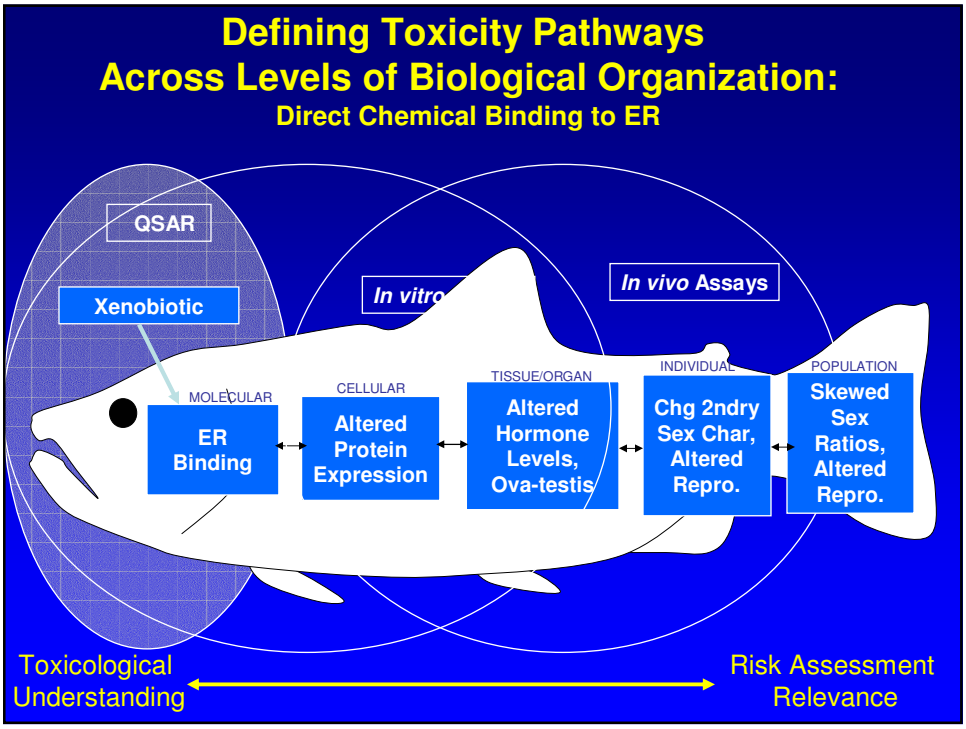
Designing a QSAR

for ER Binding

Defining Toxicity Pathways Across Levels of Biological Organization:

Direct Chemical Binding to ER



QSARs for Prioritization

What:

• Prioritize chemicals based on ability to bind ER (plausibly linked to adverse effect)

• Determine which untested chemicals should be tested in assays that will detect this activity, prioritized above very low risk chemicals for this effect

• Demonstrate how QSARs are built, for complex problems, and are useful to regulators/risk assessors

Why:

•To provide EPA with predictive tools for prioritization of testing requirements and enhanced interpretation of exposure, hazard identification and dose-response information

•Develop the means to knows what to test, when to test, how

•FQPA - Little of no data for most inerts/antimicrobials; short timeline for assessments;

Lessons Learned from early EPA exercise

1) High quality data is critical and should not be assumed

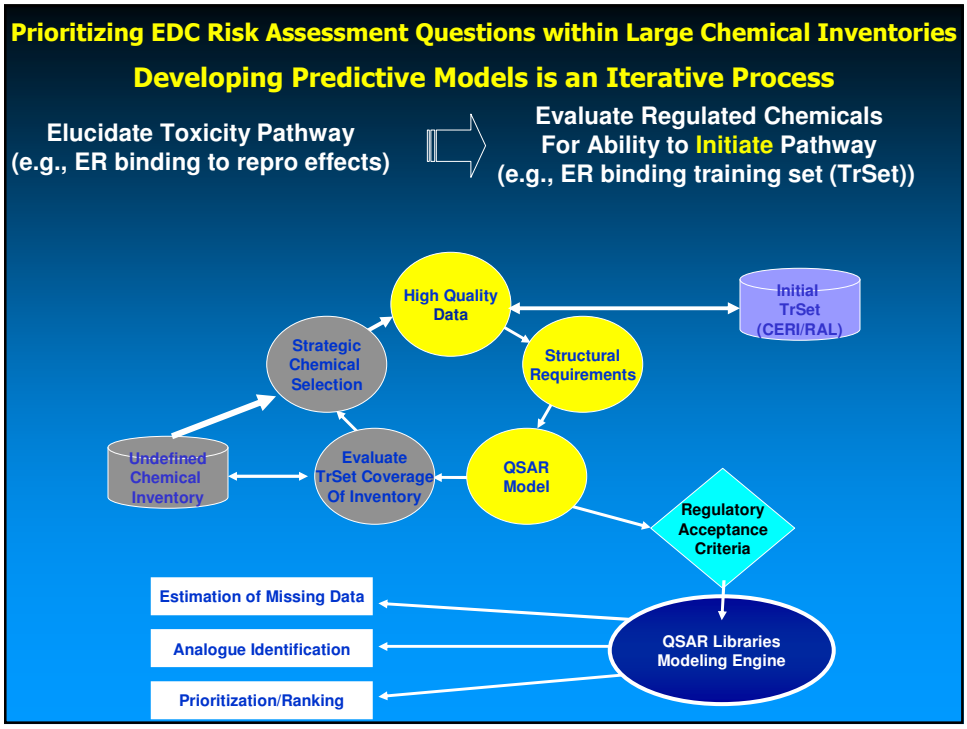
- **Models can be no better than the data upon which they are formulated**
- **Assays should be optimized to determine the adequacy for the types of chemicals found within regulatory lists**
  - **Assumption that assays adequate for high-medium potency chemicals will detect low potency chemicals warrants careful evaluation**
- **Mechanistic understanding should be sought; new information incorporated when available**
  - **Assumption that ER binding mechanism was well understood warrants careful evaluation**

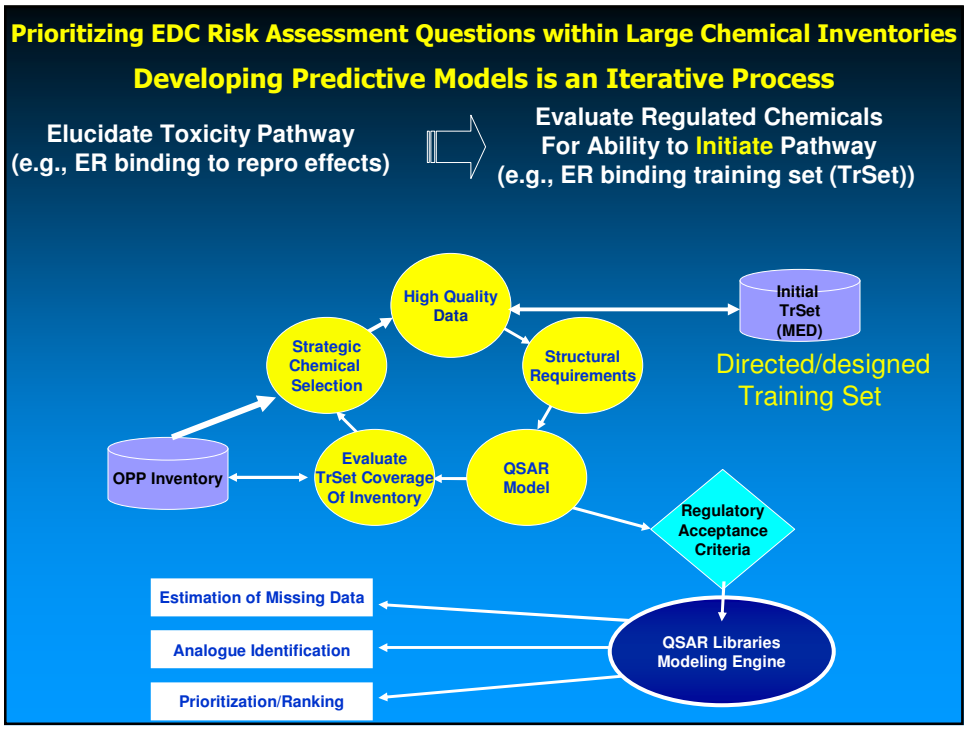
2) Defining a regulatory domain is not a trivial exercise

- **Assumption that ~6000 HPVCs would represent additional regulatory domains needs careful evaluation; regulatory lists need to be defined**
- **Structure verification is needed for all chemicals on regulatory lists**

3) Determining coverage of regulatory domain is non-trivial

- **Using a TrSet of "found" data (which included few chemicals structures found in regulatory domain) proved to be inadequate to complete QSAR development**
- **QSAR development is an iterative process that requires systematic testing within regulatory domain of interest**


Prioritizing EDC Risk Assessment Questions within Large Chemical Inventories
Developing Predictive Models is an Iterative Process
Elucidate Toxicity Pathway
(e.g., ER binding to repro effects)
Evaluate Regulated Chemicals
For Ability to Initiate Pathway
(e.g., ER binding training set (TrSet))
High Quality Data
Initial TrSet (CERI/RAL)
Strategic Chemical Selection
Structural Requirements
Undefined Chemical Inventory
Evaluate TrSet Coverage Of Inventory
QSAR Model
Regulatory Acceptance Criteria
Estimation of Missing Data
Analogue Identification
Prioritization/Ranking
QSAR Libraries Modeling Engine


Prioritizing EDC Risk Assessment Questions within Large Chemical Inventories
Developing Predictive Models is an Iterative Process
Elucidate Toxicity Pathway
(e.g., ER binding to repro effects)
Evaluate Regulated Chemicals
For Ability to Initiate Pathway
(e.g., ER binding training set (TrSet))
High Quality Data
Initial TrSet (MED)
Directed/designed Training Set
Strategic Chemical Selection
Structural Requirements
OPP Inventory
Evaluate TrSet Coverage Of Inventory
QSAR Model
Regulatory Acceptance Criteria
Estimation of Missing Data
Analogue Identification
Prioritization/Ranking
QSAR Libraries Modeling Engine

HOW to test?

