CompTox Chemistry Dashboard

Antony Williams

US EPA Office of Research and Development

EPA Tools and Resources Webinar

February 15, 2017

Problem

- Tens of thousands of chemicals are of interest to a broad spectrum of stakeholders and the Agency.
- The National Center for Computational Toxicology is looking to integrate and share our data to support computational toxicology
- Chemical structures and data are required to develop prediction models
- Data should be accessible and *Open* for reuse

Action

- Develop a publicly accessible web-based application
- Provide access to integrated data for ~740,000 chemicals – structures, properties, models, links to other agency resources
- Deliver search results to support different research needs

Impact

- Data are being used to:
  - Access toxicity data
  - Build prediction models
  - Speed chemical identification
  - Support additional software applications

- Open data allows for:
  - Integration via other public websites
  - A phone and tablet mobile application

Result

- Public access to over a decade of curation work
- Regions and states have open access to the data
- The data are being used to support Agency projects (e.g.):
  - Tire crumb
  - Hydraulic fracturing
  - Endocrine Disruption
  - Non-targeted Analysis in analytical science

Who is NCCT?

- National Center for Computational Toxicology – part of EPA's Office of Research and Development
- Research driven by EPA's *Chemical Safety for Sustainability Research Program*
  - Develop new approaches to **evaluate** the **safety** of chemicals
  - Integrate advances in biology, biotechnology, chemistry, exposure science and computer science
- Goal - To identify **chemical exposures** that may disrupt biological processes and cause adverse outcomes.

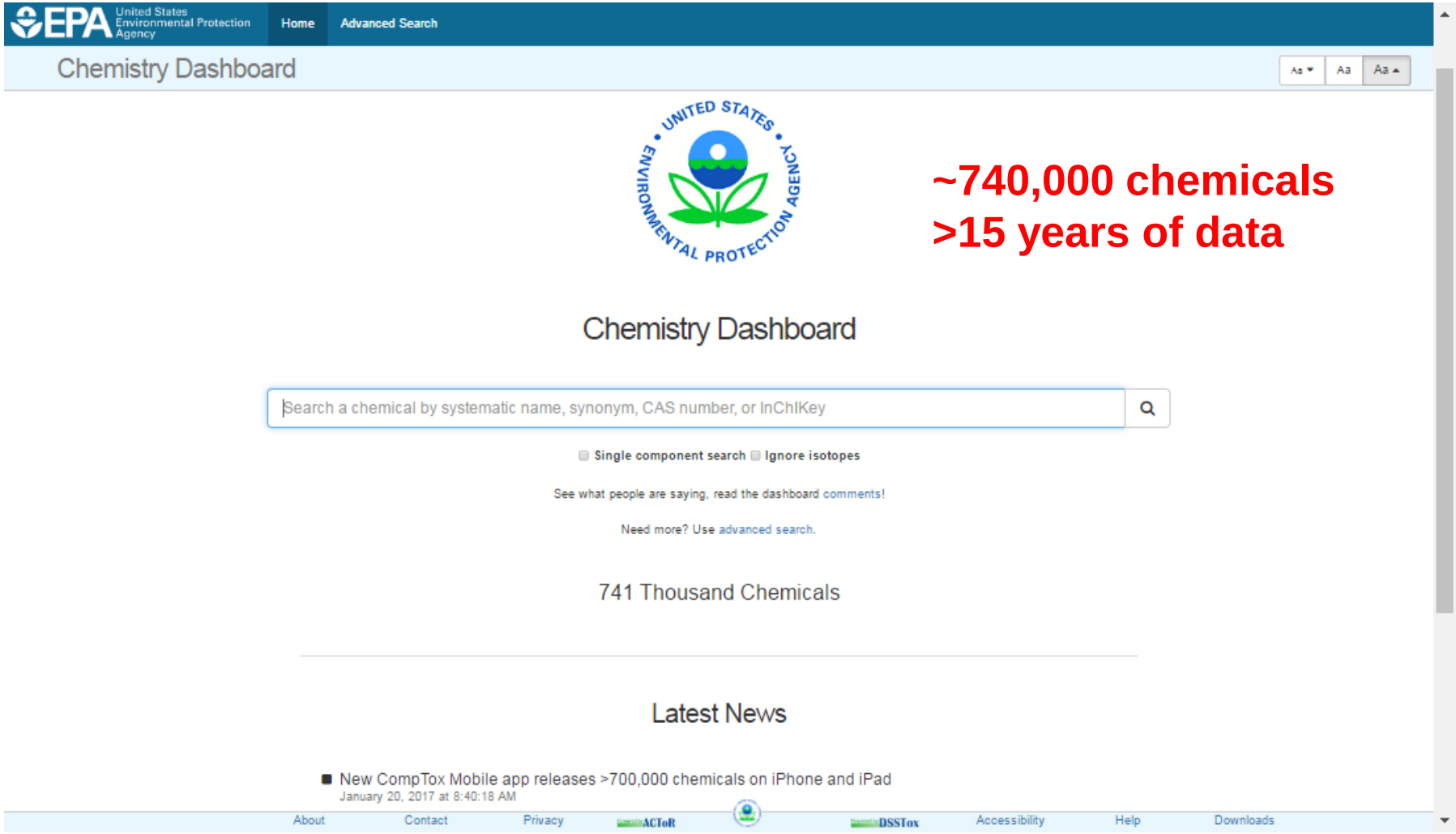
The CompTox Chemistry Dashboard – An Overview

- A **publicly accessible website** delivering access:
  - ~740,000 chemicals and related property data
  - Links to other agency websites and public data resources
  - "Literature" searches for chemicals using public resources
  - Integration to "biological assay data" for 1000s of chemicals
  - Information regarding consumer products containing chemicals
  - "Batch searching" supports searches for thousands of chemicals

