

US EPA ARCHIVE DOCUMENT



**CASE STUDY 1: Endocrine Disruptors:
Estrogen Receptor Expert System (ERES)**

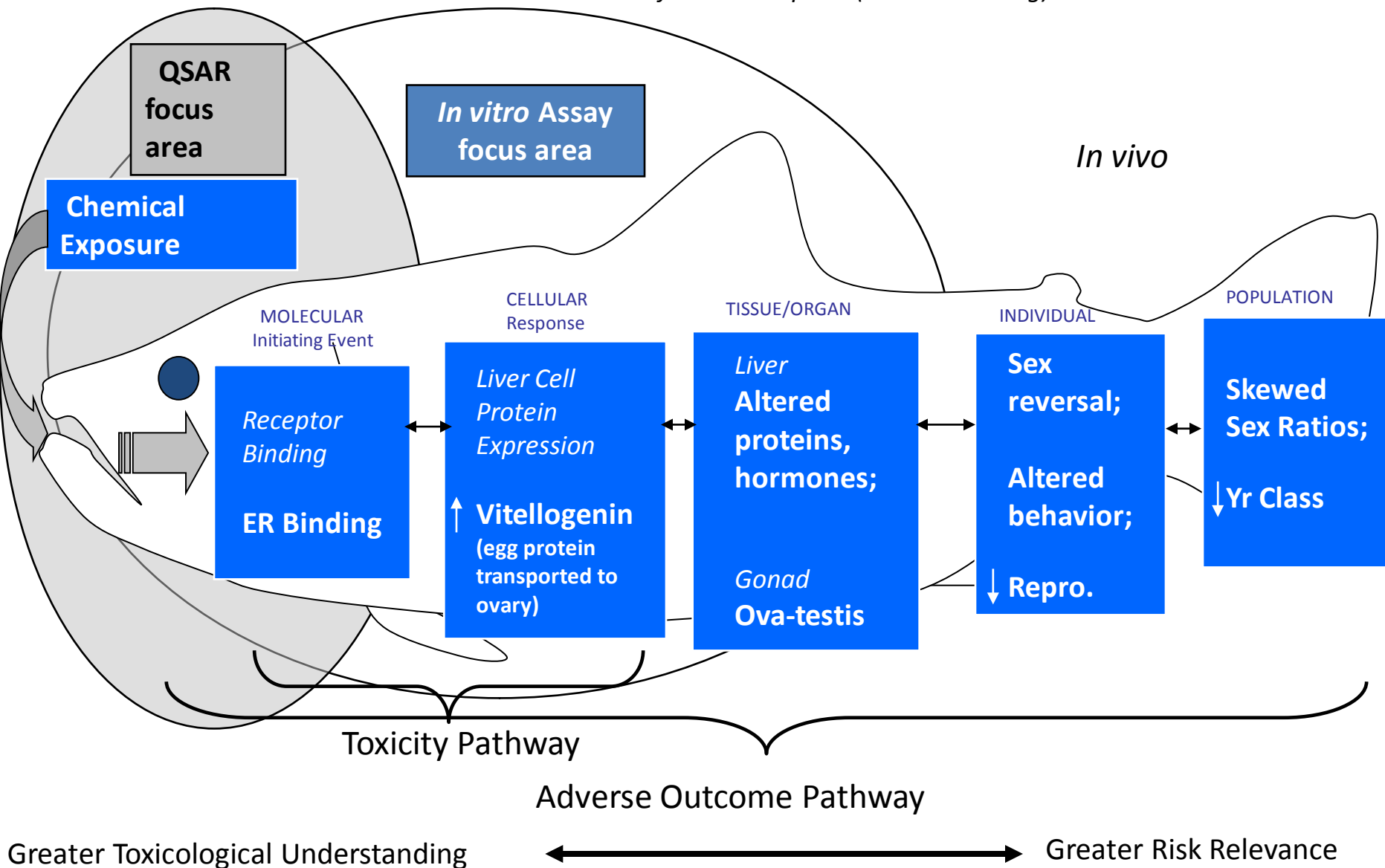
Mechanistic Basis of the Expert System

- Adverse Outcome Pathway (AOP) context
- OECD QSAR Validation Principles
- Expert System decision tree codes the measured data from *in vitro* assays designed for purpose
 - Assays optimized to detect any potential for ER interaction, including low affinity binding
 - Provides the data and expert interpretation in decision tree format; provides predictions based on extrapolation of measured data (displays training set used in prediction) so user can interpret and assess use for their purpose
 - Transparency
 - Can the QSAR estimate be explained mechanistically?
 - How reasonable is an estimate compared with data for similar chemicals?
 - Usefulness
 - Are the predictions applicable to all the chemicals of regulatory concern?
 - Does the model/expert system answer the regulatory question?

ER-mediated Reproductive Impairment Adverse Outcome Pathway

Chemical Effects Across Levels of Biological Organization

Mechanistic linkage between the risk assessment endpoint (ER-mediated reproductive impairment) and the hazard identification endpoint (MIE: ER binding)



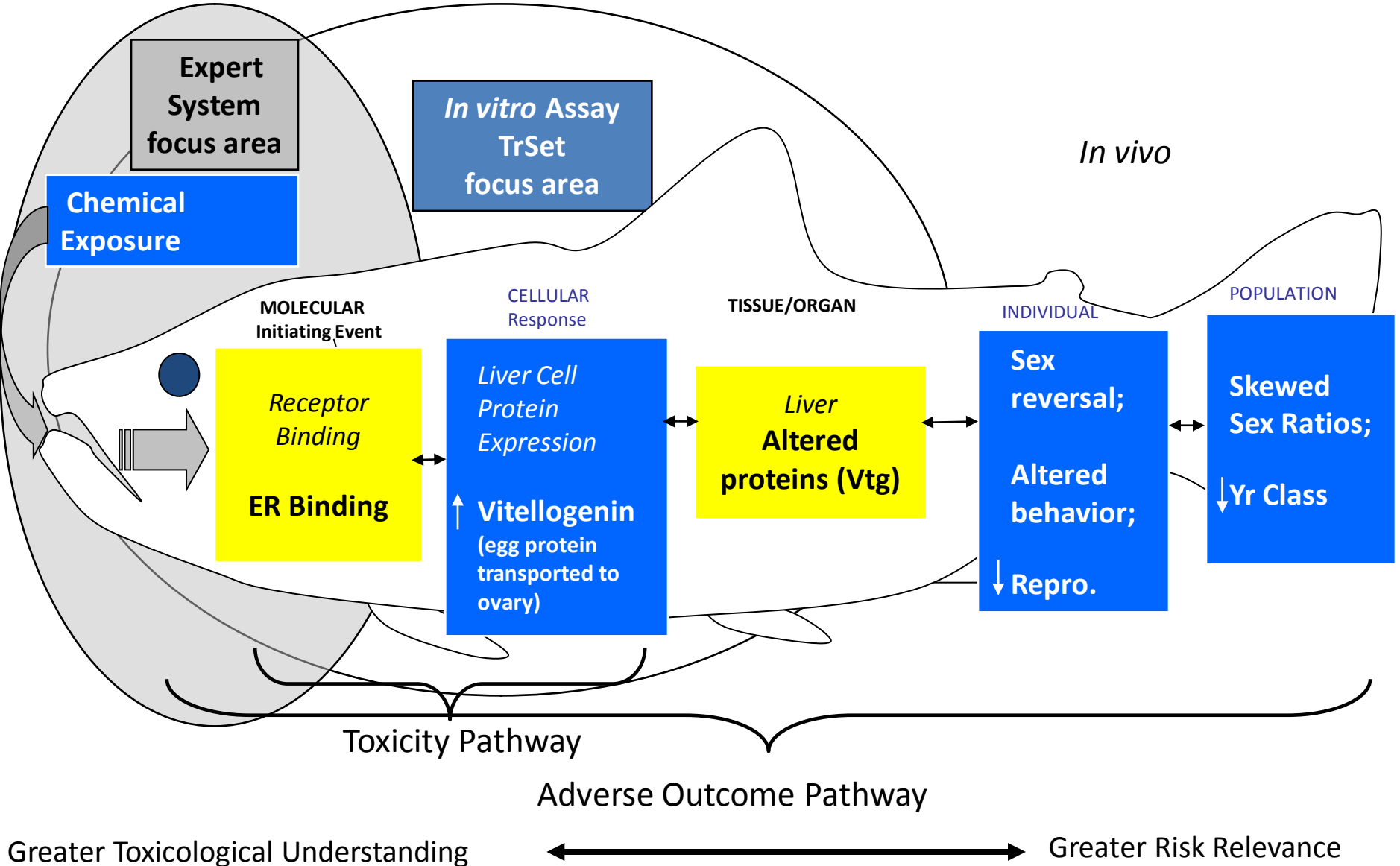
ER Binding Affinity: An Indicator of Potential Reproductive Effects

