The Problem

Adelina Voutchkova

George Washington University
Washington, D.C.

Chemicals We Encounter

1%
Pharmaceuticals

- Designed to be biologically active
- Performance criteria include a consideration of side effects
- Produced in relatively small volumes
- Well defined use scenarios

99%
Industrial chemicals

- No intentional biological activity
- Performance is divorced from toxicity
- Can be produced in multi-billion pound quantities
- Extremely diverse use scenarios

Environmental Impacts of Chemical Industry

3.93 billion lbs. of toxic chemicals were released directly to air, land, and water in 2010, 16% increase from 2009.

US EPA TRI

Fukushima Nuclear Power Plant 2011

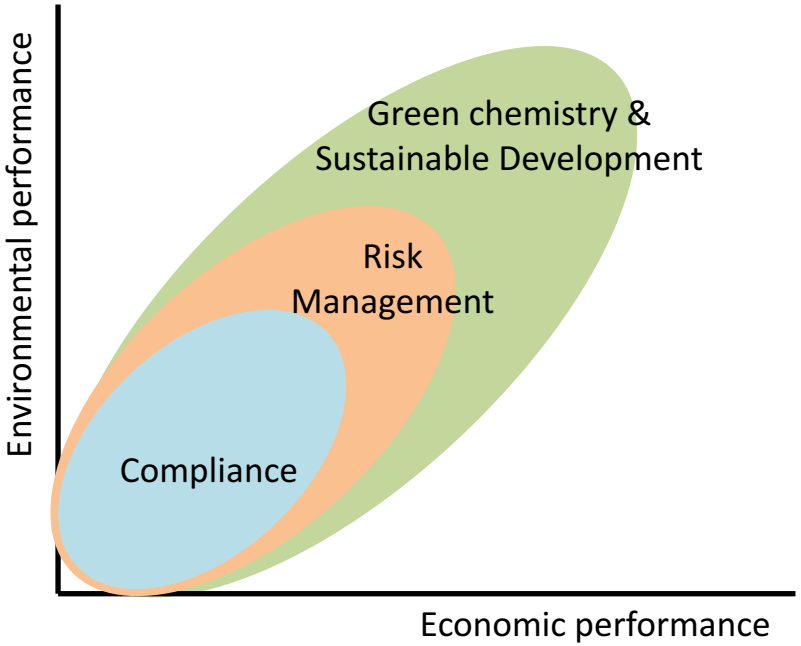
Exposure controls can and do fail

Plaquemines Parish, LA

High levels of toxic chemicals in fish cause massive fish kills

Severe weather systems and rising sea levels are providing palpable indicators of climate change

Proactive Approach

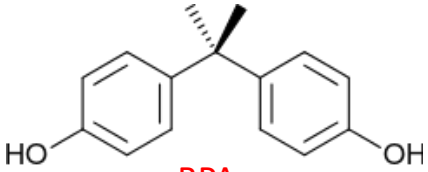


The design of chemical products and processes that reduce or eliminate the use and generation of hazardous substances

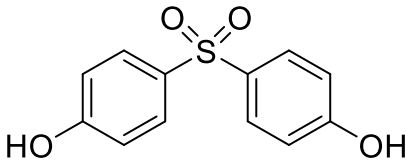
Anastas, P. and Warner, W. *Green Chemistry: Theory and Practice,* Oxford University Press: New York, 1998.

Chemical Substitutions

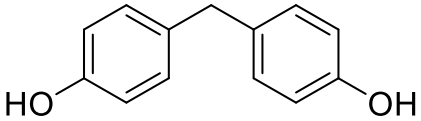
Out of the frying pan and out of the fire: replacement flame retardants now found in breast milk



BPA



BPS



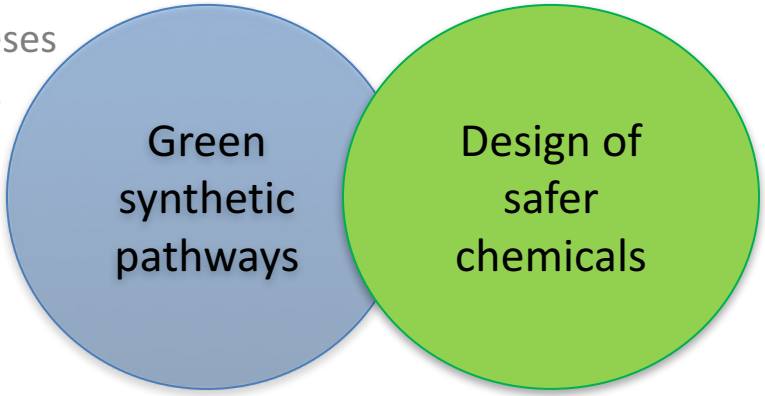
BPF

Rochester JR, Bolden AL. 2015. Environ Health Perspect 123:643–650.

Principles of Green Chemistry

the **design** of chemical products and processes that minimize the use or generation of hazardous substances.

- Waste
- Less hazardous syntheses
- Renewable feedstocks
- Catalysts
- Chemical derivatives
- Atom economy
- Safer solvents
- Energy efficiency



- Safer chemicals
- Biodegradable chemicals
- Potential for accidents

Anastas, P. and Warner, W. *Green Chemistry: Theory and Practice,* Oxford University Press: New York, 1998.

~ 90% materials' feedstocks are from petroleum

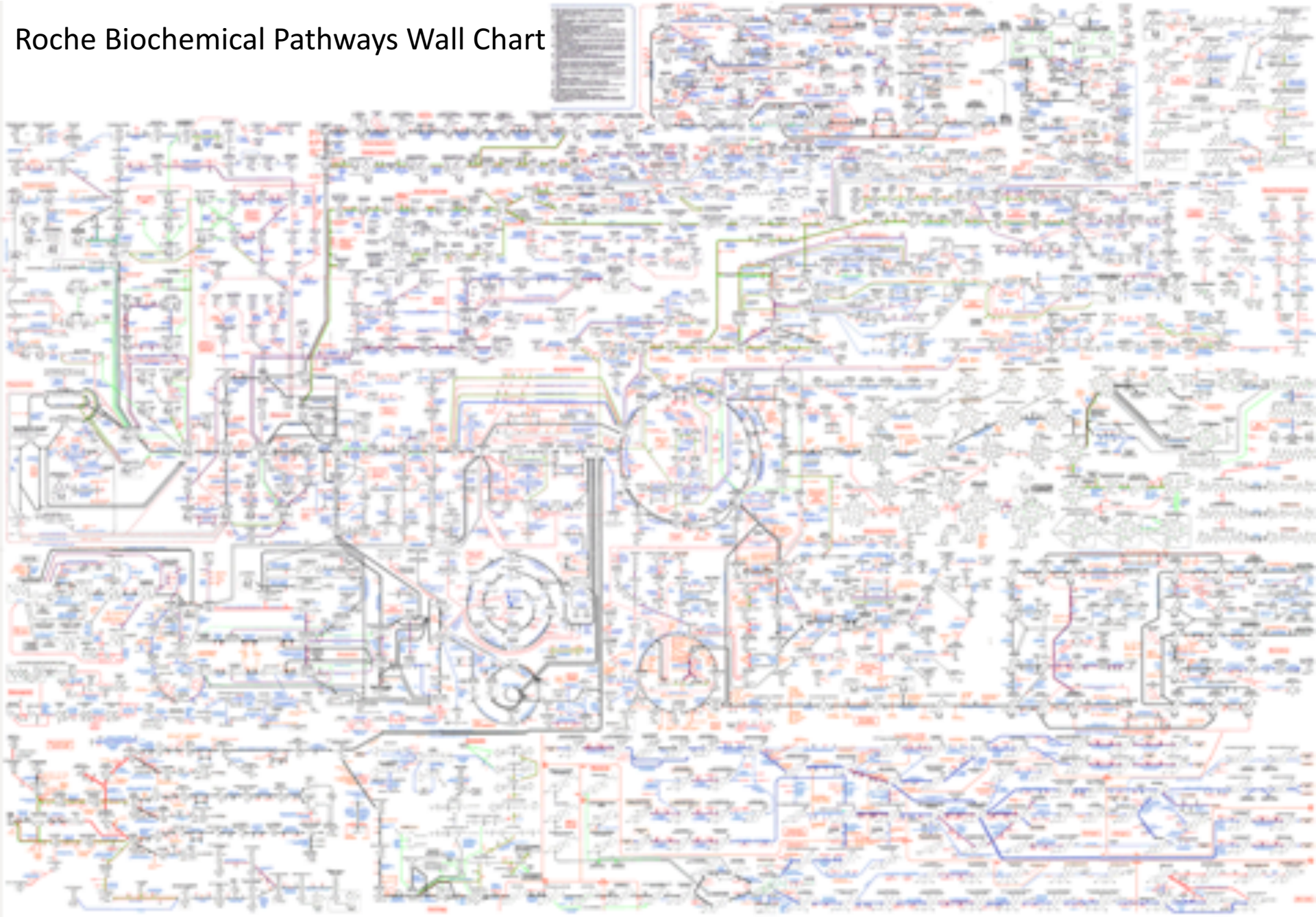
82K registered for commerce (US EPA), ca. 4% have health/safety data (e.g. MSDS), ca. 75% have not been studied at all

In practice, EPA's TSCA regulates only 5 out of 82,000+ chemicals in commerce;

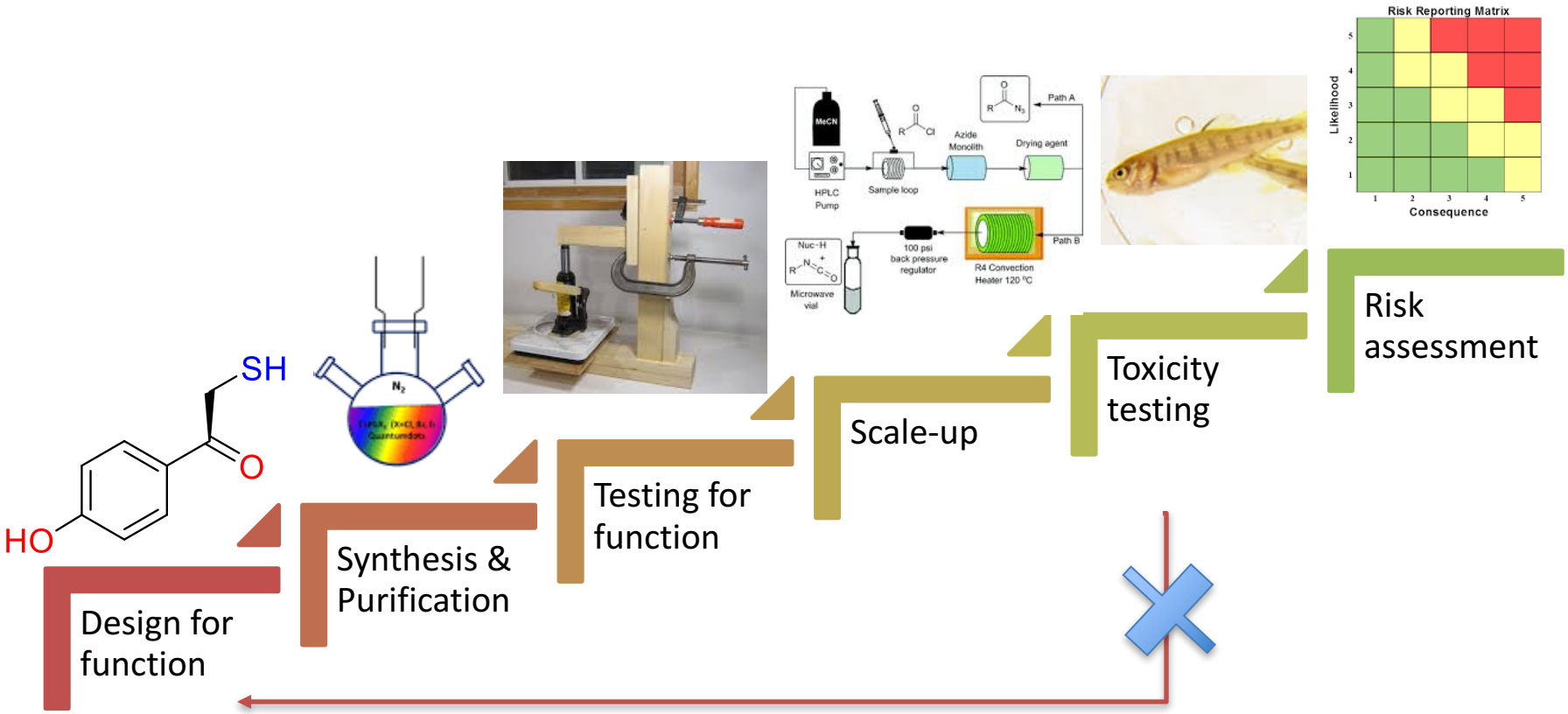
62,000 chemicals were already in commerce before 1972, are exempt from TSCA

