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

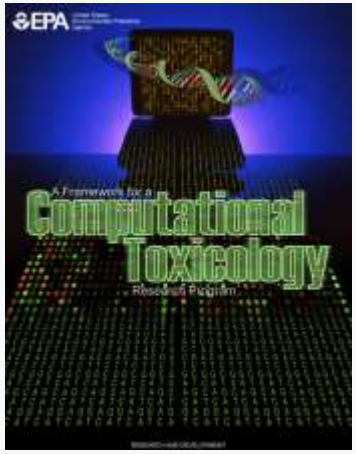
Applications of the US EPA's CompTox Dashboard to support structure identification and chemical forensics using mass spectrometry

Antony Williams[1], Andrew D. McEachran[2], Jon R. Sobus[3] and Emma Schymanski[4]

1) *National Center for Computational Toxicology, U.S. Environmental Protection Agency, RTP, NC*
2) *Oak Ridge Institute of Science and Education (ORISE) Research Participant, RTP, NC*
3) *National Exposure Research Laboratory, U.S. Environmental Protection Agency, RTP, NC*
4) *Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, Campus Belval, Luxembourg*

August 2018
ACS Fall Meeting, Boston

National Center for Computational Toxicology



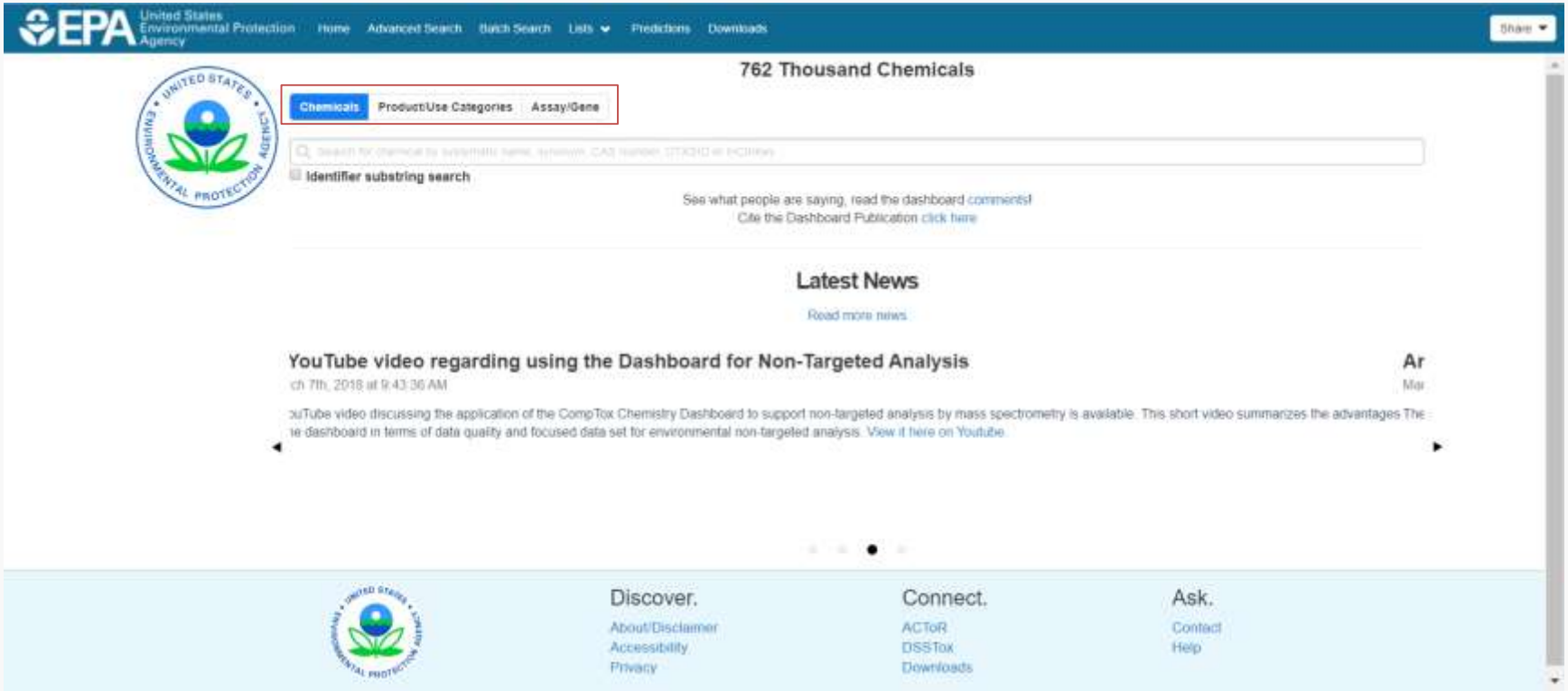
- National Center for Computational Toxicology established in 2005 to integrate:
  - High-throughput and high-content technologies
  - Modern molecular biology
  - Data mining and statistical modeling
  - Computational biology and chemistry
- Researching computational approaches to quickly evaluate the safety of chemicals for potential risk.
- Outputs: a lot of data, models, algorithms and software applications

The CompTox ~~Chemistry~~ Dashboard

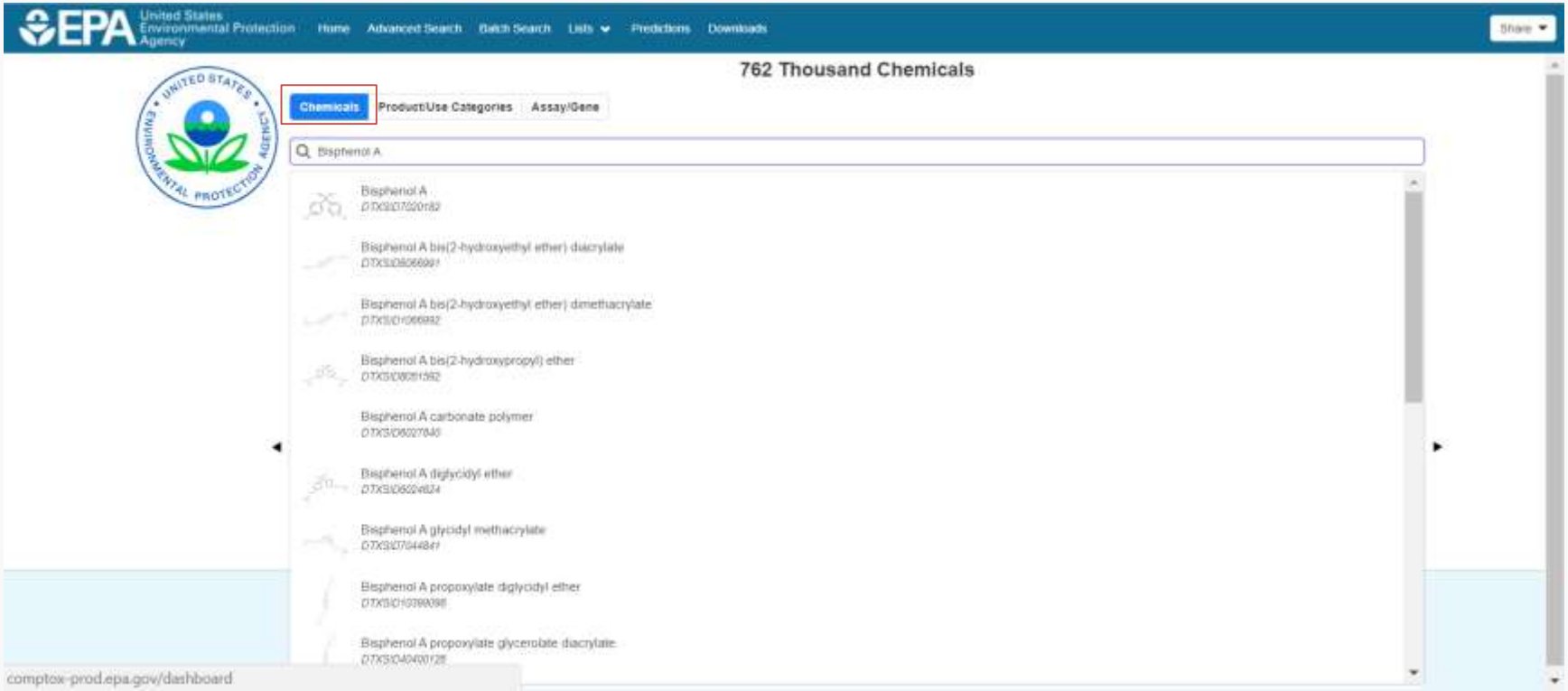
- A publicly accessible website delivering access:
  - New entry portal for all NCCT dashboards
  - ~**762,000** chemicals with related property data
  - **Searchable by chemical, product use, gene and assay (ToxCast)**
  - Experimental and predicted physicochemical property data
  - **“Bioactivity data” for the ToxCast/Tox21 project**
  - **Generalized Read-Across (GenRA) module**
  - Links to other agency websites and public data resources
  - “Literature” searches for chemicals using public resources
  - “Batch searching” for thousands of chemicals
  - DOWNLOADABLE Open Data for reuse and repurposing

