Molecular Mechanical Force-fields

Electronic energy as a parametric function of the nuclear coordinates.

Fitting the parameters to experimental or higher level computational data.

building blocks in force field methods: Atoms

Electrons are not considered as individual particles.

Bonding information, must be provided explicitly, rather than being the result of solving the electronic Schrödinger equation.

quantum aspects of the nuclear motion are also neglected: Dynamics of the atoms is treated by classical mechanics.

Molecules are described by a "ball and spring" model: atoms having different sizes and softness bonds having different lengths and stiffness

To define a force field:

specify not only the functional form but also the parameters.

Two different force fields:

identical functional forms but very different parameters.

A force field is a single entity:

It is not correct to divide the energy into its individual components.

but

Some of the terms in a force field are nearly independent of the others (particularly the bond and angle terms).

Never Mix Two different force fields.

A force field is designed for certain properties:

Most Force fields designed to reproduce structural properties.

Some

can also predict other properties, such as molecular spectra.

Transferability

means that the same set of parameters can be used for a series of related molecules.

Transferability is important

if we want to use the force field to make predictions.

Specific accurate work: develop a model specific to that molecule.

Force fields are empirical:

there is no 'correct' form for a force field.

Accuracy and computational efficiency:

to be able to calculate the first and second derivatives of the energy with respect to the atomic coordinates.

The foundation of force field methods:

molecules tend to be composed of units that are structurally similar in different molecules (Functional Groups).

comparable equilibrium geometry:

C—H bond lengths: between 1.06 and 1.10 Å

comparable force constants:

C—H stretch vibrations: between 2900 and 3300 cm⁻¹

Dividing structural features into groups:

C—H bonds attached to single-, double- or triple bonded carbon then

the variation within each of these groups becomes even smaller.

Atom type

Depends on atomic number and local environment.

MM2 Force-field

there are 71 different atom types.

MM2 Type 1:

sp3-hybridized carbon.

MM2 Type 2:

sp2 carbon bonded to another sp2-carbon (simple double bond).

MM2 Type 3:

sp2 carbon bonded to an oxygen (carbonyl group)

MM2 Type 50:

sp2 carbon in aromatic ring with delocalized bonds.

Type	Symbol	Description
1	С	sp ³ -carbon
2	С	sp ² -carbon, alkene
3	С	sp ² -carbon, carbonyl, imine
4	C	sp-carbon
22	С	cyclopropane
29	C·	radical
30	C+	carbocation
38	C	sp ² -carbon, cyclopropene
50	C	sp ² -carbon, aromatic
56	C	sp ³ -carbon, cyclobutane
57	С	sp ² -carbon, cyclobutene
58	С	carbonyl, cyclobutanone
67	C	carbonyl, cyclopropanone
68	C	carbonyl, ketene
71	C	ketonium carbon

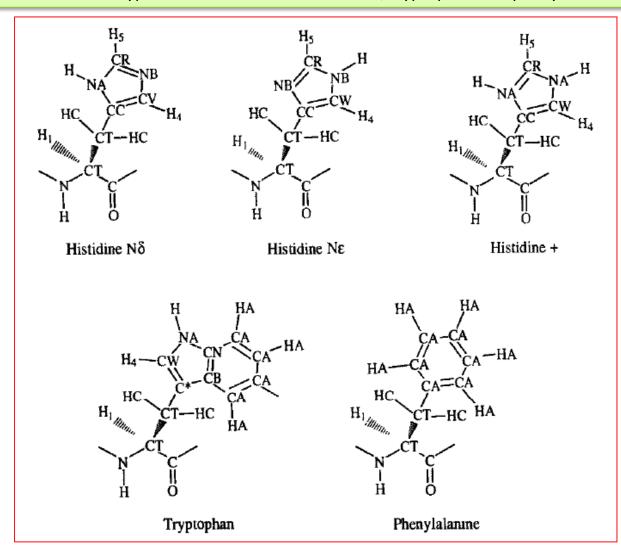
MM2 Carbon Atom types

MM2 Nitrogen and Oxygen Atom types

Туре	Symbol	Description
8	N	sp ³ -nitrogen
9	N	sp²-nitrogen, amide
10	N	sp-nitrogen
37	N	azo or pyridine (–N=)
39	N^+	sp ³ -nitrogen, ammonium
40	N	sp ² -nitrogen, pyrrole
43	N	azoxy (-N=N-O)
45	N	azide, central atom
46	N	nitro (–NO ₂)
72	N	imine, oxime (=N-)
6	О	sp³-oxygen
7	O	sp²-oxygen, carbonyl
41	O	sp²-oxygen, furan
47	O-	carboxylate
49	O	epoxy
69	O	amine oxide
70	O	ketonium oxygen

	More MI	M2 Atom types				
Туре	Symbol	Description				
15	S	sulfide (R ₂ S)				
16	S^+	sulfonium (R ₃ S ⁺)				
17	S	sulfoxide (R_2SO)				
18	S	sulfone (R_2SO_2)				
42	S	sp ² -sulfur, thiophene				
31	Ge	germanium				
32	Sn	tin				
33	Pb	lead				
34	Se	selenium				
35	Te	tellurium				
59	Mg	magnesium				
61	Fe	iron (II)				
62	Fe	iron (III)				
63	Ni	nickel (II)				
64	Ni	nickel (III)				
65	Co	cobalt (II)				
66	Co	cobalt (III)				

AMBER atom types for the amino acids histidine, tryptophan and phenylalanine.



