A Winter Short Course on Statistical Mechanics for Molecular Simulations

Lecture 2: Molecular Mechanics & Classical Force Fields

Yuan-Chung Cheng
yuanchung@ntu.edu.tw

2/2/2015

Many materials taken from Prof. A. D. MacKerell's presentation, which is gratefully acknowledged.

The Molecular Hamiltonian

$$\hat{H} = -\sum_{A} \frac{1}{2M_A}\nabla_A^2 - \sum_{i} \frac{1}{2}\nabla_i^2 + \sum_{A>B} \frac{Z_A Z_B}{R_{AB}} - \sum_{Ai} \frac{Z_A}{r_{Ai}} + \sum_{i>j} \frac{1}{r_{ij}}$$

In atomic units: 1 Hartree = 27.2114 eV = 627.509 kcal/mol

Electronic Hamiltonian

$$\hat{H}_{electronic} = \sum_{i} \hat{T}(i) - \sum_{iA} \frac{Z_A}{r_{iA}} + \sum_{i,j} \frac{1}{r_{ij}} + \sum_{A,B} \frac{Z_A Z_B}{R_{AB}}$$

Kinetic Energy of electrons

Electron-nucleus attraction

Electron-electron repulsion

Nuclear-nuclear repulsion

Solving this is a huge part of molecular simulation in chemistry (i.e. quantum chemistry), but this is not the topic that I will pursue.

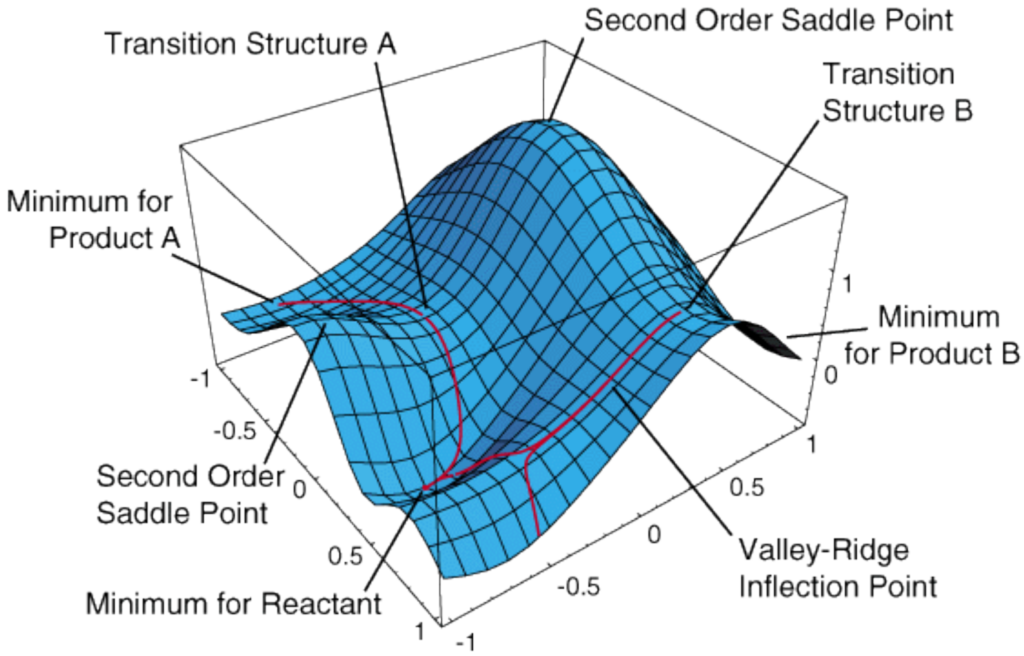
Potential Energy Surface

- QM calculations yield potential energy surface that governs nuclear motions

$$E(\mathbf{R}) = \langle \Psi | \mathbf{H}_{elec} | \Psi \rangle$$

Determines reactions/thermodynamics/…

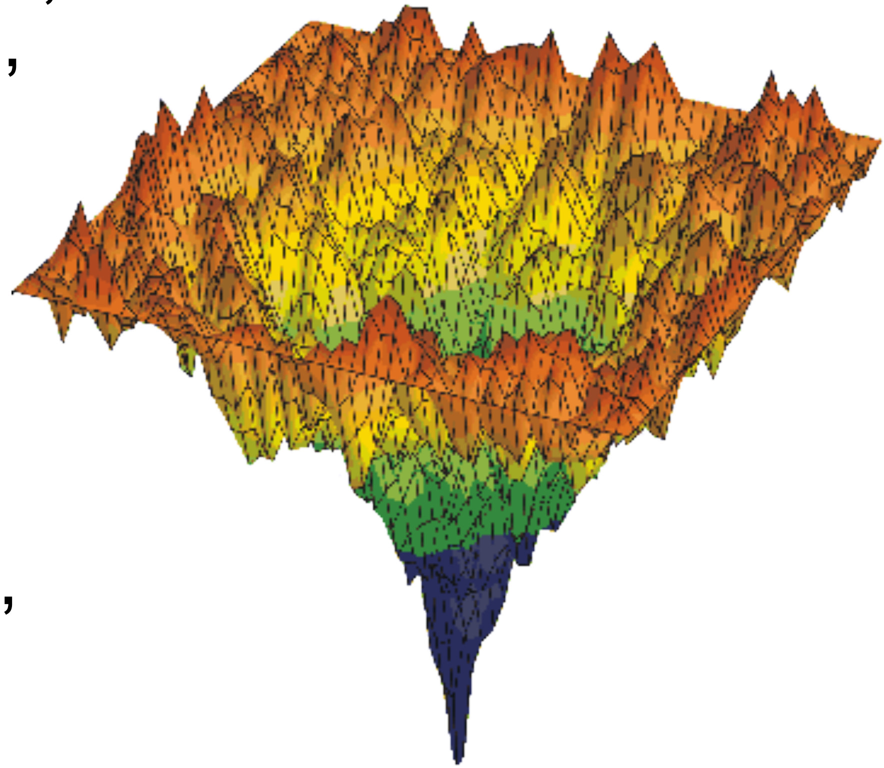
Multidimensional
Complex



Potential Energy Surface

- Large systems –
  complex energy landscape,
  many degrees of freedom,

  → full quantum
  all-electron calculations
  become infeasible!!
- We need simple classical models
- No need for all the details, anyway…

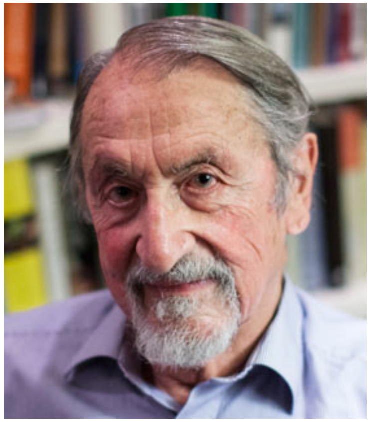


Protein folding funnel

The curse of dimensionality!!

2013 年諾貝爾化學獎

- 2013 Nobel Chemistry Prize jointly to Martin Karplus, Michael Levitt and Arieh Warshel *"for the development of multiscale models for complex chemical systems"*.
- 複雜系統多層級計算方法的建立
- 將化學實驗帶入電腦時空



Martin Karplus

Michael Levitt

Arieh Warshel

Pictures from nobelprize.org

Slide from Prof. Jhih-Wei Chu

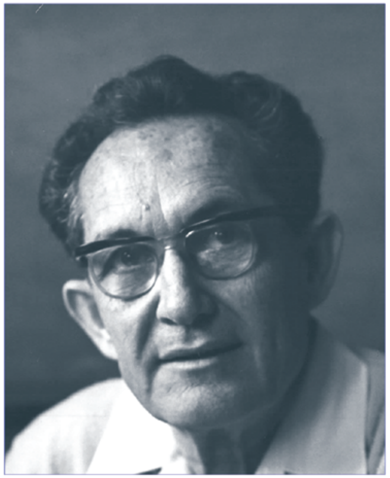
Very informative read!

Michael Levitt's Nobel Lecture

CONSISTENT FORCE-FIELD 1968

Weizmann Institute 1967-68

Michael Levitt



Shneior Lifson

Arieh Warshel

THE JOURNAL OF CHEMICAL PHYSICS VOLUME 49, NUMBER 11 1 DECEMBER 1968

Consistent Force Field for Calculations of Conformations, Vibrational Spectra, and Enthalpies of Cycloalkane and *n*-Alkane Molecules

S. Lifson and A. Warshel
Department of Chemical Physics, Weizmann Institute of Science, Rehovot, Israel

(Received 13 May 1968)

General Considerations

- Description of molecules?
- Optimization of force field parameters?
- Training set of compounds/data?
- Test set of compounds/data?
- Limitations – questions you should not ask of your force field

Overview and parameter optimization of CHARMM Force Field

Based on protocol established by

Alexander D. MacKerell, Jr , U. Maryland

See references: www.pharmacy.umaryland.edu/faculty/amackere/force_fields.htm

Especially Sanibel Conference 2003, JCC v21, 86,105 (2000)

Common All-Atom Force Fields

- Class I: Standard structural terms CHARMM, CHARMm (Accelyrs), AMBER, OPLS, ECEPP, GROMOS, SYBYL (Tripos)
- Class II: Standard + cross terms CFF95 (Accelrys), MM3, MMFF94, UFF
- Class III: Non-additive, polarizable terms QM/MM, Polarizable FF - Freisner/ Berne(Schroedinger), AMOEBA (Tinker)

They are different!!! So parameters from one cannot be used in another force field.

State of the art additive force fields are typically all-atom models

All atoms, including all hydrogens, explicitly represented in the model.

Lone pairs included on hydrogen bond acceptors in some force fields.

e.g., CHARMM22 and 27, AMBER94….03, OPLS/AA

Extended or united atom models

(omit non-polar hydrogens)

CHARMM PARAM19 (proteins)

- often used with implicit solvent models
  - ACE, EEF, GB variants
- improper term to maintain chirality
- loss of cation - pi interactions

OPLS

AMBER

GROMOS

Transition State Force Field Parameters

Same approach as standard force field parameterization

Require target data for transition state of interest: *ab initio*

Metal Force Field Parameterization

Only interaction parameters or
include intramolecular terms

Parameterization of QM atoms for QM/MM calculations

Polarizable “non-additive” force fields

Include explicit term(s) in the potential energy function to treat induction/polarization of the charge distribution by the environment. Still under development.

