http://www.orcid.org/0000-0002-2668-4821

The CompTox Chemistry Dashboard v3.0 – New Searches and Support for Bioactivity Data

Antony Williams

National Center for Computational Toxicology, U.S. Environmental Protection Agency, RTP, NC

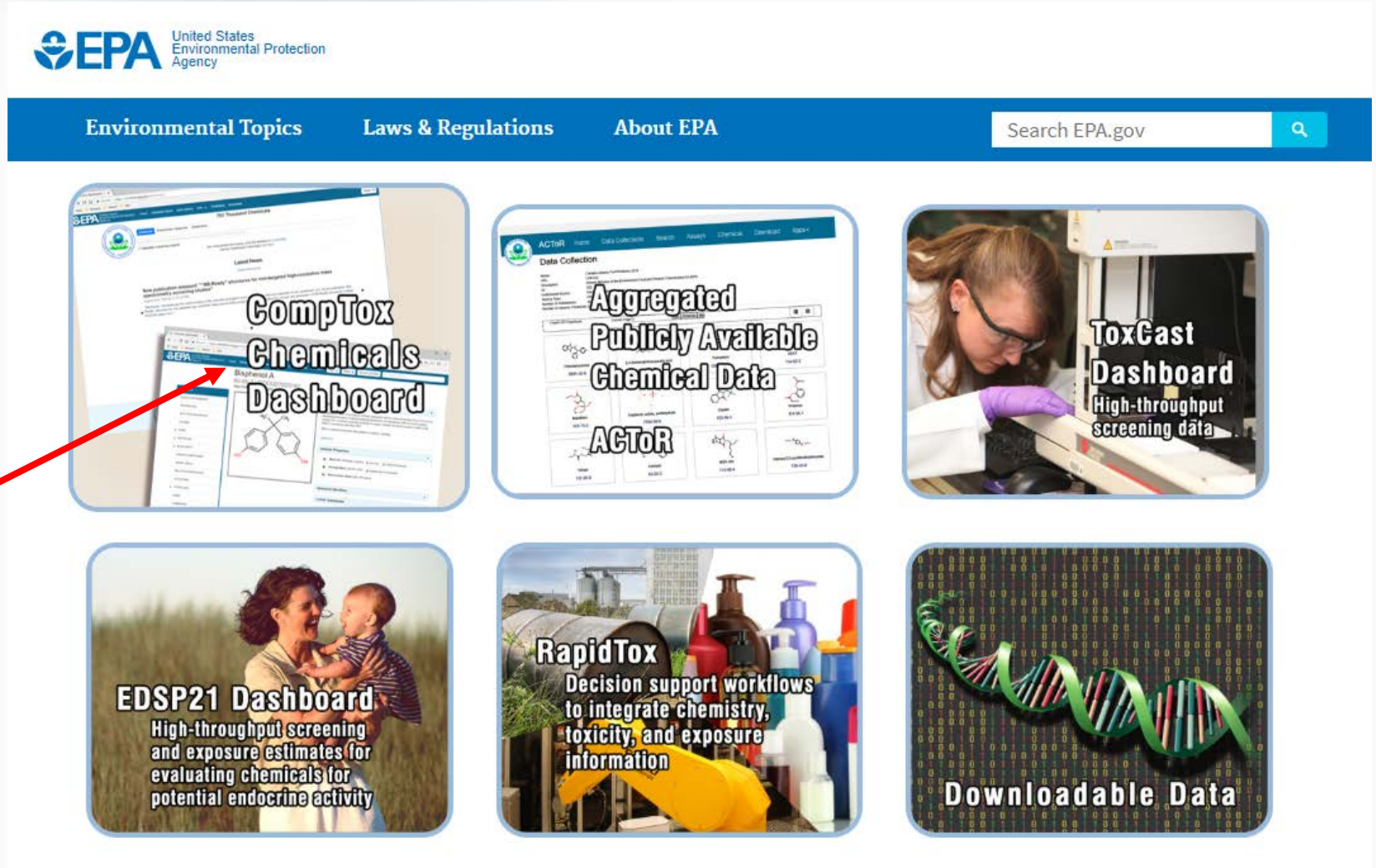
The views expressed in this presentation are those of the author and do not necessarily reflect the views or policies of the U.S. EPA

September 27th 2018
Communities of Practice

Outline – what's new in v3.0?

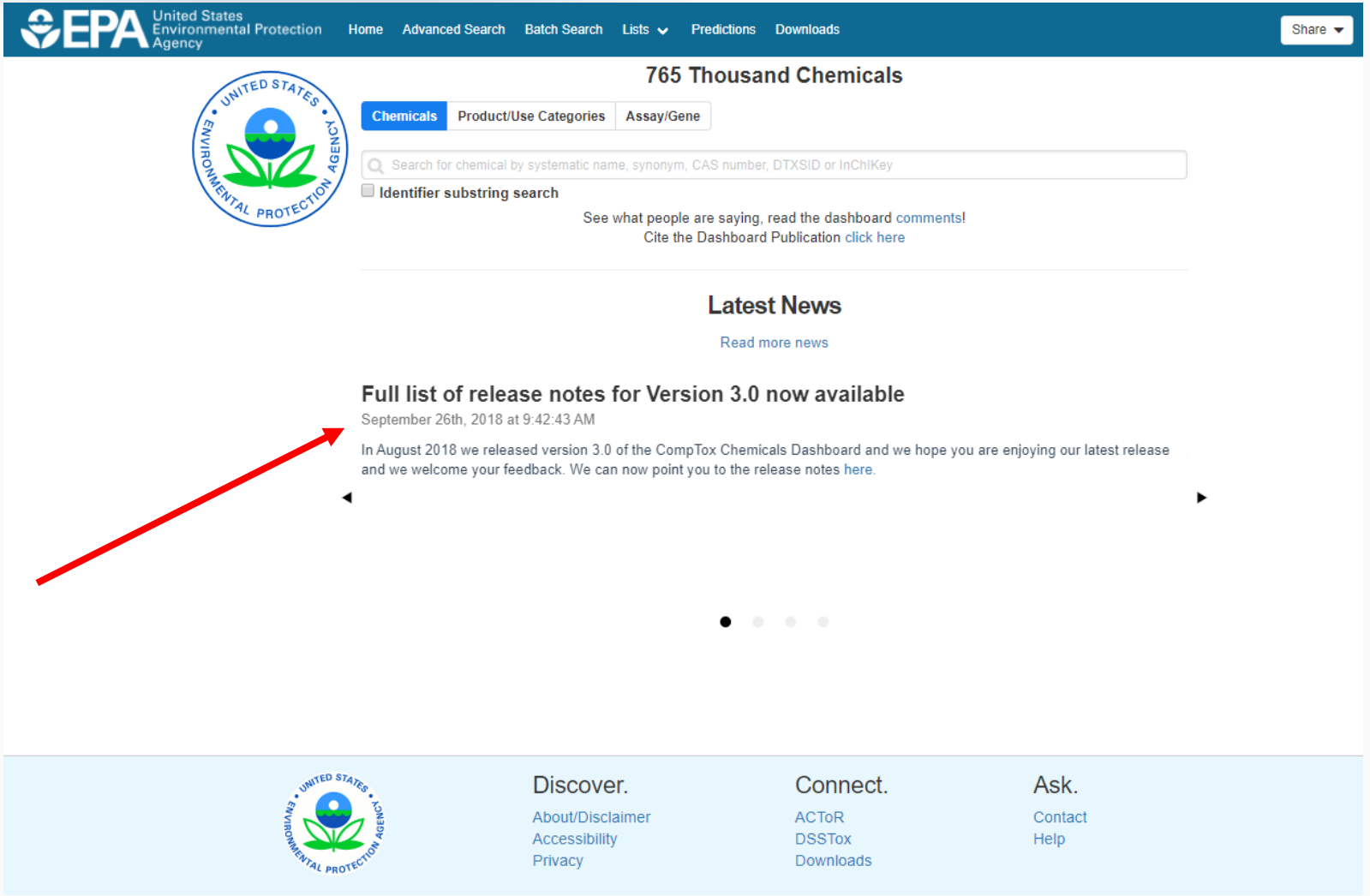
- Welcome the CompTox Portal
- New name for the dashboard
- User interface overhaul – easier navigation
- New search capabilities
- Enhanced support for bioactivity data
- New data and new lists added
- Work in progress

The CompTox Portal
https://comptox.epa.gov/



Watch for our news

https://comptox.epa.gov/dashboard/news_info



Release Notes

https://comptox.epa.gov/dashboard/comptox_release_notes.pdf

- A detailed list of new functionality and fixes

**** New Functionality**

- [ICD-2992] - Allow Display of Sets of Chemicals based on Product or Use Categories
- [ICD-3045] - Improve Submit Comment capability
- [ICD-3157] - Add Assay/Gene Search based on assays available for ToxCast
- [ICD-3174] – On-hover Zoom on structure general solution across the dashboard
- [ICD-3182] - Build Assay Chart display window for displaying one or more assay charts
- [ICD-3183] - Build multiple chart display mode
- [ICD-3242] - Include image to depict presence of isotopes
- [ICD-3324] - Add on hover display of AOP title associated with AOP Link and AOP Event
- [ICD-3339] - Create MultiStep Navigation Workflow for GenRA
- [ICD-3340] - Create GenRA layout Block 1
- [ICD-3341] - Create GenRA layout Block 2
- [ICD-3342] - Create GenRA layout Block 3
- [ICD-3355] - Create GenRA layout Block 4
- [ICD-3361] - DSSTox relationships controlled vocab table
- [ICD-3362] - Provide list of all assays that chemicals have been measured in
- [ICD-3363] - Add More Columns to Lists Page
- [ICD-3369] - Add Data Picker to Tile Mode
- [ICD-3373] - Add additional Functionality to the Assay Selection Tab
- [ICD-3374] - Create Dedicated Chemical Lists Pages associated with individual assays

Staying up with the Dashboard

https://comptox.epa.gov/dashboard/news_info

EPA United States Environmental Protection Agency

New list of pesticides added

September 19th, 2018 at 10:22:58 AM

A list of pesticides associated with the EPA Pesticide Chemical Search Database has been added to the list page. The PC codes (pesticide codes) have also been added as searchable synonyms to the chemical search page. The list is available here

Video describing how to use the "Generalized Read-Across (GenRA) module" now on YouTube

August 31st, 2018 at 12:22:02 PM

A new module describing Generalized Read-Across (GenRA) is now available here on YouTube. The video runs through the basic science behind GenRA and how to use the module in the dashboard.

New publication released: ""MS-Ready" structures for non-targeted high-resolution mass spectrometry screening studies"

August 31st, 2018 at 12:07:25 PM

"MS-Ready" structures are the basis of many of the searches to support mass spectrometry that are supported on the dashboard. Our recent publication "MS-Ready" structures for non-targeted high-resolution mass spectrometry screening studies" explains the concept and production of MS-Ready structures in detail. Read the paper here ".

Version 3 of the Dashboard released August 10th 2018

August 12th, 2018 at 2:46:01 PM

A new version of the CompTox Dashboard has been released to the community. Other than just searching for chemicals, this version includes new searches for product and use categories and assays and genes associated with ToxCast and Tox21 assays. Bioactivity curves are now viewable for assays associated with the Endocrine Disruptor Screening Program (EDSP21). A detailed list of new functionality will be forthcoming.

A Major Update to the Dashboard Releases on March 7th 2018

March 8th, 2018 at 3:49:25 PM

A major update to the dashboard has been released prior to the Society of Toxicology and American Chemical Society Spring meetings. This release brings together six months of effort in adding and curating data, major updates to the batch searching functionality and access to real time predictions for both physiochemical and toxicity endpoints. A list of release notes is available for your review. We look forward to your feedback.

A YouTube video regarding using the Dashboard for Non-Targeted Analysis

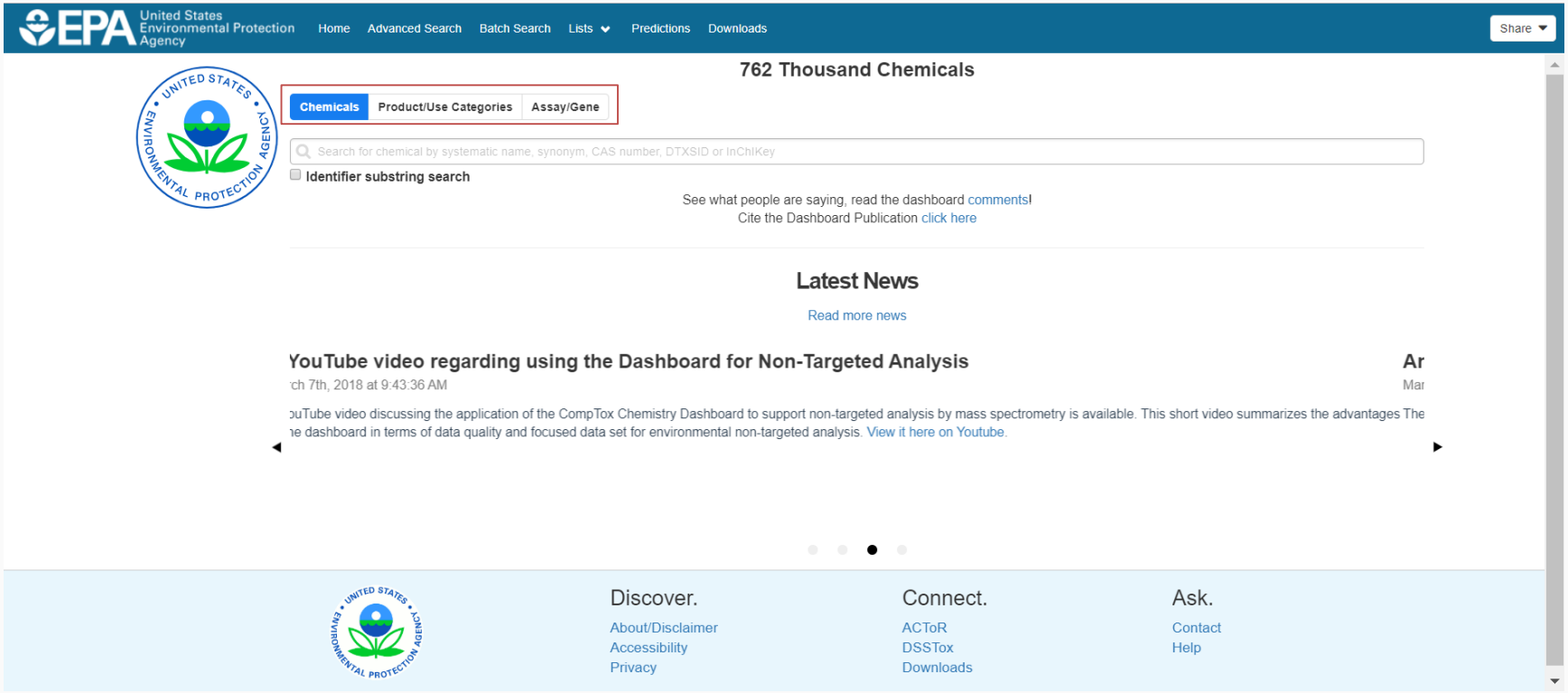
March 7th, 2018 at 9:43:36 AM

A YouTube video discussing the application of the CompTox Chemistry Dashboard to support non-targeted analysis by mass spectrometry is available. This short video summarizes the advantages of the dashboard in terms of data quality and focused data set for environmental non-targeted analysis. View it here on Youtube.

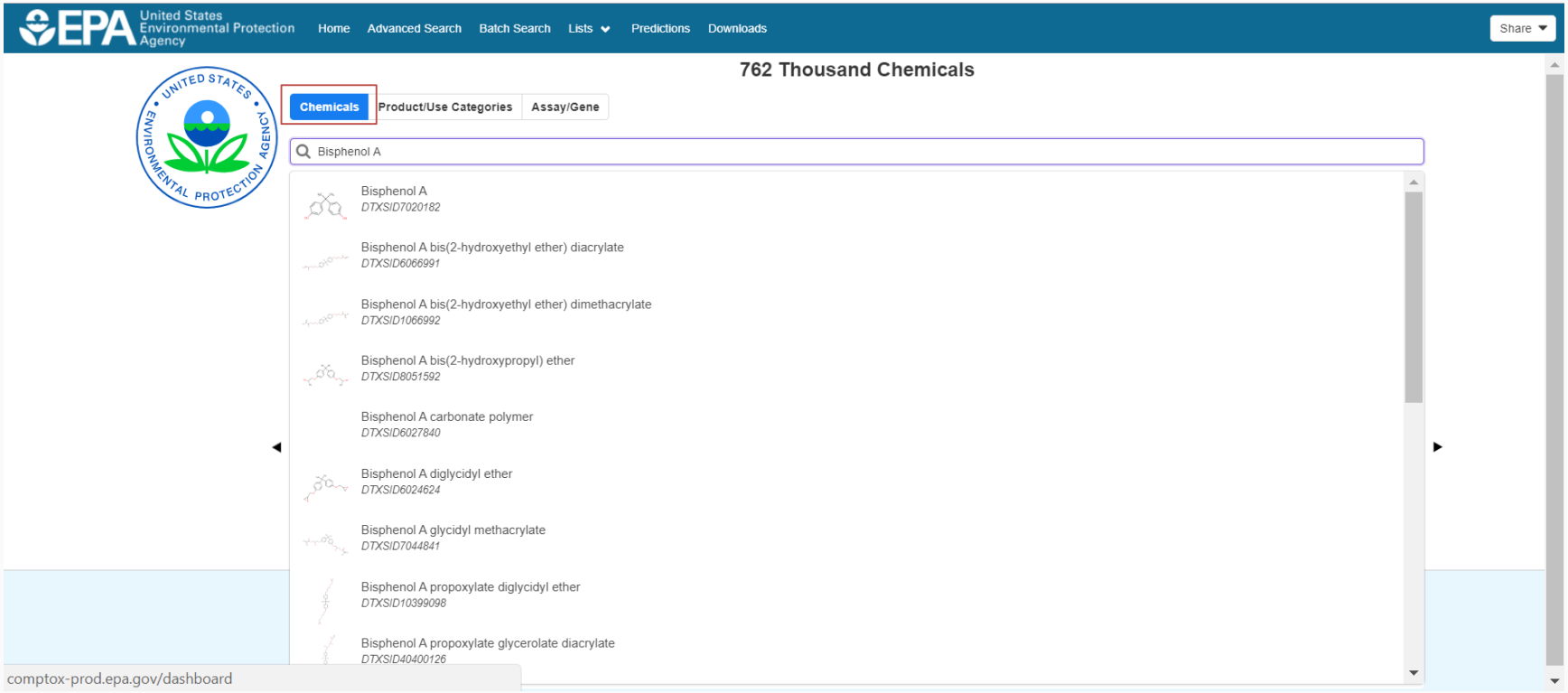
An article regarding an Excel Version of the Abstract Sifter is published.

