

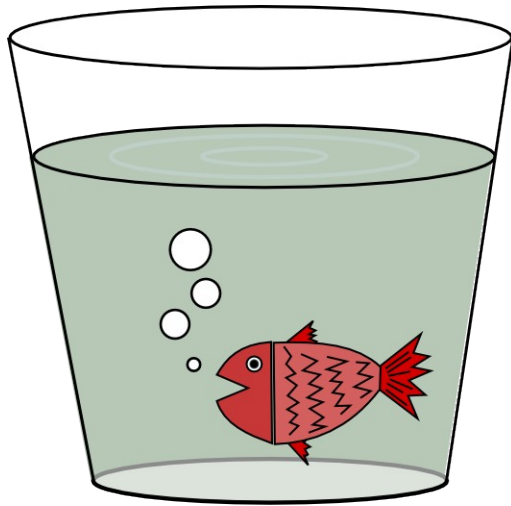
Introdução à Simulação por Dinâmica Molecular e Aplicações

Leandro Martínez
Instituto de Física de São Carlos
Universidade de São Paulo

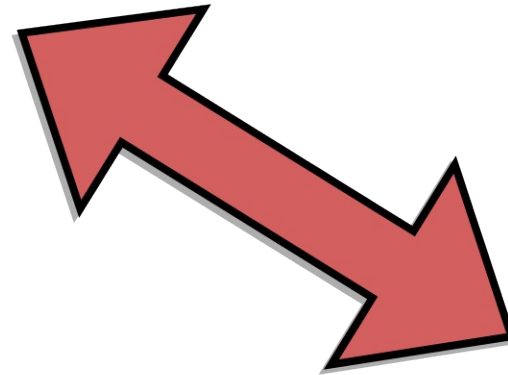
`http://limes.ifsc.usp.br`

V Escola de Modelagem Molecular em Sistema Biológicos

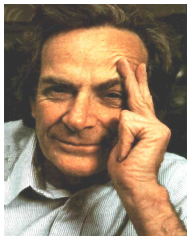
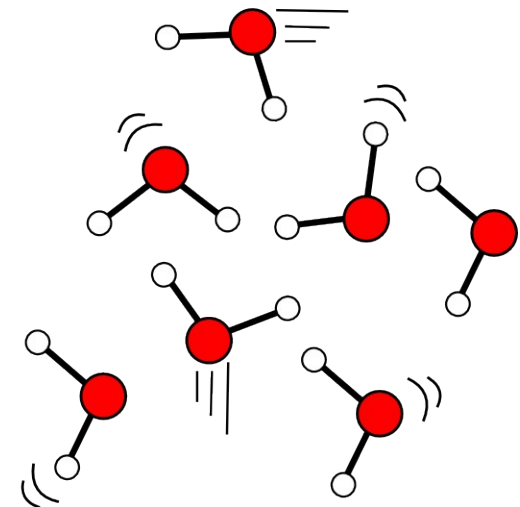
LNCC - Petrópolis, 23-27 de Agosto de 2010



« ... para nomear a
mais poderosa premissa de todas, temos que
levar adiante a ideia de que ...

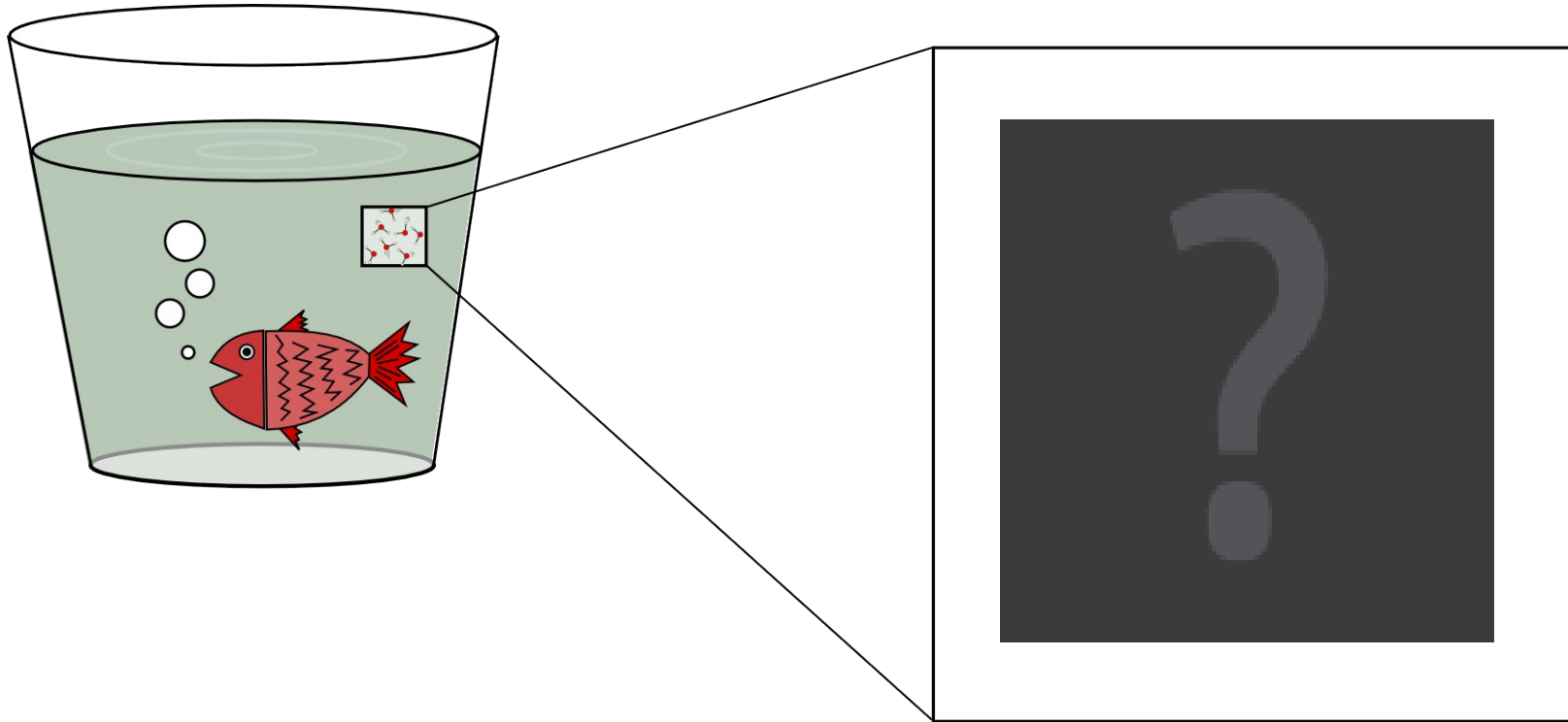


... tudo [...] pode ser compreendido a
partir do chacoalhar dos seus átomos. »



Simulações de Dinâmica Molecular:

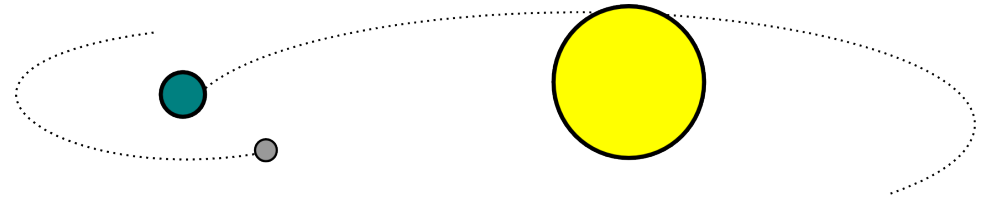
Chacoalhando os átomos.



Princípios de uma simulação

1 - Como os objetos interagem:

$$F_{i,j}(r_{i,j}) = G \frac{m_1 m_2}{r_{i,j}^2}$$



2 - Posições e velocidades iniciais: 30 de Junho de 2009, 18h 37m 54s ...

3 - Integração das equações de movimento:

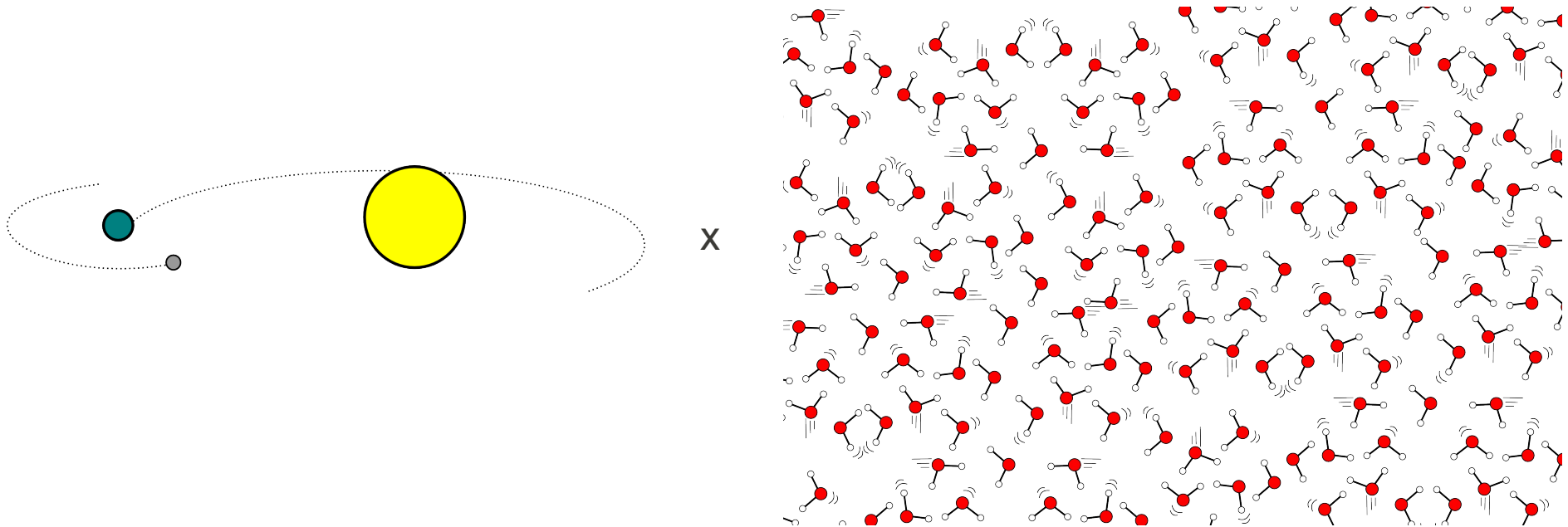
$$F(t) = f(x)$$

$$a(t) = F(t)/m$$

$$v(t+\Delta t) = v(t) + a(t)\Delta t$$

$$x(t+\Delta t) = x(t) + v(t)\Delta t + a(t)\frac{\Delta t^2}{2}$$

Princípios de uma simulação: Sistema planetários vs. Sistema molecular



Força gravitacional $\longleftrightarrow$ Interações intra- e inter-moleculares

Trajетórias precisas $\longleftrightarrow$ Propriedades médias

Energia total $\longleftrightarrow$ Temperatura e pressão

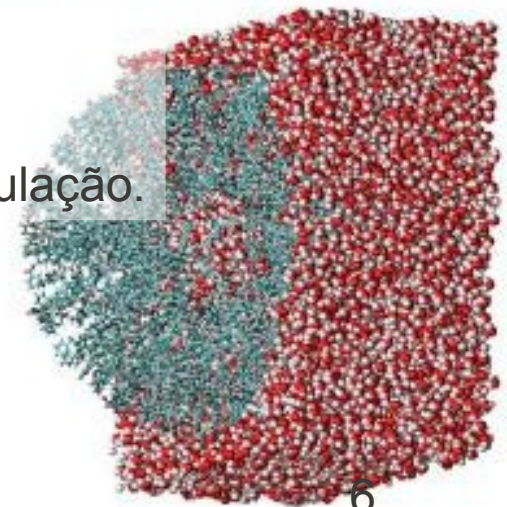
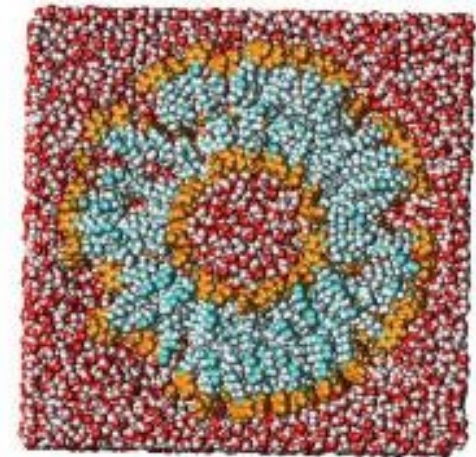
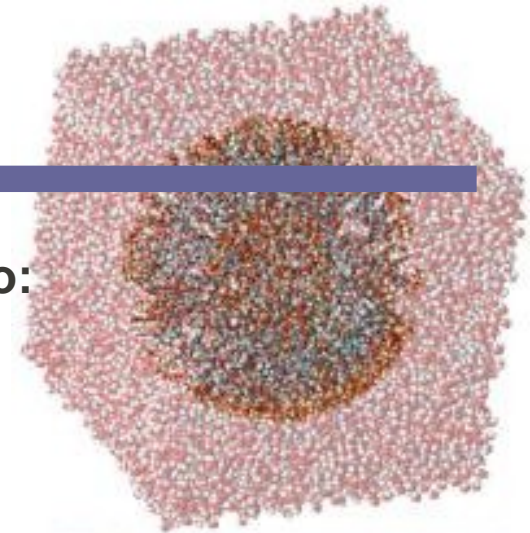
Dias, semanas, anos $\longleftrightarrow$ Nano-segundos

Para que o movimento dos átomos faça sentido é necessário:

- 1 - Representar razoavelmente bem suas interações.
- 2 - Sistema suficientemente grande para que o sistema possa ter as propriedades de um sistema macroscópico e não ter efeitos de confinamento.
- 3 - Temperatura e pressão.

Para que fazer tudo isso tenha sentido, é necessário:

- 1 - Saber o que esperar da simulação.
- 2 - Saber calcular as propriedades de interesse a partir da simulação.



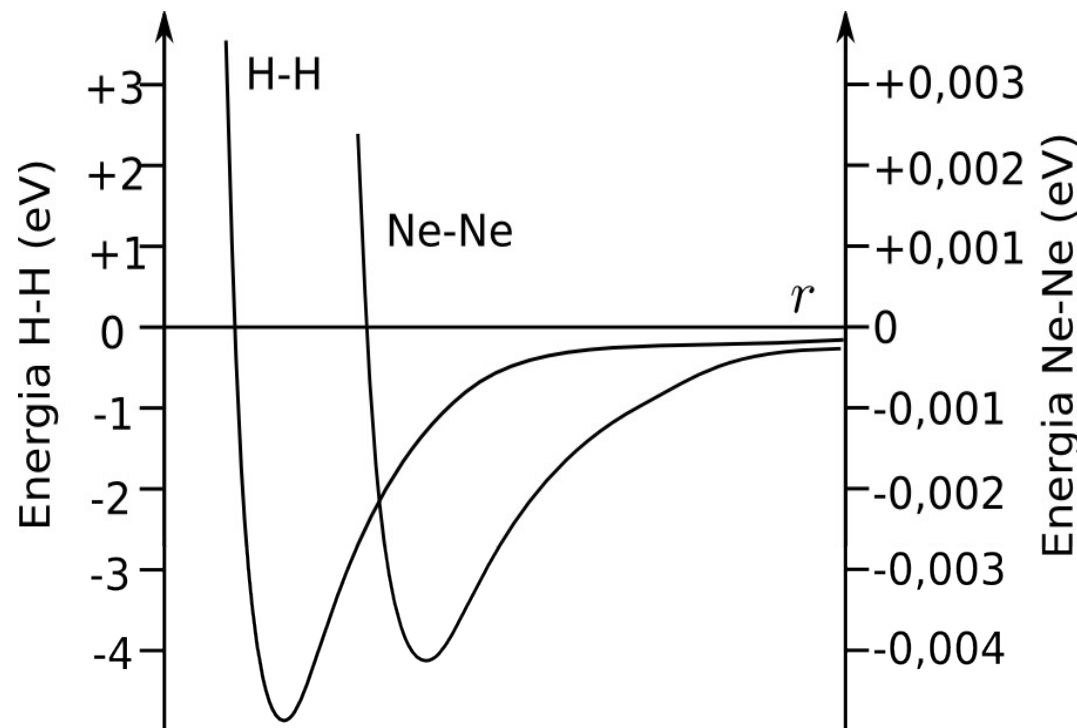
Interações

Quântico vs. Clássico

O mundo é quântico...

$$-\frac{\hbar^2}{2m}\nabla^2\Phi(\vec{x}, t) + V(\vec{x}, t)\Phi(\vec{x}, t) = i\frac{\partial}{\partial t}\Phi(\vec{x}, t)$$

... mas não tanto assim ...



