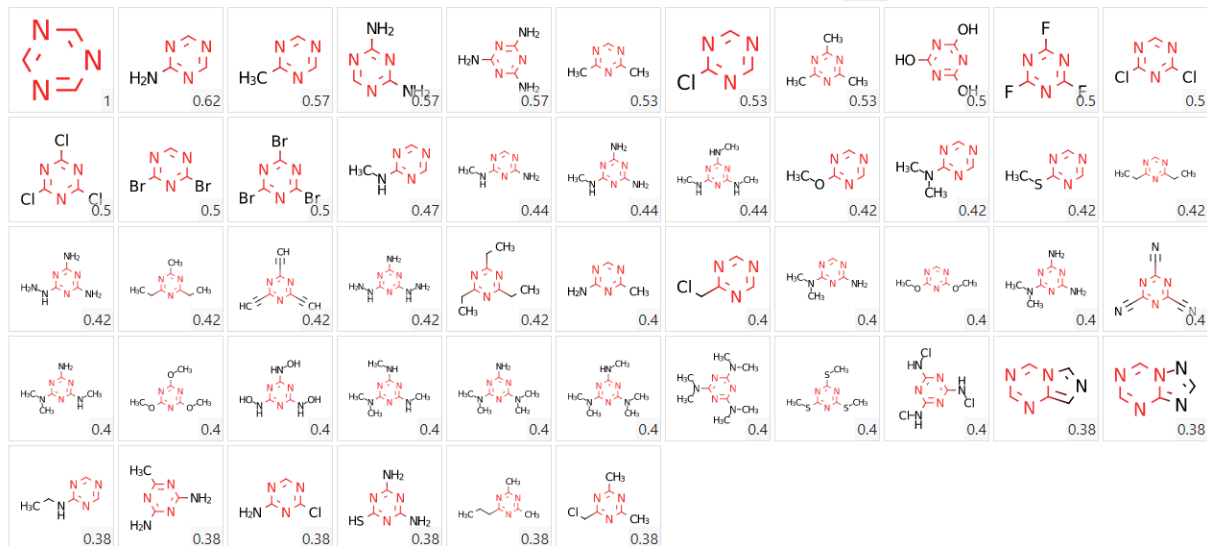
Substructure

O

Search result 2540

Show ☐ Isotopically Labeled ☐ Charged ☐ Salts or Mixtures

Sort Similarity



Search result 2540

Show ☐ Isotopically Labeled

Prototype Development

atrazine Search

100%

Select properties to predict

H T.E.S.T. 18 OPERA Search

C

N

O

S

P

F

Cl Input formula (e.g. C6 H6):

