



# Universidade Federal da Paraíba

Departamento de Química

Laboratório de Química Quântica Computacional



*Grupo de Química Computacional mais Oriental das Américas*  
L.Q.Q.C. - UFPB - João Pessoa - PB

## Mini-Curso:

# Métodos Semi-empíricos de Química Quântica

Prof. Gerd Bruno Rocha

[gbr@quimica.ufpb.br](mailto:gbr@quimica.ufpb.br) ; [gerd.rocha@gmail.com](mailto:gerd.rocha@gmail.com)

[www.quantum-chem.pro.br](http://www.quantum-chem.pro.br)



# Programa do curso.

## 1º dia:

- 1 - Aspectos teóricos dos Métodos Semi-empíricos
- 2 - Apresentação dos programas utilizados
  - 2.1 – Mopac 2012
  - 2.2 – Gabedit
- 3 – Práticas:
  - 3.1 – Reação isodésmica
  - 3.2 – Análise conformacional em ciclohexanos substituídos

# Programa do curso.

## 2º dia:

- 1 – Noções Básicas do método MOZYME.
- 2 – Práticas: Cálculo de propriedades de proteínas
  - 2.1 – Encontrando estruturas nativas de proteínas a partir do cálculo de  $\Delta H$  de formação

## **1º dia:**

- 1 - Aspectos teóricos dos Métodos Semi-empíricos
- 2 - Apresentação dos programas utilizados
  - 2.2 – Mopac 2012
  - 2.3 – Gabedit
- 3 - Práticas:
  - 3.1 – Reação isodésmica
  - 3.2 – Análise conformacional em ciclohexanos substituídos

# Métodos de química computacional

Métodos de química quântica tentam encontrar soluções da:

## Equação de Schrödinger

$$H\Psi = E\Psi$$

**Exatidão do resultado**

• Primeiros princípios

• Semi-empíricos

• Mecânica molecular

Tentam se aproximar da  
solução da equação de  
Schrödinger

Método quântico que  
usa aproximações  
empíricas

Método puramente  
empírico, não trata os  
elétrons.

**Número de átomos do sistema**

# Métodos quânticos de química computacional

## Equação de Schrödinger

$$H\Psi = E\Psi$$



**Hartree-Fock e pós-HF:**  
aproximação na interação  
eletrônica



**Funcional da densidade:**  
aproximação na  $E_{xc}$



**Métodos semi-empíricos:**  
aproximação nas integrais

# Equações de Hartree-Fock-Roothaan

Função de onda molecular → Determinante de Slater

$$\Psi = (2n!)^{-\frac{1}{2}} \begin{vmatrix} \chi_i(x_1) & \chi_j(x_1) & \cdots & \chi_k(x_1) \\ \chi_i(x_2) & \chi_j(x_2) & \cdots & \chi_k(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_i(x_{2n}) & \chi_j(x_{2n}) & \cdots & \chi_k(x_{2n}) \end{vmatrix}$$

Onde  $\chi(x)$  é o spin-orbital formado por  $\psi(r)\alpha(w)$  ou  $\psi(r)\beta(w)$ .

Forma mais compacta em função do operador de permutação  $\hat{P}$ .

$$\Psi = (2n!)^{-\frac{1}{2}} \sum_{\hat{P}} (-1)^p \hat{P} \{ \psi_1(1)\alpha(1)\psi_1(2)\beta(2) \cdots \psi_n(2n)\beta(2n) \}$$

**Molécula com 2n elétrons**

O operador de permutação só atua nas coordenadas dos elétrons, exemplo:

$$\hat{P}_{3421} \{ \psi_1(1)\alpha(1)\psi_1(2)\beta(2)\psi_2(3)\alpha(3)\psi_2(4)\beta(4) \} = \\ \psi_1(3)\alpha(3)\psi_1(4)\beta(4)\psi_2(2)\alpha(2)\psi_2(1)\beta(1)$$

1234 → 2134 → 2314 → 2341 → 3241 → 3421 = 5 permutações.

# Energia de um determinante de Slater

Energia de um determinante de Slater  $\longrightarrow \langle \Psi | \hat{H} | \Psi \rangle$  **Valor esperado**

$$\hat{H} = \hat{H}_1 + \hat{H}_2$$

$$\hat{H}_1 = \sum_i \hat{H}^{core}(i) = \sum_i \left[ -\frac{1}{2} \nabla_i^2 - \sum_A \frac{Z_A}{\bar{r}_{iA}} \right] \longrightarrow \text{um-elétron}$$

$$\hat{H}_2 = \sum_i \sum_{j>i} \frac{1}{r_{ij}} \longrightarrow \text{dois-elétron}$$

**Substituindo:**

$$\langle \Psi | \hat{H} | \Psi \rangle = \langle \Psi | \hat{H}_1 | \Psi \rangle + \langle \Psi | \hat{H}_2 | \Psi \rangle$$



# Energia de um determinante de Slater

## Valor esperado da energia

$$E = \langle \Psi | \hat{H} | \Psi \rangle = 2 \sum_i^n h_{ii} + \sum_i^n J_{ii} + \sum_i^n \sum_{j \neq i}^n (2J_{ij} - K_{ij}) = 2 \sum_i^n h_{ii} + \sum_i^n \sum_j^n (2J_{ij} - K_{ij})$$

Onde:

$$h_{ii} = \int \psi_i(1) \hat{H}^{core}(1) \psi_i(1) d\tau_1 \quad \text{Integral de energia cinética e potencial elétron-núcleo}$$

$$J_{ij} = \iint \psi_i(1) \psi_j(2) \frac{1}{r_{12}} \psi_i(1) \psi_j(2) d\tau_1 d\tau_2 \quad \text{Integral de Coulomb}$$

$$K_{ij} = \iint \psi_i(1) \psi_j(2) \frac{1}{r_{12}} \psi_j(1) \psi_i(2) d\tau_1 d\tau_2 \quad \text{Integral de troca}$$

# Equação de Hartree-Fock-Roothaan

Primeiro ponto: vamos usar o procedimento MO-LCAO em cada um dos operadores

$$\psi_i = \sum_{\mu} c_{\mu}^{(i)} \phi_{\mu}$$

$$h_{ii} = \int \psi_i(1) \hat{H}^{core}(1) \psi_i(1) d\tau_1 = \sum_{\mu} \sum_{\nu} c_{\mu}^{(i)} c_{\nu}^{(i)} \int \phi_{\mu}(1) \hat{H}^{core}(1) \phi_{\nu}(1) d\tau_1 = \sum_{\mu} \sum_{\nu} c_{\mu}^{(i)} c_{\nu}^{(i)} h_{\mu\nu}$$

$$J_{ij} = \iint \psi_i(1) \psi_j(2) \frac{1}{r_{12}} \psi_i(1) \psi_j(2) d\tau_1 d\tau_2 =$$

$$\sum_{\mu} \sum_{\nu} \sum_{\lambda} \sum_{\sigma} c_{\mu}^{(i)} c_{\nu}^{(j)} c_{\lambda}^{(i)} c_{\sigma}^{(j)} \iint \phi_{\mu}(1) \phi_{\nu}(2) \frac{1}{r_{12}} \phi_{\lambda}(1) \phi_{\sigma}(2) d\tau_1 d\tau_2 =$$

$$\sum_{\mu} \sum_{\nu} \sum_{\lambda} \sum_{\sigma} c_{\mu}^{(i)} c_{\nu}^{(j)} c_{\lambda}^{(i)} c_{\sigma}^{(j)} (\mu\nu | \lambda\sigma)$$

$$K_{ij} = \iint \psi_i(1) \psi_j(2) \frac{1}{r_{12}} \psi_j(1) \psi_i(2) d\tau_1 d\tau_2 =$$

$$\sum_{\mu} \sum_{\nu} \sum_{\lambda} \sum_{\sigma} c_{\mu}^{(i)} c_{\nu}^{(j)} c_{\lambda}^{(j)} c_{\sigma}^{(i)} \iint \phi_{\mu}(1) \phi_{\nu}(2) \frac{1}{r_{12}} \phi_{\lambda}(1) \phi_{\sigma}(2) d\tau_1 d\tau_2 =$$

$$\sum_{\mu} \sum_{\nu} \sum_{\lambda} \sum_{\sigma} c_{\mu}^{(i)} c_{\nu}^{(j)} c_{\lambda}^{(i)} c_{\sigma}^{(j)} (\mu\sigma | \nu\lambda)$$

