



DETERMINAÇÃO AUTOMATIZADA DAS COORDENADAS DO GRID BOX (X, Y, Z) PARA ANCORAGEM MOLECULAR USANDO LaBOX

Zayra V. B. da Silva¹; Samuel H. B. da Silva²; Pedro A. P. Pessoa³

¹Graduando do Curso de Licenciatura em Química do Instituto Federal de Educação, Ciência e Tecnologia do Maranhão – IFMA Campus Caxias, henrique.samuel@acad.ifma.edu.br;

²Graduanda do Curso de Licenciatura em Química do Instituto Federal de Educação, Ciência e Tecnologia do Maranhão – IFMA Campus Caxias, zayra.victoria@acad.ifma.edu.br;

³Professor orientador: Doutor em Engenharia e Ciência de Alimentos pelo PPGANEA do UNESP. Curso de Licenciatura em Química do Instituto Federal de Educação, Ciência e Tecnologia do Maranhão – IFMA Campus Caxias, pedro.pessoa@ifma.edu.br.

Palavras-Chave: LaBOX, Grid box, Ramachandran, ancoragem molecular, CHIKV.

Introdução

As ferramentas computacionais aplicadas à Química e à Biologia Molecular têm desempenhado papel central na descoberta racional de fármacos, permitindo prever interações entre proteínas e ligantes em nível atômico (Morgon; Coutinho, 2007; Santos *et al.*, 2021). Entre essas ferramentas, destaca-se o docking molecular, uma técnica amplamente utilizada para prever o modo de ligação entre uma molécula-alvo (geralmente uma proteína) e um ligante (molécula pequena de interesse) (Agu; Afiukwa, Orji *et al.*, 2023). Para que essa etapa seja realizada com precisão, é indispensável a definição correta do grid box, região tridimensional que delimita o sítio ativo da proteína onde o ligante será avaliado.

No contexto de arboviroses emergentes, o vírus Chikungunya apresenta glicoproteínas E1, E2 e E3 do envelope viral, fundamentais para o processo de fusão viral e entrada na célula hospedeira, configurando alvos estratégicos para estudos antivirais (Voss *et al.*, 2010), são alvos estruturais amplamente estudados para o planejamento racional de fármacos antivirais (Haider *et al.*, 2020). Assim, determinação das coordenadas do grid box (X, Y, Z) constitui uma das fases mais críticas do processo de docking, pois influencia diretamente a confiabilidade dos resultados obtidos (Neis, 2019). Uma delimitação incorreta dessa região pode gerar erros na previsão do modo de ligação

e comprometer a qualidade das análises (Esteves; Fernandes, Silva 2019). Nesse contexto, torna-se essencial o uso de ferramentas que auxiliem na identificação precisa do sítio ativo e na definição adequada do grid box.

O software LaBOX (RyanZR, 2024) surge como uma solução moderna e eficiente para essa finalidade. Ele foi desenvolvido para identificar automaticamente o sítio ativo das proteínas e fornecer as coordenadas ideais para configuração no AutoDock Tools (ADT) (Morris, *et al.*, 2009) e no AutoDock Vina (Trott; Olson, 2010), softwares amplamente utilizados em estudos de ancoragem molecular. Essa automatização reduz significativamente a possibilidade de erro humano e otimiza o tempo de preparação das simulações.

Considerando que muitos estudantes e pesquisadores iniciantes encontram dificuldades ao definir manualmente as coordenadas do grid box, o LaBOX apresenta-se como uma ferramenta acessível e didática, capaz de facilitar o processo e promover maior precisão nos resultados. Além disso, seu uso contribui para o aprimoramento das práticas de modelagem molecular em ambientes de ensino e pesquisa (RyanZR, 2024).

Entretanto, antes da automatização das coordenadas, é indispensável validar a qualidade estrutural da proteína. O Ramachandran plot, gerado por ferramentas como o UCSF Chimera Chimera (Pettersen, *et al.*, 2021), permite avaliar os ângulos torsionais ϕ (phi) e ψ (psi) dos resíduos de aminoácidos, assegurando que a estrutura tridimensional está conformacionalmente estável e adequada para estudos de docking.

