Chapter 2 – Molecular Modeling

Small Molecules

Part 2:

Analysis & Interactions

Molecular Electrostatic Potentials

- Critical for interaction / reactions
  - Initial interaction (longest distance)
  - Noncovalent

1. Electrostatic: charge / dipole
2. Induced dipole
3. Dispersion (van der Waals)

- Represented by energy grid
  - Electron / proton energy

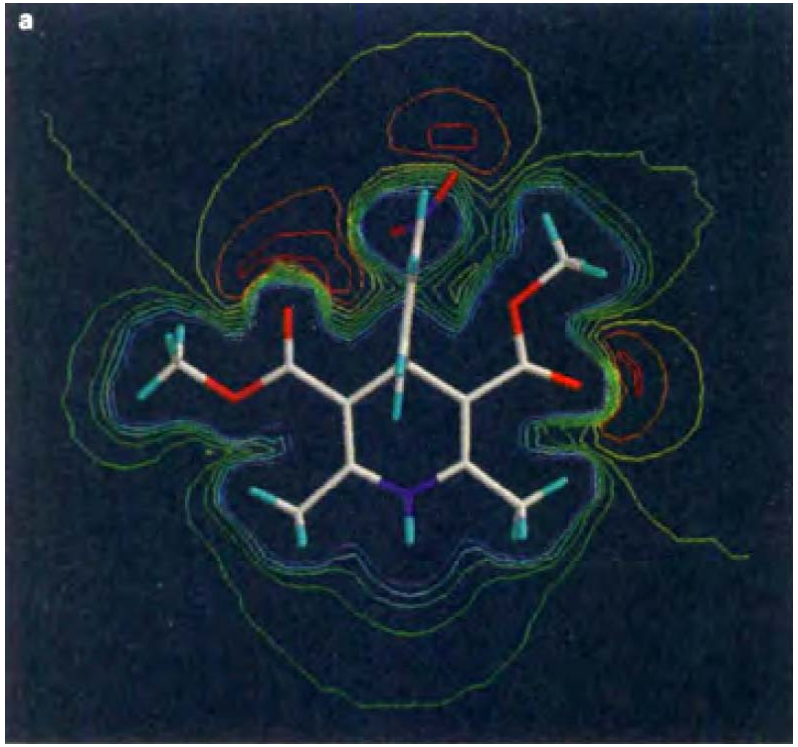
$$E = \frac{1}{4\pi\epsilon}\frac{q}{r^2}$$

Calculating MEP

- Atomic point charges
  - X-ray, QM → Electron density → Partial charge
    1. Topological calculations
       - Electro negativity
       - Bonds (connectivity, not structure)
         - Gasteiger-Hückel method (σ + delocalized π)
       - New groups of molecules must be tested by QM
    2. Quantum mechanics
       - Semiimperical or ab initio → ψ
       1. Mulliken population analysis: Atomic orbital occupancy (oldest)
       2. ESP fit method: Atomic charge fit to electron density
- Test: Dipole moment of rigid molecules

Visualising MEP

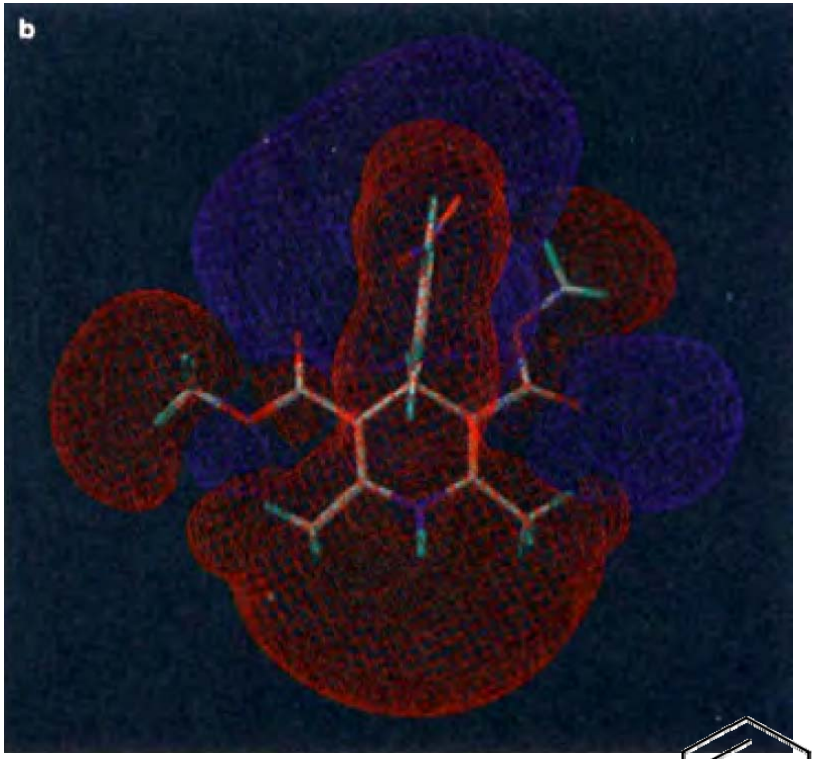
- Protons electrostatic energy in MEP
- QM: Proton and molecules wavefunction