# Equação de Hartree-Fock-Roothaan

A expressão da energia será:

$$E = 2 \sum_i^n \sum_\mu \sum_\nu c_\mu^{(i)} c_\nu^{(i)} h_{\mu\nu} + \sum_i^n \sum_j^n \sum_{\mu\nu\lambda\sigma} c_\mu^{(i)} c_\nu^{(j)} c_\lambda^{(i)} c_\sigma^{(j)} (2(\mu\nu | \lambda\sigma) - (\mu\sigma | \nu\lambda))$$

Vamos definir a quantidade:

$$P_{\mu\nu} = 2 \sum_i^n c_\mu^{(i)} c_\nu^{(i)} \quad \Rightarrow \quad \text{Matriz densidade}$$

Teremos assim:

$$E = \sum_{\mu\nu} P_{\mu\nu} h_{\mu\nu} + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} P_{\mu\nu} P_{\lambda\sigma} \left( (\mu\nu | \lambda\sigma) - \frac{1}{2} (\mu\sigma | \nu\lambda) \right)$$

Equações para todos os elétrons:

$$\sum_\nu (F_{\mu\nu} - \varepsilon_i S_{\mu\nu}) c_\nu^{(i)} = 0 \quad \therefore \mu = 1, 2, 3, \dots$$

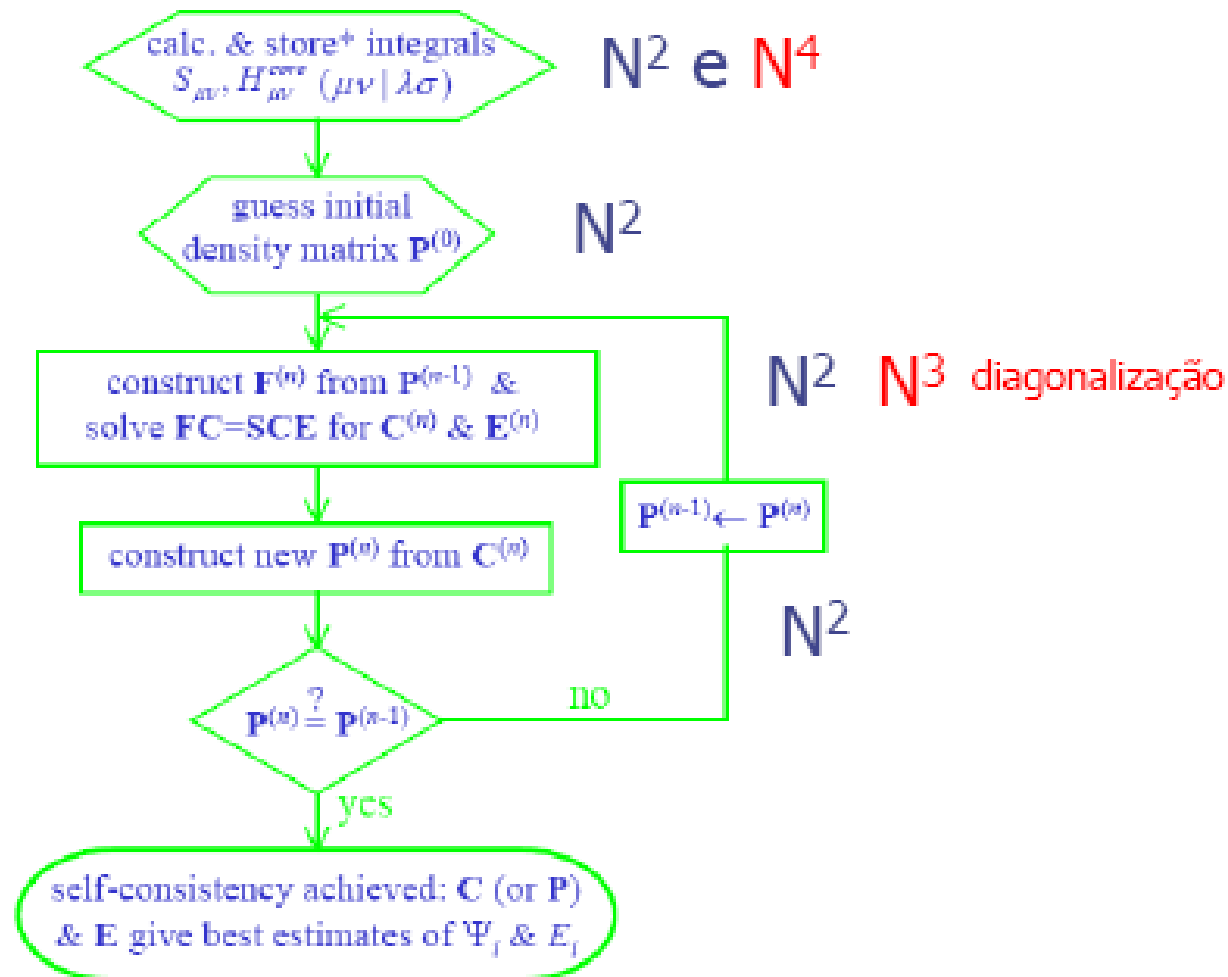
$$F_{\mu\nu} = h_{\mu\nu} + \sum_{\lambda\sigma} P_{\lambda\sigma} \left( (\mu\nu | \lambda\sigma) - \frac{1}{2} (\mu\sigma | \nu\lambda) \right)$$

**Forma matricial**

$$\mathbf{FC} = \mathbf{S} \mathbf{C} \boldsymbol{\varepsilon}$$

Equação de autovalores generalizada **11**

# Procedimento SCF



# Teoria semi-empírica

# Função de onda semi-empírica

O primeiro passo na direção de construir as equações dos método semi-empíricos é saber que nestes métodos a configuração eletrônica pode ser dividida em duas partes:

- os elétrons internos (*core*)

- os elétrons de valência.  $|\Psi^s\rangle \longrightarrow$

é escrita como um produto anti-simetrizado da parte que descreve os elétrons internos e da parte que descreve os elétrons de valência.

A justificativa para isso vem do fato de que a parte da densidade eletrônica que é mais afetada nas reações químicas vem primordialmente dos elétrons de valência (externos).

$$H_{\mu\nu}^s = \langle \phi_\mu | \left[ -\frac{1}{2} \nabla^2 - \sum_A^M \frac{Z'_A}{|\vec{R}_A - \vec{r}_i|} \right] | \phi_\nu \rangle \longrightarrow \sum_\nu^{n_{val}} F_{\mu\nu} c_{\nu i} = \sum_\nu^{n_{val}} S_{\mu\nu} c_{\nu i} \epsilon_i$$

# Aproximação ZDO (*Zero Differential Overlap*)

## *Zero differential overlap - ZDO*

$$S_{\mu_A \nu_B} = \left\langle \varphi_{\mu_A} \mid \varphi_{\nu_B} \right\rangle = \delta_{\mu \nu} \delta_{AB}$$

1. O resultado direto desta aproximação é fazer com que a matriz de *overlap* seja, agora, uma matriz identidade e, por consequência, a equação de autovalores de Hartree-Fock-Roothaan é reduzida a uma forma mais simplificada e diretamente resolvida a partir de uma simples diagonalização da matriz de Fock;
2. Outra consequência direta da aproximação ZDO é fazer com que todas as integrais de 2-elétrons que dependem do *overlap* de densidades de cargas de diferentes orbitais atômicos sejam anuladas;

$$(\mu \nu \mid \lambda \sigma) = (\mu \mu \mid \lambda \lambda) \delta_{\mu \nu} \delta_{\lambda \sigma} \longrightarrow$$

Assim, reduzem-se estas quantidades para no máximo integrais de 2-elétrons e 2-centros.

