ACToR: Aggregated Computational Toxicology Resource

Richard Judson, Ann Richard, Tommy Cathey, Tom Transue,
James Vail, Doris Smith, David Reif, Marti Wolf
U.S EPA, ORD, Computational Toxicology Research Program

research&development

ACToR is a web-based informatics platform, organized at the top level by chemical and chemical structure that is indexing, collecting, and organizing many types of data on environmental chemicals. Environmental chemicals are defined as those likely to be in the environment, including all chemicals regulated or tracked by the EPA, as well as related chemicals, such as pharmaceuticals that find their way into water sources. ACToR is indexing and linking to data from hundreds of sources, including the EPA, FDA, CDC, NIH, academic groups, other governmental agencies (state and national) and international organizations, such as the WHO. Information being indexed and gathered includes *in vivo* toxicity, *in vitro* bioassay data, use levels, exposure information, chemical structure, regulatory information and other descriptive data. Planning for the project began in mid-FY07; beta versions were available inside the EPA since early FY08, and a public version became available in December 2008. ACToR consists of a back-end database and a front-end web interface built on low-cost, publicly accessible applications and tools.

Science Question

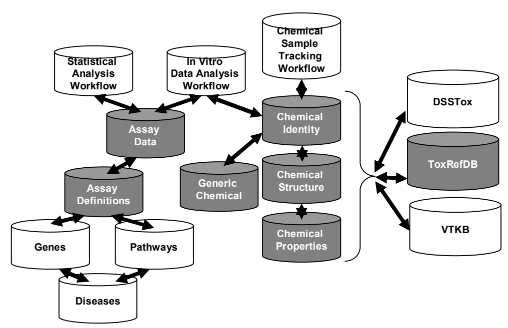
ACToR was developed to address several needs of the NCCT, the EPA and the broader community of researchers on environmental chemicals:

1. What is the collection of chemicals of most concern to the EPA?
2. What toxicology, exposure, use, in vitro and regulatory information is known for these chemicals?
3. Can this information be collected in one place to allow easy access by researchers around the world?
4. Can information on the structure, properties and activity of chemicals be tabulated for use in computational modeling efforts?

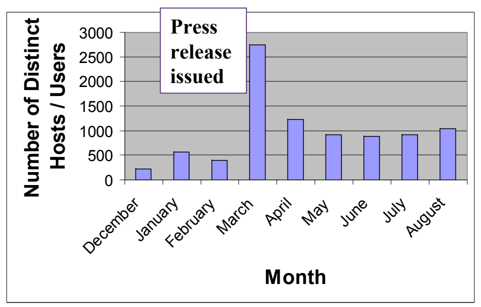
Research Goals

ACToR aims to provide a unified, centralized resource of data on environmental chemicals including toxicology, *in vitro* assay data, and chemical structure information. By gathering information on the type and location of toxicity or exposure data associated with environmental chemicals into a single, searchable, publicly accessible web-site, ACToR is providing the basis for chemical selection and screening within NCCT projects, such as ToxCastTM and Tox21. ACToR is also coordinating with the DSSTox project to incorporate quality chemical review and structure-annotation for the chemical data sets of highest interest to the various NCCT projects. In addition to its use in supporting various NCCT and EPA projects, ACToR is a publicly available EPA resource that enables other government agencies, industry, and academic researchers to quickly search and collate toxicity-related information on chemicals of interest. As such, it will promote and encourage other entities to adopt standards for chemical representation and broadly survey chemical information pertaining to toxicology resources on the Internet.

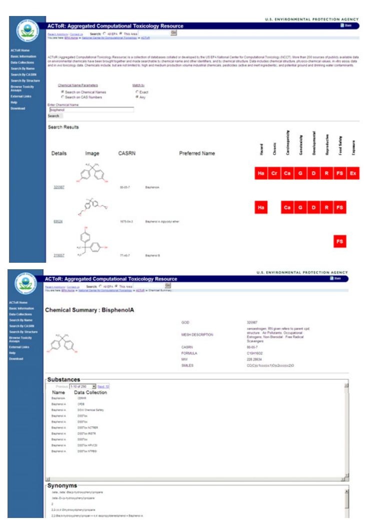
ACToR Online



The online ACToR database has the widest collection of information and can be searched chemical-by-chemical for an overview of the world's information. Currently, the database contains information on >500K unique chemicals with data from more than 400 source databases.

