

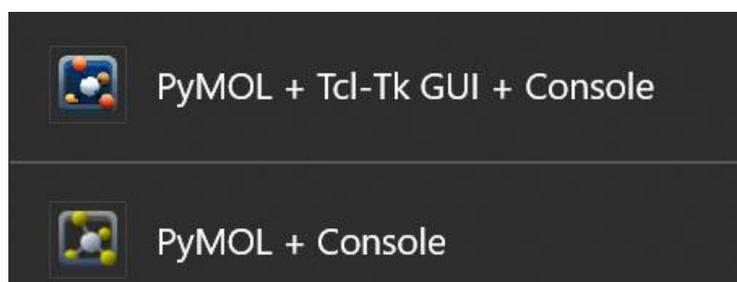
Tutorial – Docking com Vina

Aula Prática – 03/02/2022

– Re-docking utilizando o Autodock Vina

Vamos inicialmente utilizar a metodologia de reprodução de pose, chamada também de **re-docking**. A metodologia consiste em tentar reproduzir a posição adquirida pelo ligante na estrutura cristalográfica utilizando o programa de atracamento. Portanto, obtemos uma validação do protocolo e um teste do algoritmo de atracamento para o caso da proteína estudada. Ao final, podemos comparar interações e o RMSD entre pose e posição cristalográfica do ligante.

- 1) Abra o **PyMOL**.



OBS. 1: Dê preferência para o **PyMOL + Console**, assim o programa abre com mais opções nas abas.

- 2) Abra a estrutura **PDB 7L10**, um complexo da protease principal do SARS-CoV-2 (*Main protease*) com um inibidor otimizado de uma campanha de triagem virtual de 2000 compostos aprovados. Essa protease é ativa apenas em forma de dímero, porém, vamos trabalhar apenas com cadeia A para simplificar.

```
PyMOL> fetch 7L10
```

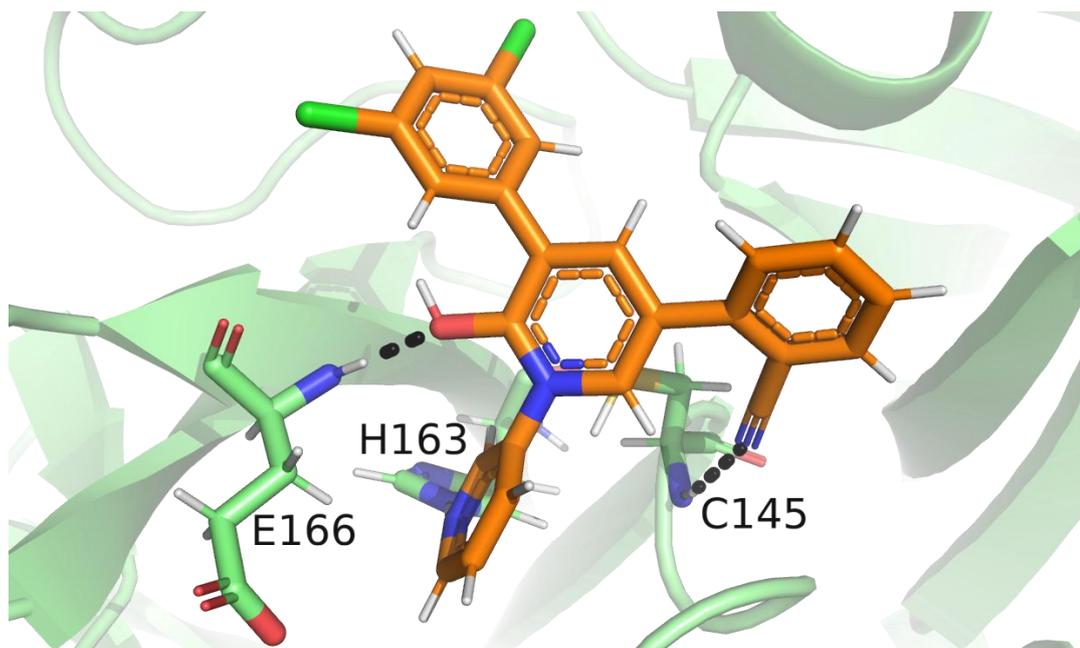
- 3) Crie um objeto contendo apenas o inibidor. Para extrair um objeto nomeado **ligante**, contendo a molécula nomeada **XEY** no arquivo PDB:

```
PyMOL> extract ligante, resn XEY  
PyMOL> show sticks, ligante  
PyMOL> util.cbam ligante
```

