Semi Empirical Force Fields and Their Limitations

Ioan Kosztin

Beckman Institute
University of Illinois at Urbana-Champaign

Potential Energy Surface (PES)

Schrödinger equation: $H_T \Psi(\boldsymbol{r},\boldsymbol{R}) = E_T \Psi(\boldsymbol{r},\boldsymbol{R})$

$H_T = K_n + K_e + V(\boldsymbol{r},\boldsymbol{R})$

Born-Oppenheimer approx. $\Psi(\boldsymbol{r},\boldsymbol{R}) = \psi(\boldsymbol{r} \mid \boldsymbol{R})\,\Theta(\boldsymbol{R})$

Equation of motion for electrons: $(K_e + V)\,\psi(\boldsymbol{r},\boldsymbol{R}) = \boxed{E_e(\boldsymbol{R})}\,\psi(\boldsymbol{r},\boldsymbol{R})$

PES

Equation of motion for nuclei:
(motion on the PES)

quantum

$[K_n + E_e(\boldsymbol{R})]\Theta(\boldsymbol{R}) = E_T\,\Theta(\boldsymbol{R})$

classical (MD)

$\boxed{m\,d^2\boldsymbol{R}/dt^2 = -\nabla E_e(\boldsymbol{R})}$

Classical Force Fields

Consist of:
1. Analytical form of the interatomic potential energy $U=E_e(\boldsymbol{R})$ as a function of the atomic coordinates of the molecule
2. Parameters which enter U

• The force field is fit to experimental data from a small set of molecules

•The molecule is considered to be a collection of atoms held together by simple elastic (harmonic) forces.

• The forces are given by the gradients of the potential energy with respect to the internal coordinates of the molecule.

Components of a Force Field

Any force field contains the necessary building blocks for the calculation of energy and force:

- a list of atom types
- a list of atomic charges
- rules for atom-types
- functional forms of the components of the energy expression
- parameters for the function terms
- rules for generating parameters that have not been explicitly defined (for some force fields)
- a defined way of assigning functional forms and parameters (for some force fields)

Force Field Energy Expression

General form:

$E = E_{bond} + E_{angle} + E_{torsion} + E_{nonbond} + E_{other}$

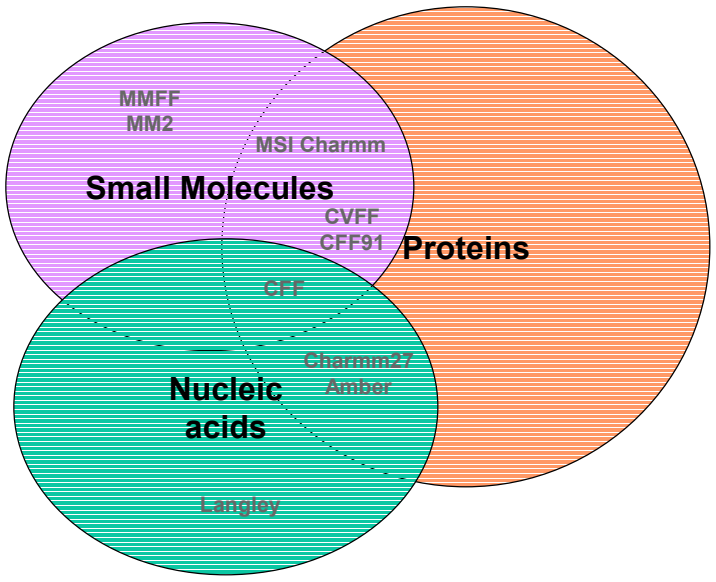
Force fields differ:

- in parameters and cross terms
- methods of parameterization

Types of Force Fields

- **Classical (first-generation) force fields**
  - **CHARMm**
  - AMBER
  - CVFF (consistent valence force field)
  - …
- Second-generation force fields
- Rule-based force fields
- Special-purpose force fields