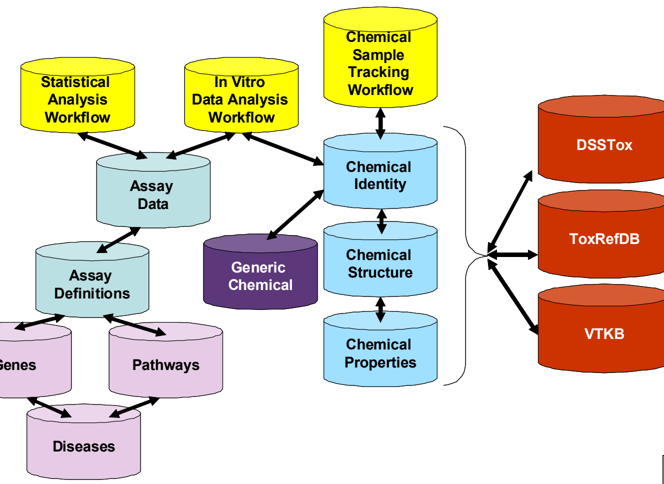


The database has a steady user base of ~ 1,000, about 50% inside the EPA.



See Live Demo

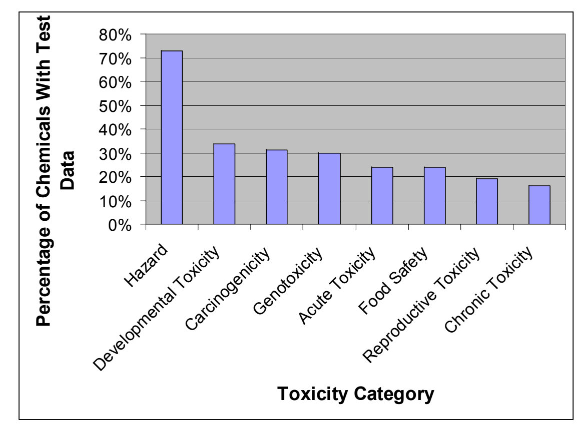
Chemical list summary and chemical data page from ACToR Online



ACToR is comprised of a set of linked databases sharing many components, in particular those containing information on chemical identity. All databases are implemented in MySQL and are filled and manipulated using open source software tools written in Perl, Java and R. Where possible, database design follows precedents set by NCBI PubChem and Entrez.

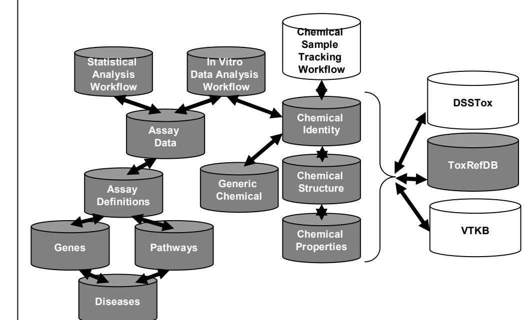
Data Sets currently in ACToR by Number of Chemicals

