

US EPA ARCHIVE DOCUMENT

QSAR Overview

P. Schmieder

EPA, Office of Research and Development,
National Health and Environmental Effect Research Lab,
Mid-Continent Ecology Division, Duluth, MN

Oct, 2008

Quantitative Structure-Activity Relationships

Assumptions

- A chemical's structure imparts properties
- A group of chemicals that produce the same biological activity (toxicity; adverse effect) have something similar about their chemistry (structure)
- Goal is to quantify 'structural similarity' imparting biological activity; identify which other chemicals may be 'similar' with the assumption that an untested chemical may produce the same activity

Chemical similarity is defined in the context of biological similarity

- Robustness Depends on:
 - Well-defined biological system; Well-characterized chemistry
 - Well-defined application –
 - Risk context - What's the question being asked - problem definition

OECD Principles for QSAR Validation

- Transparent
 - What are the assumptions?
- Defined Assay
 - Well-defined biological endpoint
 - Well-defined chemistry
- Defined Applicability Domain
 - Well-defined application – risk context
 - define the chemicals and the question the QSAR methods are being asked to address;
 - Does existing QSAR apply? Does existing QSAR knowledge-base (training set) cover the types of chemicals of regulatory concern; Is the data the QSAR was built on understood well enough to answer the regulatory question?
- Unambiguous algorithm
 - Statistical correlations (transparency – are chemical parameters used in correlations interpretable/reasonable in regard to activity?)
 - Expert Systems – on-going OECD efforts in this area
- Mechanistic interpretation
 - MOA; Toxicity Pathway to Adverse Outcome
 - Clustering/Sub-categories – similar chemistry resulting in similar activity

QSAR Assumption

Chemical
Structure/
Property



Toxicological
Potential

Δ Chemical
Structure/
Property



Δ Dose
Metric
(kinetics/
metabolism)



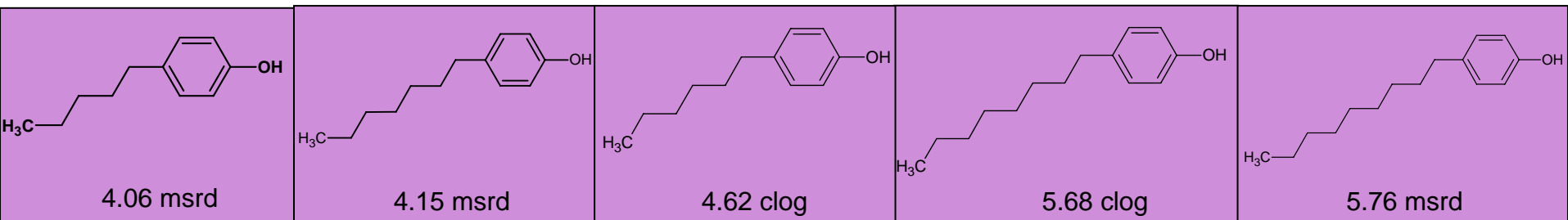
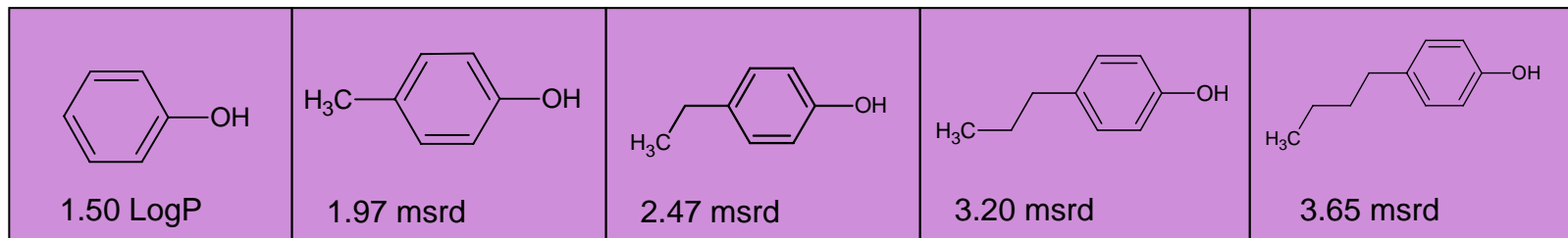
Δ Endpoint
Potency

Toxic potency is correlated to chemical concentration at the site of action

-C. Hansch

Well-defined system (chemistry and biology)

Homologous Series Alkylphenols



Well-Defined Biological System

(What do you know and what are you assuming)

- Is the chemical administered what you thought it was
 - Impurities
- Metabolism
 - Is the system used for collection of empirical data capable of xenobiotic metabolism?
 - Is what you're measuring due to parent chemical or to a metabolite?
- Kinetics
 - What do you understand about the chemical kinetics within the system?
 - Is the chemical in solution
 - Bound and unavailable
 - Loss to hydrolysis

Has chemical form and/or concentration been measured in the biological system upon which the QSAR is based

QSAR Approach

- Think beyond QSAR as a model; think of it as an approach, a tool to help understand how a chemical is most likely to interact with a biological system and what adverse effect might be the consequence of that interaction
- QSAR depends upon a well-defined biological system
- QSAR for large diverse chemicals inventories is an iterative process
- How QSAR used depends upon the regulatory context
 - Defining the regulatory domain is non-trivial; identify the exact chemicals and verify structures
 - Defining the regulatory question is essential; regulatory acceptance criteria are dependent upon the use

Risk Context

Development and use of a QSAR in regulatory risk assessment requires clear problem definition

