



MÉTODOS QUÂNTICOS SEMIEMPÍRICOS: TEORIA E PRÁTICA  
*Modelando Sistemas Biológicos com Métodos de Química  
Quântica*

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

**M.Sc. Luiz Eduardo Gomes da Cruz**

[luiz\\_eduardo\\_gc@hotmail.com](mailto:luiz_eduardo_gc@hotmail.com)

**Dr. Igor Barden Grillo**

[barden.igor@gmail.com](mailto:barden.igor@gmail.com)

**M. Sc. Júlio D. C. Maia**

[juliodaniel.carvalho@gmail.com](mailto:juliodaniel.carvalho@gmail.com)

**Prof. Gabriel A. U. Carvalho**

[urquizagabes@gmail.com](mailto:urquizagabes@gmail.com)

**Prof. Gerd Bruno Rocha**

[gbr@academico.ufpb.br](mailto:gbr@academico.ufpb.br)

[gbr@quimica.ufpb.br](mailto:gbr@quimica.ufpb.br)

[https://github.com/RochaGerd/Chemistry\\_with\\_Python](https://github.com/RochaGerd/Chemistry_with_Python)



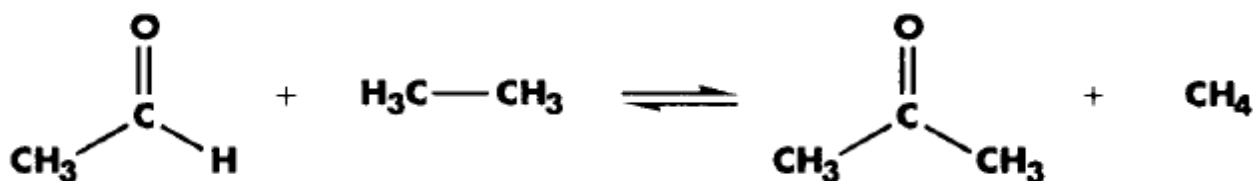
Grupo de Química Computacional mais Oriental das Américas  
L.Q.Q.C. - UFPB - João Pessoa - PB

Novembro 2024, Pirenópolis GO – Brasil

## EXERCÍCIO 1

### CALCULANDO A VARIAÇÃO DE ENTALPIA DE UMA REAÇÃO QUÍMICA, $\Delta\tilde{H}_r$

Uma reação isodésmica é aquela em que o total de cada tipo de ligação química é idêntico em ambos os lados da reação química. Sabendo disso calcule a  $\Delta\tilde{H}_r$  (entalpia de reação) para a reação a seguir, com os métodos AM1, PM3 e RM1<sup>1,2</sup> e usando o programa MOPAC<sup>3,4</sup>.



Para uma transformação qualquer, a variação da entalpia é calculada como sendo a soma da entalpia dos produtos menos a soma da entalpia dos reagentes, conforme equação a seguir.

$$\Delta\tilde{H}_r = \sum_{i=1}^{prod.} \Delta H_f^{(i)} - \sum_{j=1}^{reag.} \Delta H_f^{(j)}$$

▪ Roteiro a ser seguido:

- a) Montar todas as moléculas no programa de sua escolha, salvando-as em formato do MOPAC (.mop ou .dat).
  - Gabedit
  - Avogadro
  - Qualquer gerador online de estruturas:
    - <http://www.cheminfo.org/Chemistry/Cheminformatics/FormatConverter/index.html>.
    - <http://cdb.ics.uci.edu/cgi-bin/BabelWeb.py>
  - Outra opção para montar o input é acessar o link ([http://openmopac.net/PM7\\_accuracy/normal%20molecules.html](http://openmopac.net/PM7_accuracy/normal%20molecules.html)) e buscar pelo nome das referidas moléculas (em inglês) e salvar suas coordenadas já no formato do MOPAC.
- b) Executar cada uma das moléculas usando o programa MOPAC (com as palavras-chave: **RM1** GNORM=0.01 PRECISE XYZ), utilizando o método RM1. Para os demais

métodos mude a *keyword* **RM1** por uma das: AM1 ou PM3.

Para fazer isto, rode o MOPAC com a seguinte linha de comando:

```
$ mopac <nome-do-arquivo> &
```

onde tem <nome-do-arquivo> coloque o nome do seu arquivo de entrada sem a extensão (.mop).

c) Em seguida monte a seguinte tabela.

Você pode conferir um *script* em python no diretório desse exercício. Ele chama-se “isodesmic.py”. Para executá-lo basta digitar:

```
$ python isodesmic.py
```

Moléculas		$\Delta\tilde{H}_f$ (kcal/mol)		
		AM1	PM3	RM1
1	CH <sub>4</sub>			
2	(CH <sub>3</sub> ) <sub>2</sub> C=O			
3	H <sub>3</sub> C-CH <sub>3</sub>			
4	H <sub>3</sub> C-CHO			
$\Delta\tilde{H}_r$ (kcal/mol)				
		$\Delta\tilde{H}_r^{exp}$ (kcal/mol) = -9.90		

Para finalizar indique qual foi o método que apresentou a melhor concordância com o resultado experimental.

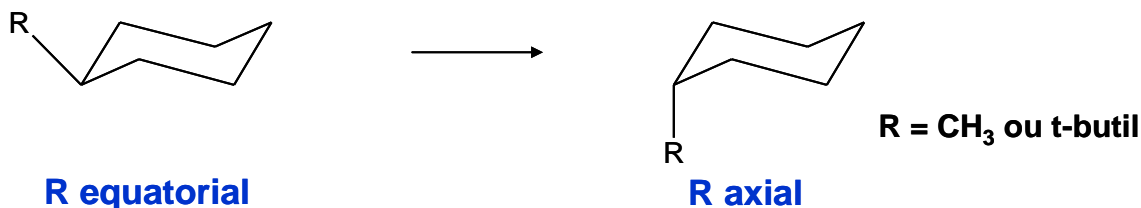
Esse exercício pode ser executado perfeitamente pelo Google Colab. Existe uma versão dele no link:

- [https://github.com/RochaGerd/Chemistry\\_with\\_Python](https://github.com/RochaGerd/Chemistry_with_Python)

## EXERCÍCIO 2 (opcional)

### CALCULANDO O $\Delta\tilde{H}_r$ DE UMA MUDANÇA CONFORMACIONAL

Neste segundo exercício vamos investigar as preferências conformacionais de alquilciclohexanos utilizando os métodos AM1, PM3 e RM1.



