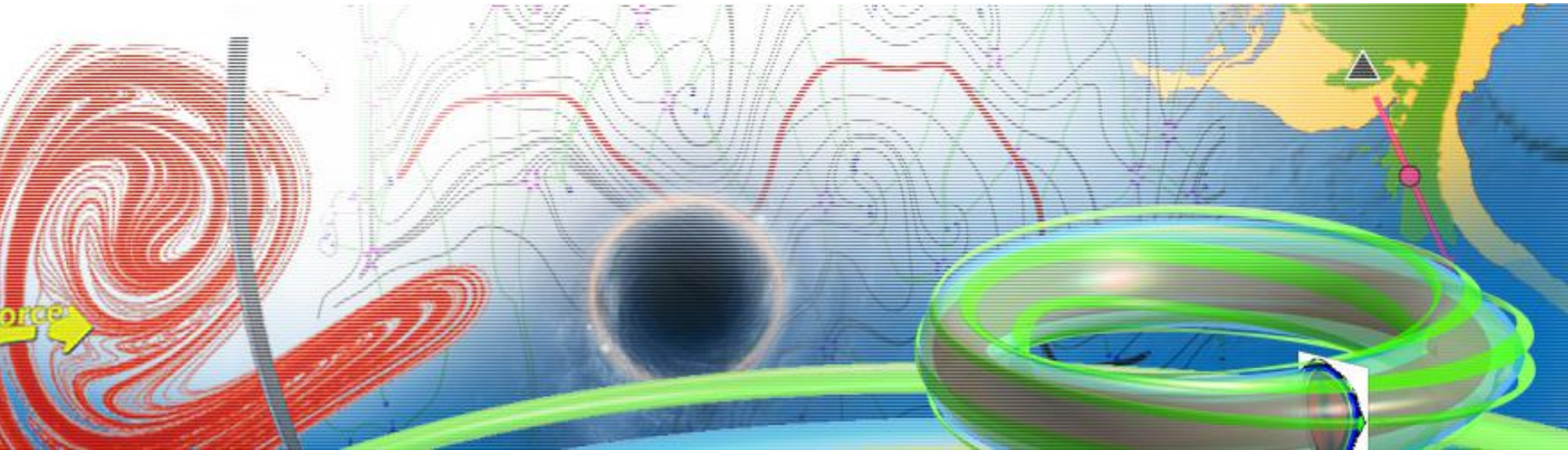
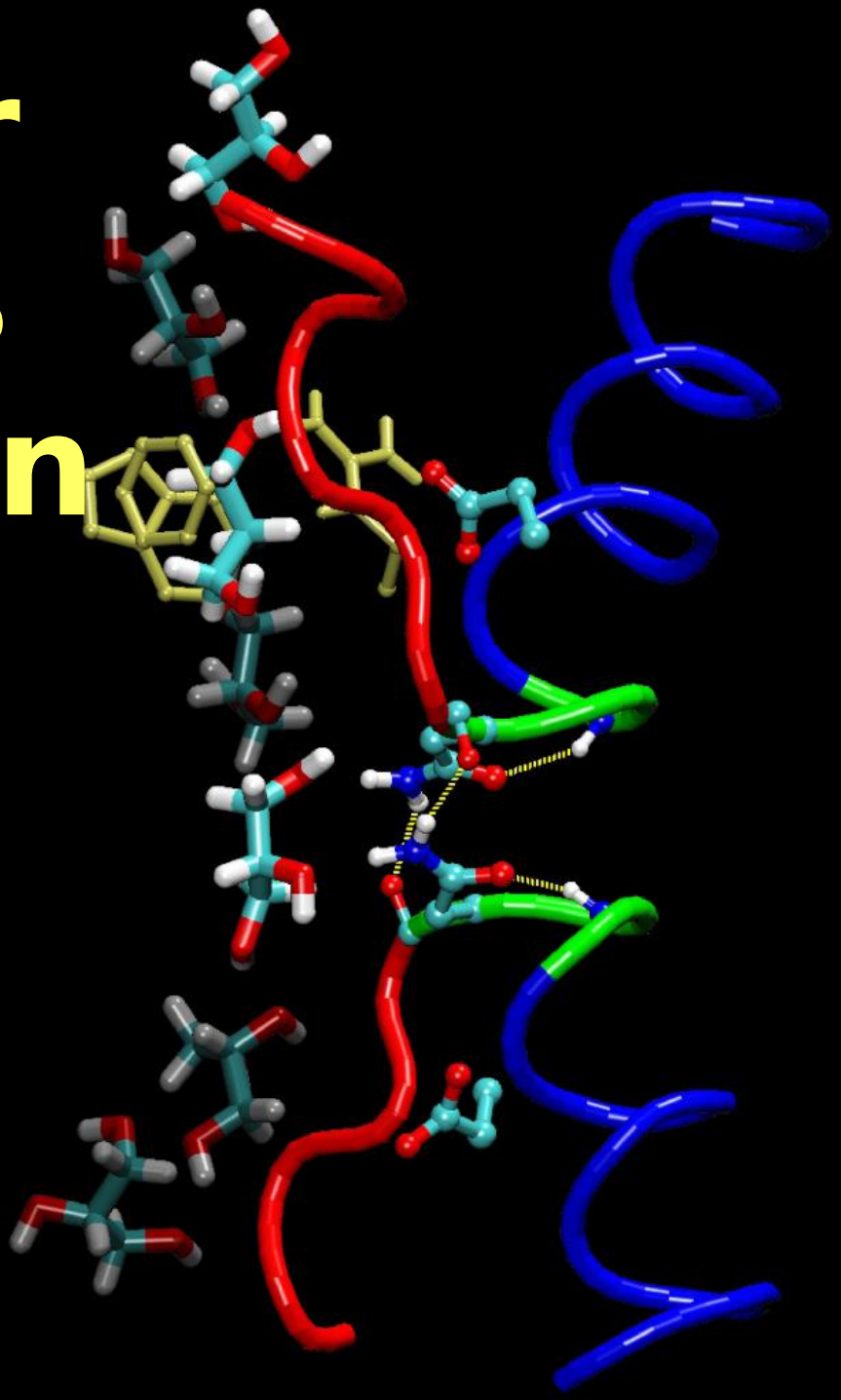
生物动力系统模拟

王冠宇 18665955633
wanggy@sustc.edu.cn



Molecular Dynamics Simulation

(a brief introduction)



Steps

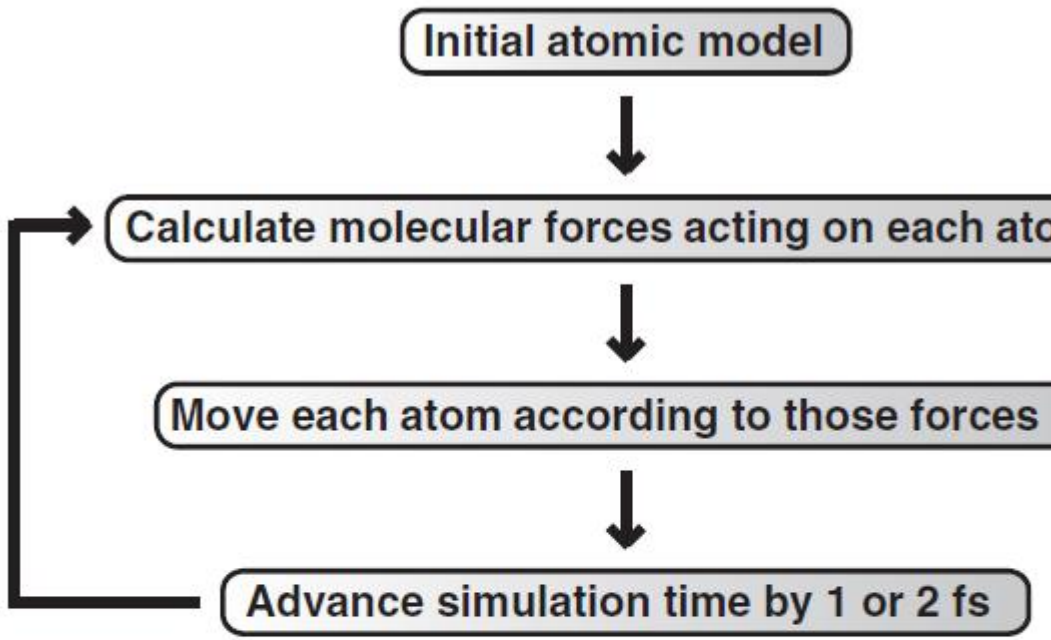


Figure 2. A schematic showing how a molecular dynamics simulation is performed. First, a computer model of the receptor-ligand system is prepared. An equation like that shown in Figure 3 is used to estimate the forces acting on each of the system atoms. The positions of the atoms are moved according to Newton's laws of motion. The simulation time is advanced, and the process is repeated many times. This figure was adapted from a version originally created by Kai Nordlund.

Give atoms initial positions $\boldsymbol{r}^{(t=0)}$, choose short Δt

↓

Get forces $\boldsymbol{F} = -\nabla V(\boldsymbol{r}^{(i)})$ and $\boldsymbol{a} = \boldsymbol{F}/m$

↓

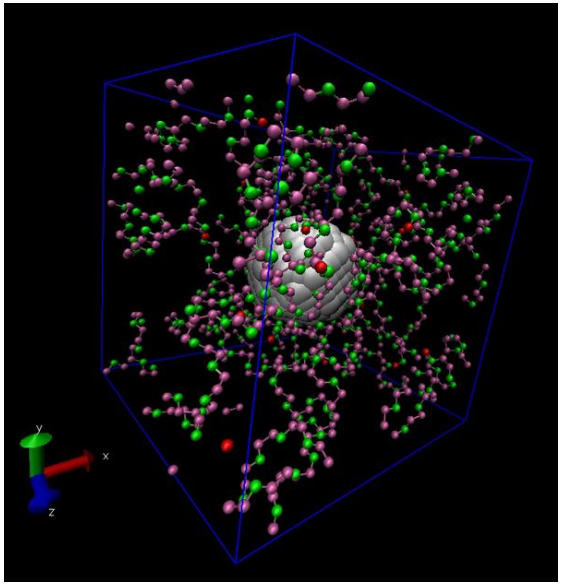
Move atoms: $\boldsymbol{r}^{(i+1)} = \boldsymbol{r}^{(i)} + \boldsymbol{v}^{(i)} \Delta t + \frac{1}{2} \boldsymbol{a} \Delta t^2 + \ldots$

↓

Move time forward: $t = t + \Delta t$

↓

Repeat as long as you need



Basic Computation

The classical MD simulations boil down to numerically integrating Newton's equations of motion for the particles (atoms, in the simplest case) which build up the investigated system:

$$m\frac{d^2\mathbf{r}_i}{dt^2} = \mathbf{F}_i(\mathbf{r}_1, \mathbf{r}_2, \ldots \mathbf{r}_N), \quad i = 1, 2, \ldots, N.$$

Here $\mathbf{r}_i$ are the position vectors and $\mathbf{F}_i$ are the forces acting upon the N particles in the system.

