Lecture 4: Molecular Mechanics

Junmei Wang

Department of Pharmacology, University of Texas Southwestern Medical Center at Dallas

Junmei.wang@utsouthwestern.edu

Lab Section For QM

1. **Gaussian Software**

   Where to run Gaussian
   - BioHPC
   - TACC (lonestar lonestar5): https://www.tacc.utexas.edu/

   Technique support: www.Gaussian.com

2. **Gaussian Input**

   Formats
   - Cartesian
   - Z-matrix
   - Mixed

   Software to generate inputs
   - Antechamber (Mulan.swmed.edu/mmfft)
   - Openbabel
   - Avogadro
   - Gaussview

Lab Section For QM

3. **Display orbitals, electron density, vibration modes etc.**
    - **Avogadro (free)**
    - Gaussview
    - VMD (free)

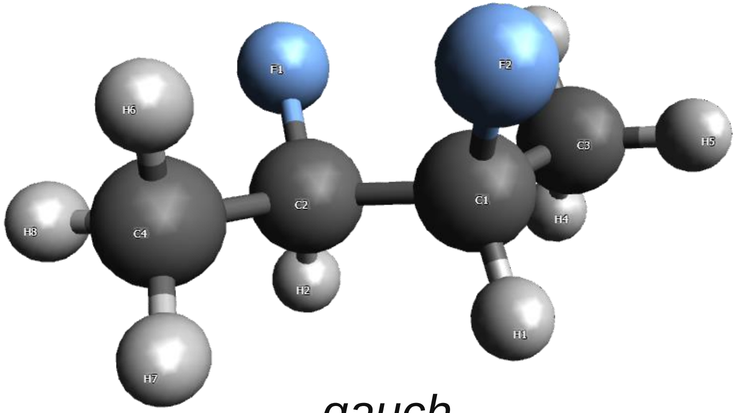
4. **Optimization**
    - How to conquer optimization failures
      Opt=calcfc,
      scf=QC

5. **Frequency Calculation**
    - ZPE
    - Polarizability
    - Scaling factor
    - About negative frequencies

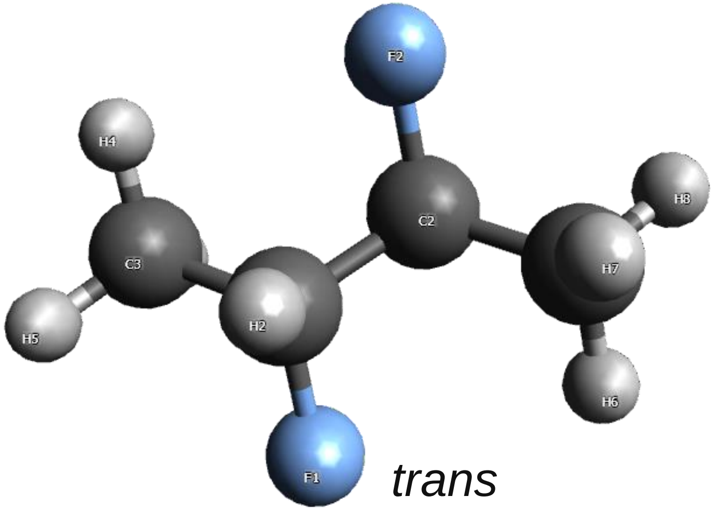
Lab Section For QM

6. Compare two conformations of a molecule

Conformation	**MP2 Energy (Hartree)**	**Thermochemistry Correction (hartree)**	**Relative Energy (kcal/mol)**
gauch	-335.86814	0.123287	0.47
trans	-335.86878	0.123182	0.0



gauch



trans

Lab Section For QM

7. **Thermochemistry property prediction (**ZPE is scaled by a factor of 0.9804**)**

- **Atomization energy** (energy difference between a molecule and its component atoms)

Molecule	**E (hartree)**	**ZPE (hartree)**	**AE (kcal/mol)**
H	-0.500273		
P	-341.196699		
PH_2	-342.509406	0.013217	187.6 (148.3)
Experiment			144.7

- **Electron affinity** (the energy released when an electron is added to a neutral molecule)

Molecule	**E (hartree)**	**ZPE (hartree)**	**AE (kcal/mol)**
PH_2^-	-342.554174	0.012412	
PH_2	-342.509406	0.013217	28.60 (1.24 ev)
Experiment			29.06 (1.26 ev)

Lab Section For QM

7. **Thermochemistry property prediction (**ZPE is scaled by a factor of 0.9804**)**

- **Ionization potential** (the energy required to remove an electron from a molecule)

Molecule	**E (hartree)**	**ZPE (hartree)**	**AE (kcal/mol)**
PH_2^+	-342.144153	0.013448	
PH_2	-342.509406	0.013217	229.35 (9.95 ev)
Experiment			226.45 (9.82 ev)

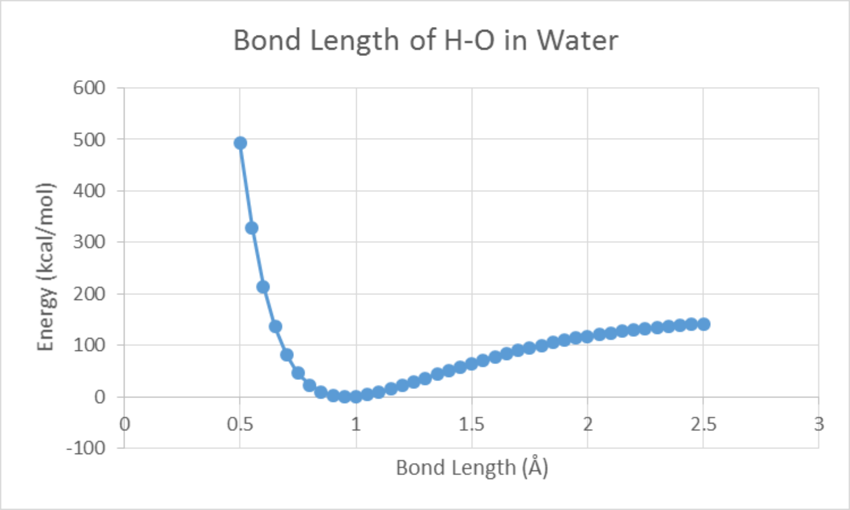
- **Proton affinity** (the energy released when a proton is added to a molecule)

Molecule	**E (hartree)**	**ZPE (hartree)**	**AE (kcal/mol)**
PH_4^+	-343.454091	0.035039	
PH_3	-343.146909	0.023784	185.7
Experiment			187.1

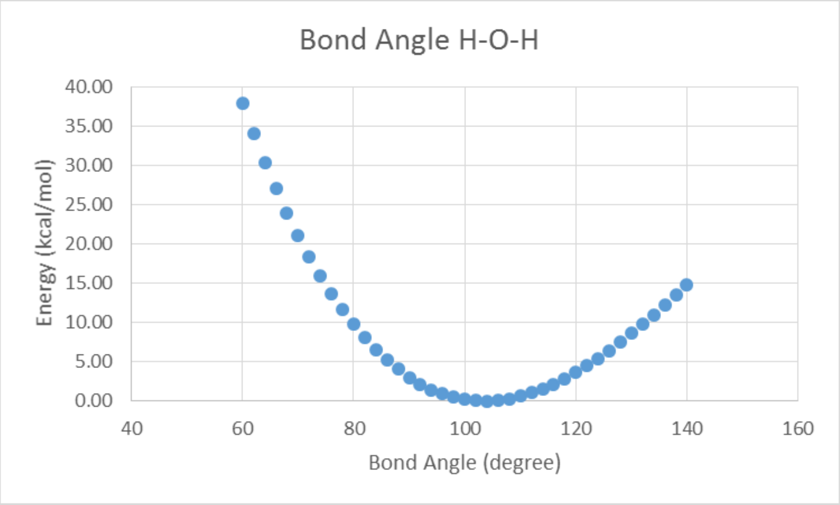
Lab Section For QM

8. Potential energy surface exploration

- Bond length



- Bond angle



Lab Section For QM

9. Reaction and Reactivity

- Enthalpy of reaction: $H^+ + H_2O \rightarrow H_3O^+$

Molecule	**E (hartree)**	**Thermal Corr (hartree)**	**AE (kcal/mol)**
H^+	0.0	0.0014167	
H_2O	-76.196848	0.025254	
H_3O^+	-76.475105	0.038471	-167.20
Experiment			-165.3 ±1.8

For H^+, only translational energy term, 3/2RT=0.889 kcal/mol is non-zero

- Enthalpy of reaction: $H_2CO \rightarrow H_2 + CO$

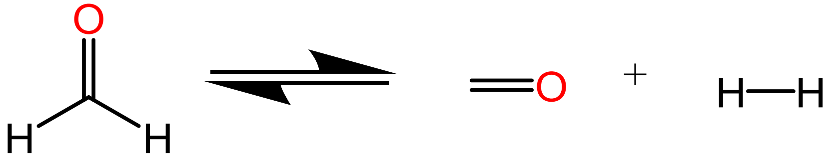
Molecule	**E (hartree)**	**Thermal Corr (hartree)**	**Total E (hartree)**
H_2	-1.144141	0.013633	-1.130508
CO	-113.021215	0.008134	-113.013081
H_2CO	-114.167747	0.031100	-114.136647
ΔH (Products - reactant)			-0.006942 (-4.4 kcal/mol)

Lab Section For QM

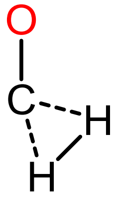
9. Reaction and Reactivity

- Conceive a transition structure
- Optimize the transition structure with keyword opt=TS
- Verify the transition structure by frequency analysis, there must has and only has one negative frequency (a characteristic feature of a transition state)
- Following the reaction path by performing **IRC** (intrinsic reaction coordinate) for the transition structure (keyword: IRC=rcfc)

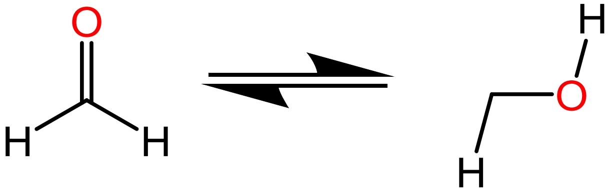
Reaction 1



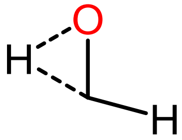
Transition structure:



Reaction 2



Transition structure:

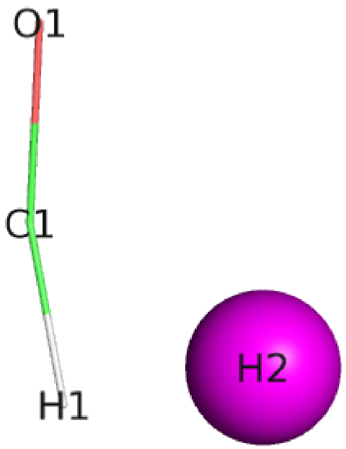


Lab Section For QM

9. **Reaction 1:** $H_2CO \rightarrow H_2 + CO$

Molecule	**E (hartree)**	**Activation Energy (kcal/mol)**
TS	-114.012907	
Reactants (H2CO)	-114.167747	97.2 (forward)
Products (H2+CO)	-114.165356	95.7 (reverse)

The trends of geometrical parameter change suggest TS is valid as “IRC I” tends to form product and “IRC II” tends to become reactant.



TS

	TS	**IRC (II)**	**IRC (I)**
C1-O1	1.18	1.16	1.18
C1-H1	1.09	2.35	1.08
H1-H2	1.35	0.83	1.40
C1-H2	1.73	1.87	1.70
∠O1-C1-H1	164.1	137.9	167.3
∠C1-H1-H2	88.9	115.6	85.6

Product Reactant

Lab Section For QM

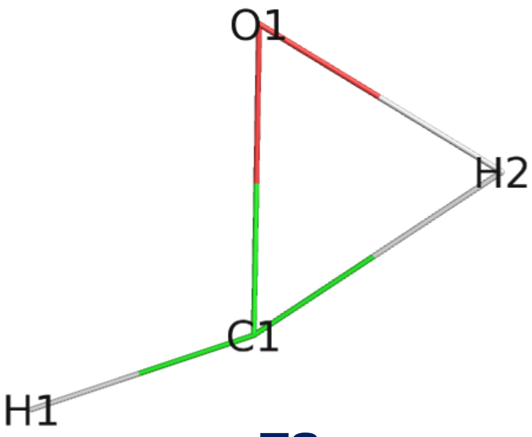
9. Reaction 2: $H_2CO \rightarrow$ H-C-O-H (trans hydrocarbene)

Molecule	E (hartree)	Activation Energy (kcal/mol)
TS	-114.019363	
Reactants (H2CO)	-114.022999	100.6 (reverse)
Products (trans HCOH)	-114.088123	43.1 (forward)

	TS	IRC (II)	IRC (I)
C1-O1	1.32	1.29	1.33
O1-H1	2.03	2.06	2.03
O1-H2	1.16	1.55	1.13
C1-H1	1.11	1.11	1.11
C1-H2	1.28	1.11	1.30
H1-H2	2.37	2.19	2.39
∠O1-C1-H1	112.9	117.9	112.5
∠O1-C1-H2	52.9	80.1	50.7
∠H1-C1-H2	165.9	162.0	163.2
∠C1-H1-H2	7.6	9.0	9.1
∠C1-O1-H1	30.2	28.4	30.4
∠C1-O1-H2	61.4	44.8	63.7

Reactant Product

The trends of geometrical parameter change suggest TS is valid as “IRC I” tends to form product and “IRC II” tends to become reactant.

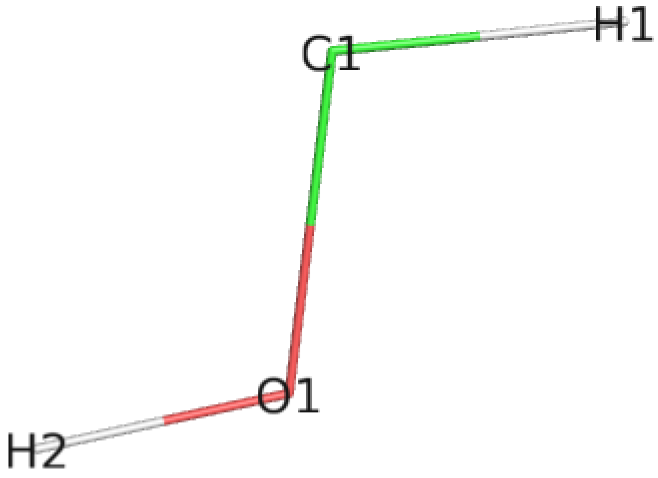


TS

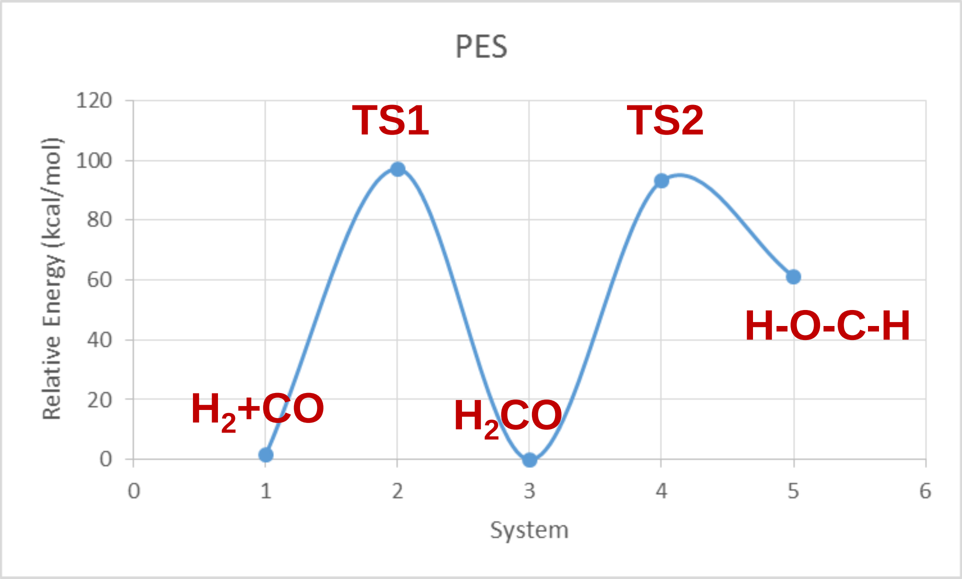
Lab Section For QM

9. Reaction and Reactivity

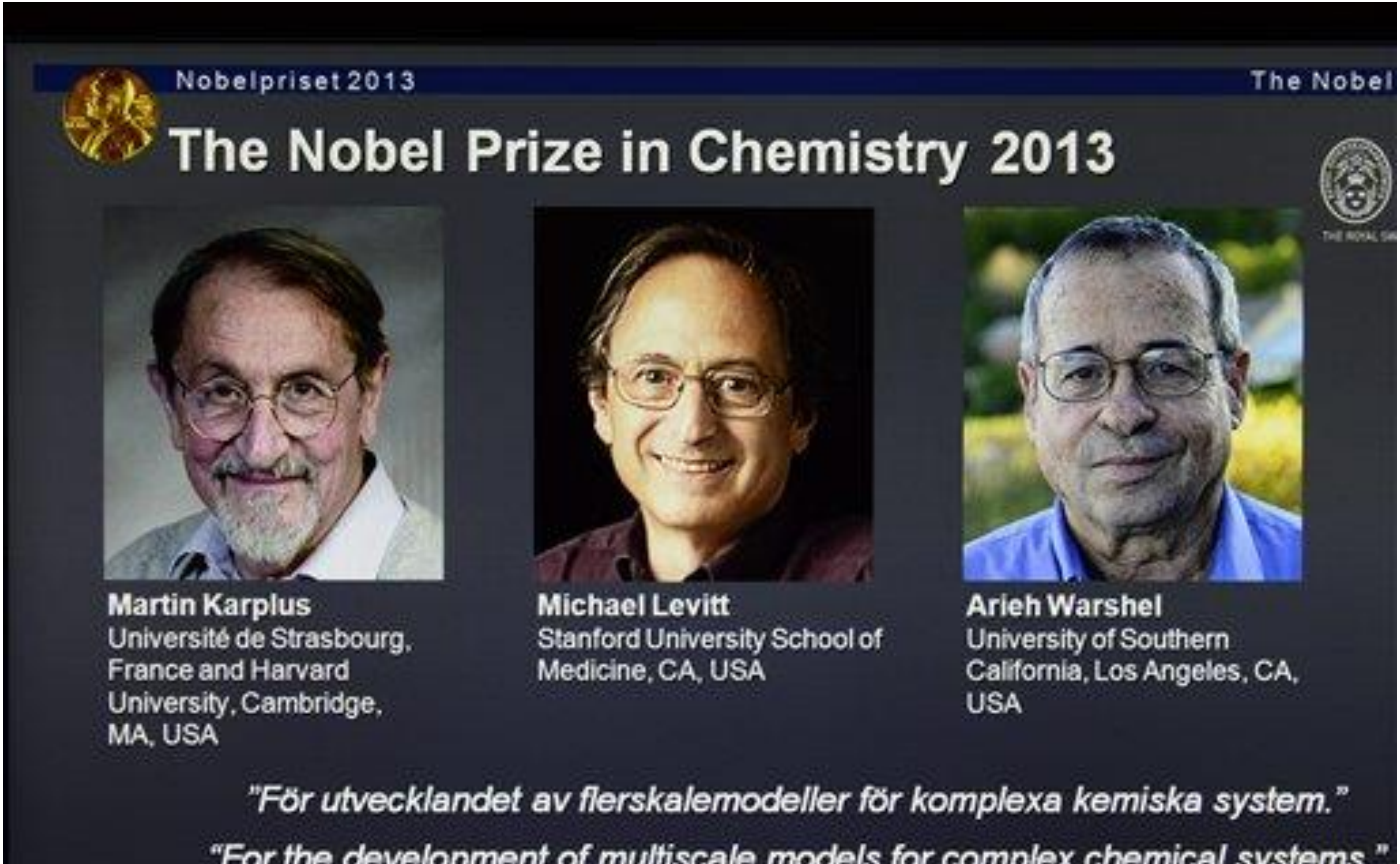
Molecule	E (hartree)	Relative Energy (kcal/mol)
Reactant (H2CO)	-114.167747	0
Product 1 (H2+CO)	-114.165356	1.5
Product 2 (H-O-C-H)	-114.070208	61.2
TS (reaction1)	-114.012907	97.2
TS (reaction2)	-114.019363	93.1



Product II
Trans-hydroxycarbene



Molecular Mechanics



Nobelpriset 2013

The Nobel

The Nobel Prize in Chemistry 2013

Martin Karplus
Université de Strasbourg, France and Harvard University, Cambridge, MA, USA

Michael Levitt
Stanford University School of Medicine, CA, USA

Arieh Warshel
University of Southern California, Los Angeles, CA, USA

"För utvecklandet av flerskalemodeller för komplexa kemiska system."

"For the development of multiscale models for complex chemical systems."

The Hierarchical Simulations of Chemistry

Hierarchical Simulations Chemistry

Multi-scale

Years
Yards

Seconds
Inches

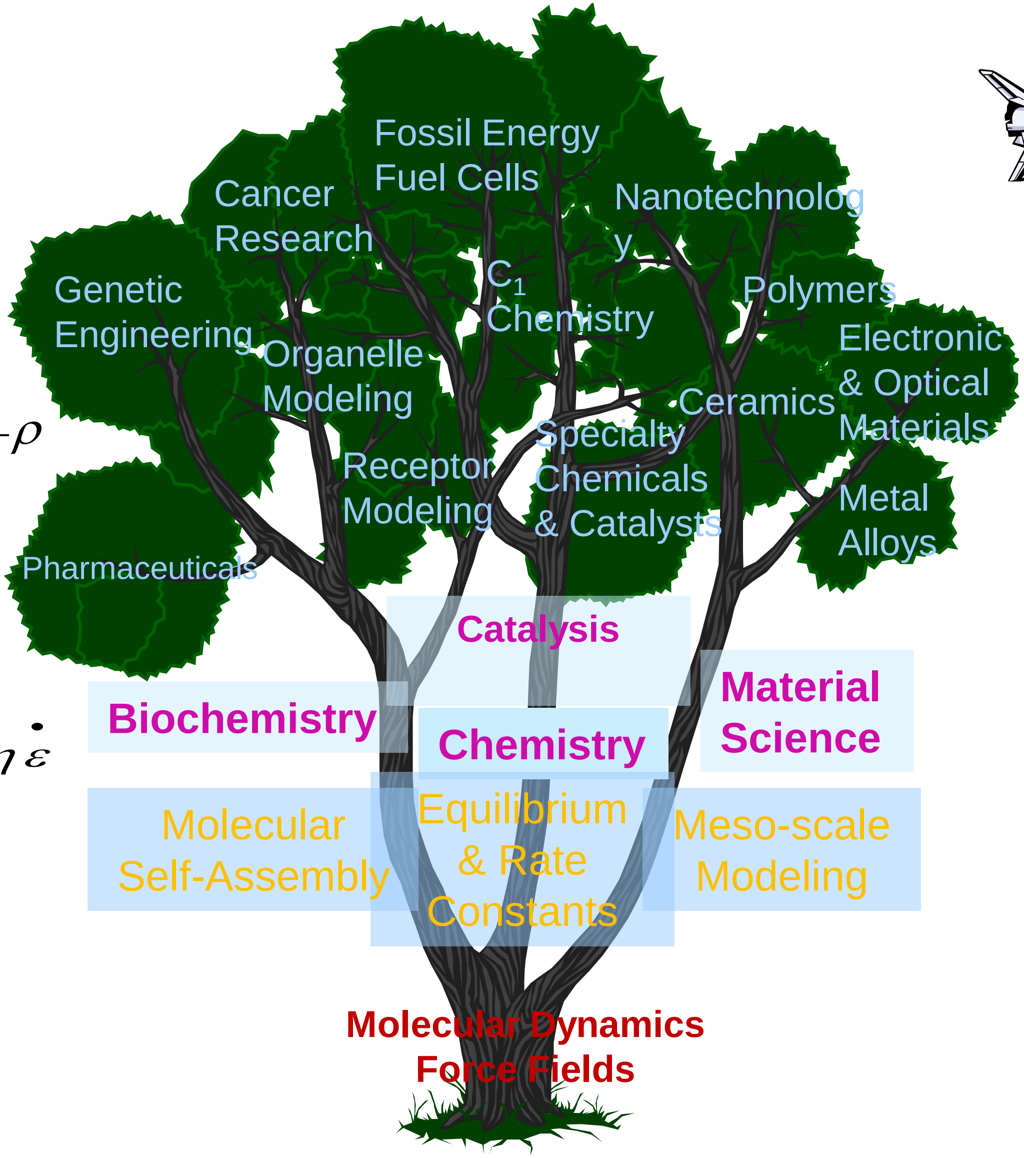
$\nabla^2\phi = -\rho$

Microseconds
Microns

$\sum = \eta\,\dot{\varepsilon}$

Picoseconds
Nanometers

Femptoseconds
Angstroms



Fossil Energy
Fuel Cells

Cancer
Research

Nanotechnolog
y

Genetic
Engineering

C_1
Chemistry

Polymers

Organelle
Modeling

Electronic
& Optical
Materials

Ceramics

Specialty
Chemicals
& Catalysts

Receptor
Modeling

Metal
Alloys

Pharmaceuticals

Catalysis

Biochemistry

Chemistry

Material Science

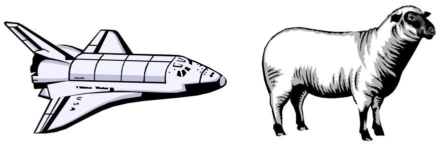
Molecular
Self-Assembly

Equilibrium
& Rate
Constants

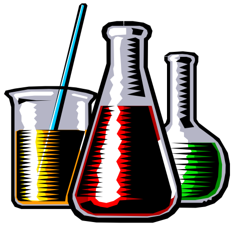
Meso-scale
Modeling

Molecular Dynamics
Force Fields

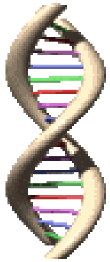
QUANTUM
MECHANICS



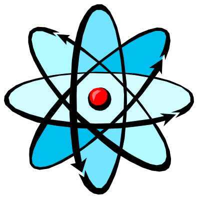
Design



Materials

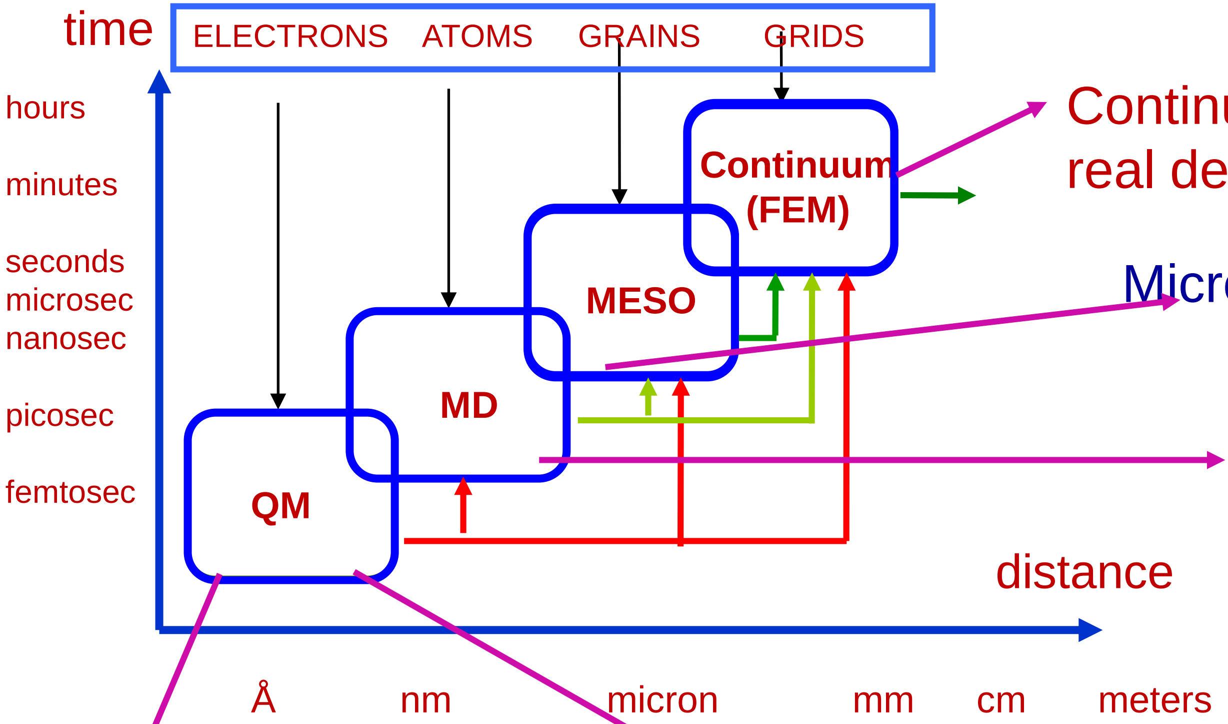


Molecules



Atoms
Electrons

Mechanical Engineer



Continuum simulations of real devices and materials

Micromechanical modeling

Deformation and Failure
(dislocations, cracks, etc.)

Transport properties

(diffusion, thermal transport, etc.)

Accurate calculations for bulk phases and molecules (EOS, dissociation curves, etc.)

