

# Force Fields for Classical Molecular Dynamics simulations of Biomolecules

Emad Tajkhorshid

Beckman Institute

Departments of Biochemistry

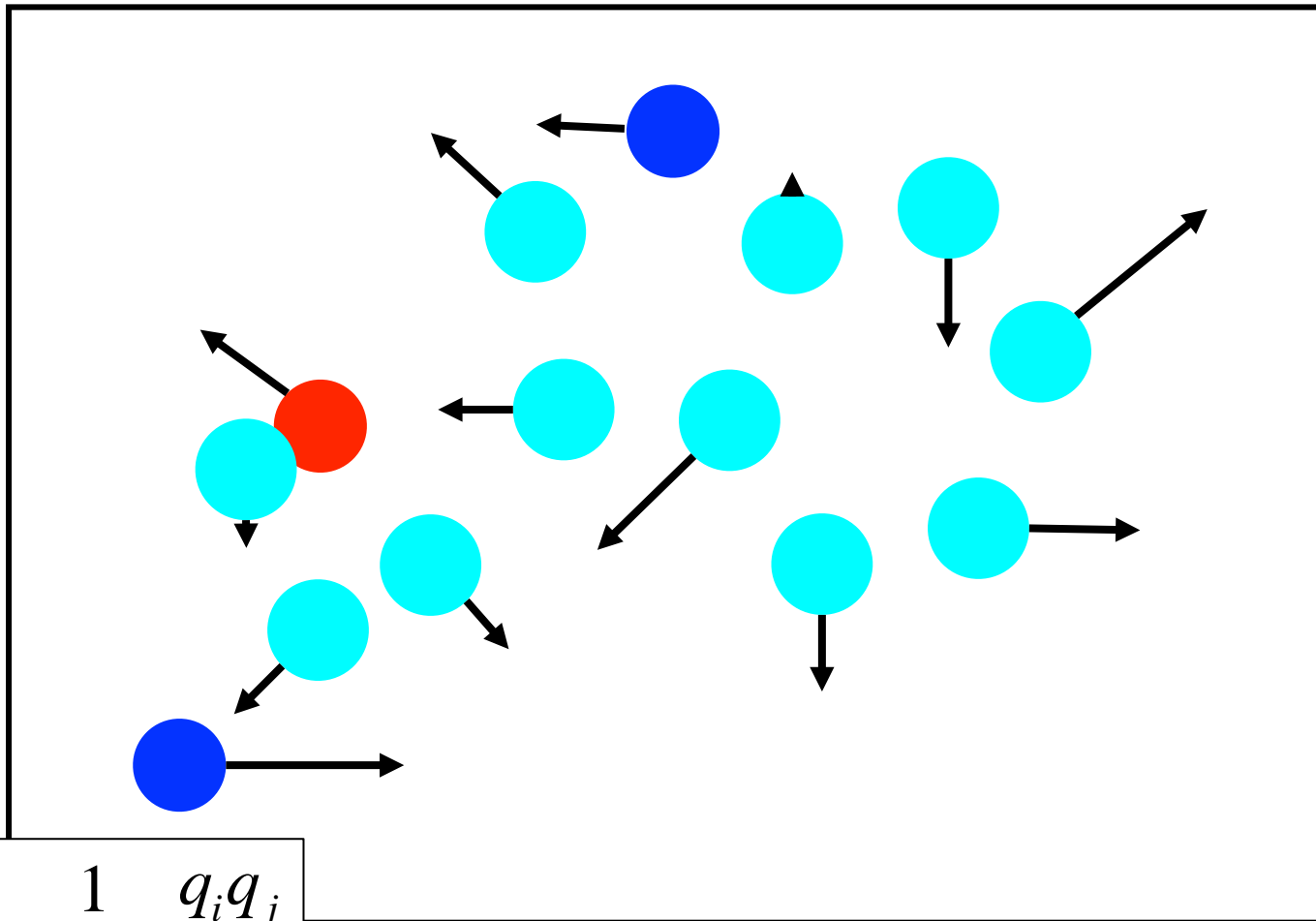
Center for Biophysics and Computational Biology

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

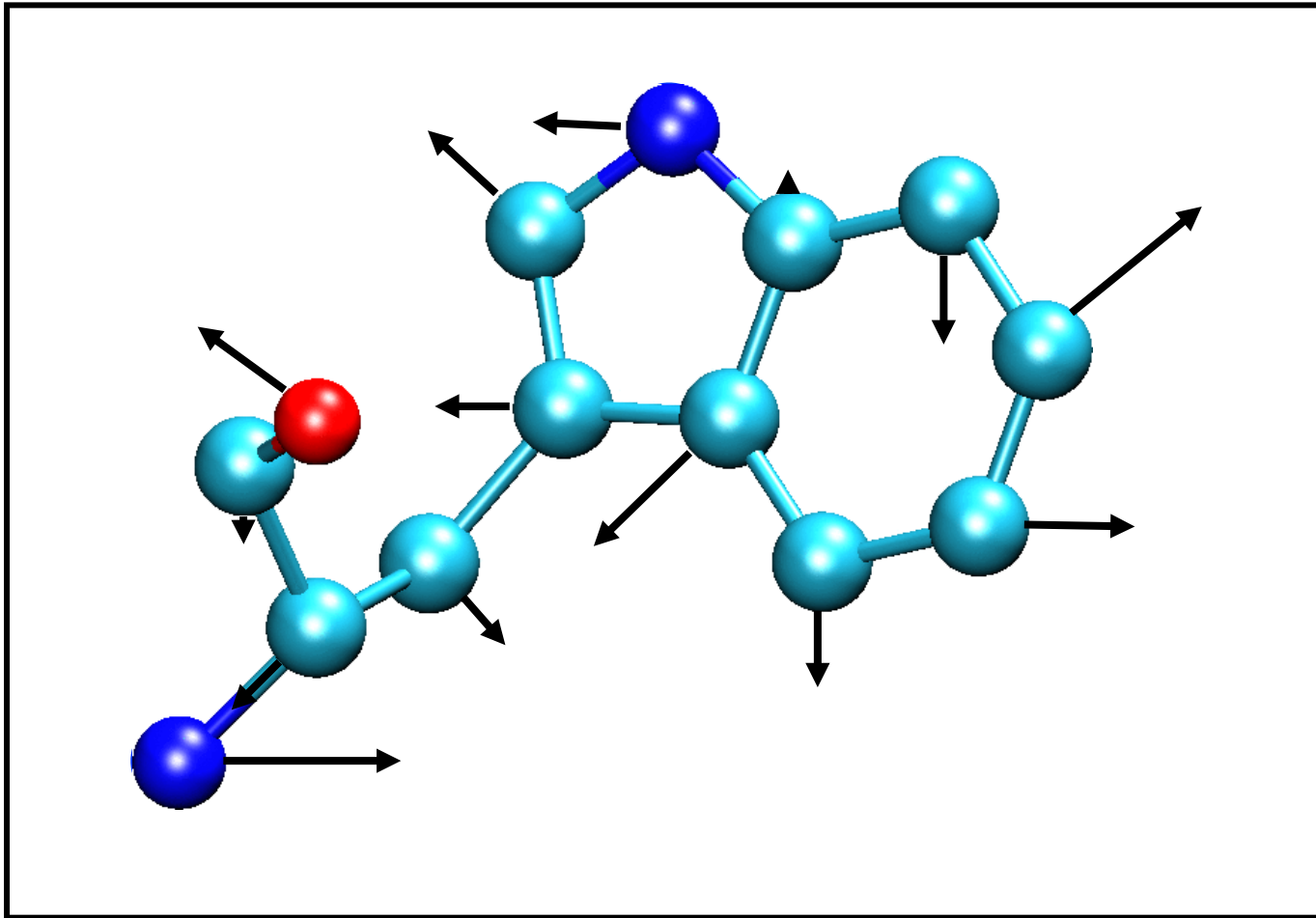


$$U(r) = \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}}$$

Coulomb interaction

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

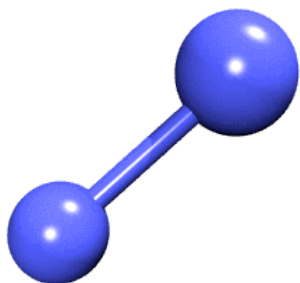
# Classical Molecular Dynamics



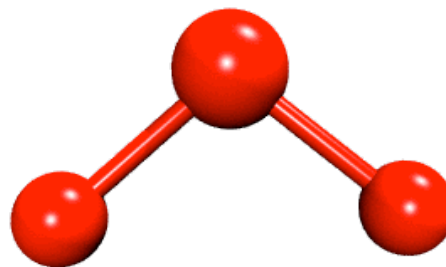
Bond definitions, atom types, atom names, parameters, ....