Quite often forces derive from *potential functions*, $U(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$, representing the potential energy of the system for the specific geometric arrangement of the particles:

$$\mathbf{F}_i(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N) = -\nabla_{\mathbf{r}_i} U(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N).$$

This form implies the conservation of the total energy $E = E_{\text{kin}} + U$, where E_{kin} is the instantaneous kinetic energy.

In some context, we use q instead of r.

The fundamental reason is that the system is a Hamiltonian system

In Hamiltonian mechanics, a classical physical system is described by a set of canonical coordinates $\boldsymbol{r} = (\boldsymbol{q}, \boldsymbol{p})$, where each component of the coordinate $\boldsymbol{q}_i, \boldsymbol{p}_i$ is indexed to the frame of reference of the system.

The time evolution of the system is uniquely defined by Hamilton's equations:

$$\frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = -\frac{\partial \mathcal{H}}{\partial \boldsymbol{q}}$$
$$\frac{\mathrm{d}\boldsymbol{q}}{\mathrm{d}t} = +\frac{\partial \mathcal{H}}{\partial \boldsymbol{p}}$$

where $\mathcal{H} = \mathcal{H}(\boldsymbol{q}, \boldsymbol{p}, t)$ is the Hamiltonian, which often corresponds to the total energy of the system. For a closed system, it is the sum of the kinetic and potential energy in the system.

$$\frac{d\boldsymbol{p}}{dt} = -\frac{\partial V}{\partial \boldsymbol{q}}$$

$$\mathcal{H} = T + V, \quad T = \frac{p^2}{2m}, \quad V = V(q).$$

Note that T is a function of p alone, while V is a function of q alone

$$\frac{d\boldsymbol{p}}{dt} = -\frac{\partial V}{\partial \boldsymbol{q}}$$

$$\boldsymbol{p} = m\boldsymbol{v} = m\frac{d\boldsymbol{q}}{dt}$$

$$m\frac{d^2\boldsymbol{q}}{dt^2} = -\frac{\partial V}{\partial \boldsymbol{q}}$$

$$\boldsymbol{F} = -\frac{\partial V}{\partial \boldsymbol{q}}$$

$$\boldsymbol{F_i} = -\frac{\partial V}{\partial \boldsymbol{q_i}}$$

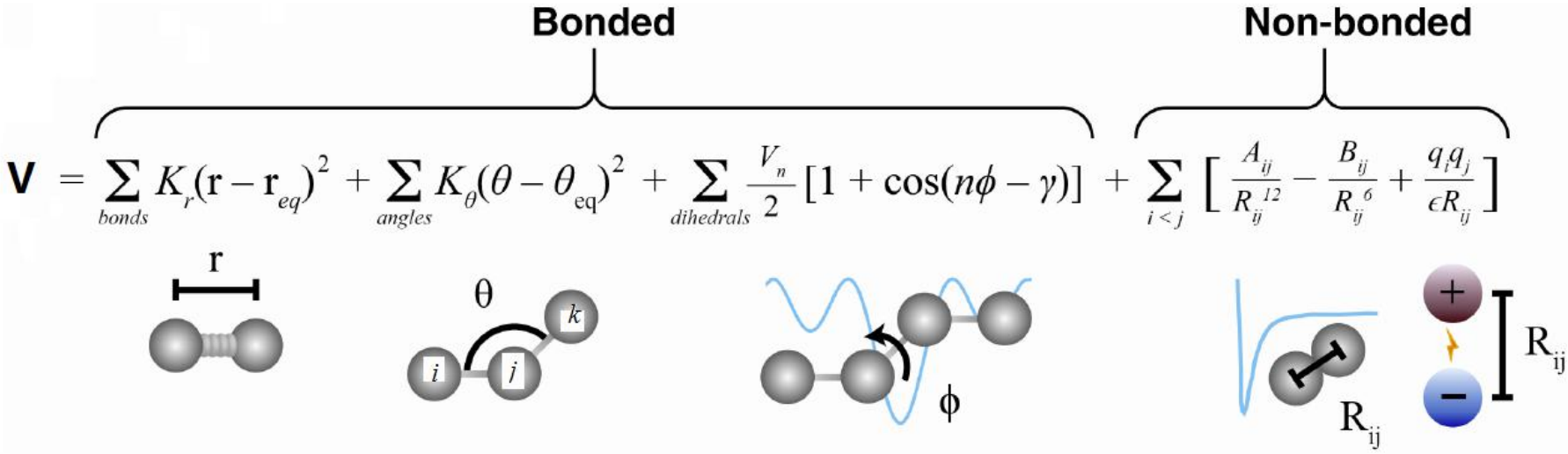
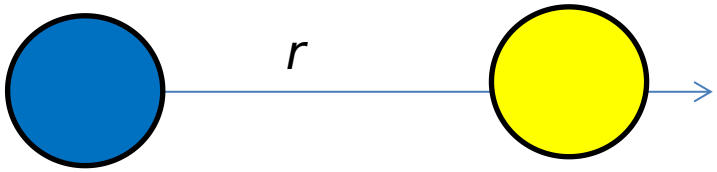
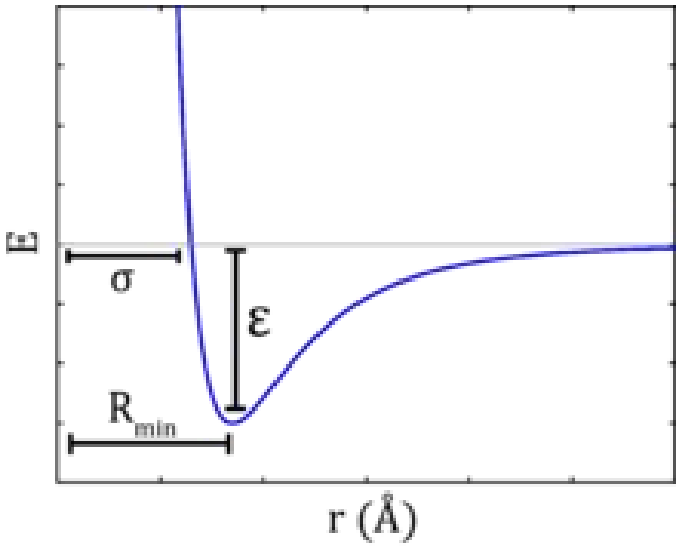


Figure 3. An example of an equation used to approximate the atomic forces that govern molecular movement. The atomic forces that govern molecular movement can be divided into those caused by interactions between atoms that are chemically bonded to one another and those caused by interactions between atoms that are not bonded. Chemical bonds and atomic angles are modeled using simple springs, and dihedral angles (that is, rotations about a bond) are modeled using a sinusoidal function that approximates the energy differences between eclipsed and staggered conformations. Non-bonded forces arise due to van der Waals interactions, modeled using the Lennard-Jones potential, and charged (electrostatic) interactions, modeled using Coulomb's law.

$$\theta = cos^{-1} \frac{(\boldsymbol{r}_k - \boldsymbol{r}_j) \cdot (\boldsymbol{r}_i - \boldsymbol{r}_j)}{|\boldsymbol{r}_k - \boldsymbol{r}_j||\boldsymbol{r}_i - \boldsymbol{r}_j|}$$

Example





Given that the potential between the two neutral atoms are:

$$V = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right]$$

Quiz

What is V when r = infinity?

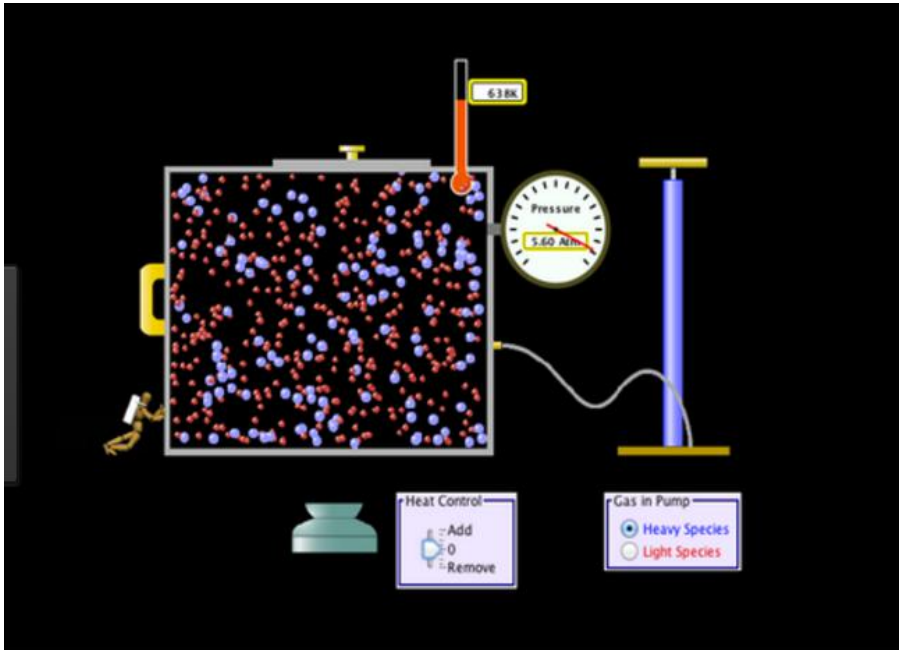
What is V when r = σ ?

What is V when r = 0 ?

What is the force acting on the yellow

Answer 0 0 $\infty F = -\frac{\partial V}{\partial r} = \frac{48\varepsilon}{r}\left[\left(\frac{\sigma}{r}\right)^{12} - \frac{1}{2}\left(\frac{\sigma}{r}\right)^{6}\right]$

Thermodynamic properties



In usual scientific researches,
thermodynamic properties
(such as temperature and pres
are given as conditions.

In MD simulations,
It is the set of atoms that is given as the starting condition.
Thermodynamic properties usually need to be calculated from the simulation.

Thermodynamic properties

Temperature

It measures kinetic energies of individual atoms.

It is actually averaged kinetic energy.