CHARMM
  Drude (MacKerell, Roux and coworkers)
  PIPF (Gao and coworkers)
  Cheq (Brooks and coworkers)
AMBER
Friesner/Berne et al. (Schrödinger Inc.)
TINKER

Class I Potential Energy function

$$E_{Total} = \sum_{bonds} k_b(b - b_0)^2 + \sum_{angles} k_\theta(\theta - \theta_0)^2$$

$$+ \sum_{dihedrals} \frac{V_n}{2}[1 + \cos(n\phi - \delta)]$$

$$+ \sum_{impropers} k_\omega(\omega - \omega_0)^2 + \sum_{Urey-Bradley} k_u(r_{1,3} - r_{1,3,0})^2$$

Non-bonded Interaction Terms

$$+ \sum_{electrostatics} (\frac{q_i q_j}{\varepsilon r_{ij}}) + \sum_{VDW} \varepsilon_{ij}[(\frac{R_{\min,ij}}{r_{ij}})^{12} - 2(\frac{R_{\min,ij}}{r_{ij}})^6]$$

From MacKerell

Intramolecular energy function and corresponding force field parameters

$$\sum_{bonds} K_b (b - b_o)^2 + \sum_{angles} K_\theta (\theta - \theta_o)^2 + \sum_{torsions} K_\phi (1 + \cos(n\phi - \delta))$$

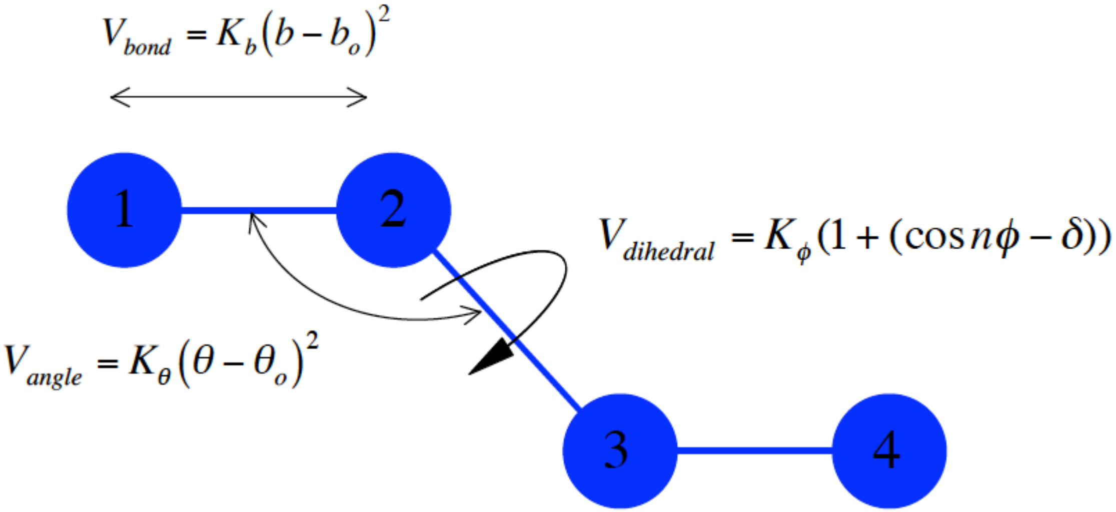
$$+ \sum_{impropers} K_\varphi (\varphi - \varphi_o)^2 + \sum_{Urey-Bradley} K_{UB} (r_{1,3} - r_{1,3,o})^2 + \sum_{\phi,\psi} V_{CMAP}$$

Equilibrium terms
- b_o: bonds
- θ_o: angles
- n: dihedral multiplicity
- δ_o: dihedral phase
- ω_o: impropers
- $r_{1,3o}$: Urey-Bradley

Force constants
- K_b: bonds
- K_θ: angles
- K_ϕ: dihedral
- K_ω: impropers
- K_{UB}: Urey-Bradley

Aka. Internal or bonded terms

Diagram of intramolecular energy terms



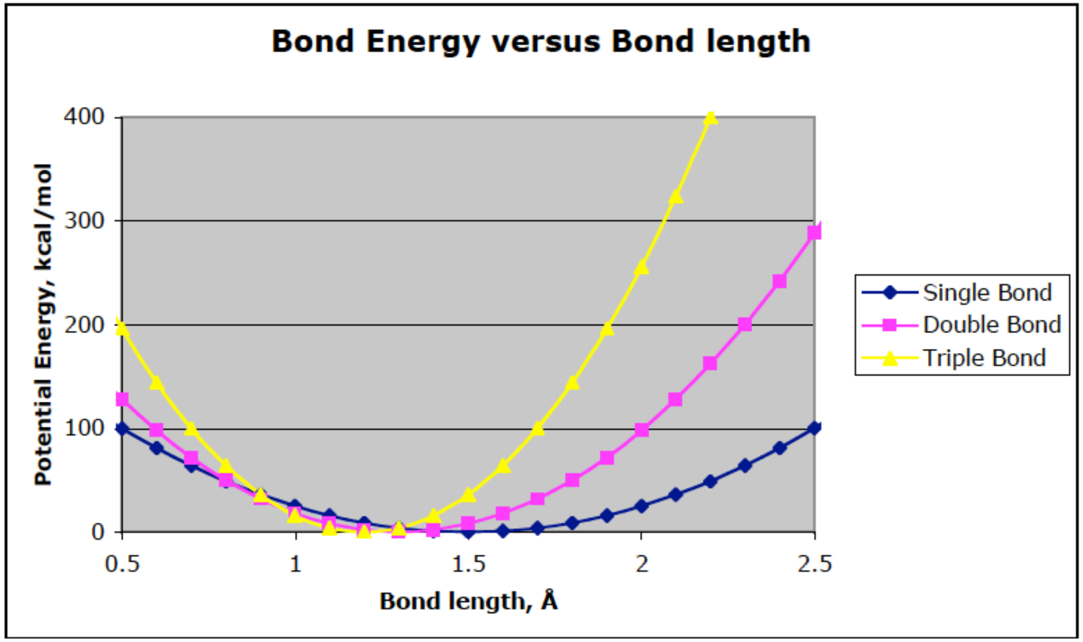
$$V_{bond} = K_b\left(b - b_o\right)^2$$

$$V_{dihedral} = K_\phi(1 + (\cos n\phi - \delta))$$

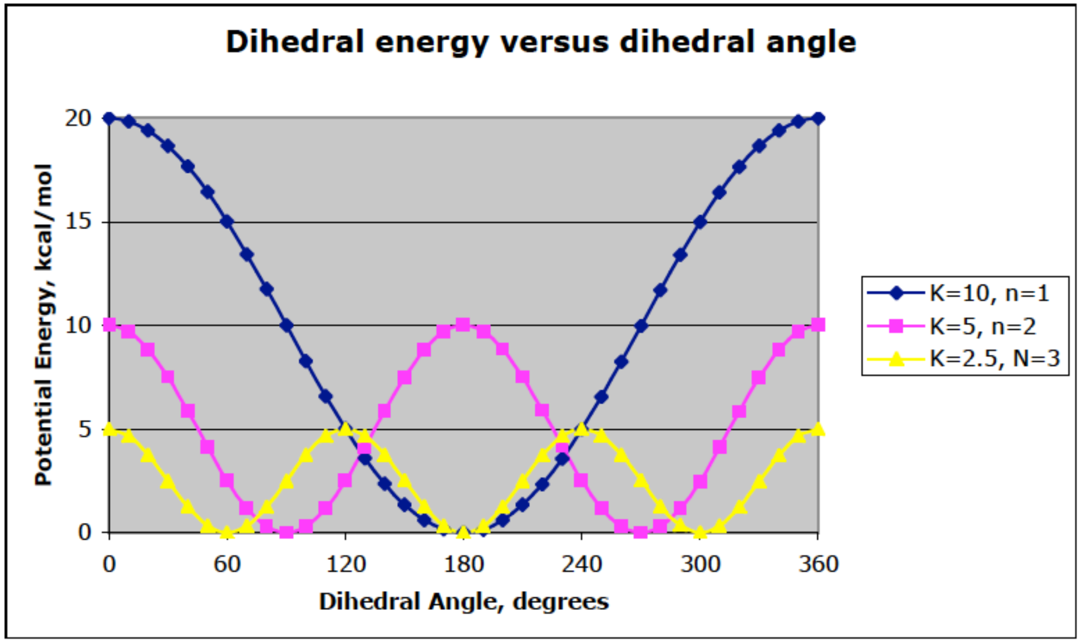
$$V_{angle} = K_\theta\left(\theta - \theta_o\right)^2$$

$$V_{bond} = K_b\left(b - b_o\right)^2$$

Chemical type	K_{bond}	b_o
C-C	100 kcal/mole/Å^2	1.5 Å
C=C	200 kcal/mole/Å^2	1.3 Å
C=-C	400 kcal/mole/Å^2	1.2 Å

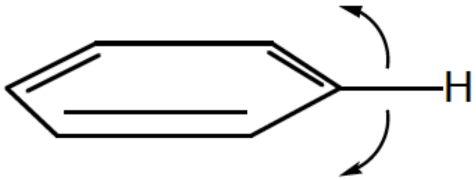


$$V_{dihedral} = K_{\phi}(1 + (\cos n\phi - \delta))$$

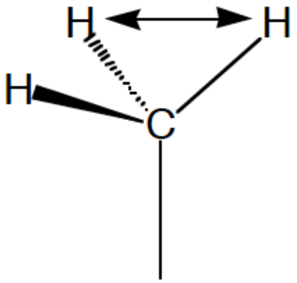


$$\delta = 0^{\circ}$$

Note use of a Fourier series for a dihedral



$$V_{improper} = K_{\varphi}\left(\varphi - \varphi_o\right)^2$$



$$V_{Urey-Bradley} = K_{UB}\left(r_{1,3} - r_{1,3o}\right)^2$$

The Urey-Bradley term captures the influence of the stretch-stretch and stretch-bend coupling terms on vibrational frequencies. It is not included in class II force fields (Amber/GROMOS, do you know why?).

See Norman Allinger, Molecular Structure: Understanding Steric and Electronic Effects from Molecular Mechanics.