- Roteiro a ser seguido:
  - a) Executar uma otimização de geometria selecionando o método RM1, usando o programa MOPAC (com as palavras-chave: **RM1** GNORM=0.01 PRECISE XYZ). Para os demais métodos mude a *keyword* **RM1** por uma das: AM1 ou PM3.
  - b) Monte a tabela a seguir com os resultados obtidos dos cálculos.

Molécula	$\Delta\tilde{H}_f$ (kcal/mol)		
	AM1	PM3	RM1
ax-metilciclohexano			
ax-t-butil-ciclohexano			
eq-metilciclohexano			
eq-t-butil-ciclohexano			
$\Delta\tilde{H}_r$ (kcal/mol)			
$\Delta\tilde{H}_r^{exp}$ [R = CH <sub>3</sub> ] = 1.7 kcal/mol			
$\Delta\tilde{H}_r^{exp}$ [R = t-butil] = 5.4 kcal/mol			

Para finalizar indique qual o método apresenta a melhor concordância com o resultado experimental.

Para usuários de Linux, na pasta desse exercício existe um script “run\_all\_ex2.sh” que executa automaticamente todos os cálculos do MOPAC. Para usá-lo você deve apenas torná-lo um

arquivo executável com o comando:

```
$ chmod u+x run_all_ex2.sh
```

Uma vez rodado todos os cálculos, você pode extrair os dados de  $\Delta\tilde{H}_f$  das moléculas através do script `read_heats_ex2.py`. Para usar esse script, você deve editá-lo e atribuir o caminho (path) onde estão seus arquivos à variável “`directory`”. Veja a seguir onde localizar essa variável:

```
# Diretório contendo os arquivos out
# Esse dado precisa ser editado
directory = "/home/rochagb/my_files/Minicurso/exercicio2"

# Lista para armazenar os dados
data = []
```

Esse script gera um arquivo “.csv” com todos os valores de  $\Delta\tilde{H}_f$  das moléculas. Basta apenas alimentar a tabela e calcular os  $\Delta\tilde{H}_r$  das reações de mudança conformacional.

Esse exercício também pode ser executado perfeitamente pelo Google Colab. Existe uma versão dele no link:

- [https://github.com/RochaGerd/Chemistry\\_with\\_Python](https://github.com/RochaGerd/Chemistry_with_Python)

## EXERCÍCIO 3

### ENCONTRANDO ESTRUTURAS NATIVAS DE PROTEÍNAS A PARTIR DO CÁLCULO DAS ENTALPIAS DE FORMAÇÃO DE *DECOYS* ( $\Delta\tilde{H}_f$ )

Proteínas são moléculas que exercem uma imensa variedade de funções nas células vivas. De dar sustentação à estrutura dos órgãos ao metabolismo, as proteínas estão envolvidas em tudo.

Para entender como funciona e o quão complicado é a estrutura de uma proteína, imagine um colar de pérolas aberto, estendido sobre uma mesa. Agora imagine quantas formas diferentes existe de se pegar este colar e *enovelá-lo* em suas mãos.

Vá além. Dê um número a cada pérola e estude, das várias e várias formas que existe de se aninhar este colar de pérolas em suas mãos, quais destas deixam a pérola número 21 voltada para fora, ou quais as que mais aproximam as pérolas 3, 9 e 37. Se estiver conseguindo seguir com o exercício mental, então você está começando a ter uma ideia do quão complexo é o problema.

Proteínas funcionam de um jeito parecido, exceto que em vez de pérolas, elas são colares feitos com 20 tipos diferentes de joias, e cada tipo de joia tem diferentes afinidades com os outros 19 tipos.

O que é notável a respeito destas moléculas é que existe uma forma particular de “enovelamento” na qual a proteína exerce sua função. Enovelamentos muito distantes do enovelamento nativo não desempenha sua função nativa corretamente e são ditas formas “desnaturadas”.

*In vivo* (e também *in vitro* para alguns sistemas), uma proteína consegue enovelar-se em tempos da ordem de microssegundos, mas para calcular esta estrutura nativa a partir somente da sequência de “joias” da proteína, levaríamos uma era geológica, a idade do universo, ou outra medida superlativa de tempo efetuando estes cálculos intermináveis.

Ao longo de seu enovelamento, uma proteína deve selecionar entre trilhões, ou mais até, de enovelamentos possíveis (chamados conformações) e encontrar a forma correta de se enovelar. Para ter uma ideia de como este processo é trabalhoso, vamos aqui tentar encontrar a estrutura nativa de três mini proteínas, cada uma delas posta entre 10 outras conformações de cada uma destas proteínas (*decoys*).

Vamos fazer isto comparando as **Entalpias de Formação** ( $\Delta\tilde{H}_f$ ) das estruturas, calculadas com o programa **MOPAC** e aplicando o método de escalonamento linear **MOZYME**<sup>5</sup>. O método semiempírico que usaremos é o **PM7**<sup>6</sup>.

A entalpia de formação ( $\Delta\tilde{H}_f$ ) não é, de forma geral e categórica, uma grandeza química que identifica a estrutura nativa de qualquer proteína, incidentalmente, no entanto, sabemos de antemão que esta grandeza discrimina bem entre conformações nativa e desnaturadas destas proteínas em questão<sup>3,7</sup>.

## CÁLCULO DA ENTALPIA DE FORMAÇÃO

O MOPAC utiliza um formato onde um conjunto de palavras-chave é colocado no cabeçalho de um conjunto de coordenadas atômicas. O MOPAC aceita este arquivo de coordenadas tanto em coordenadas internas quanto em coordenadas cartesianas, além disto, ele também aceita coordenadas no formato **PDB**, nativos do *Protein Data Bank*, o repositório mais importante de estruturas de proteínas do mundo<sup>8</sup>.