- *In vitro* Assays used to build the Expert System are along an adverse outcome pathway (AOP) ending in reproductive impairment
 - The molecular initiating event (MIE) of the pathway is ER binding (*msrd*)
 - Tissue level response key event along the pathway confirms a higher level response (*msrd*)
 - The measured data identifies which chemical structures can initiate the pathway and subsequent key event
 - The Expert System extrapolates from the measured data to predict ER binding potential of unmeasured chemicals that are within the bounds of measurement
 - The Expert System also indicates when a chemical is outside the bounds of the measured data thus accurate predictions are not possible (Unknown Binding Potential)
 - AOP context provides conceptual model useful for generating testable hypotheses (e.g., prioritization for Tier 1 screening)
 - AOP context provides decision-making rationale for the regulatory community (e.g., knowns and unknowns along the AOP)

ER-mediated Reproductive Impairment Adverse Outcome Pathway

Chemical Effects Across Levels of Biological Organization

Mechanistic linkage between the risk assessment endpoint (ER-mediated reproductive impairment) and the hazard identification endpoint (MIE: ER binding)



OECD Principles for QSAR Validation

- Well-Defined Endpoint – (*in vitro* assay domain)
 - Well-defined biological endpoint –
 - Informs important risk endpoint –
 - AOP ending in impaired reproduction provides plausible linkage of MIE to higher level adversity as basis for prioritizing chemicals for higher level assays
 - Interpreting the measurements
 - Measurement endpoint and confounding factors are discussed
 - Well-defined chemistry
 - Using *in vitro* assays that allow testing of the types of chemicals (range of properties) found on regulatory inventories
 - Understanding the chemical form and concentration in the assays
- Mechanistic interpretation
 - Being able to explain the predictions mechanistically
 - With respect to chemistry & biology in the assay system
 - Relationship of predicted parameter to regulatory question
 - Likelihood to initiate ER-mediated Reproductive Impairment AOP
 - Relationship of chemical parameters to biological activity

OECD Principles for QSAR Validation

- Defined Model Applicability Domain
 - Well-defined application
 - Regulatory question – priority setting not predicting adverse outcome
 - Expert System model domain coverage well-defined
 - Decision tree, logic rules, local *in vitro* TrSets upon which rules are based
 - Expert System model domain adequately covers the Regulatory Chemical domain
 - Fooduse pesticidal Inerts (FI) ; Antimicrobials (AM)
 - EDSP Universe

- Appropriate measures of goodness of fit, robustness, ability to predict
 - Measures appropriate for a regression model are not appropriate to evaluate an expert system logic rules
 - Series of local models, local *in vitro* TrSets, “unknown” structure compared to tested chemicals

- Unambiguous algorithm
 - Expert Systems – logic tree, rules/queries, supporting information

ER Expert System – Effects-Based Chemical Categories Approach *Chemical Similarity*

Building Effects-based Categories:

Structural similarity defined by similar biological activity

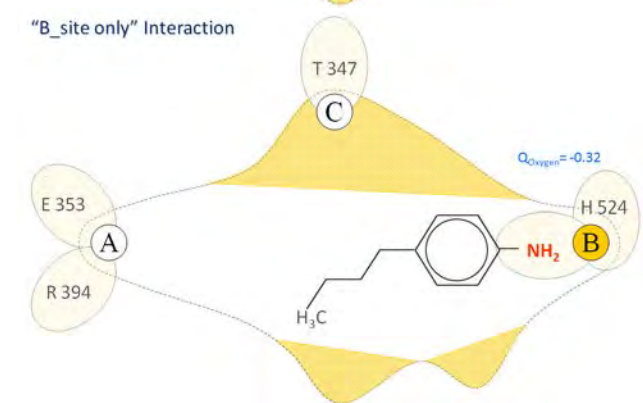
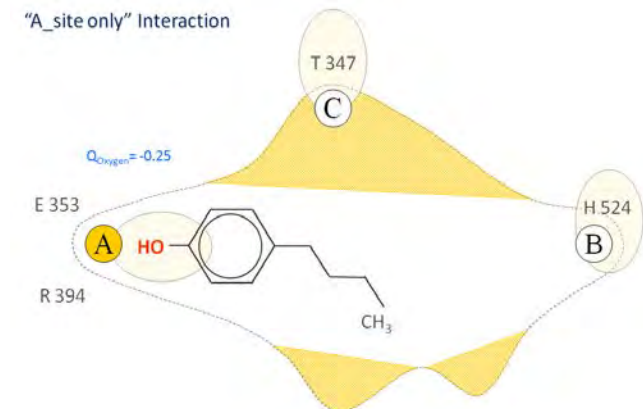
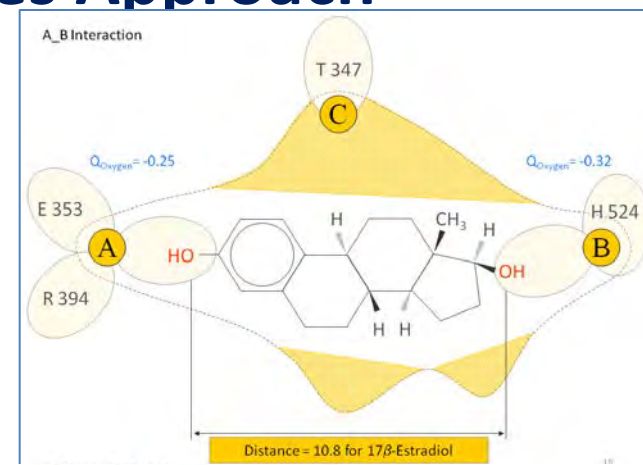
- MIE of ER-mediated AOP