- An underlying architecture for the development of future software applications

Comptox Chemistry Dashboard

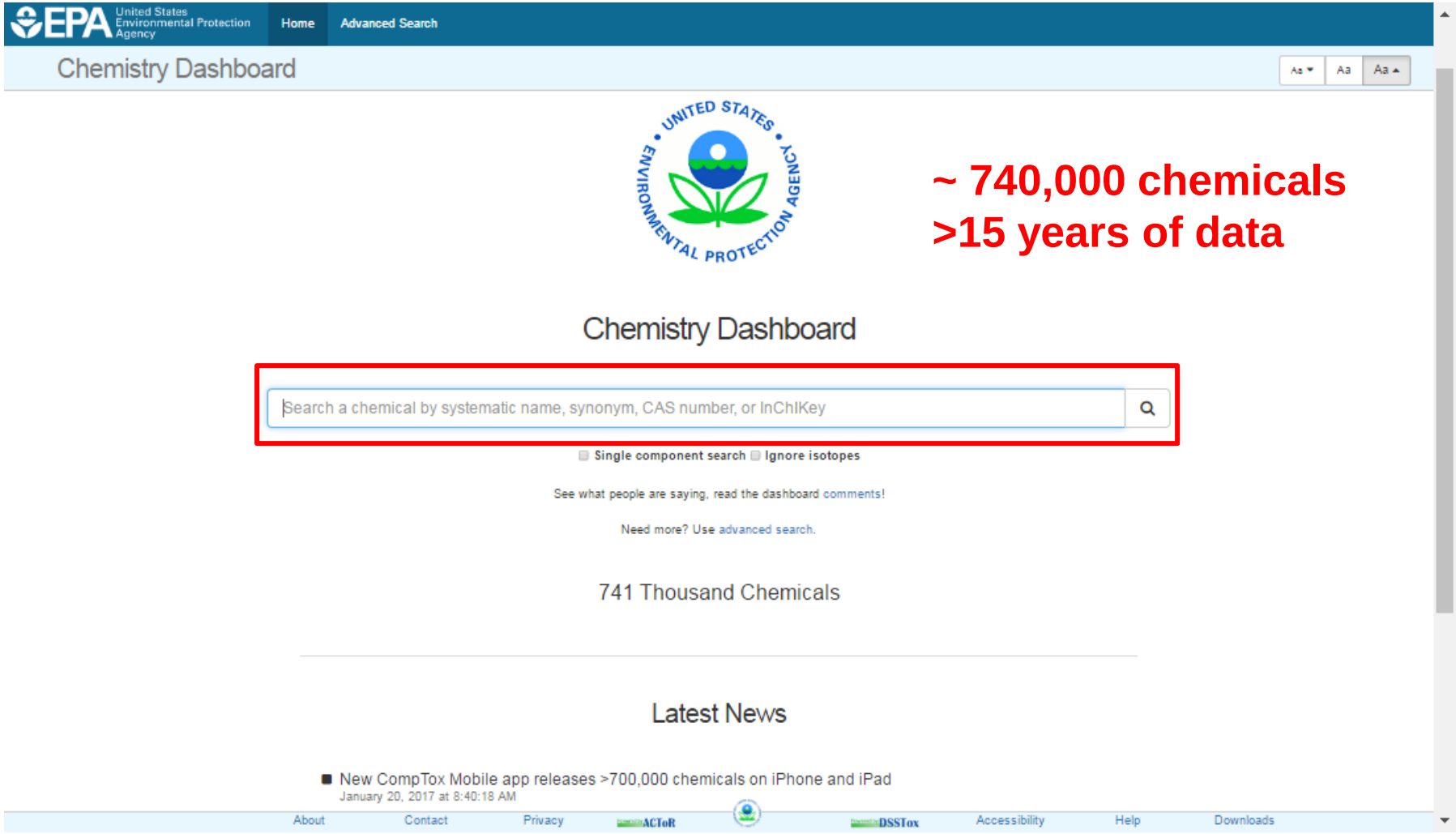
https://comptox.epa.gov



~740,000 chemicals
>15 years of data

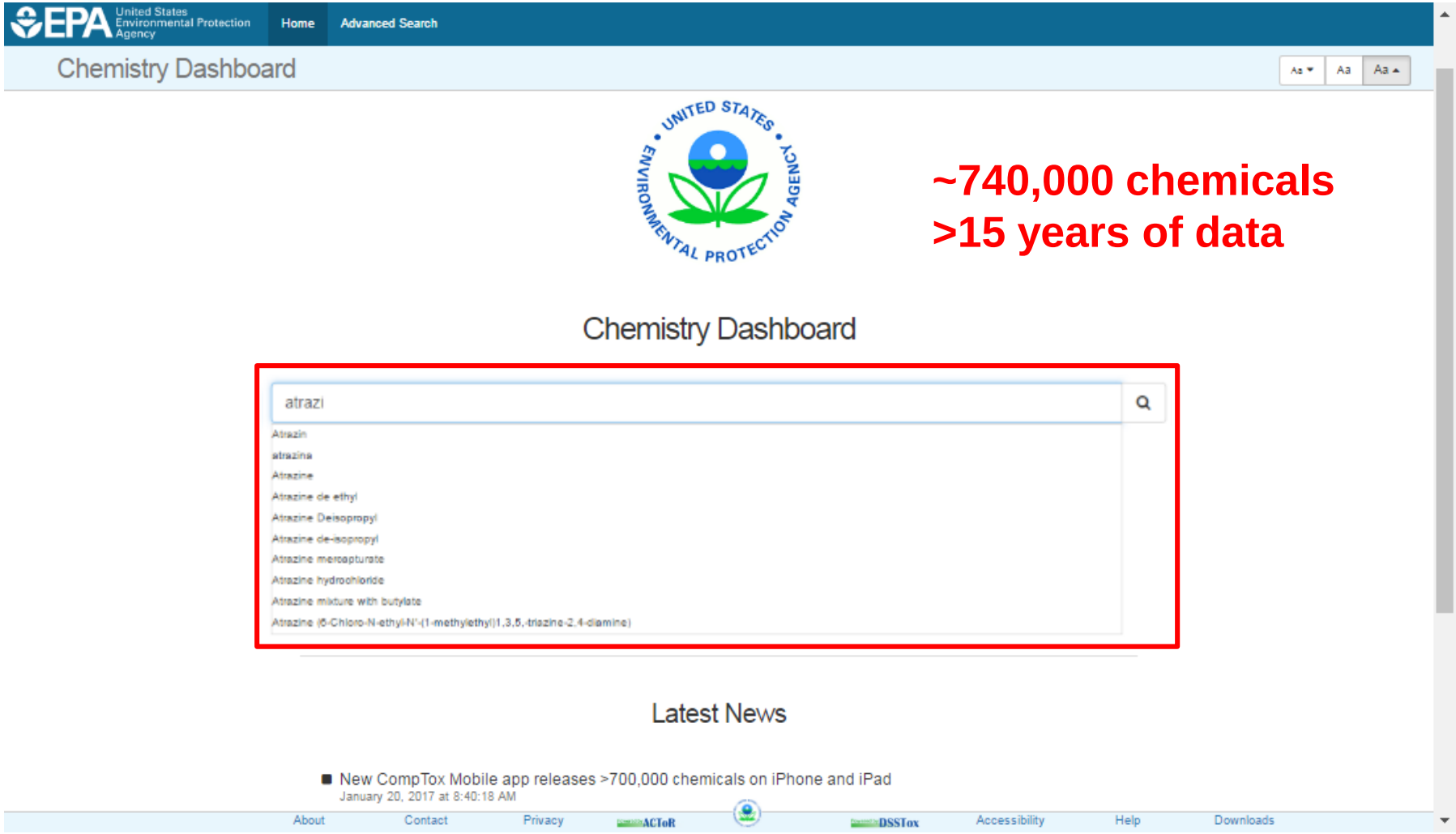
Comptox Chemistry Dashboard

https://comptox.epa.gov



Comptox Chemistry Dashboard

https://comptox.epa.gov



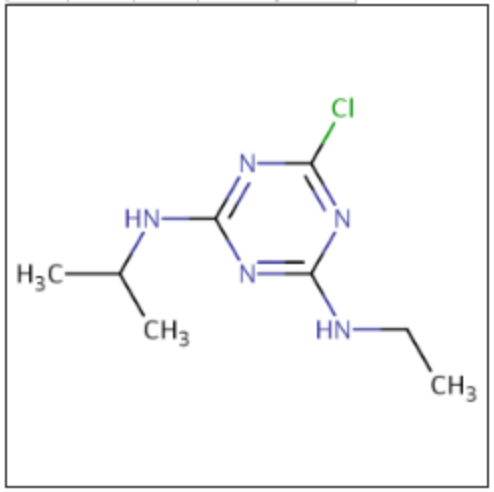
Chemical Page

- Atrazine is an herbicide of the triazine class. Atrazine is used to prevent pre- and post emergent broadleaf weeds.

Atrazine

1912-24-9 | DTXSID9020112

Searched by Approved Name: Found 1 result for 'atrazine'.



Wikipedia

Intrinsic Properties

Molecular Formula: C8H14ClN5 — Find All Chemicals

Average Mass: 215.69 g/mol

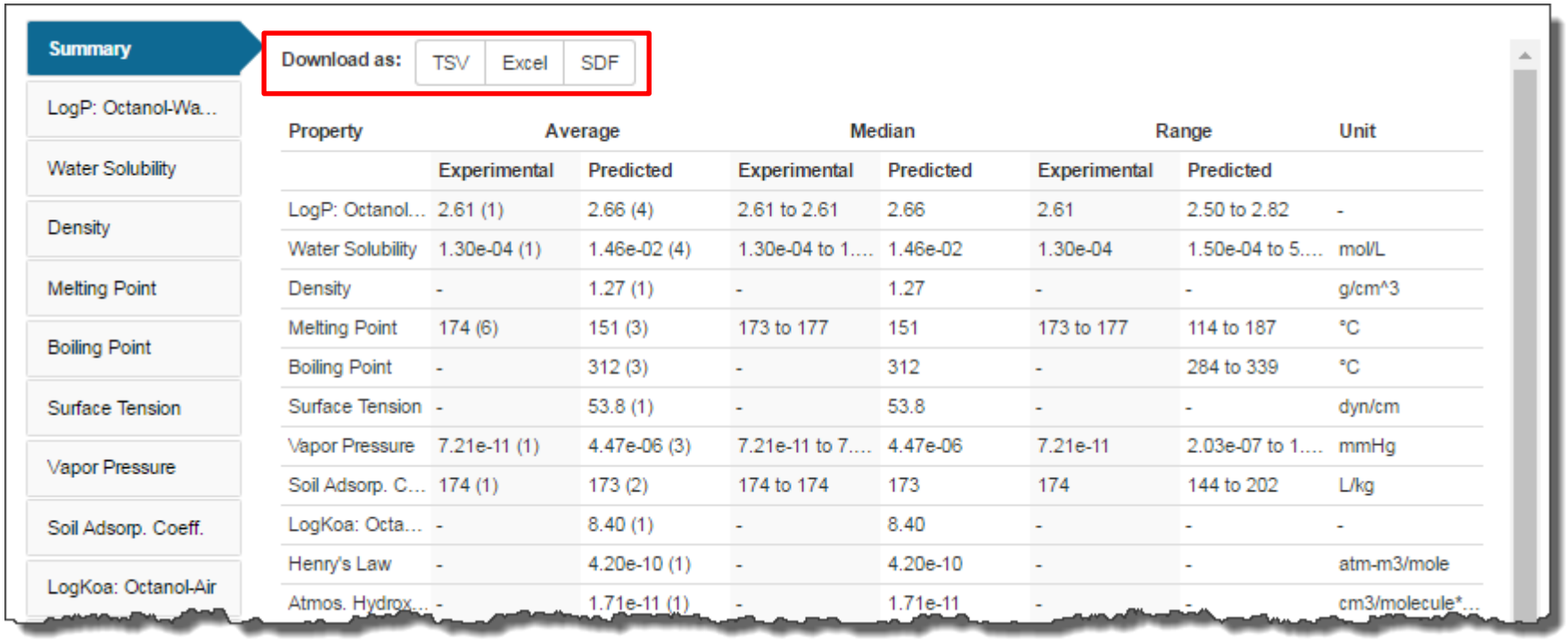
Monoisotopic Mass: 215.093773 g/mol

Structural Identifiers

Record Information

Chemical Properties | Synonyms | External Links | Product Composition | Bioassays | Exposure | Analytical | Literature | Comments

Physicochemical Properties



Summary

LogP: Octanol-Wa...

