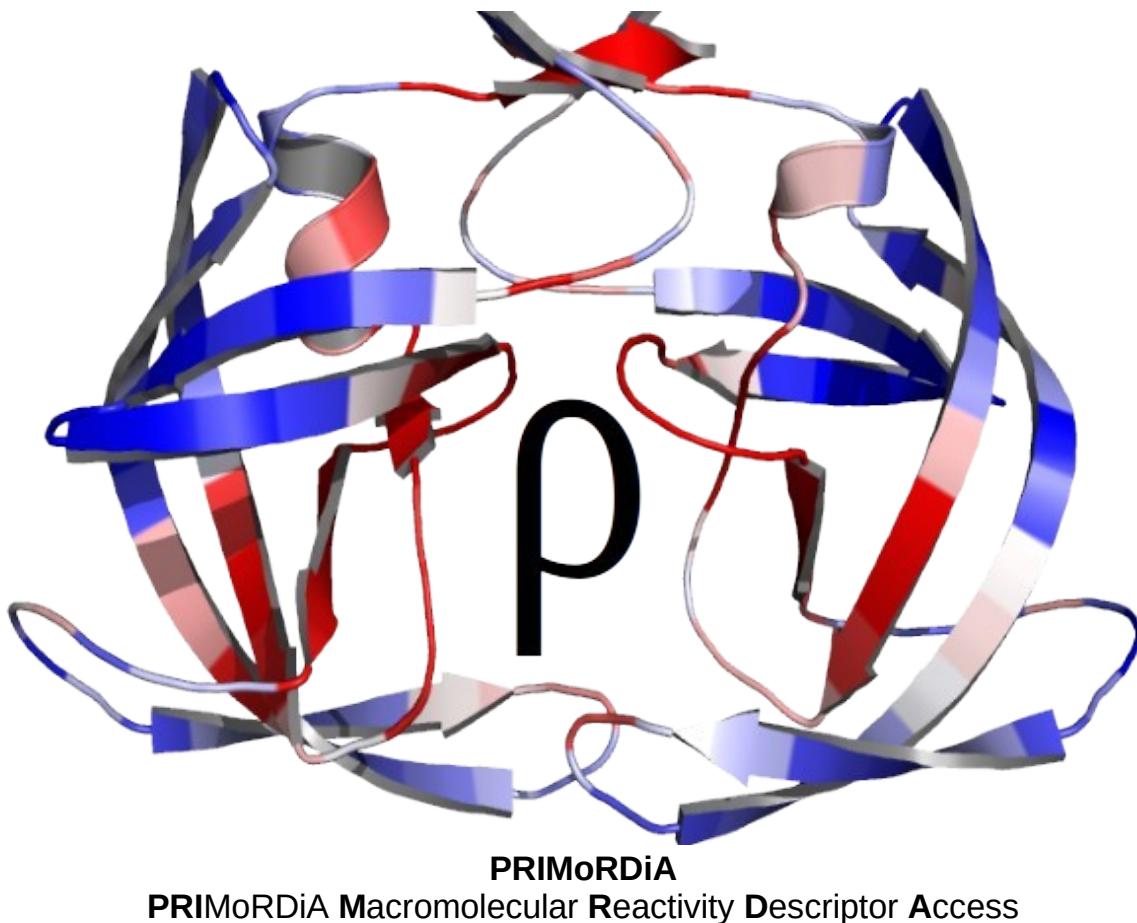

PRIMoRDiA Tutoriais: XI EMMSB

Igor Barden Grillo
(barden.igor@gmail.com)
github.com/bardenChem

15 de agosto de 2024



Sumário

1 Tutorial 1: Descritores de Banda - Aplicação para Ligante Proteína	4
1.1 Arquivos Necessários e Execução do Programa	4
1.2 Análise dos Resultados	6
2 Tutorial 2: Gerando Trajetória de Catálise Enzimática	14
2.1 O Sistema: Triose-Fosfato Isomerase	14
2.2 pDynamo e pDynamo_Scripts	15
2.3 Análise dos Resultados das Simulações	18
3 Tutorial 3: Análise de Reatividade de Caminho de Reação Enzimático	19
3.1 Análise dos Resultados	22
4 Considerações Finais	31

Introdução

Esse documento contém os tutoriais para o programa PRIMoRDiA direcionados para a XI edição da Escola de Modelagem Molecular de Sistemas Biológicos (EMMSB) do Laboratório Nacional de Computação Científica (LNCC).

O PRIMoRDiA (**PRIMoRDiA Macromolecular Reactivity Descriptors Access**) é um software escrito em *C++*, desenvolvido para o pós-processamento eficiente dos resultados de pacotes de química computacional, produzindo uma grande variedade de descritores quânticos: propriedades eletrônicas e reatividade de sistemas químicos.

O PRIMoRDiA traz implementado os principais descritores de reatividade da Teoria do DFT conceitual, incluindo as definições matemáticas para as teorias de elétrons de fronteira de Fukui e de ácidos e bases duros e moles de Pearson. O PRIMoRDiA foi desenvolvido com foco no cálculo de descritores e outras propriedades eletrônicas mais comuns para grandes moléculas que são relevantes para processos biológicos, e por isso há descritores modificados especificativamente para atender as particularidades desses sistemas.

Também oferecemos um guia de usuário com informações gerais sobre a teoria e detalhes sobre o uso e interpretação dos resultados gerados pelo programa. PRIMoRDiA foi desenvolvido na Universidade Federal da Paraíba, no Laboratório de Química Quântica e Computacional. Nossa grupo possui uma série de trabalhos já publicados, como a referência principal do programa , aplicações para a caracterização teórica de reações de catálise enzimática [1], uma revisão sobre a exploração da estrutura eletrônica de macromoléculas biológicas usando métodos semi-empírico para os cálculos quânticos[2], aplicações para ligante e proteína[3, 4], e outras aplicações[5, 6].

Os tutoriais nesse documento cobrem as funcionalidades relacionadas com o estudo da estrutura eletrônica de macromoléculas, mais especificamente proteínas, e a análise teórica de reatividade de reações enzimáticas, assim como também a obtenção de um caminho de reação enzimático para realizar o mesmo.

Utilizar o PRIMoRDiA é muito simples, consiste geralmente em duas etapas, sendo a primeira a montagem do arquivo de input, no qual o programa também ajuda a fazer, e a segunda a parte de execução dos cálculos. Por isso, os tutorias aqui apresentados tem com objetivo maior de apresentar os resultados do PRIMoRDiA e como utilizar esses dados para aplicar em problemas de pesquisa.

1 Tutorial 1: Descritores de Banda - Aplicação para Ligante Proteína

No documento de tutoriais do PRIMoRDiA encontramos 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 interesse de pesquisa significativo. No presente tutorial, serão aplicados os cálculos em complexos entre a enzima RTA (Ricin Toxic A chain) e ligantes que tem potencial inibidor já descritos experimentalmente. Mais precisamente, nesse tutorial vamos recriar resultados teóricos de um dos artigos de pesquisa do nosso grupo[4].

1.1 Arquivos Necessários e Execução do Programa

Nesse tutorial aproveitaremos os resultados de cálculos semi-empíricos dos tutoriais anteriores desse minicurso, onde vamos fazer os cálculos do PRIMoRDiA para os complexos de RTA com o ligante 19M e JP3. Para cada um desses arquivos tem um arquivo de PDB, que contém as informações relevantes ao biopolímero que vai ser utilizado pelo PRIMoRDiA para escrever descritores e representações especiais.

Para facilitar a produção de input a partir desses arquivos vamos usar o seguinte comando, na pasta onde contém os arquivos ".aux" dos complexos que serão analisados

```
computer@user:$ path/to/PRIMoRDiA_1.25v -input -op 3 -p mopac
```

Um arquivo de chamado de primordia.input é gerado na pasta. A primeira *keyword* encontrada no arquivo é "#RT normal" para indicar que o Run Type (tipo de cálculo) é do tipo comum esperado para a execução do PRIMoRDiA. O keyword "#PR" indica os que as próximas entradas na linha são parâmetros especiais, que no caso, por hora, é "eband" e "1", indicando que o limite para considerar orbitais moleculares em relação HOMO-LUMO é de 1eV. As duas linhas que começam com o inteiro "3", indicam a opção de cálculo do PRIMoRDiA especial para macromoléculas. Vamos editar o arquivo em um editor de texto para conter as informações como mostrado na caixa de texto abaixo

Listing 1: Input editado para execução do tutorial 3

```
#RT normal
#PR eband 1 extrard pymols
3 jp3_complex.aux true 0 0 jp3.pdb mopac 0 0 0 0 EW
3 19m_complex.aux true 0 0 19m_complex.pdb mopac 0 0 0 0 EW
```

Esses argumentos são detalhadamente explicados no guia do usuário, mas também são descritos um por um na lista abaixo

- "3": Indica a opção de cálculo dos descritores para o PRIMoRDiA
- "jp3_complex.aux": Nome do arquivo de output que contém a informação de estrutura eletrônica
- "true": Keyword para opção de dureza local para ser calculada na versão volumétrica
- "0": Granularidade/resolução dos arquivos de cube a serem gerados para os descritores volumétricos, zero indica que esses cálculos de grid não serão gerados.
- "0": Número máximo de orbitais moleculares a ser utilizados pelo método BD
- "jp3_coplex.pdb": Nome do PDB com as informações de referência
- "0": Coordenada do eixo x para o centro da caixa a ser gerado para os descritores volumétricos
- "0": Coordenada do eixo y para o centro da caixa a ser gerado para os descritores volumétricos
- "0": Coordenada do eixo z para o centro da caixa a ser gerado para os descritores volumétricos
- "0": Tamanho do lado da caixa a ser gerada a partir do centro dado pelas coordenadas acima
- "EW": Keyword para sinalizar o método de cálculo de descritores para macromoléculas ponderado por energia

Para a produção de arquivos cube, com dados volumétricos, é possível limitar uma área de cálculo para obter uma definição maior com menor custo computacional. Para isso é necessário definir um caixa onde o grid vai ser calculado, o centro e o tamanho dados no arquivo de input, como foi indicado acima. Nos dados originais da publicação não utilizamos os arquivos cube, por ser mais útil mapear esses descritores por átomo devido a quantidade de elementos para serem representados nas imagens.