CompTox Dashboard

https://comptox.epa.gov/dashboard



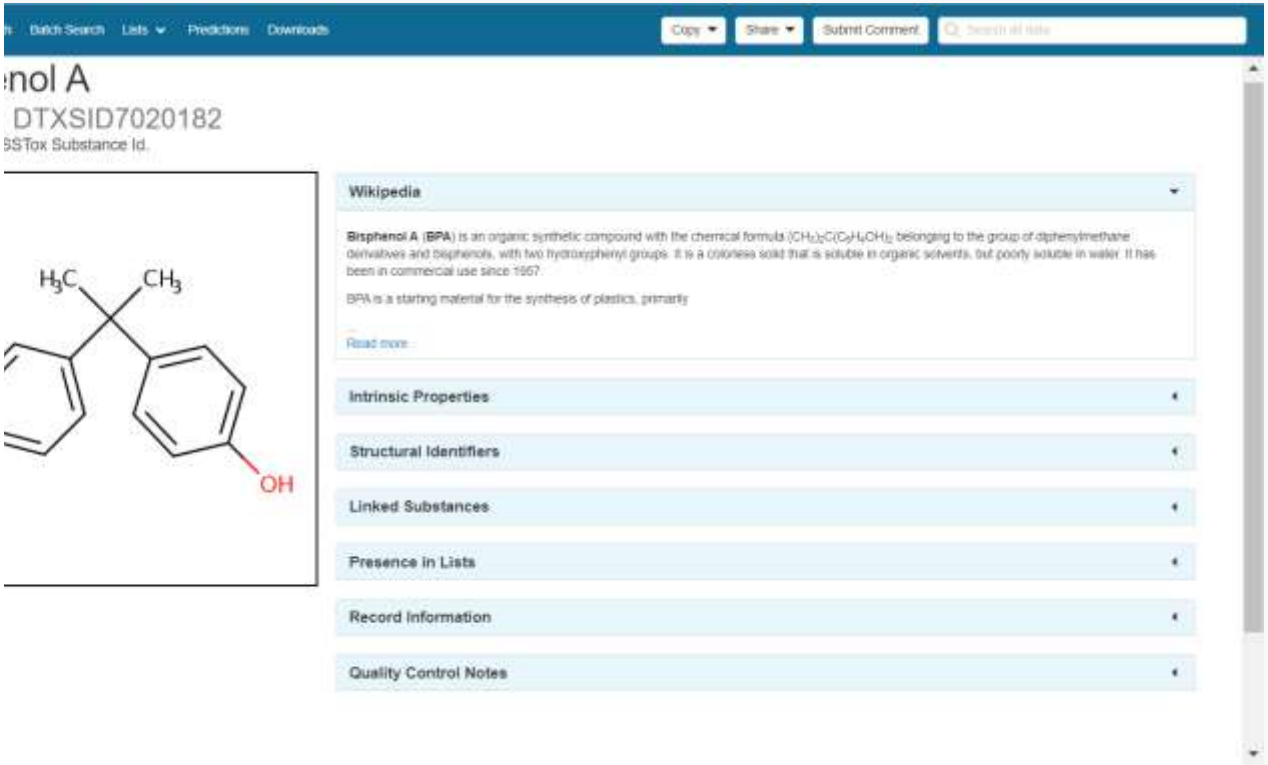
CompTox Dashboard Chemicals



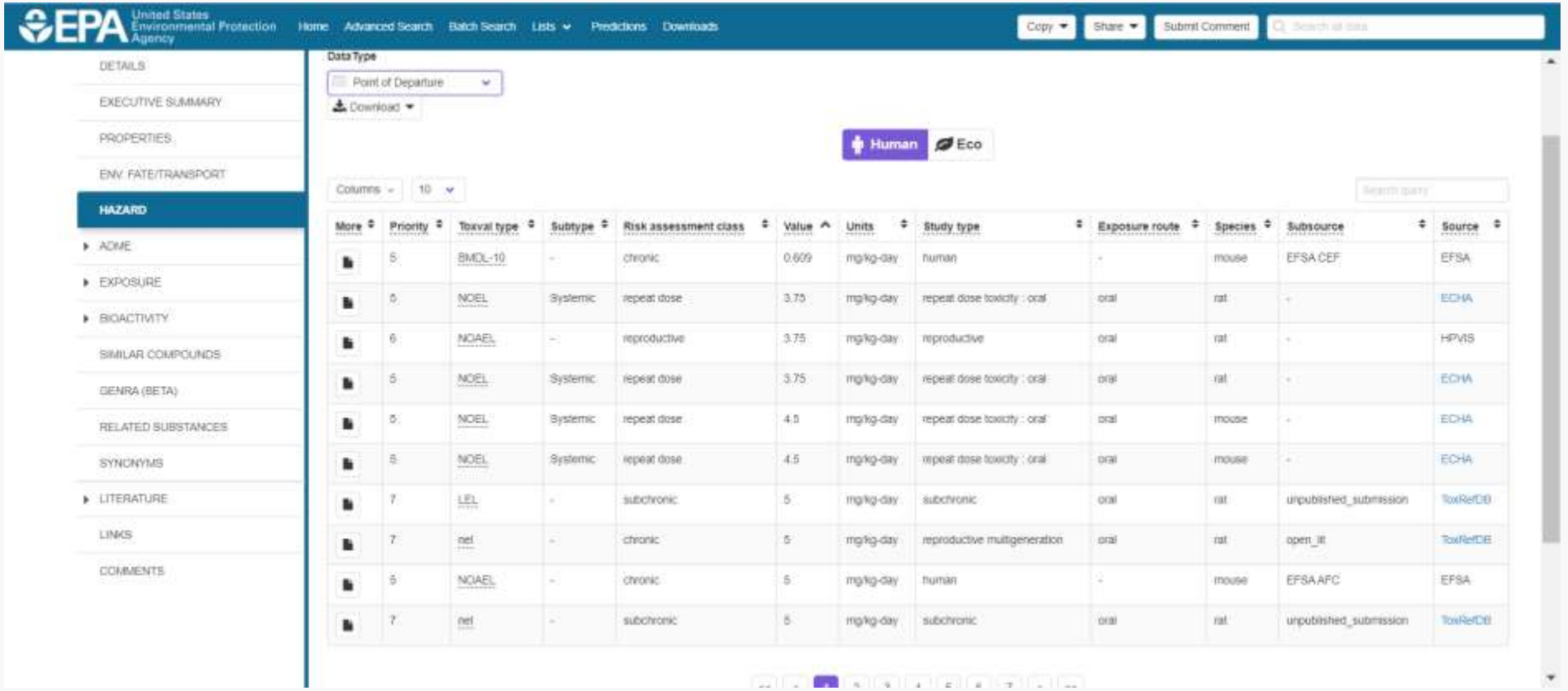
Detailed Chemical Pages

EPA United States Environmental Protection Agency

- **DETAILS**
- EXECUTIVE SUMMARY
- PROPERTIES
- ENV. FATE/TRANSPORT
- HAZARD
- ADME
- EXPOSURE
- BIOACTIVITY
- SIMILAR COMPOUNDS
- GENRA (BETA)
- RELATED SUBSTANCES
- SYNONYMS
- LITERATURE
- LINKS
- COMMENTS