Br C15H16O2

Search

Search result 5 Show ☐ Isotopically Labeled ☐ Ch

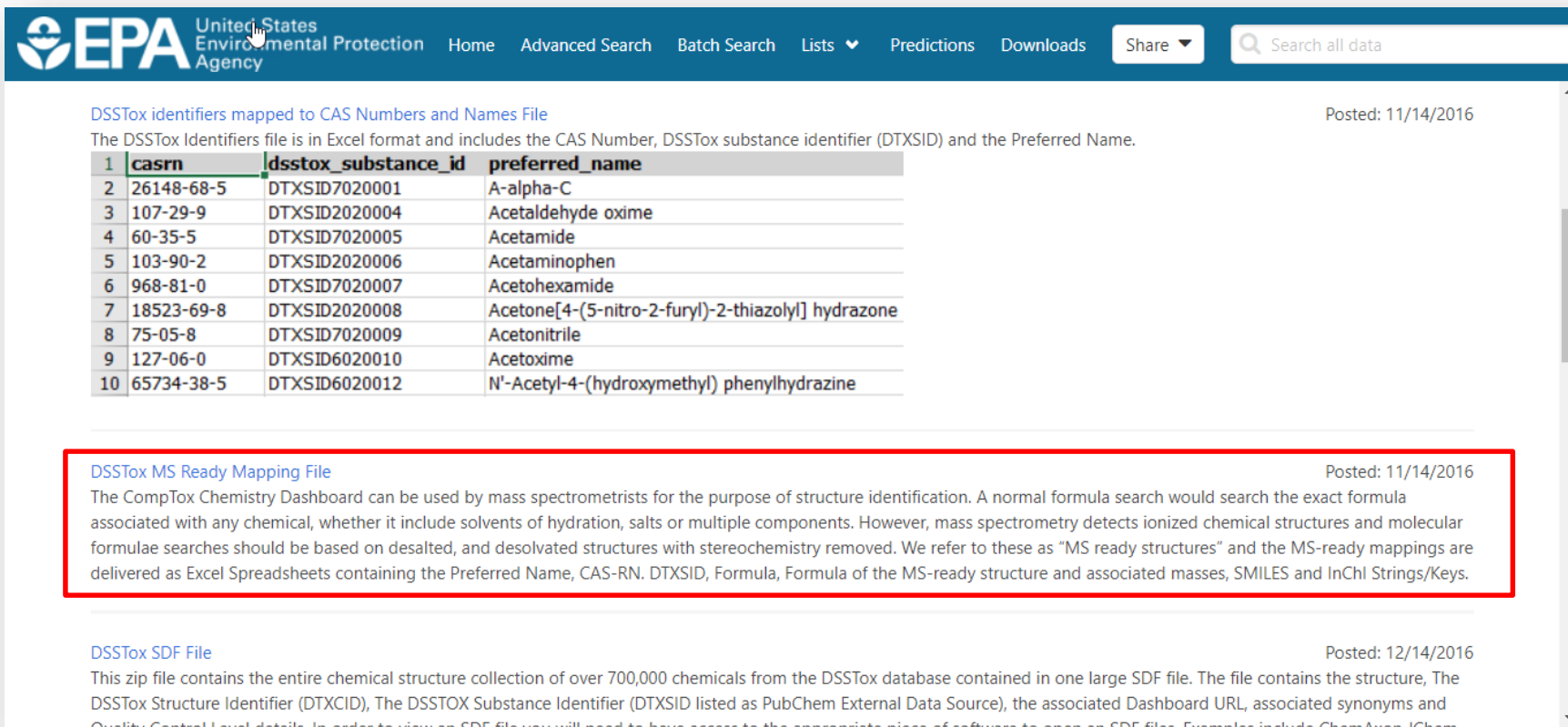
Elements per page 50

1

<https://comptox.epa.gov/dashboard/DTXSID7020182>

API services and Open Data

- Web Services <https://actorws.epa.gov/actorws/>
- Data sets also available for download..



The screenshot shows the EPA website's navigation bar with links for Home, Advanced Search, Batch Search, Lists, Predictions, Downloads, and a Share button. A search bar is also present. Below the navigation bar, there are two main sections. The first section, titled "DSSTox identifiers mapped to CAS Numbers and Names File", includes a table with 10 rows of data. The second section, titled "DSSTox MS Ready Mapping File", is highlighted with a red border and contains a detailed description of the mapping file. A third section, titled "DSSTox SDF File", is partially visible at the bottom.

DSSTox identifiers mapped to CAS Numbers and Names File Posted: 11/14/2016

The DSSTox Identifiers file is in Excel format and includes the CAS Number, DSSTox substance identifier (DTXSID) and the Preferred Name.

	casrn	dsstox_substance_id	preferred_name
2	26148-68-5	DTXSID7020001	A-alpha-C
3	107-29-9	DTXSID2020004	Acetaldehyde oxime
4	60-35-5	DTXSID7020005	Acetamide
5	103-90-2	DTXSID2020006	Acetaminophen
6	968-81-0	DTXSID7020007	Acetohexamide
7	18523-69-8	DTXSID2020008	Acetone[4-(5-nitro-2-furyl)-2-thiazolyl] hydrazone
8	75-05-8	DTXSID7020009	Acetonitrile
9	127-06-0	DTXSID6020010	Acetoxime
10	65734-38-5	DTXSID6020012	N'-Acetyl-4-(hydroxymethyl) phenylhydrazine

DSSTox MS Ready Mapping File Posted: 11/14/2016

The CompTox Chemistry Dashboard can be used by mass spectrometrists for the purpose of structure identification. A normal formula search would search the exact formula associated with any chemical, whether it include solvents of hydration, salts or multiple components. However, mass spectrometry detects ionized chemical structures and molecular formulae searches should be based on desalted, and desolvated structures with stereochemistry removed. We refer to these as "MS ready structures" and the MS-ready mappings are delivered as Excel Spreadsheets containing the Preferred Name, CAS-RN, DTXSID, Formula, Formula of the MS-ready structure and associated masses, SMILES and InChI Strings/Keys.

