

# Informatics Approaches supporting UVCB Chemicals in the US-EPA CompTox Chemicals Dashboard

***Antony Williams<sup>1</sup>, Chris Grulke<sup>1</sup> and Emma Schymanski<sup>2</sup>***

*Center for Computational Toxicology and Exposure, U.S. Environmental Protection Agency, RTP, NC  
Luxembourg Center for Systems Biomedicine, University of Luxembourg, Luxembourg, Europe*

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

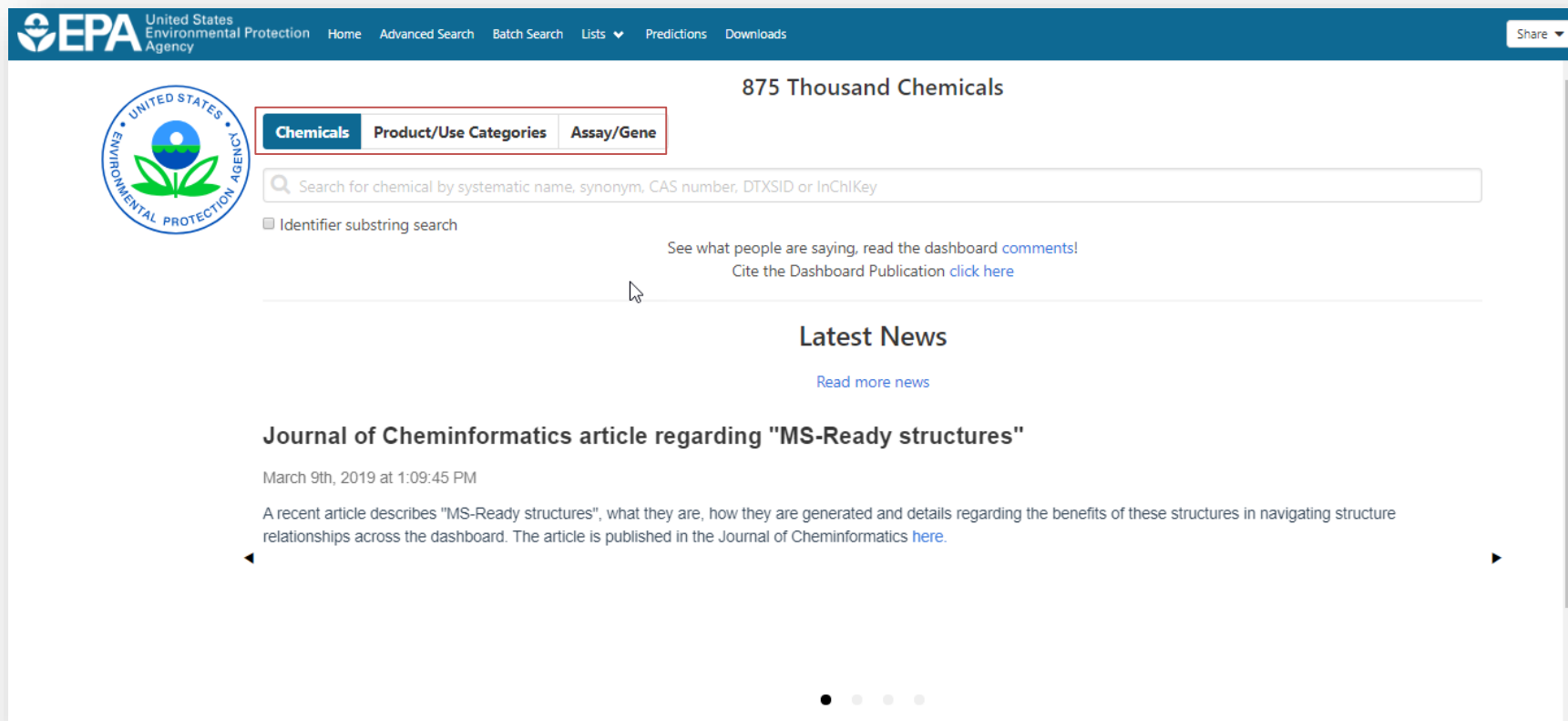
*November 2019  
SETAC, Toronto, Canada*

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

# CompTox Chemicals Dashboard

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


## 875k Chemical Substances




The screenshot shows the CompTox Chemicals Dashboard interface. At the top is a blue navigation bar with the EPA logo, the text "United States Environmental Protection Agency", and links for Home, Advanced Search, Batch Search, Lists (with a dropdown arrow), Predictions, and Downloads. A "Share" button is in the top right. Below the navigation bar is a white header area with the EPA seal on the left and the text "875 Thousand Chemicals" on the right. Under the seal are three tabs: "Chemicals" (highlighted), "Product/Use Categories", and "Assay/Gene". Below the tabs is a search bar with the placeholder text "Search for chemical by systematic name, synonym, CAS number, DTXSID or InChIKey". To the left of the search bar is a checkbox labeled "Identifier substring search". To the right of the search bar are two links: "See what people are saying, read the dashboard [comments!](#)" and "Cite the Dashboard Publication [click here](#)". Below this is a section titled "Latest News" with a link "Read more news". The first news item is titled "Journal of Cheminformatics article regarding 'MS-Ready structures'" and is dated "March 9th, 2019 at 1:09:45 PM". The text of the article states: "A recent article describes 'MS-Ready structures', what they are, how they are generated and details regarding the benefits of these structures in navigating structure relationships across the dashboard. The article is published in the Journal of Cheminformatics [here](#)." At the bottom of the news section are four small circular icons, with the first one being filled.