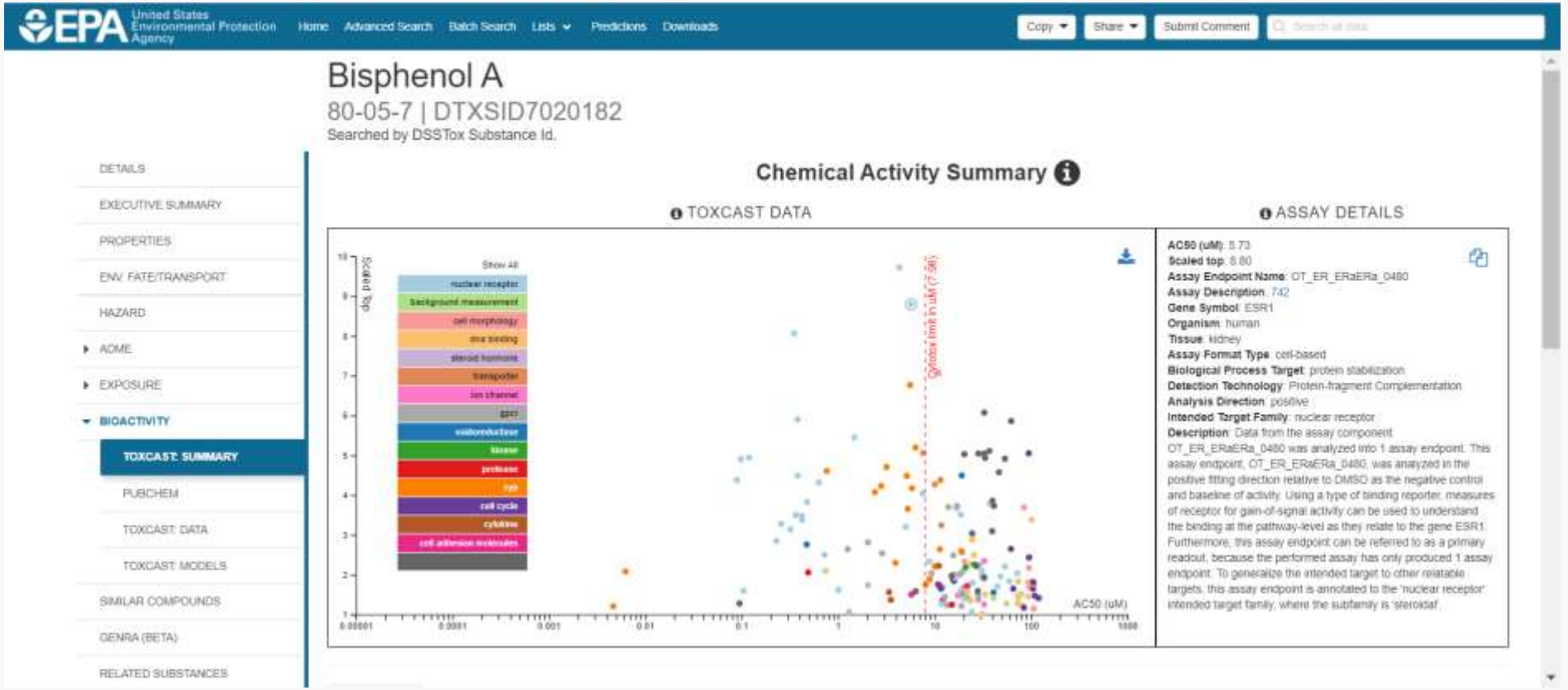


Access to Chemical Hazard Data

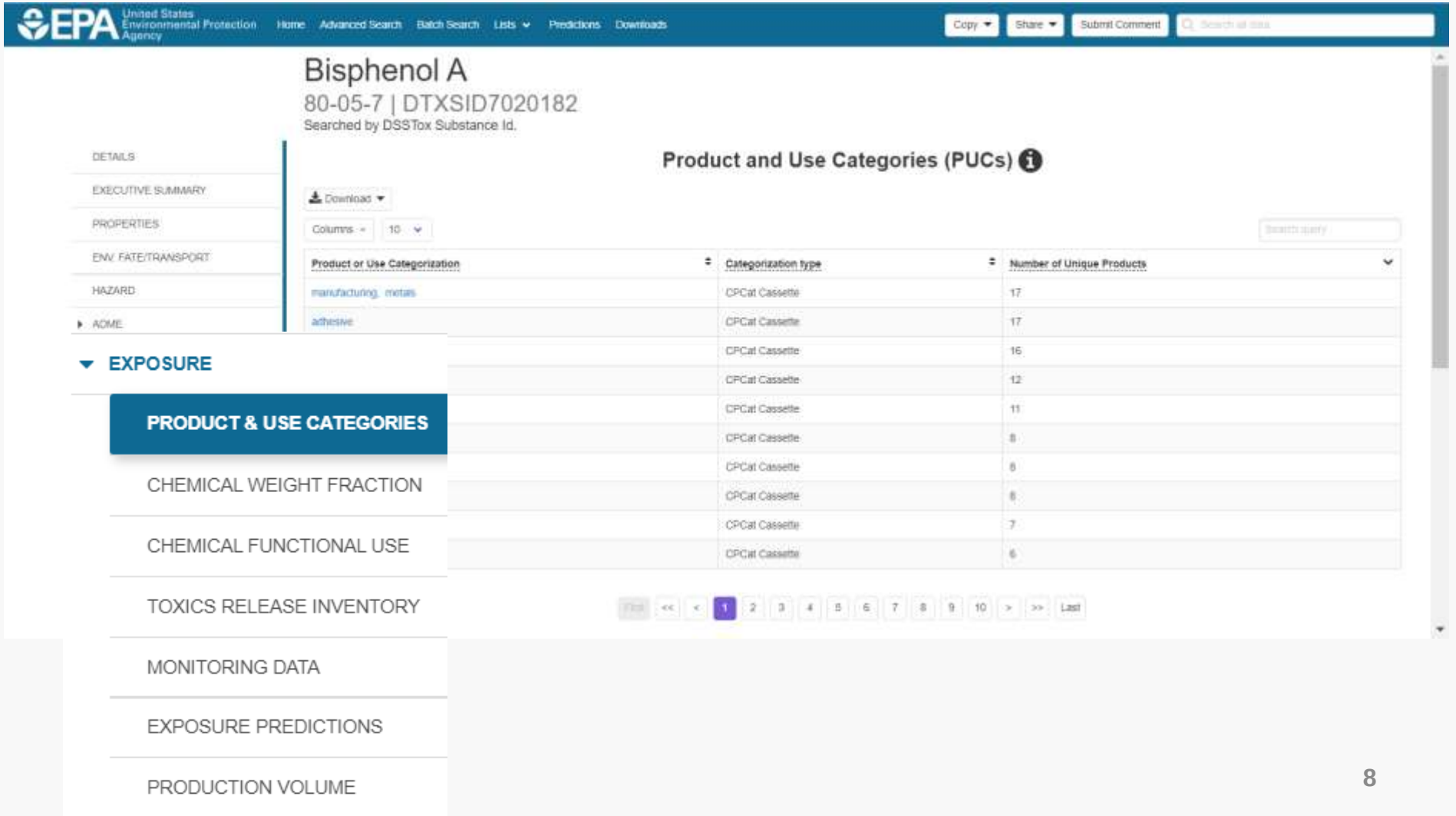


In Vitro Bioassay Screening

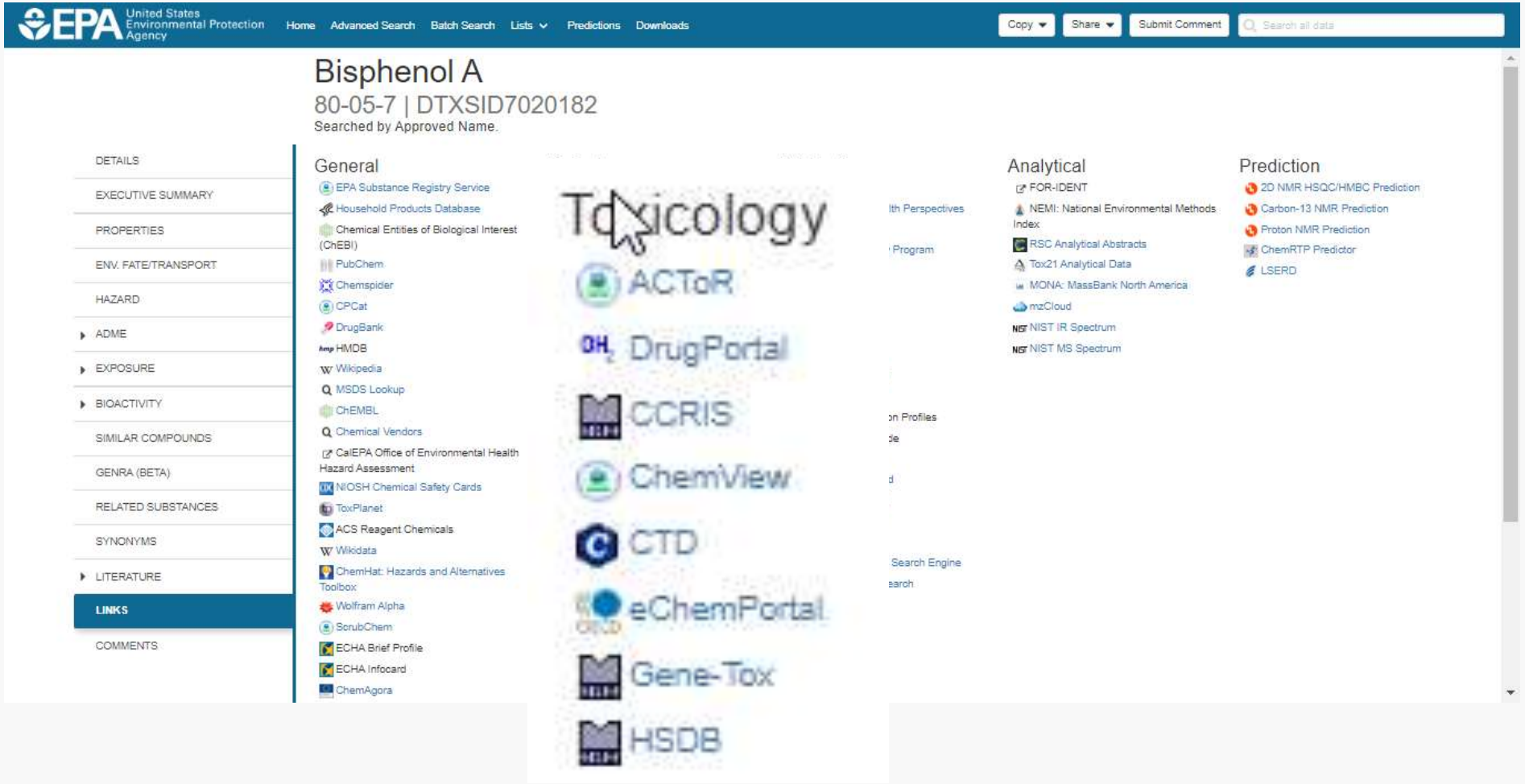
ToxCast and Tox21



Sources of Exposure to Chemicals



External Links to ~80 websites

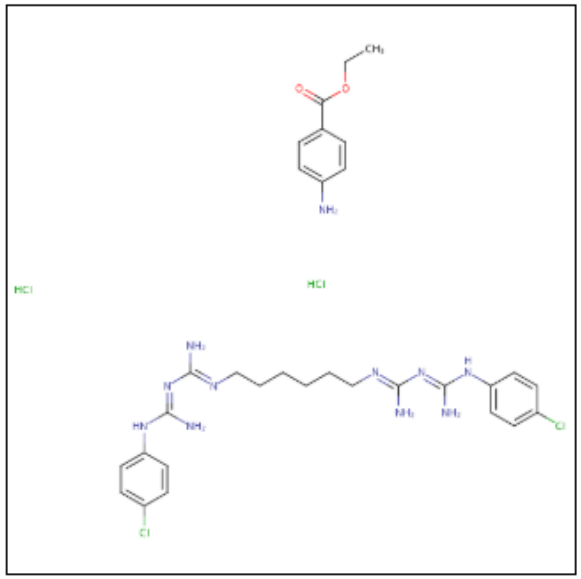


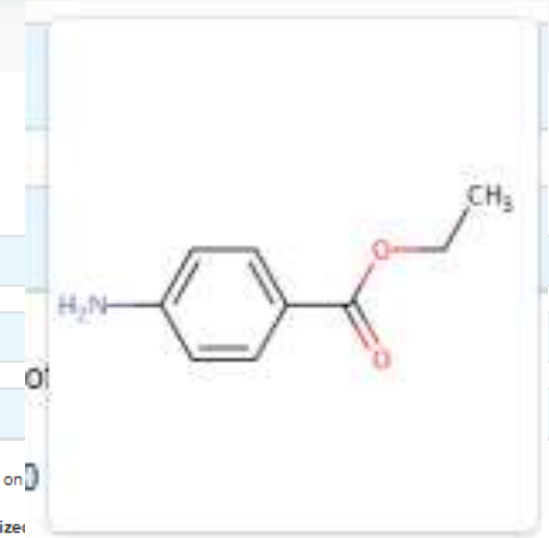
MS-Ready Mappings

Progaron

108532-15-6 | DTXSID20148579

Searched by DSSTox Substance Id.

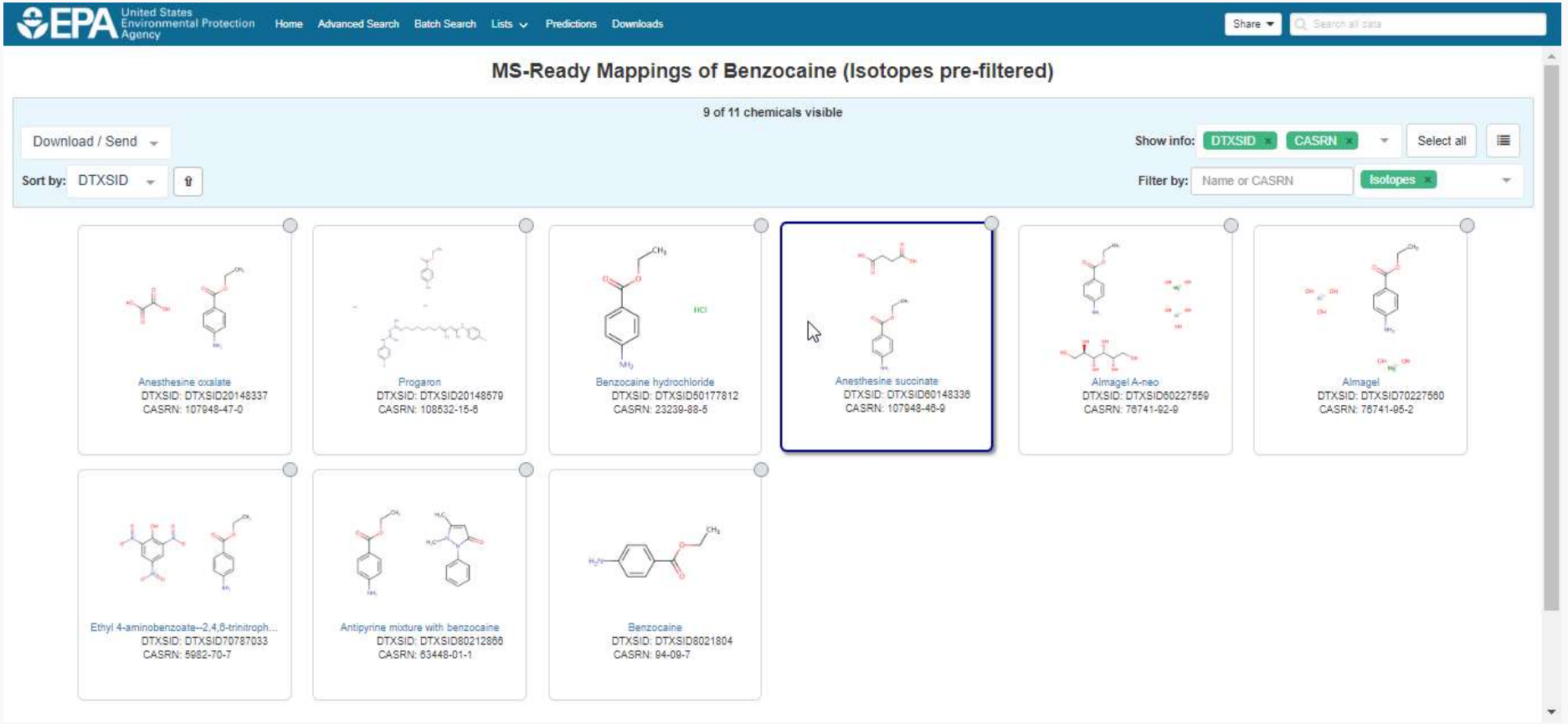




- Intrinsic Properties
- Structural Identifiers
- Linked Substances
  - Same Connectivity: 1 record (based on
  - Mixtures, Components and Neutralize
  - MS-Ready Mappings: DTXCID0013314
  - Similar Compounds: 0 records
- Presence in Lists
- Record Information
- Quality Control Notes

