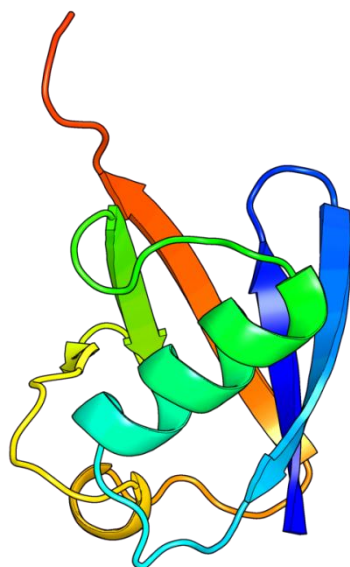




Manual de Dinâmica Molecular com Gromacs



Versão.:	0.1
Criado em.:	14/10/2024
Última atualização.:	17/01/2025

Autor.:	Rogério Ribeiro Macêdo
Lattes.:	http://lattes.cnpq.br/8806221981552346
GitHub.:	https://github.com/rogerioribeimacedo



Sumário

Introdução	3
Estrutura de diretórios	4
Preparando o sistema	5
SOLVATAÇÃO	7
Adicionando íons	8
Minimização	9
Equilíbrio	10
Ensemble NVT (temperatura)	11
Anexo 1	Erro! Indicador não definido.
Anexo 2	Erro! Indicador não definido.
Anexo 3	Erro! Indicador não definido.

INTRODUÇÃO

A escolha da ubiquitina não se fez aleatoriamente. Ela foi tema do Prêmio Nobel de Química no ano de 2004¹. Os trabalhos de Aaron Ciechanover, Avram Herskho e Irwin Rose permitiram um melhor entendimento em como as células regulam a presença de proteínas que de alguma forma estejam degradadas. Para isso, o organismo faz uso da ubiquitina que age como uma espécie de sinalizador indicando para o organismo que a proteína que recebe este sinal deve ser reciclada pelas proteossomas (os centros de reciclagem do corpo).

Outro motivo para a escolha da ubiquitina é seu pequeno tamanho, sendo composta por apenas 76 aminoácidos, como mostra a Figura 1.

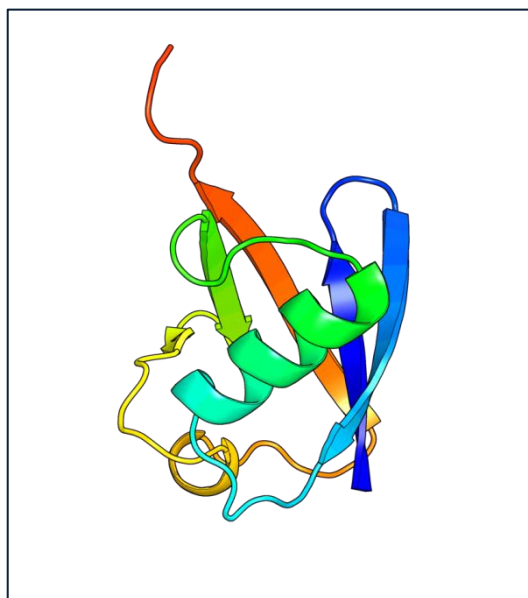


Figura 1 Ubiquitina, polipeptídeo que age como sinalizador celular para proteínas que devem ser degradadas.

Há quantidade satisfatória de estudos usando dinâmica molecular com a ubiquitina, por exemplo, o artigo: “Ubiquitin Dynamics in Complexes Reveal Molecular Recognition Mechanisms Beyond Induced Fit and Conformational Selection” (PETERS; DE GROOT, 2012). No entanto, o presente manual não pretende ser uma análise comparativa ou uma revisão bibliográfica de tais trabalhos. Também gostaria de indicar o artigo "Molecular Dynamics

¹ <https://www.nobelprize.org/prizes/chemistry/2004/popular-information/>

Simulation for All" (HOLLINGSWORTH; DROR, 2018), que fornece uma ampla gama de resultados que podem ser extraídas de simulações de dinâmica molecular.