# BASIC Search


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

 Bisphenol




Bisphenol A  
DTXSID7020182






Bisphenol A bis(2-hydroxyethyl ether) diacrylate  
DTXSID6066991




Bisphenol A bis(2-hydroxyethyl ether) dimethacrylate  
DTXSID1066992




Bisphenol A bis(2-hydroxypropyl) ether  
DTXSID8051592

Bisphenol A carbonate polymer  
DTXSID6027840



Bisphenol A diglycidyl ether  
DTXSID6024624




Bisphenol A glycidyl methacrylate  
DTXSID7044841

# Detailed Chemical Pages

**EPA** United States Environmental Protection Agency

Home Advanced Search Batch Search Lists Predictions Downloads

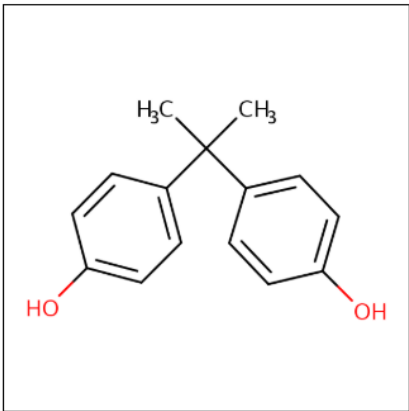
Copy Share Submit Comment Search all data



## Bisphenol A

80-05-7 | DTXSID7020182

Searched by DSSTox Substance Id.



**DETAILS**

EXECUTIVE SUMMARY

PROPERTIES

ENV. FATE/TRANSPORT

HAZARD

▶ ADME

▶ EXPOSURE

▶ BIOACTIVITY

SIMILAR COMPOUNDS

GENRA (BETA)

RELATED SUBSTANCES

SYNONYMS

▶ LITERATURE

LINKS


COMMENTS


**Wikipedia**


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


[Read more](#)


**Intrinsic Properties**


 **Molecular Formula:**  $\text{C}_{15}\text{H}_{16}\text{O}_2$

 Mol File

 Find All Chemicals

 **Average Mass:** 228.291 g/mol

 Isotope Mass Distribution

 **Monoisotopic Mass:** 228.11503 g/mol

**Structural Identifiers**

**Linked Substances**

**Presence in Lists**

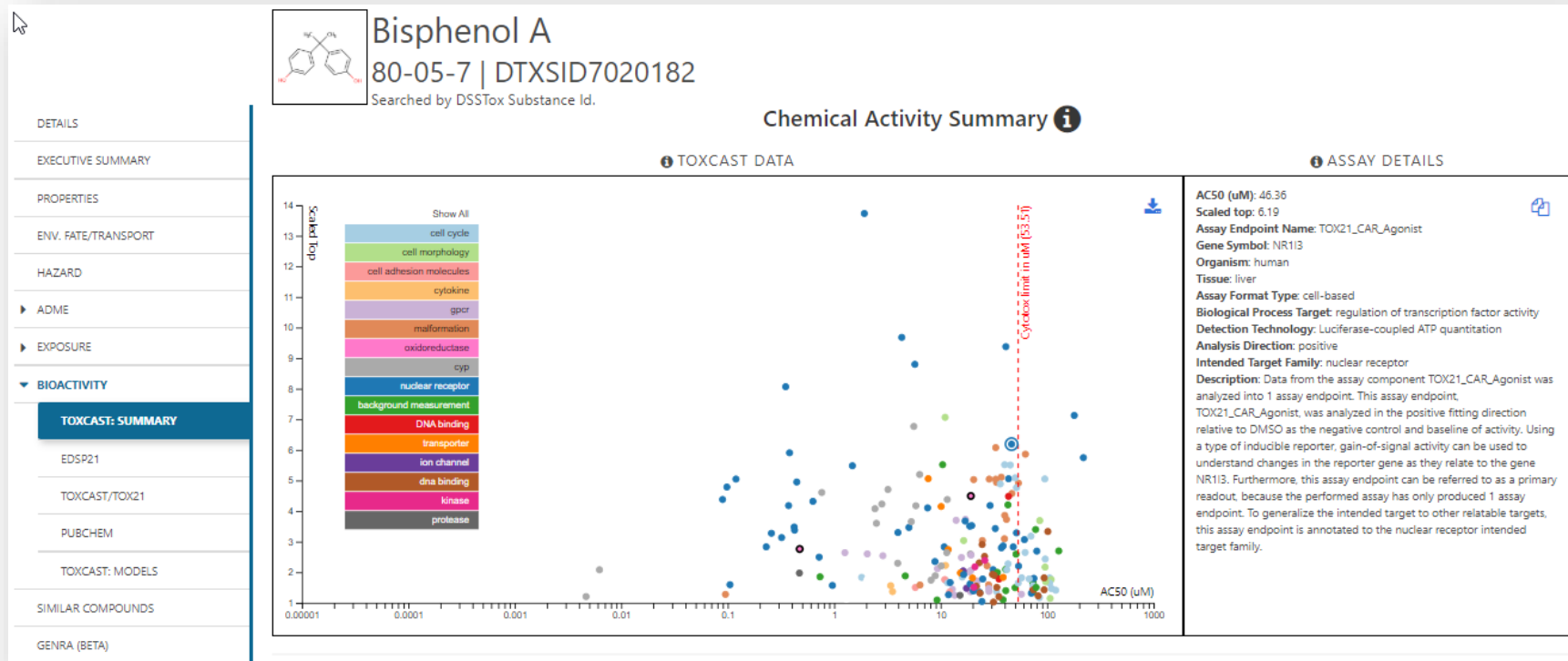
**Record Information**

**Quality Control Notes**

- ToxVal Database contains following data:
  - ~700,000 toxicity values
  - ~30 sources of data
  - ~22,000 sub-sources
  - ~5000 journals cited
  - ~70,000 literature citations

# In Vitro Bioassay Screening

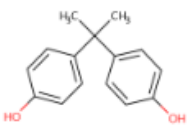
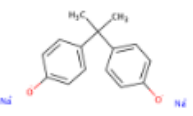
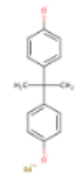
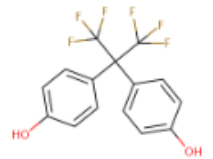
## ToxCast and Tox21



