

Modelação e Computação Científica

1º semestre, 2020-2021

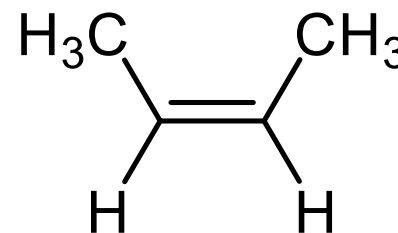
MIEC

20 de outubro de 2020

Formato SMILES: *simplified molecular-input line-entry system*

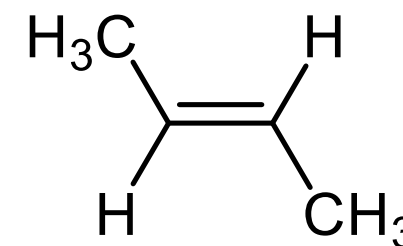
Representação de estruturas químicas através de caracteres ASCII

.	Sem ligação formal
-	Ligação simples (sem traço tb.)
=	Ligação dupla
#	Ligação tripla
\$	Ligação quádrupla
:	Ligação aromática (ou letra minúscula)
//	Estereoquímica (trans)
/\	Estereoquímica (cis)
()	Ramificações
C1 C1	Compostos cíclicos
[]	Indicar espécies carregadas e outras



cis-but-2-eno

C\C=C/C



trans-but-2-eno

C\C=C\C

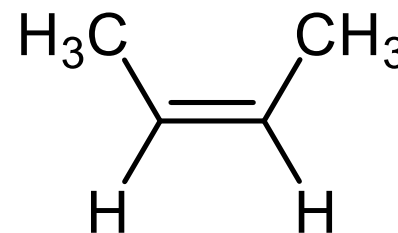
babel -i smi filein.smi -o xyz fileout.xyz --gen3D

obabel -i smi filein.smi -o xyz -O fileout.xyz --gen3D

Formato SMILES: *simplified molecular-input line-entry system*

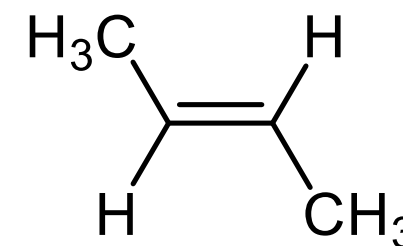
Representação de estruturas químicas através de caracteres ASCII

.	Sem ligação formal
-	Ligação simples (sem traço tb.)
=	Ligação dupla
#	Ligação tripla
\$	Ligação quádrupla
:	Ligação aromática (ou letra minúscula)
//	Estereoquímica (trans)
/\	Estereoquímica (cis)
()	Ramificações
C1 C1	Compostos cíclicos
[]	Indicar espécies carregadas e outras



cis-but-2-eno

C\C=C/C



trans-but-2-eno

C\C=C\C

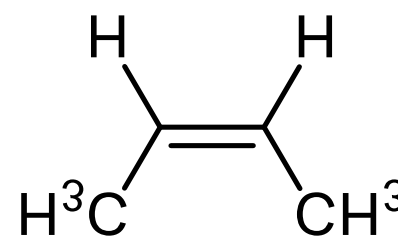
babel -i smi filein.smi -o xyz fileout.xyz --gen3D

obabel -i smi filein.smi -o xyz -O fileout.xyz --gen3D

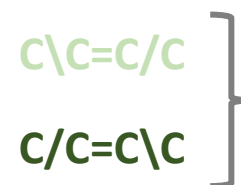
Formato SMILES: *simplified molecular-input line-entry system*

Representação de estruturas químicas através de caracteres ASCII

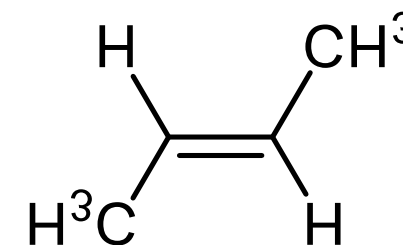
.	Sem ligação formal
-	Ligação simples (sem traço tb.)
=	Ligação dupla
#	Ligação tripla
\$	Ligação quádrupla
:	Ligação aromática (ou letra minúscula)
//	Estereoquímica (trans)
/\	Estereoquímica (cis)
()	Ramificações
C1 C1	Compostos cíclicos
[]	Indicar espécies carregadas e outras



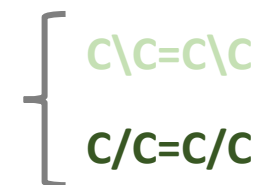
cis-but-2-eno



Não alinhado!



trans-but-2-eno



Alinhado!

equivalentes

babel -i smi filein.smi -o xyz fileout.xyz --gen3D

obabel -i smi filein.smi -o xyz -O fileout.xyz --gen3D

Formato SMILES: *simplified molecular-input line-entry system*

Representação de estruturas químicas através de caracteres ASCII

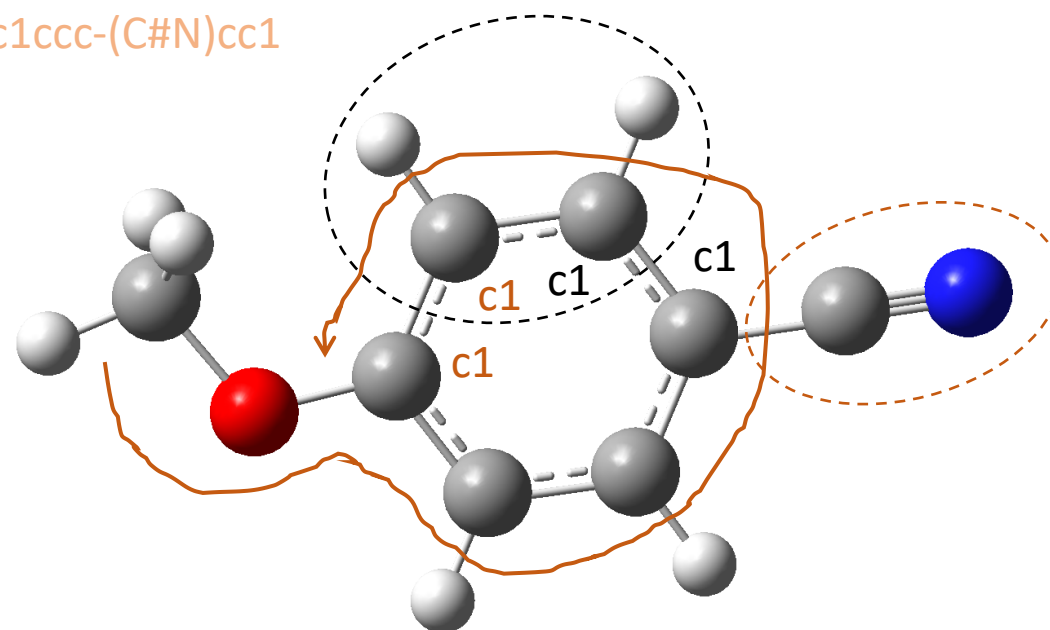
.	Sem ligação formal
-	Ligação simples (sem traço tb.)
=	Ligação dupla
#	Ligação tripla
\$	Ligação quádrupla
:	Ligação aromática (ou letra minúscula)
//	Estereoquímica (trans)
/\	Estereoquímica (cis)
()	Ramificações
C1 C1	Compostos cíclicos
[]	Indicar espécies carregadas e outras