# Tipos de Aproximações

## *Zero differential overlap - ZDO*

$$\int \varphi_{\mu_A}^* (1) \varphi_{\nu_B} (1) d\tau_1 = 0 \quad \mu_A \neq \nu_B$$

## Aproximações integrais

### CNDO

$$(\mu_A \nu_B | \lambda_C \sigma_D) = (\mu_A \mu_A | \lambda_C \lambda_C) \delta_{\mu\nu} \delta_{\lambda\sigma} \delta_{AB} \delta_{CD}$$

### INDO

$$(\mu_A \nu_B | \lambda_C \sigma_D) = (\mu_A \nu_A | \lambda_C \sigma_C) \delta_{\mu_A \nu_B} \delta_{\lambda_C \sigma_D}$$

### NDDO

$$(\mu_A \nu_B | \lambda_C \sigma_D) = (\mu_A \nu_A | \lambda_C \sigma_C) \delta_{AB} \delta_{CD}$$

Introdução de parâmetros:



Ajustados



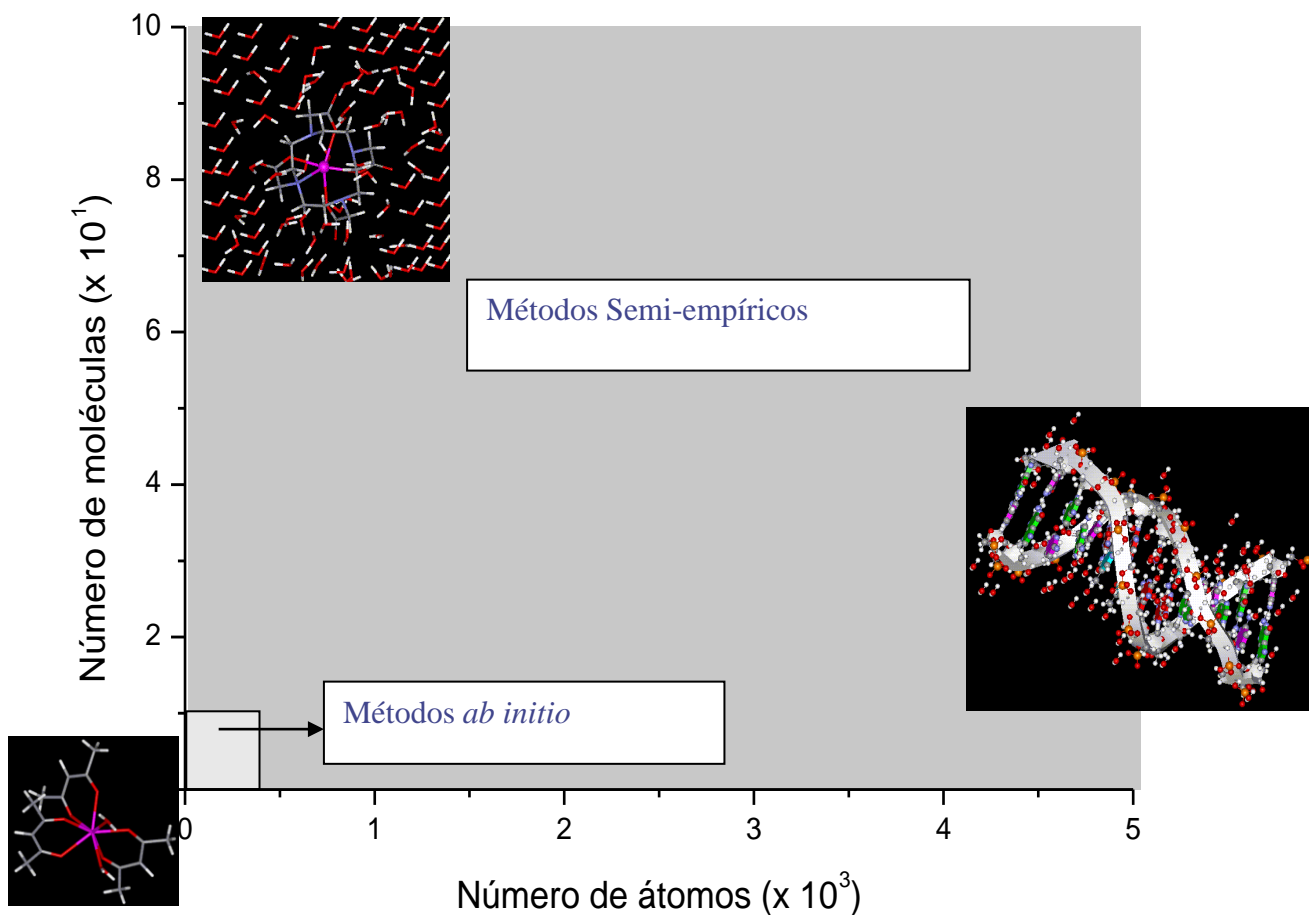
Tirados do experimento



# Semi-Empíricos

- Vantagens:
  - Para a maioria dos métodos os parâmetros são atômicos. Assim que um átomo é parametrizado, em princípio, pode ser usado em qualquer ambiente.
  - Todos os tipos de fenômenos, normalmente de interesse, podem ser modelados: reações químicas, polarizabilidades, espectros, propriedades ópticas, propriedades eletrônicas, etc.
- Limitações:
  - Exatidão em geral insuficiente.
  - Procedimento de parametrização é difícil
  - A grande maioria dos métodos não possui parâmetros para todos os átomos da tabela periódica.

# Aplicação dos Métodos Semi-empíricos



# Procedimento de Parametrização

**Etapas decisivas para o sucesso de um método semi-empírico**

- Escolha da função resposta;
- Banco de dados de moléculas
- Definição do conjunto de propriedades;
- Escolha das técnicas de parametrização;

# Função resposta e propriedades

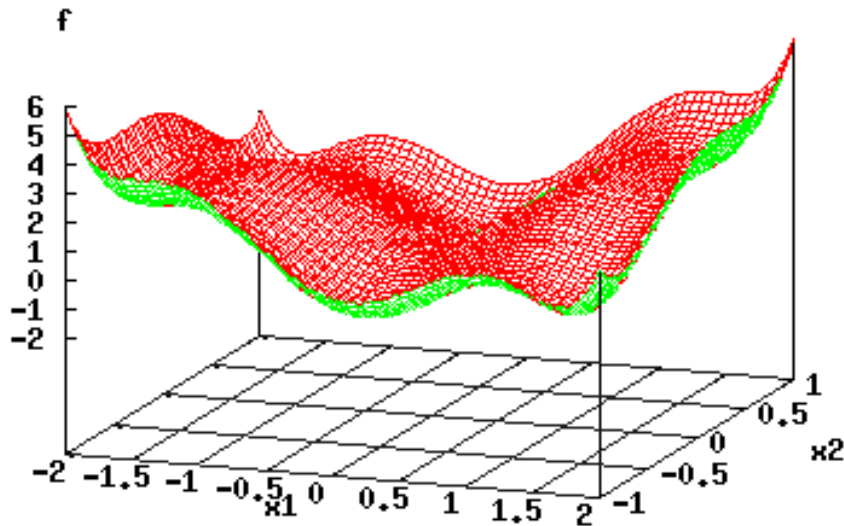
$$F^{resp} = \sum_{i=1}^n \left( X_i^{Calc} - X_i^{Exp} \right)^2 \cdot w_i^2$$

## Conjunto de pesos

Propriedade	Peso
Calor de formação ( $\Delta H_f$ )	1 kcal <sup>-1</sup> mol
Potencial de ionização ( $PI$ )	10 eV <sup>-1</sup>
Momento de dipolo ( $\mu$ )	20 D <sup>-1</sup>
Distância de ligação ( $R_{ab}$ )	100 Å <sup>-1</sup>
Ângulo ( $\theta_l$ )	2/3 grau <sup>-1</sup>
Diedros ( $\theta_d$ )	1/3 grau <sup>-1</sup>

# Otimização não-linear em muitas dimensões

O principal objetivo de usar técnicas de otimização é o de encontrar pontos extremos (máximos ou mínimos) de uma dada função de muitas variáveis.



$$f : S \subset R^n \rightarrow R$$






$$x^* \rightarrow f^* = f(x^*) \leq f(x), \quad \forall x \in S$$

$$f(x'^*) - f^* \leq \varepsilon$$

$$f(x, y) = x^2 \left( 4 - 2.1x^2 + \frac{1}{3}x^4 \right) + xy + y^2 \left( -4 + 4y^2 \right)$$

- **2 mínimos globais**  
**(0.09, -0.71) e (-0.09, 0.71)**
- **4 mínimos locais**

