

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE BIOTECNOLOGIA
CURSO DE BIOTECNOLOGIA

Análise in silico da interação entre proteínas spike de novas variantes de SARS-CoV-2 e isoformas da enzima conversora de angiotensina 2

Victor Hugo Oliveira de Andrade

Monografia apresentada à Coordenação do Curso de Biotecnologia, da Universidade Federal de Uberlândia, para obtenção do grau de Bacharel em Biotecnologia.

Uberlândia – MG
Março – 2022

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE BIOTECNOLOGIA
CURSO DE BIOTECNOLOGIA

Análise in silico da interação entre proteínas spike de novas variantes de SARS-CoV-2 e isoformas da enzima conversora de angiotensina 2

Victor Hugo Oliveira de Andrade

Jair Pereira da Cunha Junior

Daniel Ferreira De Lima Neto

Monografia apresentada à Coordenação do Curso de Biotecnologia, da Universidade Federal de Uberlândia, para obtenção do grau de Bacharel em Biotecnologia.

Uberlândia – MG
Março – 2022

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE BIOTECNOLOGIA
CURSO DE BIOTECNOLOGIA

Análise in silico da interação entre proteínas spike de novas variantes de SARS-CoV-2 e isoformas da enzima conversora de angiotensina 2

Victor Hugo Oliveira de Andrade

Jair Pereira da Cunha Junior
Instituto de Ciências Biomédicas

Carlos Roberto Prudêncio
Instituto Adolfo Lutz

Daniel Ferreira De Lima Neto
Coordenação Geral de Laboratórios de Saúde Pública - CGLAB/DAEVs/SVS/MS

Homologado pela coordenação do Curso de

Biotecnologia em __/__/__

Nilson Nicolau Júnior

Uberlândia – MG
Março – 2022

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE BIOTECNOLOGIA
CURSO DE BIOTECNOLOGIA

Análise *in silico* da interação entre proteínas spike de novas variantes de SARS-CoV-2 e isoformas da enzima conversora de angiotensina 2

Victor Hugo Oliveira de Andrade

Aprovado pela Banca Examinadora em: / / Nota: _____

Nome e assinatura do Presidente da Banca Examinadora

Uberlândia, de de

AGRADECIMENTOS

Agradeço o meu orientador Jair Pereira da Cunha Junior e meu coorientador Daniel Ferreira de Lima Neto pelo apoio científico e técnico para minha formação acadêmica e elaboração desta monografia.

Agradeço o compromisso do líder, membros e suplente da banca examinadora.

Agradeço a Universidade Federal de Uberlândia (UFU) o acesso a conteúdo científico de artigos, relatórios, livros e notícias relevantes para o meu trabalho.

Agradeço ao Instituto de Biotecnologia “Warwick Estevam Kerr” (IBTEC) e a Coordenação do Curso de Biotecnologia ao apoio e atendimento às questões burocráticas.

Agradeço a minha família e amigos pelo apoio e incentivo.

RESUMO

O novo coronavírus SARS-CoV-2 que é responsável pela pandemia de COVID-19 apresenta gravidade preocupante, alta transmissão e elevada taxa de mutação, e por isso, a proteína spike de SARS-CoV-2, que medeia a entrada do vírus na célula, e sua interação com o receptor celular ECA2 humano se tornaram alvos de estudos. Nesse contexto, são buscadas características moleculares intrínsecas à interação proteína spike de SARS-CoV-2 e ECA2 que potencialmente justificam mudanças da taxa de transmissibilidade, de gravidade e do impacto no sistema imune das variantes de SARS-CoV-2. Para isso, análises *in silico* de interação proteína-proteína vêm se mostrando metodologias eficientes para céleres conclusões científicas sobre essa interação. Foi descoberto que na interação intermolecular entre a proteína spike do SARS-CoV-2 e ECA2 os resíduos Arg403 na estrutura original de Wuhan, Arg498 e Arg403 na variante Omicron e Lys440, His505, Arg403 e Arg498 na subvariante BA.2 que deixaram o perfil da superfície interativa da spike mais positivo. Isso sugere uma seleção positiva promovendo maior estabilidade na ligação por atração de cargas, pois a superfície de ECA2 apresenta um perfil predominantemente eletronegativo.

SUMÁRIO

1. INTRODUÇÃO	1
2. OBJETIVOS	6
2.1. Objetivo geral	6
2.2. Objetivos específicos	7
3. METODOLOGIA	8
3.1. Sequências de aminoácidos das proteínas spike do SARS-CoV-2 e isoformas de ECA2	11
3.2. Geração de mutações da proteína spike do SARS-CoV-2	11
3.3. Alinhamento das sequências primárias das proteínas spike do SARS-CoV-2 e isoformas de ECA2 humanas	12
3.4. Modelagem computacional de estruturas tridimensionais (3D) das proteínas spike das variantes de SARS-CoV-2 e isoformas de ECA2	12
3.5. Avaliação da qualidade dos modelos gerados da modelagem.....	14
3.6. Sobreposição e alinhamento modelos atômicos da proteína spike de SARS-CoV-2 e isoformas de ECA2 refinados.....	14
3.7. Cálculo das raízes dos desvios quadráticos médios (RMSD) dos alinhamentos das proteínas spike do SARS-CoV-2.....	15
3.8. Interação molecular (docking molecular) entre variantes das proteínas spike de SARS-CoV-2 e isoformas de ECA2.....	16
3.9. Análise 2D das interações intermoleculares entre as proteínas spike das variantes do SARS-CoV-2 com as isoformas de ECA2	17
4. RESULTADOS E DISCUSSÃO	17

4.1.	Alinhamento das sequências de aminoácidos das proteínas spike do SARS-CoV-2	17
4.2.	Alinhamento das sequências de aminoácidos das isoformas de ECA2.....	20
4.3.	Modelagem das proteínas spike do SARS-CoV-2 e isoformas de ECA2	22
4.4.	RMSD dos alinhamentos estruturais do grupo 1 e 2 da proteína spike do SARS-CoV-2 e isoformas de ECA2	30
4.5.	Avaliação de qualidade dos modelos atômicos	33
4.5.1.	Molprobitry.....	33
4.5.2.	VERIFY3D.....	36
4.5.3.	PROCHECK.....	41
4.5.4.	ProSA	87
4.6.	Interação molecular entre variantes das proteínas spike de SARS-CoV-2 e isoformas de ECA2	95
5.	CONCLUSÃO	102
6.	REFERÊNCIAS BIBLIOGRÁFICAS.....	104

1. INTRODUÇÃO

Os coronavírus causaram 3 grandes epidemias em torno do mundo nas últimas duas décadas: a síndrome respiratória aguda grave (SARS), a síndrome respiratória do Oriente Médio (MERS) e a *coronavirus disease* 2019 (COVID-19). Os primeiros casos do novo coronavírus (SARS-CoV-2) foram relatados no dia 1 de Dezembro de 2019 em Wuhan, na China, e em Março de 2020 a Organização Mundial da Saúde (OMS) declarou que se tratava de uma pandemia (RAMANATHAN et al., 2020). De acordo com o relatório emergencial epidemiológico da OMS de 15 de Fevereiro de 2022, até o dia 13 do mesmo mês, o mundo ultrapassou a marca de 409 milhões de casos confirmados e de 5,8 milhões de mortes causadas pela COVID-19 (WHO, 2022a).

Os coronavírus da família *Coronaviridae* são vírus envelopados e de RNA de única fita sentido positivo (GORBALENYA et al., 2020), subdivididos em 4 gêneros: alfa, beta, gamma e delta (FIELDS; KNIPE; HOWLEY, 2013; “International Committee on Taxonomy of Viruses”, 1991; WOO et al., 2009), sendo somente os coronavírus alpha-hCoV-NL63, alpha-hCoV-229E, beta-HCoV-OC43, beta-HKU1 (CHIU et al., 2005; GORSE et al., 2009; JEAN et al., 2013; JEVŠNIK et al., 2012), e SARS-CoV, MERS-CoV e SARS-CoV-2, os causadores de doenças em seres humanos. Seus genomas possuem a capacidade de produzir 16 proteínas não estruturais, 9 proteínas acessórias e 4 proteínas estruturais, sendo uma das estruturais a proteína Spike (S) que é responsável por mediar a entrada do vírus na célula (PERLMAN; NETLAND, 2009). A proteína S é dividida nas subunidades S1 (13-541) responsável pela ligação dela ao receptor Enzima Conversora de Angiotensina (ECA) 2 através do receptor binding domain (RBD), e a subunidade S2 (543-1208) que medeia a fusão da membrana do vírus com a membrana da célula hospedeira (HOFFMANN et al., 2020; WALLS et al., 2020; WRAPP et al., 2020).

Os ciclos virais dos diferentes coronavírus que infectam humanos são semelhantes, diferindo nos receptores de entrada que são utilizados durante a etapa de adsorção. No caso dos SARS-CoV-2, HCoV-NL63 e SARS-CoV, os receptores ECA2 e serinoprotease transmembrana II (TMPRSS2) são os responsáveis pela entrada do vírus na célula. Após as etapas de adsorção, penetração e desnudamento da partícula viral no citoplasma da célula, se inicia a etapa de replicação viral, na qual o RNA de simples fita é replicado por um conjunto de 15 proteínas não estruturais de replicação advindas da tradução das regiões da extremidade 5', ORF1a e ORF1b, do próprio RNA viral. À medida que a síntese de proteínas não estruturais ocorre, o engendro biológico leva a formação de microambientes protetores de replicação vesiculares com RNA viral dupla fita no interior. Após a tradução de todas as proteínas virais, as proteínas estruturais transitam do retículo endoplasmático ao complexo de Golgi, onde encontram o RNA viral parcialmente encapsulado pela proteína do nucleocapsídeo (N) e isso provoca um brotamento no lúmen das vesículas secretoras, o que acarreta na exocitose dos vírions (V'KOVSKI et al., 2021).

Até o dia 31 de agosto de 2021, 4 variantes de preocupação (VOC) do SARS-CoV-2 surgiram na população humana: alfa (B.1.1.7), beta (B.1.351) gamma (P.1) e delta (B.1.617.2) (CHOI; SMITH, 2021). Podem-se observar alguns desses clados na Figura 1. Dentre as mutações que caracterizam as variantes, se destacam as mutações: A222V presente na variante delta que, quando associada à mutação D614G, pode provocar uma mudança na estabilidade da ligação entre proteína spike do SARS-CoV-2 e ECA2 (JACOB et al., 2021). A mutação D614G presente em todas as variantes de preocupação atuais está associada a uma maior taxa de replicação viral e possivelmente maior taxa de transmissão (KORBER et al., 2020; PLANTE et al., 2021; VOLZ et al., 2021).

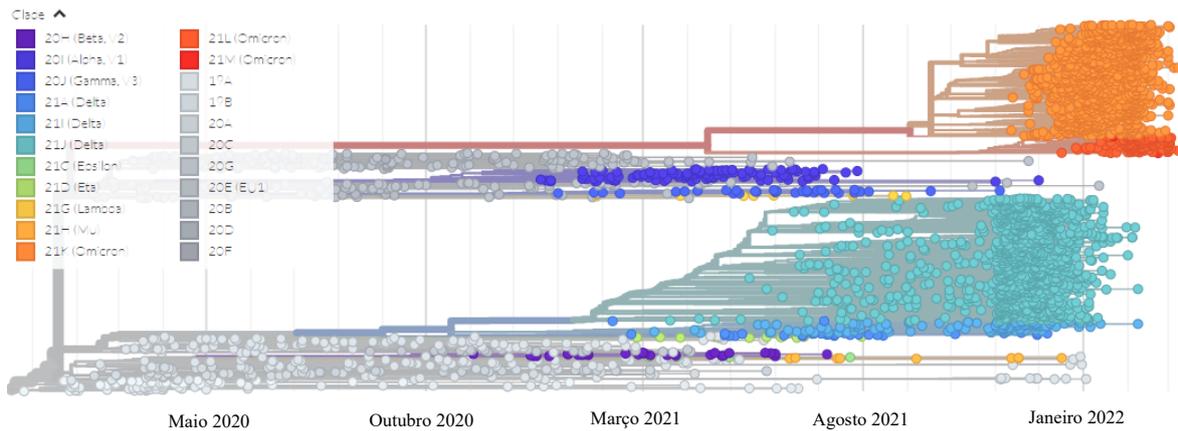


Figura 1. Representação retangular dos clados do SARS-CoV-2 de maio de 2020 até janeiro de 2022. Fonte da figura: plataforma GISAID (HADFIELD et al., 2018; SAGULENKO; PULLER; NEHER, 2018).

Em 26 de Novembro de 2021, a OMS adicionou ao catálogo de variantes de preocupação a nova variante Omicron, descoberta primeiramente na África do Sul. Essa nova variante traz consigo 3 linhagens: a BA.1 (B.1.1.529.1), BA.2 (B.1.1.529.2), e BA.3 (B.1.1.529.3) (DESINGU; NAGARAJAN; DHAMA, 2022; WANG; CHENG, 2022). A proteína spike da variante Omicron é a que mais se difere estruturalmente da molécula original de Wuhan em comparação as outras variantes já descritas, acumulando um total de 34 mutações (YANG et al., 2021). Dentre as mutações descritas, destacam as mutações N501Y, D614G, K417N e T478K, que já estavam descritas em outras variantes de preocupação (Variants of Concern – VoC) anteriores, e que juntamente com outras novas mutações da variante Omicron, potencializam os eventos de reinfecção e resistência parcial a algumas vacinas existentes. Duas das mutações mencionadas para a variante Omicron, K417N e T478K estão presentes também na variante Delta e são relacionadas respectivamente a alterações conformacionais da proteína S, levando possivelmente a uma evasão imune e mudança do perfil eletrostático da molécula, eventualmente acarretando no fortalecimento da interação RBD spike do SARS-CoV-2 e ECA2 (ARAF et al., 2022). Recentemente existem indícios de que a nova subvariante de Omicron denominada de BA.2, apresente uma maior taxa de transmissão, contudo, maiores estudos ainda são necessários para esclarecer esse fenômeno biológico (DHAWAN; PRIYANKA; CHOUDHARY, 2022).

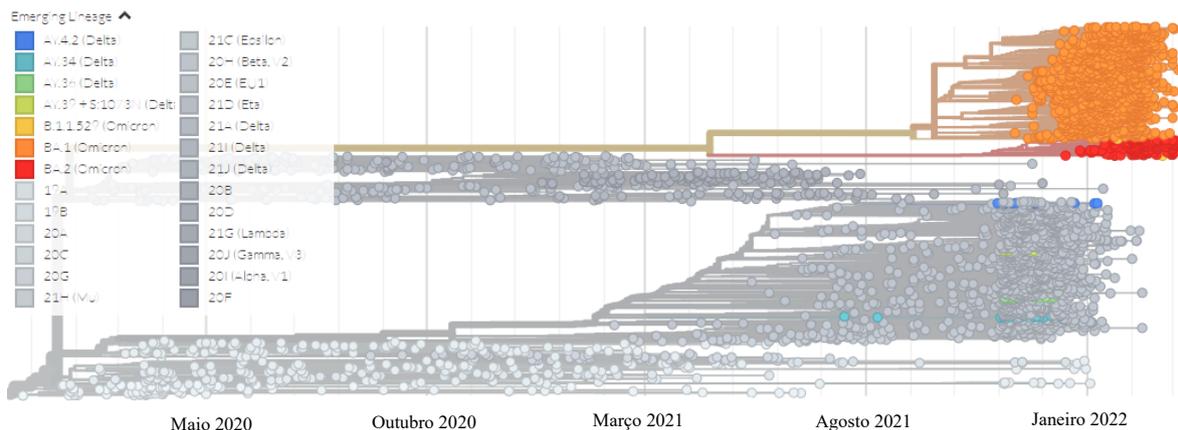


Figura 2. Representação retangular das variantes emergentes do SARS-CoV-2 de maio de 2020 até janeiro de 2022. Figura adaptada da plataforma GISAID (HADFIELD et al., 2018; SAGULENKO; PULLER; NEHER, 2018).

ECA2 humana é uma enzima carboxipeptidase associada à membrana de 805 aminoácidos que regula negativamente o sistema renina-angiotensina (responsável por controlar a pressão arterial e uma série de outras funções), convertendo angiotensina I em angiotensina 1-7, hormônio vasodilatador e anti-inflamatório e também podendo converter em angiotensina 1-9 de efeito ainda desconhecido (ALEXANDRE et al., 2020; BLUME et al., 2021; TIPNIS et al., 2000b). Maior expressão de transcritos para ECA2 foi descrita nos órgãos, vesícula biliar, coração, rim, intestino delgado, cólon, duodeno e testículo (FAGERBERG et al., 2014). A estrutura molecular de ECA2 possui 40% de identidade e 61% de similaridade com a molécula ECA 1 (TIPNIS et al., 2000a).

A codificação da ECA2 se dá pelo seu gene *ECA2* através da transcrição de 18 exons, podendo gerar por “splicing” alternativo a sua isoforma 2. Também apresenta muitos polimorfismos em seu gene, sendo esses, associados a predisposições em hipertensão e doenças cardiovasculares (BLUME et al., 2021; BURRELL et al., 2013; PATEL et al., 2014; PINHEIRO et al., 2019).

Recentemente a ECA2 passou a ser alvo de investigações científicas moleculares, tendo em vista que esta molécula é o receptor de infecção celular do vírus SARS-CoV-1,

interação (DE LAS RIVAS; FONTANILLO, 2010). As IPPs estão envolvidas em bioprocessos vitais como a replicação e transcrição de DNA, ciclos metabólicos, vias de transdução de sinais e respostas celulares e imunológicas (CATERINO et al., 2017; LIU et al., 2019). Para entender os mecanismos das IPPs, são necessários detalhes de nível atômico e molecular. Em casos que envolvam a formação de complexos transitórios de IPPs, a estrutura desses complexos proteicos pode ser resolvida por métodos experimentais, no entanto, esta abordagem se torna mais laboriosa e demorada em comparação a outras metodologias (KOZAKOV et al., 2006) e, portanto, métodos avançados de simulação/predição por ferramentas computacionais têm emergido como uma solução a este problema. Dentre as ferramentas, métodos de simulações de interação intermolecular (docking molecular) vêm sendo desenvolvidos para predizer a estrutura desses complexos com uma precisão similar as metodologias de cristalografia, difração de raio X e ressonância magnética nuclear (RITCHIE, 2008; SMITH; STERNBERG, 2002).

Portanto, neste trabalho monográfico, realizamos análises *in silico* das interações entre proteínas spike de variantes de preocupação (Variants of Concern – VoCs) do SARS-CoV-2 com isoformas humanas descritas para enzima conversora de angiotensina 2 (ECA2).

2. OBJETIVOS

2.1. Objetivo geral

Avaliar por meio de abordagem computacional (*in silico*) a interface de interação entre as proteínas spike de variantes do SARS-CoV-2 com isoformas humanas da enzima conversora de angiotensina 2 (ECA2).

2.2. Objetivos específicos

- Realizar análises comparativas das sequências primárias da proteína spike das variantes de SARS-CoV-2 em relação uma sequência de referência de genótipo original Wuhan.
- Definir as diferenças em cada molécula para inclusão das substituições e/ou adições de aminoácidos que caracterizam as mutações pontuais ou combinadas presentes na molécula spike de SARS-CoV-2;
- Realizar análises comparativas das sequências primária da ECA2 humana (código de acesso UniprotKB: Q9BYF1) com suas isoformas descritas na literatura;
- Modelar as estruturas moleculares monoméricas da proteína spike de diferentes variantes de SARS-CoV-2 contendo com mutações pontuais ou combinadas;
- Modelar as estruturas moleculares da proteína ECA2 humana e suas isoformas descritas na literatura;
- Realizar a minimização energética dos modelos atômicos;
- Refinar os modelos atômicos gerados para cada molécula, quando necessário, utilizando a plataforma Molprobit;
- Avaliar à qualidade estereoquímica dos modelos gerados na modelagem por homologia ou ab initio realizadas nos diferentes servidores computacionais;
- Efetuar por abordagem computacional as interações moleculares (docking molecular) entre as formas monoméricas da proteína spike de variantes do SARS-CoV-2 e isoformas de ECA2; e
- Elucidar e comparar as potenciais interações proteína-proteína (IPPs) estabelecidas entre diferentes combinações de proteínas spike do SARS-CoV-2 e isoformas de ECA2.

3. METODOLOGIA

O presente trabalho foi realizado seguindo passos demonstrados no fluxograma da Figura 4. Primeiramente, foram realizadas análises comparativas das sequências de proteínas de spike do SARS-CoV-2 e isoformas de ECA2 gerando um conjunto (set) de sequências, que a seguir, foram modeladas. Após a modelagem molecular, foi realizado protocolo de minimização de energia, refinamento e avaliações estereoquímicas das estruturas proteínicas moleculares. Com os modelos atômicos avaliados e refinados, foram realizadas sobreposições moleculares 3D e alinhamentos das estruturas para uma análise comparativa através dos dados de RMSD. Com os modelos referentes às proteínas spike do SARS-CoV-2, foram realizados docking molecular com as isoformas de ECA2. As interações proteína-proteína (IPPs) formadas nos complexos gerados no docking molecular foram elucidadas em diagramas 2D.

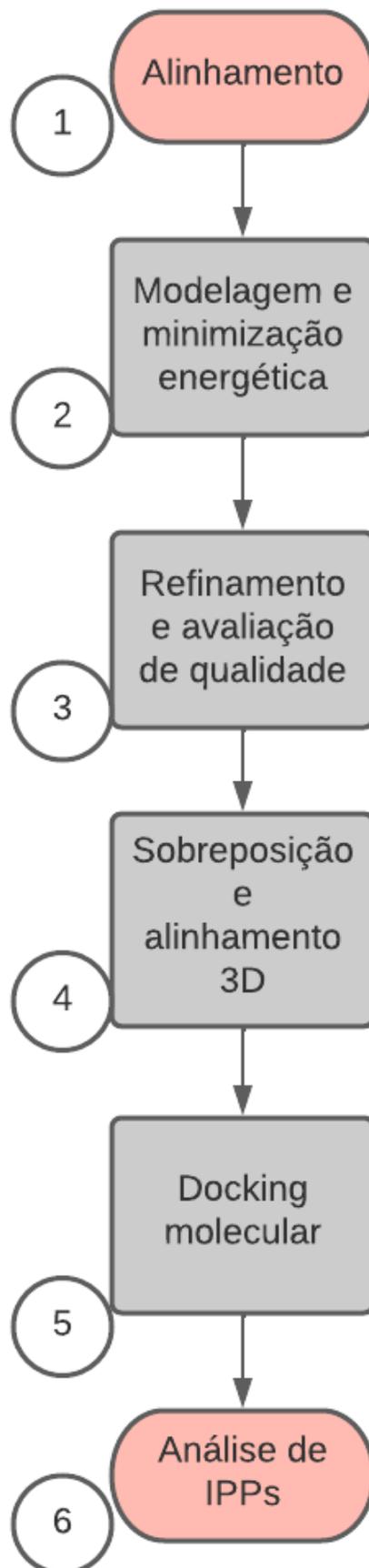


Figura 4: Fluxograma da metodologia aplicada. (1) Alinhamento das sequências primárias dos grupos 1 e 2 da proteína spike de SARS-CoV-2 e isoformas de ECA2, (2) modelagem dos modelos das variantes de SARS-CoV-

2 e isoformas de ECA2 e minimização de energia das estruturas (3) refinamento e avaliações de qualidade dos modelos gerados, (3) sobreposição e alinhamento dos modelos referentes às proteínas spike das variantes completas e mutações pontuais combinadas ou não de SARS-CoV-2, (4) docking molecular entre os modelos proteicos da spike de SARS-CoV-2 e isoformas de ECA2 e (5) análise de diagramas 2D das interações proteína-proteína (IPPs) formadas pelo docking molecular das diferentes combinações de proteínas spike do SARS-CoV-2 e isoformas de ECA2. Figura elaborada na plataforma Lucidchart (“Lucidchart”, [s.d.]

Conforme a Tabela 1, os modelos das proteínas spike do SARS-CoV-2 foram divididos em 2 grupos, sendo que, para todas as proteínas foram estabelecidas abreviações e foram gerados (quando modeladas) ou coletados os códigos PDB.

Tabela 1. Nomes dos modelos atômicos das proteínas spike do SARS-CoV-2 e suas respectivas abreviações e códigos PDB após refinamento.

Modelo grupo 1	Abreviação	Código PDB	Modelo grupo 2	Abreviação	Código PDB
Referência	S1A	7CAK-A	Referência	S2A	7K43-E
Variante de Wuhan SARS-CoV-2	S1B	D269*	Variante de Wuhan	S2B	D919*
A222V proteína spike SARS-CoV-2	S1C	D266*	Variante Alfa (B.1.1.7) proteína spike SARS-CoV-2	S2C	D920*
D614G proteína spike SARS-CoV-2	S1D	D476*	Variante Beta (B.1.351) proteína spike SARS-CoV-2	S2D	D922*
E484K proteína spike SARS-CoV-2	S1E	D490*	Variante Delta (B.1.617.2) proteína spike SARS-CoV-2	S2E	D925*
E484K, N501Y e K417Y proteína spike SARS-CoV-2	S1F	D917*	Variante Omicron (B.1.1.529) proteína spike SARS-CoV-2	S2F	D928*
			Subvariante Omicron (BA.2)	S2G	D935*

Tabela 1. Nomes dos modelos atômicos das proteínas spike do SARS-CoV-2 e suas respectivas abreviações e códigos PDB após refinamento. Respetivos dados do título, abreviação e código PDB das proteínas spike do SARS-CoV-2 das variantes de mutações pontuais ou múltiplas refinadas e modeladas tomando como referência a sequência de aminoácidos da cadeia A da estrutura 7CAK para o grupo 1 múltiplas estruturas para o grupo 2. *Códigos PDB gerados pelo PDBsum generate e, portanto, ainda não armazenados no banco de dados PDB.

Conforme a Tabela 2, foram modeladas 7 isoformas de ECA2 sendo estabelecidas respectivas abreviações e gerados códigos PDB correspondentes.

Tabela 2. Nomes dos modelos atômicos das isoformas de ECA2 e suas respectivas abreviações e códigos PDB após refinamento.

Modelo	Abreviação	Código PDB
Isoforma 1 ECA2	E1A	D932*
Isoforma 2 ECA2	E1B	D936*
Isoforma 3 ECA2	E1C	D936*
Isoforma 4 ECA2	E1D	D937*
Isoforma 5 ECA2	E1E	D966*
Isoforma 6 ECA2	E1F	D967*
Isoforma 7 ECA2	E1G	E129*

Tabela 2. Nomes dos modelos atômicos das isoformas de ECA2 e suas respectivas abreviações e códigos PDB após refinamento. Respetivos dados do título, abreviação e código PDB das isoformas de ECA2 refinadas e modeladas. *Códigos PDB gerados pelo PDBsum generate e, portanto, ainda não armazenados no banco de dados PDB.

3.1. Sequências de aminoácidos das proteínas spike do SARS-CoV-2 e isoformas de ECA2

As sequências de aminoácidos em formato FASTA da proteína spike do SARS-CoV-2 foram coletadas da cadeia A da estrutura de código PDB 7CAK (LV et al., 2020) depositada por crio microscopia eletrônica no banco de dados Protein Data Bank (PDB). Já para a proteína ECA2, as sequências das isoformas 1 (BLUME et al., 2021), isoforma 2 (BLUME et al., 2021; ONABAJO et al., 2020) e isoformas 3, 4, 5, 6 e 7 foram extraídas do banco de dados UniprotKB (BATEMAN et al., 2021). Os códigos de acesso das sequências das isoformas de ECA2 se encontram na Tabela 3.

Tabela 3. Isoformas de ECA2 e seus respectivos códigos de entrada Uniprot KB.

Isoforma de ECA2	Código de entrada UniprotKB
1	Q9BYF1-1*
2	Q9BYF1-3*
3	A0A7I2V3N4**
4	A0A7I2V4H0**
5	A0A7I2V2E9**
6	A0A7I2V3X6**
7	A0A7I2V5W5**

Tabela 3. Isoformas de ECA2 e seus respectivos códigos de entrada Uniprot KB. * Sequências com evidência experimental a nível proteico e revisadas pelo banco de dados Swiss-Prot. **Sequências com evidência experimental a nível proteico e não revisadas.

3.2. Geração de mutações da proteína spike do SARS-CoV-2

Para gerar os modelos atômicos das proteínas spike mutantes de SARS-CoV-2, as mutações descritas para as variantes de SARS-CoV-2 foram manualmente inseridas nas sequências de aminoácidos da cadeia A da sequência de aminoácidos correspondente à proteína spike do SARS-CoV-2 original de Wuhan. Para este presente trabalho, somente as mutações da Tabela 4 foram consideradas para as análises posteriores.

Tabela 4. Mutações consideradas para a modelagem das variantes de preocupação (Variants of Concern - VoC).

VoC	Mutações
Alpha	Δ H69/V70, Δ 144, L216F, N439K, Y453F, E484K, S494P, N501Y, A570D, D614G, P681H, T716I, Q836R, S982A, D1118H
Beta	L18F, A27S, D80A, D215G, K417N, E484K, N501Y, D614G, A701V.
Delta	V70F, T95I, G142D, E156G, Δ 157-158, A222V, W258L, K417N, L452R, T478K, D614G, P681R, D950N
Omicron	A67V, Δ 69-70, T95I, G142D/ Δ 143-145, Δ 211/L212I, ins214EPE, G339D, S371L, S373P, S375F, K417N,

N440K, G446S, S477N, T478K, E484K, G496S, Q498R, N501Y, Y505H, T547K, D614G, H655Y, N679K,
P681H, N764K, D796Y, N856K, Q954H, N969K, L981F
A67V, Δ69-70, T95I, ΔG142-143, ΔY144, Y145D, ΔN211, L212I, G339D, S371L, S373P, S375F, K417N,
BA.2 N440K, G446S, S477N, T478K, E484A, Q493R, G496S, Q498R, N501Y, Y505H, T547K, D614G, H655Y,
N679K, P681H, N764K, D796Y, N856K, Q954H, N969K, L981F

Tabela 4. Mutações consideradas para a modelagem das variantes de preocupação (Variants of Concern - VoC). “Ins” e “Δ” significam respectivamente inserção (ões) ou deleção (ões) de resíduos. As posições de aminoácidos são definidas com base na posição original da sequência de referência de Wuhan.

3.3. Alinhamento das sequências primárias das proteínas spike do SARS-CoV-2 e isoformas de ECA2 humanas

Foram realizados 3 alinhamentos das sequências primárias das proteínas spike do SARS-CoV-2 do grupo 1 e 2 e das isoformas de ECA2. O algoritmo selecionado foi o MUSCLE (EDGAR, 2004). Para as penalidades inerentes ao alinhamento, o parâmetro de definição de lacuna aberta foi de -2,9, de extensão da lacuna de 0 e multiplicador de Hidrofobicidade de 1,2. Foi estabelecido o número máximo de 16 iterações para o alinhamento, sendo que todas as iterações se executaram o método cluster. Com os alinhamentos, foram montados esquemas representativos.

3.4. Modelagem computacional de estruturas tridimensionais (3D) das proteínas spike das variantes de SARS-CoV-2 e isoformas de ECA2

As estruturas foram modeladas gerando 2 grupos de proteínas spike do SARS-CoV-2: o primeiro contendo 1 modelo de referência (código PDB: 7CAK) (S1A), 1 modelo da variante de Wuhan (S1B), 4 modelos de mutações pontuais não combinadas e combinadas: A222V (S1C), D614G (S1D), E484K (S1E) e E484K, N501Y e K417Y (S1F); e o segundo grupo contendo 1 modelo referência (código PDB: 7K43) (S2A), 1 modelo da variante de Wuhan (S2B) e 5 modelos de variantes da proteína spike do SARS-CoV-2: S2C (Alfa), S2D (Beta), S2E (Delta), S2F (Omicron) e S2G (BA.2). Para a modelagem do grupo 1 de proteínas spike do SARS-CoV-2 e isoformas de ECA2, foi seguido o protocolo de

modelagem por homologia e *ab initio* das plataformas I-Tasser e SwissModel (ROY; KUCUKURAL; ZHANG, 2010; WATERHOUSE et al., 2018).

Para a modelagem do grupo 2 das proteínas, foi seguido o protocolo de modelagem por homologia da plataforma YASARA (KRIEGER; VRIEND, 2014) tomando como referência os modelos da Tabela 5. Esta abordagem metodológica difere da modelagem no I-Tasser e SwissModel, pois considera múltiplos modelos atômicos como referência para modelagem por homologia.

Todas as estruturas modeladas por ambas as metodologias foram minimizadas a fim de melhorar as características estereoquímicas, já que o processo de modelagem computacional pode intrinsicamente gerar modelos com regiões de baixa qualidade estereoquímica. Somente os modelos atômicos referentes às isoformas de ECA2 não passaram pelo processo de minimização de energia.

Tabela 5. Modelos utilizados para a modelagem por homologia da estrutura de referência híbrida para o segundo grupo de proteínas spike do SARS-CoV-2 pelo software YASARA.

Cobertura	ID	Resolução	Leitura
85%	7LY2-J	2.50 A	EM glicoproteína Spike [Síndrome respiratória aguda grave coronavírus 2] <SPIKE_SARS2(1-1208)> (1041 resíduos com 0.961 pontos de qualidade), liberado 2021-04-14
88%	7K43-E	2.60 A	EM glicoproteína Spike [Síndrome respiratória aguda grave coronavírus 2] <SPIKE_SARS2(14-1211)> (1034 resíduos com 0.830 pontos de qualidade), liberado 2020-10-07
81%	7LXY-J	2.20 A	EM glicoproteína Spike [Síndrome respiratória aguda grave coronavírus 2] <SPIKE_SARS2(1-1208)> (1035 resíduos com 0.970 pontos de qualidade), liberado 2021-04-14
88%	7JJI-B	3.60 A	EM glicoproteína Spike [Síndrome respiratória aguda grave coronavírus 2] <SPIKE_SARS2(1-1273)> (1109 resíduos com 0.795 pontos de qualidade), liberado 2020-08-26
88%	7KE9-A	3.08 A	EM glicoproteína Spike [Síndrome respiratória aguda grave coronavírus 2] <SPIKE_SARS2(1-1208)> (966 resíduos com 0.785 pontos de qualidade), liberado 2020-11-04
85%	7KDG-A	3.01 A	EM glicoproteína Spike [Síndrome respiratória aguda grave coronavírus 2] <SPIKE_SARS2(1-1208)> (972 resíduos com 0.799 pontos de qualidade), liberado 2020-11-04
78%	7N8H-K	2.30 A	EM glicoproteína Spike [Síndrome respiratória aguda grave coronavírus 2] <SPIKE_SARS2(1-1208)> (999 resíduos com 0.936 pontos de qualidade), liberado 2021-07-14
86%	7LWV-B	3.12 A	EM glicoproteína Spike [Síndrome respiratória aguda grave coronavírus 2] <SPIKE_SARS2(1-1208)> (995 resíduos com 0.755 pontos de qualidade), liberado 2021-03-31
87%	6XR8-A	Desconhecido ou NMR	EM glicoproteína Spike [Síndrome respiratória aguda grave coronavírus 2] <SPIKE_SARS2(1-1273)> (1107 resíduos com 0.723 pontos de qualidade), liberado 2020-07-22
88%	7N1Q-C	3.44 A	EM glicoproteína Spike [Síndrome respiratória aguda grave coronavírus 2] <SPIKE_SARS2(1-1273)> (1104 resíduos com 0.694 pontos de qualidade), liberado 2021-07-07

Tabela 5. Modelos utilizados para a modelagem por homologia da estrutura de referência híbrida para o segundo grupo de proteínas spike do SARS-CoV-2 pelo software YASARA. Cobertura da referência, cadeia e código PDB (ID), resolução da estrutura depositada no Protein Data Bank e quantidade de resíduos e suas pontuações correspondentes.

3.5. Avaliação da qualidade dos modelos gerados da modelagem

As proteínas foram submetidas em formato PDB em 4 ferramentas para refinamento se necessário e avaliação de qualidades estereoquímicas. Todos os resultados das avaliações foram comparados ao modelo atômico de molécula de referência extraída a partir do banco de dados PDB. Foram selecionadas as proteínas cristalografadas spike 7CAK e 7K43 como proteínas de referência, sendo considerada somente a cadeia A da estrutura 7CAK para as variantes da proteína spike do SARS-CoV-2 do grupo 1 e cadeia E da estrutura 7K43 para as variantes de spike do SARS-CoV-2 do grupo 2. A estrutura 7K43 foi selecionada como referência para o grupo 2 por possuir uma taxa de alta cobertura e alta resolução (Tabela 5) molecular na modelagem por homologia utilizando sequência de Wuhan para o grupo 2. Os modelos de referência não sofreram quaisquer refinamentos ou minimizações de energia. Para o refinamento e avaliação estereoquímica, as estruturas 3D das proteínas foram submetidas na plataforma Molprobit Manchester. Após o refinamento, todos os modelos das proteínas spike do SARS-CoV-2 do grupo 1 e 2 e isoformas de ECA2 foram submetidas as avaliações de qualidade nas ferramentas VERIFY3D, PROCHECK pela plataforma SAVES v6.0 e na ferramenta ProSA por meio da plataforma ProSA-Web Protein Structure.

3.6. Sobreposição e alinhamento modelos atômicos da proteína spike de SARS-CoV-2 e isoformas de ECA2 refinados

As comparações das sequências de aminoácidos entre as estruturas das proteínas spike SARS-CoV-2 e entre as isoformas de ECA2 foram realizadas somente dentro dos grupos, entre estruturas referentes à mesma proteína e referência. Os modelos atômicos foram submetidos no software Chimera 1.15 (PETTERSEN et al., 2004) onde foi executado o

protocolo de sobreposição e alinhamento através da ferramenta MatchMaker na seção de ferramentas e comparação de estruturas. Para o alinhamento e sobreposição, todas as sequências foram confrontadas com a sequência de referência correspondente em pares, e somente o alinhamento de mais alta pontuação de cada par foi considerado. O algoritmo de alinhamento utilizado foi o de Needleman-Wunsch e a matriz BLOSUM-62. Foram considerados para os algoritmos os parâmetros como: penalidades de extensão de lacunas em 1, inclusão de estruturas secundárias em 30% com uma matriz de pontuação de 6 nos átomos de hidrogênio de estruturas em intra-hélice, -9 e 6 respectivamente nos átomos de hidrogênio e enxofre em estruturas em intra-fita e -6, -6 e 4 respectivamente nos átomos de hidrogênio, enxofre e oxigênio em outros tipos de estruturas, com penalidades e lacunas abertas para intra-hélice, intra-fita e outras estruturas respectivamente de 18, 18 e 6; e a submissão das atribuições das estruturas secundárias. Para o alinhamento, foram limitados os pares de átomos de forma inteirada para suas distâncias não excederem 2 angstroms. Além disso, os pares foram computados de forma individual após a sobreposição. Para criar o alinhamento a partir da sobreposição, foram consideradas as sequências correspondentes ao par que foi analisado com os parâmetros da distância corte entre resíduos de 5 angstroms e alinhamento do resíduo se este estiver dentro da distância corte de pelo menos um outro resíduo.

3.7. Cálculo das raízes dos desvios quadráticos médios (RMSD) dos alinhamentos das proteínas spike do SARS-CoV-2

A coleta dos dados de RMSD dos alinhamentos das proteínas spike do SARS-CoV-2 foram feitas com a ferramenta “RMSD: ca” utilizando o software Chimera (PETTERSEN et al., 2004). A ferramenta “RMSD: ca” gera um arquivo contendo os dados de RMSD por resíduos que foi transformado em uma tabela para a análise dos dados e os dados de RMSD foram apresentados em formato de gráfico utilizando o software Microsoft excel.