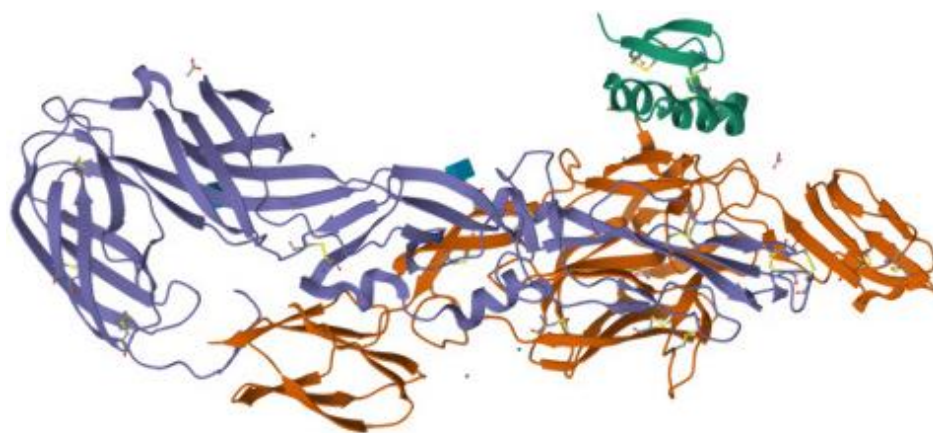
Dessa forma, este trabalho teve como objetivo validar estruturalmente as glicoproteínas E1, E2 e E3 do complexo 3N44 (Voss, *et al.*, 2010), utilizando o UCSF, e, em seguida, determinar automaticamente as coordenadas do grid box (X, Y, Z) com o software LaBOX, demonstrando sua aplicabilidade e eficiência na preparação de sistemas biomoleculares para ancoragem molecular.

As proteínas do complexo 3N44 foram preparadas com o auxílio do UCSF ChimeraX e do AutoDock Tools (ADT), enquanto sua qualidade estrutural foi avaliada no UCSF Chimera por meio da análise dos diagramas de Ramachandran, assegurando conformações adequadas para o docking. Posteriormente, o LaBOX foi empregado para identificar o sítio ativo e calcular automaticamente as coordenadas do grid box, possibilitando a obtenção de parâmetros confiáveis e reprodutíveis para as simulações de ancoragem molecular.

Material e Métodos

Neste estudo, utilizou-se como estrutura de referência o complexo 3N44, disponível no Protein Data Bank (PDB), correspondente ao conjunto das glicoproteínas E3, E2 e E1 do envelope viral CHIKV (Figura 1) (Voss *et al.*, 2010). O arquivo foi obtido no formato .pdb e aberto no software UCSF ChimeraX (Pettersen *et al.*, 2021), onde procedeu-se à separação das três glicoproteínas em arquivos independentes, com o objetivo de possibilitar a análise individual de cada subunidade durante o processo de preparação para o docking molecular.

Fig. 1: Estrutura entrelaçada das glicoproteínas E1 (lilás), E2 (laranja) e E3 (verde-esmeralda) do vírus Chikungunya



Fonte: PDB:3N44

Cada proteína foi posteriormente preparada de forma independente no AutoDock Tools (ADT). Nessa etapa, todas as moléculas de água foram removidas e foram adicionados hidrogênios polares, preservando a geometria nativa do receptor. Em seguida, foram calculadas as cargas parciais de Gasteiger e salvas no formato .pdbqt, garantindo compatibilidade com o AutoDock Vina (Trott; Olson, 2010).

Em seguida proteína foi validada no UCSF Chimera, por meio da geração do Ramachandran plot, sendo possível avaliar os ângulos torsionais ϕ (phi) e ψ (psi). Como parâmetro de qualidade, adotou-se que pelo menos 90% dos resíduos deveriam estar em regiões favorecidas, assegurando a estabilidade conformacional das proteínas (Morgon; Coutinho, 2007; Pettersen *et al.*, 2021).

Com as proteínas devidamente preparadas, cada arquivo foi submetido ao LaBOX (RyanZR, 2024), para a determinação automatizada das coordenadas centrais (X, Y, Z) e das dimensões (size_x, size_y, size_z) do grid box. Essa etapa foi conduzida por meio da

execução do script diretamente no terminal do Windows PowerShell, utilizando a sintaxe básica python LaBOX.py -l ligand.pdbqt.