Multiple ER-interaction types recognized

-chemicals can interact at different points ('A' or 'B'), depending on their properties

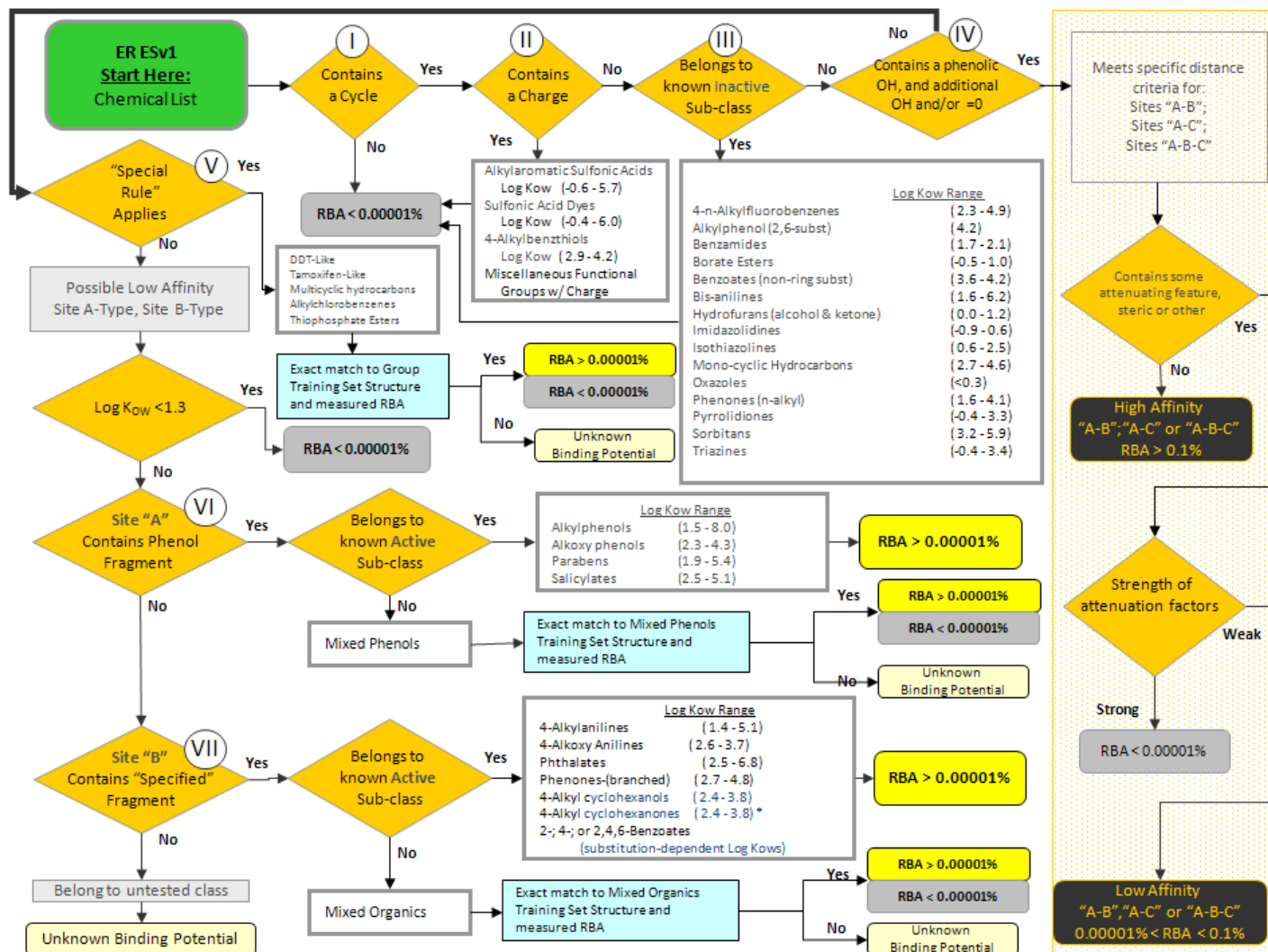
-chemicals initiating MIE at same point are 'similar'

Common biological activity within a chemical structural series is coded into ES logic rules in decision tree

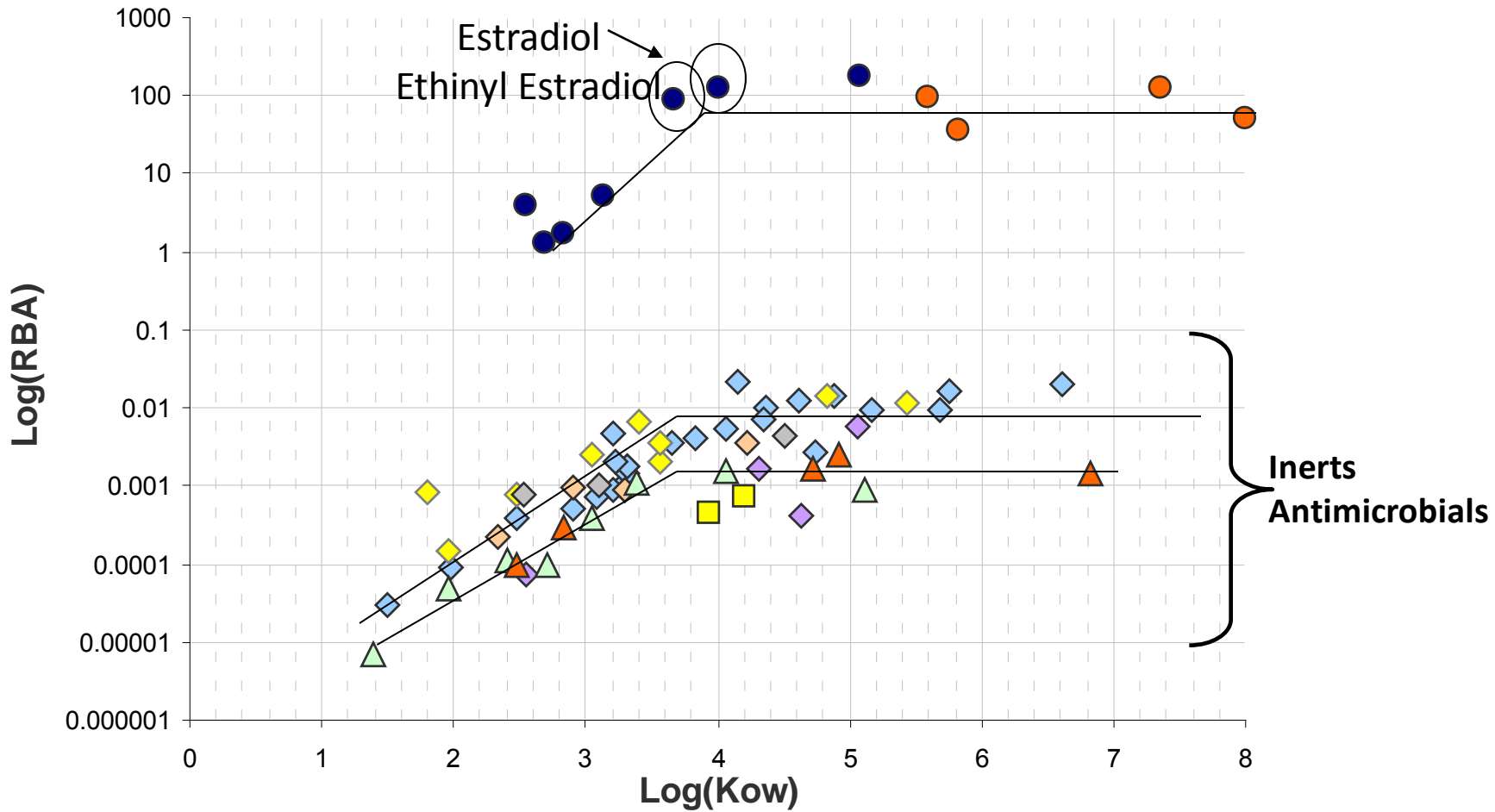


ER Expert System Decision Tree

- Expert System codes the measured data in a decision tree from the *in vitro* assays designed for purpose

