

Force Fields for Classical Molecular Dynamics simulations of Biomolecules

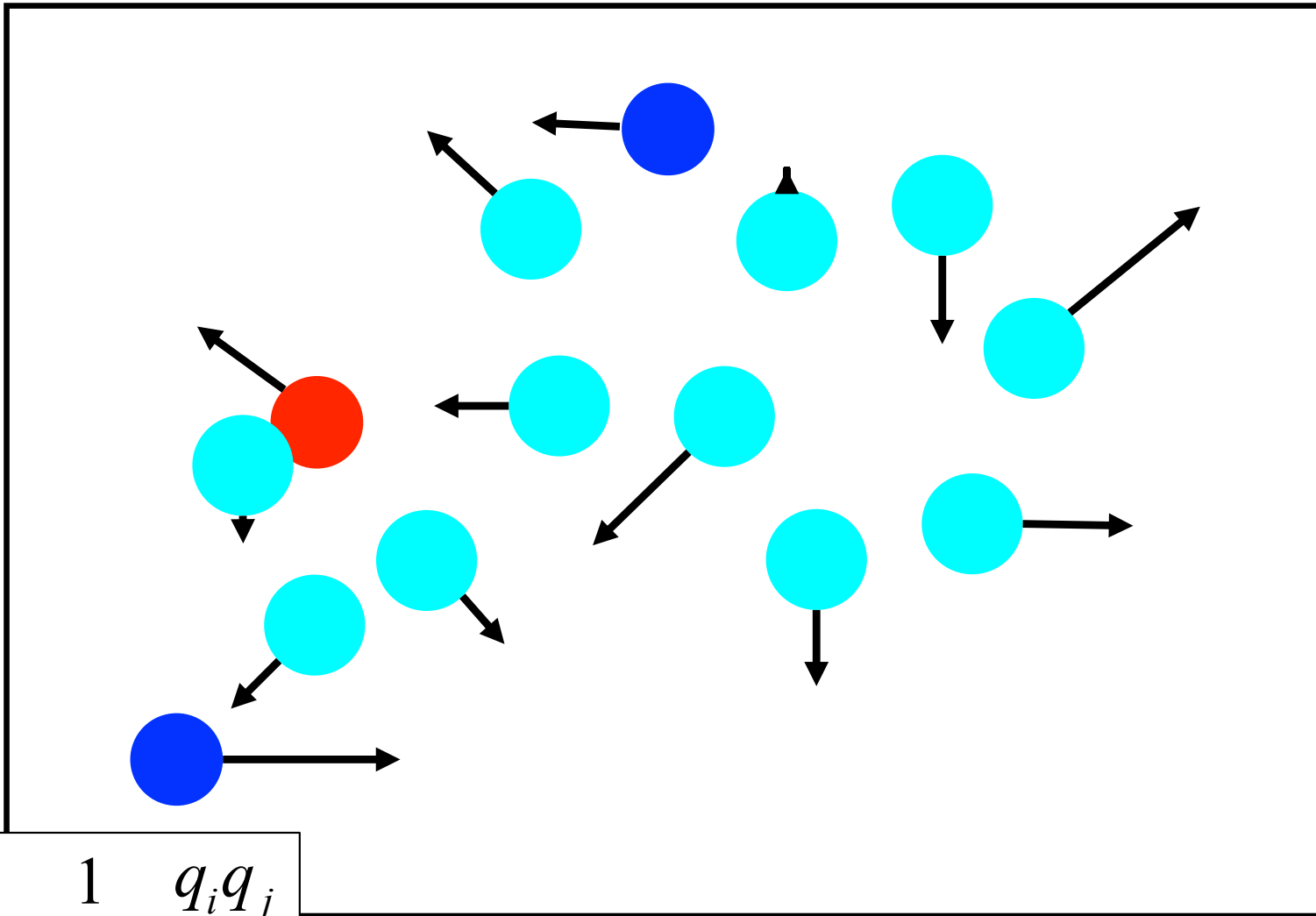
Emad Tajkhorshid

Departments of Biochemistry and Beckman Institute
Center for Biophysics and Computational Biology
University of Illinois at Urbana-Champaign

Classical Force Field Parameters

- Topology and structure files
- Parameter files
- Where do all the numbers needed by an MD code come from?
- Where to find these numbers and how to change them if needed.
- How to make topology files for ligands, cofactors, special amino acids, ...
- How to develop / put together missing parameters.

Classical Molecular Dynamics

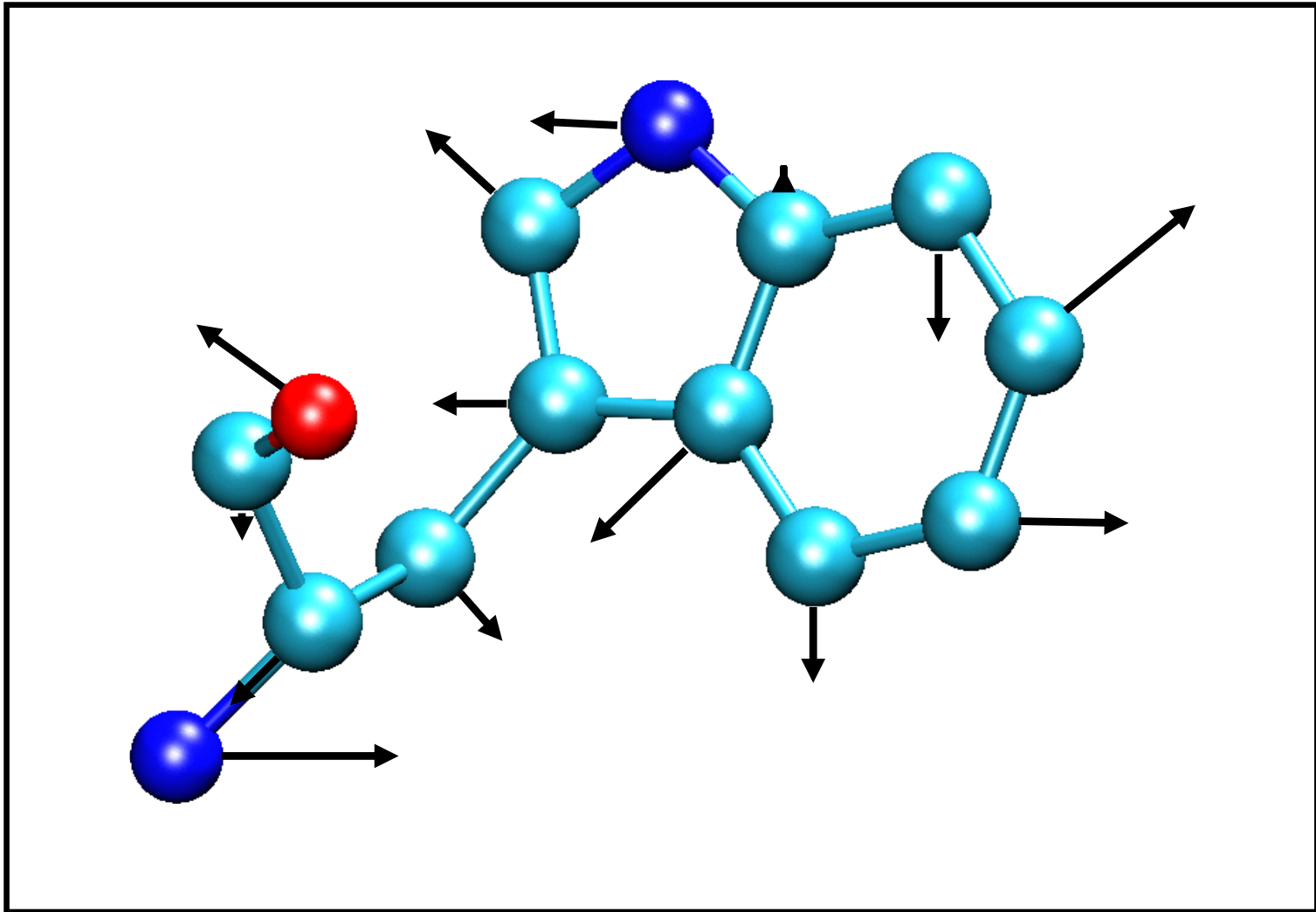


$$U(r) = \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}}$$

Coulomb interaction

$$U(r) = \epsilon_{ij} \left[\left(\frac{R_{min,ij}}{r_{ij}} \right)^{12} - \left(\frac{R_{min,ij}}{r_{ij}} \right)^6 \right]$$

