## 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



# Classical Molecular Dynamics



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

# Energy Terms Described in



## The Potential Energy Function



 $U_{bond}$  = 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

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



 $V_{bond} = K_b (b - b_o)^2$ Chemical type  $K_{\text{bond}}$ bo C-C 100 kcal/mole/Å 1.5 Å  $200 \text{ kcal/mole/Å}^{-2}$ 1.3 Å C=C 1.2 Å  $400 \text{ kcal/mole/Å}^{-2}$ C=C **Bond Energy versus Bond length** 400.0000 300.0000 200.0000 Single Bond + **Double Bond** 100.0000 **Triple Bond** 0

2.0

1.5

Bond length, Å

1.0

Potential Energy, kcal/mol

0.5

*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).

2.5



 $\delta = 0^{\circ}$ 

### **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<sub>i</sub>: 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)

$$R_{\min i,j} = R_{\min i} + R_{\min j}$$
$$\varepsilon_{i,j} = SQRT(\varepsilon_i * \varepsilon_j)$$



**Electrostatic Energy versus Distance** 

Note that the effect is long range.

From MacKerell

## Charge Fitting Strategy

CHARMM- Mulliken\*

AMBER(ESP/RESP)

Partial atomic charges



\*Modifications based on interactions with TIP3 water

## **CHARMM** Potential Function



## 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)
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

# 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
- sn2 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

## Major Recent Developments

 New set of lipid force field parameters for CHARMM (CHARMM32<sup>+</sup>)

-Pastor, B. Brooks, MacKerell

- Polarizable force field
  - -Roux, MacKerell