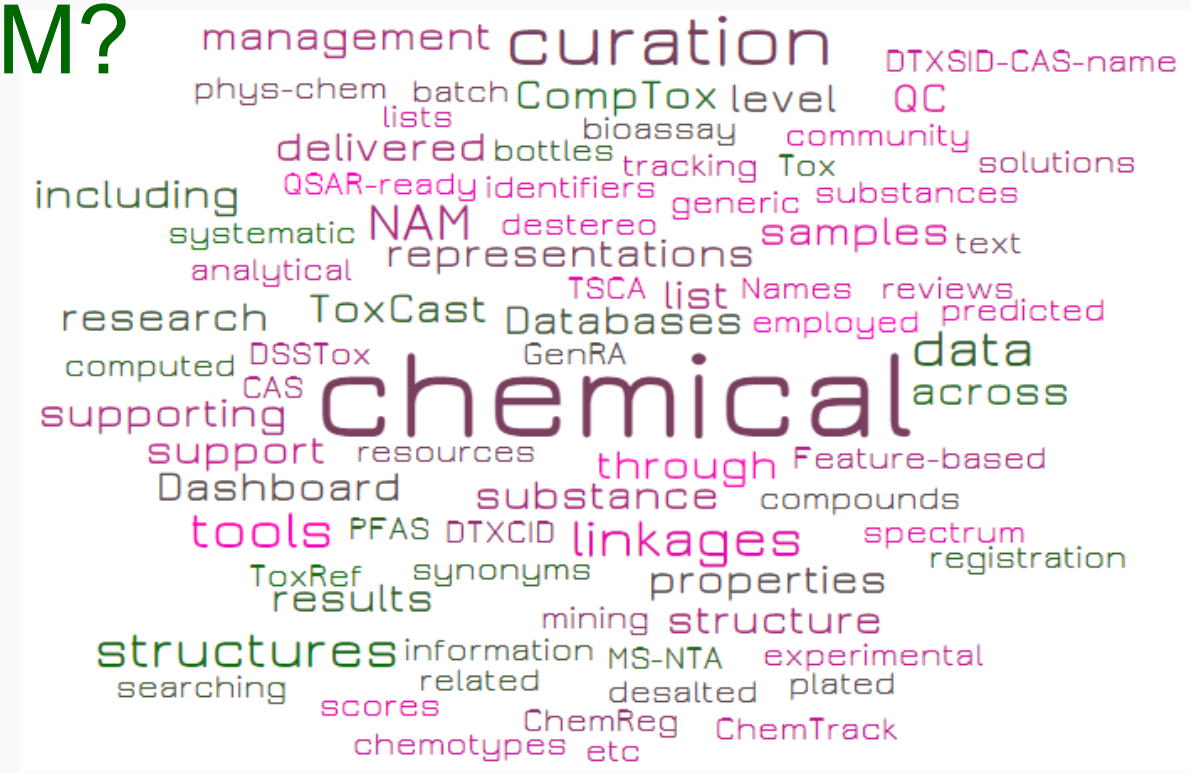
What's in a NAM?

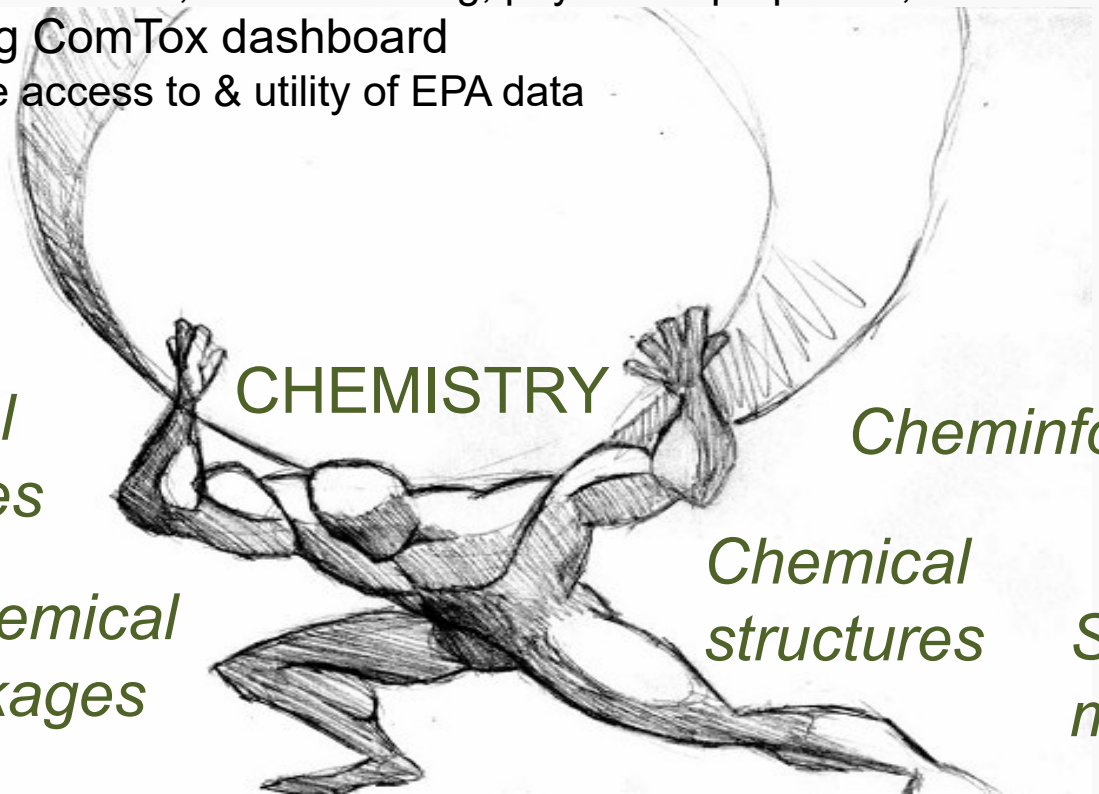
Why chemistry, of course!



Ann Richard
Computational Chemistry & Cheminformatics Branch
Chemical Characterization & Exposure Division
CCCB/ CCED/ CCTE

Challenge: Linking data resources

- ToxCast (EPA) & Tox21 (Multi-Agency)
  - screening >4000 (ToxCast) to >10K (Tox21) environmentally relevant chemicals across 10's to 100's of HTS assays
- ACToR, ExpoCast, CPCat, ToxRef DB, etc
  - meshed CAS lists, product-use database, *in vivo* reference DB
- *In Silico* models
  - exposure models, ER/AR binding, physchem properties, ADME
- Public-facing ComTox dashboard
  - facilitate access to & utility of EPA data

