

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE BIOTECNOLOGIA
CURSO DE BIOTECNOLOGIA

Análise computacional da interação entre fragmentos variáveis de cadeia única (scFv) de anticorpos modificados com a proteína do envelope do Zika vírus.

Alice Pereira Barros

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

Uberlândia – MG

Janeiro – 2023

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE BIOTECNOLOGIA
CURSO DE BIOTECNOLOGIA

Análise computacional da interação entre fragmentos variáveis de cadeia única (scFv) de anticorpos modificados com a proteína do envelope do Zika vírus.

Alice Pereira Barros

Jair Pereira da Cunha Júnior

Carlos Roberto Prudêncio

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

Uberlândia – MG

Janeiro – 2023

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE BIOTECNOLOGIA
CURSO DE BIOTECNOLOGIA

Análise computacional da interação entre fragmentos variáveis de cadeia única (scFv) de anticorpos modificados com a proteína do envelope (E) do Zika vírus.

Alice Pereira Barros

Jair Pereira da Cunha Júnior
Instituto de Ciências Biomédicas

Carlos Roberto Prudêncio
Instituto Adolfo Lutz

Homologado pela coordenação do curso de
Biotecnologia em __/__/__

Nilson Nicolau Júnior

Uberlândia – MG

Janeiro – 2023

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE BIOTECNOLOGIA
CURSO DE BIOTECNOLOGIA

Análise computacional da interação entre fragmentos variáveis de cadeia única (scFv) de anticorpos modificados com a proteína do envelope do Zika vírus.

Alice Pereira Barros

Aprovado pela Banca Examinadora em: __/__/__ Nota: __

Nome e assinatura do Presidente da Banca Examinadora

Uberlândia, 26 de janeiro de 2023

AGRADECIMENTOS

Agradeço ao meu orientador Dr. Jair Pereira da Cunha Júnior pelo suporte, apoio científico e técnico e pela confiança.

Agradeço ao meu co-orientador Dr. Carlos Roberto Prudêncio e ao Me. Hernan Hermes Monteiro da Costa por ter disponibilizado todos os materiais necessários para o meu aprendizado e execução desta monografia.

Agradeço a banca, Dr. Nilson Nicolau Júnior e Me. Hernan Hermes Monteiro da Costa, por aceitarem o convite e pelas contribuições realizadas.

Agradeço ao Victor Hugo Oliveira de Andrade pelo auxílio e colaboração durante a realização desta monografia.

Agradeço a minha família e amigos que me apoiaram e incentivaram durante todo o meu percurso.

E agradeço a Universidade Federal de Uberlândia (UFU) por disponibilizar os recursos necessários para a execução deste projeto, assim como agradeço a todos os professores pela dedicação e pelos conhecimentos compartilhados.

RESUMO

O vírus Zika (ZIKV) pertence ao gênero *Flavivirus* e possui genoma de aproximadamente 11.000 pares de bases, que codifica três proteínas estruturais - capsídeo (C), membrana (M) e envelope (E), e sete não-estruturais - NS1, NS2A, NS2B, NS3, NS4A, NS4B, NS5. A proteína E é responsável pela entrada do vírus no hospedeiro, sendo assim o principal alvo de anticorpos. Dessa maneira, a partir da bioinformática, é possível obter informações sobre as interações dos complexos ligante-alvo, a fim de descobrir novos inibidores contra o ZIKV. Sendo assim, foi analisada a interação entre a proteína E do ZIKV e fragmentos variáveis de cadeia única (scFv). Em vista disso, foram realizadas modificações no linker peptídico entre as cadeias variáveis dos anticorpos, mas que se mostraram ineficaz em promover mudanças conformacionais nos modelos atômicos. E embora as interações moleculares se limitaram aos domínios I e III da proteína E, existem algumas diferenças, de modo que são necessárias mais pesquisas para compreender se houve relação com as modificações realizadas.

Palavras-chave: Zika Vírus, Fragmentos Variáveis de Cadeia Única, Modelagem Molecular, Interações Moleculares.

SUMÁRIO

1. INTRODUÇÃO	1
2. OBJETIVOS	7
2.1 Objetivo geral	7
2.2 Objetivos específicos	7
3. METODOLOGIA	8
3.1 Modificações realizadas na sequência do scFv anti-ZIKV	8
3.2 Análise das regiões determinantes da complementariedade dos scFv 2 e 4.....	8
3.3 Determinação das regiões hipervariáveis do scFv 2 e 4	9
3.4 Modelagem molecular tridimensional (3D) dos scFv original e modificados	9
3.5 Cálculo do desvio quadrático médio (RMSD) dos alinhamentos dos fragmentos de anticorpos.....	10
3.6 Avaliação da qualidade dos modelos gerados	10
3.7 Interação molecular (<i>docking</i>) entre a proteína do envelope (E) do ZIKV	11
3.8 Análise 2D das interações intermoleculares entre a proteína E do ZIKV com o scFv...	12
4. RESULTADOS E DISCUSSÃO	13
4.1 Determinação das regiões hipervariáveis das sequências de scFv 2 e 4	13
4.2 Modificações no <i>linker</i> do scFv.....	15

4.3 Modelagem dos scFv 2 e quiméricos com <i>linker</i> original e modificado.....	17
4.4 Alinhamento estrutural dos scFv 2 e quiméricos e cálculo do RMSD.....	18
4.5 Validação dos modelos atômicos modelados	20
4.5.1 MolProbity	20
4.5.2 ProSA.....	21
4.5.3 Verify 3D	24
4.5.4 PROCHECK	27
4.6 Docking molecular entre a proteína E do ZIKV e os scFv modelados	34
4.7 Análise 2D das interações intermoleculares entre a proteína E do ZIKV e os scFv modelados.....	36
5. CONCLUSÃO	43
6. REFERÊNCIAS BIBLIOGRÁFICAS.....	44
7. ANEXO.....	55

LISTA DE TABELAS

Tabela 1 - Posição de atração para o docking molecular entre os fragmentos de anticorpos e a proteína E do ZIKV.....	11
Tabela 2 - Orientação dos aminoácidos antes e após inversão conformacional	21
Tabela 3 – Pontuação dos modelos atômicos analisados no MolProbity	21
Tabela 4 - Pontuação dos modelos atômicos com base no gráfico de Ramachandran disponibilizado no PROCHECK.....	28
Tabela 5 - Parâmetros da cadeia principal dos modelos atômicos analisados no PROCHECK	33
Tabela 6 - Parâmetros da cadeia lateral dos modelos atômicos analisados no PROCHECK	34
Tabela 7 – Scores do docking molecular entre a proteína E e os modelos atômicos	35
Tabela 8 – Resíduos da proteína E do ZIKV que realizam ligações de hidrogênio com resíduos do scFv 2 com linker original.....	38
Tabela 9 – Resíduos da proteína E do ZIKV que realizam ligações de hidrogênio com resíduos do scFv 2 com linker modificado	39
Tabela 10 – Resíduos da proteína E do ZIKV que realizam ligações de hidrogênio com resíduos do scFv quimérico com linker original	41
Tabela 11 – Resíduos da proteína E do ZIKV que realizam ligações de hidrogênio com resíduos do scFv quimérico com linker modificado.....	41

LISTA DE FIGURAS

Figura 1 – Estrutura do RNA do ZIKV.....	3
Figura 2 - Representação dos domínios da proteína E do ZIKV	4
Figura 3 - Esquema da forma imatura e madura do ZIKV	4
Figura 4 - Esquema da ligação entre a proteína E e a membrana do hospedeiro.....	5
Figura 5 – Alinhamento entre os CDRs das regiões variáveis pesada e leve do scFv 2.....	13
Figura 6 - Esquema do scFv 2.....	14
Figura 7 - Alinhamento entre os CDRs das regiões variáveis pesada e leve do scFv 4.....	15
Figura 8 - Esquema do scFv quimérico.....	15
Figura 9 - Esquema do linker original e modificado.....	16
Figura 10 - Modelos tridimensional dos scFv 2 com linker original e modificado	17
Figura 11 - Modelos atômicos dos scFv quiméricos com linker original e modificado	18
Figura 12 - Alinhamento dos scFv 2 com linker original e modificado e gráfico de RMSD	19
Figura 13 - Alinhamento dos scFv quiméricos com linker original e modificado e gráfico de RMSD.....	20
Figura 14: Gráfico do Z-score dos modelos atômicos	23
Figura 15: Gráfico de energia livre dos modelos atômicos	24

Figura 16 – Análise de compatibilidade 3D-1D do scFv 2 com linker original	25
Figura 17 - Análise de compatibilidade 3D-1D do scFv 2 com linker modificado	26
Figura 18 - Análise de compatibilidade 3D-1D do scFv quimérico com linker original.....	26
Figura 19 - Análise de compatibilidade 3D-1D do scFv quimérico com linker modificado	27
Figura 20 – Gráfico dos ângulos Chi1-Chi2 da cadeia lateral do scFv 2 com linker original	29
Figura 21 - Gráfico dos ângulos Chi1-Chi2 da cadeia lateral do scFv 2 com linker modificado	30
Figura 22 - Gráfico dos ângulos Chi1-Chi2 da cadeia lateral do scFv quimérico com linker original	31
Figura 23 - Gráfico dos ângulos Chi1-Chi2 da cadeia lateral do scFv quimérico com linker modificado.....	32
Figura 24 - Docking molecular entre os fragmentos de anticorpos e a proteína E de ZIKV	36
Figura 25 - Interações intermoleculares entre o scFv 2 com <i>linker</i> original e a proteína E do ZIKV	37
Figura 26 - Interações intermoleculares entre o scFv 2 com linker modificado e a proteína E do ZIKV	38
Figura 27 - Interações intermoleculares entre o scFv quimérico com linker original e a proteína E do ZIKV	40

Figura 28 - Interações intermoleculares entre o scFv quimérico com linker modificado e a proteína E do ZIKV	40
Figura 29 - Gráfico de Ramachandran do scFv 2 com linker original	55
Figura 30 - Gráfico de Ramachandran do scFv 2 com linker modificado	56
Figura 31 - Gráfico de Ramachandran do scFv quimérico com linker original.....	57
Figura 32 - Gráfico de Ramachandran do scFv quimérico com linker modificado.....	58
Figura 33 - Parâmetros da cadeia principal do scFv 2 com linker original.....	59
Figura 34 - Parâmetros da cadeia principal do scFv 2 com linker modificado.....	60
Figura 35 - Parâmetros da cadeia principal do scFv quimérico com linker original	61
Figura 36 - Parâmetros da cadeia principal do scFv quimérico com linker modificado.....	62
Figura 37 - Parâmetros da cadeia lateral do scFv 2 com linker original.....	63
Figura 38 - Parâmetros da cadeia lateral do scFv 2 com linker modificado	64
Figura 39 - Parâmetros da cadeia lateral do scFv quimérico com linker original.....	65
Figura 40 - Parâmetros da cadeia lateral do scFv quimérico com linker modificado.....	66

1. INTRODUÇÃO

Em 1947, o vírus Zika (ZIKV) foi inicialmente isolado de um macaco sentinela do gênero Rhesus, capturado na floresta de Zika, na Uganda, sendo caracterizada como linhagem africana. No ano seguinte, novamente o vírus foi isolado, no entanto em mosquitos *Aedes (Ae) africanus* presentes na mesma área florestal (DICK; KITCHEN; HADDOW, 1952). Na década de 50 foi realizado o primeiro isolamento do vírus em humanos na Nigéria (PETERSEN, 2016) e na década seguinte, foi descoberto uma nova linhagem na Ásia, que foi isolada em mosquitos *Ae. aegypti* na Malásia, permitindo deduzir que o ZIKV possui um ciclo urbano, devido a descoberta em um vetor urbano (SILVA, 2019; SONG, 2017). Sendo assim, a presença do ZIKV em diferentes países na Ásia e África sugere que o vírus estava amplamente distribuído em ambos continentes (WIKAN, 2016). Ao longo de 60 anos, os casos relacionados ao ZIKV eram esporádicos ou em pequenos grupos de pacientes localizados na Ásia e África, o que contribuiu para que sua gravidade fosse subestimada (SILVA, 2019; PLOURDE, 2016).

Em 2007, o ZIKV teve notoriedade ao ser notificado um grande surto de casos nas ilhas Yap da Micronésia, sendo que esta foi a primeira notificação de casos fora dos continentes africano e asiático. Desde então, a infecção por ZIKV começou a se disseminar em diferentes regiões do mundo, como Polinésia Francesa, Ilhas Cook, Ilha de Páscoa e Nova Caledônia, até que, em 2015, um grande número de casos foi notificado no Brasil, resultando na primeira evidência do vírus nas Américas (MUSSO, 2016; PLOURDE, 2016). Com o surgimento do ZIKV no Brasil, rapidamente o vírus se espalhou pelo país e, consequentemente, pelo continente, devido aos vetores *Ae. aegypti* e *Ae. Albopictus* serem altamente distribuídos na região devido a globalização (GUBLER, 2017; MUSSO, 2016; SONG, 2017).

A principal forma de transmissão do ZIKV acontece por meio da picada de mosquito do gênero *Aedes*. De forma que, no ambiente urbano, há a predominância da espécie *Ae. Aegypti* e em menores casos a espécie *Ae. albopictus* (JÚNIOR, 2015; PETERSEN, 2016; SILVA,

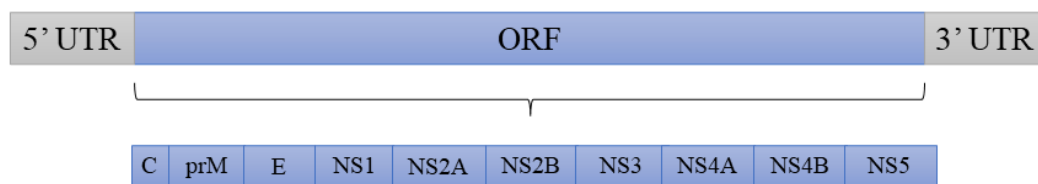
2019). Sendo assim, esses artrópodes provocam uma grande preocupação para a saúde pública, visto que são altamente distribuídos em zonas tropicais, subtropicais (*Ae. Aegypti*) e temperadas (*Ae. Albopictus*), abrangendo um extenso território para a incidência de surtos (JÚNIOR, 2015; PETERSEN, 2016; PLOURDE, 2016). Além da transmissão vetorial, existem outros modos de infecção pelo ZIKV, sendo estes por aleitamento materno, via sexual e durante a gestação (PETERSEN, 2016; PLOURDE, 2016; SILVA, 2019).

A maioria dos casos de ZIKV são assintomáticos, porém quando presentes as manifestações clínicas, há a presença de vários sintomas leves, comuns e inespecíficos, dentre eles febre, cefaleia, exantema e fadiga. Os sinais clínicos do ZIKV são similares com de outros arboviroses, como Dengue e Chikungunya, porém geralmente a infecção causada pelo Zika tem uma evolução benigna, na qual os sintomas persistem por até uma semana (CARVALHO, 2019; SILVA, 2019; SONG, 2017; PLOURDE, 2016; PETERSEN, 2016). No entanto, além dessas manifestações, a infecção por ZIKV está associada a complicações neurológicas. Durante o surto na Polinésia Francesa, houve um acréscimo de casos de Síndrome Guillain-Barré (SGB) que desenvolveram após a infecção por ZIKV (JÚNIOR, 2015, SILVA, 2019, MUSSO, 2016). Outro sinal clínico advém da capacidade do vírus de penetrar a placenta, o que pode ocasionar diversas complicações neurológicas em neonatais, como microcefalia e outras anomalias congênitas e a morte de neonatais (JÚNIOR, 2015; SONG, 2017).

O ZIKV pertence ao gênero *Flavivirus*, com capsídeo de geometria icosaédrica e seu genoma é composto por uma fita simples de ácido ribonucleico (RNA) de polaridade positiva, com aproximadamente 11kb, com porções não-traduzíveis (UTR) nas extremidades 5' e 3' e entre estas uma única matriz de leitura aberta (ORF) (Figura 1) (CARVALHO, 2019; SILVA, 2019). A partir da tradução do seu RNA, ocorre a síntese de uma única poliproteína viral, que é processada pelo o retículo endoplasmático em dois grupos de proteínas: estruturais e não estruturais. Os três tipos de proteínas estruturais no ZIKV são: capsídeo (C), membrana (M),

sendo esta gerada a partir do seu precursor, prM, e envelope (E), enquanto as sete não-estruturais consistem em NS1, NS2A, NS2B, NS3, NS4A, NS4B e NS5 (Figura 1). Dessa maneira, as proteínas estruturais são responsáveis para a formação da partícula viral, enquanto as não-estruturais estão relacionadas a replicação, empacotamento do genoma viral e a entrada na célula hospedeira (SILVA, 2019; SIROHI, 2017).

Figura 1 – Estrutura do RNA do ZIKV



Legenda: O genoma do ZIKV é composto por uma fita simples de RNA com porções não-traduzíveis (UTR) nas extremidades 5' e 3' e entre estas uma matriz de leitura aberta (ORF). A tradução do RNA viral resulta em uma única poliproteína viral, a qual possui dois grupos de proteínas: estruturais (C, capsídeo; prM, precursor da membrana; E, envelope) e não estruturais (NS1; NS2A; NS2B; NS3; NS4A; NS4B; NS5).

A proteína E está disposta em uma estrutura dimérica antiparalela no vírus e está ancorada a membrana viral, por meio de duas hélices transmembranares (TM). Essa proteína possui três domínios: DI, DII e DIII (Figura 2). O DI é dividido em três segmentos: posições 1-51, 132–192 e 280–295 promove a ligação entre os domínios II e III. Enquanto, o DII é formado por dois segmentos: posições 52–131 e 193–279 e possui uma alça de fusão ou “*fusion loop*” (FL) que é responsável pela fusão da membrana viral com a membrana do hospedeiro. Por fim, o DIII está localizado nos resíduos 296-403 e possui o sítio de ligação ao receptor, o qual permite a ligação do vírus na célula hospedeira. Com isso, tanto o DII quanto o DIII são regiões importantes para alvos terapêuticos, visto que a porção FL do DII é a sequência mais conservada entre os diferentes tipos de flavivírus e propicia uma reatividade cruzada de anticorpos, enquanto o DIII é a região menos conservada entre os flavivírus (cerca de 35% de conservação), sendo assim um candidato ideal para evitar a reatividade cruzada. Além disso, o DIII é uma

estrutura de ligação a imunoglobulinas, sendo assim importante no reconhecimento anticorpos neutralizares (AGRELLI, 2019; DAI, 2016; HEINZ, 2017; SILVA, 2019).

Figura 2 - Representação dos domínios da proteína E do ZIKV

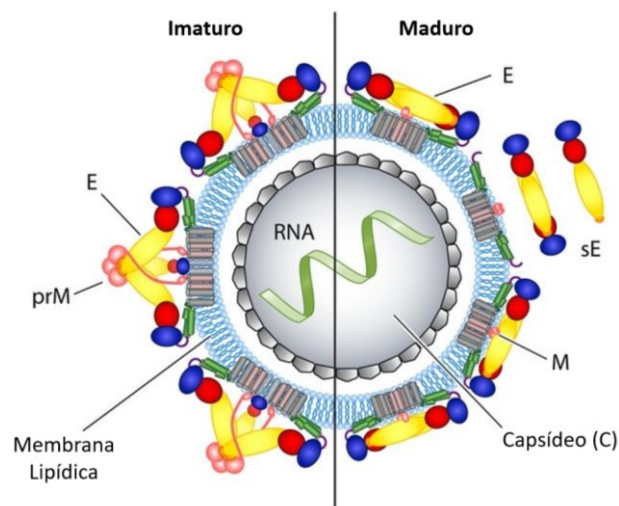


Fonte: Adaptado de Saiz et al, 2017.

Legenda: Estrutura em cristalografia de raios X da proteína E do ZIKV (ID PDB: 5JHM), com o domínio I (DI) em vermelho, domínio II (DII) em amarelo, domínio III (DIII) em azul e o loop de fusão em verde.

O ZIKV imaturo possui 60 nm e possui projeções formadas trímetros de heterodímeros da prM-E, o que torna a superfície pontiaguda. Enquanto, a forma madura possui 50 nm e uma superfície lisa, a qual é proporcionada pela orientação paralela dos homodímeros da proteína E com a membrana viral (Figura 3) (HEINZ, 2017; SIROHI, 2017).

Figura 3 - Esquema da forma imatura e madura do ZIKV



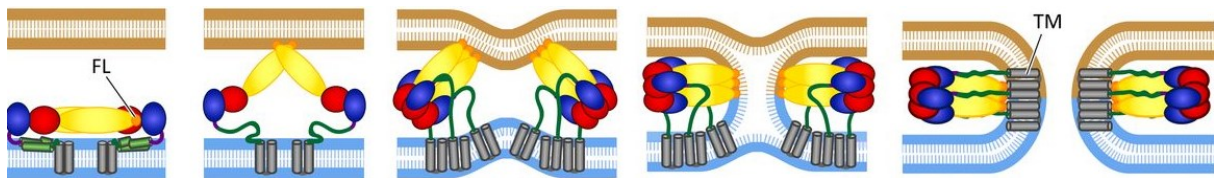
Fonte: Adaptado de Heinz et al, 2017.

Legenda: C, capsídeo; M, Membrana; prM, precursor da membrana; E, envelope; sE, proteína E solúvel.

O principal método de entrada do ZIKV nas células envolve o mecanismo de endocitose mediado por clatrina, uma vez que essa proteína é responsável pela captura de moléculas do

meio extracelular para o intracelular. Com isso, quando acontece a ligação da proteína E com o receptor da célula hospedeira, ocorre a endocitose do vírus, formando uma vesícula coberta por clatrina. Após a formação da vesícula, as clatrininas são liberadas e, dessa forma, ocorre a redução do pH no endossomo, o qual promove a dissociação do dímero da proteína E. Com essa dissociação, a ligação da região FL do DII na membrana vesicular é possibilitada. Em seguida, três cópias da proteína interagem entre si e originam um trímero em forma de grampo, com as regiões FL e TM justapostas (Figura 4) (SILVA, 2019; HEINZ, 2017; RODRIGUES, 2021).

Figura 4 - Esquema da ligação entre a proteína E e a membrana do hospedeiro



Fonte: Heinz et al, 2017.

Legenda: As cores das estruturas estão descritas na legenda da figura 3. O pH ácido do endossomo promove a dissociação do dímero da proteína E, o que permite a ligação da porção *fusion loop* (FL) no endossomo. Em seguida, há a formação de um trímero em forma de grampo da proteína E com a região FL e TM (transmembrana) justapostas. Por fim ocorre a fusão das camadas lipídicas.

Essas mudanças conformacionais da proteína E fornecem energia que possibilita a fusão das membranas lipídicas e, conseqüentemente, a liberação do RNA viral no citoplasma da célula hospedeira. Com pH neutro do retículo endoplasmático (RE), a partícula viral é montada com um trímero de prM/E, originando o ZIKV imaturo. Em seguida, essa estrutura é transportada para a rede trans do complexo de Golgi (TGN), onde o pH levemente ácido permite a clivagem do prM pela protease furina. Dessa maneira, em um pH neutro, os peptídeos de pr são liberados e ocorre a formação de homodímeros da proteína E, originando o vírus maduro e, promovendo, assim, a liberação deste no meio extracelular (AGRELLI, 2019; HEINZ, 2017; SIROHI, 2017; SMIT, 2011).

Dessa maneira, a proteína E é um dos principais alvos para anticorpos específicos ao Zika vírus, uma vez que bloqueiam a entrada do vírus no hospedeiro (FRANÇA, 2020; PACHECO,

2021). Com isso, a construção de anticorpos monoclonais contra infecções virais é uma abordagem terapêutica relevante, de forma que se pode isolar regiões biologicamente ativas do anticorpo, a fim de obter vantagens nos efeitos terapêuticos (CARDOSO, 2008; FRANÇA, 2022; ZHAO, 2016). Em vista disso, o desenvolvimento de fragmentos de anticorpos é vantajoso devido ao seu menor tamanho, o que permite a penetração em alguns órgãos inacessíveis a anticorpos inteiros (CASTRO, 2019; SOUZA, 2007). Um fragmento variável de cadeia simples (scFv) consiste na união das cadeias variáveis pesada (VH) e leve (VL) por meio de uma sequência pequena de aminoácidos, denominada *linker*. A composição e comprimento do *linker* podem afetar a estabilidade, o dobramento, a associação entre os domínios e a integridade do domínio antígeno-anticorpo (MOHAMMADI, 2016; TANG, 1996). Cada domínio variável dos anticorpos possui três regiões de hipervariabilidade, denominada de região determinante de complementariedade (CDR), os quais formam os sítios de ligação com os antígenos (CASTRO, 2019; FRANÇA, 2020, SEGER, 2010).