Water Solubility

Density

Melting Point

Boiling Point

Surface Tension

Vapor Pressure

Soil Adsorp. Coeff.

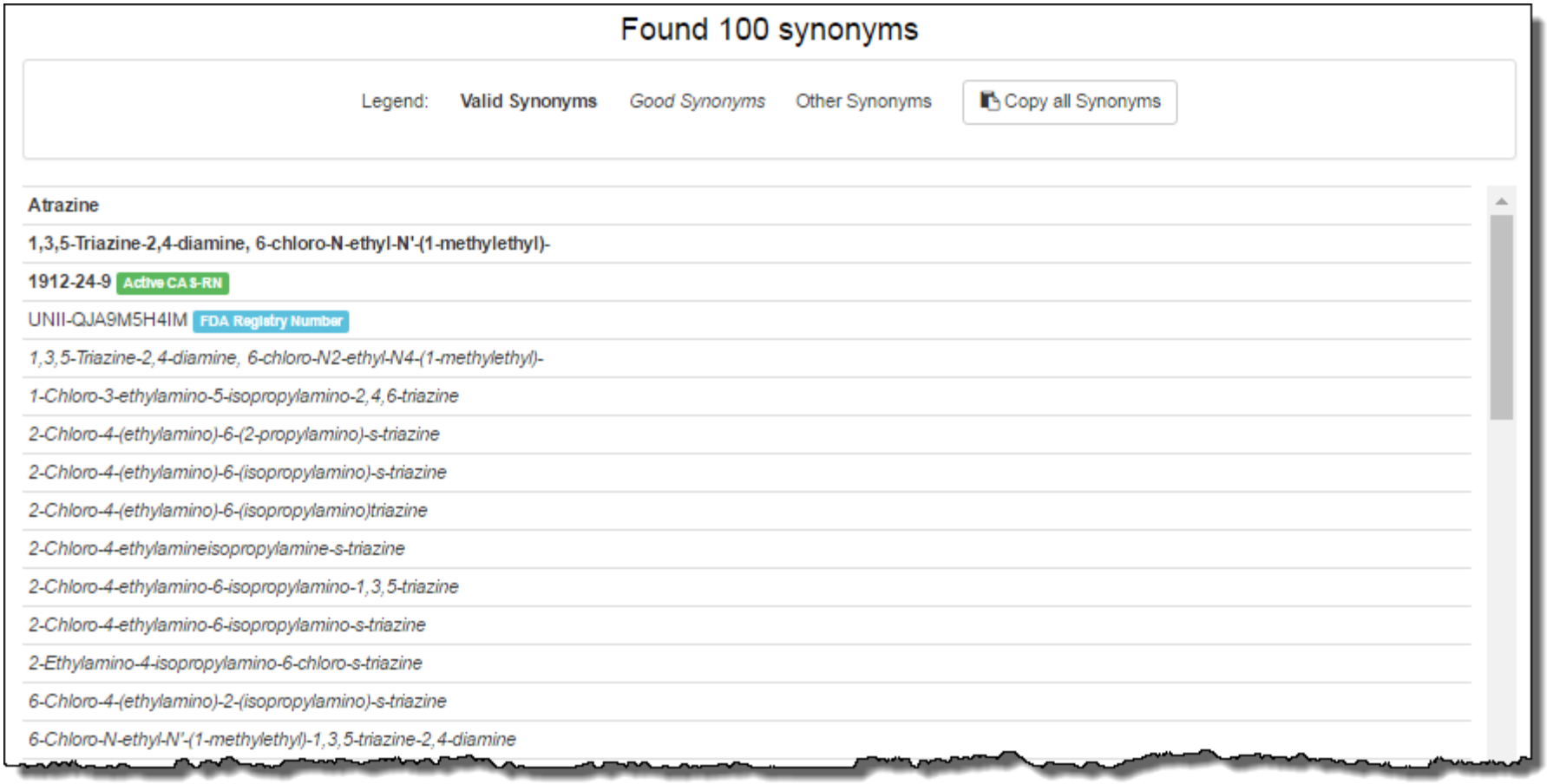
LogKoa: Octanol-Air

Download as: TSV | Excel | SDF

Property	Average		Median		Range		Unit
	Experimental	Predicted	Experimental	Predicted	Experimental	Predicted	
LogP: Octanol...	2.61 (1)	2.66 (4)	2.61 to 2.61	2.66	2.61	2.50 to 2.82	-
Water Solubility	1.30e-04 (1)	1.46e-02 (4)	1.30e-04 to 1....	1.46e-02	1.30e-04	1.50e-04 to 5....	mol/L
Density	-	1.27 (1)	-	1.27	-	-	g/cm^3
Melting Point	174 (6)	151 (3)	173 to 177	151	173 to 177	114 to 187	°C
Boiling Point	-	312 (3)	-	312	-	284 to 339	°C
Surface Tension	-	53.8 (1)	-	53.8	-	-	dyn/cm
Vapor Pressure	7.21e-11 (1)	4.47e-06 (3)	7.21e-11 to 7....	4.47e-06	7.21e-11	2.03e-07 to 1....	mmHg
Soil Adsorp. C...	174 (1)	173 (2)	174 to 174	173	174	144 to 202	L/kg
LogKoa: Octa...	-	8.40 (1)	-	8.40	-	-	-
Henry's Law	-	4.20e-10 (1)	-	4.20e-10	-	-	atm-m3/mole
Atmos. Hydrox...	-	1.71e-11 (1)	-	1.71e-11	-	-	cm3/molecule*...

Names and Identifiers

United States Environmental Protection Agency



Links to Other Resources

United States Environmental Protection Agency

General	Toxicology	Publications	Analytical
EPA Substance Re...	ACToR	Toxline	National Environme...
NIST Chemistry W...	DrugPortal	Environmental Heal...	MONA: MassBank ...
Household Product...	CCRIS	NIEHS	Tox21 Analytical Data
PubChem	ChemView	National Toxicology...	RSC Analytical Abs...
Chemspider	CTD	Google Books	FOR-IDENT
CPCat	e		
DrugBank	E		
HMDB	G		
Wikipedia	HSDB	Federal Register	
MSDS Lookup	ToxCast Dashboar...	Regulations.gov	
ChEMBL	LactMed	Springer Materials	
Chemical Vendors	ACToR PDF Report	BioCaddie DataMed	
Consumer Product...	International Toxicit...	RSC Publications	

The Office of the Federal Register (OFR) of the National Archives and Records Administration (NARA), and the U.S. Government Printing Office (GPO) jointly administer the FederalRegister.gov website.