O PRIMoRDiA tem implementado dois métodos para combinar esse orbitais moleculares:

1. Band Density (BD: sigla e Keyword utilizada dentro do programa)
2. Energy weighted (EW: sigla e Keyword utilizada dentro do programa)

O primeiro é inspirado no cálculo de densidade eletrônica, com a particularidade de considerar somente uma seleção de orbitais

moleculares a partir do HOMO, para os ocupados, e LUMO para os virtuais. O segundo é uma combinação que é ponderada pela diferença de energia dos orbitais moleculares dos orbitais de fronteira HOMO-LUMO. Há um limite para considerar esses orbitais que é definido por um valor de corte de energia definido pelo usuário.

Nesse exemplo, vamos rodar os dois sistemas com o método *EW*, que foi utilizado na publicação do nosso grupo, mas como desafio, esses cálculos podem ser feitos com o método *BD*. Finalmente, com esse input podemos colocar o programa para rodar com o seguinte comando

```
computer@user:$ path/to/PRIMORDiA_1.25v -f primordia_editado.input
```

Usamos o arquivo "primordia_editado.input", que foi editado a partir do arquivo automático gerado pelo software. O programa deve finalizar a sua execução sem erros e produzir vários arquivos no diretório, com extensão ".cube", ".pym", ".lrd" e etc. Vamos explicar os arquivos com os resultados de descritores e como analisá-los na próxima seção desse tutorial.

1.2 Análise dos Resultados

Dos resultados produzidos temos os descritores globais, locais em representação condensada (arquivos ".lrd"), locais em representação volumétrica (arquivos ".cube", não gerados nesse tutorial) e locais condensados escritos em PDB (nas pastas "RD_PDB"). Para macromoléculas é interessante observar os descritores globais por uma questão de acompanhamento da normalidade do cálculo quântico executado, pois os processos nesses sistemas tendem a ser somente locais, tendo pouca relevância a propensão de uma proteína inteira de transferir ou receber elétrons. Por isso não vamos analisar do arquivo "primordia.global".

Para esse lançamento estável do programa, PRIMORDiA 1.25v, a principal modificação é a consolidação dos scripts de automatização de análise e produção de resultados com qualidade de publicação, usando o pacote gráfico Pymol e o pacote estatístico R. Por isso, nos próximos passos do tutorial, vamos focar na produção de imagens a partir desses scripts.

No entanto, esse tipo de representação pode ser bem complicada de trabalhar, já que é uma lista de valores por átomos, ou seja, uma tabela que pode ser significativamente grande para macromoléculas. Pensando em facilitar a análise desses resultados, o PRIMORDiA escreve um arquivo de PDB para cada descritor calculado com os valores no campo do b-factor, e assim podemos usar o pacote gráfico Pymol para visualizar os descritores colorindo os átomos com uma paleta customizada. Para operar o Pymol mais automaticamente possível, o PRIMORDiA escreve um arquivo de script com o sufixo

"__pymols_pdb.pym", com um exemplo de sua estrutura mostrado na Figura 1.

```
19m_complex_pymols_pdb.pym ✘
1 preset.publication(selection='all')
2 set sphere_scale, 0.2
3 set bg_rgb, white
4 set stick_radius, 0.18
5 load 19m_complex_PDB RD/19m_complex_nucleophilicity.pdb
6 load 19m_complex_PDB RD/19m_complex_electrophilicity.pdb
7 load 19m_complex_PDB RD/19m_complex_radicality.pdb
8 load 19m_complex_PDB RD/19m_complex_netphilicity.pdb
9 load 19m_complex_PDB RD/19m_complex_hardness_Vee.pdb
10 load 19m_complex_PDB RD/19m_complex_hardness_lcp.pdb
11 load 19m_complex_PDB RD/19m_complex_fukui_pot_left.pdb
12 load 19m_complex_PDB RD/19m_complex_fukui_pot_right.pdb
13 load 19m_complex_PDB RD/19m_complex_fukui_pot_zero.pdb
14 load 19m_complex_PDB RD/19m_complex_softness_dual.pdb
15 load 19m_complex_PDB RD/19m_complex_hyper_softness.pdb
16 load 19m_complex_PDB RD/19m_complex_fukushima.pdb
17 load 19m_complex_PDB RD/19m_complex_mep.pdb
18 load 19m_complex_PDB RD/19m_complex_hardness_TFD.pdb
19 load 19m_complex_PDB RD/19m_complex_softness_avg.pdb
20 load 19m_complex_PDB RD/19m_complex_hardness_int.pdb
21 load 19m_complex_PDB RD/19m_complex_multiphilicity.pdb
22 load 19m_complex_PDB RD/19m_complex_charge.pdb
23 load 19m_complex_PDB RD/19m_complex_electron_density.pdb
24 spectrum b, blue_white_red, minimum=-0.5, maximum=0.5
25 spectrum b, white_yellow_orange_red_black, minimum=0.1, maximum=0.5
26 spectrum b, white_cyan_blue, minimum=0, maximum=0.1
27 spectrum b, white_magenta_purple_purpleblue_black, minimum=1, maximum=5
28 spectrum b, white_pink_red, minimum=0, maximum=0.1
29
```

Figura 1: Script de automação da visualização dos descritores locais condensados no Pymol escrita no PRIMORDiA e aberta em editor de texto.

Esse arquivo tem comandos para executar no Pymol modificando parâmetros de visualização e para carregar os PDBs com os descritores. Além disso ele já executa algumas paletas de cores, sendo que uma vai sobrepondo a outra, mas servem para o usuário pegar e escolher uma ou modificar seus parâmetros e cores. Vamos começar abrindo o seu Pymol, execute o comando mostrado a baixo

```
@19m_pymols_pdb.pym
```

Depois da execução desse comando, a janela do pymol deve abrir e carregar todos os arquivos de PDB e executar todos os comandos que configuram a paleta de cores baseada nos valores dos descritores (Figura 2) . O último comando de "spectrum b..." executada deve ser o que fica aparecendo, como na Figura 3.

