

Data – Experimental (Measured) and Predicted



**Finding Experimental Data
and
Information on (Q)SARs
and SMILES**

Experimental Data

Reliable Data are Essential

- ✚ **Physical and chemical properties**
 - Leads to conclusions on fate and exposure
- ✚ **Environmental fate**
 - Defines persistence and potential exposures
- ✚ **Toxicity**
 - Determines the nature of the hazard
- ✚ **Exposure**
 - Risk is dependent on the dose

Measured or Estimated Data?

- + Reliable measured / experimental data are always preferred**
- + Evaluated estimates are used when experimental data are not available**
 - Analogs can be used to evaluate reliability of estimates**
 - Judgment may be required to determine if an estimate is reasonable**

Sources of Experimental Data

 **Scientific literature (journal articles)**

 **Handbooks**

 **Encyclopedic sources**

 **Chemical catalogs**

 **Commercial databases**

 **On-Line Sources**

Other Data Sources

The P2 Framework manual

- Contains an extensive compilation of experimental data sources indexed by end point

Bench chemists within your company!!

Estimation programs and methodologies of the P2 Framework

Identifying Analogs

Why Use Analogs?

- ✚ To get relevant data when there are no experimental studies on the compound of interest
- ✚ To evaluate the reliability of estimates
- ✚ Analogs are used to assess:
 - Physical and chemical properties
 - Environmental fate data
 - Toxicity

Attributes of Good Analogs

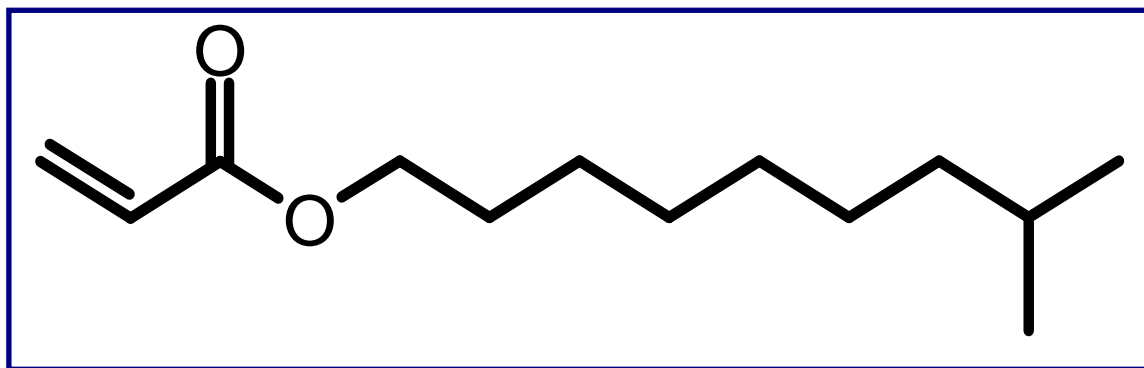
- ✚ **Good analogs are similar in the following aspects:**
 - **Size**
 - **Functional and substituent groups**
 - Same groups present (as many as possible)
 - Arrangement of groups
 - **Skeletal structure**
 - Identical arrangement of atoms in the backbone
 - **Physical properties**
 - **Biological mechanism of action**
- ✚ **Minimal number of features not present in the compound of interest**
- ✚ **Different analogs may be used for different purposes within a SF assessment**

