



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS AGRÁRIAS
DEPARTAMENTO DE ENGENHARIA DE PESCA
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA DE RECURSOS
NATURAIS

NATÁLIA FERNANDES FROTA OLIVEIRA

SIMULAÇÃO DE DINÂMICA MOLECULAR DE UM MODELO DE CAR EM
INTERAÇÃO COM O CD19, MARCADOR DE CÉLULAS CANCEROSAS

FORTALEZA

2020

NATÁLIA FERNANDES FROTA OLIVEIRA

SIMULAÇÃO DE DINÂMICA MOLECULAR DE UM MODELO DE CAR EM
INTERAÇÃO COM O CD19, MARCADOR DE CÉLULAS CANCEROSAS

Dissertação apresentada ao Programa de Pós-Graduação em Biotecnologia de Recursos Naturais da Universidade Federal do Ceará, como requisito à obtenção do título de Mestre em Biotecnologia de Recursos Naturais. Área de concentração: Biologia Estrutural e Computacional.

Orientador: Prof. Dr. Marcos Roberto Lourenzoni.

FORTALEZA

2020

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária

Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

O48s Oliveira, Natália Fernandes Frota.
Simulação de Dinâmica Molecular de um modelo de CAR em interação com o CD19, marcador de células cancerosas / Natália Fernandes Frota Oliveira. – 2020.
126 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Ciências Agrárias, Programa de Pós-Graduação em Biotecnologia de Recursos Naturais, Fortaleza, 2020.
Orientação: Prof. Dr. Marcos Roberto Lourenzoni.

1. CAR T. 2. CD19. 3. Dinâmica Molecular. I. Título.

CDD 660.6

NATÁLIA FERNANDES FROTA OLIVEIRA

SIMULAÇÃO DE DINÂMICA MOLECULAR DE UM MODELO DE CAR EM
INTERAÇÃO COM O CD19, MARCADOR DE CÉLULAS CANCEROSAS

Dissertação apresentada ao Programa de Pós-Graduação em Biotecnologia de Recursos Naturais da Universidade Federal do Ceará, como requisito à obtenção do título de Mestre em Biotecnologia de Recursos Naturais. Área de concentração: Biologia Estrutural e Computacional.

Aprovada em: ___/___/_____.

BANCA EXAMINADORA

Prof. Dr. Marcos Roberto Lourenzoni (Orientador)
Fundação Oswaldo Cruz Ceará (Fiocruz-CE)

Prof. Dr. Munir Salomão Skaf
Universidade de Campinas (UNICAMP)

Prof. Dr. Raphael Trevizani
Fundação Oswaldo Cruz Ceará (Fiocruz-CE)

A Deus.

Ao meu esposo Nilton, aos meus pais, Denise
e Ronaldo e ao meu irmão, Rhonaldo.

AGRADECIMENTOS

Agradeço, primeiramente, a **Deus**, pelo dom da vida, por iluminar meu caminho e dar-me forças para superar obstáculos diários.

À **FUNCAP**, pelo apoio financeiro com a manutenção da bolsa de auxílio.

À **Fundação Oswaldo Cruz/CE** pelos recursos cedidos.

Ao Prof. Dr. **Marcos Roberto Lourenzoni**, pela paciência, pelos inestimáveis ensinamentos e excelente orientação.

Aos professores participantes da banca examinadora Dr. **Munir Salomão Skaf** e **Raphael Trevizani** pelo tempo, pelas valiosas colaborações e sugestões.

Aos meus queridos **professores do curso de Pós-Graduação em Biotecnologia de Recursos Naturais**, da Universidade Federal do Ceará, que tanto contribuíram para minha formação.

À **Universidade Federal do Ceará**, minha casa acadêmica, a qual sou grata por tudo.

Aos meus pais, **Denise e Ronaldo**, e a meu irmão **Rhonaldo** por todo o incentivo, apoio e por sempre acreditarem que eu era capaz.

Ao meu marido, **Nilton**, pelo carinho, amor, incentivo e apoio moral nas horas mais difíceis.

Aos meus **colegas de trabalho**, por deixarem o ambiente de trabalho mais descontraído e por sempre estarem dispostos a ajudar quando requisitados.

Ao **Renato**, por ser sempre tão prestativo e por me ajudar a resolver os problemas que apareceram ao longo do caminho.

Aos **funcionários da Fundação Oswaldo Cruz/CE**, pelo suporte, pela simpatia e por terem ajudado em tudo que estivesse ao alcance.

A todos os **meus familiares**, que contribuíram de alguma forma para a realização deste trabalho.

A todos os **meus amigos irmãos**, por estarem presentes em muitos momentos da minha vida e por serem pessoas companheiras sempre que preciso.

“Talvez não tenha conseguido fazer o melhor, mas lutei para que o melhor fosse feito. Não sou o que deveria ser, mas Graças a Deus, não sou o que era antes”, Marthin Luther King.