Optical properties

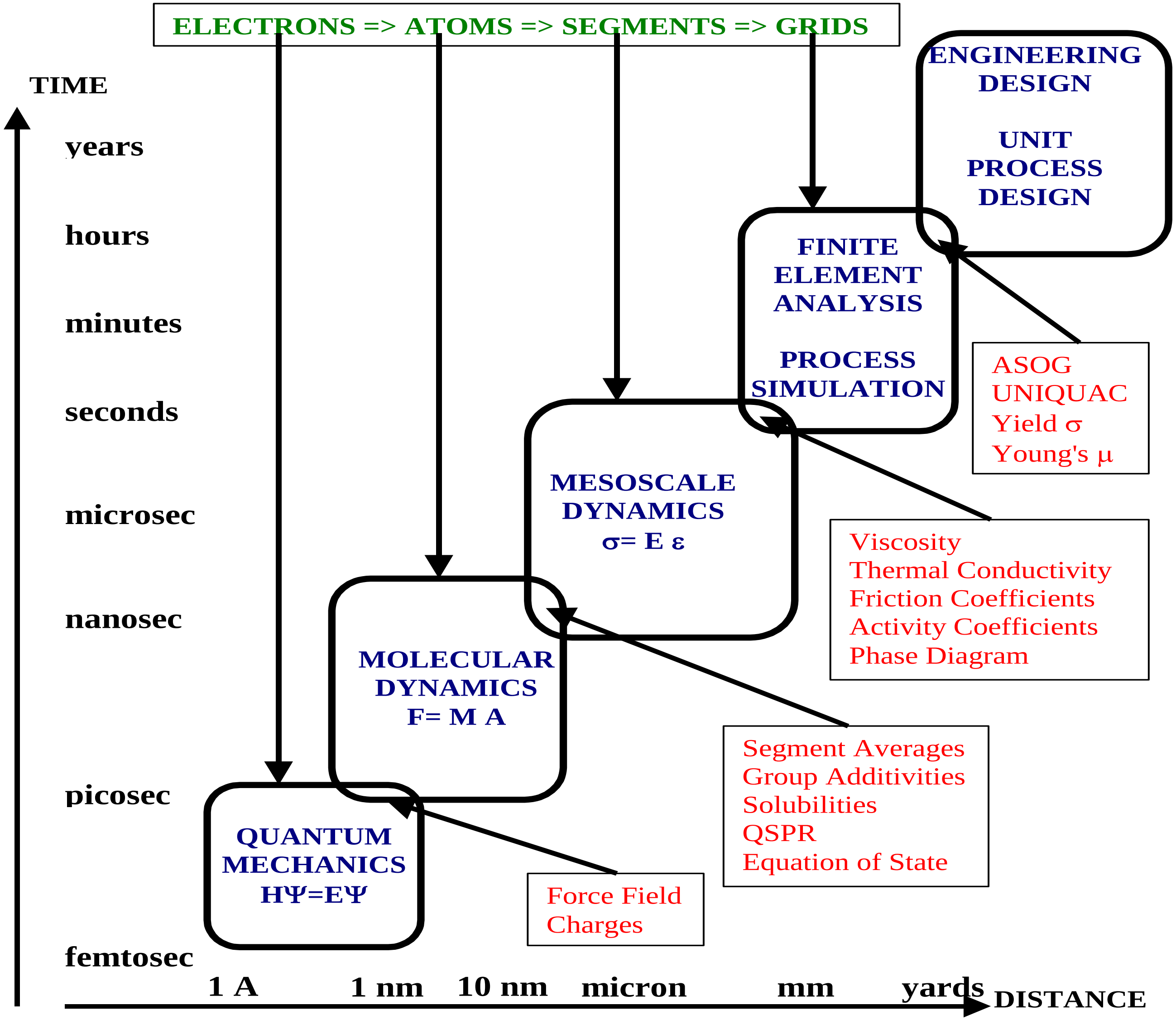
New generation reactive force fields based purely on first principles

For metals, oxides, organics.

Describes: mechanical properties, chemistry, charge transfer, atomic polarization, etc.

Chemical Engineer

Hierarchical Modeling



ELECTRONS => ATOMS => SEGMENTS => GRIDS

TIME

years

hours

minutes

seconds

microsec

nanosec

picosec

femtosec

QUANTUM MECHANICS
$H\Psi=E\Psi$

MOLECULAR DYNAMICS
F= M A

MESOSCALE DYNAMICS
$\sigma = E\ \varepsilon$

FINITE ELEMENT ANALYSIS

PROCESS SIMULATION

ENGINEERING DESIGN

UNIT PROCESS DESIGN

Force Field
Charges

Segment Averages
Group Additivities
Solubilities
QSPR
Equation of State

Viscosity
Thermal Conductivity
Friction Coefficients
Activity Coefficients
Phase Diagram

ASOG
UNIQUAC
Yield σ
Young's μ

1 A 1 nm 10 nm micron mm yards DISTANCE

Molecular Mechanics Background

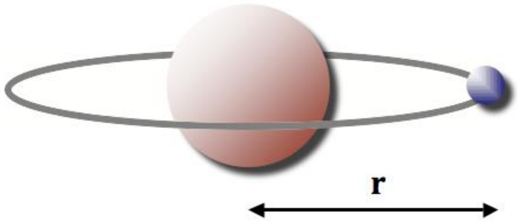
The "mechanical" molecular model was developed out of a need to describe molecular structures and properties in as practical a manner as possible.

- The great computational speed of molecular mechanics allows for its use in molecules containing thousands of atoms.

Molecular mechanics methods are based on the following principles:

- Nuclei and electrons are lumped into atom-like particles.
- Atom-like particles are spherical (radii obtained from measurements or theory) and have a net charge (obtained from theory).
- Interactions are based on springs and classical potentials.
- Interactions must be preassigned to specific sets of atoms.
- Interactions determine the **spatial distribution** of atom-like particles and their **energies**.

The Molecular Mechanics



The Born-Oppenheimer Approximation is the Foundation of Molecular Mechanics

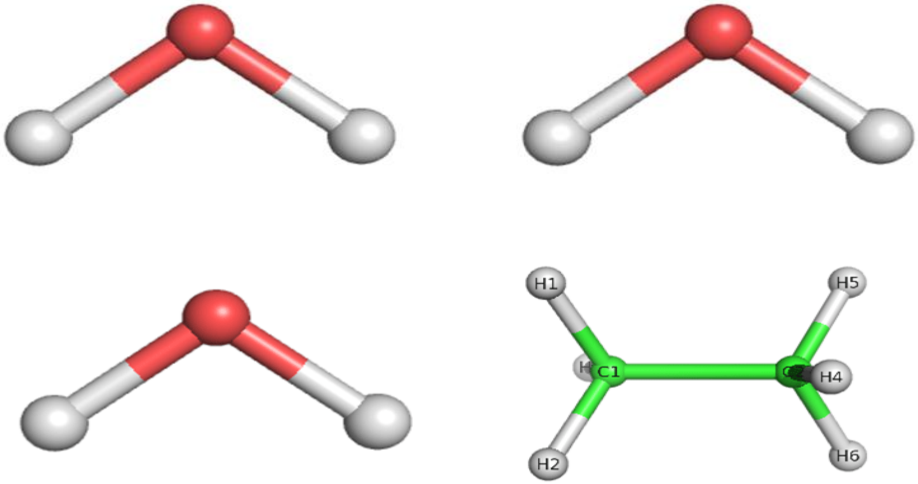
Potential Functions

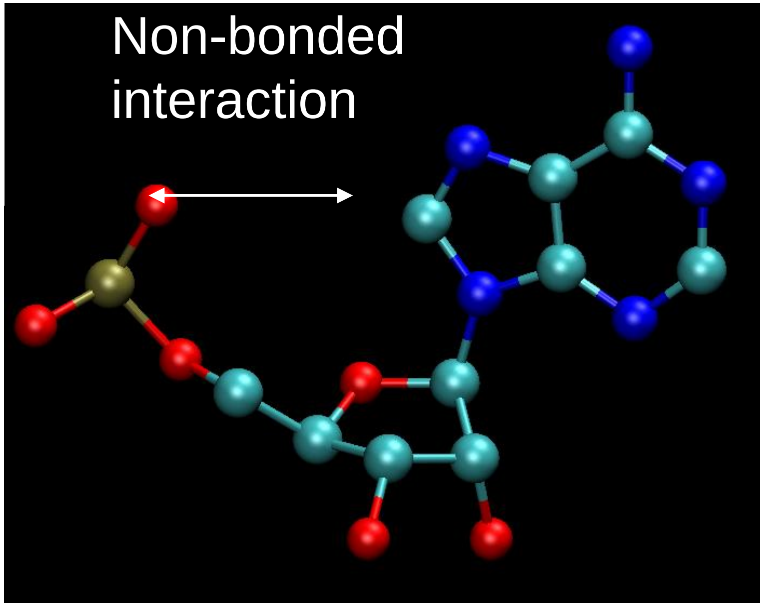
- **Molecular Mechanics**

$$\boxed{V_{potential}} = \sum_{\text{bonds}} K_r\left(r - r_{eq}\right)^2 + \sum_{\text{angles}} K_\theta\left(\theta - \theta_{eq}\right)^2 + \sum_{\text{dihedrals}} \frac{V_n}{2}\left[1 + \cos\left(n\varphi - \gamma\right)\right] +$$

$$\sum_{\substack{i<j \\ \text{van der Waals}}} \left(\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}}\right) + \sum_{\substack{i<j \\ \text{electrostatic}}} \frac{q_i q_j}{R_{ij}}$$

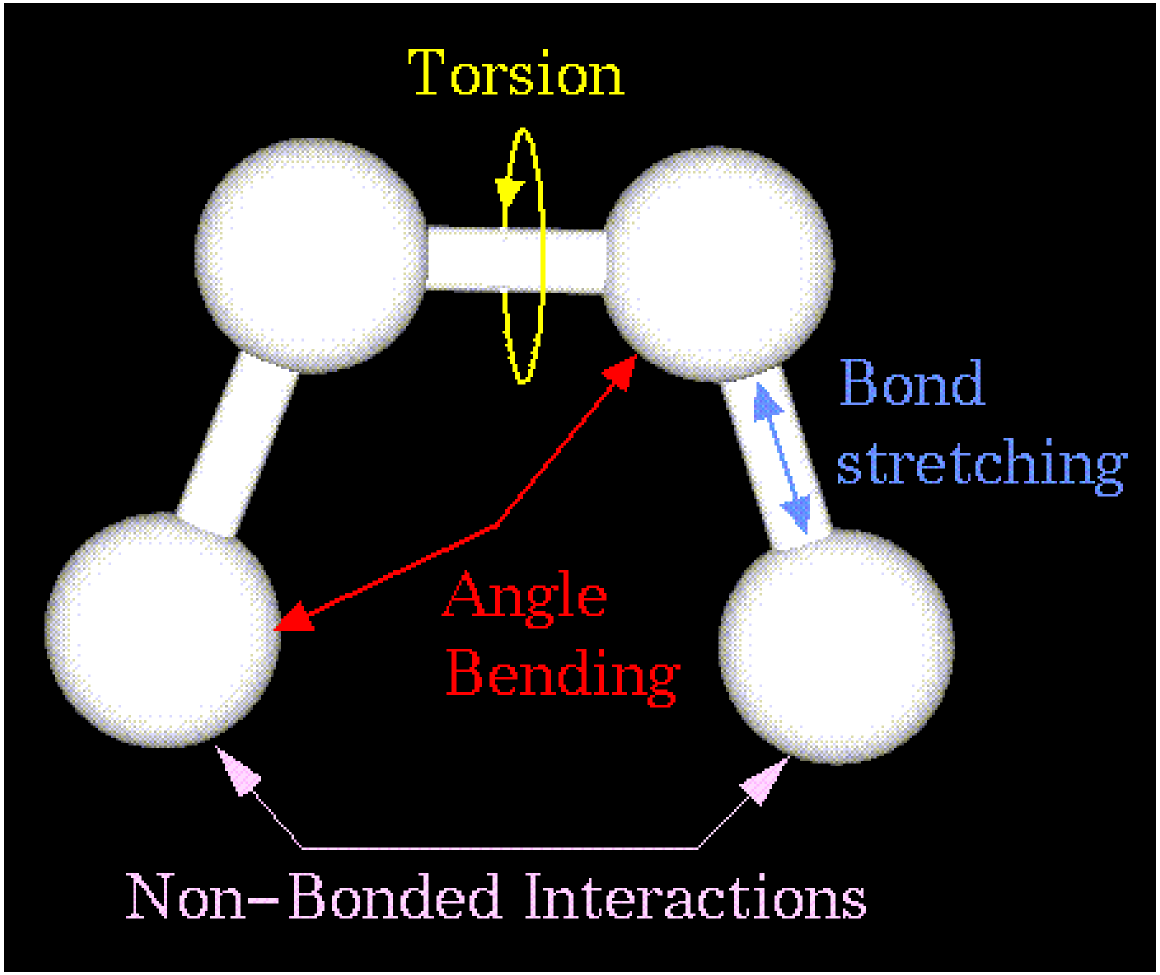
Additive
Harmonic
Effective





The Anatomy of a Molecular Mechanics Force-Field

The mechanical molecular model considers atoms as spheres and bonds as springs. The mathematics of spring deformation can be used to describe the ability of bonds to stretch, bend, and twist:



Non-bonded atoms (greater than two bonds apart) interact through van der Waals attraction, steric repulsion, and electrostatic attraction/repulsion. These properties are easiest to describe mathematically when atoms are considered as spheres of characteristic radii.

The Anatomy of a Molecular Mechanics Force-Field

The object of molecular mechanics is to predict the **energy** associated with a given conformation of a molecule. However, molecular mechanics energies have no meaning as absolute quantities. Only differences in energy between two or more conformations have meaning. A simple molecular mechanics energy equation is given by:

Total Energy =Stretching Energy +Bending Energy +Torsion Energy +Non-Bonded Interaction Energy*

Within the molecular framework, the "total energy" of a molecule is described in terms of a sum of contributions arising from **** ALL DEVIATIONS**** from "ideal" bond distances (stretch contributions), bond angles (bend contributions) and dihedral angles (torsion contributions) summarized by

$$E^{Total} = \sum_{i}^{bonds} E_i^{stretch} + \sum_{i}^{bondangles} E_i^{bend} + \sum_{i}^{dihedralangles} E_i^{torsion} + \sum_{ij}^{atompairs} E^{vanderwaals} + \sum_{ij}^{atompairs} E^{electrostatic}$$

Covalent Interactions   Non - covalent Interactions

The Anatomy of a Molecular Mechanics Force-Field

These equations together with the data (parameters) required to describe the behavior of different kinds of atoms and bonds, is called a "**FORCE FIELD**".

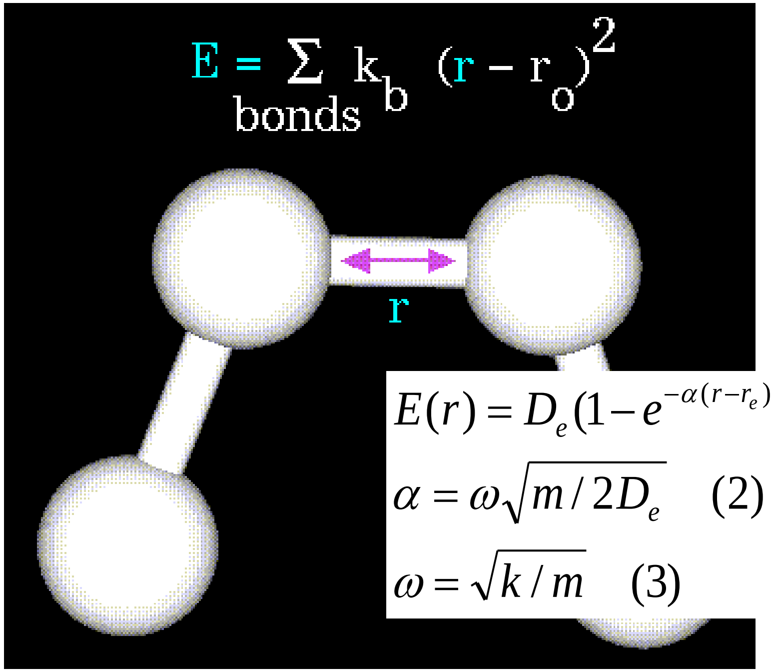
The molecular mechanics "FORCE FIELD" relates the motions, and energies of motions of atoms within the molecule. The force field is used to govern how the parts of a molecule relate to each other, that is, how each atom or group of atoms *is affected by its environment*, and how these forces contribute to the structure of the molecule.

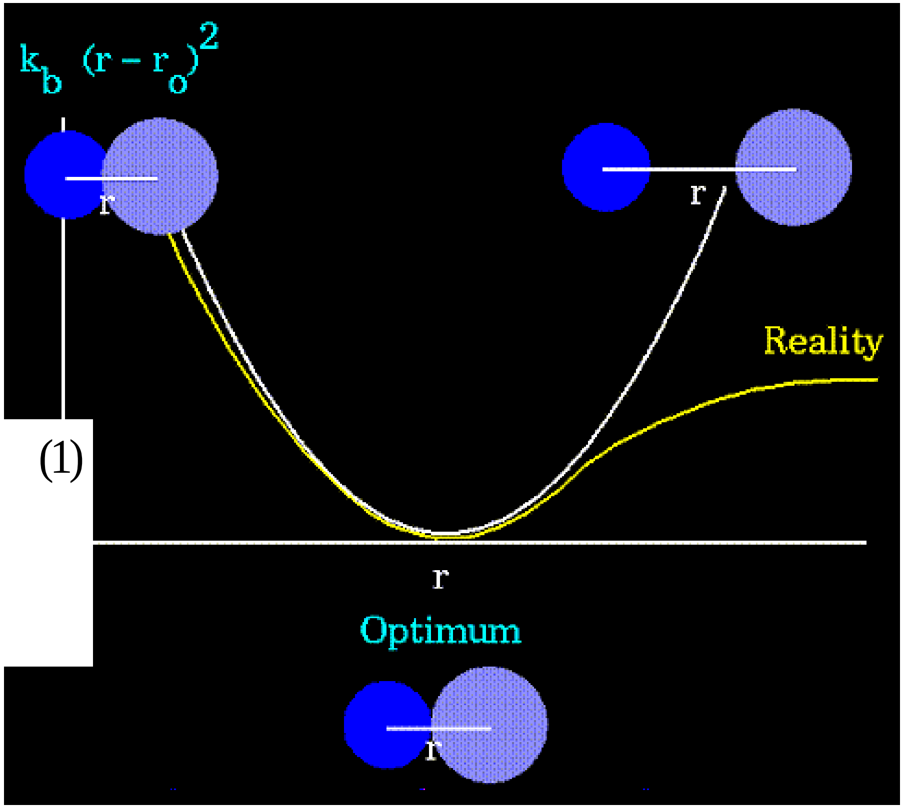
Many different kinds of force-fields have been developed over the years. Some include additional energy terms that describe other kinds of deformations. Some force-fields account for coupling between bending and stretching in adjacent bonds in order to improve the accuracy of the mechanical model.

The mathematical form of the energy terms varies from force-field to force-field. The more common forms will be described

STRETCHING ENERGY

- The stretching energy equation is based on Hooke's law. The "k_b" parameter controls the stiffness of the bond spring, while "r_o" defines its equilibrium length. Unique "k_b" and "r_o" parameters are assigned to each pair of bonded atoms based on their types (e.g. C-C, C-H, O-C, etc.). This equation estimates the energy associated with vibration about the equilibrium bond length. This is the equation of a parabola, as can be seen in the following plot





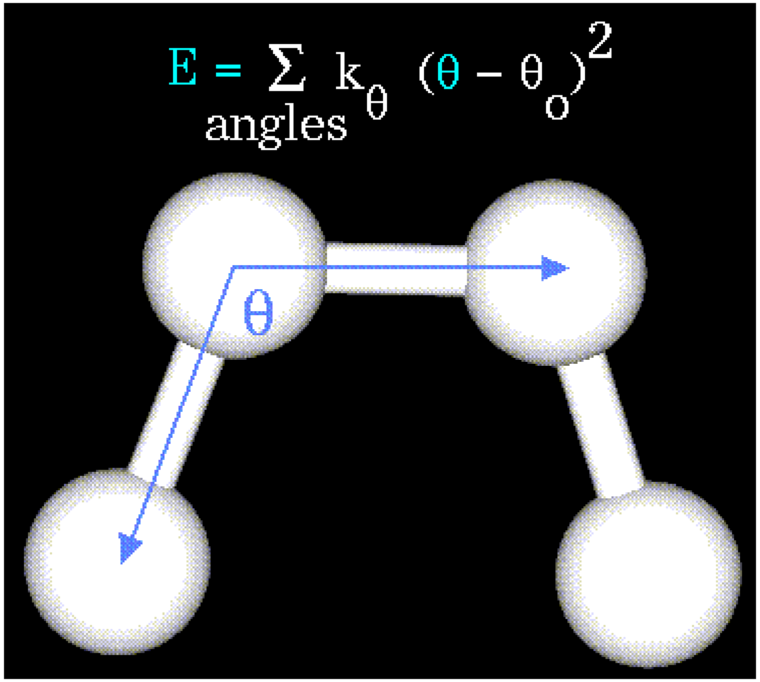
$$E(r) = D_e(1 - e^{-\alpha(r - r_e)})^2 \quad (1)$$

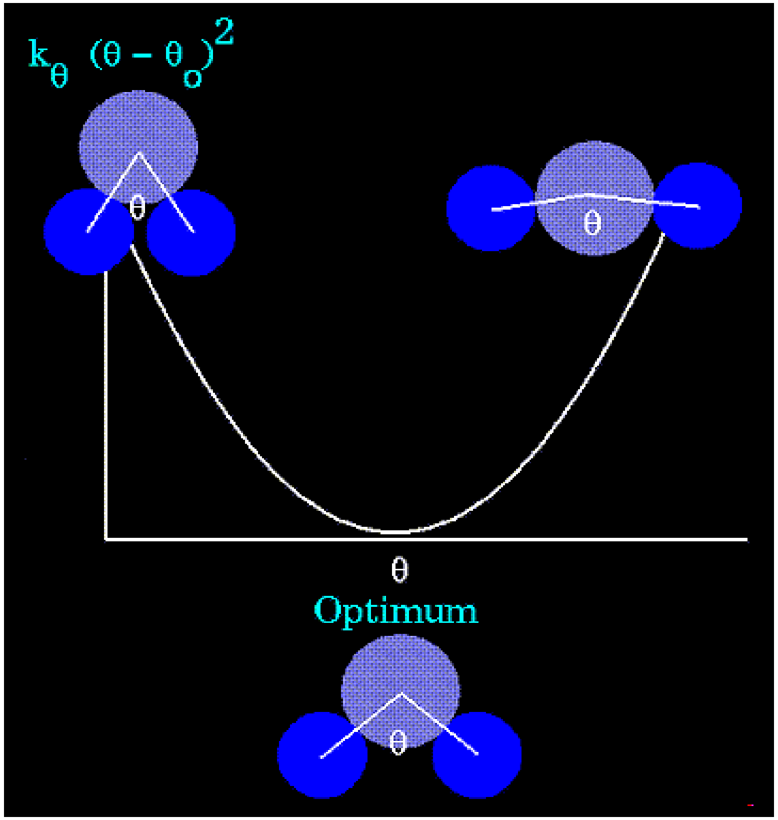
$$\alpha = \omega\sqrt{m/2D_e} \quad (2)$$

$$\omega = \sqrt{k/m} \quad (3)$$

BENDING ENERGY

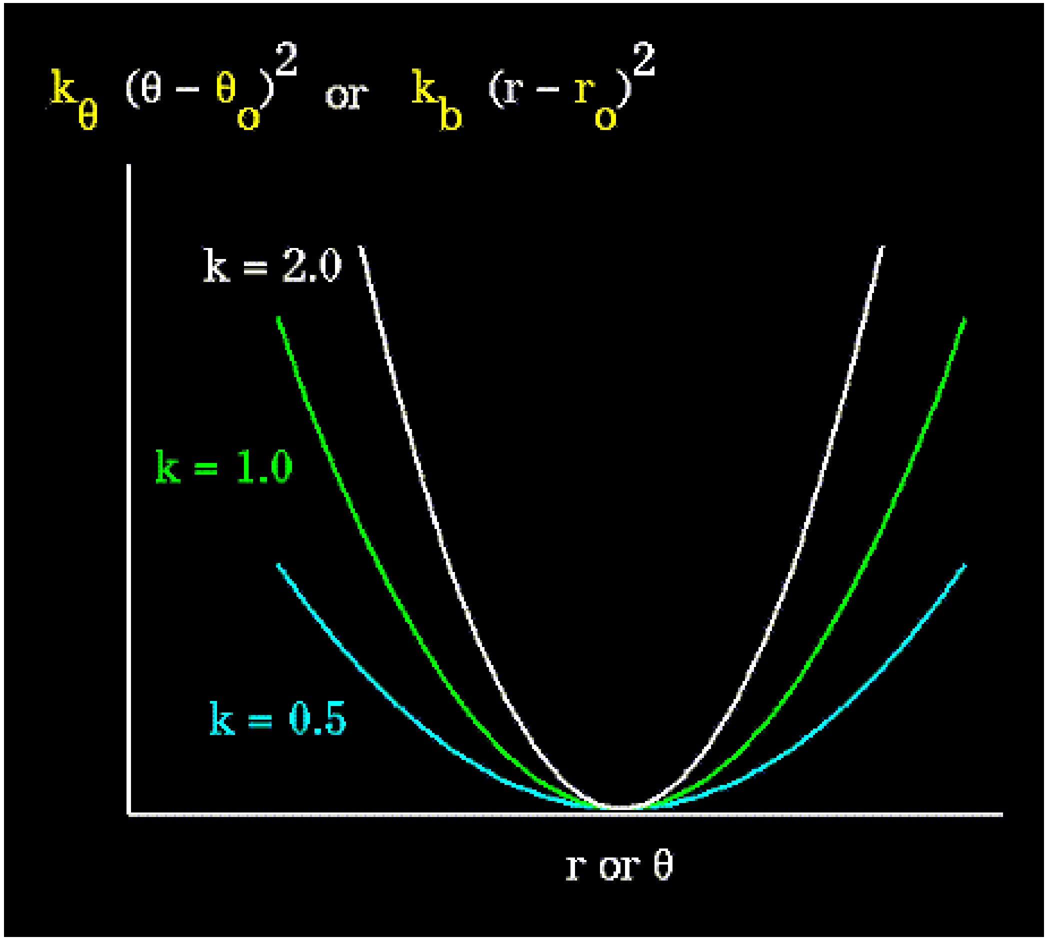
- The bending energy equation is also based on Hooke's law. The "k_θ" parameter controls the stiffness of the angle spring, while "θ_o" defines its equilibrium angle. This equation estimates the energy associated with vibration about the equilibrium bond angle





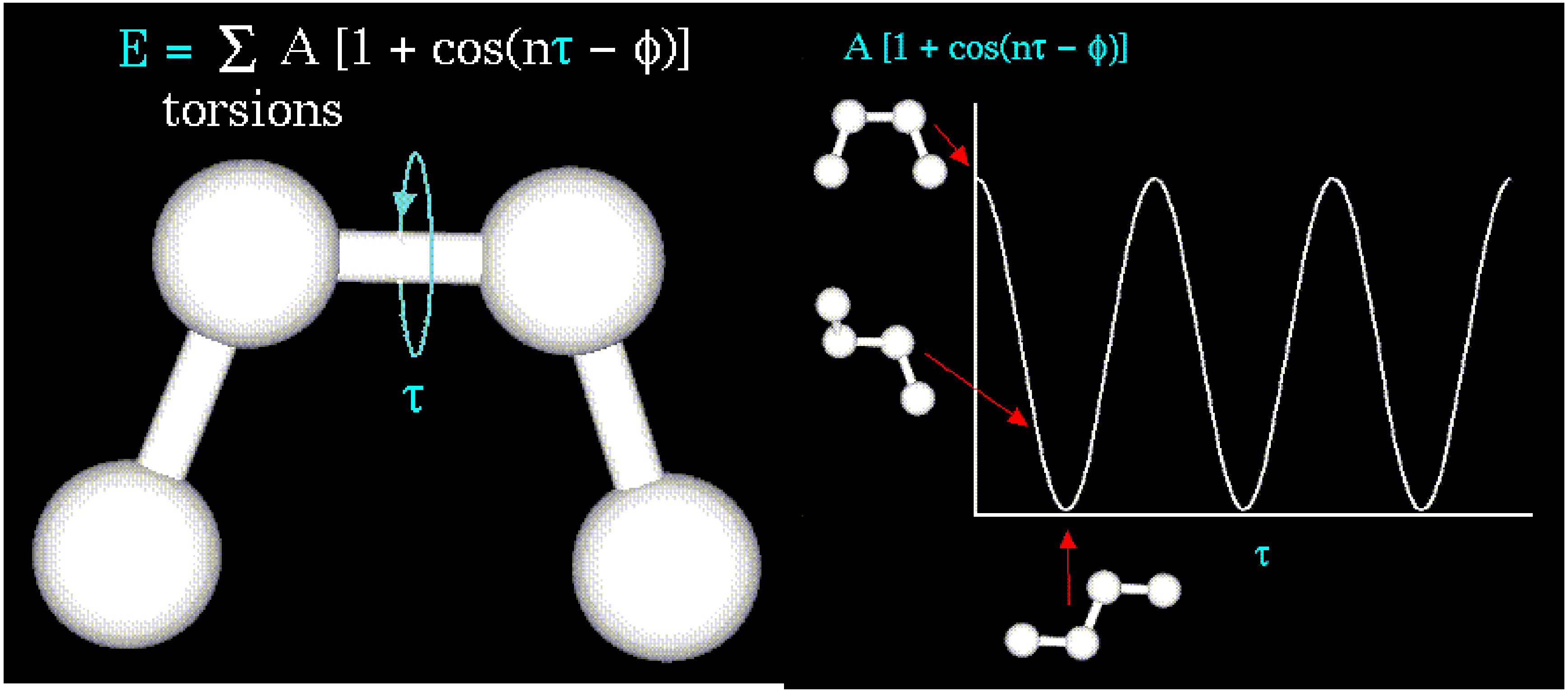
UNIQUE STRETCHING AND BENDING ENERGY

- Unique parameters for angle bending are assigned to each bonded triplet of atoms based on their types (e.g. C-C-C, C-O-C, C-C-H, etc.). The effect of the "k_b" and "k_θ” parameters is to broaden or steepen the slope of the parabola. The larger the value of "k", the more energy is required to deform an angle (or bond) from its equilibrium value. Shallow potentials are achieved for "k" values between 0.0 and 1.0. The Hookeian potential is shown in the following plot for three values of "k



TORSIONAL ENERGY

The torsion energy is modeled by a simple periodic function

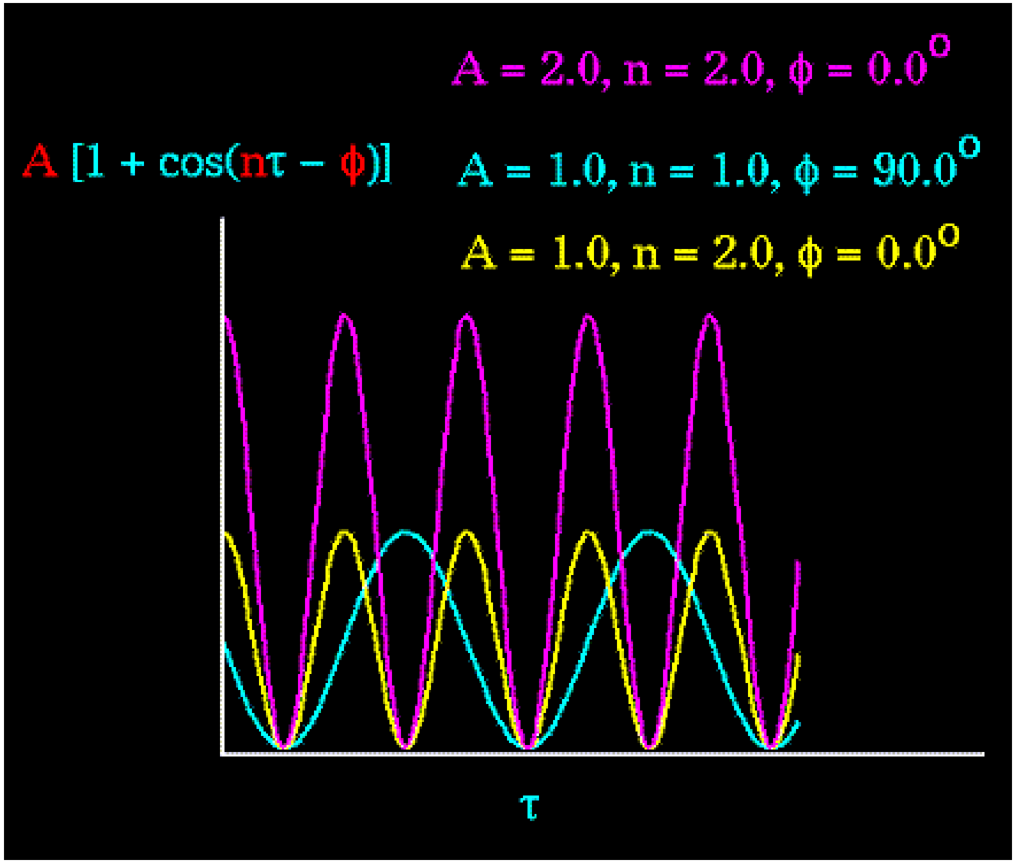


Torsional energy varies during rotation about C-C, C-N and C-O single bonds. The maximum values occur at t=0° and represent "eclipsing" interactions between atoms separated by three sigma bonds.

