

MODELAGEM MOLECULAR *IN SILICO* DE SPIKE-RBM DE VARIANTES DO SARS-COV-2

Data de submissão: 08/02/2023

Data de aceite: 03/04/2023

Emerson Willian Danzer

Universidade Tecnológica Federal do
Paraná – UTFPR
Toledo – Paraná
<https://lattes.cnpq.br/8960693239491542>

Rafael Bertolini Frigori

Docente de Física na Universidade
Tecnológica Federal do Paraná
Toledo – Paraná
<http://lattes.cnpq.br/5836878566801544>

RESUMO: As funções biológicas das proteínas são determinadas por sua estrutura terciária, ou tridimensional, porém nem sempre dispomos destas estruturas na forma cristalográfica. Durante as últimas décadas a bioinformática vem evoluindo, e, por anos a predição da estrutura tridimensional das proteínas vem sendo estudada, sendo os melhores métodos ranqueados pela CASP (Critical Assessment of Techniques for Protein Structure Prediction), um experimento de âmbito comunitário e mundial para a predição da estrutura da proteína. O software I-TASSER foi criado neste intuito, sendo ele uma pipeline que faz o uso de algoritmos para predição de estruturas

tridimensionais, ranqueado como primeiro colocado pela CASP. Neste trabalho modelamos por homologia, empregando o software I-TASSER, o Receptor Binding Motif (RBM) da proteína Spike (S) de cinco das variantes mais relevantes do vírus SARS-CoV-2. Nosso método foi validado através da comparação com estruturas cristalográficas (PDB). Posteriormente, utilizamos a ferramenta PBEQ-SOLVER, parte integrante do servidor CHARMM-GUI para modelar a distribuição espacial de cargas e resolver a equação de Poisson-Boltzmann, obtendo assim os potenciais eletrostáticos e energias livres dos modelos RBMs anteriormente modelados. Com os resultados confirmamos a acurácia dos métodos utilizados e determinamos que o aumento da infectividade do SARS-CoV-2 correlaciona-se à minimização na energia livre do RBM de sua proteína S.

PALAVRAS-CHAVE: Bioinformática; proteína; RBM; Poisson-Boltzmann; Sars-CoV-2.

IN SILICO MOLECULAR MODELING OF SPIKE-RBM OF SARS-COV-2 VARIANTS

ABSTRACT: The biological functions of proteins are determined by their tertiary or

three-dimensional structure, but these structures are not always available in crystallographic form. During the last decades, bioinformatics has been evolving, and, for years, the prediction of the three-dimensional structure of proteins has been studied, with the best methods ranked by CASP (Critical Assessment of Techniques for Protein Structure Prediction), a community and worldwide experiment to the prediction of protein structure. The I-TASSER software was created for this purpose, which is a pipeline that uses algorithms to predict three-dimensional structures, ranked first by CASP. In this work we model by homology, using the I-TASSER software, the Receptor Binding Motif (RBM) of the Spike (S) protein of five of the most relevant variants of the SARS-CoV-2 virus. Our method was validated through comparison with crystallographic structures (PDB). Subsequently, we used the PBEQ-SOLVER tool, an integral part of the CHARMM-GUI server, to model the spatial distribution of charges and solve the Poisson-Boltzmann equation, thus obtaining the electrostatic potentials and free energies of the previously modeled RBMs. With the results, we confirmed the accuracy of the methods used and determined that the increase in SARS-CoV-2 infectivity correlates with the minimization of the RBM free energy of its S protein.

KEYWORDS: Bioinformatic; modeling; protein; Poisson-Boltzmann; Sars-CoV-2.

