United States Environmental Protection Agency

GEMS Presentation, Durham, NC, 29 October 2007

Toxico-Cheminformatics: A New Frontier for Predictive Toxicology

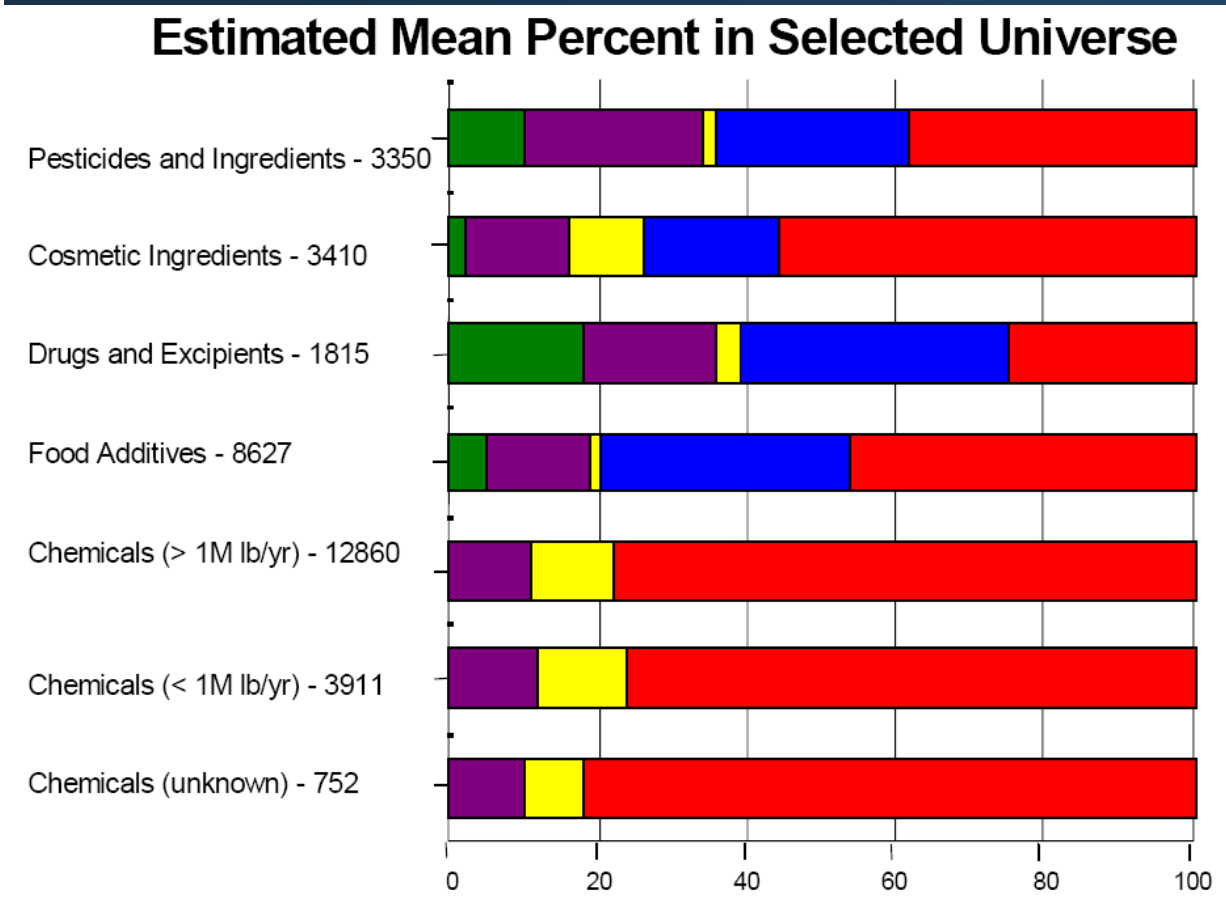
Ann Richard richard.ann@epa.gov

Office of Research and Development
National Center for Computational Toxicology

Part I.

The Problem

Environmental Chemicals: Toxicity Assessment Data Gaps



Complete Hazard Assessment Possible

Minimal Information Available

No Toxicity Information Available

Partial Hazard Assessment Possible

Some Toxicity Information Available

Strategies for Closing the Chemical Data Gap

by John S. Applegate and Katherine Baer

A Center for Progressive Reform Publication *April 2006*

National Academy of Sciences Report (2007)
Toxicity Testing in the Twenty-first Century: A Vision and a Strategy

NAS PANEL SEEKS MAJOR SHIFT IN HOW EPA ASSESSES CHEMICALS' TOXICITY

Date: June 22, 2007 -

A National Academy of Sciences (NAS) panel is calling for a major shift in how EPA assesses chemicals' toxicity, recommending that the agency base its toxicological research and regulatory processes on how substances affect biological pathways -- which send information within and between cells -- rather than so-called health endpoints, such as cancer.

The new studies envisioned by the panel would evaluate chemicals' effects on biological processes using cells or cell lines, preferably human, to examine how they react to exposure to different substances. Rather than focusing research and basing regulations on endpoints, such as a substance's apparent ability to create tumor cells or harm brain development in fetuses, EPA should center toxicity testing around "the perturbations in toxicity pathways that are expected to lead to adverse effects," the report says.

"In this framework, the goals of toxicity testing are to identify critical pathways that when perturbed can lead to adverse health outcomes and to . . . understand the effects of perturbations on human populations," says the report, *Toxicity Testing in the Twenty-first Century: A Vision and a Strategy* .

Chemical Research in Toxicology

JULY 2007
VOLUME 20, NUMBER 7

Editorial

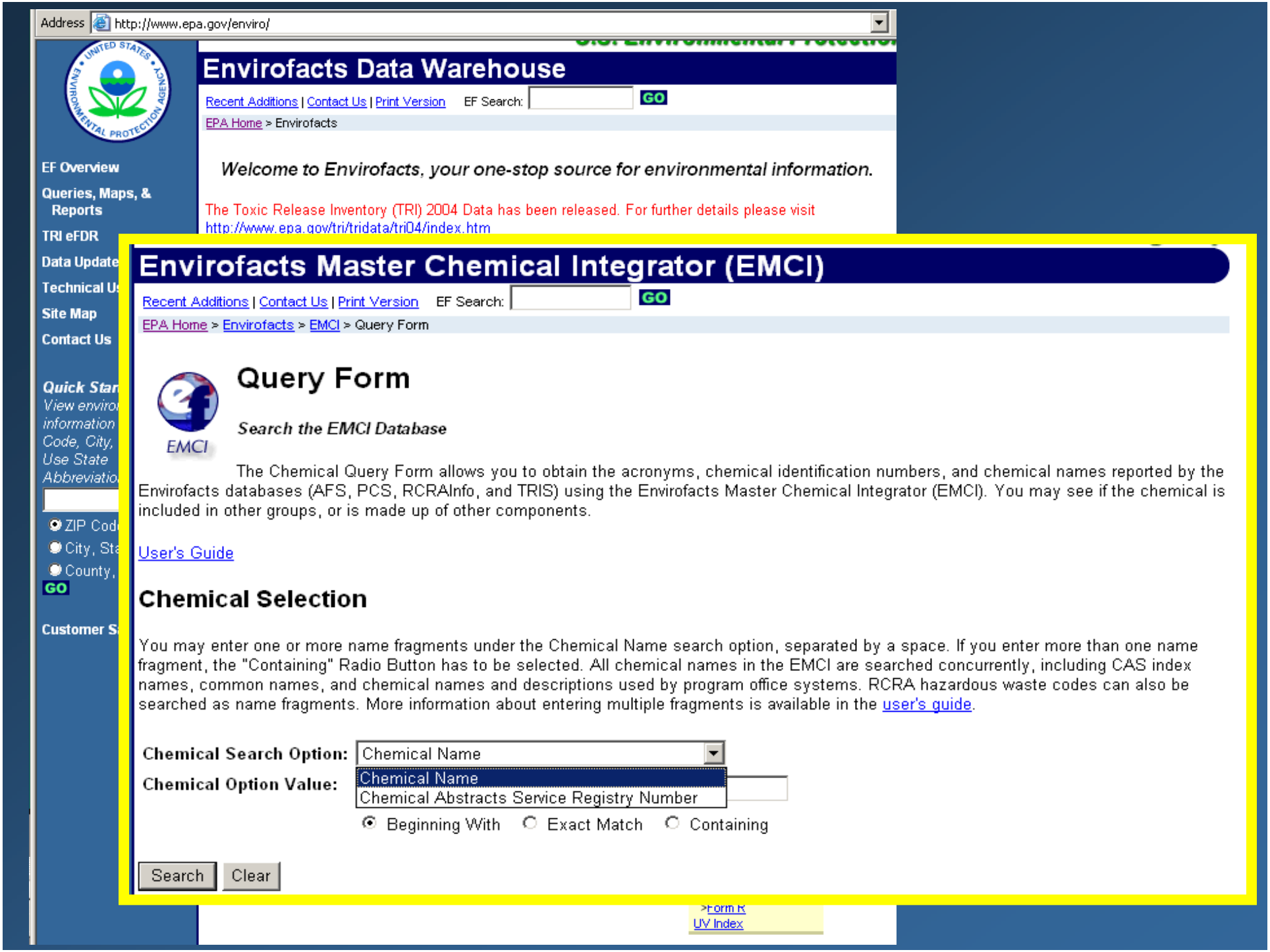
The Future of Toxicology—Wrap Up

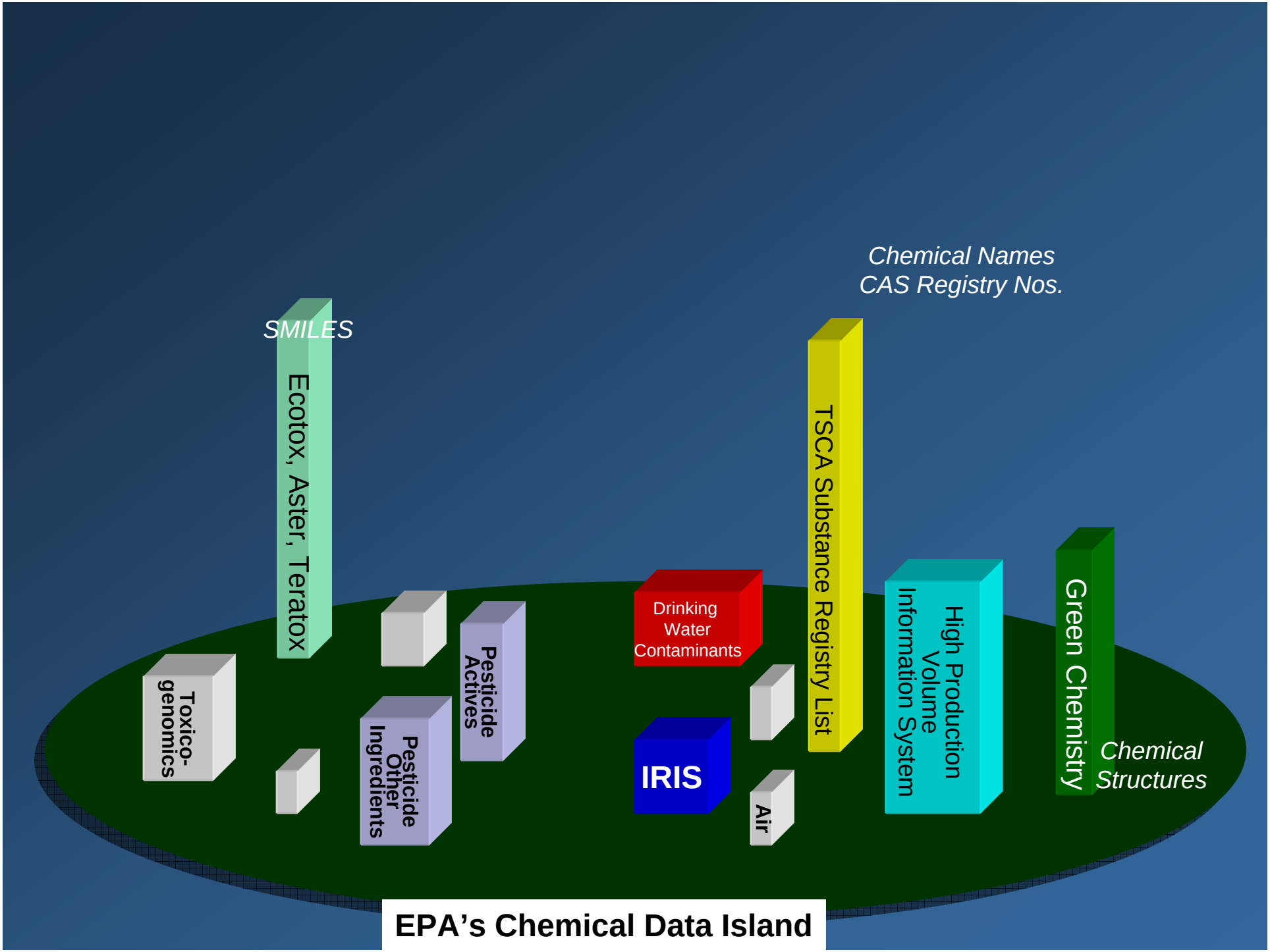
Aaron T. Jacobs and Lawrence J. Marnett*

A.B. Hancock Jr. Memorial Laboratory for Cancer Research
Departments of Biochemistry, Chemistry, and Pharmacology
Vanderbilt Institute of Chemical Biology
Center in Molecular Toxicology
Vanderbilt-Ingram Cancer Center
Vanderbilt University School of Medicine
Nashville, Tennessee 37232

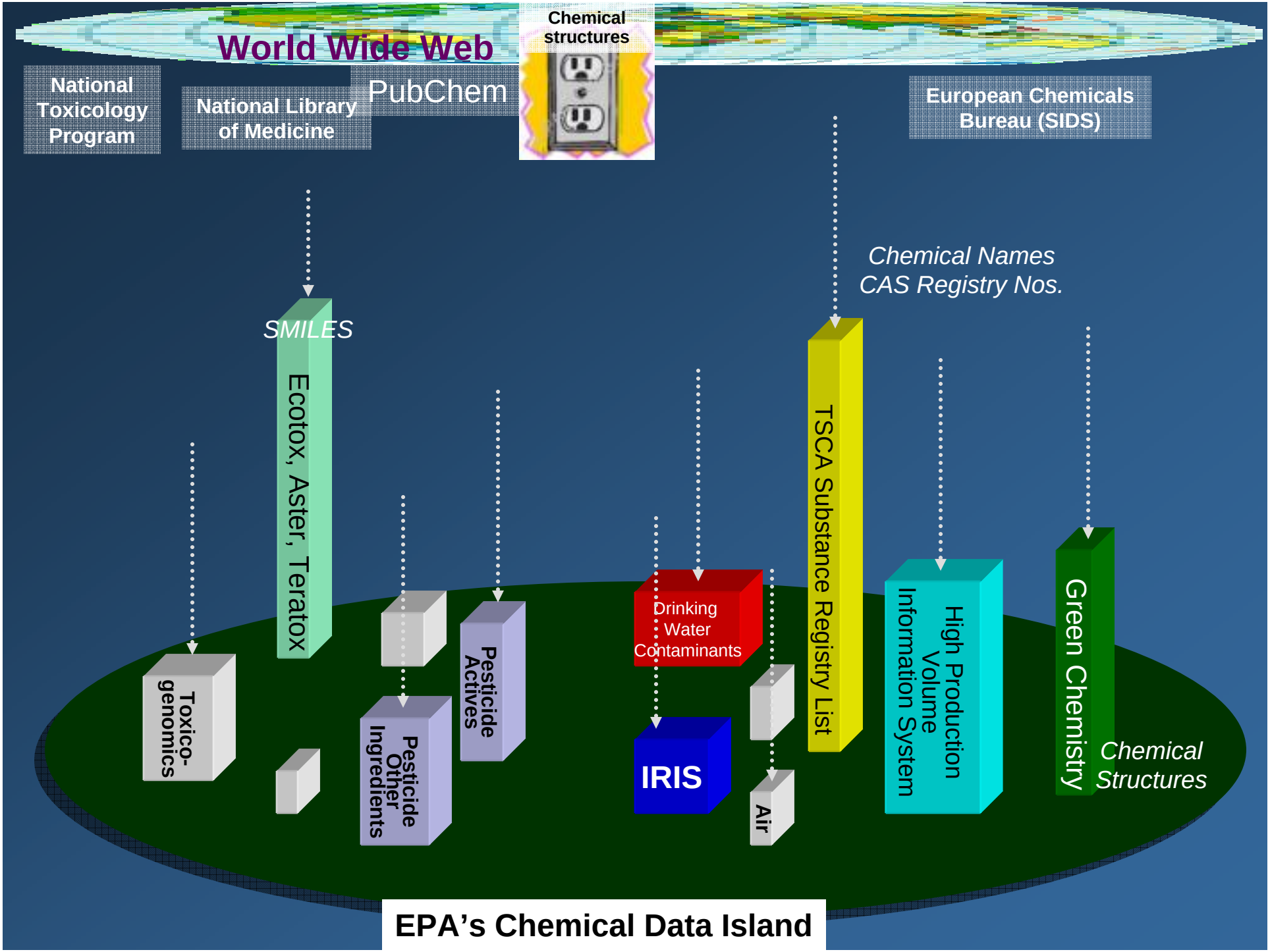
Wither toxicology? We have enjoyed a series of informative, occasionally provocative, commentaries on this subject from a paucity of compelling problems to work on that would generate a broad mandate for large-scale investment in toxicol-

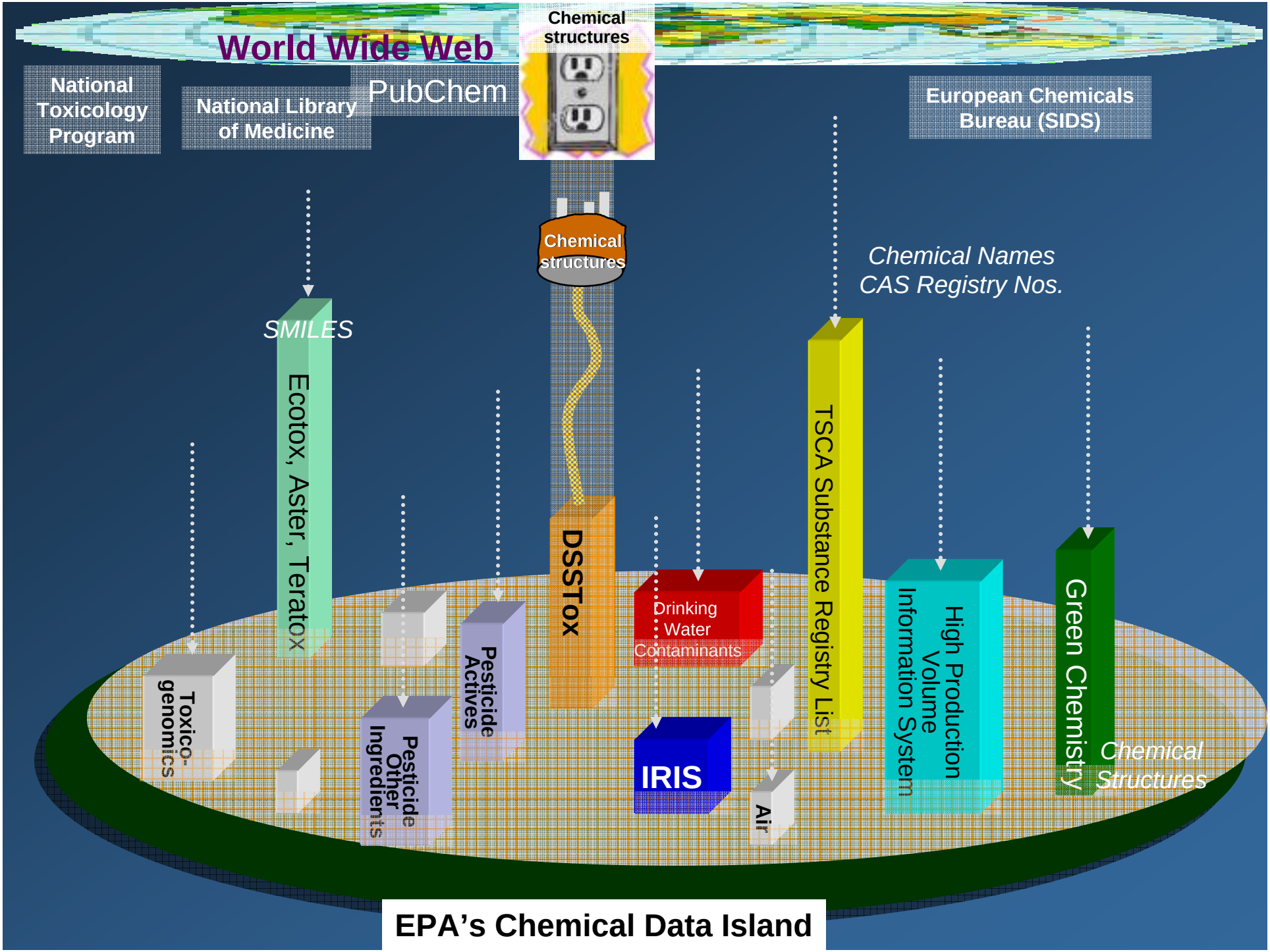
"A major focus for the future of computational toxicology will be integration and analysis of large data sets. The current state of toxicity databases is something of a mess. There are a number of databases, each with differing content, architecture, and searchability, that makes the task of integration extremely difficult."


Address http://www.epa.gov/enviro/
Envirofacts Data Warehouse
Recent Additions | Contact Us | Print Version EF Search: GO
EPA Home > Envirofacts
Welcome to Envirofacts, your one-stop source for environmental information.
The Toxic Release Inventory (TRI) 2004 Data has been released. For further details please visit
http://www.epa.gov/tri/tridata/tri04/index.htm
EF Overview
Queries, Maps, & Reports
TRI eFDR
Data Update
Technical U
Site Map
Contact Us
Quick Star
View enviro
information
Code, City,
Use State
Abbreviatio
ZIP Cod
City, Sta
County,
GO
Customer S
Envirofacts Master Chemical Integrator (EMCI)
Recent Additions | Contact Us | Print Version EF Search: GO
EPA Home > Envirofacts > EMCI > Query Form
EMCI
Query Form
Search the EMCI Database
The Chemical Query Form allows you to obtain the acronyms, chemical identification numbers, and chemical names reported by the Envirofacts databases (AFS, PCS, RCRAInfo, and TRIS) using the Envirofacts Master Chemical Integrator (EMCI). You may see if the chemical is included in other groups, or is made up of other components.
User's Guide
Chemical Selection
You may enter one or more name fragments under the Chemical Name search option, separated by a space. If you enter more than one name fragment, the "Containing" Radio Button has to be selected. All chemical names in the EMCI are searched concurrently, including CAS index names, common names, and chemical names and descriptions used by program office systems. RCRA hazardous waste codes can also be searched as name fragments. More information about entering multiple fragments is available in the user's guide.
Chemical Search Option: Chemical Name
Chemical Option Value:
Chemical Name
Chemical Abstracts Service Registry Number
Beginning With Exact Match Containing
Search Clear
>Form R
UV Index



EPA's Chemical Data Island

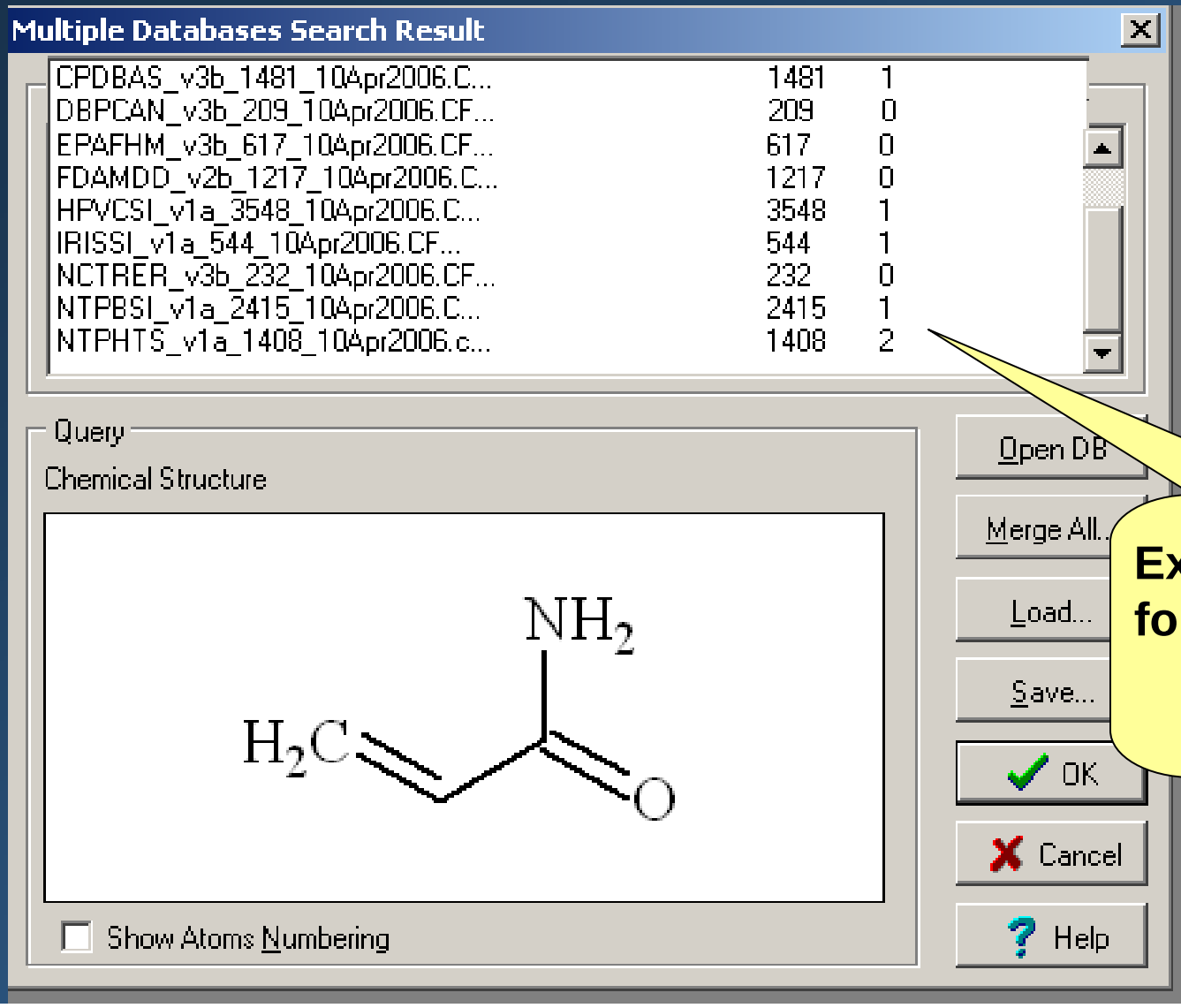

World Wide Web
Chemical structures
National Toxicology Program
National Library of Medicine
PubChem
European Chemicals Bureau (SIDS)
Chemical Names
CAS Registry Nos.
SMILES
Ecotox, Aster, Teratox
TSCA Substance Registry List
High Production Volume Information System
Green Chemistry
Chemical Structures
Toxico-genomics
Pesticide Actives
Pesticide Other Ingredients
Drinking Water Contaminants
IRIS
Air
EPA's Chemical Data Island


World Wide Web
Chemical structures
National Toxicology Program
National Library of Medicine
PubChem
European Chemicals Bureau (SIDS)
Chemical structures
Chemical Names
CAS Registry Nos.
SMILES
Ecotox, Aster, Teratox
DSSTox
TSCA Substance Registry List
Drinking Water Contaminants
High Production Volume Information System
Green Chemistry
Chemical Structures
Toxico-genomics
Pesticide Actives
Pesticide Other Ingredients
IRIS
Air
EPA's Chemical Data Island

Part II.