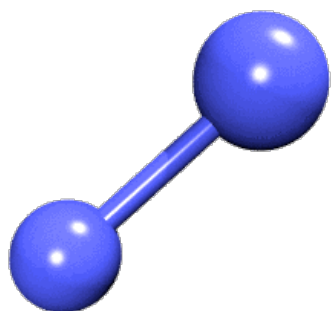
Classical Molecular Dynamics



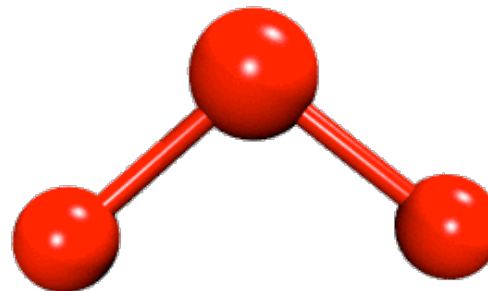
Bond definitions, atom types, atom names, parameters,

Energy Terms Described in

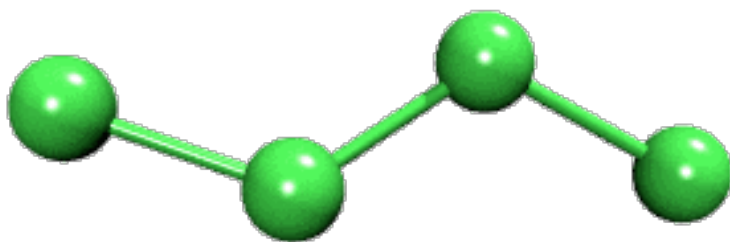
Bond



Angle



Dihedral



Improper



The Potential Energy Function

$$\begin{aligned}
 U(\vec{R}) = & \underbrace{\sum_{\text{bonds}} k_i^{\text{bond}} (r_i - r_0)^2}_{U_{\text{bond}}} + \underbrace{\sum_{\text{angles}} k_i^{\text{angle}} (\theta_i - \theta_0)^2}_{U_{\text{angle}}} + \\
 & \underbrace{\sum_{\text{dihedrals}} k_i^{\text{dihe}} [1 + \cos(n_i \phi_i + \delta_i)]}_{U_{\text{dihedral}}} + \\
 & \underbrace{\sum_i \sum_{j \neq i} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right]}_{U_{\text{nonbond}}} + \sum_i \sum_{j \neq i} \frac{q_i q_j}{\epsilon r_{ij}}
 \end{aligned}$$

U_{bond} = oscillations about the equilibrium bond length

U_{angle} = oscillations of 3 atoms about an equilibrium bond angle

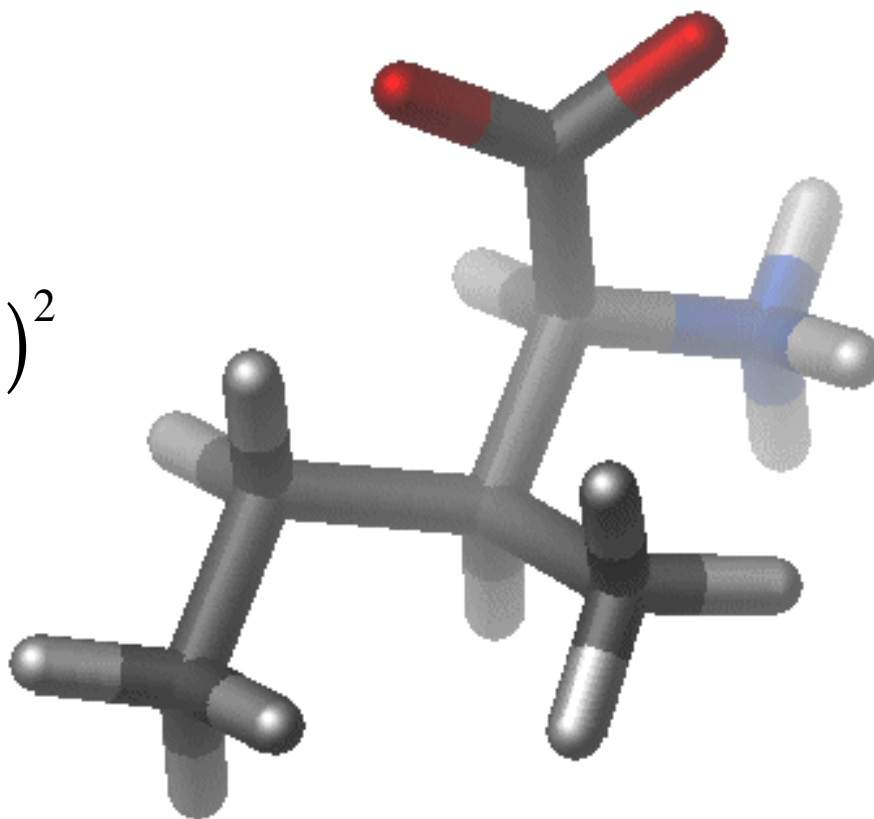
U_{dihedral} = torsional rotation of 4 atoms about a central bond

U_{nonbond} = non-bonded energy terms (electrostatics and Lenard-Jones)

