



Universidade Federal do Amazonas
Pró-Reitoria de Pesquisa e Pós-Graduação
Departamento de Apoio à Pesquisa
Programa Institucional de Bolsas de Iniciação Científica

DINÂMICA MOLECULAR DA VARIANTE P.1 DO
SARS-COV-2 E ESTUDO DA COORDENADA DE REAÇÃO
PARA DEAMINAÇÃO DA CITOSINA
(PIB-E/0035/2021)

BOLSISTA: MICAEL DAVI LIMA DE OLIVEIRA (CNPQ)

Manaus - AM
2022

MICAEL DAVI LIMA DE OLIVEIRA

DINÂMICA MOLECULAR DA VARIANTE P.1 DO
SARS-COV-2 E ESTUDO DA COORDENADA DE REAÇÃO
PARA DEAMINAÇÃO DA CITOSINA

Relatório final de pesquisa PIB-E/0035/2021
pertencente ao campo de Ciências Exatas
e da Terra, como prestação de contas ao
Programa de Bolsas de Iniciação Científica
(PIBIC) na edição 2021-2022.

Área de concentração: Química teórica

Opção: Biofísica Computacional

Orientador: Prof. Dr. Kelson Mota Teixeira de Oliveira

Colaborador: Prof. Dr. Jonathas Nunes da Silva

Manaus - AM
2022

Dedicatória

Dedicado a todos que foram acometidos pela COVID-19 principalmente durante a 2^a onda de casos da COVID-19 em Manaus. Também dedicamos a todos os pesquisadores e pessoas que se uniram para ajudar uns aos outros na luta contra o vírus SARS-CoV-2. Queria destacar que tudo o que fazemos foi para ajudar a sociedade, por mais que pareça que todos os nossos esforços pareçam ser insignificantes. Ainda que os resultados aqui obtidos não tenham impedido mais mortes pela COVID-19, precisamos continuar persistindo nesta pesquisa. E sendo assim, o principal sentido de todos os nossos esforços se resume a ajudar as pessoas de alguma forma.

Observações

Apesar do trabalho ter como título o enfoque na variante P.1, o tema foi alterado para a mais recente variante Ômicron, mas também comparamos estas duas variantes entre si. Outra observação é de que não foi possível realizar a tempo o estudo da deaminação e metilação da citosina por conta das minhas dificuldades em estimar o estados de transição e coordenada de reação (IRC) pelos softwares Gaussian 09W e GaussView 6. Até porque seria algo que demandaria muito tempo pela escassez de artigos que tenham abordado o mesmo problema. Além disso, em nenhum livro ou artigo lido ao longo deste PIBIC descreveu detalhadamente o mecanismo de deaminação e metilação da citosina, até porque existem várias reações intermediárias ainda desconhecidas pela ciência. Contudo, mesmo com o término desta pesquisa continuaremos a estudar o mecanismo destas reações, isto porque são imprescindíveis para melhor entendermos como ocorrem as alterações epigenéticas e até mesmo achar uma luz para muitas doenças caracterizadas por hipermetilação, como o câncer. Por outro lado, conseguimos utilizar a química-quântica para tentar explicar porque a mutação K417T esteve presente em várias variantes de preocupação da COVID-19. Embora no relatório final de PIBIC 2020-2021 já tivessem os resultados de Monte Carlo para as mutações que maximizam a afinidade anticorpo-antígeno, neste apresentamos novas mutações agora quando o número de alterações de aminoácidos simultâneas é 16, ao invés de 3. Isto portanto aumentou a complexidade computacional ao mesmo tempo que fornece resultados ainda mais precisos em comparação ao relatório passado.

Publicações resultantes desta pesquisa

- Infelizmente ainda não conseguimos publicar os resultados em um periódico indexado.
- Submetemos os resultados dessa pesquisa ao periódico B2 chamado de "Journal of Biomolecular Structure Dynamics", pertencente à publicadora Taylor & Francis. Os revisores já emitiram o parecer, ainda que não tenha sido rejeitado completamente foi sugerido revisões no artigo possibilitando a ressubmissão.
- Tivemos um resumo aceito no congresso nacional de química medicinal chamado de "Braz-MedChem 2022" a ser realizado em Minas Gerais. Sendo que fomos aceitos para a apresentação oral neste congresso. O título do trabalho foi "Maximization of SARS-CoV-2 neutralization by antibodies through Monte Carlo optimization".
- Mediante a importante ajuda da Prof^a Dra. Rosiane de Freitas e Msc. Clarice Santos do Instituto de Computação (IComp) da UFAM conseguimos publicar um trabalho no CLEI Electronic Journal de qualis B5 onde participei como co-autor [1];
- Tendo a Prof^a Dra. Isabelle Bezerra Cordeiro como principal coordenadora, foi submetido um projeto de pesquisa ao Edital N^o 014/2022 - PAPD/FAPEAM. O objetivo será realizar uma pesquisa ao mesmo tempo que divulgaremos a ciência à população em geral, tornando acessível e compreensível a importância de algumas técnicas de bioinformática, como docking e dinâmica molecular.

Resumo

OLIVEIRA, M. D. L. Dinâmica molecular da variante P.1 do SARS-CoV-2 e estudo da coordenada de reação para deaminação da citosina . 2022. 42 p. Iniciação Científica - Laboratório de Química Teórica e Computacional (LQTC), Universidade Federal do Amazonas, Manaus, 2022.

Várias pesquisas apontam que a última variante de preocupação Ômicron (B.1.1.529) apresenta um aumento de transmissibilidade em relação às variantes anteriores do SARS-CoV-2, porém é menos virulenta/agressiva. Desta forma, o uso de métodos computacionais como dinâmica molecular será essencial para entender suas consequências à nível molecular na interação da proteína Spike com anticorpos neutralizantes e com o receptor ACE2. Dentre os resultados obtidos a partir da dinâmica molecular em $100ns$, tivemos a redução da quantidade de ligações de hidrogênio no complexo anticorpo-antígeno. Além disso, esta nova variante resultou no aparecimento de um novo sítio proteolítico, que pode facilitar a interação ACE2-RBD e, portanto, refletir fenotipicamente em maior transmissibilidade em relação às variantes anteriores. A distribuição de carga elétrica mostrou-se mais positiva em comparação à estrutura sem mutações. Ao longo da pesquisa, percebemos que dentre os anticorpos simulados, o tipo IgA em sua forma dimérica superou o potencial de neutralização dos demais anticorpos como IgG e IgM. Os atuais diagnósticos moleculares para COVID-19 do tipo antígeno geralmente adotam a proteína Spike (S), porém os resultados aqui obtidos mostram um potencial de neutralização ainda maior por anticorpos da proteína de nucleocapsídeo (N). No decorrer desta pesquisa percebeu-se que a mutação K417T é energeticamente mais favorável e talvez por isso esteve presente na variante P.1 que infelizmente afetou a cidade de Manaus durante a grave 2^a onda. Por fim, propomos mediante o método de otimização de Monte Carlo quais substituições de aminoácidos poderiam maximizar a interação anticorpo-antígeno, o que poderá aumentar a imunogenicidade mesmo diante de novas variantes ou talvez um novo coronavírus. Desta forma, encontramos 16 (dezesseis) mutações no antígeno referente ao anticorpo IgG B38 resultando em uma afinidade maximizada de $-61,656kcal/mol$. Quanto ao anticorpo IgA-Fc1, encontramos outras mutações que maximizam o valor da afinidade ao antígeno em $-82,343kcal/mol$. Ainda que os resultados aqui obtidos sejam promissores seja para melhor entendermos à nível molecular a infecção pelo SARS-CoV-2 e o porquê de algumas mutações terem surgido ao longo do tempo, esperamos que estes resultados sejam levados adiante em estudos *in vitro*, para apenas assim confirmar nossas conclusões. Por último, realizamos uma análise qualitativa das mutações que fizeram parte das variantes que antecederam a Ômicron, e percebemos que realmente algumas mutações que afetaram a variante P.1 e Delta desapareceram, tornando o vírus menos virulento. No entanto, finalizamos ao propor a hipótese de que embora o vírus realmente tenha se enfraquecido ao longo de sua evolução, algumas mutações críticas ainda se mantiveram nas variantes mais recentes. Sendo assim, provavelmente foi devido ao aumento da quantidade de pessoas vacinadas que fez o número de óbitos decrescer de forma expressiva. Contudo, a evolução viral ainda é muito imprevisível, e por isso sempre devemos estar sempre atentos por um possível aumento de casos, e tomar as medidas que possam preservar o maior número de vidas humanas.

Palavras-chave: SARS-CoV-2, vacinas, Ômicron, dinâmica molecular, otimização de Monte Carlo, maturação de afinidade.

Agradecimentos

- Agradeço muito a Deus, por me dar forças e esperança em meio às adversidades, mesmo eu não merecendo tamanho amor.
- Aos meus pais e poucos amigos que formei ao longo da vida, com o mais profundo amor e sincera gratidão. Além disso, fico muito agradecido pela minha gatinha de estimação que me fez mais feliz mesmo nos dias tristes.
- Ao Prof^o. Dr. Kelson Mota pelos grandes ensinamentos, paciência e companheirismo sendo uma das mais importantes pessoas na minha formação científica.
- Meus sinceros agradecimentos à Prof^a Dra. Isabelle Cordeiro pelas importantes ideias de pesquisa quanto à simulação dos anticorpos IgG, IgA e IgM, além da utilização da proteína de Nucleocapsídeo (N) e seus diferentes domínios como um antígeno para maior especificidade nos diagnósticos moleculares. Além de também ter dado um grande apoio na compreensão de inúmeros tópicos em bioquímica, biofísica e imunologia.
- Sou grato ao Msc. Earle Silva pelos importantes ensinamentos no manuseio dos softwares Gaussian 09 e GaussView 6, principalmente quanto ao protocolo para estimar o estado de transição (TS) e coordenada intrínseca de reação (IRC).
- Ao Prof^o Dr. Jonathas Nunes pela inestimável ajuda em toda compreensão teórica das simulações de dinâmica molecular.
- Gratidão à Prof^a Dra. Rosiane de Freitas e Msc. Clarice Santos pelo imenso auxílio na compreensão do funcionamento de vários algoritmos de otimização utilizados ao longo da pesquisa, em principal, o método de Monte Carlo.
- Os Departamentos de Química, Genética e Física foram de grande importância por toda base de conhecimento científico ter sido construída aqui e por isso minha gratidão.
- Agradeço a todos os cientistas que estudaram em toda profundidade o SARS-CoV-2, e cujo conhecimento foi fundamental para realização desta pesquisa.
- Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela concessão financeira de código 128869/2021-6 para esta pesquisa.

Conteúdo

1	REVISÃO BIBLIOGRÁFICA	1
1.1	Proteína Spike (S) e Nucleocapsídeo (N) do SARS-CoV-2	1
1.2	A variante Ômicron	1
1.3	As sub-variantes da Ômicron	2
1.4	Os mecanismos químicos das mutações do SARS-CoV-2	3
1.5	As vacinas universais	4
1.5.1	O anticorpo IgA como principal biomarcador da infecção pelo SARS-CoV-2	4
2	OBJETIVOS	5
2.1	Objetivos específicos	5
3	FUNDAMENTOS TEÓRICOS	6
3.1	Teoria quântica molecular	6
3.1.1	Algoritmos em química computacional	6
3.2	Teoria da dinâmica molecular	9
3.3	Tipos de análise em dinâmica molecular	10
3.4	Maturação de afinidade mediante otimização de Monte Carlo	11
4	METODOLOGIA	12
4.1	Metodologias computacionais	12
4.1.1	Mutagênese e decomposição de energia	12
4.1.2	Dinâmica molecular com Desmond	12
4.1.3	Maximização da interação anticorpo-antígeno	14
4.1.4	Neutralização da proteína Spike pelos anticorpos IgG, IgA e IgM	15
4.1.5	Simulações na interação de anticorpos com a proteína de Nucleocapsídeo (N)	15
4.1.6	Descritores químico-quânticos	16
5	RESULTADOS E DISCUSSÕES	17
5.1	Simulações da variante Ômicron (B.1.1.529)	17
5.1.1	Dinâmica molecular para cristalografia da variante Ômicron	23
5.2	Proteína de nucleocapsídeo (N) nos diagnósticos moleculares	28
5.3	Dinâmica molecular dos anticorpos IgG, IgA e IgM	31
5.4	Simulações de química quântica para o entendimento do mecanismo de mutagênese viral	35
5.5	Descoberta de mutações que maximizam a interação anticorpo-antígeno	36
6	CONSIDERAÇÕES FINAIS	39
	BIBLIOGRAFIA	40

Lista de Figuras

Figura 1	Representação da reação de metilação da citosina (C) que a transforma em 5-metil-citosina (5mC). As estruturas tridimensionais foram geradas a partir do software Gaussian 09W a partir de uma otimização à nível de teoria DFT, com funcional B3LYP e bases 6-311G (2d, 2p).	3
Figura 2	Representação da reação de deaminação da citosina (C) que a transforma em uracila (U).	4
Figura 3	Caixa de solução criada com SystemBuilder no software Schrodinger Maestro 2021-2 onde o antígeno RBD com anticorpo IgG foi inserido (PDB ID: 7BZ5).	13
Figura 4	Visualização do painel de configurações para execução da maturação de afinidade anticorpo-antígeno onde é utilizado a otimização de Monte Carlo para tornar o algoritmo mais eficiente na busca pelas mutações que maximizam a afinidade do complexo proteico.	14
Figura 5	Validação estrutural de estruturas previstas no RosettaFold. O software utilizado para gerar o diagrama de Ramachandran foi o MolProbity incluído no Structure Assessment implementado na ferramenta online SWISS-Model.	15
Figura 6	Comparação entre os valores de RMSD, RMSF, Raio de giro e SASA quanto ao impacto na estabilidade estrutural frente à variante Ômicron para o complexo anticorpo-antígeno (PDB ID: 7BZ5).	19
Figura 7	Comparação da estrutura tridimensional entre a variante Ômicron e sem mutações para o complexo anticorpo-antígeno (PDB ID: 7BZ5) após 100 ns de simulação de dinâmica molecular.	21
Figura 8	Comparação entre o último quadro de dinâmica molecular a 100 ns para o complexo ACE2-RBD (PDB ID: 6M0J).	22
Figura 9	Potencial eletrostático de mutações que afetaram a região RBD da proteína Spike. A estrutura cristalográfica adotada foi o complexo entre o anticorpo IgG e a região RBD (PDB ID: 7BZ5).	22
Figura 10	Em destaque em azul temos o sítio de proteólise conforme predição no software Schrodinger Maestro 2021-2. Uma comparação entre as regiões de proteólise na variante Ômicron e nas mutações ausentes na estrutura é apresentada. A estrutura cristalográfica utilizada foi a região RBD da proteína Spike (PDB ID: 7BZ5).	23
Figura 11	Resultados da dinâmica molecular onde é comparado a estrutura cristalográfica da variante Ômicron (PDB ID: 7T9L) e o complexo ACE2-RBD ausente de mutações (PDB ID: 6M0J).	23
Figura 12	Comparativo estrutural entre o último frame da dinâmica molecular em 100 ns para o complexo ACE2-RBD sem mutações (PDB ID: 6M0J) e a cristalografia com variante Ômicron (PDB ID: 7T9L).	24

Figura 13	Alinhamento estrutural entre o último frame e o primeiro frame. Sendo que para chegarmos nessas estruturas realizamos resultado uma dinâmica molecular em 100 ns no módulo Desmond do software Schrodinger Maestro 2021-2. O complexo simulado foi o ACE2-RBD (PDB ID: 6M0J).	25
Figura 14	Análise comparativa de dinâmica molecular para diferentes domínios NCoV.	28
Figura 15	Diagrama mostrando as ligações químicas formadas entre o domínio NCoV-1 e o anticorpo IgG. A cadeia A representa o domínio do nucleocapsídeo, enquanto a cadeia pesada do anticorpo IgG é representada pela cadeia H. As setas em lilás representam as ligações de hidrogênio formadas. O diagrama foi construído a partir do último quadro de uma simulação de dinâmica molecular a 100 ns.	29
Figura 16	Diagrama de interações químicas formadas entre o domínio NCoV-2 e o anticorpo IgG (PDB ID: 7BZ5). A cadeia A representa a proteína Nucleocapsídeo, enquanto as cadeias H e L representam o anticorpo neutralizante IgG (PDB ID: 7BZ5). Esta imagem corresponde ao último quadro de uma simulação de dinâmica molecular a 100 ns.	30
Figura 17	Diagrama das interações químicas formadas entre o domínio do nucleocapsídeo (NCov3) e o anticorpo IgG (PDB ID: 7BZ5). Esta imagem corresponde à última de uma simulação de dinâmica molecular a 100 ns.	30
Figura 18	Visualização do último frame das simulações de dinâmica molecular a 100 ns para os 3 (três) domínios NCoV-1, NCoV-2 e NCoV-3 referentes à proteína nucleocapsídeo (N) do SARS-CoV-2.	31
Figura 19	(A) Comparação dos valores de RMSD contra o complexo entre RBD e anticorpos IgA1 (monômero) e IgG B38. (B) Comparação das flutuações do RMSF. (C) Comparação de valores de raio de giração. (D) Comparação dos valores SASA.	32
Figura 20	Gráfico RMSF para a forma dimérica do anticorpo IgA (PDB ID: 6UE7). Os instantâneos de estrutura correspondem ao último quadro da simulação de dinâmica molecular em 100 ns.	33
Figura 22	Visualização do último quadro de dinâmica molecular em 100ns do anticorpo IgA (PDB ID: 1OW0) e anticorpo IgG (PDB ID: 7BZ5) neutralizando a região RBD da proteína Spike. Além disso, é apresentado o resultado RMSD do alinhamento estrutural entre o último frame em relação ao primeiro.	33
Figura 21	Diagrama de interações químicas formadas no último quadro de simulações de dinâmica molecular a 100 ns para complexo entre proteína Spike e anticorpo IgA (PDB ID: 1OW0) além do anticorpo IgM (PDB ID: 2AGJ). As setas em lilás representam as ligações de hidrogênio formadas.	34
Figura 23	Mapas de potencial eletrostático (MEPS) de aminoácidos afetados pela variante Ômicron e P.1, que foram otimizados ao nível da teoria DFT com as bases gaussianas 6-311G (2d, 2p) com o funcional híbrido B3LYP no software Gaussian 09W. Todas as figuras foram geradas pelo software GaussView 6.	36
Figura 24	Comparação entre os orbitais de fronteira dos aminoácidos Lisina e Treonina que constituem a mutação K417T do SARS-CoV-2.	36

Lista de Tabelas

Tabela 1	Resultados para variação de estabilidade e afinidade contra a variante Ômicron (B.1.1.529). O complexo analisado foi ACE2-Spike (PDB ID: 7DF4). O sinal positivo denota uma diminuição, enquanto o sinal negativo representa um aumento na propriedade analisada. Os termos de energia ΔE correspondem à estabilidade e não à afinidade. Isso ocorre porque o software sempre retorna dois tipos de ΔE , mas a afinidade é zero quando o aminoácido não está interagindo diretamente com o receptor ACE2.	17
Tabela 2	Resultados de estabilidade e variação de afinidade para variante de interesse Omicron. O complexo analisado estava entre o anticorpo neutralizante Fab 253H55L e a proteína Spike em sua conformação inacessível à interação (PDB ID: 7NDA), o que torna a simulação anticorpo-antígeno mais realista.	18
Tabela 3	Resultados de dinâmica molecular a 100 ns onde foram comparados a estrutura sem mutações e o complexo contendo as mutações da variante Ômicron. O complexo estava entre o anticorpo IgG e a região RBD da proteína Spike (PDB ID: 7BZ5).	19
Tabela 4	A dinâmica molecular em 100 ns ao comparar a variante Ômicron com a estrutura sem mutações. O complexo analisado foi ACE2-RBD (PDB ID: 6M0J).	19
Tabela 5	Resultados de decomposição de energia MM-GBSA para o último quadro de dinâmica molecular a 100 ns para o complexo ACE2-RBD (PDB ID: 6M0J). Todos os valores de energia estão em kcal/mol.	20
Tabela 6	Quantitativo das interações químicas formadas entre o anticorpo IgG (PDB ID: 7BZ5) e a região RBD da proteína Spike com a variante Ômicron e na ausência.	20
Tabela 7	Ligações químicas formadas no complexo anticorpo-antígeno (PDB ID: 7BZ5) com mutações da variante Ômicron.	20
Tabela 8	Ligações químicas formadas no complexo anticorpo-antígeno (PDB ID: 7BZ5) sem mutações.	21
Tabela 9	Resultados de dinâmica molecular em 100 ns mostrando um comparativo entre a variante Ômicron e a estrutura ausente de mutações para o complexo ACE2-RBD (PDB ID: 7T9L).	24
Tabela 10	Comparativo das mutações que constituem cada uma das linhagem da Ômicron (BA.1 a BA.5), além das 4 (quatro) variantes de preocupação que a antecederam. A notação Δ representa a deleção de um aminoácido. As mutações mais críticas e que estiveram na variante P.1 ou Delta foram destacadas em vermelho, e que portanto podem estar associadas a maior virulência.	27
Tabela 11	Resultados para simulações de dinâmica molecular em 100ns no software Desmond para o complexo entre diferentes domínios do nucleocapsídeo do SARS-CoV-2 interagindo com anticorpos. São apresentados os valores médios de RMSD, RMSF, Raio de Giro e SASA.	28

Tabela 12	Análise quantitativa de ligações de hidrogênio, pontes salinas e conflitos de Van der Waals formados no último quadro do complexo anticorpo-antígeno após 100 ns de simulações de dinâmica molecular. Todas as análises foram quantificadas utilizando o módulo Protein Interaction Analysis presente no software Schrodinger Maestro 2021-2.	29
Tabela 13	Resultados para simulações de dinâmica molecular a 100ns no software Desmond para o complexo entre a região RBD da proteína Spike e anticorpos. São apresentados os valores médios de RMSD, RMSF, Raio de giro e SASA.	31
Tabela 14	Resultados de decomposição de energia MM-GBSA para os complexos anticorpo-antígeno. Todos os resultados são correspondentes ao último quadro de simulações de dinâmica molecular em 100ns. Todas as unidades de energia estão em kcal/mol.	34
Tabela 15	Docking molecular para o último quadro nas simulações de dinâmica molecular do complexo antígeno-anticorpo a 100ns. Todas as unidades de energia estão em $kcal \cdot mol^{-1}$	35
Tabela 16	Resultados dos descritores eletrônicos obtidos por otimização estrutural à nível da teoria DFT com as bases gaussianas $6 - 311G + +(2d, 2p)$ e função híbrida B3LYP. Todas as minimizações foram feitas usando o software Gaussian 09W.	35
Tabela 17	Maximização da afinidade para as combinações de mutações que maximizam a interação entre a proteína RBD e o anticorpo IgG B38 (PDB ID: 7BZ5). Todos os valores estão na unidade de $kcal/mol$	37
Tabela 18	Maximização da afinidade para 16 (dezesesseis) mutações entre a proteína RBD e o anticorpo CB6 IgG (PDB ID: 7C01). Todos os valores estão na unidade de $kcal/mol$	37
Tabela 19	Análise de afinidade para as 16 (dezesesseis) mutações que maximizam a interação entre a proteína RBD e o monômero do anticorpo IgA1-Fc (PDB ID: 1OW0). Todos os valores estão na unidade de $kcal/mol$	37
Tabela 20	Maximização da afinidade para as 16 (dezesesseis) mutações entre a proteína RBD e o anticorpo IgM (PDB ID: 2AGJ). Todos os valores estão na unidade de $kcal/mol$	37