High quality data is critical

- **Assays should be optimized to determine the adequacy for the types of chemicals on the relevant regulatory list**
  - **Test assays on low potency chemicals**
  - **Test to solubility**

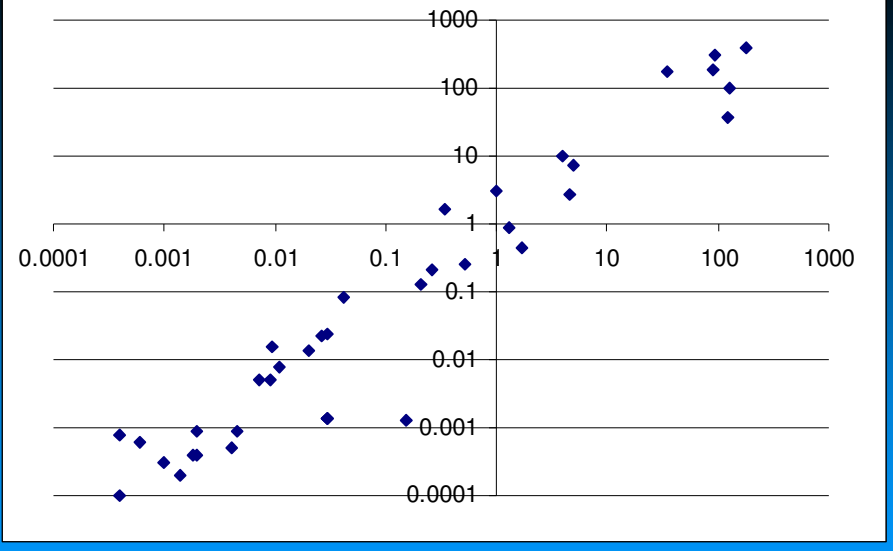
MED Database

Focus on Molecular Initiating Event

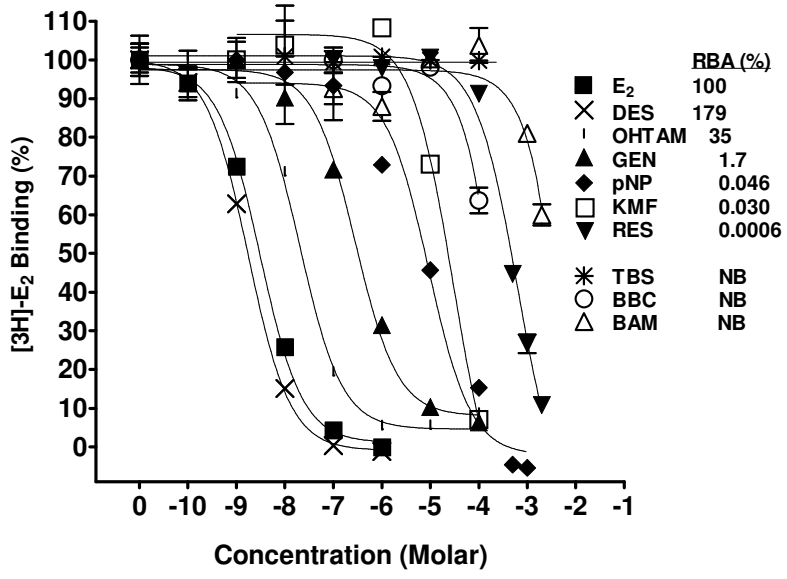
1) rtER binding is assessed using a standard competitive binding assay;

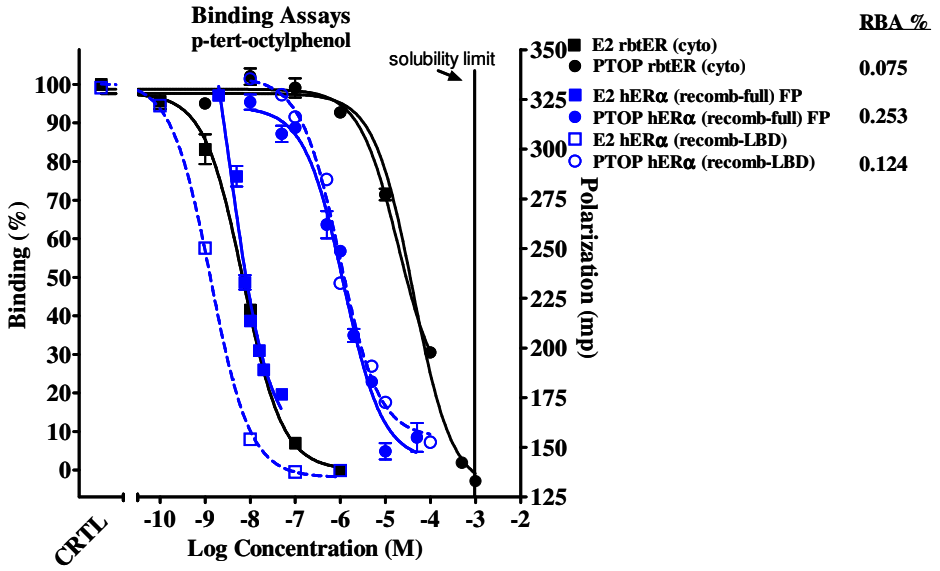
-chemicals are tested to compound solubility limit in the assay media;

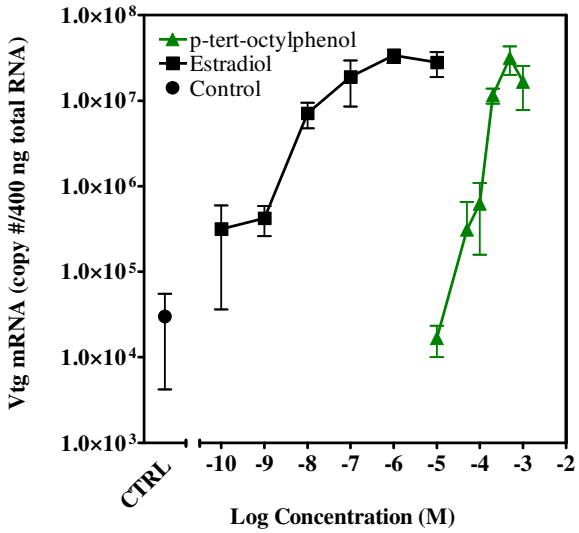
2) equivocal binding curves are interpreted using a higher-order assay (gene activation and vitellogenin mRNA production in metabolically competent trout liver slices)



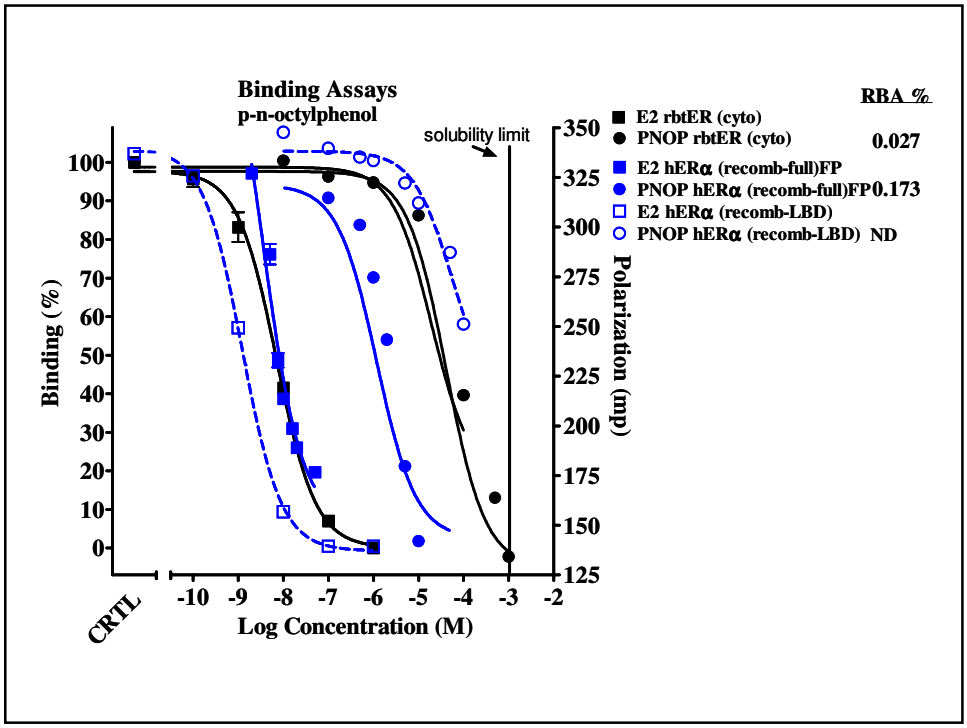
rat ER vs rainbow trout ER for 55 chemicals

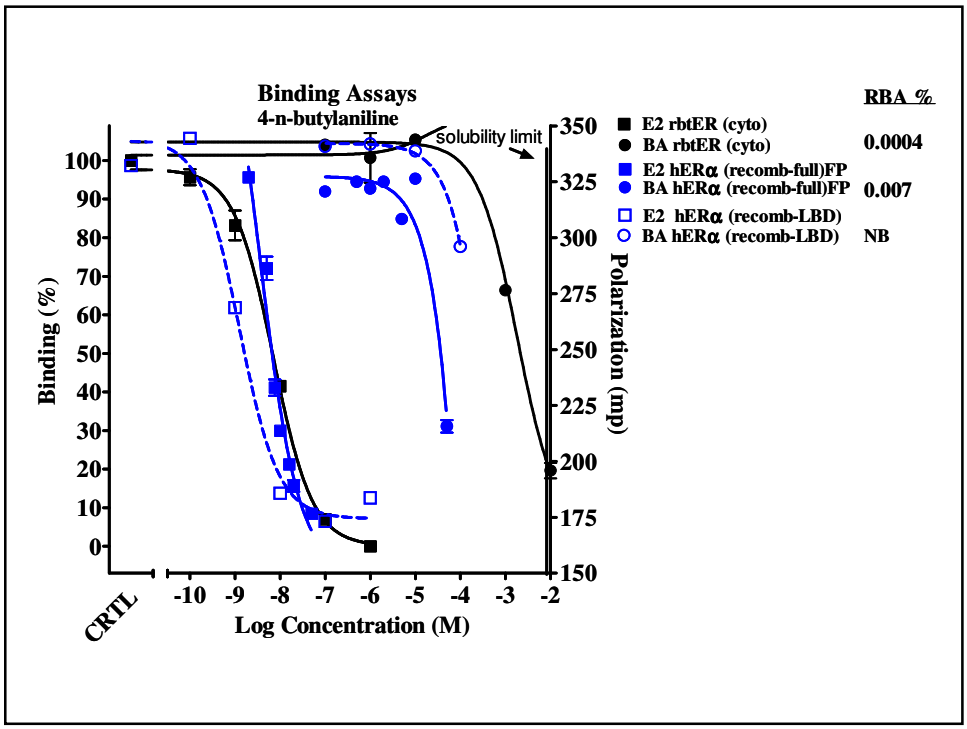




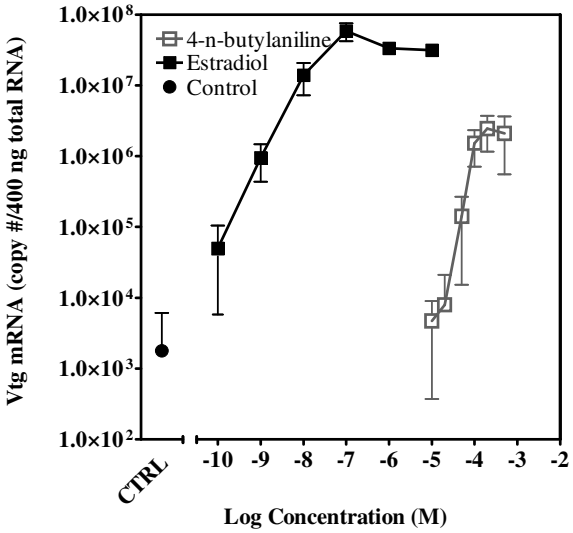


Concentration dependent vitellogenin (VTG) gene expression as VTGmRNA production in male rainbow trout liver slices exposed to **p-t-octylphenol** for 48 hrs (Mean ± STDS, n=5).