Interactions between bonded atoms

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

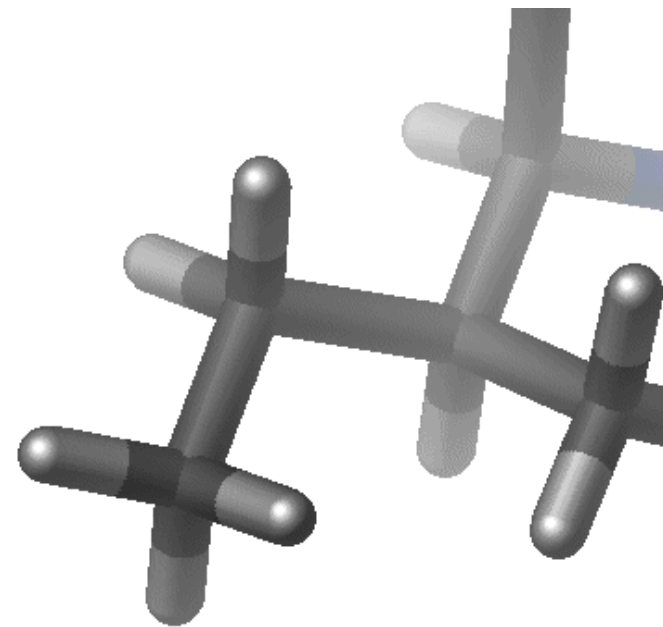
$$V_{bond} = K_b (b - b_o)^2$$



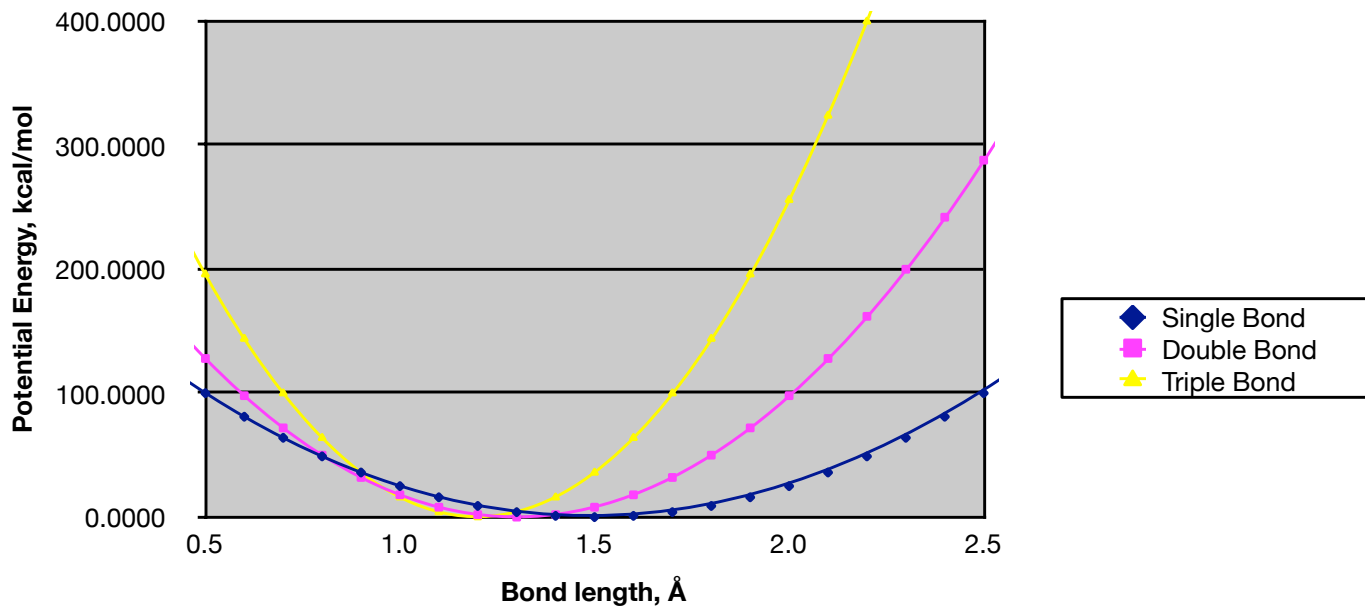
$$V_{dihedral} = K_{\phi} (1 + \cos(n\phi - \delta))$$

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

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



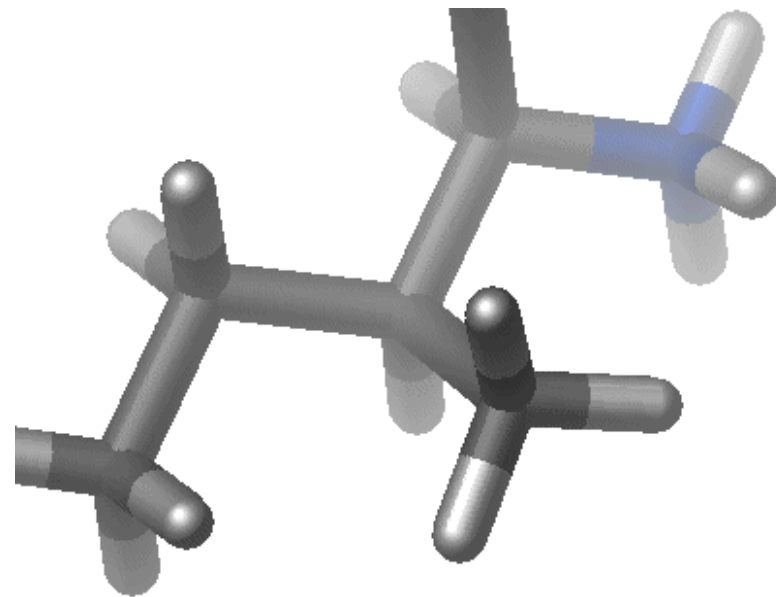
Bond Energy versus Bond length



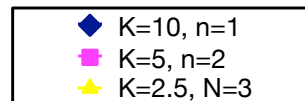
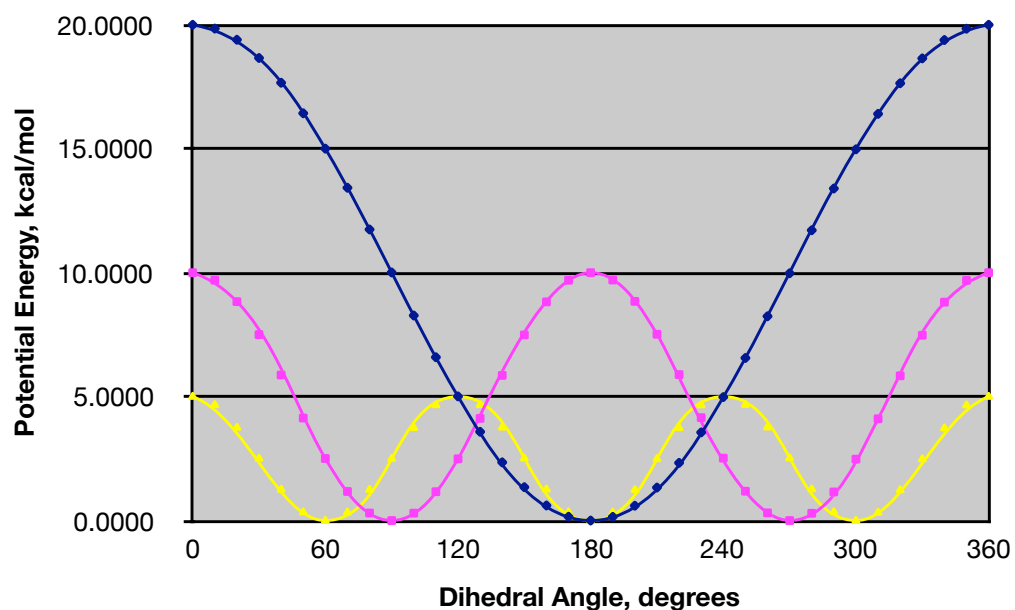
Bond angles and *improper* terms have similar quadratic forms, but with softer spring constants. The force constants can be obtained from vibrational analysis of the molecule (experimentally or theoretically).

Dihedral Potential

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



Dihedral energy versus dihedral angle



$$\delta = 0^\circ$$

Nonbonded Parameters

$$\sum_{\text{non-bonded}} \frac{q_i q_j}{4\pi D r_{ij}} + \epsilon_{ij} \left[\left(\frac{R_{\min,ij}}{r_{ij}} \right)^{12} - \left(\frac{R_{\min,ij}}{r_{ij}} \right)^6 \right]$$