C-O-C:(C:C1):C:C:C1-C#N

C-O-C1:C:C:C-(C#N):C:C1

Forma mais simples:
COc(cc1)ccc1C#N

Forma mais simples:
COc1ccc-(C#N)cc1



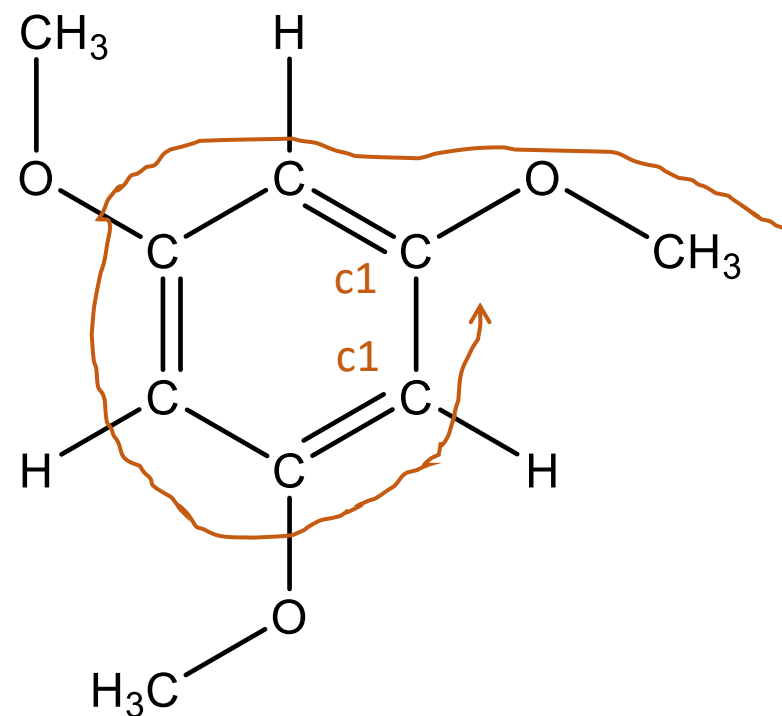
babel -i smi filein.smi -o xyz fileout.xyz --gen3D

obabel -i smi filein.smi -o xyz -O fileout.xyz --gen3D

Formato SMILES: *simplified molecular-input line-entry system*

Representação de estruturas químicas através de caracteres ASCII

.	Sem ligação formal
-	Ligação simples (sem traço tb.)
=	Ligação dupla
#	Ligação tripla
\$	Ligação quádrupla
:	Ligação aromática (ou letra minúscula)
//	Estereoquímica (trans)
/\	Estereoquímica (cis)
()	Ramificações
C1 C1	Compostos cíclicos
[]	Indicar espécies carregadas e outras



C-O-C1:C:C-(O-C):C:C-(O-C):C1

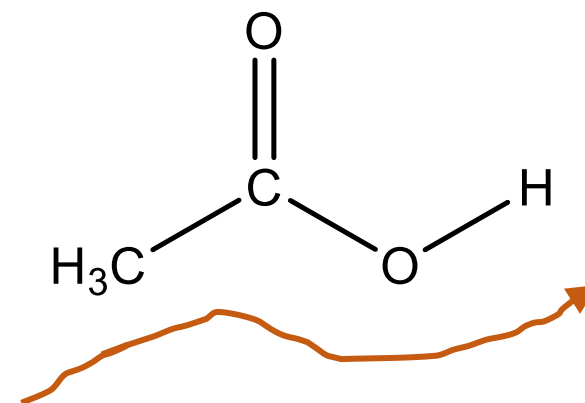
babel -i smi filein.smi -o xyz fileout.xyz --gen3D

obabel -i smi filein.smi -o xyz -O fileout.xyz --gen3D

Formato SMILES: *simplified molecular-input line-entry system*

Representação de estruturas químicas através de caracteres ASCII

.	Sem ligação formal
-	Ligação simples (sem traço tb.)
=	Ligação dupla
#	Ligação tripla
\$	Ligação quádrupla
:	Ligação aromática (ou letra minúscula)
//	Estereoquímica (trans)
/\	Estereoquímica (cis)
()	Ramificações
C1 C1	Compostos cíclicos
[]	Indicar espécies carregadas e outras



Ácido acético

C-C(=O)-O

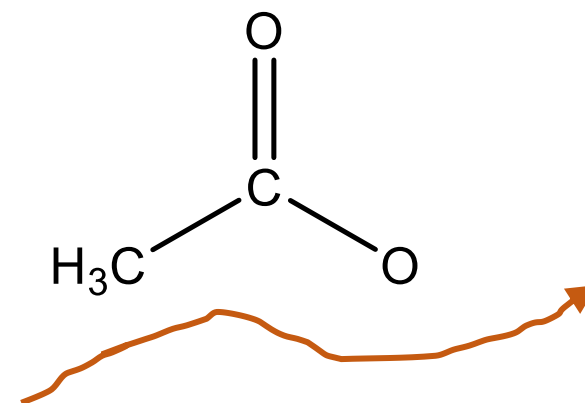
babel -i smi filein.smi -o xyz fileout.xyz --gen3D

obabel -i smi filein.smi -o xyz -O fileout.xyz --gen3D

Formato SMILES: *simplified molecular-input line-entry system*

Representação de estruturas químicas através de caracteres ASCII

.	Sem ligação formal
-	Ligação simples (sem traço tb.)
=	Ligação dupla
#	Ligação tripla
\$	Ligação quádrupla
:	Ligação aromática (ou letra minúscula)
//	Estereoquímica (trans)
/\	Estereoquímica (cis)
()	Ramificações
C1 C1	Compostos cíclicos
[]	Indicar espécies carregadas e outras



Anião acetato

C-C(=O)-[O-]

babel -i smi filein.smi -o xyz fileout.xyz --gen3D

obabel -i smi filein.smi -o xyz -O fileout.xyz --gen3D

Formato SMILES: *simplified molecular-input line-entry system*

Representação de estruturas químicas através de caracteres ASCII

.	Sem ligação formal
-	Ligação simples (sem traço tb.)
=	Ligação dupla
#	Ligação tripla
\$	Ligação quádrupla
:	Ligação aromática (ou letra minúscula)
//	Estereoquímica (trans)
/\	Estereoquímica (cis)
()	Ramificações
C1 C1	Compostos cíclicos
[]	Indicar espécies carregadas e outras

SMI

[NH4+].[Cl-]	< cloreto de amónio
[Fe+3]	< catião ferro (III)
CO	< metanol (H ₃ COH)
C-[O-]	< anião metóxido ([H ₃ CO] ⁻)
C#O	< radical formilo (HCO)
C#[O-]	< anião formilo ([HCO] ⁻)
[C-]#[O+]	< monóxido de carbono (CO)
[C-]#[OH+]	< CO protonado ([COH] ⁺)

babel -i smi filein.smi -o xyz fileout.xyz --gen3D

obabel -i smi filein.smi -o xyz -O fileout.xyz --gen3D

Mecânica Molecular

cálculo da energia potencial

Teorias

Métodos *ab initio*

Baseiam-se nas leis da Mecânica Quântica (outra UC)

- Permitem calcular a energia de um dado sistema a partir das posições nucleares integrando explicitamente o efeito dos eletrões
- Não há necessidade de parametrizações (*ab initio* ou primeiros princípios)
- A qualidade dos dados calculados é dependente das aproximações introduzidas
- Permitem estudar a formação e quebra de ligações químicas e calcular diversas propriedades eletrónicas

Teorias

Mecânica Molecular

Baseia-se nas leis da Mecânica Clássica

Permitem o cálculo da energia a partir das posições nucleares sem inclusão explícita dos elétrons

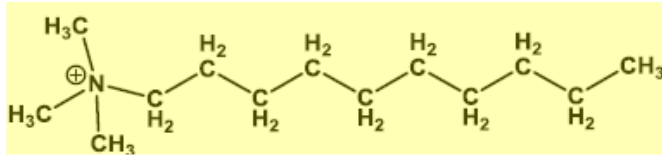
Necessitam de ser parametrizados

- Com dados experimentais
- Com dados calculados através de métodos *ab initio*

Logo, a qualidade dos dados calculados é dependente das parameterizações introduzidas

Modelos e métodos

Quânticos

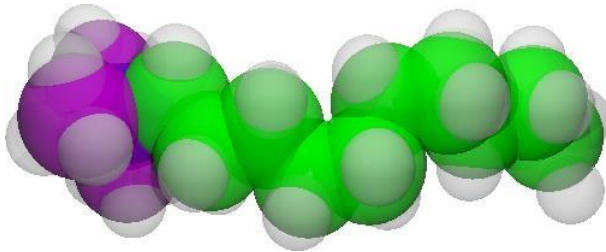


Elétrões + núcleos

159 centros; 43 ligações; ...