Saliento também que não é objetivo deste manual tratar de todas as possibilidades de análise que podem ser feitas com o Gromacs em relação à simulação de uma proteína ou qualquer outra estrutura. Portanto, nos limitaremos ao uso de configurações de uso geral (pelo menos no que tange aos conhecimentos do autor neste momento).

Os arquivos necessários para a simulação e também códigos usados para visualizar as estruturas no Pymol podem ser encontrados no seguinte repositório: https://github.com/rogerioribeiromacedo/md_ubiquitina_manual.

ESTRUTURA DE DIRETÓRIOS

Para uma melhor organização da simulação seguiremos uma estrutura de diretórios² de modo tal que cada parte da dinâmica molecular irá gerar os dados em diretórios diferentes permitindo um melhor controle do processo e também por uma questão gerenciamento de dados. Para tanto, crie a estrutura de diretórios mostrada na Figura 2 dentro da sua pasta de simulação.

```
[rogerio@laqc7 ubiquitina]$ ls -lah
total 36K
drwxr-xr-x. 9 rogerio rogerio 4.0K Oct 17 08:09 .
drwxr-xr-x. 6 rogerio rogerio 4.0K Oct 17 08:08 ..
drwxr-xr-x. 2 rogerio rogerio 4.0K Oct 17 08:27 01_preparar_ambiente
drwxr-xr-x. 2 rogerio rogerio 4.0K Oct 17 08:09 02_minimizacao
drwxr-xr-x. 2 rogerio rogerio 4.0K Oct 17 08:09 03_equilibrio_nvt
drwxr-xr-x. 2 rogerio rogerio 4.0K Oct 17 08:09 04_equilibrio_npt
drwxr-xr-x. 2 rogerio rogerio 4.0K Oct 17 08:09 05_producao
drwxr-xr-x. 2 rogerio rogerio 4.0K Oct 17 08:09 06_analises
drwxr-xr-x. 2 rogerio rogerio 4.0K Oct 17 08:09 mdp
[rogerio@laqc7 ubiquitina]$
```

Figura 2 Estrutura de diretórios da simulação.

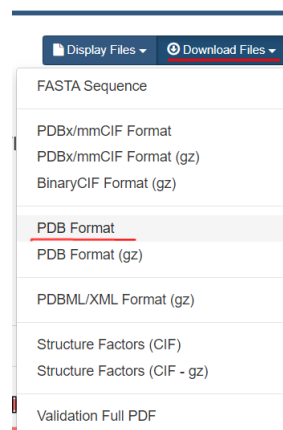
No repositório informado acima você encontrará os arquivos que contém os parâmetros da dinâmica molecular (mdp – “*molecular dynamics parameters*”). Cada arquivo representa uma fase da dinâmica e deve ser salvo no diretório “mdp”.

² Esta é a estrutura de diretório que eu pessoalmente uso. Você pode, e deve, ajustá-la às suas necessidades.

Uma observação a ser feita é em relação à versão do Gromacs sob o qual este manual foi produzido, no caso, a versão 2024.4.

PREPARANDO O SISTEMA

O primeiro passo é realizar o download da proteína em questão, para isso acesse o site [Protein Data Bank](https://www.rcsb.org/), pesquise pelo código 1UBQ (ubiquitina humana) e faça o download do arquivo pdb para dentro do diretório “01_preparar_ambiente”, nomeando como 1ubq.pdb.



A visualização da estrutura pode ser feita por uma diversidade de programas, no caso deste manual, todas as imagens são visualizadas usando o Pymol³ (escolha aquele com o qual esteja mais confortável). A Figura 3 apresenta como a nossa proteína é visualizada no Pymol. Você pode utilizar a interface do programa para chegar ao resultado, ou, pode executar a sequência de comandos encontradas no arquivo “...manual\diversos\imagem_ubiquitina\preparar_imagem_1ubq.txt”, no repositório.

