



# Computational Chemistry

---

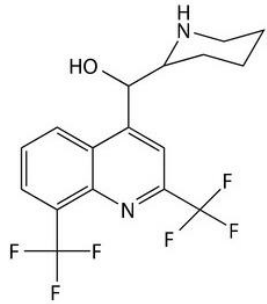
HOW DO COMPUTERS KNOW  
CHEMISTRY?

# Content

---

- ❖ Structure and function
- ❖ Mathematics in Chemistry
- ❖ Quantitative Structure Activity Relationship (QSAR)
- ❖ Components of QSAR
- ❖ Types of QSAR
- ❖ Molecular Descriptors

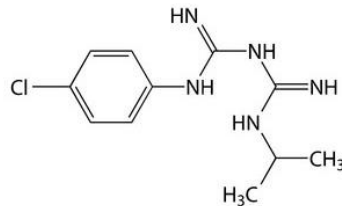
# Structure and function



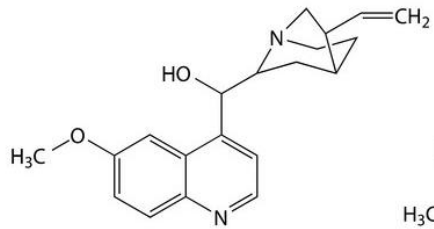
Mefloquine



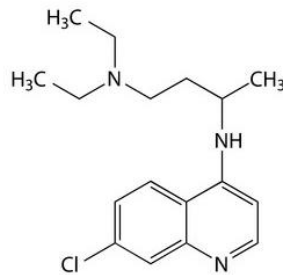
Primaquine



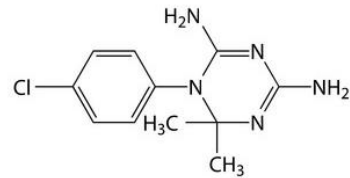
Proguanil



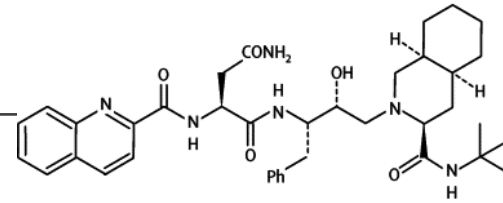
Quinine



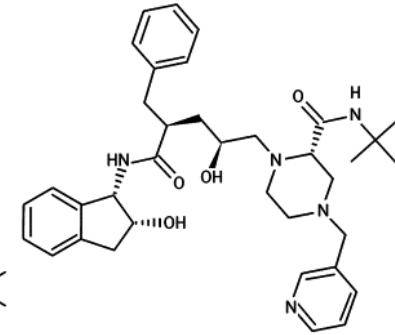
Chloroquine



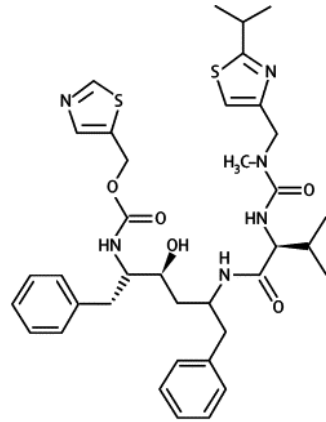
Cycloguanil



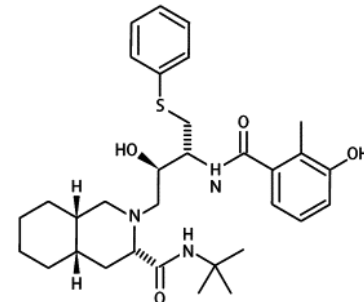