1. First we calculate the total kinetic energy

$$E_{\text{kin}} = \frac{1}{2}\sum_{i=1}^{N} m_i {v_i}^2$$

2. Then we divide N to get the average kinetic energy

E_{kin}/N

3. It is almost the temperature T

$$\frac{3}{2}k_B T = \frac{E_{kin}}{N} \qquad T = \frac{2E_{kin}}{3k_B N}$$

Thermodynamic properties

Pressure

$$PV = Nk_{\mathrm{B}}T + \frac{1}{3}\left\langle \sum_{i=1}^{N} \mathbf{r}_i \cdot \mathbf{F}_i \right\rangle$$

Where <> represents “the expectation”

For example, <T> means the average temperature.

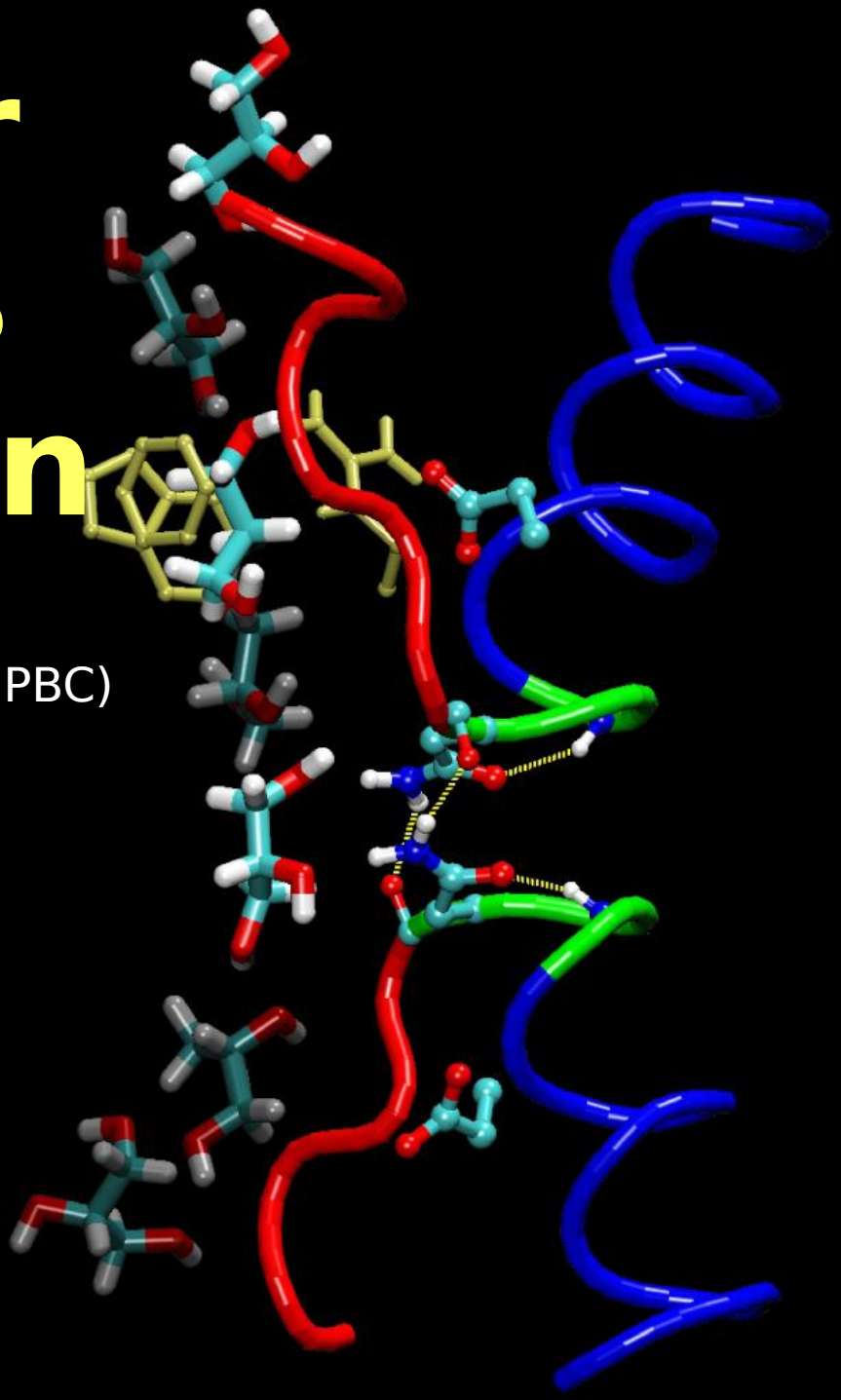
$$\left\langle \sum_{i=1}^{N} \mathbf{r}_i \cdot \mathbf{F}_i \right\rangle \neq \frac{\sum_{i=1}^{N} \mathbf{r}_i \cdot \mathbf{F}_i}{\mathrm{N}}$$

$$\left\langle \sum_{i=1}^{N} \mathbf{r}_i \cdot \mathbf{F}_i \right\rangle = \frac{\sum_{i=1}^{N} \mathbf{r}_i \cdot \mathbf{F}_i + \ldots}{\mathrm{M}}$$

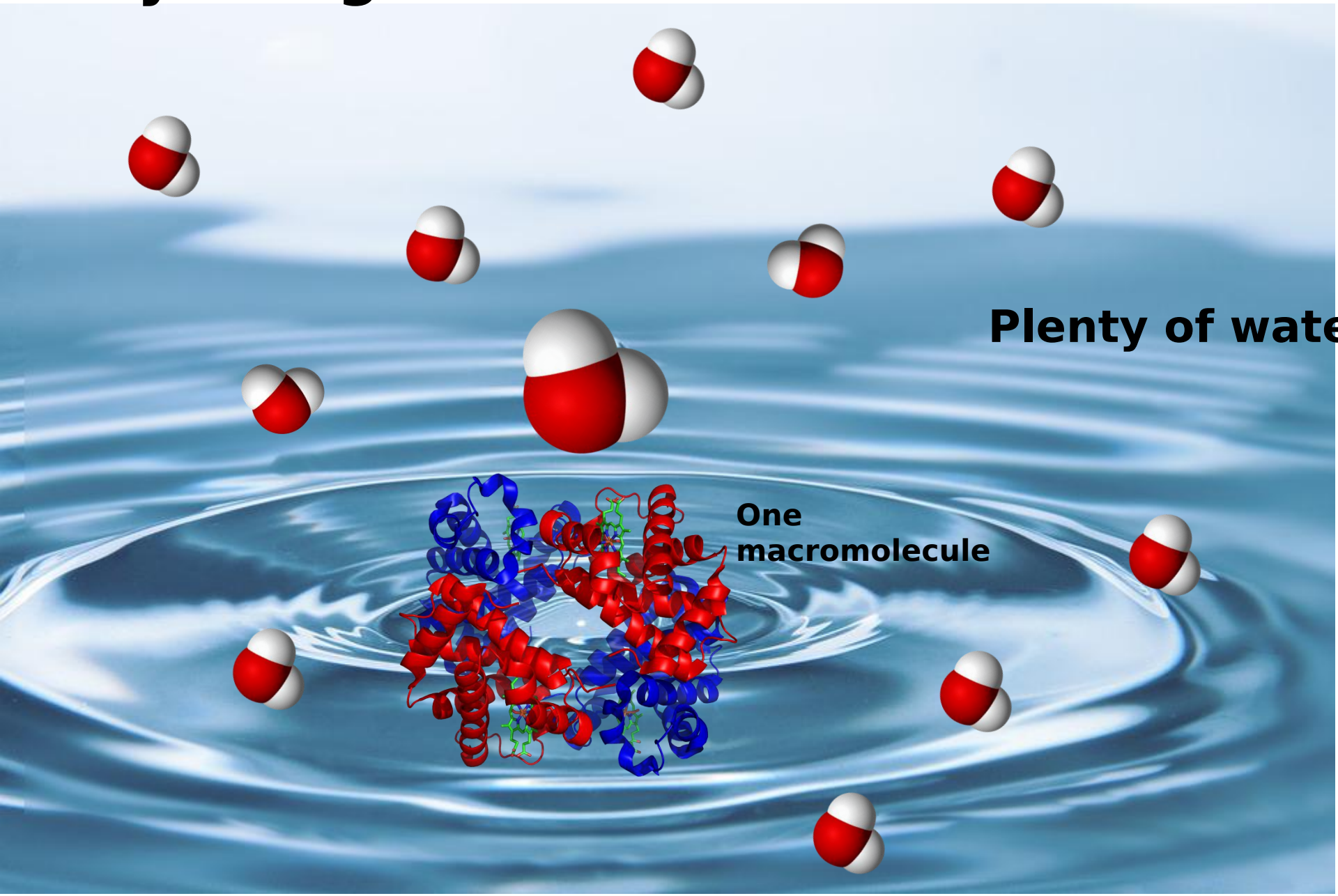
For ideal ga $\left\langle \sum_{i=1}^{N} \mathbf{r}_i \cdot \mathbf{F}_i \right\rangle = 0$, $PV = Nk_{\mathrm{B}}T$

Molecular Dynamics Simulation

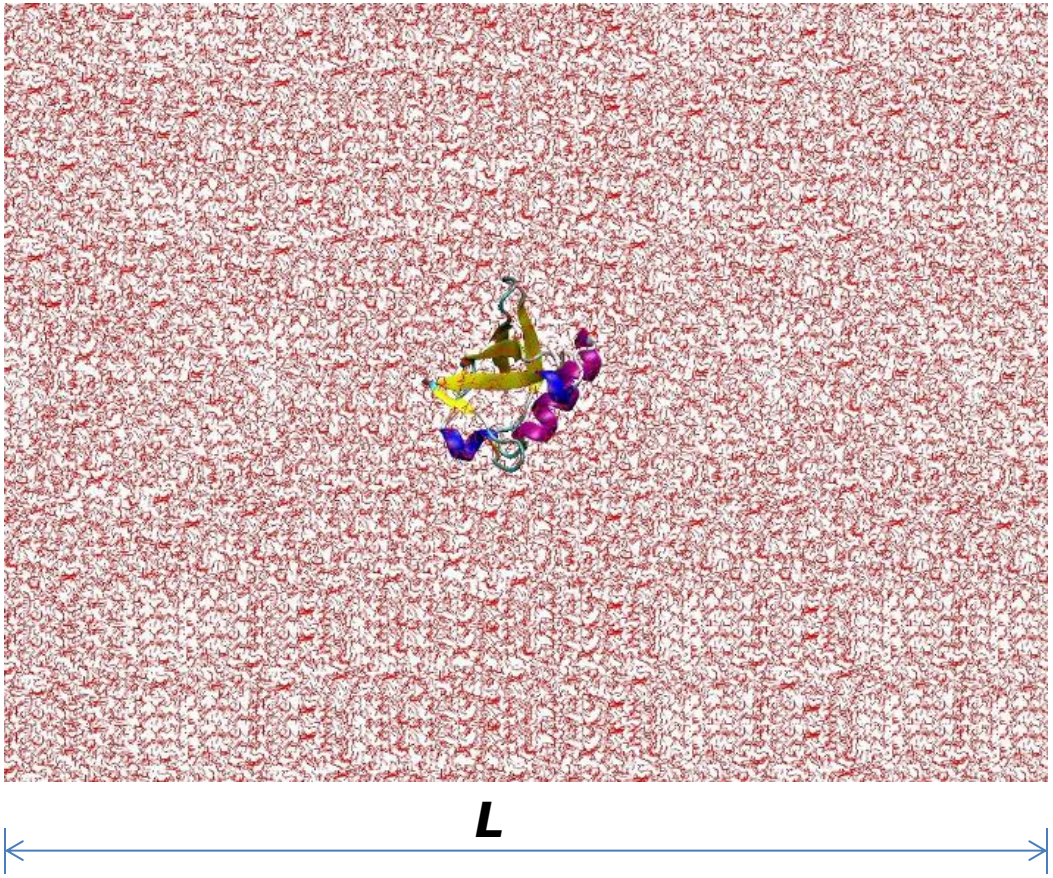
Periodic Boundary Condition (PBC)



Why using PBC?