# Parametrização do MNDO

Molecule	Reference functions			$\mu^d$
	$\Delta H_f^a$	Grad <sup>b</sup>	IPC <sup>c</sup>	
H <sub>2</sub>	X	X	X	
CH <sub>4</sub>	X	X	X	
C <sub>2</sub> H <sub>6</sub>	X	X		
C <sub>2</sub> H <sub>4</sub>	X	X	X	
C <sub>2</sub> H <sub>2</sub>	X	X	X	
	X			
	X	X		
	X		X	
C(CH <sub>3</sub> ) <sub>4</sub>	X			
N <sub>2</sub>	X	X	X	
NH <sub>3</sub>	X	X	X	X
CH <sub>3</sub> NH <sub>2</sub>	X	X		X
(CH <sub>3</sub> ) <sub>2</sub> NH	X			X
 NH	X			
N <sub>2</sub> H <sub>4</sub>	X	X		
HCN	X	X	X	X
CH <sub>3</sub> CN	X	X		
(CN) <sub>2</sub>	X			
H <sub>2</sub> O	X	X	X	X
CH <sub>3</sub> OH	X	X	X	X
(CH <sub>3</sub> ) <sub>2</sub> O	X	X	X	X
 O	X			
H <sub>2</sub> O <sub>2</sub>	X	X		
O <sub>3</sub>	X			
CO	X	X	X	X
CO <sub>2</sub>	X	X	X	
H <sub>2</sub> CO	X	X	X	X
(CH <sub>3</sub> ) <sub>2</sub> CO	X			
CH <sub>3</sub> CO	X	X		
(CHO) <sub>2</sub>	X			
HCOOH	X			X
HCOOCH <sub>3</sub>	X			

$$F^{resp} = \sum_{i=1}^n \left( X_i^{Calc} - X_i^{Exp} \right)^2 \cdot w_i^2$$

Banco de dados de propriedades

Pesos

Reference function	Weighting factor <sup>a</sup>
Heat of formation	1 kcal <sup>-1</sup> mol
Ionization potential	10 eV <sup>-1</sup>
Dipole moment	20 D <sup>-1</sup>
Gradient for bond length <sup>b</sup>	4 × 10 <sup>-6</sup> kcal <sup>-1</sup> mol Å
Gradient for bond angle <sup>b</sup>	8 × 10 <sup>-4</sup> kcal <sup>-1</sup> mol deg
Gradient for dihedral angle <sup>b</sup>	2 × 10 <sup>-3</sup> kcal <sup>-1</sup> mol deg
Bond length <sup>c</sup>	100 Å <sup>-1</sup>
Bond angle <sup>c</sup>	(2/3) deg <sup>-1</sup>
Dihedral angle <sup>c</sup>	(1/3) deg <sup>-1</sup>

1. Parâmetros para C e H
2. Parâmetros para N usando o conjunto de C, N e H
3. Parâmetro para O usando o conjunto de C, H e O

# Resultados do MNDO

## Calor de formação

Table I. Mean Absolute Errors  $\Delta(\Delta H_f)$  for Heats of Formation<sup>a</sup>

Class of compd	No.	$\Delta(\Delta H_f)$ , kcal/mol	
		MNDO	MINDO/3
All compounds	138	6.3	11.0
Hydrocarbons	58	6.0	9.7
With aromatic rings	5	1.7	12.1
With triple bonds	6	6.6	13.5
With three-membered rings <sup>b</sup>	8	5.2	11.7
With four-membered rings <sup>b</sup>	7	18.7 <sup>c</sup>	8.6
Acyclic hydrocarbons	26	3.5	6.9
Cyclic hydrocarbons	32	8.0	11.9
Bicyclic hydrocarbons with 5- and 6-membered rings	5	2.9	22.3
Nitrogen compounds (CHN)	34	6.5	17.3
Amines	11	4.1	6.9
Cyanides	8	4.6	19.6
With NN bonds	11	8.1	29.5
Oxygen compounds (CHO)	39	5.2	6.8
Alcohols, ethers	12	4.3	5.7
Aldehydes, ketones	10	4.3	6.1
Acids, esters, anhydrides	9	3.2	6.8
With OO bonds	5	7.7	3.4
Nitrogen-oxygen compounds (CHNO)	7	14.8	15.4
All compounds excluding those with 4-membered rings, <i>tert</i> -butyl groups, or NO bonds	122	5.0	10.9

# Resultados do MNDO

## Calor de isomerização e barreiras rotacionais

Table III. Heats of Isomerization  $\Delta H_{\text{isom}}$  for Closed Shell Molecules



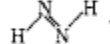
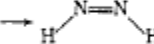
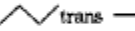
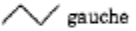
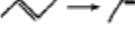

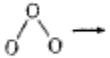













Isomerization reaction	$\Delta H_{\text{isom}}$ , kcal/mol			Isomerization reaction	$\Delta H_{\text{isom}}$ , kcal/mol		
	Calcd	Exptl <sup>a</sup>	Error		Calcd	Exptl <sup>a</sup>	Error
$\text{CH}_2=\text{CH} \rightarrow \text{CH}_2=\text{C}=\text{CH}_2$	2.9	1.3	+1.6	 $\rightarrow$ 	9.8	-5.2	+15.0
$\text{CH}_2=\text{CH} \rightarrow \triangle$	26.3	21.8	+4.5	$\text{C}_2\text{H}_5\text{NH}_2 \rightarrow (\text{CH}_3)_2\text{NH}$	6.3	7.0	-0.7
$\text{CH}_3\text{CH}=\text{CH}_2 \rightarrow \triangle$	6.3	7.8	-1.5	 $\rightarrow$ 	1.8	>0 <sup>c</sup>	
 $\rightarrow$ 	0.5	0.6 <sup>b</sup>	-0.1	$\text{CH}_3\text{CN} \rightarrow \text{CH}_3\text{NC}$	41.1	14.7	+26.4
 $\rightarrow$ 	1.1	1.1	0.0	 $\rightarrow$ 	30.5		
 $\rightarrow$ 	5.4	2.8	+2.6	$\text{C}_2\text{H}_5\text{OH} \rightarrow (\text{CH}_3)_2\text{O}$	11.8	12.2	-0.4
 $\rightarrow$ 	3.0	-1.3	+4.3	$\text{CH}_3\text{CH}=\text{O} \rightarrow \text{CH}_2=\text{CH}-\text{OH}$	7.5		
 $\rightarrow$ 	0.5	2.0	-1.5	$\text{CH}_3\text{CH}=\text{O} \rightarrow \triangle\text{O}$	26.9	27.1	-0.2
 $\rightarrow$ 	1.3	1.0	+0.3	$\text{CH}_3\text{COOH} \rightarrow \text{HCOOCH}_3$	15.6	19.7	-4.1
 $\rightarrow$ 	7.2	7.2	0.0	 $\rightarrow$ 	1.6	0.5	-1.1

Table IV. Rotational Barriers

Molecule	Barrier, kcal/mol			Ref
	Calcd	Exptl	Error	
$\text{CH}_2=\text{CH}_2$	62.5 <sup>b</sup>	65.0	-2.5	<i>a</i>
$\text{CH}_3-\text{CH}_3$	1.0	2.9	-1.9	<i>c</i>
$\text{CH}_3-\text{NH}_2$	1.1	2.0	-0.9	<i>d</i>
$\text{CH}_3-\text{OH}$	0.7	1.1	-0.4	<i>e</i>
$\text{NH}_2-\text{NH}_2$				
Cis	8.0			<i>f</i>
Trans	0.0			
$\text{HO}-\text{OH}$				
Cis	6.9	7.0	-0.1	<i>g</i>
Trans	0.0	1.1	-1.1	



# Resumo do método MNDO

- O método MNDO possibilitou a correção de muitos dos problemas que o método MINDO/3 apresentava.
- Cálculos de propriedades termodinâmicas e mecanismos de reação passaram a ser preditos com certo êxito.
- No entanto, este ainda continha algumas falhas que eram consideradas graves, como, por exemplo, a dificuldade - quase impossibilidade - de descrever ligações de hidrogênio.
- As energias de repulsão entre os átomos eram superestimadas
- Além disso, as energias de ativação obtidas com o método MNDO eram tendenciosamente superestimadas.

**Solução**



Novo método:  
AM1 (Austin Model 1)

# Método AM1

1. Introdução de gaussianas esféricas na interação núcleo-núcleo
2. Nova parametrização

$$E_N(A, B) = E_N^{MNDO}(A, B) + \frac{Z_A Z_B}{R_{AB}} \left( \sum_k a_{k_A} e^{-b_{k_A} (R_{AB} - c_{k_A})^2} + \sum_l a_{l_B} e^{-b_{l_B} (R_{AB} - c_{l_B})^2} \right)$$

- Nesta equação  $a$ ,  $b$  e  $c$  são os coeficientes ajustáveis que definem a intensidade, largura e posição destas funções Gaussianas, respectivamente.
- Os valores de  $a$ ,  $b$  e  $c$  para os átomos de H, C, N e O foram ajustados cuidadosamente utilizando intuição química.

## De que forma:

- Gaussianas atrativas eram posicionadas nas regiões onde existiam repulsões excessivas, além da distância de ligação.
- Gaussianas repulsivas foram posicionadas em distâncias internucleares pequenas.

# Método AM1

AM1: A New General Purpose Quantum Mechanical Molecular Model<sup>1</sup>

Michael J. S. Dewar,\* Eve G. Zoebisch, Eamonn F. Healy, and James J. P. Stewart

*Contribution from the Department of Chemistry, The University of Texas at Austin, Austin, Texas 78712. Received October 29, 1984*

O AM1 é um dos métodos mais citados em química computacional, até este momento com perto de 13 mil citações.