# Related Substances – Transformation Products, “Monomer-Polymer”

8 chemicals

Select all Download Send to Batch Search Relationship DTXSID PubChem CPDAT Hide chemicals that are: Filter by Name or CAS#

| Searched Chemical   | Polymer  | Polymer   | Polymer   | Predecessor: Component  |
|---|--|---|---|---|
|  <p>Bisphenol A<br/>DTXSID: DTXSID7020182<br/>PubChem: 161<br/>CPDAT: 326</p>                                | 2 related chemical structures with this substance<br><p>Formaldehyde, polymer with bisphenol A<br/>DTXSID: DTXSID3049627<br/>PubChem: 0<br/>CPDAT: 4</p>                                   | 2 related chemical structures with this substance<br><p>Bisphenol A/ Epichlorohydrin resin<br/>DTXSID: DTXSID0050479<br/>PubChem: 0<br/>CPDAT: 531</p>          | 3 related chemical structures with this substance<br><p>Bisphenol A-epichlorohydrin-polyformal...<br/>DTXSID: DTXSID9050480<br/>PubChem: 0<br/>CPDAT: 3</p> | 3 related chemical structures with this substance<br><p>Bisphenol A-epichlorohydrin-2,2',6,6'-tet...<br/>DTXSID: DTXSID70872791<br/>PubChem: 0<br/>CPDAT: 6</p> |
|  <p>Phenol, 4,4'-(1-methylethylidene)bis-, di...<br/>DTXSID: DTXSID0027480<br/>PubChem: 21<br/>CPDAT: 1</p> |  <p>Barium(2+) 4,4'-isopropylidenebispheno...<br/>DTXSID: DTXSID90240211<br/>PubChem: 6<br/>CPDAT: 0</p> |  <p>Bisphenol AF<br/>DTXSID: DTXSID7037717<br/>PubChem: 97<br/>CPDAT: 17</p> |   |   |

**What No Structures???**



# ***“UVCB” Chemical Substances***

## **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.

# Lots of UVCBS in Commerce....

1 related chemical  
structure with this  
substance

Butoxypropylene glycol  
DTXSID: DTXSID3034404  
PubChem: 0  
CPDAT: 437

0 related chemical  
structures with this  
substance

Fatty acids, tall-oil  
DTXSID: DTXSID0028088  
PubChem: 0  
CPDAT: 436



Poly(oxy-1,2-ethanediyl), alpha-tridecyl-...  
DTXSID: DTXSID8043993  
PubChem: 0  
CPDAT: 435

1 related chemical  
structure with this  
substance

Polyoxyethylene (20) oleylether  
DTXSID: DTXSID1027714  
PubChem: 0  
CPDAT: 435

0 related chemical  
structures with this  
substance

Soy oil  
DTXSID: DTXSID8027660  
PubChem: 0  
CPDAT: 435

0 related chemical  
structures with this  
substance

Bentonite  
DTXSID: DTXSID6030782  
PubChem: 0  
CPDAT: 433

0 related chemical  
structures with this  
substance

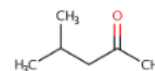
Distillates, petroleum, hydrotreated heav...  
DTXSID: DTXSID3028217  
PubChem: 0  
CPDAT: 433

1 related chemical  
structure with this  
substance

Edifas B  
DTXSID: DTXSID2020555  
PubChem: 0  
CPDAT: 433

0 related chemical  
structures with this  
substance

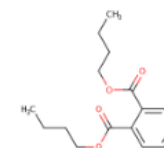
Coconut oil  
DTXSID: DTXSID8027664  
PubChem: 0  
CPDAT: 432



4-Methyl-2-pentanone  
DTXSID: DTXSID5021889  
PubChem: 173  
CPDAT: 428

0 related chemical  
structures with this  
substance

Distillates, petroleum, solvent-refined he...  
DTXSID: DTXSID9028172  
PubChem: 0  
CPDAT: 426



Dibutyl phthalate  
DTXSID: DTXSID2021781  
PubChem: 138  
CPDAT: 421

# Example UVCBs with no structures

- Aroclor 1254
- Toxaphene
- Xylenes
- Demeton
- Aroclor 1248
- Asbestos
- Technical chlordane
- Coke oven emissions
- Creosote
- Diesel engine exhaust
- Nickel refinery dust
- Nickel, soluble salts
- Refractory ceramic fibers
- Toluene diisocyanate

# What is Toxaphene?

“Complex, but reproducible mixture of at least **175** distinct C<sub>10</sub>-chloro compounds, having an approximate overall empirical formula of C<sub>10</sub>H<sub>10</sub>Cl<sub>8</sub>.; the 2 most active components are a C<sub>10</sub>H<sub>10</sub>Cl<sub>8</sub> compound and a C<sub>10</sub>H<sub>11</sub>Cl<sub>7</sub> compound which had been elucidated as 2,2,5-endo,6-exo,8,9,10-heptachlorobornane.

Produced by the chlorination of camphene to 67-69% chlorine by weight and made up of compds. of C<sub>10</sub> H<sub>8</sub> Cl<sub>10</sub>, C<sub>10</sub> H<sub>18-n</sub> Cl<sub>n</sub> (mostly polychlorobornanes) and C<sub>10</sub> H<sub>16-n</sub> Cl<sub>n</sub> (polychlorobornanes and/or polychlorotricyclenes) with n = 6 to 9”

# Toxaphene has rich data

8001-35-2 | DTXSID7021368

Searched by DSSTox Substance Id.

## Chemical Activity Summary

### Hazard

8001-35-2 | DTXSID7021368

Searched by DSSTox Substance Id.

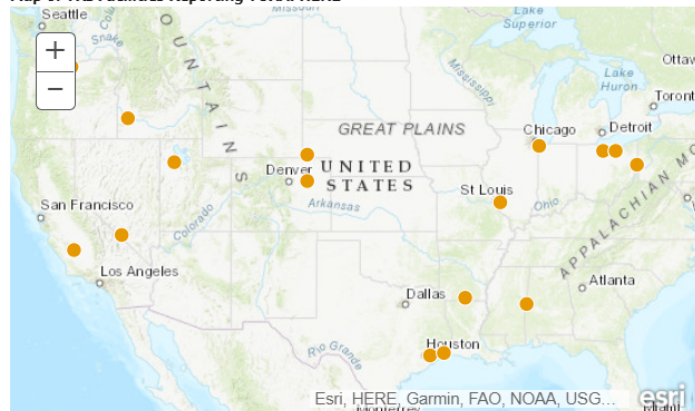
## Toxics Release Inventory

### 2015 TRI Factsheet: Chemical - TOXAPHENE, 008001352

Data Source: 2018 Updated Dataset (released October 2019)

The [Toxics Release Inventory \(TRI\)](#) tracks the management of certain toxic chemicals that may pose a threat to human health and the environment. Certain industrial facilities in the U.S. must report annually how much of each chemical is recycled, combusted for energy recovery, treated for destruction, and disposed of or otherwise released on- and off-site. This information is collectively referred to as production-related waste managed.