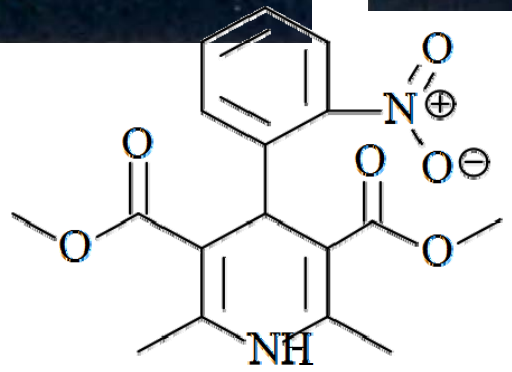
- Isocontour: (2D)

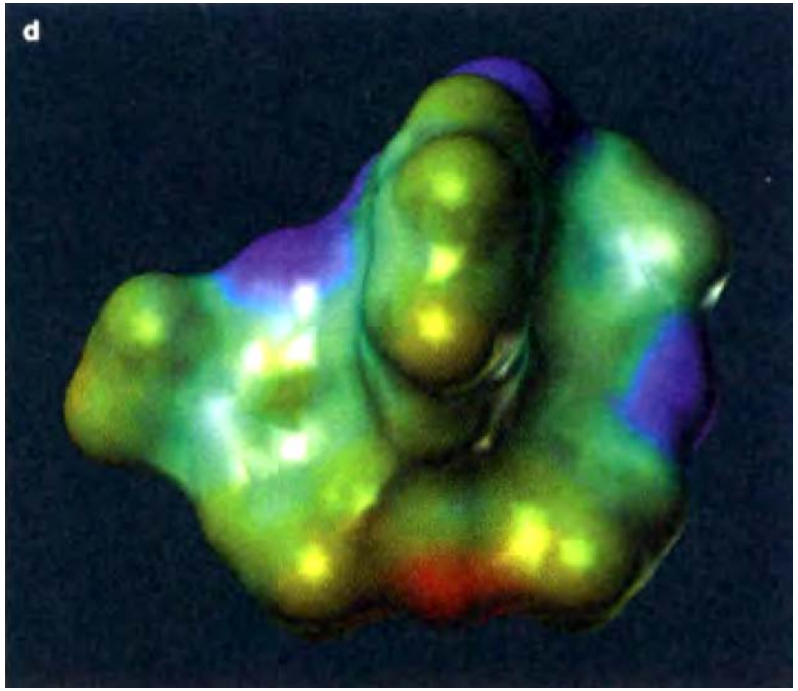
  Nifedipine

Visualising MEP



Isopotential surface

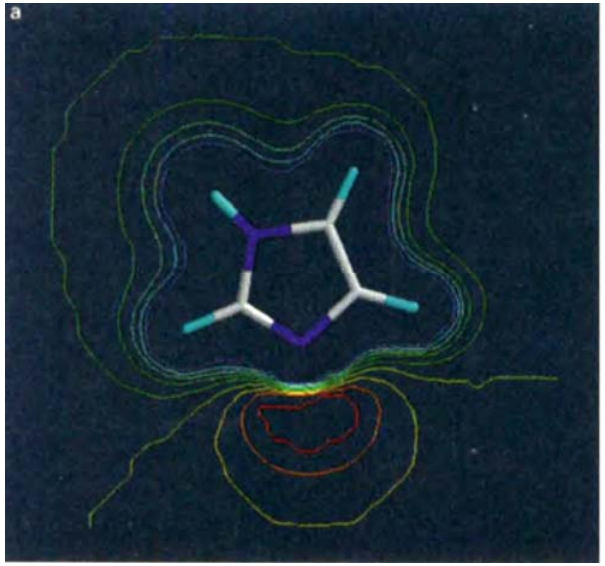




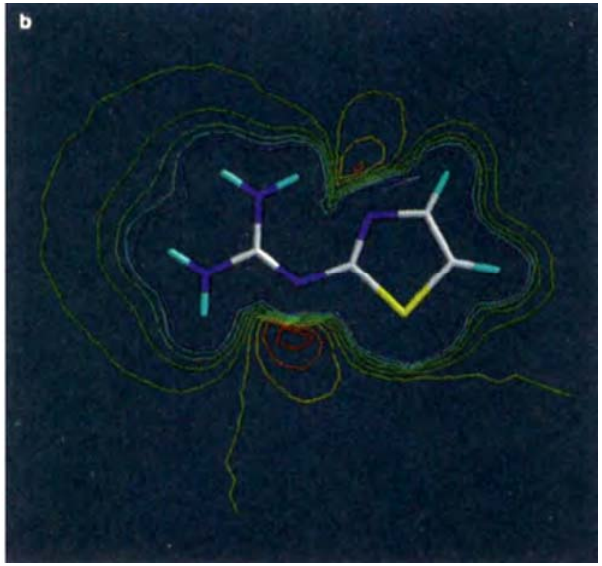
Connolly surface colour plot

MEP Superimposition

- Ligands share MEP traits
- MEP superimposition > atom – atom fit



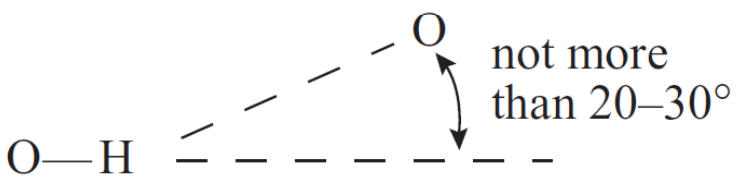
Imidazole

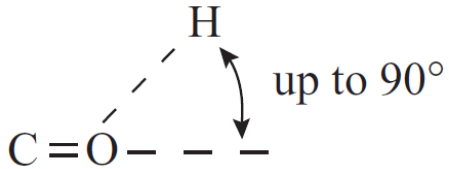


Guanidinothiazole

Molecular Interaction Fields

- Noncovalent interaction (docking)
  - Interaction energy > vdw repulsion → binding
- Target – probe interaction energy (grid) - GRID
  - Water, hydroxyl, ions etc.
  - $E_{tot} = E_{vdw} + E_{et} + E_{hb}$
    - van der Waal: Dispersion + electron overlap
    - Electrostatic: Coulomb (ε-dependent)
    - HB: Electrostatic but orientation-sensitive!