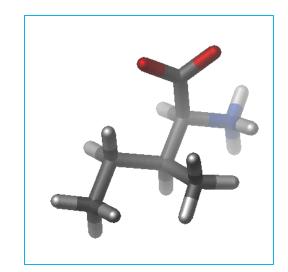
Force fields designed for specific classes of molecules use more specific atom types than force fields designed for general-purpose use.

Force Field Energy

An energy function of the nuclear coordinates:

$$E_{\text{FF}} = E_{\text{str}} + E_{\text{bend}} + E_{\text{tors}} + E_{\text{vdw}} + E_{\text{el}} + E_{\text{cross}}$$

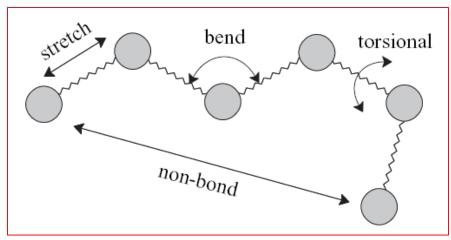
A sum of terms: each describing the energy required for distorting a molecule in a specific fashion.



geometries and relative energies can be calculated by optimization.

Stable molecules: minima on the potential energy surface

Conformational transitions: transition structure on the E_{FF} surface.



Number of Energy Terms in a Typical Force Field

10 bond terms:

2 C—C bonds and **8** C—H bonds.

18 angle terms:

1 C-C-C angle, **10** C-C-H angles and **7** H-C-H angles.

18 torsion terms:

12 H-C-C-H torsions and 6 H-C-C-C torsions.

27 non-bonded terms:

21 H-H interactions and 6 H-C interactions

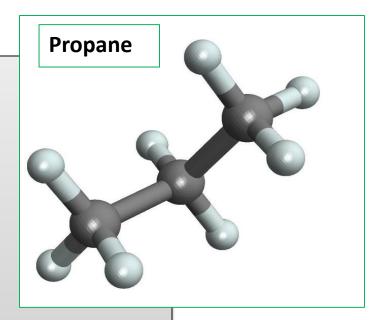
C—C bonds are symmetrically equivalent.

C—H bonds fall into two classes. (Different parameters might be used)

All angles are included even though some of them may not be independent of the others.

Thus, A sizeable number of terms are included in the force field, even for a molecule as simple as propane.

Even so, the number of terms (73) is many fewer than the number of integrals in an equivalent quantum mechanical calculation.



Stretch Energy

E_{str} is the energy function for stretching a bond between two atom types A and B.

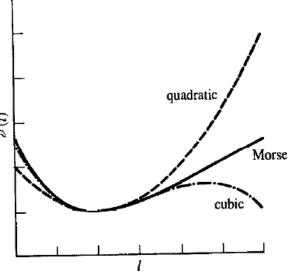
$$E_{\text{str}}(R^{\text{AB}} - R_0^{\text{AB}}) = E(0) + \frac{dE}{dR}(R^{\text{AB}} - R_0^{\text{AB}}) + \frac{1}{2}\frac{d^2E}{dR^2}(R^{\text{AB}} - R_0^{\text{AB}})^2$$

$$E_{\text{str}}(R^{\text{AB}} - R_0^{\text{AB}}) = k^{\text{AB}}(R^{\text{AB}} - R_0^{\text{AB}})^2 = k^{\text{AB}}(\Delta R^{\text{AB}})^2$$

$$E_{\rm str}(\Delta R^{\rm AB}) = k_2^{\rm AB}(\Delta R^{\rm AB})^2 + k_3^{\rm AB}(\Delta R^{\rm AB})^3 + k_4^{\rm AB}(\Delta R^{\rm AB})^4 + \cdots$$

$$E_{\text{Morse}}(\Delta R) = D(1 - e^{-\alpha \Delta R})^2$$

$$\alpha = \sqrt{\frac{k}{2D}}$$



E(0) term is normally set to zero.

Cubic anharmonicity constant k_3 is normally negative.

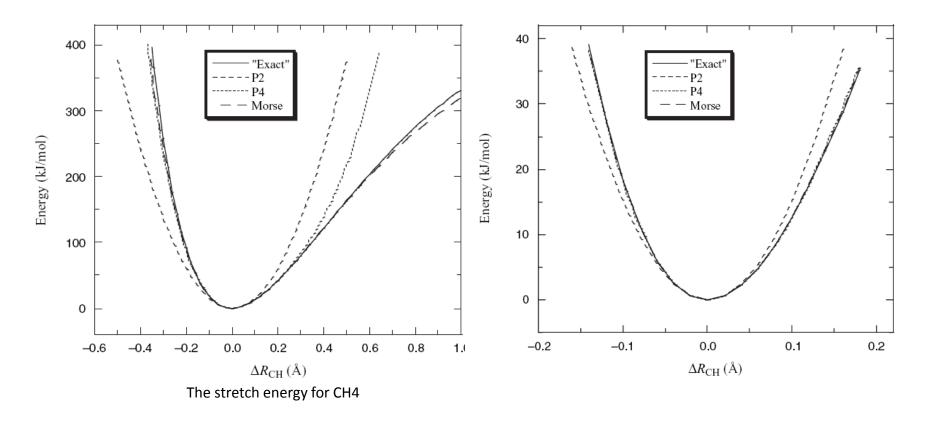
Quartic constant k_4 is normally positive.

Correct limiting behavior for a bond stretched to infinity is that the energy should converge towards the dissociation energy D.

Stretch Energy

The only important chemical region is within ~40 kJ/mol of the bottom of the curve.