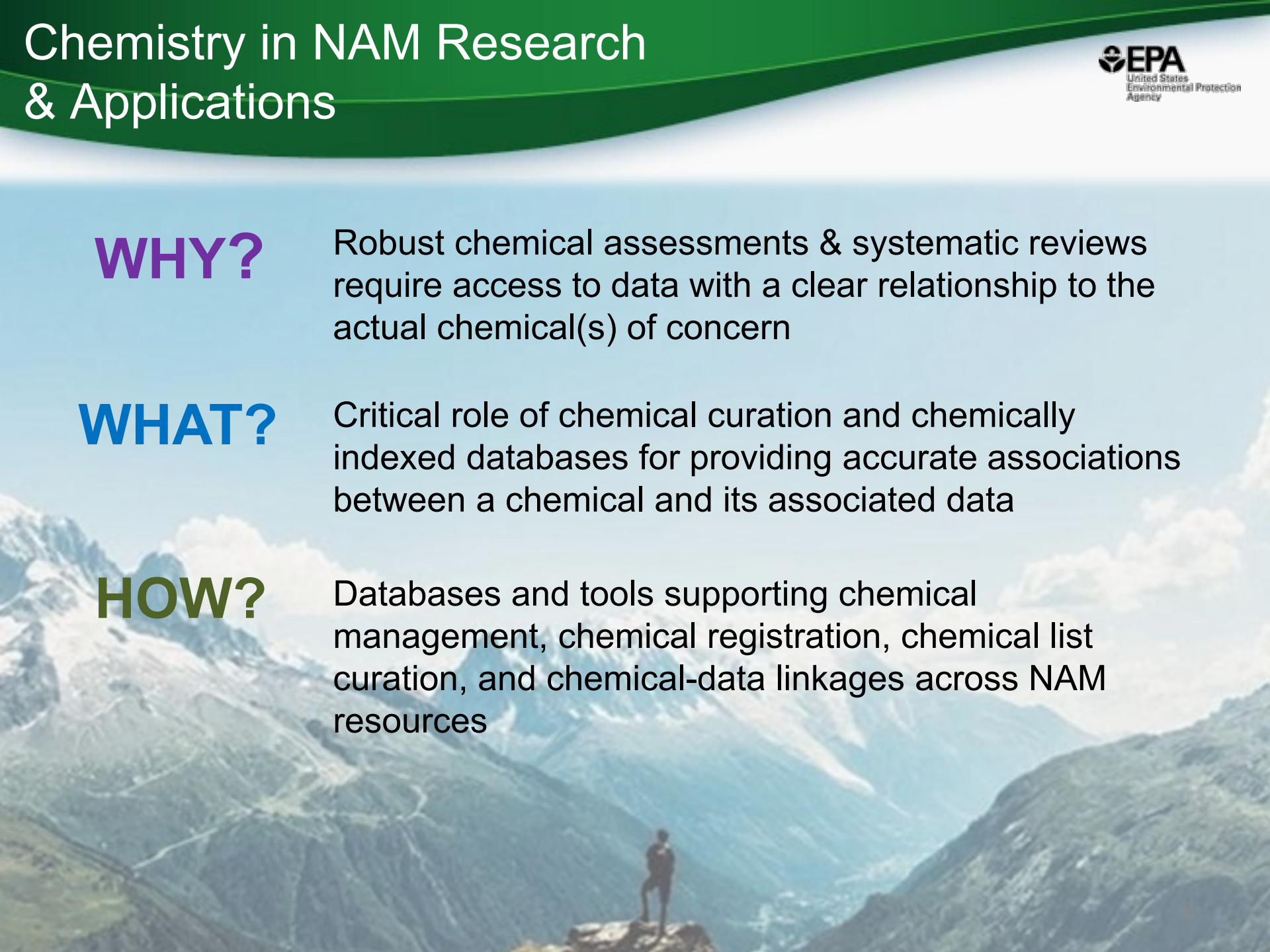


Chemistry in NAM Research & Applications

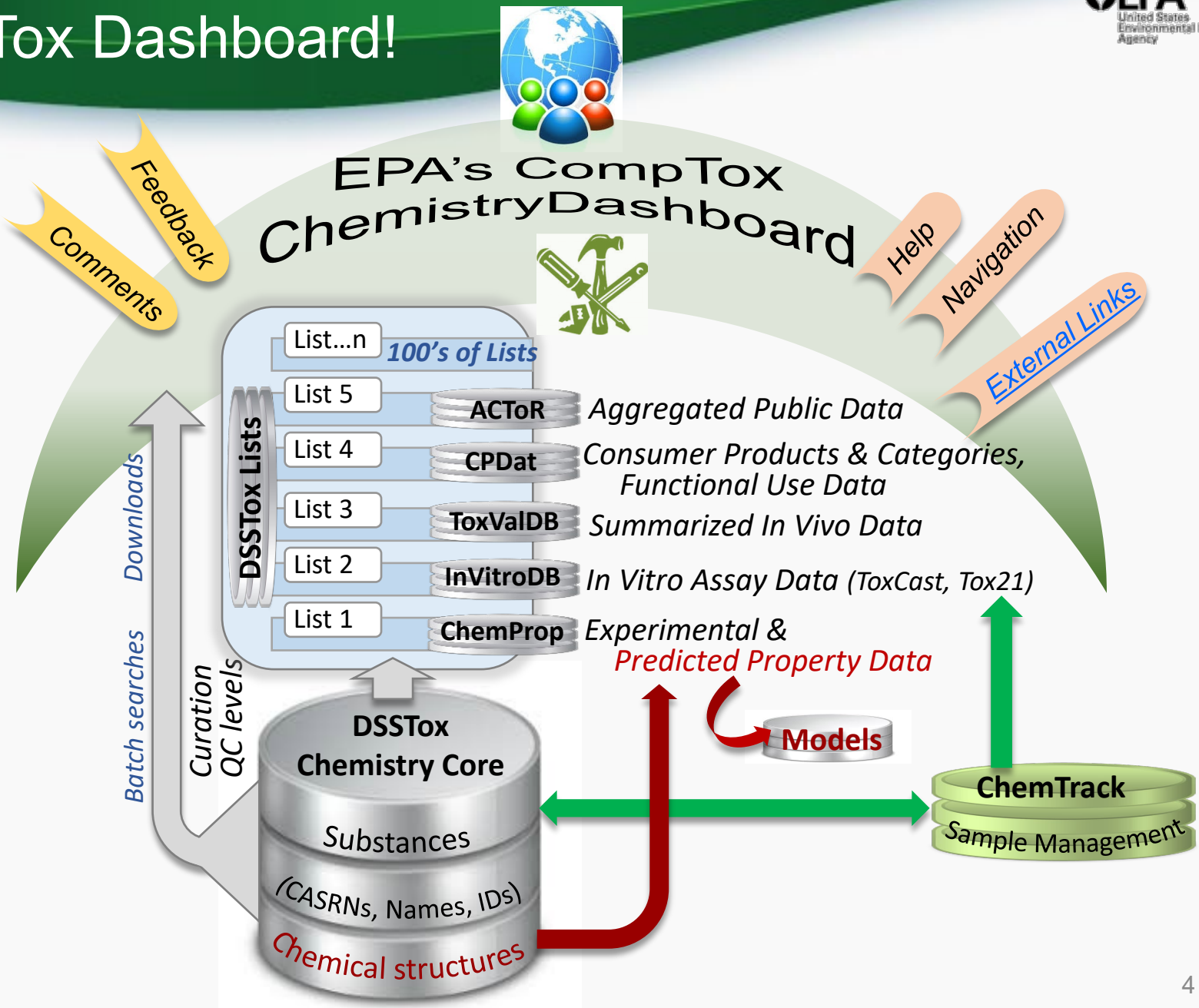
WHY? Robust chemical assessments & systematic reviews require access to data with a clear relationship to the actual chemical(s) of concern

WHAT? Critical role of chemical curation and chemically indexed databases for providing accurate associations between a chemical and its associated data

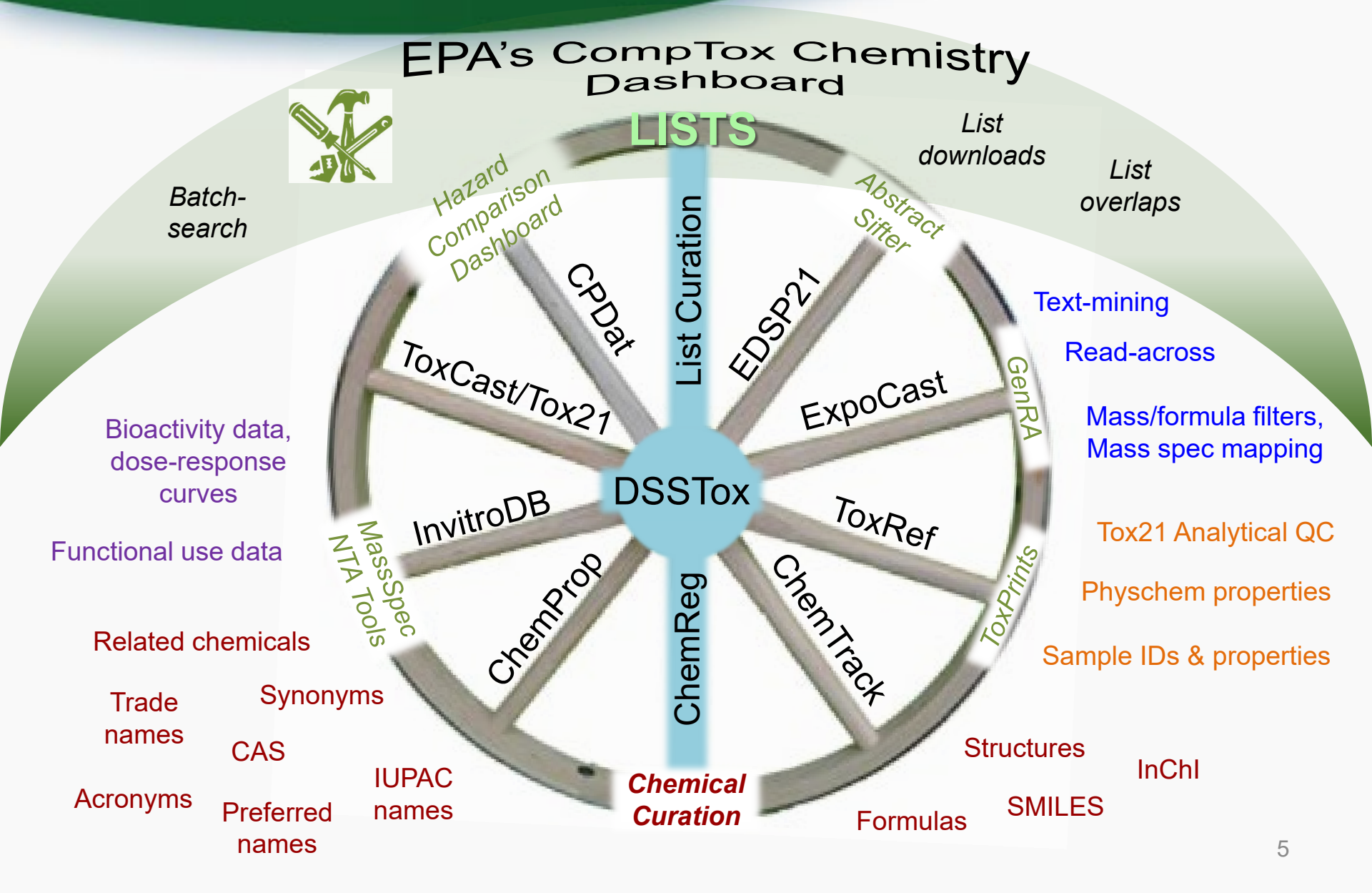
HOW? Databases and tools supporting chemical management, chemical registration, chemical list curation, and chemical-data linkages across NAM resources



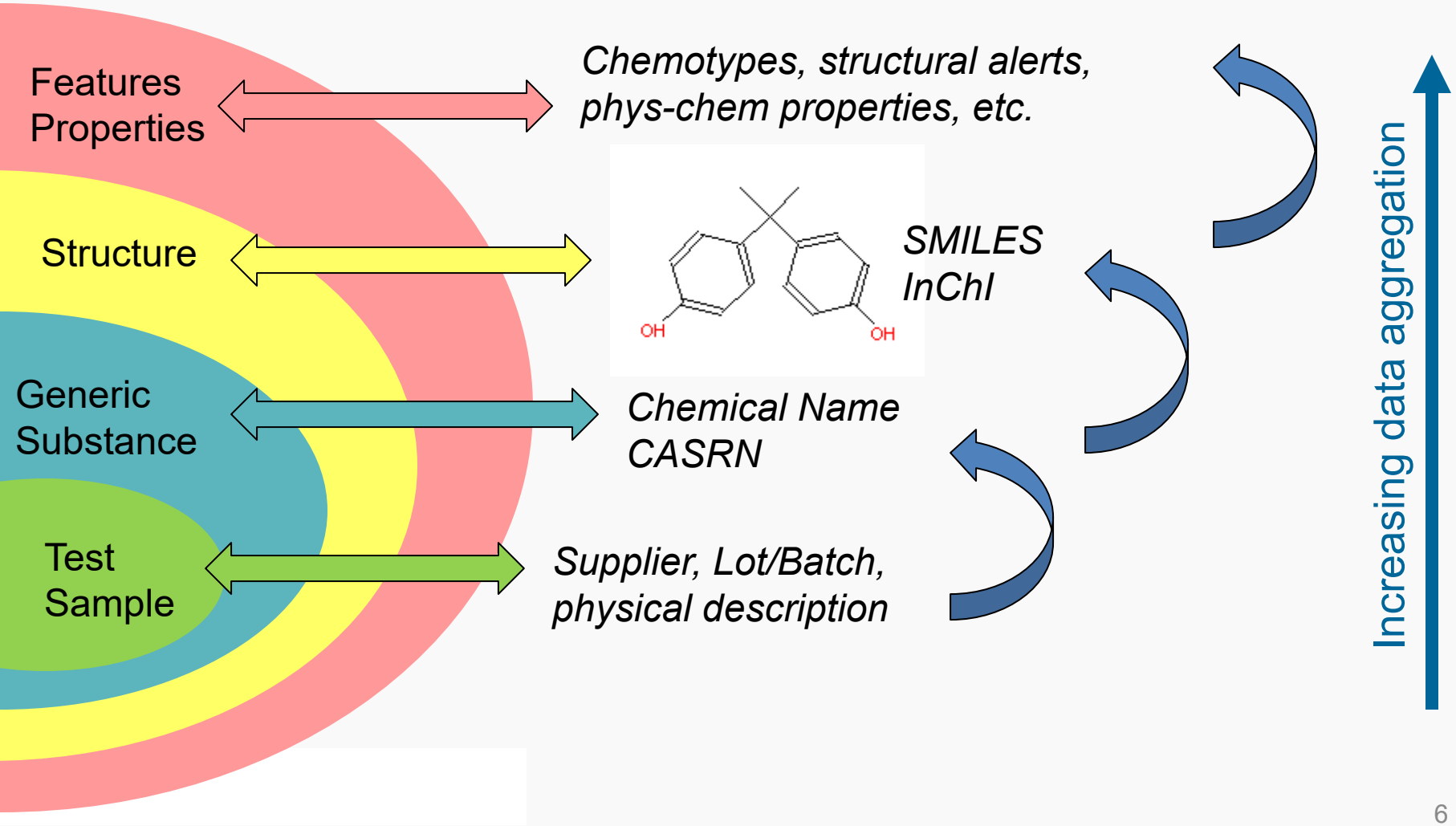
The Chemistry beast feeding the CompTox Dashboard!



DSSTox Chemical Hub



Chemical Elements to Data Integration:
Chemical representations



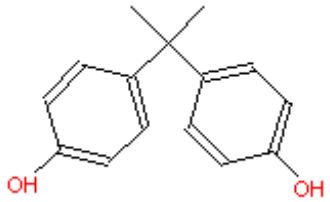
Chemical Elements to Data Integration:
Chemical representations

Features
Properties

Chemotypes, fingerprints, charges, phys-chem properties, etc.

