

Simulações computacionais de biomoléculas

Roberto K. Salinas

QBQ2505

IQ-USP

Simulações computacionais de sistemas químicos



Simulações computacionais de sistemas químicos



Simulações computacionais de sistemas químicos



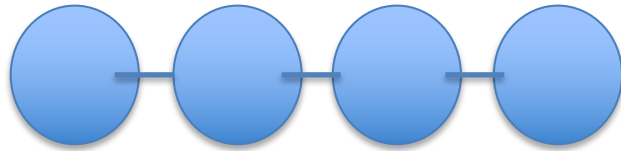
Simulações computacionais de sistemas químicos

- Simulação numérica do comportamento do sistema molecular
- Geração de um *ensemble* de estados em equilíbrio
- *Ensemble*: conjunto de configurações estatisticamente representativas
- $\langle A \rangle$: As propriedades do sistema são definidas como médias no *ensemble*

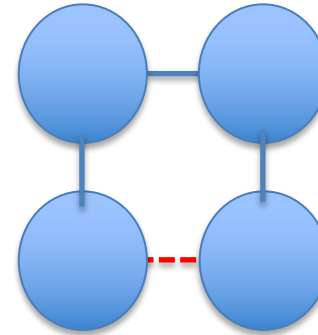
Representação da estrutura

- Depende da propriedade na qual estamos interessados
- Propriedades que não dependem dos elétrons
 - Coordenadas atômicas
 - Leis da mecânica clássica
 - Modelos “coarse grained”
 - Modelos atomísticos
- Reações químicas
 - Coordenadas dos elétrons
 - Densidade eletrônica
 - Equação de Schrödinger

Modelos “coarse grained”



$U=0$



$U=-\epsilon_0$

Propriedade de interesse, grau de exatidão e poder computacional

- Campo de força: função de energia
- Representação da molécula
- Tempo de cálculo
- Amostragem do espaço de fase (espaço conformacional)

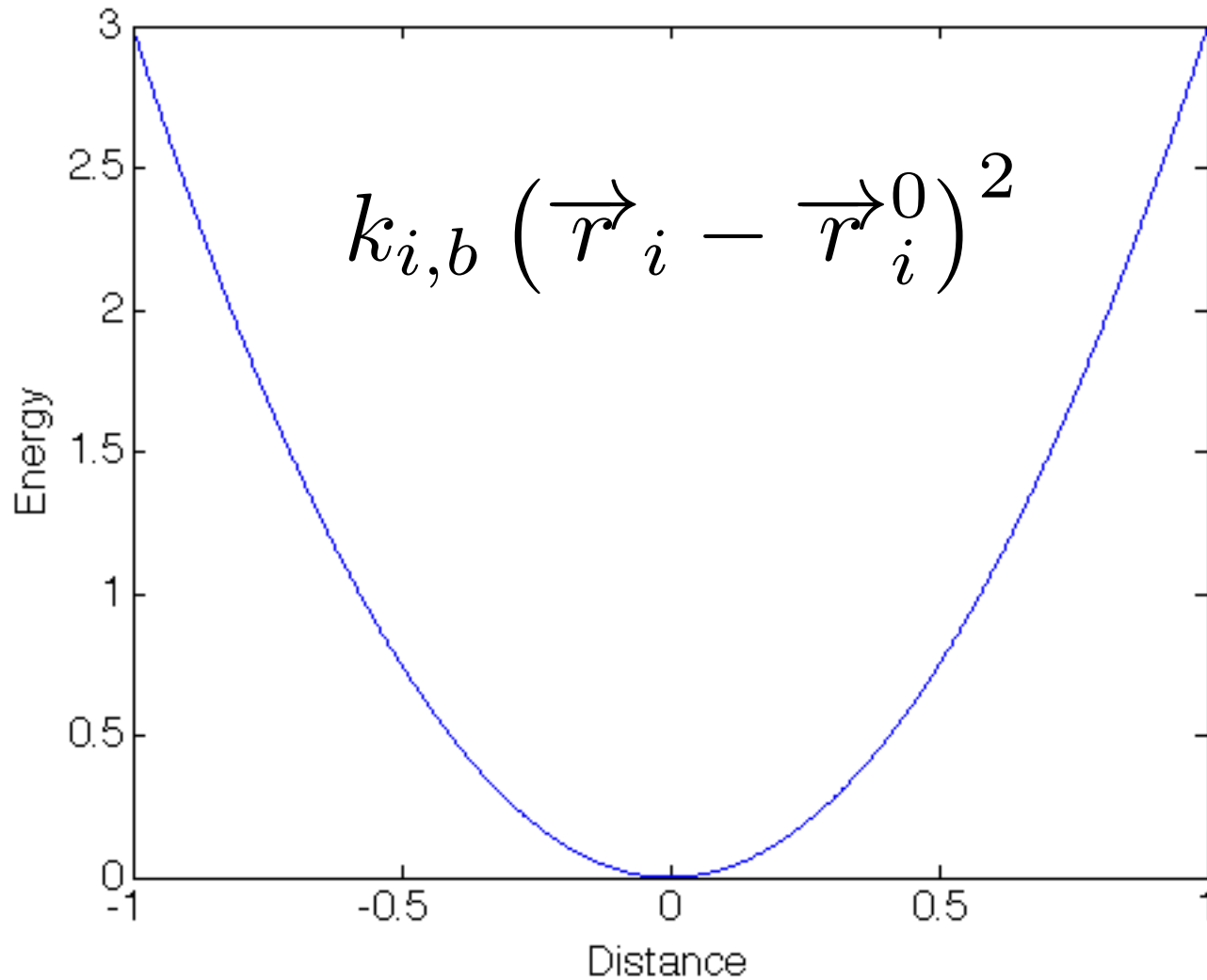
Utilidade

- Permite correlacionar o comportamento macroscópico com fenômenos microscópicos
- Obtenção de estimativas para propriedades que não podem ser determinadas experimentalmente
- Refinamento de estruturas determinadas por difração de raios-X, RMN e Cryo-EM