Fig. 2: interface de texto do LaBOX

```
PS C:\Users\henri\OneDrive\Área de Trabalho\DOCAGEM\Receptores\E1\laBOX> python LaBOX.py -l 3N40.pdb -c
```

Fonte: Autores (2025)

O software processou o arquivo do ligante e retornou as coordenadas correspondentes ao centro geométrico da região de ligação, além do tamanho ideal da caixa de busca em cada dimensão. Esses valores foram posteriormente aplicados na configuração dos arquivos de entrada do AutoDock Vina, assegurando a correta delimitação do sítio ativo. A verificação visual do posicionamento da caixa de busca foi realizada novamente no ADT, de modo a confirmar que o grid box englobava integralmente o sítio ativo de cada proteína analisada, garantindo a confiabilidade das coordenadas geradas pelo LaBOX.

Resultados e Discussão

A análise dos Ramachandran plots (Figs. 3 - 5) revelou que mais de 90% dos resíduos das glicoproteínas E1, E2 e E3 se encontram em regiões favorecidas ou permitidas, demonstrando a **estabilidade conformacional e a qualidade estrutural** das proteínas. Essa etapa foi crucial para garantir que o cálculo do grid box ocorresse sobre modelos tridimensionais corretos e sem distorções estereoquímicas.

Fig. 3: Ramachandran plot of PDB:3N44 da Glicoproteína E1.

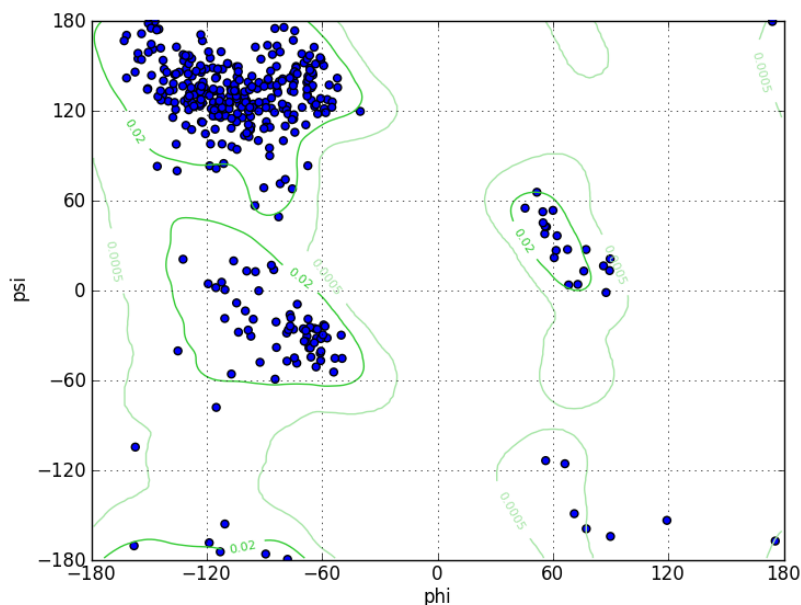


Fig. 4: Ramachandran plot of PDB:3N44 da Glicoproteína E2.

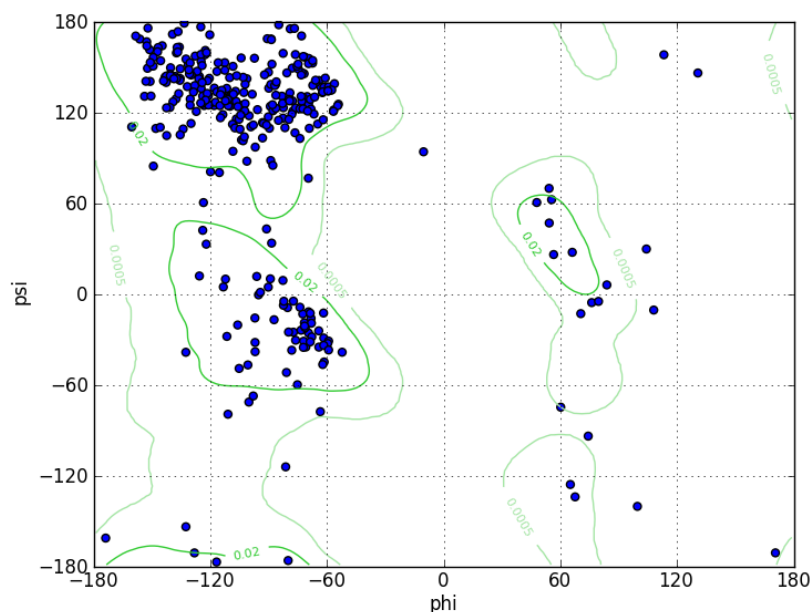
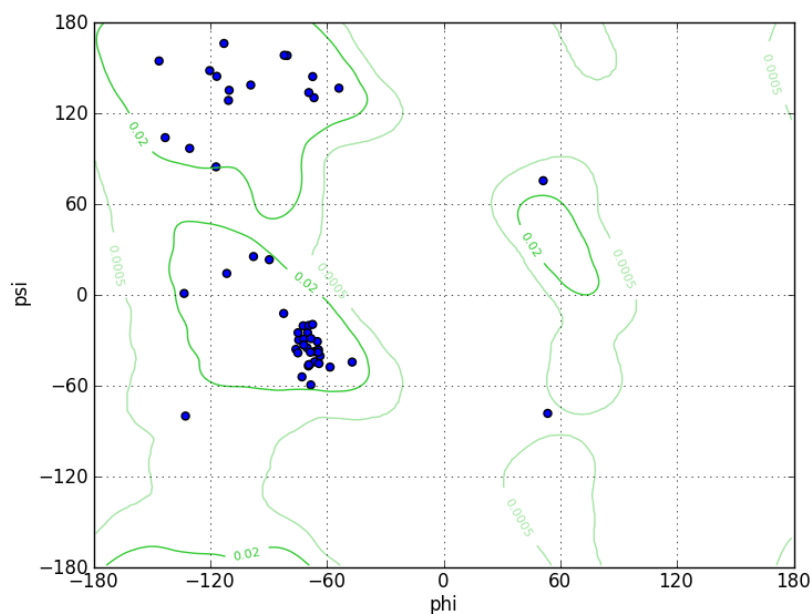


