

# *DSSTox Public Toxicity Database Network:*

## *Progress Report and New Initiatives to Expand Chemoinformatic Capabilities*

Ann Richard

Environmental Carcinogenesis Division  
National Health & Environmental Effects Research Lab  
US Environmental Protection Agency

COMPUTATIONAL TOXICOLOGY



## Chemistry-based Data Mining & Exploration:

Chemical(s)  
of concern

Chemical-  
specific  
data

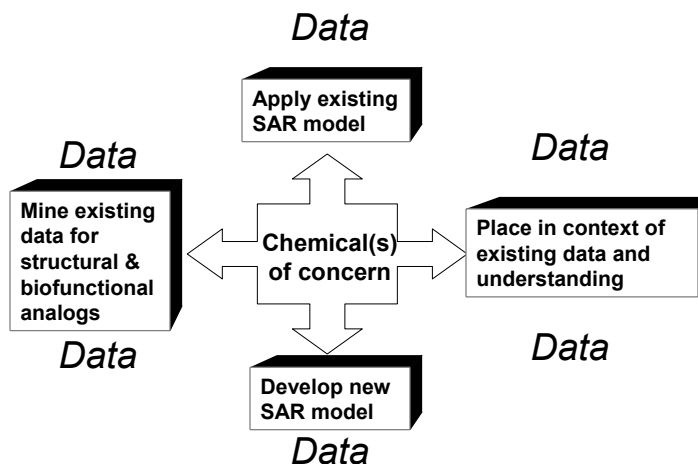
**Structure-Activity Relationships**

Structural  
analogs

Property  
analogs

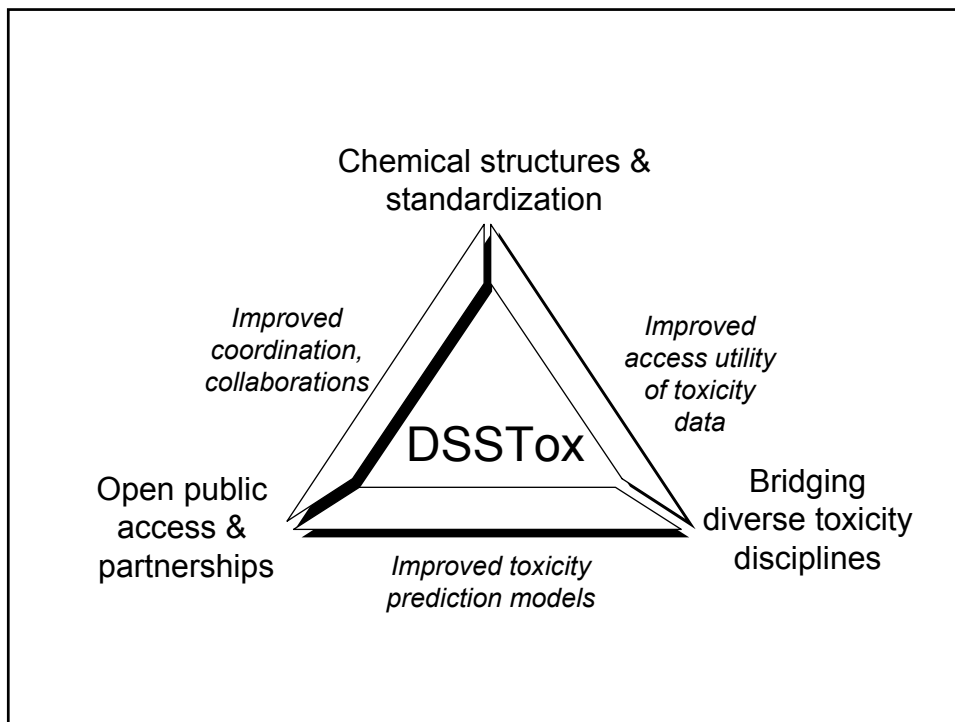
Biological/  
mechanistic  
analogs

## Structure-based Screening & Prioritization:



## Limitations of Public Toxicity Data for Use in SAR:

- Scattered sources
- Non-standard formats
- Diverse information content
- Lack of chemical structure annotation
- Cannot access full database



<http://www.epa.gov/nheerl/dsstox> **U.S. Environmental Protection Agency**

## Distributed Structure-Searchable Toxicity (DSSTox) Public Database Network

Recent Additions | Contact Us | Print Version Search:  **GO**

[EPA Home](#) > [Research & Development](#) > [National Health and Environmental Effects Research Laboratory](#) > [Distributed Structure-Searchable Toxicity \(DSSTox\) Public Database Network](#)

**The Distributed Structure-Searchable Toxicity (DSSTox) Database Network provides a community forum for publishing standard format, structure-annotated chemical toxicity data files for open public access. Project goals are to:**

- Encourage use of [DSSTox Standard Chemical Structure Fields](#) and [SDF](#) standard format files for publishing chemical toxicity databases;
- [Coordinate with outside public efforts](#) to encourage chemical structure annotation, data standardization, and open public access to toxicity databases;
- Involve the user community in the effort to migrate more public toxicity data into the DSSTox standardized format for publishing;
- Provide full, open access to toxicity data files for structure-analog searching and for facilitating development of improved models for predicting toxicity based on chemical structure.

**Distributed:** Decentralized set of standardized, field-delimited databases, each separately authored and maintained, that are able to accommodate diverse chemical toxicity data content;

**Structure-Searchable:** Standard format ([SDF](#)) structure-data files that can be readily imported into available [Chemical Relational Databases](#) and structure-searched;

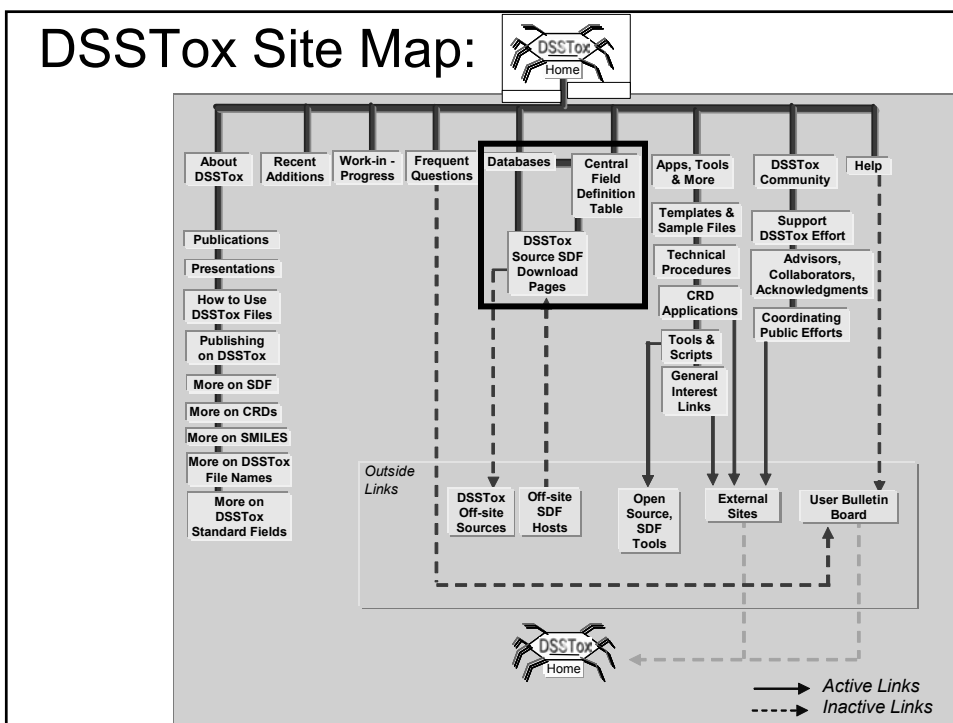
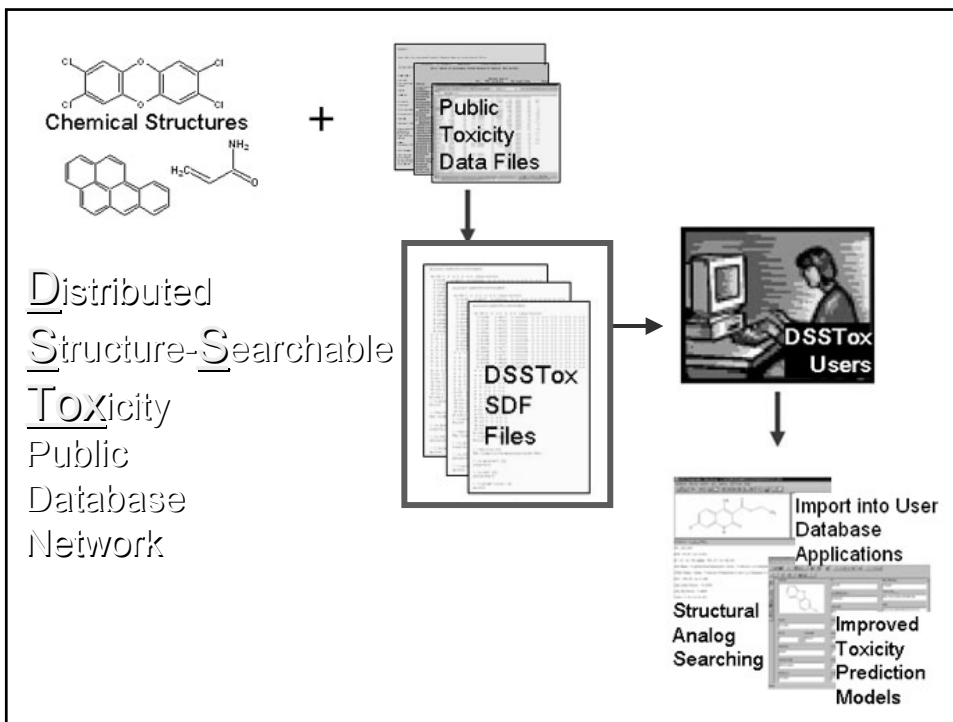
**Tox:** Toxicity data as it exists in widely disparate forms in current public databases, spanning diverse toxicity endpoints, test systems, levels of biological content, degrees of summarization, and information content.