2D dihedral energy correction map to the CHARMM 22 ϕ,ψ backbone (CMAP)

ϕ,ψ grid-based energy correction via bicubic interpolation

$$V_{CMAP} = f(\phi,\psi) = \sum_{i=1}^{4}\sum_{j=1}^{4} c_{ij}\left(\frac{\phi-\phi_L}{\Delta_\phi}\right)^{i-1}\left(\frac{\psi-\psi_L}{\Delta_\psi}\right)^{j-1}$$

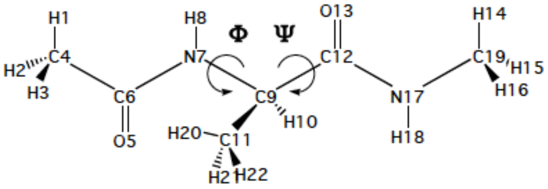
Smooth first derivatives, continuous second derivatives

Grid rectangle coefficients, c_{ij}

1) Corner grid points

2) First derivatives: $\frac{\partial f}{\partial \phi}$, $\frac{\partial f}{\partial \psi}$

3) Cross derivatives: $\frac{\partial f^2}{\partial\phi\partial\psi}$



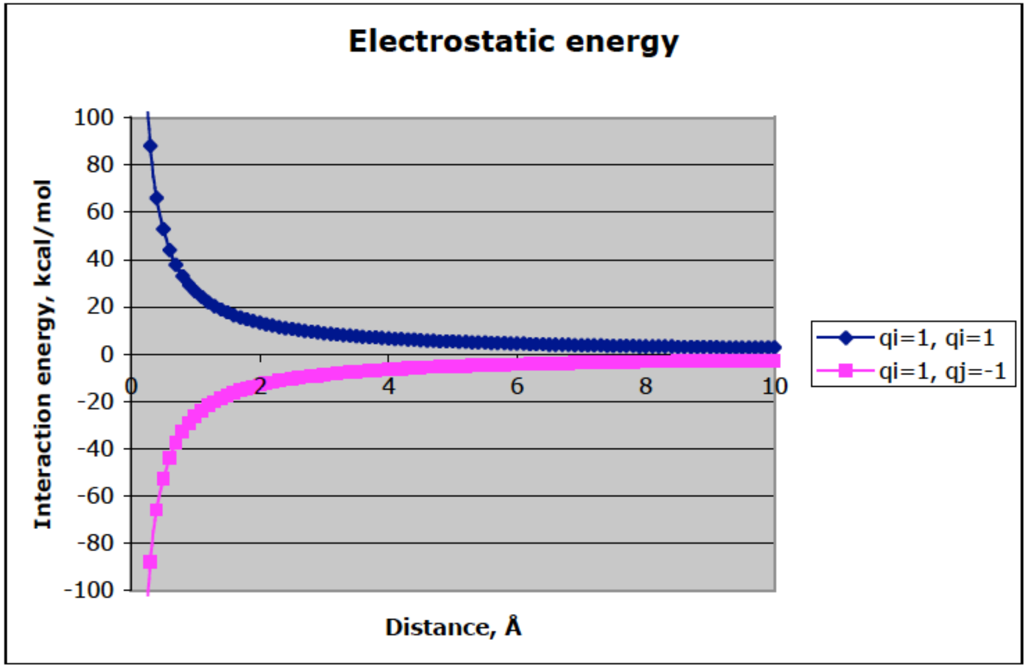
Use bicubic spline interpolation to determine derivatives

Additive intermolecular energy function and corresponding parameters

$$\sum_{nonbonded} \frac{q_i q_j}{4\pi D r_{ij}} + \varepsilon_{ij}\left[\left(\frac{R_{\min,ij}}{r_{ij}}\right)^{12} - 2\left(\frac{R_{\min,ij}}{r_{ij}}\right)^{6}\right]$$

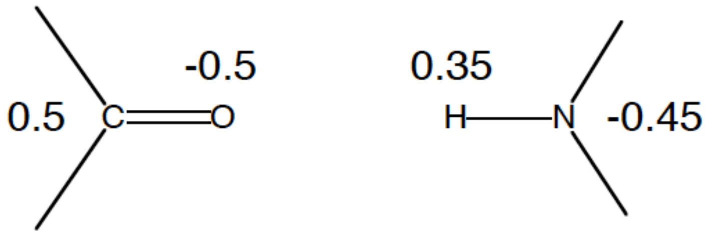
q_i: partial atomic charge
D: dielectric constant
ε: Lennard-Jones (LJ, vdW) well-depth
R_{min}: LJ radius (R_{min}/2 in CHARMM)
Combining rules (CHARMM, Amber)
$R_{min\ i,j} = R_{min\ i} + R_{min\ j}$
$\varepsilon_{i,j} = SQRT(\varepsilon_i * \varepsilon_j)$

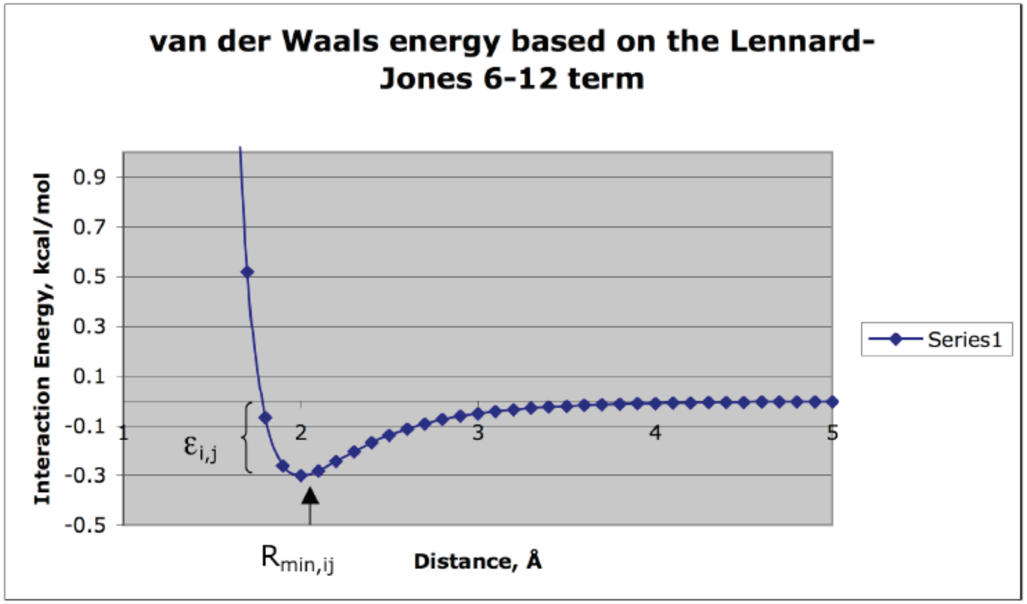
Aka. Nonbonded or external terms


Electrostatic energy
Interaction energy, kcal/mol
100
80
60
40
20
0
-20
-40
-60
-80
-100
0
2
4
6
8
10
Distance, Å
qi=1, qi=1
qi=1, qj=-1

Treatment of hydrogen bonds???

Partial atomic charges





$$\varepsilon_{ij}\left[\left(\frac{R_{\min,ij}}{r_{ij}}\right)^{12} - 2\left(\frac{R_{\min,ij}}{r_{ij}}\right)^{6}\right]$$

Alternate intermolecular terms for the electrostatic (additive) or vdW interactions

$$V_{Hbond} = \sum_{Hbonds} \varepsilon_{HB} \left[\left(\frac{R_{HB,A-H}}{r_{A-H}} \right)^{12} - \left(\frac{R_{HB,A-H}}{r_{A-H}} \right)^{10} \right] * \cos(\theta_{A-H-D})$$

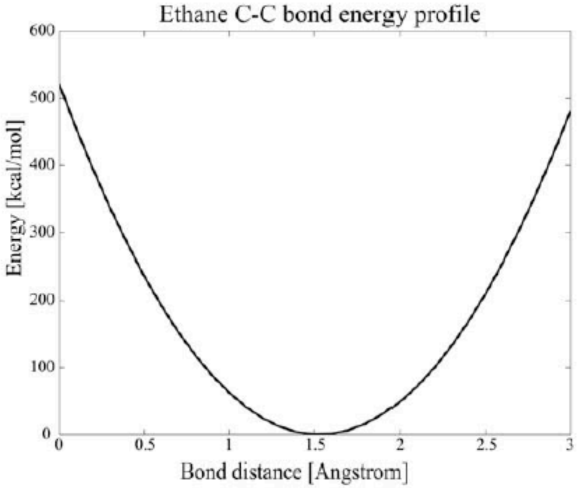
$$V_{vdw} = \sum_{vdw} \varepsilon_{ij} \left[\left(\frac{R_{\min,ij}}{r_{ij}} \right)^{9} - \left(\frac{R_{\min,ij}}{r_{ij}} \right)^{6} \right]$$

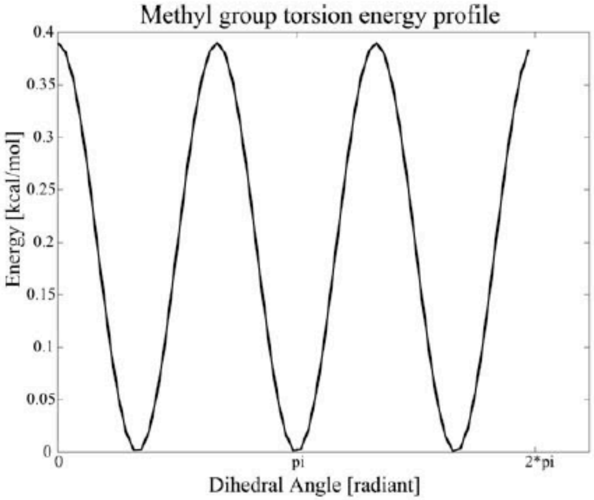
LJ-9-6

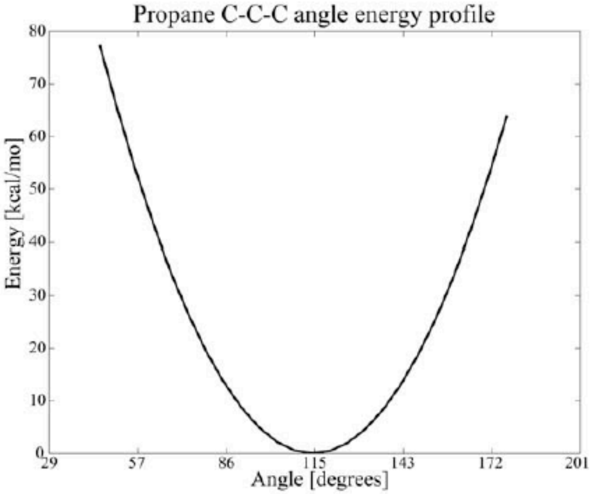
$$V_{vdw} = \sum_{vdw} \varepsilon_{ij} \left(e^{\frac{-aR_{\min,ij}}{r_{ij}}} - \left(\frac{R_{\min,ij}}{r_{ij}} \right)^{6} \right)$$