- 4) Vamos analisar as interações intermoleculares entre inibidor e enzima, para verificar quais interações temos da estrutura cristalográfica. Calcule as interações entre átomos polares do inibidor e a enzima ou moléculas de água.
- 5)
 - Selecione o ligante no menu de seleção



- No menu **Actions (A)** da seleção, clique em **Find – polar contacts – to any atoms**



Foi encontrada interações polares com a C145 (cisteína catalítica), H163 e E166.

- 6) Adicione hidrogênios (Estruturas PDB originadas de cristalografia normalmente não possuem hidrogênios).

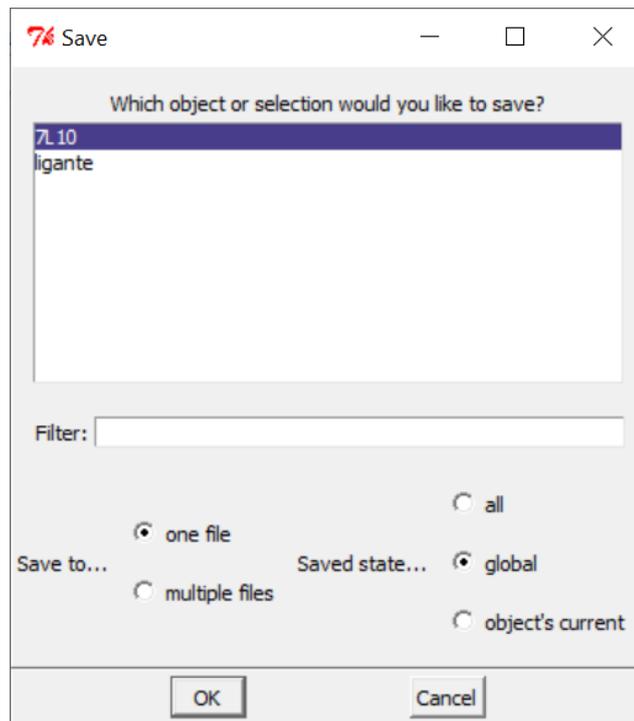
PyMOL> **h_add**

OBS. 2: Para simplificar o tutorial as protonações dos resíduos ASP, CYS, HIS, LYS, e TYR, não foram exploradas. O correto é buscar na literatura a protonação correta para o pH de ativada ou utilizar um programa/servidor de predição com o PROPKA (<https://github.com/jensengroup/propka>) ou H++ (<http://biophysics.cs.vt.edu/>).

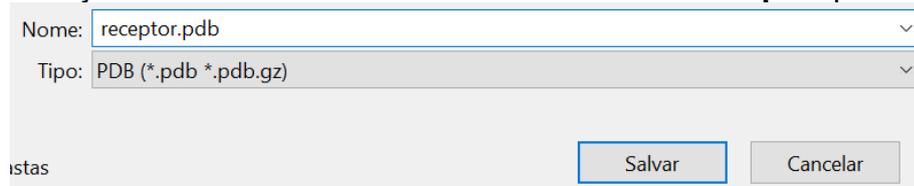
- 7) Para o atracamento vamos primeiro remover as moléculas de água (em alguns casos podemos manter algumas moléculas de águas, mas essas moléculas precisam ser analisadas com cuidado):