Chemical analogs, Read-across, SAR modeling

Structure



SMILES
InChI

Structure searching & modeling

Generic Substance

DSSTox

Chemical Name
CASRN

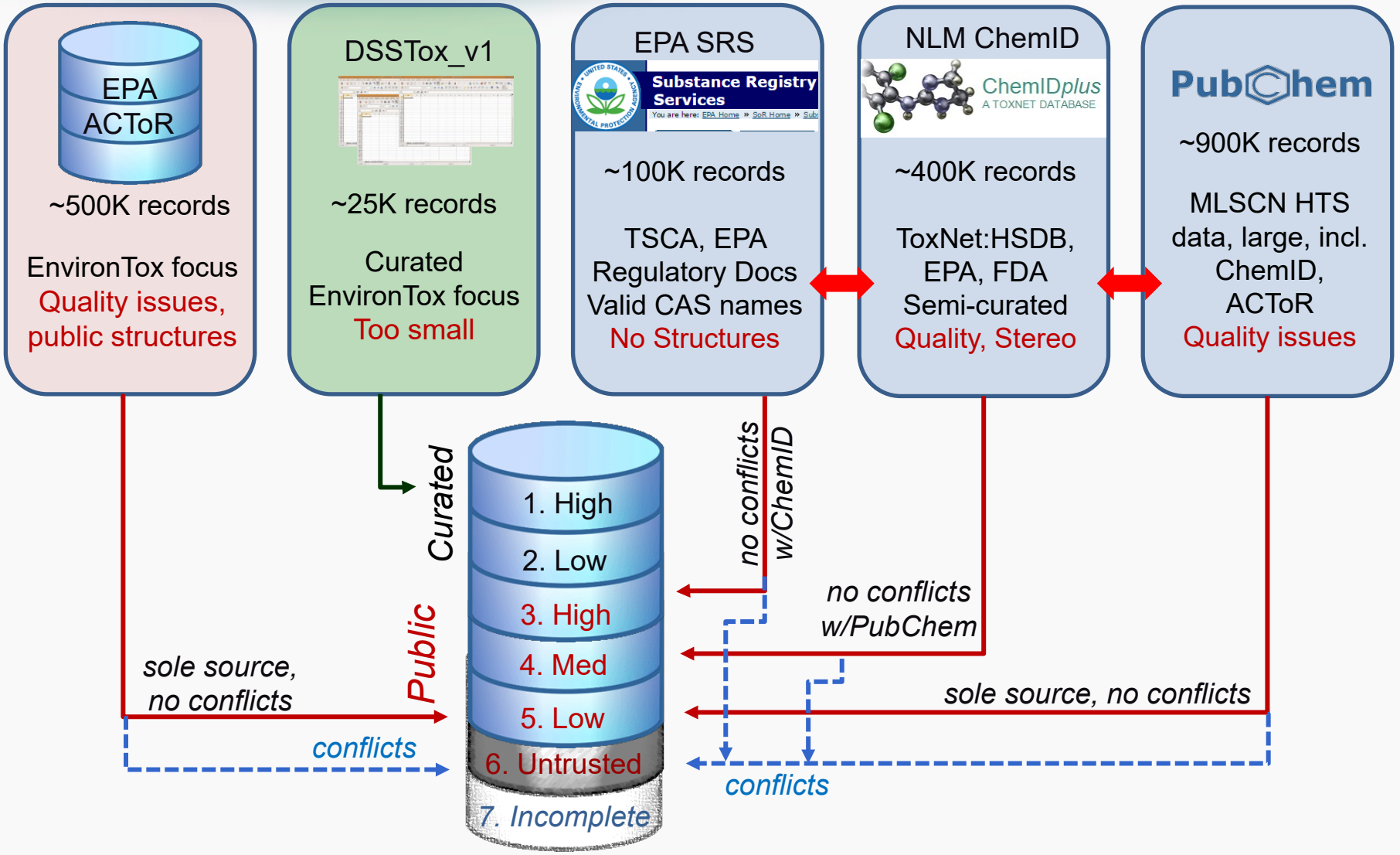
Public toxicity datasets

Test Sample

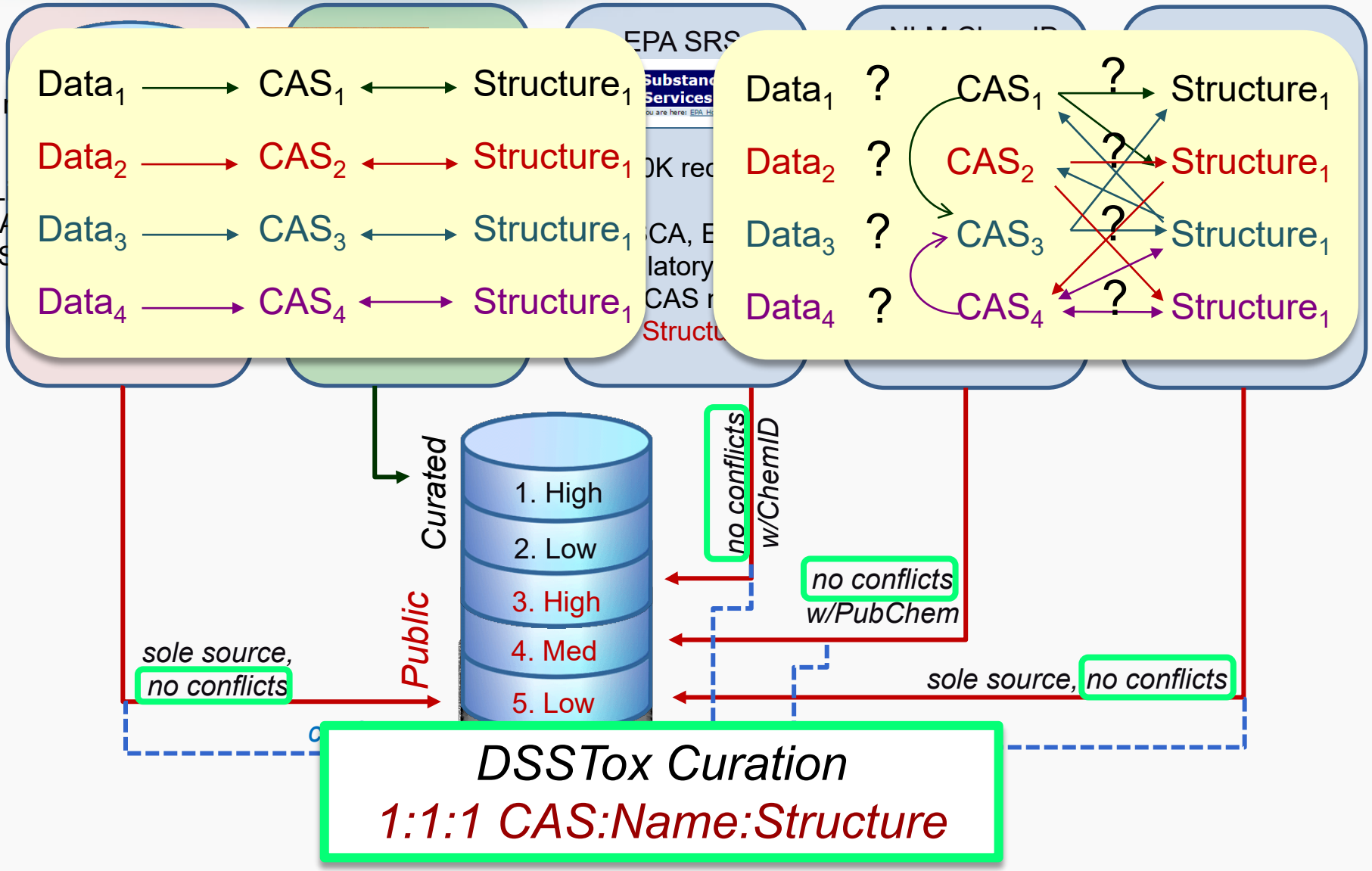
Supplier, Lot/Batch, physical description

Experimental Endpoint Data

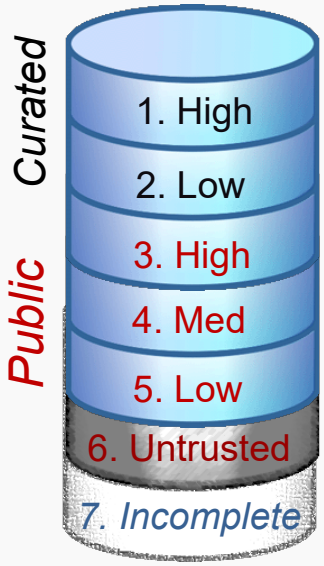
DSSTox_v2 Main Sources: Substance QC levels



DSSTox_v2 Main Sources: Substance QC levels



DSSTox_v2 QC Levels



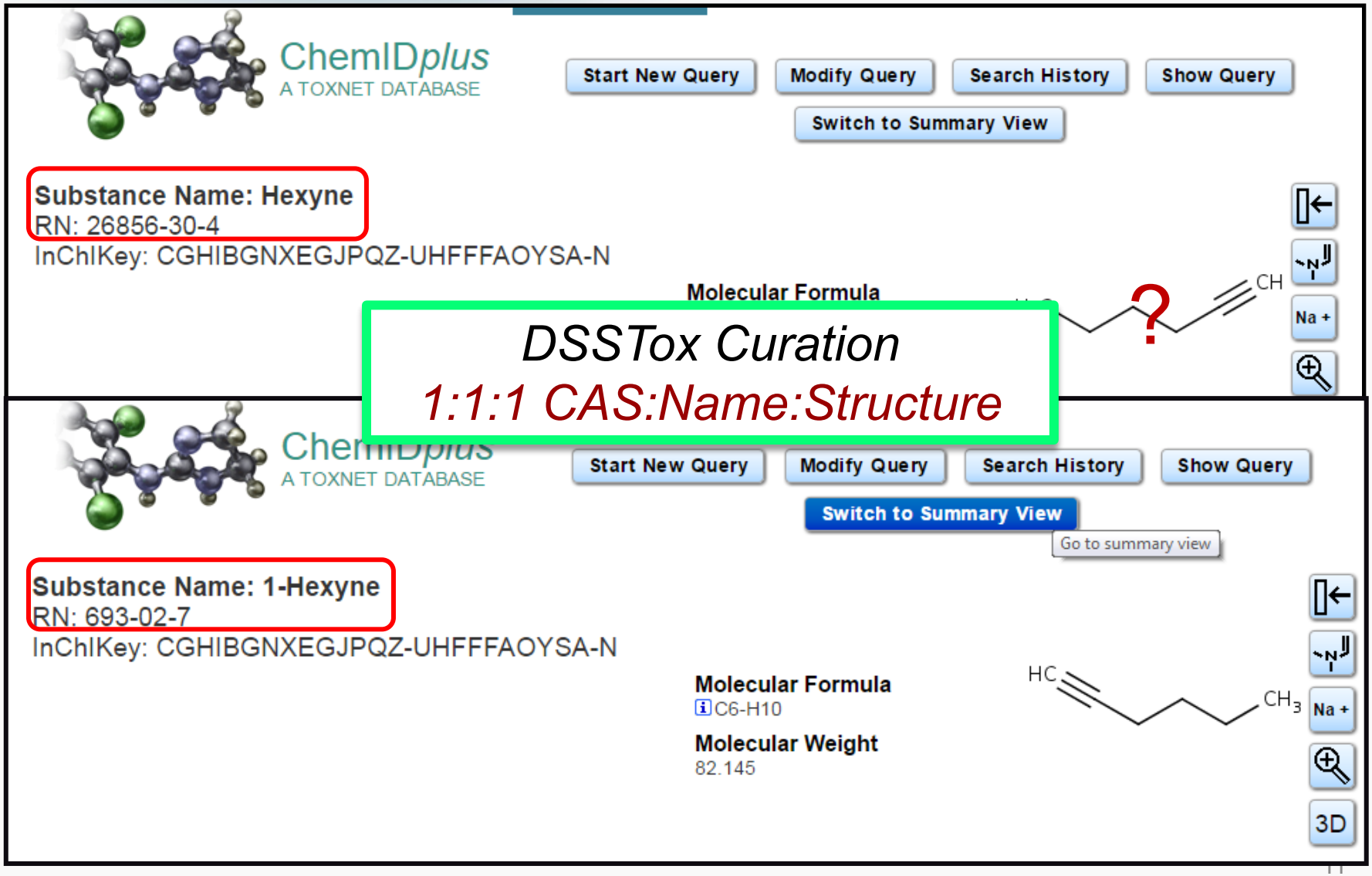
QC Level Totals (12Dec2019)