Because there is always a boundary.
And dynamics at the boundary is dramatically different from the inne

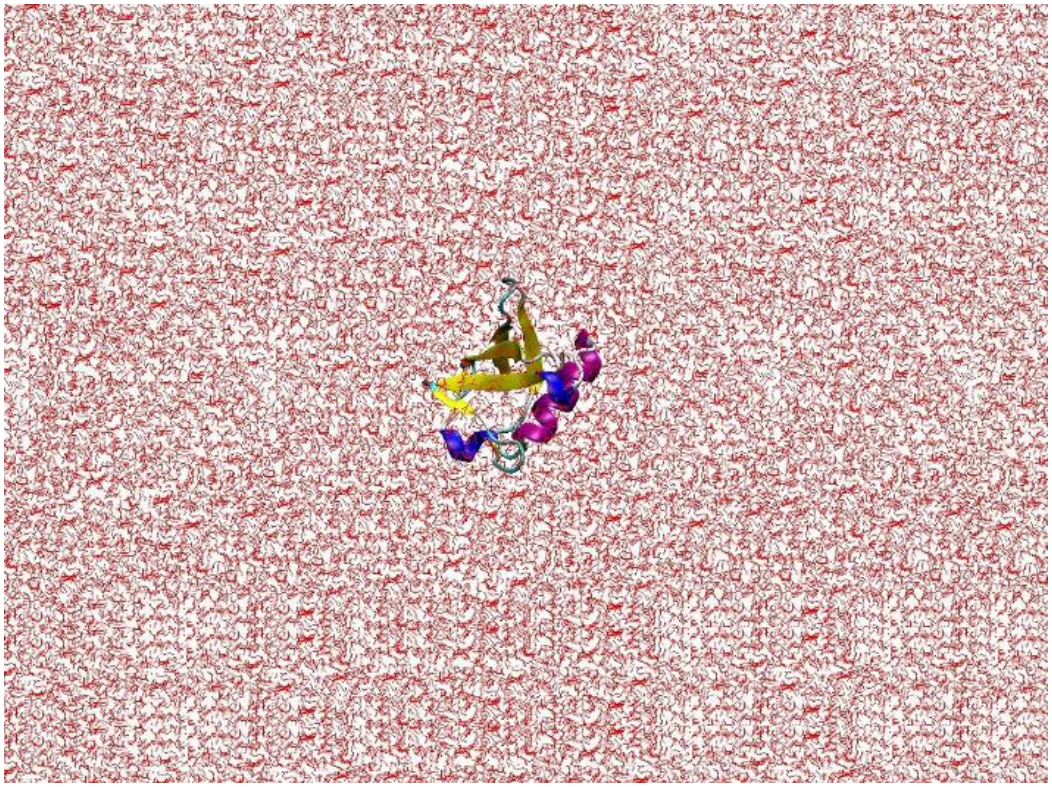


If *L* is large, then boundary effect is **relatively** small and can thus be ignored

$$\frac{L^2}{L^3} = \frac{1}{L}$$

Macroscopically observable water should be sufficient.

For example, 18 grams of water + 1 protein



$$\boldsymbol{p} = m\boldsymbol{v} = m\frac{d\boldsymbol{q}}{dt}$$

About how many water molecules in the boundary? $(10^{24})^{\frac{2}{3}} = 10^{16}$

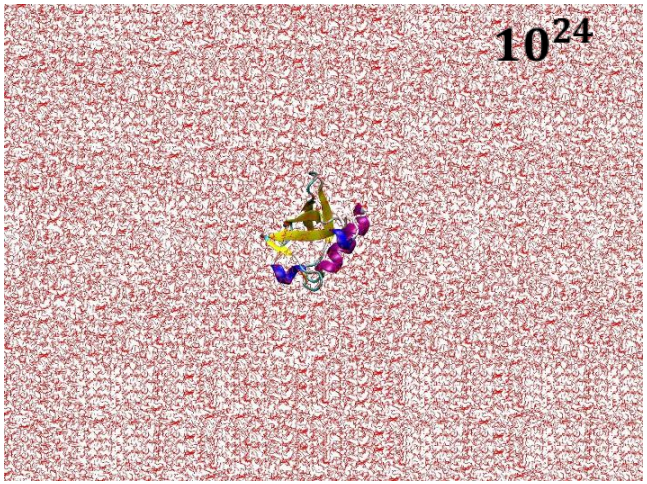
What is the percentage? $\frac{10^{16}}{10^{24}} = \frac{1}{10^{8}} = 0.00000001$

Conclusion: The boundary effect can be ignored

But, can we simulate a system with 10^{24} water molecules?

No Way!

The typical number of particles which can be handled in MD simulation nowadays is of the order 10^6



10^6

But, is the boundary effect ignorable for a system with only 10^6 wate

Water molecules in the boundary: $\vec{F} = -\frac{\partial V}{\partial r}\vec{r} = \frac{48\varepsilon}{r}\left[\left(\frac{\sigma}{r}\right)^{12} - \frac{1}{2}\left(\frac{\sigma}{r}\right)^{6}\right]$

Water molecules in total: 10^6

Percentage: $\frac{10^4}{10^6} = \frac{1}{10^2} = 0.01$

Conclusion: The boundary effect cannot be ignored

Solution:
Periodic Boundary Condition

1. **It is an infinite system (no boundary effect)**

2. **Although infinite, Computation is finite**

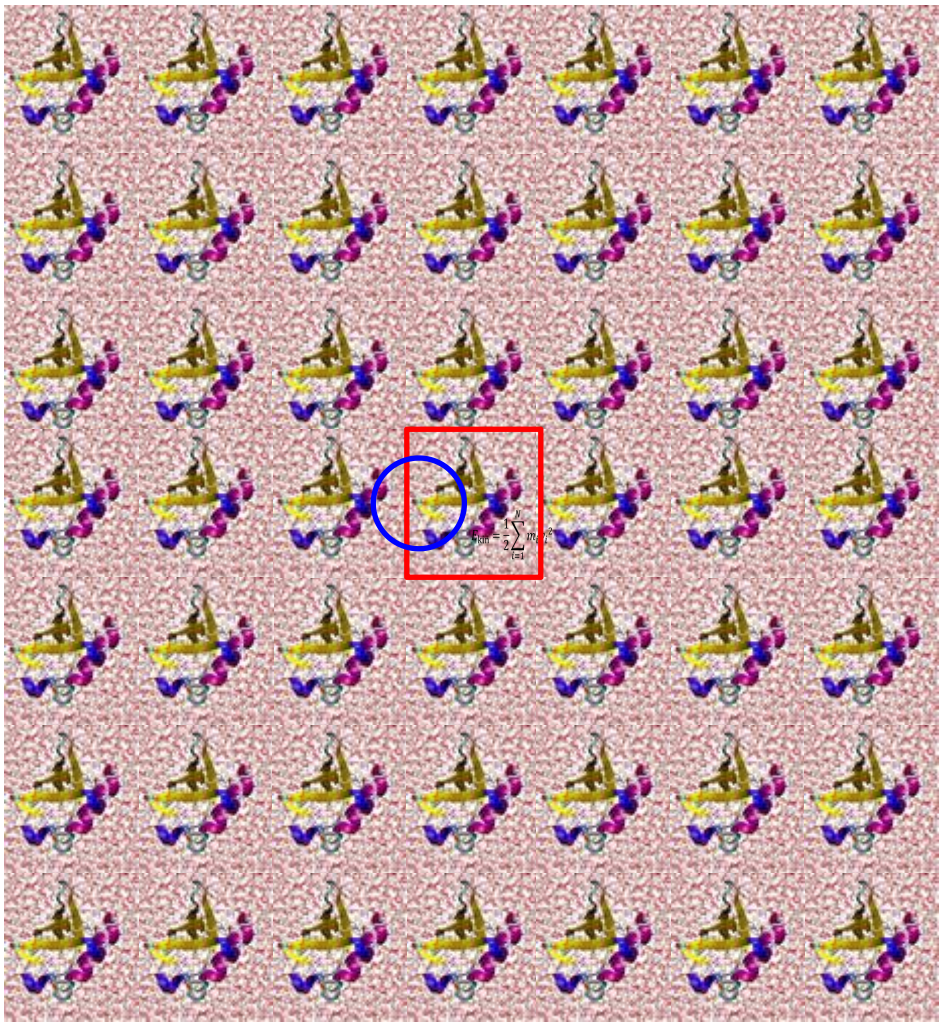
   Just consider atoms in the original box.