1 | INTRODUÇÃO

Segundo Linus Pauling (1954), “Podemos, creio eu, antecipar que o químico do futuro interessado na estrutura de proteínas, ácidos nucleicos, polissacarídeos e outras substâncias complexas com alto peso molecular passará a contar com uma nova química estrutural, envolvendo relações geométricas precisas entre os átomos nas moléculas e a aplicação rigorosa dos novos princípios estruturais, e que grandes progressos serão feitos, através desta técnica, no ataque, por métodos químicos, aos problemas da biologia e da medicina.”. Em seu discurso no prêmio Nobel de 1954, ressaltou a importância do estudo de proteínas através de uma abordagem estrutural. Atualmente, quase 70 anos depois, o avanço nesta área obteve avanços significativos devido ao advento da informática. Neste capítulo abordaremos um desses desenvolvimentos no que tange à modelagem e análise eletrostática de um segmento de uma particular proteína (Spike)

Proteínas são moléculas orgânicas complexas, compostas por aminoácidos que desempenham muitas funções importantes no corpo humano, incluindo a construção de tecidos, transporte de moléculas, resposta imune e contração muscular. Os aminoácidos são as unidades básicas das proteínas e formam ligações peptídicas. A estrutura tridimensional de uma proteína é determinada pela sequência de aminoácidos em sua cadeia polipeptídica e pela interação desses aminoácidos com o ambiente. Esta estrutura, quando no estado de mínima energia livre, chama-se estrutura nativa e é crucial para o correto funcionamento biológico da proteína (HALLING, 1995).

O processo físico que conforma a proteína a partir de uma configuração desordenada em sua estrutura nativa é designado por enovelamento proteico (ou “folding”). Vale ressaltar que alguns *foldings* ocorrem em etapas, desenvolvendo-se em locais diferentes da célula,

o que o caracteriza como um processo complexo e que demanda a ultrapassagem de diversos níveis de energia (metaestáveis), caracterizando-se assim segundo a mecânica-estatística microcanônica como uma transição de fase conformacional.

Em particular, uma nova proteína denominada “Spike” tornou-se globalmente relevante a partir de 2019, dado o surgimento da doença respiratória pandêmica denominada COVID-19 (coronavirus disease 2019). Segundo BENVENUTO et al., 2020, seu agente etiológico é o SARS-CoV-2, um vírus de RNA de cadeia simples positiva. O SARS-CoV-2 possui um envelope viral composto por diversas proteínas, sendo classificado como uma estirpe de coronavírus relacionado com a síndrome respiratória grave (SARS-CoV). Contudo, diferente do SARS-CoV, o SARS-CoV-2 possui um sítio de clivagem de furina próximo ao domínio ligante do receptor (RBD) da proteína Spike (S), tendo, portanto, alta afinidade com o receptor humano “enzima conversora da angiotensina 2 (ACE2)”. Segundo LAN et al., 2020, a diferença crucial na afinidade entre a proteína S do SARS-CoV-2 em relação à do SARS-CoV situa-se em uma pequena subunidade do RBD, conhecida como motivo ligante do receptor (RBM).

A predição de estruturas tridimensionais de proteínas vem sendo estudadas por anos. As primeiras tentativas foram realizadas por modelagem estrutural, porém eram pouco eficazes na ausência de estruturas homólogas depositadas no Protein Data Bank. Nesse ambiente, foi criado o I-TASSER, uma pipeline que faz o uso de algoritmos para predição de estruturas tridimensionais, sendo em 2020 ranqueado como primeiro colocado pela CASP (Critical Assessment of Techniques for Protein Structure Prediction) (KRYSHTAFOVYCH et al., 2019).

Dentre os softwares que ajudam a entender as interações biológicas está o CHARMM, um software utilizado para estudos de dinâmicas macromolecular com ferramentas de análise e manipulação de coordenada atômica e trajetória dinâmica. O CHARMM-GUI é uma versão deste software com uma interface gráfica online para ser amigável ao usuário, padronizar e facilitar o uso das técnicas de simulação usadas em CHARMM (JO et al., 2008).

O PBEQ-Solver é parte integrante do CHARMM-GUI, ele fornece uma interface gráfica de usuário baseada na web para ler estruturas biomoleculares, resolver as equações de Poisson-Boltzmann (PB) e visualizar interativamente o potencial eletrostático. A equação de Poisson-Boltzmann é um modelo matemático usado para descrever a distribuição do potencial eletrostático em uma solução eletrolítica, levando em conta as interações entre íons e moléculas de solvente. Tal equação relaciona a energia potencial dos íons com sua concentração e o potencial elétrico na solução, e ajuda a prever o comportamento de espécies carregadas em solução, como migração, distribuição e ligação. A equação é amplamente utilizada nas áreas de biologia computacional e engenharia química para estudar o comportamento de íons em diversos processos biológicos e industriais (HANS-JÜRGEN BUTT, KARLHEINZ GRAF, KAPPL, 2011).