CompTox Chemicals Dashboard

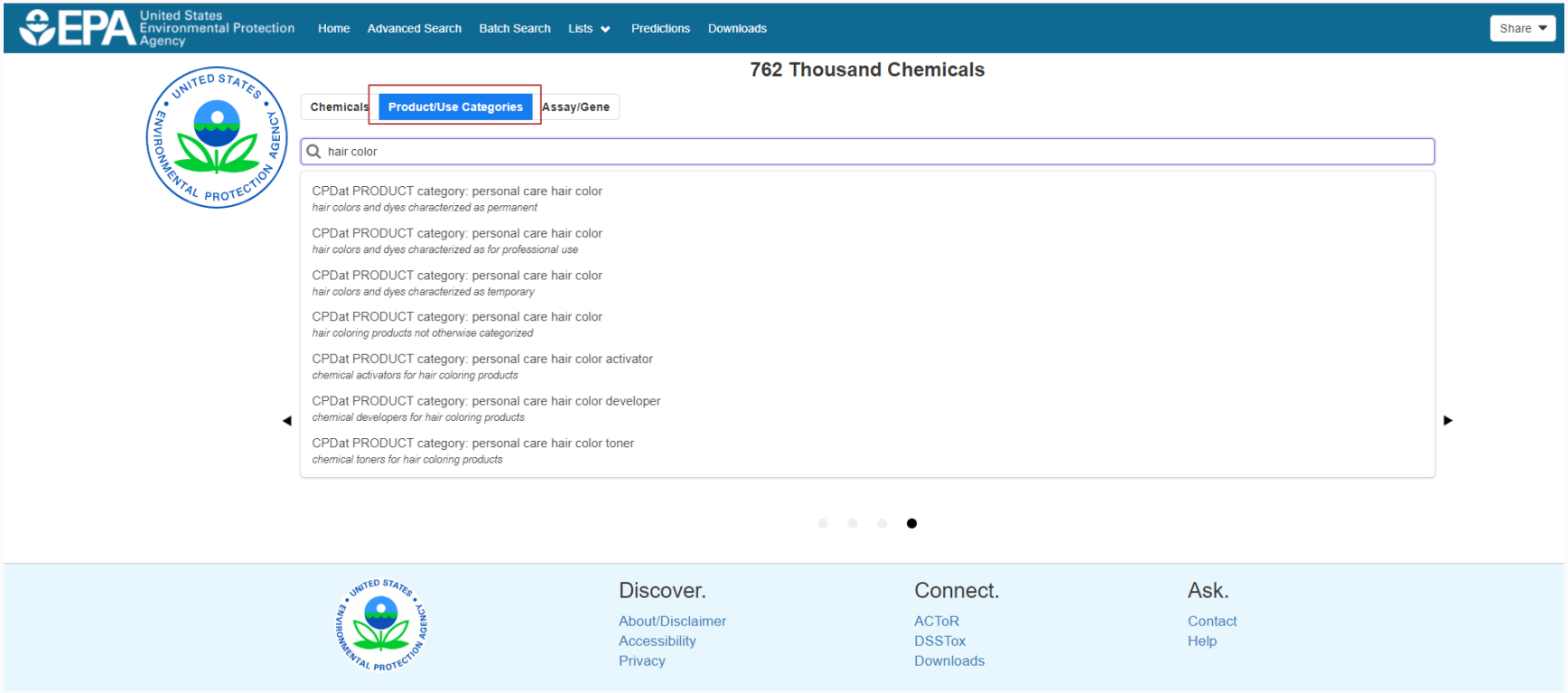
https://comptox.epa.gov/dashboard



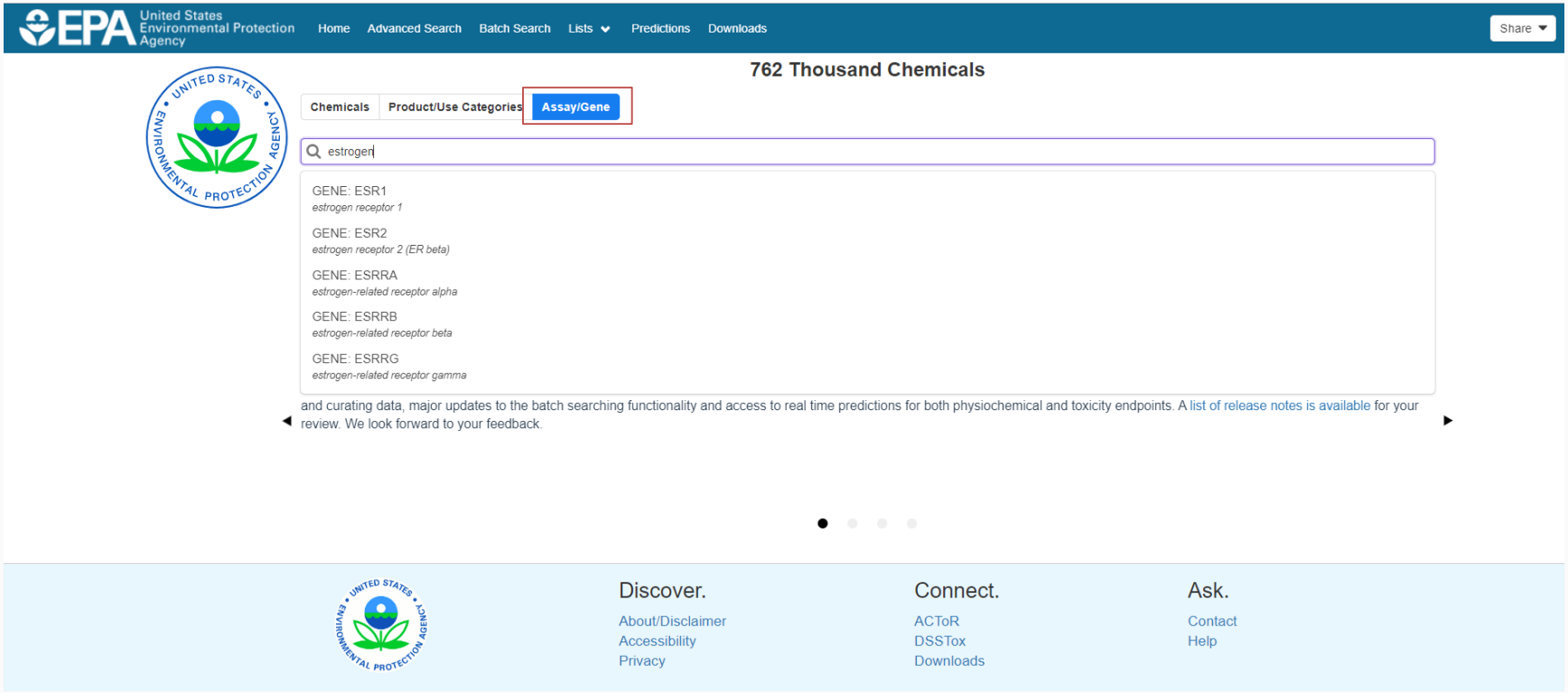
CompTox Dashboard Chemicals



CompTox Dashboard
Products and Use Categories

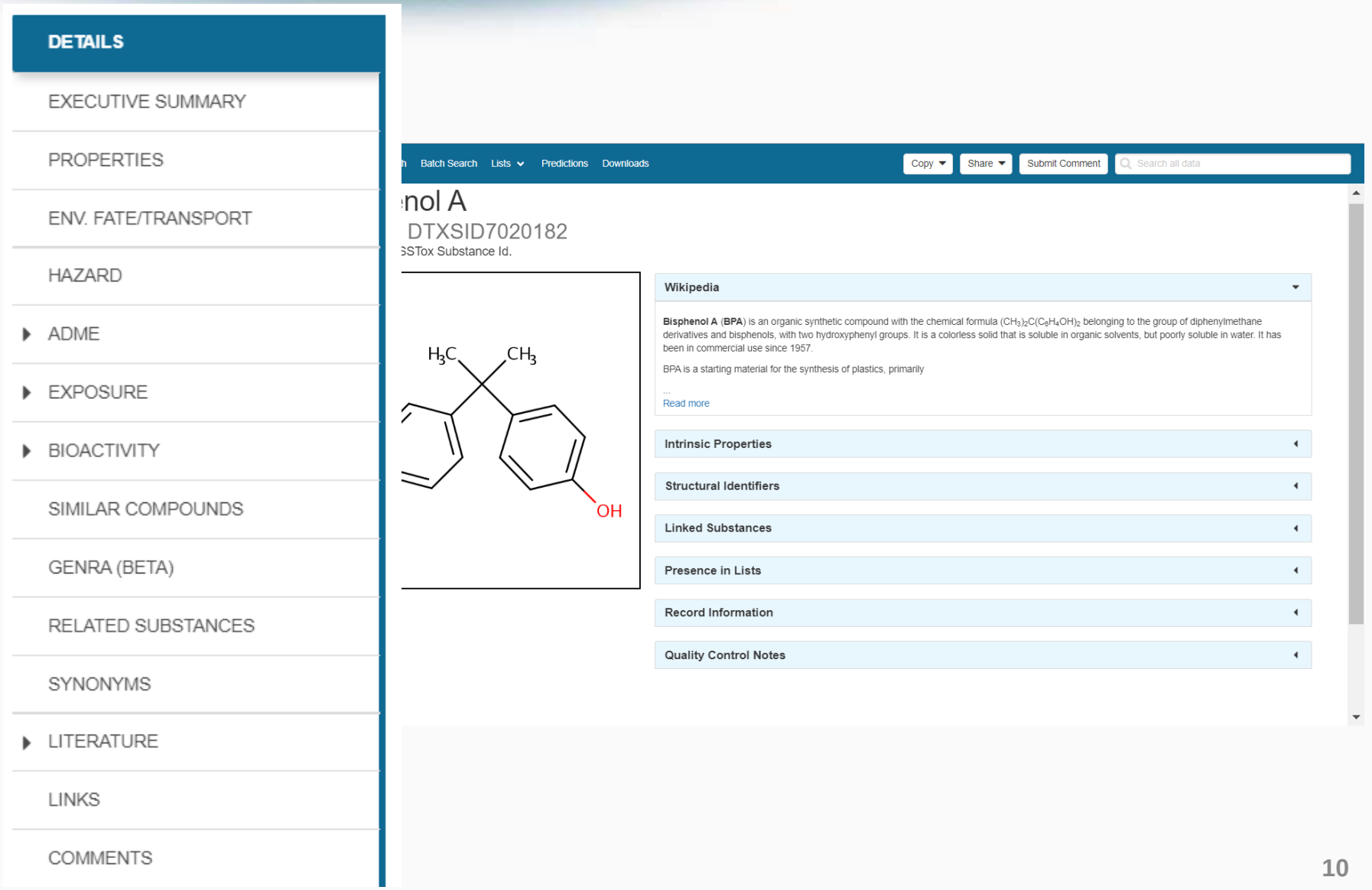


CompTox Dashboard
Assays and Genes

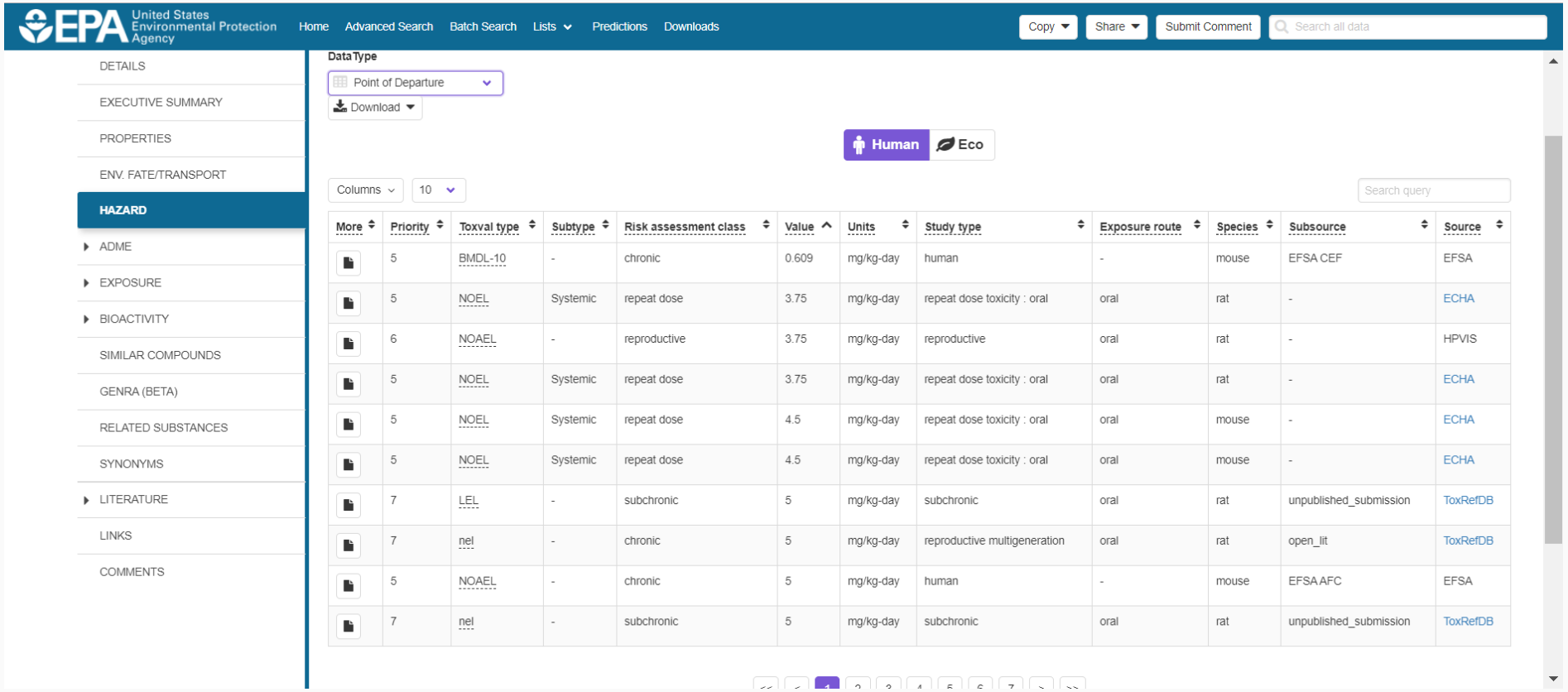


Detailed Chemical Pages
New User Interface Design

EPA United States Environmental Protection Agency



Access to Chemical Hazard Data

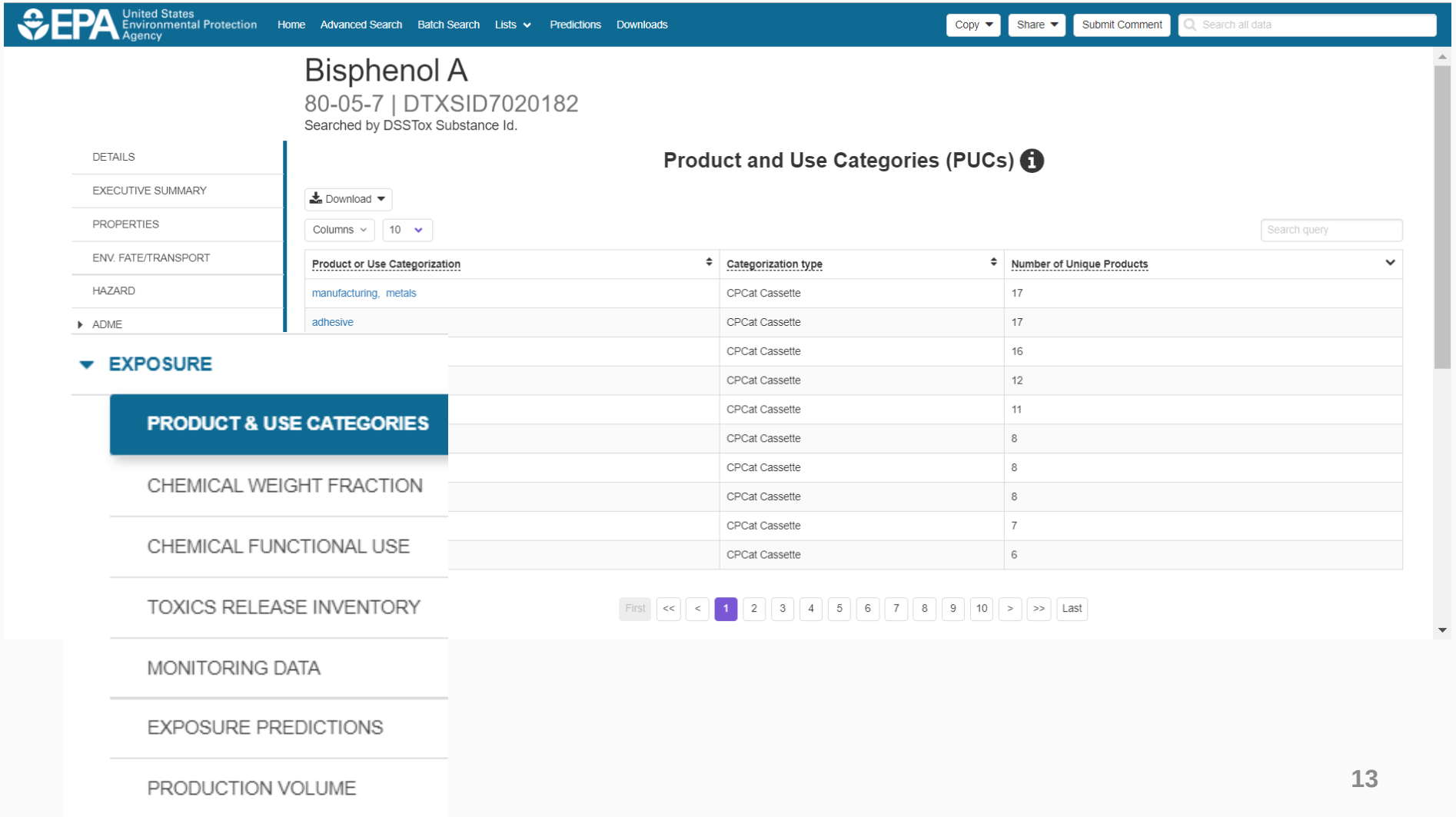


More	Priority	Toxval type	Subtype	Risk assessment class	Value	Units	Study type	Exposure route	Species	Subsource	Source
	5	BMDL-10	-	chronic	0.609	mg/kg-day	human	-	mouse	EFSA CEF	EFSA
	5	NOEL	Systemic	repeat dose	3.75	mg/kg-day	repeat dose toxicity : oral	oral	rat	-	ECHA
	6	NOAEL	-	reproductive	3.75	mg/kg-day	reproductive	oral	rat	-	HPVIS
	5	NOEL	Systemic	repeat dose	3.75	mg/kg-day	repeat dose toxicity : oral	oral	rat	-	ECHA
	5	NOEL	Systemic	repeat dose	4.5	mg/kg-day	repeat dose toxicity : oral	oral	mouse	-	ECHA
	5	NOEL	Systemic	repeat dose	4.5	mg/kg-day	repeat dose toxicity : oral	oral	mouse	-	ECHA
	7	LEL	-	subchronic	5	mg/kg-day	subchronic	oral	rat	unpublished_submission	ToxRefDB
	7	nel	-	chronic	5	mg/kg-day	reproductive multigeneration	oral	rat	open_lit	ToxRefDB
	5	NOAEL	-	chronic	5	mg/kg-day	human	-	mouse	EFSA AFC	EFSA
	7	nel	-	subchronic	5	mg/kg-day	subchronic	oral	rat	unpublished_submission	ToxRefDB

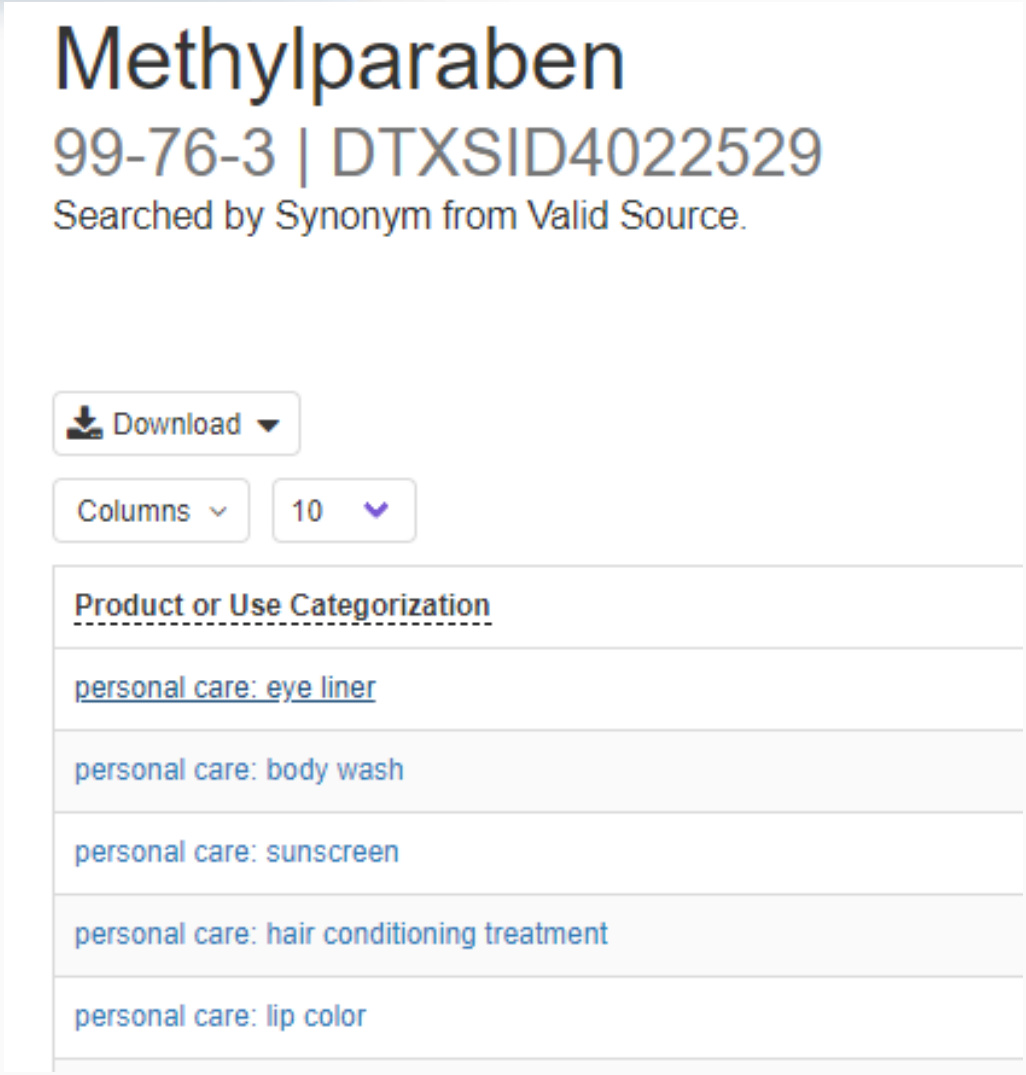
Hazard Data from “ToxVal_DB” Lots of new data added - ECOTOX

- ToxVal Database contains following data:
  - 30,050 chemicals
  - 772,721 toxicity values
  - 29 sources of data
  - 21,507 sub-sources
  - 4585 journals cited
  - 69,833 literature citations

Sources of Exposure to Chemicals



What chemicals in what product and use categories?



What chemicals in what product and use categories?

EPA United States Environmental Protection Agency

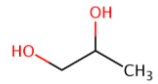
Searched by Product & Use Categories

Results for CPDat Product Category: Personal Care: Eye Liner

182 chemicals

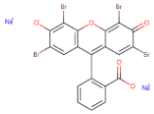
Download / Send | Show info: DTXSID × CASRN × TOXCAST × | Select all

Sort by: DTXSID | Filter by: Name or CASRN | Hide



1,2-Propylene glycol
DTXSID: DTXSID0021206
CASRN: 57-55-6
TOXCAST: 11/539

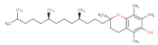
cis-Oleyl alcohol
DTXSID: DTXSID0022010
CASRN: 143-28-2
TOXCAST: 0



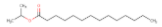
Eosin
DTXSID: DTXSID0025234
CASRN: 17372-87-1
TOXCAST: 45/302

1 related chemical structure with this substance

Polyvinylpyrrolidone
DTXSID: DTXSID0025941
CASRN: 9003-39-8
TOXCAST: 0



alpha-Vitamin E
DTXSID: DTXSID0026339
CASRN: 59-02-9
TOXCAST: 0



Isopropyl tetradecanoate
DTXSID: DTXSID0026838
CASRN: 110-27-0
TOXCAST: 1/299

0 related chemical structures with this substance

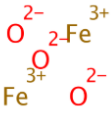
Paraffin waxes and Hydrocarbon waxes...
DTXSID: DTXSID0028115
CASRN: 63231-60-7
TOXCAST: 0

1 related chemical structure with this substance

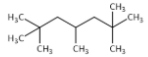
Alcohols, C16-18
DTXSID: DTXSID0028323
CASRN: 67762-27-0
TOXCAST: 1/163

0 related chemical structures with this substance

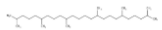
Fats and Glyceridic oils, vegetable, hyd...
DTXSID: DTXSID0028454
CASRN: 68334-28-1
TOXCAST: 0



Iron(III) oxide
DTXSID: DTXSID0029632
CASRN: 1309-37-1
TOXCAST: 0

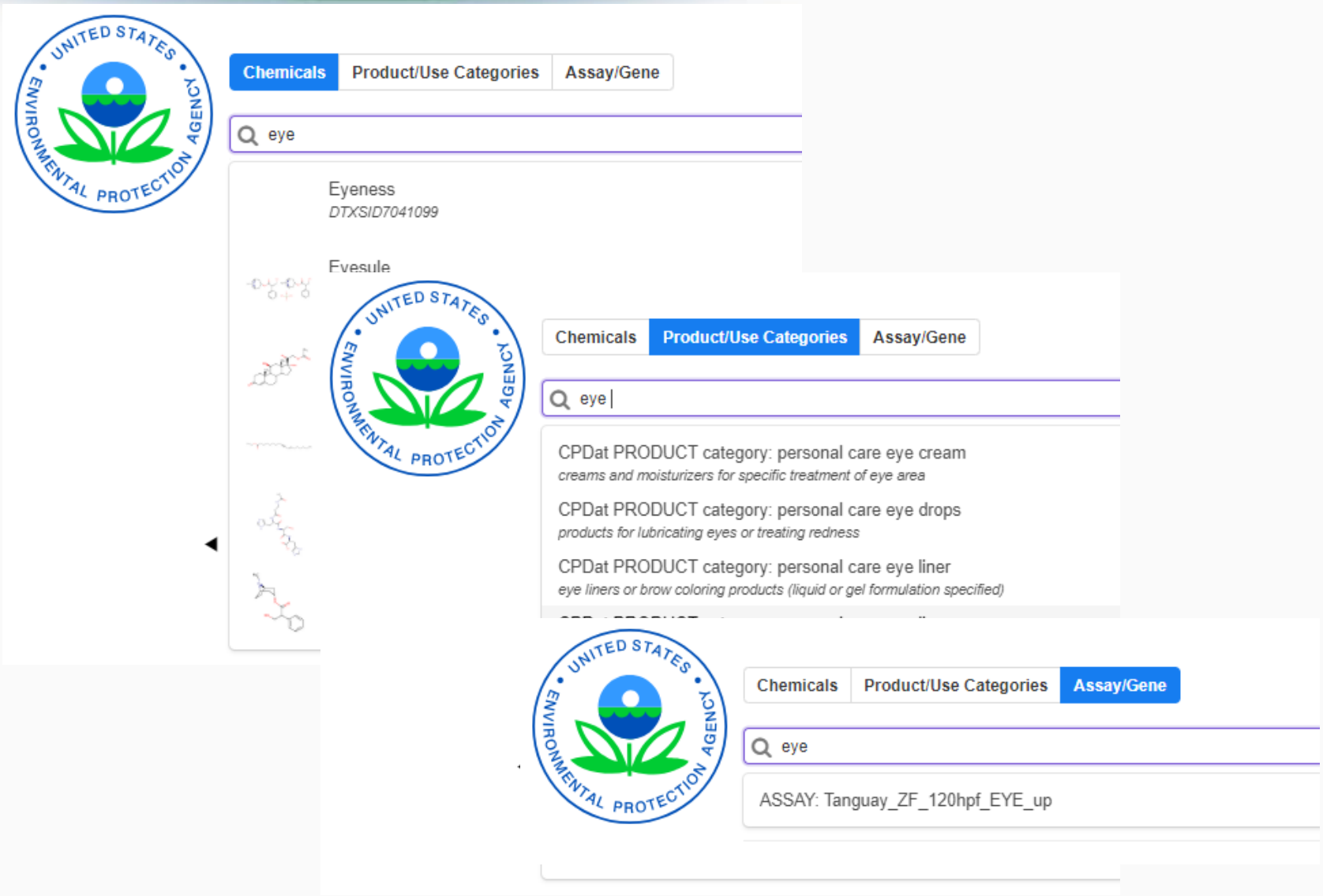


2,2,4,6,6-Pentamethylheptane
DTXSID: DTXSID0042034
CASRN: 13475-82-6
TOXCAST: 0



Squalane
DTXSID: DTXSID0046513
CASRN: 111-01-3
TOXCAST: 2/109

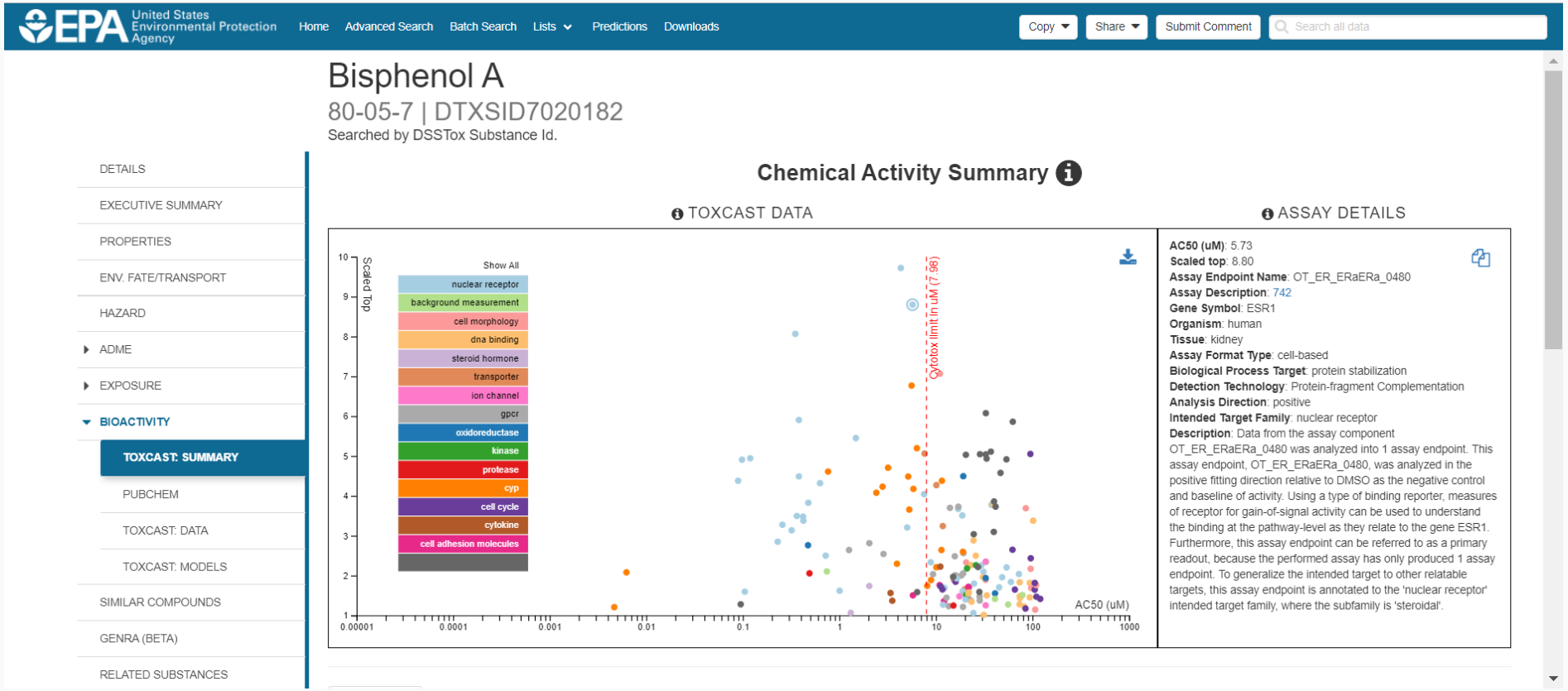
Remember home page searches
Searching for “eye”…



In Vitro Bioassay Screening

ToxCast and Tox21

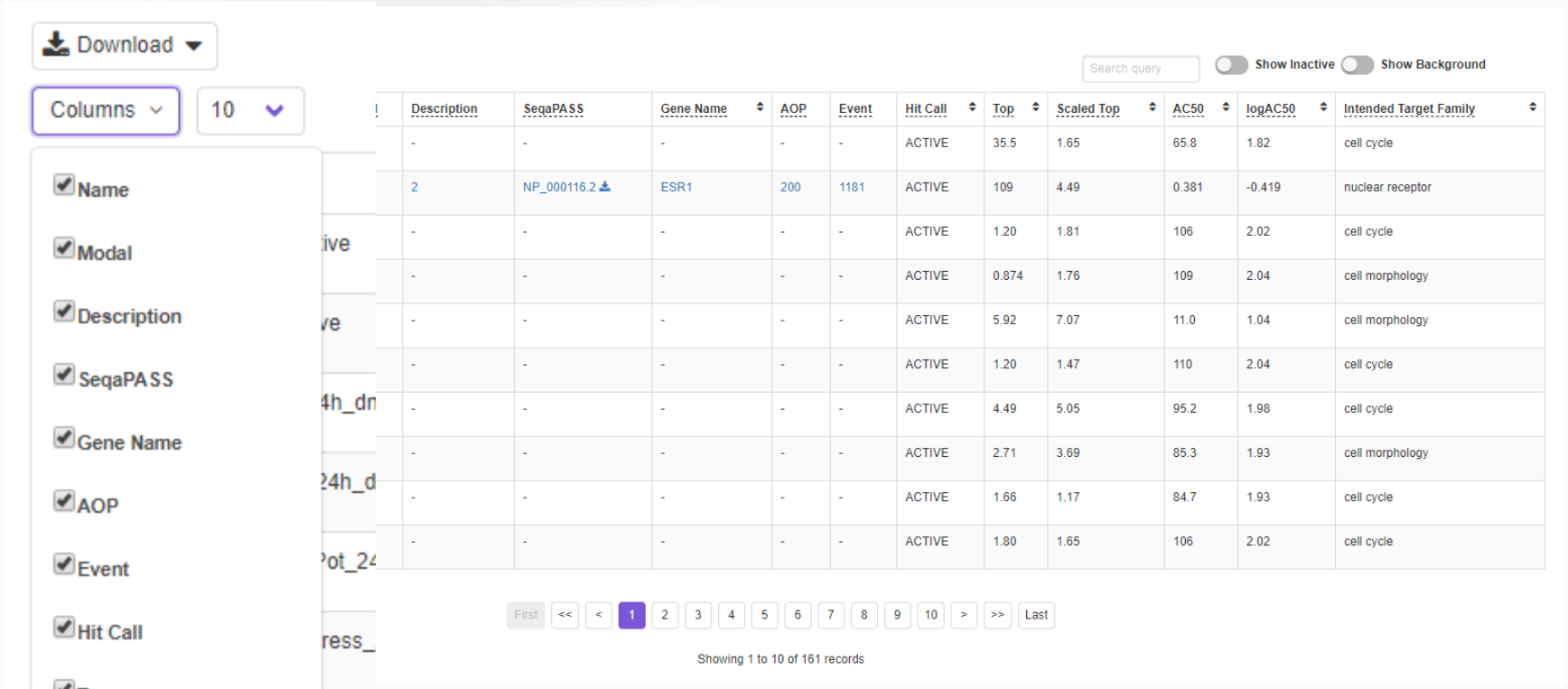
EPA United States Environmental Protection Agency



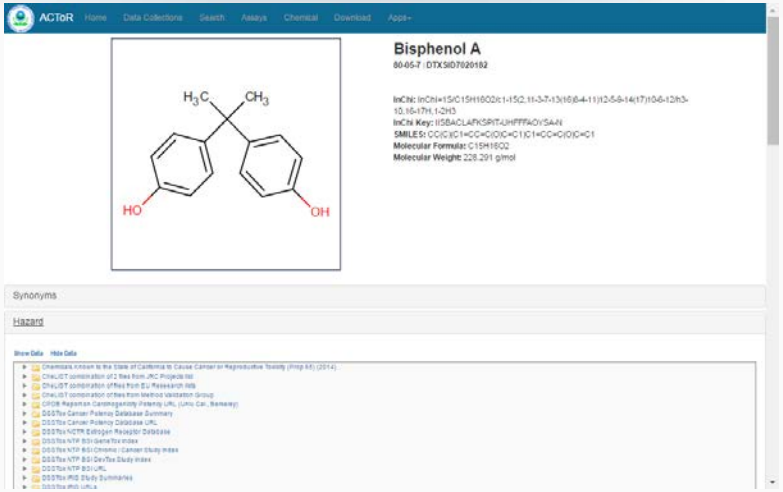
In Vitro Bioassay Screening

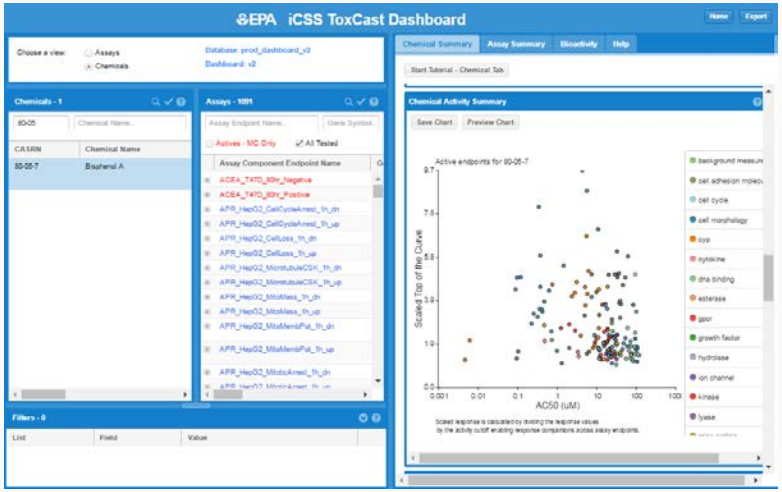
ToxCast and Tox21

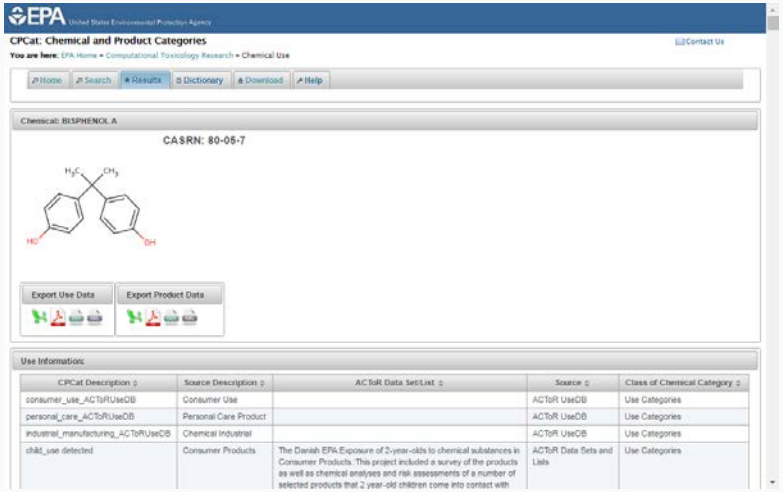
EPA United States Environmental Protection Agency

