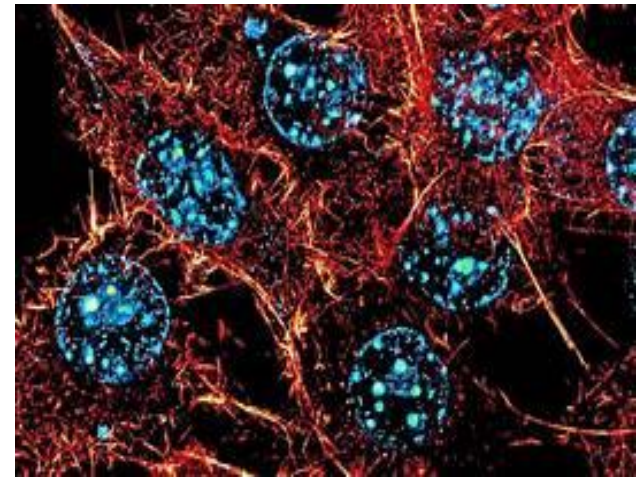
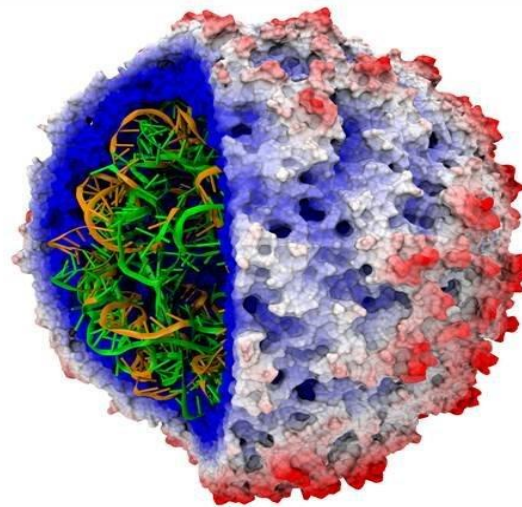
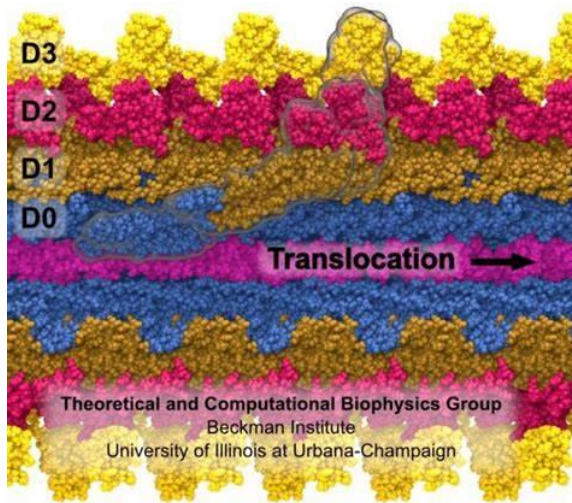


Vật lý Y sinh

Trần Thị Minh Thu
Email: ttmthu@hcmus.edu.vn



TÀI LIỆU THAM KHẢO

1. Vật lý thống kê, Nguyễn Nhật Khanh, Tủ sách Trường ĐHKHTN, 1999.
2. Mô phỏng trong Vật lý, Võ Văn Hoàng (Chủ biên) NXB Đại học Quốc gia TP HCM, 2016.
3. Understanding Molecular Simulation: From algorithms to application, Daan Frenkel and Berend Smit, Second edition, 2005, Academic press.

HÌNH THỨC THI VÀ CHẤM ĐIỂM

- Bài tập cá nhân và điểm danh tại lớp: 20%
- Báo cáo đồ án: 40%
- Điểm cuối kỳ: 40%



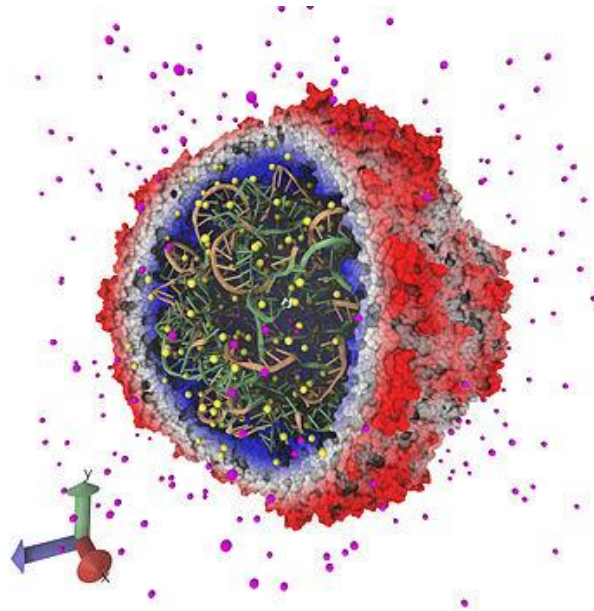
03 TC lý thuyết: 12 buổi; 04 tiết / 01 buổi

- Buổi 1+2: Giới thiệu về protein và khoa học tính toán
- Buổi 3: Phương pháp mô phỏng MD.
- Buổi 4: Kỹ thuật mô phỏng MD: thuật giải, trường lực, các phần mềm.
- Buổi 5: Kỹ thuật mô phỏng MD: thuật giải, trường lực, các phần mềm (tiếp theo).
- Buổi 6+7+8: Giới thiệu và hướng dẫn sử dụng phần mềm VMD, XMGRACE, GROMACS.
- Buổi 9+10: Giao đồ án môn học và hướng dẫn làm đồ án
- Buổi 11+12: Báo cáo đồ án

GROMACS
FAST. FLEXIBLE. FREE.