MAIS DETALHE
SISTEMAS PEQUENOS

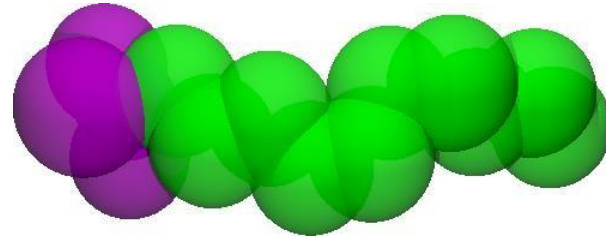
Clássicos



Todos os núcleos

44 centros; 43 ligações; ...

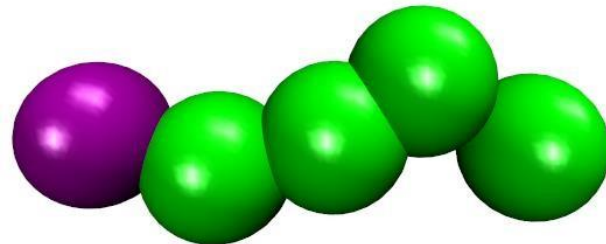
...



Núcleos unidos

14 centros; 13 ligações; ...

...



Granulação grossa

5 centros; 4 ligações; ...

...



Granulação grossa unida

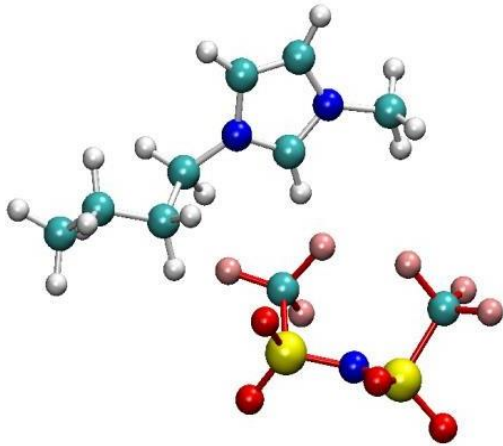
1 centro; 0 ligações; ...

MENOS DETALHE
SISTEMAS GRANDES

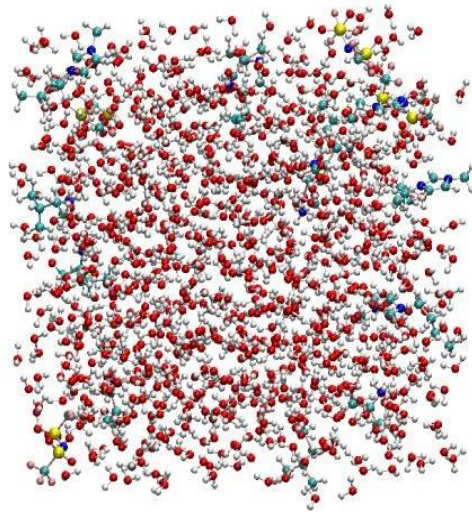
↑
COMPLEXIDADE

Tratabilidade

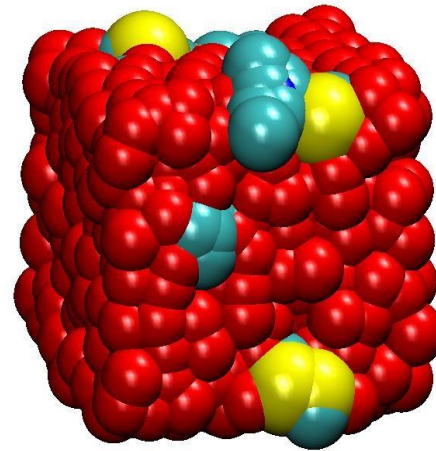
Quântico



Atomístico



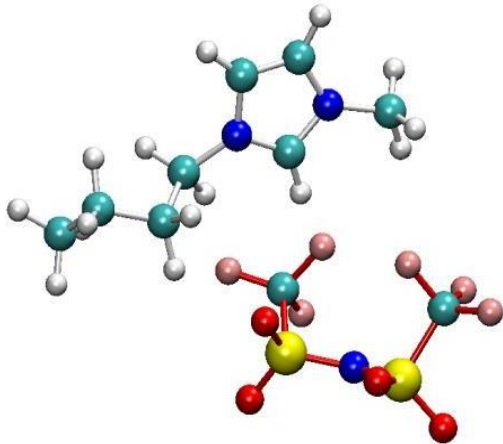
Granulação grossa



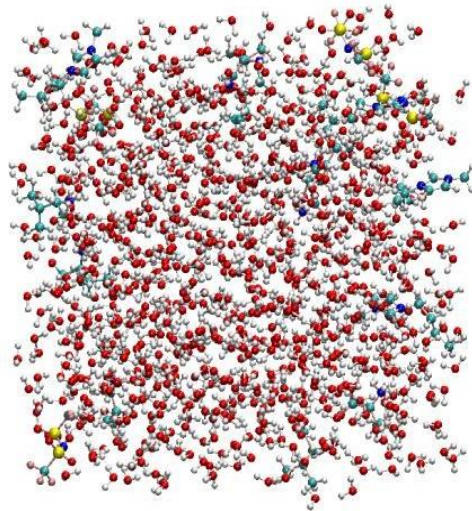
Tamanho do sistema / m

Tratabilidade

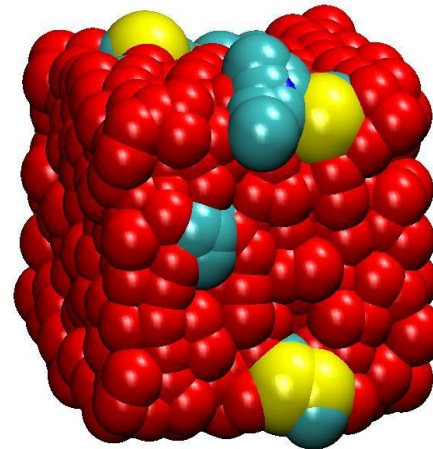
Quântico



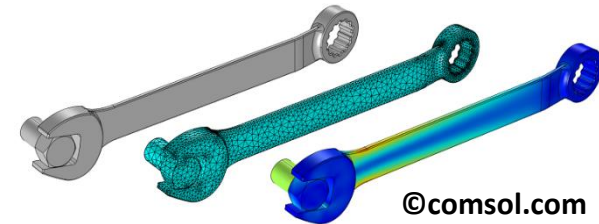
Atomístico



Granulação grossa



(outras teorias)
Métodos dos elementos finitos



...
(escala realística)



Mecânica Molecular



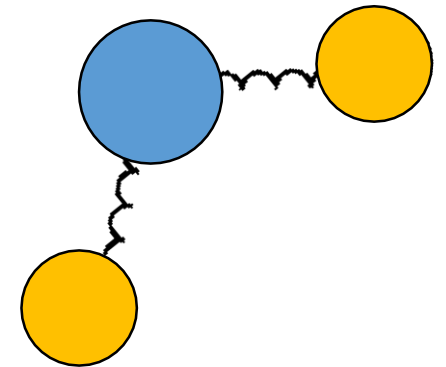
As moléculas são tratadas como *uma coleção de pesos que se ligam através de molas*.



Os *pesos* representam as posições nucleares.

