

Campos de Força Clássicos em Simulações de Dinâmica Molecular

Leandro Martínez
Instituto de Química
Universidade Estadual de Campinas

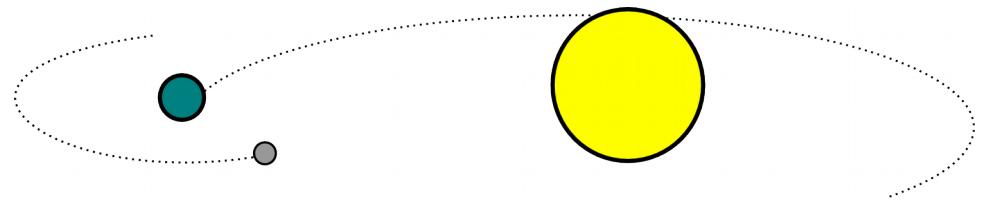
<http://leandro.iqm.unicamp.br>

VIII Escola de Modelagem Molecular em Sistemas Biológicos
Laboratório Nacional de Computação Científica
Petrópolis, RJ
Agosto de 2016

Princípios de uma simulação clássica

1 - Como os objetos interagem:

$$F_i(r_{i,j}) = G \frac{m_i m_j}{r_{i,j}^2}$$



2 - Posições e velocidades iniciais: 23 de Agosto de 2013, 10h 12m 54s ...

3 - Integração das equações de movimento Newtonianas:

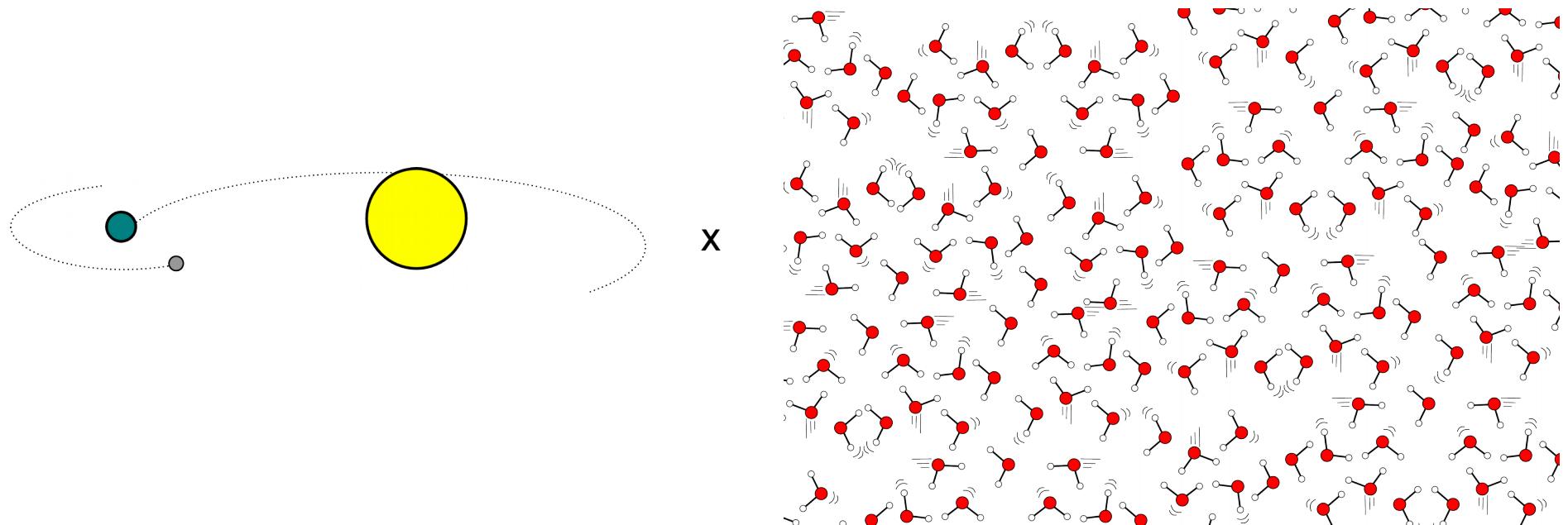
$$\vec{F}(t) = f(\vec{r})$$

$$\vec{a}(t) = \vec{F}(t)/m_i$$

$$\vec{v}(t + \Delta t) = \vec{v}(t) + \vec{a}(t)\Delta t$$

$$\vec{r}(t + \Delta t) = \vec{r}(t) + \vec{v}(t)\Delta t + \frac{1}{2}\vec{a}(t)\Delta t^2$$

Princípios de uma simulação: Sistema planetários vs. Sistema molecular



Força gravitacional \longleftrightarrow Interações intra- e inter-moleculares

Trajetórias precisas \longleftrightarrow Propriedades médias

Energia total \longleftrightarrow Temperatura e pressão

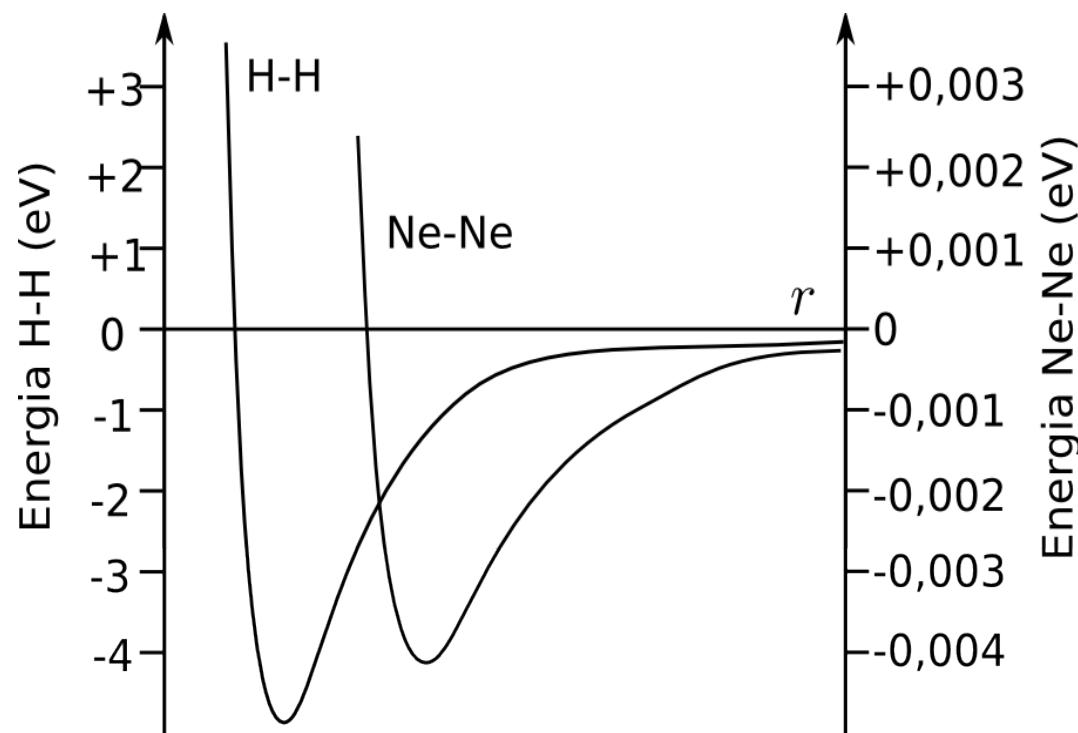
Dias, semanas, anos \longleftrightarrow Nano-segundos

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)