PyMOL> **remove resn hoh**

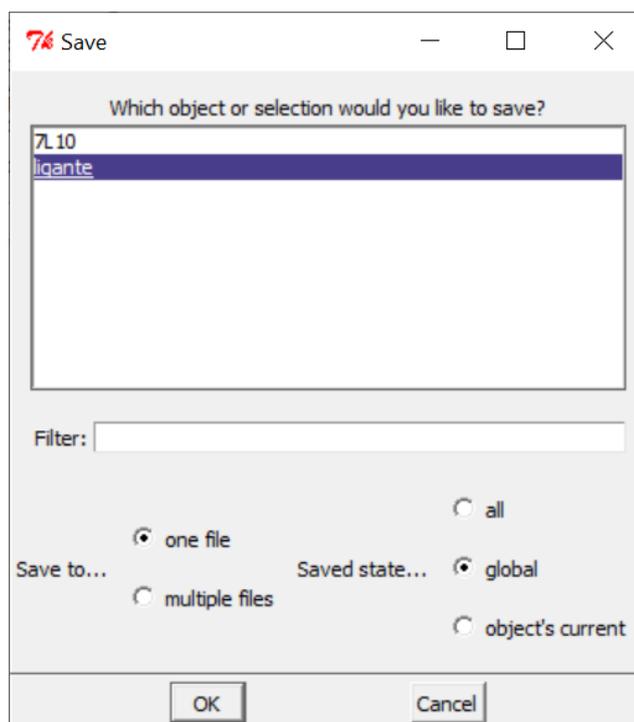
- 8) Apesar da estrutura ativa ser o dímero, vamos trabalhar apenas com cadeia A. Vamos salvar arquivos separados para o **receptor** e para o **ligante**. Em **File > Save Molecule**



Selecione o objeto 7L10. E dê **OK**. Salve com o nome de **receptor** para a proteína.



Faça o mesmo para o **ligante**, salve como **ligante**.



Nome: ▼

Tipo: ▼

stas

- 9) O Vina não faz um pré-cálculo de Grid. Porém, uma demarcação da área de docking (sítio ativo ou até mesmo a proteína toda) precisa ser estabelecida. Como estamos usando uma estrutura que já possui o ligante podemos calcular o centro de massa do ligante para ser o centro da nossa Grid.

PyMOL> **centerofmass ligante**

```
PyMOL>centerofmass ligante
Center of Mass: [ 8.417, -0.060, 22.273]
```

Anote esses números. Vamos utilizá-los logo adiante.

- 10) Podemos fechar o **PyMOL** não vamos precisar dele para os próximos passos. A preparação para o docking será feita por **linha de comando**. Já que o programa de docking que vamos utilizar, **Vina**, foi projetado para não ter uma interface e com isso ser um programa mais leve.

- 11) Para preparar os arquivos temos que **ACESSAR O SERVIDOR** no terminal de vocês:

```
$ ssh bioufmg@bioinfo.icb.ufmg.br
```

DIGITE A SENHA.

- 12) Encontre a sua pasta (com seu nome):

```
lucianna@DESKTOP-JQF7JLS:~$ ssh bioufmg@bioinfo.icb.ufmg.br
bioufmg@bioinfo.icb.ufmg.br's password:
Last login: Wed Oct 14 17:09:15 2020 from 201.17.140.105
[bioufmg@bioinfo ~]$ ls
INTRO          elisaamancio   larama         milenap
ace222         eusoujacu      larissa_cler   monic_lops
amandacpa      flavia         laryhenriques moysesmn
anna_menezes   flaviam        letbarbosa     nathalia_alv
arthurtrcf     franciscoacarmo letchris        pedrobala
barbaramas     fredericogabriel lucasduque     rayssa_soares
beatrizapgaua gabijager      luizacaixeta   renan
caizin         gabriel_ferrari macl3y         tarefa1
camilabd       gabrielrod     madumascarenhas tarefa2
camilacllc     gapmo          mari_1305       teste
ceciliahorta   jessica_alves marinak        thiagoac
clecio_alonso  jorge          markko
danikayali     karenkanda     mbarcelos
dg_sousa       key            michelle_157
[bioufmg@bioinfo ~]$
```