O PBEQ-Solver calcula o potencial eletrostático e energia livre de solvatação, a energia de interação eletrostática proteína-proteína (DNA ou RNA) e o pKa de um resíduo titulável selecionado (JO et al., 2008). Ele foi utilizado neste estudo para verificar a energia livre das moléculas geradas pelo servidor I-TASSER.

2 | MÉTODO

Para realizar a modelagem deste estudo, foram selecionadas as 5 maiores VoCs (Variants of Concern) em circulação até o momento de início do estudo segundo a Organização Mundial da Saúde. Sendo elas o vírus original WT conhecida como Wild Type (YI et al., 2020), a variante britânica B.1.1.7 (Alpha), as variantes P.1 e B.1.351 conhecidas como Gamma e Beta respectivamente e que possuem a RBM idêntica (HAUSER et al., 2021), a sequência B.1.617.2 conhecida como Delta (PLANAS et al., 2021), e por fim e mais recente B.1.1.529 conhecida como Omicron (DEJNIRATTISAI et al., 2021). A figura 1 mostra as sequências de aminoácidos presentes na RBM, e em destaque, os aminoácidos substituídos nas variantes em relação à sequência WT.

Sequência WT	SNNLDSKVGGNYNLYRLFRKSNLKPFFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPY
Sequência B.1.1.7	SNNLDSKVGGNYNLYRLFRKSNLKPFFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPT Y GVGYQPY
Sequência P.1 e B.1.351	SNNLDSKVGGNYNLYRLFRKSNLKPFFERDISTEIYQAGSTPCNGV K GFNCYFPLQSYGFQPT Y GVGYQPY
Sequência Delta	SNNLDSKVGGNYN R YRLFRKSNLKPFFERDISTEIYQAGS K PCNGVEGFNCYFPLQSYGFQPTNGVGYQPY
Sequência Ômicron	S N KLDSK V SGN N YLYRLFRKSNLKPFFERDISTEIYQAG N KPCNGV A GFNCYFPL R S S FRPT Y GV G H Q P Y

Figura 1 - Sequências RBM (Receptor Biding Motif)

Fonte: Elaborado pelo autor (2022).

2.1 I-TASSER (Iterative Threading ASSEmblY Refinement)

A predição de estruturas tridimensionais de proteínas vem sendo estudadas por anos. As primeiras tentativas foram realizadas por modelagem baseada em modelo, porém ela não funciona para proteínas que não possuem estruturas iguais/parecidas depositadas no Protein Data Bank (ŠALI & BLUNDELL, 1993; BERMAN & HENRICK & NAKAMURA, 2003). Algumas tentativas de predição sem estruturas homologas foi empregada, uma delas é o uso dos contatos resíduo-resíduo, porém não tiveram um sucesso muito grande (KINCH et al., 2011).

O I-TASSER é uma pipeline que faz o uso de algoritmos para predição de estruturas tridimensionais, sendo ranqueado em 2020 como o melhor segundo a CASP. O pipeline do I-TASSER Suite consiste em quatro etapas gerais:

1. Identificação de modelo de encadeamento: onde ocorre um procedimento para identificar proteínas que servirão de modelo em um banco de dados que tenham sua estrutura resolvida semelhante ou um enovelamento estrutural semelhante à

sequência da proteína de consulta. É usado um método de meta-threading chamado LOMETS, composto por 8 programas que reconhecem as dobras (PPAS, Env-PPAS, wPPAS, dPPAS, dPPAS2, wdPPAS, MUSTER and wMUSTER).

2. Montagem da estrutura: nesta etapa, alguns fragmentos das proteínas modelo são postos continuamente em alinhamento para criar uma conformação estrutural base para a proteína alvo. Nesta etapa é realizado uma série de simulações Monte Carlo para adequar os fragmentos para a estrutura base.

