May 24, 2016
Molecular Modeling and Simulation

Potential Energy

Agricultural Bioinformatics Research Unit,
Graduate School of Agricultural and Life
Sciences, The University of Tokyo

Tohru Terada

Schedule

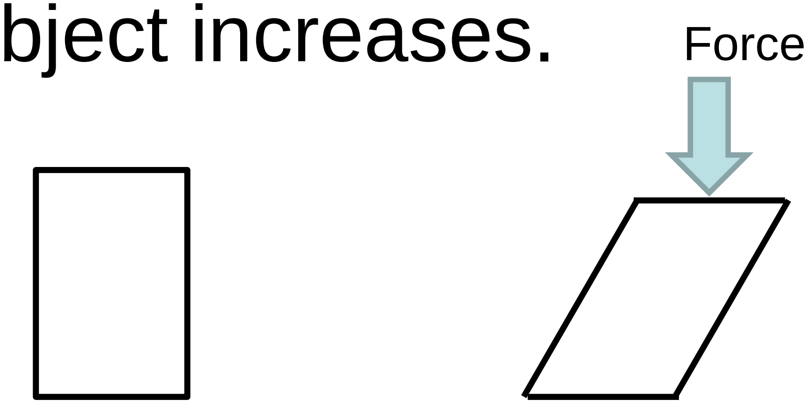
1.	May 24	Potential Energy
2.	May 31	Molecular Dynamics and Monte Carlo Methods
3.	June 7	Applications of Molecular Dynamics Methods
4.	June 14	Docking Simulations

Contents

- Practice of molecular orbital (MO) method
  - Exercise 1
- Molecular mechanics
- Energy minimization
- Practice of molecular mechanics
  - Exercise 2

Conformational energy

- When you exert a force on a object, the object is deformed and the “potential energy” of the object increases.



- Similarly, when a molecule changes its structure, its potential energy changes.

Potential energy of a molecule

- Potential energy of a molecule can be calculated by solving the Schrödinger equation approximately with molecular orbital (MO) method.

$$\hat{H}_{\mathrm{elec}} = -\sum_{i=1}^{N} \frac{1}{2} \nabla_i^2 - \sum_{i=1}^{N} \sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{1}{r_{ij}}$$

$$\hat{H}_{\mathrm{elec}} \Phi = E_{\mathrm{elec}} \Phi$$

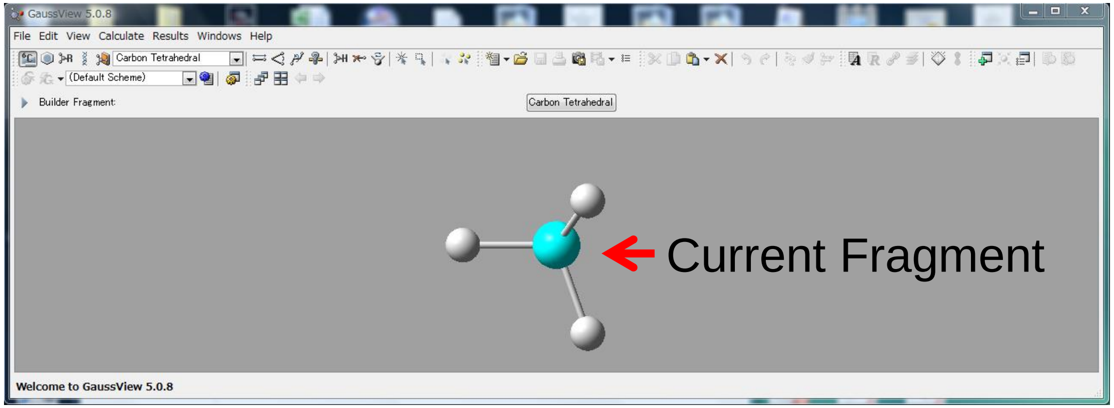
$$\boxed{E = E_{\mathrm{elec}} + \sum_{A=1}^{M} \sum_{B>A}^{M} \frac{Z_A Z_B}{R_{AB}}}$$

Potential energy of the molecule

N: Number of electrons
M: Number of atoms
Z_A: Atomic number of atom *A*
Φ: Wave function

Practice of MO method (1)

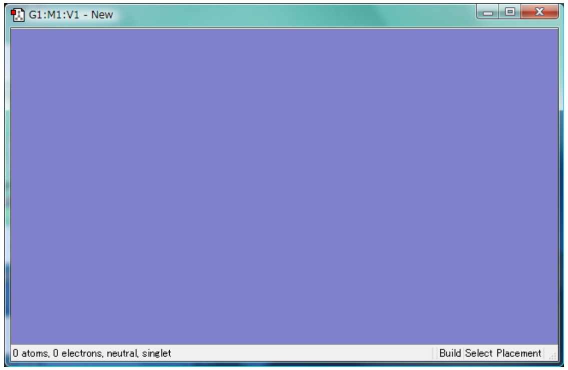
- In this lecture, we use quantum chemical calculation software, Gaussian 09W.
- Double-click the icon to start the graphic user interface (GUI) of this software, GaussView 5.0.



Current Fragment

Welcome to GaussView 5.0.8

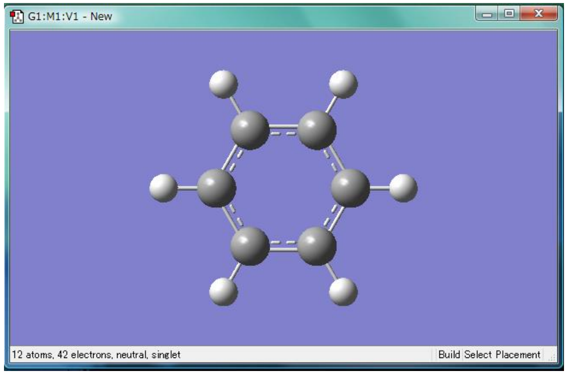
Control Panel



0 atoms, 0 electrons, neutral, singlet

Molecule View Window

Practice of MO method (2)



12 atoms, 42 electrons, neutral, singlet

1. Click Ring Fragment button in the Control Panel, select benzene, and click inside of Molecule View Window
2. Choose “Calculate” → “Gaussian Calculation Setup…” from the menu of the Control Panel
3. Set “Energy” to Job type, Set “6-31G(d)” to Basis set in the Method tab, and click “Submit.”
4. Save the input file as ”benzene.gjf” on Desktop.
5. Click “OK” when Run Gaussian Window appears.

Practice of MO method (3)

6. When the calculation finishes, click “Yes” to close Gaussian window.
7. Select “benzene.log” in the Gaussian Job Completed window and click “OK.”
8. Choose “Results” → “Summary” from the menu of the Control Panel. Potential energy is shown as E(RHF)[*].

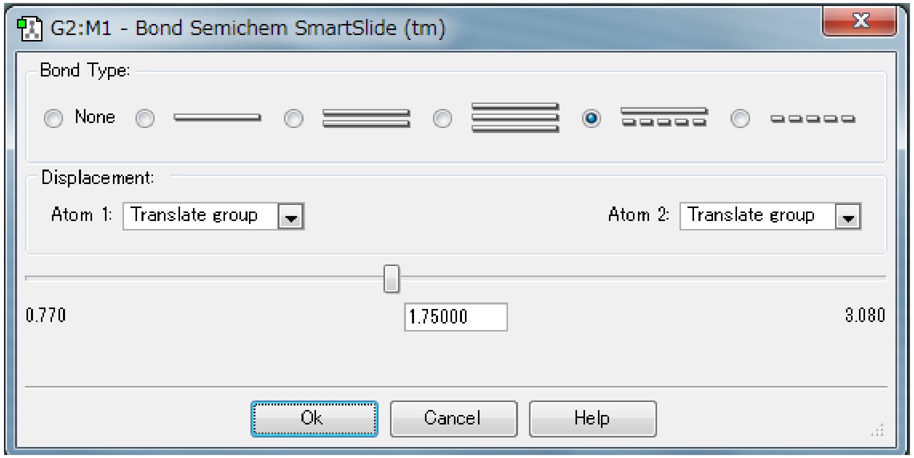
[*] 1 a.u. = 627.509 391 kcal/mol

Technical terms

- Method:
  - Hartree-Fock: the most basic *ab initio* method
  - Semi-empirical: applicable to a large system, but less accurate
  - DFT: more accurate and computationally not expensive. Electron correlation effect is partly considered.
- Basis Set:
  - STO-3G, 3-21G, 6-31G, 6-311G, etc, in order of complexity. The more complex basis set can express the molecular orbital the more accurately.
  - Diffuse (+, ++) and/or polarization [(d), (d,p), etc] functions are added as required.

Change the structure

- After clicking Modify Bond button in the Control Panel, click two neighboring carbon atoms of the benzene molecule.
- Change the bond length.



G2:M1 - Bond Semichem SmartSlide (tm)

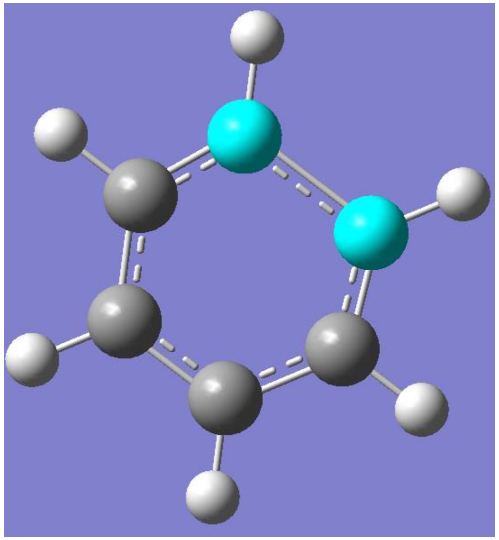
Bond Type: ○ None ○ ○ ○ ● ○

Displacement:

Atom 1: Translate group

Atom 2: Translate group

0.770 1.75000 3.080



- Calculate the potential energy.
  (Save the input file as benzene2.gjf on Desktop.)