## Problemas

1. Compostos hipervalentes não são bem reproduzidos, tanto a geometria quanto o calor de formação.
2. A geometria de dímeros da água não é predita corretamente.
3. Compostos com o grupo nitro apresentam calor de formação muito positivo.

**Solução**



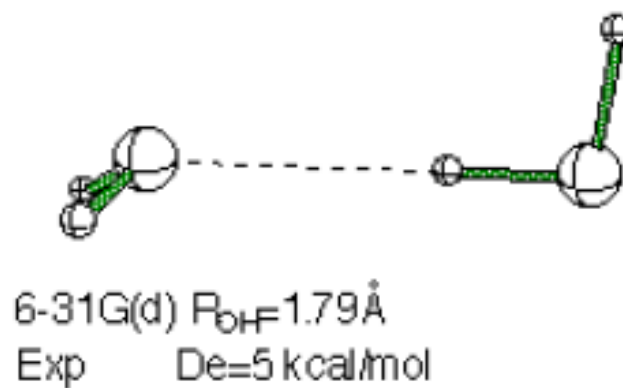
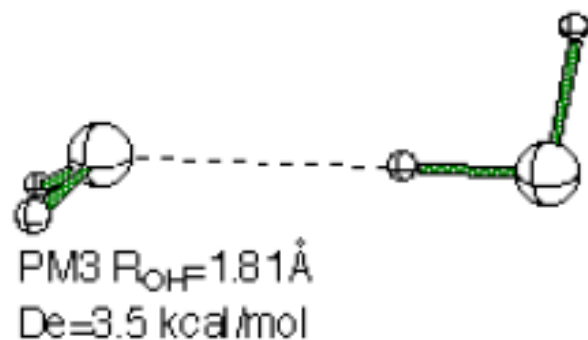
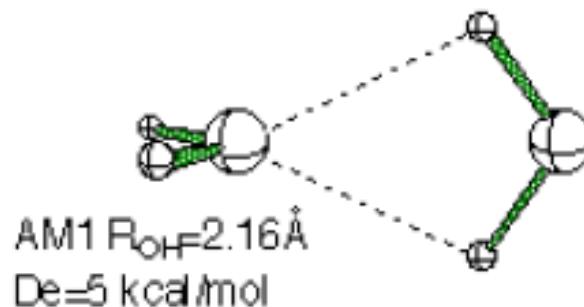
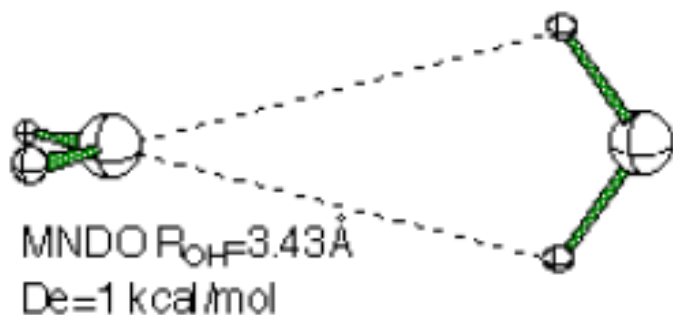
Novo método:  
PM3 (Parametric Method 3)

# Método PM3

O método PM3 é idêntico ao AM1 no formalismo e difere no procedimento de parametrização:

- Mais moléculas no banco de dados;
- Consideração de compostos hipervalentes;
- Técnicas numéricas mais robustas de parametrização;
- Mais átomos parametrizados;
- Todos os parâmetros foram ajustados.

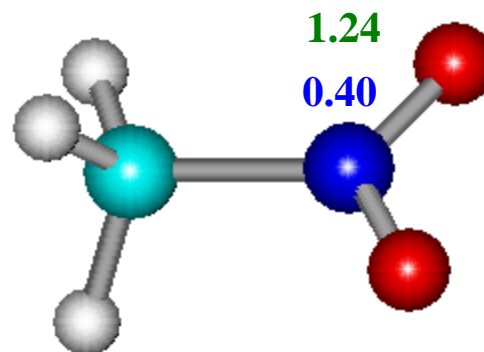
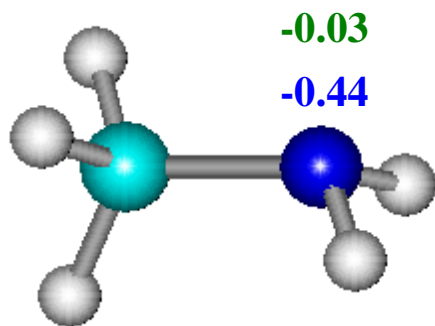
# Método MNDO, AM1 e PM3: ligação de hidrogênio



# Método PM3: limitações

## Problemas

1. Átomos de nitrogênio possuem cargas atômicas não-realistas.
2. Nitrogênios  $sp^3$  são preditos como sendo piramidais.
3. A energia de dímeros da água não é predita corretamente.
4. A função core-core apresenta mínimos espúrios que não são encontrados experimentalmente.



**Cargas PM3 nos átomos de N**

# Método MNDO, AM1 e PM3: comparação

Propriedade	MNDO	AM1	PM3
Calor de Formação (kcal/mol)	18.422	11.923	9.916
Distância de ligação ( Å)	0.066	0.053	0.065
Ângulos (graus)	6.289	5.467	5.708
Momentos dipolares (D)	0.713	0.485	0.572
Potencial de Ionização (eV)	0.921	0.715	0.752

O artigo que apresentou o PM3 possui até hoje mais de 8mil citações.

# Recife Model 1 (RM1)

Theory – NDDO/AM1

## ◆ Atoms

- C, H, N, O ← Organic Chemistry
- P, S ← Biochemistry
- F, Cl, Br, I ← Drugs
- Lanthanides ← Photonics and Medicine

◆ 191 + 252 = 443 parameters



## RM1: A Reparameterization of AM1 for H, C, N, O, P, S, F, Cl, Br, and I

GERD B. ROCHA,<sup>1</sup> RICARDO O. FREIRE,<sup>1</sup> ALFREDO M. SIMAS,<sup>1</sup> JAMES J. P. STEWART<sup>2</sup>

<sup>1</sup>Departamento de Química Fundamental, CCEN, UFPE, 50590-470, Recife PE, Brazil

<sup>2</sup>Stewart Computational Chemistry, 15210 Paddington, Colorado Springs Colorado 80921-2512

Received 25 October 2005; Accepted 27 December 2005

DOI 10.1002/jcc.20425

Published online in Wiley InterScience (www.interscience.wiley.com).

© 2006 Wiley Periodicals, Inc. J Comput Chem 27: 1101–1111, 2006

271 citations 19.07.2014 (web of Science)



# Sistemas Moleculares de Referência

## ◆ 1775 sistemas moleculares

- 1313 calores de formação;
- 232 potenciais de ionização;
- 199 momentos dipolares;
- 908 distâncias interatômicas;
- 907 ângulos de ligação e diedros.

# www.rm1.sparkle.pro.br

RM1 Semiempirical Molecular Orbital Model - Windows Internet Explorer

http://www.rm1.sparkle.pro.br/

Windows Live Bing

Novidades Perfil Email Fotos Calendário MSN Compartilhar

Favoritos Sites Sugeridos Obtenha mais comple...

