Rational Drug Design lecture 11

Łukasz Berlicki

QSAR i 3D QSAR

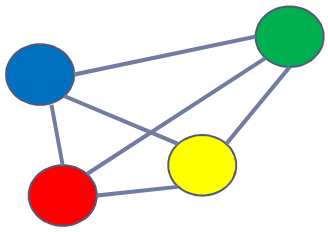
- **1964, C. Hansch i T. Fujita:** QSAR
- **1984, P. Andrews:** functional group contributions to interactions with molecular target
- **1985, P. Goodford:** GRID (binding points on protein surface)
- **1988, R. Cramer:** 3D QSAR
- **1992, H.-J. Böhm:** LUDI, interactions, docking, scoring functions
- **1997, C. Lipinski:** bioavailability, rule of five
- **1998, Ajay, W. P. Walters and M. A. Murcko; J. Sadowski i H. Kubinyi:** drug-like compounds

pharmacophore

- **pharmacophore**– a set of steric and electronic features that are necessary for optimal interaction with the molecular target.
- The pharmacophore is not a real molecule but a purely abstract creation.
- **Pharmacophoric descriptors** are places of hydrophobic, electrostatic and hydrogen bonding, which are defined as spheres or points.

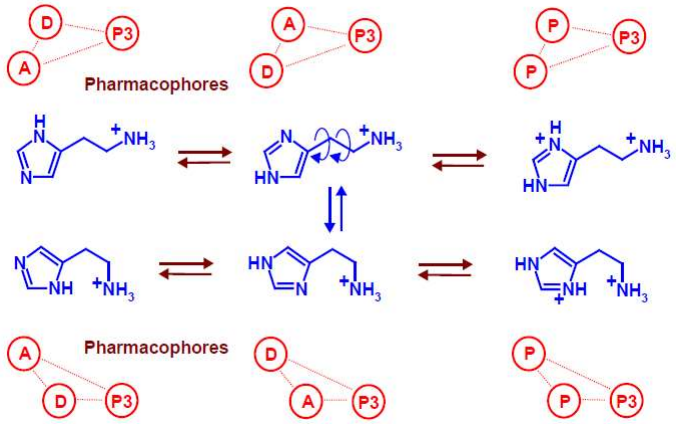
pharmacophore

- The location of individual spheres defining the pharmacophore features is defined by the distances between them.
- The pharmacophore is obtained by analysis of structures of active compounds and indicating their common features.
- The activity of new compounds is assessed by applying them to the pharmacophore.



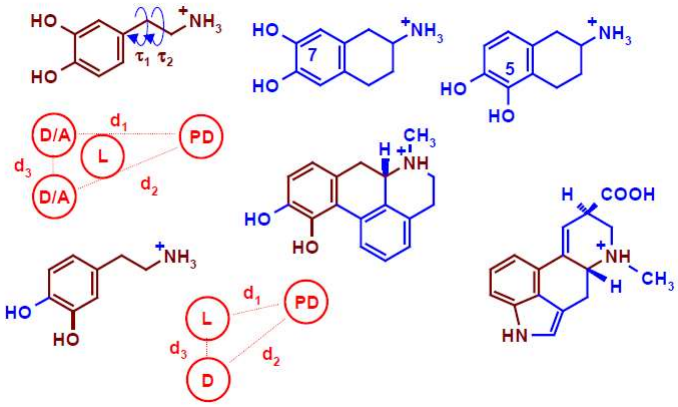
pharmacophore

- It is not possible to develop a pharmacophore based on one active structure.



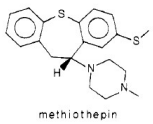
pharmacophore

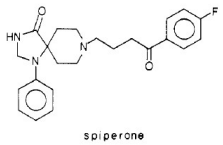
- Comparison of various active structures gives the opportunity to construct a good pharmacophore.

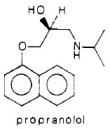


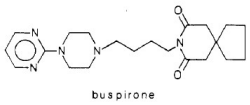
Antagonists of the 5-HT_{1A} receptor

- Known serotonin receptor antagonists and their activity.
- Most preferably, compounds with low conformational freedom should be used.





