ACToR – Aggregated Computational Toxicology Resource
&
ToxRefDB – Toxicity Reference Database

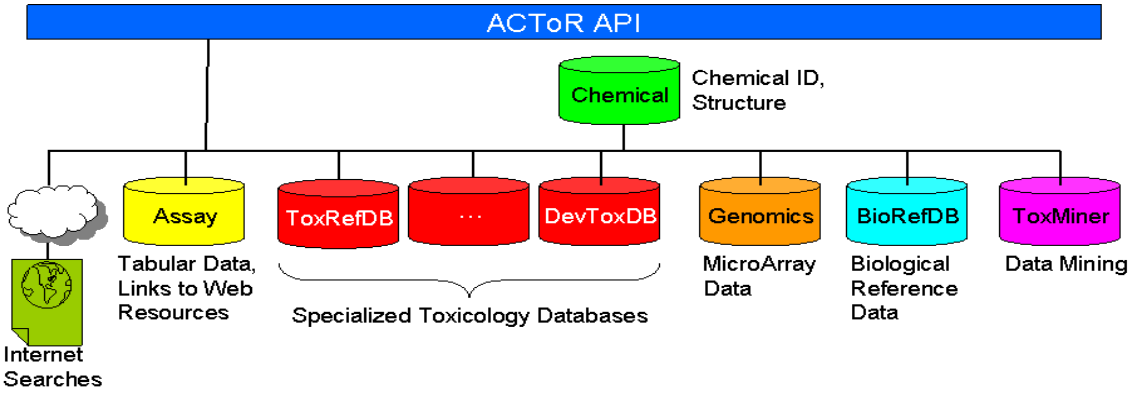
Release v2010Q1 (March 2010)
Produced by the U.S. Environmental Protection Agency,
National Center for Computational Toxicology

Table of Contents

Introduction
Overview: how to use ACToR
Overview: how to use ToxRefDB
Data Collection (ACToR Only)
Data Collection View (ACToR Only)
Generic Chemical View (ACToR & ToxRefDB)
Search by Name (ACToR Only)
Search by CAS Number (ACToR Only)
Search by Structure (ACToR Only)
Browsing Assay (ACToR Only)
Assay View (ACToR Only)
Assay Data Page (ACToR Only)
Search By Endpoint (ToxRefDB Only)
Glossary

Introduction

ACToR (Aggregated Computational Toxicology Resource) is a collection of databases collated or developed by the US EPA National Center for Computational Toxicology (NCCT). Over 580 sources of publicly available data on environmental chemicals have been brought together and made searchable by chemical names, CAS number, structure drawings, InCHI strings and SMILES. This database contains data such as hazardous testing results, chemical structure, physico-chemical values, *in vitro* assay data, exposure data, and *in vivo* toxicology data on hundreds of thousands of generic chemicals. These chemicals include, but are not limited to, high and medium production volume industrial chemicals, pesticides (active and inert ingredients), food additives and potential ground and drinking water contaminants. The overall structure of ACToR is composed of a series of domains, linked together by chemicals.



Domain	Description	Location
Chemicals	Structure, names and other basic chemical information	Main ACToR Database
Assays	Quantitative and other tabular data on chemicals	Main ACToR Database
Toxicology	*In vivo* study data from multiple domains	ACToR / ToxRefDB Database
Genomics Microarray Data	Full microarray data sets, in both original and transformed versions	Under Development
Biological Reference Data	Information on genes, proteins and pathways, downloaded from public sources	Main ACToR Database
ToxMiner	Detailed data from the ToxCast and ToxRefDB programs - used for ToxCast analyses.	Separate ToxMiner database - linked to ACToR by chemical ID

Currently, chemical toxicity data resides in a variety of specialized databases with incompatible formats and in many different locations. In the past, researchers needed to search many databases and manually organize the results in order to amass all the information on a given chemical. While this is possible for a few chemicals it is very difficult to compile comprehensive data sets on chemically-similar sets of compounds that can be explored with structure searching tools. Since the majority of chemicals in ACToR have chemical structures produced by the EPA ToxCast chemical prioritization program, ACToR will help to stimulate studies of structure-function relationships in sets of environmental chemicals. In addition, by bringing together data from a large number of sources and making the data structure-searchable, ACToR facilitates searches that transcend current databases in abundance of information and in number of generic chemicals.

The ACToR project is compiling data (both quantitative and qualitative) from a large number of sources (called data collections).Some of the sources include EPA, PubChem, other NIH, USDA and FDA databases; state, national, international and academic references. One novel data collection is ToxRefDB (Toxicity Reference Database), which includes detailed information on *in vivo* guideline study results for pesticides and other potentially toxic chemicals that has been assembled by the NCCT. ACToR is also the primary repository of data being produced by the EPA. Overall, ACToR is engineered to be flexible enough to add a variety of new data from sources with different formats in a straightforward manner. Because we have not already discovered and included all toxicology data, we are always interested in obtaining other data collections that could be incorporated into the system.

ToxRefDB was developed by the National Center for Computational Toxicology (NCCT) in partnership with EPA's Office of Pesticide Programs (OPP), to store data from in vivo animal toxicity studies. The database primarily contains pesticide registration toxicity data that used to be stored as hard-copy and scanned documents by OPP. Currently, ToxRefDB includes chronic, cancer, sub-chronic, developmental, and reproductive studies on hundreds of chemicals (many are pesticide active ingredients). Each study entered utilizes a detailed standardized vocabulary for capturing nearly every element in the database, including the toxicological outcomes. To date, ToxRefDB contains roughly 2000 studies on nearly 500 chemicals. Through this web interface the data is provided in an accessible and computable manner.

Additional background information and examples of the utility of ToxRefDB can be found in three published manuscripts:

Martin et al. (2009) "Profiling Chemicals Based on Chronic Toxicity Results from the U.S. EPA ToxRef Database" Environmental Health Perspectives doi:10.1289/ehp.0800074

Knudsen et al. (2009) "Profiling the Activity of Environmental Chemicals in Prenatal Developmental Toxicity Studies using the U.S. EPA's ToxRefDB" Reproductive Toxicology doi: 10.1016/j.reprotox.2009.03.016

Martin et al. (2009) "Profiling the Reproductive Toxicity of Chemicals from Multigeneration Studies in the Toxicity Reference Database (ToxRefDB)" Toxicological Sciences doi: 10.1093/toxsci/kfp080

Additionally, ToxRefDB provides reference toxicity data for Agency research and retrospective analyses, while maintaining its initial focus on providing toxicity endpoints for development of ToxCast predictive signatures that will be used for primary research applications. It should be noted that ToxRefDB contains only certain hazard information and does not represent all information needed for a complete risk assessment for pesticides or other chemicals. Effect designation should not be taken as determination that existing EPA risk assessments and risk management decisions need revisions. For example, in addition to studies in ToxRefDB, for purposes of registration or tolerance determination, EPA evaluated information on other mammalian toxicity effects, metabolism, aquatic life, wildlife and plant toxicity studies, and use patterns, environmental fate and persistence, and pesticide residue levels.

TOP

Overview: How to Use ACToR

In ACToR chemicals are organized into three main classes, the first two of which are modeled closely after the corresponding PubChem data model. The three main classes are:

- Substance- a unique chemical from a single "data collection"
- Compound- holds chemical structure information
- Generic Chemical- aggregates a chemical structure plus all of the corresponding substances. The common link is that all substances share the same CAS registry number.

Browsing

One can find information on ACToR information by either browsing or by searching for a particular chemical. A person can look through the ACToR database by browsing through either the data collections or the assay data.

- Assay: An assay is a collection of data for substances from one data collection.
- Data collection: A collection of chemical and assay data from a single source

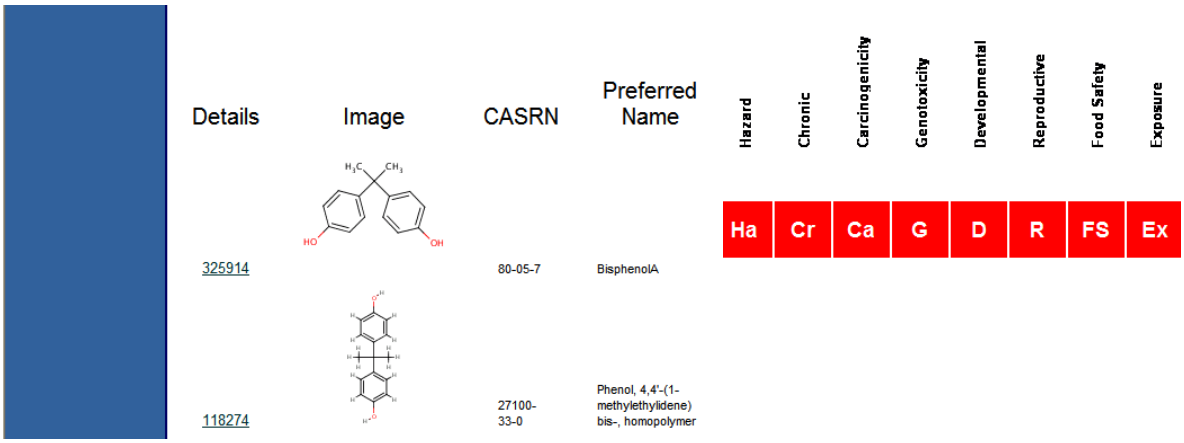
Searching

If a person has a particular chemical in mind, he or she can find it by either entering the name or CAS number or they can draw the structure.

- Search by Name- allows one to search for chemicals by their common and their many alternate names.
- Search by CAS- enables one to search by CAS number
- Search by Structure- allows a person to either draw a chemical or paste an INCHI or SMILES code onto the drawing pad to find chemicals.

After conducting a search, it is possible to go back and edit what has been entered. To do this, DO **NOT** CLICK ON THE BACK BUTTON. Clicking on the back button will erase all of the data. Instead re-click on the "Search by (name/CAS/structure)" link on the left navigation bar.

Search Results



The search results are composed of a table with all the chemicals in that data table. The columns include:

Details- contains the generic chemical id and a link to the chemical's generic chemical page
Image – Drawing of the chemical
CASRN- CAS Number
Preferred Name- the chemicals proper name

A series of red boxes will appear to the right of the "Preferred Name" column. The appearance of a red box indicates that there is toxicity data available for the substance that appears under the "Preferred Name" column. The initials within the box denote the specific type of toxicological information that is accessible with ACToR. It does not necessarily mean that the substance has tested positive for a particular effect.

H (Hazard)- This category includes substances that have been assessed for the potential to cause adverse health effects or damage to property.