Product Composition Details

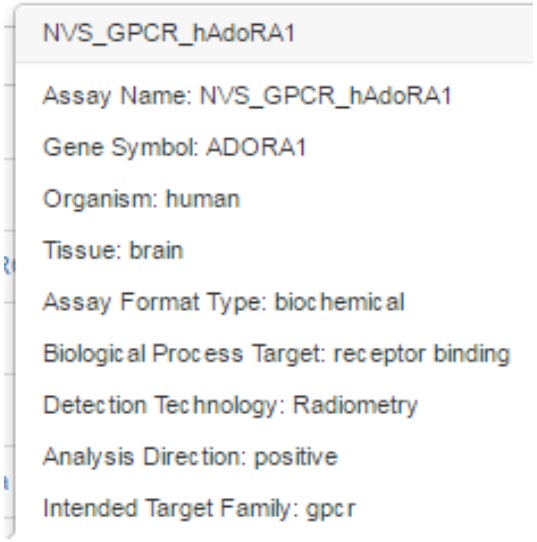
Download as: TSV | Excel

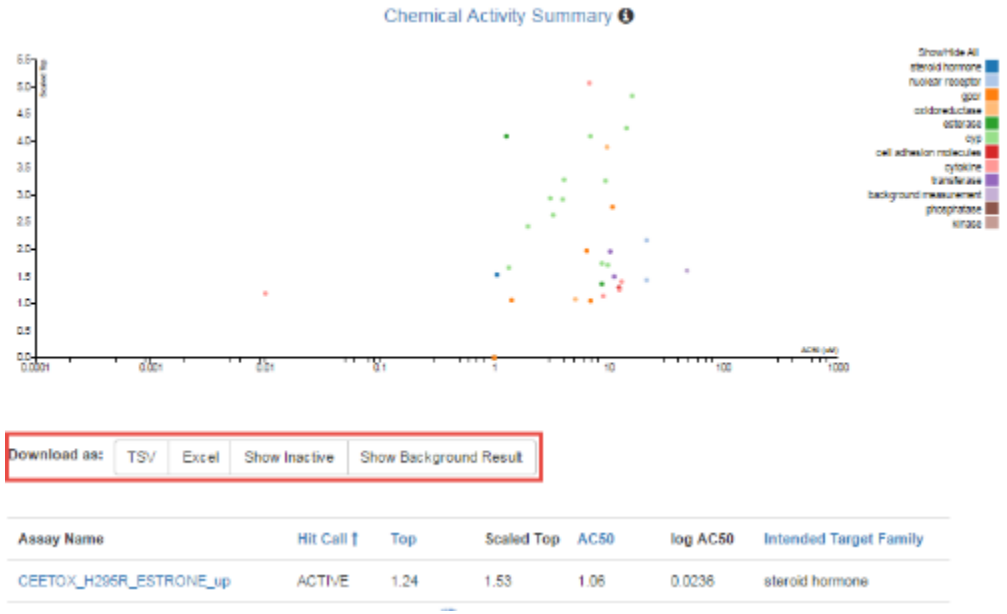
Product Composition

Product	Percent Composition ↓	Manufacturer
PS-380 ATRAZINE, 99.5% PURE	99.5%	CHEM SERVICE INC
ATRAZINE 80W	76.0%	ACETO AGRICULTURAL CHEMICALS CORP
ATRATOL 80W	71.3%	CIBA-GEIGY CORPORATION
ATRATOL 8P	8%	CIBA-GEIGY CORP
LPC-507M LAB PERFORMANCE CHECK MIXTURE IN T-B..	<1%	CHEM SERVICE INC
F2208S 2-CHLORO-4-ETHYLAMINO-S-TRIAZINE 100 UG..	<1%	CHEM SERVICE INC
ST84798, PROTURF ST AUGUSTINE WEED CONTROL	0.99%	O M SCOTT & SONS CO
SCOTTS BONUS, 26-3-3	0.94%	0. M. SCOTT & SONS CO.
SCOTTS LAWN PRO SUPER BONUS S, 21-3-3	.57%	0. M. SCOTT & SONS CO.
10-3-3 FERTILIZER + ATRAZINE	0.42%	ESTECH BRANDED FERTILIZERS, INC

In vitro Bioassay Data

- *In vitro* **bioassays** are used to determine the biological activity of a substance – **Toxcast project**
- A decade of measurements, and millions of dollars of data integrated into the dashboard





Physicochemical Properties

NEMI
National Environmental Methods Index
NWQMC

Home | About NEMI | Glossary | FAQ

View selected results | View all results | Download selected results

RESULTS: Your search returned 21 results. Back to search | Show/hide columns

Method ID	Method Source	Method Name	Analyte Name	Detection Level	Detection Level Type	Bias	Precision	Spiking Level
505	EPA-NERL	Pesticides and PCBs in Water GC-ECD	Atrazine (1912-24-9)	2.4 ug/L	MDL	85 % Rec (SL)	16.2 % RSD (SL)	5.00 ug/L
507	EPA-TSC/NERL	Pesticides in Water Using GCNPD	Atrazine (1912-24-9)	0.015 ug/L	MDL	101 % Rec (SL)	4 % RSD (SL)	0.13 ug/L
508.1	EPA-OGWDW/TSC	Chlorinated Pesticides, Herbicides, and Organohalides in Water by GCECD	Atrazine (1912-24-9)	0.003 ug/L	MDL	14 % Rec (SL)	7.14 % RSD (SL)	0.01 ug/L
525.2	EPA-NERL	Organics in Water Using GCMS	Atrazine (1912-24-9)	0.078 ug/L	MDL	109 % Rec (SL)	4.8 % RSD (SL)	0.50 ug/L

Integrated Literature Searching

EPA United States Environmental Protection Agency

PubChem

BioAssay | Compound | Substance

Google Scholar

atrazine

PubMed

PubMed comprises more than 26 million citations for biomedical literature from MEDLINE, life science journals, and online books. Citations may include links to full-text content from PubMed Central and publisher web sites.

Integrated Literature Searching

United States Environmental Protection Agency

PubChem Artic...
Google Scholar
Abstract Sifter
PubChem Patents

PubChem Articles

PUBCHEM > COMPOUND > ATRAZINE > LITERATURE > DEPOSITOR PROVIDED PUBMED CITATIONS >

Depositor Provided PubMed Citations

Download

1 to 10 of 876 1 2 3 ... 88 Date

PMID	Date	Title	Journal
26923738	2016-06-01	Atrazine blocks ovulation via suppression of Lhr and Cyp19a1 mRNA and estradiol secretion in immature gonadotropin-treated rats.	Reproductive toxicology (Elmsford, N.Y.)
27114639	2016-03-01	Atrazine exposure causes mitochondrial toxicity in liver and muscle cell lines.	Indian journal of pharmacology
26647222	2016-01-05	Endocrine-Disrupting Effects of Pesticides through Interference with Human Glucocorticoid Receptor.	Environmental science & technology
26464060	2016-01-01	Effects of Neonicotinoids on Promoter-Specific Expression and Activity of Aromatase (CYP19) in Human Adrenocortical Carcinoma (H295R) and Primary Umbilical Vein Endothelial (HUVEC) Cells.	Toxicological sciences : an official journal of the Society of Toxicology
26377646	2015-09-16	Atrazine-Mediated Disruption of Steroidogenesis in BLTK1 Murine Leydig Cells.	Toxicological sciences : an official journal of the Society of Toxicology

Integrated Literature Searching

United States Environmental Protection Agency

PubChem Articles

Google Scholar

Abstract Sifter

PubChem Patents

Select Term: Environmental AND Non-targeted Analysis AND -

Edit the Query Before Querying (73 Characters)

"Non-targeted Analysis" AND "Environmental" AND "1912-24-9" OR "Atrazine"

Submit

Integrated Literature Searching

United States Environmental Protection Agency

PubChem Articles

Google Scholar

Abstract Sifter

PubChem Patents

Select Term: Environmental AND Non-targeted Analysis AND -

Edit the Query Before Querying (73 Characters)

"Non-targeted Analysis" AND "Environmental" AND "1912-24-9" OR "Atrazine"

Google "Non-targeted Analysis" AND "Environmental" AND "1912-24-9" OR "Atrazir

Scholar About 30 results (**0.14** sec) My Citations 10

Articles
Case law
My library

Any time
Since 2017
Since 2016
Since 2013
Custom range...

Sort by relevance
Sort by date

include patents
include citations

Create alert

Environmental metabolomics: a critical review and future perspectives [PDF] academia.edu
JG Bundy, MP Davey, MR Viant - Metabolomics, 2009 - Springer
... Hence, several steps must now to taken to move **environmental** metabolomics forward, as discussed below, including issues related to ... And for a relatively rapid, quantitative and **non-targeted analysis** of the most abundant metabolites one could select the well-established ...
Cited by 306 Related articles All 9 versions Cite Save

New trends in the analytical determination of emerging contaminants and their transformation products in **environmental** waters [PDF] researchgate.net
A Agüera, MJM Bueno, AR Fernández-Alba - **Environmental** Science and ..., 2013 - Springer
... Take survey. Download PDF. **Environmental** Science and Pollution Research. **Environmental** Science and Pollution Research. ... New trends in the analytical determination of emerging contaminants and their transformation products in **environmental** waters. ...
Cited by 51 Related articles All 10 versions Cite Save

Coupling passive sampling and time of flight mass spectrometry for a better estimation of polar pesticide freshwater contamination: simultaneous target quantification ...
R Guibal, S Lissalde, A Charriau, G Poulier... - ... of Chromatography A, 2015 - Elsevier
... The proposed quantification method was validated for 43 compounds with variation of calibration slopes below 10% in **environmental** matrix. For the unvalidated compound DIA (**atrazine**-desisopropyl: an **atrazine** metabolite), interference increased the error of concentration ...
Cited by 12 Related articles All 5 versions Cite Save

Integrated Literature Searching

United States Environmental Protection Agency

PubChem Articles

Google Scholar

Abstract Sifter

PubChem Patents

Select Term:

Select a Query Term

- Select a Query Term
- Hazard
- Fate and Transport
- Metabolism/PK/PD
- Chemical Properties
- Exposure
- Mixtures

Edit the Query Before Retrieving Articles

"1912-24-9" OR "Atrazine"

Search and Count

Integrated Literature Searching

United States Environmental Protection Agency

PubChem Articles
Google Scholar
Abstract Sifter
PubChem Patents

Select Term:

Hazard

Retrieve Articles | 23 Articles (out of 23)

Edit the Query Before Retrieving Articles

("1912-24-9" OR "Atrazine") AND (NOAEL or LOEL or Rfd OR "reference concentration" OR "adverse effect level"[tiab] OR "cancer slope factor"[tiab])

Add additional query terms to filter abstracts:

Search and Count

Term 1	Term 2	Term 3	Total	PMID	PubYr	Title
0	0	0	0	26260...	2015	Using species sensitivity distribution approach to assess the risks ...
0	0	0	0	25138...	2014	Effect of atrazine and fenitrothion at no-observed-effect-levels (NO...
0	0	0	0	24797...	2014	Multigeneration reproduction and male developmental toxicity studi...
0	0	0	0	24323...	2013	Evaluation of hydroxyatrazine in the endocrine disruptor screening ...