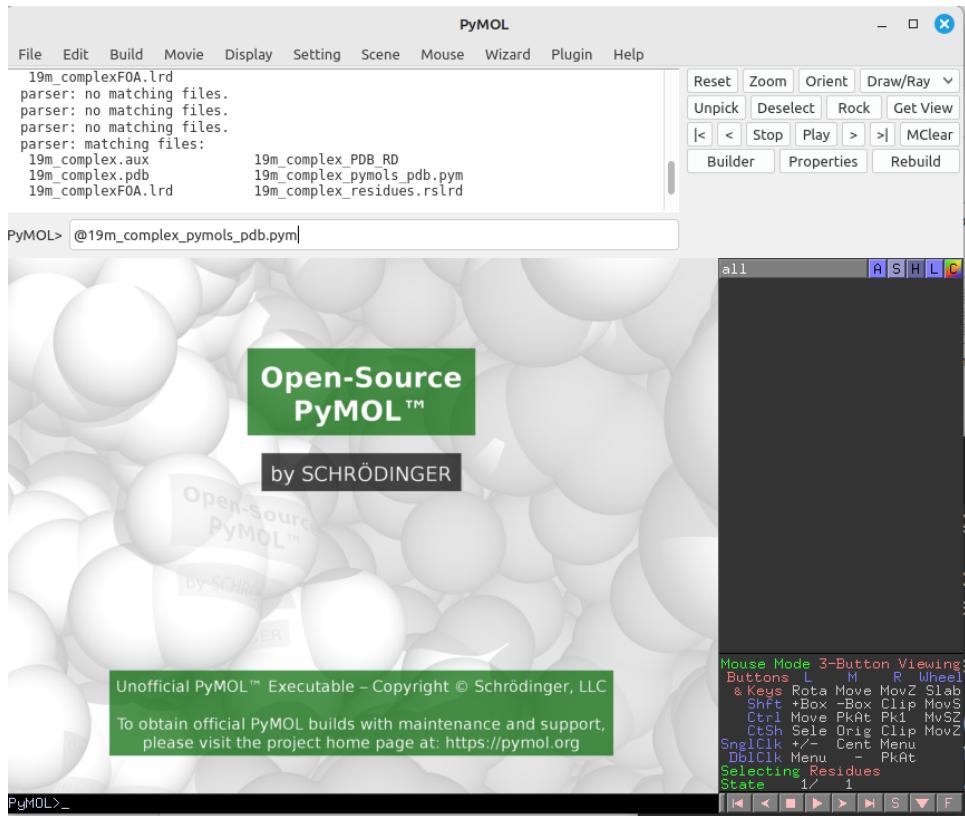


Figura 2: Janela do Pymol com o comando para executar

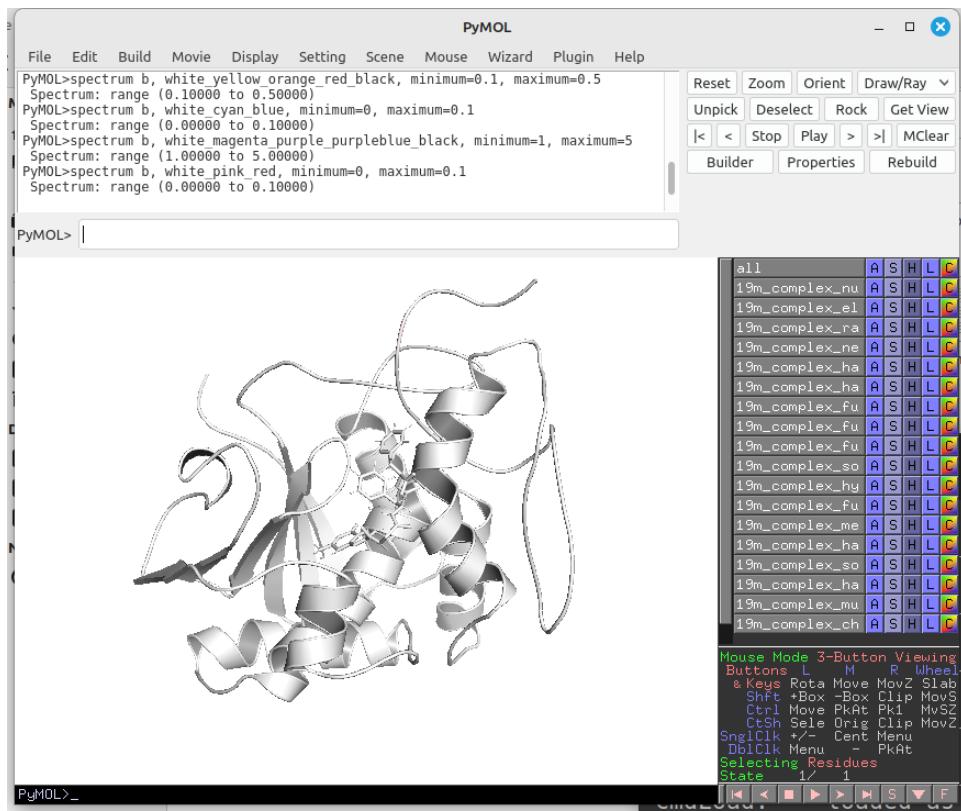


Figura 3: Descritores locais condensados para a estrutura do complexo 19M carregados no Pymol.

Na próxima imagem nós mostramos o Pymol com somente um dos objetos ativos, o 19M_complex_netphilicity, e mostrando o comando a ser executado no terminal, que foi selecionado/copiado a partir de um daqueles que foi executado quando o script foi chamado.

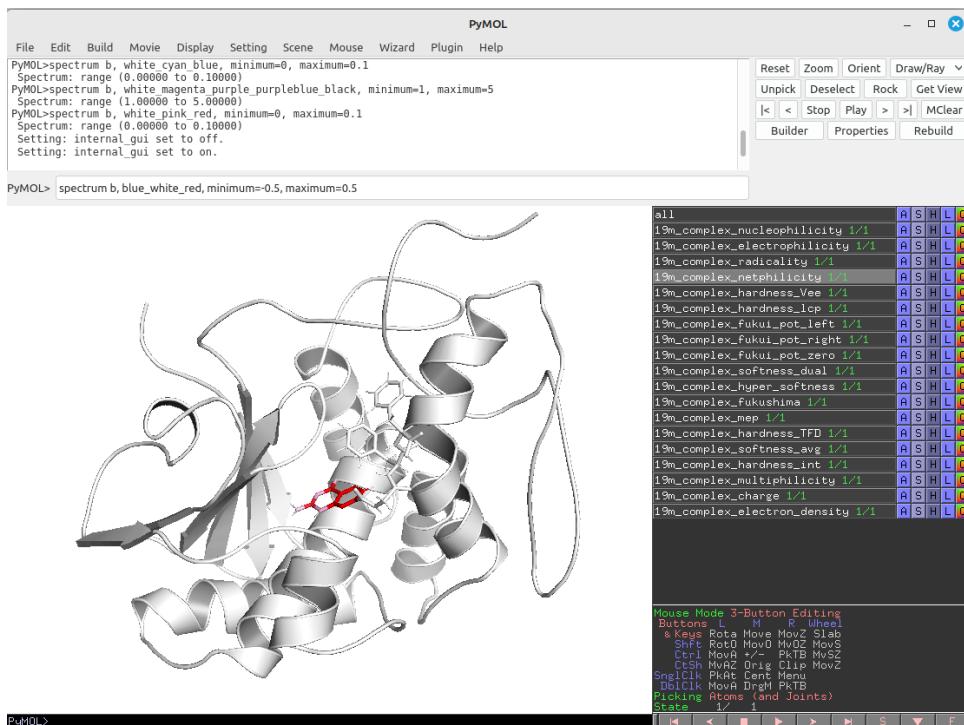


Figura 4: Janela do Pymol mostrando o objeto 19M_complex_netphilicity ativo e o comando para gerar a paleta de cores.

Esse descriptor 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 eletrofílicos, portanto, com um potencial local de nucleofilicidade, não estão destacados nessa imagem. Mas, para chegar em uma imagem semelhante a que foi para o artigo publicado, vamos analisar para esse complexo o descriptor de dureza local, mais precisamente com a aproximação de Fukui-à-esquerda, que é a função que representa a porção da densidade eletrônica mais propensa a ser transferida e que substitui a densidade eletrônica total no cálculo de dureza.

Na Figura 5, é mostrado a janela com o objeto com o descriptor de dureza mencionado já selecionado e com a paleta de cores que é geralmente escolhida pelo nosso grupo. Vamos explorar o comando mostrado na tela para gerar uma imagem com uma visualização ainda mais aprimorada que a figura que foi publicado no estudo de 2021.

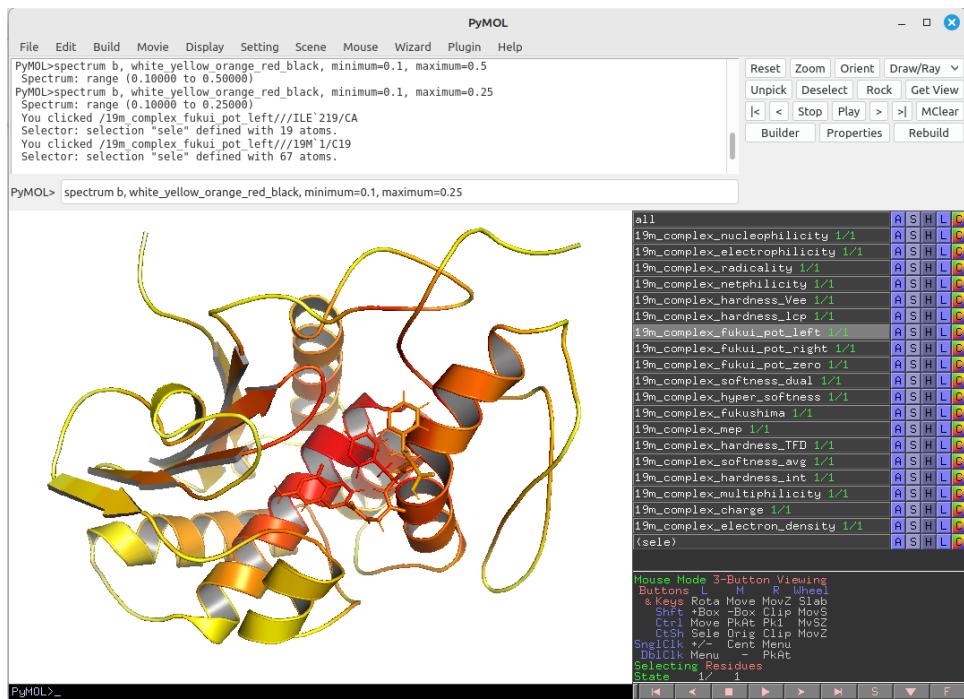


Figura 5: Janela do Pymol mostrando o objeto de potencial de Fukui Left ativo e o comando para edição da paleta de cores mostrada no campo de comando.

Vamos selecionar o ligante e ativar a visualização das esferas Figura 6 e ajustar as cores, usando como mínimo o valor de 0.15 e máximo de 0.23. Isso vai fazer com que se destaque mais a diferença entre os átomos e distinguir os mais propensos a interações duro-duro. Também, para aprimorar a renderização da imagem final, vamos mudar o modo de tracejado com o seguinte comando

```
set ray_trace_mode=1
```

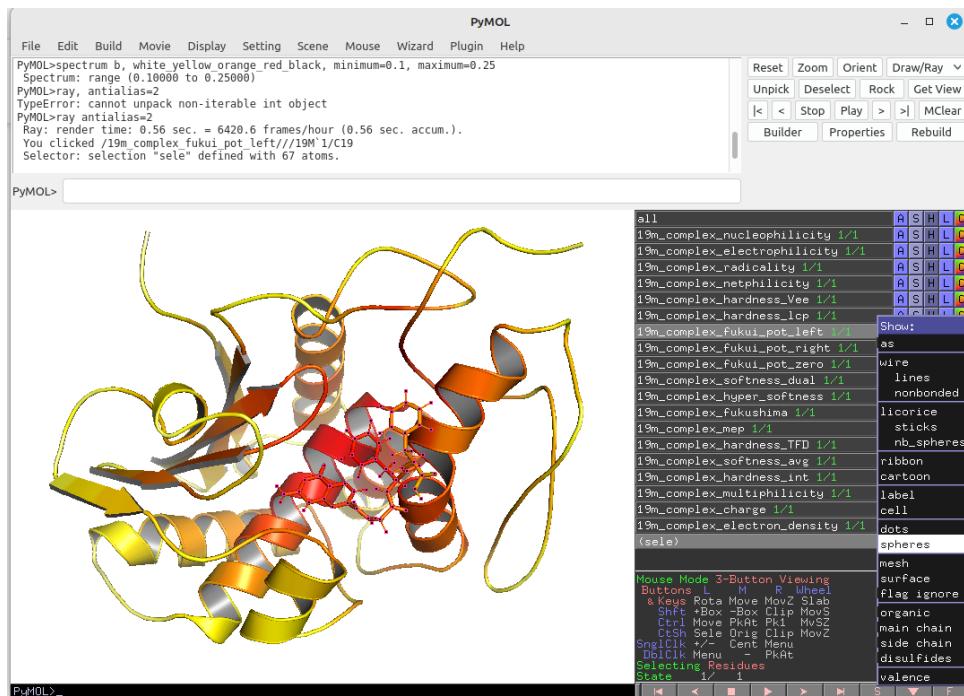


Figura 6: Janela do Pymol mostrando o ligante selecionado e o menu com opção para mostrar a representação de esferas para os átomos.