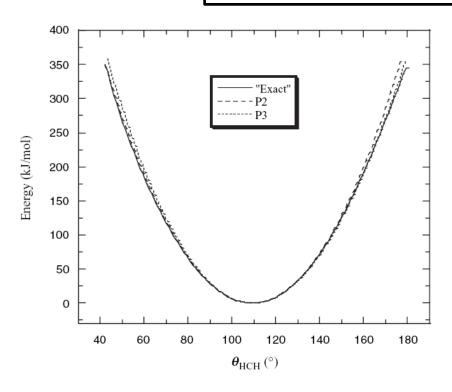
In this region, a fourth-order polynomial is essentially indistinguishable from either a Morse or the exact curve.



Bending energy

200

$$E_{\text{bend}}(\theta^{\text{ABC}} - \theta_0^{\text{ABC}}) = k^{\text{ABC}}(\theta^{\text{ABC}} - \theta_0^{\text{ABC}})^2$$



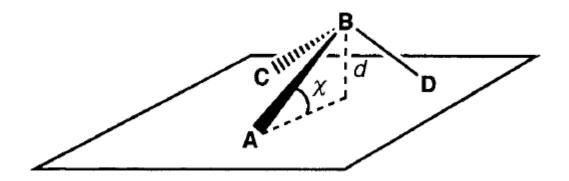
150 - "Exact" ---- P2 ----- P3 --- P3' --- P3'

The bending energy for CH4

The bending energy for H2O

Out-of-plane Bending Energy

$$E_{\text{oop}}(\chi) = k^{\text{B}} \chi^2$$
 or $E_{\text{oop}}(d) = k^{\text{B}} d^2$

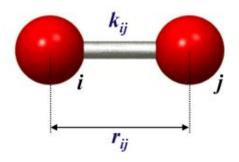


Energy cost for:

Pyramidalization of sp²-hybridized atoms.

Inversion barrier in sp³-hybridized atoms.

Forces associated with structural changes



$$V_B = k_{ij} (r_{ij} - r_0)^2$$

$$\mathbf{F}_i = -\nabla_i V_B$$

$$\mathbf{F}_i = -2k_{ij} (r_{ij} - r_0) \frac{\mathbf{r}_{ij}}{r_{ij}}$$

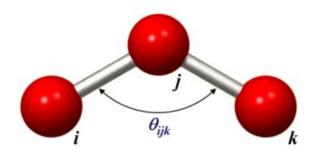
Two types of angle potentials:

1) harmonic potential (used in the CHARMM)

$$V_A = k_{ijk}^H \left(\theta_{ijk} - \theta_0\right)^2$$

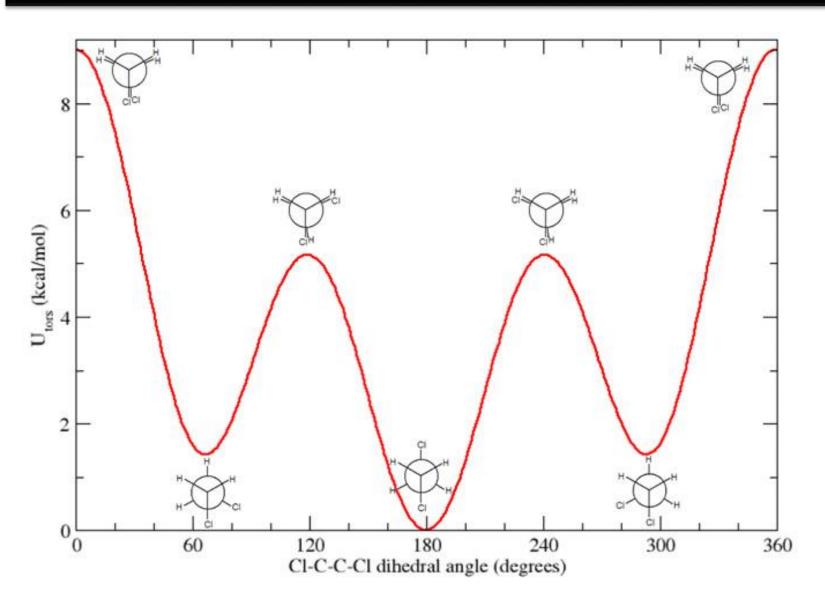
2) cosine harmonic potential (used in the GROMOS96)

$$V_A = k_{ijk}^{CH} \left(\cos \theta_{ijk} - \cos \theta_0\right)^2$$



$$\cos \theta_{ijk} = \frac{\mathbf{r}_{ij} \cdot \mathbf{r}_{kj}}{r_{ij} r_{kj}}$$

$$\mathbf{F}_{i} = -\nabla_{i} V_{A} = -\frac{\partial V_{A}}{\partial \cos \theta_{ijk}} \frac{\partial \cos \theta_{ijk}}{\partial \mathbf{r}_{i}}$$



The torsional energy is different from E_{str} and E_{hend} in three aspects:

- (1) A rotational barrier has contributions from both the non-bonded terms, and the torsional terms.
- (2) The torsional energy function must be periodic in the angle ω .
- (3) The cost in energy for rotation around a bond is often low.

Thus:

large deviations may occur, and

a Taylor expansion is not a good idea.