Campo de força

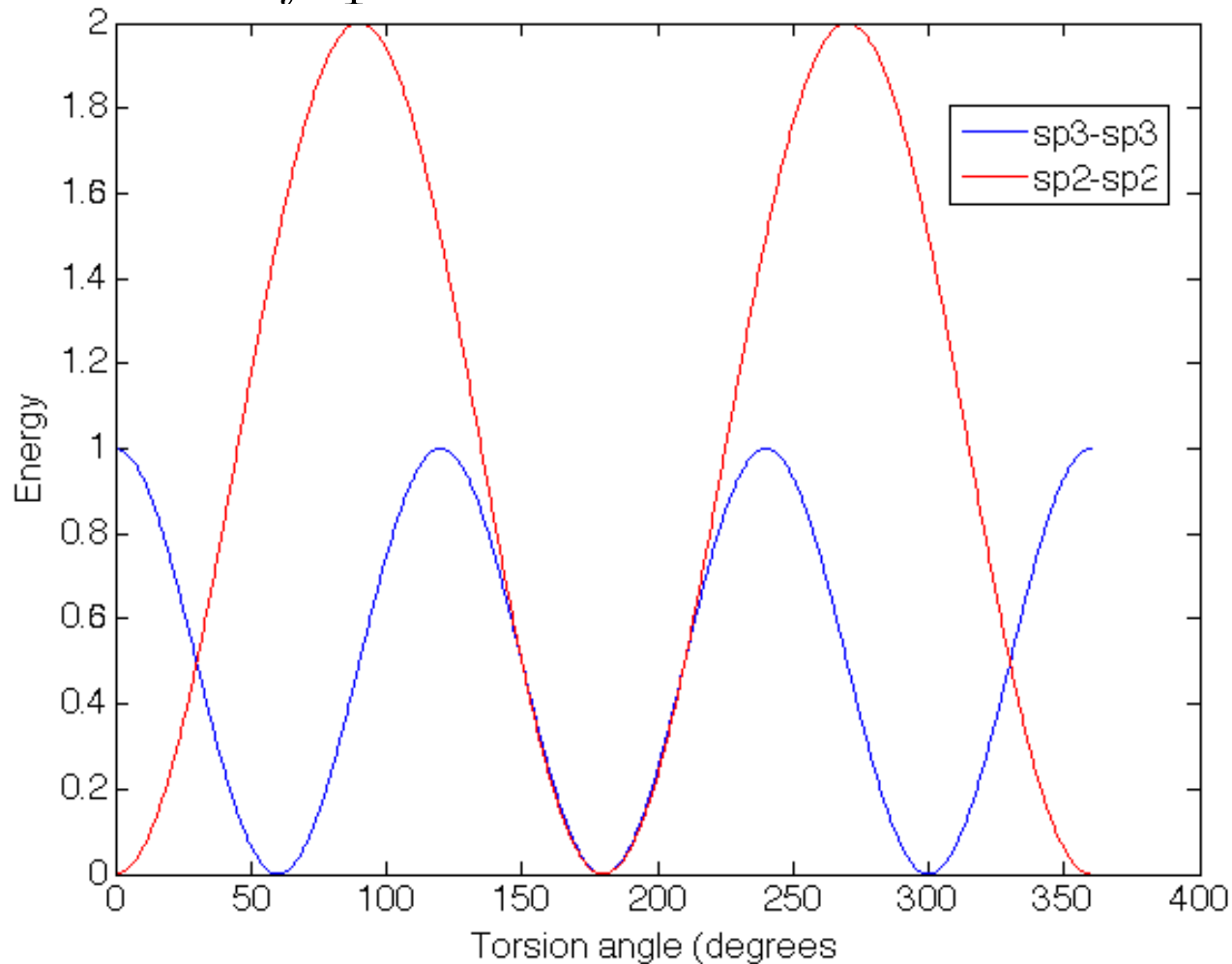
$$\begin{aligned} V(\vec{r}) = & \sum_{i=1}^N k_{i,b} (\vec{r}_i - \vec{r}_i^0)^2 + \sum_{i=1}^N k_{i,\theta} (\theta_i - \theta_i^0)^2 + \sum_{i=1}^N k_{i,\varphi} (1 + \cos(n\varphi - \gamma)) + \\ & + \sum_{i,j} \left[\frac{Cq_i q_j}{r_{i,j}} + \epsilon \left(\frac{\alpha^{12}}{r^{12}} - \frac{\alpha^6}{r^6} \right) \right] \end{aligned} \quad (1)$$

Potencial harmônico



Potencial de ângulos torsionais

$$\sum_{i=1}^N k_{i,\varphi} (1 + \cos(n\varphi - \gamma))$$



Parametrização do campo de força

- Quais são os valores para as distâncias e ângulos de equilíbrio, cargas parciais, coeficientes de van der Waals, etc. ?
- Quais são os valores para as constantes de força?
- Os parâmetros podem ser ajustados a dados experimentais (estruturas cristalográficas, entalpia de vaporização, infra-vermelho, etc.)

