

# The EPA Comptox Chemistry Dashboard: A Web-Based Data Integration Hub

***Antony Williams***

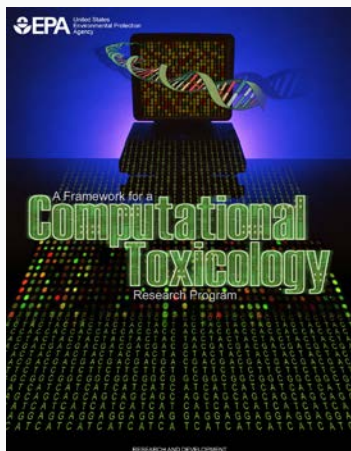
*U.S. Environmental Protection Agency, RTP, NC*

*This work was reviewed by the U.S. EPA and approved for presentation but does not necessarily reflect official Agency policy.*

*Luxembourg Centre for Systems Biomedicine*

*27<sup>th</sup> November 2017*

# National Center for Computational Toxicology

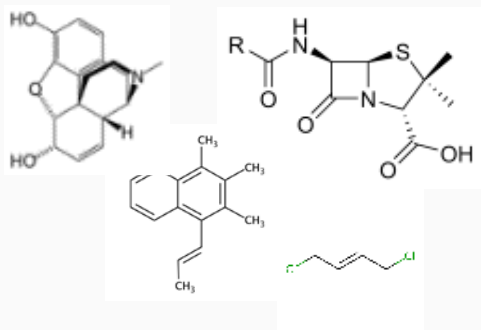


- National Center for Computational Toxicology established in 2005 to integrate:
  - High-throughput and high-content technologies
  - Modern molecular biology
  - Data mining and statistical modeling
  - Computational biology and chemistry
- Staffed by ~60 employees and contractors as part of EPA's Office of R&D
- Home of ToxCast & ExpoCast research efforts
- Key partner in U.S. Tox21 federal consortium
- Multiple cross-division collaborations (e.g. NERL, OPP, OPPT)

- Our team is broad in skills
  - Biologists, chemists and toxicologists
  - Bioinformaticians and cheminformaticians - modelers (QSAR, Deep Learning), text miners
  - Information technology team, software developers
  - A large IT support team for production applications (National Computing Center)

# Why we must innovate...

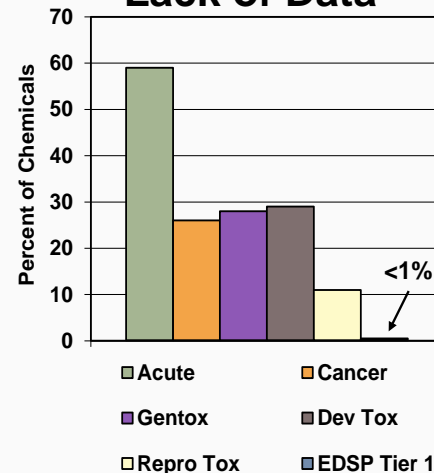
## Number of Chemicals /Combinations



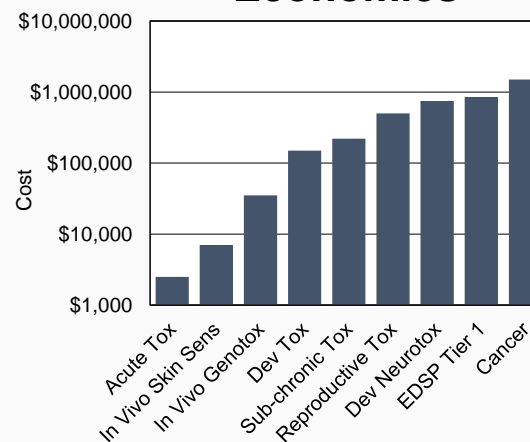
## Ethical Concerns



## Lack of Data



## Economics



# Toxic Substances Control List



Environmental Topics

Laws & Regulations

About EPA

Search EPA.gov



## TSCA Chemical Substance Inventory

CONTACT US

SHARE



- Inventory was initially published in 1979
- Second version, containing about 62k chemical substances, was published in 1982
- Continues to grow and now lists ~85k chemicals, about 15k are confidential business information

# The need for data and derivative models and algorithms

- NCCT outputs: include a lot of data, models, algorithms and software applications
- We produce Open Data – we want people to interrogate it, learn from it, develop understanding

## Toxicity Forecasting

### Advancing the Next Generation of Chemical Evaluation

EPA needs rapid and efficient methods to prioritize, screen and evaluate thousands of chemicals. EPA's Toxicity Forecaster (ToxCast) generates data and predictive models on thousands of chemicals of interest to the EPA. ToxCast uses high-throughput screening methods and computational toxicology approaches to rank and prioritize chemicals. In fact, EPA's Endocrine Disruption Screening Program (EDSP) is working to use ToxCast to rank and prioritize chemicals.



- ToxCast has data on over 1,800 chemicals from a broad range of sources including industrial and consumer products, food additives, and potentially "green" chemicals that could be safer alternatives to existing chemicals.
- ToxCast screens chemicals in over 700 high-throughput assays that cover a range of high-

## Downloadable Computational Toxicology Data

EPA's computational toxicology research efforts evaluate the potential health effects of thousands of chemicals. The process of evaluating potential health effects involves generating data that investigates the potential harm, or hazard of a chemical, the degree of exposure to chemicals as well as the unique chemical characteristics.

As part of EPA's commitment to share data, all of the computational toxicology data is publicly available for anyone to access and use.

### High-throughput Screening Data

EPA researchers use rapid chemical screening (called high-throughput screening assays) to limit the number of laboratory animal tests while quickly and efficiently testing thousands of chemicals for potential health effects.

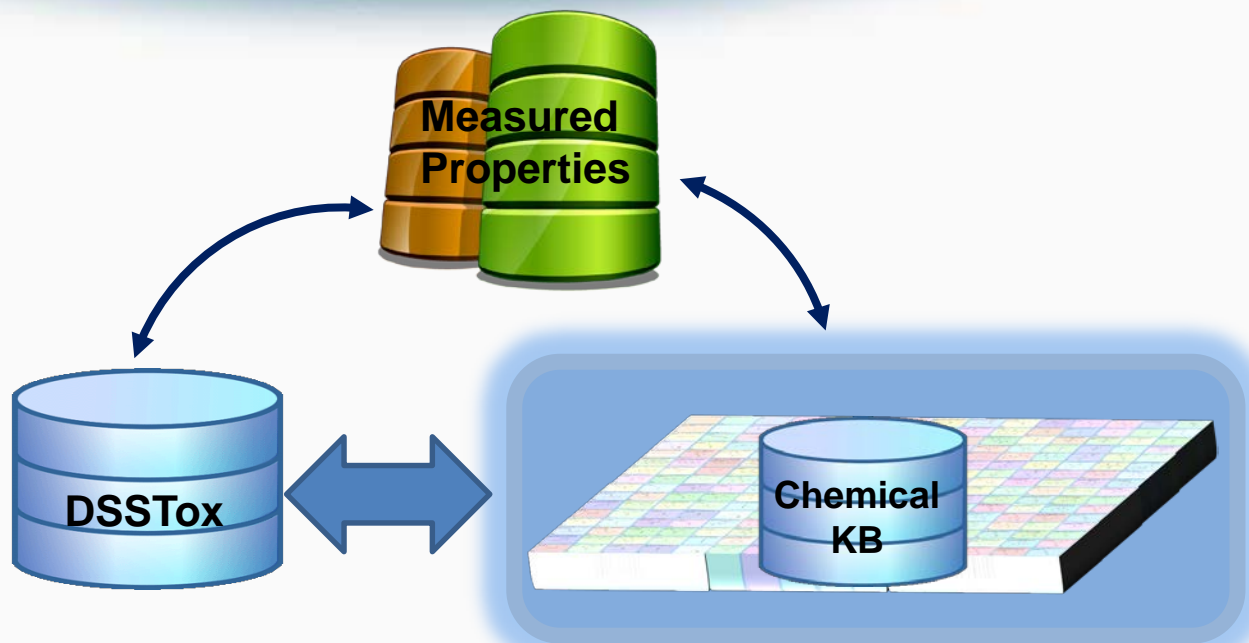
- [ToxCast Data](#): High-throughput screening data on thousands of chemicals.

### Rapid Exposure and Dose Data

EPA researchers develop and use rapid exposure estimates to predict potential exposure for thousands of chemicals.

- [High-throughput toxicokinetics data](#): It is important to link the external dose of a chemical to an internal blood or tissue concentration. This process is called toxicokinetics. EPA researchers measure the critical factors that determine the distribution

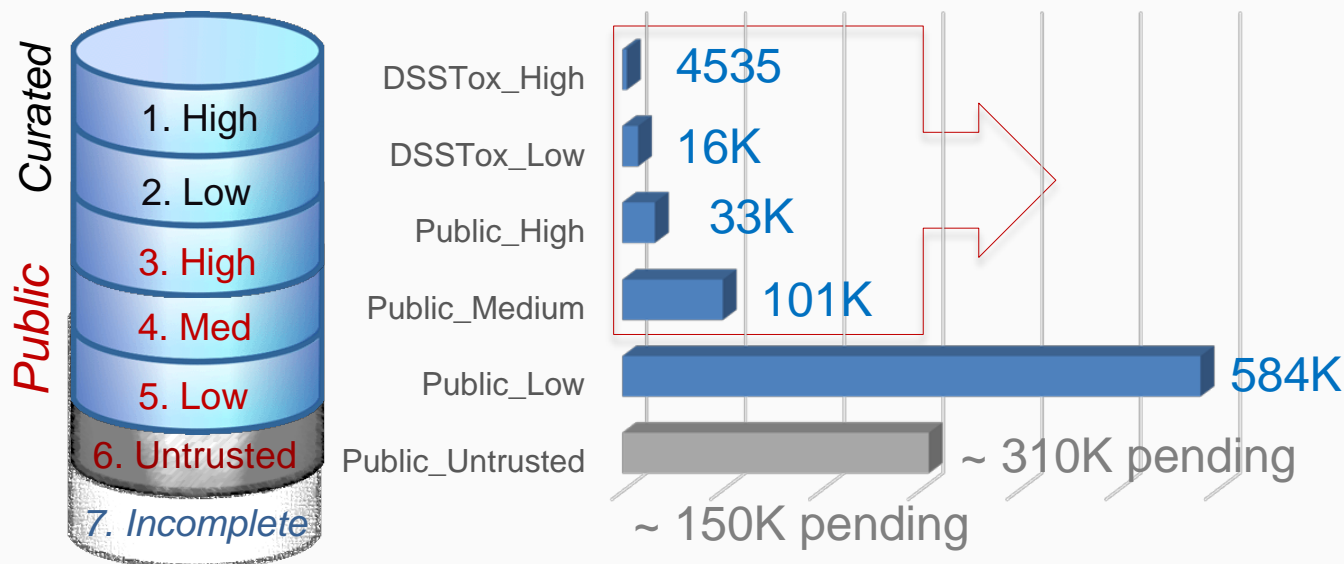
# Underpinning with chemicals





# Approximately 17 Years of Data...

## Growing with daily curation

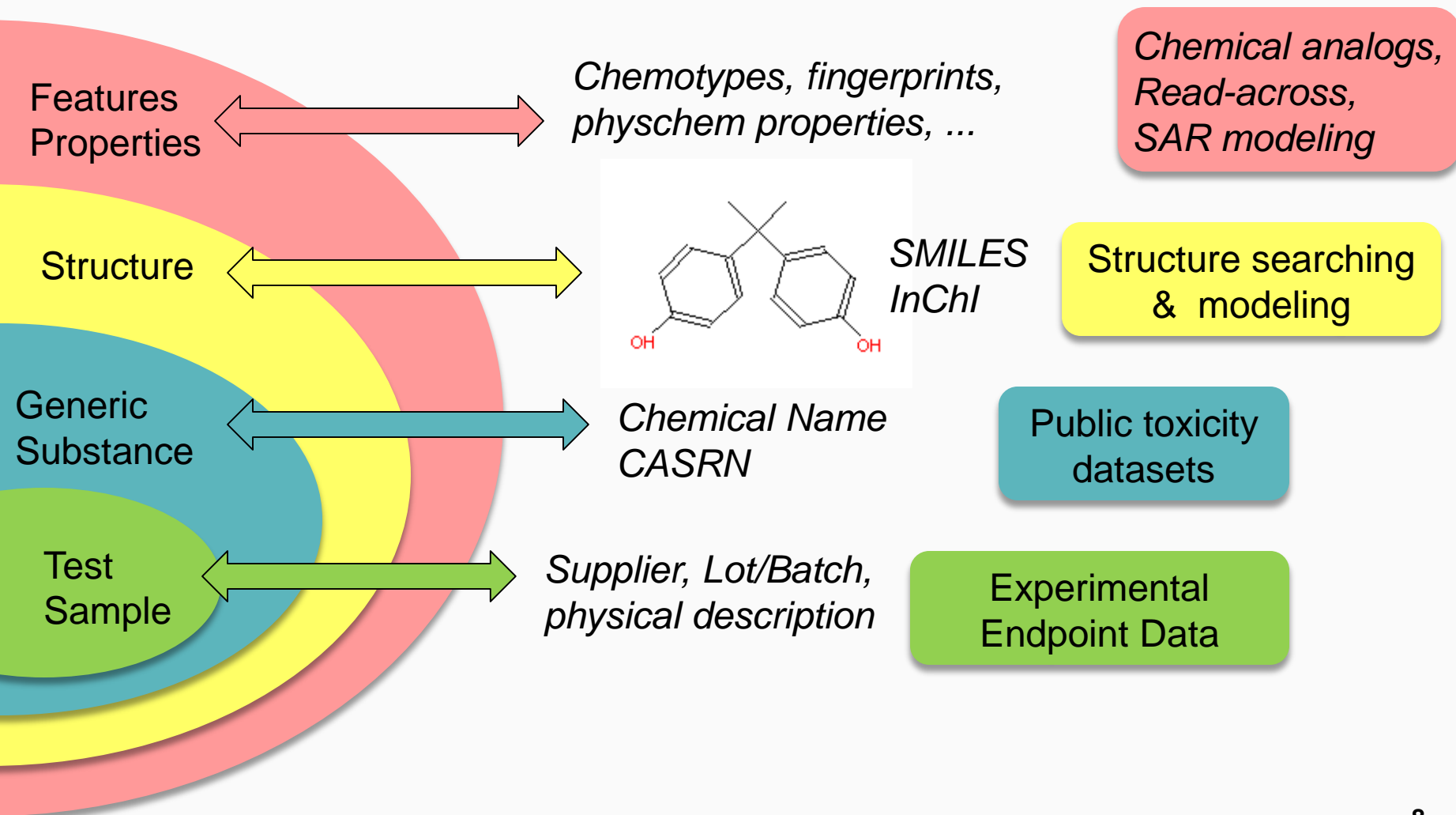


### QC Levels

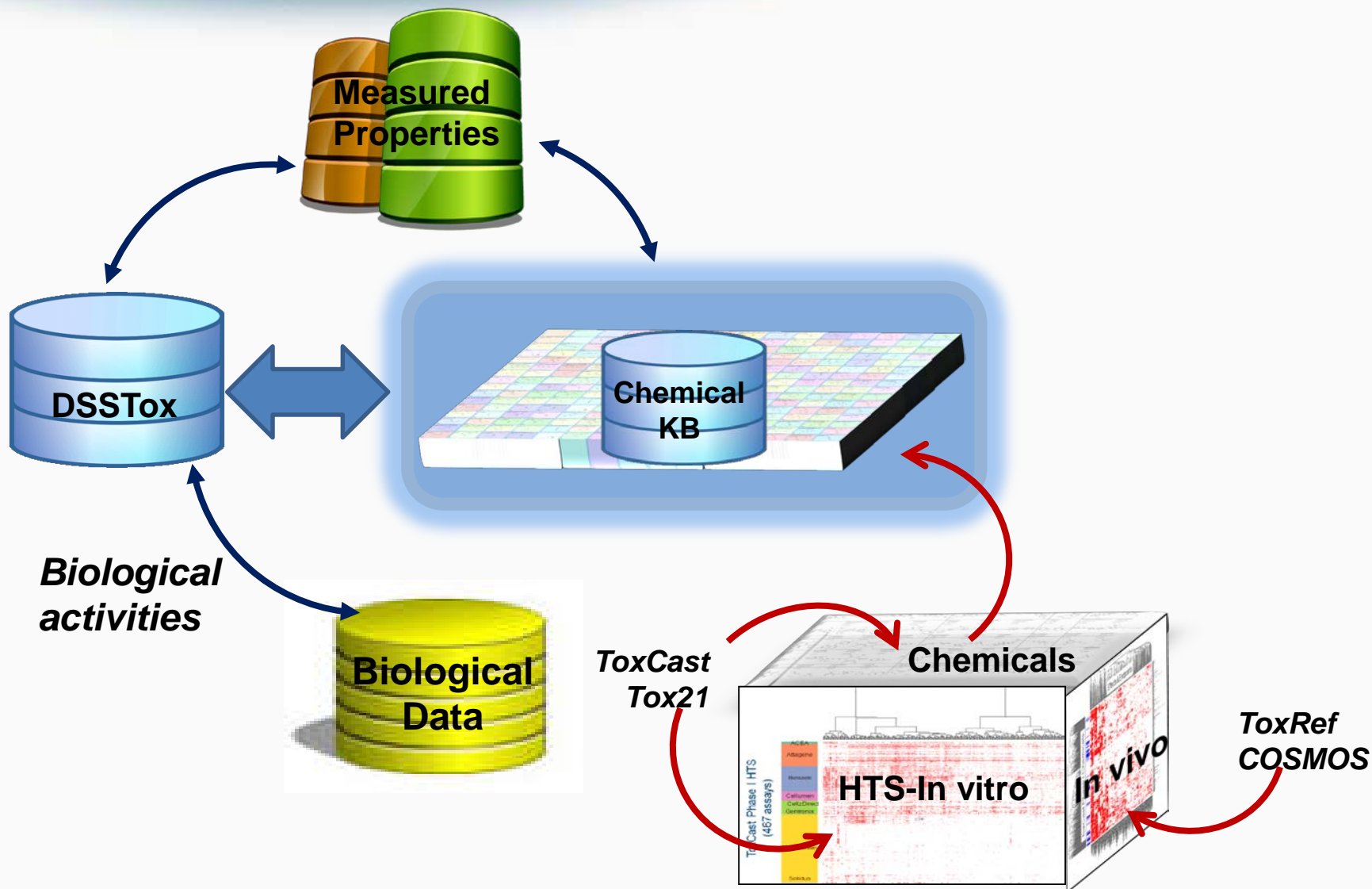
DSSTox_High:	Hand curated and validated
DSSTox_Low:	Hand curated and confirmed using multiple public sources
Public_High:	Extracted from EPA SRS and confirmed to have no conflicts in ChemID and PubChem
Public_Medium:	Extracted from ChemID and confirmed to have no conflicts in PubChem
Public_Low:	Extracted from ACToR or PubChem
Public_Untrusted:	Postulated, but found to have conflicts in public sources