Boa aproximação para estas curvas:

$$V(r) = D_e \left(1 - e^{-a(r-r_{eq})}\right)^2$$

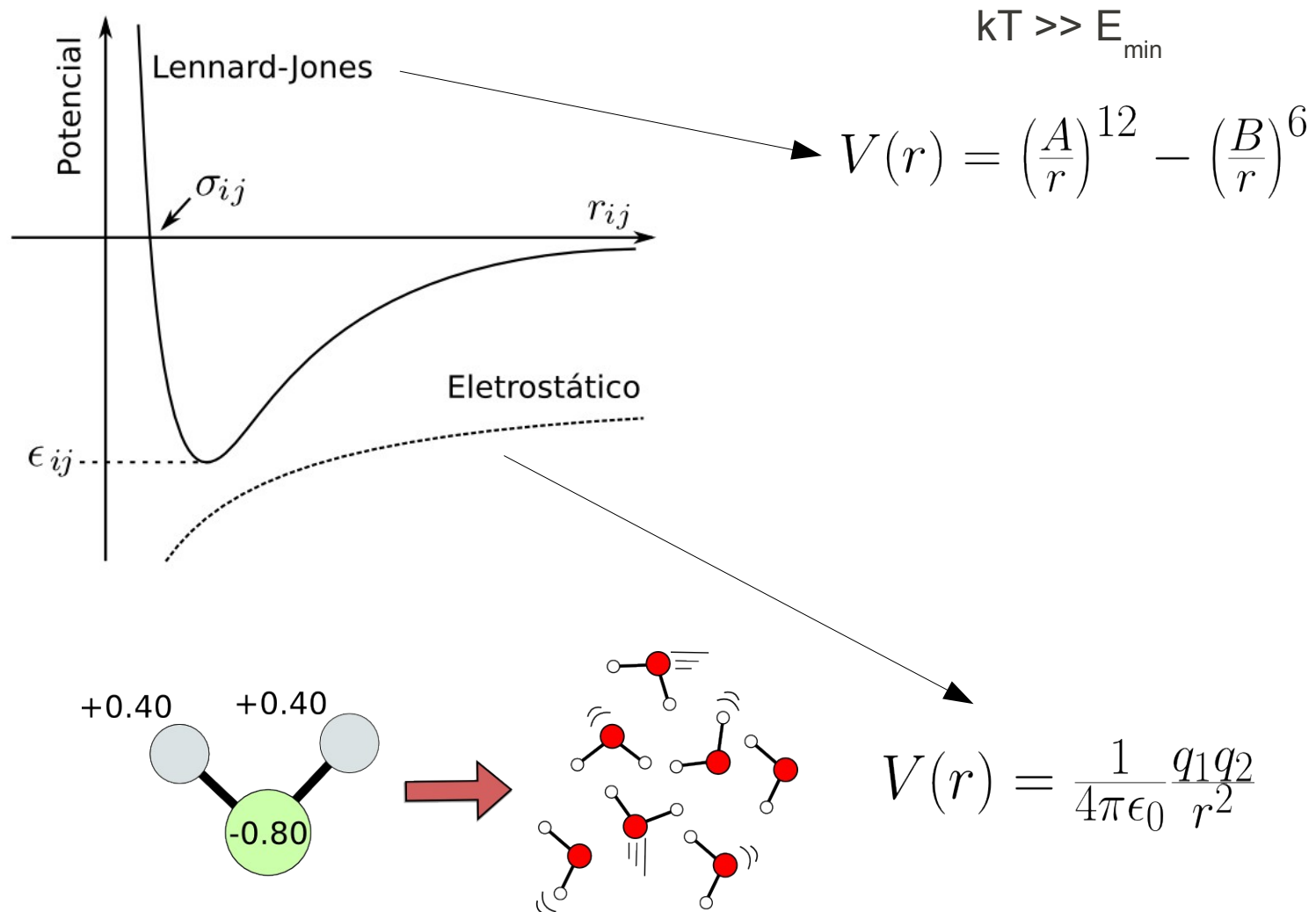
(potencial de Morse)

$$kT \sim 0,026 \text{ eV}$$

(298 K)

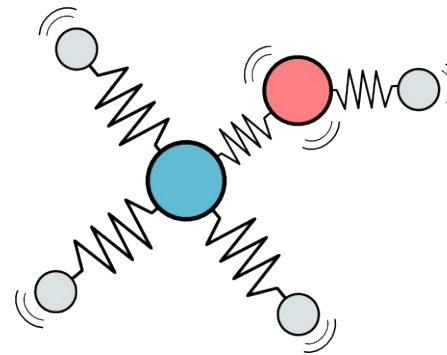
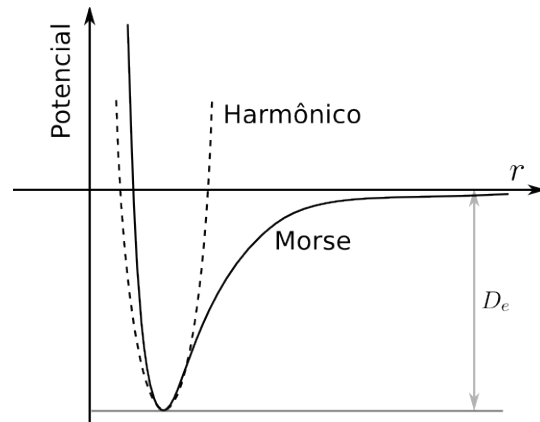
Interações inter-moleculares

Ou entre grupos distantes de uma macromolécula

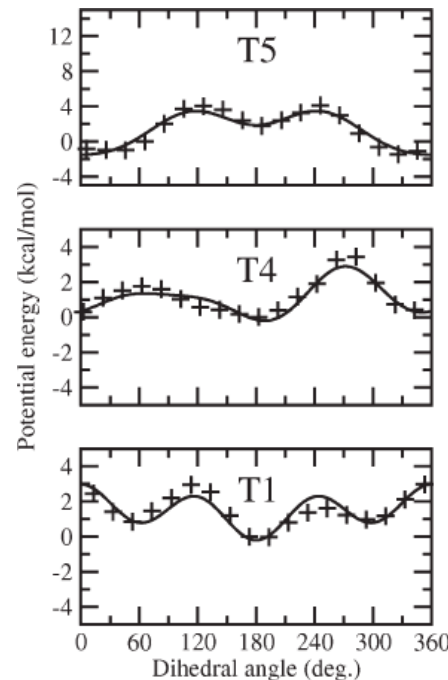
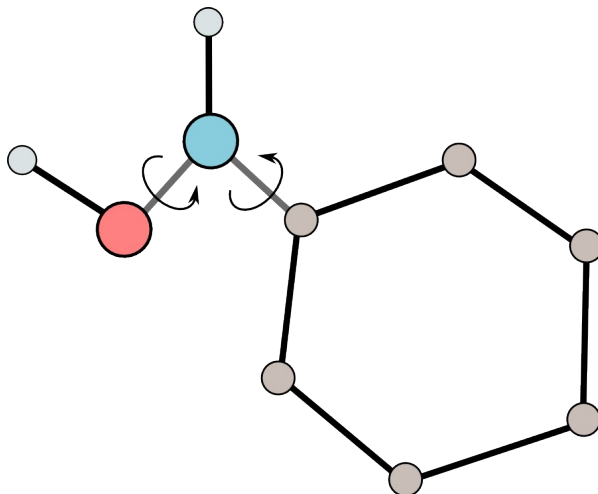
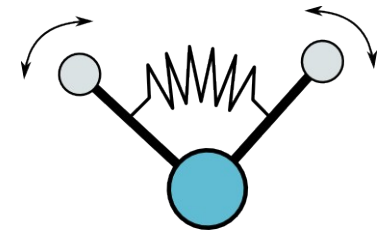


Interações intra-moleculares

Potenciais não dissociativos

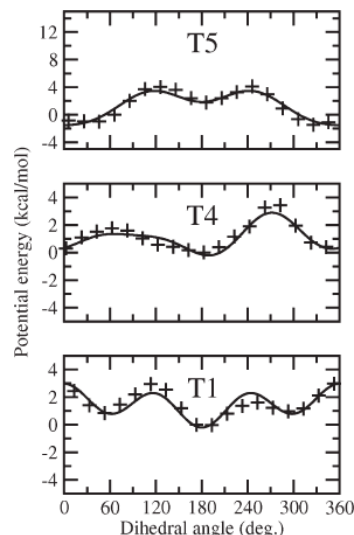
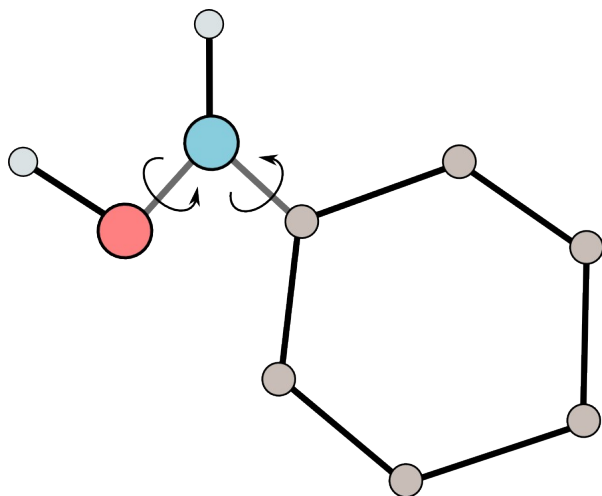


$$V(r) = k(r - r_{eq})^2$$



Somas
de senos/cossenos
que ajustam o potencial
quântico

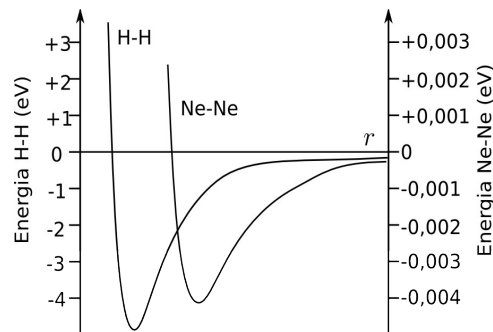
Ajuste dos parâmetros contra cálculos ab-initio



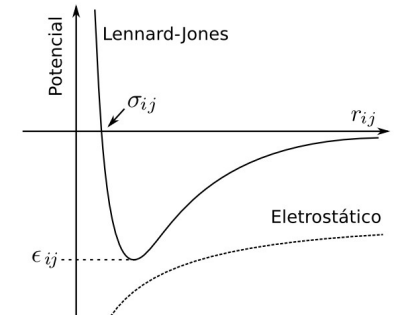
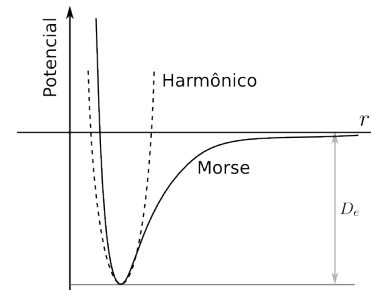
→ Energia ab-initio
inclui todas as interações
(ângulos, diedros, cargas,
vdW, etc)

$$\begin{aligned}
 V = & \sum_{\text{bonds}} K_b (r - r_{0,b})^2 + \sum_{\text{UB}} K_{\text{UB}} (S - S_{0,\text{UB}})^2 \\
 & + \sum_{\text{angle}} K_a (\theta - \theta_{0,a})^2 + \sum_{\text{dihedrals}} K_{d,n} (1 + \cos(n\chi - \delta_{d,n})) \\
 & + \sum_{\text{impropers}} K_i (\psi - \psi_{0,i})^2 \\
 & + \sum_{\text{nonbonded}} \epsilon_{ij} \left[\left(\frac{R_{\text{min},ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{\text{min},ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}},
 \end{aligned}$$

Esquema geral de uma simulação de dinâmica molecular



Cálculos ab-initio (quânticos) para poucos átomos
(ou dados experimentais)



$$F(t) = f(x)$$

$$a(t) = F(t)/m$$

$$v(t+\Delta t) = v(t) + a(t)\Delta t$$

$$x(t+\Delta t) = x(t) + v(t)\Delta t + a(t)\frac{\Delta t^2}{2}$$

Ajuste dos potenciais quânticos
com funções simples analíticas

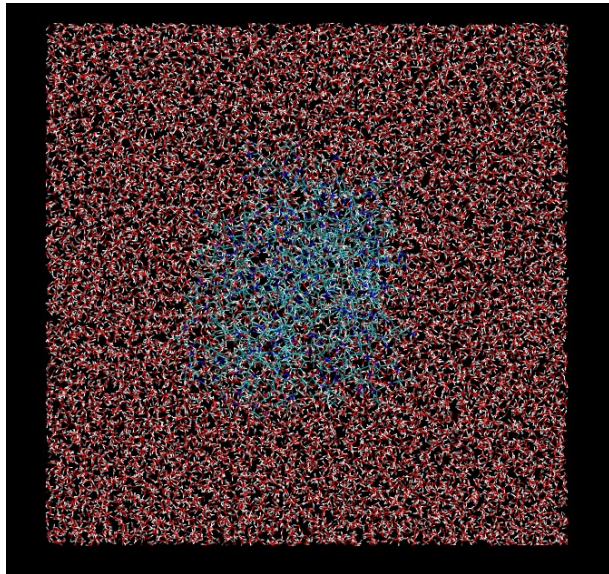
Resolução das equações
de movimento newtonianas

Sistema simulado

Sistema suficientemente grande



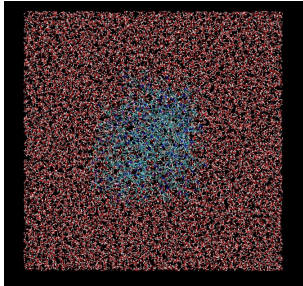
Propriedades estruturais e dinâmicas da água: 300 águas (900 átomos)



