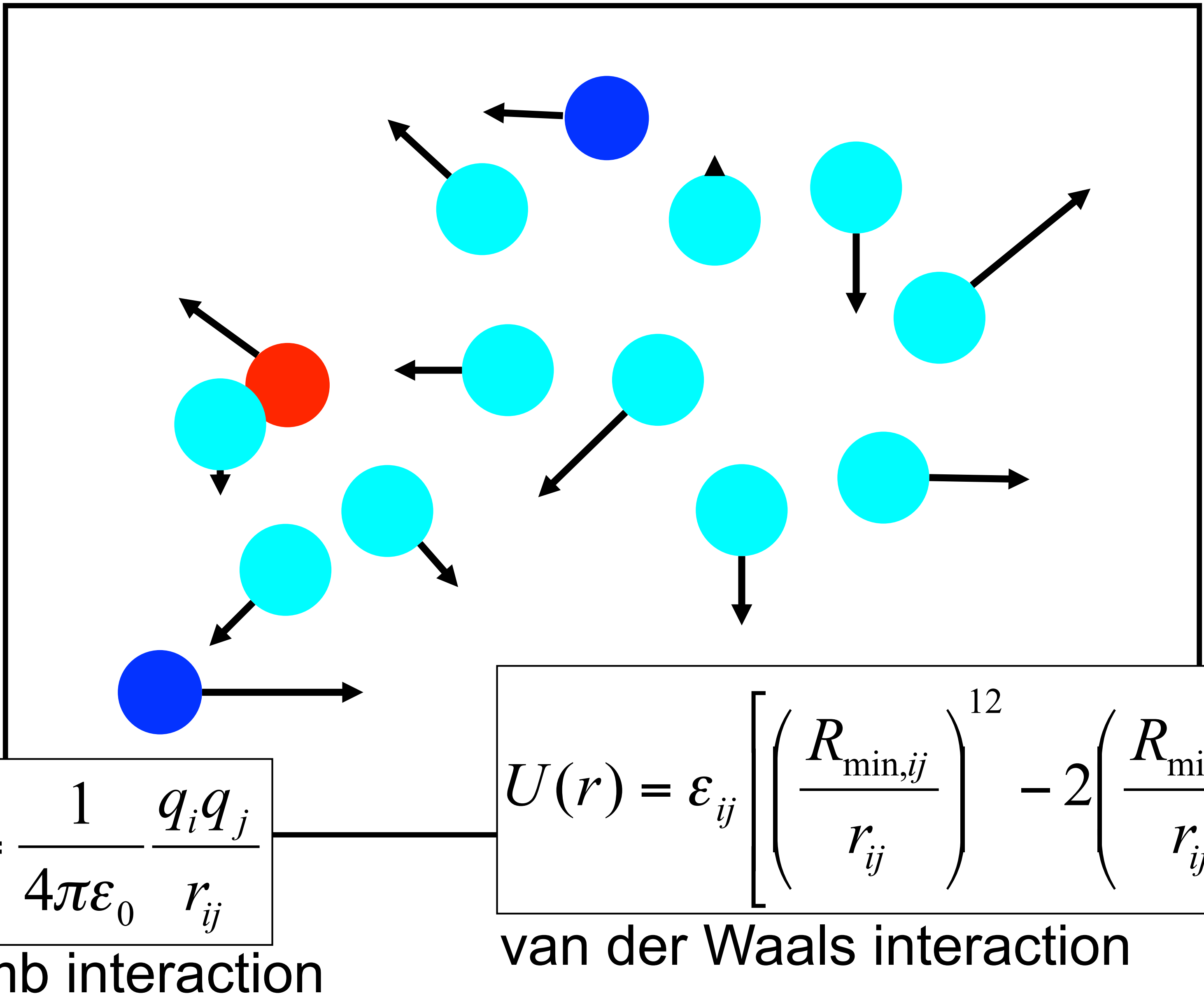
Force Fields for MD simulations

- Topology/parameter files
- Where do the numbers an MD code uses come from?
- How to make topology files for ligands, cofactors, special amino acids, …
- How to obtain/develop missing parameters.