Lista de abreviações e siglas

ACE2	Enzima conversora da angiotensina 2, do inglês “Angiotensin-converting enzyme 2”
APBS	Solucionador adaptativo de Poisson-Boltzmann, do inglês “Adaptive Poisson-Boltzmann Solver”
B.1.1.7	Linhagem da variante de preocupação Alfa que surgiu no Reino Unido.
B.1.351	Linhagem da variante de preocupação Beta que surgiu na África do Sul.
P.1	Linhagem da variante de preocupação Gama que surgiu em Manaus, Amazonas, Brasil.
B.1.617.2	Linhagem da variante de preocupação Delta que surgiu na África do Sul.
B.1.1.529	Linhagem da variante de preocupação Ômicron que surgiu em Botsuana.
B3LYP	Funcional híbrido de troca de Becke, chamado B3, combinado com o funcional de correlação, denominado LYP (Lee, Yang e Parr)
CHARMM	Química em Mecânica Macromolecular de Harvard, do inglês “Chemistry at Harvard Macromolecular Mechanics”
DFT	Teoria do Funcional da Densidade, do inglês “Density Functional Theory”
HOMO	Orbital molecular mais altamente ocupado, do inglês “Highest occupied molecular orbital”
LUMO	Orbital molecular mais fracamente ocupado, do inglês “Lowest unoccupied molecular orbital”
NPT	Ensemble com número de partículas, pressão e temperatura constantes, do inglês “Isothermal-isobaric ensemble”
MD	Dinâmica molecular, do inglês “Molecular dynamics”
MEP	Mapa de potencial eletrostático, do inglês “Molecular Electrostatic Potential”
MM-PBSA	Mecânica molecular e área de superfície de Poisson-Boltzmann, do inglês “Molecular mechanics Poisson-Boltzmann surface area”
PBC	Condições periódicas de contorno, do inglês “Periodic boundary conditions”
PDB	Banco de Dados de Proteínas, do inglês “Protein Data Bank”
RBD	Domínio de ligação ao receptor, do inglês “Receptor binding domain”
Rg	Raio de giro, do inglês “Radius of Gyration”
RMSD	Desvio quadrático médio, do inglês “Root Mean Square Deviation”

RMSF	Flutuação média quadrática, do inglês “Root Mean Square Flutuation”
SASA	Área de superfície acessível ao solvente, do inglês “solvent-accessible surface area”
WHO	Organização Mundial da Saúde, do inglês “World Health Organization”
WT	Estado nativo da proteína sem mutações ou “selvagem”, do inglês “Wild-type”
VOC	Variante de preocupação, do inglês “Variants of concern”

Lista de constantes físico-químicas

a_0	$5,29177210903(80) \times 10^{-11} m$ (Raio de Bohr)
c	$299.792.458 m \cdot s^{-1}$ (Velocidade da Luz)
e_c	$1,602176634 \times 10^{-19} C$ (Carga elétrica elementar do elétron)
E_h	$4,3597447222071(85) \times 10^{-18} J$ (Energia de Hartree)
k_B	$8,617333262 \times 10^{-5} eV \cdot K^{-1}$ (Constante de Boltzmann)
k_e	$8,9875517923(14) \times 10^9 N \cdot m^2 \cdot C^{-2}$ (Constante de Coulomb)
m_e	$9,1093837015(28) \times 10^{-31} kg$ (Massa do elétron)
m_p	$1,67262192369(51) \times 10^{-27} kg$ (Massa do próton)
m_n	$1,67492749804(95) \times 10^{-27} kg$ (Massa do nêutron)
m_H	$1,6735 \times 10^{-27} kg$ (Massa do átomo de Hidrogênio)
N_A	$6,02214076 \times 10^{23} mol^{-1}$ (Constante de Avogadro)
h	$6,62606957 \times 10^{-34} J \cdot s$ (Constante de Planck)
R	$8,31446261815324 J \cdot K^{-1} \cdot mol^{-1}$ (Constante universal dos gases)

Embora não tenhamos utilizados diretamente estas constantes, os softwares adotados nesta pesquisa precisam rotineiramente ao longo das iterações seja na dinâmica molecular, cálculos de química-quântica ou docking molecular. Os valores foram obtidos na base de dados NIST (<https://physics.nist.gov/cuu/Constants/Table/allascii.txt>) e referem-se às correções realizadas até 2018 descobertas pelo Comitê de Dados do Conselho Internacional de Ciência (CODATA) [2].

1 Revisão bibliográfica

A emergência de novas variantes do SARS-CoV-2 foi o principal motivo pelo aumento drástico no número de casos desde o início da pandemia em 2020. Dentre estas, houve a variante P.1 tendo sido a principal responsável pela grave 2^a onda de casos que sobrecarregou todo o sistema de saúde público na cidade de Manaus, e que infelizmente gerou a perda de várias vidas [3], [4].

Diante das novas variantes preocupantes do SARS-CoV-2, percebemos que grande parte delas aumenta a transmissibilidade porém também há indícios de maior letalidade. Porém este comportamento não vem ocorrendo na última variante, a Ômicron, onde há importantes indícios de uma menor virulência, ou seja, menor agressividade do vírus [5], [6]. A vacinação também vem exercendo um papel central na queda de hospitalizações. Nisto, um grupo de pesquisadores confirmou por meio de dados da vacinação em Israel que houve uma redução significativa no número de pacientes que desenvolveram a forma mais grave da doença [7].

1.1 PROTEÍNA SPIKE (S) E NUCLEOCAPSÍDEO (N) DO SARS-COV-2

O vírus SARS-CoV-2 codifica um total de quatro proteínas estruturais, a proteína spike (S), a proteína do envelope (E), a proteína da membrana (M) e a proteína do nucleocapsídeo (N). A proteína (N) participa do empacotamento de RNA e liberação de partículas virais. A proteína (N) contém dois domínios de ligação ao RNA, o domínio N-terminal (NTD) e o domínio C-terminal (CTD) [8].

Uma pesquisa realizou medições de anticorpos em pacientes infectados com SARS-CoV-2 ao reagir com o nucleocapsídeo e a proteína Spike. Assim, os anticorpos reativos contra a proteína do nucleocapsídeo apresentaram sensibilidade e especificidade de quase 100%, enquanto os anticorpos para a proteína Spike mostraram uma diminuição para 91% de sensibilidade. Portanto, foi demonstrado que os diagnósticos baseados em nucleocapsídeos foram mais sensíveis do que aqueles baseados na proteína Spike na detecção de infecção precoce [9].

1.2 A VARIANTE ÔMICRON

Primeiramente devemos lembrar que novas variantes como a Ômicron são apenas uma consequência natural da evolução do vírus, e por isso enquanto a COVID-19 não estiver totalmente erradicada, o surgimento de novas variantes é eminente. O vírus SARS-CoV-2 está evoluindo para dificultar cada vez mais a resposta do sistema imunológico humano, mas não necessariamente vem se tornando mais agressivo com o tempo [10].

A OMS vem nomeando variantes emergentes do novo coronavírus por meio de letras do alfabeto grego. A primeira a surgir foi a variante Alfa, no ano de 2020, desde então o vírus vem evoluindo até chegar as mais recentes subvariantes da Ômicron. Esta variante foi inicialmente detectada na Botsuana e África do Sul em novembro de 2021. Ainda que a variante Ômicron tenha aumentado a transmissibilidade do vírus SARS-CoV-2, vem causando sintomas menos graves do que a Delta [10].

A existência da variante Ômicron vem sendo associada à terceira onda de casos da COVID-19 no estado do Amazonas ao longo do mês de janeiro de 2022. Felizmente devido à grande taxa de pessoas vacinadas, o número de óbitos no estado vem sendo extremamente baixo [11].

A variante de preocupação B.1.1.529 (Ômicron) foi designada pela OMS (Organização Mundial da Saúde) em 26 de novembro de 2021. Essa variante possui 32 mutações sem precedentes na proteína Spike, que é justamente a principal proteína pela qual as vacinas treinam o sistema imune. É importante destacar que o impacto total destas mutações não é simplesmente aditivo entre si, e portanto é fundamental também considerar a epistasia, e portanto, como o coletivo das mutações afetam umas às outras. Esta variante foi relatada pela primeira vez na África do Sul e apresenta alguns riscos para aumento da transmissibilidade viral ou dificuldades de detecção no diagnóstico de RT-PCR [12].

Uma pesquisa modelou computacionalmente por meio de dinâmica molecular a interação entre a proteína ACE2 e a região RBD com mutações da variante Ômicron. A partir da análise, foi demonstrado que diante desta nova variante, a região RBD interage mais fortemente com a proteína ACE2 em comparação com a estrutura sem mutações. Isso foi confirmado pelo aumento no número de ligações de hidrogênio [13]. Outra pesquisa também buscou entender a modificação estrutural causada por mutações na variante Ômicron na região RBD devido à interação com anticorpos neutralizantes. Para isso, empregaram inicialmente a previsão de estrutura secundária e, posteriormente, simulações de dinâmica molecular. Assim, por meio de análise computacional percebeu-se que a variante Ômicron possui maior afinidade de ligação para o receptor ACE2 em comparação à cepa de Wuhan. Outro achado notável foi de que a região RBD era relativamente mais instável e compacta. E por último, notou-se uma afinidade de ligação menor para anticorpos monoclonais observada diante desta nova variante [14].

A maioria das mutações no vírus é neutra e, portanto, não geram alterações fenotípicas perceptíveis. No entanto, mutações que surgem devido à pressão seletiva positiva podem aumentar a aptidão viral (fitness). Essas mutações são as principais responsáveis pelo surgimento de novas variantes [15]. Entre as variantes do SARS-CoV-2, algumas são chamadas de Variantes de Preocupação (VOC), enquanto outras são apenas Variantes de Interesse (VOI), e felizmente até hoje nunca veio a surgir uma Variante de Alta Preocupação (VOHC), que aumentaria de forma significativa a virulência e patogenicidade.

1.3 AS SUB-VARIANTES DA ÔMICRON

Nos primeiros meses de 2022, uma subvariante Ômicron chamada BA.2 começou a se espalhar. Depois, o vírus evoluiu para as subvariantes BA.4 e BA.5, que são consideradas mais transmissíveis que as anteriores, porém não houve um aumento de letalidade. A subvariante BA.5 surgiu no início de julho de 2022, sendo a mais contagiosa até o momento [16]. Talvez a questão mais importante seja se as atuais vacinas estarão realmente protegidas contra as novas variantes. Embora algumas pesquisas tenham concluído que as subvariantes BA.4 e BA.5 dificultam o reconhecimento de anticorpos induzidos após a vacinação ou infecção, a imunogenicidade é pouco afetada quando se administra mais de uma dose da vacina [16].

Em essência, descobriu-se que o SARS-CoV-2 vem evoluindo para aumentar cada vez mais sua capacidade de transmissão, mas também para conseguir escapar da neutralização de anticorpos. Do ponto de vista molecular, o vírus vem aumentando sua afinidade na interação com o receptor ACE2. Em todas as variantes de preocupação, notou-se terem surgido de forma inesperada, e possivelmente com um comportamento estocástico. No entanto, o ambiente pode

umentar a probabilidade do surgimento de variantes, principalmente em consequência da baixa adesão ao distanciamento social principalmente durante os picos da pandemia [16].

Uma importante pesquisa analisou em cultivo celular como a variante Ômicron se comportou em comparação a P.1 e Delta, onde concluiu-se curiosamente que a Ômicron é menos virulenta [5]. Isso se reflete portanto em menor agressividade do vírus ao longo do seu processo evolutivo. Outra importante pesquisa também notou que na variante Ômicron houve uma diminuição no número de hospitalizações e mortes em comparação à variante anterior, Delta [6].

1.4 OS MECANISMOS QUÍMICOS DAS MUTAÇÕES DO SARS-COV-2

As mutações $C > U$ vêm sendo muito frequentes pois refletem os processos de adaptação do vírus SARS-CoV-2. Desta forma, a frequente deaminação de citosina (ver Figura 2) está conduzindo uma evolução muito rápida [17]. Outro mecanismo fundamental para a ocorrência de mutações é a metilação do DNA, mas praticamente não há nenhum estudo relacionando ao processo evolutivo do SARS-CoV-2. Em essência, a metilação é um mecanismo epigenético que consiste na inserção de um grupo metil na posição C5 da citosina, formando então 5-metil-citosina (ver Figura 1) [18]. Desta forma, o aumento da taxa de metilação do DNA, tem um papel fundamental na oncogênese [19] e na neuropsiquiatria [18], mas ainda é incerto se exerce um papel no agravamento nos sintomas da COVID-19.

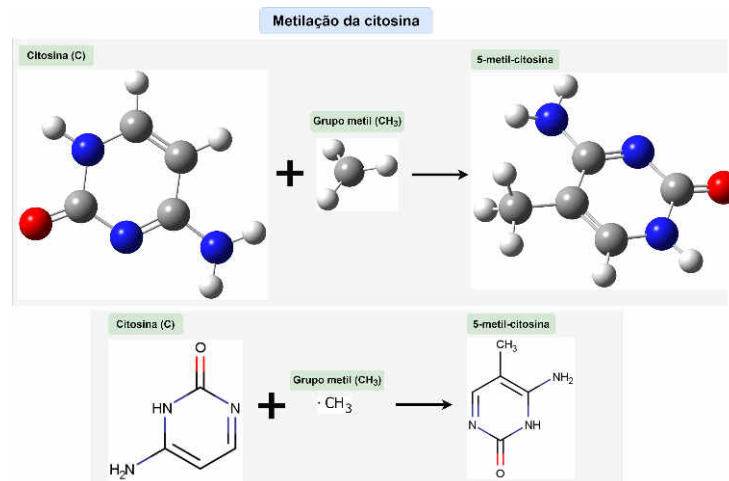


Figura 1: Representação da reação de metilação da citosina (C) que a transforma em 5-metil-citosina (5mC). As estruturas tridimensionais foram geradas a partir do software Gaussian 09W a partir de uma otimização à nível de teoria DFT, com funcional B3LYP e bases 6-311G (2d, 2p).

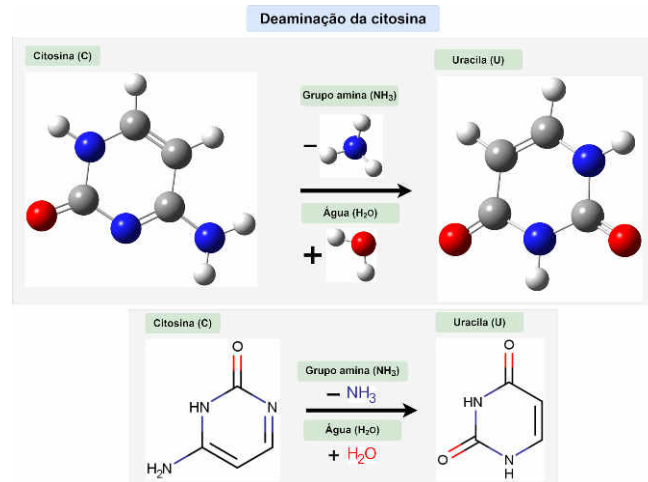


Figura 2: Representação da reação de deaminação da citosina (C) que a transforma em uracila (U).

1.5 AS VACINAS UNIVERSAIS

Um dos grandes desafios da ciência é desenvolver vacinas cuja imunogenicidade não seja afetada por variantes emergentes, ou mesmo criar vacinas antes mesmo de um possível novo coronavírus vir a existir no futuro [20]. Dentre os vários métodos em estudo há aqueles que utilizam o antígeno do primeiro SARS-CoV ou dos coronavírus que afetam outras espécies ou mesmo outras famílias de coronavírus [20]. Além de também haver pesquisas que utilizam as vacinas em nanopartículas de ferritina [20]. Uma das abordagens mais promissoras é utilizar tanto a proteína Spike como o Nucleocapsídeo (N) nas vacinas, pois isto não induz a resposta apenas de anticorpos mas também de células T [20].

1.5.1 O anticorpo IgA como principal biomarcador da infecção pelo SARS-CoV-2

Entre os anticorpos mais importantes temos o tipo IgA que prevalece na resposta inicial do corpo humano à infecção pelo SARS-CoV-2, em comparação com as concentrações de IgG e IgM. Além disso, os anticorpos IgA apresentam maior neutralização do vírus em comparação ao tipo IgG. O anticorpo IgA é predominante na proteção da mucosa respiratória como linha de defesa contra patógenos [21]. Os anticorpos IgG embora desempenhem um papel fundamental na neutralização do vírus, há pouco na literatura sobre os anticorpos secretores de IgA que são predominantes nas fases iniciais da infecção viral. Verificou-se que os monômeros de IgA para SARS-CoV-2 eram duas vezes mais potentes que os anticorpos IgG. Por outro lado, os dímeros de IgA apresentam comportamento oposto, sendo superiores aos anticorpos IgG e à forma monomérica do IgA [22].

2 Objetivos

Ao longo de toda esta pesquisa buscamos entender as consequências termodinâmicas da variante Ômicron do SARS-COV-2 diante da interação anticorpo-antígeno e ACE2-Spike. Isto porque, almejamos responder computacionalmente porque determinadas mutações virais predominam sobre outras. Mas a principal pergunta foi buscar responder se o vírus realmente vem se tornando menos virulento ao longo de sua evolução. Por último, tentamos descobrir quais mutações poderiam maximizar a afinidade anticorpo-antígeno. Em acréscimo, nos perguntamos se ao invés dos diagnósticos moleculares utilizarem a proteína Spike (S), mas na verdade utilizarem o Nucleocapsídeo (N) e seus diferentes domínios para aperfeiçoar a especificidade dos diagnósticos. Outro ponto notável foi elucidar qual classe de anticorpos neutralizantes possuem maior afinidade com o vírus. Por fim, nos propomos a descobrir quais mutações no antígeno poderiam maximizar a neutralização por anticorpos.

2.1 OBJETIVOS ESPECÍFICOS

- Simular a dinâmica molecular em $100ns$ da variante Ômicron quanto à interação da proteína Spike (S) com os anticorpos neutralizantes IgG B38, IgA1-Fc (monômero), sIgA (dímero) e IgM;
- Simular a dinâmica molecular em $100ns$ referente à variante Ômicron na interação ACE2-Spike para entender o processo de invasão celular, de modo que estas mutações foram todas aplicadas computacionalmente;
- Simular a dinâmica molecular em $100ns$ para a variante Ômicron também na interação ACE2-Spike, mas desta vez as mutações foram obtidas por cristalografia a partir da base de dados PDB;
- Compreender mediante simulações de dinâmica molecular em $100ns$ como os diferentes domínios da proteína de Nucleocapsídeo (N) interagem com anticorpos neutralizantes;
- Utilizar métodos computacionais de química-quântica para calcular os descritores eletrônicos e assim desvendar porque a mutação K417T é favorável do ponto de vista termodinâmico, e porque está presente em quase todas as variantes de preocupação (VOC) como Ômicron e P.1;
- Empregar o método de otimização de Monte Carlo para encontrar quais combinações de mutações favorecem a neutralização do antígeno (S) na interação com anticorpos.

3 Fundamentos teóricos

3.1 TEORIA QUÂNTICA MOLECULAR

Antes de adentrar nos atuais métodos computacional de química quântica molecular, precisamos revisar como foi o surgimento da mecânica quântica. Por muito tempo, químicos e físicos acreditavam que as regras que governam os sistemas microscópicos e macroscópicos eram semelhantes. Apenas no ano de 1900 que Max Planck ofereceu uma proposta inovadora de que a radiação emitida pelo corpo negro estava limitada a valores discretos, ou seja, estava “quantizada”. Conforme o século XX avançava, tornou-se cada vez mais evidente de que a quantização era não apenas uma característica da luz, mas também das partículas fundamentais das quais a matéria é constituída [23].

Desta forma, para descrever sistemas microscópicos, então, uma mecânica diferente foi necessária. Um candidato promissor foi a mecânica das ondas, já que ondas estacionárias também são quantizadas, como proposto inicialmente por De Broglie. No entanto, descobriu-se também apresentar propriedades semelhantes a partículas, e para explicar essa dicotomia, uma nova mecânica, a mecânica quântica, foi desenvolvida [23]. O operador que retorna a energia do sistema, E , possui um valor próprio sendo denominado de operador Hamiltoniano, \hat{H} . Temos então a equação de *Schrödinger* para o átomo de hidrogênio [23].

$$\hat{H}\Psi(r, R) = E\Psi(r, R) \quad (1)$$

$$\mathcal{H} = -\sum_i \frac{\hbar^2}{2m_e} \nabla_i^2 - \sum_k \frac{\hbar^2}{2m_k} \nabla_k^2 - \sum_i \sum_k \frac{e^2 Z_k}{r_{ik}} + \sum_{i<j} \frac{e^2}{r_{ij}} + \sum_{k<l} \frac{e^2 Z_k Z_l}{r_{kl}} \quad (2)$$

3.1.1 Algoritmos em química computacional

Os algoritmos em química quântica podem ajudar a prever propriedades moleculares ou espectros para posterior comparação com o experimento, nos ajudando a elucidar dados experimentais pouco claros ou mesmo modelar estados intermediários em compostos onde é impossível a observação empírica. Os diferentes métodos teóricos existentes correspondem a diferentes aproximações da equação de Schrödinger, e cuja utilização está em conformidade com o problema estudado e a infraestrutura computacional [24]. O principal objetivo da otimização estrutural de uma molécula é para que o modelo teórico se assemelhe ao máximo com a forma em que é encontrada na natureza. Desta forma, a configuração de energia mínima e mais favorável corresponde justamente ao estado mais estável do sistema [24].