**DSSTox Graphic Flowchart**

**Current Database Offerings:**  
 CPDBRM, CPDBHA, CPDBDG, CPDBPR\*,  
 DBPCAN  
 EPAFIM  
 NCTRER

\* CPDBPR last updated 29Mar04

**About DSSTox**  
 Work in Progress  
 Frequent Questions  
 Databases  
 Central Field Definition Table  
 Apps, Tools & More  
 DSSTox Community  
 Site Map  
 Glossary of Terms  
 Help




**U.S. Environmental Protection Agency**

## Distributed Structure-Searchable Toxicity (DSSTox) Public Database Network

Recent Additions | Contact Us | Print Version    Search:  **GO**

[EPA Home](#) > [Research & Development](#) > [National Health and Environmental Effects Research Laboratory](#) > [Distributed Structure-Searchable Toxicity \(DSSTox\) Public Database Network](#)



**Databases**

- **CPDBRM, CPDBHA, CPDBDO, CPDBPR: Carcinogenic Potency Database**  
**Summary Tables for Rat&Mouse, Hamster, Dog, and Non-human Primates**  
 Tumor target site incidence and TD50 potencies for 1354 chemical substances tested in rats and mouse, 80 chemical substances tested in hamsters, 5 chemicals tested in dogs, and 27 chemical substances tested in non-human primates; data reviewed and compiled from literature and NTP studies.  
*(SDF last updated 15Oct03)*
- **DBPCAN: Water Disinfection By-Products Database with Carcinogenicity Estimates**  
 Carcinogenicity estimates (high, moderate, low concern) by EPA experts using a mechanism-based analog SAR approach on a set of 209 water disinfection by-products, mostly small halogenated organics.  
*(SDF last updated 12Sep03)*
- **EPAFHM: EPA Fathead Minnow Aquatic Toxicity Database**  
 Acute toxicities of 617 chemicals tested in common assay, with mode-of-action assessments and confirmatory measures.  
*(SDF last updated 15Oct03)*
- **NCTRER: FDA's National Center for Toxicological Research - Estrogen Receptor Binding Database**  
 Estrogen receptor relative binding affinities tested in a common in vitro assay for 232 chemicals, listed with chemical class-based structure activity features.  
*(SDF last updated 17Jul03)*

**About DSSTox**

Work in Progress

Frequent Questions

Databases

Central Field Definition Table

Apps, Tools & More

DSSTox Community

Site Map

Glossary of Terms

Help

## DSSTox Toxicity Database Standards:

- Data file format (SDF)
- File naming convention
- Chemical structure information fields
- Documentation requirements
- Publishing requirements

# DSSTox SDF files

```

csChmFindW05030111462D
14 16 0 0 0 0 0 0 0999 V2000
0.1283 2.1977 0.0000 C 0 0 0 0 0 0 0 0 0 0
0.0000 0.7780 0.0000 C 0 0 0 0 0 0 0 0 0 0
1.0347 0.0000 0.0000 C 0 0 0 0 0 0 0 0 0 0
2.3261 0.5213 0.0000 C 0 0 0 0 0 0 0 0 0 0
2.4544 1.9411 0.0000 C 0 0 0 0 0 0 0 0 0 0
1.4197 2.7191 0.0000 C 0 0 0 0 0 0 0 0 0 0
3.6254 0.0000 0.0000 N 0 0 0 0 0 0 0 0 0 0
4.5318 1.0347 0.0000 C 0 0 0 0 0 0 0 0 0 0
3.8821 2.1977 0.0000 C 0 0 0 0 0 0 0 0 0 0
5.9516 1.0347 0.0000 N 0 0 0 0 0 0 0 0 0 0
6.7295 2.1977 0.0000 C 0 0 0 0 0 0 0 0 0 0
5.9516 3.4891 0.0000 C 0 0 0 0 0 0 0 0 0 0
4.5318 3.4891 0.0000 C 0 0 0 0 0 0 0 0 0 0
8.0209 2.1977 0.0000 N 0 0 0 0 0 0 0 0 0 0
1 2 1 0 0 0 0
1 6 2 0 0 0 0
2 3 2 0 0 0 0
3 4 1 0 0 0 0
4 5 1 0 0 0 0
4 7 2 0 0 0 0
5 6 1 0 0 0 0
5 9 1 0 0 0 0
7 8 1 0 0 0 0
8 9 2 0 0 0 0
8 10 1 0 0 0 0
9 13 1 0 0 0 0
10 11 2 0 0 0 0
11 12 1 0 0 0 0
11 14 1 0 0 0 0
12 13 2 0 0 0 0
M END
> <-Last Updated> (1)
5/3/01

> <-Source> (1)
http://potency.berkeley.edu/lcpdb.html

> <-Chemical> (1)
A-alpha-C

> <-CAS> (1)
26148-68-5

> <-Tested Form> (1)
neutral
    
```

## SAR Model Development "Training Sets"

- improved predictive tox models
- more comparable models
- dramatically lowered barriers to use

## Chemical Relational Database: sub-structure, text, property searching

- analog searches
- search across diverse toxicity endpoints
- search across chemical and toxicity fields

# DSSTox File Naming Standard:

EPAFHM\_v1a\_617\_15Oct03.sdf (.xls)

↓      ↓      ↓      ↓      ↓

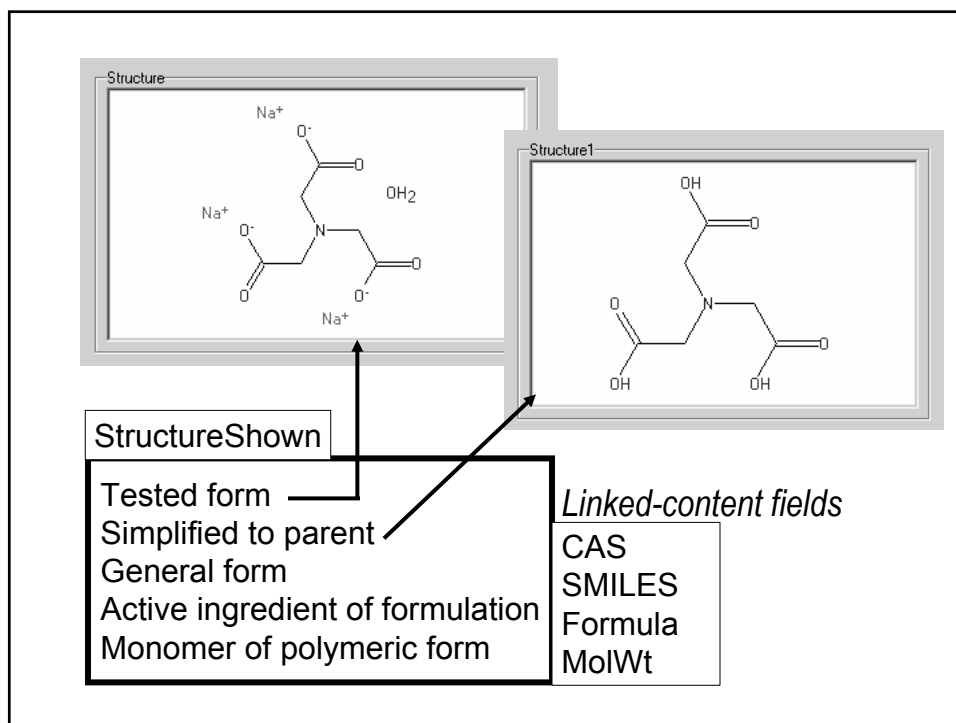
6 letter NAMEID      Number of records      Date of last modification      File type

