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

\[ E = \frac{1}{4\pi\varepsilon} \frac{q}{r^2} \]
Calculating MEP

• Atomic point charges
  • X-ray, QM $\rightarrow$ Electron density $\rightarrow$ Partial charge
    1. Topological calculations
      • Electro negativity
      • Bonds (connectivity, not structure)
        o Gasteiger-Hückel method ($\sigma$ + delocalized $\pi$)
      • New groups of molecules must be tested by QM
    2. Quantum mechanics
      • Semiimperical or ab initio $\rightarrow \psi$
        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
Molecular Interaction Fields

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

---

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

- Probe parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Methyl probe</th>
<th>Hydroxyl probe</th>
<th>Carboxyl probe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Van der Waals radius (Å)</td>
<td>1.950</td>
<td>1.650</td>
<td>1.600</td>
</tr>
<tr>
<td>Effective number of electrons</td>
<td>8</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>Polarizability (Å³)</td>
<td>2.170</td>
<td>1.200</td>
<td>2.140</td>
</tr>
<tr>
<td>Electrostatic charge</td>
<td>0.000</td>
<td>-0.100</td>
<td>-0.450</td>
</tr>
<tr>
<td>Optimal hydrogen bond energy (kcal mol⁻¹)</td>
<td>0.000</td>
<td>-3.500</td>
<td>-3.500</td>
</tr>
<tr>
<td>Hydrogen bonding radius (Å)</td>
<td>–</td>
<td>1.400</td>
<td>1.400</td>
</tr>
<tr>
<td>Number of hydrogen bonds donated</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Number of hydrogen bonds accepted</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Hydrogen bonding type</td>
<td>0</td>
<td>4</td>
<td>8</td>
</tr>
</tbody>
</table>

- 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: \( \log P_{\text{oct/wat}} = \log \left( \frac{[\text{solute}]_{\text{octanol}}}{[\text{solute}]_{\text{un-ionized}}^{\text{water}}} \right) \)
- Smaller molecules \( \rightarrow \) \( \log(P_{\text{compound}}) \)
- \( \log(P) \) (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 $\rightarrow$ (bind) $\rightarrow$ 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%
    - Standard deviation of error prediction:
      - Should be ~steady with # variables
        Few variables + Low noise = good

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

\[
Q^2 = 1 - \frac{\sum (\gamma_{obs} - \gamma_{pred})^2}{\sum (\gamma_{obs} - \gamma_{mean})^2}
\]

\[
SDEP = \sqrt{\frac{\sum (\gamma_{obs} - \gamma_{pred})^2}{N}}
\]
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 $\rightarrow$ Optimal binding partners
  - Pseudoreceptor $\neq$ 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 $\rightarrow$ L20%O $\rightarrow$ L50%O
    - Molecules not in training set

- Data quality
  - Noise
  - Reliability