- The purpose of the QSAR application must be well-defined (e.g., priority setting for testing, and chemical-specific risk assessment are two very different purposes – different acceptance criteria –
- The chemicals of regulatory concern must be defined to establish an appropriate training set for QSAR development and/or to assess appropriateness of QSAR application
 - Regulatory Domain
 - Applicability Domain of QSAR (dependent on Training Set)

A QSAR can only be as good as the underlying toxicological understanding and data it is based upon

- Toxicological activity is assessed based on a well-defined endpoint in a well-defined assay
 - e.g., chemical dosimetry –
 - if you assume parent chemical is responsible for biological activity but in fact a metabolite produced toxicity, then you're working from wrong structure
 - If you assume chemical was 100% available in your system but in fact 80% was loss due to volatility, or binding to glassware, bound to proteins, or unavailable in vehicle administered, etc

Example from Aquatic Toxicology to illustrate QSAR approach

Developing QSARs for TSCA Inventory

- Regulatory Context - TSCA Inventory ~ 40,000 chemicals from diverse chemical classes (not a small homologous series); industrial chemicals not designed for target MOA
- Question posed – what can we say about toxic potential from just the chemical structure
- Biological endpoint

Example: Aquatic Toxicology

Developing QSARs for TSCA Inventory

Biological endpoint

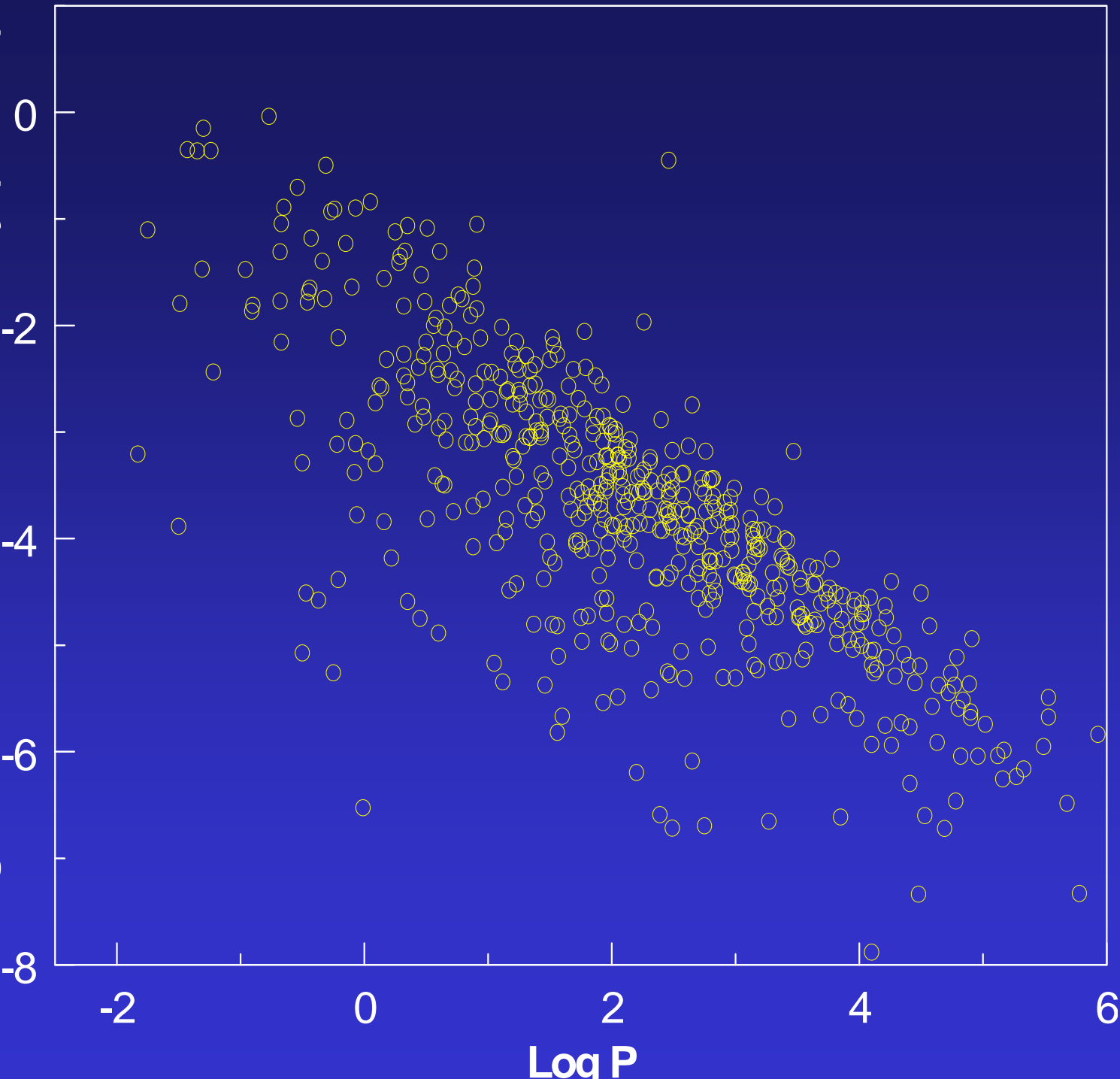
Fish Acute Toxicity - 96hr LC50; >600 chemicals

- Fathead minnow
- Flow-through test
- Endpoint - lethality
- Measured water concentrations
 - Similar to Mammalian inhalation tests
 - Steady-state; water conc is surrogate for internal conc