Protein Physics - A Course of Lectures: Alexei V. Finkelstein & Oleg B. Ptitsyn

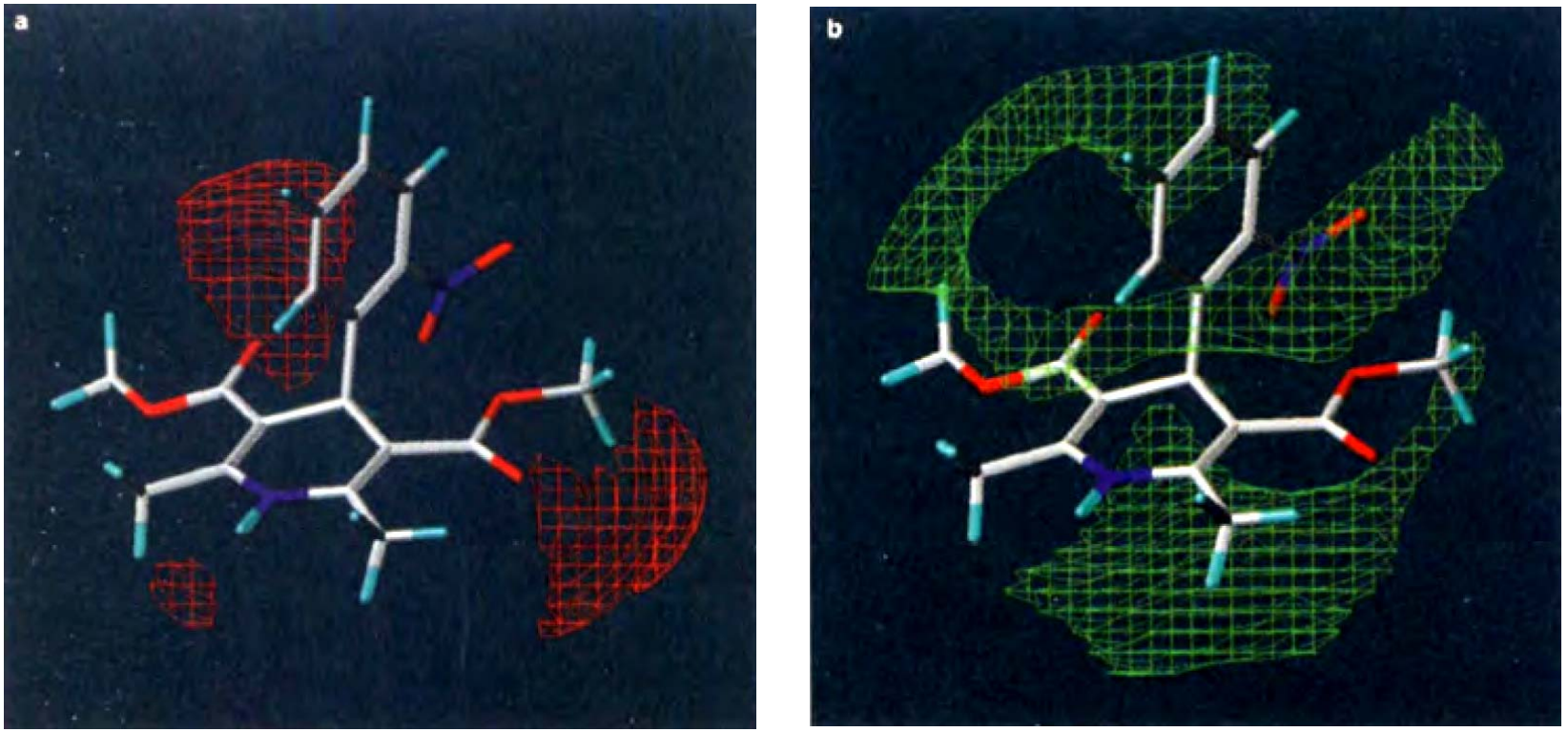
MIF Investigation: GRID

- Probe parameters

	Methyl probe	Hydroxyl probe	Carboxyl probe
Van der Waals radius (Å)	1.950	1.650	1.600
Effective number of electrons	8	7	6
Polarizabilty ($Å^3$)	2.170	1.200	2.140
Electrostatic charge	0.000	–0.100	–0.450
Optimal hydrogen bond energy (kcal mol^{-1})	0.000	–3.500	–3.500
Hydrogen bonding radius (Å)	–	1.400	1.400
Number of hydrogen bonds donated	0	1	0
Number of hydrogen bonds accepted	0	2	2