Ca (Carcinogenicity)- This category includes substances that have been assessed for the potential to cause cancer.

G (Genotoxicity)- This category includes substances that have been assessed for the potential to cause damage to genetic material, such as DNA.

D (Developmental)- This category includes substances that have been assessed for the potential to affect the development and/or growth of an organism.

R (Reproductive)- This category includes substances that have been assessed for the potential to interfere with normal reproduction, such as causing changes in fertility.

Cr (Chronic)-Substances that have tested for adverse effects by exposure to the test agent over a substantial portion of the organism's lifespan.

FS (Food Safety)- Designation in this category means that the substance has been tested as to its safety as either: an additive to food, a food ingredient or present in food packaging.

Ex (Exposure)- Exposure information, including results of biomonitoring studies, concentrations in the air, water or soil, and allowable exposure limits.

TOP

Overview: How to Use ToxRefDB

In ToxRefDB, toxicity study information is organized into four main tiers, chemical, study design, treatment group (animal) dosing information, and effects (i.e., endpoints):

- Generic Chemical- aggregates a chemical structure plus all of the corresponding substances. The common link is that all substances share the same CAS registry number.
- Study Design- captures the detailed study design information associated with a given study, including test animal species and strain, dose administration route and method, and study type, study quality, among others.
- Treatment Group & Dosing Information- captures the dose groups (primarily in mg per kg of body weight per day), gender, life-stage, generation, and sacrifice time.
- Effect & Endpoint- captures, for any treatment group, the effects and endpoints deemed to be treatment-related using a standardized effect vocabulary.

Browsing

The "ToxRefDB Home" page enables quick browsing by chemical, either by name or by CAS number.

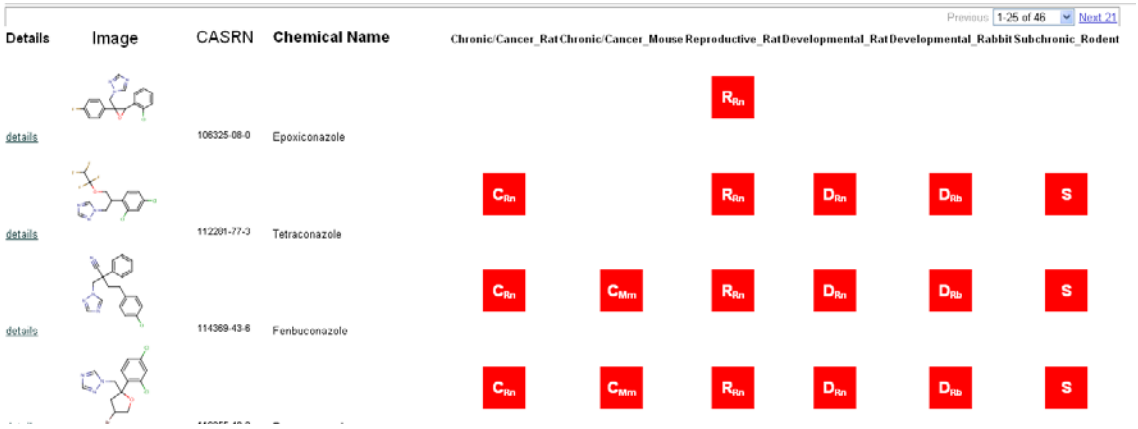
Searching

If a person has a particular chemical in mind, he or she can find it by either entering the name or CAS number or they can draw the structure.

- ToxRefDB Home: Search on Chemical Names- allows one to search for chemicals by their common and their many alternate names (See "Search by Name (ACToR Only)" for general instructions).
- ToxRefDB Home: Search on CAS Numbers - enables one to search by CAS number (See "Search by CAS Number (ACToR Only)" for general instructions)
- Search by Endpoint- allows a user to search ToxRefDB for a specific endpoint or set of endpoints across any single study type.

After conducting a search, it is possible to go back and edit what has been entered. To do this, DO **NOT** CLICK ON THE BACK BUTTON. Clicking on the back button will erase all of the data. Instead re-click on the "ToxRefDB Home" or "Search by Endpoint" page link on the left navigation bar.

Search Results



The search results are composed of a table with all the chemicals in that data table. The columns include:

Details- contains the generic chemical id and a link to the chemical's generic chemical page
Image – Drawing of the chemical
CASRN- CAS Number
Chemical Name- the chemicals proper name

A series of red boxes will appear to the right of the "Chemical Name" column. The appearance of a red box indicates that there is toxicity data available for the substance that appears under the "Preferred Name" column. The initials within the box denote the specific type of toxicological information that is accessible with ACToR. It does not necessarily mean that the substance has tested positive for a particular effect.

C_{Rn} (Chronic/Cancer Rat)- This category indicates that one or more long-term (usually 2-year) chronic and/or cancer rat studies have been conducted, entered into ToxRefDB, and considered of sufficient quality for public release.

C_{Mm} (Chronic/Cancer Mouse)- This category indicates that one or more long-term (usually 18-month to 2-year) chronic and/or cancer mouse studies have been conducted, entered into ToxRefDB, and considered of sufficient quality for public release.

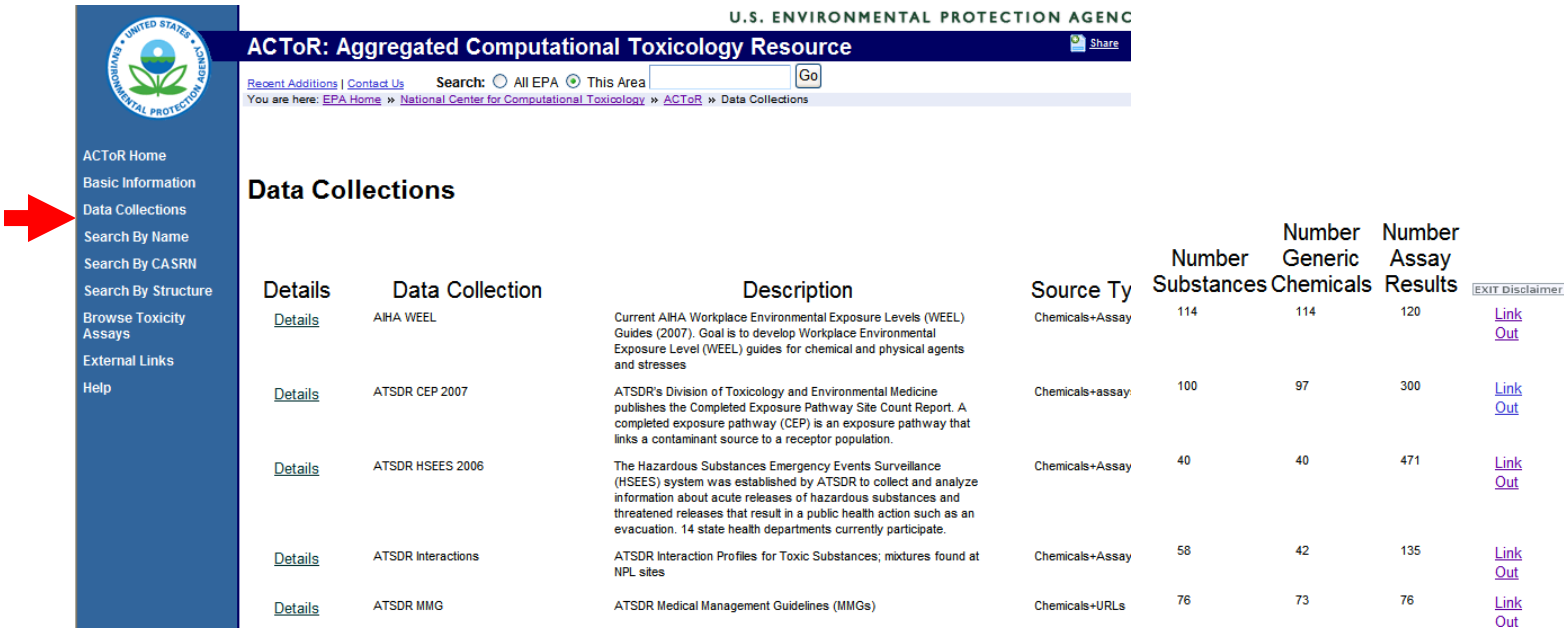
R_{Rn} (Reproductive Rat)- This category indicates that one or more multigenerational reproductive rat studies have been conducted, entered into ToxRefDB, and considered of sufficient quality for public release.

D_{Rb} (Developmental Rabbit)- This category indicates that one or more prenatal developmental rabbit studies have been conducted, entered into ToxRefDB, and considered of sufficient quality for public release.

S (Subchronic Rodent)- This category indicates that one or more medium-term (usually 90-day) sub-chronic rodent studies have been conducted, entered into ToxRefDB, and considered of sufficient quality for public release.

TOP

Data Collections (ACToR Only)



All information ACToR is organized by Data Collections. A data collection often contains substances, chemical structures and assays. The entire list of data collections in ACToR can be seen by selecting the Data Collections in the left hand navigation bar. For each collection, the following data are presented:

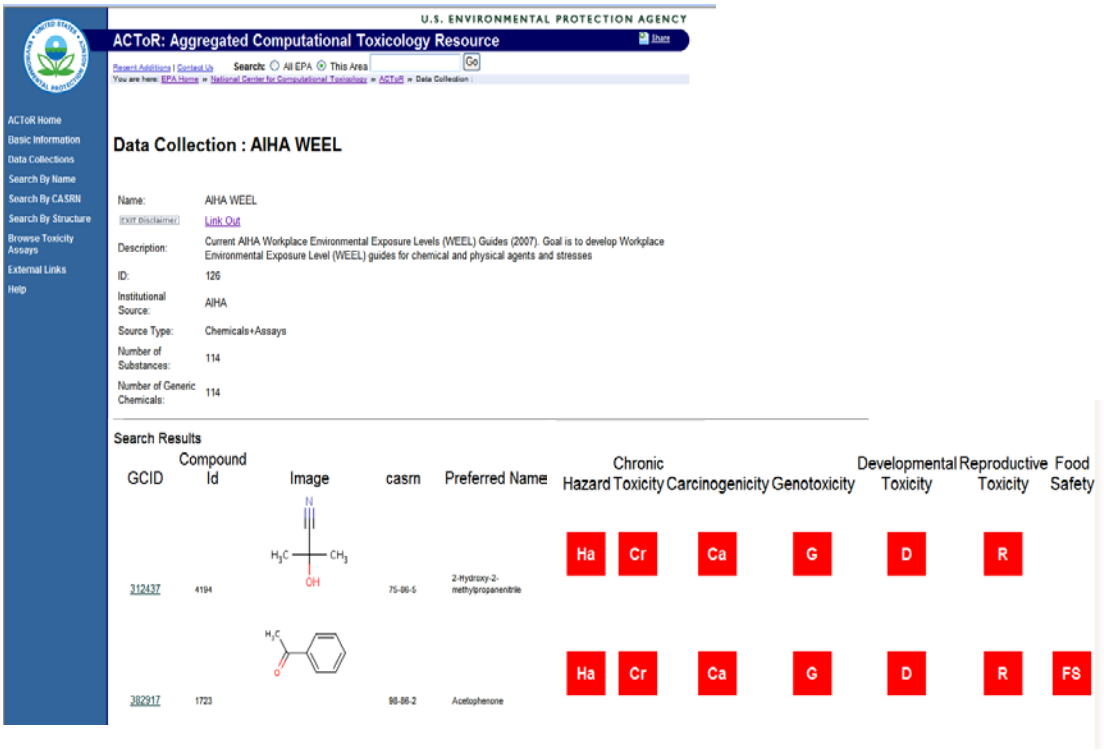
Details- contains a link to individual data collection page
Data collection- a short synopsis of the data collection
Description- short synopsis of the data collection
Source type- the type of information contained in that data collection. It can include:
- *Chemical*- chemical list of substances
- *Chemical + Assay*- a list of chemicals that have associated information
- *Chemical + Structure*- a list of chemicals and their corresponding structures
- *Chemicals +URL*- a list of chemicals having specific website for individual substances.
- *EPA target list*- target list of chemicals to investigate created by the EPA
- *Search by CASRN and name*- chemical list with names and CAS numbers

Number of substance- the number of substances within the data collection
Number of generic chemicals- the number of generic chemicals within the data collection. The number of generic chemicals may be less than number of substances if some substances do not have CAS numbers or if there are multiple substances with multiple names leaving identical CAS number.
Number of assay results- a measure of discrete assay for the substance contained in the collection.
Link out- provides a direct connection to the external website of a given data collection

To view the list of chemicals in the data collections, select the “Details” link at the left. This will take you to the Data Collection View. To navigate to the sources’ external site select the “Link Out” hyperlink at the far right in data collections table.

TOP

Data Collection View (ACToR Only)



The data collection page shows a summary of the information contained within a specific data collection. This page can be reached by clicking on data collection page and then selecting one of the "Details" links. Here, the information is divided into three parts: overview, and search results.

Overview

The top chart provides a brief overview of the data collection. This includes:

Name - name of the data collection
Link out - provides a direct path to the source's external website
Description - description of the data collection
ID - the internal id number of the data collection used for database management purposes
Institutional Source - the name of the institution that provided the data
Source Type – provides a cryptic description of the type of information contained in the data collection, e.g. a chemical list, an associated assay chemical- specific URLS, etc.
Number of Substances - the number substances in the data collection
Number of Generic Chemicals - the number of generic chemicals in the database. This number may be less than number of substance if some substances do not have CAS numbers or if there are multiple substances with the same CAS number.

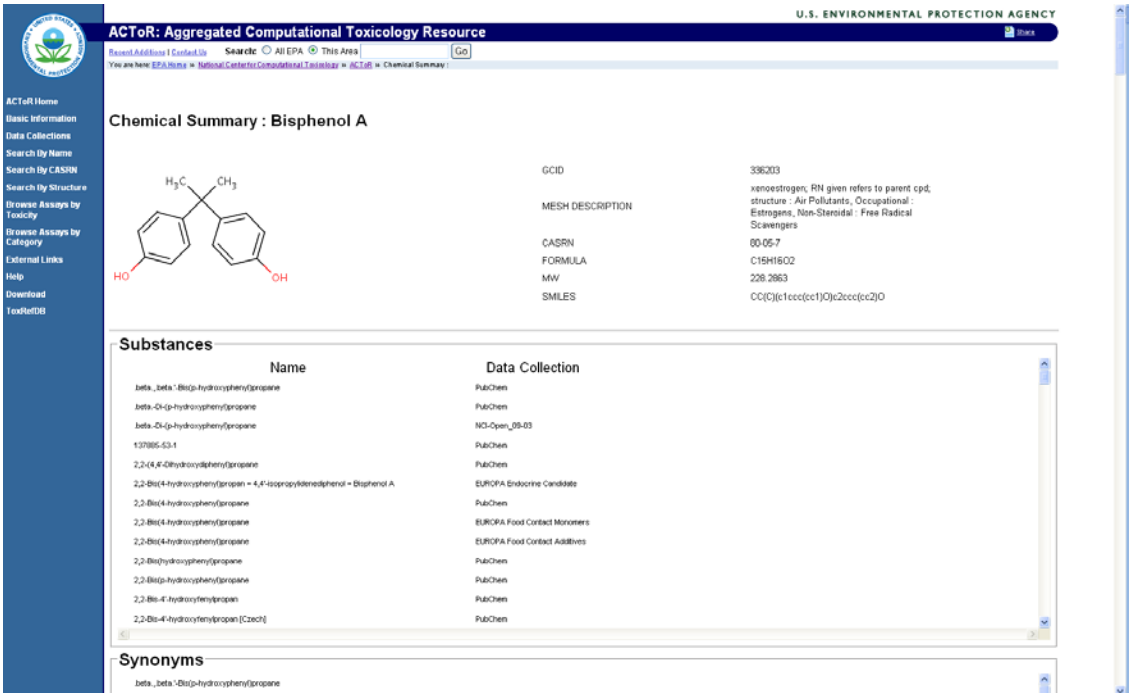
Search Results

The search results give a list of all the chemicals that appear in the data collection. For more information go to "how to use ACToR: search results".

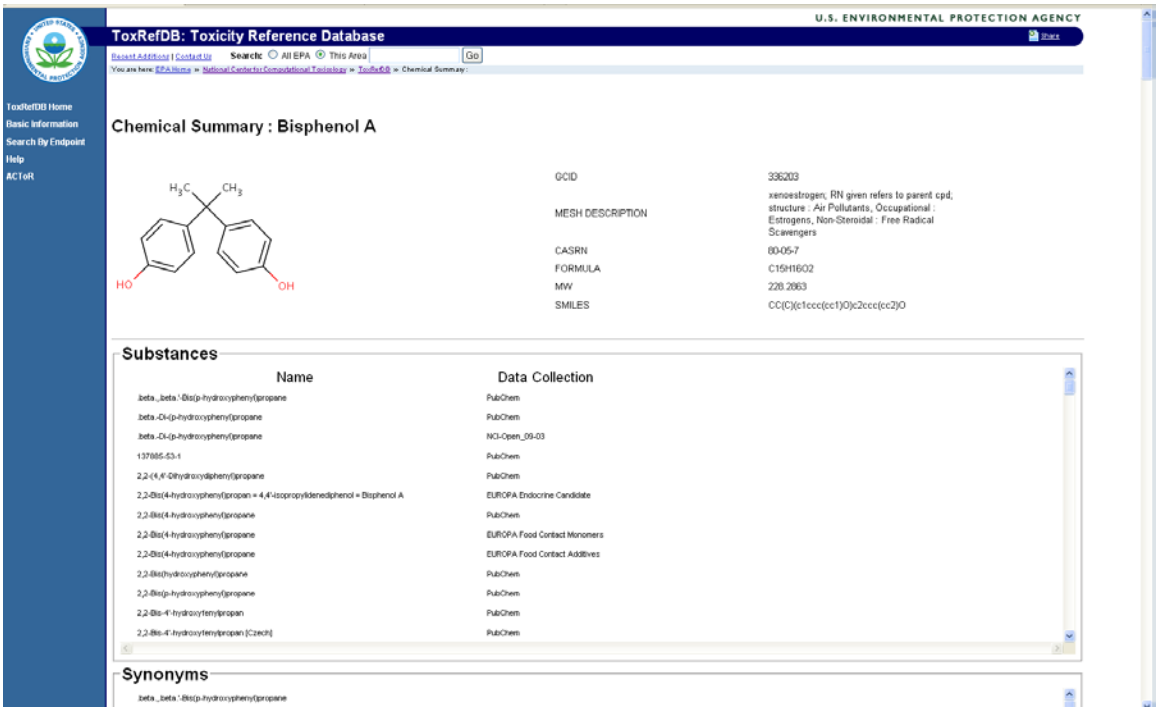
TOP

Generic Chemical View (ACToR & ToxRefDB)

ACToR View:



ToxRefDB View:



This page is accessed by clicking on "Details" for a specific chemical and provides links to all the information relating to this chemical i.e. from all data collections in the ACToR database and all study information from ToxRefDB. Data has been aggregated from all substances with a specific CASRN from all data collections. This page is divided into five main sections: Chemical Summary, Substances, Synonyms, Toxicology Data, Category Data.

Chemical Summary

GCID- Generic Chemical ID- an internal chemical ID within in ACToR
Mesh Description- When available, a brief description of the chemical that came from the MESH Database from the US Nation Library of Medicine
CASRN- The CAS **(**Chemical Abstracts Service) Registration Number
Formula- The chemical formula
MW- Molecular weight
SMILES - (Simplified Molecular Input Line Entry System) is a line notation used for representing molecules

Substance Data

The substance box provides a list of all the data collections that contain the chemical. Chemicals are aggregated by CASRN. The substance chart has the following categories:

Name- the chemicals proper name
Data Collection- the name of the data collection in which in appears

Synonyms

The synonyms box contains a list of all the CAS numbers and names used to describe the generic chemical.

Toxicology Data

The Chemical data is organized into different boxes. Each box has its own label and contains a list of assays and a summary of their results.

ToxRefDB Data – Detailed toxicological data available directly from ToxRefDB in a structured format. Information on toxicological outcomes are linked directly to the dose groups and studies (more detailed information below).
Hazard – Information on basic harm that can be caused by a chemical. Information under this category includes workplaces safety and first aid in case of exposure.
Acute Toxicity – Information on health effects due to short term exposure.
Subchronic Toxicity - Information on health effects due to intermediate term exposure.
Chronic Toxicity - Information on health effects due to long term exposure.