Fish Chronic Toxicity – 30 d Embryo/Larval
EC50; >150 chemicals

- Endpoints - growth; survival

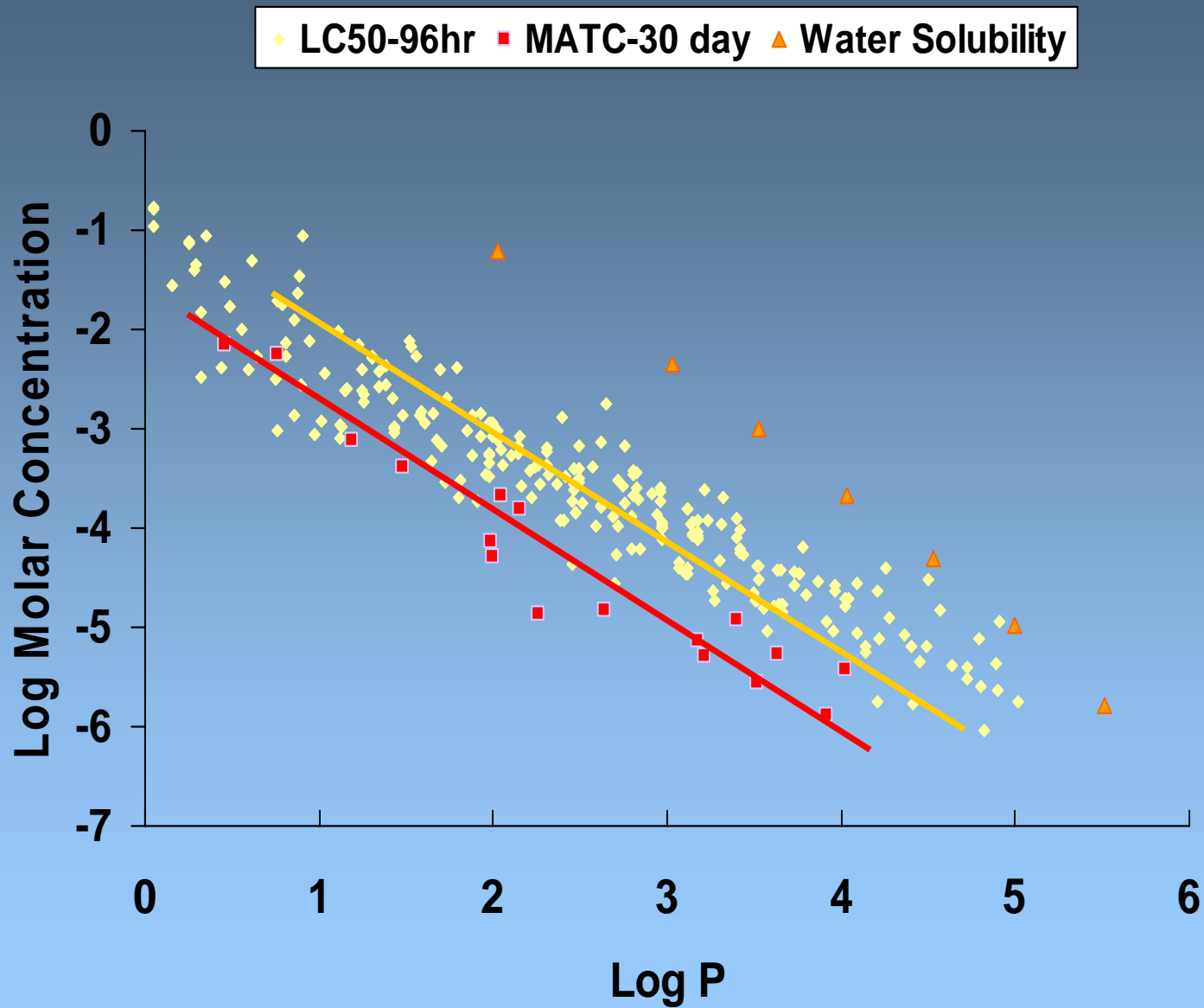
Log Fahead Minnow Molar Toxicity (1/LC50)



MOA assigned based upon biology:

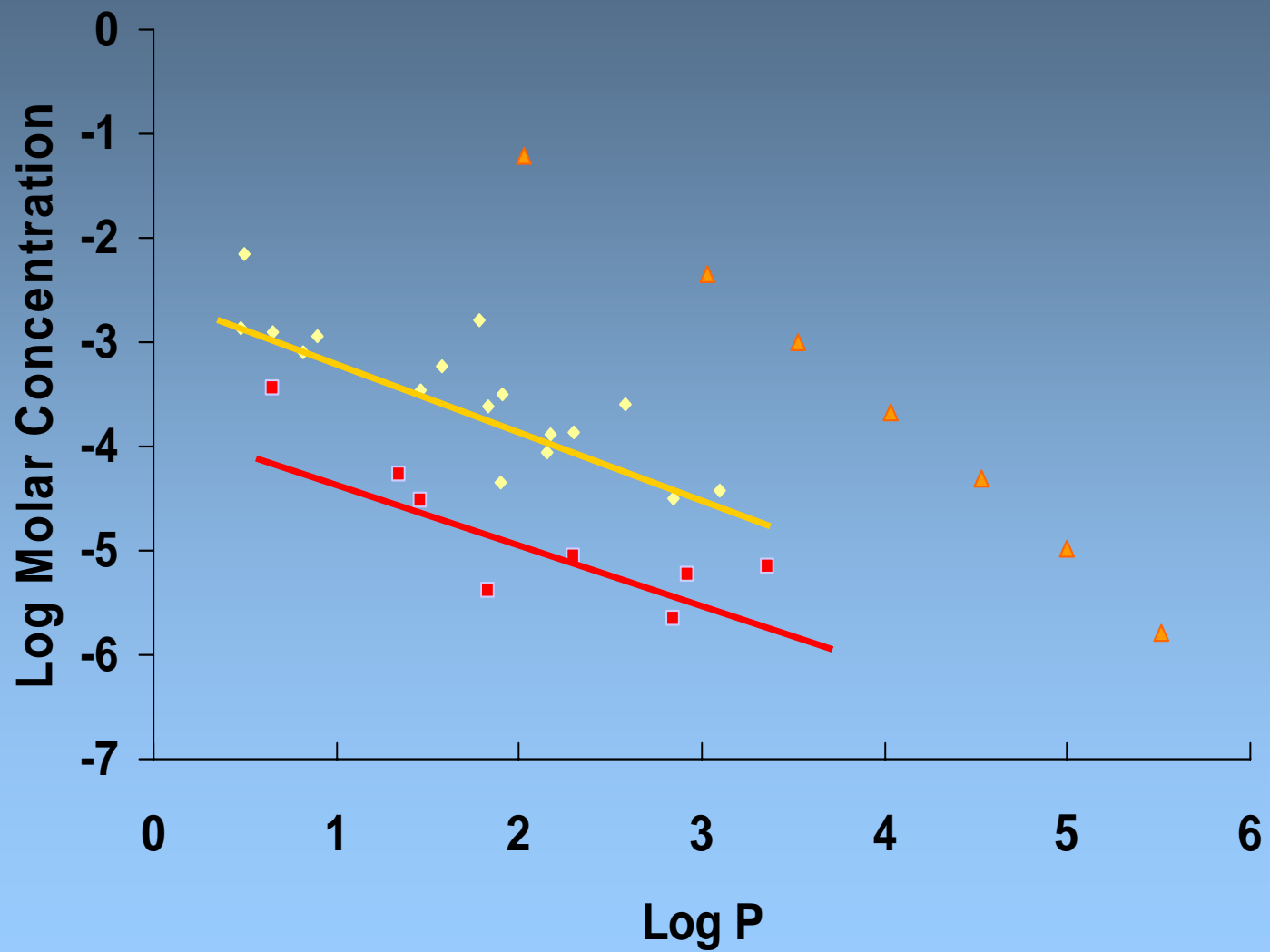
- Behavioral Observations - 96 h LC50
- Analysis of Conc-Rsp Curves - 96 h LC50
 - Ex) $24\text{h LC50}/96\text{h LC50} = 1$ (Narcosis I)
- Fish Acute Toxicity Syndromes
 - Respiratory/cardiovascular responses
- Joint Toxic Action Studies
 - Binary up to 24 chem mixtures
- Literature