Saquinavir



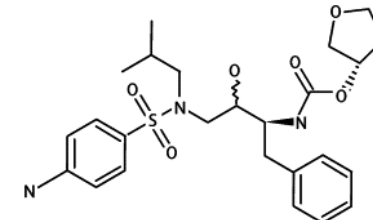
Indinavir



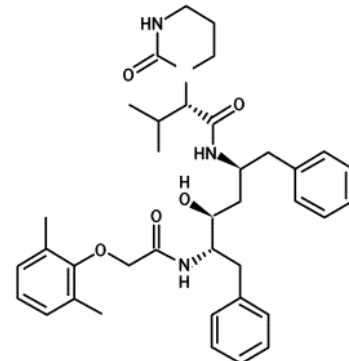
Ritonavir



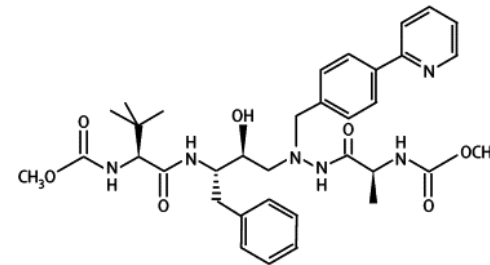
Nelfinavir



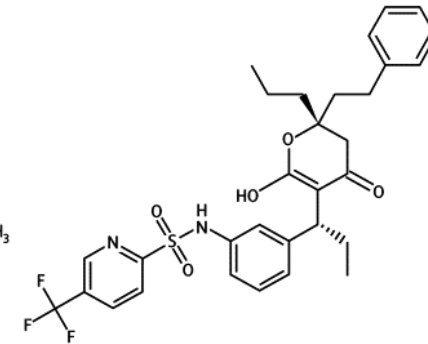
Amprenavir



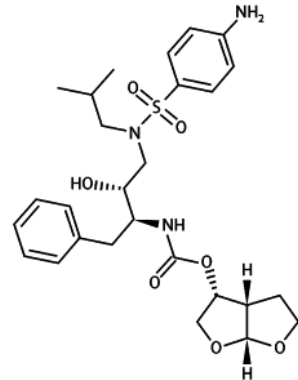
Lopinavir



Atazanavir



Tipranavir



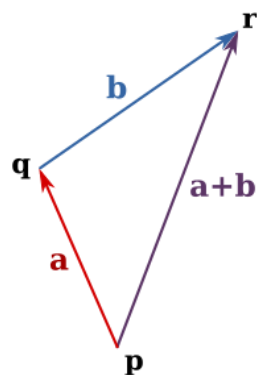
Darunavir

# Mathematics in Chemistry

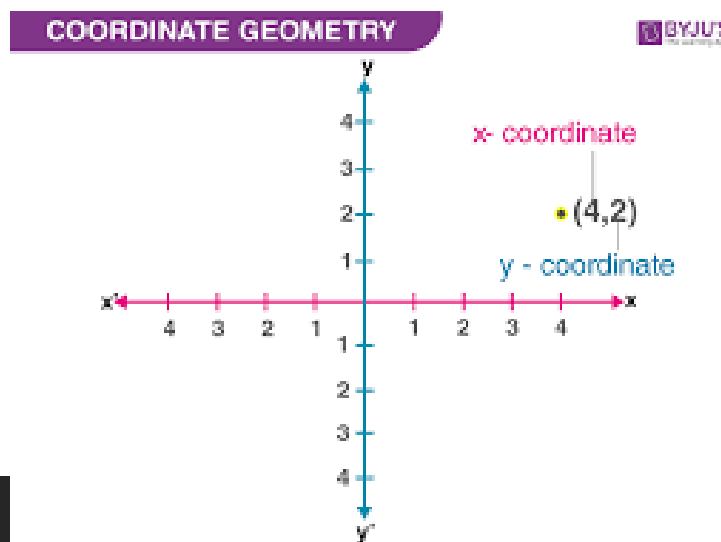
□ Linear Algebra (Matrices)

$$\begin{matrix} & \begin{matrix} 1 & 2 & \dots & n \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ \vdots \\ m \end{matrix} & \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ a_{31} & a_{32} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \end{matrix}$$