A modelagem molecular e a ancoragem molecular (*docking*) são metodologias computacionais que têm permitido a obtenção de informações sobre complexos ligante-alvo de forma similar as metodologias tradicionais, como cristalografia e difração de raio X e ressonância nuclear magnética, porém de forma mais ágil e simples. Dentro desse contexto, métodos *in silico* são importantes para o descobrimento de novos inibidores contra o Zika vírus, assim como a ampliação da compreensão sobre características físicas e bioquímicas da interação proteína-ligante (KOZAKOV, 2017; RAMHARACK, 2018; SINIGAGLIA, 2018). Além disso, essas análises *in silico* podem ser utilizadas para prever modificações que podem melhorar a afinidade entre o antígeno e o anticorpo, a fim de contribuir na sua construção (KURODA, 2012; YAMASHITA, 2018).

Portanto, neste trabalho monográfico, foi realizado modificações em um fragmento de anticorpo anti-ZIKV e com o intuito de verificar, por meio de análises *in silico*, se houve mudanças na interação com a proteína E do ZIKV.

2. OBJETIVOS

2.1 Objetivo geral

Verificar por meio de análises computacionais (*in silico*) a interação entre a proteína E do ZIKV e scFv modificados.

2.2 Objetivos específicos

- Modelar os fragmentos de anticorpo original e modificado, por meio da modelagem *ab initio*;
- Validar os modelos atômicos modelados em diferentes servidores computacionais;
- Avaliar se as modificações no *linker* entre segmentos de cadeia variável pesada e leve alteram a estrutura do fragmento de anticorpo, por meio do cálculo do Root-Mean-Square Deviation (RMSD);
- Realizar alterações nas regiões determinantes da complementariedade 3 (CDR 3) do scFv;
- Realizar a ancoragem molecular (*docking*) entre a proteína E do ZIKV e os fragmentos de anticorpos;
- Analisar por ferramentas *in silico* as interações moleculares polares entre a proteína E do ZIKV e os scFv.

3. METODOLOGIA

3.1 Modificações realizadas na sequência do scFv anti-ZIKV

As sequências de scFv utilizadas (IAL scFv 2 e scFv 4) foram disponibilizadas gentilmente pelo prof. Dr. Carlos Roberto Prudêncio e pelo Me. Hernan Hermes Monteiro da Costa, do Instituto Adolfo Lutz, São Paulo. O scFv 2 foi utilizado para as análises *in silico*, uma vez que apresentou uma diferença significativa entre a reatividade para proteína E do ZIKV e DENV nos experimentos realizado por (COSTA, 2022). No mesmo estudo, o scFv 4 obteve uma maior atividade de ligação aos antígenos do ZIKV e DENV. Com isso, modificações foram realizadas nesses fragmentos de anticorpos, a fim de gerar os modelos tridimensionais para as análises computacionais.

Portanto, a primeira modificação ocorreu no CDR 3 da região variável leve, de forma que houve a substituição da CDR 3 do scFv 2 pela CDR 3 do scFv 4, gerando um scFv quimérico. Devido ao scFv 4 previamente descrito apresentar uma maior reatividade aos antígenos de Zika Vírus em ELISA, nesta monografia foi verificado o impacto dessa modificação no sítio de ligação do fragmento de anticorpo com a proteína E de ZIKV. Outra modificação descrita neste trabalho ocorreu na sequência de aminoácidos do *linker*, na qual foi substituída as glicinas das posições 125 e 128 por duas prolínas nas respectivas posições. Esta modificação teve o intuito de verificar se modificações na região espaçadora dos scFv afetam a estrutura dos anticorpos scFv ou mesmo a ligação à proteína E.

3.2 Análise das regiões determinantes da complementariedade dos scFv 2 e 4

As regiões de determinação de complementariedade (CDRs) são sequências existentes nas regiões variáveis do anticorpo. As CDRs tem papel importante na ligação dos anticorpos aos antígenos e podem ser desenhadas para desenvolvimento de imunobiológicos, com potencial medicamentoso, visto que são regiões complementares ao antígeno, de modo que as CDR3 de cada segmento são as mais variáveis (ABBAS, 2015; RATHORE, 2019; YANG, 2009). Dessa

maneira, a análise das CDRs foi realizada combinando os servidores online ImMunoGeneTic (IMGT), Paratome, Parapred. Com isso, no IMGT, a ferramenta Domain Gap Align foi executada, de forma que esse software realiza a análise da sequência de aminoácidos por domínios. Ocorre a identificação dos domínios variável (V), constante (C) e *groove* (G) utilizando o diretório de referência de domínio IMGT, o que cria lacunas (*gaps*) de acordo com a numeração do IMGT e indica as diferenças em relação ao referencial (EHRENMANN, 2011; 2012). O programa Paratome aplica o método “*machine learning*”, por meio do alinhamento múltiplo estrutural (MSTA) com diferentes anticorpos existentes nos bancos de dados. Dessa forma, ocorre determinação das regiões de consenso, de modo que os resíduos que interagem com o antígeno são altamente semelhantes entre os anticorpos conhecidos, o que permite a identificação das CDRs da sequência desejada (KESARWANI, 2021; KUNIK, 2012). A plataforma Parapred baseia-se na predição probabilística de parátomos da sequência, sendo que o algoritmo utiliza o método de “*deep-learning*” para identificar as regiões hipervariáveis de acordo com a definição de Chothia (KESARWANI, 2021; LIBERIS, 2018; SAINI, 2021).

3.3 Determinação das regiões hipervariáveis do scFv 2 e 4

A ferramenta Basic Local Alignment Search Tool (BLAST) foi utilizada para encontrar sequências de proteínas de *Homo sapiens* (taxid:9606) homologas às regiões previamente analisadas nos servidores IMGT, Paratome e Parapred. E por meio do Clustal X, foi realizado o alinhamento entre essas regiões com o intuito de realizar a comparação das CDRs e assim, determinar os aminoácidos (intervalo de aminoácidos) que compõem cada CDR dos anticorpos scFv.

3.4 Modelagem molecular tridimensional (3D) dos scFv original e modificados

Todas as sequências foram modeladas no software de predição automática Robetta, por meio do método de “*deep-learning*” RoseTTAFold. Essa ferramenta prevê com mais precisão a estrutura tridimensional de uma proteína a partir da sua sequência de aminoácidos, devido a

uma rede neural que possui três vias. A existência de múltiplas conexões entre os caminhos da predição, permite que a rede aprenda simultaneamente relacionamentos dentro e entre sequências, distâncias e coordenadas da estrutura proteica (BAEK, 2021, 2022; FERREIRA, 2022)

3.5 Cálculo do desvio quadrático médio (RMSD) dos alinhamentos dos fragmentos de anticorpos

O cálculo do RMSD dos alinhamentos entre scFv 2 com *linker* original e modificado e entre scFv quimérico com *linker* original e modificado foi realizado pelo software Chimera UCSF. Sendo assim, foi executado o alinhamento estrutural por meio da ferramenta “MatchMaker” e, com isso, a partir do recurso “RMSD:ca”, foi possível gerar um arquivo contendo os valores de RMSD de cada resíduo. A análise dos dados foi realizada através de gráficos gerados no software Microsoft Excel (PETTERSEN, 2004).

3.6 Avaliação da qualidade dos modelos gerados

A ferramenta PROCHECK (LASKOWSKI, 1993) foi utilizada para avaliar a estereoquímica da estrutura proteica gerada, sendo que essa análise ocorre pela geometria de resíduo por resíduo e pela geometria geral da estrutura proteica (DHURIGAI, 2014). O gráfico de Ramachandran é um dos resultados disponibilizado pelo software que define quais resíduos estão em regiões energeticamente favoráveis e desfavoráveis (FILHO, 2003; PINTO, 2018). O VERIFY 3D (BOWIE, 1991) é uma plataforma que analisa a compatibilidade da estrutura tridimensional da proteína com sua própria sequência primária de aminoácidos, de forma que a atribuição da classe estrutural é baseada no ambiente e na localização das moléculas, além de ocorrer comparações com estruturas de boas qualidades (DHURIGAI, 2014; JÚNIOR, 2013). O MolProbity (CHEN, 2010) é uma ferramenta que verifica o contato de todos os átomos da estrutura, detalhando os problemas estéricos das moléculas, fornecendo as ligações de hidrogênio e as interações de van der Waals, a fim de obter uma melhor qualidade

estereoquímica do modelo (DAVIS, 2004; 2007). Para tal efeito, o software Reduce realiza adição de hidrogênios na proteína e otimiza em posições favoráveis. Além disso, essa ferramenta realiza inversões de resíduos de Asparagina (Asp), Glutamina (Gln) e Histidina (His), visto que suas cadeias laterais são facilmente desorientadas. E o programa ProSA-Web (WIEDERSTEIN, 2007) avalia a qualidade geral e local da proteína modelada, por meio das coordenadas atômicas dos carbonos- α ($C\alpha$), de forma que resultados gerados são baseados em um gráfico de *Z-score* para a estrutura geral e em um gráfico de energia para a estrutura local (FREITAS, 2019).

3.7 Interação molecular (*docking*) entre a proteína do envelope (E) do ZIKV

O *docking* molecular entre a proteína E do ZIKV e os scFv modelados foi realizado pelo programa ClusPro (KOZAKOV, 2017), o qual se baseia em três etapas. A primeira baseia-se na ancoragem do ligante em um receptor rígido, a segunda no agrupamento das 1000 estruturas de menor energia baseada no desvio quadrático médio (RMSD) e a última consiste no refinamento das estruturas utilizando a minimização de energia. O software PIPER, foi utilizado para o *docking* rígido, de forma que esse programa se fundamenta nas transformadas rápidas de Fourier para calcular a energia da interação, sem necessitar de informações adicionais sobre a estrutura do complexo proteína-proteína (KOZAKOV, 2017; SANTOS, 2019). Nos parâmetros do ClusPro, o modo selecionado foi o “Antibody Mode” e as regiões não-CDR das cadeias de anticorpos foram mascaradas automaticamente. Além disso, foram inseridos os resíduos referentes as regiões hipervariáveis como posições de atração, como demonstrado na Tabela 1. Com isso, os complexos que obtiverem os menores valores (*score*) de ancoragem serão selecionados (BANDEHPOUR, 2017; RAFIGHDOUST, 2020).

Tabela 1 - Posição de atração para o *docking* molecular entre os fragmentos de anticorpos e a proteína E do ZIKV

Fragmento de anticorpo	Posição de atração
------------------------	--------------------

scFv 2	A-24 S-25 G-26 Y-27 T-28 F-29 T-30 S-31 Y-32 D-33 I-34 F-45 E-46 W-47 M-48 G-49 W-50 M-51 N-52 P-53 N-54 S-55 G-56 N-57 T-58 G-59 Y-60 A-61 C-96 A-97 R-98 D-99 G-100 V-101 A-102 V-103 A-104 G-105 L-106 D-107 Y-108 W-109 G-110 Q-146 D-147 I-148 D-149 D-150 D-151 A-163 I-164 F-165 I-166 I-167 Q-168 E-169 A-170 T-171 T-172 L-173 V-174 P-175 F-206 C-207 L-208 Q-209 H-210 D-211 N-212 F-213 P-214 L-215 T-216 F-217 G-218
scFv quimérico	A-24 S-25 G-26 Y-27 T-28 F-29 T-30 S-31 Y-32 D-33 I-34 F-45 E-46 W-47 M-48 G-49 W-50 M-51 N-52 P-53 N-54 S-55 G-56 N-57 T-58 G-59 Y-60 A-61 C-96 A-97 R-98 D-99 G-100 V-101 A-102 V-103 A-104 G-105 L-106 D-107 Y-108 W-109 G-110 Q-146 D-147 I-148 D-149 D-150 D-151 A-163 I-164 F-165 I-166 I-167 Q-168 E-169 A-170 T-171 T-172 L-173 V-174 P-175 Y-223 C-224 Q-225 Q-226 H-227 D-228 N-229 L-230 P-231 I-232 T-233 F-234 G-235

Legenda: As letras correspondem aos aminoácidos. O número representa a posição do aminoácido. A, Alanina; C, Cisteína.; D, Aspartato; E, Glutamato; F, Fenilalanina; G, Glicina; H, Histidina; I, Isoleucina; K, Lisina; L, Leucina; M, Metionina; N, Asparagina; P, Prolina; Q, Glutamina; R, Arginina; S, Serina, T, Treonina; V, Valina; W, Triptofano; Y, Tirosina, Z, Glutamato.

3.8 Análise 2D das interações intermoleculares entre a proteína E do ZIKV com o scFv

As interações intramoleculares do complexo receptor-ligante foram analisadas na ferramenta LigPlot, pelo comando DIMPLOT (WALLACE, 1995). Esse software realiza a reposição dos resíduos do complexo tridimensional (3D) para posições equivalentes em uma estrutura bidimensional (2D) a partir de um arquivo no formato PDB. Sendo assim, foi possível analisar e verificar as distâncias em Angstroms (Å) das ligações de hidrogênio e as interações

hidrofóbicas existentes entre os aminoácidos da proteína E do ZIKV e dos fragmentos de anticorpos modelados (LASKOWSKI, 2011).

4. RESULTADOS E DISCUSSÃO

4.1 Determinação das regiões hipervariáveis das sequências de scFv 2 e 4

Modificações nas CDRs podem ser realizadas, a fim de aumentar a afinidade de ligação antígeno-anticorpos (ABBAS, 2015; RATHORE, 2019; YANG, 2009). Em vista disso, a sequência do scFv 2 foi submetida nos servidores IMGT, Paratome e Parapred, a fim de identificar as CDRs das regiões variáveis. Em seguida, realizou a busca de cinco sequências de aminoácidos homologas aos scFv, por meio do programa BLAST. Conforme demonstrado na Figura 5, executou o alinhamento entre as CDRs do anticorpo scFv2 identificadas nas ferramentas IMGT, Paratome e Parapred com as cinco sequências homologas previamente identificadas utilizando o Basic Local Alignment Search Tool (BLAST).

Figura 5 – Alinhamento entre os CDRs das regiões variáveis pesada e leve do scFv 2

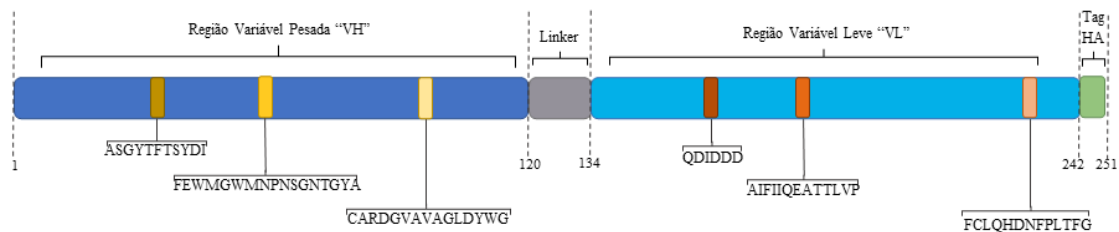


Legenda: (A) As três primeiras sequências consistem nas análises das CDRs da região variável leve do scFv 2 realizadas no ImMunoGeneTics (IMGT), no Paratome e no Parapred alinhadas com as 5 primeiras sequências de maior homologia disponibilizadas no “Basic Local Alignment Search Tool” (BLAST) (BAC02136.1; QSI97860.1; BAC02271.1; QIV65590.1 e QEP16364.1); (B) As três primeiras sequências consistem nas análises das CDRs da região variável pesada do scFv 2 realizadas no ImMunoGeneTics (IMGT), no Paratome e no Parapred alinhadas

com as 5 primeiras seqüências de maior homologia disponibilizadas no “Basic Local Alignment Search Tool” (BLAST) ((BAC01715.1; BAG62953.1; QEP27068.1; QEP13742.1 e BAC01716.1). A cor vermelha representa as CDR 1, 2 e 3 da cadeia variável leve e a cor azul representa as CDR 1, 2 e 3 da cadeia variável pesada.

Comparando o intervalo de aminoácidos das CDRs disponibilizadas nos servidores IMGT, Paratome e Parapred com as seqüências homólogas disponibilizadas no BLAST, foi determinado os resíduos de aminoácidos de cada região CDR (1,2 e 3), tanto para região da cadeia leve como cadeia pesada, indicadas na representação esquemática do constructo scFv2; Figura 6.

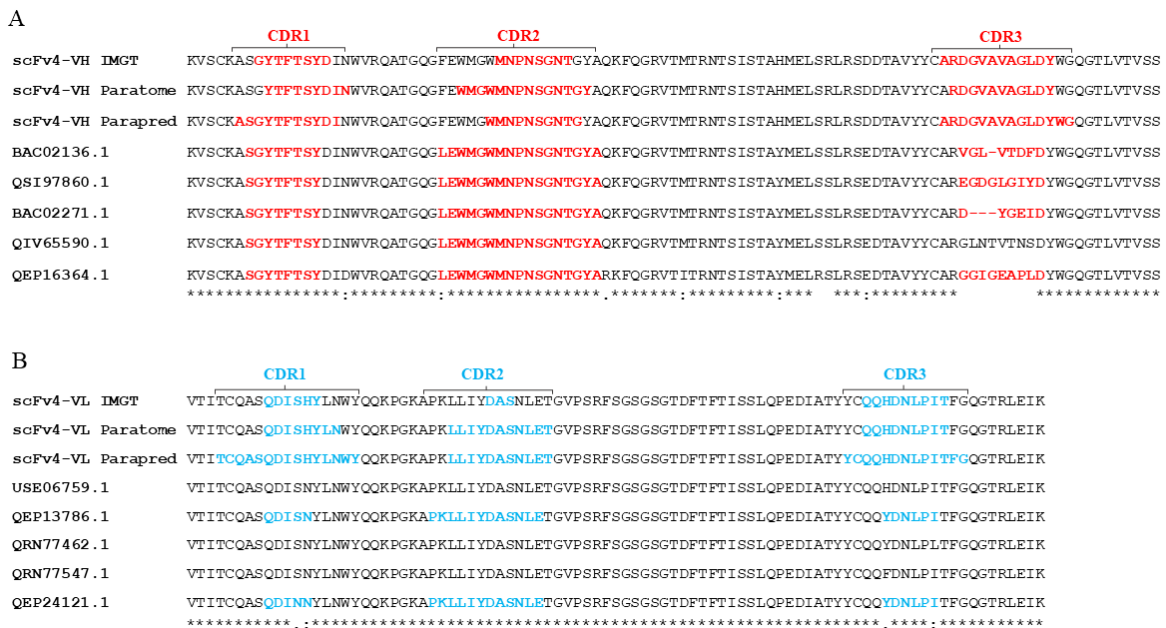
Figura 6 - Esquema do scFv 2



Legenda: (A) Esquema do scFv2 com destaque na região variável pesada (VH) em azul escuro, com as CDRs 1, 2 e 3 destacadas em tons de amarelo e suas seqüências apresentadas abaixo, com o linker em cinza, com a região variável leve (VL) em azul claro, com as CDRs 1, 2 e 3 destacadas em tons de vermelho e suas seqüências apresentadas abaixo e com a tag HA em verde.

Dessa forma, foi realizado a mesma metodologia com o scFv 4. Com isso, a Figura 7 demonstra o alinhamento das CDRs do scFv4 identificadas no IMGT, Paratome e Parapred com as cinco seqüências homólogas selecionadas no BLAST.

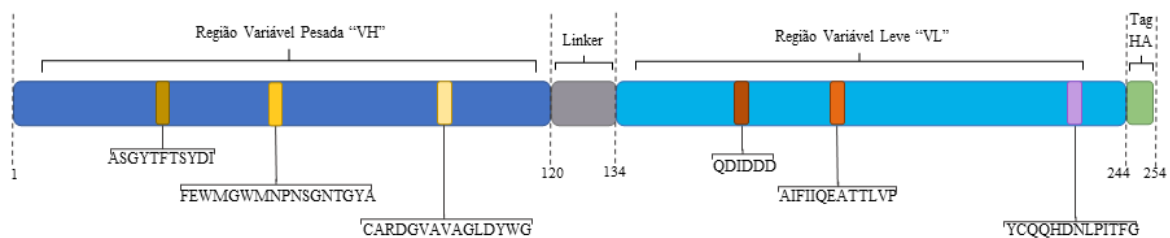
Figura 7 - Alinhamento entre os CDRs das regiões variáveis pesada e leve do scFv 4



Legenda: (A) As três primeiras sequências consistem nas análises das CDRs da região variável leve do scFv 4 realizadas no ImMunoGeneTics (IMGT), no Paratome e no Parapred alinhadas com as 5 primeiras sequências de maior homologia disponibilizadas no “Basic Local Alignment Search Tool” (BLAST) (BAC02136.1; QSI97860.1; BAC02271.1; QIV65590.1 e QEP16364.1); (B) As três primeiras sequências consistem nas análises das CDRs da região variável pesada do scFv 4 realizadas no ImMunoGeneTics (IMGT), no Paratome e no Parapred alinhadas com as 5 primeiras sequências de maior homologia disponibilizadas no “Basic Local Alignment Search Tool” (BLAST) (USE06759.1; QEP13786.1; QRN77462.1; QRN77547.1 e QEP24121.1). A cor vermelha representa as CDR 1, 2 e 3 da cadeia variável leve e a cor azul representas as CDR 1, 2 e 3 da cadeia variável pesada.

A partir dessa análise foi selecionada a CDR3 da região variável leve do scFv 4, com o intuito de substitui-la no scFv 2. Em vista disso, foi gerado um scFv quimérico como demonstrado esquematicamente na Figura 8.

Figura 8 - Esquema do scFv quimérico



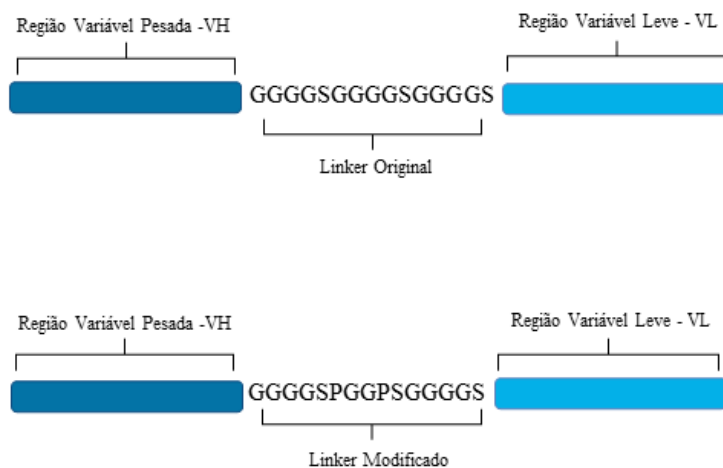
Legenda: (A) Esquema do scFv2 quimérico. A região variável pesada (VH) está indicada em azul escuro, com as respectivas CDRs 1, 2 e 3 destacadas em tons de amarelo. A sequência linker é indicada em cinza. Região variável leve (VL) representada em azul claro, com as CDRs 1 e 2 destacadas em tons de vermelho. A CDR 3 derivada do anticorpo scFv4 é demonstrada na cor lilás. As sequências de aminoácidos estão apresentadas abaixo das respectivas CDRs. A sequência C-terminal em verde indica o Tag de HA no constructo do scFv.

4.2 Modificações no linker do scFv

Modificações no *linker* de anticorpos scFv podem ser necessárias com o intuito de obter um ligante mais afim a molécula alvo (antígeno). A conformação do *linker* mais utilizada é rica em aminoácidos de glicina, $(Gly_4Ser)_n$, uma vez que a cadeia lateral do aminoácido glicina é composta por hidrogênio, o que promove uma flexibilidade e solubilidade ao fragmento de anticorpo (MOHAMMADI, 2016; TANG, 1996). Outra alternativa usada é uma conformação do *linker* rica em prolina para obter uma estrutura mais rígida e alterar as distâncias entre os domínios de anticorpos, podendo contribuir para uma maior estabilidade da molécula. Essas propriedades advêm devido a presença de uma estrutura cíclica da cadeia lateral da prolina (SPAHR, 2017; CHEN, 2013). Em vista disso, foram realizadas modificações no *linker* original do anticorpo scFv 2, conforme demonstrado na Figura 9, a fim de obter um *linker* mais flexível e possivelmente mais estável.

O *linker* original entre as regiões VH e VL do scFv possui a conformação $(Gly_4Ser)_3$ e o *linker* modificado possui a substituição das glicinas nas posições 125 e 128 por um prolina em cada posição. Essas modificações também foram realizadas no scFv quimérico.