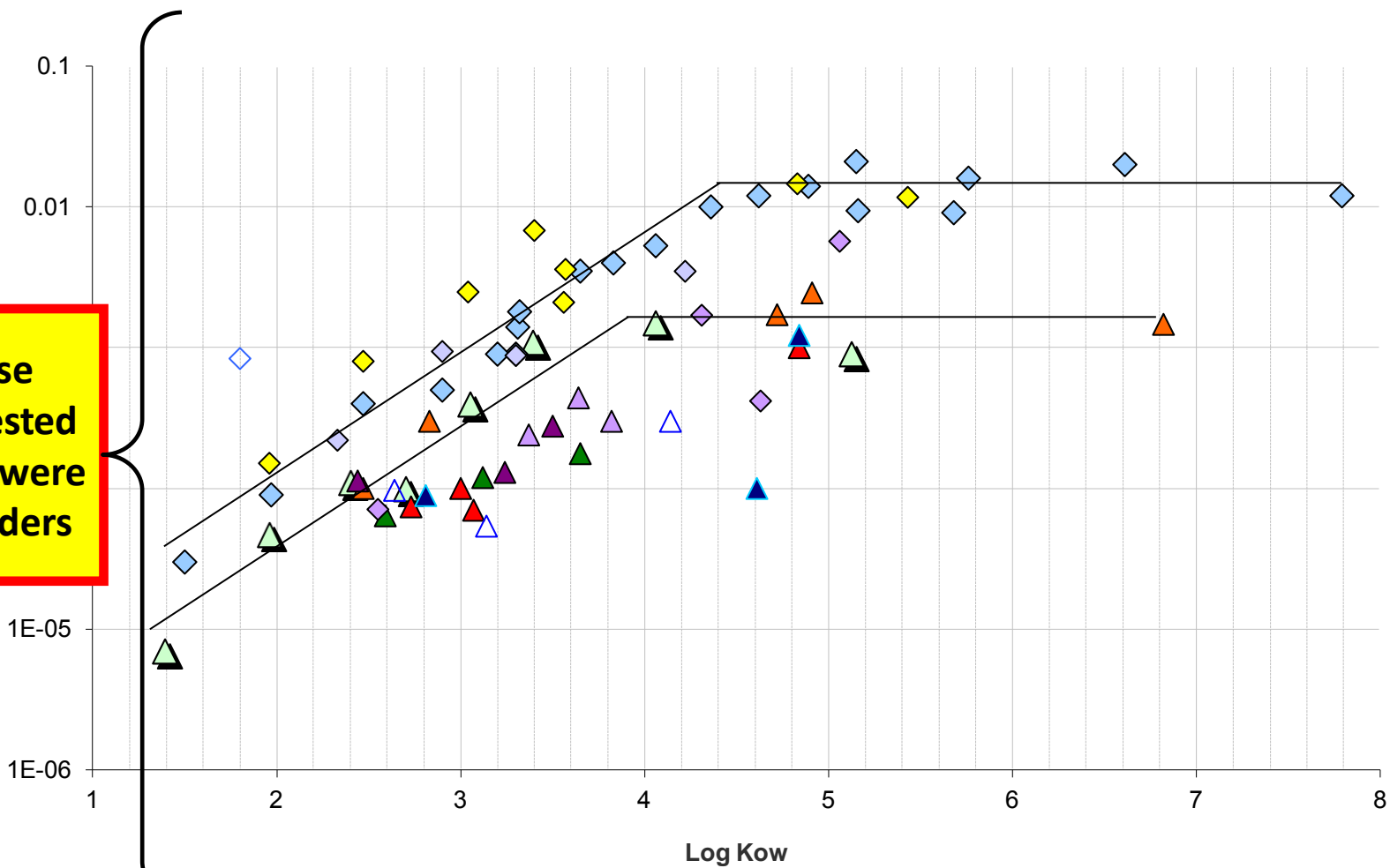


ER Binders



Trout ER Relative Binding Affinity vs. Log Kow

RBA = relative binding affinity compared to Estradiol at 100%

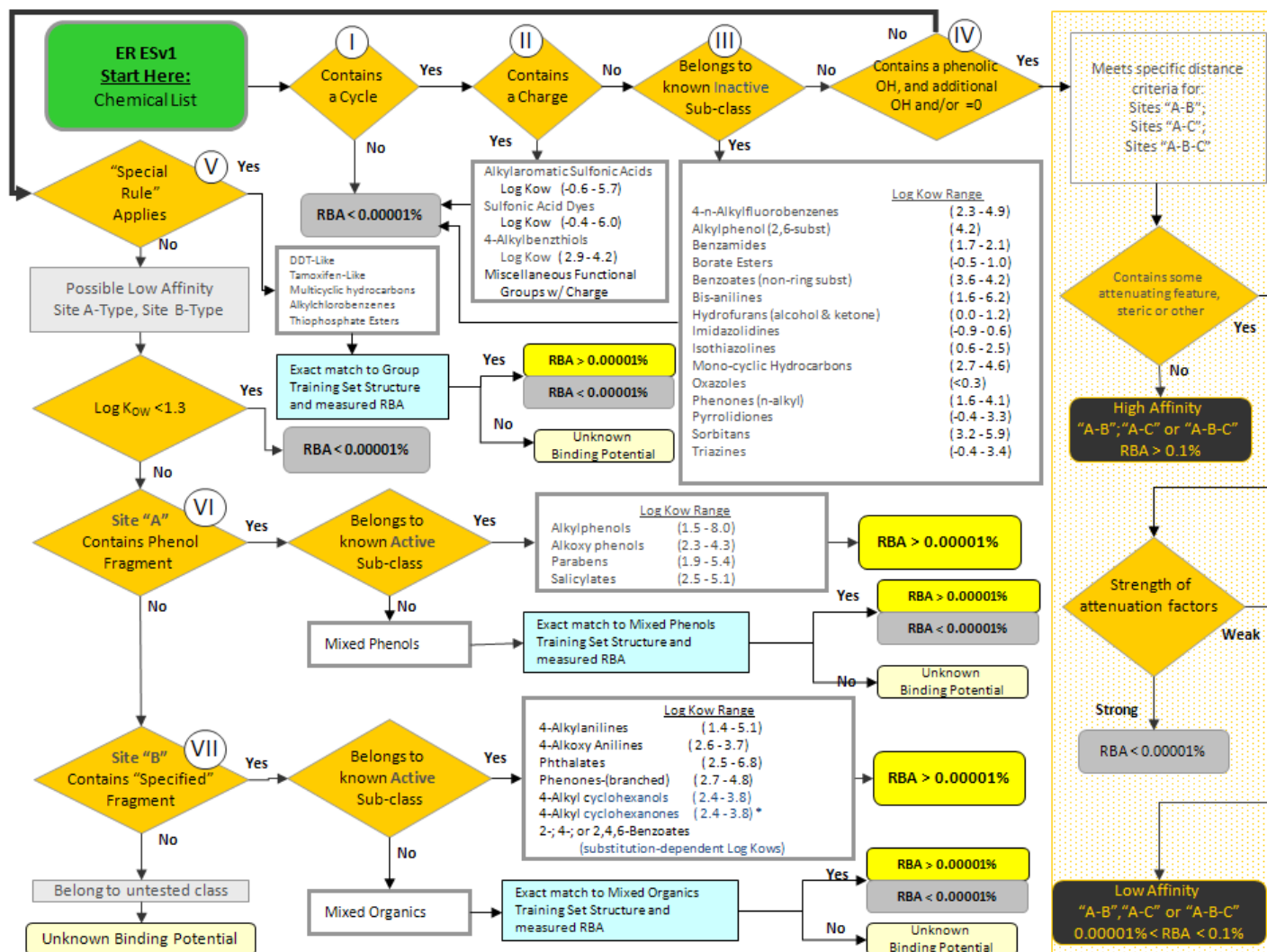


**~80 diverse chemicals tested
LogK_{ow} < 1.3 were all NON-binders**