# Chemical representation levels supporting data integration

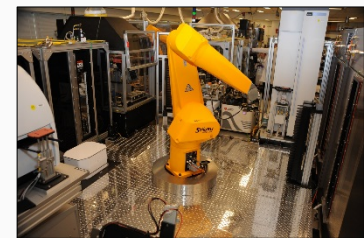
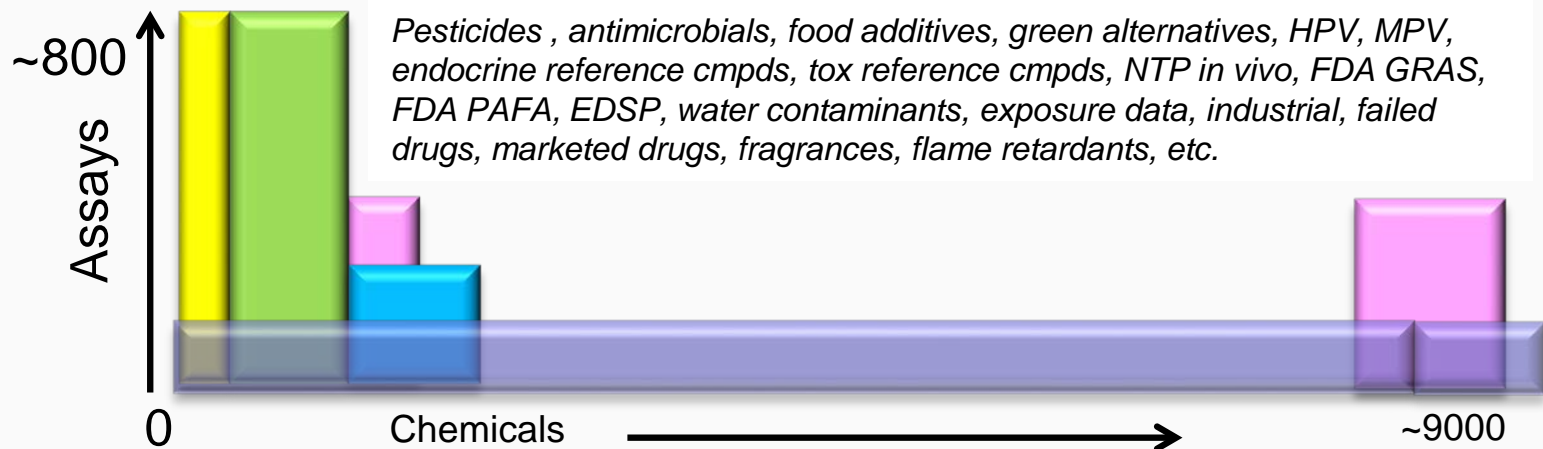


# Integrating *in vitro* and *in vivo* data

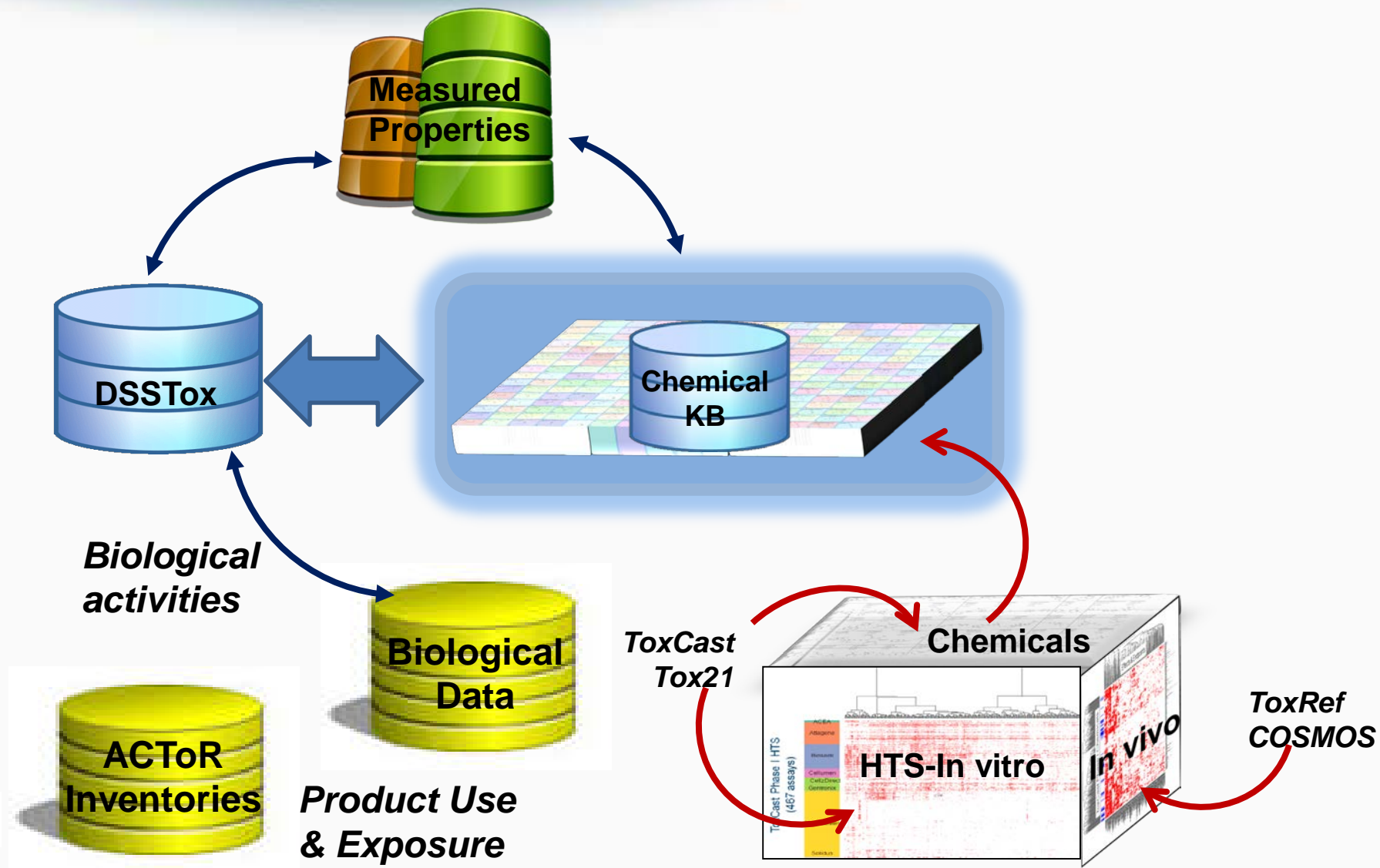


# TEN YEARS of Assay Measurements: ToxCast & Tox21

Set	Chemicals	Assays	Endpoints	Completion
ToxCast Phase I	 293	~600	~700	2011
ToxCast Phase II	 767	~600	~700	03/2013
ToxCast E1K	 800	~50	~120	03/2013
ToxCast Phase III	 ~900	~300	~300	In progress
Tox21	 ~9000	~80	~150	In progress



# Adding Product Use and Exposure



# High Throughput Measurement to Identify Exposure

## Rapid Chemical Exposure and Dose Research

EPA is responsible for ensuring the safety of thousands of chemicals. Quantitative exposure data are available for only a small fraction of registered chemicals. This type of exposure data is needed to thoroughly evaluate chemicals for potential risks to humans, wildlife and ecosystems. EPA is developing innovative methods to develop exposure estimates for thousands of chemicals to better protect human health and the environment. These innovative methods are called rapid exposure and dose assessments.

### Rapid Exposure Predictions

Rapid, also called high-throughput, exposure predictions or ExpoCast provide rapid exposure estimates for thousands of chemicals. ExpoCast quickly and efficiently looks at multiple routes of exposure to provide exposure estimates. ExpoCast uses and enhances two well-known exposure models to estimate chemical exposure.

- › [Farfield Exposure Models](#)
- › [Nearfield Exposure Models](#)

### Evaluating High-throughput Exposure Predictions

EPA is currently evaluating the effectiveness of high throughput exposure models

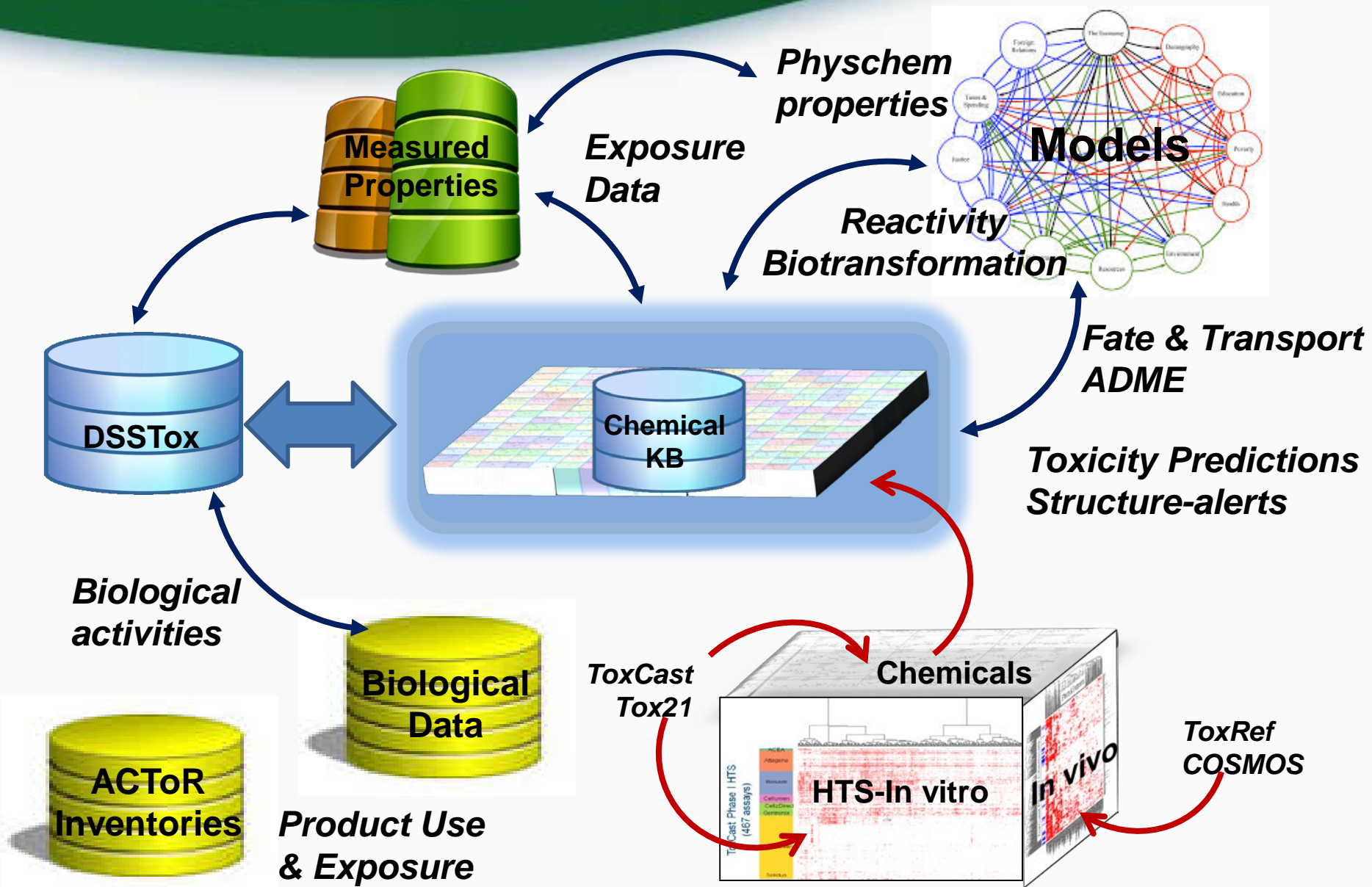


*Pictured Above: Farfield Exposure Examples*





# Building Models from the data



- Data curation, standardization and versioning is **essential**
- **Prototype** application development suffices for **research** projects
- **Production apps** need managed processes
- ODOSOS (Open Data, Open Source and Open Standards) endows many benefits




# The CompTox Chemistry Dashboard: An Overview

- A **publicly accessible website** delivering access:
  - ~760,000 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” for thousands of chemicals
- Day-to-day curation efforts for data quality


# Comptox Chemistry Dashboard

<https://comptox.epa.gov>


 United States  
Environmental Protection  
Agency

Home Advanced Search Batch Search Lists

Chemistry Dashboard As ▼ As As ▲



Chemistry Dashboard





☐ Single component search ☐ Ignore isotopes

See what people are saying, read the dashboard comments!

Need more? Use advanced search.


758 Thousand Chemicals

About Contact Privacy   Accessibility Help Downloads

**~760,000 chemicals**  
**>17 years of data**


# Comptox Chemistry Dashboard

<https://comptox.epa.gov>


 United States  
Environmental Protection  
Agency

Home Advanced Search Batch Search Lists



Chemistry Dashboard As ▼ As As ▲

 UNITED STATES  
ENVIRONMENTAL PROTECTION AGENCY

Chemistry Dashboard



- N-methyl-2-pyrrolidin-1-ylethanamine
- N-Methyl-2-pyrrolidone
- N-methyl-2-pyrrolidone dimethyl acetal
- N-Methyl-2-pyrrolidine
- N-METHYL-2-PYRROLIDINONE
- N-METHYL-2-PYRROLIDON
- N-Methyl-2-pyrrolidinone, compound with octahydro-1,3,5,7-tetranitro-1,3,5,7-tetrazocine (1:1)
- N-Methyl-2-pyrrolidone, titanium trichloride, aluminum chloride complex
- N-Methyl-2-pyrrolidone (NMP)

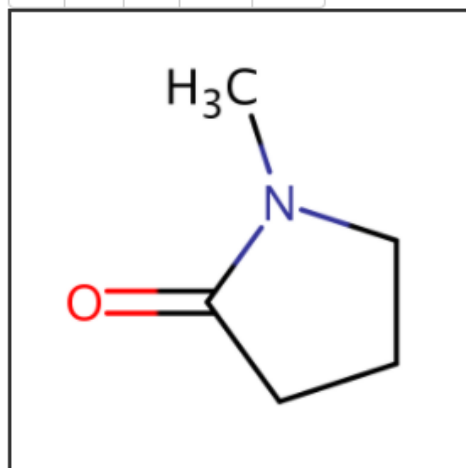
About Contact Privacy   Accessibility Help Downloads

**~760,000 chemicals**  
**>17 years of data**

## N-Methyl-2-pyrrolidone

872-50-4 | DTXSID6020856

© Searched by Synonym from Valid Source: Found 1 result for 'N-METHYLPYRROLIDONE'.