Figura 9 - Esquema do linker original e modificado

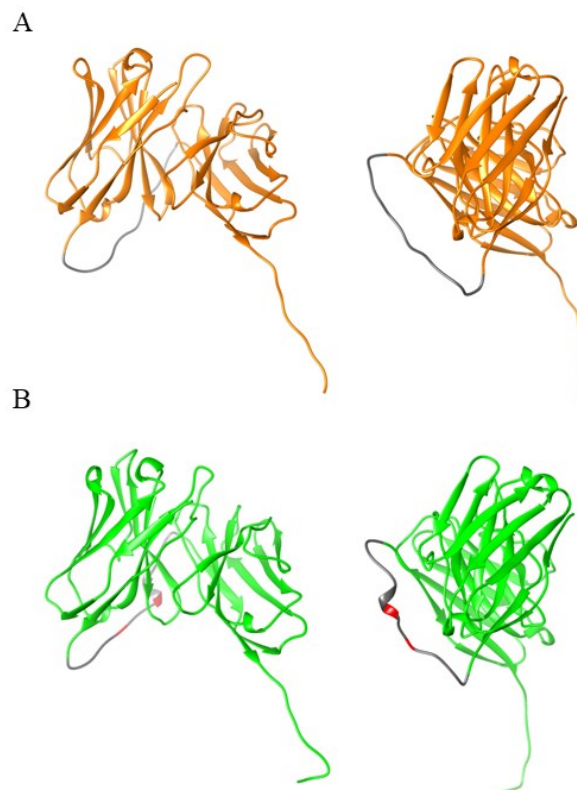


Legenda: Esquema do *linker* original e modificado do scFv. Sequência de aminoácidos do *linker* original (acima) e modificado (abaixo) entre a região variável pesada e a leve.

4.3 Modelagem dos scFv 2 e quiméricos com *linker* original e modificado

A modelagem realizada nos scFv foi pelo método *ab initio*, utilizando o servidor Robetta. Dessa forma, os modelos atômicos do scFv 2 com *linker* original e modificado estão expostos na Figura 10A e 10B, respectivamente. O *linker* modificado apresentou uma conformação levemente helicoidal na região próxima as prolinas, induzindo a rigidez para a estrutura (CHEN, 2013; JADAUN, 2020). No entanto, a adição das duas prolinas no *linker* não promoveram uma mudança conformacional substancial em toda estrutura tridimensional do scFv 2.

Figura 10 - Modelos tridimensional dos scFv 2 com linker original e modificado

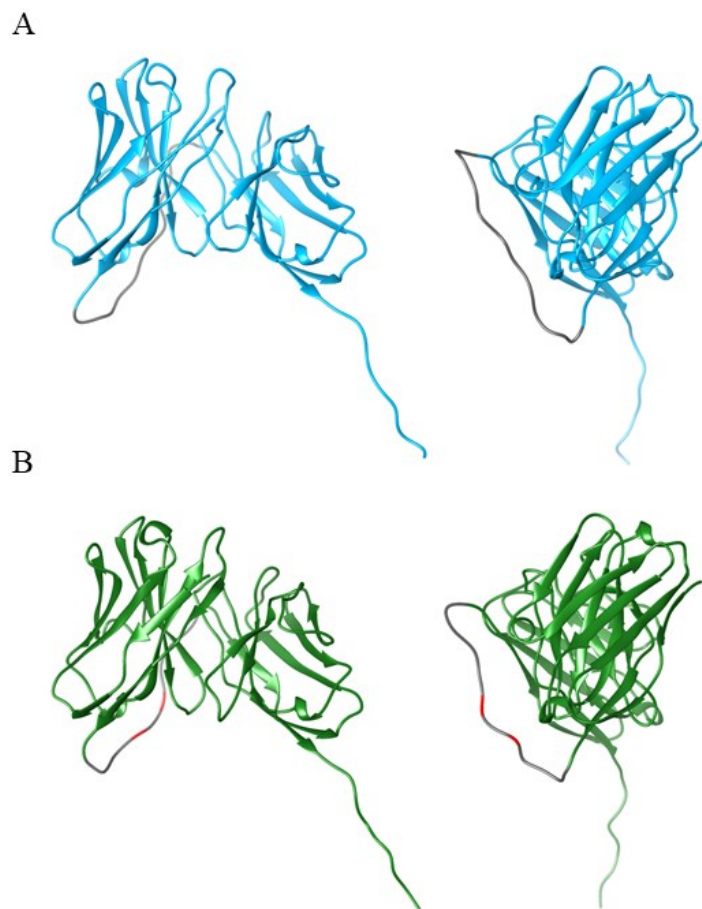


Legenda: (A) Modelo atômico do scFv 2 (em laranja) com *linker* original (destacado em cinza). (B) Modelo atômico do scFv 2 com o *linker* modificado (destacado em cinza) com prolinas identificadas em vermelho;

Com relação aos scFv quiméricos com *linker* original e modificado, na Figura 11A e 11B está exposto seus modelos atômicos respectivamente.

Assim como o scFv 2, a adição das prolinas não induziram uma mudança conformacional substancial em toda estrutura tridimensional dos scFv quiméricos. Além disso, diferentemente do scFv 2 com *linker* modificado, o scFv quimérico com o *linker* modificado não apresentou a estrutura helicoidal. Entretanto, observa-se uma diferença estrutural do *linker* modificado comparado com o original, de forma que existe a tendência de curvatura na região próxima as prolinas.

Figura 11 - Modelos atômicos dos scFv quiméricos com linker original e modificado



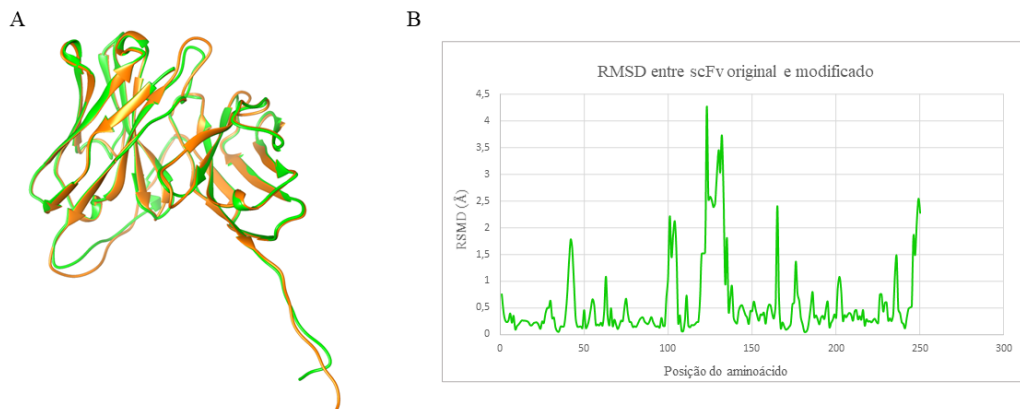
Legenda: (A) Modelo atômico do scFv quimérico (em azul claro) com *linker* original (destacado em cinza). (B) Modelo atômico do scFv quimérico (em verde escuro) com o *linker* modificado (destacado em cinza) com as prolinas identificadas em vermelho;

4.4 Alinhamento estrutural dos scFv 2 e quiméricos e cálculo do RMSD

No programa Chimera UCSF foi executado o alinhamento estrutural entre o scFv 2 com *linker* original e modificado, conforme demonstrado na Figura 12A. Com isso, foi possível

identificar os valores de RMSD de cada resíduo dos modelos e, assim plotar o gráfico RMSD x posição do aminoácido (Figura 12B).

Figura 12 - Alinhamento dos scFv 2 com linker original e modificado e gráfico de RSMD



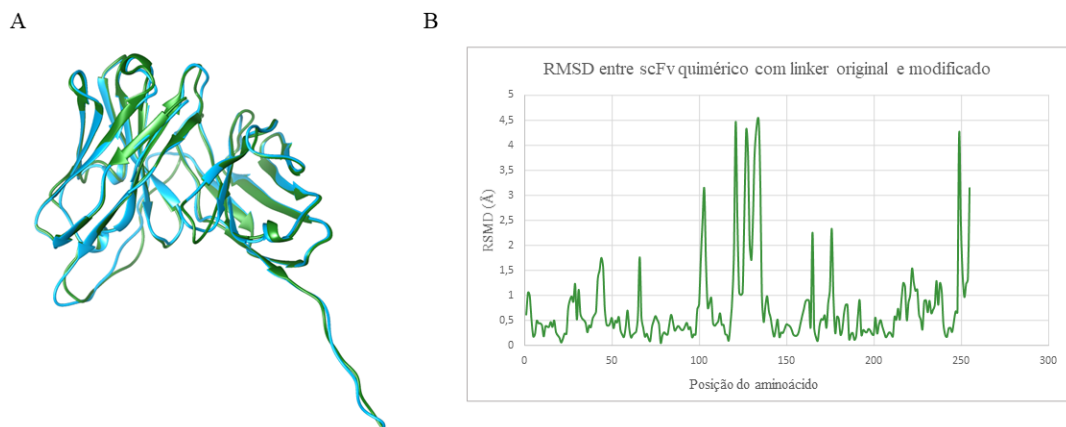
Legenda: (A) Alinhamento entre o scFv com linker original (em laranja) e modificado (em verde). (B) Gráfico do RSMD (Root Mean Square Deviation) em Ångström (Å) por resíduos dos alinhamentos entre o scFv2 com linker original e modificado.

O RMSD mede a distância média, em Ångström (Å), entre os átomos de estruturas alinhadas. Sendo assim, quanto maior for a variação de uma estrutura em relação a outra, maior será o valor do RMSD. Portanto, esse valor pode ser utilizado para verificar a similaridade entre estruturas (ALENCAR, 2021). Conforme a Figura 12B, a maioria dos valores de RMSD ficaram abaixo de 1Å, sendo que na região previamente ao *linker*, dentre os resíduos 101 a 105, obteve um valor acima de 1,5Å. As regiões que obtiveram maiores variações (acima de 2Å) foram na região do *linker*, do resíduo 123 ao 133 e na região terminal, entre os resíduos 245 a 250. Portanto, de modo geral, a modificação realizada no *linker* não foi o suficiente para modificar estruturalmente o scFv 2.

Para os scFv quiméricos foi realizado a mesma metodologia aplicada ao scFv 2. Dessa forma, a Figura 13A representa o alinhamento estrutural entre o scFv quimérico com *linker* original e modificado e a Figura 13B demonstra o gráfico RMSD em relação a posição dos aminoácidos dos scFv quiméricos. Dessa maneira, assim como o scFv 2, houve maiores variações no RMSD em três regiões, sendo elas: na região previamente ao *linker* (do resíduo

101 ao 104), no *linker* (do resíduo 120 ao 135) e na região terminal (do resíduo 249 ao 255). Porém, essas variações não foram suficientes para modificarem a conformação estrutural do scFv quimérico.

Figura 13 - Alinhamento dos scFv quiméricos com linker original e modificado e gráfico de RMSD



Legenda: (A) Alinhamento entre o scFv quimérico com linker original (em azul claro) e modificado (em verde escuro). (B) Gráfico do RMSD (Root Mean Square Deviation) entre o scFv quimérico com linker original e modificado.

4.5 Validação dos modelos atômicos modelados

A validação dos modelos atômicos gerados pelo Robetta foi validada em quatro servidores, sendo estes: MolProbity, ProSA, Verify 3D e Procheck.

4.5.1 MolProbity

No MolProbity, a análise dos contatos de todos os átomos é calculada pelo programa Probe, a qual realiza a substituição da posição do átomo por uma esfera, de forma que uma sobreposição maior que 0,4 Å entre átomos não doadores-receptores é identificado como “*serious clash*”. Com isso, os resultados do MolProbity consistem na combinação do “*clashscore*” que quantifica os confrontos (*serious clash*) por 1000 átomos, das análises dos rotâmeros e Ramachandran em um único score (MolProbity score) (CHEN, 2010; DAVIS, 2010; WILLIAMS, 2018).

Dessa maneira, a Tabela 2 apresenta os resíduos de Gln do scFv2 com *linker* original que necessitaram de inversão estereoquímica. Em relação aos outros fragmentos de anticorpos modelados, não houve necessidade de realizar inversões.

Tabela 2 - Orientação dos aminoácidos antes e após inversão conformacional

Modelo	Aminoácido	Origem	Após inversão
scFv2 com <i>linker</i> original	Gln (39)	0,19	1,1
	Gln (172)	0,19	1,1
	Gln (224)	-4,3	-1,6

Legenda: Valores da inversão conformacional dos aminoácidos em relação ao plano do eixo da cadeia principal da proteína. Número entre os parênteses representa a posição do aminoácido. Gln: Glutamina.

A Tabela 3 indica os resultados disponibilizados no MolProbity para os fragmentos de anticorpos modelados. Com isso, todos os modelos obtiveram pontuações que indicam uma alta qualidade (acima de 66%).

Tabela 3 – Pontuação dos modelos atômicos analisados no MolProbity

Modelo	MolProbity Score
scFv 2 com <i>linker</i> original	98%
scFv 2 com <i>linker</i> modificado	99%
scFv quimérico com <i>linker</i> original	99%
scFv quimérico com <i>linker</i> modificado	99%

Legenda: Valores de pontuação maiores que 66% indicam uma alta qualidade, entre 33% e 66% indicam qualidade mediana e menores que 33% indicam baixa qualidade do modelo.

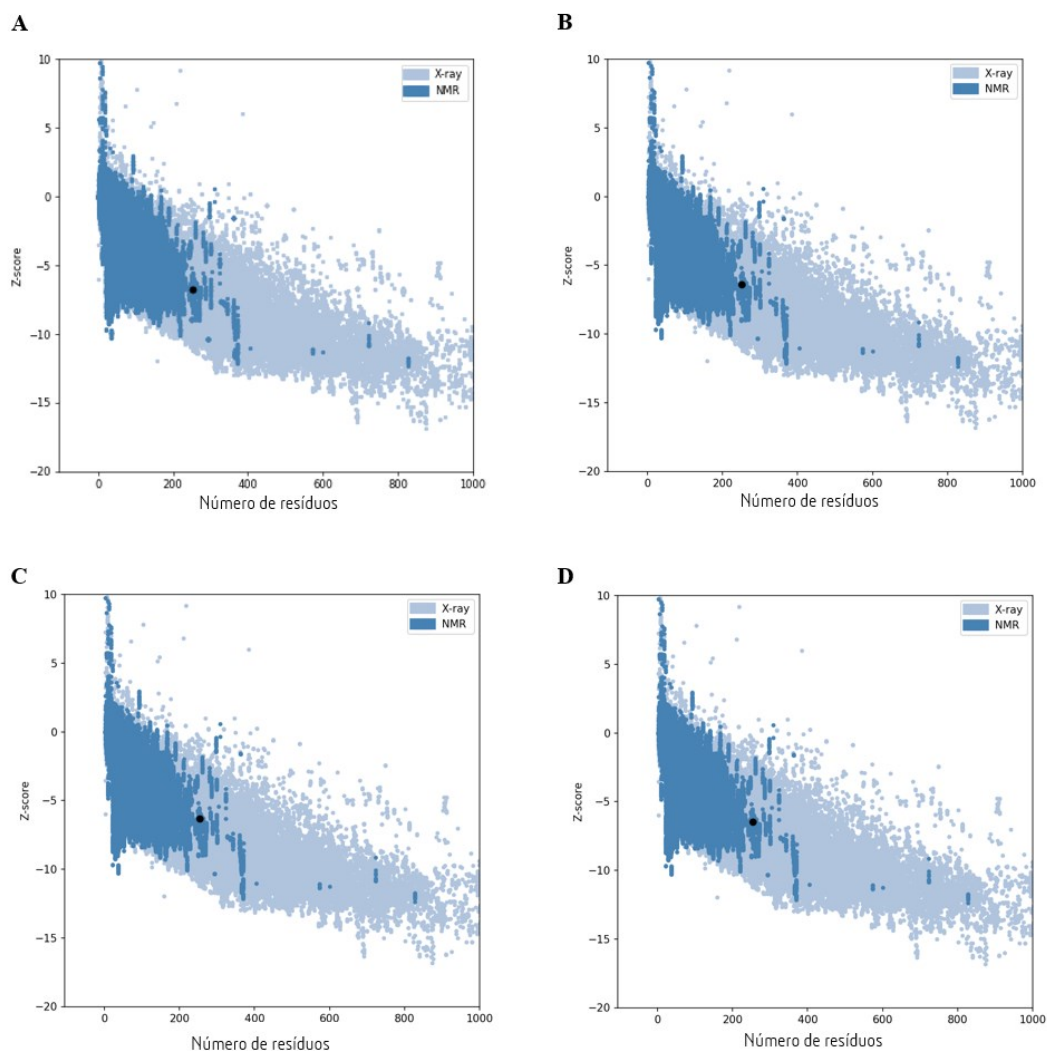
4.5.2 ProSA

O ProSA é uma ferramenta que verifica possíveis erros nas estruturas experimentais e teóricas, por meio do princípio de Boltzmann. Este princípio utiliza potenciais de força média

extraídos de estruturas conhecidas dos bancos de dados e assim realiza a predição dos potenciais da estrutura proteica desejada, conseguindo identificar mal dobramentos. Dessa maneira, o servidor utiliza a pontuação *z-score* para indicar a qualidade geral do modelo e mede o desvio da energia total da estrutura em relação a uma distribuição de energia derivada de conformações aleatórias. Esse valor é exibido em um gráfico que contém os *z-score* de todas as cadeias de proteínas determinadas experimentalmente por cristalografia de raio X e ressonância magnética nuclear (NMR) existentes nos bancos de dados (PRAJAPAT, 2014; SIPPL, 1993; WIEDERSTEIN, 2007).

A Figura 14 apresenta em forma de gráfico os *Z-score* obtidos dos modelos em relação ao número de resíduos. De modo geral, todas as estruturas modeladas possuem pontuação de alta qualidade, visto que os valores de *Z-score* estão dentro da faixa de pontuação normalmente encontrada para proteínas nativas de tamanho semelhante.

Figura 14: Gráfico do Z-score dos modelos atômicos

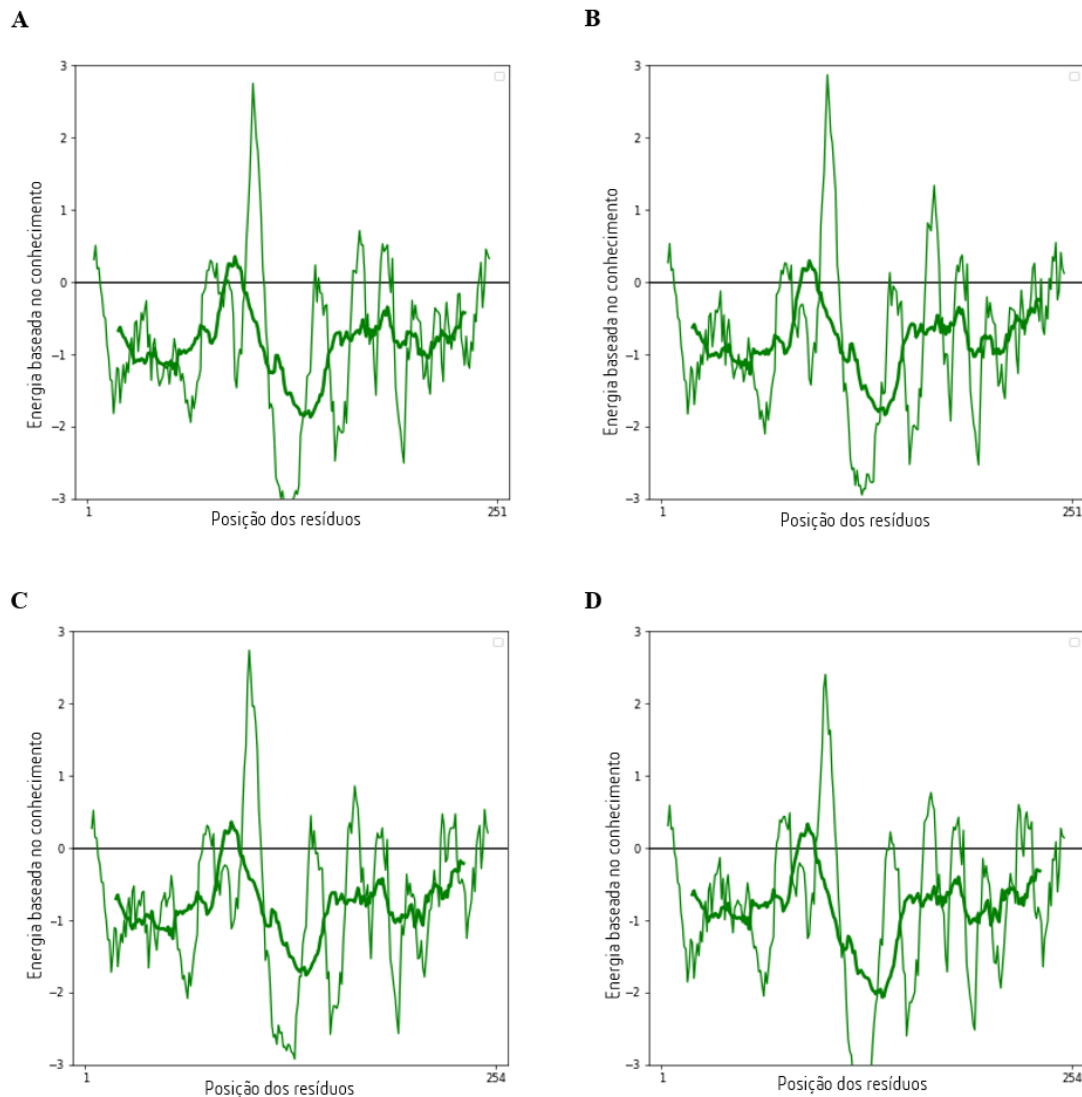


Legenda: Gráfico do Z-score esperado para modelos atômicos de alta qualidade de cristalografia de raio X e ressonância magnética nuclear com base no número de resíduos do (A) scFv 2 com *linker* original, (B) scFv 2 com *linker* modificado, (C) scFv quimérico com *linker* original e (D) scFv quimérico com *linker* modificado. 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.

O gráfico de energia livre avalia a qualidade local e é suavizado calculando energia média sobre cada fragmento de 40 e 10 resíduos, uma vez que um gráfico de energias de resíduos individuais geralmente contém grandes flutuações (PRAJAPAT, 2014; SIPPL, 1993; WIEDERSTEIN, 2007). Com isso, a Figura 15 demonstra o gráfico de energia livre em função da posição do aminoácido em fragmentos de 10 e 40 resíduos dos modelos atômicos. Os valores positivos correspondem a partes problemáticas da estrutura, portanto, como a maioria dos

resíduos obtiveram resultados negativos, os fragmentos de anticorpos modelados alcançaram bons resultados.

Figura 15: Gráfico de energia livre dos modelos atômicos



Legenda: Gráfico de energia livre em relação a posição de aminoácidos em fragmentos de 10 (em verde claro) e 40 (em verde escuro) resíduos do (A) scFv 2 com *linker* original, (B) scFv 2 com *linker* modificado, (C) scFv quimérico com *linker* original e (D) scFv quimérico com *linker* modificado.

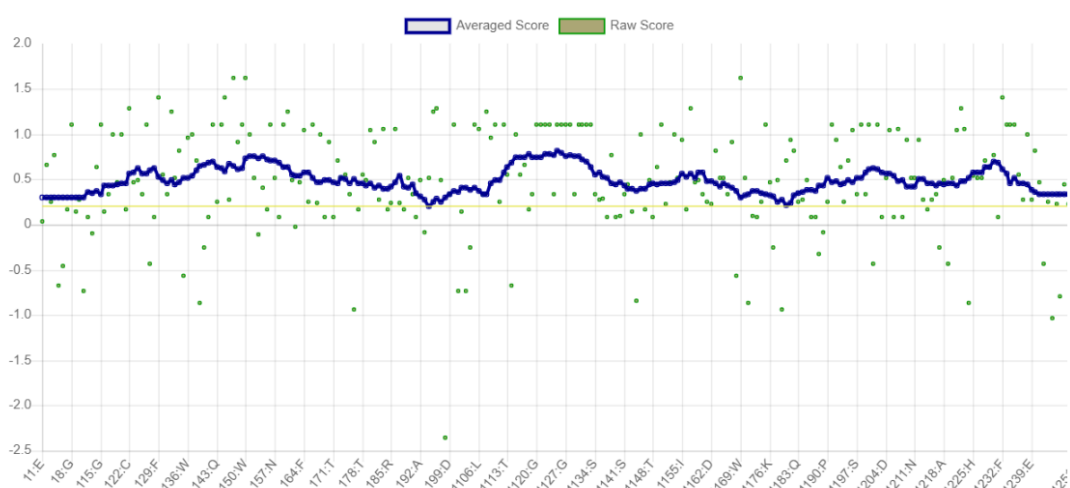
4.5.3 Verify 3D

O Verify 3D analisa a compatibilidade da estrutura tridimensional da proteína com sua sequência de aminoácidos. Com isso, o programa verifica cada posição dos resíduos em seu ambiente em uma fileira de 20 aminoácidos, de forma que a pontuação, chamada de 3D-1D, é

a preferência de cada resíduo dos 20 para esse ambiente. O ambiente é definido por três parâmetros, sendo eles: a área do resíduo inacessível, a fração da área da cadeia lateral que é coberta por átomos polares (como oxigênio e nitrogênio) e a estrutura secundária local (BOWIE, 1991; LÜTHY, 1992).