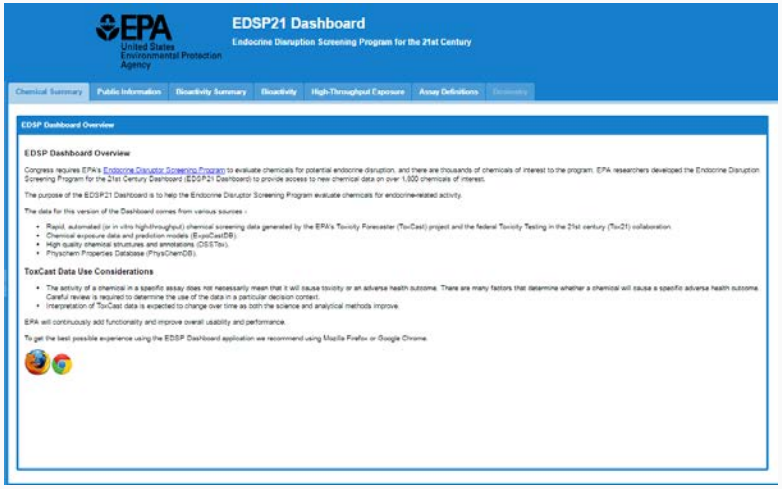


Earlier Dashboard Applications









In Vitro Bioassay Screening

ToxCast and Tox21

EPA United States Environmental Protection Agency

Download

Columns 10

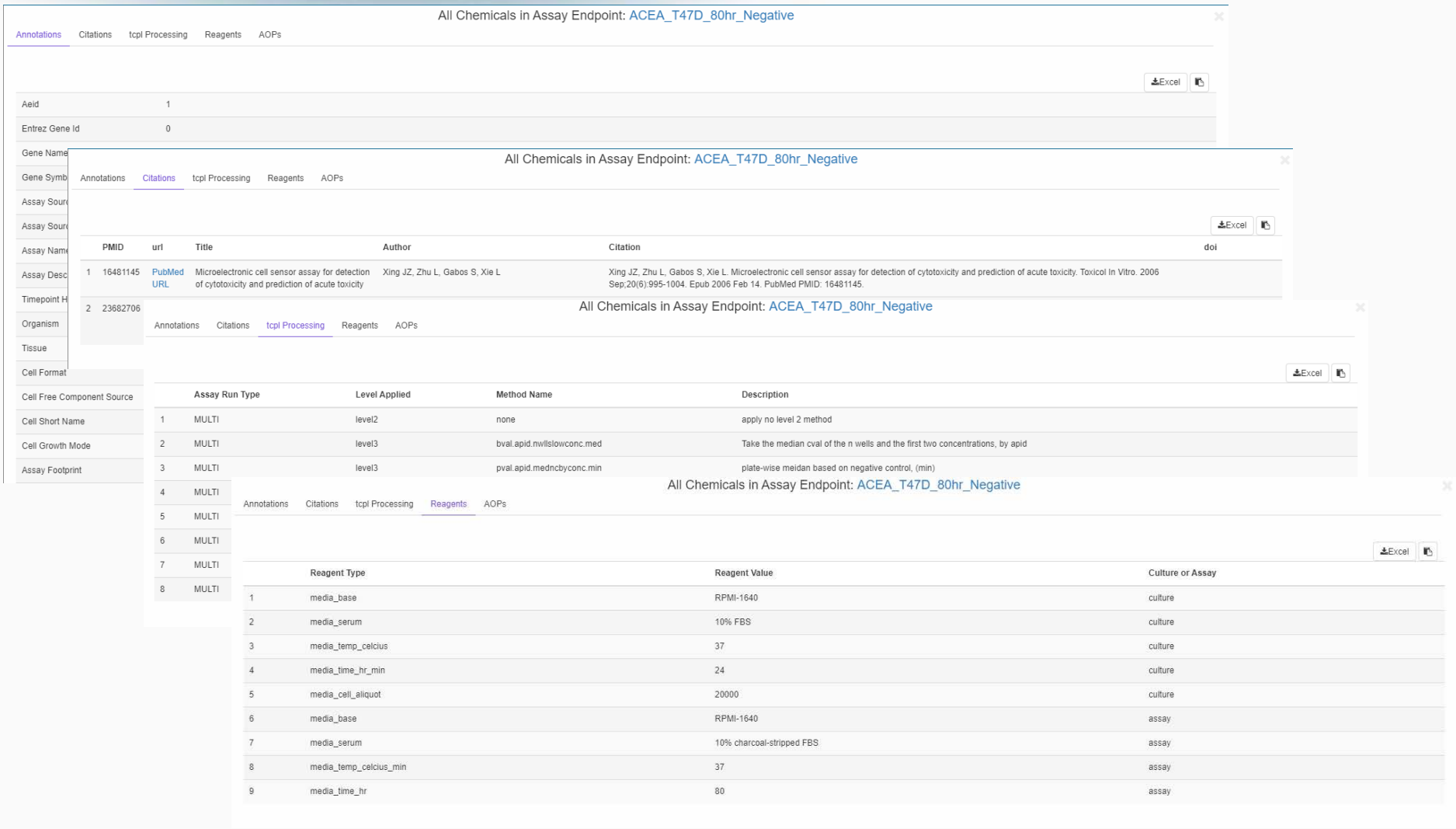
Search query | Show Inactive | Show Background

Name	Modal	Description	SeqaPASS	Gene Name	AOP	Event	Hit Call	Top	Scaled Top	AC50	logAC50	Intended Target Family
ACEA_T47D_80hr_Negative			-	-	-	-	ACTIVE	35.5	1.65	65.8	1.82	cell cycle
ACEA_T47D_80hr_Positive			NP_000116.2	ESR1	200	1181	ACTIVE	109	4.49	0.381	-0.419	nuclear receptor
APR_HepG2_CellLoss_24h_dn			-	-	-	-	ACTIVE	1.20	1.81	106	2.02	cell cycle
APR_HepG2_MitoMass_24h_dn			-	-	-	-	ACTIVE	0.874	1.76	109	2.04	cell morphology
APR_HepG2_MitoMembPot_24h_dn			-	-	-	-	ACTIVE	5.92	7.07	11.0	1.04	cell morphology
APR_HepG2_OxidativeStress_24h_up			-	-	-	-	ACTIVE	1.20	1.47	110	2.04	cell cycle
APR_HepG2_CellLoss_72h_dn			-	-	-	-	ACTIVE	4.49	5.05	95.2	1.98	cell cycle
APR_HepG2_MitoMembPot_72h_dn			-	-	-	-	ACTIVE	2.71	3.69	85.3	1.93	cell morphology
APR_HepG2_MitoticArrest_72h_up			-	-	-	-	ACTIVE	1.66	1.17	84.7	1.93	cell cycle
APR_HepG2_OxidativeStress_72h_up			-	-	-	-	ACTIVE	1.80	1.65	106	2.02	cell cycle

First << < 1 2 3 4 5 6 7 8 9 10 > >> Last

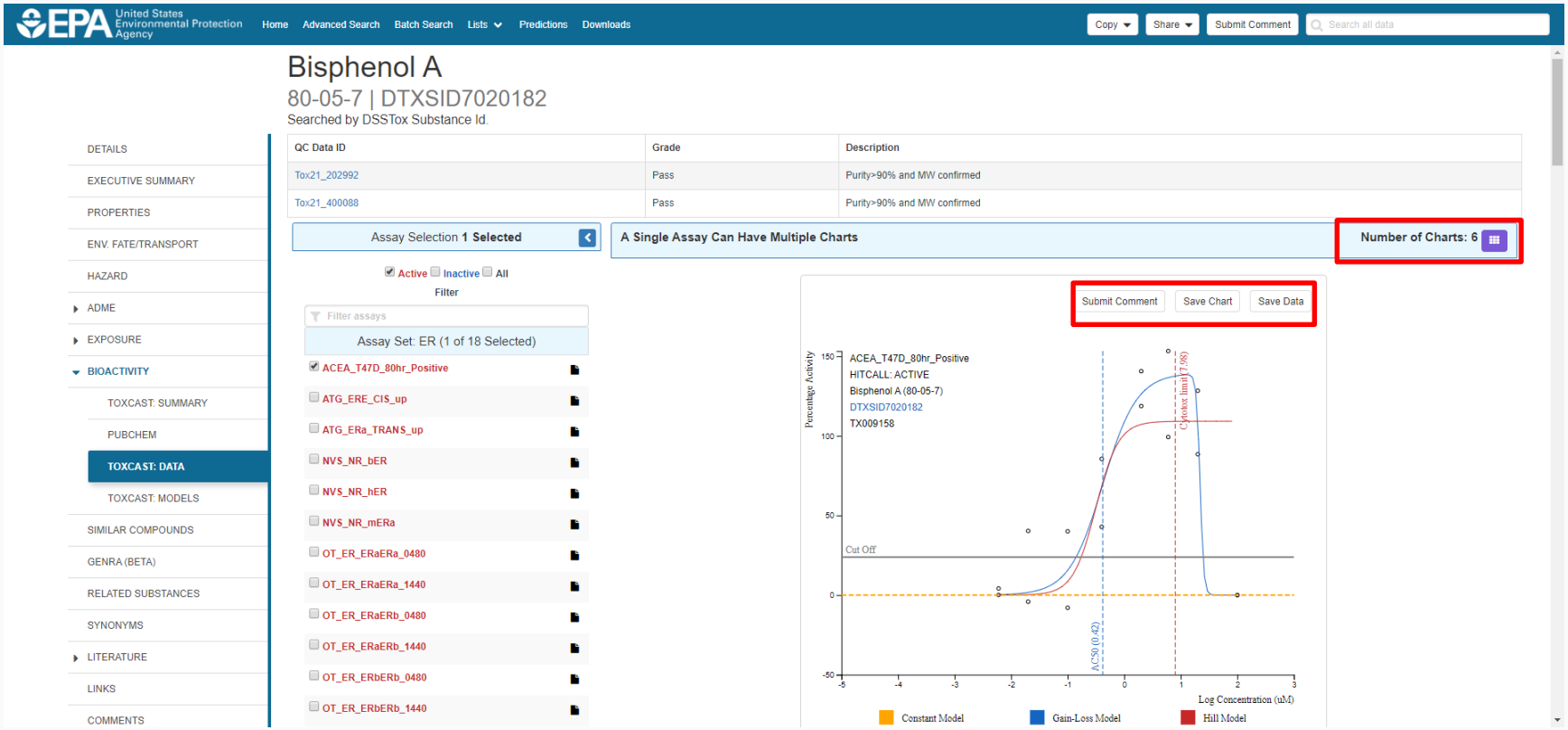
Showing 1 to 10 of 161 records

Assay Modal Details



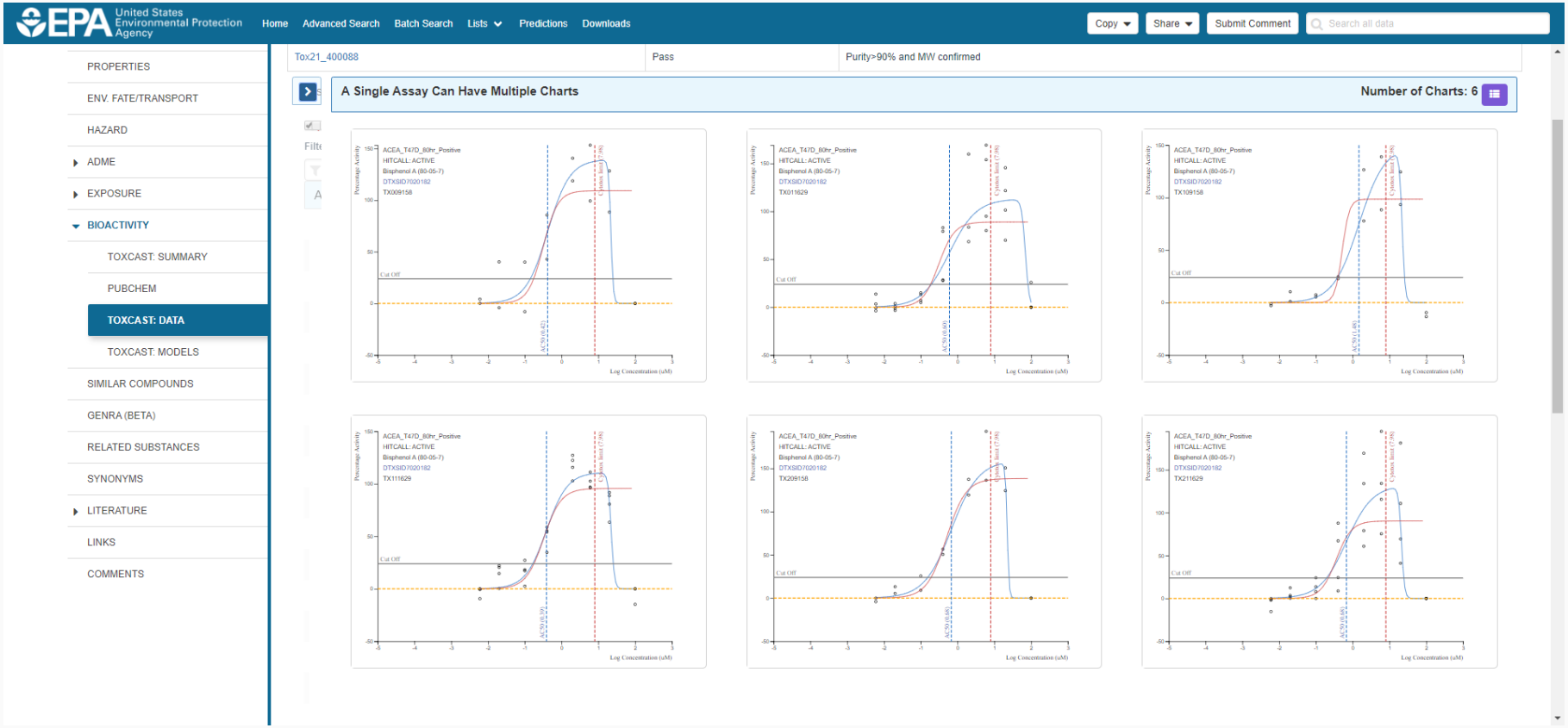
In Vitro Bioassay Screening

ToxCast and Tox21



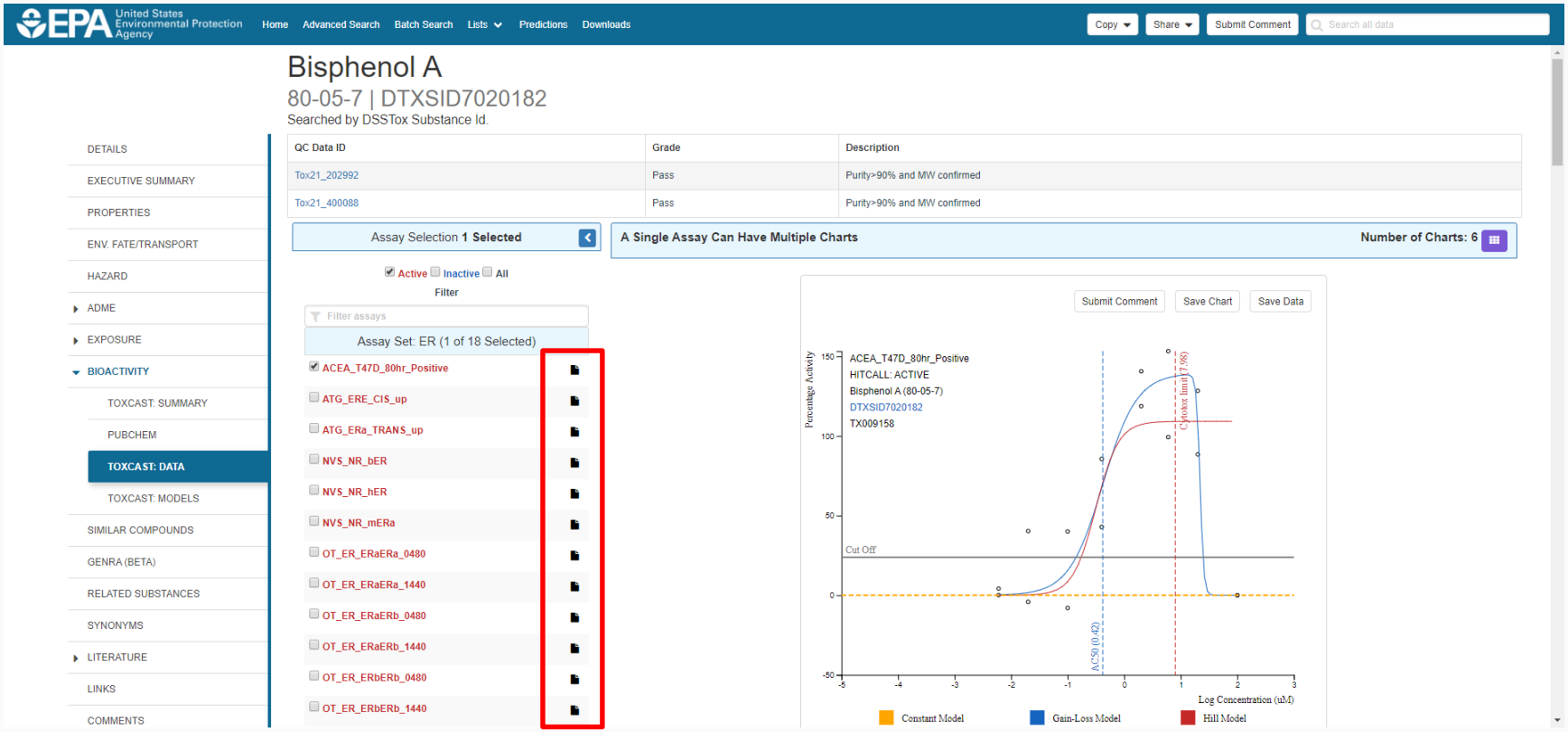
In Vitro Bioassay Screening

Multi-chart Display

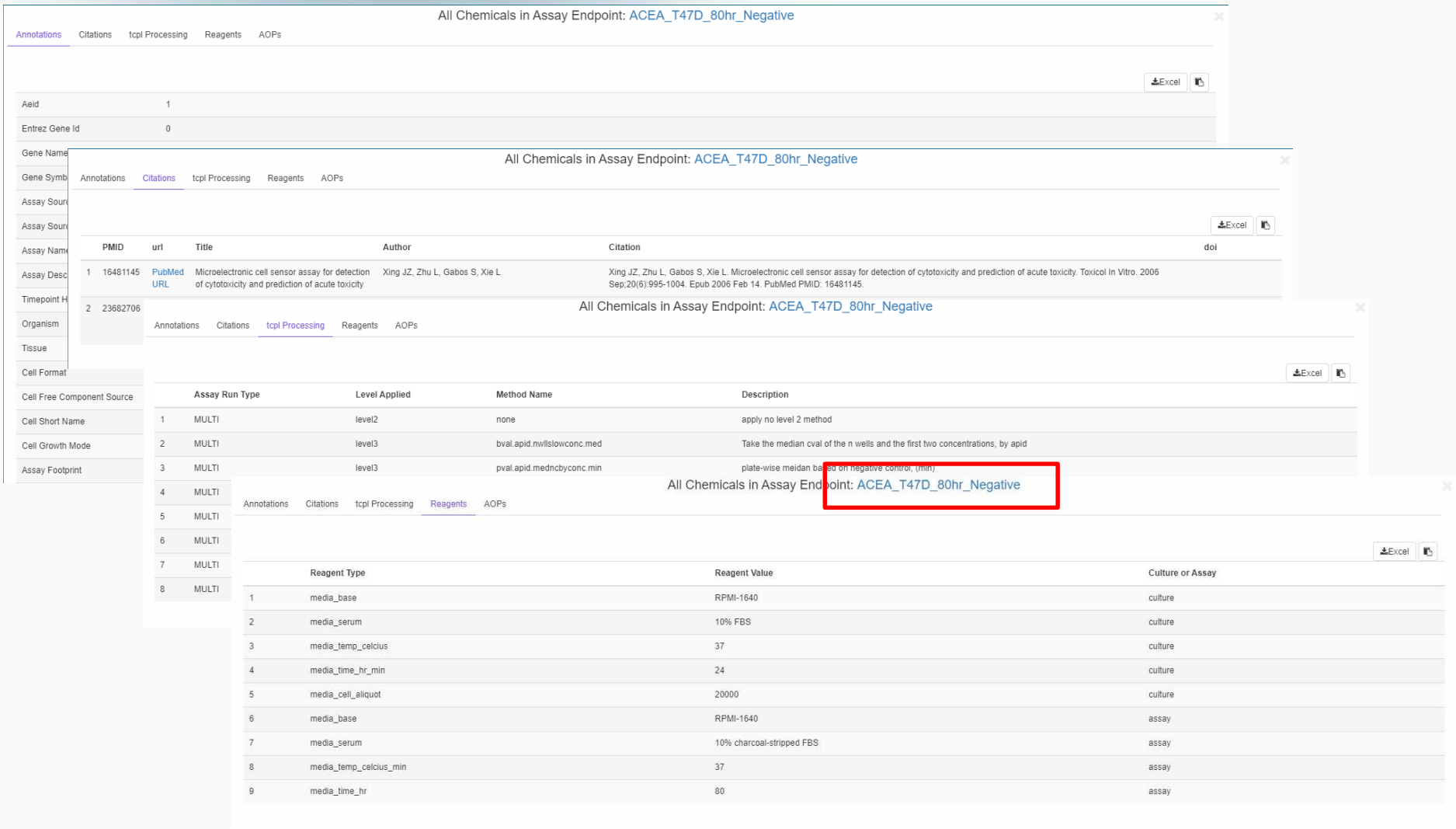


In Vitro Bioassay Screening

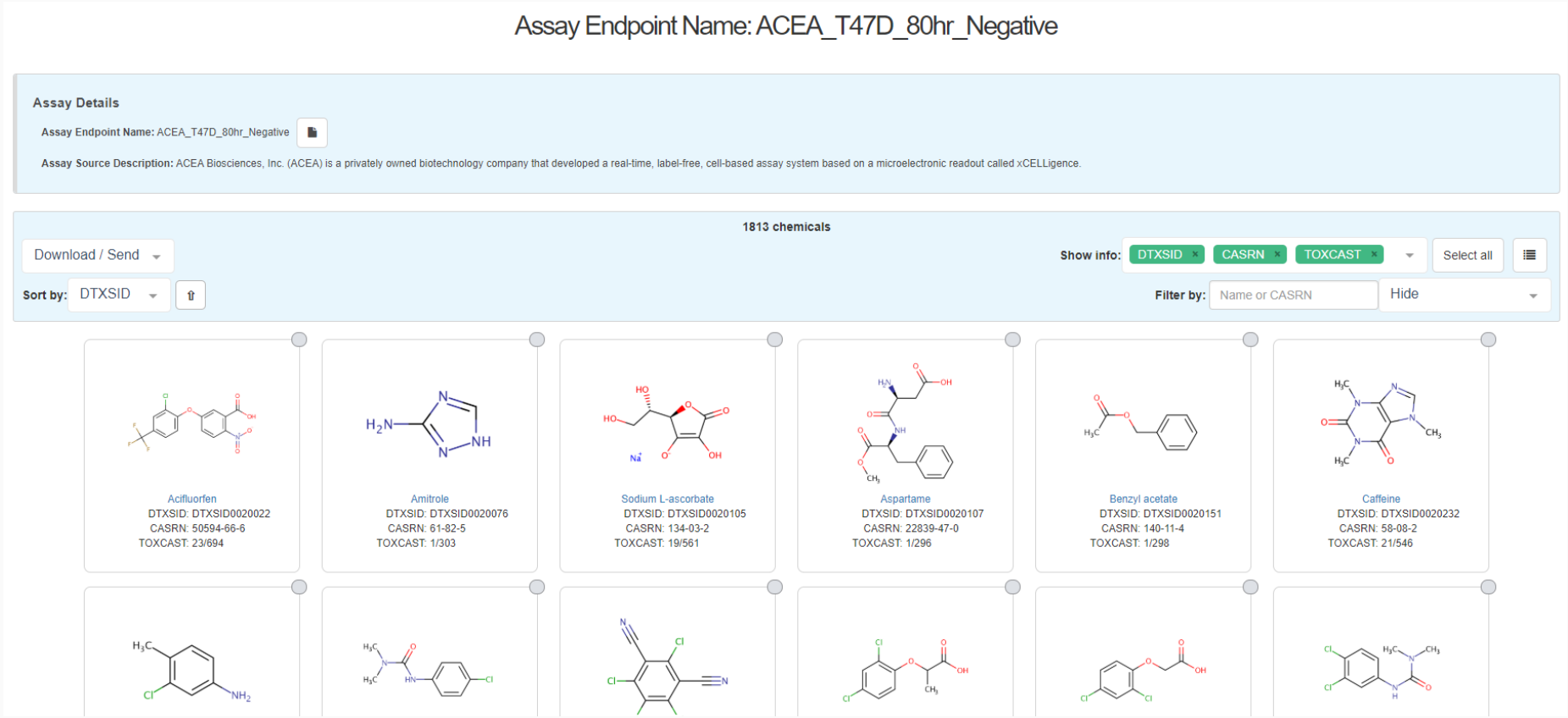
ToxCast and Tox21



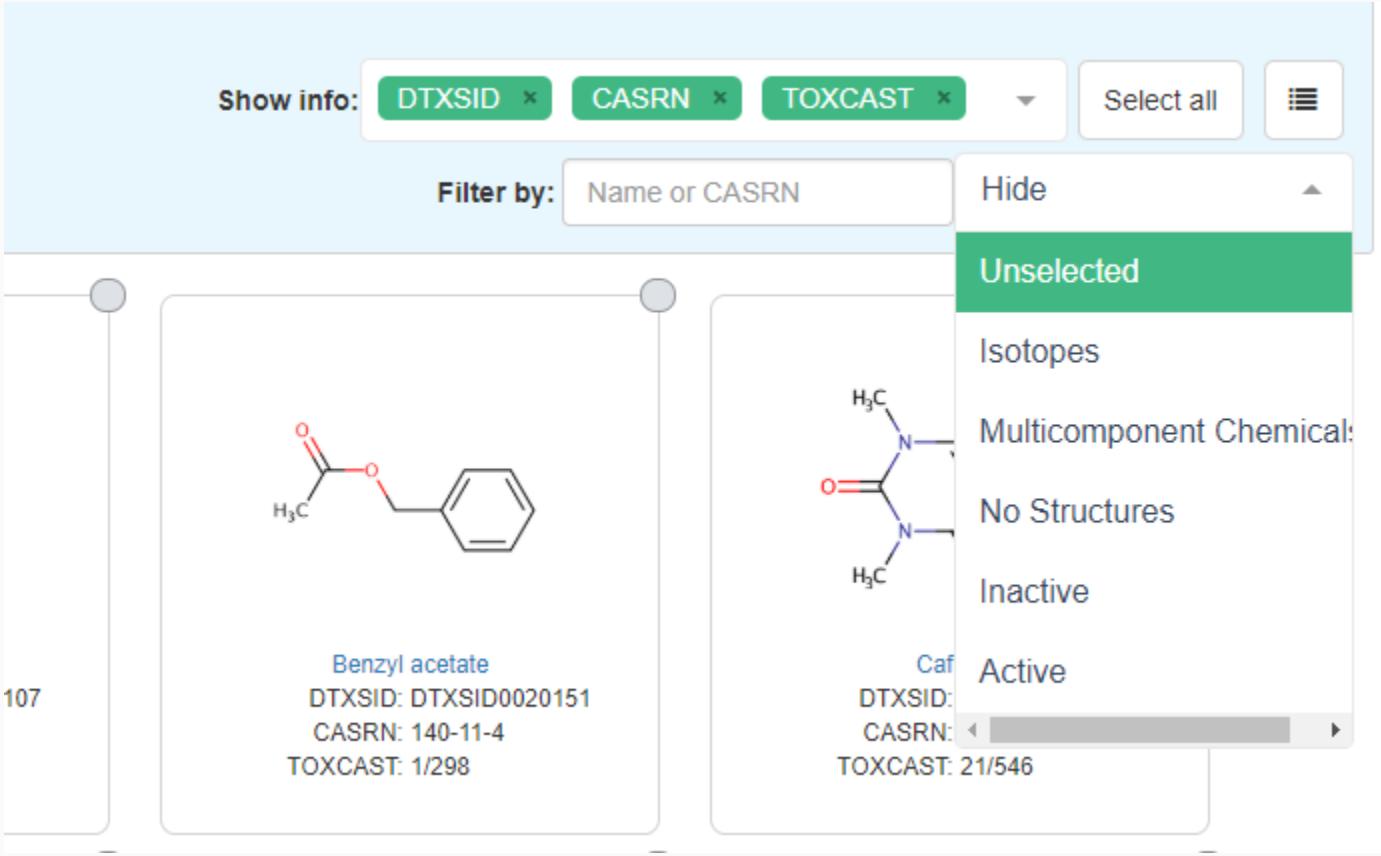
Assay Modal Details



List of Chemicals for an Assay



Choose Display Details

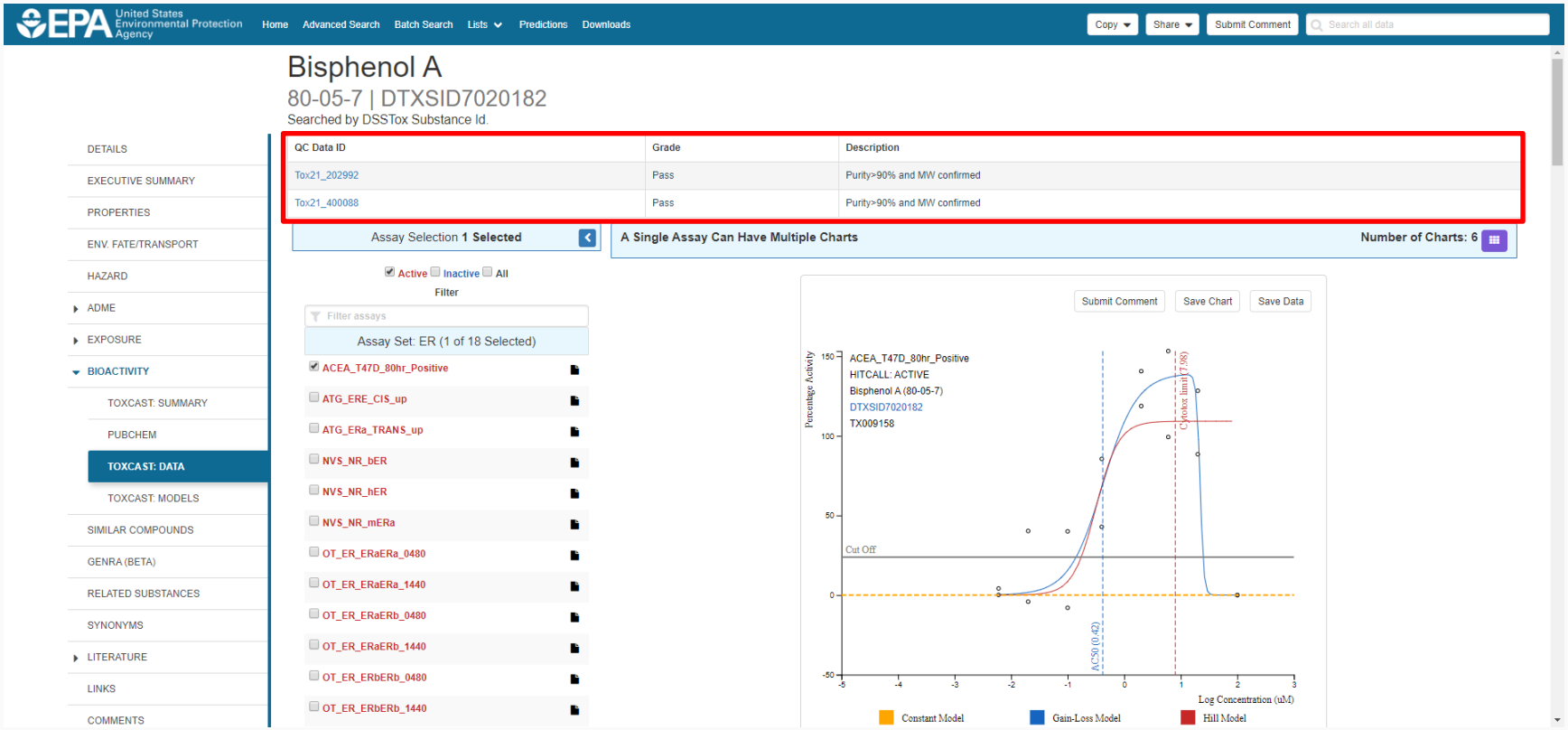


Tile/Table Mode
More flexibility in table display

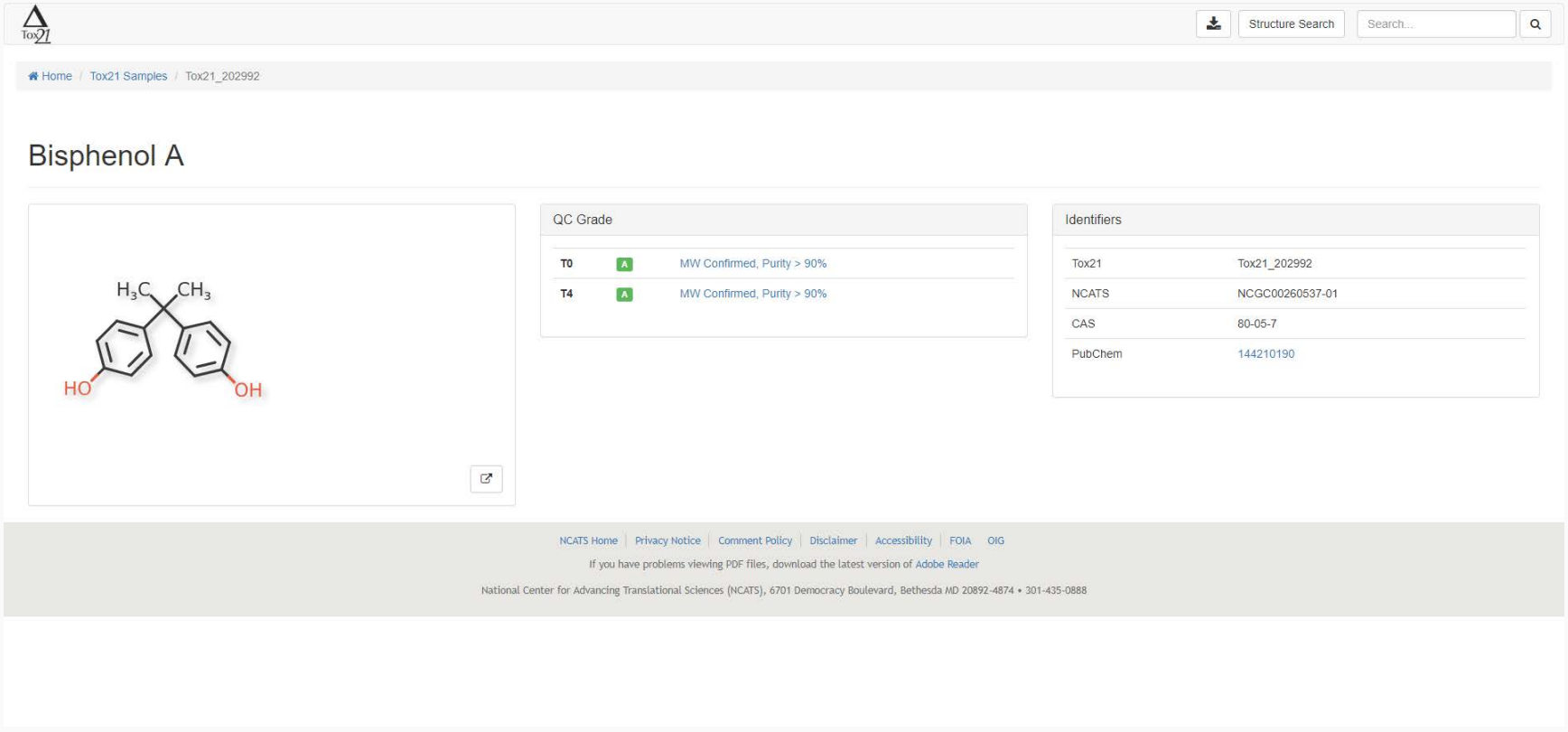
EPA United States Environmental Protection Agency