To encompass the periodicity, E_{tors} is written as a Fourier series:

$$E_{\text{tors}}(\omega) = \sum_{n=1} V_n \cos(n\omega)$$

n = 1 :rotation is periodic by 360°

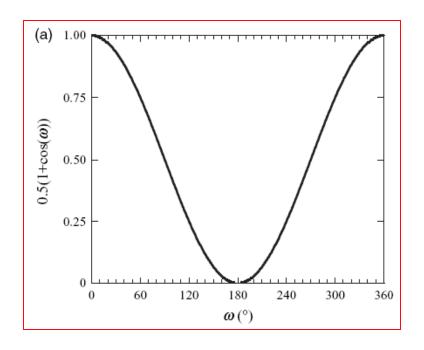
n = 2 :rotation is periodic by 180°

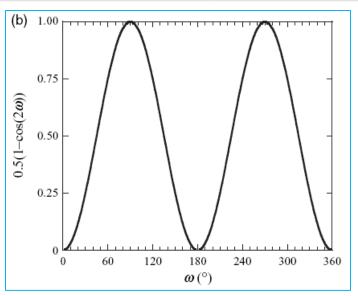
n = 3 :rotation is periodic by 120°

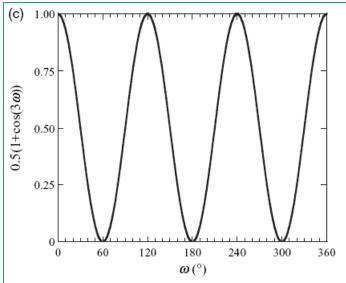
The V_n constants determine the size of the barrier for rotation.

Shift the zero point of the potential by adding a factor of one to each term:

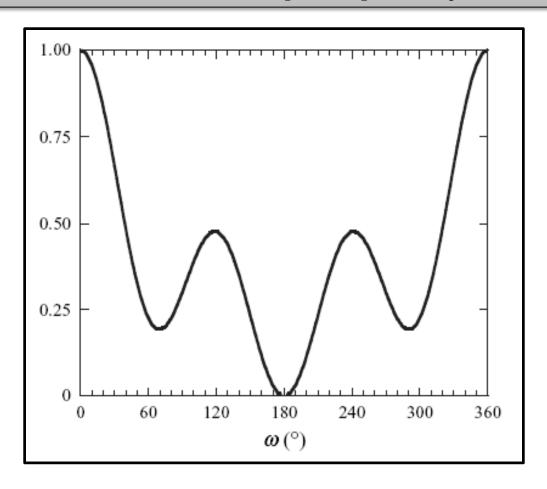
$$E_{\text{tors}}(\boldsymbol{\omega}^{\text{ABCD}}) = \frac{1}{2} V_1^{\text{ABCD}} \left[1 + \cos(\boldsymbol{\omega}^{\text{ABCD}}) \right]$$
$$+ \frac{1}{2} V_2^{\text{ABCD}} \left[1 - \cos(2\boldsymbol{\omega}^{\text{ABCD}}) \right]$$
$$+ \frac{1}{2} V_3^{\text{ABCD}} \left[1 + \cos(3\boldsymbol{\omega}^{\text{ABCD}}) \right]$$







an example with one trans and two less stable gauche minima and with a significant cis barrier, corresponding to the combination $V_1 = 0.5$, $V_2 = -0.2$, $V_3 = 0.5$

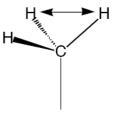


Other Energy Terms

The Improper torsional energy and Urey-Bradly



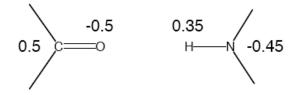
$$V_{improper} = K_{\varphi} (\varphi - \varphi_o)^2$$



$$V_{Urey-Bradley} = K_{UB} (r_{1,3} - r_{1,3o})^2$$

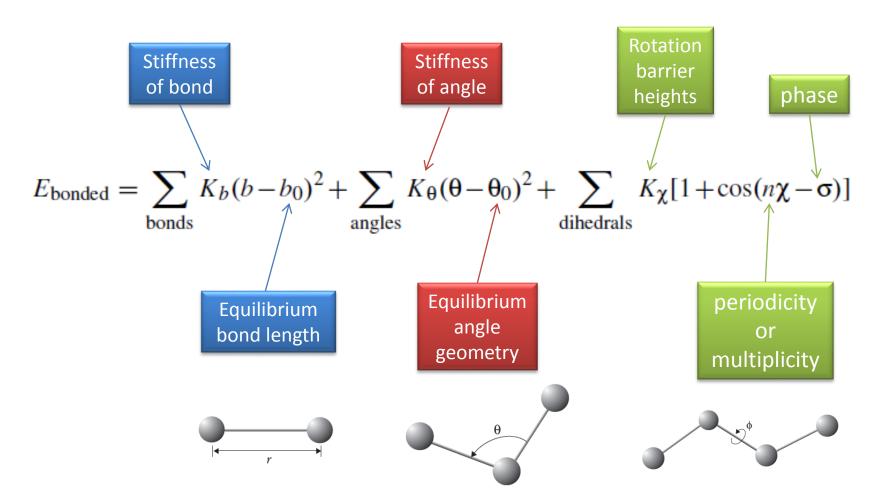
Treatment of hydrogen bonds???

$$V_{Hbond} = \sum_{Hbonds} \varepsilon_{HB} \left[\left(\frac{R_{HB,A-H}}{r_{A-H}} \right)^{12} - \left(\frac{R_{HB,A-H}}{r_{A-H}} \right)^{10} \right] * \cos(\theta_{A-H-D})$$