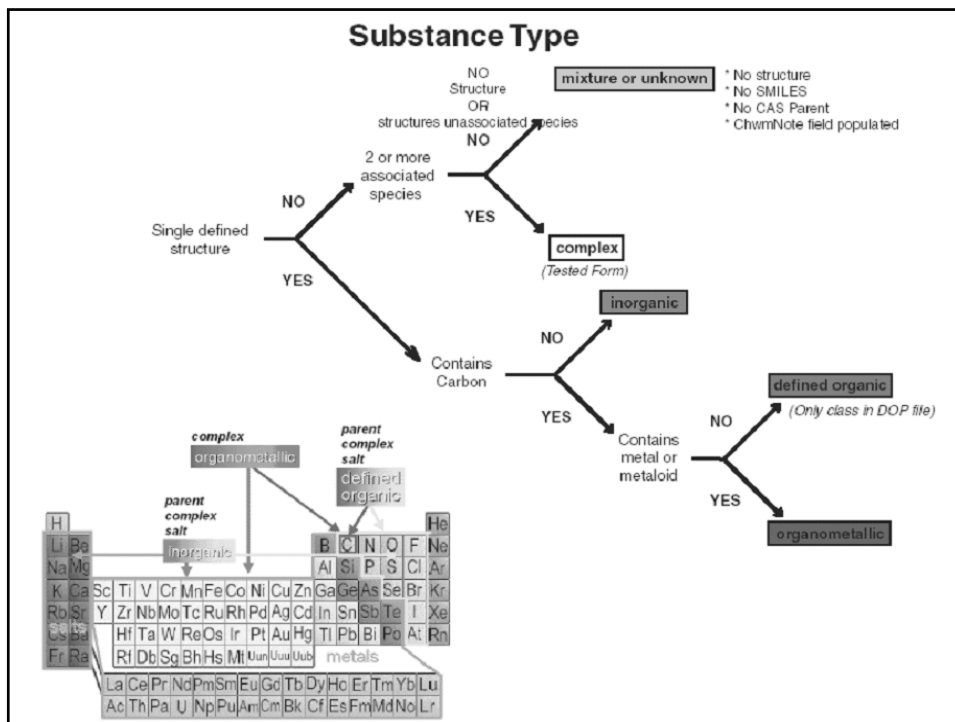
version 1: major modifications or data additions

revision a: error corrections, minor additions

## DSSTox Standard Chemical Fields:

- Structure *2D chemical structure*
- StructureShown *Description of displayed 2D structure*
  - *tested form, simplified to parent, predicted form, active ingredient of formulation, general form*
- Formula *Empirical molecular formula*
- MolWeight *Molecular weight in atomic units*
- CAS *Chem Abstracts Service No. for StructureShown*
- SMILES *Linear text notation for 2D StructureShown*
- DSSTox\_ID *Counter allows unique identification of record*
- DSSTox\_FileName *Name of file included in each record*
- ChemName *Chemical name from original data base*
- SubstanceType *Broad substance classification*
  - *defined organic, inorganic, organometallic, polymer, mixture or unknown*
- TestedForm *Tested form of chemical*
  - *parent, salt, complex, unknown or multiple forms*
- AddToParent *Salt counterions or complexed moieties*
- CAS\_TestForm *CAS No. for tested form of chemical*
- SMILES\_TestForm *SMILES code for the tested form of the chemical*
- ChemNote *Additional qualifier info for chemical fields*
  - *defined mixture characteristics, uncertainty in structure or CAS, stereochem, replicate, etc.*
- ChemCount *Counter for structure or CAS duplications in db*





## DSSTox Standard Chemical Fields:

CambridgeSoft ChemFinder Application view after SDF import

**ChemFinder [CPDBRM\_v1a\_1354\_15Oct03.dwg]**

File Edit View Text Search Record Scripts Online Window Help

CPDBRM\_v1a\_1354\_15Oct03

Structure TF

MoLID: 833  
Formula: C<sub>6</sub>H<sub>9</sub>NO<sub>6</sub>  
MolWeight: 191.14

StructureShown: tested form  
DSSTox FileName: CPDBRM\_v1a\_1354\_11Apr03

ChemName: Nitrolicacetic acid, trisodium salt, monohydrate

Substance Type: defined organic  
TestedForm: salt complex  
CAS: 18662-53-8

SMILES: N[CC(=O)O][CC(=O)O][CC(=O)O].[Na+].[Na+].[Na+]O

AddToParent: 3Na H2O  
ChemNote:

Structure CPDBRM\_DOP

StructureShown: simplified to parent  
CAS: 139-13-9

DSSTox FileName: CPDBRM\_DDP\_v1a\_1354\_11Apr03

SMILES: N[CC(=O)O][CC(=O)O][CC(=O)O]



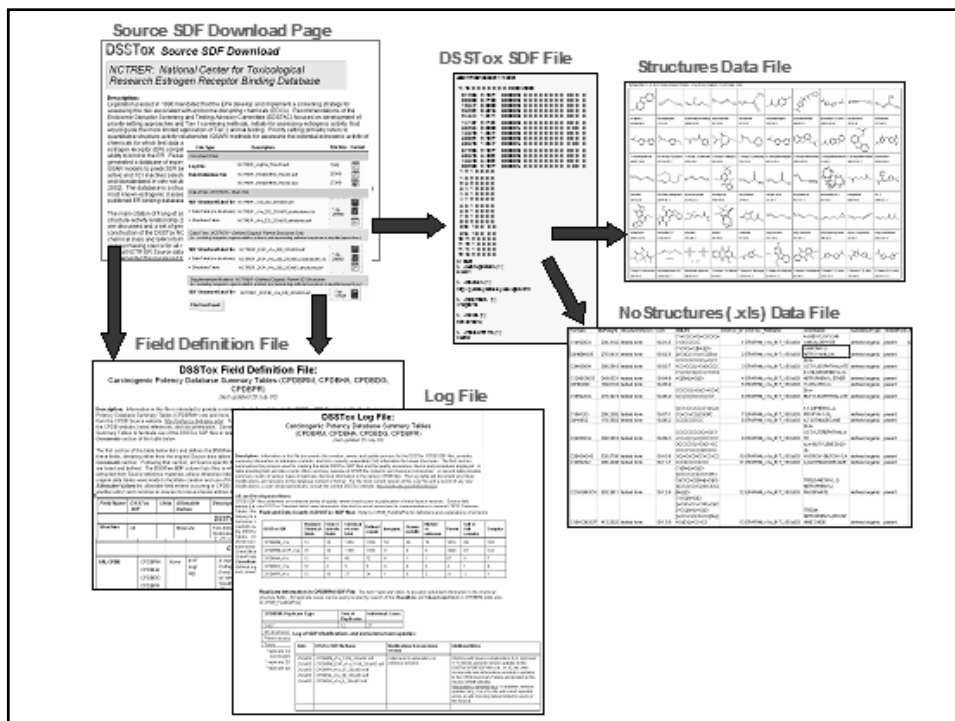
# DSSTox Source-Specific Fields:

CambridgeSoft ChemFinder Application view after SDF import

SAL CPDB neg	TD50 Rat 1760 m	TD50 Mouse 2660 m
Target Sites Rat Male kid	Target Sites Mouse Male kid	
Target Sites Rat Female kid ubl	Target Sites Mouse Female kid	
Target Sites Rat Both	Target Sites Mouse Both	
Other Species		

CPDBRM\_v1a\_1354\_15Oct03

CPDBRM\_DOP\_v1a\_1354\_15Oct03



*U.S. Environmental Protection Agency*

## Distributed Structure-Searchable Toxicity (DSSTox) Public Database Network

Recent Additions | Contact Us | Print Version | Search:

EPA Home > Research & Development > Health and Environmental Effects Research > DSSTox > DSSTox Central Field Definition Table

### Central Field Definition Table

**Indexed DSSTox SDF Files Included in Table:** Links provided to corresponding DSSTox Source SDF Download Page for each database listed by NAMEID.

**CPDB:** Carcinogenic Potency Database Summary Tables (CPDBRM, CPDBHA, CPDBDG, CPDBPR)  
**DBPCAN:** EPA Water Disinfection By-Products with Carcinogenicity Estimates  
**EPAFHM:** EPA Fathead Minnow Acute Toxicity Database  
**NCTRE:** NCTRE Estrogen Receptor Binding Database

**Central Index of DSSTox Databases**

Select a field name from the drop down list (\* denotes DSSTox Standard Chemical Fields)

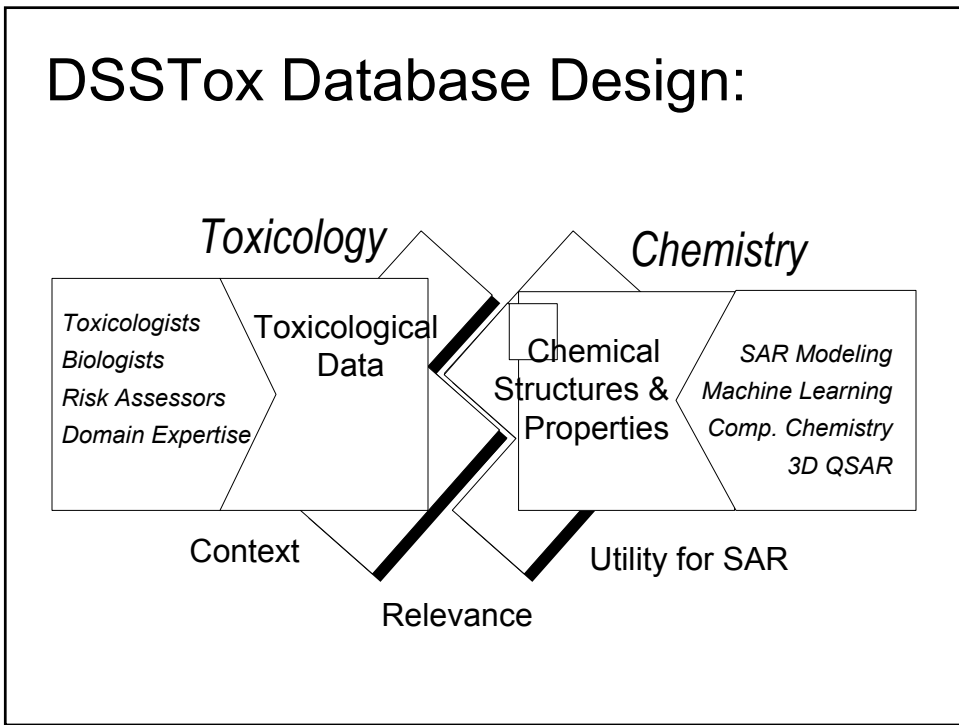
**All DSSTox fields listed alphabetically**

Field Name	Field Type	DSSTox SDF	Units	Allowable Values	Description
Activity Category <i>ER_RBA</i>		NCTRE	None	active strong/ active medium/ active weak/ slight binder/ not binder	For purposes of SAR analysis, Fara et al. (2001) divided the NCTRE data set into five main activity categories: active strong (ER_RBA > 1), active medium (1 > ER_RBA > 0.01), active weak (0.01 > ER_RBA > 1E-5), slight binder (max < 50% inhibition or ER_RBA < 1E-5) not binder (no activity with NA designation)
AddToParent	DSSTox Standard Chemical Fields	All DSSTox SDF files containing salts or complexes	None		organic and inorganic salts, entry specifies salt
AnalogCAS		DBPCAN	None	NOCAS/ #####.##.#	CAS of primary structural analog cited in SAR rationale for carcinogenic potential prediction, corresponding to <b>AnalogName</b> .