In Vitro Bioassay Screening

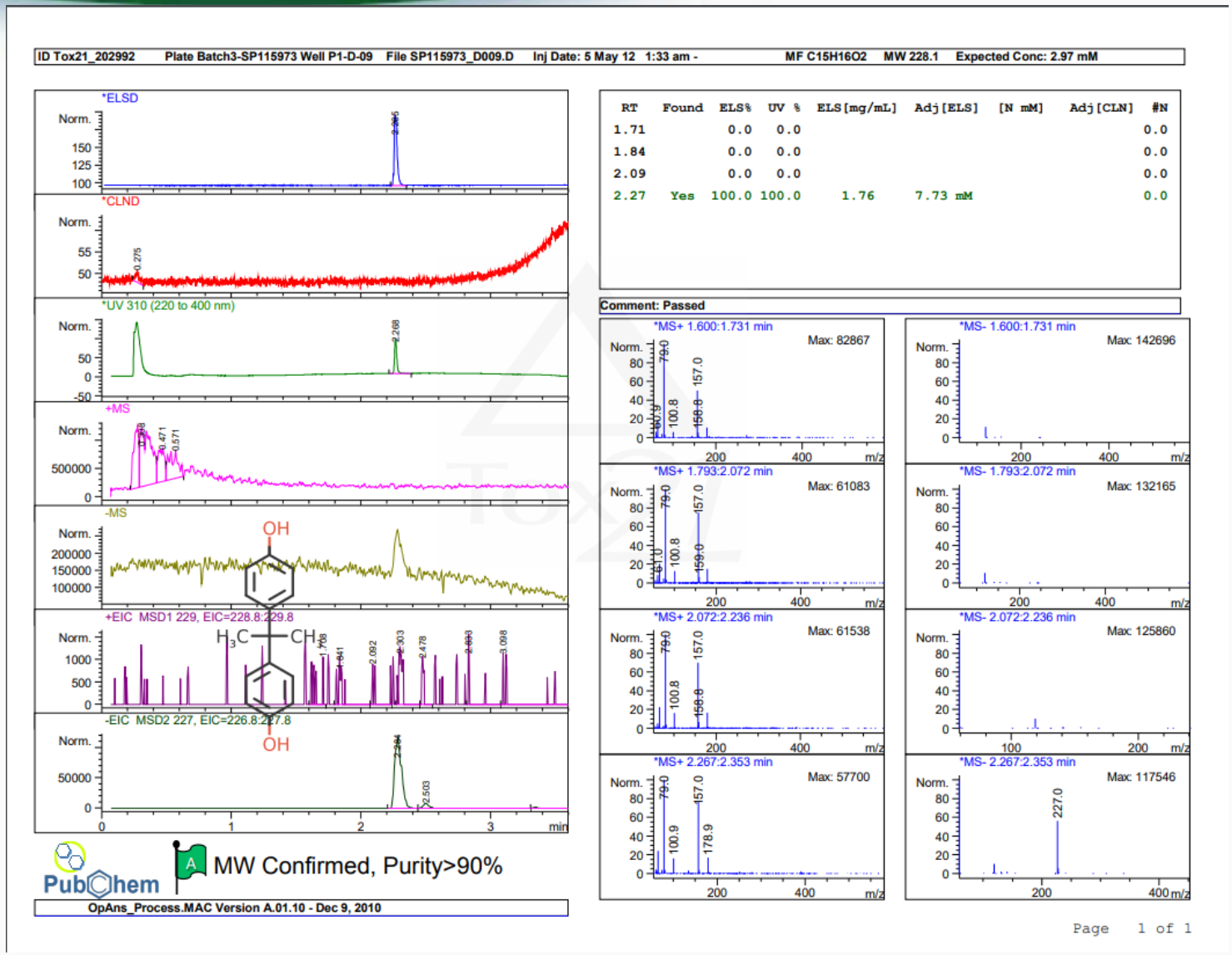
ToxCast and Tox21



Access to Analytical QC Data

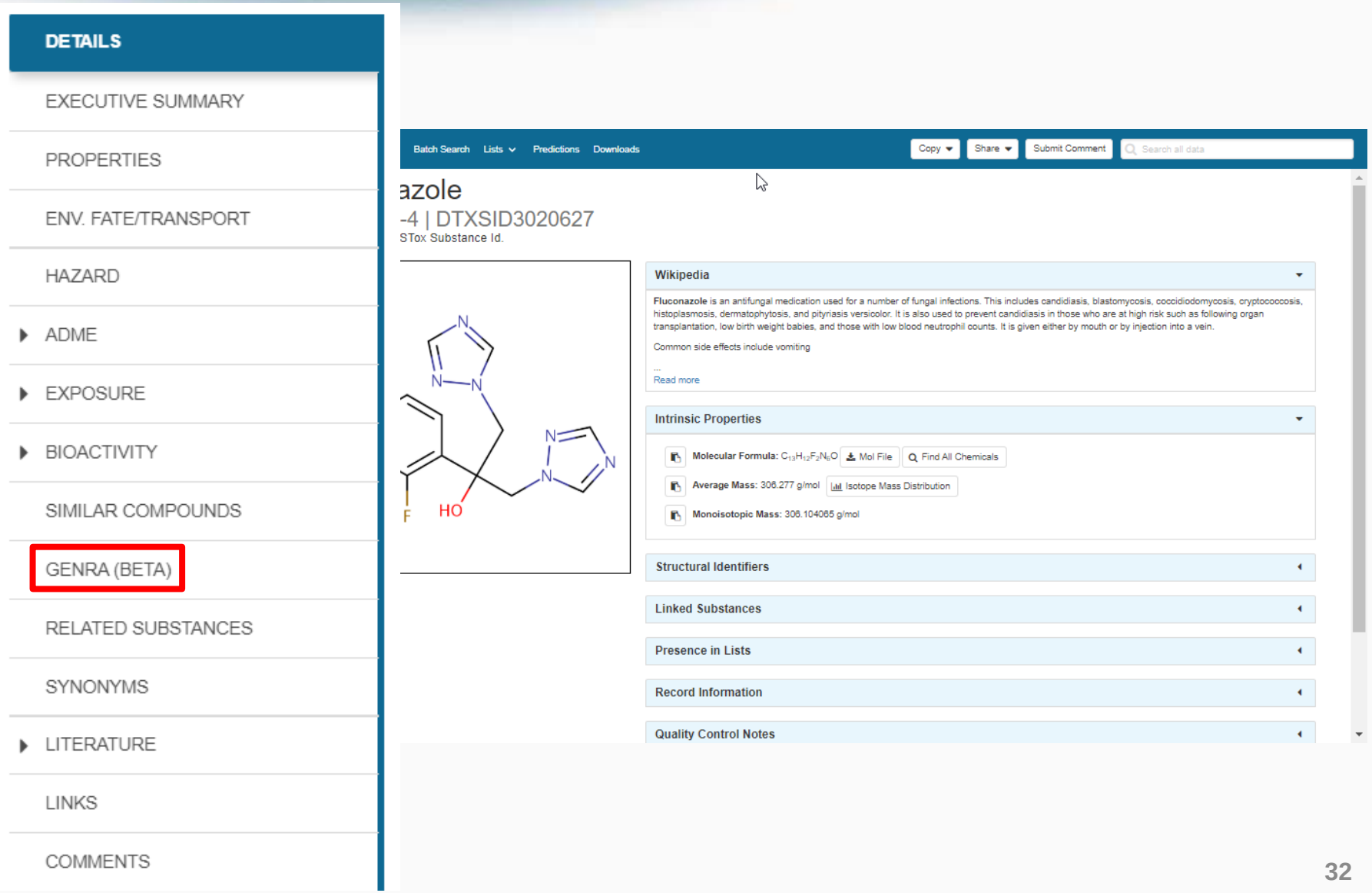


Access to Analytical QC Data

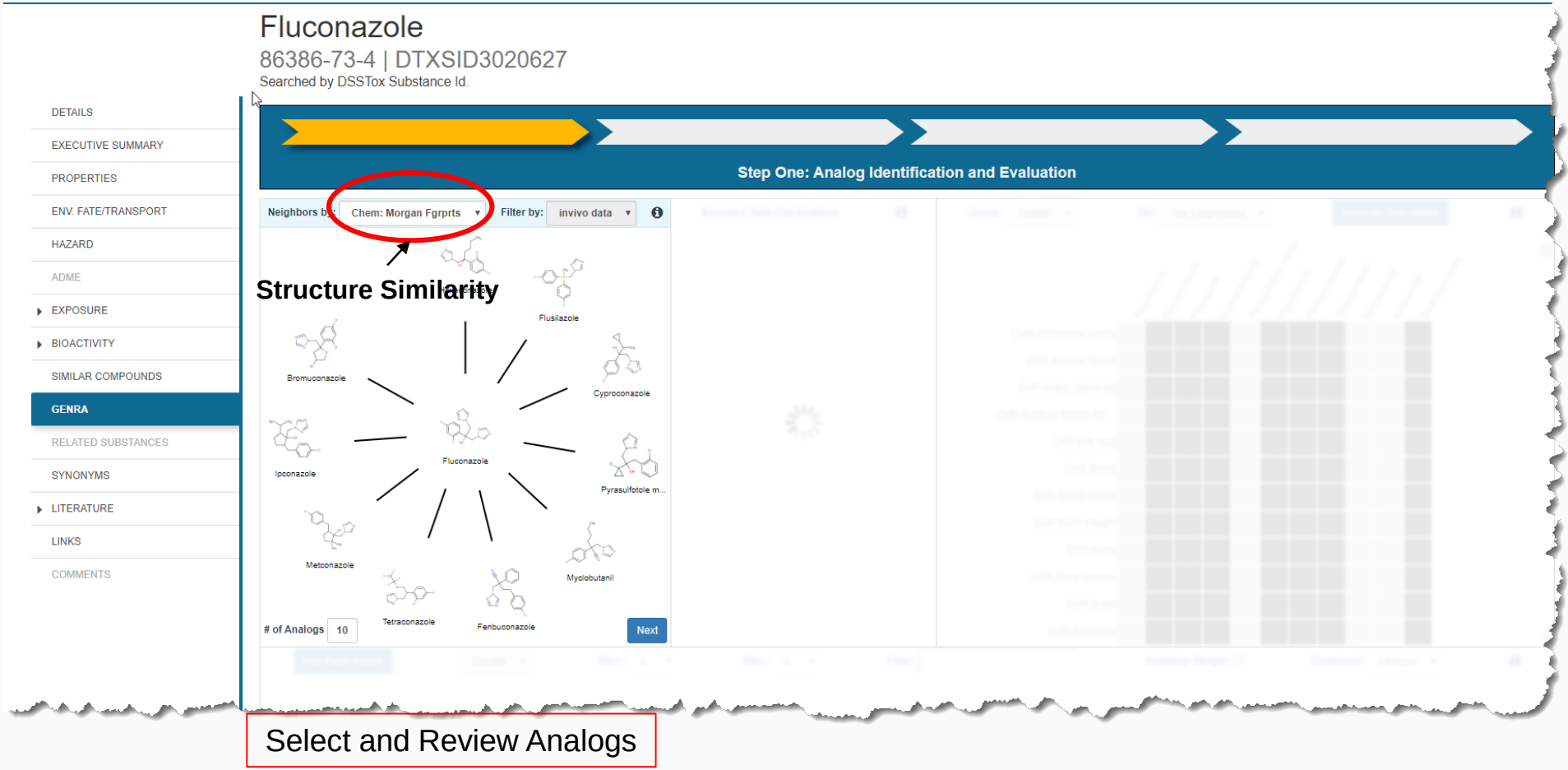


GenRA (Generalised Read-Across)

EPA United States Environmental Protection Agency

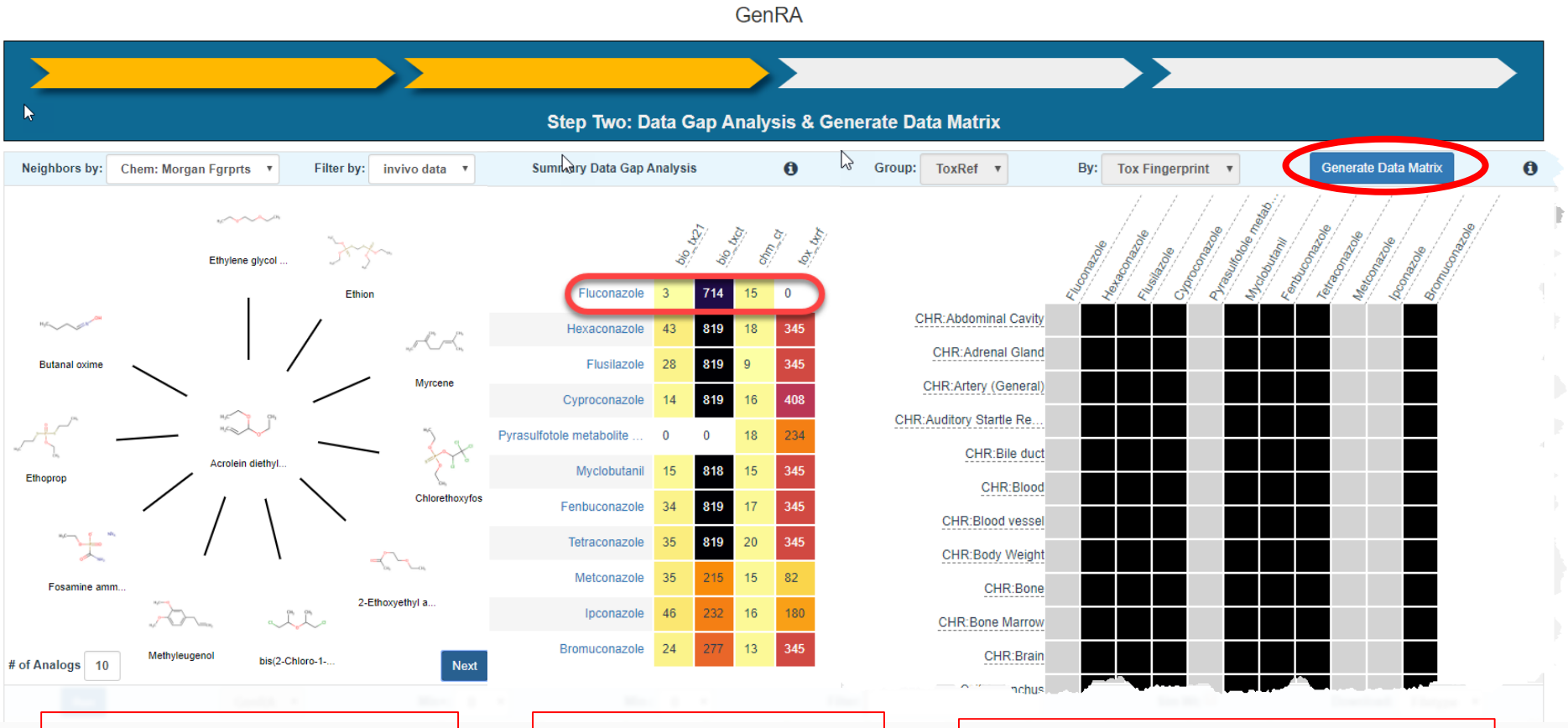


GenRA (Generalised Read-Across)



Select and Review Analogs

GenRA (Generalised Read-Across)

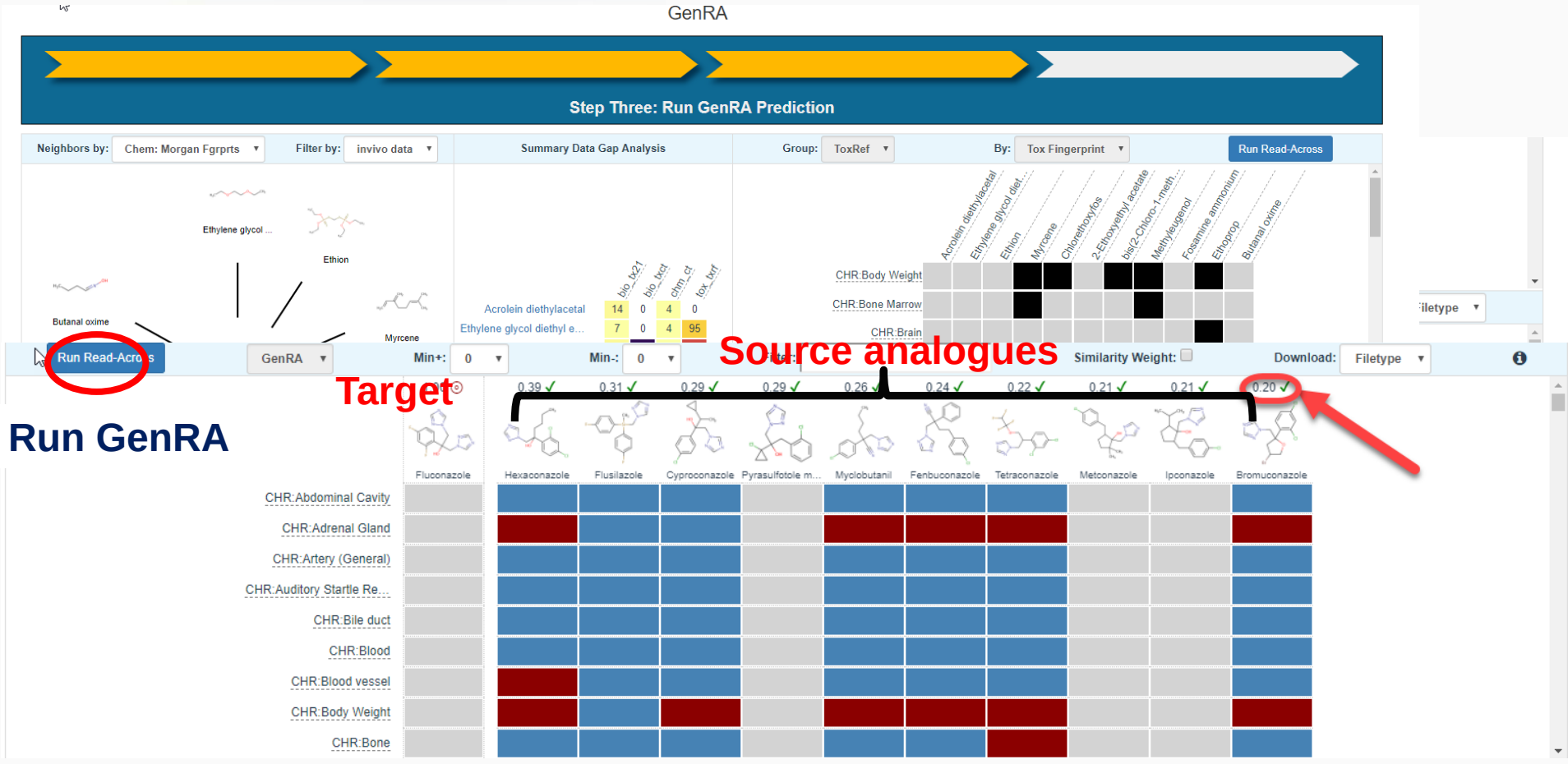


Select and Review Analogs

Review Available Data

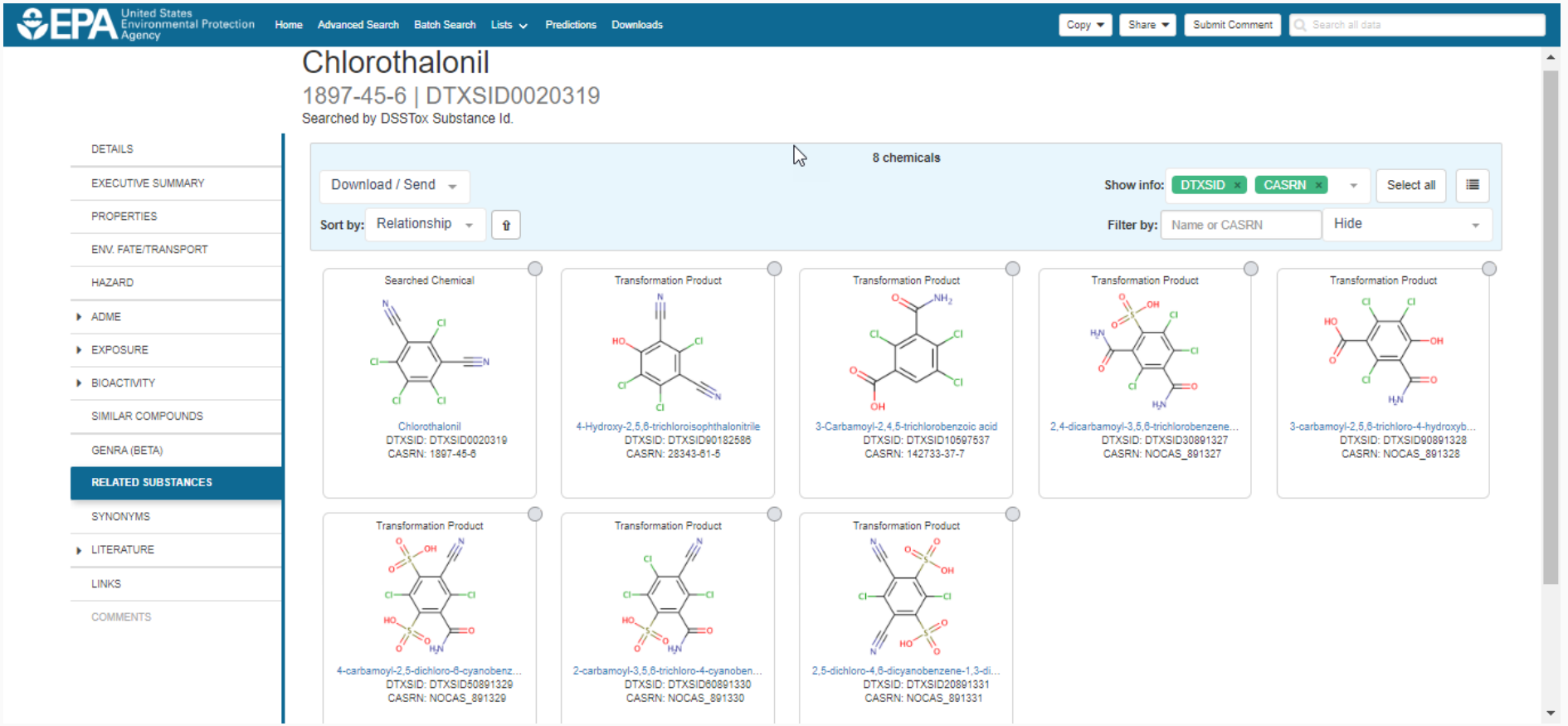
Fingerprint indicating available data

GenRA (Generalised Read-Across)