E depois, para renderizar a imagem em máxima qualidade, usamos o comando "ray", como mostrado abaixo

```
ray antialias=2
```

Depois do comando "ray", é possível estabelecer os valores em pixels da imagem e controlar isso também. A opção "antialias=2" estabelece a maior qualidade para a renderização de imagens. Depois de renderizar a imagem, a janela do Pymol fica como na ??, e pode ser salva com o comando abaixo

```
png complex_rta_19m_hardness.png
```

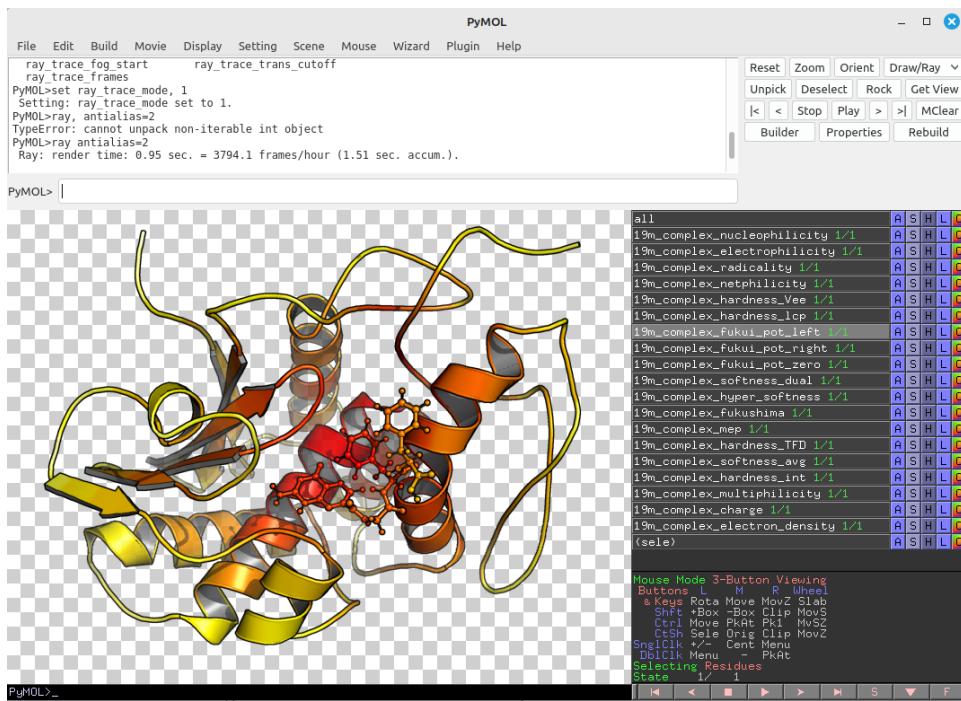


Figura 7: Janela do Pymol com a estrutura renderizada e pronta para ser salva.

O resultado é mostrado na Figura 8, em uma imagem que consegue mostras bem os átomos mais duros, baseado no potencial de Fukui, diferenciando bem os grupos farmacofóricos do ligante 19M.

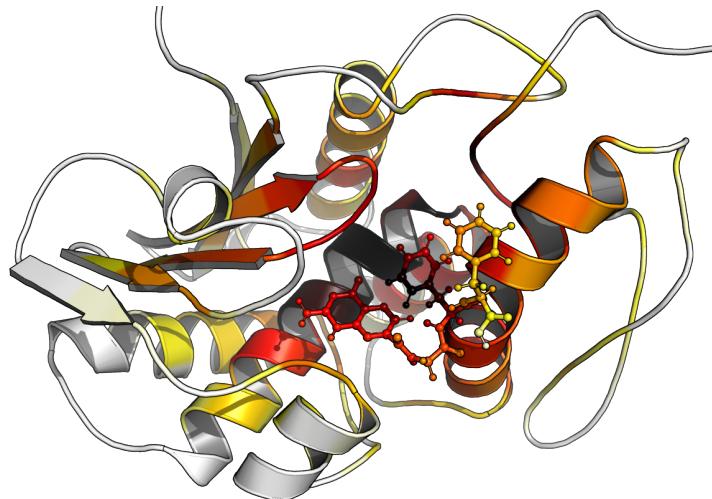


Figura 8: Imagem final renderizada do potencial de Fukui-Esquierdo para o complexo RTA-19M.

2 Tutorial 2: Gerando Trajetória de Catálise Enzimática

Um dos usos mais recorrentes do PRIMoRDiA, e portanto, da análise de estrutura eletrônica para macromoléculas, é a caracterização teórica de caminhos de reação enzimático. O que isso significa na prática é basicamente uma análise de reatividade e interações químicas contidas no sítio ativo de uma enzima com o substrato que se queira analisar, para cada etapa da reação catalisada. Isso permite obter uma visão teórica geral sobre os fenômenos químicos que ocorrem antes, durante e após esses processos tão essenciais para a manutenção dos organismos vivos, identificar os papéis de resíduos, íons, e corroborar informações de obtidas as simulações de caminhos de reação, como ordem das coordenadas de reação e sentido das transferências eletrônicas.

Apesar de tudo isso que pode ser obtido com os cálculos efetuados pelo PRIMoRDiA, é necessária a determinação prévia de uma trajetória de estruturas, que representem um caminho de reação química. Por isso, esse tutorial tem como objetivo mostrar a utilização de scripts em Python 3 desenvolvidas pelo nosso grupo para operar facilmente a biblioteca pDynamo, que permite realizar cálculos de de potencial Híbrido de Mecânica Quântica/Mecânica Molecular, que é o tipo de protocolo mais comum para simulação de reações químicas em grandes sistemas biológicos.

2.1 O Sistema: Triose-Fosfato Isomerase

Reações enzimáticas são reações químicas catalisadas por proteínas, que fazem com que transformações químicas ocorram em um tempo hábil para o funcionamento do metabolismo de organismos vivos. Portanto, as enzimas são componentes importantíssimos para a manutenção da vida, participando desde a replicação de DNA até a decomposição do açúcar.

O estudo do mecanismo dessas reações é muito importante para o desenvolvimento de drogas, entendimento de metabolismo, mecanismos de doenças e replicação de micro-organismos e aplicações industriais dessas enzimas. Em um dos estudos do nosso grupo, foi com o mecanismo da reação catalisada pela triose-fosfato-isomerase (TIM)^[5], mais especificamente a etapa mostrada no esquema abaixo na Figura 9.

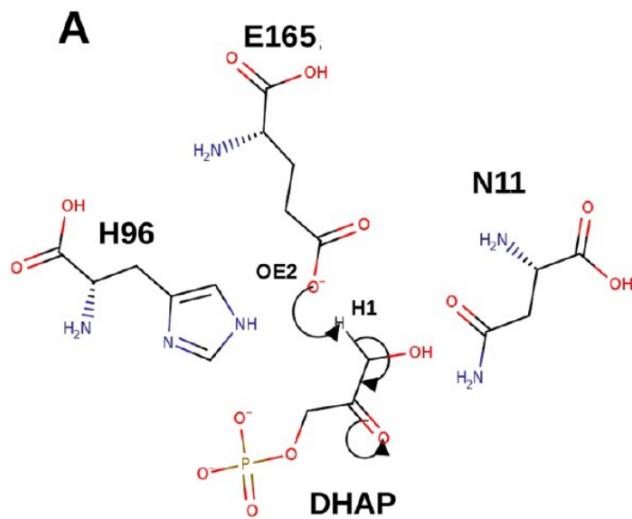


Figura 9: Representação do mecanismo de catálise enzimática estudado nesse tutorial.

Nós utilizamos esse enzima para testar a eficácia dos descritores em caracterizar teoricamente essa reação, descrevendo as principais forças, interações e efeitos do ambiente. A TIM é uma das enzimas mais caracterizadas na literatura, é uma enzima central no caminho glicolítico da interconversão do 3-fosfato D-gliceroldeído (GAP em inglês) para o fosfato dihidroxiacetona, com grande eficiência.

2.2 pDynamo e pDynamo_Scripts

O pDynamo_Scripts é uma coleção de scripts em Python 3 que encapsulam o funcionamento da biblioteca pDynamo e permitem a simulação desses sistemas complexos através de funções bem simples em Python, onde o usuário só precisa determinar o tipo de simulação e os parâmetros necessários. Esse script está ainda em desenvolvimento e tende a ter em breve uma versão que aceite um input escrito em formato de texto simples e executado como o PRIMoRDiA. Por enquanto, esses scripts vão permitir a nós a:

1. Configurar o sistema com os parâmetros do campo de força do AMBER, a partir de arquivos de topologia e coordenadas providos pelo tutorial
2. Otimizar a geometria da caixa de simulação com mecânica molecular, contendo mais de 26 mil átomos (Figura 10).
3. Recorte esférico do sistema, chamada de "pruning", onde a parte mais externa do sistema, em relação ao sítio ativo, é excluída para diminuir custos computacionais. A partir do átomo que vai ser transferido na reação, são incluídos os átomos de todos os resíduos que estejam no mínimo a 25 Å de distância (Figura 11).
4. Fixação dos átomos da camada mais externa, a partir de 20 Å

do centro reacional, deixando somente os átomos dentro desse raio com possibilidade de ter suas coordenadas atualizadas pelas simulações.

5. Otimização de geometria com mecânica molecular com esse sistema recortado e fixo.
6. Configuração do potencial híbrido QM/MM e otimização de geometria (Figura 12).
7. Simulação da varredura da coordenada de reação.
8. Refinamento da energia eletrônica no software MOPAC, com um número de átomos maior na região tratada com potencial de Mecânica Quântica.