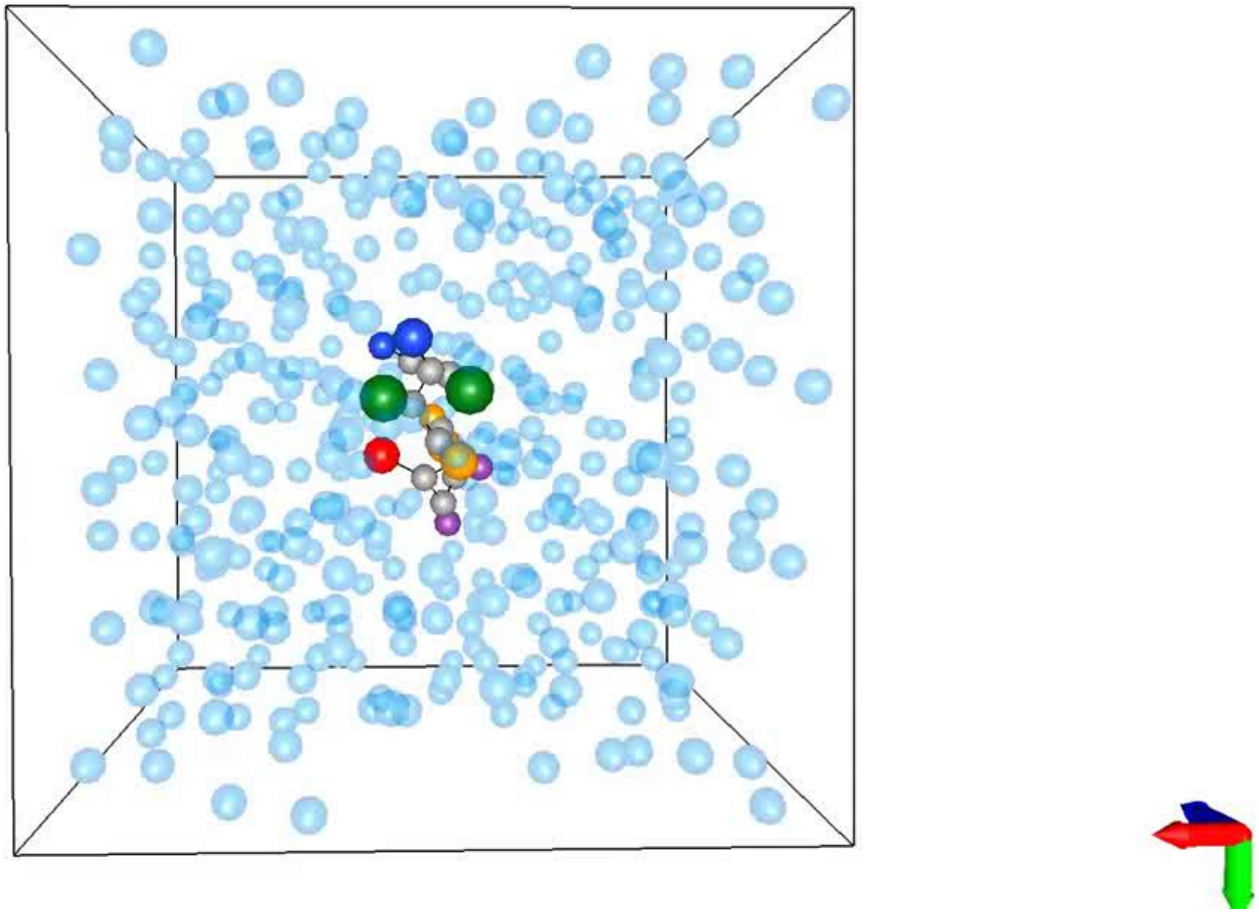
   Other atoms just copy

3. **What is the role played by other atoms?**

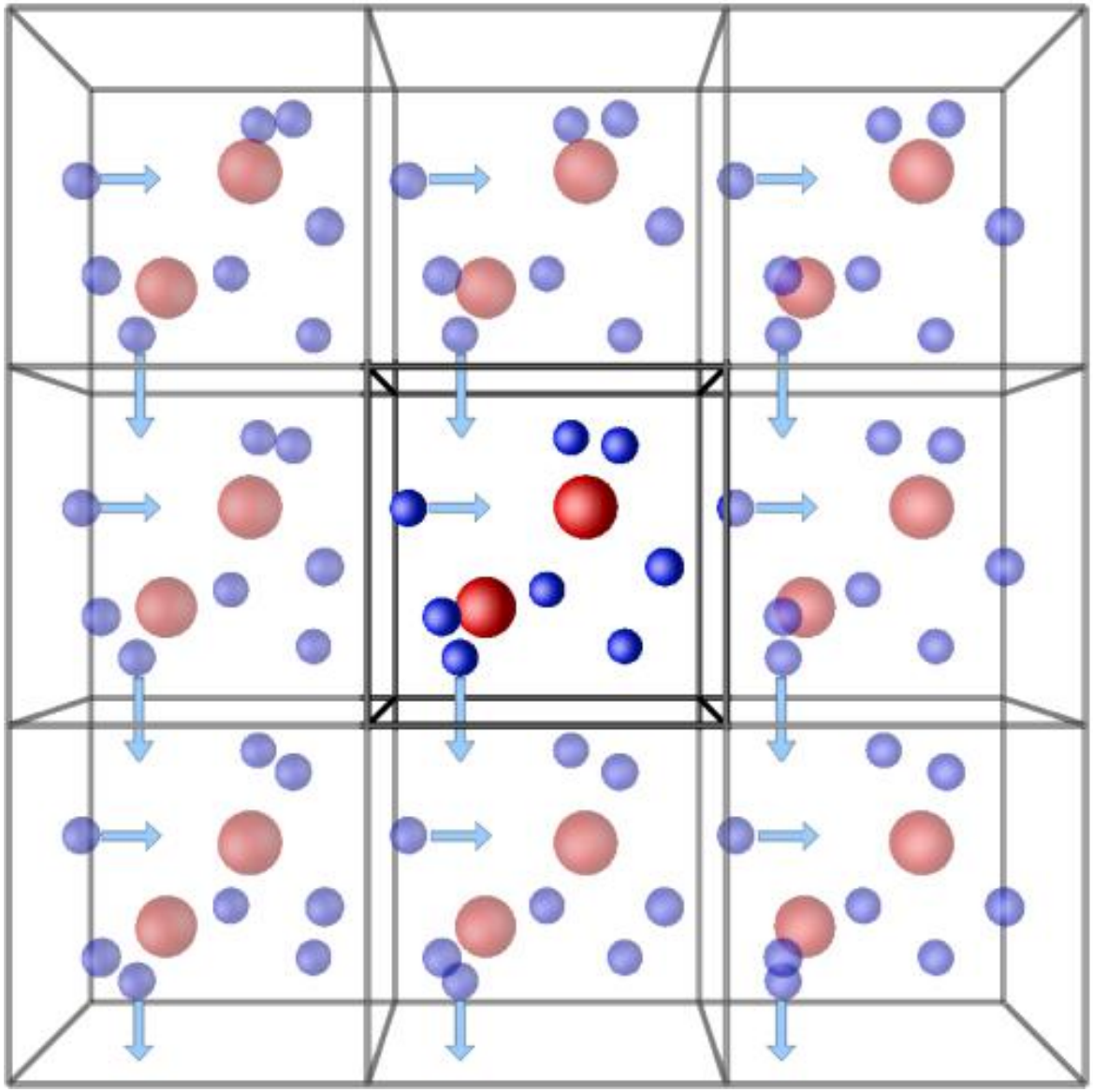
   They replace atoms in the box (此入彼出)

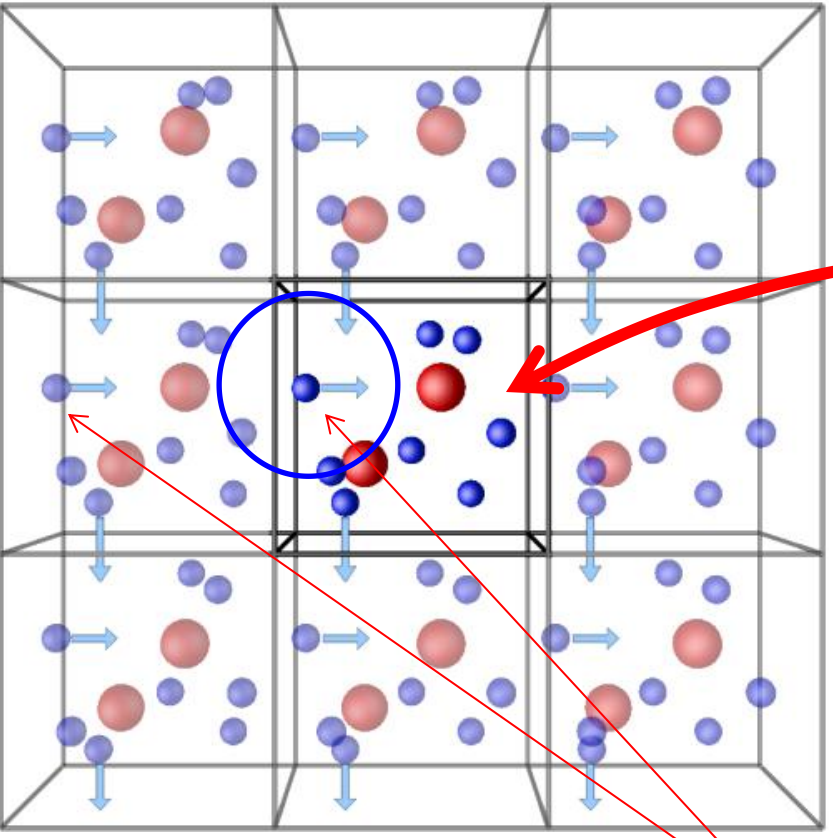
   They may exert force to atoms in the box





Solution: Periodic Boundary Condition





1. Particles in the box are mentally replicate
   infinity by rigid translation in all the thre
   Cartesian directions, completely filling t