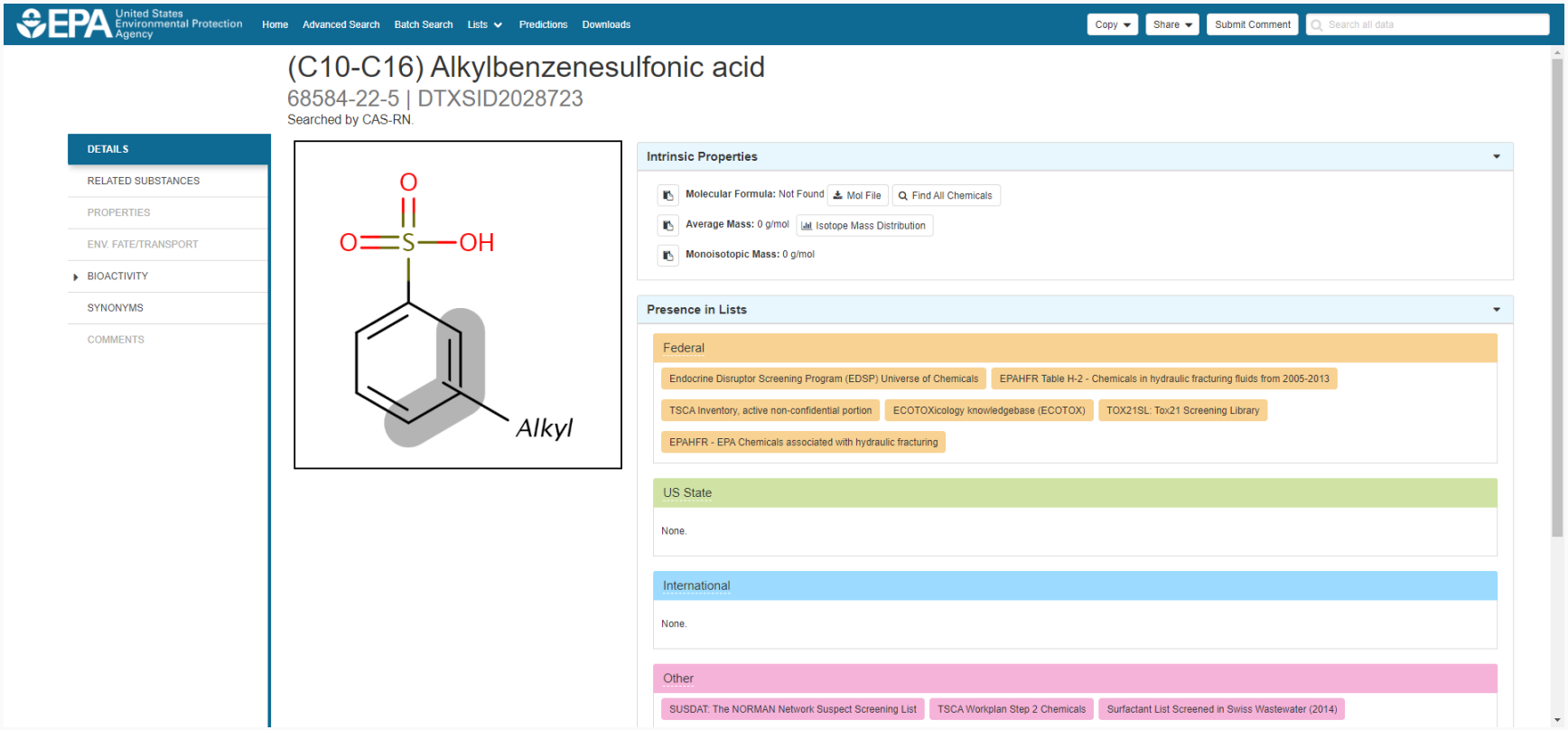


Red : Toxicity effects.
Blue: No Toxicity effects
Grey : Absence of data

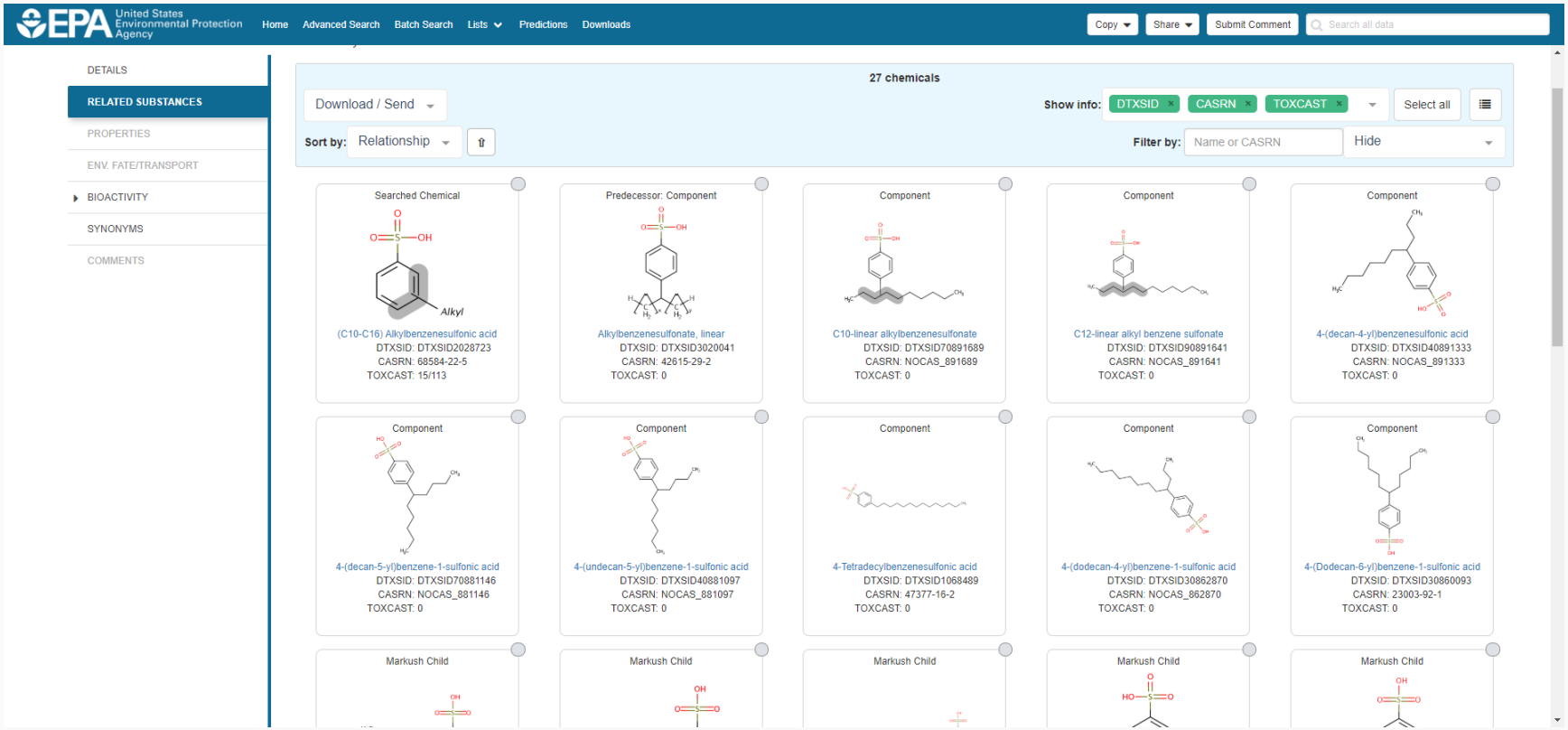
Related Substances
e.g. Transformation Products



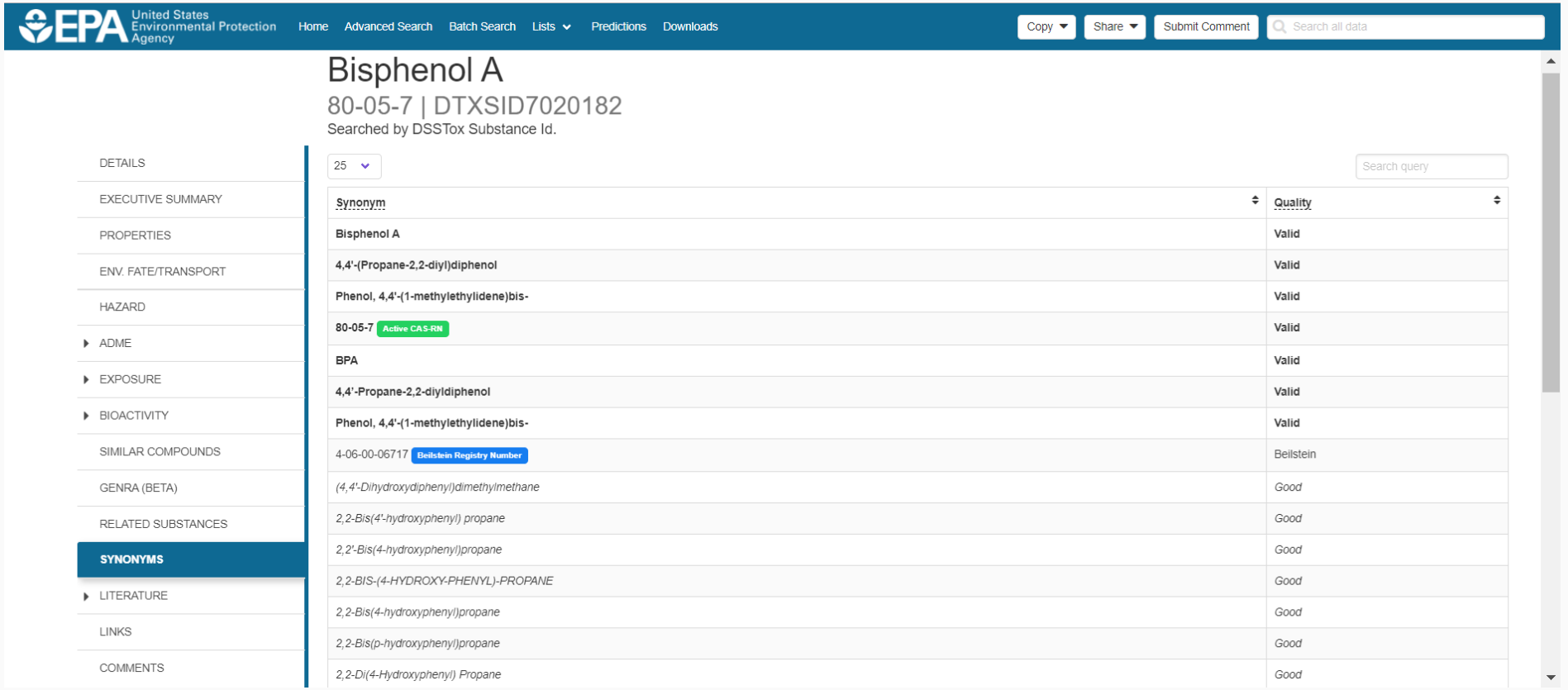
UVCB Chemicals



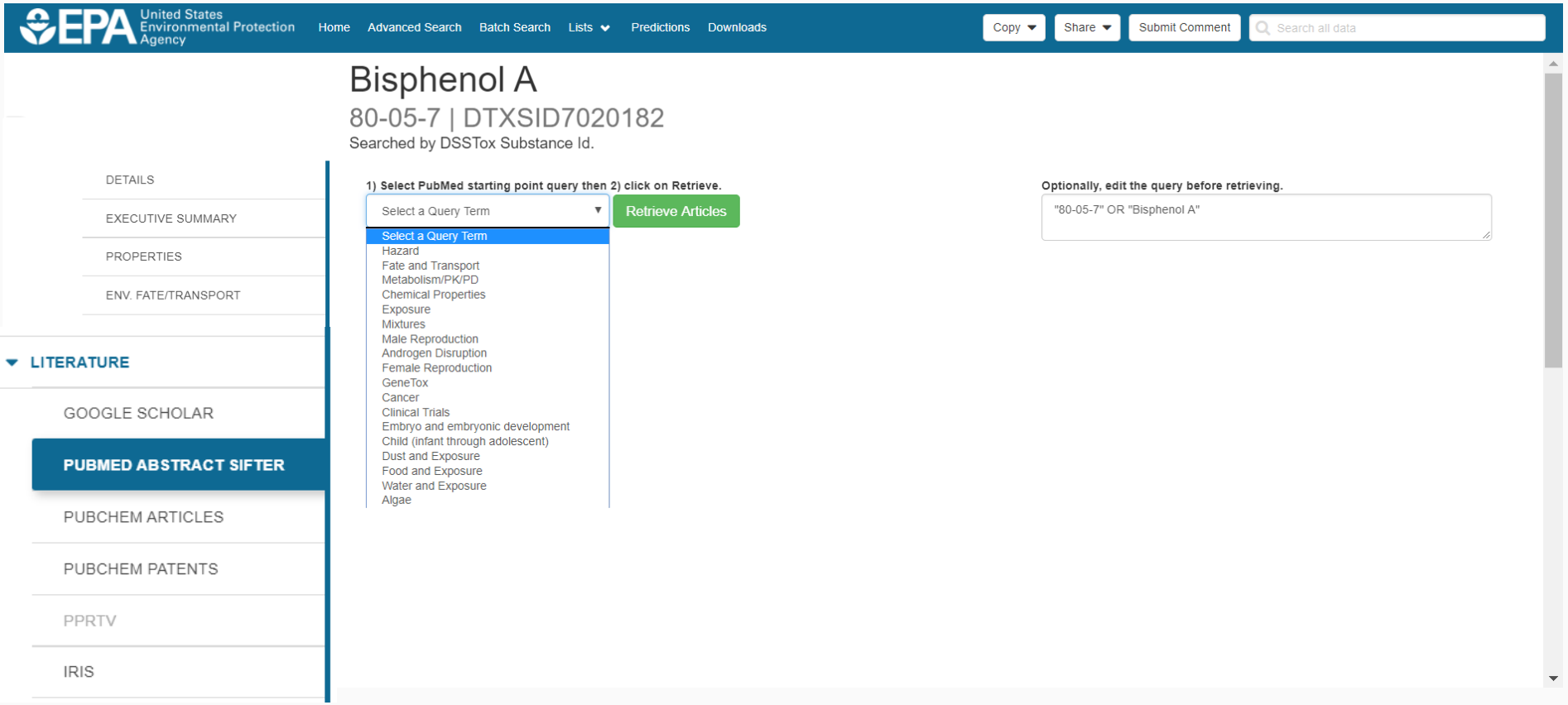
Related Substances for Markush



Identifiers to Support Searches

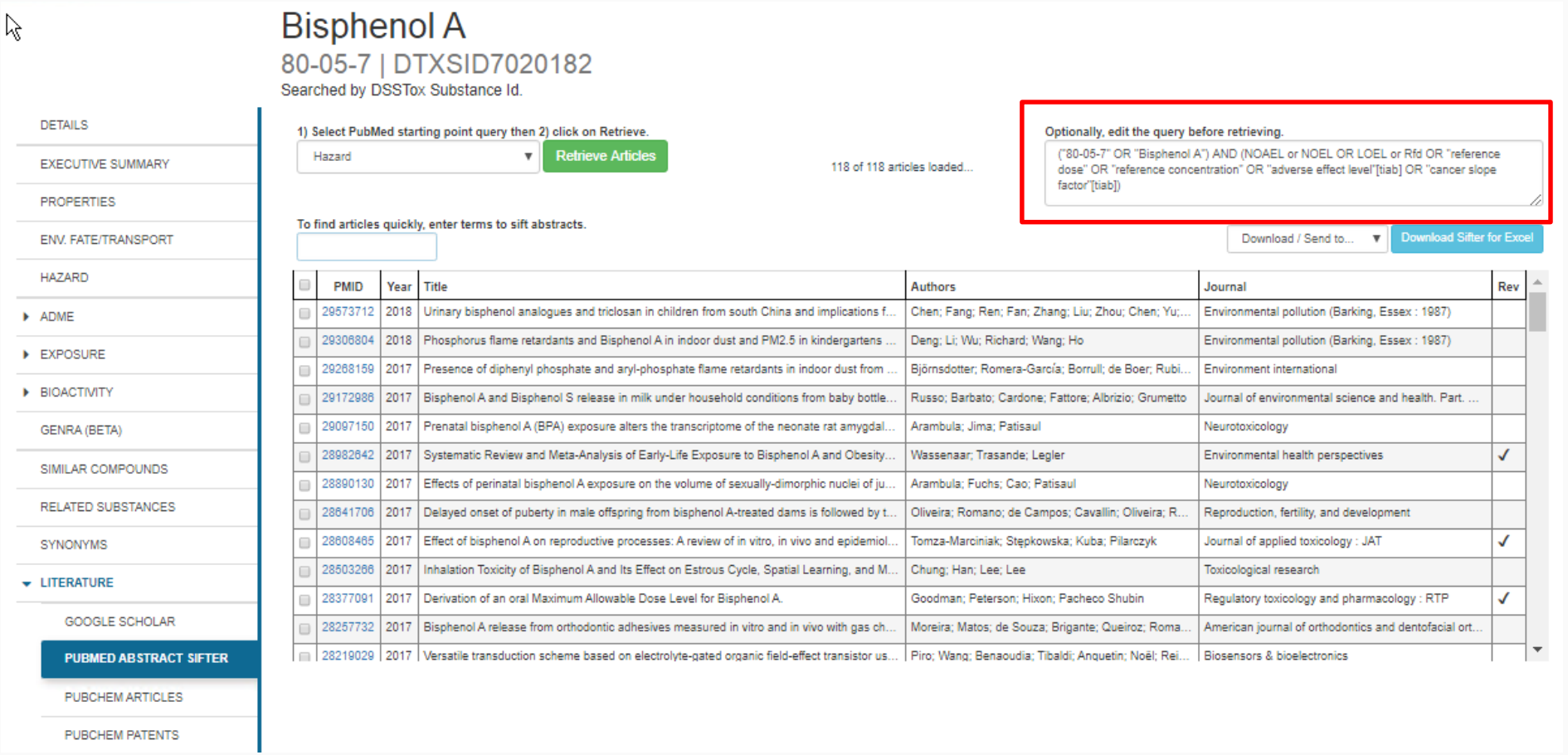


Literature Searches and Links

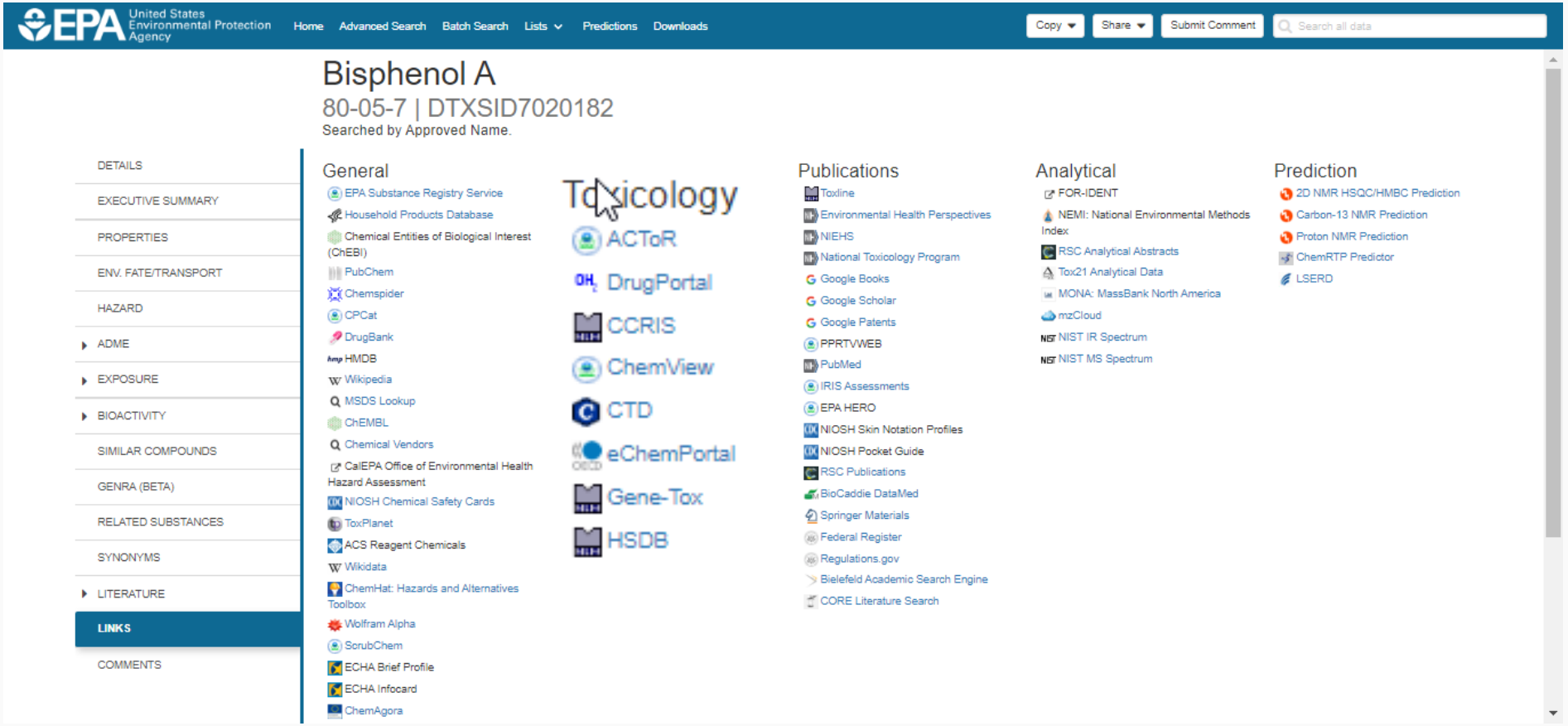


Abstract Sifter – PubMed Integration searching >28 million abstracts

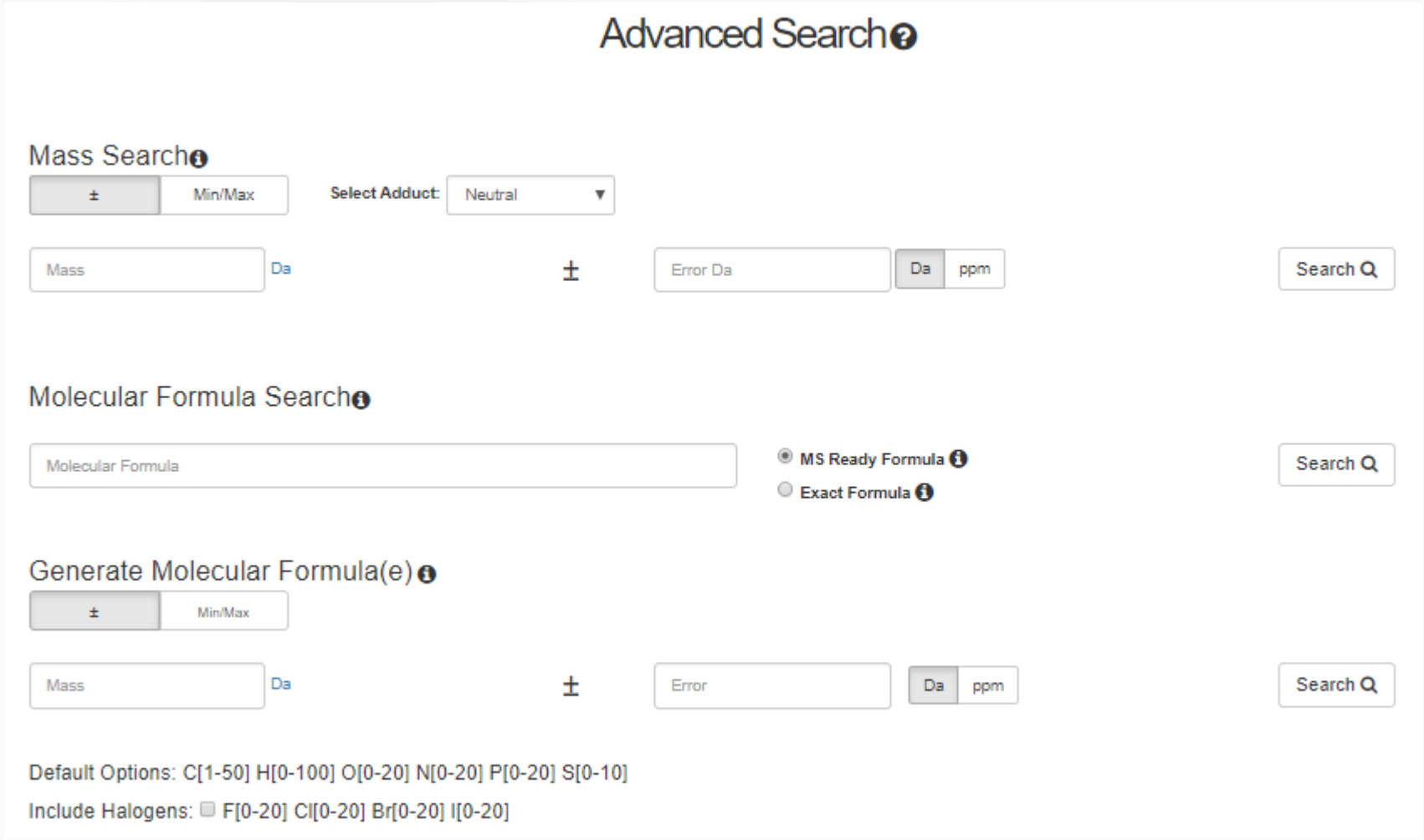
EPA United States Environmental Protection Agency



External Links to ~80 websites
Growing list of out links -



Mass and Formula Searches Supporting Mass Spectrometry



Advanced Searches
Mass Based Search

EPA United States Environmental Protection Agency

Mass Search

± | Min/Max

M

191.131 Da ± 5 Da | ppm

Search

EPA United States Environmental Protection Agency | Home | Advanced Search | Batch Search | Lists | Predictions | Downloads | Share | Search all data

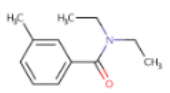
Search Results

Searched by Mass: 191.131 +/- 5.0 ppm.

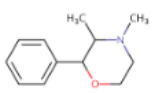
298 of 298 chemicals visible

Download / Send | Show info: DTXSID × CASRN × TOXCAST × Mass Diff × | Select all

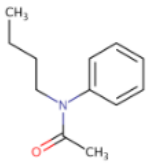
Sort by: Mass Difference | Filter by: Name or CASRN | Multicomponent Chemicals ×



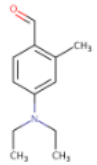
DEET
DTXSID: DTXSID2021995
CASRN: 134-62-3
TOXCAST: 14/663
Mass Diff: 0.000014



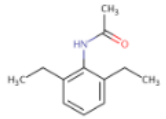
Phendimetrazine
DTXSID: DTXSID1023447
CASRN: 634-03-7
TOXCAST: 0
Mass Diff: 0.000014



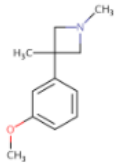
N-Butylacetanilide
DTXSID: DTXSID2042197
CASRN: 91-49-6
TOXCAST: 0
Mass Diff: 0.000014



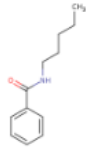
Benzaldehyde, 4-(diethylamino)-2-methyl-
DTXSID: DTXSID4059041
CASRN: 92-14-8
TOXCAST: 0
Mass Diff: 0.000014



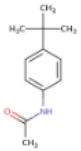
Acetanilide, 2',6'-diethyl-
DTXSID: DTXSID90168148
CASRN: 16665-89-7
TOXCAST: 0
Mass Diff: 0.000014



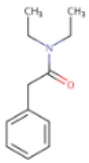
Azetidine, 1,3-dimethyl-3-(m-methoxyp...
DTXSID: DTXSID40173560
CASRN: 19832-26-9
TOXCAST: 0
Mass Diff: 0.000014



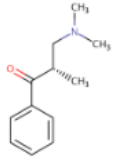
Benzamide, N-pentyl-
DTXSID: DTXSID20174196
CASRN: 20308-43-4



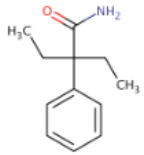
p-t-Butylacetanilide
DTXSID: DTXSID80174238
CASRN: 20330-45-4



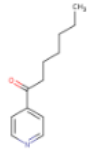
N,N-Diethylphenylacetamide
DTXSID: DTXSID00179048
CASRN: 2431-96-1



3-(Dimethylamino)-2-methylpropiophen...
DTXSID: DTXSID60180796
CASRN: 26171-50-6

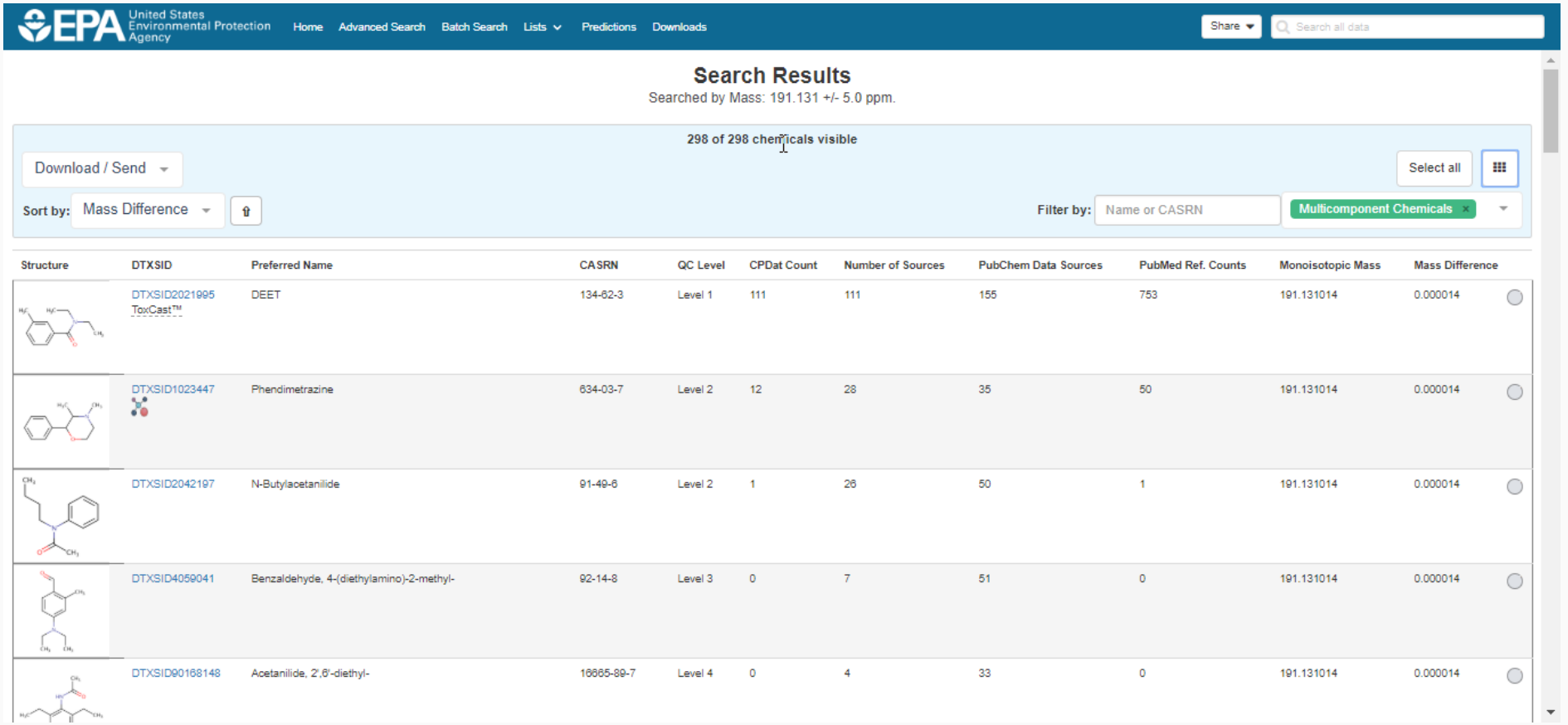


Butyramide, 2-ethyl-2-phenyl-
DTXSID: DTXSID60184653
CASRN: 30568-39-9



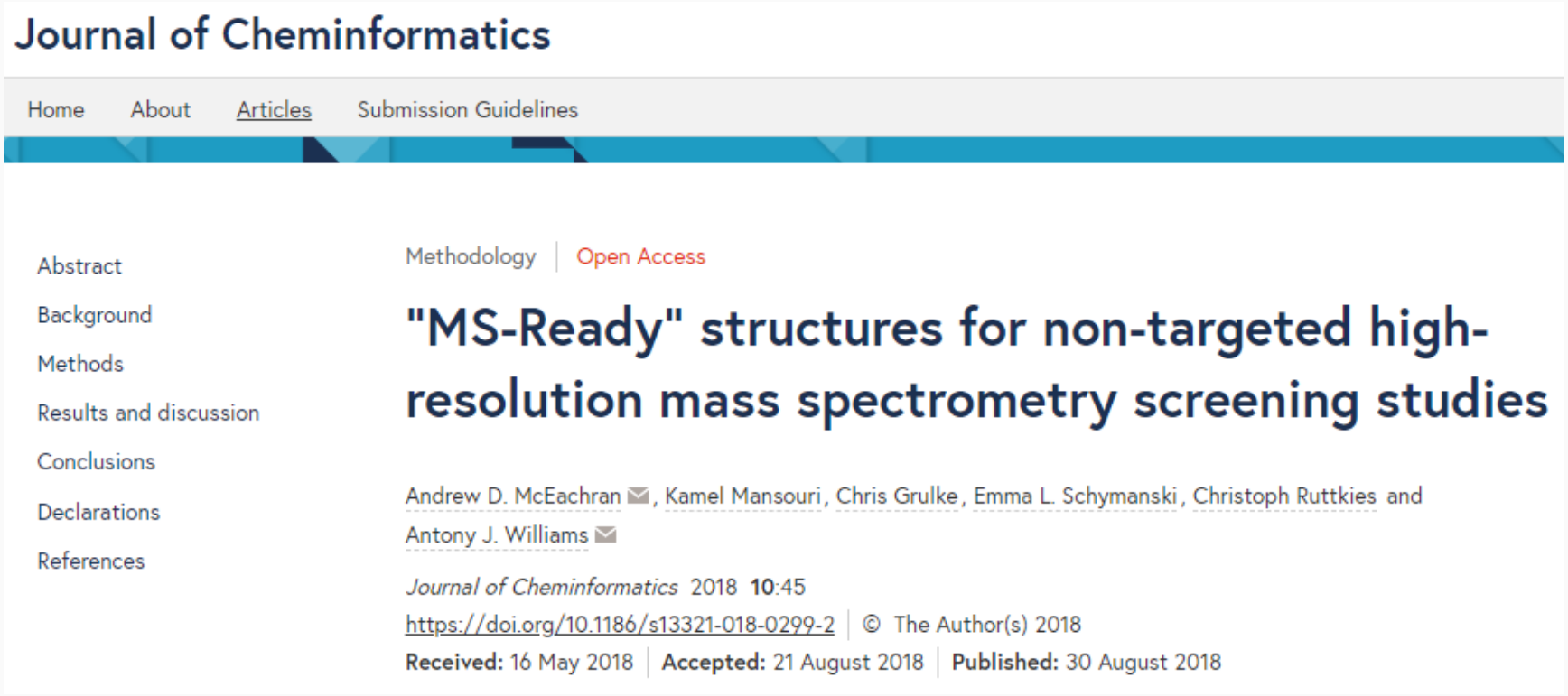
1-Heptanone, 1-(4-pyridyl)-
DTXSID: DTXSID40186594
CASRN: 32941-30-3