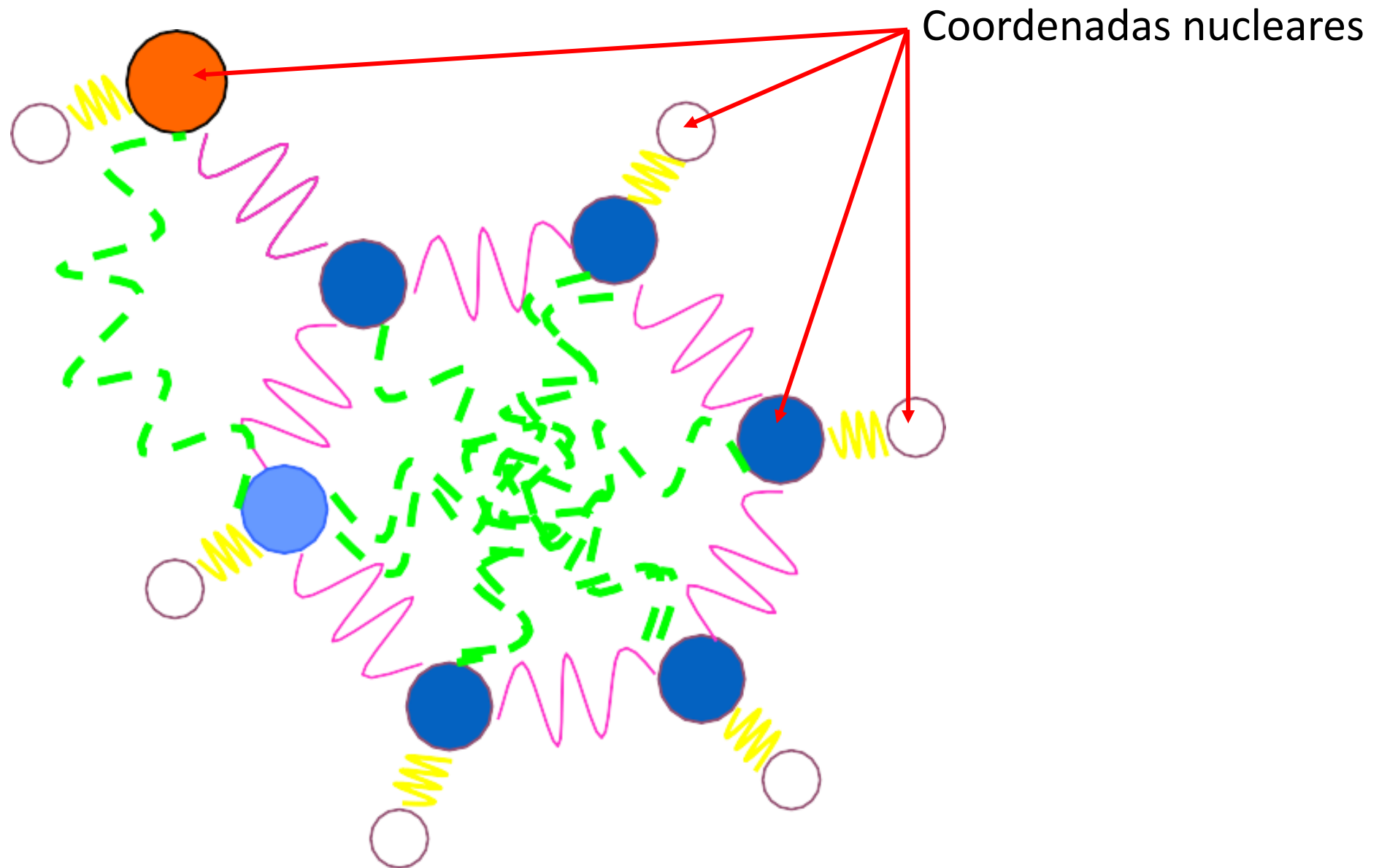


Com uma dada massa e uma dada carga



As *molas* representam ligações entre as posições nucleares.

Mecânica Molecular



Cálculo da energia de um dado sistema

Energia total = Energia cinética + Energia potencial

Cálculo da energia de um dado sistema

Energia total = **Energia cinética** + Energia potencial

$E_c = 0$ para um conjunto fixo de coordenadas

Cálculo da energia de um dado sistema

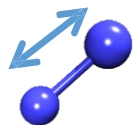
Energia total = **Energia cinética** + Energia potencial

$E_c = 0$ para um conjunto fixo de coordenadas

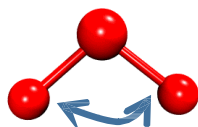
Contribuições para a Energia potencial (V)

Interações Ligantes

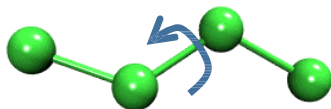
Interações Não Ligantes



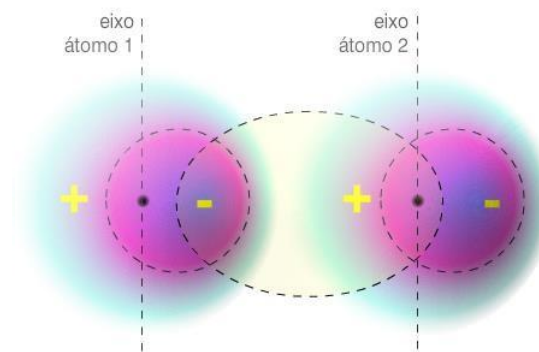
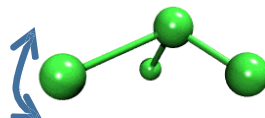
estiramento



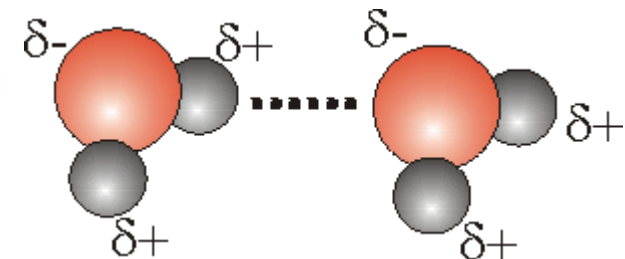
flexão



torsões em torno de ligações ou planos



Ligação de van der Waals
Polarização instantânea pela mudança na nuvem eletrônica



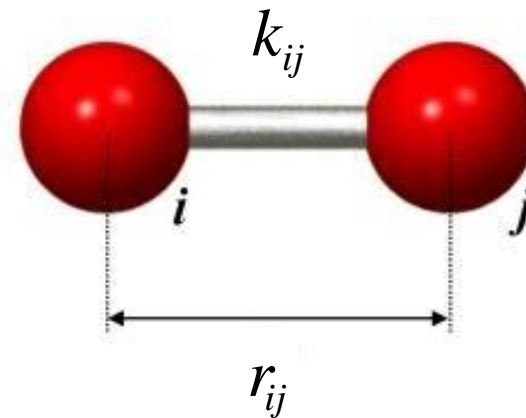
Cálculo da energia de um dado sistema

Contribuições para a Energia potencial (V)