Buckingham

Summary of Potential Terms


Ethane C-C bond energy profile


Methyl group torsion energy profile


Propane C-C-C angle energy profile

Class II force fields (e.g. MM3, MMFF, UFF, CFF)

$$
\begin{aligned}
&\sum_{bonds}\left[K_{b,2}(b-b_o)^2+K_{b,3}(b-b_o)^3+K_{b,4}(b-b_o)^4\right]\\
&+\sum_{angles}\left[K_{\theta,2}(\theta-\theta_o)^2+K_{\theta,3}(\theta-\theta_o)^3+K_{\theta,4}(\theta-\theta_o)^4\right]\\
&+\sum_{dihedrals}\left[K_{\phi,1}(1-\cos\phi)+K_{\phi,2}(1-\cos 2\phi)+K_{\phi,3}(1-\cos 3\phi)\right]\\
&+\sum_{impropers}K_{\chi}\chi^2\\
&+\sum_{bonds}\sum_{bonds'}K_{bb'}(b-b_o)(b'-b_o')+\sum_{angles}\sum_{angles'}K_{\theta\theta'}(\theta-\theta_o)(\theta'-\theta_o')\\
&+\sum_{bonds}\sum_{angles}K_{b\theta}(b-b_o)(\theta-\theta_o)\\
&+\sum_{bonds}\sum_{dihedrals}(b-b_o)\left[K_{\phi,b1}\cos\phi+K_{\phi,b2}\cos 2\phi+K_{\phi,b3}\cos 3\phi\right]\\
&+\sum_{bonds'}\sum_{dihedrals}(b'-b_o')\left[K_{\phi,b'1}\cos\phi+K_{\phi,b'2}\cos 2\phi+K_{\phi,b'3}\cos 3\phi\right]\\
&+\sum_{angles}\sum_{dihedrals}(\theta-\theta_o)\left[K_{\phi,\theta 1}\cos\phi+K_{\phi,\theta 2}\cos 2\phi+K_{\phi,\theta 3}\cos 3\phi\right]\\
&+\sum_{angles}\sum_{angles'}\sum_{dihedrals}(\theta-\theta_o)(\theta'-\theta_o')\cos\phi
\end{aligned}
$$

Merck Molecular FF: Force field for drug-like molecules

MMFF is a force field designed for pharmaceutical compounds as well as biological molecules. It may be considered one of the better general FFs, although its quality in treating proteins etc. is worse than CHARMM and other biological FFs. Therefore, MMFF is good for computing drug-receptor interactions but not for extensive minimizations etc. of proteins. The tutorial MMFF_Interaction gives an example of reading a drug molecule in Mol2 format, reading a protein structure and calculating the interaction energy. See mmff_inter_energy.inp

Limitation of additive force fields

The use of Coulomb's law with fixed atomic charges to treat the electrostatic interactions is a major simplification in current force fields. It is well known that the electron distribution of a molecule (and, thus, the atomic charges) changes as a function of the electrostatic field around the molecule. This is ignored in additive force fields. To compensate for this omission, the atomic charges are "enhanced" to mimic the polarization of molecules that occurs in a polar, condensed phase environment (e.g. aqueous solution, TIP3P water model dipole moment = 2.35 versus gas phase value of 1.85). This approximation has worked well in the current additive force fields; however, in many cases these models fail. To overcome this, next generation force fields are being developed that explicitly treat electronic polarization.

Methods to include electronic polarization in force fields

Fluctuating charge (CHEQ)

Induced dipoles (PIPF, Berne/Friesner, AMBER)

Classical Drude Oscillator

All methods require that the perturbation of the electronic distribution due to the surrounding electrostatic field be optimized in an iterative fashion. This is due to the change in the "charge distribution" of a system leading to a new electrostatic field which then requires additional re-adjustment of the charge distribution (SCF: self-consistent field calculation). Matrix diagonalization may also be used, but is frequently inaccessible due to the large number of atoms in biological systems. In the end the need to perform an SCF calculation leads to a large increase in computational demands. Special methods to minimize this limitation in MD simulations have been developed (see below).

Fluctuating Charge Model (CHEQ)

Polarization is based on the movement of charge, q, between bonded atoms i and j in response to the surrounding electrostatic field. The extent of charge movement is based on the relative electronegativity, χ, and hardness, J, of the bonded atoms. The electrostatic energy is then obtained from the Coulombic interactions between the relaxed charges.

$$V(q_{ij}) = \chi_{ij} q_{ij} + \frac{1}{2} J_{ij} q_{ij}^2$$

$$\chi_{ij} = \chi_i^{'} + \chi_j^{'} \qquad J_{ij} = J_i^{'} + J_j^{'} + 2J_{ij}^{'}$$

Electronegativity: attraction of an atom for electrons
Hardness: work needed to transfer charge (resistance to charge movement)

Induced Dipole Model

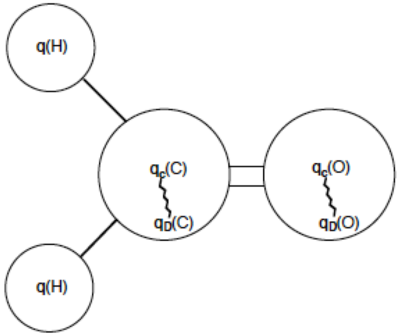
Each atom, i, carries a charge, q_i, and a dipole moment, μ_i, such that electrostatic interactions between atoms i and j include:

charge-charge interactions: $1/r_{ij}$
charge-dipole interactions: $1/r_{ij}^2$
dipole-dipole interactions: $1/r_{ij}^3$

Polarization included via relaxation of dipole moments in the electrostatic field, E_i, where α_i is the polarizability of atom i

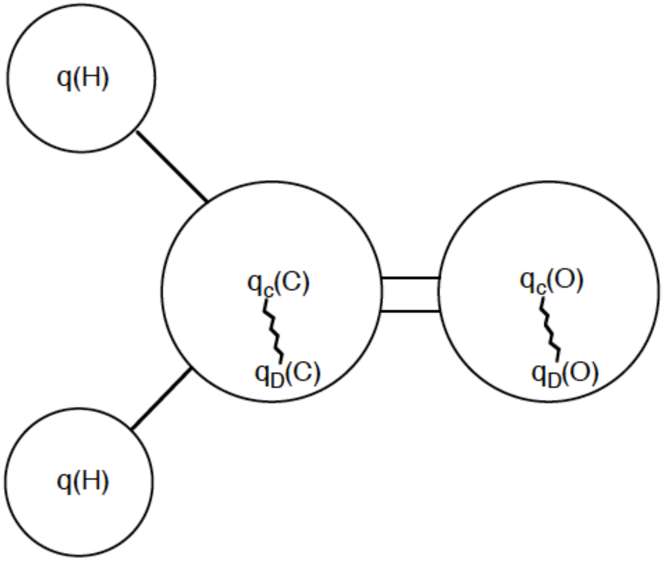
$$\mu_i = \alpha_i\left(E_i^0 + E_i^{induced}\right) = \alpha_i\left(E_i^0 + \sum_{i \neq j} T_{ij}\mu_j\right)$$

Classical Drude Oscillator



To each atom, i, add a virtual particle (Drude) attached to the atomic core via a harmonic spring and place a charge, q_D, on the Drude. The Drudes then relax their positions with respect the surrounding electrostatic field with the relative positions of the Drudes with respect to their parent atom along with the respective charges of each yielding an induced dipole moment on each atom. The electrostatic energy is then obtained from the Coulombic interactions between the atomic and Drude charges.

Classical Drude oscillator



$$\alpha(A) = q_D^2(A) / k_D$$

$$q_D(A) = \sqrt{\alpha(A) / k_D}$$

$$q_c(A) = q(A) - q_D(A)$$

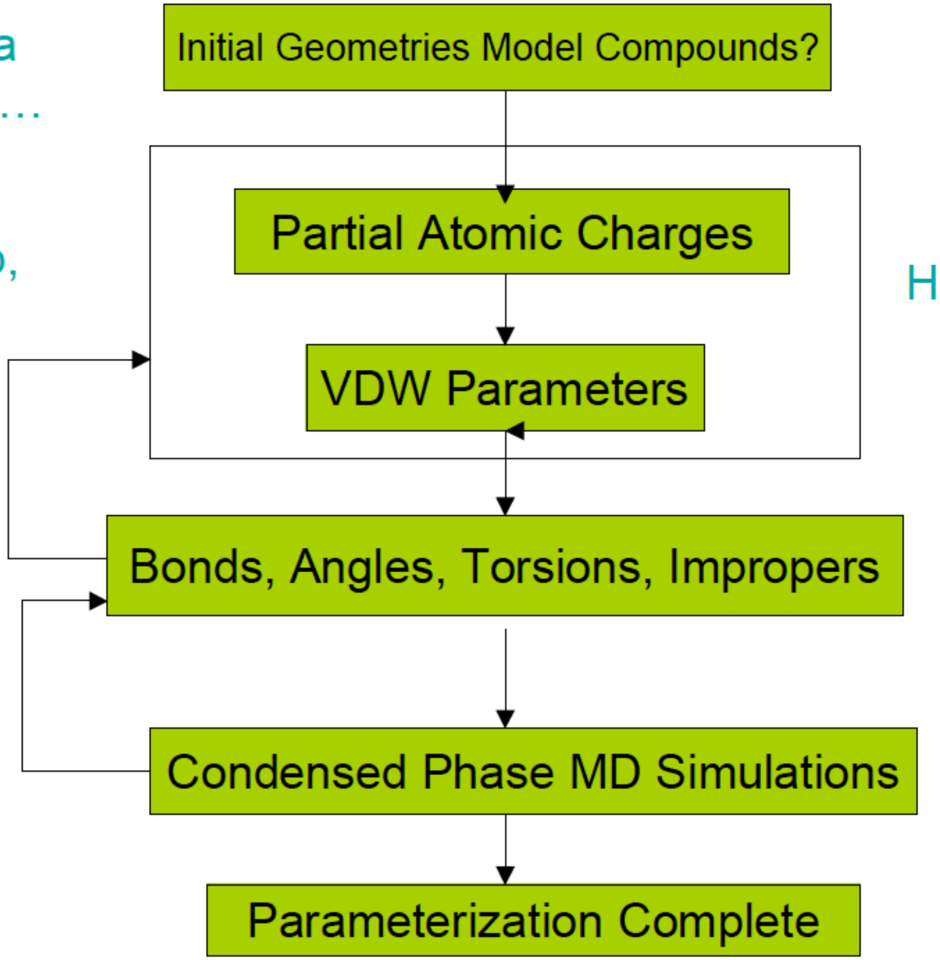
$$U_{Drude} = \sum_{A<B}^{N,N_D} \frac{q_D(A)\cdot q_c(B)}{\left|\mathbf{r}_D(A) - \mathbf{r}(B)\right|} + \sum_{A<B}^{N_D} \frac{q_D(A)\cdot q_D(B)}{\left|\mathbf{r}_D(A) - \mathbf{r}_D(B)\right|} + \frac{1}{2}\sum_{A}^{N_D} k_D \left|\mathbf{r}_D(A) - \mathbf{r}(A)\right|^2$$

Construct New Force Fields

Roadmap Charmm27 Optimization*

Exp. Data
IR,X-ray,…
Stat.Var.

