

XI Escola de Modelagem Molecular em Sistemas Biológicos
Laboratório Nacional de Computação Científica - LNCC/MCTI
Petrópolis - Rio de Janeiro - Brasil
19 a 23 de agosto de 2024

Atracamento molecular receptor-ligante e triagem virtual em larga escala


ANA LUIZA M. KARL - almkarl@lncc.br
CAMILA S. DE MAGALHÃES - camila@xerem.ufrj.br
ISABELLA A. GUEDES - isabella@lncc.br
LAURENT E. DARDENNE - dardenne@lncc.br
LETÍCIA C. DE ASSIS - lcassis@posgrad.lncc.br

Os arquivos de entrada para a prática estão disponíveis em
<https://github.com/gmmsb-lncc/emmsb-recursos/tree/main/Minicursos>

Dia 1 - *Re-docking* do complexo 1DB1

Como primeira experiência utilizando o portal DockThor-VS, faremos uma simulação de *re-docking* da vitamina D com seu receptor (PDB ID 1DB1).

1. Download dos arquivos da proteína e do ligante

- 1.1 Busque o arquivo com as coordenadas tridimensionais (.pdb) do complexo no *Protein Data Bank*: www.rcsb.org. No campo de busca em branco, digite o código do complexo 1DB1 e clique em 
- 1.2 As informações sobre o complexo serão carregadas na página. Analise a presença de ligantes e cofatores, número de cadeias e outras informações relevantes (organismo de origem, resolução do cristal, etc.).
- 1.3 Faça o download do arquivo do complexo (**Download Files** → **PDB Format**) e salve como *1db1_complex.pdb* em sua pasta pessoal¹.
- 1.4 Clique no código do ligante **VDX** para carregar a página contendo informações específicas sobre esta molécula. É possível fazer o download da estrutura 3D do ligante em sua conformação experimental presente no cristal (o que chamamos de ligante de referência) ou a ideal, *i.e.* conformação mais estável do ligante no vácuo (Download SDF File). A conformação ideal é calculada pelos programas Corina e OpenEye Omega. Neste tutorial será feito o *re-docking*, então será usada a conformação **experimental** do ligante.
- 1.5 Abra o arquivo do complexo 1DB1 e analise a estrutura tridimensional no visualizador Pymol.

Inicialmente o complexo é carregado na visualização de linhas e os átomos de oxigênio das moléculas de água são representados por **+**.
- 1.6 Visualize as estruturas secundárias da proteína (**S** → **Cartoon**).

No Pymol, é possível visualizar rapidamente as interações existentes entre ligantes e proteína. Carregue essa visualização e analise algumas interações (**A** → **Preset** → **Ligand Sites** → **Cartoon**). Com isso, apenas os resíduos da proteína que interagem com o ligante são representados como linhas. Os demais resíduos serão representados por *Cartoon*. Automaticamente são ilustradas as ligações de

¹ O arquivo já foi baixado (*1db1_complex.pdb*) no diretório *curso-docking/1db1/* disponível no website da XIEMMSB no arquivo *curso_docking.zip*.

hidrogênio formadas entre a proteína, o ligante e moléculas de água no sítio de ligação através de linhas pontilhadas amarelas.

Meça a distância de algumas ligações de hidrogênio (*Wizard -> Measurement*).

1.7 Para calibrar os parâmetros do docking, é essencial realizar um estudo de re-docking, onde o ligante é reposicionado na conformação que ele assume quando complexado com a proteína. Esse experimento é fundamental para validar o procedimento de docking, pois permite verificar se o programa é capaz de reproduzir o modo de ligação experimental utilizando os estados de ionização definidos na preparação da proteína. Para isso, o ligante deve ser extraído diretamente do arquivo PDB do complexo original.

1.7.1 No Pymol, selecione o ligante VDX pela linha de comando. Observe que os átomos do ligante serão selecionados e será criada automaticamente a entrada *1db1_ligand*:

```
select 1db1_ligand, resn VDX
```

1.7.2 Salve o arquivo do ligante no formato .pdb: *File -> Save Molecule* e escolha a entrada *1db1_ligand*.

2. Preparação da proteína

2.1 Os arquivos provenientes do PDB possuem diversas imprecisões. Além de não possuírem átomos de hidrogênio, podem ter átomos pesados faltando. Nesta etapa, vamos preparar os arquivos da proteína usando o programa **PdbThorBox** no Portal DockThor (*Docking -> Protein*).