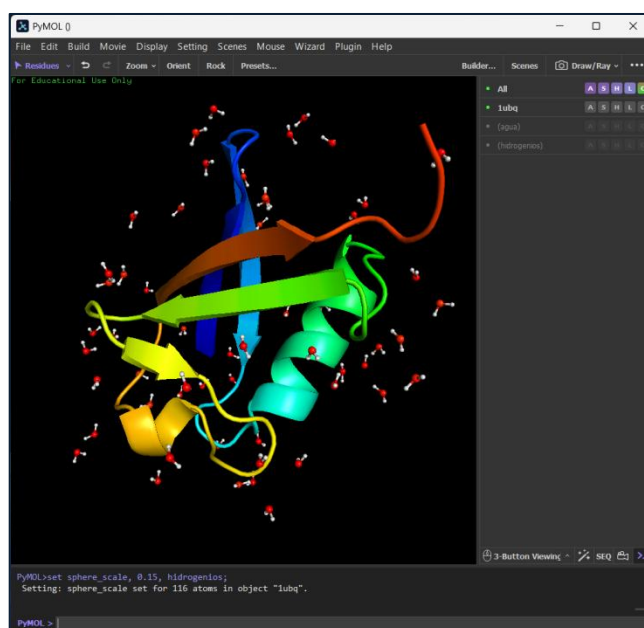


Figura 3 Ubiquitina vista no Pymol

³ <https://pymol.org/>

Como se trata de um estrutura que é resultado de uso de difração de raio-X, o arquivo .pdb pode conter elementos que não nos interessa, por exemplo, moléculas de água que são vistas na Figura 3. Precisamos remover as referências às águas presentes no arquivo. Isso pode ser feito usando um editor de texto simples, no entanto, os comandos aqui serão executados em terminal Linux, para isso usamos o comando *grep*:

```
grep -v HOH lubq.pdb > lubq_limpo.pdb
```

- “-v HOH”: retorna as linhas do arquivo que não contêm o valor, no caso, ‘HOH’. Que representa a molécula de água.
- “lubq.pdb”: arquivo de entrada que será analisado.
- “>” : indica que a saída do comando deve ser adicionada no arquivo informado.

Após a execução do comando temos então o arquivo lubq_limpo.pdb sem nenhuma referência às moléculas de água.

O próximo passo é utilizar o comando *gmx pdb2gmx* que irá gerar arquivos que serão utilizados na simulação da dinâmica molecular. Três tipos de arquivos serão gerados:

1. Arquivo de coordenadas da molécula (.gro): contém as posições dos átomos da molécula, também podendo apresentar velocidades nos eixos x, y e z.
2. Arquivos de restrições (.itp): quais átomos terão suas posições restritas durante a simulação
3. Arquivo de topologia (.top): define a molécula, apresentando informação sobre ligações, ângulos, diedros, campo de força e outros parâmetros.

O comando é apresentado abaixo: ao executá-lo, será solicitado a escolha do campo de força que deseja usar, escolha a opção referente a “*OPLS-AA/L all-atom force field (2001 aminoacid dihedrals)*” .

```
gmx pdb2gmx -f lubq_limpo.pdb -o lubq.gro -water spce
```

- “-f lubq_limpo.pdb”: arquivo de entrada
- “-o lubq.gro”: arquivo de saída com as coordenadas
- “-water spce”: modelo de água a ser usado na simulação