3.1.1.1 Teoria DFT

Em um sistemas de muitos elétrons, uma das principais abordagens de aproximação é a teoria do funcional de densidade (DFT). Desta forma, todos os graus de liberdade do sistema eletrônico

são integrados externamente, com exceção da densidade eletrônica. Apesar de tudo, a qualidade da aproximação dependerá do funcional de correlação e troca adotado. A base do DFT é a prova de Hohenberg-Kohn de que a energia eletrônica do estado fundamental pode ser determinada pela densidade de elétrons ρ [25]. Em outras palavras, existe uma correspondência individual entre a densidade eletrônica de um sistema e sua energia. Descritores moleculares eletrônicos principalmente relacionados a energia dos orbitais de fronteira, podem indicar as regiões de maior reatividade [25]. Tudo o que desconhecemos sobre o problema de muitos corpos está portanto inserido no termo de troca e correlação.

Os métodos DFT incluem os efeitos da correlação eletrônica com B3LYP embora tenha um custo de processamento menor em relação à teoria Hartree-Fock [24]. Dentre os métodos de troca e correlação existentes, há o método puro e funcionais híbridos onde incluem alguma porcentagem da correlação Hartree-Fock [24]. O conjunto de base é um conjunto de funções matemáticas para construir a função de onda da mecânica quântica em um sistema de muitos corpos. Os conjuntos de base maiores impõem menos restrições aos elétrons e aproximam com maior precisão as funções de onda [24].

3.1.1.2 Descritores químico-quânticos

O mapa de potencial eletrostático (MEP) descreve como a energia de interação em uma determinada região molecular é afetada pela distribuição da carga elétrica ocasionada por prótons e elétrons posicionados em uma dada posição r onde o potencial $V(r)$ é descrito pela [Equação 3](#) [24]. Desta forma, um potencial eletrostático negativo corresponde a alta densidade de elétrons. Por outro lado, um potencial eletrostático positivo corresponde a baixa densidade de elétrons sendo um indício da blindagem do núcleo e como consequência haverá maior probabilidade de repelir prótons [24].

$$V(r) = \sum_A \frac{Z_A}{(R_A - r)} - \int \frac{\rho(r')}{|r' - r|} dr' \quad (3)$$

Na equação anterior temos que Z_A representa a carga do núcleo A posicionado em um raio R_A enquanto r' descreve a variável de integração, enquanto $\rho(r)$ representa a função de densidade eletrônica [24]. Ao construir o MEP é possível prever regiões de maior probabilidade para ataque nucleofílico onde geralmente são destacadas em azul enquanto regiões vermelhas o ataque eletrofílico [24].

Os termos HOMO e LUMO são usados para se referir aos orbitais moleculares mais altamente e menos ocupados, respectivamente. Assim, constituem os chamados orbitais de fronteira, um local de grande reatividade e atividade espectroscópica da molécula [26]. A reatividade em química orgânica é um dos campos mais amplamente estudados, pois seu entendimento nos possibilita compreender melhor de que modo as reações químicas ocorrem. Podemos descrever a reatividade química pelo gap de energia entre os orbitais de fronteira, onde:

$$\Delta E = I - A = E_{LUMO} - E_{HOMO} \quad (4)$$

Desta forma, os índices de reatividade oriundos da teoria DFT são parâmetros indispensáveis para análises teóricas em química orgânica. Desde então, um grande número de índices de reatividade têm sido propostos ao longo do tempo. Dentre estes o potencial químico eletrônico:

$$\mu = \frac{I + A}{2} = -\frac{E_{HOMO} + E_{LUMO}}{2} \quad (5)$$

Além da eletrofilicidade ω , os índices de nucleofilicidade N e dureza η são expressos por:

$$\eta = \frac{I - A}{2} = \frac{E_{LUMO} - E_{HOMO}}{2} \quad (6)$$

Enquanto seu oposto maciez química é descrito por [27]:

$$\zeta = \frac{1}{2\eta} = \frac{1}{E_{LUMO} - E_{HOMO}} \quad (7)$$

A afinidade de ionização pode ser expressa por:

$$I = -E_{HOMO} \quad (8)$$

Enquanto isso a afinidade eletrônica é descrito por:

$$A = -E_{LUMO} \quad (9)$$

Em 1999, Robert Parr definiu o índice de eletrofilicidade:

$$\omega = \frac{\mu^2}{2\eta} \quad (10)$$

O ω é responsável por medir a energia de estabilização de uma molécula quando adquire uma quantidade de densidade de elétrons, ΔN do meio. O índice de eletrofilicidade explica a tendência de um eletrófilo adquirir uma densidade de elétrons μ além da resistência de uma molécula para trocar elétrons com o meio η [28]. Por fim, para quantificarmos a eletrofilicidade de moléculas orgânicas, também podemos expressar como função dos índices de Fukui, onde:

$$f_k^+ = q_k(N + 1) - q_k(N) \quad (11)$$

A nucleofilicidade N representa a propensão de um sistema em doar densidade de elétrons. Em 2008, foi proposto um índice empírico de nucleofilicidade para moléculas orgânicas podendo ser descrito por [29]:

$$N = E_{HOMO}(Nucléofilo) - E_{HOMO}(TCE) \quad (12)$$

A abreviação TCE refere-se ao tetracianoetileno, que é o composto de maior potencial eletrofílico [29].

3.2 TEORIA DA DINÂMICA MOLECULAR

As proteínas são as principais moléculas da vida, onde medicações e muitas doenças tem sua origem. Além disso temos mutações que afetam a conformação da proteína e alteram drasticamente sua atividade biológica. Nosso conhecimento de proteínas e suas interações deve-se muito à existência de métodos computacionais, como dinâmica molecular [30]. A energia livre de todo o sistema é a medida fundamental da estabilidade de proteínas. enquanto nossas observações de propriedades das proteínas na verdade refletem seu comportamento médio [30]. A mecânica estatística fornece a ponte entre o mundo microscópico e o macroscópico [30]. Sendo assim, a dinâmica molecular sem dúvidas é uma das mais importantes aplicações da mecânica estatística clássica, capaz de descrever uma ampla variedade de sistemas biológicos baseando-se em sua essência nas Leis de Newton para N-partículas [30]. Embora as simulações de dinâmica molecular sejam abordagens teóricas, na verdade muito se assemelham a experimentos reais [30]. Desta forma compartilham também das limitações de abordagens experimentais, tal como as medições que estão sujeitas a ruído estatístico, e por isso é imprescindível que a simulação seja a mais longa possível, o que tornará a média mais precisa e convergência dos resultados [30].

A dinâmica molecular gera interações sucessivas do sistema, integrando portanto as leis clássicas do movimento de Newton [30]. O resultado é a trajetória que descreve as posições e velocidades das partículas no sistema ao longo da simulação como função do tempo [30]. As forças são calculadas a partir do campo de força, que é um conjunto de funções matemáticas que estimam a energia potencial do sistema. Uma descrição da energia potencial total em um sistema macromolecular inclui as interações intermoleculares e as interações intramoleculares [30]. Na mecânica molecular, as funções de energia potencial são derivadas empiricamente [30].

A definição da temperatura na dinâmica molecular requer o uso do teorema da equipartição de energia. Além disso, a energia total do sistema E_{total} é aproximadamente constante ao longo da simulação [30].

$$\langle \frac{1}{2}mv^2 \rangle = \frac{1}{2}k_B T \quad (13)$$

$$F = m \cdot \alpha = -\frac{\partial V}{\partial r} = -\nabla U(x) \quad (14)$$

Na preparação da simulação, é necessário termos a topologia do sistema, a qual define não apenas pelas coordenadas tridimensionais dos átomos, mas também descreve como esses átomos estão conectados e interagem entre si [30]. Posteriormente devemos organizar os átomos da melhor maneira possível, o que é obtido por meio de uma minimização de energia do sistema [30]. Uma das primeiras etapas da dinâmica molecular constitui na adição de contra-íons para neutralizar a carga elétrica do sistema para torná-lo estável, adicionar solventes explícitos como moléculas de água. Em seguida temos a minimização de energia, onde é encontrado a conformação que está com a mínima energia potencial. A próxima etapa é semelhante à anterior, mas agora é denominada de equilíbrio pois agora a temperatura do sistema é mantida constante. E por fim, é realizado a produção das trajetórias do sistema, onde apenas agora teremos propriamente

dito o início da dinâmica molecular e cujos resultados gerados serão imprescindíveis na análise de estabilidade do sistema.

3.3 TIPOS DE ANÁLISE EM DINÂMICA MOLECULAR

Ao longo da análise de qualidade da simulação, devemos verificar a convergência da temperatura e pressão, além da energia potencial e cinética, pois caso contrário a simulação precisará ser repetida. Em resumo, é preciso verificar se o equilíbrio foi realmente alcançado na simulação [30].

Uma abordagem computacional para a entropia de uma proteína é mensurar o quão rígida é a estrutura. Dessa forma, podemos calcular os modos vibracionais de uma molécula até atingir uma frequência de ressonância [30]. Este cálculo é chamado de análise de modo normal da proteína. A área de superfície acessível ao solvente (SASA) busca medir a superfície efetiva da proteína que interage diretamente com moléculas de água [30], sendo estimada pelo algoritmo Shrake & Rupley [31]. Isto é importante, pois geralmente as regiões hidratadas são justamente onde se localiza o sítio ativo, e portanto a região de interação com o receptor.

$$\langle A \rangle_{ensemble} = \langle A \rangle_{time} \quad (15)$$

A força em um determinado átomo i ao longo de um vetor de coordenadas r_i pode ser genericamente descrita pela inclinação de V_{total} [32]:

$$F_i = -\vec{\nabla} \cdot U = m \frac{d^2 r}{dt^2} = -\frac{\partial V}{\partial r} - m\beta \frac{dr}{dt} + R(t) \quad (16)$$

A nova posição das partículas é calculada por [32]:

$$r(t + \delta t) = r(t) + \delta t \cdot v(t) + \frac{1}{2} \delta t^2 a(t) \quad (17)$$

Na equação abaixo podemos ver os principais termos que constituem o campo de força na dinâmica molecular:

$$E_{potential} = \sum_{bonds} \frac{1}{2} k_r (r - r_0)^2 + \sum_{angles} \frac{1}{2} k_\theta (\theta - \theta_0)^2 + \sum_{dihedrals} \frac{1}{2} k_\phi [1 + \cos(n\phi - \delta)]^2 + \sum_{impropers} k(\omega - \omega_0)^2 + \frac{1}{4\pi\epsilon_0} \sum_{i,j} \frac{q_i q_j}{r_{ij}^2} \quad (18)$$

Dentre os parâmetros de análise mais importantes na dinâmica molecular é o raio de giro (R_g), o RMSD e o RMSF [32]. A função do raio de giro é medir a compactação do sistema, e conseqüentemente quanto maior o R_g menor será a estabilidade do sistema, e menor será a compactação. Em relação aos valores de RMSD e RMSF, ambos medem a estabilidade estrutural do sistema. Logo, quanto maior for o RMSD e o RMSF, menor será a estabilidade. A diferença entre eles reside de que o RMSD mede o deslocamento atômico de toda proteína em função do tempo. Por outro lado, o valor do RMSF mede as flutuações atômicas médias, além de serem individuais para cada aminoácido. Resumindo, no RMSF as flutuações não são mais em função do

tempo como no RMSD, mas em relação a cada aminoácido [32]. Abaixo, podemos visualizar como são realizados os cálculos da análise de estabilidade estrutural após a conclusão das simulações de dinâmica molecular.

$$R_G^2 = \sum m_i (r_i - R_C)^2 / M \quad (19)$$

$$RMSD = \sqrt{\frac{1}{n} \sum_{i=1}^N (r_i - r_i^{ref})^2} \quad (20)$$

$$RMSF = \sqrt{\frac{1}{t} \sum_{t_j=1}^T |r_i(t_j) - r_i^{ref}|^2} \quad (21)$$

3.4 MATURAÇÃO DE AFINIDADE MEDIANTE OTIMIZAÇÃO DE MONTE CARLO

As técnicas de Monte Carlo são introduzidas para resolver problemas determinísticos de forma mais eficiente [33]. Desta forma, o método de Monte Carlo é uma ferramenta poderosa para a otimização de funções objetivo complexas. Na grande maioria das vezes as funções são determinísticas e a aleatoriedade é introduzida artificialmente para buscar de forma mais eficiente o domínio da função objetivo [33]. Portanto este método constitui uma das abordagens mais importantes da computação científica devido à sua relativa simplicidade e aplicabilidade nos mais variados campos da ciência [33].

O processo de maturação de afinidade consiste em maximizar a interação anticorpo-antígeno, o que beneficia diretamente a resposta imune do ser humano diante da infecção pelo vírus SARS-CoV-2. Por outro lado, também é possível beneficiar o vírus mediante a maximização da afinidade entre o receptor ACE2 e a proteína Spike. O processo de mutagênese pode ser feito mediante enumeração exaustiva, porém demandaria grande capacidade computacional e poderia levar a uma explosão combinatória, por simplesmente testar cada uma das mutações possíveis. Por outro lado, é possível utilizar uma otimização de Monte Carlo onde é feito um passeio aleatório em relação às varias combinações possíveis de mutações [34]. Os resultados de cada caminhada aleatória são selecionados de modo que se obtenha os melhores valores de afinidade [34]. O critério de aceitação para a próxima iteração é baseado em um valor pré-determinado para mudança de afinidade [34]. A temperatura absoluta para a probabilidade de Boltzmann também pode afetar o critério de aceitação para a próxima iteração [34]. A pesquisa pode ser encerrada se nenhum movimento aceito for encontrado em um número pré-determinado de iterações máximas, enquanto rejeitará etapas cuja afinidade mude em um valor superior ao especificado para otimizar a afinidade [34]. E por fim, também podemos definir um limite no número máximo de mutações simultâneas, restringindo assim, o espaço de busca. Conseqüentemente, não precisamos testar absolutamente todas as mutações, mas otimizar a busca para calcular apenas as que realmente possam aumentar a afinidade.

4 Metodologia

4.1 METODOLOGIAS COMPUTACIONAIS

4.1.1 *Mutagênese e decomposição de energia*

A localização e a informação exata de quais aminoácidos foram substituídos (mutações não sinônimas) e que afetaram a proteína Spike (S) que constituem a linhagem Ômicron e suas subvariantes (BA.1, BA.2, BA.2.12.1, BA.2.75, BA.4 e BA.5) foram todas obtidas a partir da plataforma CoVariants (<https://covariants.org/>) [35]. Esses dados foram imprescindíveis tanto para a dinâmica molecular, como também para a decomposição de energia MM-GBSA que foi calculado pelo algoritmo Prime [36]. Também realizamos uma comparação qualitativa da variante Ômicron e suas subvariantes em relação às 4 (quatro) VOC que a antederam, isto é, Alfa (Variante surgida no Reino Unido), Beta (Variante surgida na África do Sul), P.1 ou Gama (Variante surgida em Manaus) e Delta (Variante surgida na África do Sul).

Inicialmente, as estruturas foram preparadas utilizando o módulo "Protein Preparation Wizard"[37] integrado ao software Schrodinger Maestro 2021-2 [38], onde foi possível adicionar os átomos de hidrogênio ausentes, atribuir as cargas parciais e corrigir os problemas de valência. Posteriormente, houve um refinamento através da minimização estrutural com campo de força OPLS4 [39] para que os átomos de hidrogênio fossem otimizados para um $pH \approx 7,4$. A inserção das mutações da variante Ômicron que afetaram a região RBD da proteína Spike foi realizada utilizando também o software Schrodinger Maestro 2021-2 com o módulo "Residue and Loop Mutation". Após cada mutação, uma minimização estrutural em solvente implícito foi aplicada a cada proteína cuja cristalografia adotada foi o complexo entre o anticorpo IgG e a região RBD (PDB ID: 7BZ5) [40]. E sendo assim, todas as estruturas cristalográficas foram obtidas a partir da base de dados Protein Data Bank (PDB) [41]. Dentre as 32 mutações que afetaram a proteína Spike, aplicamos apenas 16 mutações na região RBD: G339D, S371L, S373P, S375F, K417N, N440K, G446S, S477N, T478K, E484A, Q493R, G496S, Q498R, N501Y, Y505H e T547K. Isto porque, devido a limitações computacionais seria inviável simular toda a proteína Spike, pois não haveria como alocar todos os dados na memória da GPU, por isso nos limitamos apenas à região RBD. Ao longo do trabalho também foi simulado a versão cristalográfica para a variante Ômicron (PDB ID: 7T9L), isto é, as mutações não foram inseridas computacionalmente como antes mas obtidas experimentalmente. Todos os resultados de estabilidade e variação de afinidade frente a mutações da variante de Ômicron foram obtidos pelo módulo Residue Scanning [34] integrado ao software Schrodinger Maestro 2021-2. A decomposição de energia também foi realizada para o complexo ACE2-Spike (PDB ID: 7DF4), o complexo entre o anticorpo Fab 253H55L e a proteína Spike em estado inacessível (PDB ID: 7NDA) Os resultados do potencial eletrostático foram obtidos pelo algoritmo APBS [42] integrado ao software Schrodinger Maestro 2021-2.

4.1.2 *Dinâmica molecular com Desmond*

As simulações de dinâmica molecular foram realizadas usando o algoritmo Desmond [43] implementado no software Schrodinger Maestro 2021-2. Após a inserção prévia das mutações da variante Ômicron computacionalmente, o complexo a ser simulado correspondeu a interação da

região RBD da proteína Spike e o anticorpo neutralizante IgG (PDB ID: 7BZ5) [40]. Outros complexos também foram simulados utilizando a mesma metodologia descrita, como a variante Ômicron na interação anticorpo-antígeno mas agora obtida mediante dados experimentais (PDB ID: 7T9L), e também o complexo ACE2-RBD (PDB ID: 6M0J) onde as mutações da Ômicron tiveram que ser inseridas computacionalmente.

A caixa de solvatação foi descrita pelo modelo TIP3P [44] aplicando condições periódicas de contorno (PBC) em geometria cúbica. As dimensões foram definidas e minimizadas automaticamente pelo módulo System Builder com campo de força OPLS4 [39] onde a distância de buffer foi de $10\text{\AA} \times 10\text{\AA} \times 10\text{\AA}$, que representa a distância entre as bordas na extremidade da caixa de simulação e o solvente. Em resumo, significa que a caixa de solvatação ocupou todo o complexo proteico mais um volume adicional de $(10\text{\AA})^3$. Todo o sistema foi neutralizado mediante a adição dos íons Na^+ e Cl^- a uma concentração molar fisiológica de $0,15M$. Utilizamos o ensemble NPT (isotérmico-isobárico) nas simulações, onde a temperatura de $300,0K$ se manteve constante com o auxílio do termostato de Nosé-Hoover [45] em um tempo de relaxação de $1ps$. Enquanto isso, a pressão foi fixada em $1,01325Bar$ mediante o barostato de Martyna-Tobias-Klein [46] com um tempo de relaxamento de $2ps$ e acoplamento isotrópico. As interações eletrostáticas de longo alcance também foram estimadas pelo algoritmo Particle-Mesh-Ewald (PME) [47] com um valor de corte de $9,0\text{\AA}$. Antes do início, todo o sistema foi relaxado para que apenas depois fossem gerados os dados de trajetória. O tempo de simulação foi ao longo de $100ns$ com um tempo de integração de $2fs$ realizado pelo algoritmo RESPA [48] e cujos dados de trajetória foram gravados a cada $20ps$ totalizando $1000frames$ de simulação. Devido a constantes quedas de energia, a simulações foram divididas em 5 (cinco) partes de $20ns$, o que totalizou $100ns(5 \times 20ns)$, gerando também um total de $5000frames$. Após o término das simulações todas as partes foram unidas pelo script em Python chamado de "trj_merge.py". Por fim, todas as simulações de dinâmica molecular apenas foram possíveis mediante a aceleração gráfica de uma GPU Nvidia GTX 1050 2GB com 640 CUDA cores. Enquanto isso, apenas uma pequena parte dos cálculos foram encarregados a CPU de modelo Intel Core i3 6300 com frequência de $3,6Hz$, fazendo-se o uso de apenas 1 (um) núcleo do processador.

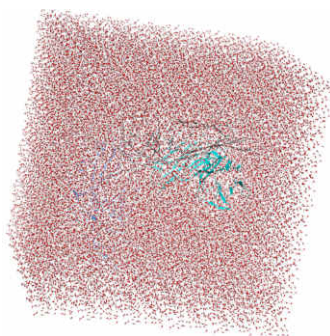


Figura 3: Caixa de solução criada com SystemBuilder no software Schrodinger Maestro 2021-2 onde o antígeno RBD com anticorpo IgG foi inserido (PDB ID: 7BZ5).

Após o término das simulações de dinâmica molecular, analisamos as interações químicas formada seja entre o anticorpo e seu antígeno, além da interação ACE2-RBD. Para isso, utilizamos apenas o último frame após transcorridos os $100ns$ de simulação. Sendo assim, novamente utilizamos o software Schrodinger Maestro 2021-2, mas agora adotamos a funcionalidade "Protein Interaction Analysis", onde o software pôde encontrar todas as interações químicas formadas em um raio máximo de $4,0\text{\AA}$. Além disso, a partir dos resultados de trajetória no formato .cms

realizamos as análises do Raio de Giro (R_g), RMSD e RMSF também no software Schrodinger Maestro 2021-2, onde adotamos sempre o Frame 0 como referência para os cálculos.

4.1.3 Maximização da interação anticorpo-antígeno

Usamos a funcionalidade "Affinity Maturation" integrada ao módulo "Residue Scanning" [34] implementado no Schrödinger Maestro 2021-2 [38], onde conseguimos realizar uma otimização Monte Carlo para encontrar as mutações que maximizam a afinidade de ligação entre o complexo anticorpo-antígeno para IgG B38 (PDB ID: 7BZ5) [40] e IgA1-Fc (PDB ID: 1OW0) [49]. A propriedade que foi otimizada foi a afinidade de interação. Escolhemos este método porque é computacionalmente viável, pois não precisamos testar as milhares de combinações possíveis de mutações. Para tornar as simulações viáveis computacionalmente, o solvente foi implicitamente considerado pelo método MM-GBSA no algoritmo Prime com campo de força OPLS4, onde este método também foi fundamental para calcular todos os valores de energia. Outra característica essencial é a maturação de afinidade, onde é feita uma busca para encontrar a afinidade de ligação ideal entre duas proteínas.

No entanto, para remover possíveis conflitos estéricos, realizamos uma minimização estrutural no algoritmo Prime nos aminoácidos próximos ao sítio de mutação dentro de um raio de corte de 5,0 angstrom. O total de combinações possíveis para os 194 aminoácidos na região RBD foi 4074 de mutações de um total de $(22 - 1)$ aminoácidos. A semente inicial foi zerada enquanto a temperatura em todas as simulações foi de 300,0K. O critério de aceitação para a próxima etapa ocorreu quando a variação de afinidade/estabilidade foi superior a 30,0. O processo foi repetido até que o número máximo de etapas fosse 200. Enquanto isso, o valor máximo para o número de mutações simultâneas foi definido para 3. Essas análises foram todas realizadas com um pH fisiológico de 7,4.