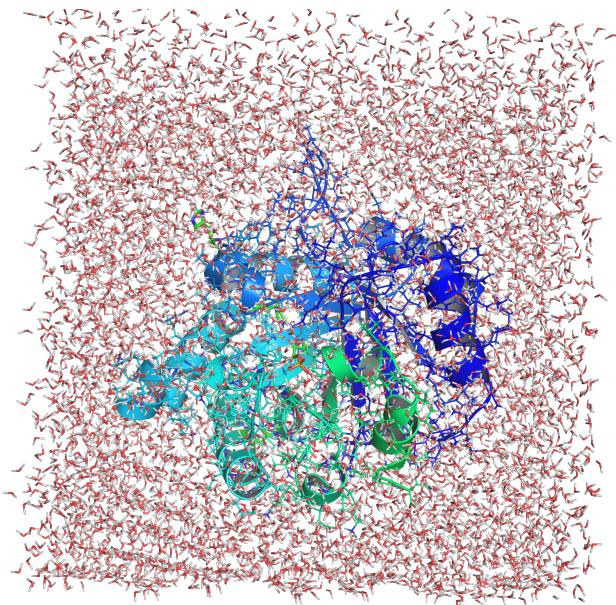


Figura 10: Caixa de simulação, contendo a proteína, substrato e moléculas de água.

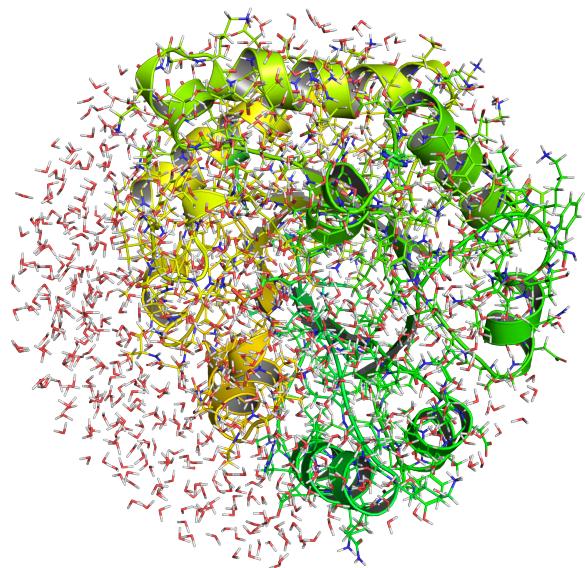


Figura 11: Sistema recortado esfericamente a partir do centro reativos, e otimizado com mecânica molecular.

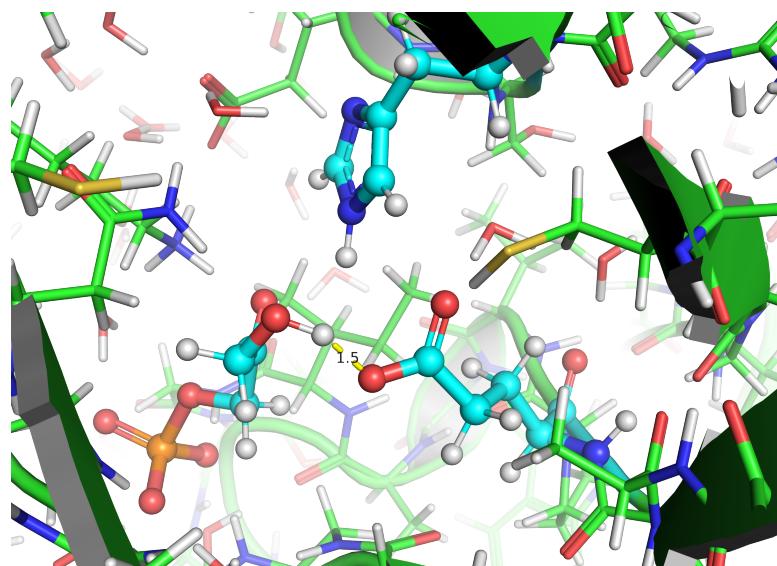


Figura 12: Sítio ativo otimizado com QM/MM, pronto para varredura de superfície relaxada.

Para esse tutorial será provido um script próprio, que é inspirado em alguns dos testes, com o nome de Reaction_TIM.py, que pode ser executado com o seguinte comando

```
computer@user:$ python3 Reaction_TIM.py
```

Com esse comando, todos as simulações já enumeradas serão executadas. No entanto, para esse tutorial, vamos fornecer os arquivos

com os resultados das simulações até a simulação de varredura de coordenada de reação, que iremos acompanhar durante o mini-curso, devido a duração das simulações de otimizações de geometria.

Mas antes, vamos analisar cada função dentro do script e entender os parâmetros necessários. (Atividade durante o mini-curso)

2.3 Análise dos Resultados das Simulações

Esse tutorial ensina a gerar uma trajetória para um caminho de reação e obter os dados de estrutura eletrônica para uma análise posterior no PRIMoRDia. No entanto, podemos verificar algumas métricas e gráficos que foram gerados pelo pDynamo_scripts, que são os mesmos dados que podem ser utilizados para estudar qualquer caminho de reação.

Durante essa aula do mini-curso, nós vamos avaliar:

1. O gráfico de energia potencial em função do deslocamento da coordenada de reação (Figura 13)
2. O gráfico de energia potencial para todos os métodos semi-empíricos no qual o sistema teve a energia refinada considerando uma região tratada por mecânica quântica maior (Figura 14) .
3. A renderização da animação da trajetória no Pymol.

Vemos que o pico de energia potencial se dá antes da transferência do hidrogênio. Essa reação, da forma como foi configurada para ser realizada, é um teste do pDynamo, e ainda não está em condições ótimas, pois a área quântica poderia ser modificada para a varredura, incluindo ou excluindo efeitos importantes. Um dos usos da análise com os descritores é justamente a confirmação dos mecanismos propostos nas análises de superfície de energia potencial.

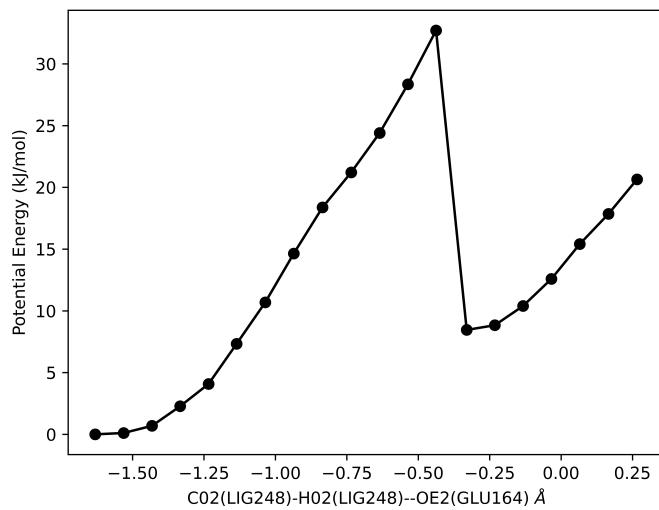


Figura 13: Variação da energia potencial durante a varredura da coordenada de reação, calculadas com potencial híbrido QM/MM com o semi-empírico RM1.

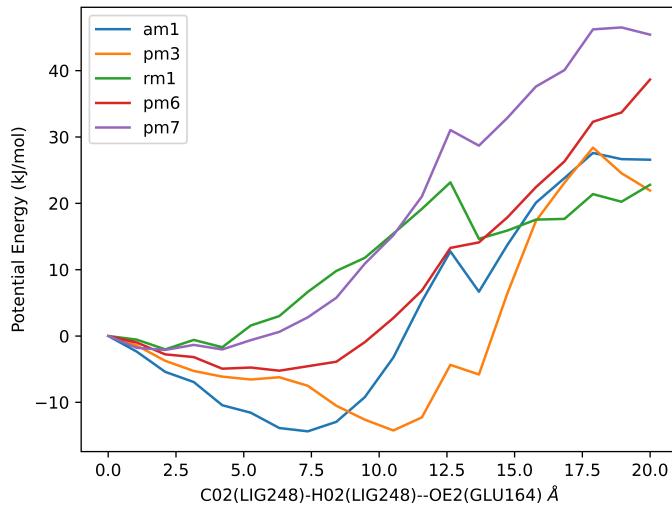


Figura 14: Refinamento de energia

3 Tutorial 3: Análise de Reatividade de Caminho de Reação Enzimático

Esse tutorial é idêntico ao tutorial 5 do repositório, com a diferença que a trajetória de reação enzimática analisada é que foi gerada no tutorial 2 do presente documento.

Os arquivos necessários para os cálculos desse tutorial estão na pasta. São arquivos de ".aux" e ".pdb", para as etapas do caminho de reação. Como foram vários métodos semi-empíricos rodados

no refinamento de energia no Mopac, foram gerados arquivos com nomes dos Hamiltonianos. Vamos renomear os arquivos de .aux que foram gerados com o RM1 e os respectivos arquivos de PBB para tirar o nome do Hamiltoniano para o PRIMoRDiA reconhecer a sequência de frames, usando os seguintes comandos

```
computer@user:$ for file in *rm1.aux;
computer@user:$   do
computer@user:$     mv "$file" "${file/_rm1.aux/.aux}";
computer@user:$ done
```

```
computer@user:$ for file in *rm1.pdb;
computer@user:$   do
computer@user:$     mv "$file" "${file/_rm1.pdb/.pdb}";
computer@user:$ done
```

Para fazer as análises especiais para a reação enzimática teremos que usar outro método de execução, que é sinalizado no PRIMoRDiA pela keyword "reaction" depois de "#RT" na primeira linha, ou seja, o Run Type passa para o modo de reação. Na pasta já tem o arquivo de input, que agora possui vários elementos importantes de citar, e está mostrado na figura abaixo

É possível perceber que tem varias outras linhas com keywords antes não usadas e somente uma entrada de arquivo ".aux" para rodar com a opção 3 de cálculo. Como esse é um cálculo para reação química, o PRIMoRDiA entende aquela entrada como valendo para todos os auxs na pasta que tenham o sufixo sinalizado, nesse caso no nosso input foi o sufixo "frame", rodando do "frame0.aux" até o "frame19.aux", obedecendo o parâmetro "dimX".