Parâmetros do campo de força

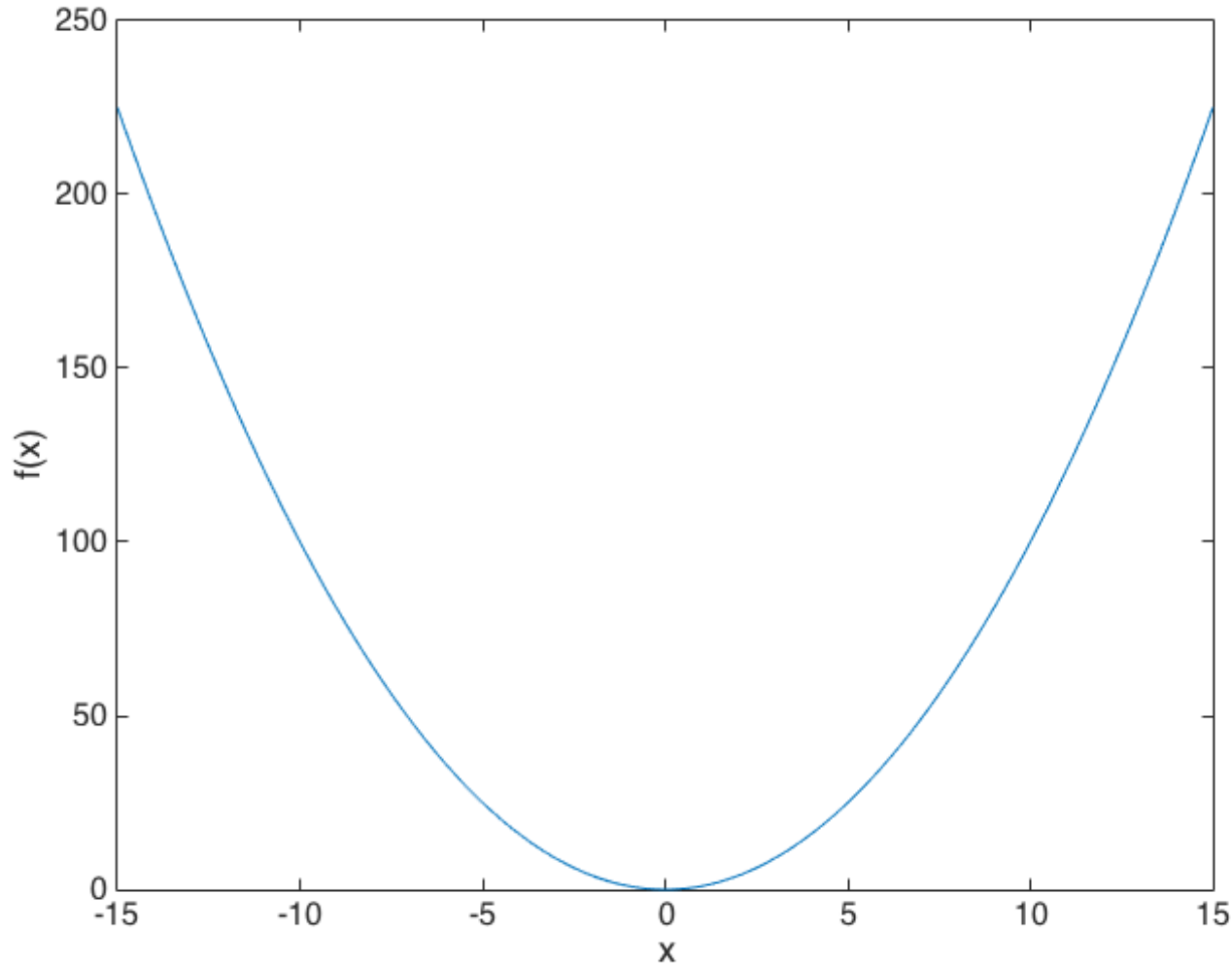
Ligação	r_0 (Å)	k (kcal.mol ⁻¹ Å ⁻²)
Csp3 – Csp3	1.523	317
Csp3 – Csp2	1.497	317
Csp2 = Csp2	1.337	690
Csp2 = O	1.208	777
C – N	1.345	719

Ângulo	θ_0 (°)	k (kcal.mol ⁻¹ grau ⁻¹)
Csp3 – Csp3 – Csp3	109.47	0.0099
Csp3 – Csp3 – H	109.47	0.0079
H - Csp3 – H	109.47	0.0070
Csp3 - Csp2 – Csp3	117.2	0.0099
Csp3 – Csp2 = Csp2	121.4	0.0121

Campos de força mais comuns

- MM2 (fase gasosa de moléculas pequenas)
- AMBER (proteínas, ácidos nucleicos)
- CHARMM (proteínas ácidos nucleicos)
- GROMOS (biomoléculas em solvente explícito)
- OPLS (biomoléculas em solvente explícito)