Quality of Analogs

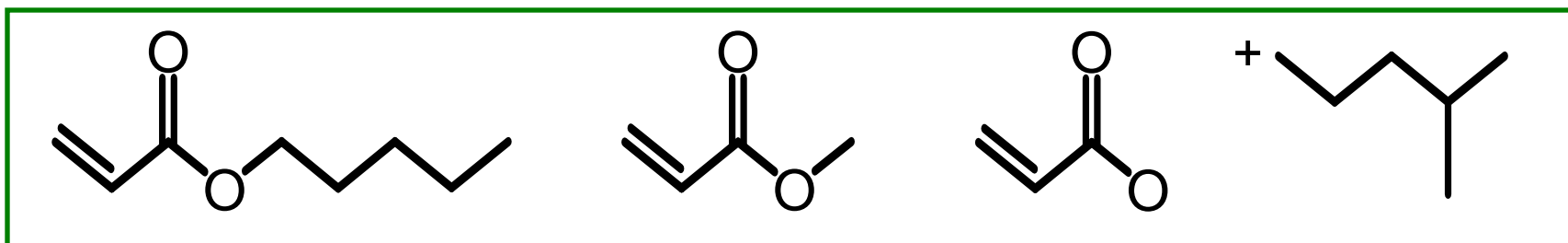
- + No analog is better than a bad one**
 - Bad analog -- provides misleading information**
 - No analog -- rely on estimation methods**
- + The best analogs may not have experimental data**

Example Analog Search

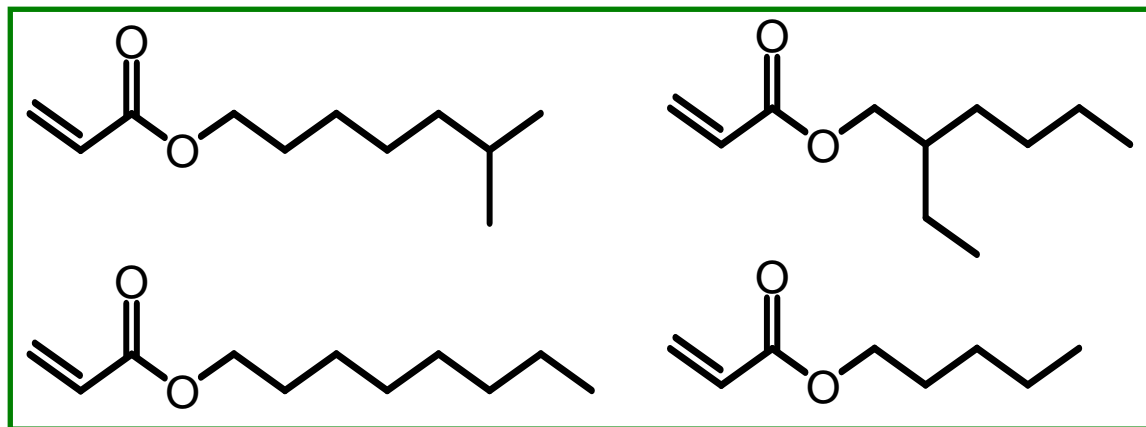
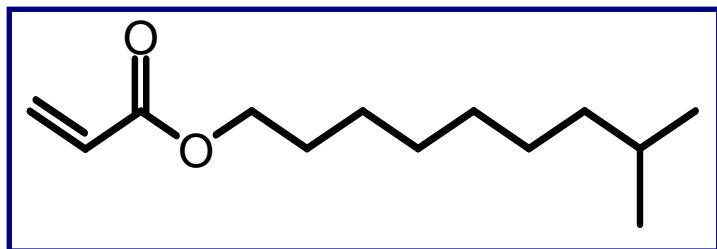
Example Chemical