Advanced Searches
Mass Based Search



MS-Ready Structures

https://jcheminf.biomedcentral.com/articles/10.1186/s13321-018-0299-2



Journal of Cheminformatics

Home About Articles Submission Guidelines

Abstract
Background
Methods
Results and discussion
Conclusions
Declarations
References

Methodology | Open Access

"MS-Ready" structures for non-targeted high-resolution mass spectrometry screening studies

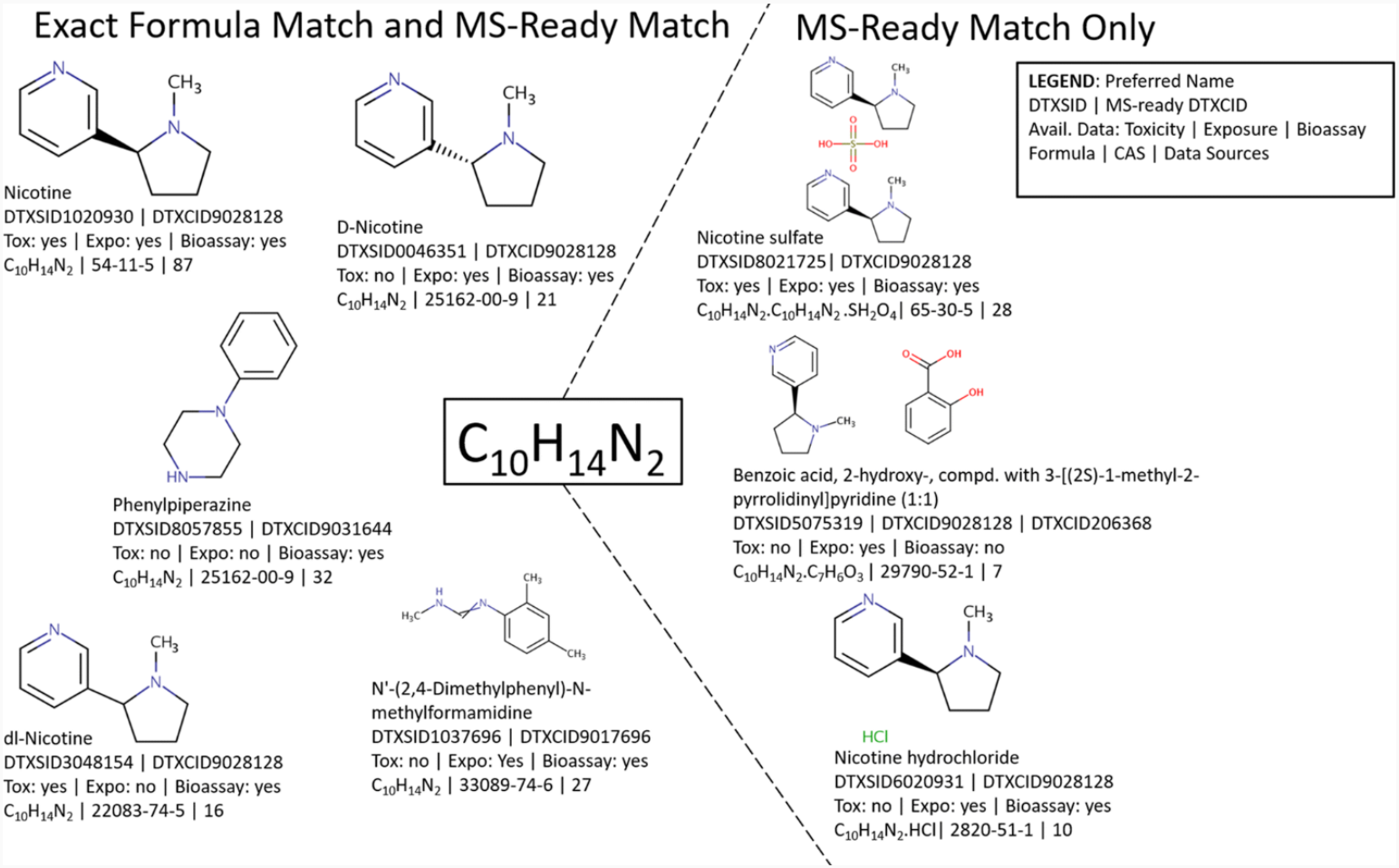
Andrew D. McEachran, Kamel Mansouri, Chris Grulke, Emma L. Schymanski, Christoph Ruttkies and Antony J. Williams

Journal of Cheminformatics 2018 **10**:45

https://doi.org/10.1186/s13321-018-0299-2 | © The Author(s) 2018

Received: 16 May 2018 | **Accepted:** 21 August 2018 | **Published:** 30 August 2018

MS-Ready Structures

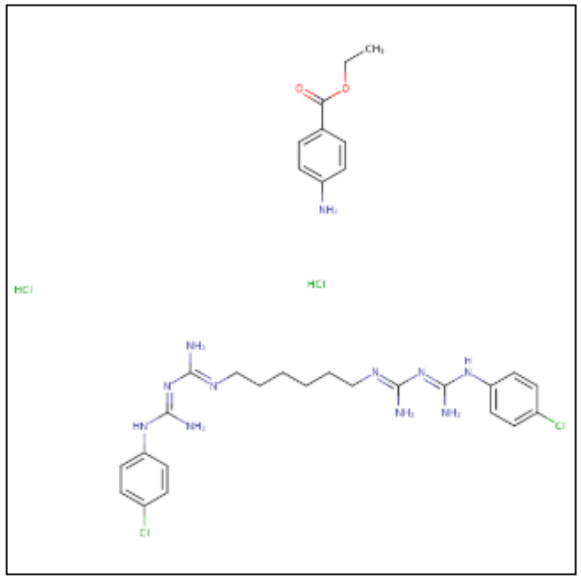


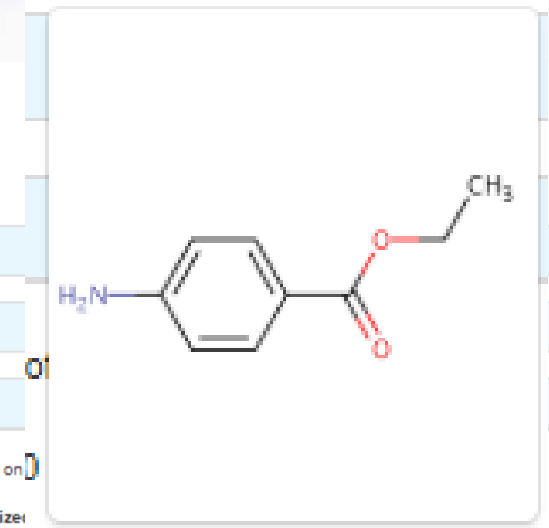
MS-Ready Mappings

Progaron

108532-15-6 | DTXSID20148579

Searched by DSSTox Substance Id.

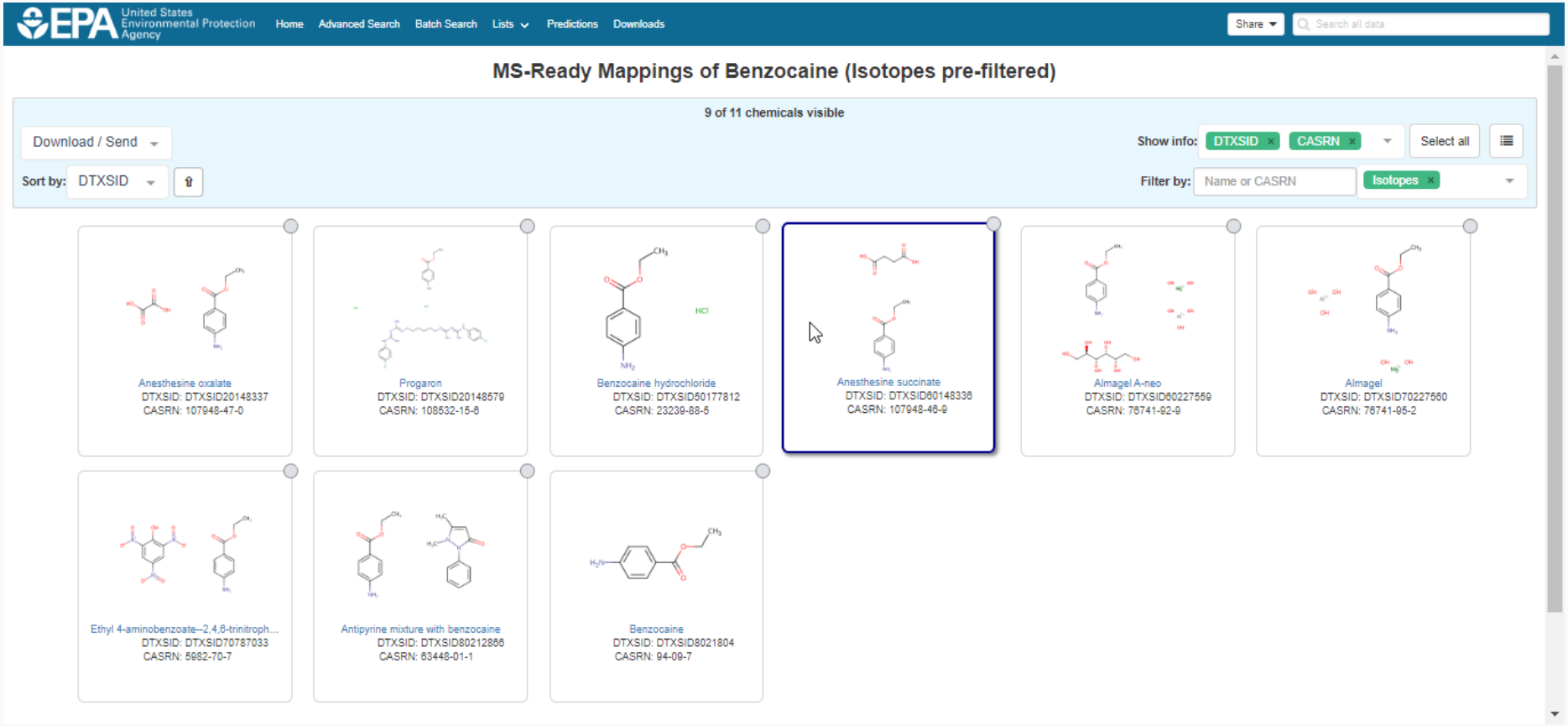




- Intrinsic Properties
- Structural Identifiers
- Linked Substances
  - Same Connectivity: 1 record (based on
  - Mixtures, Components and Neutralize
  - MS-Ready Mappings: DTXCID0013314
  - Similar Compounds: 0 records
- Presence in Lists
- Record Information
- Quality Control Notes

DTXCID301804:11 records; D

MS-Ready Mappings Set

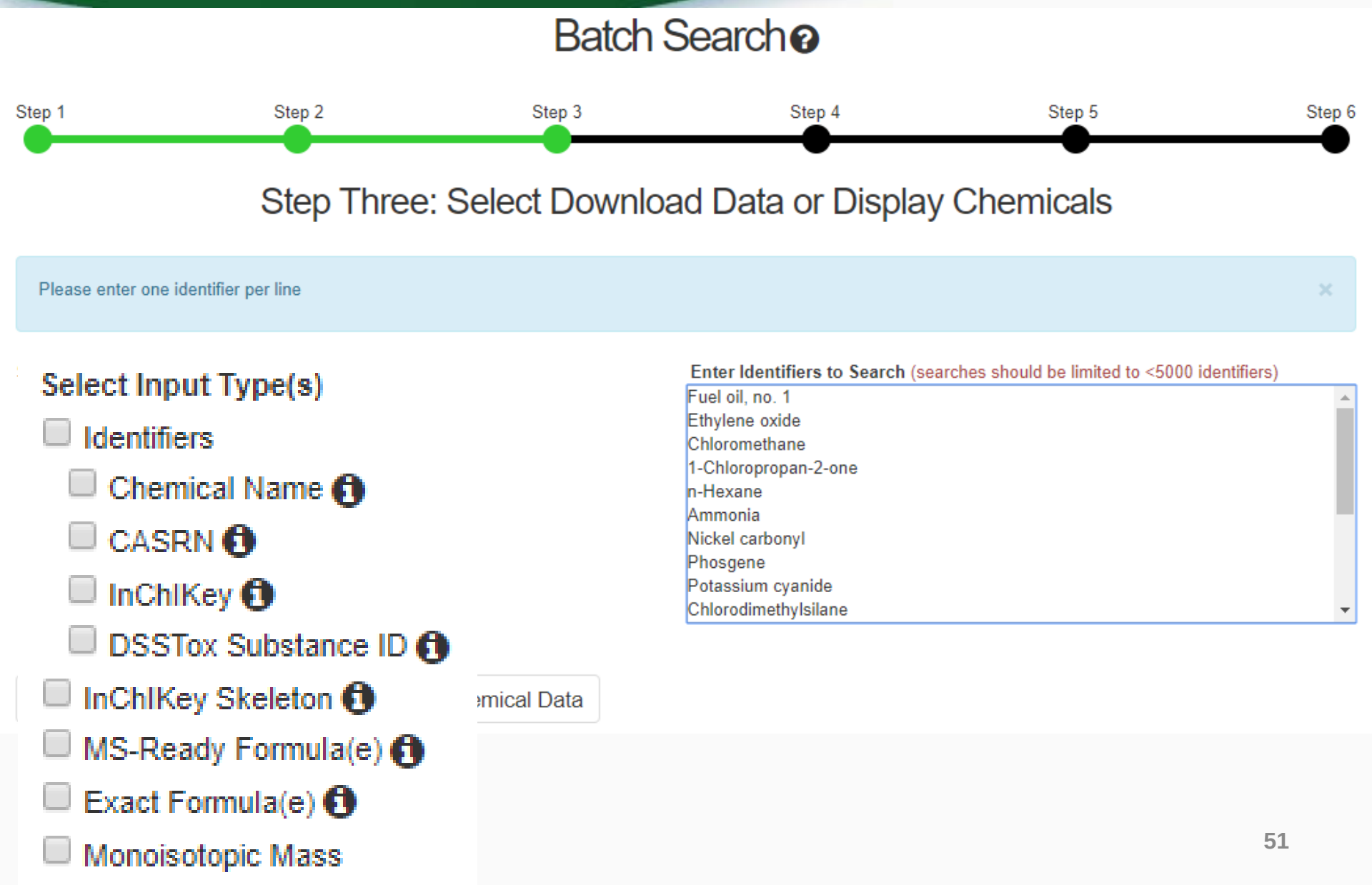


Batch Searching

- Singleton searches are useful but we work with thousands of chemicals!

- Typical questions
  - What chemicals can I get for 5000 CAS Numbers?
  - Can I get predicted properties for 1000 chemicals?
  - What is the list of chemicals for the formula $C_xH_yO_z$?
  - What is the list of chemicals for a mass +/- error ?
  - Can I get chemical lists in Excel files? In SDF files?

Batch Searching



Batch Searching

EPA United States Environmental Protection Agency

Select Output Format:

Excel

Download

Customize Results

- Select All
- Select All in Lists

Chemical Identifiers

- DTXSID
- Chemical Name
- CAS-RN
- InChIKey
- IUPAC Name

Structures

- Mol File
- SMILES
- InChI String
- MS-Ready SMILES
- QSAR-Ready SMILES

Intrinsic And Predicted Properties

- Molecular Formula
- Average Mass
- Monoisotopic Mass
- TEST Model Predictions
- OPERA Model Predictions

Presence in Lists:

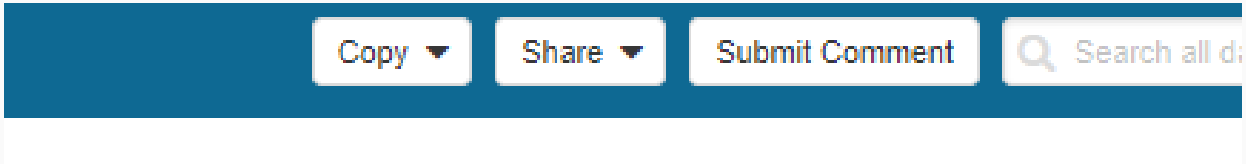
- ICCVAM test method evaluation report: in vitro ocular toxicity test methods
- 40CFR355
- A list of all PBDEs (Polybrominated diphenyl ethers)
- A list of all PCBs (Polychlorinated biphenyls)
- A list of polycyclic aromatic hydrocarbons
- Acute exposure guideline levels
- Algal Toxins
- Androgen Receptor Chemicals
- APCRA Chemicals for Prospective Analysis
- APCRA Chemicals for Retrospective Analysis
- APCRA Chemicals for Retrospective Analysis_App_List_448_Chemicals
- ATSDR Minimal Risk Levels (MRLs) for Hazardous Substances
- ATSDR Toxic Substances Portal Chemical List
- Bisphenol Compounds
- California Office of Environmental Health Hazard Assessment
- Chemicals with interesting names
- CMAP
- DNT Screening Library
- Drinking Water Suspects, KWR Water, Netherlands
- EDSP Universe
- EPA Chemicals associated with hydraulic fracturing

Excel Output

	A	B	C	D	E	F	G	H
1	INPUT	FOUND_BY	DTXSID	PREFERRED_NAME	EXPOCAST	EXPOCAST	NHANES	TOXVAL_D
2	1445-75-6	CAS-RN	DTXSID5024051	Diisopropyl methylpho	2.09e-08	Y	-	Y
3	50-00-0	CAS-RN	DTXSID7020637	Formaldehyde	1.32e-06	Y	-	Y
4	107-06-2	CAS-RN	DTXSID6020438	1,2-Dichloroethane	4.9e-06	Y	-	Y
5	57-12-5	CAS-RN	DTXSID6023991	Cyanide	-	-	-	Y
6	7550-45-0	CAS-RN	DTXSID8042476	Titanium tetrachloride	-	-	-	Y
7	79-01-6	CAS-RN	DTXSID0021383	Trichloroethylene	7.27e-06	Y	-	Y
8	121-82-4	CAS-RN	DTXSID9024142	Cyclonite	6.72e-08	Y	-	Y
9	108-05-4	CAS-RN	DTXSID3021431	Vinyl acetate	8.3e-05	Y	-	Y
10	7803-51-2	CAS-RN	DTXSID2021157	Phosphine	-	-	-	Y
11	122-66-7	CAS-RN	DTXSID7020710	1,2-Diphenylhydrazine	1.49e-07	Y	-	Y
12	101-77-9	CAS-RN	DTXSID6022422	4,4'-Methylenedianiline	6.08e-06	Y	-	Y
13	14017-34-6	CAS-RN	DTXSID90161250	Selenium difluoride	-	-	-	-
14	75-44-5	CAS-RN	DTXSID0024260	Phosgene	-	-	-	Y
15	621-64-7	CAS-RN	DTXSID6021032	N-Nitrosodipropylamine	4.55e-07	Y	-	Y
16	75-09-2	CAS-RN	DTXSID0020868	Dichloromethane	2.02e-06	Y	-	Y
17	100-41-4	CAS-RN	DTXSID3020596	Ethylbenzene	8.32e-05	Y	-	Y
18	7440-28-0	CAS-RN	DTXSID2036035	Thallium	-	-	-	Y
19	108-88-3	CAS-RN	DTXSID7021360	Toluene	8.61e-05	Y	-	Y
20	111-44-4	CAS-RN	DTXSID9020168	Bis(2-chloroethyl) ethe	2.82e-07	Y	-	Y
21	7440-42-8	CAS-RN	DTXSID3023922	Boron	-	-	-	Y
22	7440-29-1	CAS-RN	DTXSID6049800	Thorium	-	-	-	Y

How can be curate our data?

- Crowdsourcing is well proven nowadays
- Comments can be added at a record level



- Submitted comments are reviewed by administrators and responded to

Public Crowdsourced Comments

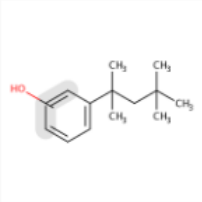
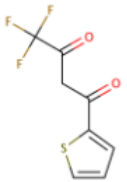
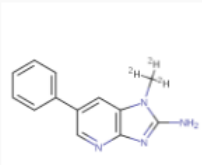
https://comptox.epa.gov/dashboard/comments/public_index

EPA United States Environmental Protection Agency

EPA United States Environmental Protection Agency | Home | Advanced Search | Batch Search | Lists | Predictions | Downloads | Share | Search all data

Crowdsourced Comments

Show 10 entries | Search:

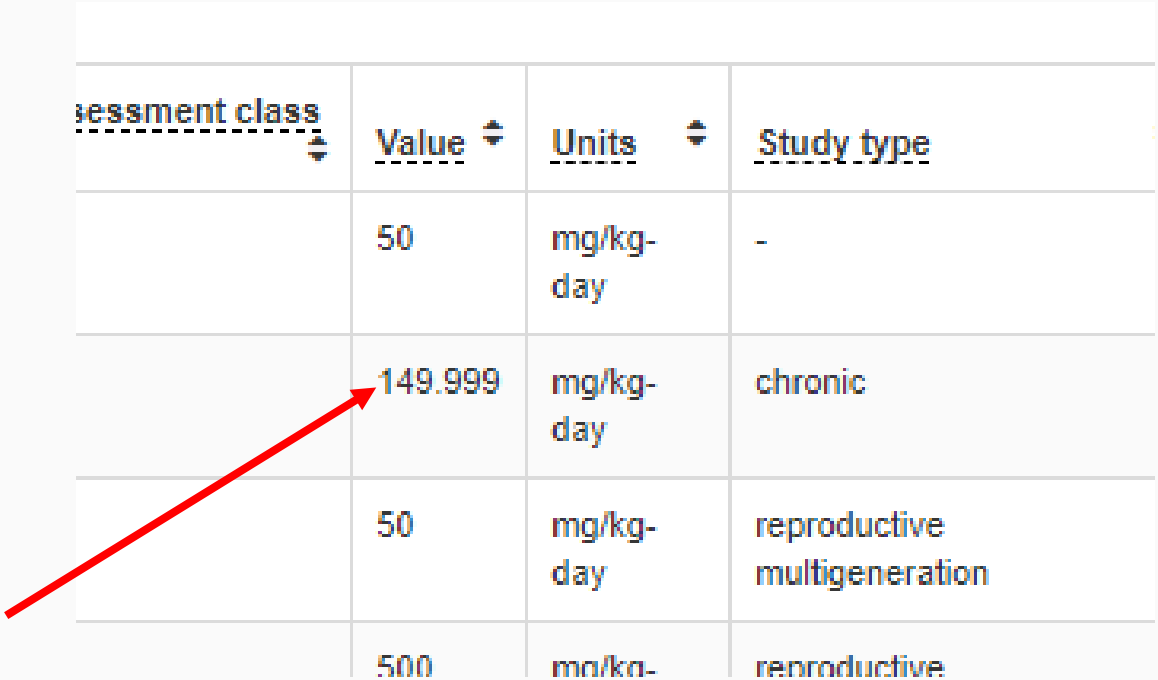
Chemical	Structure	Date	Comment	Status
(1,1,3,3-Tetramethylbutyl)phenol		2017-07-15	Octylphenol redirects here, yet the name and related chemicals are 1,1,3,3-tetramethylbutylphenol - which is only a subset of all octylphenol isomers? Is this CAS only for these alkyl isomers?	★
1,3-Butanedione, 4,4,4-trifluoro-1-(2-thienyl)-		2017-03-30	Synonym: TTFA (Any way to bank these reCAPTCHAs so I don't have to do it everytime?)	★
1-(²H₃)Methyl-6-phenyl-1H-imidazo[4,5-b]pyridin-2-amine		2017-05-06	1-(2H3)Methyl-6-phenyl-1H-imidazo[4,5-b]pyridin-2-amine 210049-13-1 | DTXSID70670097 contains an error in the empirical formula due to an error in the deuterium representation and subsequent counting	★

Comments to date

- The majority of comments to date:
  - Structure and names/CASRN do not match
  - Add additional synonyms
  - Request to add specific property data
  - Structure layout/depiction needs improving

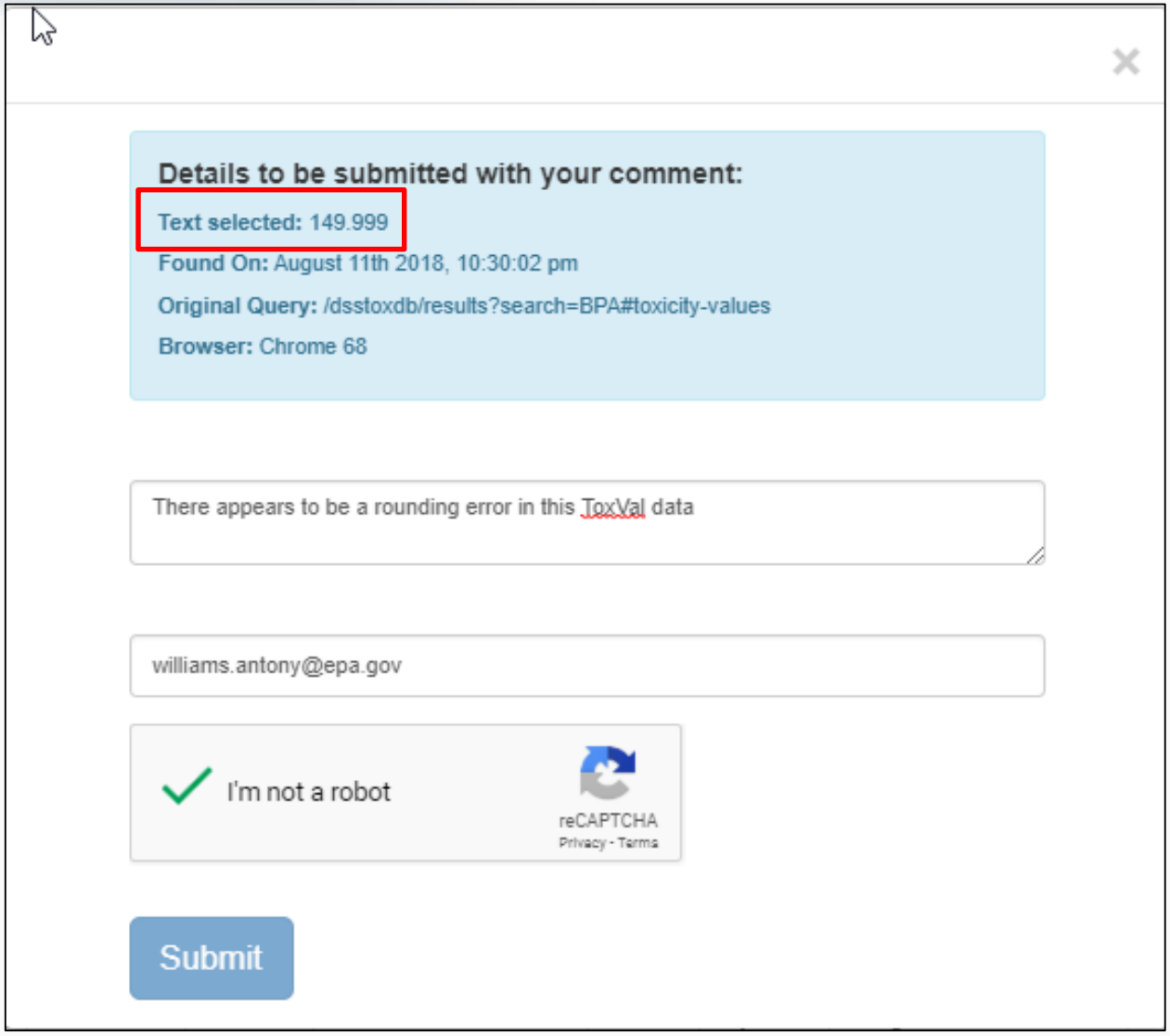
Crowdsourcing Comments
Single Cell Commenting added