Wikipedia

### Intrinsic Properties

**Molecular Formula:** C<sub>5</sub>H<sub>9</sub>NO

**Average Mass:** 99.133 g/mol

**Monoisotopic Mass:** 99.068414 g/mol

[Find All Chemicals](#)



Structural Identifiers

Related Compounds

Presence in Lists

Record Information

Chemical Properties

Env. Fate/Transport

Toxicity Values (Beta)

ADME (Beta)

Exposure

Bioassays

Similar Molecules (Beta)

Synonyms

Literature

External Links

# Chemical Properties

## Summary

Download as:

TSV

Excel

SDF

LogP: Octanol-Water

Water Solubility

Density

Flash Point

Melting Point

Boiling Point

Surface Tension

Thermal Conductivity

Vapor Pressure

Viscosity

LogK<sub>ow</sub>: Octanol-Air

Property	Average		Median		Range		Unit
	Experimental	Predicted	Experimental	Predicted	Experimental	Predicted	
LogP: Octanol-Water	-0.380 (1)	-0.329 (5)	-0.380	-0.329	-0.380	-0.494 to -0.110	-
Water Solubility	10.1 (1)	6.68 (4)	10.1	6.68	10.1	1.48 to 12.8	mol/L
Density	-	1.02 (2)	-	1.02	-	1.01 to 1.03	g/cm <sup>3</sup>
Flash Point	-	75.7 (2)	-	75.7	-	65.2 to 86.1	°C
Melting Point	-23.8 (8)	1.32e-01 (4)	-24.0	1.32e-01	-24.0 to -23.0	-10.2 to 25.9	°C
Boiling Point	203 (6)	199 (5)	204	199	202 to 204	191 to 202	°C
Surface Tension	-	33.8 (1)	-	33.8	-	-	dyn/cm
Thermal Conductivity	-	158 (1)	-	158	-	-	mW/(m <sup>2</sup> K)
Vapor Pressure	3.45e-01 (1)	5.21e-01 (4)	3.45e-01	5.21e-01	3.45e-01	1.71e-01 to 9.99e-01	mmHg
Viscosity	-	3.61 (1)	-	3.61	-	-	cP
LogK <sub>ow</sub> : Octanol-Air	-	3.84 (1)	-	3.84	-	-	-
Henry's Law	3.20e-09 (1)	9.15e-09 (1)	3.20e-09	9.15e-09	-	-	atm-m <sup>3</sup> /mole
Index of Refraction	-	1.47 (1)	-	1.47	-	-	-
Molar Weight	100.12 (1)	100.12 (1)	100.12	100.12	100.12	100.12	g/mol

Chemical Properties

Env. Fate/Transport

Toxicity Values (Beta)

ADME (Beta)

Exposure

Bioassays

Similar Molecules (Beta)

Synonyms

Literature

External Links

# Available Properties

- Solubility
  - Melting Point
  - Boiling Point
  - LogP (Octanol-water partition coefficient)
  - Atmospheric Hydroxylation Rate
  - LogBCF (Bioconcentration Factor)
  - Biodegradation Half-life
  - Henry's Law Constant
  - Fish Biotransformation Half-life
  - LogKOA (Octanol/Air Partition Coefficient)
  - LogKOC (Soil Adsorption Coefficient)
  - Vapor Pressure
- 
- Data and models are updated ~90 days

# Multiple Prediction Algorithms

Summary	LogP: Octanol-Water			
LogP: Octanol-Water				
Water Solubility				
Density				
Flash Point				
Melting Point				
Boiling Point				
Surface Tension				
Thermal Conductivity				
Vapor Pressure				
Viscosity				
LogKoa: Octanol-Air				
Henry's Law				

LogP: Octanol-Water			
	Average	Median	Range
Experimental	-0.380 (1)	-0.380	-0.380
Predicted	-0.329 (5)	-0.329	-0.494 to -0.110

Download as: [TSV](#) [Excel](#) [SDF](#)

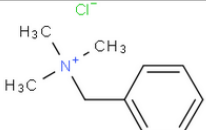
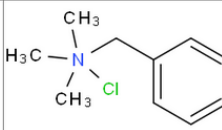
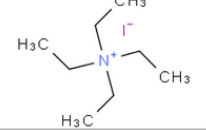
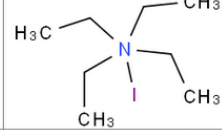
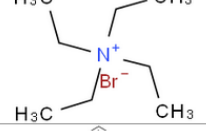

Experimental			
Source	Result		
<a href="#">PhysPropNCCT</a>	-0.380		

Predicted			
Source	Result	Calculation Details	QMRf
<a href="#">EPISUITE</a>	-0.110	Not Available	Not Available
<a href="#">NICEATM</a>	-0.494	Not Available	<a href="#">Available</a>
<a href="#">ACD/Labs Consensus</a>	-0.345	Not Available	Not Available
<a href="#">ACD/Labs</a>	-0.398	Not Available	Not Available
<a href="#">OPERA</a>	-0.300	<a href="#">OPERA Model Report</a>	<a href="#">Available</a>

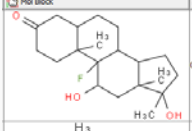
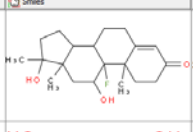
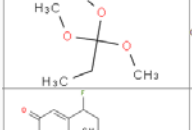
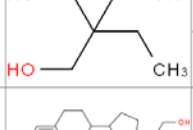
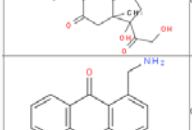
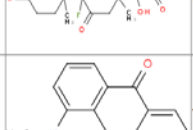




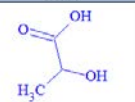
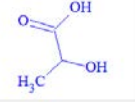
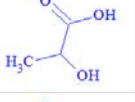
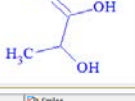
# Consuming and Curating Public Data

Public data should be curated prior to modeling

Mol Block	CAS	NAME	Smiles
	000056-93-9	BENZYL TRIMETHYL AMMONIUM CHLORIDE	
	000068-05-3	TETRAETHYL AMMONIUM IODIDE	
	000071-91-0	TETRAETHYL AMMONIUM BROMIDE	

**Covalent Halogens**

Mol Block	CAS	NAME	Smiles
	000076-43-7	FLUCYMESTERONE	
	000077-99-6	1,1,1-TRIS(4-HYDROXYETHYL)PROPANE	
	000079-60-7	CORTISONE-4A-FLUORO	
	000082-38-2	DISPERSE RED 9	

Structure	Formula	FW	CAS	NAME	MP	EgMP	ErrorMP
	C <sub>3</sub> H <sub>6</sub> O <sub>3</sub>	90.0779	000050-21-5	LACTIC ACID	1.6000000000000000e+001	2.2600000000000000e+001	5.8000000000000000e+000
	C <sub>3</sub> H <sub>6</sub> O <sub>3</sub>	90.0779	000079-33-4	L-LACTIC ACID	5.5000000000000000e+001	2.2600000000000000e+001	-3.0340000000000000e+001
	C <sub>3</sub> H <sub>6</sub> O <sub>3</sub>	90.0779	000090-02-3	2-HYDROXYPROPIONIC ACID	1.6000000000000000e+001	2.2600000000000000e+001	4.6000000000000000e+000
	C <sub>3</sub> H <sub>6</sub> O <sub>3</sub>	90.0779	010328-41-7	D-LACTIC ACID	5.5000000000000000e+001	2.2600000000000000e+001	-3.0340000000000000e+001

**Identical Chemicals**

**Mismatches**



Journal

**SAR and QSAR in Environmental Research** >

Volume 27, 2016 - Issue 11: 17th International Conference on QSAR in Environmental and Health Sciences (QSAR 2016) - Part II. Guest Editors: C.G. Barber and G.J. Myatt

258

Views

4


CrossRef citations

16

Altmetric


Articles

## An automated curation procedure for addressing chemical errors and inconsistencies in public datasets used in QSAR modelling\$

K. Mansouri, C. M. Grulke, A. M. Richard, R. S. Judson & A. J. Williams 





Pages 911-937 | Received 03 Sep 2016, Accepted 24 Oct 2016, Published online: 25 Nov 2016

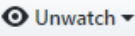


 Download citation


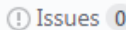

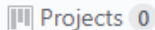


 <http://dx.doi.org/10.1080/1062936X.2016.1253611>

 Check for updates

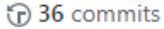
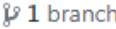



# OPERA on GitHub

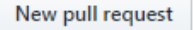
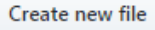
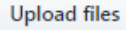


 This repository Search Pull requests Issues Marketplace Gist   


kmansouri / OPERA  1  1  0

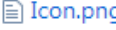

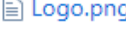
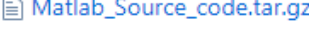
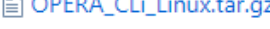
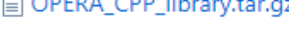
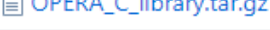
 Code  Issues 0  Pull requests 0  Projects 0  Wiki  Insights

Command line application providing QSAR models predictions as well as applicability domain and accuracy assessment for physicochemical properties and environmental fate endpoints.

 36 commits  1 branch  0 releases  1 contributor  MIT

Branch: master  New pull request  Create new file  Upload files  Find file  Clone or download

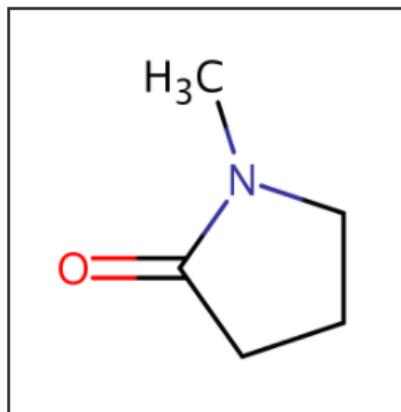
 kmansouri committed on GitHub OPERA 1.2 Windows Latest commit 731deaf on May 19

 Icon.png	OPERA 1.2 icon	3 months ago
 LICENSE	Initial commit	9 months ago
 Logo.png	Added logo and icon	9 months ago
 Matlab_Source_code.tar.gz	OPERA 1.2 MATLAB source code	3 months ago
 OPERA_CLI_Linux.tar.gz	OPERA 1.2 Linux	3 months ago
 OPERA_CPP_library.tar.gz	OPERA 1.2 C++ Library	3 months ago
 OPERA_C_library.tar.gz	OPERA 1.2 C Library	3 months ago

<https://github.com/kmansouri/OPERA.git>

# N-Methyl-2-pyrrolidone

872-50-4 | DTXSID6020856



## Model Results

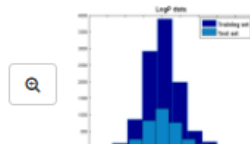
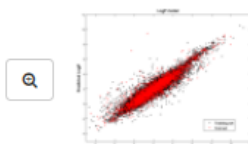
Predicted value: -0.300

Global applicability domain: **Inside** ⓘ

Local applicability domain index: 0.88 ⓘ

Confidence level: 0.81 ⓘ

## Model Performance



## Weighted KNN model

QMRf

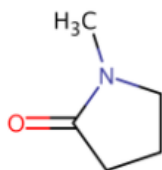
5-fold CV (75%)

Training (75%)

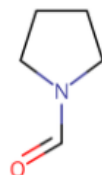
Test (25%)

Q2	RMSE	R2	RMSE	R2	RMSE
0.85	0.69	0.88	0.67	0.88	0.78

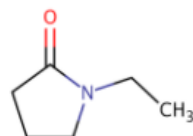
## Nearest Neighbors from the Training Set



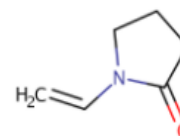
N-Methyl-2-pyrrolidone  
Measured: -0.380  
Predicted: -0.300



Prolinal  
Measured: -0.320  
Predicted: -0.226



1-Ethyl-2-pyrrolidinone  
Measured: -0.0400  
Predicted: -0.211



N-Vinyl-2-pyrrolidone  
Measured: 0.37  
Predicted: 2.65e-01

# QSAR Modeling Reporting Format

LogP (00000002).pdf - Adobe Acrobat Pro

File Edit View Window Help

Create [Icons]


Customize [Icon]

1 / 11 [Navigation Icons] 125% [Zoom]

Tools Fill & Sign Comment

**Bookmarks**

- 1.QSAR identifier
  - 1.1.QSAR identifier (title)
  - 1.2.Other related models
  - 1.3.Software coding the model
- 2.General information
  - 2.1.Date of QMRF
  - 2.2.QMRF author(s) and contact details
  - 2.3.Date of QMRF update(s)
  - 2.4.QMRF update(s)
  - 2.5.Model developer(s) and contact details
  - 2.6.Date of model development and/or publication
  - 2.7.Reference(s) to main scientific papers and/or software package
  - 2.8.Availability of information about the model
  - 2.9.Availability of another QMRF for exactly the same model
- 3.Defining the endpoint - OECD Principle 1
  - 3.1.Species
  - 3.2.Endpoint
  - 3.3.Comment on endpoint
  - 3.4.Endpoint units
  - 3.5.Dependent variable
  - 3.6.Experimental protocol
  - 3.7.Endpoint data quality and variability
- 4.Defining the algorithm - OECD Principle 2
  - 4.1.Type of model
  - 4.2.Explicit algorithm
  - 4.3.Descriptors in the model
  - 4.4.Descriptor selection
  - 4.5.Algorithm and descriptor generation
  - 4.6.Software name and version for descriptor generation
  - 4.7.Chemicals/Descriptors ratio
- 5.Defining the applicability domain - OECD Principle 3
  - 5.1.Description of the applicability domain of the model

	<b>QMRF identifier (JRC Inventory): To be entered by JRC</b>
	<b>QMRF Title:</b> LogP: Octanol-water partition coefficient prediction from the NCCT Models Suite.
	<b>Printing Date:</b> Apr 25, 2016

**1.QSAR identifier**

**1.1.QSAR identifier (title):**  
LogP: Octanol-water partition coefficient prediction from the NCCT Models Suite.

**1.2.Other related models:**  
No related models

**1.3.Software coding the model:**  
NCCT\_models V1.02  
Suite of QSAR models to predict physico-chemical properties and environmental fate of organic chemicals  
Kamel Mansouri (mansouri.kamel@epa.gov; mansourikamel@gmail.com);  
<https://comptox.epa.gov/dashboard/>

PaDEL descriptors V2.21  
Open source software to calculate molecular descriptors and fingerprints.  
Chun Wei Yap (phayapc@nus.edu.sg)  
<http://padel.nus.edu.sg/software/padeldescriptor>

MATLAB  
MATrix LABoratory is a multi-paradigm numerical computing environment and fourth-generation

# Toxicity Values

## Bioavailability Metric

Download as:

TSV

Excel

Exposure Limit

Point Of Departure

Regulatory Toxicity Value

Effect Level

Misc Hazard Information

Screening Level

Uncertainty Factor

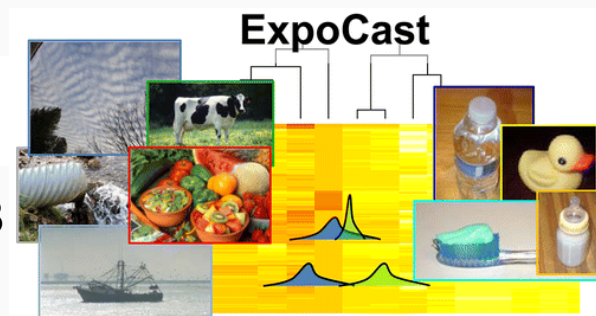
Type	Subtype	Value	Units	Study Type	Exposure Route	Study Duration	Species	Media	Details	Source
<a href="#">TD50</a>	-	20.7	mM/kg-day	-	-	-	-	-	<a href="#">DSSTox C...</a>	<a href="#">ACToR</a>
<a href="#">TD50</a>	-	2050	mg/kg-day	-	-	-	-	-	<a href="#">DSSTox C...</a>	<a href="#">ACToR</a>
<a href="#">LEL</a>	systemic	619	mg/kg-day	subchronic	oral	subchronic	mouse	-	<a href="#">Study ID: ...</a>	<a href="#">ToxRefDB</a>
<a href="#">NEL</a>	systemic	277	mg/kg-day	subchronic	oral	subchronic	mouse	-	<a href="#">Study ID: ...</a>	<a href="#">ToxRefDB</a>
<a href="#">LEL</a>	systemic	25.0	mg/kg-day	subchronic	oral	subchronic	dog	-	<a href="#">Study ID: ...</a>	<a href="#">ToxRefDB</a>
<a href="#">NEL</a>	systemic	25.0	mg/kg-day	subchronic	oral	subchronic	dog	-	<a href="#">Study ID: ...</a>	<a href="#">ToxRefDB</a>
<a href="#">LEL</a>	systemic	173	mg/kg-day	chronic	oral	chronic	mouse	-	<a href="#">Study ID: ...</a>	<a href="#">ToxRefDB</a>
<a href="#">NEL</a>	systemic	115	mg/kg-day	chronic	oral	chronic	mouse	-	<a href="#">Study ID: ...</a>	<a href="#">ToxRefDB</a>
<a href="#">LEL</a>	systemic	678	mg/kg-day	chronic	oral	chronic	rat	-	<a href="#">Study ID: ...</a>	<a href="#">ToxRefDB</a>
<a href="#">NEL</a>	systemic	283	mg/kg-day	chronic	oral	chronic	rat	-	<a href="#">Study ID: ...</a>	<a href="#">ToxRefDB</a>
<a href="#">LEL</a>	systemic	1230	mg/kg-day	subacute	oral	subacute	rat	-	<a href="#">Study ID: ...</a>	<a href="#">ToxRefDB</a>
<a href="#">NEL</a>	systemic	493	mg/kg-day	subacute	oral	subacute	rat	-	<a href="#">Study ID: ...</a>	<a href="#">ToxRefDB</a>
<a href="#">LEL</a>	systemic	2130	mg/kg-day	subacute	oral	subacute	mouse	-	<a href="#">Study ID: ...</a>	<a href="#">ToxRefDB</a>
<a href="#">NEL</a>	systemic	920	mg/kg-day	subacute	oral	subacute	mouse	-	<a href="#">Study ID: ...</a>	<a href="#">ToxRefDB</a>



## National Health and Nutrition Examination Survey

**High-Throughput Models for Exposure-Based Chemical Prioritization in the ExpoCast Project**

*Environ. Sci. Technol.*, **2013**, 47 (15), pp 8479–8488



Product & Use Categori...