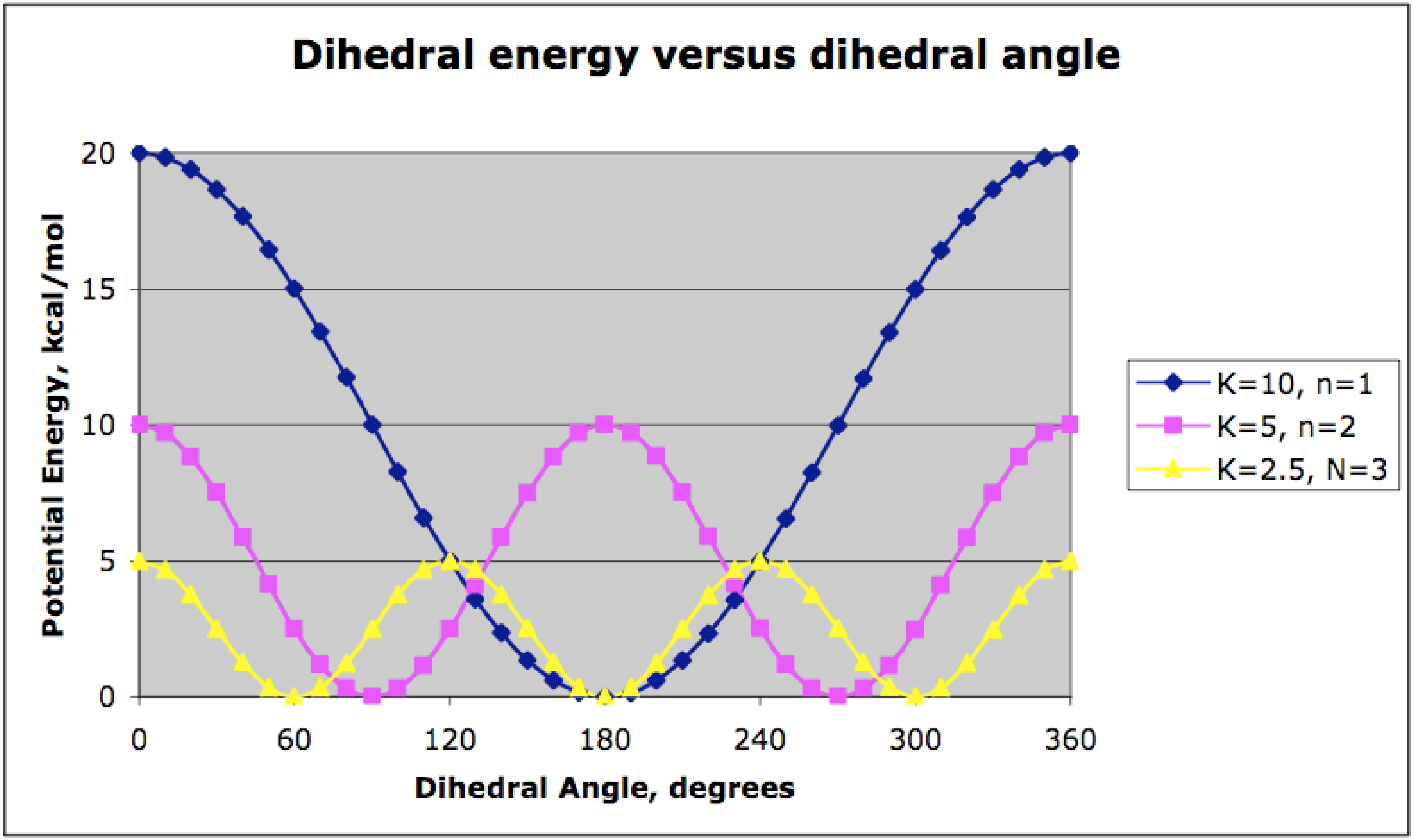
TORSIONAL ENERGY

The torsion energy is modeled by a simple periodic function

The "A" parameter controls the amplitude of the curve, the n parameter controls its periodicity, and "phi" shifts the entire curve along the rotation angle axis (tau). The parameters are determined from curve fitting. Unique parameters for torsional rotation are assigned to each bonded quartet of atoms based on their types (e.g. C-C-C-C, C-O-C-N, H-C-C-H, etc.). Torsion potentials with three combinations of "A", "n", and "phi" are shown



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

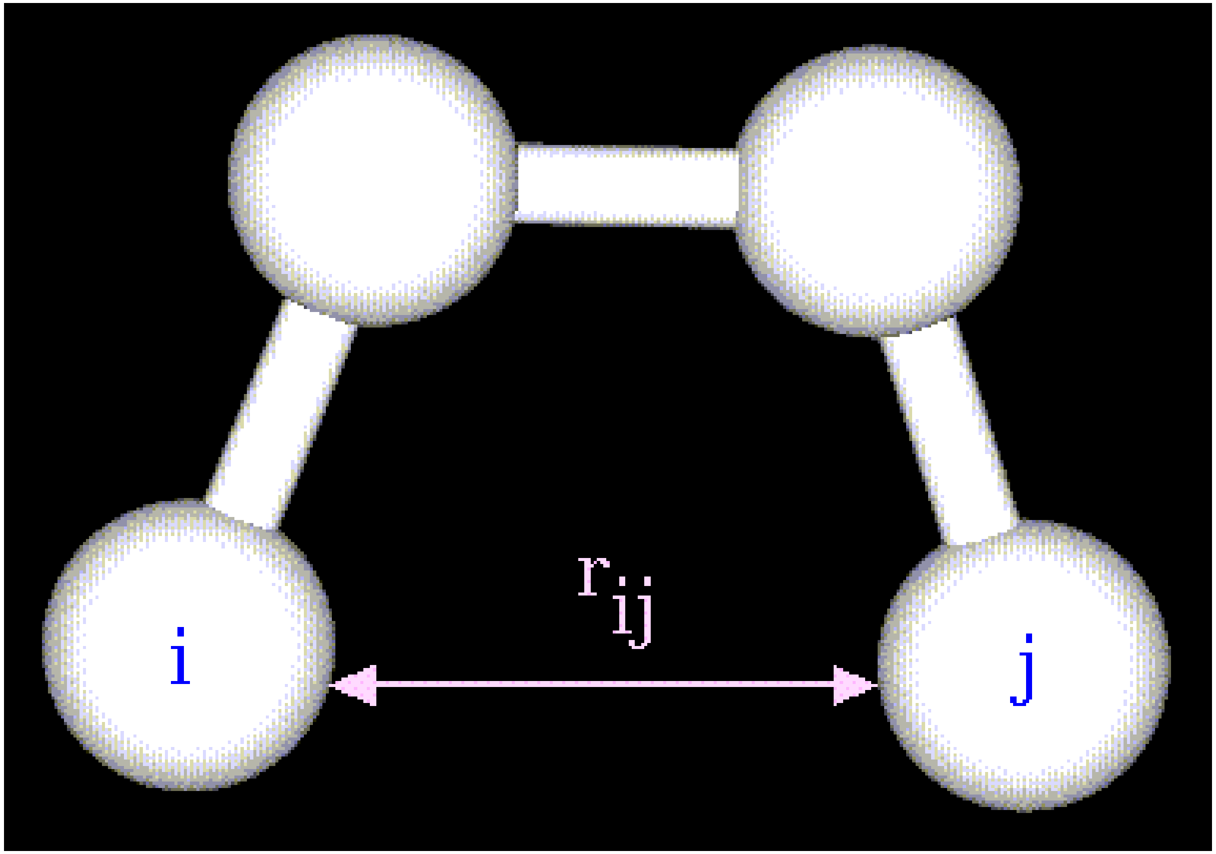


NON-COVALENT (NON-BONDED) TWO ATOM INTERACIONS

The non-bonded energy represents the pair-wise sum of the energies of all possible interacting non-bonded atoms i and j

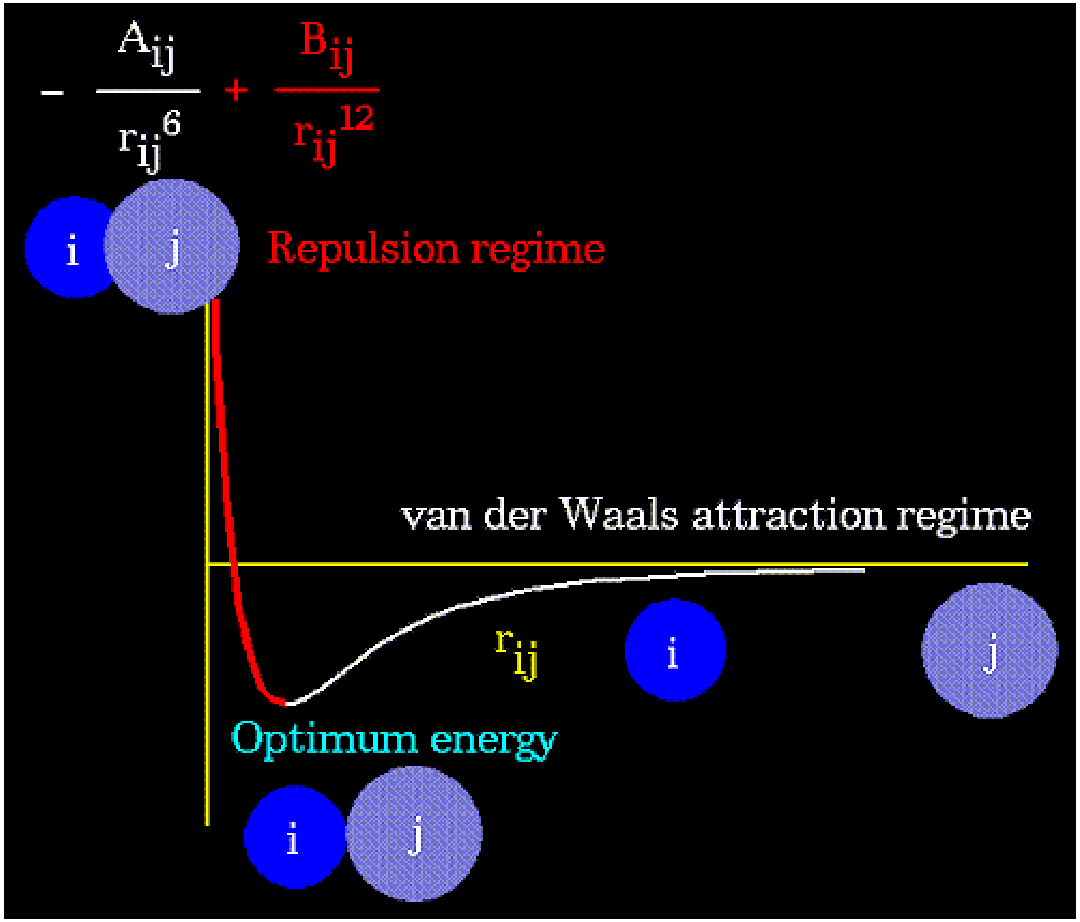
The non-bonded energy accounts for van der Waals attraction, repulsion and electrostatic interactions.

$$E = \underbrace{\sum_i \sum_j \frac{-A_{ij}}{r_{ij}^{6}} + \frac{B_{ij}}{r_{ij}^{12}}}_{\text{van der Waals term}} + \underbrace{\sum_i \sum_j \frac{q_i\, q_j}{r_{ij}}}_{\text{Electrostatic term}}$$



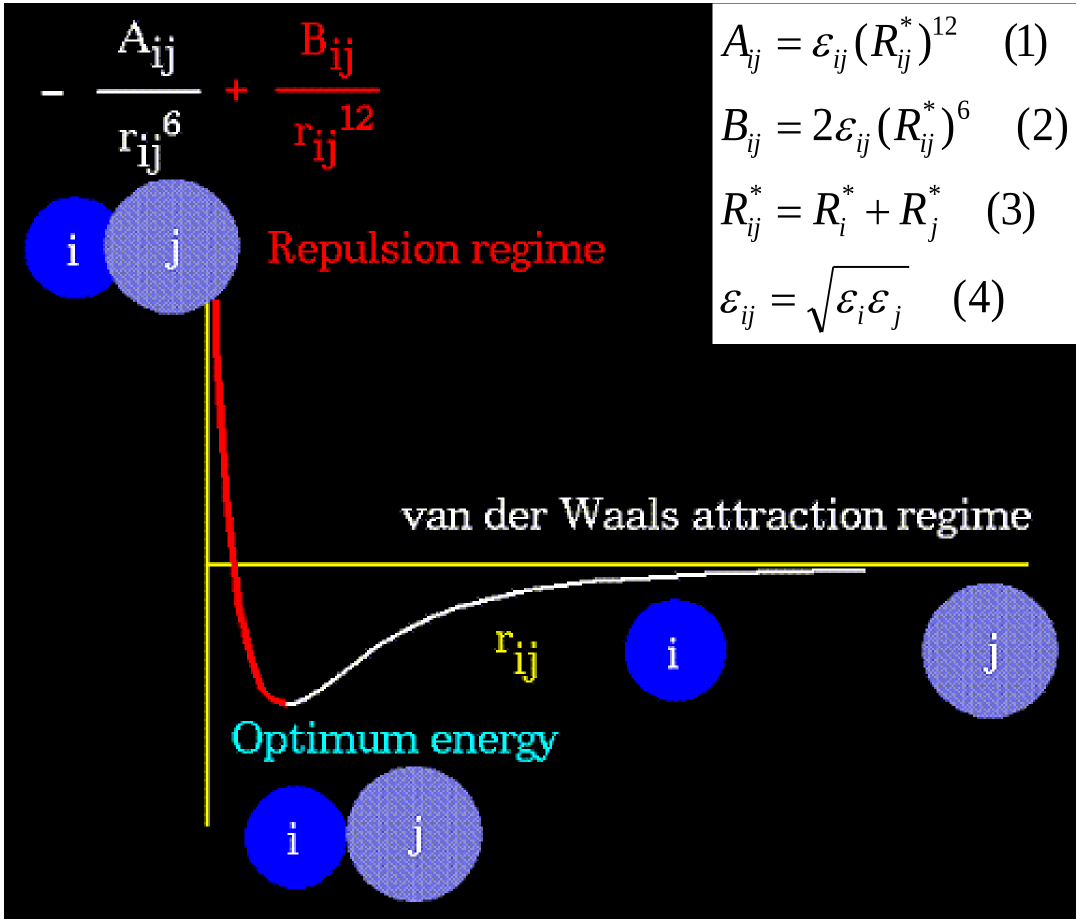
VAN DER WAALS TWO ATOM INTERACIONS

The van der Waals attraction occurs at short range, and rapidly dies off as the interacting atoms move apart by a few Angstroms. Repulsion occurs when the distance between interacting atoms becomes even slightly less than the sum of their contact radii. Repulsion is modeled by an equation that is designed to rapidly blow up at close distances ($1/r^{12}$ dependency). The energy term that describes attraction/repulsion provides for a smooth transition between these two regimes. These effects are often modeled using a 6-12 equation, as shown in the following plot



VAN DER WAALS TWO ATOM INTERACIONS

The "A" and "B" parameters control the depth and position (interatomic distance) of the potential energy well for a given pair of non-bonded interacting atoms (e.g. C:C, O:C, O:H, etc.). In effect, "A" determines the degree of "stickiness" of the van der Waals attraction and "B" determines the degree of "hardness" of the atoms (e.g marshmallow-like, billiard ball-like, etc.



$$-\frac{A_{ij}}{r_{ij}^{6}} + \frac{B_{ij}}{r_{ij}^{12}}$$

$$A_{ij} = \varepsilon_{ij}(R_{ij}^{*})^{12} \quad (1)$$

$$B_{ij} = 2\varepsilon_{ij}(R_{ij}^{*})^{6} \quad (2)$$

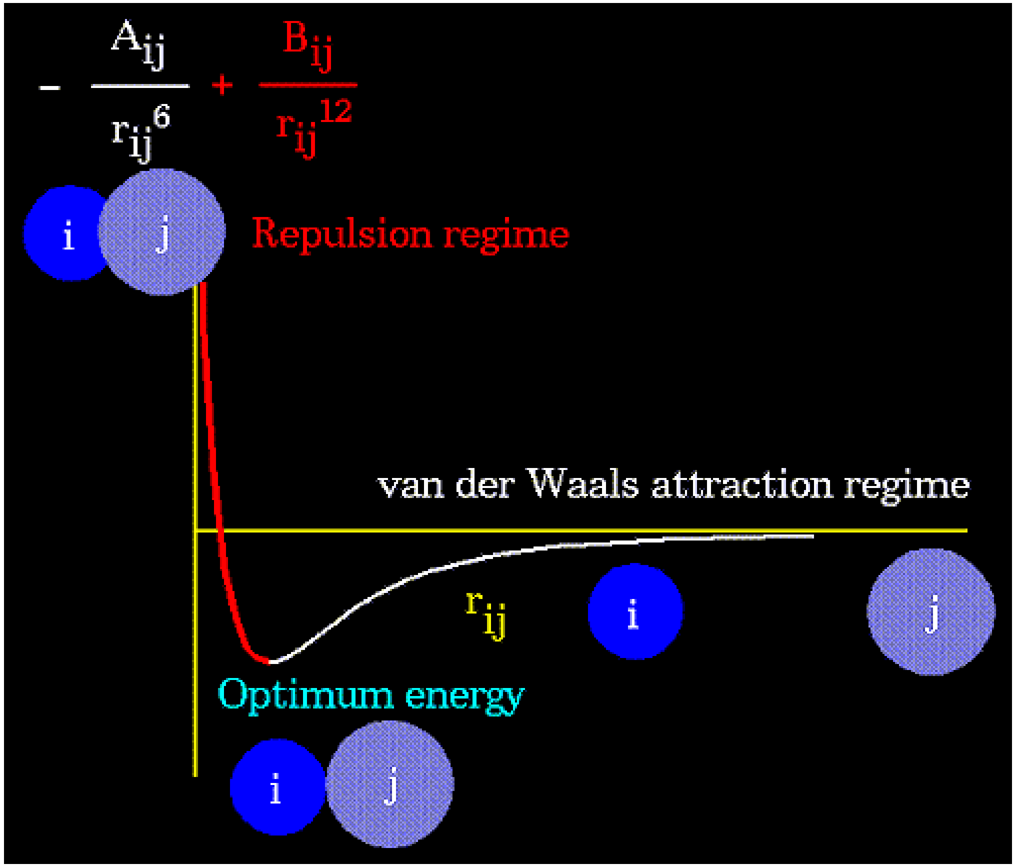
$$R_{ij}^{*} = R_{i}^{*} + R_{j}^{*} \quad (3)$$

$$\varepsilon_{ij} = \sqrt{\varepsilon_{i}\varepsilon_{j}} \quad (4)$$

The "A" parameter can be obtained from atomic polarizability measurements, or it can be calculated quantum mechanically. The "B" parameter is typically derived from crystallographic data so as to reproduce observed average contact distances between different kinds of atoms in crystals of various molecules.

VAN DER WAALS TWO ATOM INTERACIONS

TheThe "A" and "B" parameters control the depth and position (interatomic distance) of the potential energy well for a given pair of non-bonded interacting atoms (e.g. C:C, O:C, O:H, etc.). In effect, "A" determines the degree of "stickiness" of the van der Waals attraction and "B" determines the degree of "hardness" of the atoms (e.g marshmallow-like, billiard ball-like, etc.



The "A" parameter can be obtained from atomic polarizability measurements, or it can be calculated quantum mechanically. The "B" parameter is typically derived from crystallographic data so as to reproduce observed average contact distances between different kinds of atoms in crystals of various molecules.

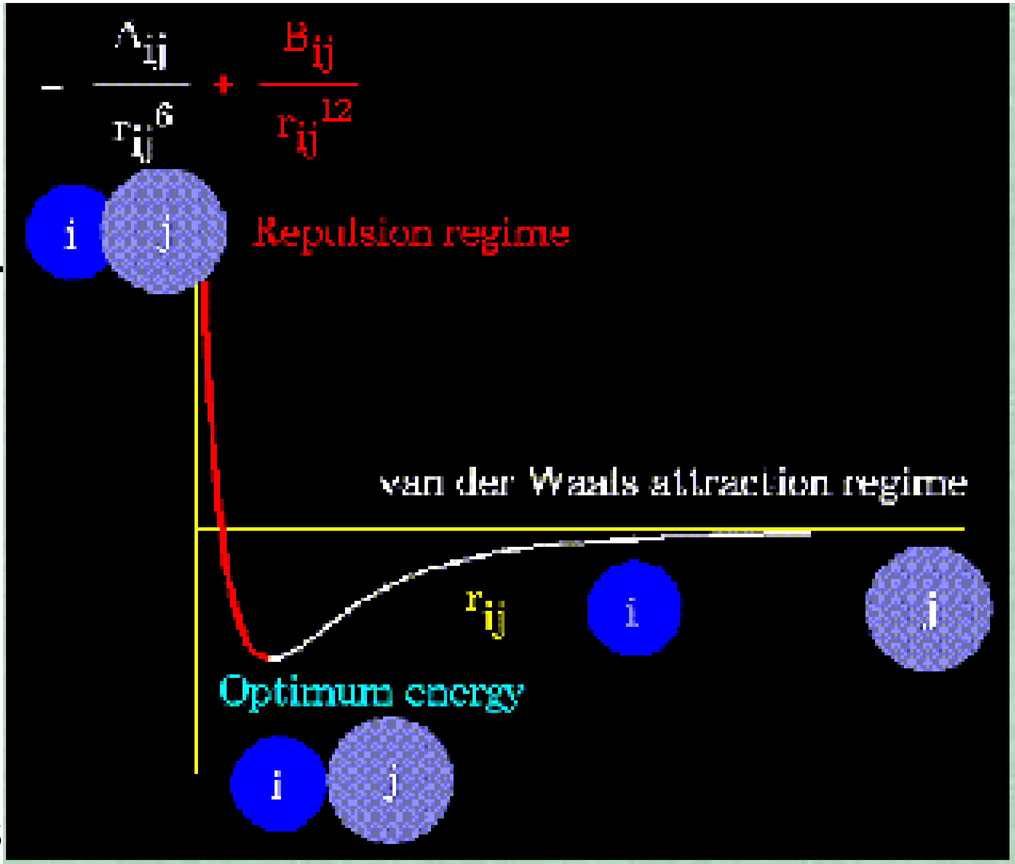
Molecular Mechanics Force Fields:

Basic Interactions and Their Models

- van der Waals attraction occurs at short range, and rapidly dies off as the interacting atoms move apart.

- Repulsion occurs when the distance between interacting atoms becomes even slightly less than the sum of their contact distance.

- Electrostatic energy dies out slowly and it can affect atoms quite far apart.



van der Waals energy as a function of atom-atom separation.

Molecular Mechanics Force Fields:
Basic Interactions and Their Models

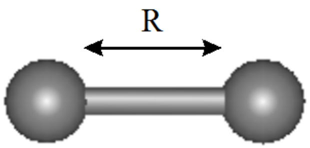
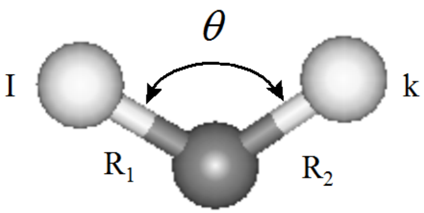
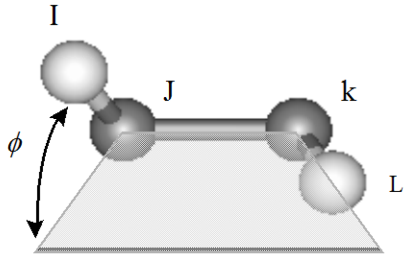
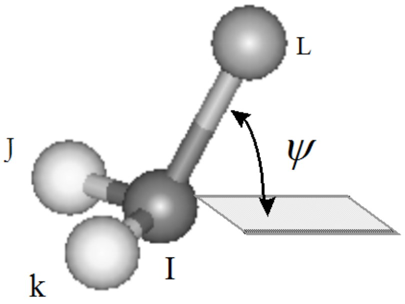
Complete Energy Function:

$$H = \sum_{atoms} \frac{p^2}{2m} + \sum_{bond-stretch} \frac{1}{2} k_r (r - r_{eq})^2 + \sum_{bond-angle-bending} \frac{1}{2} k_\theta (\theta - \theta_{eq})^2 +$$

$$\sum_{bond-rotation} \frac{v_n}{2} [1 + \cos(n\phi - \gamma)] + \sum_{S-bond} [V_0 (1 - e^{-a(r - r_0')})^2 - V_0] +$$

$$\sum_{H-bond} [V_0 (1 - e^{-a(r - r_0')})^2 - V_0] + \sum_{non-bonded} [\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} + \frac{q_i q_j}{\varepsilon_{ij} r_{ij}}]$$

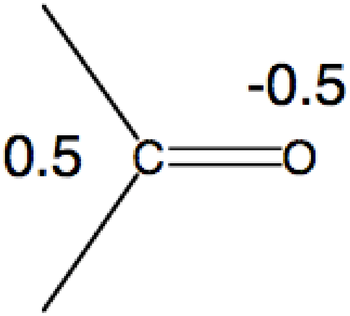
Valence Force Field

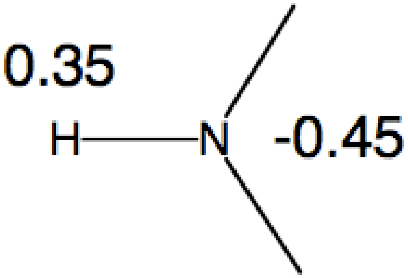
		Typical Expressions	
Description	Ilustration	Potential	Forces
Bond Stretch		$U_r = \frac{1}{2} K_b (R - R_0)^2$	$-\frac{\partial U}{\partial R} = K_b (R - R_0)$
Bond Angle bend		$U_\theta = \frac{1}{2} C (\cos\theta - \cos\theta_0)^2$	$-\frac{\partial U}{\partial \theta} = C(\cos(\theta) - \cos(\theta_0))\sin(\theta)$
Torsion		$U_\phi = \sum_{n=1}^{p} \frac{1}{2} K_{\theta,n} [1 - d\cos(n\phi)]$	$-\frac{\partial U}{\partial R} = \frac{1}{2} K_b \cdot d \cdot n \cdot \sin(n\theta)$
Inversion		$U_\psi = \frac{1}{2} C (\cos\psi - \cos\psi_0)^2$	$-\frac{\partial U}{\partial \psi} = C(\cos(\psi) - \cos(\psi_0))\sin(\psi)$

Hydrogen Bond Energy

Treatment of hydrogen bonds???