Minimização de energia

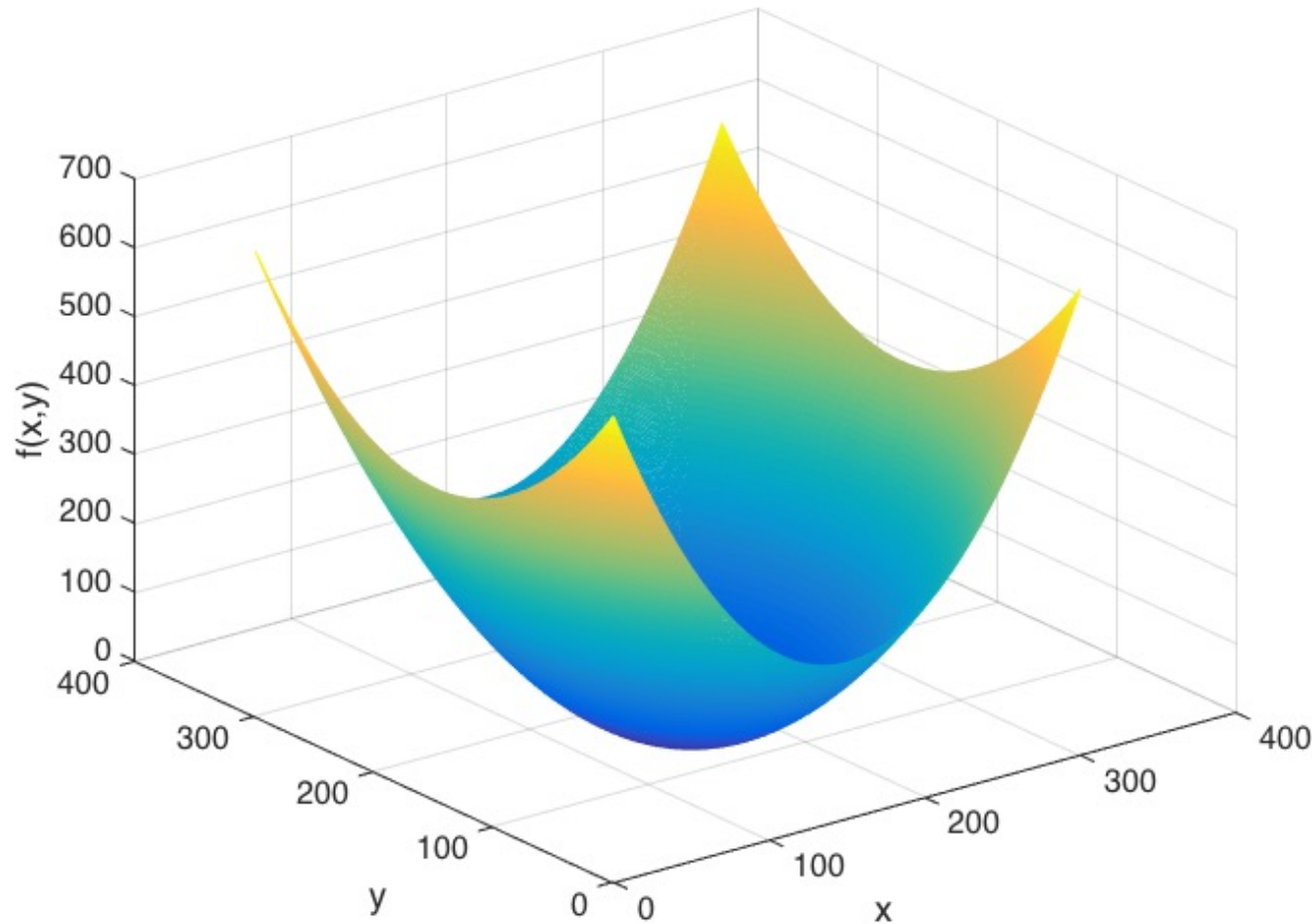


$$\frac{df(x)}{dx} = 0,$$

$$\frac{d^2 f(x)}{dx^2} > 0$$

Ponto de mínimo

Minimização de energia



$$\left(\frac{\partial f(x, y)}{\partial x} \right)_y = 0$$

$$\left(\frac{\partial f(x, y)}{\partial y} \right)_x = 0$$

$$\left(\frac{\partial^2 f(x, y)}{\partial x^2} \right)_{y^*} > 0$$

$$\left(\frac{\partial^2 f(x, y)}{\partial y^2} \right)_{x^*} > 0$$

Ponto de mínimo

Minimização de energia

$$V = V(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n)$$

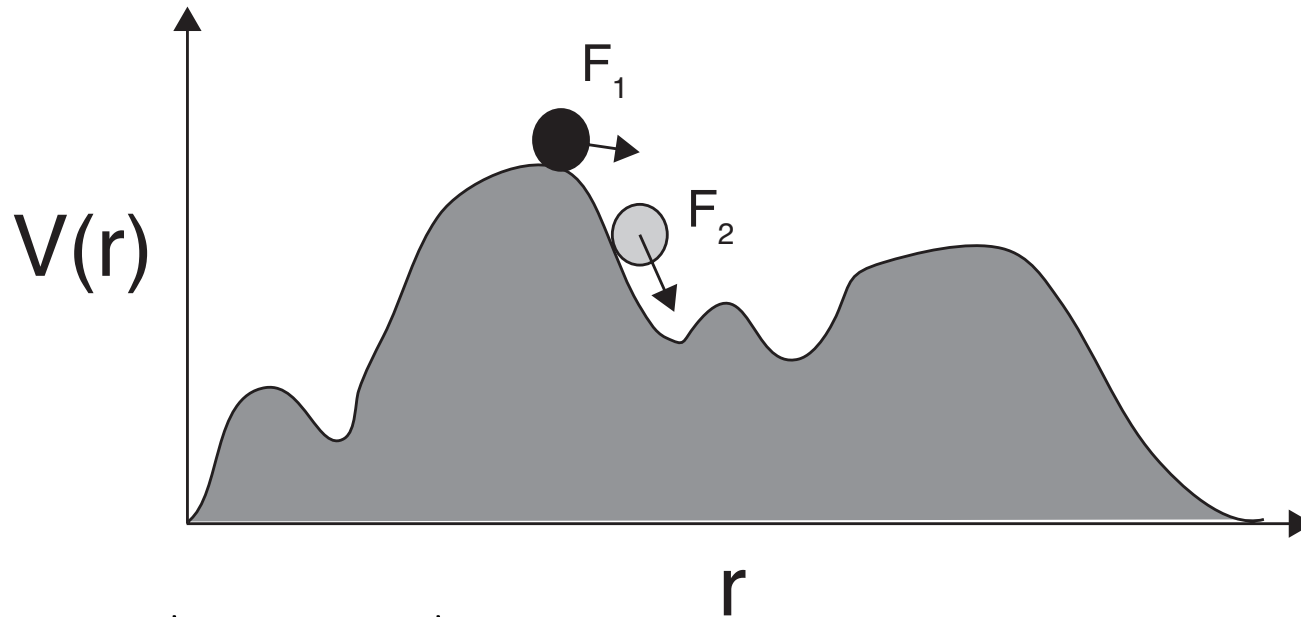
Superfície de energia multidimensional

$$\left. \begin{array}{l} \frac{\partial V}{\partial \vec{r}_i} = 0 \\ \frac{\partial^2 V}{\partial \vec{r}_i^2} > 0, \end{array} \right\} \begin{array}{l} \text{Condições que} \\ \text{satisfazem o mínimo} \\ \text{local} \end{array}$$

Como calcular a derivada?

$$\frac{\partial V}{\partial \vec{r}} = \frac{V(\vec{r} + \Delta \vec{r}) - V(\vec{r} - \Delta \vec{r})}{2\Delta \vec{r}}$$

Steepest descent



$$\vec{r}_{new} = \vec{r} + \Delta \vec{r}$$

$$\vec{s} = \frac{\vec{g}}{|\vec{g}|}$$

$$\vec{r} = \begin{bmatrix} x_1 & x_2 & \cdots & x_{n-1} & x_n \\ y_1 & y_2 & \cdots & y_{n-1} & y_n \\ z_1 & z_2 & \cdots & z_{n-1} & z_n \end{bmatrix}$$

$$\vec{g} = -\nabla V(\vec{r}) = \vec{F}(\vec{r})$$

Minimização de energia

- Em UCSF Chimera gere a estrutura de um peptídeo com a sequência KKKKPKKKK
- Em conformação de α -hélice (-57° , -47°)
- Adicione hidrogênios
- Analise o gráfico de Ramachandran
- Minimize a estrutura em 2000 passos de “steepest descent”, 100 passos de “gradient minimization”, e $\Delta r = 0.2 \text{ \AA}$
- Analise o gráfico de Ramachandran novamente