RM1 Semiempirical Molecular Orbital Model

Home RSS Print Página Segurança Ferramentas

## RM1 Semiempirical Molecular Orbital Model

中文 | 日本語 | 한국어 | ไทย

- RM1 x AM1, PM3, PM5
- RM1 software
  - MOPAC2009
  - MOPAC - earlier
  - Spartan'08
  - AMPAC
  - AMBER
  - HyperChem
  - PCGAMESS/Firefly
  - pDynamo
  - Phenix eLBOW
  - ConGENER
  - GAMESS-US
  - Gaussian 09
  - Gaussian - earlier
  - GROMACS
  - Ghemical
  - CambridgeSoft
  - WebMO
  - Winmostar
  - MATEO
  - SimuPac
- Recent articles
- Books
- RM1 research topics
  - QSAR
  - ADME-Tox
  - Barrier height
  - Calixarene
  - Cavitand
  - CI

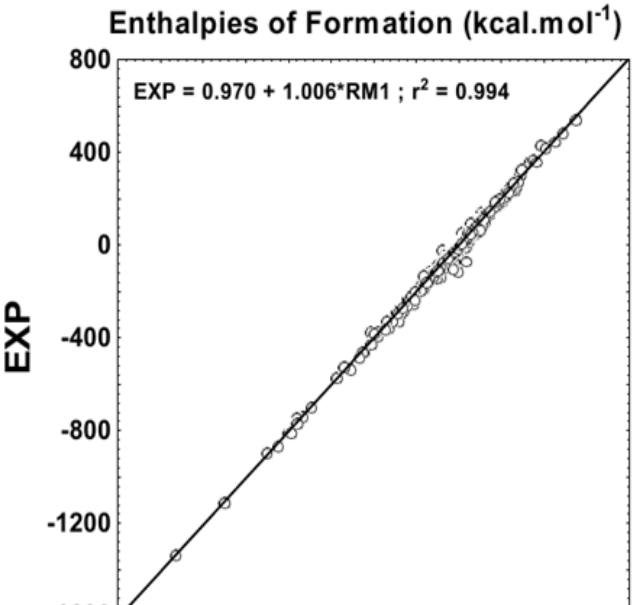
### Scholarly article

#### RM1: a Reparameterization of AM1 for H, C, N, O, P, S, F, Cl, Br, and I

Gerd Bruno Rocha, Ricardo Oliveira Freire, Alfredo Mayall Simas\*, and James J. P. Stewart.  
*Journal of Computational Chemistry* 27(10), 1101-1111, 2006

#### Enthalpies of Formation (kcal.mol<sup>-1</sup>)

EXP = 0.970 + 1.006\*RM1 ; r<sup>2</sup> = 0.994



### What MOPAC is

MOPAC (Molecular Orbital PACKage) is a semiempirical quantum chemistry program based on Dewar and Thiel's NDDO approximation. Most users use MOPAC with a [Graphical User Interface](#).

### MOPAC2012

[MOPAC2012](#) is MOPAC2009 plus the [PM7](#) and [PM7-TS](#) methods. If a bug is detected, please send a message by [E-mail](#) to MrMOPAC@OpenMOPAC.net, along with an example illustrating the bug. If you qualify for Academic not-for-profit use and already have a password for MOPAC 2009, go straight to [download](#), otherwise request a [password and download](#). For commercial and governmental prices, see [Prices](#).

Commercial users in Europe and America, please contact [CACHe Research](#) (see [Resellers](#))

Commercial users in Japan, please contact [Ryoka Systems](#)

Commercial users in Korea, please contact [KREIS I&C](#)

### PM7

PM7 is a modified form of PM6. A few errors in NDDO theory that affect large systems have been removed. All atomic and diatomic parameters were re-optimized. Average errors in organic compounds have been reduced by ~10%, and errors in large organics and [solids](#) have been significantly reduced, see [PM7 Accuracy](#).

### Reducing Computation Time

Computation time has been reduced, in some cases by over 99%, as a result of using the Intel Math Kernel Library (MKL) and by parallelizing parts of the code. For details, see [MKL](#) and [HTTP://www.quantum-chem.pro.br](http://www.quantum-chem.pro.br)

### GPU

Recently, MOPAC2012 was modified to allow NVIDIA GPU acceleration to be used, for details, see [GPU](#), this allows large systems that do not use MOZYME to run between two to just over three times faster than with just MKL and parallelization. MOZYME jobs do not run faster when a GPU chip is used. Read [Requirements for GPU MOPAC](#) before installing it.

### References and Citations Instructions and Manual

[MOPAC2012](#) [MOPAC2009](#), [PM6](#), [PM7](#)

A [manual](#) is available for MOPAC2012. Please try finding answers to questions there before sending an E-mail message requesting help. For individual topics of interest, see [Discussion Topics](#).

# Programa MOPAC

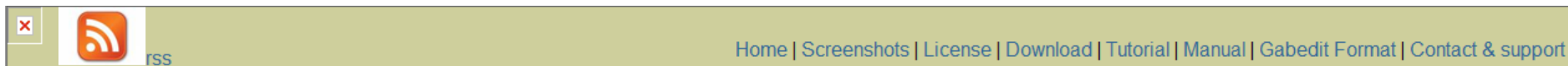
A maioria dos grandes avanços obtidos na pesquisa de métodos semi-empíricos foi essencialmente incorporada em um único programa: o **MOPAC**. Iniciado em 1983, no grupo do Prof. M. Dewar, com a participação do Prof. J. Stewart.

- 1983: MOPAC → Otimização de geometria, análise vibracional e estados de transição.
- 1984: MOPAC 2 → Conserto de erros.
- 1985: MOPAC 3 → Método AM1, cálculo de polímeros e CI (*Configuration interaction*).
- 1987: MOPAC 4 → Conserto de erros.
- 1988: MOPAC 5 → Método PM3 e caminho de reação.
- 1989: MOPAC 6 → Conserto de erros, cálculo de potencial eletrostático e extensão do PM3 para 28 elementos.
- 1993: MOPAC 7/93 → Modelo de efeito solvente, propriedades ópticas não-lineares e análises de simetria.
- 1997: MOPAC 97 → Método MNDO-d e a técnica de escalonamento linear MOZYME
- 1999: MOPAC 2000 → Alocação dinâmica de memória, moléculas grandes, sólidos, metais de transição, melhoramento no algoritmo de busca do estado de transição, cruzamento inter-sistemas e análise de zona de Brillouin.
- 2001: MOPAC 2002 → Melhoramento no pacote de solvatação, todos os átomos do grupo principal, método PM5 e parametrização para mais metais de transição.
- 2004: MOPAC 2005 → Cálculo de complexos de lantanídeos através do modelo *Sparkle/AM1*.
- 2009: MOPAC 2009 → Cálculo de biomoléculas (MOZYME) e Sólidos e cristais.

**MOPAC 2009 → Programa livre ([openmopac.net](http://openmopac.net))**

# Gabedit

<http://gabedit.sourceforge.net/>



## What is Gabedit ?

Gabedit is a graphical user interface to computational chemistry packages like Gamess-US, Gaussian, Molcas, Molpro, MPQC, OpenMopac, Orca, PCGamess and Q-Chem

It can display a variety of calculation results including support for most major molecular file formats. The advanced "Molecule Builder" allows to rapidly sketch in molecules and examine them in 3D. Graphics can be exported to various formats, including animations.

## Major features

- Gabedit can Create input file for GAMESS(US), GAUSSIAN, MOLCAS, MOLPRO, MPQC, OpenMopac, Orca, PCGamess and Q-Chem.
- Gabedit can graphically display a variety of Gamess-US, Gaussian, Molcas, Molpro, MPQC, OpenMopac, Orca, PCGamess, Q-Chem and (partially) ADF calculation results, including the following
  - Molecular orbitals.
  - Surfaces from the electron density, electrostatic potential, NMR shielding density, and other properties.
  - Surfaces may be displayed in solid, translucent and wire mesh modes. they are can be colorcoded by a separate property.
  - Contours (colorcoded), Planes colorcoded, Dipole. XYZ axes and the principal axes of the molecule.
  - Animation of the normal modes corresponding to vibrational frequencies.
  - Animation of the rotation of geometry, surfaces, contours, planes colorcoded, xyz and the principal axes of the molecule.
  - Animation of contours. Animation of planes colorcoded

Vamos usar esse programa para realizar os  
tutoriais encontrados em:  
<http://www.sparkle.pro.br/>

# Tutoriais

[www.sparkle.pro.br](http://www.sparkle.pro.br)



## Lanthanide Complexes Computational Chemistry

- [Ab Initio](#)
- [Independent Review](#)
- ▼ **Tutorial**
  - [Using MOPAC2009](#)
  - [Drawing Complexes](#)
  - [Viewing Molecules](#)
  - [UV Vis Spectra](#)
  - [IR Spectra](#)
  - [Drawing Orbitals](#)
  - [Electrostatic Map](#)
  - [pKa](#)
  - [Solvent Effects](#)
  - [Nonlinear Optics](#)
  - [Reaction Path](#)
  - [Transition State](#)
  - [Spin Density](#)
  - [Electron Density](#)
  - [Geometry](#)
  - [Solid State](#)
- ▼ **Original Articles**
  - [Lanthanum](#)
  - [Cerium](#)

[Home](#)

### Tutorial Software

- [Using MOPAC2009](#)
- [Drawing Complexes](#)
- [Viewing Molecules](#)
- [UV Vis Spectra](#)
- [IR Spectra](#)
- [Drawing Orbitals](#)
- [Electrostatic Map](#)
- [pKa](#)
- [Solvent Effects](#)
- [Nonlinear Optics](#)
- [Reaction Path](#)
- [Transition State](#)
- [Spin Density](#)
- [Electron Density](#)
- [Geometry](#)
- [Solid State](#)