Chemical Weight Fraction

Chemical Functional Use

Monitoring Data

Exposure Predictions

Download as:

TSV

Excel

National Health and Nutrition Examination Survey (NHANES) Inferences

(mg/kg-bw/day)

	Lower 95th Limit	Upper 95th Limit	Median
Ages 6-11	3.80e-05	4.92e-05	4.33e-05
Ages 12-19	2.55e-05	3.38e-05	2.93e-05
Ages 20-65	2.79e-05	3.27e-05	3.02e-05
Ages 65+	1.91e-05	2.31e-05	2.10e-05
BMI > 30	2.38e-05	2.74e-05	2.55e-05
BMI < 30	3.02e-05	3.30e-05	3.16e-05
Repro. Age Females	2.83e-05	3.31e-05	3.06e-05
Females	2.58e-05	3.03e-05	2.80e-05
Males	2.94e-05	3.37e-05	3.15e-05
Total	2.86e-05	3.08e-05	2.97e-05

Chemical Properties

Env. Fate/Transport

Synonyms

External Links

Toxicity Values (Beta)

**Exposure**

Bioassays

Similar Molecules (Beta)

Literature

Comments



# Product Composition Details

## Chemical Weight Fractions

Download as:

<a href="#">Product Name</a>	<a href="#">Product Use Category</a>	<a href="#">Minimum Weight Fraction</a>	<a href="#">Maximum Weight Fraction</a>	<a href="#">Data Type</a>	<a href="#">Source</a>
<a href="#">citristrip canadian stripping...</a>	<a href="#">home maintenance: stripper</a>	0.65	0.7	MSDS	<a href="#">Retail Product Categories/...</a>
<a href="#">citristrip stripping gel qog7...</a>	<a href="#">home maintenance: stripper</a>	0.4	0.55	MSDS	<a href="#">Retail Product Categories/...</a>
<a href="#">gumout 2 part professional...</a>	<a href="#">auto products: auto fluids a...</a>	0.3	0.4	MSDS	<a href="#">Retail Product Categories/...</a>
<a href="#">minwax water based wipe ...</a>	<a href="#">home maintenance: finish</a>	0.06	0.06	MSDS	<a href="#">Retail Product Categories/...</a>
<a href="#">10-02199- calico tip &amp; glue...</a>	<a href="#">personal care: nail polish r...</a>	0.01	0.05	MSDS	<a href="#">Retail Product Categories/...</a>
<a href="#">artificial nail remover 728 1</a>	<a href="#">personal care: nail polish r...</a>	0.01	0.05	MSDS	<a href="#">Retail Product Categories/...</a>
<a href="#">calico tip &amp; glue remover 1</a>	<a href="#">personal care: nail polish r...</a>	0.01	0.05	MSDS	<a href="#">Retail Product Categories/...</a>
<a href="#">kiss nail remover 1</a>	<a href="#">personal care: nail polish r...</a>	0.01	0.05	MSDS	<a href="#">Retail Product Categories/...</a>
<a href="#">waterborne clear wood fini...</a>	<a href="#">home maintenance: finish</a>	0.01	0.05	MSDS	<a href="#">Retail Product Categories/...</a>

[Chemical Properties](#)

[Env. Fate/Transport](#)

[Toxicity Values \(Beta\)](#)

[ADME \(Beta\)](#)

[Exposure](#)

[Bioassays](#)

[Similar Molecules \(Beta\)](#)

[Synonyms](#)

[Literature](#)

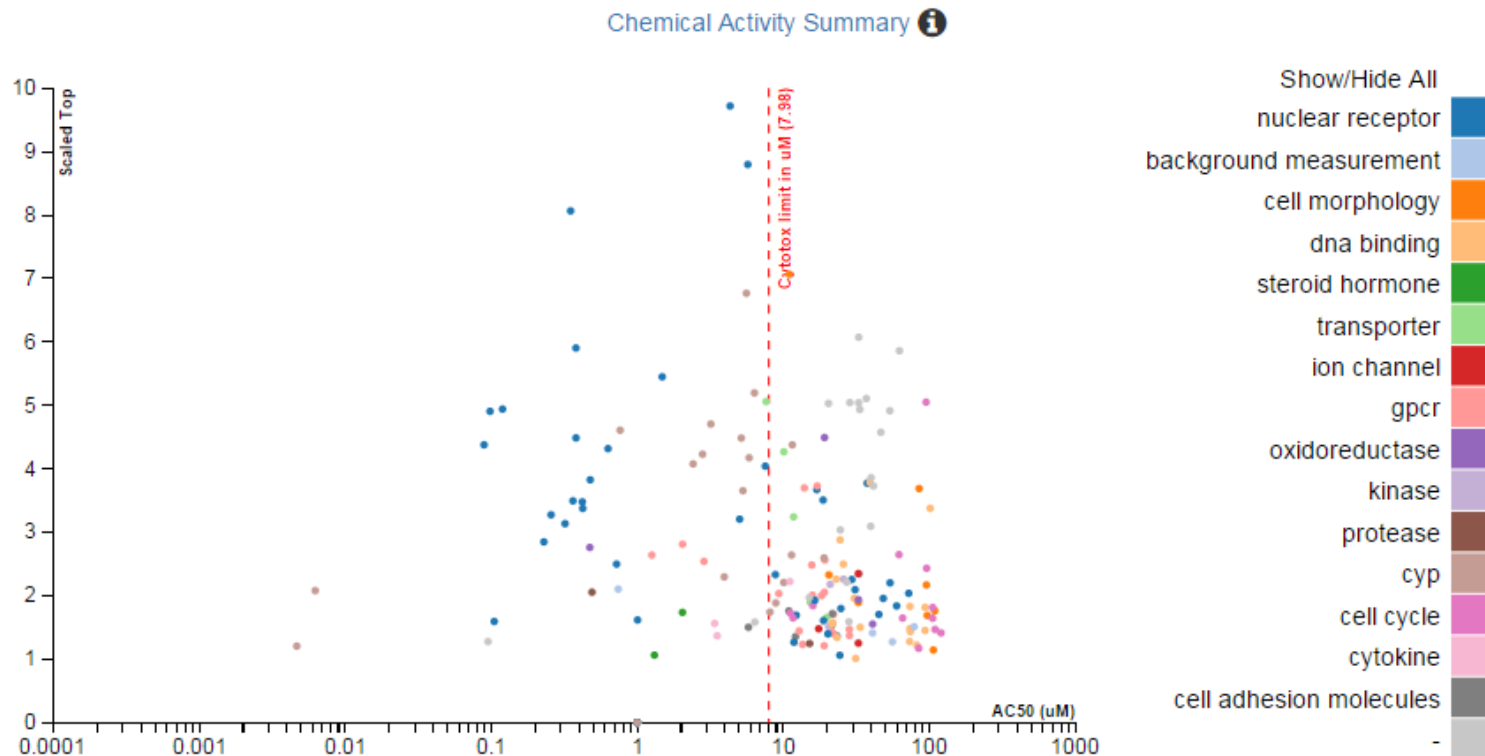
[External Links](#)

# ToxCast and Tox21 Bioassays

## ToxCast

PubChem

ScrubChem (Beta)



Chemical Properties

Env. Fate/Transport

Synonyms

External Links

Toxicity Values (Beta)

Exposure

Bioassays

Similar Molecules (Beta)

Literature

Comments

# ToxCast and Tox21 Bioassays

Download as:

TSV

Excel

Show:

Inactive

Background

Assay Name	Hit Call	Top	Scaled Top	AC50	log AC50 ↓	Intended Target Family
<a href="#">APR_Hepat_CellLoss_48hr_dn</a>	ACTIVE	1.41	1.41	120	2.08	cell cycle
<a href="#">APR_HepG2_MitoMass_24h_dn</a>	ACTIVE	0.874	1.76	109	2.04	cell morphology
<a href="#">APR_HepG2_OxidativeStress_24h_up</a>	ACTIVE	1.20	1.47	110	2.04	cell cycle
<a href="#">APR_Hepat_DNADamage_48hr_up</a>	ACTIVE	1.84	1.14	107	2.03	cell morphology
<a href="#">APR_HepG2_CellLoss_24h_dn</a>	ACTIVE	1.20	1.81	106	2.02	cell cycle
<a href="#">APR_HepG2_OxidativeStress_72h_up</a>	ACTIVE	1.80	1.65	106	2.02	cell cycle
<a href="#">ATG_HSE_CIS_up</a>	ACTIVE	1.59	3.38	102	2.01	dna binding

Chemical Properties

Env. Fate/Transport

Synonyms

External Links

Toxicity Values (Beta)

Exposure

Bioassays

Similar Molecules (Beta)


Literature

Comments

# Names and Identifiers

Found 40 synonyms

Legend: Valid Synonyms Good Synonyms Other Synonyms

 Copy all Synonyms

N-Methyl-2-pyrrolidone

1-Methylpyrrolidin-2-one

2-Pyrrolidinone, 1-methyl-

872-50-4 Active CAS-RN

2-Pyrrolidinone, 1-methyl-

1-Methyl-2-pyrrolidinone

1-Methyl-2-pyrrolidon

1-Methyl-2-pyrrolidone

1-Methyl-5-pyrrolidinone

1-Methylazacyclopentan-2-one

1-Methylpyrrolidone

1-metil-2-pirrolidona

2-Pyrrolidinone, 1-methyl-

2-Pyrrolidone, 1-methyl-

Microposit 2001

M-Pyrol

# Integrated Literature Searching

Google Scholar

PubMed Abstract S...

PubChem Articles

PubChem Patents

Select Term:

Hazard

AND

RfD OR reference dose



"RfD OR reference dose" AND "Hazard" AND "872-50-4" OR "N-Methyl-2-py



Scholar

About 22 results (0.05 sec)

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

[Solvent Substitution: An Analysis of Comprehensive \*\*Hazard\*\* Screening Indices](#)

M Debia, D Bégin, M Gérin - *Journal of occupational and ...*, 2011 - oeh.tandfonline.com

... Reliability and weighting factors were not used. Armenti and Moure-Eraso(22) used the same FHS index, using six **hazard** categories, to compare replacement options. ...

**872-50-4 N-Methyl-2-pyrrolidone** 45.4 13 1.46E-03 11 4.55 6 20 4 ...

[Related articles](#) [All 4 versions](#) [Cite](#) [Save](#)

[Effects of 'inactive'ingredients on bees](#)

CA Mullin - *Current Opinion in Insect Science*, 2015 - Elsevier

... tallow amines, organosilicone ethoxylates and co-solvents such as **N-methyl-2-pyrrolidone** (NMP) [7 ... of tools to environmentally monitor residues and determine their relative **hazard** to bees. ... Defining a benchmark or **reference dose**, particularly for bee viruses [ 51], for what can ...

[Cited by 9](#) [Related articles](#) [All 2 versions](#) [Web of Science: 6](#) [Cite](#) [Save](#) [More](#)

[Biological monitoring: exquisite research probes, risk assessment, and routine exposure measurement](#)

WM Draper - *Analytical chemistry*, 2001 - ACS Publications

ADVERTISEMENT. ...

[Cited by 19](#) [Related articles](#) [All 5 versions](#) [Web of Science: 10](#) [Cite](#) [Save](#)

[Evolution of chemical-specific adjustment factors \(CSAF\) based on recent international experience; increasing utility and facilitating regulatory acceptance](#)

VS Bhat, ME Meek, M Valcke, C English... - *Critical Reviews in ...*, 2017 - Taylor & Francis

... As shown in the right side of Figure 2, uncertainty is taken into consideration at all tiers of **hazard** assessment. ... For decades, developing "safe doses" such as the oral **reference dose** (RfD), the inhaled reference concentration (RfC) or the acceptable or tolerable daily intake (ADI ...

[All 3 versions](#) [Cite](#) [Save](#)

Chemical Properties

Env. Fate/Transport

Toxicity Values (Beta)

ADME (Beta)

Exposure

Bioassays

Similar Molecules (Beta)

Synonyms

Literature

External Links

# Integrated Literature Searching

Google Scholar

PubMed Abstract ...

PubChem Articles

PubChem Patents

Select Term:

Select a Query Term

Select a Query Term

Hazard

Fate and Transport

Metabolism/PK/PD

Chemical Properties

Exposure

Mixtures

Male Reproduction

Androgen Disruption

Female Reproduction

GeneTox

Embryo and embryonic development

Child (infant through adolescent)

Dust and Exposure

ng Articles

rrolidone" OR "N-methylpyrrolidone"

Search and Count

Chemical Properties

Env. Fate/Transport

Toxicity Values (Beta)

ADME (Beta)

Exposure

Bioassays

Similar Molecules (Beta)

Synonyms

Literature

External Links

# Integrated Literature Searching

Google Scholar

PubMed Abstract Sifter

PubChem Articles

PubChem Patents

## Edit the Query Before Retrieving Articles

("872-50-4" OR "N-Methyl-2-pyrrolidone" OR "N-methylpyrrolidone") AND (exposure OR near-field OR far-field OR SHEDS[tiab] AND ENVIRONMENTAL MONITORING)


1

0

0	0	0	0	24078144	2013	Biological monitoring and health effects of low-level exposure to N-methyl-2-pyrrolidone: a cross-sectional study.
0	0	0	0	23337464	2013	Biomonitoring of exposure to N-methyl-2-pyrrolidone in workers of the automobile industry.
0	0	0	0	19875680	2009	Quantitative risk analysis for N-methyl pyrrolidone using physiologically based pharmacokinetic and benchmark dos...
0	0	0	0	16897094	2006	Human volunteer study on the influence of exposure duration and dilution of dermally applied N-methyl-2-pyrrolidon...

Record:   1 of 16  

Title: [Not Available].

**Abstract:** N-Ethyl-2-pyrrolidone (NEP), a polar aprotic solvent, is used in many applications as substitute for the structural analogue N-methyl-2-pyrrolidone (NMP), e. g. for surface coatings, in cleaning agents and paint strippers. Monitoring studies indicate that individuals within the general public, without occupational exposure, may be exposed to NEP to an extent, which is comparable to NMP. As NMP, NEP presents a potential health hazard due to its developmental toxicity and teratogenicity. Exposure to NEP can be quantified by the determination of the excretion of its urinary metabolites 5-Hydroxy-N-ethyl-2-pyrrolidone (5-HNEP) and 2-Hydroxy-N-ethylsuccinimide (2-HESI). For the derivation of HBM values, the german Human Biomonitoring Commission (HBM commission) evaluated different toxicological endpoints and finally decided on the BMDL05 and the BMD10 for the endpoint "reduced grasp intensity" of a subchronic feeding study with rats as point of departure (POD) for further procedural steps. The resulting HBM-I and HBM-II values for the sum of the metabolites 5-HNEP and 2-HESI in the urine of children are 10 resp. 25 mg/l and in the urine of adults are 15 resp. 40 mg/l. If the HBM values are exceeded, a check-up will be necessary at first. Measurements above the HBM-II value give cause for concern, especially for pregnant women. Air meas  ents to determine the source of exposure can be useful. The possibility of skin absorption from use of

Chemical Properties

Env. Fate/Transport

Toxicity Values (Beta)

ADME (Beta)

Exposure

Bioassays

Similar Molecules (Beta)

Synonyms

Literature

External Links



# Pubmed Abstract Sifter

Select Term:

Exposure ▼

Retrieve Articles

0 Articles

Edit the Query Before Retrieving Articles

("872-50-4" OR "N-Methyl-2-pyrrolidone" OR "N-methylpyrrolidone") AND (exposure OR near-field OR far-field OR SHEDS[tiab] AND ENVIRONMENTAL MONITORING)

Add additional query terms to filter abstracts:

developmental

reproductive

occupational

Search and Count

dev...	rep...	occ...	Total	PMID	Pu...	Title
0	0	1	1	24078144	2013	Biological monitoring and health effects of low-level exposure to N-methyl-2-pyrrolidone: a cross-sectional study.
1	0	0	1	23337464	2013	Biomonitoring of exposure to N-methyl-2-pyrrolidone in workers of the automobile industry.
0	0	2	2	19875680	2009	Quantitative risk analysis for N-methyl pyrrolidone using physiologically based pharmacokinetic and benchmark dos...
0	0	1	1	16897094	2006	Human volunteer study on the influence of exposure duration and dilution of dermally applied N-methyl-2-pyrrolidon...
0	0	0	0	16362322	2005	Ambient monitoring and biomonitoring of workers exposed to N-methyl-2-pyrrolidone in an industrial facility

Record:   1 of 16  

Title: Stillbirth after occupational exposure to N-methyl-2-pyrrolidone. A case report and review of the literature.