*se não possuir pasta. Crie uma com o comando **mkdir <seunome>**

- 13) Acesse sua pasta com

```
$ cd minhapasta
```

14) Confira se você está realmente na pasta com o seu nome:

```
$ pwd
```

15) Dentro da sua pasta, crie uma nova pasta chamada docking_pratica:

```
$ mkdir docking_pratica
```

16) Acesse a pasta

```
$ cd docking_pratica
```

17) Copie o conteúdo da pasta /home/treinamento/novo_docking_pratica para cá:

```
$ cp /home/treinamento/novo_docking_pratica/* .
```

Pronto. Temos um diretório para rodar preparar os arquivos de entrada e rodar o Vina.

```
config.txt ligante.pdb receptor.pdb
```

18) Para o docking com o **Vina** precisamos dos arquivos de entrada no formato **PDBQT**. Esse formato é uma adaptação do formato PDB. Nele além das coordenadas atômicas temos informação de carga e descrição dos átomos, específico para os programas **Vina** e **Autodock4**.

Para criar o PDBQT do receptor basta:

```
$ pythonsh $mglscripsts/prepare_receptor4.py -r receptor.pdb
```

OBS. 3: pythonsh é uma versão (antiga) compilada do python para o pacote MGLTools. O caminho dele é:

```
/opt/treinamento/mgltools_x86_64Linux2_1.5.7rc1/bin/python
```

OBS. 4: \$mglscripsts é um PATH criado para não precisar digitar todo o caminho até o script. O caminho é:

```
/opt/treinamento/mgltools_x86_64Linux2_1.5.7rc1/MGLToolsPkgs/AutoDockTools/Utilities24/prepare_receptor4.py
```

19) Para criar o PDBQT do ligante basta:

```
$ pythonsh $mglscripsts/prepare_ligand4.py -l ligante.pdb -U "
```

OBS. 5: -U “ pede para que o script mantenha todos os hidrogênios do ligante. Tirando essa “flag”, apenas os hidrogênios polares permaneceriam.

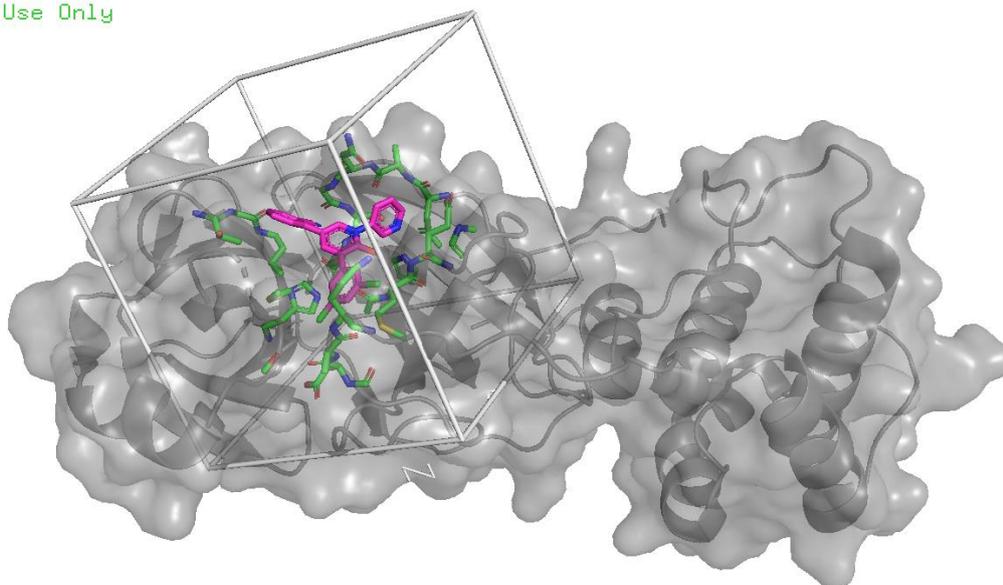
20) O último arquivo de entrada necessário é o arquivo com os parâmetros do **Vina**. É um simples arquivo de texto (vamos chamar de **config.txt**) contendo os seguintes