Force Field Variety



The CHARMm Force Field

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

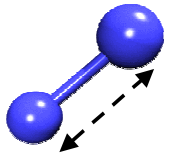
$$\underbrace{\sum_{dihedrals} k_i^{dihe}[1 + \cos(n_i\phi_i + \delta_i)]}_{U_{dihedral}} +$$

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

The CHARMm Energy Function: Bonded Terms

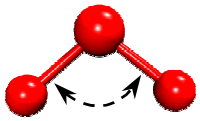
E_{bond} – bond-stretching potential energy

$$E_{bond} = \sum k_b (r - r_0)^2$$



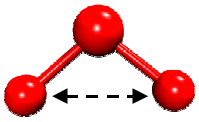
E_θ – angle-bending potential energy

$$E_\theta = \sum k_\theta (\theta - \theta_0)^2$$



E_{UB} – Urey-Bradley term

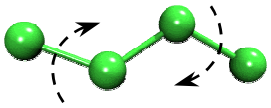
$$E_{UB} = \sum K_{UB} (S - S_0)^2$$



The CHARMm Energy Function: Bonded Terms

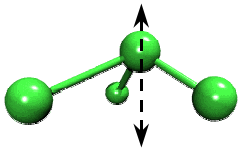
E_{dih} – dihedral angle potential energy

$$E_\phi = \sum K_\phi [1 + \cos(n\phi - \delta)]$$



E_{impr} – improper dihedral angle potential energy

$$E_\omega = \sum k_\omega (\omega - \omega_0)^2$$



The CHARMm Energy Function: Nonbonded Terms

E_{elec} - electrostatic energy

$$E_{elec} = \sum_{i>j} \frac{q_i q_j}{\varepsilon r_{ij}}$$

E_{vdw} - Van der Waals energy

$$E_{vdw} = \sum \varepsilon \left[\left(\frac{R_{\min_{ij}}}{r_{ij}} \right)^{12} - \left(\frac{R_{\min_{ij}}}{r_{ij}} \right)^{6} \right]$$

External Files

RTF – Residue Topology Files

Store information about atom mass, atom type, partial charges, connectivity, internal coordinates, residue definitions

PRM – parameter files

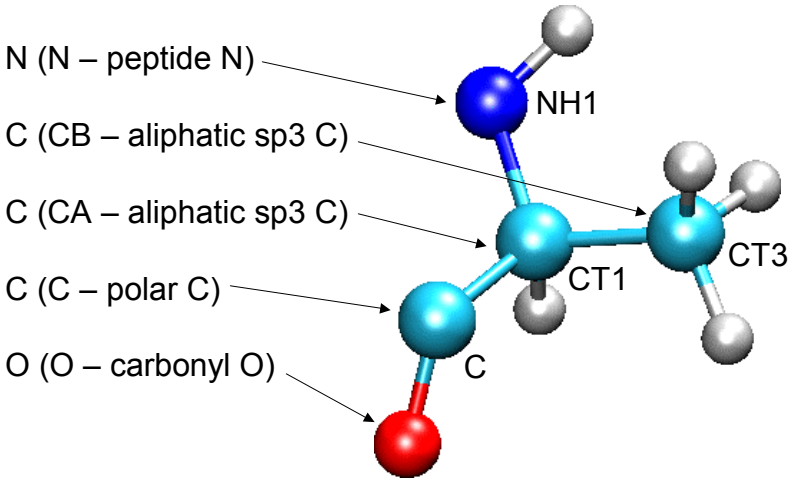
Contain parameters for force constants, equilibrium geometries, van der Waals radii, other data

PSF – Protein Structure Files

Files actually used by CHARMm, force field dependent, contain information from the RTF files, have a hierarchical organization of atoms, residues, segments