**Abstract:** N-methyl-2-pyrrolidone is a solvent that is increasingly used in a variety of industries, including petroleum refining, microelectronics, pesticide formulation, and veterinary medicine. Animal studies have demonstrated fetotoxic effects after maternal exposure to doses that have minimal to no adverse effect on the mothers. The fetotoxicity comprises resorption, stillbirth, and low birthweight and delayed ossification in surviving young. We report a human case of intrauterine growth retardation followed by fetal demise at 31 weeks gestation. The mother was a laboratory worker with no other apparent risk factors, who sustained **occupational** exposure to N-methyl-2-pyrrolidone throughout the first trimester of pregnancy. Laboratory work and solvent exposure have both previously been associated with adverse **reproductive** outcomes. Laboratories and other industries that use suspected **reproductive** toxins should have **reproductive** health policies in place that allow for decision-making based on toxicologic review, exposure assessment, and medical evaluation. These policies should allow for voluntary removal of prospective parents until environmental assessment and controls are instituted.

Chemical Properties

Env. Fate/Transport

Toxicity Values (Beta)

ADME (Beta)

Exposure

Bioassays

Similar Molecules (Beta)














Synonyms

Literature













External Links

# Links to Other Resources











## General

-  EPA Substance Re...
-  NIST Chemistry W...
-  Household Product...
-  PubChem
-  Chempider
-  CPCat
-  DrugBank
-  HMDB
-  Wikipedia
-  MSDS Lookup
-  ChEMBL
-  Chemical Vendors
-  Consumer Product...






## Toxicology

-  ACToR
-  DrugPortal
-  CCRIS
-  ChemView
-  CTD
-  EPA
-  GHS
-  HSDB
-  ToxCast Dashboar...
-  LactMed
-  ACToR PDF Report
-  International Toxicit...

## Publications

-  Toxline
-  Environmental Heal...
-  NIEHS
-  National Toxicology...
-  Google Books
-  Federal Register
-  Regulations.gov
-  Springer Materials
-  BioCaddie DataMed
-  RSC Publications

## Analytical

-  National Environme...
-  MONA: MassBank ...
-  Tox21 Analytical Data
-  RSC Analytical Abs...
-  FOR-IDENT

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.


# Example External Links...

U.S. Department of Health & Human Services [www.hhs.gov](http://www.hhs.gov)

National Institutes of Health

English | Español

## ChemHAT.org



**ECHA**  
EUROPEAN CHEMICALS AGENCY

Search the ECHA Website

About Us Regulations Addressing Chemicals of Concern Information on Chemicals Chemicals in our Life Support


N-methyl-2-pyrrolidone  
CAS: 872-50-4

ECHA > Substance Information


### How to use Substance Information

Infocards are automatically generated for each substance.


**Acute**

 Irritant  
irritates eye.

**Chronic**

 Birth defects  
developmental effects  
biological effects

**Other**

 Other  
series  
ingestion

**1-methyl-2-pyrrolidone**

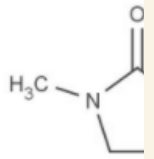
Other names: Regulated

**Substance identification**

EC / List no.: 212-001-000-000

CAS no.: 872-50-4

Mol. formula: C<sub>5</sub>H<sub>9</sub>NO



**About this substance**

This substance is not classified as hazardous.

This substance is used by consumers, by professional workers (widespread uses), in formulation or re-

**MoNA - MassBank of North America**

Spectra Downloads Upload Help

Search...

Compound Classification (provided by ClassyFire)

- Kingdom: Organic compounds
- Superclass: Organoheterocyclic compounds
- Class: Pyrrolidines
- Subclass: N-alkylpyrrolidines
- Direct Parent: N-alkylpyrrolidines

**Compound Metadata**

Name	Value
total exact mass	99.0684
SMILES	[H]C([H])([H])N1C(=O)C([H])([H])C([H])([H])C1([H])[H]
SMILES	CN1CCCC1=O
molecular formula	C <sub>5</sub> H <sub>9</sub> NO
InChIKey	SECXISVLQFMRJM-UHFFFAOYSA-N
InChI	InChI=1S/C5H9NO/c1-6-4-2-3-5(6)/h2-4H2,1H3

- Build out definitive “lists” of chemicals
  - Algal toxins
  - Poly/perfluorinated chemicals
  - Pesticides
  - Toxcast screening chemical collection
  - Public data sets:
    - NORMAN Network data collections
    - MASSBANK data

# The Collection of Lists

## Chemistry Dashboard

Aa v

Aa

Aa ^

[NIOSH Skin Notation Profiles](#)

57

The NIOSH skin notations relies on multiple skin notations to provide users a warning on the direct, systemic, and sensitizing effects of exposures of the skin to chemicals.

[NORMAN Collaborative Trial 2015 Targets and Suspects](#)

732

**NORMAN**CT15 is a compilation of all target and suspect substances reported by participants in the **NORMAN** Collaborative Trial on Non-target Screening, run by the **NORMAN** Network and described in Schymanski et al 2015, DOI: 10.1007/s00216-015-8681-7

## Chemistry Dashboard | NORMANEWS

Aa v

Aa

Aa ^

### NormaNEWS: Norman Early Warning System

Search NORMANEWS Chemicals

Q

#### List Details

**Description:** The Norman Early Warning System (NormaNEWS) is a pilot network designed to investigate the spatial and temporal distribution of newly identified contaminants of emerging concern in environmental samples through performing retrospective suspect screening on HRMS data acquired using different instrumental platforms and data processing software. The NormaNEWS pilot study was performed through recruiting eight reference laboratories with available archived HRMS data with the goal of exploring the potential of an early warning network to rapidly establish the occurrence of newly-identified contaminants of emerging concern across Europe and beyond, through the use of retrospective suspect screening employing HRMS. The pilot study was referred to as the Norman Early Warning System, abbreviated to NormaNEWS.

**Number of Chemicals:** 131

Sort Options v

Select/Deselect All

Download as:

TSV v

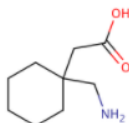
Excel v

SDF v

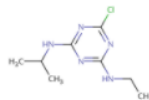
View Selected

■

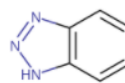
■



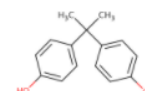
Gabapentin  
60142-96-3



Atrazine  
1912-24-9



1,2-Benzotriazole  
95-14-7



# 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

**Select Input Type(s)**

- ☒ 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	Structures	Intrinsic Properties
<input checked="" type="checkbox"/> Chemical Name	<input type="checkbox"/> Mol File	<input type="checkbox"/> Molecular Formula
<input checked="" type="checkbox"/> DTXSID	<input type="checkbox"/> SMILES	<input type="checkbox"/> Average Mass
<input type="checkbox"/> CAS-RN	<input type="checkbox"/> InChI String	<input type="checkbox"/> Monoisotopic Mass
<input type="checkbox"/> InChIKey		<input type="checkbox"/> OPERA and TEST Model Predictions
<input type="checkbox"/> IUPAC Name		

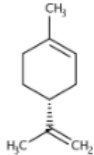

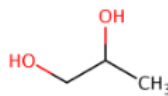
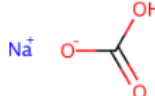
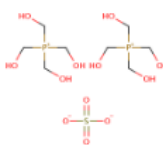
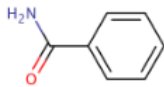



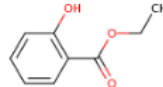
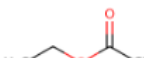

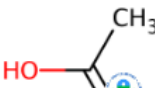
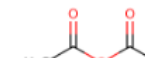
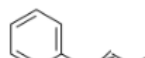
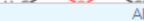
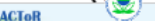
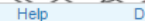
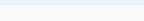
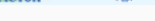
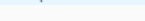


























# Access to associated data for review, modeling & download

**EPA** United States Environmental Protection Agency [Home](#) [Advanced Search](#)

## Chemistry Dashboard


[View Selected](#) [Hide Isotopes](#) [Hide Multicomponent Chemicals](#)   [Aa ▼](#) [Aa](#) [Aa ▲](#)

 <p>D-Limonene 5989-27-5</p> <input type="checkbox"/>	 <p>Potassium chloride 7447-40-7</p> <input type="checkbox"/>	 <p>1,2-Propylene glycol 57-55-6</p> <input type="checkbox"/>	 <p>Sodium bicarbonate 144-55-8</p> <input type="checkbox"/>	 <p>Tetrakis(hydroxymethyl)phospho... 55566-30-8</p> <input type="checkbox"/>
 <p>Benzamide 55-21-0</p> <input type="checkbox"/>	 <p>Methanol 67-56-1</p> <input type="checkbox"/>	 <p>1-Butanol 71-36-3</p> <input type="checkbox"/>	 <p>Propargyl alcohol 107-19-7</p> <input type="checkbox"/>	 <p>Ethyl salicylate 118-61-6</p> <input type="checkbox"/>
 <p><a href="#">About</a></p>	 <p><a href="#">Contact</a></p>	 <p><a href="#">Privacy</a></p>	 <p><a href="#">ACToR</a></p>	 <p><a href="#">DSSTox</a></p>
				
				
				
				
				
				
				
				
				
				


[About](#) [Contact](#) [Privacy](#) [ACToR](#) [DSSTox](#) [Accessibility](#) [Help](#) [Downloads](#)

# Crowdsourced Curation – HELP!

[https://comptox.epa.gov/dashboard/comments/public\\_index](https://comptox.epa.gov/dashboard/comments/public_index)

 United States  
Environmental Protection  
Agency

Home Advanced Search Batch Search Lists

Search All Data 

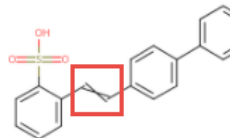
Chemistry Dashboard


Submit Comment Share Copy Aa Aa Aa

2,2'-[biphenyl]-4,4'

38775-22-3 | DTXSID7047

© Searched by DSSTox\_Substance




New Comment 

Comment

Type your comments here...


Email address


Enter your email address


☐ I'm not a robot 

reCAPTCHA  
Privacy - Terms

Submit

Find All Chemicals 





Presence in Lists

Record Information

Chemical Properties

Env. Fate/Transport

Toxicity Values (Beta)

ADME (Beta)

Exposure

Bioassays

Similar Molecules (Beta)

Synonyms

Literature

# Crowdsourced Curation – **HELP!**

[https://comptox.epa.gov/dashboard/comments/public\\_index](https://comptox.epa.gov/dashboard/comments/public_index)

Chemical Properties	Env. Fate/Transport	Toxicity Values (Beta)	ADME (Beta)	Exposure	Bioassays	Similar Molecules (Beta)	Synonyms	Literature
External Links	Comments							
<div>Add A Comment</div>								

Comment from structure source: to my knowledge the stilbene-derived fluorescent whitening agents are all trans (E) isomers, as the cis (Z) isomers are not fluorescent (although they might undergo photo-isomerisation to the cis isomers under UV light, and clothing gets yellowish again then...) . Thus I would consider the E,E form the correct one, although I don't know whether it is synthesized in a way that x % of the technical product are actually the inactive E,Z (dunno if this is still active?) or Z,Z forms and thus the undefined stereo would be correct. In the environment E,E is partly transformed to E,Z (maybe also Z,Z) and thus both isomers occur, see: <http://pubs.acs.org/doi/pdf/10.1021/es960748a> CAS number on record was 27344-41-8 which is DTXSID6036467.

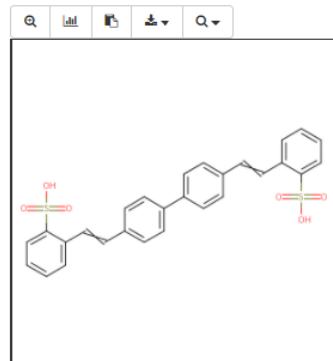
User comment posted 1 day ago

- The Question – is this a mixture of E/Z, E/E or Z/Z? Brighteners should be E/E...
- Thanks for helping us Emma!

# Curation is laborious work

38775-22-3 | DTXSID7047017

Ⓢ Searched by DSSTox\_Substance\_Id: Found 1 result for 'DTXSID7047017'.



- So, it's clear what the correct E/Z orientation is yes???

CAS Registry Number 38775-22-3

~160  ~17 

**C<sub>28</sub> H<sub>22</sub> O<sub>6</sub> S<sub>2</sub>**

Benzenesulfonic acid, 2,2'-([1,1'-biphenyl]-4,4'-diyl-di-2,1-ethenediyl) bis-

**Molecular Weight**

518.60

**Density (Predicted)**

Value: 1.414±0.06 g/cm<sup>3</sup> | Condition: Temp: 20 °C Press: 760 Torr

**pKa (Predicted)**

Value: -0.92±0.27 | Condition: Most Acidic Temp: 25 °C

**Other Names**

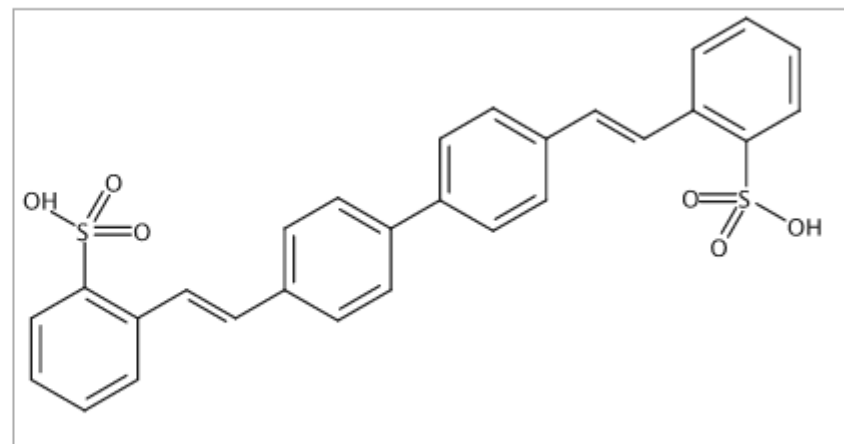
2,2'-([1,1'-Biphenyl]-4,4'-diyl-di-2,1-ethenediyl)bis[benzenesulfonic acid]

4,4'-Bis(2-sulfostyryl)biphenyl

4,4'-Bis[2-(2-sulfophenyl)ethenyl]biphenyl



C.I. 482200

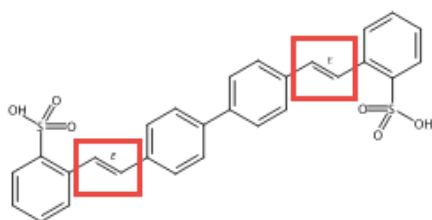
C.I. Fluorescent Brightener 351



# Curation is laborious work

1. **334756-45-5** 🔍

~5  ~1 



Double bond geometry as shown.