RESUMO

Os receptores de antígeno quiméricos (CARs) são constituídos por três componentes: um domínio extracelular, um domínio transmembrana e um domínio intracelular. O CAR pode ser introduzido em células T humanas para redirecionar a especificidade ao antígeno e melhorar a função imunoterápica. O domínio extracelular é o responsável pelo reconhecimento de antígeno, que geralmente é formado por um *single chain fragment variable* (scFv); e um espaçador (*hinge*). O transmembrana conecta os domínios extra e intracelular do CAR e, assim como o *hinge*, pode influenciar na função efetora da célula CAR T. O intracelular é a porção ativadora de células T, formada geralmente de CD3- ζ . O CD19 é expresso nas células B malignas e, por esse motivo, é um potente marcador de células cancerosas. O objetivo desse trabalho foi obter informações estruturais, através de simulação de Dinâmica Molecular, relacionadas ao mecanismo de sinalização em um modelo de CAR, inserido em um modelo de membrana de células T. Estruturas tridimensionais dos componentes do CAR foram obtidas por Modelagem Comparativa e submetidas a simulações de Dinâmica Molecular (DM). As estruturas do scFv, *hinge*, domínio transmembrana e intracelular foram unidas para formar o CAR e submetidas a DM. As DMs foram conduzidas usando o pacote GROMACS e o campo de força utilizado para descrever as interações atômicas foi o CHARMM36. As análises de distância, ângulo e PCA possibilitaram inferir um mecanismo de transdução de sinal no sistema CAR-CD19, que não foi observado no sistema CAR. A formação da interface *hinge*-scFv e a aproximação desse conjunto à membrana resulta em menor tensão na região de ligação *hinge*-transmembrana, que permite a inclinação da α -hélice do domínio transmembrana. Essa inclinação perdura de 370 ns a 600 ns, momento que é observada mudança conformacional no domínio intracelular, ratificada pela rápida transição na curva de PC1. Essa sequência de eventos propõe um mecanismo de transdução de sinal na célula CAR T, que inicia na interação entre o CD19 e o scFv e ocasiona uma mudança conformacional do domínio intracelular, disponibilizando resíduos de tirosina do CD3- ζ para fosforilação e transmissão do sinal no interior da célula CAR T. As movimentações que levam às mudanças conformacionais sequenciais no CAR-CD19 estão em concordância com a função do CAR descrita na literatura. O mecanismo proposto ratifica mudanças conformacionais no domínio intracelular que é essencial para expor os sítios de fosforilação e a função do CD3- ζ , que possui papel na sinalização.

Palavras-chave: CAR T. CD19. Dinâmica Molecular.

ABSTRACT

Chimeric antigen receptors (CARs) consist of three components: an extracellular domain, a transmembrane domain, and an intracellular domain. These receptors can be introduced into human T cells to redirect antigen specificity and improve function in passive immunotherapy. The extracellular domain is the responsible for antigen recognition, which is usually formed by a single chain fragment variable (scFv); and a spacer (*hinge*). The main function of the transmembrane domain is to connect the extra and intracellular domains of CAR and, as well as the *hinge*, can influence the effector function of the CAR T cell. The intracellular domain is the activating portion of T cells, usually formed of CD3- ζ . CD19 is expressed on malignant B cells and is therefore a potent marker of cancer cells. The objective was to obtain structural information, through molecular dynamics simulation, related to the signaling mechanism in a modeled CAR inserted in a T cell membrane model. The CAR components were modeled by Comparative Modeling and submitted to Molecular Dynamics (MD) simulations. The scFv, *hinge*, transmembrane and intracellular domain structures were joined to form the CAR and submitted to DM. DMs were conducted using the GROMACS package and the force field used to describe atomic interactions was CHARMM36. The distance, angle and PCA analyzes made it possible to infer the signal transduction mechanism in the CAR-CD19 system, which was not observed in the CAR system. The formation of the *hinge*-scFv interface and the approximation of this assembly to the membrane results in a reduction of the tension in the *hinge*-transmembrane binding region, which allows the α -helix bias of the transmembrane domain. This inclination lasts from 370 ns to \sim 600 ns, at which time a conformational change in the intracellular domain is observed, ratified by the rapid transition observed on the PC1 curve at 600 ns. This sequence of events proposes a signal transduction mechanism in the CAR T cell, dependent on the interaction between CD19 and scFv until the conformational change of the intracellular domain, providing CD3- ζ tyrosine residues for phosphorylation and signal transmission inside. of the CAR T cell. The movements leading to sequential conformational changes in CAR-CD19 are in agreement with the CAR function described in the literature. The proposed mechanism ratifies conformational changes in the intracellular domain that is essential for exposing phosphorylation sites and the function of CD3- ζ , which plays a role in signaling.

Keywords: CAR T. CD19. Molecular Dynamics.

LISTA DE FIGURAS

- Figura 1 – Modelo de reconhecimento de um complexo peptídeo-MHC pelas células T. Esta figura mostra uma molécula do MHC apresentando um peptídeo a um receptor de célula T, destacando o reconhecimento feito por dois resíduos polimórficos da molécula do MHC e um resíduo do peptídeo..... 25
- Figura 2 – Estrutura do receptor de células T. Esquema do TCR $\alpha\beta$ (à esquerda), identificando os domínios de um TCR específico a um complexo peptídeo-MHC. A porção do TCR de ligação ao antígeno é formada pelos domínios $V\beta$ e $V\alpha$. À direita, está a estrutura da porção extracelular de um TCR revelada por cristalografia de raios-X. A porção hipervariável, que forma o local de ligação peptídeo-MHC, está na região superior de $V\beta$ e $V\alpha$ 26
- Figura 3 – Componentes do complexo TCR. O complexo TCR de células T restrito ao MHC consiste no TCR $\alpha\beta$ não covalentemente ligado ao CD3 e a proteínas ζ 27
- Figura 4 – Eventos iniciais da fosforilação da tirosina durante a ativação de células T..... 29
- Figura 5 – Modelo estrutural do CAR. Estrutura de CARs de primeira geração, segunda e terceira, exemplificando os três domínios que o compõem: domínio extracelular, transmembrana e intracelular..... 31
- Figura 6 – Estrutura tridimensional do anticorpo, em formato de Y. Cada esfera representa um resíduo da estrutura do anticorpo. Azul: cadeias leves e cinza: cadeias pesadas..... 37
- Figura 7 – Representação estrutural do anticorpo monoclonal. A) Estrutura tridimensional de um anticorpo monoclonal. Em evidência fragmentos Fab e Fv. B) Representação esquemática de um anticorpo monoclonal, destacando as regiões CDRs e framework..... 38
- Figura 8 – Representação da estrutura do CD19. Os dois domínios extracelulares C2 semelhantes a Ig (Ig-like) são separados por um pequeno domínio

helicoidal não semelhante a Ig com possíveis ligações dissulfeto. O domínio citoplasmático de 242 aminoácidos, altamente conservado, possui vários resíduos de tirosina, sendo três desses resíduos chave para as funções biológicas do CD19, os que estão destacados na figura..... 39