~ 2000 new chemicals introduced to market/yr, most of which have no experimental health or safety data

Roche Biochemical Pathways Wall Chart

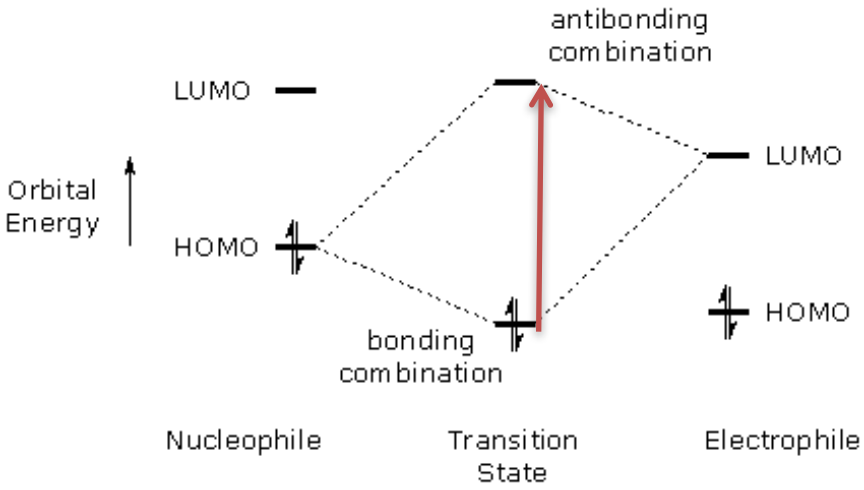


Commercial chemical development

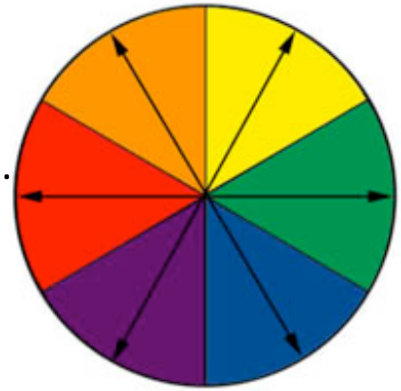


Designing for function: dyes

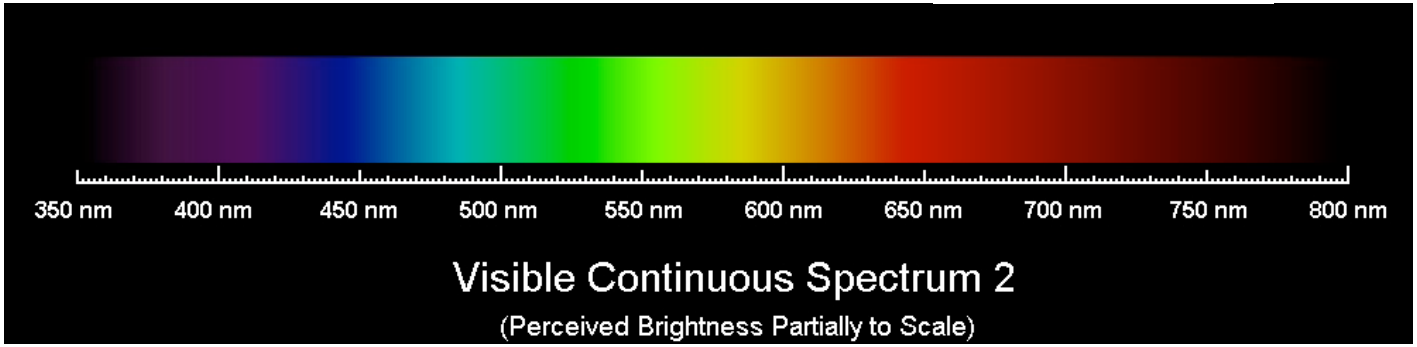
What makes a compound appear to have color?



As the HOMO/LUMO energy gap decreases, the wavelength of the absorbed color increases. To find the perceived color, look at the complementary colors on the color wheel.



$$E = hc/\lambda$$



Traditional Approach towards Minimizing Risks from Industrial Chemicals

Risk Equation, NAS 1983

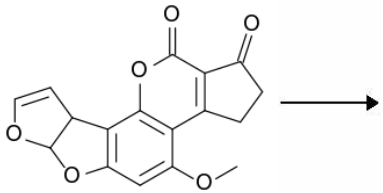
Risk = ƒ(hazard, exposure)

Probability of:

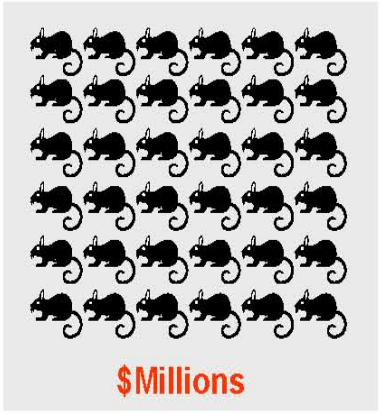
- Physical hazard – fire, chem reactivity
- Toxicity and ecotoxicity
- Global resource depletion ar destruction

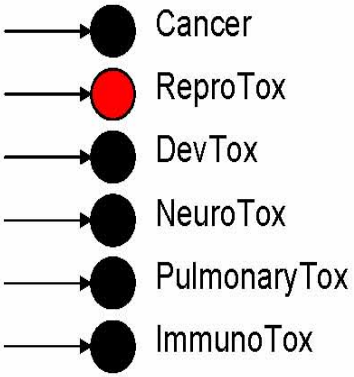
Test for hazard

Minimize exposure

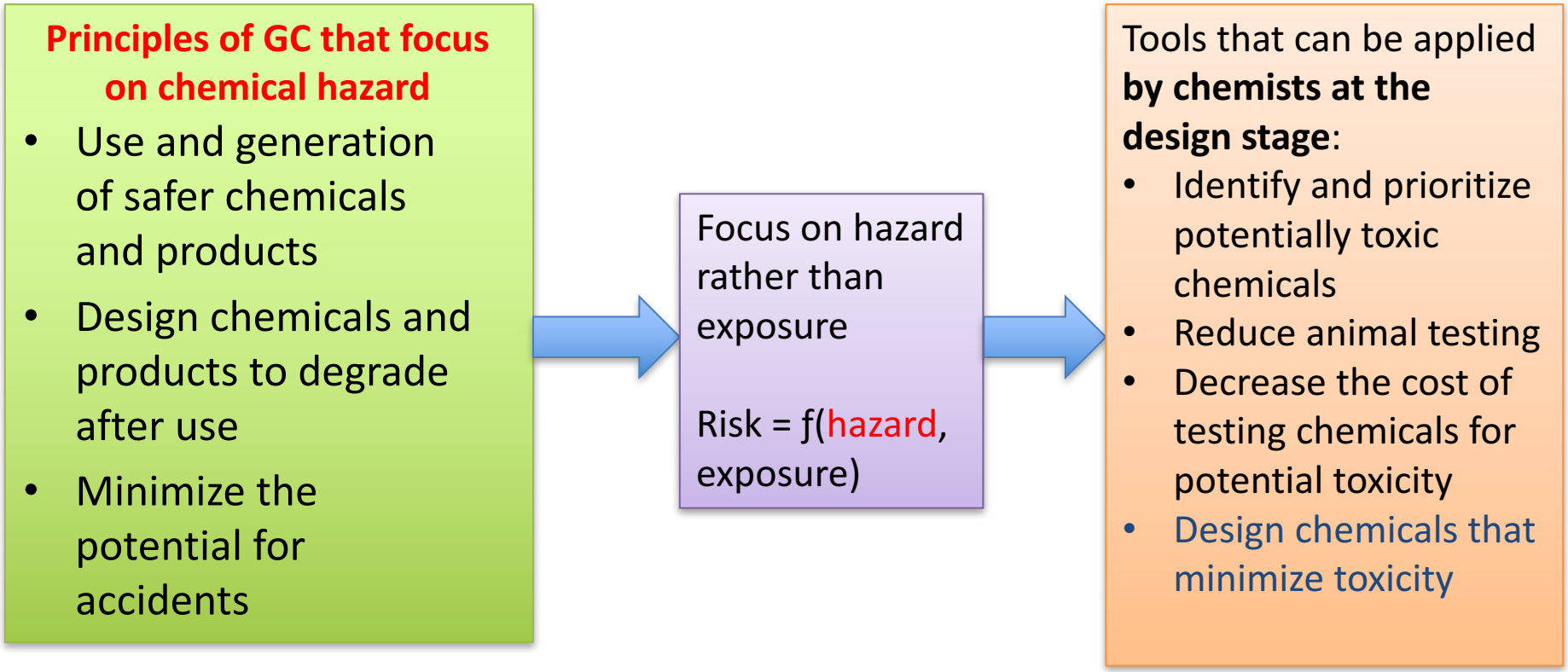


in vivo testing





Green Chemistry and Rational Design



Identification of Toxic Chemicals vs Design for Minimal Toxicity

Value of Reactive Approach

- Identify hazardous chemicals from those already in existence
- Evaluate chemical alternatives
- Carry out risk assessment

Value of Proactive Approach

- Requires that chemists consider biological activity alongside function at the design stage
- Redesign an existing chemical to minimize biological activity
- Design a new chemical that has a superior safety profile to chemicals in the market