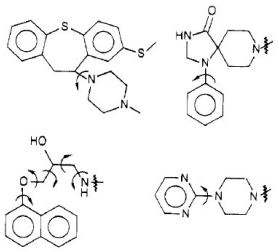


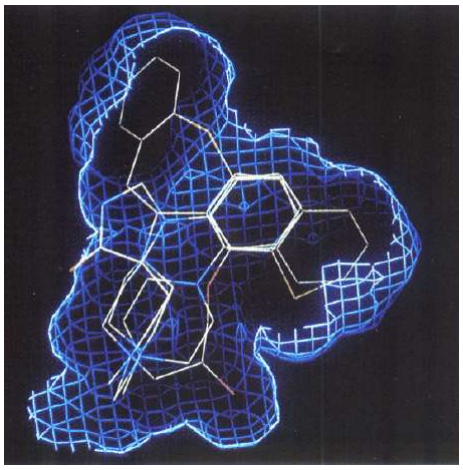


compound	pIC_{50} 5-HT_{1A}	5-HT_{1B}	5-HT_2
8-OH-DPAT	8.52	5.42	5.00
(−)-methiothepin	7.02	6.74	8.20
(+)-methiothepin	6.07	5.49	8.25
spiperone	6.91	6.00	8.67
propranolol	6.77	6.31	5.10
buspirone	7.66	4.90	5.47

Antagonists of the 5-HT_{1A} receptor

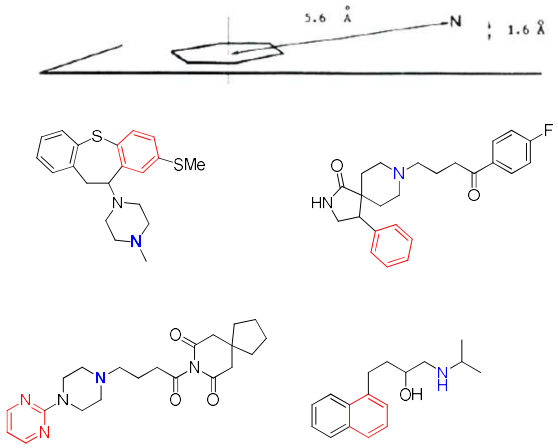
- Application of compounds with respect to conformational freedom





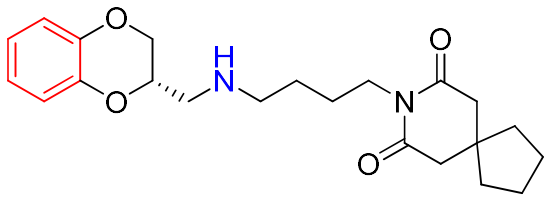
Antagonists of the 5-HT_{1A} receptor

- pharmacophore



Antagonists of the 5-HT_{1A} receptor

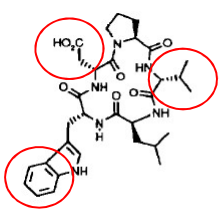
- Designing new compounds
- The obtained relationship is over 25 times stronger than the best of the training set.



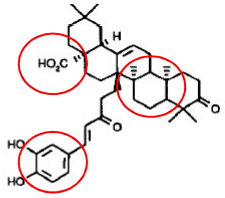
pIC_{50} = 9.2

Antagonists of the ET_A receptor

- Endothelin 1 is a peptide with vasoconstructive activity
- Endothelin receptor antagonists can be drugs for cardiovascular disease
- Design:
  - Compounds with known endothelin receptor antagonist activity.
  - The imposition of structures of active compounds indicates important structural elements.



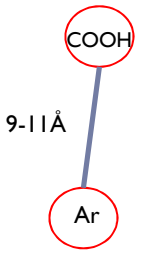
BQ 123 **1**

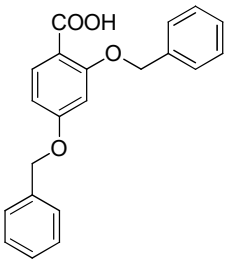


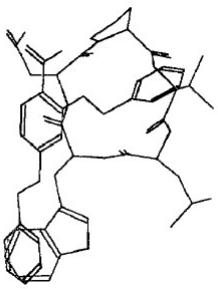
Myriceron caffeoyl ester **2**

Antagonists of the ET_A receptor

- Construction of a pharmacophore.
- Searching the database using a pharmacophore.