# Energy Terms Described in

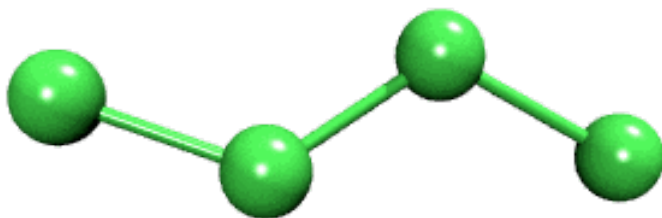
Bond



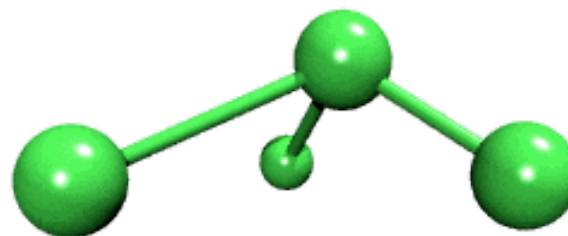
Angle



Dihedral



Improper



# The Potential Energy Function

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

$U_{\text{bond}}$  = oscillations about the equilibrium bond length

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

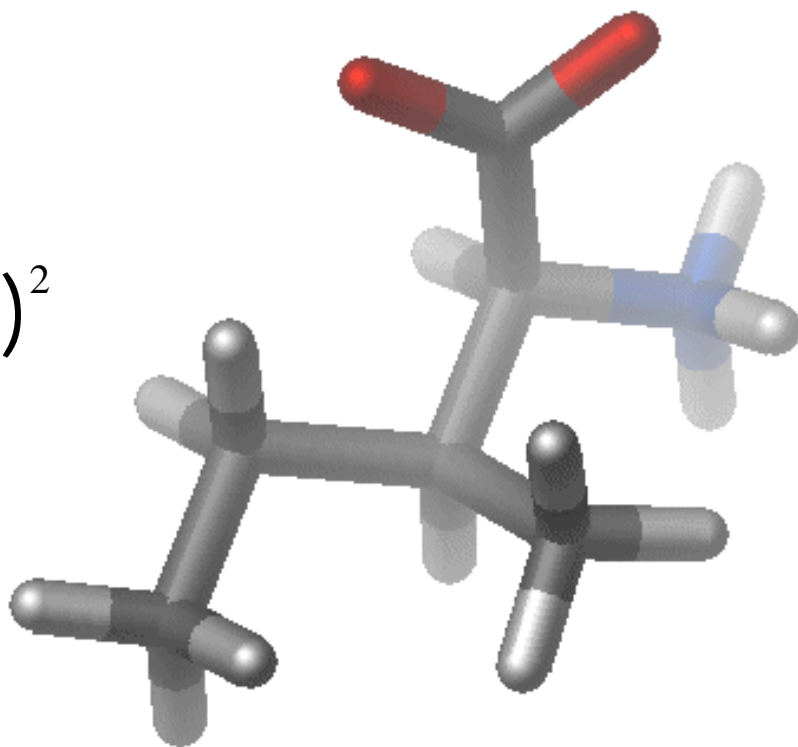
$U_{\text{dihedral}}$  = torsional rotation of 4 atoms about a central bond

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

# Interactions between bonded atoms

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

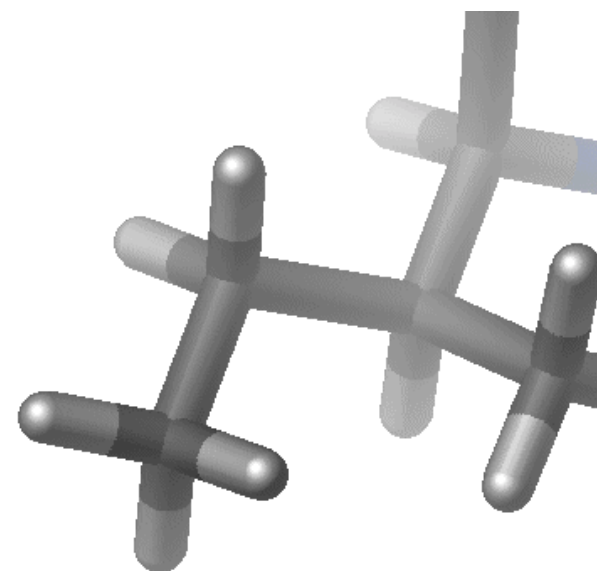
$$V_{bond} = K_b (b - b_o)^2$$



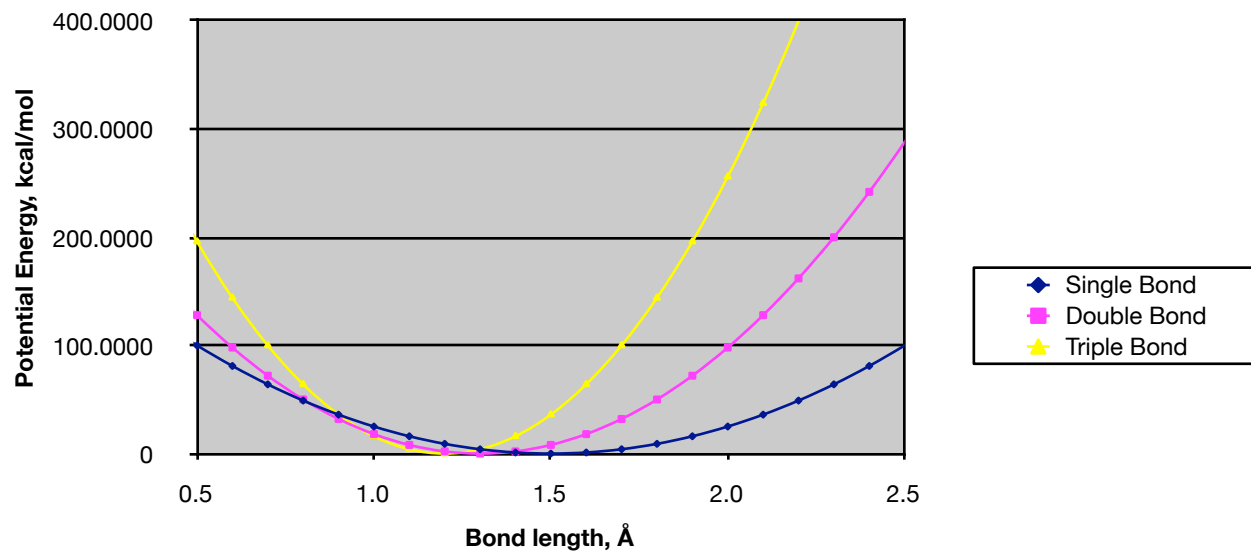
$$V_{dihedral} = K_{\phi} (1 + \cos(n\phi - \delta))$$

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

Chemical type	$K_{bond}$	$b_o$
C-C	100 kcal/mole/Å <sup>2</sup>	1.5 Å
C=C	200 kcal/mole/Å <sup>2</sup>	1.3 Å
C≡C	400 kcal/mole/Å <sup>2</sup>	1.2 Å



**Bond Energy versus Bond length**

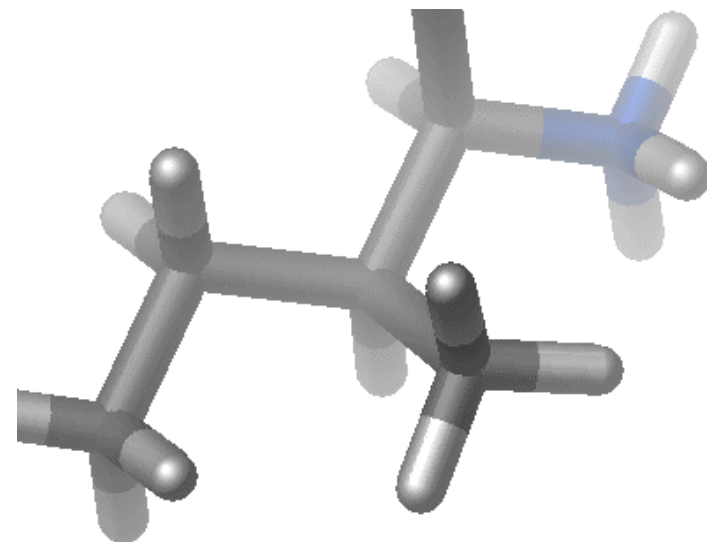


*Bond angles and improper* terms have similar quadratic forms, but with softer spring constants. The force constants can be obtained from vibrational analysis of the molecule (experimentally or theoretically).