Após a execução do comando três arquivos estarão presentes: 1ubq.gro, posre.itp e topol.top. É interessante que você abra os arquivos gerados para ter uma noção de seu conteúdo e formato. Lembrando que o **1ubq.gro** contém a estrutura de nossa proteína, e que pode ser visualizada com o Pymol (os comandos usados para a visualização do 1ubq.pdb podem ser usado para visualizar o 1ubq.gro, basta trocar a referência ao arquivo) ou outro aplicativo de sua preferência. O **topol.top** contém informações sobre qual campo de força foi escolhido, referência ao arquivo de restrição, o modelo de água utilizado e qual a composição do sistema (" [molecules]")

SOLVATAÇÃO

Nosso objetivo é simular a proteína ubiquitina em um sistema aquoso, para tanto, vamos definir as dimensões de nossa caixa de simulação e então adicionar moléculas de água. Para definir a caixa de simulação execute o comando abaixo:

```
gmx editconf -f 1ubq.gro -o 1ubq_caixa.gro -c -d 1.0 -bt cubic -box 5.0
```

- "-c": irá centralizar a proteína no centro da caixa.
- "-o 1ubq_caixa.gro": arquivo de saída com as coordenadas e a nova dimensão da caixa.
- "-d 1.0": a proteína estará a no mínimo 1 nm de distância da borda da caixa.
- "-bt cubic": define-se a caixa como um cubo.

Definida a nova caixa, estamos prontos para adicionar o solvente (água), para isso execute o comando como segue. Como resultado, será adicionado à caixa uma um conjunto de 3657 moléculas de água

```
gmx solvate -cp 1ubq_caixa.gro -cs spc216.gro -o 1ubq_solv.gro -p topol.top
```

- "-cp": a configuração da proteína que é resultado do comando anterior.
- "-cs spc216.gro": configuração do solvente usado. Uma molécula de

água (modelo de 3 pontos).

- “-o lubq_solv.gro”: a proteína com as moléculas de água.
- “-p topol.top”: arquivo de topologia.

Agora abra o arquivo topol.top e veja que temos alterações nas seções [*molecules*] e [*system*].

```
[ system ]  
; Name  
UBIQUITIN in water  
  
[ molecules ]  
; Compound      #mols  
Protein_chain_A   1  
SOL               3657
```

A Figura 4 apresenta a caixa de simulação com a ubiquitina e todas as moléculas de água adicionada simulando o solvente escolhido. Os comandos necessários podem ser encontrados em: ...manual\diversos\ubiquitina_solvatada\LEIAME_ubiquitina_solvatada.txt, no repositório.

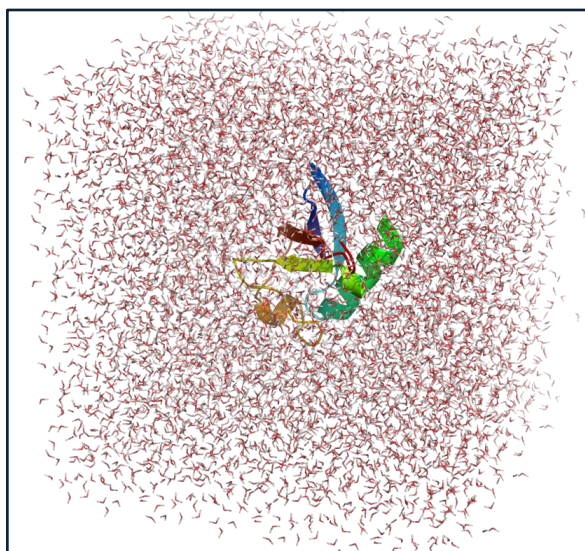


Figura 4 Sistema solvatado

ADICIONANDO ÍONS

A proteína escolhida é neutra, portanto, não há necessidade de adicionar íons para neutralizar o sistema. Essa carga zero do sistema é apresentada na saída do *gmx trajconv*. Abra o arquivo *topol.top* e verifique a última linha da seção *[atoms]*, deve encontrar um texto com “; *qtot 0*”.