Buổi 4: Kỹ thuật mô phỏng MD: thuật giải, trường lực, các phần mềm



Phương trình chuyển động

$$-\frac{dE}{dr_i} = m_i \frac{d^2 r_i}{dt^2}$$

$$E(r) = E_{stretch} + E_{bend} + E_{dihedral} + E_{electrostatic} + E_{vdW}$$

Trường lực chứa công thức và các tham số là hằng số, các trường lực khác nhau đôi chút về hàm mô tả tương tác và cách tính các thông số.

Một số trường lực phổ biến là GROMOS (Groningen và ETH Zurich), AMBER (UC Sanfrancisco), OPLS (Yale), CHARMM (Harvard và Maryland) và các biến thể từ những trường lực này.

Introduction to Molecular Dynamics

- Calculate how a system of particles evolves in time
- Consider a set of atoms with positions /velocities and the potential energy function of the system
- Predict the next positions of particles over some short time interval by solving Newtonian mechanics

Basic MD Algorithm

Set initial conditions $\mathbf{r}_i(t_0)$ and $\mathbf{v}_i(t_0)$

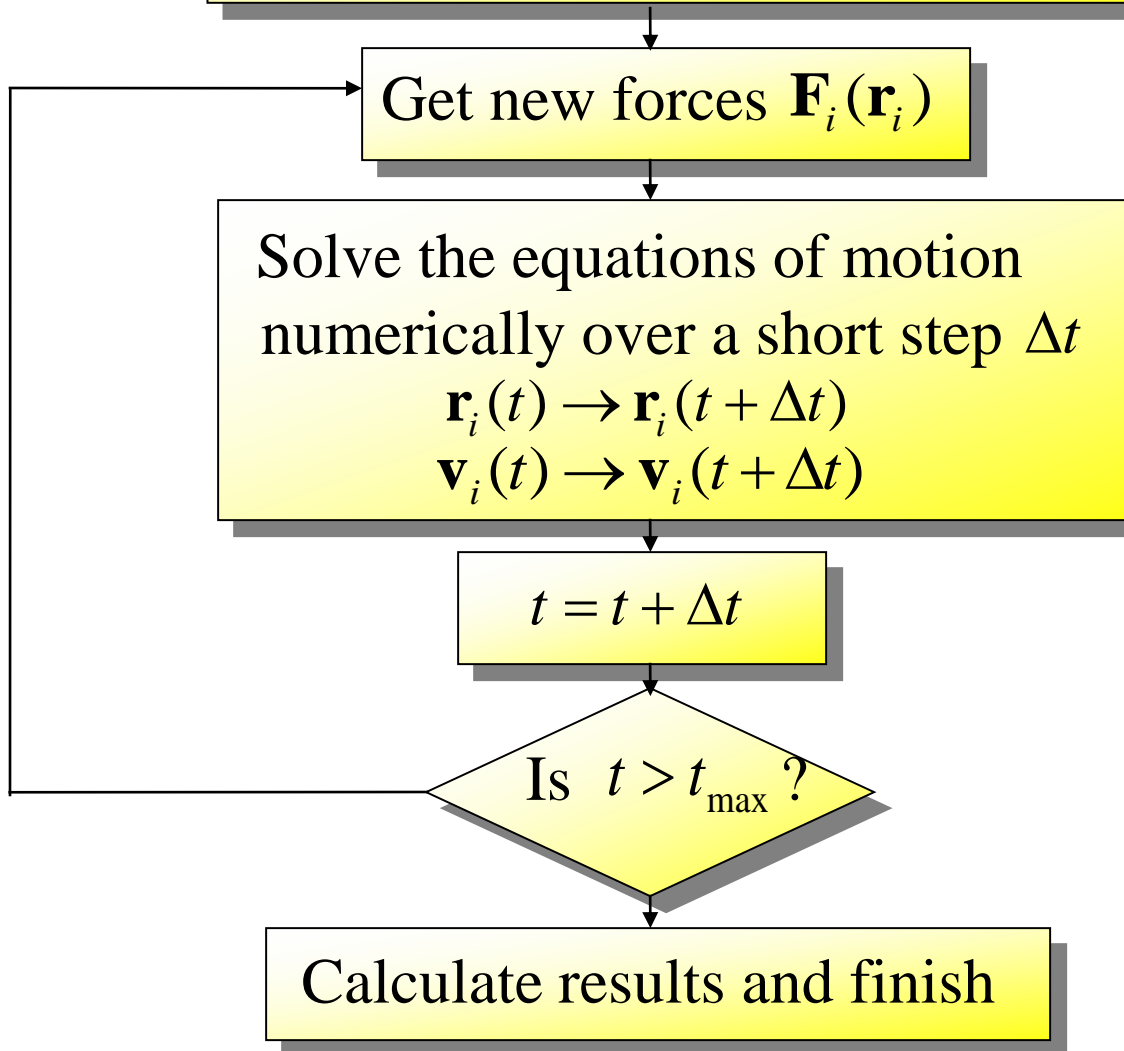
Get new forces $\mathbf{F}_i(\mathbf{r}_i)$