□ Vectors Algebra



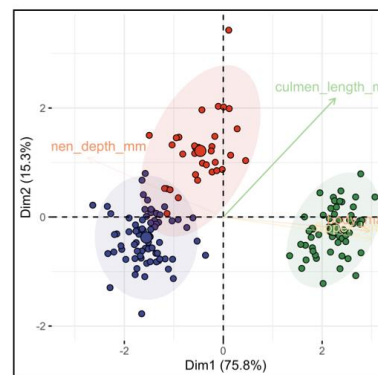
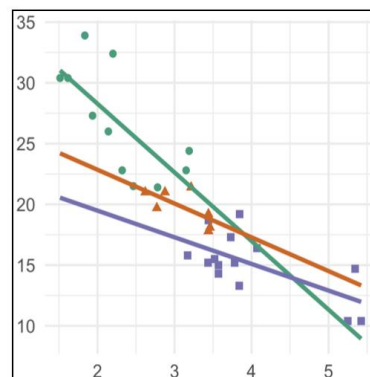
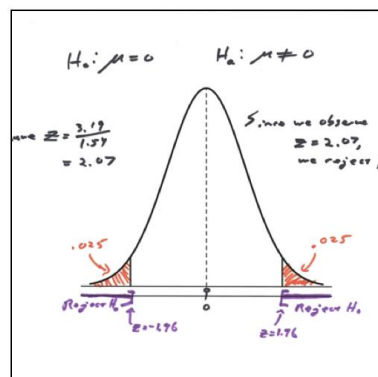
□ Coordinate Geometry



# Mathematics in Chemistry

Calculus: Derivatives and Integrations

Statistics



essential test satisfying (6) our unbiased estimator

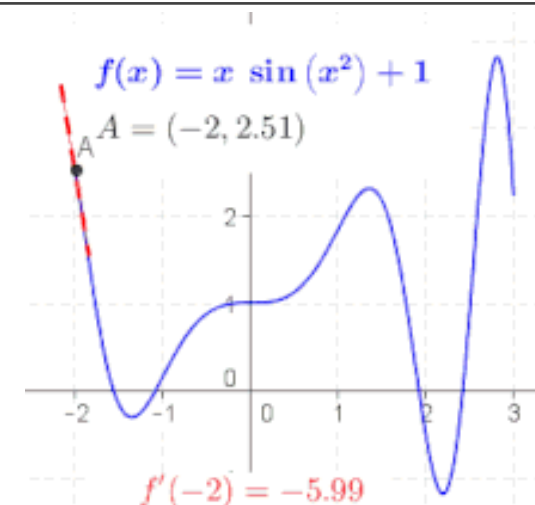
$$v = \frac{E(h_{i-1}x_1 | u_i = u)}{E(h_{i-1} | u_i = u)} = \frac{E(h_{i-1}x_1 f)}{E(h_{i-1} f)}$$

$u_i = u$ , where  $f$  is the characteristic function

$$v = \frac{\sum_{j=1}^{\infty} j k_j(u, i)}{\sum_{j=0}^{\infty} k_j(u, i)} \quad \text{for Poisson}$$

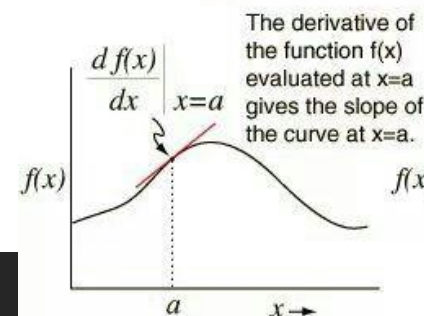
$$v = \frac{k_1(u, i)}{\sum_{j=0}^{\infty} k_j(u, i)} \quad \text{for binomial}$$

$i$  denotes the number of possible sequences  
 $\dots + x_i = u$ , and  $x_1 = j$ . For the binomial case



Derivative

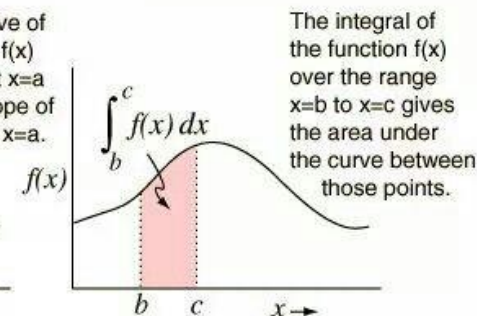
$$\frac{df(x)}{dx}$$



The derivative of the function  $f(x)$  evaluated at  $x=a$  gives the slope of the curve at  $x=a$ .