DSSTox SDF File Posted: 12/14/2016

This zip file contains the entire chemical structure collection of over 700,000 chemicals from the DSSTox database contained in one large SDF file. The file contains the structure, The DSSTox Structure Identifier (DTXCID), The DSSTOX Substance Identifier (DTXSID listed as PubChem External Data Source), the associated Dashboard URL, associated synonyms and Quality Control Level details. In order to view an SDF file you will need to have access to the appropriate piece of software to open an SDF file. Examples include ChemAxon JChem

- Data in UI, JSON and XML format

<https://actorws.epa.gov/actorws/dsstox/v02/msready?identifier=80-05-7>

<https://actorws.epa.gov/actorws/dsstox/v02/msready.json?identifier=80-05-7>

<https://actorws.epa.gov/actorws/dsstox/v02/msready.xml?identifier=80-05-7>

<https://actorws.epa.gov/actorws/dsstox/v02/msready?identifier=DTXCID60513>

<https://actorws.epa.gov/actorws/dsstox/v02/msready.json?identifier=DTXCID60513>

<https://actorws.epa.gov/actorws/dsstox/v02/msready.xml?identifier=DTXCID60513>

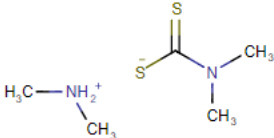
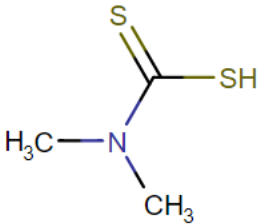
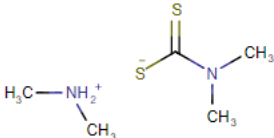
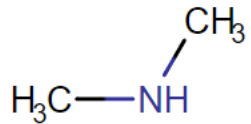
<https://actorws.epa.gov/actorws/dsstox/v02/msready?identifier=UVOFGKIRTCCKG-UHFFFAOYSA-N>

<https://actorws.epa.gov/actorws/dsstox/v02/msready.json?identifier=UVOFGKIRTCCKG-UHFFFAOYSA-N>

<https://actorws.epa.gov/actorws/dsstox/v02/msready.xml?identifier=UVOFGKIRTCCKG-UHFFFAOYSA-N>

InChIKey to DTXCIDs

<https://actorws.epa.gov/actorws/dsstox/v02/msready?identifier=UVOFGKIRTCCNKG-UHFFFAOYSA-N>

Image	DTXCID	Smiles	Image	MsReady DTXCID	MsReady SMILES
	DTXCID60513	<chem>C[NH2+]C.CN(C)C([S-])=S</chem>		DTXCID0023797	<chem>CN(C)C(S)=S</chem>
	DTXCID60513	<chem>C[NH2+]C.CN(C)C([S-])=S</chem>		DTXCID704057	<chem>CNC</chem>