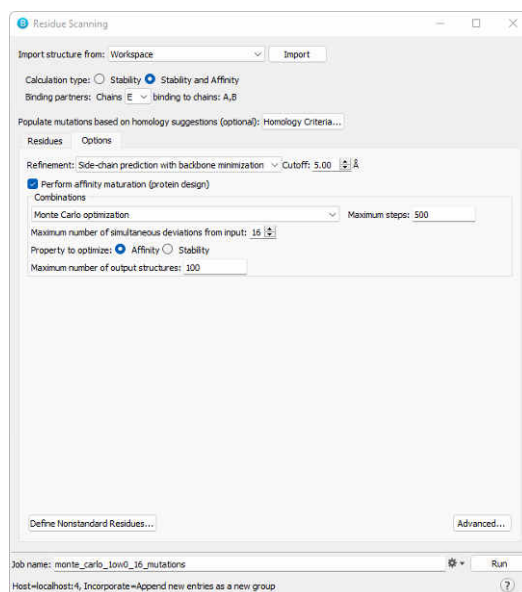


Figura 4: Visualização do painel de configurações para execução da maturação de afinidade anticorpo-antígeno onde é utilizado a otimização de Monte Carlo para tornar o algoritmo mais eficiente na busca pelas mutações que maximizam a afinidade do complexo proteico.

4.1.4 Neutralização da proteína Spike pelos anticorpos IgG, IgA e IgM

Nesta etapa, comparamos a neutralização dos anticorpos IgG, IgA e IgM em relação à proteína Spike. Em consequência dos anticorpos IgA não estarem em complexo com nenhuma proteína do SARS-CoV-2, usamos a cristalografia da forma monomérica do anticorpo IgA1-Fc (PDB ID: 1OW0) obtida isoladamente, sem estar em complexo com o antígeno, e assim conseguimos realizar o encaixe molecular entre anticorpo-antígeno para ser posteriormente submetido à dinâmica molecular. O encaixe também foi realizado com o dímero de anticorpo sIgA (PDB ID: 6UE7). O docking foi realizado utilizando o software Schrodinger Maestro 2021-2 utilizando o módulo "protein-protein docking" que utiliza o algoritmo PIPER [50]. O protocolo de ancoragem foi utilizando o modo "Standard" do software. Ele foi configurado para 70.000 rotações do anticorpo enquanto ele foi configurado para retornar 30 possíveis soluções de conformação, onde escolhemos aquela com a menor energia potencial. O encaixe com os mesmos protocolos também foi realizado para o anticorpo do tipo IgM (PDB ID: 2AGJ), como resultado da não disponibilidade de cristalografia com o complexo Spike. Por fim, realizamos simulações de docking molecular a partir do último quadro de dinâmica molecular em 100ns para quantificar a afinidade de interação no complexo anticorpo-antígeno. Utilizamos um total de 3 (três) ferramentas de encaixe: PatchDock [51], ClusPro [52] e HDock [53].

4.1.5 Simulações na interação de anticorpos com a proteína de Nucleocapsídeo (N)

A partir de cada uma das 3 (três) sequências de aminoácidos, NCoV-1, NCoV-2 e NCoV-3, foi possível modelar computacionalmente os diferentes domínios do Nucleocapsídeo (N) do SARS-CoV-2. Para tudo isso, a respectiva estrutura tridimensional foi modelada com base em técnicas de deep learning e modelagem de homologia na ferramenta RosettaFold por meio do servidor aberto Robetta (<https://robetta.bakerlab.org/>). Uma validação estrutural também foi realizada usando o diagrama de Ramachandran (ver Figura 5).

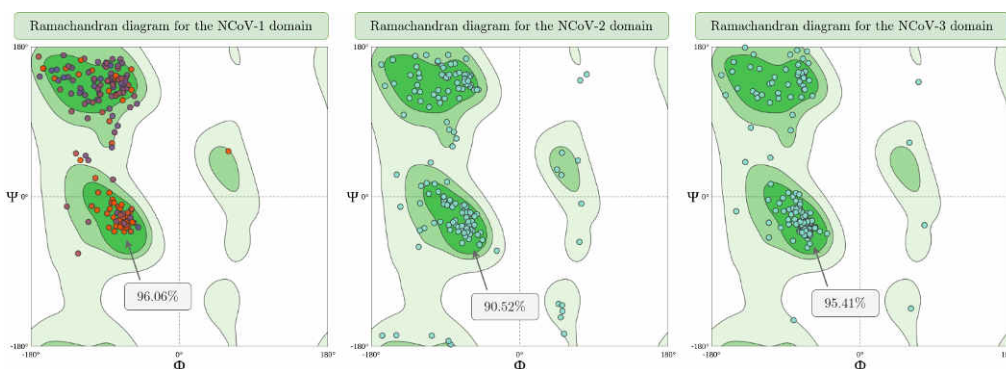


Figura 5: Validação estrutural de estruturas previstas no RosettaFold. O software utilizado para gerar o diagrama de Ramachandran foi o MolProbity incluído no Structure Assessment implementado na ferramenta online SWISS-Model.

A partir das estruturas tridimensionais obtidas anteriormente, o próximo passo foi realizar o encaixe molecular com monômero IgA (PDB ID: 1OW0), dímero IgA (PDB ID: 6UE7), IgG (PDB ID: 7BZ5) e IgM (PDB ID: 2AGJ). Desta forma, todas as estruturas cristalográficas dos anticorpos foram obtidas do banco de dados Protein Data Bank. É importante notar que o anticorpo IgG foi o único derivado do soro convalescente de pacientes acometidos por COVID-19. Para isso, foi utilizado o software Schrodinger Maestro 2021-2 com o módulo de docking proteína-

proteína com algoritmo PIPER [50] onde o ligante foi girado 70.000 vezes enquanto o número de possíveis conformações retornadas foi 30. Entre os resultados de saída, temos selecionou a primeira conformação, pois o software sempre retorna as soluções em ordem decrescente do tamanho do cluster estrutural.

4.1.6 *Descritores químico-quânticos*

Toda metodologia adotada para os cálculos de química quântica dos aminoácidos Lisina (K) e Treonina (T) referentes à mutação K417T presente na variante Ômicron e P.1, foi utilizando o software GaussView 6 para o desenho das estruturas tridimensionais, geração do arquivo de input e todas as respectivas análises de HOMO e LUMO. Enquanto isso, o Gaussian 09W [54] foi imprescindível para realização de todos os cálculos computacionais. Ambos os softwares, o Grupo de Química Teórica e Computacional (GQTC) da UFAM possui a licença acadêmica para utilizá-los.

Realizamos então a otimização da estrutura tridimensional da citosina, 5-metil-citosina, uracila e timina. Para isto, adotamos o *Job Type=Opt*, que otimizou a estrutura para o seu estado mínimo de energia, além de termos adotado um critério de convergência rígido para o término do cálculo de otimização ($\Delta Energy \leq 10^{-6} Ha$). O software foi configurado para buscar a estrutura otimizada em seu estado fundamental utilizando a teoria DFT por meio das bases gaussianas $6-311G++(2d, 2p)$ e funcional híbrido B3LYP. Por último, o número total de elétrons foi par, e portanto a otimização foi no estado de singlete onde todos os elétrons estiveram pareados.

5 Resultados e Discussões

5.1 SIMULAÇÕES DA VARIANTE ÔMICRON (B.1.1.529)

Nos resultados apresentados baixo, valores de $\Delta < 0$ denotam aumento no parâmetro analisado, enquanto que para $\Delta > 0$ indicam uma diminuição em relação ao valor inicial. Em termos de variação de afinidade (ver [Tabela 1](#)), a mutação mais preocupante foi a Q493R por induzir um aumento de $-12,916kcal/mol$ na interação ACE2-Spike. Em termos de variação de estabilidade, tivemos a mutação D796Y com aumento de $-9,371kcal/mol$.

Tabela 1: Resultados para variação de estabilidade e afinidade contra a variante Ômicron (B.1.1.529). O complexo analisado foi ACE2-Spike (PDB ID: 7DF4). O sinal positivo denota uma diminuição, enquanto o sinal negativo representa um aumento na propriedade analisada. Os termos de energia ΔE correspondem à estabilidade e não à afinidade. Isso ocorre porque o software sempre retorna dois tipos de ΔE , mas a afinidade é zero quando o aminoácido não está interagindo diretamente com o receptor ACE2.

Mutação	Δ Afinidade	Δ Estabilidade	ΔE_{vdW}	ΔE_{elec}	$\Delta G_{solvation}^{GB}$	$\Delta SASA_{total}$
A67V	0.000	-7.615	-3.878	+3.666	-1.233	-16.495
T95I	0.000	+0.114	+9.730	-1.232	-1.005	+6.559
G142D	0.000	+2.938	-5.784	-47.580	+22.736	-0.345
L212I	0.000	-3.126	+6.347	-2.513	-0.970	+0.132
G339D	+0.007	+3.102	-3.085	-28.702	-2.714	+10.575
S371L	-0.026	-1.926	-5.229	+8.146	-1.318	+15.142
S373P	+0.017	+4.626	+1.460	+28.460	+0.879	-12.429
S375F	+0.453	-2.076	-1.567	+7.862	+0.546	-5.635
K417N	+2.839	+5.941	-0.168	-57.181	+34.323	+2.389
N440K	-0.805	+4.894	+0.086	+75.551	-53.061	+37.128
G446S	+6.759	+3.971	+2.846	+6.046	-5.128	+17.030
S477N	-1.979	-2.060	+0.069	-39.271	-1.032	+7.085
T478K	-1.729	+9.863	-4.209	+36.456	-35.395	+37.358
E484A	+3.977	+11.016	+5.221	+36.161	-8.016	+0.067
Q493R	-12.916	-8.608	-5.263	+102.813	-84.100	-19.407
G496S	+5.239	-1.756	+5.255	-11.607	+7.376	-31.628
Q498R	-8.922	+23.339	+9.877	+109.482	-83.712	-4.115
N501Y	-2.567	+12.895	+12.979	+24.395	+3.097	-5.374
T547K	0.000	+4.798	-4.495	+6.689	-17.023	27.978
D614G	0.000	+2.003	+5.220	+3.139	+26.119	-32.204
H655Y	0.000	-0.999	+2.828	-2.905	+2.396	-42.956
N764K	0.000	0.313	-3.392	+40.872	-12.106	+3.528
D796Y	0.000	-9.371	-0.906	+34.054	-6.197	+19.024
N856K	0.000	+1.113	-2.976	+39.366	-12.153	-18.286
N969K	0.000	+7.910	-4.415	+4.241	-17.649	+37.732
L981F	0.000	35.053	+2.530	+18.291	-5.377	+47.936

Ao analisar o complexo anticorpo-antígeno (ver [Tabela 2](#)), nenhuma mutação apresentou diferença significativa na variação de afinidade. O maior valor obtido foi para N440K com diminuição de $+0,551kcal/mol$. Enquanto em termos de estabilidade apenas a mutação L981F induziu uma diminuição mais expressiva de $+15,713kcal/mol$.

Tabela 2: Resultados de estabilidade e variação de afinidade para variante de interesse Omicron. O complexo analisado estava entre o anticorpo neutralizante Fab 253H55L e a proteína Spike em sua conformação inacessível à interação (PDB ID: 7NDA), o que torna a simulação anticorpo-antígeno mais realista.

Mutação	Δ Afinidade	Δ Estabilidade	ΔE_{VDW}	ΔE_{elec}	$\Delta G_{solvation}^{GB}$	$\Delta SASA_{total}$
A67V	0.000	-8.813	-3.602	+6.236	+0.507	-19.054
T95I	0.000	-10.872	-3.055	-6.119	+5.872	-7.570
G142D	0.000	+0.042	-0.853	-7.744	-22.785	+38.691
L212I	0.000	+3.318	-1.026	+3.475	+1.503	+37.366
G339D	-6.824	+0.422	-2.219	-9.883	-13.713	-19.026
S371L	-1.800	-12.096	-3.124	-10.920	+2.931	-4.923
S373P	-0.406	+7.613	-0.144	26.509	+0.702	+6.602
S375F	-0.035	-9.098	+0.413	-1.785	+0.203	+9.415
K417N	-0.094	-11.186	+0.940	-55.511	+27.360	-11.117
N440K	+0.551	+6.594	-0.267	+72.860	-46.267	+61.157
G446S	+0.002	+3.253	+0.759	+9.573	-1.160	+29.960
S477N	0.000	-2.857	-0.302	-39.675	-0.487	+23.935
T478K	0.000	+2.789	-5.015	+28.982	-33.922	+10.030
E484A	0.000	+6.874	+4.861	+36.724	-8.396	-8.891
Q493R	+0.197	-6.776	-1.885	+66.954	-58.393	-1.184
G496S	0.000	+2.152	-0.762	+14.271	-3.192	+7.511
Q498R	+0.229	-4.265	-0.757	+73.436	-55.769	+31.264
N501Y	-0.001	-8.742	+3.483	+18.378	+3.936	+8.990
T547K	+0.019	-0.153	-2.368	+19.526	-32.996	-2.823
D614G	0.000	-2.995	+3.718	+4.366	+20.296	-16.673
H655Y	0.000	-5.842	+5.761	-8.862	+7.186	-0.756
N764K	0.000	+2.321	-5.125	+55.288	-23.588	+11.286
D796Y	0.000	-10.842	+2.983	+24.974	-6.447	+50.970
N856K	0.000	+7.546	-4.626	+53.158	-26.742	-30.146
N969K	0.000	+11.147	-4.072	+47.629	-14.885	+39.337
L981F	0.000	+15.713	+5.737	+2.577	-0.068	-21.768

Agora prosseguimos com a discussão dos resultados de dinâmica molecular da variante Ômicron, onde as mutações foram aplicadas computacionalmente. Notou-se que a interação com o anticorpo neutralizante IgG B38 não foi significativamente afetada, com relativamente poucas mudanças em RMSD, RMSF, Raio de giro e flutuações SASA. No entanto, no geral, mais desestabilização foi observada, algo que não foi apenas suportado pela análise RMSD (ver [Tabela 3](#) e [Figura 6](#)). A instabilidade se destacou em alguns resíduos bem específicos, o que tornou o valor médio de RMSF superior em relação à variante Ômicron. Em termos de compactação, a variante gerou uma expansão no complexo medida por um maior valor no raio de giro.

Tabela 3: Resultados de dinâmica molecular a 100 ns onde foram comparados a estrutura sem mutações e o complexo contendo as mutações da variante Ômicron. O complexo estava entre o anticorpo IgG e a região RBD da proteína Spike (PDB ID: 7BZ5).

Anticorpo	Antígeno	RMSD médio	RMSF médio	SASA médio	Rg médio
IgG B38	Wild-type	$(5.605 \pm 1.027)\text{Å}$	$(1.690 \pm 0.813)\text{Å}$	$(283.18 \pm 3.73)\text{nm}^2$	$(31.976 \pm 0.194)\text{Å}$
IgG B38	Ômicron	$(5.262 \pm 0.897)\text{Å}$	$(1.756 \pm 0.743)\text{Å}$	$(287.78 \pm 3.50)\text{nm}^2$	$(32.139 \pm 0.282)\text{Å}$

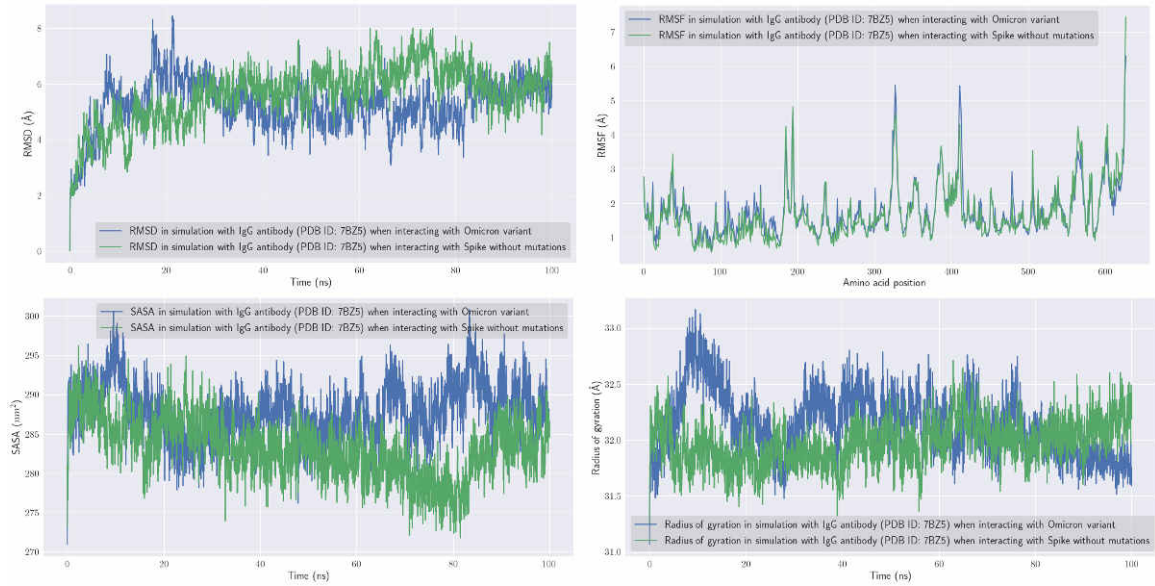


Figura 6: Comparação entre os valores de RMSD, RMSF, Raio de giro e SASA quanto ao impacto na estabilidade estrutural frente à variante Ômicron para o complexo anticorpo-antígeno (PDB ID: 7BZ5).

Em termos de análises de RMSD e RMSF, observou-se maior estabilização no complexo ACE2-RBD com a variante Ômicron (ver Tabela 4). Em termos de compactação, a variante gerou um maior raio de giro no complexo, o que denota uma pequena expansão ao longo da simulação.

Tabela 4: A dinâmica molecular em 100 ns ao comparar a variante Ômicron com a estrutura sem mutações. O complexo analisado foi ACE2-RBD (PDB ID: 6M0J).

Antígeno	RMSD médio	RMSF médio	SASA médio	Rg médio
Wild-type	$(2.935 \pm 0.430)\text{Å}$	$(1.599 \pm 0.939)\text{Å}$	$(352.77 \pm 3.57)\text{nm}^2$	$(31.132 \pm 0.207)\text{Å}$
Omicron	$(3.479 \pm 0.624)\text{Å}$	$(1.875 \pm 1.116)\text{Å}$	$(357.81 \pm 4.18)\text{nm}^2$	$(31.241 \pm 0.283)\text{Å}$

A partir dos resultados da decomposição de energia do MM-GBSA, podemos concluir que a variante Ômicron aumentou a afinidade de interação no complexo ACE2-RBD (ver Tabela 5). Isso pode ser uma indicação importante de que a entrada celular é favorecida contra a variante e, portanto, talvez haja maior transmissibilidade viral em comparação com as variantes anteriores.

Tabela 5: Resultados de decomposição de energia MM-GBSA para o último quadro de dinâmica molecular a 100 ns para o complexo ACE2-RBD (PDB ID: 6M0J). Todos os valores de energia estão em kcal/mol.

Antígeno	$\Delta G_{binding}$
Wild-type	-115.12
Ômicron	-119.97

Podemos visualizar na [Tabela 6](#) a perda de 2 (duas) pontes de hidrogênio na interação anticorpo-antígeno como consequência da variante Ômicron. Esta pode ser uma indicação importante da perda de eficácia dos anticorpos neutralizantes como resultado da variante Ômicron.

Tabela 6: Quantitativo das interações químicas formadas entre o anticorpo IgG (PDB ID: 7BZ5) e a região RBD da proteína Spike com a variante Ômicron e na ausência.

Anticorpo	Antígeno	Ligações de hidrogênio	Pontes salinas	Conflitos VdW	π -stacking
IgG B38	Wild-type	14	0	1	0
IgG B38	Ômicron	12	1	1	1

Podemos visualizar na [Tabela 7](#) as ligações de hidrogênio formadas na variante Ômicron do complexo anticorpo-antígeno, e que diante da variante Ômicron sofreu uma perda de algumas ligações, o que pode enfraquecer um pouco a neutralização do vírus.

Tabela 7: Ligações químicas formadas no complexo anticorpo-antígeno (PDB ID: 7BZ5) com mutações da variante Ômicron.

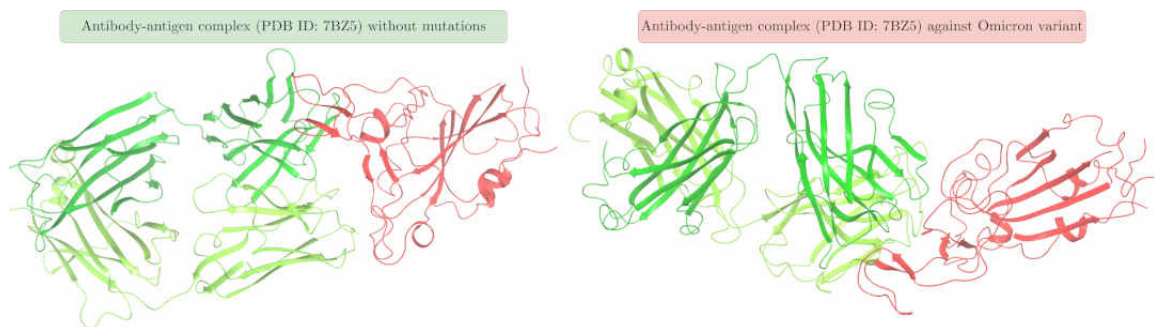
Ligação química
Spike:Asp420 → Antibody:Ser56 at 2.1Å
Spike:Arg457 → Antibody:Tyr33 at 1.9Å
Spike:Arg457 → Antibody:Ser53 at 1.8Å
Spike:Arg474 → Antibody:Ser31 at 2.0Å
Spike:Ala475 → Antibody:Arg97 at 2.1Å
Spike:Asn477 → Antibody:Gly26 at 2.2Å
Spike:Lys478 → Antibody:Glu1 at 2.1Å
Spike:Asn487 → Antibody:Arg97 at 2.0Å
Spike:Tyr489 → Antibody:Asp103 at 1.9Å
Spike:Arg498 → Antibody:Ser30 at 2.0Å
Spike:Gly502 → Antibody:Gln27 at 1.7Å
Spike:His505 → Antibody:Ile29 at 1.9Å

Por outro lado podemos visualizar na [Tabela 8](#) as ligações de hidrogênio formadas no complexo anticorpo-antígeno onde não há mutações na proteína Spike.