Carcinogenicity – Information from studies of cancer-causing ability of chemicals.
Genetic Toxicity- Information on the ability of chemicals to cause DNA damage.
Reproductive Toxicity- Information on the ability of a chemical to damage an organism's reproductive ability.
Neurotoxicity - Information on the ability of a chemical to damage nerve cells or tissues.
Developmental Neurotoxicity – Information on chemicals that cause neurological deficits during development
Immunotoxicity - determines how a chemical affects the immune system
Dermal Toxicity- includes studies about chemicals and what level of toxicity results from the substances being applied through the skin. May or may not have the skin as a target organ.
Respiratory Toxicity- contains data on how the chemicals that affect the respiratory system
Nephrotoxicity- measures to what degree that the chemical that affect the kidneys
Endocrine Effects- contains data about if and how the chemical affect hormone signaling and downstream processes
Cardiotoxicity- contains data bout how the chemical affects the heart
Ecotoxicity- includes data about how chemicals affect non-human species such as fish and amphibians
Food Safety- includes data that determines if the food can be safely used as an ingredient, additive or food wrapper
Toxicity other- contains other information about a chemical's toxicity

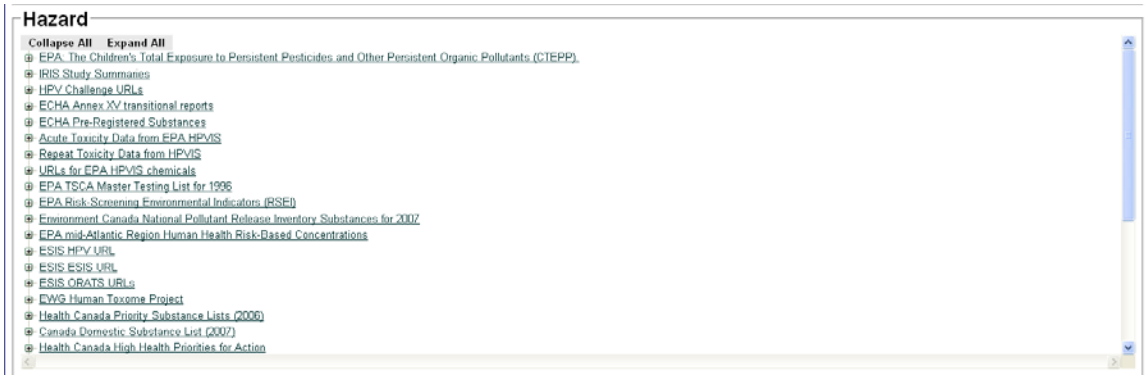
Category Data

The Chemical "Category" data is organized into different boxes. Each box has its own label and contains a list of assays and a summary of their results. These categories along with the "Toxicology Data" have been mapped to a taxonomy for accurate and consistent representations of specific data types.

Physico-Chemical Data – this is largely computed data on solubility, melting point, etc. based on chemical structure.
In Vitro – Primarily in vitro screening/profiling data coming out of the EPA's ToxCast research program, but will include processed data from Tox21 (an interagency research program for screening thousands of chemicals)
Chemical Manufacturing and Use Levels – This is data compiled by the EPA on industrial chemicals that are subject to TSCA (Toxic Substances Control Act)
Descriptive Data – a variety of tabulated descriptive data, for instance on intended use of the compound
Risk Management – This is a listing of international, U.S. Federal and state regulations and standards to which this chemical is subject.
Material Safety Data Sheet – if available, a link to the International Chemical

Safety Card, which summarizes information from the Material Safety Data Sheet.
Chemical Categories – these are chemical structure categories used, for instance, to make initial predictions of toxicity of a new compound based on similarity with compounds which have already been tested.
Pesticidal Mode of Action – If the compound is a pesticidal active ingredient, this will provide the intended biological mode of action
PubMed via MESH – If available, a link to the literature in PubChem based on the MESH term.
Notes – notes provided on the chemical for and of the data collections.
External Searches by NAME or CAS- this is a list of pre-computed URL links to search external, on-line databases based on the CASRN and preferred name of the chemical

How to move through the chart

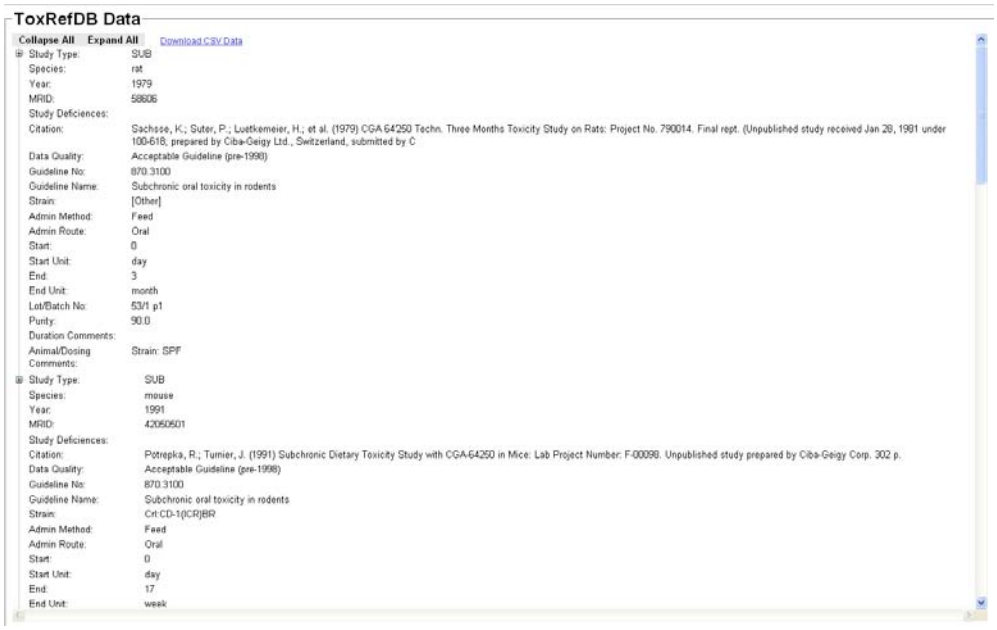


The assay names are green and underlined. Clicking on the assay names will take you to the assay view page. Clicking on the box next to the assay name will expand or collapse the result group. The result group has two columns:

Component Name- the assay component name
Value- includes number, names, text and URLs.

Each "assay" is initially collapsed for quicker viewing across large numbers of components. To expand a single assay component, click on the "+" to view the next level of information. Click "Expand All" to view all of the data under the specific category.

ToxRefDB Data
All other result groups are formatted and presented in a consistent fashion. Due to the unique structure of the data stored in ToxRefDB the information is being presented in three layers. For any given chemical with ToxRefDB data, there will be a collapsed view of the study design information for each study.



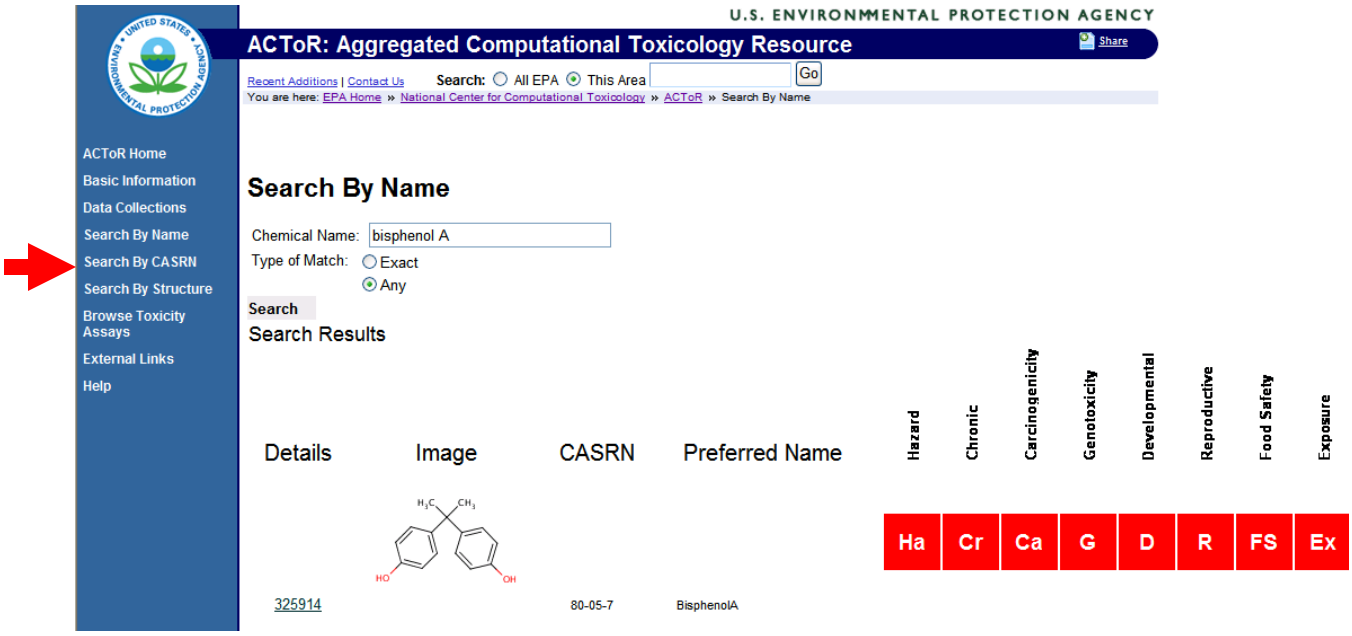
Like the other sections, to display all of the information for a given chemical press the "Expand All" button. To browse a particular study, select the "+" box and a view of the treatment groups and their associated treatment-related effects will be displayed. If the treatment group was not assigned any treatment-related effects then the text "No Effect Data Found" will be displayed. To DOWNLOAD all of the ToxRefDB data for a given chemical, select the "Download CSV Data" link and you will be able to download a zipped-up comma-separated file. This file will contain a flattened version of the data displayed in the browser. The table below provides a listing of all fields displayed with a short description.