3.8. Interação molecular (docking molecular) entre variantes das proteínas spike de SARS-CoV-2 e isoformas de ECA2

Os modelos das mutações pontuais e as variantes Wuhan, Omicron e BA.2 das proteínas spike do SARS-CoV-2 foram submetidos e confrontados para o docking molecular contra as isoformas de ECA2 na plataforma cluspro2 (KOZAKOV et al., 2017). Foi estabelecido como resíduos de atração de ambas as proteínas as posições representadas na Tabela 6.

Tabela 6. Posição de atração para o docking molecular entre a proteína spike do SARS-CoV-2 e a ECA2.

Variante	Posição de atração da cadeia D da proteína ECA2	Posição de atração da cadeia A da proteína spike do SARS-CoV-2	Referência
S2B	d-24 d-27 d-28 d-30 d-31 d-34 d-37 d-38 d-41 d-42 d-330 d-353 d-354 d-355 d-357	a-417 a-446 a-449 a-453 a-455 a-456 a-475 a-486 a-487 a-489 a-493 a-496 a-498 a-500 a-501 a-502 a-505	(LAN et al., 2020)
S2F	d-24 d-27 d-28 d-31 d-34 d-35 d-38 d-41 d-42 d-79 d-82 d-83 d-330 d-353 d-354 d-355 d-357	a-449 a-453 a-456 a-476 a-486 a-487 a-489 a-493 a-494 a-496 a-498 a-500 a-501 a-502 a-505	(MANNAR et al., 2022)
S2G	d-24 d-27 d-28 d-31 d-34 d-35 d-38 d-41 d-42 d-79 d-82 d-83 d-330 d-353 d-354 d-355 d-357	a-304 a-308 a-311 a-331 a-341 a-344 a-348 a-349 a-351 a-353 a-355 a-356 a-357 a-360*	(MANNAR et al., 2022)
S1E	d-24 d-27 d-28 d-30 d-31 d-34 d-37 d-38 d-41 d-42 d-330 d-353 d-354 d-355 d-357	a-417 a-446 a-449 a-453 a-455 a-456 a-475 a-486 a-487 a-489 a-493 a-496 a-498 a-500 a-501 a-502 a-505	(LAN et al., 2020)
S1F	d-24 d-27 d-28 d-30 d-31 d-34 d-37 d-38 d-41 d-42 d-330 d-353 d-354 d-355 d-357	a-446 a-449 a-453 a-455 a-456 a-475 a-486 a-487 a-489 a-493 a-496 a-498 a-500 a-502 a-505	(LAN et al., 2020)

Tabela 6. Posição de atração para o docking molecular entre a proteína spike do SARS-CoV-2 e a ECA2. Posições de atração para a cadeia D da isoforma 1 de ECA2 (E1A) e posições de atração para a cadeia A das proteínas spike do SARS-CoV-2 das variantes de Wuhan (S2B), Omicron (S2F), BA.2 (S2G), variante de Wuhan com mutação E484K (S1E) e variante de Wuhan com as mutações combinadas E484K, N501Y e K417Y (S1F). *Por causa do modelo atômico referente à proteína spike da subvariante BA.2 do SARS-CoV-2 estar incompleta, as posições precisaram ser adaptadas ao modelo, subtraindo 152 das posições.

Para a coleta dos modelos formados na etapa de docking molecular (complexos moleculares), foi estabelecido como prioritário o melhor modelo que apresentou um perfil balanceado entre interações eletrostáticas e hidrofóbicas. Para esse caso, o cálculo da pontuação para a seleção do melhor modelo foi com base na Equação 1:

$$E = 0,40E_{rep} + (-0,40E_{att}) + 600E_{elec} + 1,00E_{DARS} \quad (1)$$

Onde E_{rep} e E_{att} se referem respectivamente às forças repulsivas e atrativas, E_{elec} é igual à energia eletrostática e E_{DARS} é igual ao potencial referente mudança de energia livre calculada pelo sistema de isca com estruturas de referência (DARS) (CHUANG et al., 2008; KOZAKOV et al., 2017).

3.9. Análise 2D das interações intermoleculares entre as proteínas spike das variantes do SARS-CoV-2 com as isoformas de ECA2

Para a Análise 2D das interações intermoleculares entre as proteínas spike das variantes do SARS-CoV-2 com a proteína ECA2, os complexos formados na etapa de docking molecular foram submetidos em formato PDB no software LigPlot+ versão 2.2 (WALLACE; LASKOWSKI; THORNTON, 1995).

4. RESULTADOS E DISCUSSÃO

4.1. Alinhamento das sequências de aminoácidos das proteínas spike do SARS-CoV-2

Conforme demonstrado na Figura 5, são apresentados esquemas derivados do alinhamento das sequências primárias das proteínas spike do SARS-CoV-2 do com grupo 1 (3 variantes de proteínas spike com mutações pontuais A222V, D614G, E484K e 1 variante com mutações combinadas de 3 aminoácidos, K417Y, E484K e N501Y). Nessa representação esquemática, estão destacadas as regiões com as mutações de cada variante, o domínio RBD (Receptor Binding Domain) e RBM (Receptor Binding Motif) que interage com ECA2. Além disso, é possível observar que as mutações K417Y, E484K e N501Y estão mapeadas dentro da região do RBD.

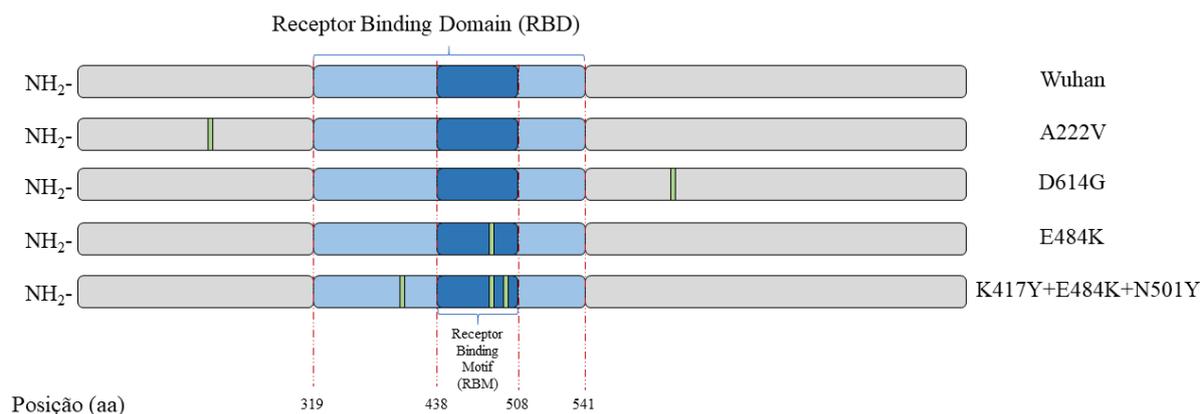


Figura 5. Representação esquemática derivada do alinhamento das sequências de aminoácidos das proteínas spike do SARS-CoV-2 do grupo 1 utilizando o algoritmo MUSCLE. Domínios e tamanho das proteínas spike do SARS-CoV-2 de Wuhan e as variantes do grupo 1 com mutações pontuais A222V, D614G, E484K e combinação K417Y, E484K e N501Y com destaque no Receptor Binding Domain (RBD) em azul claro, Receptor Binding Motif (RBM) em azul escuro e mutações de substituição de aminoácidos representadas em verde. Linhas tracejadas em vermelho fazem indicação à posição do aminoácido (aa) em relação à estrutura. Fonte do autor.

As proteínas do grupo 1 apresentam mutações pontuais cujos potenciais moleculares já foram descritos na literatura. A mutação E484K presente nas variantes Beta, Gamma e Eta promove uma resistência parcial a anticorpos neutralizantes (GARCIA-BELTRAN et al., 2021; WANG et al., 2021a; WIBMER et al., 2021; WU et al., 2021b; XIE et al., 2021), e a combinação de mutações E484K, N501Y e K417Y parece aumentar a transmissibilidade. Nesse contexto, também destacam a mutação K417Y pode reduzir a capacidade de neutralização mediada por anticorpos, uma vez que essa posição do resíduo 417 parece ser um importante alvo imunológico, e ainda a mutação N501Y que está associada a um aumento na interação entre RBD e a ECA2 (ALI; KASRY; AMIN, 2021; GU et al., 2020).

Para as proteínas spike do SARS-CoV-2 do grupo 2, as mutações de substituição, inserção e/ou deleção de aminoácidos foram destacadas na representação esquemática para cada variante de preocupação (VoC) incluindo a análise das variantes Alfa, Beta, Delta, Omicron e a subvariante BA.2, conforme demonstrado na Figura 6 e na Tabela 4. No esquema das sequências das proteínas spike estão destacados o domínio RBD (Receptor Binding Domain) e RBM (Receptor Binding Motif) que interage com ECA2, e ainda, as mutações de substituição estão representadas em verde, as de deleção em laranja e as de

inserção em roxo para cada variante. Conforme demonstrado na Tabela 4 muitas mutações foram observadas nas variantes Omicron BA.1 e BA.2 que compartilham 14 mutações no RBD.

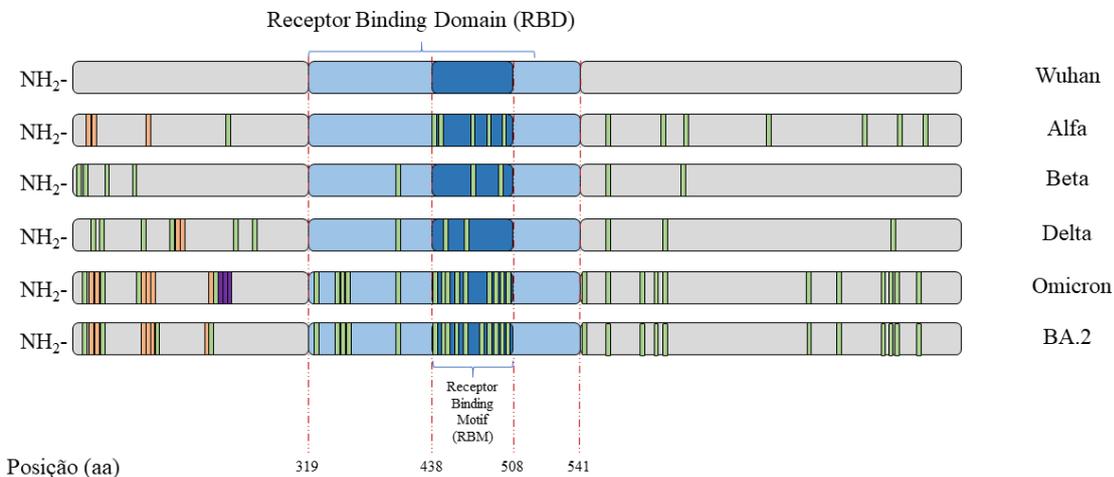


Figura 6. Representação esquemática das sequências de aminoácidos das proteínas spike do SARS-CoV-2 do grupo 2. Domínios e tamanho das proteínas spike do SARS-CoV-2 de Wuhan e as variantes do grupo 2: Alfa, Beta, Delta, Omicron e BA.2 com destaque no Receptor Binding Domain (RBD) em azul claro, Receptor Binding Motif (RBM) em azul escuro, mutações de substituição em verde, de deleção em laranja e de inserção em roxo. Linhas tracejadas em vermelho fazem indicação à posição do aminoácido (aa) em relação à estrutura. Fonte do autor.

As mutações demonstradas na Tabela 4 são mudanças estruturais descritas para a proteína spike que apresentam associação com características biológicas e patogênicas das novas variantes de SARS-CoV-2. Foi observado na COVID-19 o aumento de transmissibilidade e severidade da variante Alfa (DAVIES et al., 2021; FUNK et al., 2021; JANGRA et al., 2021), aumento de transmissibilidade, severidade e impacto no sistema imune da variante beta (CELE et al., 2021; FUNK et al., 2021; MADHI et al., 2021; PEARSON et al., [s.d.]; TEGALLY et al., 2021), aumento de transmissibilidade, severidade e impacto no sistema imune da variante Delta (PUBLIC HEALTH ENGLAND, 2021a, 2021b; SHEIKH et al., 2021a; STOWE et al., 2021) e aumento de transmissibilidade, impacto no sistema imune e uma aparente redução da severidade da variante Omicron (ALTARAWNEH et al., 2022; BAGER et al., 2022; EUROPEAN CENTRE FOR DISEASE PREVENTION AND CONTROL, 2022; FERGUSON et al., 2021; JACOBSEN et al., 2022; LYGSE et al., 2022; PUBLIC HEALTH ENGLAND, 2021c; PULLIAM et al., 2021; RÖSSLER et al.,

2022; SHEIKH et al., 2021b; STATENS SERUM INSTITUT (SSI), 2022; SURYAWANSHI et al., 2022; UK HEALTH SECURITY AGENCY (UKHSA), 2021, 2022; WHO, 2022b).

Os modelos atômicos das proteínas spike de SARS-CoV-2 com as mutações descritas foram gerados neste trabalho e foram utilizadas para avaliação do cálculo das raízes dos desvios quadráticos médios (RMSD) descritos no item 4.4.

4.2. Alinhamento das sequências de aminoácidos das isoformas de ECA2

Como mostra a Figura 7, as sequências de aminoácidos derivadas do alinhamento da estrutura primária das isoformas 1, 3, 4, 5 e 6 de ECA2 foram esquematizadas. Nessa representação, estão destacadas: a região de ligação da peptidase M2 em azul (19-606), a região homóloga a colectrina em verde (617-770) e a região de maior variabilidade entre as isoformas em cores variadas (770-805). O alinhamento das sequências de aminoácidos das isoformas de ECA2 demonstrou que na comparação entre as isoformas 1, 3, 4, 5 e 6, as diferenças estão apenas na porção C-terminal da molécula. A região C-terminal dessas isoformas de ECA2 não contempla os resíduos Gln24, Lys31, His34, Glu35, Asp38 e Lys353 envolvidos na interação molecular com a proteína spike do SARS-CoV-2 (ALI; VIJAYAN, 2020).

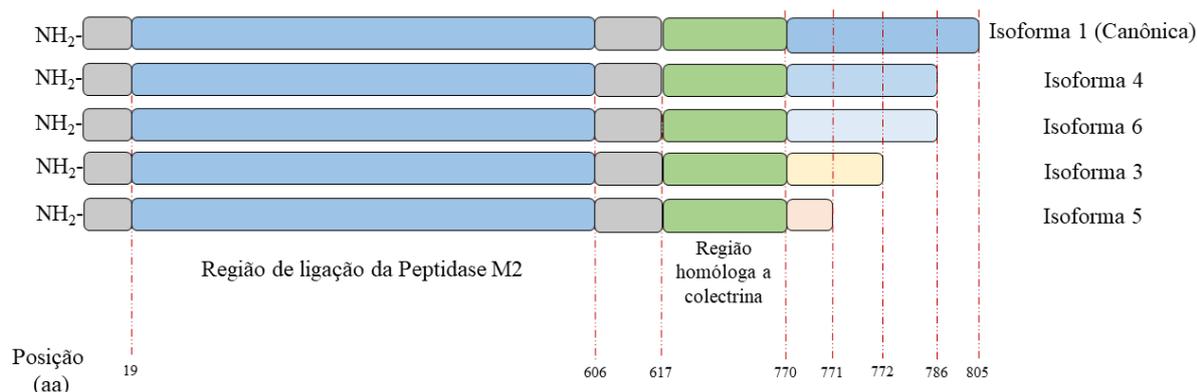


Figura 7. Representação esquemática das sequências de aminoácidos das isoformas de ECA2 (1, 3, 4, 5 e 6) em ordem decrescente de tamanho. Domínios e tamanhos da proteína ECA2 e de suas isoformas com destaque nas regiões compartilhadas: de ligação da Peptidase M2 entre as posições 19 e 606, região homóloga a colectrina entre as posições 617 e 770 e nas regiões não compartilhadas: a partir da posição de 770 aminoácidos. Linhas tracejadas em vermelho fazem indicação à posição do aminoácido (aa) em relação à estrutura. Fonte do autor.

Além disso, quando são observadas as diferenças entre as isoformas 1, 2 e 7 de ECA2 no esquema representativo que foi derivado do alinhamento das estruturas primárias dessas isoformas (Figura 8), nota-se que as isoformas 2 e 7 não possuem os resíduos Gln24, Lys31, His34, Glu35, Asp38 e Lys353, já que estas apresentam uma deleção amino terminal até os aminoácidos 365 e 605 respectivamente, em comparação a isoforma 1. Sendo assim, as isoformas 2 e 7 não apresentam os resíduos de interação para a ligação molecular entre o RBD da proteína spike do SARS-CoV-2. Portanto, neste trabalho, não foram consideradas para a realização dos docking moleculares com as proteínas spike do SARS-CoV-2 as isoformas 2 e 7 de ECA2.

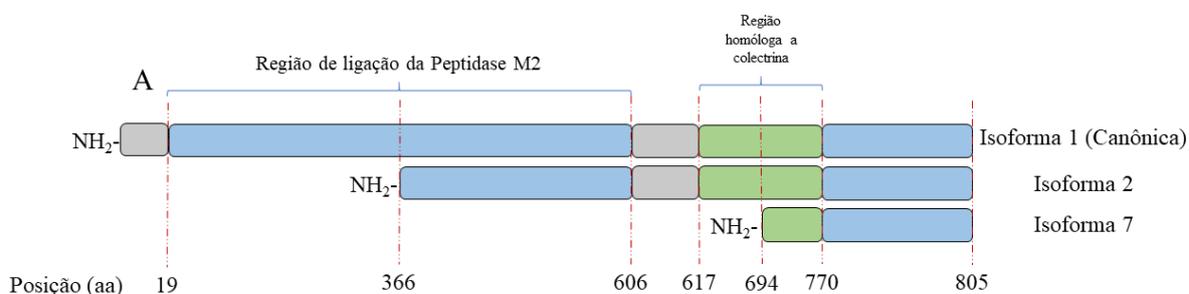


Figura 8. Representação esquemática das sequências de aminoácidos das isoformas de ECA2 (1, 2 e 7) em ordem decrescente de tamanho. Domínios e tamanhos da proteína ECA2 e de suas isoformas com destaque nas regiões compartilhadas: de ligação da Peptidase M2 entre as posições 19 e 606 e na região homóloga a colectrina entre as posições 617 e 770. Linhas tracejadas em vermelho fazem indicação a posição do aminoácido (aa) em relação à estrutura. Fonte do autor.

Na análise descritiva das sequências de aminoácidos das isoformas de ECA2, foi observado que as regiões de polimorfismos desta molécula não contemplam as regiões de interação com o RBD da proteína spike do SARS-CoV-2. Dessa forma, é provável que não haja diferenças nas interações intermoleculares entre a proteína spike do SARS-CoV-2 com as diferentes isoformas de ECA2 humanas avaliadas neste trabalho. No entanto, somente a análise na sequência primária de aminoácidos das isoformas de ECA2 não é suficiente para averiguar tal suposição.

As estruturas moleculares das isoformas de ECA2 foram modeladas e utilizadas para avaliação de RMSD descritas a seguir.

4.3. Modelagem das proteínas spike do SARS-CoV-2 e isoformas de ECA2

Como mostra a Figura 9, são apresentados os modelos atômicos das proteínas spike do SARS-CoV-2 da estrutura original de Wuhan do grupo 1 (9A), das variantes com mutação pontual A222V (9B), D614G (9C), E484K (9D) e com mutações combinadas N417Y, E484K e N501Y (9E) em forma de fitas e com acessibilidade a superfície. Conforme o esperado devido à semelhança nas sequências de aminoácidos, as variantes apresentaram em suas estruturas moleculares conformações parecidas ao modelo S1B (7CAK-A) utilizado como referência.

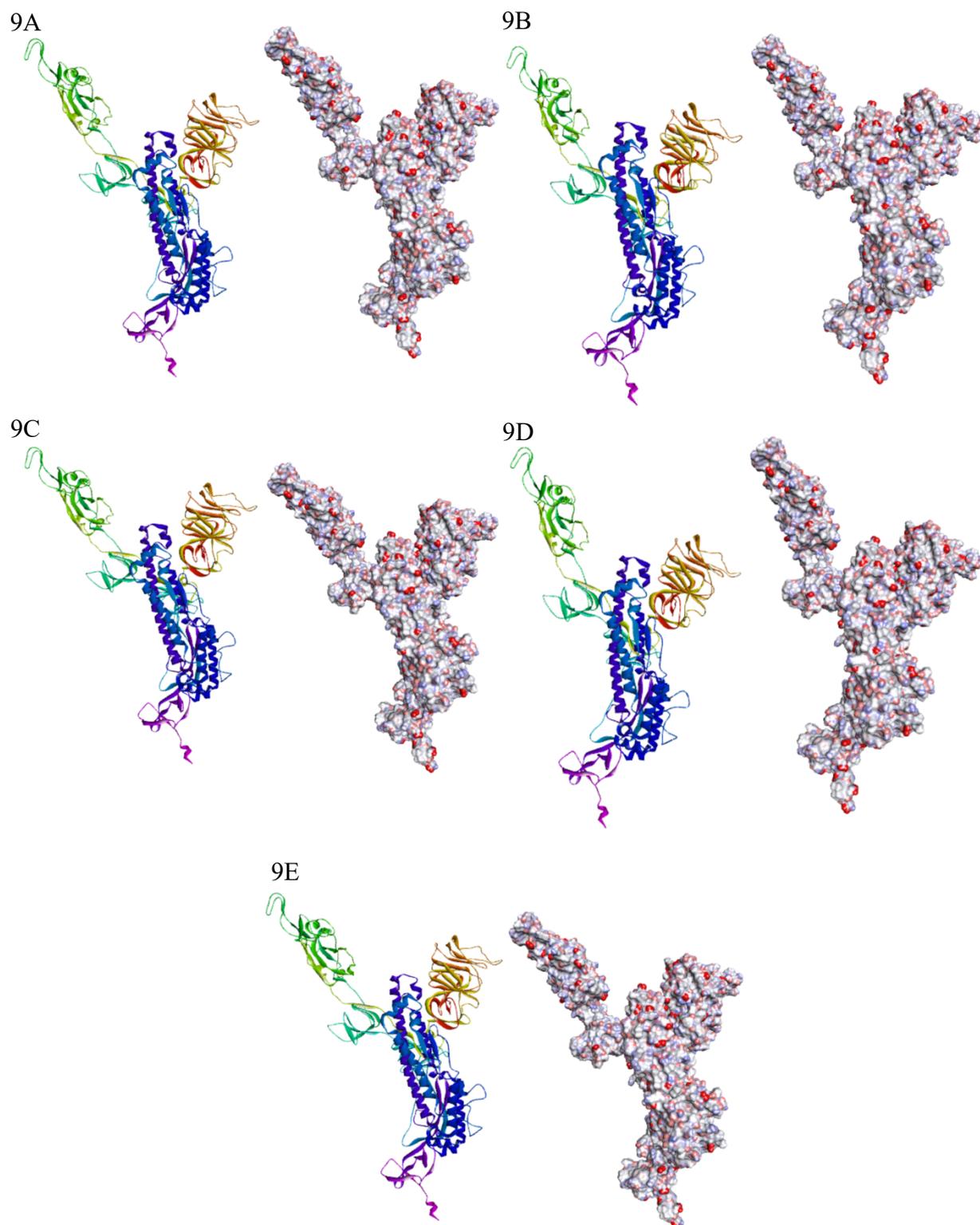


Figura 9. Representações tridimensionais em formato de fita à esquerda e com acessibilidade do solvente à superfície à direita, das proteínas spike do SARS-CoV-2 (grupo 1) das variantes de Wuhan (9A), com: mutação A222V (9B), D614G (9C), E484K (9D) e E484K, N501Y e K419Y combinadas (9E). As cores vermelha e azul das figuras da direita representam as cargas parciais dos átomos respectivamente positiva e negativa e variam de acordo com a intensidade da carga.

Como mostra a Figura 10, são apresentados os modelos atômicos das proteínas spike do SARS-CoV-2 da estrutura original de Wuhan do grupo 2 (10A), das variantes Alfa (10B),

Beta (10C), Delta (10D), Omicron (10E) e subvariante BA.2 (10F) em forma de fitas e com acessibilidade a superfície. Por mais que existam um acúmulo de mutações nas variantes Omicron e subvariante BA.2, como esperado, essas mutações não causaram mudanças conformacionais substanciais na estrutura 3D das proteínas. Somente a ausência de 145 aminoácidos na região amino terminal do modelo da proteína spike do SARS-CoV-2 da subvariante BA.2 justifica a mudança conformacional observada na figura 10F. Essa ausência está relacionada com uma limitação no processo de modelagem utilizado para essa variante da proteína spike do SARS-CoV-2.

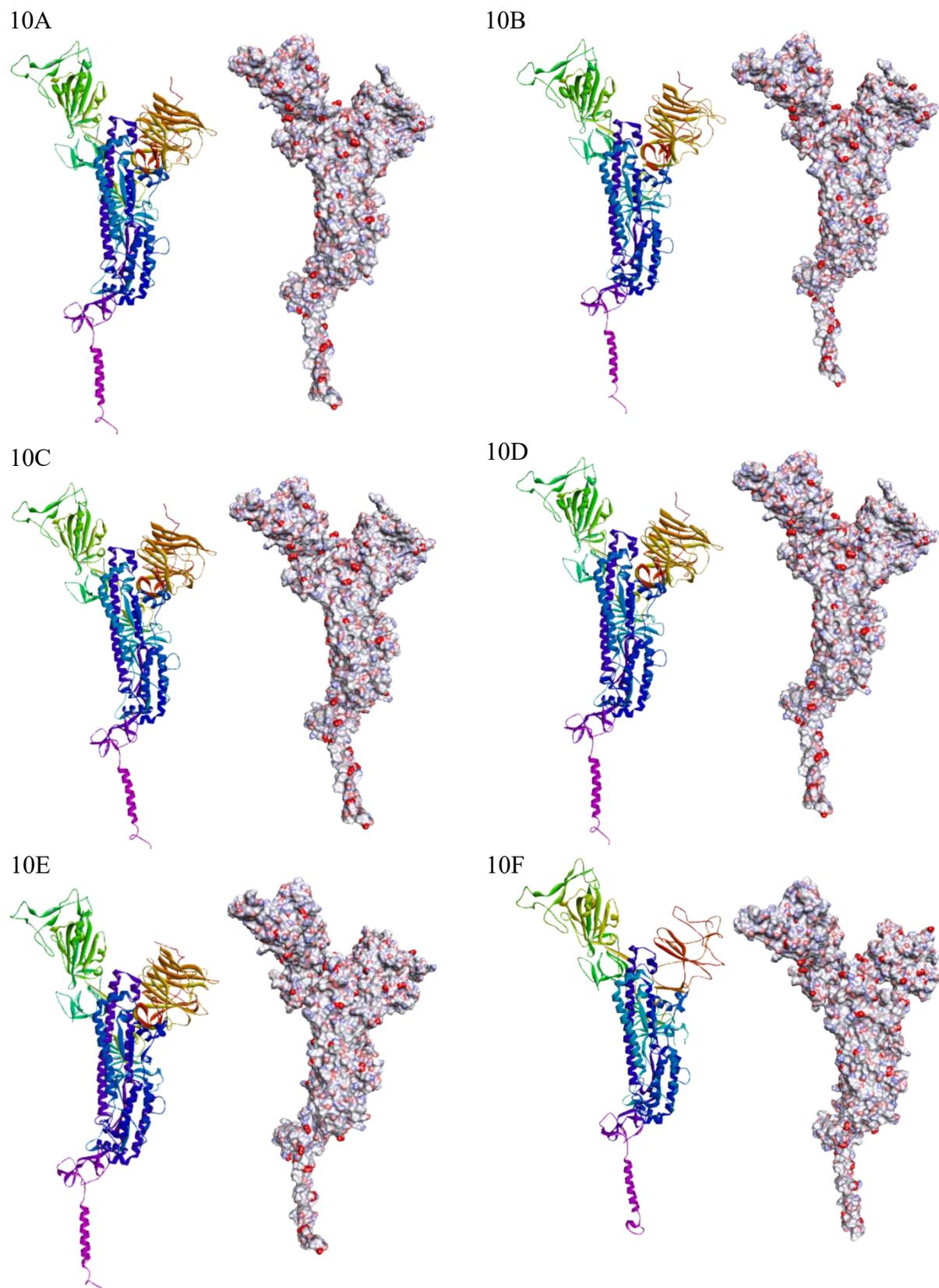


Figura 10. Representações tridimensionais em formato de fita à esquerda e com acessibilidade do solvente a superfície à direita, das proteínas spike do SARS-CoV-2 (grupo 2) das VoC de Wuhan (10A), Alfa (10B), Beta (10C), Delta (10D) e Omicron, BA.2 (10E). As cores vermelha e azul das figuras da direita representam as cargas parciais dos átomos respectivamente positiva e negativa e variam de acordo com a intensidade da carga.

Como mostra a Figura 11, são apresentados os modelos atômicos das isoformas 1 (11A), 3 (11B), 4 (11C), 5 (11D) e 6 (11E) de ECA2. Conforme o esperado devido à alta similaridade dessas isoformas, a modelagem gerou estruturas conformacionais semelhantes. Por corresponder à parte de maior variabilidade entre as isoformas, a extremidade carboxi terminal é a única região que apresentou conformação distinta. Por outro lado, causado por fatores inerentes ao processo de modelagem, a extremidade amino terminal também apresentou pequenas mudanças conformacionais entre as isoformas modeladas. Este resultado foi inesperado uma vez que as sequências primárias das isoformas de ECA2 nessa região são iguais e abordagens futuras serão empregadas para elucidar estas diferenças observadas entre os modelos estruturais destas proteínas.

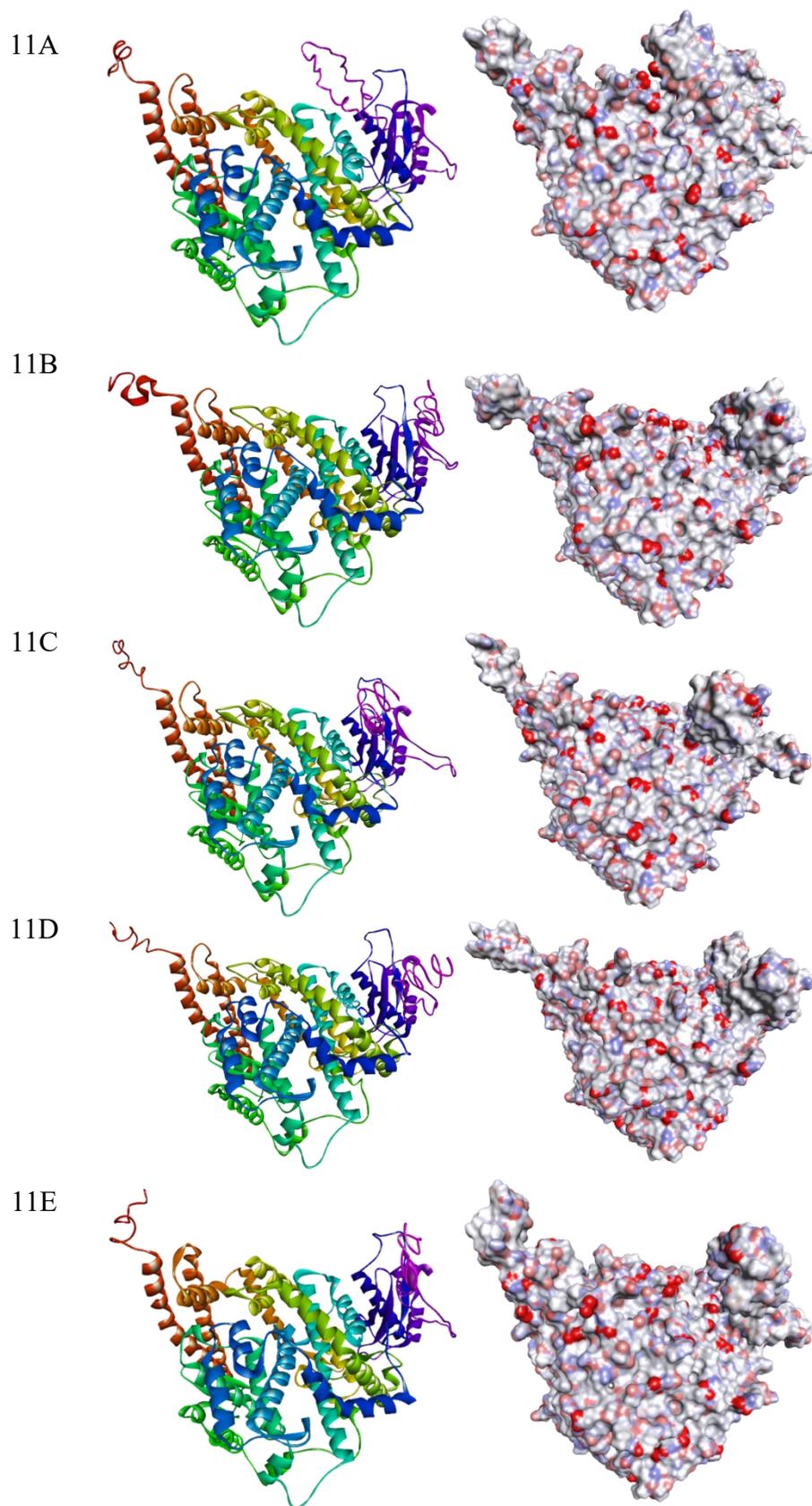


Figura 11. Representações tridimensionais em formato de fita à esquerda e com acessibilidade do solvente a superfície à direita das isoformas 1 (11A), 3 (11B), 4 (11C), 5 (11D) e 6 (11E) de ECA2 refinadas. As cores vermelha e azul das figuras da direita representam as cargas parciais dos átomos respectivamente positiva e negativa e variam de acordo com a intensidade da carga. As estruturas foram modeladas a partir da sequência de

aminoácidos das isoformas 1 (BLUME et al., 2021), isoforma 2 (BLUME et al., 2021; ONABAJA et al., 2020) e isoformas 3, 4, 5, 6 e 7 extraídas do banco de dados UniprotKB (BATEMAN et al., 2021).

Como mostra a Figura 12, são apresentados os modelos atômicos das isoformas 1 (12A), 2 (12B) e 7 (12C) de ECA2. Conforme o esperado, em função da diferença de tamanhos entre as isoformas, a modelagem gerou estruturas conformacionais muito distintas. Essas diferenças estruturais observadas entre as isoformas 1, 2 e 7 de ECA2 foram resultado da ausência de 365 aminoácidos na isoformas 2 e 695 aminoácidos na isoforma 7 em comparação a isoforma 1 (*full length*). Estas deleções descritas para as isoformas 2 e 7 estão localizadas em ambos os casos na região amino terminal da molécula ECA2 (Figura 8).

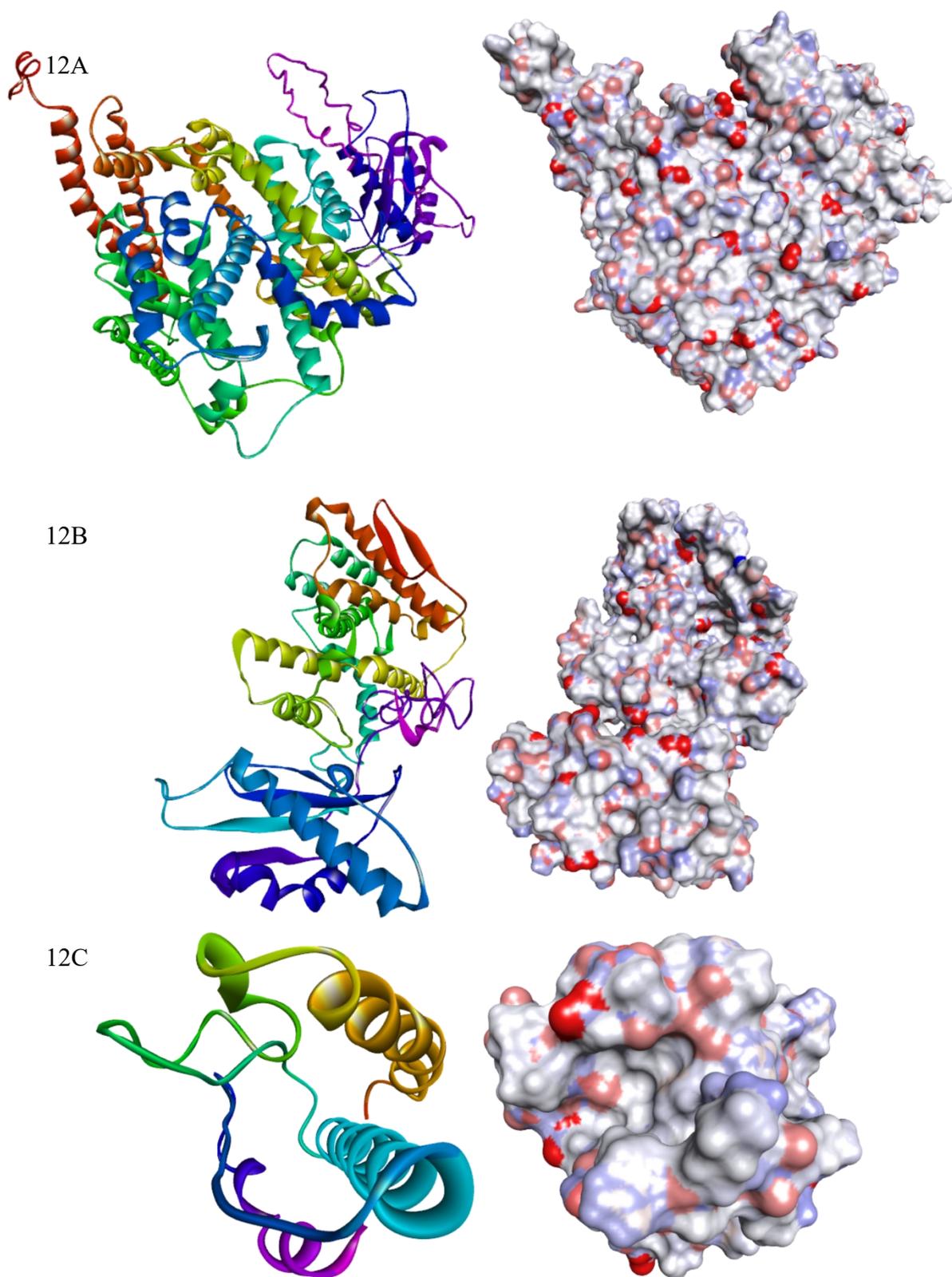


Figura 12. Representações tridimensionais em formato de fita à esquerda e com acessibilidade do solvente a superfície à direita das isoformas 1 (12A), 2 (12B) e 7 (12C) de ECA2. As cores vermelha e azul das figuras da direita representam as cargas parciais dos átomos respectivamente positiva e negativa e variam de acordo com a intensidade da carga. As estruturas foram modeladas a partir da sequência de aminoácidos das isoformas 1 (BLUME et al., 2021), isoforma 2 (BLUME et al., 2021; ONABAJO et al., 2020) e isoformas 3, 4, 5, 6 e 7 extraídas do banco de dados UniprotKB (BATEMAN et al., 2021).

Com as estruturas das proteínas spike de SARS-CoV-2 e isoforma 1 de ECA2 modeladas, foram feitas análises de RMSD, de qualidade das estruturas modeladas e a análise de docking molecular.

4.4. RMSD dos alinhamentos estruturais do grupo 1 e 2 da proteína spike do SARS-CoV-2 e isoformas de ECA2

Os dados de RMSD por resíduos dos modelos atômicos das proteínas spike do SARS-CoV-2 da estrutura original de Wuhan (S1B) e das variantes com mutações pontuais A222V (S1C), D614G (S1D), E484K (S1E), com mutações combinadas N417Y, E484K e N501Y (S1F), da variante Alfa (S2C), Beta (S2D), Delta (S2E), Omicron (S2F), da subvariante BA.2 (S2G) e das isoformas 1 (E1A), 2 (E1B), 3 (E1C), 4 (E1D), 5 (E1E), 6 (E1F) e 7 (E1G) de ECA2 foram coletados e plotados em gráfico conforme demonstrado nas Figura 13, Figura 14 e Figura 15.