Many chemicals & large chemical groups were found to be NON-binders

ER Expert System Decision Tree

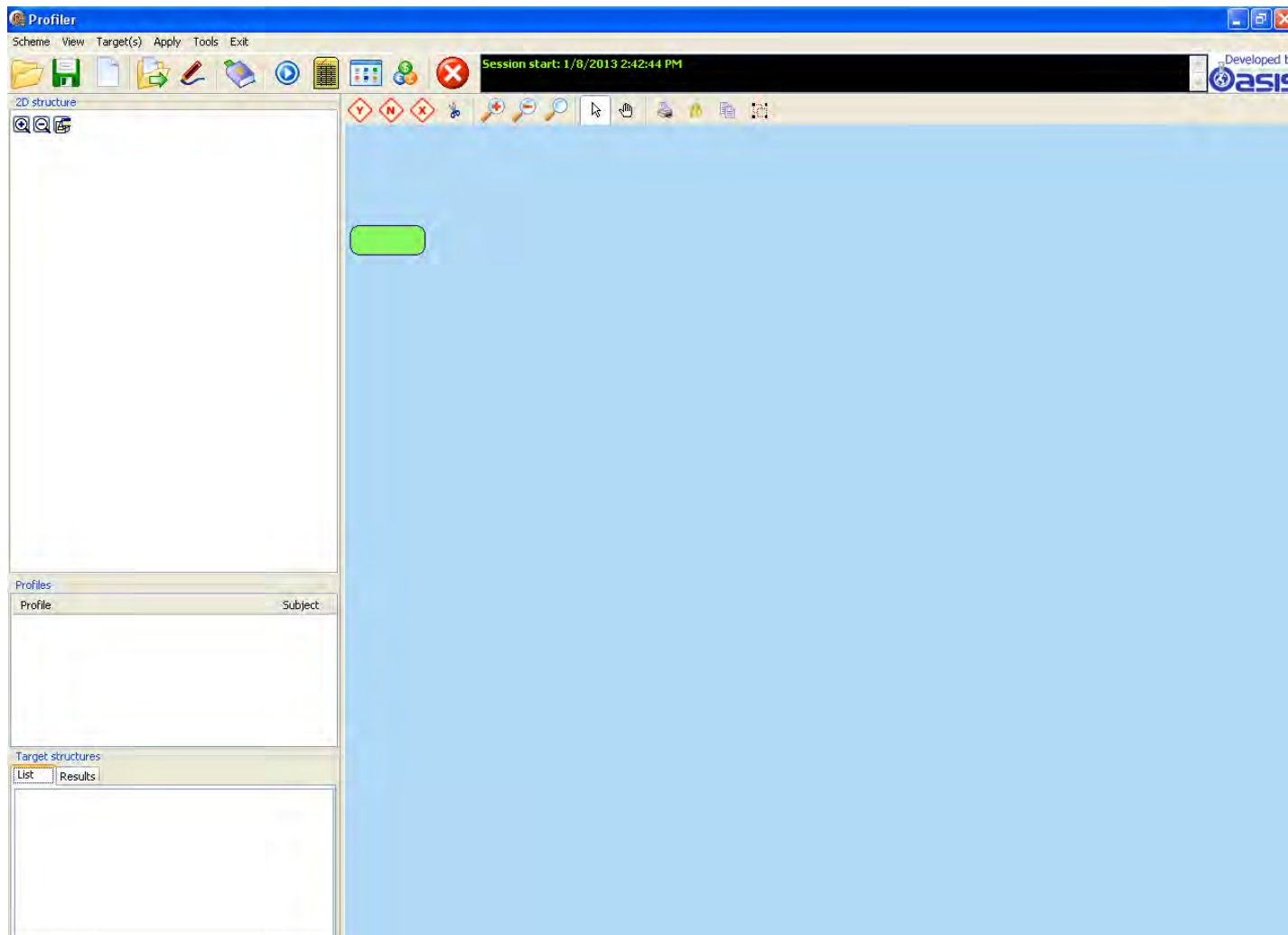
- Expert System codes the measured data in a decision tree from the *in vitro* assays designed for purpose