3. Seleção e refinamento do modelo: Nesta etapa é realizada a montagem da molécula alvo sobre a estrutura base montada, a partir disso é realizado a seleção das conformações que mais se assemelham com um alinhamento contínuo e para remover conformações com impedimento. As conformações de energia livre mais baixas são identificadas por agrupamento e então é realizada mais uma rodada de simulação de montagem para que sejam removidos os conflitos estéricos e para refinar a topologia da estrutura.

4. Anotação de função baseada na estrutura: a função da proteína alvo é inferida combinando estruturalmente os modelos 3D previstos com as proteínas de estrutura e função conhecidas no PDB.

Para estimar a qualidade das previsões da estrutura, o I-TASSER conta com uma “pontuação de confiança” chamada C-score, que varia de -5 a 2, ela é definida a partir da qualidade dos alinhamentos da estrutura simulada em relação às estruturas bases usadas para o refinamento da montagem estrutural. Para o estudo foi selecionado o modelo com maior C-score de cada OUTPUT, os respectivos valores estão representados na Tabela 1.

Sequência	C-score
Sequência WT	0.08
Sequência B.1.1.7 (Britânica)	0.13
Sequência P.1 (Brasileira) e B.1.351 (Sul Africana)	0.10
Sequência Delta (B.1.617.2)	0.16
Sequência Omicron (B.1.1.529)	0.12

Tabela 1 - C-score

Fonte: Elaborado pelo autor (2022).

Para utilizar o I-TASSER, deve ser realizado o procedimento no endereço web “<https://zhanggroup.org/I-TASSER/>” com prévio cadastro na plataforma. O INPUT é realizado escrevendo a sequência a ser modelada ou realizando o envio de um arquivo *.fasta* assim como mostrado na Figura 2.

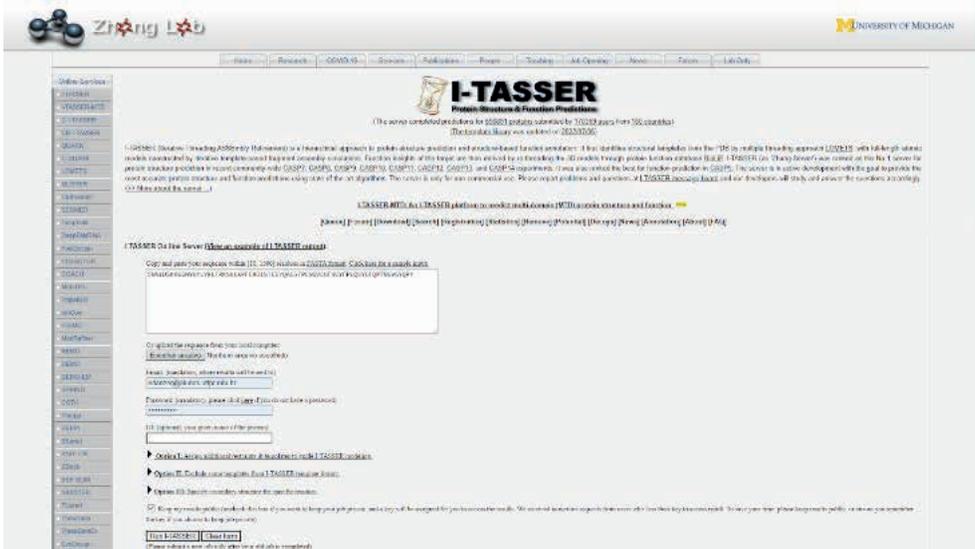


Figura 2 - Interface gráfica do I-TASSER

Fonte: Captura de tela feito pelo autor do endereço web: <https://zhanggroup.org/I-TASSER/>.

Utilizando o I-TASSER, uma a uma as seqüências foram submetidas para realização da modelagem. Para realizar a validação do estudo pelo I-TASSER foi extraído a seqüência RBM correspondente de moléculas RBD do banco de dados RCSB PDB para comparar com a molécula gerada pelo servidor e foram sobrepostas usando o software Chimera-X como demonstrado na Tabela 2.