2.2 Acesse o portal através do endereço www.dockthor.incc.br

2.3 Selecione a aba *Protein*

2.4 Carregue o arquivo do complexo 1DB1 (1db1_complex.pdb) e clique em *Send*.

2.5 Nessa etapa, caso o arquivo PDB não contenha hidrogênios, eles são adicionados conforme o estado de protonação padrão dos resíduos de aminoácidos em pH neutro: Asp e Glu carregados negativamente, Lys e Arg carregados positivamente, e His neutra (protonada em NE2).

Atenção: o programa **pdbthorbox** remove automaticamente os ligantes, cofatores e água, além de completar as cadeias laterais de resíduos de aminoácidos que possuem átomos faltando.

2.6 Visualize a proteína pelo portal (*View 3D*).

2.7 Clique em *Send to DockThor* para armazenar a estrutura da proteína preparada e prosseguir para a etapa seguinte.

2.8 Clique em *Do Not Use Cofactors*, pois para este complexo não será necessário considerar cofatores.

3. Preparação do ligante

3.1 Carregue o arquivo do ligante (1db1_ligand.pdb) e clique em *Send*.

3.2 Visualize a estrutura 3D (*View 3D*).

3.3 Habilite a opção de adicionar os átomos de hidrogênios e prepare a estrutura do ligante (marque a opção *Add hydrogens* e clique em *Send* novamente). A opção para adicionar automaticamente os átomos de hidrogênio mantém os hidrogênios já existentes, e adiciona novos caso seja necessário completar as valências dos átomos. Esta opção deve ser usada com atenção caso a estrutura do ligante esteja com algum estado de protonação pré-definido.

3.4 Visualize a estrutura do ligante preparada (*View 3D*).

3.5 Analise as ligações rotacionáveis que serão consideradas como flexíveis durante o *docking* (*Rotatable bond editor*).

3.6 Visualize na estrutura tridimensional do ligante quais são essas ligações (*View 3D*). Neste experimento todas as ligações rotacionáveis serão consideradas flexíveis.

4. Redocking

- 4.1 Já com os arquivos da proteína e do ligante preparados, podemos executar o *redocking* usando o programa **DockThor**. Na aba *DockThor* todos os principais parâmetros necessários para o *docking* estão disponíveis para serem modificados pelo usuário.
 - 4.2 O centro de coordenadas da grade para o *docking* deve ser pré-definido pelo usuário na opção *Define the binding site*, assim como a dimensão da grade. Para este experimento utilizaremos o centro da grade **X** = 10.15; **Y** = 22.37; **Z** = 34.04, com 20 Å em cada dimensão (X, Y, Z) e espaçamento (discretização) de 0.25 Å, totalizando uma grade de 531.441 pontos.
 - 4.3 Visualize a grade de energia sobreposta com a proteína na janela do visualizador NGLViewer.
 - 4.4 Visualize os parâmetros relativos ao algoritmo genético (*Select the search algorithm precision*). Utilizaremos os valores padrão do algoritmo genético (modo **Standard**).
 - 4.5 Defina um nome (*label*) para o seu projeto e adicione seu e-mail.
 - 4.6 Selecione a opção *Subscribe DockThor e-Newsletters* caso queira receber novidades do Portal DockThor.
- Obs.: É possível adicionar até cinco (05) e-mails diferentes. Você receberá um e-mail de submissão e outro quando o experimento finalizar com o link para acessar os resultados.
- 4.7 Aceite os termos de uso e clique em **Dock!** para submeter seu experimento.

5. Análise dos Resultados

- 5.1 No momento em que terminar a execução do *docking* o usuário recebe um e-mail com o endereço *web* para analisar e fazer o download dos resultados. Clique em *View Results* quando receber o e-mail.
- 5.2 Entre no *link* abaixo para analisar os resultados²:

² **Atenção:** os resultados de *docking* ficam disponíveis para acesso no portal e *download* por 30 dias a partir da data de submissão. Todos os arquivos podem ser excluídos após este período.

https://dockthor.Incc.br/v2/index.php?tab=DOCKING&page=RESULTS&jobId=emmsb2024_1db1_66b37d86af96e

- 5.3 Observe os parâmetros da análise que podem ser modificados. Mantenha os valores padrão.
- 5.4 Carregue o arquivo original do ligante (1db1_ligand.pdb) para comparar os modos de ligação encontrados pelo *docking* e a conformação observada experimentalmente.
- 5.5 Clique em *Analyse* para executar o programa **DTStatistic** e analisar automaticamente os resultados.
- 5.6 Analise os resultados de *docking* clicando em *View results interactively*.
- 5.7 Observe o valor da afinidade predita (em kcal/mol) para o melhor modo de ligação.