q_i : partial atomic charge

D : dielectric constant

ϵ : Lennard-Jones (LJ, vdW) well-depth

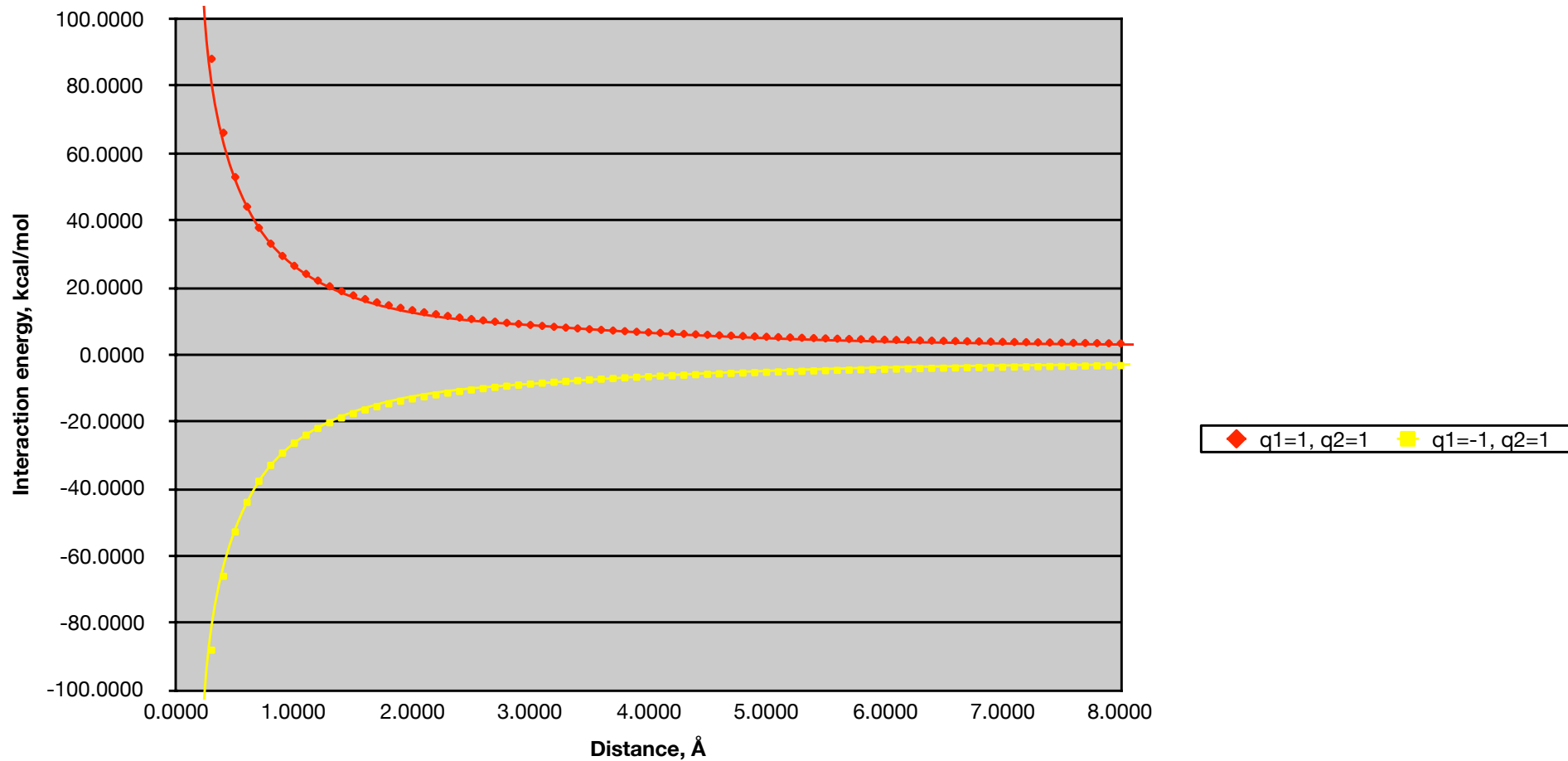
R_{\min} : LJ radius ($R_{\min}/2$ in CHARMM)

Combining rules (CHARMM, Amber)

$$R_{\min\ i,j} = R_{\min\ i} + R_{\min\ j}$$

$$\epsilon_{i,j} = \text{SQRT}(\epsilon_i * \epsilon_j)$$

Electrostatic Energy versus Distance



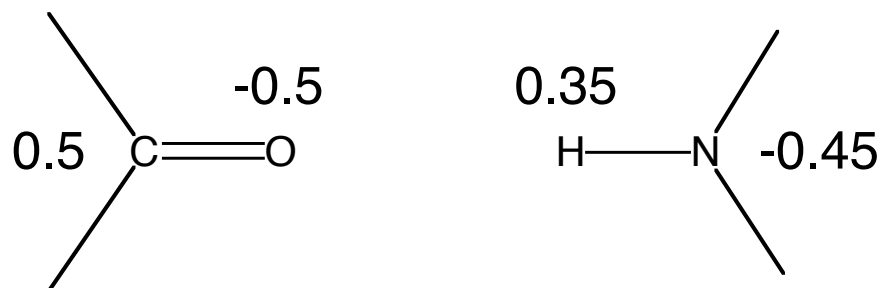
Note that the effect is long range.

Charge Fitting Strategy

CHARMM- Mulliken*

AMBER(ESP/RESP)

Partial atomic charges



*Modifications based on interactions with TIP3 water

CHARMM Potential Function

$$\begin{aligned}
 U(\vec{R}) = & \underbrace{\sum_{\text{bonds}} k_i^{\text{bond}} (r_i - r_0)^2}_{U_{\text{bond}}} + \underbrace{\sum_{\text{angles}} k_i^{\text{angle}} (\theta_i - \theta_0)^2}_{U_{\text{angle}}} + \\
 & \underbrace{\sum_{\text{dihedrals}} k_i^{\text{dihe}} [1 + \cos(n_i \phi_i + \delta_i)]}_{U_{\text{dihedral}}} + \underbrace{\sum_i \sum_{j \neq i} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right]}_{U_{\text{nonbond}}} + \sum_i \sum_{j \neq i} \frac{q_i q_j}{\epsilon r_{ij}}
 \end{aligned}$$

PDB file (points to k_i^{bond} , k_i^{angle} , k_i^{dihe})
geometry (points to k_i^{dihe})
Topology (points to $n_i \phi_i + \delta_i$)
PSF file (points to $n_i \phi_i + \delta_i$)
parameters (points to ϵ_{ij} , σ_{ij} , r_{ij})
Parameter file (points to ϵ_{ij} , σ_{ij} , r_{ij})

File Format/Structure

- The structure of a pdb file
- The structure of a psf file
- The topology file
- The parameter file
- Connection to potential energy terms

Looking at File Structures

- PDB file
- Topology file
- PSF file
- Parameter file

Parameter Optimization Strategies

Check if it has been parameterized by somebody else

Literature

Google

Minimal optimization

By analogy (direct transfer of known parameters)

Quick, starting point

Maximal optimization

Time-consuming

Requires appropriate experimental and target data

Choice based on goal of the calculations

Minimal

database screening

NMR/X-ray structure determination

Maximal

free energy calculations, mechanistic studies,
subtle environmental effects

Getting Started

- Identify previously parameterized compounds
- Access topology information – assign atom types, connectivity, and charges – **annotate changes**

CHARMM topology (parameter files)

top_all22_model.inp (par_all22_prot.inp)
top_all22_prot.inp (par_all22_prot.inp)
top_all22_sugar.inp (par_all22_sugar.inp)
top_all27_lipid.rtf (par_all27_lipid.prm)
top_all27_na.rtf (par_all27_na.prm)
top_all27_na_lipid.rtf (par_all27_na_lipid.prm)
top_all27_prot_lipid.rtf (par_all27_prot_lipid.prm)
top_all27_prot_na.rtf (par_all27_prot_na.prm)
toph19.inp (param19.inp)