Figura 9 – Representação em ribbon da porção extracelular do CD19. Cores: azul (N-terminal) ao vermelho (C-terminal). Os loops que não foram definidos na densidade eletrônica da estrutura cristalográfica (PDB 6AL5) estão mostrados em cinza. Os carboidratos observados nos locais de N-glicosilação estão mostrados como sticks em magenta; outros potenciais locais de N-glicosilação mostrados como esferas. As pontes dissulfeto estão representadas em sticks amarelas..... 40

Figura 10 – Estruturas do Fab do anticorpo B43 e do CD19 cocristalizados. A) Representação em cartoon do complexo Fab B43-CD19. CD19 está representado em cinza, cadeia pesada do Fab em azul e cadeia leve do Fab em verde. B) Interface de interação entre os domínios variáveis da cadeia leve (VL) e pesada (VH) do anticorpo B43 (superfície) e o CD19 extracelular, evidenciando os loops do epítipo, loop 1 (azul claro), loop 2 (amarelo) e loop 3 (lilás) e a região CDR H3 em linhas azuis encaixada entre os loops do epítipo. C) Interface de interação VL-VH e CD19, evidenciando as CDRs do domínio VH, CDR H1 (verde), H2 (roxo) e H3 (azul), e as alças do epítipo do CD19..... 41

Figura 11 – Representação da ligação peptídica e dos ângulos diédricos de rotação Phi e Psi..... 49

Figura 12 – Fluxograma esquemático de definição dos sistemas simulados. Abaixo de cada esquema está sua identificação e a sessão onde está descrito nos procedimentos experimentais..... 64

Figura 13 – Representação estrutural do scfv do anticorpo B43. Representação em cartoon da estrutura do scFv, destacando as regiões CDRs. L1: rosa; L2: laranja; L3: vermelho; linker: azul claro; H1: verde; H2: roxo e H3: azul... 65

Figura 14 – Representação estrutural dos modelos obtidos utilizando os programas Modeller e I-TASSER. A) Modelo de hinge e B) do domínio

transmembrana de CD8 α obtidos por Modelagem Comparativa utilizando o programa Modeller. C) Modelo do domínio intracelular, formado por 4-1BB (laranja) e CD3 ζ (verde), obtido da modelagem por threading no programa I-TASSER.....	66
Figura 15 – Perfil de RMSD para os domínios variáveis do scFv. Curva de RMSD entre as configurações estruturais do VL-VH (sobreposição e cálculo de RMSD de VL-VH), VL (sobreposição e cálculo de RMSD de VL) e VH (sobreposição e cálculo de RMSD de VH), ao longo do tempo de simulação, comparadas com a estrutura de referência inicial.....	68
Figura 16 – Resultado da análise de cluster para o scFv. Clusters de configurações de scFv ao longo da simulação, utilizando um raio de corte de 0.18 nm.....	69
Figura 17 – Representação estrutural do scFv após a sobreposição do domínio VL. A) após 100 ns (laranja); B) 200 ns (rosa) e C) 300 ns (azul claro) de simulação. O linker está destacado em cinza.....	70
Figura 18 – Perfis de RMSD para o CD19 e os domínios variáveis do scFv. Curva de RMSD entre as configurações estruturais dos componentes do scFv (sobreposição e cálculo de RMSD do VL-VH; do VL e do VH) e do CD19 (sobreposição e cálculo de RMSD do CD19), ao longo do tempo de simulação, comparadas com a estrutura inicial como referência.....	71
Figura 19 – Perfil de RMSD do complexo VLVH-CD19. Curva de RMSD das configurações estruturais do VLVH-CD19 ao longo do tempo de simulação (sobreposição e cálculo de RMSD do VLVH-CD19), comparadas com a estrutura inicial como referência.....	72
Figura 20 – Potencial de interação intermolecular (PII) scFv-CD19. PII entre o scFv e o CD19 calculado durante 600 ns, utilizando um raio de corte de 1.3 nm.....	73
Figura 21 – Representação da interface de interação scFv-CD19. Interface scFv-CD19, destacando as regiões CDR do anticorpo B43 e os loops do epítipo do CD19.....	74

Figura 22 – Potencial de interação intermolecular (PII) por resíduo. PII por resíduo calculado por programa homemade entre os resíduos dos loops do epítipo e o scFv (A, B e C) e entre os resíduos do scFv e o CD19 (D).....	75
Figura 23 – Perfil de RMSD e Raio de giro (Rg) para o hinge. A) Curva de RMSD entre as configurações estruturais do hinge, em função do tempo de simulação, comparadas com a estrutura de referência inicial. B) Rg calculado para a estrutura do hinge. Cada seta indica uma configuração durante a simulação. Amarelo: modelo; verde: em 100 ns; azul claro: em 200 ns e roxo: em 300 ns.....	76
Figura 24 – Representação estrutural do hinge. Estruturas do hinge após sobreposição de cada estrutura ao modelo. Amarelo: configuração do modelo; verde: em 100 ns; azul claro: em 200 ns e roxo: em 300 ns de simulação.....	77
Figura 25 – Perfis de RMSD para o domínio transmembrana. Curva de RMSD para as configurações estruturais do domínio transmembrana por tempo de simulação, após sobreposição com a estrutura de referência inicial nos sistemas transmembrana9-9; 12-12 e 14-14. (sobreposição e cálculo de RMSD do domínio transmembrana).....	78
Figura 26 – Perfis de RMSD para o hinge-transmembrana. Curva de RMSD entre as configurações estruturais do hinge (sobreposição e cálculo de RMSD do hinge), transmembrana (sobreposição e cálculo de RMSD do transmembrana) e hinge-transmembrana (sobreposição do domínio transmembrana e cálculo de RMSD do hinge), ao longo do tempo de simulação, comparadas com a estrutura de referência inicial.....	79
Figura 27 – Resultado da análise de cluster para o hinge-transmembrana. Clusters de configurações de hinge-transmembrana ao longo da simulação, utilizando um raio de corte de 0.6 nm.....	80
Figura 28 – Representação estrutural do hinge-transmembrana após a sobreposição do domínio transmembrana. A) no início (verde); B) em 78 ns (roxo) e C) em 200 ns (azul) de simulação.....	81
Figura 29 – Perfis de RMSD para o scFv, hinge e transmembrana. Curva de RMSD das configurações estruturais do scFv (sobreposição e cálculo de RMSD	