Partial atomic charges



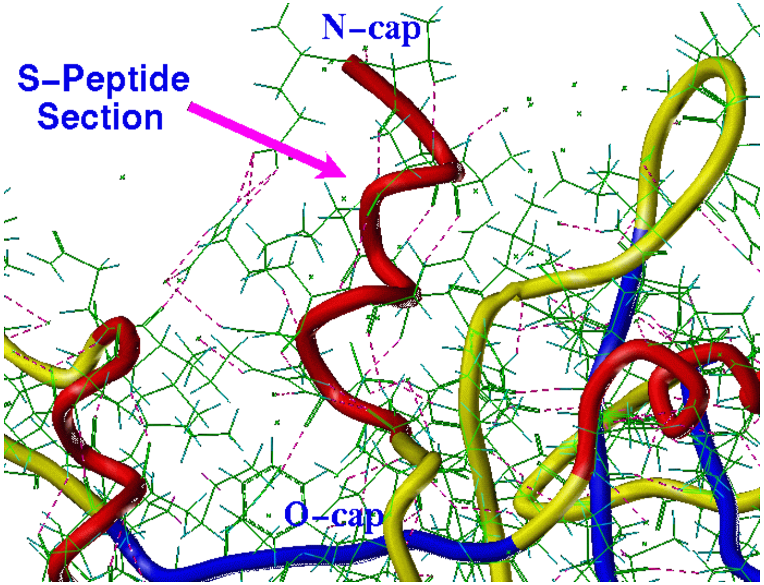


Molecular Mechanics Force Fields:

Basic Interactions and Their Models

Types of Hydrogen Bond:

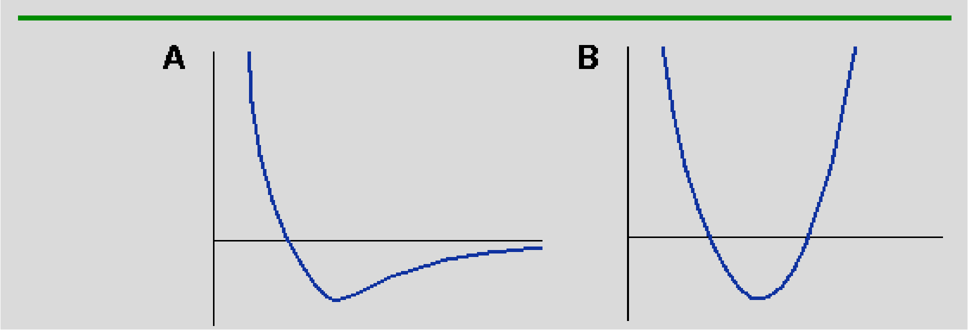
N-H … O
N-H … N
O-H … N
O-H … O



Can be modeled by

- VdW+electrostatic (AMBER)
- Modified Linard-Jones (CHARM)
- Morse potential (Prohofsky/Chen)

Morse vs Harmonic Bond term:



$V_H(r) = A/r^{12} - B/r^{6} + q_i q_j/\varepsilon_r r_{ij}$ for AMBER

$= (A/r^{12} - B/r^{10})\cos^m(\theta_{A\text{-}H\text{-}D})\cos^n(\theta_{AA\text{-}A\text{-}H})sw_1(\mathbf{r})sw_2(\theta)$ for CHARM

$= V_0(1-e^{-a(r-r0)})^2 - V_0$ for Prohofsky/Chen

Popular MMFF

Popular molecular mechanics force fields

Classical

AMBER (Assisted Model Building and Energy Refinement) - widely used for proteins and DNA

CHARMM - originally developed at Harvard, widely used for both small molecules and macromolecules

CHARMm - commercial version of CHARMM, available through Accelrys

CVFF - also broadly used for small molecules and macromolecules

GROMACS - The force field optimized for the package of the same name

GROMOS - A force field that comes as part of the GROMOS (GROningen MOlecular Simulation package), a general-purpose molecular dynamics computer simulation package for the study of biomolecular systems. GROMOS force field (A-version) has been developed for application to aqueous or apolar solutions of proteins, nucleotides and sugars. However, a gas phase version (B-version) for simulation of isolated molecules is also available

OPLS-aa, OPLS-ua, OPLS-2001, OPLS-2005 - Members of the OPLS family of force fields developed by William L. Jorgensen at Yale Department of Chemistry.

ECEPP/2 - free energy force field

Molecular Mechanics, Force Fields

Some Available Force Fields:

- CFF: (Consistent Force Field) Warshel, Lifson *et al.*; wide variety of experimental data, software for fitting force field parameters, parametrised to organic compounds, polymers, metals.
- MMFF: derived from both experimental and *ab initio* data, including HF and MP2 energies of torsion sampled structures and conformations.
- MM2/MM3/MM4: Allinger *et al.*; parametrised to heats of formations and small molecule gas phase data (particularly structures and conformational energies). Primarily for geometry optimization and prediction of thermodynamic values and IR spectral. MM3 and MM4 include hydrogen bonding.

Molecular Mechanics Force Fields For Molecular Simulations

- AMBER (Assisted Model Building with Energy Refinement) is the name of both a family of force fields developed for biomolecules by Peter Kollman, and a program for implementing them. AMBER uses harmonic stretches and bends, a cosine function for torsions, a Coulomb electrostatic interaction and a 12-6 Lennard-Jones van der Waals interaction. AMBER has been designed primarily for proteins and nucleic acids.
- CHARMM (Chemistry at HARvard Macromolecular Mechanics) is also a family of force fields and a program. CHARMM has all-atom and united atom variants and is widely used for drug molecules and macromolecules. One variant also includes the TIP3P force field for water, allowing it to be used as an explicit solvent.

Molecular Mechanics, Force Fields

- The GROMOS (GROningen MOlecular Simulation computer program package) force field and package were developed for biomolecular systems at the University of Groningen and at ETH in Zurich. GROMOS uses a united atom approach to fragments within biomolecules. There are both aqueous and gas phase versions.
- GROMACS (GROningen MAchine for Chemical Simulations) is the free molecular simulation “engine” that has grown out of GROMOS and can also support most of the other available force fields. Indeed AMBER, CHARM and GROMOS were all primarily developed for molecular dynamics.

Molecular Mechanical Force Field Parameterization

Molecular Mechanical Force Field Parameterization

1. Why is MMFF important?

- Inaccurate forces => non-physical motions
- Central to free energy calculations, molecular docking, etc.

2. What are characteristics of a good MMFF?

- Transferability (not specially tailored to each molecule)
- Accuracy
- Chemical space (wide applicability)

3. Why is it extremely difficult to develop MMFF ?

- Force field parameter space is extremely rough
- Large chemical space (10^{60} molecules)
- Many chemical functional groups (> 100)
- Need to model numerous molecular properties at various conditions (temperature, pressure, etc.)

Human intervention is inevitable in force field parameterization.

Molecular Mechanics Force Fields:

Parameterization

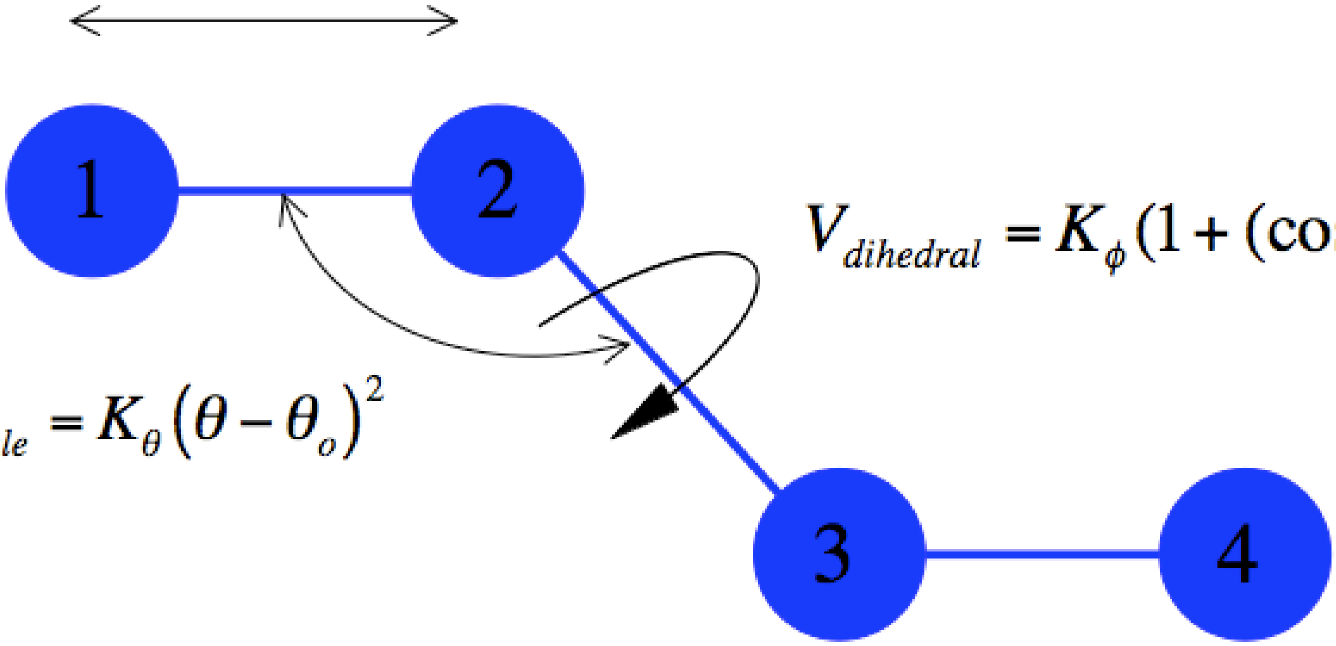
- In addition to the functional form of the potentials, a force field typically defines a set of parameters for each of a number of atom or particle types that correspond to different atoms and bonding patterns in commonly simulated molecules.

- The parameter set includes values for atomic mass and partial charge for individual atoms, and equilibrium bond lengths and angles for pairs, triplets, and quadruplets of bonded atoms.

- Preparation for a molecular dynamics simulation involves assigning an atom or particle type to each atom or particle in the molecules of interest.

- Although many molecular simulations involve biological macromolecules such as proteins, DNA, and RNA, the parameters for given atom types are generally derived from observations on small organic molecules that are more tractable for experimental study and quantum calculation.

Diagram of intramolecular energy terms

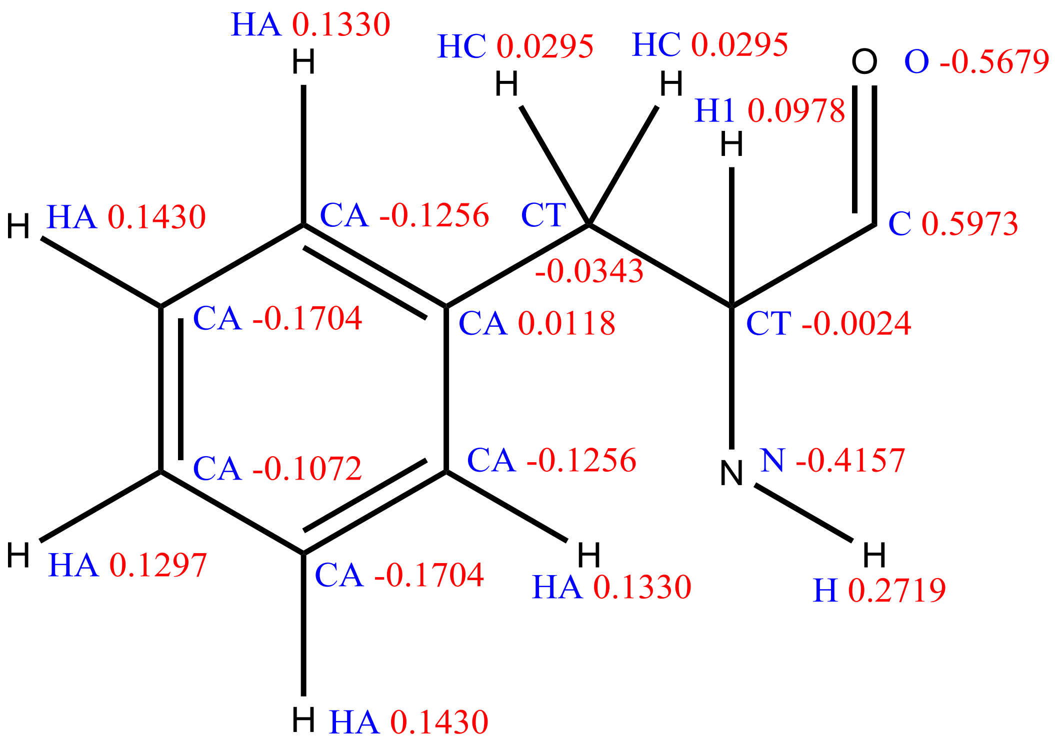


$$V_{bond} = K_b\left(b - b_o\right)^2$$

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

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

AMBER Force Field For Phenylalanine Residue



Atom Types: 9
N H CT H1 HC HA C O

Bond Parameters (A-B): 10
For example CA-HA
$k_r = 367.0$, $r_{eq} = 1.080$

Angle Parameters (A-B-C): 15
For example CA-CA-CA
$k_\theta = 63.0$, $\theta_{eq} = 120.0$

Dihedral Parameters (A-B-C-D): 21
For example CA-CA-CA-CA
$V_2 = 3.625$, $n = 2$, $\gamma = 180.0$

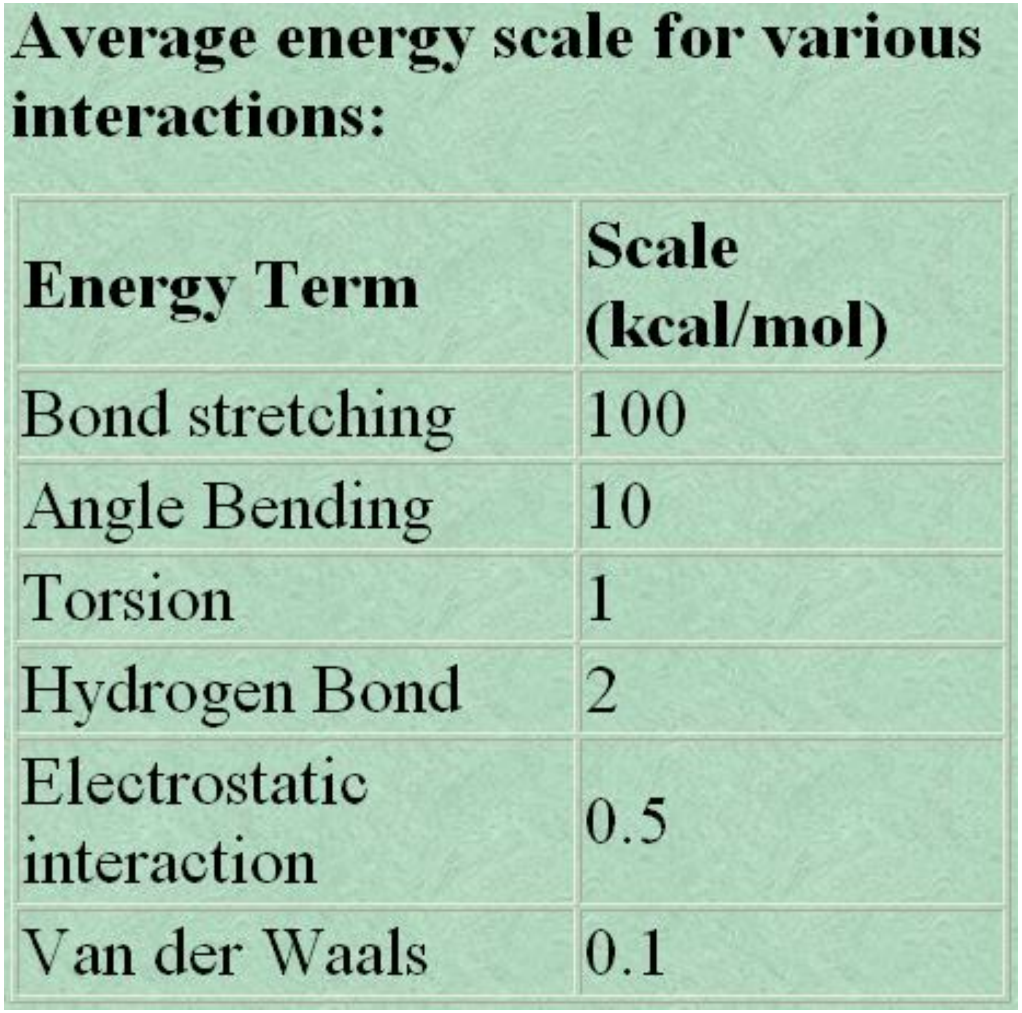
A, **B**, **C** and **D** are **atom types.**

$$V_{potential} = \sum_{\text{bonds}} K_r \left(r - r_{eq}\right)^2 + \sum_{\text{angles}} K_\theta \left(\theta - \theta_{eq}\right)^2 + \sum_{\text{dihedrals}} \frac{V_n}{2}\left[1 + \cos\left(n\varphi - \gamma\right)\right] + \sum_{\substack{i<j \\ \text{van der Waals}}} \left(\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}}\right) + \sum_{\substack{i<j \\ \text{electrostatic}}} \frac{q_i q_j}{R_{ij}}$$

Molecular Mechanics Force Fields:

Basic Interactions and Their Models

Concept of energy scale is Important for molecular Modeling



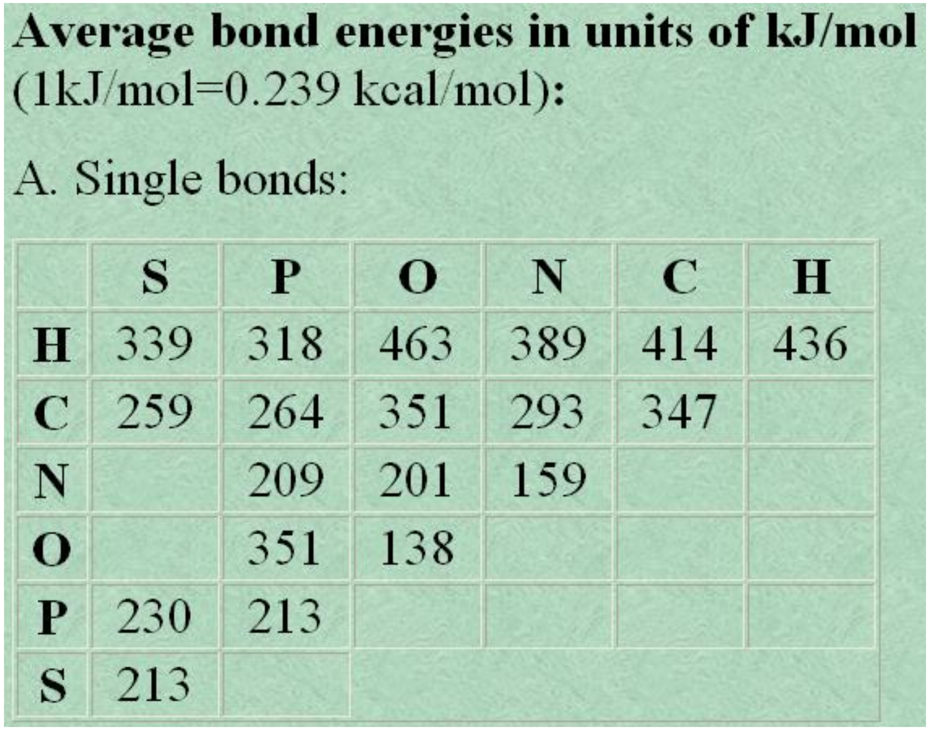
Average energy scale for various interactions:

Energy Term	**Scale (kcal/mol)**
Bond stretching	100
Angle Bending	10
Torsion	1
Hydrogen Bond	2
Electrostatic interaction	0.5
Van der Waals	0.1

Molecular Mechanics Force Fields:

Basic Interactions and Their Models

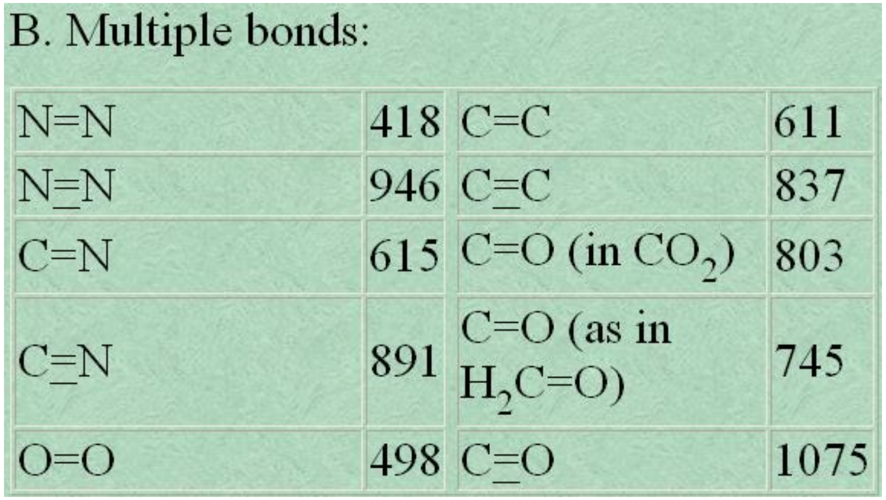
Concept of energy scale is Important for molecular modeling



Average bond energies in units of kJ/mol (1kJ/mol=0.239 kcal/mol):

A. Single bonds:

	S	**P**	**O**	**N**	**C**	**H**
H	339	318	463	389	414	436
C	259	264	351	293	347	
N		209	201	159		
O		351	138			
P	230	213				
S	213					



B. Multiple bonds:

N=N	418	C=C	611
N≡N	946	C≡C	837
C=N	615	C=O (in CO_2)	803
C≡N	891	C=O (as in $H_2C=O$)	745
O=O	498	C≡O	1075

An iterative approach is required to obtain self-consistent parameters

Intramolecular ⟷ Intermolecular

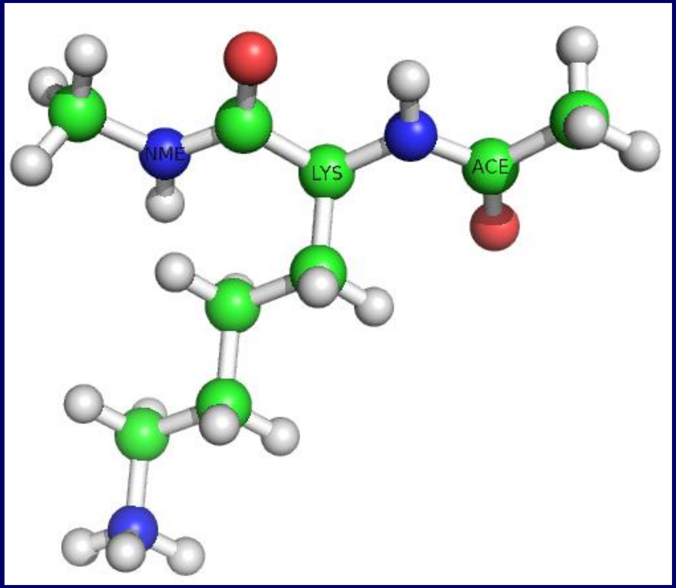
The nonbond/intermolecular parameters will impact the resulting geometries, vibrations and conformational energies. Thus, it is necessary to apply an iterative approach

Philosophy of AMBER Force Field Parameterization

- **Model Compounds**
  Building blocks of macromolecules
  Small, representative

- **Reference Data**
  QM data: optimized geometries, conformational energies, interaction energies, electric moments, electrostatic potentials, electron densities

  Experimental data: crystal/NMR structures, vibrational frequencies, pure liquid/solid properties, solvation free energies, NMR data such as J-J couplings, order parameters



NME-LYS-ACE

Strategies of AMBER Force Field Parameterization

1. Partial Charge

   HF/6-31G* RESP (derive charges to reproduce *ab initio* ESP)

2. Van der Waals

   Reproduce bulk properties, such as density and heat of vaporization, hydration free energies

3. Bond length and bond angle parameters

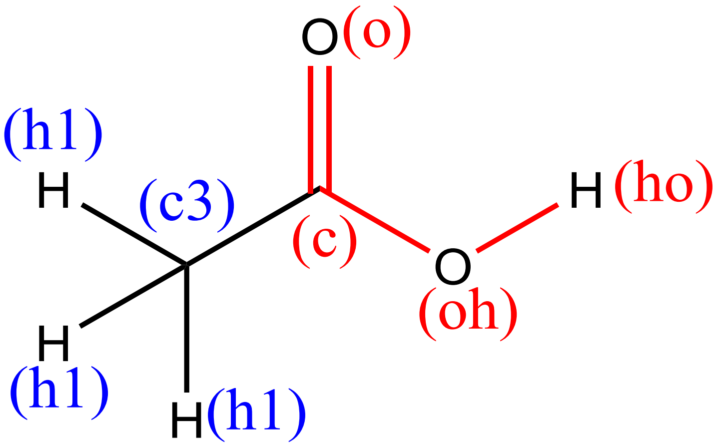
   Experiments or high-level *ab initio* calculations

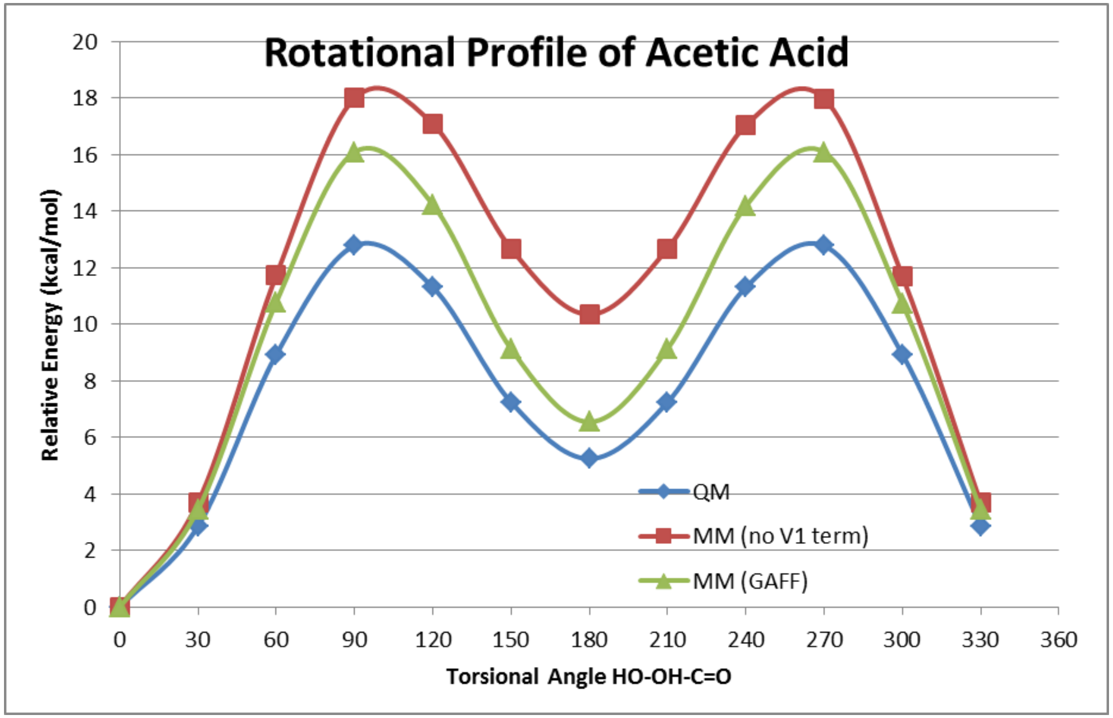
4. Torsional angle parameters

   Experimental or high-level *ab initio* relative energies and rotational profiles

Torsional Angle Parameterization: An Example

$$V_{tor} = \sum_{n=1}^{N} \frac{V_n}{2}(1+\cos(n\phi-\gamma))$$





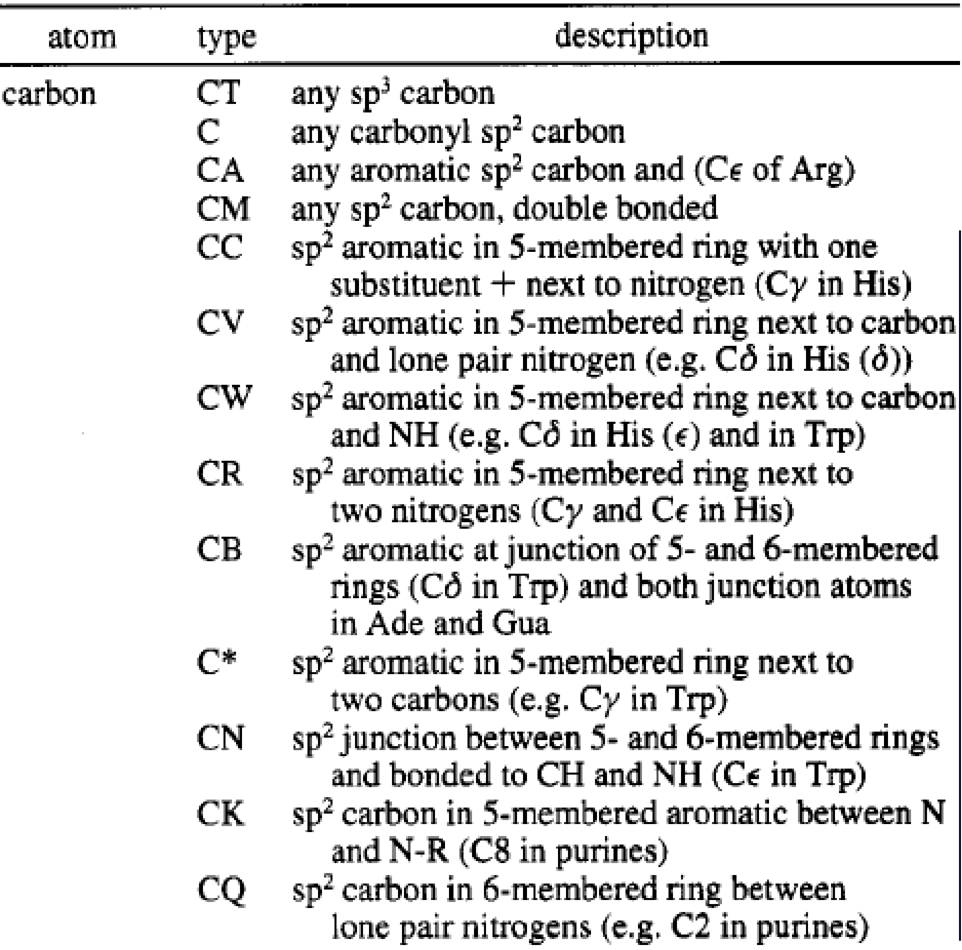
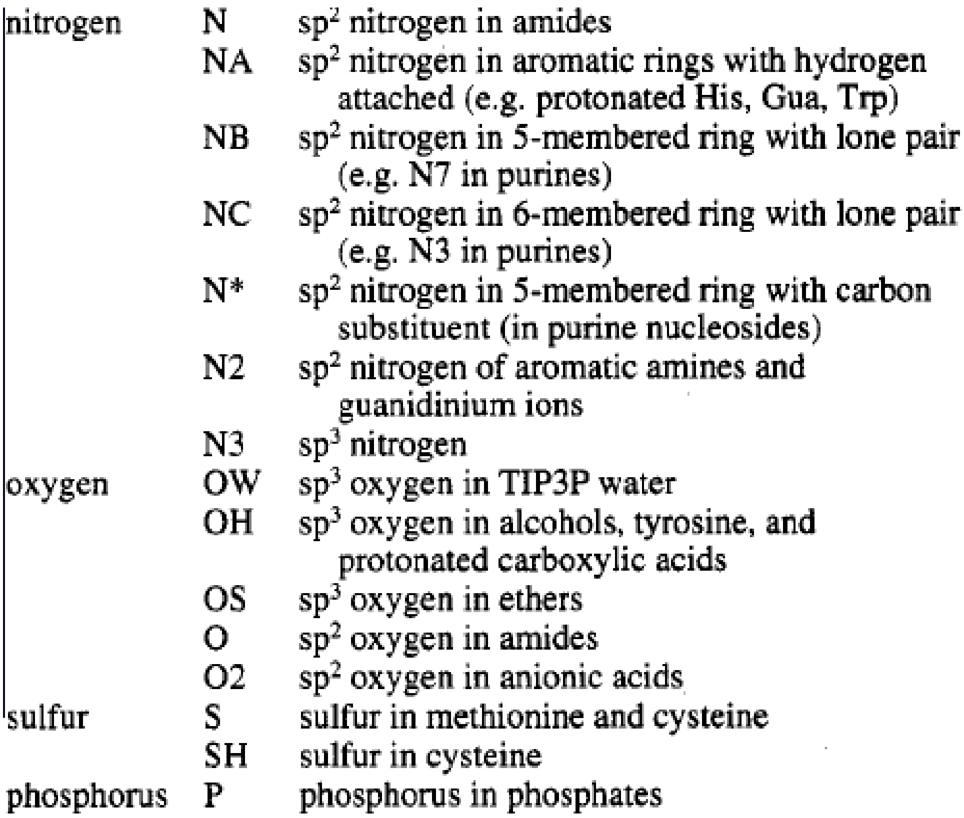
Torsional parameter to be optimized: **ho-oh-c=o**