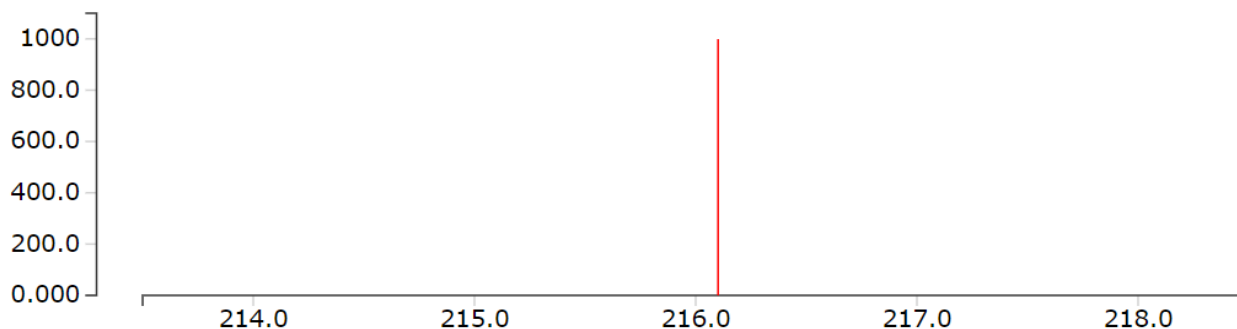
MassBank mapping to Dashboard

MassBank Record: EA028808

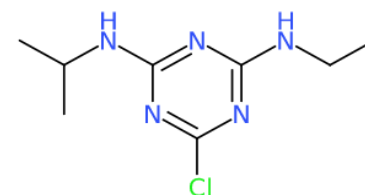
[Home](#) | [Search](#) | [Record Index](#) | [Data Privacy](#) | [Imprint](#) | MassBank ID:

Atrazine; LC-ESI-ITFT; MS2; CE: 15%; R=15000; [M+H]⁺

Mass Spectrum



Chemical Structure



CH\$NAME: Atrazine

CH\$NAME: 6-chloro-N-ethyl-N'-isopropyl-1,3,5-triazine-2,4-diamine

CH\$NAME: 6-chloranyl-N4-ethyl-N2-propan-2-yl-1,3,5-triazine-2,4-diamine

CH\$COMPOUND_CLASS: N/A; Environmental Standard

CH\$FORMULA: [C8H14ClN5](#)

CH\$EXACT_MASS: 215.0932

CH\$SMILES: c1(nc(nc(n1)Cl)NCC)NC(C)C

CH\$IUPAC: InChI=1S/C8H14ClN5/c1-4-10-7-12-6(9)13-8(14-7)11-5(2)3/h5H,4H2,1-3H3,(H2,10,11,12,13,14)

CH\$LINK: CAS [1912-24-9](#)

CH\$LINK: CHEBI [15930](#)

CH\$LINK: KEGG [C06551](#)

CH\$LINK: PUBCHEM [CID:2256](#)

CH\$LINK: INCHIKEY [MXWJVTOOROXGIU-UHFFFAOYSA-N](#)

CH\$LINK: CHEMSPIDER [2162](#)

CH\$LINK: COMPTOX [DTXSID9020112](#)

***When data is
openly shared...***

NORMAN Suspect List Exchange

<https://www.norman-network.com/?q=node/236>

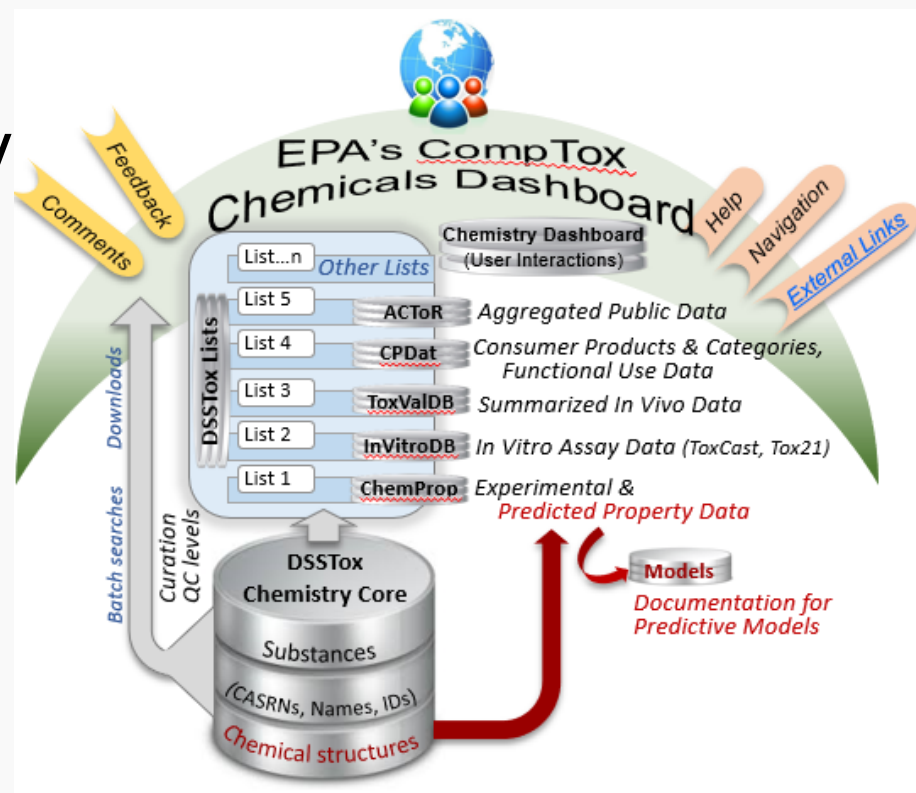
NORMAN

Network of reference laboratories, research centres and related organisations for monitoring of emerging environmental substances