Dessa maneira, a análise no Verify 3D do scFv 2 com *linker* original está representada na Figura 16. Dessa forma, o modelo atômico exibiu uma alta qualidade, visto que apresentou 100% dos resíduos com pontuação 3D-1D $\geq 0,2$.

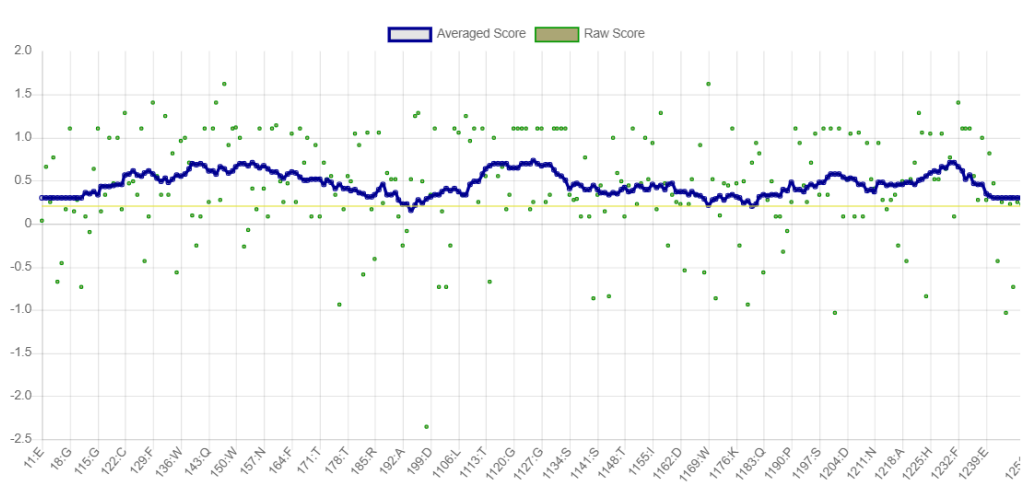
Figura 16 – Análise de compatibilidade 3D-1D do scFv 2 com *linker* original



Legenda: O segmento azul representa o resultado médio e os pontos verdes representam o resultado bruto da pontuação 3D-1D em função dos resíduos. Modelos com alta qualidade devem apresentar pelo menos 80% dos resíduos com pontuação 3D-1D $\geq 0,2$.

Na Figura 17 está representado a análise de compatibilidade 3D-1D do scFv 2 com *linker* modificado. Dessa forma, o modelo atômico exibiu uma alta qualidade, visto que apresentou 99,6% dos resíduos com pontuação 3D-1D $\geq 0,2$.

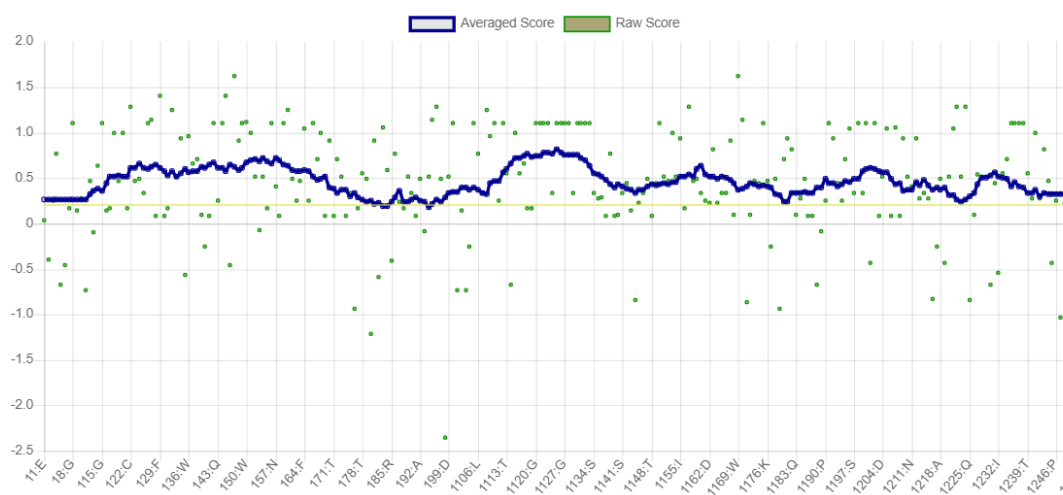
Figura 17 - Análise de compatibilidade 3D-1D do scFv 2 com *linker* modificado



Legenda: O segmento azul representa o resultado médio e os pontos verdes representam o resultado bruto da pontuação 3D-1D em função dos resíduos. Modelos com alta qualidade devem apresentar pelo menos 80% dos resíduos com pontuação 3D-1D $\geq 0,2$.

Na Figura 18 está representado a análise de compatibilidade 3D-1D do scFv quimérico com *linker* original. Dessa forma, o modelo atômico exibiu uma alta qualidade, visto que apresentou 98,82% dos resíduos com pontuação 3D-1D $\geq 0,2$.

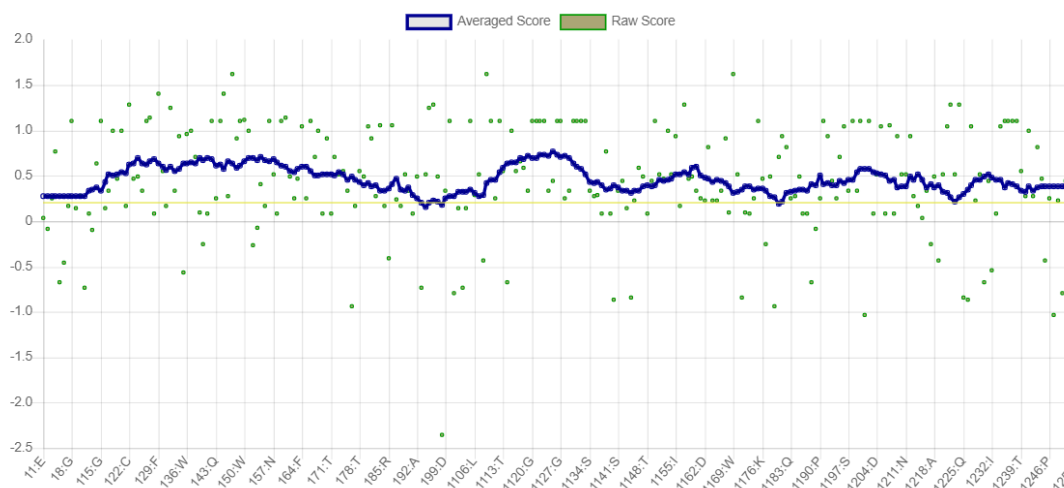
Figura 18 - Análise de compatibilidade 3D-1D do scFv quimérico com *linker* original



Legenda: O segmento azul representa o resultado médio e os pontos verdes representam o resultado bruto da pontuação 3D-1D em função dos resíduos. Modelos com alta qualidade devem apresentar pelo menos 80% dos resíduos com pontuação 3D-1D $\geq 0,2$.

Na Figura 19 está representado a análise de compatibilidade 3D-1D do scFv quimérico com *linker* modificado. Dessa forma, o modelo atômico exibiu uma alta qualidade, visto que apresentou 98,82% dos resíduos com pontuação 3D-1D $\geq 0,2$.

Figura 19 - Análise de compatibilidade 3D-1D do scFv quimérico com *linker* modificado



Legenda: O segmento azul representa o resultado médio e os pontos verdes representam o resultado bruto da pontuação 3D-1D em função dos resíduos. Modelos com alta qualidade devem apresentar pelo menos 80% dos resíduos com pontuação 3D-1D $\geq 0,2$.

4.5.4 PROCHECK

A ferramenta PROCHECK analisa e avalia qualidade estereoquímica da estrutura proteica disponibilizando gráficos que fornecem detalhes sobre os ângulos diedros, comprimentos e ângulos de ligação da cadeia principal. Os resultados fornecidos pelo PROCHECK advêm da comparação com estruturas de alta qualidade disponibilizadas nos bancos de dados (DYM, 2006; LASKOWSKI, 1993).

O gráfico de Ramachandran é gerado no PROCHECK e tem o intuito de analisar e identificar erros nos ângulos diedros, phi (ϕ), psi (ψ). Esses ângulos de torção estão localizados respectivamente nas ligações N-C α e C α -C da cadeia principal dos aminoácidos, de forma que suas orientações obedecem aos limites impostos pela estereoquímica. Sendo assim, o gráfico de Ramachandran possui regiões favoráveis e não favoráveis de acordo com as combinações

possíveis dos ângulos ϕ e ψ (LASKOWSKI, 1993; MORRIS, 1992; RAMACHANDRAN, 1963; RAMAKRISHNAN, 1965).

A Tabela 4 representa a porcentagem de resíduos nas regiões mais favoráveis dos scFv modelados, de modo que os modelos foram comparados com 118 estruturas com pelo menos 2 Å de resolução e fator R de no máximo 20%. Embora, os modelos atômicos não alcançaram no mínimo 90% dos resíduos nas regiões mais favoráveis, todos obtiveram valores próximo a 90%, o que pode indicar uma qualidade satisfatória para esses modelos.

Tabela 4 - Pontuação dos modelos atômicos com base no gráfico de Ramachandran disponibilizado no PROCHECK

Modelo	Score
scFv 2 com linker original	87,3%
scFv 2 com linker modificado	88,7%
scFv quimérico com linker original	86,9%
scFv quimérico com linker modificado	87,4%

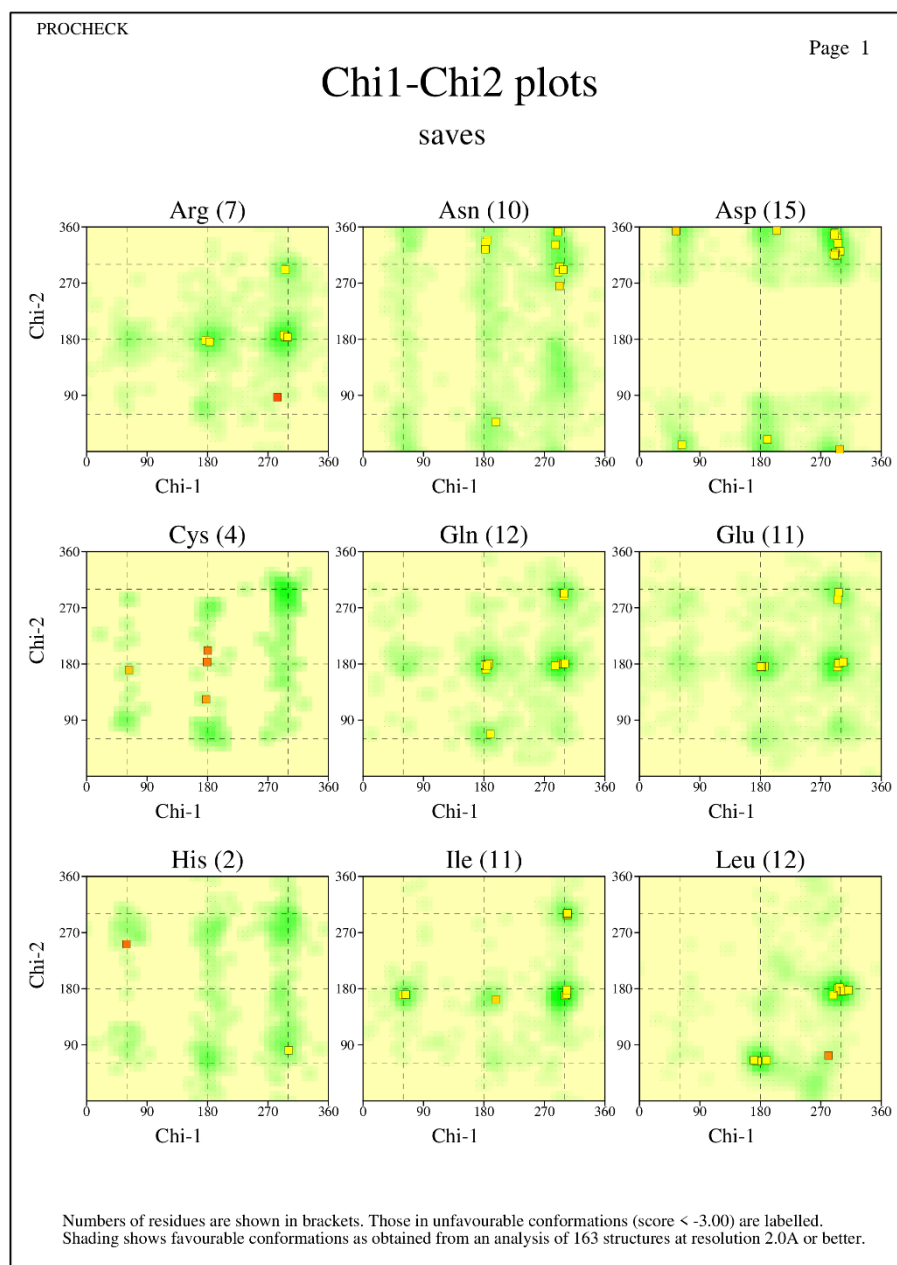
Legenda: Modelos com alta qualidade devem apresentar pelo menos 90% dos resíduos nas regiões mais favoráveis de acordo com as combinações possíveis dos ângulos phi (ϕ) e psi (ψ).

Assim como os ângulos da cadeia principal, ϕ e ψ , influenciam na conformação da proteína, os ângulos de torção da cadeia lateral dos aminoácidos, chi (χ), também possuem um papel importante nessa conformação, devido às forças de atração e repulsão presentes nos átomos. Esses ângulos dependem da quantidade de átomos existente na cadeia lateral do aminoácido, com isso o PROCHECK considera somente os ângulos chi-1 (χ_1) e chi-2 (χ_2) para gerar os gráficos (BORGUESAN, 2016; COWELL, 2004).

As Figuras 20 a 23 demonstram os gráficos dos ângulos χ_1 e χ_2 das cadeias laterais dos aminoácidos do scFv 2 com *linker* original, scFv 2 com *linker* modificado, scFv quimérico com *linker* original e scFv quimérico com *linker* modificado, respectivamente. Sendo assim, os modelos foram comparados com 163 estruturas com resolução de 2 Å ou mais, o que permitiu

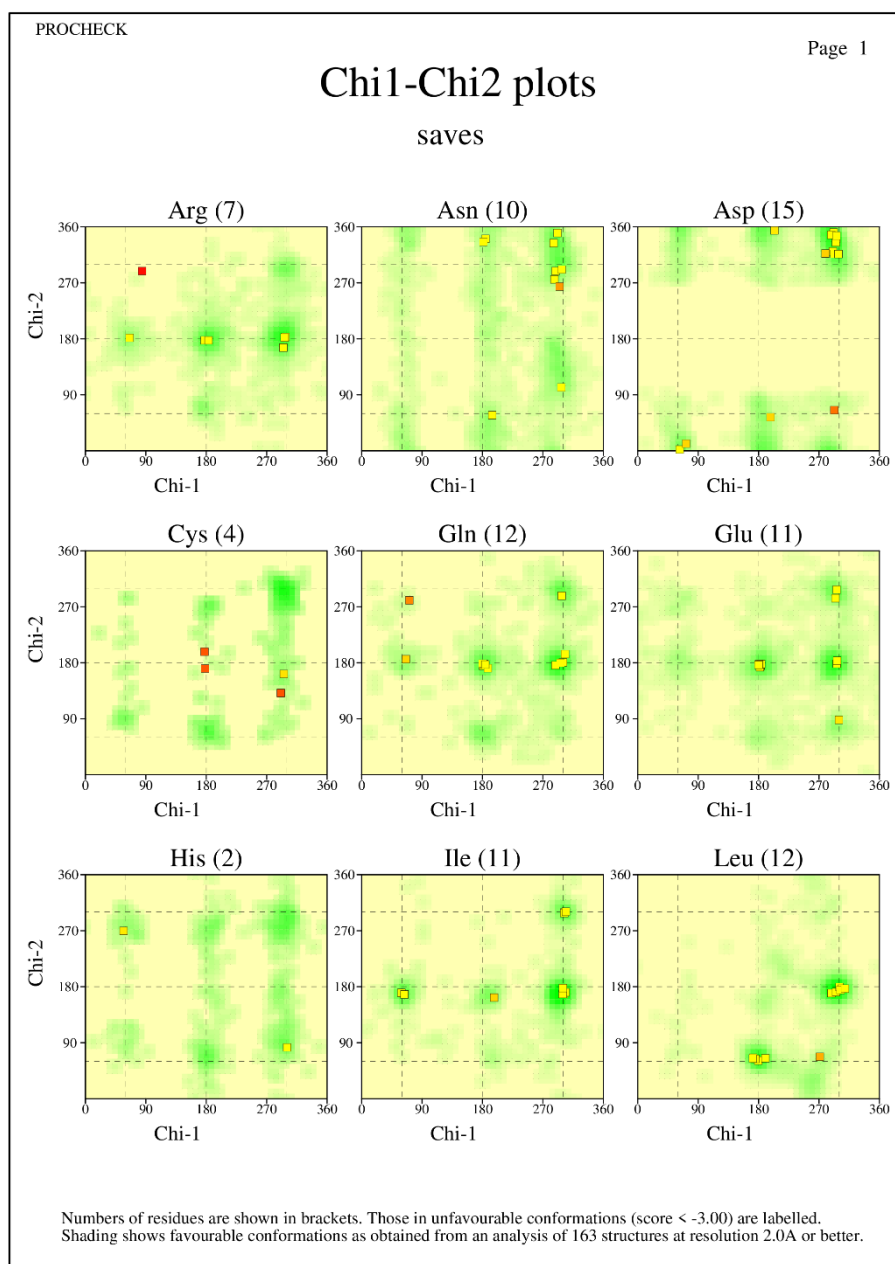
identificar quais resíduos estão em uma conformação desfavorável ($score < -3,00$). Logo, nenhum resíduo da cadeia lateral dos modelos modelados apresentaram uma conformação desfavorável energeticamente.

Figura 20 – Gráfico dos ângulos Chi1-Chi2 da cadeia lateral do scFv 2 com *linker* original



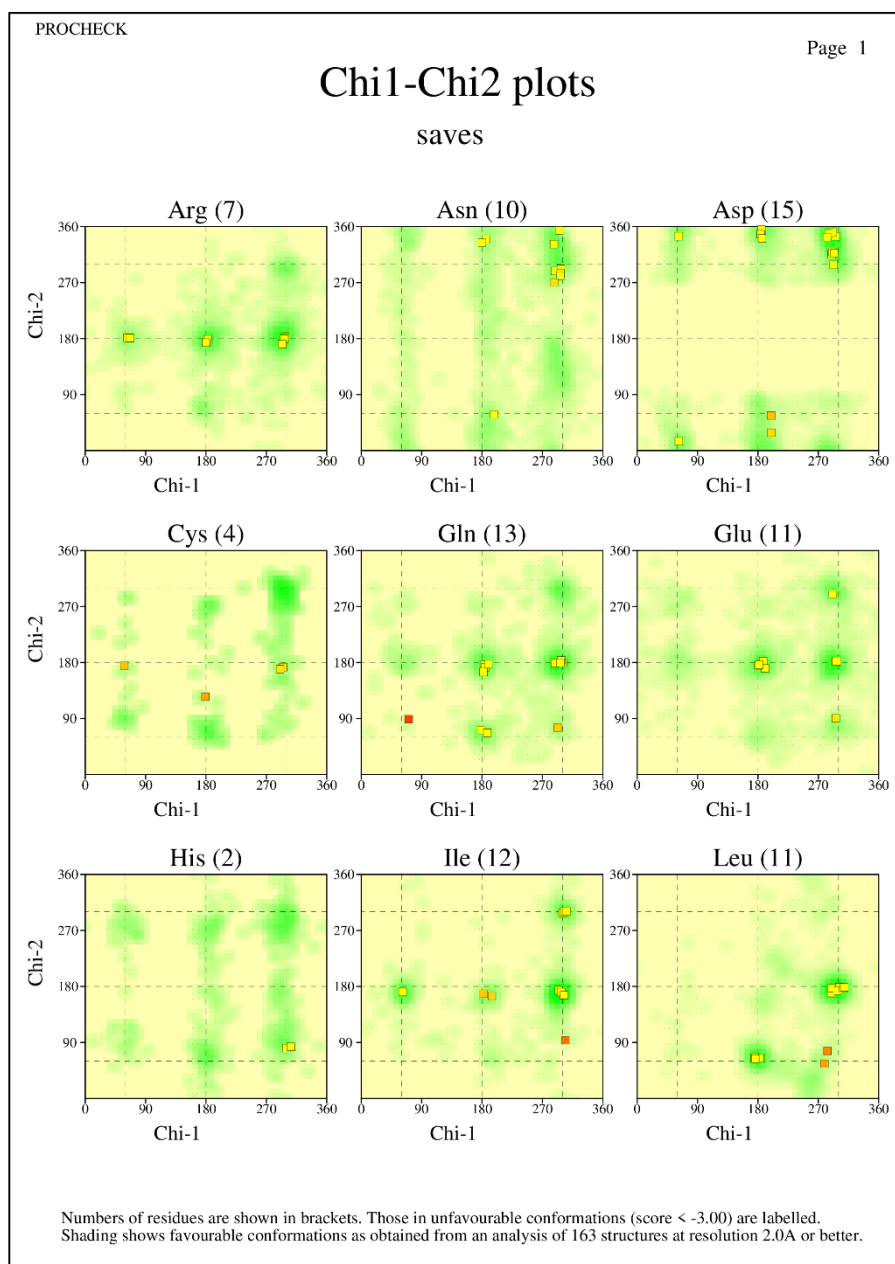
Legenda: Ângulos Chi1-Chi2 da cadeia lateral por tipos de resíduos: Alanina (Ala), Arginina (Arg), Asparagina (Asn), Ácido Aspártico (Asp), Cisteína (Cys), Glutamina (Gln), Ácido Glutâmico (Glu), Glicina (Gly) e Histidina (His) do modelo atômico. A quantidade de resíduos do mesmo tipo é mostrada em parênteses. Os resíduos destacados e identificados indicam uma pontuação menor que -3,00, indicando uma conformação desfavorável energeticamente.