do scFv), hinge (sobreposição e cálculo de RMSD do hinge) e transmembrana (sobreposição e cálculo de RMSD do transmembrana) por tempo de simulação, comparadas com a estrutura de referência inicial.	82
Figura 30 – Resultado da análise de cluster para o scFv-hinge-transmembrana. Clusters de configurações de scFv-hinge-transmembrana ao longo da simulação, utilizando um raio de corte de 0.6 nm.....	83
Figura 31 – Potencial de interação intermolecular (PII) entre o scFv e a membrana (scFv-membrana) e RMSD do hinge em relação ao scFv. A) PII entre o scFv e a membrana calculado durante 500 ns, utilizando um raio de corte de 1.3 nm. B) RMSD do hinge em relação ao scFv (scFv-hinge).....	84
Figura 32 – Representação estrutural do scFv-hinge-transmembrana após sobreposição do domínio transmembrana. Conformações do scFv-hinge-transmembrana inserido em membrana após 200 (verde) e 500 ns (roxo)...	85
Figura 33 – Perfis de RMSD para os domínios transmembrana e intracelular. Curva de RMSD das configurações estruturais dos domínios intracelular (sobreposição e cálculo de RMSD do intracelular) e transmembrana (sobreposição e cálculo de RMSD do transmembrana), ao longo do tempo de simulação, comparadas com a estrutura de referência inicial.....	86
Figura 34 – Perfis de RMSD para os componentes do CAR: scFv, hinge, transmembrana e intracelular. A) Curva de RMSD das configurações estruturais do scFv (sobreposição e cálculo de RMSD do scFv), hinge (sobreposição e cálculo de RMSD do hinge), transmembrana (sobreposição e cálculo de RMSD do transmembrana) e intracelular (sobreposição e cálculo de RMSD do intracelular); B) Curva de RMSD das configurações estruturais do hinge após a sobreposição do domínio transmembrana (Transmembrana-hinge) ao longo do tempo de simulação, comparadas com a estrutura de referência inicial.....	87
Figura 35 – Resultado da análise de cluster para o CAR. Clusters de configurações do CAR ao longo da simulação, utilizando um raio de corte de 0.6 nm.....	88
Figura 36 – Distância e PII entre o scFv e a membrana. A) Distância entre o centro de massa do scFv e o carbono alpha (C α) do resíduo Trp301 da α -hélice do	

	domínio transmembrana (Distância A). B) Distância projetada em z do centro de massa do scFv e o C α do resíduo Trp301 do domínio transmembrana (Distância B). C) PII scFv-membrana. D) Ângulo β formado entre os vetores M e O. E) Representação das Distâncias A e B e dos vetores M e O, formando o ângulo β	89
Figura 37 – Raio de giro (Rg) do hinge e PII entre o hinge e a membrana (hinge-membrana) e, o domínio transmembrana (hinge-transmembrana). A) Rg do hinge ao longo do tempo. B) PII hinge-membrana. C) PII hinge-transmembrana.....		90
Figura 38 – Distância C e D e ângulo α , entre os vetores M e N. A) Distância entre os C α dos resíduos Trp301 e Leu319 do domínio transmembrana (Distância C). B) Distância em z entre as Trp301 e Leu319 do domínio transmembrana projetada no eixo z dos resíduos (Distância D). C) Ângulo α formado entre os vetores M e N. D) Representação das Distâncias C e D e dos vetores M e N, formando o ângulo α		91
Figura 39 – Representação estrutural do CAR em membrana. Conformações do CAR inserido em membrana após 200 (verde), 600 (azul claro), 1000 (roxo) e 1200 ns (amarelo), sobrepondo-se a estrutura do domínio transmembrana.		92
Figura 40 – Projeções dos componentes do CAR (PC1 e PC2) por tempo de simulação. PC1 e PC2 da estrutura do A) hinge; B) do domínio transmembrana; C) do domínio intracelular por tempo de simulação.....		93
Figura 41 – Projeções da simulação do CAR nos subespaços comuns definidos por PC1/PC2. PC1 e PC2 da estrutura do A) hinge; B) do domínio transmembrana; C) do domínio intracelular. O círculo azul representa o conjunto de configurações até 600 ns e o círculo vermelho representa o conjunto de configurações a partir de 900 ns.....		94
Figura 42 – Perfis de RMSD para os componentes do CAR: scFv, hinge, transmembrana e intracelular. A) Curva de RMSD das configurações estruturais do scFv (sobreposição e cálculo de RMSD do scFv), transmembrana (sobreposição e cálculo de RMSD do transmembrana) e intracelular (sobreposição e cálculo de RMSD do intracelular ao longo do		

tempo de simulação, comparadas com a estrutura de referência inicial. B) RMSD em função do tempo de simulação para o hinge. A curva preta mostra o RMSD utilizando como referência para sobreposição a estrutura inicial (hinge inicial) e a curva vermelha, tendo a estrutura final como referência (hinge final).....	95
Figura 43 – Distância entre o scFv e a membrana e PII scFv-hinge. A) Distância entre o centro de massa do scFv e carbono alpha (C α) do Trp560 da α -hélice do domínio transmembrana (Distância A). B) Distância em z do centro de massa do scFv e do C α do Trp560 do domínio transmembrana (Distância B). C) PII entre o scFv e o hinge (scFv-hinge). D) Ângulo β formado entre os vetores M e O.....	97
Figura 44 – Raio de giro (Rg) do hinge e PII entre o hinge e a membrana (hinge-membrana) e, o domínio transmembrana (hinge-transmembrana). A) Rg do hinge ao longo do tempo. B) PII hinge-membrana. C) PII hinge-transmembrana.....	98
Figura 45 – Distância C e D e ângulo α , entre os vetores M e N. A) Distância entre os C α dos resíduos Trp560 e Leu578 do domínio transmembrana (Distância C). B) Distância em z entre as Trp301 e Leu319 do domínio transmembrana projetada no eixo z dos resíduos (Distância D). C) Ângulo α formado entre os vetores M e N.....	100
Figura 46 – Projeções dos componentes do CAR (PC1 e PC2) por tempo de simulação. PC1 e PC2 da estrutura A) do hinge; B) do domínio transmembrana; C) do domínio intracelular por tempo de simulação.....	102
Figura 47 – Projeções da simulação do CAR interagindo com o CD19 nos subespaços comuns definidos por PC1/PC2. PC1 e PC2 da estrutura A) do hinge; B) do domínio transmembrana. C) do domínio intracelular. O círculo azul representa o conjunto de configurações até 200 ns e o círculo vermelho representa o conjunto de configurações a partir de 600 ns.....	103
Figura 48 – Representação estrutural do complexo CAR-CD19. Conformações CAR-CD19 em A) 400; B) 600 e C) 1000 ns.....	105