Integral

$$\int f(x) dx$$

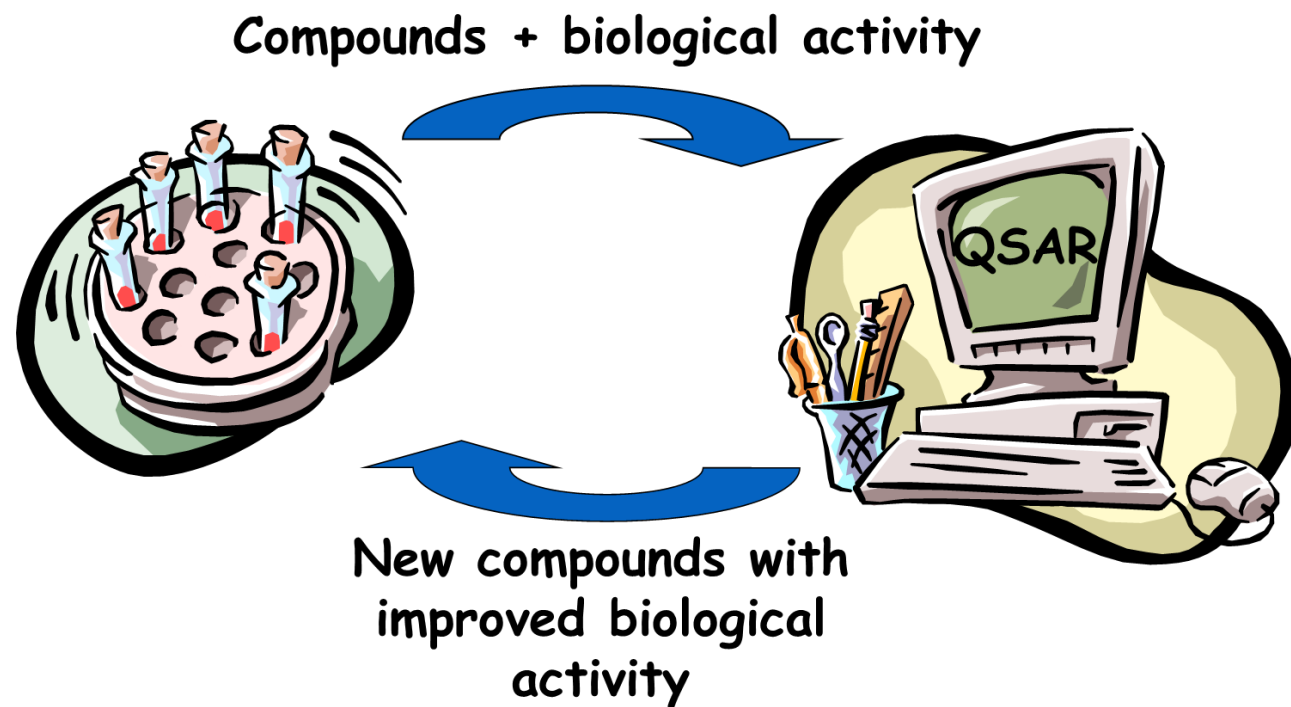


The integral of the function  $f(x)$  over the range  $x=b$  to  $x=c$  gives the area under the curve between those points.

# Quantitative Structure Activity Relationship

---

- ❑ The number of compounds required for synthesis in order to place 10 different groups in 4 positions of benzene ring is  $10^4$ .
- ❑ Solution: synthesize a small number of compounds and from their data derive rules to predict the biological activity of other compounds



# Components of QSAR

**C** = molar concentration  
that causes a certain  
biological effect

values of the  
regression  
coefficients

95% confidence intervals  
of the coefficients and  
the constant term

---

$$\text{Log } 1/C = 1.15 (\pm 0.2) \pi - 1.46 (\pm 0.4) \sigma^+ + 7.82 (\pm 0.2)$$

logarithms of  
reciprocal values are  
the correct scaling

lipophilicity  
parameter

electronic  
parameter

constant  
term

$$(n = 22; r = 0.945; s = 0.196; F = 78.6; Q^2 = 0.841; s_{\text{PRESS}} = 0.238)$$

number of  
compounds

correlation coefficient  $r$ ;  
measure of the relative  
quality of a model

Fisher value;  
measure of the  
statistical significance

standard deviation  $s$ ;  
measure of the absolute  
quality of a model

standard deviation  
of crossvalidation  
predictions and  
squared crossvalidation  
regression coefficient;  
both are measures of  
internal predictivity