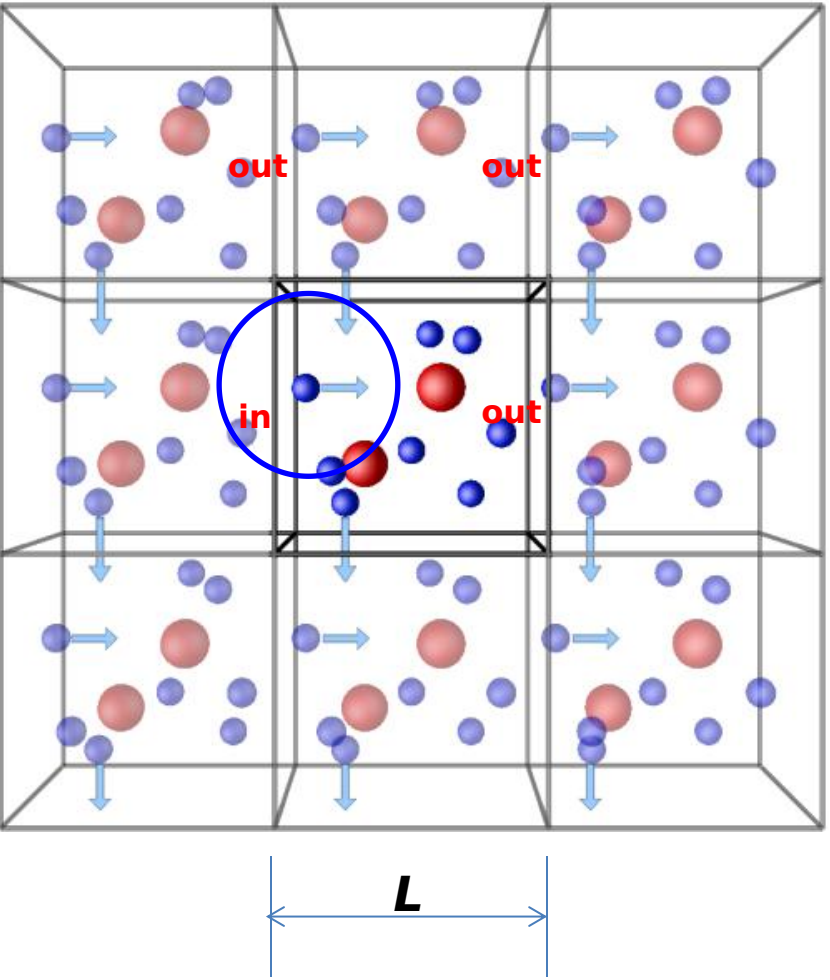
2. All the “image” particles move solidary wi
   “original” particle from the simulated box.
   When a particle enters or leaves the simu
   region, an image particle leaves or enters
   region, such that the number of particles
   simulation region is always conserved.

3. Only particles in the box are simulated.
   Other particles just copy.

Question: which particle needs I to calculate force on it?

4. How to calculate the force? There are so many particles acting on it!!!

5. Use a cut-off circle with radius R.



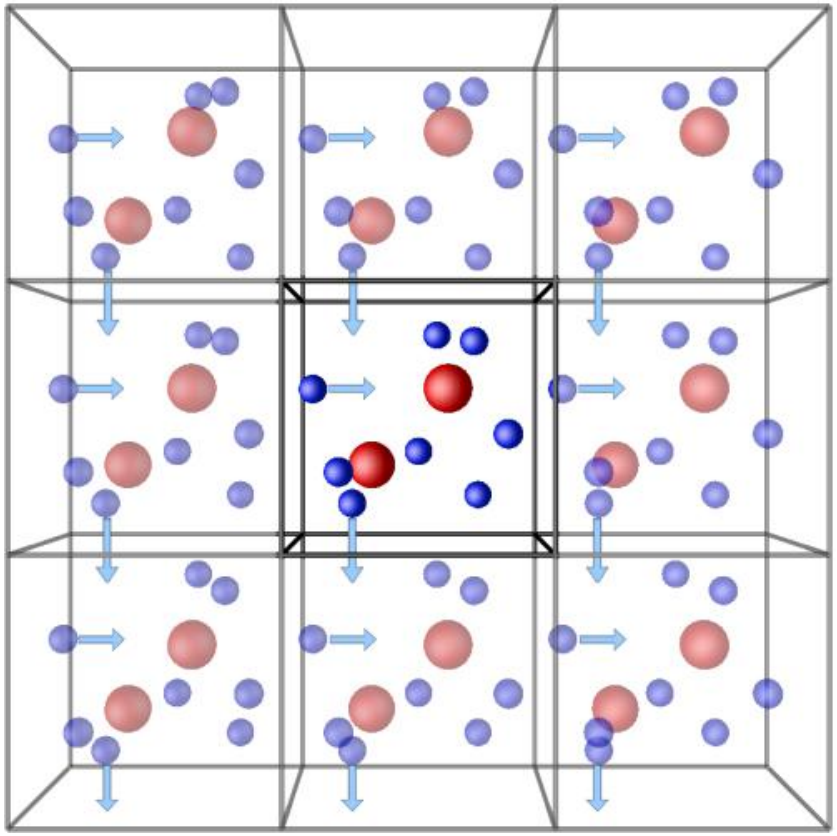
A trick:
By carefully choosing the radius R,
only one image particle locates within the

The particle within the circle
is not necessarily the one in the simulation

consider only the closest and neglect

Criterion for choosing R

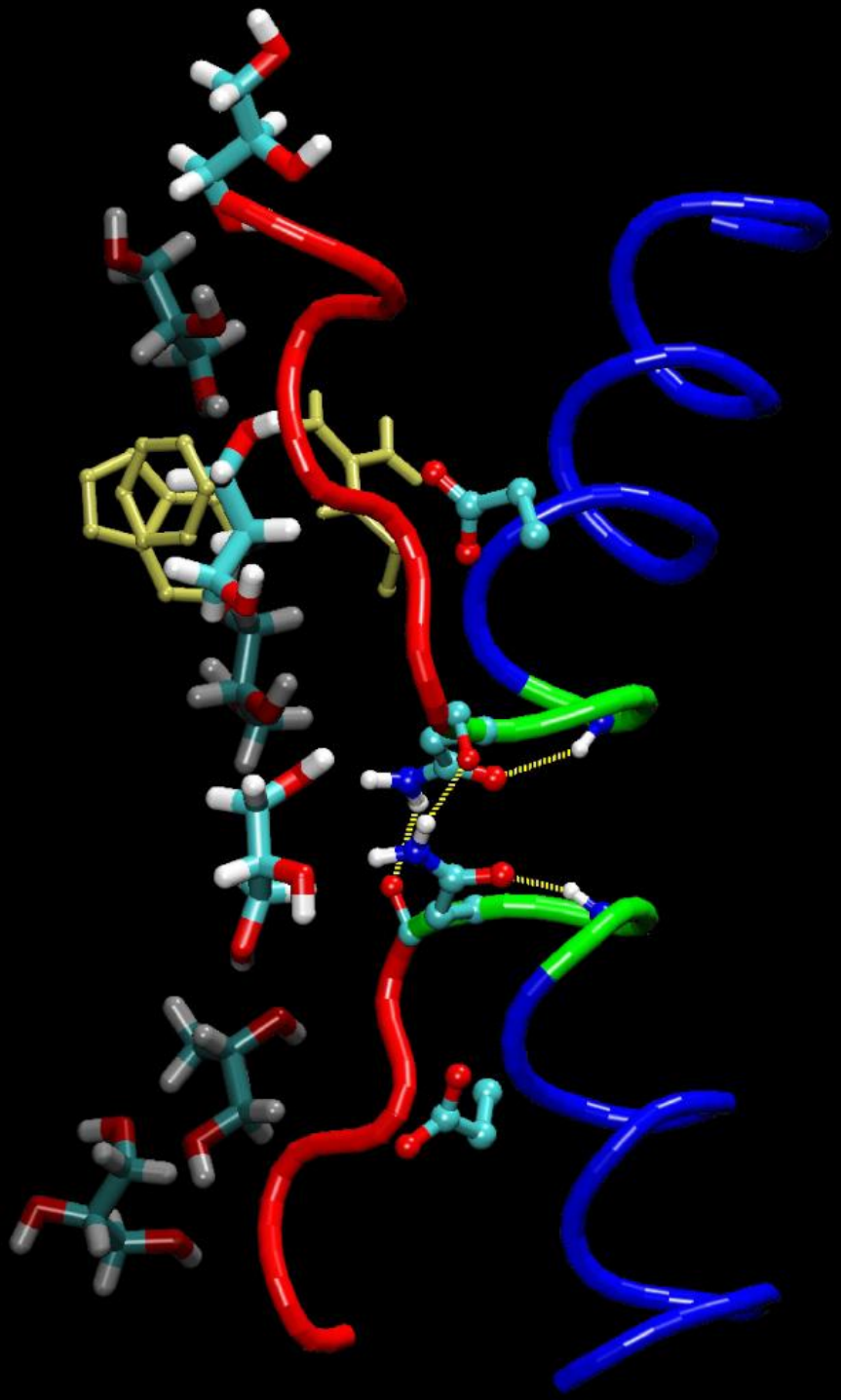
$$2R < L$$



6. The surface effects are thus virtually eliminated
   (because the system is essentially infinite)
7. What really amazing is that:
   Although the system is infinite,
   one needs only to calculate N particles
   Others just copy.

Molecular Dynamics Simulation

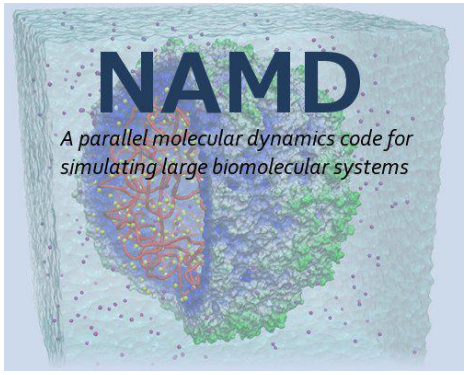
Software

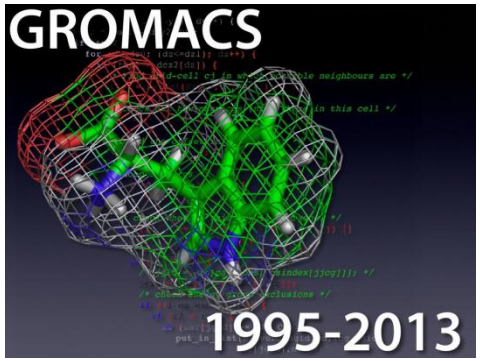


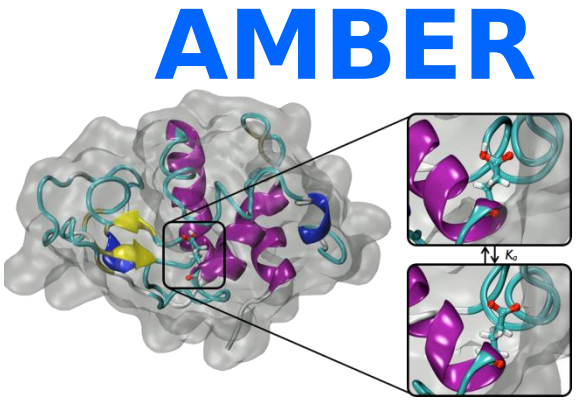
Illustration

. . .