Figura 49 – Mudanças conformacionais na região de ligação do hinge ao domínio transmembrana, α -hélice do domínio transmembrana e no domínio intracelular. A) Distância E entre o hinge e a α -hélice do domínio transmembrana. B) Ângulo θ , formado entre os vetores P e N. C) Ângulo α , formado entre os vetores M e N. D) PC1 e PC2 da estrutura do domínio intracelular por tempo de simulação. E) Representação dos vetores M, N e P, e dos ângulos α e θ 111

LISTA DE TABELAS

Tabela 1	– Sequências de aminoácidos dos domínios variáveis do anticorpo B43 e do linker que compõem o scFv com a respectiva numeração dos resíduos e dos átomos de acordo com o arquivo do sistema em formato PDB submetido à simulação de DM.....	56
Tabela 2	– Sequência de resíduos dos componentes do CAR comercial Kymriah® usados na modelagem dos domínios espaçador (hinge), transmembrana e intracelular do CAR anti-CD19 em estudo.....	58
Tabela 3	– Numeração dos resíduos e sequência de aminoácidos das CDRs do anticorpo B43 de acordo com o arquivo PDB 6AL5 submetido à simulação de DM.....	66
Tabela 4	– Resultados obtidos no programa MolProbity para cada modelo obtido de scFv, hinge, transmembrana e intracelular.....	67
Tabela 5	– Sequência de resíduos dos componentes do CAR comercial Kymriah® usados na modelagem dos domínios espaçador (hinge), transmembrana e intracelular do CAR anti-CD19 em estudo.....	67
Tabela 6	– Numeração dos resíduos e sequência de aminoácidos das CDRs do anticorpo B43 de acordo com o arquivo PDB 6AL5 submetido à simulação de DM.....	68
Tabela 7	– Resultados obtidos no programa MolProbity para cada modelo obtido de scFv, hinge, transmembrana e intracelular.....	71
Tabela 8	– Sequência de resíduos dos componentes do CAR comercial Kymriah® usados na modelagem dos domínios espaçador (hinge), transmembrana e intracelular do CAR anti-CD19 em estudo.....	78
Tabela 9	– Numeração dos resíduos e sequência de aminoácidos das CDRs do anticorpo B43 de acordo com o arquivo PDB 6AL5 submetido à simulação de DM.....	82
Tabela 10	– Resultados obtidos no programa MolProbity para cada modelo obtido de scFv, hinge, transmembrana e intracelular.....	88

Tabela 11 – Sequência de resíduos dos componentes do CAR comercial Kymriah® usados na modelagem dos domínios espaçador (hinge), transmembrana e intracelular do CAR anti-CD19 em estudo.....	96
--	----

LISTA DE ABREVIATURAS E SIGLAS

3D	Tridimensional
ADCC	Citotoxicidade Celular Dependente de Anticorpo
BLASTp	<i>Basic Local Alignment Search Tool protein</i>
C α	Carbono alfa
CAR	Receptores de Antígeno Quiméricos
CD19	<i>Cluster of Differentiation 19</i>
CDR	Regiões Determinantes de Complementariedade
CH	Domínio Constante da Cadeia Pesada
CL	Domínio Constante da Cadeia Leve
DM	Dinâmica Molecular
EM	Esclerose Múltipla
Fab	<i>Fragment antigen-binding</i>
Fc	Fragmento cristalizável
FDA	<i>Food and Drug Administration</i>
Fiocruz/CE	Fundação Oswaldo Cruz/Ceará
Fv	Fragmento variável
GROMACS	<i>Groningen Machine for Chemical Simulation</i>
Kcal	Quilocaloria
LLC	Leucemia Linfocítica Crônica
LLA	Leucemia Linfoblástica Aguda
mAb	Anticorpo monoclonal
Nm	Nanômetro
NPT	Número de partículas, Pressão e Temperatura
Ns	Nanossegundo
NVT	Número de partículas, Volume e Temperatura
PCA	Análise de Componentes Principais
PDB	<i>Protein Data Bank</i>
PII	Potencial de Interação Intermolecular
Rcut	Raio de Corte
RMSD	<i>Root Mean Square Deviation</i>
scFv	Fragmento Variável de Cadeia Única

SUS	Sistema Único de Saúde
VH	Domínio Variável da Cadeia Pesada
VL	Domínio Variável da Cadeia Leve

LISTA DE SÍMBOLOS

α	Alpha
δ	Delta
ε	Épsilon
γ	Gama
κ	Capa
\pm	Mais-menos
μ	Mi
λ	Lambda
σ	Sigma
θ	Teta
%	Porcentagem
®	Marca Registrada
ζ	Zeta