Classical Molecular Dynamics



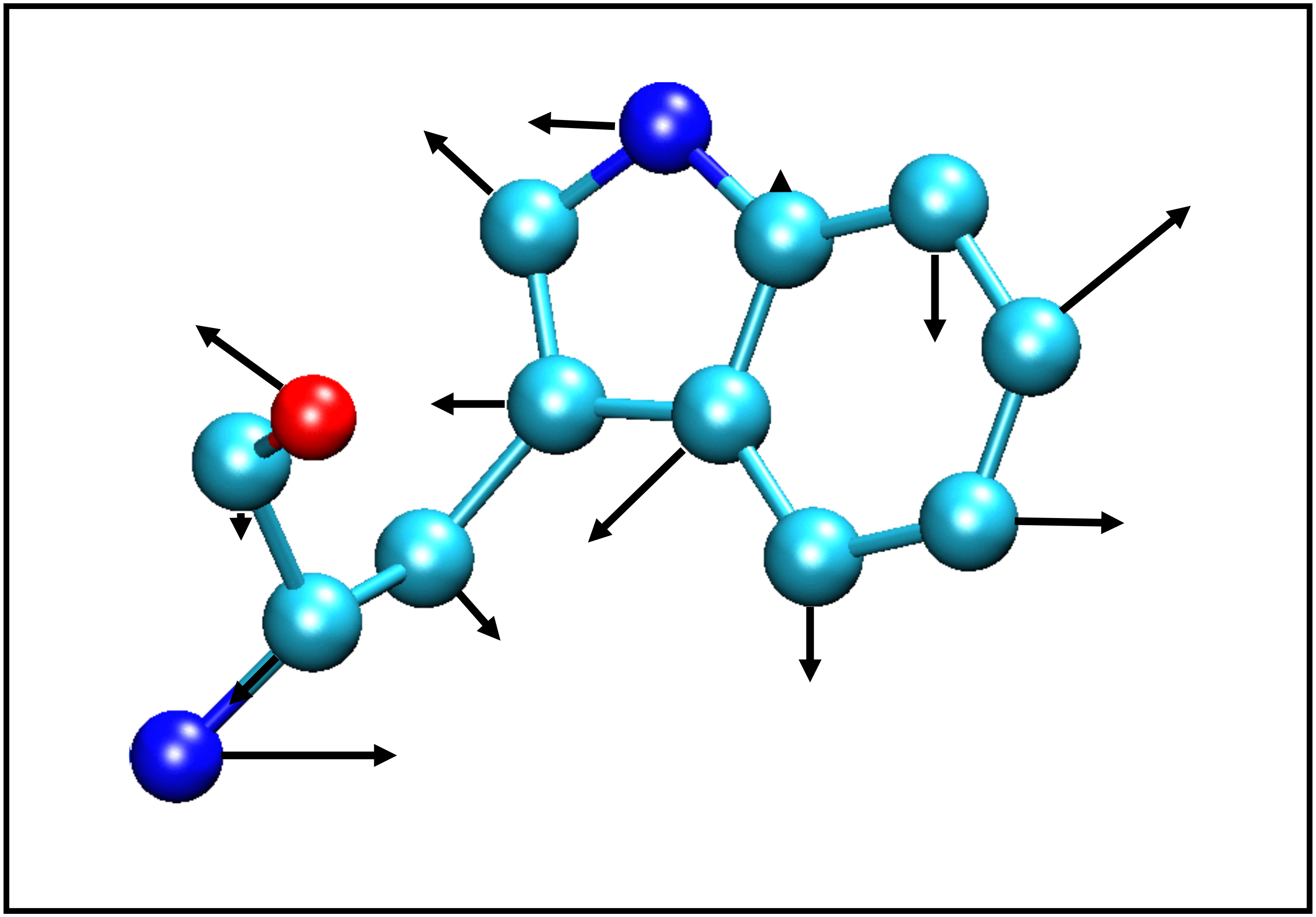
$$U(r) = \frac{1}{4\pi\varepsilon_0}\frac{q_i q_j}{r_{ij}}$$