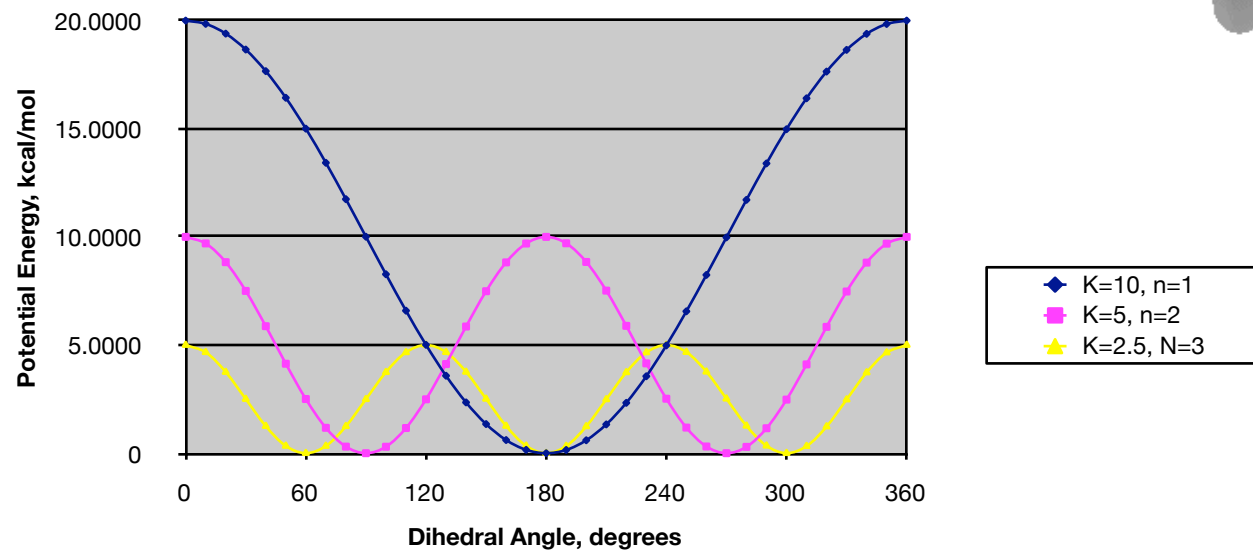


# Dihedral Potential

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



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

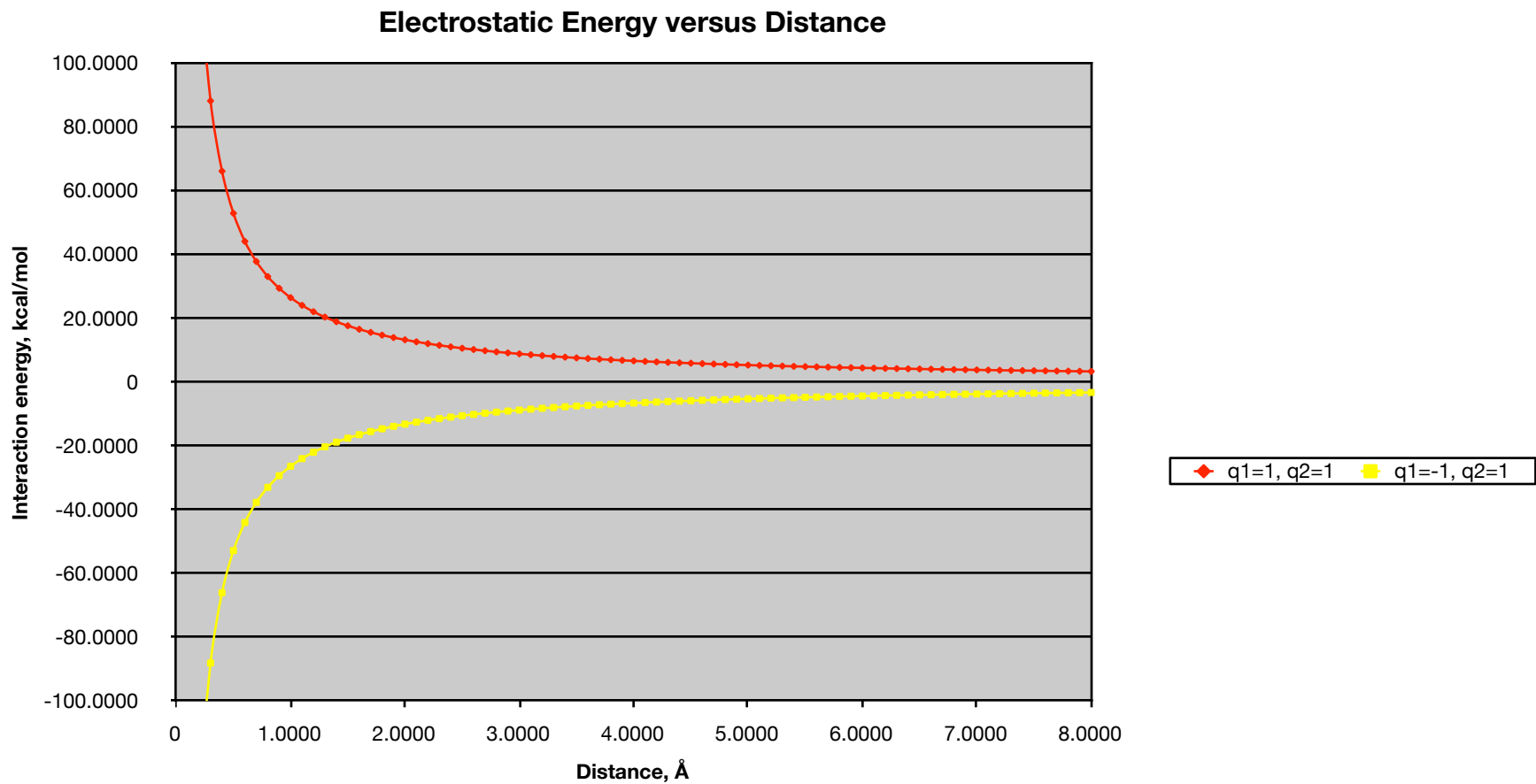
$\epsilon$ : Lennard-Jones (LJ, vdW) well-depth

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

Combining rules (CHARMM, Amber)

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

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



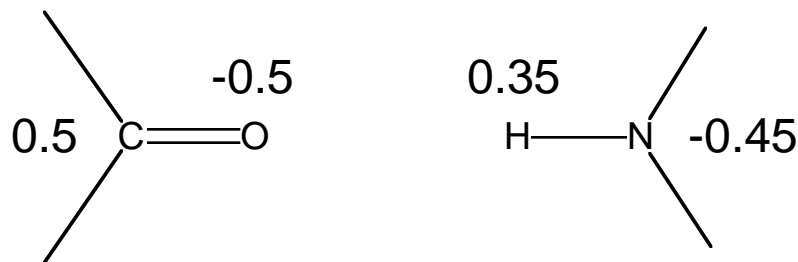
Note that the effect is long range.