Binding Assays
p-n-octylphenol
solubility limit
Binding (%)
Polarization (mp)
Log Concentration (M)
CRTL
RBA %
E2 rbtER (cyto)
PNOP rbtER (cyto) 0.027
E2 hERα (recomb-full)FP
PNOP hERα (recomb-full)FP 0.173
E2 hERα (recomb-LBD)
PNOP hERα (recomb-LBD) ND

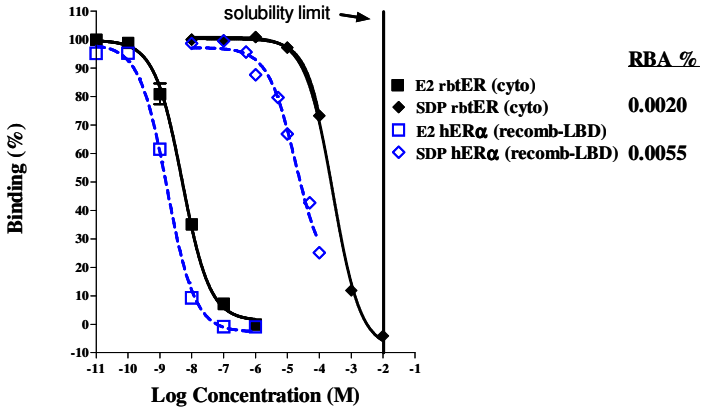

Binding Assays
4-n-butylaniline
solubility limit
Binding (%)
Polarization (mp)
Log Concentration (M)
CRTL
RBA %
E2 rbtER (cyto)
BA rbtER (cyto) 0.0004
E2 hERα (recomb-full)FP
BA hERα (recomb-full)FP 0.007
E2 hERα (recomb-LBD)
BA hERα (recomb-LBD) NB

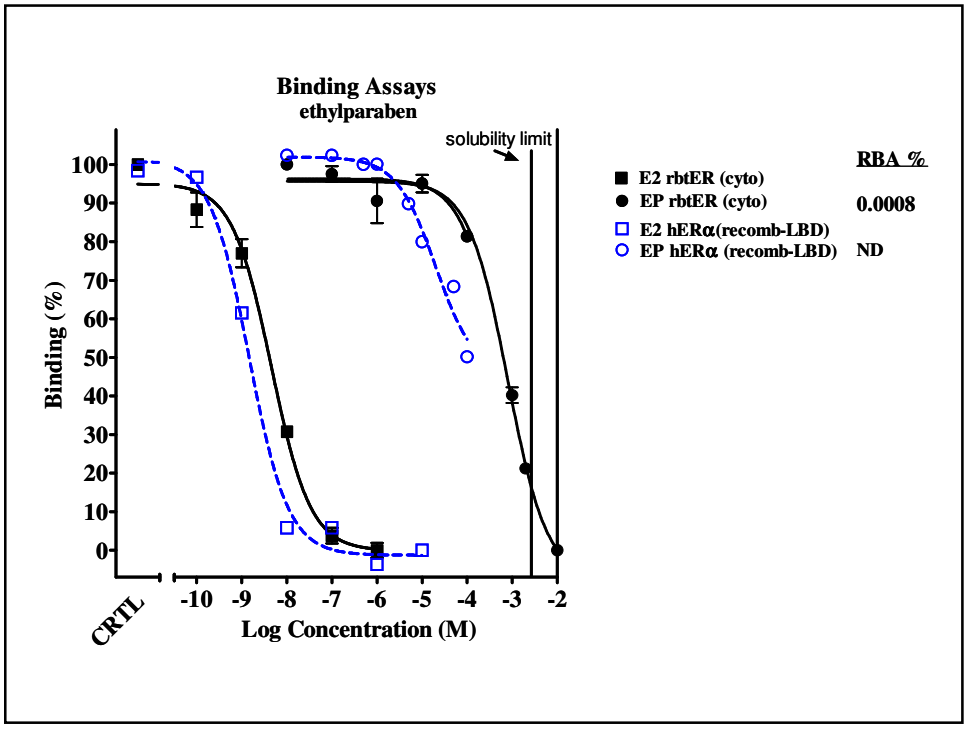
VTG Gene Activation

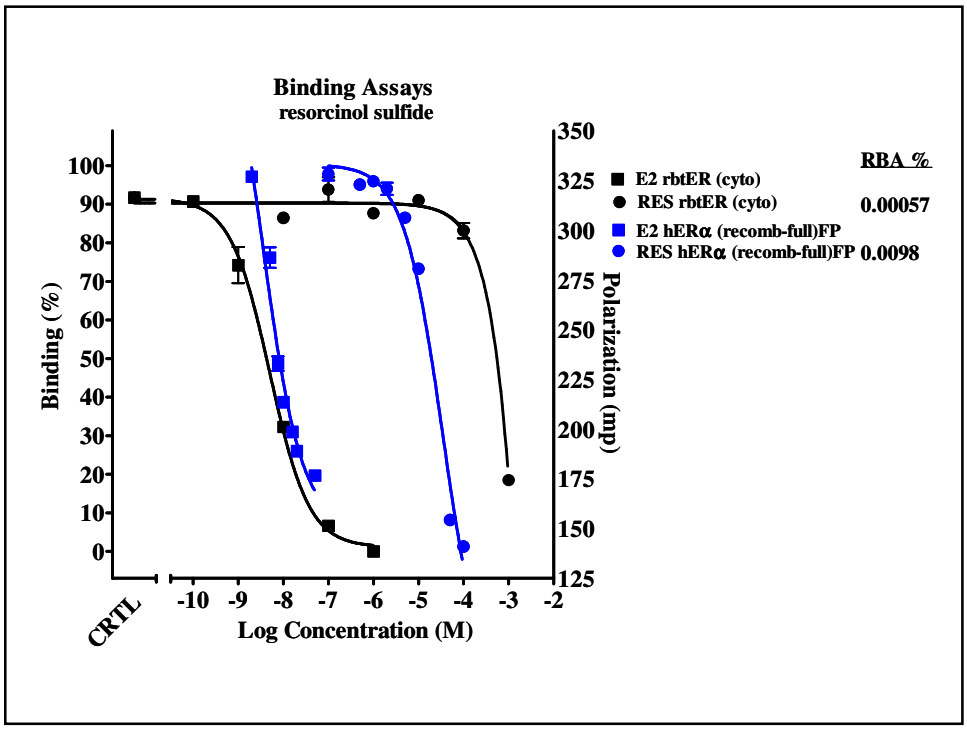


4-n-butylaniline
(Mean ± STDS, n=5)

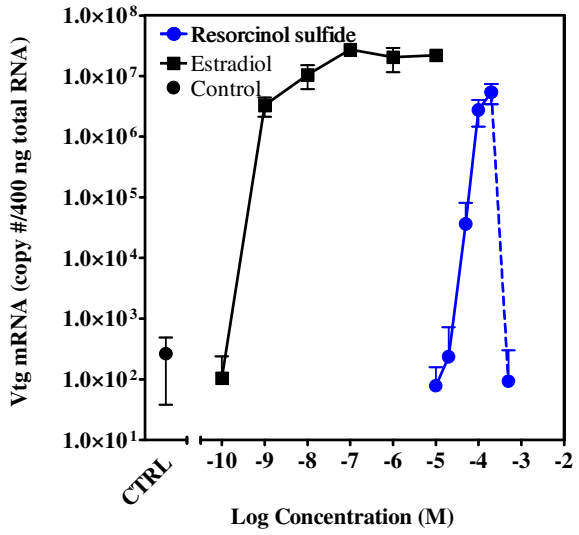
Binding Assays
4,4'-sulfonyldiphenol




Binding Assays
ethylparaben
solubility limit
Binding (%)
Log Concentration (M)
CRTL
E2 rbtER (cyto)
EP rbtER (cyto)
E2 hERα(recomb-LBD)
EP hERα (recomb-LBD)
RBA %
0.0008
ND


Binding Assays
resorcinol sulfide
Binding (%)
Polarization (mp)
Log Concentration (M)
CRTL
E2 rbtER (cyto)
RES rbtER (cyto)
E2 hERα (recomb-full)FP
RES hERα (recomb-full)FP
RBA %
0.00057
0.0098

VTG Gene Activation



resorcinol sulfide

(Mean ± STDS, n=5; dashed line indicates toxic concentrations).

WHAT to test?

Data collected needs to address the problem

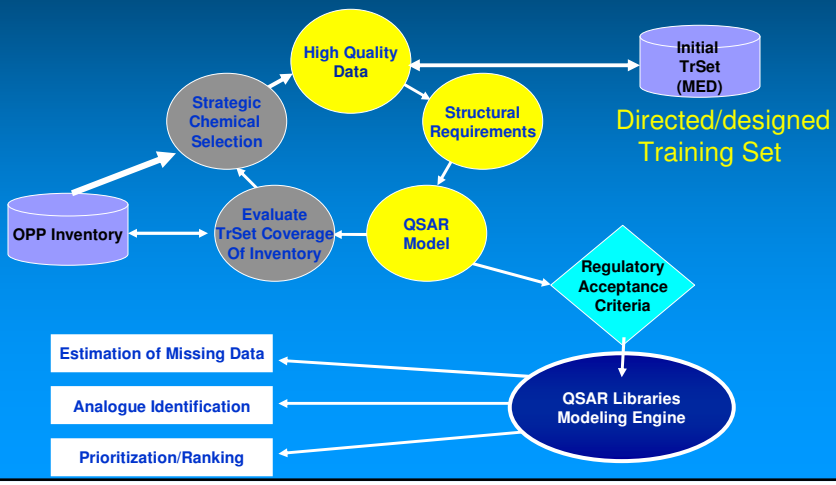
- **Expand training set to cover types of chemicals on the relevant regulatory lists**

Prioritizing EDC Risk Assessment Questions within Large Chemical Inventories

Developing Predictive Models is an Iterative Process

Elucidate Toxicity Pathway
(e.g., ER binding to repro effects)

Evaluate Regulated Chemicals
For Ability to Initiate Pathway
(e.g., ER binding training set (TrSet))



2) Defining a regulatory domain is not a trivial exercise

3) Determining coverage of regulatory domain is non-trivial

- Using a TrSet of "found" data (which included few chemicals structures found in regulatory domain) proved to be inadequate to complete QSAR development
- QSAR development is an iterative process that requires systematic testing within regulatory domain of interest

Define the Problem:
Food Use Pesticide Inerts

List included:

937 entries
-(36 repeats + 8 invalid CAS#)
893 entries

893 entries = 393 discrete + 500 non-discrete substances
(44% discrete : 56% non-discrete)

393 discrete chemicals include:
- organics
- inorganics
- organometallics

500 non-discrete substances include:
- 147 polymers of mixed chain length
- 170 mixtures
- 183 undefined substances

OPP Chemical Inventories

Chemical Category	Total	Discrete	Defined Mixtures	Polymers	Undefined Substance
Food Use Inerts	893	393	170	147	183
Antimicrobials	224	169	27	6	22
Sanitizers	104	69	10	19	6
Antimicrobials + Sanitizers	299	211	35	25	28
HPV IUR 2002	2708	1605	284	50	769
Total Inerts* (OPP website, Aug 2004)	2891	1462	155	579	695
Registered Pesticide Active Ingredients*	1110	873	33	10	194

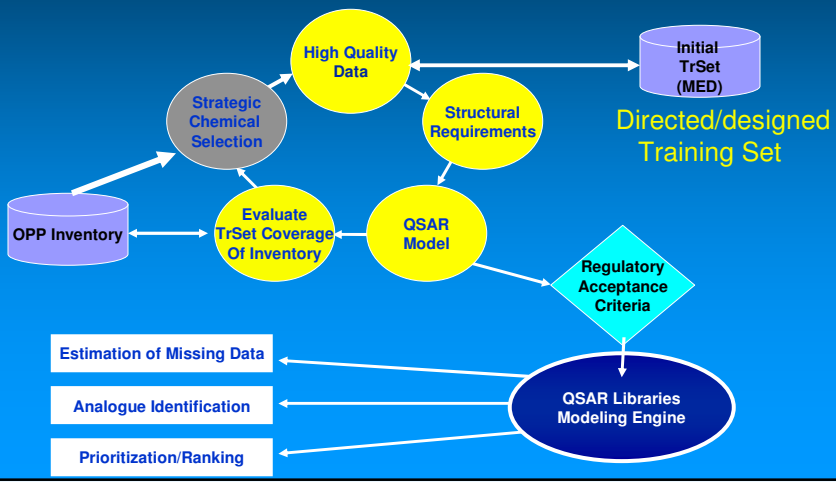
* Structure verification in progress

Prioritizing EDC Risk Assessment Questions within Large Chemical Inventories

Developing Predictive Models is an Iterative Process

Elucidate Toxicity Pathway
(e.g., ER binding to repro effects)