~ 200K substances in top 4 QC quality bins

QC Levels

DSSTox_High:	Hand curated - highest confidence from definitive source
DSSTox_Low:	Hand curated and confirmed using multiple public sources
Public_High:	Extracted from EPA SRS and confirmed to have no conflicts in ChemID and PubChem
Public_Medium:	Extracted from ChemID and confirmed to have no conflicts in PubChem
Public_Low:	Extracted from ACToR or PubChem (single source, no conflicts)
Public_Untrusted:	Postulated, but found to have conflicts in public sources

Substance-structure ambiguity



ChemReg: *Substance Registration*

EPA United States Environmental Protection Agency

ACToR-DSSTox Chemical Regis

View/Edit a Single Record | Structure Search | Browse/Curate Records

Ambiguous Synonym matched null
You are viewing the record associated with DTXSID3031864
CASRN: 1763-23-1

PFOS

Other Possible Records

CAS	Reason
DTXSID80108992	Ambiguous Synonym matched null

Systematic Nam
MolFormula:
MolWeight:
InChI Key:
Smiles:
DTXSID:

Calculate from

Substance_ID:
CAS:
Name:
Substance Type
QC Level:
Data Source:
QC Notes:

Synonyms (20) — Synonyms (20)

Other Cas (1) — Other CAS (1)

	CAS-RN	Relationship	Source	Comments
●	132324-11-9	Deleted	STN(DSSTox)	

Delete Selected | Add Other Cas

Successor Substances (8) — Successor Substances (8)

	CAS-RN	Relationship	Source	Struct	Casrn	Comments
●	45298-90-6	Component	STN(DSSTox)	✓	■	Sulfonate
●	2795-39-3	Salt Form	STN(DSSTox)	✓	■	Potassium Salt
●	29457-72-5	Salt Form	STN(DSSTox)	✓	■	Lithium Salt
●	56773-42-3	Salt Form	STN(DSSTox)	✓	■	Tetraethylammonium Salt
●	29081-56-9	Salt Form	STN(DSSTox)	✓	■	NH4+ salt
●	4021-47-0	Salt Form	STN(DSSTox)	✓	■	Sodium Salt
●	111873-33-7	Salt Form	STN(DSSTox)	✓	■	N,N,N-Tributylbutylammon
●	91036-71-4	Salt Form	STN(DSSTox)	✓	■	Magnesium salt

Delete Selected | Add Related Cas

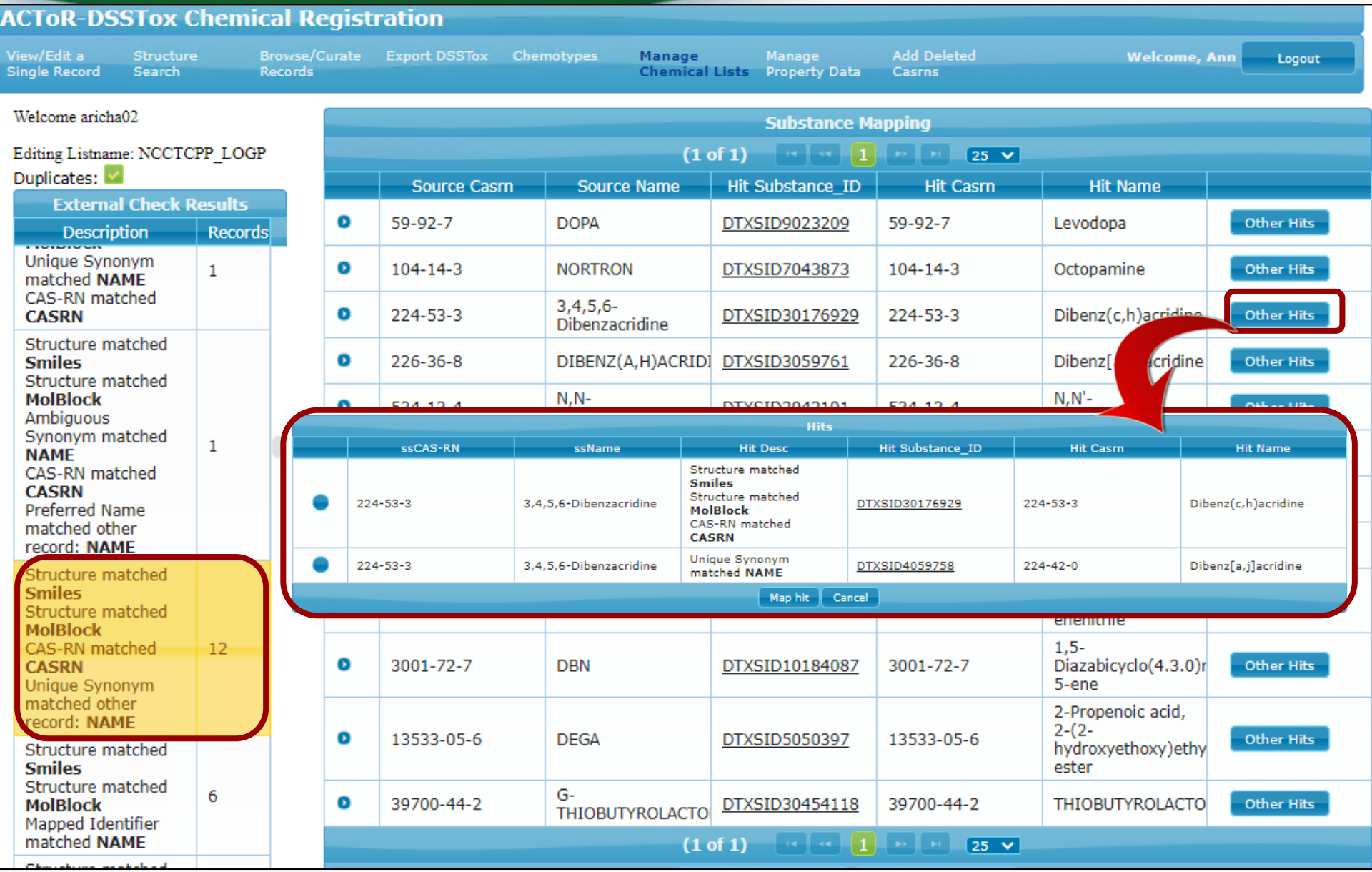
Predessessor Substances (2) — Predecessor Substances (2)

Relationship	DTXSID	CAS-RN	Source	Comments
Predecessor: Component	DTXSID20872963	NOCAS_872963	STN(DSSTox)	PFOS
Predecessor: Component	DTXSID40880545	64202-77-3	STN(DSSTox)	

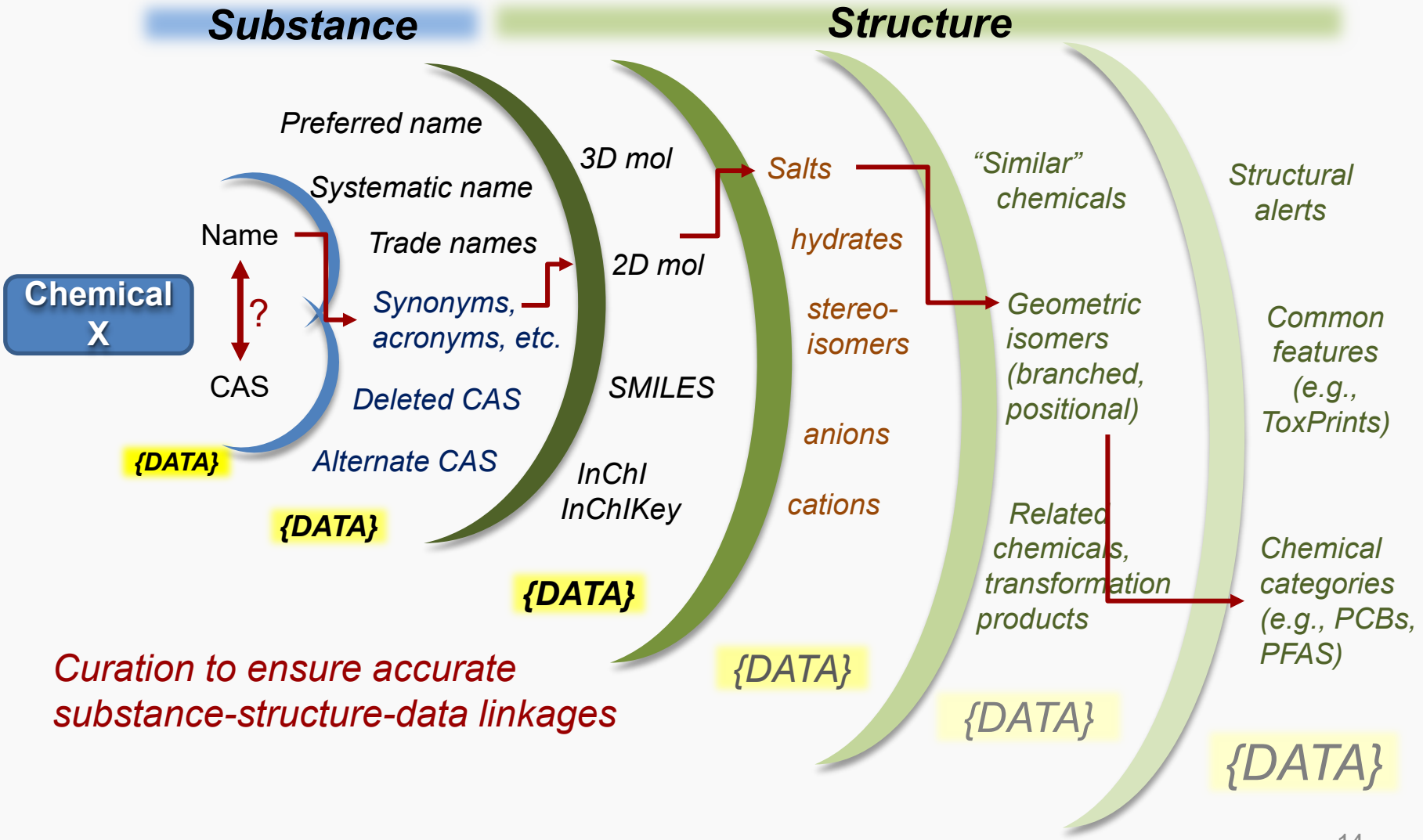
Associated Lists (290) — Associated Lists (290)