External Files

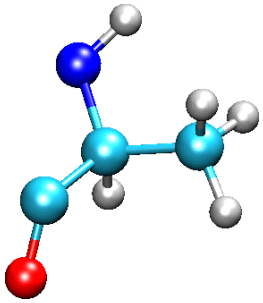
Atoms have different environments



RTF Files

```
MASS     10 HS     1.00800 H ! thiol hydrogen
MASS     11 HE1    1.00800 H ! for alkene; RHC=CR
MASS     20 C     12.01100 C ! carbonyl C, peptide backbone
MASS     21 CA    12.01100 C ! aromatic C
MASS     22 CT1   12.01100 C ! aliphatic sp3 C for CH
------- (missing data here)
RESI ALA          0.00
GROUP
ATOM N    NH1    -0.47  !     |
ATOM HN   H       0.31  !  HN-N
ATOM CA   CT1     0.07  !     |     HB1
ATOM HA   HB      0.09  !     |    /
GROUP                   !  HA-CA--CB-HB2
ATOM CB   CT3    -0.27  !     |    \
ATOM HB1  HA      0.09  !     |     HB3
ATOM HB2  HA      0.09  !   O=C
ATOM HB3  HA      0.09  !     |
GROUP                   !
ATOM C    C       0.51
ATOM O    O      -0.51
BOND CB CA  N  HN  N  CA
BOND C  CA  C  +N  CA HA  CB HB1  CB HB2  CB HB3
DOUBLE O  C
IMPR N -C CA HN  C CA +N O
DONOR HN N
ACCEPTOR O C
IC -C   CA   *N   HN    1.3551 126.4900  180.0000 115.4200  0.9996
IC -C   N    CA   C     1.3551 126.4900  180.0000 114.4400  1.5390
IC N    CA   C    +N    1.4592 114.4400  180.0000 116.8400  1.3558
IC +N   CA   *C   O     1.3558 116.8400  180.0000 122.5200  1.2297
IC CA   C    +N   +CA   1.5390 116.8400  180.0000 126.7700  1.4613
IC N    C    *CA  CB    1.4592 114.4400  123.2300 111.0900  1.5461
IC N    C    *CA  HA    1.4592 114.4400 -120.4500 106.3900  1.0840
IC C    CA   CB   HB1   1.5390 111.0900  177.2500 109.6000  1.1109
IC HB1  CA   *CB  HB2   1.1109 109.6000  119.1300 111.0500  1.1119
IC HB1  CA   *CB  HB3   1.1109 109.6000 -119.5800 111.6100  1.1114
```

Alanine

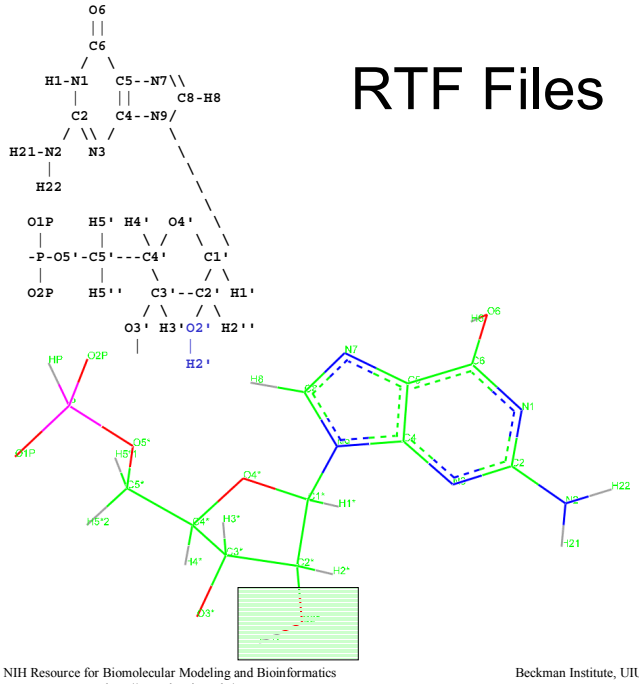


RTF Files: H_2O & Shake

```
RESI TIP3          0.000
! tip3p water model
! generate using noangle nodihedral
GROUP
ATOM OH2   OT      -0.834
ATOM H1    HT       0.417
ATOM H2    HT       0.417
BOND OH2 H1 OH2 H2 H1 H2
! the last bond is needed for shake
ANGLE H1 OH2 H2                ! required
ACCEPTOR OH2
```