Field Name	Description
CASRN	Chemical Abstract Service registry number
Chemical Name	Preferred chemical name
Study ID	ToxRefDB study identifier
Citation	Study citation (primarily based on EPA's OPPIN system)
MRID	OPP's unique study identifier
Year	Study Year (primarily year of study report)
Data Quality	Data quality (primarily assessed by adherence to guideline)
Study Deficiencies	Study-level deficiencies and comments (free-text; primarily as stated in review of study)
Study Type	Study type short name (CHR='Chronic';MGR='Reproductive',DEV='Development';SUB='Subchronic')
Guideline No	EPA/OPPTS harmonized guideline number
Guideline Name	EPA/OPPTS harmonized guideline name
Species	Species tested in study
Strain	Strain tested in study
Admin Method	Dose administration method (e.g., gavage, feed, water, etc.)
Admin Route	Dose administration route (e.g., oral, inhalation, etc.)
Start	Start of study dosing
Start Unit	Start of study units

End	End of study dosing
End Unit	End of study units
Lot/Batch No	Test material lot and batch number
Purity	Test material purity (%)
Duration Comments	Comments specific to the duration of the study (free-text)
Animal/Dosing Comments	Specifies test animal strain and dosing information
Treatment Group ID	ToxRefDB treatment group id
Generation	Generation or life-stage of treatment group
Dosing Period	Dosing period (e.g., initial-to-terminal, interim sacrifice, second mating, etc.)
Gender	Gender of treatment-group
Dose Level	Dose level for set of treatment groups
Dose	Administered dose (primarily mg/kg-bw/day)
Dose Unit	Dosing units (primarily mg/kg-bw/day)
Dosing Duration	Duration of dosing
Duration Unit	Dosing duration units
N	Number of animals per treatment group
Effect ID	ToxRefDB effect id
Effect Type	Type or Class of Effect (e.g., organ weight, gross pathology, neoplastic pathology, etc.)
Effect Target	Target of effect (e.g., organ, analyte, system, etc.)
Effect Description	Description of effect based on effect type and target
Direction	Direction of effect: Increase or Decrease
Target Site	Specific site of effect (e.g., cell type, region, etc.)
Effect Description (Free-Text)	Detailed description of effect (where not exact with standard vocabulary)
LOAEL	Assigned study-level LOAEL
Endpoint Category	High-level categories specific to each study type

TOP

Search by Name (ACToR Only)



To search for a chemical, type in the full or partial name in the text box. Select either "exact match" or "any match". Exact match will find the chemical whose name matches what you typed in. "Any" match will find matches that are similar to what you typed in. The search is performed against all of the synonyms that have been compiled for each generic chemical. When one clicks the search button, a chemical list chart appears with the results. Note that the search by name program does not accept SMILES or InCHI notation. To use SMILES or InCHI see the Search by Structure page.

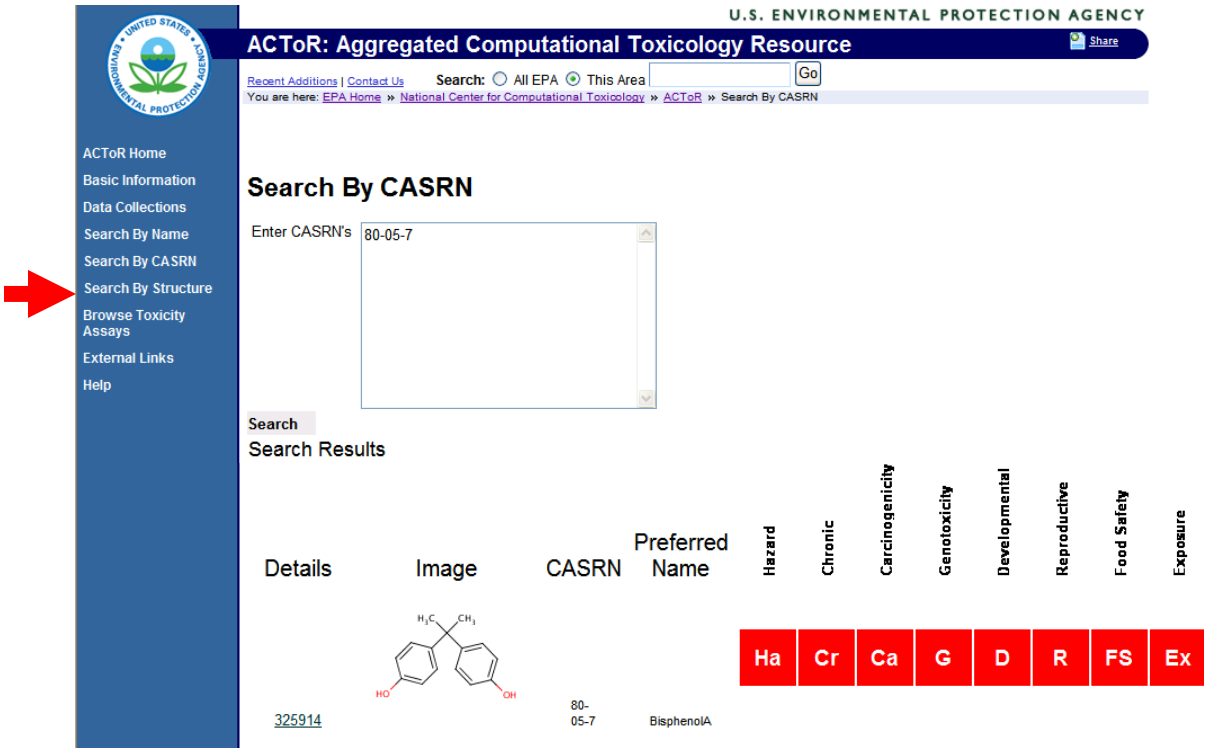
Search Results

See "How to use ACToR: Search Results".

After conducting a search, one can go back and edit what they entered. To do this, DO **NOT** CLICK ON THE BACK BUTTON. Clicking on the back button will erase all of your data. Instead re-click on the "Search by Name" button on the left navigation bar.

TOP

Search by CAS Number (ACToR Only)



Using CAS numbers is another way of locating and identifying chemicals. To find a chemical using a CAS number, type in one or more CAS numbers in the text box, separated by either commas or new lines. After search is completed, a standard chemical chart will appear.

Some examples of number in CAS format are:
7439-92-1
7440-50-8
79-34-5
39001-02-0

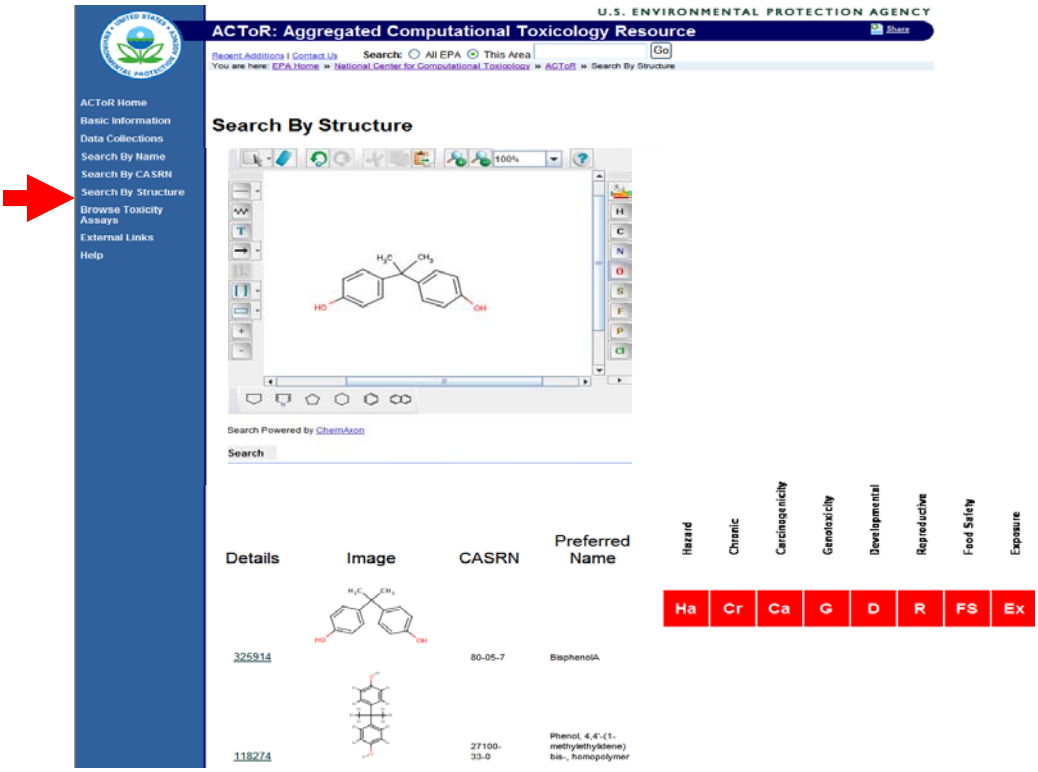
Search Results

See "How to use ACToR: Search Results".

A search may be edited after being submitted. To do this, DO **NOT** CLICK ON THE BACK BUTTON. Clicking on the back button will erase all of your data. Instead re-click on the "Search by CAS" button on the left navigation bar.

TOP

Search by Structure (ACToR Only)



Drawing Board

There are two major ways to construct a molecule

1. The typical way to construct a molecule is to select a template (see 1), bonds (arrows 2), and atoms (see 3 and 4). First, select a template then click on the canvas (see 1). Then, click on button 2 and select the bond type. To attach a bond, place the cursor over the molecule until a purple circle appears. If there is a need to connect two molecules, then click and hold the left mouse button and drag the other end of the bond to the other molecule until another purple circle appears before letting go of the left mouse button. To add atoms, either click on one of the "quick add" buttons (3) or select button 4. Button 4 causes a small window to appear with the periodic table on it. Select an element and then click close. Clicking on the "Query" tab, gives some more options that do not appear on the periodic table.

2. The fastest way to make molecules is to copy the molecule's SMILES or InCHI string and click on button 5.

For a more in-depth tutorial for this program, click on button 6 on the upper right hand corner. This takes you to the ChemAxon Help Site.