Data & Data Linkages

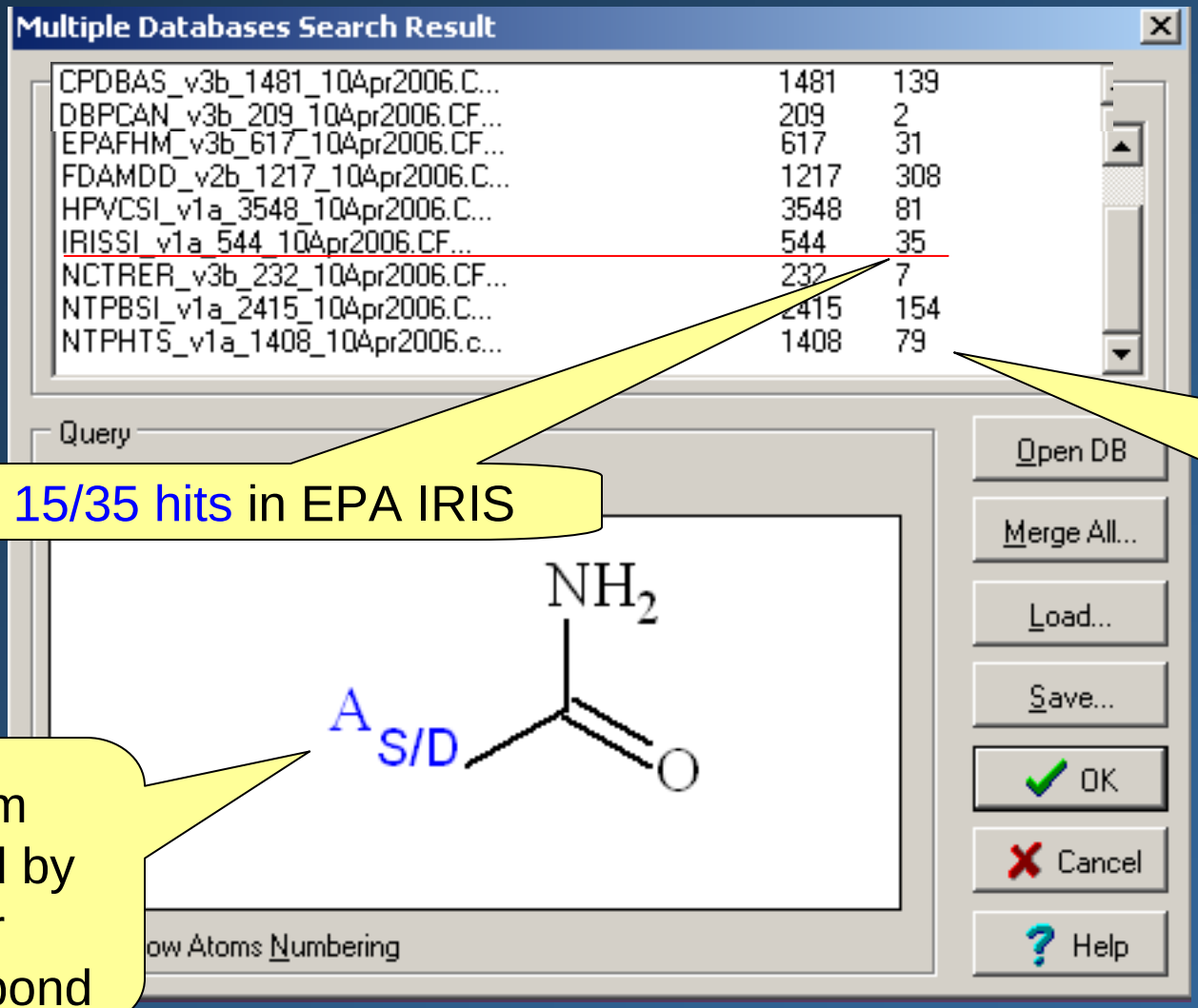
Structure Searching Across 9 Diverse Toxicity Databases:



Exact match search for Acrylamide:

Result: 6 hits

Generalized Sub-Structure Searching Across 9 Diverse Toxicity Databases:



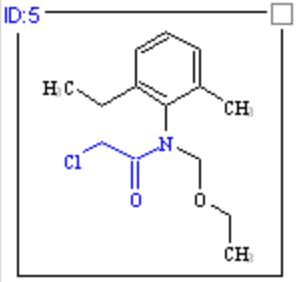
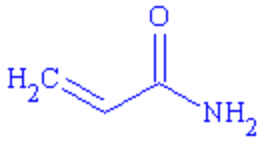
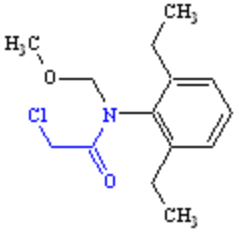
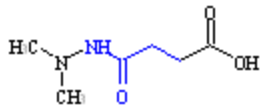
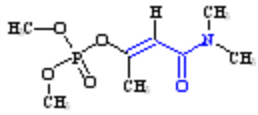
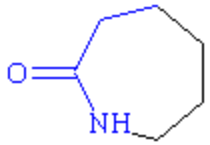
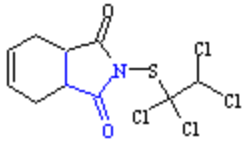
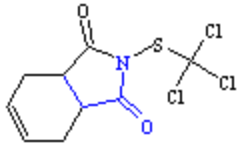
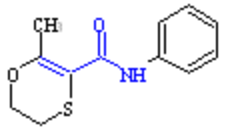
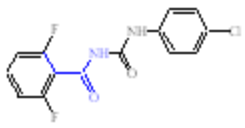
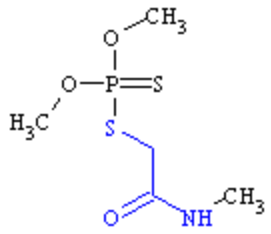
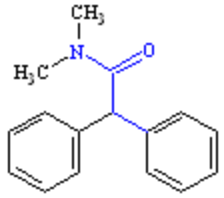
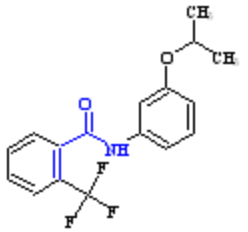
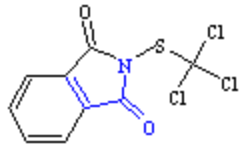
View 15/35 hits in EPA IRIS

Result: 836 hits

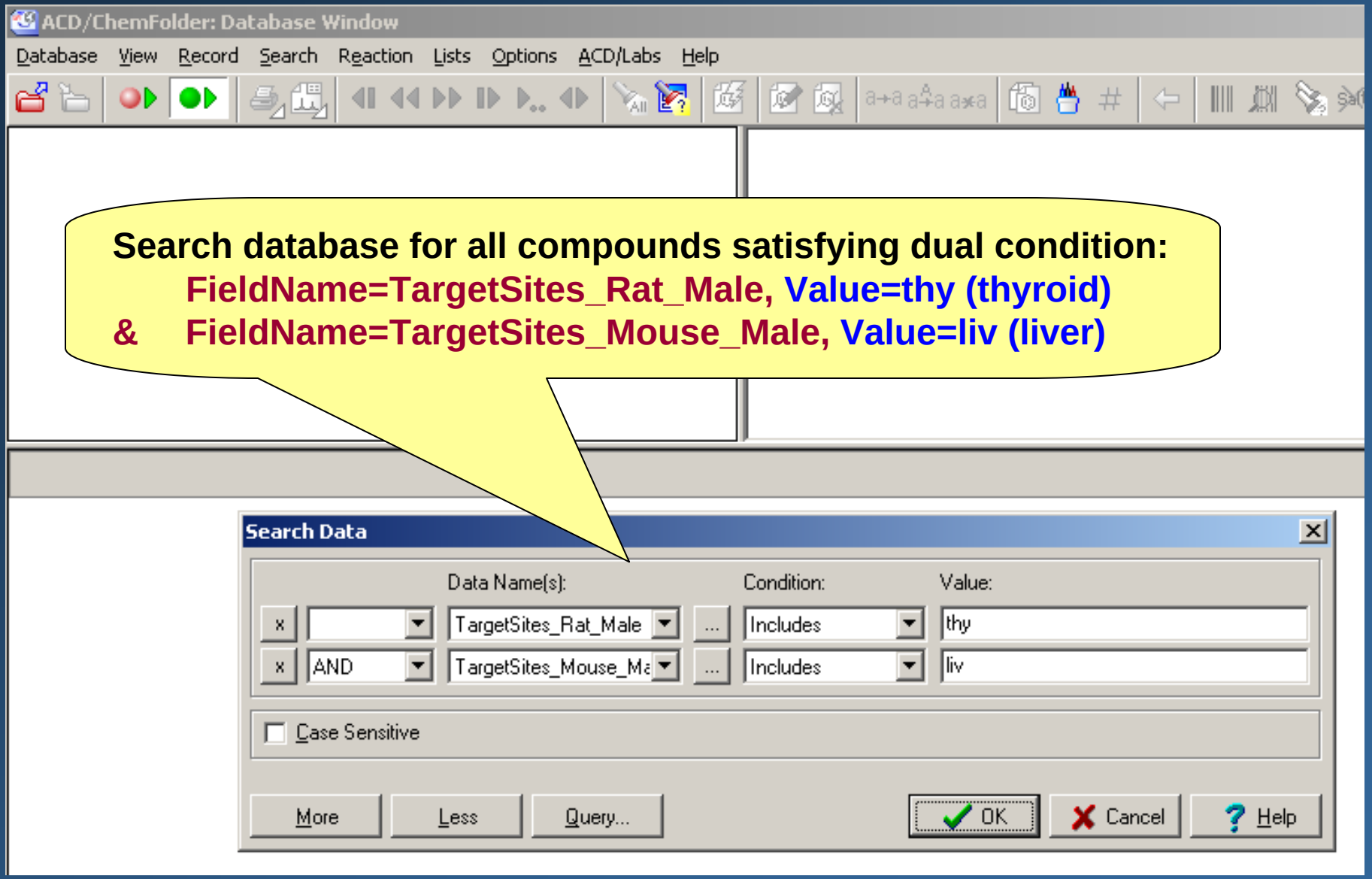
Any atom attached by single or double bond

Generalized Sub-Structure Searching Across 9 Diverse Toxicity Databases:

15/35 Hits in IRIS containing acrylamide-like moiety

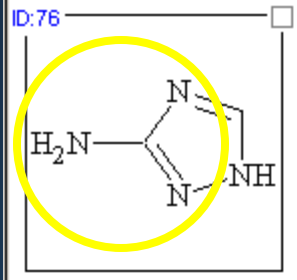
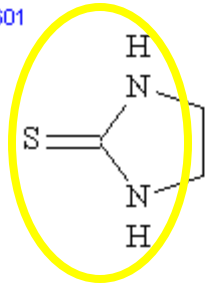
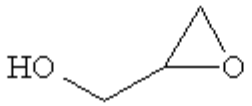
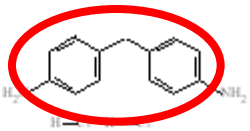
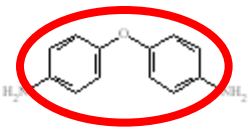
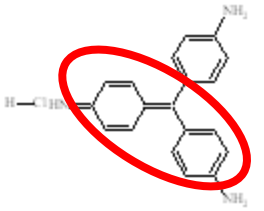
ID:5 	ID:12 	ID:16 	ID:17 	ID:58 
Acetochlor 34256-82-1	Acrylamide 79-06-1	Alachlor 15972-60-8	Daminozide 1596-84-5	Bentazon 25057-89-0
ID:71 	ID:99 	ID:100 	ID:101 	ID:108 
Bidrin 141-66-2	Caprolactam 105-60-2	Captafol 2425-06-1	Captan 133-06-2	Carboxin 5234-68-4
ID:206 	ID:210 	ID:229 	ID:266 	ID:268 
Diflubenzuron 35367-38-5	Dimethoate 60-51-5	Diphenamid 957-51-7	Flutolanil 66332-96-5	N-(Trichloromethylthio)phthali 133-07-3

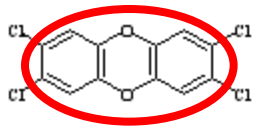
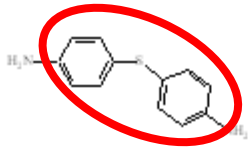
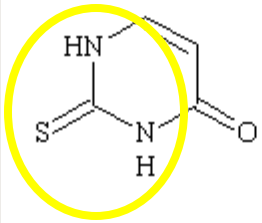
Relational Biological Content Searching: Carcinogenic Potency Database – All Species (CPDBAS_1481)



ACD/ChemFolder: Database Window - [D:\DSSTOX_MISC\ACD_CHEMFOLDERFILES_10APR2006\CPDBAS_V3B_1481_10APR2006.CFD]

Database View Record Search Reaction Lists Options ACD/Labs Help

ID:76 	ID:601 	ID:666 	ID:871 	ID:1094 	ID:1247 
76 CPDBAS_v3b_1481_10A	601 CPDBAS_v3b_1481_10	666 CPDBAS_v3b_1481_10	871 CPDBAS_v3b_1481_10	1094 CPDBAS_v3b_1481_1	1247 CPDBAS_v3b_1481_1
3-Aminotriazole	Ethylene thiourea (ETU)	Glycidol	4,4'-Methylenedianiline dihyd	4,4'-Oxydianiline	C.I. Basic red 9 monohydroc
negative	positive	positive	positive	positive	positive
thy	thy	ezy lgi mgl nrv per ski sto thy	liv thy	liv thy	ezy liv ski sub thy
pit thy	thy	cli hmo mgl nrv orc sto thy	thy	liv thy	ezy sub thy
liv	liv pit thy	hag liv lun ski sto	adr liv thy	hag liv	liv
liv	liv pit thy	hag mgl ski sub ute	hmo liv thy	hag liv thy	adr liv

ID:1315 	ID:1344 	ID:1347 
1315 CPDBAS_v3b_1481_1	1344 CPDBAS_v3b_1481_1	1347 CPDBAS_v3b_1481_1
2,3,7,8-Tetrachlorodibenzo-p	4,4'-Thiodianiline	Thiouracil
negative	positive	thy
orc thy	ezy lgi liv thy	thy
liv lun	ezy thy ute	liv
liv	liv thy	liv
liv thy	liv thy	

Search Results: 9 hits / 1481 total

TargetSites_Rat_Male = thyroid

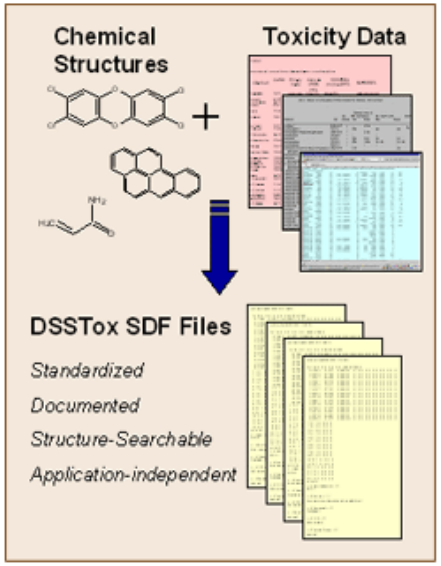
& TargetSites_Mouse_Male = liver

DSSTox

Distributed Structure-Searchable Toxicity (DSSTox) Database Network is a project of EPA's Computational Toxicology Program, helping to build a public data foundation for improved structure-activity and predictive toxicology capabilities. The DSSTox website provides a public forum for publishing downloadable, structure-searchable, standardized chemical structure files associated with toxicity data. More>

EPA DSSTox Structure-Browser v1.0 GO

DSSTox Structure-Browser information Page



- DSSTox Graphic Flowchart
- DSSTox Project Goals
- DSSTox Publications

Recent Additions: 27 September 2007

- TOXCST: Research Chemical Inventory for EPA's ToxCast Program - Updated to v2a

Recent Additions: 28 August 2007

***Launch of DSSTox Structure-Browser v1.0:

- A new structure-search capability for published DSSTox Data Files, allows users to search by DSSTox Standard Chemical Fields and includes options for:
  - **Text Search:** Chemical Name, CAS RN, InChI, Formula
  - **Structure Search (*Exact, Substructure, Similarity*):** SMILES or Structure Drawing Tool entry

***Revised Standard ID Fields for all DSSTox files:

- Modified Record, File, Chemical, and Substance ID fields to index all unique DSSTox structures and substances, also with respect to file record and version

DSSTox Data Files: Details>

CPDBAS_v4a_1481_15Jun2007 **New content
DBPCAN_v4a_209_15Jun2007
EPAFHM_v4a_617_15Jun2007
FDAMDD_v3a_1216_25Jul2007
HPVCSI_v2a_3548_15Aug2007**New content
IRISTR_v1a_544_28Jul2007 **New file
NCTRER_v4a_232_15Jun2007
NTPBSI_v2a_2293_24Aug2007 **Updated content
NTPHTS_v1a_1408_25Jul2007
TOXCST_v2a_320_25Sep2007 **Updated

More on Data File Types

DSSTox Standard Chemical Fields:

STRUCTURE

DSSTox_CID

STRUCTURE_Formula

STRUCTURE_MolecularWeight

STRUCTURE_ChemicalType

STRUCTURE_TestedForm_DefinedOrganic

STRUCTURE_ChemicalName_IUPAC

STRUCTURE_SMILES

STRUCTURE_Parent_SMILES

STRUCTURE_InChI

STRUCTURE_Shown

- tested chemical,
- general form of chemical,
- active ingredient of formulation,
- representative isomer in mixture,
- representative component in mixture,
- monomer of polymer

, simplified to parent

- defined organic
- inorganic
- organometallic

- parent,
- salt Na, Cl, etc
- complex HCl, H2O, mesylate, etc

DSSTox_Generic_SID

TestSubstance_ChemicalName

TestSubstance_CASRN

TestSubstance_Description

- single chemical compound
- macromolecule
- mixture or formulation
- unspecified or multiple forms

ChemicalNote

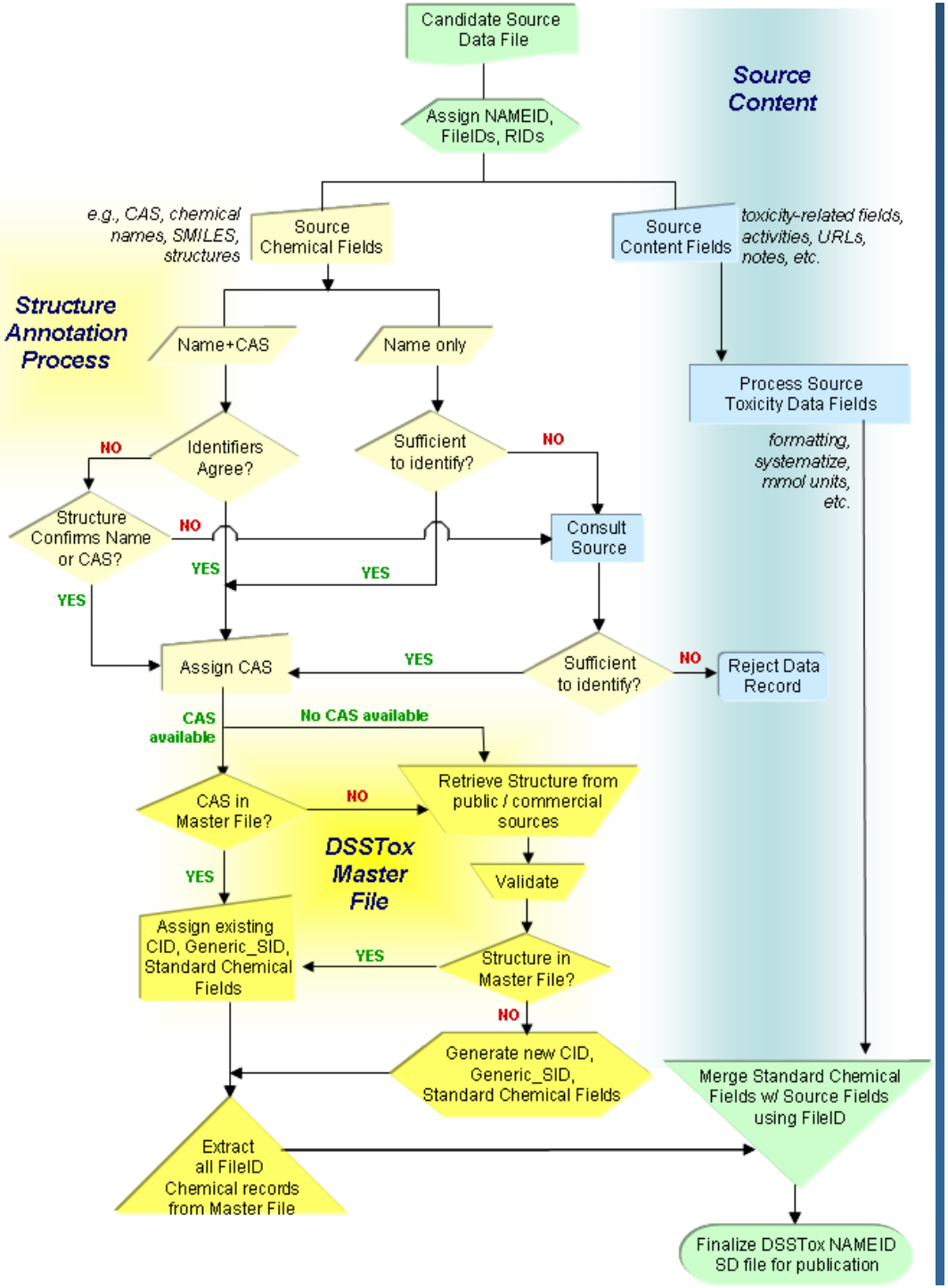
- description of mixture
- nature & CAS of components
- tautomers
- stereochemistry

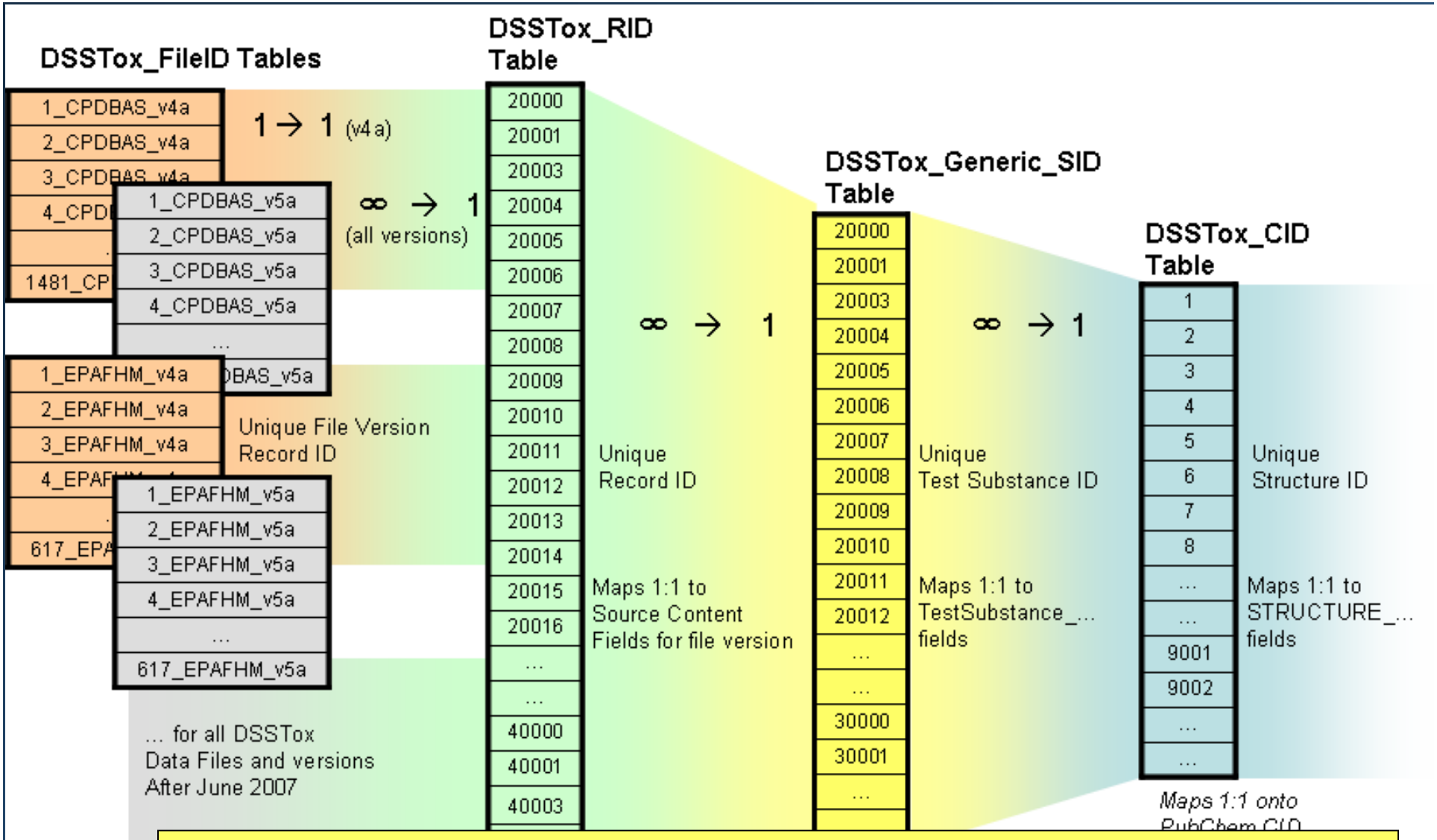
DSSTox_RID

DSSTox_FileID

DSSTox Chemical Quality Assurance Procedures:

- Chemical identification
- Structure annotation
- Internal consistency
- Quality review





Current DSSTox Master File

- > 35 data files and chemical inventories structure-indexed
- > 9300 unique substances with quality-reviewed chemical annotation
- > 7400 unique chemical structures indexed
- > 20,000 data records spanning all data files and inventories

NAMEID	version #records date	Expanded DSSTox Data File Title & Description	Status
CPDBAS	v4a 1481 15Jun2007	**Carcinogenic Potency Database Summary Tables - All Species:** Tumor target site incidence, TD50 potencies, summary activity calls for rat, mouse, hamster, dog, and/or non-human primate; data reviewed and compiled from literature and NTP studies.	*Updated*
DBPCAN	v4a 209 15Jun2007	**EPA Water Disinfection By-Products with Carcinogenicity Estimates Database:** 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.	*Updated*
EPAFHM	v4a 617 15Jun2007	**EPA Fathead Minnow Acute Toxicity Database:** Acute toxicities of 617 chemicals tested in common assay, with mode-of-action assessments and confirmatory measures.	*Updated*
FDAMDD Return to Top	v3a 1216 25Jul2007	**FDA Center for Drug Evaluation & Research - Maximum (Recommended) Daily Dose Database:** Maximum (recommended) daily dose (MRDD) values for 1217 pharmaceuticals in mg/kg-body weight (bw)/day, converted to mmol and normalized to dataset; MRDD values extracted from public literature sources.	*Updated*
HPVCSI	v2a 3548 15Aug2007	**EPA High Production Volume Challenge Program Structure-Index Locator File:** Compiled structures for three chemical lists provided on EPA HPV Challenge Program website; each record includes reference index to dated list.	*Updated*
IRISTR	v1a 544 15Aug2007	**EPA Integrated Risk Information System (IRIS) Toxicity Review Data File:** Compiled structures for EPA IRIS website with chemical-specific URLs linking to risk assessment summary data pages for 544 chemical substances.	*New*
NCTRER	v4a 232 15Jun2007	**FDA National Center for Toxicological Research (NCTR) - 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.	*Updated*
NTPBSI Return to Top	v2a 2293 24Aug2007	**National Toxicology Program (NTP) On-line Chemical Bioassay Database Structure-Index Locator File:** Compiled structures for the NTP On-line Database with chemical-specific URLs linking to NTP study summary pages; file includes fields for each of 4 main bioassay study areas with indicator values specifying presence or absence of study data for the chemical substance record.	*Updated*
NTPHTS	v2a 1408 25Jul2007	**National Toxicology Program (NTP) High-Throughput Screening Project Structure-Index File:** Compiled structures for set of 1408 NTP chemical substances provided to the NIH Chemical Genomics Center for HTS bioassay testing and to PubChem (PubChem_CIDs and PubChem_SIDs included in NTPHTS_v2a file); NCGC HTS bioassay data are being deposited into PubChem and can be retrieved with these PubChem chemical CID and SID record listings.	*Updated*
TOXCST	v1a 340 03Aug2007	**Research Chemical Inventory for EPA's ToxCast™ Program Structure-Index File:** Compiled structures for set of 340 chemical substances that are candidates for Phase I High-throughput screening (HTS) testing within the EPA ToxCast program. File will be updated with links to PubChem CIDs and SIDs for retrieving assay data, and with updates to chemical inventory as Program moves to Phase II and beyond.	*New*

CPDBAS_v5a_1547

STRUCTURE
DSSTox_RID
DSSTox_CID
DSSTox_Generic_SID
DSSTox_FileID
STRUCTURE_Formula
STRUCTURE_MolecularWeight
STRUCTURE_ChemicalType
STRUCTURE_TestedForm
_DefinedOrganic
STRUCTURE_Shown
TestSubstance_ChemicalName
TestSubstance_CASRN
TestSubstance_Description
ChemicalNote
STRUCTURE_ChemicalName
_IUPAC
STRUCTURE_SMILES
STRUCTURE_Parent_SMILES
STRUCTURE_InChI
StudyType
Endpoint
Species

Mutagenicity_SAL_CPDB
TD50_Rat_mg
TD50_Rat_mmol
TD50_Rat_Note
TargetSites_Rat_Male, Female, Both Sexes
TD50_Mouse_mg
TD50_Mouse_mmol
TD50_Mouse_Note
TargetSites_Mouse_Male, Female, Both Sexes
TD50_Hamster_mg
TD50_Hamster_mmol
TD50_Hamster_Note
TargetSites_Hamster_Male, Female, Both Sexes
TD50_Dog_mg
TargetSites_Dog
TD50_Rhesus_mg
TargetSites_Rhesus
TD50_Cynomolgus_mg
TargetSites_Cynomolgus
TD50_Dog_Rhesus_Cynomolgus_Note
ActivityCategory_SingleCellCall
ActivityCategory_MultiCellCall
ActivityCategory_MultiCellCall_Details
ToxicityNote
NTP_TechnicalReport
Website_URL

0
1

multisite active;
multisex active;
multispecies active
multisex inactive;
multispecies inactive

adrenal gland;
bone;
clitoral gland;
esophagus;
ear/Zymbal's gland;
gall bladder;
harderian gland;
hematopoietic system;
kidney;
large intestine;
liver;
lung;
mesovarium;
mammary gland;
mixture;
myocardium;
nasal cavity
nervous system;
oral cavity
ovary;
pancreas;
peritoneal cavity;
pituitary gland;
preputial gland;
prostate;
skin;
small intestine;
spleen;
stomach;
subcutaneous tissue;
bearing animals;
nd;
adder;
system.

CPDBAS_v5a_1547_25Oct2007

	ActivityCategory_		ActivityCategory_MultiCellCall_Details**			
Call	SingleCellCall	MultiCellCall**	multisite***	multisex	multispecies	Total Incidences* (CPDBAS_v5a)
Active (1)	1					224
	1	1	✓			81
	1	1		✓		113
	1	1			✓	8
	1	1	✓	✓		123
	1	1	✓		✓	27
	1	1		✓	✓	37
	1	1	✓	✓	✓	193
	1	1	Total MultiCellCall Incidences (1/1)			582
Inactive (0)	0		✓			169
	0	0	✓	✓		266
	0	0	✓		✓	15
	0	0	✓	✓	✓	288
	0	0	Total MultiCellCall Incidences (0/0)			569

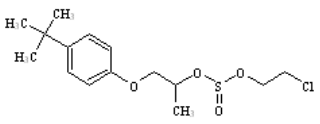
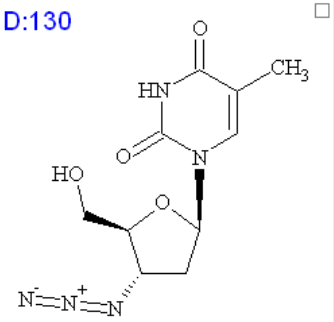
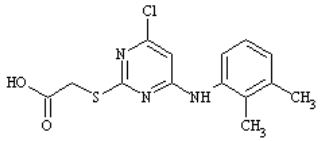
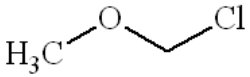
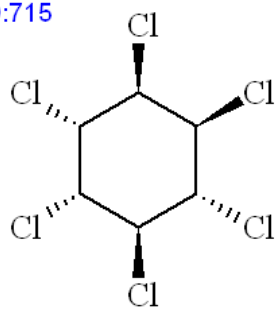
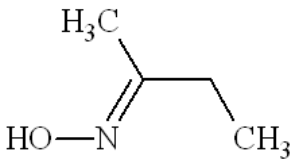
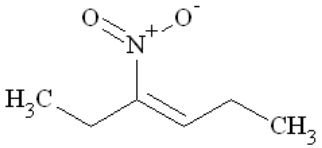
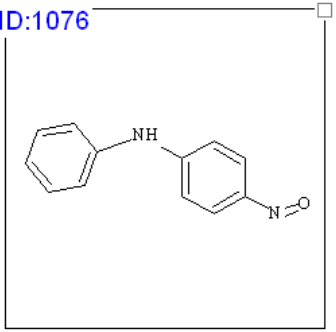
Are "**SingleCellCall Actives"** sufficient evidence?

If chemical is "**MultiCell Active**", it most likely crosses target site, sex and species.

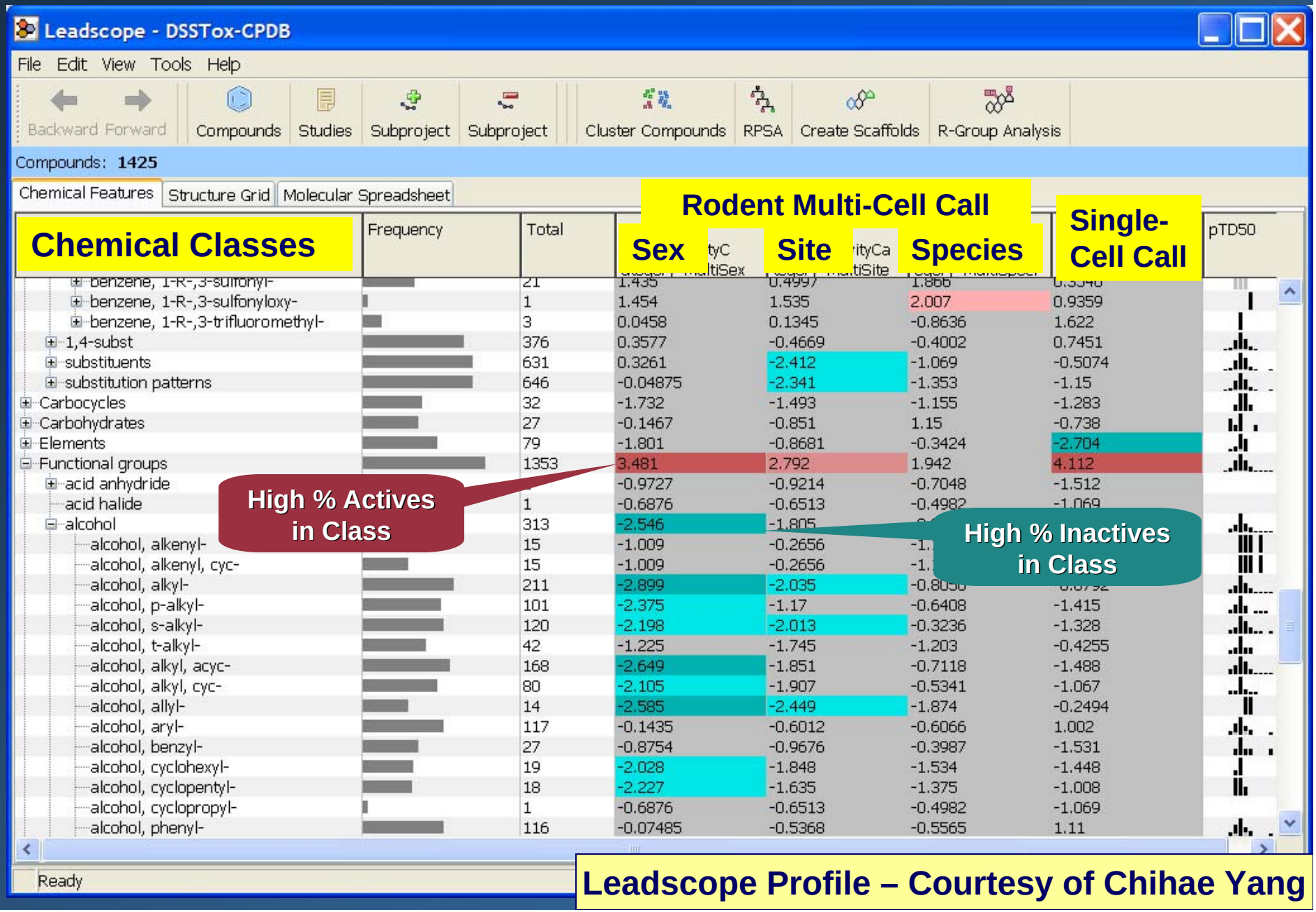
If chemical is "**MultiCell Inactive**", it most likely crosses sex and species.

CPDBAS_v5a_1547

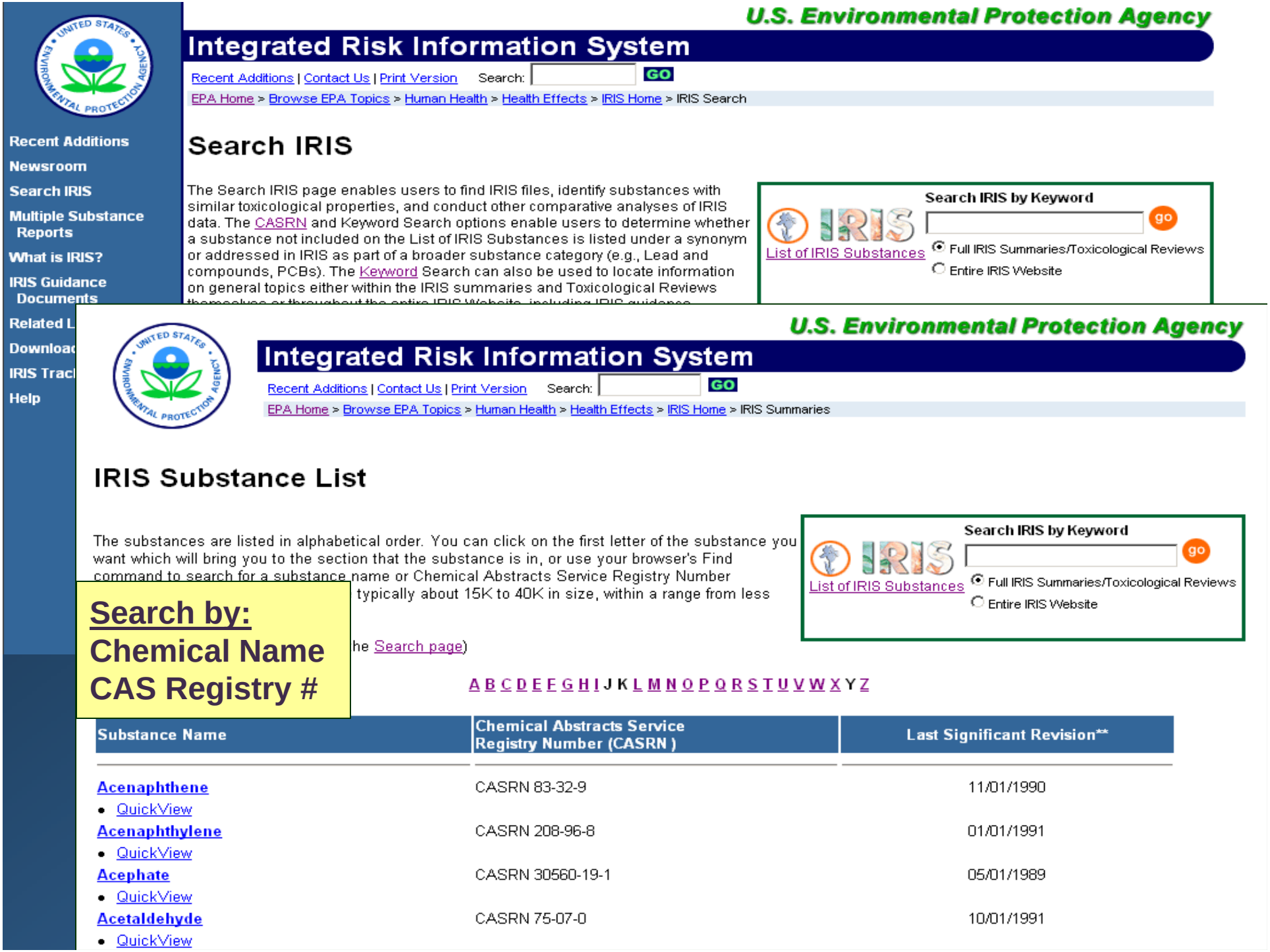
	ActivityCategory_		ActivityCategory_MultiCellCall_Details**			
Call	SingleCellCall	MultiCellCall**	multisite***	multisex	multispecies	Total Incidences* (CPDBAS_v5a)
Active (1)	1					224
	1	1	✓			81
	1	1		✓		113
	1	1				
	1	1				
	1	1				
	1	1				
	1	1				
	1	1				
	1	1				
Inactive (0)	0					
	0	0				
	0	0				
	0	0				
	0	0				

ID:98 	ID:130 	ID:299 	ID:316 
Aramite 140-57-8	AZT 30516-87-1	[4-Chloro-6-(2,3-xylidino)-2-p 50892-23-4	Chloromethyl methyl ether 107-30-2
ID:715 	ID:912 	ID:1000 	ID:1076 
alpha-1,2,3,4,5,6-Hexachloro 319-84-6	Methylethylketoxime 96-29-7	3-Nitro-3-hexene 4812-22-0	p-Nitrosodiphenylamine 156-10-5

Chemical-Biological Profiling of CPDBAS Activities



Leadscope Profile – Courtesy of Chihae Yang


U.S. Environmental Protection Agency
Integrated Risk Information System
Recent Additions | Contact Us | Print Version Search: GO
EPA Home > Browse EPA Topics > Human Health > Health Effects > IRIS Home > IRIS Search
Recent Additions
Newsroom
Search IRIS
Multiple Substance Reports
What is IRIS?
IRIS Guidance Documents
Related L
Download
IRIS Trac
Help
Search IRIS
The Search IRIS page enables users to find IRIS files, identify substances with similar toxicological properties, and conduct other comparative analyses of IRIS data. The CASRN and Keyword Search options enable users to determine whether a substance not included on the List of IRIS Substances is listed under a synonym or addressed in IRIS as part of a broader substance category (e.g., Lead and compounds, PCBs). The Keyword Search can also be used to locate information on general topics either within the IRIS summaries and Toxicological Reviews
Search IRIS by Keyword
go
List of IRIS Substances
Full IRIS Summaries/Toxicological Reviews
Entire IRIS Website
U.S. Environmental Protection Agency
Integrated Risk Information System
Recent Additions | Contact Us | Print Version Search: GO
EPA Home > Browse EPA Topics > Human Health > Health Effects > IRIS Home > IRIS Summaries
IRIS Substance List
The substances are listed in alphabetical order. You can click on the first letter of the substance you want which will bring you to the section that the substance is in, or use your browser's Find command to search for a substance name or Chemical Abstracts Service Registry Number
typically about 15K to 40K in size, within a range from less
he Search page)
Search IRIS by Keyword
go
List of IRIS Substances
Full IRIS Summaries/Toxicological Reviews
Entire IRIS Website
Search by:
Chemical Name
CAS Registry #
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
Substance Name
Chemical Abstracts Service Registry Number (CASRN)
Last Significant Revision**
Acenaphthene
CASRN 83-32-9
11/01/1990
QuickView
Acenaphthylene
CASRN 208-96-8
01/01/1991
QuickView
Acephate
CASRN 30560-19-1
05/01/1989
QuickView
Acetaldehyde
CASRN 75-07-0
10/01/1991
QuickView

U.S. Environmental Protection Agency

Integrated Risk Information System

Recent Additions | Contact Us Search: All EPA IRIS [Go]

You are here: EPA Home » Human Health » IRIS » IRIS Summaries » IRIS Quickviews

Aniline Quickview (CASRN 62-53-3)

Quickview Navigation

- view Anilin

Health assessm
comprehensive
Regional Offices

Disclaimer: Thi
the Full IRIS Sur

For definitions

Status of

File First On-Lin
Last Significant

Category (sec
Oral RfD Assess
Inhalation RfC As
Carcinogenicity As

Chronic Health Hazard Assessments for Noncarcinogenic Effects

Carcinogenicity Assessment for Lifetime Exposure

Weight of Evi

Weight of Evider

B2 (Probable hum

Weight of Evider

Induction of tumo
genetic toxicologi

This may be a syn

Principal and

- 20-26

Confidence in

- Study
- Databa
- RfC --

Quantitative Estimate of Carcinogenic Risk from Oral Exposure

Oral Slope Factor(s)	Extrapolation Method
5.7×10^{-3} per mg/kg-day	Linearized multistage procedure, extra risk

Drinking Water Unit Risks
1.6×10^{-7} per ug/L

Risk Level	Concentration
E-4 (1 in 10,000)	6×10^{2} ug/L
E-5 (1 in 100,000)	6×10^{1} ug/L
E-6 (1 in 1,000,000)	6 ug/L

Dose-Response Data (Carcinogenicity, Oral Exposure)

Tumor Type: Spleen, combined fibrosarcoma, stromal sarcoma, capsular sarcoma and hemangiosarcoma
Test Species: Rat/CD-F, male
Route: Oral, Diet
Reference: CIIT, 1982

Home
About DSSTox
Work in Progress
Frequent Questions
Structure Data Files
Central Field Definition Table
Apps, Tools & More
DSSTox Community
Site Map
Glossary of Terms
Help

SDF Download Page:

IRISTR: EPA Integrated Risk Information System (IRIS) Toxicity Review Data *Database File*

** *New DSSTox* Database File *28July2007:*
- Replaces previously published IRISSI_v1a Structure-Index Locater File
- Added 34 new IRIS Source-Specific data fields.

New Users: For general inform… and their use in Chemical Relation…

Description: *The following des…*

EPA's Integrated Risk Infor… National Center for Environ… found in the environment. IR… substances for use in risk a… training in toxicology, but w… following categories:

- Oral reference doses
- Hazard identification,

File Types	Description	File Size	Format
Documentation Files			
Log File	IRISTR_LogFile_28Jul2007.pdf (PDF, 4 pp.)	38KB	PDF
Field Definition File	IRISTR_FieldDefFile_28Jul2007.pdf (PDF, 8 pp.)	75KB	PDF
	IRISTR_FieldDefFile_28Jul2007.doc	204KB	DOC
Data Files: *IRISTR*			
SDF Structure Data File	IRISTR_v1a_544_28Jul2007.sdf	*.zip 1.1 MB	SDF
• Data Table (no structures)	IRISTR_v1a_544_28Jul2007_nostructures.xls		XLS
• Structures Table	IRISTR_v1a_544_28Jul2007_structures.pdf (PDF, 11pp., 1.8MB)		PDF

File Error Report

IRISTR SDF Fields (55 total)

DSSTox Standard Chemical Fields (18)

DSSTox Standard Toxicity Fields (3)

Oral_RfD_Assessed
Oral_RfD_CriticalEffects
Oral_RfD_mg_per_kg_day
Oral_RfD_mmol_per_kg_day
Oral_RfD_Notes
Oral_RfD_Confidence

Inhalation_RfC_Assessed
Inhalation_RfC_CriticalEffects
Inhalation_RfC_mg_per_m3
Inhalation_RfC_mmol_per_m3
Inhalation_RfC_Notes
Inhalation_RfC_Confidence

WtOfEvidence_Cancer_Assessed
WtOfEvidence_Cancer_Concern
WtOfEvidence_1986GuidelineCategories
WtOfEvidence_UpdatedGuidelineUsed
WtOfEvidence_Cancer_Narrative

DrinkingWater_OralSlope_Assessed
DrinkingWater_PrecursorEffect_TumorType
DrinkingWater_OralSlopeFactor_mg_per_kg_day
DrinkingWater_OralSlopeFactor_mmol_per_kg_day
DrinkingWater_ExtrapolationMethod_Notes
DrinkingWater_UnitRisk_microg_per_L
DrinkingWater_UnitRisk_micromol_per_L
DrinkingWater_StudyRoute

Inhalation_UnitRisk_Assessed
Inhalation_PrecursorEffect_TumorType
Inhalation_UnitRisk_microg_per_m3
Inhalation_UnitRisk_micromol_per_m3
Inhalation_StudyRoute
Inhalation_ExtrapolationMethod_Notes

TotalAssessments
Note_IRISTR
Website_URL (field contains chemical-specific links to IRIS "Quick View

IRISTR Toxicity Review Areas	Totals_v1a*
Oral RfDs	357
Inhalation RfCs	70
Weight of Evidence Characterizations	243
Oral Slope Factors/Drinking Water Unit Risks	76
Air Unit Risks	54
Total Assessments	800

Number of IRISTR Toxicity Review Assessments per Chemical	Totals_v1a**
0	(39)
1	323
2	101
3	54
4	22
5	5
Total Assessments	800

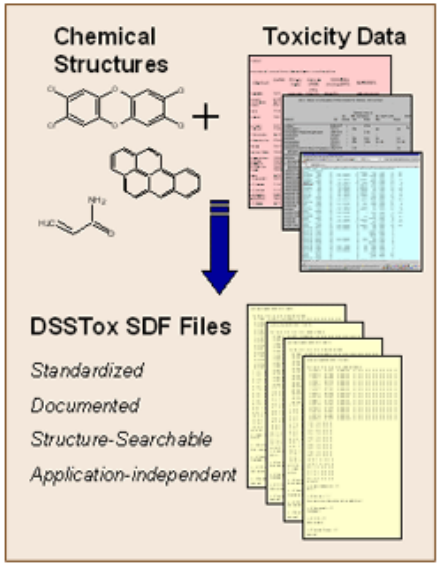
About DSSTox
Work in Progress
Frequent Questions
Structure Data Files
Central Field Definition Table
Apps, Tools & More
DSSTox Community
Site Map
Glossary of Terms
Help

DSSTox

Distributed Structure-Searchable Toxicity (DSSTox) Database Network is a project of EPA's Computational Toxicology Program, helping to build a public data foundation for improved structure-activity and predictive toxicology capabilities. The DSSTox website provides a public forum for publishing downloadable, structure-searchable, standardized chemical structure files associated with toxicity data. More>

EPA DSSTox Structure-Browser v1.0 GO

DSSTox Structure-Browser information Page



- DSSTox Graphic Flowchart
- DSSTox Project Goals
- DSSTox Publications

Recent Additions: 27 September 2007

- TOXCST: Research Chemical Inventory for EPA's ToxCast Program - Updated to v2a

Recent Additions: 28 August 2007

***Launch of DSSTox Structure-Browser v1.0:

- A new structure-search capability for published DSSTox Data Files, allows users to search by DSSTox Standard Chemical Fields and includes options for:

  - **Text Search:** Chemical Name, CAS RN, InChI, Formula
  - **Structure Search (*Exact, Substructure, Similarity*):** SMILES or Structure Drawing Tool entry