Coulomb interaction

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

van der Waals interaction

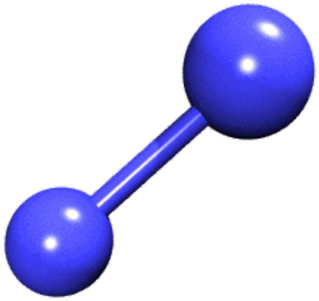
Classical Molecular Dynamics



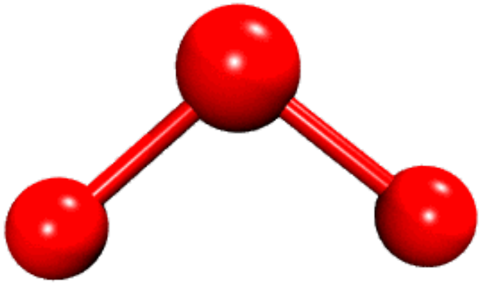
Bond definitions, atom types, atom names, parameters, ….

Energy Terms Described in the CHARMm Force Field

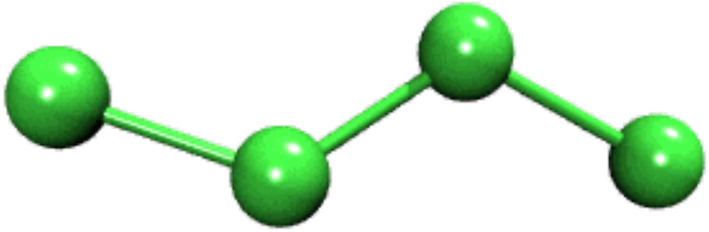
Bond



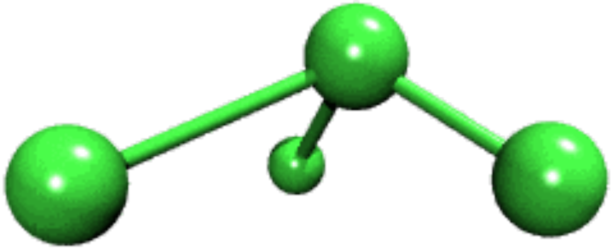
Angle



Dihedral



Improper



The Potential Energy Function

$$U(\vec{R}) = \underbrace{\sum_{bonds} k_i^{bond}(r_i - r_0)^2}_{U_{bond}} + \underbrace{\sum_{angles} k_i^{angle}(\theta_i - \theta_0)^2}_{U_{angle}} +$$

$$\underbrace{\sum_{dihedrals} k_i^{dihe}[1 + \cos(n_i\phi_i + \delta_i)]}_{U_{dihedral}} +$$

$$\underbrace{\sum_i \sum_{j \neq i} 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \sum_i \sum_{j \neq i} \frac{q_i q_j}{\epsilon r_{ij}}}_{U_{nonbond}}$$

U_{bond} = oscillations about the equilibrium bond length

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

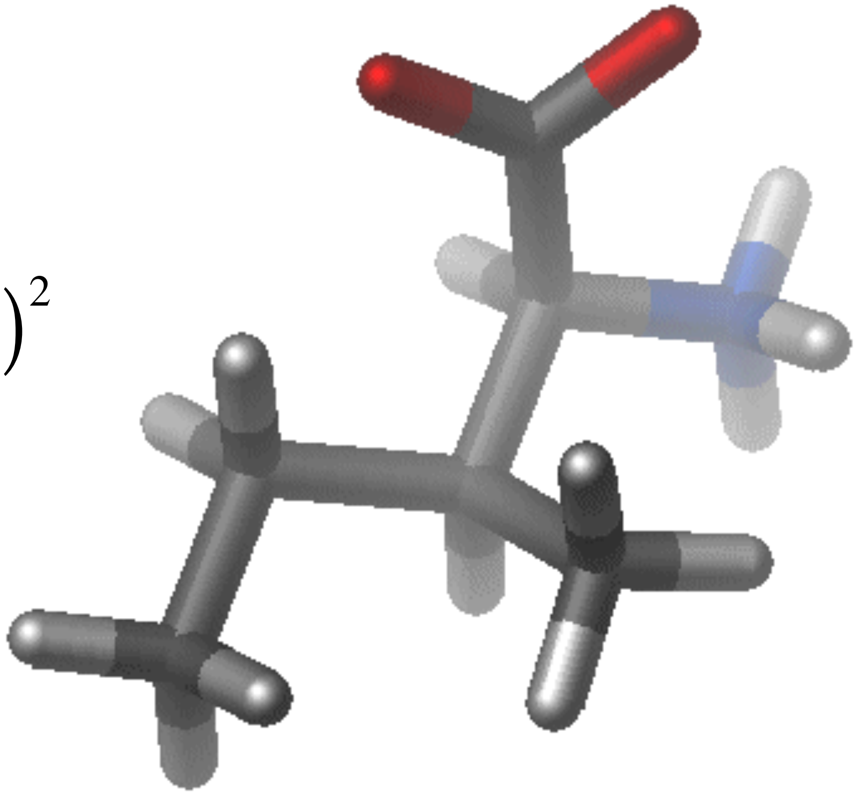
$U_{dihedral}$ = torsional rotation of 4 atoms about a central bond