Structural energy

- Examine change of energy by changes in
  - Bond lengths
  - Angles
  - Dihedral angles
  - Interatomic distances between atoms that are not covalently bonded

using model molecules.

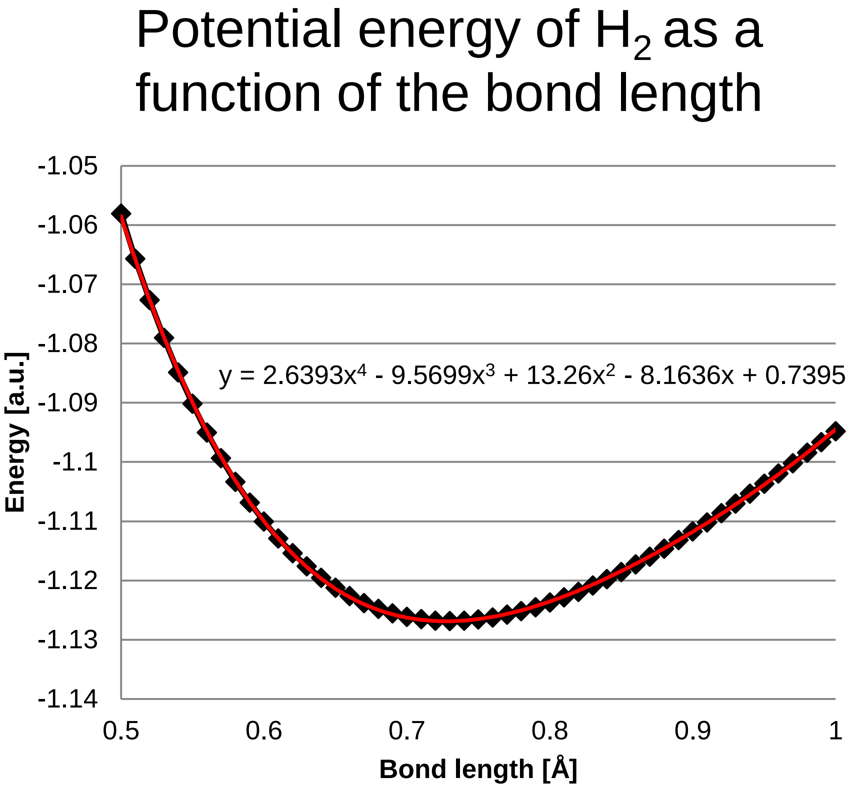
Diatomic molecule (1)

1. Download H2.gjf from this lecture’s web page.
2. Choose “すべてのプログラム” → “Gaussian 09W” → “Gaussian 09W” from Start menu to start Gaussian 09W.
3. Choose “File” → “Open” from the menu to open H2.gjf.
4. Choose “File” → “Exit & Run” from the menu to start the calculation.
5. Save the output file as H2.out on Desktop.

Diatomic molecule (2)

6. When the calculation has finished, choose “File” → “Exit” from the menu.
7. Start GaussView 5.0. Choose “File” → “Open” from the menu of the Control Panel to open H2.out. (Set ”Gaussian Output Files (*.out *.log)” to ファイルの種類.)
8. Choose “Results” → “Scan” from the menu of the Control Panel.
9. Right-click within the Scan plot window, choose “Save Data” from the menu, and save the data as H2_scan.txt on Desktop.

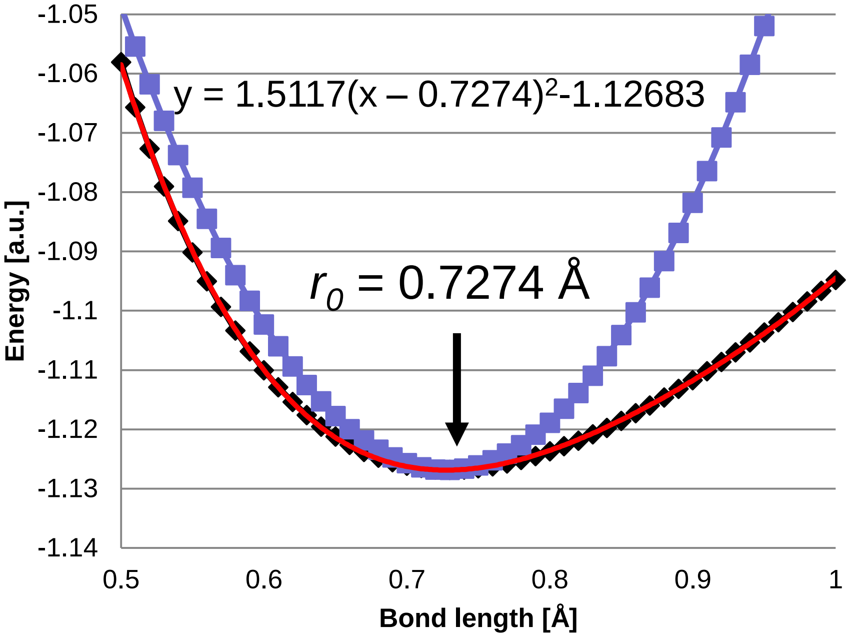
Diatomic molecule (3)



- Open H2_scan.txt with Excel and plot the potential energy against the bond length.
- You find the plot is well approximated with a fourth-order polynomial.
- Potential energy can be calculated at low computational cost with this polynomial function.

Potential energy function

Potential energy of H_2 as a function of the bond length



- The bond length thermally fluctuates about the length at the energy minimum.
- In biomolecular simulations, the temperature is about 300 K.
  →$kT = 0.6\ \mathrm{kcal\ mol}^{-1} = 10^{-3}$ a.u.
- The width of the fluctuation is about 0.04 Å.

$$E(r_0+\Delta r)=E(r_0)+\left.\frac{\partial E(r)}{\partial r}\right|_{r_0}\Delta r+\frac{1}{2}\left.\frac{\partial^2 E(r)}{\partial r^2}\right|_{r_0}\Delta r^2+O(\Delta r^3)$$

$$E(r_0+\Delta r)-E(r_0)\approx k(r-r_0)^2$$

Approximated with quadratic function

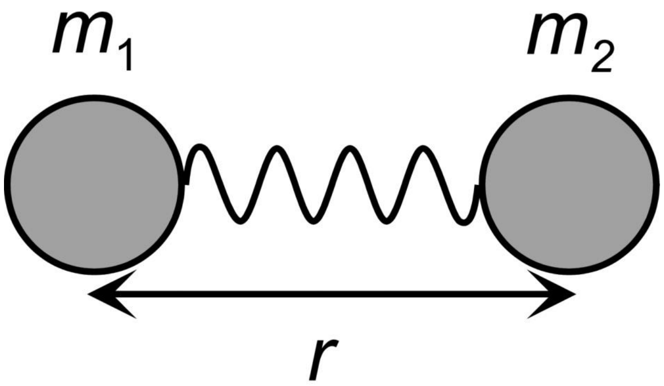
Taylor series

Taylor series of *f*(*x*) in the neighborhood of *p*.
$\Delta x = x - p$.

$$f(p+\Delta x) = f(p) + \left.\frac{df(x)}{dx}\right|_{x=p} \Delta x + \frac{1}{2!}\left.\frac{d^2 f(x)}{dx^2}\right|_{x=p} \Delta x^2 + \cdots + \frac{1}{k!}\left.\frac{d^k f(x)}{dx^k}\right|_{x=p} \Delta x^k + \cdots$$

$$= \underbrace{f(p) + \left.\frac{df(x)}{dx}\right|_{x=p} \Delta x + \frac{1}{2!}\left.\frac{d^2 f(x)}{dx^2}\right|_{x=p} \Delta x^2}_{\text{Truncated at the second order}} + \underbrace{O\left(\Delta x^3\right)}_{\substack{\text{Error} \\ \text{(order of } \Delta x^3\text{)}}}$$