Na Figura 13, são mostrados os dados de RMSD em angstrom em função da posição de resíduo das comparações dos modelos moleculares das proteínas spike do SARS-CoV-2 da estrutura original de Wuhan do grupo 1 (S1B) com as das variantes com mutação pontual A222V (S1C), D614G (S1D), E484K (S1E) e com mutações combinadas N417Y, E484K e N501Y (S1F). Os dados de RMSD das comparações S1B x S1C (modelo original de Wuhan x variante spike com A222V) e S1B x S1D (modelo original de Wuhan x variante spike com D614G) foram baixos (entre 0 e 0,4 Å), enquanto para as comparações S1B x S1E (modelo original de Wuhan x variante spike com E484K) e S1B x S1F (modelo original de Wuhan x variante spike com K417Y, E484K e N501Y) foi observado maior mudança no perfil (>0,4 Å). A mudança no perfil de RMSD causada pela presença da mutação E484K corrobora com seu potencial na redução da neutralização por anticorpos na COVID-19 (GREANEY et al., 2021; JANGRA et al., 2021; WANG et al., 2021b; WEISBLUM et al., 2020; WU et al.,

2021a). Por outro lado, a mudança no perfil de RMSD apresentada pela presença das mutações K417Y e N501Y na comparação S1B x S1F se relaciona com a capacidade dessas mutações em aumentar a afinidade de ligação com ECA2 (WROBEL et al., 2022).

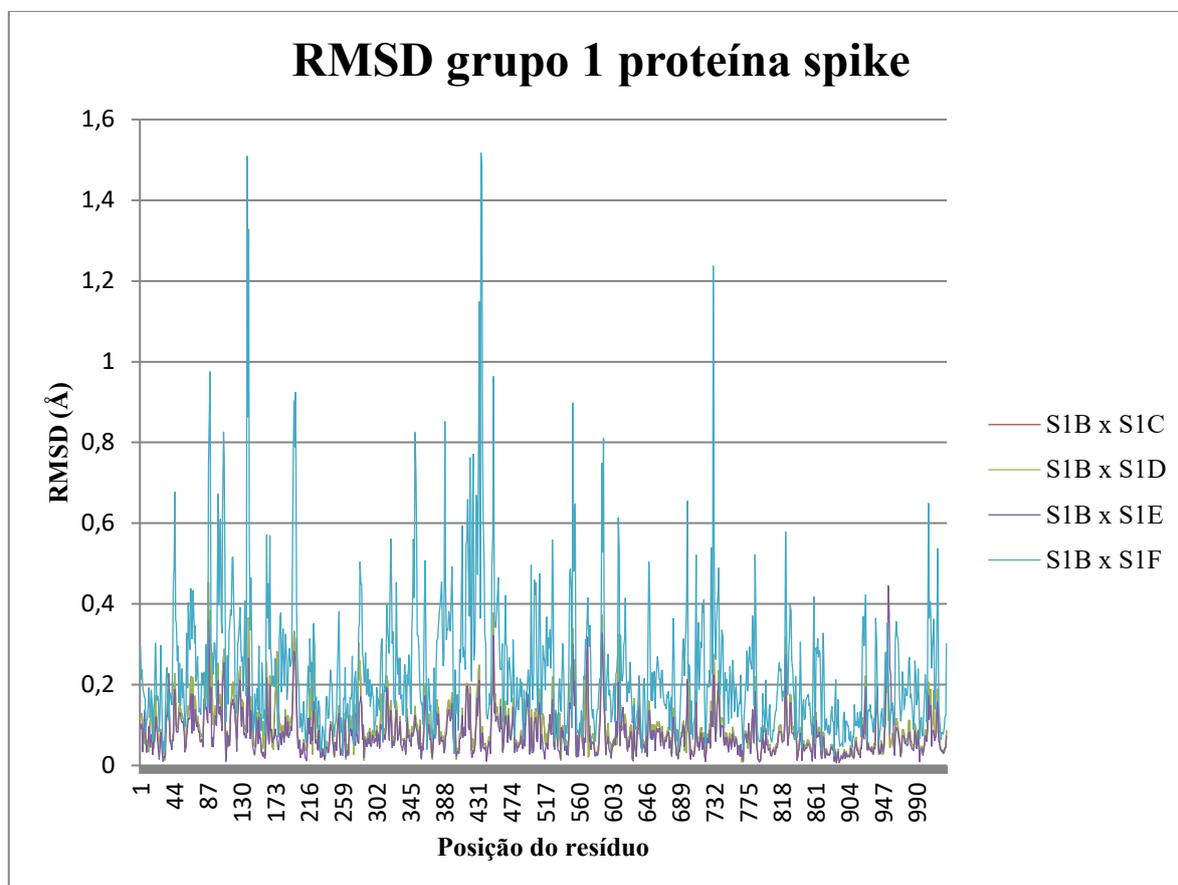


Figura 13. Desvios quadráticos médios (RMSD) em Å por resíduos dos alinhamentos entre a proteína spike do SARS-CoV-2 de Wuhan do grupo 1 (S1B) com as variantes de mutação pontual A222V (S1C), D614G (S1D), E484K (S1E) e a combinação E484K, N501Y e K417Y (S1F).

Conforme a Figura 14, são mostrados e plotados os dados de RMSD em angstrom em função da posição de resíduo das comparações dos modelos moleculares das proteínas spike do SARS-CoV-2 da estrutura original de Wuhan do grupo 2 (S2B) com as das variantes Alfa (S1C), Beta (S1D), Delta (S1E), Omicron (S1F) e da subvariante BA.2 (S1G). O perfil de RMSD das combinações S2B x S2F e S2B x S2G apresentaram diferenças substanciais em comparação com as demais combinações, isso se deve ao grande acúmulo de mutações nas variantes Omicron e a subvariante BA.2. Adicionalmente, foi possível constatar altíssimos valores de RMSD nas análises comparativas S2B x S2F e S2B x S2G fora da região da RBD,

e altos valores de RMSD dentro da região RBD (319-541 em relação à sequência original de Wuhan) da proteína spike.

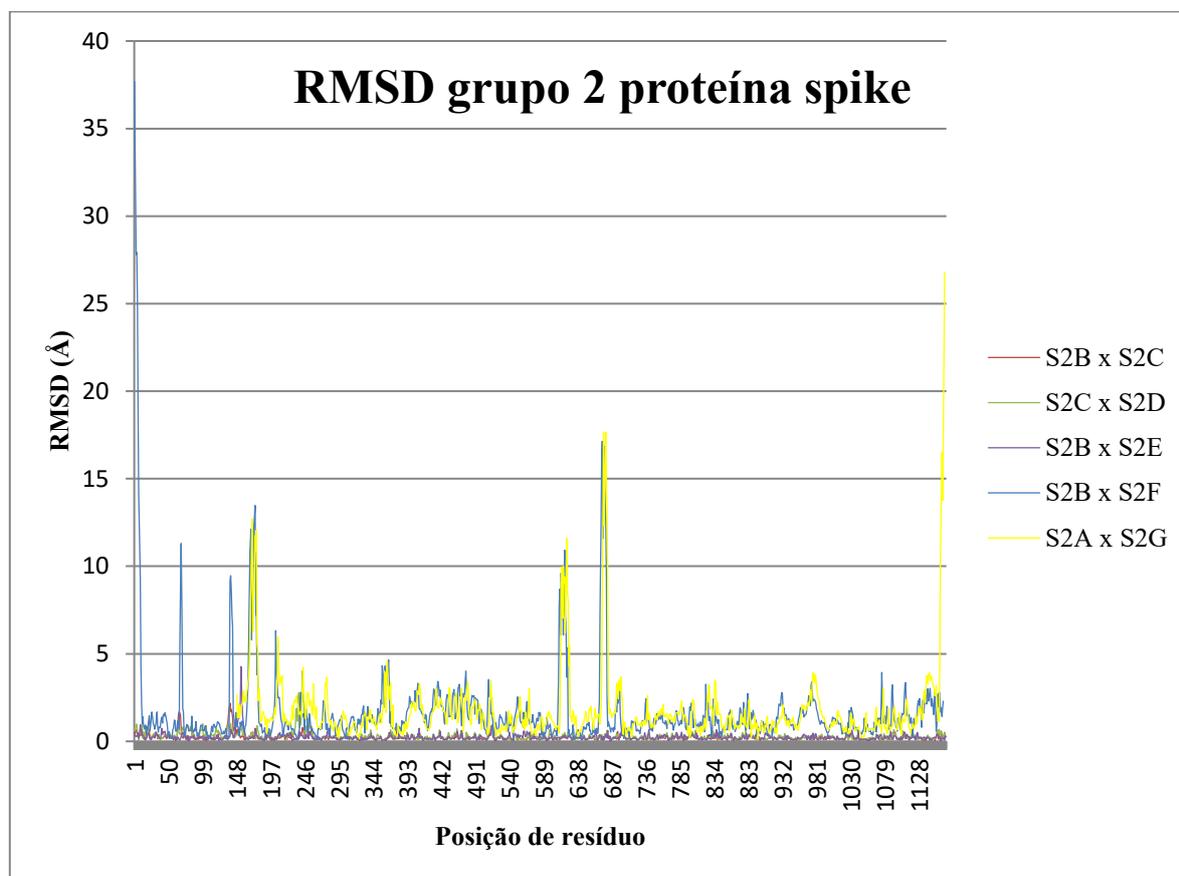


Figura 14. Desvios quadráticos médios (RMSD) em Å por resíduos dos alinhamentos entre a proteína spike do SARS-CoV-2 de Wuhan do grupo 2 (S2B) com as variantes Alfa (S2C), Beta (S2D), Delta (S2E), Omicron (S2F) e BA.2 (S2G).

Conforme a Figura 15 são mostrados e plotados os dados de RMSD em angstrom em função da posição de resíduo das comparações dos modelos moleculares da isoforma 1 (E1A) de ECA2, com as demais isoformas 2 (E1B), 3 (E1C), 4 (E1D), 5 (E1E), 6 (E1F) e 7 (E1G). Os perfis de RMSD de todas as comparações foram semelhantes entre si na região de interação da ECA2, que compreende as regiões ao redor dos resíduos Gln24, Lys31, His34, Glu35, Asp38 e Lys353 (ALI; VIJAYAN, 2020), com a proteína spike do SARS-CoV-2. Isso sugere que não deve haver diferenças estruturais substanciais na estrutura das moléculas de ECA2, e, portanto, as interações intermoleculares entre a proteína spike do SARS-CoV-2 e as isoformas de ECA2 devem ser similares entre as isoformas desta molécula, envolvendo em especial os mesmos resíduos de aminoácidos na interface de interação proteína-proteína.

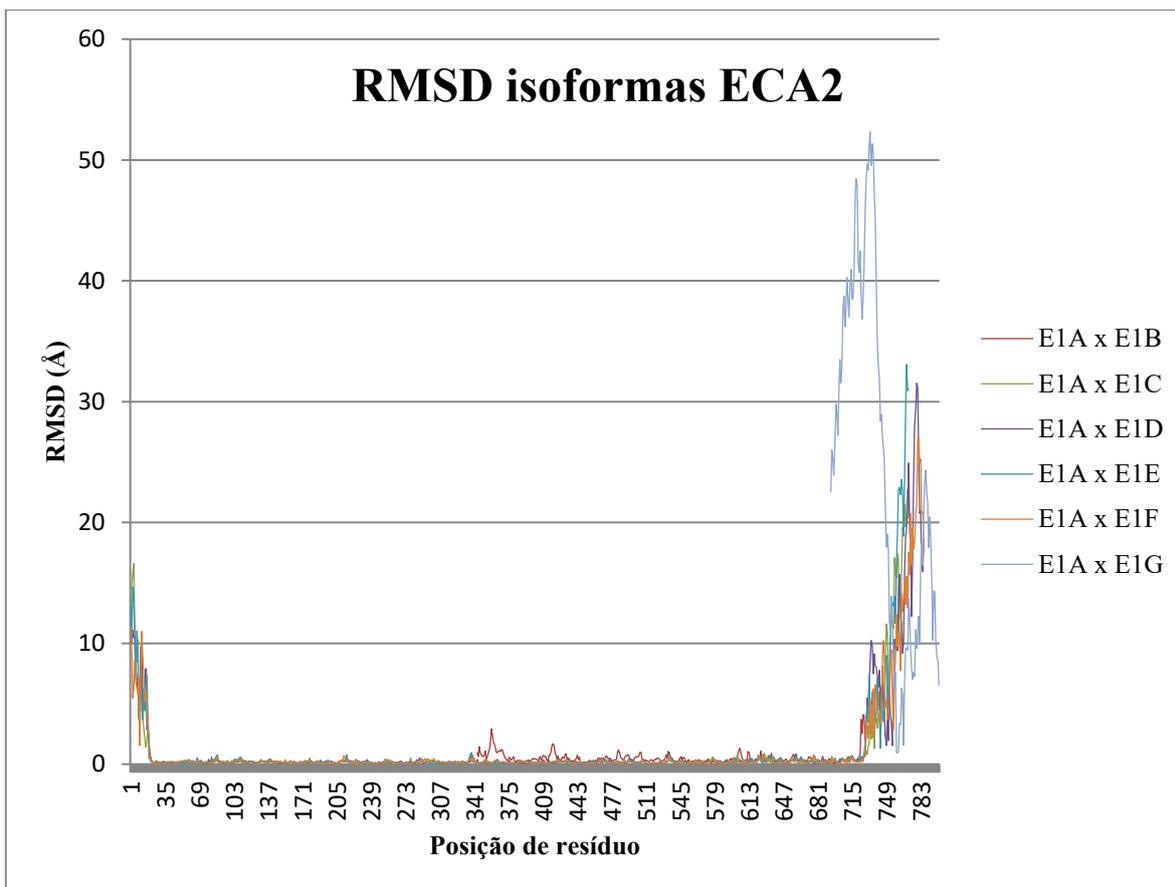


Figura 15. Desvios quadráticos médios (RMSD) por resíduos dos alinhamentos entre a isoforma 1 (E1A) contra as isoformas 2 (E1B), 3 (E1C), 4 (E1D), 5 (E1E), 6 (E1F) e 7 (E1G) de ECA2.

Com os modelos atômicos selecionados das proteínas spike do SARS-CoV-2 e isoforma 1 da ECA2, foram realizadas avaliações de qualidade e docking molecular.

4.5. Avaliação de qualidade dos modelos atômicos

4.5.1. Molprobit

A ferramenta Molprobit consiste em uma plataforma de navegador para analisar a qualidade estereoquímica global e local de proteínas e ácidos nucleicos. Para isso, o serviço inicia com uma adição de hidrogênios na molécula e inversão de resíduos, seguida de uma avaliação de todos os contatos entre os átomos. A adição de hidrogênio tanto polar quanto não polar é um importante pré-requisito para a posterior análise de contato, pois a maioria das estruturas cristalografadas ainda não incorporam átomos de hidrogênio. Para tal, a plataforma se utiliza do software REDUCE (WORD et al., 1999a) que adiciona e otimiza hidrogênios nas

posições favoráveis, com exceção de H em moléculas de água. Além disso, a plataforma age em inversões de resíduos de Asn, Gln e His, que comumente são identificados com erros em orientações, devido à densidade eletrônica ter sido utilizada para definir as suas orientações corretas, variável essa que nem sempre é a ideal para definir orientações de aminoácidos em uma cadeia peptídica. Após esta etapa, segue-se para a análise de todos os contatos entre os átomos por meio do software PROBE (WORD et al., 1999b) que permite identificar a existência de clash através da construção de uma esfera com diâmetro de 0,5 angstroms que rola na superfície de van der Waals dos átomos medindo a sobreposição de átomos não ligados. Um contato entre átomos não doadores-aceptores com uma sobreposição de 0,4 angstroms caracteriza um choque grave (“serious clash”) estereoquímico. Dentre os resultados, a plataforma gera a pontuação de clash que equivale ao número de confrontos ou serious clashes por 1000 átomos e a pontuação Molprobity que junta as informações de rotâmeros, gráficos de Ramachandran e pontuações de clash, normalizando essas variáveis a mesma unidade de medida de resolução de estruturas, Angstroms (CHEN et al., 2010).

Com a ferramenta Molprobity, se obteve o refinamento com inversões de resíduos de asparagina, glutamina e histidina.

Conforme a

Tabela 7, são detalhados os resíduos dos modelos atômicos das proteínas spike do SARS-CoV-2 das variantes com mutação pontual E484K (S1E), com mutações combinadas N417Y, E484K e N501Y (S1F), da estrutura original de Wuhan (S2B) e da variante Omicron (S2F), da subvariante BA.2 (S2G) e Isoforma 1 de ECA2 (E1A) que necessitam de inversões estereoquímicas, suas orientações de origem e após a inversão.

Tabela 7. Aminoácidos e respectivas orientações antes e após refinamento.

	Modelo	Aminoácidos	Origem	Após inversão
Proteínas spike	S1E	Asn (196)	-2,4	-0,78

do SARS-CoV-2 grupo 1	S1F	Asn (61), Asn (196), Asn (282), Asn (616)	-1,1; -1,8; -1,5; -0,51	0; -0,52; 0; 0
Proteínas spike do SARS-CoV-2 grupo 2	S2B	Asn (17), Asn (282)	-1,5; -0,67	0,023; -0,17
	S2F	-	-	-
	S2G	-	-	-
Proteínas ECA2	E1A	Asn (49), Gln (76), Gln (96), Gln (442)	-2,2; -0,69; -0,24; 0,82	3,5; 0,98; 3,2; 1,9

Tabela 7. Aminoácidos e respectivas orientações antes e após refinamento. Inversão conformacional dos aminoácidos em relação ao plano do eixo da cadeia principal dos modelos atômicos das proteínas spike do SARS-CoV-2 das variantes com mutação pontual E484K (S1E), com mutações combinadas N417Y, E484K e N501Y (S1F), da estrutura original de Wuhan (S2B) e da variante Omicron (S2F), da subvariante BA.2 (S2G) e Isoforma 1 de ECA2.

Conforme a Tabela 8, foram detalhadas as pontuações de clash e Molprobability das estruturas moleculares das proteínas spike do SARS-CoV-2 das variantes com mutação pontual E484K (S1E), com mutações combinadas N417Y, E484K e N501Y (S1F), da estrutura original de Wuhan (S2B) e da variante Omicron (S2F), da subvariante BA.2 (S2G) e isoforma 1 de ECA2 (E1A). Somente o modelo E1A apresentou pontuações de baixa qualidade. Isso provavelmente se deve a não realização do protocolo de minimização de energia da estrutura molecular da isoforma 1 de ECA2. As limitações computacionais de processamento de dados nos impediram de realizar esta abordagem de minimização de energia da estrutura para isoforma 1 de ECA2 até o presente momento.

Tabela 8. Pontuações de clash e de Molprobability relativas para os modelos atômicos da proteína spike do SARS-CoV-2 e isoforma 1 de ECA2.

	Modelo	Pontuação de clash (%)	Pontuação Molprobability (%)
Proteínas spike do SARS-CoV-2 grupo 1	S1E	100	100
	S1F	100	100
Proteínas spike do SARS-CoV-2 grupo 2	S2B	99	99
	S2F	99	99
	S2G	99	100
Isoforma 1 de ECA2	E1A	31	12

Tabela 8. Pontuações de clash e de Molprobability relativas para os modelos atômicos da proteína spike do SARS-CoV-2 e isoforma 1 de ECA2. Proteínas spike do SARS-CoV-2 das variantes com mutação pontual E484K (S1E), com mutações combinadas N417Y, E484K e N501Y (S1F), da estrutura original de Wuhan (S2B) e da variante Omicron (S2F), da subvariante BA.2 (S2G) e Isoforma 1 de ECA2. Valores de pontuação p de clash e molprobability $p > 66\%$ indicam uma alta qualidade, $33\% < p < 66\%$ indicam qualidade mediana e $p < 33\%$ indicam baixa qualidade do modelo.

4.5.2. VERIFY3D

A ferramenta VERIFY3D é capaz de determinar a compatibilidade dos modelos atômicos (3D) com suas próprias sequências de aminoácidos (1D). Para isso, o software se baseia em um sistema de pontuação que se refere à condição mais favorável estereoquimicamente do resíduo no ambiente e é calculado pela soma de 3 parâmetros de cada um dos resíduos em faixas de 20 aminoácidos. Os parâmetros consistem em: área do resíduo inacessível ao solvente, fração da área da cadeia lateral coberta por átomos polares como O e N, e a estrutura secundária local. Por fim, os parâmetros são comparados com modelos de alta qualidade e então atribuído uma pontuação 3D-1D (BOWIE; LÜTHY; EISENBERG, 1991; LUTHY; BOWIE; EISENBERG, 1992).

Com a ferramenta VERIFY3D, se obteve uma avaliação da compatibilidade 3D-1D dos modelos atômicos das estruturas referentes à proteína spike do SARS-CoV-2 de referência 7CAK-A (Figura 16A) e 7K43-E (Figura 17A), variante de mutação pontual E484K (Figura 16B), de mutações combinadas E484K, N501Y e K417Y (Figura 16C), estrutura molecular de Wuhan do grupo 2 (Figura 17A), variante Omicron (Figura 17B) e BA.2 (Figura 17C) e isoforma 1 de ECA2 (Figura 18). Os dados das pontuações foram sintetizados na Tabela 9 e Tabela 10. Pontuações médias 3D-1D $\geq 0,2$ indicam alta qualidade do segmento e a contemplação de pelo menos 80% dos segmentos acima desse valor apontam para uma avaliação positiva do modelo.

Como aponta a Figura 16, as pontuações de compatibilidade 3D-1D dos modelos atômicos das proteínas spike do SARS-CoV-2 do grupo 1 da referência 7CAK-A (16A), e das variantes com mutação pontual E484K (16B) e combinadas K417Y, E484K e N501Y (16C) são plotadas em gráficos em função dos resíduos. É viável aferir que os modelos apresentaram perfis semelhantes entre si.

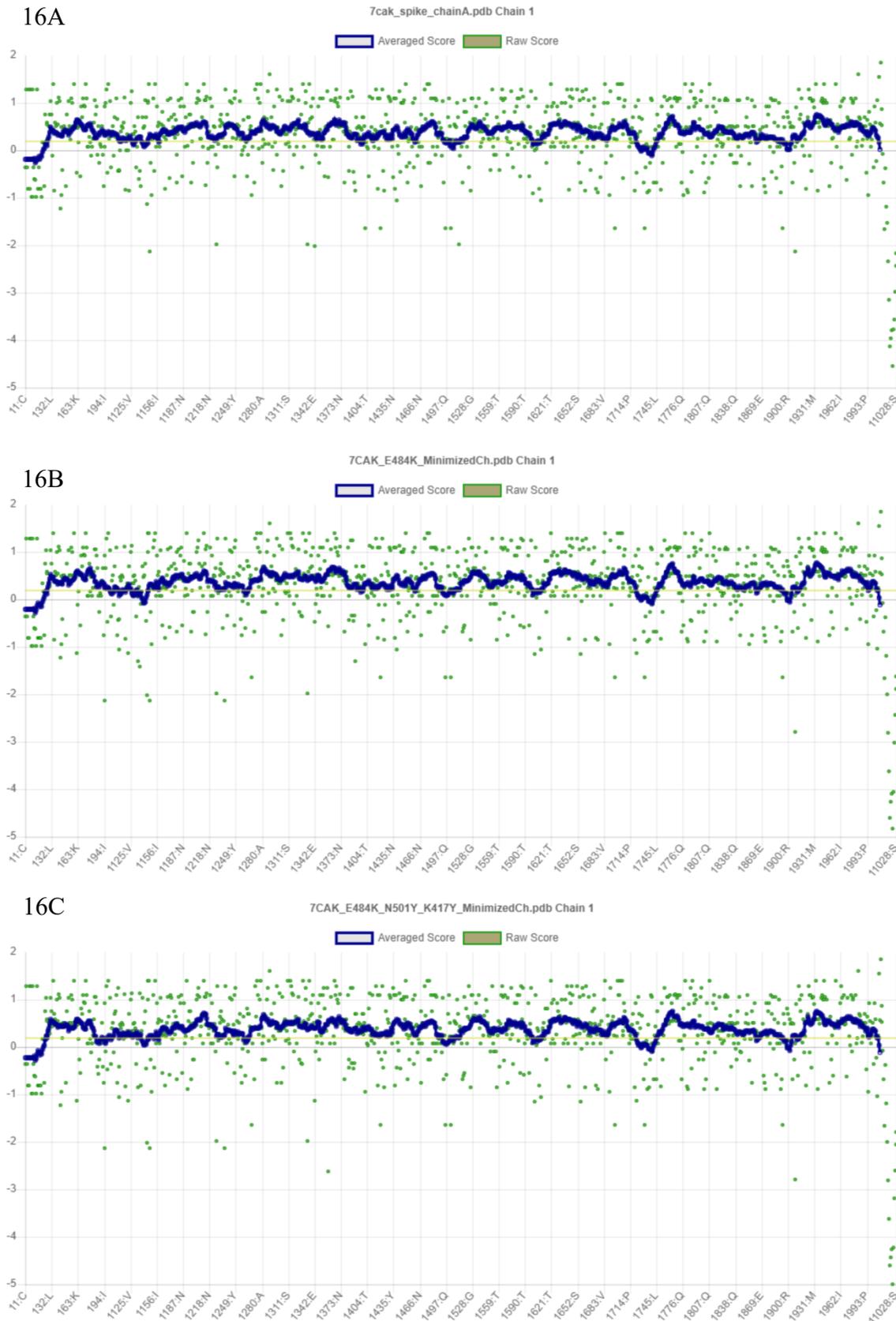


Figura 16. Análise de compatibilidade 3D-1D dos modelos atômicos das proteínas spike do SARS-CoV-2 do grupo 1 da referência 7CAK-A (16A), e das variantes com mutação pontual E484K (16B) e combinadas K417Y,

E484K e N501Y (16C). O segmento azul representa o resultado médio e os pontos verdes representam o resultado bruto da pontuação 3D-1D em função dos resíduos.

Como aponta a Figura 17, as pontuações de compatibilidade 3D-1D dos modelos atômicos das proteínas spike do SARS-CoV-2 do grupo 1 da referência 7K43-E (17A), da estrutura original de Wuhan (17B), da Omicron (17C) e da BA.2 (17D) são plotadas em gráficos em função dos resíduos. Os modelos das proteínas spike do SARS-CoV-2 de referência e da Omicron apresentaram um perfil de baixa compatibilidade 3D-1D na região carboxi terminal, enquanto é viável aferir que os modelos das proteínas spike do SARS-CoV-2 da referência e da subvariante BA2 apresentaram perfis semelhantes.

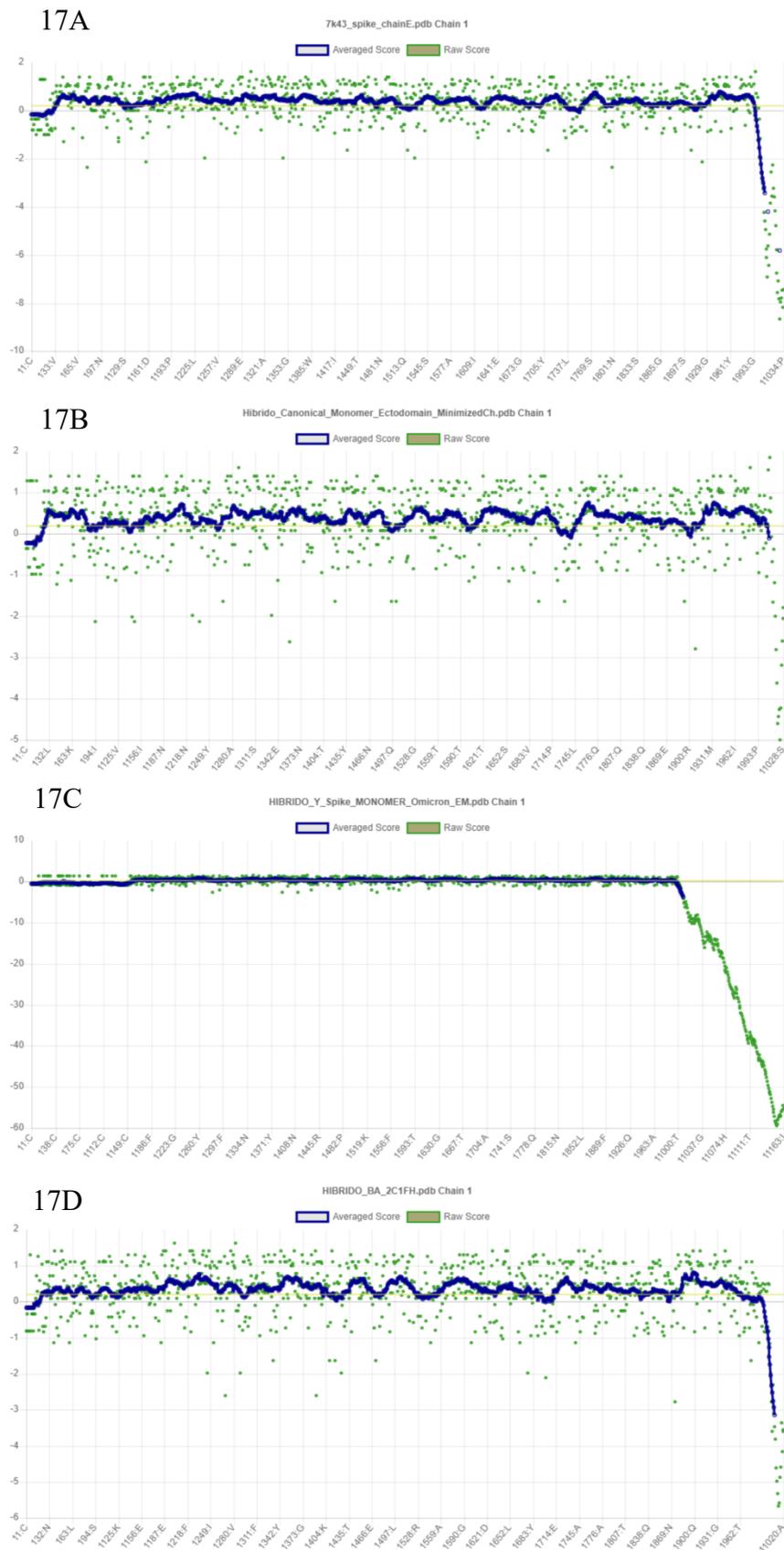


Figura 17. Análise de compatibilidade 3D-1D dos modelos atômicos das proteínas spike do SARS-CoV-2 da referência 7K43-E (17A), das variantes de Wuhan (17B), da Omicron (17C) e da BA.2 (17D). O segmento azul representa o resultado médio e os pontos verdes representam o resultado bruto da pontuação 3D-1D.

Como aponta a Figura 18, as pontuações de compatibilidade 3D-1D do modelo atômico da Isoforma 1 de ECA2 são plotadas em gráficos em função dos resíduos.

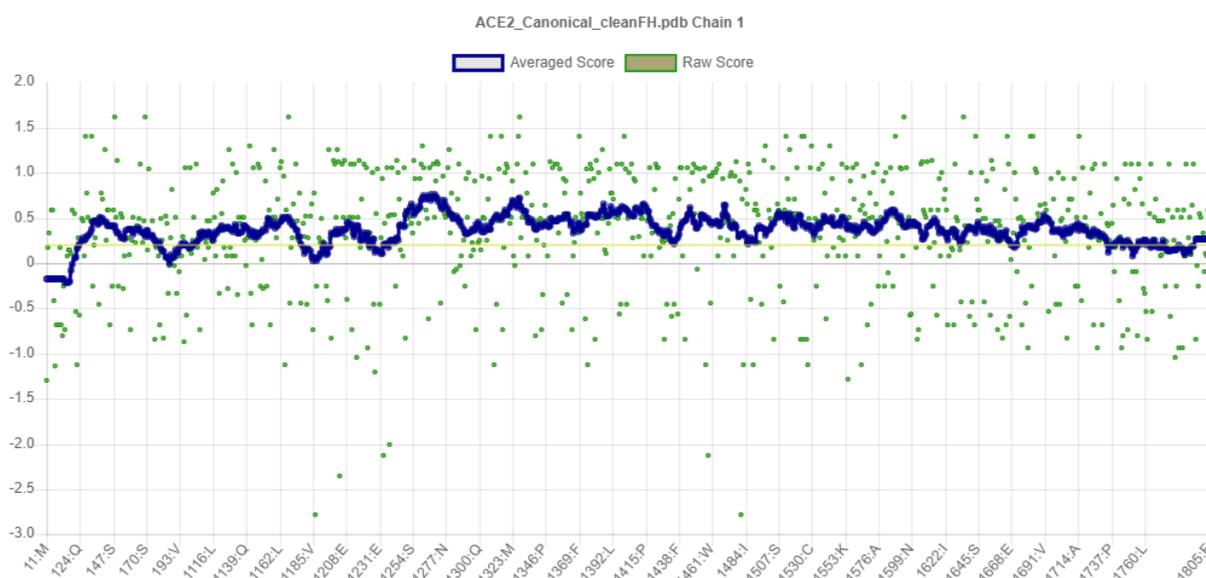


Figura 18. Análise de compatibilidade 3D-1D do modelo atômico referente isoforma 1 de ECA2. O segmento azul representa o resultado médio e os pontos verdes representam o resultado bruto da pontuação 3D-1D.

Como mostra a Tabela 9, os resultados gerais quantitativos das frequências relativas de segmentos com pontuação 3D-1D $\geq 0,2$ dos modelos atômicos das proteínas spike do SARS-CoV-2 do grupo 1 da referência 7CAK-A (A), das variantes com mutação pontual E484K (E), combinadas K417Y, E484K e N501Y (F) e do grupo 2 da referência 7K43-E (A), da variante Omicron (F) e da subvariante BA.2 (G) são mostrados. Somente o modelo da proteína spike da variante Omicron não apresentou uma qualidade 3D-1D satisfatória.

Tabela 9. Frequência relativa de segmentos com pontuação 3D-1D $\geq 0,2$ das estruturas 3D modeladas das proteínas spike do SARS-CoV-2.

Grupo	Frequência (%) de segmentos de pontuação (3D-1D) $\geq 0,2$ do modelo				
	Referência (A)	B	E	F	G
Proteína spike do SARS-CoV-2 grupo 1	87,16	---	84,73	85,28	---
Proteína spike do SARS-CoV-2 grupo 2	84,53	85,60	---	65,18	80,00

Tabela 9. Frequência relativa de segmentos com pontuação 3D-1D $\geq 0,2$ das estruturas 3D modeladas das proteínas spike do SARS-CoV-2. Proteínas spike do SARS-CoV-2 do grupo 1 da referência 7CAK-A (A), das variantes com mutação pontual E484K (E), combinadas K417Y, E484K e N501Y (F) e Proteínas spike do SARS-CoV-2 do grupo 2 da referência 7K43-E (A), da variante Omicron (F) e da subvariante BA.2 (G). Frequências de pelo menos 80% indicam que o modelo apresentou uma compatibilidade satisfatória de sua conformação tridimensional com a sequência de aminoácidos.

Como mostra a Tabela 10, os resultados gerais quantitativos das frequências relativas de segmentos com pontuação 3D-1D $\geq 0,2$ do modelo atômico da Isoforma 1 de ECA2 são mostrados. Somente o modelo da proteína spike da variante Omicron não apresentou uma qualidade 3D-1D satisfatória.

Tabela 10. Frequência relativa de segmentos com pontuação 3D-1D $\geq 0,2$ da estrutura 3D modelada das isoforma 1 de ECA2.

Proteína	Frequência (%) de segmentos de pontuação (3D-1D) $\geq 0,2$ do modelo (%)
Isoforma 1 de ECA2	87,08

Tabela 10. Frequência relativa de segmentos com pontuação 3D-1D $\geq 0,2$ da estrutura 3D modelada das isoforma 1 de ECA2. Frequências de pelo menos 80% indicam que o modelo apresentou uma compatibilidade satisfatória de sua conformação tridimensional com a sequência de aminoácidos.

4.5.3. PROCHECK

A ferramenta PROCHECK gera um conjunto de gráficos que descrevem e avaliam a qualidade de características estereoquímicas de cada resíduo de modelos proteicos, tais como: ângulos de torsão diedros phi (ϕ), psi (ψ), ômega (ω), Chi1 e Chi2, distâncias das ligações e desvios das raízes quadráticas médias (RMS) das distâncias entre os átomos de um mesmo plano, permitindo examinar posições de resíduos desfavoráveis estereoquimicamente. A avaliação da qualidade se dá através de uma comparação dos dados da estereoquímica entre o modelo atômico submetido e estruturas de alta qualidade de mesma resolução e é apontada pelos gráficos fornecidos que estabelecem limites para os parâmetros de um modelo atômico de qualidade (LASKOWSKI et al., 1993, 1996). Para a análise dos ângulos diedros, essa ferramenta gera o gráfico de Ramachandran. O Ramachandran é um gráfico ilustrativo onde é possível analisar uma combinação de dois ângulos de torsão, sendo comumente usado para os ângulos de torsão phi (ϕ) e psi (ψ) (Figura 19) dos resíduos de um modelo proteico. A partir da análise, é viável inferir esta combinação de ângulos de torsão obedece aos limites impostos pela estereoquímica. O gráfico de Ramachandran apresenta regiões permitidas e não permitidas nas quais são preenchidas pela combinação de par ordenado dos ângulos de torsão

envolvidos, em graus, dos resíduos. Os ângulos de torção ϕ e ψ se encontram respectivamente nas ligações N-C α e C α -C (carboxila) de um aminoácido (RAMACHANDRAN; RAMAKRISHNAN; SASISEKHARAN, 1963).