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

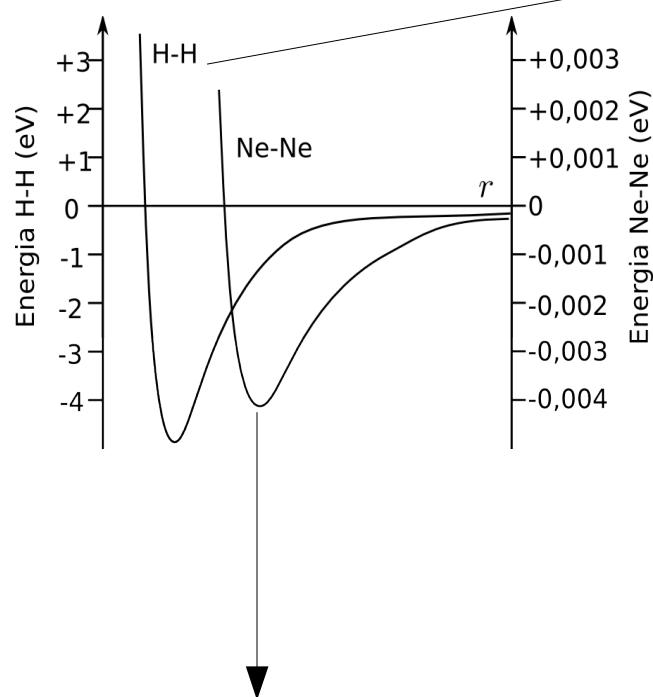
(298 K)

Quântico vs. Clássico

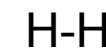
O mundo é quântico...

... mas não tanto assim ...

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

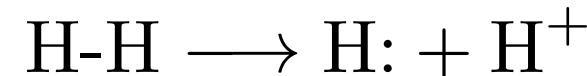
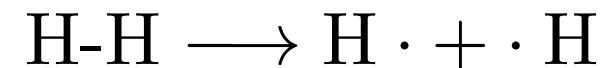


Não há quebra de ligação.
Sem ambiguidade.



Quebra de ligação covalente.

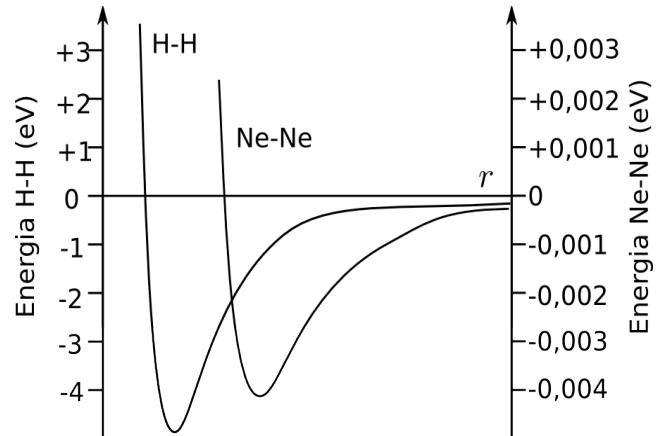
Homolítica ou Heterolítica?



Superfícies de energia diferentes
para cada mecanismo. Representação
clássica muito limitada.

Interações

Campos de força clássicos



Interações intermoleculares:

Podem ser razoavelmente bem representadas usando potenciais simples, para qualquer distância.

“Interações não-ligadas”

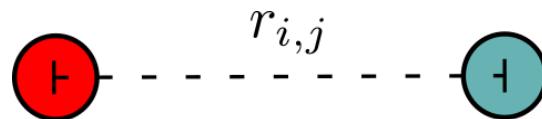
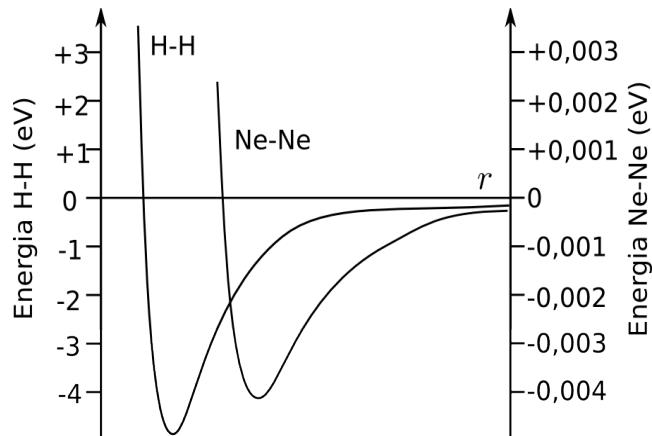
Interações intramoleculares:*

Potenciais específicos que para a representação de deslocamentos intramoleculares, sem ruptura de ligações químicas.

“Interações ligadas”

*Ou entre grupos distantes na mesma molécula.

Interações não-ligadas

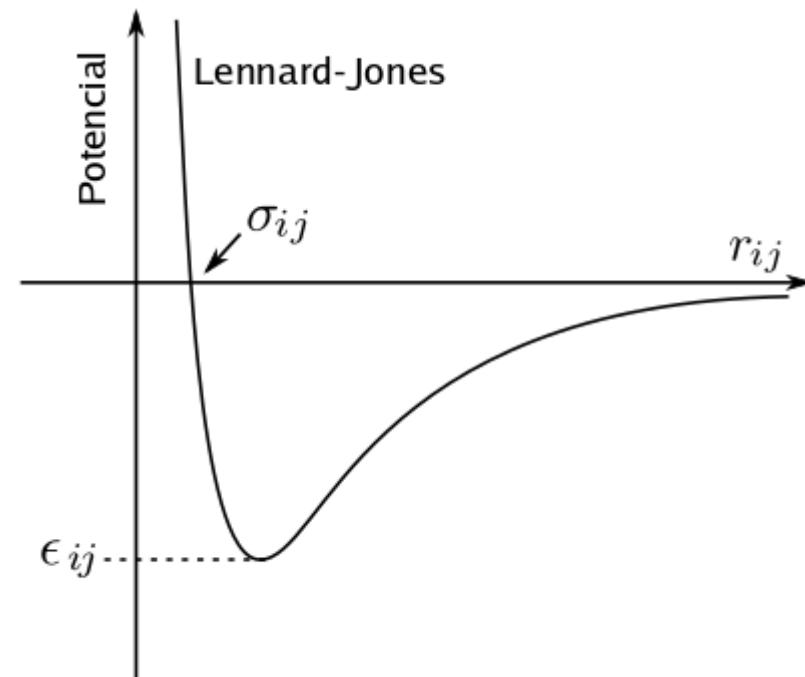


Cada “classe” de átomo possui parâmetros ϵ_i e σ_i que, combinados por pares, determinam os ϵ e σ específicos da interação.

Forças de London, ou de van der Waals:

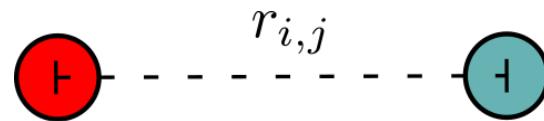
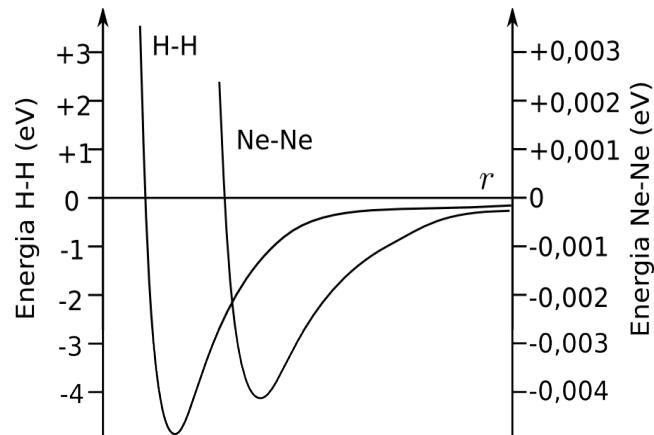
Interações resultantes das flutuações da densidade eletrônica, presentes em todas as moléculas, e dependentes da polarizabilidade:

$$V_{LJ}(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right]$$



Interações

Interações não-ligadas



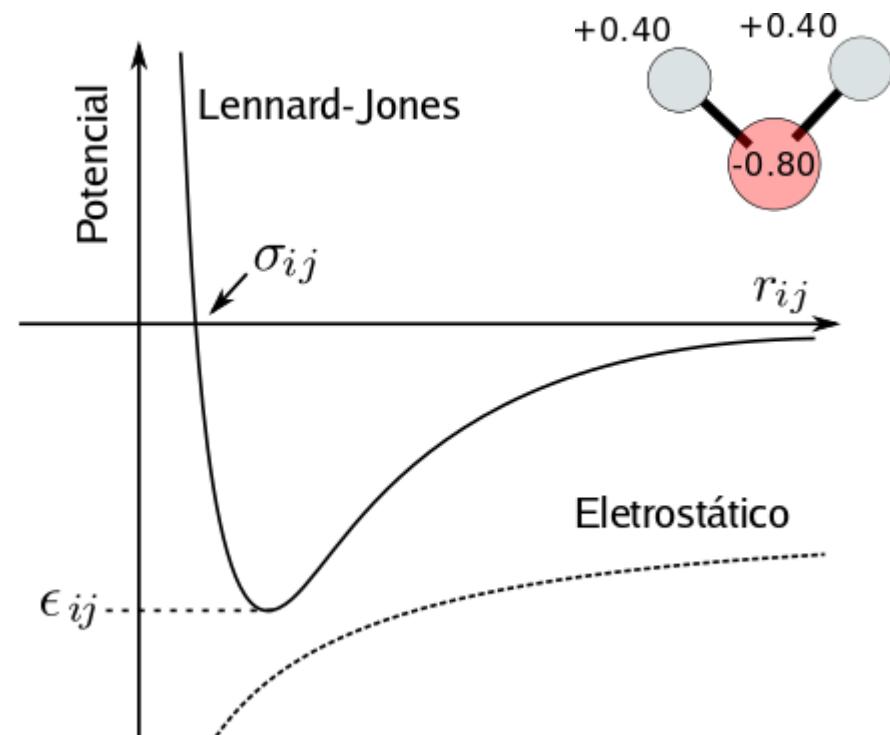
Cada “tipo” de átomo possui uma carga parcial característica.

As cargas são mais sensíveis ao ambiente molecular que a polarizabilidade.

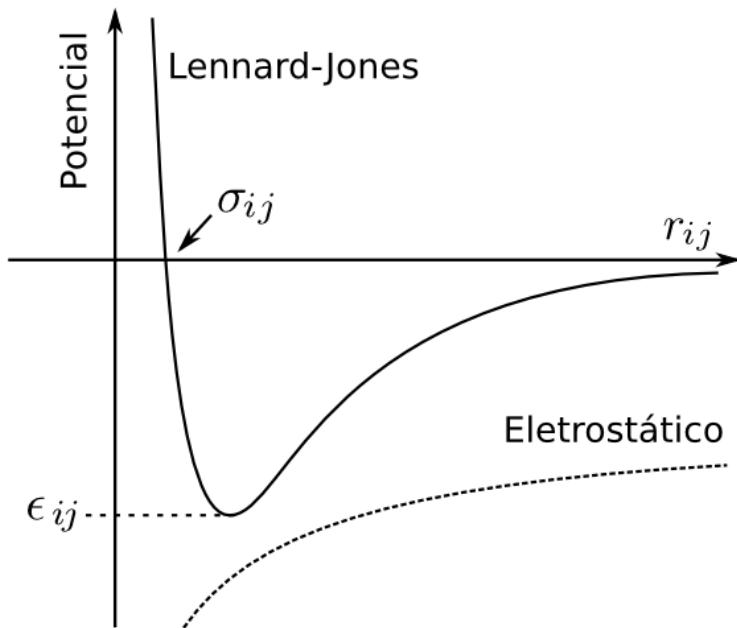
Interações eletrostáticas:

Resultantes diretas das cargas, ou cargas parciais dos átomos:

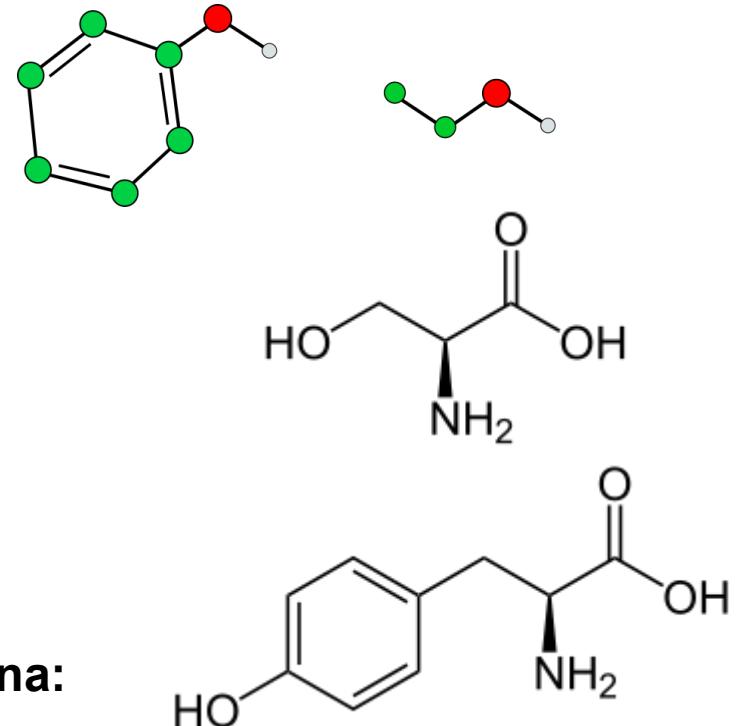
$$V_e(r) = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{r}$$



Interações não-ligadas



Exemplo:

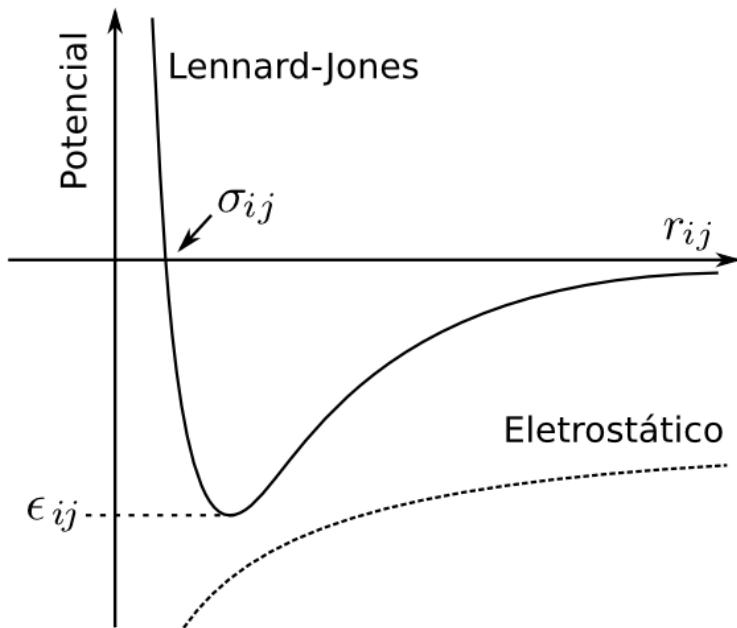


Oxigênios das cadeias laterais de Serina e Tirosina:
(campo de força CHARMM)