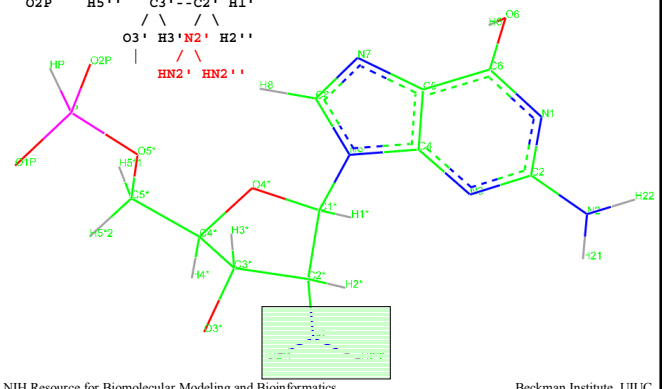

RTF Files

```
RESI GUA          -1.00  !
ATOM P    P        1.50  !
ATOM O1P  ON3     -0.78  !
ATOM O2P  ON3     -0.78  !
ATOM O5'  ON2     -0.57  !
ATOM C5'  CN8B    -0.08  !
ATOM H5'  HN8      0.09  !
ATOM H5'' HN8      0.09  !
GROUP                    !
ATOM C4'  CN7      0.16  !
ATOM H4'  HN7      0.09  !
ATOM O4'  ON6B    -0.50  !
ATOM C1'  CN7B     0.16  !
ATOM H1'  HN7      0.09  !
GROUP                    !
ATOM N9   NN2B    -0.02  !
ATOM C4   CN5      0.26  !
ATOM N2   NN1     -0.68  !
ATOM H21  HN1      0.32  !
ATOM H22  HN1      0.35  !
ATOM N3   NN3G    -0.74  !
ATOM C2   CN2      0.75
ATOM N1   NN2G    -0.34
ATOM H1   HN2      0.26
ATOM C6   CN1      0.54
ATOM O6   ON1     -0.51
ATOM C5   CN5G     0.00
ATOM N7   NN4     -0.60
ATOM C8   CN4      0.25
ATOM H8   HN3      0.16
GROUP
ATOM C2'  CN7B     0.14
ATOM H2'' HN7      0.09
ATOM O2'  ON5     -0.66
ATOM H2'  HN5      0.43
GROUP
ATOM C3'  CN7      0.01
ATOM H3'  HN7      0.09
ATOM O3'  ON2     -0.57
```



Modified RTF File

```
RESI GNA         -1.00  !
ATOM P     P       1.50  !
ATOM O1P   ON3    -0.78  !
ATOM O2P   ON3    -0.78  !
ATOM O5'   ON2    -0.57  !
ATOM C5'   CN8B   -0.08  !
ATOM H5'   HN8     0.09  !
ATOM H5''  HN8     0.09  !
GROUP                    !
ATOM C4'   CN7     0.16  !
ATOM H4'   HN7     0.09  !
ATOM O4'   ON6B   -0.50  !
ATOM C1'   CN7B    0.16  !
ATOM H1'   HN7     0.09  !
GROUP                    !
ATOM N9    NN2B   -0.02  !
ATOM C4    CN5     0.26  !
ATOM N2    NN1    -0.68  !
ATOM H21   HN1     0.32  !
ATOM H22   HN1     0.35  !
ATOM N3    NN3G   -0.74  !
ATOM C2    CN2     0.75
ATOM N1    NN2G   -0.34
ATOM H1    HN2     0.26
ATOM C6    CN1     0.54
ATOM O6    ON1    -0.51
ATOM C5    CN5G    0.00
ATOM N7    NN4    -0.60
ATOM C8    CN4     0.25
ATOM H8    HN3     0.16
GROUP
ATOM C2'   CN7B    0.14
ATOM H2''  HN7     0.09
ATOM N2'   NN1    -0.66
ATOM HN2'  HN1     0.21
ATOM HN2'' HN1     0.22
GROUP
ATOM C3'   CN7     0.01
ATOM H3'   HN7     0.09
ATOM O3'   ON2    -0.57
```