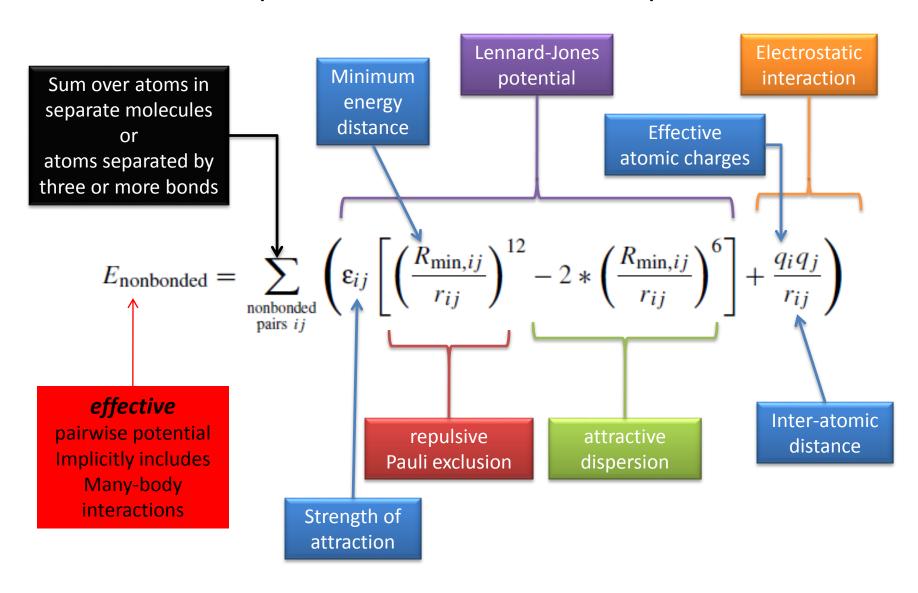


Bonded terms (Internal or intra-molecular interactions):

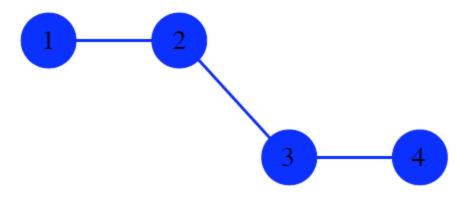
stretching of bonds bending of angles rotation of dihedrals



Nonbonded terms (external or intermolecular interactions):



Example of nonbond exclusions



nonbond (intermolecular) interactions between bonded atoms are treated with special rules

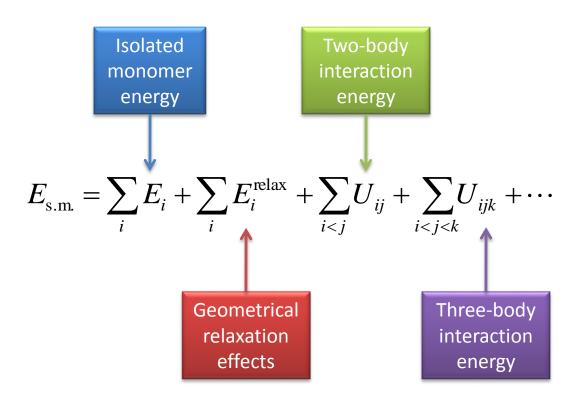
1,2 interactions: 0

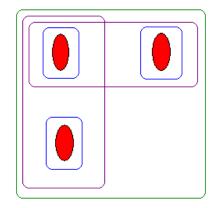
1,3 interactions: 0

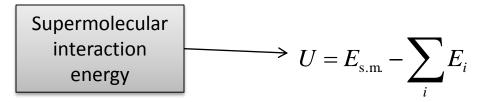
1,4 interactions: 1 or scaled

> 1,4 interactions: 1

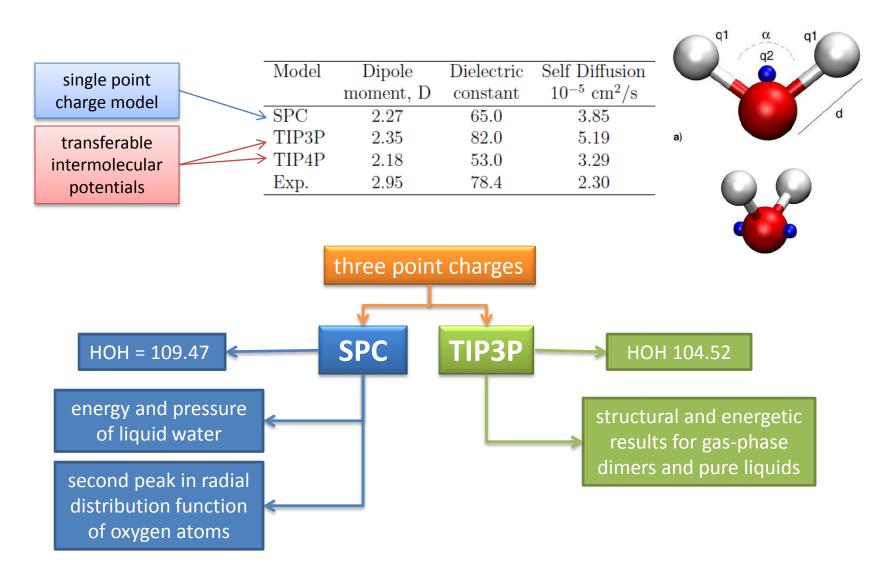
Many-body interactions





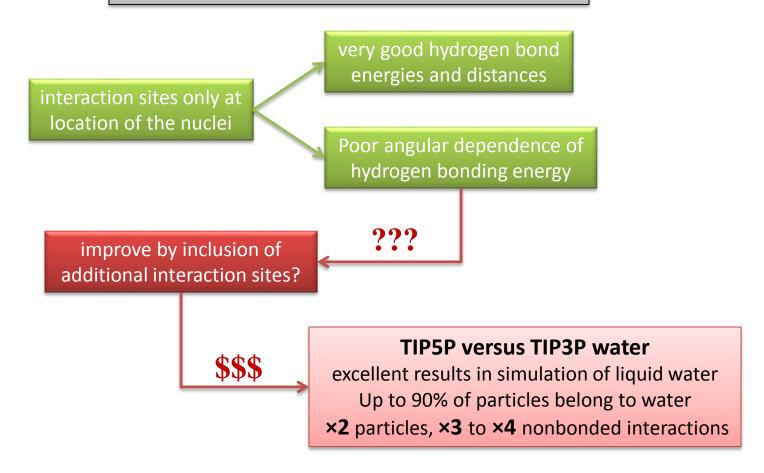


Modeling Water at the Molecular Level



no explicit term for hydrogen bonding

biologically important hydrogen bonds are handled by the combination of the LJ and Coulomb terms.