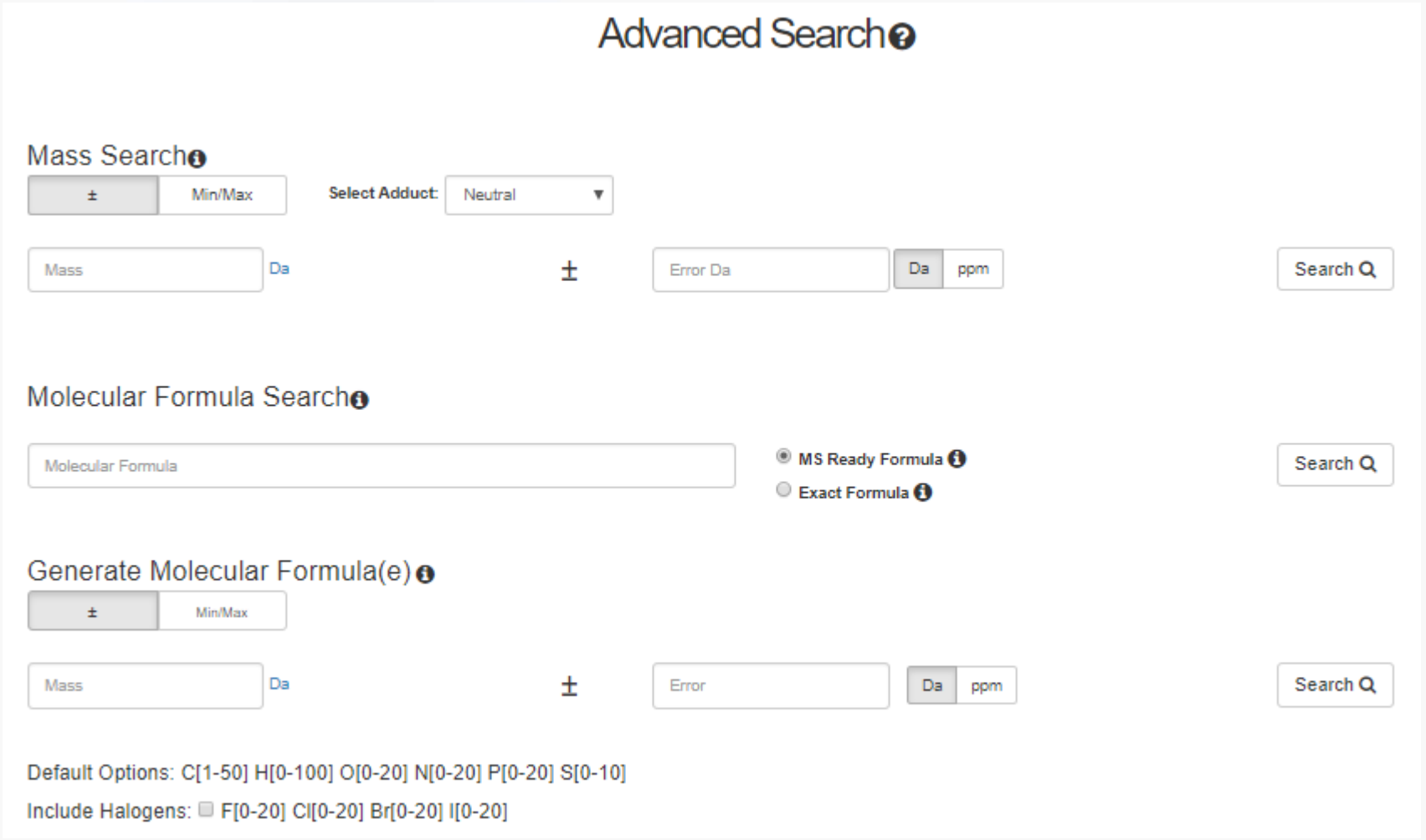
DTXCID301804:11 records; D

MS-Ready Mappings Set

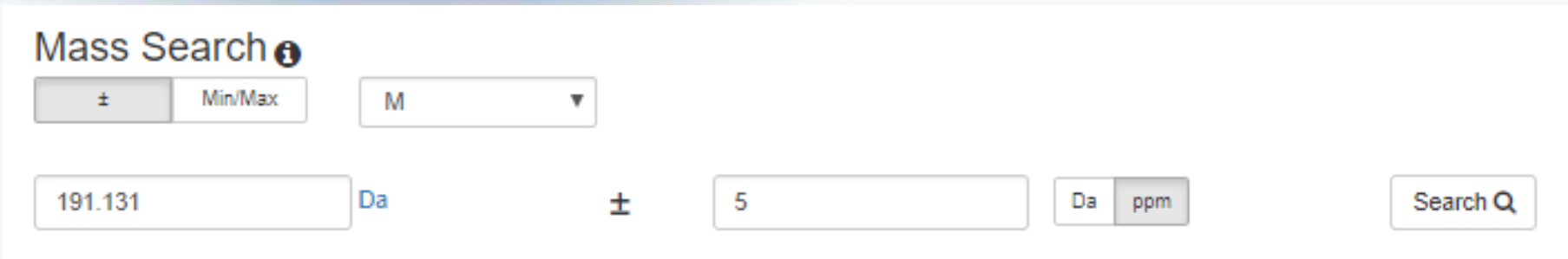


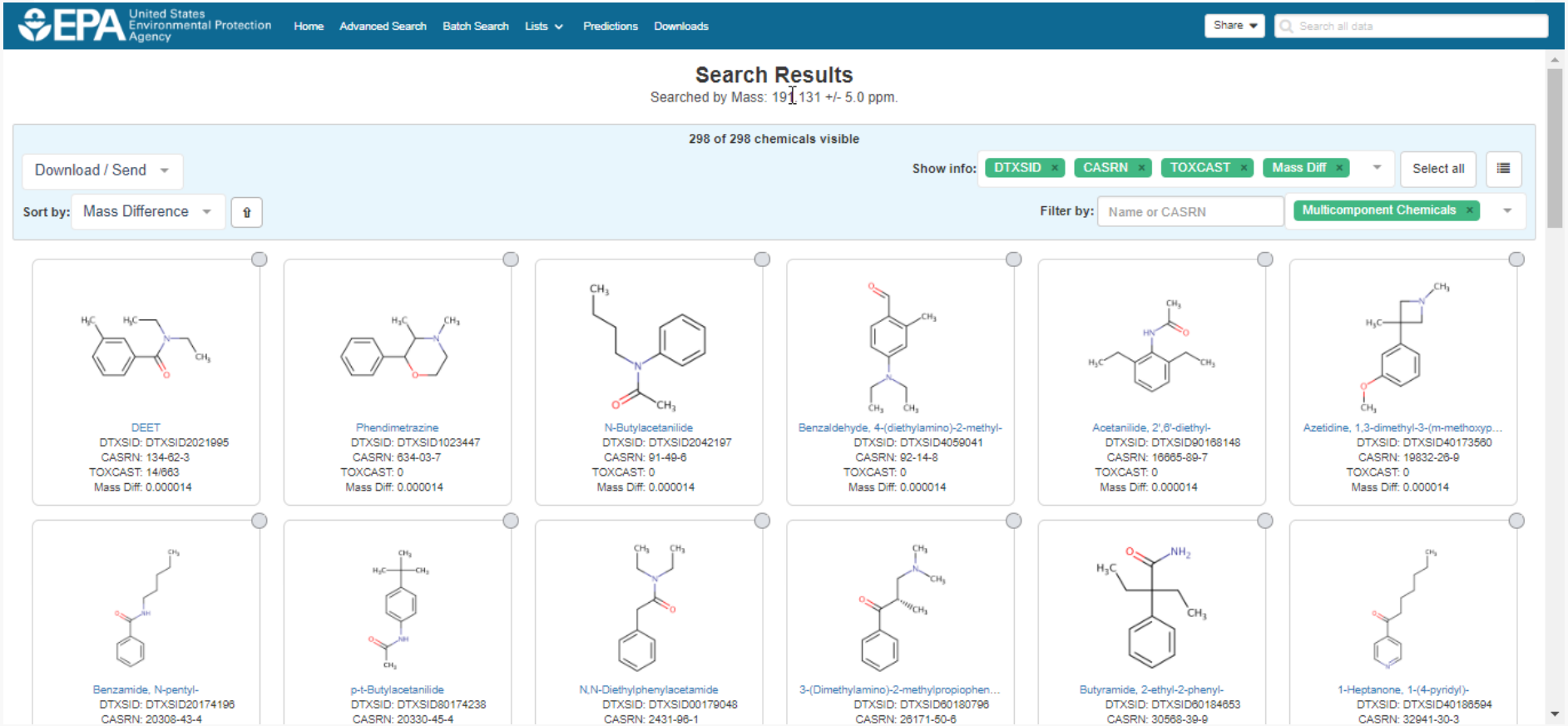
Mass and Formula Searches Supporting Mass Spectrometry

EPA United States Environmental Protection Agency

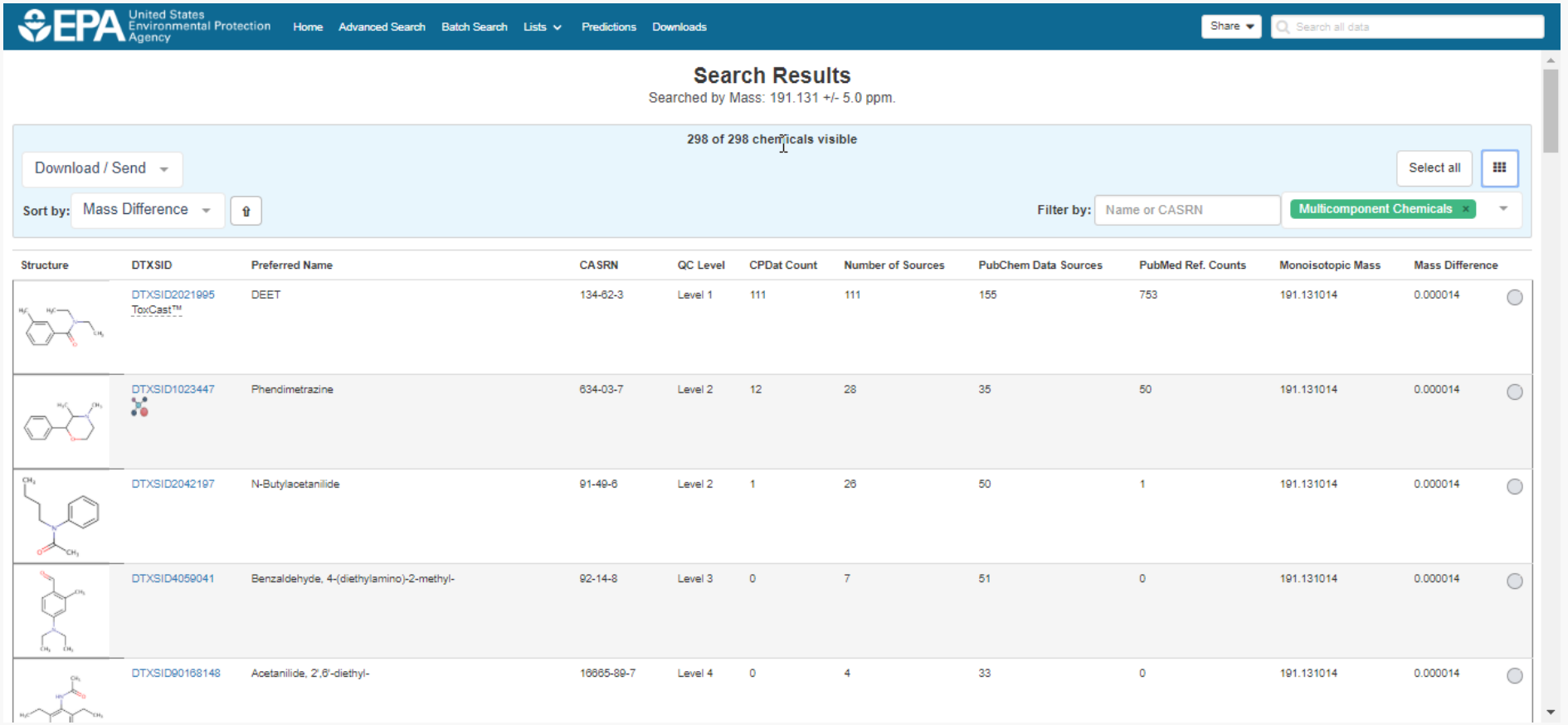


Advanced Searches
Mass Based Search





Advanced Searches
Mass Based Search

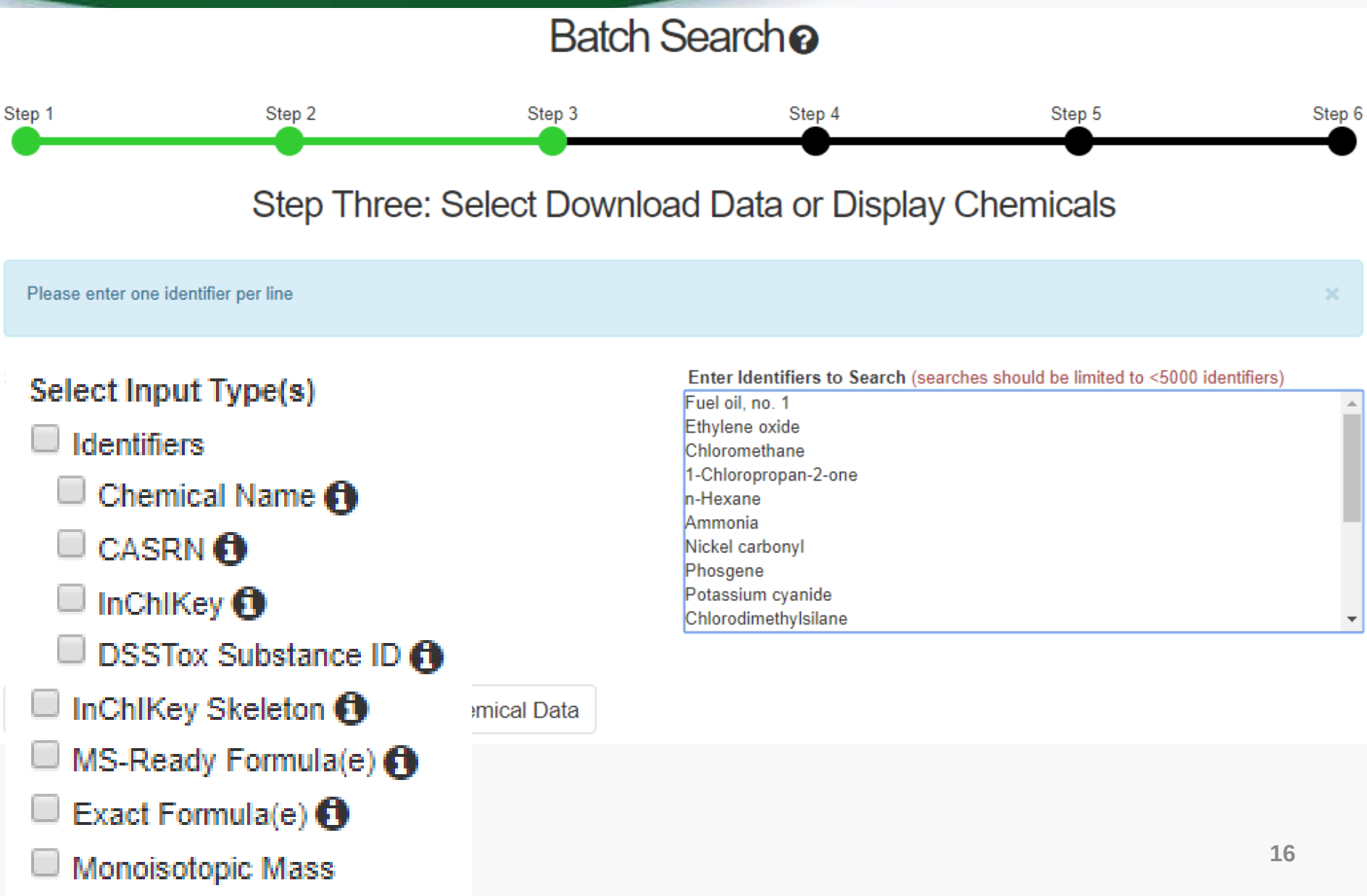


Batch Searching

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

- Typical questions
  - 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

INPUT	FOUND_BY	DTXCID_IN	DATA_SOU	TOXVAL_D	TOXCAST_	TOXCAST_	NUMBER_C	PUBCHEM_	STO
C6H12O3	MS Ready I	DTXCID701	91	Y	0.36	2/562	24	83	Y
C6H12O3	MS Ready I	DTXCID003	67	Y	0.36	1/276	376	80	Y
C6H12O3	MS Ready I	DTXCID106	65	Y	4.42	5/113	6	77	Y
C6H12O3	MS Ready I	DTXCID105	45	Y	0.0	0/163	3	94	-
C6H12O3	MS Ready I	DTXCID901	38	Y	-	-	14	110	Y
C6H12O3	MS Ready I	DTXCID402	34	Y	0.0	0/113	-	53	Y
C6H12O3	MS Ready I	DTXCID202	31	Y	-	-	-	36	Y
C6H12O3	MS Ready I	DTXCID202	30	-	2.54	7/276	-	54	-
C6H12O3	MS Ready I	DTXCID109	26	Y	-	-	-	46	-
C6H12O3	MS Ready I	DTXCID202	24	Y	0.0	0/113	-	47	-
C6H12O3	MS Ready I	DTXCID303	22	Y	-	-	-	89	-
C6H12O3	MS Ready I	DTXCID302	20	Y	-	-	2	25	Y
C6H12O3	MS Ready I	DTXCID407	19	Y	-	-	12	62	-
C6H12O3	MS Ready I	DTXCID704	17	Y	-	-	-	64	-
C6H12O3	MS Ready I	DTXCID704	16	Y	-	-	3	49	-