Figura 21 - Gráfico dos ângulos Chi1-Chi2 da cadeia lateral do scFv 2 com *linker* modificado



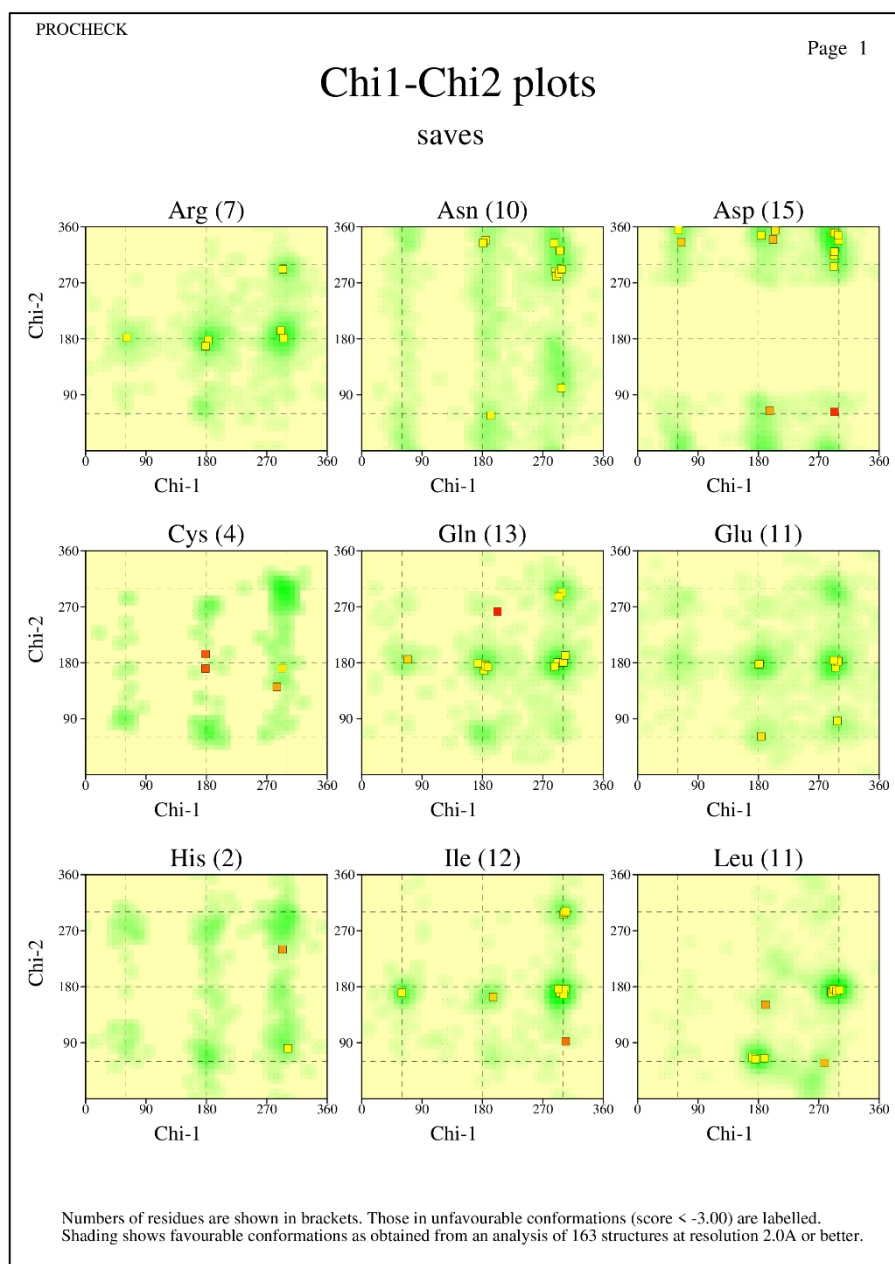
Legenda: Ângulos Chi1-Chi2 da cadeia lateral por tipos de resíduos: Alanina (Ala), Arginina (Arg), Asparagina (Asn), Ácido Aspártico (Asp), Cisteína (Cys), Glutamina (Gln), Ácido Glutâmico (Glu), Glicina (Gly) e Histidina (His) do modelo atômico. A quantidade de resíduos do mesmo tipo é mostrada em parênteses. Os resíduos destacados e identificados indicam uma pontuação menor que -3,00, indicando uma conformação desfavorável energeticamente.

Figura 22 - Gráfico dos ângulos Chi1-Chi2 da cadeia lateral do scFv quimérico com *linker* original



Legenda: Ângulos Chi1-Chi2 da cadeia lateral por tipos de resíduos: Alanina (Ala), Arginina (Arg), Asparagina (Asn), Ácido Aspártico (Asp), Cisteína (Cys), Glutamina (Gln), Ácido Glutâmico (Glu), Glicina (Gly) e Histidina (His) do modelo atômico. A quantidade de resíduos do mesmo tipo é mostrada em parênteses. Os resíduos destacados e identificados indicam uma pontuação menor que -3,00, indicando uma conformação desfavorável energeticamente.

Figura 23 - Gráfico dos ângulos Chi1-Chi2 da cadeia lateral do scFv quimérico com *linker* modificado



Legenda: Ângulos Chi1-Chi2 da cadeia lateral por tipos de resíduos: Alanina (Ala), Arginina (Arg), Asparagina (Asn), Ácido Aspártico (Asp), Cisteína (Cys), Glutamina (Gln), Ácido Glutâmico (Glu), Glicina (Gly) e Histidina (His) do modelo atômico. A quantidade de resíduos do mesmo tipo é mostrada em parênteses. Os resíduos destacados e identificados indicam uma pontuação menor que -3,00, indicando uma conformação desfavorável energeticamente.

Outros gráficos disponibilizados pelo PROCHECK são em relação aos parâmetros da cadeia principal, sendo eles: qualidade do gráfico do Ramachandran, desvio dos ângulos ômega (ω), contatos ruins por 100 resíduos, desvio do ângulo zeta (ζ), desvio de energia de ligações de hidrogênio (H) e fator-G, sendo este um indicador estrutural de proteínas que utiliza parâmetros estereoquímicos. Portanto, esses valores indicam como a estrutura analisada se

compara com outras estruturas refinadas em uma resolução semelhante (LASKOWSKI, 1993; YASUDA, 2022).

Em vista disso, a Tabela 5 apresenta os parâmetros da cadeia principal de todos os modelos atômicos modelados. Em relação a porcentagem de resíduos nas regiões mais favoráveis, ao desvio do ângulo ω , aos contatos ruins por 100 resíduos e ao desvio de energia de ligações de H, todos os modelos ficaram dentro dos parâmetros de alta qualidade. Enquanto, ao desvio do ângulo zeta (ζ) e fator-G, os modelos obtiveram melhor valor dentro o esperado para modelos de alta qualidade.

Tabela 5 - Parâmetros da cadeia principal dos modelos atômicos analisados no PROCHECK

Modelo	Porcentagem de resíduos nas regiões mais favoráveis	Desvio do ângulo ω	Contatos ruins por 100 resíduos	Desvio do ângulo zeta	Desvio de energia de ligações de hidrogênio	Fator-G geral
scFv 2 com linker original	Dentro	Dentro	Dentro	Melhor	Dentro	Melhor
scFv 2 com linker modificado	Dentro	Dentro	Dentro	Melhor	Dentro	Melhor
scFv quimérico com linker original	Dentro	Dentro	Dentro	Melhor	Dentro	Melhor
scFv quimérico com linker modificado	Dentro	Dentro	Dentro	Melhor	Dentro	Melhor

Os ângulos da cadeia lateral podem assumir três tipos de conformações de baixa energia, gauche negativo: -60° ; gauche positivo: $+60^\circ$ e trans $\pm 180^\circ$ (HRUBY, 2000). Diante disso, PROCHECK disponibiliza gráficos dos parâmetros da cadeia lateral, sendo eles: desvio do χ_1 gauche negativo (g^-), desvio do χ_1 trans, desvio do χ_1 gauche positivo (g^+), desvio dos ângulos χ_1 agrupados e desvio padrão dos ângulos Chi-2 (χ_2) trans.

A Tabela 6 exibe os parâmetros da cadeia lateral de todos os fragmentos de anticorpos modelados. No que se refere aos desvios dos ângulos χ_1^- , χ_1^+ , χ_1^- agrupados e χ_2^+ , todos os modelos apresentaram qualidades melhores que o esperado de estruturas com alta qualidade.

Tabela 6 - Parâmetros da cadeia lateral dos modelos atômicos analisados no PROCHECK

Modelo	Desvio do Gauche negativo Chi-1	Desvio do Trans Chi-1	Desvio do Gauche positivo Chi-1	Desvio dos Agrupados Chi-1	Desvio do Trans Chi-2
scFv 2 com linker original	Melhor	Melhor	Melhor	Melhor	Melhor
scFv 2 com linker modificado	Melhor	Melhor	Melhor	Melhor	Melhor
scFv quimérico com linker original	Melhor	Melhor	Melhor	Melhor	Melhor
scFv quimérico com linker modificado	Melhor	Melhor	Melhor	Melhor	Melhor

4.6 Docking molecular entre a proteína E do ZIKV e os scFv modelados

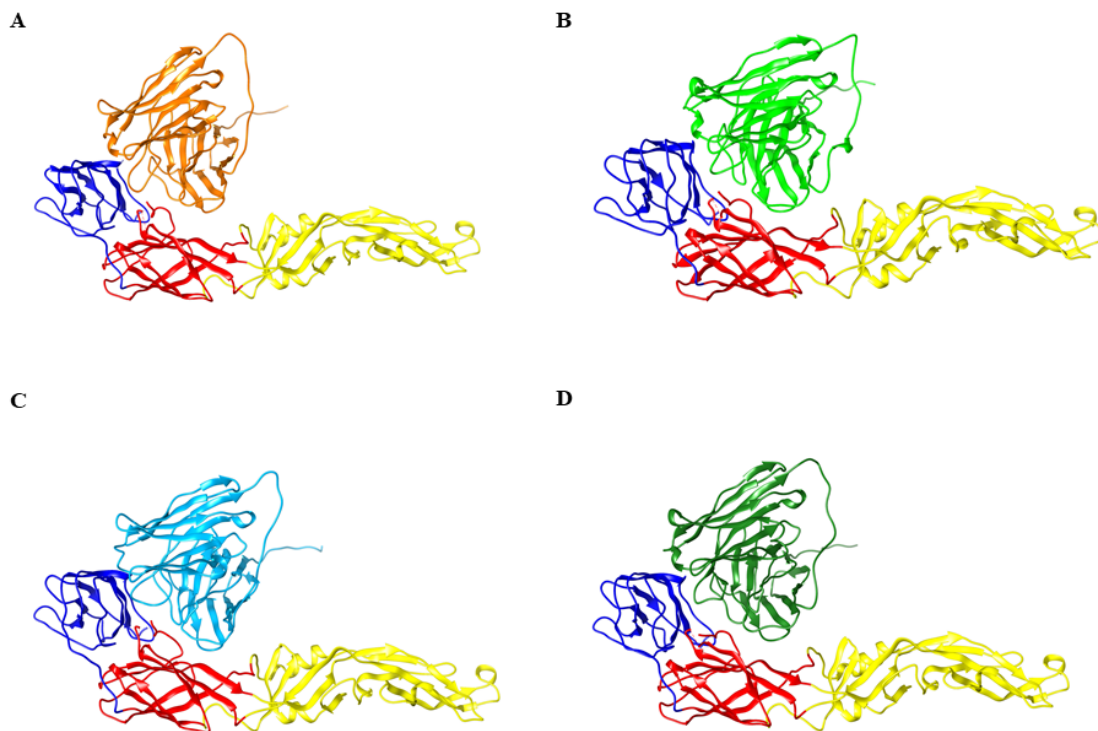
Após a validação dos modelos atômicos foi realizado a ancoragem molecular com a proteína E do ZIKV, disponibilizada no Protein Data Bank (PDB) (ID: 5JHM). Com isso, o Cluspro disponibiliza os dez melhores *clusters*, de forma que, para cada ancoragem, foram selecionados os que obtiveram os melhores *scores* (BRENKE, 2012; NEMATLI, 2021), como representado na Tabela 7.

Tabela 7 – Scores do *docking* molecular entre a proteína E e os modelos atômicos

Complexos com a proteína E	Amostra	Score
scFv 2 com <i>linker</i> original	Centro	-515.2
	Menor Energia	-658.6
scFv 2 com <i>linker</i> modificado	Centro	-510.1
	Menor Energia	-643.1
scFv quimérico com <i>linker</i> original	Centro	-508.6
	Menor Energia	-539.3
scFv quimérico com <i>linker</i> modificado	Centro	-485.8
	Menor Energia	-599.3

Conforme a Figura 24, todos os fragmentos de anticorpos modelados realizaram a interação entre os domínios I e III da proteína E do ZIKV. Em vista disso, o DIII possui grande importância no desenvolvimento de medicamentos contra o ZIKV, uma vez que esse domínio interage com os receptores celulares presentes na célula hospedeira. Dessa forma, o DIII é dos principais alvos de anticorpos neutralizantes contra o ZIKV, pois sua inibição impede a internalização e, conseqüentemente, a proliferação do vírus no hospedeiro. Além disso, esse domínio consiste na região menos conservada entre os flavivírus, sendo assim anticorpos que reagem com essa região possuem menos chances de promover uma reatividade cruzada, diferentemente de anticorpos contra o DII que são menos neutralizantes, devido ao DII ser mais conservado (GALLICHOTTE, 2019; SHARMA, 2020; YU, 2017).

Figura 24 - Docking molecular entre os fragmentos de anticorpos e a proteína E de ZIKV



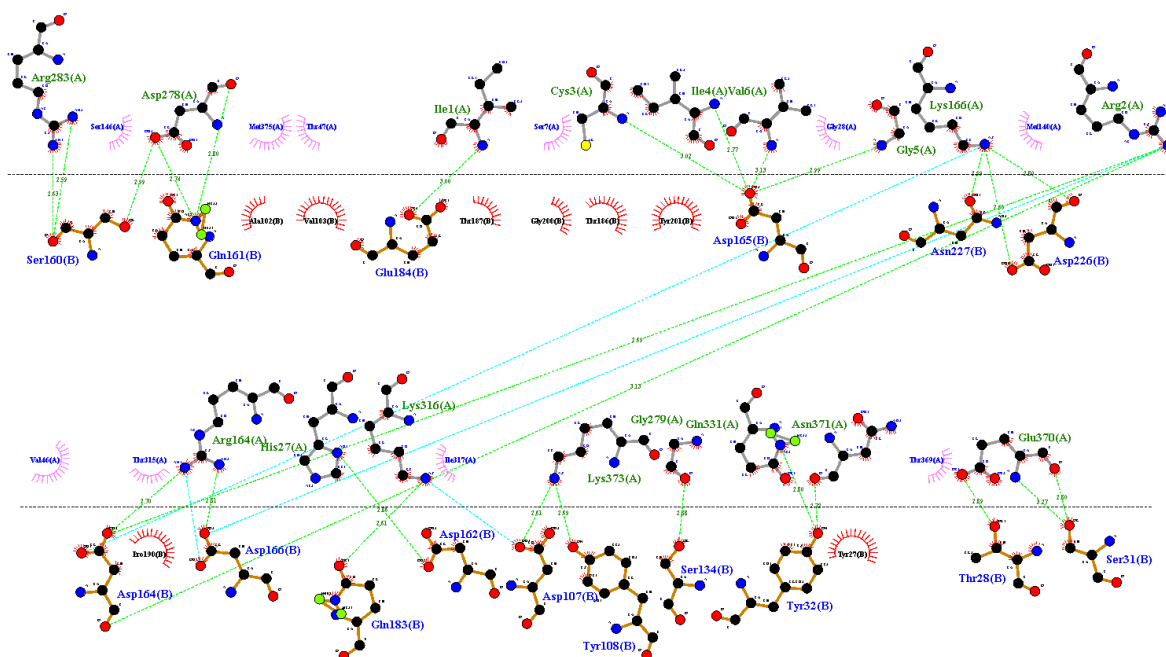
Legenda: Ancoragem molecular entre a proteína E do ZIKV (DI em vermelho, DII em amarelo e DIII em azul) com (A) scFv 2 com linker original (em laranja); (B) scFv 2 com linker modificado (em verde); (C) scFv quimérico com linker original (em azul claro); e (D) scFv quimérico com linker original (em verde escuro).

4.7 Análise 2D das interações intermoleculares entre a proteína E do ZIKV e os scFv modelados

A partir dos arquivos selecionados no ClusPro foi realizada a análise 2D das interações intermoleculares, por meio da ferramenta DIMPLOT disponível no LigPlot. Esse programa permite identificar quais resíduos participam de ligações de hidrogênio e interações hidrofóbicas, de forma que as distâncias (em Å) das ligações de hidrogênio estão expostas na interface. Sendo assim, a partir do LigPlot é possível comparar as interações entre diferentes ligantes para um mesmo alvo (WALLACE, 1995).

A Figura 25 demonstra as interações entre o scFv 2 com *linker* original (cadeia B) e a proteína E do ZIKV (cadeia A) e a Figura 26 as interações entre o scFv 2 com *linker* modificado

Figura 26 - Interações intermoleculares entre o scFv 2 com *linker* modificado e a proteína E do ZIKV



Legenda: Na parte superior do pontilhado estão descritos os resíduos da cadeia A, a qual representa a proteína E do ZIKV e na parte inferior são os resíduos da cadeia B, a qual representa o scFv 2 com *linker* modificado. As linhas tracejadas em verdes representam as ligações de hidrogênio, em azul claro as pontes salinas e em rosa e vermelho as interações hidrofóbicas.

Tabela 8 – Resíduos da proteína E do ZIKV que realizam ligações de hidrogênio com resíduos do scFv 2 com *linker* original

Resíduos da Proteína E do ZIKV	Domínio da proteína E	Resíduos do scFv 2 com <i>linker</i> original
Arg283	DI	Asp162
Lys166	DI	Tyr201
Ile1	DI	Glu184
Arg164	DI	Asp166
Arg2	DI	Asp165
Lys316	DIII	Asp107
Lys373	DIII	Try32 Asp107 Tyr108
Thr369	DIII	Asn54
Glu367	DIII	Asn54
Ile317	DIII	Glu1
Thr315	DIII	Tyr108
Phe314	DIII	Tyr27
Thr313	DIII	Tyr27
Asp371	DIII	Tyr32
Glu370	DIII	Ser31 Thr28 Thr30

Legenda: Arginina (Arg), Asparagina (Asn), Aspartato (Asp), Cisteína (Cys), Fenilalanina (Phe), Glicina (Gly), Glutamato (Glu), Isoleucina (Ile), Lisina (Lys), Serina (Ser), Tirosina (Tyr), Treonina (Thr). Os números correspondem a posição do aminoácido na proteína. Domínio I (DI). Domínio III (DIII).

Tabela 9 – Resíduos da proteína E do ZIKV que realizam ligações de hidrogênio com resíduos do scFv 2 com *linker* modificado

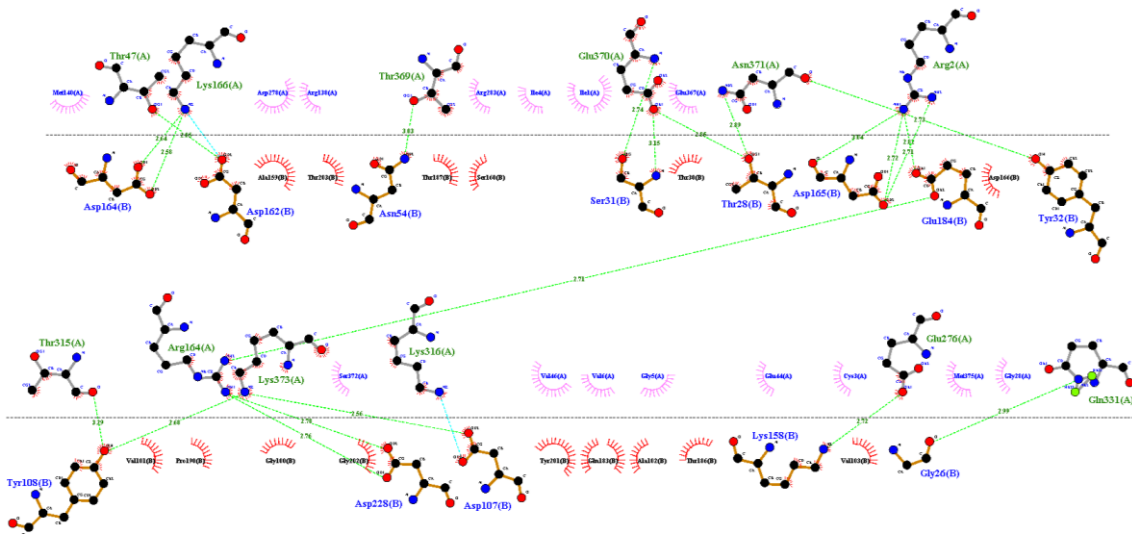
Resíduos da Proteína E ZIKV	Domínio da proteína E	Resíduos do scFv 2 com <i>linker</i> modificado
Arg283	DI	Ser160 *
Ile1	DI	Glu184
Cys3	DI	Asp165
Ile4	DI	Asp165
Val6	DI	Asp165
Gly5	DI	Asp165
Lys166	DI	Asn227 Asp226 *
Arg2	DI	Asp164 *
Arg164	DI	Asp164 Asp166
His27	DI	Asp162
Asp278	DII	Ser160 Gln161*
Gly279	DII	Ser134
Gln331	DIII	Tyr32
Asn371	DIII	Tyr32
Glu370	DIII	Ser31 Thr28

Legenda: Arginina (Arg), Asparagina (Asn), Aspartato (Asp), Cisteína (Cys), Fenilalanina (Phe), Glicina (Gly), Glutamato (Glu), Glutamina (Gln), Isoleucina (Ile), Lisina (Lys), Serina (Ser), Tirosina (Tyr), Treonina (Thr), Valina (Val). Os números correspondem a posição do aminoácido na proteína. Os aminoácidos destacados com asterisco (*) realizaram mais de uma ligação de hidrogênio. Domínio I (DI). Domínio II (DII). Domínio III (DIII).

A Figura 27 demonstra as interações entre o scFv quimérico com *linker* original (cadeia B) e a proteína E do ZIKV (cadeia A) e a Figura 28 as interações entre o scFv quimérico com *linker* modificado (cadeia B) e a proteína E do ZIKV (cadeia A), de forma que estão destacados os resíduos de interação que formam ligações de hidrogênio e interações hidrofóbicas.

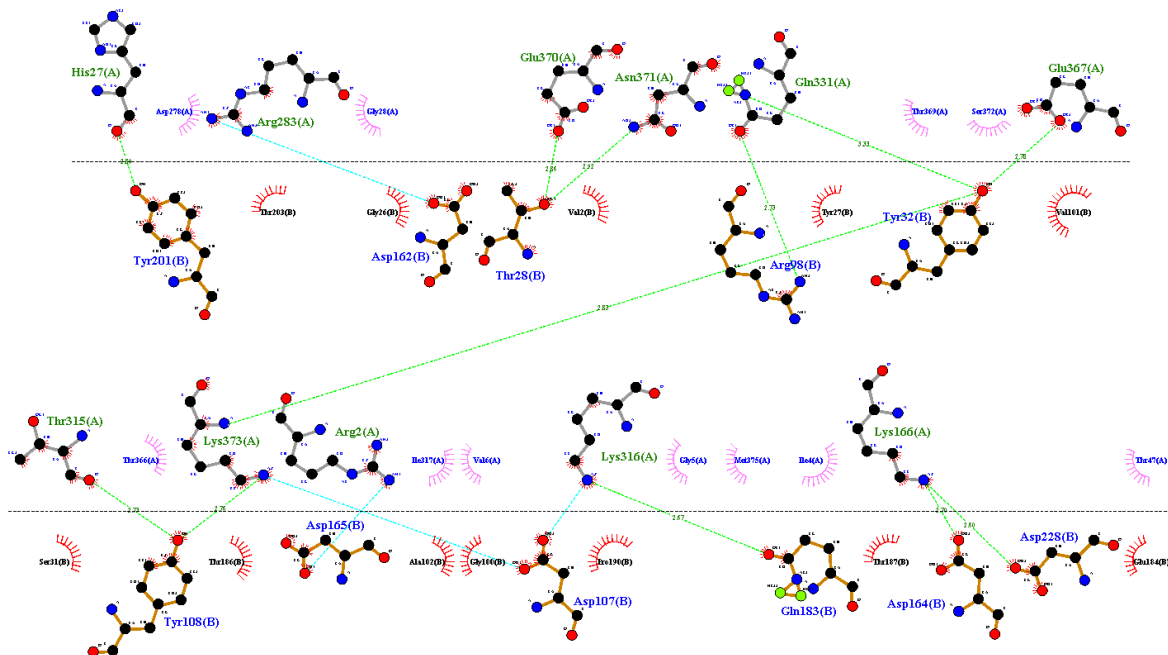
As Tabelas 10 e 11 apresentam os resíduos dos scFv 2 com *linker* original e modificado que realizam ligações de hidrogênio com a proteína E do ZIKV. Com isso, comparando as interações expostas nas Figuras 27 e 28 e os resíduos nas Tabelas 10 e 11, verifica uma alteração no perfil das ligações de hidrogênio, de forma que o scFv quimérico com *linker* original apresentou uma maior quantidade de ligações de hidrogênio.

Figura 27 - Interações intermoleculares entre o scFv quimérico com linker original e a proteína E do ZIKV



Legenda: Na parte superior do pontilhado estão descritos os resíduos da cadeia A, a qual representa a proteína E do ZIKV e na parte inferior são os resíduos da cadeia B, a qual representa o scFv quimérico com linker original. As linhas tracejadas em verdes representam as ligações de hidrogênio, em azul claro as pontes salinas e em rosa e vermelho as interações hidrofóbicas.

Figura 28 - Interações intermoleculares entre o scFv quimérico com linker modificado e a proteína E do ZIKV



Legenda: Na parte superior do pontilhado estão descritos os resíduos da cadeia A, a qual representa a proteína E do ZIKV e na parte inferior são os resíduos da cadeia B, a qual representa o scFv quimérico com linker modificado. As linhas tracejadas em verdes representam as ligações de hidrogênio, em azul claro as pontes salinas e em rosa e vermelho as interações hidrofóbicas.