Dinâmica molecular

- Superfície de energia potencial de uma molécula contém muitos mínimos locais
- Quando os mínimos locais são conhecidos, podemos utilizar mecânica estatística para calcular propriedades
- Por DM podemos gerar configurações do sistema e calcular propriedades
- Podemos acompanhar flutuações conformacionais em função do tempo (dinâmica)

Dinâmica molecular

- Medidas experimentais são médias no *ensemble*
- Uma trajetória de DM gera novas configurações em função do tempo

- Média no tempo:

$$\langle A \rangle = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_{t=0}^{\tau} A(\vec{r}, \vec{p}) dt$$

- Hipótese ergódica:

$$\langle A \rangle_{ensemble} = \langle A \rangle_{tempo}$$

Dinâmica molecular

- Devemos acompanhar o sistema por tempo longo o suficiente tal que

$$\langle A \rangle_{ensemble} = \langle A \rangle_{tempo}$$

- Conjuntos de conformações são determinados sequencialmente
- Determinístico: configuração no tempo futuro é predita a partir da configuração atual

Como gerar novas configurações?

Algoritmo de Verlet

$$\vec{F}_i = \frac{V(\vec{r}_i + \Delta \vec{r}_i) - V(\vec{r}_i - \Delta \vec{r}_i)}{2\Delta \vec{r}_i}$$

$$\frac{\vec{F}_i}{m_i} = \frac{d^2 \vec{r}_i}{dt^2} = \frac{\vec{r}_i(t + \Delta t) + \vec{r}_i(t - \Delta t) - 2\vec{r}_i(t)}{2\Delta t^2}$$

$$\vec{r}_i(t + \Delta t) = 2\vec{r}_i - \vec{r}_i(t - \Delta t) + 2\Delta t^2 \frac{\vec{F}_i}{m_i}$$

Hamiltoniano clásico

$$H(p_i, r_i) = \sum_{i=1}^N \frac{1}{2m_i} p_i^2 + V(r_i)$$

$$E_{\text{tot}} = E_{\text{pot}} + E_{\text{kin}}$$

Como calcular as velocidades?

$$v_i(t) = \frac{r_i(t + \Delta t) - r_i(t - \Delta t)}{2\Delta t}$$

Algoritmo de Velocity-Verlet

$$r_i(t + \Delta t) = r_i(t) + \frac{dr_i}{dt} \Delta t + \frac{\Delta t^2}{2} \frac{F_i}{m_i}$$

$$v_i(t + \Delta t) = v(t) + \frac{\Delta t}{2m_i} (F_i(t) + F_i(t + \Delta t))$$

Temperatura e velocidades


Teoria cinética dos gases

$$\frac{3}{2}kT = \frac{m \langle v^2 \rangle}{2}$$

Distribuição gaussiana de velocidades

$$p_v = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{v^2}{2\sigma}} \quad \sigma = \left(\frac{kT}{m} \right)^{\frac{1}{2}}$$

Step by step

- (i) definir a composição do sistema, número e tipo de átomos, massas, cargas parciais, potenciais de interação
 - (ii) Definir as posições e as velocidades iniciais ($t=0$)
 - (iii) Definir o Δt , e o número de integrações (duração da simulação)
 - (iv) Calcular a força em cada átomo para $t=0$
 - (v) Integração:
 - (a) Calcular as posições em $t+\Delta t$ (velocity-verlet)
 - (b) Calcular as forças em $t+\Delta t$ (slide 24)
 - (c) Calcular as velocidades em $t+\Delta t$ (velocity-verlet)
 - (d) Armazenar posições e velocidades, energia, etc.
 - (e) incrementar o Δt
- 

Etapas de uma simulação

- Minimização de energia para evitar forças muito grandes
- Etapa de aquecimento (p. ex. 10 a 300 K)
- Etapa de equilíbrio (temp. mantida constante por re-scaling); duração depende do tamanho da molécula
- Etapa produtiva (sem modificação temperatura)

Questões chave

- Qual é o Δt adequado?
- Qual é a configuração inicial?
- Quanto tempo eu preciso simular?
- Verificar se a energia total é conservada
- Verificar se a temperatura permaneceu estável

Análise

- Séries temporais

- Médias

$$\langle X \rangle = \frac{1}{n_t} \sum_{n=1}^{n_t} X_n$$

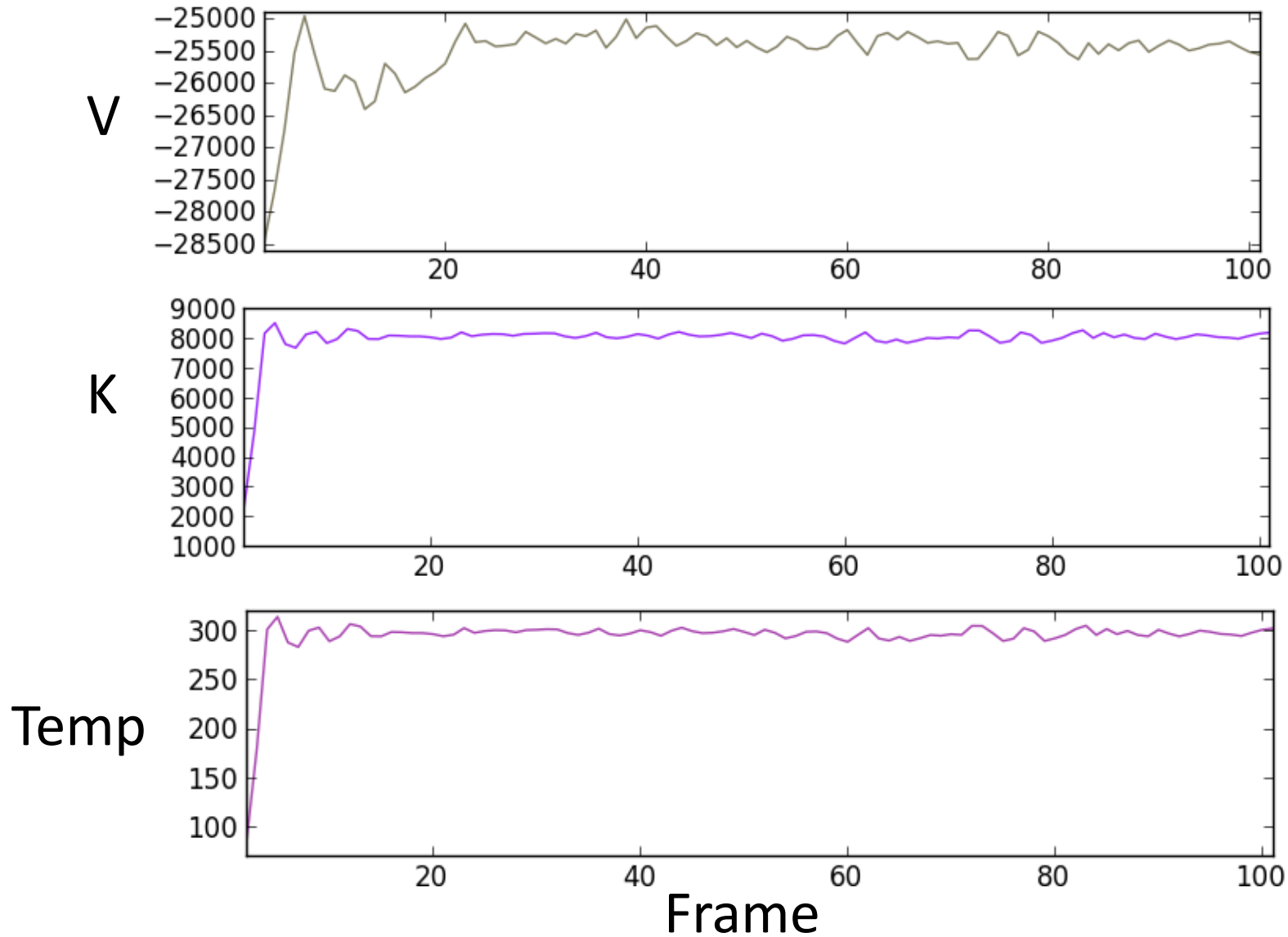
- Flutuações (variança)