Hydrogen bonding type	0	4	8

- Metal ion coordination
- Ligands (same receptor): Common IF

MIF Investigation: GRID

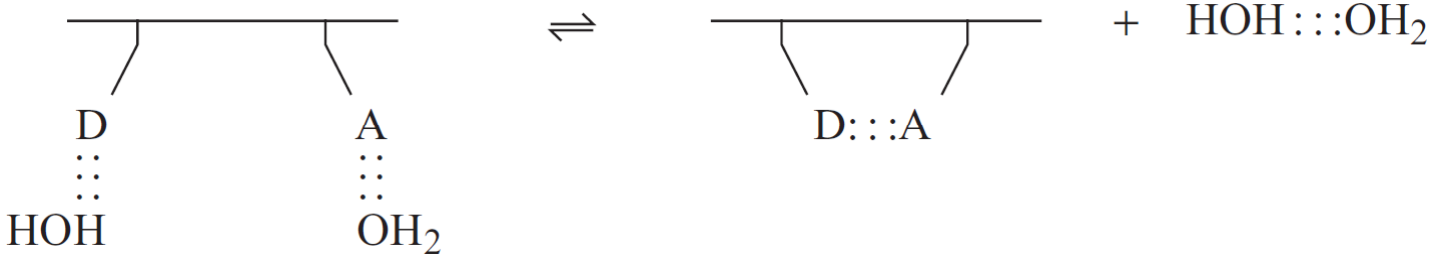


Hydroxyl: -3.5 kcal mol^{-1} Methyl: -1.4 kcal mol^{-1}

Isopotential on nifedipine

Hydrophobic Interactions

- No simple calculation
- Arises from increased entropy of solvent
- $G=H-TS$
- Emperical: $\log P_{oct/wat} = \log\left(\frac{[solute]_{octanol}}{[solute]_{water}^{un-ionized}}\right)$
- Smaller molecules $\rightarrow \log\left(P_{compound}\right)$
- log(P) (1D) < Hydrophobic field (3D)