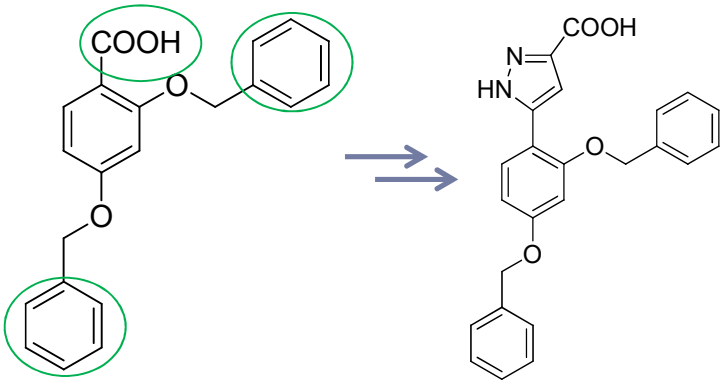




IC_{50} = 9 μM

Antagonists of the ET_A receptor

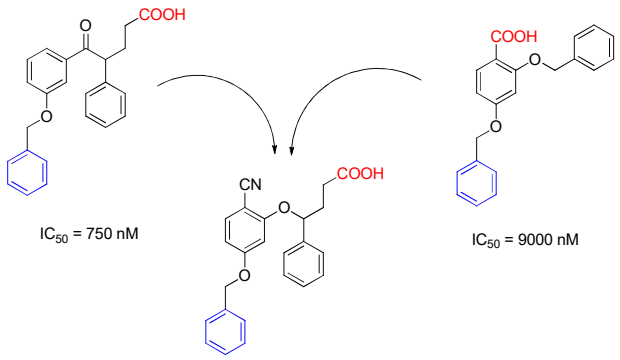
- Optimization of substituents



IC_{50} = 1 μM

Antagonists of the ET_A receptor

- The combination of structures found compounds gave a ligand with very high activity



3D QSAR - CoMFA

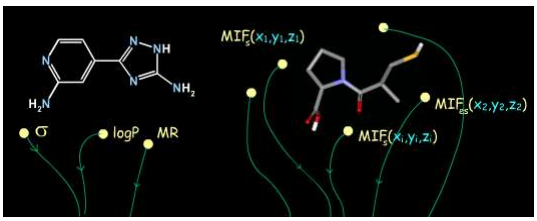
- **Comparative Molecular Field Analysis – CoMFA**

- Assumptions:
  - Activity directly related to structural properties
  - The compound does not form a covalent bond with the molecular target
  - The same mechanism of action and mode of binding to the molecular target

CoMFA

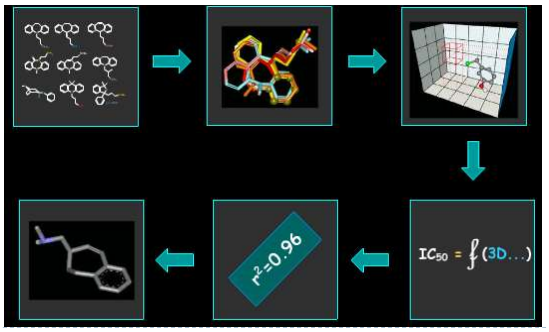
- QSAR:
  - Independent descriptors
  - Descriptors determine the properties of the compound
  - A relatively small number of descriptors

- 3D QSAR, CoMFA
  - Dependent descriptors
  - Descriptors determine field values near the compound
  - A very large number of descriptors



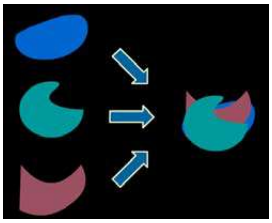
CoMFA

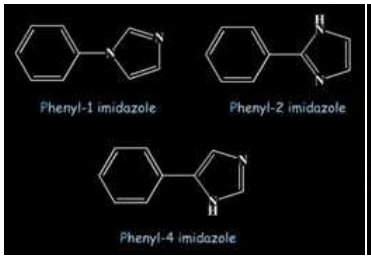
- Steps of 3D QSAR procedure
  - Proposition of the binding mechanism (structure of the active site?)
  - Compounds geometry
  - Superimposition of compounds and placing points in the grid
  - Calculation of interactions of molecules with probes at each point of the grid
  - Regression
  - Validation
  - predictions

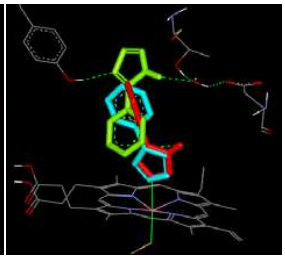


Superimposition of structures

- A good superimposition of active compound structures is key to obtaining the right result
- It is not always evident how to superimpose compounds.