Suspect Screening and Non-Targeted Analysis Workflow

EPA United States Environmental Protection Agency

Suspect Screening

- Raw Samples
- Extracted Samples
- Raw Features
- "Molecular Features"
- DSSTox Chemical Database
- Matched Formulas
- Mapped Structures
- Prioritized Structures (using ToxPi)
- Confirmed Structures (using ToxCast standards)
- Predicted Concentrations

Non-Targeted Analysis

- Processed Features
- Prioritized Features
- Predicted Formulas
- Candidate Structures
- Sorted Structures
- Predicted Retention Times
- Predicted Mass Spectra
- Predicted/Observed Functional Use
- Predicted/Observed Media Occurrence
- Methodological Concordance
- Top Candidate Structure(s)

Color Key

Red = Analytical Chemistry

Blue = Data Processing & Analysis

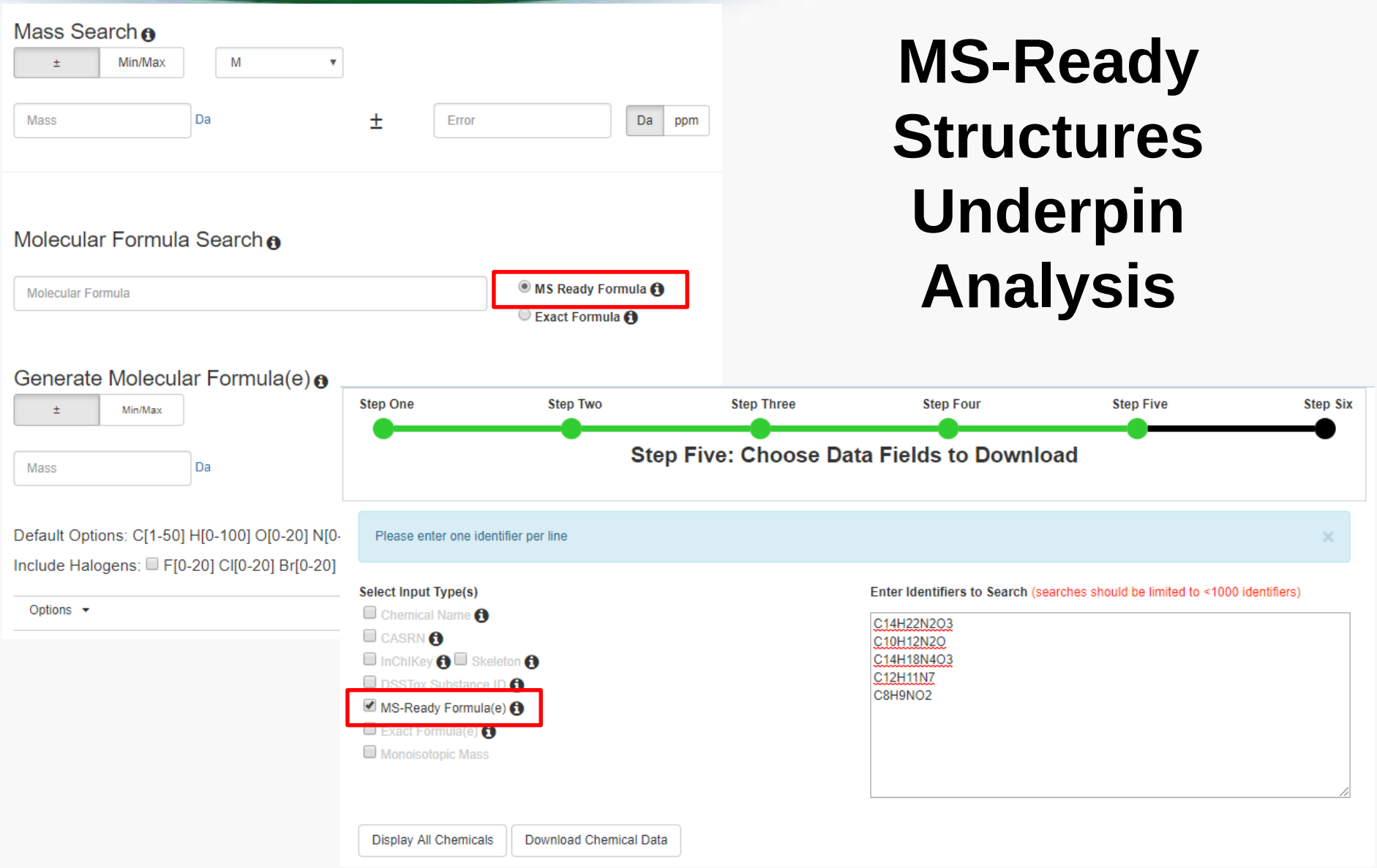
Purple = Mathematical & QSPR Modeling

Green = Informatics & Web Services

Content from J. Sobus, US-EPA-NERL

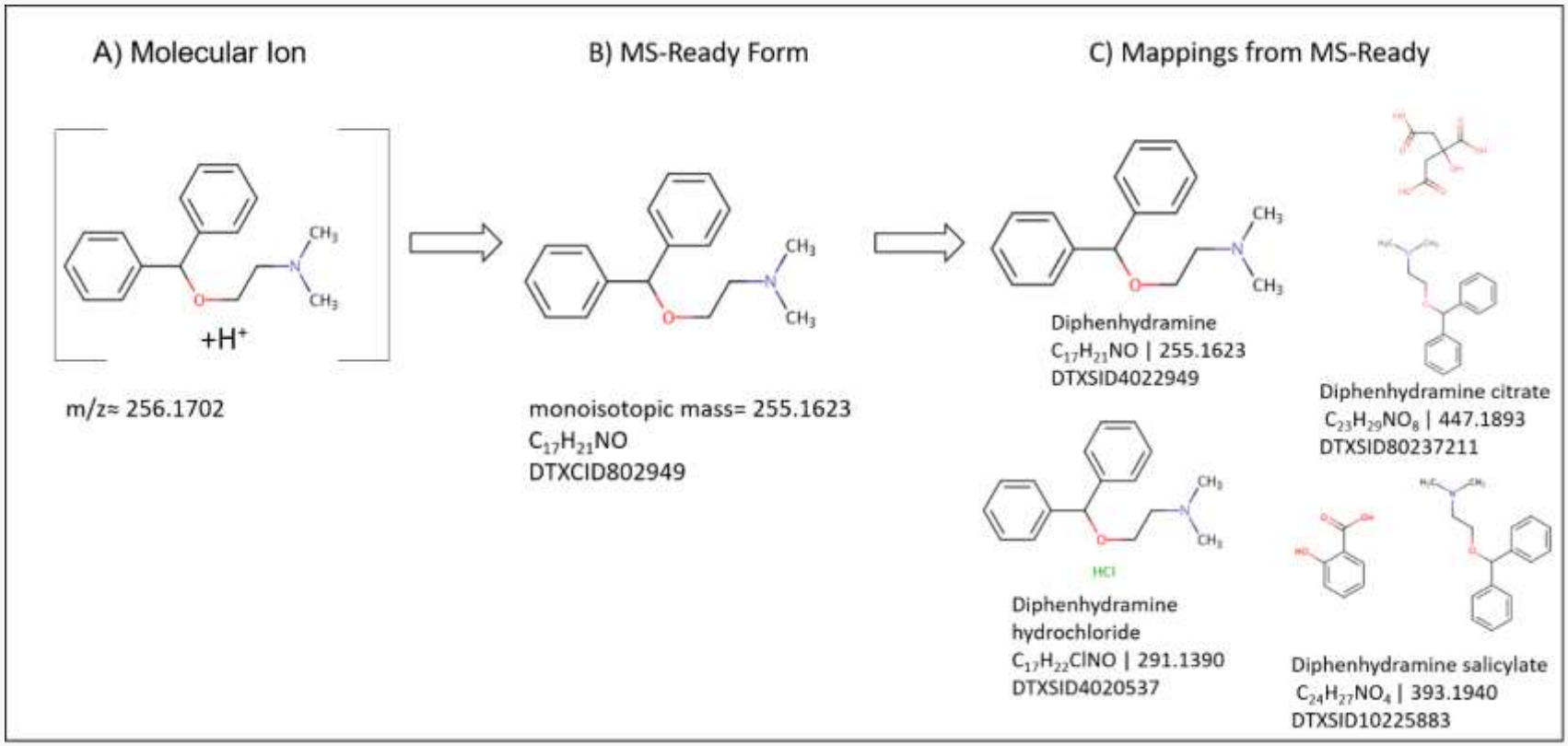
The Dashboard to Support MS-Analysis

EPA United States Environmental Protection Agency



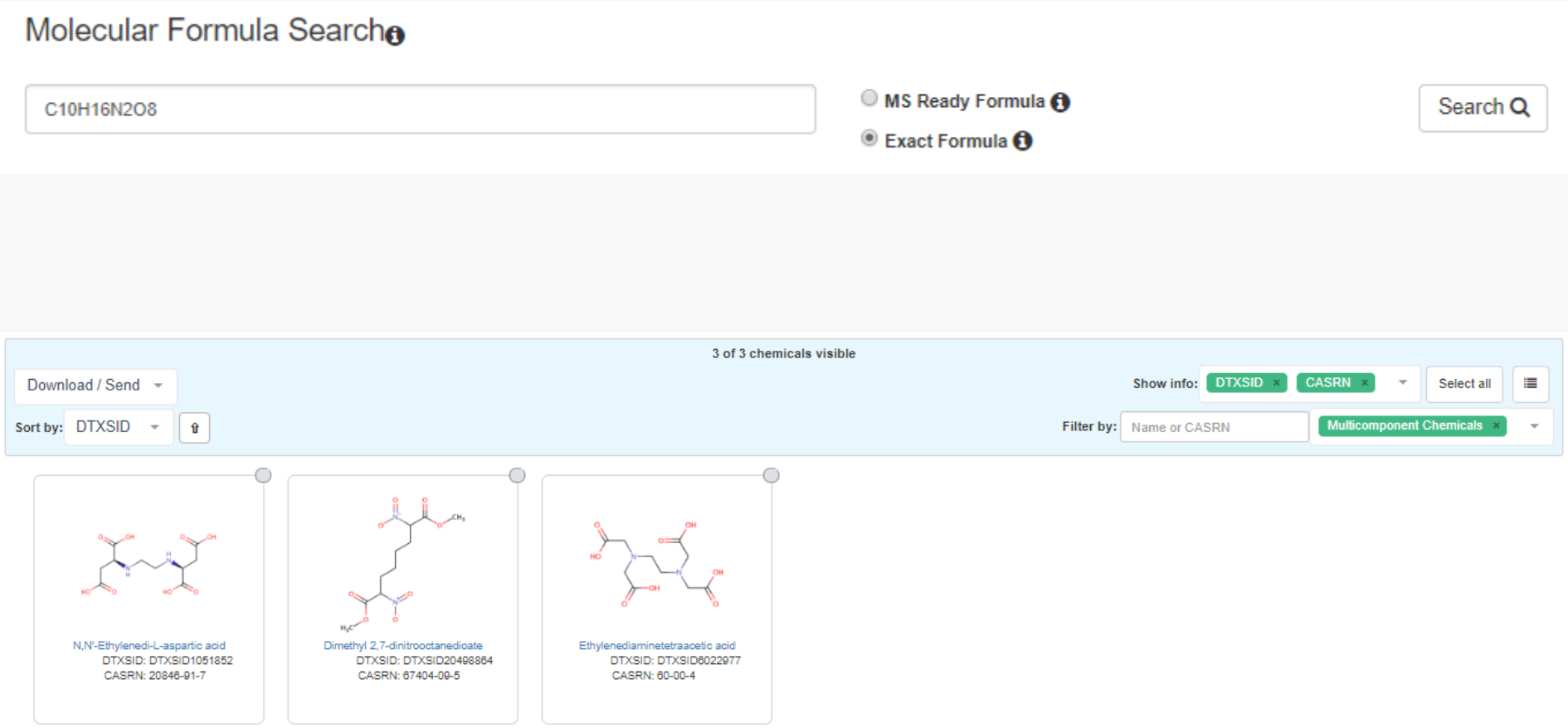
MS-Ready Structures Underpin Analysis

Specific Data-Mappings "MS-Ready Structures"