Na sua pasta de entrada para esse exercício deve haver três subpastas, cada uma delas contém dez estruturas de proteínas em formato PDB. Estes conjuntos representam conformeros de uma mesma proteína de mesma composição atômica e sequência de aminoácidos.

Os arquivos de entrada do MOPAC devem ter a extensão “.mop”. Para fazer um arquivo “.mop” da estrutura PDB, simplesmente faça uma cópia do arquivo PDB, salve-a com o outro nome e a extensão “.mop”. Escolha uma estrutura de cada grupo e faça um arquivo “.mop” de cada uma. Em seguida, substitua o “REMARK” na primeira linha de cada arquivo “.mop” pelas seguintes palavras-chaves:

```
PM7 1SCF MOZYME XYZ PL T=1D EPS=78.4 RSOLV=1.3 PDB +  
CUTOFF=9.0 DISP(1.0)
```

A figura a seguir mostra como deve ficar as primeiras linhas do arquivo de entrada .mop.

```

PM7 GEO-OK 1SCF MOZYME XYZ PL T=1D EPS=78.4 RSOLV=1.3 PDB +
CUTOFF=9.0 LET DISP(1.0)
1CTF
MOZYME calculation
ATOM      1  N   GLU      1      23.454      7.293     -0.766      1.00      0.00      N
ATOM      2  H1  GLU      1      22.573      6.902     -1.103      1.00      0.00      H
ATOM      3  H2  GLU      1      24.053      7.451     -1.567      1.00      0.00      H
ATOM      4  H3  GLU      1      23.271      8.192     -0.338      1.00      0.00      H
ATOM      5  CA  GLU      1      24.108      6.362      0.201      1.00      0.00      C
ATOM      6  HA  GLU      1      25.080      6.790      0.458      1.00      0.00      H
ATOM      7  CB  GLU      1      23.298      6.263      1.516      1.00      0.00      C
ATOM      8  HB2 GLU      1      22.997      7.263      1.829      1.00      0.00      H
ATOM      9  HB3 GLU      1      22.386      5.698      1.328      1.00      0.00      H
ATOM     10  CG  GLU      1      24.051      5.612      2.686      1.00      0.00      C
ATOM     11  HG2 GLU      1      23.319      5.350      3.453      1.00      0.00      H
ATOM     12  HG3 GLU      1      24.522      4.688      2.354      1.00      0.00      H
ATOM     13  CD  GLU      1      25.110      6.550      3.294      1.00      0.00      C
ATOM     14  OE1 GLU      1      26.038      6.972      2.564      1.00      0.00      O
ATOM     15  OE2 GLU      1      25.011      6.880      4.500      1.00      0.00      O
ATOM     16  C   GLU      1      24.420      5.007     -0.475      1.00      0.00      C
ATOM     17  O   GLU      1      25.339      4.957     -1.288      1.00      0.00      O
ATOM     18  N   PHE      2      23.643      3.940     -0.224      1.00      0.00      N
ATOM     19  H   PHE      2      22.898      4.044      0.450      1.00      0.00      H
ATOM     20  CA  PHE      2      23.790      2.599     -0.833      1.00      0.00      C
ATOM     21  HA  PHE      2      24.466      2.670     -1.687      1.00      0.00      H
ATOM     22  CB  PHE      2      24.415      1.618      0.181      1.00      0.00      C
ATOM     23  HB2 PHE      2      23.642      1.305      0.883      1.00      0.00      H
ATOM     24  HB3 PHE      2      24.736      0.734     -0.367      1.00      0.00      H
ATOM     25  CG  PHE      2      25.607      2.115      0.984      1.00      0.00      C
ATOM     26  CD1 PHE      2      25.562      2.098      2.393      1.00      0.00      C
ATOM     27  HD1 PHE      2      24.678      1.740      2.905      1.00      0.00      H
ATOM     28  CE1 PHE      2      26.663      2.549      3.144      1.00      0.00      C
-----

```

Após preparar os arquivos desta execução, vamos precisar executar o MOPAC com cada arquivo deste. Para fazer isto, rode o MOPAC com a seguinte linha de comando (no Linux):

```
$ nohup mopac <nome-do-arquivo> &
```

Par os usuários de Linux, o comando `nohup` coloca o processo para rodar no *background*, evitando que ele interrompa se o terminal for fechado. O símbolo “&” destrava o terminal para que você possa continuar operando o sistema enquanto o programa roda. Para acompanhar o processo do cálculo, é necessário acompanhar o preenchimento do arquivo de saída `<nome-do-arquivo>.out` usando comando `tail` da seguinte forma:

```
$ tail -f <nome-do-arquivo>.out | grep "DELTAE"
```

Este comando composto irá mostrar as últimas linhas do arquivo de saída, mas apenas aquelas que tiverem a *string* “DELTAE”. A visualização será atualizada para cada nova linha escrita no arquivo de saída, isto permitirá que o progresso do cálculo seja acompanhado à medida em que ele é realizado. Quando o DELTAE atingir um valor muito próximo de zero, o cálculo terá convergido, caso em que o `tail` irá parar de atualizar a escrita das linhas. Neste caso deve-se suspender o `tail` com um CTRL+C.



Para se obter o valor da entalpia de formação da estrutura deve-se aplicar o comando `cat` da seguinte forma:

```
$ cat <nome-do-arquivo>.out | grep "HEAT"
```

O programa `cat` deve exibir as linhas do arquivo que contenham a *string* HEAT, neste caso, apenas uma que diz:

```
FINAL HEAT OF FORMATION =      XXXX.XXXX KCAL/MOL =  YYYY.YYYY KJ/MOL
```

Onde os valores XXXX.XXXX e YYYY.YYYY representam os calores de formação em kcal/mol e kJ/mol respectivamente. É importante que, ao fazer as análises, tenha-se o cuidado de escolher os valores de mesmas unidades para todas as estruturas. Em geral, o valor de calor de formação é negativo.