$$\langle \Delta X^2 \rangle = (\langle X - \langle X \rangle \rangle)^2$$

- RMS deviation of the coordinates with respect to the initial structure (\mathbf{r}_0)

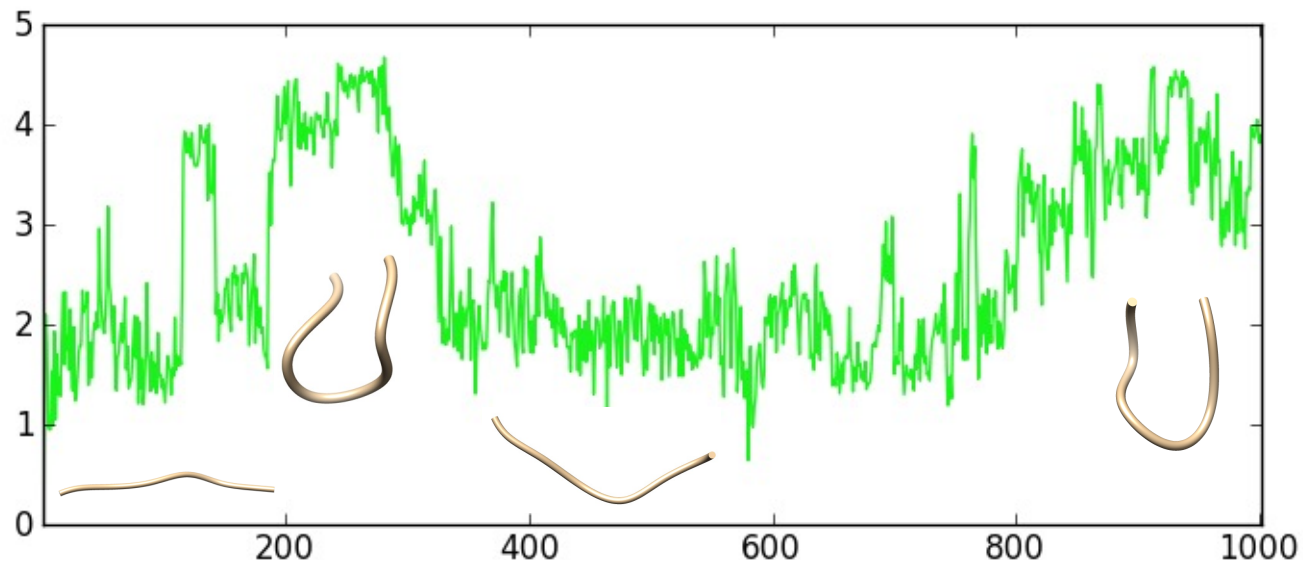
$$\sigma = \sqrt{\frac{\sum_{i=1}^N (\vec{r}_i - \vec{r}_0)^2}{N}}$$