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

Interactions between bonded atoms

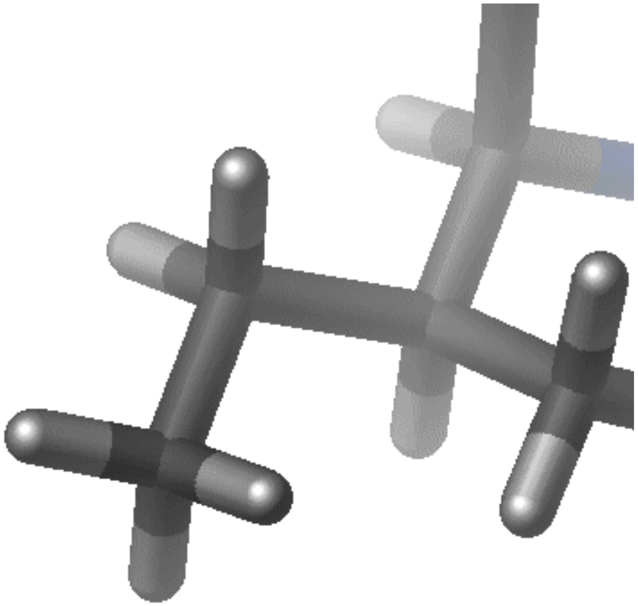
$$V_{angle} = K_{\theta}\left(\theta - \theta_o\right)^2$$