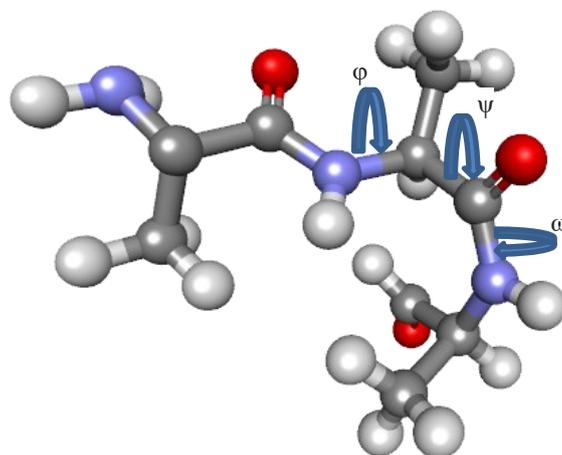
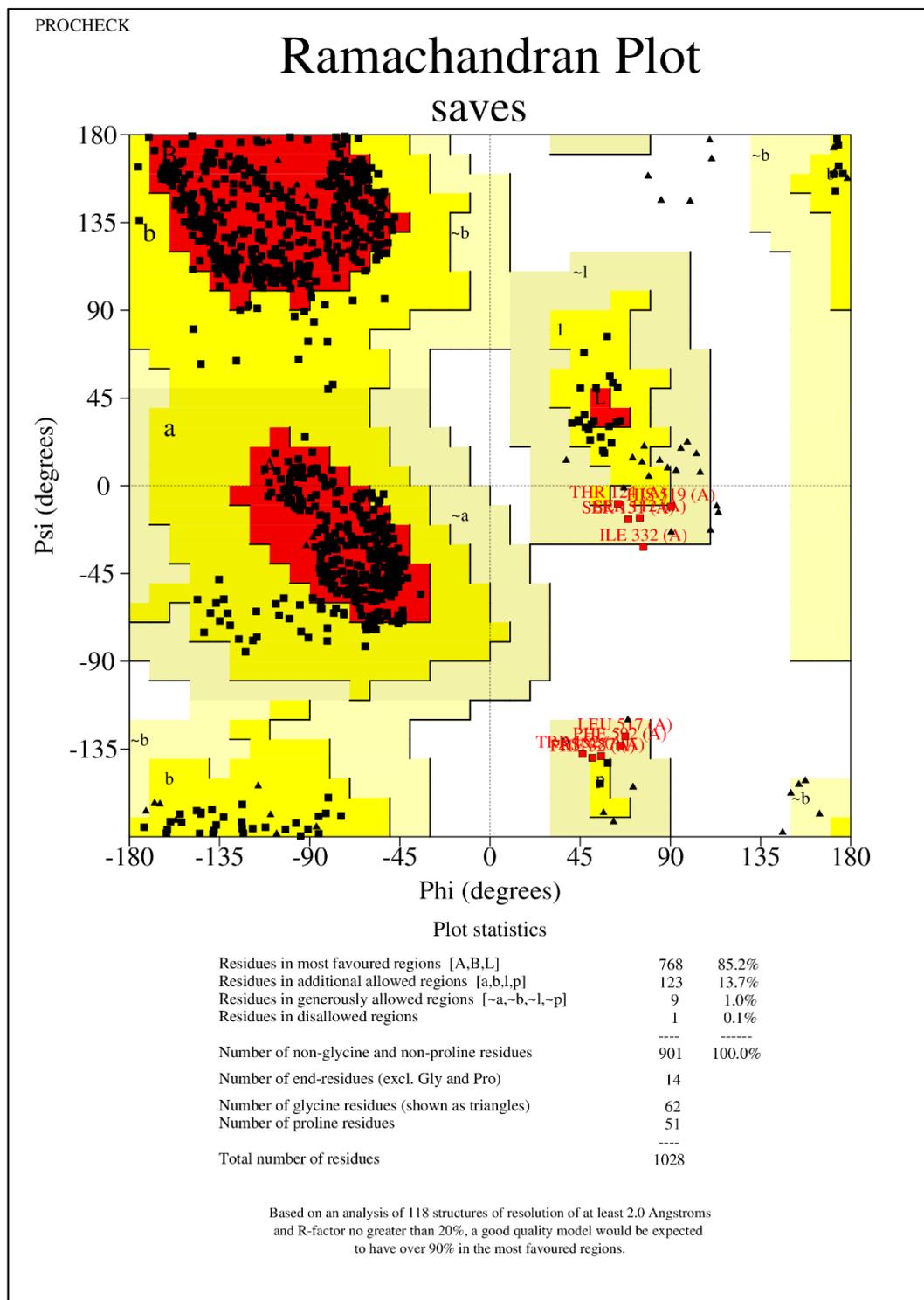


Figura 19. Representação esquemática em forma de esferas e bastões de três alaninas formando uma cadeia por duas ligações peptídicas com destaque nos ângulos de torção phi (ϕ), psi (ψ) e ômega (ω). Esferas em cinza, azul, cinza claro e vermelho representam respectivamente os átomos de carbono, nitrogênio, hidrogênio e oxigênio, e os bastões representam as ligações intramoleculares simples e duplas. Fonte do autor. A figura foi construída através de ferramentas de modelagem molecular do software Discovery Studio Visualizer (“Discovery Studio Visualizer 17.2.0.16349”, [s.d.]).

De acordo com a Figura 20, são disponibilizados e avaliados as combinações de ângulos de torção phi (ϕ) e psi (ψ) dos resíduos do modelo atômico da proteína spike do SARS-CoV-2 de referência 7CAK-A. Comparando os dados dos ângulos de torção phi (ϕ) e psi (ψ) com 118 estruturas com pelo menos 2 angstroms de resolução e fator R de no máximo 20%, o modelo não apresentou uma alta qualidade, uma vez que a porcentagem de resíduos nas regiões mais favoráveis não superou o valor de 90%.



saves_01.ps

Figura 20. Gráfico Ramachandran do modelo atômico 7CAK-A. As áreas destacadas de vermelho ao branco representam as combinações dos ângulos diedros ϕ (phi) e ψ (psi) de mais favoráveis a não permitidas dos

resíduos representados pelos pontos pretos (dentro de regiões favoráveis) e vermelhos (fora de regiões favoráveis).

De acordo com a Figura 21, são disponibilizados e avaliados as combinações de ângulos de torção phi (φ) e psi (ψ) dos resíduos do modelo atômico da proteína spike do SARS-CoV-2 com mutação pontual E484K. Comparando os dados dos ângulos de torção phi (φ) e psi (ψ) com 118 estruturas com pelo menos 2 angstroms de resolução e fator R de no máximo 20%, o modelo não apresentou uma alta qualidade, uma vez que a porcentagem de resíduos nas regiões mais favoráveis não superou o valor de 90%.

De acordo com a Figura 22, são disponibilizados e avaliados as combinações de ângulos de torção phi (ϕ) e psi (ψ) dos resíduos do modelo atômico da proteína spike do SARS-CoV-2 com mutações combinadas K417Y, E484K e N501Y. Comparando os dados dos ângulos de torção phi (ϕ) e psi (ψ) com 118 estruturas com pelo menos 2 angstroms de resolução e fator R de no máximo 20%, o modelo não apresentou uma alta qualidade, uma vez que a porcentagem de resíduos nas regiões mais favoráveis não superou o valor de 90%.

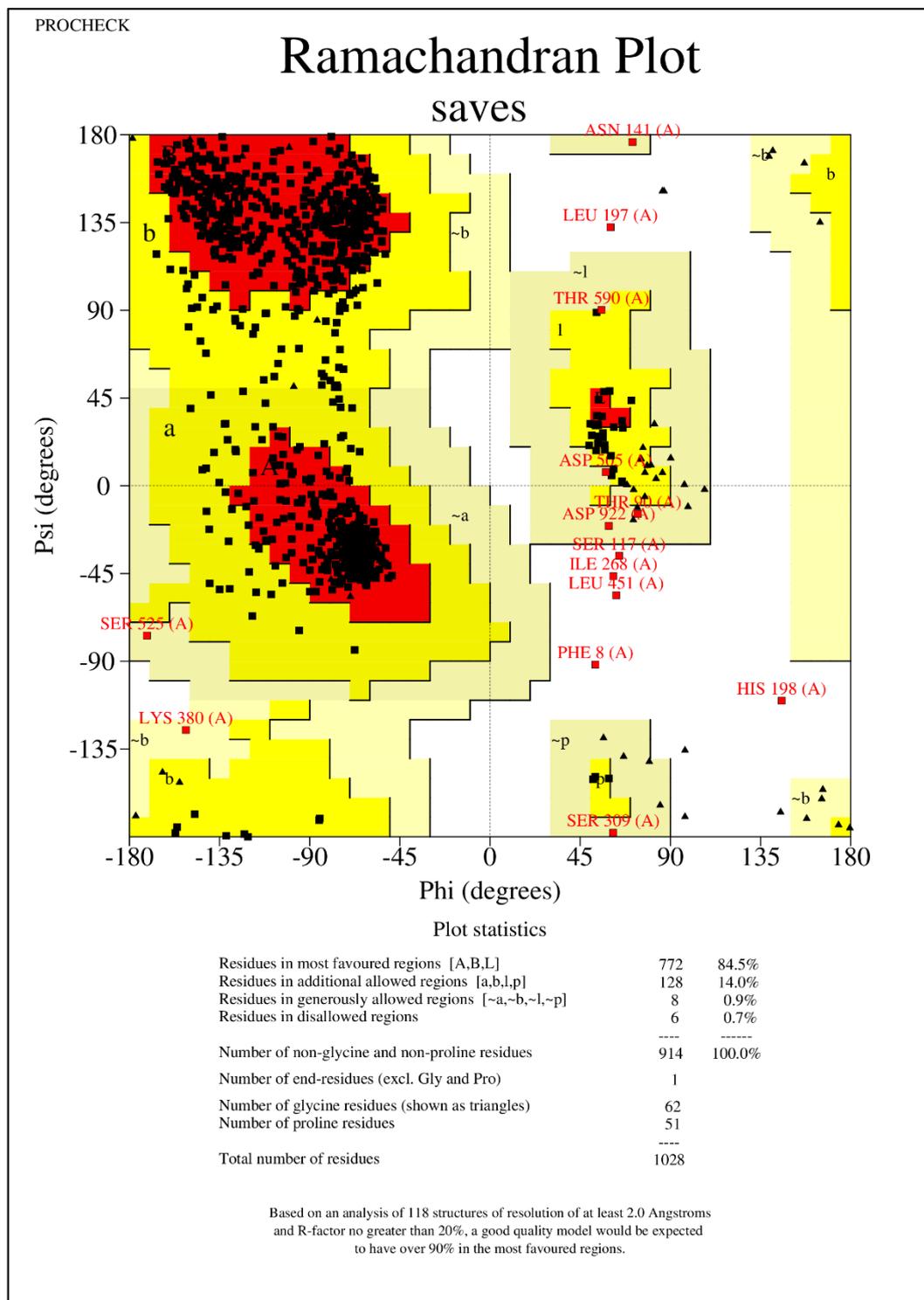


Figura 22. Gráfico Ramachandran do modelo atômico referente à proteína spike do SARS-CoV-2 da variante de com as mutações K417Y+E484K+N501Y do grupo 1. As áreas destacadas de vermelho ao branco representam as combinações dos ângulos diedros de mais favoráveis a não permitidas dos resíduos representados pelos pontos pretos (dentro de regiões favoráveis) e vermelhos (fora de regiões favoráveis).

De acordo com a Figura 23, são disponibilizados e avaliados as combinações de ângulos de torção phi (ϕ) e psi (ψ) dos resíduos do modelo atômico da proteína spike do SARS-CoV-2 de referência 7K43-E. Comparando os dados dos ângulos de torção phi (ϕ) e psi (ψ) com 118 estruturas com pelo menos 2 angstroms de resolução e fator R de no máximo 20%, o modelo apresentou uma alta qualidade, uma vez que a porcentagem de resíduos nas regiões mais favoráveis superou o valor de 90%.

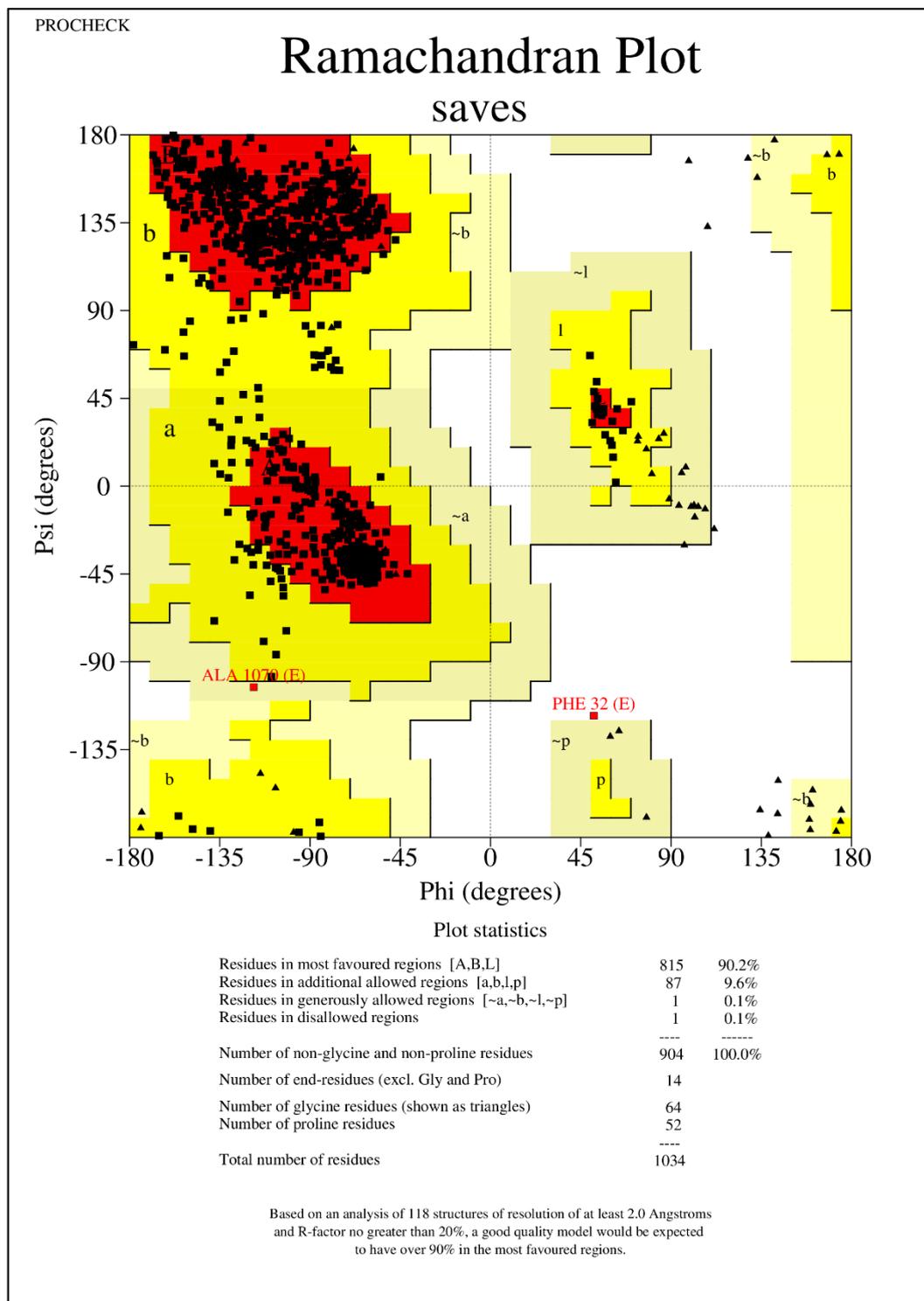


Figura 23. Gráfico Ramachandran do modelo atômico referente à proteína spike do SARS-CoV-2 da estrutura de referência do grupo 2 7K43-E. As áreas destacadas de vermelho ao branco representam as combinações dos ângulos diedros de mais favoráveis a não permitidas dos resíduos representados pelos pontos pretos (dentro de regiões favoráveis) e vermelhos (fora de regiões favoráveis).

De acordo com a Figura 24, são disponibilizadas e avaliadas as combinações de ângulos de torção phi (φ) e psi (ψ) dos resíduos do modelo atômico da proteína spike do SARS-CoV-2 original de Wuhan do grupo 2. Comparando os dados dos ângulos de torção phi (φ) e psi (ψ) com 118 estruturas com pelo menos 2 angstroms de resolução e fator R de no máximo 20%, o modelo não apresentou uma alta qualidade, uma vez que a porcentagem de resíduos nas regiões mais favoráveis não superou o valor de 90%.

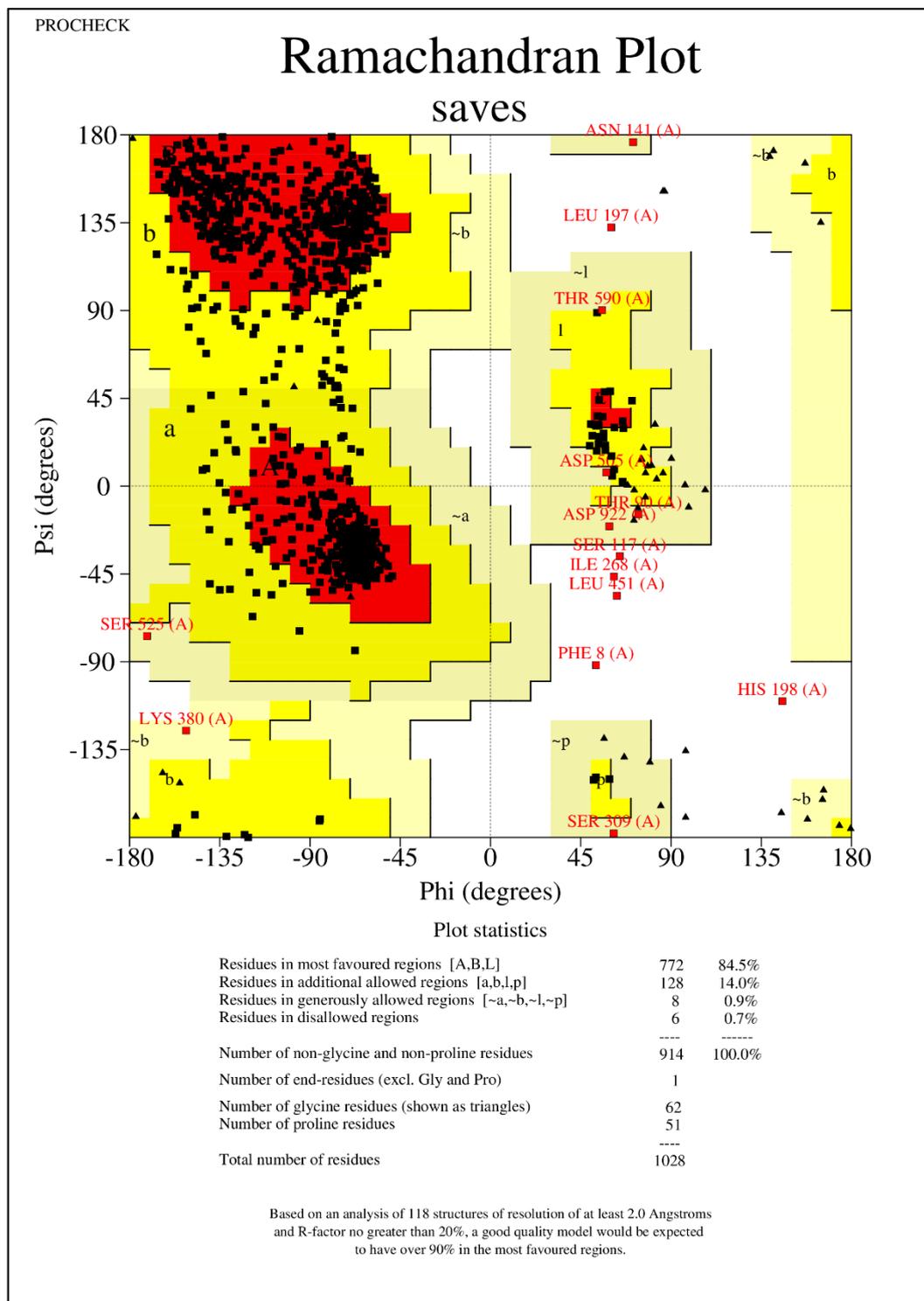


Figura 24. Gráfico Ramachandran do modelo atômico referente à proteína spike do SARS-CoV-2 da VoC de Wuhan do grupo 2. As áreas destacadas de vermelho ao branco representam as combinações dos ângulos diedros de mais favoráveis a não permitidas dos resíduos representados pelos pontos pretos (dentro de regiões favoráveis) e vermelhos (fora de regiões favoráveis).

De acordo com a Figura 25, são disponibilizadas e avaliadas as combinações de ângulos de torção phi (φ) e psi (ψ) dos resíduos do modelo atômico da proteína spike do SARS-CoV-2 da VoC Omicron. Comparando os dados dos ângulos de torção phi (φ) e psi (ψ) com 118 estruturas com pelo menos 2 angstroms de resolução e fator R de no máximo 20%, o modelo não apresentou uma alta qualidade, uma vez que a porcentagem de resíduos nas regiões mais favoráveis não superou o valor de 90%.

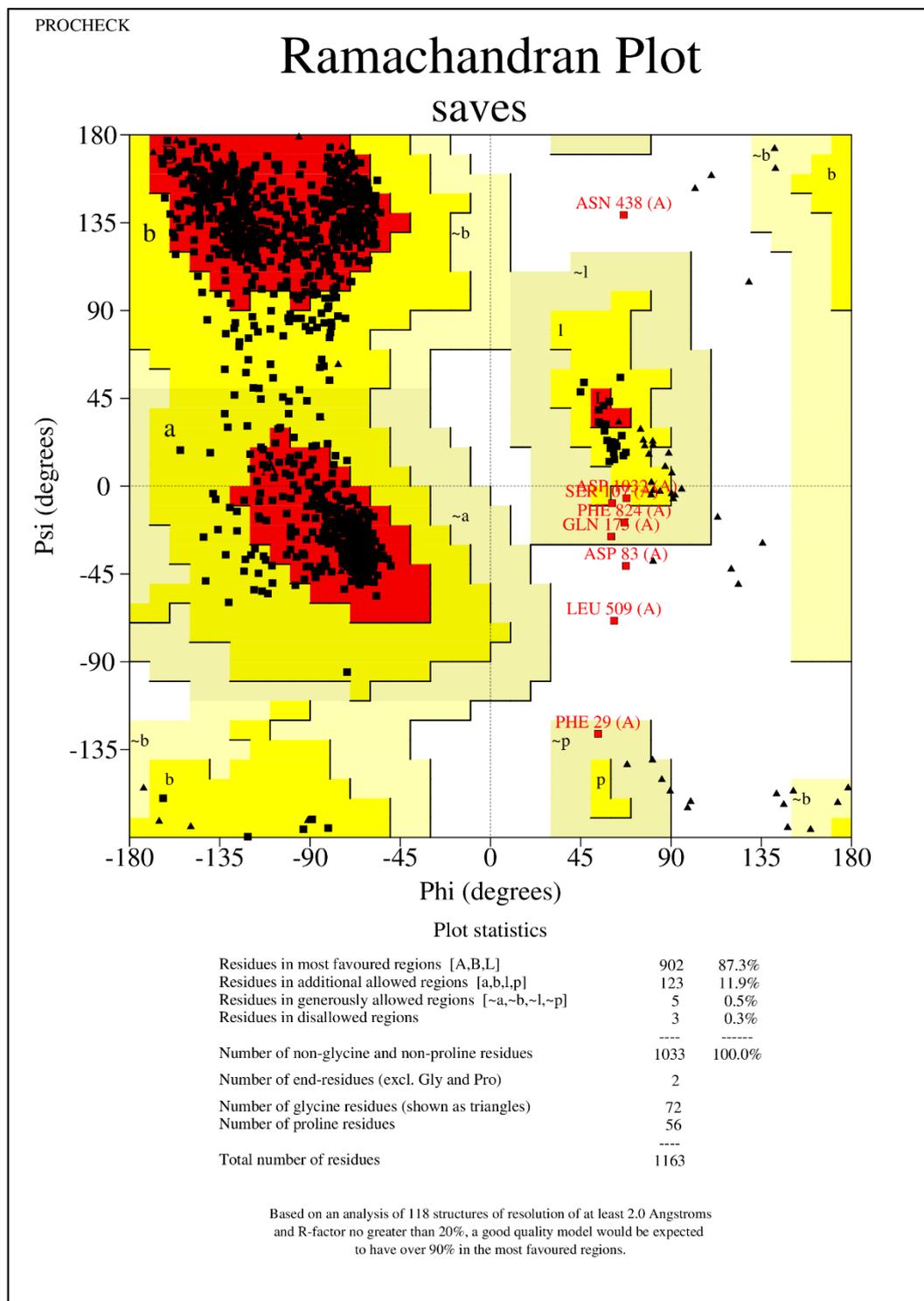
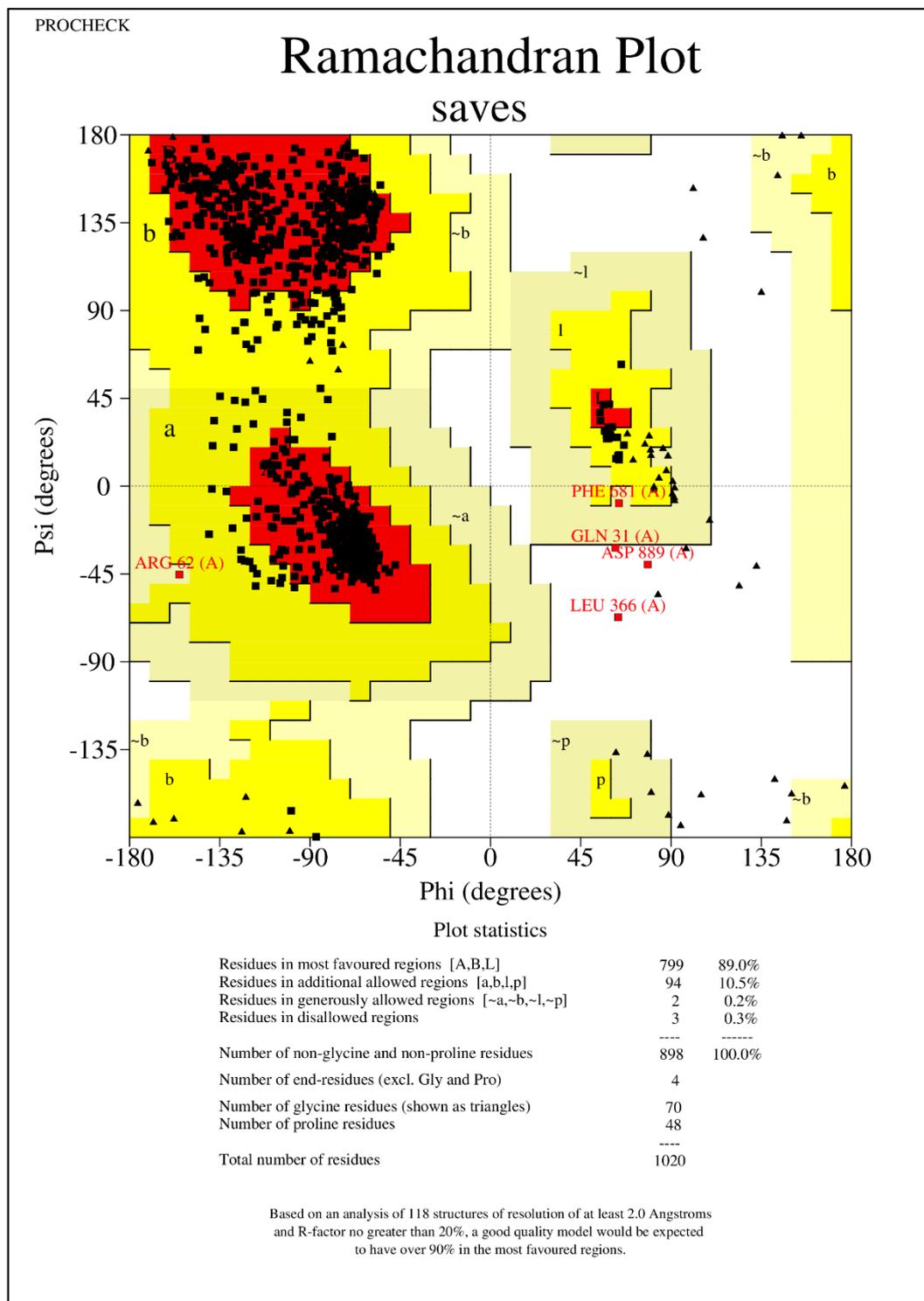


Figura 25. Gráfico Ramachandran do modelo atômico referente à proteína spike do SARS-CoV-2 da VoC Omicron. As áreas destacadas de vermelho ao branco representam as combinações dos ângulos diedros de mais favoráveis a não permitidas dos resíduos representados pelos pontos pretos (dentro de regiões favoráveis) e vermelhos (fora de regiões favoráveis).

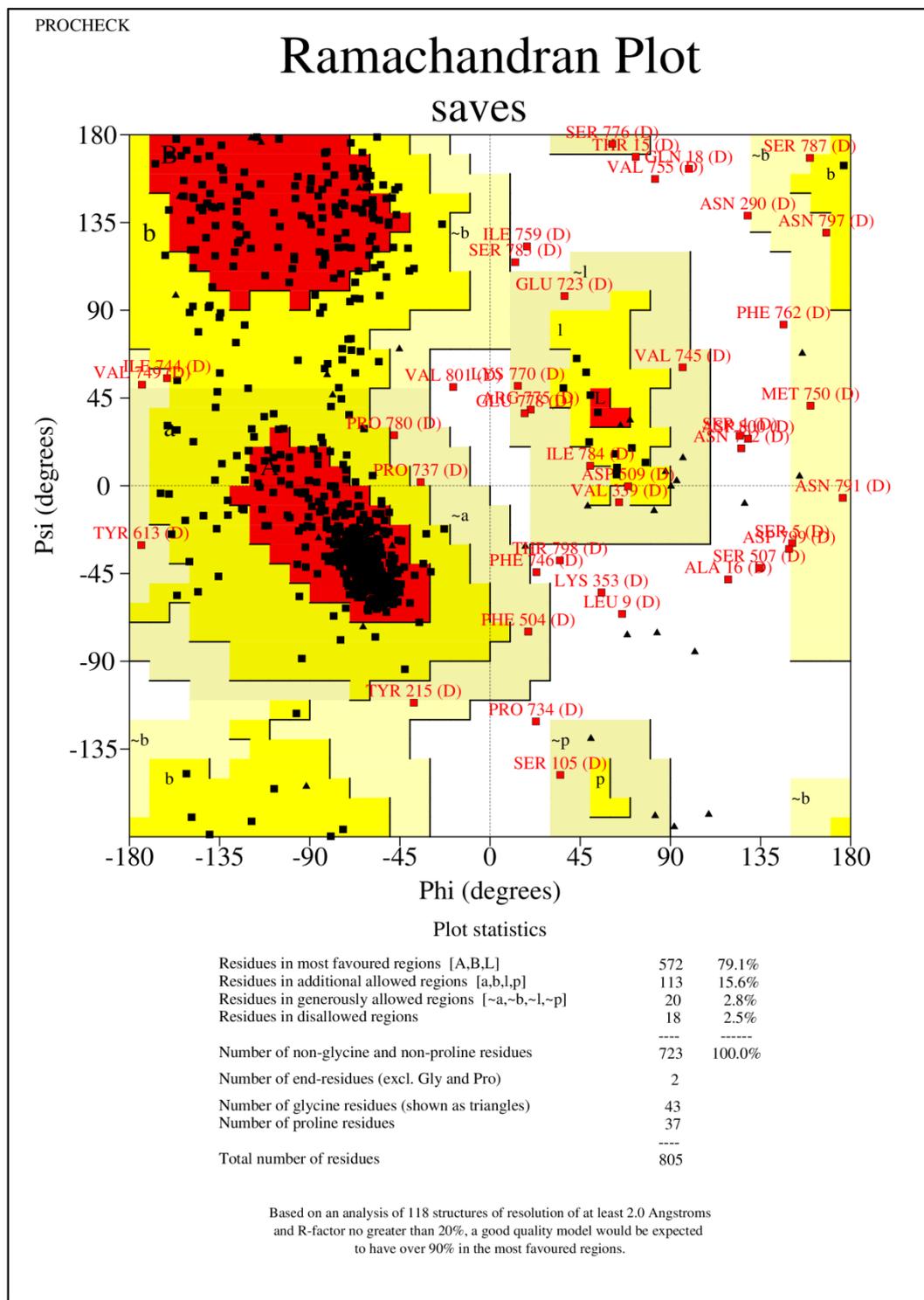
De acordo com a Figura 26, são disponibilizadas e avaliadas as combinações de ângulos de torção phi (ϕ) e psi (ψ) dos resíduos do modelo atômico da proteína spike do SARS-CoV-2 da subvariante BA.2. Comparando os dados dos ângulos de torção phi (ϕ) e psi (ψ) com 118 estruturas com pelo menos 2 angstroms de resolução e fator R de no máximo 20%, o modelo não apresentou uma alta qualidade, uma vez que a porcentagem de resíduos nas regiões mais favoráveis não superou o valor de 90%.



saves_01.ps

Figura 26. Gráfico Ramachandran do modelo atômico referente à proteína spike do SARS-CoV-2 da VoC BA.2. As áreas destacadas de vermelho ao branco representam as combinações dos ângulos diedros de mais favoráveis a não permitidas dos resíduos representados pelos pontos pretos (dentro de regiões favoráveis) e vermelhos (fora de regiões favoráveis).

De acordo com a Figura 27, são disponibilizadas e avaliadas as combinações de ângulos de torção phi (ϕ) e psi (ψ) dos resíduos do modelo atômico isoforma da ECA2. Comparando os dados dos ângulos de torção phi (ϕ) e psi (ψ) com 118 estruturas com pelo menos 2 angstroms de resolução e fator R de no máximo 20%, o modelo não apresentou uma alta qualidade, uma vez que a porcentagem de resíduos nas regiões mais favoráveis não superou o valor de 90%.



saves_01.ps

Figura 27. Gráfico Ramachandran do modelo atômico referente à isoforma 1 de ECA2. As áreas destacadas de vermelho ao branco representam as combinações dos ângulos diedros de mais favoráveis a não permitidas dos resíduos representados pelos pontos pretos (dentro de regiões favoráveis) e vermelhos (fora de regiões favoráveis).

Conforme a Tabela 11, são detalhados os dados da porcentagem de resíduos nas regiões mais favoráveis, regiões adicionais permitidas, regiões generosamente permitidas e regiões não permitidas das estruturas moleculares das proteínas spike do SARS-CoV-2 das variantes com mutação pontual E484K (S1E), com mutações combinadas N417Y, E484K e N501Y (S1F), da estrutura original de Wuhan (S2B) e da variante Omicron (S2F), da subvariante BA.2 (S2G) e isoforma 1 de ECA2 (E1A).

Tabela 11. Caracterização dos resíduos com base no gráfico de Ramachandran das proteínas spike do SARS-CoV-2 do grupo 1 e 2 e das isoformas de ECA2.

Grupos	Abreviações	Resíduos (%) nas:			
		Regiões mais favoráveis	Regiões adicionais permitidas	Regiões generosamente permitidas	Regiões não permitidas
Proteína spike do SARS-CoV-2 grupo 1	S1A	85,2	13,7	1,0	0,1
	S1E	83,0	14,7	1,5	0,8
	S1F	84,5	14,0	0,9	0,7
Proteína spike do SARS-CoV-2 grupo 2	S2A	90,2	9,6	0,1	0,1
	S2B	84,5	14,0	0,9	0,7
	S2F	87,3	11,9	0,5	0,3
	S2G	89,0	10,5	0,2	0,3
Isoformas de ECA2	E1A	79,1	15,6	2,8	2,5

Tabela 11. Caracterização dos resíduos com base no gráfico de Ramachandran das proteínas spike do SARS-CoV-2 do grupo 1 e 2 e das isoformas de ECA2. Porcentagem de resíduos que devido a combinação de ângulos phi (ϕ) e psi (ψ) pertencem as regiões mais favoráveis, regiões adicionais permitidas, regiões generosamente permitidas e regiões não permitidas das proteínas de referência 7CAK-A (S1A), variante de Wuhan (S2B) e variantes de mutações pontuais E484K (S1E), K417Y, E484K e N501Y (S1F) do grupo 1, de referência 7K43-E (S2A), Omicron (S2F) e BA.2 (S2G) do grupo 2 e isoforma 1 (E1A) de ECA2.

Os aminoácidos responsáveis por reduzirem a porcentagem de resíduos nas regiões mais favoráveis nas figuras dos gráficos Ramachandran dos ângulos de torção phi (ϕ) e psi (ψ) anteriormente descritos, foram detalhados nas Figura 28, Figura 29 e Figura 30.

Como indicado na Figura 28, os resíduos destacados e identificados dos modelos atômicos de proteína spike do SARS-CoV-2 de referência 7CAK-A (28A) do grupo 1 e de variantes de mutações pontuais E484K (28B) e mutações combinadas K417Y, E484K e N501Y (28C) apresentaram uma pontuação conformacional $<-3,00$ que aponta para a baixa

qualidade residual em relação aos ângulos phi (φ) e psi (ψ). Para calcular a pontuação, foi estabelecida uma comparação com 163 estruturas com 2 angstroms de resolução ou mais.

Como demonstrado na Tabela 12, os resíduos identificados dos modelos atômicos da proteína spike do SARS-CoV-2 de referência do grupo 1 7CAK-A (S1A) e de variantes de mutações pontuais E484K (S1E) e mutações combinadas K417Y, E484K e N501Y (S1F) apresentaram uma baixa pontuação conformacional em relação aos ângulos phi (ϕ) e psi (ψ), o que indica uma baixa qualidade residual. Desconsiderando o modelo referência, apenas os resíduos Asn358 do modelo S1F e Gln498 do modelo S1E da proteína spike do SARS-CoV-2 do grupo 1 estão inseridos no RBD (319-541 aminoácidos em relação a sequência de Wuhan) e estão representados na tabela. Isso sugere que possíveis erros conformacionais nesses resíduos não devem causar grandes interferências negativas no docking molecular.

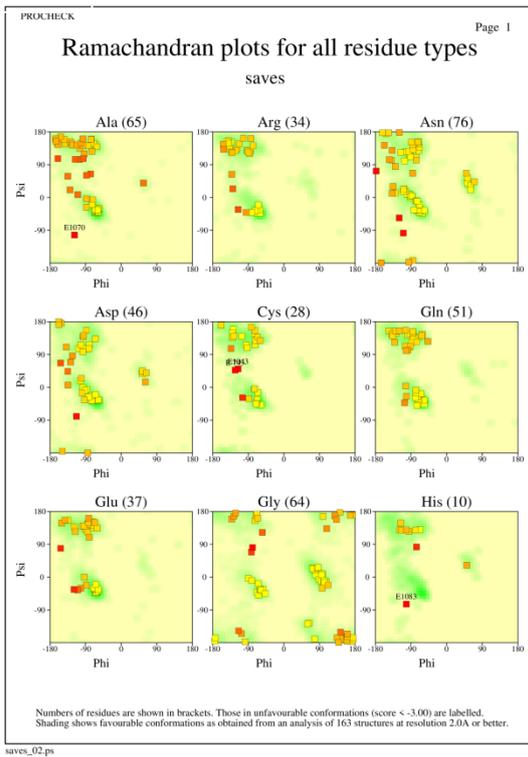
Tabela 12. Posições dos resíduos cuja conformação é desfavorável pelo gráfico de Ramachandran nos modelos da proteína spike do SARS-CoV-2 do primeiro grupo.

Resíduo	Posição do resíduo na proteína spike do SARS-CoV-2 do grupo 1		
	S1A	S1E	S1F
Ala	1080	89, 685, 951	685, 951
Arg			
Asn	164, 1135	130, 141, 270	130, 358, 571
Asp	663	698, 922	922
Cys	590, 1126	732, 924	732, 924
Gln	564, 580, 901	498, 592	952
Glu	324		
Gly	857	646	172, 646
His	146, 519, 1083	198, 964	198, 569, 964

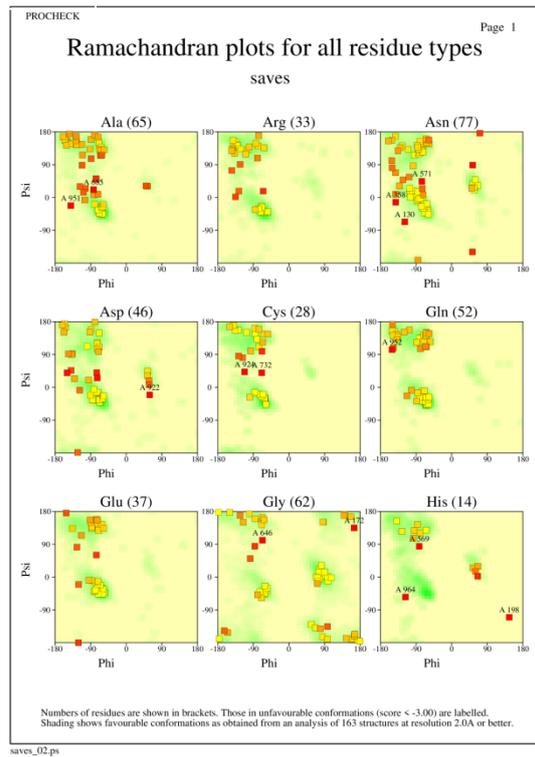
Tabela 12. Posições dos resíduos cuja conformação é desfavorável pelo gráfico de Ramachandran nos modelos da proteína spike do SARS-CoV-2 do primeiro grupo. Tabela de dados do gráfico de Ramachandran por resíduos dos modelos atômicos das variantes da proteína spike de referência 7CAK-A (S1A) e variantes de mutações pontuais E484K (S1E) e K417Y, E484K e N501Y (S1F). Os resíduos descritos devido a sua combinação de ângulos phi e psi, pertencem a área não permitida do gráfico de Ramachandran.