Vamos analisar os outros parâmetros do input

Listing 2: Input editado para execução do tutorial 3

```
#RT reaction
#PR eband 3 Rscript pymols
#Reaction RC1 137 138 73
#Reaction dimX 20
#TRJ residues 3 4 9
3 frame.aux true 30 10 frame.pdb mopac 0 0 0 0 EW
```

Segunda linha

- Segunda linha
 - #PR: flag para sinalizar que serão dados parâmetros para o PRIMoRDiA
 - eband: keyword sinalizando que o próximo parâmetro é um valor de energia de corte para considerar os orbitais moleculares

res

- Rscript: keyword importante par gerar os scripts de análise dos descritores durante a trajetória de reação
- pymols: keyword para o output de scripts para o Pymol
- Terceira linha
 - #Reaction: flag para sinalizar que vai ser passado parâmetros para análise de coordenada de reação
 - RC1: keyword sinalizando que vais ser passado o número dos átomos que compõem a coordenada de reação
 - 9: número do átomo de carbono C1 (C02 no PDB) do GAP
 - 107:número do átomo de hidrogênio H1 (H02 no PDB) ligado ao DHAP
 - 51: número do átomo de oxigênio OE2 do GLU
- Quarta Linha
 - #Reaction: flag para sinalizar que vai ser passado parâmetros para análise de coordenada de reação
 - dimX : keyword para sinalizar que o próximo valor na linha da o número de etapas da trajetória
 - 20 : número de etapas da trajetória
- Quinta Linha
 - #TRJ: flag para sinalizar que vai ser passado parâmetros tratar com o conjunto de frames que compõem a trajetória de reação
 - residues : keyword indicando que os próximos valores inteiros passados nessa linha são o número dos resíduos no PBD que se deseja analisar para a dada trajetória.
 - 3: número no PDB da histidina catalítica
 - 4: número no PDB do Glutamato catalítico
 - 9: número no PDB do resíduo representando o substrato.
- Sexta Linha
 - 3 : opção de cálculo do PRIMoRDiA para macromoléculas
 - frame.aux: nome base para os arquivos, antes do ponto o prefixo sem o número que indica o frame, e depois do ponto a extensão do arquivo
 - true : keyword sinalizando que nenhum métodos de dureza local foi pedido para cálculo da representação volumétrica

- 30 : finura do grid para representação volumétrica, sempre deixar em zero até uma nova atualização do software
- 10 : número máximo de orbitais moleculares de fronteira para ser utilizado se o método de combinação de orbitais for BD
- frame.pdb: nome base para os arquivos de pdb, mesma lógica para os arquivos que contém a estrutura eletrônica
- mopac : nome do programa de onde foi calculada a estrutura eletrônica
- 0 : parâmetro não relevante com grid 0
- 0 : parâmetro não relevante com grid 0
- 0 : parâmetro não relevante com grid 0
- 0 : parâmetro não relevante com grid 0
- EW : keyword para o método de combinação de orbitais moleculares

3.1 Análise dos Resultados

Para analisar todos os descritores e outras propriedades de energia, o PRIMoRDiA cria um script para executar no R que por sua vez produz gráficos dessas quantidades variando em função das frames. Depois da execução deve aparecer arquivos com sufixo ".R", um deles com nome "frame0_reaction_analysis.R" e outro "frame0_residuos_analysis.R". Vamos executar o primeiro com o Rscript, como no comando abaixo

```
computer@user:$ Rscript frame0_reaction_analysis.R
```

Depois desse comando, o R deve produzir diversos arquivos de imagem com propriedades globais e locais para cada átomo da coordenada de reação. No primeiro gráfico da Figura 15 HOF é o calor de formação (heat of formation) que o Mopac calcula, e é uma das melhores formas de variação de energia em uma reação química simulada com método semi-empírico. Lembrado que todas essas propriedades são em referência ao primeiro valor, ou seja, cada ponto uma variação ao estado inicial. Nessas simulações uma estrutura que normalmente se deseja determinar é a estrutura de transição, que é determinada pelo ponto máximo de variação de energia, nesse caso na penúltima etapa. No entanto, nesse nosso cálculo com PM6, a área quântica foi bem menor do que a utilizada no na nossa publicação e o HOF não foi um bom indicador do estado de transição. Já a energia eletrônica mostra um primeiro máximo no ponto zero da coordenada de reação, que é o momento que o próton está a mesma distância do carbono e do oxigênio. Em geral, os outros descritores globais indicam um aumento da moleza química (softness) e consequente diminuição da dureza (hardness),

aumento do potencial químico e da eletrofilicidade total indicam que o sistema tende a receber transferência de elétrons.

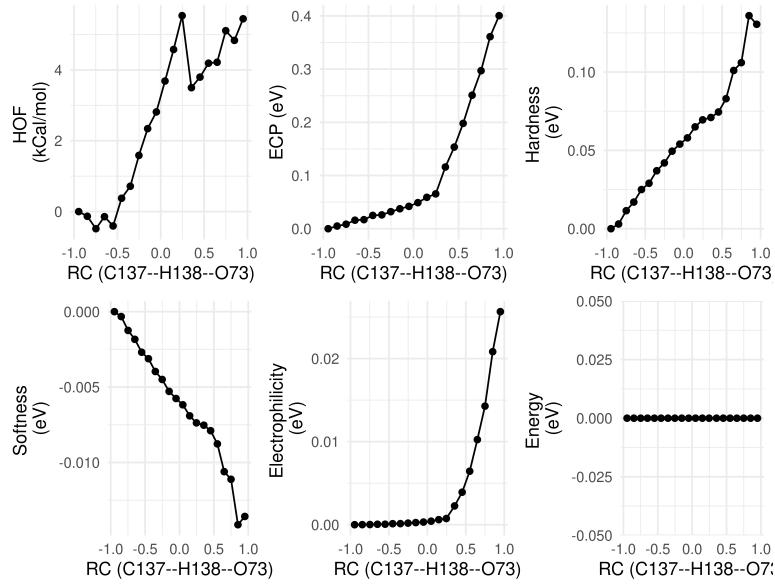


Figura 15: Descritores globais calculados para cada etapa da trajetória de reação.

Na Figura 16, nós mostramos os descritores para o átomo de oxigênio do resíduo GLU. No script gerado pelo PRIMORDiA nós analisamos 9 propriedade por átomo da coordenada de reação. Para o oxigênio é interessante ressaltar algumas, como o decréscimo na densidade eletrônica, atingindo um mínimo no estado de transição e depois retomando o valor nos produtos. A carga parcial aumenta também, indicando que há uma transferência de elétrons do oxigênio para o carbono, e que depois o sistema em volta faz uma "devolução" de carga para o oxigênio, mas que no final ainda fica com menos, já que agora o grupo carboxílico está protonado e portanto neutro. Já para a dureza local esse átomo apresenta comportamentos diferentes para cada tipo de método de cálculo, para interpretar o que cada um significa é importante entender cada modelo. Brevemente, a dureza baseada no LCP tem mais correção com o descritor global de dureza, indicando que esse tinha a propensão de interagir por processo duro-duro e depois foi diminuindo esse comportamento conforme foi transferindo carga.

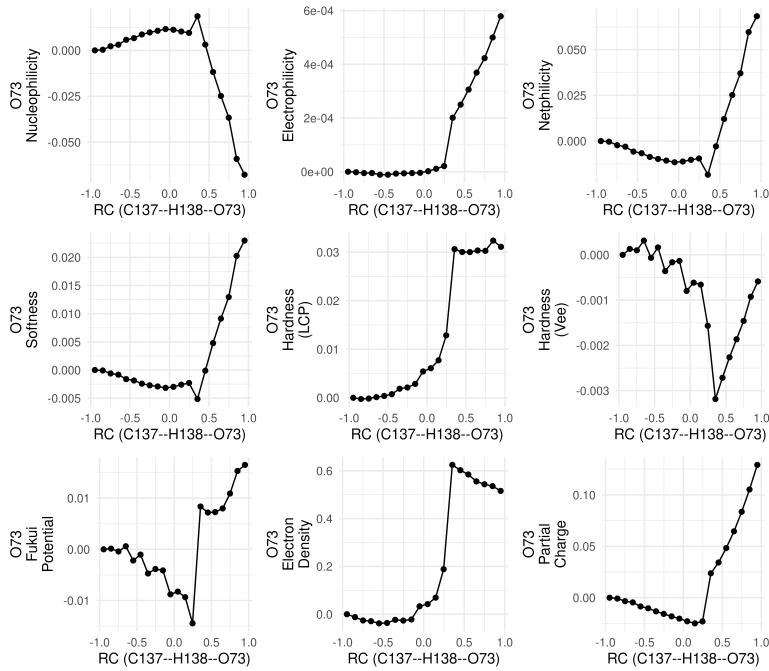


Figura 16: Descritores locais calculados para o átomo O2E do grupo carboxílico do Glutamato catalítico para a trajetória de reação.

Os descritores para o carbono C1 estão na Figura 17. Tem bastante coisa interessante para analisar nesse caso. Primeiro, como que para dois métodos de dureza o valor máximo condiz com o estado de transição e se assemelha bastante com o perfil de energia da reação. Já que a interação dominante para transferência de prótons é a duro-duro, o carbono assume o valor máximo para esse descritor, em dois métodos Vee e LCP. O método de dureza baseado no potencial de Fukui da a força exercida nos núcleos em respeito da variação do número de elétrons, o que tem uma interpretação mais diferente da geral sobre dureza comum.