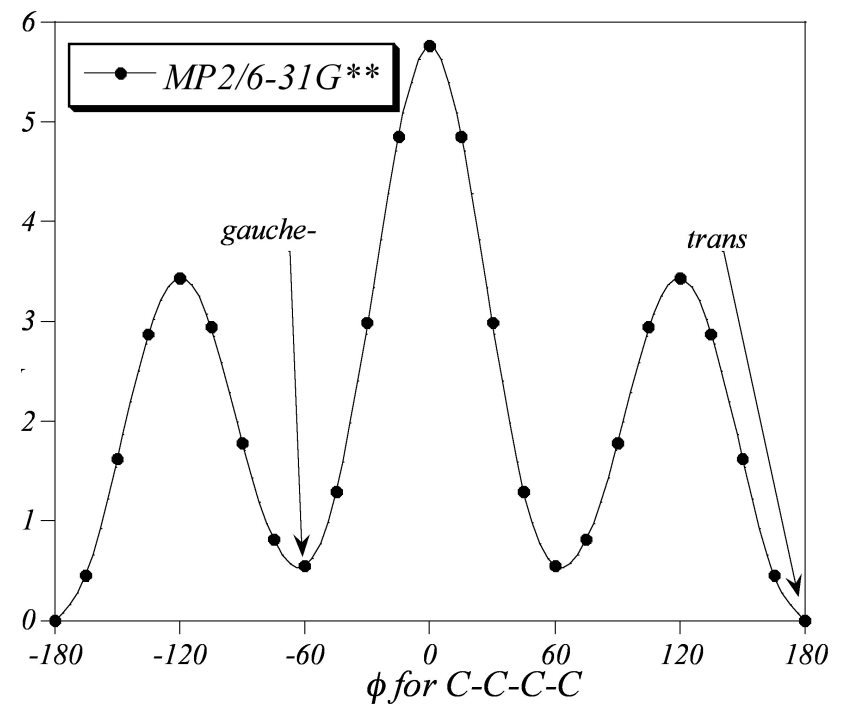
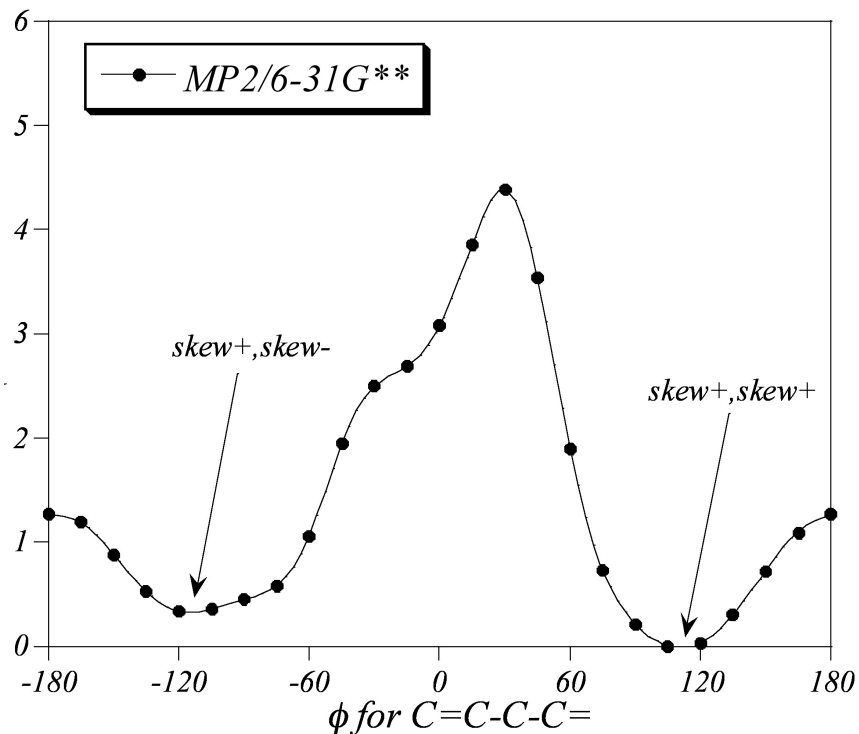
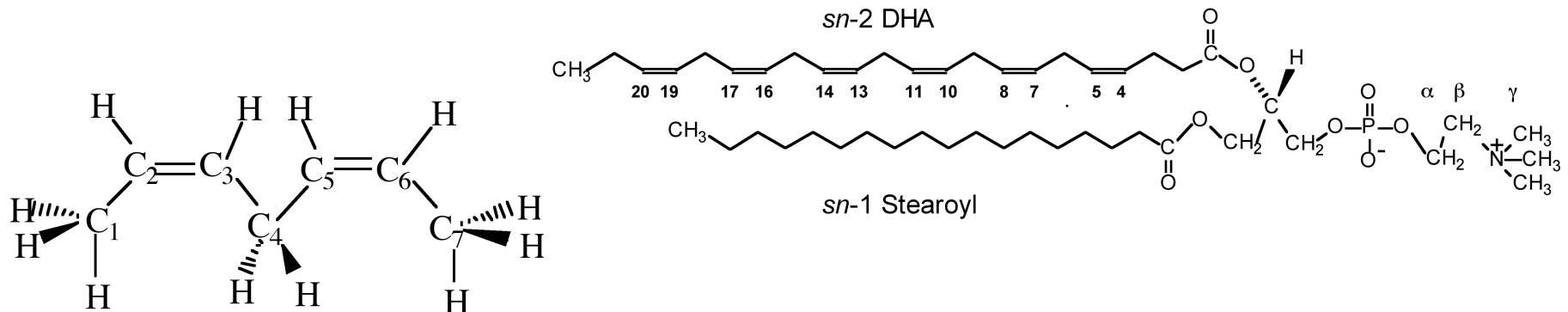
NA and lipid force fields have new LJ parameters for the alkanes, representing increased optimization of the protein alkane parameters. Tests have shown that these are compatible (e.g. in protein-nucleic acid simulations). For new systems is suggested that the new LJ parameters be used. Note that only the LJ parameters were changed; the internal parameters are identical

Partial Charge Assignment

- Most important aspect for ligands
- Different force fields might take different philosophies
 - AMBER: RESP charges at the HF/6-31G level
 - Overestimation of dipole moments
 - Easier to set up
 - CHARMM: Interaction based optimization
 - TIP3P water representing the environment
 - Could be very difficult to set up
- Conformation dependence of partial charges
- Lack of polarization
- Try to be consistent within the force field
- pKa calculations for titratable residues

