Identifying candidate reference chemicals for *in vitro* testing of the retinoid pathway from public data and literature mining

Nancy C. Baker[1], Jocylin D. Pierro[2], Laura W. Taylor[2], Thomas Knudsen[2]

[1] Leidos, RTP, NC; [2] CCTE, EPA, RTP, NC

Nancy C. Baker I baker.nancy@epa.gov I 0000-0002-8351-9435

www.epa.gov

Introduction

Reference chemicals are essential to calibrate *in vitro* assays and establish confidence in their results. The retinoid pathway is critical in embryonic development and chemical perturbation of this pathway is known to lead to adverse outcomes. *In vitro* testing of chemicals against the major targets of the retinoid pathway can help predict activity for environmental chemicals. Using public data sources, we assembled a set of candidate reference chemicals for ten targets of the retinoid pathway:

- Retinol binding protein (serum) – RBP
- Stimulated by retinoic acid 6 (STRA6)
- Cellular retinol binding protein (CRBP)
- Cellular retinoic acid binding protein (CRABP)
- Retinoic acid 4-hydroxylase (CYP26)
- Retinol dehydrogenase (RDH)
- Retinal dehydrogenase (RALDH)
- Retinoic acid receptors (RARA, RARB, RARG)

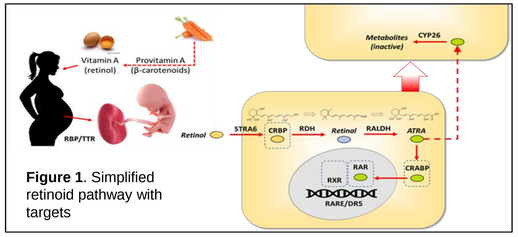


Figure 1. Simplified retinoid pathway with targets

Methods

Structured databases of in vitro testing results were searched for candidate chemicals, including: ToxCast / Tox21; ChEMBL; Protein Data Bank (PDB); and Biomedical literature concepts in EPA LitDB.

In order to establish confidence in our candidate chemicals, we used the PubMed Abstract Sifter[1] to assess the literature connections to adverse outcomes associated with retinoid pathway disruption.

Candidate chemical compilation

Hits in ToxCast RARa assays. The results were filtered to chemicals with AC50 under 2 micromola
Curation steps were used to remove chemical hits with caution flags.

Assay	DSSToxID	Chemical Name	AC50 (uM)
ATG_RARa_trans_up	DTXSID3023556	Retinol	0.0763
	DTXSID2022880	Danazol	0.7338
	DTXSID9020453	Dieldrin	0.7704
	DTXSID8024151	Imazalil	0.9084
	DTXSID8041379	2,6-Di-tert-butyl-4-methoxyphenol	1.0705
	DTXSID9037539	Endosulfan I	1.3836
	DTXSID2032500	Triflumizole	1.4526
	DTXSID4038922	Tetrabromobisphenol A bis(2-hydroxyethyl) ether	1.5032
	DTXSID7047279	CP-532623	1.5610
	DTXSID4021559	2,6-Di-tert-butyl-4-nitrophenol	1.8870
NVS_NR_hRAR_Antagonist	DTXSID2026523	Symclosene	0.2638
	DTXSID8041909	Dysprosium(III) chloride	0.8540
NVS_NR_hRARa_Agonist	DTXSID7021239	all-trans-Retinoic acid	0.1258
	DTXSID9022310	Daidzein	0.9369
	DTXSID0023901	Bentazone	1.8986
	DTXSID8037706	Potassium perfluorooctanesulfonate	1.9118

US EPA ToxCast and Tox21 assays data were extracted from the EPA Chemicals Dashboard (comptox.epa.gov) for RARA, RARB, RARG and the retinoid pathway (ATG_DR5 and Tox21_RAR). A stringent cutoff of 2.0 micromolar AC50 was applied in order to retrieve only the most potent hits.

Figure 2. Sample data from ToxCast.

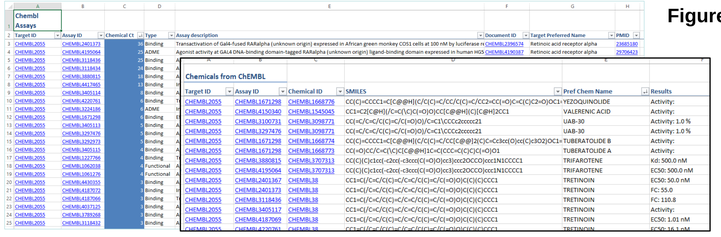


Figure 3. Sample data from ChEMBL.