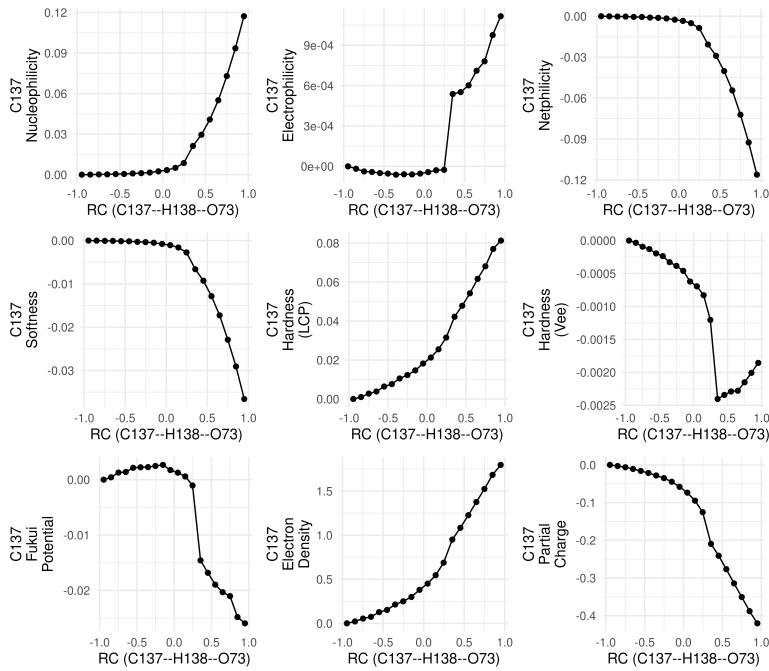


Figura 17: Descritores locais calculados para o átomo C1 do substrato para a trajetória de reação.

Os descritores para o átomo transferido, o hidrogênio do DHAP (H1), se encontra na figura de baixo. Os descritores baseados nos orbitais moleculares, como eletrofilicidade, nucleofilicidade e netfilicidade mostraram variação mas com valores muito baixos. Agora é interessante como a densidade eletrônica vai diminuindo, se tornando cada vez mais um próton, um hidrogênio carregado positivamente. A dureza local calculado pela aproximação do potencial de Fukui, foi a que apresentou maior coerência com o processo, com um aumento das interações eletrostáticas próximo ao estado de transição, diminui e atinge o máximo nos produtos, onde a carga parcial alcança o maior valor, sendo que esse átomo tende a ser transferido pelo mesmo tipo de interação.

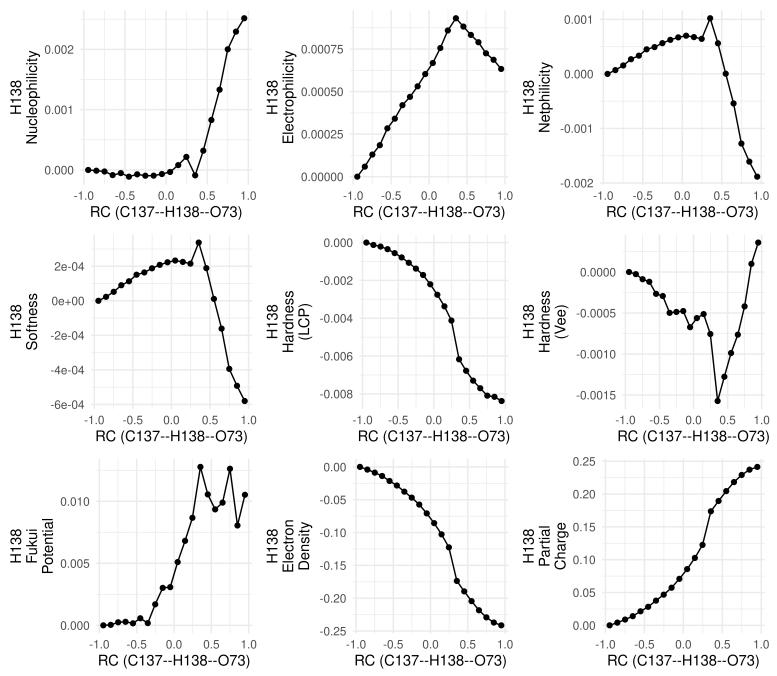


Figura 18: Descritores locais calculados para o átomo H1 do substrato para a trajetória de reação.

Como no Tutorial 3, vamos gerar imagens dos descritores para alguns pontos da trajetória. O PRiMoRDiA escreve para cada arquivo um script de Pymol para automatizar a visualização dos descritores. Abra o Pymol e execute o arquivo com o prefixo "frame0" e sufixo ".pym". Como mostrado na Figura 18 e no caixa de comando abaixo.

```
@frame0_pymols.pym
```

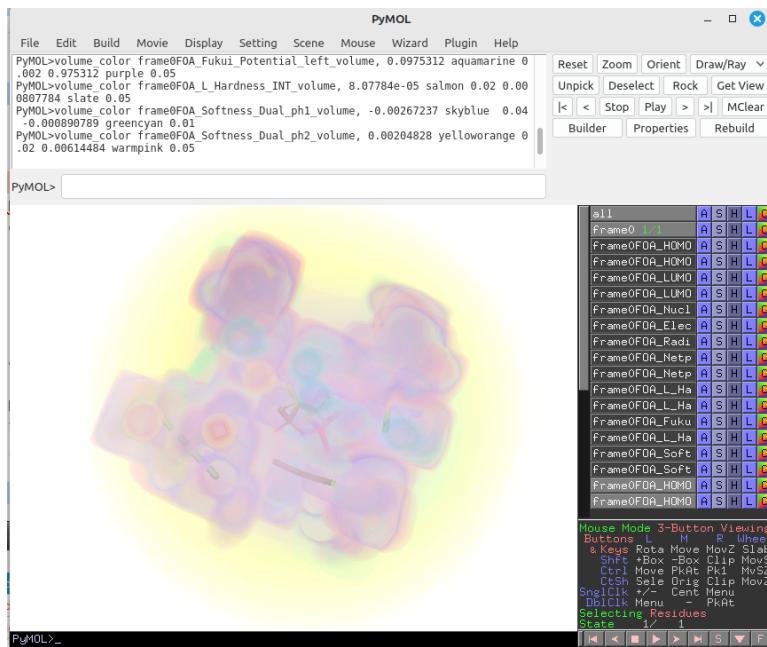


Figura 19: Janela do Pymol com os resultados dos descritores carregados e o objeto do descritor de .

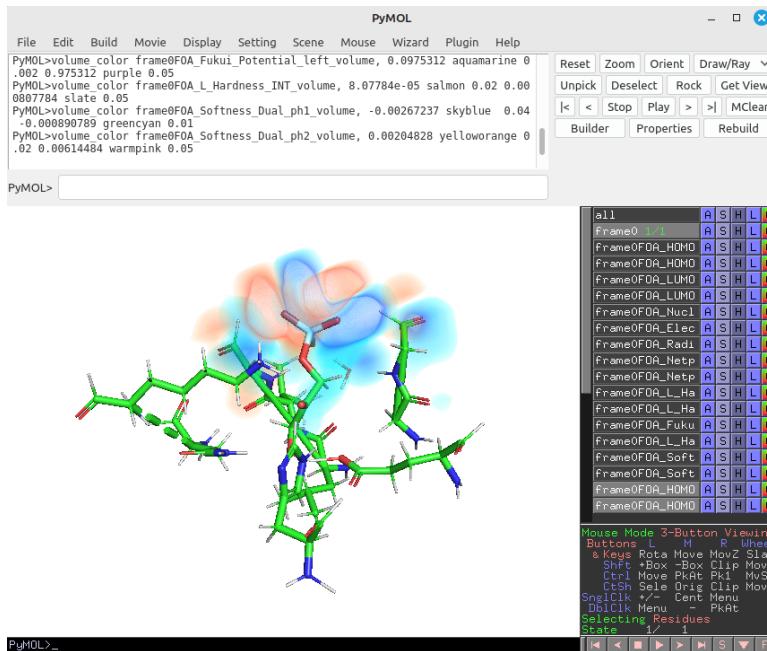


Figura 20: Janela do Pymol com a representação volumétrica do orbital HOMO.

Esses descritores em representação volumétrica são esteticamente apelativos, mas apresentam um custo computacional alto e podem apresentar uma visualização mais confusa quando o restante da estrutura da proteína é apresentada. Para gerar representações de descritores centrada nos átomos pode ser utilizado o outro script

gerado também pelo PRIMoRDiA, com o seguinte comando no software pymol

```
@frame0_pymols_pdb.pym
```

As próximas figuras foram resultado do processo de renderização no pymol para o descriptor de dureza local com o segundo método disponível no PRIMoRDiA. O objeto com a proteína inteira foi colorido com os carbonos em verde e escondido todas as outras representação que não fossem cartoon para não atrapalhar a visualização do objeto com o descriptor. Para a palheta escolhida o comando utilizado foi

```
spectrum b, white_yellow_orange_red_black, minimum=0.1,maximum=0.2
```

Esses valores podem sofrer pequenos ajustar para melhorar o destaque dos átomo com o maior valor de dureza. Depois, com o comando "ray" seguido do "png" as imagens foram salvas. A Figura 21 mostra a dureza local para a primeira etapa da coordenada. Durante o tutorial, abriremos outras estruturas e faremos as renderizações dos outros descritores.