Solve the equations of motion
numerically over a short step Δt
 $\mathbf{r}_i(t) \rightarrow \mathbf{r}_i(t + \Delta t)$
 $\mathbf{v}_i(t) \rightarrow \mathbf{v}_i(t + \Delta t)$

$t = t + \Delta t$

Is $t > t_{\max}$?

Calculate results and finish



Simulation Setup

Simulation Cell

- Boundary Condition
- Constructing neighboring cells

Initial atom velocities

MD Time step

Temperature Control

Simulation Cell

- usually using orthogonal cells

Open boundary

- for a molecule or nanocluster in vacuum
- not for a continuous medium

Fixed boundary

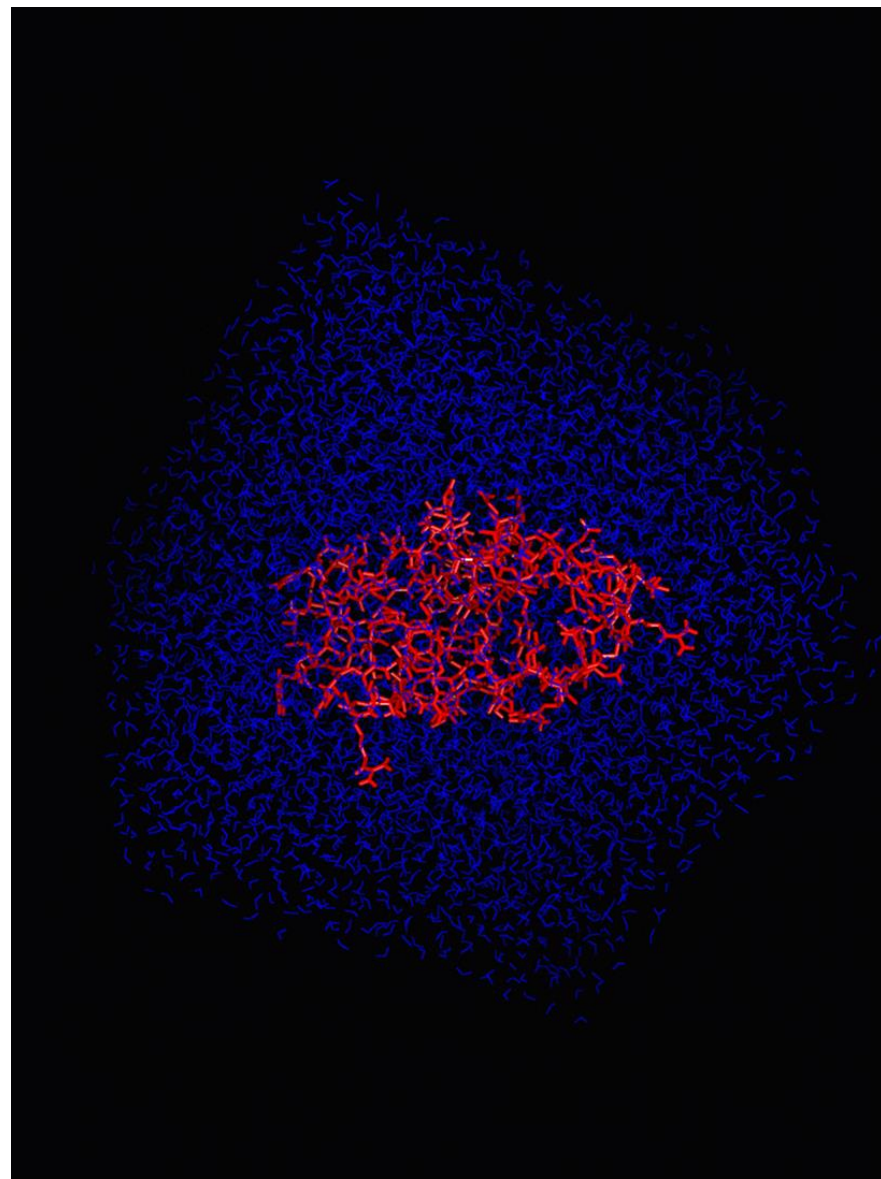
- fixed boundary atoms
- completely unphysical