# Data Structure in QSAR

---

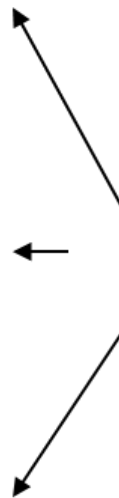
X-variable

Y-variable (activity data in QSAR)

2  
4  
1  
.  
.  
.



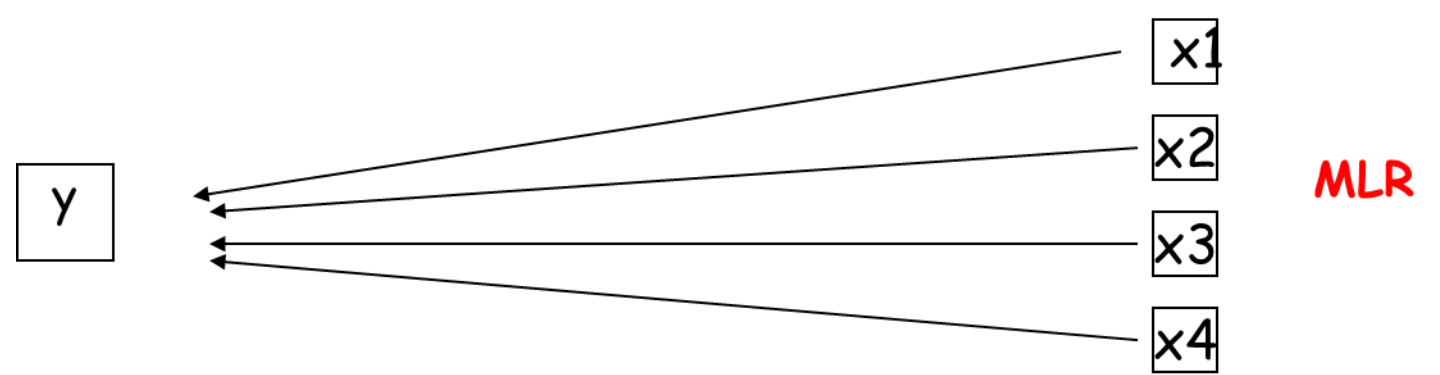
7  
6  
8  
.  
.  
.