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

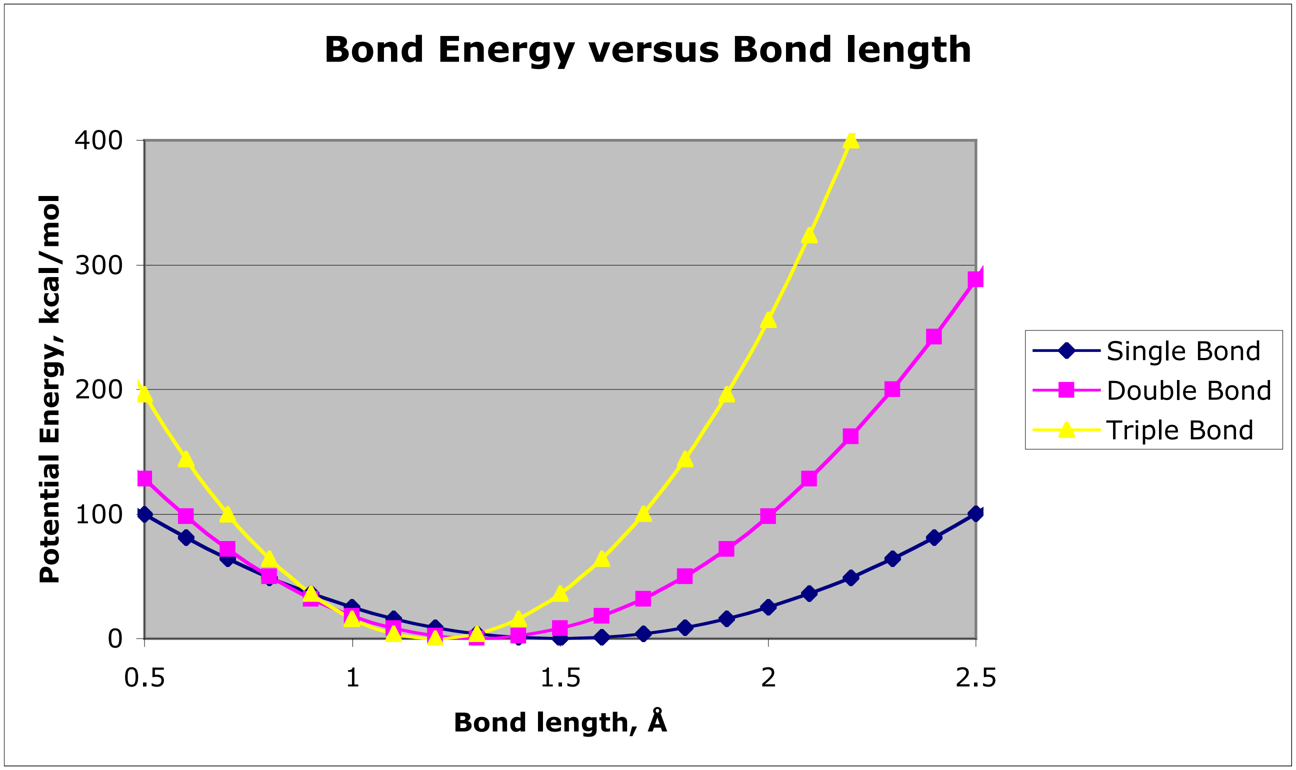


$$V_{dihedral} = K_{\phi}(1 + \cos(n\phi - \delta))$$

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

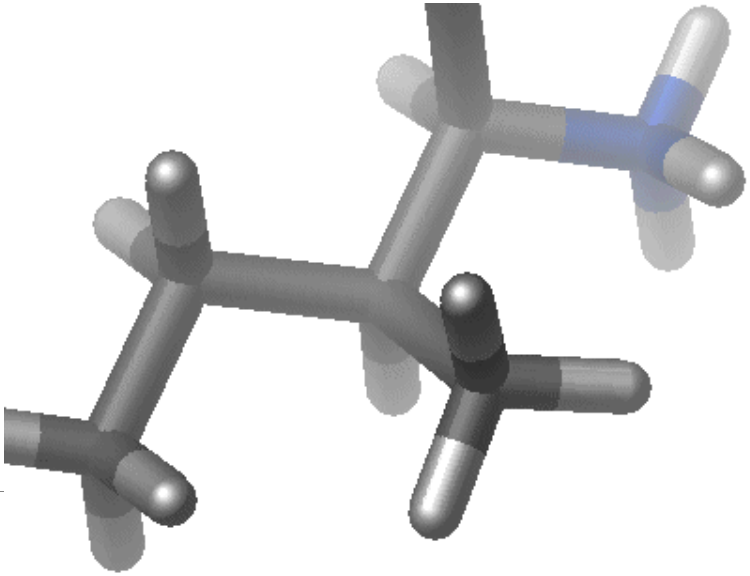


Chemical type	K_{bond}	b_o
C-C	100 kcal/mole/Å ^{2}	1.5 Å
C=C	200 kcal/mole/Å ^{2}	1.3 Å
C=C	400 kcal/mole/Å ^{2}	1.2 Å