Periodic boundary conditions

- obtaining bulk properties

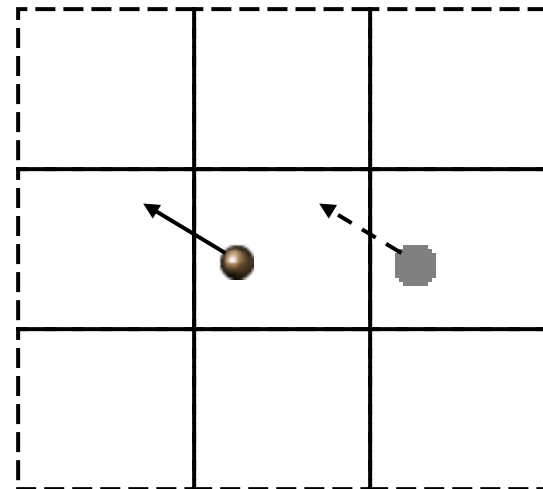
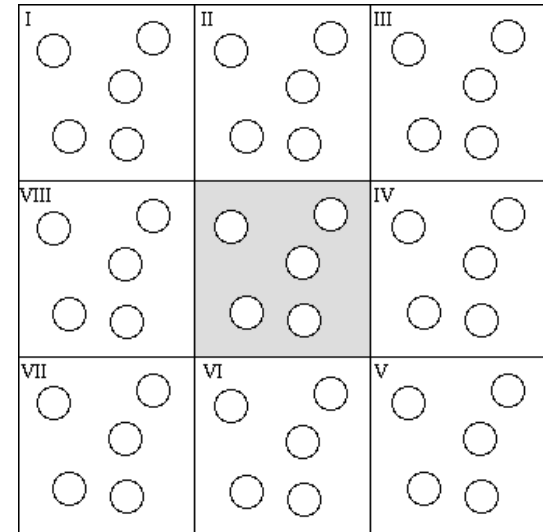
Molecules in Solution

- In real situations, a molecule is rarely isolated. In biological systems, proteins, RNA, and DNA are immersed in a sea of water molecules
- To accurately portray the effect of the solvent molecules on a system, the solvent molecules must be free flowing
- How do we establish computational boundaries while keeping a realistic solvent simulation?



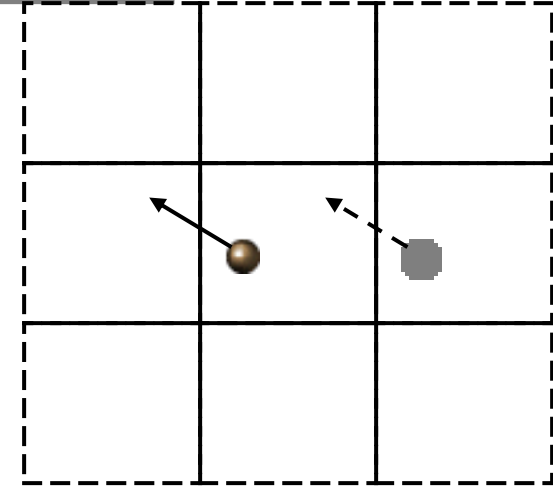
Periodic Boundary Conditions

- Simulate a segment of molecules in a larger solution by having repeatable regions
- Potential calculations are run only on each atom's closest counterpart in the 27 cubes
- When an atom moves off the edge, it reappears on the other side (like in asteroids)



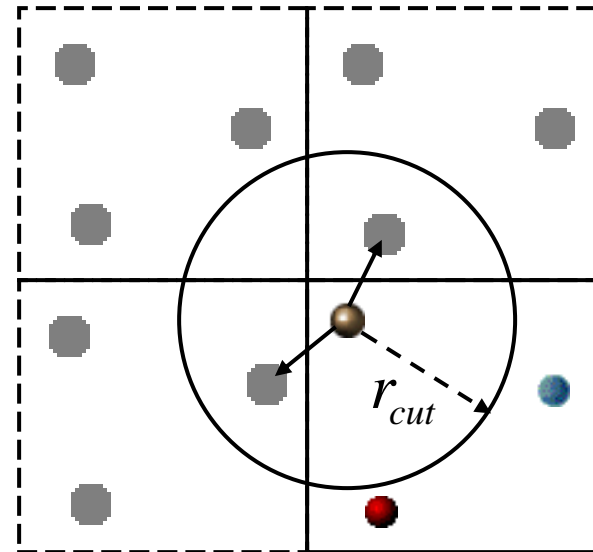
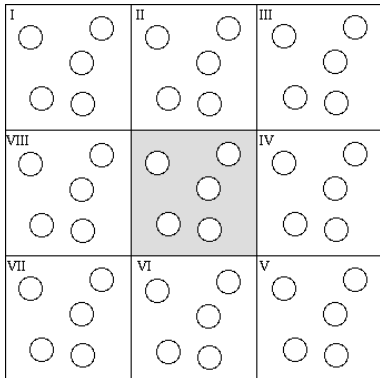
Periodic boundary conditions