Simulation

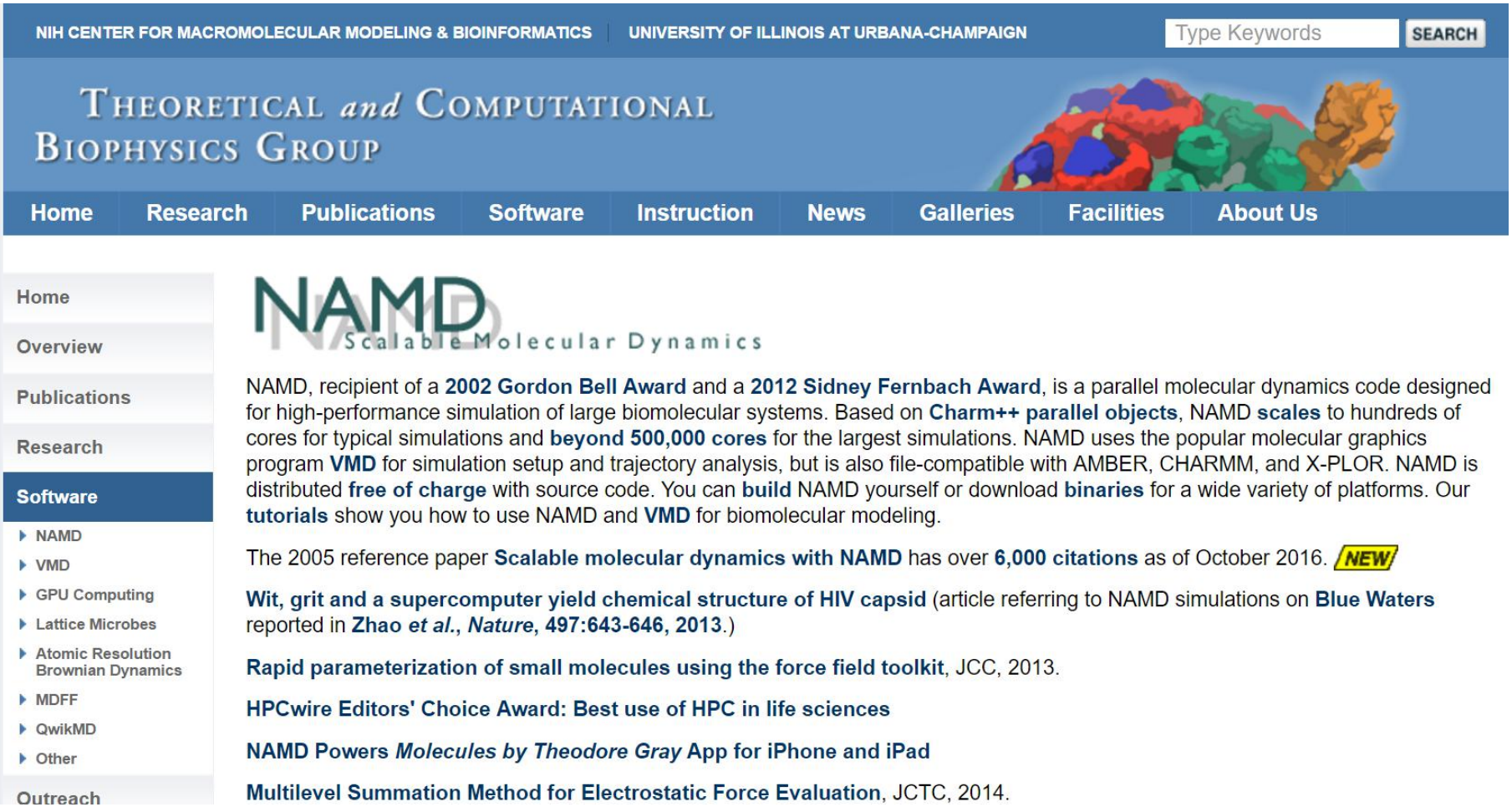






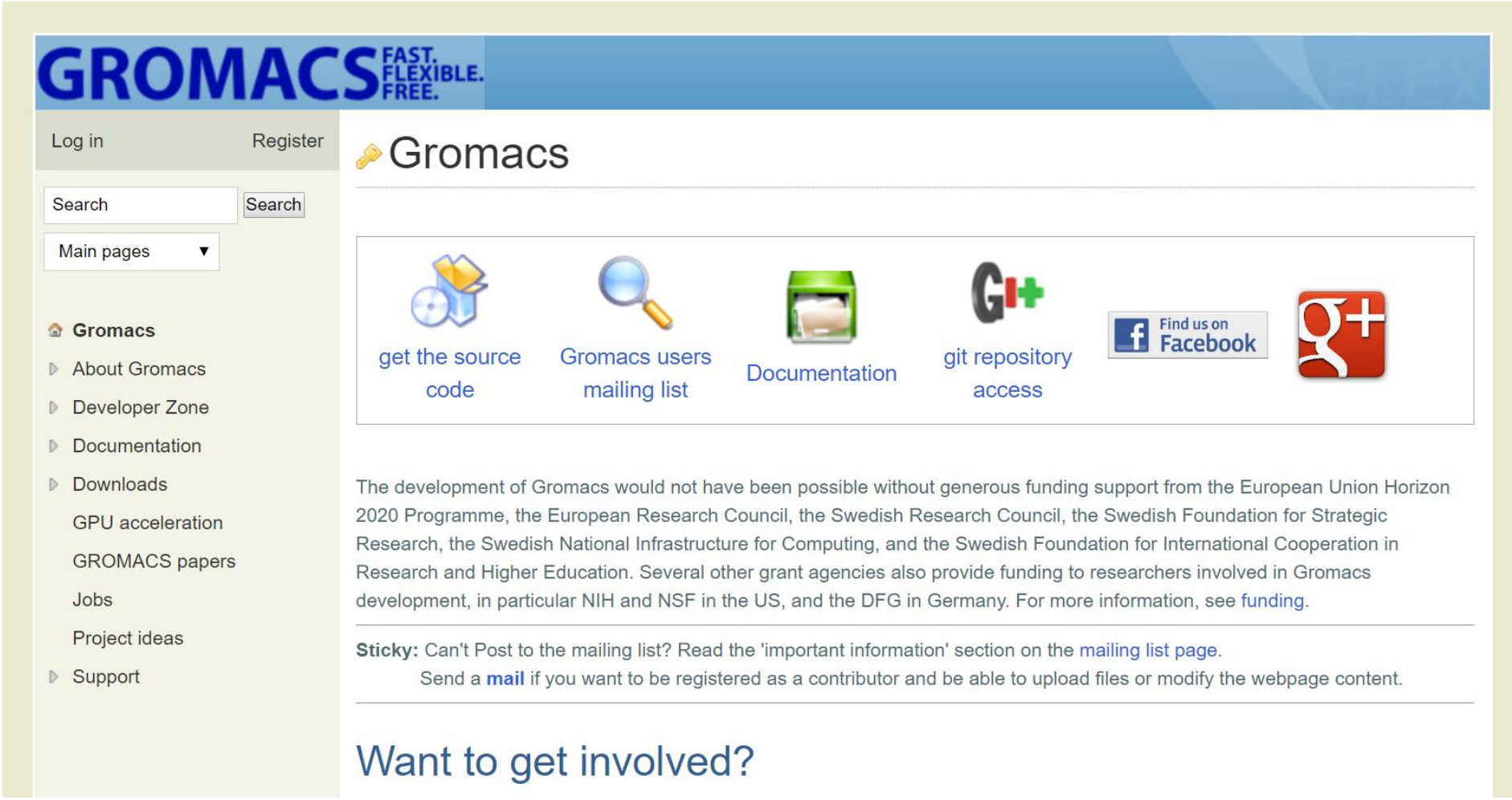
MD Simulation Software

http://www.ks.uiuc.edu/Research/namd/



MD Simulation Software

http://www.gromacs.org/

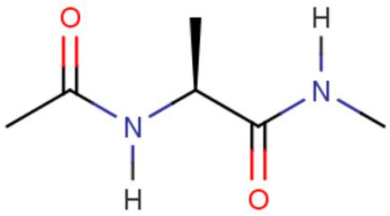


MD Simulation Software:

AMBER: Assisted Model Building and Energy Refinement

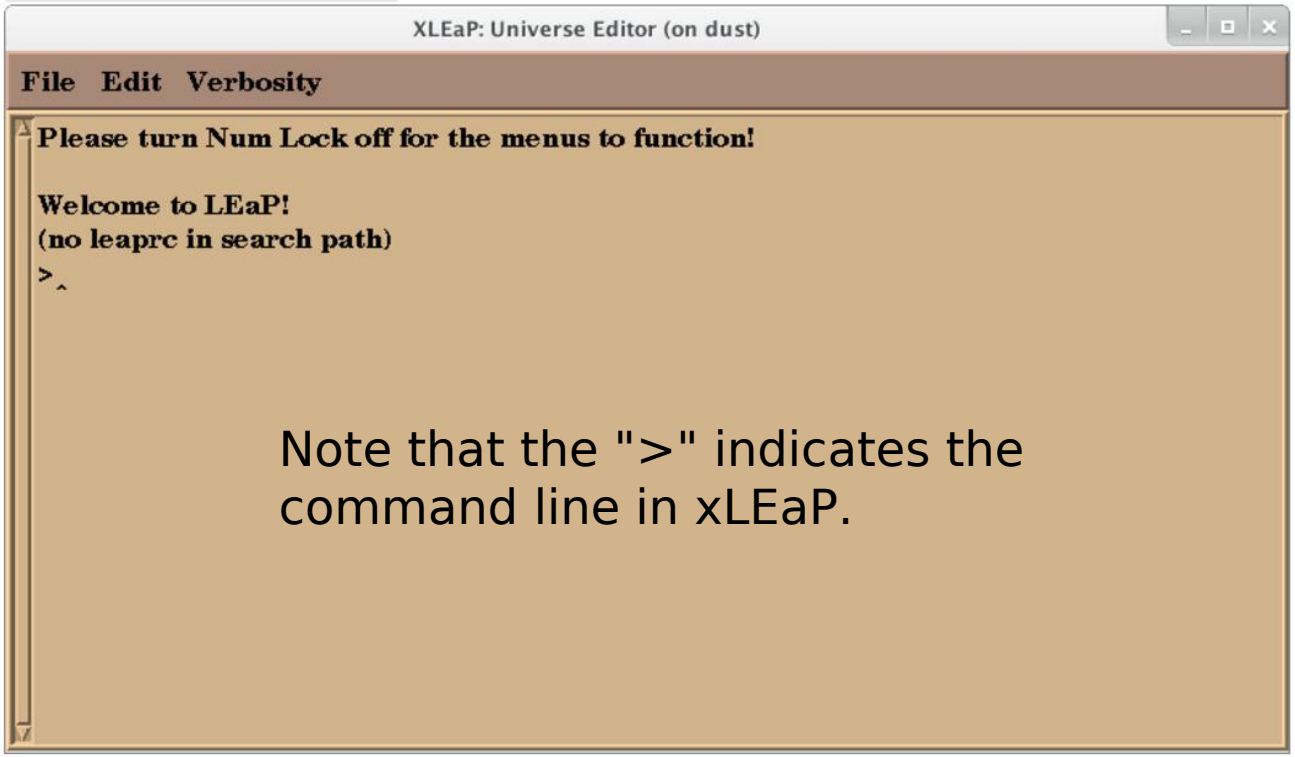
http://ambermd.org/

Example 1: build the following molecule in the preparatory program called xLEaP for simulation in AMBER