Dia 2 – Docking utilizando cofatores e Exploração dos estados de Protonação da Proteína

Parte 1 - Docking utilizando cofatores - 2WX2

Esta prática abordará o processo de docking com a inclusão de cofatores usando o programa DockThor, proporcionando uma compreensão prática de como esses componentes essenciais podem influenciar as interações entre ligantes e proteínas.

1. Preparação da proteína

1.1 Abra o arquivo *2wx2_complex.pdb* com o Pymol para visualizar o ligante, a proteína e os cofatores.

1.2 Alguns programas de atracamento molecular, como o DockThor, precisam que os cofatores e moléculas de água estejam em arquivos separados. Faremos essa separação utilizando a linha de comando do programa Pymol, que nos permite fazer muitas seleções úteis. Salve em diferentes arquivos as seguintes moléculas relativas à **cadeia A**:

(a) ligante (*fluconazole* – **TPF**) – salve como *2wxw_ligand.pdb*

```
select 2wx2_ligand, resn TPF & chain A
```

(b) grupo heme sem o átomo de ferro (**HEM**) – salve como *2wxw_cofactor1.pdb*

```
select 2wx2_cofactor1, resn HEM & not name FE & chain A
```

(c) ferro (**FE**) – salve como *2wxw_cofactor2.pdb*

```
select 2wx2_cofactor2, name FE & chain A
```

(d) água estrutural 2214 (**HOH**) – salve como *2wxw_water.pdb*

```
select 2wx2_water, resn HOH & resi 2214
```

1.3 Salve os quatro arquivos (*i.e.*, ligante, heme, ferro e água) no formato .pdb selecionando a opção *Multiple Files (File -> Save Molecule)*.

1.4 Prepare o arquivo da proteína no Portal DockThor. Não será necessário alterar nenhum estado de protonação.

1.5 Faça download dos arquivos preparados (*Download*).

1.6 Analise os arquivos gerados. Observe que o nome dos arquivos de saída da proteína foram alterados para um identificador padrão no formato “**protein_ID**”. ID corresponde a um número identificador gerado aleatoriamente.

Arquivos	Descrição
*.in	Coordenadas atômicas, cargas parciais e tipo dos átomos da proteína.
*_prep.pdb	Proteína preparada no formato PDB (sem HETATM).

1.6.1 Abra o arquivo da proteína preparada no Pymol e analise as alterações realizadas.

```
pymol protein_ID_prep.pdb &
```

2. Preparação dos cofatores

2.1 Carregue os três arquivos de cofatores no portal e prepare-os adicionando os átomos de hidrogênio:

(a) 2wx2_cofactor1.pdb

(b) 2wx2_cofactor2.pdb

(c) 2wx2_water.pdb

2.2 Faça o *download* dos arquivos resultantes da preparação dos cofatores (*Download*).

2.3 Analise os arquivos gerados (diretório OUTPUT/).

Arquivos	Descrição
cofactor_log.txt	Arquivo contendo o log da preparação do(s) cofator(s).
cofactor_mapfile.csv	Tabela que relaciona o identificador aleatório de cada cofator com seu nome original.
*.format	Arquivo original do cofator em determinado formato (e.g. pdb).

*.top	Possui as coordenadas internas, cargas parciais, tipos de átomos e outras informações do cofator. Será usado para o <i>docking</i> .
new_*.sdf	Arquivo do mesmo formato do original contendo o cofator preparado. Este arquivo só é gerado quando se usa a opção para adicionar hidrogênios.

3. Preparação do ligante

- 3.1.1 Carregue o arquivo do ligante (*2wxw_ligand.pdb*).
- 3.1.2 Visualize a estrutura antes de realizar a preparação (*View 3D*).
- 3.1.3 Habilite a opção de adicionar os átomos de hidrogênios e prepare a estrutura do ligante.
- 3.1.4 Visualize a estrutura do ligante preparada (*View 3D*).
- 3.1.5 Faça o *download* dos arquivos resultantes da preparação do ligante (*Download*).
- 3.1.6 Analise os arquivos gerados (diretório OUTPUT/).