Heat Vap,
Rmin,…



Initial Geometries Model Compounds?

Partial Atomic Charges

VDW Parameters

Bonds, Angles, Torsions, Impropers

Condensed Phase MD Simulations

Parameterization Complete

QM/MP2/6-31G*
Barriers, bonds,…

HF/6-31G* hydrated groups, TIP3W

Self-consistent iteration

*based on MacKerell, JCC v21, 86,105 (2000)

Extension of the additive CHARMM force fields for drug like molecules

1) Decompose molecule into molecular fragments

2) Identify molecular fragments already in the CHARMM force fields

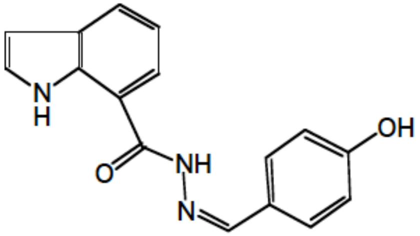
3) Create RTF information for full molecule and molecular fragments (ie. Model compounds) not available (toppar stream file).

4) Identify missing parameters, obtain initial guesses for the new parameters based on analogy to available parameters and place in the toppar stream file.

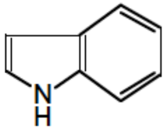
5) Optimize new parameters based on QM data
   i) Geometries and vibrational spectra at MP2/6-31G* (MP2/6-31+G* for anions)
   ii) Conformational energies for rotation of selected dihedrals at MP2/6-31G* (MP2/6-31+G* for anions)
   iii) Partial atomic charges based on reproduction of HF/6-31G* water-model compound interaction energies

6) Perform tests to reproduce experimental data on new molecule if available (structures of many small molecules are available in the Cambridge Structural Database).

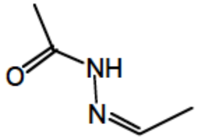
Deconstruct target molecule into molecular fragments for parameter assignment and optimization



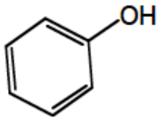
A



B



C

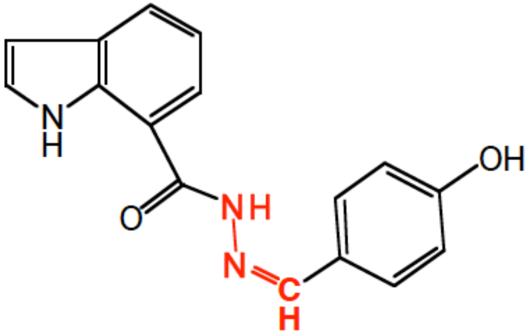


Needs new parameters

A) Indole
B) Hydrazine (model compound 1)
C) Phenol

Linking model compounds: When creating a covalent link between model compounds move the charge on deleted H into the carbon to maintain integer charge (i.e. methyl (q_C=-0.27, q_H=0.09) to methylene (q_C=-0.18, q_H=0.09)

Identify internal parameters to be optimized.
Only optimize new parameters!



Bonds (list doesn't include lipid-protein alkane nomenclature differences)
NH1-NR1, NR1-CEL1
Angles
NR1-NH1-H, NR1-NH1-C, NH1-NR1-CEL1
NR1-CEL1-CTL3, NR1-CEL1-HEL1
Dihedrals
CTL3-C-NH1-NR1, C-NH1-NR1-CEL1, O-C-NH1-NR1,
NH1-NR1-CEL1-HEL1, NH1-NR1-CEL1-CTL3
H-NH1-NR1-CEL1, NR1-CEL1-CTL3-HAL3

Let CHARMM identify missing parameters during IC and energy calls. Add explicit terms if wildcards are used for dihedrals to increase quality of agreement. ONLY include new parameters; do NOT optimize available parameters as this will negatively impact other aspects of the force field. If necessary, create a new atom type for a selected atom to allow for new parameters to be required and optimized.

Parameters by analogy versus optimized parameters

In the following slides various aspects of the parameter optimization process will be given. In slides with results, data labeled "Analogy" represent the results for parameters obtained by analogy to other parameters while the optimized results are those following optimization of the parameters.

Intermolecular Optimization Target Data

A number are methods are available to obtain the charges and LJ parameters as shown below. For the charges, CHARMM is based on the reproduction of QM minimum interaction energies and geometries along with dipole moments. Final tests are performed to reproduce condensed phase properties, although such data is typically not available for drug-like molecules.

- Local/Small Molecule
  - Experimental
    - Interaction enthalpies (MassSpec)
    - Interaction geometries (microwave, crystal)
    - Dipole moments
  - Quantum mechanical
    - Mulliken Population Analysis
    - Electrostatic potential (ESP) based
      - CHELPG (g03: POP=(CHELPG,DIPOLE))
      - Restricted ESP (AMBER)
    - Dimer Interaction Energies and Geometries (OPLS, CHARMM)
    - Dipole moments

- Global/condensed phase (all experimental)
  - Pure solvents (heats of vaporization, density, heat capacity, isocompressibility)
  - Aqueous solution (heats/free energies of solution, partial molar volumes)
  - Crystals (heats of sublimation, lattice parameters, interaction geometries)

CHARMM Partial Atomic Charge Determination

Additive Models: account for lack of explicit inclusion of polarizability via "overcharging" of atoms.

Adjust charges to reproduce HF/6-31G* minimum interaction energies and distances between the model compound and water

- scale target HF/6-31G* interaction energies
  - 1.16 for polar neutral compounds
  - 1.0 for charged compounds
- Empirical distances should be ~0.2 Å shorter the HF/6-31G*

Empirical Dipole moments should be ~10 to 20% large than HF/6-31G* values

For a particular force field do NOT change the QM level of theory for determination of electrostatic parameters. This is necessary to maintain consistency with the remainder of the force field. Thus, use HF/6-31G* for CHARMM additive force fields

LJ (vdw) parameters

Direct transfer from available parameters is generally adequate

Test via
- Heat of vaporization
- Density (Molecular Volume)
- Partial molar volume
- Crystal simulations

For details of LJ parameter optimization see Chen, Yin and MacKerell, JCC, 23:199-213 (2002)

Intramolecular optimization target data

Listed below are the types of target data for the internal parameters. For most drug molecules the amount of experimental data is minimal, requiring the use of QM data. (MP2/6-31G* or MP2/6-31+G* for anions). However, for geometries it is often possible to do surveys of the Cambridge Structural Database for a type of linkage to obtain target geomtries.

Geometries (equilibrium bond, angle, dihedral, UB and improper terms)
  microwave, electron diffraction, *ab initio*
  small molecule x-ray crystallography (CSD)
  crystal surveys of geometries

Vibrational spectra (force constants)
  infrared, raman, *ab initio*

Conformational energies (force constants)
  microwave, *ab initio*

Bonds and angles for model compound B

In gen_model_b.inp, look at geometries after minimization using the IC FILL, IC PRINT commands and compare data with target data. Alternatively, the QUICK commands may be used to obtain the CHARMM geometries for comparison.

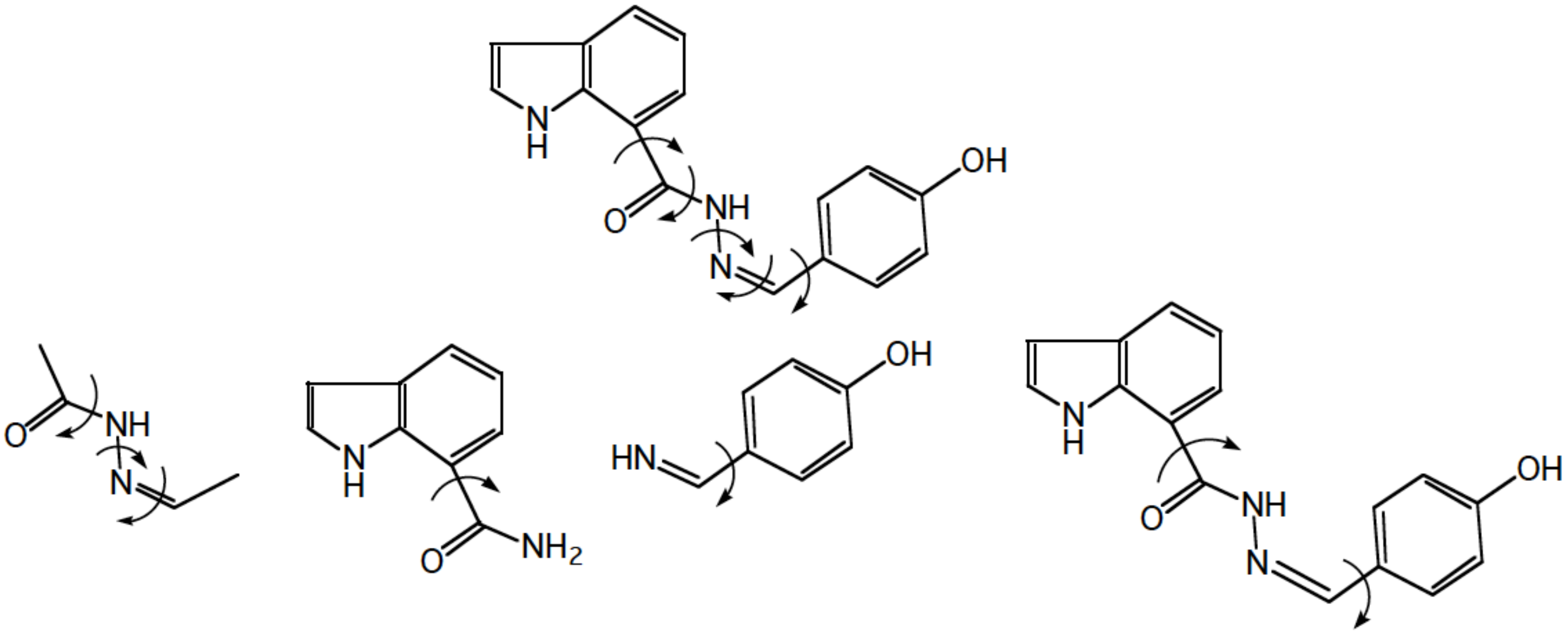
	MP2/6-31G*		CSD		Analogy	Optimized
Bond lengths	1	2	1	2		
C-N[a]	1.385	1.382	1.37±0.03	1.35±0.01	1.342	1.344
N-N	1.370	1.366	1.38±0.02	1.37±0.01	1.386	1.365
N=C	1.289	1.290	1.29±0.02	1.28±0.01	1.339	1.289
Angles						
C-N-N	120.8	122.4	120.7±5.8	119.7±2.9	124.5	121.4
N-N=C	116.0	116.6	114.5±5.3	115.8±1.6	119.6	115.6
N=C-C	119.9	120.0	120.7±4.7	121.2±2.2	122.4	121.0

The MP2/6-31G* results are for the 1) all-trans and 2) 0˚, 180˚, 180˚ global minimum energy structures. The Cambridge structural database results represent mean±standard deviation for all structures with R-factor < 0.1 and 1) the N7 and C10 sites undefined and 2) the N7 and C10 sites explicitly protonated. A) Not optimized as part of the present study.