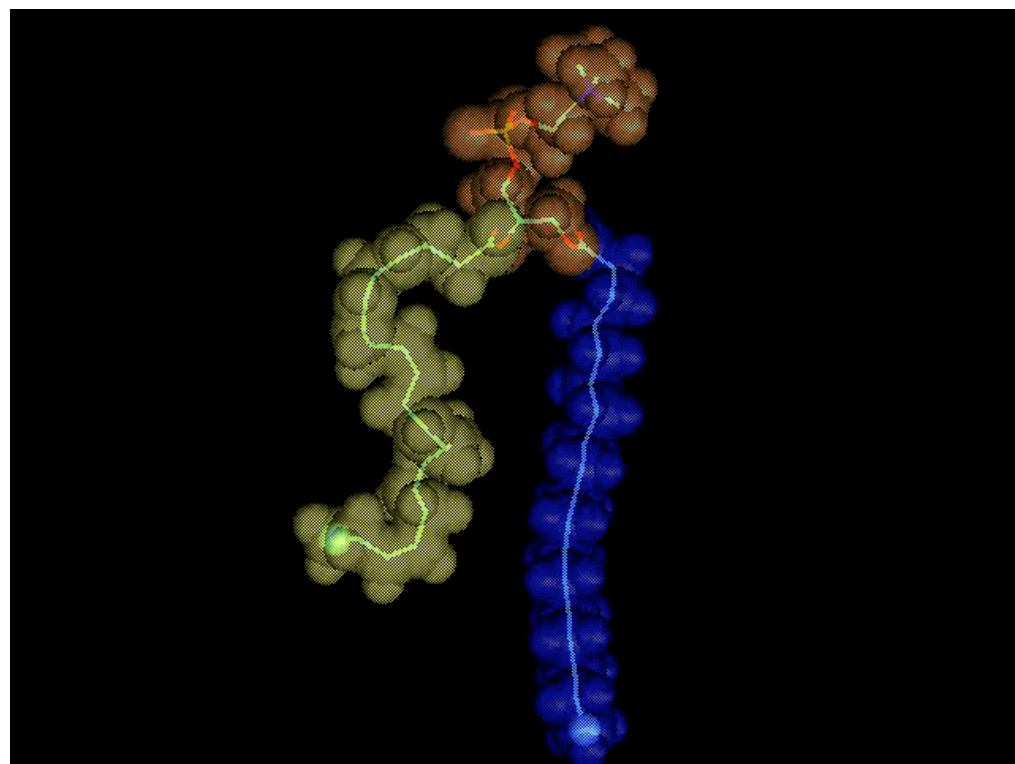
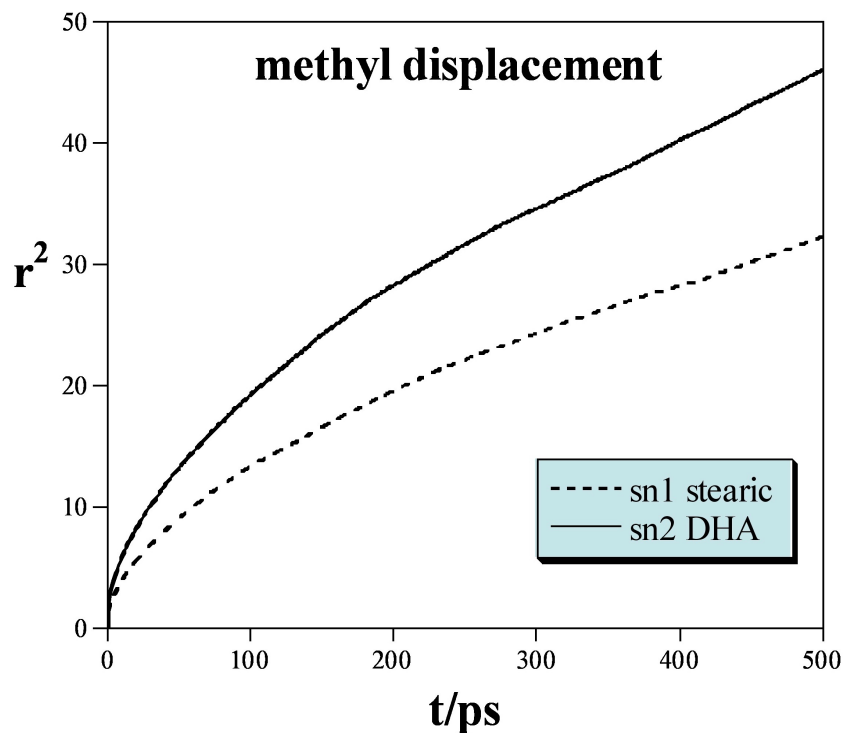
Parameterization of unsaturated lipids

- All C=C bonds are cis, what does rotation about neighboring single bonds look like?



Dynamics of saturated vs. polyunsaturated lipid chains

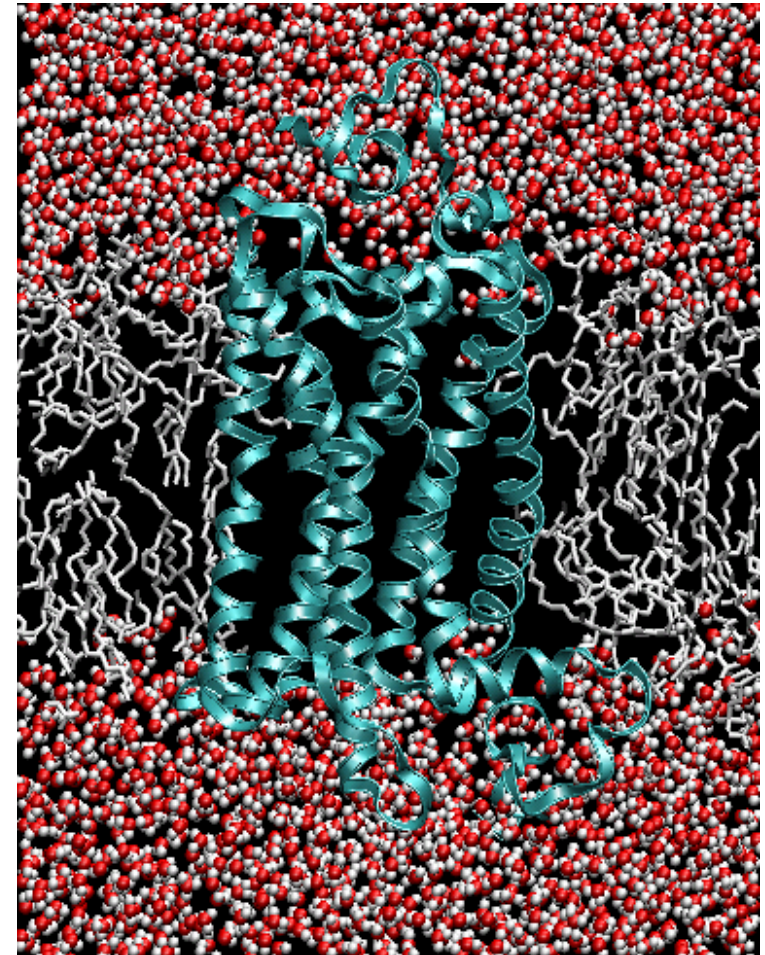
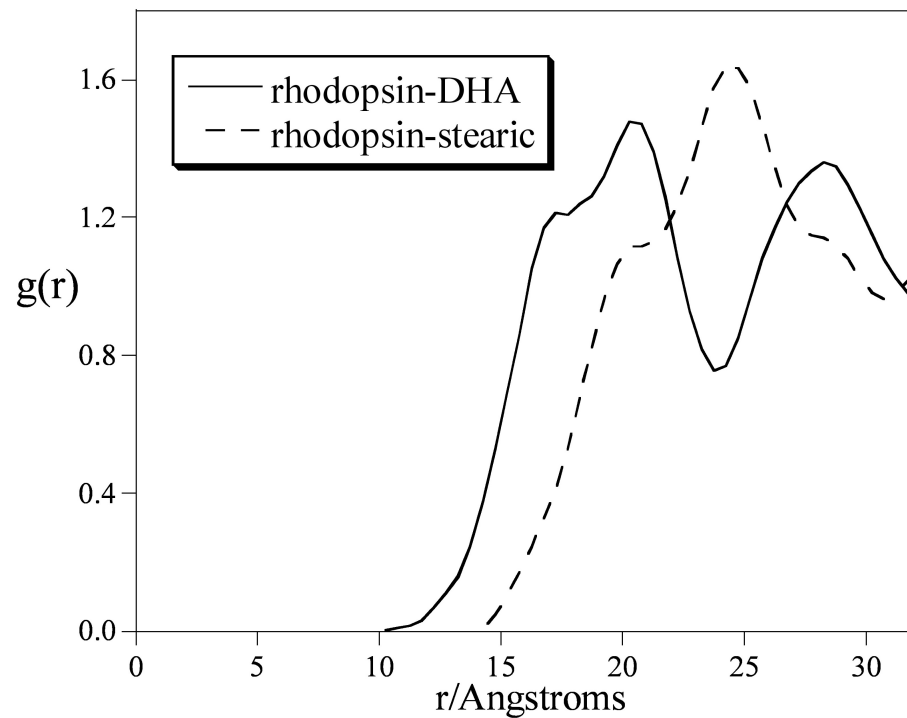
- *sn*1 stearic acid = blue
- *sn*2 DHA = yellow
- 500 ps of dynamics



Movie courtesy of Mauricio Carrillo Tripp

Lipid-protein interactions

- Radial distribution around protein shows distinct layering of acyl chains



Courtesy of Scott Feller, Wabash College

Major Recent Developments

- New set of lipid force field parameters for CHARMM (CHARMM32+)
 - Pastor, B. Brooks, MacKerell
- Polarizable force field
 - Roux, MacKerell