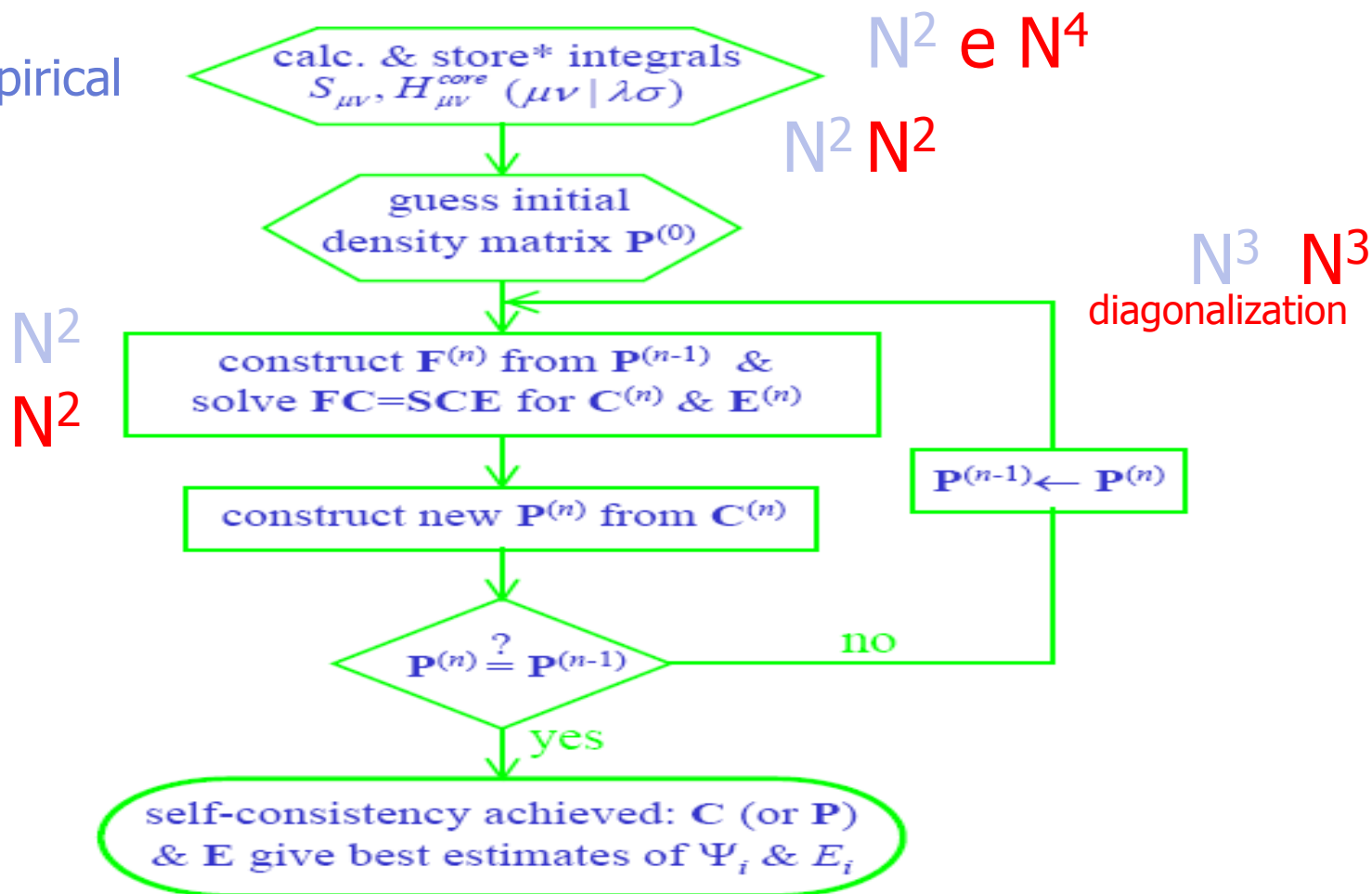
# **Algoritmos de linear scaling na CPU e na memória para o tratamento quântico de proteínas**

# Where is the bottleneck?

## Complexidades para algumas partes do cálculo SCF convencional

Ab initio

Semiempirical





# Natureza do problema (localidade em química quântica)

A interação Coulômbica depende de:  $r^{-1}$

Em se tratando de matrizes densidade:

$$\rho(r,r') = \exp[-\sqrt{E_{\text{gap}}} |r-r'|]$$

Apenas os elementos da matriz densidade que pertencem a átomos próximos possuem valor significativo.

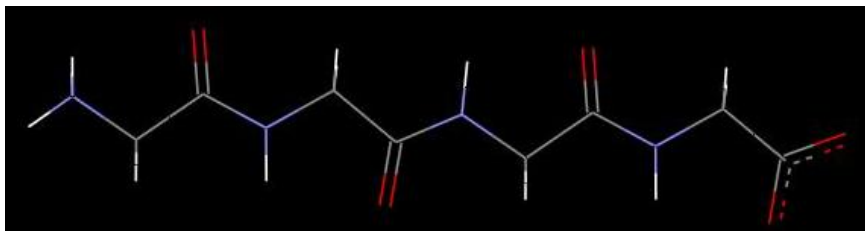
- Raios de corte
- Gera matrizes esparsas

# Uso de raios de corte

TABLE I  
Bond-orders between terminal nitrogen and backbone atoms in GLY-GLY-GLY-GLY.

Bond to	Order
C <sub>2</sub>	1.017090
C <sub>3</sub>	0.015249
N <sub>4</sub>	0.004989
C <sub>5</sub>	0.001219
C <sub>6</sub>	0.000086
N <sub>7</sub>	0.000031 → 7.2Å
C <sub>8</sub>	0.000009
C <sub>9</sub>	0.000001
N <sub>10</sub>	0.000000
C <sub>11</sub>	0.000000
C <sub>12</sub>	0.000000

GLY-GLY-GLY-GLY has the structure  $\text{H}_2\text{N}-\text{CH}_2-\text{CO}-\text{NH}-\text{CH}_2-\text{CO}-\text{NH}-\text{CH}_2-\text{COOH}$



7-9 Å → Valor comumente usado

Definido o raio de corte a esparsidade da matriz densidade está praticamente determinada

TABLE IV  
Effect of varying cutoff.

Cutoff (Å)	$\Delta H_i$ (kcal/mol)	Diff. <sup>a</sup>	Gradient <sup>b</sup> (kcal/mol/Å)	Time (s)
3.0	-2071.41	+72.99	3.629	570
4.0	-2152.08	-7.68	1.146	1078
5.0	-2135.00	+9.40	0.689	1493
6.0	-2141.87	+2.53	0.596	1955
7.0	-2143.56	+0.84	0.723	2267
8.0	-2143.51	+0.89	0.351	2360
9.0	-2144.13	+0.27	0.230	2583
10.0	-2144.33	+0.07	0.250	2840
12.0	-2144.16	+0.24	0.170	2882
14.0	-2144.33	+0.07	0.131	3056
16.0	-2144.40	0.00	0.084	3157
18.0	-2144.40	0.00	0.061	3198
20.0	-2144.41	-0.01	0.038	3219
22.0	-2144.39	+0.01	0.026	3257
24.0	-2144.40	0.00	0.014	3277
26.0	-2144.40	0.00	0.000	3273

<sup>a</sup>Difference between  $\Delta H_i$  at given cutoff and  $\Delta H_i$  at 26 Å.

<sup>b</sup>Root-mean-square difference of gradient components at given cutoff and gradient at 26 Å.

# Orbitais moleculares localizados

$$\mathbf{FC} = \mathbf{S} \mathbf{C} \epsilon$$

As soluções da eq. de Fock (os autovetores  $\rightarrow$  C) geram orbitais canônicos que se espalham sobre toda a molécula.

Os orbitais canônicos não são apropriados para o tratamento computacional de sistemas grandes  $\rightarrow$  não refletem a natureza local do problema.