Arquivos	Descrição
ligand_log.txt	Arquivo contendo o log da preparação do(s) ligante(s).
ligand_mapfile.csv	Tabela que relaciona o identificador aleatório de cada ligante com seu nome original.
*.sdf	Arquivo original do ligante no formato SDF.
*.top	Possui as coordenadas internas, cargas parciais, tipos de átomos e outras informações do ligante. Será usado para o <i>docking</i> .
new_*.sdf	Arquivo do ligante preparado no formato SDF. Este arquivo só é gerado quando se usa a opção para adicionar hidrogênios.

- 3.1.7 Prossiga para a etapa de configuração do *docking* clicando em *Send to docking*.

4. Re-docking

4.1 Já com os arquivos da proteína, do ligante e dos cofatores preparados, podemos executar o *re-docking* usando o programa **DockThor**.

4.2 Verifique os arquivos que estão sendo considerados (Item 1. *Check your docking input files*).

4.3 Uma das formas de definir o centro do sítio de ligação é utilizar as coordenadas atômicas de um átomo de referência. Em experimentos de *re-docking*, costumamos selecionar o átomo mais central do ligante de referência. Abaixo, mostramos como obter os valores das coordenadas X, Y, Z de um átomo de referência com o programa Pymol.

4.3.1 Abra o arquivo do complexo no Pymol.

4.3.2 Dê zoom no ligante de referência:

```
zoom resn TPF
```

4.3.3 Escolha a opção de selecionar átomos e clique no átomo C1 do ligante.

4.3.4 Utilize os comandos abaixo para obter as coordenadas do átomo selecionado:

```
1. from pymol import stored
2. stored.pos = []
3. iterate_state 1, (sele), stored.pos.insert(0,(x,y,z))
4. label sele, ("%5.5s, %5.5s, %5.5s") % stored.pos.pop()
```

4.3.5 As coordenadas obtidas foram: **X:** 3.91 **Y:** 15.54 **Z:** 15.40, com uma grade com discretização de 0.25 Å e 22 Å em cada dimensão (X, Y, Z), totalizando 704.969 pontos.

4.3.6 Visualize a grade de energia sobreposta com a proteína e os cofatores.

4.3.7 Utilizaremos os valores padrão do algoritmo genético (**Standard**).

4.3.8 Identifique seu experimento e submeta o *docking*.

5. Análise dos Resultados

5.1 Entre no *link* abaixo para analisar os resultados³:

https://dockthor.lncc.br/v2/index.php?tab=DOCKING&page=RESULTS&jobId=emmsb2024_2wx2_cof_66bf4c50dfa47

5.2 Carregue o arquivo original do ligante (*2wx2_ligand.pdb*) para comparar os modos de ligação encontrados pelo *docking* e a conformação observada experimentalmente.

5.3 Faça o *download* e analise os arquivos gerados pelo programa **DockThor** (diretório *RESULTS*).

Arquivos	Descrição
dockthor.out	Armazena todo o log do programa, incluindo o tempo de execução.
*_run_X.pdb	Soluções líderes de cada <i>cluster</i> obtido na execução X. <u>Obs.</u> : Estes arquivos são fornecidos apenas simulações com apenas um ligante.
*_run_X.log	Resumo das energias dos líderes de cada <i>cluster</i> da execução X. <u>Obs.</u> : Estes arquivos são fornecidos apenas simulações com apenas um ligante.
result-ligand_ID.mol2	Soluções líderes de cada <i>cluster</i> dentre todas as execuções do algoritmo genético. ID é o identificador do ligante.
result-ligand_ID.log	Resumo das energias dos líderes de cada <i>cluster</i> . ID é o identificador do ligante.
parameters.txt	Arquivo dos parâmetros utilizados no <i>docking</i> .
bestranking.csv	Tabela contendo o melhor modo de ligação (selecionado conforme a energia total, <i>Etotal</i>) de cada ligante, ordenados segundo a predição da afinidade (<i>Score</i> , dado em kcal/mol).
bestranking.mol2	Arquivo contendo a estrutura 3D do melhor modo de ligação (selecionado de acordo com a energia total, <i>Etotal</i>) de cada

³ **Atenção:** os resultados de *docking* ficam disponíveis para acesso no portal e *download* por 20 dias a partir da data de submissão. Todos os arquivos são excluídos após este período.

	ligante, ordenados segundo a predição da afinidade (<i>Score</i> , dado em kcal/mol).
--	--

5.4 Submeta o *docking* novamente sem utilizar cofatores e compare os resultados com o anterior (i.e., usando cofatores).