SUMÁRIO

1	INTRODUÇÃO	25
1.1	<i>Biologia de linfócitos T</i>	25
1.2	<i>Receptores de antígenos quiméricos (CARs).....</i>	30
1.3	<i>Anticorpos.....</i>	36
1.4	<i>Receptor CD19.....</i>	38
1.5	<i>Simulação de Dinâmica Molecular (DM)</i>	42
1.6	<i>Importância e justificativa</i>	45
2	OBJETIVOS	47
2.1	<i>Objetivos gerais</i>	47
2.2	<i>Objetivos específicos.....</i>	47
3	METODOLOGIA	48
3.1	<i>Modelagem Molecular</i>	48
3.2	<i>Simulação de Dinâmica Molecular e análises</i>	51
3.3	<i>Dinâmica essencial de proteínas.....</i>	54
4	PROCEDIMENTOS EXPERIMENTAIS.....	56
4.1	<i>Modelagem dos componentes do CAR</i>	56
4.1.1	<i>Modelagem do scFv anti-CD19</i>	56
4.1.2	<i>Modelagem do hinge de CD8α</i>	57
4.1.3	<i>Modelagem do domínio transmembrana de CD8α.....</i>	57
4.1.4	<i>Modelagem do domínio intracelular formado por 4-1BB e CD3-ζ.....</i>	58
4.1.5	<i>Validação de cada modelo obtido</i>	58
4.2	<i>Simulação de Dinâmica Molecular e análises</i>	59
4.2.1	<i>Dinâmica Molecular do scFv.....</i>	59
4.2.2	<i>Dinâmica Molecular do scFv-CD19.....</i>	59
4.2.3	<i>Dinâmica Molecular do hinge</i>	60
4.2.4	<i>Dinâmica Molecular do domínio transmembrana.....</i>	60
4.2.5	<i>Modelagem e Dinâmica Molecular do hinge-transmembrana</i>	61
4.2.6	<i>Modelagem e Dinâmica Molecular do scFv-hinge-transmembrana</i>	61
4.2.7	<i>Dinâmica Molecular transmembrana-intracelular.....</i>	62
4.2.8	<i>Modelagem e Dinâmica Molecular do CAR</i>	62
4.2.9	<i>Dinâmica Molecular do CAR-CD19</i>	62
4.2.10	<i>Definição dos sistemas simulados</i>	63

4.3	<i>Análise de Componentes Principais (PCA)</i>	64
5	RESULTADOS	65
5.1	<i>Modelagem dos componentes do CAR</i>	65
5.1.1	<i>Modelagem do scFv</i>	65
5.1.2	<i>Modelagem do hinge, do domínio transmembrana e intracelular</i>	66
5.2	<i>Simulação de Dinâmica Molecular (DM) e análises</i>	67
5.2.1	<i>Dinâmica Molecular do scFv</i>	67
5.2.2	<i>Dinâmica Molecular do scFv-CD19</i>	70
5.2.3	<i>Dinâmica Molecular do hinge de CD8α</i>	75
5.2.4	<i>Dinâmica Molecular do domínio transmembrana</i>	77
5.2.5	<i>Modelagem e Dinâmica Molecular do hinge-transmembrana</i>	78
5.2.6	<i>Modelagem e Dinâmica Molecular do scFv-hinge-transmembrana</i>	81
5.2.7	<i>Dinâmica Molecular do transmembrana-intracelular</i>	85
5.2.8	<i>Modelagem e Dinâmica Molecular do CAR</i>	86
5.2.9	<i>Dinâmica Molecular do CAR-CD19</i>	94
6	DISCUSSÃO	103
7	CONCLUSÃO	110
	REFERÊNCIAS	111
	ANEXO A- ANÁLISE DA QUALIDADE GEOMÉTRICA DOS MODELOS GERADOS NO PROGRAMA MOLPROBITY	121
	ANEXO B- GRÁFICO RAMACHANDRAN GERADO PELO PROGRAMA MOLPROBITY PARA CADA MODELO OBTIDO	122
	ANEXO C - GRÁFICO RAMACHANDRAN GERADO PELO PROGRAMA PROCHECK PARA CADA MODELO OBTIDO	124

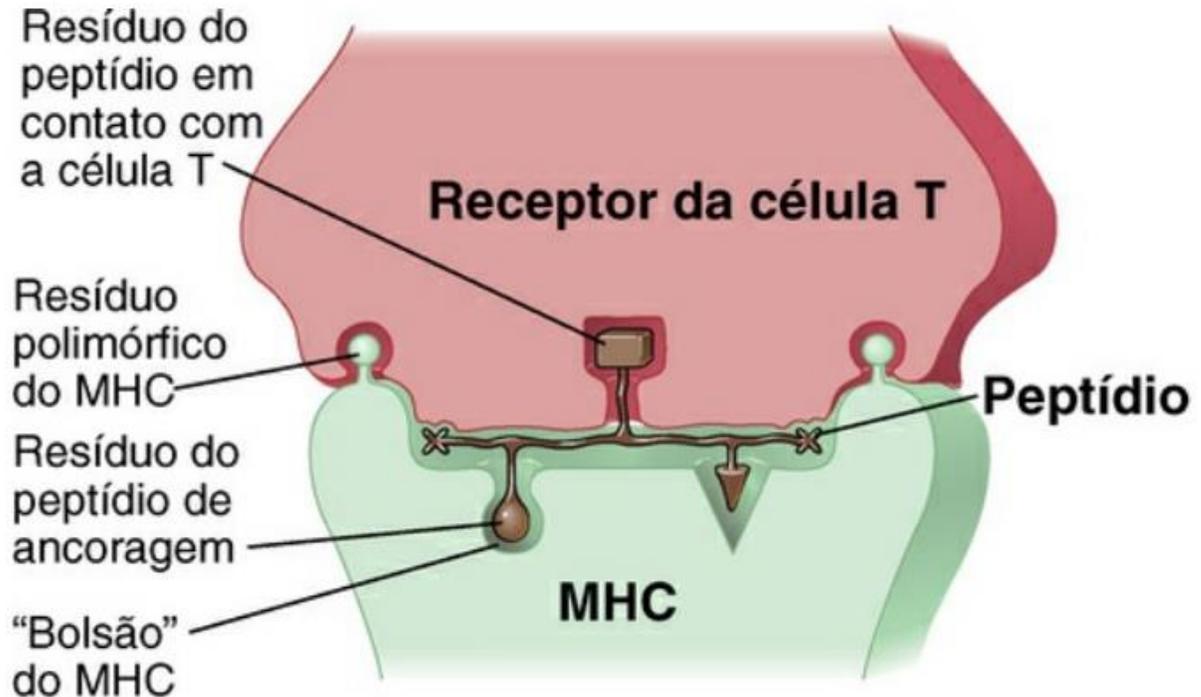
1 INTRODUÇÃO

1.1 Biologia de linfócitos T

Os linfócitos T são células brancas mononucleares do sangue derivadas de células-tronco da medula óssea que sofreram maturação no timo e são semeadas nos tecidos linfóides periféricos e no conjunto de células circulantes (VALLEJO *et al.*, 2004). O marcador molecular das células T é o complexo receptor de células T (TCR), um dímero formado por duas cadeias polipeptídicas polimórficas ligadas à membrana e conectadas por ligação dissulfeto. O TCR é específico a um ligante antigênico e está associado a uma coleção de proteínas citoplasmáticas não polimórficas ligadas à membrana, conhecida como CD3 (CLEVERS *et al.*, 1988).