A Brief Overview of The Force Field Toolkit (ffTK)

Dr. Christopher G. Mayne
Tajkhorshid Group
February 13th, 2012

Available Methods for Obtaining Parameters

Analogy (Re-use of parameters from similar structures)

ParamChem <https://www.paramchem.org/>

SwissParam <http://swissparam.ch/>

Zoete *et al.*; *J. Comp. Chem.* 32(11) **2011**, pp.2359-2368

MATCH <http://brooks.chem.lsa.umich.edu/software>

Yesselman *et al.*; *J. Comp. Chem.* 33(2) **2012**, pp.189-202

Development

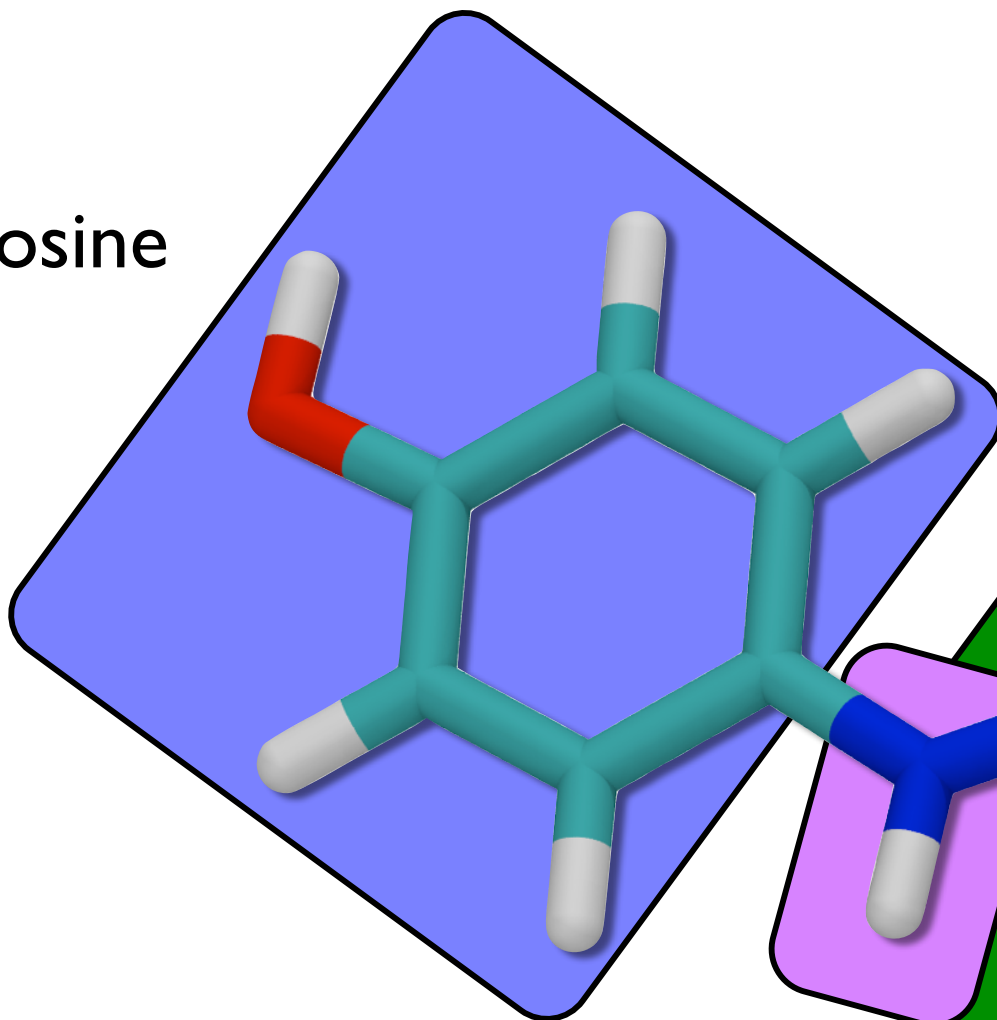
ParaTool <http://www.ks.uiuc.edu/Research/vmd/plugins/paratool/>

GAAMP <http://gaamp.lcrc.anl.gov/para-fit.html>

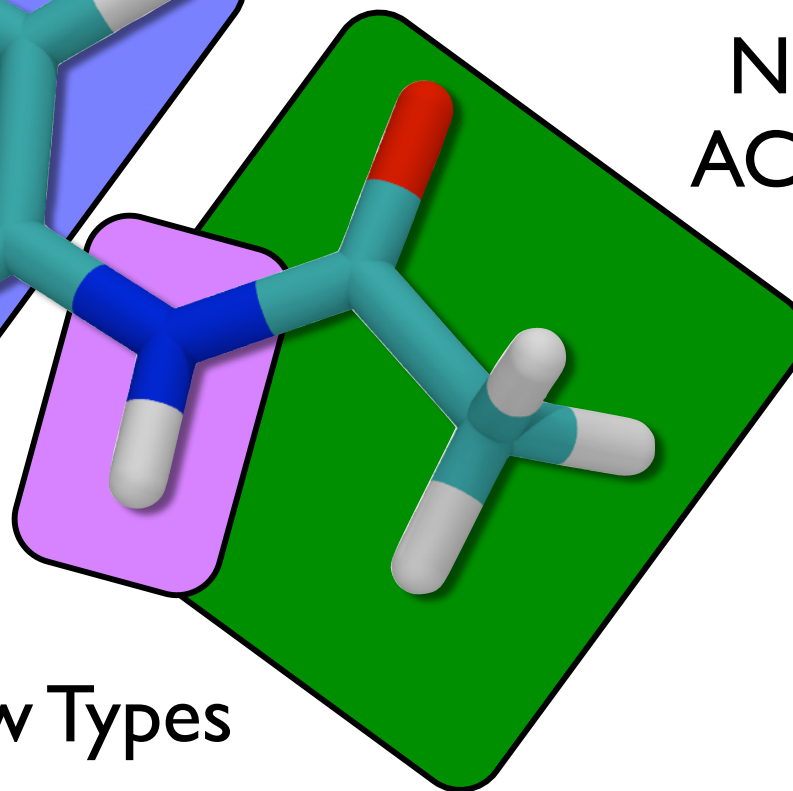
ffTK <http://www.ks.uiuc.edu/Research/vmd/plugins/fftk/>