Análise

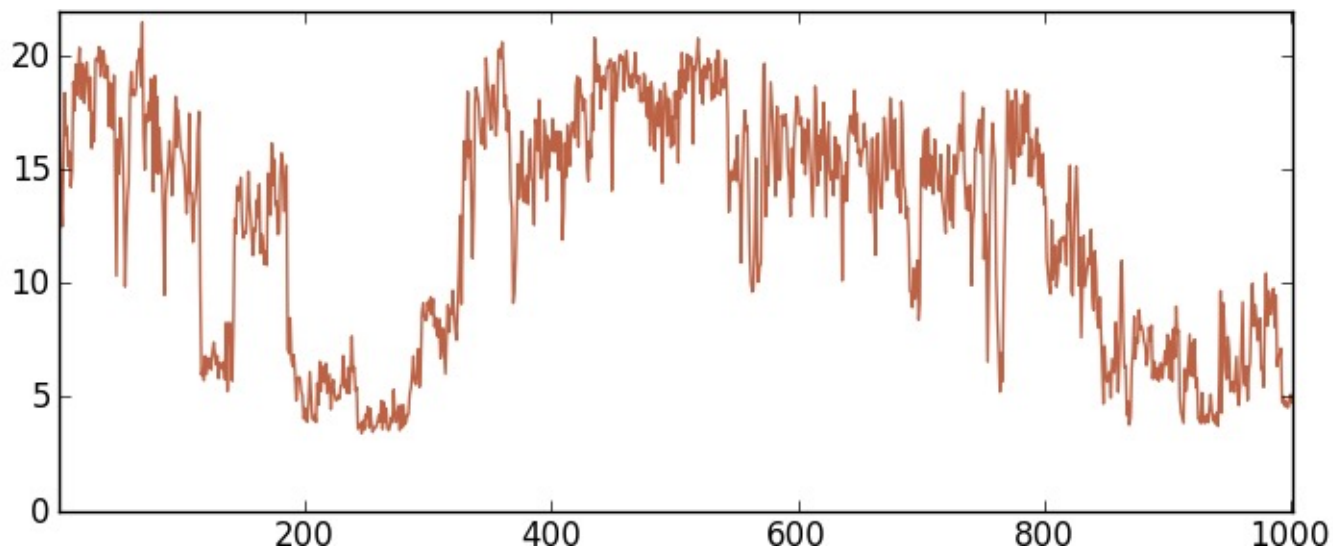


Análise

RMSD

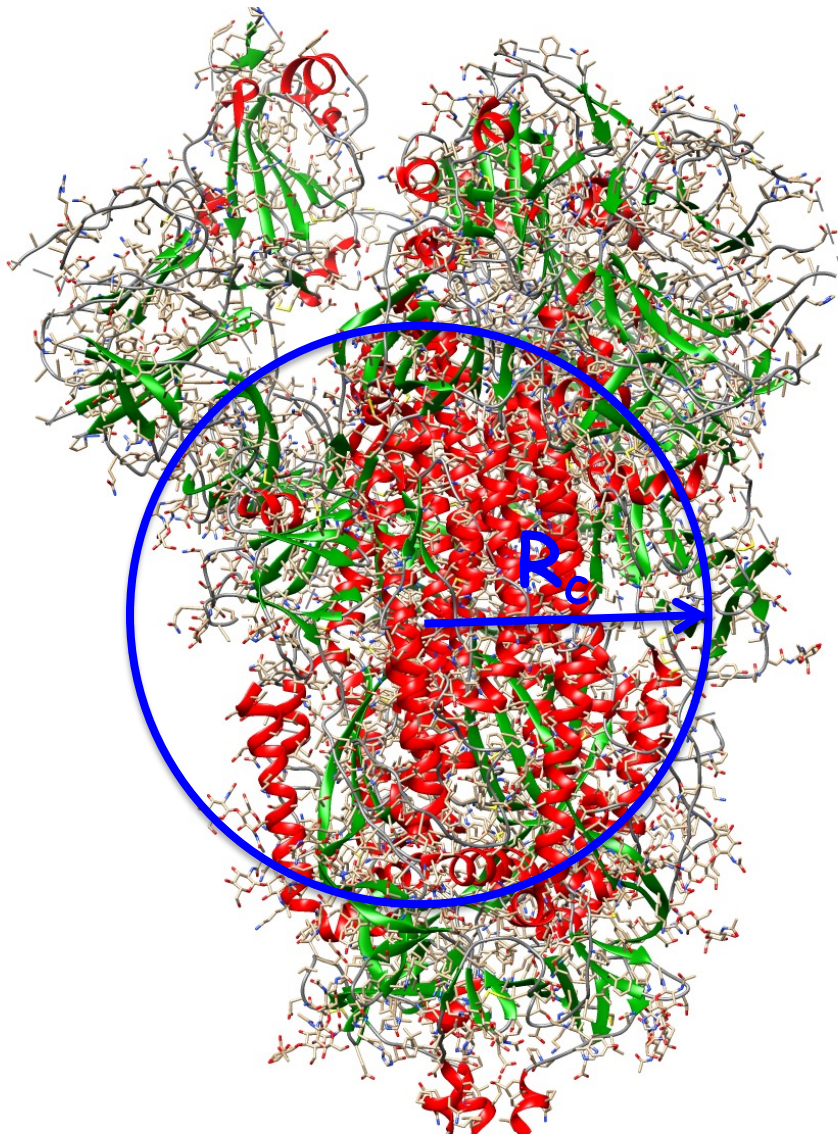


$d_{\text{end-to-end}}$ (Å)



Frame

Cálculo da energia de interação entre átomos não ligados



$$\sum_{i,j} \left[\frac{C q_i q_j}{r_{i,j}} + \epsilon \left(\frac{\alpha^{12}}{r^{12}} - \frac{\alpha^6}{r^6} \right) \right]$$