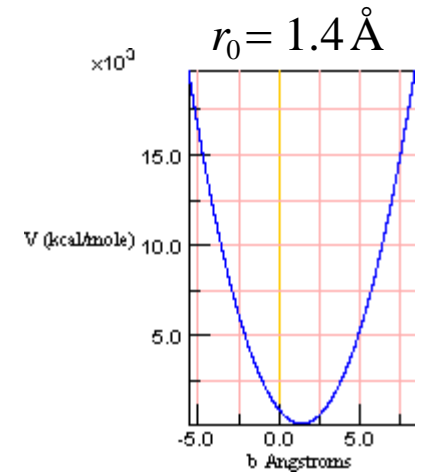
Ligações

$$\sum k_{ij}(r_{ij} - r_0)^2$$

Interações Ligantes



$$k_{ij} = 400 \text{ ((kcal/mol)/\AA^2)}$$

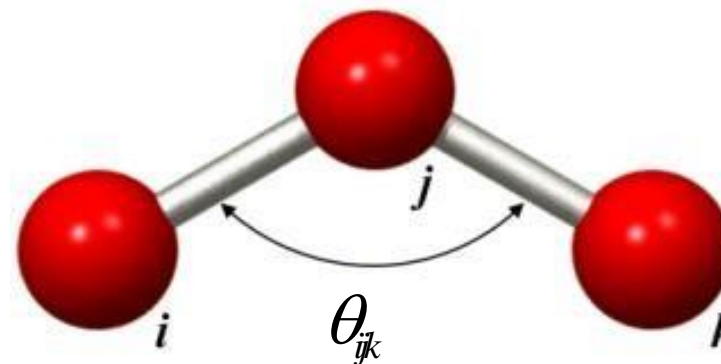


Função harmónica

Ligação	r_0 (Å)	k_{ij} ((kcal/mol)/Å ²)
C-C	1.523	317
C=C	1.337	690

Ângulos

$$\sum k_{ijk}(\theta_{ijk} - \theta_0)^2$$



Outras funções

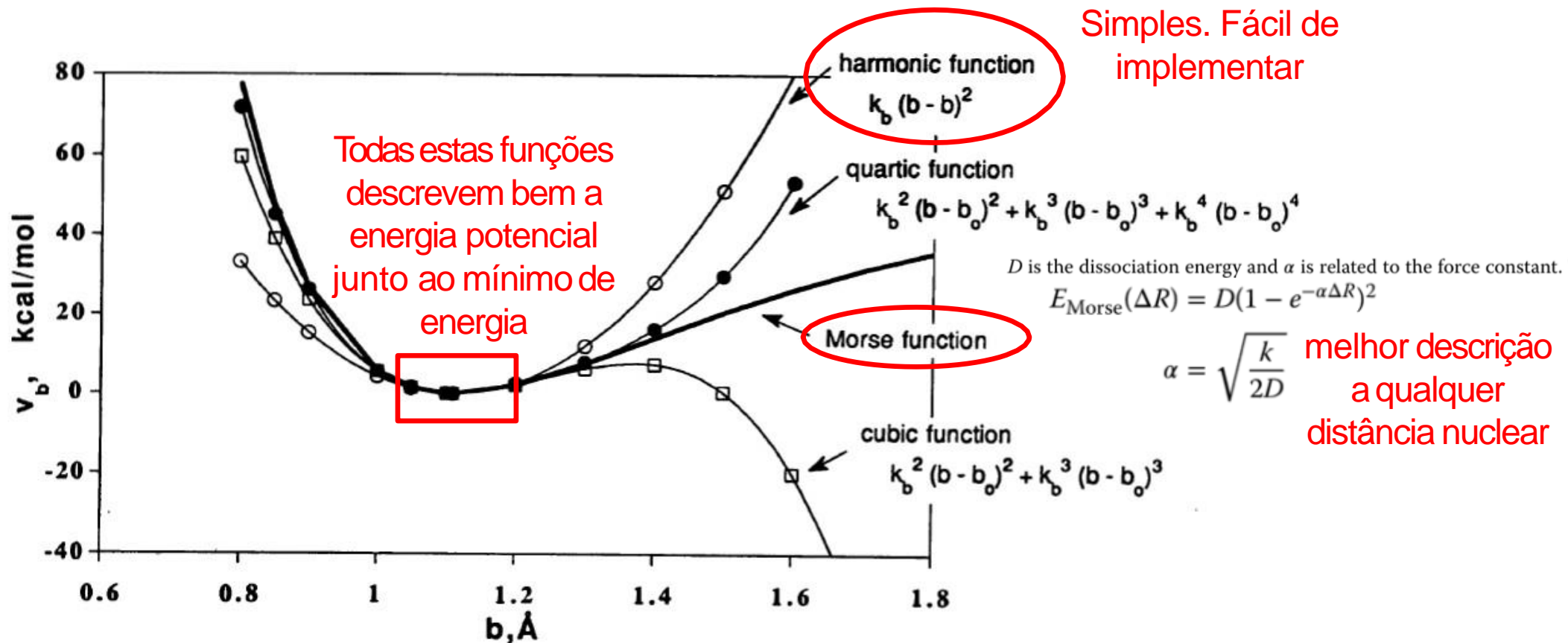


Figure 3 Schematic of a Morse function and the related harmonic, cubic, and quartic potentials (Eqs. [3] and [4]). When the bond length is increased beyond the point of the minimum, the harmonic potential rises too steeply. The cubic term corrects for the anharmonicity locally, but at longer distances turns and goes catastrophically to negative infinity. The quartic potential remains a good approximation over a relatively large range and is always attractive at large distances.

Outras funções (C-H no CH_4)