	“tipo”	“classe”	carga
SERINA:	ATOM OG	OH1	-0.66
TIROSINA:	ATOM OH	OH1	-0.54
		ϵ_{OH1}	σ_{OH1}
OH1	-0.152100	1.770000	! ALLOW
			ALC ARO
	“classe”		

$$\sigma_{i,j} = \frac{\sigma_i + \sigma_j}{2}$$

Interações não-ligadas



SERINA:	ATOM	OG	OH1	-0.66
TIROSINA:	ATOM	OH	OH1	-0.54
	OH1	-0.152100	1.770000 ! ALLOW	ALC ARO

$$\sigma_{i,j} = \frac{\sigma_i + \sigma_j}{2}$$

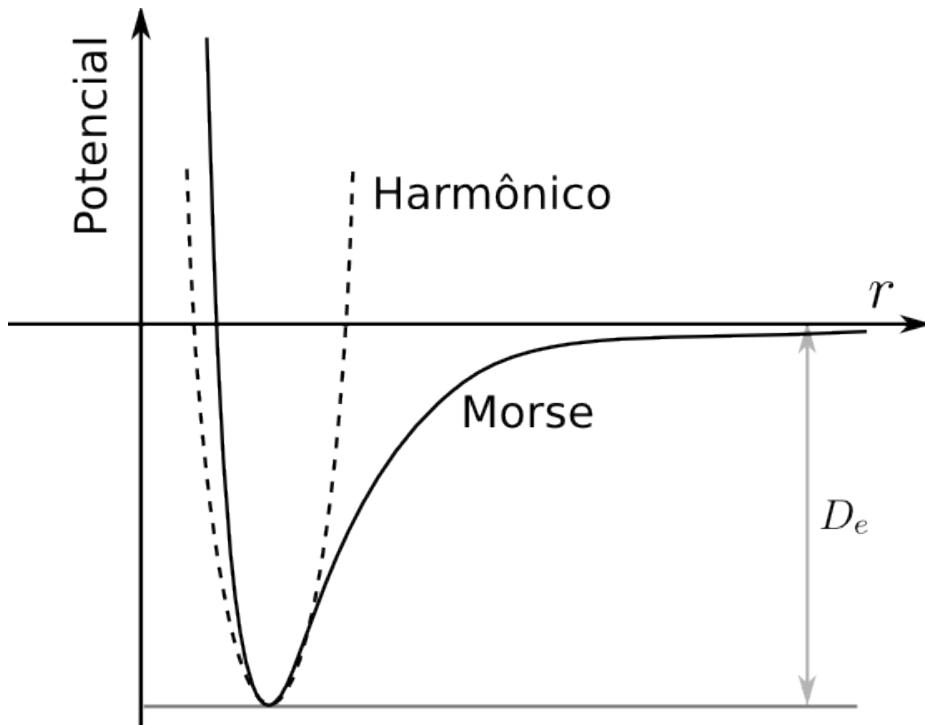
$$V_{\text{nonbonded}} = \sum_{\text{pares}} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}}$$

Parâmetros σ e ϵ definidos por pares de “classes” de átomos (mais gerais).

Cargas parciais definidas por “tipo” de átomo, geralmente bastante específicas.

Interações ligadas

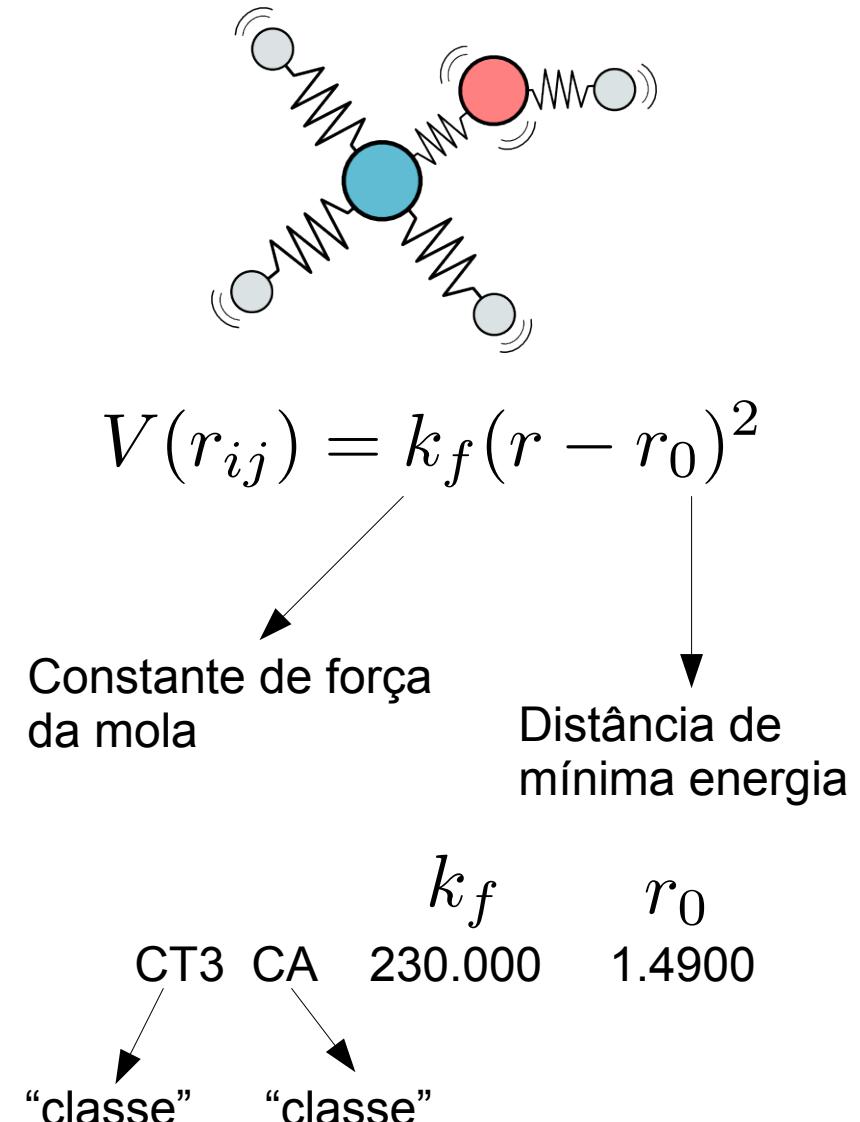
Potenciais não dissociativos - ligações covalentes



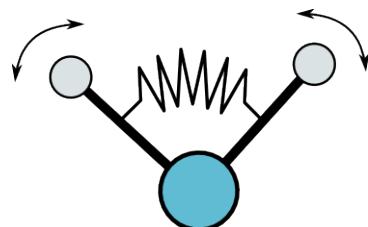
Potencial harmônico é uma boa aproximação para a região de mínimo de energia potencial.

Ligações covalentes são raramente rompidas na temperatura ambiente:

$$D_e \gg kT$$



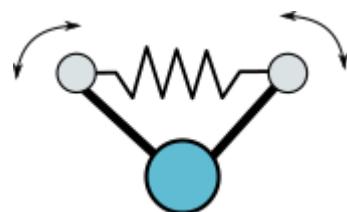
Interações ligadas



Ângulos: $V(\theta) = k_\theta (\theta - \theta_0)^2$

$$\theta = \theta(\vec{r}_1, \vec{r}_2, \vec{r}_3)$$

Potencial harmônico dependente explicitamente do ângulo.