Como indicado na Figura 29, os resíduos destacados e identificados dos modelos atômicos referentes à cadeia E da estrutura de referência (código PDB: 7K43) (29A), da estrutura molecular da proteína spike original de Wuhan do grupo 2 (29B), da variante Omicron (29C) e da subvariante BA.2 (29D) apresentaram uma pontuação conformacional $< -3,00$ que aponta para a baixa qualidade residual em relação aos ângulos phi (ϕ) e psi (ψ). Para calcular a pontuação, foi estabelecida uma comparação com 163 estruturas com resolução de 2 angstroms ou mais.

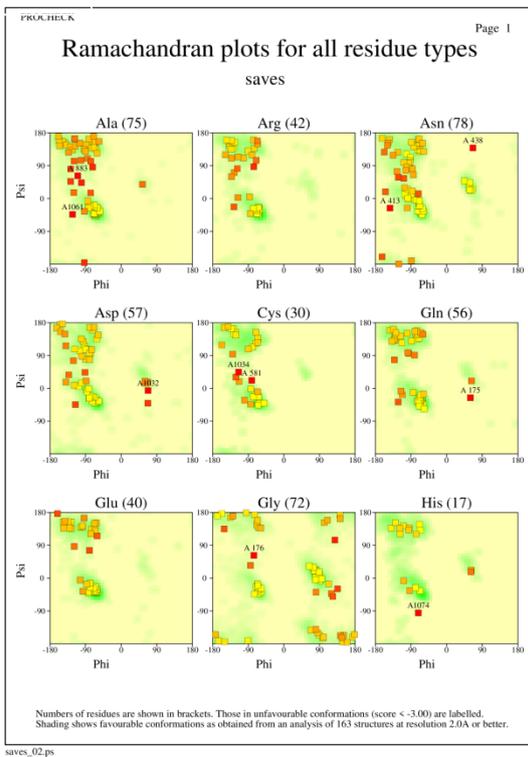
29A



29B



29C



29D

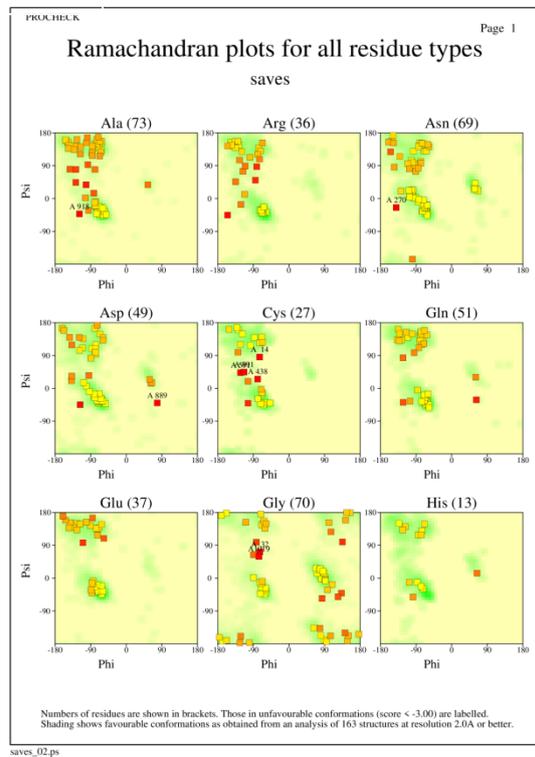


Figura 29. Gráficos Ramachandran individuais dos ângulos de torção phi (ϕ) e psi (ψ) por tipos de resíduos: Ala, Arg, Asn, Asp, Cys, Gln, Glu, Gly e His dos modelos atômicos referentes à cadeia E da estrutura de referência (código PDB: 7K43) (29A), da estrutura molecular da proteína spike original de Wuhan do grupo 2 (29B), da variante Omicron (29C) e da subvariante BA.2 (29D). O número de resíduos de mesmo tipo é

mostrado em parênteses. Resíduos destacados e identificados indicam uma pontuação < -3,00 e, portanto, uma conformação desfavorável energeticamente.

Como apresentado na Tabela 13, os resíduos identificados dos modelos atômicos da proteína spike do SARS-CoV-2 de referência 7K43 (S2A), da estrutura molecular original da proteína de Wuhan do grupo 2 (S2B), da variante omicron (S2F) e da subvariante BA.2 (S2G) apresentaram uma baixa pontuação conformacional em relação aos ângulos phi (ϕ) e psi (ψ), o que indica uma baixa qualidade residual. Desconsiderando o modelo referência, apenas os resíduos Asn358 do modelo S2B, Asn413 e Asn 438 do modelo S2F e Asn270 do modelo S2G da proteína spike do SARS-CoV-2 do grupo 2 estão inseridos no RBD (319-541 aminoácidos em relação a sequência de Wuhan) e estão representados na tabela. Isso sugere que possíveis erros conformacionais nesses resíduos não devem causar grandes interferências negativas no docking molecular.

Tabela 13. Posições dos resíduos cuja conformação é desfavorável pelo gráfico de Ramachandran dos ângulos diedros phi (ϕ) e psi (ψ) nos modelos da proteína spike do SARS-CoV-2 do segundo grupo.

Resíduo	Posição do resíduo na proteína spike do SARS-CoV-2 do grupo 2			
	S2A	S2B	S2F	S2G
Ala	1070	685, 951	833, 1061	918
Arg				
Asn		130, 358, 571	413, 438	270
Asp		922	1032	889
Cys	743, 1043	732, 924	581, 1034	14, 438, 591, 891
Gln		952	175	
Glu				
Gly		172, 646	176	32, 1019
His	1083	198, 569, 964	1074	

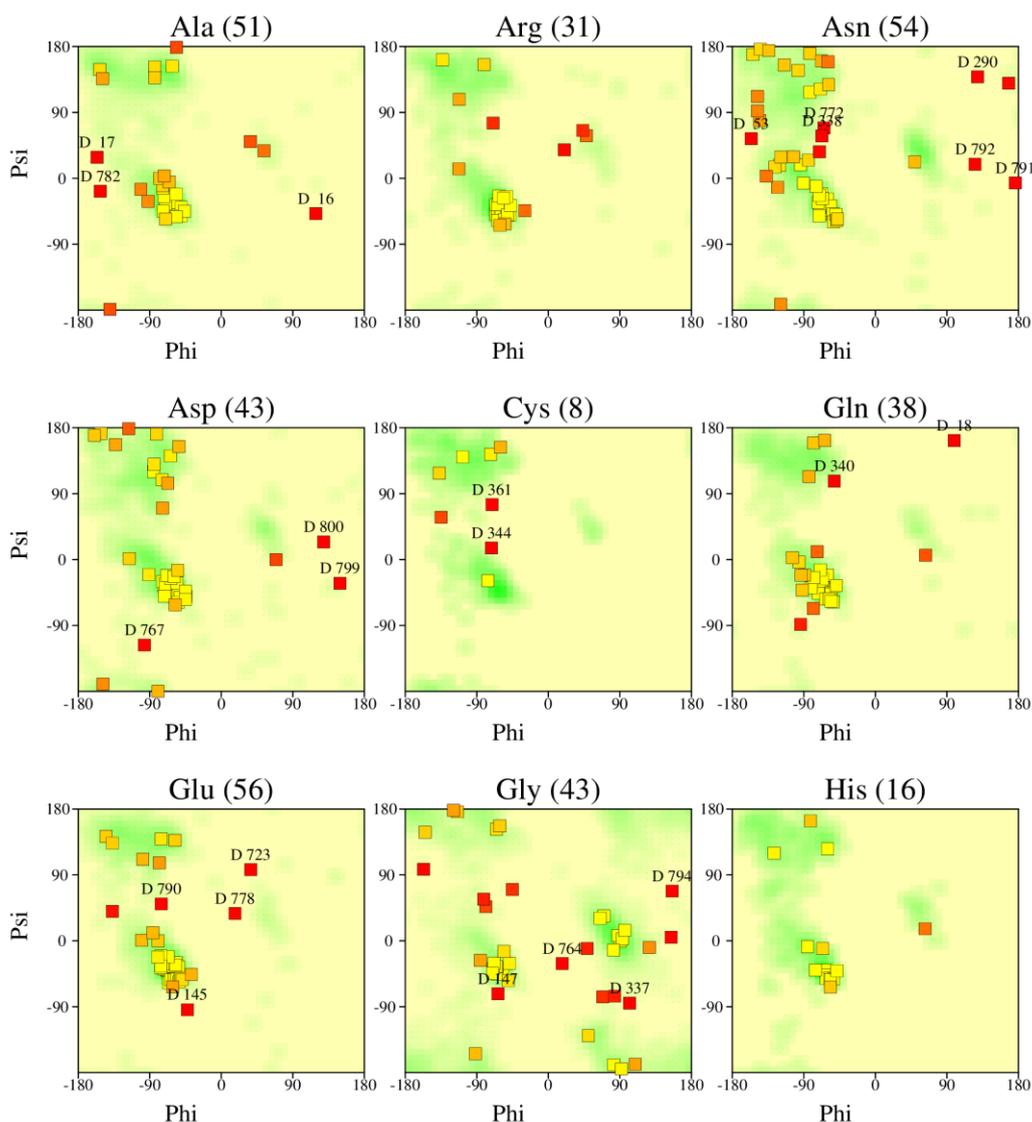
Tabela 13. Posições dos resíduos cuja conformação é desfavorável pelo gráfico de Ramachandran dos ângulos diedros phi (ϕ) e psi (ψ) nos modelos da proteína spike do SARS-CoV-2 do segundo grupo. Tabela de dados do gráfico de Ramachandran por resíduos dos modelos atômicos da proteína spike do SARS-CoV-2 de referência 7K43 (S2A), da estrutura molecular da proteína spike do SARS-CoV-2 original de Wuhan do grupo 2 (S2B), da variante omicron (S2F) e da subvariante BA.2 (S2G). Os resíduos descritos devido a sua combinação de ângulos phi e psi, pertencem a área não permitida do gráfico de Ramachandran.

Como indicado na Figura 30, os resíduos destacados e identificados do modelo atômico da Isoforma 1 de ECA2 apresentaram uma pontuação conformacional <-3,00 que aponta para a baixa qualidade residual em relação aos ângulos phi (ϕ) e psi (ψ). Para calcular

a pontuação, foi estabelecida uma comparação com 163 estruturas com resolução de 2 angstroms ou mais.

Ramachandran plots for all residue types

saves



Numbers of residues are shown in brackets. Those in unfavourable conformations (score < -3.00) are labelled. Shading shows favourable conformations as obtained from an analysis of 163 structures at resolution 2.0Å or better.

saves_02.ps

Figura 30. Gráficos Ramachandran individuais dos ângulos de torção phi (ϕ) e psi (ψ) por tipos de resíduos: Ala, Arg, Asn, Asp, Cys, Gln, Glu, Gly e His do modelo atômico da isoforma 1 de ECA2. O número de resíduos de mesmo tipo é mostrado em parênteses. Resíduos destacados e identificados indicam uma pontuação < -3,00 e, portanto, uma conformação desfavorável.

Como apontado na Tabela 13, os resíduos identificados dos modelos atômicos da proteína spike do SARS-CoV-2 de referência 7K43-E (S2A), da estrutura molecular original da proteína de Wuhan do grupo 2 (S2B), da variante omicron (S2F) e da subvariante BA.2 (S2G) apresentaram uma baixa pontuação conformacional em relação aos ângulos de torção phi (ϕ) e psi (ψ), o que indica uma baixa qualidade residual. Dos indicados, os resíduos Ala16, Ala17, Gln18, Asn53, Asn338, Gly337 e Cys344 da Isoforma 1 de ECA2 estão inseridos nas regiões de interação com proteína spike do SARSCoV-2. Isso sugere que possíveis erros conformacionais nesses resíduos podem causar interferências negativas nas interações, com a presença no docking molecular de resíduos que realmente não interagem com a proteína spike do SARS-CoV-2 ou com a ausência de resíduos que deveriam participar das interações, prejudicando a precisão do docking molecular.

Tabela 14. Posições dos resíduos cuja conformação é desfavorável pelo gráfico de Ramachandran dos ângulos diedros phi (ϕ) e psi (ψ) no modelo da isoforma 1 de ECA2.

Resíduo	Posição do resíduo na isoforma 1 de ECA2
Ala	16, 17, 782
Arg	
Asn	53, 290, 338, 772, 791, 792
Asp	767, 799, 800
Cys	344, 361
Gln	18, 340
Glu	152, 723, 778, 790
Gly	147, 337, 764, 796
His	

Tabela 14. Posições dos resíduos cuja conformação é desfavorável pelo gráfico de Ramachandran dos ângulos diedros phi (ϕ) e psi (ψ) no modelo da isoforma 1 de ECA2. Devido à combinação de ângulos phi e psi, os resíduos descritos se encontram em regiões não permitidas do gráfico de Ramachandran.

Através do gráfico de Ramachandran, também é possível verificar a estereoquímica de ângulos de torção da cadeia lateral. Como mostra a Figura 31, os ângulos de torção Chi são representados em um modelo atômico do aminoácido lisina. Esses ângulos variam em número dependendo do tamanho da cadeia lateral do aminoácido. Para a elaboração do gráfico de

Ramachandran dos ângulos de torção da cadeia lateral, somente os ângulos de torção Chi1 e Chi2 são considerados.

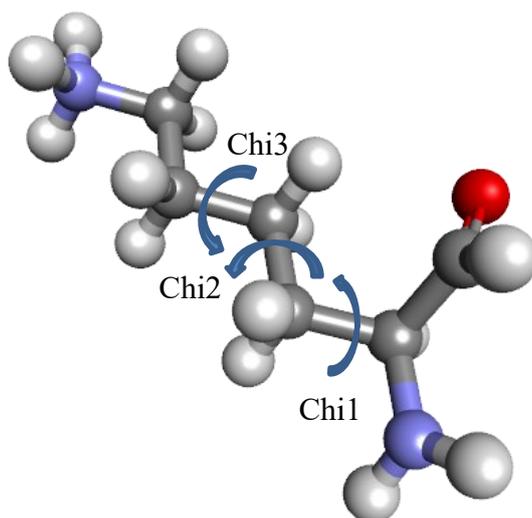
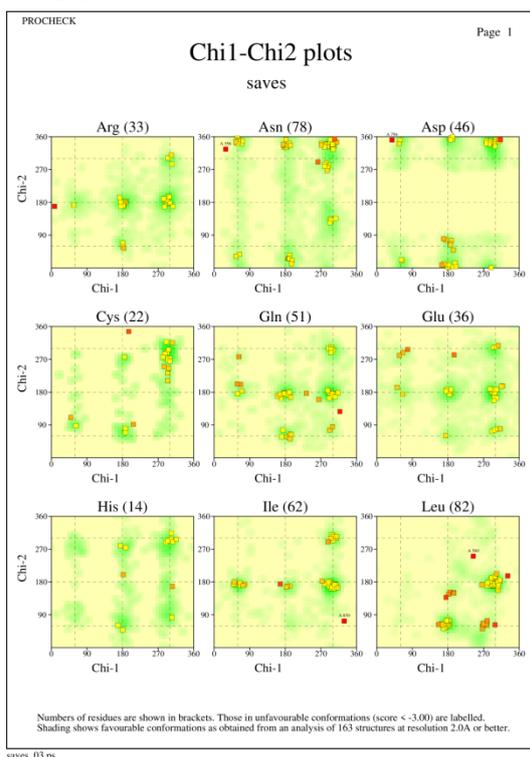


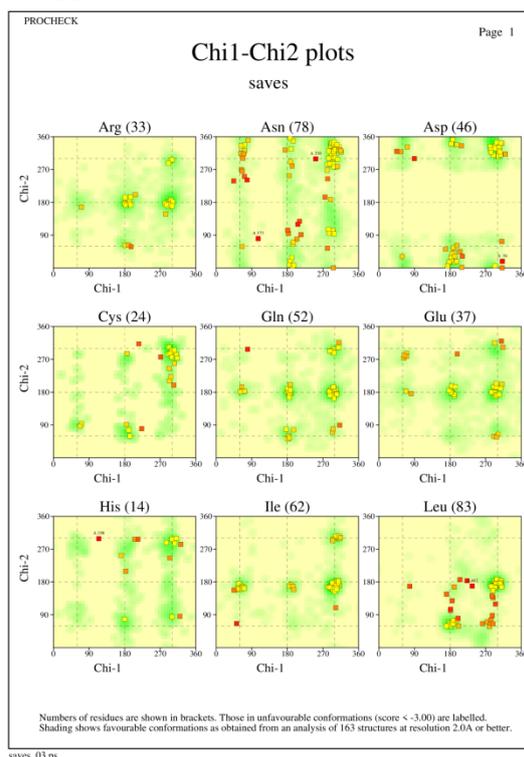
Figura 31. Representação esquemática em forma de esferas e bastões do aminoácido lisina com destaque nos ângulos de torção Chi1, Chi2, Chi3... da cadeia lateral. Esferas em cinza, azul, cinza claro e vermelho representam respectivamente os átomos de carbono, nitrogênio, hidrogênio e oxigênio e os bastões representam as ligações intramoleculares simples e duplas. Fonte do autor. A figura foi construída através das ferramentas de modelagem molecular do software Discovery Studio Visualizer (“Discovery Studio Visualizer 17.2.0.16349”, [s.d.]).

Como aponta a Figura 32, são demonstrados os gráficos de Ramachandran dos ângulos de torção Chi1 e Chi2 da cadeia lateral dos aminoácidos dos modelos atômicos das proteínas spike do SARS-CoV-2 da referência 7CAK-A (32A) e das variantes de mutação pontual E484K (32B) e combinadas K417Y, E484K e N501Y (32C). Os resíduos que pertencem as regiões não permitidas são destacados e identificados pela posição e cadeia proteica. Para calcular a pontuação, foi estabelecida uma comparação com 163 estruturas com resolução de 2 angstroms ou mais.

32A



32B



32C

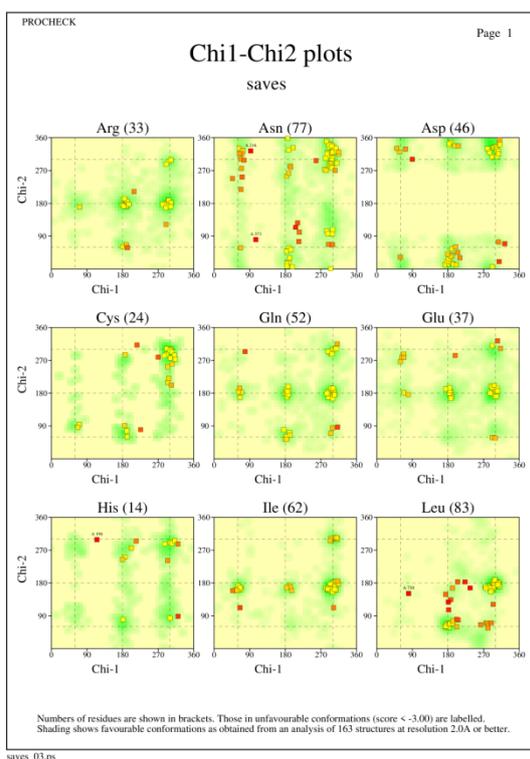


Figura 32. Ângulos das ligações Chi1-Chi2 da cadeia lateral por tipos de resíduos: Ala, Arg, Asn, Asp, Cys, Gln, Glu, Gly e His dos modelos das proteínas spike do SARS-CoV-2 de referência 7CAK-A (32A) e das variantes de mutação pontual E484K (32B) e combinadas K417Y, E484K e N501Y (32C). O número de resíduos de mesmo tipo é mostrado em parênteses. Resíduos destacados e identificados indicam uma pontuação < -3,00 e, portanto, uma conformação desfavorável energeticamente.

Como apresenta a Tabela 15, os resíduos identificados dos modelos atômicos da proteína spike do SARS-CoV-2 de referência 7CAK-A (32A) e das variantes de mutação pontual E484K (32B) e combinadas K417Y, E484K e N501Y (32C) apresentaram uma baixa pontuação conformacional em relação aos ângulos de torção Chi1 e Chi2, o que indica uma baixa qualidade residual. Desconsiderando a referência, somente os resíduos Asn330, Asn373 e Leu 467 do modelo S1E e Asn216 e Asn373 do modelo S1F estão inseridos no RBD (319-541 aminoácidos em relação à sequência de Wuhan) e estão representados na tabela. Isso sugere que possíveis erros conformacionais nesses resíduos não devem causar grandes interferências negativas no docking molecular.

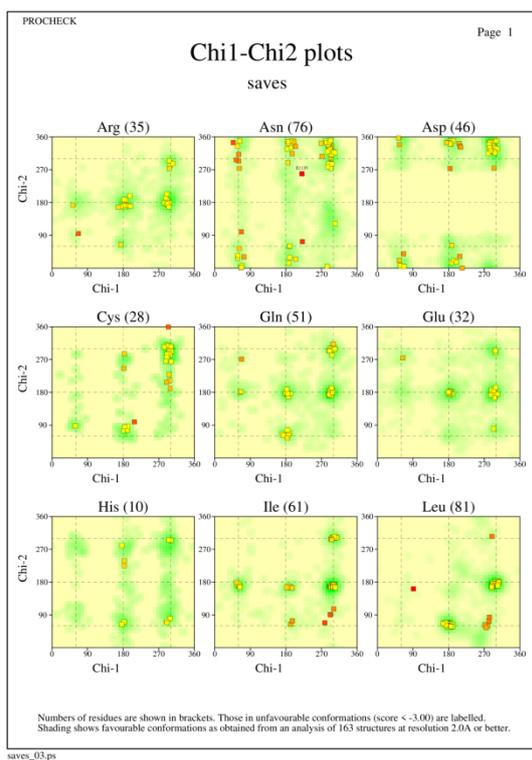
Tabela 15. Posições dos resíduos cuja conformação é desfavorável pelo gráfico de Ramachandran dos ângulos de Chi1-Chi2 nos modelos da proteína spike do SARS-CoV-2 do primeiro grupo.

	Grupo 1 proteína spike SARS-CoV-2		
	S1A	S1E	S1F
Arg			
Asn	556	330, 373	216, 373
Asp	769	54	
Cys			
Gln			
Glu			
His		198	198
Ile	850		
Leu	560	467	730

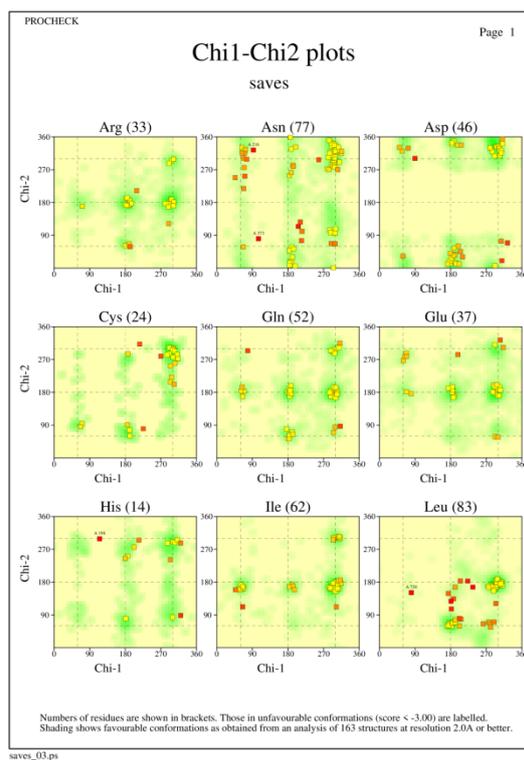
Tabela 15. Posições dos resíduos cuja conformação é desfavorável pelo gráfico de Ramachandran dos ângulos de Chi1-Chi2 nos modelos da proteína spike do SARS-CoV-2 do primeiro grupo. Tabela de dados do gráfico de Ramachandran por resíduos dos ângulos de torção Chi1-Chi2 dos modelos atômicos da proteína spike de referência 7CAK-A (S1A) e as variantes de mutações pontuais E484K (S1E) e K417Y, E484K e N501Y (S1F). Devido à combinação de ângulos Chi1-Chi2, os resíduos descritos se encontram em regiões não permitidas do gráfico de Ramachandran.

Como aponta a Figura 33, são demonstrados os gráficos de Ramachandran dos ângulos de torção Chi1 e Chi2 da cadeia lateral dos aminoácidos dos modelos atômicos referentes à cadeia E da proteína spike do SARS-CoV-2 de referência 7CAK-A (32A) e das variantes de mutação pontual E484K (32B) e combinadas K417Y, E484K e N501Y (32C). Os resíduos que pertencem às regiões não permitidas são destacados e identificados com posição e cadeia proteica. Para calcular a pontuação, foi estabelecida uma comparação com 163 estruturas com resolução de 2 angstroms ou mais.

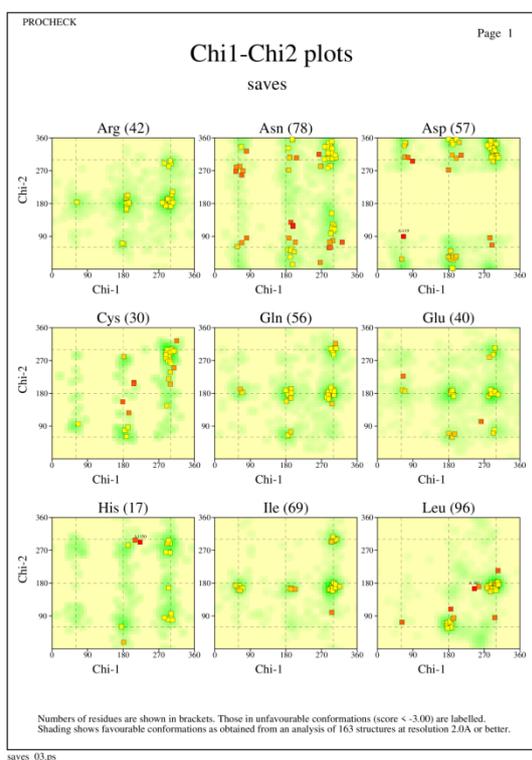
33A



33B



33C



33D

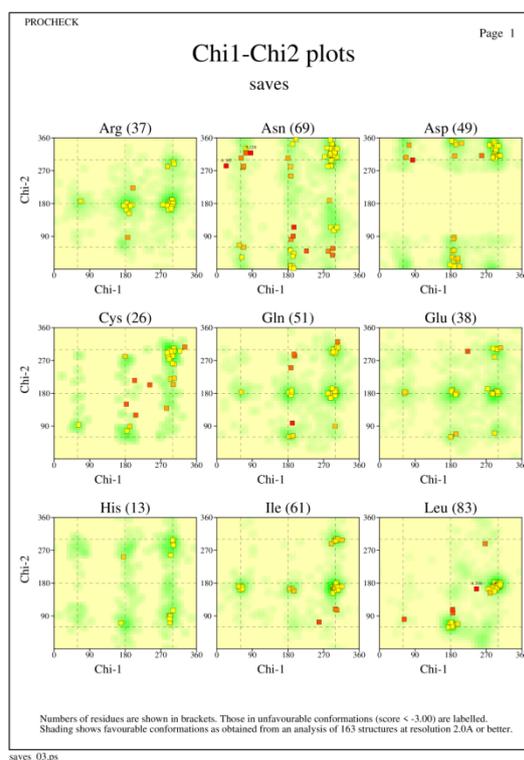


Figura 33. Ângulos das ligações Chi1-Chi2 da cadeia lateral por tipos de resíduos: Ala, Arg, Asn, Asp, Cys, Gln, Glu, Gly e His dos modelos atômicos referentes à cadeia E da proteína spike do SARS-CoV-2 de referência (código PDB: 7K43) (33A), da estrutura molecular original de Wuhan do grupo 2 (33B), da variante Omicron (33C) e da subvariante BA.2 (33D). O número de resíduos de mesmo tipo é mostrado em parênteses. Resíduos

destacados e identificados indicam uma pontuação < -3,00 e, portanto, uma conformação desfavorável energeticamente.

Como apresenta a Tabela 16, os resíduos identificados dos modelos atômicos da proteína spike do SARS-CoV-2 de referência 7K43 (S2A), da estrutura molecular original da proteína de Wuhan do grupo 2 (S2B), da variante omicron (S2F) e da subvariante BA.2 (S2G) apresentaram uma baixa pontuação conformacional em relação aos ângulos de torção Chi1 e Chi2, o que indica uma baixa qualidade residual. Desconsiderando a referência, somente os resíduos Asn373 do modelo S2B, Asp419 e Leu381 do modelo S2F e Asn165 e Leu238 do modelo S2G estão inseridos no RBD (319-541 aminoácidos em relação à sequência de Wuhan) e estão representados na tabela. Isso sugere que possíveis erros conformacionais nesses resíduos não devem causar grandes interferências negativas no docking molecular.

Tabela 16. Posições dos resíduos cuja conformação é desfavorável pelo gráfico de Ramachandran dos ângulos de Chi1-Chi2 nos modelos da proteína spike do SARS-CoV-2 do segundo grupo.

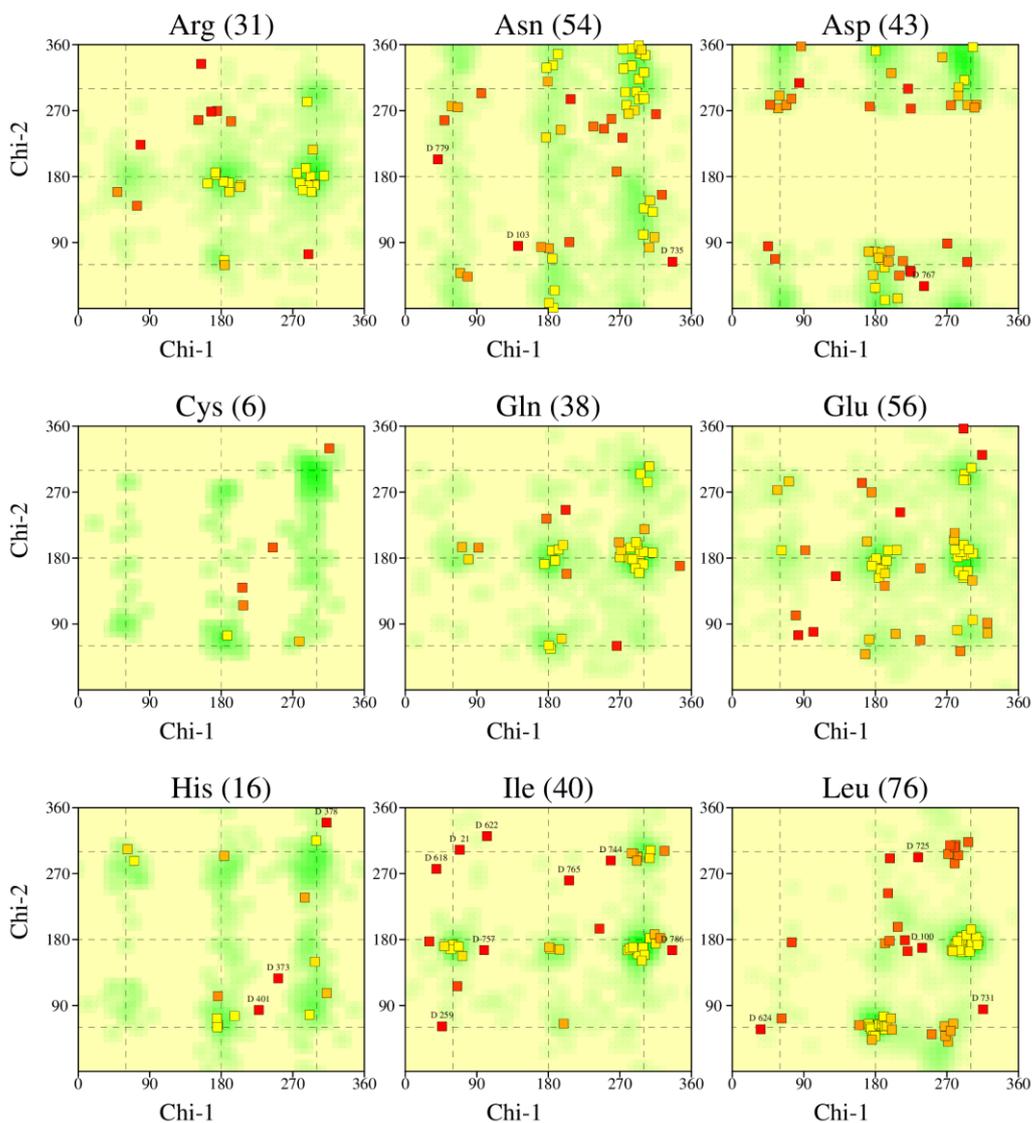
Grupo 2 proteína spike SARS-CoV-2				
	S2A	S2B	S2F	S2G
Arg				
Asn	1135	216, 373		128, 165
Asp			419	
Cys				
Gln				
Glu				
His		198	1150	
Ile				
Leu		730	381	238

Tabela 16. Posições dos resíduos cuja conformação é desfavorável pelo gráfico de Ramachandran dos ângulos de Chi1-Chi2 nos modelos da proteína spike do SARS-CoV-2 do segundo grupo. Tabela de dados do gráfico de Ramachandran por resíduos dos ângulos de torção Chi1-Chi2 dos modelos atômicos referentes à cadeia E da estrutura de referência (código PDB: 7K43) (S2A), da estrutura molecular original de Wuhan do grupo 2 (S2B), variante Omicron (S2C) e subvariante BA.2 (S2D). Devido à combinação de ângulos Chi1-Chi2, os resíduos descritos se encontram em regiões não permitidas do gráfico de Ramachandran.

Como aponta a Figura 34, são demonstrados os gráficos de Ramachandran dos ângulos de torção Chi1 e Chi2 da cadeia lateral dos aminoácidos do modelo atômico da Isoforma 1 de ECA2. Os resíduos que pertencem às regiões não permitidas são destacados e identificados com posição e cadeia proteica. Para calcular a pontuação, foi estabelecida uma comparação com 163 estruturas com resolução de 2 angstroms ou mais.

Chi1-Chi2 plots

saves



Numbers of residues are shown in brackets. Those in unfavourable conformations (score < -3.00) are labelled. Shading shows favourable conformations as obtained from an analysis of 163 structures at resolution 2.0Å or better.

saves_03.ps

Figura 34. Ângulos das ligações Chi1-Chi2 da cadeia lateral por tipos de resíduos: Ala, Arg, Asn, Asp, Cys, Gln, Glu, Gly e His do modelo atômico referente à isoforma 1 de ECA2. O número de resíduos de mesmo tipo é

mostrado em parênteses. Resíduos destacados e identificados indicam uma pontuação < -3,00 e, portanto, uma conformação desfavorável energeticamente.

Como apresenta a Tabela 17, os resíduos identificados do modelo atômico da isoforma 1 de ECA2 apresentaram uma baixa pontuação conformacional em relação aos ângulos de torção Chi1 e Chi2, o que indica uma baixa qualidade residual. Dos indicados, somente o resíduo Gly21 que está no sítio de interação da ECA2 com a proteína spike do SARS-CoV-2 aparece na tabela. Isso sugere que possíveis erros conformacionais nesse resíduo não devem causar grandes interferências negativas no docking molecular.

Tabela 17. Posições dos resíduos cuja conformação é desfavorável pelo gráfico de Ramachandran dos ângulos de Chi1-Chi2 nos modelos da isoforma 1 de ECA2.

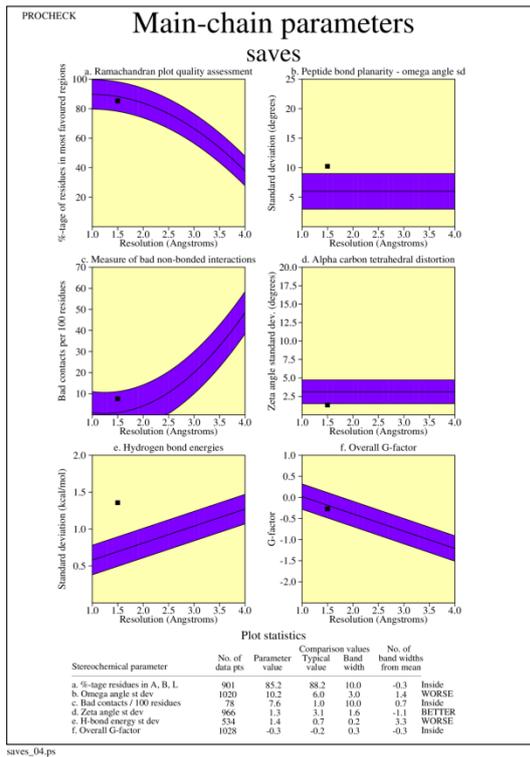
Resíduo	Posição do resíduo na isoforma 1 de ECA2
Ala	
Arg	103, 735, 779
Asn	767
Asp	
Cys	
Gln	
Glu	373, 378, 401
Gly	21, 259, 618, 622, 744, 757, 765, 786
His	100, 624, 725, 731

Tabela 17. Posições dos resíduos cuja conformação é desfavorável pelo gráfico de Ramachandran dos ângulos de Chi1-Chi2 nos modelos da isoforma 1 de ECA2. Tabela de dados do gráfico de Ramachandran por resíduos dos ângulos de torção Chi1-Chi2 do modelo atômico referentes à isoforma 1 de ECA2. Devido à combinação de ângulos Chi1-Chi2, os resíduos descritos se encontram em regiões não permitidas do gráfico de Ramachandran.

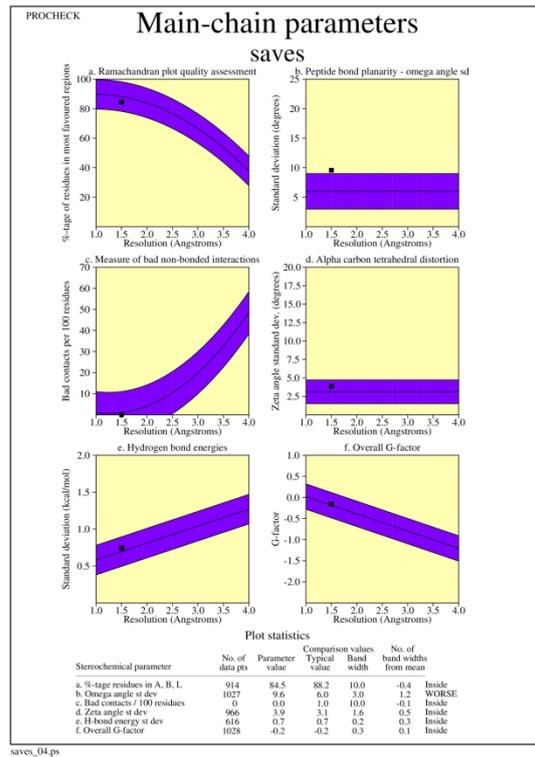
Através das análises do PROCHECK, foi viável aferir e avaliar os parâmetros da cadeia principal dos modelos atômicos.

Como mostra a Figura 35, os parâmetros de porcentagem de resíduos nas regiões mais favoráveis, desvio do ângulo ômega, contatos ruins/100 resíduos, desvio do ângulo zeta, desvio de energia de ligações de hidrogênio (H) e fator-G geral dos modelos atômicos das proteínas spike do SARS-CoV-2 da referência 7CAK-A (35A) e das variantes de mutação pontual E484K (35B) e combinadas K417Y, E484K e N501Y (35C) foram avaliados quantitativa e qualitativamente. Esses parâmetros são plotados em gráficos em função da resolução.