# 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  
**geometry**  
 Topology  
 PSF file  
**parameters**  
 Parameter file

# 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

# Structure of a PDB file

index	resname	name	chain	resid	X	Y	Z	segname
ATOM	22	N	ALA B	3	-4.073	-7.587	-2.708	1.00 0.00 BH
ATOM	23	HN	ALA B	3	-3.813	-6.675	-3.125	1.00 0.00 BH
ATOM	24	CA	ALA B	3	-4.615	-7.557	-1.309	1.00 0.00 BH
ATOM	25	HA	ALA B	3	-4.323	-8.453	-0.704	1.00 0.00 BH
ATOM	26	CB	ALA B	3	-4.137	-6.277	-0.676	1.00 0.00 BH
ATOM	27	HB1	ALA B	3	-3.128	-5.950	-0.907	1.00 0.00 BH
ATOM	28	HB2	ALA B	3	-4.724	-5.439	-1.015	1.00 0.00 BH
ATOM	29	HB3	ALA B	3	-4.360	-6.338	0.393	1.00 0.00 BH
ATOM	30	C	ALA B	3	-6.187	-7.538	-1.357	1.00 0.00 BH
ATOM	31	O	ALA B	3	-6.854	-6.553	-1.264	1.00 0.00 BH
ATOM	32	N	ALA B	4	-6.697	-8.715	-1.643	1.00 0.00 BH
ATOM	33	HN	ALA B	4	-6.023	-9.463	-1.751	1.00 0.00 BH
ATOM	34	CA	ALA B	4	-8.105	-9.096	-1.934	1.00 0.00 BH
ATOM	35	HA	ALA B	4	-8.287	-8.878	-3.003	1.00 0.00 BH
ATOM	36	CB	ALA B	4	-8.214	-10.604	-1.704	1.00 0.00 BH
ATOM	37	HB1	ALA B	4	-7.493	-11.205	-2.379	1.00 0.00 BH
ATOM	38	HB2	ALA B	4	-8.016	-10.861	-0.665	1.00 0.00 BH
ATOM	39	HB3	ALA B	4	-9.245	-10.914	-1.986	1.00 0.00 BH
ATOM	40	C	ALA B	4	-9.226	-8.438	-1.091	1.00 0.00 BH
ATOM	41	O	ALA B	4	-10.207	-7.958	-1.667	1.00 0.00 BH

000

10 20 30 40 50 60 70

>>> It is an ascii, fixed-format file <<<

“No connectivity information”

# 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

[www.pharmacy.umaryland.edu/faculty/amackere/force\\_fields.htm](http://www.pharmacy.umaryland.edu/faculty/amackere/force_fields.htm)

# A Brief Overview of The Force Field Toolkit (*ffTK*)

Dr. Christopher G. Mayne  
Tajkhorshid Group  
February 13<sup>th</sup>, 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

---

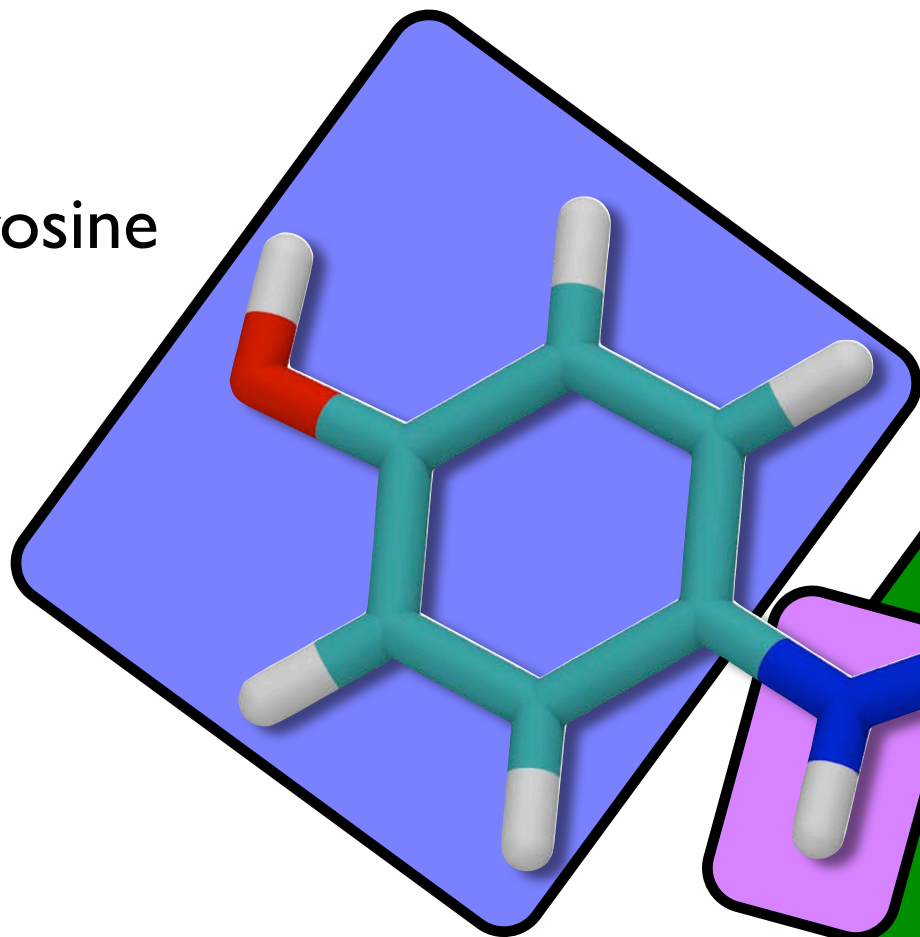
**ParaTool** <http://www.ks.uiuc.edu/Research/vmd/plugins/paratool/>

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

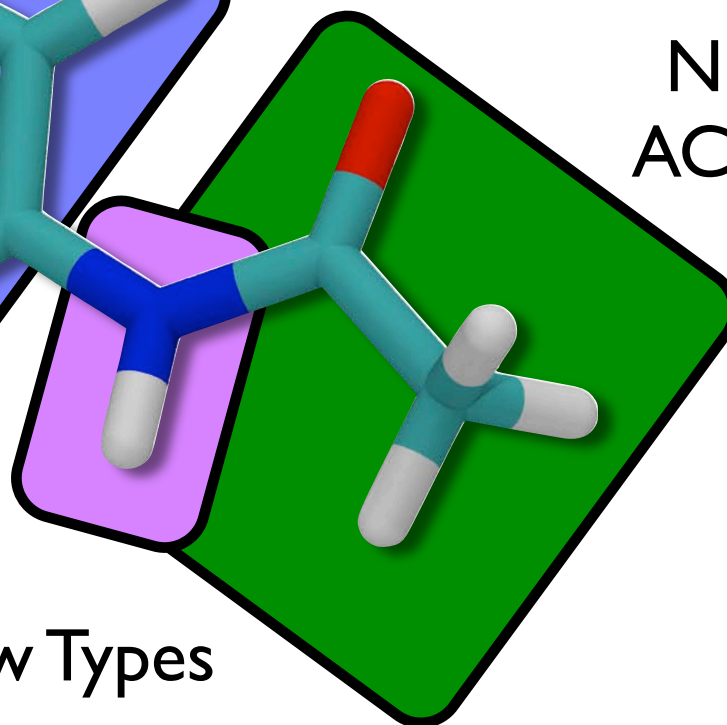
**ffTK** <http://www.ks.uiuc.edu/Research/vmd/plugins/fftk/>