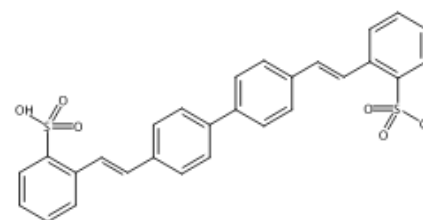
**C<sub>28</sub> H<sub>22</sub> O<sub>6</sub> S<sub>2</sub>**

Benzenesulfonic acid, 2,2'-((1E)-[1,1'-biphenyl]-4,4'-diylidene)-2,1-ethenediyl]bis- (9CI)

► **Key Physical Properties**

2. **38775-22-3** 🔍

~160  ~17 



**C<sub>28</sub> H<sub>22</sub> O<sub>6</sub> S<sub>2</sub>**

Benzenesulfonic acid, 2,2'-([1,1'-biphenyl]-4,4'-diylidene)-2,1-ethenediyl]bis-

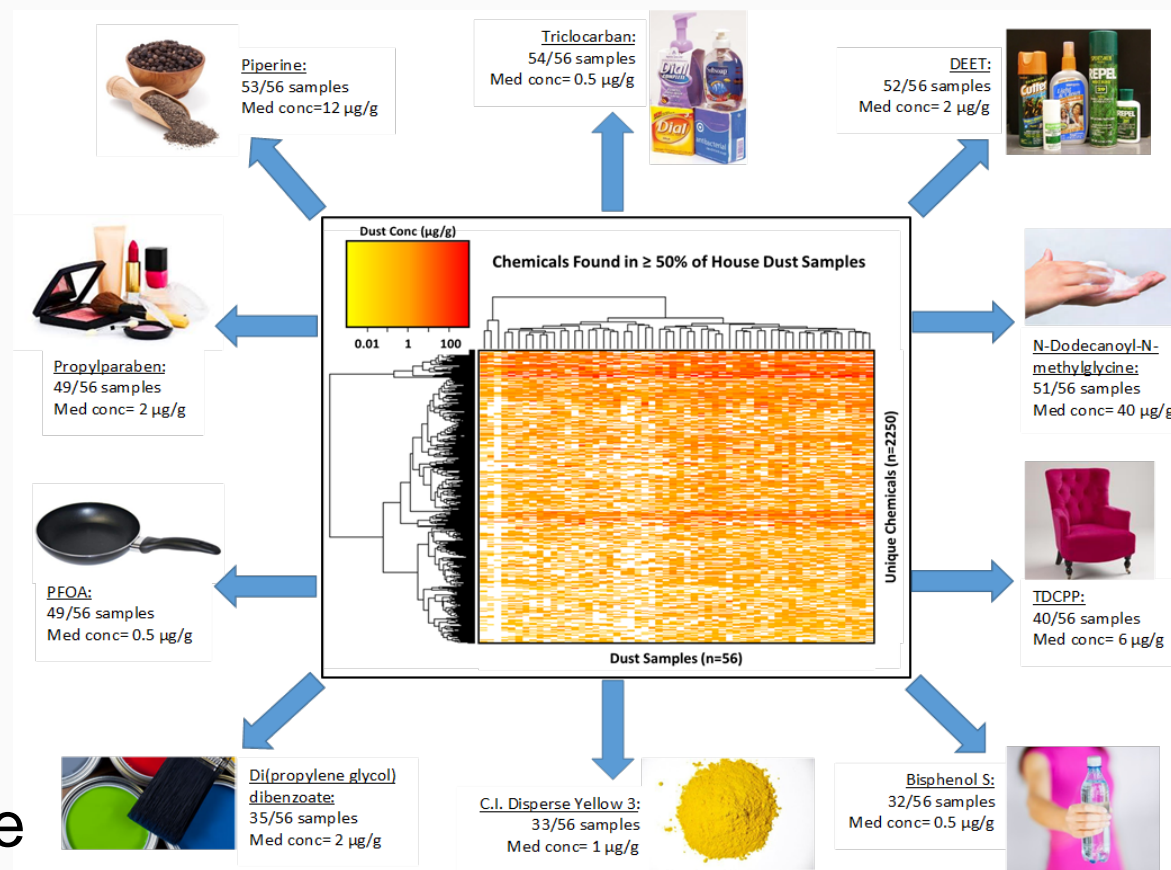
► **Key Physical Properties**

Regulatory Information

- I have a 1000 CAS Numbers (or Names) – is there data available?
  - Has any Toxcast data been run?
  - Are there Toxicity Data values available?
  - Are there predicted exposure data (via Expocast)?
  - Can I get predicted physchem data for my model?
- Identifying chemicals by analytical chemistry

# Suspect Screening Analysis Chemicals in House Dust

- Chemical confirmation from ToxCast
- Detection frequency
- External calibration for concentrations
- Cluster analysis (homes/chemicals)
- Chemical functional use

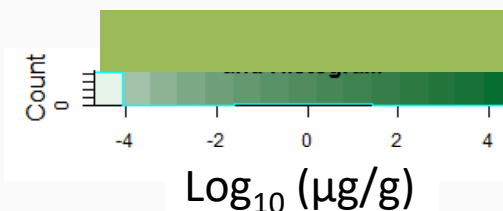
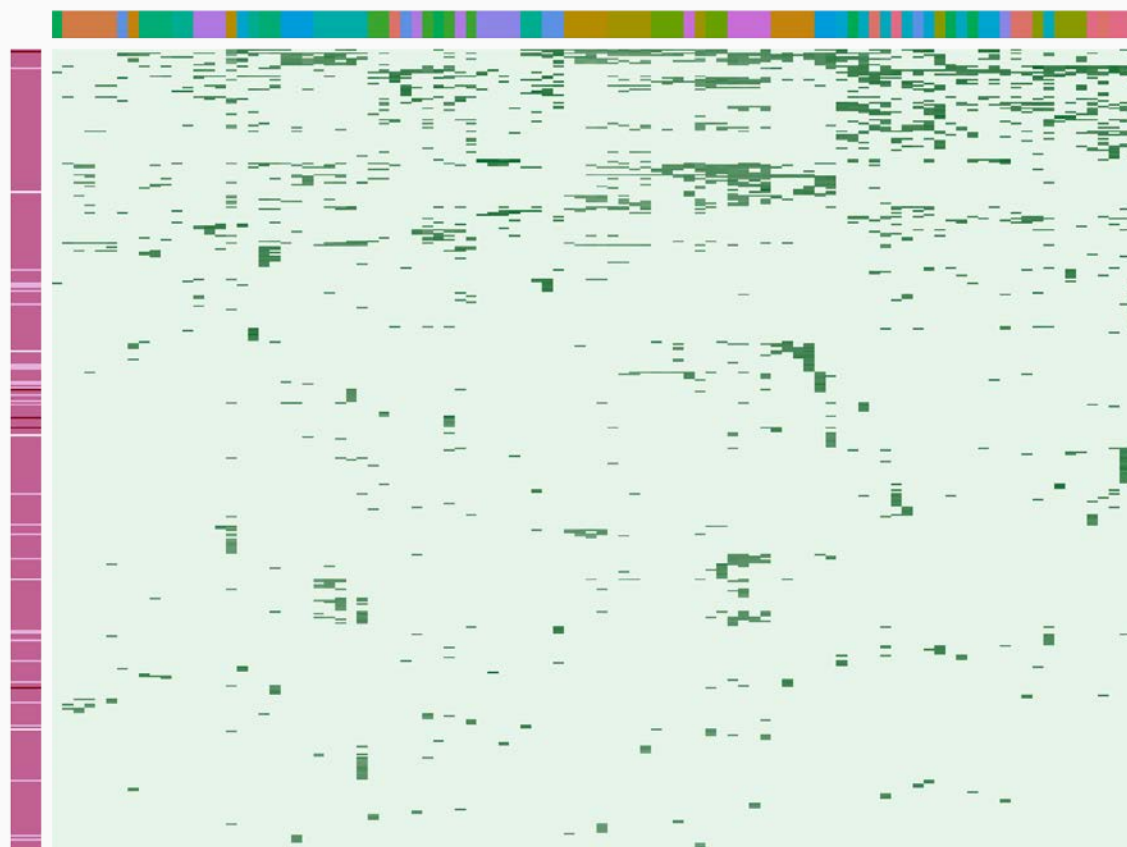




# Improving Exposure Estimates – Characterizing Commercial Products

423 ToxCast and/or Commonly Occurring Chemicals\*

100 Consumer Products and Articles of Commerce



- Air freshener
- Baby soap
- Carpet
- Carpet padding
- Cereals
- Cotton clothing
- Deodorant
- Fabric upholstery
- Glass cleaners
- Hand soap
- Indoor house paint
- Lipstick
- Plastic children's toys
- Shampoo
- Shaving cream
- Shower curtain
- Skin lotion
- Sunscreen
- Toothpaste
- Vinyl upholstery

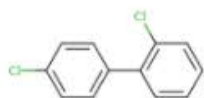
- GCXGC-MS with DCM Extraction
- 1606 tentative and confirmed chemical identifications

- Common Chemical (n>19)
- ToxCast
- Flame Retardant
- Potent ER

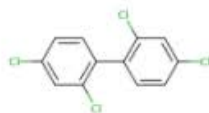
- Sometimes the simplest of questions are difficult to answer!
  - What is the list of CAS Numbers for all PCBs?
  - Can I get an SDF file of all PCBs?
  - Do you have predicted properties for all PCBs?
  - What toxicity data is available for individual PCBs?
  - Have you measured ToxCast data for any PCBs?
  - Can I get all PCBs listed in an Excel Spreadsheet?

# Chemical “Families”

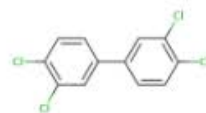
Download as: [TSV](#) [Excel](#) [SDF](#)



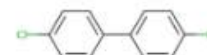
2,4'-Dichlorobiphenyl  
34863-43-7



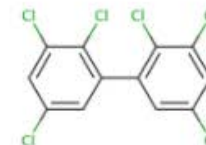
2,2',4,4'-Tetrachlorobiphenyl  
2437-79-8



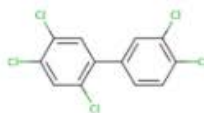
3,3',4,4'-Tetrachlorobiphenyl  
32598-13-3



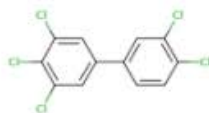
4,4'-Dichlorobiphenyl  
2050-68-2



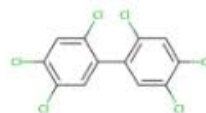
2,2',3,3',5,5'-Hexachlorobiphenyl  
35694-04-3



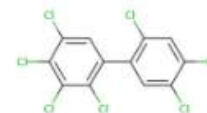
2,3',4,4',5-Pentachlorobiphenyl  
31508-00-6



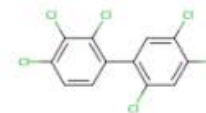
3,3',4,4',5-Pentachlorobiphenyl  
57465-28-8



2,2',4,4',5,5'-Hexachlorobiphenyl  
35065-27-1



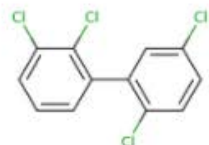
2,2',3,4,4',5,5'-Heptachlorobiphenyl  
35065-29-3



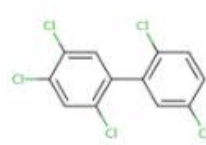
2,2',3,4,4',5'-Hexachlorobiphenyl  
35065-26-2



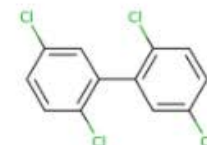
2,2',3,5,6-Pentachlorobiphenyl  
36379-99-6



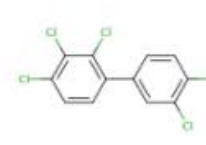
2,2',3,5'-tetrachlorobiphenyl  
41464-39-5



2,2,4,5,5'-Pentachlorobiphenyl  
37680-73-2



2,2',5,5'-Tetrachlorobiphenyl  
35693-99-3



2,3,3',4,4'-Pentachlorobiphenyl  
32598-14-4

# One click download

	A	B	C	D	E	F	G	H	I
1	DTXSID	CASRN	PREFERRED NAME	IUPAC NAME	MOLECULAR FORMULA	MONOISOTOPIC MASS	AVERAGE MASS	SMILES	INCHI KEY
2	DTXSID9074779	70362-45-7	PCB 045	2,2',3,6-Tetrachloro-1,1'-biphenyl	C12H6Cl4	289.92236	291.98001	<chem>C1C1=CC=C(C1)C(=C1C1)C1=CC=CC=C1C1</chem>	VHGHZT
3	DTXSID9074777	38444-73-4	PCB 019	2,2',6-Trichloro-1,1'-biphenyl	C12H7Cl3	255.96133	257.54001	<chem>C1C1=CC=CC(C1)=C1C1=C(C1)C=CC=C1</chem>	MXVJRBBC
4	DTXSID9074228	74472-39-2	2,3',4',5',6-Pentachlorobiphenyl	2,3',4',5',6-Pentachlorobiphenyl	C12H5Cl5	323.88339	326.42001	<chem>C1C1=CC(=CC(C1)=C1C1)C1=C(C1)C=CC=C1C1</chem>	WAZUWHG
5	DTXSID9074226	74472-37-0	2,3,4,4',5-Pentachlorobiphenyl	2,3,4,4',5-Pentachlorobiphenyl	C12H5Cl5	323.88339	326.42001	<chem>C1C1=CC=C(C=C1)C1=C(C1)C(C1)=C(C1)C(C1)=C1</chem>	SXZSFWHC
6	DTXSID9074224	74472-35-8	2,3,3',4,6-Pentachlorobiphenyl	2,3,3',4,6-Pentachlorobiphenyl	C12H5Cl5	323.88339	326.42001	<chem>C1C1=CC=CC(=C1)C1=C(C1)C(C1)=C(C1)C=C1C1</chem>	XGQBSVVY
7	DTXSID9074222	74472-33-6	2,3,3',6-Tetrachlorobiphenyl	2,3,3',6-Tetrachlorobiphenyl	C12H6Cl4	289.92236	291.98001	<chem>C1C1=CC=CC(=C1)C1=C(C1)C(C1)=CC=C1C1</chem>	WZNAMGY
8	DTXSID9074220	74338-23-1	2,3',5',6-Tetrachlorobiphenyl	2,3',5',6-Tetrachlorobiphenyl	C12H6Cl4	289.92236	291.98001	<chem>C1C1=CC(=CC(C1)=C1)C1=C(C1)C=CC=C1C1</chem>	HDULUCZR
9	DTXSID9074199	68194-11-6	2,3,4',5,6-Pentachlorobiphenyl	2,3,4',5,6-Pentachlorobiphenyl	C12H5Cl5	323.88339	326.42001	<chem>C1C1=CC=C(C=C1)C1=C(C1)C(C1)=CC(C1)=C1C1</chem>	ZDDZPDTV
10	DTXSID9074197	68194-08-1	2,2',3,4',6-Hexachlorobiphenyl	2,2',3,4',6-Hexachlorobiphenyl	C12H4Cl6	357.84442	360.85999	<chem>C1C1=CC(C1)=C(C(C1)=C1)C1=C(C1)C(C1)=CC=C1C1</chem>	RPPNJBZN
11	DTXSID9074195	68194-04-7	2,2',4,6-Tetrachlorobiphenyl	2,2',4,6-Tetrachlorobiphenyl	C12H6Cl4	289.92236	291.98001	<chem>C1C1=CC=C(C(C1)=C1)C1=C(C1)C=CC=C1C1</chem>	WVHNUGR
12	DTXSID9074193	60233-25-2	2,2',3,4',6-Pentachlorobiphenyl	2,2',3,4',6-Pentachlorobiphenyl	C12H5Cl5	323.88339	326.42001	<chem>C1C1=CC(C1)=C(C(C1)=C1)C1=C(C1)C(C1)=CC=C1</chem>	GOFFZTAP
13	DTXSID9074191	60145-23-5	2,2',3,4,4',5,6-Heptachlorobiphenyl	2,2',3,4,4',5,6-Heptachlorobiphenyl	C12H3Cl7	391.80544	395.31	<chem>C1C1=CC(C1)=C(C(C1)=C1)C1=C(C1)C(C1)=C(C1)C(C1)=C1</chem>	RXRLRYZU
14	DTXSID9074149	41411-61-4	2,2',3,4,5,6-Hexachlorobiphenyl	2,2',3,4,5,6-Hexachlorobiphenyl	C12H4Cl6	357.84442	360.85999	<chem>C1C1=C(C=C1)C1=C(C1)C(C1)=C(C1)C(C1)=C1C1</chem>	RUEIBQJF
15	DTXSID9074147	40186-70-7	2,2',3,3',4,5',6-Heptachlorobiphenyl	2,2',3,3',4,5',6-Heptachlorobiphenyl	C12H3Cl7	391.80544	395.31	<chem>C1C1=CC(C1)=C(C1)C(=C1)C1=C(C1)C(C1)=C(C1)C=C1C1</chem>	KJBDZJFS
16	DTXSID9074145	39635-33-1	3,3',4,5,5'-Pentachlorobiphenyl	3,3',4,5,5'-Pentachlorobiphenyl	C12H5Cl5	323.88339	326.42001	<chem>C1C1=CC(=CC(C1)=C1)C1=CC(C1)=C(C1)C(C1)=C1</chem>	MXVAYAXI
17	DTXSID9074143	38444-76-7	2,3',6-Trichlorobiphenyl	2,3',6-Trichlorobiphenyl	C12H7Cl3	255.96133	257.54001	<chem>C1C1=CC=CC(=C1)C1=C(C1)C=CC=C1C1</chem>	VQOFJPFY
18	DTXSID9074141	38380-05-1	2,2',3,3',4,6-Hexachlorobiphenyl	2,2',3,3',4,6-Hexachlorobiphenyl	C12H4Cl6	357.84442	360.85999	<chem>C1C1=CC=C(C1)C(=C1C1)C1=C(C1)C(C1)=C(C1)C=C1</chem>	OKBJVIVE
19	DTXSID9073599	65510-45-4	2,2',3,4,4'-Pentachlorobiphenyl	2,2',3,4,4'-Pentachlorobiphenyl	C12H5Cl5	323.88339	326.42001	<chem>C1C1=CC(C1)=C(C=C1)C1=C(C1)C(C1)=C(C1)C=C1</chem>	LACXVZHA
20	DTXSID9073541	52744-13-5	2,2',3,3',5,6-Hexachlorobiphenyl	2,2',3,3',5,6-Hexachlorobiphenyl	C12H4Cl6	357.84442	360.85999	<chem>C1C1=CC(=C(C1)C(C1)=C1)C1=C(C1)C(C1)=CC=C1C1</chem>	UUTNFLRS
21	DTXSID9073410	16606-02-3	2,4',5-Trichlorobiphenyl	2,4',5-Trichlorobiphenyl	C12H7Cl3	255.96133	257.54001	<chem>C1C1=CC=C(C=C1)C1=C(C1)C=CC(C1)=C1</chem>	VAHKBZSA
22	DTXSID80873557	36559-22-5	2,2',3,4'-Tetrachlorobiphenyl	2,2',3,4'-Tetrachlorobiphenyl	C12H6Cl4	289.92236	291.98	<chem>C1C1=CC(C1)=C(C=C1)C1=CC=CC(C1)=C1C1</chem>	ALFHIHDQ
23	DTXSID8074780	61798-70-7	PCB 131	2,2',3,3',4,6-Hexachlorobiphenyl	C12H4Cl6	357.84442	360.85999	<chem>C1C1=CC(C1)=C(C1)C(C1)=C1C1=CC=CC(C1)=C1C1</chem>	WDLTVNW
24	DTXSID8074239	74472-51-8	2,3,3',4,5,5',6-Heptachlorobiphenyl	2,3,3',4,5,5',6-Heptachlorobiphenyl	C12H3Cl7	391.80544	395.31	<chem>C1C1=CC(=CC(C1)=C1)C1=C(C1)C(C1)=C(C1)C(C1)=C1C1</chem>	ZUTDUGMH
25	DTXSID8074237	74472-49-4	2,2',3,4,5,6,6'-Heptachlorobiphenyl	2,2',3,4,5,6,6'-Heptachlorobiphenyl	C12H3Cl7	391.80544	395.31	<chem>C1C1=CC=CC(C1)=C1C1=C(C1)C(C1)=C(C1)C(C1)=C1C1</chem>	FGDPOTMH
26	DTXSID8074235	74472-47-2	2,2',3,4,4',5,6-Heptachlorobiphenyl	2,2',3,4,4',5,6-Heptachlorobiphenyl	C12H3Cl7	391.80544	395.31	<chem>C1C1=CC(C1)=C(C=C1)C1=C(C1)C(C1)=C(C1)C(C1)=C1C1</chem>	DJEUXBQA
27	DTXSID8074233	74472-44-9	2,3,3',4',5,6-Hexachlorobiphenyl	2,3,3',4',5,6-Hexachlorobiphenyl	C12H4Cl6	357.84442	360.85999	<chem>C1C1=CC=C(C=C1C1)C1=C(C1)C(C1)=CC(C1)=C1C1</chem>	ZAGRQXMH