35A



35B



35C

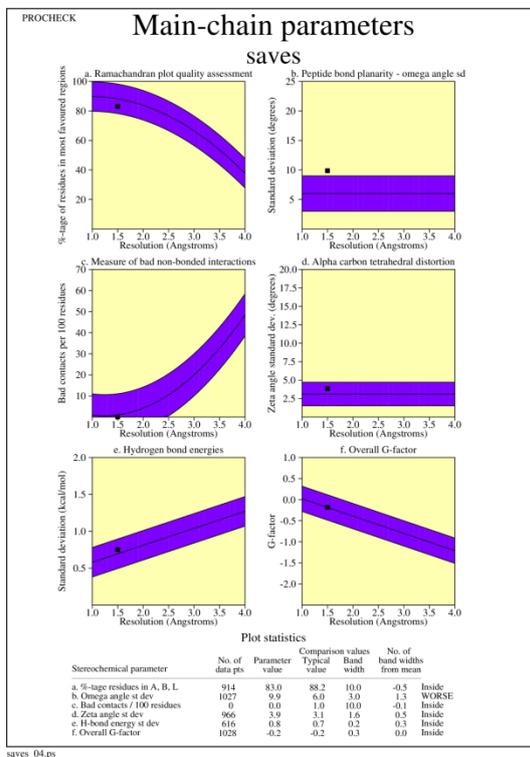
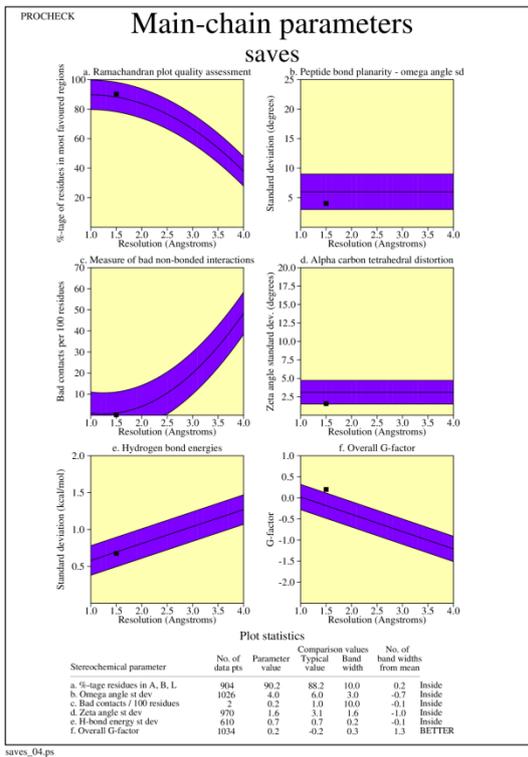


Figura 35. Parâmetros da cadeia principal dos modelos atômicos das proteínas spike do SARS-CoV-2 da referência 7CAK-A (35A) e das variantes de mutação pontual E484K (35B) e combinadas K417Y, E484K e N501Y (35C). Os pontos pretos representam os resultados do modelo e as partes coloridas em índigo elétrico representam a região esperada para um modelo de alta qualidade. Pontos dentro das regiões destacadas indicam uma alta qualidade do modelo. Valores de G-factor (f) acima da região destacada são considerados resultados

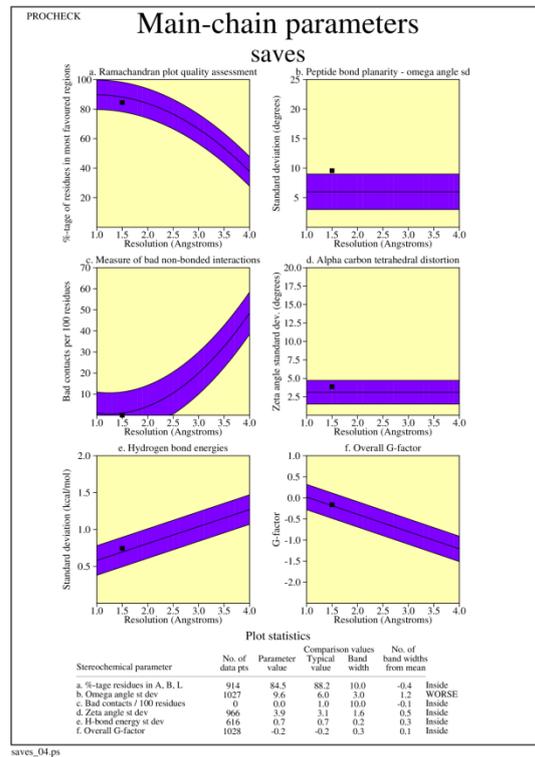
melhores e abaixo piores que o esperado para uma alta qualidade do modelo. Para as demais variáveis (a, b, c, d, e), valores acima da região destacada são considerados piores e abaixo melhores que o esperado para uma alta qualidade do modelo.

Como mostra a Figura 36, os parâmetros de porcentagem de resíduos nas regiões mais favoráveis, desvio do ângulo ω , contatos ruins/100 resíduos, desvio do ângulo zeta, desvio de energia de ligações de hidrogênio (H) e fator-G geral dos modelos atômicos da proteína spike do SARS-CoV-2, cadeia E da estrutura de referência (código PDB: 7K43) (36A), da estrutura molecular original de Wuhan do grupo 2 (36B), da variante Omicron (36C) e da subvariante BA.2 (36D) foram avaliados quantitativa e qualitativamente. Esses parâmetros são plotados em gráficos em função da resolução.

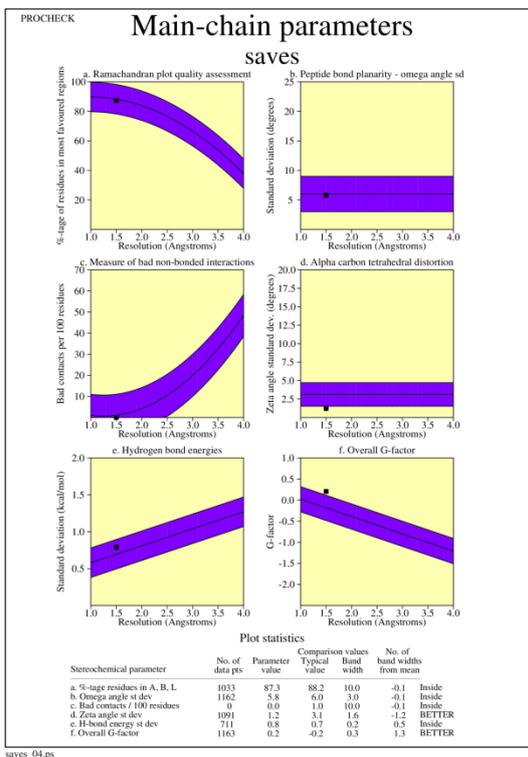
36A



36B



36C



36D

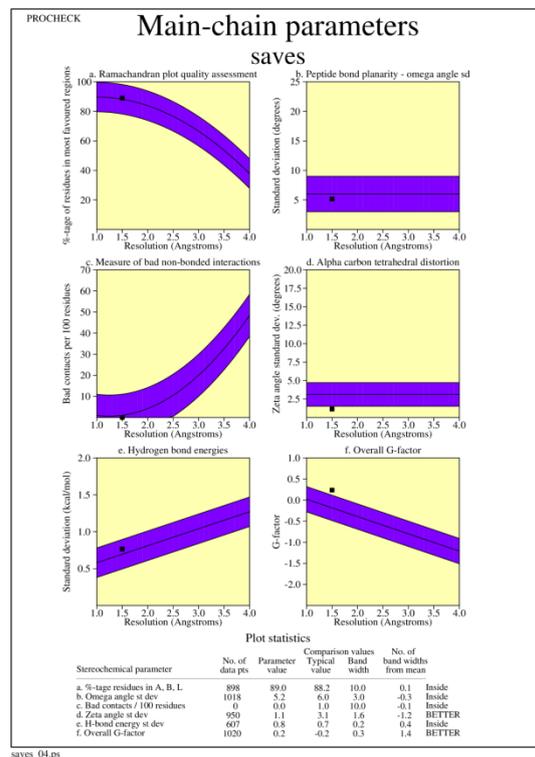
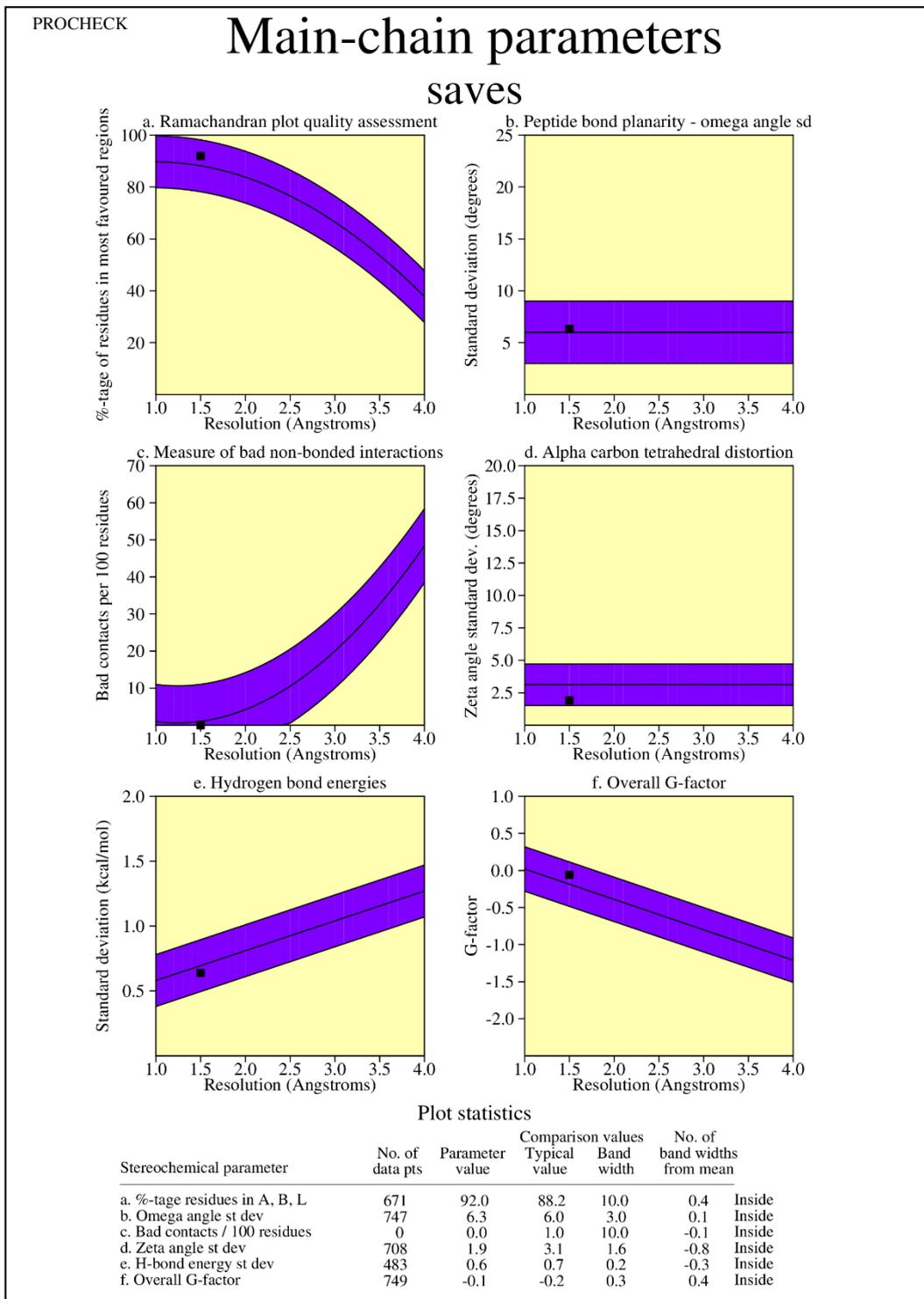


Figura 36. Parâmetros da cadeia principal dos modelos atômicos da proteína spike do SARS-CoV-2 referentes à cadeia E da estrutura de referência (código PDB: 7K43) (36A), da estrutura molecular original de Wuhan do grupo 2 (36B), variante Omicron (36C) e subvariante BA.2 (36D). Os pontos pretos representam os resultados do modelo e as partes coloridas em índigo elétrico representam a região esperada para um modelo de alta

qualidade. Pontos dentro das regiões destacadas indicam uma alta qualidade do modelo. Valores de G-factor (f) acima da região destacada são considerados resultados melhores e abaixo piores que o esperado para uma alta qualidade do modelo. Para as demais variáveis (a, b, c, d, e), valores acima da região destacada são considerados piores e abaixo melhores que o esperado para uma alta qualidade do modelo.

Como mostra a Figura 37, os parâmetros de porcentagem de resíduos nas regiões mais favoráveis, desvio do ângulo ω , contatos ruins/100 resíduos, desvio do ângulo zeta, desvio de energia de ligações de hidrogênio (H) e fator-G geral do modelo atômico da isoforma 1 de ECA2 foram avaliados quantitativa e qualitativamente. Esses parâmetros são plotados em gráficos em função da resolução.



saves_04.ps

Figura 37. Parâmetros da cadeia principal do modelo atômico da isoforma 1 de ECA2. Os pontos pretos representam os resultados do modelo e as partes coloridas em índigo elétrico representam a região esperada para um modelo de alta qualidade. Pontos dentro das regiões destacadas indicam uma alta qualidade do modelo. Valores de G-factor (f) acima da região destacada são considerados resultados melhores e abaixo piores que o

esperado para uma alta qualidade do modelo. Para as demais variáveis (a, b, c, d, e), valores acima da região destacada são considerados piores e abaixo melhores que o esperado para uma alta qualidade do modelo.

Como mostra a Tabela 18, os parâmetros da cadeia principal de porcentagem de resíduos nas regiões mais favoráveis, desvio do ângulo ômega, contatos ruins/100 resíduos, desvio do ângulo zeta, desvio de energia de ligações de hidrogênio (H) e fator-G geral dos modelos atômicos da proteína spike do SARS-CoV-2 da referência 7CAK-A (S1A) e das variantes de mutação pontual E484K (S1E), combinadas K417Y, E484K e N501Y (S1F), da referência 7K43-E (S2A), da estrutura molecular original de Wuhan do grupo 2 (S2B), da variante Omicron (S2F), da subvariante BA.2 (S2G) e Isoforma 1 de ECA2 (E1A) foram avaliados qualitativamente. Até mesmo parâmetros da cadeia principal referentes ao desvio do ângulo ômega e desvio de energia de ligações de hidrogênio dos modelos de referência (S1A e S2A) foram observados como piores que o esperado para um modelo de alta qualidade.

Tabela 18. Análise qualitativa dos parâmetros da cadeia principal dos modelos atômicos da proteína spike do SARS-CoV-2 e da isoforma 1 de ECA2.

Modelo atômico	Qualidade (dentro, melhor ou pior que o esperado para alta qualidade)					
	Porcentagem de resíduos nas regiões mais favoráveis	Desvio do ângulo ômega	Contatos ruins/100 resíduos	Desvio do ângulo zeta	Desvio de energia de ligações de H	Fator-G geral
S1A	Dentro	Pior	Dentro	Melhor	Pior	Dentro
S1E	Dentro	Pior	Dentro	Dentro	Dentro	Dentro
S1F	Dentro	Pior	Dentro	Dentro	Dentro	Dentro
S2A	Dentro	Dentro	Dentro	Dentro	Dentro	Melhor
S2B	Dentro	Pior	Dentro	Dentro	Dentro	Dentro
S2F	Dentro	Dentro	Dentro	Melhor	Dentro	Melhor
S2G	Dentro	Dentro	Dentro	Melhor	Dentro	Melhor
E1A	Dentro	Dentro	Dentro	Melhor	Dentro	Melhor

Tabela 18. Análise qualitativa dos parâmetros da cadeia principal dos modelos atômicos da proteína spike do SARS-CoV-2 e da isoforma 1 de ECA2. Análise qualitativa dos modelos referentes a proteína spike do SARS-CoV-2: de referência 7CAK-A (S1A), da variante com mutação pontual E484K (S1E) e com mutações combinadas K417Y, E484K e N501Y (S1F) do grupo 1, de referência 7K43 (S2A), estrutura original de Wuhan (S2B), variante Omicron (S2C) e subvariante BA.2 (S2D) do grupo 2; e isoforma 1 de ECA2.

Pelas análises do PROCHECK, também foi possível descrever parâmetros da cadeia lateral dos modelos atômicos. Dentre os parâmetros, se encontram os ângulos de gauche positivo (g+), gauche negativo (g-) e gauche trans que foram esquematizados no aminoácido cisteína para exemplificar as suas localizações em relação aos átomos na Figura 38.

A partir das figuras Figura 39, Figura 40 e Figura 41, foi possível descrever parâmetros relacionados à cadeia lateral dos modelos da proteína spike do SARS-CoV-2: de referência 7CAK-A (S1A), da variante com mutação pontual E484K (S1E) e com mutações combinadas K417Y, E484K e N501Y (S1F) do grupo 1, de referência 7K43 (S2A), estrutura original de Wuhan (S2B), variante Omicron (S2F) e subvariante BA.2 (S2G) do grupo 2; e isoforma 1 de ECA2. Os parâmetros consistem no desvio em graus dos ângulos gauche negativo (a), trans (b), positivo (c), e desvio padrão agrupado dos ângulos Chi-1 (d) e desvios padrão dos ângulos trans Chi-2 da cadeia lateral. Os ângulos de gauche positivo (g+), gauche negativo (g-) e gauche trans foram esquematizados no aminoácido cisteína para exemplificar as suas localizações em relação aos átomos (Figura 38).

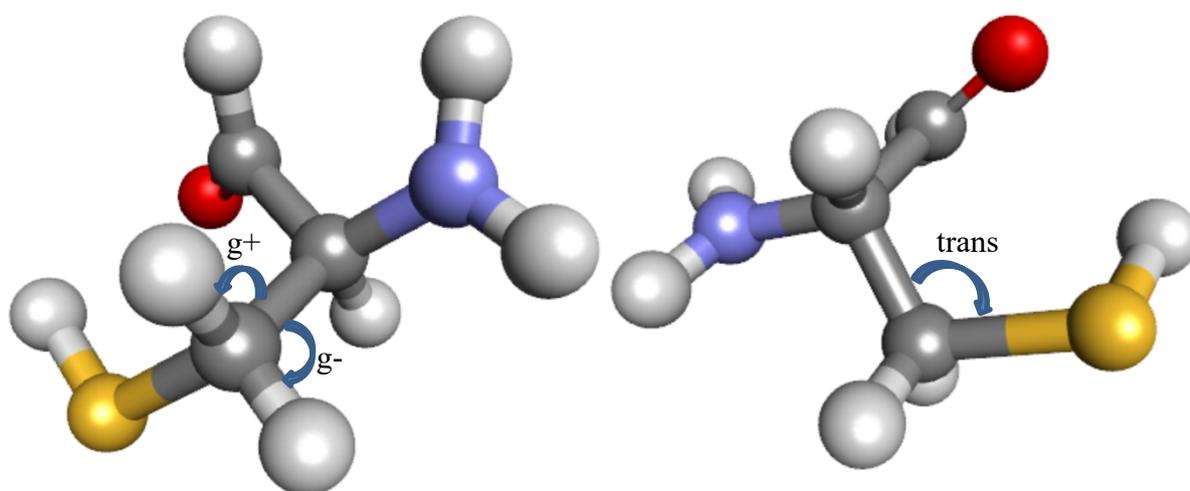
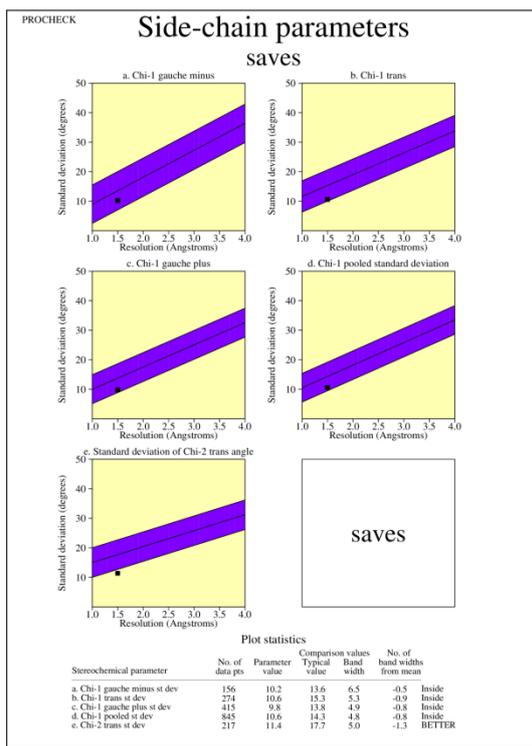


Figura 38. Representação esquemática em forma de esferas e bastões do aminoácido cisteína de frente (esquerda) e por trás (direita) com destaque nos ângulos de gauche positivo (g+), gauche negativo (g-) e gauche trans da cadeia lateral. Esferas em cinza, azul, cinza claro, vermelhas e amarelas representam respectivamente os átomos de carbono, nitrogênio, hidrogênio, oxigênio e enxofre e os bastões representam as ligações intramoleculares simples e duplas. Fonte do autor. A figura foi construída através das ferramentas de modelagem molecular do software Discovery Studio Visualizer (“Discovery Studio Visualizer 17.2.0.16349”, [s.d.]).

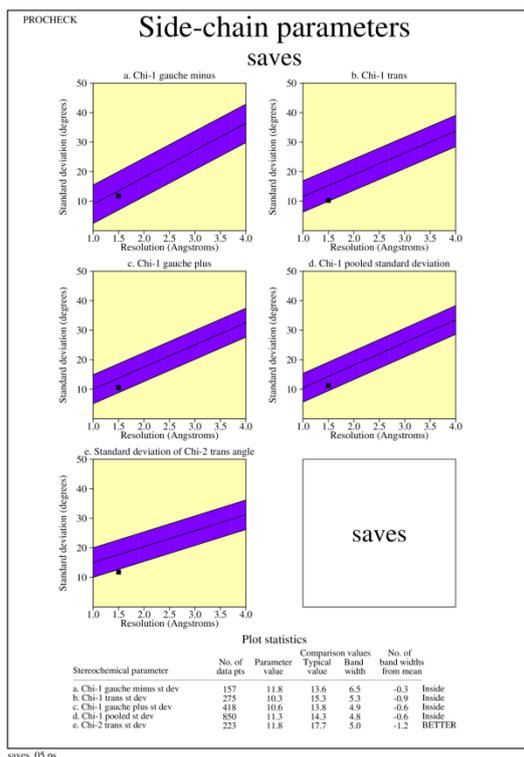
Como apresenta a Figura 39, os parâmetros da cadeia lateral de desvio em graus dos ângulos gauche negativo (a), trans (b), positivo (c), desvio padrão agrupado dos ângulos Chi-1 (d) e desvios padrão dos ângulos trans Chi-2 da cadeia lateral dos modelos atômicos das proteínas spike do SARS-CoV-2 da referência 7CAK-A (39A) e das variantes de mutação pontual E484K (39B) e combinadas K417Y, E484K e N501Y (39C) foram avaliados

quantitativa e qualitativamente. Esses parâmetros são plotados em gráficos em função da resolução.

39A



39B



39C

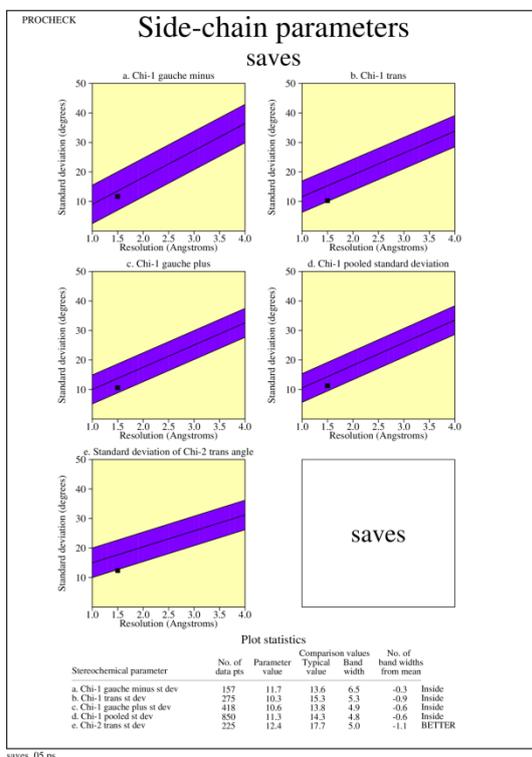
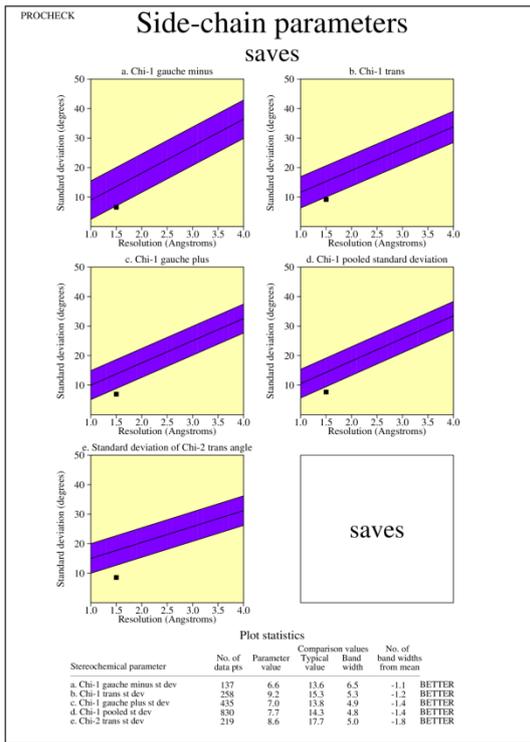


Figura 39. Parâmetros da cadeia lateral de desvio em graus dos ângulos gauche negativo (a), trans (b), positivo (c), e desvio padrão agrupado dos ângulos Chi-1 (d) e desvios padrão dos ângulos trans Chi-2 da cadeia lateral (e) dos modelos atômicos das proteínas spike do SARS-CoV-2 da referência 7CAK-A (39A) e das variantes de mutação pontual E484K (39B) e combinadas K417Y, E484K e N501Y (39C). Os pontos pretos representam os resultados do modelo e as partes coloridas em índigo elétrico representam a região esperada para um modelo de

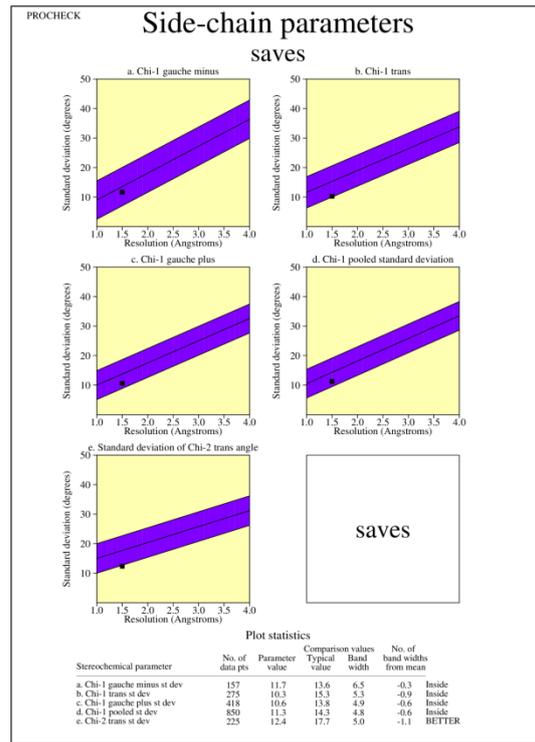
alta qualidade. Valores acima da região são considerados piores e abaixo melhores que o esperado para uma alta qualidade do modelo.

Como apresenta a Figura 40, os parâmetros da cadeia lateral de desvio em graus dos ângulos gauche negativo (a), trans (b), positivo (c), desvio padrão agrupado dos ângulos Chi-1 (d) e desvios padrão dos ângulos trans Chi-2 da cadeia lateral dos modelos atômicos da proteína spike do SARS-CoV-2 de referência 7K43-E (40A), da estrutura molecular original de Wuhan do grupo 2 (40B), variante Omicron (40C) e subvariante BA.2 (40D) foram avaliados quantitativa e qualitativamente. Esses parâmetros são plotados em gráficos em função da resolução.

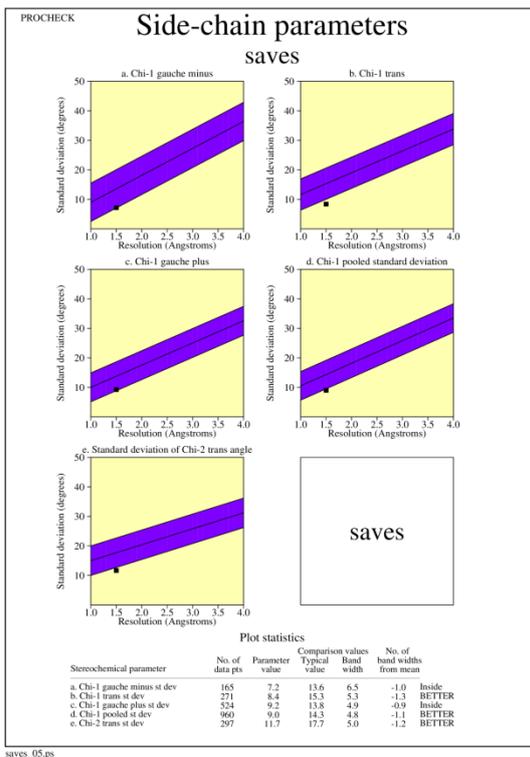
40A



40B



40C



40D

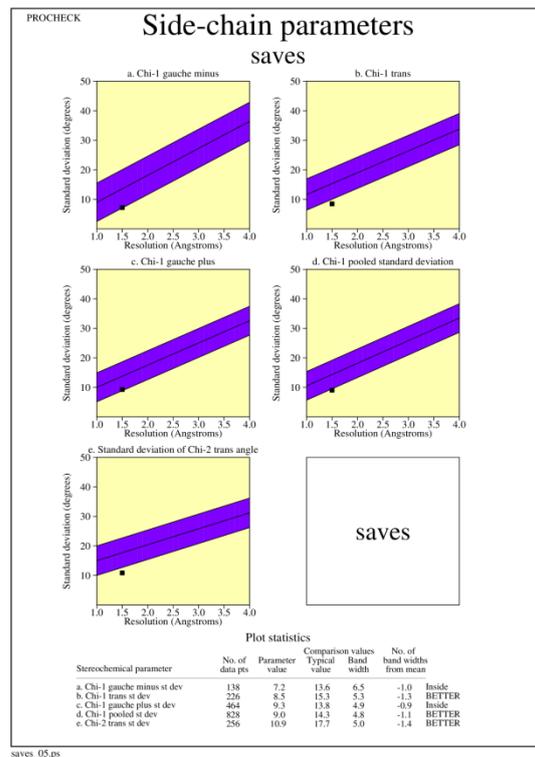
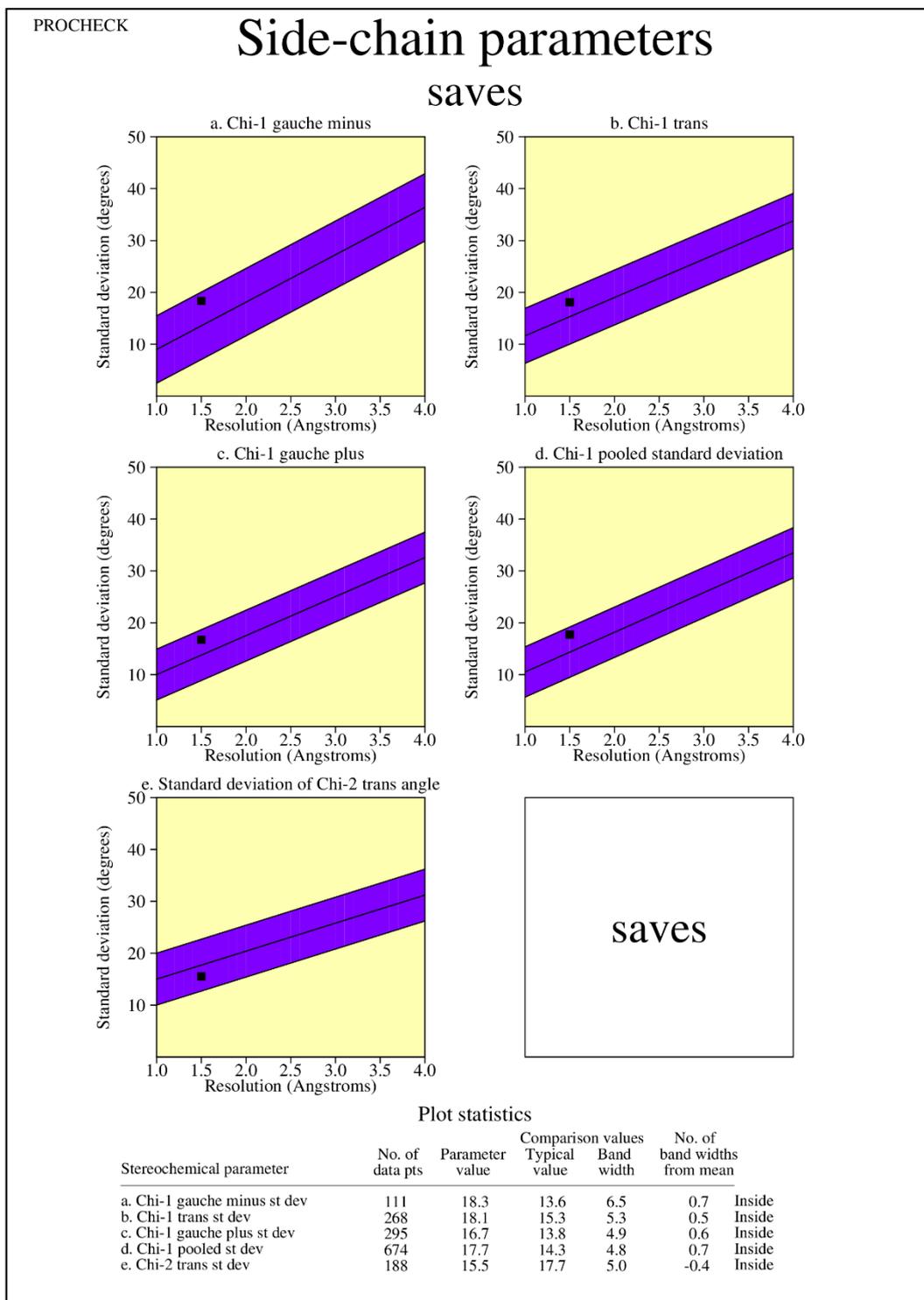


Figura 40. Parâmetros da cadeia lateral de desvio em graus dos ângulos gauche negativo (a), trans (b), positivo (c), desvio padrão agrupado dos ângulos Chi-1 (d) e desvios padrão dos ângulos trans Chi-2 da cadeia lateral (e) dos modelos atômicos das proteínas spike do SARS-CoV-2 da referência 7K43-E (40A), da estrutura molecular original de Wuhan do grupo 2 (40B), variante Omicron (40C) e subvariante BA.2 (40D). Os pontos pretos representam os resultados do modelo e as partes coloridas em índigo elétrico representam a região esperada para

um modelo de alta qualidade. Valores acima da região são considerados piores e abaixo melhores que o esperado para uma alta qualidade do modelo.

Como apresenta a Figura 41, os parâmetros da cadeia lateral de desvio em graus dos ângulos gauche negativo (a), trans (b), positivo (c), desvio padrão agrupado dos ângulos Chi-1 (d) e desvios padrão dos ângulos trans Chi-2 da cadeia lateral do modelo atômico da isoforma de ECA2 foram avaliados quantitativa e qualitativamente. Esses parâmetros são plotados em gráficos em função da resolução.



saves_05.ps

Figura 41. Parâmetros da cadeia lateral de desvio em graus dos ângulos gauche negativo (a), trans (b), positivo (c), desvio padrão agrupado dos ângulos Chi-1 (d) e desvios padrão dos ângulos trans Chi-2 da cadeia lateral (e) do modelo atômico da isoforma 1 de ECA2. Os pontos pretos representam os resultados do modelo e as partes coloridas em índigo elétrico representam a região esperada para um modelo de alta qualidade. Valores acima da região são considerados piores e abaixo melhores que o esperado para uma alta qualidade do modelo.

Como apresenta a Tabela 19, são inferidas qualificações em dentro, melhor ou pior que o esperado para alta qualidade dos modelos atômicos da proteína spike do SARS-CoV-2 da referência 7CAK-A (S1A) e das variantes de mutação pontual E484K (S1E), combinadas K417Y, E484K e N501Y (S1F), da referência 7K43-E (S2A), da estrutura molecular original de Wuhan do grupo 2 (S2B), da variante Omicron (S2F), da subvariante BA.2 (S2G) e Isoforma 1 de ECA2 (E1A) em relação aos desvios dos ângulos de gauche negativo Chi-1, trans Chi1, gauche positivo Chi-1, agrupados Chi-1 e trans Chi-2. Todos os modelos apresentaram qualidades no mínimo satisfatórias, destacando os desvios dos ângulos trans Chi-2 que demonstraram na maioria das estruturas, qualidades melhores que o esperado.

Tabela 19. Análise qualitativa dos parâmetros da cadeia lateral dos modelos atômicos da proteína spike do SARS-CoV-2 e da isoforma 1 de ECA2.

Modelo atômico	Qualidade dos desvios dos ângulos (dentro, melhor ou pior que o esperado para alta qualidade)				
	Gauche negativo Chi-1	Trans Chi-1	Gauche positivo Chi-1	Agrupados Chi-1	Trans Chi-2
S1A	Dentro	Dentro	Dentro	Dentro	Melhor
S1E	Dentro	Dentro	Dentro	Dentro	Melhor
S1F	Dentro	Dentro	Dentro	Dentro	Melhor
S2A	Melhor	Melhor	Melhor	Melhor	Melhor
S2B	Dentro	Dentro	Dentro	Dentro	Melhor
S2F	Dentro	Melhor	Dentro	Melhor	Melhor
S2G	Dentro	Melhor	Dentro	Melhor	Melhor
E1A	Dentro	Dentro	Dentro	Dentro	Dentro

Tabela 19. Análise qualitativa dos parâmetros da cadeia lateral dos modelos atômicos da proteína spike do SARS-CoV-2 e da isoforma 1 de ECA2. Parâmetros da cadeia lateral de desvio em graus dos ângulos gauche negativo, trans, positivo, desvio padrão agrupado dos ângulos Chi-1 e desvios padrão dos ângulos trans Chi-2 da cadeia lateral. Dados referentes às proteínas spike do SARS-CoV-2 de referência 7CAK-A (S1A), da variante com mutação pontual E484K (S1E) e com mutações combinadas K417Y, E484K e N501Y (S1F) do grupo 1, de referência 7K43 (S2A), estrutura original de Wuhan (S2B), variante Omicron (S2C) e subvariante BA.2 (S2D) do grupo 2; e isoforma 1 de ECA2.

4.5.4. ProSA

A ferramenta ProSA permite um reconhecimento de possíveis erros em modelos teóricos e experimentais. Utilizando potenciais baseados no conhecimento da força média pelo princípio de Boltzmann (SIPPL, 1990) extraídos de um banco de dados de estruturas conhecidas, o software é capaz de compara-los à sequência de aminoácidos a ser avaliada e assim predizer os potenciais da força média da estrutura. Além disso, o programa calcula a

energia da estrutura por meio do potencial de distância de pares e o potencial da exposição de resíduos ao solvente (SIPPL, 1993; WIEDERSTEIN; SIPPL, 2007).

O Z-score é uma ferramenta para se medir o desvio da média de um determinado padrão (Equação 2):

$$Zscore = \frac{x-\mu}{\sigma} \quad (2)$$

Quando o padrão se refere à energia total esperada para uma molécula, o valor de Z-score indica o desvio da energia total de uma determinada amostra em comparação a esse padrão (SIPPL, 1993, 1995).