Experimental Measurements of Toxicity *In Vivo*

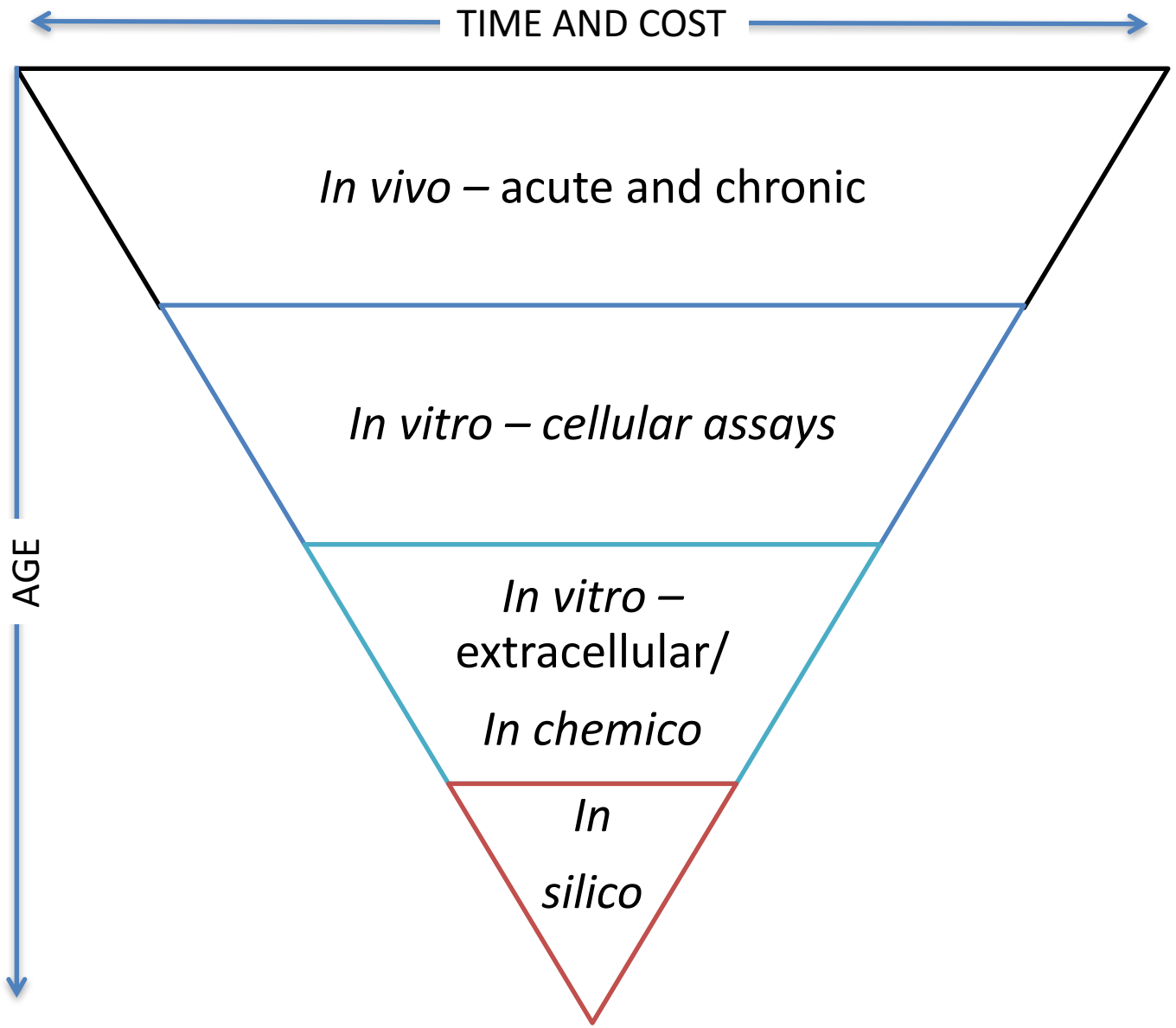
Can broadly classify toxicity tests based on *length of exposure*

- Acute Toxicity test
  - Drop dead testing
  - Time = 2 days (invertebrates) to 4 d. (fish)
    - LD_{50}
    - LC_{50}
    - TL_m (median tolerance dose)
    - EC_{50} (effective concentration)
      - Lose equilibrium, sit on bottom → "ecologically" dead
  - Not very ecologically relevant but quick, relatively cheap (but still ~$700-1,200 per test)

- Chronic toxicity test
  - Growth, reproduction
  - More ecologically relevant data but takes longer, more expensive
  - Shows effect at much lower dose
  - Test requires much more "baby-sitting"
    - NOEC (No Effect Concentration)
    - LOEC (Low Effect Concentration)

Part 3: **Using *in silico* models to predict ecotoxicity**

Toxicity (Hazard) Estimation Methods



Hazard prediction vs chemical design

Chemists/engineers consider biological activity alongside function at the design stage

In silico toxicity (hazard) estimation tools

- Identify hazardous chemicals from those already in existence
- evaluate chemical alternatives
- risk assessment

In silico Chemical Design

- Often there are no existing chemicals/materials that have desired function and no biol. activity
- In this case we must design from scratch (de novo) chemical that is functional and has a superior safety profile to chemicals in the market

Predictive Models

- What is a predictive (QSAR) model?

  A model whose *primary* purpose is for prediction of a quantitative outcome (as opposed to inference)

- Mechanistic interpretation of the model is considered necessary
- To make a good model that predicts well on future samples, you need to know a lot about
  - The predictors (physical meaning, relation to each other)
  - The biological data (accuracy, mechanisms, method)

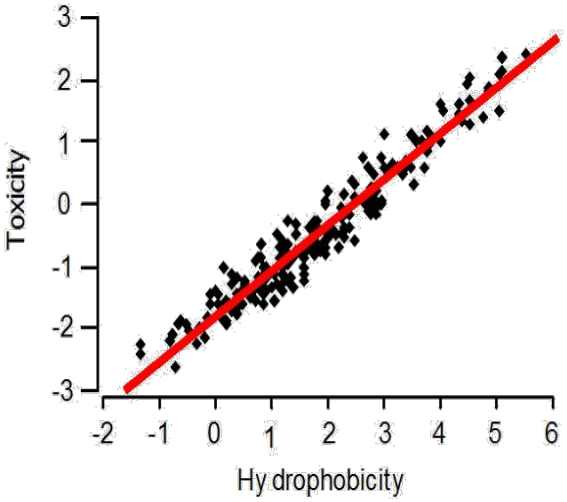
What Are They **Not** Good For?

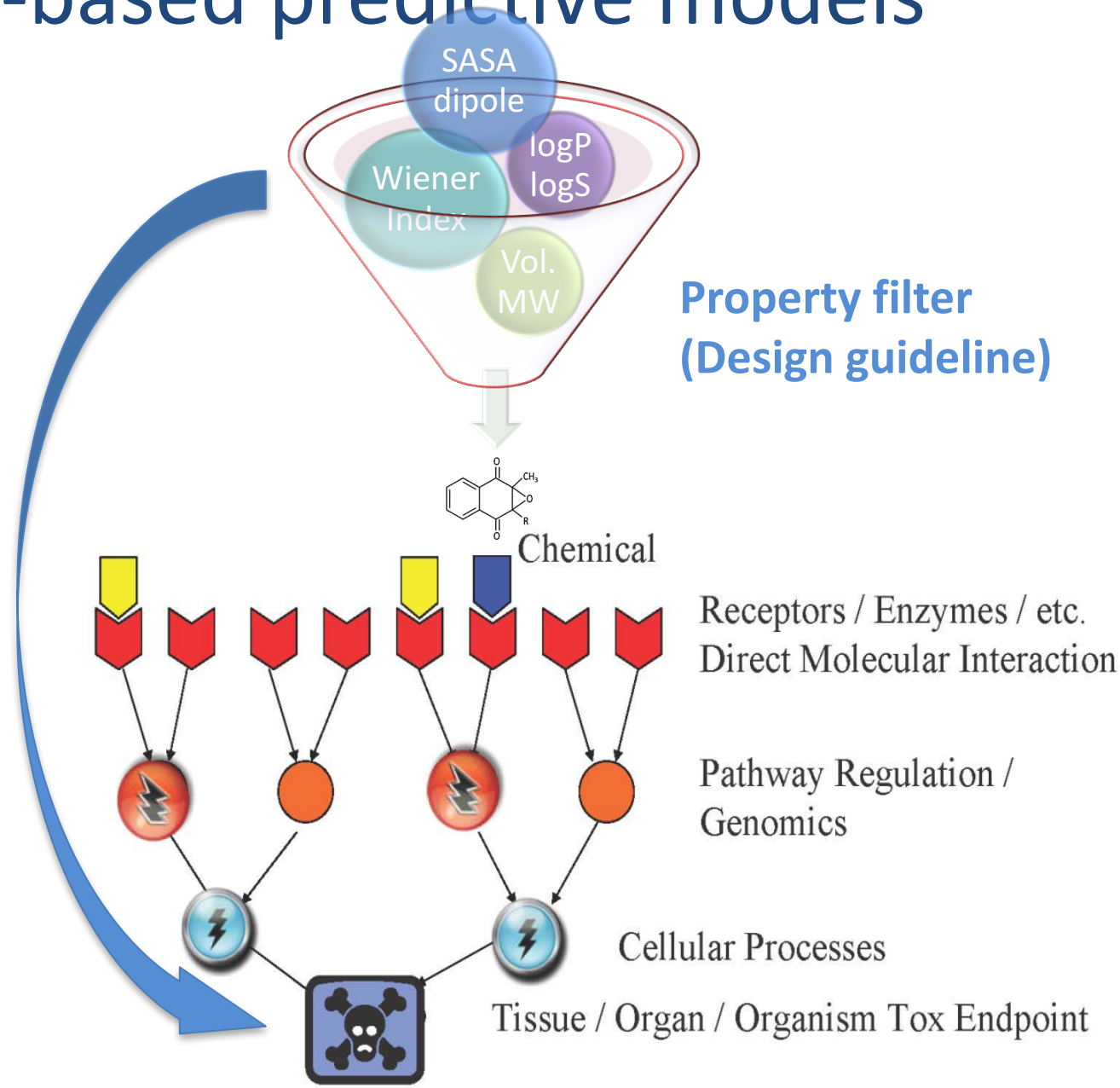
- An example:

  A toxicologist collects some data from a small animal trial and wants a model that would use gene expression data to predict toxicological response. There were about 54K predictors and data was collected on ~20 animals.

- Problem?
- Solution?
  - the mechanism is understood, effort must be made to first identify the mechanistically-relevant set of predictors

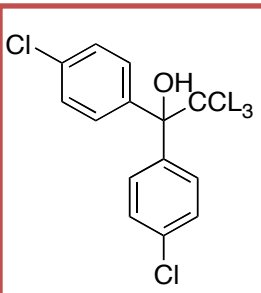
Property-based predictive models



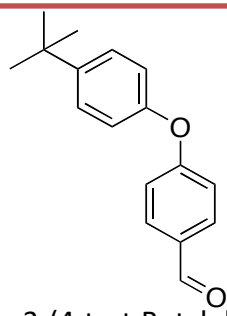


1

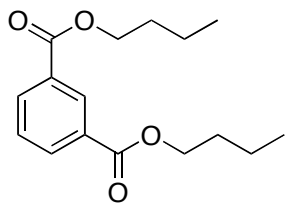
Why use properties and not structure?