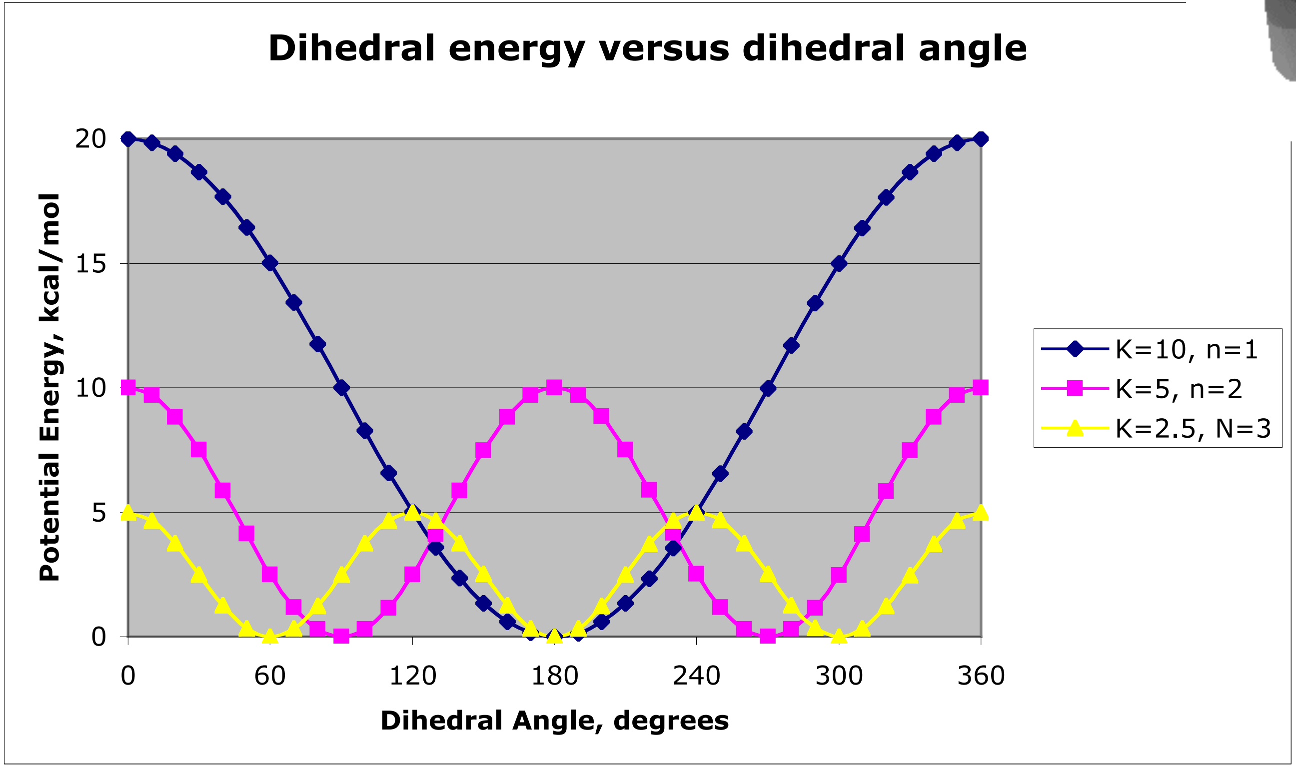


Bond angles and *improper* terms have similar quadratic forms, but with softer spring constants. The force constants can be obtained from vibrational analysis of the molecule (experimentally or theoretically).