MINIMIZAÇÃO

Nosso sistema agora está pronto para que iniciemos o processo de dinâmica molecular. Mas precisamos garantir que o sistema como um todo não apresente estruturas que estejam uma sobre a outra, que as distâncias entre moléculas sejam apropriadas, ou seja, vamos realizar um procedimento de minimização da energia potencial. Isso irá garantir que o sistema possa passar para as próximas etapas.

Primeiramente verifique se no diretório “mdp” existe uma arquivo chamado *em.mdp*, nele encontram-se os parâmetros necessários ao processo de minimização. Então, acesse a pasta 02_minimização e execute o comando abaixo, ele irá gerar os dados iniciais para o procedimento.

```
gmx grompp -f ../mdp/em.mdp
-c ../01_preparar_ambiente/lubq_solv.gro
-p ../01_preparar_ambiente/topol.top -o em.tpr -pp em.top -po em.mdp
```

Executado o comando, o próximo passo é executar o procedimento de minimização propriamente dito com o comando:

```
gmx mdrun -deffnm em &
```

Após a finalização do processo podemos então extrair os dados da evolução da energia potencial para verificar se o sistema atingiu um estado de equilíbrio nos permitindo continuar o processo. Para isso vamos usar o comando “*gmx energy*”, a sintaxe do comando é vista logo abaixo.

```
echo "Potential" | gmx energy -f em.edr -o em_energia_potencial.xvg
```

- echo "Potential": sem essa expressão um menu é apresentado e o usuário precisa escolher qual a opção de energia que será extraída.
- "-f em.edr": arquivo que o Gromacs usa para armazenar os dados sobre energia
- "-o <arquivo>": o arquivo de saída. O Gromacs salva em formato para Xmgrace. Mas pode ser visualizado por outras aplicações, porém poderá ser necessário excluir as primeiras linhas.
- "-xvg none": se desejar que o arquivo de saída seja gerado sem valores de cabeçalho no arquivo.

A Figura 5 mostra o gráfico que é construído com os dados do comando anterior. Podemos perceber que a energia potencial atinge um estado de equilíbrio após, aproximadamente, 7.500 passos. Portanto, todas as moléculas de água e a nossa proteína estão posicionadas de modo a minimizar a energia potencial do sistema.

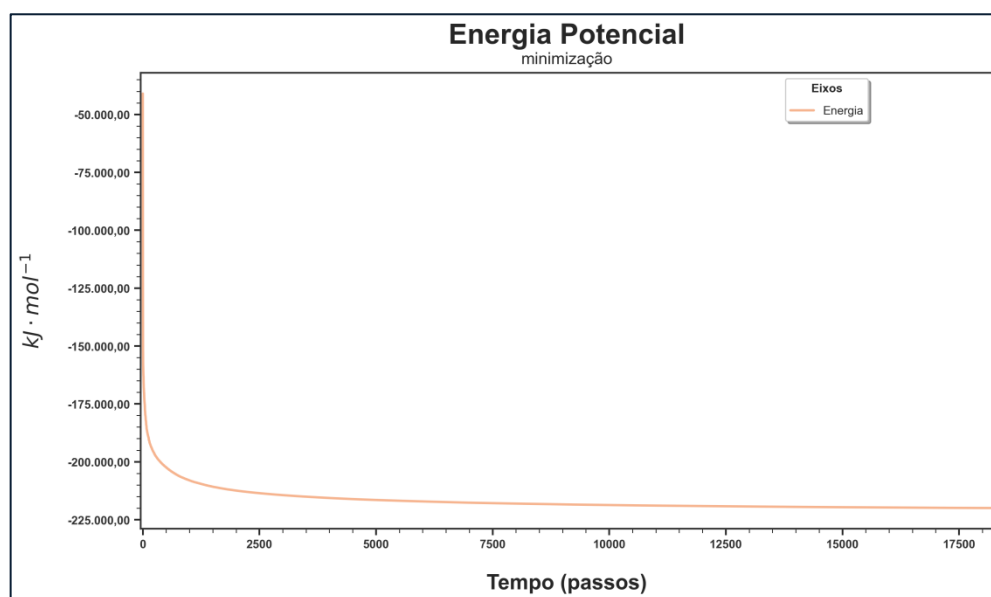


Figura 5 Energia Potencial (minimização)