Comparison of dynamics



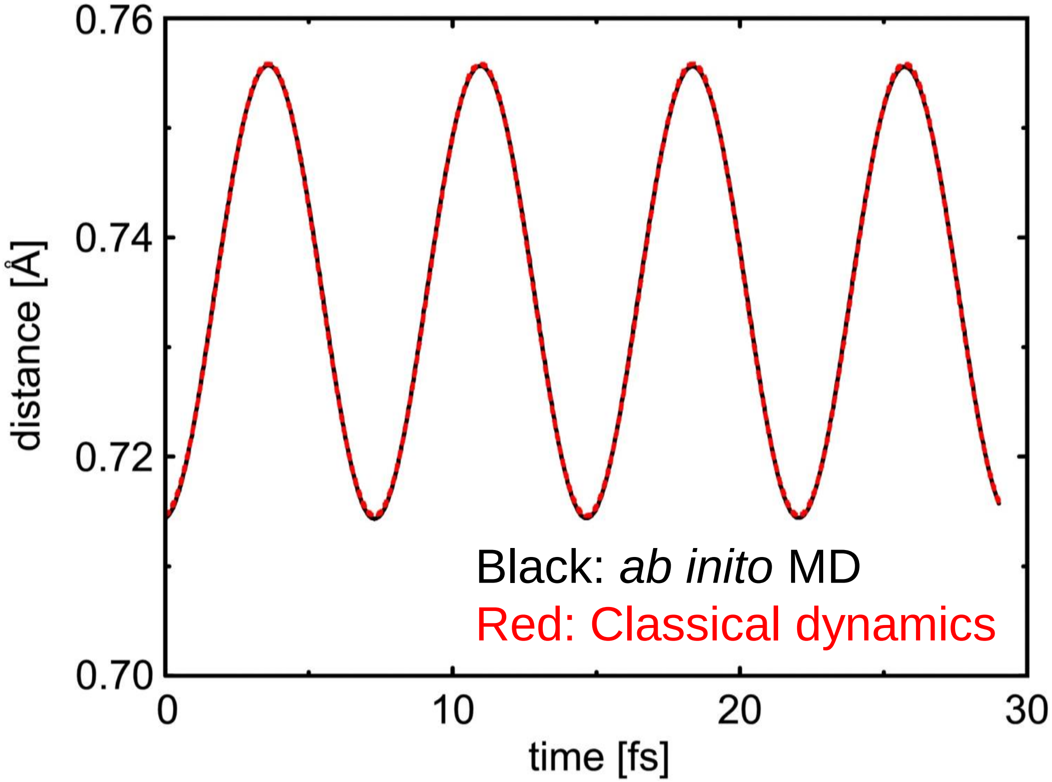
$$E(r) = k(r - r_0)^2$$

$$F(r) = -\frac{\partial E(r)}{\partial r} = -2k(r - r_0)$$

$$\mu \ddot{r} = -2k(r - r_0)$$

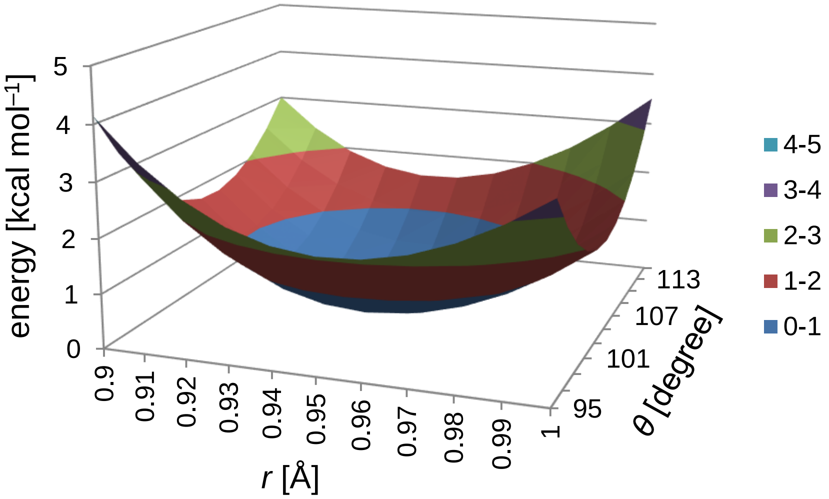
$$\omega = \sqrt{\frac{2k}{\mu}}, \quad \mu = \frac{m_1 m_2}{m_1 + m_2}$$

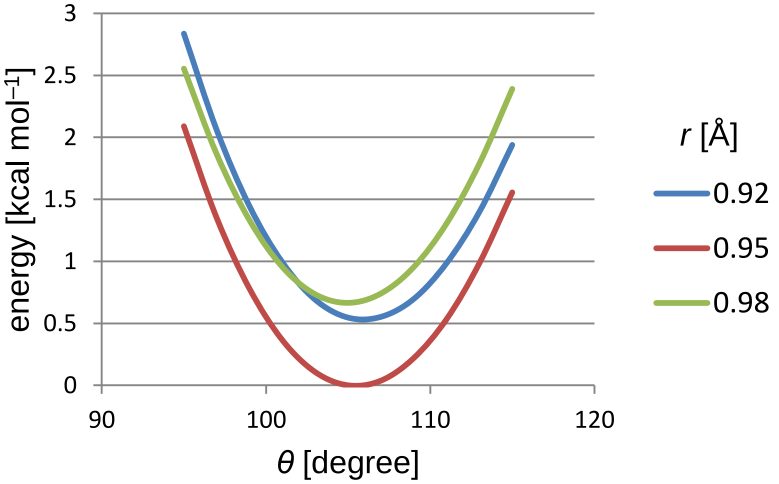
$$r(t) = A\cos(\omega t + \delta) + r_0$$



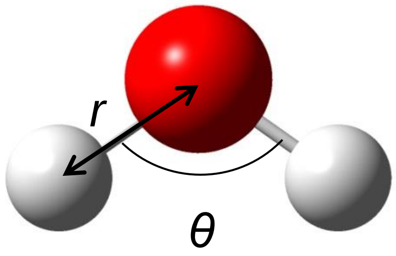
The trajectory of classical dynamics based on the potential energy function is almost the same as that of *ab inito* MD.

Energy of three-atom molecule





Model: water

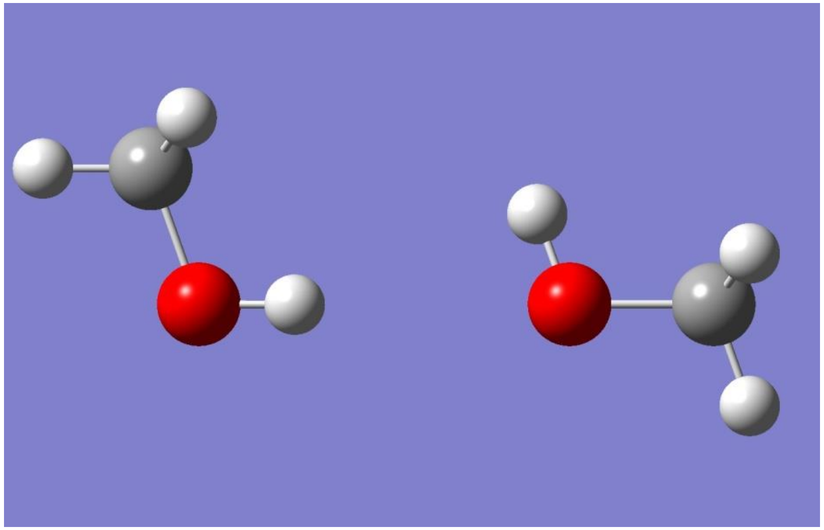


- The energy function is approximated as the sum of *r*-dependent and *θ*-dependent terms.

$$E(r,\theta) \approx k_b (r - r_0)^2 + k_a (\theta - \theta_0)^2$$

Intermolecular interaction (1)

1. Start GaussView 5.0. Click the icon of Element Fragment in the Control Panel, and select “O” to show “Oxygen Tetravalent” in Current Fragment window.
2. Click on two different points within Molecule View window.
3. Select “C” from the Element Fragments to show “Carbon Tetrahedral” in the Current Fragment window. Click on one of hydrogen atoms of H_2O to replace it with CH_3.
4. Similarly, replace one hydrogen atom of another H_2O with CH_3



Intermolecular interaction (2)

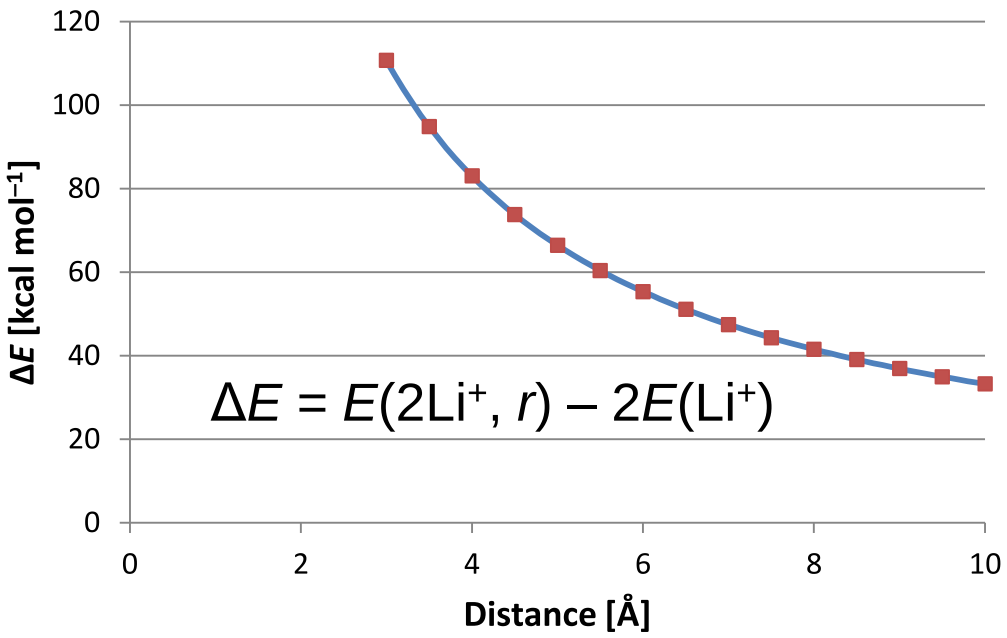
5. Choose “Calculate” →“Gaussian Calculation Setup…” from the menu of Control Panel. Set Job type to “Energy”, set Basis set of the Method tabt to “6-31G(d)”, and click “Submit.” Save the input file as “methanol2.gjf” on Desktop.
6. Similarly, calculate energy for a single molecule. Save the input file as “methanol1.gjf” on Desktop.
7. Calculate interaction energy as:

$$\Delta E = E_{\mathrm{AB}} - \left(E_{\mathrm{A}} + E_{\mathrm{B}}\right)$$

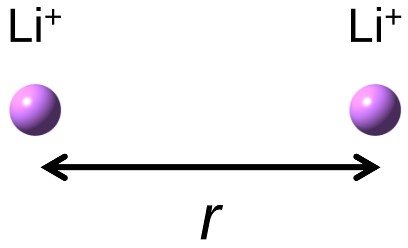
$$\Delta E = -230.0688122 - [2 \times (-115.0334869)]$$

$$= -0.0018384 \text{ a.u.} = -1.15 \text{ kcal mol}^{-1}$$

Interaction energy of ions



$\Delta E = E(2\mathrm{Li}^+, r) - 2E(\mathrm{Li}^+)$



The interaction energy ΔE is given by the E_{elec} term with $q_i = q_j = 1$.