An Example: Acetaminophen

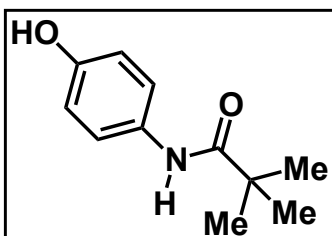
Tyrosine



NMA or
ACE-patch



New Types



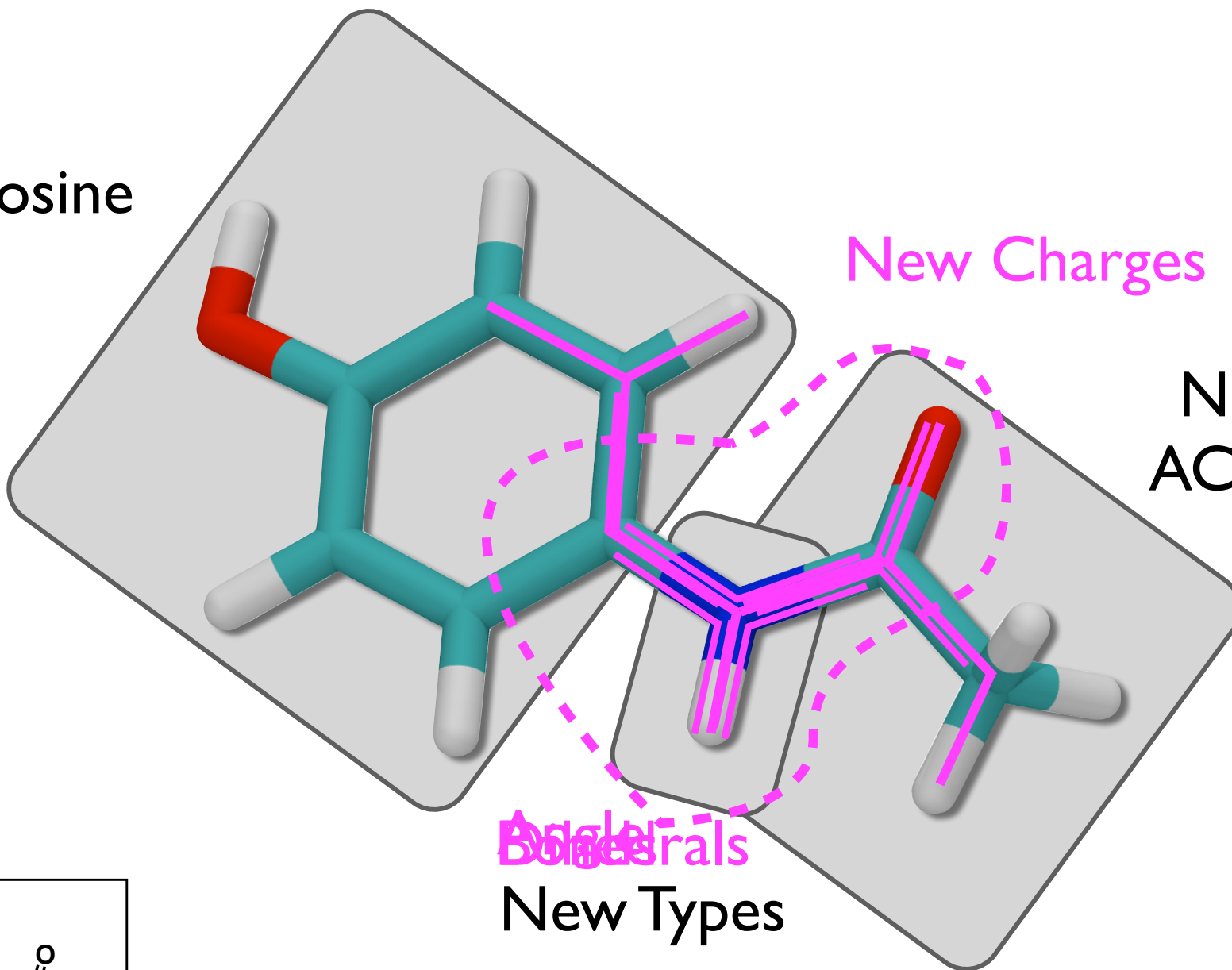
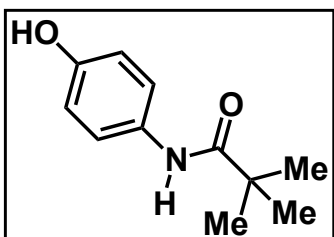
An Example: Acetaminophen

Tyrosine

New Charges

NMA or
ACE-patch

Bonds
Angles
New Types



ffTK Facilitates the Parameterization Workflow

PSF/PDB

Find Missing Parameters

~~SYSTEM PREPARATION~~

Geometry Optimization (QM)

Water Interaction En. (QM)

~~CHARGES~~

Charge Optimization

Frequency Calculation (QM)

~~BONDS & ANGLES~~

Transform Hessian

Torsion Scan (QM)

~~DIPEDRALS / TORSIONS~~

Torsion Optimization

Calculation

build init PAR

update PDB

update PSF

update PAR

update PAR

Action

PAR File

ffTK Interface

file dialog buttons

entry boxes to hold variables to separate tabs action buttons

The image shows a screenshot of the ffTK interface with several annotations. At the top, there are tabs for different tasks: BuildPar, Opt. Geometry, Water Int., Opt. Charges (selected), Calc. Bonded, Scan Torsions, and Opt. Torsions. Below the tabs is the 'Input' section, which contains several entry boxes and buttons. The 'PSF File' box contains '/Users/cmayne/Desktop/test11/01-sysprep/pacp.psf' and has a 'Browse' button next to it. The 'PDB File' box contains '/Users/cmayne/Desktop/test11/02-geoopt/pacp-opt.pdb' and also has a 'Browse' button. The 'Residue Name' box contains 'PACP' and has a 'Resname From TOP' button. To the right of these boxes are buttons for 'Load PSF/PDB', 'Label Atoms', and a dropdown menu for 'Name'. Below the input section is a 'Parameter Files' section, which contains a list of files: '/Users/cmayne/Desktop/test11/common/pacp-init.par' and '/Users/cmayne/Desktop/test11/common/par_all22_prot.inp'. To the right of this list are buttons for 'Add', 'Delete', and 'Clear'. At the bottom of the input section are 'NAMD binary' and 'Output LOG' boxes. The 'NAMD binary' box contains 'namd2' and has a 'Browse' button. The 'Output LOG' box contains 'tmp.log' and has a 'SaveAs' button. At the bottom of the interface is a treeview with four items: 'Charge Constraints', 'QM Target Data', 'Advanced Settings', and 'Results'. Annotations with arrows point to various parts of the interface: 'entry boxes to hold variables' points to the PSF File box; 'file dialog buttons' points to the Browse buttons; 'action buttons' points to the Load PSF/PDB, Label Atoms, and Name buttons; 'treeview boxes to hold lists' points to the Parameter Files list; 'action menus' points to the Add, Delete, and Clear buttons; and 'back' points to the treeview items. There is also an arrow labeled 'arrow' pointing to the left side of the interface.

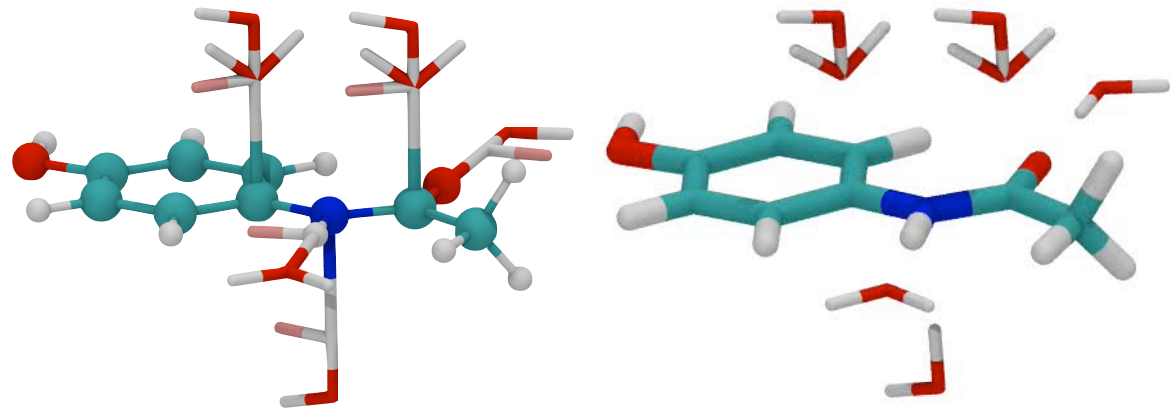
ffTK 1.0 is Available in VMD 1.9.1

Setup necessary QM calculations

Visualize calculation input/output

Multidimensional optimization
of developing parameters

Read/Write files (pdb, par, log, gau)



Full documentation online:

<http://www.ks.uiuc.edu/Research/vmd/plugins/fftk/>

Tutorial is under development