Class I force fields

Intramolecular (internal, bonded terms)

$$\begin{split} &\sum_{bonds} K_b \big(b - b_o\big)^2 + \sum_{angles} K_\theta \big(\theta - \theta_o\big)^2 + \sum_{torsions} K \phi \big(1 + \cos(n\phi - \delta)\big) \\ &+ \sum_{impropers} K_\phi \left(\phi - \phi_o\right)^2 + \sum_{Urey-Bradley} K_{UB} \big(r_{1\,3} - r_{1\,3,o}\big)^2 \end{split}$$

Intermolecular (external, nonbonded terms)

$$\sum_{nonbonded} \frac{q_i q_j}{4 \pi D r_{ij}} + \varepsilon_{ij} \left[\left(\frac{R_{\min,ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{\min,ij}}{r_{ij}} \right)^{6} \right]$$

CHARMM AMBER OPLS ECEPP GROMOS

Class II force fields

$$\sum_{bonds} \left[K_{b,2} (b - b_o)^2 + K_{b,3} (b - b_o)^3 + K_{b,4} (b - b_o)^4 \right]$$

$$+ \sum_{one los} \left[K_{\theta,2} (\theta - \theta_o)^2 + K_{\theta,3} (\theta - \theta_o)^3 + K_{\theta,4} (\theta - \theta_o)^4 \right]$$

+
$$\sum_{\text{dihedrals}} [K_{\phi,1}(1-\cos\phi) + K_{\phi,2}(1-\cos2\phi) + K_{\phi,3}(1-\cos3\phi)]$$

$$+\sum_{impropers}K_{\chi}\chi^{2}$$

$$+\sum_{bonds}\sum_{angles}K_{b\theta}(b-b_o)(\theta-\theta_o)$$

$$+\sum_{bonds}\sum_{dihedrals}(b-b_o)\left[K_{\phi,b1}\cos\phi+K_{\phi,b2}\cos2\phi+K_{\phi,b3}\cos3\phi\right]$$

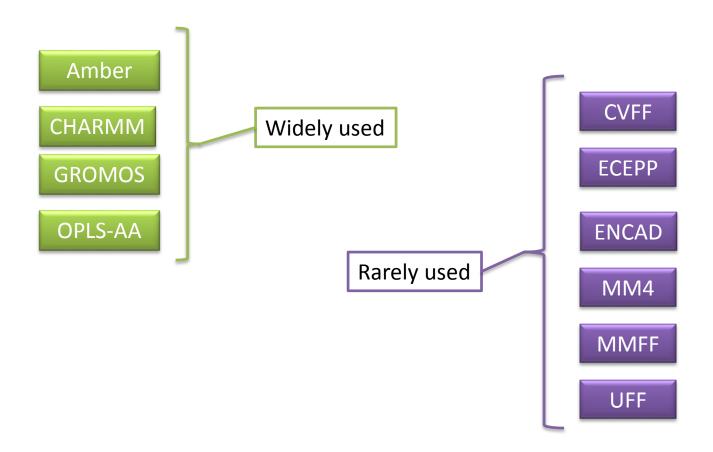
$$+ \sum_{bonds'} \sum_{dihedrals} (b' - b_o') [K_{\phi, b'1} \cos \phi + K_{\phi, b'2} \cos 2\phi + K_{\phi, b'3} \cos 3\phi]$$

$$+\sum_{\textit{angles dihedrals}}\sum_{\textit{dihedrals}}(\theta-\theta_o)\big[K_{\phi,\theta\,1}\cos\!\phi+K_{\phi,\theta\,2}\cos2\phi+K_{\phi,\theta\,3}\cos3\phi\big]$$

$$+\sum_{angles\ angles'\ dihedrals} \sum_{dihedrals} (\theta - \theta_o)(\theta' - \theta_o')\cos\phi$$