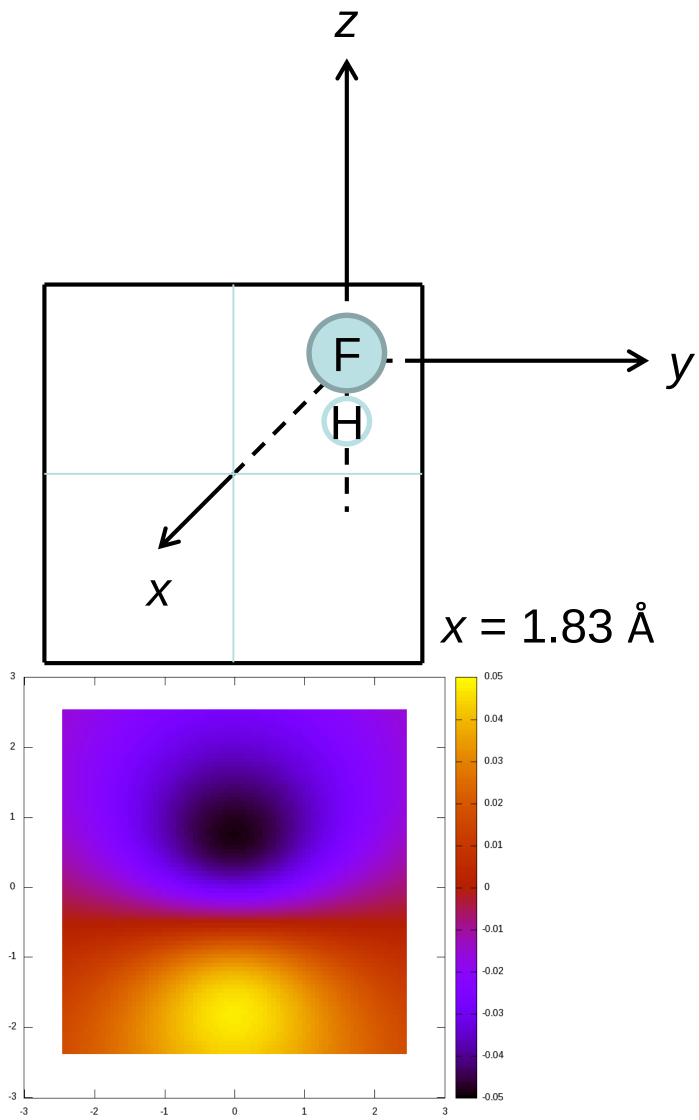
$$E_{\mathrm{elec}} = \frac{(q_i e)(q_j e) N_{\mathrm{A}}}{4\pi\varepsilon_0 r_{ij} (10^{-10}\,\mathrm{m})} = (332.06\,\mathrm{kcal\,mol^{-1}}) \frac{q_i q_j}{r_{ij}}$$

Electrostatic potential (1)

- Electrostatic energy of a particle is calculated as the product of electrostatic potential (ESP) at the position and the charge of the particle.
- ESP is calculated from the charge distribution.

$$\varphi(\mathbf{r}) = \int d\mathbf{r}' \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}$$

- The right figure shows the ESP of a HF molecule.

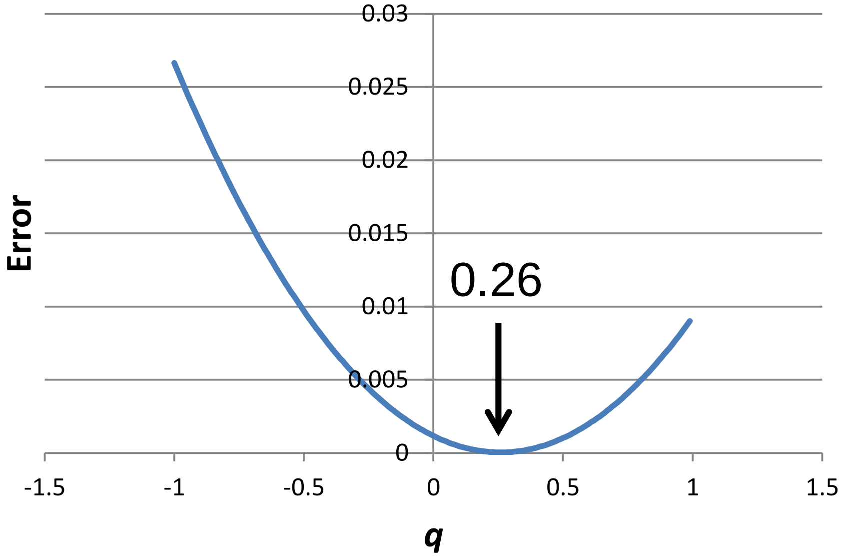


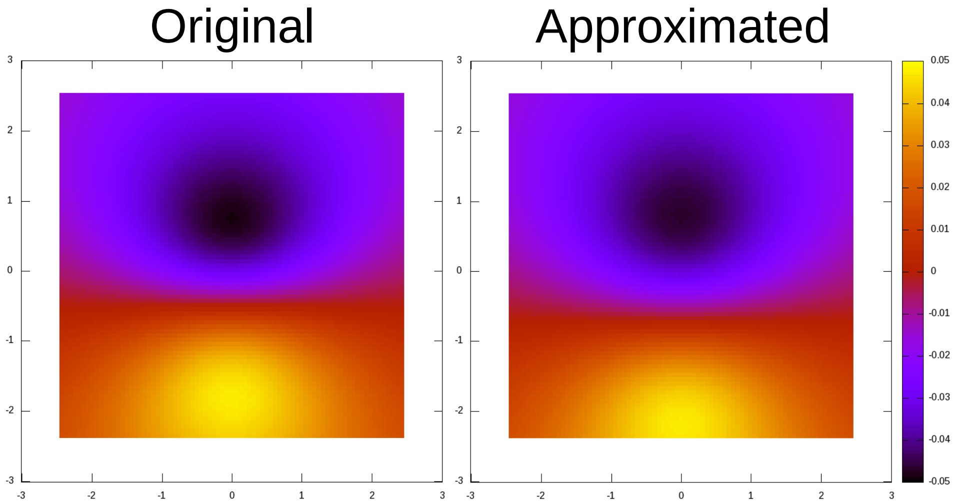
Electrostatic potential (2)

- ESP generated by a molecule is approximated by that generated by point charges placed at the centers of its atoms.

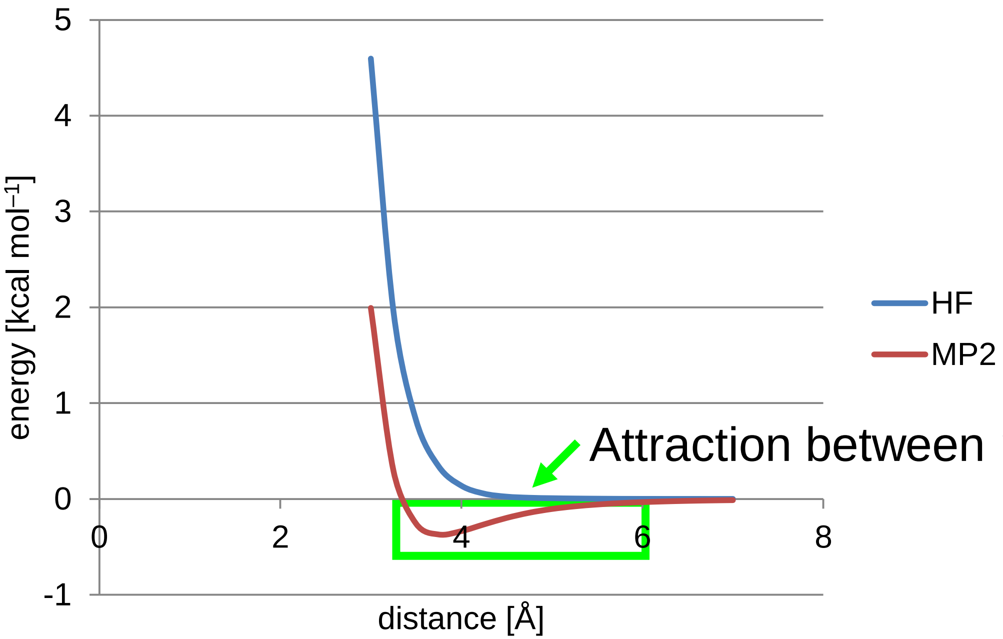
$$\varphi(\mathbf{r}) \approx \sum_i \frac{q_i}{|\mathbf{r}-\mathbf{r}_i|}$$

- Charges are determined to minimize the error between the original and the approximated ESPs.
    - Charge of H: 0.26
    - Charge of F: –0.26