Automation of the ER Expert System

Built through collaborative effort between EPA, OECD and LMC

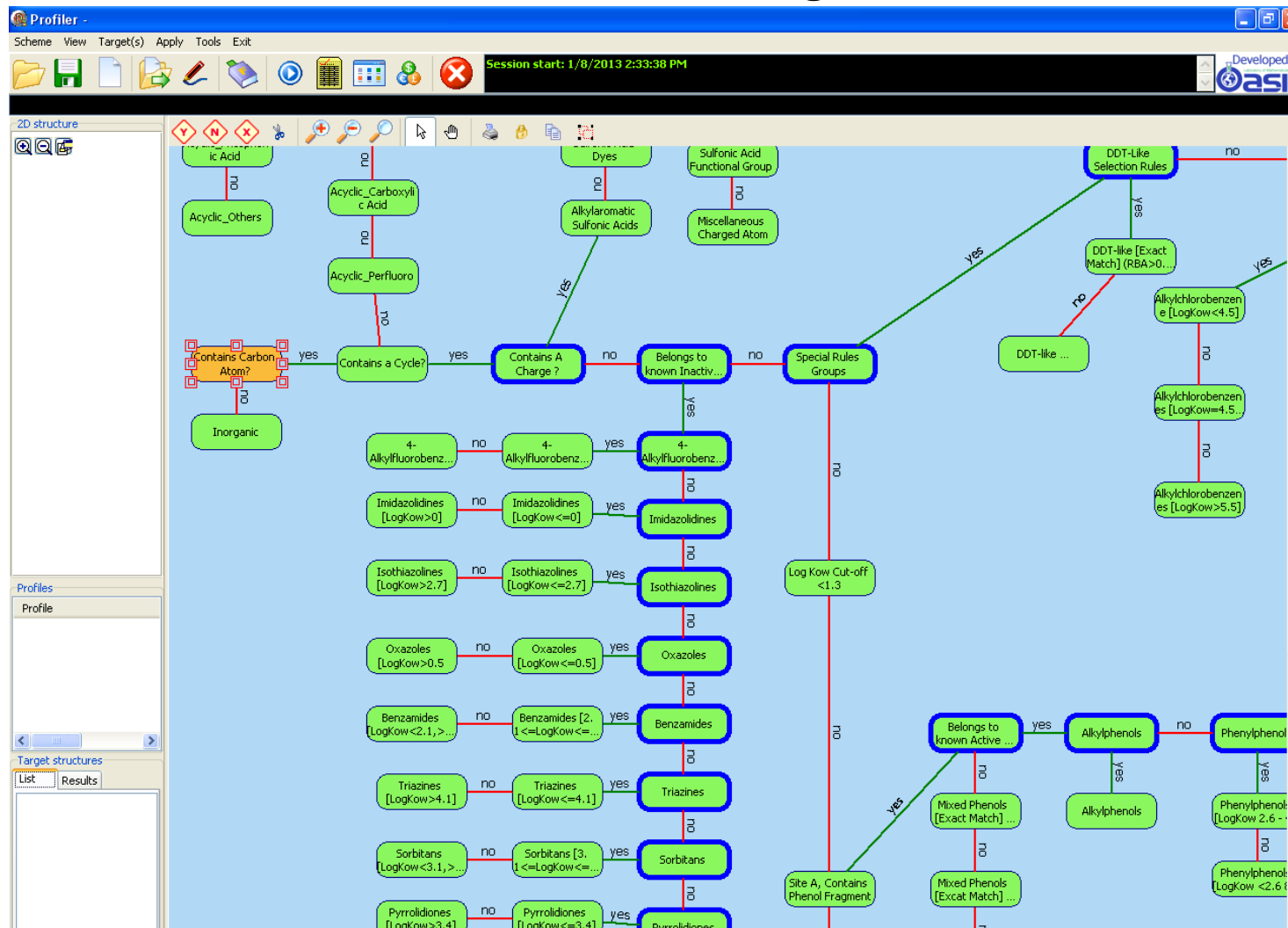
<http://www.qsartoolbox.org/>



Automation of the ER Expert System

Decision Tree

Yes/No decision-based dendroid logic scheme



Automation of the ER Expert System

Single chemical profiling....

The screenshot displays the Profiler software interface. On the left, the 2D structure of a chemical compound is shown, which is 4-(2-hydroxyphenyl)phenol. The main window features a complex decision tree for classification. The tree starts with 'Contains Carbon Atom?' and branches through various categories like 'Acyclic_Carboxylic Acid', 'Acyclic_Perfluoro', 'Contains a Cycle?', 'Contains A Charge?', 'Belongs to known Inactive', 'Special Rules (Groups)', 'Log Kow Cut-off <1.3', 'Site A, Contains Phenol Fragment', and 'Belongs to known Active'. The final classification is 'Mixed Phenols [Exact Match]'. The interface includes a menu bar (Scheme, View, Target(s), Apply, Tools, Exit), a toolbar, and a status bar showing the target structure: c1(O)c(Cc2ccc(O)cc2)cccc1. The bottom left panel shows the 'Profiles' section with 'Profile: Unknown Binding Potential' and 'Subject: Parent'. The 'Target structures' list contains the same SMILES string.

Automation of the ER Expert System

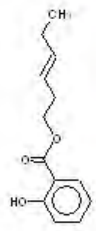
- Or batch profiling, e.g., inventory list

The screenshot displays the Profiler software interface. On the left, a 2D chemical structure is shown, which is a substituted imidazolidine ring with a chlorine atom, two methyl groups, and a bromine atom. Below the structure, the 'Profiles' section shows 'RBA < 0.00001%' and 'Parent'. The 'Target structures' list includes various chemical entities such as '3,5,7-Triaza-1-azoniatricyclo(3.3.1.1) (superscript 3)', 'Copper', '2H-Cyclopent(d)isothiazol-3(4H)-one', '1H-Pyrazole-1-methanol', 'Bromochloro-5-ethyl-5-methyl-2,4-imidazolidinedi...', '1,3-Dichloro-5-ethyl-5-methylhydantoin', 'Nonanoic acid, sulfonyl ester, sodium salt', 'Methanol', '1-Octanesulfonic acid, 2-sulfino', '1,2-Octanesulfonic acid', and 'Didecyl dimethyl ammonium carbonate and didecyl...'. The main area features a decision tree with nodes like 'Acidic Acid', 'Acyclic Others', 'Acyclic Carboxylic Acid', 'Acyclic Perfluoro', 'Inorganic', 'Contains Carbon Atom?', 'Contains a Cycle?', 'Contains A Charge?', 'Alkylaromatic Sulfonic Acids', 'Dyes', 'Sulfonic Acid Functional Group', 'Miscellaneous Charged Atom', 'Belongs to known Inactive...', 'Special Rules Groups', 'Log Kow Cut-off < 1.3', 'Alkylfluorobenz...', 'Imidazolidines [LogKow > 0]', 'Imidazolidines [LogKow <= 0]', 'Isothiazolines [LogKow > 2.7]', 'Isothiazolines [LogKow <= 2.7]', 'Oxazoles [LogKow > 0.5]', 'Oxazoles [LogKow <= 0.5]', 'Benzamides [LogKow > 2.1, > ...]', 'Benzamides [2.1 <= LogKow <= ...]', 'Triazines [LogKow > 4.1]', 'Triazines [LogKow <= 4.1]', 'Sorbitans [LogKow < 3.1, > ...]', 'Sorbitans [3.1 <= LogKow <= ...]', 'Pyrrolidones [LogKow > 3.4]', 'Pyrrolidones [LogKow <= 3.4]', 'Belongs to known Active ...', 'Mixed Phenols [Exact Match] ...', and 'Site A, Contains Phenol Fragment'. A green arrow labeled 'yes' points from the 'Special Rules Groups' node to a 'DDT-like' label.

The path followed through yes/no questions in decision tree to the final decision point is displayed.

Sch

2D



Profiles

Profile

RBA>0.00001%

3-Alkylanilines

Multi subst. Alkylanilines

Target structures

List Results

33059051 - 2-HYDROXY-4-...

34689468 - Phenol,_methy...

35285699 - Benzoic_acid,...

51115630 - Benzoic_acid,...

51363645 - disodecyl_pher...

56803373 - tert-Butylphen...