Com a análise feita pela ferramenta ProSA foi possível obter informações sobre a qualidade geral (Z-score) de energia livre das proteínas spike da variante de SARS-CoV-2 de referência 7CAK-A, com mutação pontual E484K, e combinadas K417Y, E484K e N501Y de referência 7K43-E, original de Wuhan e variantes de preocupação Omicron e subvariante BA.2. São demonstrados que os modelos que não passaram pelo protocolo de minimização de energia apresentaram regiões de alto valor energético nos gráficos de energia livre das Figura 43, Figura 45 e Figura 47.

Conforme a Figura 42, são demonstrados em forma de gráfico os Z-score obtido e esperado em função do número de resíduos da proteína dos modelos atômicos das proteínas spike da variante de SARS-CoV-2 de referência 7CAK-A (Figura 42A), com mutação pontual E484K (Figura 42B), e combinadas K417Y, E484K e N501Y (Figura 42C).

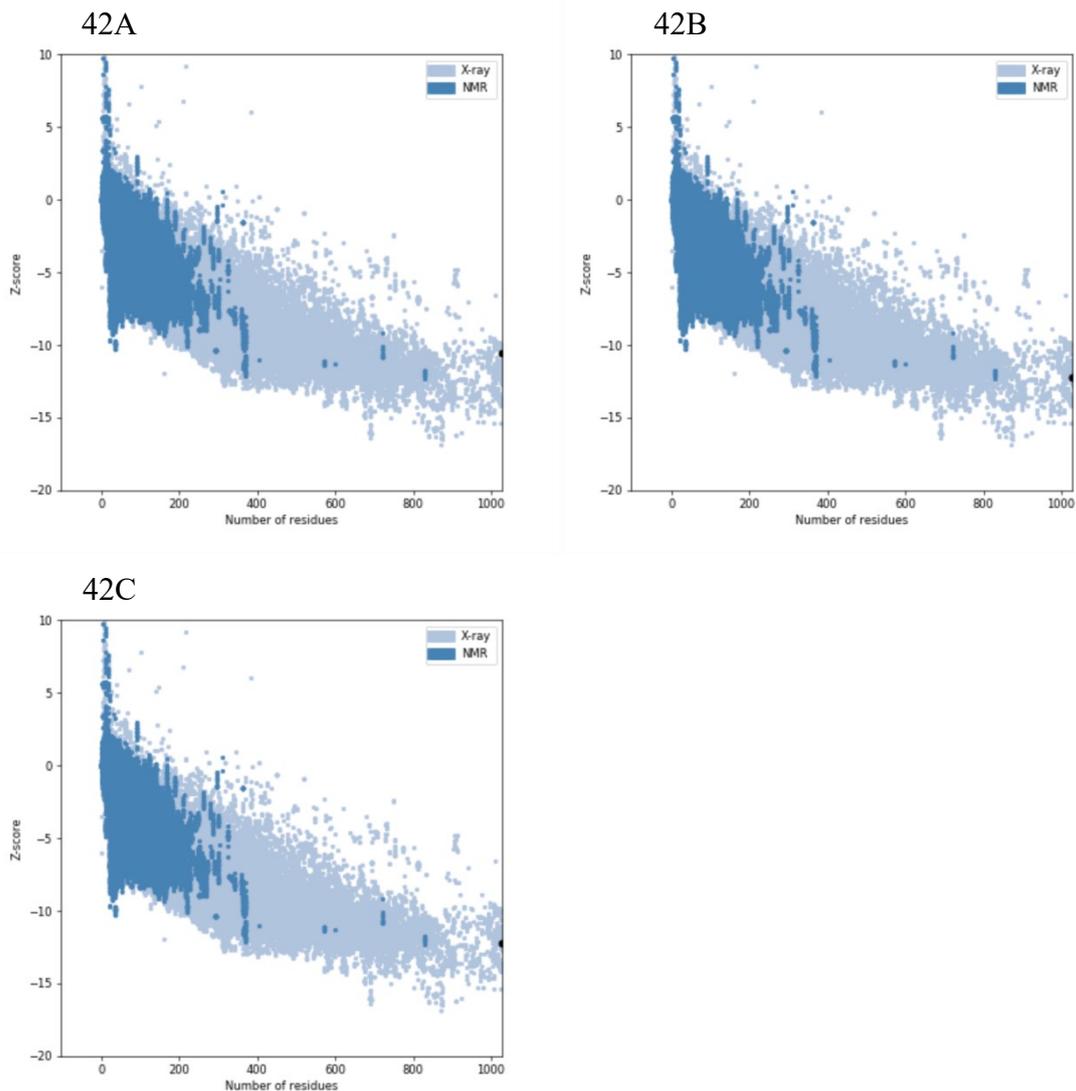


Figura 42. Z-score esperado para modelos atômicos de alta qualidade de cristalografia de raio X e ressonância magnética nuclear (NMR) com base no número de resíduos dos modelos atômicos das proteínas spike do SARS-CoV-2 da referência 7CAK-A (43A) e das variantes de mutação pontual E484K (43B) e combinadas K417Y, E484K e N501Y (43C). Regiões coloridas em azul-azul e azul-azul claro indicam uma alta qualidade esperada para o modelo de cristalografia de raio X e NMR respectivamente. O ponto preto aponta o resultado do modelo.

Conforme a Figura 43, são demonstrados os dados de energia livre em janelas de 10 e (verde claro) e 40 (verde escuro) aminoácidos em função do tamanho da molécula dos modelos atômicos das proteínas spike da variante de SARS-CoV-2 de referência 7CAK-A (Figura 42A), com mutação pontual E484K (Figura 42B), e combinadas K417Y, E484K e N501Y (Figura 42C). Os dados de energia livre da Figura 42A apresentou 4 regiões, identificadas pelos picos, de erros estruturais. Esses erros devem estar relacionados a modelos de estrutura energeticamente desfavoráveis, uma vez que não foi executado o protocolo de minimização de energia no modelo de referência 7CAK-A.

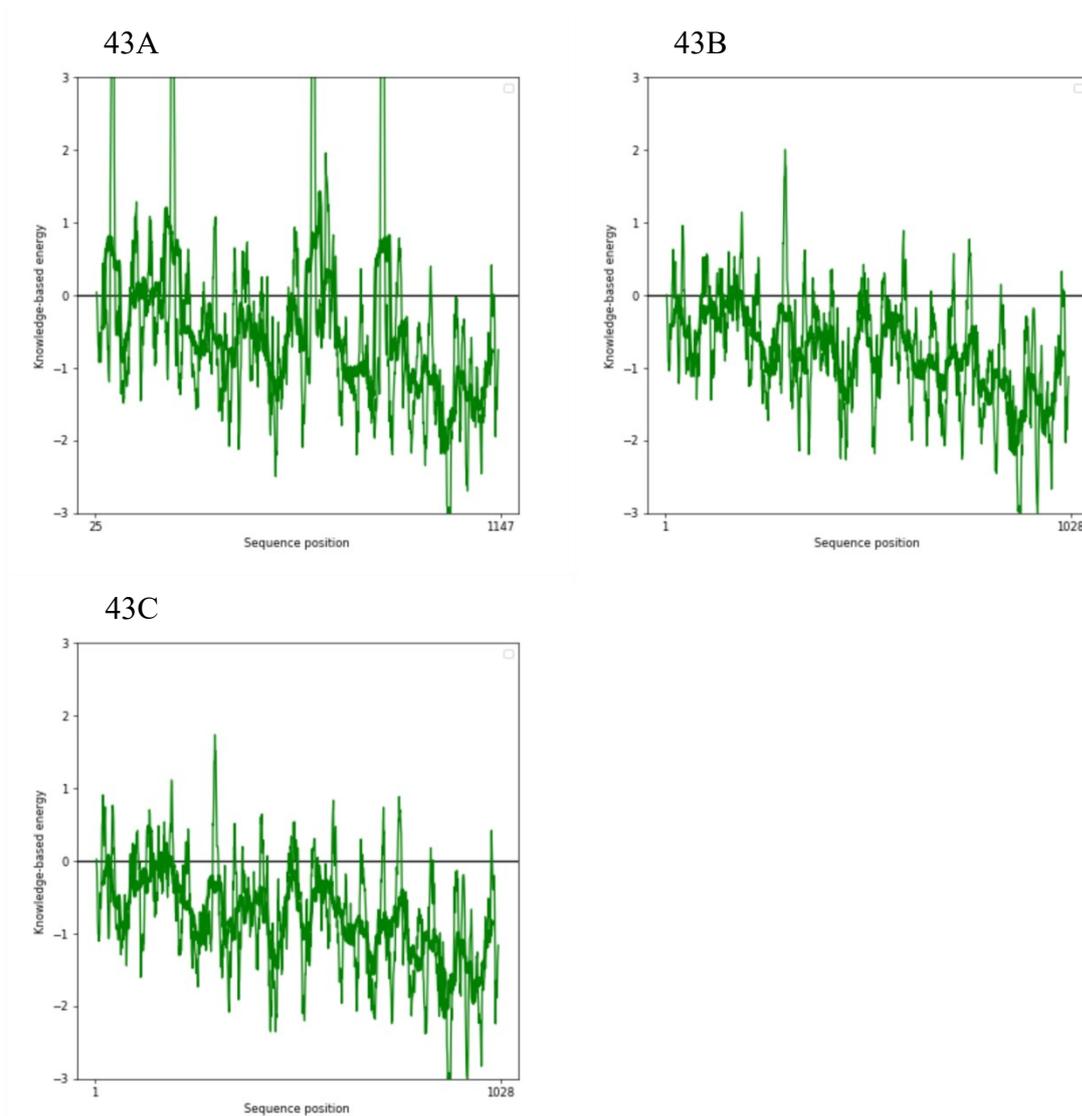


Figura 43. Gráfico de energia livre em seqüências agrupadas de 10 em 10 (verde claro) e 40 em 40 (verde escuro) aminoácidos dos modelos atômicos das proteínas spike do SARS-CoV-2 da referência 7CAK-A (43A) e das variantes de mutação pontual E484K (43B) e combinadas K417Y, E484K e N501Y (43C). Valores positivos indicam erros intrínsecos à conformação da região.

Conforme a Figura 44, são demonstrados em forma de gráfico os Z-score obtido e esperado em função do número de resíduos da proteína dos modelos atômicos das proteínas spike da variante de SARS-CoV-2 de referência 7K43-E (Figura 44A), original de Wuhan (Figura 44B), e variantes de preocupação Omicron (Figura 44C) e subvariante BA.2 (Figura 44D). Os dados de Z-score das estruturas se assemelham a referência, indicando uma qualidade satisfatória dos modelos.

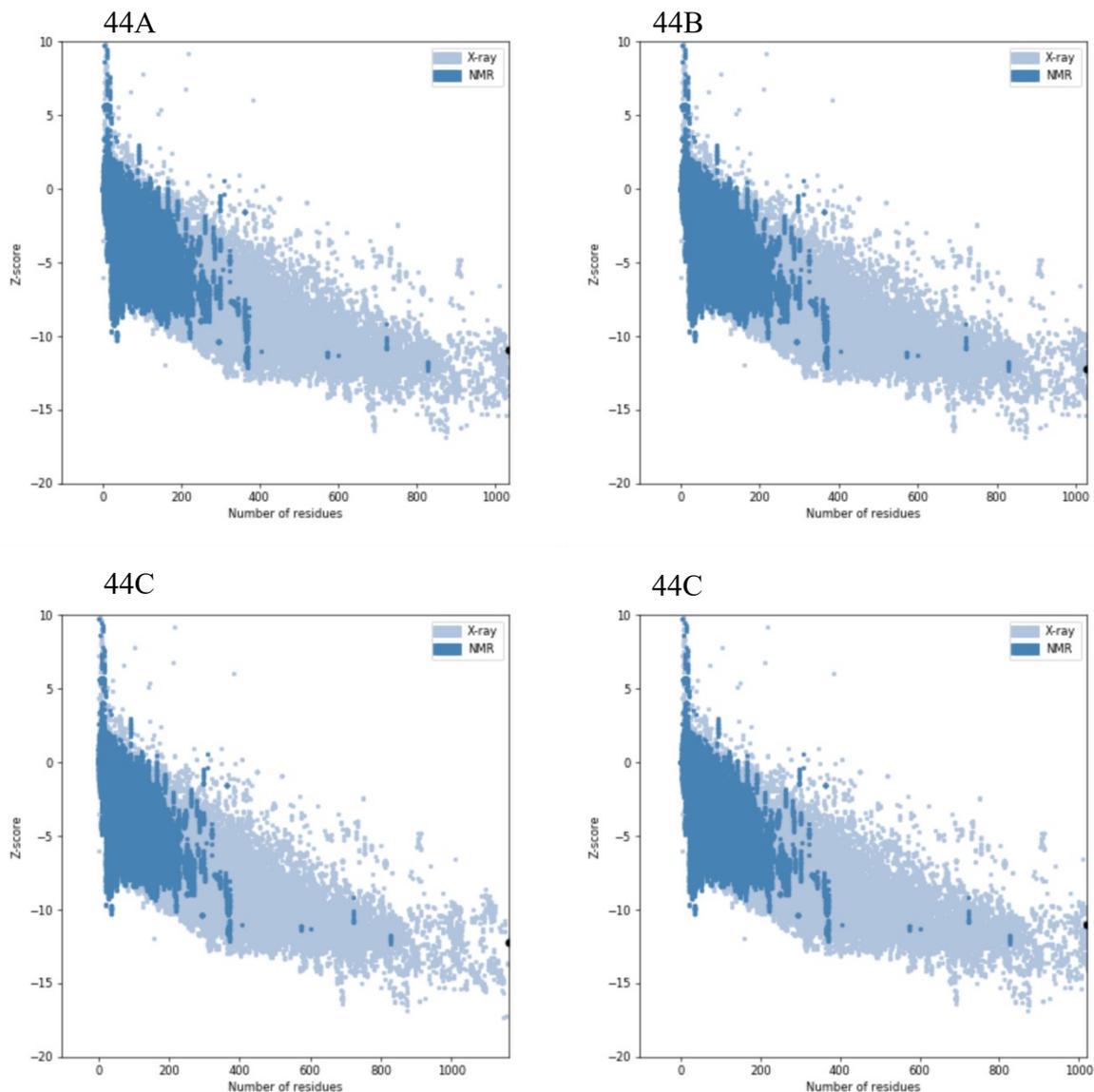


Figura 44. Z-score esperado para modelos atômicos de alta qualidade de cristalografia de raio X e ressonância magnética nuclear (NMR) com base no número de resíduos dos modelos atômicos das proteínas spike do SARS-CoV-2 da referência 7K43-E (44A), da estrutura molecular original de Wuhan do grupo 2 (44B), variante Omicron (44C) e subvariante BA.2 (44D). Regiões coloridas em azul-escuro e azul-claro indicam uma alta qualidade esperada para o modelo de cristalografia de raio X e NMR respectivamente. O ponto preto aponta o resultado do modelo.

Conforme a Figura 45, são demonstrados os dados de energia livre em janelas de 10 e (verde claro) e 40 (verde escuro) aminoácidos em função do tamanho da molécula dos modelos atômicos das proteínas spike da variante de SARS-CoV-2 de referência 7K43-E (Figura 44A), original de Wuhan (Figura 44B), e variantes de preocupação Omicron (Figura 44C) e subvariante BA.2 (Figura 44D). Os dados de energia livre da Figura 45A apresentou 3 regiões, identificadas pelos picos, de erros estruturais. Esses erros devem estar relacionados a

modelos de estrutura energeticamente desfavoráveis, uma vez que não foi executado o protocolo de minimização de energia no modelo de referência 7K43-E.

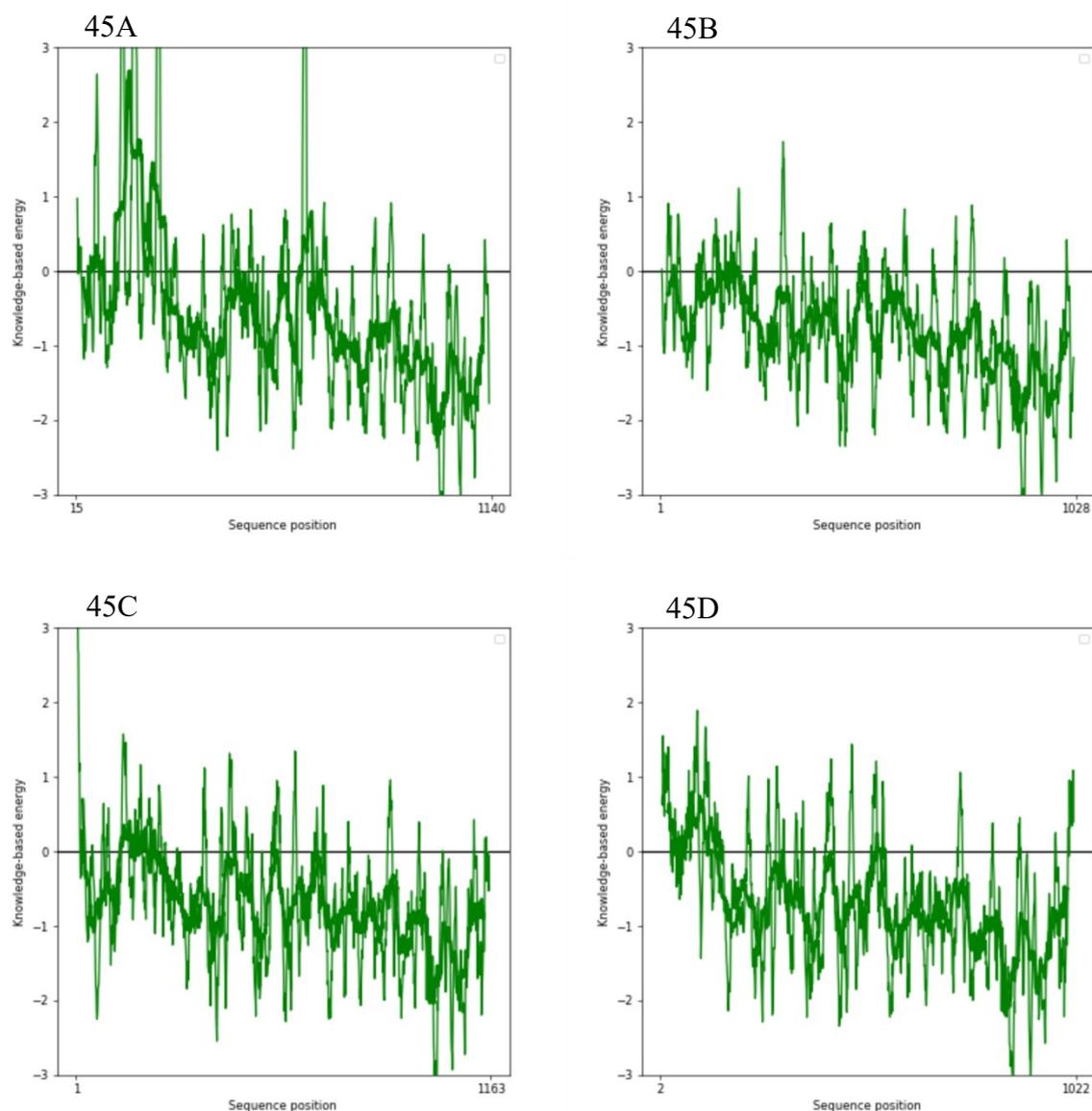


Figura 45. Gráfico de energia livre em seqüências agrupadas de 10 em 10 (verde claro) e 40 em 40 (verde escuro) aminoácidos dos modelos atômicos das proteínas spike do SARS-CoV-2 da referência 7K43-E (45A), da estrutura molecular original de Wuhan do grupo 2 (45B), variante Omicron (45C) e subvariante BA.2 (45D). Valores positivos indicam erros intrínsecos à conformação da região.

Conforme a Figura 46, é demonstrado em forma de gráfico o Z-score obtido e esperado em função do número de resíduos da proteína dos modelos atômicos da isoforma de ECA2.

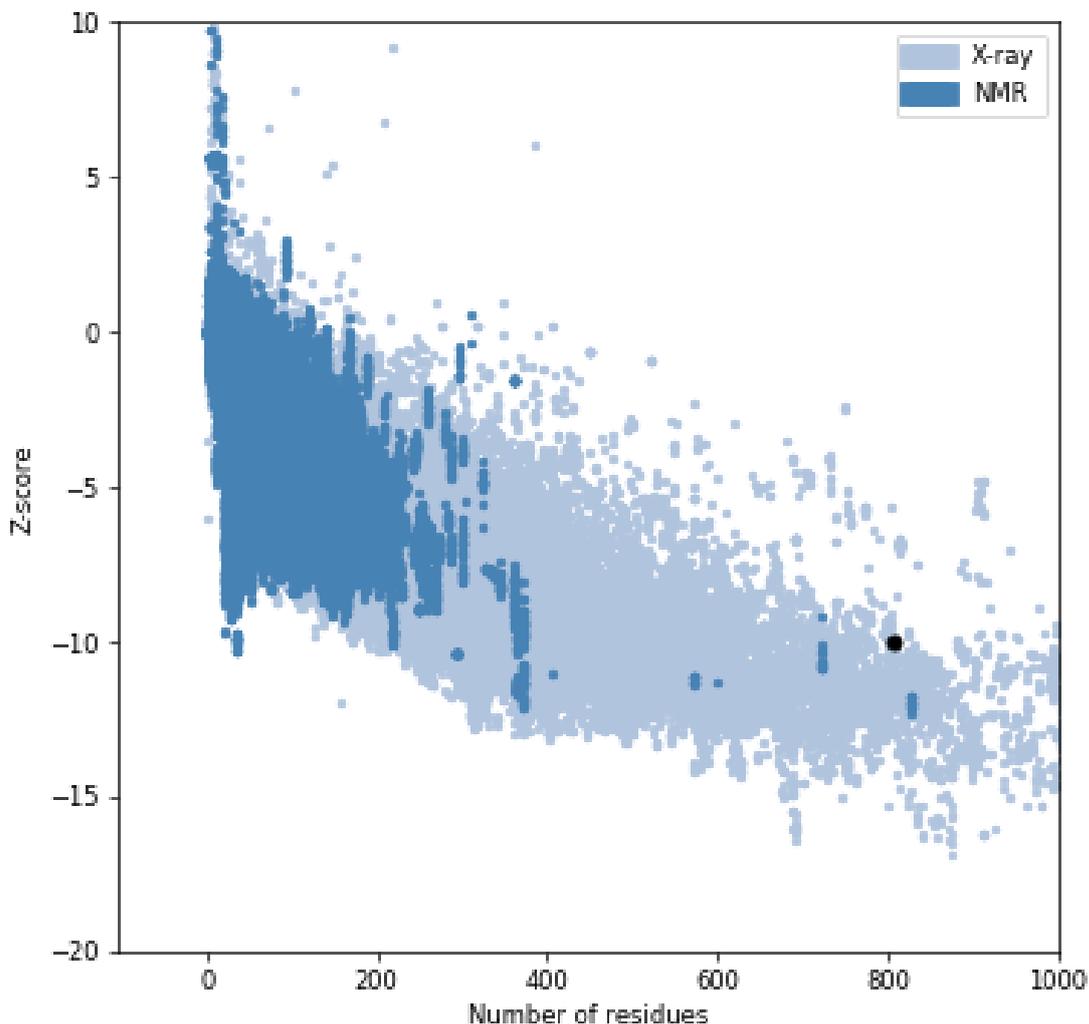


Figura 46. Z-score esperado para modelos atômicos de alta qualidade de cristalografia de raio X e ressonância magnética nuclear (NMR) com base no número de resíduos do modelo atômico da isoforma 1 de ECA2. Regiões coloridas em azul-azul e azul-azul claro indicam uma alta qualidade esperada para o modelo de cristalografia de raio X e NMR respectivamente. O ponto preto aponta o resultado do modelo.

Conforme a Figura 47, é demonstrado os dados de energia livre em janelas de 10 e (verde claro) e 40 (verde escuro) aminoácidos em função do tamanho da molécula do modelo atômico da isoforma 1 de ECA2. Os dados de energia livre da figura apresentaram 2 regiões, identificadas pelos picos, de erros estruturais. Esses erros devem estar relacionados a modelos de estrutura energeticamente desfavoráveis, uma vez que não foi executado o protocolo de minimização de energia no modelo da isoforma 1 de ECA2.

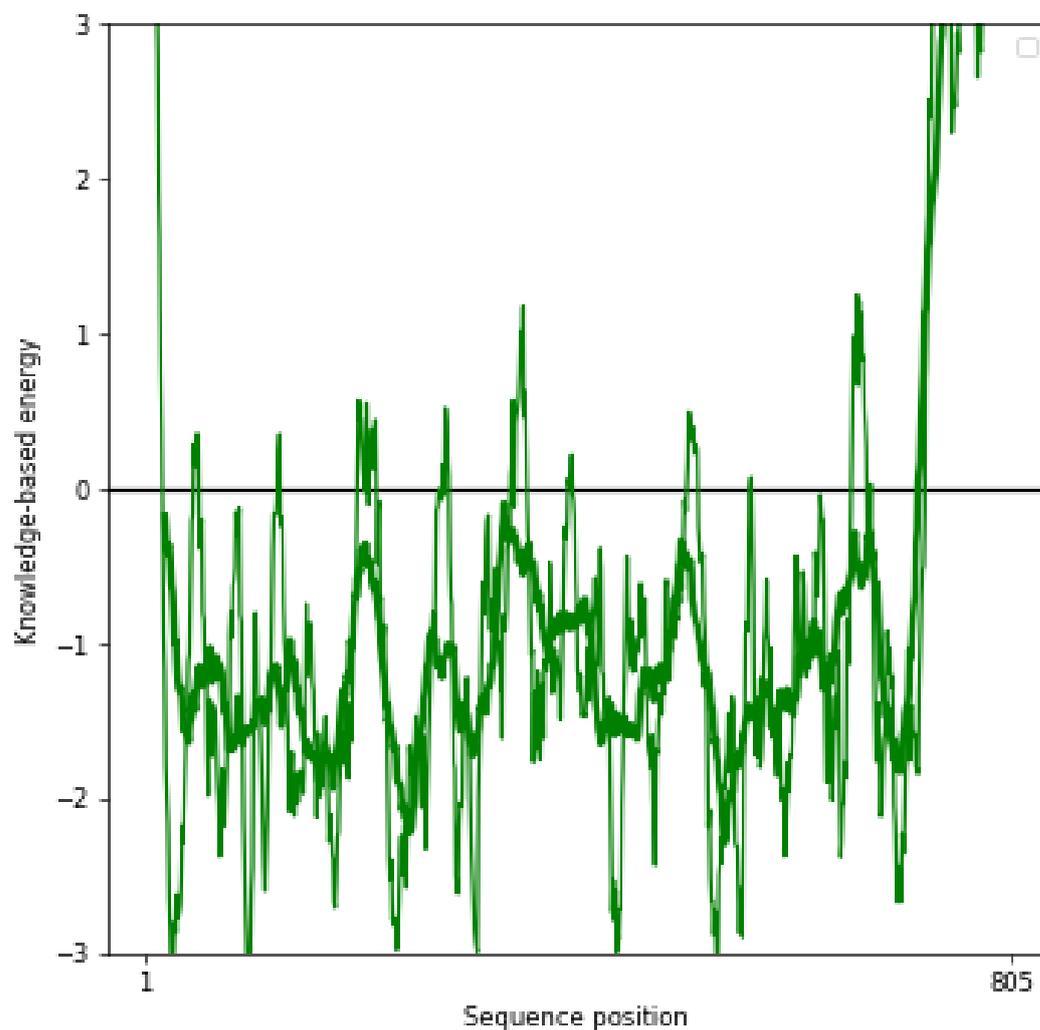


Figura 47. Gráfico de energia livre em sequências agrupadas de 10 em 10 (verde claro) e 40 em 40 (verde escuro) aminoácidos do modelo atômico da isoforma da ECA1. Valores positivos indicam erros intrínsecos à conformação da região.

Como descritos na Tabela 20, os dados de Z-score das proteínas spike das variantes de SARS-CoV-2 de referência 7CAK-A, com mutação pontual E484K, e combinadas K417Y, E484K e N501Y de referência 7K43-E, original de Wuhan e variantes de preocupação Omicron e subvariante BA.2, foram próximos ao Z-score das estruturas de referência, indicando uma qualidade satisfatória dos modelos.

Tabela 20. Valores de Z-score para os modelos atômicos das proteínas spike do SARS-CoV-2 do grupo 1 e 2 e isoforma 1 de ECA2.

Proteína spike SARS-CoV-2	Z-score	Proteína ECA2	Z-score
S1A	-10,56	E1A	-10,02
S1E	-12,24		
S1F	-12,2		
S2A	-10,91		
S2B	-12,2		
S2F	-12,23		
S2G	-10,99		

Tabela 20. Valores de Z-score para os modelos atômicos das proteínas spike do SARS-CoV-2 do grupo 1 e 2 e isoforma 1 de ECA2. Valores positivos ou negativos demonstram variação do padrão da média energética de moléculas de alta qualidade e que possuem o mesmo tamanho e resolução.

Com os modelos atômicos selecionados das proteínas spike do SARS-CoV-2 e isoforma 1 da ECA2, foram realizadas interações intermoleculares.

4.6. Interação molecular entre variantes das proteínas spike de SARS-CoV-2 e isoformas de ECA2

A partir de um arquivo PDB de um complexo com as interações intermoleculares definidas, o software Ligplot+ é capaz de elucidar as interações em um formato intuitivo. Através de sua ferramenta DIMPLOT, o programa é capaz de criar um diagrama 2D das IPPs, onde é possível observar duas interfaces, que se referem a interface de cada cadeia que está participando de interações intermoleculares. Nas interfaces, é possível observar os resíduos que participam de ligações de hidrogênio e interações hidrofóbicas, além de suas cadeias correspondentes. Ainda, para as interações de hidrogênio, é possível coletar as distâncias das ligações em angstroms (WALLACE; LASKOWSKI; THORNTON, 1995).

Por causa do direcionamento aplicado no docking molecular através dos resíduos de atração estabelecidos, os resíduos de interação que foram identificados se limitaram a região do RBD da proteína spike do SARS-CoV-2.

Conforme apresentado na Figura 48, é demonstrado o docking molecular das ligações de hidrogênio entre as proteínas spike de SARS-CoV-2 original de Wuhan (cadeia A) e a

alterar o perfil de interações intermoleculares dos aminoácidos em ambas as proteínas da interface de interação do complexo e484.

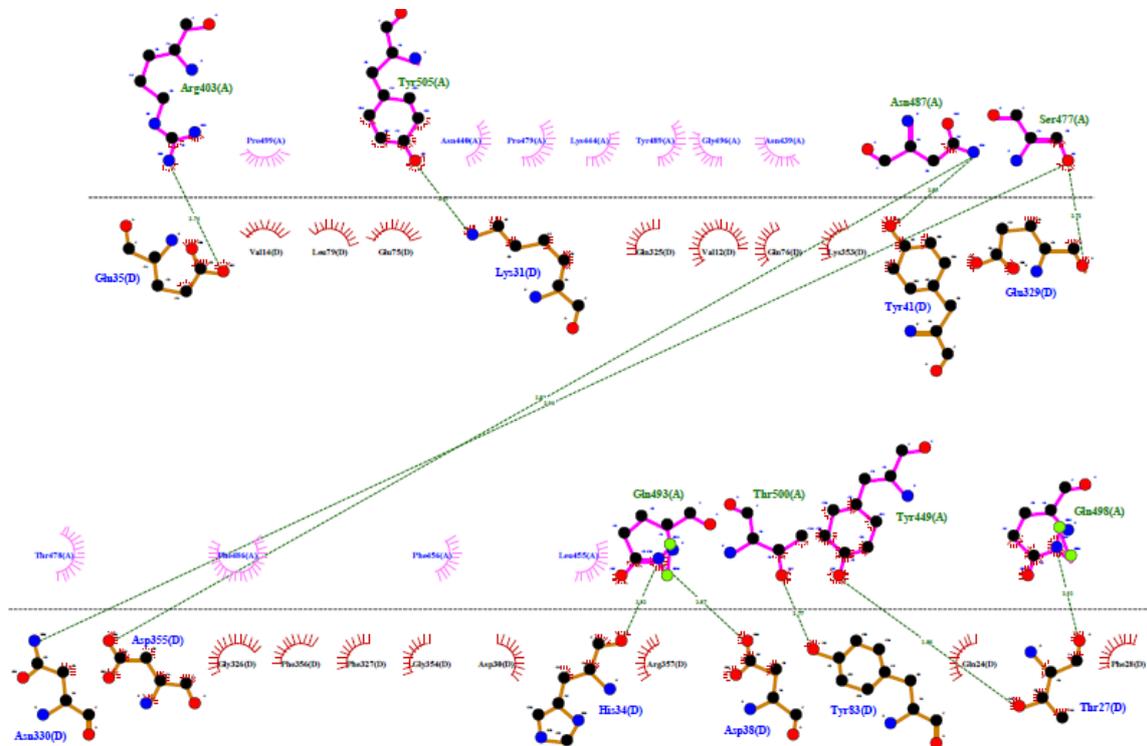


Figura 49. Representação esquemática em 2D do docking molecular entre a proteína spike do SARS-CoV-2 da variante com a mutação pontual E484K do grupo 1 (S1E) e a isoforma 1 de ECA2 (E1A) (Código PDBsum do complexo: e744). Na parte superior em relação ao pontilhado horizontal se encontram os resíduos de interação da cadeia A da proteína S1E e na parte inferior os resíduos de interação da cadeia D da proteína E1A. As linhas tracejadas em verde representam as interações hidrofílicas entre os resíduos, e os demais resíduos representam as interações hidrofóbicas.

Comparando as interfaces de interação formada nos complexos de códigos PDBsum e824 (variante da proteína spike do SARS-CoV-2 com 3 mutações K417Y, E484K e N501Y interagindo com a isoforma 1 de ECA2) e e744 (proteína spike original do SARS-CoV-2 de Wuhan interagindo com a isoforma 1 de ECA2) das respectivas Figura 50 e Figura 48, também são relatadas mudanças no perfil de ligações de hidrogênio. São demonstradas na interface de interação molecular do complexo e484 as ligações de hidrogênio entre os resíduos respectivamente da proteína spike do SARS-CoV-2 e da isoforma de ECA2: Gly496/Asp30, Tyr505/Lys68, Gln498/Tyr83, Tyr449/Asp30, Lys444/Val12 e Lys444/Ala13 que estão ausentes na interface de interação molecular do complexo e744 e compartilhados

entre ambos os complexos. Isso aponta que a combinação de mutações na molécula spike do SARS-CoV-2, K417Y, E484K e N501Y, foi capaz de alterar o perfil de interações intermoleculares dos aminoácidos em ambas as proteínas da interface de interação do complexo e824. Tyr505/Glu75, Ser477/Glu329, Gln493/Asp38, Gln498/Thr27 e Thr500/Tyr83

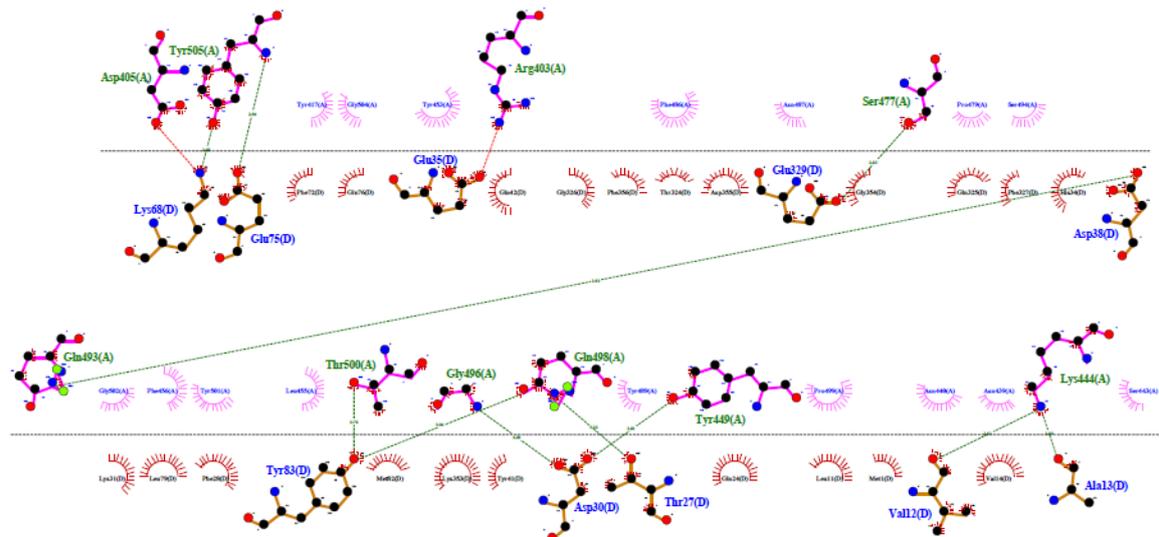


Figura 50. Representação esquemática em 2D do docking molecular entre a proteína spike do SARS-CoV-2 da variante com as mutações pontuais K417Y, E484K e N501Y do grupo 1 (S1F) e a isoforma 1 de ECA2 (E1A) (Código PDBsum do complexo: e824). Na parte superior em relação ao pontilhado horizontal se encontram os resíduos de interação da cadeia A da proteína S1F e na parte inferior os resíduos de interação da cadeia D da proteína E1A. As linhas tracejadas em verde representam as interações hidrofílicas entre os resíduos, e os demais resíduos representam as interações hidrofóbicas.

Comparando as interfaces de interação formada nos complexos de códigos PDBsum e825 (proteína spike do SARS-CoV-2 da variante Omicron interagindo com a isoforma 1 de ECA2) e e744 (proteína spike original do SARS-CoV-2 de Wuhan interagindo com a isoforma 1 de ECA2) das respectivas Figura 51 e Figura 48, também são relatadas mudanças no perfil de ligações de hidrogênio. São demonstradas na interface de interação molecular do complexo e484 as ligações de hidrogênio entre os resíduos respectivamente da proteína spike do SARS-CoV-2 e da isoforma de ECA2: Thr491/Arg355, Arg489/Asn49, Asn478/Tyr83, Tyr444/His34, Tyr440/Gln42, Ser487/Gln42, Ser487/Asp38, Arg394/Asp38, Tyr464/Thr27 e Ala466/Gln24 e Thr491/Arg357 que estão ausentes na interface de interação molecular do complexo e744. Isso aponta que as mutações na molécula spike do SARS-CoV-2 da variante

Omicron foram capazes de alterar completamente o perfil de interações intermoleculares dos aminoácidos em ambas as proteínas da interface de interação do complexo e825.

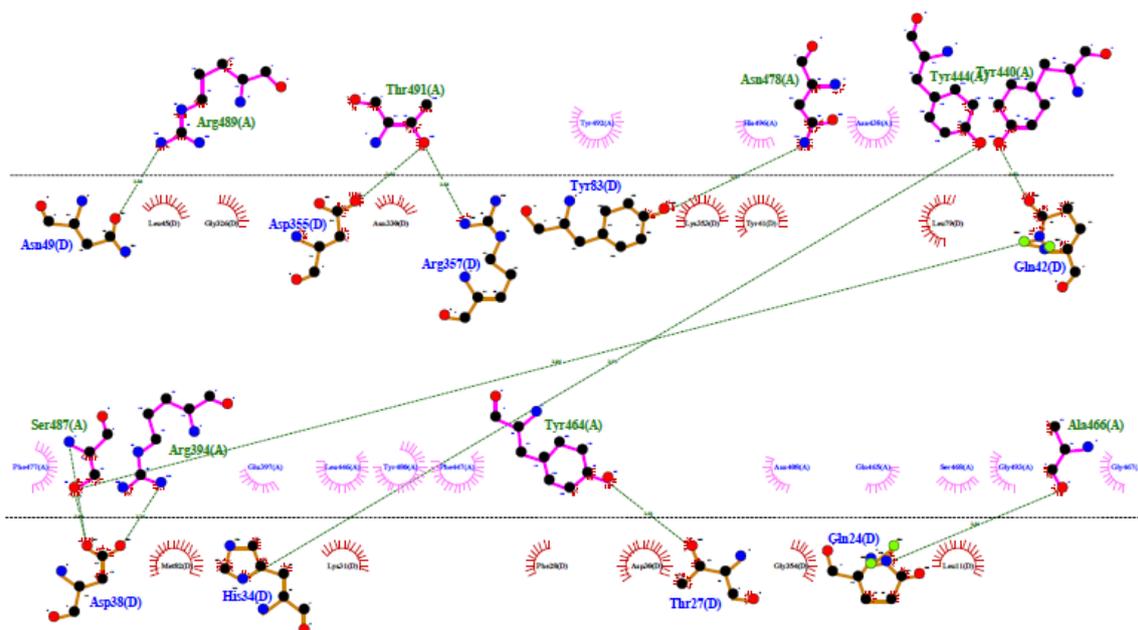


Figura 51. Representação esquemática em 2D do docking molecular entre a proteína spike do SARS-CoV-2 da variante Omicron do grupo 2 (S2F) e a isoforma 1 de ECA2 (E1A) (Código PDBsum do complexo: e825). Na parte superior em relação ao pontilhado horizontal se encontram os resíduos de interação da cadeia A da proteína S2E e na parte inferior os resíduos de interação da cadeia D da proteína E1A. As linhas tracejadas em verde representam as interações hidrofílicas entre os resíduos, e os demais resíduos representam as interações hidrofóbicas.