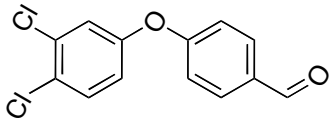
Dicofol (Kelthane)
115-32-2
logP: 6.07
LUMO: -0.467 eV
LC_{50}: 0.603 mg/L



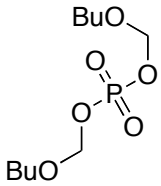
3-(4-tert-Butylphenoxy)b
69770-23-6
logP: 5.93
LUMO: -0.530 eV
LC_{50}: 0.370 mg/L



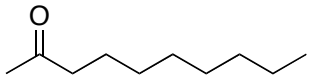
phthalate
3126-90-7
logP: 5.53
LUMO: -0.757 eV
LC_{50}: 0.900 mg/L



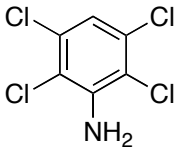
bisphenol
79124-76-8
logP: 5.49
LUMO: -0.728 eV
LC_{50}: 0.300 mg/L



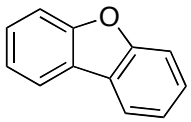
78-51-3
logP: 4.09
LUMO: 0.817 eV
LC_{50}: 11.2 mg/L



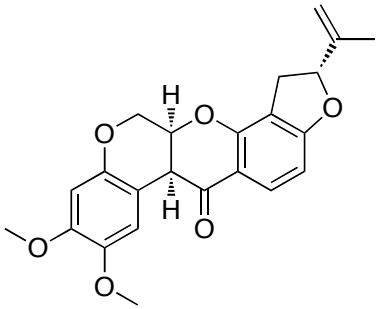
112-12-9
logP: 4.09
LUMO: 0.884 eV
LC_{50}: 1.50 mg/L



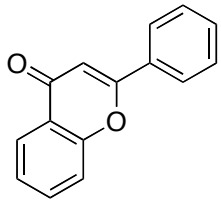
3481-20-7
logP: 4.10
LUMO: -0.507 eV
LC_{50}: 0.270 mg/L



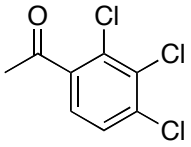
132-64-9
logP: 4.12
LUMO: -0.410 eV
LC_{50}: 1.50 mg/L



83-79-4
logP: 4.10
LUMO: -0.477 eV
LC_{50}: 0.0052 mg/L



13608-87-2
logP: 3.57
LUMO: -0.683 eV
LC_{50}: 2.00 mg/L



525-82-6
logP: 3.56
LUMO: -0.776 eV
LC_{50}: 3.50 mg/L

75-47-8
logP: 3.54
LUMO: -0.738 eV
LC_{50}: 2.90 mg/L

Assessing Model Accuracy

- How well does a regression model perform? Answering this question depends on how we want to use the model. Possible goals are:
  - To understand the relationship between the predictor and the response.
  - To use the model for prediction of unknowns

- In either case, we can use several of different measures to evaluate model performance. We will focus on two:
  - Coefficient of determination (R^2)
  - Root mean square error (RMSE)

Root Mean Squared Error (RMSE)

- RMSE measures the average deviation of an observation to the best-fit plane, i.e. square root of MSE
- RMSPE measures the average deviation of an observation to its predicted value for the test or cross-validation set

$$RMSPE = \sqrt{\frac{\sum_{i=1}^{n^*} (y_i - \hat{y}_i)^2}{n^*}}$$

n* = the number of observations in the test or cross-validation set

Assessing Model Predictive Power: *External Validation*

- Compute R^2 and RMSE on data for which the model was ***not*** built (i.e. a test set or cross-validation set).
  - For a held-out set of data, R^2 is commonly referred to as Q^2

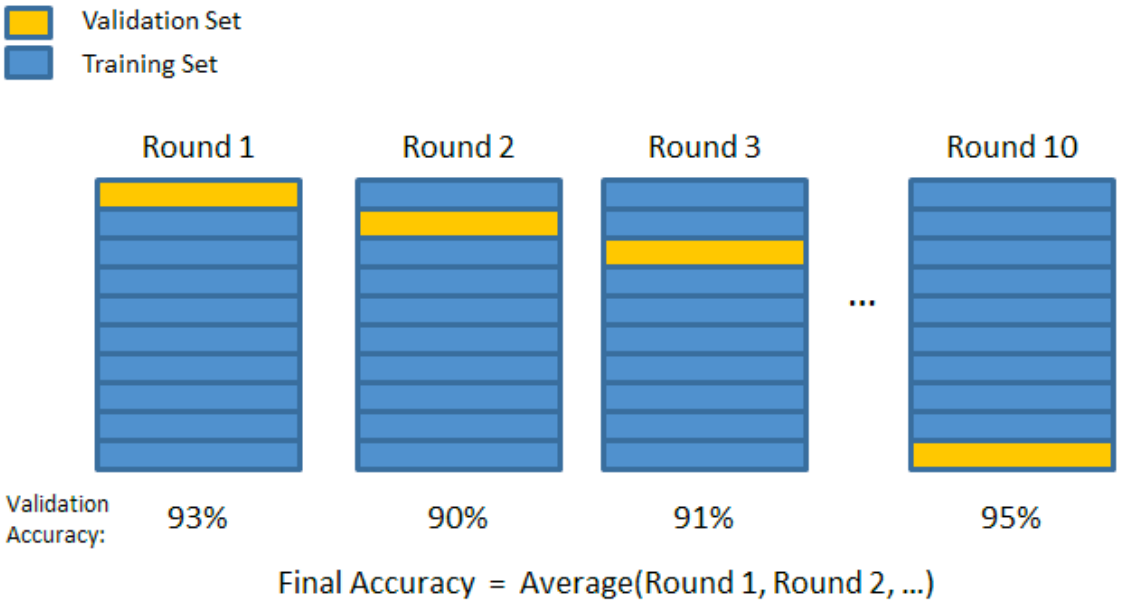
- RMSE measures the average deviation of an observation to the best-fit plane, i.e. square root of MSE
- RMSEP measures the average deviation of an observation to its predicted value for the test or cross-validation set

$$RMSPE = \sqrt{\frac{\sum_{i=1}^{n^*} (y_i - \hat{y}_i)^2}{n^*}}$$

n^* = the number of observations in the test or cross-validation set

K-fold Cross Validation: *Internal Validation*

- Here, we randomly split the data into *K* blocks of roughly equal size
- We leave out the first block of data and fit a model, which is used to predict the held-out block
- We continue this process until we have predicted all *K* hold-out blocks
- The final performance is based on the hold-out predictions



Sample Results

	Training Data		Validation Data	
	RMSE	**R^2**	**RMSEP**	**Q^2**
Linear Reg	5.23	0.691	4.53	0.742

- A reason you may see differences: multicollinearity
  - Multicollinearity in the predictors can produce somewhat unstable solutions for each resample
  - When the data are slightly changed, the model can drastically change

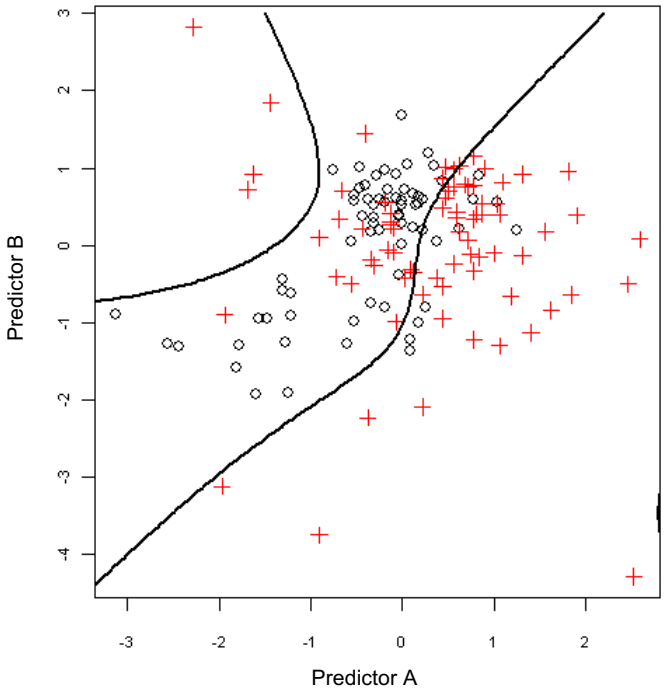
Common Pitfalls of QSARs

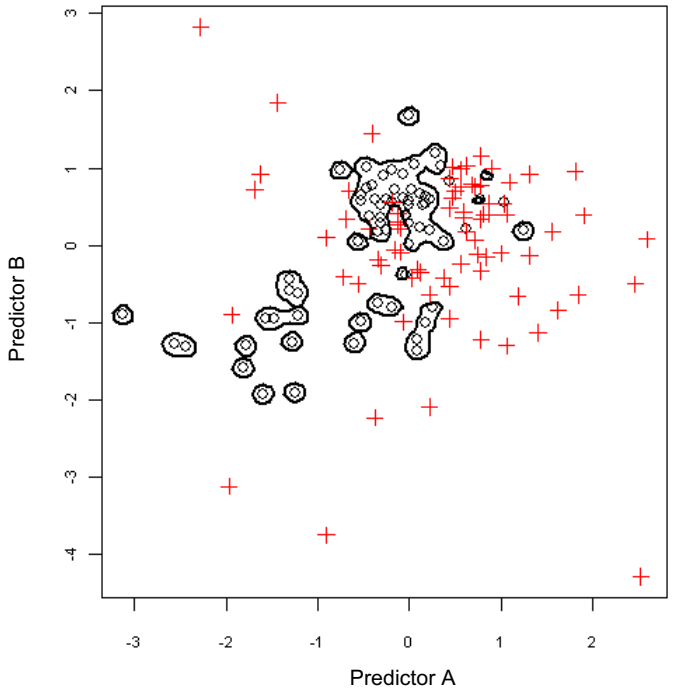
- **Over-fitting:** occurs when a model has extremely good prediction for the training data but predicts poorly when
  - the data are slightly perturbed
  - new data are used (external validation)
  - Number of descriptors used should not be > 1/5 number of compounds in training set
- **Imaginary relationships:** Complex regression and classification models assume that there are patterns in the data.
  - Without some control many models can find very intricate relationships between the predictor and the response
  - These patterns may not be valid for the entire population.

Over-Fitting

The plots below show classification boundaries for two models built on the same data

Which one is over-fit?





Defining the Applicability Domain

- *Extent of extrapolation*: What types of chemicals can be reliably predicted?
- Describe the training set
- Test the applicability
- Mechanistic applicability

QSAR Commandments

1. An unambiguous algorithm
2. A defined applicability domain
3. Appropriate measures of goodness of fit, robustness and predictivity are used
   1. R^2, RMSE ($R^2 > 0.6$)
   2. Q^2, RMSEP ($R^2 > 0.5$)
4. Model validation: quantitative assessment of model robustness and its predictive power
5. Cardinal rules:
   - n>4k (n= # of compounds, k = # of descriptors)
   - Each descriptor must have significance level p <0.05
   - Pairwise correlation among descriptors < 0.9

Gramatica, P. QSAR and Combinatorial Sci. 2007, 26, 694.
Tropsha, A.; Gramatica, P.; Gombar, V. QSAR and Combinatorial Sci. 2003, 22, 69.

Comparing QSAR Models

Table 2 Predictive tool summary

	ADMET	CADRE-AT	ECOSAR	KATE	TEST
Free-ware?	No	No	Yes	Yes	Yes
Statistical method	2D ANNE	Classification system	Class-specific linear regression	Class-specific linear regression	Consensus model
AD definition	Molecular descriptor space	Molecular descriptor space	Log *P* range and class categorization concerns	Log *P* range and class categorization concerns	Molecular descriptor space
Training set size	490	565	1000s[a]	535	823
Training set species	Pimephales promelas	Pimephales promelas	All OCSPP approved species	Oryzias latipes, pimephales promelas	Pimephales promelas
Output	LC_{50}	Toxicity category ($n = 4$)	LC_{50}	LC_{50}	LC_{50}
# of chemicals in the AD[b]	78	80	61	35	57

[a] The exact number of compounds is not available. [b] Number of chemicals in the validation set ($N = 83$) that are in the AD of each model. OCSPP – office of chemical safety and pollution prevention; 2D ANNE -two-dimensional artificial neural network ensemble. AD – applicability domain.