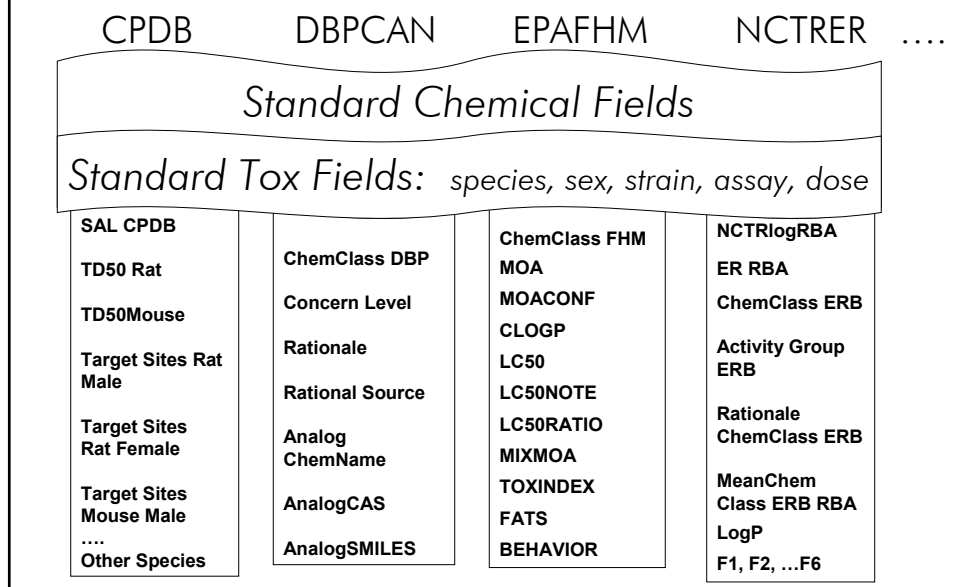
**Link to Standard Chemical Field Definitions Table**

**Link to DSSTox database containing field**

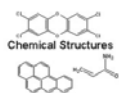
Page cont...  
 NAMEID | listed above  
 If a field is Fields gen Field Defin  
 In the tabl cannot oc explained entry is a the numbe files, we a separated distinguish



## Integrating Diverse Databases from a Chemical Structure Perspective:



## DSSTox Project Feasibility: *Direct Benefits to Data Source Collaborators*



*Chemical structure annotation*  
*Construct standardized DSSTox data file*



*Main-author of DSSTox "publication"*  
*Reach audience of potential users/collaborators*  
*Control presentation of data*



*Source DB can now be linked with DSSTox DB library and structure searched.*

# Migrate More Public Toxicity Data into DSSTox Standard Format: Phase II, III, ...

- FDA Max Recommended Therapeutic Dose - Pharmaceuticals
- NCTR Androgen, Thyroid, and Endocrine Disruption Databases
- NTP Immunotox battery tests
- NTP Rodent carcinogenicity bioassays, subchronic bioassays, developmental, repro, immuno, etc.
- UniLever Skin Sensitization database
- DNA Intercalators Database
- EPA's High Production Volume (HPV) chemicals structure index file
- EPA's Integrated Risk Information System (IRIS) structure index file
- ICVAM databases on LD50, skin sensitization, local lymph node assay, skin corrosivity, endocrine disruption, etc
- EPA pesticide ecotoxicity database
- Developmental toxicity database (literature – FDA, TOPKAT)

http://www.fda.gov/cder/Offices/OPS\_IO/MRTD.htm

www.fda.gov/cder/Offices/OPS\_IO/MRTD.htm

**FDA** U.S. Food and Drug Administration Department of Health and Human Services  
**CENTER FOR DRUG EVALUATION AND RESEARCH**

FDA Home Page | CDER Home Page | CDER Site Info | Contact CDER | What's New @ CDER

CDER Home About CDER Drug Information Regulatory Guidance CDER Calendar Specific Audiences CDER Archives

Search  GO powered by Google™

### Maximum Recommended Therapeutic Dose (MRTD) Database

FDA's Center for Drug Evaluation and Research, Office of Pharmaceutical Science, Informatics and Computational Safety Analysis Staff's Maximum Recommended Therapeutic Dose (MRTD) database contains values for 1,235 pharmaceuticals listed in *Martindale: The Extra Pharmacopoeia* (1973, 1983, and 1993) and *The Physicians' Desk Reference* (1995 and 1999).

Most of the MRTD values in the database were determined from p route of exposure and daily treatments, usually for 3 - 12 months. I dose treatment regimens to achieve desired pharmacological effects the MRTD database were antineoplastics and anesthetics and were When separate MRTDs were reported for different routes of expo database. In addition, some pharmaceuticals have different MRTD elderly patients. In this situation only MRTD values for the average

Pharmaceuticals that are administered orally are usually tested over as mg/day. We converted the mg/day unit to mg/kg-body weight (t kg. In contrast, the dose unit for most antineoplastic drug MRTDs i bw/day using the formula  $mg/kg-bw/day = mg/m^2/1.73$  for an average reported in parts per million (ppm) which were converted to mg/kg mg/kg-bw/day for an average 60 kg adult. MRTD values for the 12 1000 mg/kg-bw/day.

Comments or corrections should be sent to [kennard@cder.fda.gov](mailto:kennard@cder.fda.gov)

#### Maximum Recommended Therapeutic Dose (MRTD) Database (1,235 pharmaceuticals)

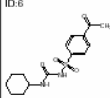
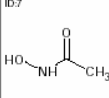
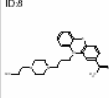
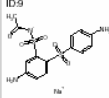
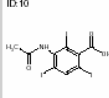
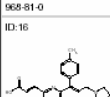
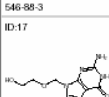
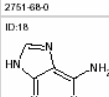
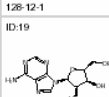
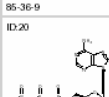
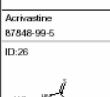
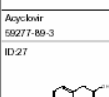
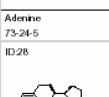
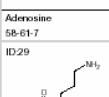
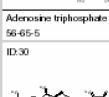
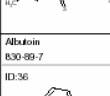
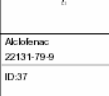
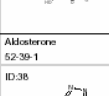
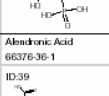
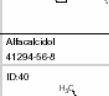
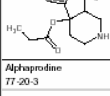
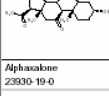
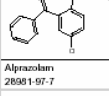
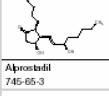
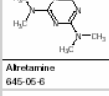
A B C D E F-G H-K L M N O P Q-R S T U-Z

MRTD (mg/kg-bw/day)	Generic Chemical Name	Chemical Structure (SMILES Code)
3.00000	Acemetacin	C1=C(C)C=C(C)C(=O)N(C)C(=O)C2=CC=C(C)C=C2C3=CC(=O)OCC(C)=O
0.20000	Acenocoumarol	C1=CC=C(C)C(=O)C1(O)C(C)C(C)C(=O)C2=CC=C(C)N(C)=O
50.00000	Acetaminophen	C1=CC=C(C)C(N)C(=O)C=C1
5.00000	Acetamide	C1=CC=C(C)C(N)C(=O)C=C1
16.70000	Acetazolamide	N1N=C(S)N(C)C(=O)S(=O)N(C)C(=O)C=C1
25.00000	Acetobenzamide	C2=C(C)C(C)C(=O)C(=O)C(S)C(=O)N(C)C(=O)N(C)C(=O)C=C2
16.70000	Acetylsalicylic acid	CC(=O)N(O)
10.00000	Acetophenone	C1=CC=C(C)C(=O)N(C)C(=O)C(C)C(=O)C(C)C(=O)C=C1
66.70000	Acetosalicylic acid	C2=C(N)C(=O)C(S)C(=O)C(C)C(=O)N(C)C(=O)N(C)C(=O)C=C2