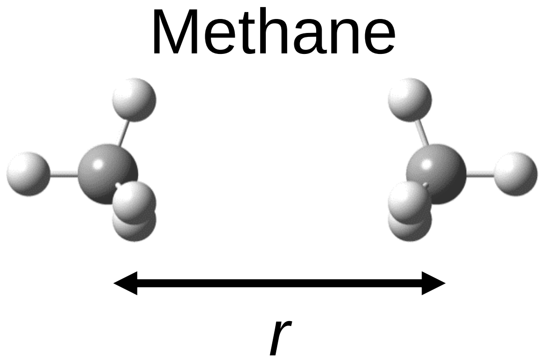




Van der Waals interactions



Attraction between methane molecules

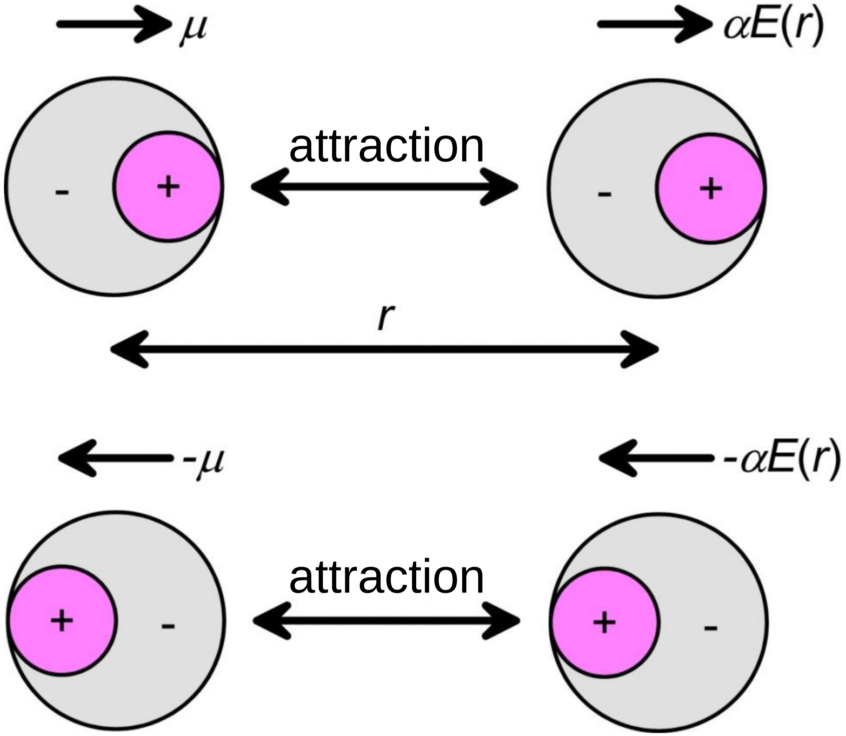


The use of the method that considers the electronic correlation effect is necessary to calculate the attraction between nonpolar molecules.

Origin of the attraction

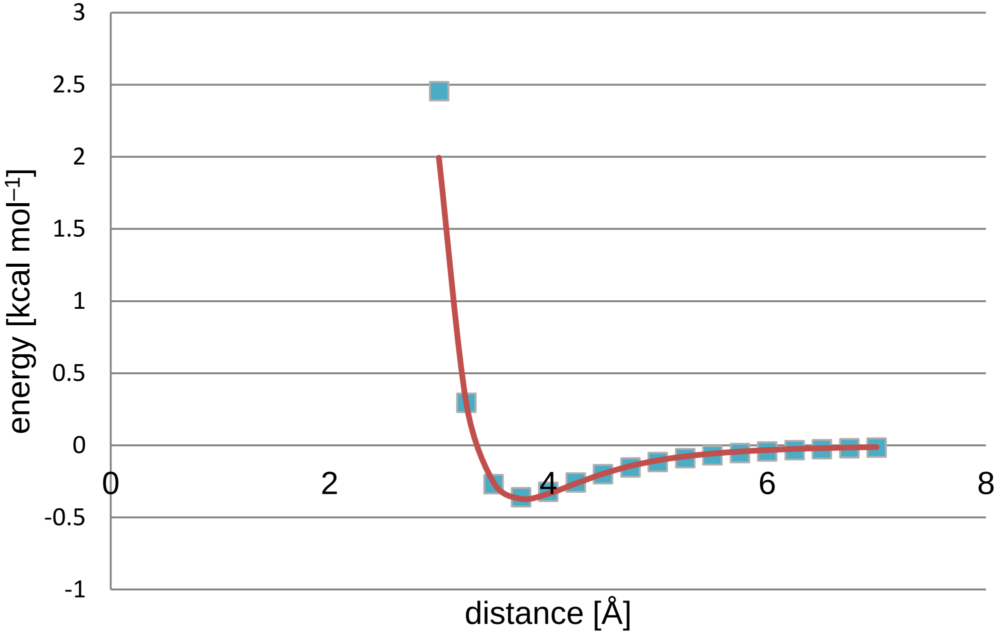
Fluctuation of the electron cloud produce an instantaneous electric dipole moment even on a nonpolar molecule, which induces an electric dipole on nearby molecules.

The interaction between the dipole and the induced dipole is always attractive.



$<\mu>=0$ $<E(r)>=0$

Van der Waals energy function



$$E_{\mathrm{vdW}}(r) = 4\varepsilon\left[\underbrace{\left(\frac{\sigma}{r}\right)^{12}}_{\text{Repulsion}} - \underbrace{\left(\frac{\sigma}{r}\right)^{6}}_{\text{Attraction}}\right]$$

At $r = \sigma$, $E_{\mathrm{vdW}} = 0$

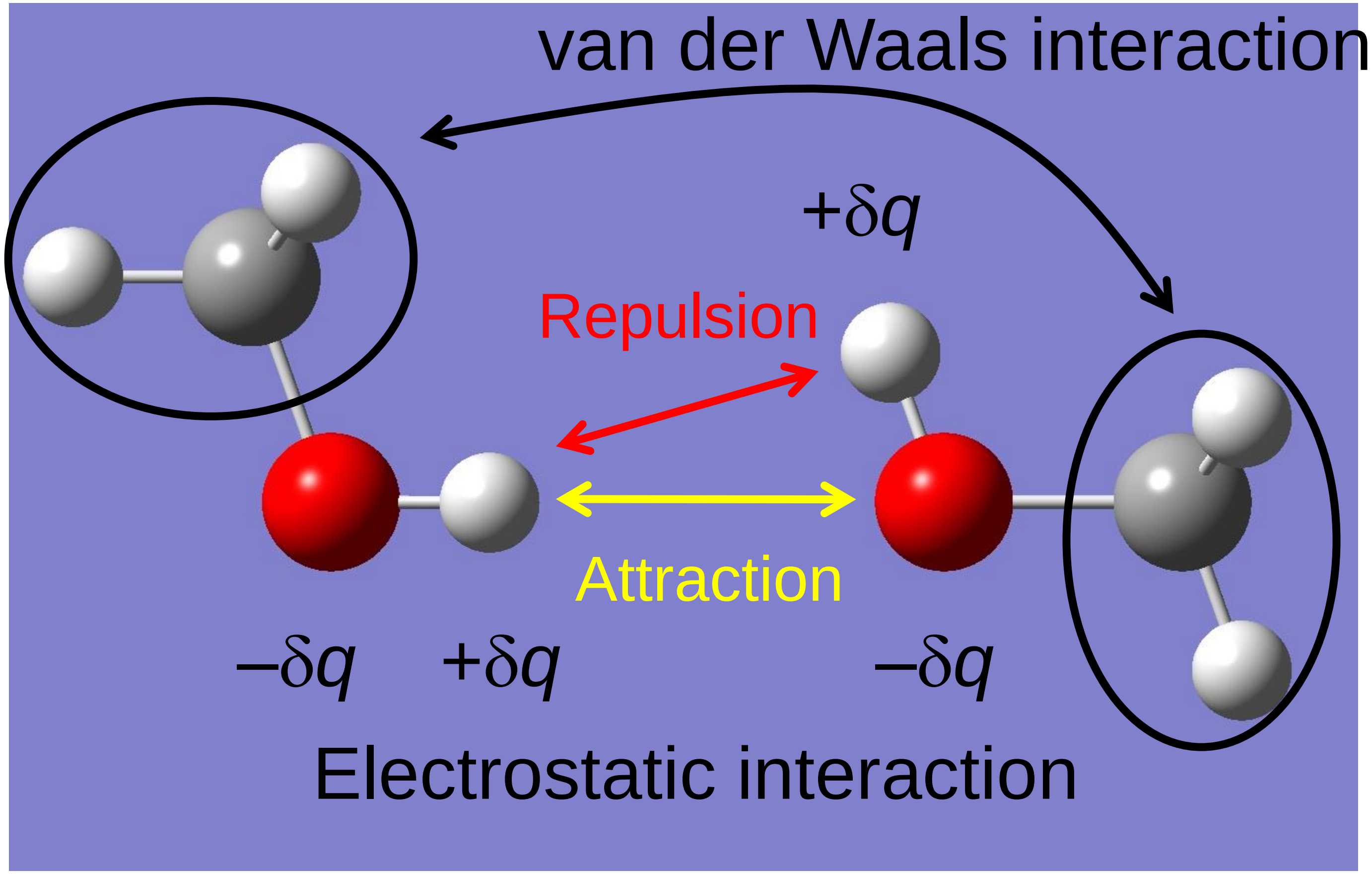
At $r = 2^{1/6}\sigma$, $E_{\mathrm{vdW}} = -\varepsilon$