Fig. 5: Ramachandran plot of PDB:3N44 da Glicoproteína E3.



Fonte: Autores (2025)

A execução do LaBOX permitiu a obtenção das coordenadas centrais (X, Y, Z) e dos tamanhos (size_x, size_y, size_z) correspondentes às regiões de ligação das glicoproteínas E3, E2 e E1. O método mostrou-se eficiente e reprodutível, eliminando a subjetividade comum em delimitações manuais de grid box.

Fig. 6: Visão das coordenadas e dimensões calculadas pelo LaBOX

```
PS C:\LaBOX> python LaBOX.py -l E1.pdb -c.txt
LaBOX.py
└o Grid Box Center: X 14.208 Y 9.885 Z 23.711
    Grid Box Size : W 116.994 H 238.884 D 108.98
PS C:\LaBOX> python LaBOX.py -l E2.pdb -c.txt
LaBOX.py
└o Grid Box Center: X 23.34 Y -15.65 Z 18.605
    Grid Box Size : W 156.946 H 189.02 D 102.31
PS C:\LaBOX> python LaBOX.py -l E3.pdb -c.txt
LaBOX.py
└o Grid Box Center: X 55.349 Y -12.741 Z 33.196
    Grid Box Size : W 52.562 H 64.354 D 62.704
```

Fonte: Autores (2025)

A Tabela 1 apresenta um exemplo dos resultados obtidos após a análise com o LaBOX. Os valores foram posteriormente inseridos nos arquivos de configuração do **AutoDock Vina**, garantindo que o volume de busca abrangesse a totalidade do sítio ativo.

Tabela 1: Exemplo de coordenadas e dimensões calculadas pelo LaBOX

Proteína	Center X (Å)	Center Y (Å)	Center Z (Å)	Size X (Å)	Size Y (Å)	Size Z (Å)
E1	14.208	9.885	23.771	116.994	238.884	108.98
E2	23.34	-15.65	16.605	156.946	189.02	102.31
E3	55.349	-12.741	33.196	52.562	64.354	62.704

Fonte: Autores (2025)

Os resultados demonstraram que o uso do LaBOX otimiza a etapa de preparação para *docking* molecular, uma vez que o cálculo automatizado das coordenadas reduz o risco de erro humano e padroniza a área de busca. Além disso, a verificação visual no ADT (Figuras: 7 - 9) confirmou o posicionamento correto da caixa de busca sobre os resíduos do sítio ativo, validando o método.

Essa definição precisa do grid box é fundamental para que o ligante seja avaliado dentro do ambiente catalítico mais favorável, maximizando as chances de formação de complexos proteína-ligante energeticamente estáveis. Assim, a integração entre a validação estrutural (via Chimera) e a automatização das coordenadas (via LaBOX) representa uma abordagem metodológica robusta para estudos de ancoragem molecular.