- Highlight an alphanumeric text string



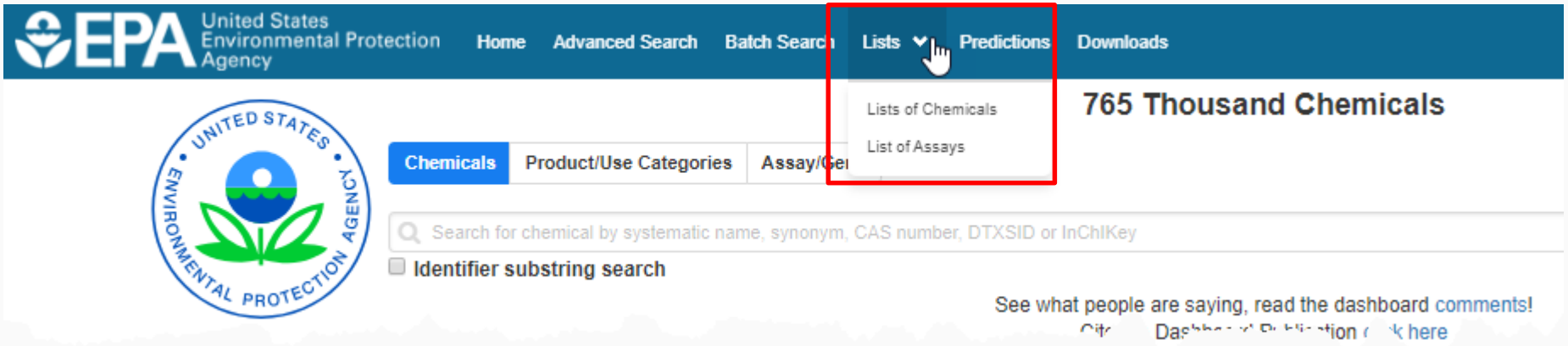
ssessment class	Value	Units	Study type
	50	mg/kg-day	-
	149.999	mg/kg-day	chronic
	50	mg/kg-day	reproductive multigeneration
	500	mg/kg-	reproductive

Crowdsourcing Comments



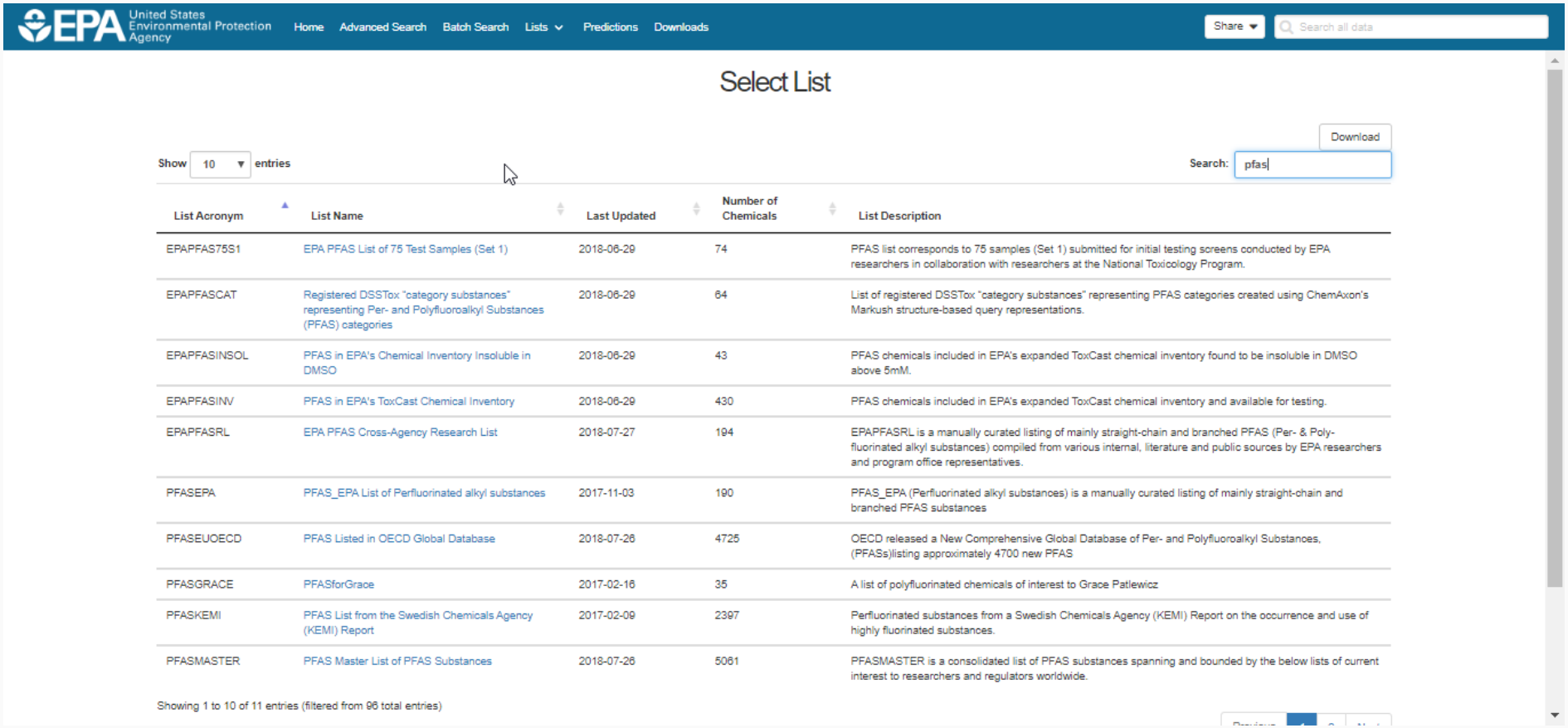
Lists of Lists

- Lists of chemicals – ca. 100 lists
- List of ToxCast/Tox21 assays



11 PFAS Lists

http://comptox-prod.epa.gov/dashboard/chemical_lists



List Acronym	List Name	Last Updated	Number of Chemicals	List Description
EPAPFAS75S1	EPA PFAS List of 75 Test Samples (Set 1)	2018-06-29	74	PFAS list corresponds to 75 samples (Set 1) submitted for initial testing screens conducted by EPA researchers in collaboration with researchers at the National Toxicology Program.
EPAPFASCAT	Registered DSSTox "category substances" representing Per- and Polyfluoroalkyl Substances (PFAS) categories	2018-06-29	64	List of registered DSSTox "category substances" representing PFAS categories created using ChemAxon's Markush structure-based query representations.
EPAPFASINSOL	PFAS in EPA's Chemical Inventory Insoluble in DMSO	2018-06-29	43	PFAS chemicals included in EPA's expanded ToxCast chemical inventory found to be insoluble in DMSO above 5mM.
EPAPFASINV	PFAS in EPA's ToxCast Chemical Inventory	2018-06-29	430	PFAS chemicals included in EPA's expanded ToxCast chemical inventory and available for testing.
EPAPFASRL	EPA PFAS Cross-Agency Research List	2018-07-27	194	EPAPFASRL is a manually curated listing of mainly straight-chain and branched PFAS (Per- & Poly-fluorinated alkyl substances) compiled from various internal, literature and public sources by EPA researchers and program office representatives.
PFASEPA	PFAS_EPA List of Perfluorinated alkyl substances	2017-11-03	190	PFAS_EPA (Perfluorinated alkyl substances) is a manually curated listing of mainly straight-chain and branched PFAS substances
PFASEUOECD	PFAS Listed in OECD Global Database	2018-07-26	4725	OECD released a New Comprehensive Global Database of Per- and Polyfluoroalkyl Substances, (PFASs)listing approximately 4700 new PFAS
PFASGRACE	PFASforGrace	2017-02-16	35	A list of polyfluorinated chemicals of interest to Grace Patlewicz
PFASKEMI	PFAS List from the Swedish Chemicals Agency (KEMI) Report	2017-02-09	2397	Perfluorinated substances from a Swedish Chemicals Agency (KEMI) Report on the occurrence and use of highly fluorinated substances.
PFASMASTER	PFAS Master List of PFAS Substances	2018-07-26	5061	PFASMASTER is a consolidated list of PFAS substances spanning and bounded by the below lists of current interest to researchers and regulators worldwide.

Showing 1 to 10 of 11 entries (filtered from 96 total entries)

The OECD List of PFAS

http://www.oecd.org/chemicalsafety/portal-perfluorinated-chemicals/

EPA United States Environmental Protection Agency

The OECD releases a new list of PFASs

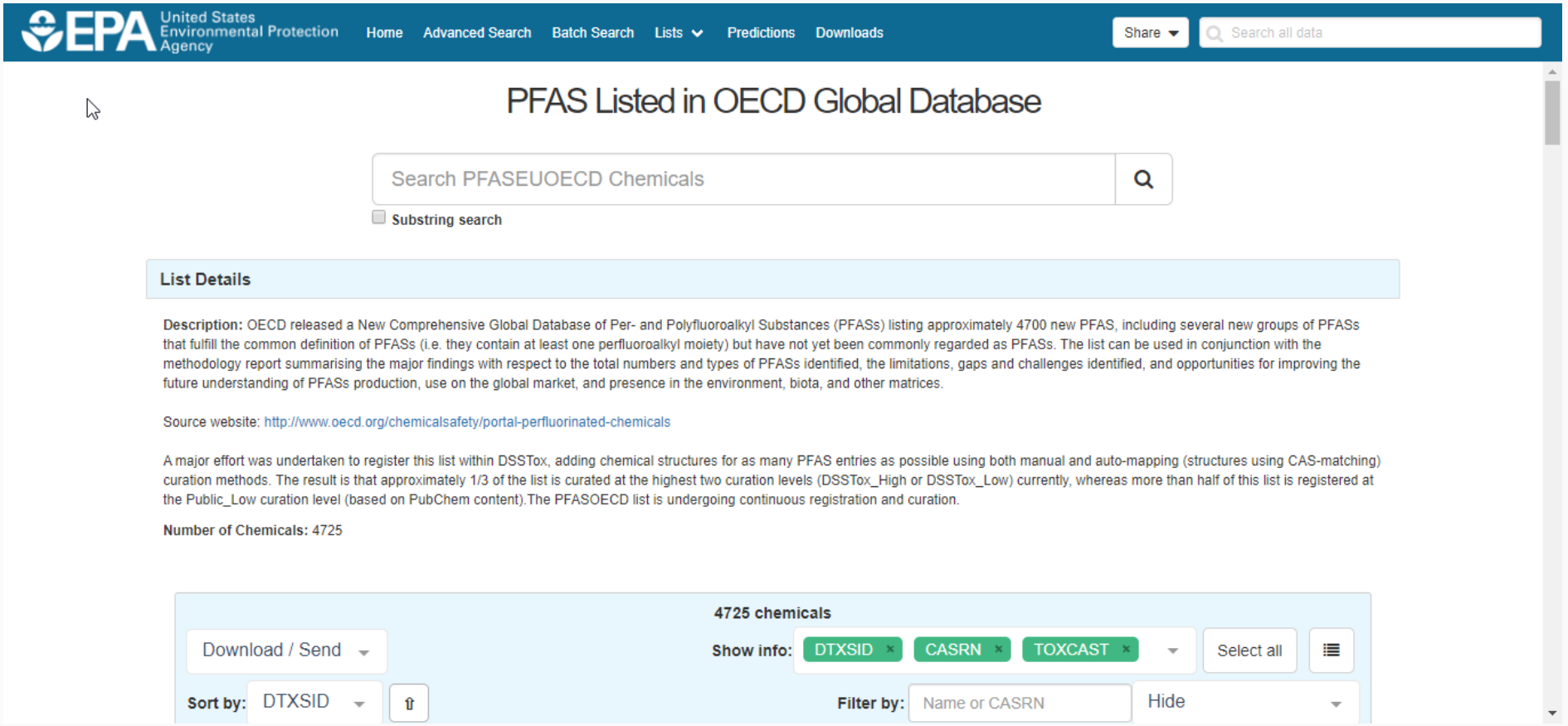
The OECD releases a new list of Per- and Polyfluoroalkyl Substances (PFASs) based on a comprehensive analysis of information available in the public domain. In total, 4730 PFAS-related CAS numbers have been identified and categorised in this study, including several new groups of PFASs that fulfil the common definition of PFASs (i.e. they contain at least one perfluoroalkyl moiety) but have not yet been commonly regarded as PFASs.

This work has been conducted under the OECD/UN Environment Global PFC Group in support of the Strategic Approach to International Chemicals Management (SAICM) and shifting to safer alternatives for PFASs.

The New Comprehensive Global Database of Per- and Polyfluoroalkyl Substances (PFASs) comes with a methodology report also detailing the major findings with respect to the total numbers and types of PFASs identified, the limitations, gaps and challenges identified in the development of the new list, and opportunities for improving the future understanding of PFASs production, use on the global market, and presence in the environment, biota, and other matrices.

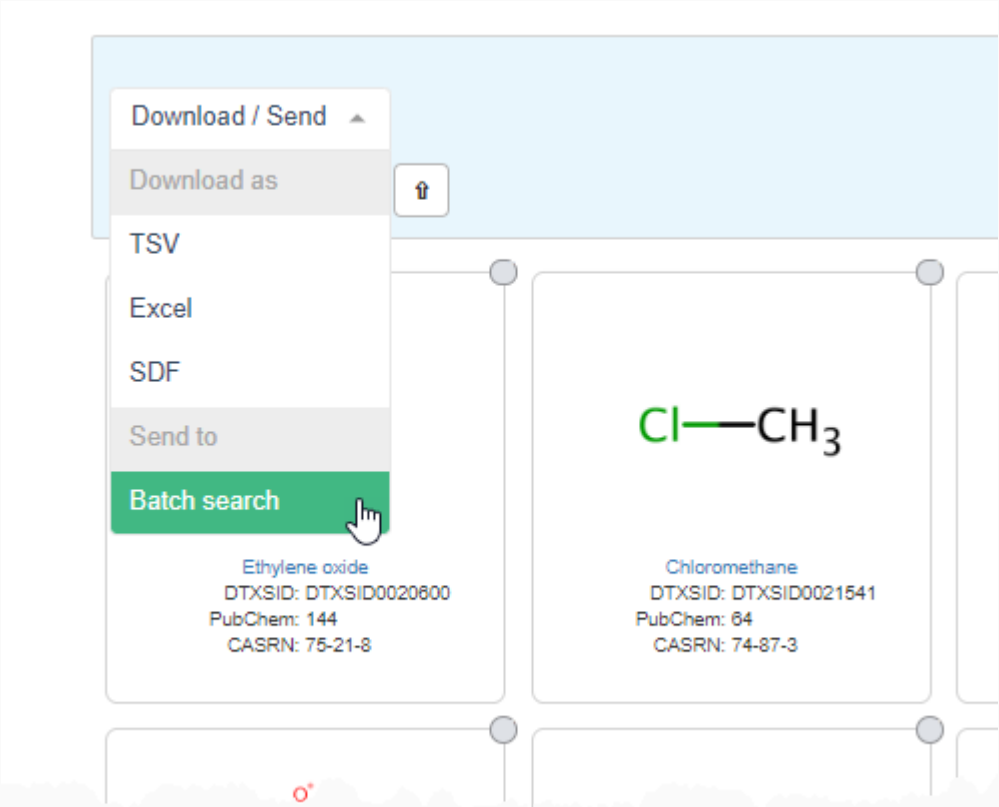
The OECD List of PFAS

http://www.oecd.org/chemicalsafety/portal-perfluorinated-chemicals/

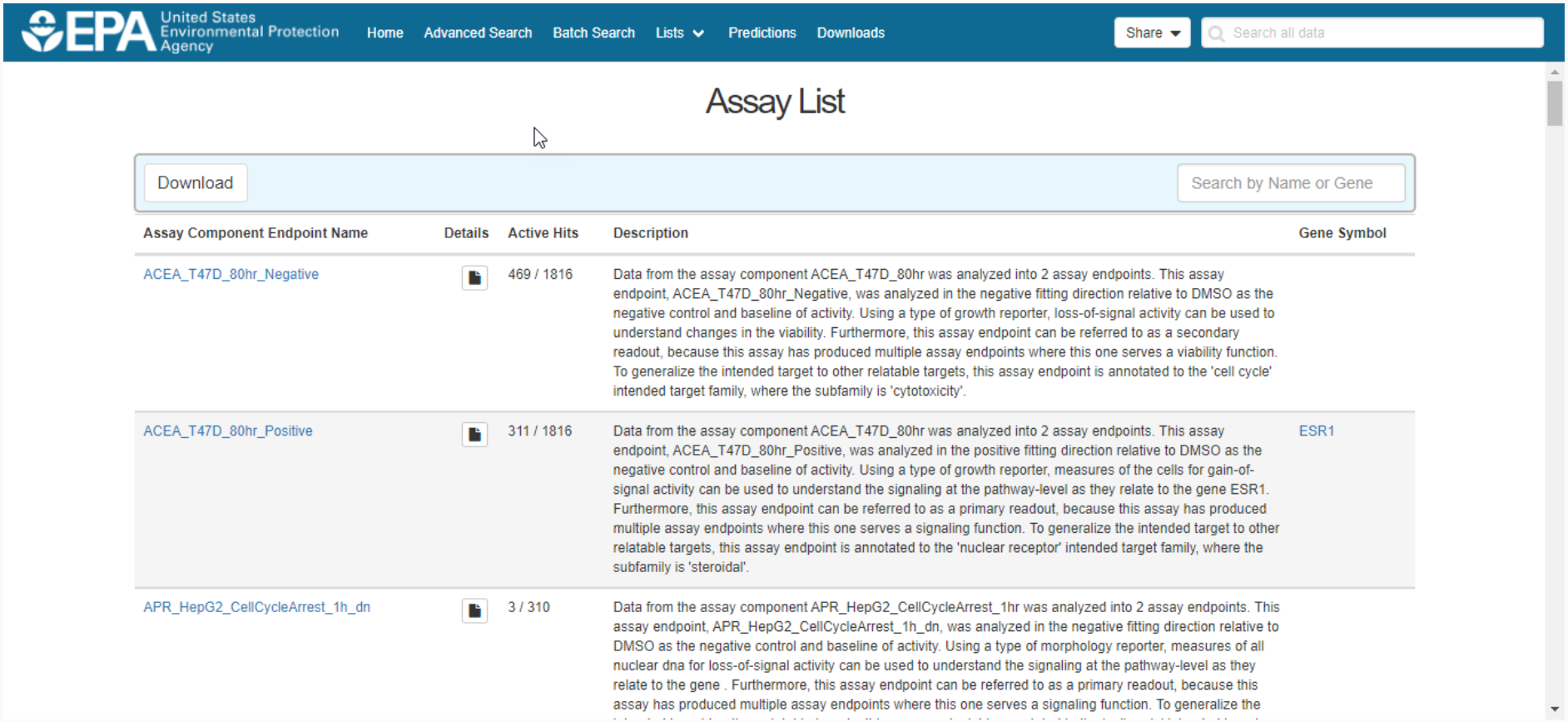


Want data for a list???

- Simply send to Batch and choose data…



List of Assays



EPA United States Environmental Protection Agency

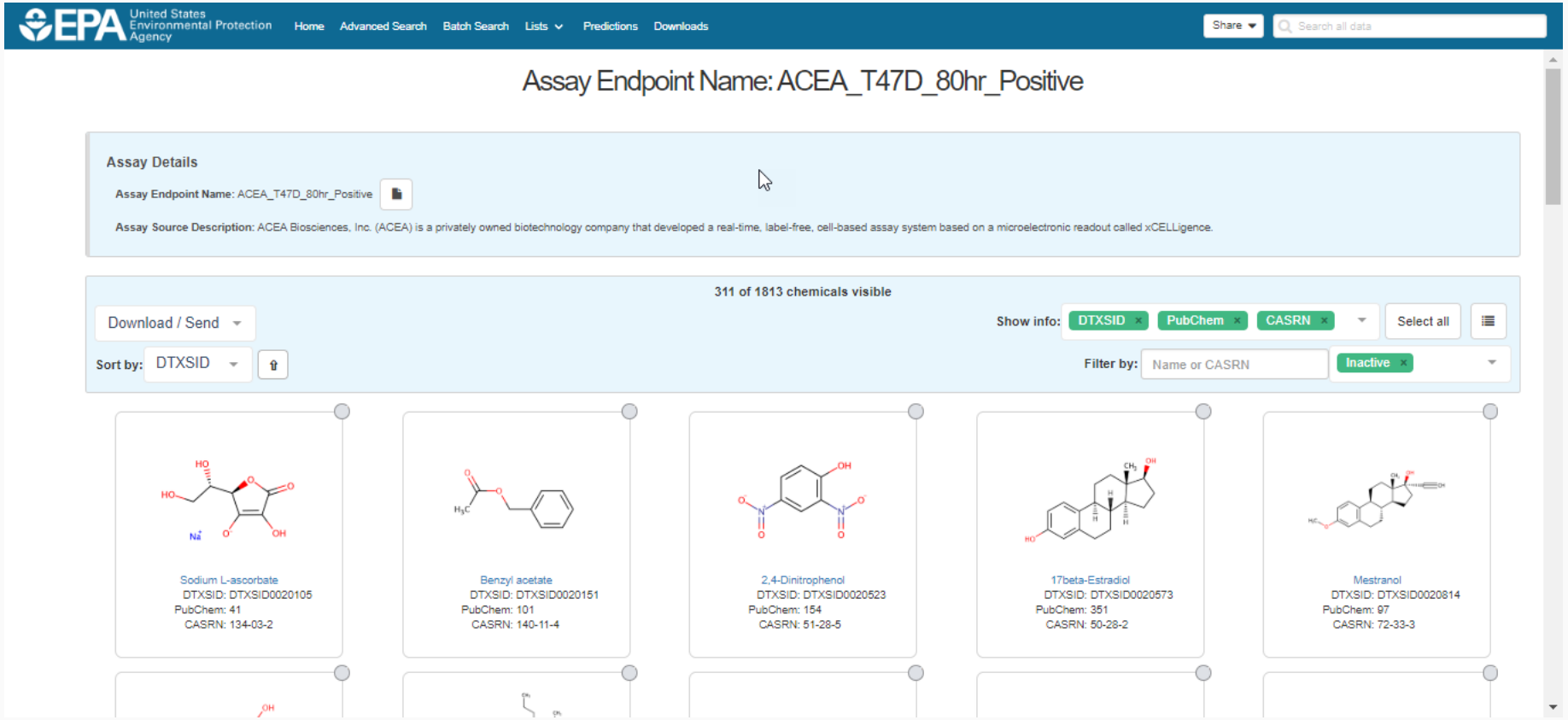
Home | Advanced Search | Batch Search | Lists | Predictions | Downloads | Share | Search all data

Assay List

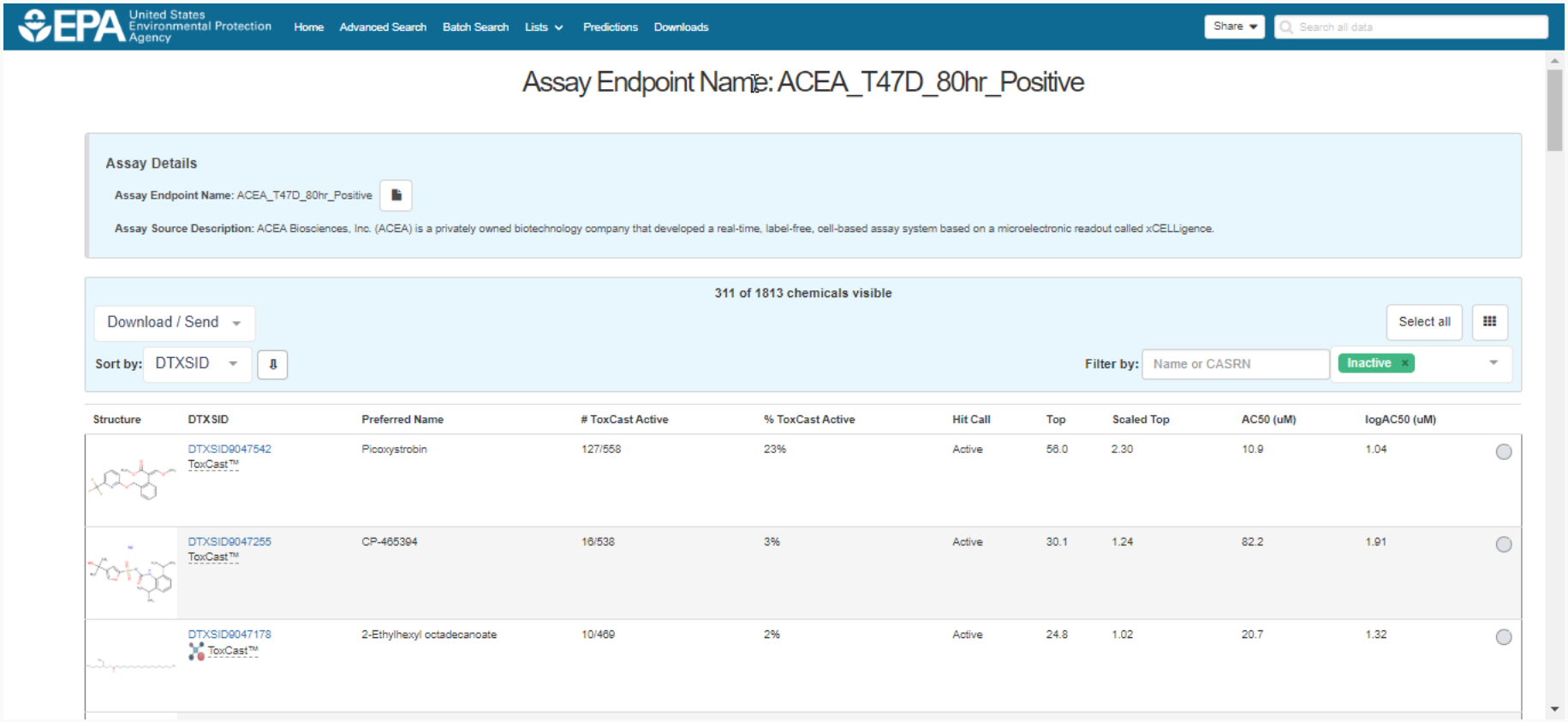
Download | Search by Name or Gene

Assay Component Endpoint Name	Details	Active Hits	Description	Gene Symbol
ACEA_T47D_80hr_Negative		469 / 1816	Data from the assay component ACEA_T47D_80hr was analyzed into 2 assay endpoints. This assay endpoint, ACEA_T47D_80hr_Negative, was analyzed in the negative fitting direction relative to DMSO as the negative control and baseline of activity. Using a type of growth reporter, loss-of-signal activity can be used to understand changes in the viability. Furthermore, this assay endpoint can be referred to as a secondary readout, because this assay has produced multiple assay endpoints where this one serves a viability function. To generalize the intended target to other relatable targets, this assay endpoint is annotated to the 'cell cycle' intended target family, where the subfamily is 'cytotoxicity'.	
ACEA_T47D_80hr_Positive		311 / 1816	Data from the assay component ACEA_T47D_80hr was analyzed into 2 assay endpoints. This assay endpoint, ACEA_T47D_80hr_Positive, was analyzed in the positive fitting direction relative to DMSO as the negative control and baseline of activity. Using a type of growth reporter, measures of the cells for gain-of-signal activity can be used to understand the signaling at the pathway-level as they relate to the gene ESR1. Furthermore, this assay endpoint can be referred to as a primary readout, because this assay has produced multiple assay endpoints where this one serves a signaling function. To generalize the intended target to other relatable targets, this assay endpoint is annotated to the 'nuclear receptor' intended target family, where the subfamily is 'steroidal'.	ESR1
APR_HepG2_CellCycleArrest_1h_dn		3 / 310	Data from the assay component APR_HepG2_CellCycleArrest_1hr was analyzed into 2 assay endpoints. This assay endpoint, APR_HepG2_CellCycleArrest_1h_dn, was analyzed in the negative fitting direction relative to DMSO as the negative control and baseline of activity. Using a type of morphology reporter, measures of all nuclear dna for loss-of-signal activity can be used to understand the signaling at the pathway-level as they relate to the gene . Furthermore, this assay endpoint can be referred to as a primary readout, because this assay has produced multiple assay endpoints where this one serves a signaling function. To generalize the	