# An Example: Acetaminophen

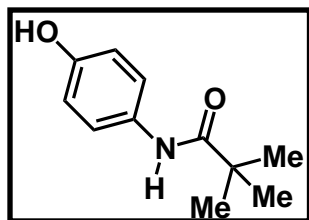
Tyrosine



NMA or  
ACE-patch

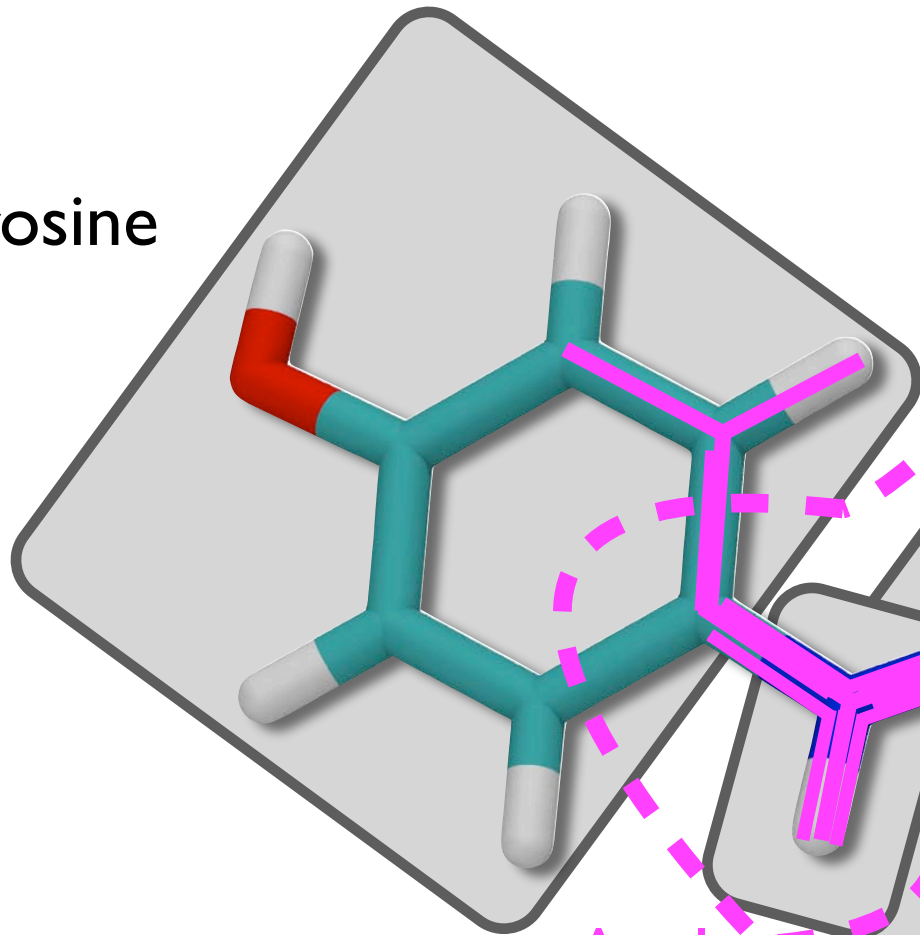


New Types



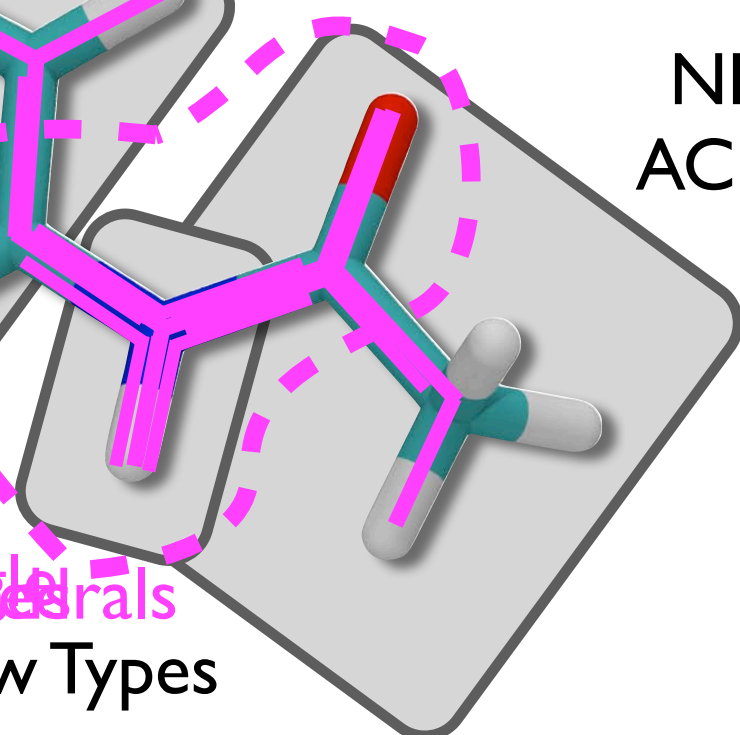
# An Example: Acetaminophen

Tyrosine

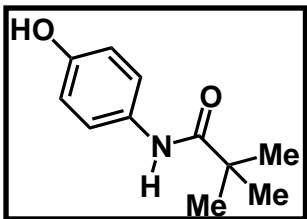


New Charges

NMA or  
ACE-patch



Angles  
Bonds  
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)

~~LIBRARY / TORSIONS~~  
Torsion Optimization

build init PAR

update PDB

update PSF

update PAR

update PAR

Calculation

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 is a tabbed interface with tabs: 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 input fields and buttons. The 'PSF File' field contains the path '/Users/cmayne/Desktop/test11/01-sysprep/pacp.psf' and has a 'Browse' button next to it. The 'PDB File' field contains the path '/Users/cmayne/Desktop/test11/02-geoopt/pacp-opt.pdb' and also has a 'Browse' button. The 'Residue Name' field contains 'PACP' and has a 'Resname From TOP' button. To the right of these fields are buttons for 'Load PSF/PDB' and 'Label Atoms', with a dropdown menu for 'Name' below the 'Label Atoms' button. Below the input fields is a section for 'Parameter Files (both pre-defined and in-progress)' which contains a list of two 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'. Below the parameter files is a section for 'NAMD binary' and 'Output LOG'. The 'NAMD binary' field contains 'namd2' and has a 'Browse' button. The 'Output LOG' field 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 and PDB file fields; 'file dialog buttons' points to the 'Browse' buttons; 'action buttons' points to the 'Load PSF/PDB' and 'Label Atoms' buttons; 'action menus' points to the 'Name' dropdown; 'treeview boxes to hold lists' points to the parameter files list; and 'arrow' points to the left side of the interface.