O TCR reconhece apenas epítopos da molécula nativa que foram fragmentados (ou processados) e apresentados nas superfícies das células não linfóides apresentadoras de antígeno. O epítipo geralmente é um peptídeo apresentado na superfície das células apresentadoras de antígeno (APCs) que estão ligadas a moléculas do complexo principal de histocompatibilidade (MHC) de classe I (T CD8⁺) ou classe II (T CD4⁺) (Figura 1). Assim, o ligante do TCR é um complexo bimolecular formado pelo peptídeo antigênico e pelo MHC, uma conformação única reconhecida por um TCR específico. Moléculas do MHC podem ligar-se e apresentar somente peptídeos, não outras estruturas químicas, e é por isso que a maioria das células T reconhece apenas peptídeos. As moléculas do MHC são altamente polimórficas, portanto variações nessas moléculas entre indivíduos influenciam a ligação a peptídeos e o reconhecimento de células T. Uma única célula T é capaz de reconhecer um peptídeo específico, apresentado por apenas uma molécula de MHC. Assim, a apresentação do peptídeo antigênico ligado à molécula de MHC é a base para a restrição das respostas imunes mediadas por células T (ABBAS *et al.*, 2015).

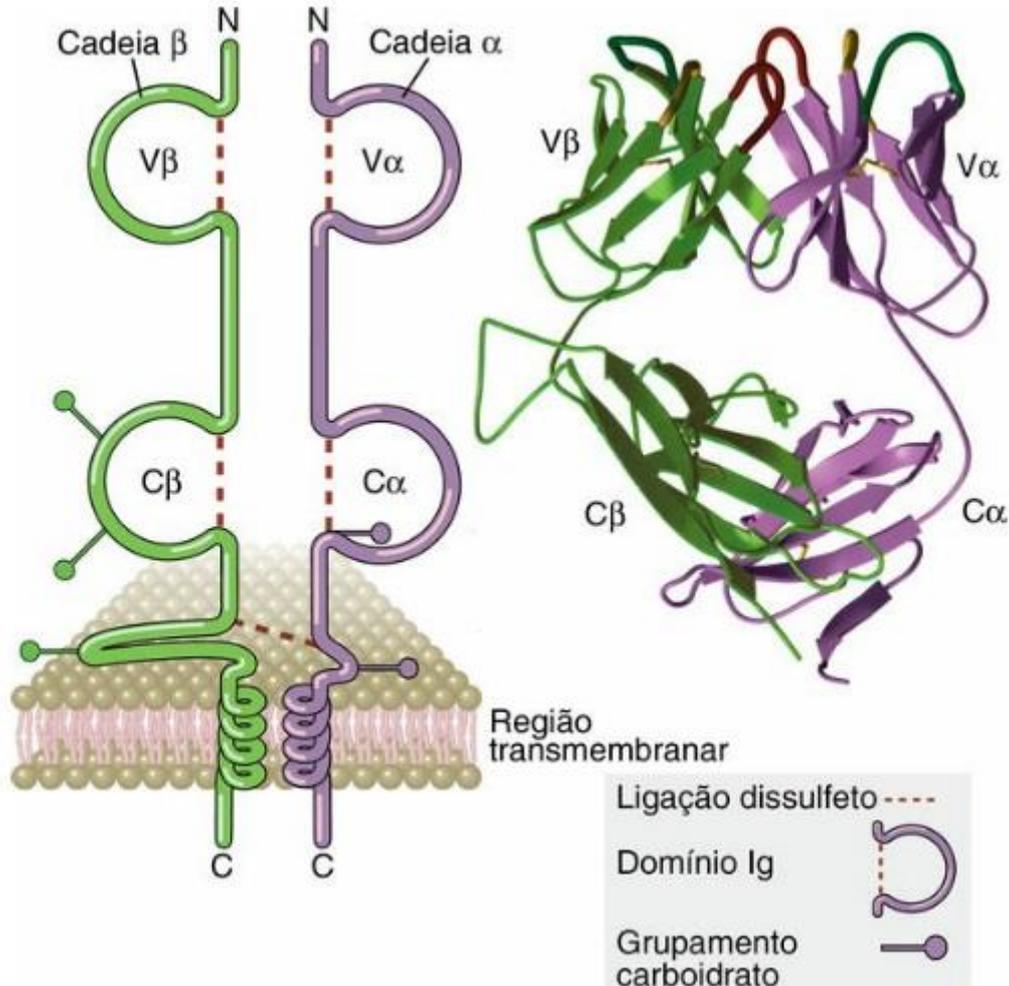
Figura 1 - Modelo de reconhecimento de um complexo peptídeo-MHC pelas células T. Esta figura mostra uma molécula do MHC apresentando um peptídeo a um receptor de célula T, destacando o reconhecimento feito por dois resíduos polimórficos da molécula do MHC e um resíduo do peptídeo.



Fonte: ABBAS *et al.* (2015).

Os receptores de antígeno das células T auxiliares, restritas ao MHC CD4⁺, e de linfócitos T CD8⁺ citotóxicos (CTL) são heterodímeros que consistem em duas cadeias polipeptídicas transmembranares, designadas TCR α e β , covalentemente ligadas umas às outras por uma ponte dissulfeto entre os resíduos extracelulares de cisteína. Estas são chamadas de células T $\alpha\beta$. Um tipo menos comum de TCR é composto por cadeias TCR γ e δ , e as células nas quais são expressos são chamadas células T $\gamma\delta$. Cada cadeia TCR α e β consiste de uma porção N-terminal do tipo imunoglobulina (Ig) de domínio variável (V), um domínio constante tipo Ig (C), uma região hidrofóbica transmembrana, e uma região citoplasmática curta de 5 a 12 resíduos (Figura 2).