Evaluate Regulated Chemicals
For Ability to Initiate Pathway
(e.g., ER binding training set (TrSet))



Original ER Binding Training Sets

- Initial focus of ER binding data sets from 1990s - 2004:
  - Steroids, anti-estrogens *(high potency binders)*
  - Organochlorines
  - Alkylphenols

	CERI hER	NCTR rER	MED rtER	Food Use Inerts	Anti-microbial	HPV Inerts	HPV TSCA
Steroid, Anti-E2, OrganoCl	150 (30%)	91 (40%)	37	2 (<1%)	2 (1%)	6 (1%)	178 (3%)
Alkyl-phenols	35 (7%)	13 (6%)	22	3 (1%)	7 (3%)	6 (1%)	71 (1%)
Covered groups as % of total	37%	46%		2%	4%	2%	4%

Building New Training Sets

- New inventories
  - Food Use Inerts
  - Antimicrobials and Sanitizers
  - HPV inerts
  - Total Inerts
  - HPV TSCA chemicals

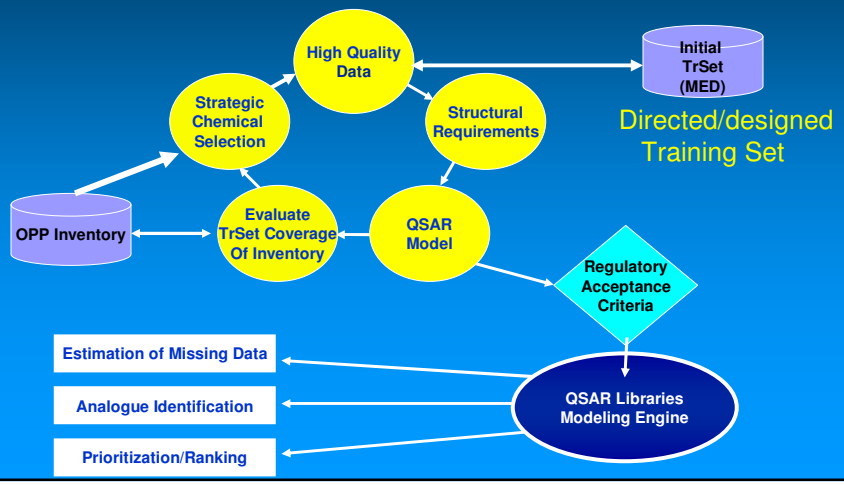
	CERI (hER)	NCTR (rER)	ORD-MED (rtER)	Food Use Inerts	A/S	HPV Inerts	HPV TSCA
Acyclics	3 (0.6%)	6 (2.6%)	22 (10%)	230 (59%)	121 (57%)	291 (65%)	2655 (41%)
Aromatic Sulfates	4 (0.8%)	1 (0.4%)	15	88 (22%)	6 (3%)	15 (3%)	347 (5%)

Prioritizing EDC Risk Assessment Questions within Large Chemical Inventories

Developing Predictive Models is an Iterative Process

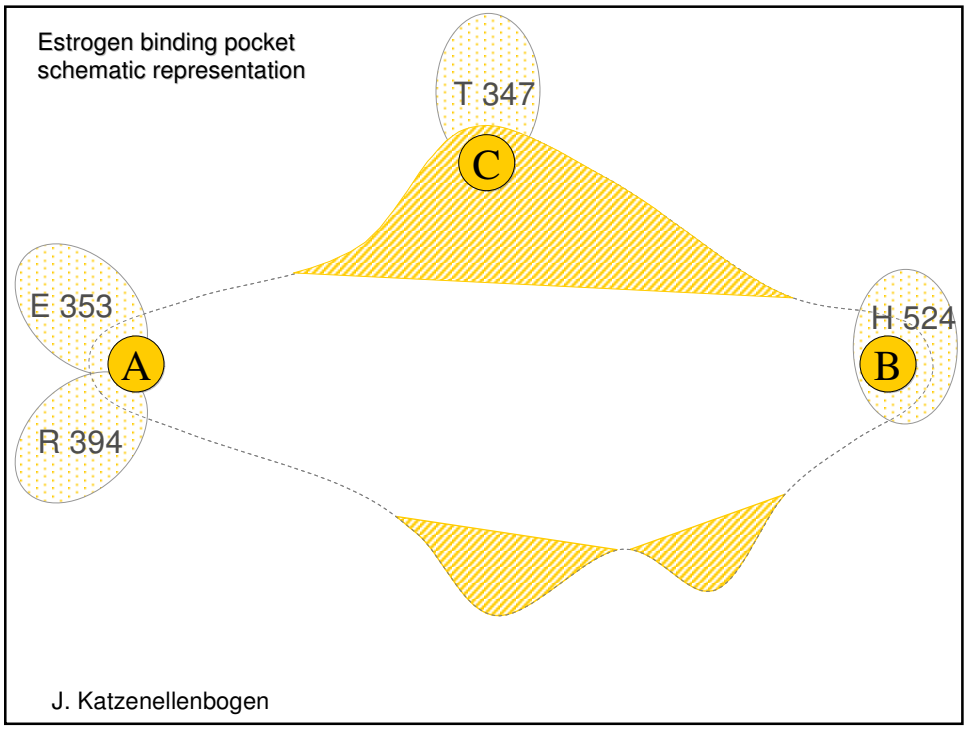
Elucidate Toxicity Pathway
(e.g., ER binding to repro effects)

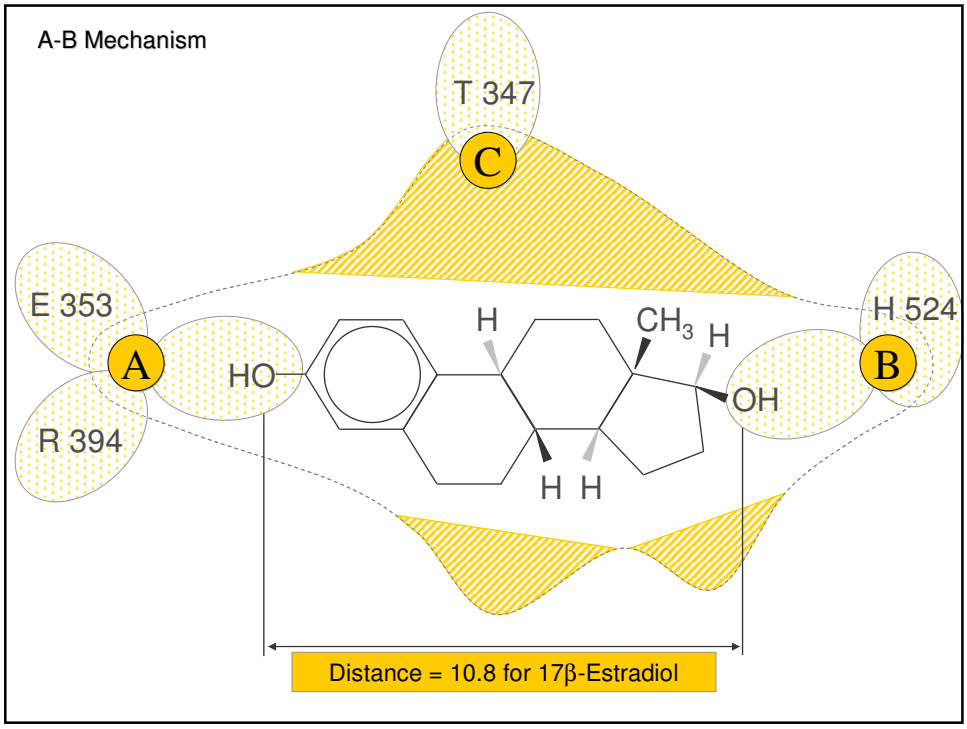
Evaluate Regulated Chemicals
For Ability to Initiate Pathway
(e.g., ER binding training set (TrSet))