Tabela 10 – Resíduos da proteína E do ZIKV que realizam ligações de hidrogênio com resíduos do scFv quimérico com *linker* original

Resíduos da Proteína E ZIKV	Domínio da proteína E	Resíduos do scFv quimérico com <i>linker</i> original
Thr47	DI	Asp162
Lys166	DI	Asp164 *
Arg2	DI	Asp165 * Glu184
Arg164	DI	Asp228 * Glu184
Glu276	DII	Lys158
Gln331	DIII	Gly26
Thr315	DIII	Tyr108
Lys373	DIII	Tyr108 Asp107
Thr369	DIII	Asn54
Glu370	DIII	Ser31* Thr28
Asn371	DIII	Thr28 Tyr32

Legenda: Arginina (Arg), Asparagina (Asn), Aspartato (Asp), Cisteína (Cys), Fenilalanina (Phe), Glicina (Gly), Glutamato (Glu), Glutamina (Gln), Isoleucina (Ile), Lisina (Lys), Serina (Ser), Tirosina (Tyr), Treonina (Thr), Valina (Val). Os números correspondem a posição do aminoácido na proteína. Os aminoácidos destacados com asterisco (*) realizam mais de duas ligações de hidrogênios. Domínio I (DI). Domínio II (DII). Domínio III (DIII).

Tabela 11 – Resíduos da proteína E do ZIKV que realizam ligações de hidrogênio com resíduos do scFv quimérico com *linker* modificado

Resíduos da Proteína E ZIKV	Domínio da proteína E	Resíduos do scFv quimérico com <i>linker</i> modificado
His27	DI	Tyr201
Lys166	DI	Asp228 Asp164
Asn371	DIII	Thr28
Gln331	DIII	Arg98 Tyr32
Glu367	DIII	Tyr32
Glu370	DIII	Thr28
Lys316	DIII	Gln183
Lys373	DIII	Tyr108 Tyr32
Thr315	DIII	Tyr108

Legenda: Arginina (Arg), Asparagina (Asn), Aspartato (Asp), Glutamato (Glu), Glutamina (Gln), Histidina (His), Lisina (Lys), Tirosina (Tyr), Treonina (Thr). Os números correspondem a posição do aminoácido na proteína. Domínio I (DI). Domínio III (DIII).

Como descrito no Tópico 4.3 e comprovado pelas análises 2D das interações, a ligação dos fragmentos de anticorpos se limitaram principalmente nas regiões do DI e DIII da proteína E. Com isso, a interação nesse sítio é vantajosa, visto que impede a inserção do vírus na membrana do hospedeiro e, conseqüentemente, bloqueia a trimerização da proteína E (SHARMA, 2020).

Analisando todas as interfaces de interações, esperava-se um acréscimo de ligações de hidrogênio para o scFv quimérico com *linker* modificado, visto que o scFv 2 com mesmo *linker* apresentou um aumento nessas ligações. Portanto, não se pode afirmar que esse acréscimo advém da adição das prolinas, sendo assim necessário uma maior apuração de estudos para compreender esse comportamento molecular.

A substituição do CDR 3 VL do scFv 2 pelo CDR 3 VL do scFv 4 tem o intuito de verificar se houve modificações na ligação do scFv quimérico com epítomos da proteína E. Embora, o scFv 2 com *linker* original apresente mais ligações de hidrogênio quando comparado com o scFv quimérico com *linker* original, não houve mudanças substanciais nos valores das ancoragens. Dessa forma, outras interações, as quais não foram avaliadas nesta monografia, podem estar auxiliando na interação entre o fragmento de anticorpo e a proteína E do ZIKV.

5. CONCLUSÃO

Após a apresentação das análises realizadas, pode-se concluir que:

- A substituição de duas glicinas por duas prolinas do *linker* dos scFv não foram suficientes para promover mudanças conformacionais na estrutura do anticorpo exceto pela região do *linker*, de maneira que não houve grandes alterações nos valores de RMSD.
- A modelagem realizada no RoseTTAFold apresentou qualidade satisfatória em todos os métodos de validação realizados;
- As ancoragens moleculares realizadas na proteína E do ZIKV se limitaram principalmente nos domínios I e III;
- Houve diferenças nas interações moleculares, de forma que o scFv 2 com *linker* modificado apresentou mais ligações de hidrogênio, porém o mesmo não foi observado no scFv quimérico com mesmo *linker*. Então, são necessárias mais pesquisas para compreender se tal comportamento advém do *linker* modificado.
- O scFv 2 com *linker* original apresentou mais ligações de hidrogênio quando comparado com o scFv quimérico com *linker* original, porém não houve mudanças substanciais nos valores das ancoragens.

6. REFERÊNCIAS BIBLIOGRÁFICAS

ABBAS, A. K., LICHTMAN, A. H., PILLAI, S. **Imunologia Celular e Molecular**. 8ª Edição. Saunders Elsevier, 2015.

AGRELLI, A., et al. ZIKA Virus Entry Mechanisms in Human Cells. **Infection, Genetics and Evolution**, vol. 69, p. 22–29, abril de 2019.

ALENCAR, W. L. M. **Interações de ftalocianinas de Co, Cu e não metálicas com estruturas externas de SARS COV-2 utilizando Docking e dinâmica molecular**. 2021. 86 f. Tese (Doutorado em Engenharia de Recursos Naturais da Amazônia) - Instituto de Tecnologia, Universidade Federal do Pará, Belém, 2021

BAEK, M., et al. Accurate Prediction of Protein Structures and Interactions Using a Three-Track Neural Network. **Science**, vol. 373, no 6557, p. 871–76, agosto de 2021.

BAEK, M., et al. Deep Learning and Protein Structure Modeling. **Nature Methods**, vol. 19, no 1, p. 13–14, janeiro de 2022.

BANDEHPOUR, Mojgan, et al. Bioinformatics Evaluation of Novel Ribosome Display-Selected Single Chain Variable Fragment (ScFv) Structure with Factor H Binding Protein through Docking. **Journal of Theoretical and Computational Chemistry**, vol. 16, nº 03, p. 175002, Mai. de 2017.

BORGUESAN, B. **GARTS: um algoritmo genético baseado no método de seleção por torneio restrito adaptativo para o problema de predição de estruturas 3D de proteínas**. 2016. Dissertação (Mestrado em Ciência da Computação). Universidade Federal do Rio Grande do Sul - Instituto de Informática, Porto Alegre, 2016.

BOWIE, J. U., et al. A Method to Identify Protein Sequences That Fold into a Known Three-Dimensional Structure. **Science**, vol. 253, no 5016, p. 164–70, julho de 1991.

BRENKE, R., et al. Application of asymmetric statistical potentials to antibody-protein docking. **Bioinformatics**. vol 28, no 20, p. 2608-2614, outubro de 2012.

CARVALHO, L. R. Análise da presença do Zika vírus em pacientes sintomáticos e sem sintomas clínicos para arboviroses. Dissertação (Mestrado em Biociências). Unesp. São José do Rio Preto. 2019;

CASTRO, N. F. L. Desenvolvimento de novos anticorpos monoclonais humanos ANTI-CD3. 2019. 109 f., il. Tese (Doutorado em Patologia Molecular) - Universidade de Brasília, Brasília, 2019.

CHEN, V. B., et al. MolProbity: all-atom structure validation for macromolecular crystallography. **Acta Crystallographica Section D Biological Crystallography**, vol. 66, no 1, p. 12–21, janeiro de 2010.

CHEN, X., et al. Fusion Protein Linkers: Property, Design and Functionality. **Advanced Drug Delivery Reviews**, vol. 65, no 10, p. 1357–69, outubro de 2013.

COWELL, S. M., et al. Exploring Ramachandran and Chi Space: Conformationally Constrained Amino Acids and Peptides in the Design of Bioactive Polypeptide Ligands. **Current Medicinal Chemistry**, vol. 11, no 21, p. 2785–98, novembro de 2004.

DAI, L., et al. Structures of the Zika Virus Envelope Protein and Its Complex with a Flavivirus Broadly Protective Antibody. **Cell Host & Microbe**, vol. 19, n° 5, p. 696–704, Mai. de 2016.

DAVIS, I. W., et al. MolProbity: All-Atom Contacts and Structure Validation for Proteins and Nucleic Acids. **Nucleic Acids Research**, vol. 35, no Web Server, p. W375–83, maio de 2007.

DAVIS, I. W., et al. MolProbity: Structure Validation and All-Atom Contact Analysis for Nucleic Acids and Their Complexes. **Nucleic Acids Research**, vol. 32, p. W615–W19, Jul. de 2004.

DHURIGAI, N.; DANIEL, R. R.; AUXILIA, L. R. Structure determination of leghemoglobin using homology modeling. **Int. J. Curr. Microbiol. App. Sci**, vol 3, n° 10, p. 177-187, 2014.

DONKOR, E. S., et al. Bioinformatics with basic local alignment search tool (BLAST) and fast alignment (FASTA). **Journal of Bioinformatics and Sequence Analysis**, vol. 6, n° 1, p. 1–6, abril de 2014.

DYM, O., et al. Detection of Errors in Protein Models. **International Tables for Crystallography**. Volume F: Crystallography of Biological Macromolecules, p 520– 530, 2006.

EHRENMANN, F. IMGT/DomainGapAlign: The IMGT® Tool for the Analysis of IG, TR, MH, IgSF, and MhSF Domain Amino Acid Polymorphism. **Immunogenetics**, vol. 882, p. 605–33, 2012.

EHRENMANN, F., et al. IMGT/DomainGapAlign: IMGT Standardized Analysis of Amino Acid Sequences of Variable, Constant, and Groove Domains (IG, TR, MH, IgSF, MhSF). **Cold Spring Harbor Protocols**, vol. 2011, n° 6, p. 737-49, junho de 2011.

FERREIRA, F. L. **Caracterização funcional de proteínas hipotéticas de *Leptospira interrogans* como adesinas e potenciais antígenos vacinais e para diagnóstico**. 2021. Tese (Doutorado em Bioquímica) - Instituto de Química, Universidade de São Paulo, São Paulo, 2022.

FILHO, O. A. S. et al. Modelagem de Proteínas Por Homologia. **Química Nova**, vol. 26, no 2, p. 253–59, Mar. de 2003.

FRANÇA, R. K. A. O. **Desenvolvimento e caracterização de anticorpos humanos com potencial neutralizante do vírus Zika**. 2020. xv, 114 f., il. Dissertação (Mestrado em Biologia Molecular) - Universidade de Brasília, Brasília, 2020.

FREITAS, L. B. N. **Análise molecular da adaptação de insetos herbívoros aos cardenólídeos de calotropis procera: aspectos bioquímicos e ecológicos**. 2019. 75f. Dissertação (Mestrado em Bioquímica), Universidade Federal do Ceará, Fortaleza, 2019.

GALLICHOTTE, E. N., et al. Role of Zika Virus Envelope Protein Domain III as a Target of Human Neutralizing Antibodies. **MBio**, vol. 10, no 5, p. e01485-19, outubro de 2019.

GUBLER, D. J., et al. History and Emergence of Zika Virus. **The Journal of Infectious Diseases**, vol. 216, n° 10, p. S860–67, dezembro de 2017.

HEINZ, F. X., et al. The Antigenic Structure of Zika Virus and Its Relation to Other Flaviviruses: Implications for Infection and Immunoprophylaxis. **Microbiology and Molecular Biology Reviews**, vol. 81, n° 1, p. e00055-16, março de 2017.

HRUBY, V.J., et al. Conformational and topographical considerations in designing agonist peptidomimetics from peptide leads. **Curr Med Chem**, vol 7, no 9, p. 945-70, setembro de 2000.

JADAUN, J. S., et al. Chimeric Enzyme Designing for the Synthesis of Multifunctional Biocatalysts. **Biomass, Biofuels, Biochemicals, Elsevier**. p. 119–43, 2020.

JEANMOUGIN, F., et al. Multiple Sequence Alignment with Clustal X. **Trends in Biochemical Sciences**, vol. 23, n° 10, p. 403–05, outubro de 1998.

JOHNSON, M., et al. NCBI BLAST: A Better Web Interface. **Nucleic Acids Research**, vol. 36, n° Web Server, p. W5–9, maio de 2008.

JÚNIOR, N. N. **Diferenças estruturais e "docking" receptor-ligante da proteína E7 do vírus do papiloma humano (hpv) de alto e baixo riscos para o câncer cervical**. 2013. Tese (Doutorado em Genética). Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2013.

JÚNIOR, V. L. P., et al. Vírus Zika: revisão para clínicos. **Acta Med Port.** vol. 28, n° 6, p. 760–5, 2015.

KESARWANI, V., et al. Identification of Unique Peptides for SARS-CoV-2 Diagnostics and Vaccine Development by an In Silico Proteomics Approach. **Frontiers in Immunology**, vol. 12, p. 725240, setembro de 2021.

KOZAKOV, D., et al. The ClusPro Web Server for Protein–Protein Docking. **Nature Protocols**, vol. 12, n° 2, p. 255–278, fev. de 2017.

KUNIK, V., et al. Paratome: An Online Tool for Systematic Identification of Antigen-Binding Regions in Antibodies Based on Sequence or Structure. **Nucleic Acids Research**, vol. 40, n° W1, p. W521–24, julho de 2012.

KURODA, D., et al. Computer-Aided Antibody Design. **Protein Engineering Design and Selection**, vol. 25, n° 10, p. 507–22, outubro de 2012.

LARKIN, M. A., et al. Clustal W and Clustal X Version 2.0. **Bioinformatics**, vol. 23, no 21, p. 2947–48, novembro de 2007.

LASKOWSKI, R. A., et al. LigPlot+: Multiple Ligand–Protein Interaction Diagrams for Drug Discovery. **Journal of Chemical Information and Modeling**, vol. 51, n° 10, p. 2778–86, Out. de 2011.

LASKOWSKI, R. A., et al. PROCHECK: a program to check the stereochemical quality of protein structures. **Journal of Applied Crystallography**, vol. 26, n° 2, p. 283–91, abril de 1993.

LIBERIS, E., et al. Parapred: Antibody Paratope Prediction Using Convolutional and Recurrent Neural Networks. **Bioinformatics**, vol. 34, n° 17, p. 2944–50, setembro de 2018.

LÜTHY, R., et al. Assessment of Protein Models with Three-Dimensional Profiles. **Nature**, vol. 356, no 6364, p. 83–85, março de 1992.

MOHAMMADI, M., et al. Insilico Analysis of Three Different Tag Polypeptides with Dual Roles in ScFv Antibodies. **Journal of Theoretical Biology**, vol. 402, p. 100–06, agosto de 2016.

MORRIS, A. L., et al. Stereochemical Quality of Protein Structure Coordinates. **Proteins: Structure, Function, and Genetics**, vol. 12, no 4, p. 345–64, abril de 1992.

MUSSO, D., et al. Zika Virus. **Clinical Microbiology Reviews**, vol. 29, n° 3, p. 487–524, julho de 2016.

NEMATI, A. S., et al. Designing a new multi epitope-based vaccine against COVID-19 disease: an immunoinformatic study based on reverse vaccinology approach. In Review, 19 de fevereiro de 2021.

PACHECO, A. R. Respostas imunes induzidas por vacinas de DNA contra Zika e neuroadaptação de um isolado viral em camundongos para futuros testes de vacinas. 2021. 99

f. Dissertação (Mestrado em Biologia Celular e Molecular) - Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro, 2021.

PETERSEN, L. R., et al. Zika Virus. **New England Journal of Medicine**, vol. 374, n° 16, p. 1552–63, abril de 2016.

PETTERSEN, E. F., et al. UCSF Chimera - A Visualization System for Exploratory Research and Analysis. **Journal of Computational Chemistry**, vol. 25, no 13, p. 1605–12, outubro de 2004.

PINTO, I. A. **Identificação *in silico* de potenciais inibidores da via wnt canônica em adenocarcinoma mamário humano**. 2018. 63 f. Dissertação (Mestrado). Universidade Federal de Alfenas, Alfenas, 2018.

PLOURDE, A. R., et al. A Literature Review of Zika Virus. **Emerging Infectious Diseases**, vol. 22, n° 7, p. 1185–92, julho de 2016.

PRAJAPAT, R., et al. Recognition of Errors in the Refinement and Validation of Three-Dimensional Structures of AC1 Proteins of Begomovirus Strains by Using ProSA-Web. **Journal of Viruses**, vol. 2014, p. e752656, janeiro de 2014.

RAFIGHDOUST, H., et al. Bioinformatics prediction and experimental validation of VH antibody fragment interacting with Neisseria meningitidis factor H binding protein. **Iranian Journal of Basic Medical Sciences**, vol. 23, no 8, p. 1053–58, Ago. de 2020.

RAMACHANDRAN, G. N., et al. Stereochemistry of Polypeptide Chain Configurations. **Journal of Molecular Biology**, vol. 7, no 1, p. 95–99, julho de 1963.

RAMAKRISHNAN, C., et al. Stereochemical Criteria for Polypeptide and Protein Chain Conformations. **Biophysical Journal**, vol. 5, no 6, p. 909–33, novembro de 1965.

RAMHARACK, P., et al. Zika Virus NS5 Protein Potential Inhibitors: An Enhanced in Silico Approach in Drug Discovery. **Journal of Biomolecular Structure and Dynamics**, vol. 36, n° 5, p. 1118–33, abr. de 2018.

RATHORE, A. S., et al. designing Antibody against Highly Conserved Region of Dengue Envelope Protein by in Silico Screening of ScFv Mutant Library. **PLOS ONE**. vol. 14, no 1, p. e0209576, janeiro de 2019.

RODRIGUES, J. V. S. **Avaliação do potencial efeito antiviral de compostos 1,2,3-triazólicos derivados da vanilina contra o Zika vírus**. 2021. 81 f. Dissertação (Mestrado em Biologia Celular e Estrutural) - Universidade Federal de Viçosa, Viçosa. 2021.

ROY, A., et al. I-TASSER: A Unified Platform for Automated Protein Structure and Function Prediction. **Nature Protocols**, vol. 5, n° 4, p. 725–38, Abr. de 2010.

SAINI, S., et al. Affinity maturation of cross-reactive CR3022 antibody against the receptor binding domain of SARS-CoV-2 via in silico site-directed mutagenesis. preprint, In Review, 2 de abril de 2021.

SAIZ, J. C., et al. Zika Virus: What Have We Learnt Since the Start of the Recent Epidemic?. **Frontiers in Microbiology**, vol. 8, Agosto de 2017.

SANTOS, J. C. **Análise da tetramerização da proteína CD38 por meio de docking molecular**. 2019. Dissertação (Mestrado em Bioinformática), Universidade de São Paulo, São Paulo, 2019.

SEGER, J., et al. Decreased Production of TNF-Alpha by Lymph Node Cells Indicates Experimental Autoimmune Encephalomyelitis Remission in Lewis Rats. **Memórias Do Instituto Oswaldo Cruz**, p. 263, maio de 2010.

SHARMA, N., et al. Small Molecule Inhibitors Possibly Targeting the Rearrangement of Zika Virus Envelope Protein. **Antiviral Research**, vol. 182, p. 104876, outubro de 2020.

SILVA, C. S. **Desenvolvimento e Caracterização de Linhagem Celular Expressando Estavelmente um Replicon Repórter do Vírus Zika**. Dissertação. 2019. 69f. (Mestrado em Biociências e Biotecnologia em Saúde) – Instituto Aggeu Magalhães, Fundação Oswaldo Cruz, Recife, 2019.

SINIGAGLIA, A., et al. In Silico Approaches to Zika Virus Drug Discovery. **Expert Opinion on Drug Discovery**. vol. 13, nº 9, p. 825–835, set. de 2018.

SIPPL, M. J. Recognition of Errors in Three-Dimensional Structures of Proteins. **Proteins: Structure, Function, and Genetics**, vol. 17, no 4, p. 355–62, dezembro de 1993.

SIROHI, D., et. al. Zika Virus Structure, Maturation, and Receptors. **The Journal of Infectious Diseases**. vol. 216, nº10, p. S935–S944, Dez. de 2017.

SONG, B. H., et al. Zika Virus: History, Epidemiology, Transmission, and Clinical Presentation. **Journal of Neuroimmunology**, vol. 308, p. 50–64, julho de 2017.

SOUZA, G. R. L. **Construção de uma biblioteca de anticorpos (scFv) apresentada em fagos para seleção, análise e caracterização de antígenos do carrapato bovino (Boophilus microplus)**. 2007. 97 f. Tese (Doutorado em Ciências Biológicas) - Universidade Federal de Uberlândia, Uberlândia, 2007.

SPAHR, C., et al. High-Resolution Mass Spectrometry Confirms the Presence of a Hydroxyproline (Hyp) Post-Translational Modification in the GGGGP Linker of an Fc-Fusion Protein. **MAbs**, vol. 9, no 5, p. 812–19, julho de 2017.

TANG, Y., et al. Selection of Linkers for a Catalytic Single-Chain Antibody Using Phage Display Technology. **Journal of Biological Chemistry**, vol. 271, no 26, p. 15682–86, junho de 1996.

VIEIRA, G. C. **Modelagem molecular e imunodeteção de DNA Metiltransferases 2 de Drosophilídeos: uma abordagem evolutiva da enigmática DNMT2**. 2015. 63 f. Dissertação (Mestrado). Universidade Federal do Rio Grande do Sul, Porto Alegre, 2018.

WALLACE, A. C., et al. LIGPLOT: A Program to Generate Schematic Diagrams of Protein-Ligand Interactions. **Protein Engineering, Design and Selection**, vol. 8, no 2, p. 127–34, Fev. de 1995.

WIEDERSTEIN, M., et al. ProSA-Web: Interactive Web Service for the Recognition of Errors in Three-Dimensional Structures of Proteins. **Nucleic Acids Research**, vol. 35, n° Web Server, p. W407–W10, Mai. de 2007.

WIKAN, N., et al. Zika virus: history of a newly emerging arbovirus. **The Lancet Infectious Diseases**, v. 16, n° 7, p. e119-e126, 2016.

WILLIAMS, C. J., et al. MolProbity: More and better reference data for improved all-atom structure validation. **Protein Science: A Publication of the Protein Society**, vol. 27, no 1, p. 293–315, janeiro de 2018.

YAMASHITA, T. Toward Rational Antibody Design: Recent Advancements in Molecular Dynamics Simulations. **International Immunology**, vol. 30, n° 4, p. 133–40, abril de 2018.

YANG, H. Y., et al. Construction of a Large Synthetic Human ScFv Library with Six Diversified CDRs and High Functional Diversity. **Molecules and Cells**, vol. 27, no 2, p. 225–35, fevereiro de 2009.

YASUDA, T., et al. Structural Validation by the G -Factor Properly Regulates Boost Potentials Imposed in Conformational Sampling of Proteins. **Journal of Chemical Information and Modeling**, vol. 62, no 14, p. 3442–52, julho de 2022.

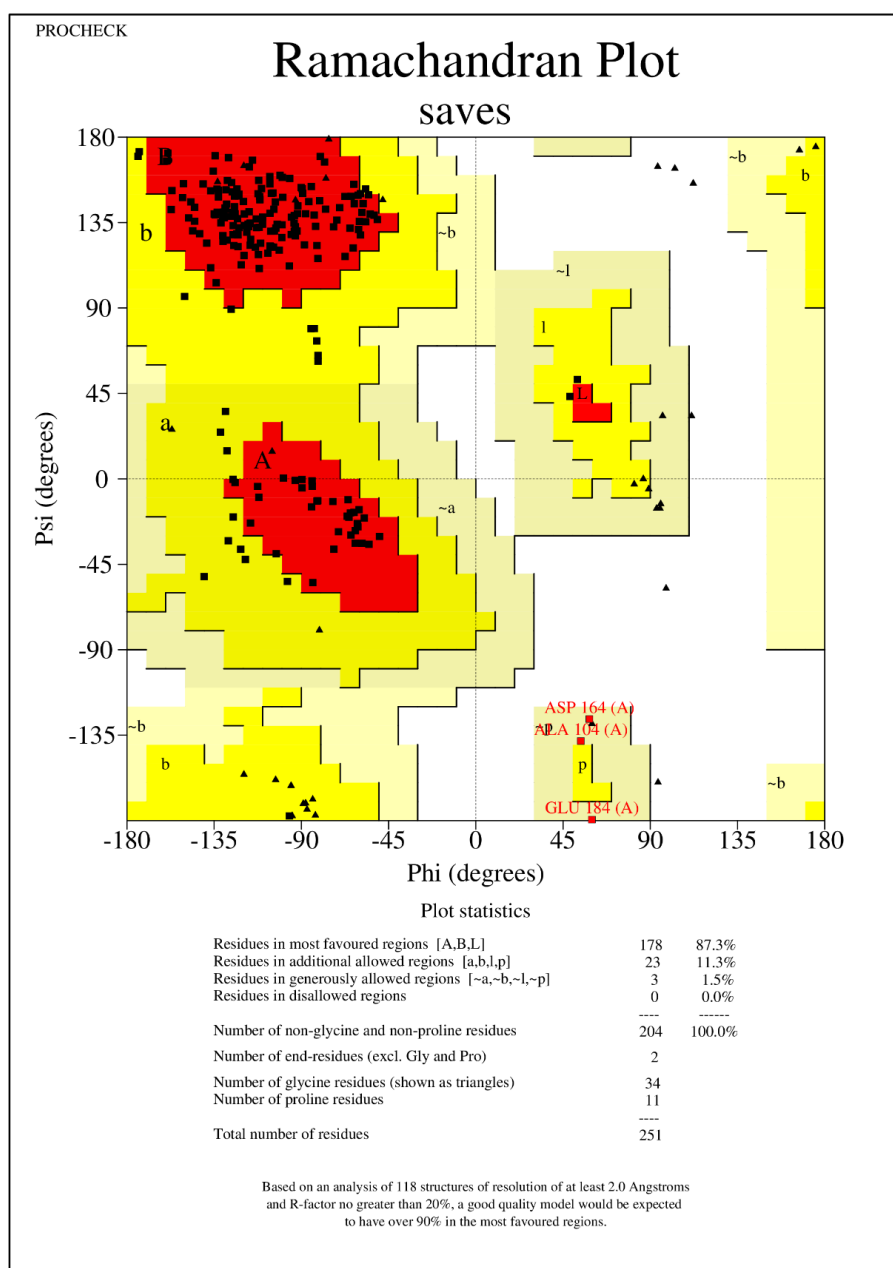
YU, L., et al. Delineating antibody recognition against Zika virus during natural infection. **JCI Insight**, vol. 2, no 12, p. e93042, junho de 2017.