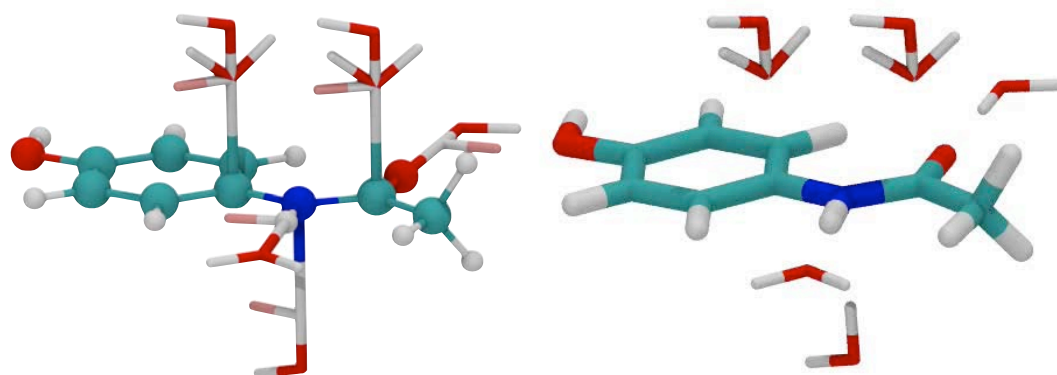
# ffTK 1.0 is Available in VMD 1.9.1

Setup necessary QM calculations

Visualize calculation input/output

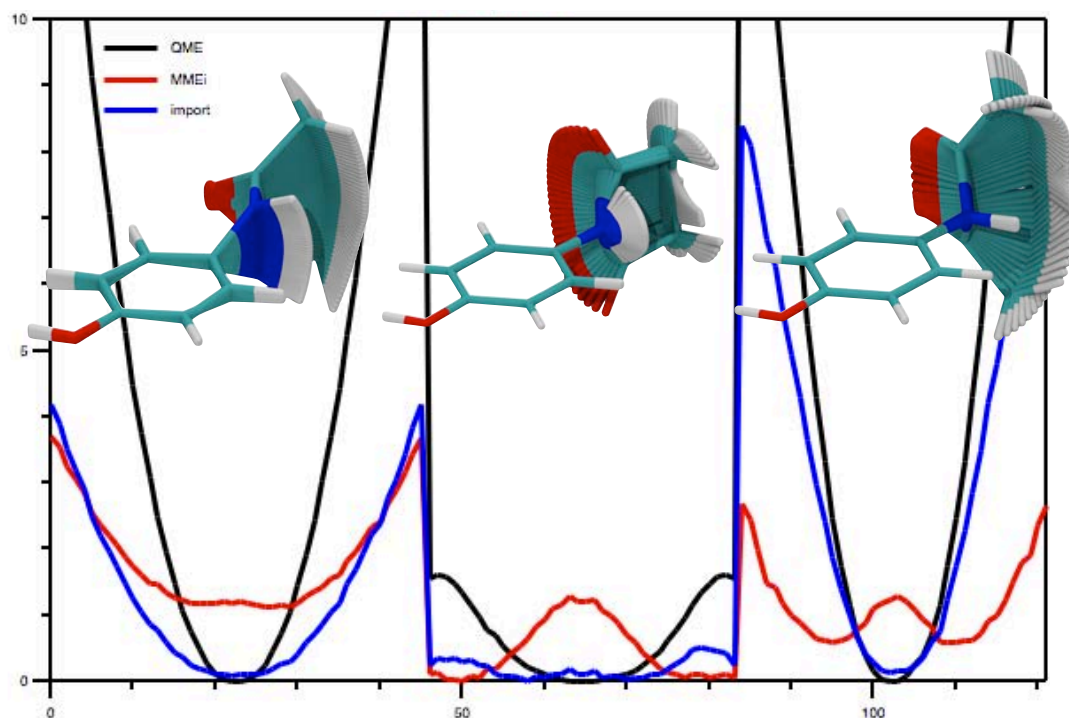
Multidimensional optimization  
of developing parameters

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



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

Tutorial is under development

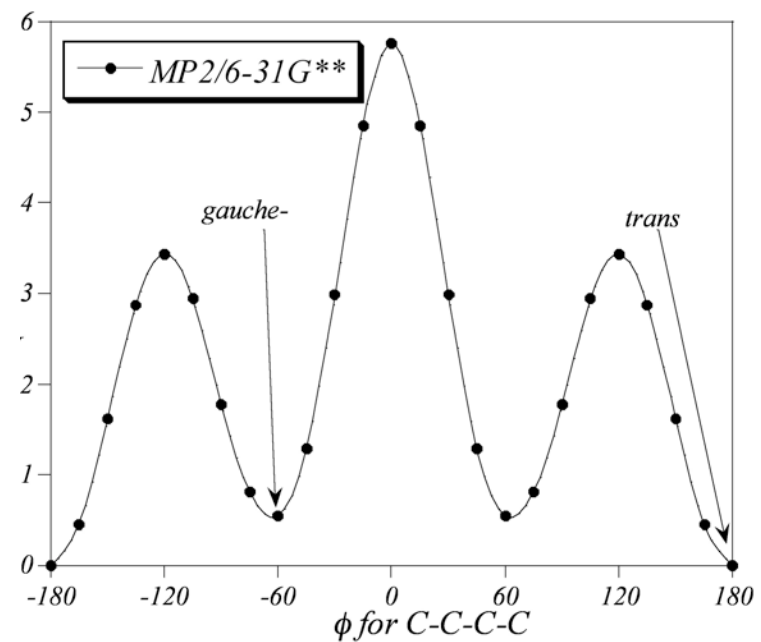
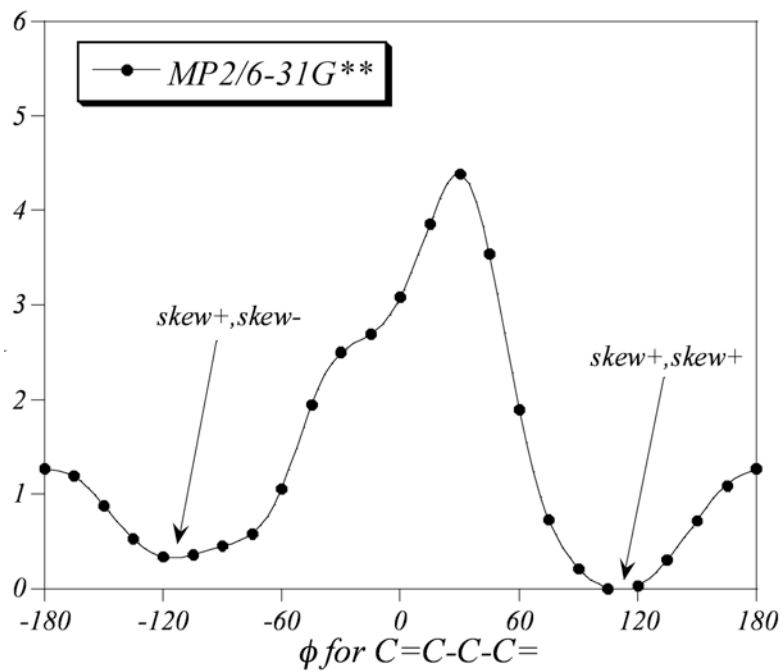
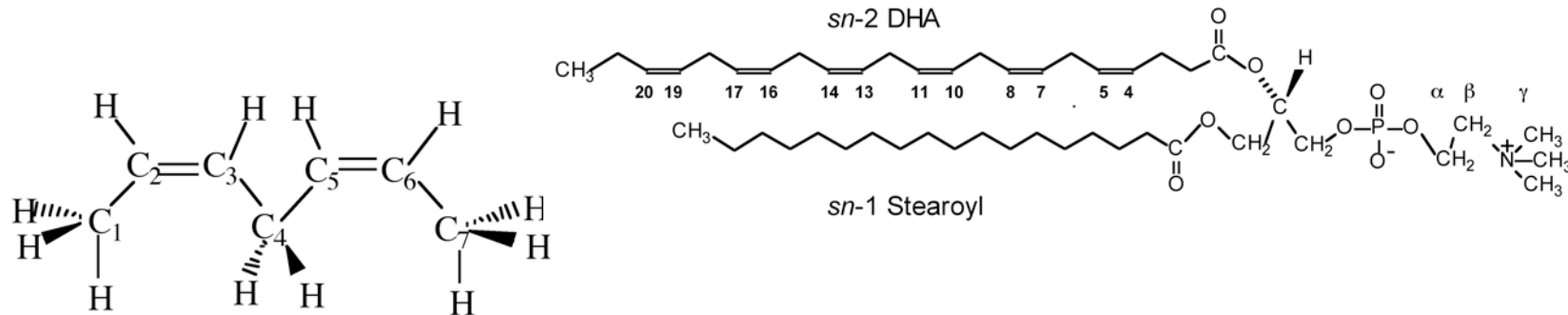