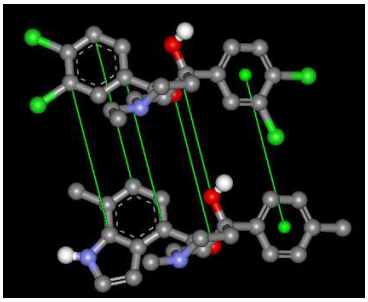


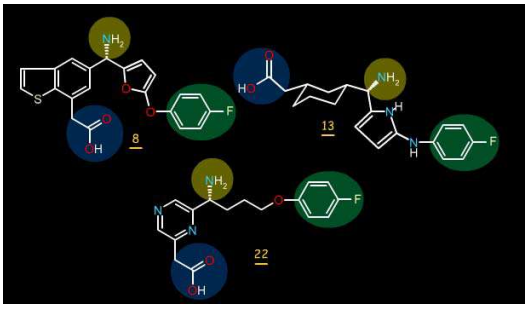




Superimposition of structures

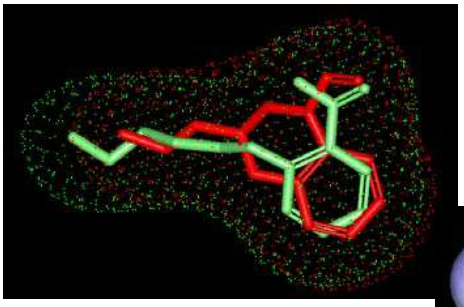
- Superimposition of the corresponding atoms
- Superimposition of pharmacophores

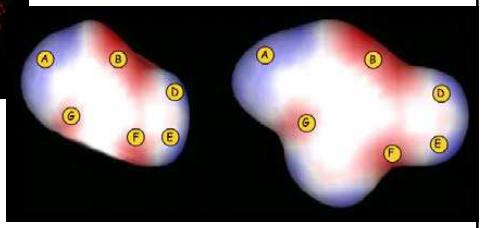




Superimposition of structures

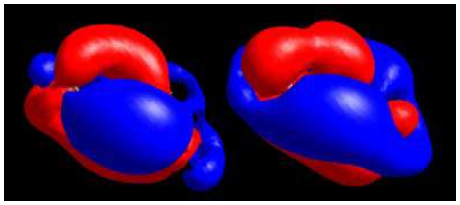
- Superimposition of compounds shapes
- Superimposition according to molecular fields

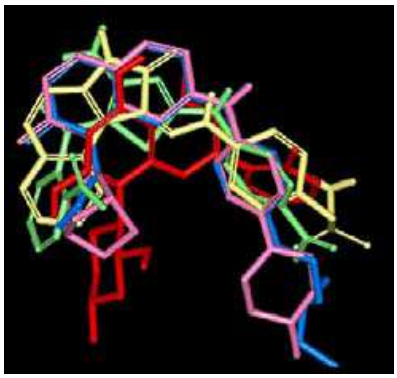




Superimposition of structures

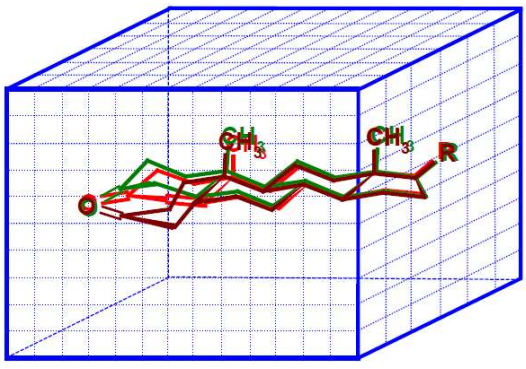
- Superimposition according to electrostatic potential
- Superimposition according to the binding method (ideal case).