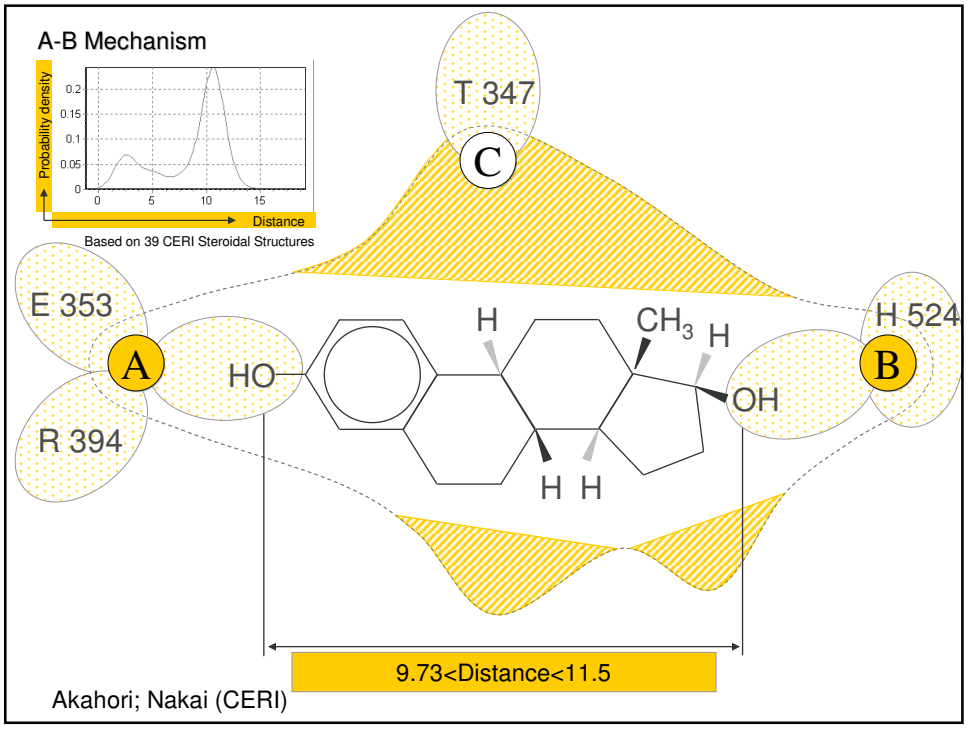


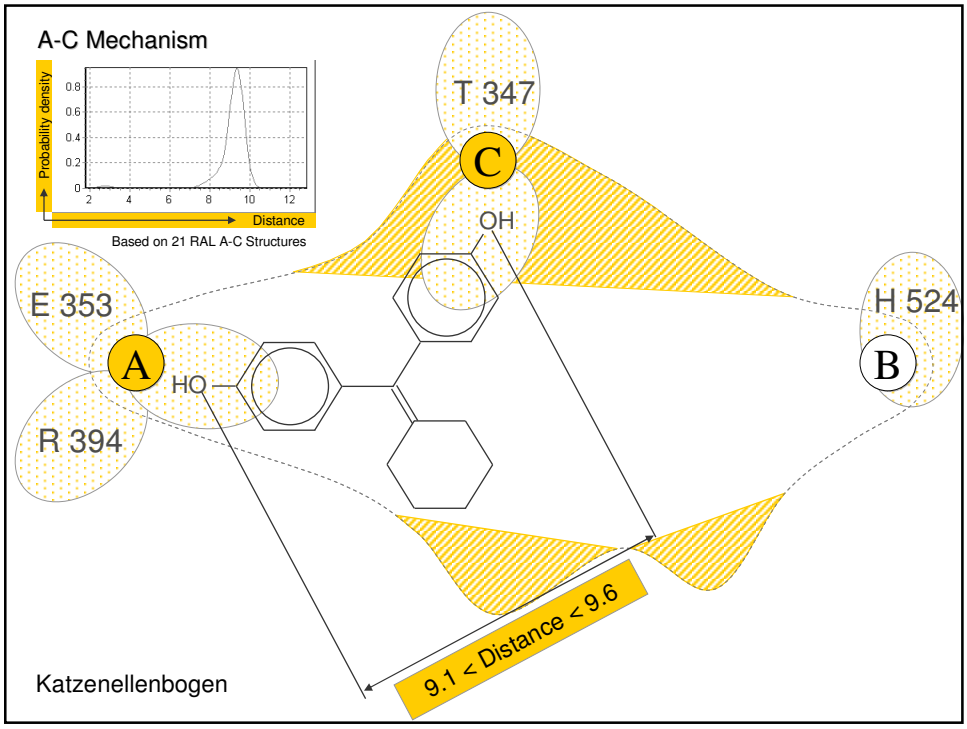
QSAR Principles for ER interactions

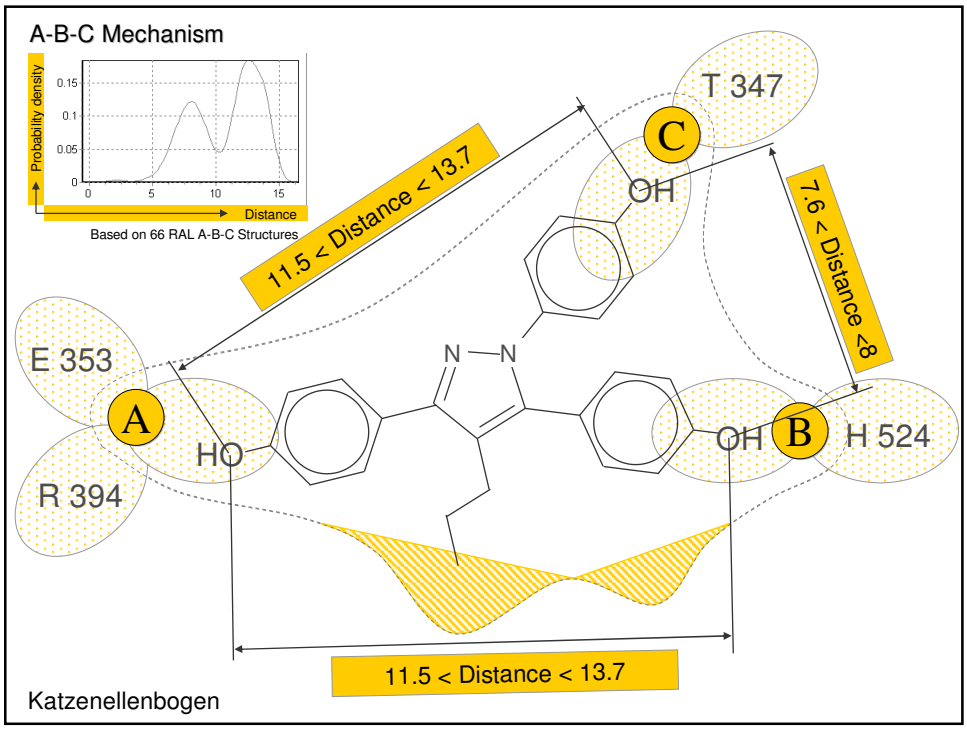
- Chemical are “similar” if they produce the same biological action from the same initiating event
  - Not all chemicals bind ER in same way, i.e., not all “similar”
  - ER binders are “similar” if they have the same type of interaction within the receptor
- QSARs require a well-defined/well understood biological system; assay strengths and limitations understood
- QSARs for large list of diverse chemicals
  - require iterative process – test, hypothesize, evaluate, new hypothesis, test again, etc.
  - to gain mechanistic understanding to group similar acting chemicals; build model within a group

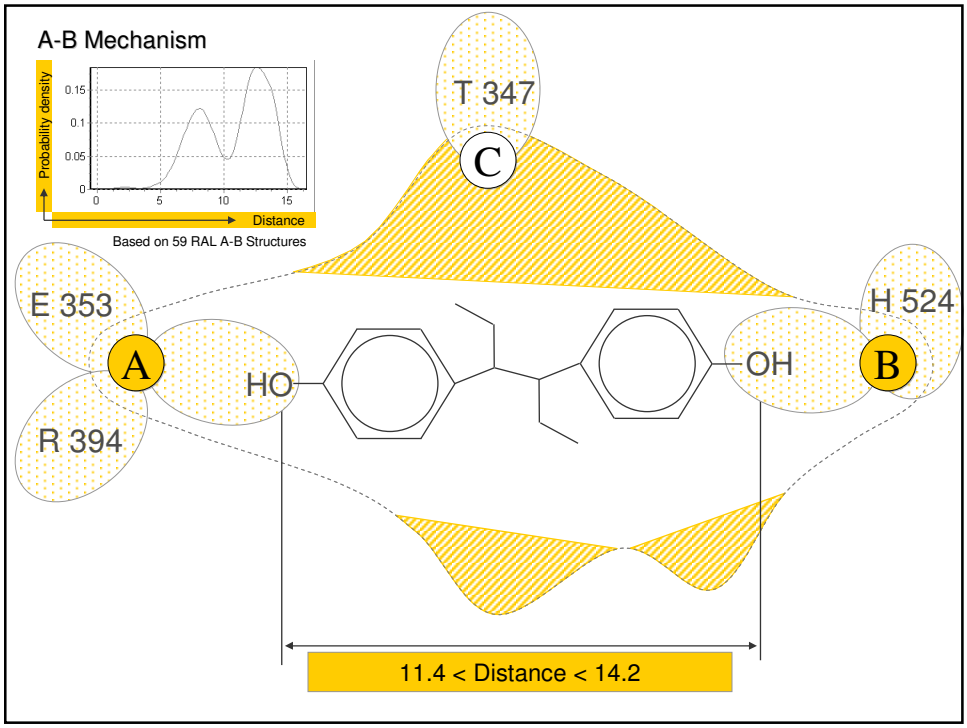



A-B Mechanism
T 347
C
E 353
A
R 394
HO
H
CH3 H
OH
H H
H 524
B
Distance = 10.8 for 17β-Estradiol

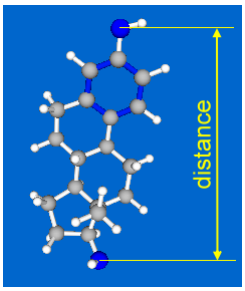

A-B Mechanism
Probability density
0.2
0.15
0.1
0.05
0
0
5
10
15
Distance
Based on 39 CERI Steroidal Structures
T 347
C
E 353
A
R 394
HO
H
CH3 H
OH
H H
H 524
B
9.73<Distance<11.5
Akahori; Nakai (CERI)


A-C Mechanism
Probability density
0.8
0.6
0.4
0.2
0
2
4
6
8
10
12
Distance
Based on 21 RAL A-C Structures
T 347
C
OH
E 353
A
HO
R 394
H 524
B
9.1 < Distance < 9.6
Katzenellenbogen


A-B-C Mechanism
Probability density
0.15
0.1
0.05
0
0
5
10
15
Distance
Based on 66 RAL A-B-C Structures
T 347
C
OH
11.5 < Distance < 13.7
7.6 < Distance <8
E 353
A
HO
N
N
R 394
OH
B
H 524
11.5 < Distance < 13.7
Katzenellenbogen



Hypothesis testing



***Hypothesis**: Chemicals with interatomic distance between O-atoms satisfying distance criteria for a binding type have the potential to bind ER based on electronic interactions.*

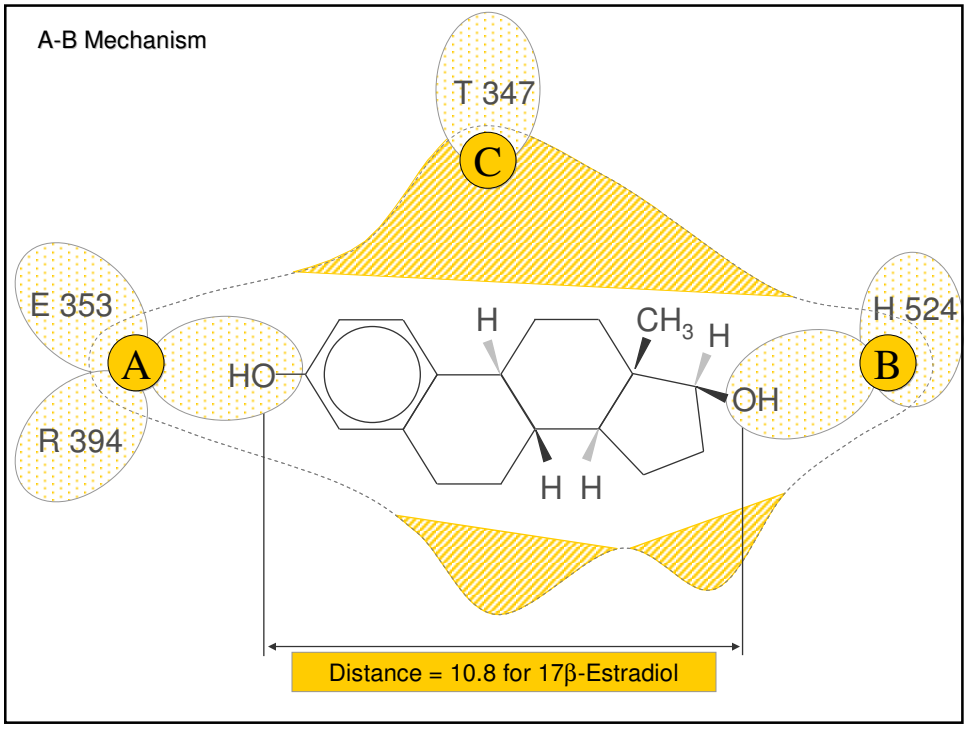
- *Hypothesize structural parameter(s) associated with toxicity*
- *Select chemicals that satisfy the hypothesis*
- *Test, and confirm or modify hypothesis*