Nonpolar Narcotic Toxicants

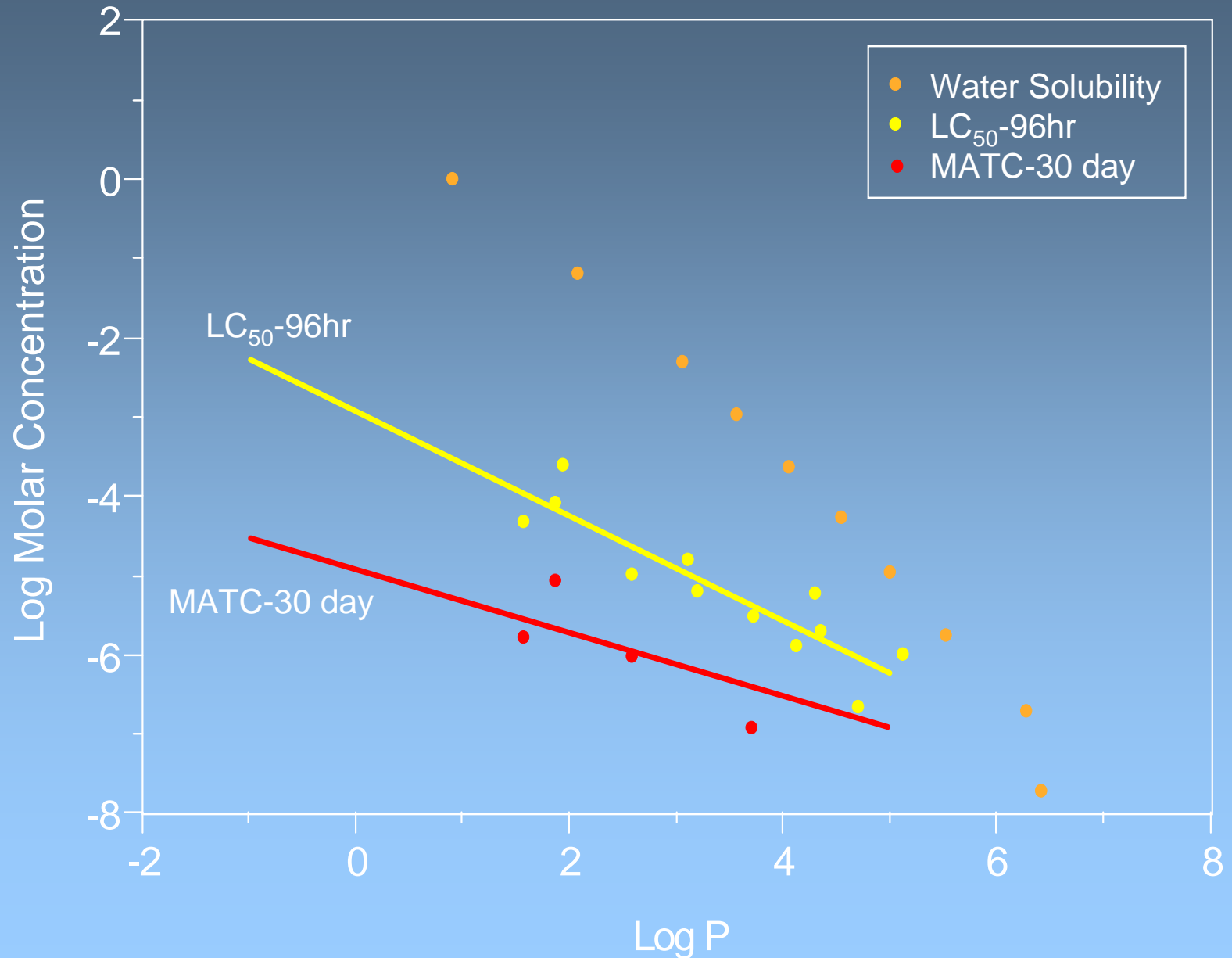


Polar Narcotic Toxicants

◆ LC50-96hr ■ MATC-30 day ▲ Water Solubility

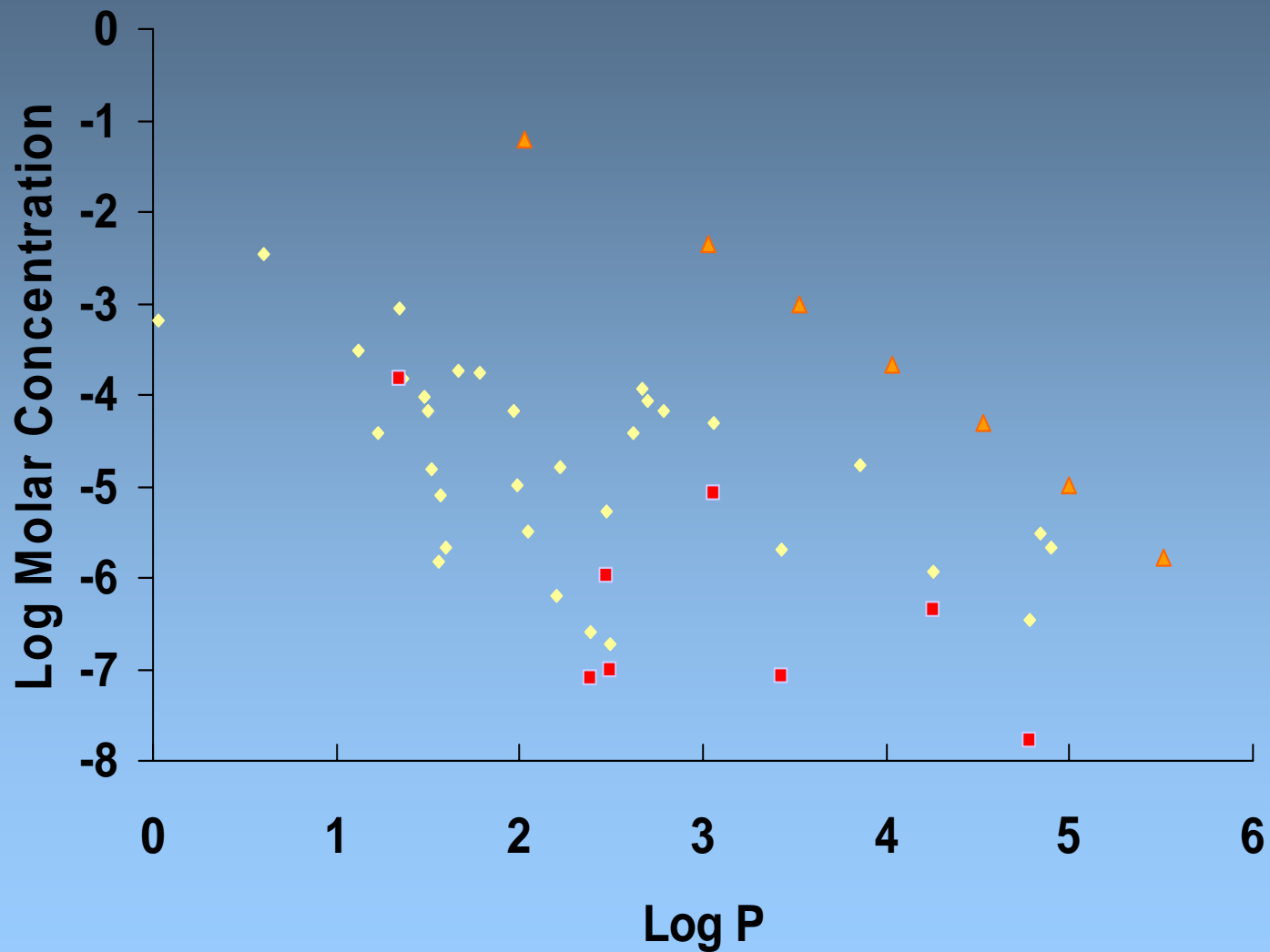


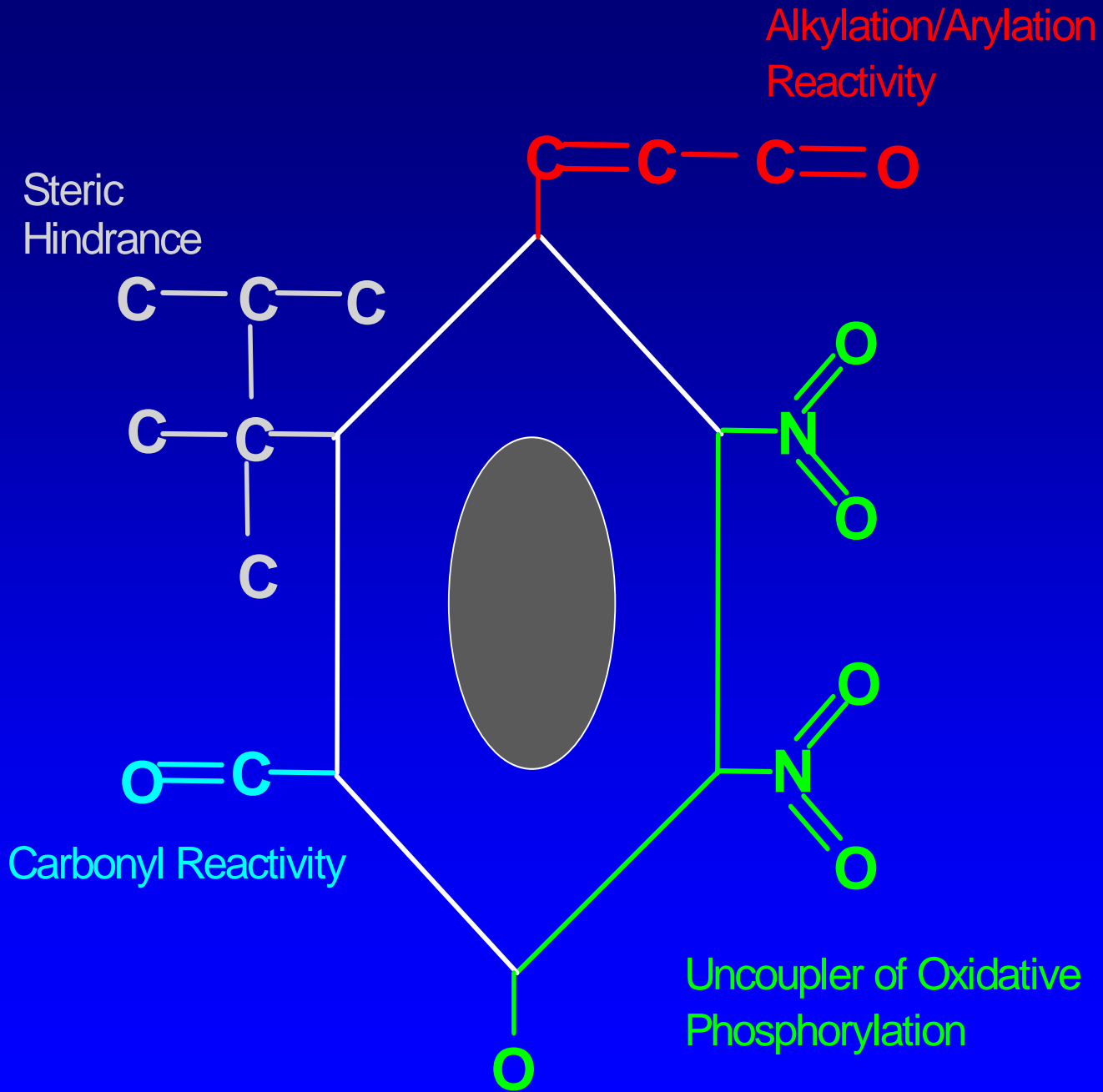
Oxidative Phosphorylation Uncouplers



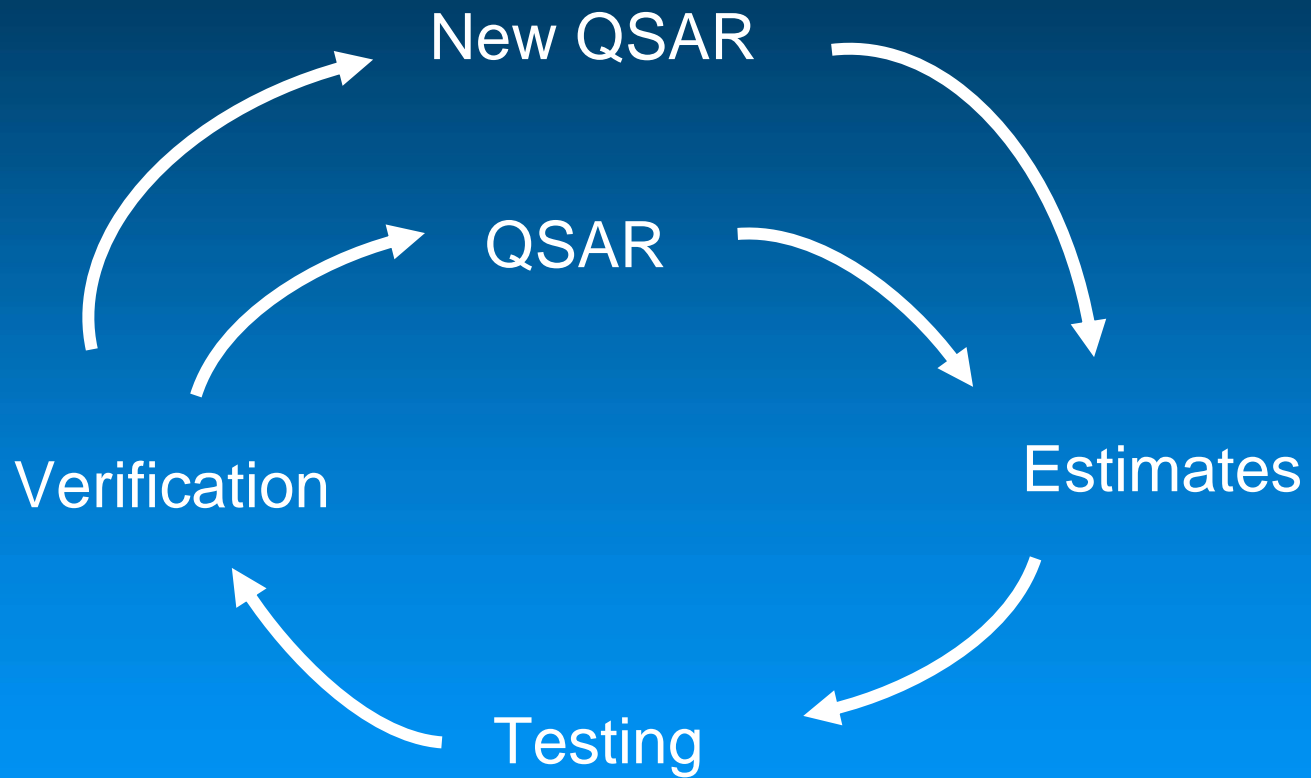
Reactive Toxicants

◆ LC50-96hr ■ MATC-30 day ▲ Water Solubility





QSAR Iterative Process