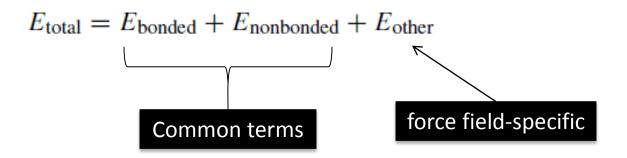
MM3 MMFF UFF CFF

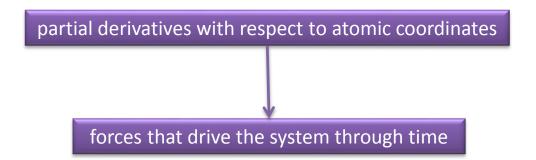
Comparison of Protein Force Fields



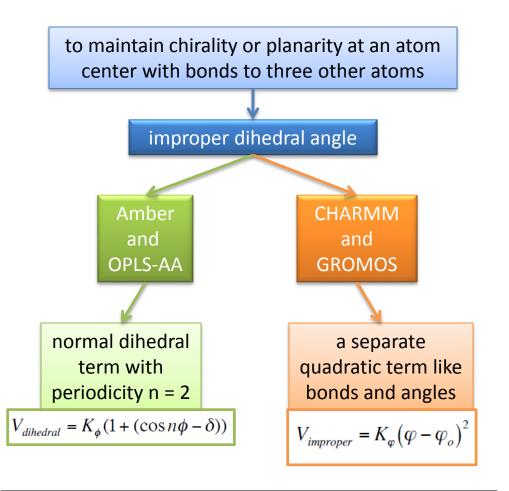
Force Fields Functional Forms

a differentiable function of atomic coordinates to represent energy of system





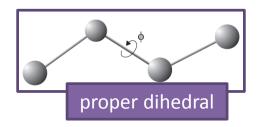
force field-specific terms: improper dihedral angle

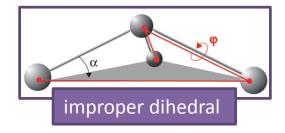


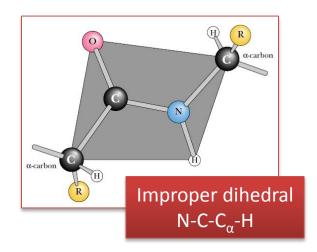
In the case of the GROMOS:

No hydrogen atoms bonded to aliphatic carbons.

Impropers preserve chirality at these carbon centers.



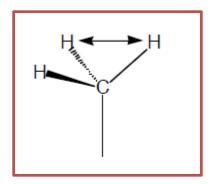




force field-specific terms: Urey-Bradly angle term

Specific to CHARMM force field

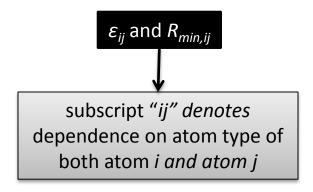
treats the two terminal atoms in an angle
(i.e., 1,3 atoms)
with a quadratic term that
depends on the
atom—atom distance

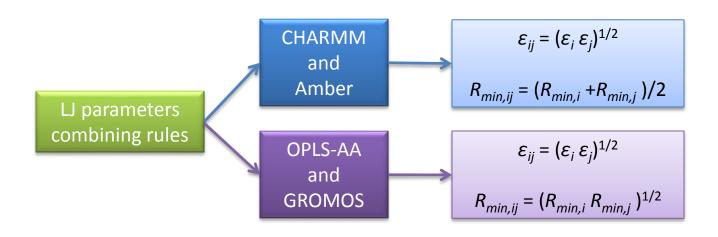


$$V_{Urey-Bradley} = K_{UB} \left(r_{1,3} - r_{1,3o} \right)^2$$

difference in the nonbonded parts

LJ parameters combining rules





difference in the nonbonded parts

handling of 1,4-nonbonded interactions

Amber

1,4-LJ interactions by ½
Coulomb interactions by 1.2

OPLS-AA

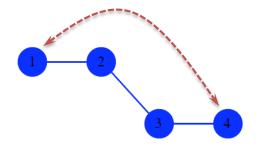
a scale factor of ½ to both interactions

GROMOS

a case-by-case approach for polar or ionic or aromatic situations

CHARMM

not scaling 1,4-nonbonded interactions special 1,4-LJ parameters for a few pairs



nonbonded interactions prevents the lowering of rotational barriers caused by electronic polarization.

$$H_3C$$
 H_{1111}
 CH_3
 H_3
 CH_3
 H_3
 CH_3

depth and shape of free energy contours of alanine dipeptide in water are substantially different for Amber, CHARMM, and OPLS-AA.

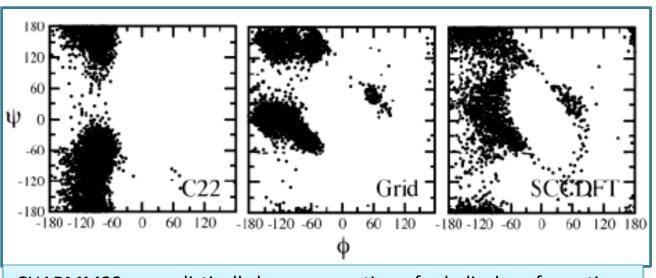
Improved Treatment of the Protein Backbone

Specific to CHARMM force field

2D dihedral energy correction map (CMAP)

$$f(\phi, \psi) = \sum_{i=1}^{4} \sum_{j=1}^{4} c_{ij} \left(\frac{\phi - \phi_L}{\Delta_{\phi}} \right)^{i-1} \left(\frac{\psi - \psi_L}{\Delta_{\psi}} \right)^{j-1}$$

a grid-based interpolation procedure for the energy difference between LMP2/cc-pVQZ(-g) and CHARMM22 dipeptide energy surfaces



CHARMM22: unrealistically large proportion of π -helical conformations

Table 1 Supported force fields, methods, and academic cost of the Amber, CHARMM, GRO-MACS, NAMD, and Tinker MD software packages

	Amber v. 9	CHARMM v. c32b2	GROMACS v. 3.3.2	NAMD v. 2.6	Tinker v. 4.2
Amber	X	X	X	X	X
CHARMM		X		X	X^a
GROMOS			X		
OPLS-AA		X	X	X	X
Thermostat ^b		X	X	c	X
Barostat ^b		X	X	X	
PME	X	X	X	X	X
LRC	X	X	X		
Parallel ^d	X	X	X	X	
Academic cost	\$400	\$600	\$0	\$0	\$0

All data in the table is from the software packages' respective web sites and the documentation available therein.