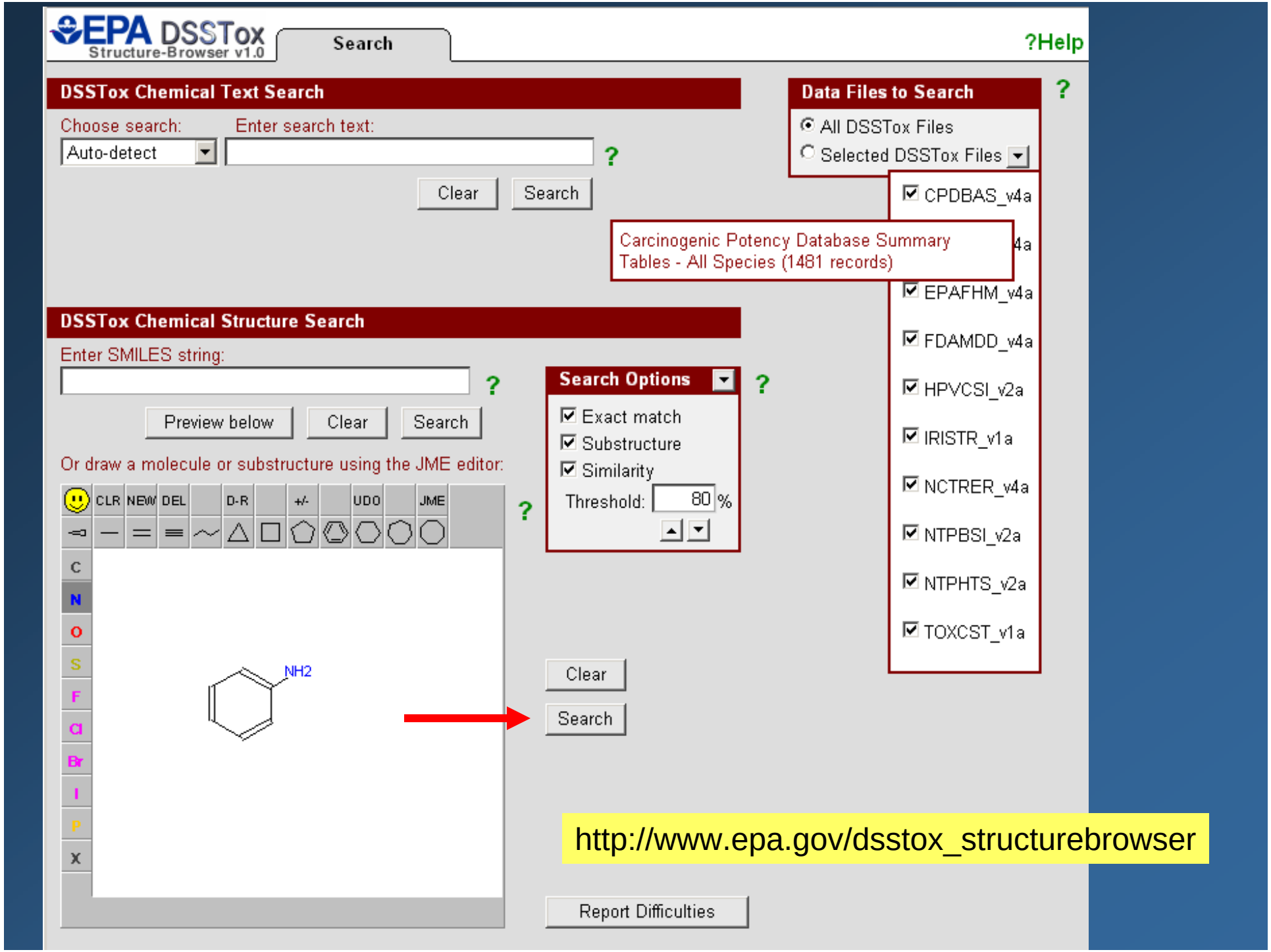
***Revised Standard ID Fields for all DSSTox files:

- Modified Record, File, Chemical, and Substance ID fields to index all unique DSSTox structures and substances, also with respect to file record and version

DSSTox Data Files: Details>

CPDBAS_v4a_1481_15Jun2007 **New content
DBPCAN_v4a_209_15Jun2007
EPAFHM_v4a_617_15Jun2007
FDAMDD_v3a_1216_25Jul2007
HPVCSI_v2a_3548_15Aug2007**New content
IRISTR_v1a_544_28Jul2007 **New file
NCTRER_v4a_232_15Jun2007
NTPBSI_v2a_2293_24Aug2007 **Updated content
NTPHTS_v1a_1408_25Jul2007
TOXCST_v2a_320_25Sep2007 **Updated

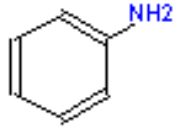
More on Data File Types


EPA DSSTox
Structure-Browser v1.0
Search
?Help
DSSTox Chemical Text Search
Choose search:
Auto-detect
Enter search text:
Clear
Search
Data Files to Search
All DSSTox Files
Selected DSSTox Files
CPDBAS_v4a
Carcinogenic Potency Database Summary Tables - All Species (1481 records)
EPAFHM_v4a
FDAMDD_v4a
HPVCSI_v2a
IRISTR_v1a
NCTRER_v4a
NTPBSI_v2a
NTPHTS_v2a
TOXCST_v1a
DSSTox Chemical Structure Search
Enter SMILES string:
Preview below
Clear
Search
Search Options
Exact match
Substructure
Similarity
Threshold: 80 %
Or draw a molecule or substructure using the JME editor:
CLR NEW DEL D-R +/- UDO JME
C N O S F Cl Br I P X
NH2
Clear
Search
Report Difficulties
http://www.epa.gov/dsstox_structurebrowser

EPA DSSTox Structure-Browser v1.0

Search | File Incidences | ?Help

Search Results Summary for DSSTox Substances - File Breakdown Incidences

Query	Results Type	Hits	Display
Structure: 	Exact matches	1	Details
	Substructures	1130	Details
	Similarity > 80%	8	Details

?

DSSTox File ?	Total#Records	Exact matches	Substructures	Similarity > 80%
CPDBAS_v4a	1481	1	309	7
DBPCAN_v4a	209	-	1	-
EPAFHM_v4a	617	1	102	1
FDAMDD_v4a	1216	-	261	-
HPVCSI_v2a	3548	1	216	4
IRISTR_v1a	544	1	68	2
NCTRER_v4a	232	-	10	-
NTPBSI_v2a	2293	1	494	5
NTPHTS_v2a	1408	1	313	3
TOXCST_v1a	340	-	80	-
Total Unique Substance Hits		**1**	**1130**	**8**
Total Substance Hits - All Files		**6**	**1854**	**22**

?

EPA DSSTox Structure-Browser v1.0

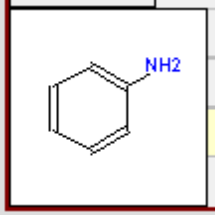
Search | File Incidences | **Search Details** | ?Help

Details

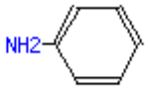
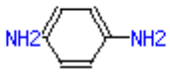
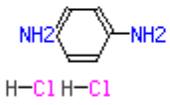
Output Options

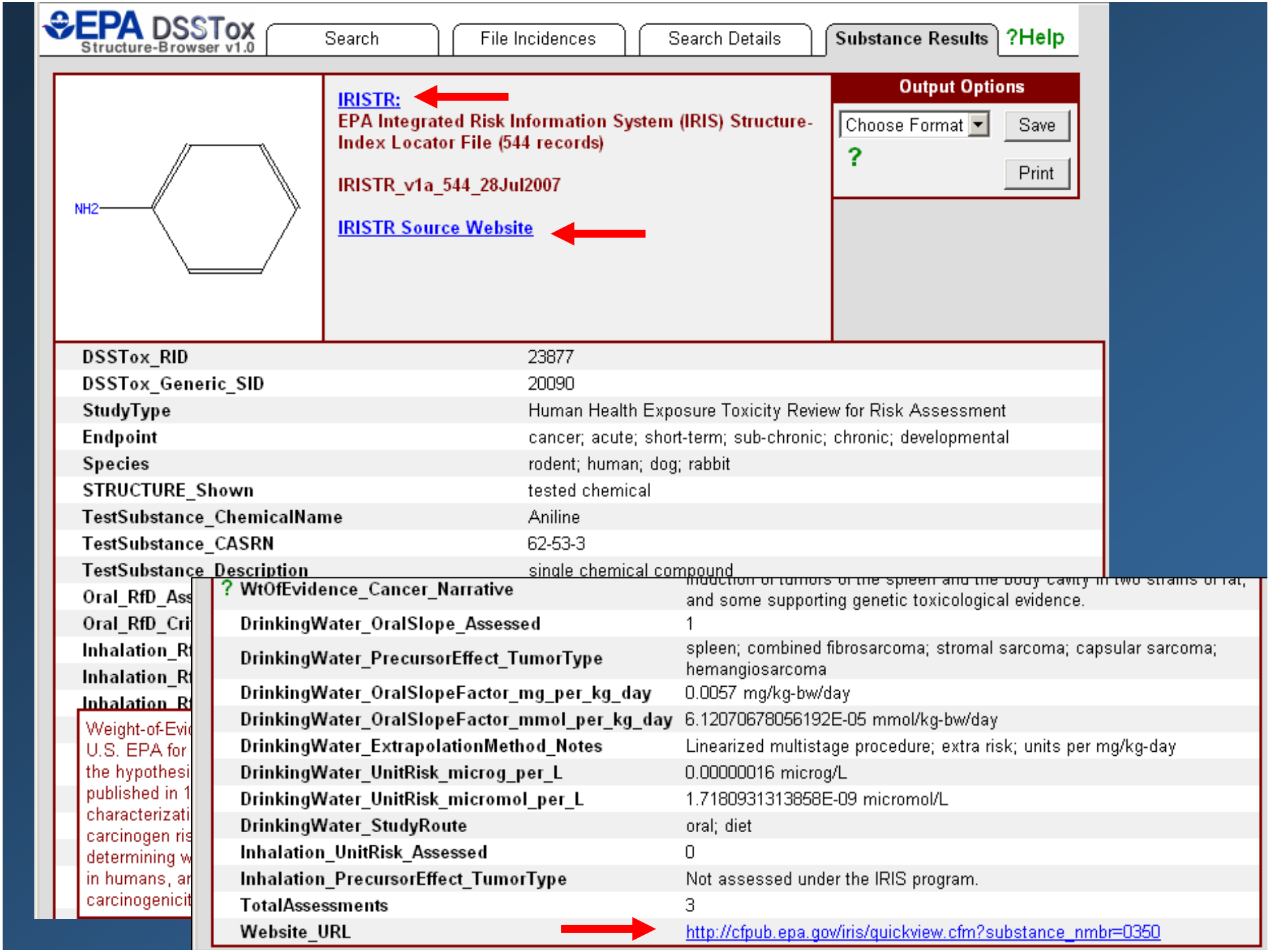
Choose Format | Save | Print

?

Query	Results Type	Hits	Display
Structure: 	Exact matches	1	Details
	Substructures	1130	Details
	Similarity > 80%	8	Details

?

? DSSTox Substance ID	? Similarity Score%	Structure Match	Substance Name	CASRN	Substance Description	Details (Data Files)
20090	100		Aniline	62-53-3	single chemical compound	CPDBAS EPAFHM HPVCSI IRISTR NTPBSI NTPHTS
20091	94.1		Aniline.HCl	142-04-1	single chemical compound	CPDBAS NTPBSI
21138	94.1		p-Phenylenediamine	106-50-3	single chemical compound	CPDBAS HPVCSI
21141	88.8		p-Phenylenediamine.2HCl	624-18-0	single chemical compound	CPDBAS NTPBSI


EPA DSSTox
Structure-Browser v1.0
Search
File Incidences
Search Details
Substance Results
?Help
NH2
IRISTR:
EPA Integrated Risk Information System (IRIS) Structure-Index Locator File (544 records)
IRISTR_v1a_544_28Jul2007
IRISTR Source Website
Output Options
Choose Format
Save
Print
DSSTox_RID 23877
DSSTox_Generic_SID 20090
StudyType Human Health Exposure Toxicity Review for Risk Assessment
Endpoint cancer; acute; short-term; sub-chronic; chronic; developmental
Species rodent; human; dog; rabbit
STRUCTURE_Shown tested chemical
TestSubstance_ChemicalName Aniline
TestSubstance_CASRN 62-53-3
TestSubstance_Description single chemical compound
? WtOfEvidence_Cancer_Narrative
and some supporting genetic toxicological evidence.
DrinkingWater_OralSlope_Assessed 1
DrinkingWater_PrecursorEffect_TumorType spleen; combined fibrosarcoma; stromal sarcoma; capsular sarcoma; hemangiosarcoma
DrinkingWater_OralSlopeFactor_mg_per_kg_day 0.0057 mg/kg-bw/day
DrinkingWater_OralSlopeFactor_mmol_per_kg_day 6.12070678056192E-05 mmol/kg-bw/day
DrinkingWater_ExtrapolationMethod_Notes Linearized multistage procedure; extra risk; units per mg/kg-day
DrinkingWater_UnitRisk_microg_per_L 0.00000016 microg/L
DrinkingWater_UnitRisk_micromol_per_L 1.7180931313858E-09 micromol/L
DrinkingWater_StudyRoute oral; diet
Inhalation_UnitRisk_Assessed 0
Inhalation_PrecursorEffect_TumorType Not assessed under the IRIS program.
TotalAssessments 3
Website_URL http://cfpub.epa.gov/iris/quickview.cfm?substance_nmbr=0350

Aniline Quickview (CASRN 62-53-3)

Quickview Navigation

+ Expand navigation:

- view Aniline Summary

Health assessment information on a chemical substance is included in IRIS only after a comprehensive review of toxicity data by U.S. EPA health scientists from several Program Offices, Regional Offices, and the Office of Research and Development.

Disclaimer: This QuickView represents a snapshot of key information. We suggest that you read the Full IRIS Summary to put this information into complete context.

For definitions of terms in the IRIS Web site, refer to the IRIS Glossary.

Status of Data for Aniline

File First On-Line: 09/07/1988
Last Significant Revision: 11/01/1990

Category (section)	Status	Last Revised
Oral RfD Assessment	No data	
Inhalation RfC Assessment	On-line	12/01/1993
Carcinogenicity Assessment	On-line	02/01/1994

Chronic Health Hazard Assessments for Noncarcinogenic Effects

IRIS Home
Recent Additions
Newsroom
Search IRIS
Multiple Substance Reports
What is IRIS?
IRIS Guidance Documents
Related Links
Download IRIS
IRIS Track
Help

IRIS Substance List

The substances are listed in alphabetical order. You can click on the first letter of the substance you want which will bring you to the section that the substance is in, or use your browser's Find command to search for a substance name or Chemical Abstracts Service Registry Number (CASRN). These substance files are typically about 15K to 40K in size, within a range from less than 10K up to about 120K.

(To search the IRIS database, go to the Search page)

List of IRIS Substances

Search IRIS by Keyword go

Full IRIS Summaries/Toxicological Reviews

Entire IRIS Website

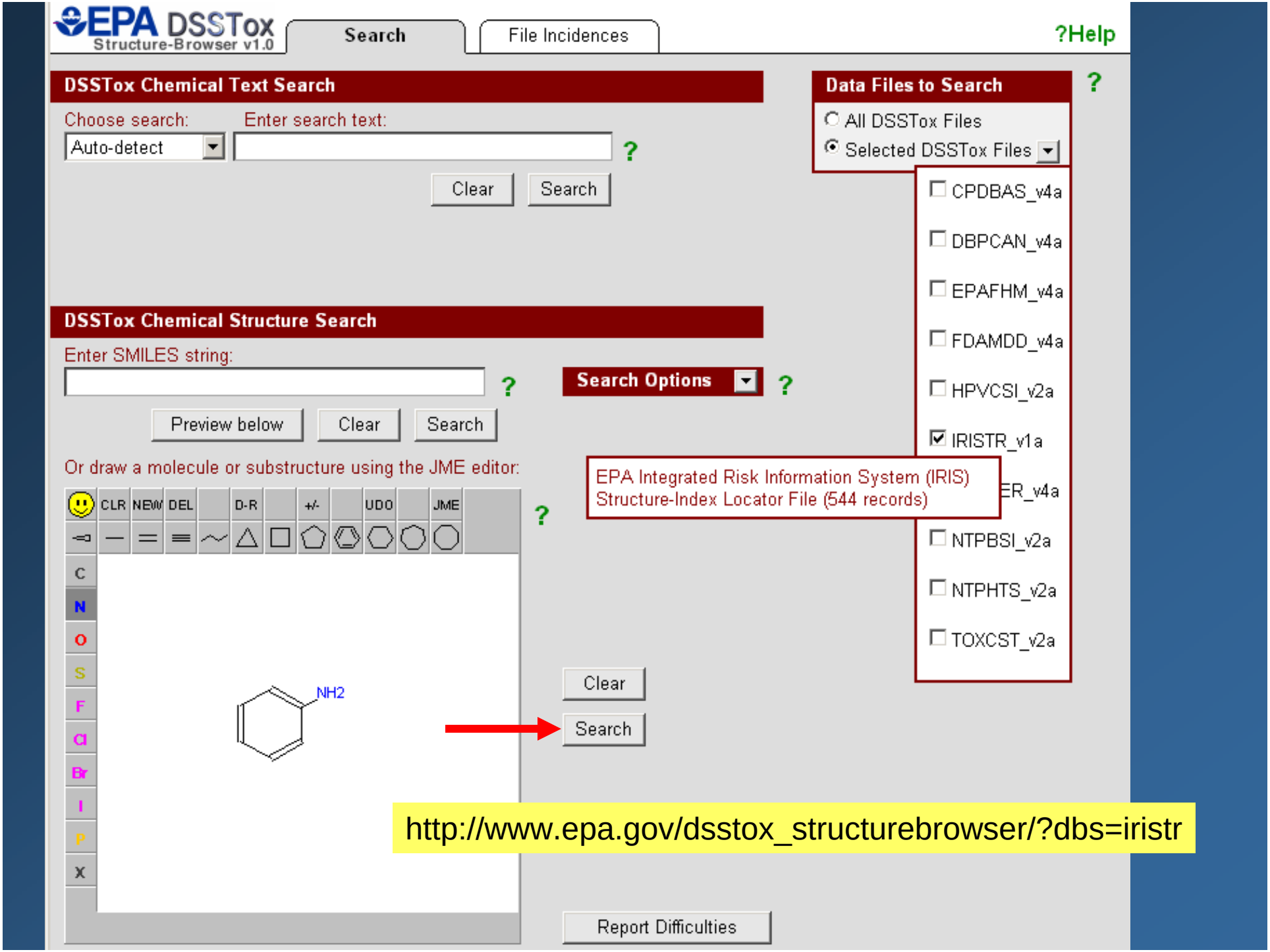
Structure-search

EPA DSSTox GO Structure-Browser v1.0

You will need Adobe Reader to view some of the files on this page. See EPA's PDF page to learn more.

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

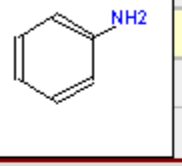
Substance Name	Chemical Abstracts Service Registry Number (CASRN)	Last Significant Revision*
Acenaphthene	CASRN 83-32-9	11/01/1990
• QuickView		
Acenaphthylene	CASRN 208-96-8	01/01/1991
• QuickView		
Acephate	CASRN 30560-19-1	05/01/1989
• QuickView		
Acetaldehyde	CASRN 75-07-0	10/01/1991
• QuickView		
Acetochlor	CASRN 34256-82-1	09/01/1993


EPA DSSTox
Structure-Browser v1.0
Search
File Incidences
?Help
DSSTox Chemical Text Search
Choose search:
Auto-detect
Enter search text:
Clear
Search
Data Files to Search
All DSSTox Files
Selected DSSTox Files
CPDBAS_v4a
DBPCAN_v4a
EPAFHM_v4a
FDAMDD_v4a
HPVCSI_v2a
IRISTR_v1a
NTPBSI_v2a
NTPHTS_v2a
TOXCST_v2a
DSSTox Chemical Structure Search
Enter SMILES string:
Search Options
Preview below
Clear
Search
Or draw a molecule or substructure using the JME editor:
CLR
NEW
DEL
D-R
+/-
UDO
JME
NH2
EPA Integrated Risk Information System (IRIS)
Structure-Index Locator File (544 records)
Clear
Search
Report Difficulties
http://www.epa.gov/dsstox_structurebrowser/?dbs=iristr

EPA DSSTox Structure-Browser v1.0

Search | File Incidences | Search Details | ?Help

Details

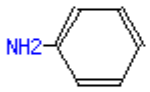
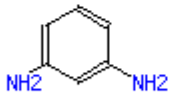
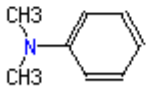
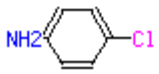
Query	Results Type	Hits	Display
Structure: 	Exact matches	1	Details
	Substructures	68	Details
	Similarity > 80%	2	Details

?

Output Options

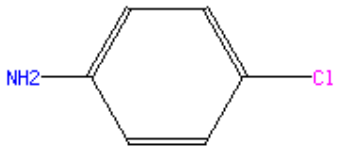
Choose Format | Save | Print

?

? DSSTox Substance ID	? Similarity Score%	Structure Match	Substance Name	CASRN	Substance Description	Details (Data Files)
20090	100		Aniline	62-53-3	single chemical compound	IRISTR
21137	88.8		m-Phenylenediamine	108-45-2	single chemical compound	IRISTR
20507	69.5		N,N-Dimethylaniline	121-69-7	single chemical compound	IRISTR
20295	61.5		p-Chloroaniline	106-47-8	single chemical compound	IRISTR

EPA DSSTox Structure-Browser v1.0

Search | File Incidences | Search Details | **Substance Results** | ?Help



IRISTR:
EPA Integrated Risk Information System (IRIS) Structure-Index Locator File (544 records)

IRISTR_v1a_544_28Jul2007

IRISTR Source Website

Output Options

Choose Format | Save

?

Print

DSSTox_RID	23963
DSSTox_Generic_SID	20295
StudyType	Human Health Exposure Toxicity Review for Risk Assessment
Endpoint	cancer; acute; short-term; sub-chronic; chronic; developmental
Species	rodent; human; dog; rabbit
STRUCTURE_Shown	tested chemical
TestSubstance_ChemicalName	p-Chloroaniline
TestSubstance_CASRN	106-47-8
TestSubstance_Description	single chemical compound
Oral_RfD_Assessed	1
Oral_RfD_CriticalEffects	nonneoplastic lesions splenic capsule
Oral_RfD_mg_per_kg_day	0.004 mg/kg-bw/day
Oral_RfD_mmol_per_kg_day	3.13549656467158E-05 mmol/kg-bw/day
Oral_RfD_Notes	LOAEL (Lowest observed adverse effect level): 12.5 mg/kg-day
Oral_RfD_Confidence	Low
Inhalation_RfC_Assessed	0
Inhalation_RfC_CriticalEffects	Not assessed under the IRIS program.
WtOfEvidence_Cancer_Assessed	0
WtOfEvidence_1986GuidelineCategories	Not assessed under the IRIS program.
DrinkingWater_OralSlope_Assessed	0
DrinkingWater_PrecursorEffect_TumorType	Not assessed under the IRIS program.
Inhalation_UnitRisk_Assessed	0
Inhalation_PrecursorEffect_TumorType	Not assessed under the IRIS program.
TotalAssessments	1
Website_URL	http://cfpub.epa.gov/iris/quickview.cfm?substance_nmbr=0320

p-Chloroaniline Quickview (CASRN 106-47-8)

- view p-Chloroaniline Summary

Health assessment information on a chemical substance is included in IRIS only after a comprehensive review of toxicity data by U.S. EPA health scientists from several Program Offices, Regional Offices, and the Office of Research and Development.

Disclaimer: This QuickView represents a snapshot of key information. We suggest that you read the Full IRIS Summary to put this information into complete context.

For definitions of terms in the IRIS Web site, refer to the IRIS Glossary.

Status of Data for p-Chloroaniline

File First On-Line: 08/22/1988
Last Significant Revision: 08/22/1988

Category (section)	Status	Last Revised
Oral RfD Assessment	On-line	02/01/1995
Inhalation RfC Assessment	No data	
Carcinogenicity Assessment	No data	

Chronic Health Hazard Assessments for Noncarcinogenic Effects

Reference Dose for Chronic Oral Exposure (RfD)

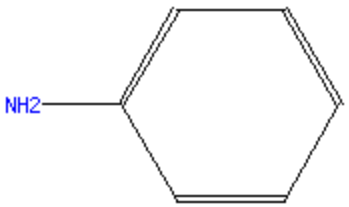
Critical Effect	Point of Departure	UF	MF	RfD
Nonneoplastic lesions of splenic capsule	LOAEL : 12.5 mg/kg-day	3000	1	4×10^{-3} mg/kg-day

The Point of Departure listed serves as a basis from which the Oral RfD was derived. See Discussion of Conversion Factors and Assumptions for more details.


EPA IRIS Website
DSSTox Structure-Browser

EPA DSSTox Structure-Browser v1.0

Search | File Incidences | Search Details | **Substance Results** | ?Help



IRISTR:
EPA Integrated Risk Information System (IRIS) Structure-Index Locator File (544 records)

IRISTR_v1a_544_28Jul2007

IRISTR Source Website

Output Options

Choose Format | Save | Print

?

PubChem

Links directly to chemical data page for Aniline in PubChem

DSSTox_RID	23877
DSSTox_Generic_SID	20090
StudyType	Human Health Exposure Toxicity Review for Risk Assessment
Endpoint	cancer; acute; short-term; sub-chronic; chronic; developmental
Species	rodent; human; dog; rabbit
STRUCTURE_Shown	tested chemical
TestSubstance_ChemicalName	Aniline
TestSubstance_CASRN	62-53-3
TestSubstance_Description	single chemical compound
Oral_RfD_Assessed	0
Oral_RfD_CriticalEffects	Not assessed under the IRIS program.
Inhalation_RfC_Assessed	1
Inhalation_RfC_CriticalEffects	mild spleen toxicity
Inhalation_RfC_mg_per_m3	0.001 mg/m3
Inhalation_RfC_mmol_per_m3	1.07380820711613E-05 mmol/m3
Inhalation_RfC_Notes	NOAEL (No observed adverse effect level) HEC (Human Equivalent Concentration): 3.4 mg/m3
Inhalation_RfC_Confidence	Low
WtOfEvidence_Cancer_Assessed	1
WtOfEvidence_Cancer_Concern	Medium
WtOfEvidence_1986GuidelineCategories	B2; Probable human carcinogen - based on sufficient evidence of carcinogenicity in animals
WtOfEvidence_Cancer_Narrative	Induction of tumors of the spleen and the body cavity in two strains of rat; and some supporting genetic toxicological evidence.

