

ACToR: Aggregated Computational Toxicology Resource Richard Judson, Ann Richard, Tommy Cathey, Tom Transue,

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

Science Question

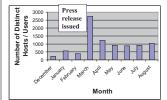
- ACTOR was developed to address several needs of the NCCT, the EPA and the broader community of researchers on
- What is the collection of chemicals of most concern to the EPA?
- What toxicology, exposure, use, in vitro and regulatory information is known for these chemicals?
- Can this information be collected in one place to allow easy access by researchers around the world?
- 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 structureannotation 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



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.

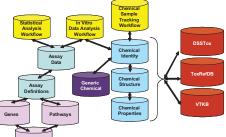


The database has a steady user base of ~ 1,000, about 50% inside the EPA.



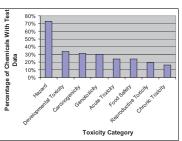
Chemical list summary and chemical

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 loxicity, 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 no love-cost, publicly accessible applications and tools.

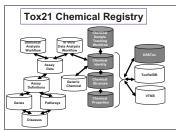


Data Sets currently in ACToR by Number of Chemicals

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.



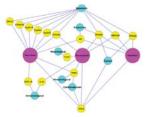
Summary of toxicity data coverage in ACToR for 17,414 chemicals nominated for the Tox21 program



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

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)

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

- Statistical analysis of ToxCast Phase I has been greatly aided by the organization provided by ToxMiner.
- ACToR Online is becoming widely used as a source of information on a variety of environmental chemicals.
- 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)