Tabela 8: *Ligações químicas formadas no complexo anticorpo-antígeno (PDB ID: 7BZ5) sem mutações.*

Ligação química
Spike:Asp405 → Antibody:Tyr94 at 2.0Å
Spike:Thr415 → Antibody:Ser56 at 1.9Å
Spike:Tyr431 → Antibody:Gly54 at 2.4Å
Spike:Tyr449 → Antibody:Ser31 at 1.8Å
Spike:Leu455 → Antibody:Tyr33 at 2.2Å
Spike:Arg457 → Antibody:Ser53 at 2.3Å
Spike:Tyr473 → Antibody:Ser31 at 1.7Å
Spike:Ala475 → Antibody:Ile28 at 2.1Å
Spike:Asn487 → Antibody:Gly26 at 1.9Å
Spike:Tyr489 → Antibody:Asn32 at 1.8Å
Spike:Gly496 → Antibody:Ser30 at 1.9Å
Spike:Gly502 → Antibody:Gly28 at 1.7Å
Spike:Tyr505 → Antibody:Gln90 at 2.4Å
Spike:Tyr505 → Antibody:Leu91 at 1.9Å

Também é possível termos uma visualização na [Figura 7](#) quanto às alterações estruturais em decorrência da variante Ômicron.

**Figura 7:** *Comparação da estrutura tridimensional entre a variante Ômicron e sem mutações para o complexo anticorpo-antígeno (PDB ID: 7BZ5) após 100 ns de simulação de dinâmica molecular.*

No complexo ACE2-RBD (ver [Figura 8](#)), podemos observar maiores mudanças conformacionais na ausência de mutações, o que leva à conclusão de que a variante Ômicron tornou a estrutura mais estável e rígida.

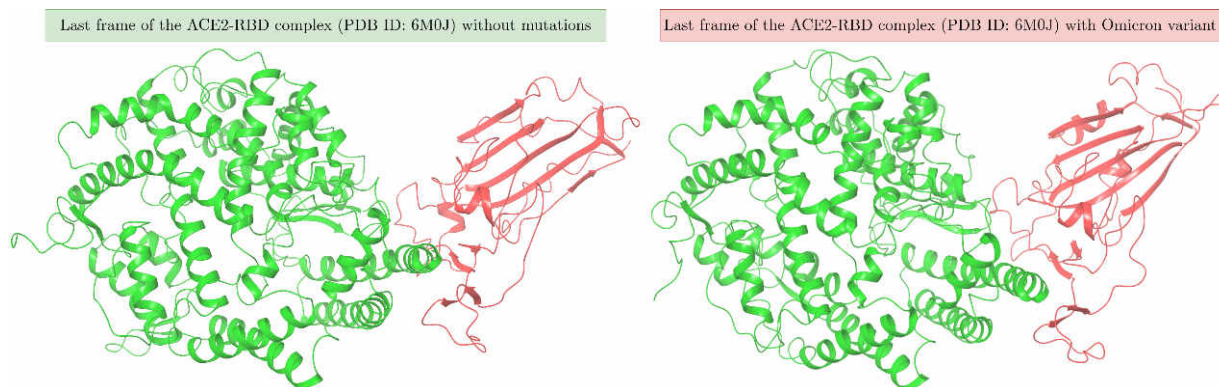


Figura 8: Comparação entre o último quadro de dinâmica molecular a 100 ns para o complexo ACE2-RBD (PDB ID: 6M0J).

No geral, podemos perceber na [Figura 9](#) de que a distribuição de carga contra a variante Ômicron tornou-se mais positiva em comparação com a estrutura wild-type. Ao analisar a carga elétrica líquida na variante Ômicron, observou-se uma distribuição de carga mais positiva em $+11eV$ para o complexo anticorpo-antígeno (PDB ID: 7BZ5), enquanto na ausência de mutações a carga foi de $+3eV$. No complexo ACE2-RBD (PDB ID: 6M0J) contra a Ômicron apresentou uma carga de $-18eV$, enquanto na ausência de mutações esse valor foi de $-21eV$. Portanto, reforçando a hipótese de que a variante torna a distribuição de carga mais positiva.

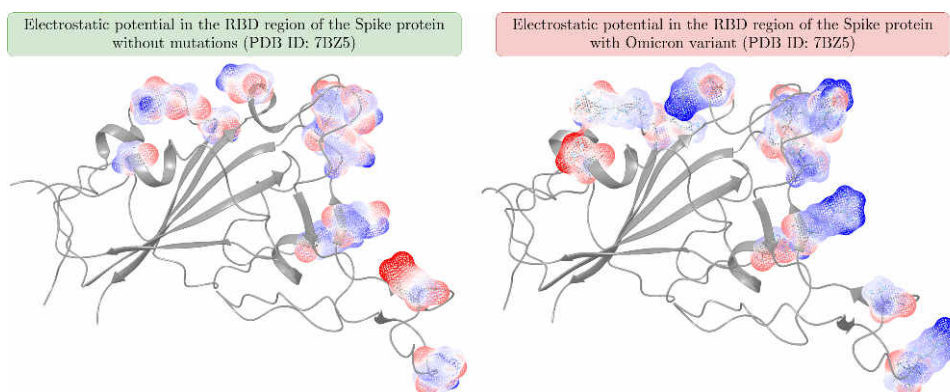


Figura 9: Potencial eletrostático de mutações que afetaram a região RBD da proteína Spike. A estrutura cristalográfica adotada foi o complexo entre o anticorpo IgG e a região RBD (PDB ID: 7BZ5).

Notou-se na [Figura 10](#) que o sítio de proteólise foi afetado pela variante Ômicron. Em outras palavras, houve uma região proteolítica formada pelos resíduos Phe338, Asp339 e Glu340, que antes não existiam, apareceu antes da variante emergente. Isso pode ser um indicativo de maior probabilidade de interação com o receptor ACE2, o que talvez possa refletir em maior transmissibilidade.



Figura 10: Em destaque em azul temos o sítio de proteólise conforme predição no software Schrodinger Maestro 2021-2. Uma comparação entre as regiões de proteólise na variante Ômicron e nas mutações ausentes na estrutura é apresentada. A estrutura cristalográfica utilizada foi a região RBD da proteína Spike (PDB ID: 7BZ5).

5.1.1 Dinâmica molecular para cristalografia da variante Ômicron

Podemos notar na [Figura 11](#) que o RMSD diante da variante Ômicron foi ligeiramente superior. Isto pode refletir menor estabilidade estrutural (ver [Figura 12](#)) em consequência da Ômicron. Em termos do raio de giro, a variante Ômicron novamente foi superior, novamente refletindo instabilidade estrutural. Quanto aos valores de exposição ao solvente SASA, nota-se também maiores valores diante da Ômicron. Isso é preocupante pelo menos a sub-variante BA.1 da Ômicron, pois um maior valor SASA favorece haver maiores regiões de interação com o receptor ACE2. Contudo, o ideal seríamos comparar os resultados aqui obtidos com a variante anterior, Delta. Infelizmente nos limitamos a comparar apenas com a cepa inicial surgida em Wuhan. Porém, foram publicados estudos pioneiros que mostraram que a variante Ômicron possui uma virulência menor comparada à variante Delta e P.1 [5].

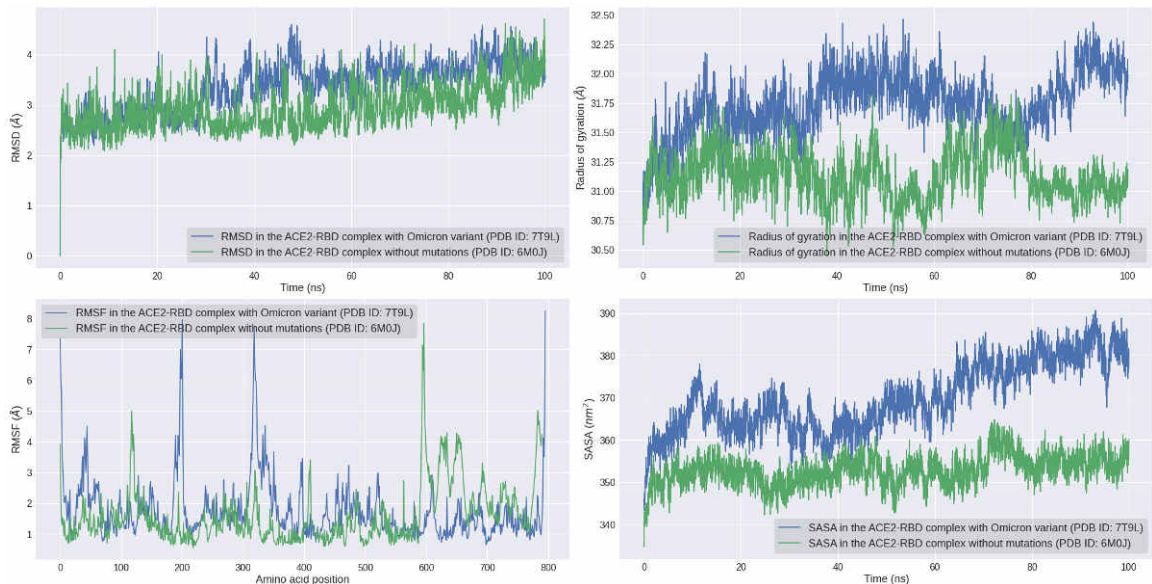


Figura 11: Resultados da dinâmica molecular onde é comparado a estrutura cristalográfica da variante Ômicron (PDB ID: 7T9L) e o complexo ACE2-RBD ausente de mutações (PDB ID: 6M0J).

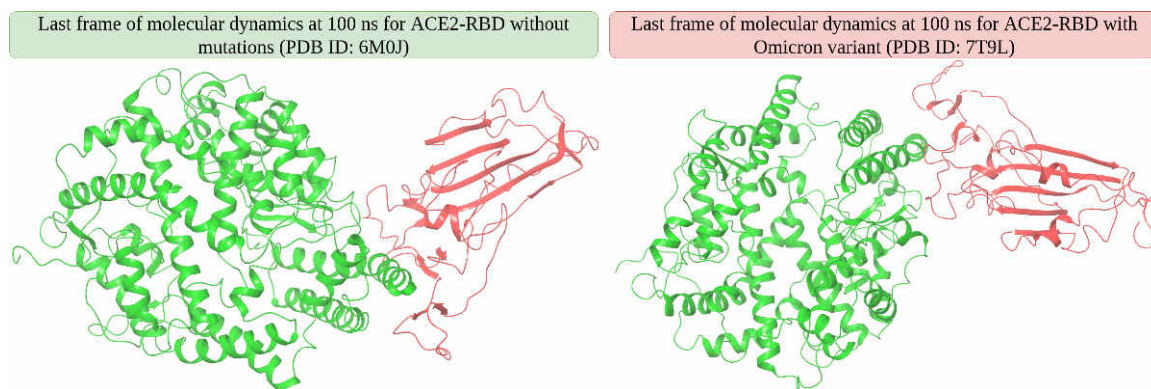


Figura 12: Comparativo estrutural entre o último frame da dinâmica molecular em 100 ns para o complexo ACE2-RBD sem mutações (PDB ID: 6M0J) e a cristalografia com variante Ômicron (PDB ID: 7T9L).

Por fim, podemos visualizar na [Tabela 9](#) os resultados obtidos para o complexo ACE2-RBD quanto ao impacto da variante Ômicron.

Tabela 9: Resultados de dinâmica molecular em 100 ns mostrando um comparativo entre a variante Ômicron e a estrutura ausente de mutações para o complexo ACE2-RBD (PDB ID: 7T9L).

Antígeno	RMSD médio	RMSF médio	SASA médio	Rg médio
Wild-type	$(2.935 \pm 0.430)\text{Å}$	$(1.599 \pm 0.939)\text{Å}$	$(352.77 \pm 3.57)\text{nm}^2$	$(31.132 \pm 0.207)\text{Å}$
Ômicron	$(3.351 \pm 0.490)\text{Å}$	$(1.742 \pm 0.970)\text{Å}$	$(369.87 \pm 8.23)\text{nm}^2$	$(31.742 \pm 0.262)\text{Å}$

Também buscamos realizar um comparativo a partir de simulações de dinâmica molecular entre as variantes P.1 e Ômicron. Nisto quantificamos a estabilidade estrutural do complexo ACE2-RBD (PDB ID: 6M0J) diante das duas variantes (ver [Figura 13](#)). Na literatura ainda não há consenso se o aumento de afinidade é diretamente proporcional à instabilidade ou não. De qualquer forma, consideramos esta hipótese como verdadeira na interpretação dos resultados. Nisto, percebeu-se que o valor de RMSD da P.1 foi inferior à Ômicron, o que pode se traduzir em menor afinidade pelo receptor ACE2. No entanto, foi justamente a variante P.1 que mais tirou vidas em Manaus, e a conclusão deveria ser uma maior afinidade pelo receptor. Independente dos resultados, sabemos que a dinâmica molecular por si só ainda não consegue explicar totalmente o comportamento viral, por haver outros fatores desde moleculares até ambientais.

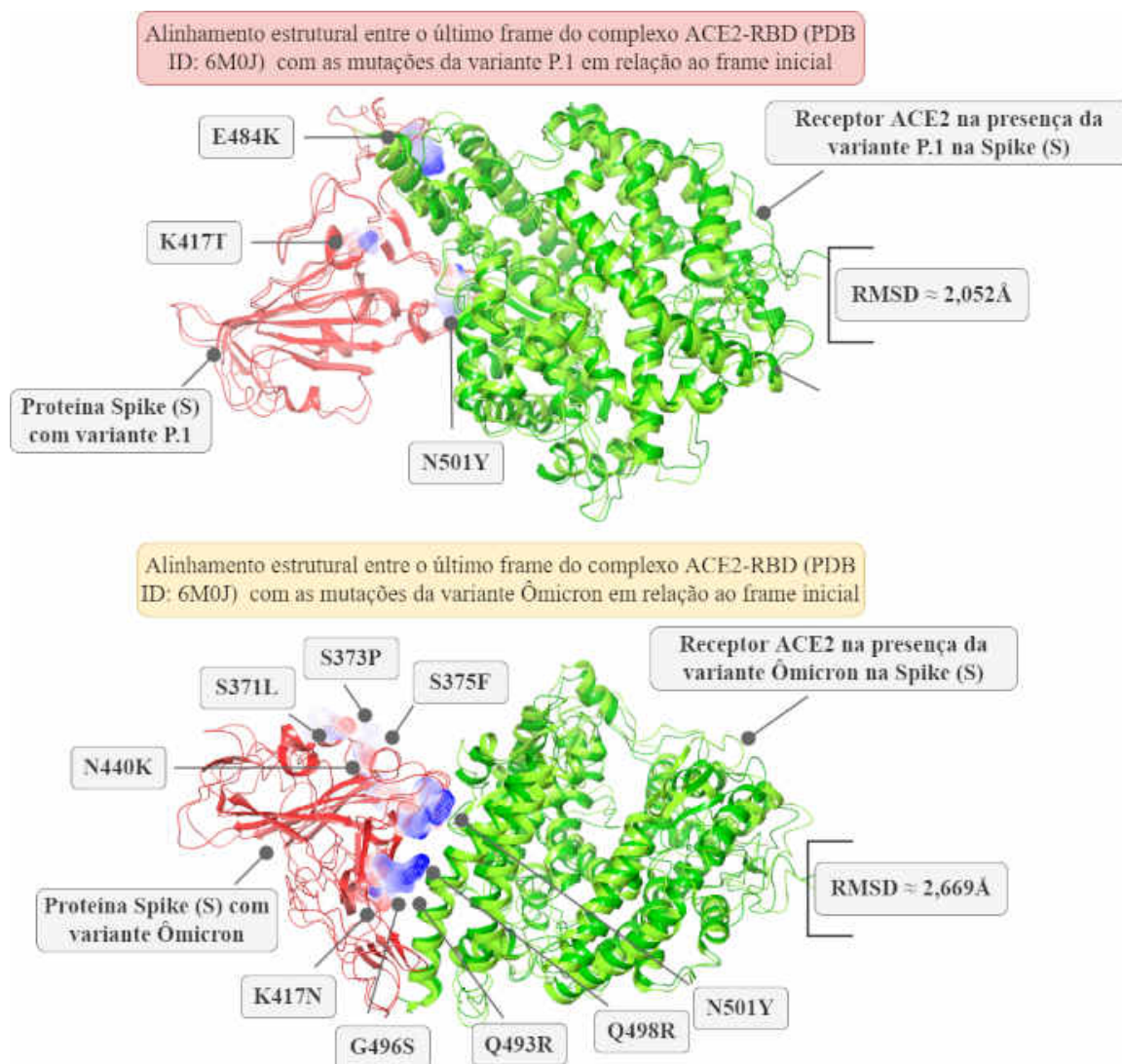


Figura 13: Alinhamento estrutural entre o último frame e o primeiro frame. Sendo que para chegarmos nessas estruturas realizamos resultado uma dinâmica molecular em 100 ns no módulo Desmond do software Schrodinger Maestro 2021-2. O complexo simulado foi o ACE2-RBD (PDB ID: 6M0J).

Ao analisar as mutações que constituem cada subvariante da Ômicron (ver ??), nota-se que determinadas mutações embora tenham desaparecido de algumas linhagens, elas ressurgiram nas linhagens seguintes. Em contrapartida, houveram mutações que se repetiram ao longo de todas as sublinhagens da Ômicron, e que portanto tendem a ser as mais preocupantes. Dentre estas, tivemos: G339D, S373P, S375F, K417N, T478K, E484A, Q498R, N501Y, Y505H, D614G, H655Y, N679K, P681H, N764K, D796Y, Q954H e N969K.

Algumas destas mutações estiveram previamente existentes em variantes de preocupações anteriores, o que levanta a questão de porque o vírus chegou a essas mutações específicas. Porém se de acordo com a literatura o vírus vem se tornando menos virulento [5], [6], e por isso na teoria algumas mutações deveriam desaparecer. Isto nos leva a concluir que o vírus embora tenha enfraquecido ao longo do tempo, mas foi principalmente o ser humano que diante da vacinação tornou o sistema imune mais preparado para enfrentar o vírus. A partir dos resultados não foi possível concluir se a variante Ômicron realmente dificulta a neutralização do antígeno

por anticorpos em comparação às variantes P.1 ou Delta, mas apenas em relação à cepa original de Wuhan, pois foram utilizados softwares diferentes nestas simulações, o que impossibilitou uma comparação justa.

De qualquer forma, esses resultados refletem uma pressão seletiva positiva, onde o vírus busca se beneficiar aumentando a estabilidade estrutural na interação ACE2-Spike. Um ponto curioso é que nem todas as mutações contribuem necessariamente para uma maior estabilidade ou afinidade na interação ACE2-Spike. Isso pode ser uma indicação do comportamento estocástico das mutações, embora a evolução natural favoreça algumas mutações em detrimento de outras.

Inicialmente organizamos de forma qualitativa as mutações que constituem cada uma das variantes de preocupação (ver [Tabela 10](#)). Esta etapa foi importante para entendermos quais mutações da variante P.1 permaneceram ao longo da evolução viral. Dentre as 12 (doze) que mutações que estiveram presentes na proteína Spike (S) da variante P.1, 8 (oito) foram exclusivas apenas da P.1. Consequentemente pelo vírus ter perdido grande partes das mutações presentes na P.1 ao evoluir para Delta e depois para Ômicron, podemos explicar a diminuição no número de hospitalizações e mortes, em grande parte, pela queda da virulência. Contudo, é imprescindível lembrar que o número cada vez mais de pessoas vacinadas certamente vem dificultando que os pacientes desenvolvam a forma mais grave da COVID-19, mesmo que algumas mutações críticas da P.1 ainda permaneçam até hoje.

Percebe-se também que na variante Ômicron houveram mutações reversas, isto é, que após terem surgido uma vez, desapareceram na linhagem seguinte, porém apareceram novamente. Além disso, podemos notar que grande parte das mutações da variante Ômicron afetaram justamente a região RBD, que está compreendida entre os aminoácidos Thr333 e Gly526. Sabemos que é justamente nesta região onde há maior interação com o receptor ACE2.

Uma questão central nesta pesquisa foi entender à nível molecular as causas do aumento de transmissibilidade e virulência do SARS-CoV-2. Conforme resultados encontrados na literatura, percebeu-se que na última variante, Ômicron, uma maior taxa de transmissão não necessariamente se refletiu em maior agressividade do vírus. Seja na literatura mas também a partir dos resultados obtidos ao longo desta pesquisa, realmente se percebe um aumento de afinidade na interação ACE2-Spike e dificuldade para neutralização de anticorpos. No entanto, ainda não sabemos a verdadeira causa de uma maior virulência.

Tabela 10: Comparativo das mutações que constituem cada uma das linhagem da Ômicron (BA.1 a BA.5), além das 4 (quatro) variantes de preocupação que a antecederam. A notação Δ representa a deleção de um aminoácido. As mutações mais críticas e que estiveram na variante P.1 ou Delta foram destacadas em vermelho, e que portanto podem estar associadas a maior virulência.