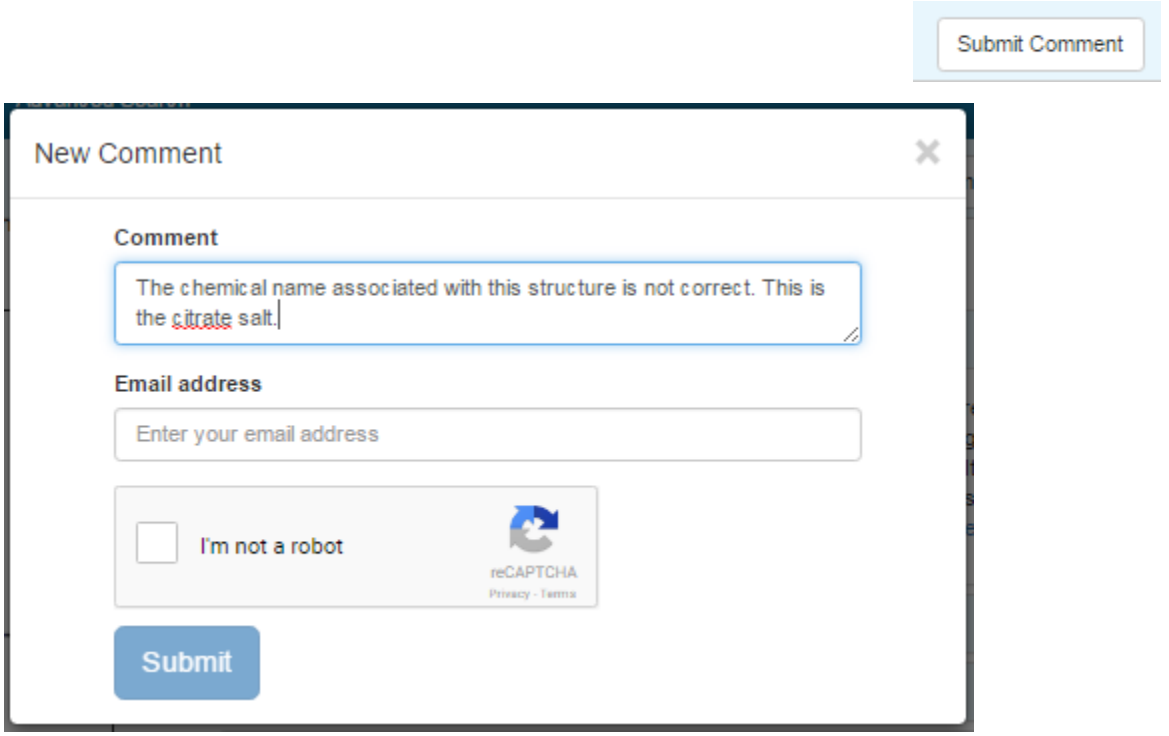
Record: 1 of 23

Title: Using species sensitivity distribution approach to assess the risks of commonly detected agricultural pesticides to Australia's tropical freshwater ecosystems.

Abstract: To assess the potential impacts of agricultural pesticides on tropical freshwater ecosystems, the present study developed temperature-specific, freshwater species protection concentrations (i.e., ecotoxicity threshold values) for 8 pesticides commonly detected in Australia's tropical freshwaters. Because relevant toxicity data for native tropical freshwater species to assess the ecological risks were mostly absent, scientifically robust toxicity data obtained at ≥20 °C were used for ecologically relevant taxonomic groups representing primary producers and consumers. Species sensitivity distribution (SSD) curves were subsequently generated for predicted chronic exposure using Burrlioz 2.0 software with mixed chronic and converted acute data relevant to exposure conditions at ≥20 °C. Ecotoxicity threshold values for tropical freshwater ecosystem

Crowdsourced Data Curation

- Maintaining high-quality data is a challenge
- Every user can contribute to improving the data!



“UVCBs” - Unknown or Variable Composition, Complex Reaction Products and Biological Materials

Poly (acrylamide-co-acrylic acid), partial sodium salt

62649-23-4 | DTXSID1049722

Searched by Approved Name: Found 1 result for 'Poly (acrylamide-co-acrylic acid), partial sodium salt'.

Record Information

Citation

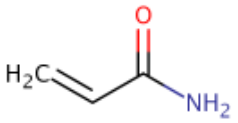
U.S. Environmental Protection Agency. Chemistry Dashboard. https://comptox.epa.gov/dashboard/DTXSID1049722 (accessed February 12, 2017), Poly (acrylamide-co-acrylic acid), partial sodium salt

Data Quality 80%

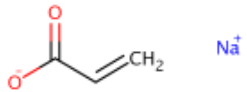
Manually curated and no conflicts in multiple public sources

Related Chemicals

Found 2 chemicals



Acrylamide
79-06-1



2-Propenoic acid, sodium salt
7446-81-3

Batch Searching for Data for Thousands of Chemicals

- What are these chemicals?

Chemicals Used in the Hydraulic Fracturing Process in Pennsylvania
Prepared by the Department of Environmental Protection
Bureau of Oil and Gas Management

Updated June 10, 2010

Chemical	Product Name
2,2-Dibromo-3-Nitrilopropionamide	Bio Clear 1000/Bio Clear 2000/ Bio-Clear 200/BioRid20L/ EC6116A
2-methyl-4-isothiazolin-3-one	X-Cide 207
5-chloro-2-methyl-4-isothiazolin-3-one	X-Cide 207
Acetic Acid	Fe-1A Acidizing Composition/ Packer Inhibitor
Acetic Anhydride	Fe-1A Acidizing Composition
Acetylene	GT&S Inc./ Airco
Alcohol Ethoxylated	C12-16 NE-200
Alkyl benzene sulfonic acid	Tetrolite AW0007/ FR-46
Ammonia (aqueous)	FAW-5
Ammonium Bifluoride	ABF 37%
Ammonium Persulfate	AP Break
Ammonium Bisulfite	Techni-Hib 604/ Fe OXCLEAR/ Packer Inhibitor
Ammonium chloride	Salt Inhibitor
Ammonium Salt (alkylpolyether	

Batch Searching for Data for Thousands of Chemicals

United States Environmental Protection Agency

Select Input Type(s)

- [x] Chemical Name
- [ ] CAS-RN
- [ ] InChIKey
- [ ] DSSTox Substance ID
- [ ] Exact Molecular Formula

Enter Identifiers to Search

Sodium l-glutamate
4-Hydroxy-3-nitrophenylarsonic acid
4-Nitro-1,2-phenylenediamine
Methoxypromazine
1-Phenyl-3-methyl-5-pyrazolone
1-Phenyl-2-thiourea
Phenylbutazone
1,4-Benzenediamine
Prednisolone
Probenecid