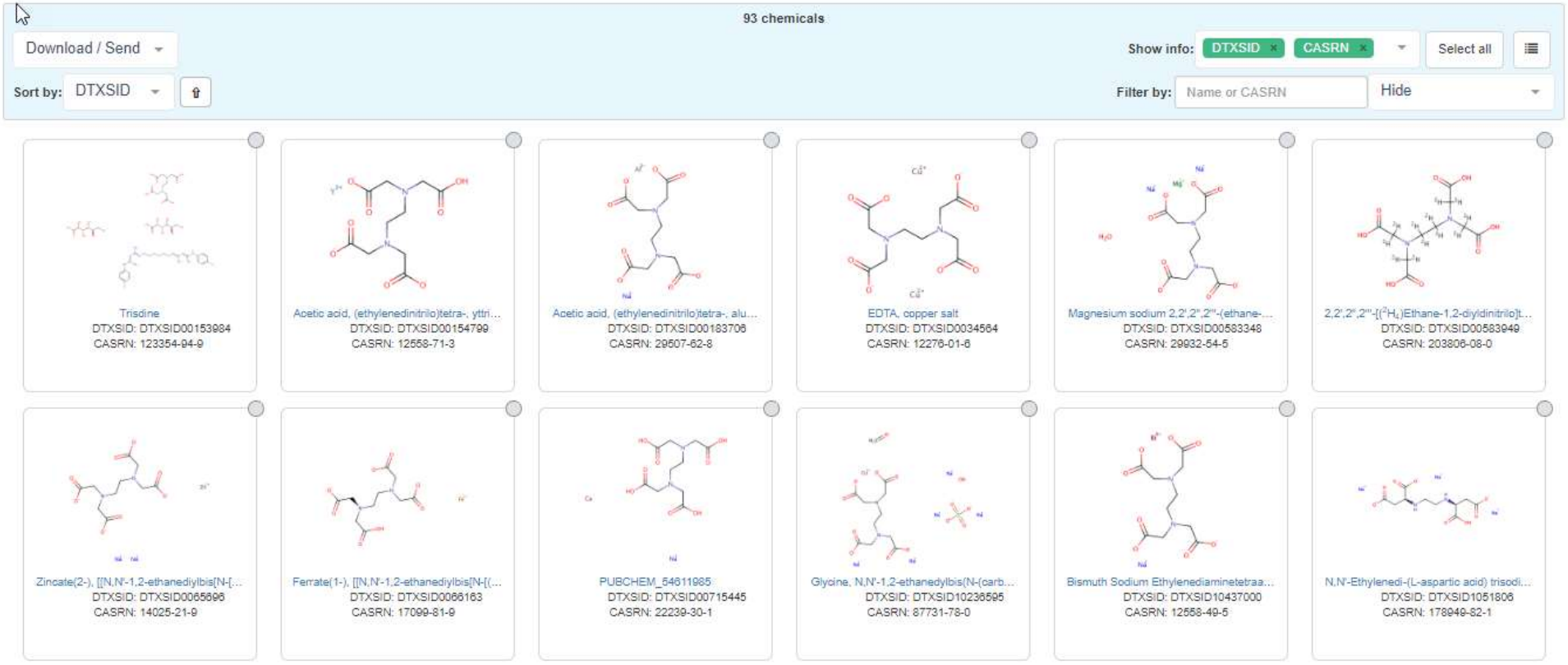
MS-Ready Mappings

- Input Formula: C10H16N2O8



MS-Ready Mappings

- **Same** Input Formula: C10H16N2O8
- MS Ready Formula Search: 93 Chemicals



MS-Ready Mappings

- 93 chemicals returned in total
  - Only 7 of the 93 are single component chemicals
  - Only 4 of the 93 are non-isotope-labeled
  - 3 are neutral compounds and 1 is charged

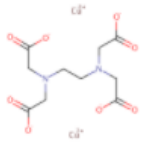
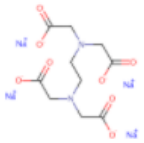
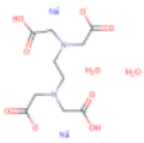
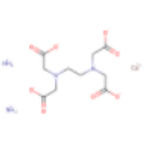
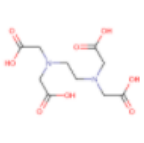
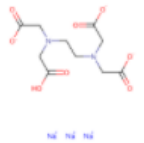
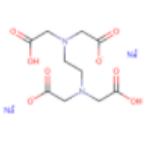
Complexity to Simplicity
93 Chemicals – 7 in EPAHFR

EPA United States Environmental Protection Agency

	INPUT	DTXCID_INDIVIDU	FORMULA_	SMILES_I	DTXSID	CASRN	EXPOCAST_M	EXPOCAST	DATA_SOURCI	TOXVAL_D	TOXCAST_	TOXCAST	# OF PUBMED	PUBCHEM	EPAHFR
2	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID6022977	60-00-4	7.96e-05	Y	71	Y	2.65	3/113	25251	158	Y
3	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID9027073	139-33-3	-	-	41	Y	-	-	25251	56	Y
4	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID3026350	64-02-8	-	-	37	Y	-	-	-	57	Y
5	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID7020556	150-38-9	-	-	30	Y	-	-	-	33	Y
6	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID5049609	67989-88-2	-	-	20	Y	-	-	-	8	Y
7	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID5049576	6381-92-6	-	-	19	Y	-	-	25251	31	Y
8	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID0034564	12276-01-6	-	-	11	-	-	-	-	8	Y
9	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID5027774	15708-41-5	-	-	48	Y	1.98	6/303	241	53	
10	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID2036409	62-33-9	4.64e-06	Y	37	Y	0.0	0/64	25251	42	-
11	C10H16N2O8	DTXCID00197424	C10H16N2(	OC(=O)C(	DTXSID1051852	20846-91-7	-	-	36	Y	-	-	89	25	-
12	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID6042107	15375-84-5	-	-	25	Y	-	-	97	25	-
13	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID3036442	5964-35-2	-	-	23	Y	-	-	25251	25	-
14	C10H16N2O8	DTXCID00197424	C10H16N2(	OC(=O)C(	DTXSID1051806	178949-82-1	-	-	22	Y	-	-	-	5	-
15	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID0065696	14025-21-9	-	-	22	Y	-	-	-	43	-
16	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID9027813	20824-56-0	-	-	21	Y	-	-	-	12	-
17	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID9027815	21265-50-9	-	-	20	Y	-	-	241	24	-
18	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID5058272	17421-79-3	-	-	19	Y	-	-	25251	25	-
19	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID3058612	2001-94-7	-	-	18	Y	-	-	25251	19	-
20	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID8027820	22473-78-5	-	-	16	Y	-	-	-	11	-
21	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID8058324	17572-97-3	-	-	15	-	-	-	-	36	-
22	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID8028343	67859-51-2	-	-	14	Y	-	-	-	5	-
23	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID4051328	13235-36-4	-	-	14	-	-	-	-	18	-
24	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID6070980	68015-77-0	-	-	14	Y	-	-	-	13	-
25	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID9058317	15934-01-7	-	-	11	-	-	-	-	5	-
26	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID0066163	17099-81-9	-	-	11	-	-	-	241	14	-
27	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID1068988	54959-35-2	-	-	11	-	-	-	241	14	-
28	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID5074266	60816-63-9	-	-	11	-	-	-	1	10	-
29	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID4048197	39208-15-6	-	-	10	-	-	-	-	28	-
30	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID2065830	14931-83-0	-	-	10	-	-	-	47	9	-
31	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID70189997	36499-65-7	-	-	10	-	-	-	25298	26	-
32	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID7051420	61916-40-3	-	-	9	-	-	-	-	4	-
33	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID2051425	73513-47-0	-	-	8	Y	-	-	-	3	-
34	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID7051426	73637-19-1	-	-	8	Y	-	-	-	5	-
35	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID2051427	73637-20-4	-	-	8	Y	-	-	-	-	-
36	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID3058741	10378-23-1	-	-	8	Y	-	-	-	31	-
37	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID6065925	15708-48-2	-	-	8	-	-	-	-	19	-
38	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID20217976	6766-87-6	-	-	8	-	-	-	-	13	-
39	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID5065807	14689-29-3	-	-	7	-	-	-	-	12	-
40	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID6069408	60544-70-9	-	-	7	-	-	-	-	12	-
41	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID00153984	123354-94-9	-	-	7	-	-	-	2	6	-
42	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID70190705	37209-61-3	-	-	7	-	-	-	6	9	-
43	C10H16N2O8	DTXCID902977	C10H16N2(	OC(=O)CI	DTXSID7051424	67401-50-7	-	-	6	-	-	-	-	4	-

Complexity to Simplicity
93 Chemicals – 7 in the list

EPA United States Environmental Protection Agency

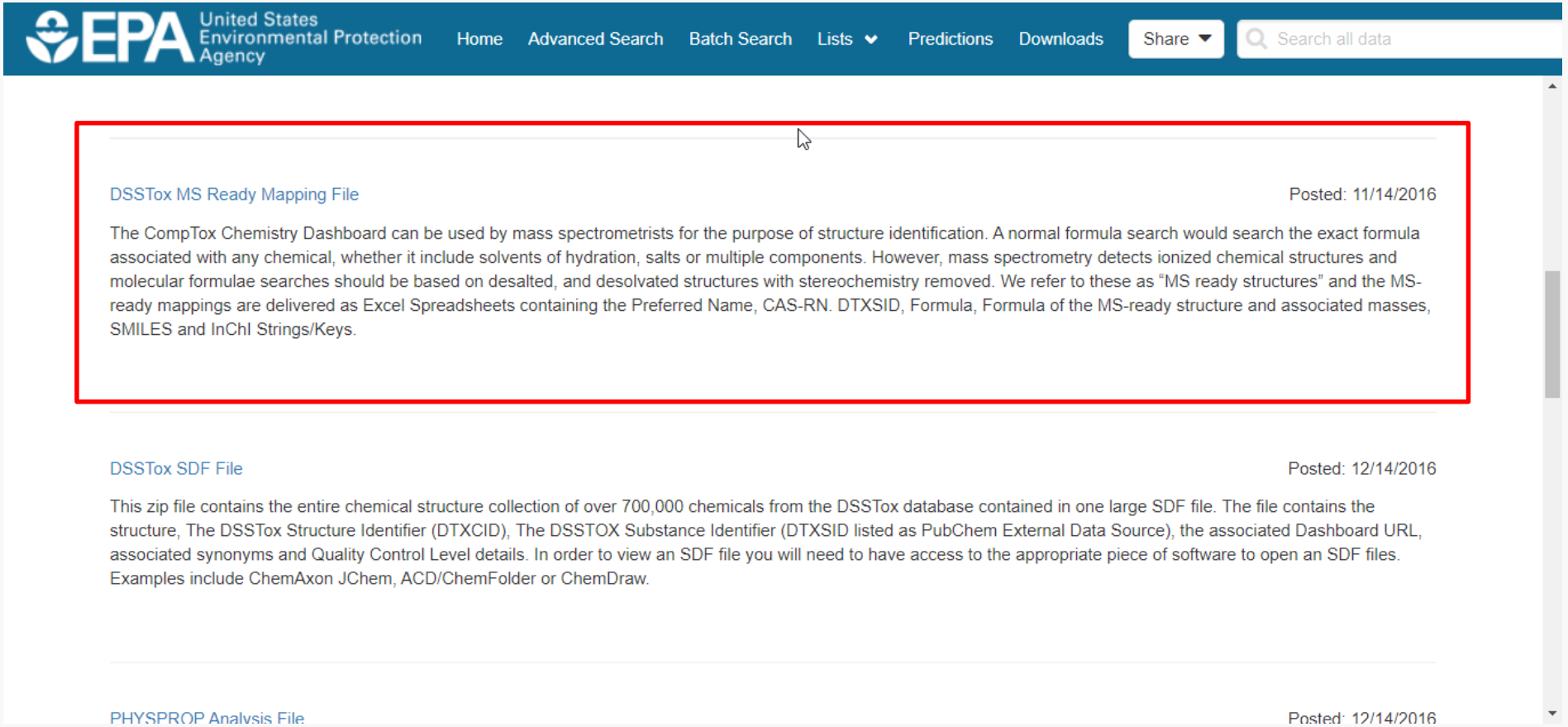
Structure	DTXSID	Preferred Name	CASRN	QC Level	CPDat Count	Number of Sources	PubChem Data Source	PubMed Data Source	Monoisotopic Mass
	DTXSID0034564	EDTA, copper salt	12276-01-6	Level 1	10	11	8	0	413.918561
	DTXSID3026350	Ethylenediaminetetraacetic acid tetrasodium salt	64-02-8	Level 1	1227	37	57	0	380.018442
	DTXSID5049576	Disodium ethylenediaminetetraacetate dihydrate	6381-92-6	Level 1	93	19	31	25251	372.075683
	DTXSID5049609	Ethylenediaminetetraacetic acid, diammonium copper salt	67989-88-2	Level 2	9	20	8	0	387.057712
	DTXSID6022977	Ethylenediaminetetraacetic acid	60-00-4	Level 1	346	71	158	25251	292.090665
	DTXSID7020556	Trisodium ethylenediaminetetraacetate	150-38-9	Level 1	85	30	33	0	358.036498
	DTXSID9027073	Ethylenediaminetetraacetic acid, disodium salt	139-33-3	Level 1	1358	41	56	25251	336.054554