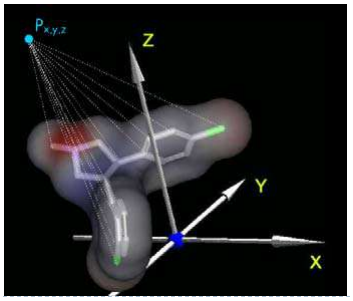
grid

- Place the superimposed compounds in a grid of points of appropriate size
- Calculations are carried out only for grid points



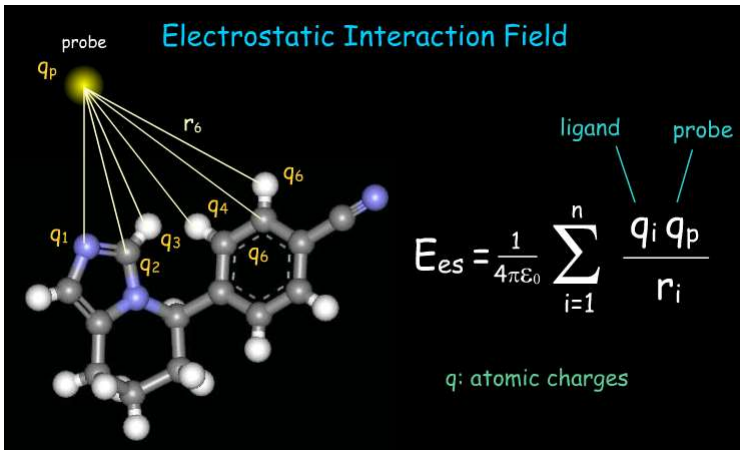
Interactions with molecular probes

- Electrostatic probe (proton, H^+) - electrostatic interactions
- Hydrophobic probe (CH_3) - Van der Waals interaction
- Multiatomic probes (OH, NH_2, COOH, etc.)