(Minimum)

$\sigma = 3.34$ Å

$\varepsilon = 0.36$ kcal mol^{-1}

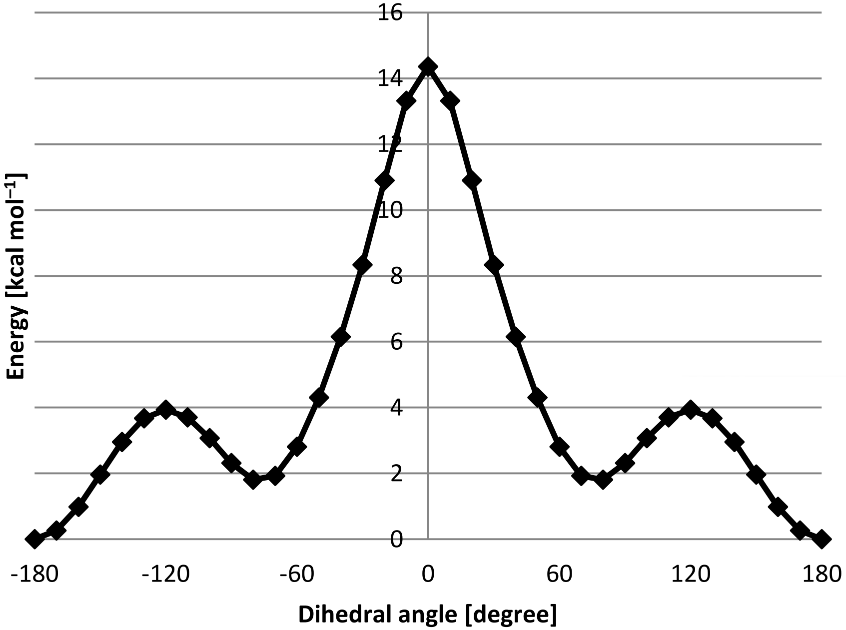
Component of interactions

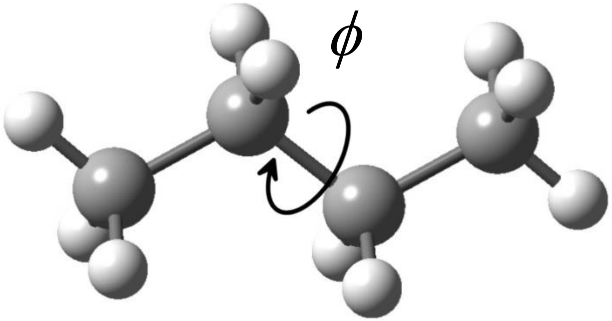


Dihedral angle energy

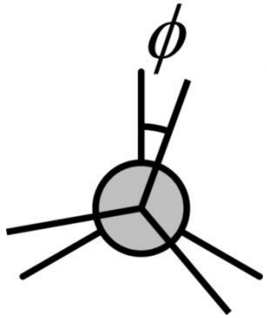
Model: butane

Energy of butane as a function of dihedral angle





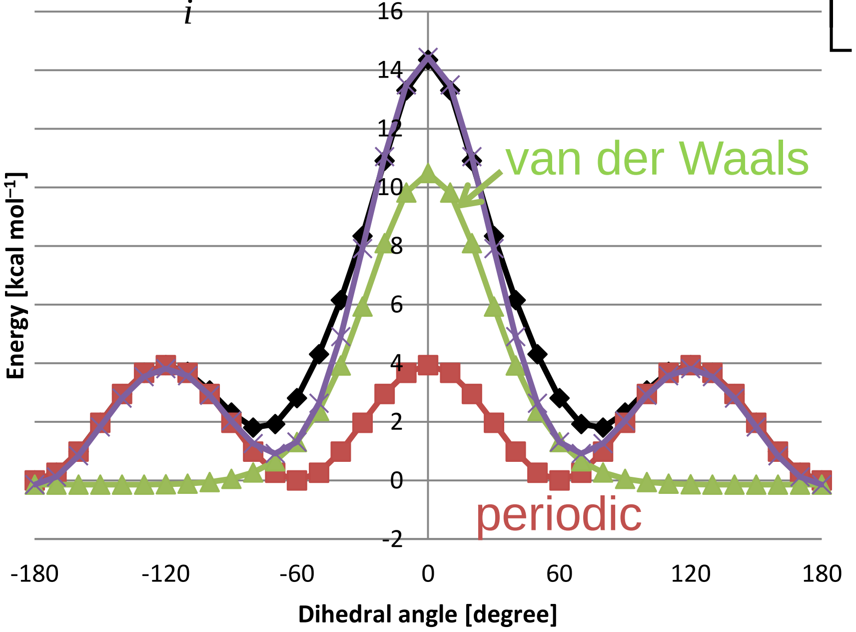
Gaussian job file: butane.gjf

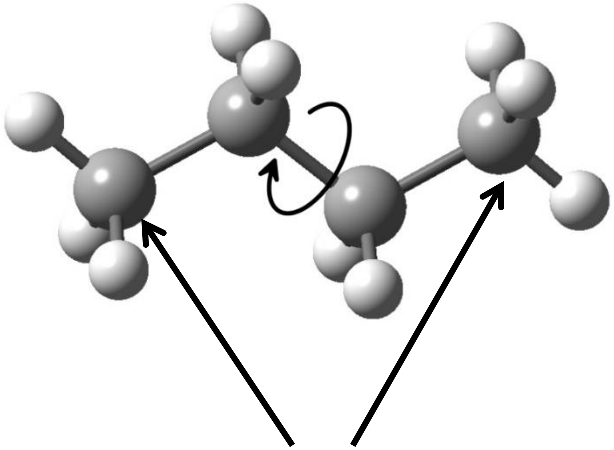


Dihedral angle energy function

- Energy function of butane is expressed as the sum of a periodic function and van der Waals energy function.

$$E(\phi) = \sum_i k_i \left[1 + \cos\left(n_i \phi - \delta_i^0\right)\right] + 4\varepsilon \left[\left(\frac{\sigma}{r_{14}}\right)^{12} - \left(\frac{\sigma}{r_{14}}\right)^{6}\right]$$





Only the van der Waals interaction between the first and fourth methyl groups are considered.

Exercise 1

- Determine n, δ, ε, and σ of the potential energy function of butane.
  - Download “butane_scan.xlsx” from this lecture’s web page.
  - Choose “Plot”→“Plot Molecular Property” from the menu of Scan Plot window and select “Bond” from the pull-down menu and set atom indices to “1” and “4” to obtain the list of the distances between the first and fourth carbon atoms.

Molecular mechanics

- QM calculation demands large computational cost.

⇩

- The computational cost can be reduced by using potential energy function that reproduces QM energy values.

$$E(\mathbf{r}) = \overset{\text{①Bond}}{\boxed{\sum_b k_b \left(r_b - r_b^0\right)^2}} + \overset{\text{②Angle}}{\boxed{\sum_a k_a \left(\theta_a - \theta_a^0\right)^2}} + \overset{\text{③Dihedral}}{\boxed{\sum_d k_d \left[1 + \cos\left(n_d \phi_d - \delta_d^0\right)\right]}}$$

$$+ \underset{\text{④Non-bonded}}{\boxed{\sum_{ij} \left\{ 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6} \right] + \frac{q_i q_j}{r_{ij}} \right\}}}$$

Force field parameters

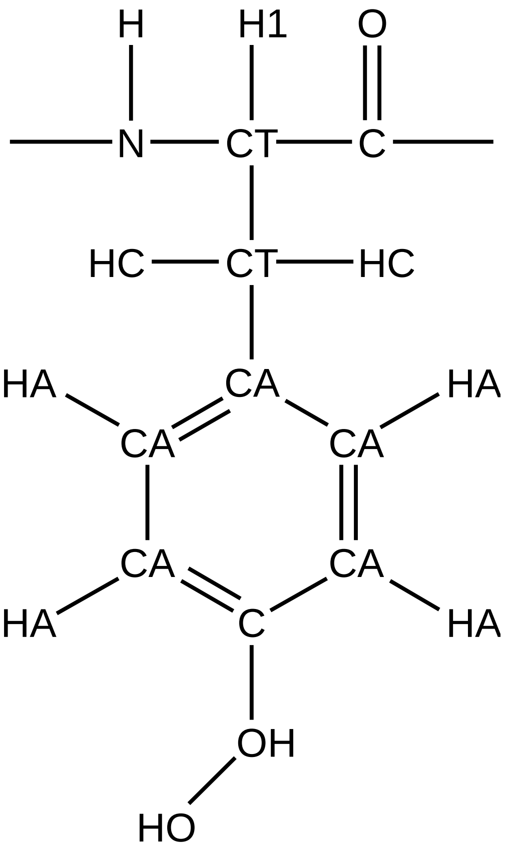
- Force field parameters
  - Parameters used in the potential energy function. (equilibrium bond length, force constant, partial charge, etc.)
- Ab initio force field parameters
  - Derived using QM results.
- Empirical force field parameters
  - Determined so as to reproduce the structure and thermodynamic quantities.