ABC	826	EPA OPPIN Inerts	3532	GLC AirToxics 2002	190	SRS_RCRA_F_Waste	9
AIHA WEEL	114	EPA OPPIN Pesticide Active	835	HCPSL (1999)	174	SRS_RCRA_P_Waste	120
Arizona Pest Use	73	EPA OW MCL	74	HCPSL (2006)	361	SRS_RCRA_T_Char	40
ATSDR CEP 2007	97	EPA OW UCMR 1and2 Water	73	Health Canada DSL Structures	9505	SRS_RCRA_U_Waste	251
ATSDR HSEES 2006	40	EPA OW UCMR2	10	Health Canada High Health	516	SRS_SARA_110	391
ATSDR Interactions	42	EPA OW WQC Nonpriority	30	Health Canada Low Priority	762	SRS_SARA_302A	406
ATSDR MMG	73	EPA OW WQC Organoleptic	23	Health Canada Moderate Health	631	SRS_SDWA_NPDWR	177
ATSDR MRL	164	EPA OW WQC Priority	122	HEI	29	SRS_SDWA_NSDWR	10
ATSDR PDN Doc	16	EPA PCCL	529	HERA	263	SRS_TRI	665
ATSDR PDN Status	58	EPA Pesticide Active	405	HMDB	2206	SRS_TSCA_12_b_Export	799
ATSDR PHS	200	EPA Pesticide Fate	188	INCHEM CICAD	268	SRS_TSCA_4a_Test_Rules	80
ATSDR ToxFaq	218	EPA PestUsage	29	INCHEM EHC	1121	SRS_TSCA_4_TERM	33
ATSDR ToxGuides	28	EPA PFS Biopesticide	159	INCHEM HSG	158	SRS_TSCA_4_Tests	296
ATSDR ToxProfile	284	EPA PFS Conventional	105	INCHEM IARC	1254	SRS_TSCA_5A_SNUR	560
AU NPI	79	EPA PFS Reduced Risk	72	INCHEM PD	93	SRS_TSCA_5e_Consent_Orders	200
AU NPI Threshold	79	EPA RBC	639	IRAC MOA	220	SRS_TSCA_6A_CCCR	10
Cal Pesticide Residues	41	EPA Reregistration	626	ITER TERA	639	SRS_TSCA_6_Unreasonable_Risk	21
Cal PestUse	195	EPA RSEI	403	KIDB	843	SRS_TSCA_8A_CAIR	52
CalEPA	676	EPA SRS Structures	31916	MDCH	118	SRS_TSCA_8A_INFO	259
CalEPA Pyrethroids	12	EPA Title III 2006	1328	MESH	61625	SRS_TSCA_8A_PAIR	931
CalOEHHA	274	EPA TRI	631	MHLW JP	236	SRS_TSCA_8c_SARS	52
CalOEHHA Air	120	EPA TRI Glycols	3457	MINN FactSheets	22	SRS_TSCA_8D_HSDR_a	342
CalOEHHA Water	62	EPA TRI PBT	17	MINN HealthRisk	122	SRS_TSCA_8D_HSDR_c	250
CARB	39	EPA TRI Releases 1988	252	MINN HRE Air	87	SRS_TSCA_8D_HSDR_d	414
CARB BusStudy	18	EPA TRI Releases 1989	254	MINN HRE Water	136	SRS_TSCA_8D_TERM	332
CERCLA	275	EPA TRI Releases 1990	259	MITI Japan Structures	14486	SRS_TSCA_Inventory	65513
CERHR	41	EPA TRI Releases 1991	264	MUTBRS	846	SRS_TTO	113
CPDB	1509	EPA TRI Releases 1992	258	NCI-Open_09-03	122719	SRS_UCMR_1999	26
CYP IUPUI	245	EPA TRI Releases 1993	256	NHANES	196	TC CWA	27
CYP NCI CTEP	387	EPA TRI Releases 1994	277	NHEXAS Arizona	54	TC ERG2008	215
Danish EPA Structures	82055	EPA TRI Releases 1995	461	NHEXAS Maryland	35	TC WaterReactive	77
DanishEPA Coformulants	19	EPA TRI Releases 1996	463	NHEXAS Region5	22	Tox21 New Nominations	185
DEA DrugsOfConcern	35	EPA TRI Releases 1997	459	NIAID HIV DrugDev	72	ToxCast Phase I (320)	309
DEA NFLIS2007	25	EPA TRI Releases 1998	493	NIOSH CARC	126	ToxCast_1408	1331
DEA Schedule I	118	EPA TRI Releases 1999	498	NIOSH ChemHazard	455	ToxRefDB	307
DEA Schedule II	52	EPA TRI Releases 2000	498	NIOSH ERSHDB	40	TSCATS	6438
DEA Schedule III	74	EPA TRI Releases 2001	493	NIOSH Immediately Dangerous	381	UK Allergens	26
DEA Schedule IV	69	EPA TRI Releases 2002	490	NIOSH NOES	427	UK Cosmetics Provisional	59
DEA Schedule V	8	EPA TRI Releases 2003	490	NIOSH PEL	385	UK Cosmetics Restricted	72
DOD Gulf Pest	12	EPA TRI Releases 2004	478	NLM TOXNET CCRIS	9301	UK EA Inventory	225
DOD Wildlife Tox	32	EPA TRI Releases 2005	481	NLM TOXNET DART	15049	UK FlashPoint	33
DOE GreenhouseGases	12	EPA TRI Releases 2006	489	NLM TOXNET GENETOX	3205	UK Munitions Sea	11