V_2 =2.3, γ= 180
V_1 =1.9, γ= 0.0

Density at 25 ^{o}C:

Exp: 1.0446
GAFF: 1.1282
GAFF2: 1.0635

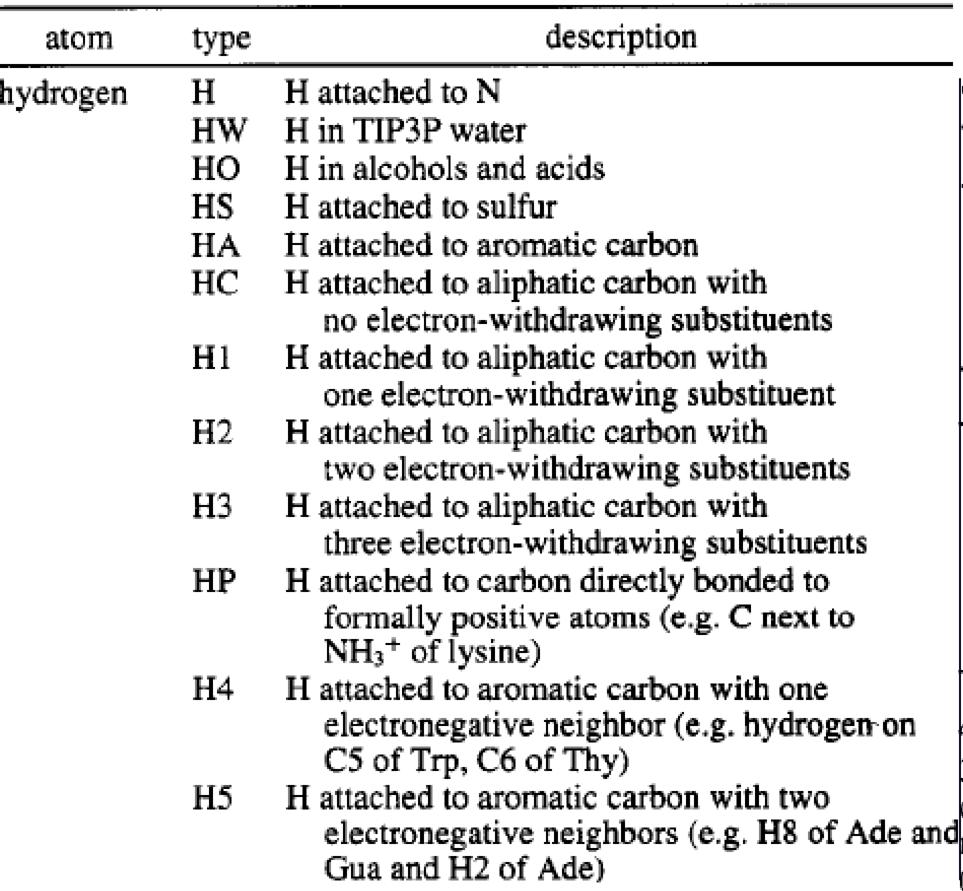
Molecular Mechanics Force Fields:
Atom Types

atom	type	description
carbon	CT	any sp^3 carbon
	C	any carbonyl sp^2 carbon
	CA	any aromatic sp^2 carbon and (Cϵ of Arg)
	CM	any sp^2 carbon, double bonded
	CC	sp^2 aromatic in 5-membered ring with one substituent + next to nitrogen (Cγ in His)
	CV	sp^2 aromatic in 5-membered ring next to carbon and lone pair nitrogen (e.g. Cδ in His (δ))
	CW	sp^2 aromatic in 5-membered ring next to carbon and NH (e.g. Cδ in His (ϵ) and in Trp)
	CR	sp^2 aromatic in 5-membered ring next to two nitrogens (Cγ and Cϵ in His)
	CB	sp^2 aromatic at junction of 5- and 6-membered rings (Cδ in Trp) and both junction atoms in Ade and Gua
	C*	sp^2 aromatic in 5-membered ring next to two carbons (e.g. Cγ in Trp)
	CN	sp^2 junction between 5- and 6-membered rings and bonded to CH and NH (Cϵ in Trp)
	CK	sp^2 carbon in 5-membered aromatic between N and N-R (C8 in purines)
	CQ	sp^2 carbon in 6-membered ring between lone pair nitrogens (e.g. C2 in purines)
nitrogen	N	sp^2 nitrogen in amides
	NA	sp^2 nitrogen in aromatic rings with hydrogen attached (e.g. protonated His, Gua, Trp)
	NB	sp^2 nitrogen in 5-membered ring with lone pair (e.g. N7 in purines)
	NC	sp^2 nitrogen in 6-membered ring with lone pair (e.g. N3 in purines)
	N*	sp^2 nitrogen in 5-membered ring with carbon substituent (in purine nucleosides)
	N2	sp^2 nitrogen of aromatic amines and guanidinium ions
	N3	sp^3 nitrogen
oxygen	OW	sp^3 oxygen in TIP3P water
	OH	sp^3 oxygen in alcohols, tyrosine, and protonated carboxylic acids
	OS	sp^3 oxygen in ethers
	O	sp^2 oxygen in amides
	O2	sp^2 oxygen in anionic acids
sulfur	S	sulfur in methionine and cysteine
	SH	sulfur in cysteine
phosphorus	P	phosphorus in phosphates

AMBER: J. Am. Chem. Soc. 117, 5179-5197

Molecular Mechanics Force Fields:
Atom Types and Standard Parameters



atom	type	description
hydrogen	H	H attached to N
	HW	H in TIP3P water
	HO	H in alcohols and acids
	HS	H attached to sulfur
	HA	H attached to aromatic carbon
	HC	H attached to aliphatic carbon with no electron-withdrawing substituents
	H1	H attached to aliphatic carbon with one electron-withdrawing substituent
	H2	H attached to aliphatic carbon with two electron-withdrawing substituents
	H3	H attached to aliphatic carbon with three electron-withdrawing substituents
	HP	H attached to carbon directly bonded to formally positive atoms (e.g. C next to NH_3^+ of lysine)
	H4	H attached to aromatic carbon with one electronegative neighbor (e.g. hydrogen on C5 of Trp, C6 of Thy)
	H5	H attached to aromatic carbon with two electronegative neighbors (e.g. H8 of Ade and Gua and H2 of Ade)

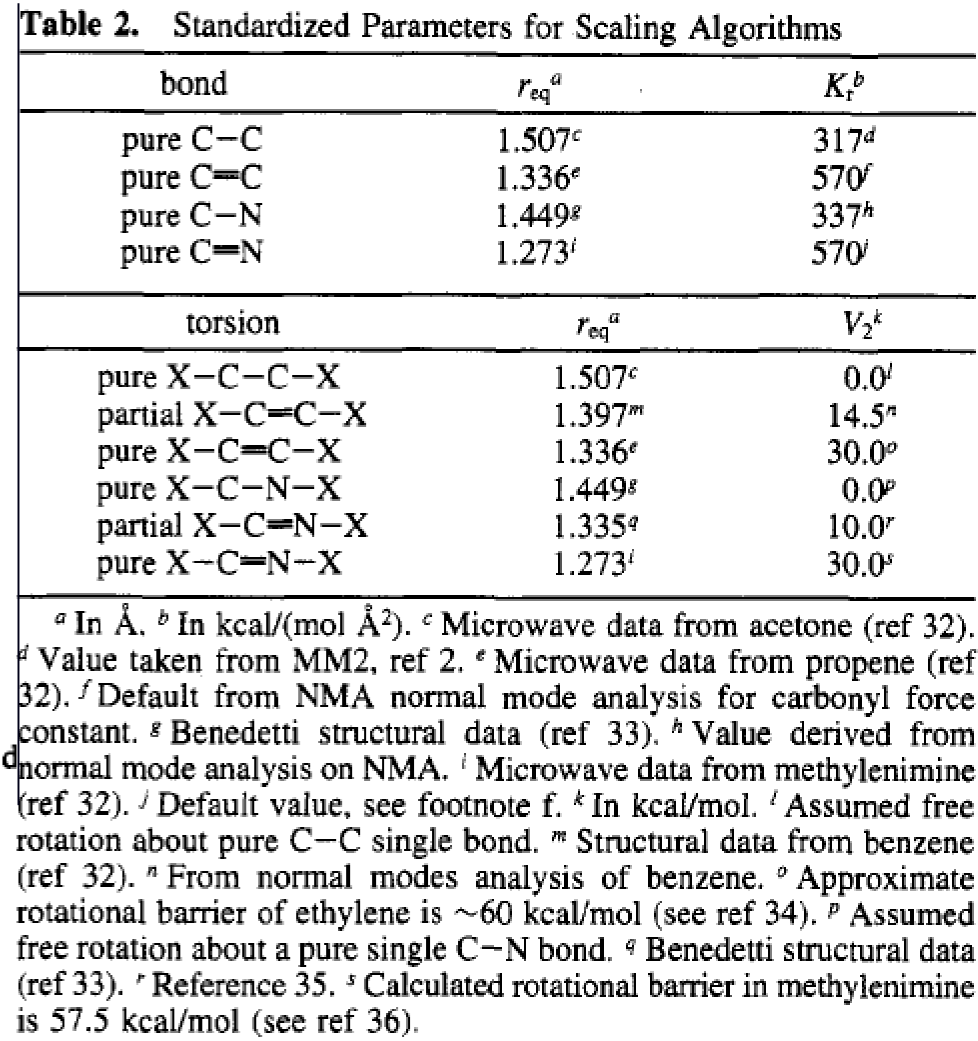


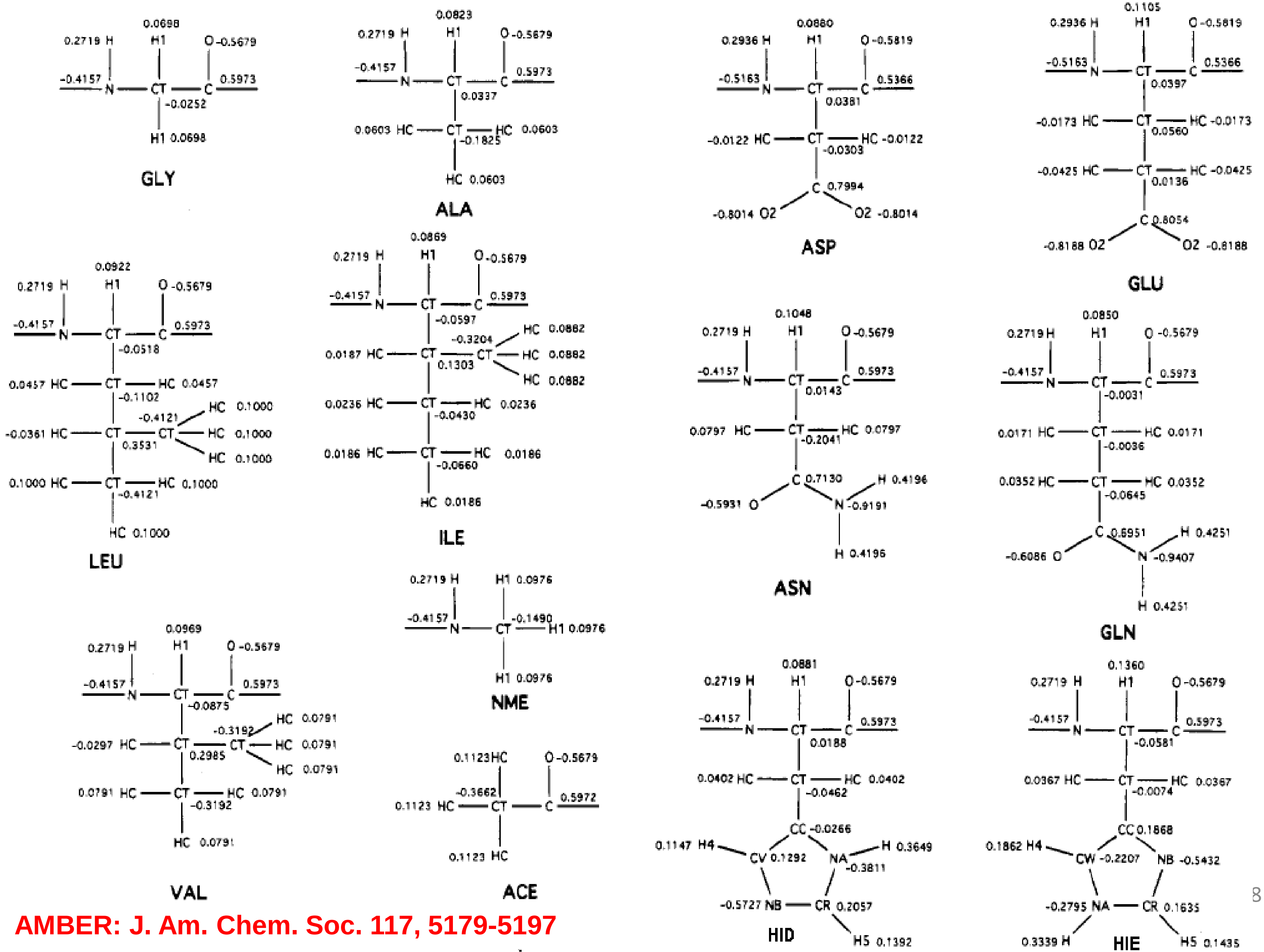
Table 2. Standardized Parameters for Scaling Algorithms

bond	r_{eq}[a]	K_r[b]
pure C−C	1.507[c]	317[d]
pure C=C	1.336[e]	570[f]
pure C−N	1.449[g]	337[h]
pure C=N	1.273[i]	570[j]
torsion	**r_{eq}[a]**	**V_2[k]**
pure X−C−C−X	1.507[c]	0.0[l]
partial X−C=C−X	1.397[m]	14.5[n]
pure X−C=C−X	1.336[e]	30.0[o]
pure X−C−N−X	1.449[g]	0.0[p]
partial X−C=N−X	1.335[q]	10.0[r]
pure X−C=N−X	1.273[i]	30.0[s]

[a] In Å. [b] In kcal/(mol Å^2). [c] Microwave data from acetone (ref 32). [d] Value taken from MM2, ref 2. [e] Microwave data from propene (ref 32). [f] Default from NMA normal mode analysis for carbonyl force constant. [g] Benedetti structural data (ref 33). [h] Value derived from normal mode analysis on NMA. [i] Microwave data from methylenimine (ref 32). [j] Default value, see footnote f. [k] In kcal/mol. [l] Assumed free rotation about pure C−C single bond. [m] Structural data from benzene (ref 32). [n] From normal modes analysis of benzene. [o] Approximate rotational barrier of ethylene is ~60 kcal/mol (see ref 34). [p] Assumed free rotation about a pure single C−N bond. [q] Benedetti structural data (ref 33). [r] Reference 35. [s] Calculated rotational barrier in methylenimine is 57.5 kcal/mol (see ref 36).

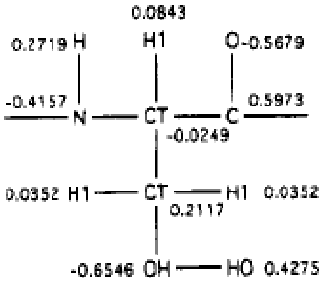
AMBER: J. Am. Chem. Soc. 117, 5179-5197

Molecular Mechanics Force Fields: Atom Partial Charge

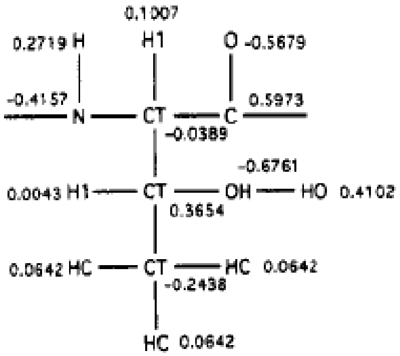


AMBER: J. Am. Chem. Soc. 117, 5179-5197

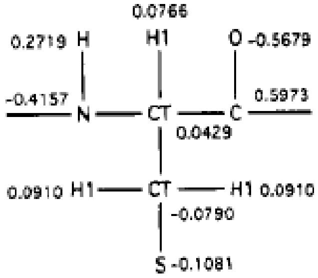
Molecular Mechanics Force Fields: Atom Partial Charge



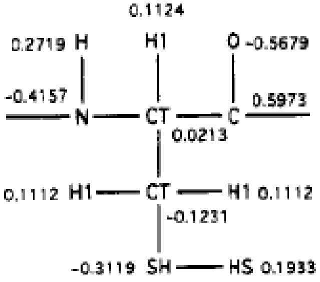
SER



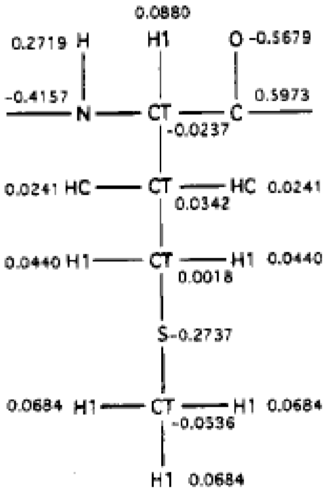
THR



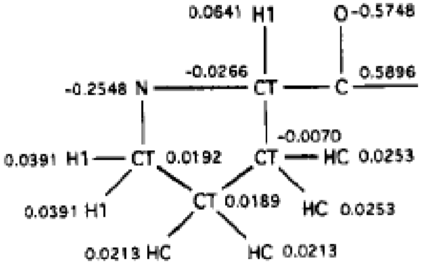
CYX



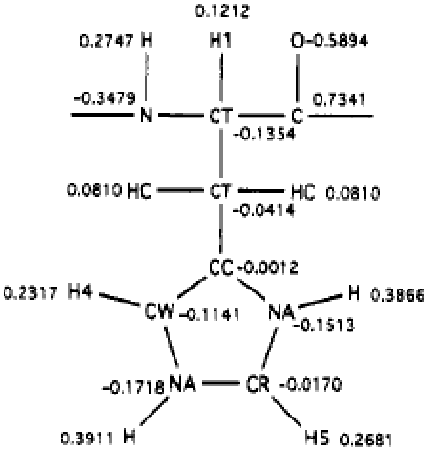
CYS



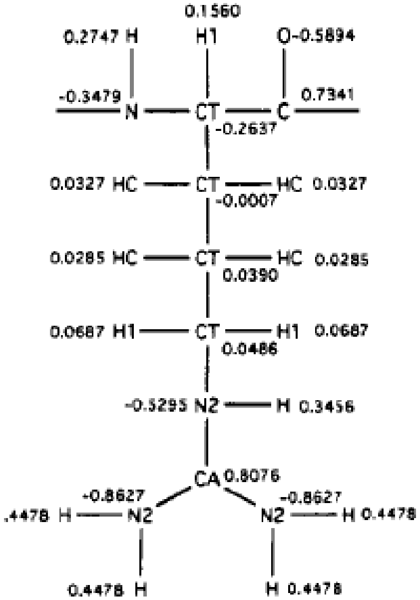
MET



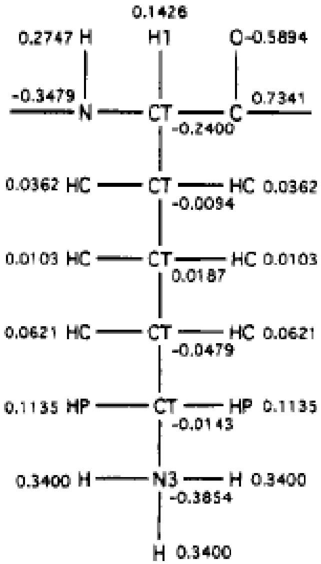
PRO



HIP

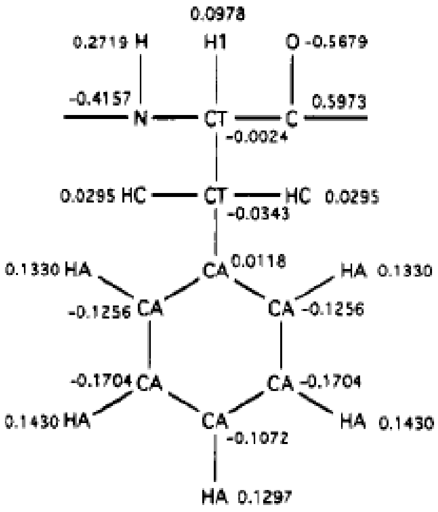


ARG

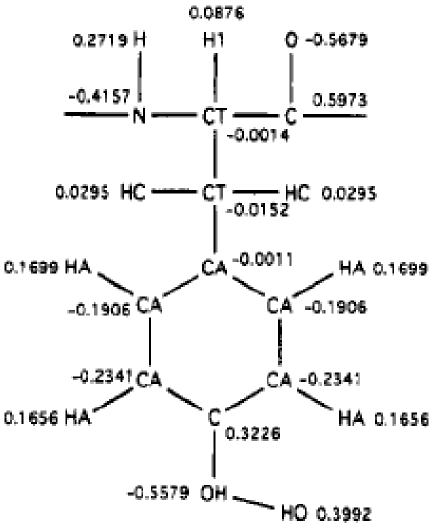


LYS

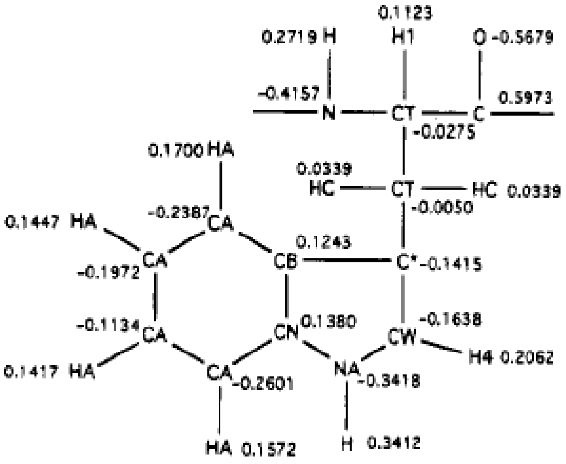
Molecular Mechanics Force Fields: Atom Partial Charge