Chemical similarity is defined in the context of biological similarity

in vivo -

- acute aquatic tox (water expo)/inhal tox – steady-state; water conc/vapor conc represents internal conc for MOAs where Log P is sufficient for potency prediction;
- longer term chronic endpoints with complex toxicokinetics and toxicodynamics are not readily amenable to QSAR potency predictions – analog techniques;

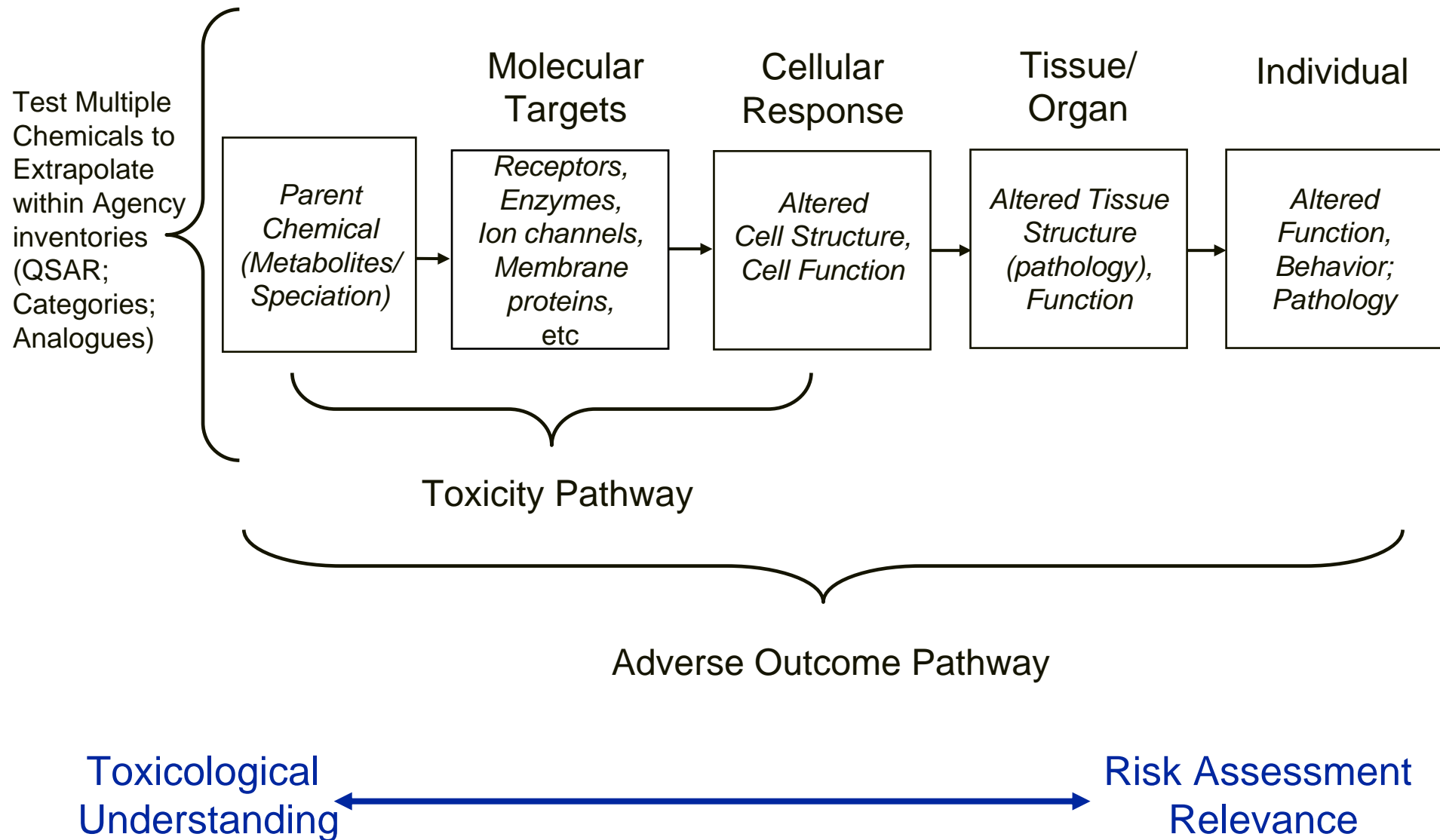
Toxic potency is correlated to chemical concentration at the site of action

-C. Hansch

in vitro -

- closer to site of chem/biol interaction so good place for QSAR focus, but requires linkage to adverse effect
- acceptable uncertainty (e.g., linkage in vitro to adverse outcome) is defined by risk assessors and depends upon the risk question, regulatory authority, data poor vs data rich chemicals, etc