EQUILÍBRIO

Com o sistema em equilíbrio em termos da orientação do solvente que estamos usando e a relação com a proteína escolhida, passamos à fazer de equilíbrio. Nesta, define-se a temperatura com a qual iremos simular o nosso sistema e, também, a pressão que permite atingir uma dada densidade que represente as condições experimentais desejadas.

Ensemble NVT (temperatura)

O ensemble NVT quer dizer: número constante de partículas, volume e temperatura. É também conhecido como canônico. O objetivo é fazer com que o sistema atinja um patamar de temperatura que no caso é 25 °C ou 298,15 Kelvin. Para isso, acesse a pasta 03_equilibrio_nvt e execute os comandos abaixo. Para esse caso, você precisará da presença do arquivo “nvt.mdp” na pasta mdp.

```
gmx grompp -f ../mdp/nvt.mdp -c ../02_minimizacao/em.gro -  
r ../02_minimizacao/em.gro -p ../01_preparar_ambiente/topol.top -o  
nvt.tpr -pp nvt.top -po nvt.mdp  
  
gmx mdrun -deffnm nvt &  
• Não esqueça o símbolo “&”. Ele indica ao Linux que a aplicação  
será executada e então o prompt de comando liberado.
```

Após a conclusão do processo podemos, também, avaliar a evolução da temperatura. Para isso, usamos o comando “*gmx energy*”.

```
echo "Temperature" | gmx energy -f nvt.edr -o nvt_temperatura.svg
```

A Figura 6 mostra o gráfico da evolução da temperatura e como esperado há variações durante o processo, no entanto, a temperatura está em torno de uma média não apresentando nenhuma grande discrepância. Sendo assim, o sistema está equilibrado a uma dada temperatura.

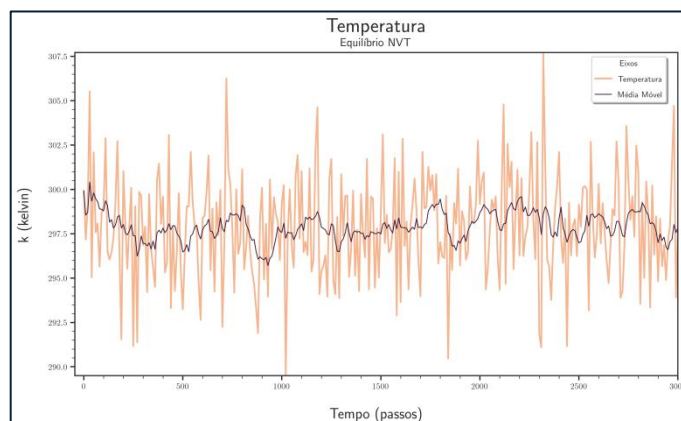


Figura 6 Temperatura (NVT)

Ensemble NPT (pressão)

O próximo passo é realizar o equilíbrio em relação à pressão. O ensemble NPT quer dizer: número constante de partículas, pressão e temperatura. É necessário agora fazer com que o sistema seja equilibrado para que possa estar a uma dada pressão, que no caso será 1 bar e também a temperatura será mantida a mesma do processo anterior. Usaremos o arquivo “*npt.mdp*” que deve estar na pasta *mdp*. Acesse a pasta 04_equilibrio_npt e execute os comandos abaixo.

```
gmx grompp -f ../mdp/npt.mdp -c ../03_equilibrio_nvt/nvt.gro -
t ../03_equilibrio_nvt/nvt.cpt -r ../03_equilibrio_nvt/nvt.gro -
p ../01_preparar_ambiente/topol.top -o npt.tpr -pp npt.top -po npt.mdp

gmx mdrun -deffnm npt &
```

- Não esqueça o símbolo “&”.