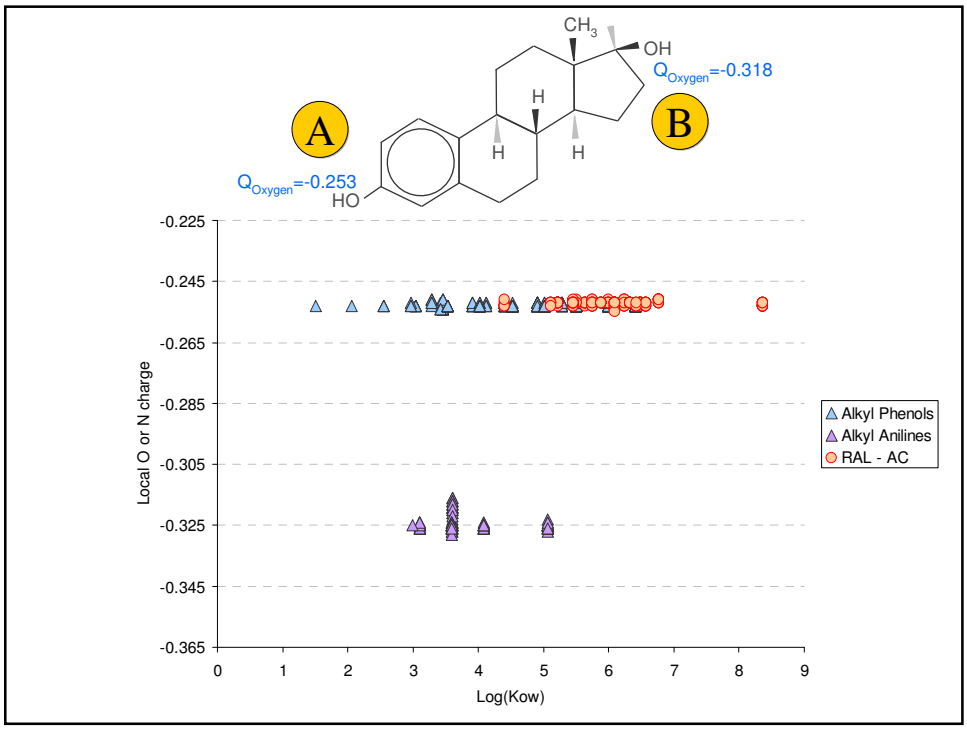
- Because acyclics are > 50% of inventories, what is the possibility that any acyclics satisfy criteria of high affinity binding types?
- Selected acyclics for testing that met A_B distance; no binders found (charged cmpds – apparent binding but no activation)
- As suspected, most OPP chemicals could not be evaluated with the A_B or A_C mechanism models;
- Need to refine ER binding hypotheses to investigate additional binding types
  - Chemicals interact with ER in more than one way, influencing data interpretation and model development;
  - Need to group chemicals by like activity, then attempt to model as a group that initiate action through same chemical-biological interaction mechanism, and should have common features
  - Find common features and predict which other untested chemicals may have similar activity – prioritize for testing

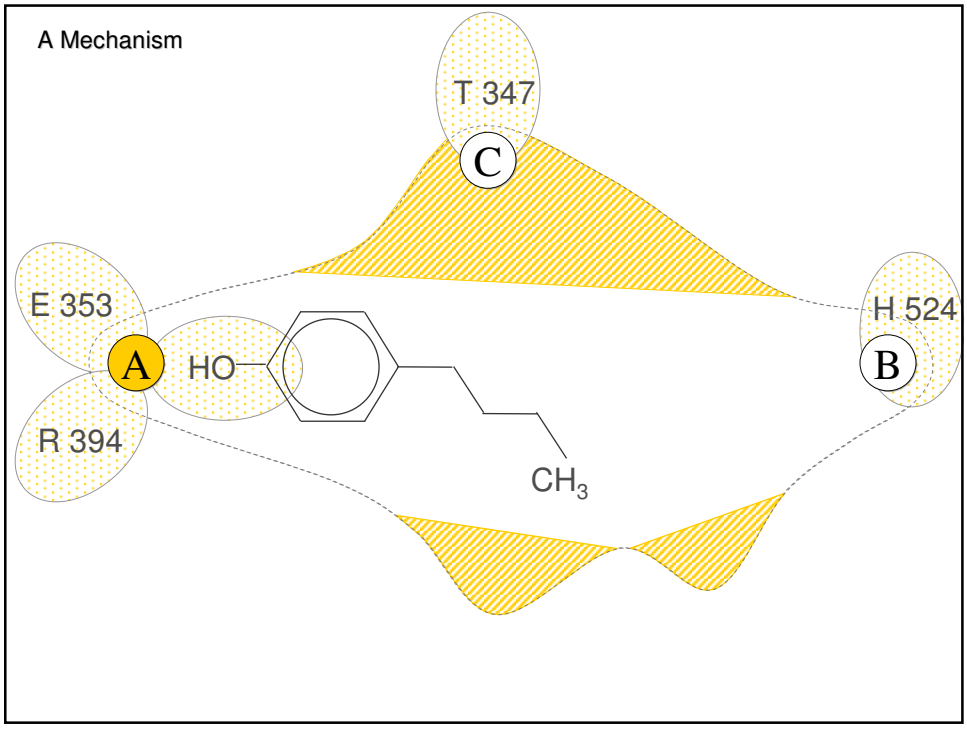
HOW to interpret test results?

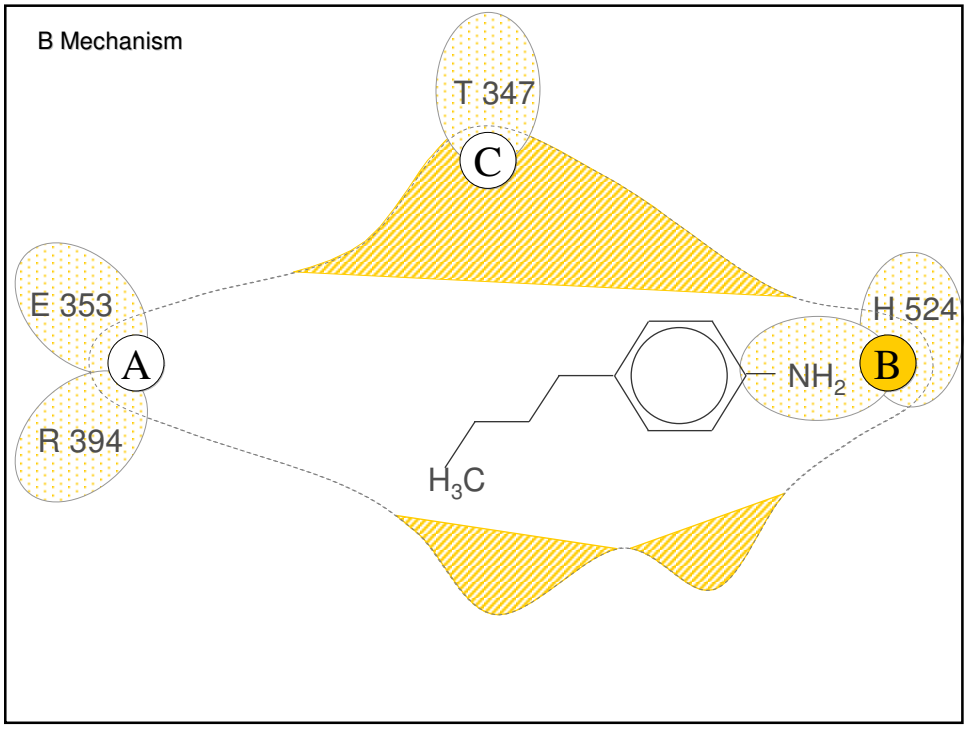
High quality data is critical

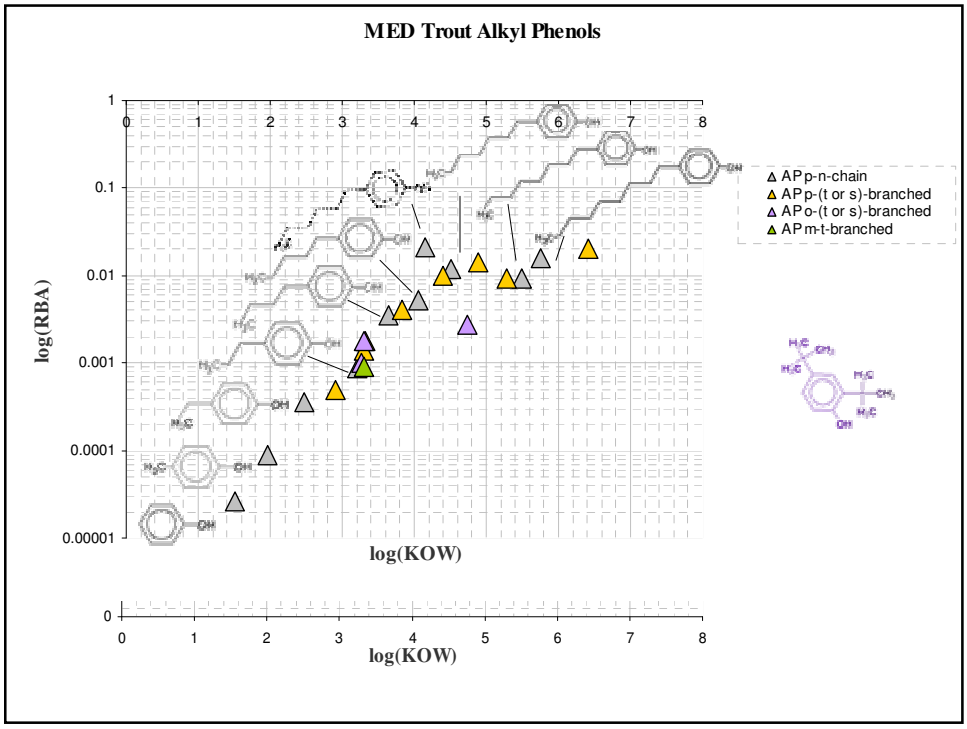
- **ER binding hypotheses refined**
  - **Chemicals interact with ER in more than one way, influencing data interpretation and model development**

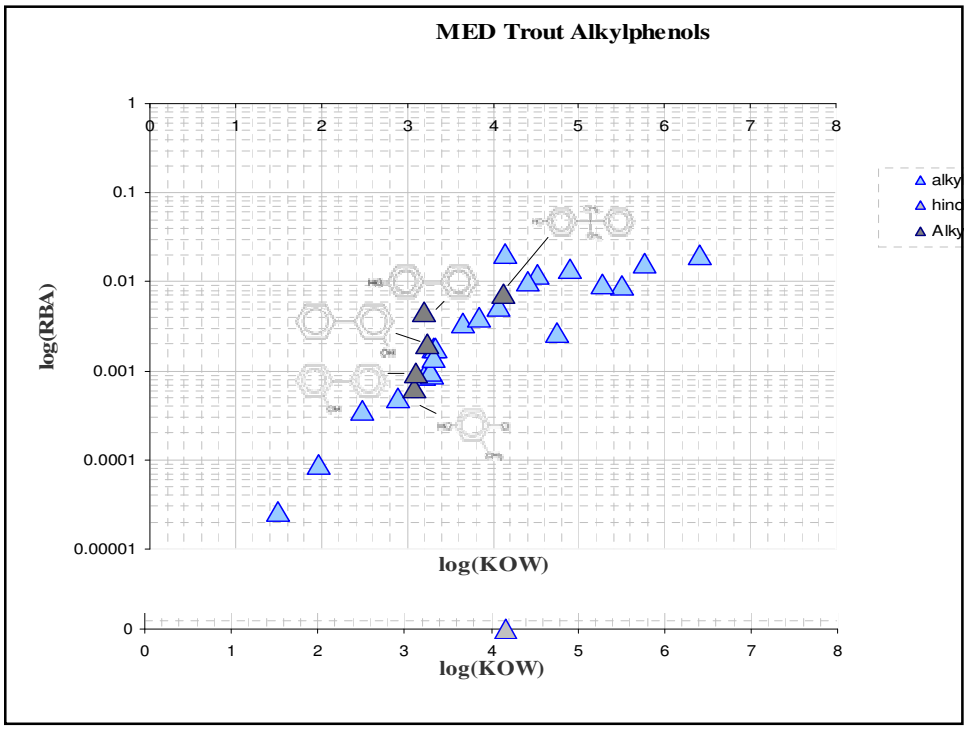

A-B Mechanism
T 347
C
E 353
A
R 394
HO
H
CH3 H
OH
H H
H 524
B
Distance = 10.8 for 17β-Estradiol