Adverse Outcome Pathway and Chemical Extrapolation (QSAR)

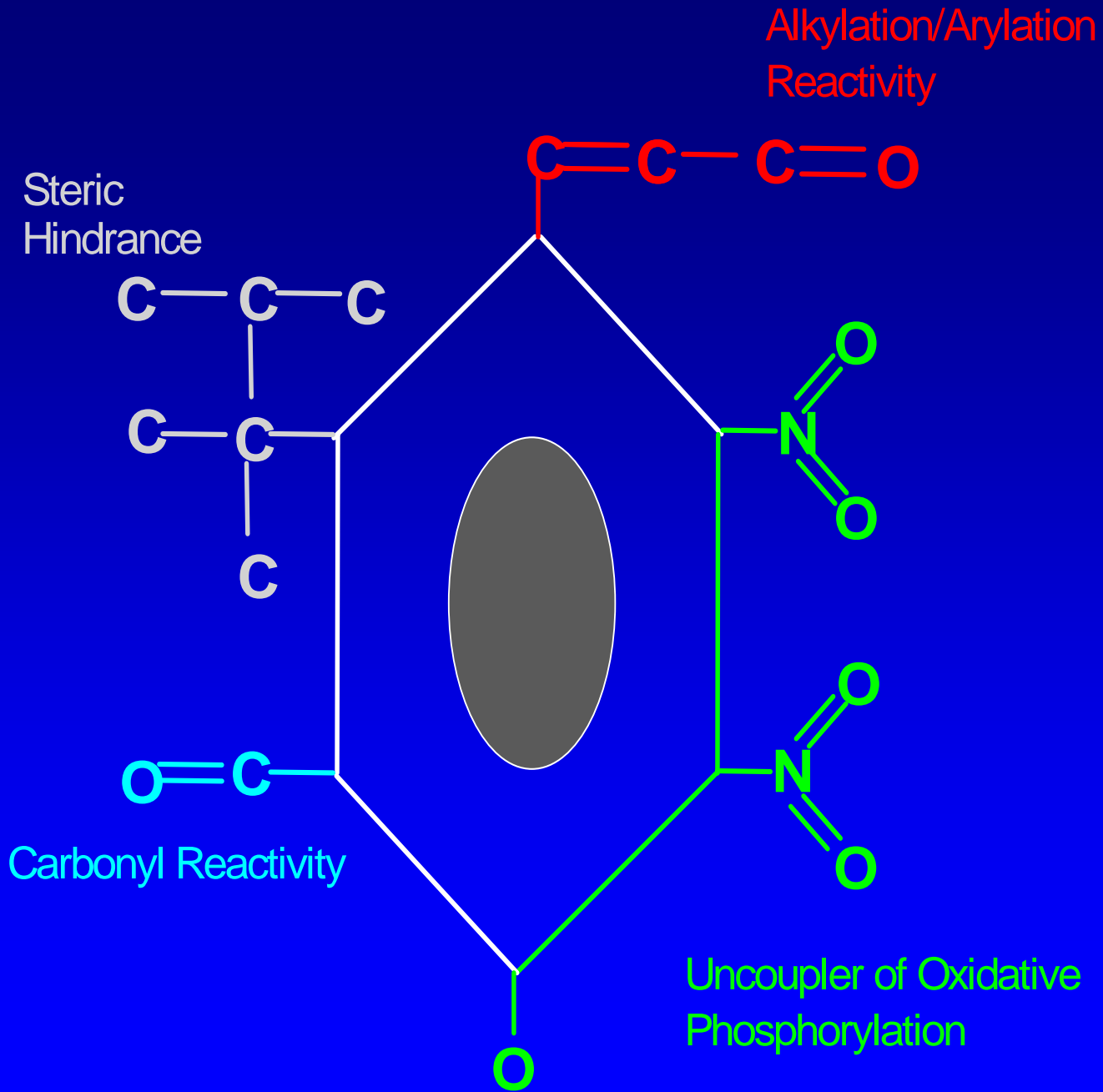


Risk Assessment Roles for QSAR

- **Developing Chemical Categories**
 - clustering chemical behavior for interpolation
- **Classification and Prioritization**
 - broad hazard identification profiles (PBT)
- **Hypothesis-Driven Strategic Testing**
 - sequential testing to minimize **unused** test data
- **Estimated Values for Untested Chemicals**
 - extend the universe of initial risk assessments

Structural Alerts

- Identification of atom configurations which generally will lead to specific behavior
- Created by distilling “rules of thumb” from compilations of measured test data
- Rules based on chemical substructures should be considered hypotheses
- All “rules of thumb” have important exceptions outside specific domains



Expert Systems

- Describing chemical behavior globally requires numerous “rules” and models
- Substructure rules can be arranged in a hierarchy of nested “If..Then..Else” queries
- Computers are ideally suited to executing complex hierarchies objectively
- Expert judgement can be encoded in the hierarchy to use rules and models properly

QSAR Predictions

- Predictions from QSAR actually involve two different aspects
- The most important is the prediction of the **mechanisms** underlying the behavior
- The second is the prediction of the intensity of the predicted behavior
- Because different parameters are needed, a single statistical model is seldom robust

QSAR Assumptions

- Chemical similarity (e.g., analogues) is defined in the context of biological similarity (i.e., analogous biological activity)
 - “Similar” chemicals, by definition, invoke the same toxicity pathway (within a specified biological model)
 - (Q)SARs are developed from a known or hypothesized “mode/mechanism” of action; hypothesis is tested to refine the models