# Partial Charge Assignment

- 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

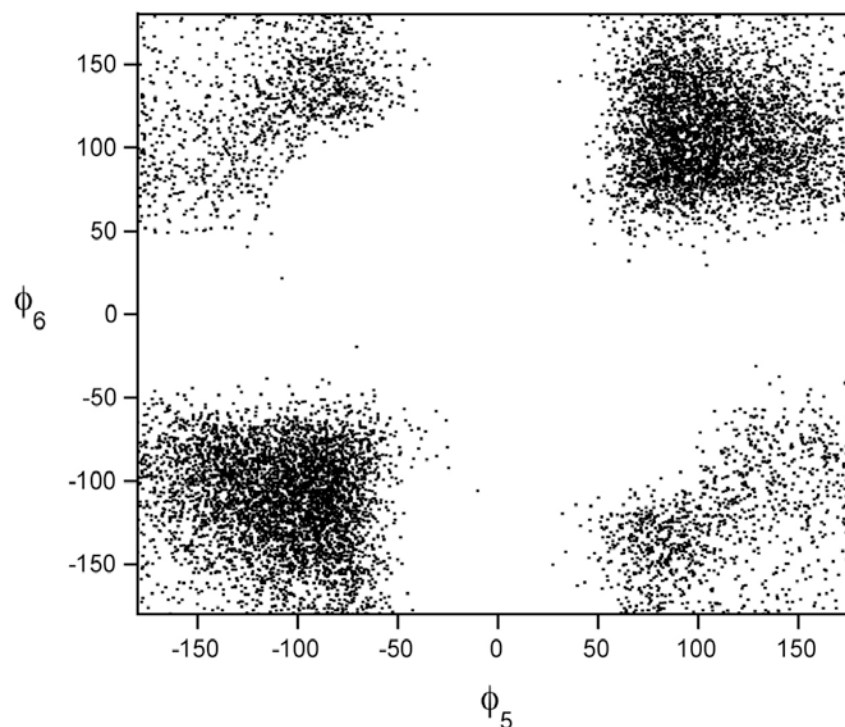
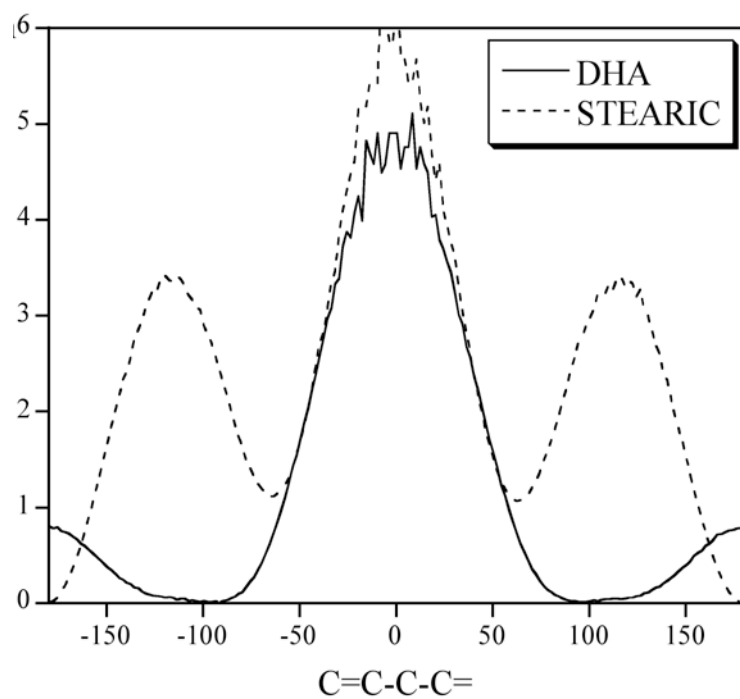
- All C=C bonds are cis, what does rotation about neighboring single bonds look like?



Courtesy of Scott Feller, Wabash College

# DHA conformations from MD

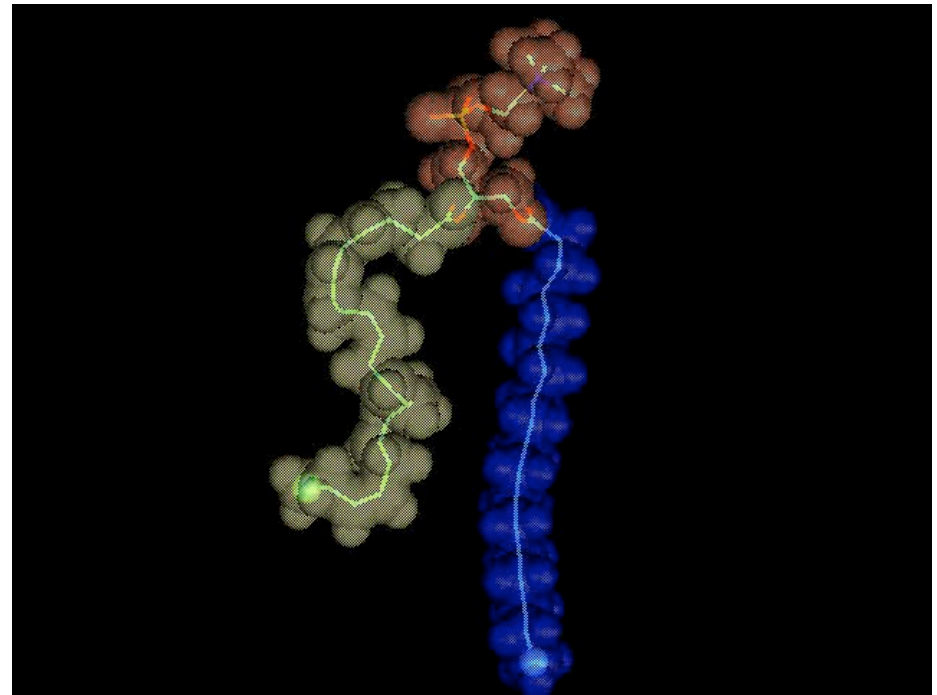
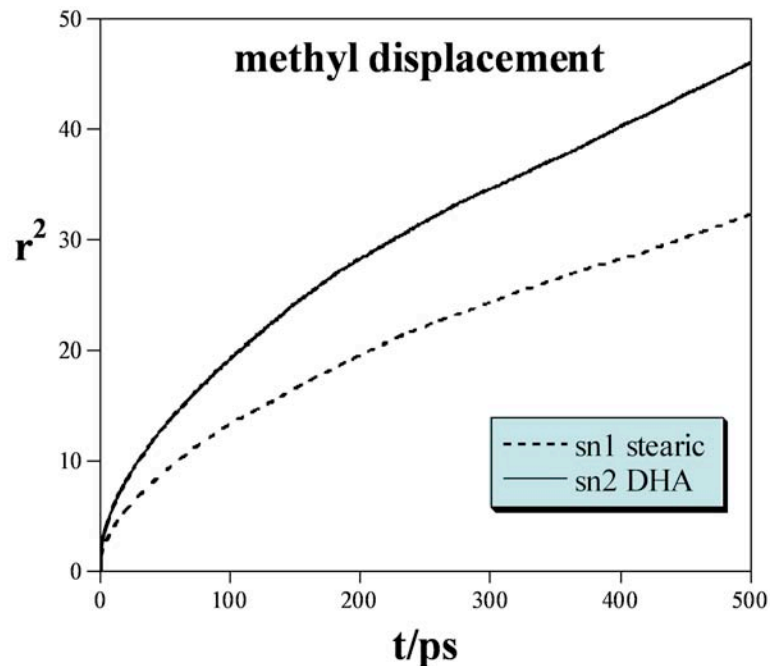
- rotational barriers are extremely small
- many conformers are accessible w/ short lifetimes