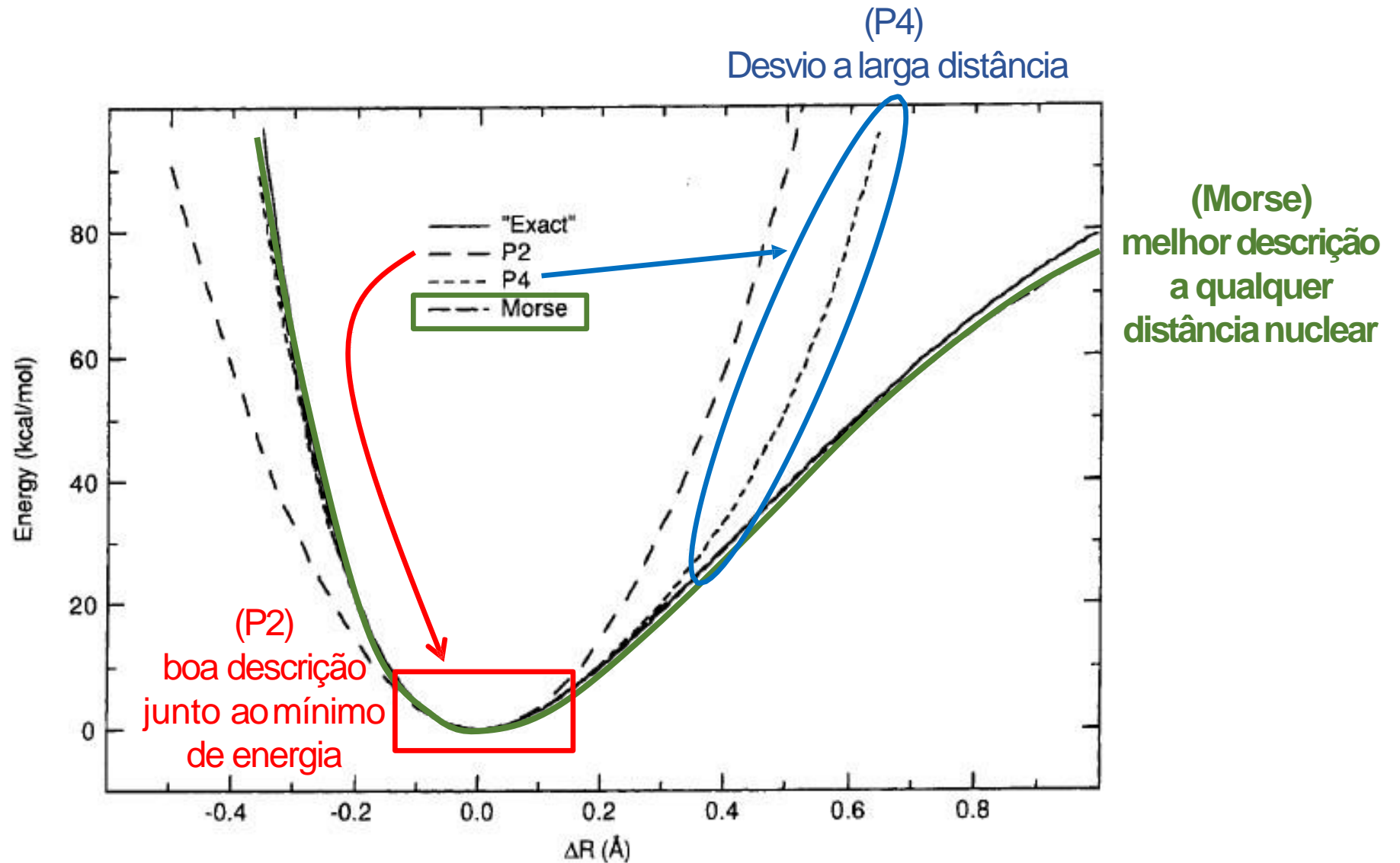


Figure 2.1 The stretch energy for CH_4

Cálculo da energia de um dado sistema

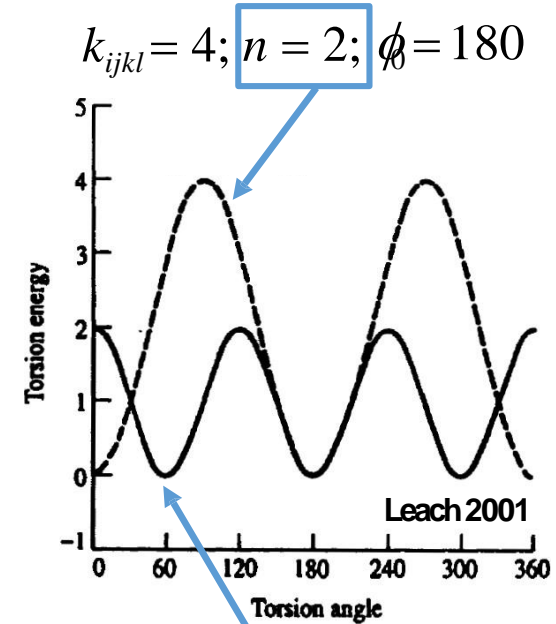
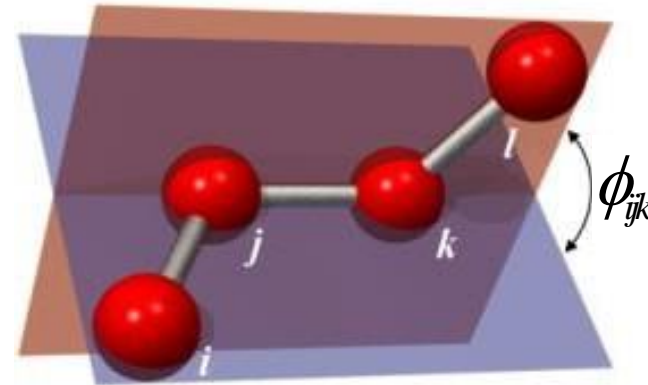
Contribuições para a Energia potencial (V)

Interações Ligantes

Torsões 1-4

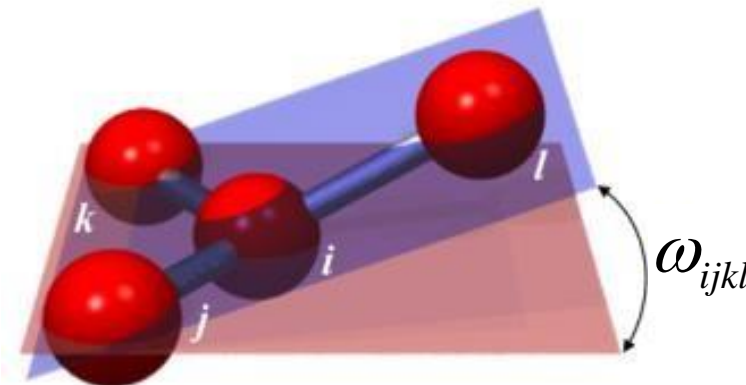
Diedros

$$\sum k_{ijkl}(1 + \cos(n\phi_{ijkl} - \phi_0))$$

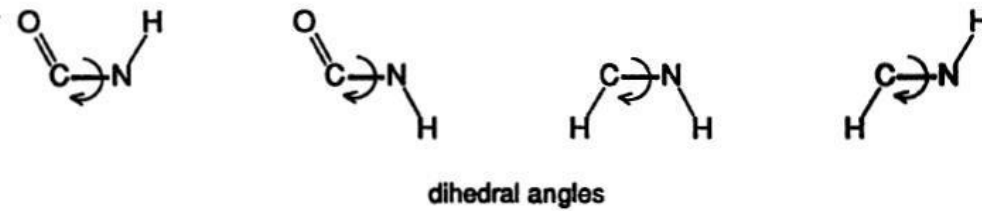
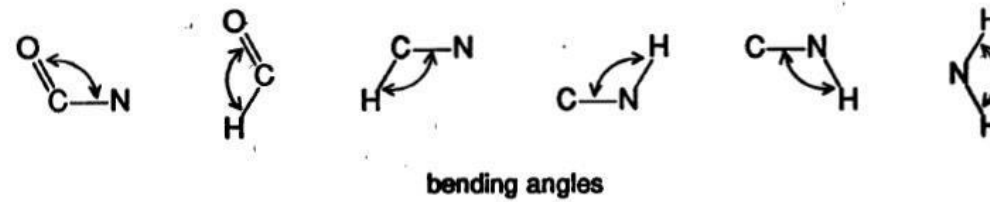
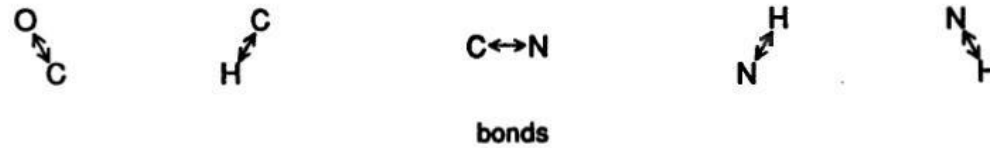
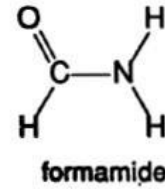


Diedros impróprios

$$\sum k_{ijkl}(\omega_{ijkl} - \omega_0)^2$$



Exemplo: Interações ligantes na formamida



Cálculo da energia de um dado sistema

Contribuições para a Energia potencial (V)

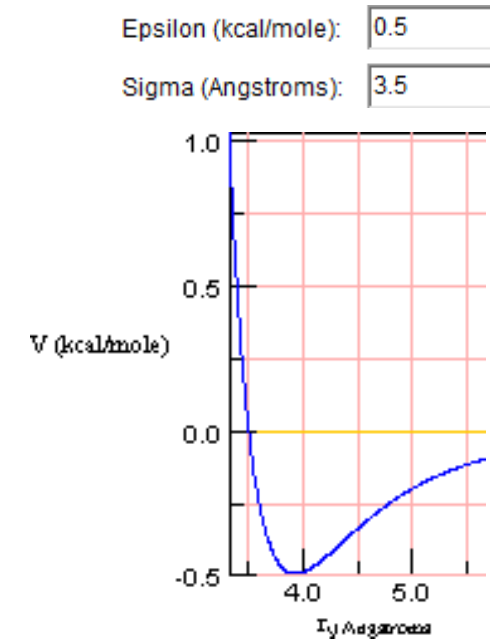
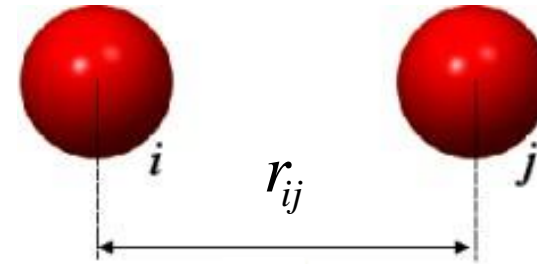
Interações Não Ligantes