NH1-NR1 from 400/1.38 to 550/1.36, NR1=CEL1 from 500/1.342 to 680/1.290: C-NH1-NR1 from 50.0/120.0 to 50.0/115.0, NH1- NR1-CEL1 from 50.0/120.0 to 50.0/115.0, NR1-CEL1-CT3 from 48.0/123.5 to 48.0/122.5. For planar systems keep the sum of the equilibrium angle parameters equal to 360.0

Dihedral optimization based on QM potential energy surfaces (HF/6-31G* or MP2/6-31G*).

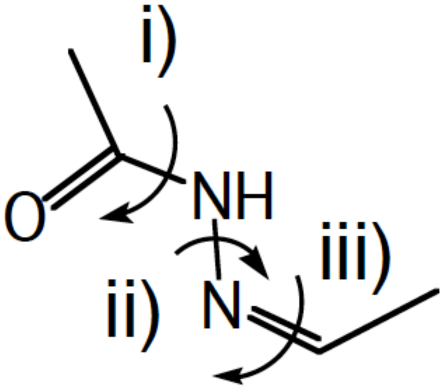
Final optimization of selected dihedrals (typically those containing only non-hydrogen atoms along a rotatable bond) are based on the reproduction of QM potential energy surfaces. This assures that both the relative energy and location of minima are correctly treated as are the barriers to rotation.



Note that additional model compounds may be required.

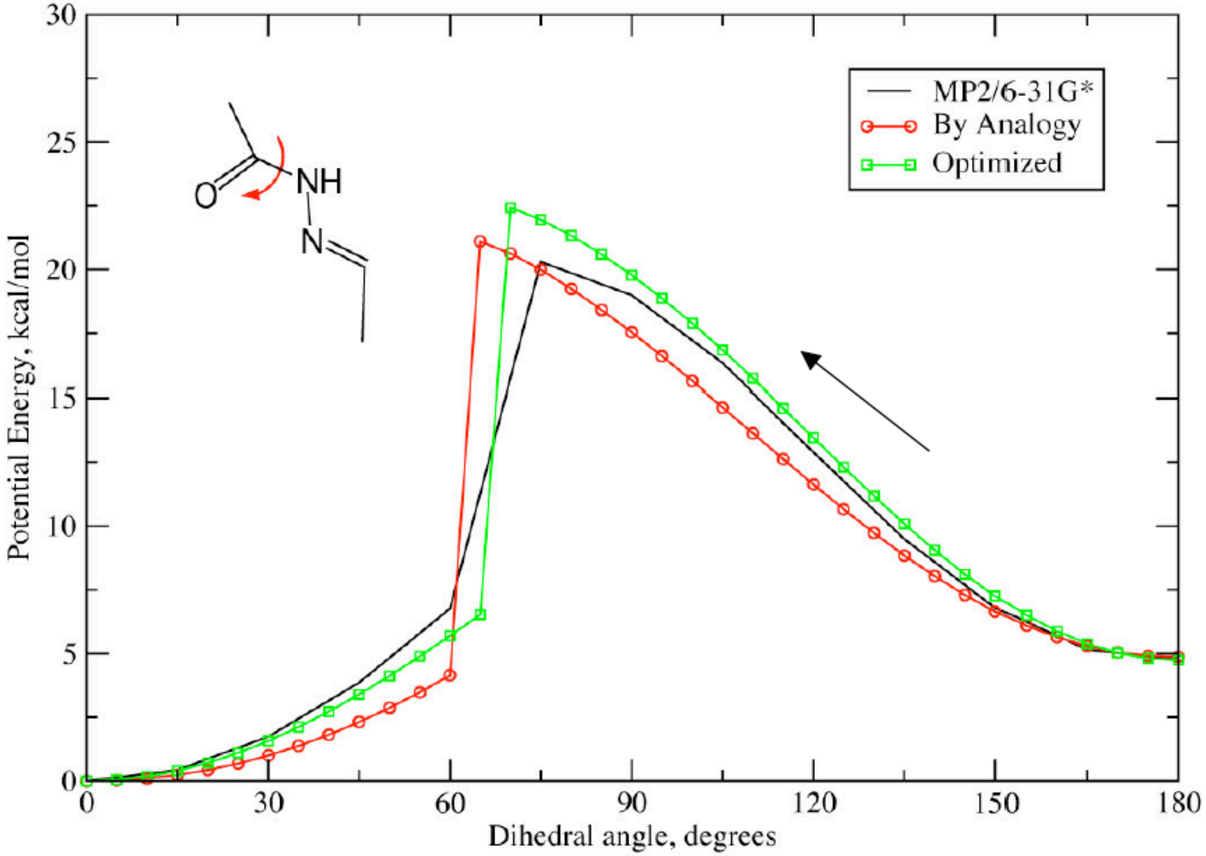
Potential energy surfaces on compounds with multiple rotatable bonds.

Run model_b_surf_all_one.inp followed by model_b_surf_all_two.inp to obtain energy surfaces



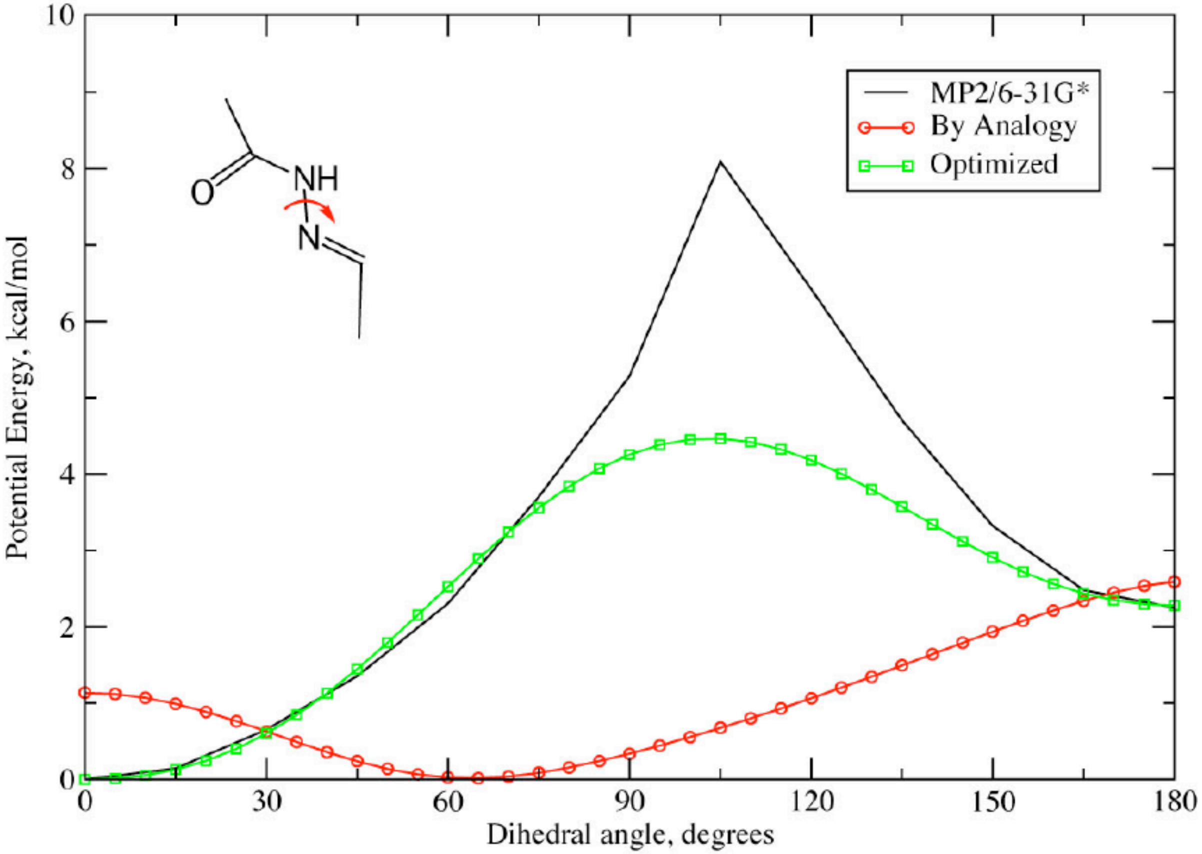
1) Full geometry optimization
2) Constrain n-1 dihedrals to minimum energy values or trans conformation
3) Sample selected dihedral surface
4) Repeat for all rotatable bonds
5) Repeat 2-4 using alternate minima if deemed necessary