Complete a tabela a seguir com os valores calculados de entalpia de formação e procure aquela estrutura dentro de cada grupo que tem a menor entalpia de formação. Estas são as estruturas nativas.

Estrutura	$\Delta\tilde{H}_f$ (kcal/mol)		
	1CTF	1R69	1SN3
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

Da mesma forma que no exercício 2, na pasta “~/exercicio3/” existem dois *scripts* (`run_all_ex3.sh` e `read_heats_ex3.py`) para se automatizar as execuções e coleta dos dados de  $\Delta\tilde{H}_f$ . No caso deste exercício, deve-se editar estes *scripts* para adaptar as novas configurações de pasta e arquivos.

Esse exercício também pode ser executado perfeitamente pelo Google Colab. Existe uma versão dele no link:

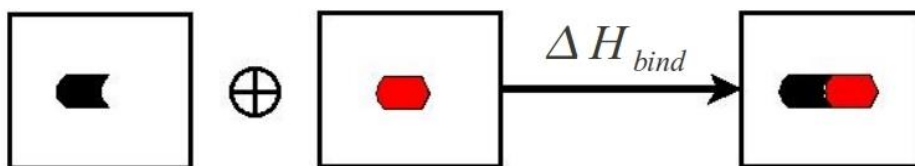
- [https://github.com/RochaGerd/Chemistry\\_with\\_Python](https://github.com/RochaGerd/Chemistry_with_Python)

## EXERCÍCIO 4

### ENTALPIA DE INTERAÇÃO PROTEÍNA-LIGANTE ( $\Delta\tilde{H}_{bind}$ ) (PARTE 1)

Neste exercício, veremos como estimar a variação de entalpia de um sistema devido à interação entre uma proteína e um ligante, em solução, o  $\Delta\tilde{H}_{bind}$ .

Para uma transformação qualquer, a variação da entalpia é calculada como sendo a soma da entalpia dos produtos menos a soma da entalpia dos reagentes. Por exemplo, para a transformação abaixo:



$$\Delta\tilde{H}_r = \sum_{i=1}^{prod.} \Delta H_f^{(i)} - \sum_{j=1}^{reag.} \Delta H_f^{(j)}$$

No caso da variação de entalpia da reação de *binding* entre um ligante e um receptor (enzima) para formação do complexo enzima-ligante teremos:

$$\Delta\tilde{H}_{bind} = \Delta\tilde{H}_f^{(complexo)} - \left( \Delta\tilde{H}_f^{(ligante)} + \Delta\tilde{H}_f^{(enzima)} \right)$$

A entalpia de formação (FINAL HEAT OF FORMATION) pode ser conseguida com o MOPAC, seguindo os procedimentos detalhados anteriormente.

#### NOTA:

Note que variação de entalpia não significa dizer que a reação ocorre, ou não. Em alguns casos quando a entropia não tem uma contribuição muito expressiva, este pode ser o caso, mas em geral a variável correta nesse caso é a variação da energia livre de Gibbs de *binding* ( $\Delta\tilde{G}_{bind}$ ), cujo cálculo é mais complexo e não será abordado aqui.

Adicionalmente, as estruturas apresentadas aqui são meramente para fins didáticos e não foram equilibradas com métodos apropriadamente quantitativos de dinâmica molecular e podem não ser estruturas apropriadamente representativas de *ensembles* termodinâmicos bem definidos. A neutralidade do sistema não foi considerada e nenhum cálculo de equilíbrio em solvente explícito foi realizado. Esse aspecto será explorado no exercício 5, que é a parte 2 deste exercício 4.

## CÁLCULO DAS ENTALPIAS DE FORMAÇÃO

Na pasta do exercício devem haver quatro arquivos do tipo PDB: `2az8.pdb`, `alvo.pdb`, `complexo.pdb` e `ligante.pdb`. O primeiro é a estrutura original do complexo enzima-ligante em questão. Foi colocada nessa pasta por motivos de possíveis comparações. Os três últimos, são o que os nomes indicam, que na sequência são: a enzima sem o ligante, o complexo enzima-ligante e o ligante isolado.

Agora crie um arquivo do tipo “.mop” para cada molécula. Esses inputs serão os sistemas solvatados que para os quais usaremos as seguintes palavras-chave:

```
PM7 1SCF MOZYME XYZ PL T=1D EPS=78.4 RSOLV=1.3 PDB +  
CUTOFF=9.0 DISP(1.0)
```

Execute o MOPAC em cada arquivo de entrada, um de cada vez, usando o comando do exercício anterior:

```
$ nohup mopac <nome-do-arquivo> &
```

Se desejar, acompanhe o progresso do cálculo usando o comando:

```
$ tail -f <nome-do-arquivo>.out | grep "DELTAE"
```

Quando o DELTAE for abaixo de 0.005 kcal/mol, o cálculo terá convergido, e o `tail` vai parar

de atualizar a saída. Você deve sair do `tail` pressionando CTRL+C.

Para se obter o valor da entalpia de formação da estrutura, aplique o comando `cat` como visto no exercício 3:

```
$ cat <nome-do-arquivo>.out | grep "HEAT"
```

Isto irá exibir o resultado final dos cálculos e a grandeza que queremos ver:

```
FINAL HEAT OF FORMATION =      XXXX.XXXX KCAL/MOL =  YYYY.YYYY KJ/MOL
```

Onde os valores XXXX.XXXX e YYYY.YYYY representam os calores de formação em kcal/mol e kJ/mol respectivamente.

Complete a tabela abaixo com os valores calculados de entalpia de formação e calcule a variação da entalpia de interação em solução.

Lembre-se, preste bastante atenção na ordem em que realiza as subtrações. Inverter a ordem significa dizer que você está calculando a variação de entalpia da reação inversa àquela que desejava. Trocar acidentalmente a ordem de subtração das variações é uma fonte bastante comum de erro.