#### Map of TRI Facilities Reporting TOXAPHENE




#### Quick Facts for 2015

|   | Chemical           | United States     |
|---|--------------------|-------------------|
| <b>Number of TRI Facilities:</b>  | 16                 | 22,286            |
| <b>Total Production-Related Waste Managed:</b>                                | 200.2 thousand lbs | 28.9 billion lbs  |
| <b><a href="#">Total On-site and Off-site Disposal or Other Releases:</a></b> | 15.0 thousand lbs  | 3.4 billion lbs   |
| <b><a href="#">Total On-site:</a></b>   | 14.0 thousand lbs  | 2.9 billion lbs   |
| • <a href="#">Air:</a>  | 25 lbs             | 690.4 million lbs |
| • <a href="#">Water:</a>  | 0 lbs              | 199.3 million lbs |
| • <a href="#">Land:</a>   | 14.0 thousand lbs  | 2.0 billion lbs   |
| <b><a href="#">Total Off-site:</a></b>  | 1.0 thousand lbs   | 511.4 million lbs |

TOXAPHENE ranks 239 out of 499 chemicals reported to TRI in 2015 (Rank 1 = highest releases)

# UVCB: Mapped Chemicals

## Polymer individual components

 United States Environmental Protection Agency

Home Advanced Search Batch Search Lists Predictions Downloads

Copy Share Submit Comment Search all data

### 1-Hexene, 3,3,4,4,5,5,6,6,6-nonafluoro-, polymer with ethene and 1,1,2,2-tetrafluoroethene

68258-85-5 | DTXSID50880596

Searched by DSSTox Substance Id.

4 chemicals

Select all Download Send to Batch Search Relationship CASRN DTXSID Mono.Mass Hide chemicals that are: Filter by Name or CASRN

DETAILS

EXECUTIVE SUMMARY

PROPERTIES

ENV. FATE/TRANSPORT

HAZARD

ADME

EXPOSURE

BIOACTIVITY

SIMILAR COMPOUNDS

GENRA (BETA)

**RELATED SUBSTANCES**

SYNONYMS

LITERATURE

LINKS

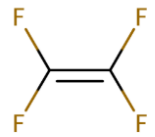
COMMENTS

Searched Chemical

3 related chemical structures with this substance

1-Hexene, 3,3,4,4,5,5,6,6,6-nonafluoro-, ...  
CASRN:68258-85-5  
DTXSID:DTXSID50880596  
Mono.Mass:0

Monomer



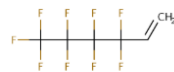
Tetrafluoroethylene  
CASRN:116-14-3  
DTXSID:DTXSID6021325  
Mono.Mass:99.993613

Monomer

$H_2C=CH_2$

Ethylene  
CASRN:74-85-1  
DTXSID:DTXSID1026378  
Mono.Mass:28.0313

Monomer



1H,1H,2H-Perfluoro-1-hexene  
CASRN:19430-93-4  
DTXSID:DTXSID6047575  
Mono.Mass:246.009104



#### Discover.

[About/Disclaimer](#)  
[Accessibility](#)  
[Privacy](#)

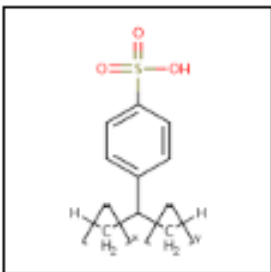
#### Connect.

[ACToR](#)  
[DSSTox](#)  
[Downloads](#)

#### Ask.

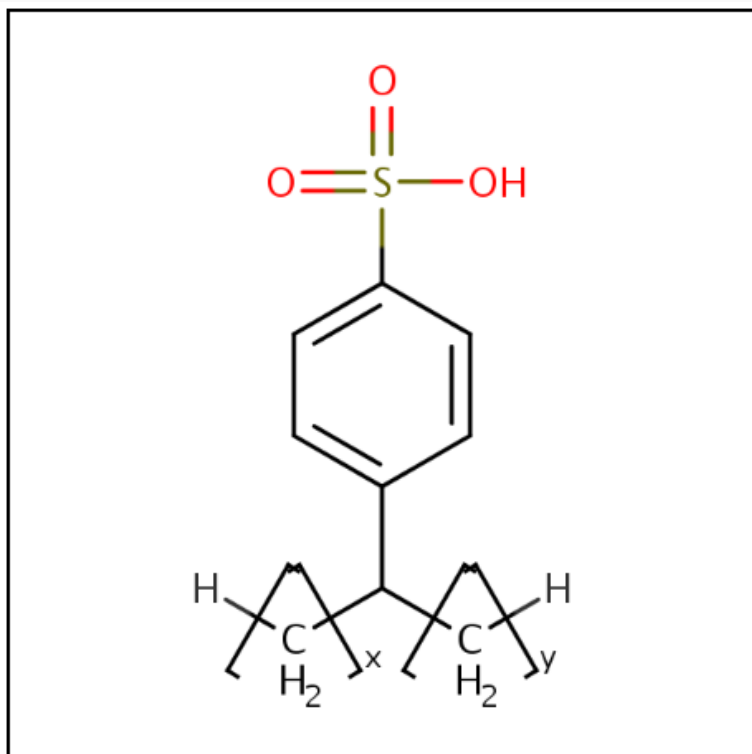
[Contact](#)  
[Help](#)

# Linear Alkylbenzenesulfonates



Alkylbenzenesulfonate, linear  
42615-29-2 | DTXSID3020041

Searched by DSSTox Substance Id.