B.1.1.7	B.1.351	P.1	B.1.617.2	BA.1	BA.2	BA.2.12.1	BA.2.75	BA.4	BA.5
-	-	L18F	-	-	-	-	-	-	-
-	-	T19R	-	-	T19I	T19I	T19I	T19I	T19I
-	-	T20N	-	-	-	-	-	-	-
-	-	-	-	-	Δ L24	Δ L24	Δ L24	Δ L24	Δ L24
-	-	-	-	-	Δ P25	Δ P25	Δ P25	Δ P25	Δ P25
-	-	P26S	-	-	Δ P26	Δ P26	Δ P26	Δ P26	Δ P26
-	-	-	-	-	A27S	A27S	A27S	A27S	A27S
-	-	-	-	A67V	-	-	-	-	-
Δ H69	-	-	-	Δ H69	-	-	-	Δ H69	Δ H69
Δ V70	-	-	-	Δ V70	-	-	-	Δ V70	Δ V70
-	D80A	-	-	-	D80A	-	-	-	-
-	-	-	-	T95I	-	-	-	-	-
-	-	D138Y	-	-	-	-	-	-	-
-	-	-	G142D	Δ G142	G142D	G142D	G142D	G142D	G142D
-	-	-	-	Δ V143	-	-	-	-	-
Δ Y144	-	-	-	Δ Y144	-	Δ Y144	Δ Y144	Δ Y144	Δ Y144
-	-	-	-	Y145D	-	-	-	-	-
-	-	-	-	-	-	-	W152R	-	-
-	-	-	Δ E156	-	-	-	-	-	-
-	-	-	Δ F157	-	-	-	F157L	-	-
-	-	-	R168G	-	-	-	-	-	-
-	-	R190S	-	-	-	-	-	-	-
-	-	-	-	-	-	-	I210V	-	-
-	-	-	-	Δ N211	-	-	-	-	-
-	-	-	-	L212I	-	-	-	-	-
-	-	-	-	-	V213G	V213G	V213G	V213G	V213G
-	D215G	-	-	-	-	-	-	-	-
-	Δ L241	-	-	-	-	-	-	-	-
-	Δ L242	Δ L242	-	-	-	-	-	-	-
-	Δ A243	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	G257S	-	-
-	-	-	-	G339D	G339D	G339D	G339D	G339D	G339D
-	-	-	-	S371L	S371F	S371F	S371F	S371F	S371F
-	-	-	-	S373P	S373P	S373P	S373P	S373P	S373P
-	-	-	-	S375F	S375F	S375F	S375F	S375F	S375F
-	-	-	-	-	T376A	T376A	T376A	T376A	T376A
-	-	-	-	-	D405N	D405N	D405N	D405N	D405N
-	-	-	-	-	R408S	R408S	R408S	R408S	R408S
-	K417N	K417T	-	K417N	K417N	K417N	K417N	K417N	K417N
-	-	-	-	N440K	N440K	N440K	N440K	N440K	N440K
-	-	-	-	G446S	-	-	G446S	-	-
-	-	-	L452R	-	-	L452Q	-	L452R	L452R
-	-	-	-	S477N	S477N	S477N	S477N	S477N	S477N
-	-	-	T478K	T478K	T478K	T478K	T478K	T478K	T478K
-	E484K	E484K	-	E484A	E484A	E484A	E484A	E484A	E484A
-	-	-	-	-	-	-	F486V	F486V	F486V
-	-	-	-	Q493R	Q493R	Q493R	R494Q	-	-
-	-	-	-	G496S	-	-	-	-	-
-	-	-	-	Q498R	Q498R	Q498R	Q498R	498R	Q498R
N501Y	N501Y	N501Y	-	N501Y	N501Y	N501Y	N501Y	N501Y	N501Y
-	-	-	-	Y505H	Y505H	Y505H	Y505H	Y505H	Y505H
-	-	-	-	T547K	-	-	-	-	-
A570D	-	-	-	-	-	-	-	-	-
D614G	D614G	D614G	D614G	D614G	D614G	D614G	D614G	D614G	D614G
-	-	H655Y	-	H655Y	H655Y	H655Y	H655Y	H655Y	H655Y
-	-	-	-	N679K	N679K	N679K	N679K	N679K	N679K
P681H	-	-	P681R	P681H	P681H	P681H	P681H	P681H	P681H
-	A701V	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	S704L	-	-
T716I	-	-	-	-	-	-	-	-	-
-	-	-	-	N764K	N764K	N764K	N764K	N764K	N764K
-	-	-	-	D796Y	D796Y	D796Y	D796Y	D796Y	D796Y
-	-	-	-	N856K	-	-	-	-	-
-	-	-	D950N	-	-	-	-	-	-
-	-	-	-	Q954H	Q954H	Q954H	Q954H	Q954H	Q954H
-	-	-	-	N969K	N969K	N969K	N969K	N969K	N969K
-	-	-	-	L981F	-	-	-	-	-
S982A	-	-	-	-	-	-	-	-	-
-	-	D1027I	-	-	-	-	-	-	-
D1118H	-	-	-	-	-	-	-	-	-
-	-	V1176F	-	-	-	-	-	-	-

5.2 PROTEÍNA DE NUCLEOCAPSÍDEO (N) NOS DIAGNÓSTICOS MOLECULARES

Um dos grandes desafios das simulações teóricas é prever com precisão como o anticorpo se encaixa no antígeno. No entanto nem sempre teremos resultados que serão corroborados por estudos experimentais. Por outro lado, acreditamos que os resultados obtidos podem, de alguma forma, ajudar a entender as diferentes reatividades dos domínios do nucleocapsídeo. Na análise do domínio NCoV-2, dentre as flutuações de RMSF, destaca-se picos mais elevados na proteína nucleocapsídeo em comparação com o anticorpo neutralizante. Isso significa mudanças conformacionais substanciais no nucleocapsídeo enquanto o anticorpo IgG permaneceu mais rígido.

Tabela 11: Resultados para simulações de dinâmica molecular em 100ns no software Desmond para o complexo entre diferentes domínios do nucleocapsídeo do SARS-CoV-2 interagindo com anticorpos. São apresentados os valores médios de RMSD, RMSF, Raio de Giro e SASA.

Anticorpo	Domínio	RMSD médio	RMSF médio	SASA médio	Rg médio
IgG B38	NCov-1	$(4.072 \pm 0.257)\text{Å}$	$(1.845 \pm 1.232)\text{Å}$	$(299.54 \pm 5.25)\text{nm}^2$	$(26.886 \pm 0.127)\text{Å}$
IgG B38	NCov-2	$(14.969 \pm 2.480)\text{Å}$	$(4.129 \pm 3.853)\text{Å}$	$(328.96 \pm 8.76)\text{nm}^2$	$(29.596 \pm 0.563)\text{Å}$
IgG B38	NCov-3	$(23.179 \pm 7.057)\text{Å}$	$(6.503 \pm 10.566)\text{Å}$	$(334.47 \pm 7.12)\text{nm}^2$	$(34.982 \pm 1.196)\text{Å}$

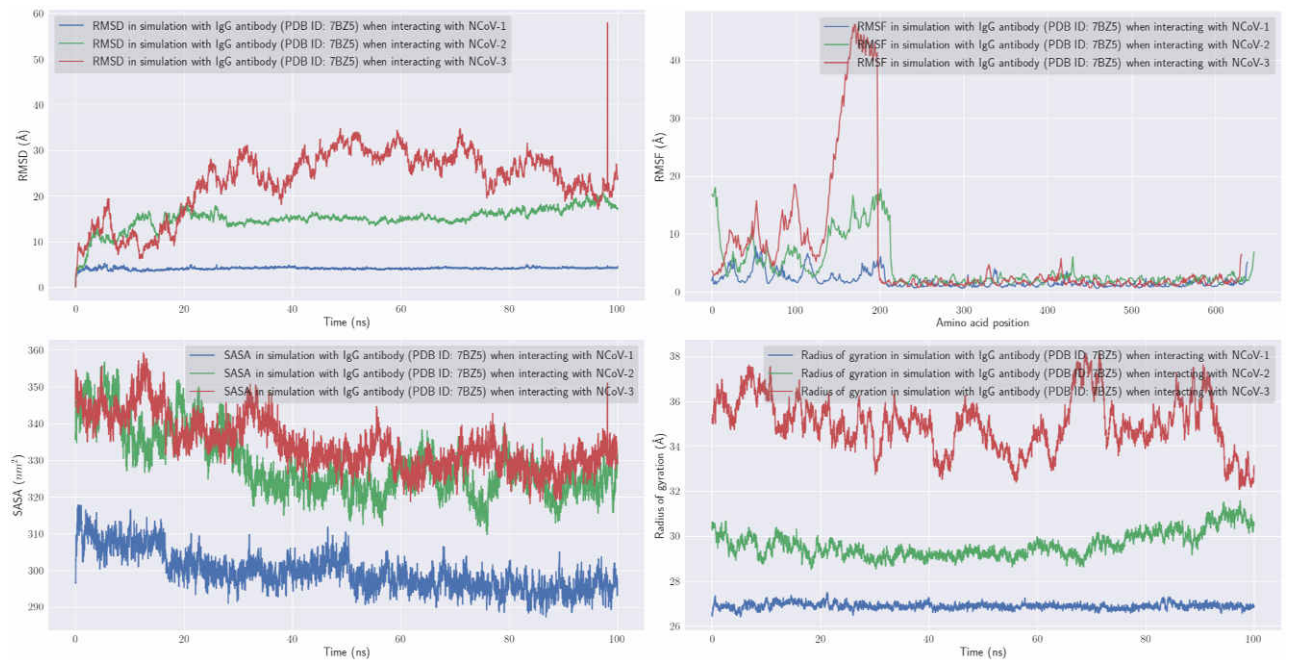


Figura 14: Análise comparativa de dinâmica molecular para diferentes domínios NCoV.

Dentre as interações químicas formadas entre o domínio NCoV-1 e o anticorpo IgG, tivemos um total de 15 ligações de hidrogênio e 2 pontes salinas. Por outro lado, ao analisar o domínio NCoV-2, houve uma diminuição para 10 ligações de hidrogênio e 1 ponte salina. Isso pode representar que a afinidade por anticorpos neste domínio é menor do que o domínio NCoV-1. Além disso, o comportamento desordenado do nucleocapsídeo foi mais notável no domínio NCoV-2 devido a flutuações significativas de RMSF.

Tabela 12: Análise quantitativa de ligações de hidrogênio, pontes salinas e conflitos de Van der Waals formados no último quadro do complexo anticorpo-antígeno após 100 ns de simulações de dinâmica molecular. Todas as análises foram quantificadas utilizando o módulo Protein Interaction Analysis presente no software Schrodinger Maestro 2021-2.

Anticorpo	Antígeno	Ligações de Hidrogênio	Ponte salinas	Conflitos VdW
IgG B38	NCoV-1	15	2	1
IgG B38	NCoV-2	10	1	0
IgG B38	NCoV-3	12	1	0

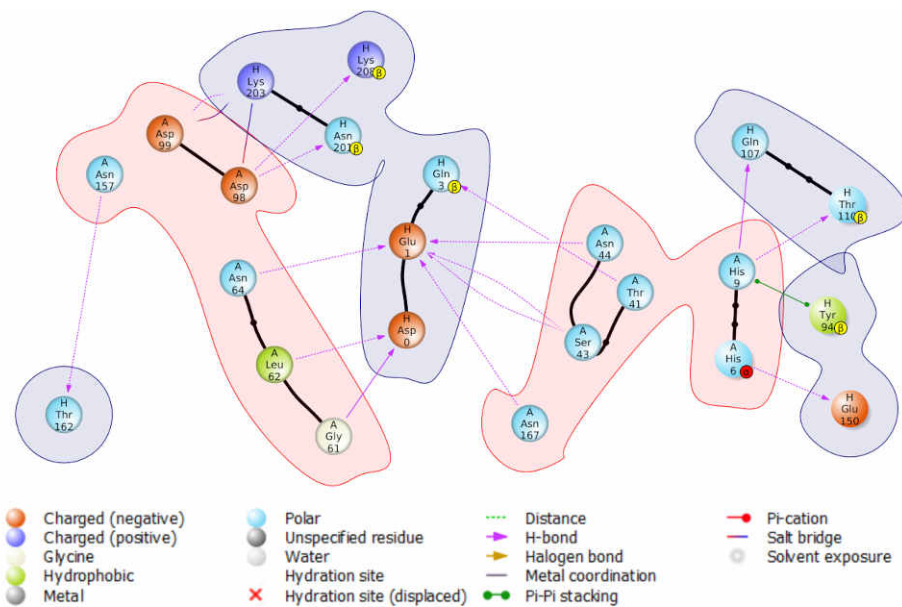


Figura 15: Diagrama mostrando as ligações químicas formadas entre o domínio NCoV-1 e o anticorpo IgG. A cadeia A representa o domínio do nucleocapsídeo, enquanto a cadeia pesada do anticorpo IgG é representada pela cadeia H. As setas em lilás representam as ligações de hidrogênio formadas. O diagrama foi construído a partir do último quadro de uma simulação de dinâmica molecular a 100 ns.

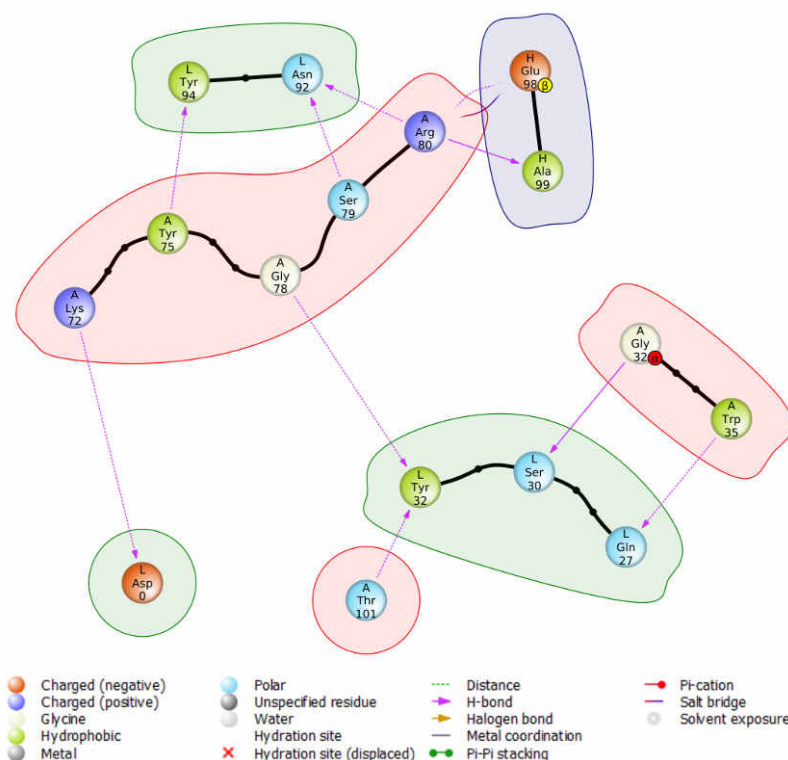


Figura 16: Diagrama de interações químicas formadas entre o domínio NCoV-2 e o anticorpo IgG (PDB ID: 7BZ5). A cadeia A representa a proteína Nucleocapsídeo, enquanto as cadeias H e L representam o anticorpo neutralizante IgG (PDB ID: 7BZ5). Esta imagem corresponde ao último quadro de uma simulação de dinâmica molecular a 100 ns.

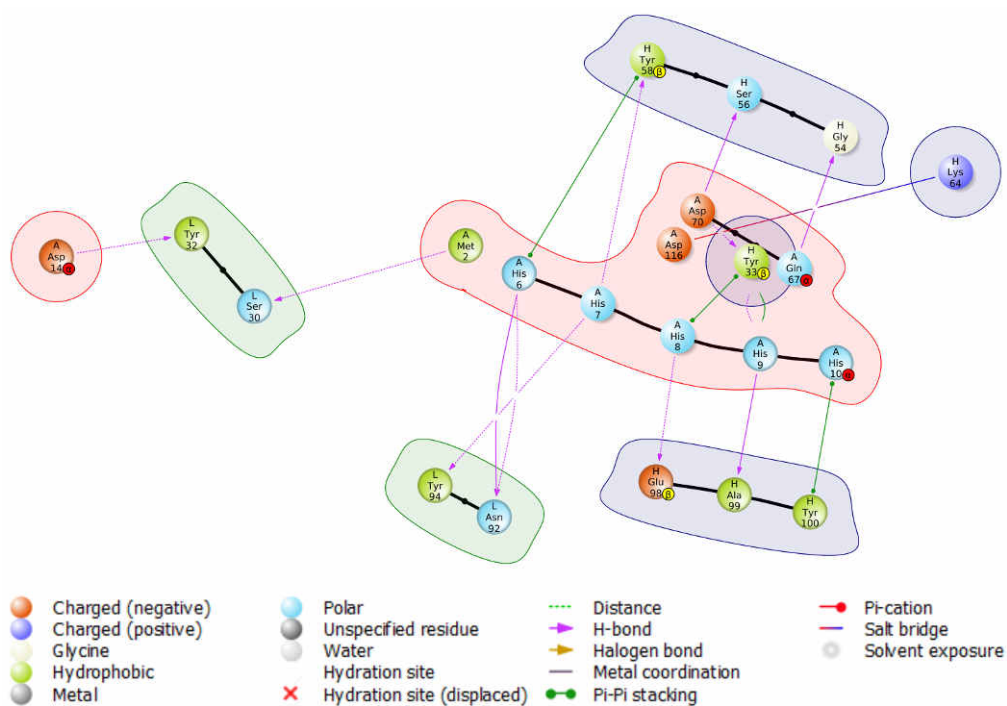


Figura 17: Diagrama das interações químicas formadas entre o domínio do nucleocapsídeo (NCov3) e o anticorpo IgG (PDB ID: 7BZ5). Esta imagem corresponde à última de uma simulação de dinâmica molecular a 100 ns.

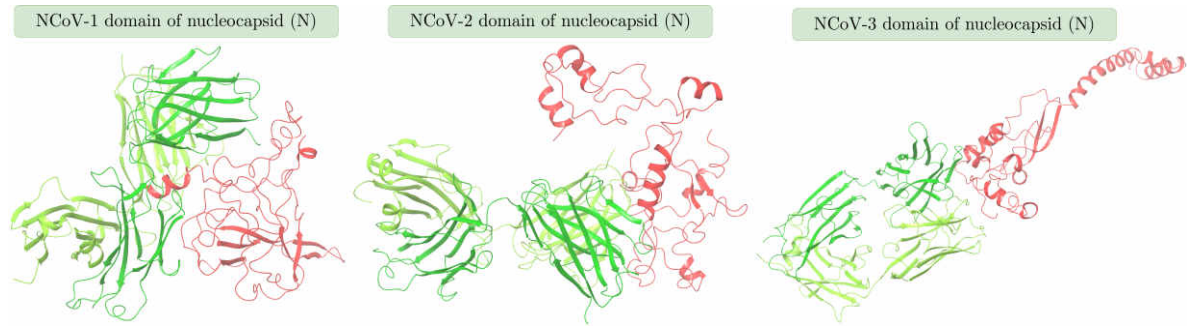


Figura 18: Visualização do último frame das simulações de dinâmica molecular a 100 ns para os 3 (três) domínios NCoV-1, NCoV-2 e NCoV-3 referentes à proteína nucleocapsídeo (N) do SARS-CoV-2.

5.3 DINÂMICA MOLECULAR DOS ANTICORPOS IGG, IGA E IGM

A partir dos resultados de dinâmica molecular, a estabilidade estrutural foi quantificada (ver [Tabela 13](#) e [Figura 19](#)) ao longo de 100 ns. Assim, apenas os resultados de RMSD e Raio de giro suportam a hipótese de maior estabilidade da IgA monomérica em relação à IgG. Enquanto os valores de RMSF e SASA negam essa hipótese. De qualquer forma, o docking molecular também foi realizado com 3 (três) ferramentas do último quadro para que possamos chegar a um consenso. Assim, é possível entender melhor os diferentes potenciais neutralizantes entre os tipos IgG e IgA na interação com o Spike. Notou-se que as flutuações no anticorpo dimérico IgA foram extremamente altas, o que deve ser um indicativo de altas alterações conformacionais devido à maior afinidade de interação com a proteína Spike.

Tabela 13: Resultados para simulações de dinâmica molecular a 100ns no software Desmond para o complexo entre a região RBD da proteína Spike e anticorpos. São apresentados os valores médios de RMSD, RMSF, Raio de giro e SASA.

Anticorpo	Antígeno	RMSD médio	RMSF médio	SASA médio	Rg médio
IgG B38	Wild-type	$(5.605 \pm 1.027)\text{Å}$	$(1.690 \pm 0.813)\text{Å}$	$(283.18 \pm 3.73)\text{nm}^2$	$(31.976 \pm 0.194)\text{Å}$
IgA1-Fc (monômero)	Wild-type	$(4.972 \pm 0.921)\text{Å}$	$(2.105 \pm 0.961)\text{Å}$	$(301.97 \pm 5.55)\text{nm}^2$	$(26.991 \pm 0.316)\text{Å}$
IgA (dímero)	Wild-type	$(31.697 \pm 5.168)\text{Å}$	$(6.706 \pm 5.456)\text{Å}$	$(546.09 \pm 4.28)\text{nm}^2$	$(43.890 \pm 0.991)\text{Å}$
IgM	Wild-type	$(7.152 \pm 1.540)\text{Å}$	$(3.585 \pm 1.314)\text{Å}$	$(301.93 \pm 4.17)\text{nm}^2$	$(33.086 \pm 0.566)\text{Å}$

Pode-se notar na [Figura 20](#) que é somente após uma determinada região de aminoácidos que as flutuações se tornam significativas no dímero do anticorpo IgA. Essa observação é notável ao longo das análises de RMSF, pois é possível visualizar a contribuição individual de cada aminoácido. Assim optamos pelo gráfico RMSF, pois parece ser o único que revela algo fundamental na interpretação do dímero de IgA. Os demais gráficos mostram apenas grandes picos de flutuações para o anticorpo IgA. Sendo assim, se compararmos apenas o primeiro conjunto (cadeia A e B) de aminoácidos do anticorpo dímero IgA em complexo com a região Spike RBD (cadeia E), vemos que as flutuações quase coincidem com o monômero IgA. No entanto, as flutuações começam a se tornar extremamente altas ao analisar precisamente no segundo conjunto de cadeias de anticorpos. Portanto, o dímero induz alterações conformacionais substanciais no segundo conjunto de cadeias (cadeia F e G) que compõem o anticorpo IgA. Assim, parece que a existência do dímero é a principal causa de maiores flutuações. O dímero conseqüentemente

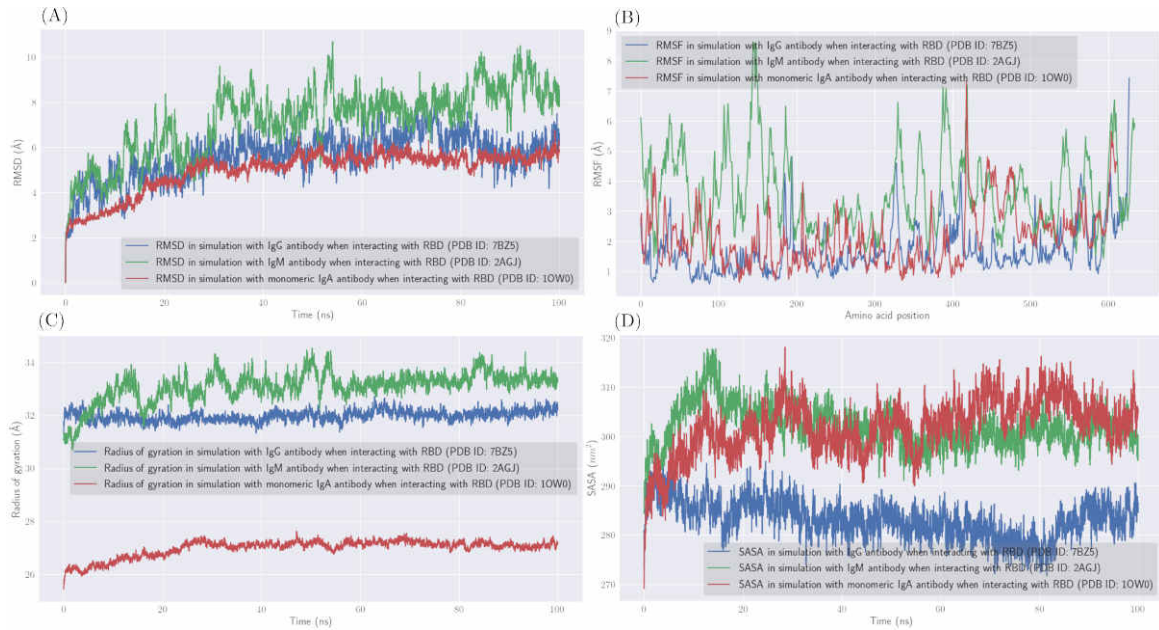


Figura 19: (A) Comparação dos valores de RMSD contra o complexo entre RBD e anticorpos IgA1 (monômero) e IgG B38. (B) Comparação das flutuações do RMSF. (C) Comparação de valores de raio de giração. (D) Comparação dos valores SASA.

parece levar ao aparecimento de uma conformação inexistente no monômero, talvez uma forma ainda mais relaxada do anticorpo IgA. Acima de tudo, a quantidade significativa de flutuações no dímero do anticorpo IgA representa maior flexibilidade e mobilidade estrutural ao longo da interação com a região RBD do vírus.