NCBI

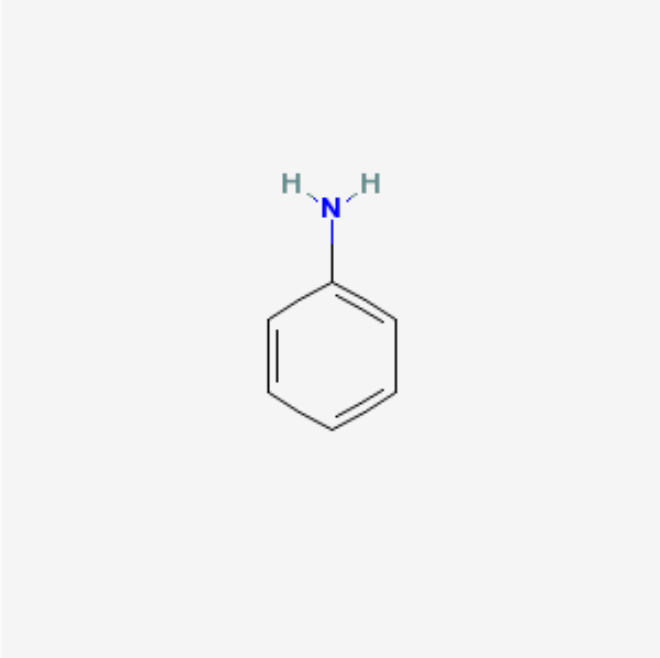
PubChem

Information on biological activities of small molecules

HOME | SEARCH | SITE MAP | PubMed | Entrez | Structure | GenBank | PubChem | Help

Search PubChem Compound GO

Compound Summary:



CID: 6115

BioActivity: Summary

All: 53 Links
Active: 6 Links
Inactive: 45 Links
Inconclusive: 1 Link

Protein Structures: 3 Links

Protein Sequences: 15 Links

NLM Toxicology: Link

Substances:

All: 270 Links
Same: 28 Links
Mixture: 242 Links

Related Compounds:

Same, Connectivity: 13 Links

Similar Compounds: 164 Links

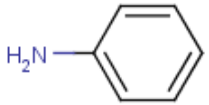
Structure Search

MeSH | Synonyms | Properties | Descriptors | Category | Exports

United States National Library of Medicine

ChemIDplus Advanced

NAME: Aniline
RN: 62-53-3



Basic Information

Full Record

Structure

Names & Synonyms

Formulas

Classification Codes

Registry Numbers

Toxicity

Physical Properties

For more information about this substance, you may select from the the links below.

File Locator

AIDSLINE	AIDS Citations from MEDLINE
CANCERLIT	CANCER LITerature from Medline
CCRIS	NCI Chem Carcino Res Info Sys
DART/ETIC	Developmental and Reprod.Tox.
DSL	Domestic Sub. List of Canada
EINECS	EU Inv of Exist. Comm. Chem Sub
EMIC	Env. Mutagen Info. Center
GENETOX	EPA GENetic TOXicology
HSDB	Hazardous Substances Data Bank
Haz-Map	Occ. Exposure to Haz. Agents
IRIS	EPA Integrated Risk Info. System
ITER	International Tox. Est. for Risk
MEDLINE	MEDical literature onLINE
MeSH	Medical Subject Headings File
PubChem	PubChem
RTECS	Reg. of Toxic Eff. of Chem. Sub.
TOXLINE Core	NLM TOXLINE Core from MEDLINE
TOXLINE Special	NLM TOXLINE Special on TOXNET
TOXMAP	NLM Enviro. Health e-Maps
TRI2000	EPA Toxics Release Inv. 2000

Search Navigation

Start New Query

Modify Query

Show Query

Search History

Structure Similarity Search

Structure Salt/Parent Search

Transfer Structure

Basic ChemIDplus Search

NCBI

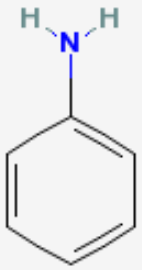
PubChem

Information on biological activities of small molecules

HOME | SEARCH | SITE MAP | PubMed | Entrez | Structure | GenBank | PubChem | Help

Search PubChem Compound GO

Compound Summary:



CID: 6115

BioActivity: Summary

All: 53 Links
Active: 6 Links
Inactive: 45 Links
Inconclusive: 1 Link

Protein Structures: 3 Links

Protein Sequences: 15 Links

NLM Toxicology: Link

Substances:

All: 270 Links
Same: 28 Links
Mixture: 242 Links

Related Compounds:

Same, Connectivity: 13 Links

Similar Compounds: 164 Links

Structure Search

MeSH | Synonyms | Properties | Descriptors | Category | Exports

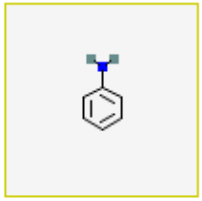
NCBI

PubChem

Information on biological activities of small molecules

HOME | SEARCH | SITE MAP | PubMed | Entrez | Structure | GenBank | PubChem | Help

BioActivity Summary: -- Original CID: 6115



Total BioAssay Count: 47

Total Substance Count: 4

Total Compound Count: 1

Tested: 1
Active: 1
Inactive: 1

Revise Substance Selection:

Select Active
Add Active
Add Tested
Add Similar Substances

Revise BioAssay Selection:

Select Active
Add Active

Structure Activity Analysis

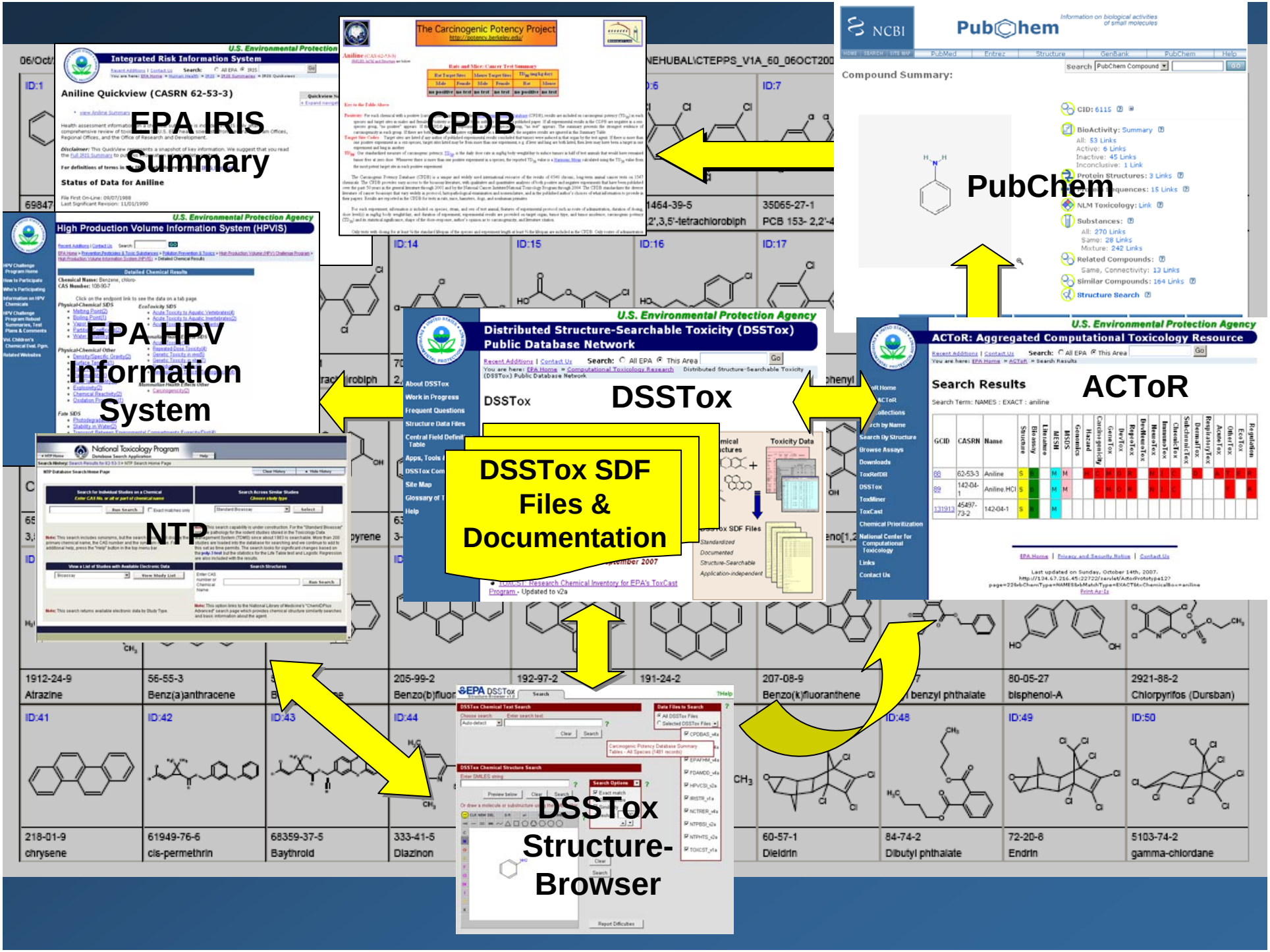
BioAssay Data Table

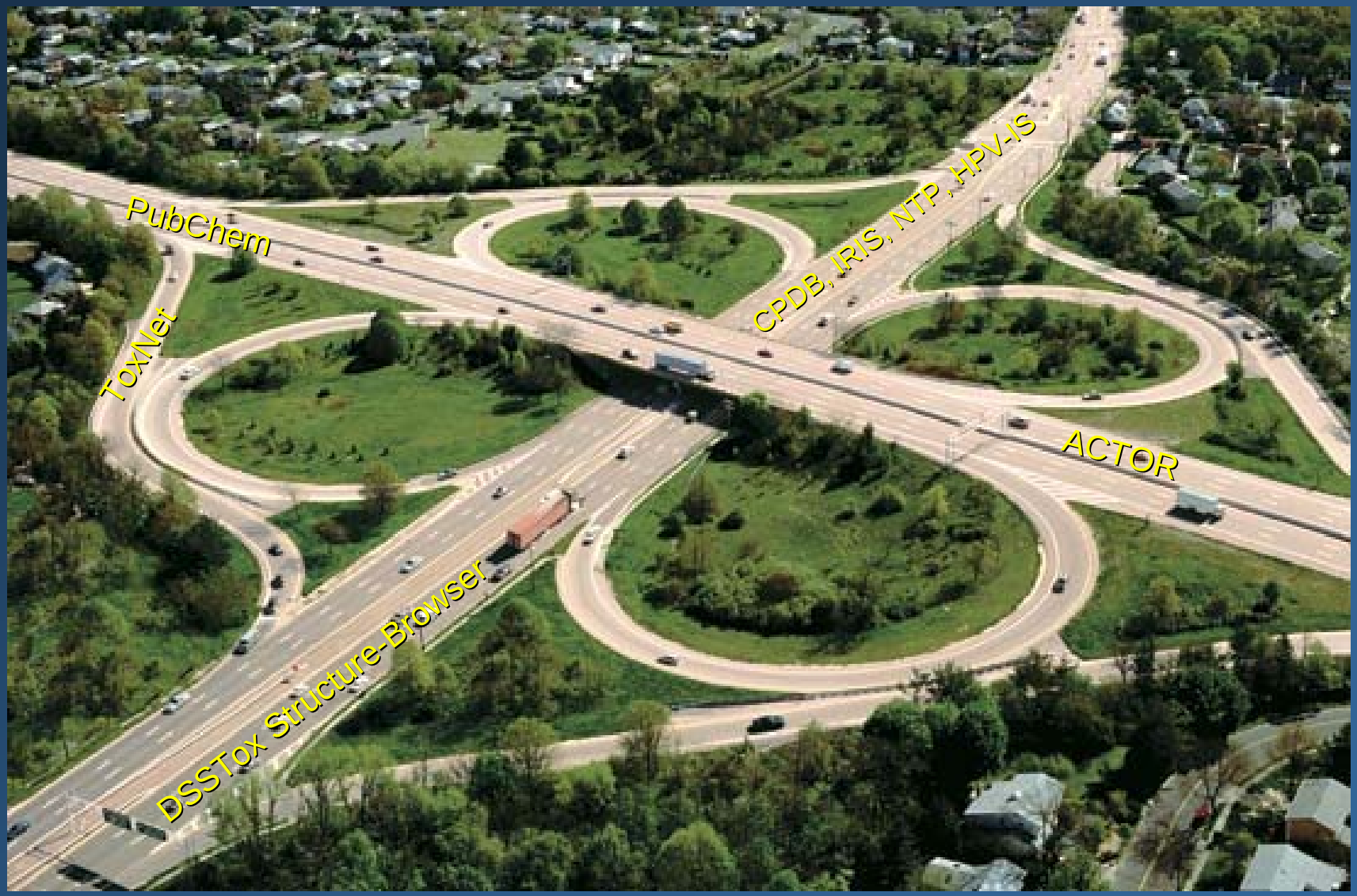
		AID						
42	☐	426		1		1	Confirmatory	Cell Viability - Jurkat
43	☐	421		1		1	Confirmatory	Cell Viability - BJ
44	☐	356				1	Confirmatory	EPA Fathead Minnow Acute Toxicity database (EPAFHM)
45	☐	352		1		1	Other	Carcinogenic Potency Database (CPDBAS)
46	☐	256		1		1		NCI In Vivo Anticancer Drug Screen. Data for tumor model L1210 Leukemia
47	☐	248		1		1		

Tot

#		AID	Active	Inactive	Discrepant	Tested	Outcome Method	
1	☐	175	1	1	1	1		NCI Yea
2	☐	167	1	1	1	1		NCI Yeast Anticancer Drug Screen. Data for the bub3 strain

Current DSSTox files will be deposited in PubChem, providing links to DSSTox files (e.g., IRISTR) and Source web pages (EPA IRIS)


EPA IRIS Summary
CPDB
PubChem
EPA HPV Information System
DSSTox
ACToR
DSSTox SDF Files & Documentation
NTP
DSSTox Structure-Browser

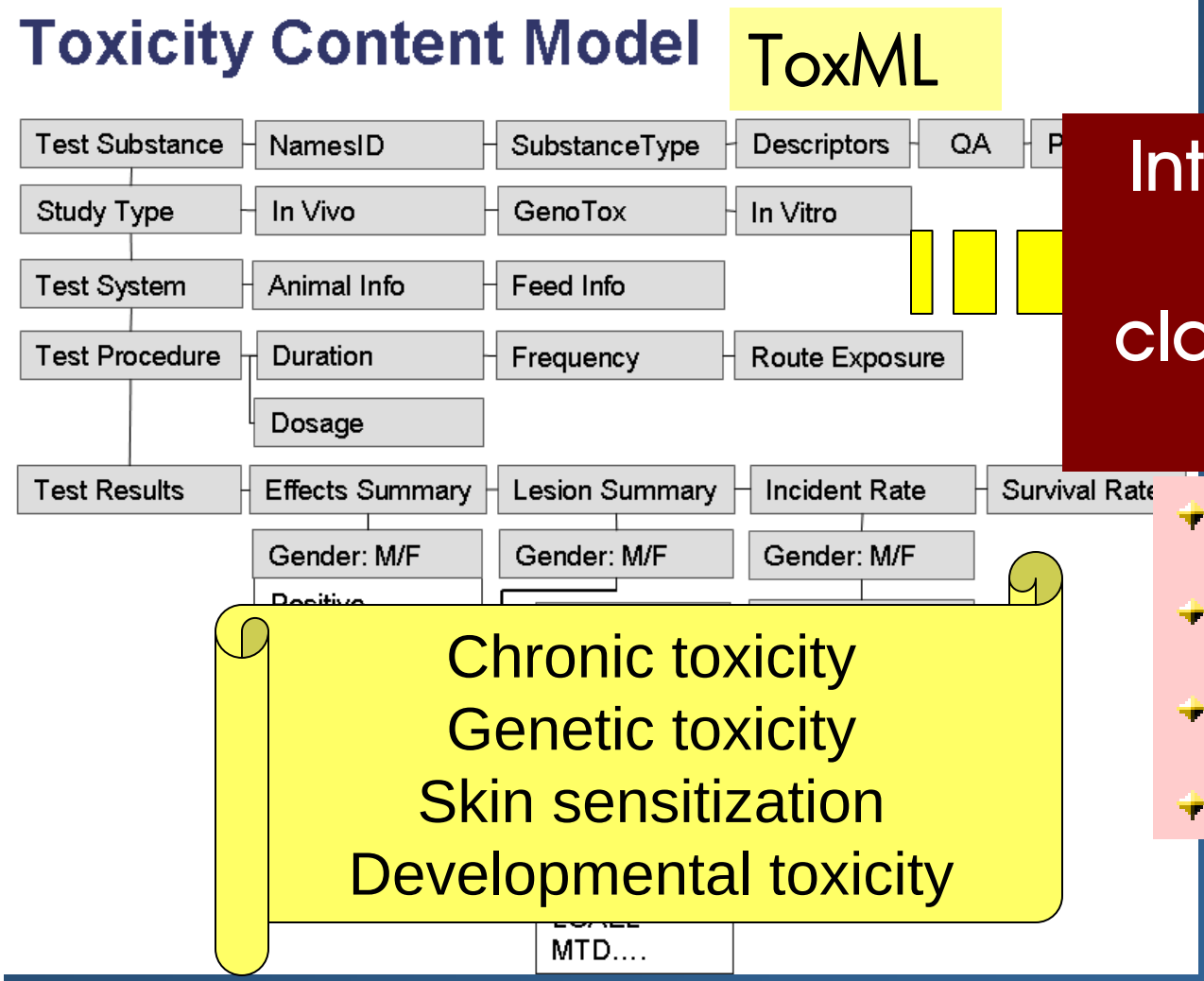

PubChem
ToxNet
CPDB, IRIS, NTP , HPV-IS
ACTOR
DSSTox Structure-Browser

Part III.

Data Models

Toxicity Experimental Data → Summary Data:

ToxML / LIST Collaborations: FDA CDER/CFSAN



Intermediate toxicity classifications for SAR

+

−

- Activity categories
- Potency categories
- Mode of action categories
- Summary calls

Chronic toxicity
Genetic toxicity
Skin sensitization
Developmental toxicity

LeadScope

ILSI Developmental Toxicity Database Project

Sponsor: Health Canada

ILSI Coordinator: Beth Julien

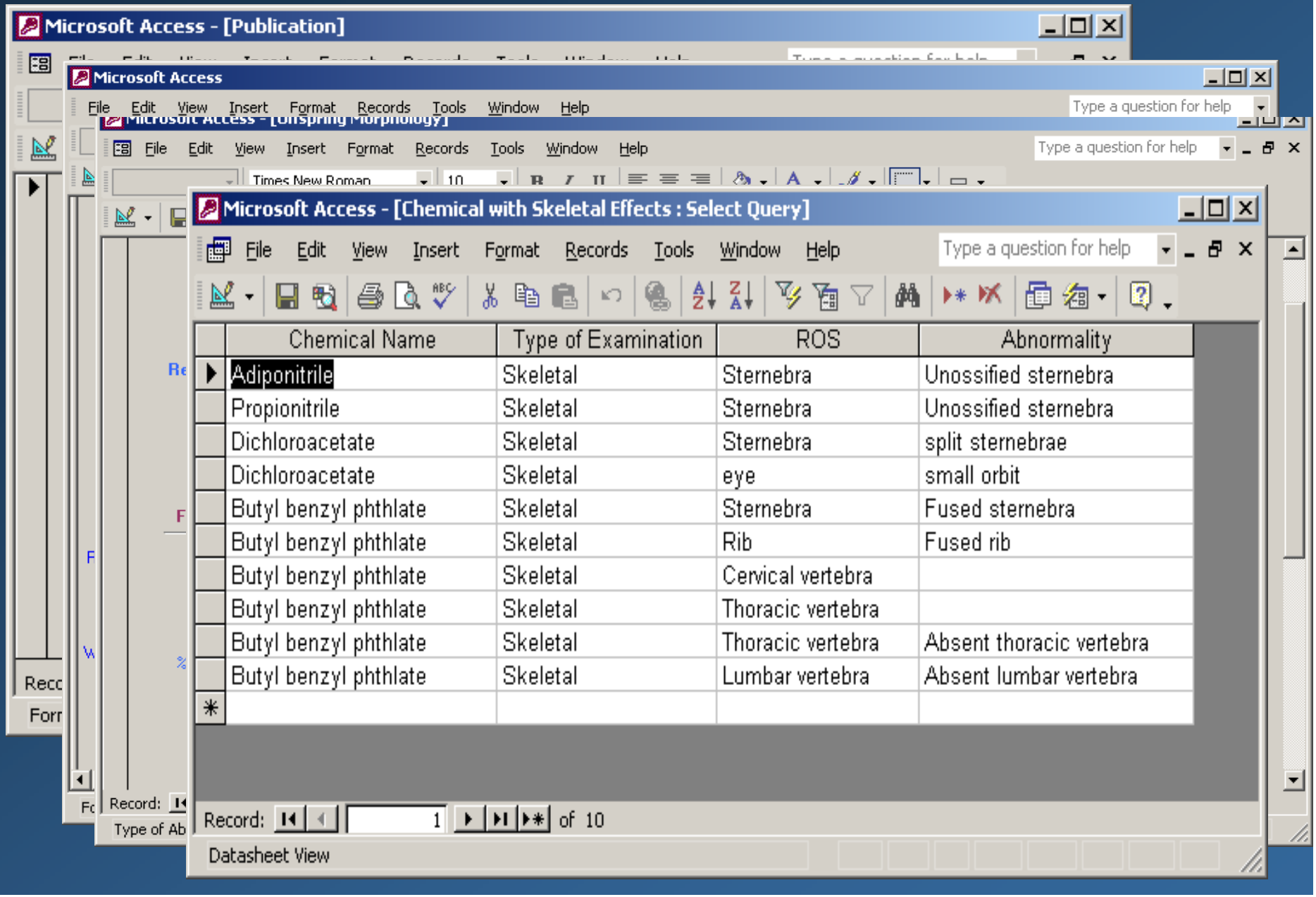
John DeSesso Mitretek Systems
Paul Foster NIEHS
Bette Meek Health Canada
Phil Mirkes Texas A & M
Ann Richard USEPA NCCT
Jennifer Seed USEPA OPPT
Dave Wise – Merck
Chihae Yang – Leadscope
Tom Collins, Mary Shackelford, Yan Gu – FDA CFSAN
Ed Matthews and Dan Benz – FDA CDER
Don Stump – WIL Research
Calvin Willhite – California EPA
Cynthia Van Landingham - ENVIRON

What is the current state of SAR Prediction methods for Developmental Toxicity?

MultiCase
DEREK
TOPKAT

Unsatisfactory

ILSI Developmental Toxicity Database Project



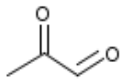
FDA Toxicity Databases

FDA SAR Genetox Database

- H
  fo
- In
- S
  - S

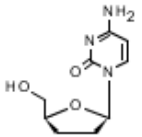
Structures | Spreadsheet | Graphs

Structure / ID

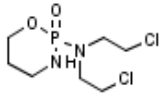


LS-139762-SAR

LS-1658-SAR



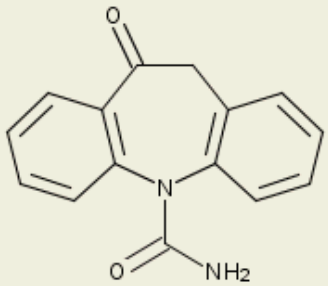
LS-1645-SAR



LS-1302-SAR

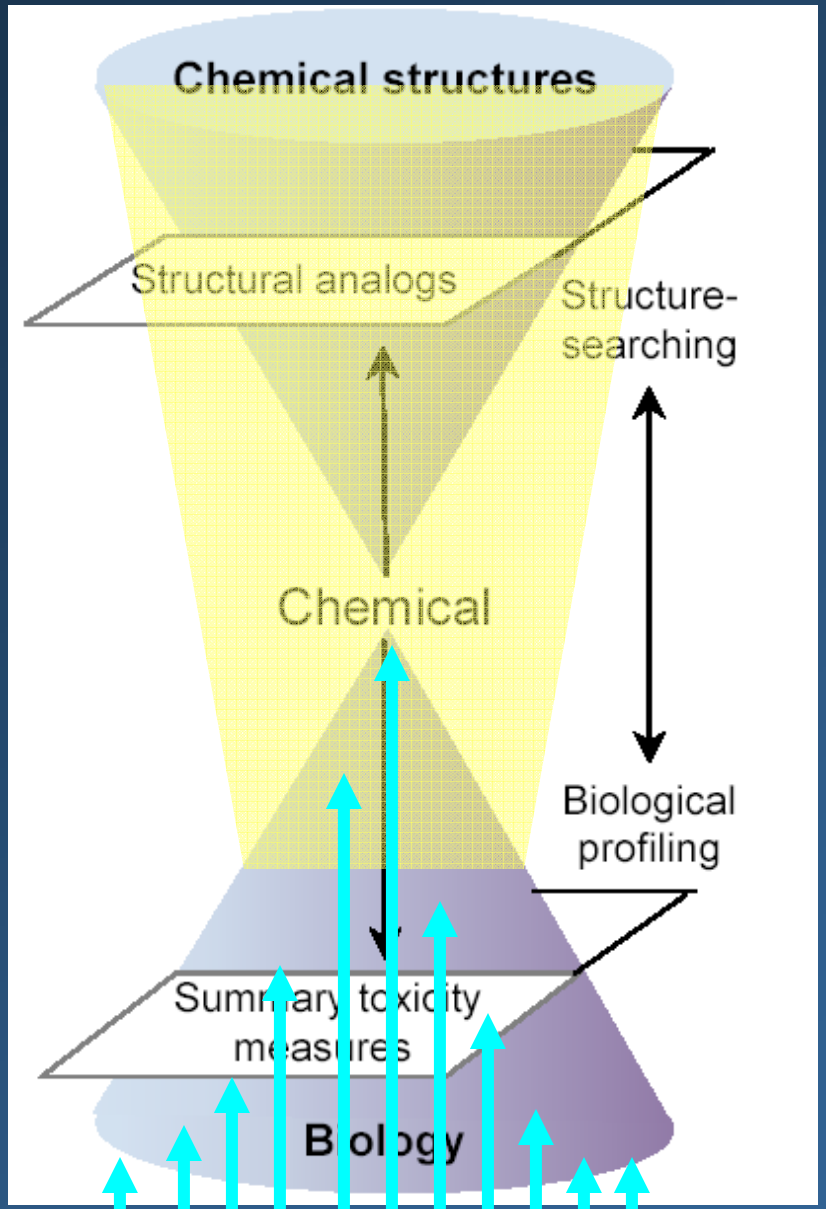
Compound Details - LS-60365-SAR

Print | Print Preview | Open Structure with Editor | Copy Structure | Find | Close



Field	Data
LS-60365-SAR	SAR Structures Created: 11-20-2006 09:16:54 Modified: 11-20-2006 09:16:54
Datasets	
Numeric Data	
LeadScope-calculated	
Molecular Weight	252.3
Parent Molecular Weight	252.3
Rotatable Bonds	1.0
Hydrogen Bond Acceptors	2.0
Hydrogen Bond Donors	1.0
Lipinski Score	0.0
ALogP	2.64
Polar Surface Area	63.4
Parent Atom Count	19.0
Leadscope	
FDA 2006 SAR Genetox Database	
Bacterial Mutation	1.0
Salmonella	1.0
TA100	1.0
TA98	0.0
TA1535	0.0
TA1537	0.0
Mammalian Mutation	0.0
ivt Chrom. Abs.	1.0
invivo Micronucleus	0.0
invivo Micronucleus Rat	0.0
Text Data	
Leadscope	
FDA 2006 SAR Genetox Database	
Bacterial Mutation	Positive
Salmonella	Positive
TA100	Positive
TA98	Negative
TA1535	Negative
TA1537	Negative
Mammalian Mutation	Negative
ivt Chrom. Abs.	Positive
invivo Micronucleus	Negative
invivo Micronucleus Rat	Negative

DSSTox Summary Toxicity Data Files



ToxML

Yang et al (2006) Landscape of current toxicity databases and database standards. *Curr Opinion Drug Discov Develop* 9(1),124-133.