Molécula	$\Delta\tilde{H}_f$ (kcal/mol)
alvo	
ligante	
complexo	
$\Delta\tilde{H}_{bind}$ (kcal/mol)	

Esse exercício também pode ser executado perfeitamente pelo Google Colab. Existe uma versão dele no link:

- [https://github.com/RochaGerd/Chemistry\\_with\\_Python](https://github.com/RochaGerd/Chemistry_with_Python)

## EXERCÍCIO 5

### ENTALPIA DE INTERAÇÃO PROTEÍNA-LIGANTE ( $\Delta\tilde{H}_{bind}$ ) (PARTE 2)

Neste exercício, mais uma vez veremos como estimar a variação de entalpia de um sistema devido à interação entre uma proteína e um ligante, em solução. No caso desse exercício vamos tentar reproduzir os dados do trabalho de Santos, A.R. e colaboradores<sup>9</sup>.

No estudo em questão<sup>9</sup>, Santos, A.R. e colaboradores avaliaram as interações entre a subunidade da toxina ricina A (RTA) e um conjunto de seis inibidores (ver figura a seguir) usando métodos de química quântica semiempíricos e de uma estratégia híbrida QM/MM-ONIOM. Os pesquisadores calcularam entalpias de ligação ( $\Delta\tilde{H}_{bind}$ ) e descritores químicos quânticos de reatividade. Eles compararam esses cálculos com os dados de concentração inibitória ( $IC_{50}$ ). As geometrias para todos os complexos RTA-ligante foram obtidas a partir de simulações clássicas de dinâmica molecular em meios aquosos<sup>9</sup>.

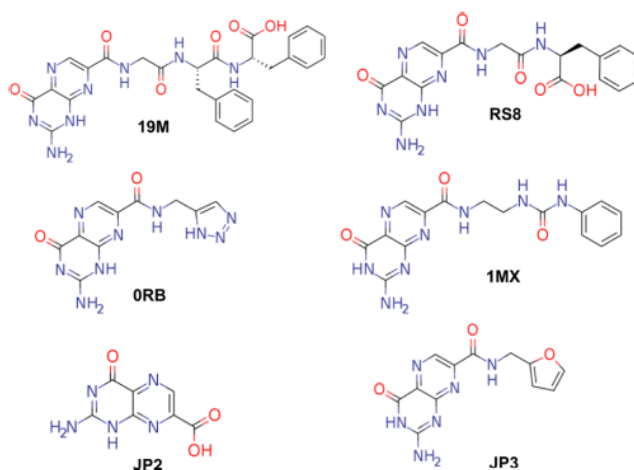


Figura – Estruturas dos seis inibidores da RTA da Ricina<sup>9</sup>.

Os resultados mostraram que os cálculos de energia *single-point* de  $\Delta\tilde{H}_{bind}$  com os métodos semiempíricos PM6-DH+, PM6-D3H4 e PM7, e ONIOM QM/MM, correlacionaram-se bem com os dados de  $IC_{50}$ . No entanto, a correlação diminuiu significativamente quando  $\Delta\tilde{H}_{bind}$  foi calculado após a otimização da geometria completa com todos os métodos semiempíricos. Para uma discussão aprofundada desses detalhes de qual geometria utilizar para um cálculo quântico de  $\Delta\tilde{H}_{bind}$  pode ser vista no trabalho de Ryde, U. e colaborador<sup>10</sup>.

Os cálculos dos descritores de reatividade usando o programa PRIMoRDiA<sup>11-14</sup> indicaram que a sobreposição molecular e as interações eletrostáticas são importantes para a afinidade entre os ligantes e

o *pocket* de ligação da RTA. As estruturas eletrônicas desses complexos enzima-ligante serão avaliadas nos próximos exercícios deste minicurso<sup>10</sup>.

Informações adicionais sobre estudos teóricos/computacionais envolvendo a enzima RTA da ricina podem ser vistas nos estudos<sup>15,16</sup>.

## CÁLCULO DAS ENTALPIAS DE FORMAÇÃO

Na pasta do exercício 5 deve haver seis pastas (0RB\_complexo, 1MX\_complexo, 19M\_complexo, JP2\_complexo, JP3\_complexo e RS8\_complexo). Em cada uma das pastas existem quatro arquivos: `protein.mop`, `XXX.mop`, `XXX_complex.mop` e `XXX_complex.pdb`, onde XXX é o código de três letras de cada um dos complexos.

Execute o MOPAC em cada arquivo de entrada, um de cada vez, usando o comando do exercício anterior:

```
$ nohup mopac <nome-do-arquivo> &
```

Se desejar, acompanhe o progresso do cálculo usando o comando:

```
$ tail -f <nome-do-arquivo>.out | grep "DELTAE"
```

Quando o DELTAE for abaixo de 0.005 kcal/mol, o cálculo terá convergido, e o `tail` vai parar de atualizar a saída. Você deve sair do `tail` pressionando CTRL+C.

Para se obter o valor da entalpia de formação da estrutura, aplique o comando `cat` como visto no exercício 4:

```
$ cat <nome-do-arquivo>.out | grep "HEAT"
```

Isto irá exibir o resultado final dos cálculos e a grandeza que queremos ver:

```
FINAL HEAT OF FORMATION =      XXXX.XXXX KCAL/MOL =  YYYY.YYYY KJ/MOL
```

onde os valores XXXX.XXXX e YYYY.YYYY representam os calores de formação em kcal/mol e kJ/mol respectivamente.

Complete a tabela a seguir com os valores calculados de entalpia de formação e calcule a variação da entalpia de interação ( $\Delta\tilde{H}_{bind}$ ) em solução.

$$\Delta\tilde{H}_{bind} = \Delta\tilde{H}_f^{(complexo)} - \left( \Delta\tilde{H}_f^{(ligante)} + \Delta\tilde{H}_f^{(enzima)} \right)$$

OBS1: se seu computador tem igual ou mais de 4 cores, dá para submeter dois ou até três cálculos ao mesmo tempo. Para isso, use o comando:

```
$ nohup mopac <nome-do-arquivo> &
```

OBS2: Uma sugestão importante é comprimir os arquivos gerados nesse exercício, pois os arquivos “.out” e “.aux” ocupam bastante espaço em disco. Como um exemplo para isto, você pode executar o comando a seguir na pasta do complexo JP3. Repita isto para todos os outros sistemas.

```
$ tar -cvxf JP3_complexo.tar.gz *.out *.pdb *.mop *.aux *.arc
```

Inibidor da Ricina	$\Delta\tilde{H}_f$ (kcal/mol)			$\Delta\tilde{H}_{bind}$ (kcal/mol)	IC <sub>50</sub>
	Complexo	Ligante	Proteína		
<b>19M</b>					15
<b>RS8</b>					20
<b>0RB</b>					70
<b>1MX</b>					209
<b>JP2</b>					230
<b>JP3</b>					380

Tabela para preenchimento com os dados produzidos no exercício 5. Os valores de IC<sub>50</sub> foram retirados das suas respectivas fontes originais que podem ser obtidas na publicação de Santos, A. R e colaboradores<sup>9</sup>.