Dihedral Potential



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

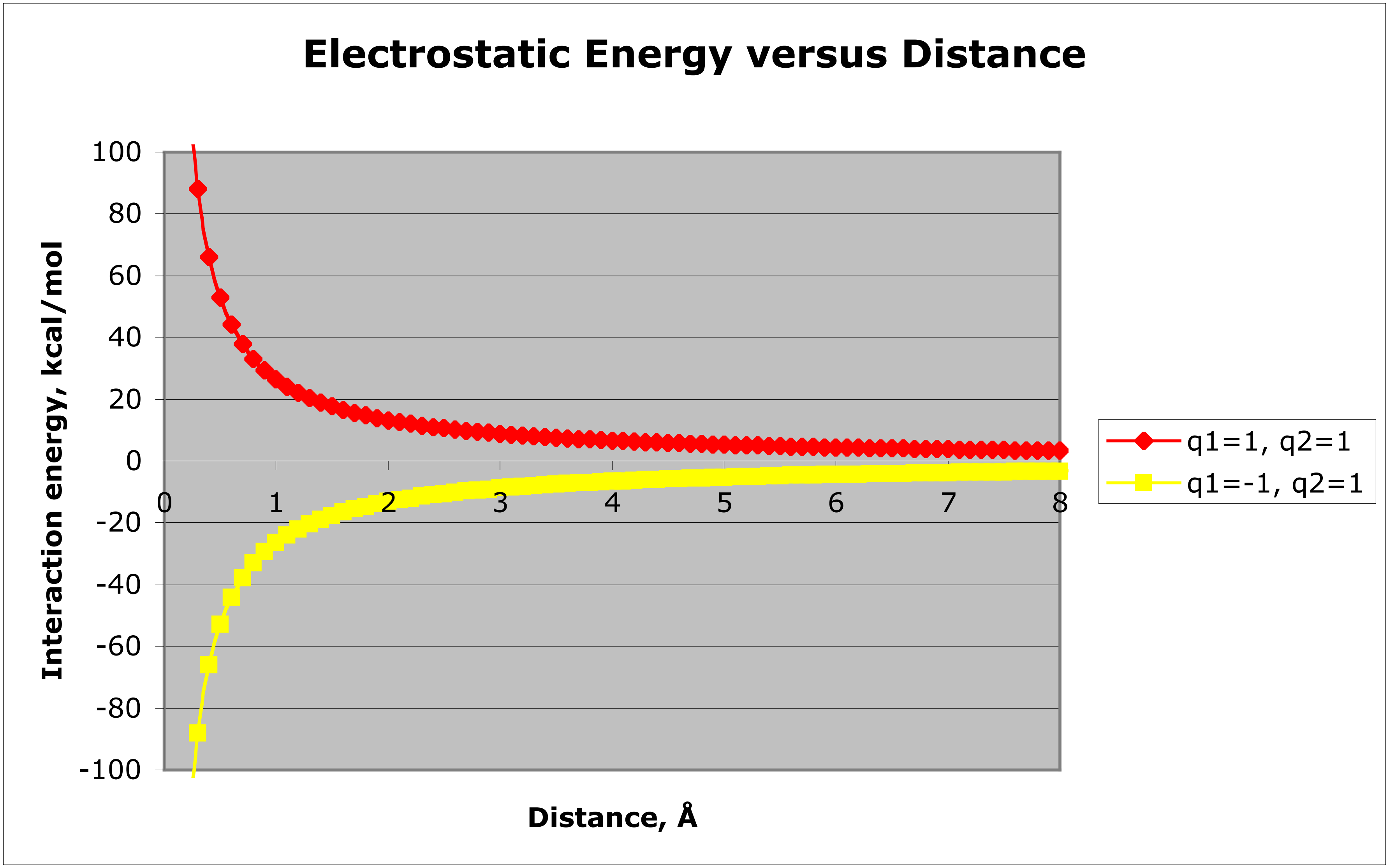


$$\delta = 0^{\circ}$$

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



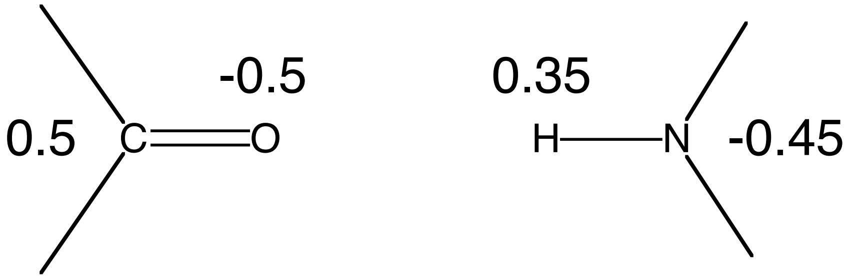
Note that the effect is long range.