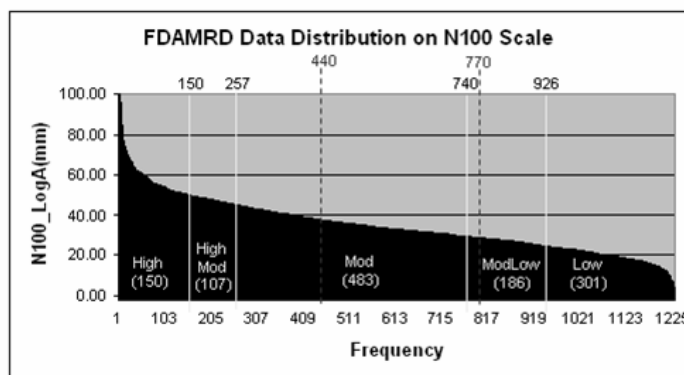
**FDAMDD:**  
 FDA's Maximum Daily  
 (recommended therapeutic)  
 Dose Database

✓ 1227 chemicals  
 ✓ 2173 CAS identified

- CAS\_Other
- SMILES\_nostereo
- SMILES\_stereo
- StudyType
- Species
- Endpoint
- Therap\_Cat
- Dose\_MRDD\_mg
- Dose\_MRDD\_mm
- LOG10\_MRDD\_mm
- N100\_MRDD\_mm
- Activity\_Category

ID:6  Acetolenside 968-61-0	ID:7  Acetylsalicylic acid 546-88-3	ID:8  Acetphenazine 2751-68-0	ID:9  Acetosalone 126-12-1	ID:10  Acetic acid 65-36-9
ID:16  Acrivastine 87946-99-5	ID:17  Acyclovir 59277-89-3	ID:18  Adenine 73-24-5	ID:19  Adenosine 58-61-7	ID:20  Adenosine triphosphate 56-65-5
ID:26  Albutin 830-89-7	ID:27  Alclerac 22131-79-9	ID:28  Alkosterone 52-39-1	ID:29  Alendronic Acid 66376-36-1	ID:30  Allicaltidol 41294-56-8
ID:36  Alphaprodine 77-20-3	ID:37  Alphacabone 23930-19-0	ID:38  Alprazolam 28981-97-7	ID:39  Alprostadil 745-65-3	ID:40  Alretamine 645-06-6
ID:46  Amifostine 29537-89-6	ID:47  Amikacin 37517-28-6	ID:48  Antrocaproic acid 60-52-2	ID:49  Aminoblic acid, 4- 54-62-6	ID:50  Aminoglutethimide 125-84-6

**FDAMDD:**  
 Distribution of N100\_MRDD\_mm values and  
 Activity\_Category cutoffs for 1227 chemicals



**N100 Scale:**  
 High (100-50)  
 HighMod (50-45)  
 Mod (45-30)  
 ModLow (30-25)  
 Low (25-0)

**MCASE (approx):**  
 High (100-37.8)  
 Mod (37.8-29.3)  
 Low (29.3-0)

# NTP/IMP: National Toxicology Program Immunotox Battery

- 50 chemicals
- 12 immunotox measures
- comparison to rodent carcinogenicity response

FUNDAMENTAL AND APPLIED TOXICOLOGY 10, 2-19 (1988)

### METHODS EVALUATION

#### Development of a Testing Battery to Assess Chemical-Induced Immunotoxicity: National Toxicology Program's Guidelines for Immunotoxicity Evaluation in Mice

MICHAEL I. LUSTER,\* ALBERT E. MUNSON,† PETER T. THOMAS,‡  
MICHAEL P. HOLSAPPLE,† JAMES D. FENTERS,‡ KIMBER L. WHITE, JR.,†  
LLOYD D. LAUER,§ DORI R. GERMOLEC,\* GARY J. ROSENTHAL,\* AND JACK H. DEANS§

\*Systemic Toxicology Branch, National Toxicology Program, National Institute of Environmental Health Sciences, P.O. Box 12231, Research Triangle Park, North Carolina 27709; †Department of Pharmacology and Toxicology, Medical College of Virginia, Virginia Commonwealth University, Richmond, Virginia 23298; ‡Life Sciences Division, IIT Research Institute, 10 West 35th Street, Chicago, Illinois 60616; and §Chemical Industry Institute of Toxicology, Department of Cell Biology, P.O. Box 12137, Research Triangle Park, North Carolina 27709

O1	O2	O3	O4	O5	O6	O7	O8	O9	O10
 Sedative 76-05-8	 4-Chloro-1,2-dichloro-benzene 118-06-1	 4-Chloro-1,2-dichloro-benzene 205-30-4	 Isoniazid 7064-02-1	 Nifedipine 1446-86-8	 Nifedipine 90-07-6	 Nifedipine 90-07-6	 Nifedipine 90-07-6	 Nifedipine 90-07-6	 Nifedipine 90-07-6
 Sedative 93-58-4	 4-Chloro-1,2-dichloro-benzene 118-06-1	 4-Chloro-1,2-dichloro-benzene 205-30-4	 Isoniazid 7064-02-1	 Nifedipine 1446-86-8	 Nifedipine 90-07-6	 Nifedipine 90-07-6	 Nifedipine 90-07-6	 Nifedipine 90-07-6	 Nifedipine 90-07-6
 Sedative 93-58-4	 4-Chloro-1,2-dichloro-benzene 118-06-1	 4-Chloro-1,2-dichloro-benzene 205-30-4	 Isoniazid 7064-02-1	 Nifedipine 1446-86-8	 Nifedipine 90-07-6	 Nifedipine 90-07-6	 Nifedipine 90-07-6	 Nifedipine 90-07-6	 Nifedipine 90-07-6
 Sedative 93-58-4	 4-Chloro-1,2-dichloro-benzene 118-06-1	 4-Chloro-1,2-dichloro-benzene 205-30-4	 Isoniazid 7064-02-1	 Nifedipine 1446-86-8	 Nifedipine 90-07-6	 Nifedipine 90-07-6	 Nifedipine 90-07-6	 Nifedipine 90-07-6	 Nifedipine 90-07-6
 Sedative 93-58-4	 4-Chloro-1,2-dichloro-benzene 118-06-1	 4-Chloro-1,2-dichloro-benzene 205-30-4	 Isoniazid 7064-02-1	 Nifedipine 1446-86-8	 Nifedipine 90-07-6	 Nifedipine 90-07-6	 Nifedipine 90-07-6	 Nifedipine 90-07-6	 Nifedipine 90-07-6

FUNDAMENTAL AND APPLIED TOXICOLOGY 18, 200-210 (1992)

### Risk Assessment in Immunotoxicology

#### I. Sensitivity and Predictability of Immune Tests

MICHAEL I. LUSTER,\* CHRISTOPHER PORTER,† D. GAYLA PAIT,\*<sup>1</sup> KIMBER L. WHITE, JR.,‡  
CHRIS GENNINGS,‡ ALBERT E. MUNSON,§ AND GARY J. ROSENTHAL,\*

\*Systemic Toxicology Branch and †Statistics and Biomathematics Branch, National Institute of Environmental Health Sciences, NIEHS, Research Triangle Park, North Carolina 27709; ‡Department of Biostatistics and Department of Pharmacology and Toxicology, Medical College of Virginia/Virginia Commonwealth University, Richmond, Virginia 23298

## Coordinating Public Efforts:



- ACD/Labs (Advanced Chemistry Development) ChemFolder Public Databases
- Cambridge-Soft's ChemFinder.Com Chemical Search Website
- FDA (Food & Drug Administration) Center for Drug Evaluation & Research
- ILSI (International Life Sciences Institute) SAR Toxicity Database Project, in collaboration with LHASA, Lmt.LIST (LeadScope In Silico Tox) Focus Group
- LIST (LeadScope In Silico Tox) Focus Group
- MGED: MIAMI-Tox
- NCI (National Cancer Institute) Public Data Outreach – Structure Web Browser
- NIEHS's National Center for Toxicogenomics
- NLM (National Library of Medicine) TOXNET
- NTP (National Toxicology Program) On-line Public Databases
- SRC (Syracuse Research Corporation) PBT-Profiler and Analog Search Tools

ChemFinder.Com  
Database & Internet Searching

Online Training: ChemDraw 8.0  
November 15, 1999

NCI Structure-Browser

### Advanced Chemistry Development

Web Librarian DBs

• The NTP/IMP database includes a number of public databases...  
• The NTP/IMP database includes a number of public databases...  
• The NTP/IMP database includes a number of public databases...

### ISSCentral – Freeware CRD

ISSCentral – Freeware CRD

SEARCH – Query Molecular Database

ISSCentral provides an environment for making up your own CRD files. Basic operations like adding, deleting, saving, and opening CRD files are supported. More complex operations like creating a query, saving, and opening a query are also supported. For more information on how to use the system please refer to the online help file located by the @ symbol on the right hand side of this page.

### EPA/Syracuse Res Corp

EPA/Syracuse Res Corp

Parachlor, Bisphenol-A, and Toxic Profile Both rated for Organic Chemicals On-Line

**PBT Profiler**  
A Component of CPPT's P2 Framework  
Assessing Chemicals in the Absence of Data

**ChemFinder.com**  
Database & Internet Searching

ChemStore.com ChemFinder.com  
ChemNow.com ChemCAB.com  
CambridgeSoft.com

**CALL ME NOW**

**Online Training : ChemDraw 8.0**  
CHEMICAL STRUCTURE DRAWING STANDARD REGISTER  
November 19, 2003 Join from your desk, classroom, or office

Enter a Chemical Name, CAS Number, Molecular Formula or Weight.  
Use \* for partial names (e.g. ben\*).  
Search here for free. For professional searching, use ChemINDEX.