# How Did We Do This? DSSTox

## ACToR-DSSTox Chemical Registration

[View/Edit a  
Single Record](#)

[Structure  
Search](#)

[Browse/Curate  
Records](#)

[Export DSSTox](#)

[Chemotypes](#)

[Manage  
Chemical Lists](#)

[Manage  
Property Data](#)

[Add Deleted  
Casrns](#)

Welcome, antony

[Logout](#)

Substance\_ID: DTXSID5024267

CAS: 1336-36-3

Name: Polychlorinated biphenyls

Substance Type: Mixture/Formulation

QC Level: DSSTox\_High

Data Source: STN(DSSTox)

QC Notes: biphenyl with multiple (unknown number) chlorines attached at unknown locations

Compound\_ID:

Chemical Shown:

No Structure

Private Notes:

Source of CAS-Compound:

Double Stereo:

Chiral Stereo:

Chemical Form:

► Synonyms (31)

► Other Cas (0)

▼ Successor Substances (209)

	CAS-RN	Relationship	Source	Struc	Casrn	Comments
●	32774-16-6	is a Representative Isomer of this	STN(DSSTox)	✓	■	structure shown 3,3',4,4',5,5'
●	2051-60-7	is a Representative Isomer of this	Public	■	■	
●	2051-61-8	is a Representative Isomer of this	Public	■	■	
●	2051-62-9	is a Representative Isomer of this	Public	■	■	
●	13029-08-8	is a Representative Isomer of this	Public	■	■	
●	16605-91-7	is a Representative Isomer of this	Public	■	■	
●	25569-80-6	is a Representative Isomer of this	Public	■	■	
●	33284-50-3	is a Representative Isomer of this	Public	■	■	

# Relationship Mappings

- Various relationship mappings can be established. To this point all are manual.

Relationship	Source
is a Representative Isomer of this	STN(DSSTox) ▼
is a Representative Component of this substance	
is a Mixture Component of this substance	
is a Monomer of this substance	
is an Active Ingredient of this substance	
is a Representative Isomer of this substance	
is a General Form of this substance	
is a Transformation Product of this substance	

- In progress – **metabolite** mappings (building metabolism competence into high-throughput assay)





The screenshot shows the EPA website's navigation bar with links for Environmental Topics, Laws & Regulations, and About EPA. Below this is the TSCA Chemical Substance Inventory page. The page has a sidebar with links to TSCA Inventory Home, About the Inventory, Access the Inventory, and Policy and Guidance. The main content area features the title "Chemical Substances of Unknown or Variable Composition, Complex Reaction Products and Biological Materials (UVCB Substance) on the TSCA Inventory" and a brief description: "This paper is a compendium of information related to the broad class of chemical substances referred to as UVCBs for the Toxic Substances Control Act (TSCA) Chemical Substance Inventory. These chemical substances cannot be represented by unique structures and molecular formulas."

Environmental Topics   Laws & Regulations   About EPA   Search EPA.gov

## TSCA Chemical Substance Inventory

CONTACT US   SHARE   [f](#)   [t](#)   [p](#)   [e](#)

- TSCA Inventory Home
- About the Inventory
- Access the Inventory
- Policy and Guidance

### Chemical Substances of Unknown or Variable Composition, Complex Reaction Products and Biological Materials (UVCB Substance) on the TSCA Inventory

This paper is a compendium of information related to the broad class of chemical substances referred to as UVCBs for the Toxic Substances Control Act (TSCA) Chemical Substance Inventory. These chemical substances cannot be represented by unique structures and molecular formulas.

- UVCB chemical examples
  - Surfactants with undefined composition
  - Petroleum Distillates
  - Gelatins, hydrozylates
  - Formaldehyde, reaction products with diethanolamine
  - Fatty acids, linseed-oil, compds. with triethylamine



# Managing UVCB Relationships

## Alkylbenzenesulfonate, linear

42615-29-2 | DTXSID3020041

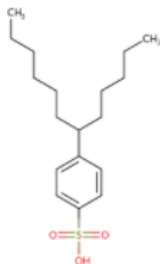
**i** Searched by Synonym: Found 1 result for  
'Linear alkylbenzene sulfonate'.

### Presence in Lists

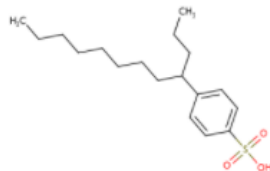
Surfactant List Screened in Swiss Wastewater (2014)

#### Surfactant List Screened in Swiss Wastewater (2014)

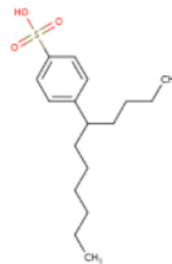
EAWAGSURF is a list of surfactants screened in Swiss wastewater effluents as part of a 2014 study. Structures/mixtures are being progressively curated and linked (Schymanski/Williams). Further details in Schymanski et al 2014, DOI: 10.1021/es4044374



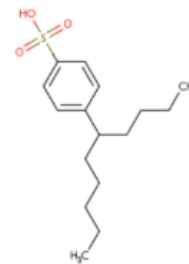
4-(Dodecan-6-yl)benzene-1-sulfon...  
23003-92-1



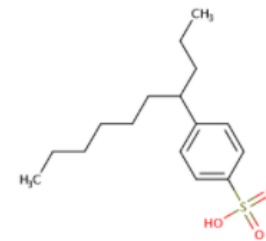
4-(dodecan-4-yl)benzene-1-sulfoni...  
NOCAS\_862870



C11-LAS  
NOCAS\_881097



4-(decan-5-yl)benzene-1-sulfonic ...  
NOCAS\_881146



4-(decan-4-yl)benzenesulfonic acid  
NOCAS\_891333

# Deliver Data and Models for Reuse

## Toxicity ForeCaster (ToxCast™) Data

EPA's most updated, publicly available high-throughput toxicity data on thousands of chemicals. This data is generated through the EPA's ToxCast research effort. ToxCast is part of the Toxicology in the 21st Century (Tox21) federal collaboration. All data is available for download and includes the following data sets. The release date and version names for the data sets are provided in the table below.

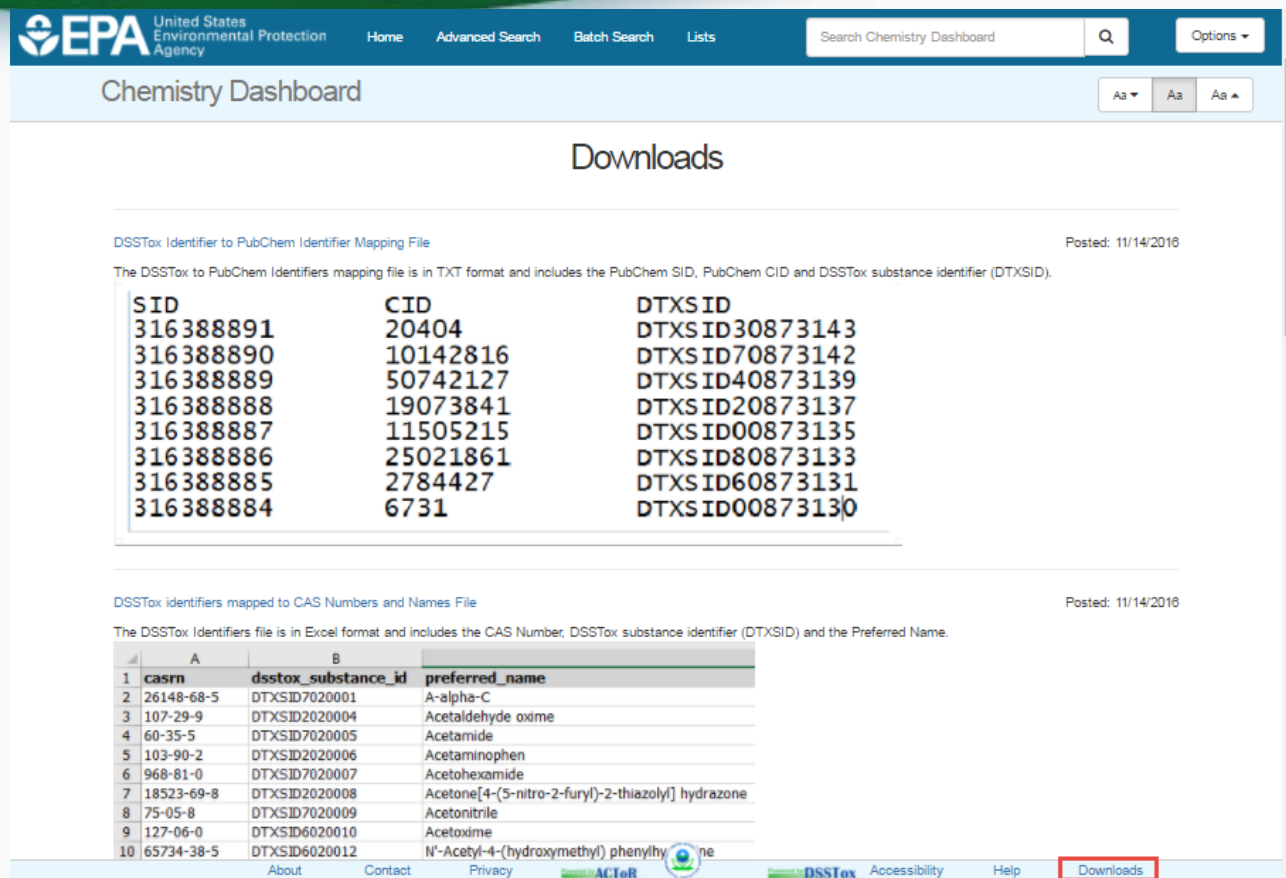
As part of EPA's commitment to share data, all of the computational toxicology data is publicly available for anyone to access and use. EPA's computational toxicology data is considered "open data", and thus all of the data below are free of all copyright restrictions, and fully and freely available for both non-commercial and commercial use.



- **ToxCast & Tox21 Chemicals:** A list of all chemicals screened, along with descriptions.
- **ToxCast & Tox21 High-throughput Assays:** Information about the hundreds of assays used to screen the chemicals.
- **ToxCast & Tox21 Summary Data:** Summary hit calls for all chemicals and concentration response curves for all assays.
- **MySQL Database:** Downloadable database that provides access to EPA's analysis of the ToxCast and Tox21 high-throughput screening data.
- **R Package:** The computer programming package used to process and model all EPA ToxCast and Tox21 chemical screening data.
- **Data Spreadsheet:** Spreadsheet that provides access to EPA's analysis of the ToxCast and Tox21 high-throughput screening data.
- **Concentration Response Curves:** Concentration response curves for all ToxCast & Tox21 assays.
- **Collaborative Estrogen Receptor Activity Prediction Project Data:** Data and supplemental files from CERAPP, a large-scale modeling project.

<https://www.epa.gov/chemical-research/toxicity-forecaster-toxcasttm-data>

# Delivering our Chemistry Data



The screenshot shows the EPA Chemistry Dashboard with a 'Downloads' section. It features two download links: one for a mapping file (DTXS ID to PubChem Identifier) and another for a file mapping DTXS identifiers to CAS numbers and names. Both files are in TXT or Excel format. The page includes a search bar, navigation links, and a footer with various utility links.

**Downloads**

[DTXS Identifier to PubChem Identifier Mapping File](#) Posted: 11/14/2016

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

SID	CID	DTXS ID
316388891	20404	DTXS ID30873143
316388890	10142816	DTXS ID70873142
316388889	50742127	DTXS ID40873139
316388888	19073841	DTXS ID20873137
316388887	11505215	DTXS ID00873135
316388886	25021861	DTXS ID80873133
316388885	2784427	DTXS ID60873131
316388884	6731	DTXS ID00873130

[DTXS identifiers mapped to CAS Numbers and Names File](#) Posted: 11/14/2016

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

	A	B	
	casn	dsstox_substance_id	preferred_name
1	26148-68-5	DTXSID7020001	A-alpha-C
2	107-29-9	DTXSID2020004	Acetaldehyde oxime
3	60-35-5	DTXSID7020005	Acetamide
4	103-90-2	DTXSID2020006	Acetaminophen
5	968-81-0	DTXSID7020007	Acetohexamide
6	18523-69-8	DTXSID2020008	Acetone[4-(5-nitro-2-furyl)-2-thiazolyl] hydrazone
7	75-05-8	DTXSID7020009	Acetonitrile
8	127-06-0	DTXSID6020010	Acetoxime
9	65734-38-5	DTXSID6020012	N'-Acetyl-4-(hydroxymethyl) phenylhydrazone

- Various types of data at FTP download site:  
[ftp://newftp.epa.gov/COMPTOX/Sustainable\\_Chemistry\\_Data/Chemistry\\_Dashboard](ftp://newftp.epa.gov/COMPTOX/Sustainable_Chemistry_Data/Chemistry_Dashboard)

# SDF Download

ACD/Spectrus DB: Database Window - [C:\USERS\AW...OWNLOADS\CHEMISTRYDASHBOARD-ADVANCEDSEARCH\_2017-08-14\_12-14-55.SDF]

Database View Record Search Lists Plates Options ACD/Labs Help

1 (ID:1) 2 (ID:2) 3 (ID:3) 4 (ID:4) 5 (ID:5) 6 (ID:6) 7 (ID:7)

8 (ID:8) 9 (ID:9) 10 (ID:10) 11 (ID:11) 12 (ID:12) 13 (ID:13) 14 (ID:14)

15 (ID:15) 16 (ID:16) 17 (ID:17) 18 (ID:18) 19 (ID:19) 20 (ID:20) 21 (ID:21)

22 (ID:22) 23 (ID:23) 24 (ID:24) 25 (ID:25) 26 (ID:26) 27 (ID:27) 28 (ID:28)


ID: 1 A: 1/89 B: 89 Last Updated: 14/08/2017 12:16 Single DB

1-ChemSketch 2-Database 3-Processor

# Open Data Reuse on iOS

## CompTox Mobile

[View in iTunes](#)


 This app is designed for both iPhone and iPad

**Free**  
Category: Productivity  
Released: Jan 16, 2017  
Version: 1.0  
Size: 267 MB  
Language: English  
Seller: Kirill Blinov  
© 2017 Molecule Apps,  
2017 EPA  
Rated 4+

**Compatibility:** Requires iOS 6.0 or later. Compatible with iPhone, iPad, and iPod touch.

**Customer Ratings**  
We have not received enough ratings to display an average for the current version of this application.

**More by Kirill Blinov**



NMR

[View in Mac App Store](#)

### Description

Find chemical structure instantly by exact mass (m/z), <sup>13</sup>C NMR chemical shifts, structure name or CAS Registry Number in a database of about 720,000 EPA CompTox structures.