ChEMBL was searched for each of the targets and records downloaded to Excel for evaluation and scanning for chemical entities with identified names. For candidate chemicals, the publication was retrieved and reviewed.

Figure 4. Sample data from PDB.

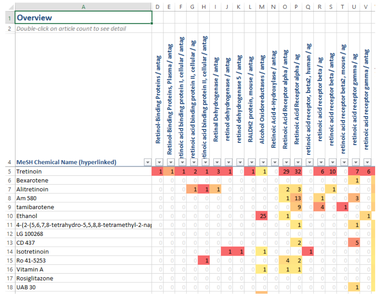
Sample of Protein Data Bank records for Cellular Retinol Binding Protein (CRBP)

PDB Link	Structure Title	PMID	Year Dep	Ligands Used
1CBQ	CRYSTAL STRUCTURE OF CELLULAR RETINOIC-ACID-BINDING PROTEINS I AND II IN COMPLEX WITH ALL-TRANS-RETINOIC ACID AND A SYNTHETIC RETINOID	7704533	1994	6-(2,3,4,5,6,7-HEXAHYDRO-2,4,4-TRIMETHYL-1-METYLENINDEN-2-YL)-3-METHYLHEXA-2,4-DIENOIC ACID / PHOSPHATE ION
1CBR	CRYSTAL STRUCTURE OF CELLULAR RETINOIC-ACID-BINDING PROTEINS I AND II IN COMPLEX WITH ALL-TRANS-RETINOIC ACID AND A SYNTHETIC RETINOID	7704533	1994	RETINOIC ACID
6E5L	Crystal structure of human cellular retinol binding protein 1 in complex with abnormal-cannabidiol (abn-CBD)	30721022	2018	(1'R,2'R)-5'-methyl-6-pentyl-2'-(prop-1-en-2-yl)-1',2',3',4'-tetrahydro[1,1'-biphenyl]-2,4-diol
6E5S	Crystal structure of holo retinal-bound domain-swapped dimer Q108K:T51D mutant of human Cellular Retinol Binding Protein II	31557439	2018	RETINAL
6E5T	Crystal structure of human cellular retinol binding protein 1 in complex with abnormal-cannabidiorcin (Abn-CBDO)	30721022	2018	(1'R,2'R)-5',6-dimethyl-2'-(prop-1-en-2-yl)-1',2',3',4'-tetrahydro[1,1'-biphenyl]-2,4-diol

PDB was searched for each target and results downloaded and reviewed.

Figure 5. Sample data from LitDB literature term database

EPA's database of MeSH terms extracted from PubMed was queried for each of the targets and co-annotations of chemicals and antagonism or agonism. The results were loaded to Excel and reviewed.

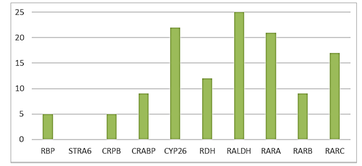


Results summary

The process produced approximately 220 putative reference chemicals. The results differed among the targets.

- For STRA6, no putative reference chemicals were identified.
- RALDH has many candidates, but this enzyme family has many members that share substrates.
- Targets of interest to drug discovery have the most in vitro testing activity, but fewer environmental chemicals are represented.

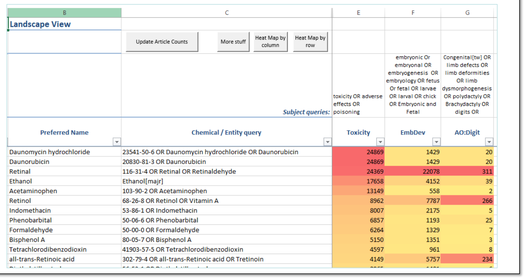
Figure 6. Candidate reference chemical counts by target



PubMed searches with the Abstract Sifter tool were performed and yielded 137 (~ 75%) chemicals showed connections to developmental toxicity.

As an example, 55 chemicals had literature on limb defects.

Figure 7. Abstract Sifter article counts for toxicity, developmental toxicity, and limb defects



Next steps

CYP26 candidate reference chemicals

5-7 compounds for purchase and evaluation

Tox21 Cross-Partner Project 13 is exploring development of a high-throughput assay for CYP26 bioactivity. The candidate chemical collection is undergoing review to winnow the list for this proof-of-concept study. Considerations for selections to the shortlist are:

- Strong supportive data
- Environmental interest
- Known developmental toxicant
- Chemical diversity

U.S. Environmental Protection Agency
Office of Research and Development

This poster does not necessarily represent U.S. EPA policy.

References:

1) Baker N et al. F1000 Res. 2017 Dec 21;6. pii: Chem Inf Sci-2164.