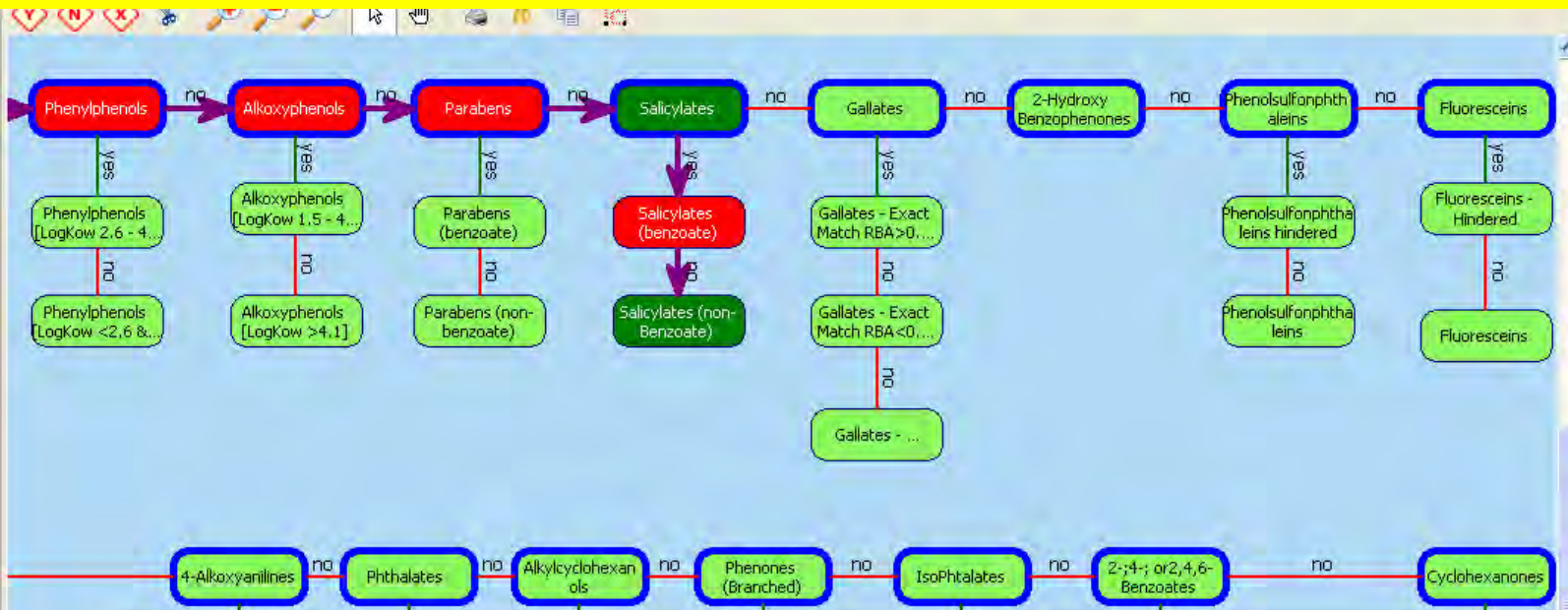
65405778 - cis-3-Hexenyl...

65652417 - Phosphoric_aci...

68555588 - 3-Methyl-2-but...

79915745 - beta.-Isopropo...

84852153 - 4-Nonylphenol,

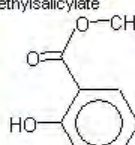
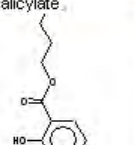
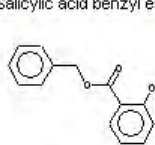
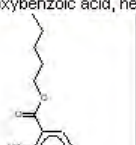
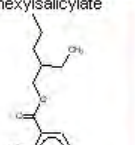
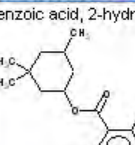


Salicylates (non Benzoate), ID: 38

Properties Reference Target Metabolism

Training Set Boundaries

Grid List

Methylsalicylate  RBA (%) 0.000071 KOWWIN LogKow 2.6	Butyl salicylate  RBA (%) 0.00042 KOWWIN LogKow 4.08	Salicylic acid benzyl ester  RBA (%) 0.0017 KOWWIN LogKow 4.31	2-Hydroxybenzoic acid, he2-Ethylhexylsalicylate  RBA (%) 0.0057 KOWWIN LogKow 5.06	  RBA (%) 0.00031 KOWWIN LogKow 5.97
Benzoic acid, 2-hydroxy-,  RBA (%) 0.000062 KOWWIN LogKow 6.16				

CAS 65405-77-8
LogKow 4.28

cis-3-Hexenyl salicylate

#	Status	ID	Fragmen	Fragmen
1	Satisfied	1	c1(C=O	Not four

Properties Reference Target

Caption

4-Alkylanilines 1.3<LogKow<5.1

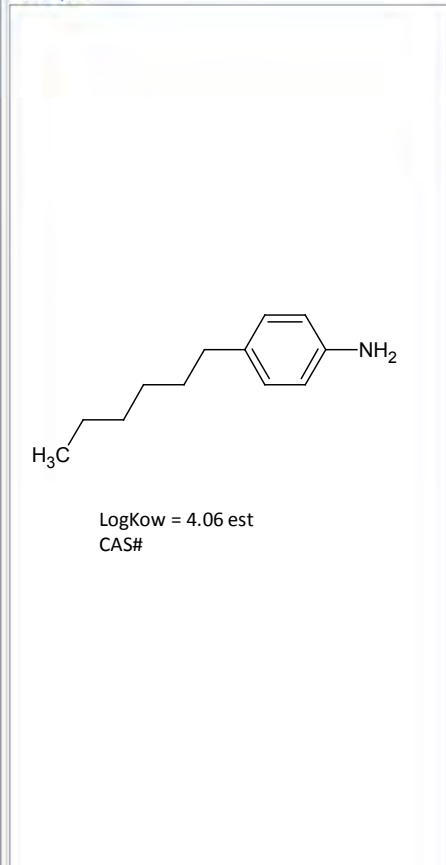
YES Category Name

RBA>0.00001%

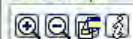
NO Category Name

Label to assign to the target if it doesn't meet the criteria

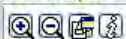
Description



YES Example

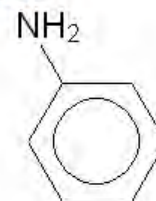


NO Example



Boundaries Training Set

Chemical name	SMILES	RBA (%)	LogKow	LogKow ...	Acronym	CAS No.
Aniline	<chem>c1(N)cccc1</chem>	<0.000001	0.90	msrd	ANL	62533
4-Methylaniline, p-Toluidine	<chem>c1(N)ccc(C)cc1</chem>	0.000007	1.39	msrd	PT	106490
4-Ethylaniline	<chem>c1(N)ccc(CC)cc1</chem>	0.000047	1.96	msrd	EA	589162
4-Propylaniline	<chem>c1(N)ccc(CCC)cc1</chem>	0.00011	2.40	msrd	PA	2696846
4-tert-butylaniline	<chem>C(C)(C)(C)c1ccc(N)cc1</chem>	0.000098	2.70	msrd	TBA	769926
4-n-butylaniline	<chem>c1(N)ccc(CCCC)cc1</chem>	0.00040	3.05	msrd	BA	104132
4-n-amylaniline	<chem>c1(N)ccc(CCCCC)cc1</chem>	0.0011	3.39	msrd	AAN	33228443
4-hexylaniline	<chem>c1(N)ccc(CCCCCC)cc1</chem>	0.0015	4.06	ClogP	HAL	33228454
4-n-octylaniline	<chem>c1(N)ccc(CCCCCCCC)cc1</chem>	0.00089	5.12	ClogP	OA	16245797