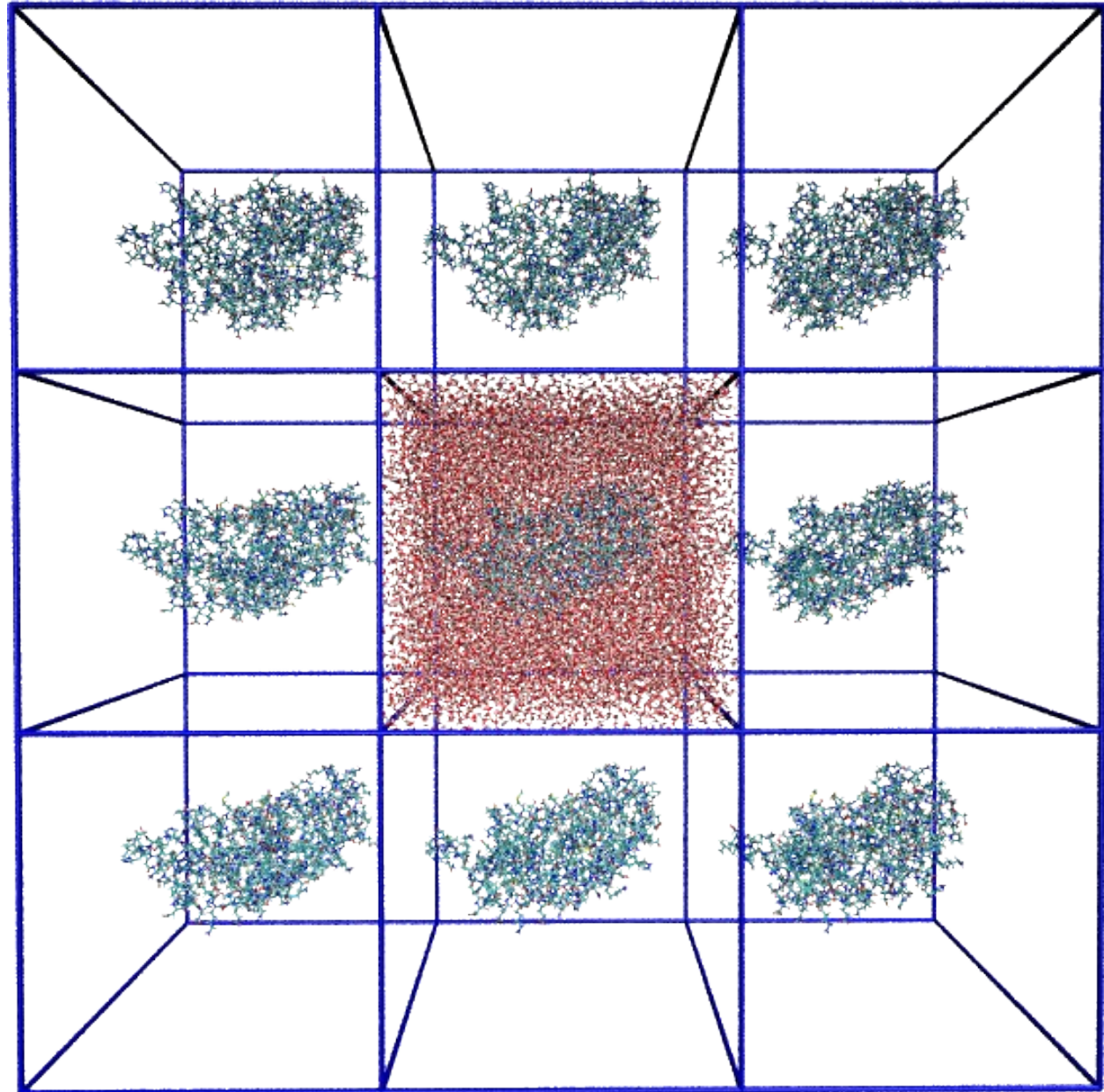
Propriedades de uma proteína em água:
1 proteína (5000 átomos) + 20 mil águas
(60 000 átomos).

Sistema

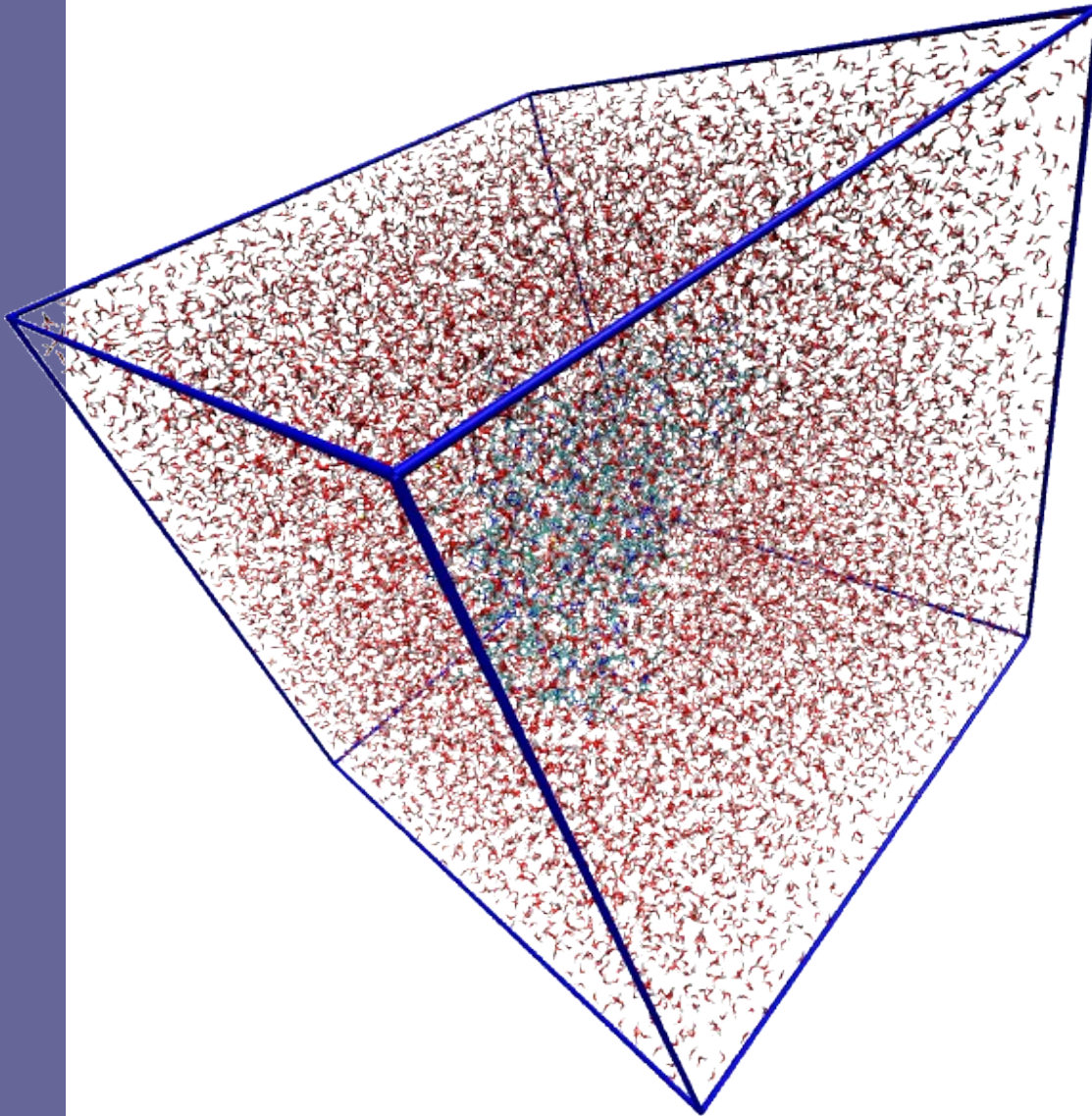
Sistema suficientemente grande



Condições
periódicas
de contorno



Temperatura e pressão



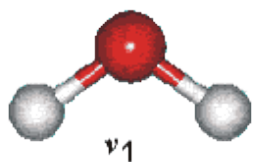
Controle sobre as velocidades:

$$T = \frac{2}{3} \frac{1}{k} \left\langle \frac{1}{2} m v^2 \right\rangle$$

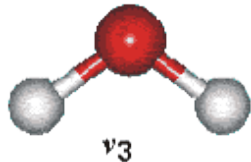
Pressão: Densidade da caixa.

Expectativas

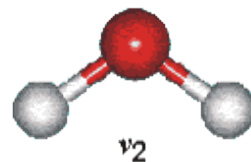
O que esperar da simulação



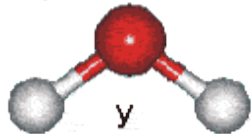
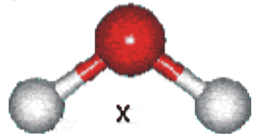
symmetric stretch



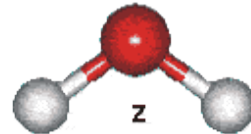
asymmetric stretch



bend

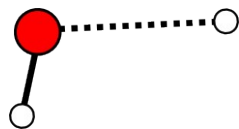


librations

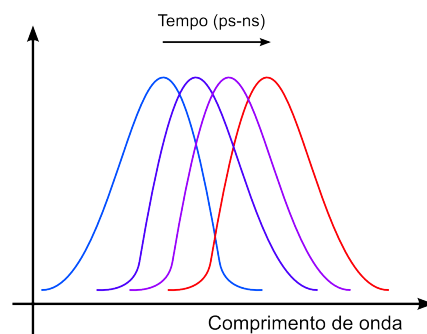


Estiramentos / torções: 100%!!
100% parametrização...

Rotação, libração, difusão: propriedades coletivas do líquido, começa a ser interessante.



Quebra de ligação?



Espectroscopia
resolvida no tempo

Enovelamento (muito tempo)



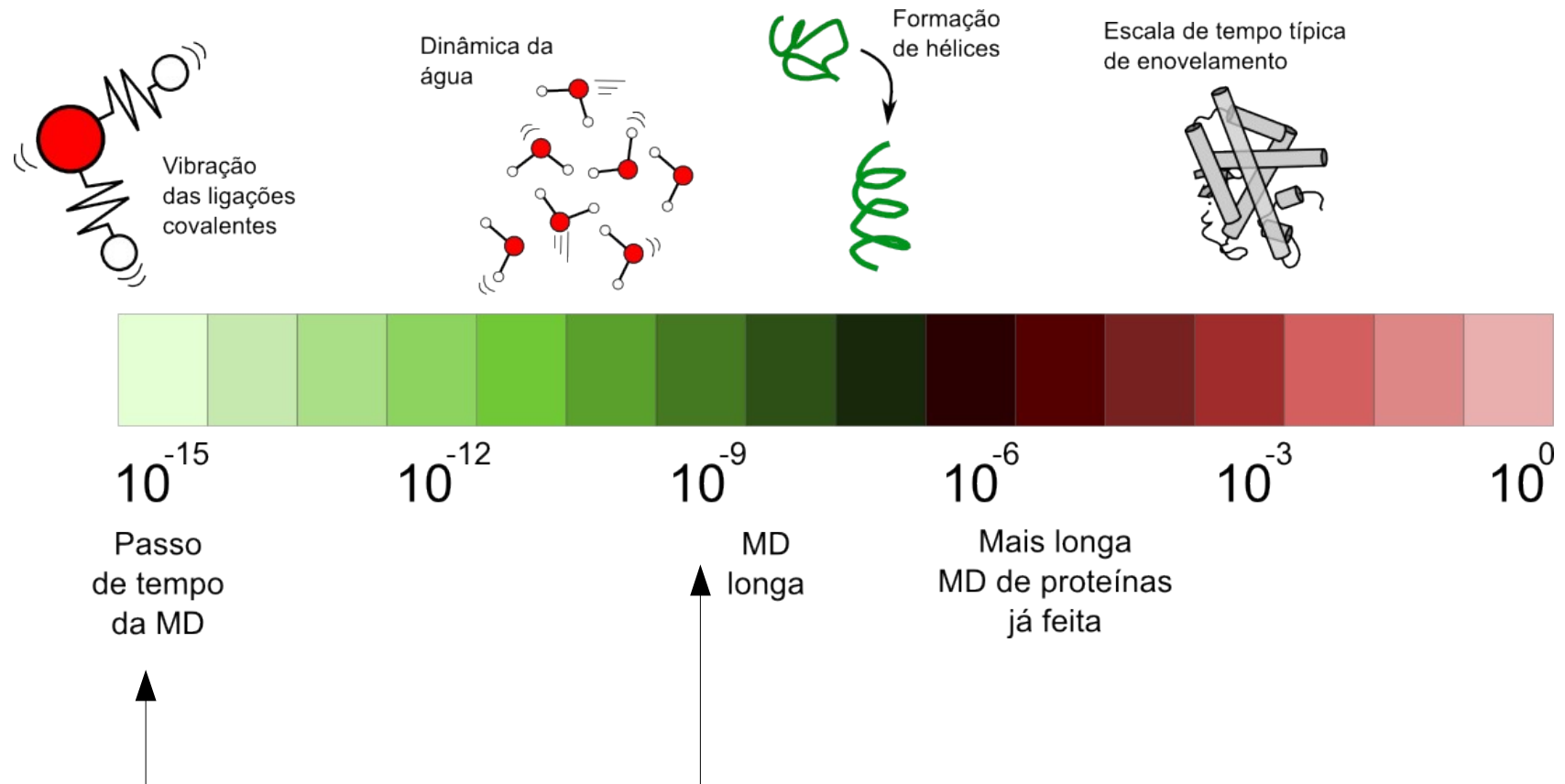
Saber calcular as propriedades de interesse a partir da simulação.



?



Expectativas



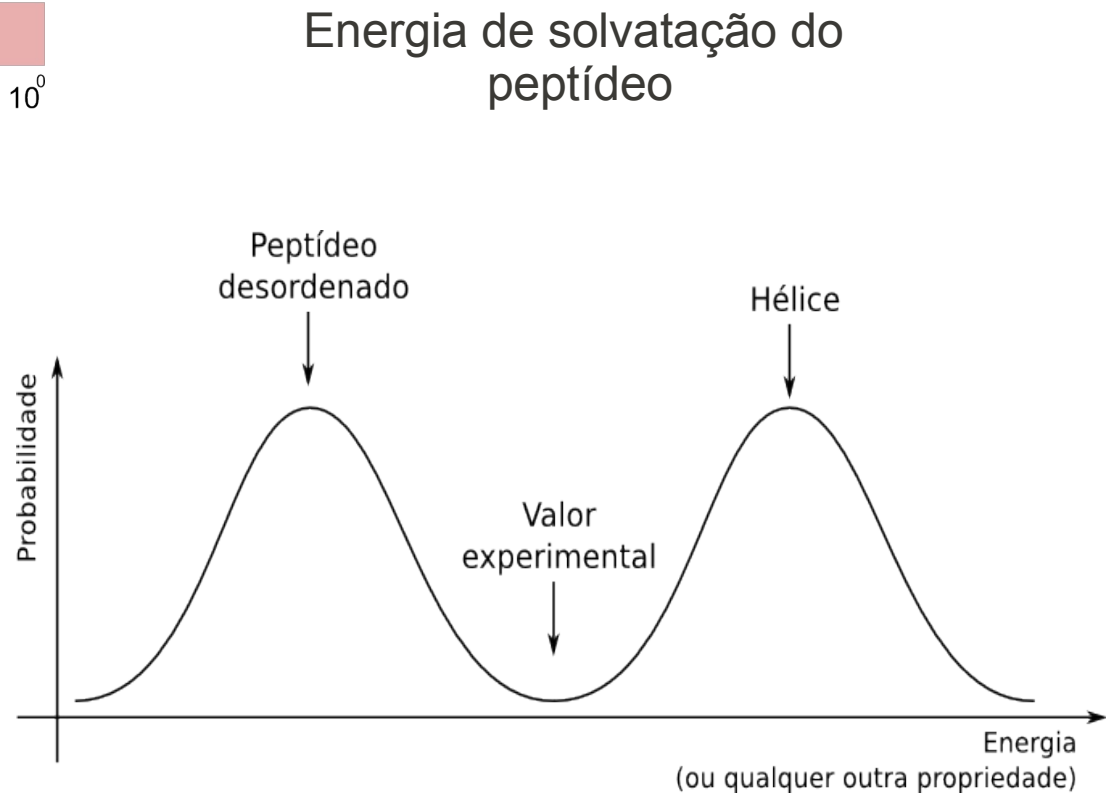
Para simular 1 ns são necessários da ordem 1 000 000 de ciclos.

Com ~ 100 000 átomos (uma proteína + 30 mil águas):

1 mês em um único processador moderno.