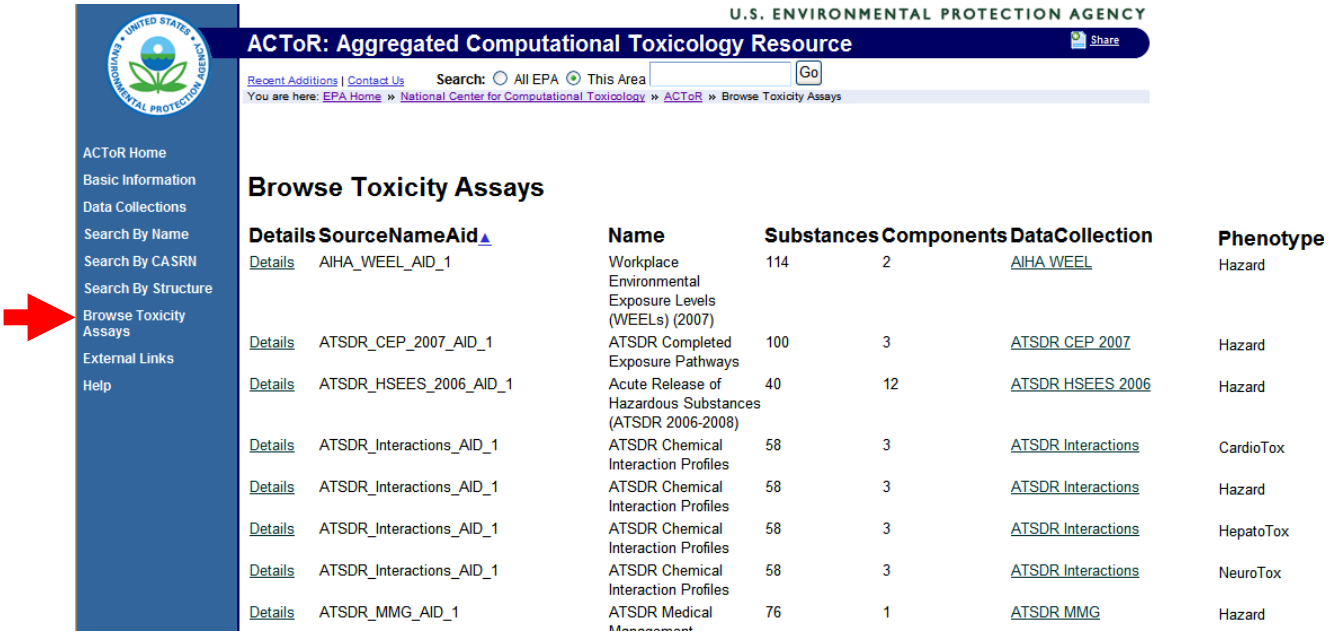
Search Results

See "How to use ACToR: Search Results".

After conducting a search, one can go back and edit what they entered. To do this, DO **NOT** CLICK ON THE BACK BUTTON. Clicking on the back button will erase all of your data. Instead re-click on the "Search by Structure" button on the left navigation bar.

TOP

Browsing Assays

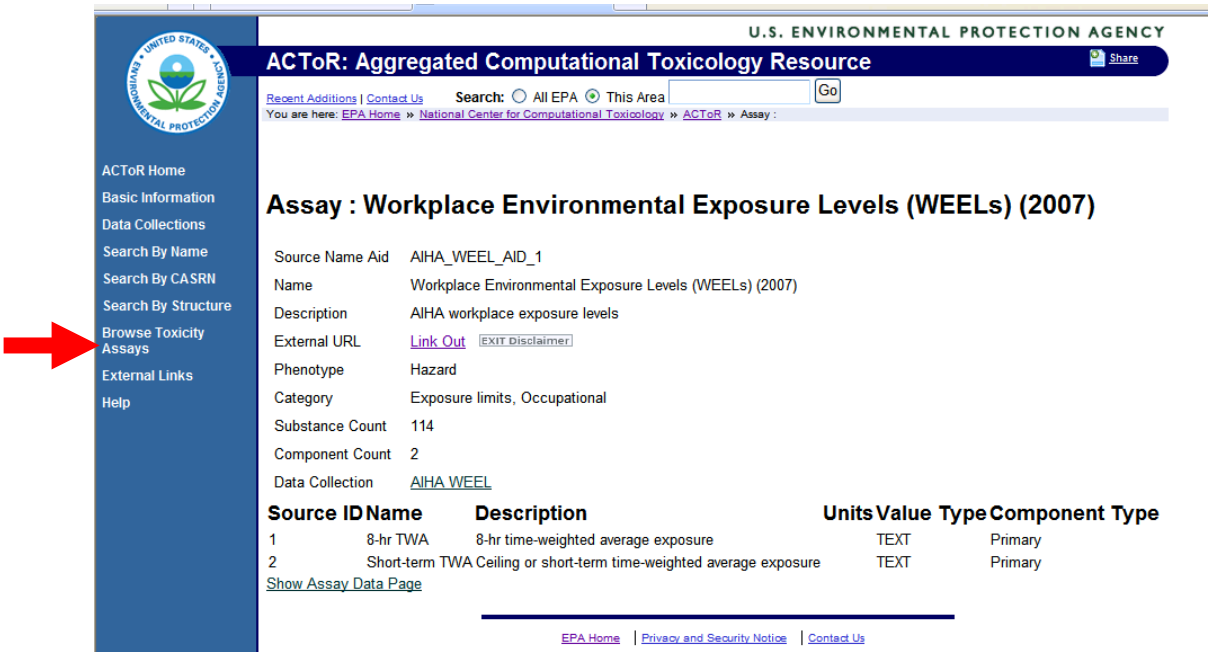


The browse assay page can be found by clicking on “Browse Toxicity Assay” in the left navigation button.
This page contains a collection of all the assays on the ACToR database. Assays are classified under one category but may have multiple phenotypes. Assays are organized alphabetically by assay name. Assays differ from data collections in that a data collection is all the information from a single source, while an assay is a collection of data arranged into a chart. So a single source can carry multiple assays.

Details- a link to a page that contains more details about a particular assay
Source Name AID- the location of the assay
Name- the name of the assay
Substances- the number of substances in the data
Components- the number of assay components
Data Collection- the name of the Data Collection
Category- The name of the Assays Categories
Phenotype- The name of the Assay Phenotypes

TOP

Assay View



The assay view gives a more in depth view of each assay. One can find this page by clicking on the "Details" link on the browse assay page. There are two sections to this page: the overview and the assay component chart.

Overview

Source Name AID- the location of the assay
Name- the name of the assay
Description- a brief summary of what the data entails
External URL- a link to the source
Components Count- the number of components
Data Collection- the name of the data collection

The above may be repeated if the assay has multiple phenotypes.

Assay Component Chart

Source ID	Name	Description	Units	Value Type	Component Type
1	8-hr TWA	8-hr time-weighted average exposure		TEXT	Primary
2	Short-term TWA	Ceiling or short-term time-weighted average exposure		TEXT	Primary

Show Assay Data Page

This section gives information about the type of data in the assay data. The assay component chart includes:

Source ID- the number assigned to the assay component within the assay
Name- the name of the assay component
Description- text description of the assay component
Units- what units are used for the assay component
Value Type- the type of values. Values include
    Float- numbers with decimals. This includes standard decimal notation,

exponential notation and scientific notation.
Integer- whole numbers such as -2, 0, 1, 2
Categorical- contains categories such as colors
Boolean- true or false
URL- contains the URL from a hyperlink
Text- is in text form but is not a category, URL or True or False

Component type- tells whether this is a primary or a secondary value. The values include

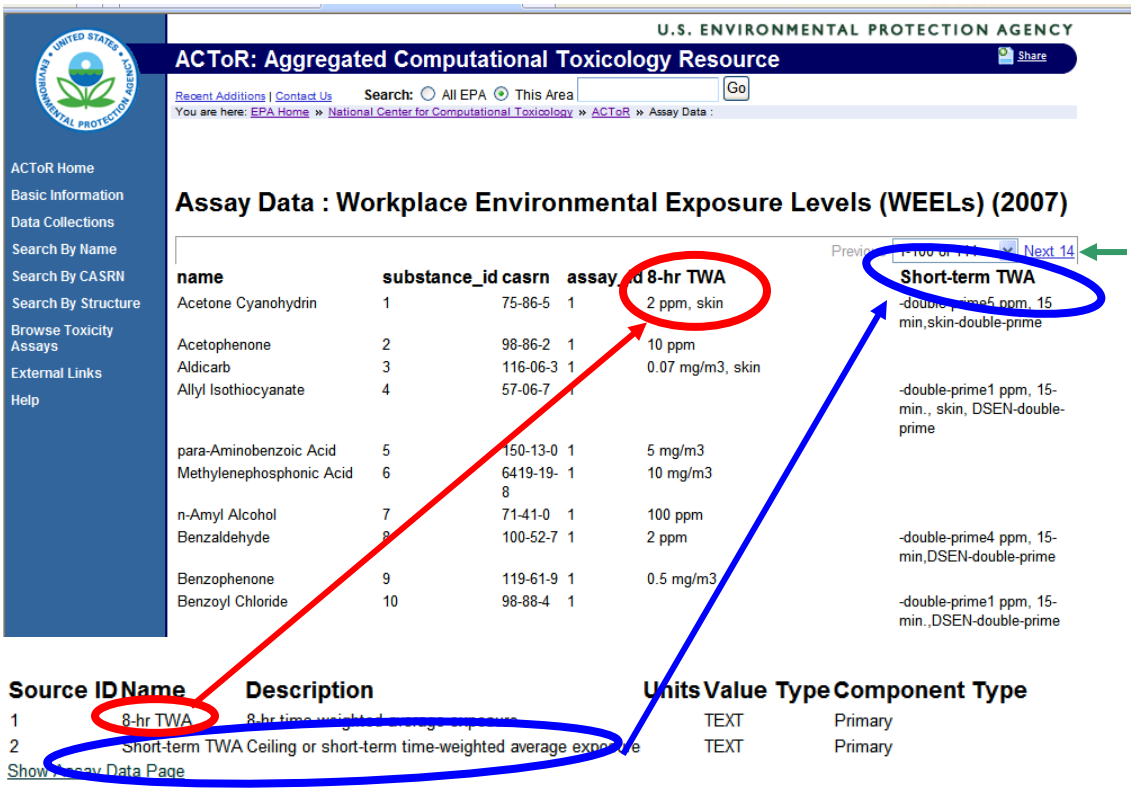
Primary- a number that one would computer with
Modifier- includes Variance and SD
Variance- the distribution of a sample (Statistics)
SD- standard deviation (Statistics)
Annotation- additional comments

Show Assay Data Page

This section contains a link that takes one to the Assay Data Page.

TOP

Assay Data Page



This page contains the assay data. In this chart, the columns headers are the assay components and the rows are the substance. The cells of this table are the assay results. The assay table can have more than one row or entry for the same substance, and elements in the data matrix can be empty.

One can find this page by clicking the link “Show Assay Data” on the Assay View page.