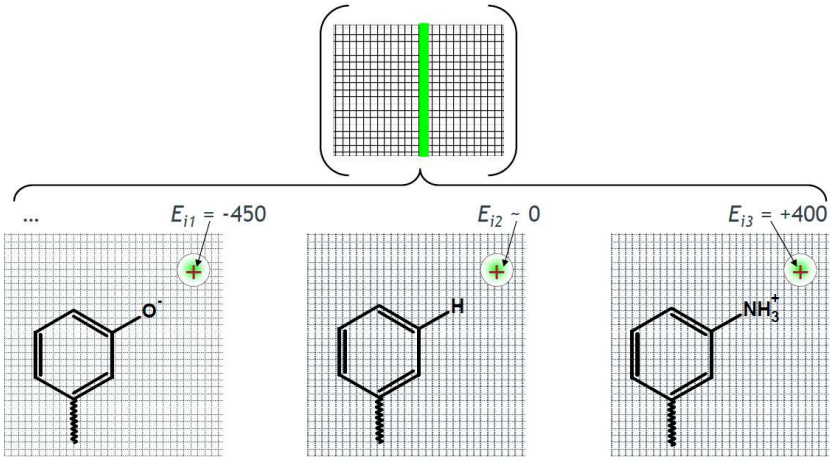


Electrostatic interactions

- Coulomb's law

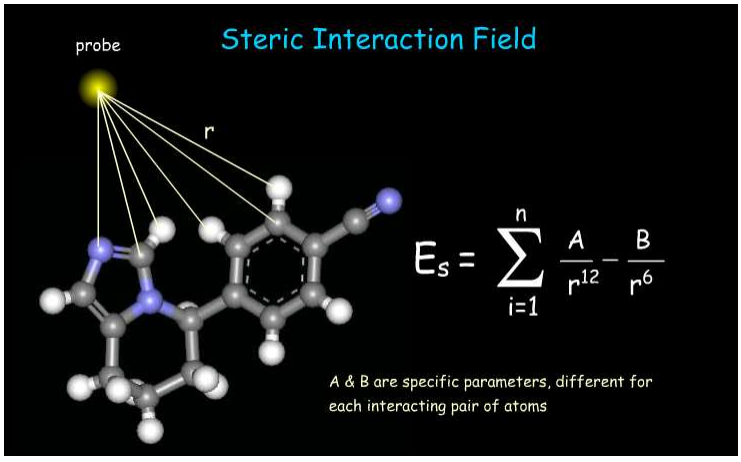


Electrostatic interactions



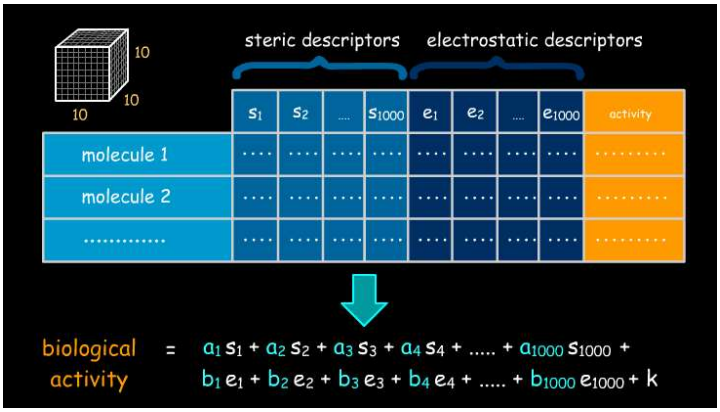
Hydrophobic interactions

- Lenard-Jones's potential



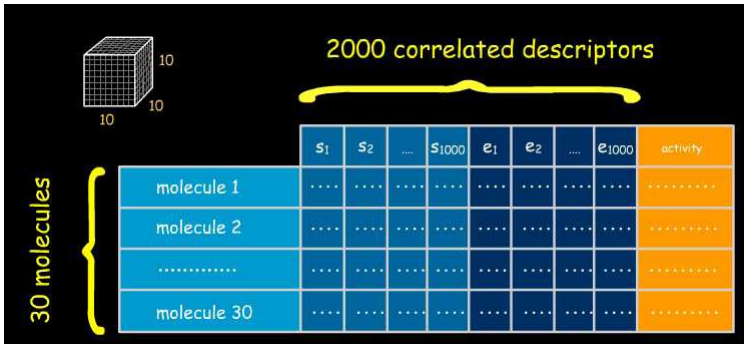
CoMFA

- CoMFA analysis correlates the obtained molecular field (values for grid points) with the activity of compounds



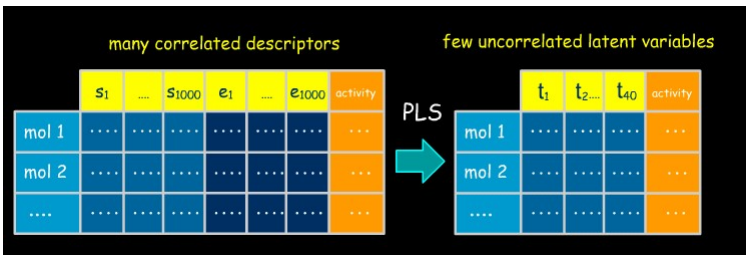
Correlation of variables

- A significant excess of descriptors relative to the number of compounds
- The number of compounds should be 3-5 greater than the number of descriptors



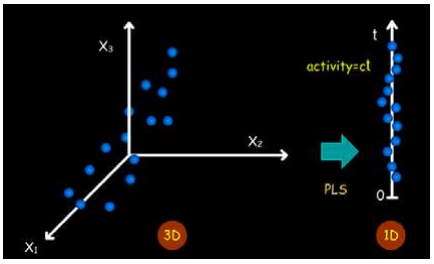
PLS

- PLS – **p**artial **l**east **s**quares
- Determine a small number of hidden independent variables t (latent variables, LV)



PLS

- The factors **t** are determined on the basis of a linear combination of descriptors;
- The factors **t** have no structural significance;
- QSAR equation:
  - **Activity = $c_1t_1 + c_2t_2 + c_3t_3 + \ldots$**



Validation

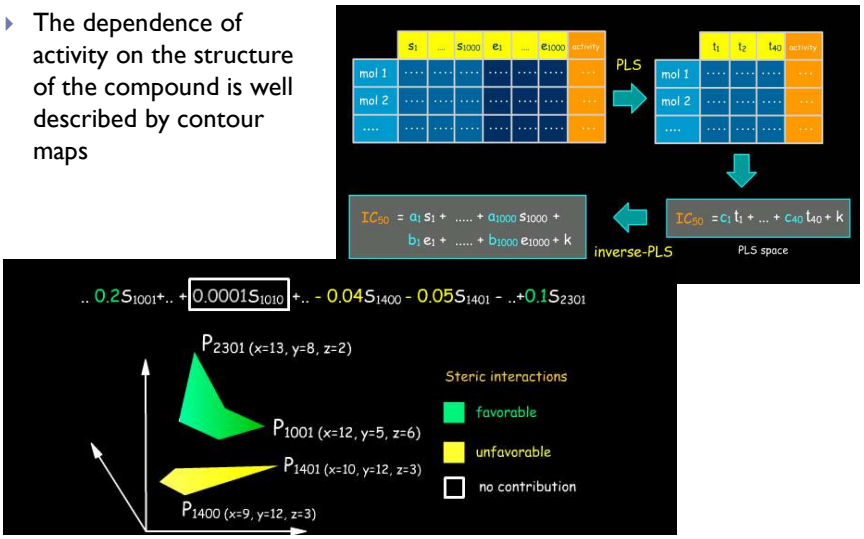
- The obtained QSAR model should be validated (evaluated), using a set of test compounds (different from the compounds used to build the model)

$$q^2 = 1 - \frac{PRESS}{SD} = 1 - \frac{\sum_{i=1}^{n}(Y_a - Y_p)^2}{\sum_{i=1}^{n}(Y_a - Y_m)^2}$$