Display All Chemicals | Download Chemical Data

Select Output Format

Excel

Customize Results

- [ ] Select All

Chemical Identifiers

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

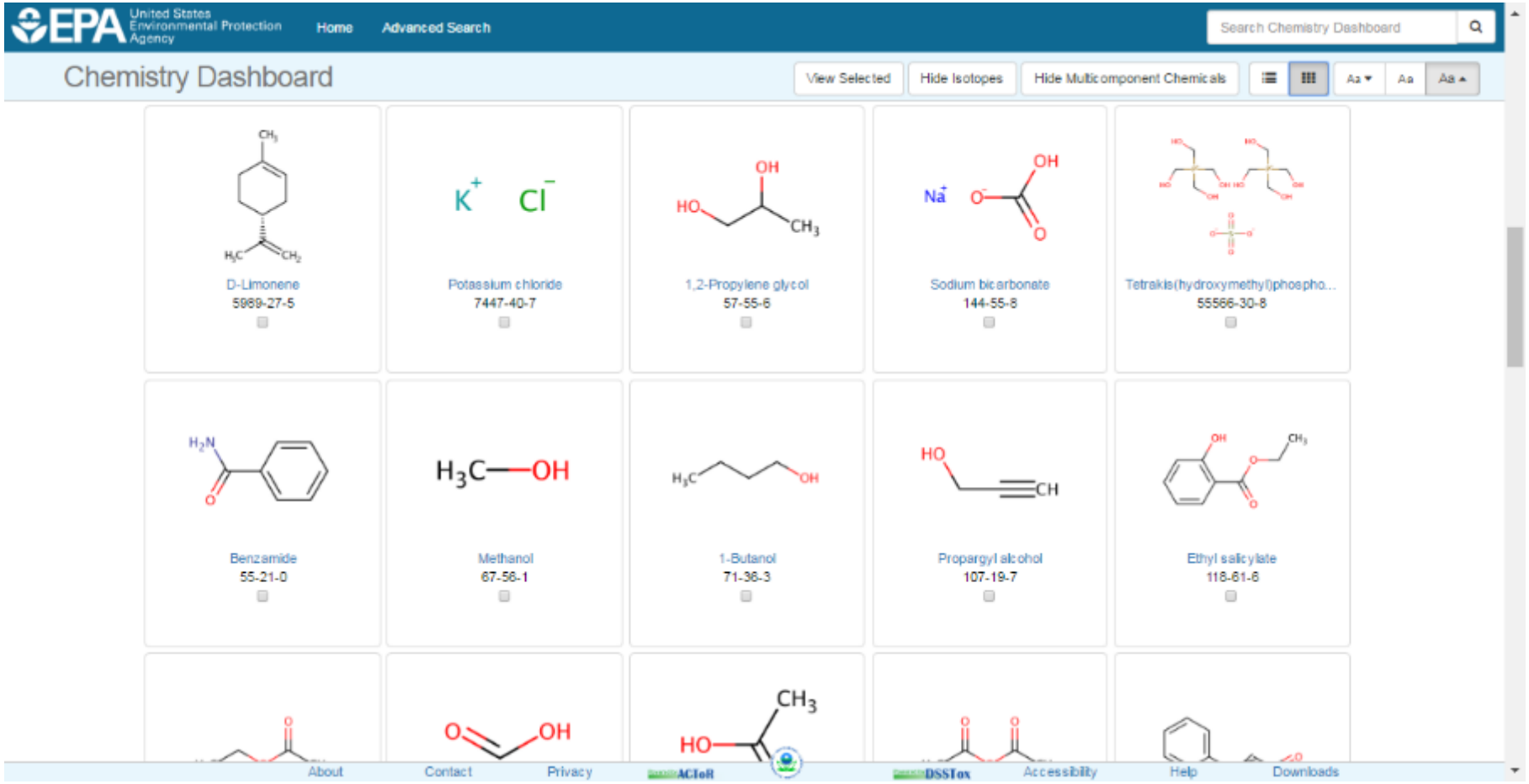
Structures

- [ ] Mol File
- [ ] SMILES
- [ ] InChI String

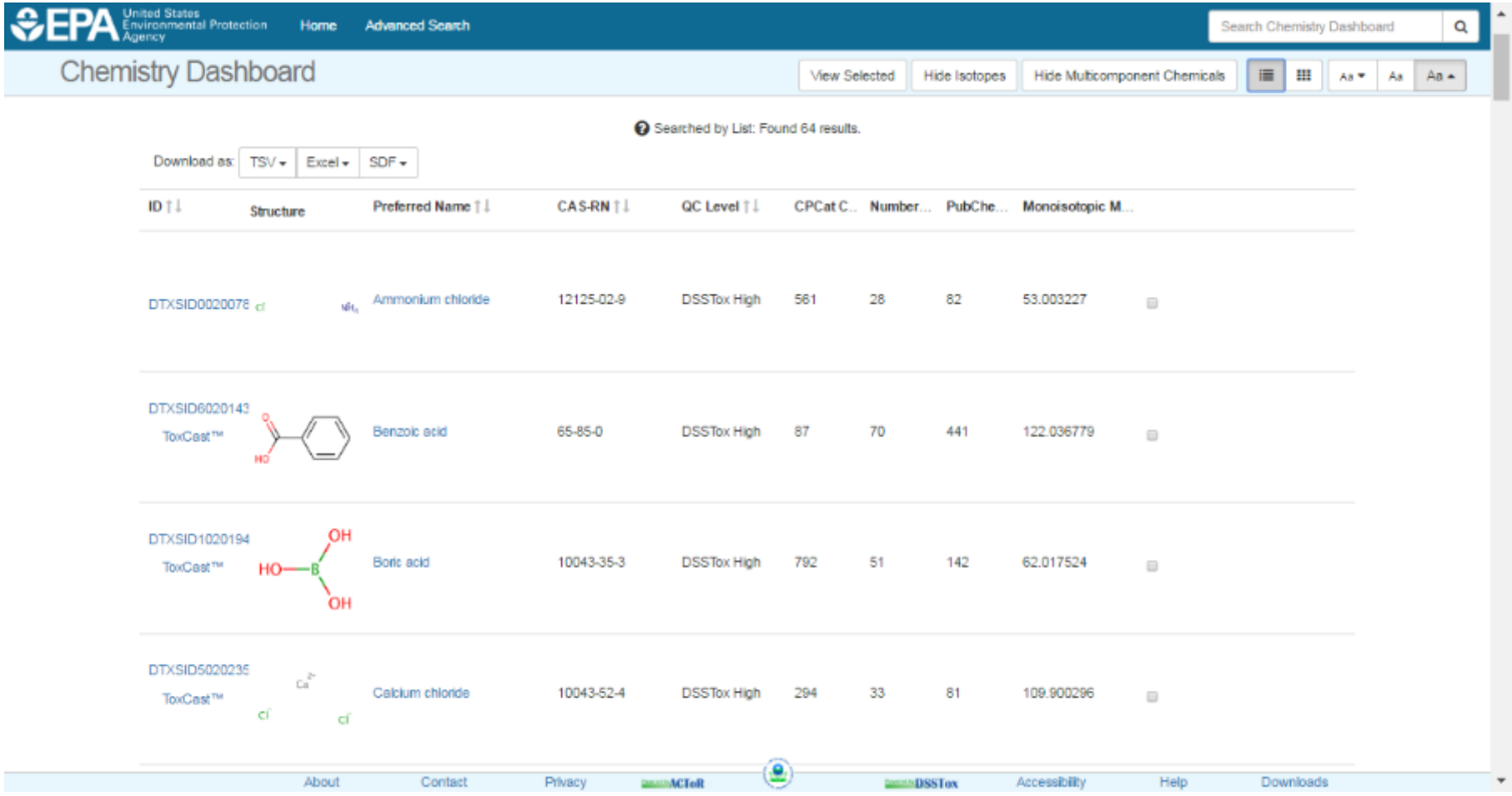
Intrinsic Properties

- [ ] Molecular Formula
- [ ] Average Mass
- [ ] Monoisotopic Mass
- [ ] OPERA and TEST Model Predictions