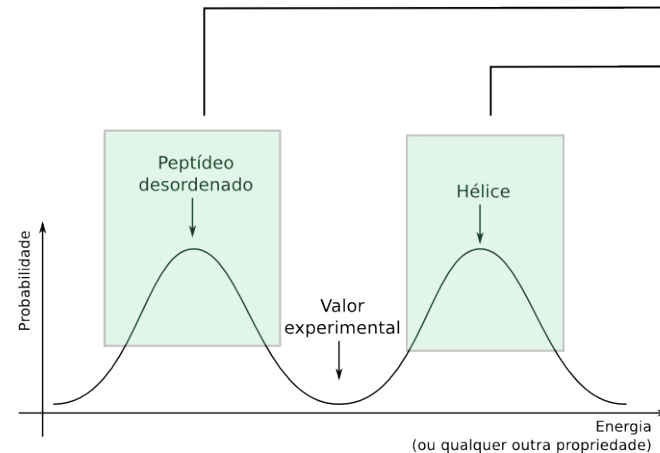
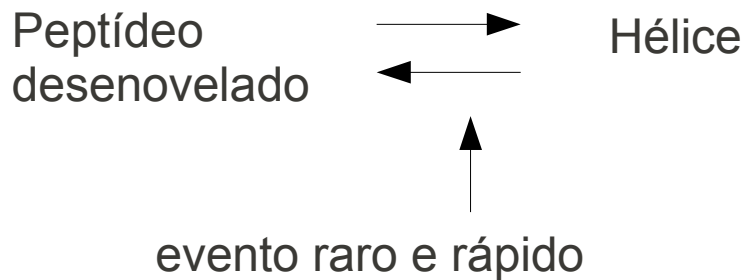
Amostragem

ística, estadística, estadística, estadística, estadística, estadística, estadística, estadís



Amostragem

ística, estatística, estatística, estatística, estatística, estatística, estatística, estatística



The Effect of Motional Averaging on the Calculation of NMR-Derived Structural Properties

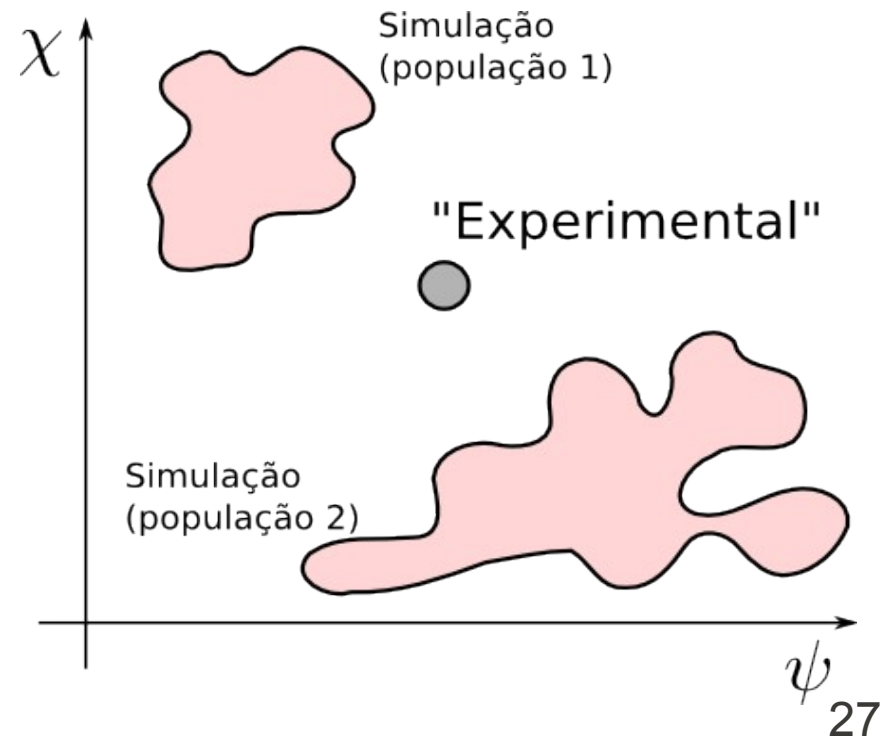
Xavier Daura,¹ Iris Antes,^{1,2} Wilfred F. van Gunsteren,¹ Walter Thiel,² and Alan E. Mark^{1*}

¹Laboratory for Physical Chemistry, Swiss Federal Institute of Technology (ETH), Zürich, Switzerland

²Institute of Organic Chemistry, University of Zürich, Zürich, Switzerland

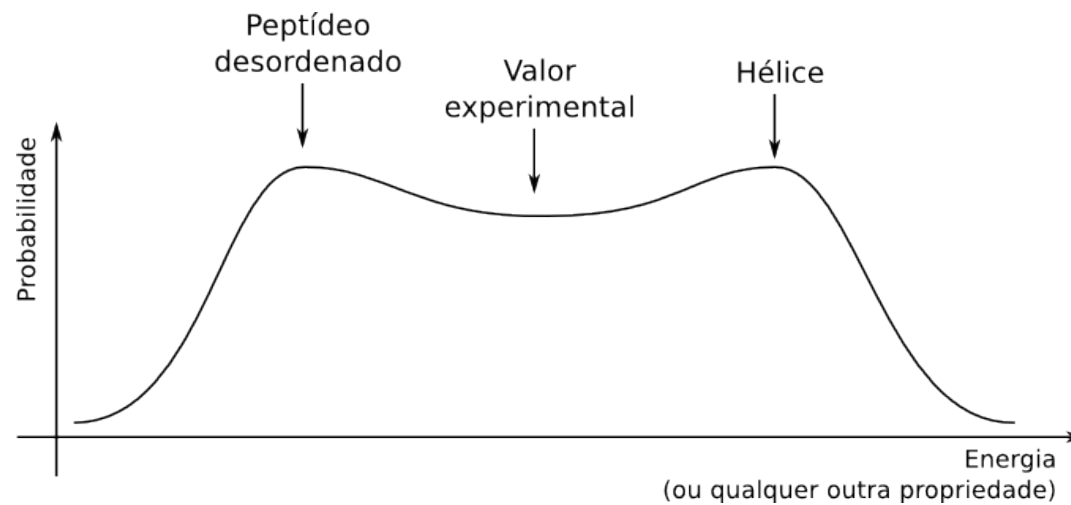
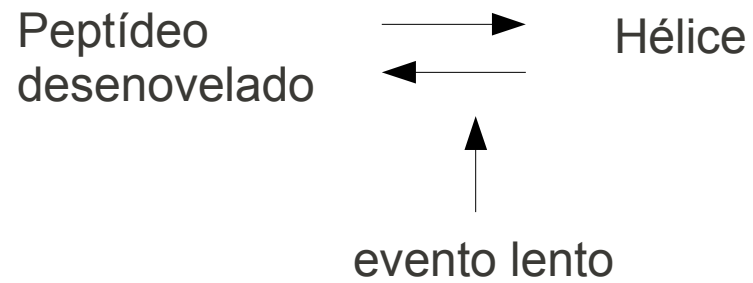
ABSTRACT The effect of motional averaging when relating structural properties inferred from nuclear magnetic resonance (NMR) experiments to molecular dynamics simulations of peptides is considered. In particular, the effect of changing populations of conformations, the extent of sampling, and the sampling frequency on the estimation of nuclear Overhauser effect (NOE) inter-proton distances, vicinal

This is further complicated by the fact that the NMR effect itself is weak. That is, even at high field there is only a slight excess of aligned spins. Experiments depend on the combined signal of a large collection of individual molecules accumulated on a millisecond time scale. In short, NMR data represents a complex time and ensemble average of a non-linear property. Based on the NMR signal alone it is not possible to solve for the spatial arrange-



Amostragem

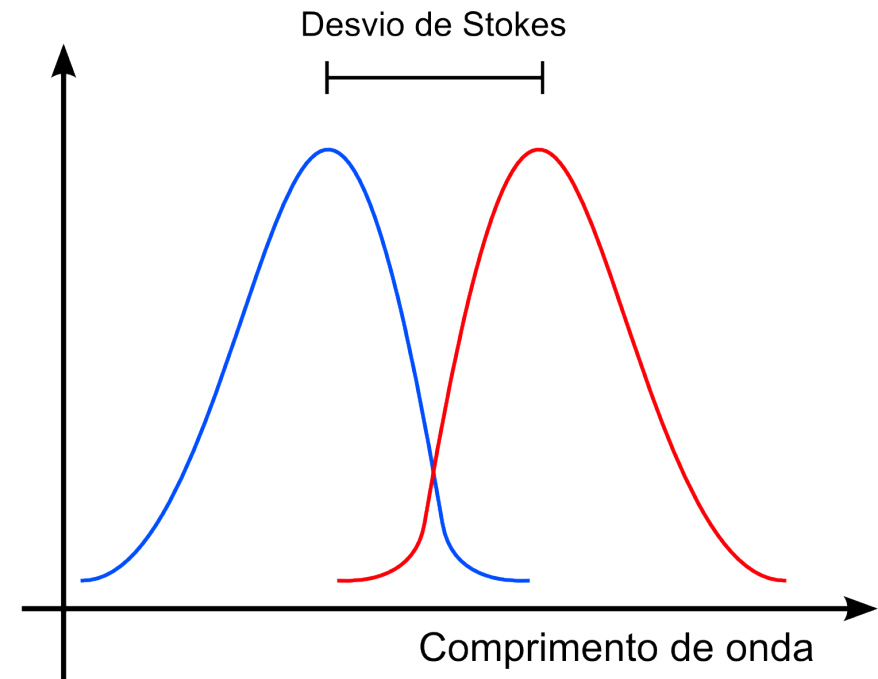
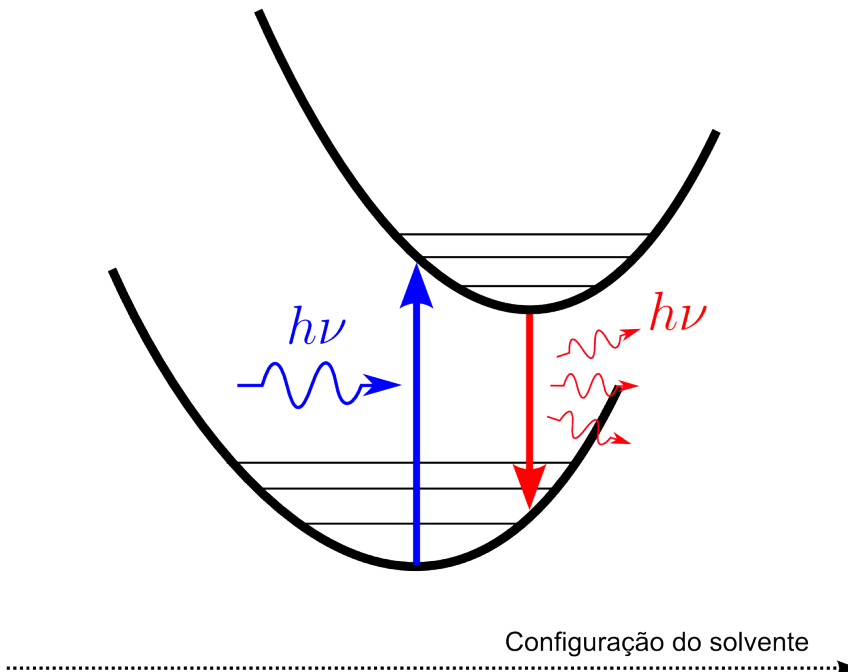
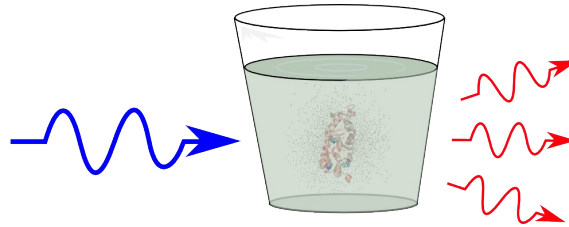
ística, estadística, estadística, estadística, estadística, estadística, estadística, estadís



Simulação e Experimento

Quando é possível ter estatística, comparação direta com experimentos:

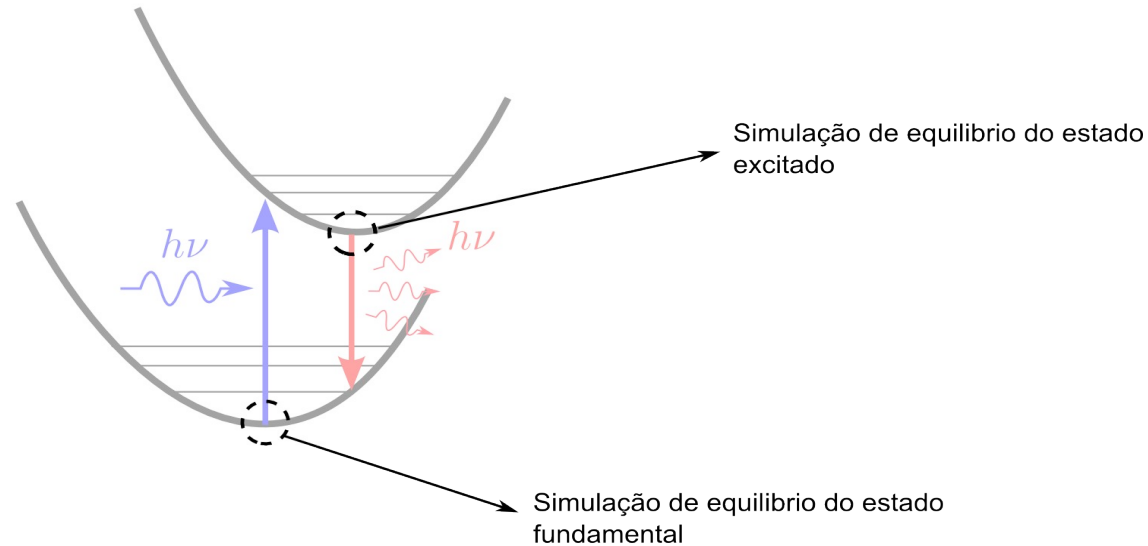
Exemplo: Desvio de Stokes



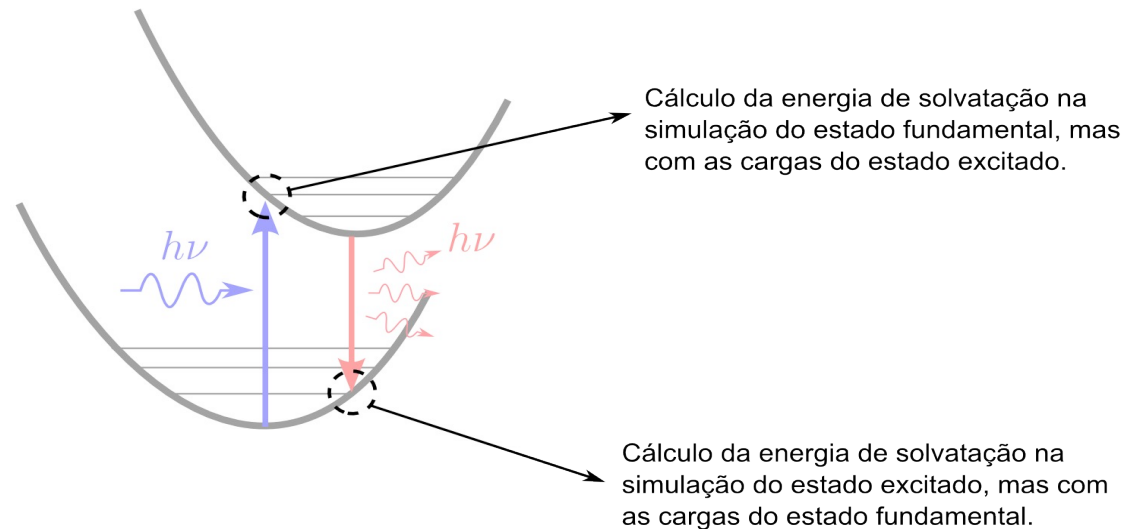
Quando é possível ter estatística, comparação direta com experimentos:

Exemplo: Desvio de Stokes

Solvente relaxado



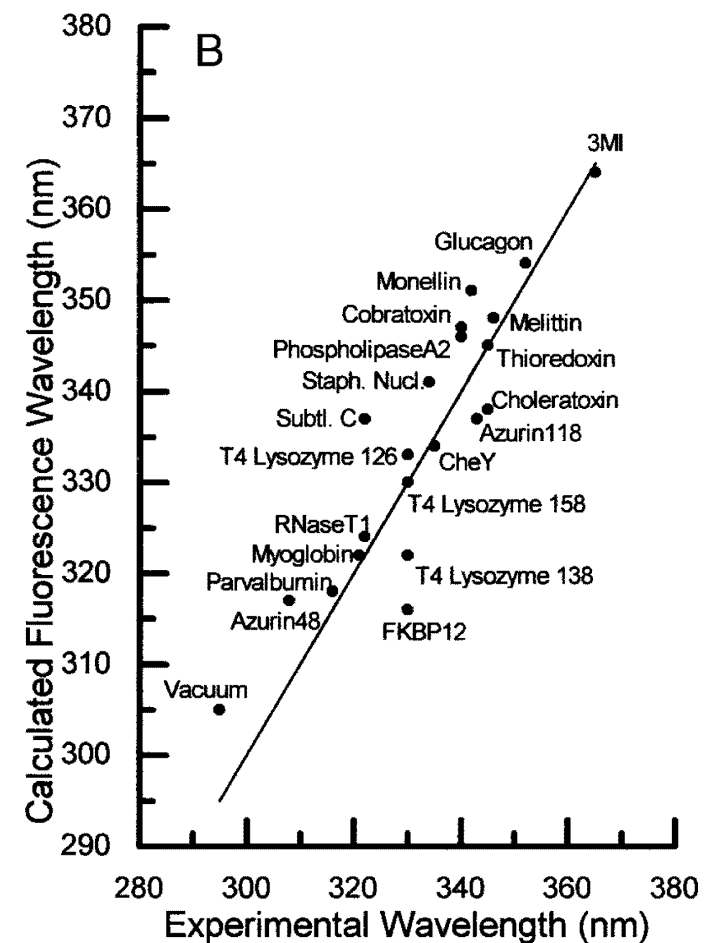
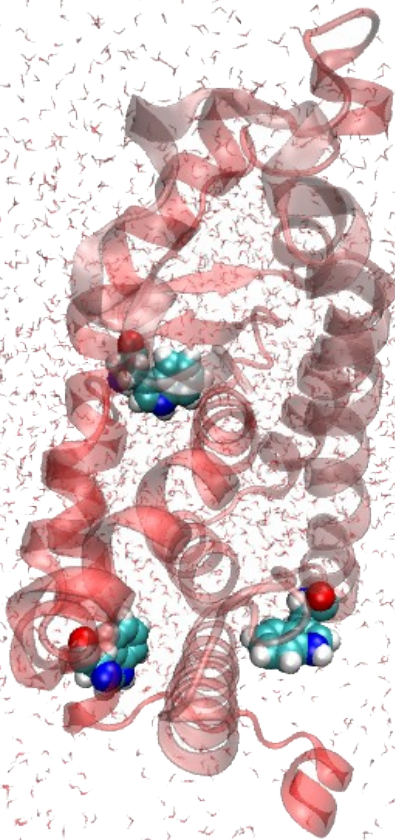
Solvente não-relaxado



Simulação e experimento

ística, estadística, estadística, estadística, estadística, estadística, estadística, estadís

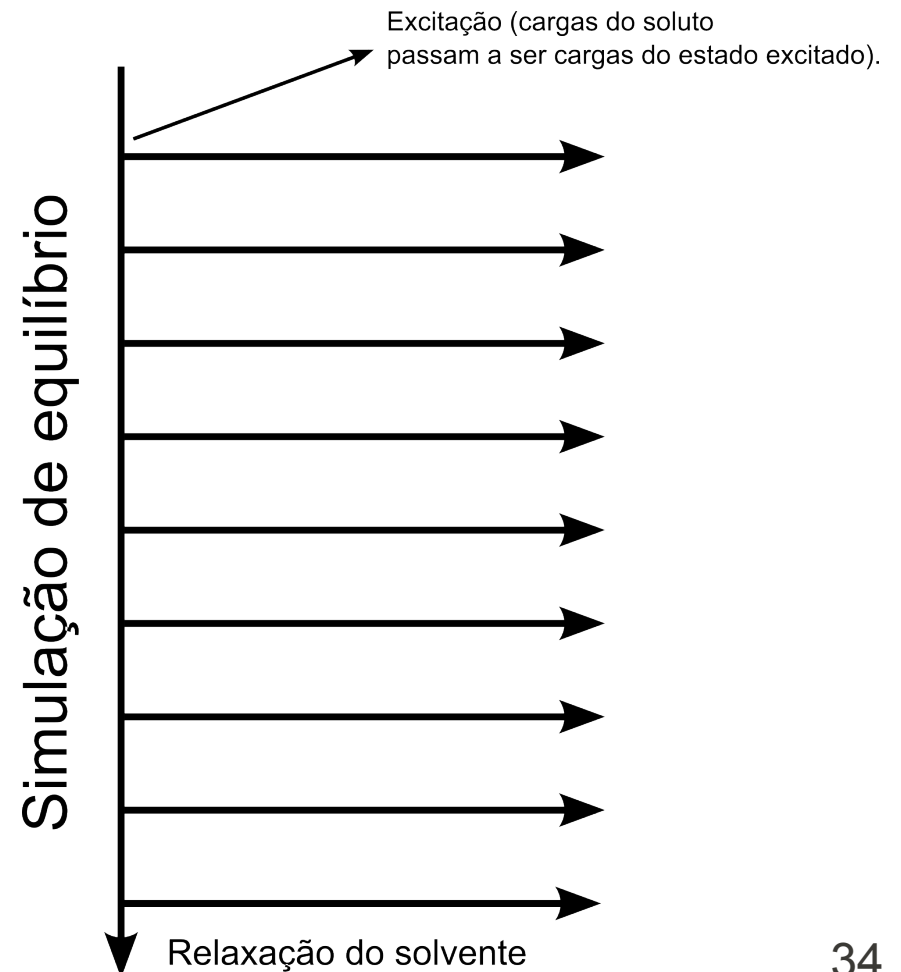
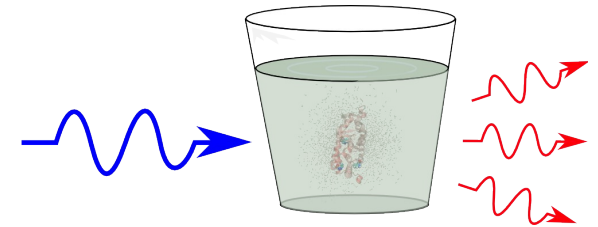
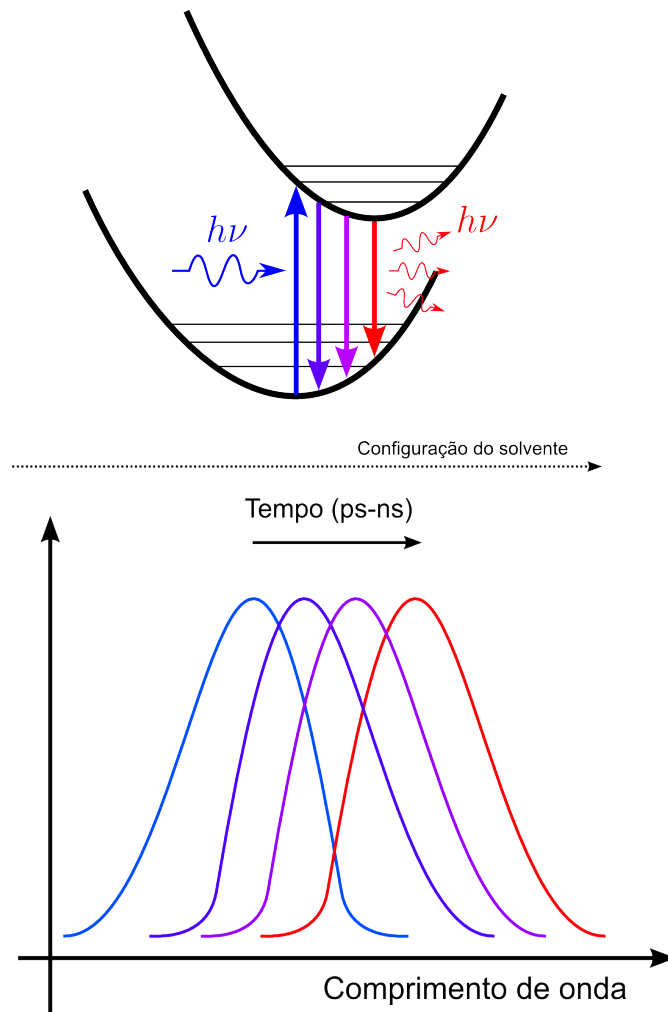
Quando é possível ter estatística, comparação direta com experimentos:



Simulação e experimento

Quando é possível ter estatística, comparação direta com experimentos:

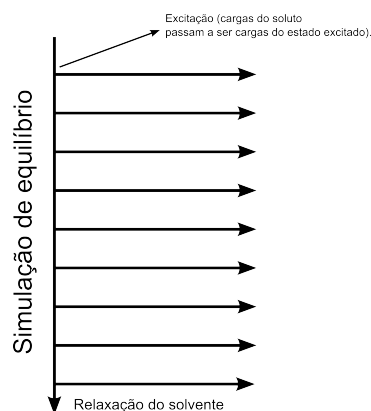
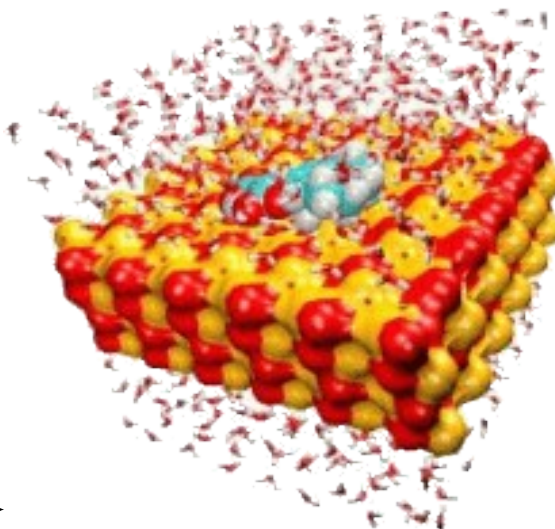
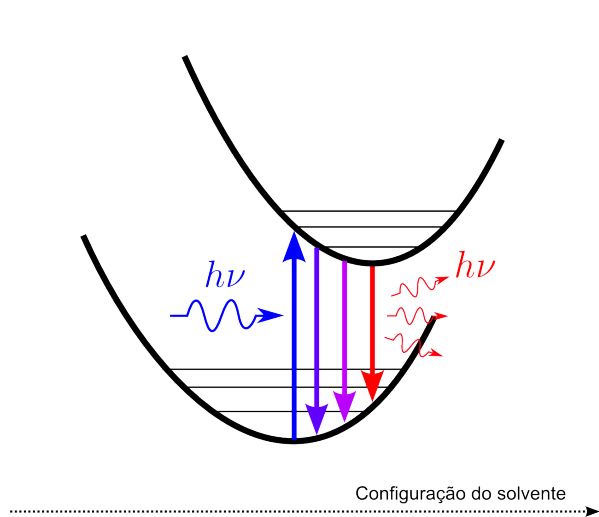
Exemplo: Desvio de Stokes resolvido no tempo



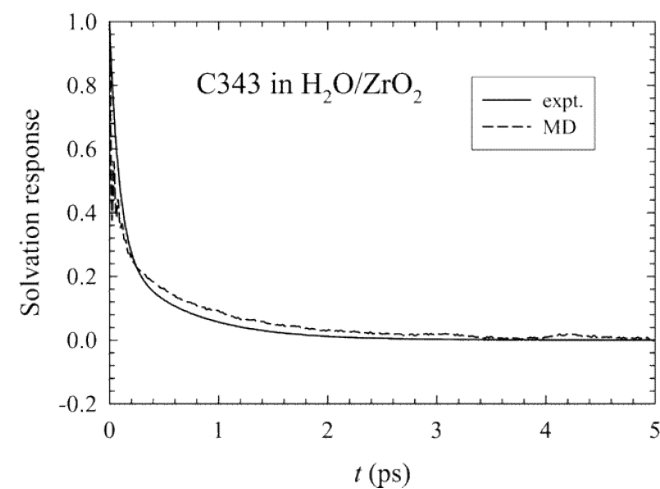
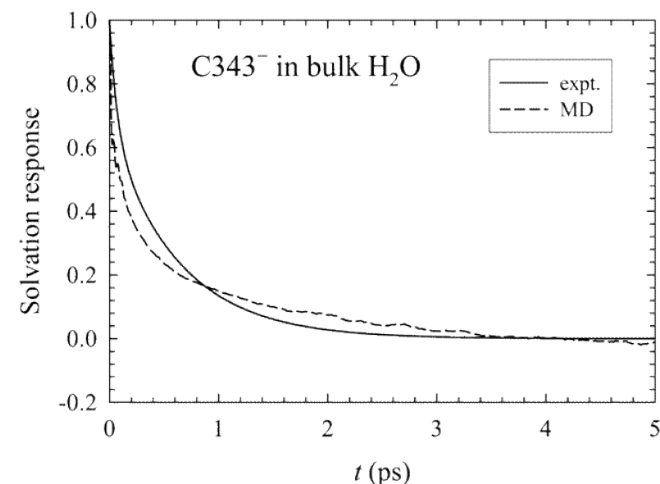
Simulação e experimento

Quando é possível ter estatística, comparação direta com experimentos:

Exemplo: Desvio de Stokes resolvido no tempo.



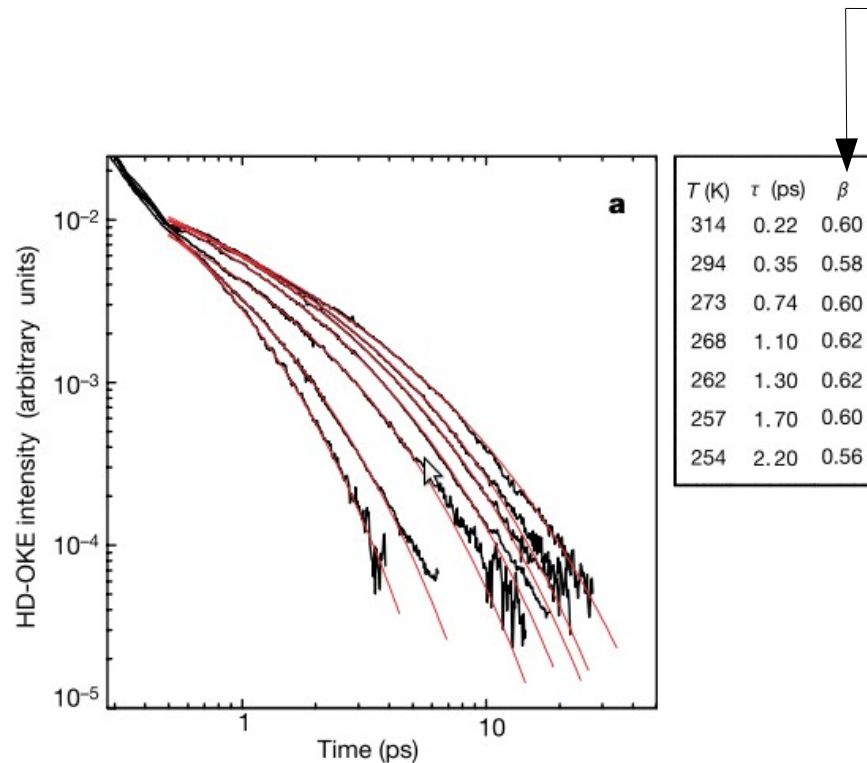
Diferentes estados de protonação tem diferentes espectros.