Fig. 7: Coordenadas do Gridbox na Glicoproteína E1 size 40x40x40

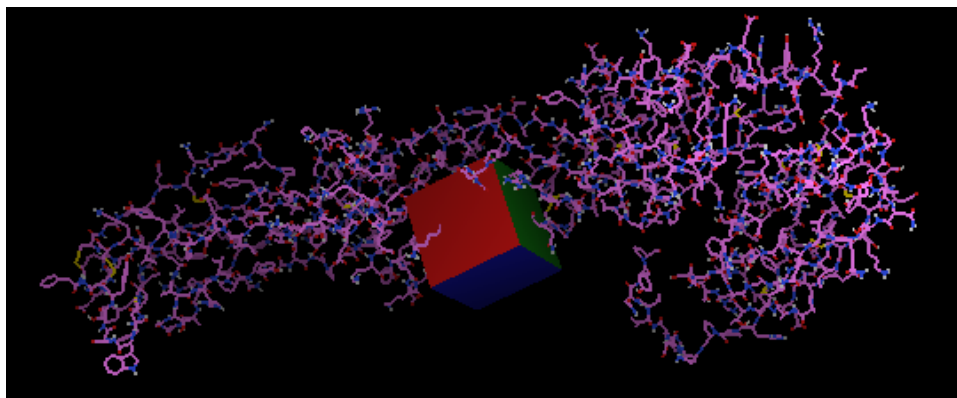


Fig. 8: Coordenadas do Gridbox na Glicoproteína E2 size 40x40x40

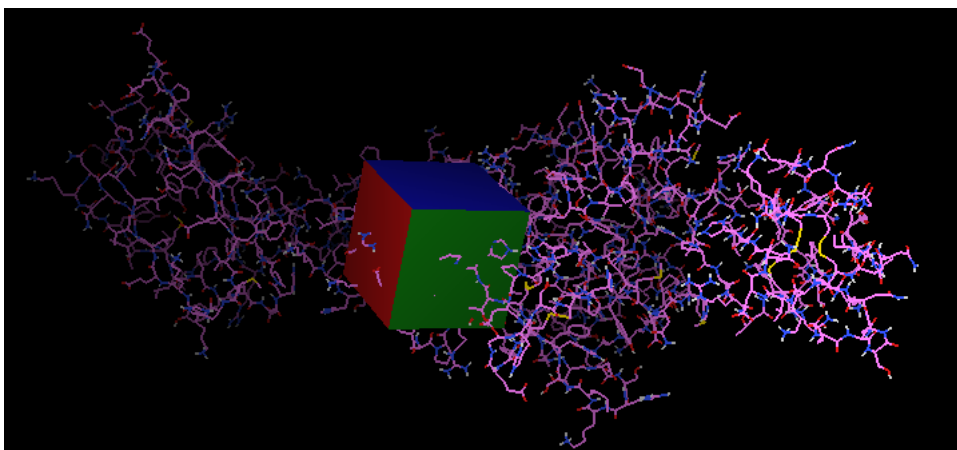
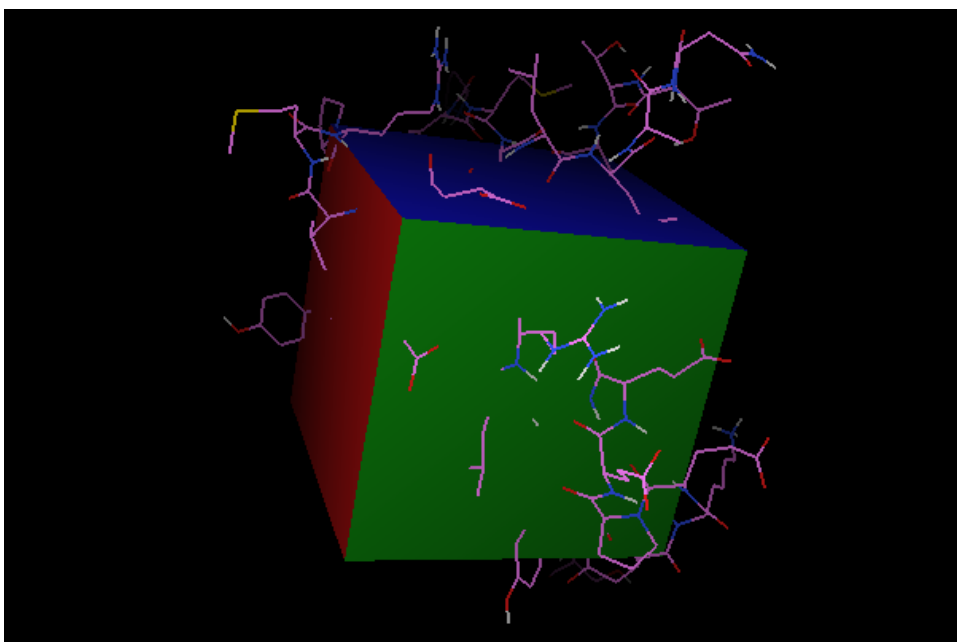