^aNo support for the additional grid-based CHARMM force field energy term.

^bAll packages provide means for thermostating and barostating. "X" indicates support for the recommended thermostating and barostating methods.

^cLacks recommended thermostat, but such thermostating can be achieved in constant temperature–constant pressure simulations through use of the Langevin piston barostat.

^dAbility to run MD on multiple CPUs in parallel.

J. Chem. Theory Comput. 2010, 6, 459-466

Journal of Chemical Theory and Computation

Implementation of the CHARMM Force Field in GROMACS: Analysis of Protein Stability Effects from Correction Maps, Virtual Interaction Sites, and Water Models

Pär Bjelkmar,† Per Larsson,† Michel A. Cuendet,‡ Berk Hess,† and Erik Lindahl*,†

Center for Biomembrane Research, Department of Biochemistry & Biophysics, Stockholm University, SE-106 91 Stockholm, Sweden, and Molecular Modeling Group, Swiss Institute of Bioinformatics, CH-1015 Lausanne, Switzerland

Received October 16, 2009

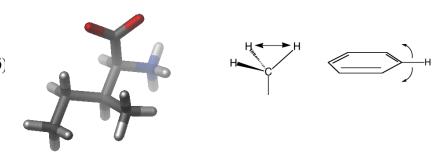
Table 2. Average ϕ , ψ Differences (deg) from Crystal Structure

	CHARMM27			CHARMM27 + CMAP				
system (1IGD)	$\langle \Delta \phi \rangle$	$\langle \Delta \phi \rangle$	$\langle \Delta \psi \rangle$	$\langle \Delta \psi \rangle$	$\langle \Delta \phi \rangle$	$\langle \Delta \phi \rangle$	$\langle \Delta \psi \rangle$	$\langle \Delta \psi \rangle$
				All Res	sidues			
TIP3P	0.5	16.5	-10.8	19.8	-1.2	9.3	-4.6	9.8
CHARMM TIP3P	2.8	14.1	-12.0	17.3	-1.4	9.1	-5.6	8.9
TIP4P	2.7	13.5	-12.3	18.6	-0.5	9.5	-6.8	10.3
TIP3P vsites	2.7	13.4	-16.3	22.4	-2.2	8.9	-3.4	8.5
OBC	3.6	14.4	-17.2	21.0	-0.9	12.7	-6.4	15.7
OBC vsites	2.7	14.0	-20.8	24.9	-1.7	11.6	-7.0	15.3
	Helical Residues							
TIP3P	4.1	8.0	-3.7	7.0	0.3	3.4	-0.1	3.5
CHARMM TIP3P	4.2	7.9	-3.8	6.9	0.1	3.1	-0.2	3.6
TIP4P	3.9	8.4	-3.7	7.3	0.3	3.7	0.1	3.5
TIP3P vsites	4.6	8.2	-4.5	7.8	0.2	2.7	-0.4	3.4
OBC	6.3	8.6	-7.4	9.0	2.7	5.6	-1.9	4.1
OBC vsites	7.1	8.9	-7.5	8.8	3.4	5.8	-2.0	3.9
	Sheet Residues							
TIP3P	6.6	11.5	-12.5	17.6	-1.4	8.6	-4.3	7.0
CHARMM TIP3P	5.3	10.5	-10.2	16.8	-1.6	8.8	-4.1	6.9
TIP4P	4.9	8.7	-10.5	17.4	-1.4	9.4	-4.5	7.9
TIP3P vsites	5.3	9.9	-14.5	20.1	-0.8	6.9	-2.5	6.4
OBC	5.7	11.0	-16.5	21.5	-1.4	6.9	-1.9	8.8
OBC vsites	5.0	10.5	-18.2	23.1	-0.7	6.4	-2.6	9.5

CHARMM All-Atom Additive Force-Field

Intramolecular (internal, bonded terms)

$$\sum_{bonds} K_b (b - b_o)^2 + \sum_{angles} K_\theta (\theta - \theta_o)^2 + \sum_{torsions} K_\phi (1 + \cos(n\phi - \delta)) + \sum_{impropers} K_\phi (\varphi - \varphi_o)^2 + \sum_{Urey-Bradley} K_{UB} (r_{1,3} - r_{1,3,o})^2 + \sum_{\phi,\psi} V_{CMAP}$$



dihedral energy correction map:

$$V_{CMAP} = f(\phi, \psi) = \sum_{i=1}^{4} \sum_{j=1}^{4} c_{ij} \left(\frac{\phi - \phi_{L}}{\Delta_{\phi}} \right)^{i-1} \left(\frac{\psi - \psi_{L}}{\Delta_{\psi}} \right)^{j-1}$$

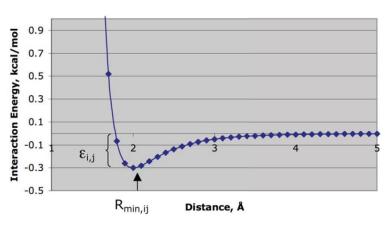
H1 H8
$$\Phi$$
 Ψ || H14 H16 H18 H18 H18 H18 H18

Intermolecular (external, nonbonded terms)

$$\sum_{nonbonded} \frac{q_i q_j}{4\pi D r_{ij}} + \varepsilon_{ij} \left[\left(\frac{R_{\min,ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{\min,ij}}{r_{ij}} \right)^{6} \right]$$

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

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



Extension of additive CHARMM force fields for drug like molecules

CGenFF paper:

J. Comp. Chem. 31 (2010) 671