Access to associated data for review, modeling & download

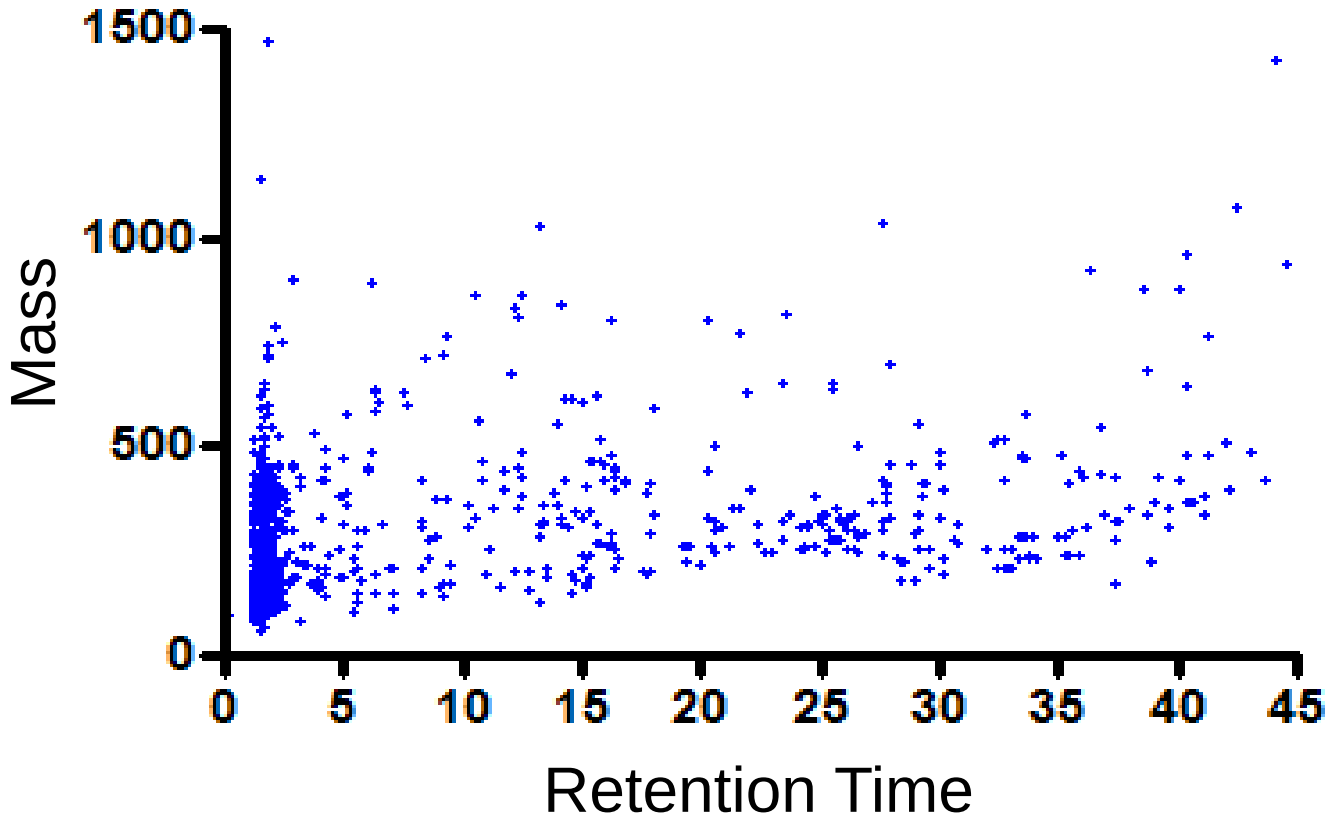


Access to associated data for review, modeling & download

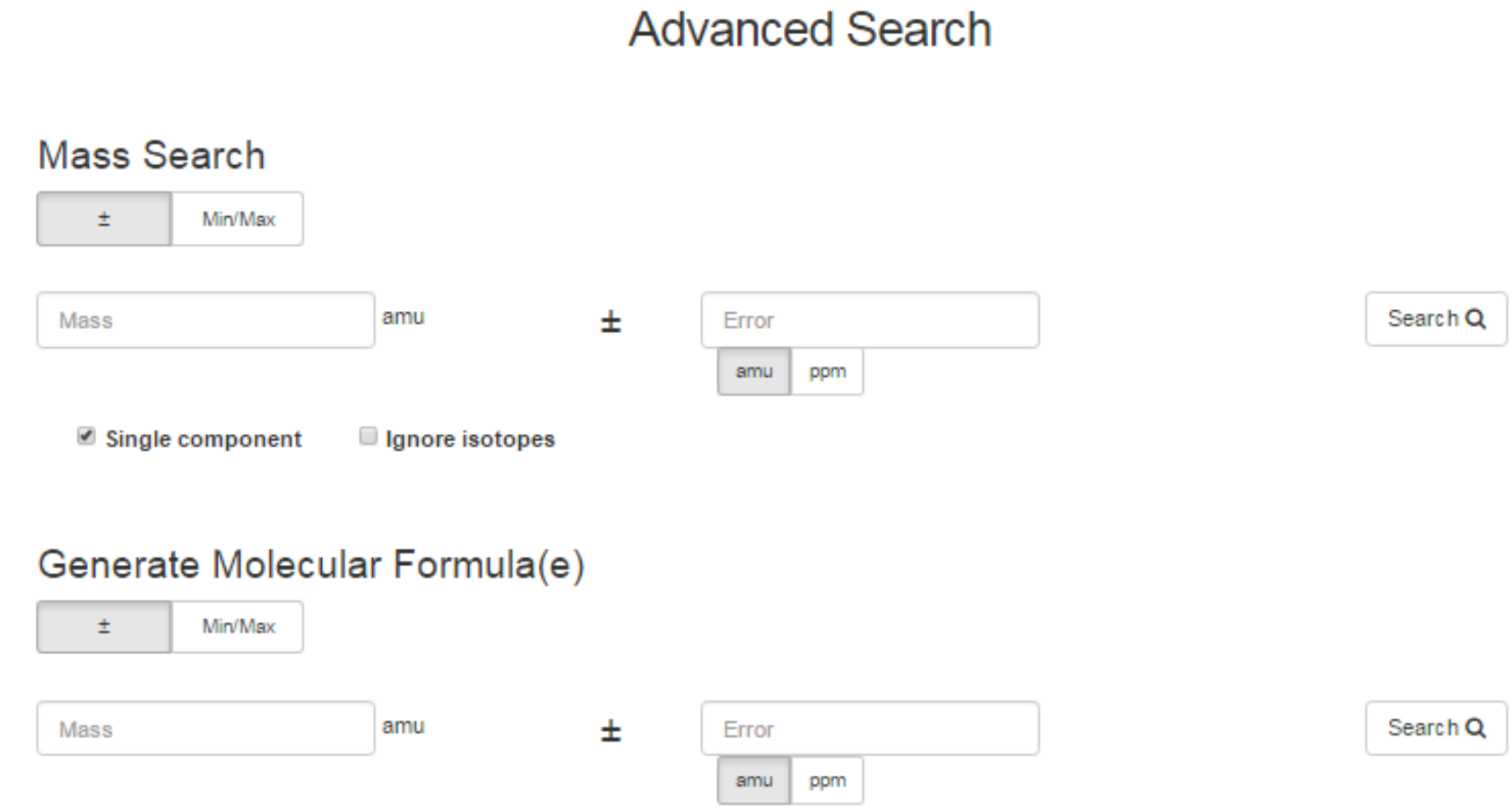


Non-targeted Analysis

~1000 Peaks in an American Health Homes Dust Sample



Searches for Specific Purposes



Data Available for Download

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

Downloads

DSSTox Identifier to PubChem Identifier Mapping File — Posted: 11/14/2016

The DSSTox to PubChem Identifiers mapping file is in TXT format and includes the PubChem SID, PubChem CID and DSSTox substance identifier (DTXSID).

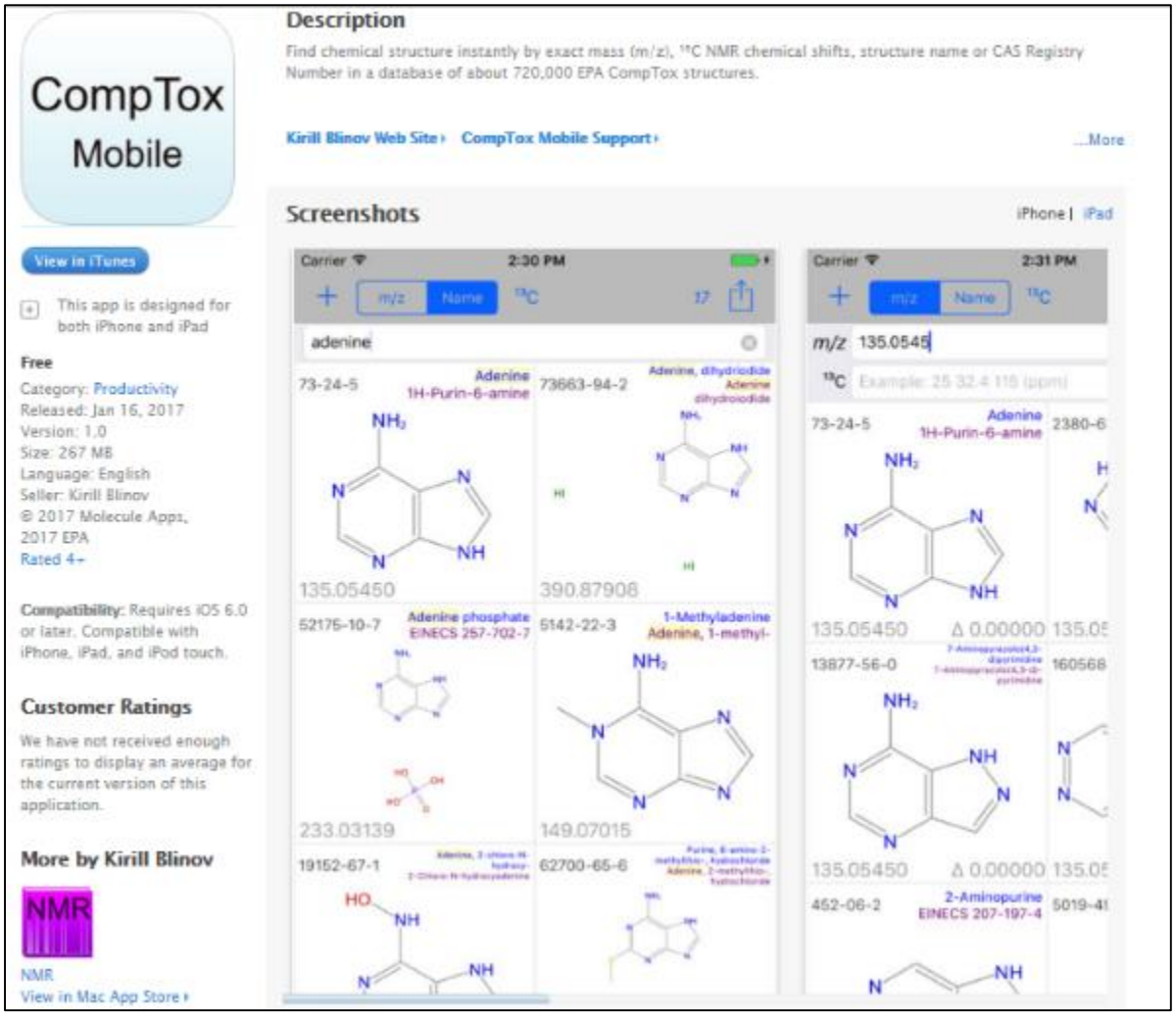
SID	CID	DTXSID
316388891	20404	DTXSID30873143
316388890	10142816	DTXSID70873142
316388889	50742127	DTXSID40873139
316388888	19073841	DTXSID20873137
316388887	11505215	DTXSID00873135
316388886	25021861	DTXSID80873133
316388885	2784427	DTXSID60873131
316388884	6731	DTXSID00873130