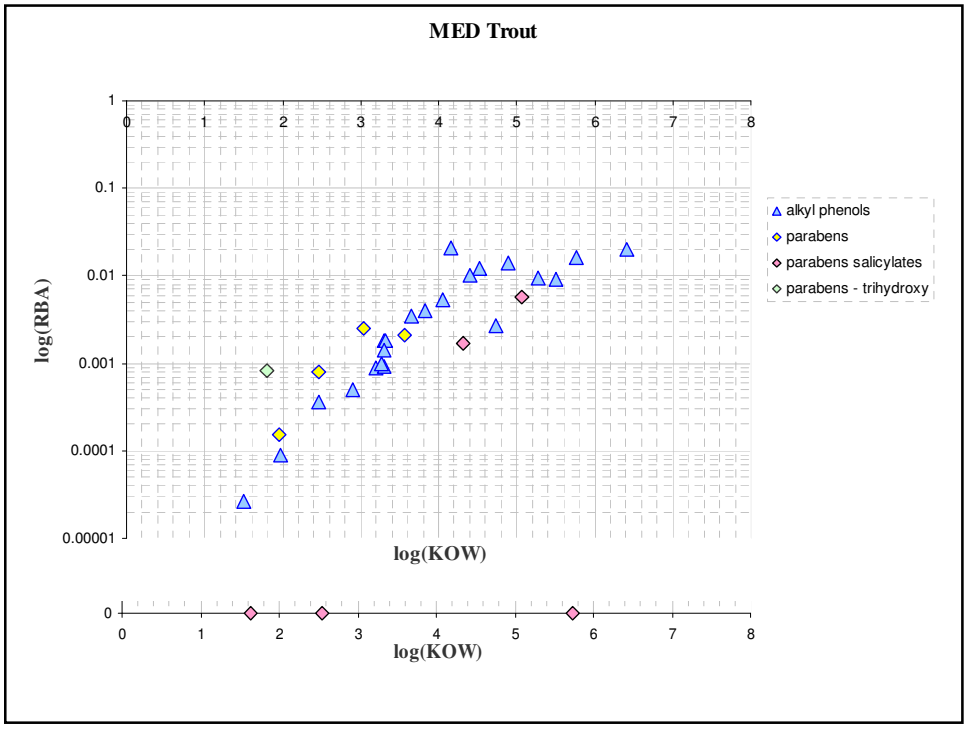

CH3
OH
QOxygen=-0.318
A
B
H
H
H
QOxygen=-0.253
HO
-0.225
-0.245
-0.265
-0.285
-0.305
-0.325
-0.345
-0.365
Local O or N charge
0
1
2
3
4
5
6
7
8
9
Log(Kow)
Alkyl Phenols
Alkyl Anilines
RAL - AC

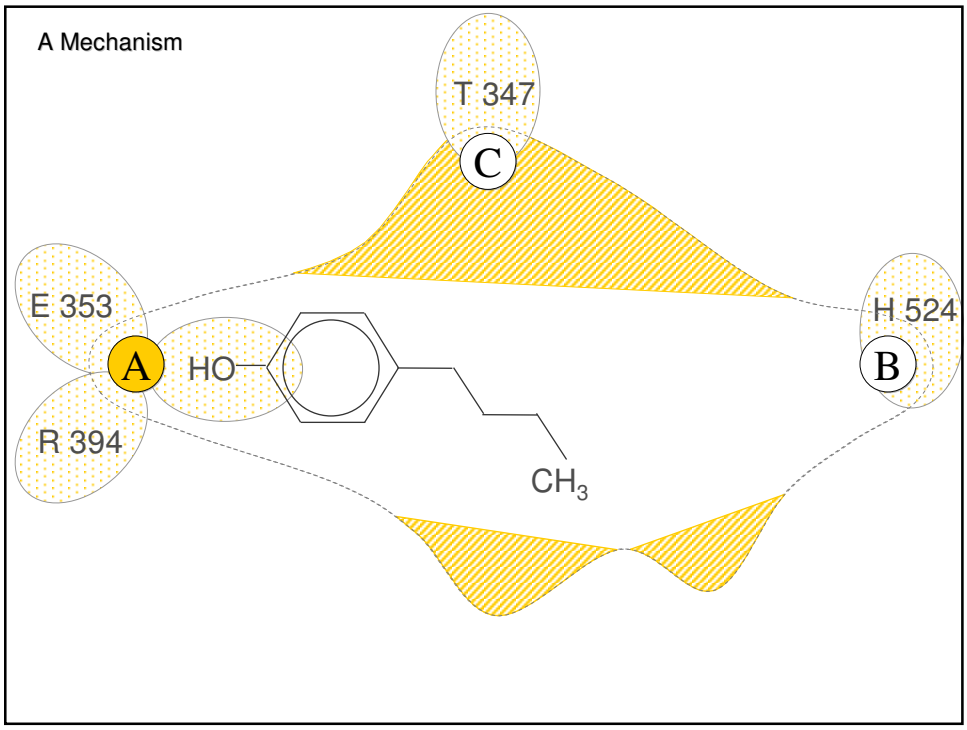

A Mechanism
T 347
C
E 353
A
HO
R 394
H 524
B
CH3

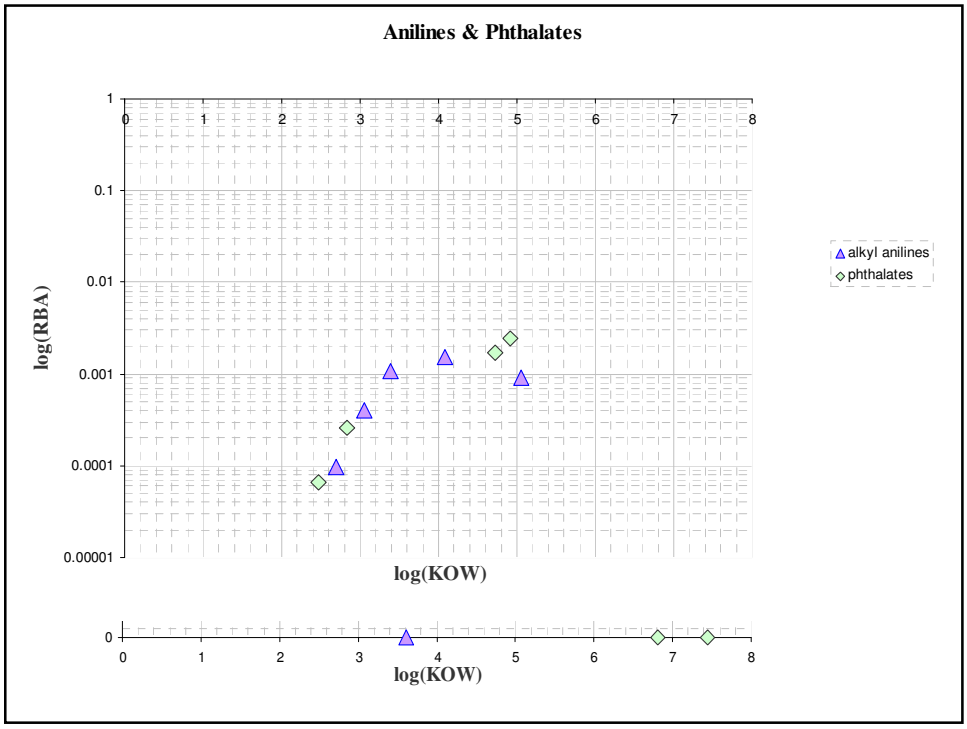

B Mechanism
T 347
C
E 353
A
R 394
H 524
B
NH2
H3C

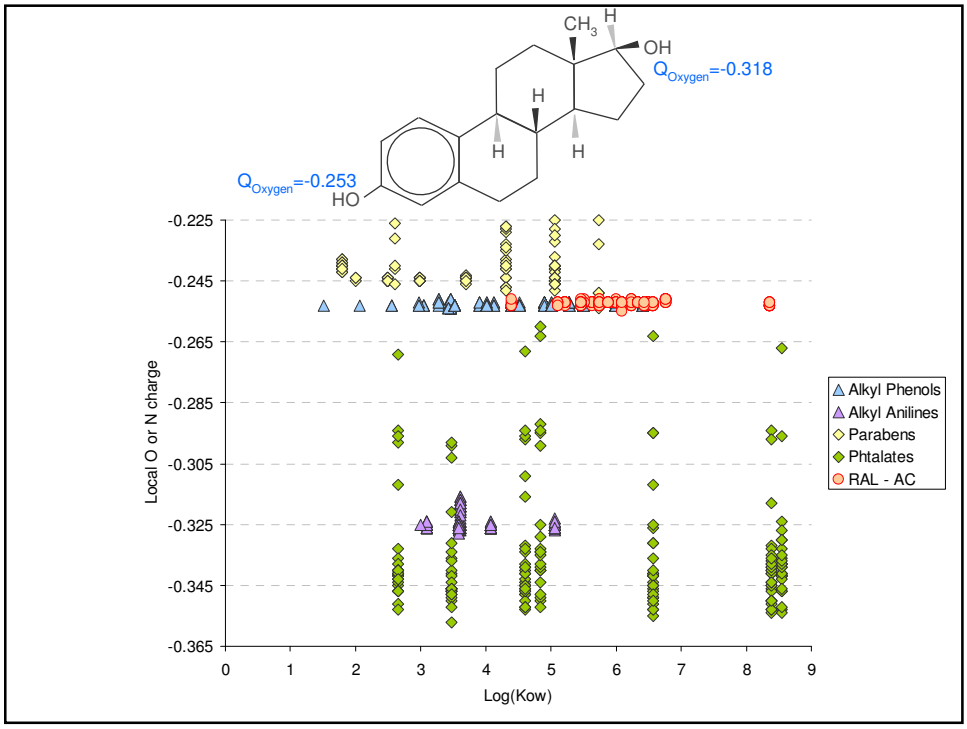

MED Trout Alkyl Phenols
log(RBA)
log(KOW)
log(KOW)
APp-n-chain
APp-(t or s)-branched
APo-(t or s)-branched
APm-t-branched