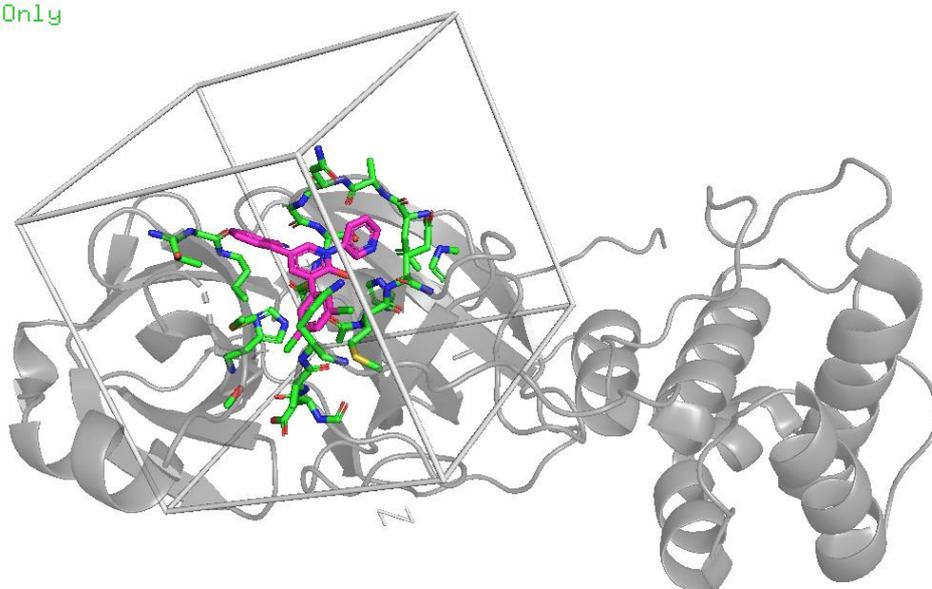
parâmetros:

```
receptor = receptor.pdbqt      #Nome do arquivo PDBQT do receptor
ligand = ligante.pdbqt        #Nome do arquivo PDBQT do ligante
center_x = 8.417              #Coordenada X do centro da caixa (aqui o centro do ligante)
center_y = -0.06              #Coordenada Y do centro da caixa (aqui o centro do ligante)
center_z = 22.273            #Coordenada Z do centro da caixa (aqui o centro do ligante)
size_x = 25.00                #Tamanho da caixa para o eixo X
size_y = 25.00                #Tamanho da caixa para o eixo Y
size_z = 25.00                #Tamanho da caixa para o eixo Z
out = docked_ligante.pdbqt    #Nome do arquivo de saída com as poses do docking
log = docked.log               #Arquivo de log com informações da saída do docking
num_modes = 10                #Quantidade de poses
```

Center_x, center_y e center_z são as coordenadas do centro de massa do ligante calculados anteriormente no Pymol. Isso quer dizer que o programa terá sua “região de busca” centrada nessas coordenadas. Para delimitar uma região que compreenda parte da proteína ou um sítio de interesse temos os parâmetros size_x, size_y e size_z, que no nosso caso, estende **25 Å** para cada eixo desse ponto formando uma caixa cúbica nessa região da proteína. A região para docking será então:

For Educational Use Only





21) Agora que temos todos os arquivos necessário, podemos rodar o **Vina**. Utilizamos o comando:

```
$ vina --config config.txt
```