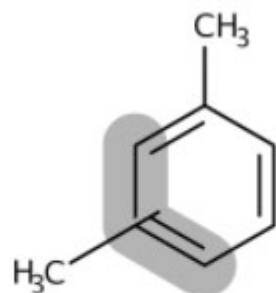


- Ambiguous representations of *some* structures are possible
- “Markush” representations can be valuable for describing and mapping chemicals

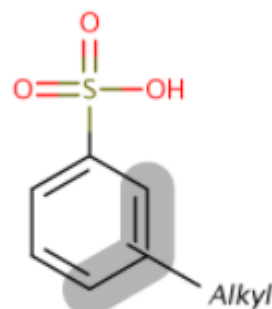


# “Markush Structures”

[https://en.wikipedia.org/wiki/Markush\\_structure](https://en.wikipedia.org/wiki/Markush_structure)



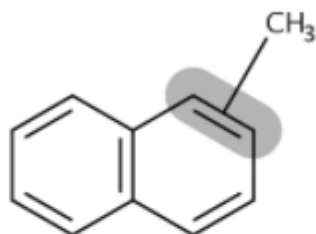
Xylenes  
1330-20-7



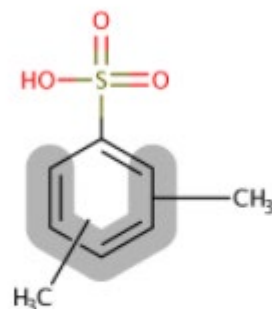
(C10-C16) Alkylbenzenesulfonic acid  
68584-22-5



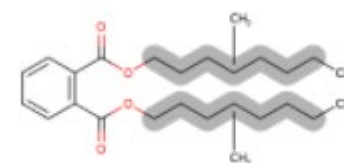
n-Nonylphenol  
25154-52-3



Methylnaphthalene  
1321-94-4



Sodium xylenesulfonate  
1300-72-7



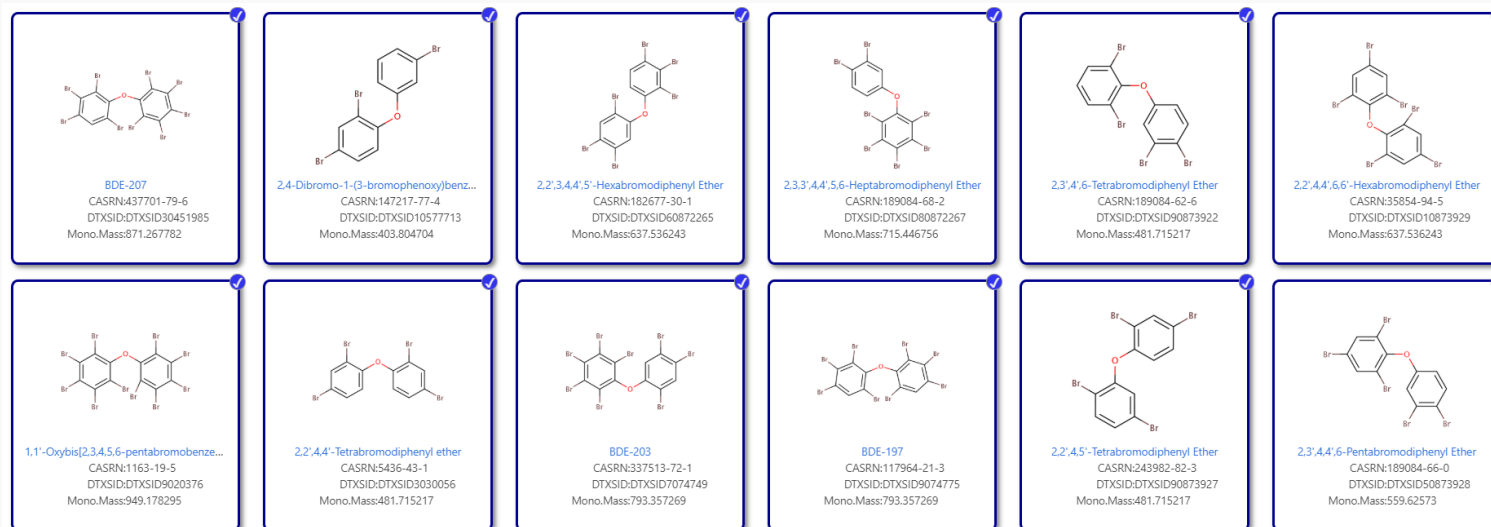
Diisononyl phthalate  
28553-12-0

# Chemical “grouping” – PCBs, PBDEs

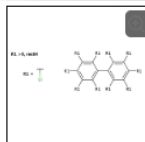
## PCBs



## PBDEs



# Some Markush chemicals are bounded













Polychlorinated biphenyls

1336-36-3 | DTXSID5024267

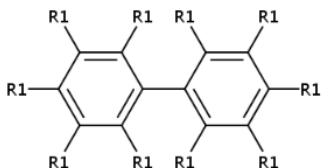
Searched by Expert Validated Synonym.

## Previously 1-1 mapping

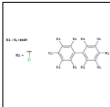
### ▼ Successor Substances (209)

|   | CAS-RN     | Relationship                       | Source        |
|---|------------|------------------------------------|---------------|
|    | 32774-16-6 | is a Representative Isomer of this | STN(DSSTox) ▼ |
|    | 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 ▼      |
|  | 34883-43-7 | is a Representative Isomer of this | Public ▼      |
|  | 34883-39-1 | is a Representative Isomer of this | Public ▼      |

R1 >0, restH



# Now Markush Enumeration



## Polychlorinated biphenyls

### 1336-36-3 | DTXSID5024267

Searched by Expert Validated Synonym.