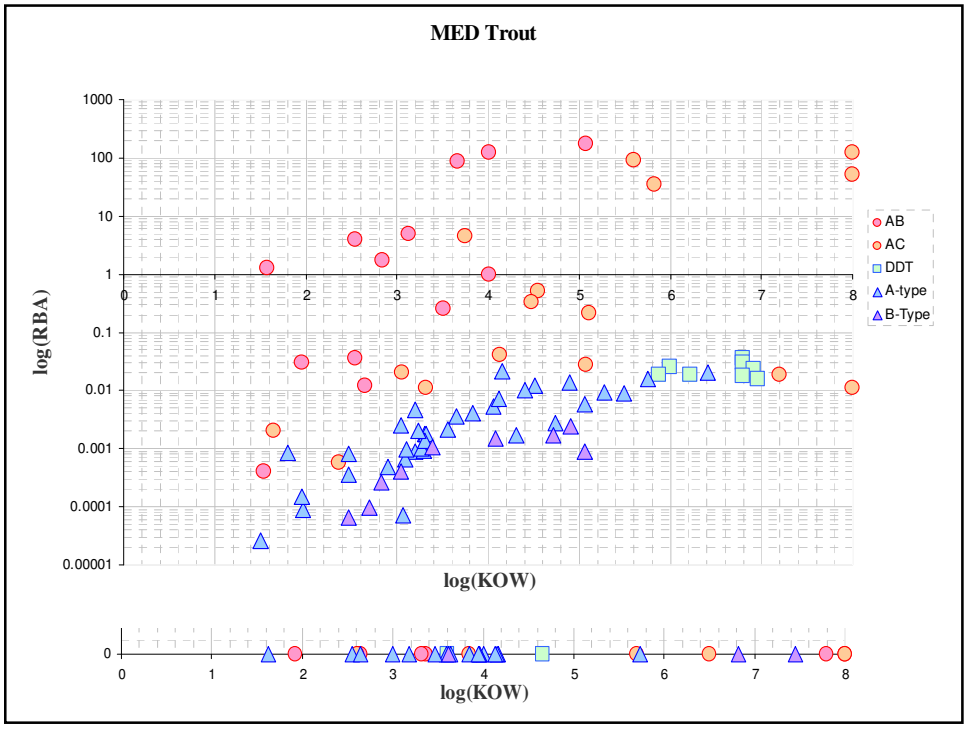

MED Trout Alkylphenols
log(RBA)
log(KOW)
log(KOW)

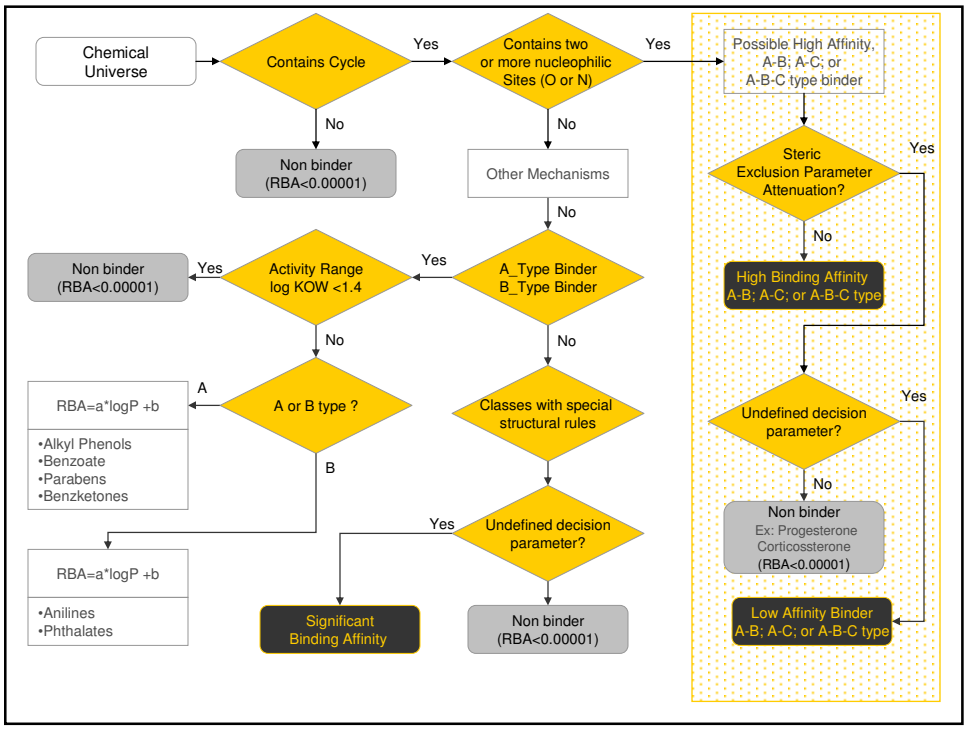

MED Trout
log(RBA)
log(KOW)
log(KOW)
alkyl phenols
parabens
parabens salicylates
parabens - trihydroxy

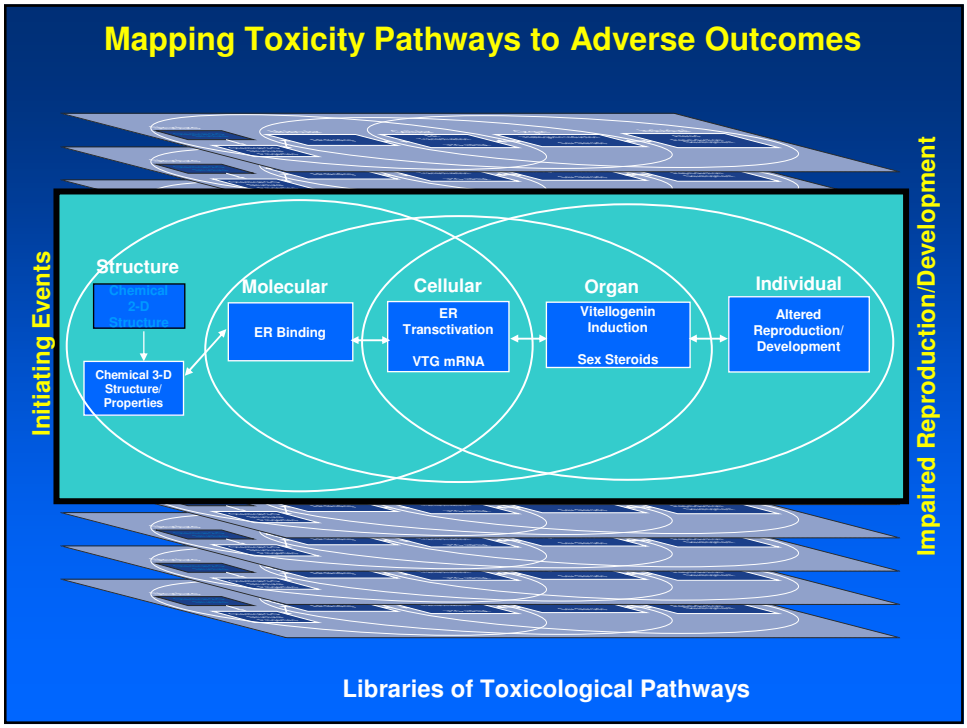

A Mechanism
T 347
C
E 353
A
HO
R 394
H 524
B
CH3

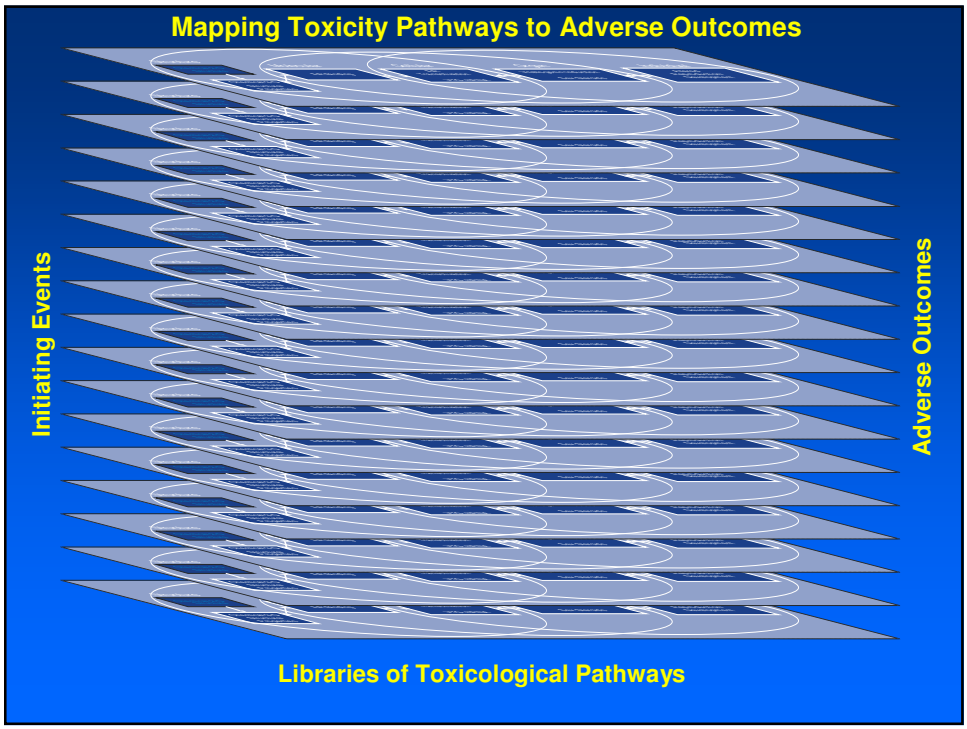

Anilines & Phthalates
log(RBA)
log(KOW)
log(KOW)
alkyl anilines
phthalates


CH3
H
OH
QOxygen=-0.318
H
H
H
QOxygen=-0.253
HO
Local O or N charge
Log(Kow)
Alkyl Phenols
Alkyl Anilines
Parabens
Phtalates
RAL - AC


MED Trout
log(RBA)
log(KOW)
log(KOW)
AB
AC
DDT
A-type
B-Type


Chemical Universe
Contains Cycle
Yes
Contains two or more nucleophilic Sites (O or N)
Yes
Possible High Affinity, A-B; A-C; or A-B-C type binder
No
Non binder (RBA<0.00001)
No
Other Mechanisms
No
Steric Exclusion Parameter Attenuation?
Yes
No
High Binding Affinity A-B; A-C; or A-B-C type
Non binder (RBA<0.00001)
Yes
Activity Range log KOW <1.4
Yes
A_Type Binder B_Type Binder
No
No
RBA=a*logP +b
•Alkyl Phenols
•Benzoate
•Parabens
•Benzketones
A
A or B type ?
Classes with special structural rules
Undefined decision parameter?
Yes
No
Non binder Ex: Progesterone Corticossterone (RBA<0.00001)
B
RBA=a*logP +b
•Anilines
•Phthalates
Yes
Undefined decision parameter?
Significant Binding Affinity
Non binder (RBA<0.00001)
Low Affinity Binder A-B; A-C; or A-B-C type


Mapping Toxicity Pathways to Adverse Outcomes
Initiating Events
Impaired Reproduction/Development
Structure
Chemical 2-D Structure
Chemical 3-D Structure/ Properties
Molecular
ER Binding
Cellular
ER Transctivation
VTG mRNA
Organ
Vitellogenin Induction
Sex Steroids
Individual
Altered Reproduction/ Development
Libraries of Toxicological Pathways


Mapping Toxicity Pathways to Adverse Outcomes
Initiating Events
Adverse Outcomes
Libraries of Toxicological Pathways

Acknowledgements:

MED – J. Denny, R. Kolanczyk, B. Sheedy, M. Tapper;
SSC – C. Peck; B. Nelson; T. Wehinger, B. Johnson; L. Toonen; R. Maciewski
NRC Post-doc: H. Aladjov
Bourgus University - LMC: O. Mekenyan, and many others
Chemicals Evaluation Research Institute (CERI), Japan - Y. Akahori, N. Nakai
EPA/NERL-Athens: J. Jones
EPA/OPP:
EFED - S. Bradbury, J. Holmes
RD - B.Shackleford, P. Wagner
AD - J. Housenger, D. Smegal
HED – L. Scarano
Mentors: G. Veith, L. Weber, and J.M. McKim, III