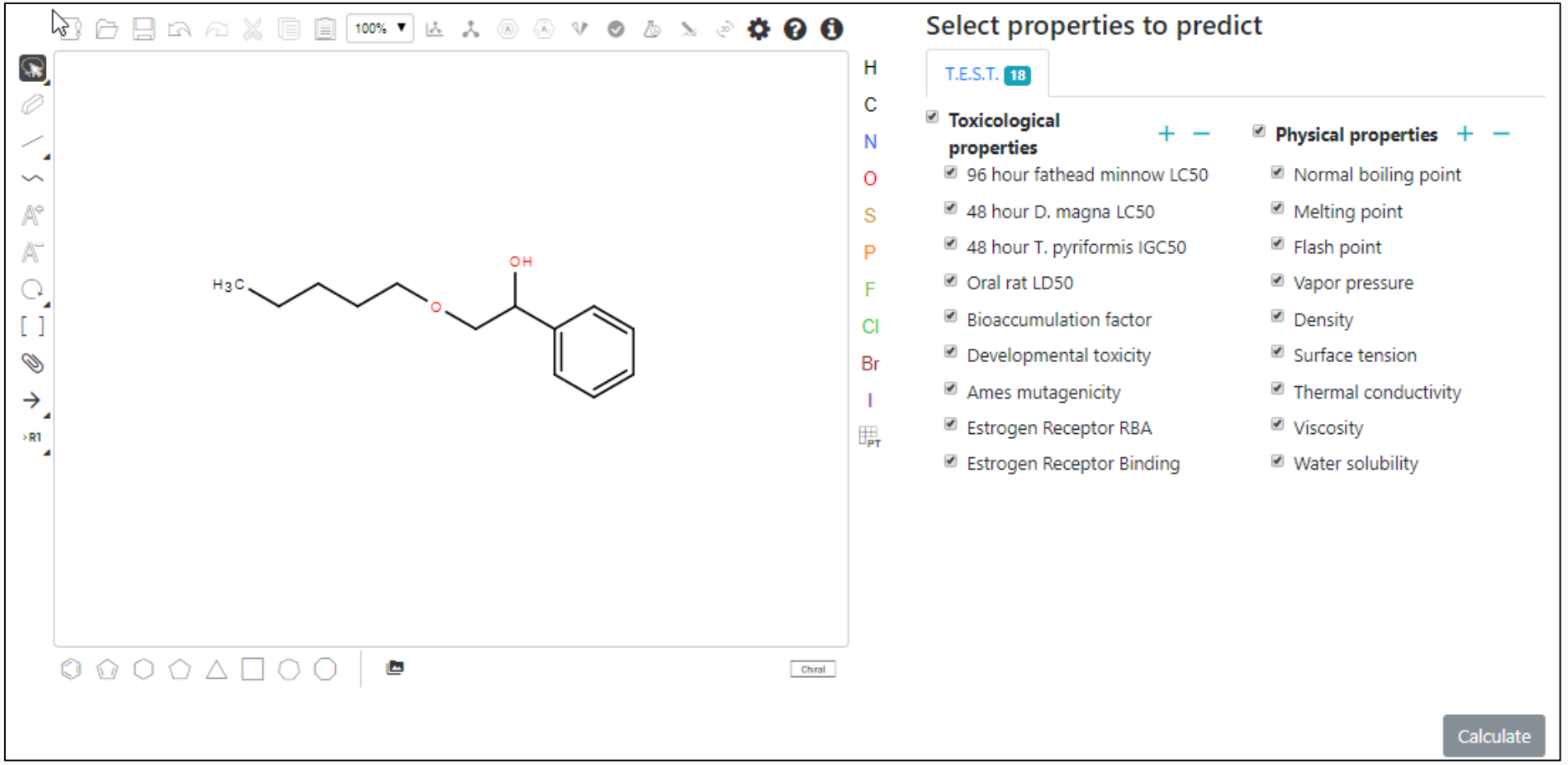
Select an Assay to Navigate Tile View



Select an Assay to Navigate Table View



Real-Time Predictions



Real-Time Predictions

Property	Experimental Value	Prediction				
		Consensus	Hierarchical clustering	Single model	Group contribution	Nearest neighbor
96 hour fathead minnow LC50		4.477 -Log10(mol/L) 6.954 mg/L	4.195 -Log10(mol/L) 13.288 mg/L	3.994 -Log10(mol/L) 21.110 mg/L	3.478 -Log10(mol/L) 69.224 mg/L	6.238 -Log10(mol/L) 0.120 mg/L
48 hour D. magna LC50		4.398 -Log10(mol/L) 8.328 mg/L	3.877 -Log10(mol/L) 27.677 mg/L	4.039 -Log10(mol/L) 19.026 mg/L	4.084 -Log10(mol/L) 17.173 mg/L	5.593 -Log10(mol/L) 0.532 mg/L
48 hour T. pyriformis IGC50		4.063 -Log10(mol/L) 18.039 mg/L	3.731 -Log10(mol/L) 38.668 mg/L		3.386 -Log10(mol/L) 85.610 mg/L	5.070 -Log10(mol/L) 1.773 mg/L
Oral rat LD50		1.758 -Log10(mol/kg) 3640.950 mg/kg	1.982 -Log10(mol/kg) 2172.756 mg/kg			1.533 -Log10(mol/kg) 6101.245 mg/kg
Bioaccumulation factor		1.797 Log10 62.700	2.202 Log10 159.310	1.287 Log10 19.346	1.181 Log10 15.157	2.520 Log10 330.834
Developmental toxicity		false	false	false		true
Ames mutagenicity		false	false			false
Estrogen Receptor RBA		-3.075 Log10 $8.418*10^{-4}$	-3.078 Log10 $8.356*10^{-4}$	-3.720 Log10 $1.907*10^{-4}$		-2.427 Log10 0.004
Estrogen Receptor Binding		true	true	true	false	true

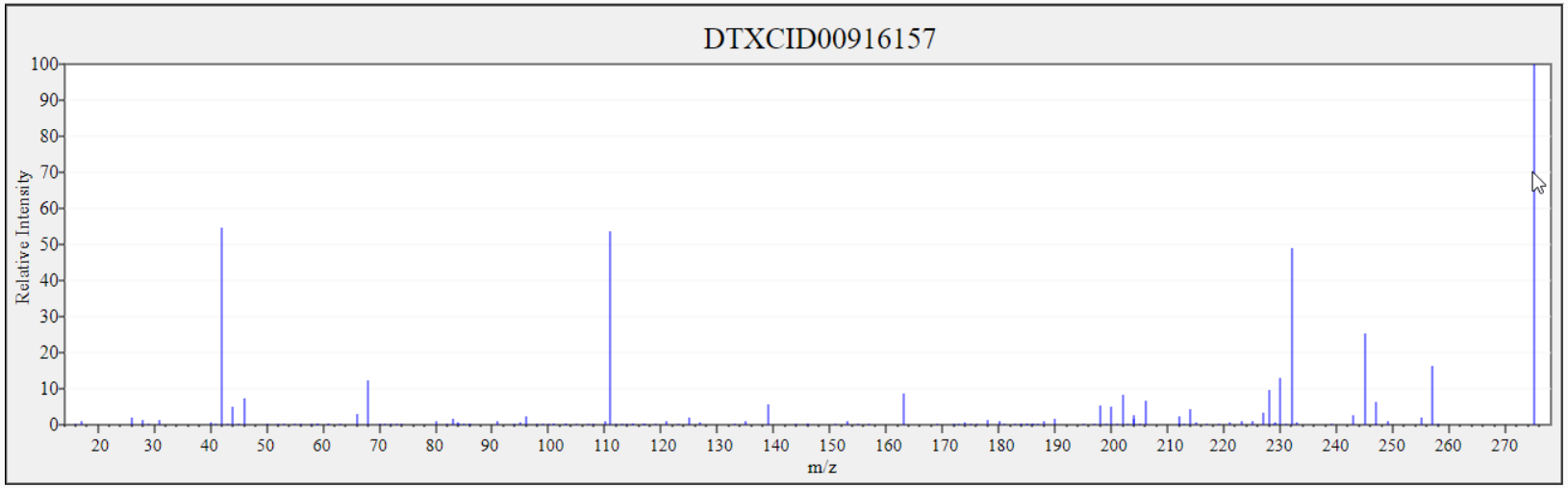
Work in Progress

- CFM-ID
  - Viewing and Downloading pre-predicted spectra
  - Search spectra against the database
- Structure/substructure/similarity search
- pKa prediction

Predicted Mass Spectra

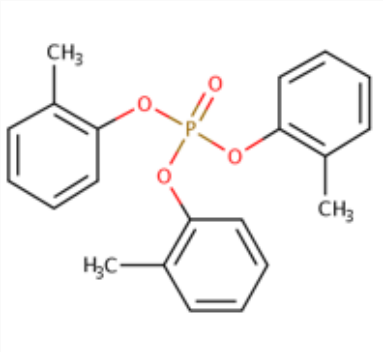
http://cfmid.wishartlab.com/

- MS/MS spectra prediction for ESI+, ESI-, and EI
- Predictions generated and stored for >700,000 structures, to be accessible via Dashboard

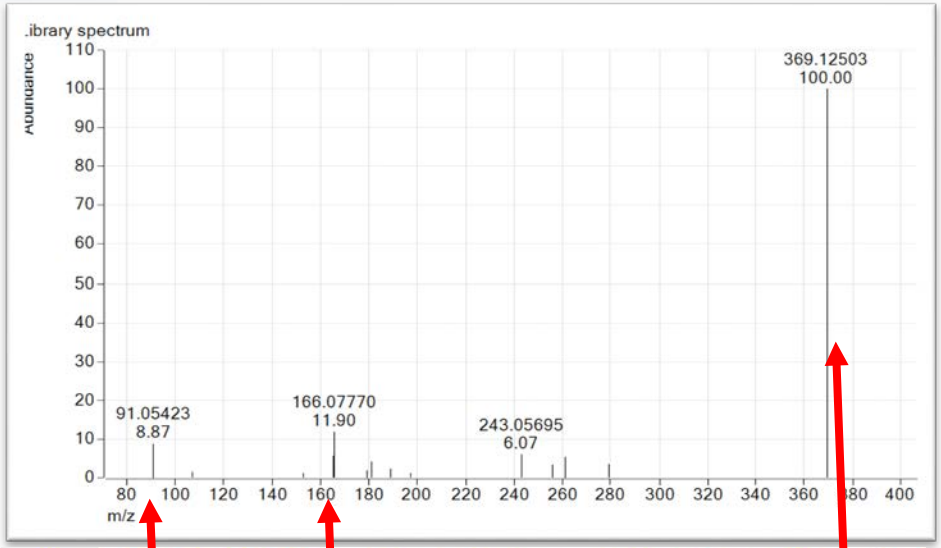


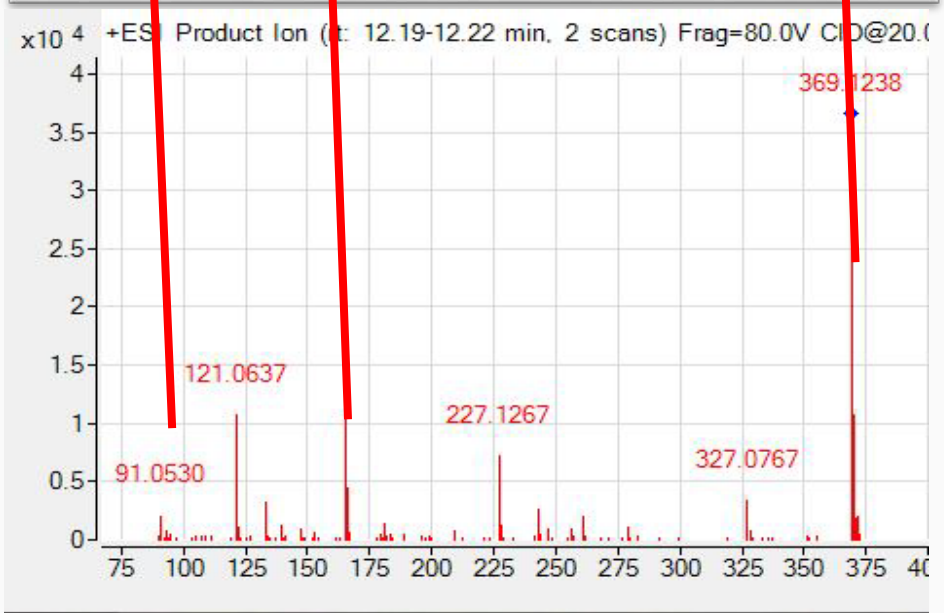
Predicted Mass Spectra

Library Fragmentation Spectra (20eV)



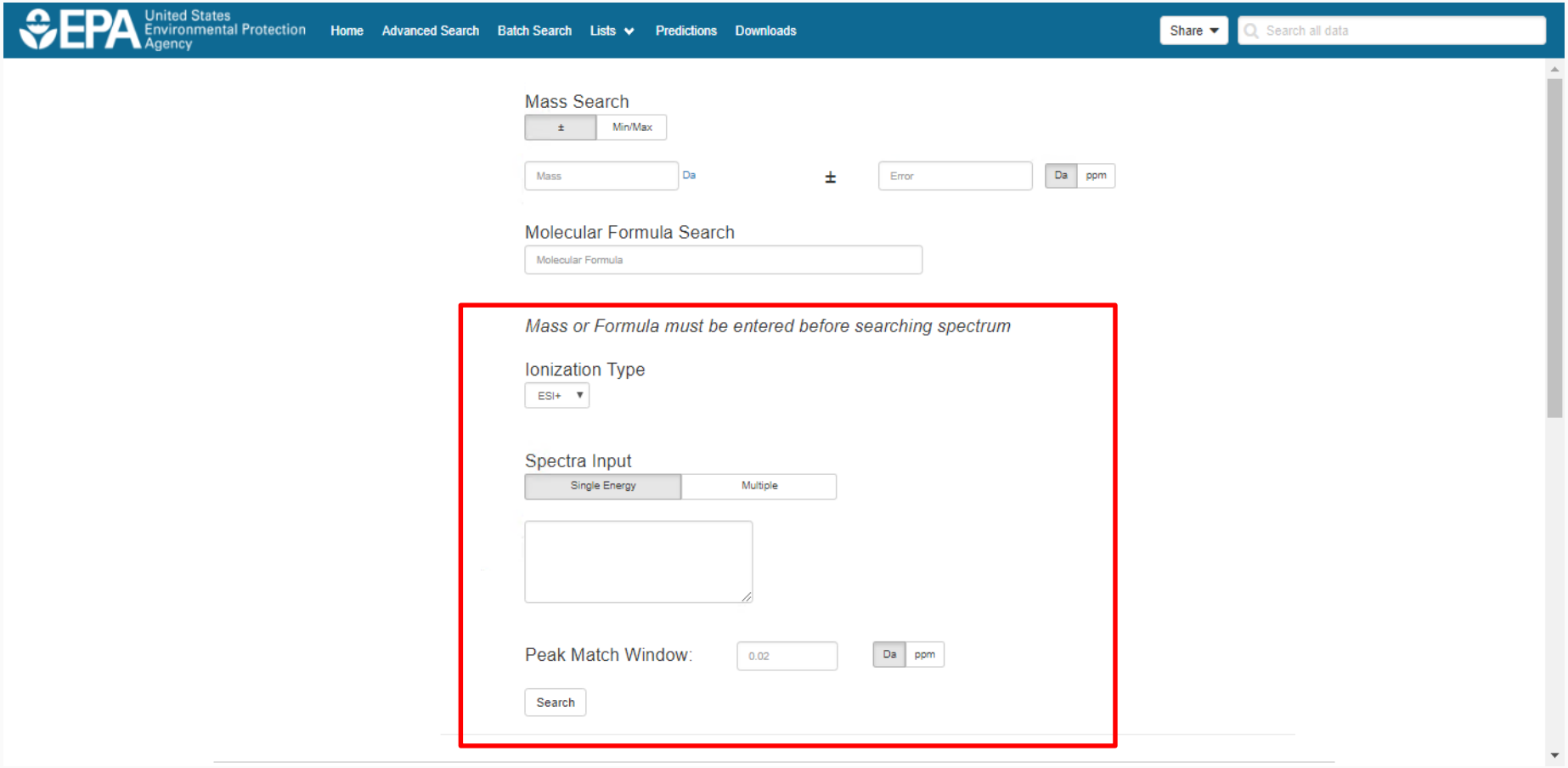
Observed Fragmentation Spectra (20eV)



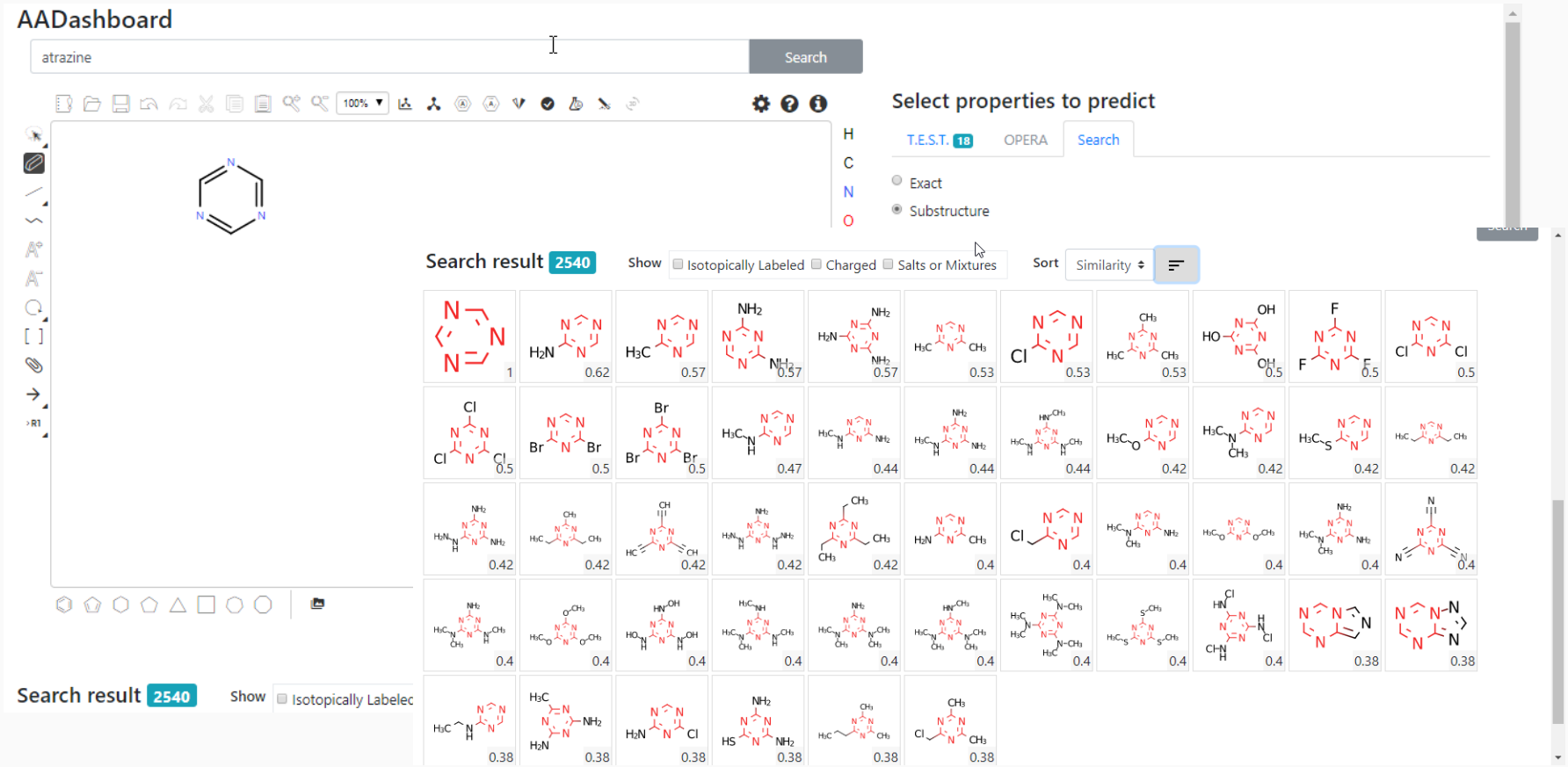


Match Score

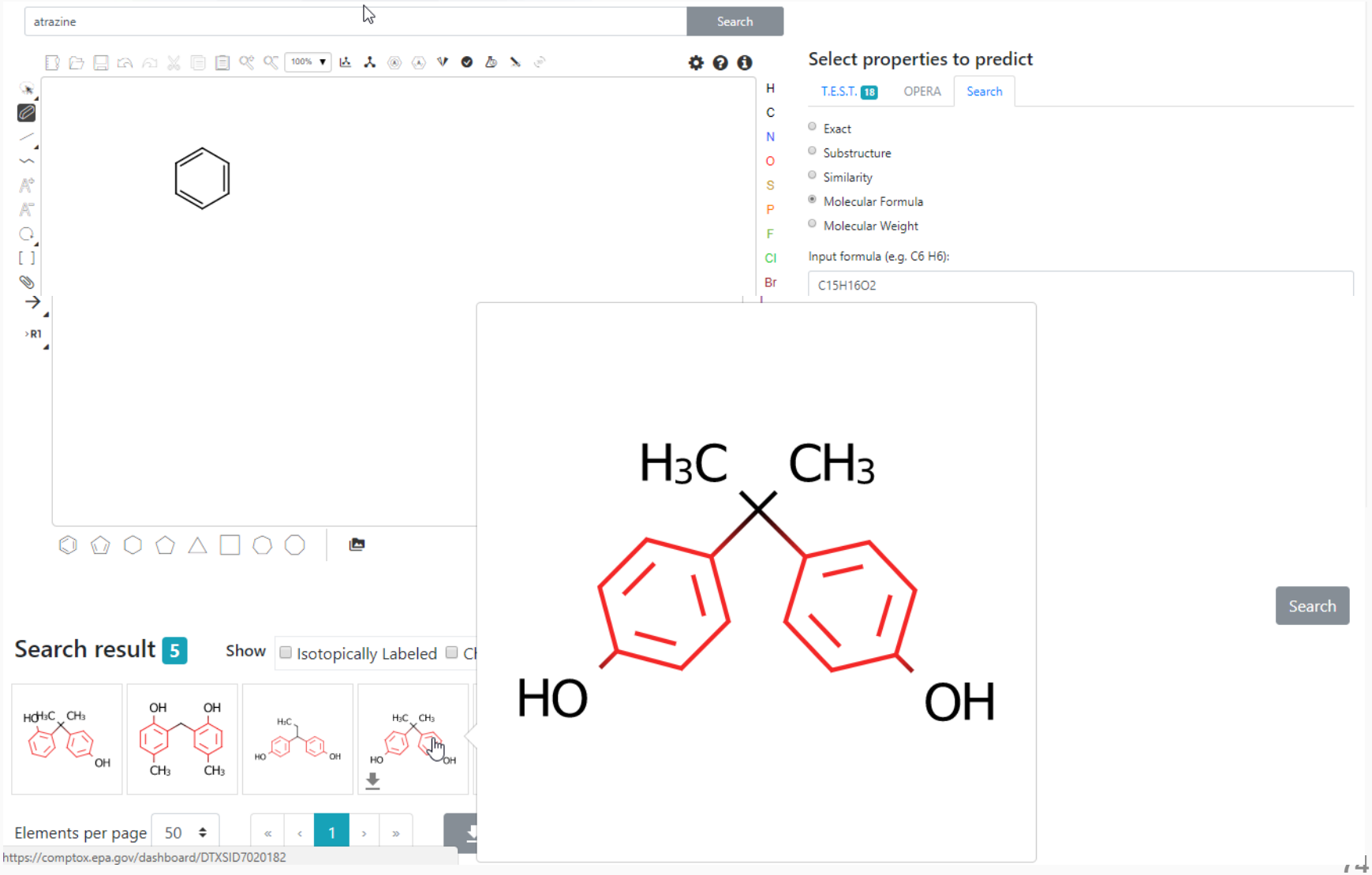
Search Expt. vs. Predicted Spectra



Prototype Development

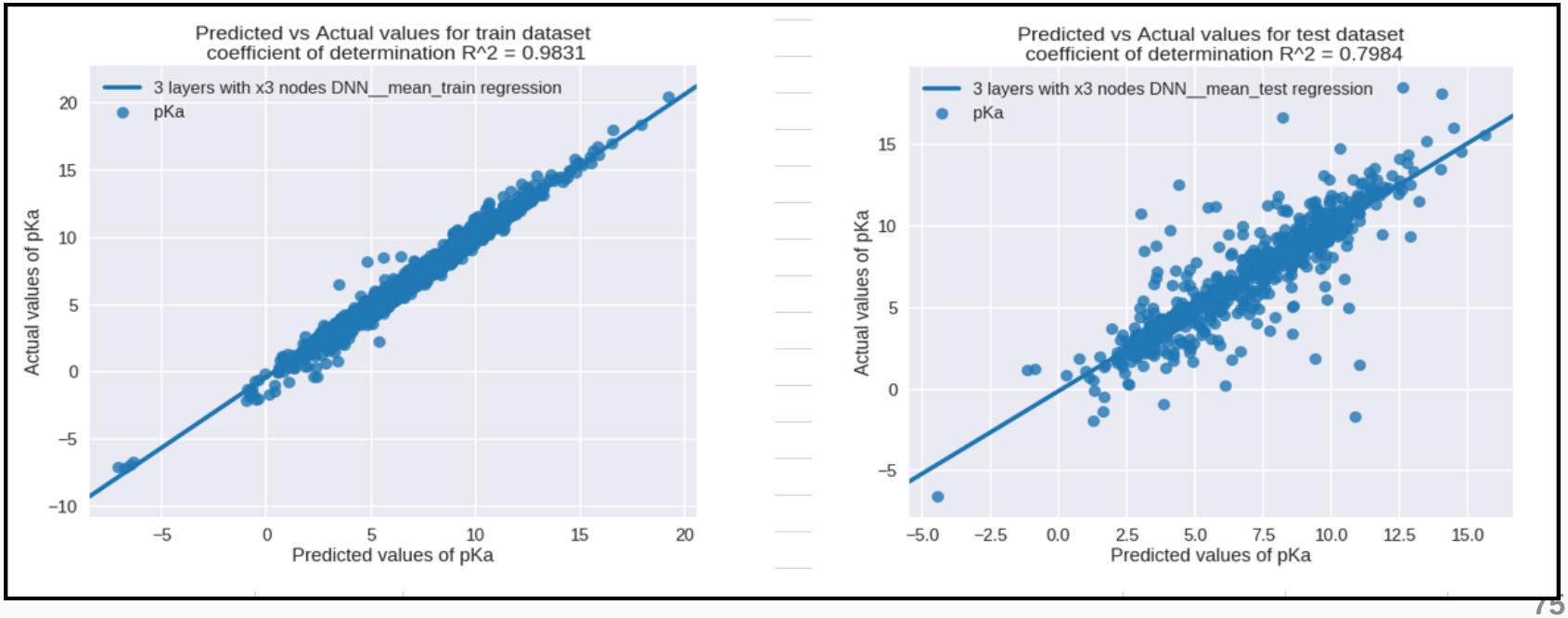


Prototype Development



pKa Prediction Model

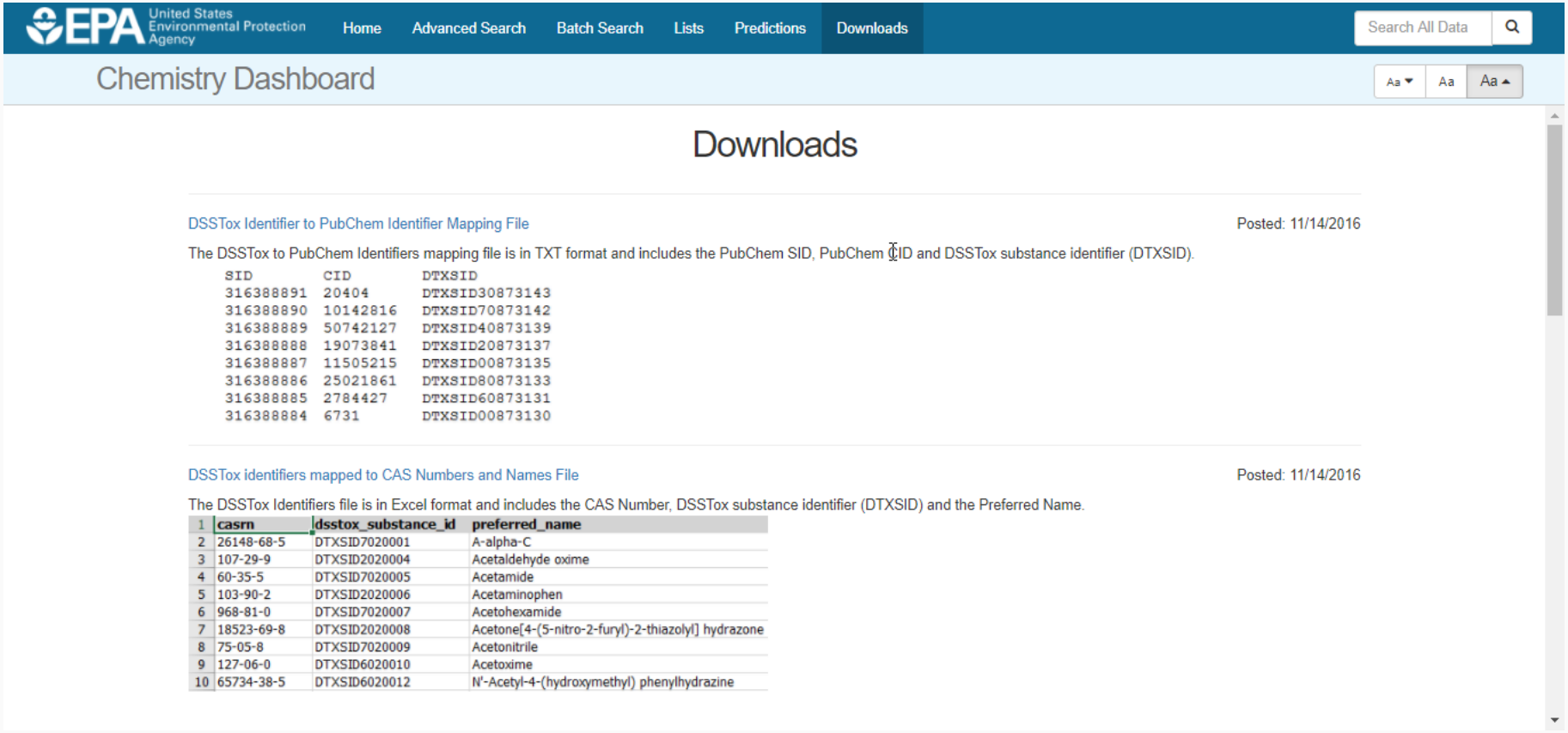
- pKa prediction models based on Open Data Set of 8000 chemicals – acidic, basic and amphoteric chemicals



NCCT “InvitroDB_v3”

- The last public release of ToxCast data (invitroDB_v2) was in 3rd Quarter of 2015
- The next release invitroDB_v3 is Fall 2018
- Data includes new assays, new chemicals, new pipelining, results of data curation
- Data will also release via CompTox Dashboard
- Data will be available at https://www.epa.gov/chemical-research/exploring-toxcast-data-downloadable-data

Downloadable Data Being Updated



Conclusion

- The CompTox Chemistry Dashboard provides access to data for ~765,000 chemicals
- An expanding list of data types and sources has been integrated
- New searches based on Product Use and Categories and Assay and Gene
- The chemical lists of interest grows with each release

- Next release scheduled for Fall 2018 with InvitroDB_v3 data – more chemicals, more assays

How is it built?

https://jcheminf.springeropen.com/articles/10.1186/s13321-017-0247-6

EPA United States Environmental Protection Agency

Williams et al. J Cheminform (2017) 9:61
DOI 10.1186/s13321-017-0247-6

Journal of Cheminformatics

DATABASE Open Access

The CompTox Chemistry Dashboard: a community data resource for environmental chemistry

Antony J. Williams[1*], Christopher M. Grulke[1], Jeff Edwards[1], Andrew D. McEachran[2], Kamel Mansouri[1,2,4], Nancy C. Baker[3], Grace Patlewicz[1], Imran Shah[1], John F. Wambaugh[1], Richard S. Judson[1] and Ann M. Richard[1]

Acknowledgments

- Our NCCT CompTox Chemical Dashboard Development and IT Team
- The NCCT Team of Scientists
- NERL scientists - Mass Spectrometry
- Kamel Mansouri – OPERA models
- Todd Martin – TEST predictions

Contact

Antony Williams
US EPA Office of Research and Development
National Center for Computational Toxicology (NCCT)
Williams.Antony@epa.gov
ORCID: https://orcid.org/0000-0002-2668-4821