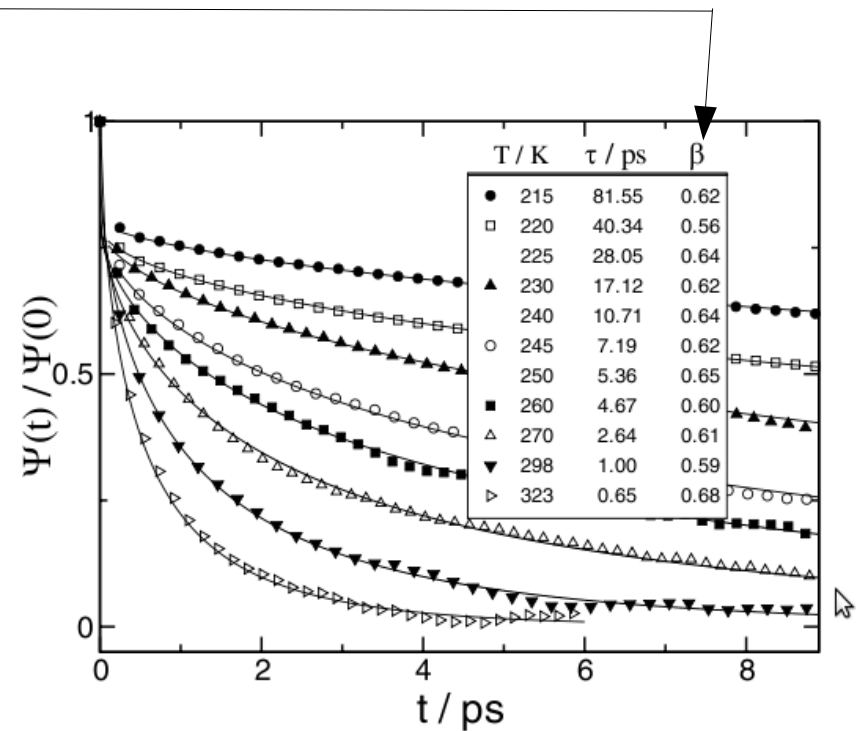
Simulação e experimento

atística, estadística, estadística, estadística, estadística, estadística, estadística, estat

Quando é possível ter estatística, comparação direta com experimentos:



Structural relaxation in supercooled water.
Righini e col. Nature, 428, 18, 2004.



Optical kerr effect in supercooled water.
Skaf e Sonoda, Phys. Rev. Lett. 2005.

Efeito Kerr: Variação do índice de refração pela ação de um campo elétrico. Associado à taxa de relaxação do dipolo coletivo do líquido (medidas acima).

ística, estadística, estadística, estadística, estadística, estadística, estadística, estadís

NA VIDA REAL

Sistemas pequenos e rápidos

- Toda a estatística que quiser
- Propriedades são calculadas e comparadas com valores experimentais
- A simulação é um “sistema real” em que qualquer experimento pode ser feito.
- “Olhar” para a simulação tem papel secundário.

Sistemas grandes e/ou lentos

- Estatística muito limitada, logo:
 - Raramente uma propriedade pode ser calculada com confiabilidade.
 - Eventos interessantes ocorrem raramente.

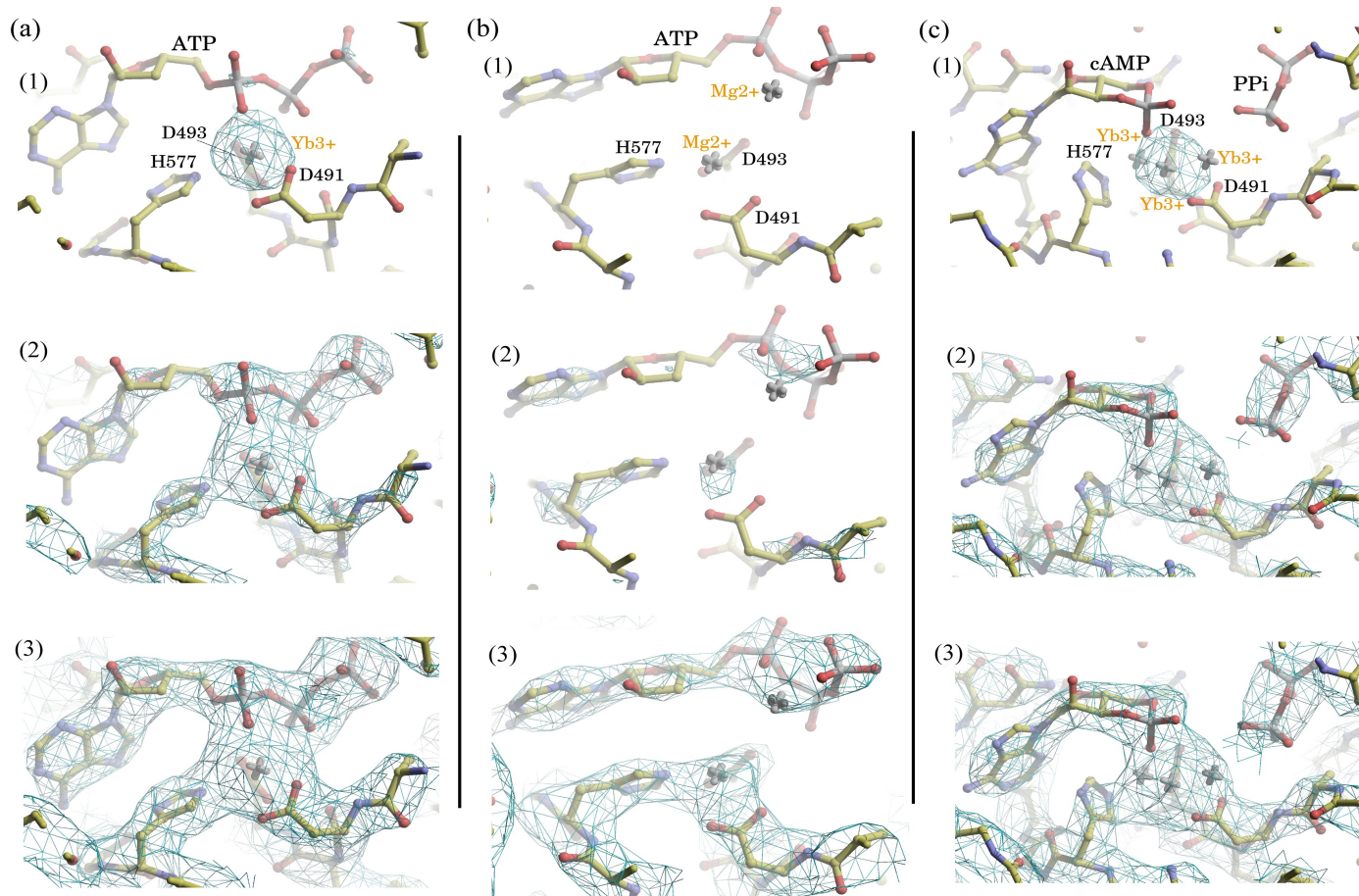
ISSO SERVE PARA ALGUMA COISA?

Sistemas Grandes (biologia)

Quando e por que simulações com pouca estatística servem para alguma coisa:

Sistemas biológicos: Quase toda a informação estrutural tem pouca “estatística”...

Cristal, 4°C, Tampão Fosfato, 3.5Å de resolução, Yb³⁺ no lugar de Mg²⁺...



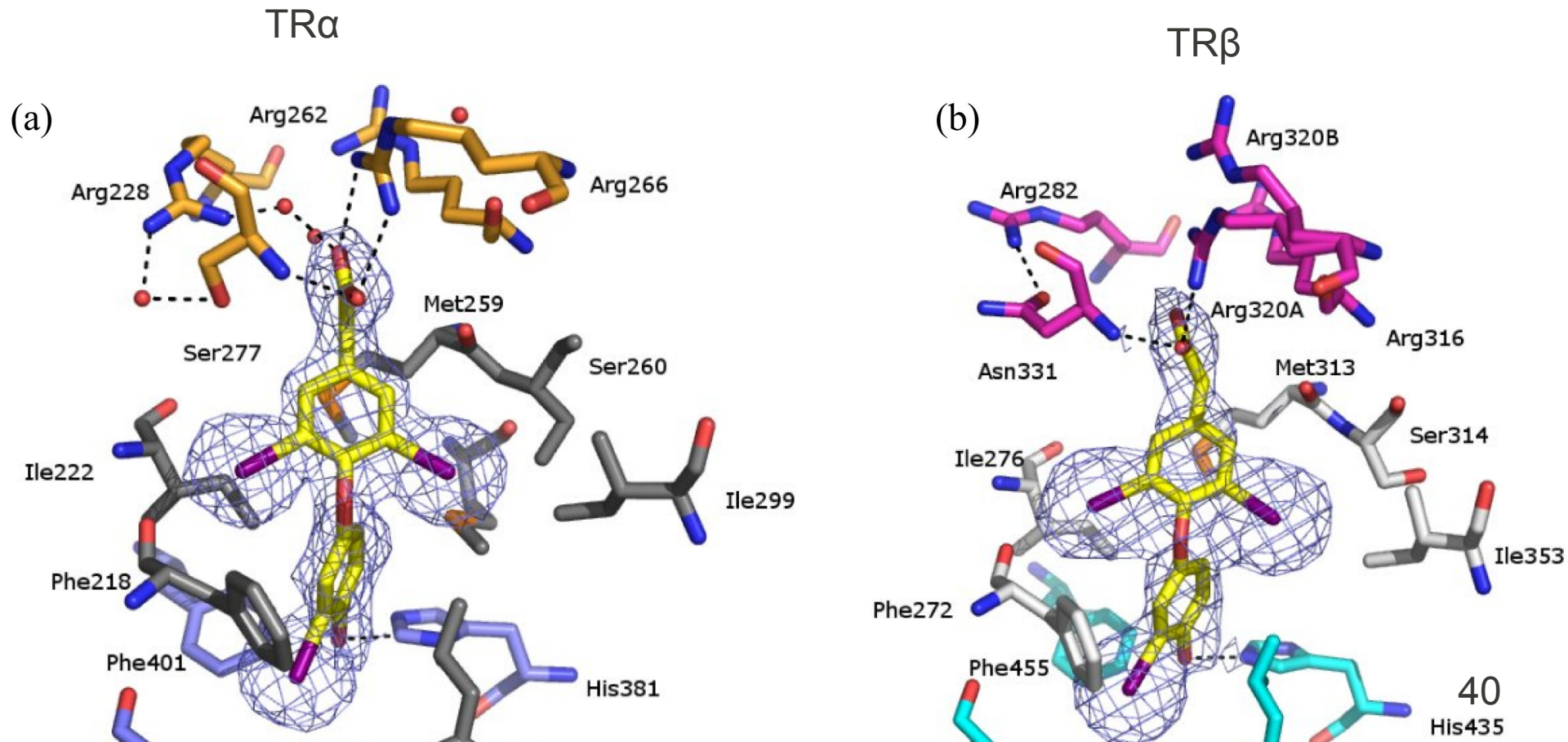
*Fator de Edema
do Antrax
Tang e col.
Nature, 2002.*

Quando e por que simulações com pouca estatística servem para alguma coisa:

Sistemas biológicos: Quase toda a informação estrutural tem pouca estatística...

TRIAC é um hormônio natural que se liga preferencialmente à isoforma β do receptor do hormônio tireoideano (e isso é bom)

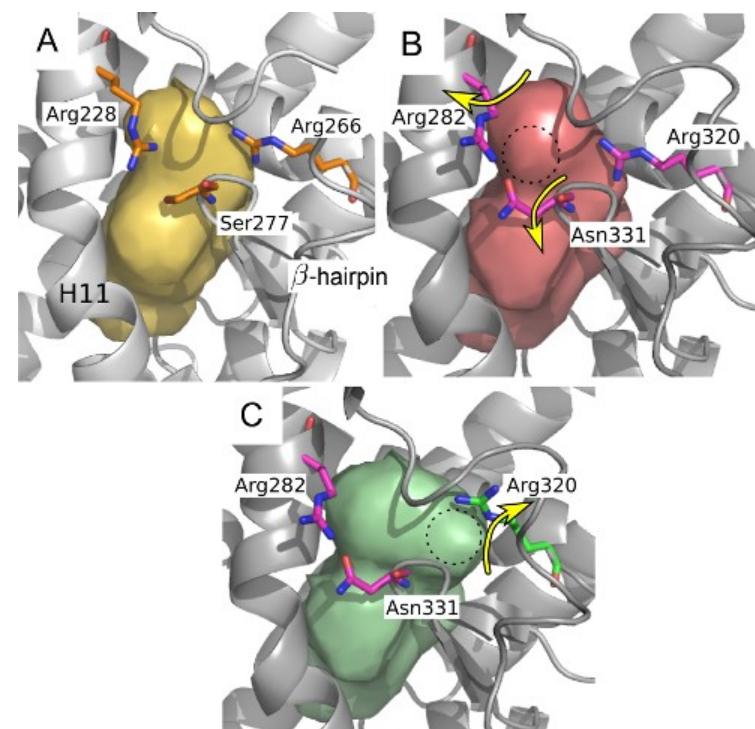
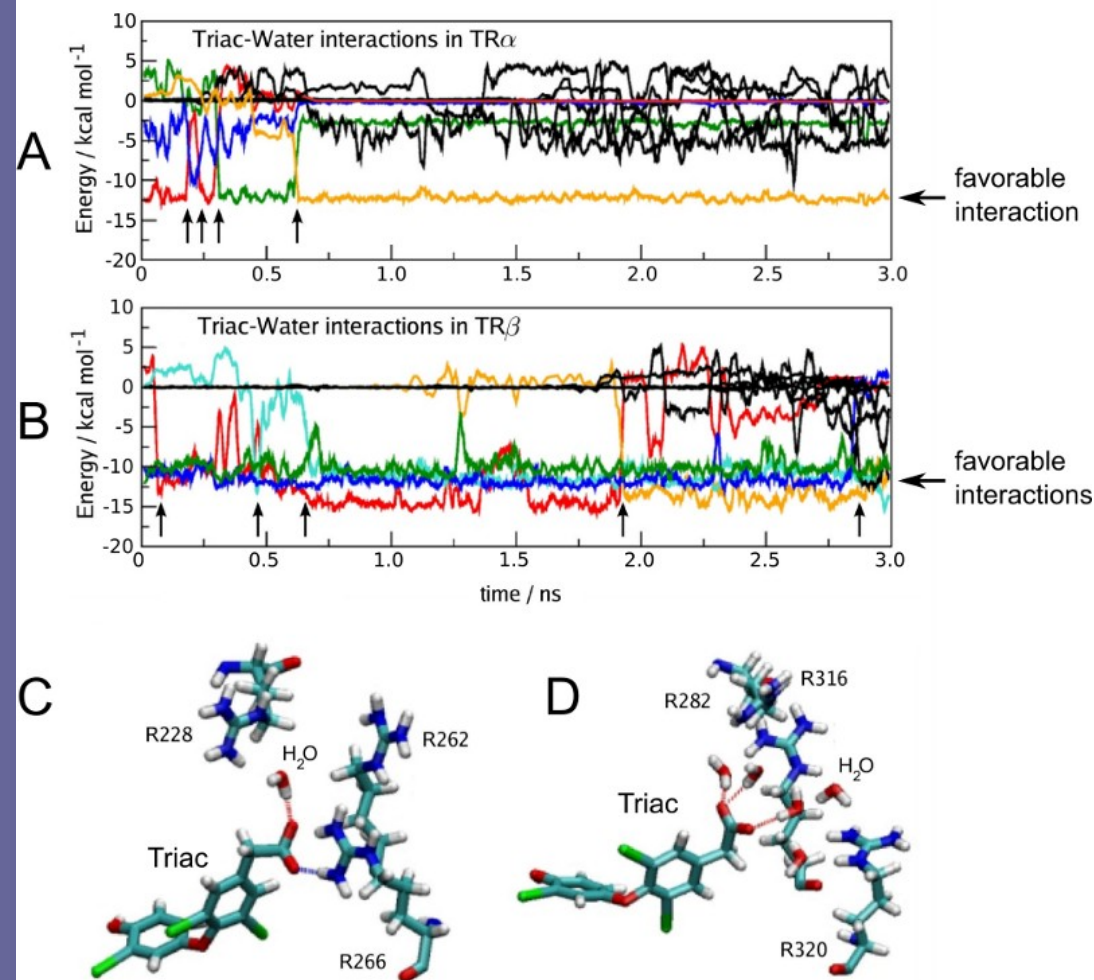
Estruturas cristalográficas:



Quando e por que simulações com pouca estatística servem para alguma coisa:

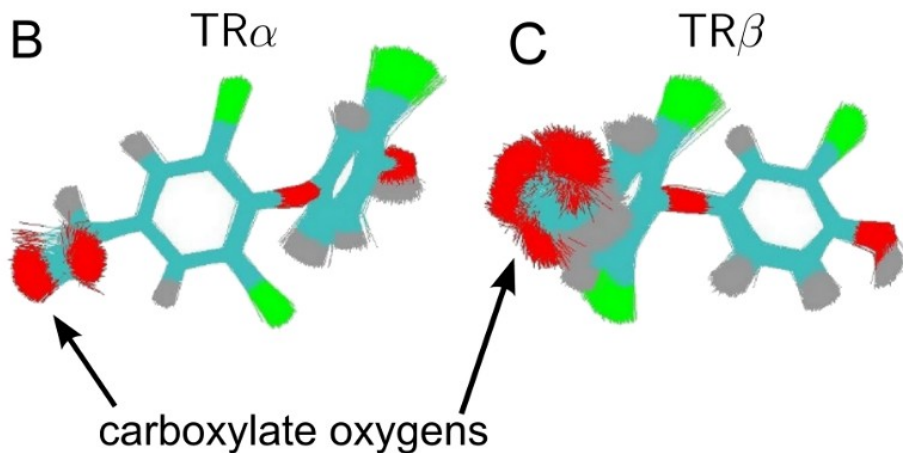
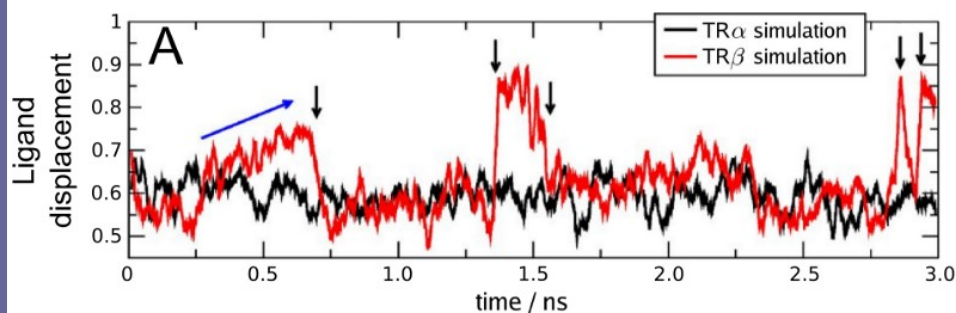
Múltiplas moléculas de água entram no sítio de ligação

De volta à estrutura: uhm... de fato a cavidade era maior no β (+50-100 Å³)

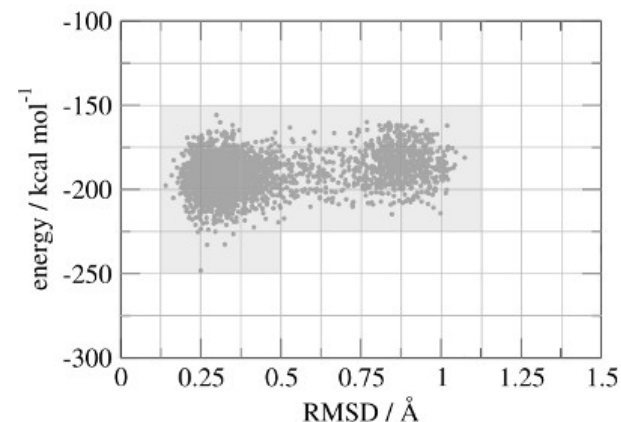


... mas por que TR β seletivo?

Quando e por que simulações com pouca estatística servem para alguma coisa:



Triac tem maior mobilidade na cavidade do TR β e ganha, portanto, entropia conformacional



Estimated conformational entropy gain ~ 0.4 kcal / mol
Experimental $\Delta\Delta G \sim 0.5$ kcal / mol

► Mutação S/N inverte a seletividade, confirmação experimental.

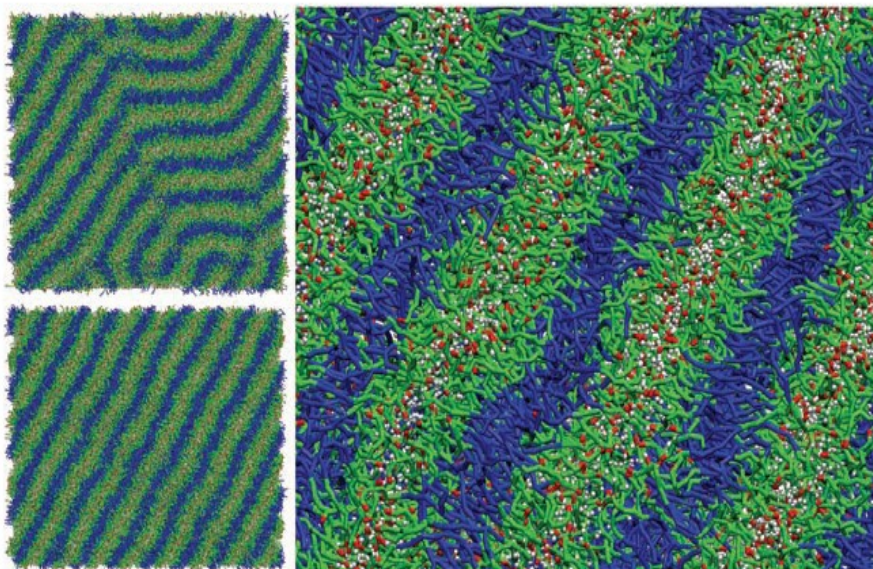


Fig. 1. Images taken from a MD simulation of an aqueous surfactant solution. The simulation system contains approximately 800,000 CG particles and represents a solution with 80% surfactant $C_{12}E_6$ and 20% water. The snapshots were taken after 100 ns (**top**) and ~500 ns (**bottom and right**) of MD simulation, starting from an appropriately hydrated hexagonal phase. The surfactant hydrophobic tails are shown in blue, the head groups in green and red, and the water in white. The phase transformation from initial hexagonal to final lamellar phase passes through a dislocated structure (top) before achieving the equilibrium structure (bottom). The image on the right shows expanded detail from the final structure. The CG parameters used in the MD simulations were taken from (14).

Transição de fase hexagonal
p/ lamelar passa por estrutura
deslocada (top)

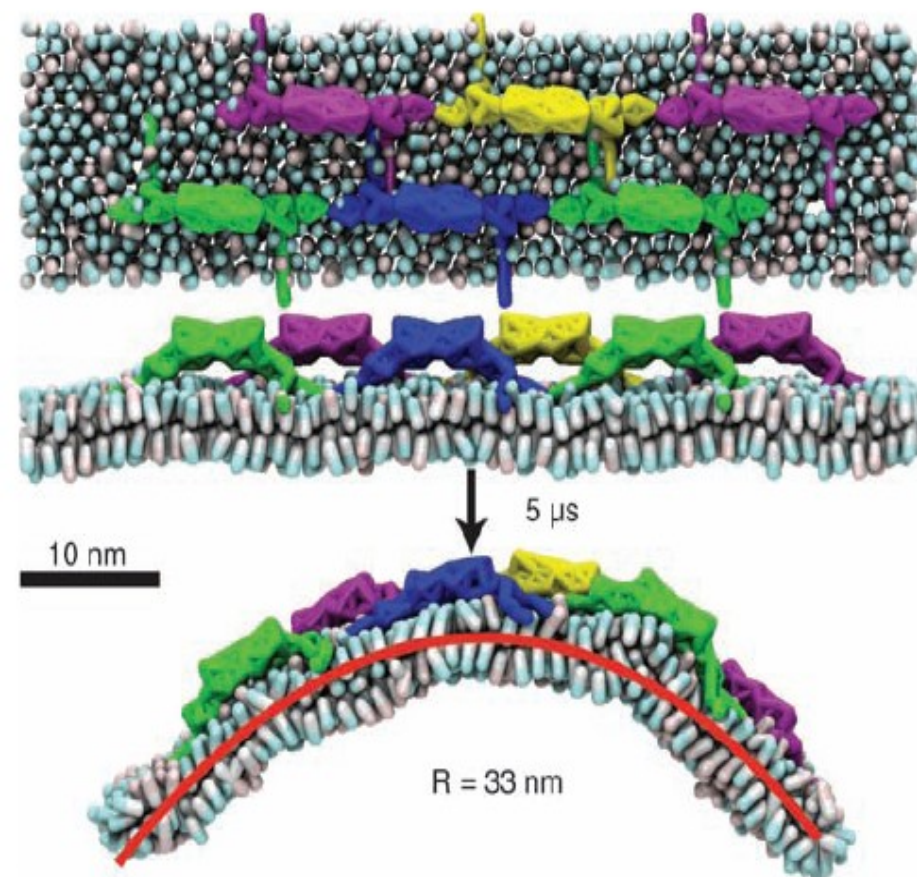
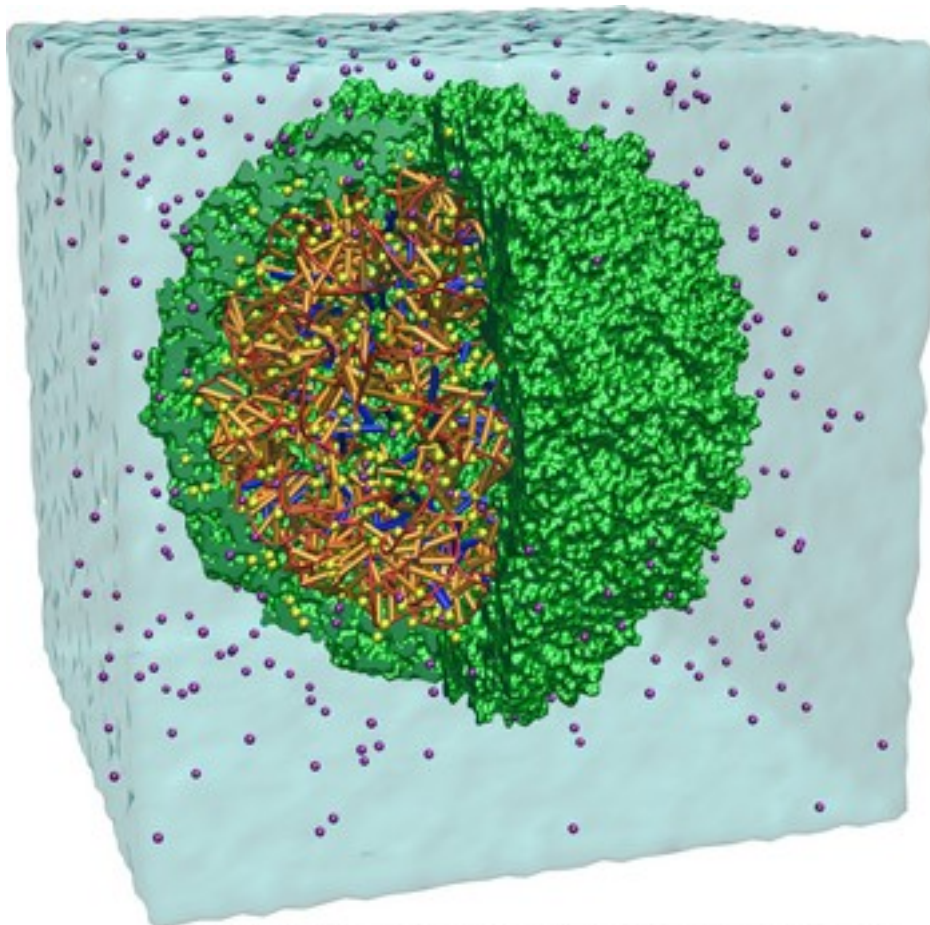


Fig. 2. Images taken from a MD simulation by Arkhipov *et al.*, (17) of biomembrane sculpting by protein-BAR domains. The simulation shown in the figure was carried out using a box with dimensions 100 by 16 by 50 nm and would correspond to a system of 10 million atoms in a fully atomistic representation. However, the computation was done using a shape-based CG model



Theoretical and Computational Biophysics Group
Beckman Institute
University of Illinois at Urbana-Champaign

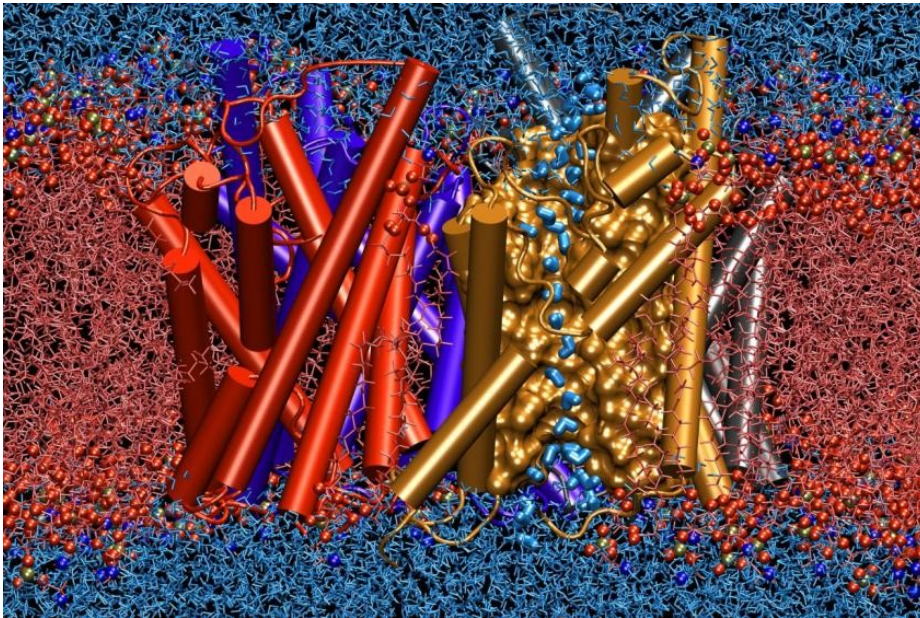
Simulação de um Vírus
(um organismo completo!)

- 1 milhão de átomos
- 50 ns

"A estrutura é estável."

"Sem o RNA, a estrutura não é estável."