PDB em comparação com a variante predita	7RKV (Jette et. al.) em marrom.	7EDJ (Yang et. al.) em marrom.	7LYM (Gobeil et. al.) em marrom.	7SBL (Zhang et. al.) em marrom.	7WPE (Yin et. al.) em marrom.
	Seqüência RBM WT em azul.	Seqüência RBM B.1.1.7 em azul.	Seqüência RBM P.1 e B.1.351 em azul.	Seqüência RBM B.1.617.2 em azul.	Seqüência RBM B.1.1.529 em azul.
Visualização 3D					
RMSD	RMSD entre 68 pares de átomos podados é 1,217 Angstroms; (em todos os 71 pares: 1,285)	RMSD entre 69 pares de átomos podados é 0.893 Angstroms; (em todos os 71 pares: 1,050)	RMSD entre 68 pares de átomos podados é 0,918 Angstroms; (em todos os 71 pares: 1,063)	RMSD entre 70 pares de átomos podados é 0,804 Angstroms; (em todos os 71 pares: 0,876)	RMSD entre 60 pares de átomos podados é 0,915 Angstroms; (em todos os 71 pares: 1.500)

Tabela 2 - PDBs sobrepostos

Fonte: Elaborado pelo autor (2022).

O RMSD (root-mean-square deviation) na bioinformática indica a distância média entre os átomos de proteínas sobrepostas e é comumente utilizado para medir a similaridade entre duas proteínas. Seu valor é determinado pela Eq. 1.

$$\text{RMSD} = \sqrt{\frac{1}{N} \sum_{i=1}^N \delta_i^2} \quad (1)$$

Onde δ indica a distância de um átomo i e uma estrutura referência. Para cada átomo alinhado das proteínas sobrepostas, foi calculado o quadrado da distância entre cada um deles, a soma de todos os quadrados, a divisão da soma pelo número de átomos e por fim, calcula-se a raiz quadrada do valor. Resultando assim em um valor com unidade Angstrom, que equivale a 10^{-10}m , unidade comum na biologia estrutural.

2.2 Equação de Poisson-Boltzmann

No eletromagnetismo clássico a Equação de Poisson (Eq. 2) é uma equação diferencial parcial (EDP) que rege o potencial eletrostático (V) e descreve seu valor em todos os pontos do espaço como uma função da densidade das cargas elétricas (ρ).

$$\nabla^2 V = \frac{-\rho_v}{\epsilon} \quad (2)$$

Essa equação assume (ρ) estacionário, logo, é incapaz de lidar com a movimentação (flutuações de posição) de natureza térmica de íons em uma solução biológica, cuja distribuição de cargas é de natureza estatística. Para lograr essa modelagem, foi proposta a equação de Poisson-Boltzmann (PBEQ), em que a densidade de cargas, torna-se uma função da concentração iônica (c_i^0), do trabalho eletrostático para mover os íons no meio (W), e da temperatura de equilíbrio do meio circundante (T) apresentado na Eq. 3:

$$c_i = c_i^0 * e^{\frac{-W_i}{k_B T}} \quad (3)$$

Desta forma, a densidade de cargas na equação de Poisson é modificada segundo as regras da Mecânica Estatística, considerando a distribuição estatística de Boltzmann, e assim, a densidade de cargas elétricas (ρ) da Eq. 2 torna-se a Eq. 4:

$$\rho_e = e(c^+ - c^-) = c_0 e * \left[e^{\frac{-eV(x,y,z)}{k_B T}} - e^{\frac{eV(x,y,z)}{k_B T}} \right] \quad (4)$$

Assim, a teoria de Poisson-Boltzmann é uma boa aproximação na maioria das condições fisiológicas, especialmente para íons monovalentes e para potenciais de superfície que não são muito grandes. Próximo à superfície carregada, o tamanho finito dos grupos iônicos da superfície e dos contra-íons pode levar a desvios dos resultados de Poisson-Boltzmann.