van der Waals

Potencial de
Lennard-Jones

$$\sum \sum \epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right]$$

repulsão a curta distância FORTE atração a longa distância FRACA



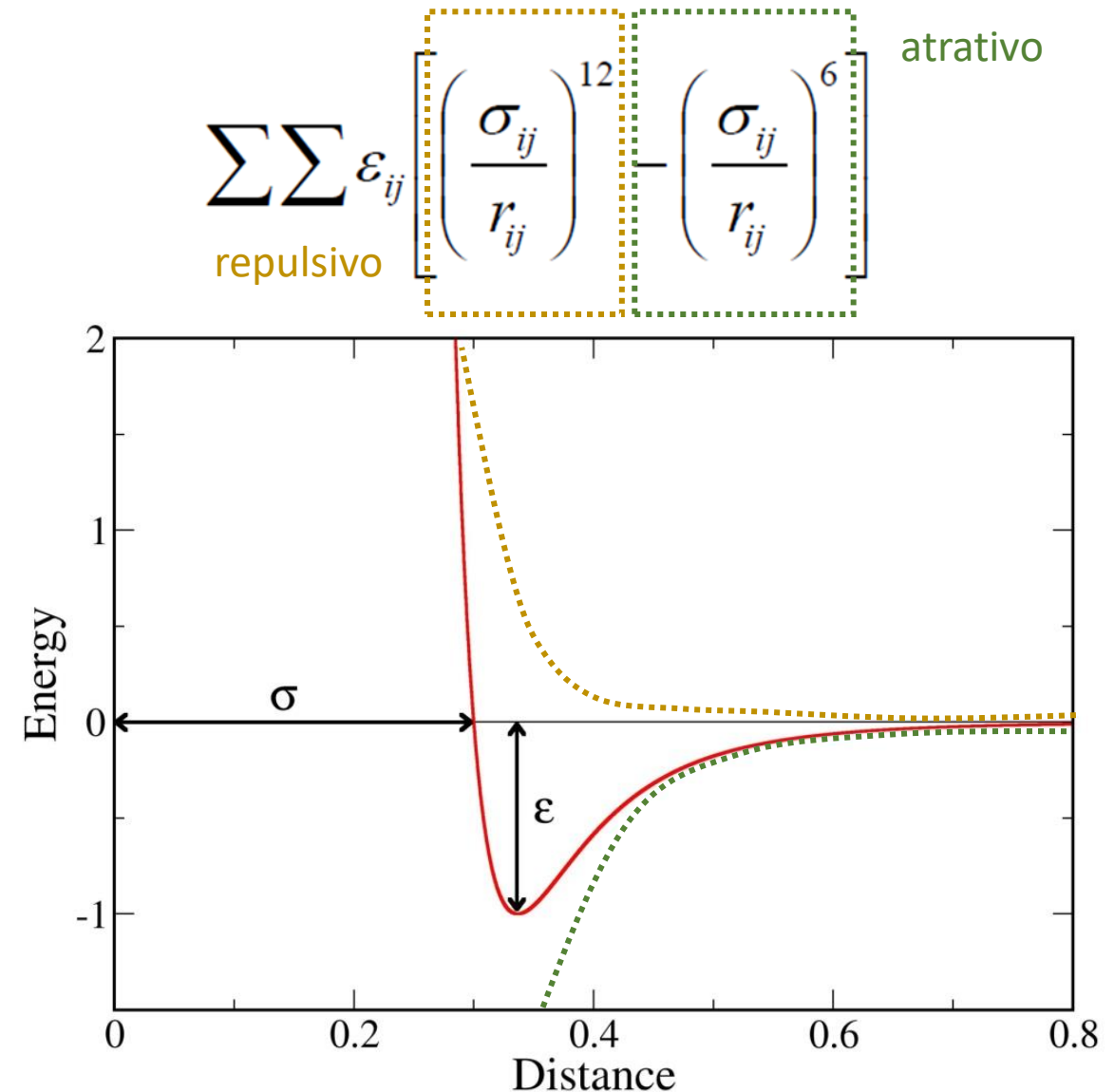
Repulsão e dispersão combinadas – o potencial de Lennard-Jones

Os parâmetros de Lennard-Jones são normalmente estimados com base em propriedades de átomos isolados (por exemplo, tamanho do carbono para σ_{CC})

Mas para interações entre dois átomos diferentes, por exemplo, entre O e H no etanol, precisamos de σ_{OH} . Nestes casos, usam-se as denominações **regras de combinação**, a partir dos parâmetros derivados a partir de átomos isolados. As regras mais comuns são as de **Lorentz-Berthelot**:

$$\sigma_{ij} = \frac{\sigma_{ii} + \sigma_{jj}}{2} \quad \epsilon_{ij} = \sqrt{\epsilon_{ii} \times \epsilon_{jj}}$$

As regras de combinação são empíricas mas ajudam a reduzir o número de parâmetros usados.



Cálculo da energia de um dado sistema

Contribuições para a Energia potencial (V)

Interações Não Ligantes

van der Waals

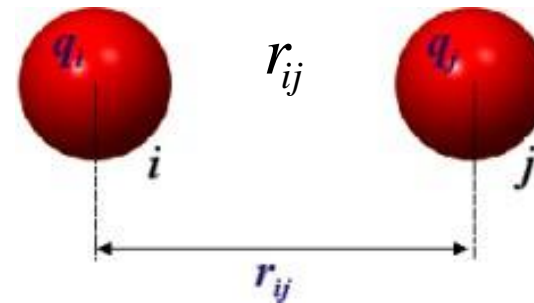
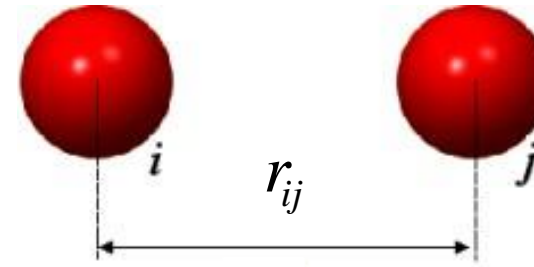
Potencial de
Lennard-Jones

$$\sum \sum \varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right]$$

repulsão a curta distância FORTE atração a longa distância FRACA

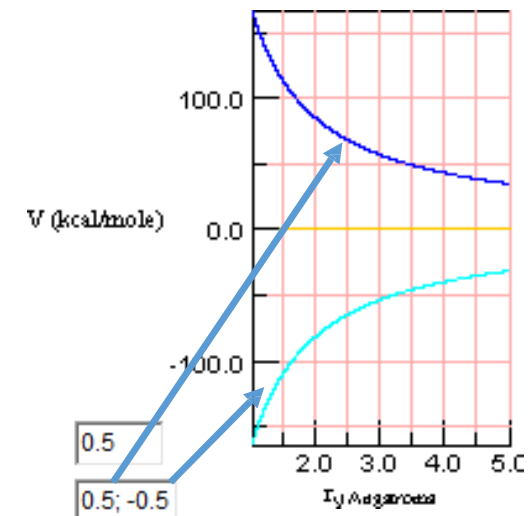
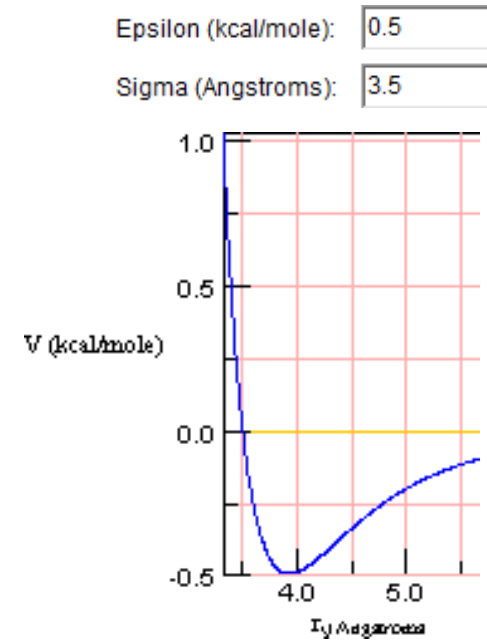
Eletrostáticas

$$\sum \sum \frac{q_i q_j}{r_{ij}}$$



q_i (esu):

q_j (esu)