Model Compound 1, Surface 1



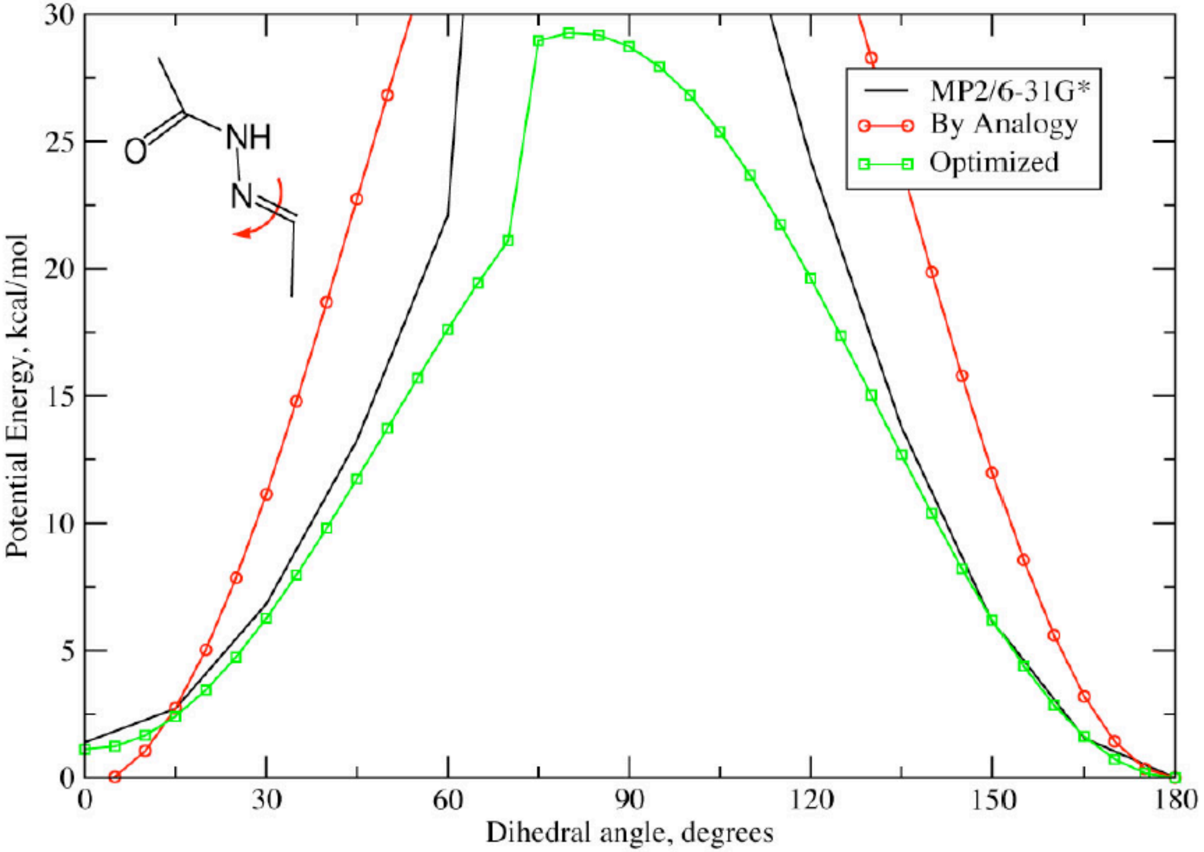
Note that the potential energy surface about a given torsion is the sum of the contributions from ALL terms in the potential energy function, not just the dihedral term. This is the reason why parameter optimization is an iterative process as described above.

Model Compound 1, Surface 2



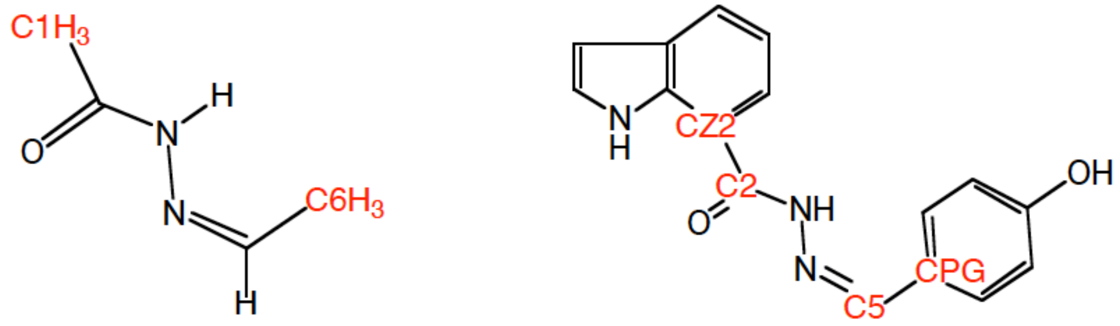
Note the emphasis on fitting the low energy region of the surface as this region is sampled in MD simulations. However, if studies are targeting rotation about that bonds this emphasis must be taken into account when interpreting results.

Model Compound 1, Surface 3



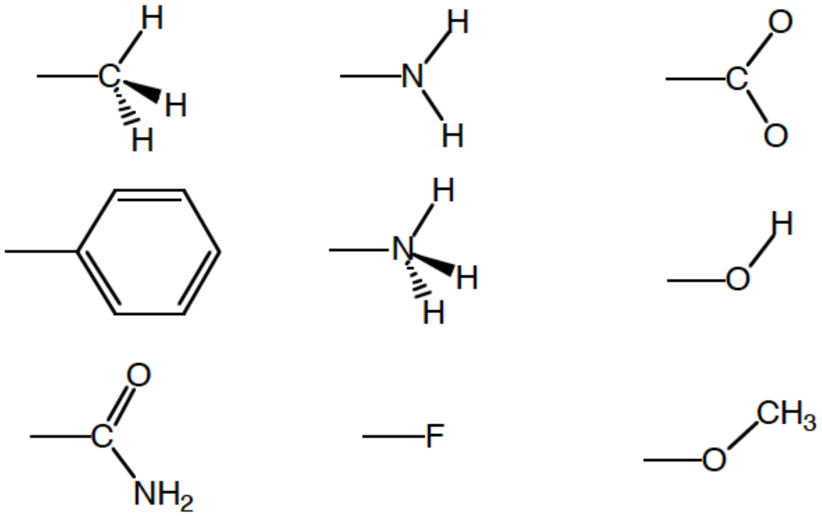
Creation of full drug compound

1) Rename phenol atom types to avoid conflicts with indole (add P to atom type)
2) Delete model 1 terminal methyls, indole and phenol HZ2 and HPG hydrogens, respectively, and perform charge adjustments
    i) Move HZ2 charge (0.115) into CZ2 (-0.115 -> 0.000) total charge on deleted C1 methyl (0.00) onto CZ2 (0.00 -> 0.00)
    ii) Move HPG charge (0.115) into CPG (-0.115 -> 0.000) and move total charge on the C6 methyl (0.18) onto CPG (0.00 -> 0.18)
4) Add parameters by analogy (use CHARMM error messages)
5) Generate IC table (IC GENErate)
6) Generate cartesian coordinates based on IC table (check carefully!)



Lead Optimization

Addition of simple functional groups is generally straightforward once the full compound parameters have been optimized.



Class I potential energy function

$$V_T = \sum_{bonds} K_b (b - b_0)^2 + \sum_{angles} K_\theta (\theta - \theta_0)^2 + \sum_{dihedrals} K_\phi [1 + \cos(n\phi - \delta)]$$

$$+ \sum_{1,3 pairs} K_{ub} (S - S_0)^2 + \sum_{improper} K_w (w - w_0)^2$$

$$+ \sum_{nonbonded} \varepsilon_{ij} \left[\left(\frac{R_{\min,ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{\min,ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4 \pi D r_{ij}}$$

Amber
CHARMM
GROMOS
OPLS

Molecular Mechanics

- Simplest type of calculation
    - Used when systems are very large and approaches that are more accurate become too costly (in time and memory)

- Does not use any quantum mechanics instead uses parameters derived from experimental or *ab initio* data
    - Uses information like bond stretching, bond bending, torsions, electrostatic interactions, van der Waals forces and hydrogen bonding to predict the energetics of a system
    - The energy associated with a certain type of bond is applied throughout the molecule. This leads to a great simplification of the equation

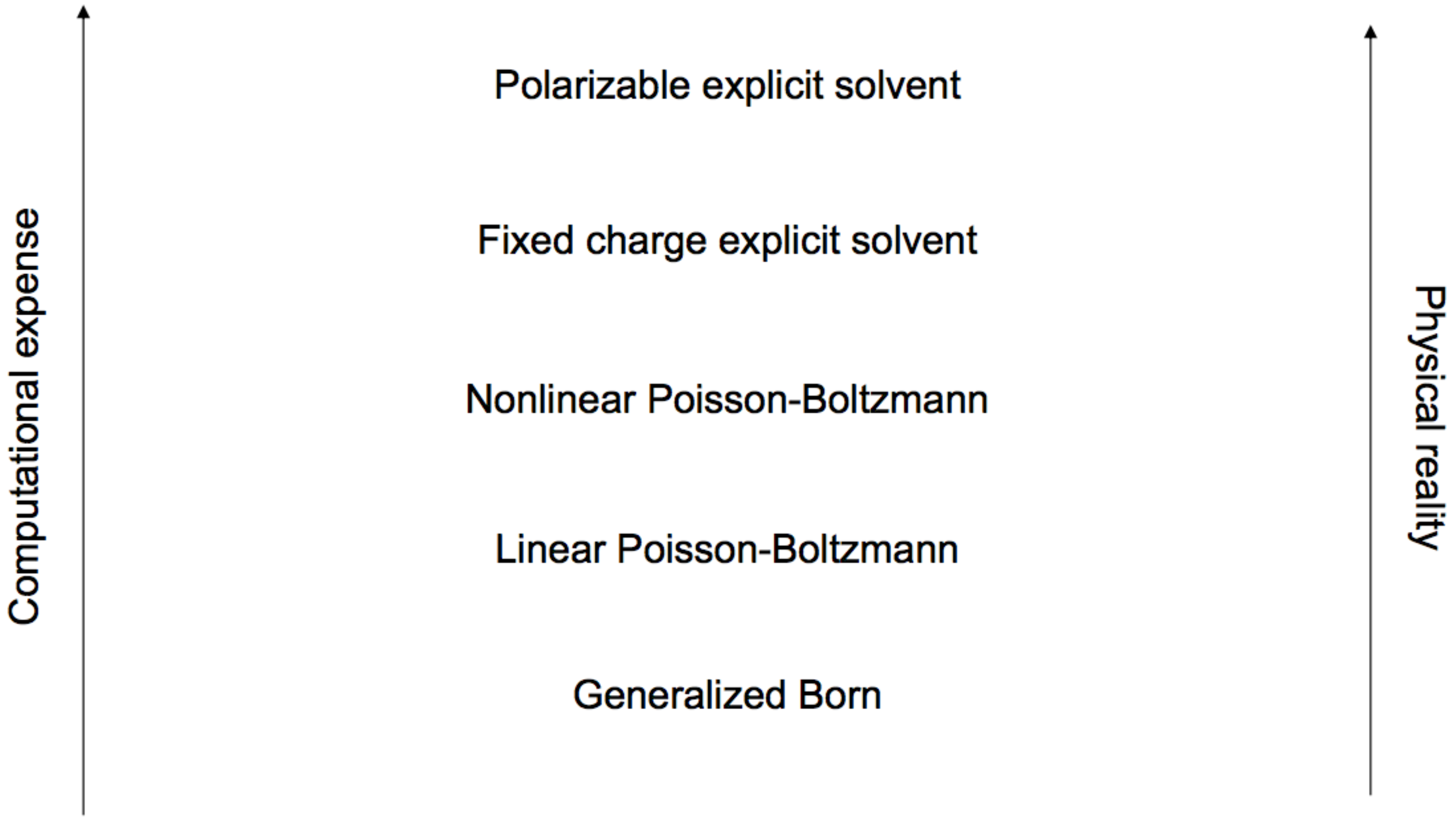
- It should be clarified that the energies obtained from molecular mechanics do not have any physical meaning, but instead describe the difference between varying conformations (type of isomer). Molecular mechanics can supply results in heat of formation if the zero of energy is taken into account.