Ao quantificar as interações químicas formadas no complexo anticorpo-antígeno (ver [Figura 21](#)), poderemos entender qual tipo de anticorpo teve maior potencial para neutralizar a proteína Spike. Desta forma, foi possível quantificar um total de 9 (nove) ligações de hidrogênio formadas no complexo anticorpo-antígeno com o tipo IgA (monômero) referente ao último quadro da dinâmica molecular. Por outro lado, para o tipo IgG, foram formadas um total de 14 (catorze) ligações de hidrogênio. Esta é, portanto, uma indicação de maior seletividade e interação com o antígeno pelo tipo IgG. A forma dimérica do anticorpo IgA também foi analisada, com aumento substancial para 20 (vinte) ligações de hidrogênio entre a cadeia proteica Spike e as cadeias do anticorpo. Outro destaque é a formação de uma ponte salina entre o aminoácido Lys356 da proteína Spike e Glu422 pertencente ao anticorpo IgA a uma distância de 2,4 Å. Portanto, refletindo algo muito importante, que a forma dimérica do anticorpo IgA tem maior potencial de neutralização, em comparação com a forma monomérica, mas também com o tipo IgG.

Podemos notar que o valor de alinhamento RMSD (ver [Figura 22](#)) para o anticorpo IgG foi menor do que para o anticorpo IgA, indicando maior estabilidade estrutural no complexo anticorpo-antígeno. É importante destacar que valores mais favoráveis de energia livre na interação denotam maior estabilidade estrutural. No entanto, é muito complexo definir se maior estabilidade representa maior ou menor interação anticorpo-antígeno ou se o potencial neutralizante do anticorpo. Afinal, um sistema mais estável indica maior formação de ligações intermoleculares que potencializam a neutralização de anticorpos.

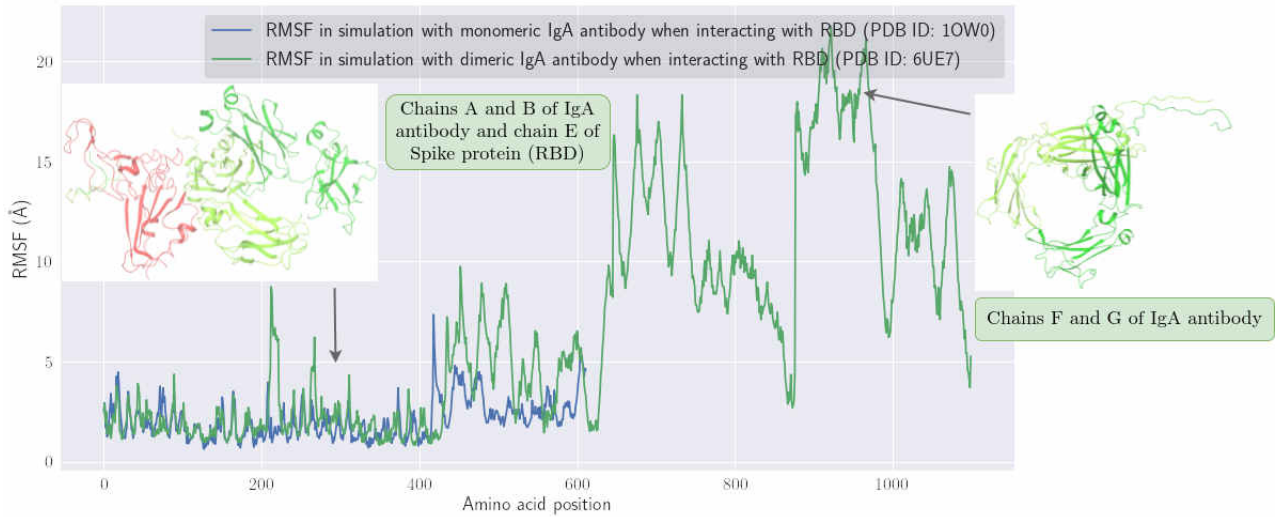


Figura 20: Gráfico RMSF para a forma dimérica do anticorpo IgA (PDB ID: 6UE7). Os instantâneos de estrutura correspondem ao último quadro da simulação de dinâmica molecular em 100 ns.

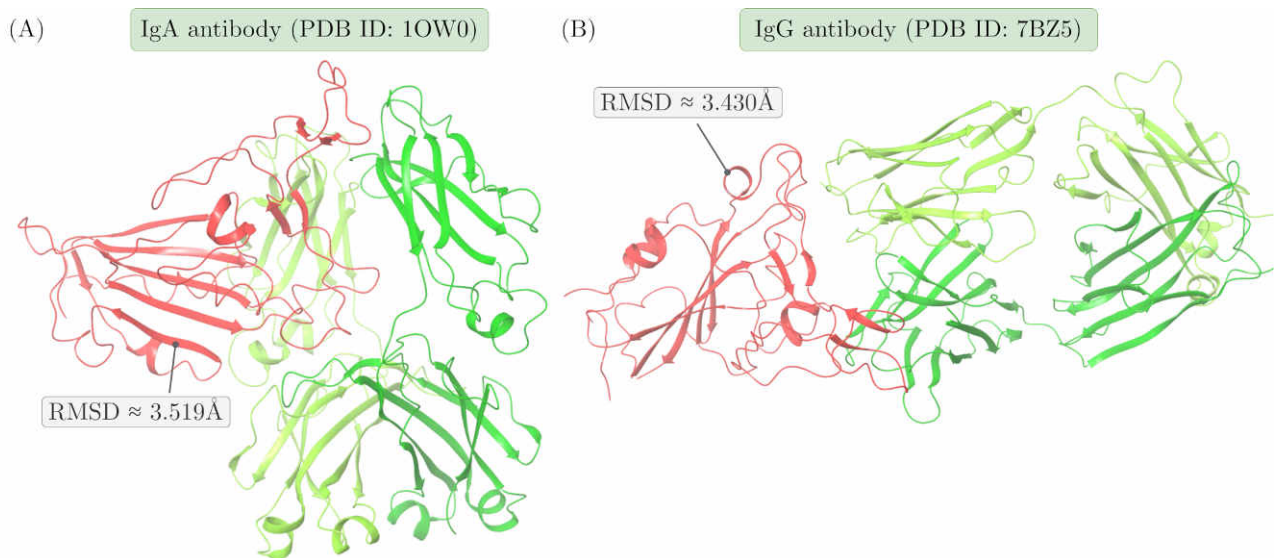


Figura 22: Visualização do último quadro de dinâmica molecular em 100ns do anticorpo IgA (PDB ID: 1OW0) e anticorpo IgG (PDB ID: 7BZ5) neutralizando a região RBD da proteína Spike. Além disso, é apresentado o resultado RMSD do alinhamento estrutural entre o último frame em relação ao primeiro.

Através da decomposição da energia MM-GBSA com o último quadro da dinâmica molecular (ver Tabela 14), notou-se que a forma dimérica do anticorpo IgA realmente forneceu valores mais favoráveis na energia livre de Gibbs. Em segundo lugar, temos a forma monomérica de IgA que apresentou resultados quase semelhantes ao tipo IgG. Por fim, tivemos o anticorpo IgM com potencial neutralizante extremamente baixo. Os resultados teóricos estão, portanto, em conformidade com o que foi relatado experimentalmente onde a IgA dimérica se destaca com um potencial de neutralização muito superior às demais [22]. É importante notar que esses valores de energia livre são relativos, por isso apenas podemos realizar uma comparação qualitativa entre os anticorpos consistente com os resultados experimentais.

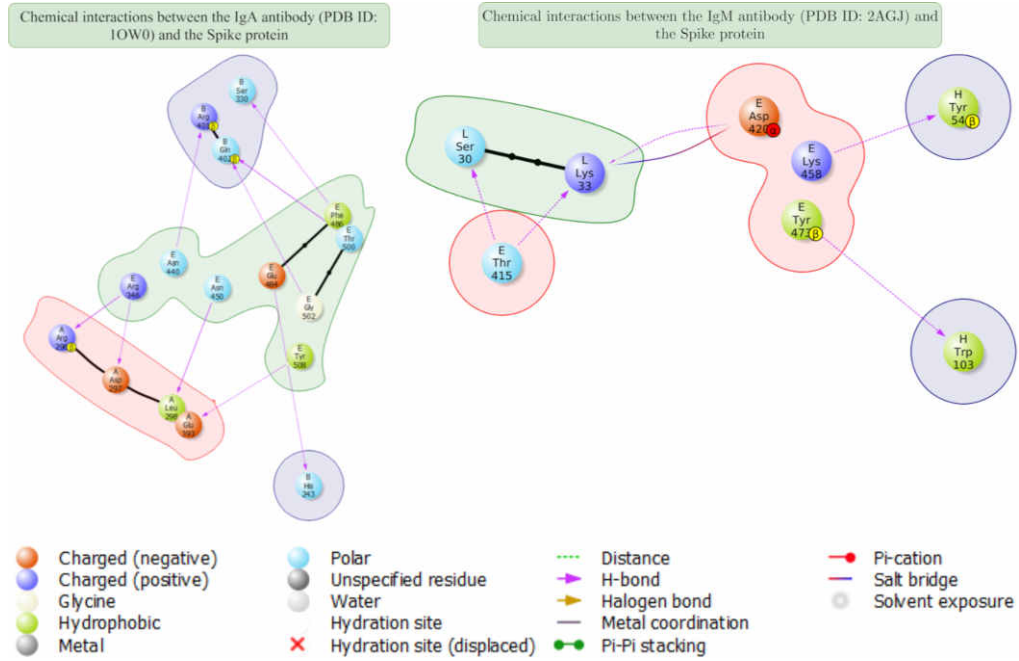


Figura 21: Diagrama de interações químicas formadas no último quadro de simulações de dinâmica molecular a 100 ns para complexo entre proteína Spike e anticorpo IgA (PDB ID: 1OW0) além do anticorpo IgM (PDB ID: 2AGJ). As setas em lilás representam as ligações de hidrogênio formadas.

Tabela 14: Resultados de decomposição de energia MM-GBSA para os complexos anticorpo-antígeno. Todos os resultados são correspondentes ao último quadro de simulações de dinâmica molecular em 100ns. Todas as unidades de energia estão em kcal/mol.

Anticorpo	Antígeno	$\Delta G_{binding}$
IgG B38	Wild-type	-105.06
IgA1-Fc (monomer)	Wild-type	-106.90
sIgA (dimer)	Wild-type	-184.27
IgM	Wild-type	-25.61

A interpretação baseada apenas na dinâmica molecular é muito abstrata e confusa, por isso realizamos docking molecular para ter uma ideia melhor da mudança de afinidade entre os diferentes anticorpos. A partir da análise de docking molecular (ver Tabela 15), o anticorpo IgG apresentou maior afinidade em 3 (três) ferramentas, Prodigy [55], HDock [53] e PatchDock [51]. Enquanto isso o tipo IgA (monômero) foi maior apenas na ferramenta ClusPro [52]. Esta é uma indicação de que o tipo IgG tem maior potencial neutralizante, mas o resultado se inverte totalmente ao compararmos com o IgA (dímero), em que absolutamente foi superior em todas as 4 (quatro) ferramentas adotadas. Isso é um importante indício de que a forma dimérica do anticorpo IgA é a principal resposta e também a primeira defesa no organismo humano diante da infecção pelo vírus SARS-CoV-2.

Tabela 15: *Docking molecular para o último quadro nas simulações de dinâmica molecular do complexo antígeno-anticorpo a 100ns. Todas as unidades de energia estão em kcal · mol⁻¹.*

Anticorpo	Prodigy	HDock	PatchDock	ClusPro
IgG B38	-14.6	-386.00	12616	-861.4
IgA1-Fc (monômero)	-10.7	-313.35	11218	-928.7
sIgA (dímero)	-17.8	-587.44	13598	-1458.9
IgM	-8.4	-244.94	12512	-768.2

5.4 SIMULAÇÕES DE QUÍMICA QUÂNTICA PARA O ENTENDIMENTO DO MECANISMO DE MUTAGÊNESE VIRAL

A partir da análise dos descritores químico-quânticos (ver [Tabela 16](#)), podemos observar que a transição de Lisina (K417) para Treonina (417T) proporciona maior afinidade eletrônica, potencial de ionização e eletronegatividade. Todos esses são fatores que podem contribuir para uma maior interação na interface ACE2-RBD. A mutação K417T vem ocorrendo tanto na mais recente variante Ômicron, mas também esteve presente na variante P.1/Gama que surgiu em Manaus. É importante destacar que esta variante foi a principal responsável pela grave 2^a onda de casos que sobrecarregou todo o sistema de saúde público nesta cidade.

Tabela 16: *Resultados dos descritores eletrônicos obtidos por otimização estrutural à nível da teoria DFT com as bases gaussianas 6-311G++(2d,2p) e função híbrida B3LYP. Todas as minimizações foram feitas usando o software Gaussian 09W.*

Descritor	Lisina	Treonina
Energia mínima	-13529.8117eV	-11930.5981
Momento de dipolo ($\vec{\mu}$)	0.902509Debye	4.266042Debye
E_{HOMO}	-6.6521eV	-7.4124eV
E_{LUMO}	-0.40218eV	-0.82831eV
E_{HOMO-1}	-7.0959eV	-8.1397eV
E_{LUMO+1}	-0.33633eV	-0.17959eV
ΔE	6.2499eV	6.5841eV
Afinidade eletrônica	0.40218eV	0.82831eV
Afinidade de ionização	6.6521eV	7.4124eV
Dureza química (η)	3.1250eV	3.2920eV
Maciez química (η^{-1})	0.16000eV	0.15188eV ⁻¹
Potencial químico (μ)	-3.5271eV	-4.1203eV
Índice de eletrofilicidade (ω)	1.9905eV	2.5785eV
Eletronegatividade (χ)	3.5271eV	4.1203eV
Índice de nucleofilicidade (N)	2.7165eV	1.9562eV

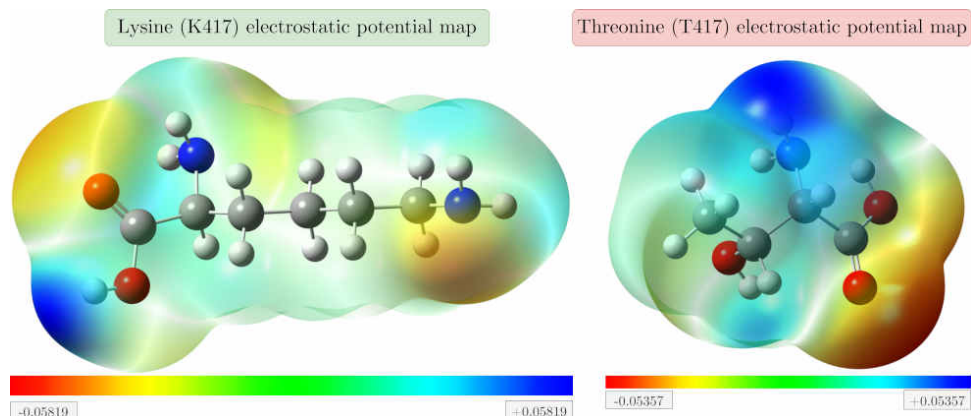


Figura 23: Mapas de potencial eletrostático (MEPS) de aminoácidos afetados pela variante Ômicron e P.1, que foram otimizados ao nível da teoria DFT com as bases gaussianas 6-311G (2d, 2p) com o funcional híbrido B3LYP no software Gaussian 09W. Todas as figuras foram geradas pelo software GaussView 6.

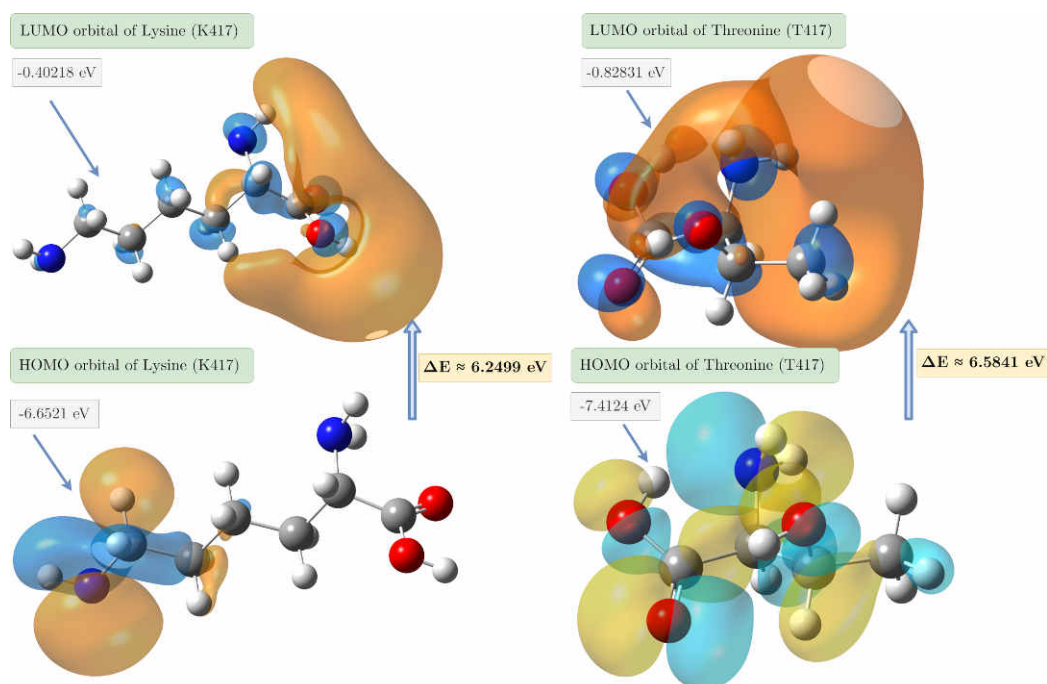


Figura 24: Comparação entre os orbitais de fronteira dos aminoácidos Lisina e Treonina que constituem a mutação K417T do SARS-CoV-2.

5.5 DESCOBERTA DE MUTAÇÕES QUE MAXIMIZAM A INTERAÇÃO ANTICORPO-ANTÍGENO

O número de mutações simultâneas pode afetar criticamente a aparência de certas mutações sobre outras. É, portanto, um problema de otimização onde devemos procurar o menor número de mutações biologicamente aceitáveis que maximizem a afinidade de interação anticorpo-antígeno. Assim, fazendo as modificações corretas no antígeno (ver Tabela 17, Tabela 18, Tabela 19 e Tabela 20), podemos melhorar a eficácia das vacinas atuais, por exemplo.

A partir do método de Monte Carlo encontramos um total de 16 (dezesseis) mutações no antígeno que resultaram em uma afinidade maximizada de $-61,656 \text{ kcal/mol}$ entre o anticorpo

IgG B38 e a proteína Spike (PDB ID: 7BZ5). Dentre estas, tivemos: G339Q, T345R, F347C, S349C, I358L, N360D, Y380R, G416R, S438F, L455R, F490L, Y508I, L518M, A520V e P527V. Por outro lado, em relação ao anticorpo IgA1-Fc (PDB ID: 1OW0) encontramos que a afinidade maximizada com o antígeno seria de $-82,343 \text{ kcal/mol}$. De modo que as mutações foram: T333V, N370M, S373E, S399R, I402Q, D405W, Y421S, N422H, Y449L, S469N, E471T, P491Y, S494P, Y495R, F497H e F515R.

Tabela 17: Maximização da afinidade para as combinações de mutações que maximizam a interação entre a proteína RBD e o anticorpo IgG B38 (PDB ID: 7BZ5). Todos os valores estão na unidade de kcal/mol.

Categoria	Mutações	Δ Afinidade	Δ Estabilidade	Δ Energia	Δ SASA _{total}
Afinidade maximizada	G339Q, T345R, F347C, S349C, I358L, N360D, Y380R, G416R, S438F, L455R, F490L, Y508I, L518M, A520V, P527V	-61.656	-55.479	-187.629	-171.800

Tabela 18: Maximização da afinidade para 16 (dezesesseis) mutações entre a proteína RBD e o anticorpo CB6 IgG (PDB ID: 7C01). Todos os valores estão na unidade de kcal/mol.

Categoria	Mutações	Δ Afinidade	Δ Estabilidade	Δ Energia	Δ SASA _{total}
Afinidade maximizada	L335A, F347Y, A352Y, K356E, V367I, T385H, V395I, I402G, G416Y, N437T, L441V, S443T, V445G, G476H, F490Y, Q506M	-31.442	-14.280	-6.490	-23.829

Tabela 19: Análise de afinidade para as 16 (dezesesseis) mutações que maximizam a interação entre a proteína RBD e o monômero do anticorpo IgA1-Fc (PDB ID: 1OW0). Todos os valores estão na unidade de kcal/mol.

Categoria	Mutações	Δ Afinidade	Δ Estabilidade	Δ Energia	Δ SASA _{total}
Afinidade maximizada	T333V, N370M, S373E, S399R, I402Q, D405W, Y421S, N422H, Y449L, S469N, E471T, P491Y, S494P, Y495R, F497H, F515R	-82.343	-62.944	-117.064	-170.163

Tabela 20: Maximização da afinidade para as 16 (dezesesseis) mutações entre a proteína RBD e o anticorpo IgM (PDB ID: 2AGJ). Todos os valores estão na unidade de kcal/mol.