DTXRID	Source Identifier	Inventory Code
DTXRID8031865	Perfluorooctane Sulfonate, PFOS	NTPBSI
DTXRID0040753	Perfluorooctane sulfonic acid	TOXCAST

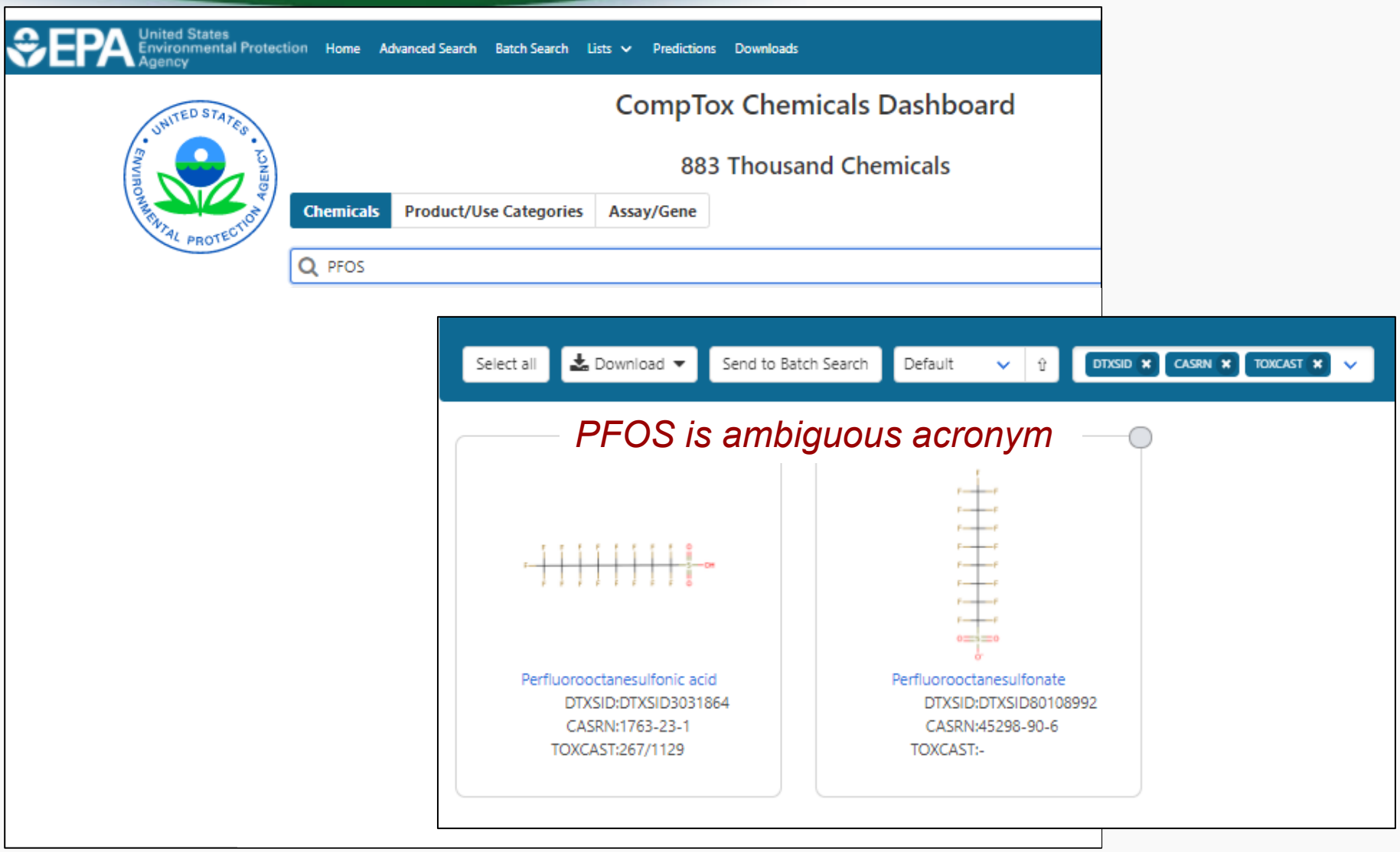
ChemReg: *List Curation*



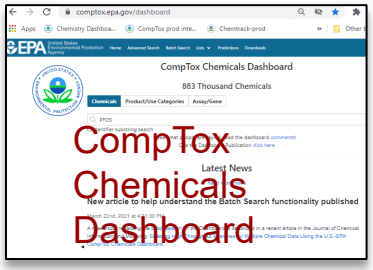
Chemical-data gathering for systematic reviews



Chemical-data gathering for systematic reviews: *e.g., PFOS*

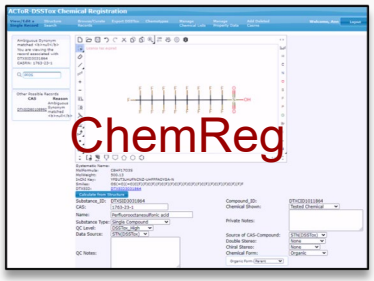


Chemical-data gathering for systematic reviews: *e.g., PFOS*

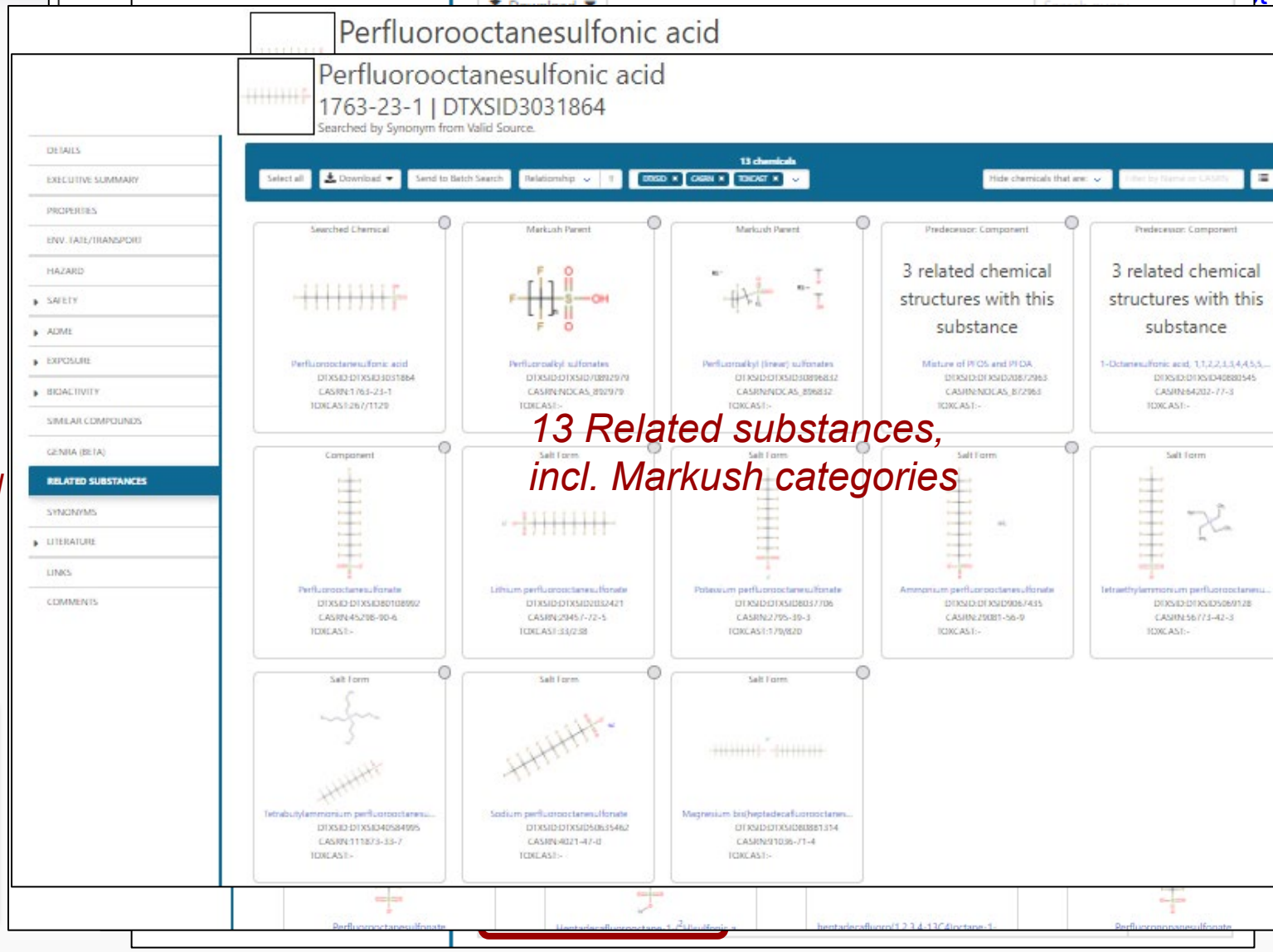


CompTox Chemicals Dashboard

Curation QC Level
Lists
Synonyms
Similar cmpds
Related cmpds



ChemReg



13 Related substances, incl. Markush categories

Curated Lists: e.g., PFAS

United States Environmental Protection Agency | Home

Curated substance

Download | Columns

List Acronym	List Name
PFASNORDIC	PFAS: Nordic PFAS
PFASNTREV19	PFAS: PFAS in Non 2019)
PFASOECD	PFAS: Listed in OE
PFASOECDNA	NORMAN: List of P Nikiforos Alygizaki
PFASSTRUCT	PFAS|EPA: PFAS str 2020)
PFASTRI	PFAS: PFAS to the Program by the Na
PFASTRIER	PFAS Community-
TONYPFASDASH	TONY'S Mockup o

PFAS|EPA: PFAS structures in DSSTox (update August 2020)

Search PFASSTRUCT Chemicals

Identifier substring search

PFAS structure file:

8163 substances with structures

List Details

Description: List consists of all DTXSID records with a structure assigned, and using a set of substructural filters based on community input. The substructural filters (visible here) are designed to be simple, reproducible and transparent, yet general enough to encompass the largest set of structures having sufficient levels of fluorination to potentially impart PFAS-type properties.