DOE GreenhouseGases Production	24	EPA TSCA MTL	562	NLM TOXNET HSDB	5115	UK PestUsage	32
DOE IndoorAir	57	EPA TSCA Structures	52083	NLM TOXNET Toxicology	70760	UK Pharm Use	50
DOW Chemical Safety	129	EPA TSCA Sunset	193	NOAA MarineResources	515	UK Pharm Water	16
DrugBank	2219	ESIS ESIS	2604	NOAA Pharm Kow	198	UK PPDB	998
DSSTox	7817	ESIS HPV	2767	NOAA Quick ref	87	UK Space	126
DSSTox EPAFHM	617	ESIS LPV	7802	NPIC	326	UMD MOA	33
DSSTox FDAMDD	1214	ESIS NLP	702	NPRI List 2007	309	UN IPCC Greenhouse Gases	22
DSSTox GEOGSE	849	ESIS ORATS	141	NPRI List 2008	309	US FDA DRUGS	4808
DSSTox HPVCSI	3548	ESIS PBT	127	NPRI Releases 2007	292	USDA FS HERA	35
DSSTox HPVISD	1006	EU Detergent DID	102	NTP DevTox	93	USDA NOP Crops	60
DSSTox IRISTR	537	EU EDEXIM Part1	116	NTP ImmunoTox	54	USDA NOP Livestock	30
DSSTox NCTRER	228	EU EDEXIM Part2	37	NTP Nominations	146	USDA NOP Processing	75
DSSTox NTPBSI	2198	EU EDEXIM Part3	31	NTP ReproTox	87	USDA Pest Properties	332
DSSTox NTPHTS	1354	EU Fragrances	2434	NTP ToxCarcLong	513	USDA Pesticide Residues	232
EDGE2	20	EUROPA CosIng Dyes	162	NTP ToxCarcShort	92	USGS Biosolids	24
EINECS Structures	52928	EUROPA CosIng Preserv	141	OECD Categories	114	USGS HBSL	415
Environment Canada DSL (2007)	22017	EUROPA CosIng UVFilters	33	OECD HPV	4128	USGS Pest Fish	27
EPA AEGL Final	39	EUROPA DanishEPA QSAR	134	OECD HPVC Structures	2856	USGS Pest GroundWater	83
EPA AEGL Holding	33	EUROPA Dyes Assessed	49	OPCW	56	USGS Pest Sediment	31
EPA AEGL Interim	192	EUROPA Dyes Provisional	63	OSHA EPA Occupational Database	750	USGS Pest Stream	83
EPA AEGL Proposed	17	EUROPA ED Status	108	OSHA HealthGuidelines	165	USGS Pharm Soil	19
EPA AEGL SecondPriority	371	EUROPA ED Water	23	PCN	1813	USGS SWQA	17
EPA CARC	400	EUROPA Endocrine Candidate	489	PCP Ingredients	1402	USGS VOCs Aquifer	55
EPA CCL1	48	EUROPA Food Contact Additives	482	PubChem	475265	USGS Wastewater Minn	91
EPA CCL2	40	EUROPA Food Contact Monomers	298	RAISMetadata	641	USITC Commodity 2001	8
EPA CCL3	93	EUROPA Food Contact Provisional	24	RAISToxProfiles	75	USITC Dyes	177
EPA Child PK	44	EUROPA Food Residues	736	RAISToxValues	712	USITC FlavorFragrances 1997	30
EPA DWSHA	201	EUROPA FoodPest 1996	8	REACH Intentions	42	USITC Intermed DutyFree	1346
EPA ECOTOX	8835	EUROPA FoodPest 1997	13	REACH Intentions Submitted	16	USITC Pharma DutyFree	7502
EPA EDC73	73	EUROPA FoodPest 1998	19	REACH PreRegister	116657	WASH PBT	75
EPA EFED MOA	1642	EUROPA FoodPest 1999	19	RiskIE	3525	WHO CPH	559
EPA HAP 1996 NATA	23	EUROPA FoodPest 2000	19	RoC	217	WHO Discontinued Pesticides	285
EPA HAP 1999 NEI	162	EUROPA FoodPest 2001	33	Rotterdam AppendixV	176	WHO Drug Interactions	19
EPA HAP 2002 NEI	162	EUROPA FoodPest 2002	38	Rotterdam PIC	81	WHO Water Guidelines	90
EPA HAP AcuteRisk	307	EUROPA FoodPest 2003	39	Scorecard	3107	Wikipedia	1542
EPA HAP ARP	138	EUROPA PestDiet 2001	34	SIDS	395	Wildlife Toxicity Assessment	32
EPA HAP ChronicRisk	248	EUROPA PestDiet 2002	40	SRS_CAA112_b_HAP	169	WISC AirToxics	613
EPA HAP FactSheets	173	EUROPA PestDiet 2003	41	SRS_CAA112_b_HON	385	WISC GRN Wells2001	311
EPA HPV	2539	EUROPA Supplements	107	SRS_CAA_109	5		
EPA HPV Challenge	1973	EUROPA VetMed	28	SRS_CAA_111	477		
EPA HPV Challenge Structures	8059	EWG TOXOME	444	SRS_CAA_112R	139		
EPA HPVIS	1094	EXTOXNET	188	SRS_CAA_602	89		