A seguinte tela vai aparecer:

```
#####  
# If you used AutoDock Vina in your work, please cite: #  
# #  
# O. Trott, A. J. Olson, #  
# AutoDock Vina: improving the speed and accuracy of docking #  
# with a new scoring function, efficient optimization and #  
# multithreading, Journal of Computational Chemistry 31 (2010) #  
# 455-461 #  
# DOI 10.1002/jcc.21334 #  
# Please see http://vina.scripps.edu for more information. #  
#####  
Detected 4 CPUs  
Reading input ... done.  
Setting up the scoring function ... done.  
Analyzing the binding site ... done.  
Using random seed: 1941012439  
Performing search ...  
0% 10 20 30 40 50 60 70 80 90 100%  
|----|----|----|----|----|----|----|----|----|----|  
|*****|  
done.  
Refining results ... done.  
  
mode | affinity | dist from best mode  
 | (kcal/mol) | rmsd l.b. | rmsd u.b.  
-----+-----+-----+-----  
1 | -8.7 | 0.000 | 0.000  
2 | -8.6 | 1.052 | 2.447  
3 | -8.2 | 1.623 | 5.531  
4 | -8.2 | 1.723 | 5.610  
5 | -8.1 | 1.900 | 2.074  
6 | -8.0 | 2.909 | 6.183  
7 | -8.0 | 3.143 | 6.524  
8 | -7.9 | 3.091 | 6.643  
9 | -7.8 | 4.194 | 6.756  
10 | -7.5 | 3.382 | 7.239  
Writing output ... done.
```

OBS. 6: O programa Vina é aleatório e a cada rodada ele escolhe um número aleatório, dado pelo parâmetro chamado seed. Portanto, os resultados podem ser diferentes em cada rodada.

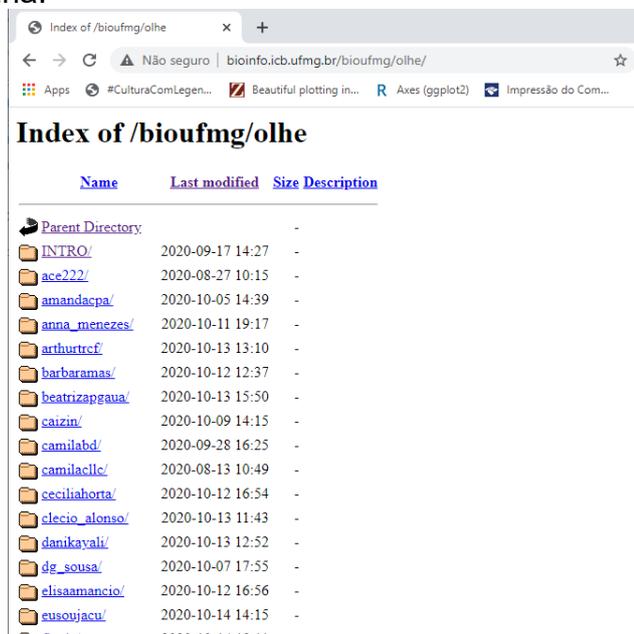
22) Pronto! Já temos os modos calculados pelo **Vina**.

Para verificar os resultados, vamos baixar o arquivo **docked_ligante.pdbqt** para o nosso computador e abrir pelo PyMOL, junto com o ligante cristalográfico **ligante.pdb**.