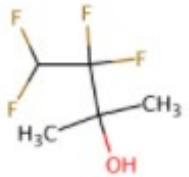
Number of Chemicals: 8163

Enflurane
DTXSID:DTXSID1020562
CASRN:13838-16-9
TOXCAST:0/235

1,1,1,2-Tetrafluoroethane
DTXSID:DTXSID1021324
CASRN:811-97-2
TOXCAST:-

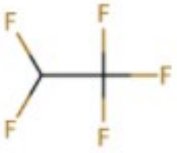
Tetrafluoroethylene
DTXSID:DTXSID6021325
CASRN:116-14-3
TOXCAST:-

2H-Perfluoro-2-propanol
DTXSID:DTXSID1022134
CASRN:920-66-1
TOXCAST:-

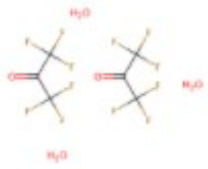


2-Methyl-3,3,4,4-tetrafluoro-2-butanol
DTXSID:DTXSID9022283
CASRN:29553-26-2
TOXCAST:-

Fulvestrant
DTXSID:DTXSID4022369
CASRN:129453-61-8
TOXCAST:157/924



Pentafluoroethane
DTXSID:DTXSID1024251
CASRN:354-33-6
TOXCAST:-



Hexafluoroacetone sesquihydrate
DTXSID:DTXSID7025392
CASRN:13098-39-0
TOXCAST:-

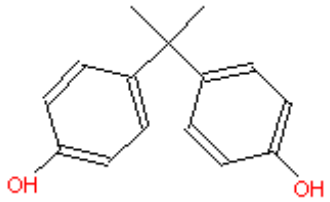
Chemical Elements to Data Integration:
Chemical representations

Features
Properties

Chemotypes, fingerprints, charges, phys-chem properties, etc.

Chemical analogs, Read-across, SAR modeling

Structure



SMILES
InChI

Structure searching & modeling

Generic Substance

DSSTox

Chemical Name
CASRN

Public toxicity datasets

Test Sample

Supplier, Lot/Batch, physical description

Experimental Endpoint Data

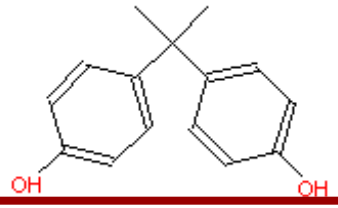
Chemical Elements to Data Integration:
Chemical representations

Features
Properties

Chemotypes, fingerprints, charges, phys-chem properties, etc.

Chemical analogs, Read-across, SAR modeling

Structure



SMILES
InChI

Structure searching & modeling

Generic Substance

DSSTox

Chemical Name
CASRN

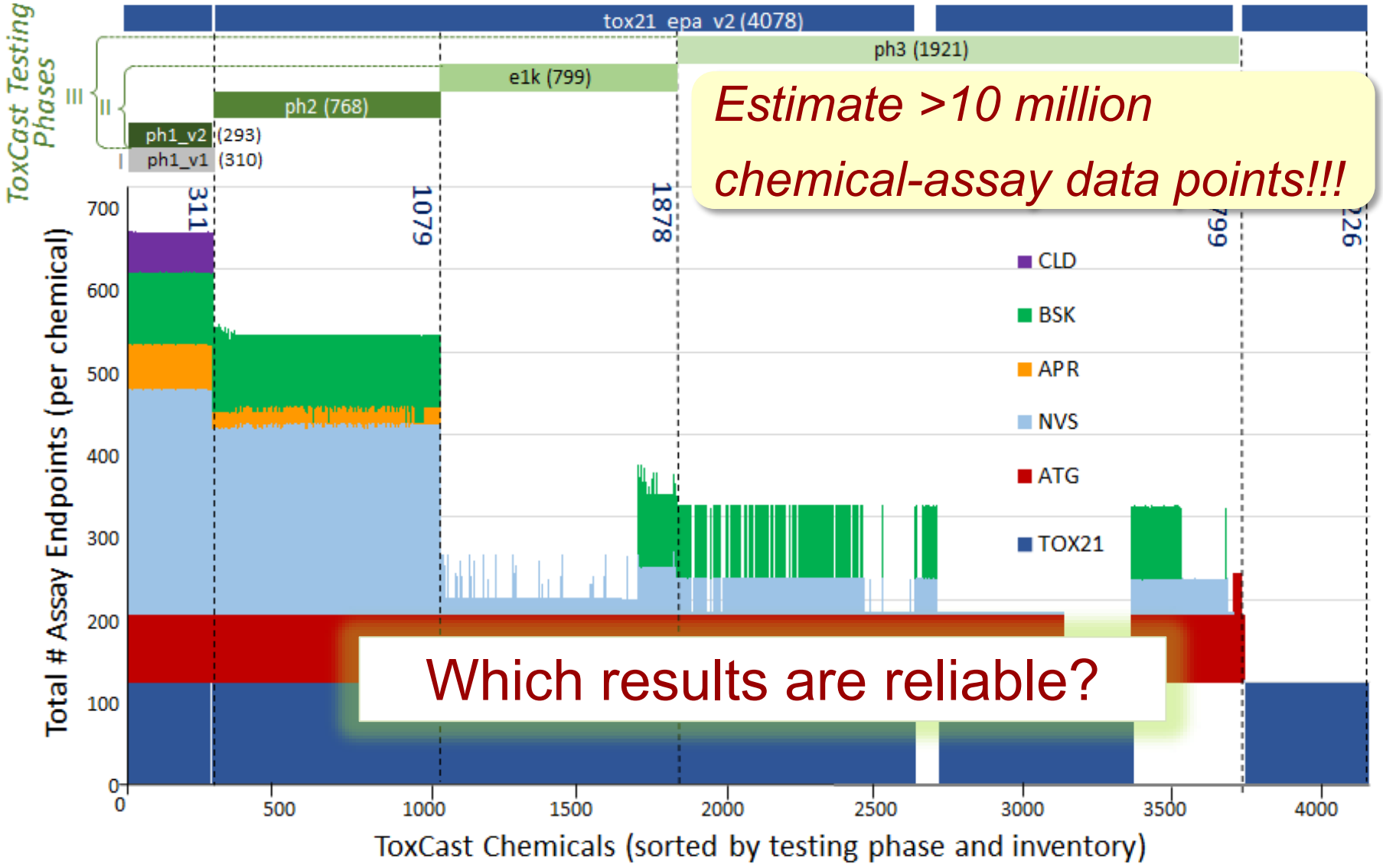
Public toxicity datasets

Test Sample

Supplier, Lot/Batch, physical description

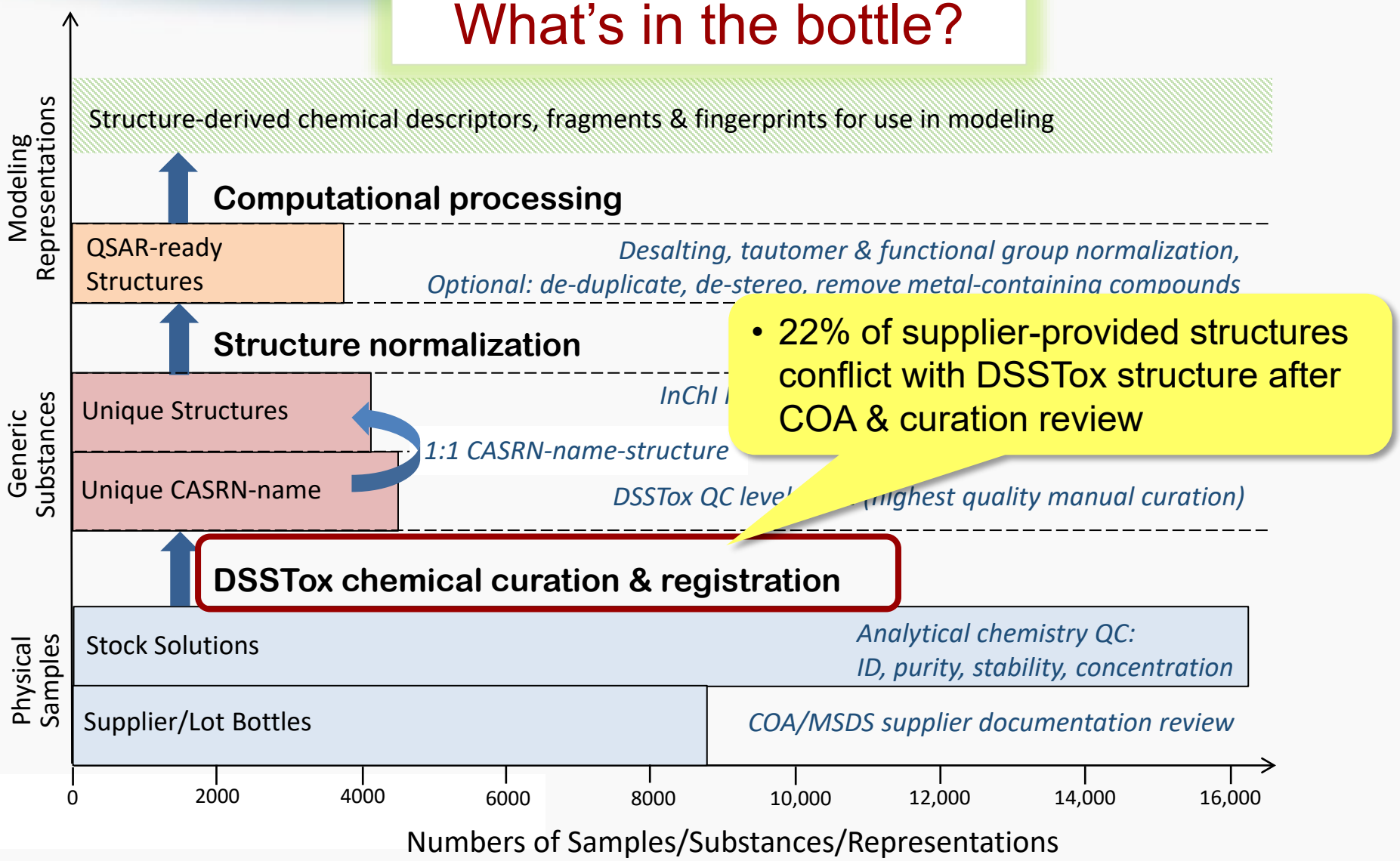
Experimental Endpoint Data

ToxCast/To21 HTS data



Chemical & Data Quality Issues

ToxCast Chemical Library: Quality Control Steps



ChemTrack:
Chemical Sample Management

- *Curators manually review COAs & MSDSs and map each original supplier bottle to a DTXSID*
- *Neat and solution tracking for >53K bottle codes!*

chemtrack Inventory Partnerships Ordering

Bottle Inventory

Uploaded MOSAIC Files 48 | Curated Bottles 53607 | Uncurated Bottles 3360

Columns 10 Export

Barcode	CoA Summary Id	Barcode type	DTXSID	Compound name	Casrn	Qty available	Concentration (mM)	Vendor
Filter by Barc	Filter b	Filter by Barcode ty	Filter by DTXSI	Filter by Compou	Filte	Filter	Filter by C	Filter
0073019474	7083	EPA_Vial_MLR96_(1.4mL)	DTXSID2047648	3-PHENYLPROPYL ACETATE	122-72-5	1000ul	20	Sigma Chemical Company
0073019475	7065	EPA_Vial_MLR96_(1.4mL)	DTXSID6047604	PHENETHYL BUTYRATE	103-52-6	955ul	20	Sigma Chemical Company
0073019476	7089	EPA_Vial_MLR96_(1.4mL)	DTXSID5021621		116-53-0	1000ul	20	Sigma Chemical Company

What can analytical QC data tell us?