- Y_a – actual value, Y_p - predicted value, Y_m - average of observed activities
- PRESS = sum of squares of differences,
- SD = sum of squares of deviations from the mean

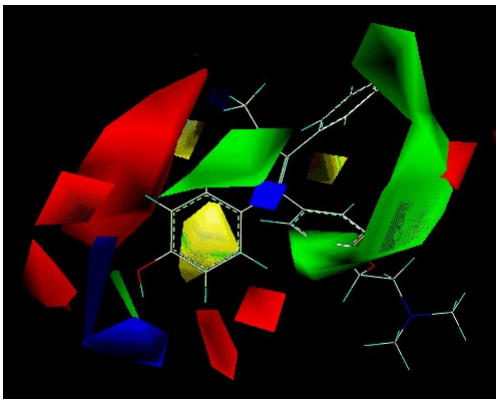
Contour maps

- The dependence of activity on the structure of the compound is well described by contour maps



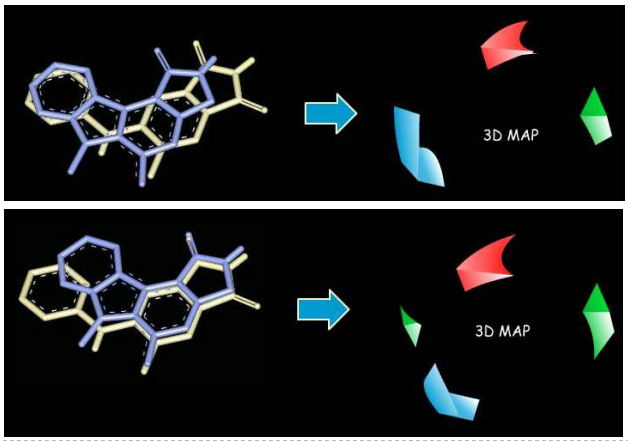
Contour maps

- Contour maps show regions with favorable and unfavorable effects on biological activity
- Steric:
  - Favorable (green)
  - Unfavorable (yellow)
- Electrostatic:
  - Positive charge beneficial (blue)
  - Negative charge favorable (red)



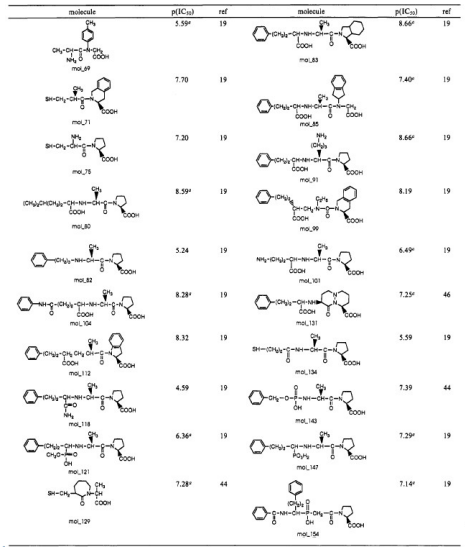
Contour maps

- The image obtained on the CoMFA contour maps depends on the fit of the molecules.



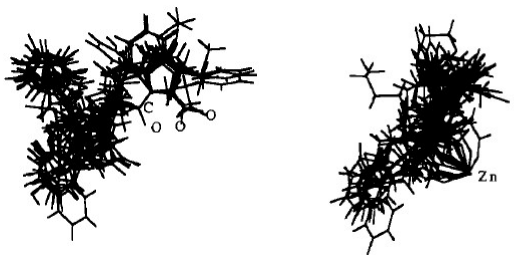
CoMFA - example

- ACE inhibitors
- 68 various compounds with known ACE activity in the training set