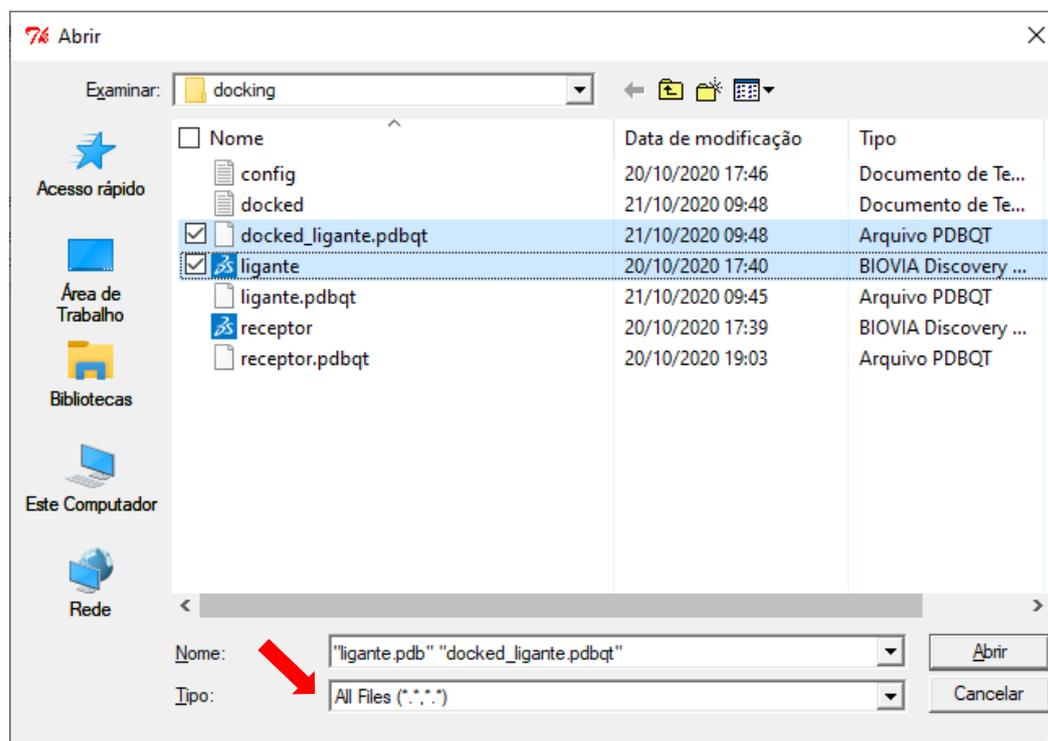
Para isso utilizamos o site

bioinfo.icb.ufmg.br/biufmg/olhe

Digite usuário e senha.