Comparando as interfaces de interação formada nos complexos de códigos PDBsum e826 (proteína spike do SARS-CoV-2 da subvariante Omicron BA.2 interagindo com a isoforma 1 de ECA2) e e744 (proteína spike original do SARS-CoV-2 de Wuhan interagindo com a isoforma 1 de ECA2) das respectivas Figura 52 e Figura 48, também são relatadas mudanças no perfil de ligações de hidrogênio. São demonstradas na interface de interação molecular do complexo e484 as ligações de hidrogênio entre os resíduos respectivamente da proteína spike do SARS-CoV-2 e da isoforma de ECA2: Tyr349/Tyr41, Thr348/Asp355, Lys288/Glu329, Lys288/Glu329, Tyr301/Asp38, Arg251/Asp38, His353/Lys353, Gly350/Gly354, Gly344/Tyr41, Arg346/Arg355, Asn325/Leu11, Asn265/Asp30 Tyr337/Gln76 e Thr348/Arg357 que estão ausentes na interface de interação molecular do complexo e744. Isso aponta que as mutações na molécula spike do SARS-CoV-2 da variante

Omicron BA.2, foram capazes de alterar completamente e de forma mais intensa o perfil de interações intermoleculares dos aminoácidos em ambas as proteínas da interface de interação do complexo e826.

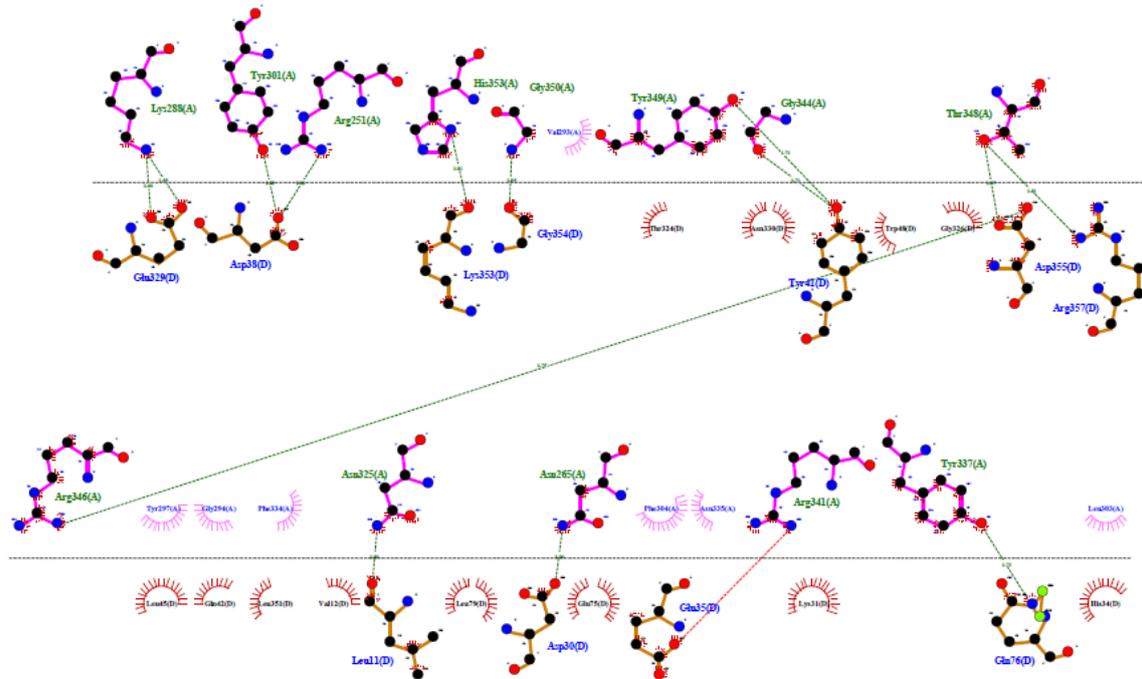


Figura 52. Representação esquemática em 2D do docking molecular entre a proteína spike do SARS-CoV-2 da variante BA.2 do grupo 2 (S2G) e a isoforma 1 de ECA2 (E1A) (Código PDBsum do complexo: e826). Na parte superior em relação ao pontilhado horizontal se encontram os resíduos de interação da cadeia A da proteína S2F e na parte inferior os resíduos de interação da cadeia D da proteína E1A. As linhas tracejadas em verde representam as interações hidrofílicas entre os resíduos, e os demais resíduos representam as interações hidrofóbicas. Devido a proteína spike do SARS-CoV-2 da variante BA.2 modelada estar incompleta, deve-se considerar adicionar 152 aminoácidos, que estão faltantes na extremidade amino-terminal, à qualquer posição de resíduo antes de estabelecer uma comparação direta com as outras interações moleculares de outros complexos.

Conforme mostra a Tabela 21, a posição dos resíduos, originais e adaptados, do modelo da proteína spike de SARS-CoV-2 variante Omicron que interagem por ligações de hidrogênio com a isoforma 1 da ECA2 são demonstrados para um melhor entendimento. Uma vez que há 9 lacunas de aminoácidos na região “up stream” em relação ao RBD desse modelo no alinhamento em comparação à sequência de aminoácidos da proteína spike do SARS-CoV-2 original de Wuhan, os mesmos resíduos de ambas as sequências são representados por posições diferentes.

Tabela 21. Posição de resíduos originais e adaptadas à sequência original de Wuhan que realizam ligações de hidrogênio no docking molecular entre os modelos da proteína spike da variante Omicron do SARS-CoV-2 e isoforma 1 de ECA2

Posição de resíduo no modelo S2F	Posição de resíduo adaptado à sequência original de Wuhan
Arg489	Arg498
Thr491	Thr500
Thr491	Thr500
Asn478	Asn487
Tyr444	Tyr453
Tyr440	Tyr449
Ser487	Ser496
Ser487	Ser496
Ser487	Ser496
Arg394	Arg403
Tyr464	Tyr473
Ala466	Ala475

Tabela 21. Posição de resíduos originais e adaptadas à sequência original de Wuhan que realizam ligações de hidrogênio no docking molecular entre os modelos da proteína spike da variante Omicron do SARS-CoV-2 e isoforma 1 de ECA2. Foram adicionadas 9 unidades a cada posição que corresponde a lacunas de aminoácidos no alinhamento entre a proteína spike da variante Omicron de SARS-CoV-2 e isoforma 1 de ECA2.

Conforme mostra a Tabela 22, a posição dos resíduos, originais e adaptados, do modelo da proteína spike de SARS-CoV-2 subvariante BA.2 que interagem por ligações de hidrogênio com a isoforma 1 da ECA2 são demonstrados para um melhor entendimento. Uma vez que há uma falta de 152 aminoácidos na região amino terminal desse modelo em relação à sequência de aminoácidos da proteína spike do SARS-CoV-2 original de Wuhan, os mesmos resíduos de ambas as sequências são representados por posições diferentes.

Tabela 22. Posição de resíduos originais e adaptadas que realizam ligações de hidrogênio no docking molecular entre os modelos da proteína spike da subvariante BA.2 do SARS-CoV-2 e isoforma 1 de ECA2

Posição de resíduo no modelo S2G	Posição de resíduo adaptado à sequência original de Wuhan
Lys288	Lys440
Lys288	Lys440
Tyr301	Tyr453
Arg251	Arg403
His353	His505
Gly350	Gly502
Tyr349	Tyr501
Gly344	Gly496
Thr348	Thr500
Thr348	Thr500
Arg346	Arg498
Asn325	Asn477
Asn265	Asn417
Tyr337	Tyr489

Tabela 22. Posição de resíduos originais e adaptadas que realizam ligações de hidrogênio no docking molecular entre os modelos da proteína spike da subvariante BA.2 do SARS-CoV-2 e isoforma 1 de ECA2. Foram adicionadas 152 unidades a cada posição que corresponde a lacunas de aminoácidos no alinhamento entre a proteína spike da subvariante Omicron BA.2 de SARS-CoV-2 e isoforma 1 de ECA2.

Com base nos resíduos de interação estabelecidos nas ligações de hidrogênio, é possível notar uma seleção evolutiva positiva de aminoácidos de cargas positivas na proteína spike do SARS-CoV-2 com os resíduos e posições em relação a sequência de Wuhan original Arg403 na referência de Wuhan, Arg498 e Arg403 na variante Omicron e Lys440, His505, Arg403 e Arg498 na subvariante BA.2. Uma vez que a superfície da ECA2 é predominantemente negativa (PRABAKARAN; XIAO; DIMITROV, 2004), a presença dos resíduos descritos para proteína spike nas variantes virais possivelmente conferem maior estabilidade nas interações intermoleculares, em especial em interações eletrostáticas, favorecendo a interação de ECA2 e da proteína spike do SARS-CoV-2. Esse mesmo efeito de aumento da estabilidade foi observado também na interação molecular entre os resíduos Asp38 da ECA2 com Lys353 da proteína NL63-CoV, responsável por mediar a fusão do coronavírus NL63 na célula (WU et al., 2009).

5. CONCLUSÃO

Com base nos dados apresentados neste trabalho monográfico, foi possível concluir que:

- As mutações pontuais A222V e D614G inseridas de forma individual nas sequências dos modelos atômicos de proteína spike do SARS-CoV-2 não promoveram grandes mudanças nos perfis de RMSD (desvio das raízes quadráticas médias – root mean square deviation) quando comparados com a sequência original de Wuhan. Com isso, foi concluído que essas mutações não provocaram tantas mudanças quanto a mutação pontual E484K e a combinação K417Y, E484K e N501Y nas interações de docking molecular com a isoforma 1 de ECA2. Contudo, não foi utilizado um método estatístico para se avaliar a significância das mudanças nos perfis de RMSD entre os modelos.

- A região de polimorfismo das isoformas de ECA2 descritas na literatura só diferem entre si na região carboxi terminal. Além disso, a região de interação com a proteína spike do SARS-CoV-2 que se encontra em uma região diferente é idêntica entre as isoformas. Portanto, as isoformas de ECA2 não devem apresentar diferenças nas interações moleculares com a proteína spike do SARS-CoV-2.

- Quando comparados com a sequência original de Wuhan, as proteínas spike do SARS-CoV-2 das variantes de preocupação Omicron e sua subvariante BA.2 apresentaram valores de RMSD superiores aos valores das outras variantes Alfa, Beta e Delta. Por conseguinte, somente as sequências das proteínas spike do SARS-CoV-2 das variantes Omicron e sua subvariante BA.2 foram consideradas para o docking molecular.

- A qualidade estereoquímica geral das moléculas da proteína spike do SARS-CoV-2 foram semelhantes às suas referências de comparação, indicando que esses modelos apresentaram qualidade satisfatória. Já a estrutura da isoforma 1 de ECA2 não apresentou bons resultados de qualidade estereoquímica e isso pode estar relacionado à não execução do protocolo de minimização de energia desse modelo.

- As redes de interações moleculares das variantes da proteína spike do SARS-CoV-2 revelaram uma seleção evolutiva positiva de resíduos carregados positivamente. Uma vez que a superfície da ECA2 é predominantemente negativa, essa seleção aponta uma tendência evolutiva de interações eletrostáticas mais fortes entre a proteína spike do SARS-CoV-2 e ECA2.

6. REFERÊNCIAS BIBLIOGRÁFICAS

ALEXANDRE, J. et al. Renin-angiotensin-aldosterone system and COVID-19 infection.

Annales d'Endocrinologie, v. 81, n. 2–3, p. 63–67, jun. 2020.

<https://doi.org/10.1016/j.ando.2020.04.005>

ALI, A.; VIJAYAN, R. Dynamics of the ACE2–SARS-CoV-2/SARS-CoV spike protein interface reveal unique mechanisms. **Scientific Reports**, v. 10, n. 1, p. 14214, 26 dez. 2020.

<https://doi.org/10.1038/s41598-020-71188-3>

ALI, F.; KASRY, A.; AMIN, M. The new SARS-CoV-2 strain shows a stronger binding affinity to ACE2 due to N501Y mutant. **Medicine in Drug Discovery**, v. 10, p. 100086, jun. 2021.

<https://doi.org/10.1016/j.medidd.2021.100086>

ALTARAWNEH, H. N. et al. Protection against the Omicron Variant from Previous SARS-CoV-2 Infection. **New England Journal of Medicine**, p. 1–3, 9 fev. 2022.

ARAF, Y. et al. Omicron variant of SARS-CoV-2: Genomics, transmissibility, and responses to current COVID-19 vaccines. **Journal of Medical Virology**, n. December 2021, p. 1–8, 23 jan. 2022.

<https://doi.org/10.1002/jmv.27588>

BAGER, P. et al. Reduced Risk of Hospitalisation Associated With Infection With SARS-CoV-2 Omicron Relative to Delta: A Danish Cohort Study. **Preprints with THE LANCET**,

p. 27, 2022. <http://dx.doi.org/10.2139/ssrn.4008930>

BATEMAN, A. et al. UniProt: the universal protein knowledgebase in 2021. **Nucleic Acids Research**, v. 49, n. D1, p. D480–D489, 8 jan. 2021.

<https://doi.org/10.1093/nar/gkaa1100>

BLUME, C. et al. A novel ACE2 isoform is expressed in human respiratory epithelia and is upregulated in response to interferons and RNA respiratory virus infection. **Nature Genetics**,

v. 53, n. 2, p. 205–214, 11 fev. 2021. <https://doi.org/10.1038/s41588-020-00759-x>

BOWIE, J. U.; LÜTHY, R.; EISENBERG, D. A method to identify protein sequences that

fold into a known three-dimensional structure. **Science**, v. 253, n. 5016, p. 164–170, 1991.

<https://doi.org/10.1126/science.1853201>

BURRELL, L. M. et al. The ACE2 gene: its potential as a functional candidate for cardiovascular disease. **Clinical Science**, v. 124, n. 2, p. 65–76, 1 jan. 2013.

<https://doi.org/10.1042/CS20120269>

CATERINO, M. et al. Protein-protein interaction networks as a new perspective to evaluate distinct functional roles of voltage-dependent anion channel isoforms. **Molecular**

BioSystems, v. 13, n. 12, p. 1–37, 2017. <https://doi.org/10.1039/C7MB00434F>

CELE, S. et al. Escape of SARS-CoV-2 501Y.V2 from neutralization by convalescent plasma. **Nature**, v. 593, n. 7857, p. 142–146, 6 maio 2021. [https://doi.org/10.1038/s41586-021-](https://doi.org/10.1038/s41586-021-03471-w)

[03471-w](https://doi.org/10.1038/s41586-021-03471-w)

CHEN, V. B. et al. MolProbity : all-atom structure validation for macromolecular crystallography. **Acta Crystallographica Section D Biological Crystallography**, v. 66, n. 1, p. 12–21, 1 jan. 2010. <https://doi.org/10.1107/S0907444909042073>

CHIU, S. S. et al. Human coronavirus NL63 infection and other coronavirus infections in children hospitalized with acute respiratory disease in Hong Kong, China. **Clinical Infectious Diseases**, v. 40, n. 12, p. 1721–1729, 2005. <https://doi.org/10.1086/430301>

CHOI, J. Y.; SMITH, D. M. SARS-CoV-2 Variants of Concern. **Yonsei Medical Journal**, v. 62, n. 11, p. 961–968, 2021. <https://doi.org/10.3349/ymj.2021.62.11.961>

CHUANG, G. Y. et al. DARS (Decoys As the Reference State) potentials for protein-protein docking. **Biophysical Journal**, v. 95, n. 9, p. 4217–4227, 2008.

<https://doi.org/10.1529/biophysj.108.135814>

DAVIES, N. G. et al. Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. **Science**, v. 372, n. 6538, p. 0–10, 9 abr. 2021.

<https://doi.org/10.1126/science.abg30>

DE LAS RIVAS, J.; FONTANILLO, C. Protein-protein interactions essentials: Key concepts

to building and analyzing interactome networks. **PLoS Computational Biology**, v. 6, n. 6, p. 1–8, 2010. <https://doi.org/10.1371/journal.pcbi.1000807>

DESINGU, P. A.; NAGARAJAN, K.; DHAMA, K. Emergence of Omicron third lineage BA.3 and its importance. **Journal of Medical Virology**, n. November 2021, p. 2–4, 23 jan. 2022. <https://doi.org/10.1002/jmv.27601>

DHAWAN, M.; PRIYANKA; CHOUDHARY, O. P. Emergence of Omicron sub-variant BA.2: Is it a matter of concern amid the COVID-19 pandemic? **International Journal of Surgery**, v. 99, n. January, p. 1–3, mar. 2022. <https://doi.org/10.1016/j.ijssu.2022.106581>
Discovery Studio Visualizer 17.2.0.16349. , [s.d.].

EDGAR, R. C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. **Nucleic Acids Research**, v. 32, n. 5, p. 1792–1797, 2004.
<https://doi.org/10.1093/nar/gkh340>

EUROPEAN CENTRE FOR DISEASE PREVENTION AND CONTROL. **Assessment of the further spread and potential impact of the SARS-CoV-2 Omicron variant of concern in the EU / EEA , 19th update Risk assessed ECDC: Stockholm**. [s.l: s.n.].

FAGERBERG, L. et al. Analysis of the Human Tissue-specific Expression by Genome-wide Integration of Transcriptomics and Antibody-based Proteomics. **Molecular & Cellular Proteomics**, v. 13, n. 2, p. 397–406, fev. 2014. <https://doi.org/10.1074/mcp.M113.035600>

FERGUSON, N. et al. **Report 50: Hospitalisation risk for Omicron cases in England** Imperial College Report. [s.l: s.n.]. Disponível em:

<<https://www.imperial.ac.uk/media/imperial-college/medicine/mrc-gida/2021-12-22-COVID19-Report-50.pdf>>.

FIELDS, B. N.; KNIPE, D. M.; HOWLEY, P. M. **Coronaviridae**. **Fields Virology**. Philadelphia: [s.n.].

FUNK, T. et al. Characteristics of SARS-CoV-2 variants of concern B.1.1.7, B.1.351 or P.1:

data from seven EU/EEA countries, weeks 38/2020 to 10/2021. **Eurosurveillance**, v. 26, n. 16, p. 1–10, 22 abr. 2021. <https://doi.org/10.2807/1560-7917.ES.2021.26.16.2100348>

GARCIA-BELTRAN, W. F. et al. Multiple SARS-CoV-2 variants escape neutralization by vaccine-induced humoral immunity. **Cell**, v. 184, n. 1–9, p. 2372–2383, abr. 2021. <https://doi.org/10.1016/j.cell.2021.03.013>

GORBALENYA, A. E. et al. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. **Nature Microbiology**, v. 5, p. 536–544, 2020. <https://doi.org/10.1038/s41564-020-0695-z>

GORSE, G. J. et al. Human coronavirus and acute respiratory illness in older adults with chronic obstructive pulmonary disease. **Journal of Infectious Diseases**, v. 199, n. 6, p. 847–857, 2009. <https://doi.org/10.1086/597122>

GREANEY, A. J. et al. Comprehensive mapping of mutations in the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human plasma antibodies. **Cell Host & Microbe**, v. 29, n. 3, p. 463–476.e6, mar. 2021. <https://doi.org/10.1016/j.chom.2021.02.003>

GU, H. et al. Adaptation of SARS-CoV-2 in BALB/c mice for testing vaccine efficacy. **Science**, v. 369, n. 6511, p. 1603–1607, 25 set. 2020. <https://doi.org/10.1126/science.abc4730>

HADFIELD, J. et al. Nextstrain: real-time tracking of pathogen evolution. **Bioinformatics**, v. 34, n. 23, p. 4121–4123, 1 dez. 2018. <https://doi.org/10.1093/bioinformatics/bty407>

HOFFMANN, M. et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. **Cell**, v. 181, n. 2, p. 271–280, 2020.

International Committee on Taxonomy of Viruses. Disponível em:

<<https://talk.ictvonline.org/>>. Acesso em: 1 nov. 2021.

<https://doi.org/10.1016/j.cell.2020.02.052>

JACOB, J. J. et al. Evolutionary Tracking of SARS-CoV-2 Genetic Variants Highlights an Intricate Balance of Stabilizing and Destabilizing Mutations. **mBio**, v. 12, n. 4, p. 1–11, 2021.

JACOBSEN, H. et al. **Results of Studies Evaluating the Impact of SARS- CoV-2 Variants**

of Concern on COVID-19 Vaccines : An Ongoing Systematic Review Neutralization

Plots. [s.l: s.n.]. <https://doi.org/10.1128/mBio.01188-21>

JANGRA, S. et al. SARS-CoV-2 spike E484K mutation reduces antibody neutralisation. **The Lancet Microbe**, v. 2, n. 7, p. e283–e284, 2021.

JEAN, A. et al. Severity and outcome associated with human coronavirus OC43 infections among children. **Pediatric Infectious Disease Journal**, v. 32, n. 4, p. 325–329, 2013.

[https://doi.org/10.1016/S2666-5247\(21\)00068-9](https://doi.org/10.1016/S2666-5247(21)00068-9)

JEVŠNIK, M. et al. Coronavirus infections in hospitalized pediatric patients with acute respiratory tract disease. **BMC Infectious Diseases**, v. 12, n. 1, p. 1, 2012.

<https://doi.org/10.1186/1471-2334-12-365>

KANEHISA, M. KEGG: Kyoto Encyclopedia of Genes and Genomes. **Nucleic Acids Research**, v. 28, n. 1, p. 27–30, 1 jan. 2000. <https://doi.org/10.1093/nar/28.1.27>

KANEHISA, M. Toward understanding the origin and evolution of cellular organisms.

Protein Science, v. 28, n. 11, p. 1947–1951, 9 nov. 2019. <https://doi.org/10.1002/pro.3715>

KANEHISA, M. et al. KEGG: integrating viruses and cellular organisms. **Nucleic Acids Research**, v. 49, n. D1, p. D545–D551, 8 jan. 2021. <https://doi.org/10.1093/nar/gkaa970>

KORBER, B. et al. Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. **Cell**, v. 182, n. 4, p. 812–827, 2020.

<https://doi.org/10.1016/j.cell.2020.06.043>

KOZAKOV, D. et al. PIPER: An FFT-Based Protein Docking Program with Pairwise Potentials. **Proteins: Structure, Function and Bioinformatics**, n. 65, p. 392–406, 2006.

<https://doi.org/10.1002/prot.21117>

KOZAKOV, D. et al. The ClusPro web server for protein-protein docking. **Nature Protocols**, v. 12, n. 2, p. 255–278, 2017. <https://doi.org/10.1038/nprot.2016.169>

KRIEGER, E.; VRIEND, G. YASARA View - molecular graphics for all devices - from smartphones to workstations. **Bioinformatics (Oxford, England)**, v. 30, n. 20, p. 2981–2982,

2014. <https://doi.org/10.1093/bioinformatics/btu426>

LAN, J. et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. **Nature**, v. 581, n. 7807, p. 215–220, 14 maio 2020.

<https://doi.org/10.1038/s41586-020-2180-5>

LASKOWSKI, R. A. et al. PROCHECK: a program to check the stereochemical quality of protein structures. **Journal of Applied Crystallography**, v. 26, p. 283–291, 1993.

<https://doi.org/10.1107/S0021889892009944>

LASKOWSKI, R. A. et al. AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. **Journal of Biomolecular NMR**, v. 8, n. 4, p. 477–486, 1996. <https://doi.org/10.1007/BF00228148>

LIU, X. et al. Detection of protein complexes from multiple protein interaction networks using graph embedding. **Artificial Intelligence in Medicine**, v. 96, p. 107–115, 2019.

<https://doi.org/10.1016/j.artmed.2019.04.001>

Lucidchart. Disponível em: <<https://www.lucidchart.com/>>. Acesso em: 22 mar. 2022.

LUTHY, R.; BOWIE, J. U.; EISENBERG, D. Assessment of protein models with three-dimensional profiles. **Nature**, v. 359, p. 83–85, 1992.

<https://doi.org/10.1107/S0021889892009944>

LV, Z. et al. Structural basis for neutralization of SARS-CoV-2 and SARS-CoV by a potent therapeutic antibody. **Science**, v. 369, n. 6510, p. 1505–1509, 18 set. 2020.

<https://doi.org/10.1007/BF00228148>

LYNGSE, F. P. et al. Transmission of SARS-CoV-2 Omicron VOC subvariants BA.1 and BA.2: Evidence from Danish Households. **medRxiv**, p. 1–38, 2022.

<https://doi.org/10.1016/j.artmed.2019.04.001>

MADHI, S. A. et al. Efficacy of the ChAdOx1 nCoV-19 Covid-19 Vaccine against the B.1.351 Variant. **New England Journal of Medicine**, v. 384, n. 20, p. 1885–1898, 20 maio 2021. <https://doi.org/10.1056/NEJMoa2102214>

MANNAR, D. et al. SARS-CoV-2 Omicron variant: Antibody evasion and cryo-EM structure of spike protein–ACE2 complex. **Science**, v. 375, n. 6582, p. 760–764, 2022.

<https://doi.org/10.1126/science.abn7760>

ONABAJO, O. O. et al. Interferons and viruses induce a novel truncated ACE2 isoform and not the full-length SARS-CoV-2 receptor. **Nature Genetics**, v. 52, n. 12, p. 1283–1293, 19 dez. 2020. <https://doi.org/10.1038/s41588-020-00731-9>

PATEL, S. K. et al. From gene to protein—experimental and clinical studies of ACE2 in blood pressure control and arterial hypertension. **Frontiers in Physiology**, p. 1–12, 24 jun. 2014. <https://doi.org/10.3389/fphys.2014.00227>

PEARSON, C. A. et al. Estimates of severity and transmissibility of novel South Africa SARS-CoV-2 variant 501Y.V2. [s.d.].

PERLMAN, S.; NETLAND, J. Coronaviruses post-SARS: Update on replication and pathogenesis. **Nature Reviews Microbiology**, v. 7, n. 6, p. 439–450, 2009.

<https://doi.org/10.1038/nrmicro2147>

PETTERSEN, E. F. et al. UCSF Chimera - A visualization system for exploratory research and analysis. **Journal of Computational Chemistry**, v. 25, n. 13, p. 1605–1612, 2004.

<https://doi.org/10.1002/jcc.20084>

PINHEIRO, D. S. et al. The combination of ACE I/D and ACE2 G8790A polymorphisms reveals susceptibility to hypertension: A genetic association study in Brazilian patients. **PLOS ONE**, v. 14, n. 8, p. 1–15, 20 ago. 2019. <https://doi.org/10.1371/journal.pone.0221248>

PLANTE, J. A. et al. Spike mutation D614G alters SARS-CoV-2 fitness. **Nature**, v. 592, n. 7852, p. 116–121, 1 abr. 2021. <https://doi.org/10.1038/s41586-020-2895-3>

PRABAKARAN, P.; XIAO, X.; DIMITROV, D. S. A model of the ACE2 structure and function as a SARS-CoV receptor. **Biochemical and Biophysical Research**

Communications, v. 314, n. 1, p. 235–241, 2004. <https://doi.org/10.1016/j.bbrc.2003.12.081>

PUBLIC HEALTH ENGLAND. **SARS-CoV-2 variants of concern and variants under**

investigation in England Technical briefing 16. [s.l: s.n.].

PUBLIC HEALTH ENGLAND. **SARS-CoV-2 variants of concern and variants under investigation in England Technical briefing 12.** [s.l: s.n.].

PUBLIC HEALTH ENGLAND. **SARS-CoV-2 variants of concern and variants under investigation in England Technical briefing: Update on hospitalisation and vaccine effectiveness for Omicron VOC-21NOV-01 (B.1.1.529).** [s.l: s.n.].

PULLIAM, J. R. C. et al. Increased risk of SARS-CoV-2 reinfection associated with emergence of the Omicron variant in South Africa. **medRxiv**, p. 1–43, 2021.

<https://doi.org/10.1101/2021.11.11.21266068>

RAMACHANDRAN, G. N.; RAMAKRISHNAN, C.; SASISEKHARAN, V.

Stereochemistry of polypeptide chain configurations. **Journal of Molecular Biology**, v. 7, n. 1, p. 95–99, 1963. [https://doi.org/10.1016/S0022-2836\(63\)80023-6](https://doi.org/10.1016/S0022-2836(63)80023-6)

RAMANATHAN, K. et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. **The Lancet**, v. 395, n. January, p. 497–506, 2020.

[https://doi.org/10.1016/S0140-6736\(20\)30183-5](https://doi.org/10.1016/S0140-6736(20)30183-5)

RITCHIE, D. Recent Progress and Future Directions in Protein-Protein Docking. **Current Protein & Peptide Science**, v. 9, n. 1, p. 1–15, 2008.

<https://doi.org/10.2174/138920308783565741>

RÖSSLER, A. et al. Neutralization profile of Omicron variant convalescent individuals.

medRxiv, 2022. <https://doi.org/10.1101/2022.02.01.22270263>

ROY, A.; KUCUKURAL, A.; ZHANG, Y. I-TASSER: A unified platform for automated protein structure and function prediction. **Nature Protocols**, v. 5, n. 4, p. 725–738, 2010.

<https://doi.org/10.1038/nprot.2010.5>

SAGULENKO, P.; PULLER, V.; NEHER, R. A. TreeTime: Maximum-likelihood phylodynamic analysis. **Virus Evolution**, v. 4, n. 1, p. 1–9, 1 jan. 2018.

<https://doi.org/10.1093/ve/vex042>

SHANNON, P. et al. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. **Genome Research**, v. 13, n. 11, p. 2498–2504, nov. 2003. <https://doi.org/10.1101/gr.1239303>

SHEIKH, A. et al. SARS-CoV-2 Delta VOC in Scotland: demographics, risk of hospital admission, and vaccine effectiveness. **The Lancet**, v. 397, n. 10293, p. 2461–2462, 2021a. [https://doi.org/10.1016/S0140-6736\(21\)01358-1](https://doi.org/10.1016/S0140-6736(21)01358-1)

SHEIKH, A. et al. Edinburgh Research Explorer Severity of Omicron variant of concern and vaccine effectiveness against symptomatic disease: national cohort with nested test negative design study in Scotland *****PRE-PRINT-NOT PEER REVIEWED*****. **The University of Edinburgh - Preprint**, 2021b.

SIPPL, M. J. Calculation of conformational ensembles from potentials of mean force. **Journal of Molecular Biology**, v. 213, n. 4, p. 859–883, jun. 1990. [https://doi.org/10.1016/S0022-2836\(05\)80269-4](https://doi.org/10.1016/S0022-2836(05)80269-4)

SIPPL, M. J. Recognition of errors in three-dimensional structures of proteins. **Proteins: Structure, Function, and Bioinformatics**, v. 17, n. 4, p. 355–362, 1993. <https://doi.org/10.1002/prot.340170404>

SIPPL, M. J. Knowledge-based potentials for proteins. **Current Opinion in Structural Biology**, v. 5, n. 2, p. 229–235, abr. 1995. [https://doi.org/10.1016/0959-440X\(95\)80081-6](https://doi.org/10.1016/0959-440X(95)80081-6)

SMITH, G. R.; STERNBERG, M. J. E. Prediction of protein-protein interactions by docking methods. **Current Opinion in Structural Biology**, v. 12, n. 1, p. 28–35, 2002. [https://doi.org/10.1016/S0959-440X\(02\)00285-3](https://doi.org/10.1016/S0959-440X(02)00285-3)

STATENS SERUM INSTITUT (SSI). Now, an Omicron variant, BA.2, accounts for almost half of all Danish Omicron-cases. **Statens Serum Institut (SSI)**, p. 3868, 2022.

STOWE, J. et al. Effectiveness of COVID-19 vaccines against hospital admission with the Delta (B.1.617.2) variant. **Public Health England**, v. 37, n. 1, p. <https://khub.net/web/phe-national/public-library>, 2021.

SURYAWANSHI, R. K. et al. **Limited Cross-Variant Immunity after Infection with the SARS-CoV-2 Omicron Variant Without Vaccination.** [s.l: s.n.].

<https://doi.org/10.1038/s41586-022-04865-0>

TEGALLY, H. et al. Detection of a SARS-CoV-2 variant of concern in South Africa. **Nature**, v. 592, n. 7854, p. 438–443, 15 abr. 2021. <https://doi.org/10.1038/s41586-021-03402-9>

TIPNIS, S. R. et al. A human homolog of angiotensin-converting enzyme: Cloning and functional expression as a captopril-insensitive carboxypeptidase. **Journal of Biological Chemistry**, v. 275, n. 43, p. 33238–33243, 2000a. <https://doi.org/10.1074/jbc.M002615200>

TIPNIS, S. R. et al. A Human Homolog of Angiotensin-converting Enzyme. **Journal of Biological Chemistry**, v. 275, n. 43, p. 33238–33243, out. 2000b.

<https://doi.org/10.1074/jbc.M002615200>

UK HEALTH SECURITY AGENCY (UKHSA). **SARS-CoV-2 variants of concern and variants under investigation in England Technical briefing 35.** [s.l: s.n.].

UK HEALTH SECURITY AGENCY (UKHSA). **COVID-19 vaccine surveillance report.**

[s.l: s.n.]. Disponível em:

<https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1054071/vaccine-surveillance-report-week-6.pdf>.

V'KOVSKI, P. et al. Coronavirus biology and replication: implications for SARS-CoV-2.

Nature Reviews Microbiology, v. 19, n. 3, p. 155–170, 28 mar. 2021.

<https://doi.org/10.1038/s41579-020-00468-6>

VOLZ, E. et al. Evaluating the Effects of SARS-CoV-2 Spike Mutation D614G on Transmissibility and Pathogenicity. **Cell**, v. 184, n. 1, p. 64–75, 2021.

<https://doi.org/10.1016/j.cell.2020.11.020>

WALLACE, A. C.; LASKOWSKI, R. A.; THORNTON, J. M. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. **Protein engineering**, v. 8, n. 2, p.

127–34, fev. 1995. <https://doi.org/10.1093/protein/8.2.127>

WALLS, A. C. et al. Structure, function and antigenicity of the SARS-CoV-2 spike glycoprotein. **A Cell Press**, v. 181, n. 2, p. P281- 292.E6, 2020.

<https://doi.org/10.1016/j.cell.2020.02.058>

WANG, L.; CHENG, G. Sequence analysis of the emerging SARS-CoV-2 variant Omicron in South Africa. **Journal of Medical Virology**, v. 94, n. 4, p. 1728–1733, 27 abr. 2022.

<https://doi.org/10.1002/jmv.27516>

WANG, P. et al. Antibody resistance of SARS-CoV-2 variants B.1.351 and B.1.1.7. **Nature**, v. 593, n. 7857, p. 130–135, 2021a. <https://doi.org/10.1038/s41586-021-03398-2>

WANG, Z. et al. mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating variants. **Nature**, v. 592, n. 7855, p. 616–622, 22 abr. 2021b.

<https://doi.org/10.3410/f.739524179.793585051>

WATERHOUSE, A. et al. SWISS-MODEL: Homology modelling of protein structures and complexes. **Nucleic Acids Research**, v. 46, n. W1, p. W296–W303, 2018.

<https://doi.org/10.1093/nar/gky427>

WEISBLUM, Y. et al. Escape from neutralizing antibodies by SARS-CoV-2 spike protein variants. **eLife**, v. 9, p. 1, 28 out. 2020. <https://doi.org/10.7554/eLife.61312>

WHO. **COVID-19 weekly epidemiological update**World Health Organization. [s.l: s.n.].

Disponível em: <<https://www.who.int/publications/m/item/covid-19-weekly-epidemiological-update>>.

WHO. **Weekly epidemiological update on COVID-19 - 8 February 2022**Weekly

epidemiological update on COVID-19 Edition 78. [s.l: s.n.]. Disponível em:

<<https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19---8-february-2022>>.

WIBMER, C. K. et al. SARS-CoV-2 501Y.V2 escapes neutralization by South African COVID-19 donor plasma. **Nature Medicine**, v. 27, n. 4, p. 622–625, 2 abr. 2021.

<https://doi.org/10.1038/s41591-021-01285-x>

WIEDERSTEIN, M.; SIPPL, M. J. ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. **Nucleic Acids Research**, v. 35, n. SUPPL.2, p. W407–W410, 2007. <https://doi.org/10.1093/nar/gkm290>

WOO, P. C. Y. et al. Comparative Analysis of Complete Genome Sequences of Three Avian Coronaviruses Reveals a Novel Group 3c Coronavirus. **Journal of Virology**, v. 83, n. 2, p. 908–917, 2009. <https://doi.org/10.1128/JVI.01977-08>

WORD, J. M. et al. Asparagine and glutamine: Using hydrogen atom contacts in the choice of side-chain amide orientation. **Journal of Molecular Biology**, v. 285, n. 4, p. 1735–1747, 1999a. <https://doi.org/10.1006/jmbi.1998.2401>

WORD, J. M. et al. Visualizing and quantifying molecular goodness-of-fit: Small-probe contact dots with explicit hydrogen atoms. **Journal of Molecular Biology**, v. 285, n. 4, p. 1711–1733, 1999b. <https://doi.org/10.1006/jmbi.1998.2400>

WRAPP, D. et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. **Science**, v. 367, n. 6483, p. 1260–1263, 2020. <https://doi.org/10.1126/science.abb2507>

WROBEL, A. G. et al. Evolution of the SARS-CoV-2 spike protein in the human host. **Nature Communications**, v. 13, n. 1, p. 1–7, 4 dez. 2022. <https://doi.org/10.1038/s41467-022-28768-w>

WU, K. et al. Crystal structure of NL63 respiratory coronavirus receptor-binding domain complexed with its human receptor. **Proceedings of the National Academy of Sciences**, v. 106, n. 47, p. 19970–19974, 24 nov. 2009. <https://doi.org/10.1073/pnas.0908837106>

WU, K. et al. mRNA-1273 vaccine induces neutralizing antibodies against spike mutants from global SARS-CoV-2 variants. **Angewandte Chemie International Edition**, 6(11), 951–952., 2021a. <https://doi.org/10.1101/2021.01.25.427948>

WU, K. et al. Serum Neutralizing Activity Elicited by mRNA-1273 Vaccine. **New England Journal of Medicine**, v. 384, n. 15, p. 1468–1470, 15 abr. 2021b. <https://doi.org/10.1056/NEJMc2102179>

XIE, X. et al. Neutralization of SARS-CoV-2 spike 69/70 deletion, E484K and N501Y variants by BNT162b2 vaccine-elicited sera. **Nature Medicine**, v. 27, n. 4, p. 620–621, 8 abr. 2021. <https://doi.org/10.1038/s41591-021-01270-4>

YANG, Q. et al. Structural Analysis of the SARS-CoV-2 Omicron Variant Proteins. **Research**, v. 2021, p. 1–4, 28 dez. 2021. <https://doi.org/10.34133/2021/9769586>