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.

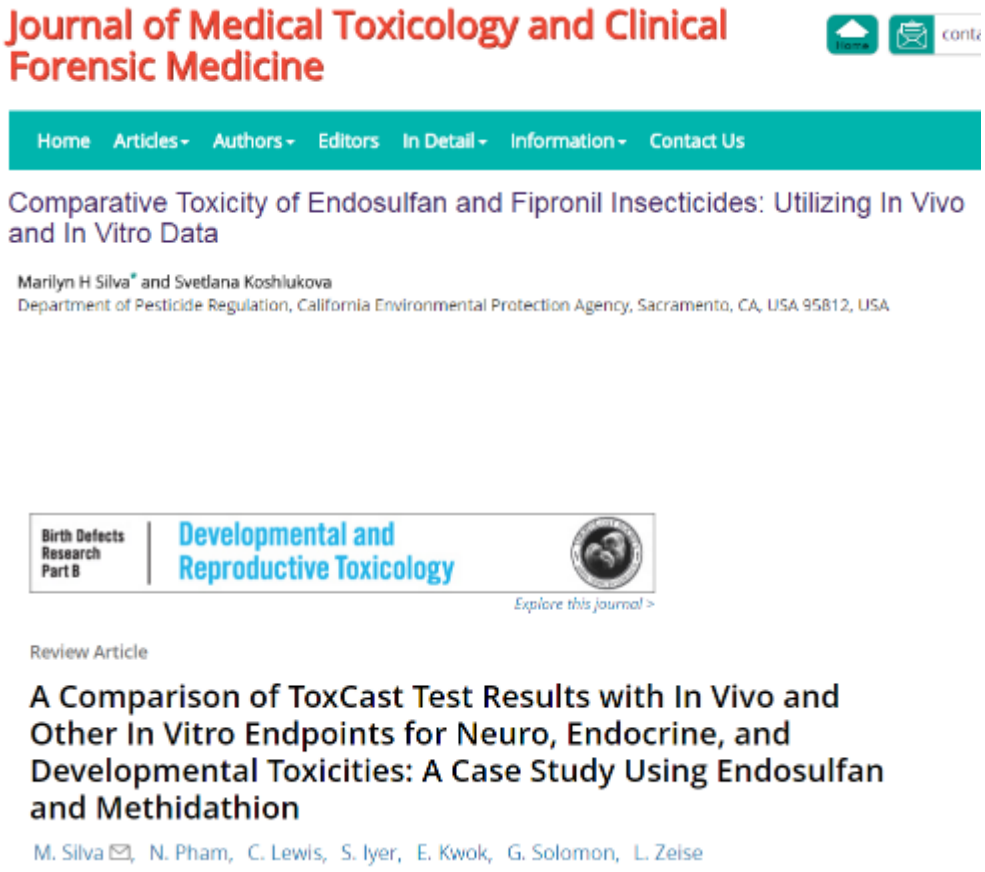
	A	B	
1	casrn	dsstox_substance_id	preferred_name
2	26148-68-5	DTXSID7020001	A-alpha-C
3	107-29-9	DTXSID2020004	Acetaldehyde oxime
4	60-35-5	DTXSID7020005	Acetamide
5	103-90-2	DTXSID2020006	Acetaminophen
6	968-81-0	DTXSID7020007	Acetohexamide

ca. 720,000 structures on an iPhone or iPad



Real World Applications

- Direct Access to data for researchers



Journal of Medical Toxicology and Clinical Forensic Medicine

Home Articles Authors Editors In Detail Information Contact Us

Comparative Toxicity of Endosulfan and Fipronil Insecticides: Utilizing In Vivo and In Vitro Data

Marilyn H Silva* and Svetlana Koshlukova

Department of Pesticide Regulation, California Environmental Protection Agency, Sacramento, CA, USA 95812, USA

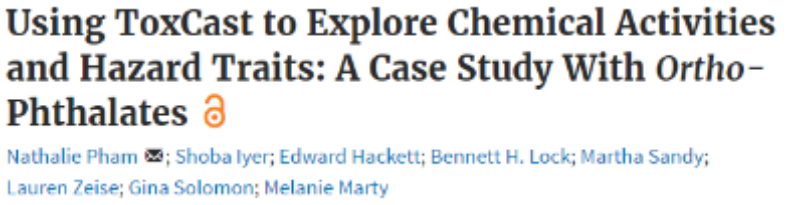
Birth Defects Research Part B | Developmental and Reproductive Toxicology

Explore this journal >

Review Article

A Comparison of ToxCast Test Results with In Vivo and Other In Vitro Endpoints for Neuro, Endocrine, and Developmental Toxicities: A Case Study Using Endosulfan and Methidathion

M. Silva, N. Pham, C. Lewis, S. Iyer, E. Kwok, G. Solomon, L. Zeise



Using ToxCast to Explore Chemical Activities and Hazard Traits: A Case Study With *Ortho*-Phthalates

Nathalie Pham; Shoba Iyer; Edward Hackett; Bennett H. Lock; Martha Sandy; Lauren Zeise; Gina Solomon; Melanie Marty

Real World Applications

- From chemical names to chemical data (COMMON NEED!)

Anthracene		120-12-7
Aromatic oil		
9,10-Anthracenedione, 2-ethyl	2-Ethylanthracene-9,10-dione	84-51-5
Arobenzene		103-33-3
Benz(e)acenaphthylene	Acephenanthrylene	201-06-9
Benzaldehyde, 3-hydroxyl-4-methoxy	3-Hydroxy-4-methoxy-benzaldehyde	621-59-0
Benz(a)anthracene		56-55-3
Benzene		71-43-2
Benzene, 1,3-bis(1-methylethenyl)-	1,3-bis(1-methylethenyl)benzene; 1,3-Diisopropenylbenzene	3748-13-8
Benzene, 1,4-bis(1-methylethenyl)-	1,4-Bis(1-methylethenyl)benzene	1605-18-1
1,4-Benzenediamine, N,N'-diphenyl-	N,N'-Diphenyl-p-phenylenediamine	74-31-7

Future Work

- Continue expansion and curation of data and types.
- Provide “programmatic access” to all data – connect to other Agency resources and allow other scientists to integrate their scientific applications.
- Integrate algorithms that allow for real-time predictions.
- Continue to assemble and enhance chemical lists and data for specific projects. Make available to Agency researchers and for public use.

Conclusion

- The CompTox Chemistry Dashboard provides access to data for ~740,000 chemicals
- An **Integration Hub** integrating multiple data sources: experimental and predicted property data, bioassay data, links to public and agency resources
- Data downloads allows for reuse in other systems and integration of resources to support research
- <1 year since initial release but already an important resource supporting environmental science

Contact

Antony Williams

US EPA Office of Research and Development

National Center for Computational Toxicology (NCCT)

919-541-1033

Williams.Antony@epa.gov