Searching batches
Formula (or mass) searching

EPA United States Environmental Protection Agency

	A	B	C	D	E	F	G
1	INPUT	DTXSID	CASRN	PREFERRED NAME	MOL FORMULA	MONOISOTOPIC MASS	DATA SOURCES
2	C14H22N2O3	DTXSID2022628	29122-68-7	Atenolol	C14H22N2O3	266.163042576	46
3	C14H22N2O3	DTXSID0021179	6673-35-4	Practolol	C14H22N2O3	266.163042576	32
4	C14H22N2O3	DTXSID4048854	841-73-6	Bucolome	C14H22N2O3	266.163042576	20
5	C14H22N2O3	DTXSID1045407	13171-25-0	Trimetazidine dihydrochloride	C14H24Cl2N2O3	338.116398	19
6	C14H22N2O3	DTXSID0045753	56715-13-0	R-(+)-Atenolol	C14H22N2O3	266.163042576	19
7	C14H22N2O3	DTXSID2048531	5011-34-7	Trimetazidine	C14H22N2O3	266.163042576	14
8	C14H22N2O3	DTXSID10239405	93379-54-5	Esatenolol	C14H22N2O3	266.163042576	12
9	C14H22N2O3	DTXSID50200634	52662-27-8	N-(2-Diethylaminoethyl)-2-(4-hydroxyphenoxy)acetamide	C14H22N2O3	266.163042576	7
10	C14H22N2O3	DTXSID4020111	51706-40-2	dl-Atenolol hydrochloride	C14H23ClN2O3	302.1397203	6
11	C14H22N2O3	DTXSID1068693	51963-82-7	Benzenamine, 2,5-diethoxy-4-(4-morpholinyl)-	C14H22N2O3	266.163042576	5
12	C18H34N2O6S	DTXSID3023215	154-21-2	Lincomycin	C18H34N2O6S	406.213757997	35
13	C18H34N2O6S	DTXSID7047803	859-18-7	Lincomycin hydrochloride	C18H35ClN2O6S	442.1904357	22
14	C18H34N2O6S	DTXSID20849438	1398534-62-7	PUBCHEM_71432748	C18H35ClN2O6S	442.1904357	1
15	C10H12N2O	DTXSID1047576	486-56-6	Cotinine	C10H12N2O	176.094963014	40
16	C10H12N2O	DTXSID8075330	50-67-9	Serotonin	C10H12N2O	176.094963014	22
17	C10H12N2O	DTXSID8044412	2654-57-1	4-Methyl-1-phenylpyrazolidin-3-one	C10H12N2O	176.094963014	18
18	C10H12N2O	DTXSID80165186	153-98-0	Serotonin hydrochloride	C10H13ClN2O	212.0716407	11
19	C10H12N2O	DTXSID2048870	29493-77-4	(4R,5S)-4-methyl-5-phenyl-4,5-dihydro-1,3-oxazol-2-amine	C10H12N2O	176.094963014	10
20	C10H12N2O	DTXSID10196105	443-31-2	6-Hydroxytryptamine	C10H12N2O	176.094963014	9
21	C10H12N2O	DTXSID90185693	31822-84-1	1,4,5,6-Tetrahydro-5-phenoxypyrimidine	C10H12N2O	176.094963014	7
22	C10H12N2O	DTXSID40178777	2403-66-9	2-Benzimidazolepropanol	C10H12N2O	176.094963014	7
23	C10H12N2O	DTXSID80157026	13140-86-8	N-Cyclopropyl-N'-phenylurea	C10H12N2O	176.094963014	6
24	C10H12N2O	DTXSID30205607	570-14-9	4-Hydroxytryptamine	C10H12N2O	176.094963014	6
25	C14H18N4O3	DTXSID5023900	17804-35-2	Benomyl	C14H18N4O3	290.137890456	68
26	C14H18N4O3	DTXSID3023712	738-70-5	Trimethoprim	C14H18N4O3	290.137890456	51
27	C14H18N4O3	DTXSID40209671	60834-30-2	Trimethoprim hydrochloride	C14H19ClN4O3	326.1145682	8
28	C14H18N4O3	DTXSID70204210	55687-49-5	Benzenemethanol, 4-((2,4-diamino-5-pyrimidinyl)methyl)-2,	C14H18N4O3	290.137890456	5
29	C14H18N4O3	DTXSID20152671	120075-57-2	6-Methoxy-4-(3-(N,N-dimethylamino)propylamino)-5,8-quin	C14H18N4O3	290.137890456	4
30	C14H18N4O3	DTXSID30213742	63931-79-3	1H-1,2,4-Benzotriazepine-3-carboxylic acid, 4,5-dihydro-4-	C14H18N4O3	290.137890456	3
31	C14H18N4O3	DTXSID30219608	69449-07-6	2,4-Pyrimidinediamine, 5-((3,4,5-trimethoxyphenyl)methyl)-	C14H20N4O4	308.14845514	3
32	C14H18N4O3	DTXSID20241155	94232-27-6	L-Aspartic acid, compound with 5-((3,4,5-trimethoxyphenyl	C18H25N5O7	423.175398165	3
33	C14H18N4O3	DTXSID80241156	94232-28-7	L-Glutamic acid, compound with 5-((3,4,5-trimethoxypheny	C19H27N5O7	437.191048229	3
34	C14H18N4O3	DTXSID20143781	101204-93-7	1H-Pyrido(2,3-e)-1,4-diazepine-2,3,5-trione, 4-(2-(diethylam	C14H18N4O3	290.137890456	3
35	C12H11N7	DTXSID6021373	396-01-0	Triamterene	C12H11N7	253.107593382	52
36	C12H11N7	DTXSID00204465	5587-93-9	Ampyrimine	C12H11N7	253.107593382	7
37	C12H11N7	DTXSID5064621	7300-26-7	Benzenamine, 4-azido-N-(4-azidophenyl)-	C12H9N7	251.091943318	4
38	C12H11N7	DTXSID00848025	90293-82-6	Sulfuric acid--6-phenylpteridine-2,4,7-triamine (1/1)	C12H13N7O4S	351.074973101	1
39	C12H11N7	DTXSID50575293	92310-83-3	(1E)-N-Phenyl-1,2-bis(1H-1,2,4-triazol-1-yl)ethan-1-imine	C12H11N7	253.107593382	1
40	C8H9NO2	DTXSID2020006	103-90-2	Acetaminophen	C8H9NO2	151.063328534	75
41	C8H9NO2	DTXSID6026567	134-20-3	Methyl 2-aminobenzoate	C8H9NO2	151.063328534	50

Downloadable Data



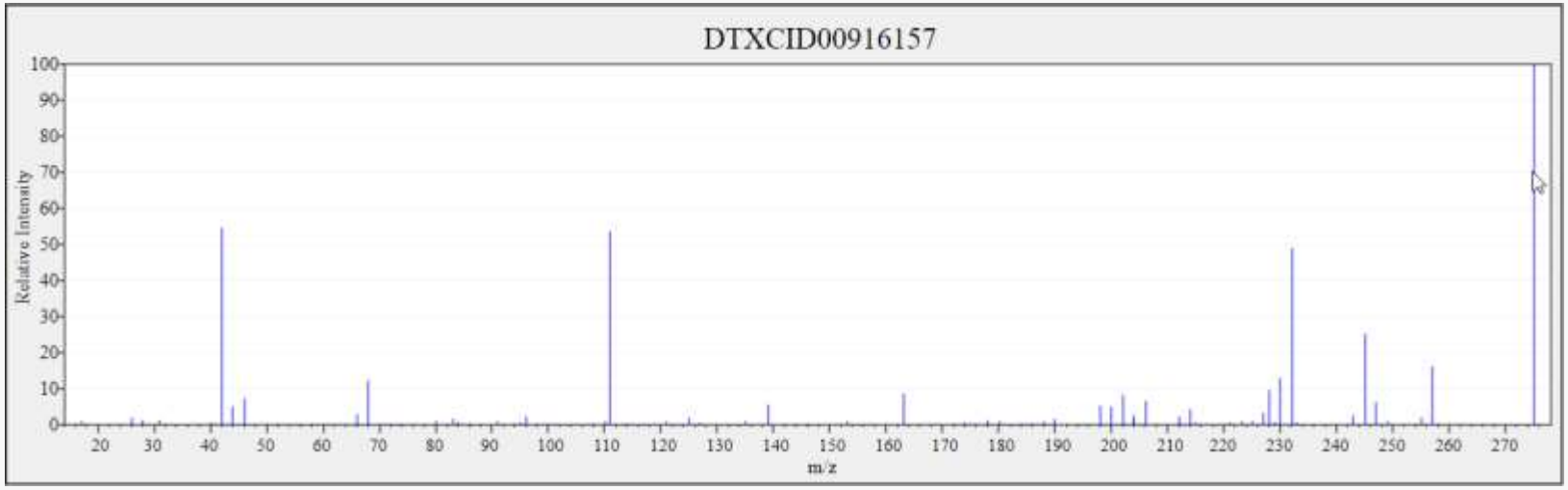
Work in Progress

- CFM-ID
  - Viewing and Downloading pre-predicted spectra
  - Search spectra against the database
- Retention Time Index Prediction
- Structure/substructure/similarity search
- Generation of MS-ready structures:
  - Upload file, download results
  - Service based generation

Predicted Mass Spectra

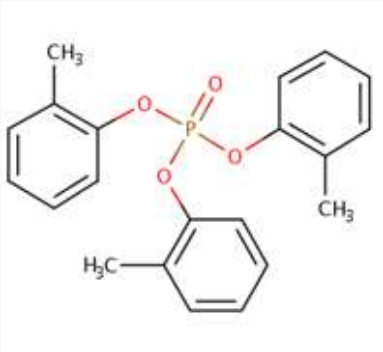
http://cfmid.wishartlab.com/

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

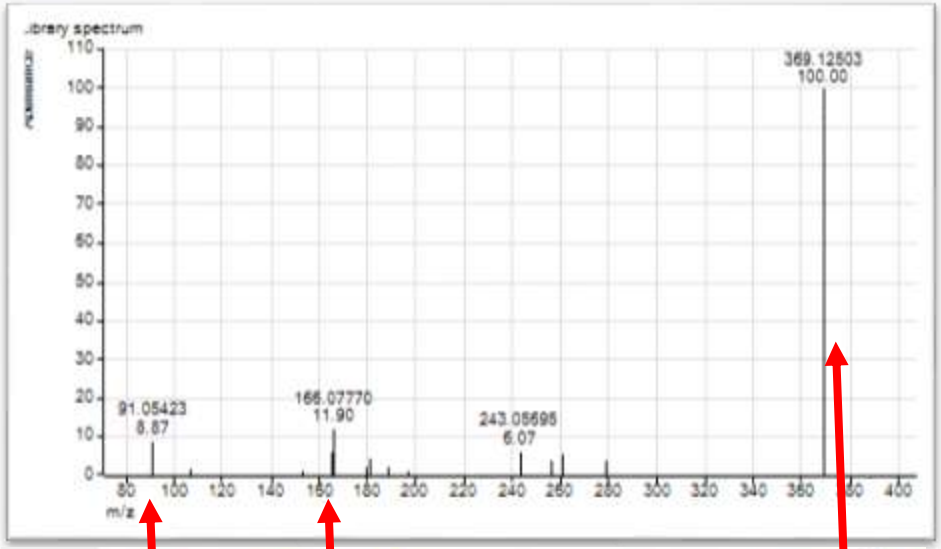


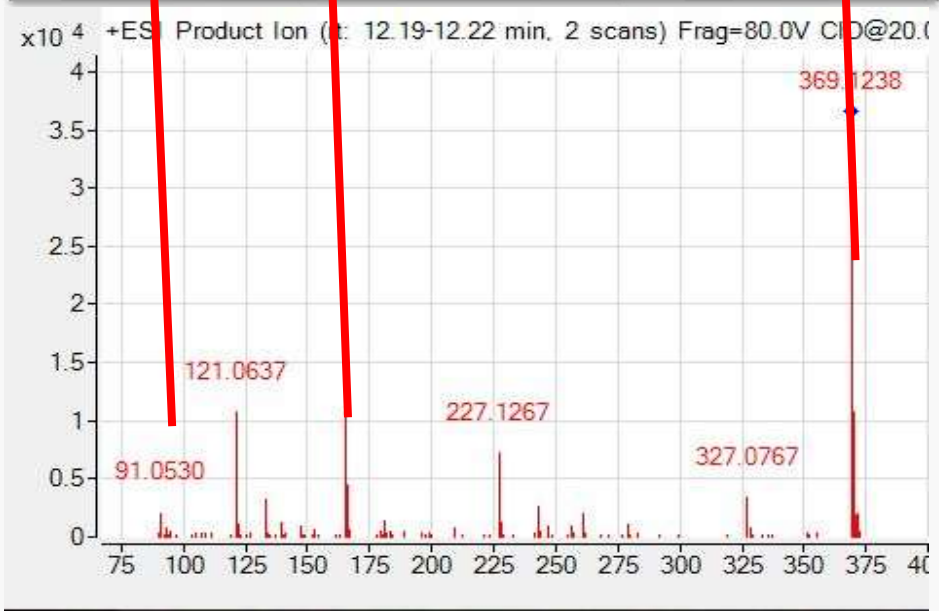
Predicted Mass Spectra

Library Fragmentation Spectra (20eV)