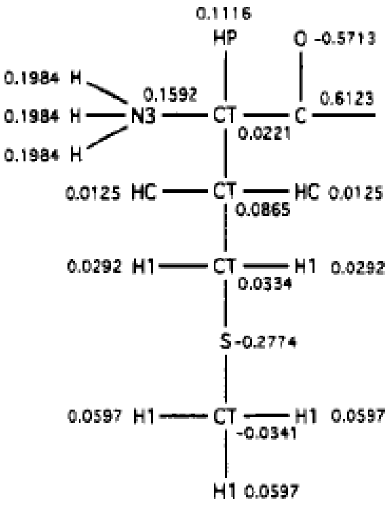
PHE



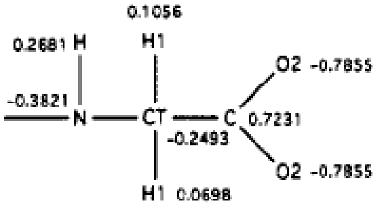
TYR



TRP



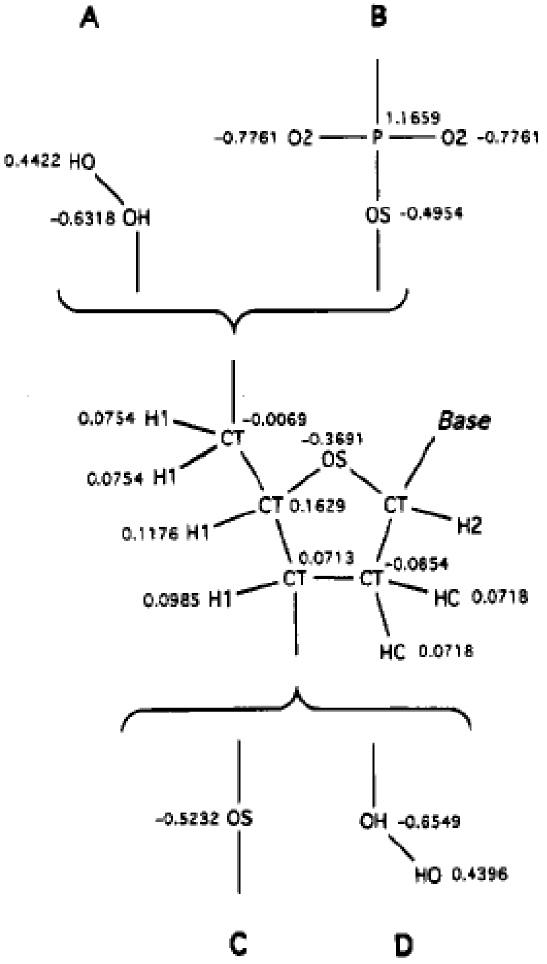
MET_nt

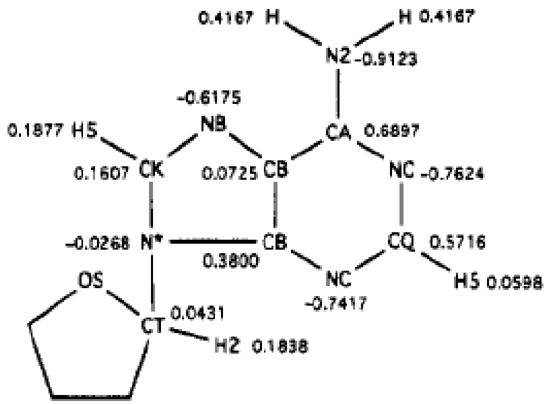


GLY_ct

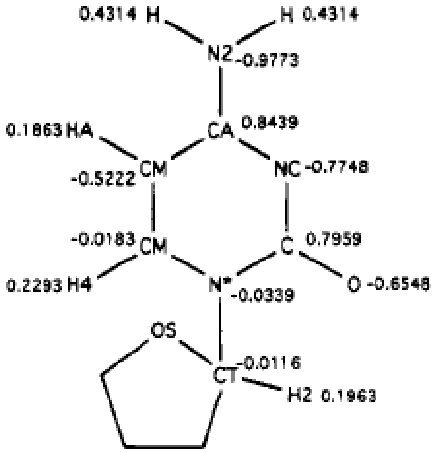
AMBER: J. Am. Chem. Soc. 117, 5179-5197

Molecular Mechanics Force Fields: Atom Partial Charge

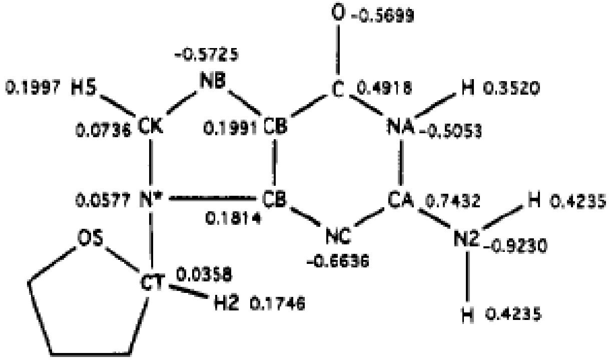




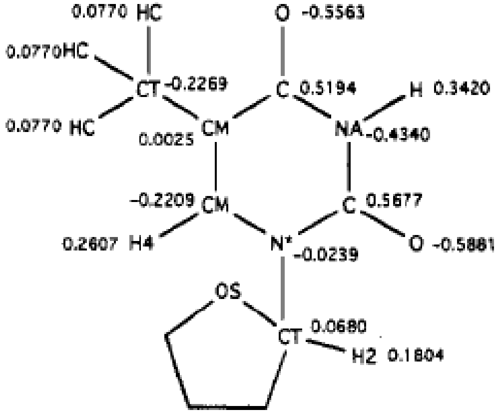
ADE



CYT



GUA

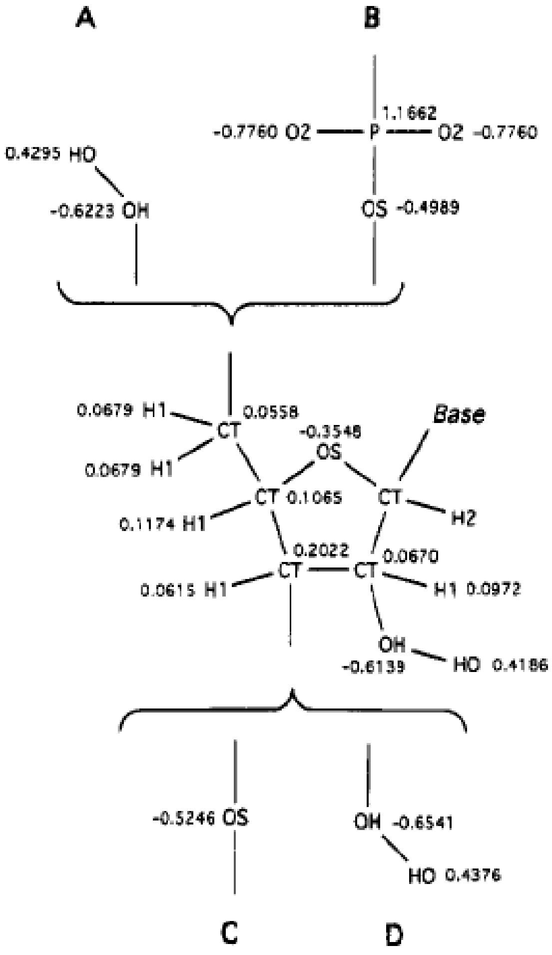


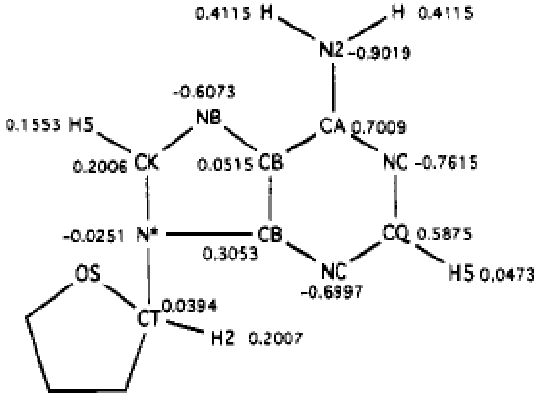
Charge for DNA

THY

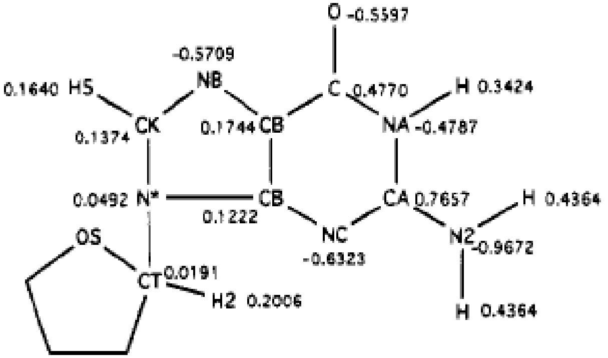
AMBER: J. Am. Chem. Soc. 117, 5179-5197

Molecular Mechanics Force Fields: Atom Partial Charge

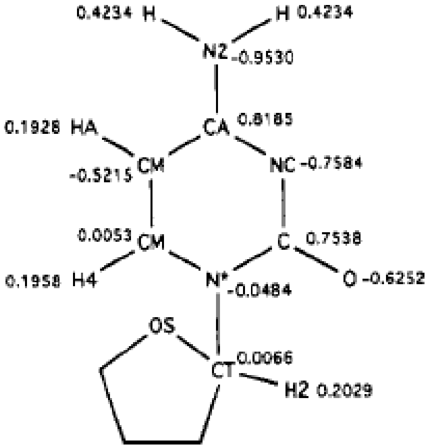




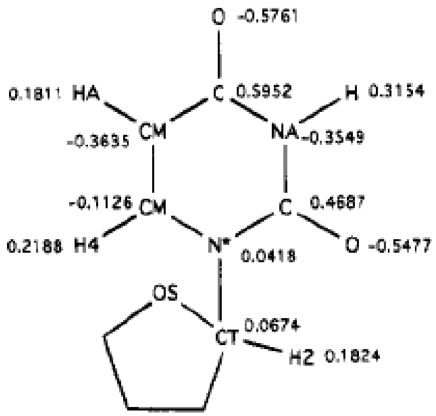
ADE



GUA



CYT

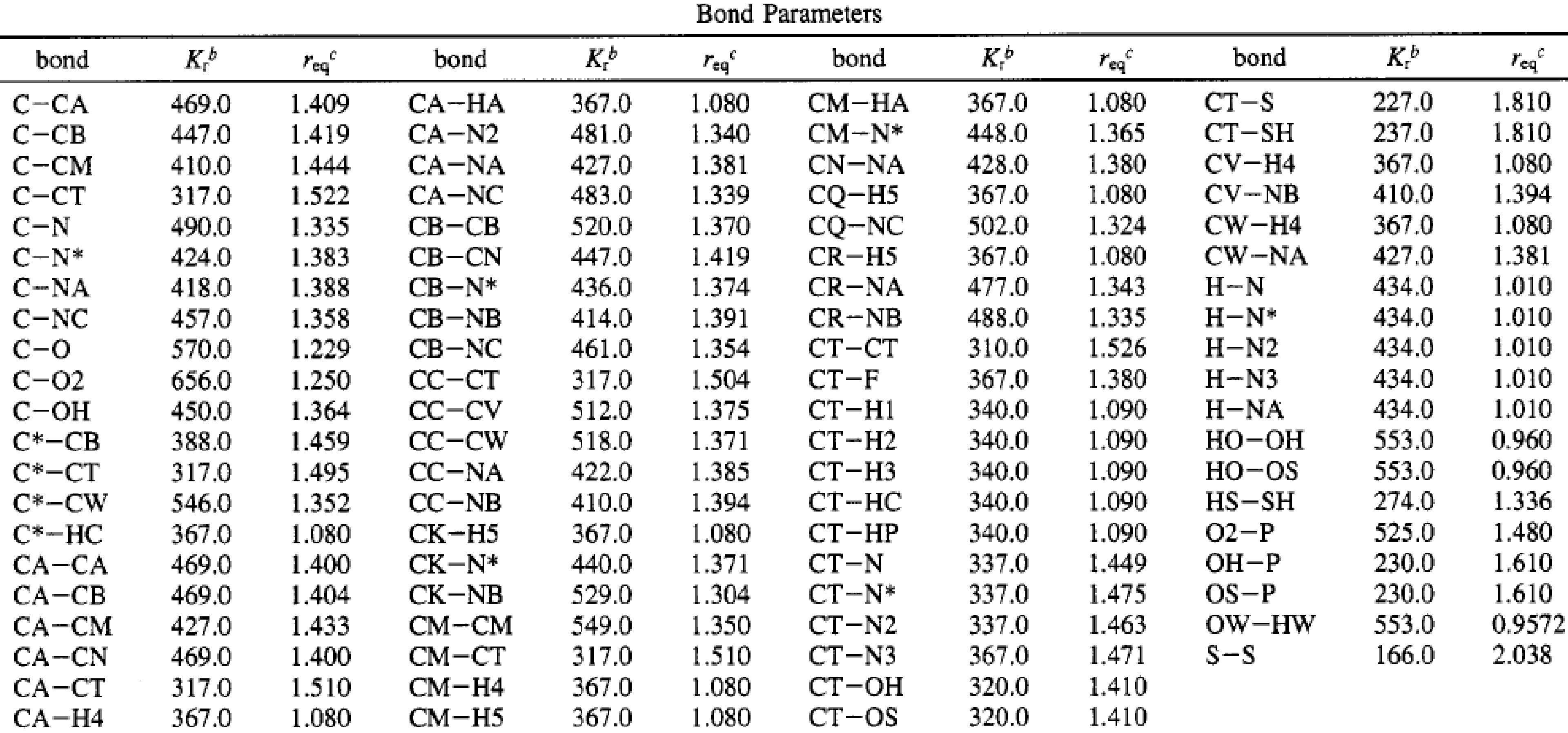


URA

Charge for RNA

AMBER: J. Am. Chem. Soc. 117, 5179-5197

Molecular Mechanics Force Fields: Bond Parameters

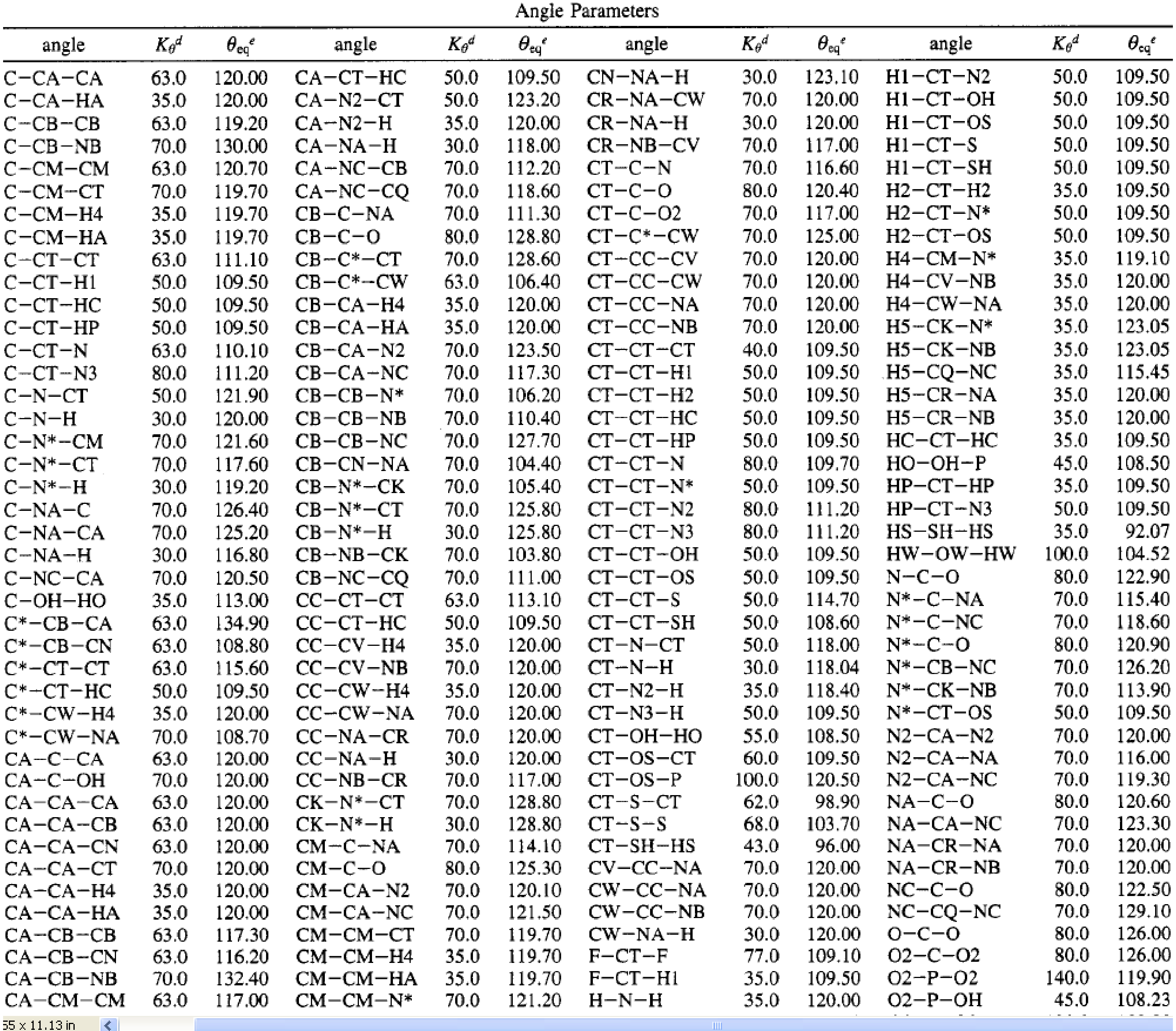


Bond Parameters

bond	K_r^b	r_{eq}^c	bond	K_r^b	r_{eq}^c	bond	K_r^b	r_{eq}^c	bond	K_r^b	r_{eq}^c
C−CA	469.0	1.409	CA−HA	367.0	1.080	CM−HA	367.0	1.080	CT−S	227.0	1.810
C−CB	447.0	1.419	CA−N2	481.0	1.340	CM−N*	448.0	1.365	CT−SH	237.0	1.810
C−CM	410.0	1.444	CA−NA	427.0	1.381	CN−NA	428.0	1.380	CV−H4	367.0	1.080
C−CT	317.0	1.522	CA−NC	483.0	1.339	CQ−H5	367.0	1.080	CV−NB	410.0	1.394
C−N	490.0	1.335	CB−CB	520.0	1.370	CQ−NC	502.0	1.324	CW−H4	367.0	1.080
C−N*	424.0	1.383	CB−CN	447.0	1.419	CR−H5	367.0	1.080	CW−NA	427.0	1.381
C−NA	418.0	1.388	CB−N*	436.0	1.374	CR−NA	477.0	1.343	H−N	434.0	1.010
C−NC	457.0	1.358	CB−NB	414.0	1.391	CR−NB	488.0	1.335	H−N*	434.0	1.010
C−O	570.0	1.229	CB−NC	461.0	1.354	CT−CT	310.0	1.526	H−N2	434.0	1.010
C−O2	656.0	1.250	CC−CT	317.0	1.504	CT−F	367.0	1.380	H−N3	434.0	1.010
C−OH	450.0	1.364	CC−CV	512.0	1.375	CT−H1	340.0	1.090	H−NA	434.0	1.010
C*−CB	388.0	1.459	CC−CW	518.0	1.371	CT−H2	340.0	1.090	HO−OH	553.0	0.960
C*−CT	317.0	1.495	CC−NA	422.0	1.385	CT−H3	340.0	1.090	HO−OS	553.0	0.960
C*−CW	546.0	1.352	CC−NB	410.0	1.394	CT−HC	340.0	1.090	HS−SH	274.0	1.336
C*−HC	367.0	1.080	CK−H5	367.0	1.080	CT−HP	340.0	1.090	O2−P	525.0	1.480
CA−CA	469.0	1.400	CK−N*	440.0	1.371	CT−N	337.0	1.449	OH−P	230.0	1.610
CA−CB	469.0	1.404	CK−NB	529.0	1.304	CT−N*	337.0	1.475	OS−P	230.0	1.610
CA−CM	427.0	1.433	CM−CM	549.0	1.350	CT−N2	337.0	1.463	OW−HW	553.0	0.9572
CA−CN	469.0	1.400	CM−CT	317.0	1.510	CT−N3	367.0	1.471	S−S	166.0	2.038
CA−CT	317.0	1.510	CM−H4	367.0	1.080	CT−OH	320.0	1.410			
CA−H4	367.0	1.080	CM−H5	367.0	1.080	CT−OS	320.0	1.410			

AMBER: J. Am. Chem. Soc. 117, 5179-5197

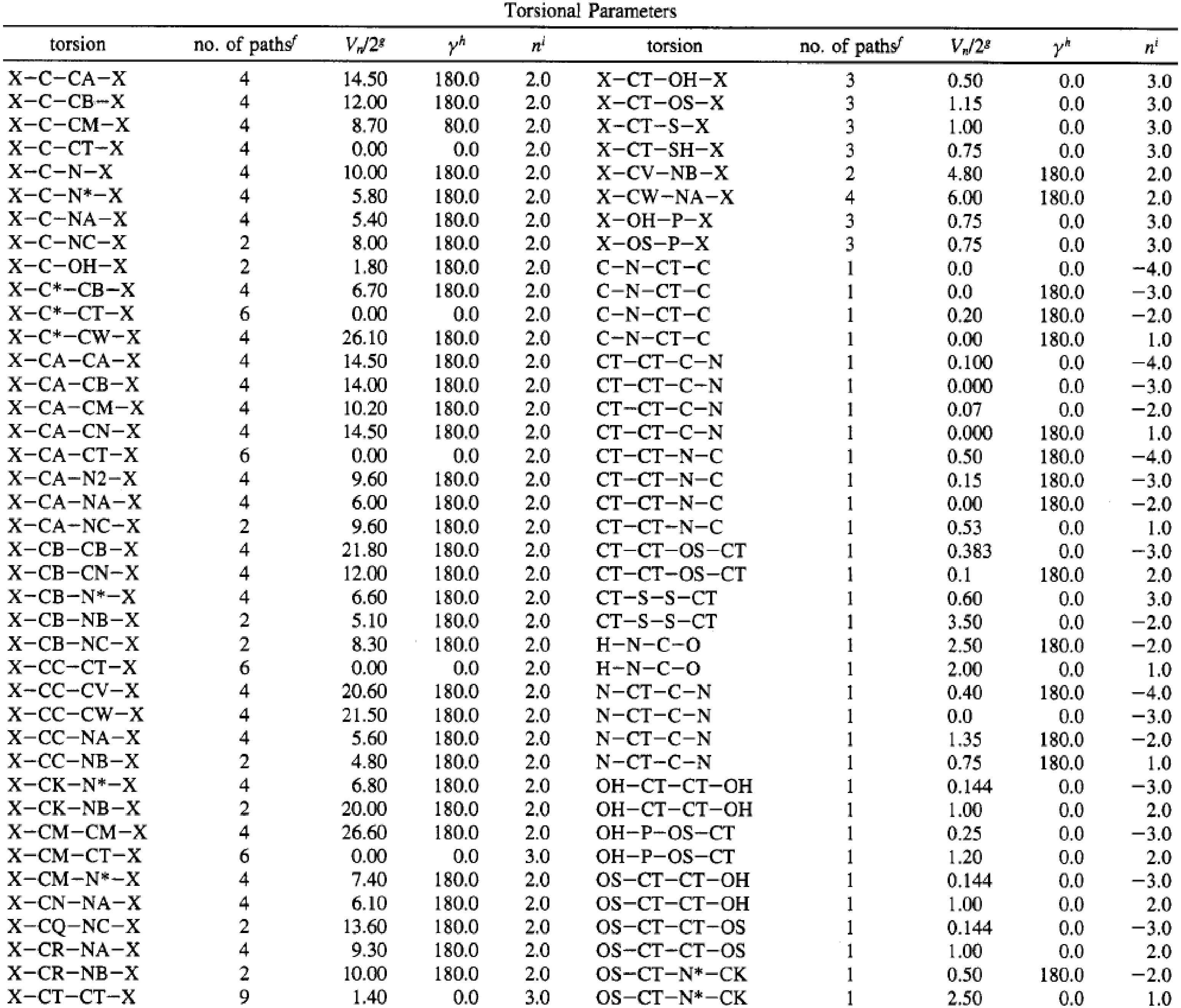
Molecular Mechanics Force Fields: Bond Parameters



Angle Parameters

angle	$K_\theta{}^d$	$\theta_{eq}{}^e$	angle	$K_\theta{}^d$	$\theta_{eq}{}^e$	angle	$K_\theta{}^d$	$\theta_{eq}{}^e$	angle	$K_\theta{}^d$	$\theta_{eq}{}^e$
C-CA-CA	63.0	120.00	CA-CT-HC	50.0	109.50	CN-NA-H	30.0	123.10	H1-CT-N2	50.0	109.50
C-CA-HA	35.0	120.00	CA-N2-CT	50.0	123.20	CR-NA-CW	70.0	120.00	H1-CT-OH	50.0	109.50
C-CB-CB	63.0	119.20	CA-N2-H	35.0	120.00	CR-NA-H	30.0	120.00	H1-CT-OS	50.0	109.50
C-CB-NB	70.0	130.00	CA-NA-H	30.0	118.00	CR-NB-CV	70.0	117.00	H1-CT-S	50.0	109.50
C-CM-CM	63.0	120.70	CA-NC-CB	70.0	112.20	CT-C-N	70.0	116.60	H1-CT-SH	50.0	109.50
C-CM-CT	70.0	119.70	CA-NC-CQ	70.0	118.60	CT-C-O	80.0	120.40	H2-CT-H2	35.0	109.50
C-CM-H4	35.0	119.70	CB-C-NA	70.0	111.30	CT-C-O2	70.0	117.00	H2-CT-N*	50.0	109.50
C-CM-HA	35.0	119.70	CB-C-O	80.0	128.80	CT-C*-CW	70.0	125.00	H2-CT-OS	50.0	109.50
C-CT-CT	63.0	111.10	CB-C*-CT	70.0	128.60	CT-CC-CV	70.0	120.00	H4-CM-N*	35.0	119.10
C-CT-H1	50.0	109.50	CB-C*-CW	63.0	106.40	CT-CC-CW	70.0	120.00	H4-CV-NB	35.0	120.00
C-CT-HC	50.0	109.50	CB-CA-H4	35.0	120.00	CT-CC-NA	70.0	120.00	H4-CW-NA	35.0	120.00
C-CT-HP	50.0	109.50	CB-CA-HA	35.0	120.00	CT-CC-NB	70.0	120.00	H5-CK-N*	35.0	123.05
C-CT-N	63.0	110.10	CB-CA-N2	70.0	123.50	CT-CT-CT	40.0	109.50	H5-CK-NB	35.0	123.05
C-CT-N3	80.0	111.20	CB-CA-NC	70.0	117.30	CT-CT-H1	50.0	109.50	H5-CQ-NC	35.0	115.45
C-N-CT	50.0	121.90	CB-CB-N*	70.0	106.20	CT-CT-H2	50.0	109.50	H5-CR-NA	35.0	120.00
C-N-H	30.0	120.00	CB-CB-NB	70.0	110.40	CT-CT-HC	50.0	109.50	H5-CR-NB	35.0	120.00
C-N*-CM	70.0	121.60	CB-CB-NC	70.0	127.70	CT-CT-HP	50.0	109.50	HC-CT-HC	35.0	109.50
C-N*-CT	70.0	117.60	CB-CN-NA	70.0	104.40	CT-CT-N	80.0	109.70	HO-OH-P	45.0	108.50
C-N*-H	30.0	119.20	CB-N*-CK	70.0	105.40	CT-CT-N*	50.0	109.50	HP-CT-HP	35.0	109.50
C-NA-C	70.0	126.40	CB-N*-CT	70.0	125.80	CT-CT-N2	80.0	111.20	HP-CT-N3	50.0	109.50
C-NA-CA	70.0	125.20	CB-N*-H	30.0	125.80	CT-CT-N3	80.0	111.20	HS-SH-HS	35.0	92.07
C-NA-H	30.0	116.80	CB-NB-CK	70.0	103.80	CT-CT-OH	50.0	109.50	HW-OW-HW	100.0	104.52
C-NC-CA	70.0	120.50	CB-NC-CQ	70.0	111.00	CT-CT-OS	50.0	109.50	N-C-O	80.0	122.90
C-OH-HO	35.0	113.00	CC-CT-CT	63.0	113.10	CT-CT-S	50.0	114.70	N*-C-NA	70.0	115.40
C*-CB-CA	63.0	134.90	CC-CT-HC	50.0	109.50	CT-CT-SH	50.0	108.60	N*-C-NC	70.0	118.60
C*-CB-CN	63.0	108.80	CC-CV-H4	35.0	120.00	CT-N-CT	50.0	118.00	N*-C-O	80.0	120.90
C*-CT-CT	63.0	115.60	CC-CV-NB	70.0	120.00	CT-N-H	30.0	118.04	N*-CB-NC	70.0	126.20
C*-CT-HC	50.0	109.50	CC-CW-H4	35.0	120.00	CT-N2-H	35.0	118.40	N*-CK-NB	70.0	113.90
C*-CW-H4	35.0	120.00	CC-CW-NA	70.0	120.00	CT-N3-H	50.0	109.50	N*-CT-OS	50.0	109.50
C*-CW-NA	70.0	108.70	CC-NA-CR	70.0	120.00	CT-OH-HO	55.0	108.50	N2-CA-N2	70.0	120.00
CA-C-CA	63.0	120.00	CC-NA-H	30.0	120.00	CT-OS-CT	60.0	109.50	N2-CA-NA	70.0	116.00
CA-C-OH	70.0	120.00	CC-NB-CR	70.0	117.00	CT-OS-P	100.0	120.50	N2-CA-NC	70.0	119.30
CA-CA-CA	63.0	120.00	CK-N*-CT	70.0	128.80	CT-S-CT	62.0	98.90	NA-C-O	80.0	120.60
CA-CA-CB	63.0	120.00	CK-N*-H	30.0	128.80	CT-S-S	68.0	103.70	NA-CA-NC	70.0	123.30
CA-CA-CN	63.0	120.00	CM-C-NA	70.0	114.10	CT-SH-HS	43.0	96.00	NA-CR-NA	70.0	120.00
CA-CA-CT	70.0	120.00	CM-C-O	80.0	125.30	CV-CC-NA	70.0	120.00	NA-CR-NB	70.0	120.00
CA-CA-H4	35.0	120.00	CM-CA-N2	70.0	120.10	CW-CC-NA	70.0	120.00	NC-C-O	80.0	122.50
CA-CA-HA	35.0	120.00	CM-CA-NC	70.0	121.50	CW-CC-NB	70.0	120.00	NC-CQ-NC	70.0	129.10
CA-CB-CB	63.0	117.30	CM-CM-CT	70.0	119.70	CW-NA-H	30.0	120.00	O-C-O	80.0	126.00
CA-CB-CN	63.0	116.20	CM-CM-H4	35.0	119.70	F-CT-F	77.0	109.10	O2-C-O2	80.0	126.00
CA-CB-NB	70.0	132.40	CM-CM-HA	35.0	119.70	F-CT-H1	35.0	109.50	O2-P-O2	140.0	119.90
CA-CM-CM	63.0	117.00	CM-CM-N*	70.0	121.20	H-N-H	35.0	120.00	O2-P-OH	45.0	108.23

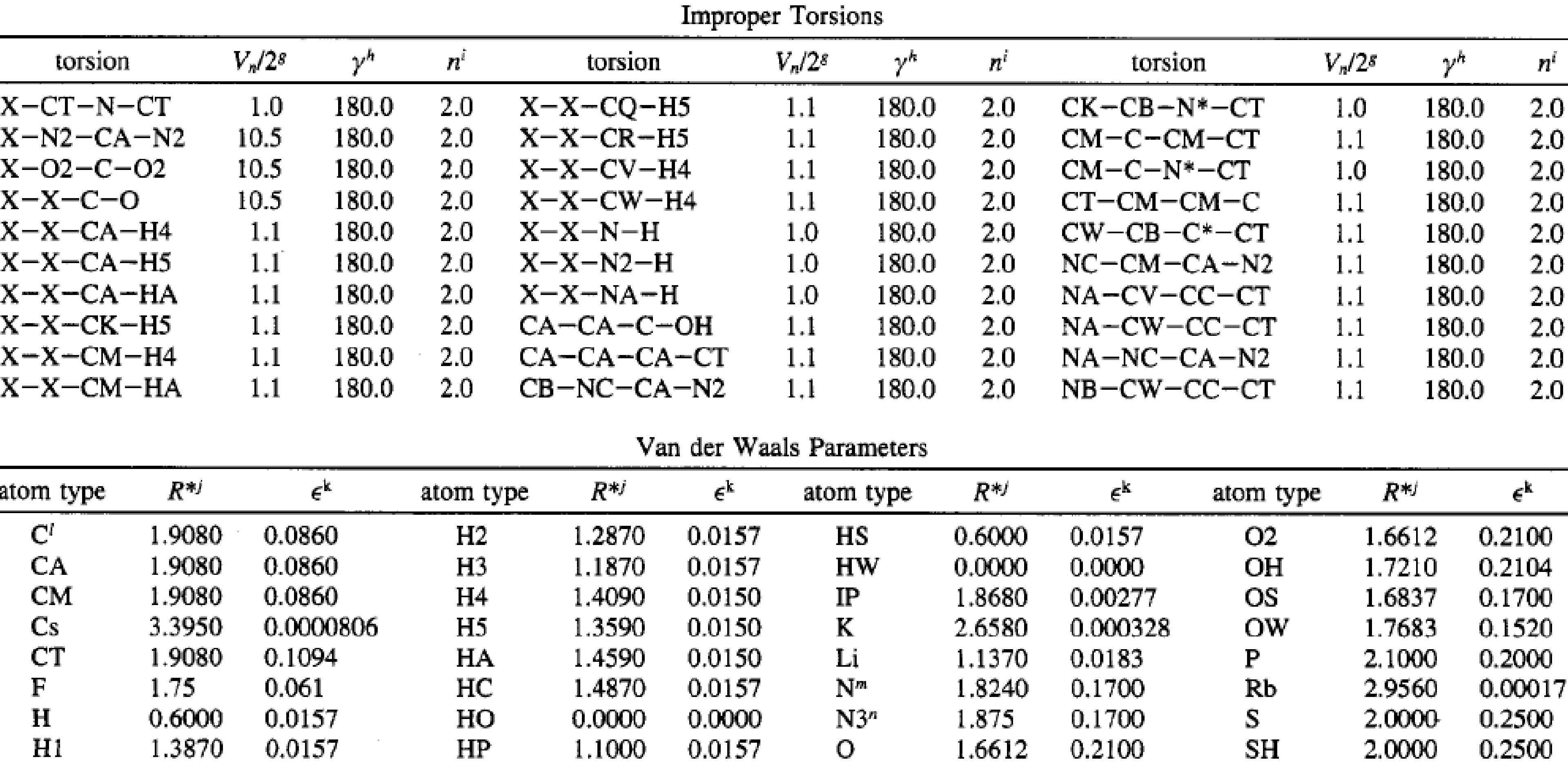
Molecular Mechanics Force Fields: Bond Parameters