[Kirill Blinov Web Site](#) • [CompTox Mobile Support](#) • [...More](#)

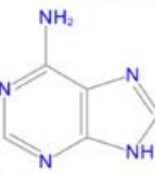
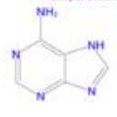

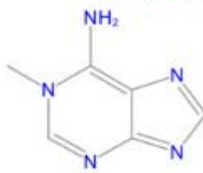
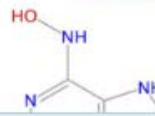

### Screenshots

iPhone | iPad

Carrier 2:30 PM

+ m/z Name <sup>13</sup>C 17

adenine

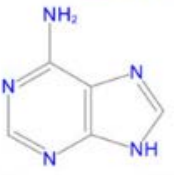

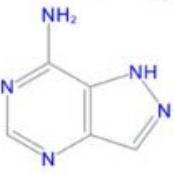
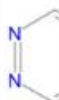

73-24-5	Adenine 1H-Purin-6-amine	73663-94-2	Adenine, dihydrodioxide Adenine dihydrodioxide
			
135.05450		390.87908	
52175-10-7	Adenine phosphate EINECS 257-702-7	5142-22-3	1-Methyladenine Adenine, 1-methyl-
			
233.03139		149.07015	
19152-67-1	Adenine, 2-chloro-N-hydroxy- 2-Chloro-N-hydroxyadenine	62700-65-6	Purine, 6-amino-2-methylthio-, hydrochloride Adenine, 2-methylthio-, hydrochloride
			

Carrier 2:31 PM

+ m/z Name <sup>13</sup>C

m/z 135.0545

<sup>13</sup>C Example: 25 32.4 115 (ppm)

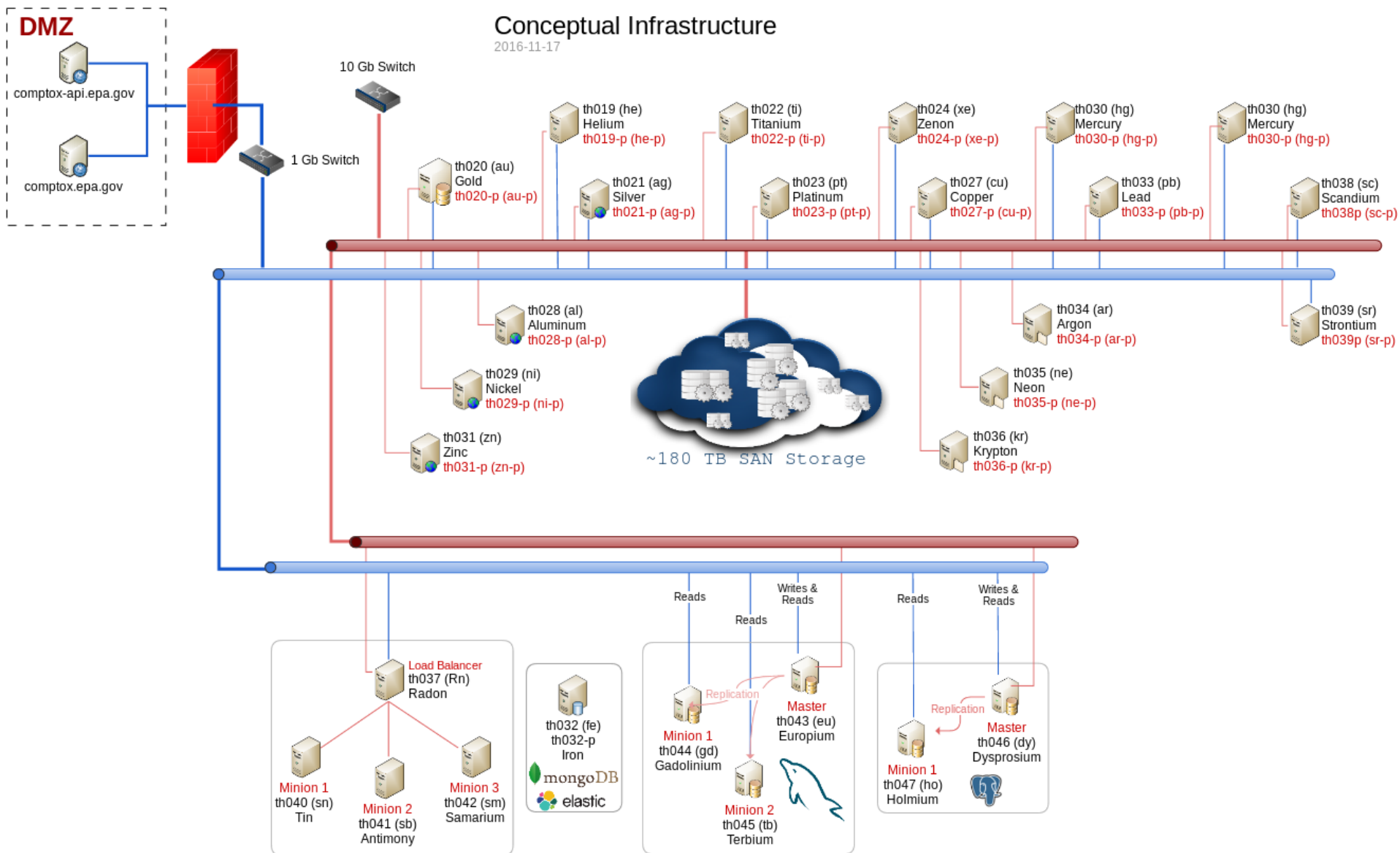
73-24-5	Adenine 1H-Purin-6-amine	2380-6	
			
135.05450	Δ 0.00000	135.05	
13877-56-0	7-Aminopurazole(4,3-dipyridine) 7-Aminopurazole(4,3-dipyridine)	160568	
			
135.05450	Δ 0.00000	135.05	
452-06-2	2-Aminopurine EINECS 207-197-4	5019-45	
			



# Computational Infrastructure and Processes

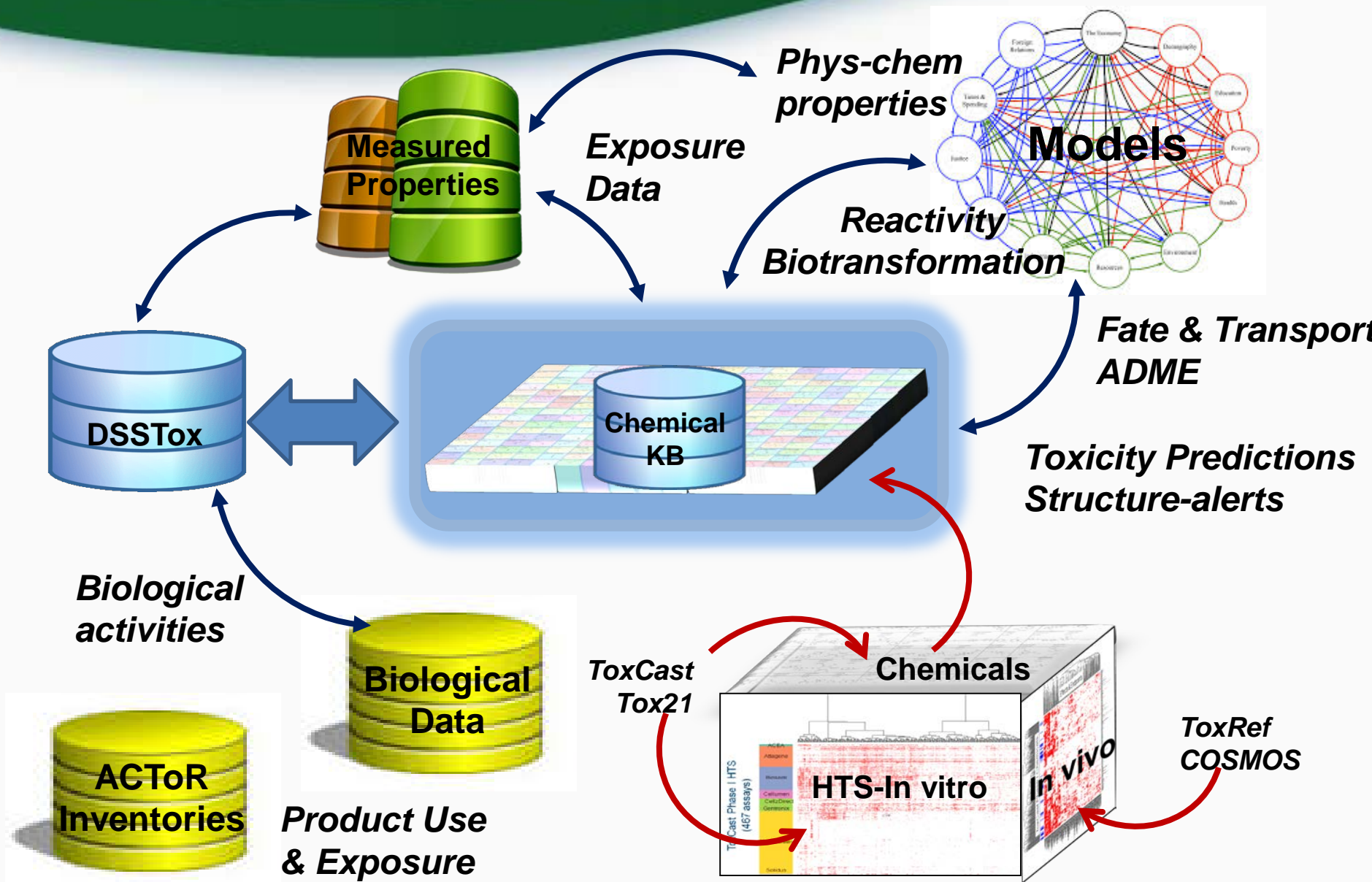
- 10 years of development – heterogeneous
  - MySQL, Java, R-code, Matlab, multiple Javascript environments, 100s of web services
  - Multiple projects, loose connectivity
- Migration to homogeneity for production Apps
  - MySQL
  - Java
  - JQuery/JavaScript
  - Versioned microservices
- Sprint methodology with Atlassian management tools
  - **Confluence** – knowledge management
  - **Jira** – ticketing/stories (work assignments)
  - **Fisheye** – collaborative peer code review
  - **Bitbucket** – code version control system

# Computational Infrastructure





# We're not done yet...



# OPERA Services in development

```
{
  "meta": {
    "statusCode": 200,
    "status": "success",          // [success | fail | error]
    "success": true,              // [true | false]
    "message": "string",         // A specific informational message...
    "warnings": "string",        // A warning message like: "LEADING ZEROS STRIPPED"
    "selfUrl": "http://api.zn.epa.gov/dsstox/chemicals", // The URL for THIS resource.
  },
  "data": [{
    "dsstox_cid": "DTXCID101",
    "inchi": "InChI=1S/C11H9N3/c12-10-6-5-8-7-3-1-2-4-9(7)13-11(8)14-10/h1-6H,(H3,12,13,14)",
    "inchiKey": "FJTNLJLPLJDTRM-UHFFFAOYSA-N",
    "model_name": "OPERA_LogP",
    "model_version": "v1.1.4",
    "predicted_property_id": 45433098
    "predicted_value": 2.53385
    "global_ad": 1
    "local_ad": 0.865944
    "confidence_score": 0.618723
  }],
}
```

# T.E.S.T services (ALPHA)

← ⓘ 🔒 [https://comptox.epa.gov/dashboard/web-test/WS?smiles=ClC\(Cl\)\(Cl\)Cl](https://comptox.epa.gov/dashboard/web-test/WS?smiles=ClC(Cl)(Cl)Cl)

JSON Raw Data Headers

Save Copy

```
uuid: "919325ac-14e5-45b5-b932-e401b79edf8c"
predictionTime: 1502741333903
software: "T.E.S.T (Toxicity Estimation Software Tool)"
softwareVersion: "4.2"
condition: "25°C"
endpoint: "Water solubility at 25°C"
method: "Consensus"
▼ predictions:
  ▼ 0:
    id: "C_1502741333903"
    smiles: "ClC(Cl)(Cl)Cl"
    expVal: "2.288"
    expValMass: "792.473"
    predVal: "2.186"
    predValMass: "1003.411"
```

# T.E.S.T services (ALPHA)

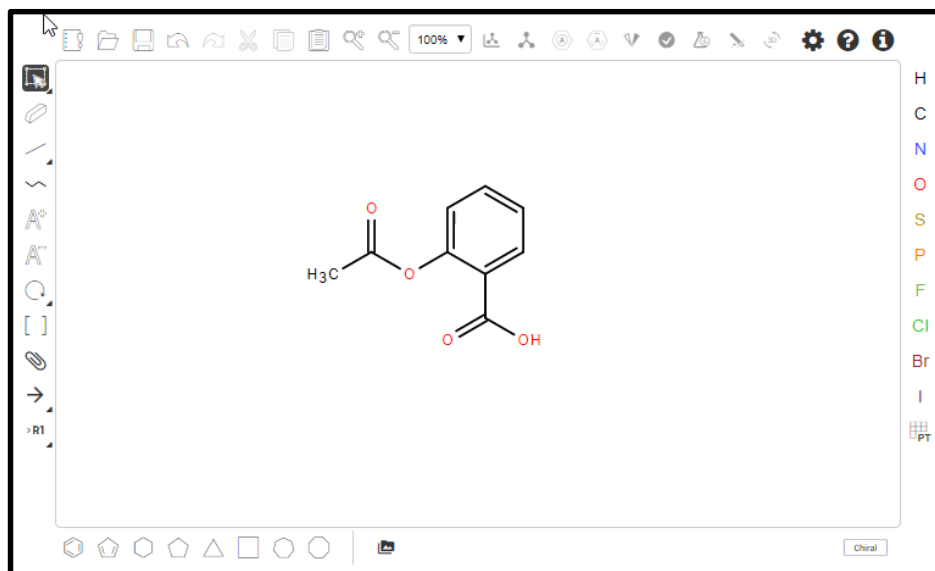
- 96hr fathead minnow 50% lethal concentration (LC50)
- 48hr daphnia magna 50% lethal concentration (LC50)
- Tetrahymena pyriformis 50% growth inhibition conc. (IGC50)
- Oral rat 50% lethal dose (LD50)
- Bioconcentration Factor (BCF)
- Developmental Toxicity (DevTox)
- Ames Mutagenicity (Mutagenicity)
- Normal boiling point, Flash point, Melting point
- Surface tension, Viscosity, Water Solubility
- Thermal Conductivity, Vapor Pressure, Density

# Real time predictions

## Chemistry Dashboard

Aa ▾ Aa Aa ▴

Search a chemical by systematic name, synonym, CAS number, or InChIKey



### Select properties to predict

T.E.S.T. 18

OPERA

EPI Suite

#### ☒ TOXICITY + -

- ☒ Fathead minnow LC50 (96 hr) ⓘ
- ☒ Daphnia magna LC50 (48 hr) ⓘ
- ☒ T. pyriformis IGC50 (48 hr)
- ☒ Oral rat LD50
- ☒ Bioaccumulation factor
- ☒ Estrogen Receptor RBA

#### ☒ PHYS\_CHEM + -

- ☒ Normal boiling point
- ☒ Vapor pressure at 25°C
- ☒ Melting point
- ☒ Flash point
- ☒ Density
- ☒ Surface tension at 25°C

#### ☒ ACTIVITY + -

- ☒ Developmental Toxicity
- ☒ Mutagenicity
- ☒ Estrogen Receptor Binding

Calculate

# Real time predictions

## Chemistry Dashboard

Aa ▼ Aa Aa ▲

Provider: TEST

Download table as: [TSV](#) [Excel](#)

Property	Experimental Value	Prediction				
		Consensus	Hierarchical clustering	Single model	Group contribution	Nearest neighbor
Fathead minnow LC50 (96 hr)		3.309 Log10(mol/L) 88.471 mg/L	3.286 Log10(mol/L) 93.341 mg/L	2.901 Log10(mol/L) 226.071 mg/L	3.240 Log10(mol/L) 103.561 mg/L	3.808 Log10(mol/L) 28.034 mg/L
Daphnia magna LC50 (48 hr)		3.271 Log10(mol/L) 96.599 mg/L	3.956 Log10(mol/L) 19.942 mg/L	3.858 Log10(mol/L) 24.999 mg/L	1.998 Log10(mol/L) 1808.159 mg/L	
T. pyriformis IGC50 (48 hr)		2.411 Log10(mol/L) 698.990 mg/L			1.982 Log10(mol/L) 1876.118 mg/L	2.840 Log10(mol/L) 260.425 mg/L
Oral rat LD50	2.955 Log10(mol/kg) 199.840 mg/kg	2.637 Log10(mol/kg) 415.666 mg/kg	2.942 Log10(mol/kg) 206.130 mg/kg			2.332 Log10(mol/kg) 838.203 mg/kg
Bioaccumulation factor		0.031 Log10 1.073	0.166 Log10 1.465	-0.161 Log10 0.690	-0.242 Log10 0.572	0.361 Log10 2.295
Estrogen Receptor RBA						-3.293 Log10 5.089*10 <sup>-4</sup>
Developmental Toxicity		true	true	true		
Mutagenicity						

- Continuous updating of lists and ongoing curation
- Integration to other agency databases – ECOTOX, ChemView
- Structure and substructure searching in development
- Release of Analytical QC data for ToxCast
- Further Non-Targeted Analysis support – spectral search
- Open API and web services
- “CompTox ~~Chemistry~~ Dashboard” will integrate other dashboards over time (i.e. EDSP21 and ToxCast)



# Acknowledgements



Credit: the Research Triangle Foundation

## **EPA-RTP**

*An enormous team of  
contributors from NCCT*

*and collaborators from  
NERL  
NHERL  
NRMRL*

## **Antony Williams**

US EPA Office of Research and Development

National Center for Computational Toxicology (NCCT)

[Williams.Antony@epa.gov](mailto:Williams.Antony@epa.gov)

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