From MacKerell

Charge Fitting Strategy

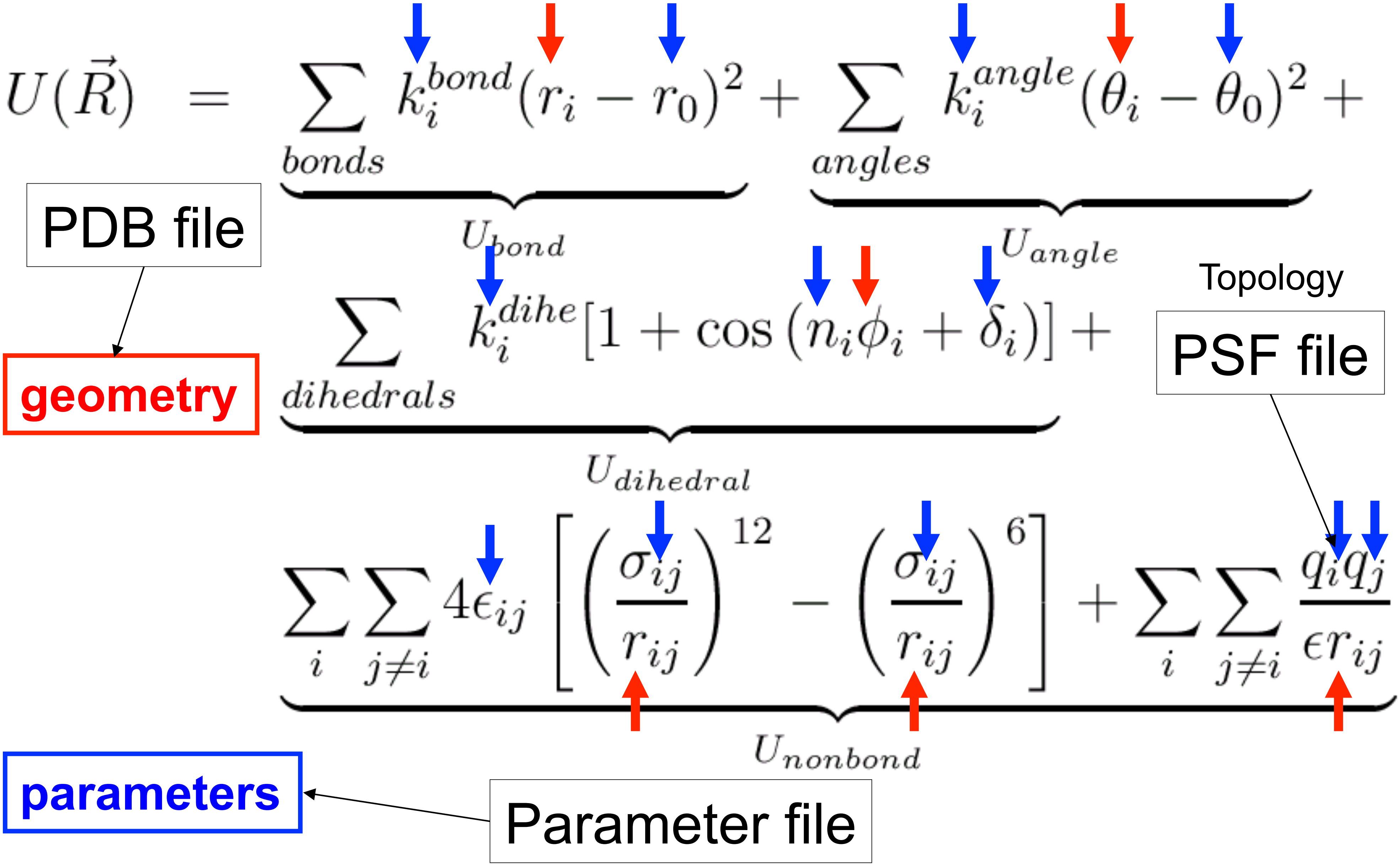
CHARMM- Mulliken* AMBER(ESP/RESP)

Partial atomic charges



*Modifications based on interactions with TIP3 water

CHARMM Potential Function



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

PDB file

geometry

Topology

PSF file

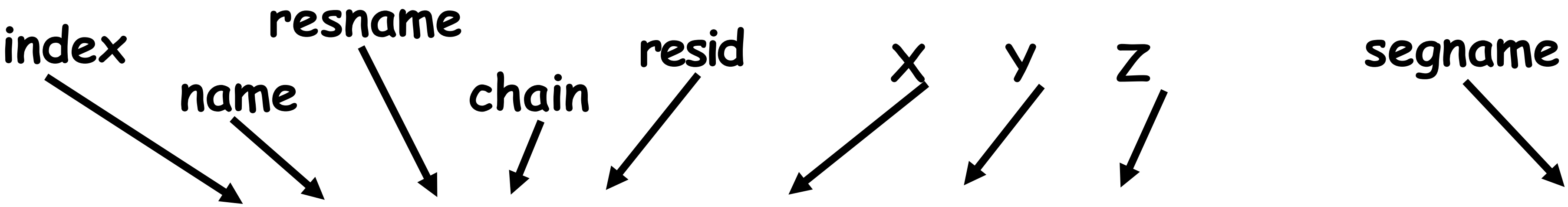
parameters

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 **name** **resname** **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
00000000000000000000000000000000000000000000000000000000000000000000000000
        10        20        30         40         50         60         70
```

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

“No connectivity information”

Checking 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 (i.e. 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