No caso de estudo em tela, a resolução numérica da PBEQ para os RBMs permite determinar suas energias livres de solvatação. Segundo critérios de estabilidade

termodinâmicos, espera-se que quanto menor forem tais energias livres, maiores serão as interações dos RBMs com a proteína ACE2. Dessa forma as variantes mais infecciosas, e, portanto, biologicamente mais adaptadas aos humanos, teriam menor energia livre. Para efetuar esse cálculo, empregamos o módulo PBEQ-Solver do servidor CHARMM-GUI, uma versão gráfica online do CHARMM, que é um pacote de softwares desenvolvido em Harvard para simulações de dinâmicas macromoleculares. Por sua vez o PBEQ-Solver fornece uma interface gráfica para ler estruturas biomoleculares, resolver a equação de Poisson-Boltzmann (PB) e visualizar interativamente o potencial eletrostático. Para utilizá-lo submetemos os modelos RBM (Tabela 1) no endereço web “<https://charmm-gui.org/?doc=input/pbeqsolver>” com prévio cadastro na plataforma.

3 | RESULTADOS

Comparando as sequências geradas pelo I-Tasser, verificou-se que a sequência B.1.1.7 possui 98,59% de similaridade com a sequência WT, a sequência P.1 (Brasileira), B.1.351 (Sul Africana) e a sequência Delta (B.1.617.2) possuem 97,18% de similaridade com a sequência WT, sendo a sequência Omicron (B.1.1.529) a mais diferente, com 85,92% de similaridade com a sequência WT.

A partir do cálculo da equação de Poisson-Boltzmann, foi possível identificar a energia livre de solvatação das sequências modeladas, na tabela 3, apresentando os valores adquiridos da média de energia de solvatação livre adquirida pelo PBEQ-Solver.

Sequência	Energia Livre de Solvatação
Sequência WT	-581.5225 kcal/mol
Sequência B.1.1.7 (Britânica)	-568.1241 kcal/mol
Sequência P.1 (Brasileira) e B.1.351 (Sul Africana)	-625.4217 kcal/mol
Sequência Delta (B.1.617.2)	-685.0464 kcal/mol
Sequência Omicron (B.1.1.529)	-973.2539 kcal/mol

Tabela 3 - Resultados do PBEQ-SOLVER

Tabela 4 - Fonte: Elaborado pelo autor (2022).

4 | CONCLUSÃO

Durante a pandemia de Covid-19 houve uma urgência de estudos relacionados ao Sars-Cov-2, vírus responsável pela doença. Uma vez que podemos não dispor de modelos isolados de moléculas tridimensionais para ser estudados, analisamos a viabilidade e análise da modelagem das sequências RBM de cinco variantes preocupantes, estrutura está responsável diretamente pela ligação do vírus à ACE2 humana. Há uma diferença significativa na variante Omicron, com 85,92% de similaridade em comparação à Wild-

Type, enquanto as outras superam o patamar de 90% de similaridade. Ao modelar as sequências de aminoácidos com o I-TASSER, o C-score resultante dos modelos melhores ranqueados não foi menor que 0. Somado a isso, o RMSD da sobreposição dos fragmentos RBM cristalográficos com os modelos anteriormente selecionados foram próximos de 1 Ångström, o que indica uma boa paridade na estrutura das moléculas sobrepostas. Com a verificação da qualidade da modelagem realizada, podemos efetuar o cálculo da equação de Poisson-Boltzmann, que confirmou a hipótese de que as variantes mais infecciosas teriam as menores energias livres. Por mais que haja outros fatores que interfiram na infectividade de uma variante, haver uma menor energia livre no segmento que realiza a ligação do vírus com o receptor humano, indica que a variante está mais adaptada a sua função, com grande destaque para o resultado gerado pela variante Omicron. Estudos com a modelagem destes importantes fragmentos podem levar a um foco preciso em pesquisas para medicamentos e vacinas, ajudando assim a reduzir custos e a aumentar a assertividade dos mesmos.

REFERÊNCIAS

BENVENUTO, D. et al. **The 2019-new coronavirus epidemic: Evidence for virus evolution.** Journal of Medical Virology, v. 92, n. 4, 7 fev. 2020.