Torsional Parameters

torsion	no. of paths^f	$V_n/2^g$	γ^h	n^i	torsion	no. of paths^f	$V_n/2^g$	γ^h	n^i
X–C–CA–X	4	14.50	180.0	2.0	X–CT–OH–X	3	0.50	0.0	3.0
X–C–CB–X	4	12.00	180.0	2.0	X–CT–OS–X	3	1.15	0.0	3.0
X–C–CM–X	4	8.70	80.0	2.0	X–CT–S–X	3	1.00	0.0	3.0
X–C–CT–X	4	0.00	0.0	2.0	X–CT–SH–X	3	0.75	0.0	3.0
X–C–N–X	4	10.00	180.0	2.0	X–CV–NB–X	2	4.80	180.0	2.0
X–C–N*–X	4	5.80	180.0	2.0	X–CW–NA–X	4	6.00	180.0	2.0
X–C–NA–X	4	5.40	180.0	2.0	X–OH–P–X	3	0.75	0.0	3.0
X–C–NC–X	2	8.00	180.0	2.0	X–OS–P–X	3	0.75	0.0	3.0
X–C–OH–X	2	1.80	180.0	2.0	C–N–CT–C	1	0.0	0.0	−4.0
X–C*–CB–X	4	6.70	180.0	2.0	C–N–CT–C	1	0.0	180.0	−3.0
X–C*–CT–X	6	0.00	0.0	2.0	C–N–CT–C	1	0.20	180.0	−2.0
X–C*–CW–X	4	26.10	180.0	2.0	C–N–CT–C	1	0.00	180.0	1.0
X–CA–CA–X	4	14.50	180.0	2.0	CT–CT–C–N	1	0.100	0.0	−4.0
X–CA–CB–X	4	14.00	180.0	2.0	CT–CT–C–N	1	0.000	0.0	−3.0
X–CA–CM–X	4	10.20	180.0	2.0	CT–CT–C–N	1	0.07	0.0	−2.0
X–CA–CN–X	4	14.50	180.0	2.0	CT–CT–C–N	1	0.000	180.0	1.0
X–CA–CT–X	6	0.00	0.0	2.0	CT–CT–N–C	1	0.50	180.0	−4.0
X–CA–N2–X	4	9.60	180.0	2.0	CT–CT–N–C	1	0.15	180.0	−3.0
X–CA–NA–X	4	6.00	180.0	2.0	CT–CT–N–C	1	0.00	180.0	−2.0
X–CA–NC–X	2	9.60	180.0	2.0	CT–CT–N–C	1	0.53	0.0	1.0
X–CB–CB–X	4	21.80	180.0	2.0	CT–CT–OS–CT	1	0.383	0.0	−3.0
X–CB–CN–X	4	12.00	180.0	2.0	CT–CT–OS–CT	1	0.1	180.0	2.0
X–CB–N*–X	4	6.60	180.0	2.0	CT–S–S–CT	1	0.60	0.0	3.0
X–CB–NB–X	2	5.10	180.0	2.0	CT–S–S–CT	1	3.50	0.0	−2.0
X–CB–NC–X	2	8.30	180.0	2.0	H–N–C–O	1	2.50	180.0	−2.0
X–CC–CT–X	6	0.00	0.0	2.0	H–N–C–O	1	2.00	0.0	1.0
X–CC–CV–X	4	20.60	180.0	2.0	N–CT–C–N	1	0.40	180.0	−4.0
X–CC–CW–X	4	21.50	180.0	2.0	N–CT–C–N	1	0.0	0.0	−3.0
X–CC–NA–X	4	5.60	180.0	2.0	N–CT–C–N	1	1.35	180.0	−2.0
X–CC–NB–X	2	4.80	180.0	2.0	N–CT–C–N	1	0.75	180.0	1.0
X–CK–N*–X	4	6.80	180.0	2.0	OH–CT–CT–OH	1	0.144	0.0	−3.0
X–CK–NB–X	2	20.00	180.0	2.0	OH–CT–CT–OH	1	1.00	0.0	2.0
X–CM–CM–X	4	26.60	180.0	2.0	OH–P–OS–CT	1	0.25	0.0	−3.0
X–CM–CT–X	6	0.00	0.0	3.0	OH–P–OS–CT	1	1.20	0.0	2.0
X–CM–N*–X	4	7.40	180.0	2.0	OS–CT–CT–OH	1	0.144	0.0	−3.0
X–CN–NA–X	4	6.10	180.0	2.0	OS–CT–CT–OH	1	1.00	0.0	2.0
X–CQ–NC–X	2	13.60	180.0	2.0	OS–CT–CT–OS	1	0.144	0.0	−3.0
X–CR–NA–X	4	9.30	180.0	2.0	OS–CT–CT–OS	1	1.00	0.0	2.0
X–CR–NB–X	2	10.00	180.0	2.0	OS–CT–N*–CK	1	0.50	180.0	−2.0
X–CT–CT–X	9	1.40	0.0	3.0	OS–CT–N*–CK	1	2.50	0.0	1.0

Molecular Mechanics Force Fields: Bond and Non-Bonded Parameters



Improper Torsions

torsion	$V_n/2^g$	γ^h	n^i	torsion	$V_n/2^g$	γ^h	n^i	torsion	$V_n/2^g$	γ^h	n^i
X−CT−N−CT	1.0	180.0	2.0	X−X−CQ−H5	1.1	180.0	2.0	CK−CB−N*−CT	1.0	180.0	2.0
X−N2−CA−N2	10.5	180.0	2.0	X−X−CR−H5	1.1	180.0	2.0	CM−C−CM−CT	1.1	180.0	2.0
X−O2−C−O2	10.5	180.0	2.0	X−X−CV−H4	1.1	180.0	2.0	CM−C−N*−CT	1.0	180.0	2.0
X−X−C−O	10.5	180.0	2.0	X−X−CW−H4	1.1	180.0	2.0	CT−CM−CM−C	1.1	180.0	2.0
X−X−CA−H4	1.1	180.0	2.0	X−X−N−H	1.0	180.0	2.0	CW−CB−C*−CT	1.1	180.0	2.0
X−X−CA−H5	1.1	180.0	2.0	X−X−N2−H	1.0	180.0	2.0	NC−CM−CA−N2	1.1	180.0	2.0
X−X−CA−HA	1.1	180.0	2.0	X−X−NA−H	1.0	180.0	2.0	NA−CV−CC−CT	1.1	180.0	2.0
X−X−CK−H5	1.1	180.0	2.0	CA−CA−C−OH	1.1	180.0	2.0	NA−CW−CC−CT	1.1	180.0	2.0
X−X−CM−H4	1.1	180.0	2.0	CA−CA−CA−CT	1.1	180.0	2.0	NA−NC−CA−N2	1.1	180.0	2.0
X−X−CM−HA	1.1	180.0	2.0	CB−NC−CA−N2	1.1	180.0	2.0	NB−CW−CC−CT	1.1	180.0	2.0

Van der Waals Parameters

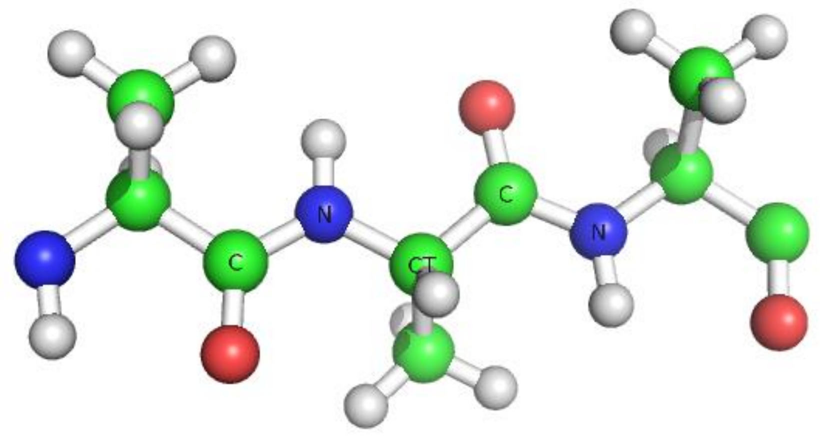
atom type	R^{*j}	ϵ^k	atom type	R^{*j}	ϵ^k	atom type	R^{*j}	ϵ^k	atom type	R^{*j}	ϵ^k
C^l	1.9080	0.0860	H2	1.2870	0.0157	HS	0.6000	0.0157	O2	1.6612	0.2100
CA	1.9080	0.0860	H3	1.1870	0.0157	HW	0.0000	0.0000	OH	1.7210	0.2104
CM	1.9080	0.0860	H4	1.4090	0.0150	IP	1.8680	0.00277	OS	1.6837	0.1700
Cs	3.3950	0.0000806	H5	1.3590	0.0150	K	2.6580	0.000328	OW	1.7683	0.1520
CT	1.9080	0.1094	HA	1.4590	0.0150	Li	1.1370	0.0183	P	2.1000	0.2000
F	1.75	0.061	HC	1.4870	0.0157	N^m	1.8240	0.1700	Rb	2.9560	0.00017
H	0.6000	0.0157	HO	0.0000	0.0000	$N3^n$	1.875	0.1700	S	2.0000	0.2500
H1	1.3870	0.0157	HP	1.1000	0.0157	O	1.6612	0.2100	SH	2.0000	0.2500

AMBER: J. Am. Chem. Soc. 117, 5179-5197

Molecular Mechanical Force Field Parameterization *For* Biomolecules

Torsional Angle Parameterization

$$V_{potential} = \sum_{bonds} K_r (r - r_{eq})^2 + \sum_{angles} K_\theta (\theta - \theta_{eq})^2 + \boxed{\sum_{dihedrals} \frac{V_n}{2}[1 + \cos(n\phi - \gamma)]} + \sum_{i<j} \left\{ \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} \right] + \frac{q_i q_j}{\varepsilon R_{ij}} \right\}$$



10 GA runs with an internal paralleled GA to reproduce LMP2/cc-pvtz(-f) energies of dipeptides

	Model 1			**Model 2**		
	Fourier (n)	$V_n/2$	Phase angle	Fourier	$V_n/2$ (n)	Phase angle
C-N-CT-C	1	0.748	0	1	0.552	0
	2	0.719	0	2	0.662	0
	3	0.199	0	3	0.316	0
N-CT-C-N	1	0.319	0	1	0.109	0
	2	1.194	180	2	1.184	180
	3	0.013	0	3	0.000	0

How Well Do AMBER FF Reproduce QM Energies

	FF94	FF96	FF99	FF99SB	FF03	M1	M2
ALA	0.79	0.82	0.77	0.72	0.82	0.86	0.87
ARG	0.80	0.80	0.81	0.75	0.81	0.81	0.82
ASN	0.68	0.77	0.65	0.64	0.75	0.79	0.79
ASP	0.79	0.73	0.81	0.76	0.78	0.77	0.78
CYS	0.74	0.78	0.71	0.72	0.67	0.79	0.80
GLN	0.73	0.81	0.70	0.69	0.80	0.83	0.83
GLU	0.76	0.80	0.73	0.69	0.82	0.80	0.80
GLY	**0.81**	**0.80**	**0.75**	**0.85**	**0.87**	**0.86**	**0.87**
HIS	0.72	0.83	0.66	0.68	0.62	0.84	0.85
ILE	0.86	0.88	0.84	0.83	0.80	0.89	0.90
LEU	0.82	0.87	0.78	0.80	0.84	0.88	0.89
LYS	0.93	0.94	0.91	0.91	0.88	0.95	0.95

QM: LMP2/cc-pVTZ (-f), 20 ϕ-ψ maps

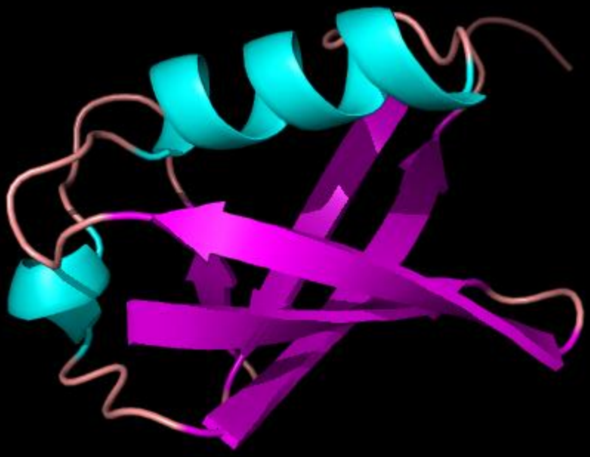
How Well Do AMBER FF Reproduce QM Energies - Continued

	FF94	FF96	FF99	FF99SB	FF03	Model1	Model2
MET	0.81	0.86	0.78	0.79	0.82	0.88	0.88
PHE	0.77	0.81	0.75	0.72	0.81	0.83	0.84
PRO	0.89	0.89	0.91	0.90	0.89	0.90	0.90
SER	0.78	0.81	0.74	0.78	0.77	0.83	0.83
THR	0.73	0.77	0.67	0.68	0.74	0.78	0.78
TRP	0.76	0.81	0.75	0.72	0.77	0.80	0.81
TYR	0.69	0.75	0.63	0.68	0.56	0.73	0.74
VAL	0.87	0.88	0.85	0.85	0.76	0.88	0.89
AUE	2.23	2.11	2.48	2.67	2.34	**2.01**	**1.99**
RMSE	2.94	2.78	3.18	3.44	3.05	**2.68**	**2.65**
R^2	0.79	0.81	0.77	0.76	0.78	**0.83**	**0.84**

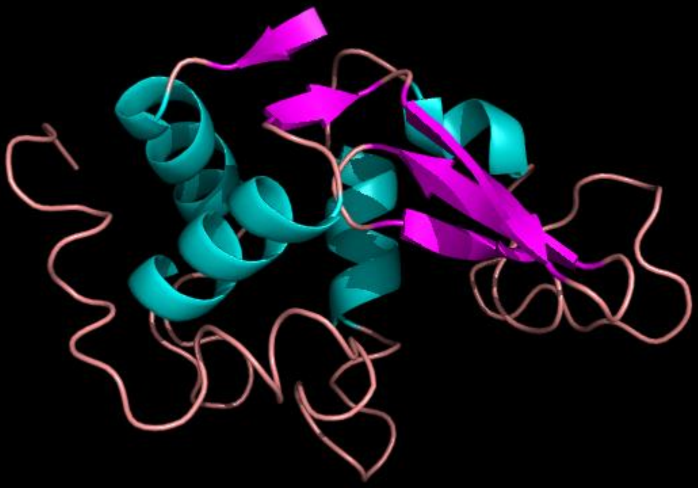
Evaluation of Force Field Models

1. NMR order parameters S^2
2. Protein folding
3. ~~J-couplings of small peptides~~
4. ~~Decoy analysis~~

S^2 Prediction – Measure Amide N-H Vibrational Motion



1UBQ

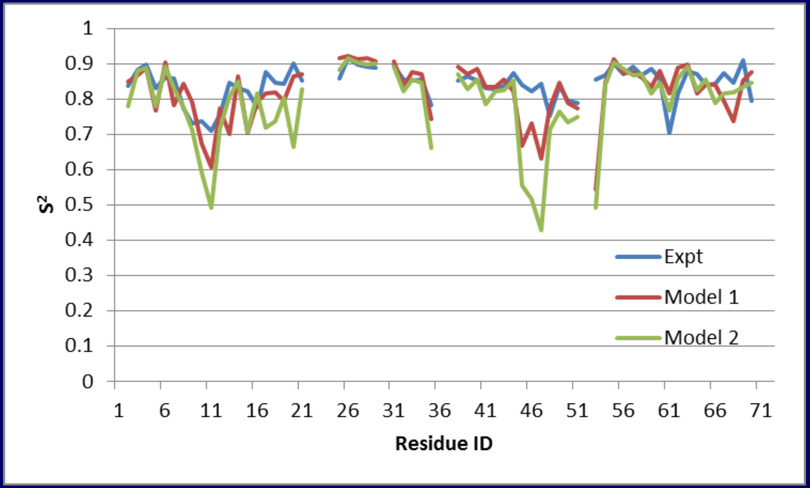


6LYT – hen egg white lysozyme

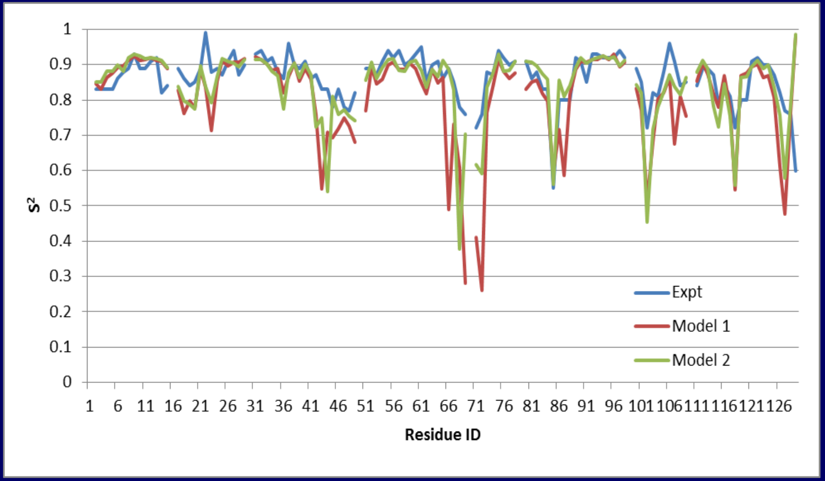
- 25 nano second MD simulations
- Autocorrelation function was calculated up to 4ns
  (the mean of the second half was taken as S^2)
- Experimental data
  Ubiquitin: JACS, 117, 12562; lysozyme: Biochemistry, 34, 4041

Backbone RMSD Plot And Order Parameter S^2 of 1UBQ And 6YLT

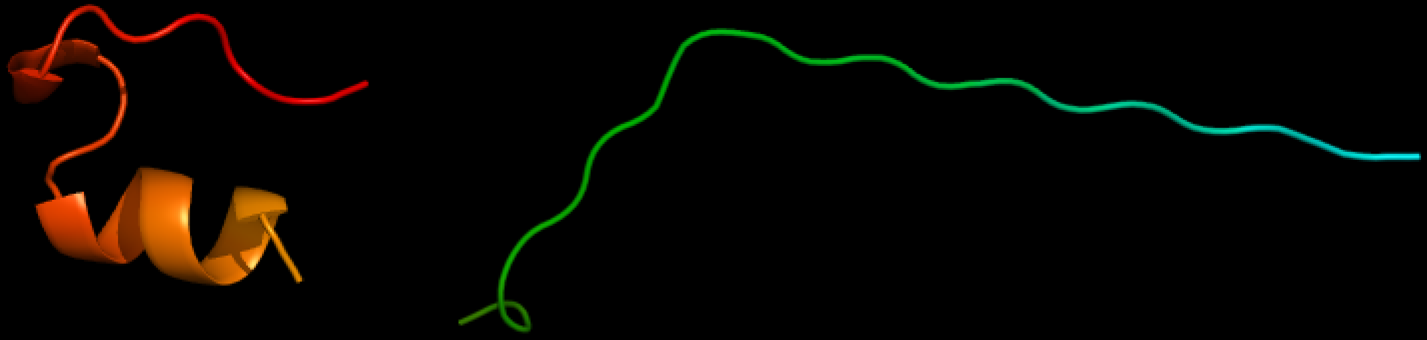
ubiquitin



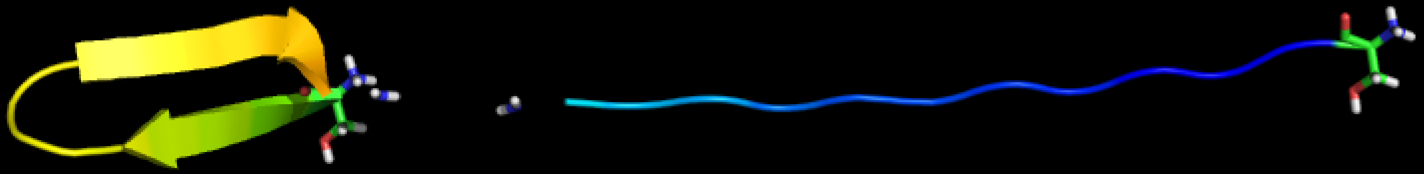
Lysozyme



Protein Folding



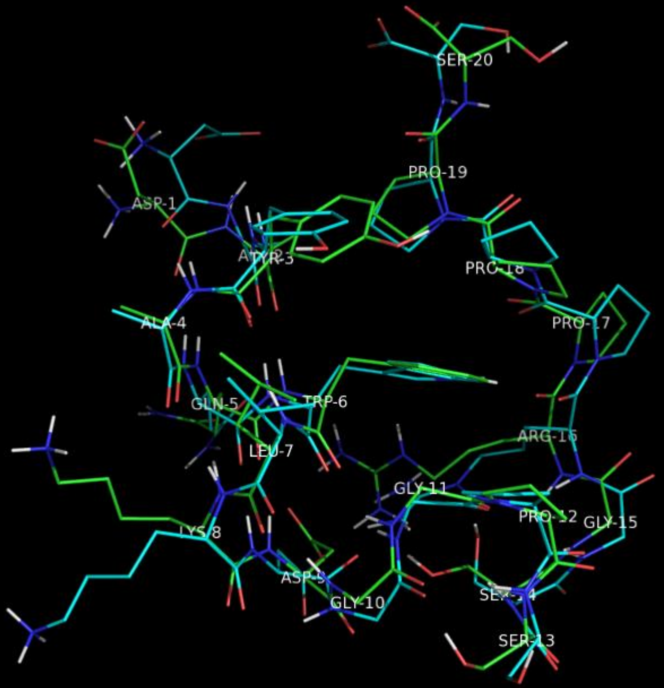
2OJF (Trpcage)

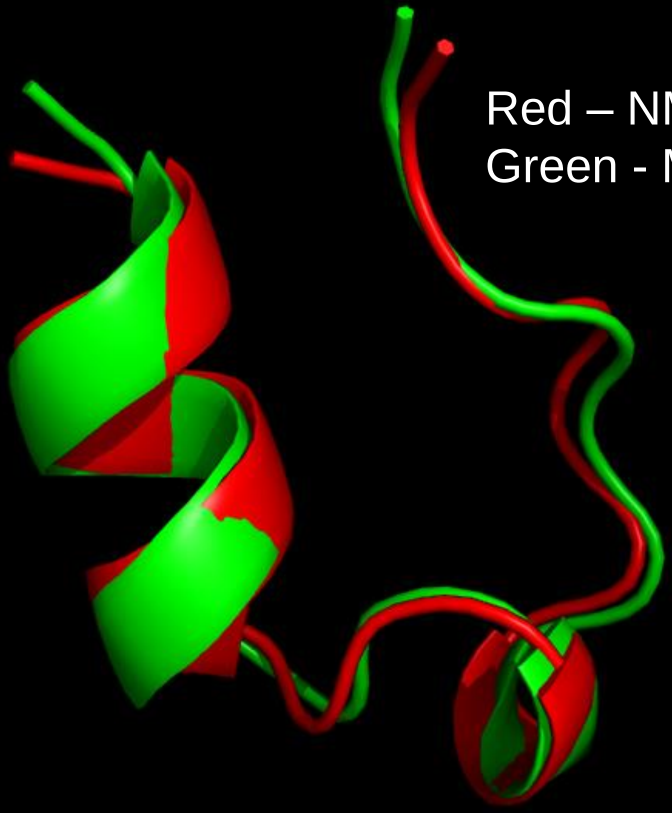


1LE1 (tryptophan zipper 2)

1. Starting from fully extended conformation
2. GBMD at 298 K

Comparison of the “Best” MD And NMR Structures for Trpcage

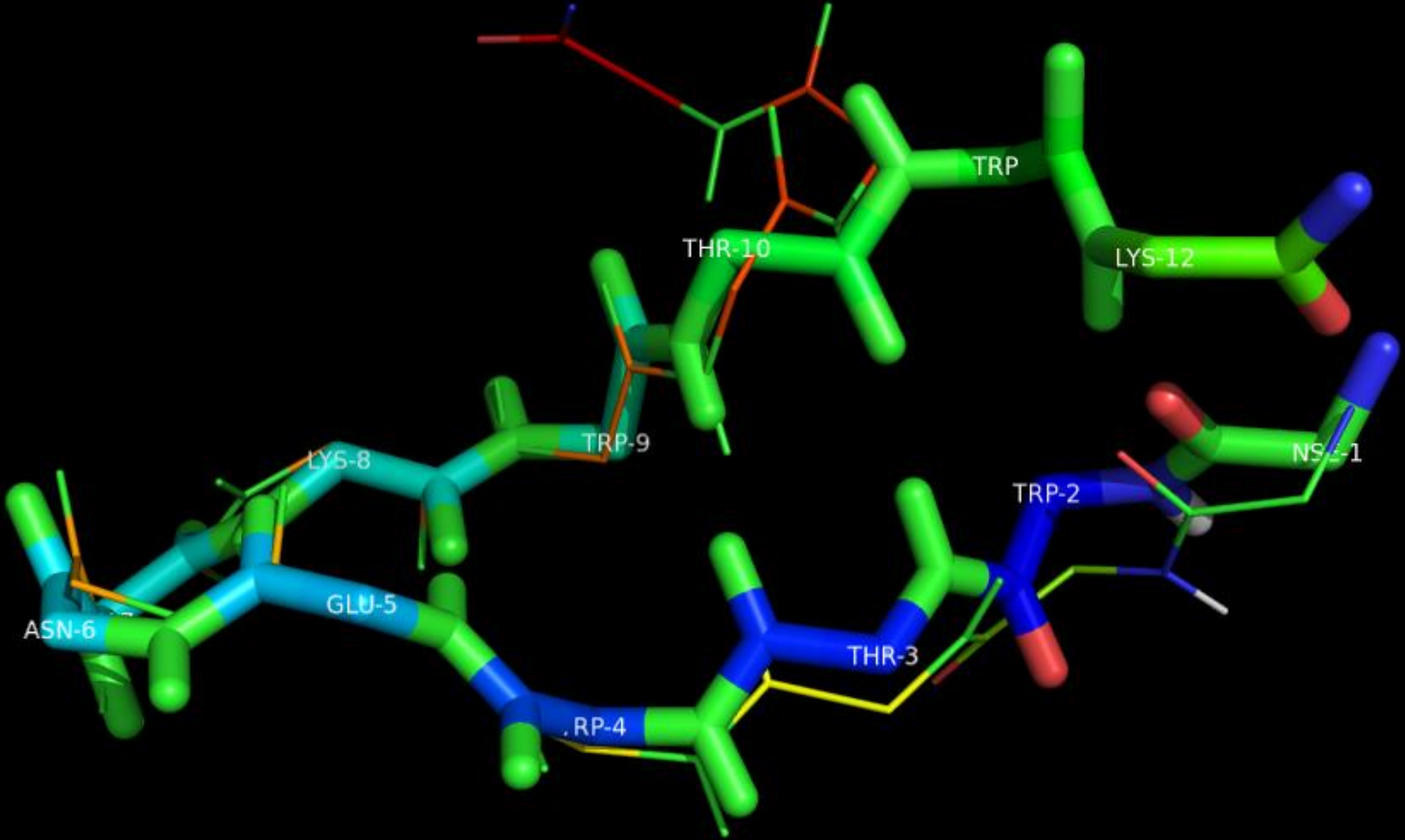




Red – NMR
Green - MD

The MD structure was obtained with Model 2, main chain RMSD is 0.80 Å
Simmerling et al’s folding achieved an RMSD of 1.1 Å (JACS, 124, 11258-11259)

Comparison of the “Best” MD And NMR Structures of Tryptophan Zipper 2

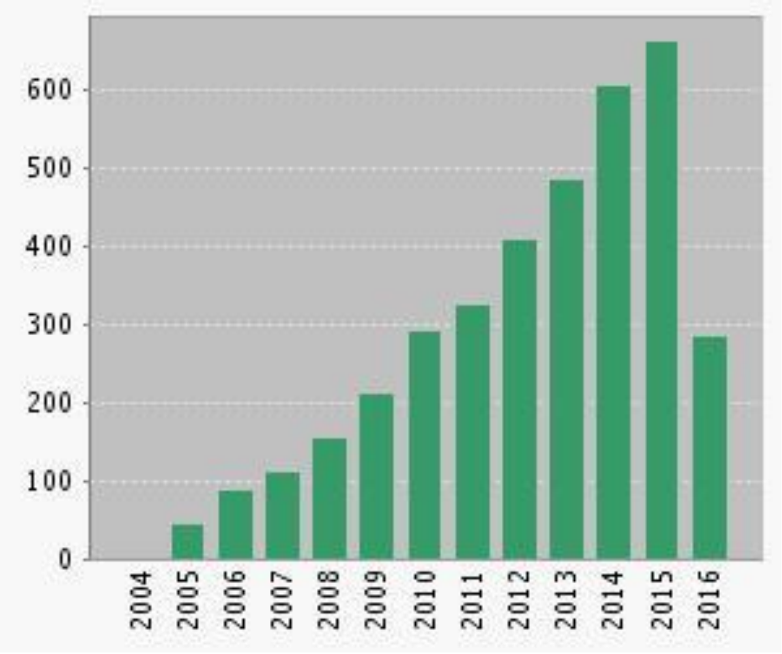


The MD structure was obtained with Model 2, main chain RMSD is 0.79 Å

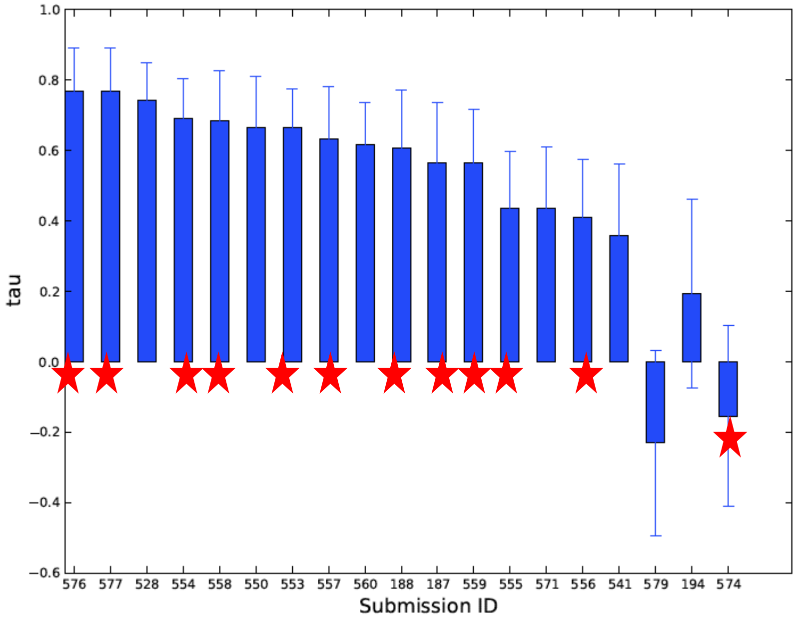
Molecular Mechanical Force Field Parameterization For Organic / Drug Molecules

GAFF – A Successful, Widely Used MMFF

Citations in Each Year *



SAMPL4 Host-Guest Challenge **



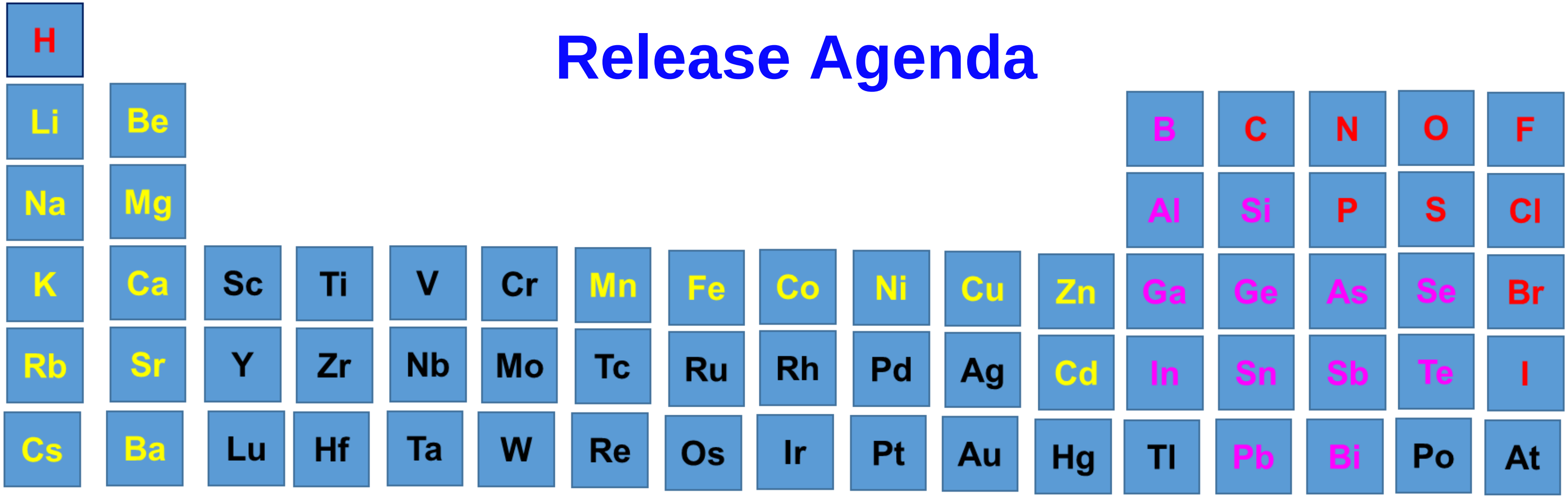
*** webofknowledge.com**

**** Courtesy: Wei Yang, FSU**

Correlation Coefficients Between the Expt. and Predicted Binding Free Energies of 23 Guest molecules of Cucurbit Uril and Octaacid

GAFF was used in 14 out of 21 submissions.