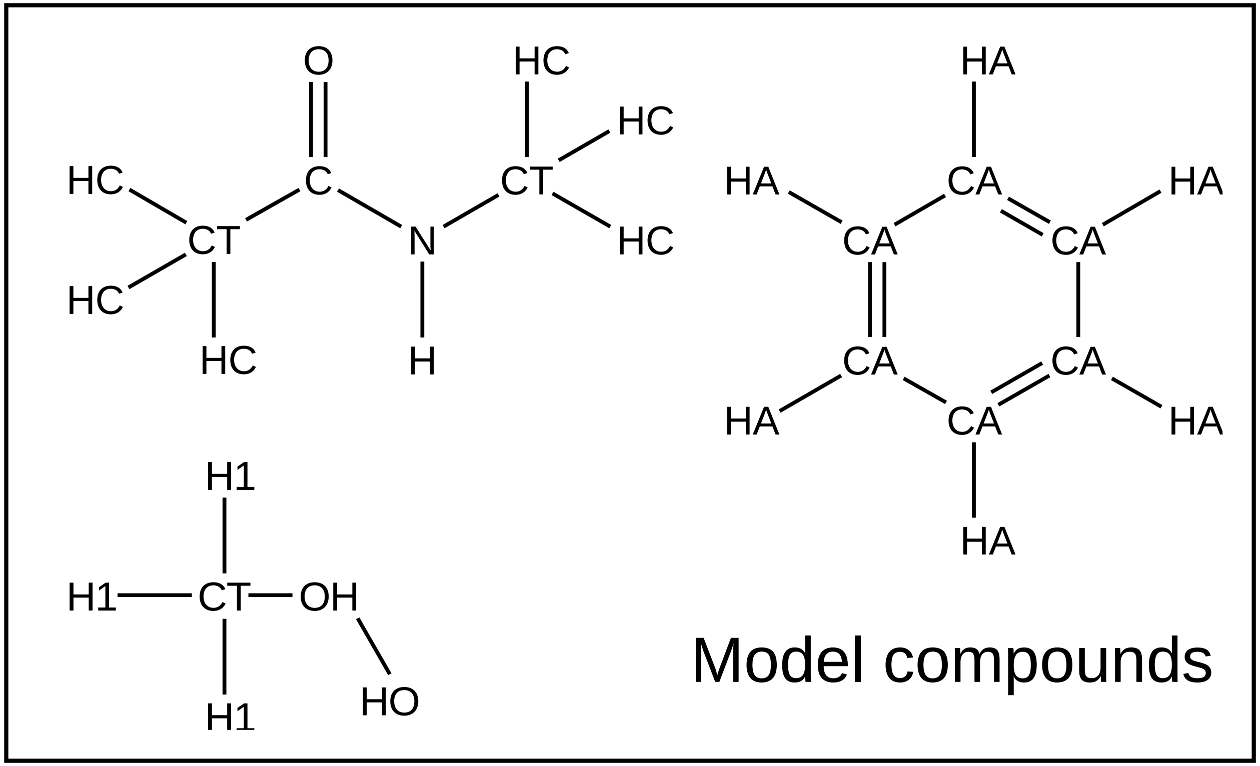
Problems and solutions (1)

- A biomacromolecule is composed of a large number of atoms.
  - It is difficult to apply the QM calculation to the whole molecule.

- Atoms are classified into atom types according to their chemical properties (type of hybridization or substituents). The same parameters are assigned to the atoms of the same atom type.
- Parameters are determined for model compound.

Atom types and model compounds



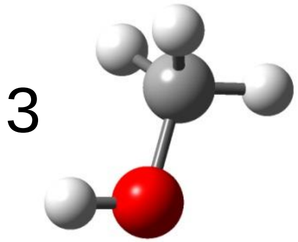


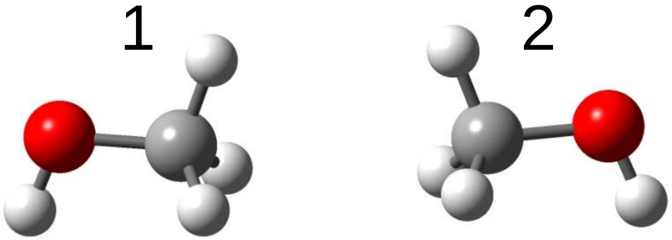
Model compounds

Jorgensen & Tirado-Rives, *J. Am. Chem. Soc.* **110**, 1657 (1988)

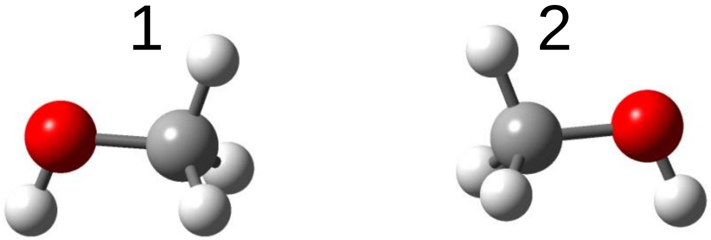
Problems and solutions (2)

- In condensed (liquid) phase, molecules are close to each other. The interaction between two molecules is affected by a third molecule near the molecules.
  - The parameters determined in gas phase are not applicable.





Gas phase

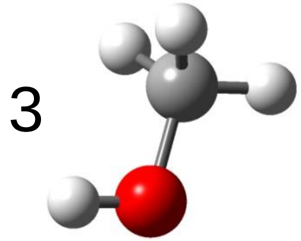


Condensed phase

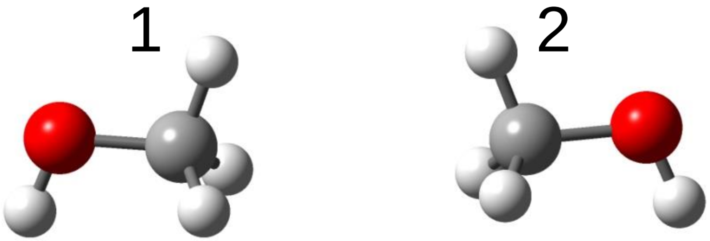
Effective potential energy

$$
\begin{aligned}
E(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3) &= E(\mathbf{r}_1,\mathbf{r}_2)+E(\mathbf{r}_1,\mathbf{r}_3)+E(\mathbf{r}_2,\mathbf{r}_3)+\Delta E(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3) \\
&= E^{\mathrm{eff}}(\mathbf{r}_1,\mathbf{r}_2)+E^{\mathrm{eff}}(\mathbf{r}_1,\mathbf{r}_3)+E^{\mathrm{eff}}(\mathbf{r}_2,\mathbf{r}_3)
\end{aligned}
$$

$E(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)$: energy of three-molecule system
$E(\mathbf{r}_1,\mathbf{r}_2)$: energy of two-molecule system
$E^{\mathrm{eff}}(\mathbf{r}_1,\mathbf{r}_2)$: effective pairwise potential



Effect of the rigorous many-body interaction is incorporated in the effective pairwise potential.

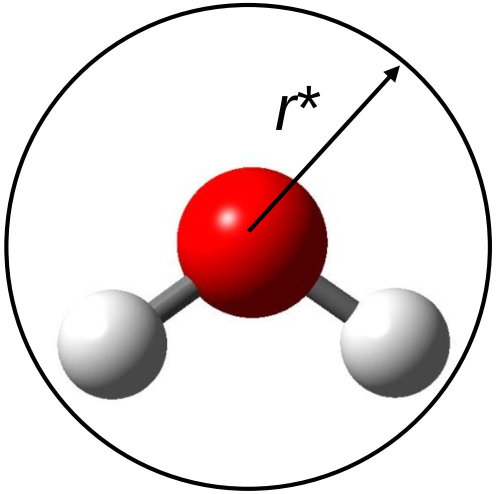


Condensed phase

Parameters are determined to reproduce experiments.

Water models (1)

	SPC	TIP3P
r(OH)	1.0	0.9572
$\angle$HOH	109.47	104.52
r^*	1.7766	1.7683
ε	0.1554	0.1520
q_H	0.41	0.417



Van der Waals interaction is calculated only between oxygen atoms.

r(OH) [Å], $\angle$HOH [degree]
r^* [Å^{12} kcal mol^{-1}], $2r^* = \sqrt[6]{2}\sigma$
ε [Å^6 kcal mol^{-1}]
$q_O = -2q_H$

Jorgensen *et al. J. Chem. Phys.* **79**, 926 (1983)

Water models (2)

	SPC	TIP3P	Expt.
d	0.971	0.982	0.997
ΔH_{vap}	10.77	10.45	10.51
C_p	23.4	16.8	17.99
α	58	41	25.7
κ	27	18	45.8

d: density [g cm^{-3}], ΔH_{vap}: heat of vaporization [kcal mol^{-1}]
C_p: molar isobaric specific heat [cal mol^{-1} K]
α: coefficient of thermal expansion [10^{-5} K^{-1}]
κ: compressibility [10^{-6} atm^{-1}]
All values are at 25 ˚C, 1 atm.

Jorgensen *et al. J. Chem. Phys.* **79**, 926 (1983)

Force field parameter sets

- Force field parameter sets are distributed along with molecular simulation software.
- AMBER
  - http://www.ambermd.org/
- CHARMM
  - http://www.charmm.org/
- GROMOS, GROMACS
  - http://www.igc.ethz.ch/gromos/
  - http://www.gromacs.org/

Energy minimization (1)

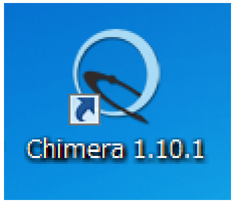
- Changes the conformation to the one having lower potential energy.
- Also called “structure optimization.”
- Must be done in advance of the molecular dynamics simulation to avoid atomic overlaps.