Test Set and Regulatory Categories

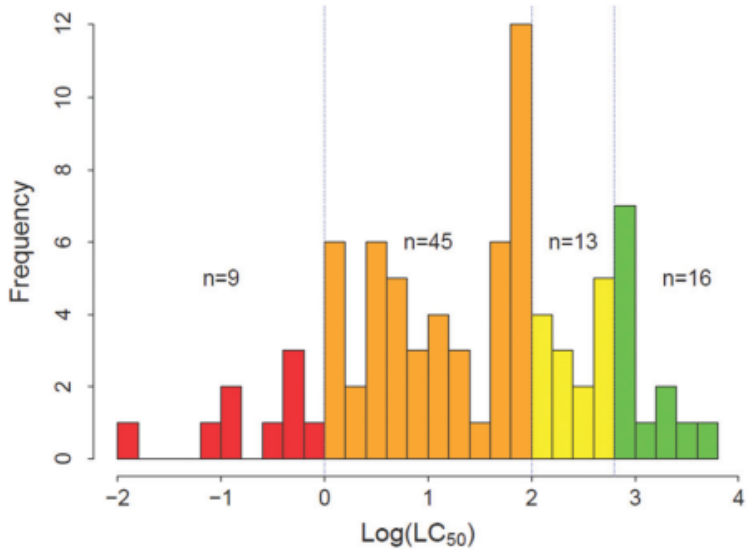


Fig. 1 The distribution of log(LC_{50}) thresholds (mg L^{-1}).

Assess 83 chemicals external to the training

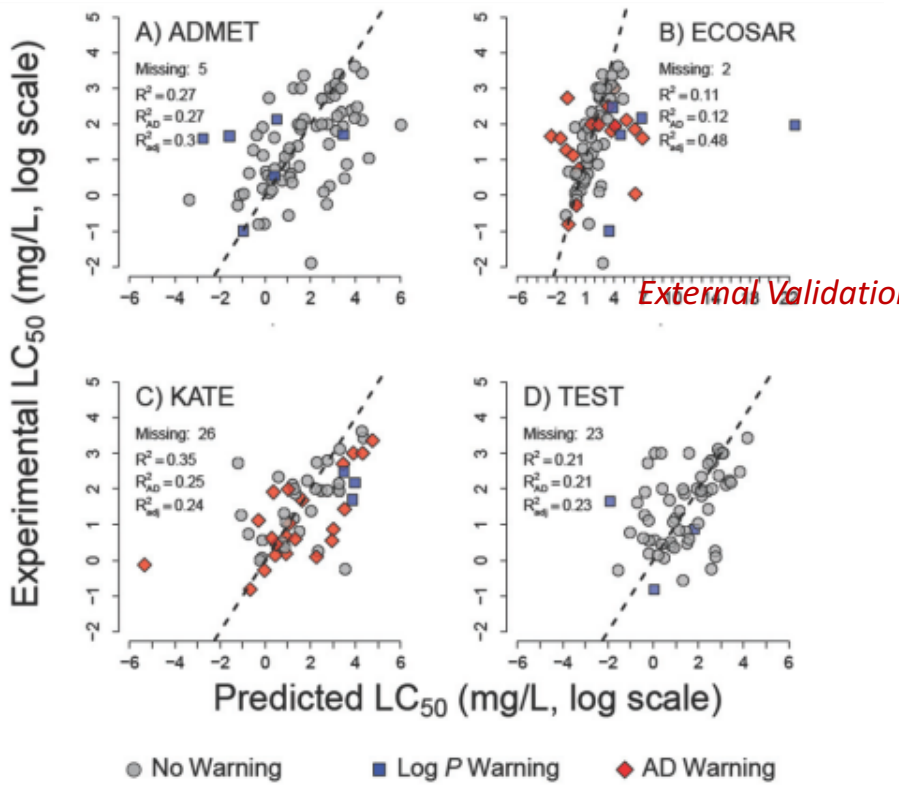
Accuracy and results

Table 3 Tool performance and comparison summary statistics based on all 83 chemicals in the testing data set

Measures of predictive accuracy	ADMET	CADRE-AT	ECOSAR	KATE	TEST
Total accuracy (%)[a]	53%	83%	51%	58%	48%
Predictive power (%)[b]	49%	80%	49%	40%	35%
Number of missing predictions	5	3	2	26	23
Coefficient of variance (R^2)	0.27	NA[c]	0.11[d]	0.35	0.21
RMSE (log scale)	1.60	NA[c]	2.94[d]	1.47	1.32
% within 1 regulatory category	80.8	92.5	85.2	85.5	88.3
% within a factor of 2 (%)	25.6	NA[c]	25.9	26.3	30.0
% within a factor of 5 (%)	48.7	NA[c]	54.3	47.4	50.0
% within a factor of 10 (%)	57.7	NA[c]	63.0	64.9	63.3
% within a factor of 100 (%)	80.8	NA[c]	76.5	82.5	85.0
% within a factor of 1000 (%)	91.0	NA[c]	86.4	94.7	98.3

[a] Total accuracy is the fraction of chemicals assessed by each tool for which the predicted LC_{50} falls within the same regulatory category as the measured LC_{50}. [b] Similar to total accuracy, predictive power measures the total number of correct category assignments. However, lack of prediction is treated as an incorrect assignment. [c] Cannot be calculated; software tool provides regulatory category designation only. [d] Parametric correlation might provide poor estimate of covariance due to extreme outliers. RMSE – root mean squared error.

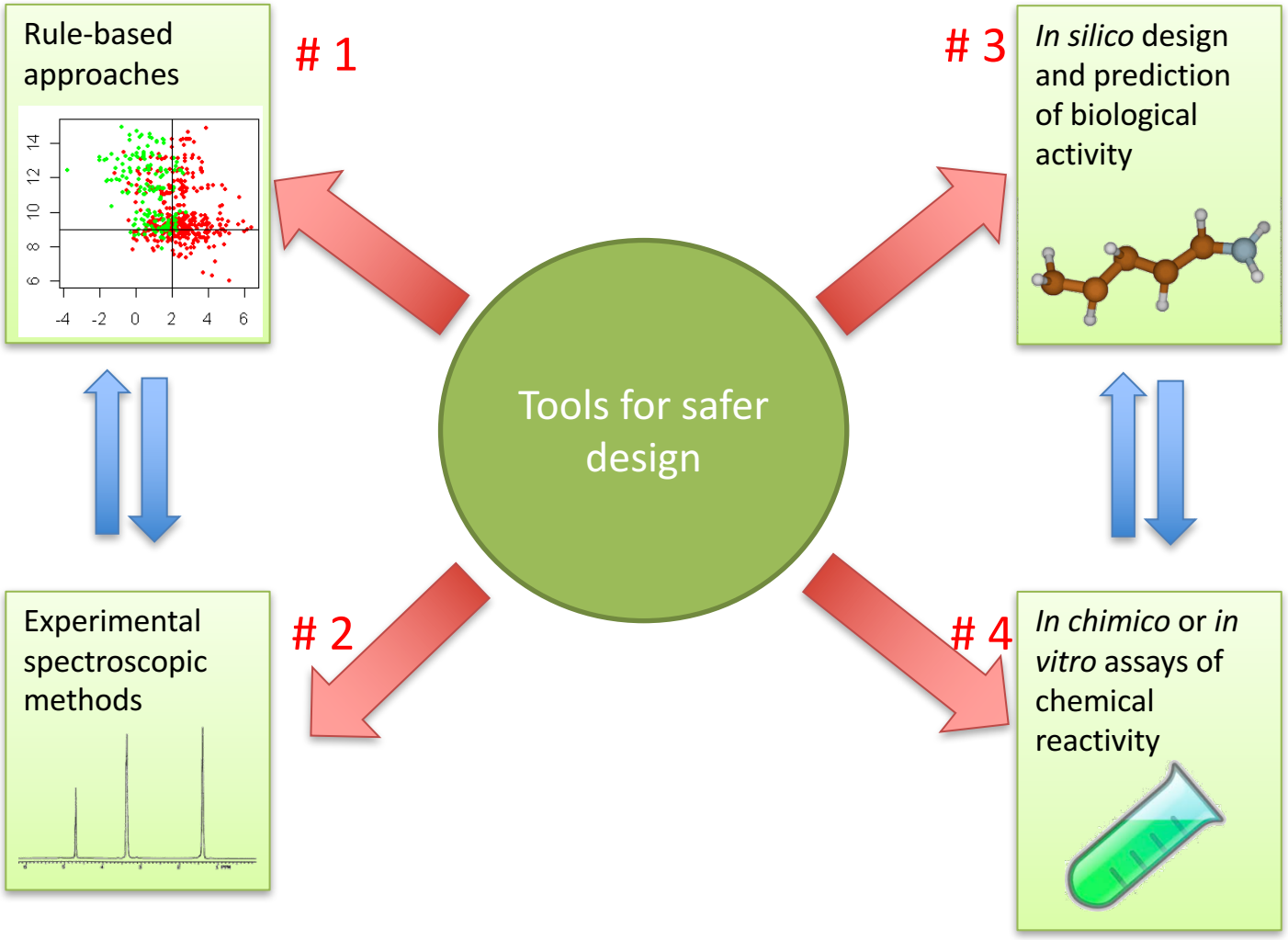
Predicted vs Measured LC_{50}



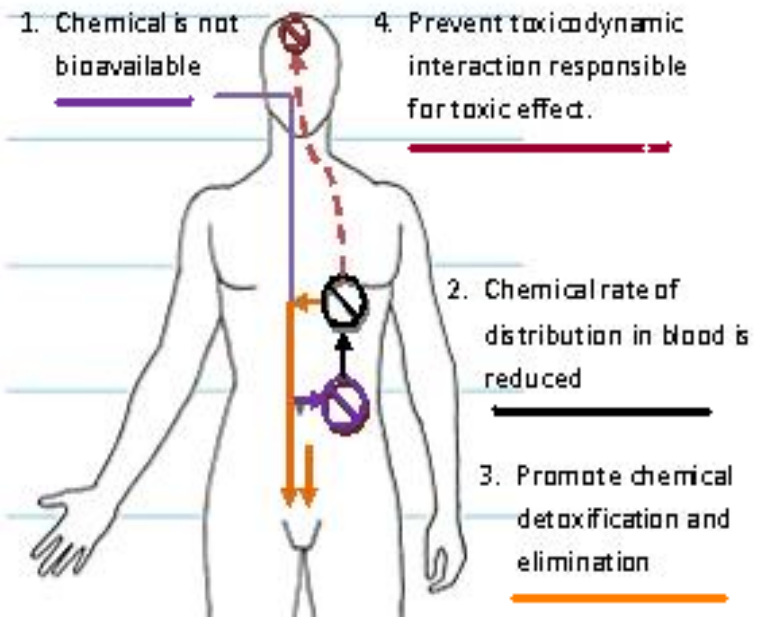
External Validation

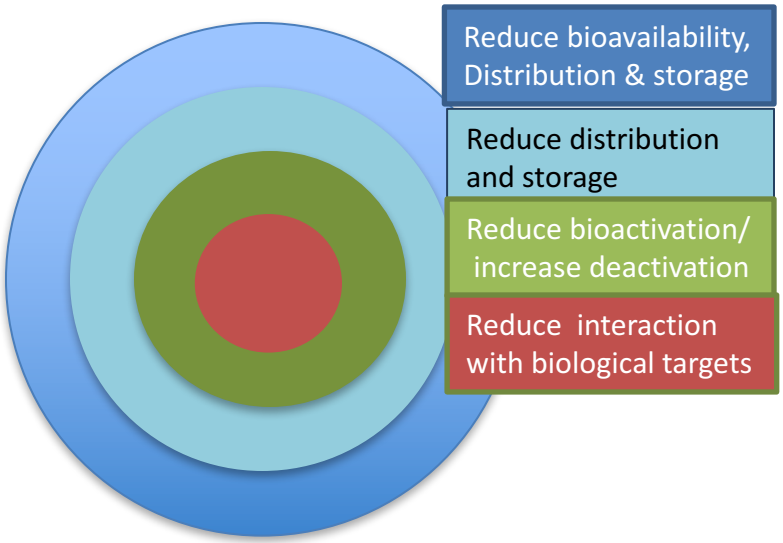
Fig. 3 Correlations between predicted and experimental LC_{50} values on log for (A) ADMET predictor, (B) ECOSAR, (C) KATE, and (D) TEST. Red: chemicals that lie outside the AD; blue: log *P* estimates used by the tool are >1 log unit below Marvin log $D_{7.4}$ estimates. R^2_{AD}: coefficient of determination for chemicals inside AD; R^2_{adj} is the coefficient of determination for chemicals inside the AD and without log *P* warnings.