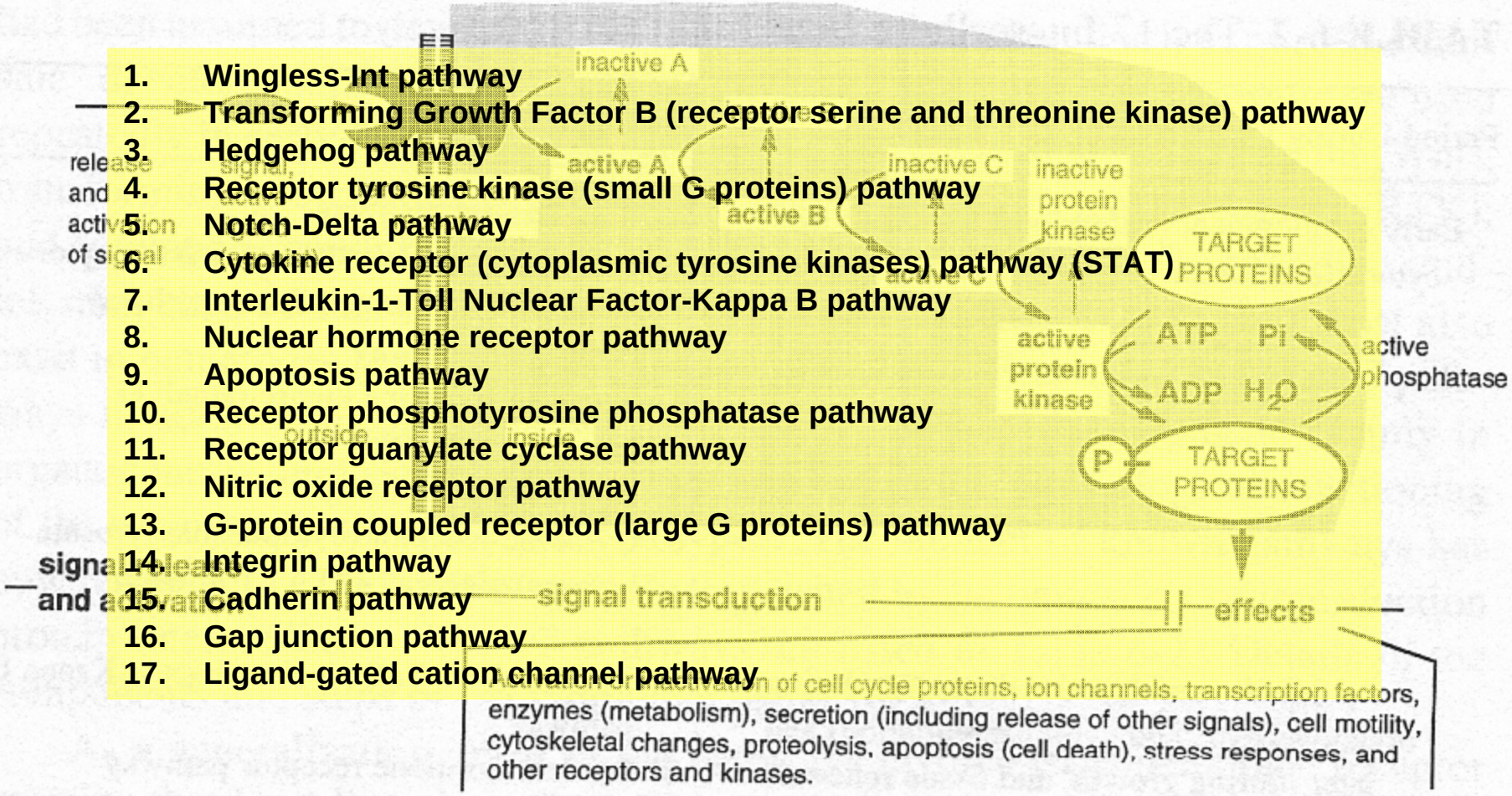
Yang et al (2006) The art of data mining the minefields of toxicity databases to link chemistry to biology. *Curr Comput-Aided Drug Design*, 2(2), 135-150.

Part IV.

Toxicity Profiling

Scientific Frontiers in Developmental Toxicology & Risk Assessment - National Academy of Sciences, 2000

1. **Wingless-Int pathway**
2. **Transforming Growth Factor B (receptor serine and threonine kinase) pathway**
3. **Hedgehog pathway**
4. **Receptor tyrosine kinase (small G proteins) pathway**
5. **Notch-Delta pathway**
6. **Cytokine receptor (cytoplasmic tyrosine kinases) pathway (STAT)**
7. **Interleukin-1-Toll Nuclear Factor-Kappa B pathway**
8. **Nuclear hormone receptor pathway**
9. **Apoptosis pathway**
10. **Receptor phosphotyrosine phosphatase pathway**
11. **Receptor guanylate cyclase pathway**
12. **Nitric oxide receptor pathway**
13. **G-protein coupled receptor (large G proteins) pathway**
14. **Integrin pathway**
15. **Cadherin pathway**
16. **Gap junction pathway**
17. **Ligand-gated cation channel pathway**



Chemical “Probes” of biological systems:

“We find that the connection between structure and biological response is not symmetric, with biological response better at predicting chemical structure than vice versa.”

D. Covell and coworkers
NCI Developmental Therapeutics Program
J Chem Inf Model (2006) 46:430-437

Predictive Toxicology

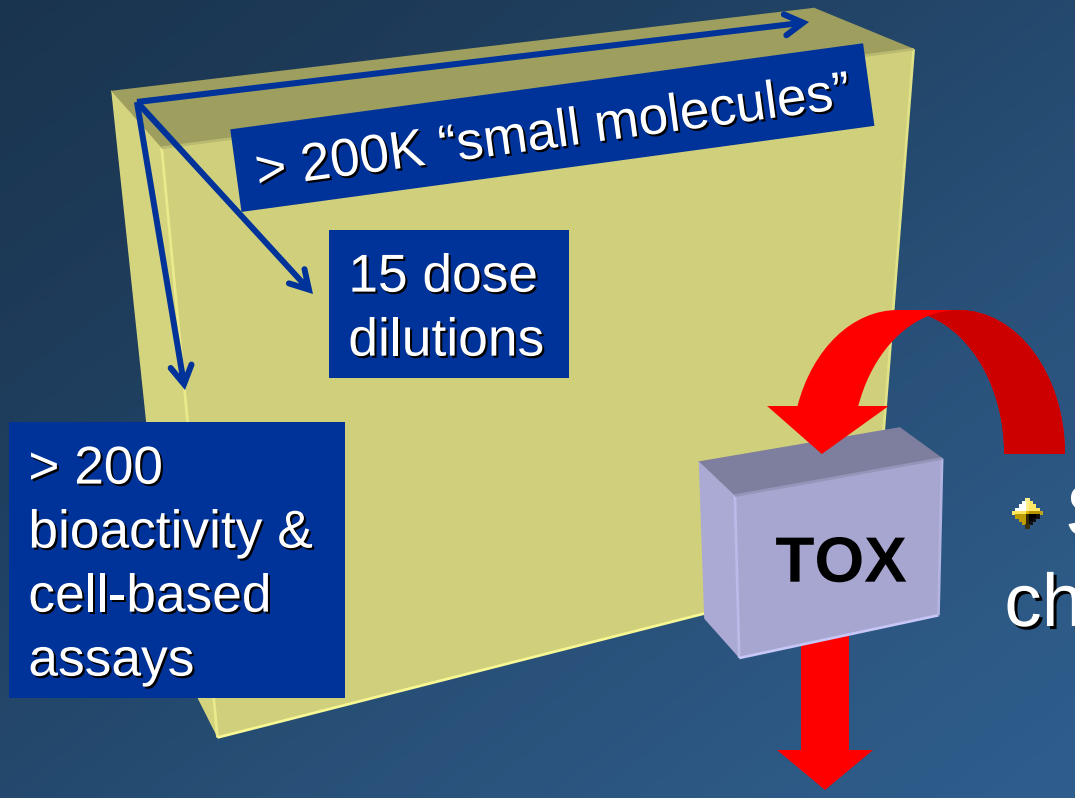
SAR Predictions based solely on chemical structures & properties

Toxicity predictions based on in vitro, HTS or gene expression profiles

NIH/NCGC Roadmap: nihroadmap.nih.gov

Small Molecules High-Throughput Screening Initiative

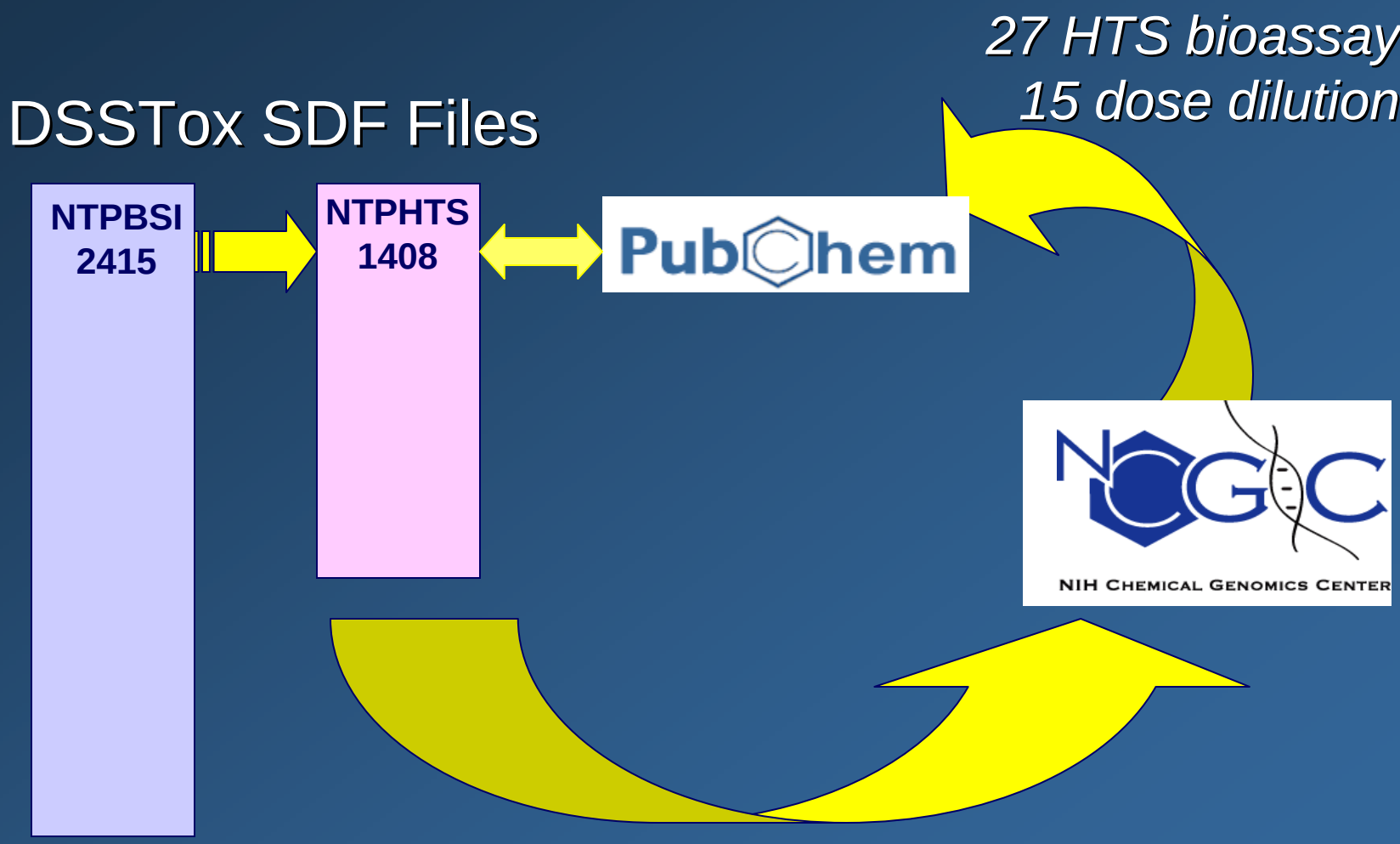
NIH Chemical Genomics Center



- Sample "toxicity" chemical space:
  - NTP chemicals
  - EPA pesticides/high interest compounds
  - DSSTox

- Reference dataset of toxicity-related chemicals with structures & bioactivity profiles

NTP High-Throughput Testing Program in Collaboration with NCGC



NCBI | PubChem | National Library of Medicine

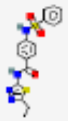
HOME | SEARCH | SITE MAP | PubMed | Entrez | Structure | GenBank | PubChem | Help

BioAssay Results

BioAssay ID (AID): 360
Source: NCGC
Name: Glucocerebrosidase

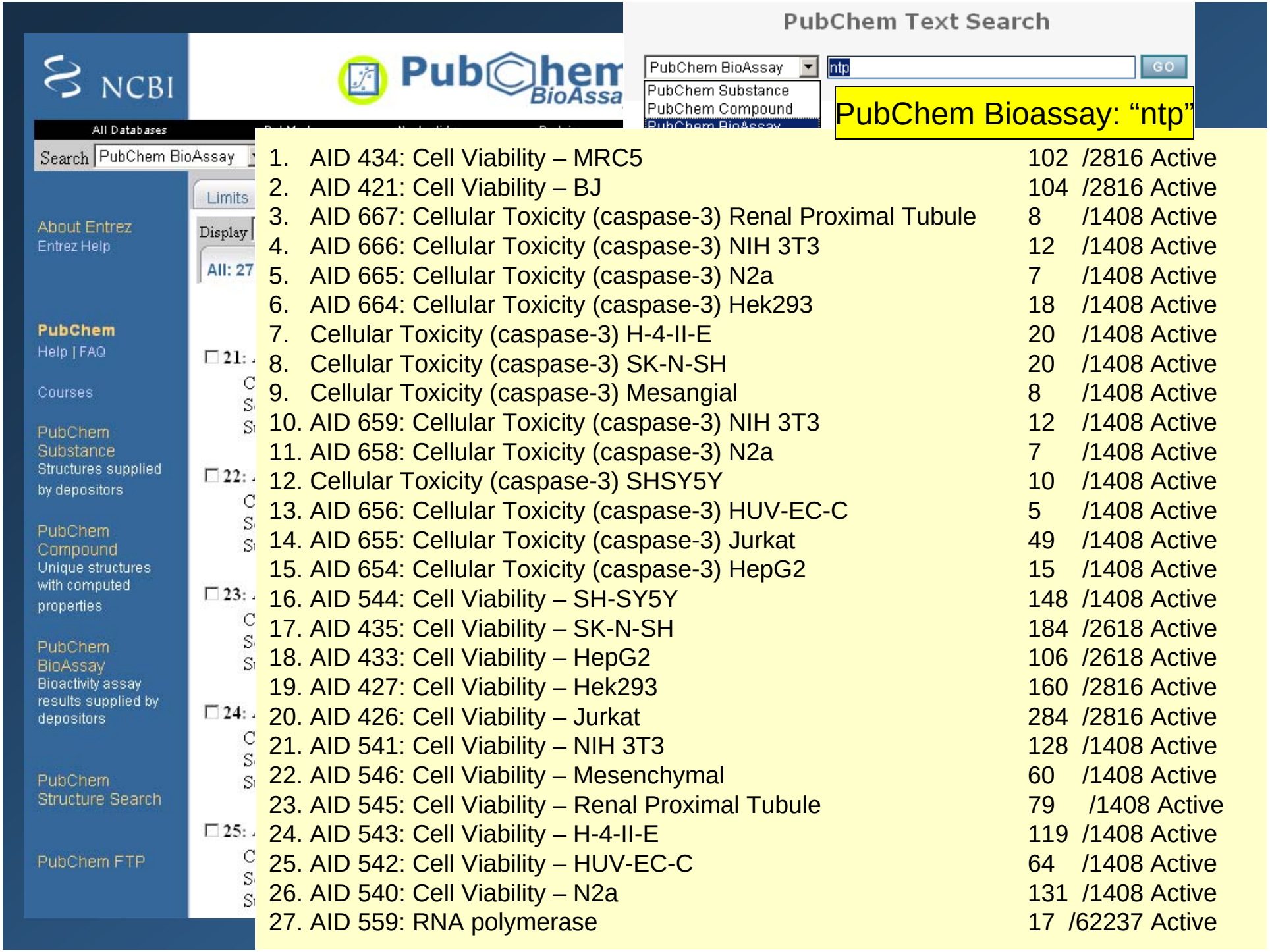
Total 48125 compounds found (48125 unique), 20 displayed: Next page

AC50, Hill Coefficient

Structure	PubChem SID	PubChem CID	Outcome	Activity Score	Submitter	Submission Date	Activity Direction	Activity Qualifier	Qualified AC50	Log of AC50	Hill Coefficient	Curve R2	Data Type	Compound Type
	4243169	3237927	Active	72	ncgc	19 Jan 2006	decreasing	=	6.06e-008	-7.22	0.87	1	qHTS	NIHSMR
	4264637	2210290	Active	71	ncgc	19 Jan 2006	decreasing	=	7e-008	-7.16	0.66	1	qHTS	NIHSMR

Full Concentration Response Curve

Compound QC	Curve Fit Model	Hill S0	Hill Sinf	Hill dS	Log AC50 Std Error	Curve Chi2f	Excluded Points	Number of Points	Activity at 4.925nM (%)	Activity at 24.623nM (%)	Activity at 0.123uM (%)	Activity at 0.615uM (%)	Activity at 3.077uM (%)	Activity at 15.386uM (%)	Activity at 0.077mM (%)
QC'd by DPI	4pHill (AC50,n,S0,Sinf)	-1.23	-100.1	98.92	0.02	0.5	{}	7	-11.3	-31.9	-65.8	-88.4	-96.4	-99.6	-100.2
QC'd by DPI	4pHill (AC50,n,S0,Sinf)	11.41	-107.2	118.6	0.1	2.04	{}	7	-5.5	-30.2	-56.9	-84.4	-100	-103.7	-105



1.	AID 434: Cell Viability – MRC5	102 /2816 Active
2.	AID 421: Cell Viability – BJ	104 /2816 Active
3.	AID 667: Cellular Toxicity (caspase-3) Renal Proximal Tubule	8 /1408 Active
4.	AID 666: Cellular Toxicity (caspase-3) NIH 3T3	12 /1408 Active
5.	AID 665: Cellular Toxicity (caspase-3) N2a	7 /1408 Active
6.	AID 664: Cellular Toxicity (caspase-3) Hek293	18 /1408 Active
7.	Cellular Toxicity (caspase-3) H-4-II-E	20 /1408 Active
8.	Cellular Toxicity (caspase-3) SK-N-SH	20 /1408 Active
9.	Cellular Toxicity (caspase-3) Mesangial	8 /1408 Active
10.	AID 659: Cellular Toxicity (caspase-3) NIH 3T3	12 /1408 Active
11.	AID 658: Cellular Toxicity (caspase-3) N2a	7 /1408 Active
12.	Cellular Toxicity (caspase-3) SHSY5Y	10 /1408 Active
13.	AID 656: Cellular Toxicity (caspase-3) HUV-EC-C	5 /1408 Active
14.	AID 655: Cellular Toxicity (caspase-3) Jurkat	49 /1408 Active
15.	AID 654: Cellular Toxicity (caspase-3) HepG2	15 /1408 Active
16.	AID 544: Cell Viability – SH-SY5Y	148 /1408 Active
17.	AID 435: Cell Viability – SK-N-SH	184 /2618 Active
18.	AID 433: Cell Viability – HepG2	106 /2618 Active
19.	AID 427: Cell Viability – Hek293	160 /2816 Active
20.	AID 426: Cell Viability – Jurkat	284 /2816 Active
21.	AID 541: Cell Viability – NIH 3T3	128 /1408 Active
22.	AID 546: Cell Viability – Mesenchymal	60 /1408 Active
23.	AID 545: Cell Viability – Renal Proximal Tubule	79 /1408 Active
24.	AID 543: Cell Viability – H-4-II-E	119 /1408 Active
25.	AID 542: Cell Viability – HUV-EC-C	64 /1408 Active
26.	AID 540: Cell Viability – N2a	131 /1408 Active
27.	AID 559: RNA polymerase	17 /62237 Active

kNN Consensus QSAR Modeling of NTP-HTS Data

Hao Zhu, Ivan Rusyn, and Alexander Tropsha
School of Pharmacy & School of Public Health, UNC-Chapel Hill

NTP HTS 1408

Carcinogenic Potency Database 1481

270

300+ Chemical Descriptors Only
(9 models > 0.7 validation cutoff)

300+ Chemical + 6 HTS "Descriptors"
(34 models > 0.7 validation cutoff)

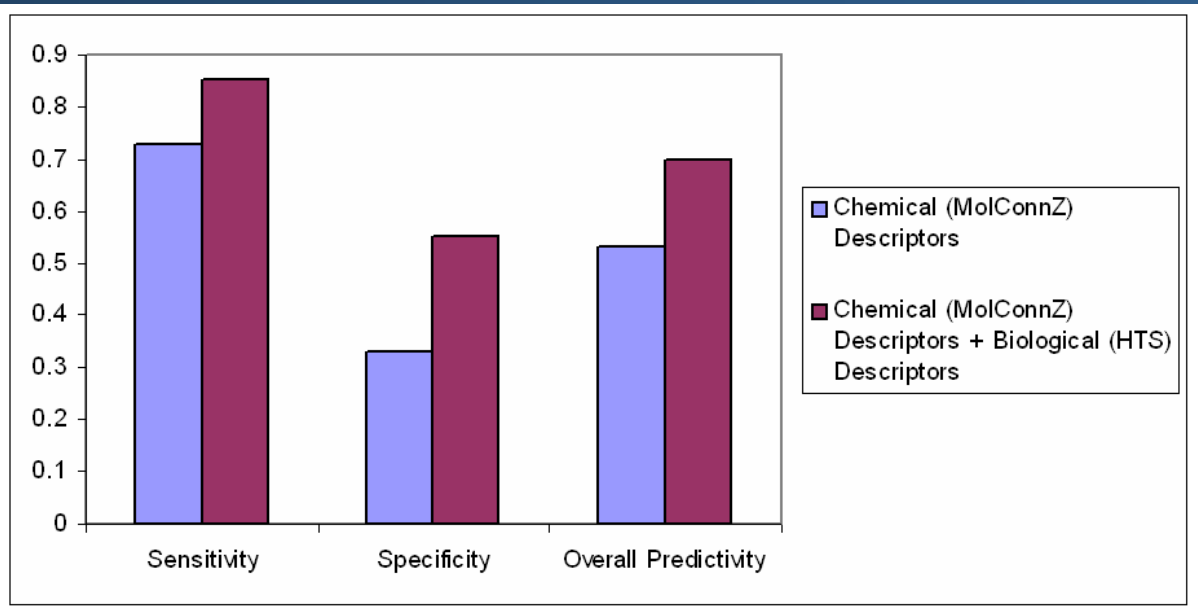
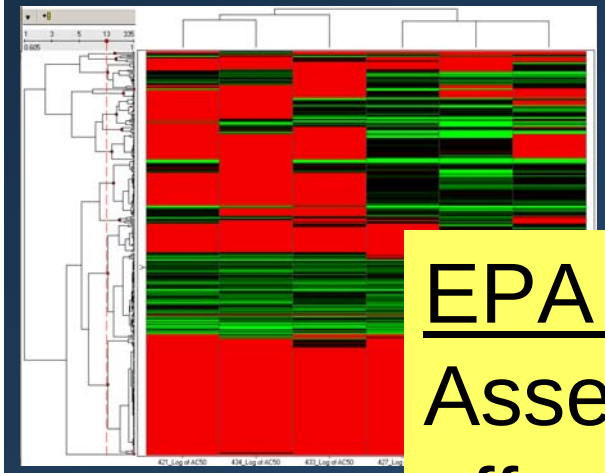
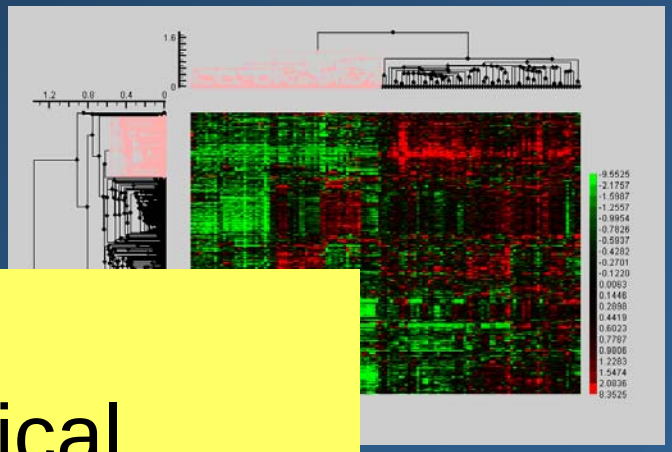


Figure 3. Comparison of the results from *k*NN QSAR models using two types of descriptors.