18662-53-8 Search

**ChemFinder.com**  
Public site offers CAS, Name, and Structure-searching capabilities

**Nitrioltriacetic acid trisodium salt monohydrate** [18662-53-8]  
**Synonyms:** hamshire nta na3; nitrioltriacetic acid trisodium salt monohydrate; N,N-bis(carboxymethyl)glycine trisodium salt monohydrate; perma klear nt; nitrioltriacetate monohydrate; nta sodium hydrate; trilon a92; Trisodium ni monohydrate; trisodium salt of nitrioltriacetic acid, monohydrate;

More information about the chemical is available at:  
**Chemical Online Order (1)**  
Available Chemicals Exchange  
Information about this particular compound

**Tools**  
BUY AT CHEMACK.COM  
VIEW CHEMDRAW STRUCT  
VIEW CHEM3D MODEL  
CAS RN Lookup  
THE MERCK INDEX  
NCI DATABASE

**Op**  
ADD  
ADD/CM

**Health (3)**  
(B)(e) TRIAGE Chemical Studies Database  
Berkeley Carcinogenic Potency Database  
National Toxicology Program (NTP) publications  
Information about this particular compound  
NTP Chemical Health and Safety Data  
Information about this particular compound  
UMCP Partial list of teratogens  
DSSTox: CPDBRM

**The Carcinogenic Potency Project**  
Lois Swirsky Gold, Ph.D., Director

**U.S. Environmental Protection Agency**  
**Distributed Structure-Searchable Toxicity (DSSTox)**  
Contact Us | First Version Search  
EPA Home | DSSTox Home | SDF Download Page | CPDBRM, CPDBHA, CPDBDO, CPDBPR: Carcinogenic Potency Database Summary Tables for Rats and Mice, Hamster, Dogs, and Non-Human Primates  
**DSSTox Source SDF Download Page**  
CPDBRM, CPDBHA, CPDBDO, CPDBPR: Carcinogenic Potency Database Summary Tables for Rats and Mice, Hamster, Dogs, and Non-Human Primates  
Description: The CPDB Summary Tables list summarized results for experiments on 1370 substances in the Carcinogenic Potency Database (CPDB). These Summary Tables represent one of many possible summarizations of the data in the CPDB. The CPDB, which continues to be expanded, includes detailed results and analyses of more than 5000 chronic, long term carcinogenesis bioassays reported in over 1200 papers in the general literature and more than 400 Technical Reports of the National Cancer Institute/National Toxicology Program. Details

**Formula** C<sub>6</sub>H<sub>8</sub>NNa<sub>3</sub>O<sub>7</sub> **Molecular Weight** 275.10101  
**CAS RN** 18662-53-8 **Melting Point (°C)** ~410  
**ACX Number** X1005204-9 **Boiling Point (°C)**  
**Density** **Vapor Density**  
**Refractive Index** **Vapor Pressure**  
**Evaporation Rate** **Water Solubility** >=10 g/100 mL at  
**Flash Point (°C)** **EPA Code**  
**DOT Number** **RTECS** AJ1070000  
**Comments** White crystalline powder

**Web-based Structure-Searching of DSSTox SDF Files:**

NCI Structure Browser  
NLM ToxNet  
ACD Web-Librarian  
SRC/EPA PBT Profiler

**Structure-search displays chemical record in DSSTox: NCTREX**

**User seeks to learn more about NCTREX and its data fields**

**DSSTox Central Field Definition Table**  
(Last updated 11 November 03)  
Indexed DSSTox SDF Files Included in Table:  
This page provides corresponding DSSTox Source SDF Download Page for each database listed by NCTREX.

**NCTREX**  
NCTREX: Carcinogenic Potency Database Summary Tables (CPDBRM, CPDBHA, CPDBDO, CPDBPR)  
NCTREX: EPA Water Disinfection By-Product with Carcinogenicity Data  
NCTREX: EPA Radon Mitigation Acute Toxicity Database  
NCTREX: NCTR Estrogen Receptor Binding Database

The table below contains an alphabetical listing of central field names contained in all DSSTox SDF files currently online on this website. DSSTox Standard Chemical Fields are included in this listing but are separately designated. For each field listed in this table, the DSSTox SDF files in which the field is contained is listed under the column DSSTox SDF Download Page. The DSSTox SDF Download Page containing full reference documentation for that database. This copy of this table provides abbreviated content for each separate NCTREX field file (NCTREX-CPDBRM, EP-ATB, etc.) reference document offered for download on each DSSTox Download Page for that field.

It is noted that indicated to be a DSSTox Standard Chemical Field (yellow highlighted), a link is provided to the file on DSSTox SDF files. For more complete information, a user is encouraged to be the link on DSSTox SDF files. For more complete information, a user is encouraged to be the link on DSSTox SDF files.

In the table below, Abbreviated Values for allowable field entries occurring in DSSTox SDF files, separated by slashes (/) for categorical fields and underlined and contain a space for non-categorical fields (e.g. chemical symbols). These are explained in the Description section. Abbreviated values are: 1-10. The journal symbol (J) indicates the value is not a number; the possible additional values are listed in the Description section. The possible additional values are separated by a single slash (/) in the description, and not a space to list them.

References cited in the Description column of the table reference sections to the table (indicated by NCTREX).

**DSSTox Source SDF Download**  
NCTREX: National Center for Toxicological Research Estrogen Receptor Binding Database  
Description: Legislation passed in 1998 mandated that the EPA develop and implement a screening strategy for assessing the risk associated with endocrine disrupting chemicals (EDCs). Recognizing the importance of the Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC) focused on development of priority-setting approaches and Tier 1 screening methods, initially for assessing endocrine activity, that would guide the more limited application of Tier 2 animal testing. Priority setting primarily refers to quantitative structure-activity relationship (QSAR) methods for assessing the potential endocrine activity of chemicals for which test data are unavailable. Included on this list are 14 endocrine molecules in the NCTREX database.

Field Name	File	DSSTox SDF
Ability to Bind to Estrogen Receptor	CPDBRM	CPDBRM
Ability to Bind to Estrogen Receptor	CPDBHA	CPDBHA
Ability to Bind to Estrogen Receptor	CPDBDO	CPDBDO
Ability to Bind to Estrogen Receptor	CPDBPR	CPDBPR
Ability to Bind to Estrogen Receptor	CPDBRM	CPDBRM
Ability to Bind to Estrogen Receptor	CPDBHA	CPDBHA
Ability to Bind to Estrogen Receptor	CPDBDO	CPDBDO
Ability to Bind to Estrogen Receptor	CPDBPR	CPDBPR
Ability to Bind to Estrogen Receptor	CPDBRM	CPDBRM
Ability to Bind to Estrogen Receptor	CPDBHA	CPDBHA
Ability to Bind to Estrogen Receptor	CPDBDO	CPDBDO
Ability to Bind to Estrogen Receptor	CPDBPR	CPDBPR



● Online Access to the EPA DSSTox databases


The EPA has recently announced their Distributed Structure-Searchable Toxicity (DSSTox) Database Network which provides a community forum for publishing standard format, structure-annotated chemical toxicity data files for open public access. ACD/Labs has enhanced these files by using our [PhysChem predictors](#) and [systematic naming software](#) to add these parameters to the databases. These databases have also been made available online using [ACD/Labs Web Librarian](#) technology. For online access use the links listed at [Samples & databases page for Web Librarian](#).



File Name: **epac**  
Download Datab:

## Web Librarian DBs

We have made the numerous DSSTox databases available in a number of flexible formats. Below are listed four links for each database.

- The **DSSTox description link** will take you directly to the DSSTox webpage from which the source files are obtained and where any modifications to the databases will be made available.
- The **SDF file link** will enable you to download these files directly from the DSSTox website.
- The **CFD file** includes the source files in [ChemFolder](#) format and a ChemFolder form to use with the database for more visual appeal.
- The **online database links** provide direct online access to the databases via our [ACDWeb Librarian](#) interface. This interface allows structure, substructure and alphanumeric text based searching of the databases. For details of how to use the system please review the online help file indicated by the  at the top right hand side of the page.



Explore

Database

Search Records

ACD/Web Librarian™



Search Conditions



### Main Search Form

Multiple Database Search:

Current Database Search:

Directory Search:

Note:

Formula (example: C10 Cl(1-10) F(0) N):

Exact:

Formula Weight (example: 120.3 or 20..50):

Structure/Reaction Search:

Show Structure/Reaction:

- Flavours
- Food Additives
- Mole Chemical Corporation
- Molecule of the Month
- Standard Agent Database by National Cancer Institute
- Combi NMR Databases
- EPA
  - EPA Distributed Structure-Searchable Toxicity (DSSTox) Database
    - Carcinogenic Potency Database (CPDB)
    - EPA Fathead Minnow Acute Toxicity (EPAFHM)
    - National Center for Toxicological Research (NCTR) Estrogen
    - Water Disinfection By-Products with Carcinogenicity Estimates
  - Private Chromatography Applications Database
  - Public Chromatography Applications Database

Name	Type	Size	Modified
Summary ...	ACD/CHEMF...	305 KB	Tue Mar 23 ...
Summary ...	ACD/CHEMF...	537 KB	Tue Mar 23 ...
Summary ...	ACD/CHEMF...	420 KB	Tue Mar 23 ...
Summary ...	ACD/CHEMF...	2376 KB	Tue Mar 23 ...
Summary ...	ACD/CHEMF...	2202 KB	Tue Mar 23 ...