1. Start xLEaP now with the **xleap** command

```
$ xleap
```



2. **source** (load) a force field

a set of parameter and equations for use in molecular mechanics

$$V_{\mathrm{AMBER}} = \sum_{i}^{n_{\mathrm{bonds}}} b_i(r_i - r_{i,\mathrm{eq}})^2 + \sum_{i}^{n_{\mathrm{angles}}} a_i(\theta_i - \theta_{i,\mathrm{eq}})^2 + \sum_{i}^{n_{\mathrm{dihedrals}}} \sum_{n}^{n_{i,\max}} (V_{i,n}/2)[1 + \cos(n\phi_i - \gamma_{i,n})]$$
$$+ \sum_{i<j}^{n_{\mathrm{atoms}}}{}' \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}}\right) + \sum_{i<j}^{n_{\mathrm{atoms}}}{}' \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}},$$

```
> source leaprc.protein.ff14SB
```

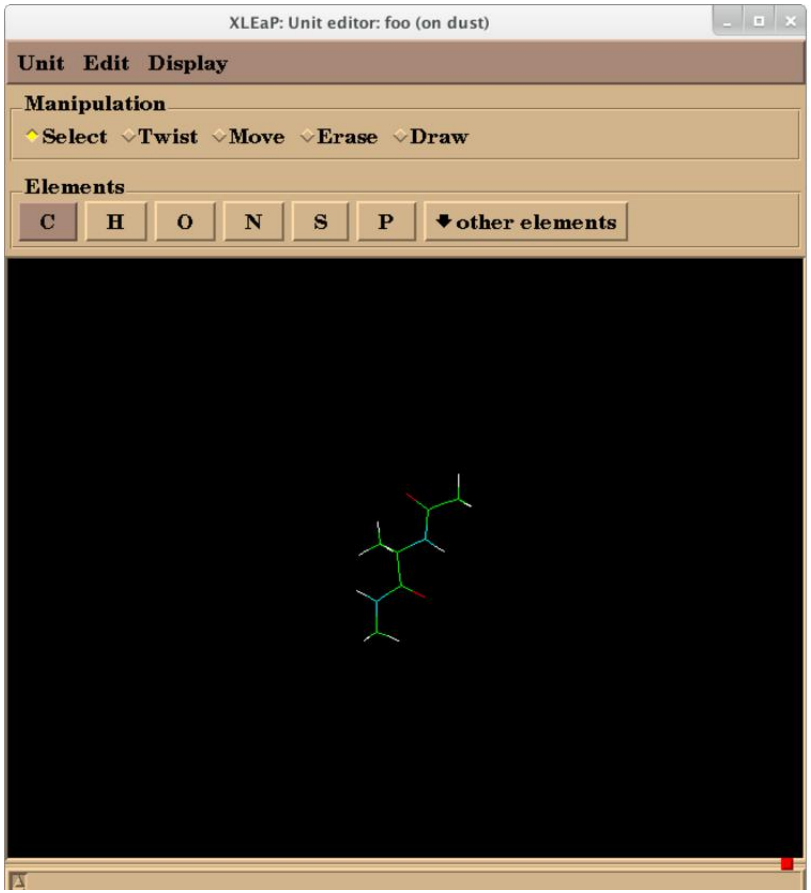
Name of the force field

3. Use **sequence** to create a new unit called **foo** out of the **ACE**, **ALA**, and **NME** units

```
> foo = sequence { ACE ALA NME }
```

4. Use the **edit** command to view the st

```
> edit foo
```



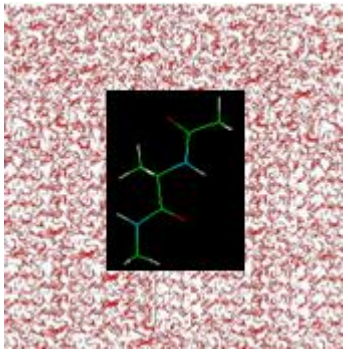
5. Solvate the system with the **solvatebox** command.

```
> solvatebox foo TIP3PBOX 10.0
```

Solvate the molecule by adding TIP3P water molecules

TIP3PBOX specifies the type of water box to solvate with.

10.0 indicates that the molecule should have a buffer of at least 10 Angstroms between alanine dipeptide and the periodic box wall.



MD simulation is performed under PBC

6. Use the **saveamberparm** command to save the prmtop and inpcrd file

```
> saveamberparm foo prmtop inpcrd
```

Prmtop: parameter/topology file

```
%VERSION  VERSION_STAMP = V0001.000  DATE = 06/30/15  11:44:23
%FLAG TITLE
%FORMAT(20a4)
ACE
%FLAG POINTERS
%FORMAT(10I8)
    1912       9    1902       9      25      11      43      24       0       0
    2619     633       9      11      24      13      21      20      10       1
       0       0       0       0       0       0       0       1      10       0
       0
%FLAG ATOM_NAME
%FORMAT(20a4)
HH31CH3 HH32HH33C   O   N   H   CA  HA  CB  HB1 HB2 HB3 C   O   N   H   CH3 HH31
HH32HH33O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2
O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1
H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O
H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2
O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1
H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O
H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2
O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1
H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O
H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2
O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1
H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O
H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2
O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1
H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O
H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2
O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1
H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O
H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2
O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1
H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O
H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2
O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1  H2  O   H1
```

....

You need not to know details

```
%FLAG POINTERS
%FORMAT(10i8)  NATOM,    NTYPES, NBONH,  MBONA,  NTHETH, MTHETA,
               NPHIH,    MPHIA,  NHPARM, NPARM,  NNB,    NRES,
               NBONA,    NTHETA, NPHIA,  NUMBND, NUMANG, NPTRA,
               NATYP,    NPHB,   IFPERT, NBPER,  NGPER,  NDPER,
               MBPER,    MGPER,  MDPER,  IFBOX,  NMXRS,  IFCAP,
               NUMEXTRA, NCOPY
  NATOM   : total number of atoms
  NTYPES  : total number of distinct atom types
  NBONH   : number of bonds containing hydrogen
  MBONA   : number of bonds not containing hydrogen
  NTHETH  : number of angles containing hydrogen
  MTHETA  : number of angles not containing hydrogen
  NPHIH   : number of dihedrals containing hydrogen
  MPHIA   : number of dihedrals not containing hydrogen
  NHPARM  : currently not used
  NPARM   : used to determine if addles created prmtop
  NNB     : number of excluded atoms
  NRES    : number of residues
  NBONA   : MBONA + number of constraint bonds
  NTHETA  : MTHETA + number of constraint angles
  NPHIA   : MPHIA + number of constraint dihedrals
```

```
%FLAG ATOM_NAME
%FORMAT(20a4)   (IGRAPH(i), i=1,NATOM)
  IGRAPH : the user-specified atoms names
```

Inpcrd: coordinate/restart file

Atom 1 **Atom 2**

```
ACE
   1912
  13.6813322  13.1481714  15.2733473  13.6813322  14.2381714  15.2733488
  13.1675952  14.6020204  16.1631727  13.1675902  14.6020234  14.3835287
  15.1087512  14.7889664  15.2733457  16.0719112  14.0255774  15.2733421
  15.2367066  16.1178202  15.2733455  14.4144513  16.7043316  15.2733474
  16.5345934  16.7620967  15.2733444  17.0889272  16.4637102  16.1631639
  17.3426356  16.3690139  14.0412007  16.8045928  16.6695344  13.1421471
  18.3118152  16.8671068  14.0675580  17.4898714  15.2890438  14.0319637
  16.3940071  18.2775899  15.2733501  15.2819757  18.8008742  15.2733549
  17.5273845  18.9830547  15.2733512  18.4183327  18.5073334  15.2733483
  17.5273863  20.4320551  15.2733549  16.4999074  20.7959063  15.2733591
  18.0411296  20.7959028  16.1631770  18.0411212  20.7959068  14.3835300
  27.8556240  25.5899380  22.4329930  27.7750290  24.6753990  22.1621640
  27.4420660  26.0866450  21.7269640  25.9874280  27.3678400  24.3139200
  25.2116380  26.8878840  24.6037770  26.6387210  26.6884050  24.1395360
  20.0040110  27.6798210  25.1914960  19.1724990  27.4246920  25.5911470
  20.6121090  27.7508980  25.9272880  19.9841370  31.7138300  25.3872980
  20.2869410  30.8236040  25.2083200  20.7284960  32.2708520  25.1595230
  18.4132060  22.6400100  18.9757350  17.5248760  22.9816070  18.8737230
  18.6973250  22.4400420  18.0838180  16.3765070  25.2405750  24.8597050
  16.9481180  24.4727950  24.8600690  16.3766610  25.5394510  23.9503660
  25.4283270  30.0435590  24.3209360  25.3591550  29.1025530  24.1598490
  26.1903130  30.1337000  24.8931930  17.6189940  20.6490110  26.6465980
  17.4392780  19.7515860  26.9268740  17.2170020  21.1961980  27.3212930
  26.4588440  22.8505970  24.5478290  26.8353260  21.9777850  24.4351960
  26.3607950  23.1864670  23.6568730  20.5893450  21.2231600  20.0095730
  20.6960200  22.1235330  20.3164490  19.6785810  21.1749570  19.7190380
```

▪ ▪ ▪ ▪ ▪ ▪

6. To quit xLEaP use **quit.**

```
> quit
```

All the above steps are just for preparati
Now let’s get into the real meat.

1. Minimization

2. Heating with constant volume and temperature (NVT) for 20ps from 0K to 300K

3. Production MD with constant pressure and temperature (NPT) at 300K and 1atm for 60ps