Lembre-se, preste bastante atenção na ordem em que realiza as subtrações. Inverter a ordem significa dizer que você está calculando a variação de entalpia da reação inversa àquela que desejava. Trocar acidentalmente a ordem de subtração das variações é uma fonte bastante comum de erro.

Para finalizar, faça o gráfico de correlação com os dados de IC<sub>50</sub> vs  $\Delta\tilde{H}_{bind}$  para ver se essas duas quantidades se correlacionam bem.

Esse exercício também pode ser executado perfeitamente pelo Google Colab. Existe uma versão dele no link:

- [https://github.com/RochaGerd/Chemistry\\_with\\_Python](https://github.com/RochaGerd/Chemistry_with_Python)



## EXERCÍCIO 6

### ANALISANDO A ESTRUTURA ELETRÔNICA DE COMPLEXOS ENZIMA-LIGANTE: DESCRITORES DE REATIVIDADE

No documento de tutoriais do PRIMoRDiA (<https://github.com/igorChem/PRIMoRDiA1.0v>) pode ser encontrado no tutorial de número 3 como executar os cálculos e as análises de descritores quânticos de reatividade para sistemas macromoleculares de relevância biológica, que obrigatoriamente precisam ter sua estrutura tridimensional codificada em um arquivo de PDB. No exemplo do tutorial original, os cálculos exploram todas as opções possíveis para sistemas de teste, polipeptídeos de aproximadamente 300 átomos, que não possuem um interesse de pesquisa significativo.

Neste exercício, vamos usar os resultados produzidos no exercício 5 e calcular descritores de reatividade dos complexos enzima-ligante 19M e JP3. Este exercício é mais uma forma de reproduzir os dados do trabalho de Santos, A.R. e colaboradores<sup>9</sup>.

Para isso, vamos usar o programa PRIMoRDiA<sup>11</sup>, que é de domínio público e pode ser baixado através do endereço (<https://github.com/igorChem/PRIMoRDiA1.0v>). Outra fonte de consulta é o repositório ([https://github.com/bardenChem/XI\\_EMMSB\\_PRIMoRDiA](https://github.com/bardenChem/XI_EMMSB_PRIMoRDiA)).

### PREPARAÇÃO DO INPUT PARA O PRIMoRDiA

Copie os arquivos `19M_complexo.tar.gz` e `JP3_complexo.tar.gz` (salvos no exercício 5) para o diretório de sua escolha. Uma dica é usar os diretórios correspondentes aos referidos complexos do exercício 5, que no caso são: `19M_complexo` e `JP3_complexo`.

Alternativamente, nos diretórios `19M_complexo` e `JP3_complexo` existem arquivos “.aux.gz”, que são arquivos comprimidos dos arquivos “.aux” de cada complexo enzima-ligante. Você pode usá-los, caso não tenha acesso aos arquivos salvos no exercício 5.

Entre em cada um dos diretórios e crie um arquivo de *input* para o PRIMoRDiA como está mostrado a seguir. Salve como `primordia_jp3.inp`. Note que você deve fazer o mesmo para o complexo 19M.

```
#RT normal
```

```
#PR eband 5 extrard pymols
```

```
3 jp3_complex.aux true 0 0 jp3_complex.pdb mopac 0 0 0 0 EW
```

Perceba que você precisa dos arquivos “.pdb” de cada complexo enzima-ligante para rodar apropriadamente o PRIMoRDiA.

## **EXECUTANDO O PRIMoRDiA PARA OS COMPLEXOS RTA-JP3 E RTA-19M**

Para executar o programa PRIMoRDiA para cada um dos complexos, use a seguinte linha de comando.

```
$ PRIMoRDiA_1.25v -f primordia_jp3.inp &
```

Esse comando vai produzir vários arquivos e diretórios.

## **VISUALIZANDO RESULTADOS DO PRIMoRDiA COM O PYMOL PARA OS COMPLEXOS RTA-JP3 E RTA-19M**

Nessa seção, você vai visualizar os resultados de descritores de reatividade usando o programa PyMol.

Para a análise dos resultados no Pymol, primeiro abra o software PyMol e execute o comando na caixa de comandos do PyMol:

```
@19m_complex_pymols_pdb.pym
```

O arquivo com a terminação “.pym” é um script do PyMol que automatiza o carregamento de vários arquivos com o cálculo dos descritores de reatividade por átomo e pré-configurações de paleta de cores. Ao executar esse script, você deve perceber que vários arquivos “.pdb” foram carregados, um para cada descritor de reatividade local calculado pelo programa PRIMoRDiA.

Você deve ativar o objeto “netphilicity” (quarto objeto de cima para baixo) e executar o seguinte comando no PyMol:

```
spectrum b, blue_white_red, minimum=-0.01,maximum=0.01
```

Esse descritor mostra as regiões propensas a receber um ataque nucleofílico nos valores positivos, pintados em tons de vermelho, portanto, demonstra o potencial local de eletrofilicidade desses átomos.

Nos valores negativos, pintados em tons de azul, não aparecem nessa imagem pois os átomos propensos a ataques eletrofilicos, portanto, com um potencial local de nucleofilicidade, não estão destacados nessa imagem.