Substructure search strategies

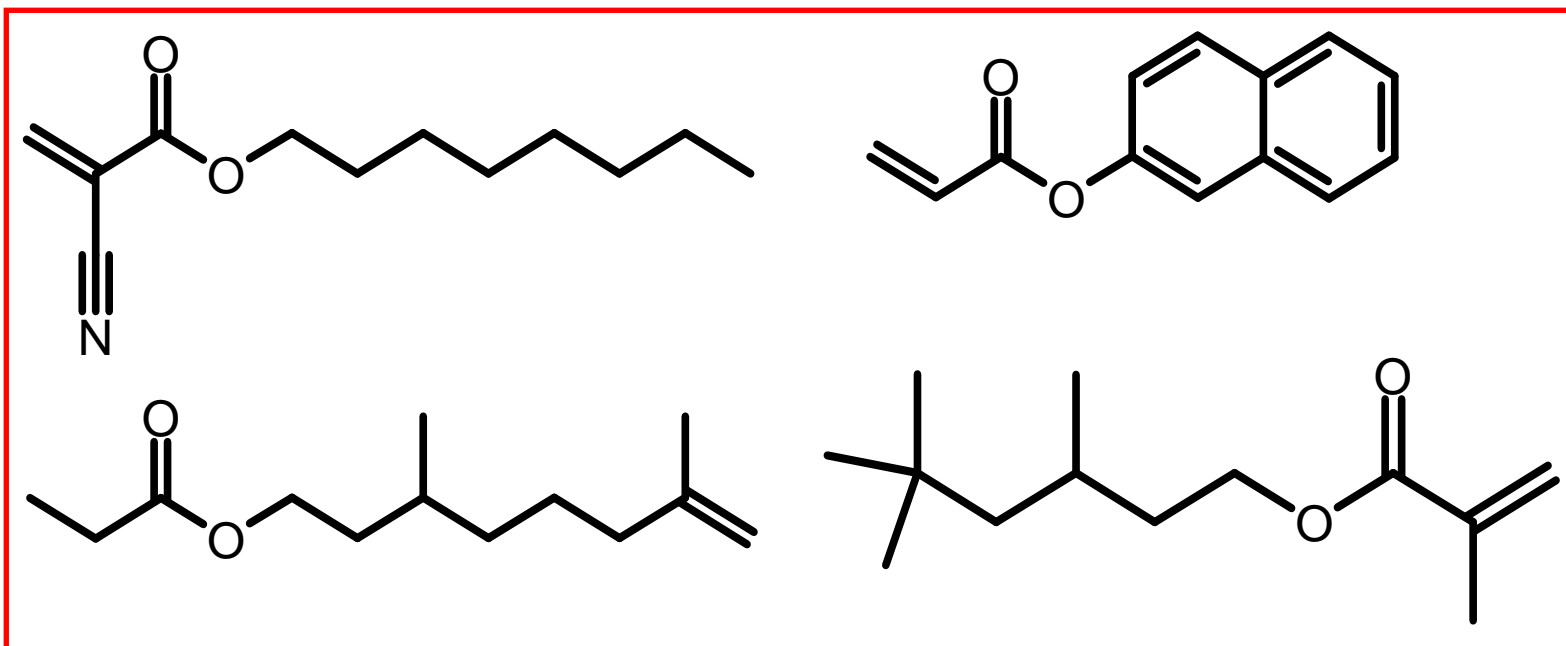


Good Analogs



- ✚ Contain similar functional and substituent groups
- ✚ Are of similar size
- ✚ Have similar biological mechanisms of action
- ✚ Do not contain significant features that are not present in the compound of interest

Poor Analogs



- + Have different functional or substituent groups
- + Have a different arrangement of functional groups
- + Have significant features that are not present in the compound of interest
- + May have different biological mechanisms of action

Substructure Searching Hints

- ✚ Too many hits in the substructure search?
 - Increase the amount of detail in the substructure
- ✚ Too few hits in the substructure search?
 - Reduce the amount of detail in the substructure
- ✚ Be wary of “percent similarity” results from search engines to choose analogs
 - Only provides what molecules have in common, not what they *do not* have in common

In Summary

 **Analog searching and selection is as much an art as a science**

- **Must use professional judgment**

 **Help is available!**

- **The P2 Framework**
- **SF P2 Partnerships**
- **Analog Identification Methodology (AIM) discussed in a later presentation**

Introduction to SAR/QSAR (Quantitative) Structure Activity Relationships

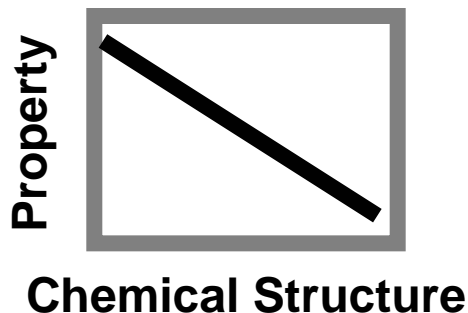
Introduction to SAR/QSAR

✚ Structure Activity Relationship (SAR)