Quando e por que simulações com pouca estatística servem para alguma coisa:



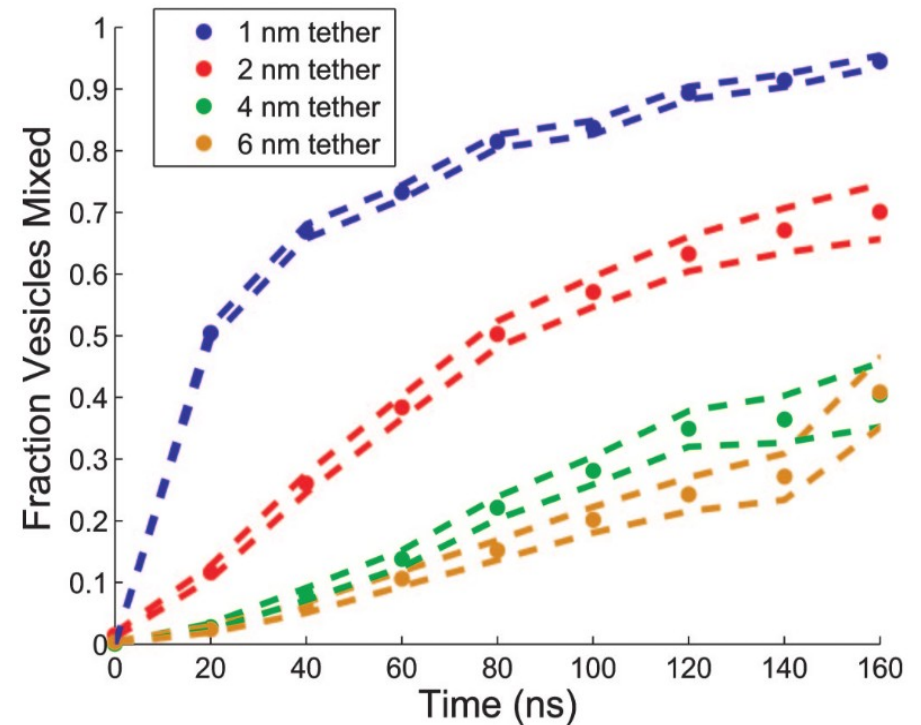
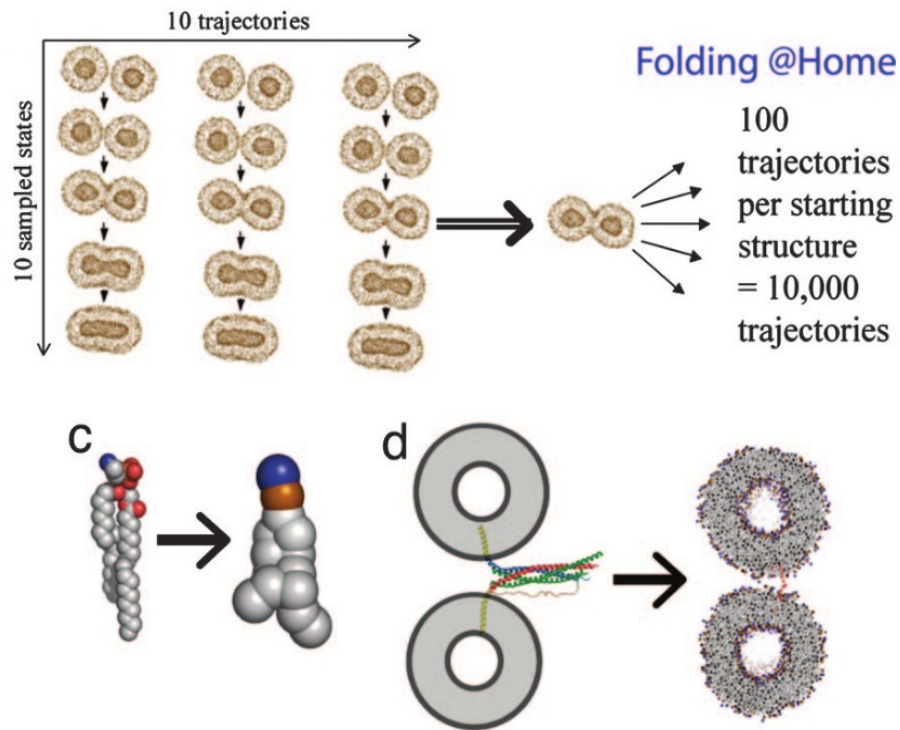
- Exclusão de ions. - 9 eventos observados

Schulten and col.
Science, 296, 525, 2002.⁴⁵

Quando a estatística se encontra com
os sistemas grandes e lentos

Quando a estatística se encontra com os sistemas grandes e lentos:

Folding@Home



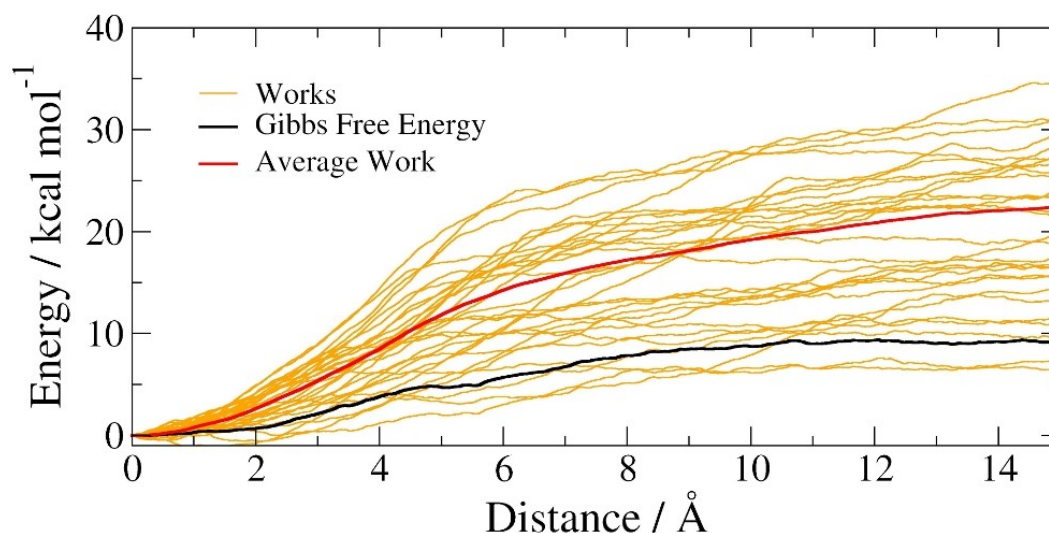
Quando a estatística se encontra com os sistemas grandes e lentos:

Dissociation@Unicamp

$$e^{-\beta\Delta G} = \langle e^{-\beta W} \rangle$$

Equilíbrio Não-equilíbrio

30 simulações de 4 ns (por enquanto)



60.000 átomos, 20 dias por simulação em
4 processadores.

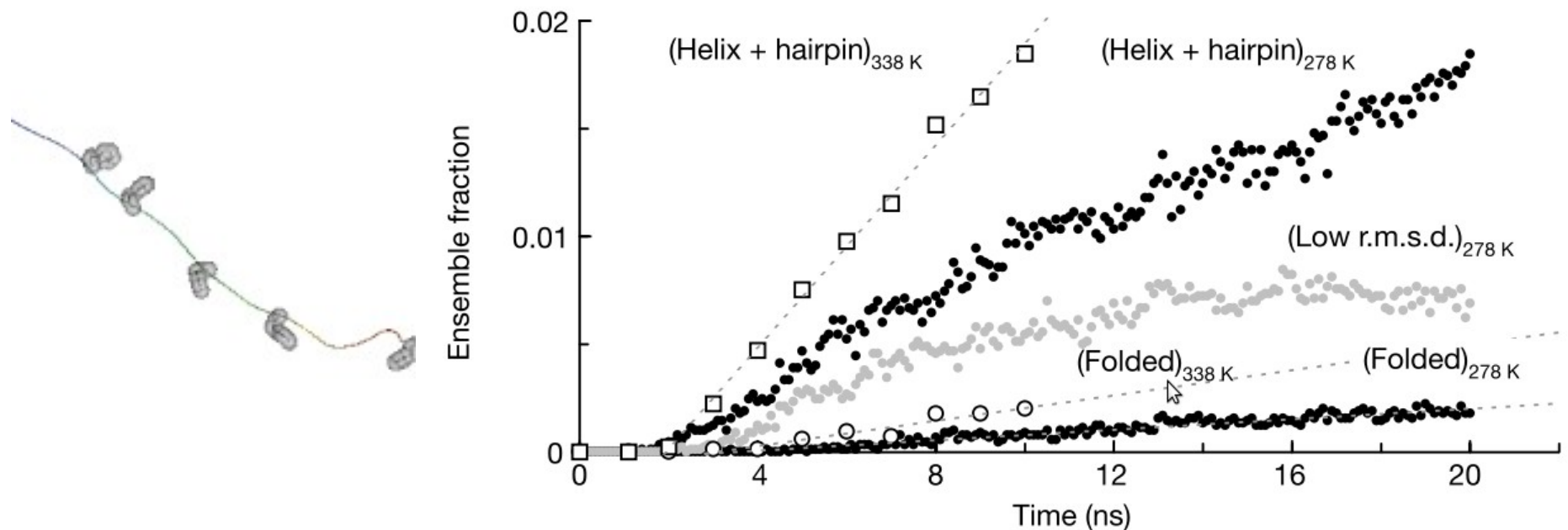
$$\Delta G \sim 8 \text{ kcal mol}^{-1}$$

Paulo C. T. Souza,
Martinez, Skaf, 2009.



Quando a estatística se encontra com os sistemas grandes e lentos:

Folding@home



Tempo médio de enovelamento experimental: $7.5 \pm 3.5 \mu\text{s}$

Tempo médio obtido na simulação: $6 \mu\text{s}$

Como NÃO fazer a mesma coisa:

Not folding@home

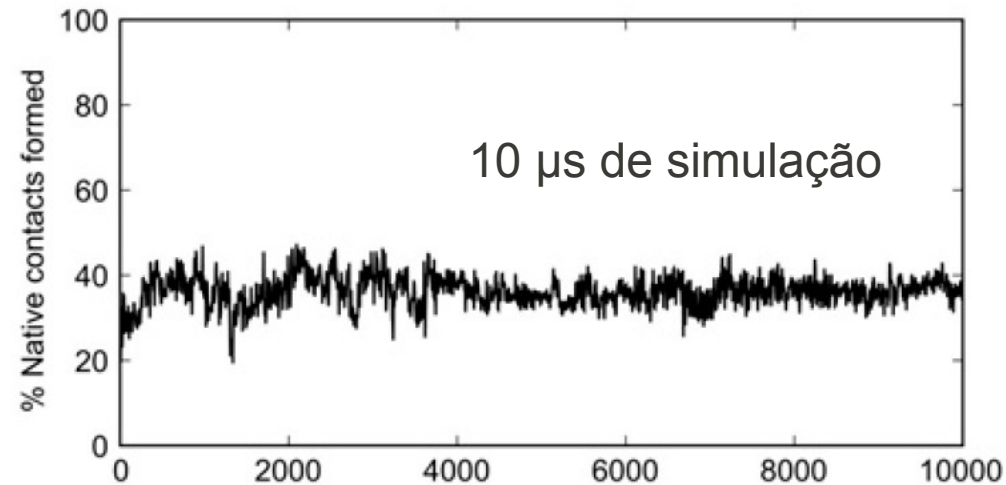
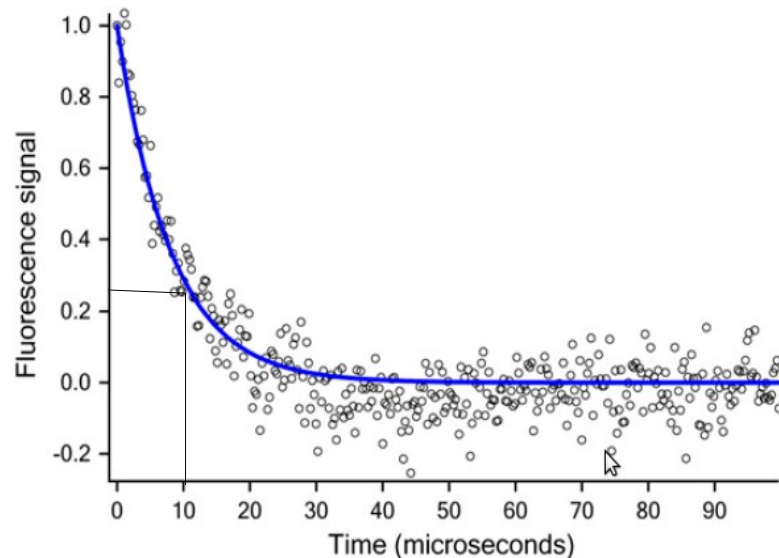


FIGURE 1 Experimental relaxation kinetics of Fip35 with a time resolution of 280 ns at 352 K.

"This 10 μ s is one the longest MD simulation performed to date..."

"A sufficiently accurate force field should, in principle, fold a protein in a single trajectory, on a time-scale similar to the experiment."

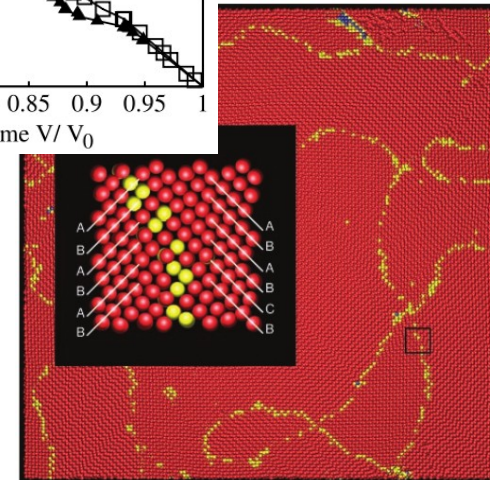
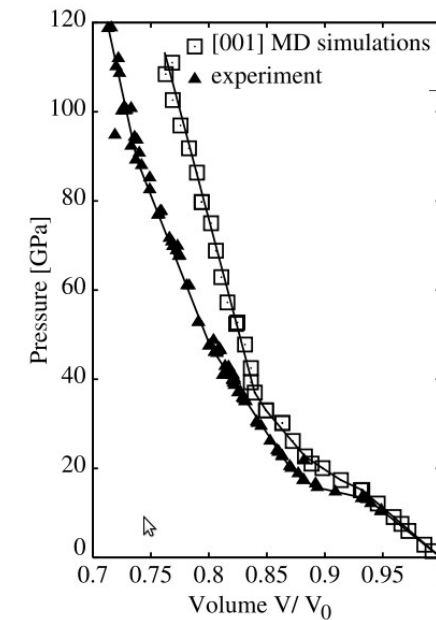
com 70% de chances...

Schulten e col.
Biophys J L75, 2008

Sistemas grandes

Quando a estatística se encontra com os sistemas grandes e lentos:

Propagação de ondas de choque em metais: 8 milhões de átomos.



*Kadau et al,
Science, 296, 16815, 2002.*

Resumo

Conclusões

- MD usa interações clássicas que são aproximações das interações quânticas
- Simulação de milhares de átomos, por tempos "longos"
- Sistemas "pequenos": Muitas repetições podem ser feitas. Médias da simulação correspondem a observáveis experimentais.
- Sistemas "grandes": Simulações são um suporte adicional para o conjunto de informações experimentais.



Presente/Futuro:

- Cada vez sistemas maiores podem ser tratados como "pequenos"
- Melhora progressiva na representação das interações, inclusive pela introdução da MQ.

Agradecimentos:

Munir S. Skaf (Unicamp)
Milton T. Sonoda (UFU)
Igor Polikarpov (IFSC-USP)

Paulo C. T. Souza
Rodrigo Silveira
Lucimara R. Martins

Financiamento

Fapesp / CNPq / CAPES