Cálculo da energia de um dado sistema

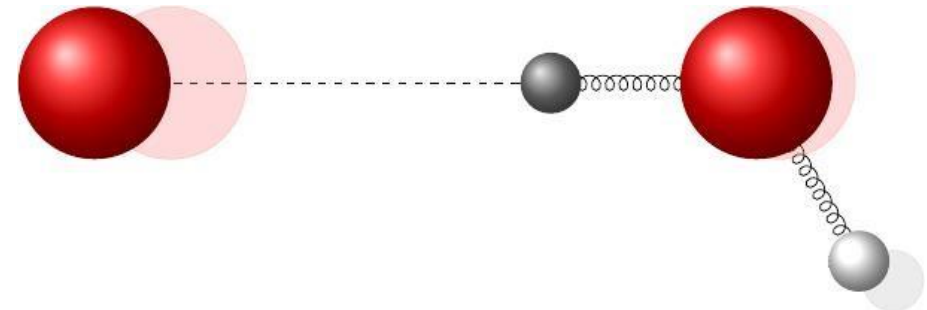
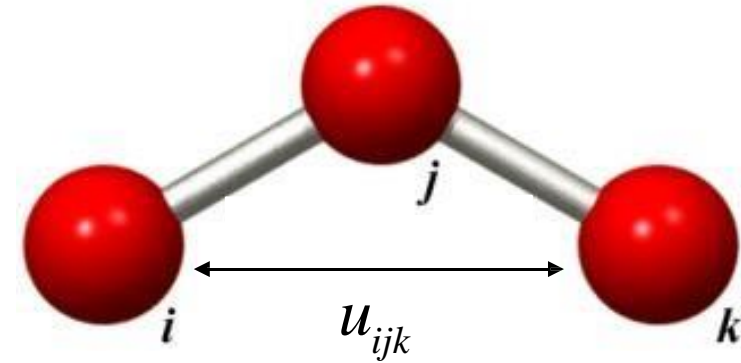
Contribuições para a Energia potencial (V)

Interações Ligantes / Não Ligantes

Outras

Urey-Bradley 1-3 $\sum k_{ik} (u_{ik} - u_0)^2$

importante, por exemplo, no caso de alcanos cíclicos, pois os ângulos dos carbonos sp^3 nestes compostos são mais fechados que em alcanos lineares



Ligações de hidrogénio

Cálculo da energia de um dado sistema

Contribuições para a Energia potencial (V) = Interações Ligantes + Interações Não Ligantes

$$V_{tot,R} = V_r + V_\theta + V_\phi + V_\omega + V_{vdW} + V_{elec.} + V_{...} + \dots$$

Este conjunto de equações denomina-se **CAMPO DE FORÇAS**

Parametrização geralmente realizada com o auxílio de cálculos *ab initio* (geometrias, cargas pontuais, etc.)
No caso dos parâmetros vdW usa-se muitas vezes informação experimental obtida para líquidos puros (densidade, etc.)

Campos de forças (dependem da parametrização e dos termos incluídos)

MM+	apropriado para o cálculo de propriedades termodinâmicas de pequenas moléculas apolares ou com baixa polaridade. Inclui o termo de Urey-Bradley e, ao contrário dos abaixo indicados, pode usar dipolos no lugar de cargas pontuais para o cálculo das interações eletrostáticas
OPLS	apropriado para reproduzir as propriedades de líquidos e de macromoléculas. Aconselhado para estudar moléculas em solução onde as interações não ligantes assumam um papel muito relevante
AMBER	apropriado para cálculos envolvendo macromoléculas (proteínas e ác. nucleicos)
UFF	Interessante pela sua versatilidade pois cobre todos os átomos da Tabela Periódica dos Elementos

Prós e contras dos métodos de Mecânica Molecular

Prós:

1. As funções para a obtenção da energia potencial são simples e de rápida resolução;
2. Podem ser tratadas moléculas ou sistemas moleculares de grandes dimensões;
3. Originam bons resultados para sistemas semelhantes aos que foram utilizados na parametrização do campo de forças;
4. A conectividade entre átomos é preservada pois não há quebra nem formação de ligações;
5. Tratam razoavelmente as interações do tipo van der Waals (nem sempre acontece com métodos baseados nas leis da mecânica quântica);
6. Permitem o cálculo de frequências vibracionais e previsões de desvios de RMN;
7. Adequados para explorar o espaço conformacional de moléculas grandes de forma rápida;
8. Ideais para estudos de dinâmica molecular (ver adiante).

Contras:

1. Não há um campo de forças único e nem sempre é fácil decidir qual o mais apropriado;
2. O número de parâmetros aumenta rapidamente com o número de elementos químicos tratados/parametrizados;
3. É necessário saber qual o tipo de átomo e de ligação para todos os elementos numa molécula (e.g. C_{sp} ; C_{sp^2} ; C_{sp^3} ; etc.);
4. Não permite o estudo de ligações ou interações pouco usuais (ausência de parametrizações);
5. Podem originar “falsos mínimos”;
6. Não podem ser usados para o estudo de formação e quebra de ligações (não é possível localizar estruturas referentes a estados de transição, logo não é possível o estudo de reações químicas).

AMBER General Force Field for organic molecules (Version 1.4, March 2010)

c 12.01	0.616	Sp2 C carbonyl group
c1 12.01	0.360	Sp C
c2 12.01	0.360	Sp2 C
c3 12.01	0.878	Sp3 C
...		
h1 1.008	0.135	H bonded to aliphatic carbon with 1 electrwd. group
h2 1.008	0.135	H bonded to aliphatic carbon with 2 electrwd. group
...		
ho 1.008	0.135	Hydroxyl group
...		
o 16.00	0.434	Oxygen with one connected atom
oh 16.00	0.465	Oxygen in hydroxyl group
os 16.00	0.465	Ether and ester oxygen
ow 16.00	0.465	Oxygen in water
...		
...		

Ver conteúdo do ficheiro



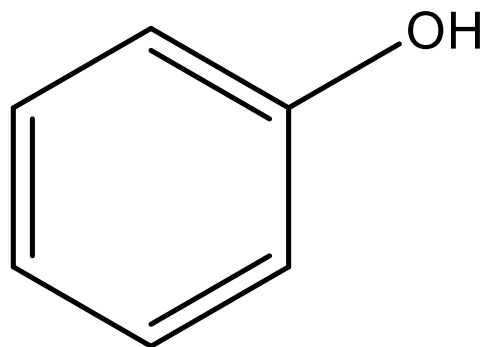
gaff.dat

Calcular energia com open babel (obenergy)

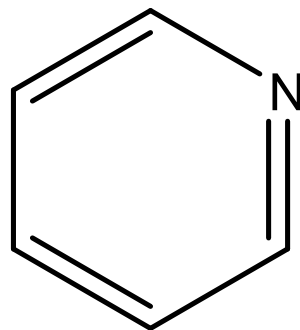
obenergy -ff *"forcefield"* **molecula.xyz**

O-C1:C:C:C:C:C1

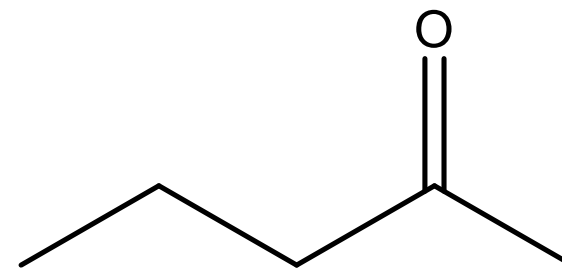
babel -i smi fenol.smi -o xyz fenol.xyz --gen3D



fenol



piridina



2-pentanona

Campos de força (*force fields*) disponíveis

GAFF	General Amber Force Field.
Ghemical	Ghemical force field.
MMFF94	MMFF94 force field.
MMFF94s	MMFF94s force field.
UFF	Universal Force Field.

Para informação adicional:

<https://openbabel.org/wiki/Obenergy>

Usage: *obenergy* [*options*] <filename>