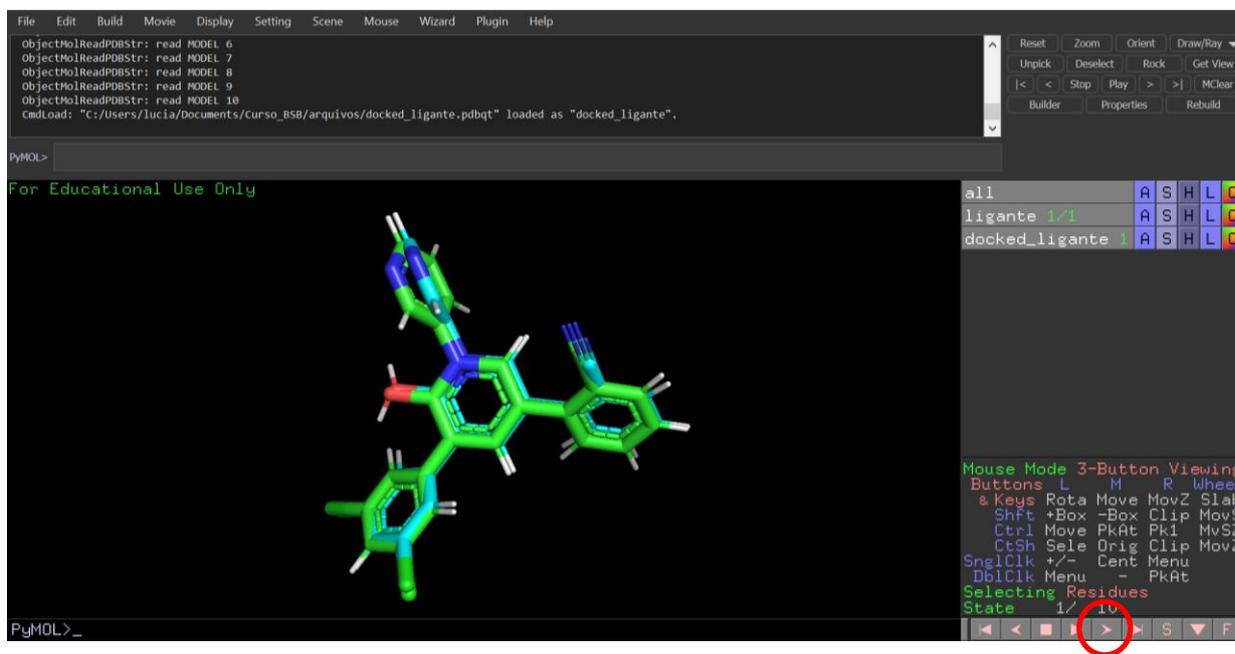


No PyMOL vamos em **File > Open** navegar até a pasta onde estão os arquivos:



Caso o arquivo PDBQT não aparecer, mude o **Tipo** para **All Files**. E selecione os arquivos.

As poses abrem todos juntas na visualização **docked_ligante.pdbqt**. Para ir passando por elas clique no botão **>** no canto direito inferior.



23) Olhando todas as **poses**, a que mais se aproximou do ligante cristalográfico foi a primeira (melhor valor de pontuação **-8.7 kcal/mol**). Ou seja, temos aqui um exemplo de sucesso no **re-docking**, onde o programa encontra a pose cristalográfica e ranqueia como sendo o melhor de todos os modos calculados.

24) Para confirmarmos o sucesso do **re-docking**, calculamos o valor de RMSD das primeiras poses. Se o valor do RMSD estiver abaixo de 2.0 Å para a primeira pose, temos um sucesso tanto da função de pontuação quando do algoritmo de busca. Porém, se a primeira pose estiver acima de 2.0 Å, enquanto as poses de colocação próxima possuírem RMSD abaixo, temos uma falha na pontuação.

Voltando ao **TERMINAL**, ou seja, fora do **PyMOL**. Vamos primeiro separar cada uma das poses com o programa **vina_split**:

```
$ vina_split --input docked_ligante.pdbqt
```

```
docked_ligante_ligand_01.pdbqt
docked_ligante_ligand_02.pdbqt
docked_ligante_ligand_03.pdbqt
docked_ligante_ligand_04.pdbqt
docked_ligante_ligand_05.pdbqt
docked_ligante_ligand_06.pdbqt
docked_ligante_ligand_07.pdbqt
docked_ligante_ligand_08.pdbqt
docked_ligante_ligand_09.pdbqt
docked_ligante_ligand_10.pdbqt
```

E usando um script do próprio **MGLTools**, vamos calcular o RMSD ente ligante e pose:

```
$ pythonsh $mglscripsts/compute_rms_between_conformations.py -f  
ligante.pdq -s docked_ligante_ligand_01.pdbqt
```

Para o re-docking ser considerado um **sucesso**, a pose de **melhor score (a primeira colocada)** tem que alcançar o RMSD **abaixo de 2.0 Å**. No caso ilustrado a pose de menor score atingiu esse critério. Porém, para casos que isso não acontece tem que se ter cuidado para olhar todo o grupo de poses ranqueadas com cuidado para não descartar verdadeiros positivos.

OBS. 7: A verificação de RMSD só pode ser feita quando temos um ligante cristal para comparar. Em um experimento de *docking* com ligantes novos esse parâmetro não pode ser calculado. Porém, as interações formadas entre proteína e ligante podem (e devem) ser analisadas e exploradas.

25) Vamos analisar as interações intermoleculares entre essa e a enzima. Abra o arquivo do receptor. Em **File > Open** navegar até a pasta onde estão os arquivos. Selecione o arquivo **receptor.pdb**

Selecione **a pose** no menu de seleção. No menu **Actions (A)** da seleção, clique em

Find – polar contacts – to any atoms

Alguma interação foi igual a vista entre ligante do cristal e receptor?

Você pode fazer a mesma coisa para o ligante do cristal e mudar a cor no menu **Color (C)**.

26) Salve (**File > Save Session**) como **Mpro_dock1.pse**. Feche o **PyMOL**.