Sample-specific

- purity
- age of sample
- supplier variability
- plating errors

Compound - specific

- volatility
- reactivity
- limited solubility in DMSO
- adheres to plastic

Tox21

Identity
- Wrong mass
- Different chemical than expected

Tox21

Purity
- Positive mass ID
- Contaminants detected

Tox21

Concentration
- Positive mass ID
- Less than expected (<30%)

Tox21

Stability
- T0 Pass → T4 Fail
- Depends on age of sample

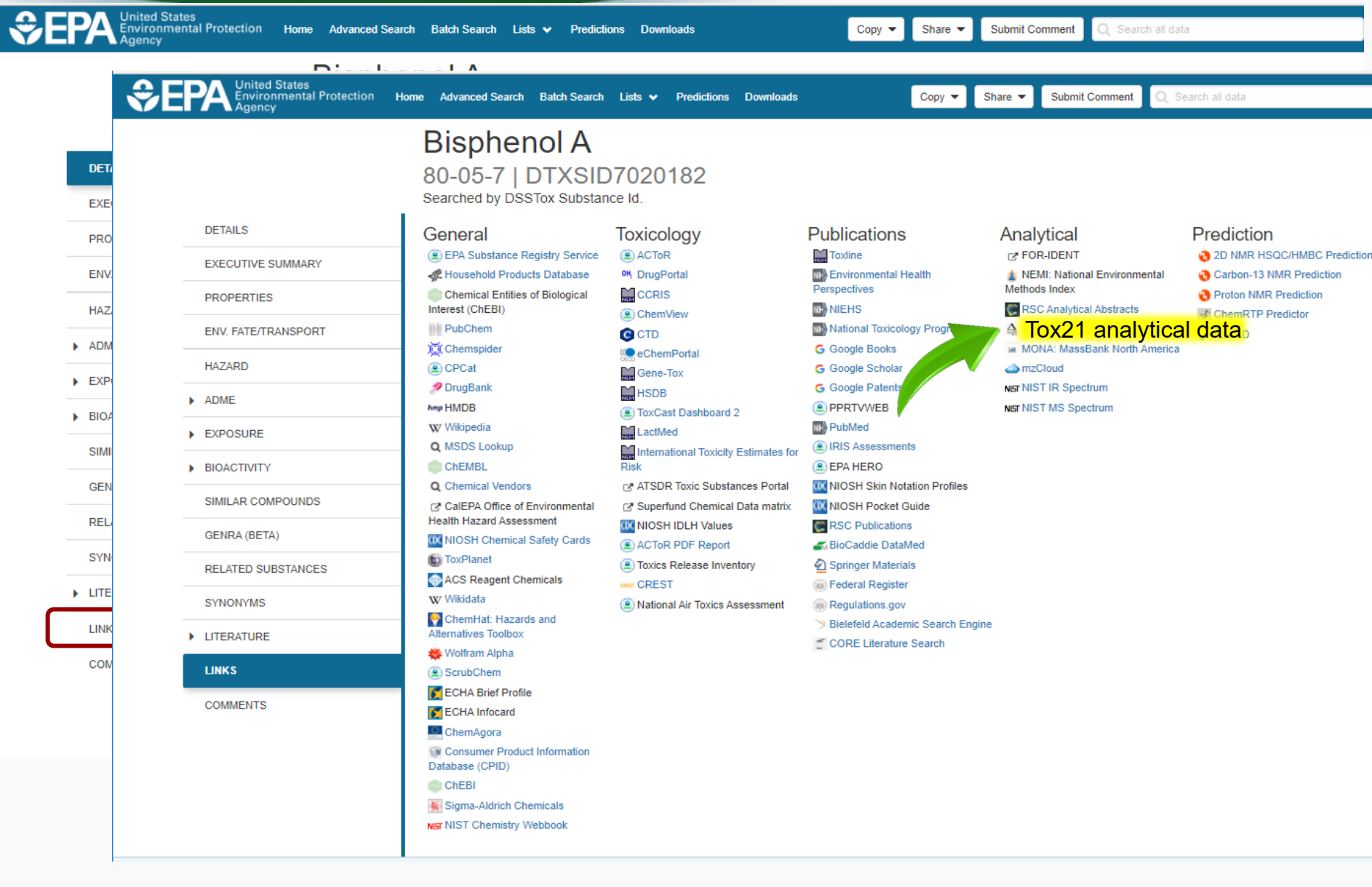
Tox21

Mixture/Isomers
- Positive mass ID
- Contaminants detected

Impact on HTS results

- Is bioactivity due to cmpd or contaminant?
- Is lack of activity due to cmpd not present or present at low concentration?

Some cmpds have QC data …



At the Tox21 ID soln level…

EPA United States Environmental Protection Agency

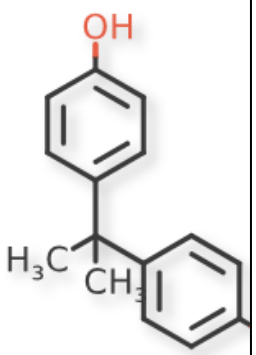
Secure | https://tripod.nih.gov/tox21/samples/80-05-7

Apps | Chemistry Dashboard | CompTox (zn) Dashb | ChemReg_v0.9.6 | ChemTrack | Confluence - NCCT K | JIRA | Other bookmark