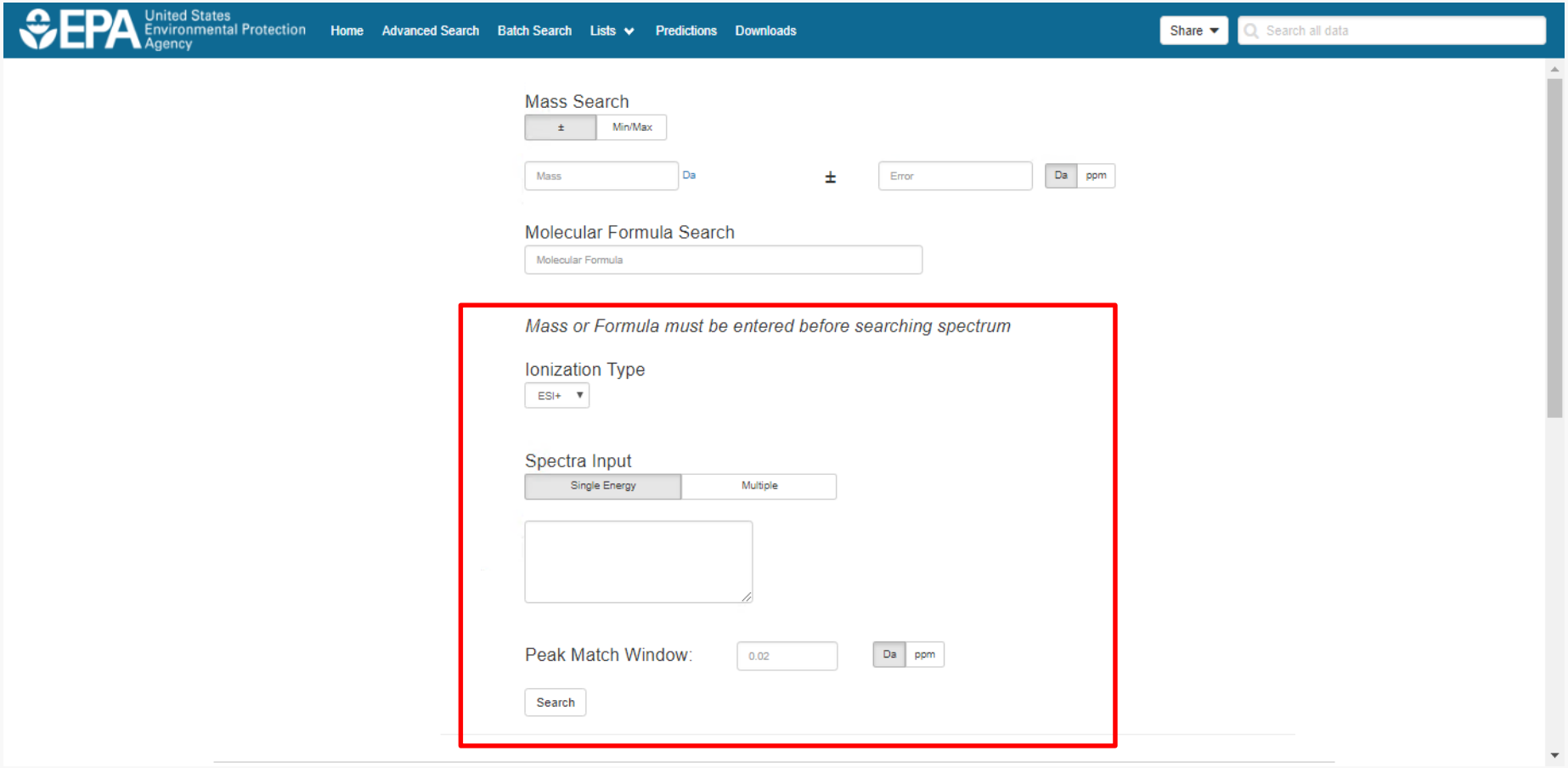
Observed Fragmentation Spectra (20eV)



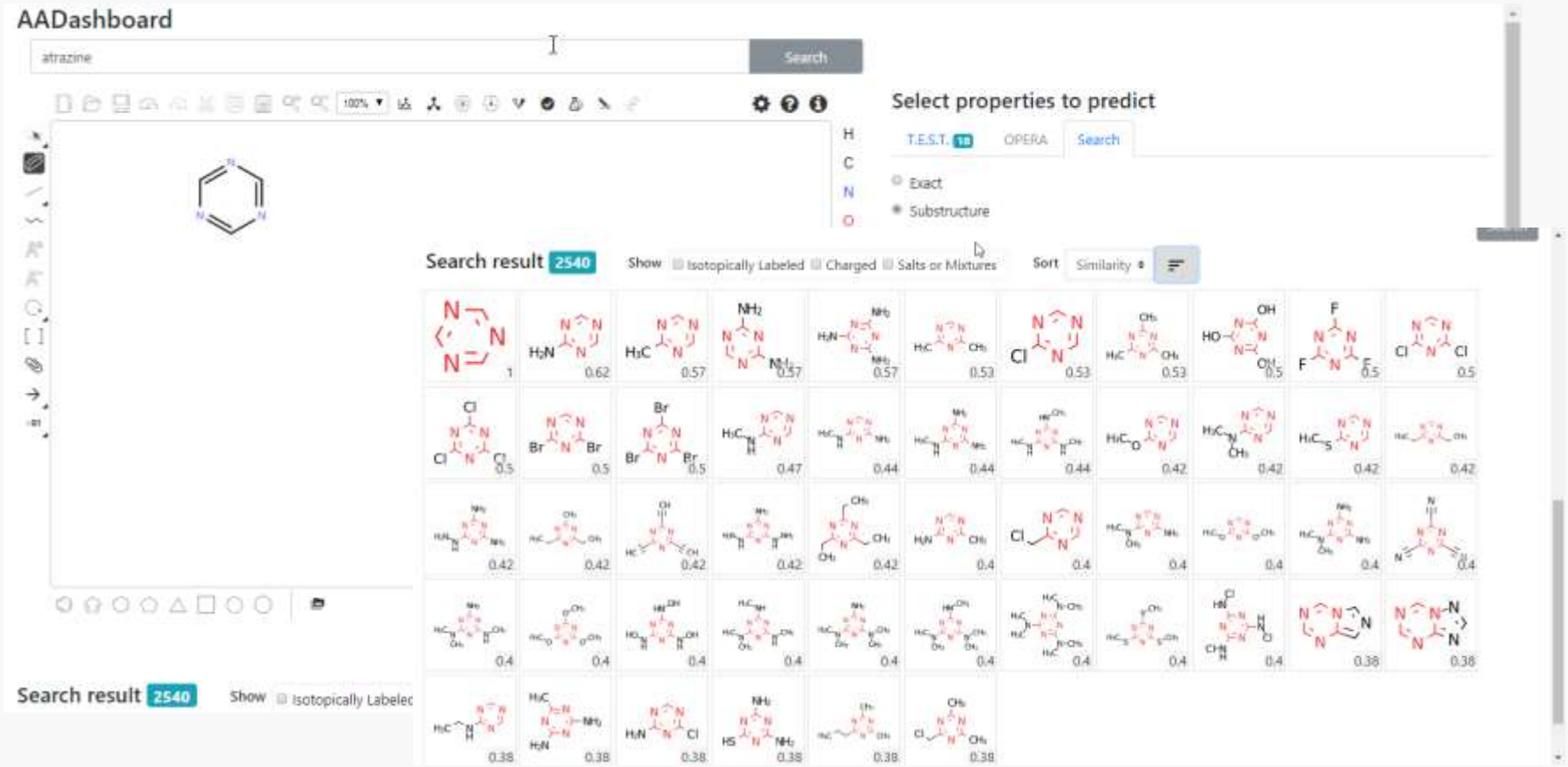


Match Score

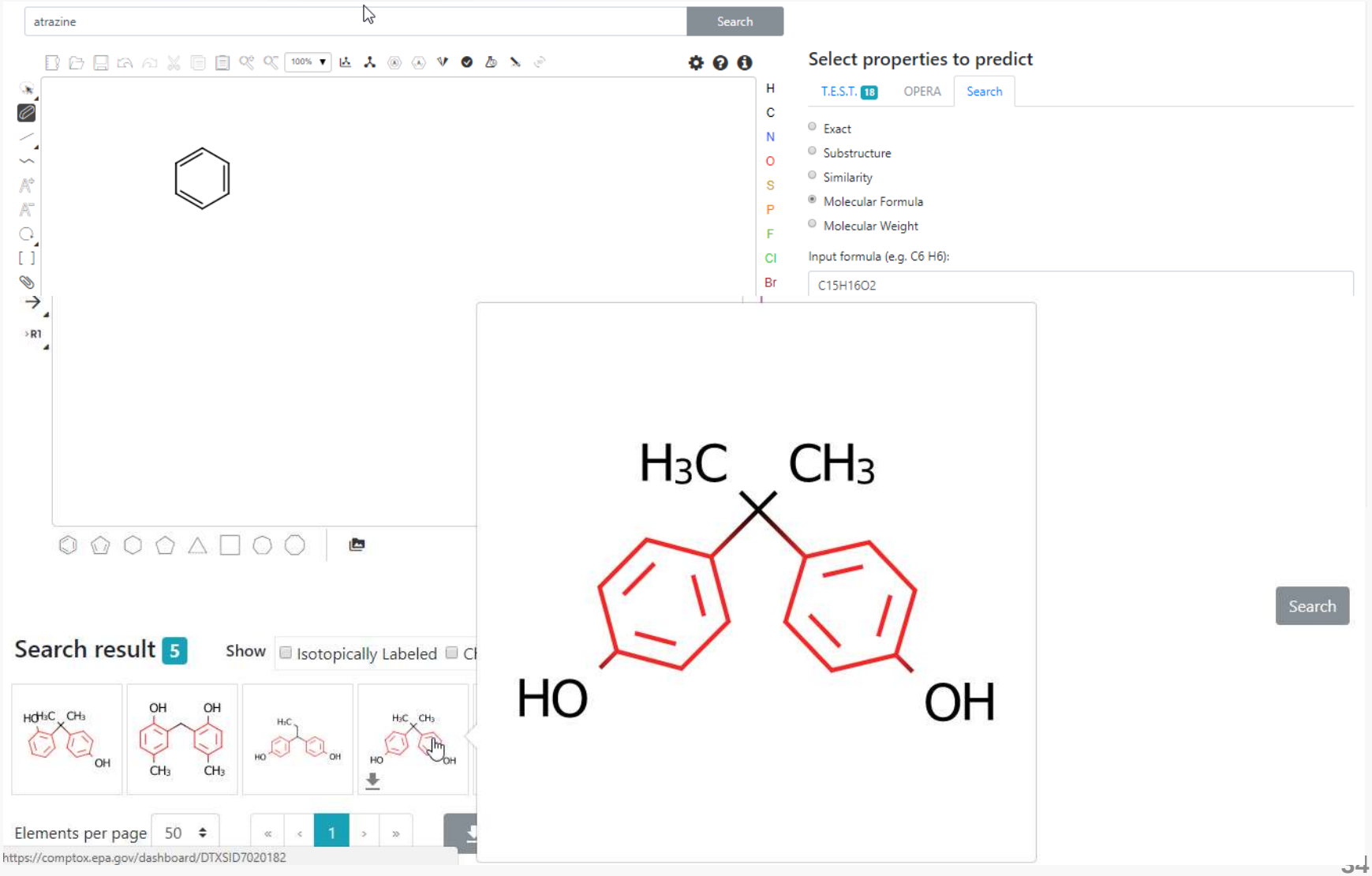
Search Expt. vs. Predicted Spectra



Prototype Development



Prototype Development



Conclusion

- The CompTox Chemistry Dashboard provides access to data for ~760,000 chemicals
- Multiple prediction models available for data gap filling
  - OPERA models and TEST models – PhysChem and Tox endpoints
  - Models based on *in vitro* data – classification models
  - Generalized Read-Across development in progress
- Real time prediction models rollout has started
- Web services available for some physchem and toxicity endpoints
- 2 years development as a CompTox Integration Hub

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