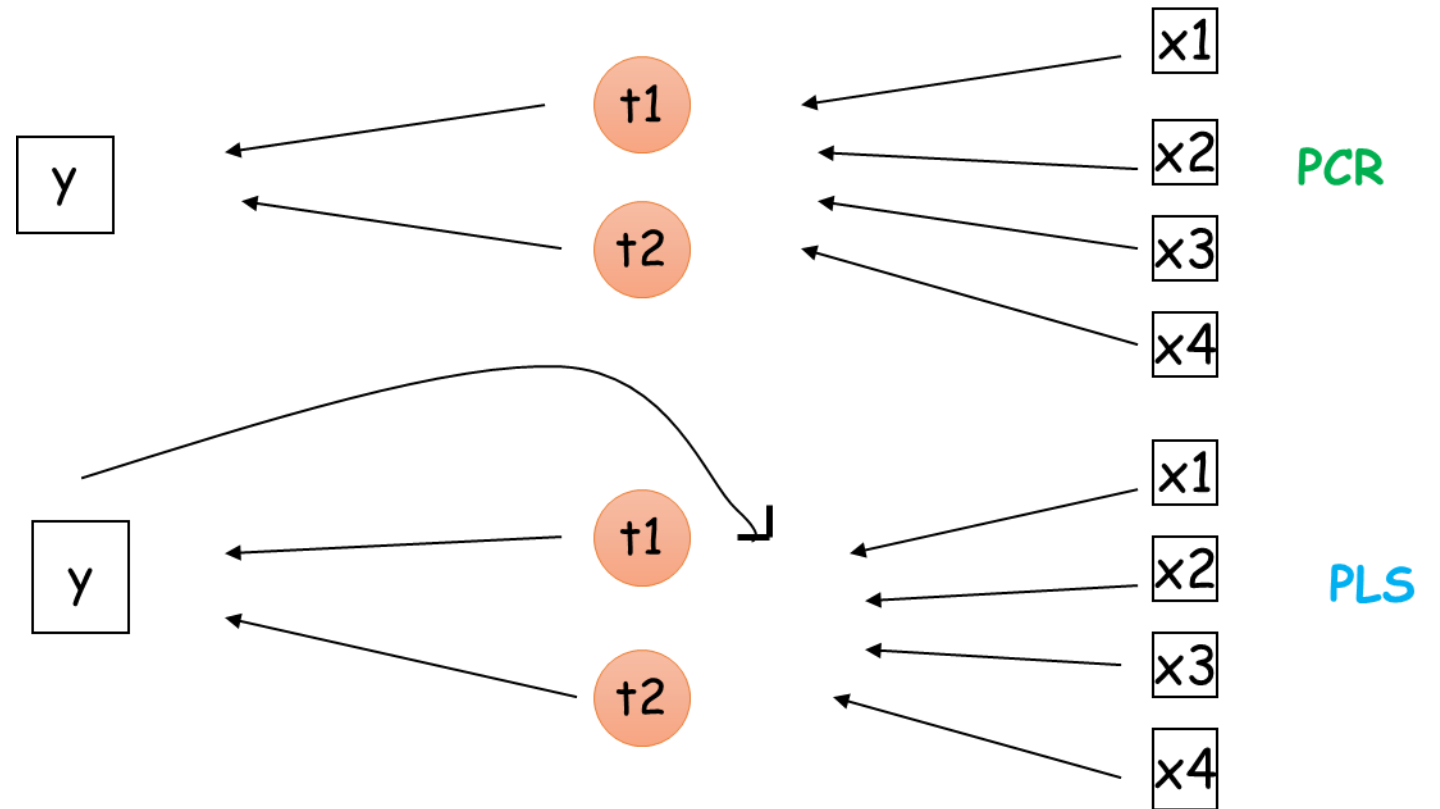
Objects, same number  
in x and y-column