DETAILS

EXECUTIVE SUMMARY

PROPERTIES

ENV. FATE/TRANSPORT

HAZARD

ADME

► EXPOSURE

BIOACTIVITY

SIMILAR COMPOUNDS

**RELATED SUBSTANCES**

SYNONYMS

► LITERATURE

LINKS

COMMENTS

Deselect all

Download

Send to Batch Search

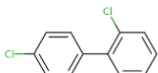
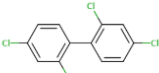
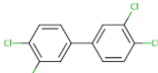
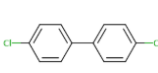
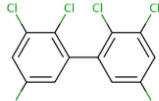
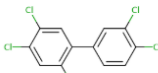
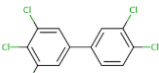
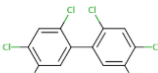
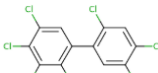
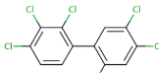
Relationship

18 of 211 chemicals selected

CASRN X DTXSID X Mono.Mass X

Unselected X

Filter by Name or CASRN

|   |  |   |  |   |
|---|--|---|--|---|
| <div>Markush Child</div> <div></div> <div>2,4'-Dichlorobiphenyl<br/>CASRN:34883-43-7<br/>DTXSID:DTXSID0022511<br/>Mono.Mass:222.000306</div> | <div>Markush Child</div> <div></div> <div>2,2',4,4'-Tetrachlorobiphenyl<br/>CASRN:2437-79-8<br/>DTXSID:DTXSID0022513<br/>Mono.Mass:289.922361</div> | <div>Markush Child</div> <div></div> <div>3,3',4,4'-Tetrachlorobiphenyl<br/>CASRN:32598-13-3<br/>DTXSID:DTXSID5022514<br/>Mono.Mass:289.922361</div> | <div>Markush Child</div> <div></div> <div>4,4'-Dichlorobiphenyl<br/>CASRN:2050-68-2<br/>DTXSID:DTXSID0022515<br/>Mono.Mass:222.000306</div> | <div>Markush Child</div> <div></div> <div>2,2':3,3':5,5'-Hexachlorobiphenyl<br/>CASRN:35694-04-3<br/>DTXSID:DTXSID4030045<br/>Mono.Mass:357.844416</div> |
| <div>Markush Child</div> <div></div>  | <div>Markush Child</div> <div></div>   | <div>Markush Child</div> <div></div>  | <div>Markush Child</div> <div></div>   | <div>Markush Child</div> <div></div>  |

# UVCB: Complex Surfactants


DETAILS


RELATED SUBSTANCES

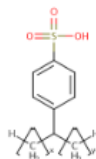
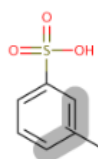
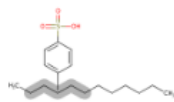
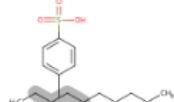
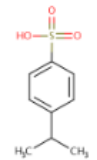
PROPERTIES

COMMENTS

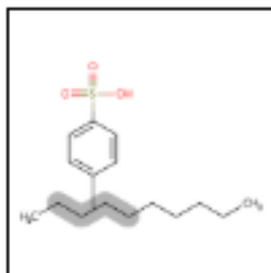
25 chemicals

Select all Download Send to Batch Search Relationship  DTXSID PubChem CPDAT

Hide chemicals that are: Filter by Name or CASRN 

|  |   |   |   |   |
|--|---|---|---|---|
| <p>Searched Chemical</p>  <p>Alkylbenzenesulfonate, linear<br/>DTXSID: DTXSID3020041<br/>PubChem: 0<br/>CPDAT: 83</p> | <p>Component</p>  <p>(C10-C16) Alkylbenzenesulfonic acid<br/>DTXSID: DTXSID2028723<br/>PubChem: 0<br/>CPDAT: 132</p> | <p>Component</p> <p>2 related chemical structures with this substance</p> <p>Benzenesulfonic acid, C10-16-alkyl deriv...<br/>DTXSID: DTXSID7028724<br/>PubChem: 0<br/>CPDAT: 21</p>                   | <p>Component</p> <p>2 related chemical structures with this substance</p> <p>Benzenesulfonic acid, C10-13-alkyl deriv...<br/>DTXSID: DTXSID3029784<br/>PubChem: 0<br/>CPDAT: 179</p>                | <p>Component</p> <p>4 related chemical structures with this substance</p> <p>Benzenesulfonic acid, dodecyl-, branched...<br/>DTXSID: DTXSID4049583<br/>PubChem: 0<br/>CPDAT: 22</p>                     |
| <p>Component</p> <p>1 related chemical structure with this substance</p> <p>Benzenesulfonic acid, C10-16-alkyl deriv...<br/>DTXSID: DTXSID7099894<br/>PubChem: 0<br/>CPDAT: 52</p>                     | <p>Component</p> <p>1 related chemical structure with this substance</p> <p>Benzenesulfonic acid, mono-C9-17-branched...<br/>DTXSID: DTXSID20100697<br/>PubChem: 0<br/>CPDAT: 16</p>                  | <p>Component</p>  <p>C12-linear alkyl benzene sulfonate<br/>DTXSID: DTXSID90891641<br/>PubChem: 0<br/>CPDAT: 0</p> | <p>Component</p>  <p>C10-linear alkylbenzenesulfonate<br/>DTXSID: DTXSID70891689<br/>PubChem: 0<br/>CPDAT: 0</p> | <p>Markush Child</p>  <p>4-Isopropylbenzenesulfonic acid<br/>DTXSID: DTXSID1044932<br/>PubChem: 42<br/>CPDAT: 0</p> |

# UVCB: Complex Surfactants

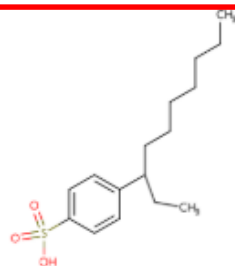


C10-linear alkylbenzenesulfonate

NOCAS\_891689 | DTXSID70891689

Searched by DSSTox Substance Id.