YUAN, X., et al. Ab Initio Protein Structure Prediction Using Pathway Models. **Comparative and Functional Genomics**, vol. 4, n° 4, p. 397–401, Jul. de 2003.

ZHAO, H., et al. Structural Basis of Zika Virus-Specific Antibody Protection. **Cell**, vol. 166, n° 4, p. 1016–27, agosto de 2016.

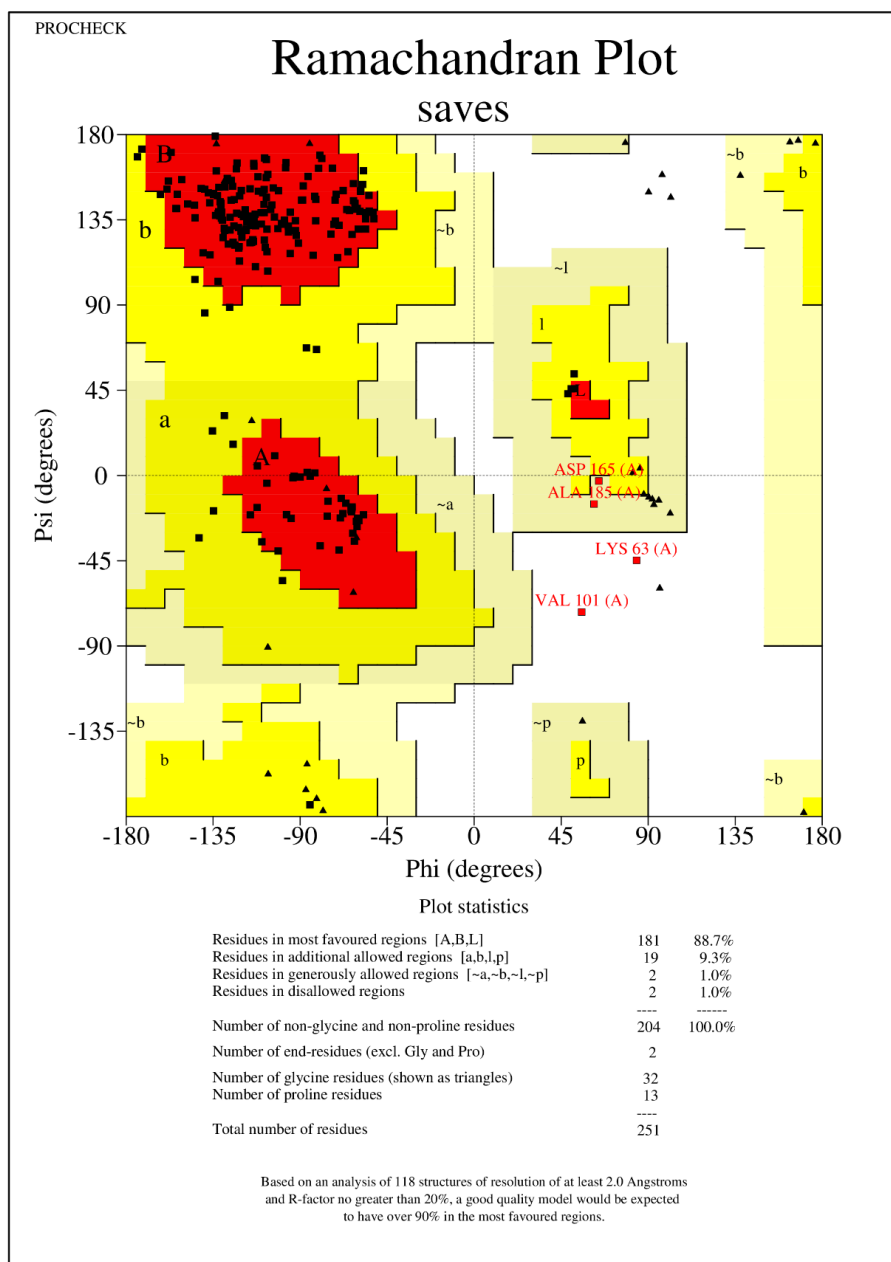
ANEXO

Figura 29 - Gráfico de Ramachandran do scFv 2 com linker original



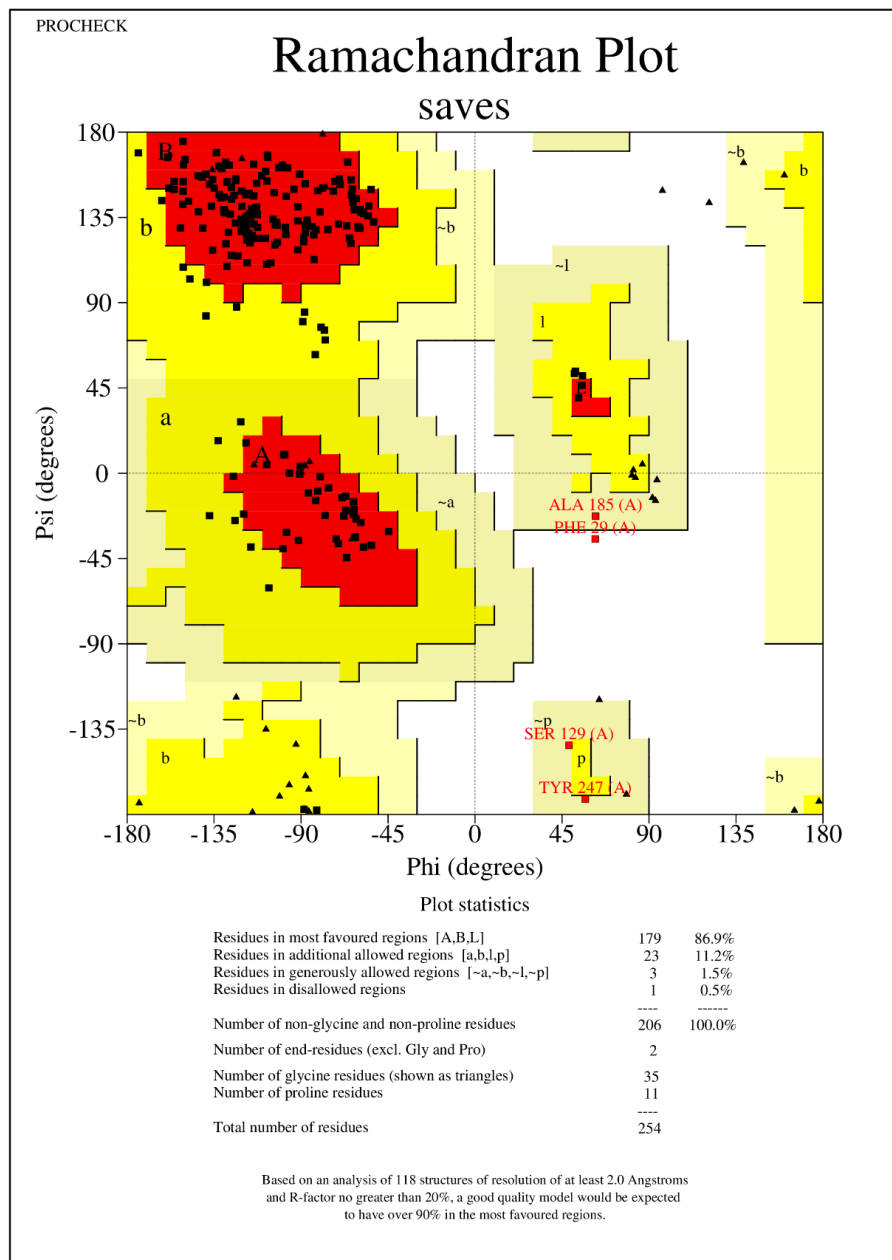
Legenda: As áreas destacadas em vermelho representam as combinações dos ângulos diedros ϕ (phi) e ψ (psi) mais favoráveis, em amarelo as regiões favoráveis, em amarelo claro as regiões menos favoráveis e em branco as regiões desfavoráveis. Os pontos pretos indicam os resíduos do modelo em regiões favoráveis e os vermelhos em regiões desfavoráveis.

Figura 30 - Gráfico de Ramachandran do scFv 2 com *linker* modificado



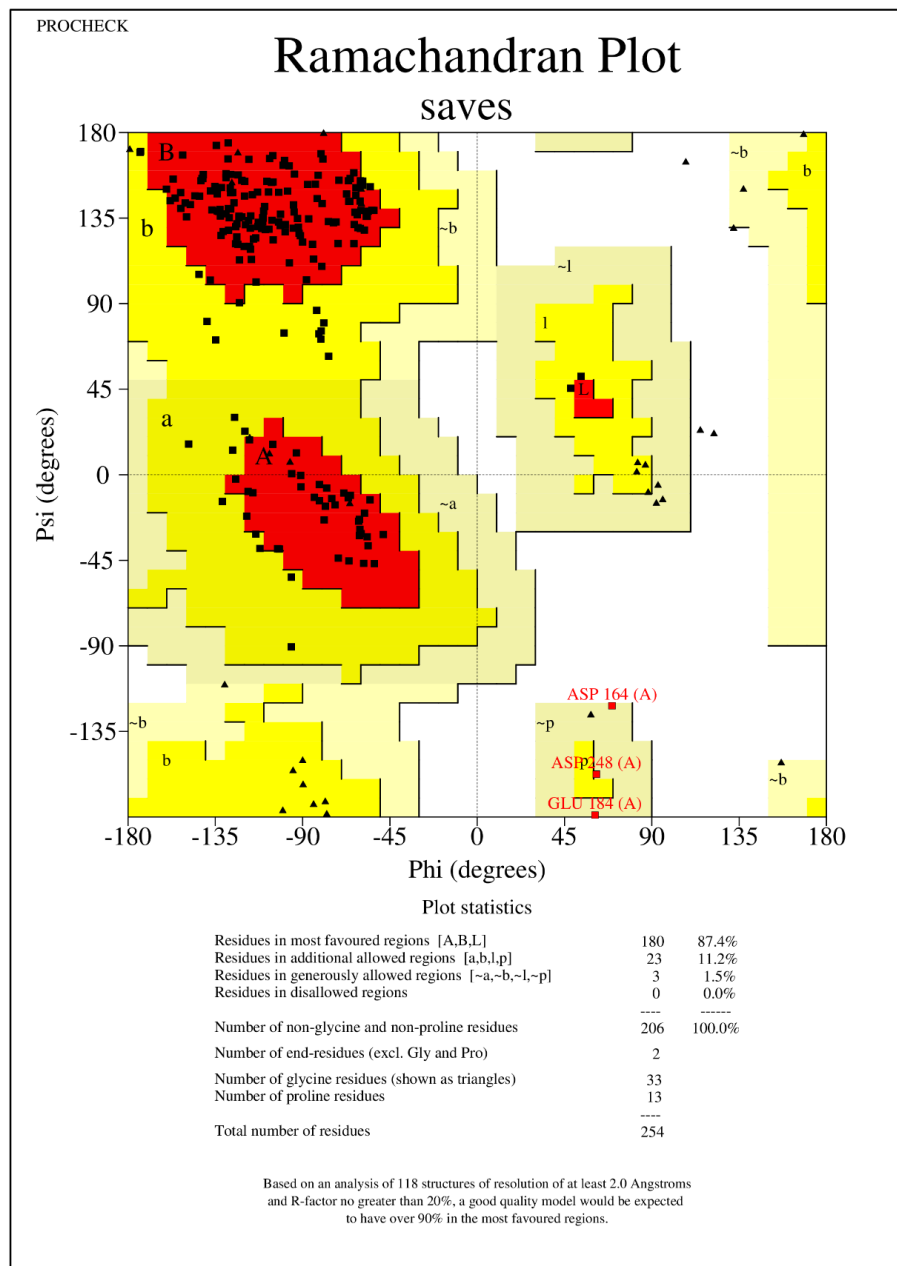
Legenda: As áreas destacadas em vermelho representam as combinações dos ângulos diedros ϕ (phi) e ψ (psi) mais favoráveis, em amarelo as regiões favoráveis, em amarelo claro as regiões menos favoráveis e em branco as regiões desfavoráveis. Os pontos pretos indicam os resíduos do modelo em regiões favoráveis e os vermelhos em regiões desfavoráveis.

Figura 31 - Gráfico de Ramachandran do scFv quimérico com *linker* original



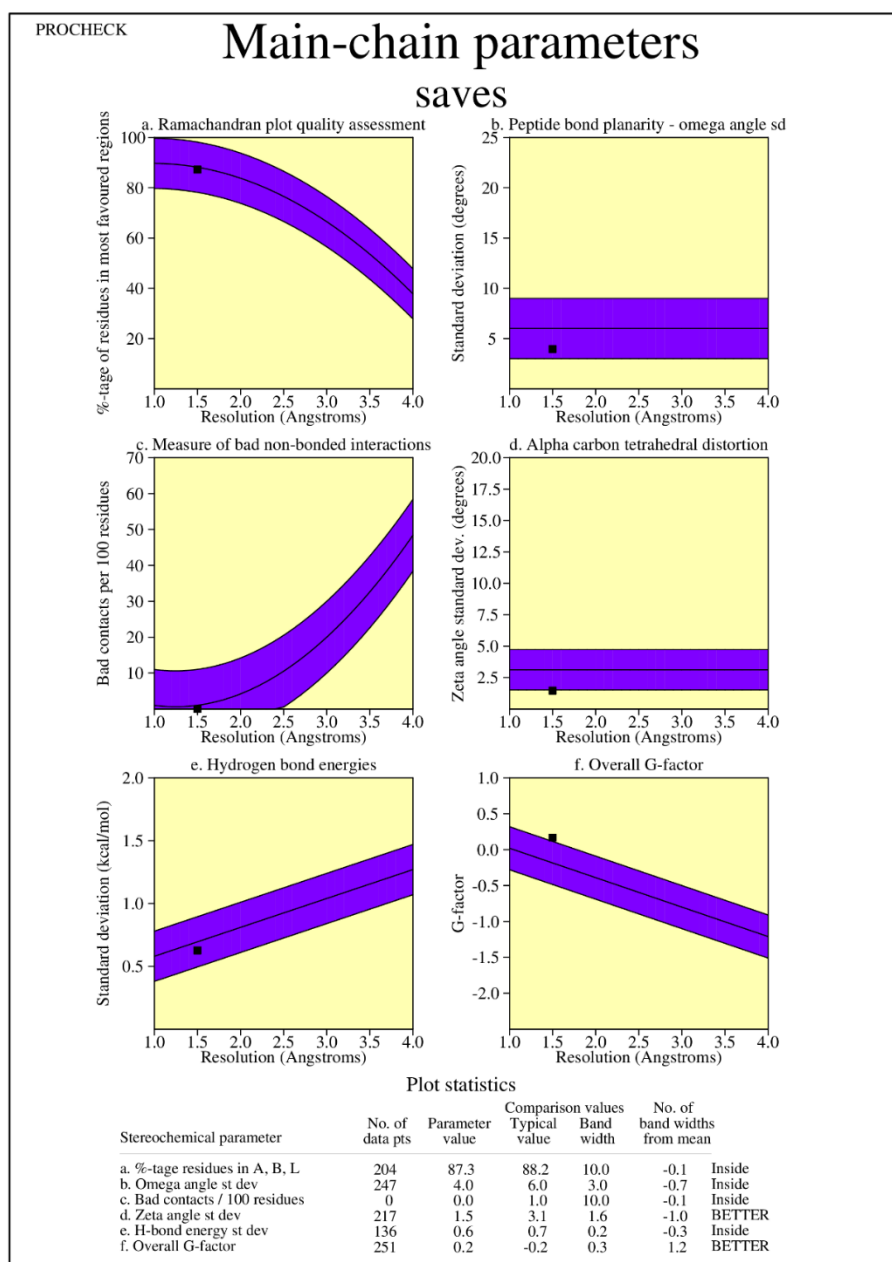
Legenda: As áreas destacadas em vermelho representam as combinações dos ângulos diedros ϕ (phi) e ψ (psi) mais favoráveis, em amarelo as regiões favoráveis, em amarelo claro as regiões menos favoráveis e em branco as regiões desfavoráveis. Os pontos pretos indicam os resíduos do modelo em regiões favoráveis e os vermelhos em regiões desfavoráveis.

Figura 32 - Gráfico de Ramachandran do scFv quimérico com *linker* modificado



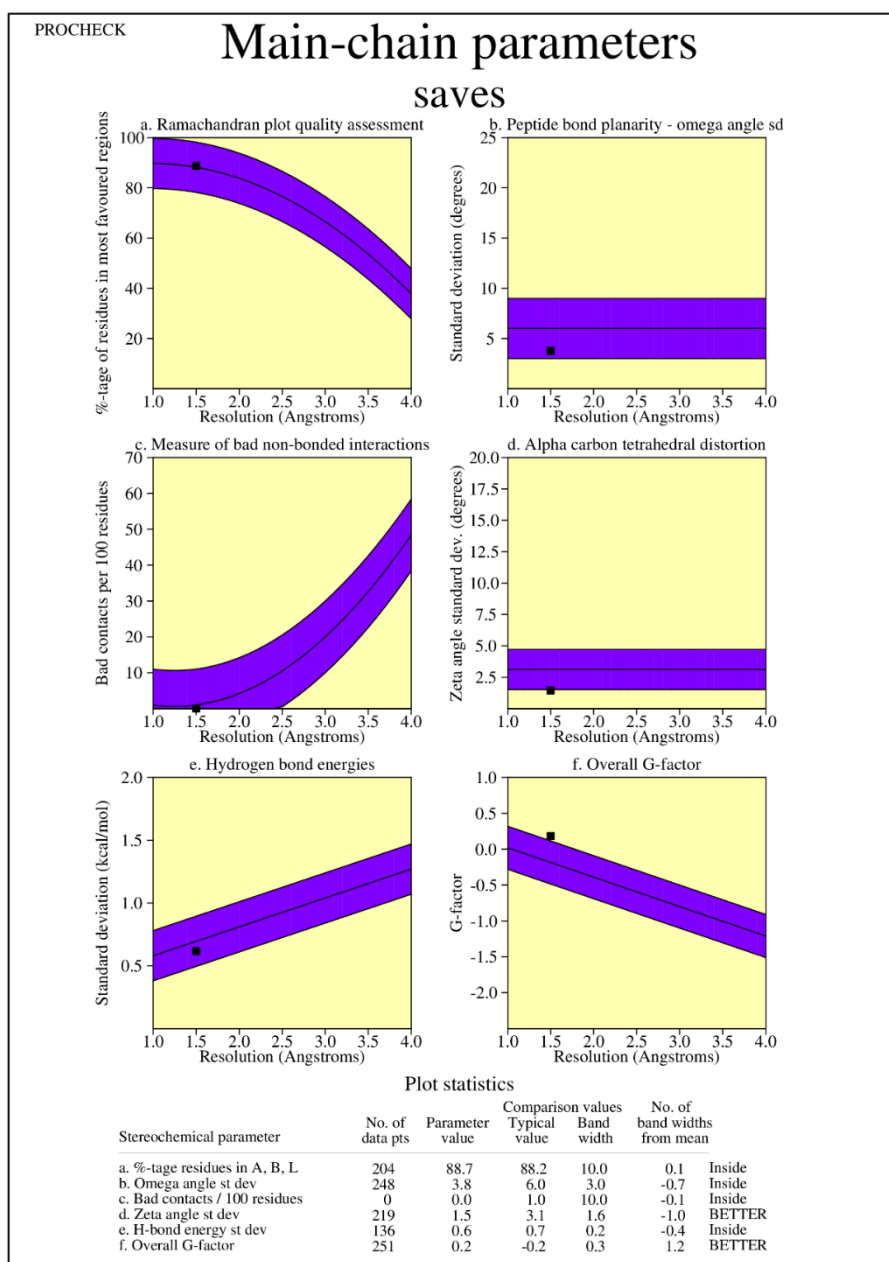
Legenda: As áreas destacadas em vermelho representam as combinações dos ângulos diedros ϕ (phi) e ψ (psi) mais favoráveis, em amarelo as regiões favoráveis, em amarelo claro as regiões menos favoráveis e em branco as regiões desfavoráveis. Os pontos pretos indicam os resíduos do modelo em regiões favoráveis e os vermelhos em regiões desfavoráveis.

Figura 33 - Parâmetros da cadeia principal do scFv 2 com *linker* original



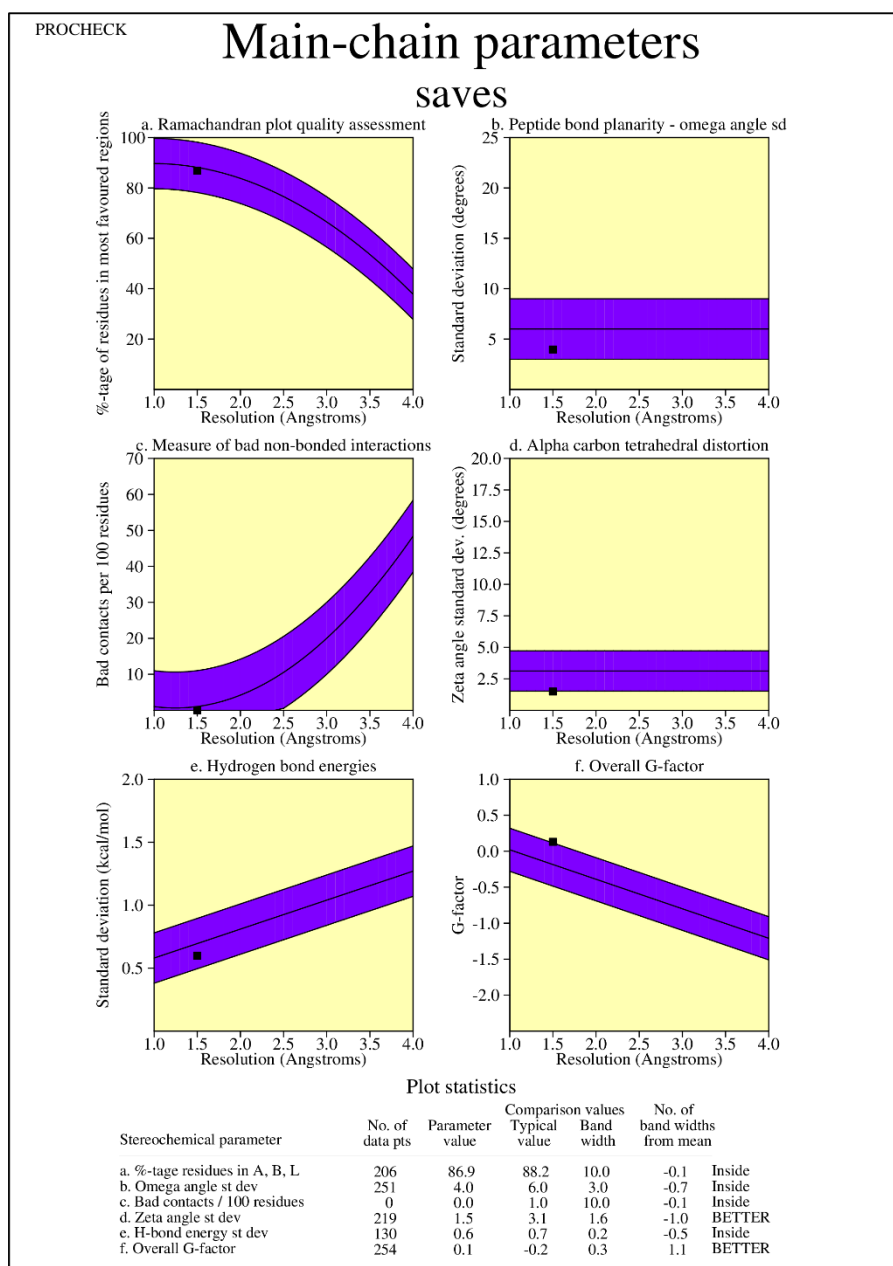
Legenda: (a) Gráfico da qualidade do Ramachandran; (b) Gráfico do desvio do ângulo ômega (ω); (c) Gráfico dos contatos ruins por 100 resíduos; (d) Gráfico do desvio do ângulo zeta (ζ); (e) Gráfico do desvio de energia de ligações de hidrogênio (H); (f) Gráfico do fator-G geral. A faixa azul escura representa os resultados das estruturas com alta qualidade. A linha central é um ajuste de mínimos quadrados para a tendência média em função da resolução (Å) e a largura da banda de cada lado corresponde a uma variação de um desvio padrão em torno da média. Os pontos pretos representam os resultados do modelo. Valores de Fator-G (f) acima da região destacada são considerados resultados melhores e abaixo piores. Para as demais variáveis (a, b, c, d, e), valores acima da região destacada são considerados piores e abaixo melhores.