Developing Tools That Enable the Rational Design of Safer Chemicals



The physiological “gates” of chemical exposure

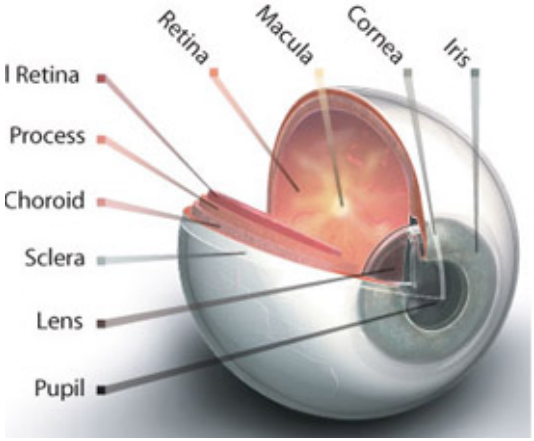


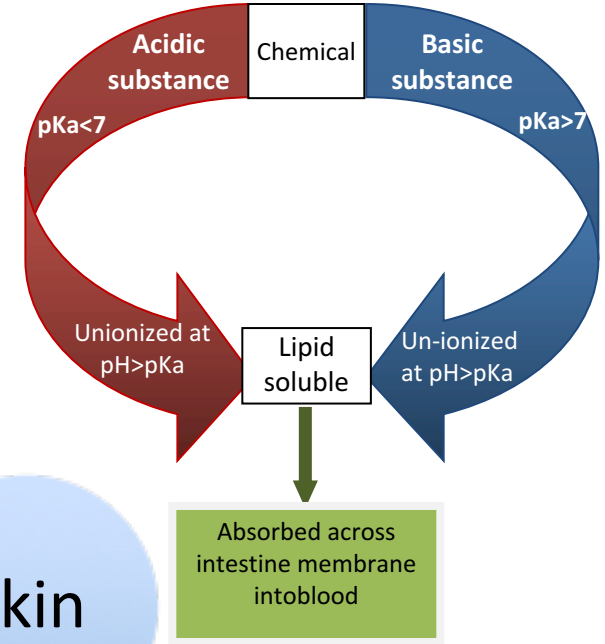


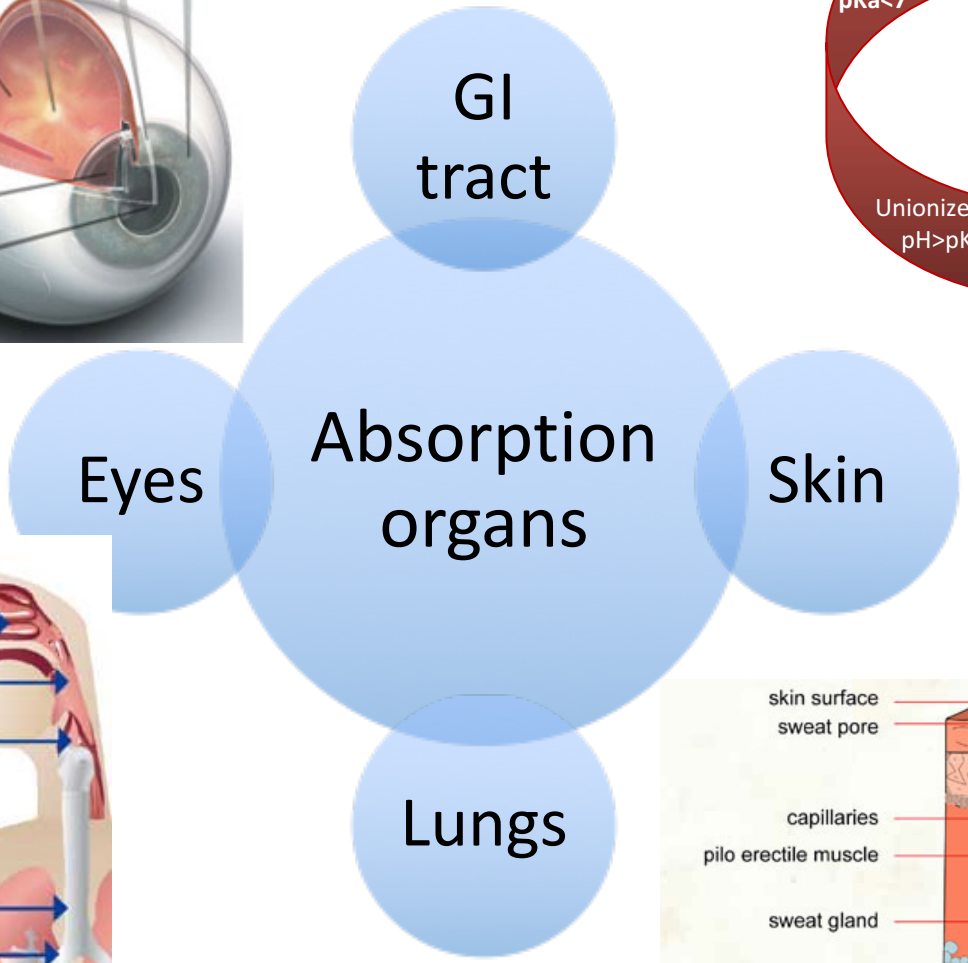
Mechanistic layers of reducing toxicological hazard.

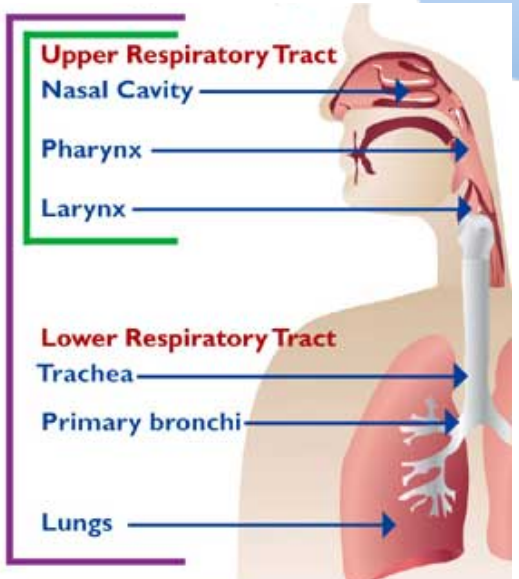
Voutchkova, A.; Osimitz, T.; Anastas, T. *Chem. Rev*. 2010, 110, 5845

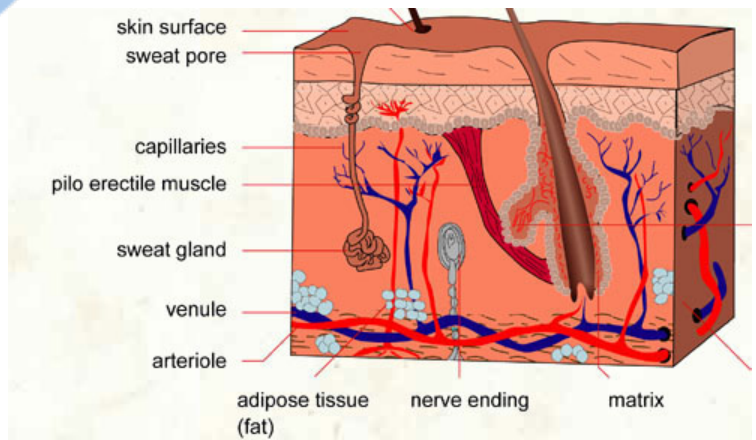
Property-based guidelines for bioavailability



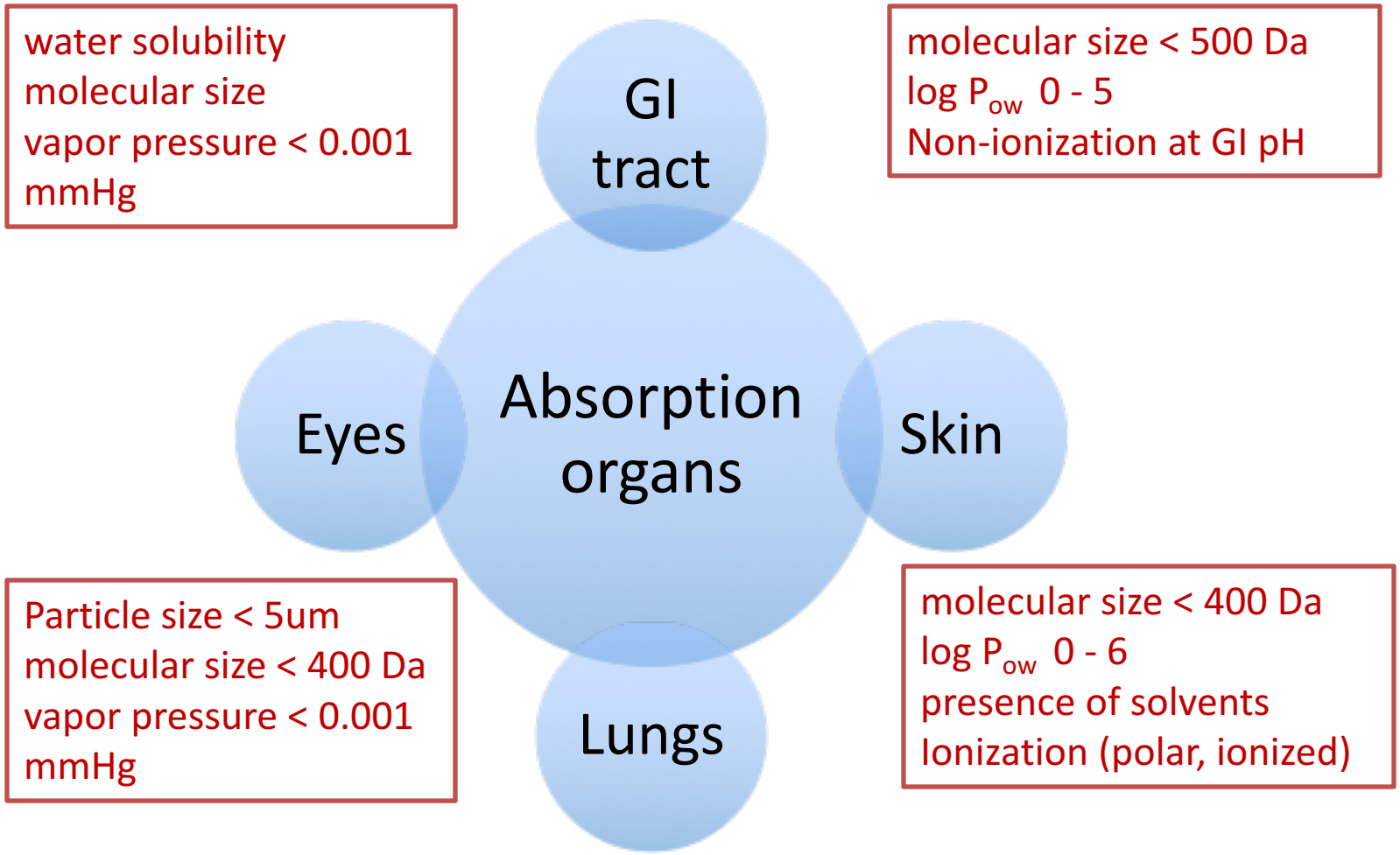








Property-based guidelines for bioavailability



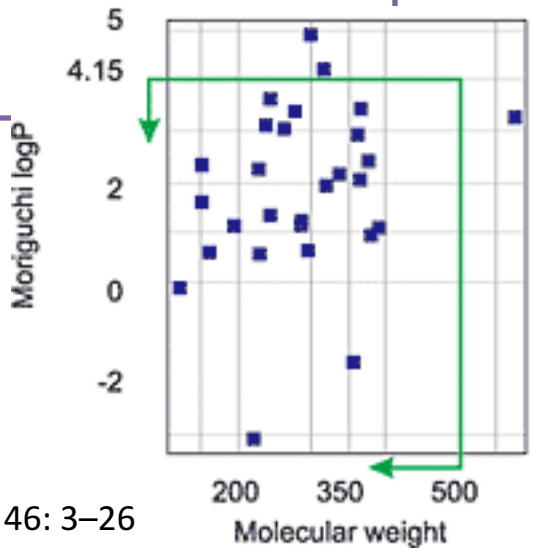
Voutchkova, A.; Osimitz, T.; Anastas, T. Chem. Rev. 2010, 110, 5845