5.5 Entre no *link* abaixo para analisar os resultados⁴:

https://dockthor.Incc.br/v2/index.php?tab=DOCKING&page=RESULTS&jobId=emmsb2024_2wx2_noCof_66b3cf0e62be3

Parte 2 - Estados de Protonação da Proteína – 1HXW

Nesta simulação, nosso objetivo é realizar o *docking* do ligante ritonavir, um inibidor altamente flexível da proteína HIV-1 protease, definindo o correto estado de protonação desta enzima. Para isso, testaremos diferentes protocolos, alterando principalmente o número de ligações rotacionáveis do ligante e alguns parâmetros do algoritmo genético.

1. Estado de protonação da proteína

- 1.1 Faça o download do arquivo do complexo 1HXW no PDB e salve como *1hwx_complex.pdb*.
- 1.2 Visualize as interações do ligante **RIT** com a proteína usando o programa Pymol.
- 1.3 Observe o estado de protonação dos resíduos Asp25 (cadeia A) e Asp25 (cadeia B).
- 1.4 Separe o arquivo do ligante (*1hwx_ligand.pdb*).
- 1.5 Utilizaremos o servidor PDB2PQR para definir os estados de protonação da enzima HIV1-protease através da predição de pKa em pH 7. Acesse o endereço: <https://server.poissonboltzmann.org/pdb2pqr>
- 1.6 Carregue o arquivo da proteína *1hwx_complex.pdb*.
- 1.7 Selecione a opção *Add/keep chain ids in the PQR file*.
- 1.8 Defina pH = 7.

⁴ **Atenção:** os resultados de *docking* ficam disponíveis para acesso no portal e *download* por 20 dias a partir da data de submissão. Todos os arquivos são excluídos após este período.

1.9 Submeta a análise (*Submit*).


1.10 Abra o arquivo *.propka* e observe os valores de pKa preditos para o resíduo de aminoácido Asp25 das cadeias A e B.

1.11 Retorne para a página de resultados e salve o arquivo de formado *.pqr* como **1hwx_protein_pdb2pqr.pdb** (clique com o botão direito do *mouse* Salvar link como...).

1.12 Abra o arquivo preparado da proteína (*1hwx_protein_pdb2pqr.pdb*) no Pymol. Verifique os hidrogênios adicionados e os estados de protonação do resíduo de aminoácido Asp25 das cadeias A e B.

1.13 Carregue o arquivo preparado da proteína no portal DockThor.

1.14 Observe os estados de protonação do resíduo Asp25 das cadeias A e B.

Obs.: Verifique a nomenclatura dos estados de protonação possíveis clicando no símbolo .

1.15 Salve o estado atual da proteína e passe para a etapa de preparação do ligante (*Send to DockThor*). Não utilizaremos cofatores neste experimento.

1.16 Carregue o ligante *1hwx_ligand.pdb* e adicione os átomos de hidrogênio.

1.17 Utilize como centro da grade de energia: **X**: 13.503 **Y**: 24.140 **Z**: 4.392 e dimensões 22 Å.

1.18 Mantenha os parâmetros padrão do algoritmo genético (**Standard**).

1.19 Submeta o *docking*.

1.20 Entre no *link* abaixo para analisar os resultados. Carregue o arquivo do ligante de referência para calcular o RMSD.

https://dockthor.incc.br/v2/index.php?tab=DOCKING&page=RESULTS&jobId=emmsb2024_1hwx_pdb2pqr_66c634c256348

1.21 Compare os resultados acima com os resultados do *redocking* (endereço abaixo) sem definir o estado de protonação correto do resíduo Asp25.

https://dockthor.incc.br/v2/index.php?tab=DOCKING&page=RESULTS&jobId=emmsb2024_1hwx_noProt_66c6339073cbd

Dia 3 – Alterando os parâmetros do *docking* e entendendo erros de interpolação

Parte 1 - Alterando os parâmetros do *docking* - 1HXW

1. Ligações rotacionáveis do ligante

- 1.1 Carregue o arquivo da proteína do complexo 1HXW preparada pelo PDB2PQR ou prepare o arquivo da proteína proveniente do PDB alterando o estado de protonação do Asp25 da cadeia B (ASPN2).
- 1.2 Não utilizaremos cofatores. Vá para a etapa de preparação do ligante.
- 1.3 Prepare o arquivo do ligante (*1hwx_ligand.pdb*), adicionando os hidrogênios.
- 1.4 Visualize os graus de liberdade torcionais do ligante. Realizaremos experimentos com variação da flexibilidade do ligante, conforme os itens abaixo:
 - (a) *Docking Rígido*: desmarque todas as caixas referentes às 19 torções do ligante.
 - (b) *Docking flexível*: deixe todas as caixas marcadas.