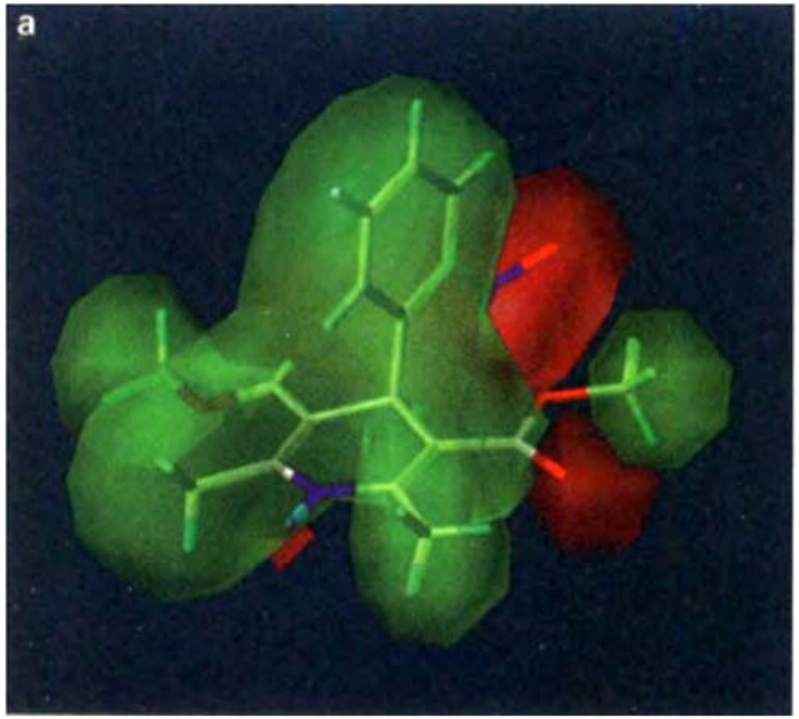


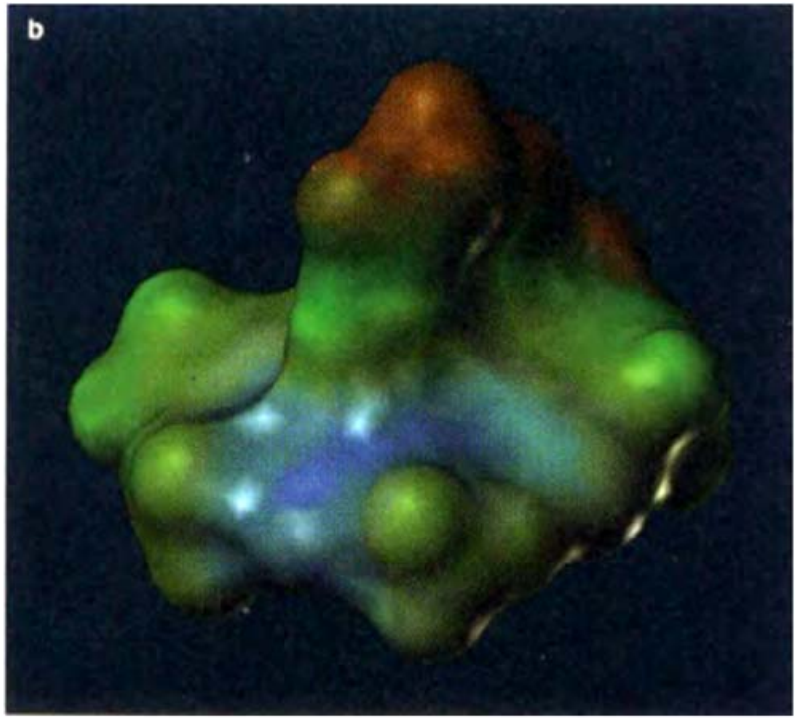
Protein Physics - A Course of Lectures: Alexei V. Finkelstein & Oleg B. Ptitsyn

Hydrophobic Field: HINT & MOLCAD

- Empiric hydrophobic fragment constants
  - Partition experiments
  - Hydrophobicity, hydrophilicity & vdw
- HINT: Hydropathic field
  - Constants, Connolly & distance function
- Empiric data: contour level estimation