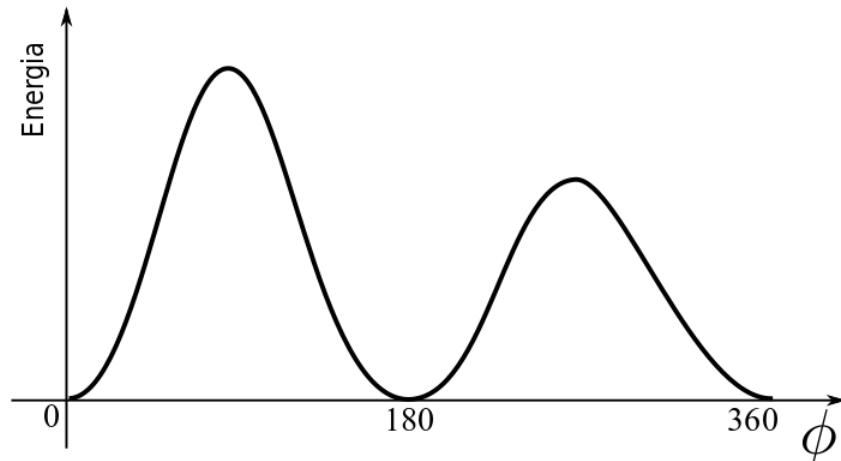
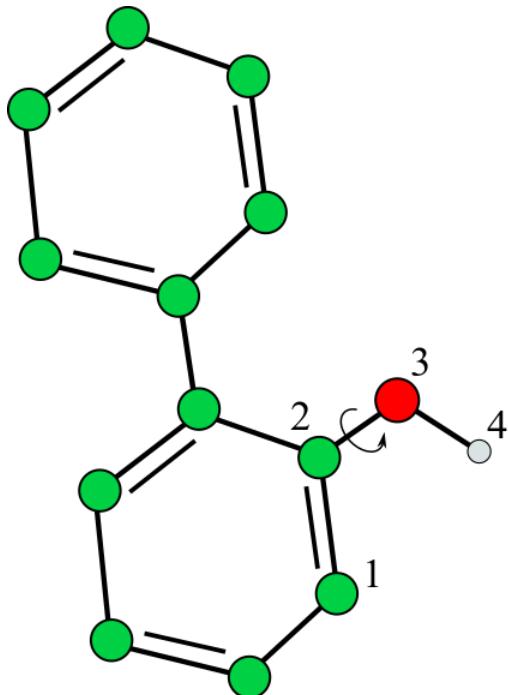


Potencial de Urey-Bradley

$$V(d) = k_{\text{UB}} (d - d_0)^2$$
$$d = d(\vec{r}_1, \vec{r}_3)$$

Ligaçāo entre os átomos 1 e 3 do ângulo. Introduz a correlação entre o comprimento de ligação e o potencial angular - importante para um refinamento da descrição das vibrações. Geralmente não é importante em simulações de biomoléculas.

Diedros:

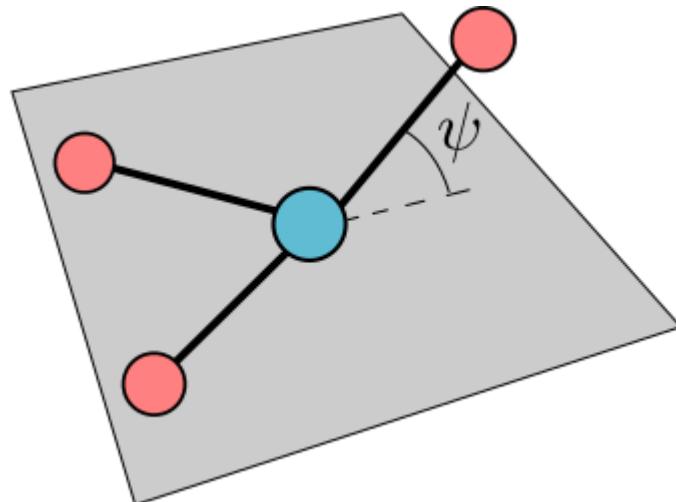
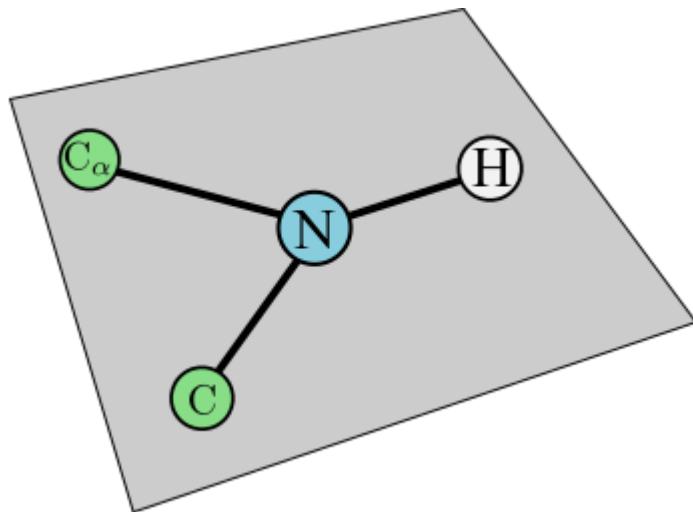


$$V(\phi) = \sum_i^{N_\phi} K_\phi^i [1 + \cos(n_i \phi - \delta_i)]$$
$$\phi = \phi(\vec{r}_1, \vec{r}_2, \vec{r}_3, \vec{r}_4)$$

Soma de vários termos deste tipo para cada diedro.

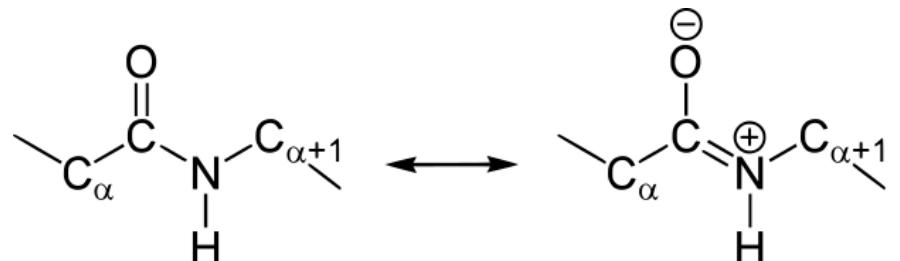
Determinantes para o equilíbrio conformacional!

Diedros impróprios



Manter a planaridade de estruturas

Exemplo: Ligação peptídica.



$$V(\psi) = K_\psi (\psi - \psi_0)^2$$

$$\psi = \psi(r_1, r_2, r_3, r_4)$$

Definição do campo de força

$$\begin{aligned} V = & \sum_{\text{bonds}} K_b(r - r_0)^2 + \sum_{\text{angle}} K_a(\theta - \theta_0)^2 + \\ & + \sum_{\text{UB}} K_{\text{UB}}(S - S_0)^2 + \sum_{\text{dihedrals}} K_\phi [1 + \cos(n\phi - \delta)] + \\ & + \sum_{\text{impropers}} K_i(\psi - \psi_0)^2 + \\ & + \sum_{\text{pares}} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}} \end{aligned}$$

Parâmetros: $K_b, r_0, K_a, \theta_0, K_{\text{UB}}, S_0, K_\phi, n, \delta, K_i, \phi_0, \epsilon, \sigma, q$

Definição do campo de força

$$K_b, r_0, K_a, \theta_0, K_{\text{UB}}, S_0, K_\phi, n, \delta, K_i, \phi_0, \epsilon, \sigma, q$$

Topologia:

```
RESI CYS          0.00
GROUP
ATOM N   NH1    -0.47  |   |
ATOM HN  H      0.31   | HN-N
ATOM CA  CT1    0.07   |   |
ATOM HA  HB     0.09   |   |
GROUP
ATOM CB  CT2    -0.11  |   |
ATOM HB1 HA     0.09   |   |
ATOM HB2 HA     0.09   |   O=C
ATOM SG  S      -0.23  !   |
ATOM HG1 HS     0.16   |
GROUP
ATOM C   C      0.51
ATOM O   O     -0.51
BOND CB CA    SG CB   N HN   N  CA
BOND C  CA    C +N  CA HA   CB HB1
BOND CB HB2   SG HG1
DOUBLE O  C
IMPR N -C CA HN   C CA +N O
```

Parâmetros:

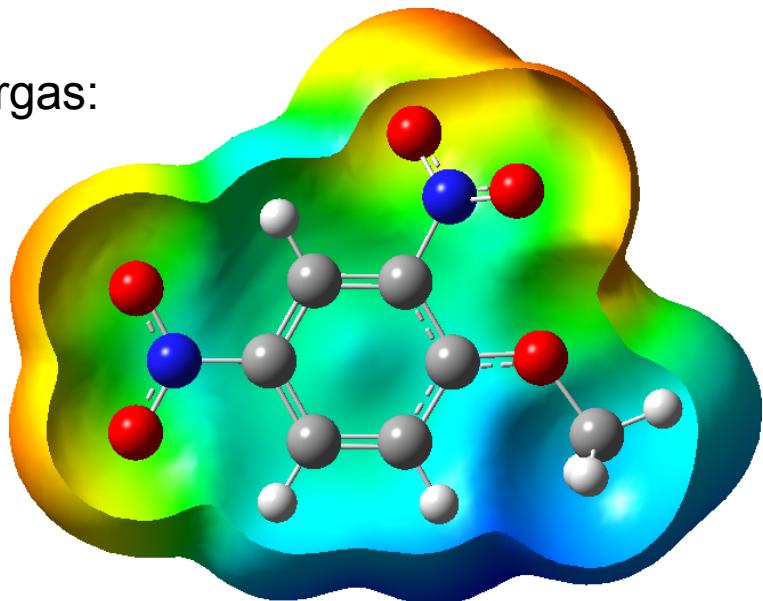
```
BONDS
!
!V(bond) = Kb(b - b0)**2
!
!Kb: kcal/mole/A**2
!b0: A
!
!atom type Kb           b0
!
!Carbon Dioxide
CST  OST   937.96    1.1600 ! JES
C    C     600.000   1.3350 ! ALLOW ARO HEM
                                ! Heme vinyl substituent (KK, from propene (JC
CA   CA    305.000   1.3750 ! ALLOW ARO
                                ! benzene, JES 8/25/89
CE1  CE1    440.000   1.3400 !
                                ! for butene; from propene, yin/adm jr., 12/95
CE1  CE2    500.000   1.3420 !
                                ! for propene, yin/adm jr., 12/95
CE1  CT2    365.000   1.5020 !
```

De onde vieram todos esses parâmetros?

Cada classe de átomos tem seus parâmetros, cada tipo suas cargas, cada molécula sua topologia...

Ajuste dos parâmetros usando cálculos ab-initio

Cargas:

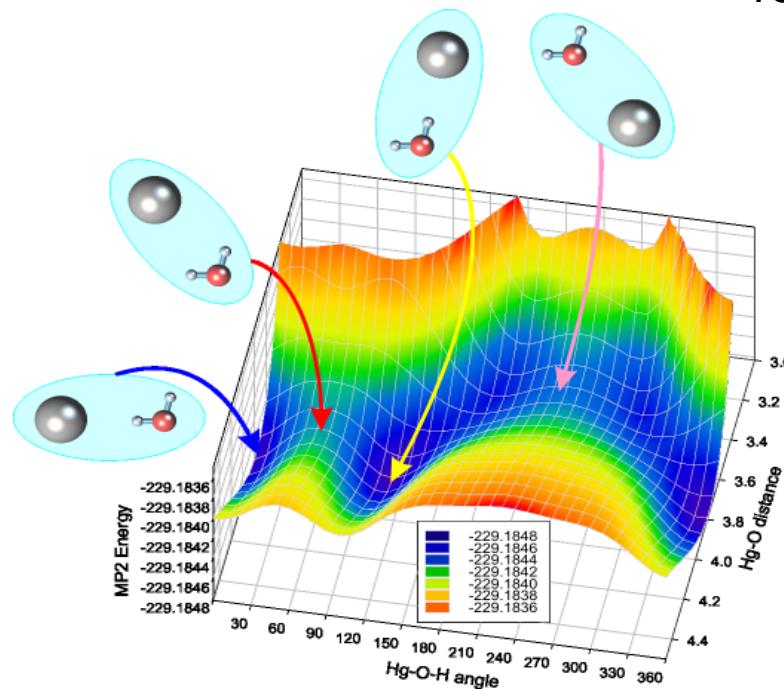
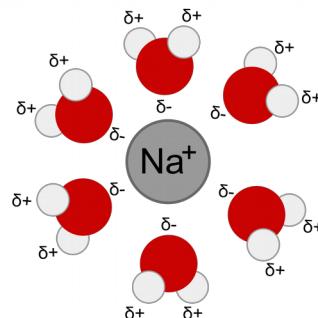
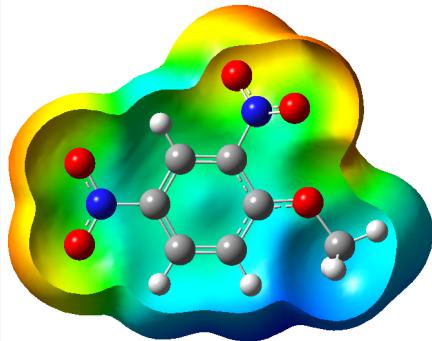


Potencial eletrostático ao redor da estrutura, calculado quanticamente

Quais são as cargas pontuais (cargas parciais) dos átomos, que melhor reproduzem este potencial eletrostático?

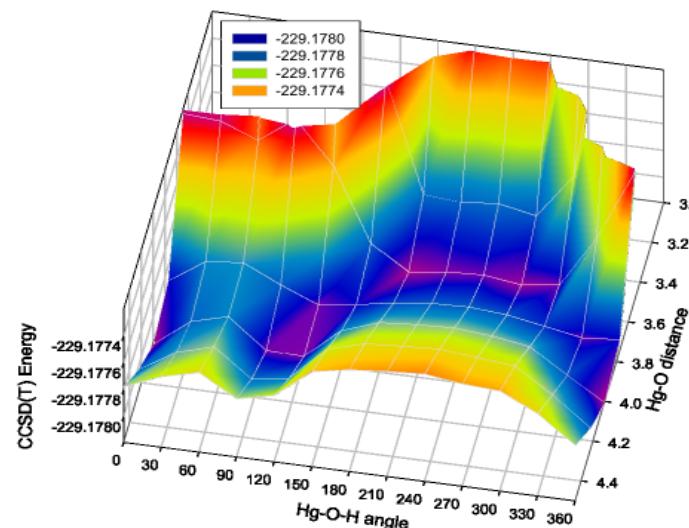
Ajuste do potencial eletrostático quântico: determinação das cargas parciais

Ajuste dos parâmetros usando cálculos ab-initio

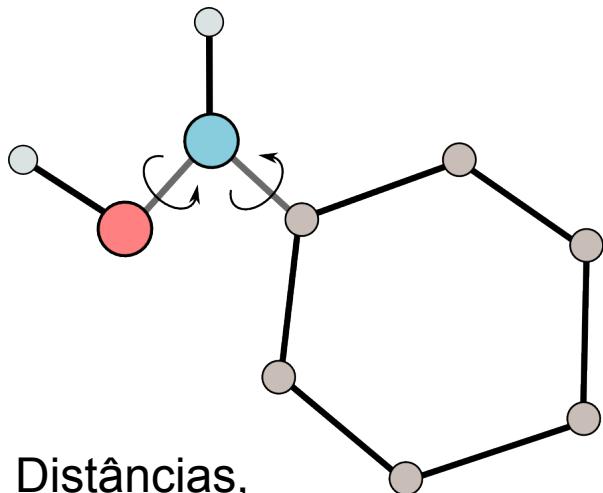


Parâmetros de Lennard-Jones:

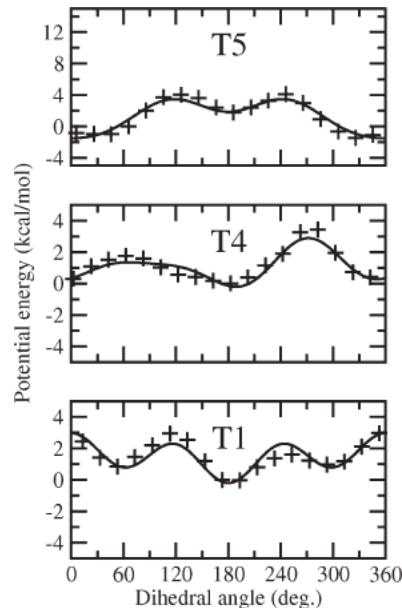
Dadas as cargas, calcular a energia potencial quântica para diferentes distâncias e orientações intermoleculares: ajustar os parâmetros para reproduzir a superfície de energia potencial.