Mas, as soluções localizadas podem ser conseguidas a partir de uma transformação unitária  $\rightarrow N^3$

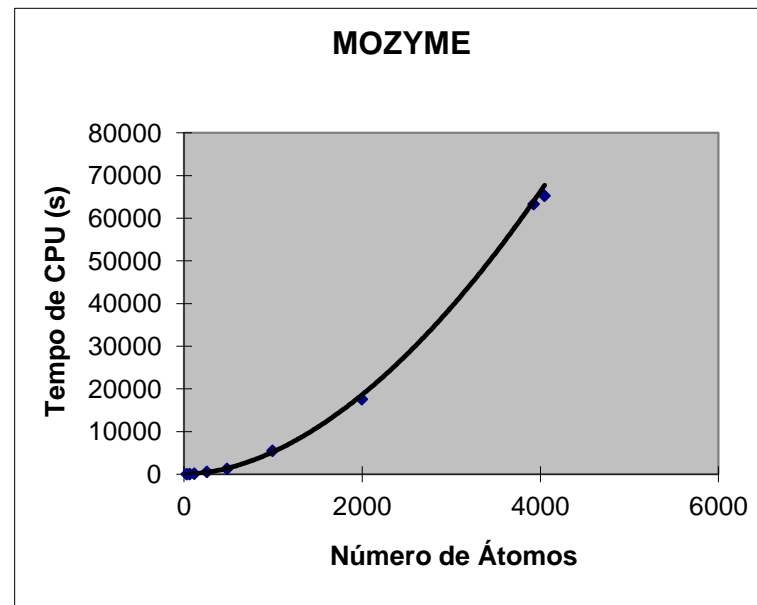
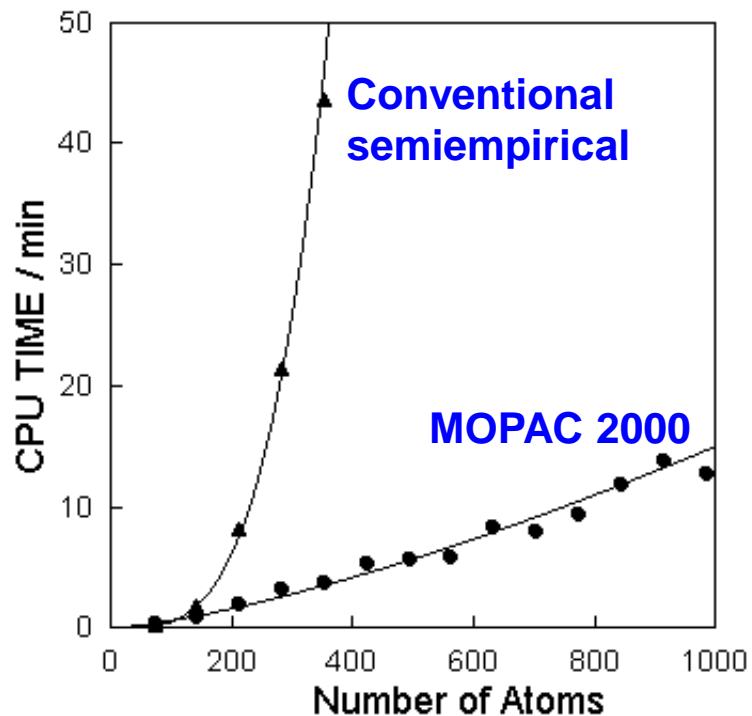
Para resolver isso eu já inicio o cálculo com orbitais moleculares localizados.

# MOZYME

Application of Localized Molecular Orbitals to the Solution of Semiempirical Self-Consistent Field Equations. J. J. P. Stewart. Int. J. Quantum Chem., 58, 1996, 133-146.

- Aplica raios de corte para evitar interações de longa distância.
- Começa com orbitais localizados ortonormais de acordo com estruturas de Lewis;
- Para obter o SCF → deixar os orbitais localizados se expandirem;
- Como? Aniquilando todos os elementos da matriz de Fock que conectam os orbitais ocupados com os virtuais. (seguidas rotações de Jacobi).

# MOZYME



18,000 atoms

70 years  
conventional  
semiempirical

=

1 day  
MOPAC 2000

Escalonamento na CPU  $N^{1.9}$

# O Problema do Protein Folding

## Assessment of Semiempirical Quantum Mechanical Methods for the Evaluation of Protein Structures

Andrew M. Wollacott<sup>†</sup> and Kenneth M. Merz, Jr.\*

**Abstract:** The ability to discriminate native structures from computer-generated misfolded ones is key to predicting the three-dimensional structure of a protein from its amino acid sequence. Here we describe an assessment of semiempirical methods for discriminating native protein structures from decoy models. The discrimination of decoys entails an analysis of a large number of protein structures and provides a large-scale validation of quantum mechanical methods and their ability to accurately model proteins. We combine our analysis of semiempirical methods with a comparison of an AMBER force field to discriminate decoys in conjunction with a continuum solvent model. Protein decoys provide a rigorous and reliable benchmark for the evaluation of scoring functions, not only in their ability to accurately identify native structures but also to be computationally tractable to sample a large set of non-native models.

*J. Chem. Theory Comput.* **2007**, *3*, 1609–1619

# O Problema do Protein Folding

**Table 1.** Select Decoy Sets Used

decoy set	PDB	description	type	$N_{\text{decoys}}$	$N_{\text{res}}$	$N_{\text{atoms}}$	rmsd range (Å)	%H/%E <sup>b</sup>
4-state-reduced	1ctf	C-terminal domain of ribosomal protein L7/L12	X-ray	630	68	1005	2.1–9.8	53/26
	1r69	N-terminal domain of phage 434 repressor	X-ray	675	63	997	2.2–9.4	70/0
	1sn3	scorpion toxin variant 3	X-ray	660	65	948	2.5–10.3	12/22
	2cro	phage 434 Cro protein	X-ray	674	65	1081	2.0–9.5	66/0
	4pti	trypsin inhibitor	X-ray	687	58	892	2.8–10.7	16/24
	4rxn	rubredoxin	X-ray	677	54	794	2.5–9.2	0/20
Rosetta	1gb1	immunoglobulin binding domain of streptococcal protein G	NMR	999	54	823	3.1–18.0	28/33
	1hsn	high mobility group protein I box	NMR	999	62	1014	4.1–17.6	79/0
	1orc	Cro repressor (mutant)	X-ray	999	56	877	4.0–14.1	46/27
	1pgx	protein G (B2 domain)	X-ray	999	57	873	3.4–20.4	26/46
	1uxd	fructose repressor DNA-binding domain	NMR	999	43	690	2.2–12.5	81/0
	2fow	RNA binding domain of ribosomal protein LII	NMR	999	66	1009	4.0–21.6	52/9
	1hc8 <sup>a</sup>	RNA binding domain of ribosomal protein LII	X-ray	999	66	1009	4.0–21.6	53/9
	1r69	N-terminal domain of phage 434-repressor	X-ray	999	61	976	3.1–15.5	70/0

<sup>a</sup> 1hc8 is the X-ray structure for the 2fow decoy set. <sup>b</sup> Percentage of helices and sheets in the native.

**Table 5.** Energy Decomposition for All-Atom Minimized Decoy Sets as Calculated with DivScore Using the PM3 Hamiltonian

decoy type	system	rank <sub>tot</sub>	rank <sub>H<sub>F</sub></sub> <sup>a</sup>	gap <sub>H<sub>F</sub></sub> <sup>d</sup>	rank <sub>solv</sub> <sup>b</sup>	gap <sub>solv</sub> <sup>d</sup>	rank <sub>LJ6</sub> <sup>c</sup>	gap <sub>LJ6</sub> <sup>d</sup>
4-state-reduced	1ctf	1	4	44.62	471	570.51	1	−22.04
	1r69	1	2	19.53	594	559.40	1	−58.94
	1sn3	1	1	−63.47	569	613.26	5	22.06
	2cro	1	16	72.01	524	515.20	1	−59.28
	4pti	1	16	95.99	440	455.73	25	58.75
	4rxn	1	1	−32.07	646	925.86	10	34.17
Rosetta	1gb1	1	3	4.07	976	677.13	1	−13.66
	1hsn	15	154	152.73	857	447.77	1	−1.42
	1orc	1	64	183.35	812	570.97	7	6.07
	1pgx	1	1	−10.12	978	627.83	1	−60.63
	1uxd	45	21	118.11	972	343.64	1	−23.04
	2fow	5	56	173.82	894	618.72	4	16.15
	1hc8	1	24	124.08	920	644.35	1	−12.19
	1r69	1	1	−10.07	977	828.78	1	−141.21

<sup>a</sup> Heat of formation. <sup>b</sup> Solvation energy. <sup>c</sup> Dispersive term of the classic LJ6–12 potential. <sup>d</sup> Energy gaps are reported as the energetic difference between the native structure and the decoy with the lowest value for the given energetic term.

# Programa do curso.

## 2º dia:

Prática: Cálculo de propriedades de proteínas

- Encontrando estruturas nativas de proteínas a partir do cálculo de  $\Delta H$  de formação;



**FIM**