■ An atom moving out of boundary comes back on the other side



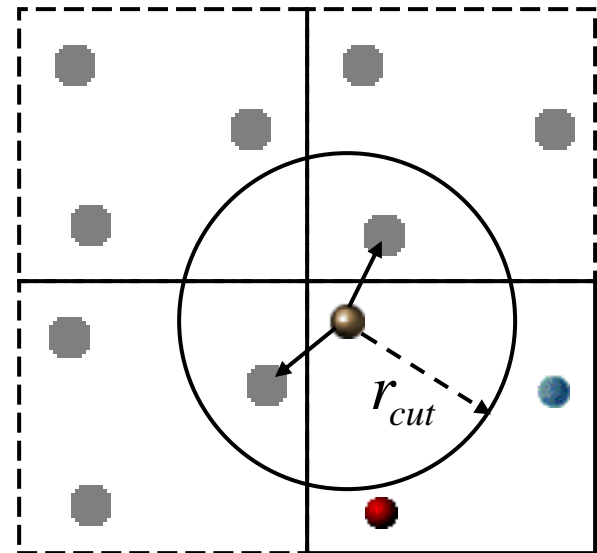
■ considered in force calculation

$$r_{cut} < \frac{L}{2}$$



Cutoff Methods

- Ideally, every atom should interact with every other atom
- This creates a force calculation algorithm of quadratic order
- We may be able to ignore atoms at large distances from each other without suffering too much loss of accuracy



Simulation Setup

Simulation Cell

- Boundary Condition
- Constructing neighboring cells

Initial atom velocities

MD Time step

Temperature Control

Khởi tạo vận tốc ban đầu

Phân bố Maxwell-Boltzmann

- Xác suất tìm thấy hạt với vận tốc \mathbf{v}

$$P(\mathbf{v}_x) = \left(\frac{m}{2\pi k_B T} \right)^{1/2} \exp \left(-\frac{1}{2} m \mathbf{v}_x^2 / k_B T \right)$$

- Khởi tạo vận tốc ở nhiệt độ T

$$\frac{3}{2} k_B T = \frac{1}{2} m \mathbf{v}^2$$

Simulation Setup

Simulation Cell

- Boundary Condition
- Constructing neighboring cells

Initial atom velocities

MD Time step

Temperature Pressure Control

Simulation Setup

Simulation Cell

- Boundary Condition
- Constructing neighboring cells

Initial atom velocities

MD Time step

Temperature, Pressure Control

Cân bằng nhiệt độ

Thuật toán Berendsen:

- Hệ ban đầu được thiết lập tại nhiệt độ T_0 sau đó thay đổi nhiệt độ của hệ theo hàm mũ để ổn định nhiệt độ tại T_0 .
- Để hiệu chỉnh nhiệt độ của hệ luôn ở giá trị T_0 , tốc độ thay đổi nhiệt độ của hệ được cho bởi phương trình:

$$\frac{dT}{dt} = \frac{T_0 - T}{\tau}$$

T_0 : nhiệt độ cố định mà ta mong muốn,

T : nhiệt độ tức thời tại bước thời gian,

τ : là hằng số điều chỉnh.

Giải thuật Berendsen trong cân bằng nhiệt độ cho tập các cấu hình của hệ không tuân theo phân bố chính tắc, do đó giải thuật này chỉ phù hợp để cân bằng hệ trước khi mô phỏng.

Cân bằng nhiệt độ

Thuật toán V-rescale:

- Thêm vào phương trình hiệu chỉnh nhiệt một thông số ngẫu nhiên để đảm bảo phân bố của động năng luôn chính xác bằng cách tự hiệu chỉnh chính nó.
- Phương trình hiệu chỉnh theo thuật toán v-rescale:

$$dK = (K - K_0) \frac{dt}{\tau_T} + 2 \sqrt{\frac{K K_0}{N_f}} \frac{dW}{\tau_T}$$

K : động năng, τ_T : thời gian hồi phục, N_f : bậc tự do của hệ và dW : nhiễu Wiener.

Tạo ra các cấu hình tuân theo tập hợp thống kê chính tắc → được sử dụng nhiều trong mô phỏng động học phân tử.

Cân bằng áp suất

- **Giải thuật Berendsen:** được sử dụng mặc định trong Gromacs. Thuật toán Berendsen dựa trên ý tưởng thay đổi tọa độ và các vector hộp trong mỗi bước mô phỏng để tái tạo lại áp suất thực. Công thức hiệu chỉnh tương tự như cân bằng nhiệt độ:

$$\frac{dP}{dt} = \frac{P_0 - P}{\tau_P}$$

P_0 : áp suất cố định mà ta mong muốn,

P : áp suất tức thời tại bước thời gian,

τ_P : là hằng số điều chỉnh.

Giải thuật Berendsen trong cân bằng áp suất cũng cho tập các cấu hình của hệ không tuân theo phân bố chính tắc, do đó giải thuật này chỉ phù hợp để cân bằng hệ trước khi mô phỏng.

Cân bằng áp suất

- **Giải thuật Parrinello-Rahman:**

- Ý tưởng: thay đổi phương trình chuyển động của các hạt bằng cách thêm một thông số đặc trưng làm thay đổi kích thước hộp mô phỏng.
- Các vector hộp mô phỏng cũng thay đổi theo một phương trình chuyển động.
- Các cấu hình sau khi mô phỏng của hệ thỏa điều kiện tập hợp thống kê chính xác tốt hơn Berendsen nên giải thuật này có thể sử dụng trong mô phỏng động học phân tử.