Courtesy of Shalayna Lair, University of Texas at El Paso

OTHER ISSUES

Solvation, electrostatics…

Solvation Models



Computational expense

Polarizable explicit solvent

Fixed charge explicit solvent

Nonlinear Poisson-Boltzmann

Linear Poisson-Boltzmann

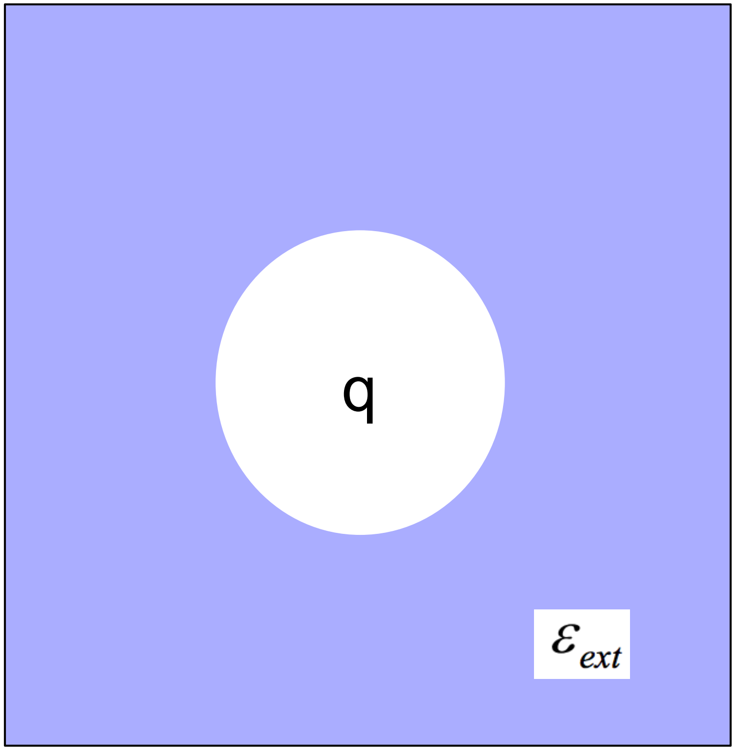
Generalized Born

Physical reality

Distance-dependent dielectric

Surface area based models

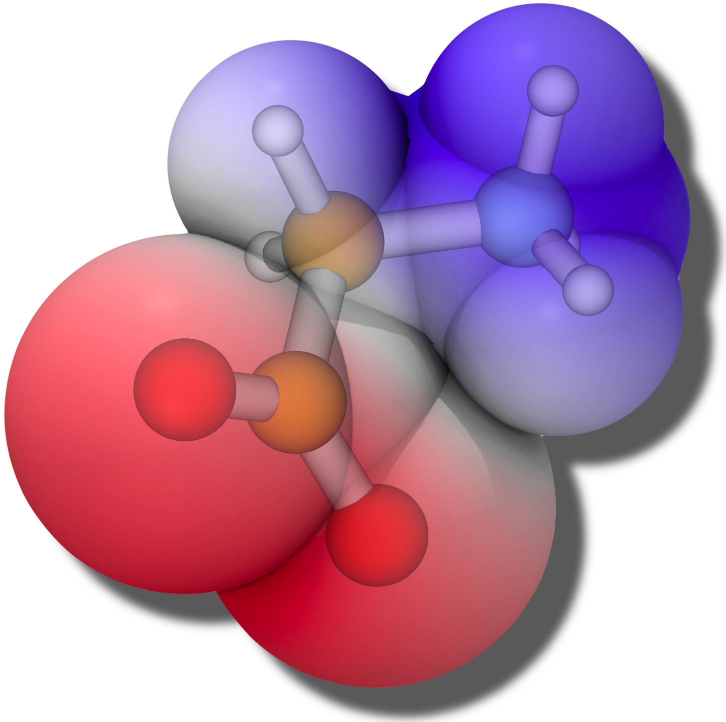
Born Model



$$\Delta G_{solv} = -\frac{1}{2}\left(1 - \frac{1}{\varepsilon_{ext}}\right)\frac{q^2}{R}$$

- Solvent modeled as continuum dielectric medium
- Solvation free energy of a charge easily calculated
- No molecular details, assumes instantaneous solvant relaxation…
- Can be generalized…

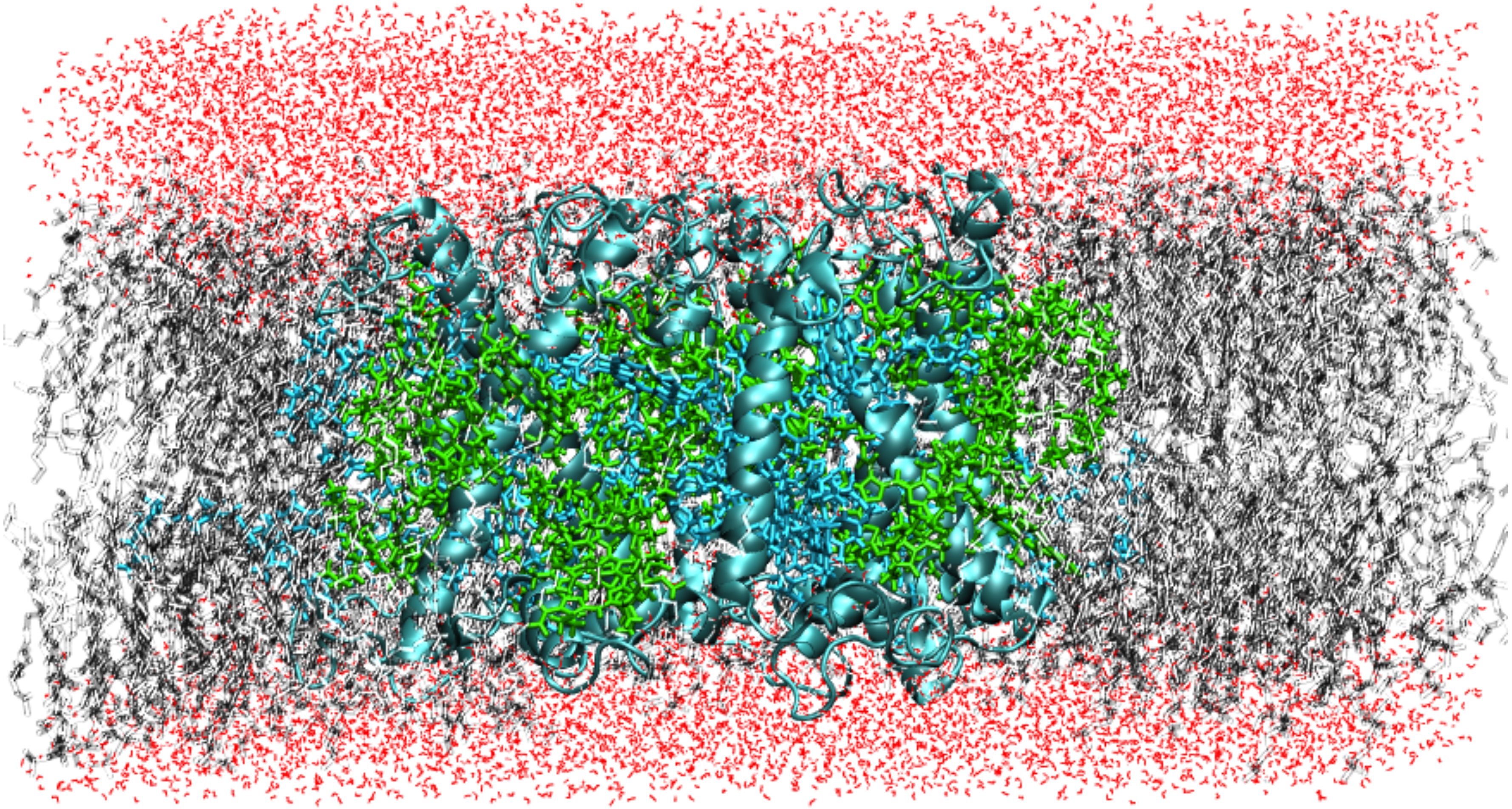
Born Model



- Solvent modeled as continuum dielectric medium
- Solvation free energy of a charge easily calculated
- No molecular details, assumes instantaneous solvant relaxation…
- Can be generalized…

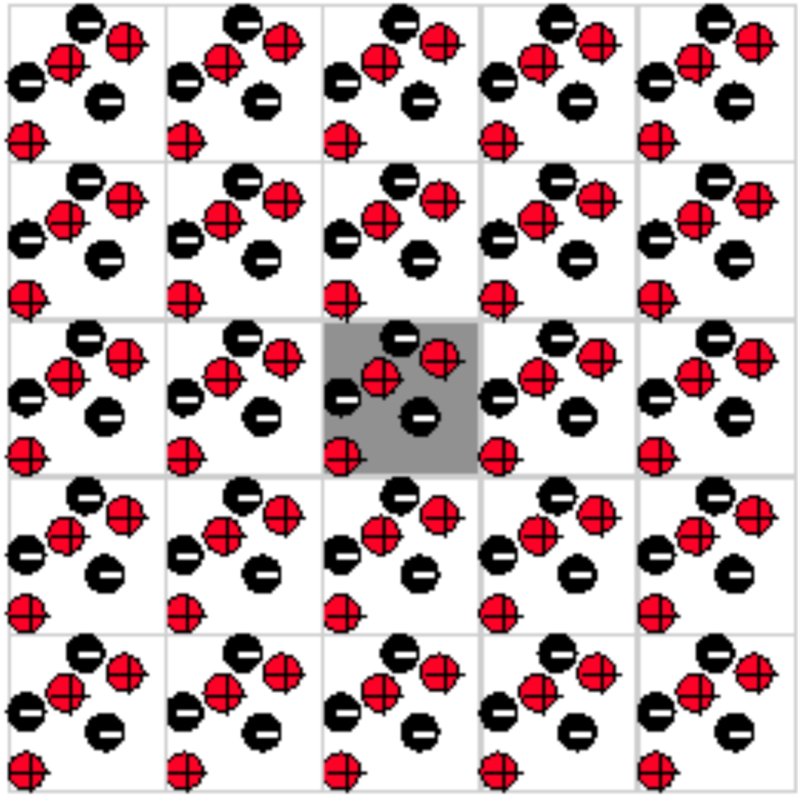
https://chemistry.osu.edu/~herbert/projects/PCM.html

Explicit Solvent Model



Long range electrostatic interactions can be troublesome!

Ewald Sum



Particle-mesh Ewald: combine short-range cut-off and FFT for long-range part on a mesh.