Depois ative o objeto `fukui_pot_left` (sétimo de cima para baixo) e executar o comando:

```
spectrum b, white_yellow_orange_red_black, minimum=0.15,maximum=0.20
```

Esse comando vai colorir os átomos compatível com os valores da coluna *b-factor*, que no caso do cálculo que foi feito, agora, possui valores do descritor de reatividade `fukui_pot_left`. Assim, cores mais escuras possuem maior valor de `fukui_pot_left`. Você pode “jogar” com os valores de mínimo e máximo para aumentar ou diminuir o contraste dos valores calculados. Para um maior contraste, use valores de mínimo e máximo mais próximos.

Agora, você vai ajustar o tipo de tracejado ao redor da molécula dessa forma:

```
set ray_trace_mode, 1
```

Após, você deve renderizar a molécula em alta qualidade com o comando do PyMol:

```
ray antialias=2
```

Finalmente, você deve salvar a figura renderizada, como:

```
png complex_rta_19m_hardness.png
```

Você deve realizar o mesmo procedimento para o complexo JP3-RTA. Após executar esses dois procedimentos, você deve apontar qual ligante interage melhor com o sítio ativo da RTA, com base nos valores do descritor de reatividade dureza local.

Esse exercício também pode ser executado pelo Google Colab, sem a parte de visualização com o PyMol. Existe uma versão dele no link:

- [https://github.com/RochaGerd/Chemistry\\_with\\_Python](https://github.com/RochaGerd/Chemistry_with_Python)

## EXERCÍCIO 7

### CALCULANDO AS INTERAÇÕES NÃO-COVALENTES PARA SOLVATAÇÃO DE ENZIMAS

Neste exercício vamos explorar cálculos realizados no estudo conduzido por Rocha, R.E.O e colaboradores<sup>17</sup>. O estudo usou uma combinação de técnicas computacionais como simulações de dinâmica molecular, análises farmacofóricas e cálculos de mecânica quântica para comparar os monômeros de protease principal (Mpro) do SARS-CoV e do SARS-CoV-2. O estudo revelou diferenças importantes em seu movimento, *druggability* e atividade catalítica, oferecendo *insights* sobre potenciais estratégias de desenvolvimento de medicamentos para a COVID-19.

Assim, neste exercício será possível calcular as superfícies do NCI (*Non-Covalent Interactions*)<sup>18</sup> na interface  $M^{\text{Cov1}}_{\text{Pro}}\text{-água}$  e  $M^{\text{Cov2}}_{\text{Pro}}\text{-água}$ . Logo, serão mapeadas as interações intermoleculares presentes nos sistemas.

Os cálculos vão ser feito dentro do diretório `exercicio7/Sars_Cov_1_e_2/`. Acesse o diretório `exercicio7/Sars_Cov_1_e_2/`

```
$ cd exercicio7/Sars_Cov_1_e_2/
```

Você pode observar que existem arquivos “.nci” que são os *inputs* para o NCIPLOT das conformações da  $M^{\text{Cov1}}_{\text{Pro}}$  fechada, semiaberta e aberta, respectivamente, A, J e N, para a  $M^{\text{Cov2}}_{\text{Pro}}$  são os arquivos B, G e O. Os arquivos “.xyz” contém as coordenadas dos sistemas em estudo e os arquivos “gnuplot” serão utilizados para gerar gráficos de cada sistema.

O comando para executar o cálculo de NCIPLOT é mostrado na sequência.

```
$ nciplot A.nci > A.out
```

É possível observar o resultado desse comando utilizando o programa de visualizações, o VMD. Ao final do cálculo serão gerados arquivos “.vmd”. Use o seguinte comando para visualizar os resultados:

```
$ vmd -e A.vmd
```

Após terminar os cálculos do NCIPLOT, execute o GNUPLOT com o seguinte comando para

cada sistema calculado:

```
$ gnuplot A_gnuplot_nci
```

É possível observar o resultado utilizando o Okular, para isso, utilize o comando:

```
$ okular A.png
```

Use o comando mostrado na sequência para unir as imagens dos resultados e facilitar sua análise e visualizar o arquivo unidas.png:

```
$ montage A.png J.png N.png B.png G.png O.png -tile 2x3 -geometry +5+5 \
        unidas.png
```

Por fim, tente visualizar os arquivos “.vmd” e “unidas.png” de modo simultâneo e diante dos resultados observados, marque as alternativas corretas nas questões 1 e 2 e responda a pergunta da questão 3:

- 1) Qual a principal diferença nas interações com o solvente entre os modelos  $M^{\text{Cov}^1}_{\text{Pro}}$  e  $M^{\text{Cov}^2}_{\text{Pro}}$ ?
  - a) Ambos os modelos apresentam interações favoráveis com o solvente nas conformações estudadas.
  - b) A  $M^{\text{Cov}^2}_{\text{Pro}}$  favorece interações hidrofóbicas, enquanto o  $M^{\text{Cov}^1}_{\text{Pro}}$  favorece interações polares.
  - c) A  $M^{\text{Cov}^2}_{\text{Pro}}$  favorece interações polares, enquanto o  $M^{\text{Cov}^1}_{\text{Pro}}$  favorece interações hidrofóbicas. Sendo principalmente mais presente esta diferença quando comparadas as conformações abertas e fechadas.
  - d) Não há diferença significativa nas interações com o solvente entre os dois modelos.

Resposta:

- 2) Qual das seguintes afirmações sobre as conformações dos modelos é a mais correta?
  - a) Ambas as conformações são igualmente favoráveis.
  - b) A  $M^{\text{Cov}^2}_{\text{Pro}}$  favorece conformações mais fechadas.

c) A  $M^{\text{Cov1}}_{\text{Pro}}$  favorece conformações mais abertas.

d) A  $M^{\text{Cov2}}_{\text{Pro}}$  favorece conformações mais abertas.

Resposta:

3) Por que o Sars-Cov-2 era tão transmissível em relação ao Sars-Cov-1?