Ajuste dos parâmetros usando cálculos ab-initio:



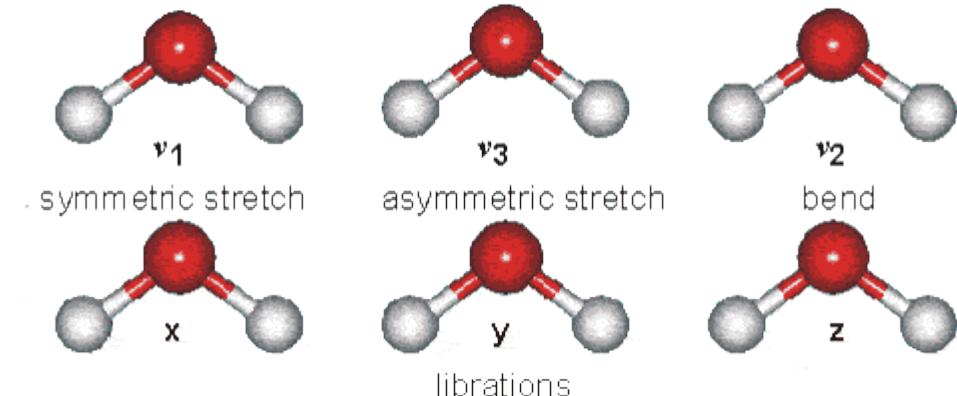
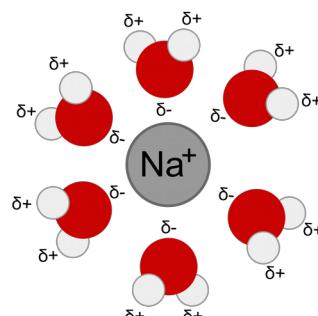
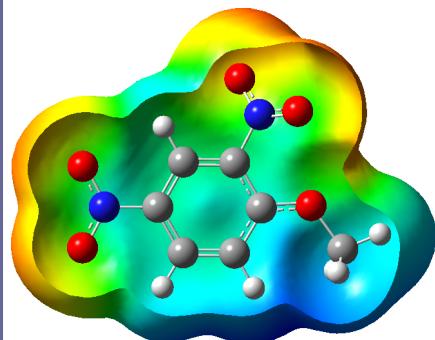
Distâncias,
constantes
de força de
ligações, etc.



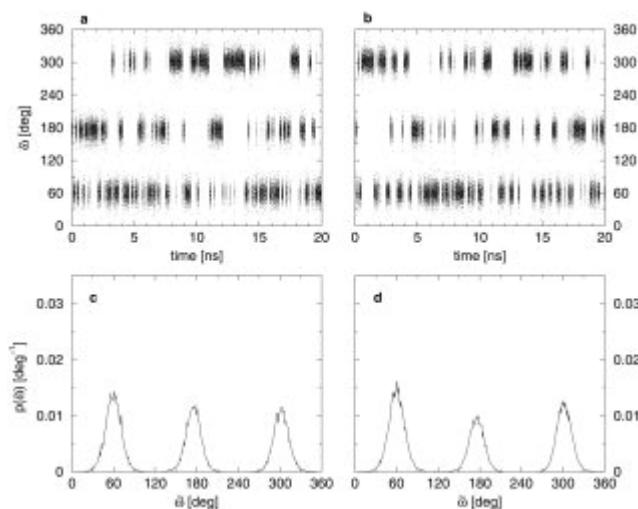
→ Energia ab-initio
inclui todas as interações
(ângulos, diedros, cargas,
vdW, etc)

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

Ou, alternativamente, usando dados experimentais

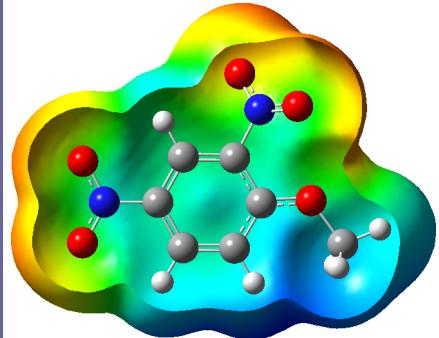


Parâmetros para ligações covalentes e ângulos podem ser obtidos a partir de dados de espectroscopia vibracional.



A parametrização de diedros pode ser validada comparando as populações de conformeres em um simulação com experimentos de RMN (por exemplo).

Ou, alternativamente, usando dados experimentais



$$\Delta H_{\text{vap}} = H_g - H_l$$

$$H_g = U_g + pV_g$$

$$H_l = U_l + pV_l$$

$$\Delta H_{\text{vap}} = U_g - U_l + p(V_g - V_l)$$

$$V_l \ll V_g \quad pV_g = RT/n$$

U_g : Energia intramolecular - conformacional

U_l : Energia potencial média do líquido

1. Simulação da molécula em fase gasosa; 2. Simulação do líquido

Campos de força populares

AMBER, CHARMM, OPLS, GROMOS

(AMBER, CHARMM e GROMOS também nomes de programas de simulação)

Filosofia de parametrização diferente.

Uso ou não de dados experimentais; Validação com dados experimentais; etc.

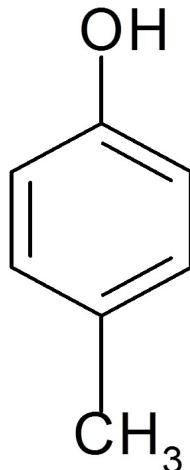
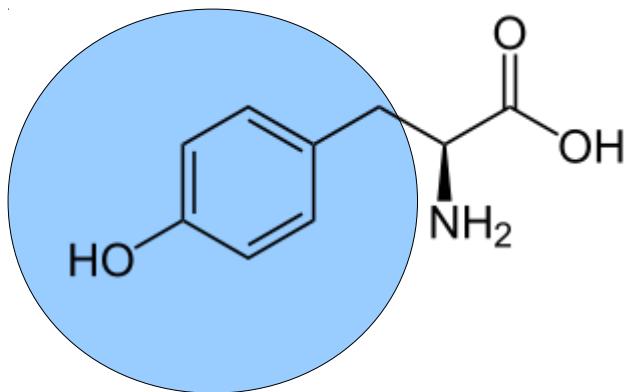
Essencialmente não-compatíveis entre si. Só se souber muito bem o que está fazendo (AMBER e OPLS tem semelhanças, por exemplo).

Características desejáveis, e supostamente satisfeitas:

- Os parâmetros são transferíveis entre grupos similares.
- O campo de força de um sistema complexo pode ser obtido da soma de suas partes.

Campos de força populares

TRANSFERÍVEIS



Grupos similares:

Mesmas classes de átomos,
mesmos parâmetros.

Geralmente as cargas devem
ser recalculadas.

Exemplo: Os parâmetros para os aminoácidos são independentes da proteína simulada (inclusive as cargas). Não é necessário parametrizar nada para simular uma nova proteína que não tenha complicações.

Campos de força populares

ADITIVIDADE

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

A forma funcional é uma soma de termos.

Mais que isso, assume-se a auto-consistência do campo de força.