The charts vary with every assay. However, all of them contain the following:

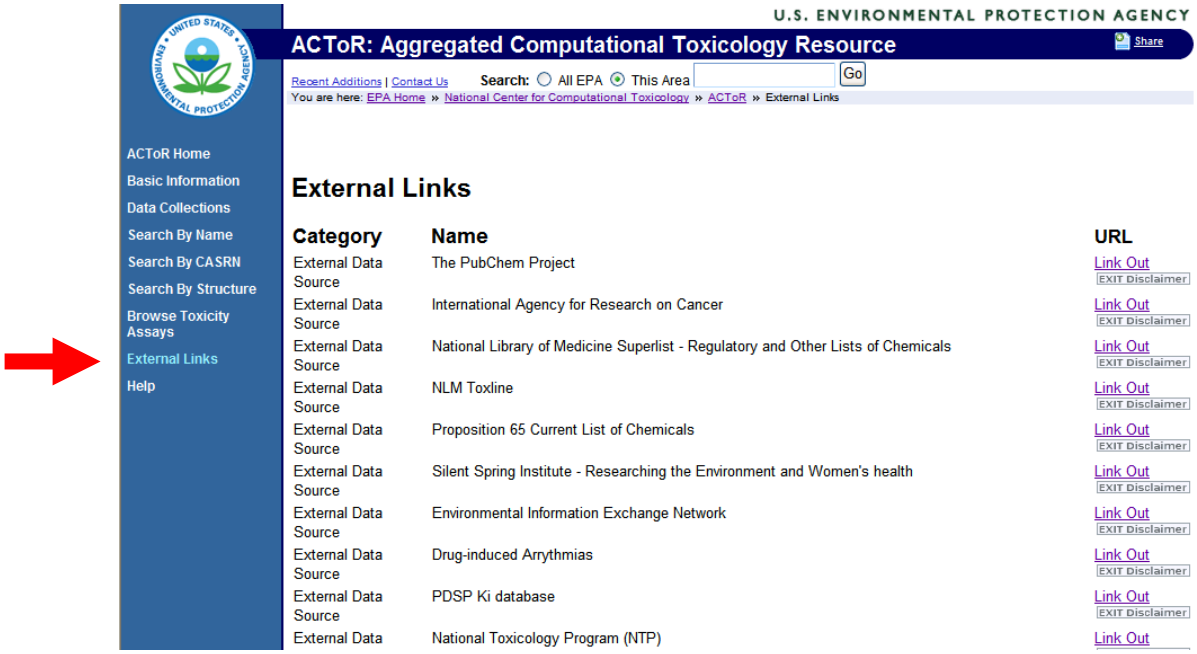
Name- the name of the chemical
Substance ID- the substance ID number
Assay ID- the ID number for the assay within the data collection

The Remaining column headings are assay components. The component names can also be found on the Assay View Page’s assay component chart.

Sometime, is the assay table is very long, there will be multiples “pages”. One can learn how to navigate through the chart by going to here.

TOP

External Links



This page contains a link to other large databases that contain information about chemicals.

Category: are all external data sources
Name: name of the website
URL: the hyperlink to the data

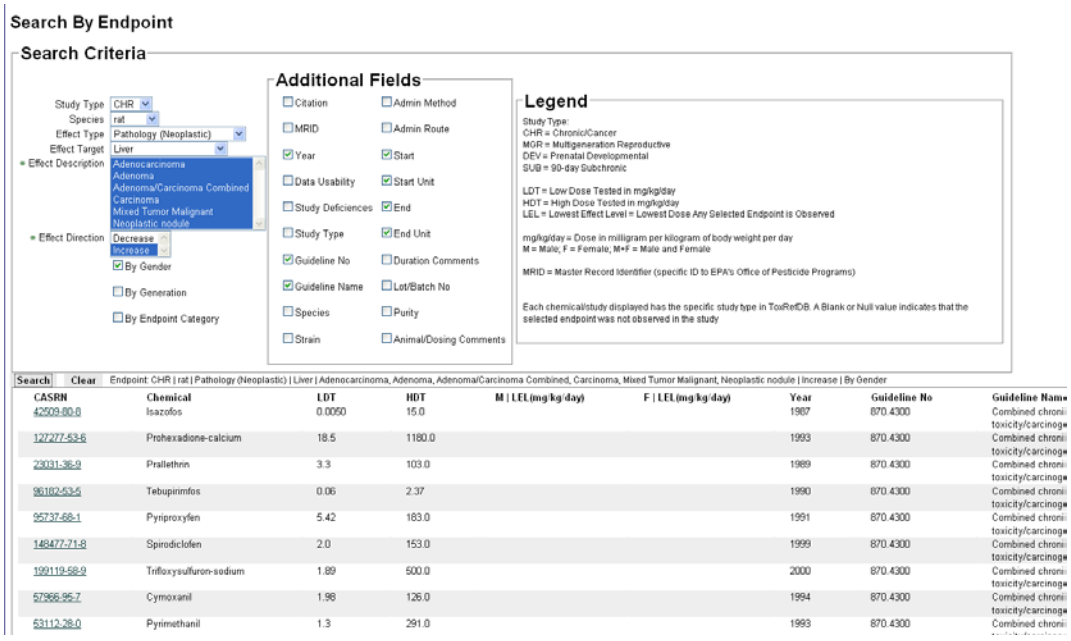
TOP

Search By Endpoint (ToxRefDB Only)

In ToxRefDB, toxicity study information is organized into four main tiers, chemical, study design, treatment group (animal) dosing information, and effects (i.e., endpoints). The "Search By Endpoint" page allows for flexible querying of the database for specific effects or sets of effects of interest. Since ToxRefDB captures on the treatment-related effects, the absence of an effect in the database is assumed to be a negative. This assumption is not always true and this should be considered whenever retrieving data from ToxRefDB.

Based on the standardized vocabulary, there is a defined set of terms for capturing the effect information on any given study. This vocabulary is constantly growing and changing based on feedback from subject matter experts. The top tier is the study type which consists of subchronic (SUB), chronic/cancer (CHR), reproductive (MGR), and developmental (DEV) studies. Each of these studies has been tested in rat, mouse and/or rabbit. Each unique effect is determined based on the combination of the Effect Type, Effect Target, Effect Description, and Effect Direction.

In the example below, the Chronic/Cancer (CHR) rat studies were selected and queried for whether or not the studies observed liver tumors [Pathology (Neoplastic) | Liver | Adenocarcinoma, Adenoma, Adenoma/Carcinoma Combined, Carcinoma, Mixed Tumor Malignant, Neoplastic nodule | Increase]. As described in the legend, a blank or null value means that the selected effects were not observed in the study. If an effect was observed than the lowest effect level (LEL: Lowest Dose Any Selected Endpoint is Observed) across the study is displayed.



In addition to selecting the specific set of endpoints to query, the user can choose to split the endpoint out “By Gender”, “By Generation” or “By Endpoint Category”. This will independently evaluate the LEL for each gender, generation and endpoint category, or any combination of the three parameters. In the example above, “By Gender” is selected and therefore the LEL (in mg/kg/day) for rat liver tumors is provided for both Male (M) and Female (F). The user can also select “Additional Fields”, which provide study-level details for each returned study and include:

Citation: Reference/citation information for the study (primarily pulled directly from the EPA’s OPPIN Database)

MRID: Master Record Identifier (specific unique study ID to EPA's Office of Pesticide Programs)

Year: Primarily the year of the final report or publication year

Data Usability: Evaluation criteria of data quality based on study guideline adherence (only Acceptable-Guideline and Acceptable-Non-guideline studies are made public)

Study Deficiencies: Free-text explanation of any study guideline deficiencies or comments regarding the curation of the data into the database

Study Type: Generic study type grouped based on similar guidelines and specific study types

Guideline No: OPPTS Harmonized Health Effects guideline number; although many of the older studies were performed under the older guideline system, each study entered into ToxRefDB is mapped to the new guideline numbers as a unique value for representing study types.

Guideline Name: Study type name associated with the guideline number

Species: Tested species; will be same as the selected value

Strain: Tested strain; due to the difficulty in deriving a unique, non-redundant list of strain that can be easily mapped to legacy data many have been listed as [other] and free-text values are displayed under the “Animal/Dosing Comments”

Admin Method: Dose administration method used in the study, primarily in the feed, water or via gavage

Admin Route: Dose administration route, primarily oral

Start & Start Unit: Time point in which dosing began; for a systemic toxicity study this is usually at time 0, but for developmental studies this is usually on a specific gestation day (GD) and for reproductive study the duration start is indicated by the time-premating.

End & End Unit: Time point in which dosing ends; for a systemic toxicity study this is usually at an absolute time (e.g., 2-years), but for developmental studies this is usually on a specific gestation day (GD) and for reproductive study the duration end is indicated by the last generation (e.g., 2 generation).

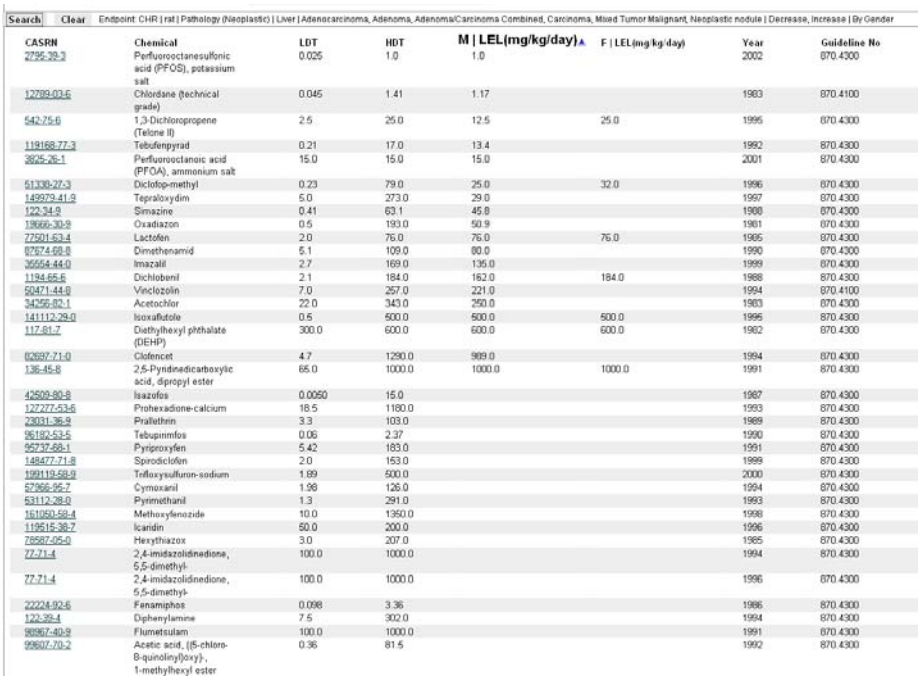
Duration Comments: Free-text explanation of any dosing or study duration information that is unable to be captured in other fields

Lot/Batch No: Test material lot and batch number, when provided

Purity: Test material purity

Animal/Dosing Comments: Free-text explanation of strain information or additional explanation of the dosing regimen not captured in the available fields

Once the “Search Criteria” is selected the user can select the “Search” button. While the search is ongoing a “Search In Progress” text will display (Do not continue to hit “Search). To the right of the “Search” and “Clear” buttons, the endpoint criteria will be displayed as shown in the screenshot below. Once the data is returned from the search into the table the user can sort on any field by selecting the column header label. The “arrow” to the right of the column header when ‘up’ means the data has been sorted in ascending order and when ‘down’, descending order.