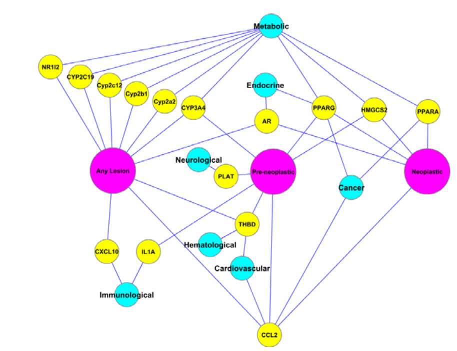
Summary of toxicity data coverage in ACToR for 17,414 chemicals nominated for the Tox21 program

ToxMiner



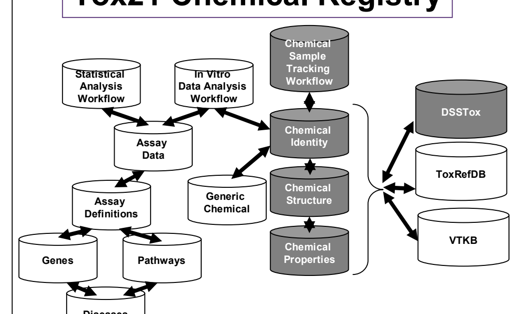
The ToxMiner database merges all data from ToxCast assays, endpoint data from ToxRefDB and chemical identity information from the main ACToR database.

The statistical analysis workflow components of the ToxMiner database manage the input and results for all statistical machine learning analyses run for ToxCast



An example result from ToxMiner showing genes (yellow) correlated with stages of rat liver disease from ToxCast and ToxRefDB (purple) and human disease from the literature (blue)

Tox21 Chemical Registry



The Tox21 Chemical Registry is managing all nominations for chemicals to be run in the Tox21 and ToxCast Phase II assays. Currently, we have about 15,000 chemicals nominated from dozens of organizations inside and outside of the EPA. Chemicals are automatically annotated as having appropriate physico-chemical properties, being simple or complex mixtures, being salts or metals, etc. Groups of chemicals are exported to be ordered and their status through ordering, receiving, solubilization and plating is tracked.

Results/Conclusions

The ACToR databases have proven to be flexible and scalable enough to handle all of the types of data imported as part of the ACToR and ToxCast projects.

Impact and Outcomes

1. Statistical analysis of ToxCast Phase I has been greatly aided by the organization provided by ToxMiner.
2. ACToR Online is becoming widely used as a source of information on a variety of environmental chemicals.
3. The Tox21 workflow database has enabled us to manage nominations from dozens of organizations and to rationalize the process of selecting thousands of chemicals to push into the Tox21 and ToxCast Phase II assays.

Future Directions

FY09

- **Initial public deployment.**
- **Release of version 2, including refined chemical structure information.**
- Develop workflow for tabularization of data buried in text reports.
- Integrate all ToxCast and ToxRefDB data.
- Quarterly releases with new data.

FY10

- Quarterly releases with new data.
- Implementation of a process to gather tabular data on priority chemicals from text reports.
- Survey sources of chemical use and exposure data and import any remaining sources.
- Develop flexible query interface and data download process.
- Develop process to extract data from open literature.

FY11 and FY12

- Quarterly releases with new data

References

- R. Judson, A. Richard, D. Dix, K. Houck, F. Elloumi, M. Martin, T. Cathey, T.R. Transue, R. Spencer, M. Wolf, "ACToR – Aggregated Computational Toxicology Resource", Toxicology and Applied Pharmacology, Vol. 233, 7-13 (2008).
- R. Judson, A. Richard, D.J. Dix, K. Houck, M. Martin, R. Kavlock, V. Dellarco, T. Henry, T. Holderman, P. Sayre, S. Tan, T. Carpenter, E. Smith, "The Toxicity Data Landscape for Environmental Chemicals", Environmental Health Perspectives, Vol. 117, 685 (2008)

This poster does not necessarily reflect EPA policy. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.