Resposta:

## **FEEDBACK do Minicurso apresentado no X SeedMol**

Agradecemos se você puder emitir sua impressão sobre o minicurso. Você pode acessar um formulário no link a seguir:

[https://docs.google.com/forms/d/e/1FAIpQLSfDG19IfrM4H-62gRcNP017g6aihaJuy022ARTrFLVD7\\_kAOA/viewform?usp=sf\\_link](https://docs.google.com/forms/d/e/1FAIpQLSfDG19IfrM4H-62gRcNP017g6aihaJuy022ARTrFLVD7_kAOA/viewform?usp=sf_link)

## REFERÊNCIAS

1. Rocha GB, Freire RO, Simas AM, Stewart JJP. RM1: A reparameterization of AM1 for H, C, N, O, P, S, F, Cl, Br, and I. *J Comput Chem*. 2006;27(10):1101-1111. doi:10.1002/jcc.20425
2. Lima N, Rocha G, Freire R, Simas A. RM1 Semiempirical Model: Chemistry, Pharmaceutical Research, Molecular Biology and Materials Science. *J Braz Chem Soc*. 2018;30(4):683-716. doi:10.21577/0103-5053.20180239
3. Maia JDC, Urquiza Carvalho GA, Manguiera Jr. CP, Santana SR, Cabral LAF, Rocha GB. GPU linear algebra libraries and GPGPU programming for accelerating MOPAC semiempirical quantum chemistry calculations. *J Chem Theory Comput*. 2012;8(9):3072-3081. doi:10.1021/ct3004645
4. Stewart JJP. MOPAC web site. openmopac.net. Published 2022. Accessed March 30, 2022.
5. Stewart JJP. Application of localized molecular orbitals to the solution of semiempirical self-consistent field equations. *Int J Quantum Chem*. 1996;58(2):133-146. doi:10.1002/(SICI)1097-461X(1996)58:2<133::AID-QUA2>3.0.CO;2-Z
6. Stewart JJP. Optimization of parameters for semiempirical methods VI: more modifications to the NDDO approximations and re-optimization of parameters. *J Mol Model*. 2013;19(1):1-32. doi:10.1007/s00894-012-1667-x
7. Urquiza-Carvalho GA, Fragoso WD, Rocha GB. Assessment of semiempirical enthalpy of formation in solution as an effective energy function to discriminate native-like structures in protein decoy sets. *J Comput Chem*. 2016;37(21):1962-1972. doi:10.1002/jcc.24415
8. Berman HM, Westbrook J, Feng Z, et al. The Protein Data Bank. *Nucleic Acids Res*. 2000;28:235-242.
9. Rocha-Santos A, Chaves EJF, Grillo IB, de Freitas AS, Araújo DAM, Rocha GB. Thermochemical and Quantum Descriptor Calculations for Gaining Insight into Ricin Toxin A (RTA) Inhibitors. *ACS Omega*. 2021;6(13):8764-8777. doi:10.1021/acsomega.0c02588
10. Ryde U, Söderhjelm P. Ligand-Binding Affinity Estimates Supported by Quantum-Mechanical Methods. *Chem Rev*. 2016;116(9):5520-5566. doi:10.1021/acs.chemrev.5b00630
11. Grillo IB, Urquiza-Carvalho GA, Rocha GB. PRIMoRDiA: A Software to Explore Reactivity



and Electronic Structure in Large Biomolecules. *J Chem Inf Model.* 2020;60(12):5885-5890. doi:10.1021/acs.jcim.0c00655

12. Grillo IB, Rocha GB. PRIMoRDiA. 2020. <https://github.com/igorChem/PRIMoRDiA1.0v>.
13. Grillo IB, Urquiza-Carvalho GA, Rocha GB. Quantum chemical descriptors as a modeling framework for large biological structures. In: *Chemical Reactivity*. Elsevier; 2023:59-88. doi:10.1016/B978-0-32-390259-5.00009-3
14. Grillo IB, Urquiza-Carvalho GA, Rocha GB. Quantum chemical descriptors based on semiempirical methods for large biomolecules. *J Chem Phys.* 2023;158(20). doi:10.1063/5.0132687
15. Chaves EJJ, Padilha IQM, Araújo DAM, Rocha GB. Determining the Relative Binding Affinity of Ricin Toxin A Inhibitors by Using Molecular Docking and Nonequilibrium Work. *J Chem Inf Model.* 2018;58(6):1205-1213. doi:10.1021/acs.jcim.8b00036
16. Chaves EJJ, Gomes da Cruz LE, Padilha IQM, Silveira CH, Araujo DAM, Rocha GB. Discovery of RTA ricin subunit inhibitors: a computational study using PM7 quantum chemical method and steered molecular dynamics. *J Biomol Struct Dyn.* February 2021:1-19. doi:10.1080/07391102.2021.1878058
17. Rocha REO, Chaves EJJ, Fischer PHC, et al. A higher flexibility at the SARS-CoV-2 main protease active site compared to SARS-CoV and its potentialities for new inhibitor virtual screening targeting multi-conformers. *J Biomol Struct Dyn.* May 2021:1-21. doi:10.1080/07391102.2021.1924271
18. Boto RA, Peccati F, Laplaza R, et al. NCIPLoT4: Fast, Robust, and Quantitative Analysis of Noncovalent Interactions. *J Chem Theory Comput.* 2020;16(7):4150-4158. doi:10.1021/acs.jctc.0c00063

## **Links úteis**

- <https://www.embnet.org/wp/quick-guides/>
- <https://chemozart.com>
- <https://web.chemdoodle.com/demos/2d-to-3d-coordinates/>
- [http://uspex-team.org/online\\_utilities/zmatrix/](http://uspex-team.org/online_utilities/zmatrix/)
- <http://www.cheminfo.org/Chemistry/Cheminformatics/FormatConverter/index.html>.
- <http://cdb.ics.uci.edu/cgi-bin/BabelWeb.py>
- [http://openmopac.net/PM7\\_accuracy/normal%20molecules.html](http://openmopac.net/PM7_accuracy/normal%20molecules.html)
- <https://pymolwiki.org/index.php/PLoS>
- [https://pymolwiki.org/index.php/Main\\_Page](https://pymolwiki.org/index.php/Main_Page)