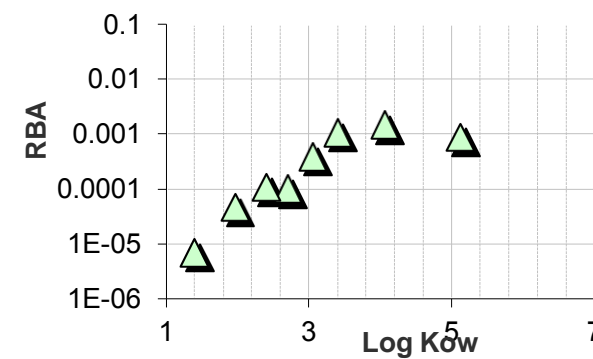
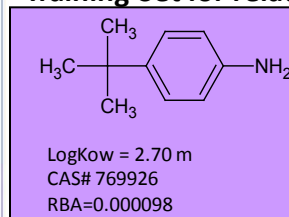


Details

Training Set for: Aniline (RBA<0.00001%), and 4,n-alkylanilines 1.3<LogKow<5.1

<p>LogKow = 0.90 m CAS# 62533 RBA<0.000001%</p>	<p>LogKow = 1.39 m CAS# 106-49-0 RBA = 0.000007</p>	<p>LogKow = 1.96 m CAS# 589-16-2 RBA = 0.000047</p>	<p>LogKow = 2.40 m CAS# 2696-84-6 RBA = 0.00011</p>	<p>LogKow = 3.05 m CAS# 104-13-2 RBA = 0.00040</p>
<p>LogKow=3.39 m CAS# 33228443 RBA = 0.0011</p>	<p>LogKow = 4.06 est CAS# 33228454 RBA=0.0015</p>	<p>LogKow = 5.12 est CAS# 16245-79-7 RBA = 0.00089</p>		

Training Set for related group: 4,t-alkylanilines:

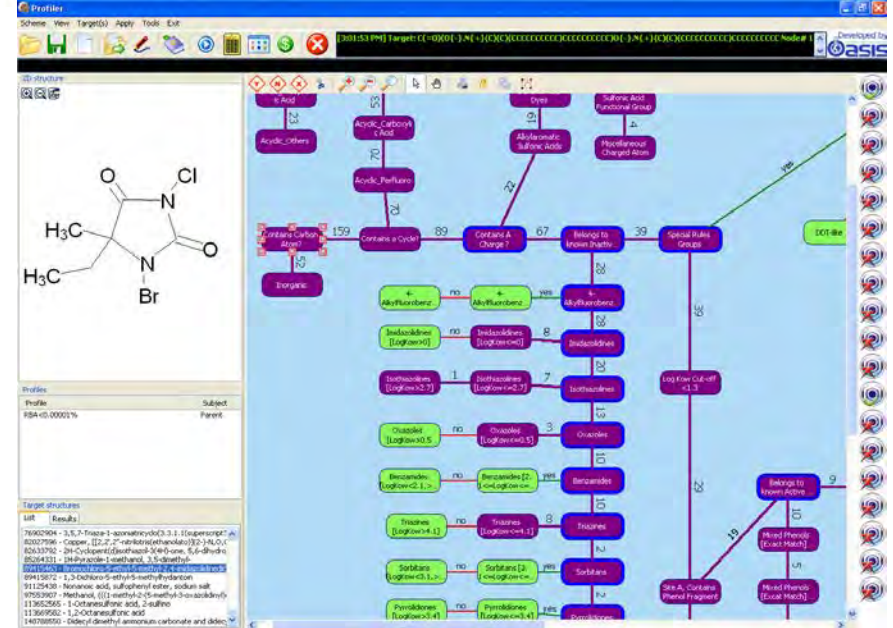
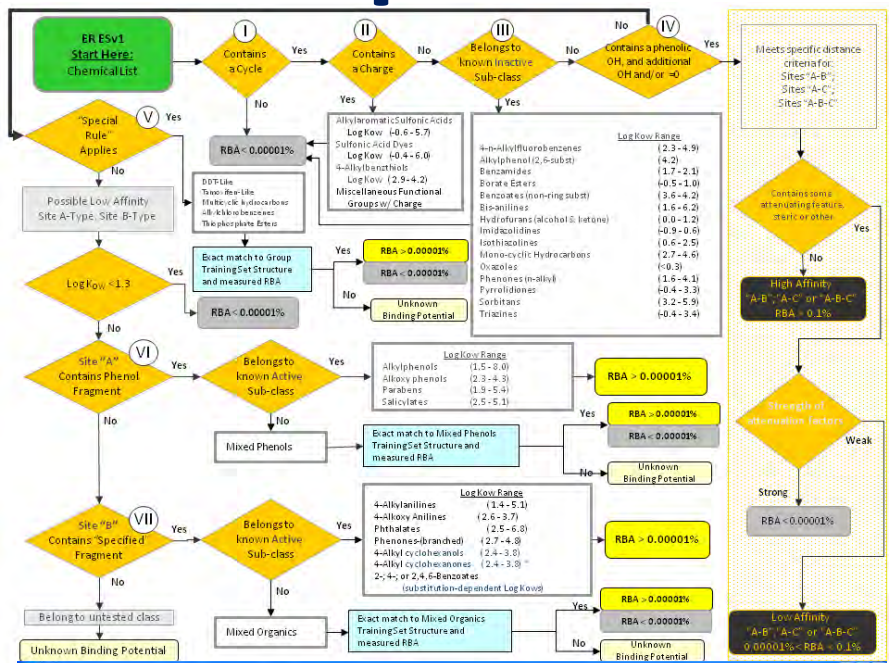


Notes on Interpretation: Aniline and a series of 4,n-alkylanilines were tested for rainbow trout ER affinity and gene activation. Aniline (LogKow=0.90) did not bind while 4,n-methyl-, ethyl-, propyl-, butyl-, pentyl-, and octylaniline all bound rER. Additionally, 4,t-butylaniline also bound rER. All 8 alkylanilines also induced VTG mRNA in trout liver slices to varying extents, with at least one concentration significantly different from controls. Alkylanilines of LogKow 3 to 4 yielded the most potent VTG induction. Therefore, 4-alkylanilines of LogKow between 1.3 and 5.2 are predicted to bind ER, with a good degree of confidence; 4,t-substituted alkylanilines in this LogKow range are also likely to bind.

Alkylanilines of Log Kow greater than that of p,n-octylaniline (5.12), or with ortho- or meta-alkylsubstituents are outside of the domain of the training set.

Expert System Development

- 2009 SAP review
 - Model Domain - **ER ESv1**
 - Regulatory Domains
 - Coverage: Food-use Inerts (FI) and Antimicrobials (~95% NON-Binders; ~5% prioritized Binders)
- 2013 SAP review:
 - Automated ER ESv1**
 - Expanding Domain Coverage with *in vitro* testing; build effects-based chemical categories:
 - Non Food-use inerts (NFI)
 - Evaluate ES Coverage of EDSP Universe = 71%; (~5% prioritized Binders)
- Additional Work:
 - Complete **ER ESv2**
 - Expand Domain (*in vitro* testing) to cover remaining EDSP Universe; build **ER ESv3**



Thank you!

Pat Schmieder

Molecular & Cellular Mechanisms Research Branch Chief

Mid-Continent Ecology Division, Duluth, MN

National Health and Environmental Effects Research Laboratory

Office of Research and Development

U. S. Environmental Protection Agency