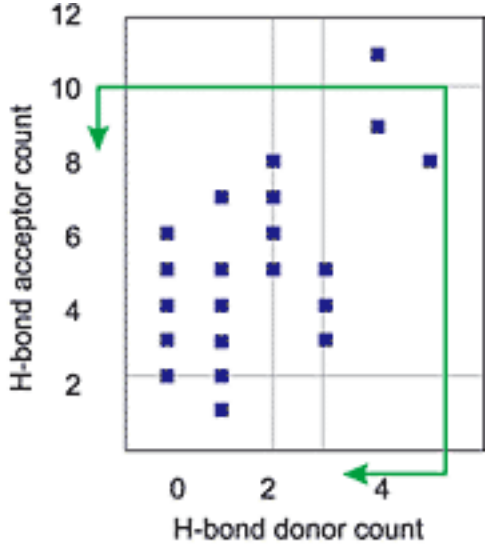
Lipinski rules for drug likeness
Oral Bioavailability

Lipinski, 1997
~90% of drugs on the market have the following properties in common:

Lipinski's Rules for Druglikeness

1. Not more than 5 hydrogen bond donors
2. Not more than 10 hydrogen bond acceptors
3. Molecular weight under 160-480 D
4. Octanol-water coefficient (logP) < 5
5. 20-70 atoms
6. Molecular refractivity from 40-130 m^3/mol.
7. At least one N or O
8. Less than 6 rings



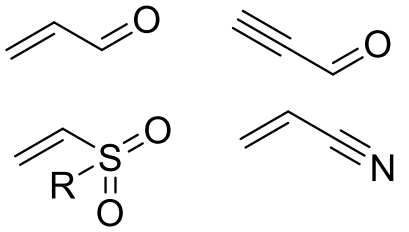


Lipinski, C. et al, *Adv. Drug Del. Rev.* 2001, 46: 3–26

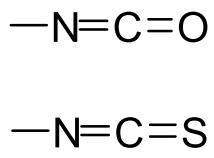
1

Rule-based approaches: chemical reactivity

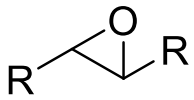
Examples of electrophilic toxicophores:



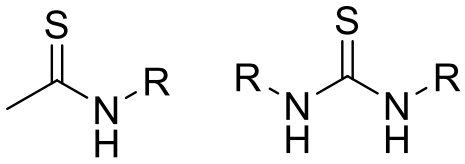
Michael acceptors
Mutagenicity, carcinogenicity, hepatoxicity, neurotoxicity, hematotoxicity



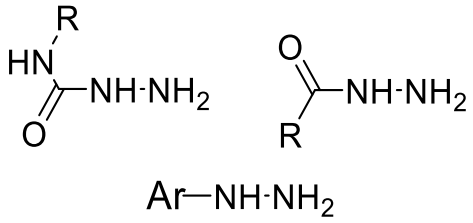
Isocyanates
Mutagenicity, carcinogenicity, respiratory sensitization, asthma



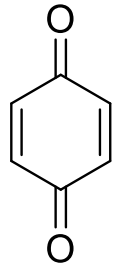
Epoxides
Mutagenicity, carcinogenicity, respiratory sensitization



Thiocarboxamides, thioureas
Thyroid gland toxicity, hypothyroidism

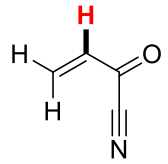


Hydrazides, semicarbazides
Developmental toxicity, osteolathyrism, blood dyscrasias, cancer, autoimmune disease.



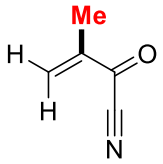
Quinones
Mutagenicity, carcinogenicity, Oxidative stress

Structural Modifications That Influence Biological Activity



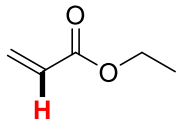
Acrylonitrile

Possible human carcinogen



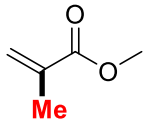
Methylacrylonitrile

Non-carcinogenic



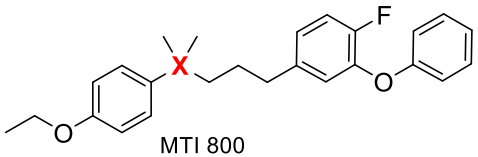
Ethyl acrylate

Carcinogenic



Methyl methacrylate

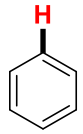
Non-carcinogenic



MTI 800

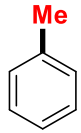
X = C" fish LC_{50} = 3 mg/L

X = Si, no mortality at 50 mg/L



Benzene

Carcinogenic

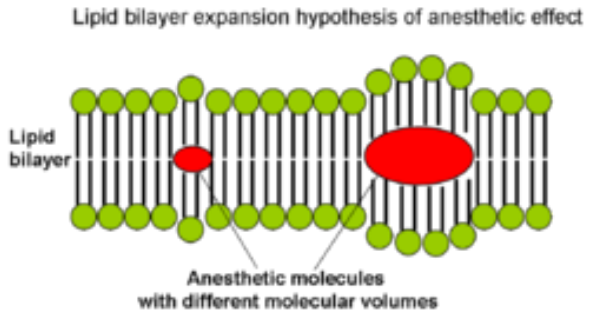


Toluene

Less carcinogenic

Pharmacodynamics/Toxicodynamics

Baseline toxicity (Narcosis)



- Bulky and hydrophobic (organic) molecules accumulate inside the cell membrane causing its distortion and expansion
- This reversibly alters function of membrane ion channels, thus providing anesthetic effect.
- Actual chemical structure of molecule not important, but its molecular volume and hydrophobicity play the major role

Reactive toxicity

1. Covalent interactions
2. Receptor binding
3. Non-covalent interactions

Baseline Narcosis: Octanol-water partition coefficient (log *P*)

$$log\ P_{oct/wat} = log\left(\frac{[solute]_{octanol}}{[solute]_{water}^{un-ionized}}\right)$$

Strongly orrelated to:
- Bioavailability
- Bioaccumulation
- Narcosis

***In Silico* methdos:**
- Group contribution (CLOGP, ALOGP, KOWWIN): r^2 = 0.90-0.95
- Molecular topology methdos (VLOGP, QikProp): r^2 = 0.90-0.98
- Free energy of solvation methods

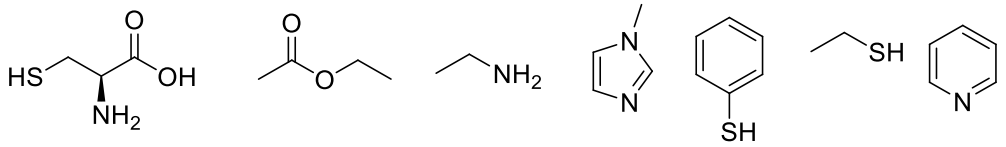
Experimental methods:
- Shake flask
- HPLC (OECD 117)
- Electrokinetic chromatography

- 10
hydrophilic

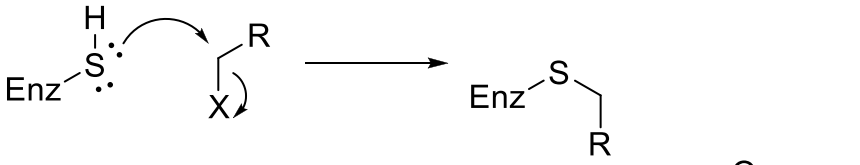
+10
lipophilic

Reactive toxicity: Covalent Interactions

Model biological nucleophiles

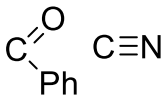


Mechanisms

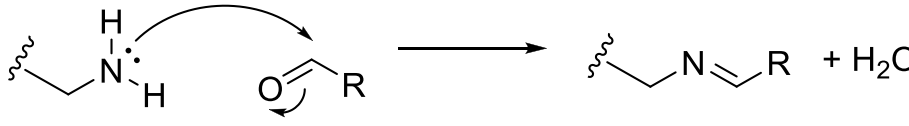


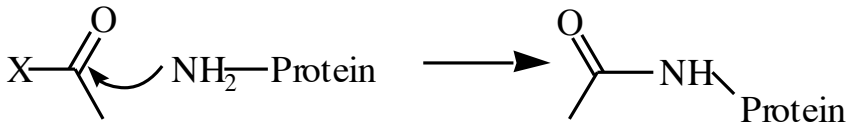
X = I, Br, Cl

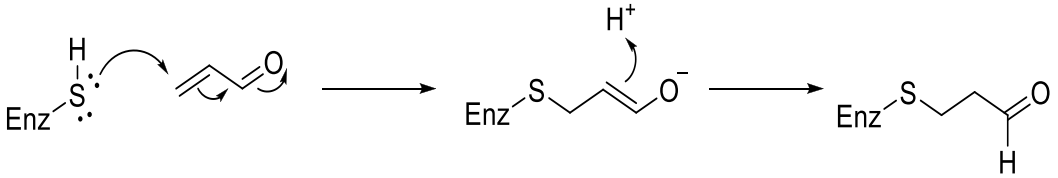
R = C=C; Ph, C≡C



C≡N







C=O NO_2 C≡C C≡N SO_2

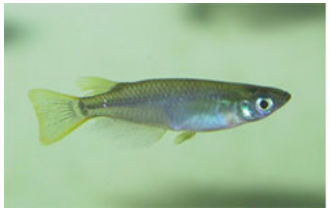
Application to Design Guidelines for Aquatic Toxicity

Agglomerated biological data that is prone to experimental error

Fathead minnow
LC_{50}, 96-h assay

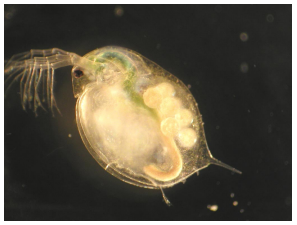
U.S. E.P.A.