Tích tích phân phương trình chuyển động: Giải thuật Verlet

Tọa độ của hạt ở thời điểm $t + \Delta t$ được biểu diễn theo tọa độ, vận tốc và gia tốc tại thời điểm t áp dụng khai triển Taylor như sau [195]:

$$r_i(t + \Delta t) \approx r_i(t) + \dot{r}_i(t)\Delta t + \frac{1}{2}\ddot{r}_i(t)\Delta t^2 + \frac{1}{3!}\dddot{r}_i(t)\Delta t^3 + O(\Delta t^4) \quad (3.4)$$

Ở đây chỉ sử dụng khai triển đến bậc 2, i là vị trí số hạt trong hệ và i chạy từ 1 đến tổng số hạt mà hệ có. Sử dụng $\dot{r}_i(t) = v_i(t)$ và $\ddot{r}_i(t) = \frac{F_i}{m_i}$. Công thức (3.6) có thể viết lại:

$$r_i(t + \Delta t) \approx r_i(t) + \Delta t v_i(t) + \frac{\Delta t^2}{2m_i} F_i(t) \quad (3.5)$$

Khai triển tương tự cho $r_i(t - \Delta t)$, ta có:

$$r_i(t - \Delta t) \approx r_i(t) - \Delta t v_i(t) + \frac{\Delta t^2}{2m_i} F_i(t) \quad (3.6)$$

Cộng hai phương trình (2.7) và (2.8), ta được:

$$r_i(t + \Delta t) + r_i(t - \Delta t) \approx 2r_i(t) + \frac{\Delta t^2}{m_i} F_i(t) \quad (3.7)$$

Phương trình (3.7) là công thức chính của giải thuật Verlet. Như vậy, nếu cho tọa độ ban đầu của hệ $r_i(0), \dots, r_N(0)$ và vận tốc ban đầu $v_i(0), \dots, v_N(0)$, phương trình (3.4) có thể tạo ra quỹ đạo chuyển động với khoảng thời gian bất kì. Vận tốc các hạt của hệ có thể tính được qua (3.8):

$$v_i(t) \approx \frac{1}{2\Delta t} [r_i(t + \Delta t) - r_i(t - \Delta t)] \quad (3.8)$$

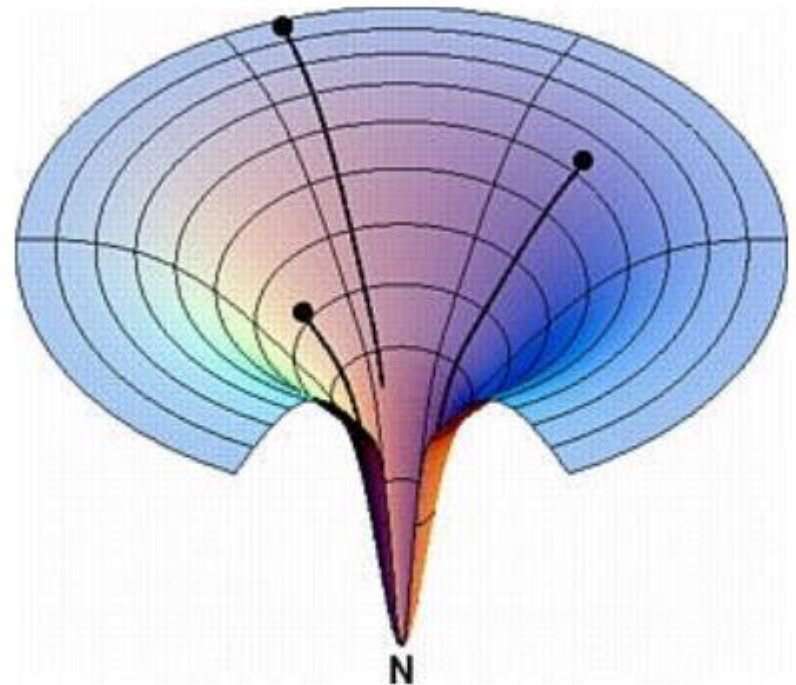
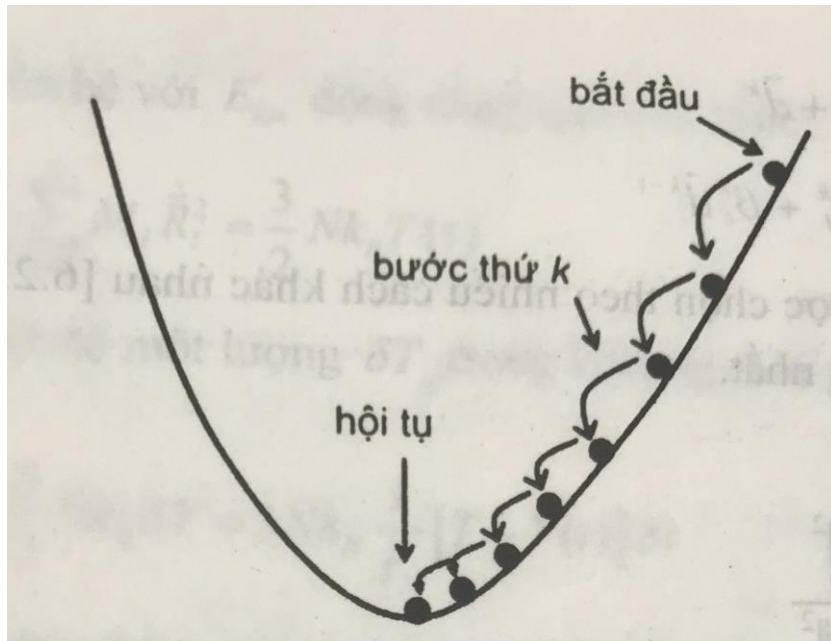
Hệ protein trước khi mô phỏng MD cần:

1. Cực tiểu hóa năng lượng
2. Cân bằng nhiệt độ
3. Cân bằng áp suất

Cực tiểu hóa năng lượng

Mục đích:

- Ổn định lại các phân tử nước,
- Tránh hiện tượng chồng lấn nguyên tử và đứt gãy các liên kết.



Cực tiểu hóa năng lượng

Phương pháp Steepest descent

- Gradient \vec{G}_I của thế tương tác tương ứng với tọa độ nguyên tử I có dạng:

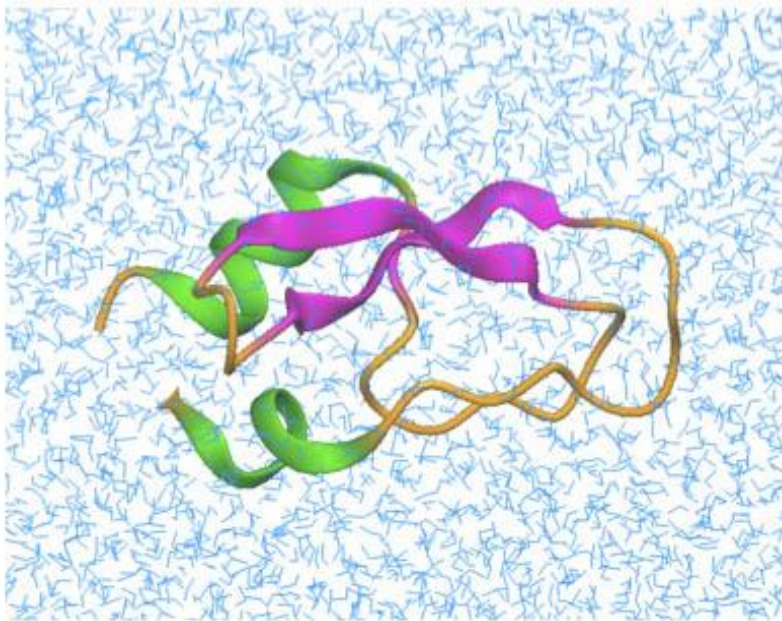
$$\vec{G}_I = \vec{\nabla}_I U(\{\vec{R}_I\}) = \frac{\partial}{\partial \vec{R}_I} U(\{\vec{R}_I\})$$

- Gradient \vec{G}_I chỉ hướng di chuyển của nguyên tử I sao cho tốc độ tăng của thế tương tác là nhanh nhất.
- Năng lượng của hệ giảm nếu di chuyển I theo chiều ngược lại với \vec{G}_I , cấu trúc của protein tương ứng cực tiểu địa phương gần nhất được tìm bằng cách lặp phương trình sau:

$$\vec{R}_I^{k+1} = \vec{R}_I^k - \gamma \vec{G}_I^k$$

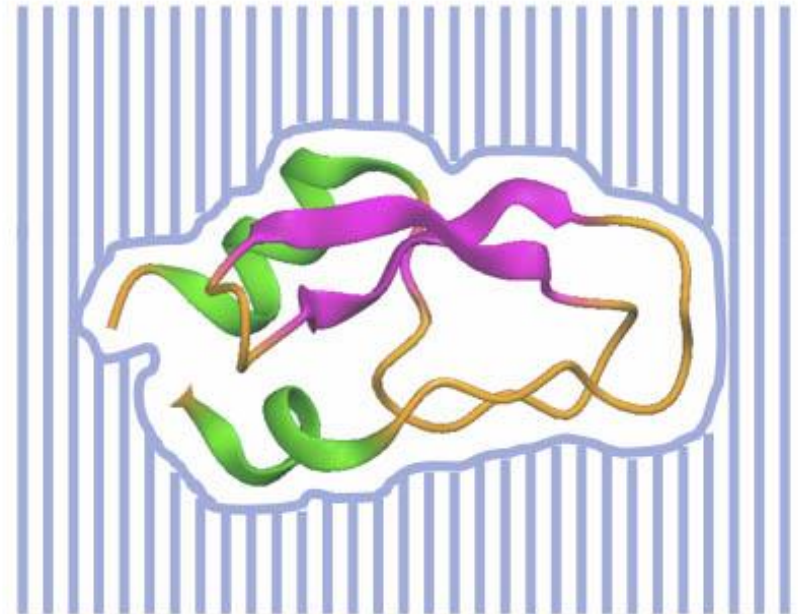
Mô hình nước trong mô phỏng

Explicit Water



$$H = H_{\text{prot}} + H_{\text{prot-wat}} + H_{\text{wat}}$$

Implicit Solvent



$$H = H_{\text{prot}} + \Delta G_{\text{solv}}$$

Mô hình nước tường minh

Mô hình nước tường minh sử dụng các phân tử nước đã được mô hình hóa.