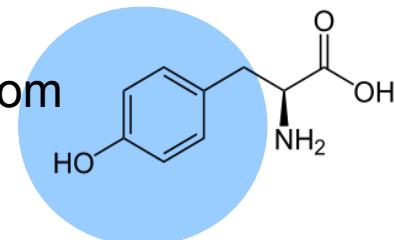
Isto é, moléculas (ou partes de) podem ser combinadas para formar um sistema complexo, e o potencial será a soma dos potenciais das partes.

(fácil de entender dado que a forma funcional é uma soma de termos, mas poderia não funcionar. É uma suposição, razoável em muitos casos, que funcione).

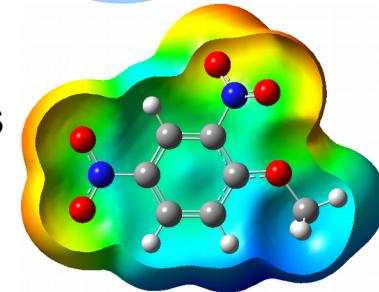
Ex: Uma molécula muito polarizável em solventes de polaridades muito diferentes. Pode falhar.

Parametrização de uma nova molécula

1. Obter parâmetros de LJ, ligações, ângulos por analogia de grupos com estruturas já parametrizadas.



2. Geralmente as cargas devem ser recalculadas, usando os protocolos adequados para os campos de força em uso.

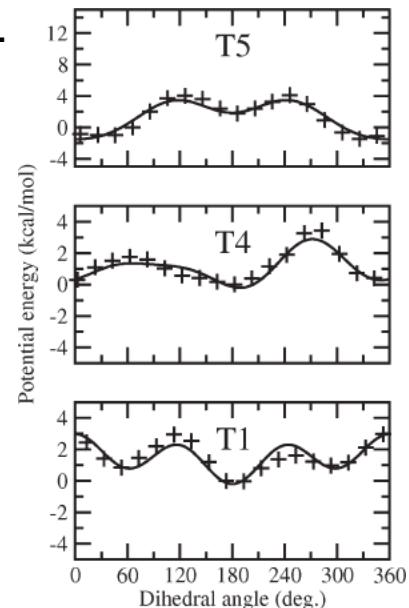


3. Geralmente alguns diedros devem ser reparametrizados.

Muito cuidado com casos especiais!

Ex. Carboidratos: parece facilmente transferível, mas não é!

Ex. Moléculas inusuais, muito polarizáveis. Muito carregadas. Líquidos iônicos, etc.



A qualidade do resultado depende da qualidade da parametrização.

Parametrização de uma nova molécula

Ferramentas de auxílio (automatizações):

Antechamber

Antechamber is a set of auxiliary programs for molecular mechanic (MM) studies. This software package is devoted to solve the following problems during the MM calculations: (1) recognizing the atom type; (2) recognizing bond type; (2) judging the atomic equivalence; (3) generating residue topology file; (4) finding missing force field parameters and supplying reasonable and similar substitutes. As an accessory module in the AMBER 7 and AMBER 8 packages, antechamber can generate input automatically for most organic molecules in a database. The algorithms behind these manipulations may be useful outside the Amber family of programs as well.

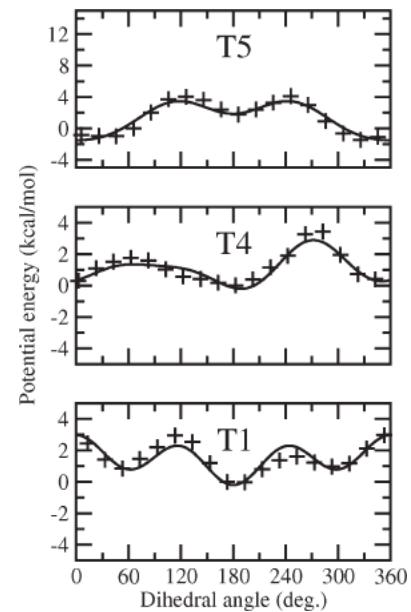
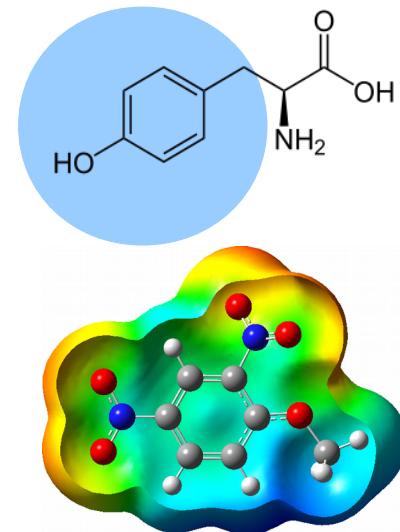
CGenFF

Welcome to the CHARMM General Force Field for organic molecules. If you want to perform simulations on a system that contains a small organic molecule using the CHARMM force field, you've come to the right place. The CGenFF program will provide you with parameters for your specific molecule, which you can read into your simulation software together with the main CGenFF topology and parameter files and parts or all of the CHARMM36 biomolecular force field. Before starting the simulation, please take a moment to examine the penalty scores in the output of the CGenFF program.

SwissParam

This service provides topology and parameters for small organic molecules compatible with the CHARMM all atoms force field, for use with CHARMM and GROMACS.

These parameters are intended for "drug design" type calculations, like docking, minimizations, rapid binding energy estimation (LIECE), etc... Deeper parameterizations are required for applications in which a very fine description of the molecule is required, (e.g. normal mode calculations...). SwissParam results can then be used as a starting point for such a procedure.



Campos de força clássicos especiais

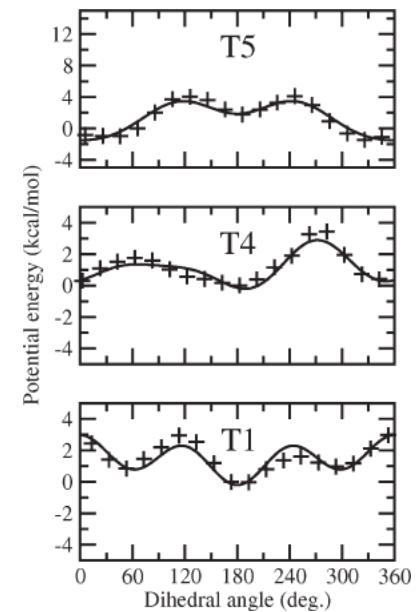
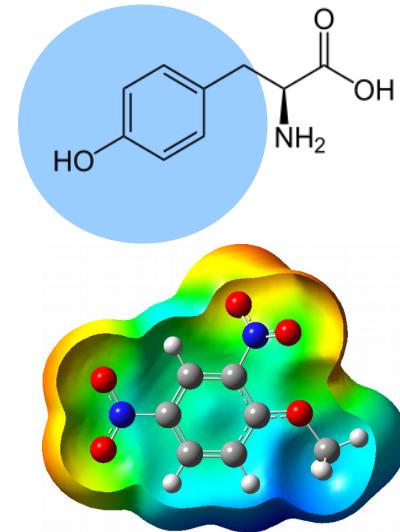
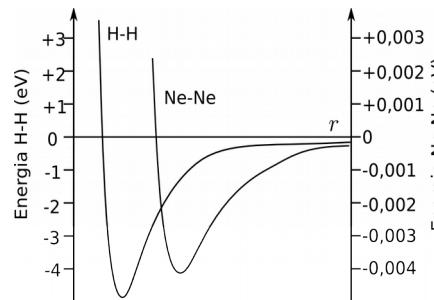
- Reativos (ReaxFF)

- Polarizáveis (líquidos iônicos)

- Coarse-Grained (Sistemas muito grandes)

- Baseados na estrutura (potenciais harmônicos ligando Cas), para estudos de enovelamento.

- Solvente implícito



Agradecimentos:

Organização do EMMSB (Paulo, Laurent, etc.)

FAPESP/CNPq

Leandro Martínez
<http://leandro.iqm.unicamp.br>