Markush Child



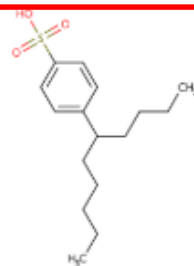
4-(Decan-3-yl)benzene-1-sulfonic acid

DTXSID: DTXSID20859618

PubChem: 0

CPDAT: 0

Markush Child



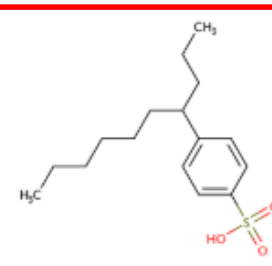
4-(decan-5-yl)benzene-1-sulfonic acid

DTXSID: DTXSID70881146

PubChem: 0

CPDAT: 0

Markush Child



4-(decan-4-yl)benzenesulfonic acid

DTXSID: DTXSID40891333

PubChem: 0

CPDAT: 0

# ***Chemical Lists and Categories***

# EPAHFR: Hydraulic Fracturing

Home Advanced Search Batch Search Lists ▼ Predictions Downloads

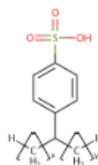
Lists of Chemicals

List of Assays

## WATER|EPA; Chemicals associated with hydraulic fracturing

☐ Identifier substring search

List

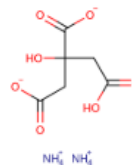


Alkylbenzenesulfonate, linear  
DTXSID: DTXSID3020041  
PubChem: 0  
CPDAT: 83

Cl

NH<sub>4</sub>

Ammonium chloride  
DTXSID: DTXSID0020078  
PubChem: 82  
CPDAT: 260

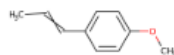


Diammonium citrate  
DTXSID: DTXSID5020079  
PubChem: 19  
CPDAT: 18

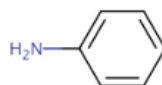
OH

NH<sub>4</sub>

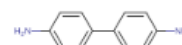
Ammonium hydroxide  
DTXSID: DTXSID4020080  
PubChem: 83  
CPDAT: 857



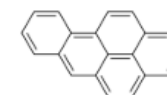
Anethole  
DTXSID: DTXSID4020086  
PubChem: 59  
CPDAT: 33



Aniline  
DTXSID: DTXSID8020090  
PubChem: 284  
CPDAT: 80



Benzidine  
DTXSID: DTXSID2020137  
PubChem: 112  
CPDAT: 15



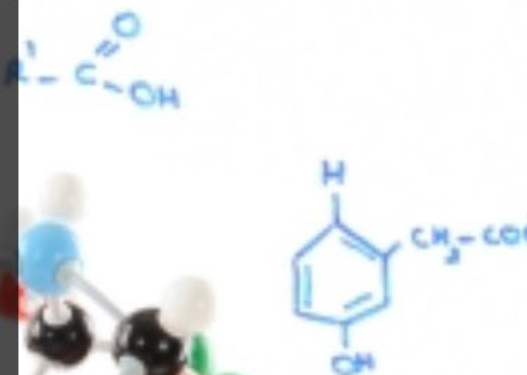
Benzo[a]pyrene  
DTXSID: DTXSID2020139  
PubChem: 161  
CPDAT: 81



## TSCA Chemical Substance Inventory

CONTACT US

The Toxic Substances Control Act (TSCA) Chemical Substance Inventory contains all existing chemical substances manufactured, processed, or imported in the United States that do not qualify for an exemption or exclusion under TSCA. [This may be your starting place for interaction with EPA on TSCA regulatory matters.](#)



Nucleophiles (Nu):



# TSCA Inventory List

## Active Non-confidential portion

### EPA|TSCA: TSCA Inventory, active non-confidential portion

☐ Identifier substring search

#### List Details

11919 of 23514 chemicals visible

Select all

Download

Send to Batch Search

Default

↑

CASRN

DTXSID

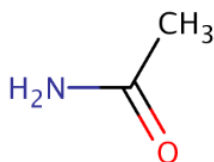
Mono.Mass

No Structures

Filter by Name or CASRN

⌵

Unselected  
Isotopes  
Multicomponent Chemicals  
No Structures

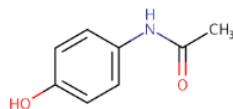


Acetamide

CASRN:60-35-5

DTXSID:DTXSID7020005

Mono.Mass:59.037114

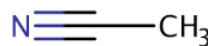


Acetaminophen

CASRN:103-90-2

DTXSID:DTXSID2020006

Mono.Mass:151.063329

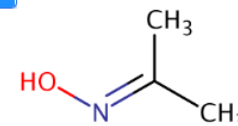


Acetonitrile

CASRN:75-05-8

DTXSID:DTXSID7020009

Mono.Mass:41.026549



Acetoxime

CASRN:127-06-0

DTXSID:DTXSID6020010

Mono.Mass:73.052764

**Over 11,500 out of 23,500 have NO STRUCTURES**

# ***Batch Searching***

- Singleton searches are useful but people generally want data on LOTS of chemicals!
- Typical questions
  - What is the list of chemicals for the formula  $C_xH_yO_z$
  - What is the list of chemicals for a mass +/- error
  - Can I get chemical lists in Excel files? In SDF files?
  - Can I include properties in the download file?










# Batch Search Names

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

Step Five: Choose Data Fields to Download

Please enter one identifier per line

Select Input Type(s)

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

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

Buprenorphine  
Codeine  
Dextromethorphan  
Dihydrocodeine  
Dihydromorphine  
Ethylmorphine  
Fentanyl  
Heroin  
Hydrocodone  
Hydromorphone







Display All Chemicals Download Chemical Data

Excel  
Download






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

# Add Other Data of Interest






## Chemical Identifiers

- ☒ DTXSID 
- ☒ Chemical Name 
- ☐ DTXCID 
- ☒ CAS-RN 
- ☒ InChIKey 
- ☐ IUPAC Name 

## Structures

- ☐ Mol File 
- ☐ SMILES 
- ☐ InChI String 
- ☒ MS-Ready SMILES 
- ☐ QSAR-Ready SMILES 

## Intrinsic And Predicted Properties






- ☒ Molecular Formula 
- ☐ Average Mass 
- ☒ Monoisotopic Mass 
- ☐ TEST Model Predictions 
- ☐ OPERA Model Predictions 

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



# Related Substance Relationships

## Enhanced Data Sheets

- ☐ MetFrag Input File (Beta) 
- ☐ ToxPrint single fingerprints 
- ☐ Abstract Sifter Input File (Beta) 
- ☐ Synonyms and Identifiers 
- ☒ Related Substance relationships 
- ☐ ToxPrint fingerprints (ChemoTyper) 
- ☐ Associated ToxCast Assays 