- MOLCAD: Lipophilic potential surface
  - Prediction/optimization of: ligand a/o receptor
  - Conformational change (partition)
    - Test molecules must be rigid





Hydrophobic field map

Lipophilic potential (Connolly)

Nifedipine

Pharmacophore Identification

- Pharmacophore →(bind)→ Enzyme/receptor
  - Sterically consistent elements
    - Pharmacophoric elements
      - HB donors & receptors, ringsystems, flexibility…
      - Selection?
- Superimposition
  - Selection
    - Activity
    - Antagonist – agonist
    - Conformation energy

Superimposition

- Atom-by-atom
  - Root-mean-squared minimisation
  - Result sensitive to pharmacophoric element of interest
- Active analog approach
  - Congeneric conformational changes allowed
  - Speed increase with rigidity
  - Only small molecules (subunits)
  - Defines distance span from extremes of rigidity
    - Reduces possible congeneric conformations

Superimposition

- Rapid pairwise superposition (SEAL)
  - Compairs relative distance dissimilar molecules
  - Yields information of global shape
- Other methods
  - Superposition of HB donors/acceptors
  - FlexS
    - Flexible superpositioning
  - Superimposition of molecular fields
    - Charge, hydrophobicity, vdw
    - Grid point weights according to structure-activity relationship
    - Search template: most rigid congener

3D QSAR

(3D Quantitative structure-activity relationship)

- Characterised compounds
  - Structure + Biological activity
- Correlation with field properties
  - vdW, electrostatic, lipophilicity…
- Next-generation compounds
- Biological data
  - *in vitro*
  - Common binding mode
  - Diffusion
  - Inactive compounds
  - Large activity span (3 orders)

Comparative Molecular Field Analysis

- Relies on field properties (grid)
- Steric + electrostatic interactions
- No entropic or hydrophobic effects
  - Knowledge of binding mode is needed
- Statistics
  - Partial Least Square
  - More energies than compounds
    - Some less important
    - Linear combinations
    - Leave-one-out cross validation

→

Comparative Molecular Field Analysis

- Statistics (LOO)
  - Predicting activity one molecule from set (without it)
    - Square of crossvalidated correlation coefficient:
      - Should be more than 50%

$$Q^2 = 1 - \frac{\sum \left(\gamma_{obs} - \gamma_{pred}\right)^2}{\sum \left(\gamma_{obs} - \gamma_{mean}\right)^2}$$

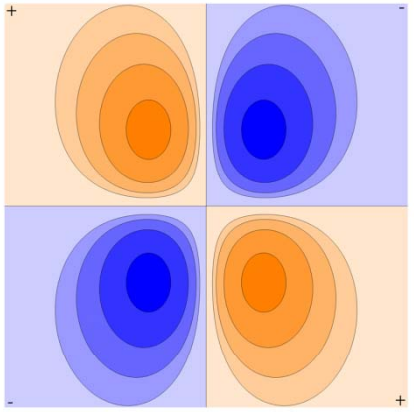
    - Standard deviation of error prediction:
      - Should be ~steady with # variables
        Few variables + Low noise = good

$$SDEP = \sqrt{\frac{\sum \left(\gamma_{obs} - \gamma_{pred}\right)^2}{N}}$$

  - Scrambling test
    - Set is mixed (activity/molecule)
      - Bad predictions → model may be ok

CoMFA related methods

- Comparative Molecular Similarity Indices Analysis
  - No field potentials → Gaussian functions
    - Easier interpretation + No cut-off values needed
- Graphical Retrieval and Information Display – General Optimal Linear PLS Estimation
  - Classifies variables (energies)
    - Only helpful variables are considered in final run
- Alignment-independent methods
  - Inertia, dipole, quadropole moments



3D QSAR

- Short coming: Induced fit
  - Flexible amino acids + Flexible HB donors/acceptors
- Pseudoreceptor models
  - Pharmacophore → Optimal binding partners
  - Pseudoreceptor ≠ receptor (structure)
- Receptor-based 3D QSAR
  - Docking + CoMFA
  - Good at identifying binding pocket

3D QSAR interpretation & reliability

- Visualisation
  - Recognition of activity specific regions
- Model
  - Must be verified (Q^2 & SDEP)
    - LOO → L20%O → L50%O
    - Molecules not in training set
- Data quality
  - Noise
  - Reliability