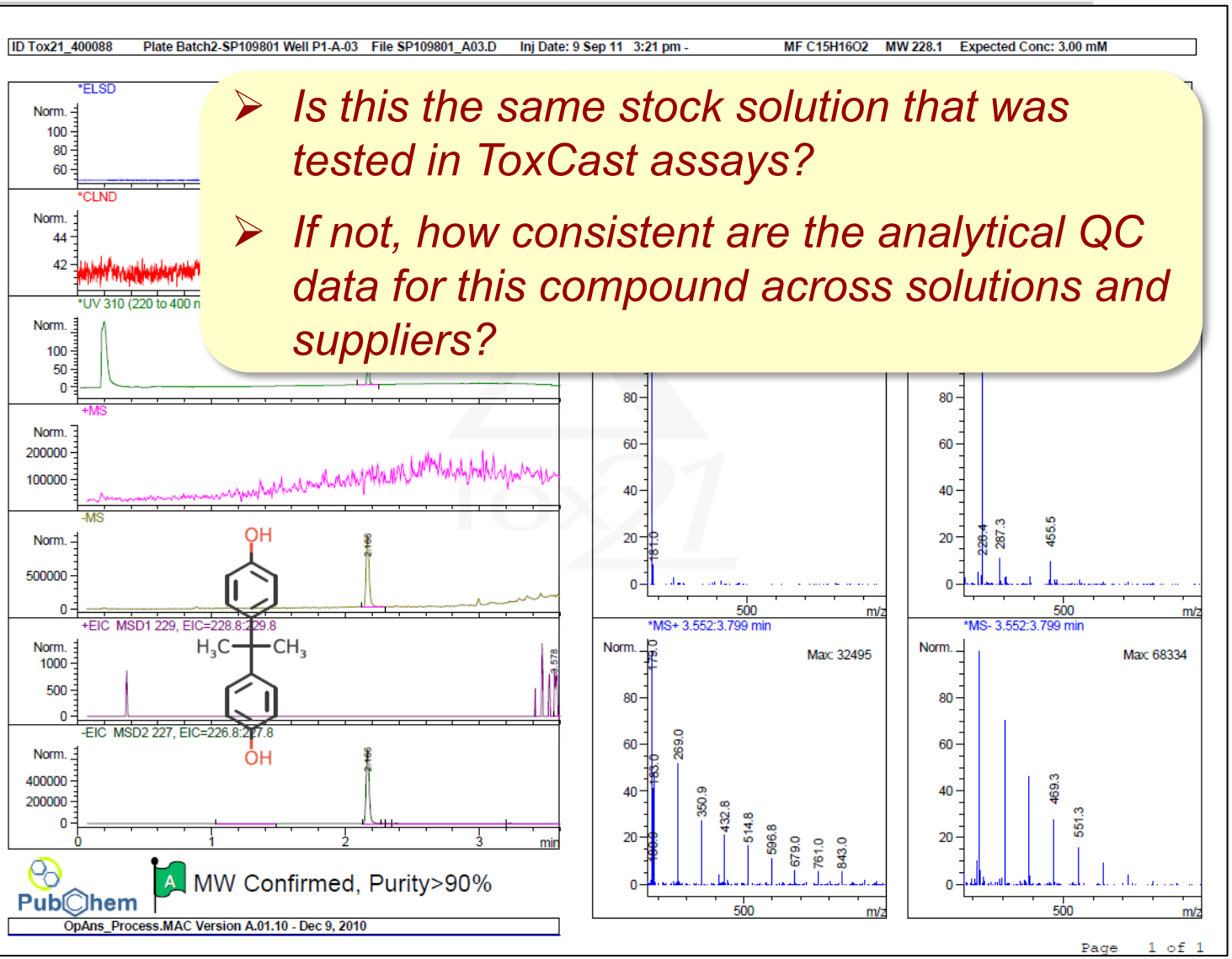
Tox21

Home / Tox21 Samples / Tox2

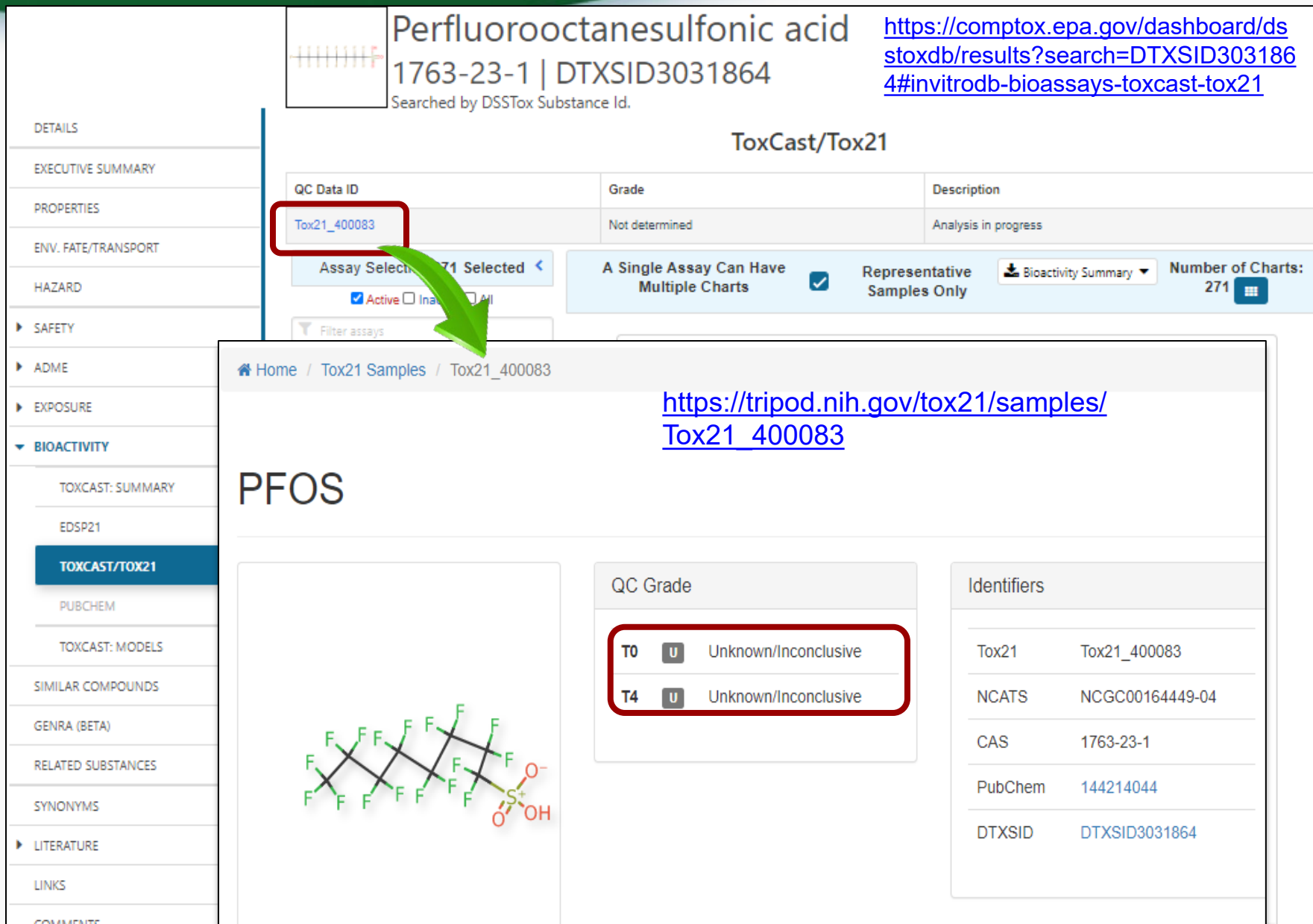
Bisphenol A



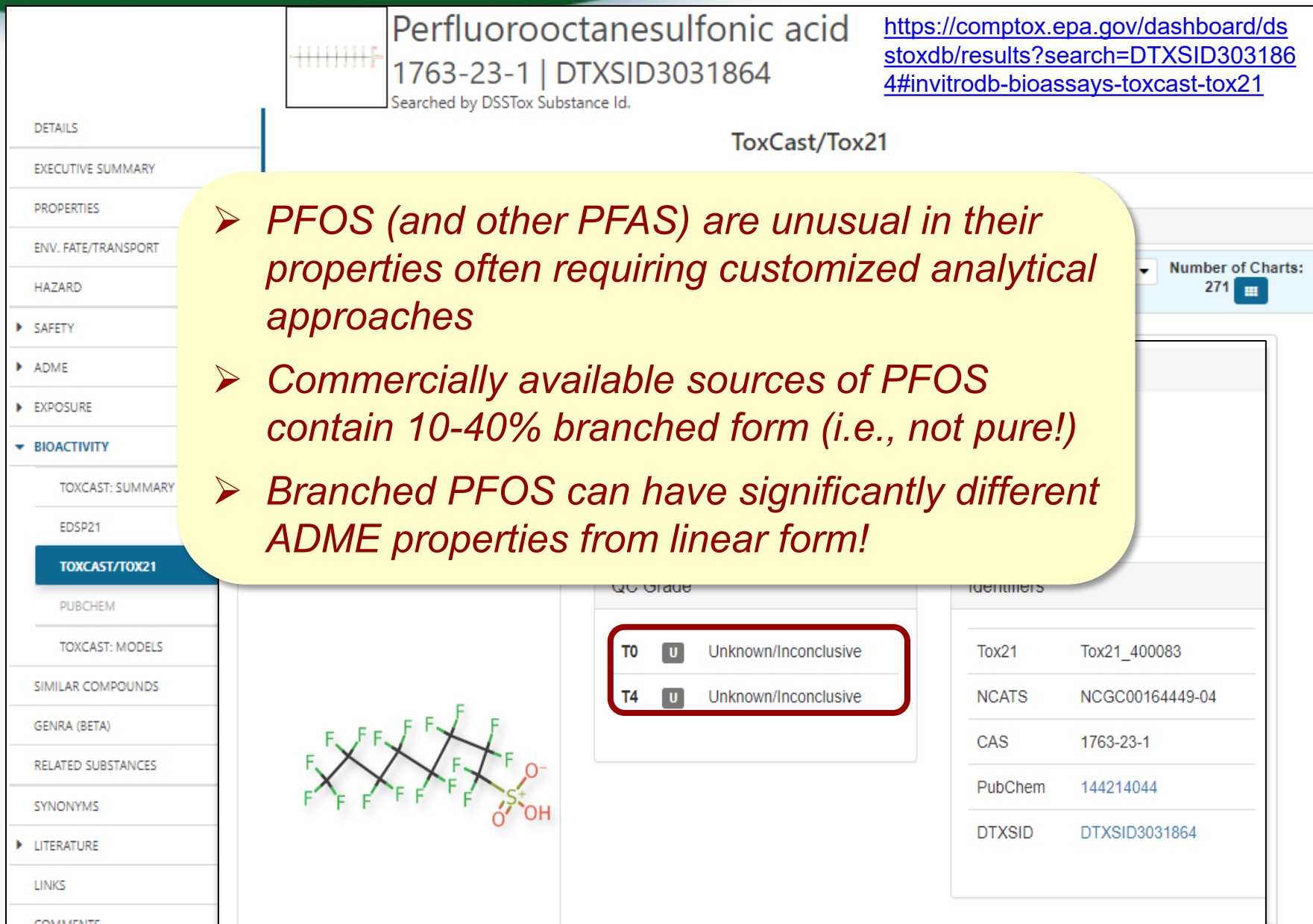
- *Is this the same stock solution that was tested in ToxCast assays?*
- *If not, how consistent are the analytical QC data for this compound across solutions and suppliers?*



Some don't …

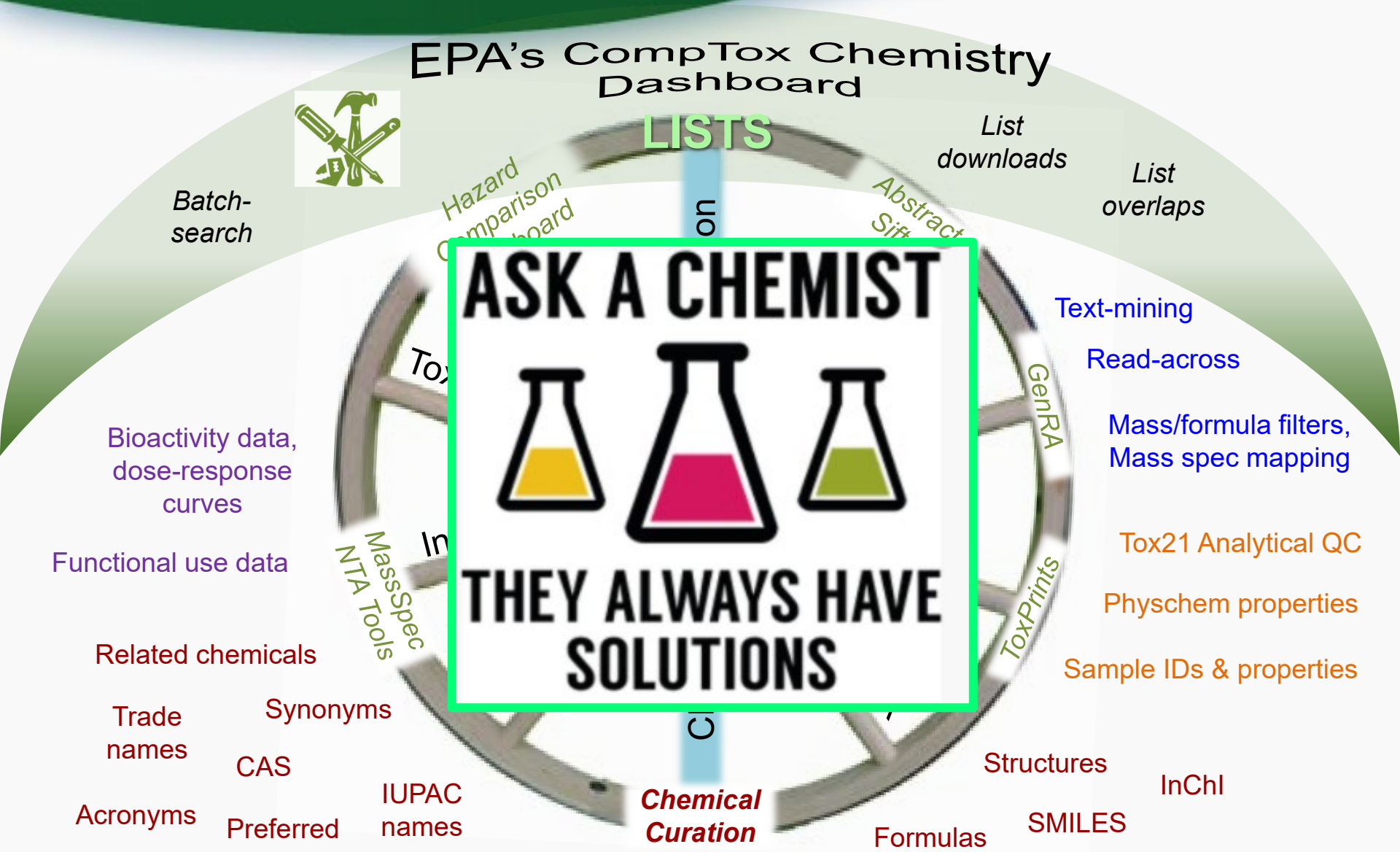

Perfluorooctanesulfonic acid
1763-23-1 | DTXSID3031864
Searched by DSSTox Substance Id.
https://comptox.epa.gov/dashboard/dsstoxdb/results?search=DTXSID3031864#invitrodb-bioassays-toxcast-tox21
DETAILS
EXECUTIVE SUMMARY
PROPERTIES
ENV. FATE/TRANSPORT
HAZARD
SAFETY
ADME
EXPOSURE
BIOACTIVITY
TOXCAST: SUMMARY
EDSP21
TOXCAST/TOX21
PUBCHEM
TOXCAST: MODELS
SIMILAR COMPOUNDS
GENRA (BETA)
RELATED SUBSTANCES
SYNONYMS
LITERATURE
LINKS
ToxCast/Tox21
QC Data ID
Grade
Description
Tox21_400083
Not determined
Analysis in progress
A Single Assay Can Have Multiple Charts
Representative Samples Only
Bioactivity Summary
Number of Charts: 271
Home / Tox21 Samples / Tox21_400083
https://tripod.nih.gov/tox21/samples/Tox21_400083
PFOS
QC Grade
T0 U Unknown/Inconclusive
T4 U Unknown/Inconclusive
Identifiers
Tox21 Tox21_400083
NCATS NCGC00164449-04
CAS 1763-23-1
PubChem 144214044
DTXSID DTXSID3031864

But sometimes additional chemical knowledge is important …



- *PFOS (and other PFAS) are unusual in their properties often requiring customized analytical approaches*
- *Commercially available sources of PFOS contain 10-40% branched form (i.e., not pure!)*
- *Branched PFOS can have significantly different ADME properties from linear form!*

DSSTox Chemical Hub



Acknowledgements:

- Chemistry Team & DSSTox Curators
  Chris Grulke (ChemReg)
  Antony Williams (CompTox Chemicals Dashboard)
  Indira Thillainadarajah, Saku Sivasupramaniam, Brian Meyer, Vicente Samano, Linda Adams
- ToxCast
  Richard Judson, Katie Paul-Friedman, Keith Houck (retired) and cast of many
- ChemTrack
  Kathy Coutros
- Tox21 Analytical Team (NCATS, NTP)
  Ruili Huang, Christopher LeClaire, Suramya Waidyanatha

This work was reviewed by EPA and approved for publication but does not necessarily reflect official Agency policy.