BERMAN, H.; HENRICK, K.; NAKAMURA, H. **Announcing the worldwide Protein Data Bank.** Nature Structural & Molecular Biology, v. 10, n. 12, p. 980–980, dez. 2003.

DEJNIRATTISAI, W. et al. **Omicron-B.1.1.529 leads to widespread escape from neutralizing antibody responses.** 7 dez. 2021.

GOBEIL, S. M.-C. et al. **Effect of natural mutations of SARS-CoV-2 on spike structure, conformation, and antigenicity.** Science, v. 373, n. 6555, 6 ago. 2021.

HALLING, P. J. **Proteins: Structures and molecular properties (2nd edition).** by Thomas E. Creighton, W. H. Freeman, New York, 1992, xiii + 512 pp, price £22.95. ISBN 0-7167-7030-X. Journal of Chemical Technology AND Biotechnology, v. 62, n. 1, p. 105–105, jan. 1995.

HANS-JÜRGEN BUTT; KARLHEINZ GRAF; KAPPL, M. **Physics and chemistry of interfaces.** Weinheim: Wiley-Vch, 2011.

HAUSER, B. M. et al. **Rationally designed immunogens enable immune focusing to the SARS-CoV-2 receptor binding motif.** 16 mar. 2021.

JETTE, C. A. et al. **Broad cross-reactivity across sarbecoviruses exhibited by a subset of COVID-19 donor-derived neutralizing antibodies.** Cell Reports, v. 36, n. 13, p. 109760, set. 2021.

JO, S. et al. **CHARMM-GUI: A web-based graphical user interface for CHARMM.** Journal of Computational Chemistry, v. 29, n. 11, p. 1859–1865, 19 mar. 2008.

JO, S. et al. **PBEQ-Solver for online visualization of electrostatic potential of biomolecules.** Nucleic Acids Research, v. 36, n. suppl_2, p. W270–W275, 28 maio 2008.

KINCH, L. N. et al. **CASP9 target classification.** Proteins, v. 79 Suppl 10, n. Suppl 10, p. 21–36, 2011.

KRYSHTAFOVYCH, A. et al. **Critical assessment of methods of protein structure prediction (CASP)—Round XIII.** Proteins: Structure, Function, and Bioinformatics, v. 87, n. 12, p. 1011–1020, 23 out. 2019.

LAN, J. et al. **Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor.** Nature, v. 581, 30 mar. 2020.

PAULING, L. **Nobel Lecture.** NobelPrize.org. Nobel Prize Outreach AB 2023. Disponível em: <<https://www.nobelprize.org/prizes/chemistry/1954/pauling/lecture/>>. Acessado em 7 de fev. de 2023.

PLANAS, D. et al. **Reduced sensitivity of SARS-CoV-2 variant Delta to antibody neutralization.** Nature, p. 1–7, 8 jul. 2021.

ŠALI, A.; BLUNDELL, T. L. **Comparative Protein Modelling by Satisfaction of Spatial Restraints.** Journal of Molecular Biology, v. 234, n. 3, p. 779–815, dez. 1993.

WORLD HEALTH ORGANIZATION. **Tracking SARS-CoV-2 variants.** Disponível em: <<https://www.who.int/activities/tracking-SARS-CoV-2-variants>>.

YANG, T.-J. et al. **Effect of SARS-CoV-2 B.1.1.7 mutations on spike protein structure and function.** Nature Structural & Molecular Biology, v. 28, n. 9, p. 731–739, 12 ago. 2021.

YI, C. et al. **Key residues of the receptor binding motif in the spike protein of SARS-CoV-2 that interact with ACE2 and neutralizing antibodies.** Cellular & Molecular Immunology, v. 17, n. 6, p. 621–630, 15 maio 2020.

YIN, W. et al. **Structures of the Omicron spike trimer with ACE2 and an anti-Omicron antibody.** Science, v. 375, n. 6584, p. 1048–1053, 4 mar. 2022.

ZHANG, J. et al. **Membrane fusion and immune evasion by the spike protein of SARS-CoV-2 Delta variant.** Science, 26 out. 2021.