Figura 34 - Parâmetros da cadeia principal do scFv 2 com *linker* modificado



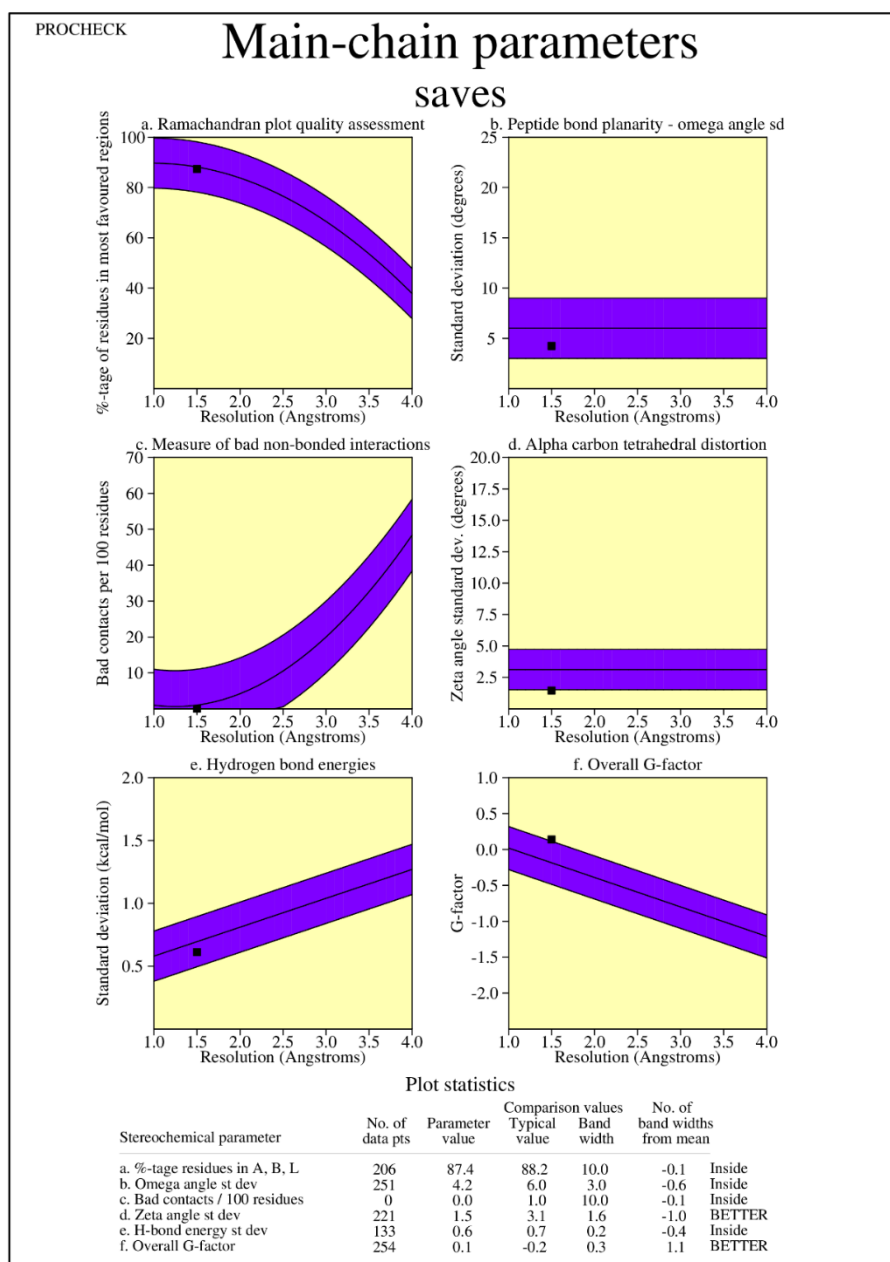
Legenda: (a) Gráfico da qualidade do Ramachandran; (b) Gráfico do desvio do ângulo ômega (ω); (c) Gráfico dos contatos ruins por 100 resíduos; (d) Gráfico do desvio do ângulo zeta (ζ); (e) Gráfico do desvio de energia de ligações de hidrogênio (H); (f) Gráfico do fator-G geral. A faixa azul escura representa os resultados das estruturas com alta qualidade. A linha central é um ajuste de mínimos quadrados para a tendência média em função da resolução (Å) e a largura da banda de cada lado corresponde a uma variação de um desvio padrão em torno da média. Os pontos pretos representam os resultados do modelo. Valores de Fator-G (f) acima da região destacada são considerados resultados melhores e abaixo piores. Para as demais variáveis (a, b, c, d, e), valores acima da região destacada são considerados piores e abaixo melhores.

Figura 35 - Parâmetros da cadeia principal do scFv quimérico com *linker* original



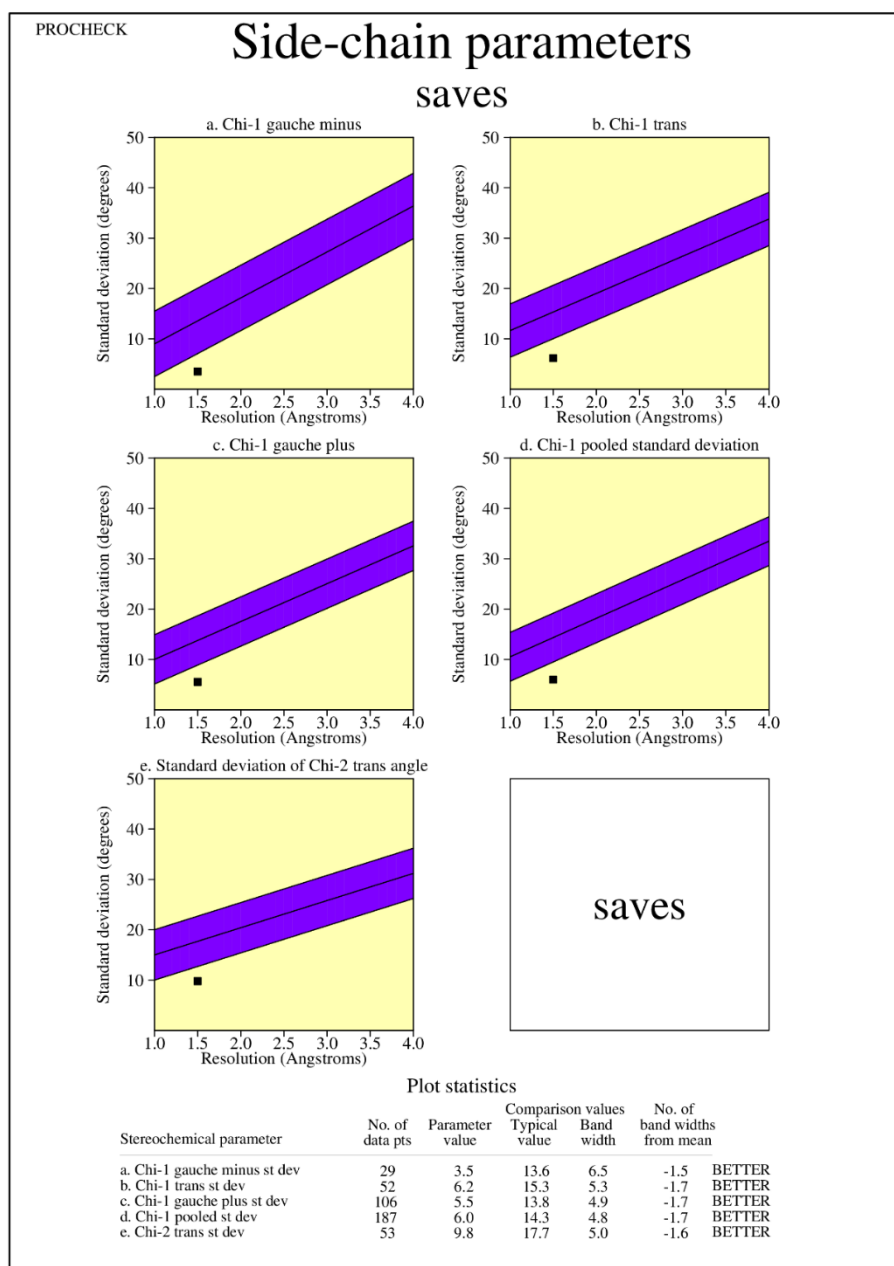
Legenda: (a) Gráfico da qualidade do Ramachandran; (b) Gráfico do desvio do ângulo ômega (ω); (c) Gráfico dos contatos ruins por 100 resíduos; (d) Gráfico do desvio do ângulo zeta (ζ); (e) Gráfico do desvio de energia de ligações de hidrogênio (H); (f) Gráfico do fator-G geral. A faixa azul escura representa os resultados das estruturas com alta qualidade. A linha central é um ajuste de mínimos quadrados para a tendência média em função da resolução (Å) e a largura da banda de cada lado corresponde a uma variação de um desvio padrão em torno da média. Os pontos pretos representam os resultados do modelo. Valores de Fator-G (f) acima da região destacada são considerados resultados melhores e abaixo piores. Para as demais variáveis (a, b, c, d, e), valores acima da região destacada são considerados piores e abaixo melhores.

Figura 36 - Parâmetros da cadeia principal do scFv quimérico com *linker* modificado



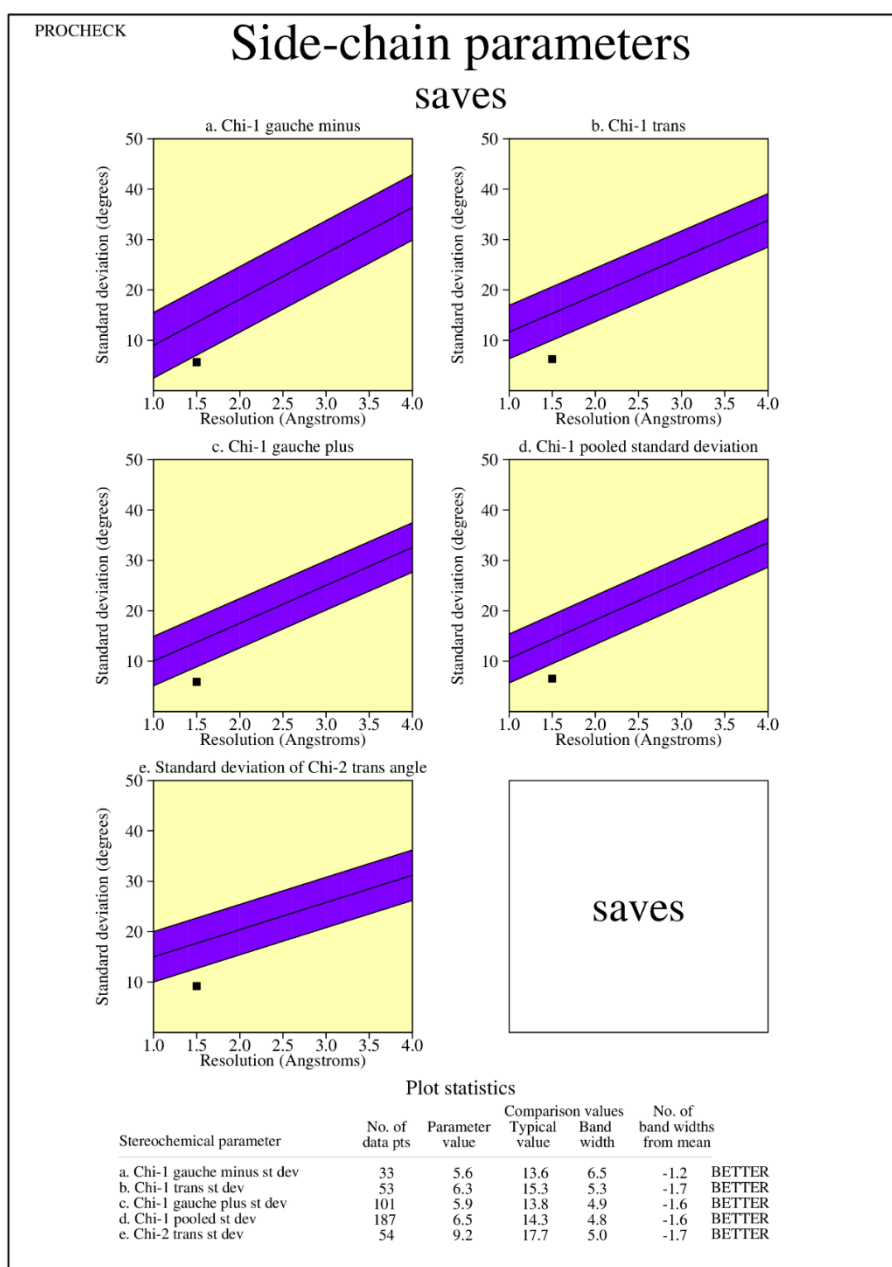
Legenda: (a) Gráfico da qualidade do Ramachandran; (b) Gráfico do desvio do ângulo ômega (ω); (c) Gráfico dos contatos ruins por 100 resíduos; (d) Gráfico do desvio do ângulo zeta (ζ); (e) Gráfico do desvio de energia de ligações de hidrogênio (H); (f) Gráfico do fator-G geral. A faixa azul escura representa os resultados das estruturas com alta qualidade. A linha central é um ajuste de mínimos quadrados para a tendência média em função da resolução (Å) e a largura da banda de cada lado corresponde a uma variação de um desvio padrão em torno da média. Os pontos pretos representam os resultados do modelo. Valores de Fator-G (f) acima da região destacada são considerados resultados melhores e abaixo piores. Para as demais variáveis (a, b, c, d, e), valores acima da região destacada são considerados piores e abaixo melhores.

Figura 37 - Parâmetros da cadeia lateral do scFv 2 com *linker* original



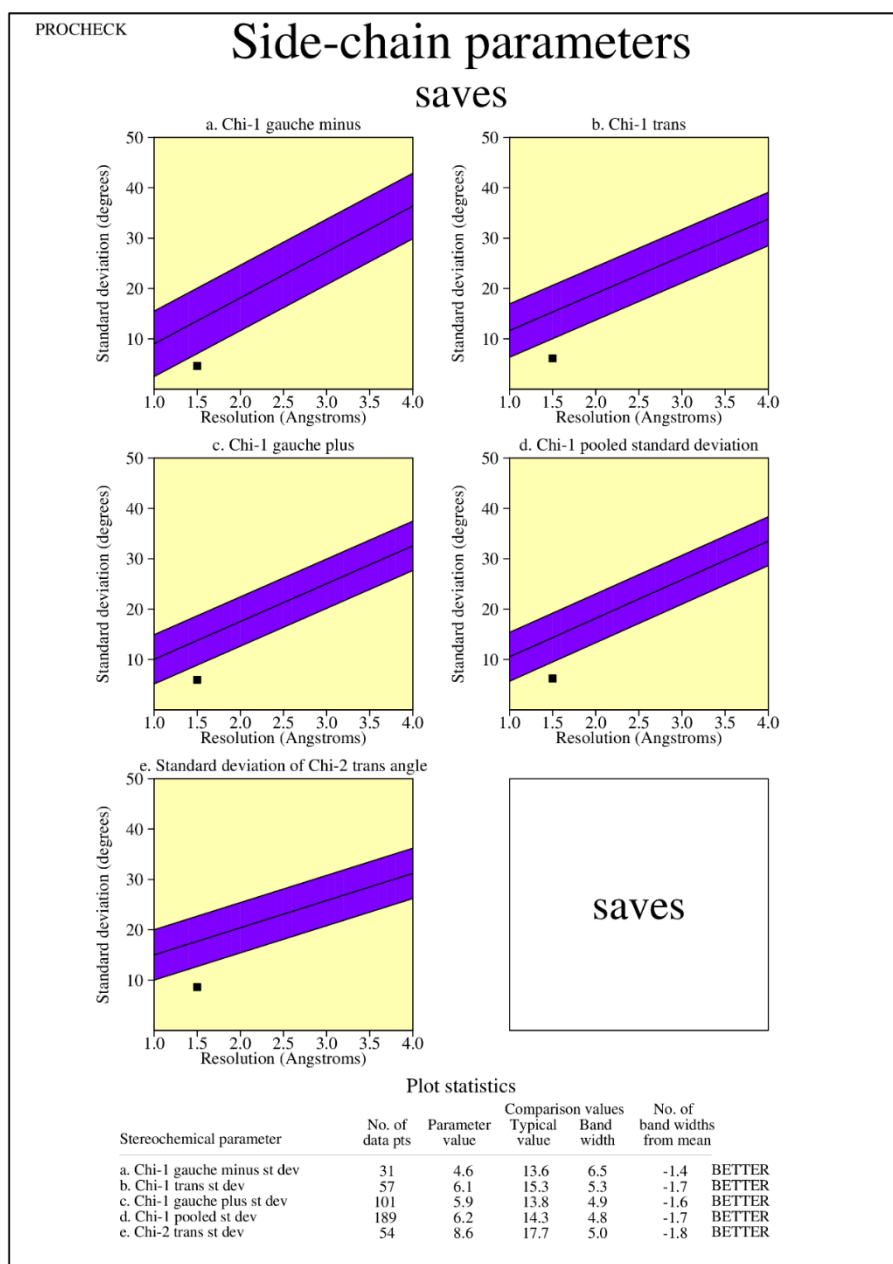
Legenda: Os gráficos mostram os desvios dos ângulos (a) Chi-1 gauche negativo; (b) Chi-1 trans; (c) Chi-1 gauche positivo; (d) Chi-1 agrupado; (e) Chi-2 trans. A linha central é um ajuste de mínimos quadrados para a tendência média em função da resolução (Å) e a largura da banda de cada lado corresponde a uma variação de um desvio padrão em torno da média. Os pontos pretos representam os resultados do modelo. Valores acima da região são considerados piores e abaixo melhores.

Figura 38 - Parâmetros da cadeia lateral do scFv 2 com *linker* modificado



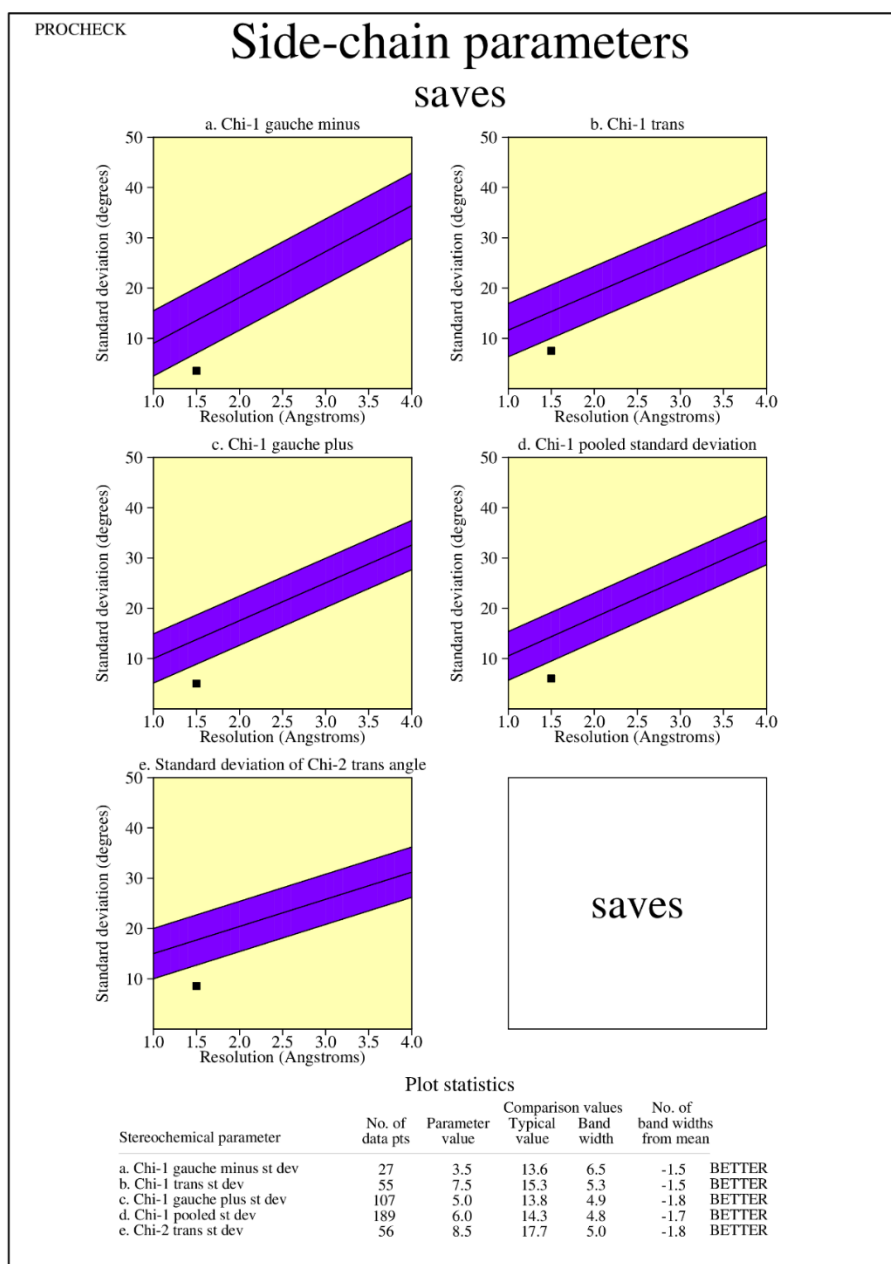
Legenda: Os gráficos mostram os desvios dos ângulos (a) Chi-1 gauche negativo; (b) Chi-1 trans; (c) Chi-1 gauche positivo; (d) Chi-1 agrupado; (e) Chi-2 trans. A linha central é um ajuste de mínimos quadrados para a tendência média em função da resolução (Å) e a largura de cada lado corresponde a uma variação de um desvio padrão em torno da média. Os pontos pretos representam os resultados do modelo. Valores acima da região são considerados piores e abaixo melhores.

Figura 39 - Parâmetros da cadeia lateral do scFv quimérico com *linker* original



Legenda: Os gráficos mostram os desvios dos ângulos (a) Chi-1 gauche negativo; (b) Chi-1 trans; (c) Chi-1 gauche positivo; (d) Chi-1 agrupado; (e) Chi-2 trans. A linha central é um ajuste de mínimos quadrados para a tendência média em função da resolução (Å) e a largura da banda de cada lado corresponde a uma variação de um desvio padrão em torno da média. Os pontos pretos representam os resultados do modelo. Valores acima da região são considerados piores e abaixo melhores.

Figura 40 - Parâmetros da cadeia lateral do scFv quimérico com *linker* modificado



Legenda: Os gráficos mostram os desvios dos ângulos (a) Chi-1 gauche negativo; (b) Chi-1 trans; (c) Chi-1 gauche positivo; (d) Chi-1 agrupado; (e) Chi-2 trans. A linha central é um ajuste de mínimos quadrados para a tendência média em função da resolução (Å) e a largura da banda de cada lado corresponde a uma variação de um desvio padrão em torno da média. Os pontos pretos representam os resultados do modelo. Valores acima da região são considerados piores e abaixo melhores.