Após cada uma das duas modificações acima, siga os passos abaixo:

- 1.4.1 Clique em *Apply* para salvar as configurações de flexibilidade e passe para a próxima etapa (*Send to DockThor*).
- 1.4.2 Na aba *Docking*, digite seu e-mail, os valores para o centro da grade: **X**: 13.503 **Y**: 24.140 **Z**: 4.392 e 20 Å em cada dimensão.
- 1.4.3 Mantenha a configuração padrão do algoritmo genético (**Standard**).
- 1.4.4 Defina um nome para o experimento, identificando se a simulação é de *docking rígido* ou *flexível*.
- 1.4.5 Submeta o experimento de *docking*.
- 1.4.6 Entre nos *links* abaixo para analisar os resultados de (a) e (b), respectivamente:

https://dockthor.incc.br/v2/index.php?tab=DOCKING&page=RESULTS&jobId=emmsb2024_1hwx_rig_66ba4a348c9a8

https://dockthor.Incc.br/v2/index.php?tab=DOCKING&page=RESULTS&jobId=emmsb2024_1hwx_flex_66ba4b1b7a61c

2. Parâmetros do algoritmo genético (AG)

- 2.1.1 Prepare os arquivos da proteína HIV-1 protease e do ligante ritonavir (*1hwx_ligand.pdb*) conforme o protocolo anterior.
- 2.1.2 Não é necessário incluir nenhum cofator.
- 2.1.3 Consideraremos o ligante totalmente flexível nestas simulações.
- 2.1.4 Na aba *Docking*, utilize os valores para o centro da grade: X: 13.503, Y: 24.140, Z: 4.392 e 20 Å em cada dimensão.
- 2.1.5 Utilizaremos valores distintos para o parâmetro *Population Size* do algoritmo genético. Mude o valor deste parâmetro para cada um dos itens abaixo.
 - (a) 100
 - (b) 250
 - (c) 500
- 2.1.6 Dê um nome para cada experimento de *docking* identificando cada configuração utilizada e submeta os experimentos.
- 2.1.7 Entre nos *links* abaixo para analisar os resultados de (a), (b) e (c), respectivamente:

https://dockthor.Incc.br/v2/index.php?tab=DOCKING&page=RESULTS&jobId=emmsb2024_1hwx_p100_66ba4efd34be7

https://dockthor.Incc.br/v2/index.php?tab=DOCKING&page=RESULTS&jobId=emmsb2024_1hwx_p250_66ba51182456f

https://dockthor.Incc.br/v2/index.php?tab=DOCKING&page=RESULTS&jobId=emmsb2024_1hwx_p500_66ba51e6df068

- 2.1.8 Agora, vamos alterar o parâmetro *Number of Evaluations*. Repita os passos em 2.1.1 e 2.1.4.
- 2.1.9 Mude o parâmetro *Population Size* e deixe-o fixo em 500 para todos os experimentos.

2.1.10 Mude o valor do parâmetro *Number of Evaluations* de acordo com cada um dos itens abaixo. Após cada uma das alterações, submeta os experimentos de *docking* identificando-os.

(a) 100000

(b) 250000

(c) 500000

2.1.11 Entre nos *links* abaixo para analisar os resultados de (a), (b) e (c), respectivamente:

https://dockthor.lncc.br/v2/index.php?tab=DOCKING&page=RESULTS&jobId=emmsb2024_1hwx_av100_66ba599c324ca

https://dockthor.lncc.br/v2/index.php?tab=DOCKING&page=RESULTS&jobId=emmsb2024_1hwx_av250_66ba5aad7bc43

https://dockthor.lncc.br/v2/index.php?tab=DOCKING&page=RESULTS&jobId=emmsb2024_1hwx_av500_66ba5bbcb05a7

Parte 2 - Erro de interpolação - 3HH4

1. Configurações da grade de energia (espaço de busca) - 3HH4

A grade de energia determina a região do receptor que será considerada como sítio de ligação e, conseqüentemente, limita o espaço de busca em que as diversas conformações do ligante serão geradas. Os parâmetros da grade de energia influenciam nos resultados do *docking*: quanto maior a aproximação utilizada, menos acurados serão os resultados. Neste experimento, aprenderemos a selecionar o centro da grade de energia e utilizaremos diversas configurações. Utilizaremos uma proteína complexada com um ligante pequeno (benzeno, **BNZ**) para avaliar melhor diferentes configurações da grade de energia (código PDB: 3HH4).