Categoria	Mutações	Δ Afinidade	Δ Estabilidade	Δ Energia	Δ SASA _{total}
Afinidade maximizada	T333A, N343G, R346G, N354H, K356M, Y369H, S383Q, G416R, K424Y, T430N, S443V, V445T, G446V, N460H, Q493S	-39.014	-30.479	+64.941	-157.911

Acreditamos que para melhorar a interação dos anticorpos com o antígeno devemos adotar o valor de afinidade maximizado. Podemos notar que os anticorpos IgG e IgA apresentaram diferentes combinações de mutações que maximizam a afinidade de interação. Isso é um indicativo de seletividade, ou seja, um determinado conjunto de mutações apenas beneficiará a neutralização de um anticorpo em detrimento de outro. De qualquer forma, ter realizado simulações para ambos os tipos de anticorpos poderia contribuir para uma resposta mais seletiva dependendo das mutações escolhidas para o antígeno.

Por fim, é importante notar que as ferramentas que predizem $\Delta\Delta G$ são baseadas apenas em uma estrutura estática, a partir da qual a variação de energia é calculada como resultado da mutação. Embora seja promissor usar o método de Monte Carlo, idealmente, para resultados mais consistentes, o número de iterações deve ser o maior possível. Outra abordagem pioneira no método de Monte Carlo seria aplicar mutações em anticorpos que interagem com a proteína Spike. Isso seria importante para aumentar a sensibilidade dos diagnósticos, mas também para aumentar a eficácia dos anticorpos neutralizantes.

Ao aumentarmos o número de mutações simultâneas no antígeno também aumenta-se proporcionalmente a complexidade computacional. Por isso, o número de mutações simultâneas adotado se deve simplesmente a limitações computacionais do hardware utilizado. De qualquer forma, ainda que tivéssemos capacidade computacional o suficiente, o nosso grande receio é de que ao haver muitas mutações simultâneas, as células da memória imune não fossem mais capazes de reconhecer uma infecção real pelo vírus SARS-CoV-2. Curiosamente mesmo aumentando o número de aumentações simultâneas, isso não necessariamente se refletiu em maiores valores de afinidade anticorpo-antígeno. Portanto, a quantidade de mutações simultâneas adotada foi justamente aquela que se traduziu em maiores valores de afinidade. Ao longo da pesquisa buscou-se simplesmente por tentativa e erro ajustar o número de mutações simultâneas que acarretassem no maior valor possível de afinidade anticorpo-antígeno.

Dentre todos os resultados obtidos ao longo desta pesquisa, a aplicação do método de Monte Carlo ao SARS-CoV-2 é certamente o mais promissor e importante que obtivemos até então. Nas otimizações de Monte Carlo na busca de mutações que maximizem a afinidade anticorpo-antígeno, ficamos extremamente preocupado quanto à quantidade de mutações simultâneas. Infelizmente não obtivemos sucesso em prever novas mutações do vírus SARS-CoV-2. Isto porque partimos de uma hipótese incorreta de que a evolução viral necessariamente seria direcionada para uma maximização da afinidade ACE2-Spike. Contudo, aparentemente a evolução do SARS-CoV-2 não segue um padrão observável, e que portanto parece ser realmente cego. No entanto, houve um importante artigo publicado na Science [56] que a partir da análise por inteligência artificial de mais 6 milhões de sequenciamentos do vírus pôde então prever uma subvariante da Ômicron. Isso portanto é o ponto de partida para prever a sequência de um novo coronavírus que ainda virá a existir. Apesar de tudo, ainda não conseguimos encontrar formas de implementar essa ideia.

6 Considerações finais

- A variante Ômicron causou o surgimento de um novo sítio proteolítico, que pode favorecer a interação ACE2-Spike em relação à cepa de Wuhan.
- Outro resultado notável foi de que diante da variante Ômicron houveram a perda de 2 (duas) ligações de hidrogênio na interação anticorpo-antígeno, o que pode enfraquecer a resposta de anticorpos neutralizantes.
- O anticorpo IgA em sua forma dimérica mostrou-se ter o maior potencial de neutralização da região RBD do SARS-CoV-2, comparado a sua forma monomérica, além dos tipos IgM e IgG.
- Ao longo da pesquisa percebemos que transição do aminoácido Lisina (K417) para Treonina (417T) gerou um aumento em diversos descritores químico-quânticos, como afinidade eletrônica, potencial de ionização e eletronegatividade. Isto pode ser um indício do porquê a mutação K417T esteve presente na variante P.1 que assolou Manaus.
- O domínio NCoV-2 da proteína de nucleocapsídeo (N) mostrou-se ter grande potencial de interação com anticorpos em comparação à proteína Spike (S) do vírus SARS-CoV-2. Isto pode ser um importante indício de que diagnósticos do tipo antígeno para COVID-19 terão maior especificidade se utilizarem a proteína (N).
- Obtemos no decorrer da pesquisa que as mutações G339Q, T345R, F347C, S349C, I358L, N360D, Y380R, G416R, S438F, L455R, F490L, Y508I, L518M, A520V e P527V no antígeno resultaram em uma afinidade maximizada de $-61,656\text{kcal/mol}$ para o anticorpo IgG B38.
- Em relação ao anticorpo IgA1-Fc encontramos as mutações T333V, N370M, S373E, S399R, I402Q, D405W, Y421S, N422H, Y449L, S469N, E471T, P491Y, S494P, Y495R, F497H e F515R que maximizam a afinidade com o antígeno em $-82,343\text{kcal/mol}$.
- Por fim, ao comparar as mutações que constituem a variante Ômicron com outras VOC anteriores, como P.1 e Delta, percebeu-se que algumas mutações permaneceram ao longo do processo evolutivo. Consequentemente sugerimos a hipótese de que embora o vírus SARS-CoV-2 tenha se tornado menos virulento, mas foi devido principalmente às crescentes campanhas de vacinação que tornou o sistema imune do ser humano mais forte e preparado diante da infecção pela COVID-19.
- Apesar de tudo, ainda seria imprescindível confirmar estes resultados em ensaios experimentais, para que apenas assim possamos levar estes resultados adiante na esperança de realmente ajudar as pessoas.

Bibliografia

- [1] C. De Souza, M. Oliveira, J. A. Bessa, K. Mota e R. De Freitas, «3D structural prediction, analysis and validation of SARS-CoV-2 protein molecules», *CLEI Electron. J.*, vol. 25, n.º 2, mai. de 2022.
- [2] E. Tiesinga, P. J. Mohr, D. B. Newell e B. N. Taylor, «CODATA recommended values of the fundamental physical constants: 2018», vol. 93, n.º 2, jun. de 2021. DOI: [10.1103/revmodphys.93.025010](https://doi.org/10.1103/revmodphys.93.025010).
- [3] F. G. Naveca, V. Nascimento, V. C. de Souza et al., «COVID-19 in Amazonas, Brazil, was driven by the persistence of endemic lineages and P.1 emergence», en, *Nat. Med.*, vol. 27, n.º 7, pp. 1230–1238, jul. de 2021.
- [4] N. R. Faria, T. A. Mellan, C. Whittaker et al., «Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil», en, *Science*, vol. 372, n.º 6544, pp. 815–821, mai. de 2021.
- [5] R. Suzuki, D. Yamasoba, I. Kimura et al., «Attenuated fusogenicity and pathogenicity of SARS-CoV-2 Omicron variant», en, *Nature*, vol. 603, n.º 7902, pp. 700–705, mar. de 2022.
- [6] T. Nyberg, N. M. Ferguson, S. G. Nash et al., «Comparative analysis of the risks of hospitalisation and death associated with SARS-CoV-2 omicron (B.1.1.529) and delta (B.1.617.2) variants in England: a cohort study», en, *Lancet*, vol. 399, n.º 10332, pp. 1303–1312, abr. de 2022.
- [7] H. Rossman, S. Shilo, T. Meir, M. Gorfine, U. Shalit e E. Segal, «COVID-19 dynamics after a national immunization program in Israel», en, *Nat. Med.*, vol. 27, n.º 6, pp. 1055–1061, jun. de 2021.
- [8] R. Arya, S. Kumari, B. Pandey et al., «Structural insights into SARS-CoV-2 proteins», *J. Mol. Biol.*, vol. 433, n.º 2, p. 166 725, jan. de 2021.
- [9] P. D. Burbelo, F. X. Riedo, C. Morishima et al., «Sensitivity in Detection of Antibodies to Nucleocapsid and Spike Proteins of Severe Acute Respiratory Syndrome Coronavirus 2 in Patients With Coronavirus Disease 2019», *The Journal of Infectious Diseases*, vol. 222, n.º 2, pp. 206–213, mai. de 2020, ISSN: 0022-1899. DOI: [10.1093/infdis/jiaa273](https://doi.org/10.1093/infdis/jiaa273).
- [10] J. R. C. Pulliam, C. van Schalkwyk, N. Govender et al., «Increased risk of SARS-CoV-2 reinfection associated with emergence of Omicron in South Africa», *Science*, vol. 376, n.º 6593, mai. de 2022.
- [11] L. Ferrante, L. H. Duczmal, E. Capanema et al., «Dynamics of COVID-19 in Amazonia: A history of government denialism and the risk of a third wave», *Prev. Med. Rep.*, vol. 26, n.º 101752, p. 101 752, abr. de 2022.
- [12] *SARS-CoV-2 variants of concern and variants under investigation: Technical Briefing 29*, Accessed in 06 December, 2021. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1036501/Technical_Briefing_29_published_26_November_2021.pdf.

- [13] C. S. Lupala, Y. Ye, H. Chen, X. Su e H. Liu, «Mutations in RBD of SARS-CoV-2 Omicron variant result stronger binding to human ACE2 protein», *bioRxiv*, 2021. DOI: [10.1101/2021.12.10.472102](https://doi.org/10.1101/2021.12.10.472102).
- [14] P. Ranjan, Neha, C. Devi, K. A. Devar e P. Das, «The influence of new SARS-CoV-2 variant Omicron (B.1.1.529) on vaccine efficacy, its correlation to Delta Variants: a computational approach», *bioRxiv*, 2021. DOI: [10.1101/2021.12.06.471215](https://doi.org/10.1101/2021.12.06.471215).
- [15] J. Flint, V. R. Racaniello, G. F. Rall e A. M. Shalka, *Principles of Virology*, 4^a ed. American Society for Microbiology, 2015, vol. 2, ISBN: 978-1-55581-934-7.
- [16] Q. Wang, Y. Guo, S. Iketani et al., «Antibody evasion by SARS-CoV-2 Omicron subvariants BA.2.12.1, BA.4 and BA.5», *Nature*, vol. 608, n.º 7923, pp. 603–608, ago. de 2022.
- [17] R. Matyášek e A. Kovařík, «Mutation Patterns of Human SARS-CoV-2 and Bat RaTG13 Coronavirus Genomes Are Strongly Biased Towards C>U Transitions, Indicating Rapid Evolution in Their Hosts», *Genes*, vol. 11, n.º 7, 2020, ISSN: 2073-4425. DOI: [10.3390/genes11070761](https://doi.org/10.3390/genes11070761).
- [18] L. D. Moore, T. Le e G. Fan, «DNA methylation and its basic function», *Neuropsychopharmacology*, vol. 38, n.º 1, pp. 23–38, jan. de 2013.
- [19] M. Ehrlich, «DNA methylation in cancer: too much, but also too little», *Oncogene*, vol. 21, n.º 35, pp. 5400–5413, ago. de 2002.
- [20] K. O. Saunders, E. Lee, R. Parks et al., «Neutralizing antibody vaccine for pandemic and pre-emergent coronaviruses», *Nature*, vol. 594, n.º 7864, pp. 553–559, mai. de 2021. DOI: [10.1038/s41586-021-03594-0](https://doi.org/10.1038/s41586-021-03594-0).
- [21] D. Sterlin, A. Mathian, M. Miyara et al., «IgA dominates the early neutralizing antibody response to SARS-CoV-2», *Science Translational Medicine*, vol. 13, n.º 577, 2021. DOI: [10.1126/scitranslmed.abd2223](https://doi.org/10.1126/scitranslmed.abd2223).
- [22] Z. Wang, J. C. C. Lorenzi, F. Muecksch et al., «Enhanced SARS-CoV-2 neutralization by dimeric IgA», *Science Translational Medicine*, vol. 13, n.º 577, eabf1555, 2021. DOI: [10.1126/scitranslmed.abf1555](https://doi.org/10.1126/scitranslmed.abf1555).
- [23] C. J. Cramer, «Essential of computational chemistry: theories and models», em. John Wiley & Sons, 2004, p. 105, ISBN: 0-470-09181-9.
- [24] J. B. Foresman e E. Frisch, «Exploring Chemistry with Electronic Structure Methods», em, 3^a ed. Gaussian, Inc. A, 2015, ISBN: 978-1-935522-03-4.
- [25] F. Jensen, «Introduction to Computational Chemistry», em, 3^a ed. John Wiley & Sons, 2017, p. 515, ISBN: 978-1-1188-2599-0.
- [26] P. Atkins e R. Friedman, «Molecular quantum mechanics», em, 6^a ed. New York: Oxford University Press, 2011, pp. 295–334, ISBN: 978-0-19-954142-3.
- [27] L. R. Domingo, M. Ríos-Gutiérrez e P. Pérez, «Applications of the Conceptual Density Functional Theory Indices to Organic Chemistry Reactivity», *Molecules*, vol. 21, n.º 6, 2016, ISSN: 1420-3049. DOI: [10.3390/molecules21060748](https://doi.org/10.3390/molecules21060748).
- [28] R. G. Parr, L. V. Szentpály e S. Liu, «Electrophilicity Index», *Journal of the American Chemical Society*, vol. 121, n.º 9, pp. 1922–1924, 1999. DOI: [10.1021/ja983494x](https://doi.org/10.1021/ja983494x).
- [29] L. R. Domingo, E. Chamorro e P. Pérez, «Understanding the Reactivity of Captodative Ethylenes in Polar Cycloaddition Reactions. A Theoretical Study», *The Journal of Organic Chemistry*, vol. 73, n.º 12, pp. 4615–4624, 2008. DOI: [10.1021/jo800572a](https://doi.org/10.1021/jo800572a).

- [30] D. Frenkel e B. Smit, «Understanding molecular simulations - From algorithms to applications», em, 2^a ed. Academic Press, 2002, ISBN: 0-12-267351-4.
- [31] A Shrake e J. A. Rupley, «Environment and exposure to solvent of protein atoms. Lysozyme and insulin», en, *J. Mol. Biol.*, vol. 79, n.º 2, pp. 351–371, set. de 1973.
- [32] D. C. Rapaport, «The art of molecular dynamics simulation», em, 2^a ed. Cambridge University Press, 2004, pp. 5–6, ISBN: 978-0-511-19448-1.
- [33] D. P. Kroese, T. Brereton, T. Taimre e Z. I. Botev, «Why the Monte Carlo method is so important today», *WIREs Computational Statistics*, vol. 6, n.º 6, pp. 386–392, jun. de 2014. DOI: [10.1002/wics.1314](https://doi.org/10.1002/wics.1314).
- [34] H. Beard, A. Cholleti, D. Pearlman, W. Sherman e K. A. Loving, «Applying Physics-Based Scoring to Calculate Free Energies of Binding for Single Amino Acid Mutations in Protein-Protein Complexes», *PLOS ONE*, vol. 8, n.º 12, dez. de 2013. DOI: [10.1371/journal.pone.0082849](https://doi.org/10.1371/journal.pone.0082849).
- [35] S. Elbe e G. Buckland-Merrett, «Data, disease and diplomacy: GISAID’s innovative contribution to global health», *Global Chall.*, vol. 1, n.º 1, pp. 33–46, jan. de 2017.
- [36] J. Li, R. Abel, K. Zhu, Y. Cao, S. Zhao e R. A. Friesner, «The VSGB 2.0 model: a next generation energy model for high resolution protein structure modeling», en, *Proteins*, vol. 79, n.º 10, pp. 2794–2812, out. de 2011.
- [37] G. M. Sastry, M. Adzhigirey, T. Day, R. Annabhimoju e W. Sherman, «Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments», *J. Comput. Aided Mol. Des.*, vol. 27, n.º 3, pp. 221–234, mar. de 2013.
- [38] Schrödinger, LLC, «Schrödinger Release 2021-2: Maestro», 2021.
- [39] C. Lu, C. Wu, D. Ghoreishi et al., «OPLS4: Improving Force Field Accuracy on Challenging Regimes of Chemical Space», *Journal of Chemical Theory and Computation*, vol. 17, n.º 7, pp. 4291–4300, jun. de 2021. DOI: [10.1021/acs.jctc.1c00302](https://doi.org/10.1021/acs.jctc.1c00302). URL: <https://doi.org/10.1021/acs.jctc.1c00302>.
- [40] Y. Wu, F. Wang, C. Shen et al., «A noncompeting pair of human neutralizing antibodies block COVID-19 virus binding to its receptor ACE2», *Science*, vol. 368, n.º 6496, pp. 1274–1278, jun. de 2020. DOI: [10.1126/science.abc2241](https://doi.org/10.1126/science.abc2241).
- [41] S. K. Burley, C. Bhikadiya e C. Bi, «RCSB Protein Data Bank: powerful new tools for exploring 3D structures of biological macromolecules for basic and applied research and education in fundamental biology, biomedicine, biotechnology, bioengineering and energy sciences», *Nucleic Acids Res.*, vol. 49, n.º D1, pp. D437–D451, jan. de 2021.
- [42] E. Jurrus, D. Engel, K. Star et al., «Improvements to the APBS biomolecular solvation software suite», *Protein Science*, vol. 27, n.º 1, pp. 112–128, 2018. DOI: <https://doi.org/10.1002/pro.3280>.
- [43] K. J. Bowers, E. Chow, H. Xu et al., «Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters», em *Proceedings of the 2006 ACM/IEEE Conference on Supercomputing*, Tampa, Florida: Association for Computing Machinery, 2006, ISBN: 0769527000. DOI: [10.1145/1188455.1188544](https://doi.org/10.1145/1188455.1188544).
- [44] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey e M. L. Klein, «Comparison of simple potential functions for simulating liquid water», *The Journal of Chemical Physics*, vol. 79, n.º 2, pp. 926–935, 1983. URL: <https://doi.org/10.1063/1.445869>.

- [45] D. J. Evans e B. L. Holian, «The nose-hoover thermostat», *J. Chem. Phys.*, vol. 83, n.º 8, pp. 4069–4074, out. de 1985.
- [46] G. J. Martyna, D. J. Tobias e M. L. Klein, «Constant pressure molecular dynamics algorithms», *J. Chem. Phys.*, vol. 101, n.º 5, pp. 4177–4189, set. de 1994.
- [47] T. Darden, D. York e L. Pedersen, «Particle mesh Ewald: An N log(N) method for Ewald sums in large systems», *The Journal of Chemical Physics*, vol. 98, n.º 12, pp. 10089–10092, 1993. URL: <https://doi.org/10.1063/1.464397>.
- [48] M. Tuckerman, B. J. Berne e G. J. Martyna, «Reversible multiple time scale molecular dynamics», *J. Chem. Phys.*, vol. 97, n.º 3, pp. 1990–2001, ago. de 1992.
- [49] A. B. Herr, E. R. Ballister e P. J. Bjorkman, «Insights into IgA-mediated immune responses from the crystal structures of human FcRI and its complex with IgA1-Fc», *Nature*, vol. 423, n.º 6940, pp. 614–620, mai. de 2003. DOI: [10.1038/nature01685](https://doi.org/10.1038/nature01685). URL: <https://doi.org/10.1038/nature01685>.
- [50] D. Kozakov, R. Brenke, S. R. Comeau e S. Vajda, «PIPER: An FFT-based protein docking program with pairwise potentials», *Proteins: Structure, Function, and Bioinformatics*, vol. 65, n.º 2, pp. 392–406, 2006. DOI: <https://doi.org/10.1002/prot.21117>.
- [51] «PatchDock and SymmDock: servers for rigid and symmetric docking»,
- [52] «The ClusPro web server for protein–protein docking»,
- [53] Y. Yan, H. Tao, J. He e S.-Y. Huang, «The HDOCK server for integrated protein–protein docking», *Nat. Protoc.*, vol. 15, n.º 5, pp. 1829–1852, mai. de 2020.
- [54] M. J. Frisch, G. W. Trucks, H. B. Schlegel et al., *Gaussian 09W*, Gaussian Inc. Wallingford CT, 2009.
- [55] L. C. Xue, J. P. Rodrigues, P. L. Kastritis, A. M. Bonvin e A. Vangone, «PRODIGY: a web server for predicting the binding affinity of protein–protein complexes», *Bioinformatics*, btw514, ago. de 2016.
- [56] F. Obermeyer, M. Jankowiak, N. Barkas et al., «Analysis of 6.4 million SARS-CoV-2 genomes identifies mutations associated with fitness», *Science*, vol. 376, n.º 6599, pp. 1327–1332, jun. de 2022. DOI: [10.1126/science.abm1208](https://doi.org/10.1126/science.abm1208). URL: <https://doi.org/10.1126/science.abm1208>.

- Setembro (2021) → Agosto (2022) • Levantamento bibliográfico sobre aspectos moleculares do vírus SARS-CoV-2 e surgimento de novas variantes e linhagens. Leitura e aprendizado na literatura sobre a virulência e imunogenicidade viral diante das variantes.
- Setembro (2021) → Novembro (2021) • Simulação da dinâmica molecular em 100ns da variante Ômicron do SARS-CoV-2 para o complexo anticorpo-antígeno.
- Setembro (2021) → Novembro (2021) • Simulação da dinâmica molecular em 100ns da variante Omicron do SARS-CoV-2 para o complexo ACE2-RBD com mutações inseridas computacionalmente.
- Novembro (2021) → Dezembro (2021) • Simulação da dinâmica molecular em 100ns da variante Ômicron do SARS-CoV-2 para o complexo ACE2-RBD a partir de mutações obtidas mediante cristalografia.
- Dezembro(2021) → Janeiro (2022) • Simulação da interação da proteína Spike com os anticorpos neutralizantes IgG, IgA e IgM mediante dinâmica molecular em 100ns.
- Janeiro (2022) → Fevereiro (2022) • Utilização do método de otimização de Monte Carlo para elucidar as mutações que maximizam a neutralização da proteína Spike por anticorpos.
- Fevereiro (2022) → Março (2022) • Comparação do potencial de neutralização da proteína de Nucleocapsídeo (N) em relação à proteína Spike (S).
- Março (2022) → Maio (2022) • Estudo dos descritores de química-quântica na mutação K417T. Análise qualitativa de todas as mutações que se fizeram presente desde a variante alfa até a última subvariante da Ômicron.
- Maio (2022) → Agosto (2022) • Conclusões sobre possíveis implicações da estabilidade conformacional para o aumento da transmissibilidade viral
- Agosto (2022) • Conclusão do relatório final de Iniciação Científica. (atividade obrigatória)
- Novembro (2022) • Preparação da apresentação final para o Congresso de Iniciação Científica. (atividade obrigatória)