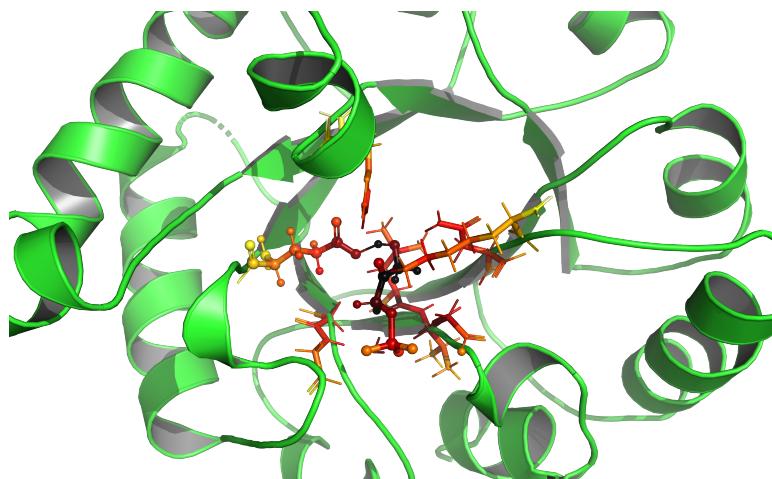


Figura 21: Dureza local para os átomos da enzima tratados com mecânica quântica para a estrutura corresponde ao primeiro passo da trajetória de reação.

O PRIMoRDiA também escreve scripts especiais do R para a análise dos descritores para resíduos específicos, indicados no input, produzindo gráficos de barra com a média e o desvio padrão, como na figura abaixo para o descriptor de eletrofilicidade local na Figura 22, nucleofilicidade local na Figura 23.

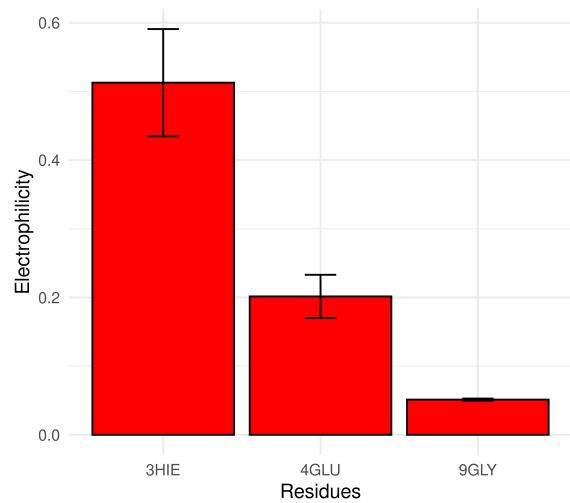


Figura 22: Gráfico de barras para a eletrofilicidade nos resíduos indicados durante a trajetória.

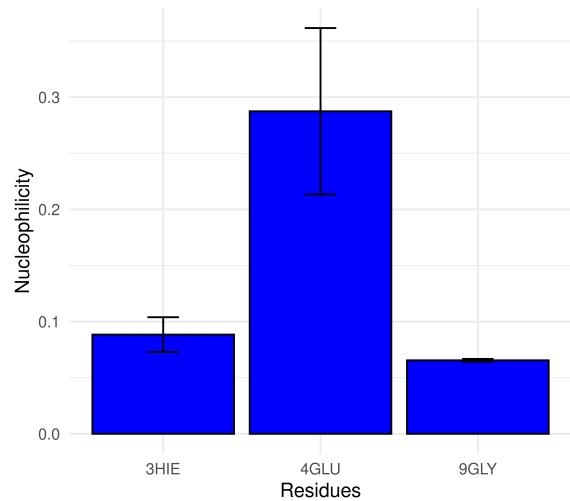


Figura 23: Gráfico de barras para a Nucleofilicidade nos resíduos indicados durante a trajetória.

Também, é acompanhado a variação desses valores durante a trajetória com o cálculo de médias móveis, como na Figura 24 é mostrado para densidade eletrônica em cada resíduo indicado no input, para a dureza local na Figura 25 e moleza local??.

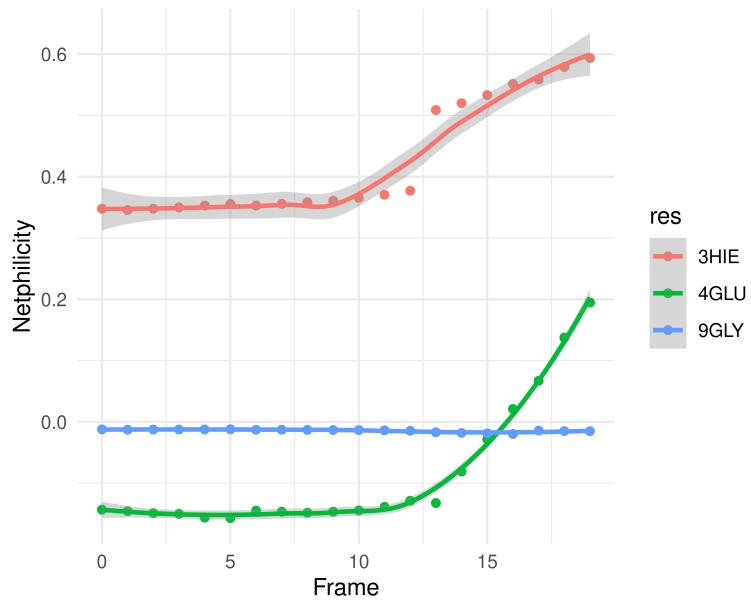


Figura 24: Médias móveis para a densidade eletrônica para os resíduos indicados no input para análise da trajetória.

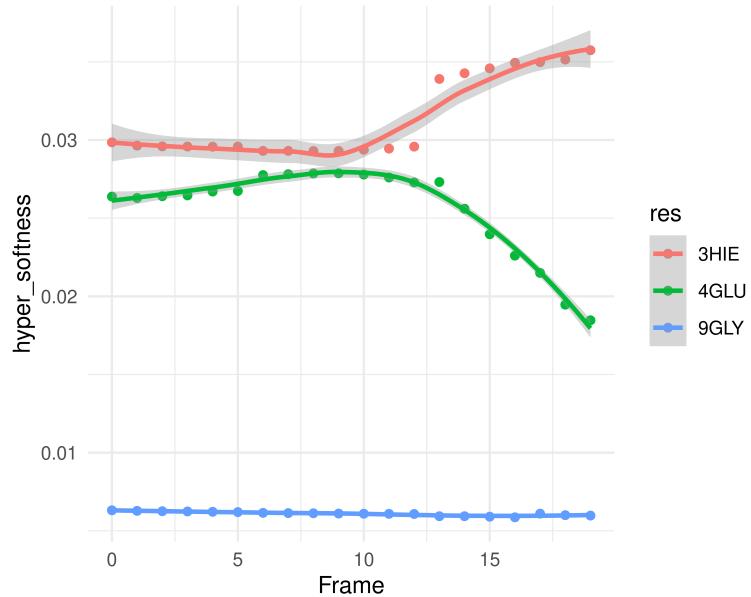


Figura 25: Médias móveis para o descritor de Hiper moleza local para os resíduos indicados no input para análise da trajetória.

Há diversas análises que podemos fazer baseados nos descritores de reatividade, que por sua vez também são numerosos, para cada átomo envolvido e para cada resíduo. O PRIMORDiA fornece uma forma automática de dispor todos esses resultados da forma mais pronta para análise possível.

Agradecimentos

We gratefully thanks the Brazilian science/education institutes by the physical structure and financial support: INCT-INAMI, CNPq,CAPES,FAPESQ-PB, PRONEX-FACEPE and FINEP. Universidade Federal da Paraíba (UFPB) and the computer resources of Centro Nacional de Processamento de Alto Desempenho em São Paulo (CENAPAD-SP). Research project Bioinformática Estrutural de Proteínas: Modelos, Algoritmos e Aplicações Bioteclógicas (Edital Biologia Computacional 51/2013, processo AUXPE1375/2014 da CAPES). G.B.R. acknowledges support from the Brazilian National Council for Scientific and Technological Development (CNPq grant no. 309761/2017-4).

Referências

- [1] I. B. Grillo, G. A. Urquiza-Carvalho, and G. B. Rocha, “PRI-MoRDiA: A Software to Explore Reactivity and Electronic Structure in Large Biomolecules,” *Journal of Chemical Information and Modeling*, vol. 60, pp. 5885–5890, Dec. 2020.
- [2] I. B. Grillo, G. A. Urquiza-Carvalho, and G. B. Rocha, “Quantum chemical descriptors based on semiempirical methods for large biomolecules,” *The Journal of Chemical Physics*, vol. 158, p. 201001, May 2023.
- [3] I. B. Grillo, G. A. Urquiza-Carvalho, E. J. F. Chaves, and G. B. Rocha, “Semiempirical methods do Fukui functions: Unlocking a modeling framework for biosystems,” *Journal of Computational Chemistry*, vol. 41, pp. 862–873, Apr. 2020.
- [4] A. Rocha-Santos, E. J. F. Chaves, I. B. Grillo, A. S. De Freitas, D. A. M. Araújo, and G. B. Rocha, “Thermochemical and Quantum Descriptor Calculations for Gaining Insight into Ricin Toxin A (RTA) Inhibitors,” *ACS Omega*, vol. 6, pp. 8764–8777, Apr. 2021.
- [5] I. B. Grillo, G. A. Urquiza-Carvalho, J. F. R. Bachega, and G. B. Rocha, “Elucidating Enzymatic Catalysis Using Fast Quantum Chemical Descriptors,” *Journal of Chemical Information and Modeling*, vol. 60, pp. 578–591, Feb. 2020.
- [6] I. B. Grillo, J. F. R. Bachega, L. F. S. M. Timmers, R. A. Caceres, O. N. De Souza, M. J. Field, and G. B. Rocha, “Theoretical characterization of the shikimate 5-dehydrogenase reaction from Mycobacterium tuberculosis by hybrid QC/MM simulations and quantum chemical descriptors,” *Journal of Molecular Modeling*, vol. 26, p. 297, Nov. 2020.