RTF Files

- Differences between RTF's for single molecules and residues
  - First residue converted to **N-terminus** NH_3^+
  - Last residue converted to **C-terminus** $CO{-}O^-$
- Special atoms

  `-C, -O, +N, +H, +CA`
  - Refers to atoms in residues preceding (-) or following (+)

Parameter Files

```
BONDS
!V(bond) = Kb(b - b0)**2
!Kb: kcal/mole/A**2
!b0: A
!atom type Kb          b0
!
C    C     600.000     1.3350 ! ALLOW ARO HEM
                ! Heme vinyl substituent (KK, from propene (JCS))
CA   CA    305.000     1.3750 ! ALLOW   ARO
                ! benzene, JES 8/25/89
------------------------- (missing data here)---------------
ANGLES
!V(angle) = Ktheta(Theta - Theta0)**2
!V(Urey-Bradley) = Kub(S - S0)**2
!Ktheta: kcal/mole/rad**2
!Theta0: degrees
!Kub: kcal/mole/A**2 (Urey-Bradley)
!S0: A
!
!atom types     Ktheta    Theta0   Kub     S0
!
CA   CA   CA    40.000    120.00   35.00   2.41620 ! ALLOW   ARO
                ! JES 8/25/89
CE1  CE1  CT3    48.00    123.50
```


AMBER Force Field

$$E_{\text{pot}} = \underbrace{\sum_{b} K_2(b-b_0)^2}_{(1)} + \underbrace{\sum_{\theta} H_\theta(\theta-\theta_0)^2}_{(2)} + \underbrace{\sum_{\phi} \frac{V_n}{2}[1+\cos(n\phi-\phi_0)]}_{(3)}$$
$$+ \underbrace{\sum \varepsilon[(r^*/r)^{12} - 2(r^*/r)^6]}_{(4)} + \underbrace{\sum q_i q_j/\varepsilon_{ij} r_{ij}}_{(5)} + \underbrace{\sum\left[\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}}\right]}_{(6)}$$

Consistent Valence Force Field

$$E_{\text{pot}} = \sum_{b} D_b[1 - e^{-\alpha(b-b_0)}] + \sum_{\theta} H_\theta(\theta - \theta_0)^2 + \sum_{\phi} H_\phi[1 + s\cos(n\phi)]$$

(1) (2) (3)

$$+ \sum_{\chi} H_\chi \chi^2 + \sum_{b}\sum_{b'} F_{bb'}(b - b_0)(b' - b'_0) + \sum_{\theta}\sum_{\theta'} F_{\theta\theta'}(\theta - \theta_0)(\theta' - \theta'_0)$$

(4) (5) (6)

$$+ \sum_{b}\sum_{\theta} F_{b\theta}(b - b_0)(\theta - \theta_0) + \sum_{\phi} F_{\phi\theta\theta'}\cos\phi(\theta - \theta_0)(\theta' - \theta'_0) + \sum_{\chi}\sum_{\chi'} F_{\chi\chi'}\chi\chi'$$

(7) (8) (9)

$$+ \sum \varepsilon[(r^*/r)^{12} - 2(r^*/r)^6] + \sum q_i q_j / \varepsilon r_{ij}$$

(10) (11)

The Power of the Force Field Approach

1) Force field-based simulations can handle large systems, are several orders of magnitude faster (and cheaper) than quantum-based calculations.

2) The analysis of the energy contributions can be done at the level of individual or classes of interactions.

3) Modification of the energy expression to bias the calculation.

Limitations of the Force Field Approach

Applications beyond the capability of classical force field methods include:

- Electronic transitions (photon absorption).
- Electron transport phenomena.
- Proton transfer (acid/base reactions).