- The relationship between the chemical structure of a molecule and its activity

✚ Quantitative Structure Activity Relationship (QSAR)

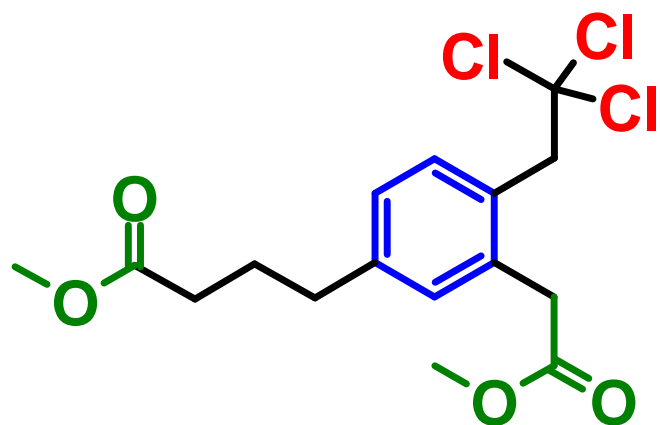
- Provides numeric value for the relationship



QSAR Equations

✚ Using a physical chemical property

✚ Using molecular descriptors

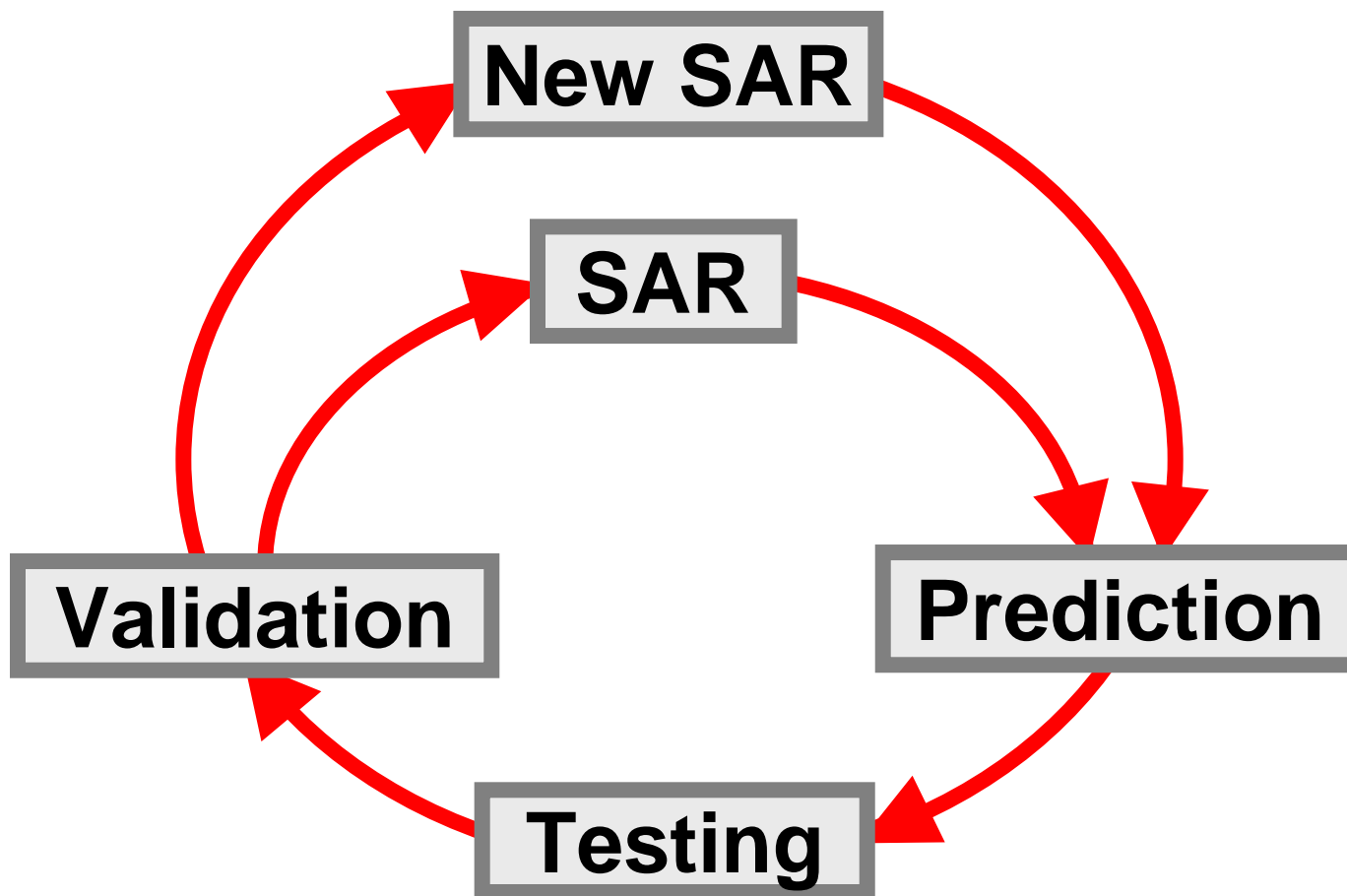


Fragment	#	Value
Aromatic	1	2.2
Ester	2	4.4
Chlorine	3	-1.1

Expert Systems

- ✚ **A computerized system to mimic the thinking and reasoning of human experts**
 - **Formalized, codified, and organized system of structure-activity relationships**
 - **Rules based**
 - **Provide qualitative results**
 - **Low, moderate, or high**
 - **Certain, probable, or improbable**

SAR/QSAR Development

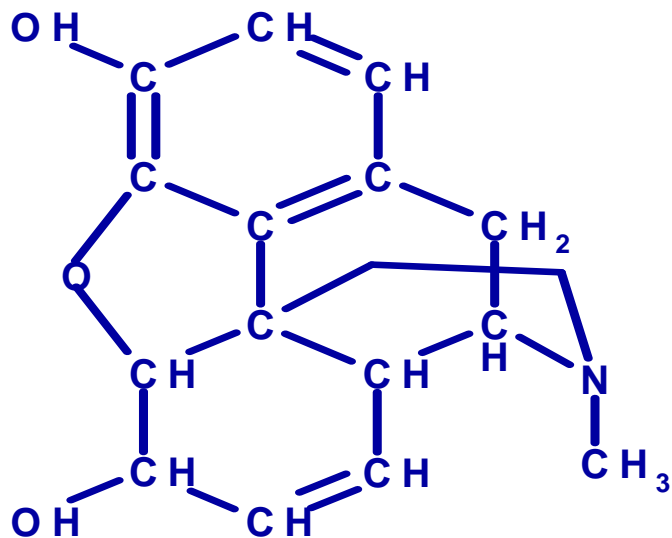


SMILES (Simplified Molecular Input Line Entry System)

SMILES

Structures will need to be converted to SMILES notation for structural entries into various QSAR programs

The purpose of SMILES is to go from this...



..... to this:

Oc1ccc2CC(N3C)C4C=CC(O)C5Oc1c2C45CC3

SMILES – Help is Available!!

- SMILES Guides
 - P2 Framework Manual, Appendix C
 - EPI Software Help System
 - SMILES Tutorial
www.daylight.com/dayhtml/smiles/
- Chemical drawing programs that produce SMILES (cut and paste)
 - ChemDraw
 - ChemSketch (Free version at www.acdlabs.com/download/chemsk.html)
- EPI imports “mol” files (and other popular formats) produced from drawing programs (ISISDraw, ChemDraw, ChemSketch, Alchemy, HyperChem, & Molecular Presentation Graphics, etc.) and gives SMILES in results screen