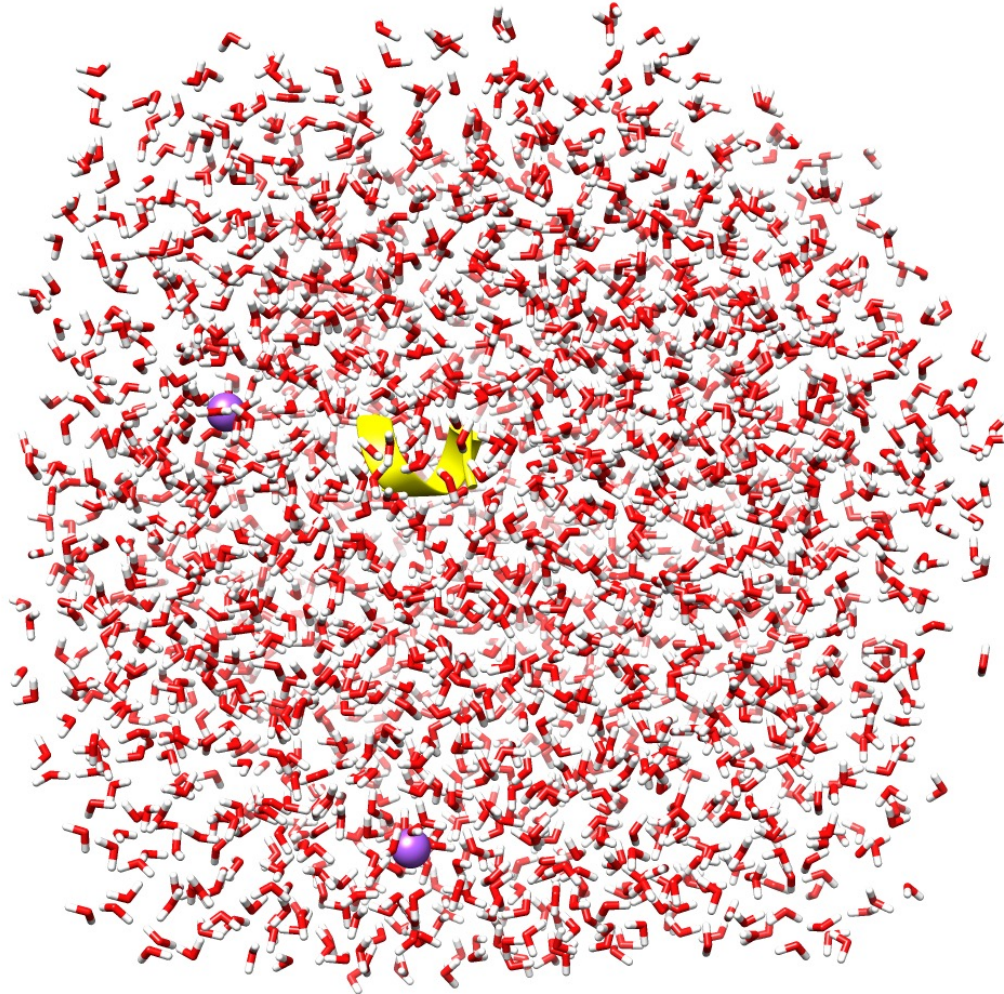
$R_c = \text{cutoff}$

$$S(r) = \begin{cases} 1 & (r < R_c) \\ 0 & (r > R_c) \end{cases}$$

$$V_{nb}' = V_{nb} S(r)$$

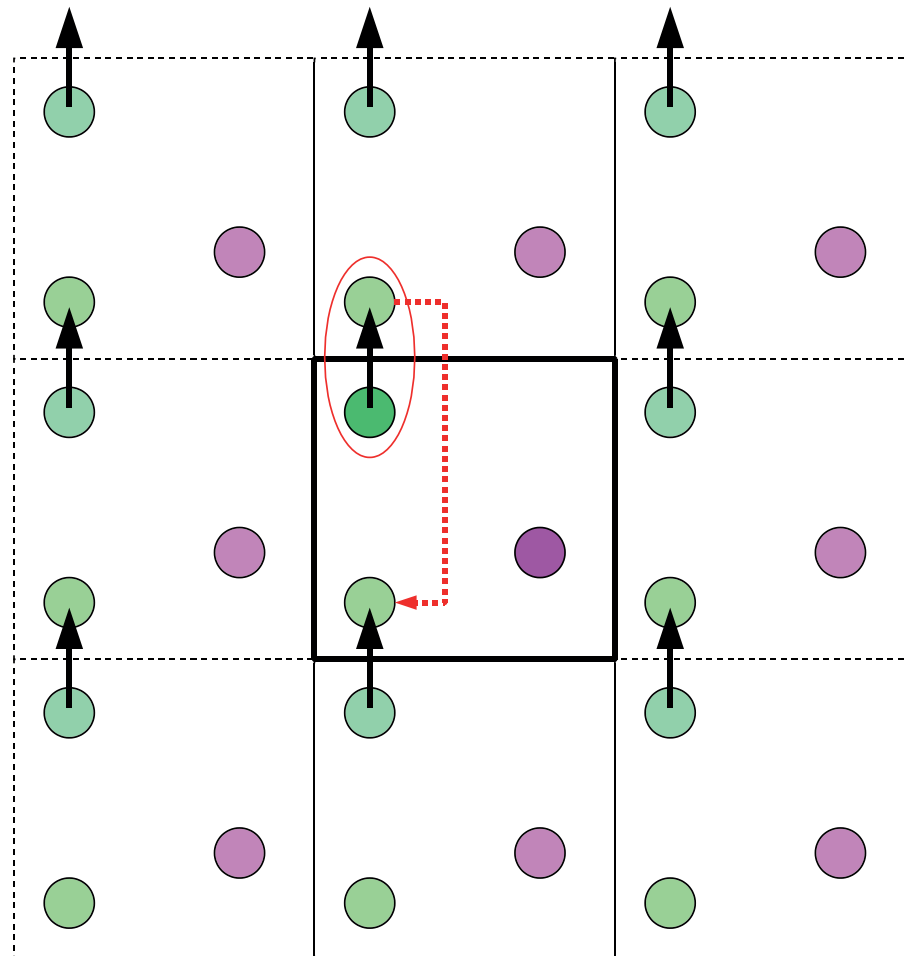
Truncation method

Periodic boundary conditions (PBC)



Periodic boundary conditions (PBC)

- Cubic, orthorhombic, hexagonal, octahedral, dodecahedral, etc
- Simetria translacional
- Comprimento da caixa



Ewald summation

- Permite calcular interações entre átomos não ligados em sistemas periódicos sem necessidade de efetuar truncagem (*cutoff*)
- Largamente utilizada em simulações de biomoléculas
- A carga elétrica líquida da caixa deve ser zero

$$S_p = \frac{1}{2} \sum_n \sum_{i,j(i \neq j)} \frac{\lambda_i \lambda_j}{|r_i - r_j + t_n|^p}$$

n é o número de caixas; *i* e *j* pares atômicos

Comparação com dados experimentais

- *ensemble* microcanônico (V,E,N)
- *ensemble* canônico (N,V,T)
- *ensemble* isobárico-isotérmico (N,P,T)

ensemble NPT

- Controle da temperatura
- Controle da pressão
- Velocidades e coordenadas são re-escaladas

$$\xi_{Temp} = \sqrt{1 + \frac{\Delta t}{\tau_T} \left(\frac{T_B}{T} - 1 \right)}$$

$$\xi_{pressão} = \sqrt{1 - \frac{\Delta\beta}{\tau_P} (P_B - P)}$$

T e P são a temperatura e a pressão em um dado instante, T_B e P_B a temperatura e pressão de referência, τ_T e τ_P são constantes de acoplamento com o banho

Bibliografia

- Martin J. Field, “A practical introduction to the simulation of molecular systems”, 2^a edição, Cambridge University Press (2007)
- Leach, A. R., Molecular Modelling: Principles and Applications, 2a ed., Pearson-Prentice Hall (2001)