Release Agenda



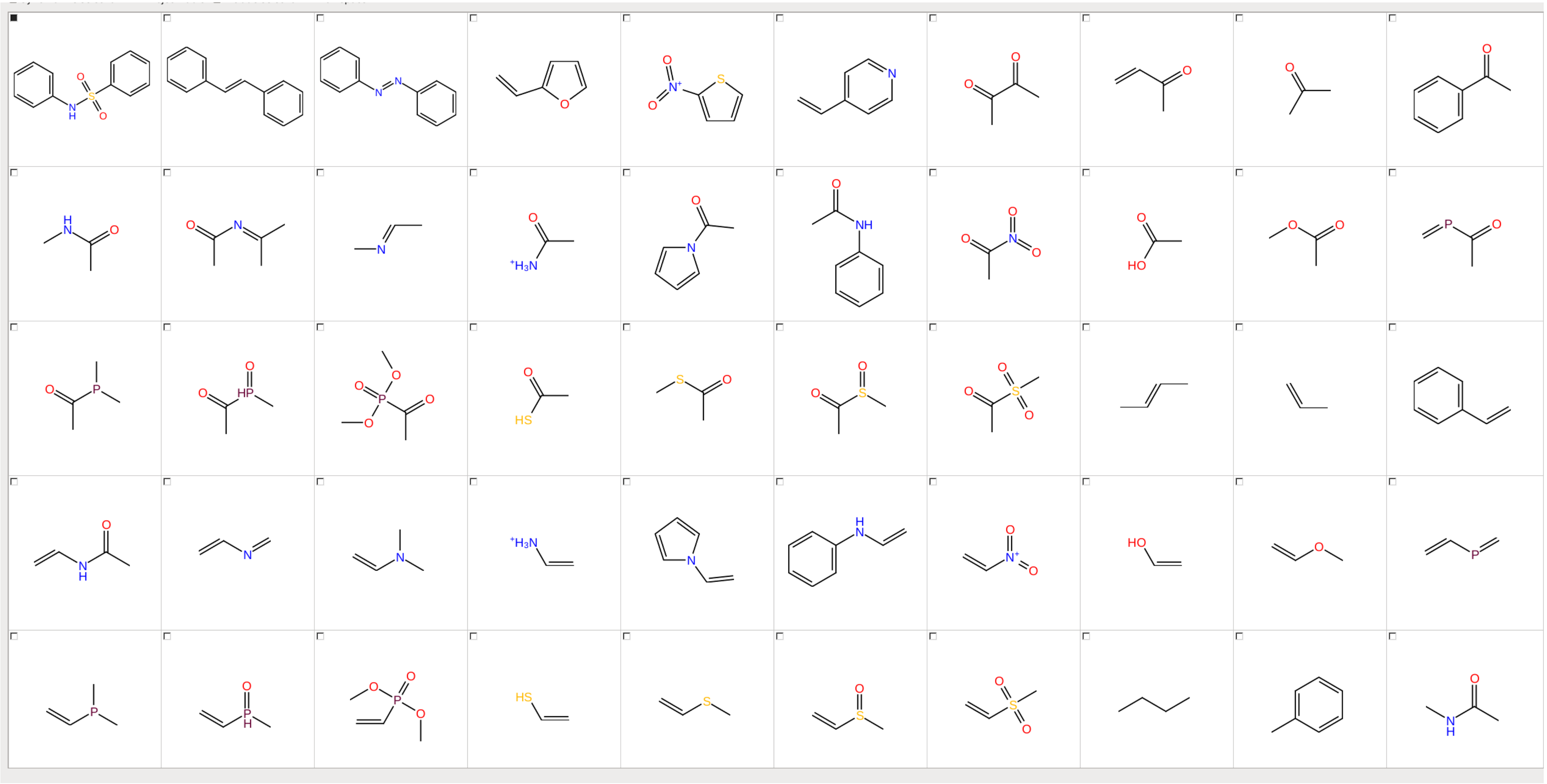
H																
Li	Be											B	C	N	O	F
Na	Mg											Al	Si	P	S	Cl
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I
Cs	Ba	Lu	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At

1. **Basic version released with AMBER16**
   10 basic elements

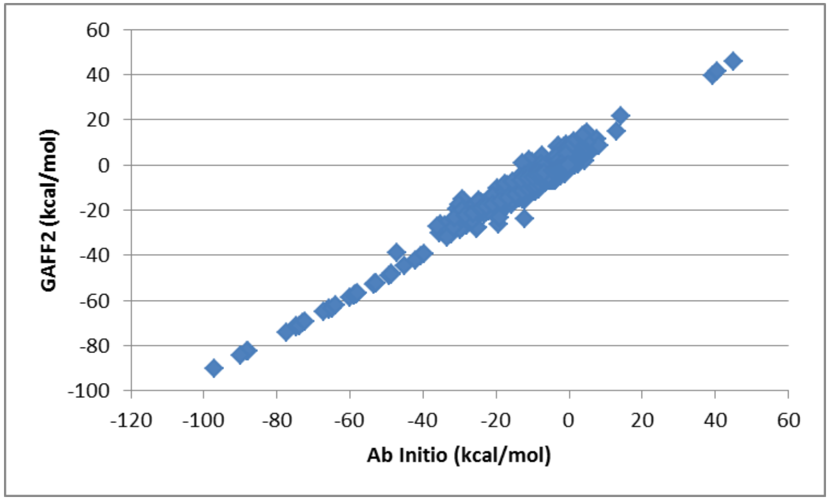
2. **Extended version released before September 2016**
   Other nonmetallic atom types

3. **Full version in 2017**
   Metallic elements

Diversity of Model Compounds For Torsional Angle Parameterization

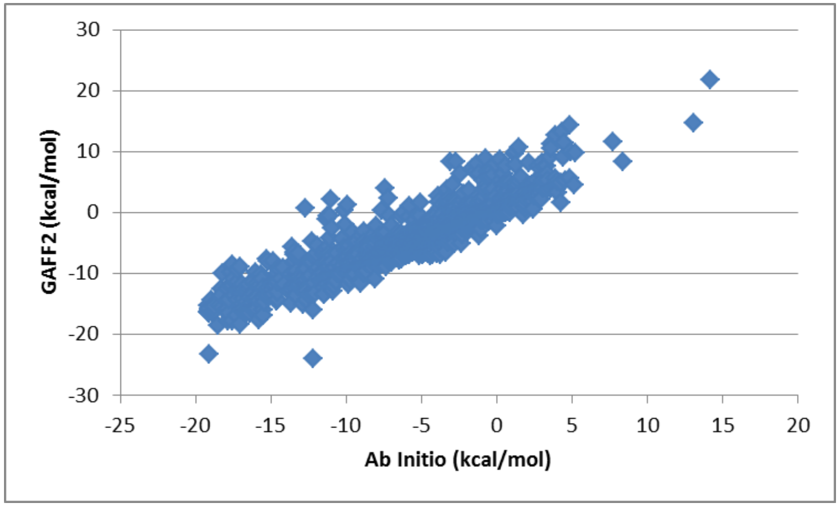


Performance of GAFF2 in Reproducing *Ab Initio* Interaction Energies



AUE = 1.10 kcal/mol
RMSE = 2.11 kcal/mol
ASE = -0.80 kcal/mol
Number of data points: 2954

Excluding Charged molecules

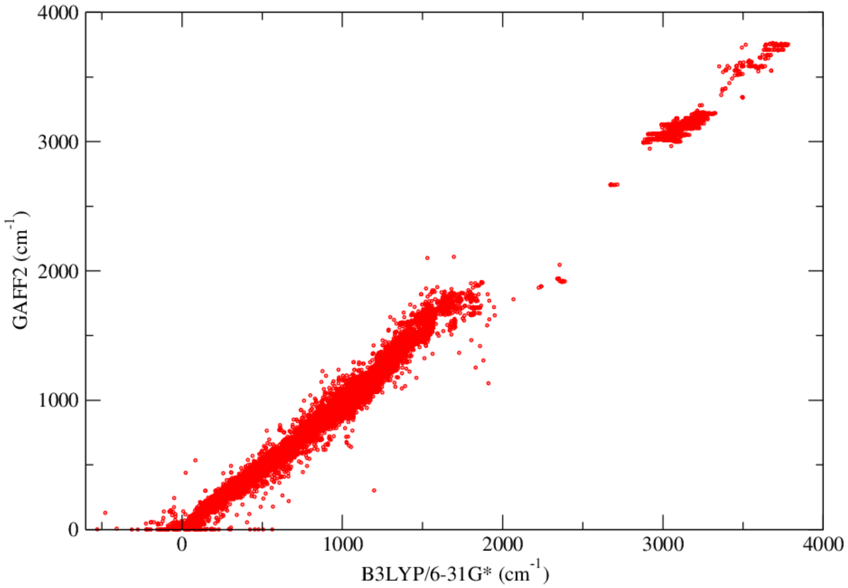


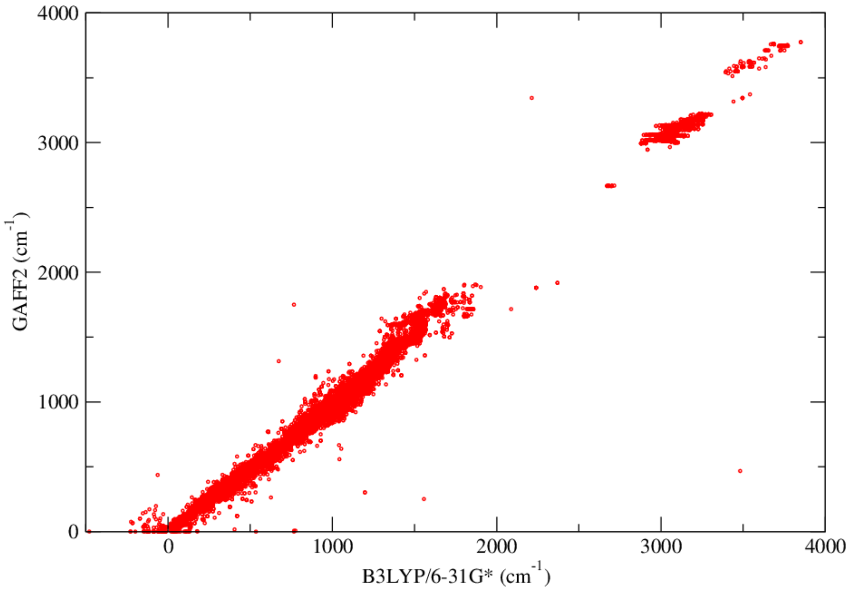
AUE = 0.97 kcal/mol
RMSE = 1.84 kcal/mol
ASE = -0.68 kcal/mol
Number of data points: 2834

AUE: Average unsigned error
RMSE: root-mean-squared error
ASE: average signed error

Bond Stretching And Bending Force Constants

$$V_{bond+angle} = \sum_{bonds} \boxed{K_r}\left(r - r_{eq}\right)^2 + \sum_{angles} \boxed{K_\theta}\left(\theta - \theta_{eq}\right)^2$$





Training Set

#data : 22407

GAFF2: AUE: 51.4, RMSE: 67.2 cm^{-1}

GAFF : AUE: 88.4, RMSE: 122.9 cm^{-1}

Test Set

#data : 29003

GAFF2: AUE: 48.4, RMSE: 63.0 cm^{-1}

GAFF : AUE: 88.0, RMSE: 123.3 cm^{-1}

Strategies of GAFF2 Parameterization

1. Partial Charge
2. **van der Waals**
    - Reproduce bulk properties, such as density and heat of vaporization
    - Reproduce high level interaction energies at MP2/aug-cc-pVTZ or above levels
3. Bond length and bond angle parameters
4. Torsional angle parameters

Objective Function

- Genetic algorithm is applied to optimize the r and ε to maximize Fitness

- Total fitness function calls: 200000 to 1 million

$$\boldsymbol{Fitness} = w_{IE}\frac{1}{RMSE\_IE} + w_{d}\frac{1}{APE\_d} + w_{Hvap}\frac{1}{RMSE\_Hvap}$$

Single point calculations using the *ab initio* optimized structures

MD simulations for liquids

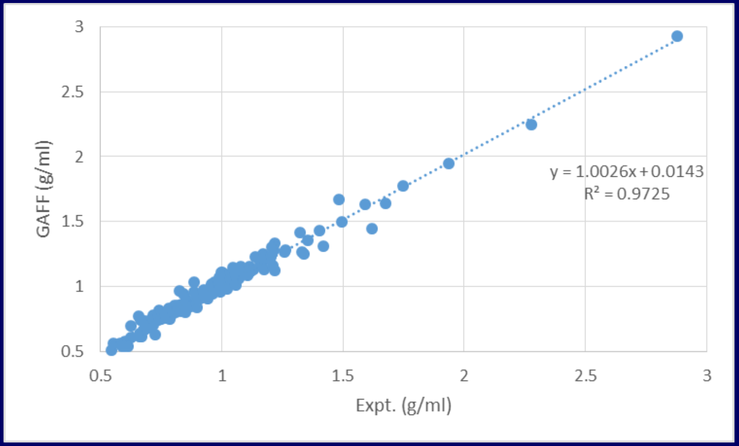
MD simulations for liquids

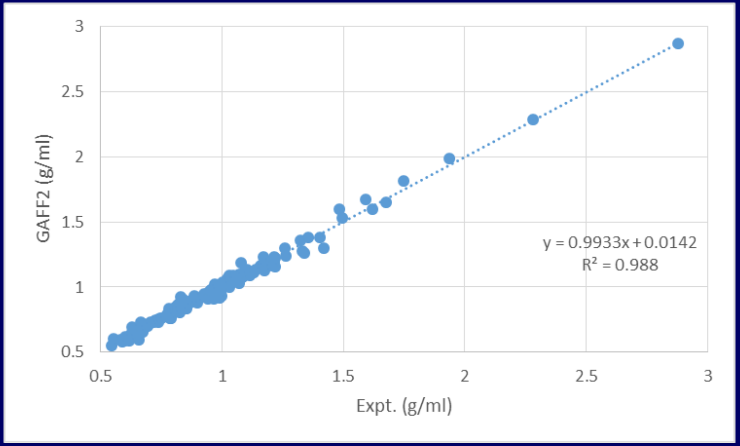
Liquid Densities

GAFF

GAFF2

Training Set : 158

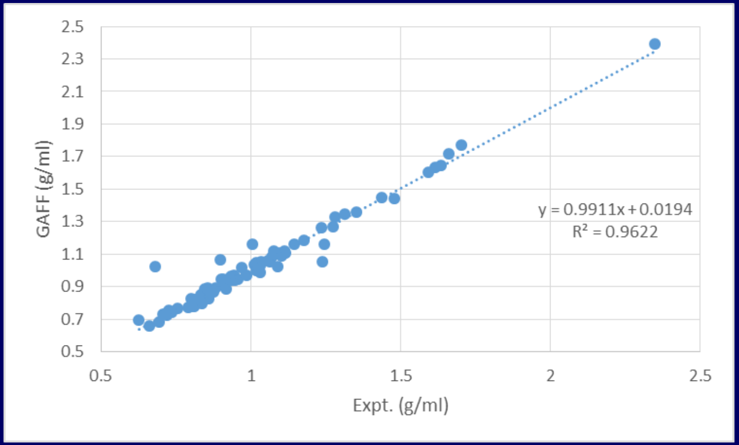


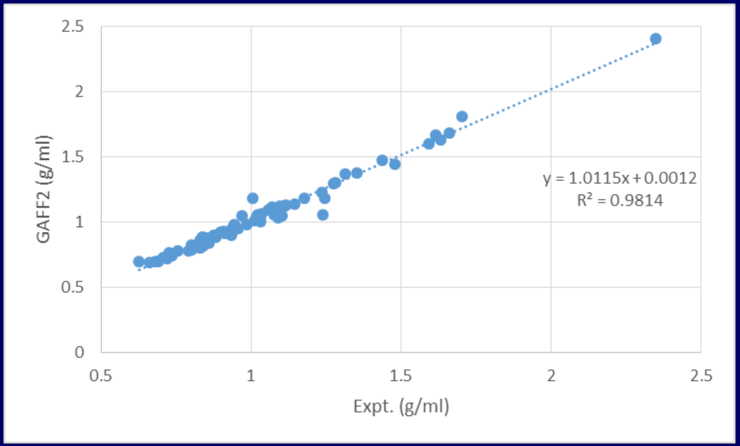


APE = 4.17%, RMSAPE = 5.56%

APE = **2.51** %, RMSAPE = 3.43%

Test Set : 83





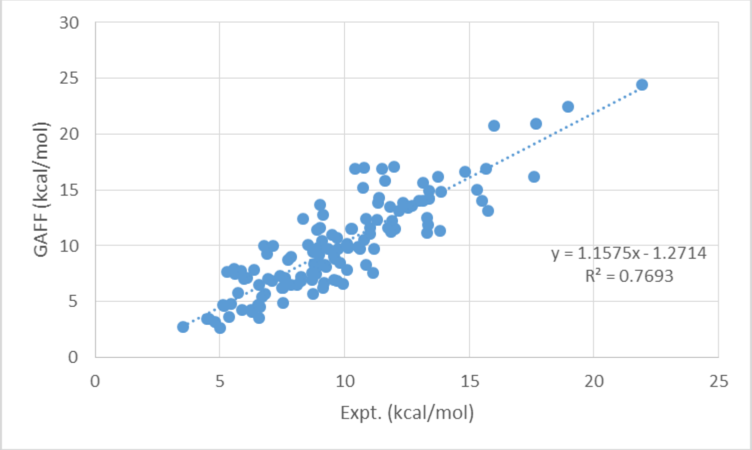
APE = 3.28%, RMSAPE = 6.92%

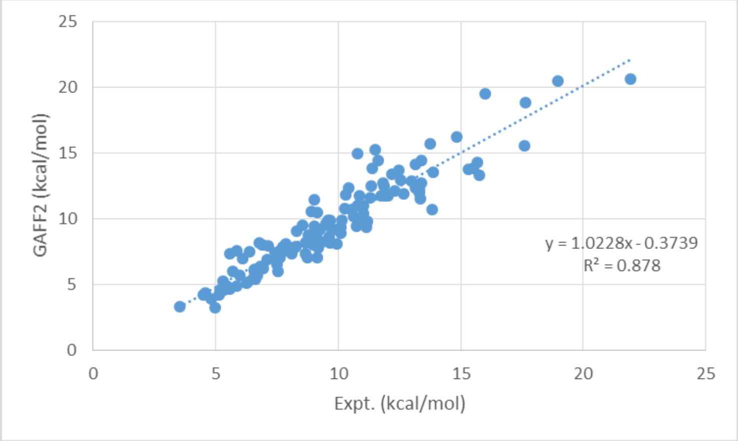
APE = **2.72**%, RMSAPE = 3.93%

Heat of Vaporization

GAFF **GAFF2**

Training Set : 129

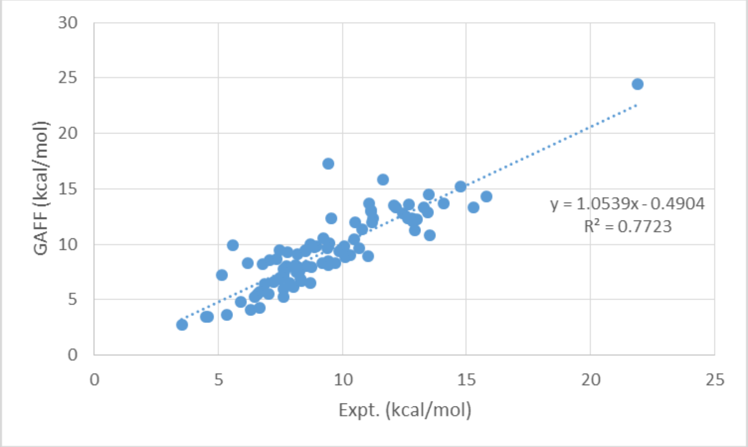


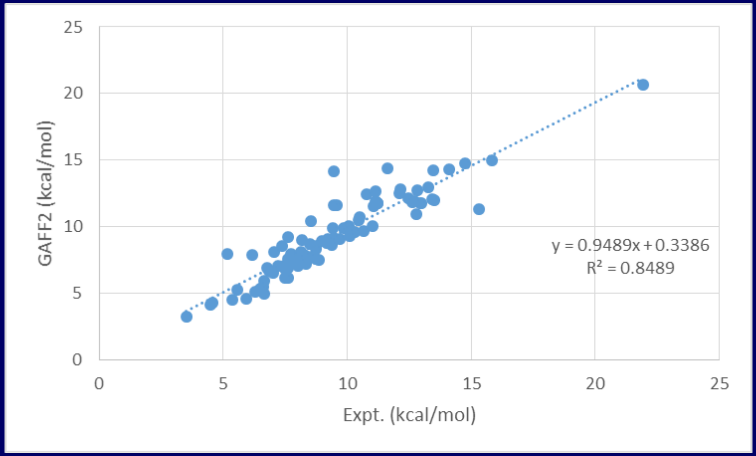


AUE = 1.69, RMSE = 2.13 kcal/mol

AUE = 0.97, RMSE = 1.11 kcal/mol

Test Set : 89





AUE = 1.28, RMSE = 1.67 kcal/mol

AUE = 0.86, RMSE = 1.18 kcal/mol

Polarizable Molecular Mechanical Force Field

Limitation of additive force fields

The use of Coulomb's law with fixed atomic charges to treat the electrostatic interactions is a major simplification in current force fields.

It is well known that the electron distribution of a molecule (and, thus, the atomic charges) changes as a function of the electrostatic field around the molecule. This is ignored in additive force fields.

To compensate for this omission, the atomic charges are "enhanced" to mimic the polarization of molecules that occurs in a polar, condensed phase environment (e.g. aqueous solution, TIP3P water model dipole moment = 2.35 versus gas phase value of 1.85).

Polarizable Force Field Development

1. Dipole Interaction – Based Polarizable Force Field

- Describe electrostatic interaction in a more physical fashion
- Suitable to study heterogeneous systems, dielectric continually changed system, highly charged systems.

$$V_{potential} = \sum_{\text{bonds}} K_r \left(r - r_{eq}\right)^2 + \sum_{\text{angles}} K_\theta \left(\theta - \theta_{eq}\right)^2 + \sum_{\text{dihedrals}} \frac{V_n}{2}\left[1 + \cos\left(n\varphi - \gamma\right)\right] +$$

$$\sum_{\substack{i<j \\ \text{van der Waals}}} \left(\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}}\right) + \sum_{\substack{i<j \\ \text{electrostatic}}} \frac{q_i q_j}{R_{ij}} \quad \boxed{+V_{pol}}$$

Dipole Interaction Models – Thole Scheme

$$\mathrm{V_{pol}} = -1/2 \sum_{p} \mu_p E_p{}^0$$

Dipole interaction model

$$\mu_p = \alpha_p [\boldsymbol{E_p^0} - \sum_{q \neq p}^{N} \boldsymbol{T}_{pq}]$$

$$E_p^0 = \sum_{q \neq p} \frac{q_q \mathbf{r}_{pq}}{r_{pq}^3}$$

Dipole field tensor

$$\boldsymbol{T}_{pq} = \frac{f_e}{r_{pq}^3} \boldsymbol{I} - \frac{3 f_t}{r_{pq}^5} \begin{bmatrix} x^2 & xy & xz \\ yx & y^2 & yz \\ zx & zy & z_2 \end{bmatrix}$$

Linear

$$v = r_{pq} / \left[a(\alpha_p \alpha_q)^{1/6} \right]$$

$$if\,(v >= 1) \quad f_e = 1.0, f_t = 1.0$$

$$if\,(v < 1) \quad f_e = 4v^3 - 3v^4, f_t = v^4$$

Exponential

$$v = r_{pq} / \left[a(\alpha_p \alpha_q)^{1/6} \right]$$

$$f_e = 1 - \left(\frac{v^2}{2} + v + 1 \right) \exp(-v)$$

$$f_t = 1 - (\frac{1}{6} v^3 + \frac{1}{2} v^2 + v + 1) \exp(-v)$$

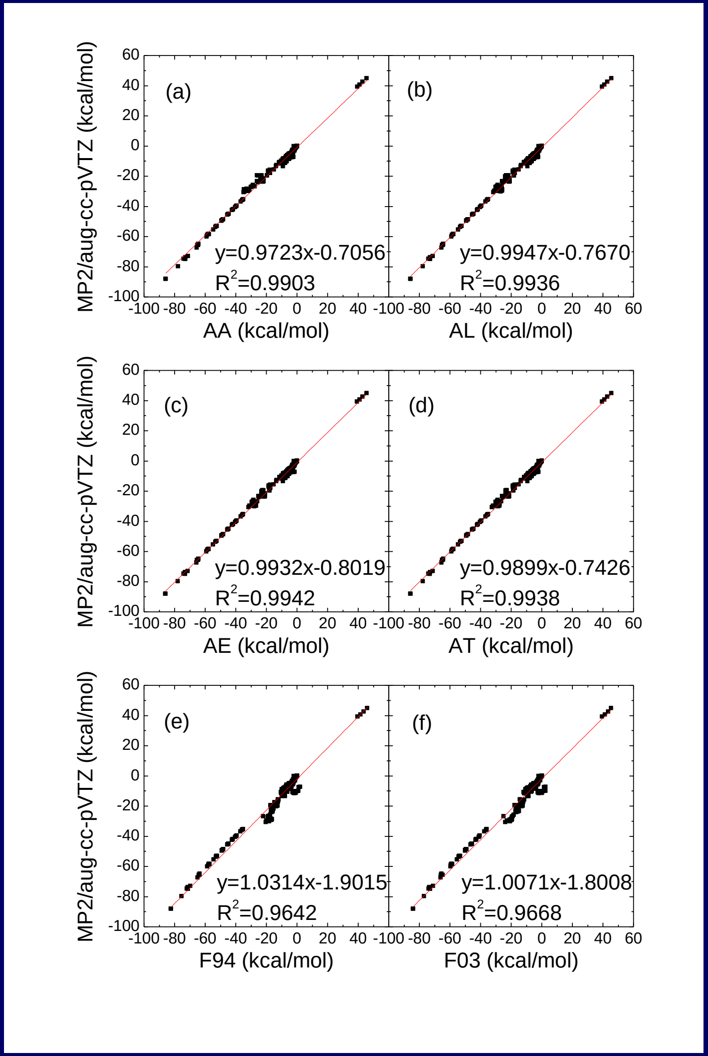
Amoeba-like

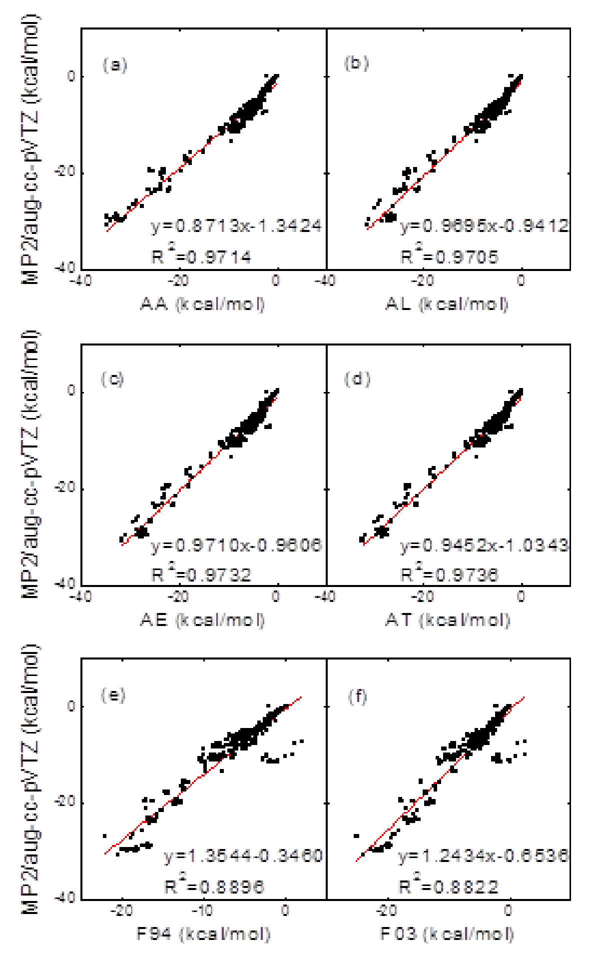
$$v = r_{pq} / \left[a(\alpha_p \alpha_q)^{1/6} \right]$$

$$f_e = 1 - \exp(-v^3)$$

$$f_t = 1 - (v^3 + 1) \exp(-v^3)$$

Comparison of Polarizable and Additive Force Fields





Linear: AUE = 1.02, RMSE = 1.56, Amoeba-like: AUE=1.05, RMSE=1.66

J. Phys. Chem. B, 115, 3091-3099**, 2011**; J. Phys. Chem. B, 115, 3100-3111