Correlating Domain Outputs

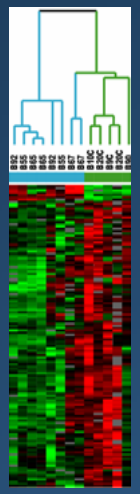


CellularAss

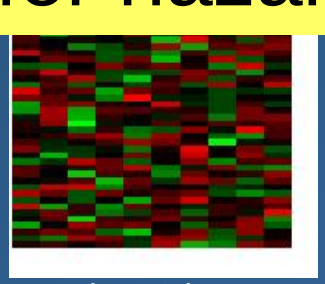


cal properties

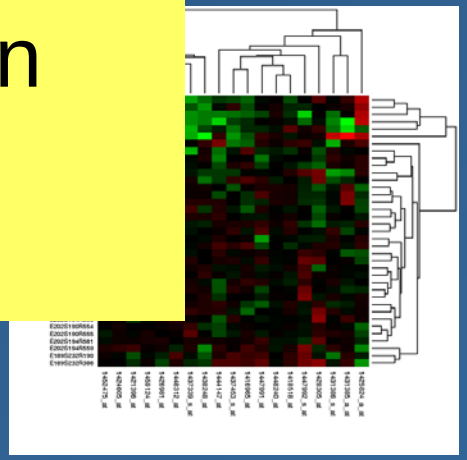
EPA ToxCast Goal:
Assess broad biological effect patterns and correlate with the patterns of known toxicants for forecasting potential for hazard



Biochemical Assays



Genomic Signatures

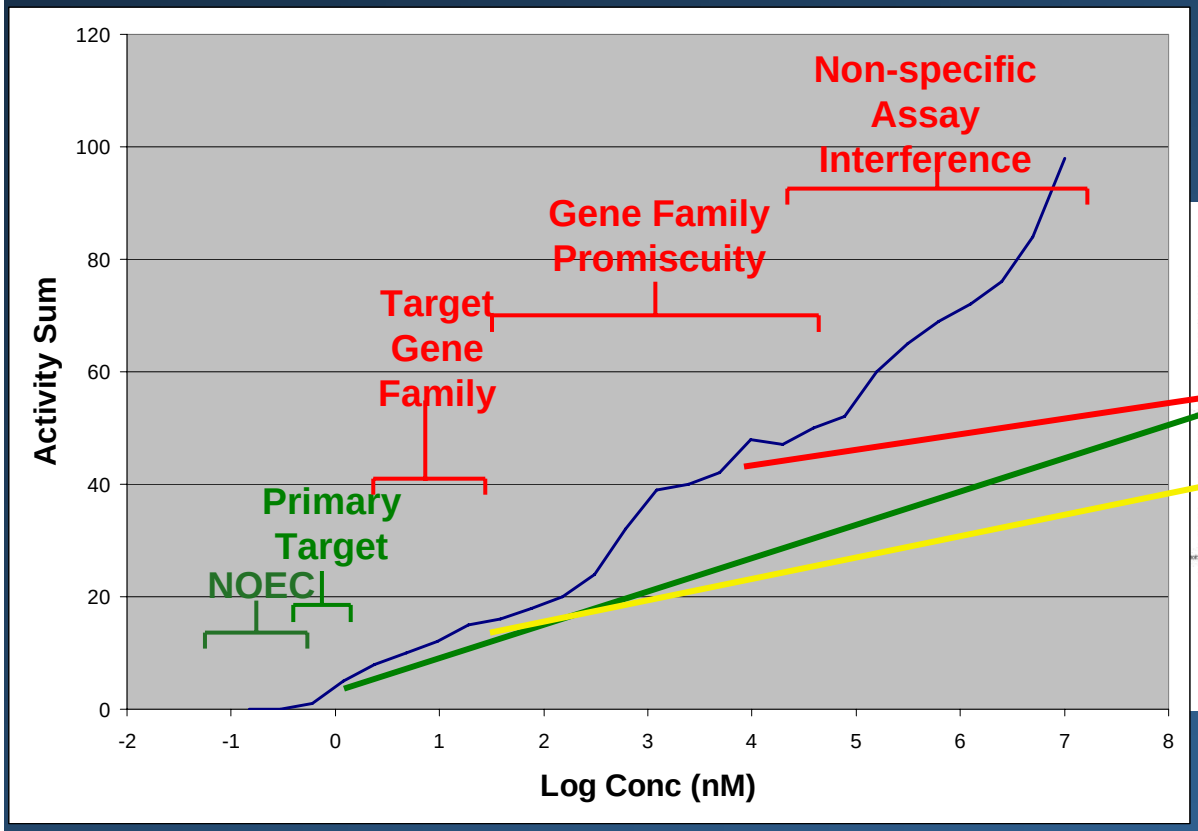


Toxicology Endpoints

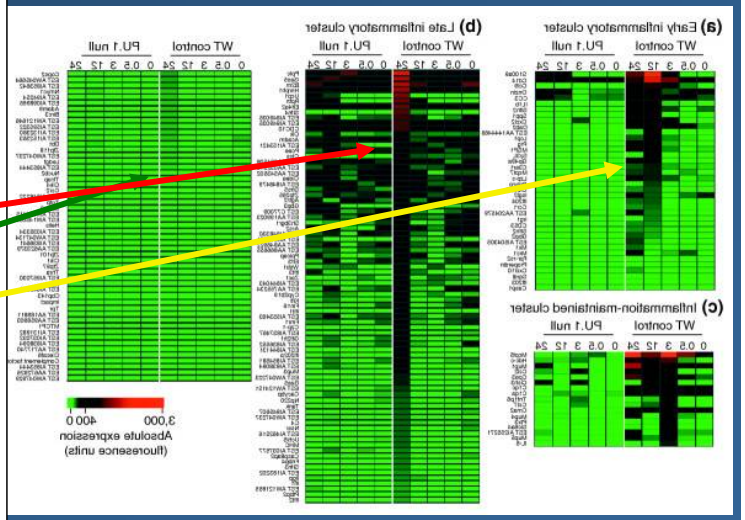
EPA ToxCast Program

- HTS screening of 320 pesticide actives and other compounds in hundreds of bioassays; 1408 total compounds for NCGC screening
- ToxRefDB - High quality reference data for registered pesticides entered into toxicity data model
- DSSTox structure annotation, overlaps with toxicity databases, chemical selection criteria
- ACToR integrated data warehouse & analysis system to support prediction modeling efforts

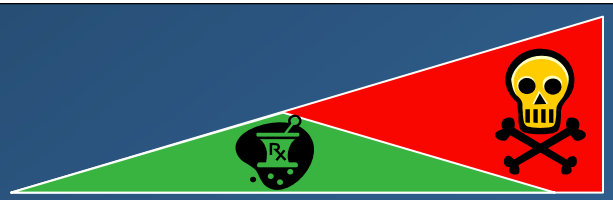
Bioactivity Profiling of Pharmaceuticals vs Environmental Chemicals



HTS Target Assay Panels

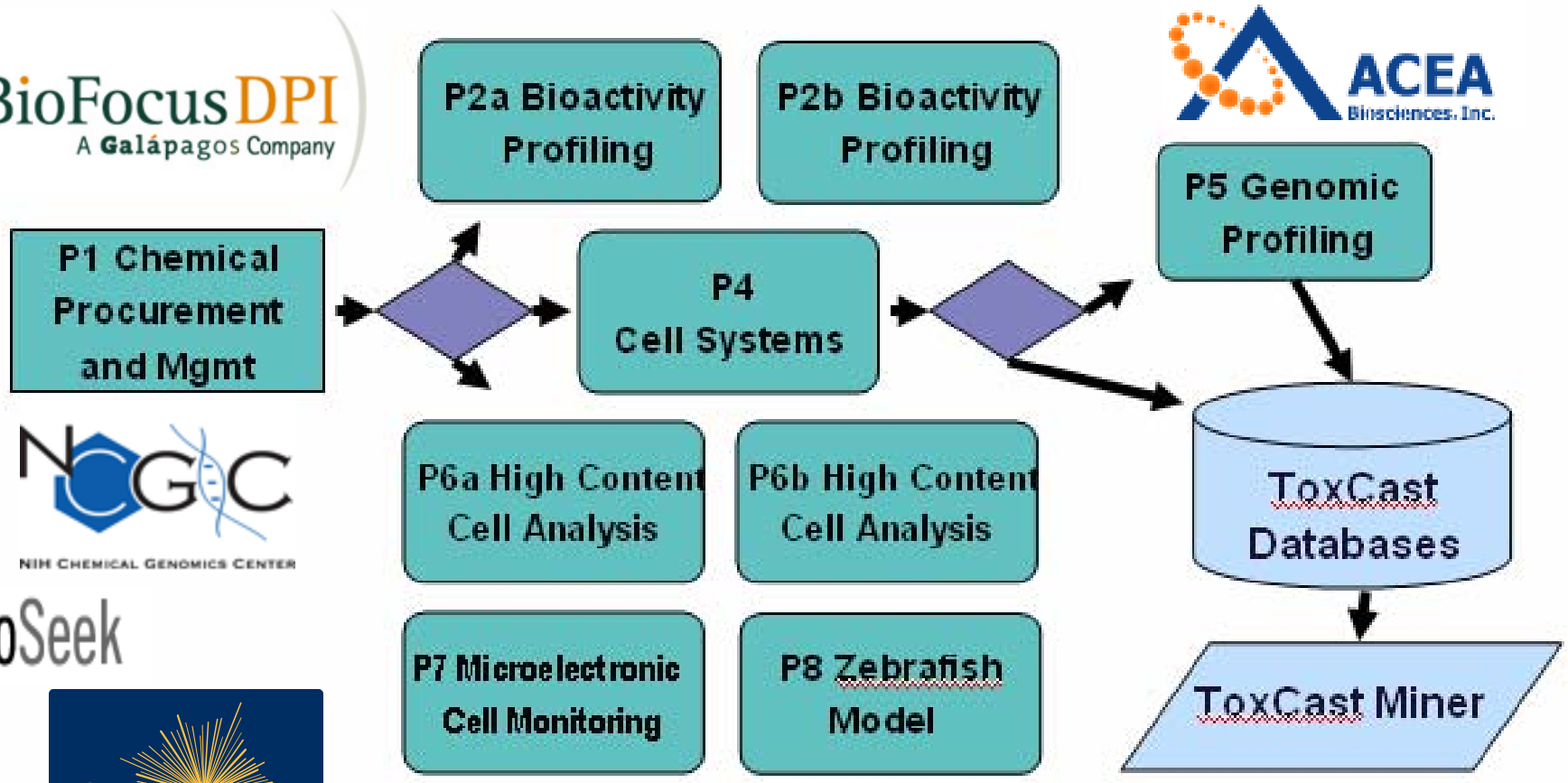


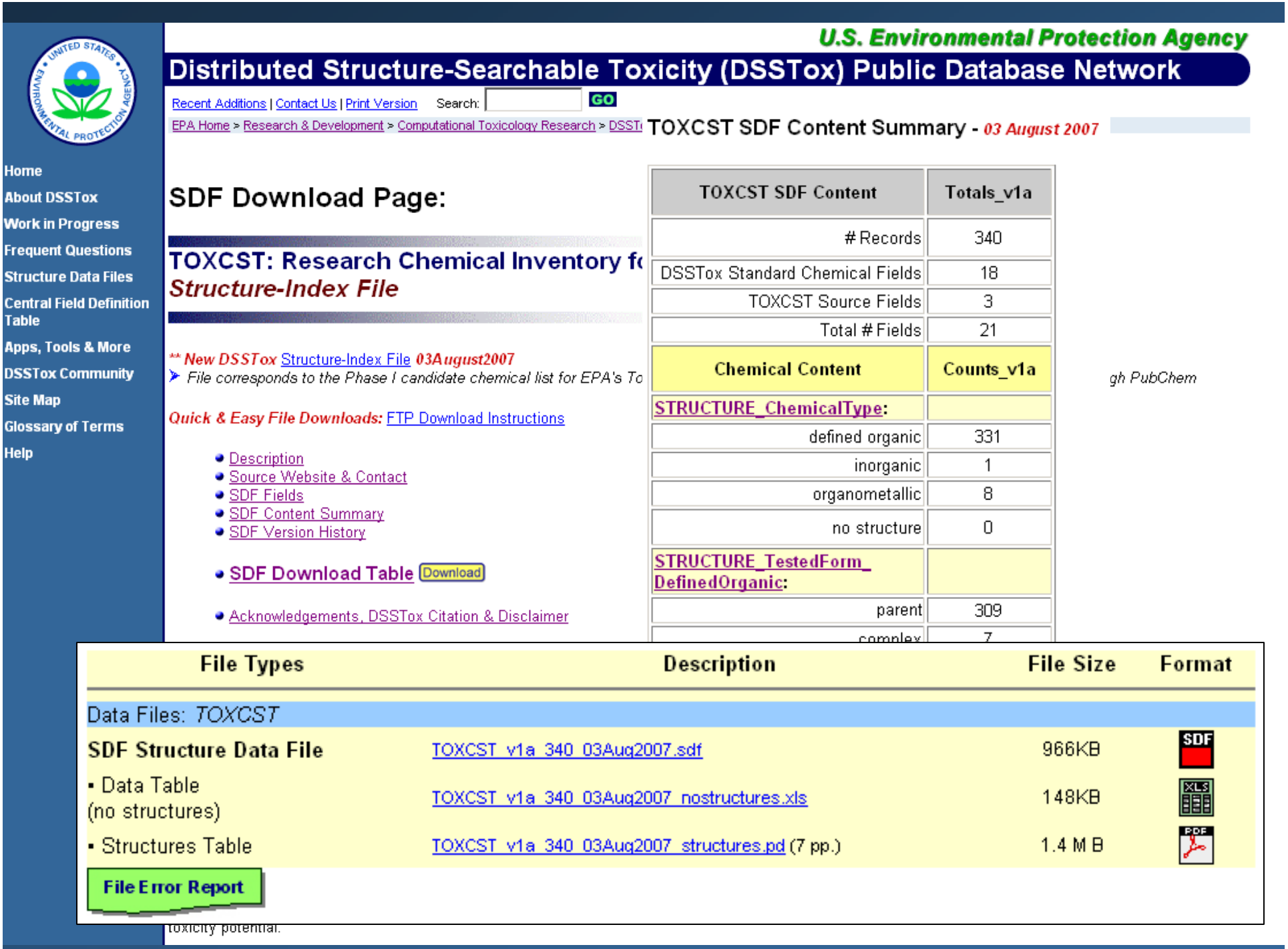
Heat map profiles of activity in different assays (columns) arranged by compound (rows). Inactive (green) to highly active (red).



Slide courtesy of Keith Houck, EPA NCCT

ToxCast Contracts for Data Generation





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 > Computational Toxicology Research > DSSTo

TOXCST SDF Content Summary - 03 August 2007

Home
About DSSTox
Work in Progress
Frequent Questions
Structure Data Files
Central Field Definition Table
Apps, Tools & More
DSSTox Community
Site Map
Glossary of Terms
Help

SDF Download Page:

TOXCST: Research Chemical Inventory fo
Structure-Index File

** New DSSTox Structure-Index File 03August2007
File corresponds to the Phase I candidate chemical list for EPA's To

Quick & Easy File Downloads: FTP Download Instructions

- Description
- Source Website & Contact
- SDF Fields
- SDF Content Summary
- SDF Version History
- SDF Download Table Download
- Acknowledgements, DSSTox Citation & Disclaimer

TOXCST SDF Content	Totals_v1a
# Records	340
DSSTox Standard Chemical Fields	18
TOXCST Source Fields	3
Total # Fields	21
Chemical Content	**Counts_v1a**
STRUCTURE_ChemicalType:	
defined organic	331
inorganic	1
organometallic	8
no structure	0
STRUCTURE_TestedForm_DefinedOrganic:	
parent	309
complex	7

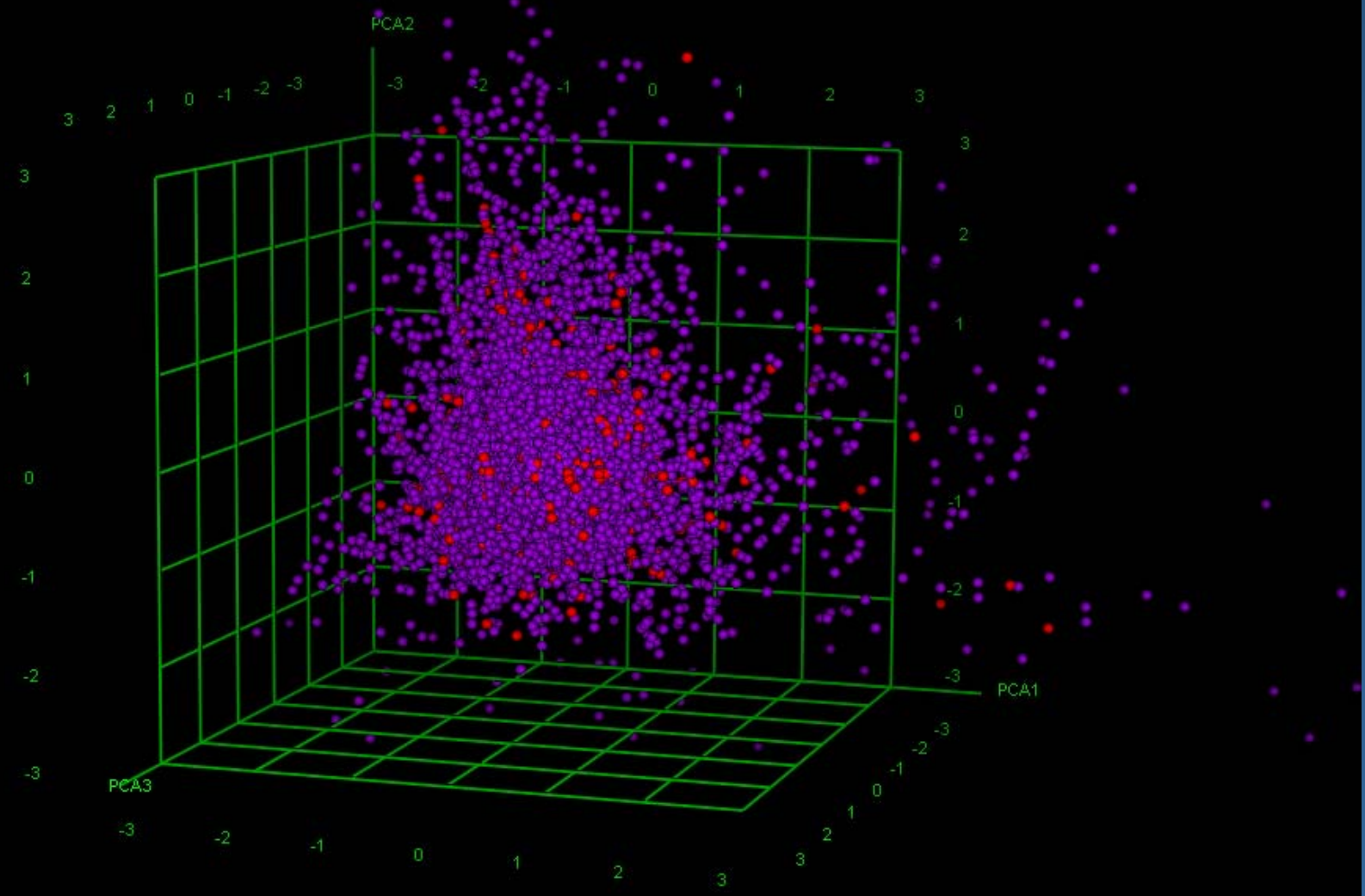
gh PubChem

File Types	Description	File Size	Format
Data Files: *TOXCST*			
SDF Structure Data File	TOXCST_v1a_340_03Aug2007.sdf	966KB	SDF
▪ Data Table (no structures)	TOXCST_v1a_340_03Aug2007_nostructures.xls	148KB	XLS
▪ Structures Table	TOXCST_v1a_340_03Aug2007_structures.pd (7 pp.)	1.4 MB	PDF

File Error Report

toxicity potential.

- DSSTox Master File (>7400 structures)
- EPA Pesticidal Actives (>800 structures)



MOE – 20 descriptors (physchem prop, atom counts);
3 PCA's account for 2/3s of variance

Slide courtesy of Rocky Goldsmith, EPA NCCT

EPA ToxCast
ToxRefDB - Toxicity Reference Database:

Source:	EPA's Office of Pesticide Programs (OPP)
Format:	Data Evaluation Record (DER)
Chemical:	Conventional Pesticide Active Ingredients (~800)
Purity:	Technical Grade (>90%)
Dosing:	Primarily Orally Administered (based on availability and use pattern of pesticide)
Study Type:	Subchronic Toxicity (Rodents and Non-Rodents)
	Prenatal Developmental Toxicity
	Reproduction and Fertility Effects (2-generation)
	Chronic Toxicity (rat, mouse, and dog)
	Carcinogenicity (rat and mouse)
	Developmental Neurotoxicity
	Immunotoxicity

Targeted Toxicological Data Collection

Data Collection Results

> 4000 DER (2500 studies)

for over 400 pesticides

Toxicological Schema and Lexicon Development

- ToxML compatibility and interoperability
- Standardized fields and vocabulary
  - Study Type (OPPTS/OECD Test Guidelines)
  - Data Usability (Data Quality) Code
  - Animal Info (Species, Strain, Sex Category)
  - Treatment Group Category (Adult, Offspring, etc.)
  - Endpoint Category (Systemic, Maternal, etc.)
  - Effect Descriptors (Type, Target, and Description Vocabulary)

Courtesy of Matt Martin, EPA NCCT

DATA EVALUATION RECORD

STUDY TYPE: Combined Chronic/Oncogenicity Study in Rats
OPPTS Number: 870.4300 OPP Guideline Number: §83-5

TXR#: 0050178
DP BARCODE: D264893 SUBMISSION CODE: S575895
P.C. CODE: 000586 TOX. CHEM. NO.: None

TEST MATERIAL (PURITY): Bifenazate (90.2-92.5% a.i.)

SYNONYMS: D2341; UCC-D2341 Technical

CITATION: Ivett, J.L. (1999) 104-Week Combined Chronic Dietary Toxicity and Oncogenicity Study in Rats with D2341. Covance Laboratories, Inc., Vienna, VA. Laboratory Study ID No. 798-229, March 19, 1999. MRID 45076504. Unpublished.

SPONSOR: Uniroyal Chemical Company, Inc., 74 Amity Road, Bethany, Connecticut

EXECUTIVE SUMMARY: In a rat combined chronic/oncogenicity study (MRID 45076504), bifenazate (90.2-92.5% a.i., Lot/Batch # PP159945) was administered in the diet to Sprague-Dawley Crl:CD®BR rats (50/sex/group) for up to 104 weeks at nominal doses of 0, 20, 80, or 200 ppm (males) and 160 ppm (females) (equivalent to 0/0, 1.0/1.2, 3.9/4.8, and 9.7/9.7 mg/kg/day [M/F], respectively). An additional 10 rats/sex/dose were sacrificed at 53 weeks.

Mortality, clinical signs, clinical chemistry, urinalysis, organ weights, and macroscopic findings for both sexes at all doses were unaffected by treatment with bifenazate. No treatment-related effects were observed in any parameter in the 20 ppm dose group.

Only minor and/or transient toxicologically significant differences from controls were detected after treatment with bifenazate at 80 or 200/160 ppm (M/F) rats for 104 weeks.

BIFENAZATE **Combined Chronic/Oncogenicity (§83-5)**

During the first 52 to 77 weeks of the study in both sexes at the high-dose, reductions ($p \leq 0.05$) in mean body weights (↓3-9%), mean body weight gains (↓6-17%), and mean total food consumption (↓3-7%) were observed. The decreases in overall (weeks 1-104) body weight gain (↓9%), overall total food consumption (↓4%), and overall food efficiency (↓22%) observed in the high-dose females were not statistically significant. In the high-dose females, decreased ($p \leq 0.05$) erythrocyte, hemoglobin, and hematocrit (↓5-10%) values were detected at weeks 13, 26, and 52; These values were within historical control values. The decreases in these parameters observed in the high-dose females at weeks 78 and/or 105 were not statistically significant (↓6-9%).

Chronic toxicity was characterized in the high-dose animals only as relatively minor decreases ($p \leq 0.05$) in body weight, body weight gain, and food consumption during the first 52 to 77 weeks of the study with no statistically significant differences in food efficiency. There was only a 5-10% difference from concurrent controls in hematological parameters in the high-dose females (the values were all within historical controls) and the microscopic finding of increased severity of pigment in the spleen observed in both sexes was transient, occurring only in the 9-10 rats/sex/dose sacrificed at 53 weeks. In addition, the lens opacity finding is equivocal.

The chronic LOAEL was 200 ppm for males and 160 ppm for females (equivalent to 9.7 mg/kg/day, each) based on decreases ($p \leq 0.05$) in body weight, body weight gain, food consumption and hematological parameters. The chronic NOAEL is 80 ppm for males and females (equivalent to 3.9 mg/kg/day for males and 4.8 mg/kg/day for females).

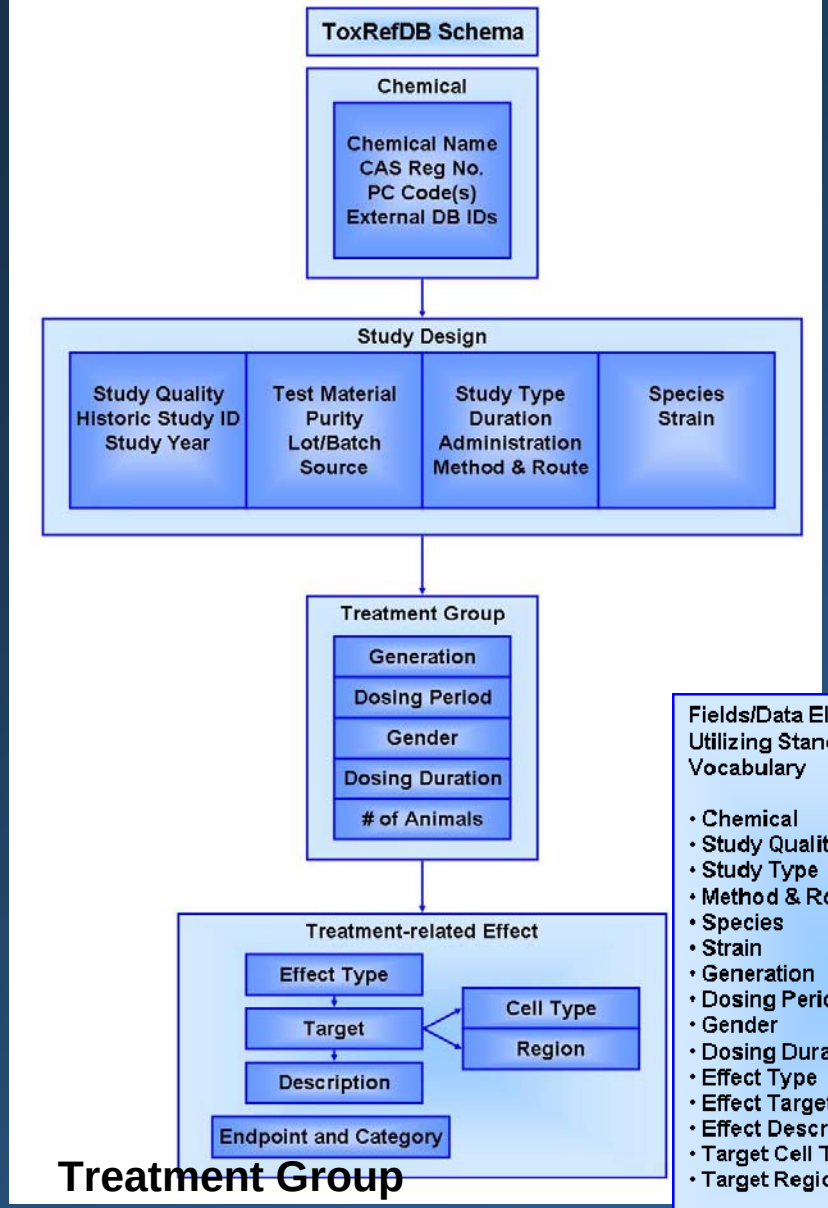
No treatment-related neoplastic changes were observed and although it appears that the animals could have tolerated a higher dose, dose selection in the 2-year study was carefully based on the significant results at 200 and 400 ppm in the 90-day feeding study in rats. Therefore the carcinogenic potential of the test substance was considered to be negative at acceptable dose levels.