Search | Clear | Endpoint: CHR | rat | Pathology (Neoplastic) | Liver | Adenocarcinoma, Adenoma, Adenoma/Carcinoma Combined, Carcinoma, Mixed Tumor Malignant, Neoplastic nodule | Decrease, Increase | By Gender

CASRN	Chemical	LDT	HDT	M | LEL(mg/kg/day)	F | LEL(mg/kg/day)	Year	Guideline No
2795-39-3	Perfluorooctanesulfonic acid (PFOS), potassium salt	0.025	1.0	1.0		2002	870.4300
12789-03-6	Chlordane (technical grade)	0.045	1.41	1.17		1983	870.4100
542-75-6	1,3-Dichloropropene (Telone II)	2.5	25.0	12.5	25.0	1995	870.4300
119168-77-3	Tebufenpyrad	0.21	17.0	13.4		1992	870.4300
3825-26-1	Perfluorooctanoic acid (PFOA), ammonium salt	15.0	15.0	15.0		2001	870.4300
51338-27-3	Diclofop-methyl	0.23	79.0	25.0	32.0	1996	870.4300
149979-41-9	Tepraloxydim	5.0	273.0	29.0		1997	870.4300
122-34-9	Simazine	0.41	63.1	45.8		1988	870.4300
19666-30-9	Oxadiazon	0.5	193.0	50.9		1981	870.4300
77501-63-4	Lactofen	2.0	76.0	76.0	76.0	1995	870.4300
87674-68-8	Dimethenamid	5.1	109.0	80.0		1990	870.4300
35554-44-0	Imazalil	2.7	169.0	135.0		1999	870.4300
1194-65-6	Dichlobenil	2.1	184.0	162.0	184.0	1988	870.4300
50471-44-8	Vinclozolin	7.0	257.0	221.0		1994	870.4100
34256-82-1	Acetochlor	22.0	343.0	250.0		1983	870.4300
141112-29-0	Isoxaflutole	0.5	500.0	500.0	500.0	1996	870.4300
117-81-7	Diethylhexyl phthalate (DEHP)	300.0	600.0	600.0	600.0	1982	870.4300
62697-71-0	Clofencet	4.7	1290.0	989.0		1994	870.4300
136-45-8	2,5-Pyridinedicarboxylic acid, dipropyl ester	65.0	1000.0	1000.0	1000.0	1991	870.4300
42509-80-8	Isazofos	0.0050	15.0			1987	870.4300
127277-53-6	Prohexadione-calcium	18.5	1180.0			1993	870.4300
23031-36-9	Prallethrin	3.3	103.0			1989	870.4300
96182-53-5	Tebupirimfos	0.06	2.37			1990	870.4300
95737-68-1	Pyriproxyfen	5.42	183.0			1991	870.4300
148477-71-8	Spirodiclofen	2.0	153.0			1999	870.4300
199119-58-9	Trifloxysulfuron-sodium	1.89	500.0			2000	870.4300
57966-95-7	Cymoxanil	1.98	126.0			1994	870.4300
53112-28-0	Pyrimethanil	1.3	291.0			1993	870.4300
161050-58-4	Methoxyfenozide	10.0	1360.0			1998	870.4300
119515-38-7	Icaridin	50.0	200.0			1996	870.4300
78587-05-0	Hexythiazox	3.0	207.0			1985	870.4300
77-71-4	2,4-imidazolidinedione, 5,5-dimethyl-	100.0	1000.0			1994	870.4300
77-71-4	2,4-imidazolidinedione, 5,5-dimethyl-	100.0	1000.0			1996	870.4300
22224-92-6	Fenamiphos	0.098	3.36			1986	870.4300
122-39-4	Diphenylamine	7.5	302.0			1994	870.4300
98967-40-9	Flumetsulam	100.0	1000.0			1991	870.4300
99607-70-2	Acetic acid, [(5-chloro-8-quinolinyl)oxy]-, 1-methylhexyl ester	0.36	81.5			1992	870.4300

The hyperlinked CAS numbers allow the user to go directly to the “Generic Chemical” page and look in more detail at the effects and additional data on the chemical. As with other section of ACToR and ToxRefDB, if the user would like to return to the “Search By Endpoint” page with returned search information use the “Search By Endpoint” link on the sidebar and do NOT press the back button.

TOP

Glossary

Active ingredients

An active ingredient is a substance in a drug which has some pharmaceutical or pesticide values. This is the opposite for inactive ingredients, which are only carriers that allow that body to processes the active ingredients better.

ACToR

(Aggregated Computational Toxicology Resource) is a collection of databases collated or developed by the EPA National Center for Computational Toxicology

Assay

An assay is a collection of data for substances from one data collection. Currently, an assay can be thought of a simple table. An assay falls into one data type category but may have multiple phenotypes. An assay can have more than one row or entry for the same substance, and elements in the data matrix can be empty.

Assay category

Assays are organized into a number of categories that describe the broad type of data presented. Several of these categories describe the level of biological organization being probed, while others describe the class of information being presented. The current sets of categories are:

PhysicoChemical- physical and chemical properties (in vitro and/or in silico)
Biochemical- chemical processes in living organisms that are non-cell based
Genomics- gene expression values or signatures
Cellular- cell-based assay
Tissue- tissue slice assay
Organ- focus on organs
Organism- focus on organisms (animal testing)
In vivo toxicology (tabular primary)- tabulated results from primary animal-based studies of chemical effects
In vivo toxicology (study listing primary)- primary studies are available but have not been tabulated
In vivo toxicology (tabular secondary)-tabulated data from secondary sources for in vivo toxicology studies
In vivo toxicology (summary calls)- derived summary determinations of risk
In vivo toxicology (summary report via URL)- links to text reports on the web for which specific data values are not directly accessible in tabular form
General Descriptive information- a brief description of the chemical
Regulation- listings of chemicals that fall under specific environmental laws, government mandates, or standards

Chemical Category

Chemical category is a listing of structural or uses categories, often intended for prioritization efforts

Chemical Summary URL- link to chemical summaries
Chemical Use Level- the amount of chemicals produced or used
Pesticidal mode of action (MoA)- explains how to drug or pesticide works and what the chemical targets

Assay component

An assay component defines one column or element of an assay. A component has a unique ID, a name, a description, a data type, and optionally units.

Assay phenotype

Some assays are characterized by toxicology phenotypes. This allows one to organize the data in ACToR into broad toxicity areas. The current set of phenotypes are:

Hazard – Information on basic harm that can be caused by a chemical. Information under this category includes workplaces safety and first aid in case of exposure.
Acute Toxicity – Information on health effects due to short term exposure.
Subchronic Toxicity - Information on health effects due to intermediate term exposure.
Chronic Toxicity - Information on health effects due to long term exposure.
Carcinogenicity – Information from studies of cancer-causing ability of chemicals.
Genetic Toxicity- Information on the ability of chemicals to cause DNA damage.
Reproductive Toxicity- Information on the ability of a chemical to damage an organism's reproductive ability.
Neurotoxicity - Information on the ability of a chemical to damage nerve cells or tissues.
Developmental Neurotoxicity – Information on chemicals that cause neurological deficits during development
Immunotoxicity - determines how a chemical affects the immune system
Dermal Toxicity- includes studies about chemicals and what level of toxicity results from the substances being applied through the skin. May or may not have the skin as a target organ.
Respiratory Toxicity- contains data on how the chemicals that affect the respiratory system
Nephrotoxicity- measures to what degree that the chemical that affect the kidneys
Endocrine- contains data about if and how the chemical affect hormone signaling and downstream processes
Cardiotoxicity- contains data bout how the chemical affects the heart
Ecotoxicity- includes data about how chemicals affect non-human species such as fish and amphibians

Food Safety- includes data that determines if the food can be safely used as an ingredient, additive or food wrapper
Toxicity other- contains other information about a chemical's toxicity
PK/metabolism- Information on pharmacokinetics and metabolism of xenobiotic chemicals

Assay result

An assay result is one data point for a single substance and a single assay component.

Assay types

There are two main types of assays: phenotypes and categories.

CAS

CAS (Chemical Abstract Services) Registry Number (for more information)

Some examples of number in CAS format are:
7439-92-1
7440-50-8
79-34-5
59325
39001-02-0
59001050

Chemical

A chemical is defined by a unique chemical ID in the database and can be either a substance or a compound.

Chemical structure

Diagram of a chemical- can be used to search for information about chemicals.

Compound

A compound is an entity with a chemical ID and chemical structure information, which may be a 2 or 3 dimensional molfile or a string representation. This can be SMILES or InCHII.

Data collection

A data collection is at minimum a set of substances with corresponding CAS registry numbers and names. Additional information may include chemical structures and assays. As mentioned above, a generic chemical links together data from many data collections on all substances that share a common CAS registry number. All data is initially compiled as part of a set of Data Collections.

Exposure data

Exposure studies, measure the amount of a substance that people and animals are exposed to. This data does not explain how the person was exposed to it or if it causes health problems.

Generic chemical

A generic chemical aggregates all data from all data collections for substances with a single given CAS number. It will have links to one or more substances and all of their related assay data, as well as all synonyms derived from the substances.

HPV and MPV

HPV stands for High Production Volume industrial chemicals and MPV for Medium Production Volume industrial chemicals

Includes the chemical structure and calculated physical chemical properties of compounds produced or imported into the United States

InCHI

The IUPAC International Chemical Identifier (InChI™) is a non-proprietary identifier for chemical substances (for more information).

Inert ingredient

An inert ingredient means any substance other than an active ingredient. Inert ingredients tend to be carriers for the active ingredients. (for more information)

In vitro

An experiment that is performed outside of a living organisms (for examples test tubes)

In vivo

Experimentation done on or inside of living organisms- other wise known as animal testing

MESH

is the U.S. National Library of Medicine's controlled vocabulary used for indexing articles for MEDLINE/ PubMed. MESH terminology provides a consistent way to retrieve information that may use different terminology for the same concepts.

SMILES

SMILES (Simplified Molecular Input Line Entry System) is a *line notation* (a typographical method using printable characters) for entering and representing molecules and reactions. (for more information)

Substance

A substance is an entity with a chemical ID, one or more names (including a CAS number) and potentially a URL pointing to primary data. One special name for the substance is the “source name sid” which is a unique alphanumeric label from the source, which allows a unique link back to the source.

TOP