1.1 Faça o download do arquivo do complexo 3HH4 no PDB e salve como *3hh4_complex.pdb*. Prepare os arquivos da proteína e do ligante. Não é necessário alterar nenhum estado de protonação da proteína nem utilizar cofatores.

1.2 Defina o centro da grade de energia: $X = 27.47$, $Y = 7.29$, $Z = 4.22$

1.3 Utilize uma grade de energia reduzida, com 14Å em cada dimensão e os seguintes valores de discretização da grade de energia:

- (i) 0.5 Å;
- (ii) 0.25 Å;
- (iii) 0.145 Å.

1.4 Altere os parâmetros do algoritmo genético: 50000 avaliações e 100 indivíduos na população inicial. Desative a opção de suavização do potencial de vdW (*Soft vdW*).

1.5 Compare os valores encontrados dos potenciais de Coulomb e van der Waals com os calculados pelo campo de força MMFF94 sem a aproximação por grade de energia (Coulomb = -0,43 kcal/mol, vdW = -10,31 kcal/mol, Einter = -10,74 kcal/mol).

- **0.5 Å:**

[https://dockthor.lncc.br/v2/index.php?tab=DOCKING&page=RESULTS
&jobId=emmsb2024_3hh4_i_66ba8a3209251](https://dockthor.lncc.br/v2/index.php?tab=DOCKING&page=RESULTS&jobId=emmsb2024_3hh4_i_66ba8a3209251)

- **0.25 Å:**

[https://dockthor.lncc.br/v2/index.php?tab=DOCKING&page=RESULTS
&jobId=emmsb2024_3hh4_ii_66ba8b61312c6](https://dockthor.lncc.br/v2/index.php?tab=DOCKING&page=RESULTS&jobId=emmsb2024_3hh4_ii_66ba8b61312c6)

- **0.154 Å:**

[https://dockthor.lncc.br/v2/index.php?tab=DOCKING&page=RESULTS
&jobId=emmsb2024_3hh4_iii_66ba8c4079fe9](https://dockthor.lncc.br/v2/index.php?tab=DOCKING&page=RESULTS&jobId=emmsb2024_3hh4_iii_66ba8c4079fe9)

Dia 4 – Triagem Virtual em Larga Escala – 2I78

Esta prática abordará o processo de triagem virtual usando o programa DockThor, proporcionando uma compreensão prática de como identificar compostos com melhor afinidade por um alvo específico, auxiliando na seleção de potenciais substâncias ativas.

1. *Re-docking*

Submeta um experimento de *re-docking* para calibrar os parâmetros da simulação e validar o protocolo adotado para a proteína dipeptidil peptidase IV (DPP4, código PDB 2I78).

- Utilize apenas a cadeia B; remova as outras cadeias do arquivo .pdb da proteína (arquivo: 2i78_chainB_protein.pdb)
- Não é necessário alterar os estados de protonação da proteína.
- Não é necessário adicionar cofatores.
- Determine o centro da grade de energia de uma destas formas:
 - o Átomo central do ligante de referência (KIQ)
 - o Resíduo de aminoácido importante no sítio de ligação

(Dica: use o PyMOL para obter as coordenadas do centro do eseoço de busca (grade) como feito no passo 11 da prática do dia 3:

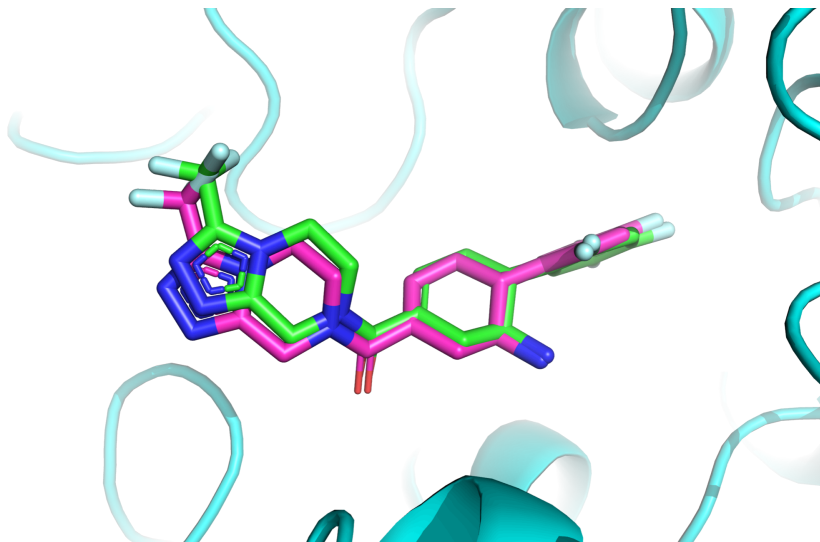
```
GRID SETTINGS
center x = -18.1
center y = -6.44
center z = 61.03
total size x = 22
total size y = 22
total size z = 22
discretization = 0.25
```