Courtesy of Scott Feller, Wabash College

# Dynamics of saturated vs. polyunsaturated lipid chains

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

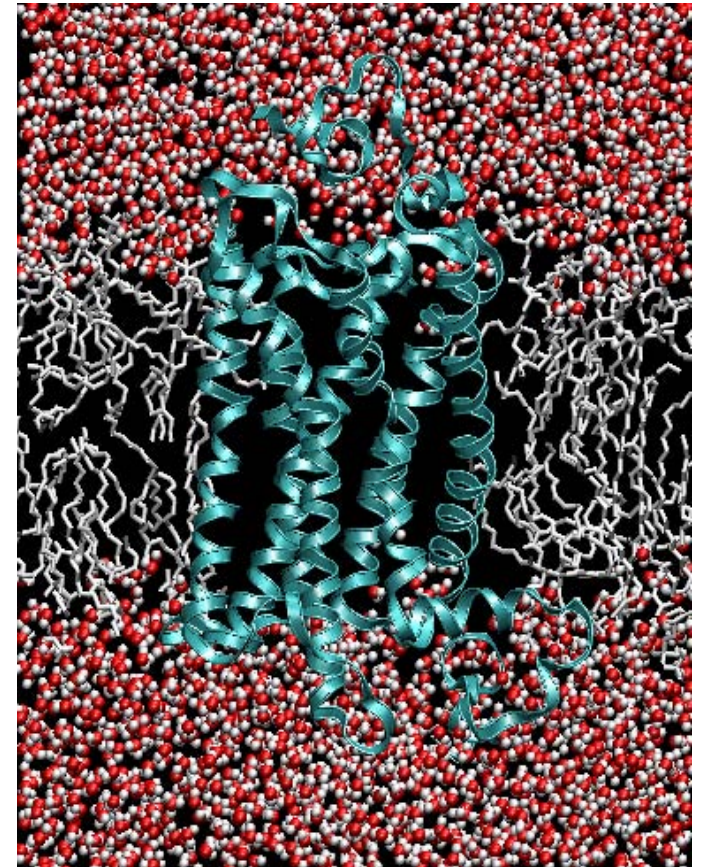
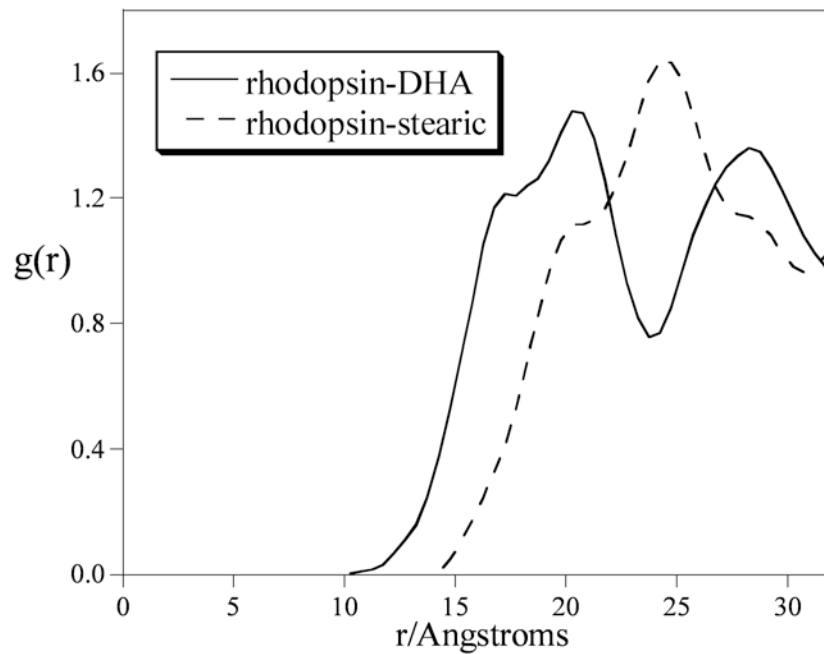


*Movie courtesy of Mauricio Carrillo Tripp*

Courtesy of Scott Feller, Wabash College

# Lipid-protein interactions

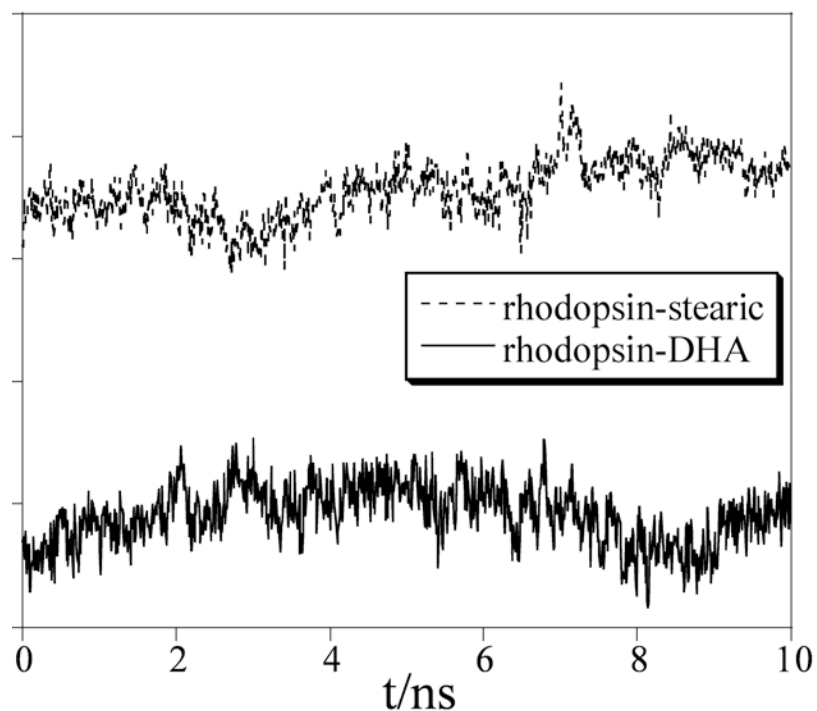
- Radial distribution around protein shows distinct layering of acyl chains



Courtesy of Scott Feller, Wabash College

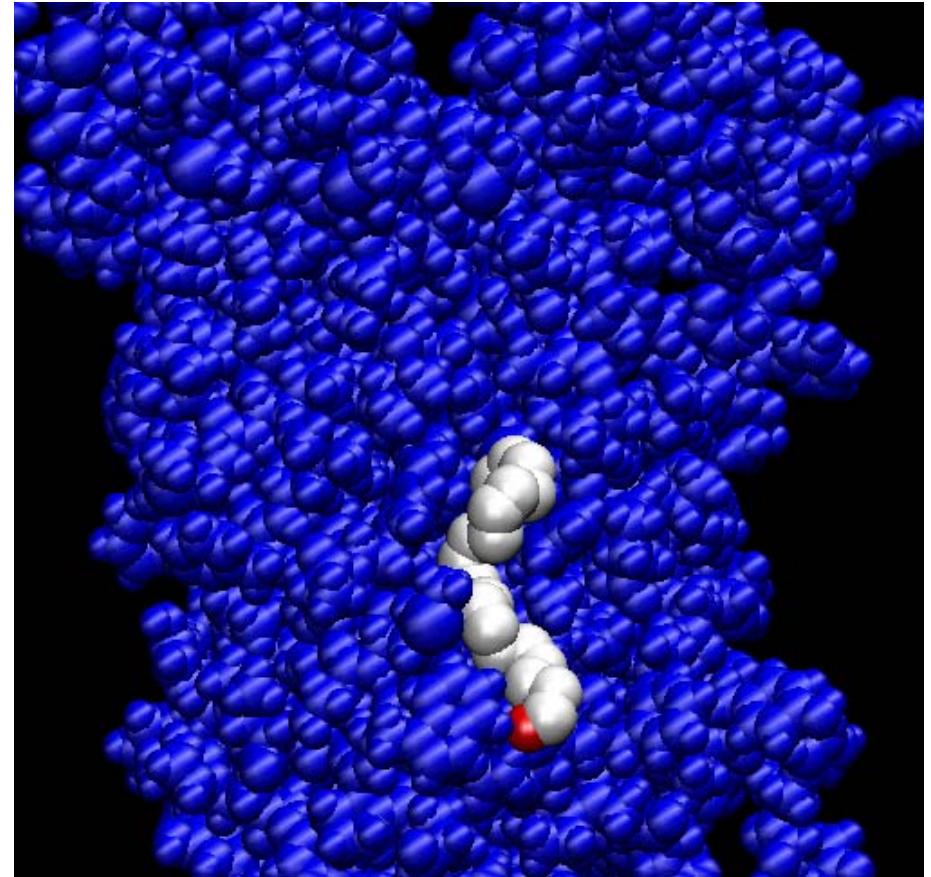
# Lipid-protein interactions

- Decomposition of non-bonded interaction shows rhodopsin is strongly attracted to unsaturated chain
- All hydrophobic residues are stabilized by DHA



<u>resname</u>	$U_{DHA}$	$U_{stearic}$	<u>ratio</u>
PHE	-44.9	-22.6	2.0
ILE	-30.0	-10.1	3.0
VAL	-24.0	-9.6	2.5
LEU	-23.1	-13.0	1.8
MET	-22.8	-9.7	2.4
TYR	-18.6	-10.4	1.8
ALA	-11.4	-3.0	3.8
TRP	-10.3	-2.4	4.2

# Origin of protein:DHA attraction



- Flexibility of the DHA chain allows solvation of the rough protein surface to occur with little intra-molecular energy cost



# Major Recent Developments

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