Fig. 9: Coordenadas do Gridbox na Glicoproteína E3 size 40x40x40



Fonte: Autores (2025)

Os resultados demonstraram que o uso do LaBOX otimiza a etapa de preparação para *docking* molecular, uma vez que o cálculo automatizado das coordenadas reduz o risco de erro humano e padroniza a área de busca. Além disso, a verificação visual no ADT confirmou o posicionamento correto da caixa de busca sobre os resíduos do sítio ativo, validando o método.

Conclusões

O uso do LaBOX mostrou-se uma estratégia prática e eficiente para a determinação automática das coordenadas centrais (X, Y, Z) e das dimensões do *grid box* em estudos de *docking* molecular. A metodologia aplicada neste trabalho, a partir da estrutura 3N44 e da análise das glicoproteínas E3, E2 e E1, demonstrou a reprodutibilidade e a confiabilidade da ferramenta. Dessa forma, o LaBOX representa um recurso importante para a padronização e agilização da preparação de sistemas biomoleculares, contribuindo para maior precisão nos resultados de simulações computacionais.

Agradecimentos

Agradeço à Fundação de Amparo à Pesquisa e ao Desenvolvimento Científico e Tecnológico do Maranhão (FAPEMA) pela bolsa concedida e ao IFMA Campus Caxias pela infraestrutura disponibilizada, que contribuíram direta e indiretamente para a realização deste estudo

Referências

AGU, P. C., AFIUKWA, C. A., ORJI, O. U. *et al.* Molecular docking as a tool for the discovery of molecular targets of nutraceuticals in diseases management. *Sci Rep* **13**, 13398 (2023). <https://doi.org/10.1038/s41598-023-40160-2>

BÖCK, F. C.; HELFER, G. A.; COSTA, A. B.; DESSUY, M. B.; FERRÃO, M. F. Rapid Determination of Ethanol in Sugarcane Spirit Using Partial Least Squares Regression Embedded in Smartphone. **Food Analytical Methods**, 11(4), 1951-1957, 2018.

ESTEVES, M. E. A.; FERNANDES, T. V. A.; SILVA, M. L. DA. A importância da preparação de estruturas na metodologia de docking molecular. *BioInfo*, [s.l.], 2019.



HAIDER, N. *et al.* Basic reproduction number of Chikungunya virus transmitted by *Aedes* mosquitoes. *Emerging Infectious Diseases*, v. 26, n. 10, p. 2429-2431, 2020.

MORGON, N. H.; COUTINHO, K. *Métodos de química teórica e modelagem molecular*. São Paulo: Livraria da Física, 2007.

MORRIS, G. M. *et al.* AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of Computational Chemistry*, v. 30, n. 16, p. 2785-2791, 2009.

NEIS, A. Tutorial: Encaixando Molecular com o AutoDock Tools. Medium, 20 out. 2019. Disponível em: <https://medium.com/omixdata/tutorial-docking-molecular-com-o-autodock-tools-cc285496af33>.

PETTERSEN, E. F. *et al.* UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Science*, v. 30, n. 1, p. 70-82, jan. 2021.

RYANZR. LaBOX: A Grid Box Calculation Tool for Molecular Docking. Versão 1.0, 2024.

TROTT, O.; OLSON, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, v. 31, n. 2, p. 455-461, 2010.

VIVAS-REYES, R.; NAVARRO, D.; CORTES, L. E. Exploring Emergent Properties in Chemistry Education: A Philosophical Perspective on the Molecular Revolution. **J. Chem. Educ.** 101(10), 4173–4181, 2024.

VOSS, J. E. *et al.* Glycoprotein organization of Chikungunya virus particles revealed by X-ray crystallography. *Nature*, v. 468, n. 7324, p. 709-712, 2010.