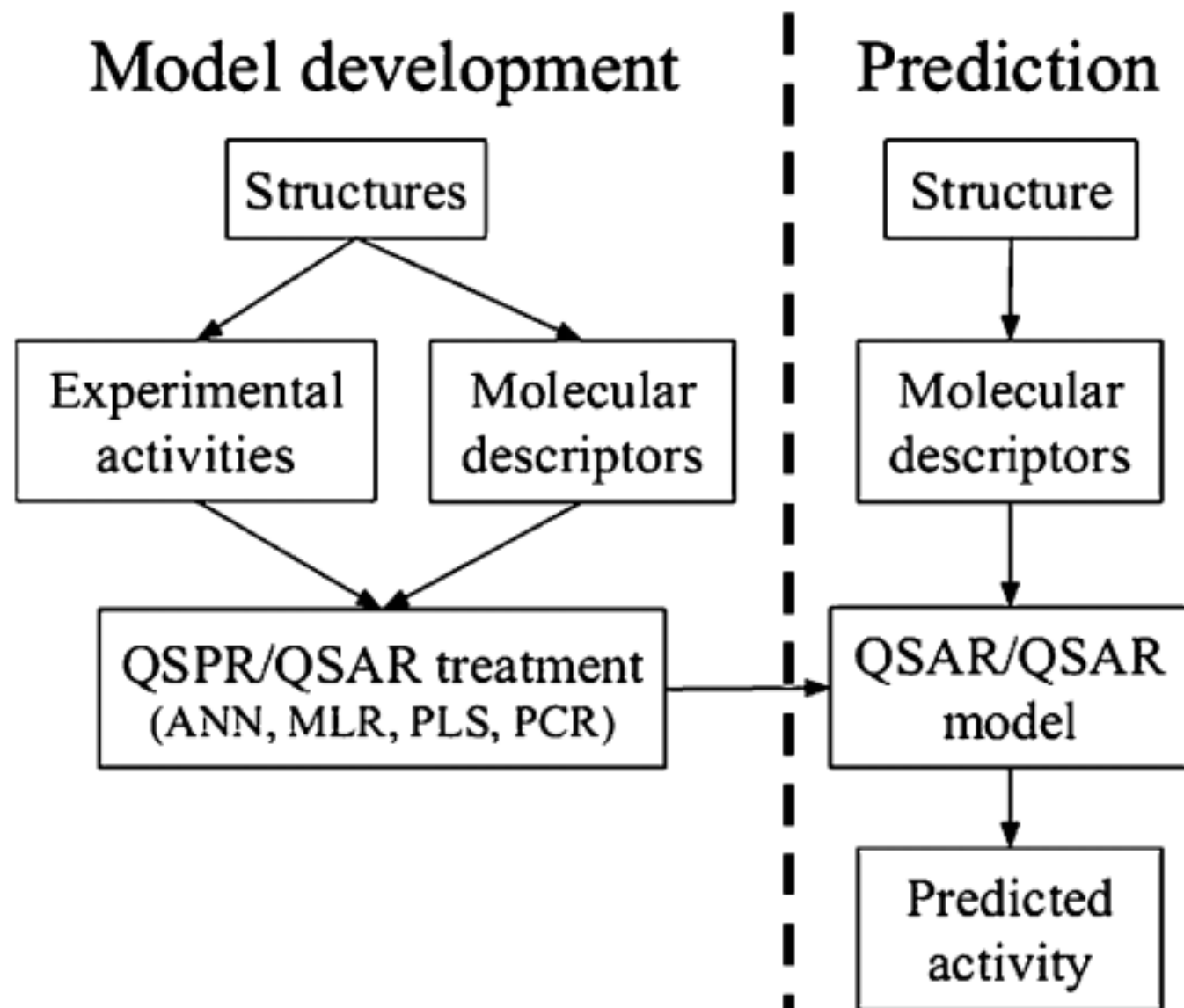
# Mathematics in QSAR



## Chemometrics in QSAR



# Steps in QSAR



# Dimensionality in QSAR

---

□ 1D-QSAR

□ 2D-QSAR

□ 3D-QSAR (Receptor-dependent and Receptor-independent)

□ 4D-QSAR (Receptor-dependent and Receptor-independent)

□ 5D-QSAR → Induced Fit effect

□ 6D-QSAR → Solvation effects

# Molecular Descriptors

Molecular Property	Corresponding Interaction	Parameters
Lipophilicity	hydrophobic interactions	$\log P$ , $\pi$ , $f$ , $R_M$ , $\chi$
Polarizability	van-der-Waals interactions	MR, parachor, MV
Electron density	ionic bonds, dipol-dipol interactions, hydrogen bonds, charge transfer interactions	$\sigma$ , $R$ , $F$ , $\kappa$ , quantum chemical indices
Topology	steric hindrance geometric fit	$E_S$ , $r_V$ , $L$ , $B$ , distances, volumes

# Type of Molecular Descriptors

---

- ❑ 0D-descriptors (i.e. constitutional descriptors, count descriptors)
- ❑ 1D-descriptors (i.e. list of structural fragments, fingerprints)
- ❑ 2D-descriptors (i.e. graph invariants)
- ❑ 3D-descriptors (i.e. quantum-chemical descriptors, size, steric, surface and volume)
- ❑ 4D-descriptors (i. e. GRID or CoMFA methods, Volsurf)

# What should a descriptor be like?

---

- ☐ Should have structural interpretation
- ☐ Should have good correlation with at least one property
- ☐ Should preferably discriminate among isomers
- ☐ Should be possible to apply to local structure
- ☐ Should possible to generalize to "higher" descriptors
- ☐ Should be simple

# What should a descriptor be like?

---

- ☐ Should not be based on experimental properties
- ☐ Should not be trivially related to other descriptors
- ☐ Should be possible to construct efficiently
- ☐ Should use familiar structural concepts
- ☐ Should change gradually with gradual change in structures
- ☐ Should have the correct size dependence, if related to the molecule size