671 chemicals



Japanese medaka
LC_{50}, 96-h assay

285 chemicals



Daphnia magna
EC_{50}, 48-h assay
Japan Ministry of Environment

363 chemicals

4 categories guided by EPA thresholds of concern for acute aquatic toxicity (LC_{50}/EC_{50}:)

<1 mg/L < 0.0067 mmol/L	1–100 mg/L 0.0067 - 1.49 mmol/L	100–500 mg/L 1.49-3.32 mmol/L	> 500 mg/L >3.32 mmol/L

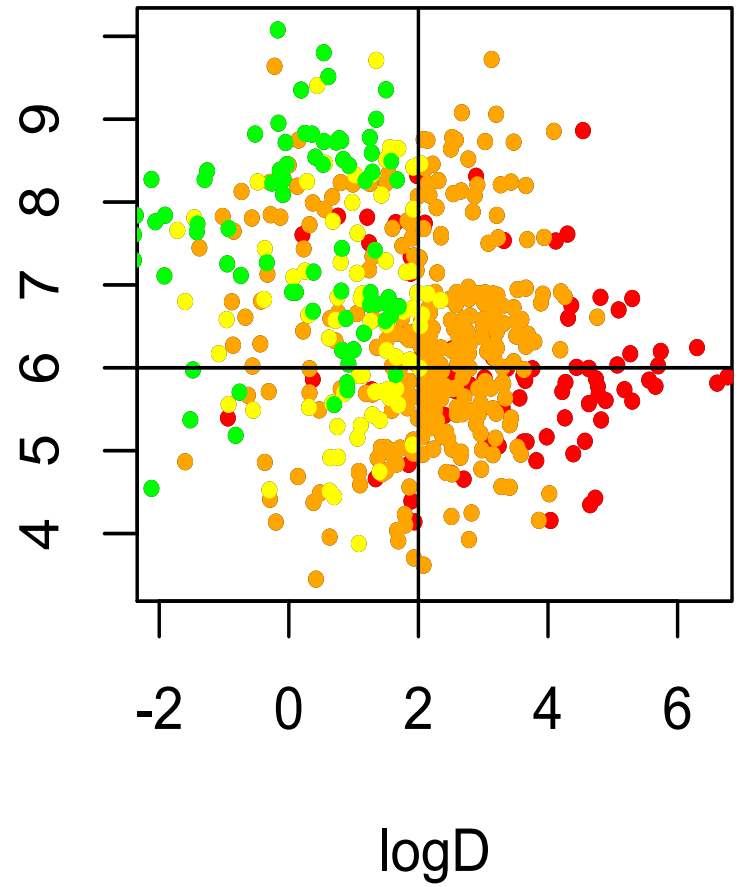
Rule of 2 for reduced aquatic toxicity

logD: logP at pH 7.4 (biological)
dE: HOMO-LUMO gap

85% of the compounds that have low acute aquatic toxicity concern have a logD < 1.7 and ΔE > 6 eV

58

Compounds that meet these criteria are **10 times more likely to have low acute aquatic toxicity** compared to compounds that do not meet these criteria. These results are mechanistically rationalized.

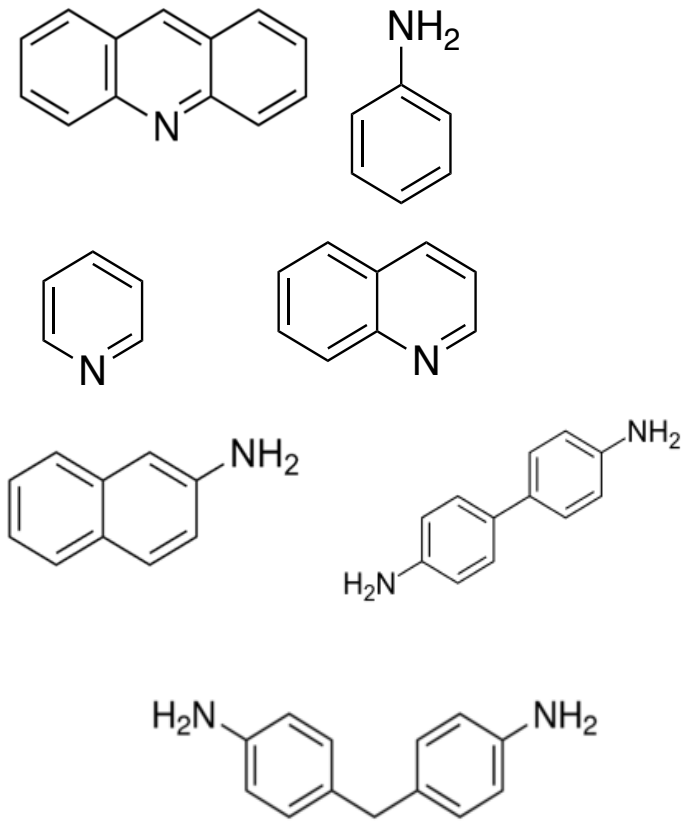


Take-away: how to go about considering ecotoxicity of a chemical

- Is it a known compound? If so, identify all available experimental data (in vivo and in vitro) for all ecotoxicity endpoints
- If it is a new compound/material, consider predictive methods:
  - Predict logP and dE
  - Determine if likely safe to aquatic species (rule of 2)
  - Consider likelihood of reactivity (reactive f-n groups)
  - Consider bioavailability

Case Study: Selecting Safer Aromatic Amine

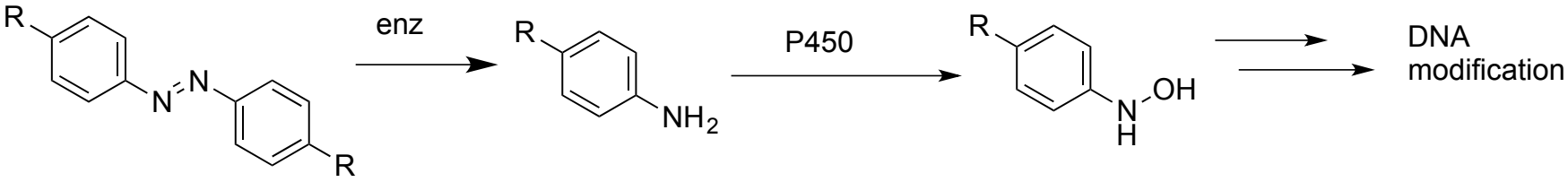
- Aromatic amines are widely used in dyeing industry as a starting material for manufacturing of different types of azo dyes
- Used in chemical industry for manufacturing petrol and diesel fuel, varnishes, and antioxidants.
- Widely used in some metal-coating multifunctional compositions for motor, transmission and industrial oils.
- Some aromatic amines can be find application in plastic, textile and rubber industries.
- Used in production of cloths, rubber and plastics.
- Formed during the thermal treatment of foods with high protein content. They can also be transferred from food packaging materials into foodstuffs



- **Toxicity:** However, they are known to be toxic, associated with methaemoglobinemia, agranulocytosis, aplastic anaemia, hepatotoxicity, skin hypersensitivity and increased risk of mutagenicity.
- Mechanism:

1. oxidation of the aromatic ring *ortho* or *para* to the aniline nitrogen

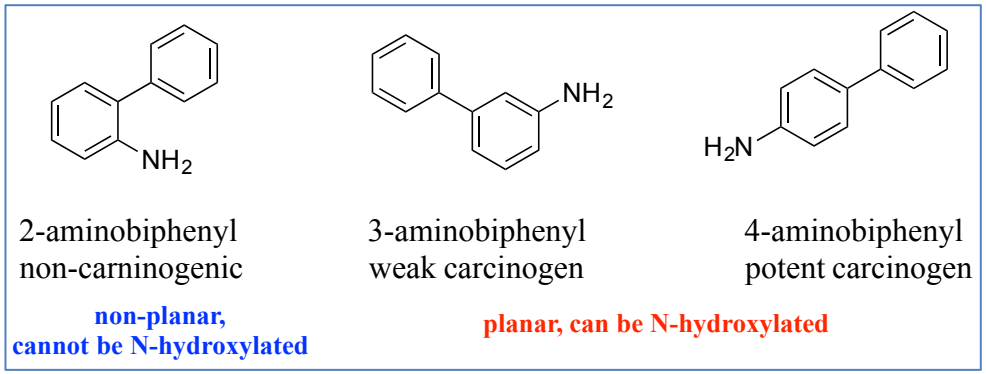
Diazo dyes and primary aromatic amines: mutagenicity



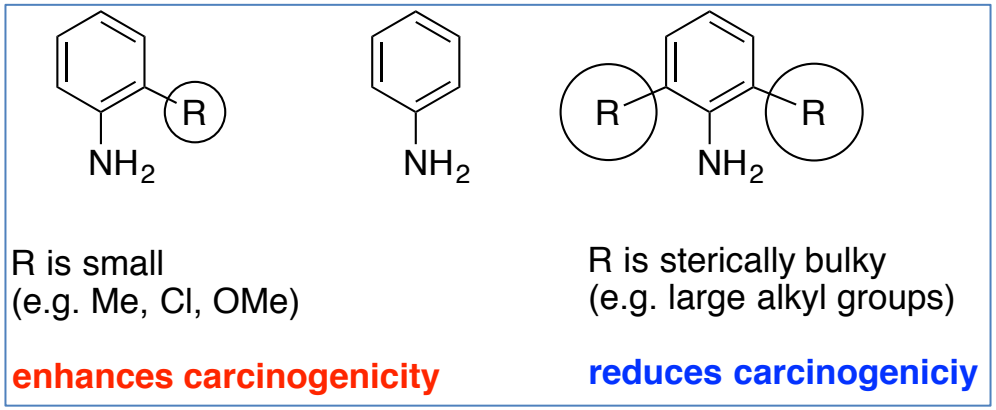
2. Oxidation of the aniline nitrogen to hydroxylamine, nitroso, nitro and related species. the hydroxylamine species undergo acetylation or sulfation to deliver a good leaving group which leads to reactive metabolites

Is it possible to develop a “safe” diazo dye?

– Decrease N-hydroxylation of aromatic amine



– Reduce the electrophilic reactivity of the SA by steric or electronic effects.



Goal of Case Study

You are provided with a list of PAAs and the acute and chronic ecotoxicity data associated with them, in addition to predicted data on biodegradation. From these compounds select 3 candidates to propose as ***safer*** alternatives.

Report back: Prepare 2 min report (

Biodegradation resources

- EpiSUITE BIOWIN (US EPA)
- http://eawag-bbd.ethz.ch/predict/

Workshop References

Kostal, J.; Voutchkova-Kostal, A. M.; Zimmerman, J. B.; Anastas, P., Computational Approaches for Molecular Design for Reduced Toxicity - A Case Study: Acute Toxicity to the Fathead Minnow (P. promelas). *Proc. Nat. Acad. Sci.* **2013,** *In Review.*

Voutchkova-Kostal, A.; An, N.; Van Der Mei, F.; Patent Application No. 61/836,430, 2013.

Kostal, J.; Voutchkova-Kostal, A. M.; Weeks, B.; Zimmerman, J. B.; Anastas, P. T., A Free Energy Approach to the Prediction of Olefin and Epoxide Mutagenicity and Carcinogenicity. *Chem. Res. Tox.* **2012** *25* (12), 2780–2787.

Voutchkova-Kostal, A. M.; Kostal, J.; K., C.; Brooks, B. W.; Zimmerman, B.; Anastas, P., Rational Molecular Design for Reduced Chronic Aquatic Toxicity *Green Chemistry* **2012,** *14*, 1001-1008.

Voutchkova, A. M.; Kostal, J.; Steinfeld, J. B.; Emerson, J. W.; Brooks, B. W.; Anastas, P.; Zimmerman, B., Towards rational molecular design: derivation of property guidelines for reduced acute aquatic toxicity. *Green Chemistry* **2011,** *13* (9), 2373-2379.

Voutchkova, A. M.; Osimitz, T. G.; Anastas, P. T., Toward a Comprehensive Molecular Design Framework for Reduced Hazard. *Chem Rev* **2010,** *110* (10), 5845-5882.