## Analog Search Component to Toxicity Estimation

### Using the PBT Profiler

[Information needed](#)

[Examples](#)

[Interpreting Results](#)

[What's new?](#)

### Related Links

[About PBTs](#)

[PBT Strategy](#)

[TRI PBT Project](#)

[P2 Framework](#)

[Links & Contacts](#)



Comments

Persistent, Bioaccumulative, and Toxic Profiles Estimated for Organic Chemicals On-Line

**PBT Profiler**  
A Component of OPPT's  
P2 Framework  
*Assessing Chemicals in  
the Absence of Data*

About

Methodology

Criteria

Anonymity & Security

Definitions

Terms of Use

Chemicals That

Can't be Profiled

The PBT Profiler was developed as a voluntary screening tool to identify Pollution Prevention opportunities for chemicals without experimental data.

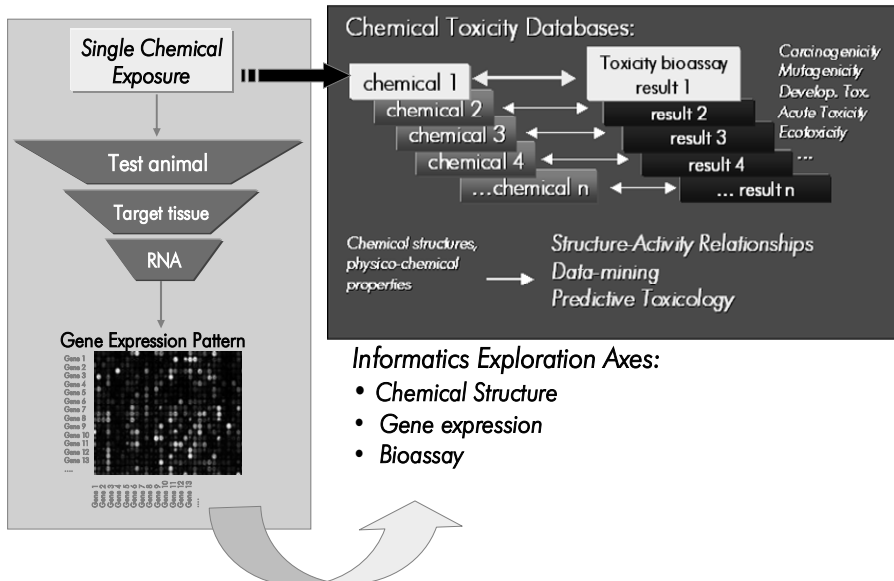
Users of the PBT Profiler acknowledge that they have read and accept the [Terms of Use](#)

**Start the PBT Profiler**

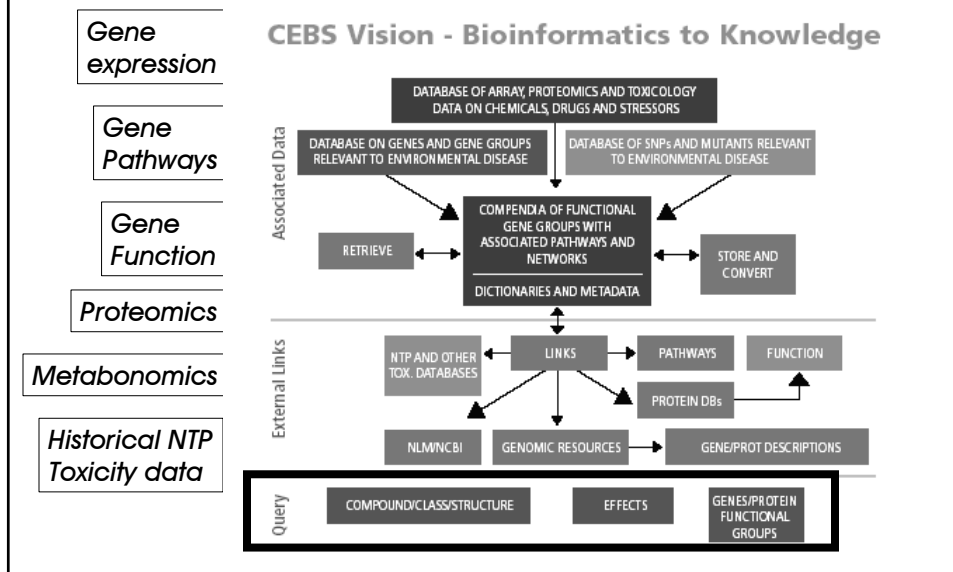
*Developed by the [Environmental Science Center](#) under contract to the [Office of Pollution Prevention and Toxics, U.S. Environmental Protection Agency](#)*

*Computer Resources Donated by [Syracuse Research Corporation](#) Ver 1.203 Last Updated May 20, 2003*

## Bioinformatics ..... meets Chemoinformatics



# Chemical Effects in Biological System Knowledge-Base (CEBS)

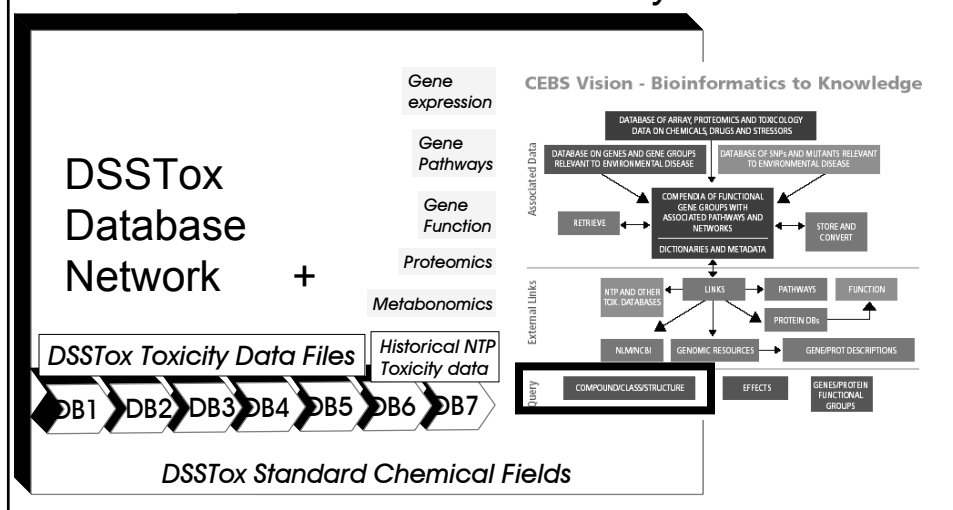


## DSSTox / CEBS Collaboration:

Part 1 – DSSTox Annotation of CEBS

Part 2 – Link CEBS to DSSTox Database Network

Part 3 – Add structural searchability to CEBS



## Chemical Relational Searching Capability:

Database status: 250251 open structures ready for searching.  
Mail: [Wolfe.D.Boyd@nih.gov](mailto:Wolfe.D.Boyd@nih.gov) for bug reports, comments and questions (and CC to Marc.C.Nicklaus if you like).

Start Search: [Reset]

Query Type: [Negate] Query Data Value

NSC Number(s) [Editor] [(No options)]

CAS Number(s) [Editor] [(No options)]

Formula [Editor] [Other elements allowed]

Molecular Weight Range [Editor] [(No options)]

Exact Structure [All molecules] [Browse]

Connect query fields by: AND OR XOR

Max. number of hits and search time: 100 hits, 90 seconds [Background job named]

Output Format: HTML Table with Samples preferably 3D

Output Sort: NSC Number

Page loads: 05380 Queries: 01/01

NCI CACTVS Structure Browser:  
<http://cactus.nci.nih.gov/>

Database status: 250251 open structures ready for searching.  
Mail: [Wolfe.D.Boyd@nih.gov](mailto:Wolfe.D.Boyd@nih.gov) for bug reports, comments and questions (and CC to Marc.C.Nicklaus if you like).