Energy minimization (2)

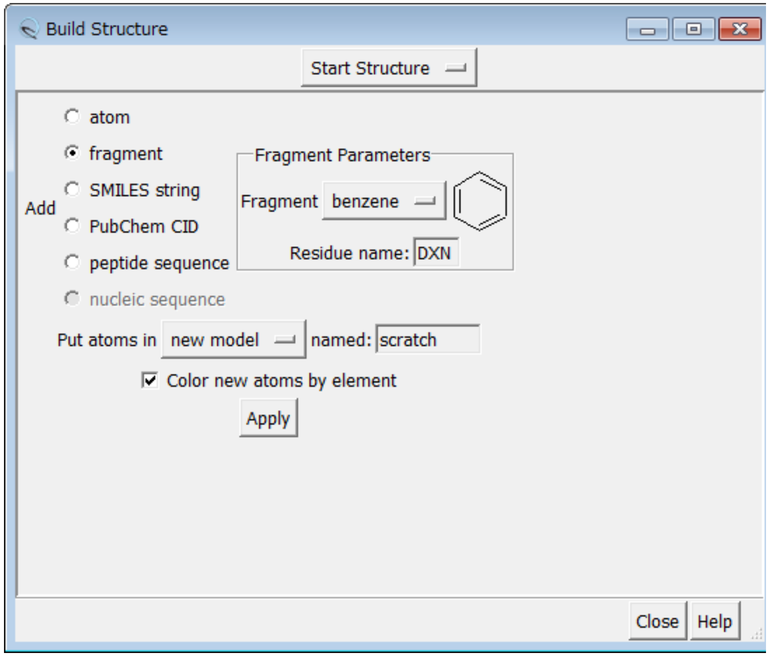
- First-order algorithm
  - Steepest descent method
    Simple, but many steps are required until convergence.
  - Conjugate gradient method
    When the energy function is a quadratic form of *n* variables, the minimum is reached in *n* steps.
- Second-order algorithm
  - Newton-Raphson法
    Quick to converge, but calculation of Hessian ($\nabla^2 E$) requires a high computational cost.

Modeling a small molecule (1)

1. Start UCSF Chimera 1.10.1



2. Choose “Tools”→“Structure Editing”→ “Build Structure.”
3. Choose “Start Structure” and Add “fragment” and select “benzene” for the fragment.
4. Enter “DXN” (dibenzo-*p*-dioxin) in the box of Residue name and click “Apply.”



Add: ● fragment

Fragment: benzene

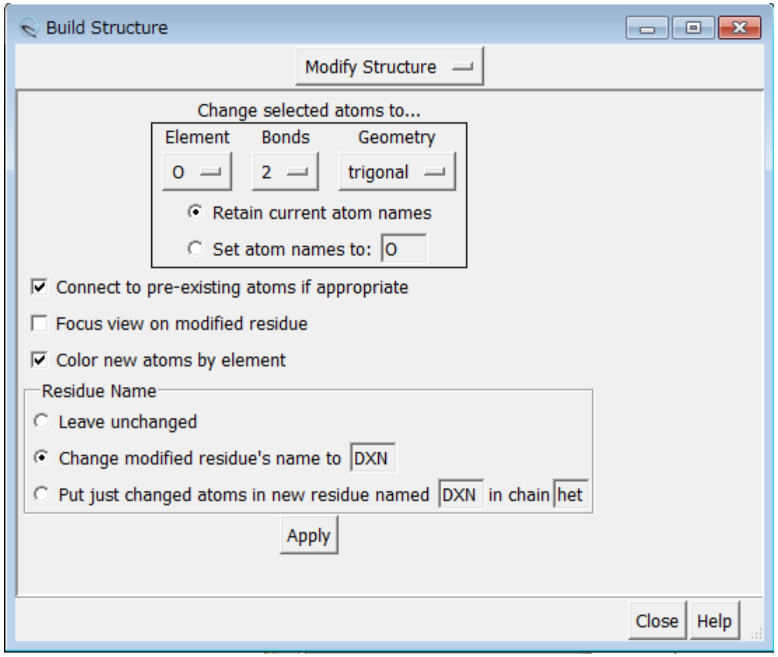
Residue name: DXN

Put atoms in: new model, named: scratch

☑ Color new atoms by element

Modeling a small molecule (2)

5. Select a hydrogen atom by clicking it while pressing the Ctrl key.
6. Change into the “Modify structure” mode, set as shown in the right, and “Apply.” →The selected atom is changed to oxygen.



Change selected atoms to...

Element	Bonds	Geometry
O	2	trigonal

● Retain current atom names

○ Set atom names to: O

☑ Connect to pre-existing atoms if appropriate

☐ Focus view on modified residue

☑ Color new atoms by element

Residue Name

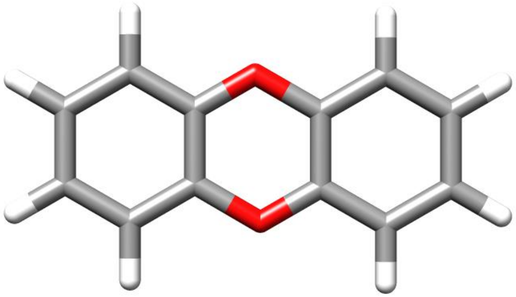
○ Leave unchanged

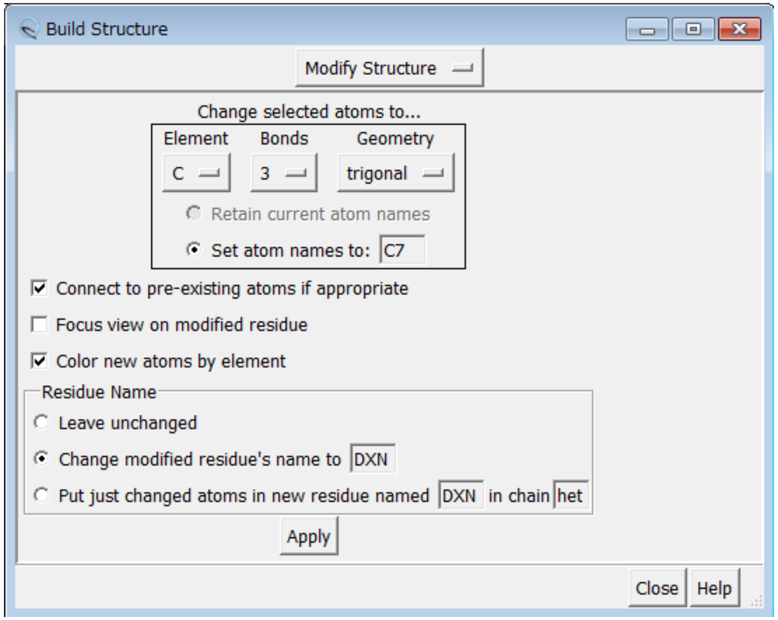
● Change modified residue's name to DXN

○ Put just changed atoms in new residue named DXN in chain het

Modeling a small molecule (3)

7. Select the hydrogen atoms bonded to the oxygen atom.
8. Set as shown in the right and "Apply."
   →Changed to carbon.
9. Using this function, generated the model of dibenzo-*p*-dioxin.





Change selected atoms to...

Element	Bonds	Geometry
C	3	trigonal

○ Retain current atom names
● Set atom names to: C7

☑ Connect to pre-existing atoms if appropriate
☐ Focus view on modified residue
☑ Color new atoms by element

Residue Name
○ Leave unchanged
● Change modified residue's name to DXN
○ Put just changed atoms in new residue named DXN in chain het

Energy minimization

1. Choose “Tools”→“Structure Editing”→“Add Charge.”
2. Select “AMBER ff99SB” and “AM1-BCC” for Standard residues and Other residues, respectively, and click “OK.”
3. Select “+0” for Net Charge and “OK.”
4. Choose “Tools”→“Structure Editing”→ “Minimize Structure” and click “minimize.”
5. Find the energy value becomes smaller in “Favorites”→“Reply Log.”

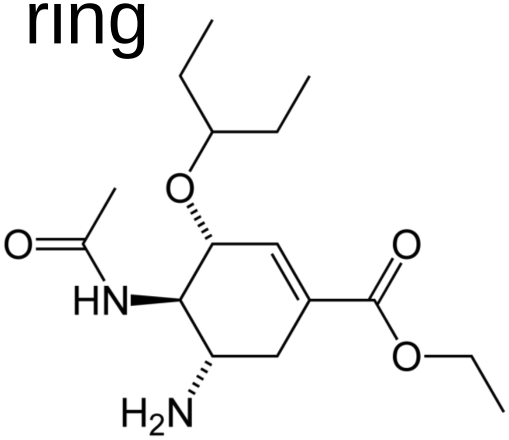
SMILES

- SMILES is a representation of the molecular structure.
- Select “Add SMILES string” in the Build Structure window in the “Start structure” mode.
- Examples of SMILES string.
  - benzene: c1ccccc1
    (Lower cases are used for aromatic atoms)
  - dibenzo-*p*-dioxin: c13ccccc1Oc2ccccc2O3
  - alanine: [N+][C@@H](C)C(=O)[O-]

http://www.daylight.com/dayhtml/doc/theory/theory.smiles.html

Exercise 2

- Generate this molecule and do energy minimization.
  - Use SMILES “C1CCCC=C1” to generate the ring



Oseltamivir: an antiviral drug
(trade name: Tamiflu)

- Save the image of the energy-minimized structure in the PNG format.

Assignment submission instructions

- Attach the Excel file of Exercise 1 and the image file of Exercise 2 to an email.
- Put the energy values before and after the energy minimization in the body of the email.
- Put “Molecular Modeling” in the Subject field.
- Be sure to put your name and ID card number (if you are a student) in the body of the email.
- Send the email to Prof. Tohru Terada (tterada@iu.a.u-tokyo.ac.jp).