(c)

2. Analise os resultados

https://dockthor.lncc.br/v2/index.php?tab=DOCKING&page=RESULTS&jobId=emmsb2024_2i78_redocking_66c74f9251750

- o Você pode inserir a estrutura cristalográfica do ligante KIQ como molécula de referência para obter valores de RMSD (arquivo: 2i78_chainB_ligand.pdb).



Redocking do ligante KIQ na estrutura da proteína DPP4. A estrutura cristalográfica está colorida em magenta e a melhor pose encontrada no redocking está colorida em verde. Note a similaridade, referente a RMSD: 0.853 Å.

3. Validação para triagem

Para validar o protocolo utilizado em experimentos de triagem virtual, é essencial realizar uma triagem virtual "teste" que inclua compostos conhecidos como ativos e um conjunto de compostos inativos. Um protocolo de triagem virtual adequado deve ser capaz de identificar o maior número possível de compostos ativos entre aqueles com melhor predição de afinidade. Na ausência de um número suficientemente grande de compostos inativos comprovados experimentalmente, utilizam-se compostos "*decoys*", que são potencialmente inativos.

Nesta etapa de validação, faremos a triagem virtual de um conjunto com 10 compostos ativos (identificador "ChEMBL") e 90 compostos *decoys* (identificador "ZINC") provenientes do conjunto DUD-E para a proteína dipeptidil peptidase IV (DPP4).

3.1 Prepare o arquivo da proteína.

3.2 Carregue todos os arquivos .mol2 de todos os 100 compostos (dataset.mol2) e prepare-os sem adicionar hidrogênios (os compostos já vem preparados do portal DUD-E).

3.3 Utilize os mesmos parâmetros definidos no experimento de *re-docking*. Na configuração do algoritmo genético, mantenha a opção "Virtual Screening" selecionada.

3.4 Entre no *link* abaixo para analisar os resultados:

https://dockthor.lncc.br/v2/index.php?tab=DOCKING&page=RESULTS&jobId=emmsb2024_2i78_virtualScreening_66c7526949904

Referências

- I. A. Guedes, C. S. de Magalhães, L. E. Dardenne. **Receptor-ligand Molecular Docking**. Biophysical Reviews, **2014**.
- C. S. de Magalhães, D. M. Almeida, H. J. C. Barbosa, L. E. Dardenne. **A Dynamic Niching Genetic Algorithm Strategy for Docking of Highly Flexible Ligands**. Information Sciences, **2014**.
- **Bioinformática: da Biologia à Flexibilidade Moleculares**. Endereço para download: <https://www.ufrgs.br/bioinfo/ebook>
- PDB - H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne. **The Protein Data Bank**. Nucleic Acids Research, 28: 235-242, **2000**.
- Almeida, D.M., **DockThor: Implementação, Aprimoramento e Validação de um Programa de Atracamento Molecular Receptor-Ligante**. Laboratório Nacional de Computação Científica – Petrópolis-RJ, **2011**.
- Mysinger, M.M., Carchia, M., Irwin, J.J., Schoichet, B.K. **Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking**. Med. Chem., **2012**.

Links Interessantes

- PropKa – cálculo de pKa de resíduos de aminoácidos de proteínas: <http://propka.ki.ku.dk/>
- PDB2PQR - preparação de proteína como adição de hidrogênios, cálculo de cargas parciais (vários campos de forças) e cálculo de pKa dos resíduos: http://nbc-222.ucsd.edu/pdb2pqr_1.9.0/

- Corina - calcula a estrutura 3D de pequenas moléculas, a partir do SMILE ou do desenho da estrutura em 2D: https://www.molecular-networks.com/online_demos/corina_demo_interactive
- Informações sobre o formato PDB: <http://www.wwpdb.org/documentation/format32/v3.2.html>