The submitted study is classified as **Acceptable\ Guideline** (§83-5) and does satisfy the guideline requirements for a chronic toxicity study (§83-1) and a carcinogenicity study (§83-2) in rats.

COMPLIANCE: Signed and dated GLP, Quality Assurance, Data Confidentiality, and Flagging statements were provided.

Courtesy of Matt Martin, EPA NCCT

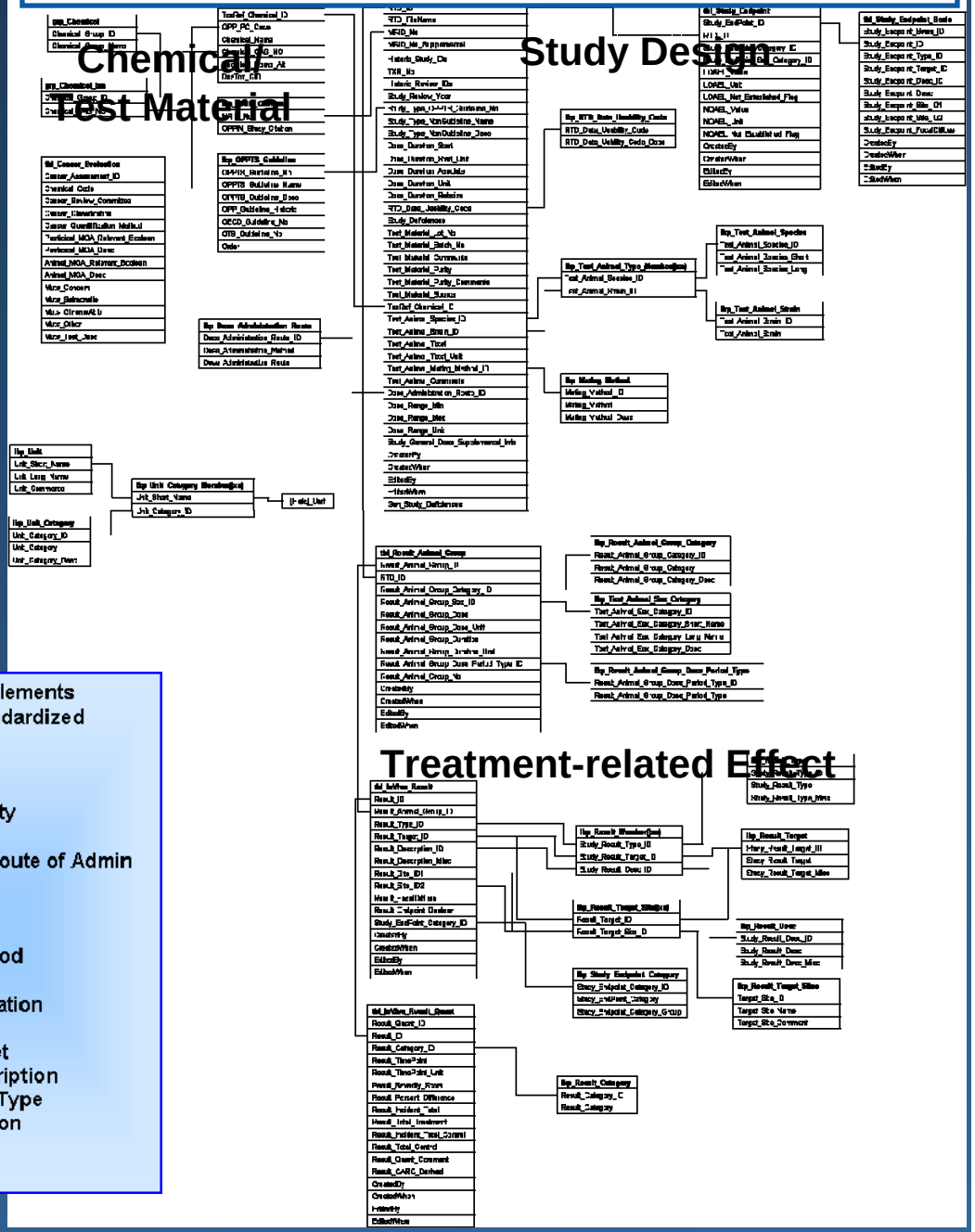
ToxRefDB Effects Relationship Diagram



Treatment Group

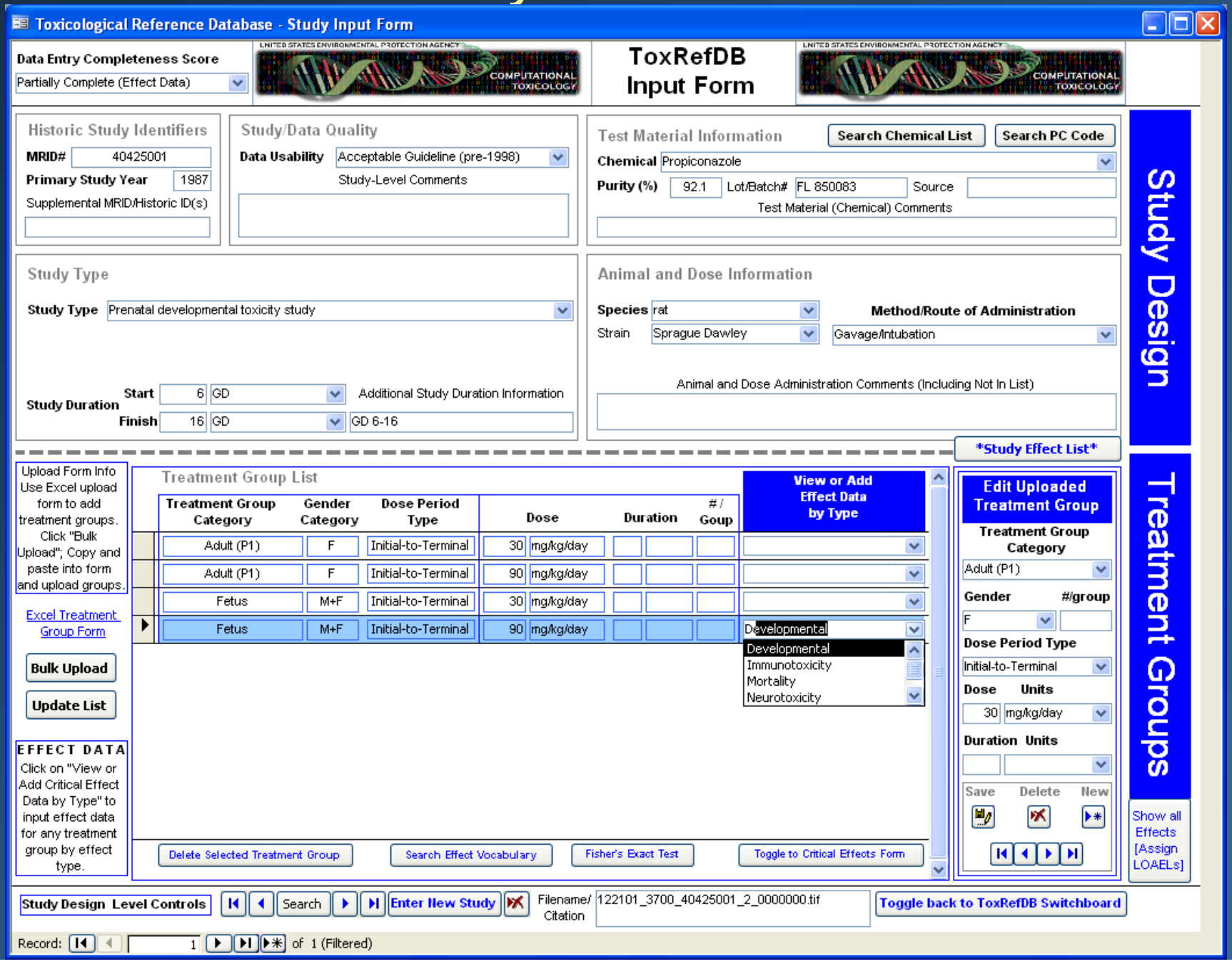
Fields/Data Elements Utilizing Standardized Vocabulary

- Chemical
- Study Quality
- Study Type
- Method & Route of Admin
- Species
- Strain
- Generation
- Dosing Period
- Gender
- Dosing Duration
- Effect Type
- Effect Target
- Effect Description
- Target Cell Type
- Target Region



Courtesy of Matt Martin, EPA NCCT

EPA ToxCast Toxicity Reference Database



ToxRefDB Data Entry Status

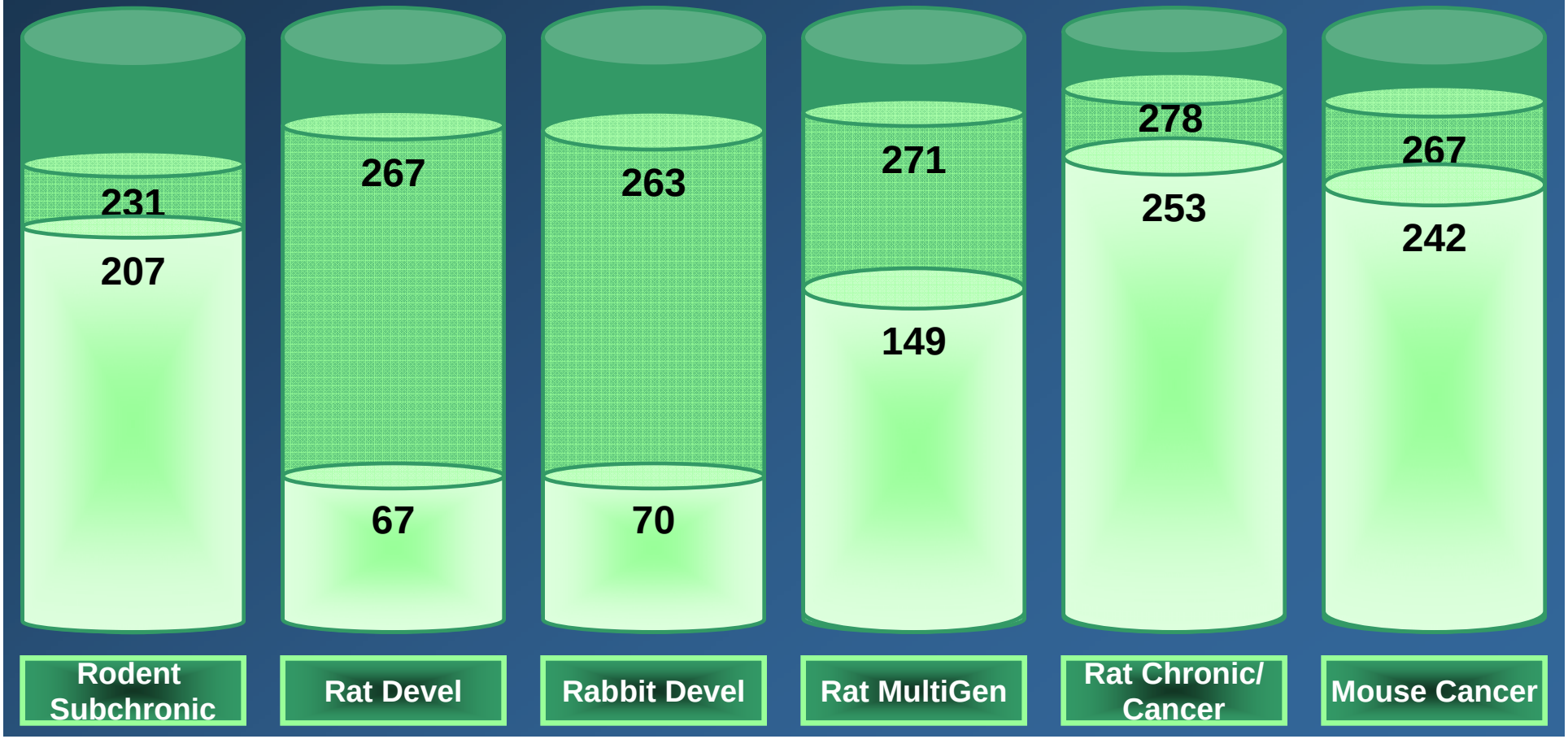
ToxCast Chemicals	320
Unique Chemicals	307
Pesticide Actives	291

Current as of October 12, 2007

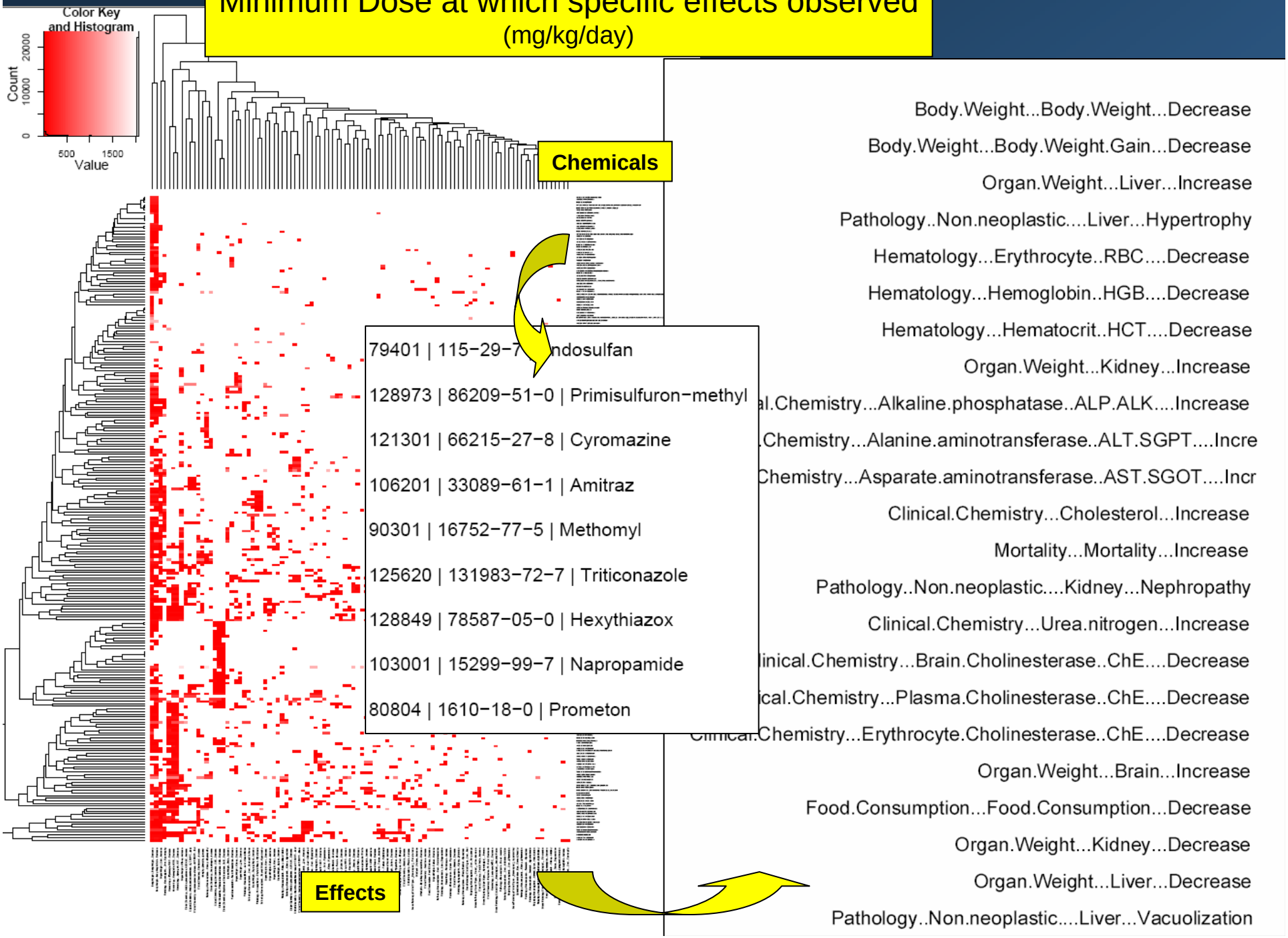
LEGEND

- Total ToxCast Unique Chemicals (307)
- # of Chemicals with Data Coverage
- # of Chemicals with Complete Data Entry

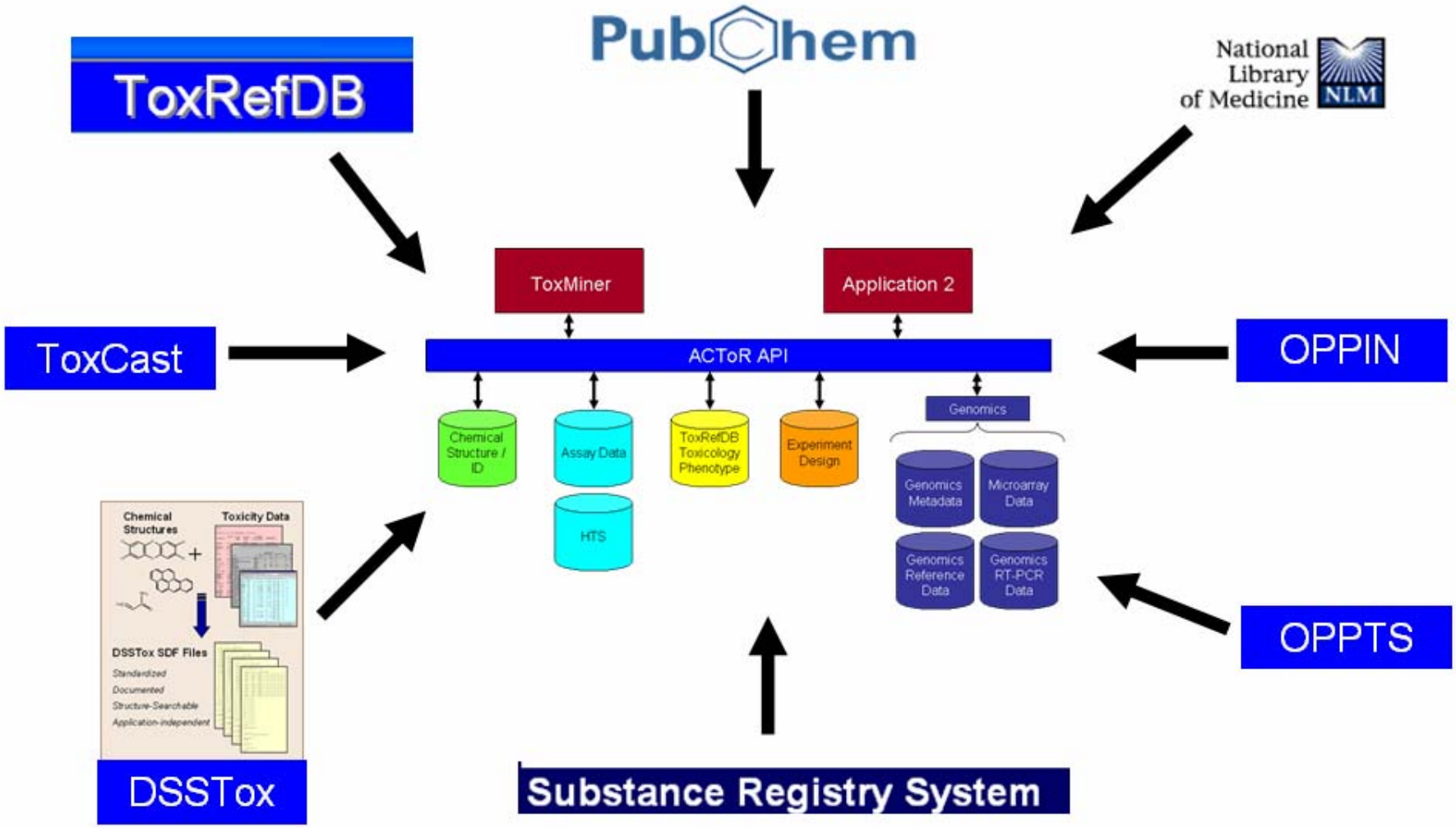
Study Type



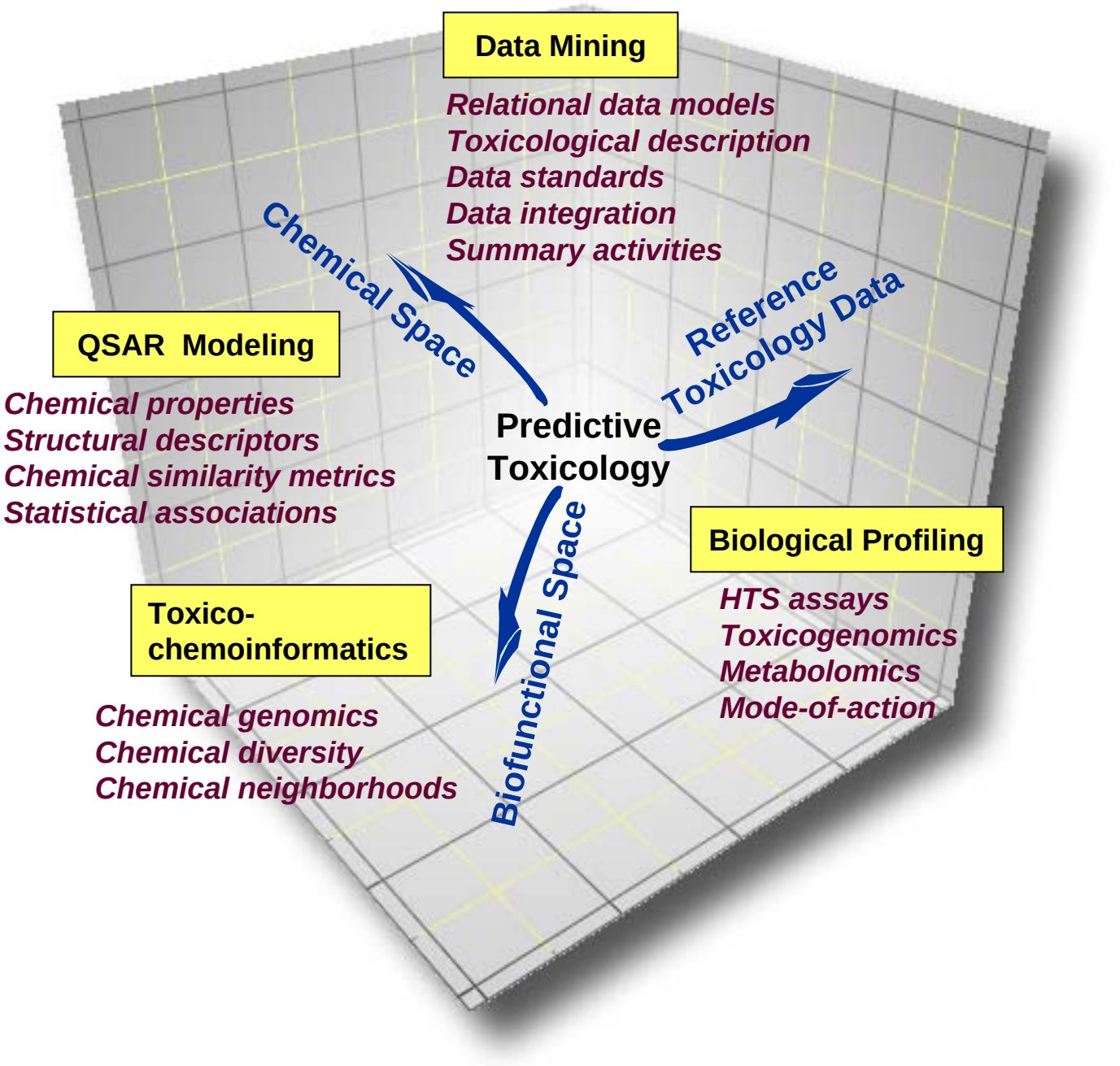
Minimum Dose at which specific effects observed
(mg/kg/day)



ACToR Draws from Diverse Data Sources



Courtesy of Richard Judson, EPA National Center for Computational Toxicology


Data Mining
Relational data models
Toxicological description
Data standards
Data integration
Summary activities
Chemical Space
QSAR Modeling
Chemical properties
Structural descriptors
Chemical similarity metrics
Statistical associations
Reference Toxicology Data
Predictive Toxicology
Biofunctional Space
Biological Profiling
HTS assays
Toxicogenomics
Metabolomics
Mode-of-action
Toxico-chemoinformatics
Chemical genomics
Chemical diversity
Chemical neighborhoods

Acknowledgements:

- **EPA DSSTox Team:**
  Maritja Wolf (DSSTox) and Tom Transue (Structure-browser) – Lockheed Martin
  ClarLynda Williams, NCSU Bioinformatics Graduate Student
- **NIEHS/NTP – NTP HTS Program:**
  Ray Tice, Cynthia Smith, Beth Bowden, Brad Collins
- **NCCT ToxCast Team:**
  Robert Kavlock (Director, NCCT)
  David Dix (ToxRefDB, HTS, Genomics)
  Keith Houck (HTS)
  Matt Martin (ToxRefDB)
  Richard Judson (ACToR)
  Rocky Goldsmith (Modeling)
- **ToxML:** Chihae Yang – Leadscope
- **UNC-CH School of Pharmacy:** Alex Tropsha, Hao Zhu, Ivan Rusyn
- **PubChem Project:** Steve Bryant, Jane Tseng, Yanli Wang
- **Carcinogenic Potency Project:** Lois Gold
- **NIH Molecular Libraries Initiative:** Chris Austin, Ajit Jadhav (NIH/NCGC)