- Năm 1974, Stillinger đã tạo ra mô hình nước đầu tiên có tên là ST2 dựa trên mô hình 4 điểm tích điện .
- Mô hình SPC và TIP3P được xây dựng bởi Berendsen và Jorgensen cùng các cộng sự. Hai mô hình này sử dụng ba điểm cố định để biểu diễn phân tử nước.
- Mô hình TIP4P thêm một điểm thuộc đường phân giác của góc H-O-H.

Các mô hình nước được xây dựng sao cho mô hình các điện tích điểm phù hợp với dạng của thế năng tương tác giữa hai phân tử nước.

Mô hình nước tường minh

Thế năng tương tác này là tổng của tương tác tĩnh điện giữa các cặp nguyên tử mang điện tích trái dấu và thế năng Lenard-Jones giữa các nguyên tử O được xác định thông qua năng lượng đôi của hai phân tử nước m và n :

$$\epsilon_{mn} = \sum_i^m \sum_j^n \frac{q_i q_j e^2}{r_{ij}} + \frac{A}{r_{OO}^{12}} - \frac{C}{r_{OO}^6}$$

- Các giá trị điện tích của nguyên tử Hydro, hằng số A và C được chọn sao cho kết quả tính toán phù hợp nhất với thực nghiệm.
- Gromacs hỗ trợ các mô hình nước SPC, SPC/E, TIP3P, TIP4P, TIP5P

Mô hình nước tường minh

	SPC	TIP3P	BF	TIPS2	TIP4P
$r(\text{OH}), \text{\AA}$	1.0	0.9572	0.96	0.9572	0.9572
HOH, deg	109.47	104.52	105.7	104.52	104.52
A	629.4	582.0	560.4	695.0	600.0
C	625.5	595.0	837.0	600.0	610.0
$q(\text{O})$	-0.82	-0.834	0.0	0.0	0.0
$q(\text{H})$	0.41	0.417	0.49	0.535	0.52
$q(\text{M})$	0.0	0.0	-0.98	-1/07	-1/04
$r(\text{OM}), \text{\AA}$	0.0	0.0	0.15	0.15	0.15

Mô hình nước không tường minh

- Giả lập môi trường tương tác giữa các phân tử protein và các phân tử nước bởi một trường trung bình được mô tả bởi một hàm thế tương tác chỉ phụ thuộc vào tọa độ của các nguyên tử protein, G^{solv} .
- Mô hình nước không tường minh bỏ qua độ nhớt của chất lỏng nên cấu hình protein thường nhiều hơn vì quá trình chuyển đổi cấu trúc xảy ra nhanh hơn.
- Các mô hình dung môi không tường minh được tính toán dựa trên hình thức luận Born tổng quát và được sử dụng trong phần mềm Gromacs là Still, HCT và OBC.
- Theo mô hình Born, năng lượng tự do G^{solv} là tổng của các số hạng bao gồm G^{np} là năng lượng hòa tan tự do gây ra bởi các tương tác không gây ra bởi điện tích hay moment lưỡng cực và G^{pol} là năng lượng tự do tạo bởi tương tác tĩnh điện giữa protein và nước. Có thể biểu diễn năng lượng tự do G^{solv} qua hai số hạng này như sau:

$$G^{solv} = G^{np} + G^{pol}$$

Mô hình nước không tường minh

G^{np} là năng lượng cần thiết để di dời các nguyên tử nước nhằm tạo nên không gian giữa protein cộng với tương tác van der Waals giữa protein và nước. G^{np} tỉ lệ với diện tích tiếp xúc của protein và nước và được biểu diễn bởi:

$$G^{np} = \gamma SA$$

γ : Hằng số bán thực nghiệm có thứ nguyên của sức căng bề mặt

SA: Diện tích tiếp xúc của protein và nước

Mô hình nước không tường minh

- G^{pol} được xác định từ phương trình Born tổng quát:

$$G^{pol} = \left(1 - \frac{1}{\varepsilon}\right) \sum_{i=1}^n \sum_{j>i}^n \frac{q_i q_j}{\sqrt{r_{ij}^2 + b_i b_j \exp\left(\frac{-r_{ij}^2}{4b_i b_j}\right)}}$$

- q_i và q_j : điện tích của nguyên tử thứ i và j
- r_{ij} : là khoảng cách giữa hai nguyên tử i và j
- ε : là hằng số điện môi
- b_i, b_j : là bán kính Born của các nguyên tử i, j , có độ dài bằng khoảng cách từ tâm nguyên tử đến bề mặt tiếp xúc (các mẫu nước khác nhau có cách tính bán kính Born khác nhau)