Wastewater Suspect List based on Swedish Product Data	Wastewater Suspect List Original File with Mapped DTXSIDs (12/02/2019)	KEMIWSUS InChIKeys (12/02/2019)	A prioritized list of 1,123 substances relevant for wastewater based on Swedish product registry data, including scores. Provided by Stellan Fischer, KEMI.
Algal toxins list from CompTox	ALGALTOX XLSX, CSV (14/02/2019) CompTox ALGALTOX List	ALGALTOX InChIKeys (14/02/2019)	List of algal toxins (generated during blooms) from the CompTox Chemicals Dashboard.
CCL 4 Chemical Candidate List	CCL4 XLSX, CSV (14/02/2019) CompTox CCL4 List	CCL4 InChIKeys (14/02/2019)	Contaminants that are not (yet) regulated in the USA but are known or anticipated to occur in public water systems; from CompTox.
Hydrogen Deuterium Exchange (HDX) Standard Set	HDXNOEX XLSX, CSV (14/02/2019) CompTox HDXNOEX List CompTox HDXEXCH List	HDXNOEX InChIKeys (14/02/2019)	Environmental standard set used to investigate hydrogen deuterium exchange in small molecule HRMS (Ruttkies et al. submitted). HDXEXCH list also contains observed deuterated species.
Neurotoxicants Collection from Public Resources	NEUROTOXINS XLSX, CSV (14/02/2019) CompTox NEUROTOXINS List	NEUROTOXINS InChIKeys (14/02/2019)	A list of neurotoxicants compiled from public resources, details on CompTox and Schymanski <i>et al.</i> (submitted).
Statins Collection from Public Resources	STATINS XLSX, CSV (14/02/2019) CompTox STATINS List	STATINS InChIKeys (14/02/2019)	A list of statins (lipid-lowering medications) compiled from public resources, details on CompTox.
Synthetic Cannabinoids and Psychoactive Compounds	SYNTHCANNAB XLSX, CSV (14/02/2019) CompTox SYNTHCANNAB List	SYNTHCANNAB InChIKeys (14/02/2019)	A list of synthetic cannabinoids and psychoactive compounds assembled from public resources, from CompTox.

Conclusion

- Dashboard access to data for ~875,000 chemicals
- MS-Ready data facilitates structure identification
- Related metadata facilitates candidate ranking
- Relationship mappings and chemical lists of great utility
- Curation and mutual sharing of chemical lists is important (e.g. NORMAN)



Acknowledgements



Credit: the Research Triangle Foundation

- *An enormous team of contributors from NCCT, especially the IT software development team*
- *Our curation team for their care and focus on data quality*
- *Many public domain databases and open data contributors*
- *Emma Schymanski, University of Luxembourg for coordinating curation of NORMAN lists*
- *Jerry Zeigenbaum and Andrew McEachran, Agilent Inc., for example spectra to test CFM-ID matching*

Antony Williams

NCCT, US EPA Office of Research and Development,

Williams.Antony@epa.gov

ORCID: <https://orcid.org/0000-0002-2668-4821>

Williams et al. *J Cheminform* (2017) 9:61
DOI 10.1186/s13321-017-0247-6


 Journal of Cheminformatics

DATABASE

Open Access

The CompTox Chemistry Dashboard: a community data resource for environmental chemistry



Antony J. Williams^{1*} , Christopher M. Grulke¹, Jeff Edwards¹, Andrew D. McEachran², Kamel Mansouri^{1,2,4}, Nancy C. Baker³, Grace Patlewicz¹, Imran Shah¹, John F. Wambaugh¹, Richard S. Judson¹ and Ann M. Richard¹

<https://doi.org/10.1186/s13321-017-0247-6>