Após a finalização do processo podemos, como no item anterior, avaliar a evolução da pressão, e também, a densidade. Para extrair os dados execute os comandos abaixo:

```
echo "Pressure" | gmx energy -f npt.edr -o npt_pressao.xvg
echo "Density" | gmx energy -f npt.edr -o npt_densidade.xvg
```

A Figura 7 apresenta o resultado do equilíbrio com o ensemble NPT.

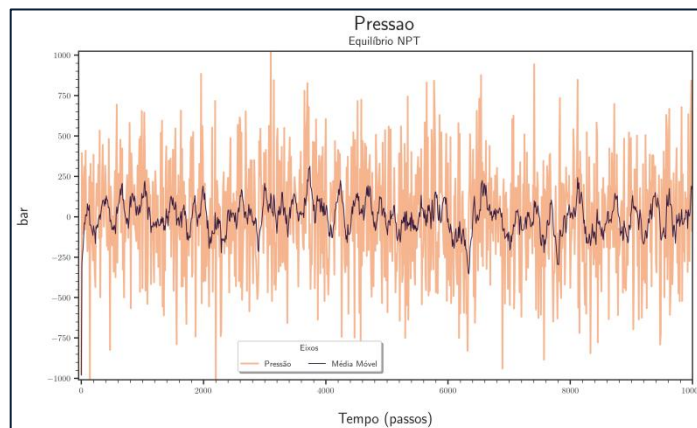


Figura 7 Pressão (NPT)

A figura 8 apresenta o resultado de variação da densidade do sistema.

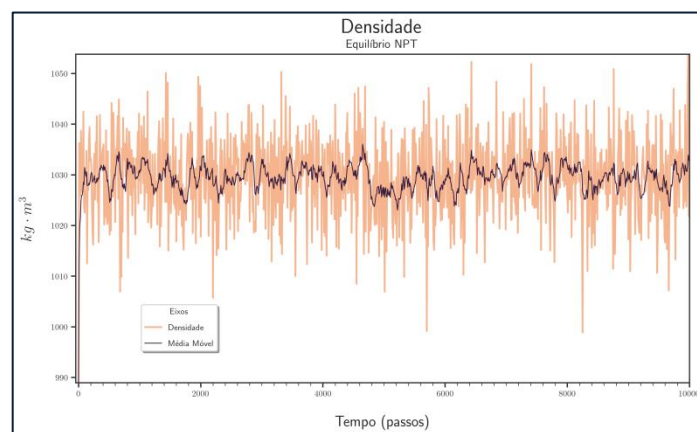


Figura 8 Densidade (NPT)

PRODUÇÃO

Neste momento temos o sistema equilibrado a uma dada temperatura e pressão, portanto, estamos pronto para executar a produção da dinâmica molecular. Para este passo precisamos do arquivo “*md.mdp*”. Portanto, entrar na pasta 05_producao e executar os comandos abaixo:

```
gmx grompp -f ../mdp/md.mdp -c ../04_equilibrio_npt/npt.gro -
t ../04_equilibrio_npt/npt.cpt -p ../01_preparar_ambiente/topol.top -o
md_20.tpr -pp md_20.top -po md_20.mdp
```



```
gmx mdrun -deffnm md_20 &
```

- Não esqueça o símbolo "&".



REFERÊNCIAS

HOLLINGSWORTH, S. A.; DROR, R. O. Molecular Dynamics Simulation for All. *Neuron*, 99, n. 6, p. 1129-1143, 2018.

PETERS, J. H.; DE GROOT, B. L. Ubiquitin Dynamics in Complexes Reveal Molecular Recognition Mechanisms Beyond Induced Fit and Conformational Selection. *PLOS Computational Biology*, 8, n. 10, p. e1002704, 2012.