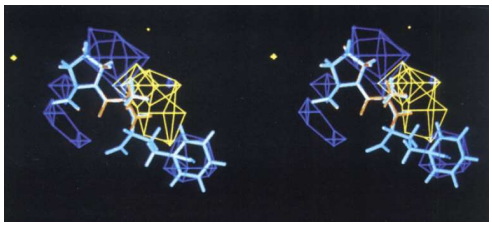
CoMFA - example

- ACE inhibitors
- The compounds were superimposed according to the proposed mechanism of action

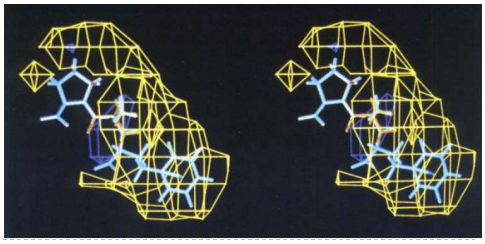


CoMFA - example

- ACE inhibitors
- Electrostatic maps:
  - Yellow (positive charge)
  - Blue (negative charge)



- Steric maps:
  - Yellow (preferred)
  - Blue (unfavorable)



COMSIA

- **CoMSIA - Co**mparative **M**olecular **S**imilarity **I**ndex **A**nalysis
  - Molecular probes:
    - electrostatic
    - hydrophobic
    - Hydrogen bond donor
    - Hydrogen bond acceptor
  - Potentials are not calculated on the basis of energy but similarity indexes
  - The method is less sensitive to wrong superimposition of molecules

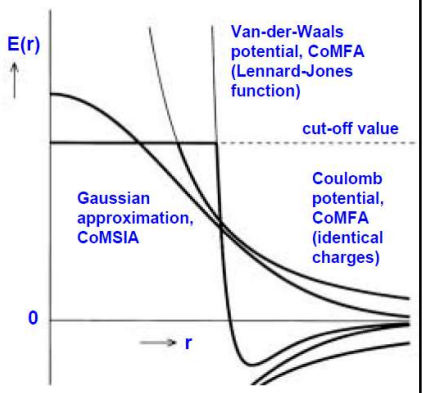
COMSIA

- Coulomb's potential

$$E_C = \sum_{i=1}^{n} \frac{q_i q_j}{D r_{ij}}$$

- Lennard-Jones's potential

$$E_{vdW} = \sum_{i=1}^{n} \left(A_j r_{ij}^{-12} - C_j r_{ij}^{-6} \right)$$



- Similarity indexes

$$A_F = -\sum_{i=1}^{m}\sum_{j=1}^{n} w_{ij} e^{-\alpha r_{ij}^2} \quad ; \quad w_{ij} = w_E q_i q_j + w_S V_i V_j + \ldots$$

3D QSAR

- Advantages
  - The three-dimensional structure of molecules is taken into account.
  - It is possible to create a model for diverse relationships.
  - Steric, electrostatic and hydrogen bonds are taken into account.
  - We get three-dimensional maps of favorable and unfavorable regions.
- Disadvantages
  - Uncertainty about bioactive conformation
  - Uncertainty about different ligand binding methods
  - Big risk of accidental correlations
  - Applicable only for *in vitro* data

Other 3D QSAR methods

- **HQSAR (Hologram QSAR):** compound structures are encoded as binary strings. The method does not require the superimposition of molecules and gives the chance to analyze large data sets. The presence of functional groups or molecular fragments in a given place forms a descriptor.
- **GRIND (GRid INdependent Descriptors).** The method uses a combination of several simplified molecular fields. Descriptors are based on three-dimensional structures but independent of the orientation of compounds in space.
- **QuaSAR (Quasi-atomistic SAR).** Pseudo-receptors are defined, i.e. pseudoatomic coating around superimposed compounds (no mesh). Pseudoatoms are possible, e.g. HB type (hydrogen bond).

Other 3D QSAR methods

- **4D, 5D,** 6D-QSAR (multidimensional QSAR). The methods may include:
  - many conformations and protonation states of the same ligand (4D);
  - hypothetical ligand and receptor (5D) fit states;
  - hypothetical ligand states including solvent effects (6D).
- **RD-QSAR** (Receptor-Dependent QSAR). The method takes into account the structure of the receptor. For example, it is possible to predict toxicity of compounds based on the structures of estrogen, androgen and cytochrome P450 receptors. It uses the 5D-QSAR method: the inductive effects of ligand-receptor matching are more realistic.

Summary

- Pharmacophors show important common elements of active compounds.
- 3DQSAR models show variable elements important for activity.