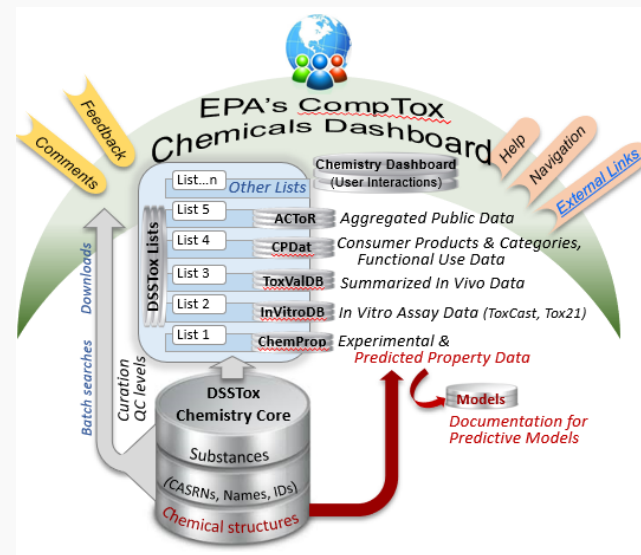
- ☐ EPA: National-Scale Air
- ☐ EPA: PPRTV Chemical R
- ☐ EPA: Provisional Adviso
- ☐ EPA: Safer Choice Chem

Selecting this checkbox provides a separate Excel worksheet containing the relationship between two chemicals. The output file includes the DTXSIDs and names/CASRN between the input list and the related chemical. Relationships include, for example, polymer, components, salt form, transformation product and other relationships.

|    | A       | B             | C              | D                      | E              |                              |
|----|---------|---------------|----------------|------------------------|----------------|------------------------------|
| 1  | INPUT   | DTXSID        | PREFERRED_NAME | HAS_RELATIONSHIP_WITH  | RELATED_DTXSID | RELATED_PREFERRED_NAME       |
| 2  | xylenes | DTXSID2021446 | Xylenes        | Transformation Product | DTXSID40176394 | N-Benzoylalanine             |
| 3  | xylenes | DTXSID2021446 | Xylenes        | Component              | DTXSID6026298  | m-Xylene                     |
| 4  | xylenes | DTXSID2021446 | Xylenes        | Component              | DTXSID3021807  | o-Xylene                     |
| 5  | xylenes | DTXSID2021446 | Xylenes        | Component              | DTXSID2021868  | p-Xylene                     |
| 6  | xylenes | DTXSID2021446 | Xylenes        | Predecessor: Component | DTXSID9021421  | Xylenes; defined mixture 1   |
| 7  | xylenes | DTXSID2021446 | Xylenes        | Predecessor: Component | DTXSID7021447  | Xylenes; defined mixture 2   |
| 8  | xylenes | DTXSID2021446 | Xylenes        | Predecessor: Component | DTXSID30891529 | Total Petroleum Hydrocarbons |
| 9  | xylenes | DTXSID2021446 | Xylenes        | Markush Child          | DTXSID3021807  | o-Xylene                     |
| 10 | xylenes | DTXSID2021446 | Xylenes        | Markush Child          | DTXSID6026298  | m-Xylene                     |
| 11 | xylenes | DTXSID2021446 | Xylenes        | Markush Child          | DTXSID2021868  | p-Xylene                     |

# Conclusion

- Building an integrated hub for environmental chemistry data to serve computational toxicology
- Transparent access to data and models – file downloads, SQL data dumps and web services
- Expansion of functionality to serve all data streams generated by NCCT across the agency & community
- Data **QUALITY** is a key focus - ongoing curation
- Ongoing API development will provide enhanced access to data streams





# Acknowledgements



Credit: the Research Triangle Foundation

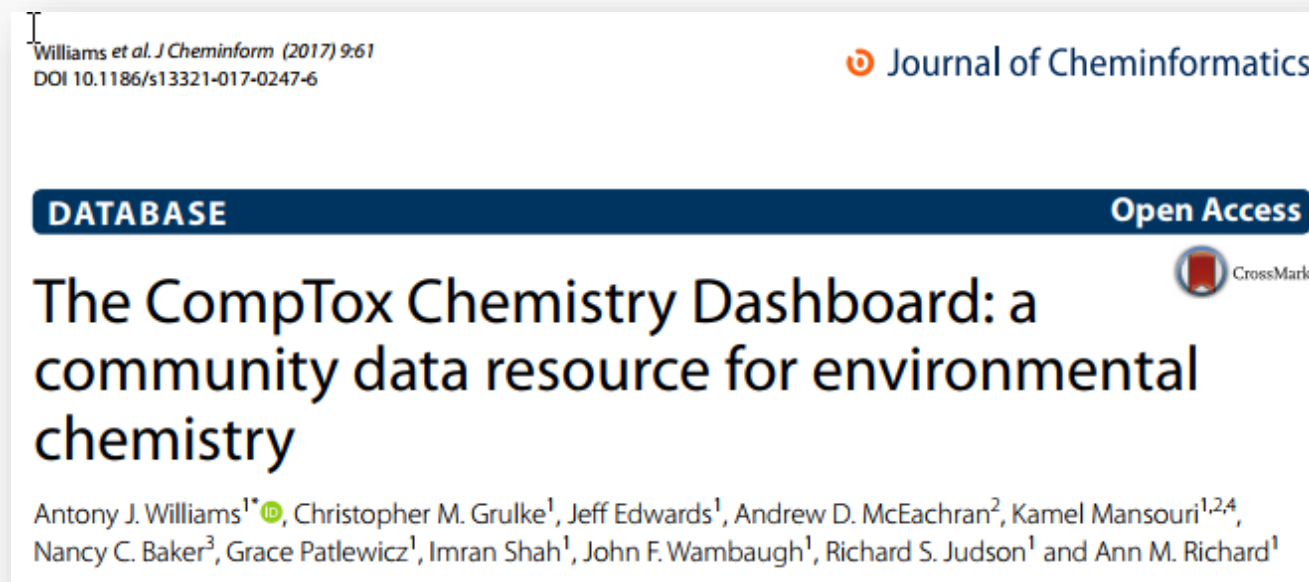
## EPA-RTP

- *An enormous team of contributors from CCTE, especially the IT software development team*
- *Our curation team for their care and focus on data quality*
- *Multiple centers and laboratories across the EPA*
- *Many public domain databases and open data contributors*

## Antony Williams

CCTE, US EPA Office of Research and Development,  
[Williams.Antony@epa.gov](mailto:Williams.Antony@epa.gov)

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



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