3D Structure Editor by P. Bitt (NOVARTIS)

3D Constraints for This Substructure Query

Distance (Å) [Value Range]

Atom [Atom]

Distance (Å) [Value Range]

Atom [Atom]

Distance (Å) [Value Range]

Atom [Atom]

Distance (Å) [Value Range]

Atom [Atom]

Global Query Parameters for ALL 3D structure queries

Highlight Matched Substructure on Display

No Match of Aromatic Bonds on Single Double Bonds

Testament about PSS search

Allow Overlay of Multiple Fragments

Enforce Ring System Equality

Use exact charges

Transfer to Query Form [Clear Form and Editor]

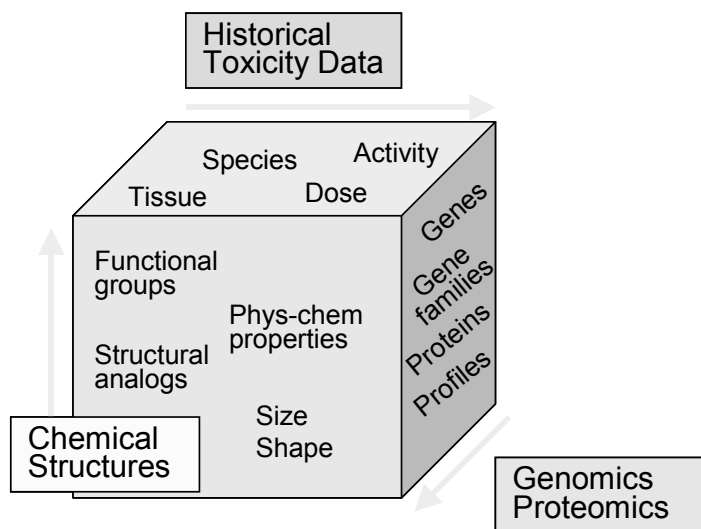
Create a Public Domain Browser for CEBS:

- NIH/NCI CACTVS Structure Browser
- JChem/Marvin
- LHASA VITIC

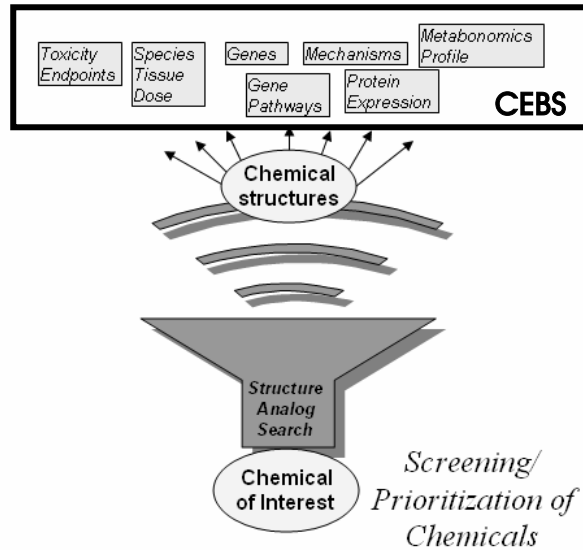
Chemistry Standards:

- INChI chemical codes
- IUPAC Names

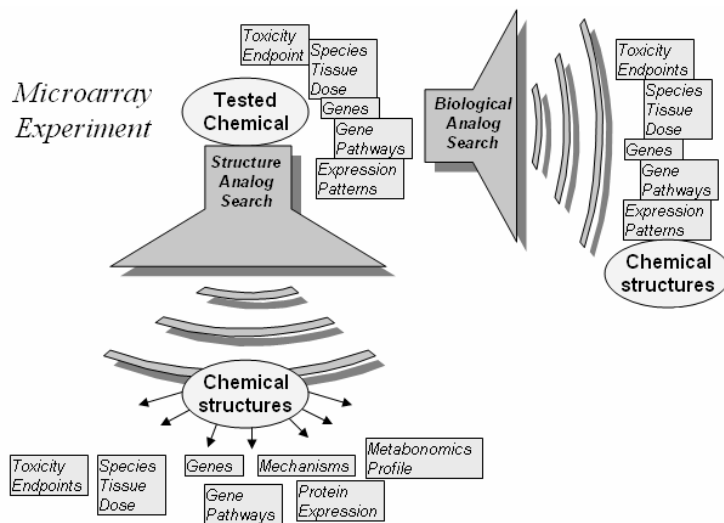
## CEBS Chemo-bioinformatics: Expanded Relational Search Domains



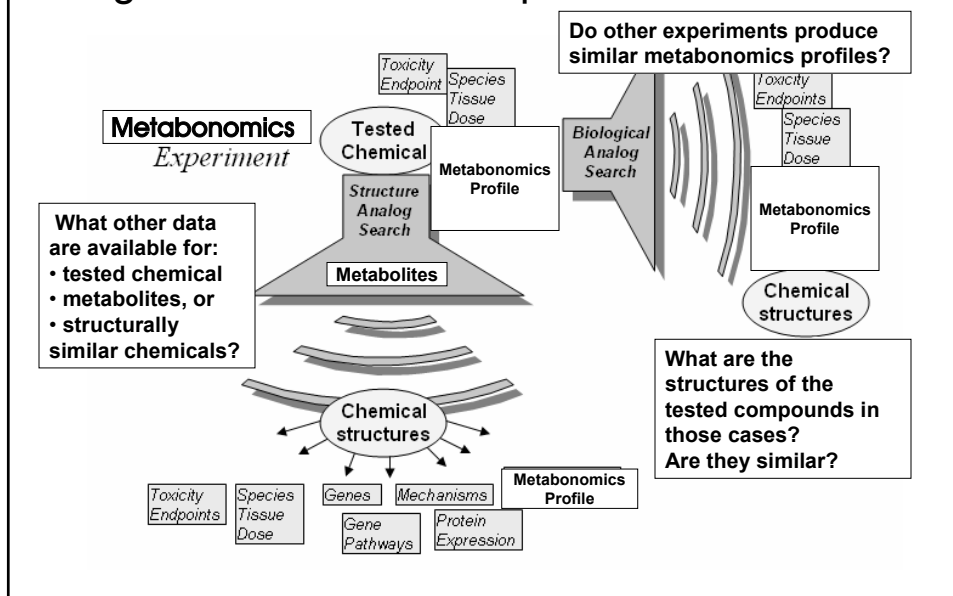
# Gathering Data on Chemical Analogues



# Exploring Broader Associations from Single Microarray Experiment



## Exploring Broader Associations from Single Metabonomics Experiment



## DSSTox Status & Future Directions:

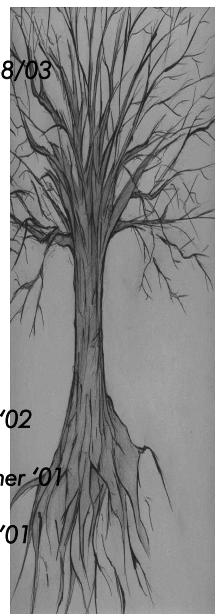
- ✓ Outside Coordination:
  - *Tox standards efforts (ToxML, MIAMI-Tox)*
  - *Chemistry standards efforts (InCHI, IUPAC names)*
  - *Other public efforts (NCI, TOXNET, PBT, etc)*
- ✓ LHASA UK Tox DB – DSSTox public data distributor
- ✓ FDA, EPA, EU – assist in publishing non-proprietary toxicity data from internal databases
- ✓ EPA CompTox Augmented Funding Award – 1 yr:  
*Expand DSSTox public database offerings*
- ✓ EPA CompTox New Start Proposal – CEBS collaboration: *Undergoing review*

## DSSTox Collaborators/Advisors/Acknowledgements:

• Cancer Potency Data Base - rodent carcinogenicity.....	Lois Swirsky Gold, Thomas Stone
• EPA - Ecotoxicity (fathead minnow, Teratox).....	Chris Russom
• EPA/OPPT,OW – DBP cancer assessment.....	Yin-ak Woo, Mary Manibusan
• FDA/NCTR - Estrogen receptor binding data base.....	Weida Tang, Hong Fang
• FDA/CDER – Maximum recommended dose drugs .....	Dan Benz, Ed Matthews, Joe Contrera
• EPA/OPP, Ecotox – pesticides .....	Brian Montague, Pauline Wagner
• NTP Gene-tox data; IRIS .....	Errol Zeiger, Zeiger Consulting
• Nat Ctr Toxicogenomics: CEBS .....	Mike Waters, Ray Tennant
• GlaxoSK - GeneTox/NTP Salmonella database.....	Neal Carriello, Vijay Gombar
• NIEHS/NTP Rodent carcinogenicity, etc .....	Skip Eastin, Doug Bristol
• Developmental toxicity .....	Vijay Gombar (GlaxoSK), Orest Macina
• ICVAM Toxicity databases .....	Ray Tice, Marc Jackson, ILS
• Unilever Skin Sensitization Database.....	Martin Barrett, Marlin Consulting
• ZEBET Acute Toxicity Database .....	Julie Penzotti, Rational Discovery
• Tulane/Xavier Univ – Endocrine Disruption .....	Tom Wiese
• NCI – SDF tools, CACTVS structure browser.....	Marc Nicklaus
• LeadScope – SDF/XML converter, FDA carcinogens .....	Chihae Yang
• SDF Viewer application .....	Thomas Harrocks, Intuitive Software Solutions
• ACD – ChemFolder, WebBase .....	Antony Williams, M Hachey, G Shear
• CambridgeSoft – ChemFinder application .....	Rich Talbot
• EPA scientific advisors .....	Stephen Nesnow, Adam Swank
• EPA/CSC – web development, IT .....	Brian Garges/ D Kanipe, D Marshall

## DSSTox Development Team:

- ClarLynda Williams – Coordinator, Lead Project Technician  
– NCCU/EPA COOP Student Trainee 12/00-12/02; EPA 12/02-8/03
- Jamie Burch  
– NCCU/EPA COOP Student Trainee 11/02-2/04
- Brian Rogers  
– NCCU/EPA COOP Student Trainee 1/04-present
- Marty Wolf  
– EPA Senior Environmental Employment Program 8/04-present
- Audrey Evans  
– ECO/EPA Summer Student Trainee, Summer '03
- Todd Stewart  
– EPA-UNC Student COOP, Spring/Summer 02
- Nina Fields  
– Shaw Univ. High School Minority Mentoring Program, Summer '02
- James Beidler  
– EPA Summer Student Employee, Warren-Wilson College, Summer '01
- Daniel Ohuoba  
– Shaw Univ. High School Minority Mentoring Program, Summer '01
- Adam Swank  
– ECD, EPA



Optional DSSTox Website Tour

<http://www.epa.gov/nheerl/dsstox/>