Figura 2 - Estrutura do receptor de células T. Esquema do TCR $\alpha\beta$ (à esquerda), identificando os domínios de um TCR específico a um complexo peptídeo-MHC. A porção do TCR de ligação ao antígeno é formada pelos domínios $V\beta$ e $V\alpha$. À direita, está a estrutura da porção extracelular de um TCR revelada por cristalografia de raios-X. A porção hipervariável, que forma o local de ligação peptídeo-MHC, está na região superior de $V\beta$ e $V\alpha$.



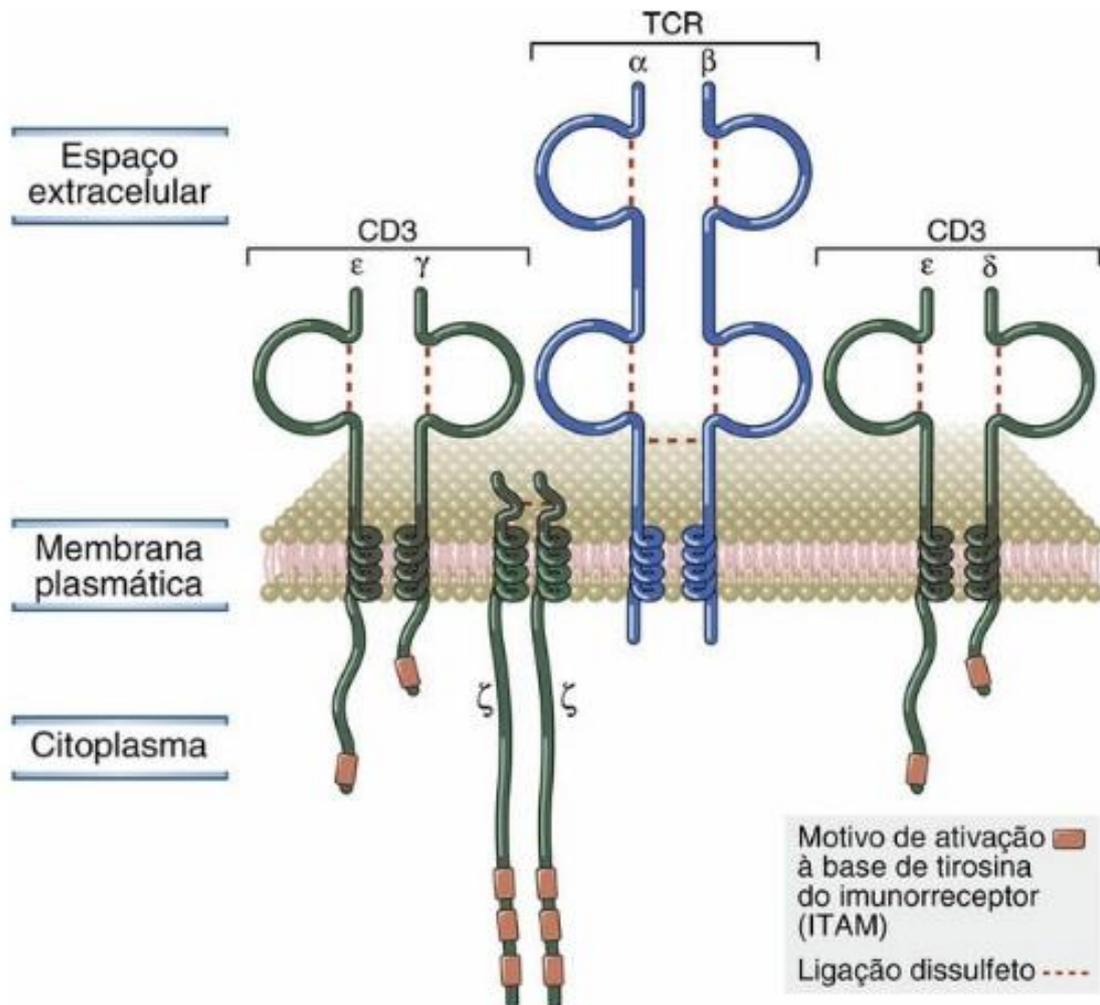
Fonte: ABBAS *et al.* (2015).

As regiões V das cadeias de TCR α e β contêm trechos curtos de aminoácidos que concentram a variabilidade entre diferentes TCRs, e estes formam as regiões hipervariáveis ou determinantes da complementariedade (CDRs). Três CDRs para a cadeia α e três CDRs para a cadeia β em conjunto formam a região do TCR que reconhece especificamente os complexos peptídeo-MHC. Da mesma forma, os anticorpos (imunoglobulinas) possuem os domínios variáveis da cadeia leve (VL) e pesada (VH) que contêm as regiões CDR, responsáveis pelo reconhecimento do antígeno (ABBAS *et al.*, 2015).

Devido às regiões citoplasmáticas do TCR serem muito pequenas para realizar a transdução de sinais, moléculas associadas fisicamente ao TCR funcionam como transdutoras de sinal pelo complexo receptor de antígenos. As proteínas CD3 e ζ estão associadas de forma

não covalente ao heterodímero TCR $\alpha\beta$ para formar o complexo TCR (Figura 3). Quando o TCR reconhece o antígeno, estas proteínas associadas levam à transdução de sinal que resulta na ativação da célula T. As proteínas CD3 e a cadeia ζ são idênticas em todas as células T, independentemente da especificidade, o que é consistente com o seu papel na sinalização e não no reconhecimento do antígeno (ABBAS *et al.*, 2015).

Figura 3 - Componentes do complexo TCR. O complexo TCR de células T restrito ao MHC consiste no TCR $\alpha\beta$ não covalentemente ligado ao CD3 e a proteínas ζ .



Fonte: ABBAS *et al.* (2015).

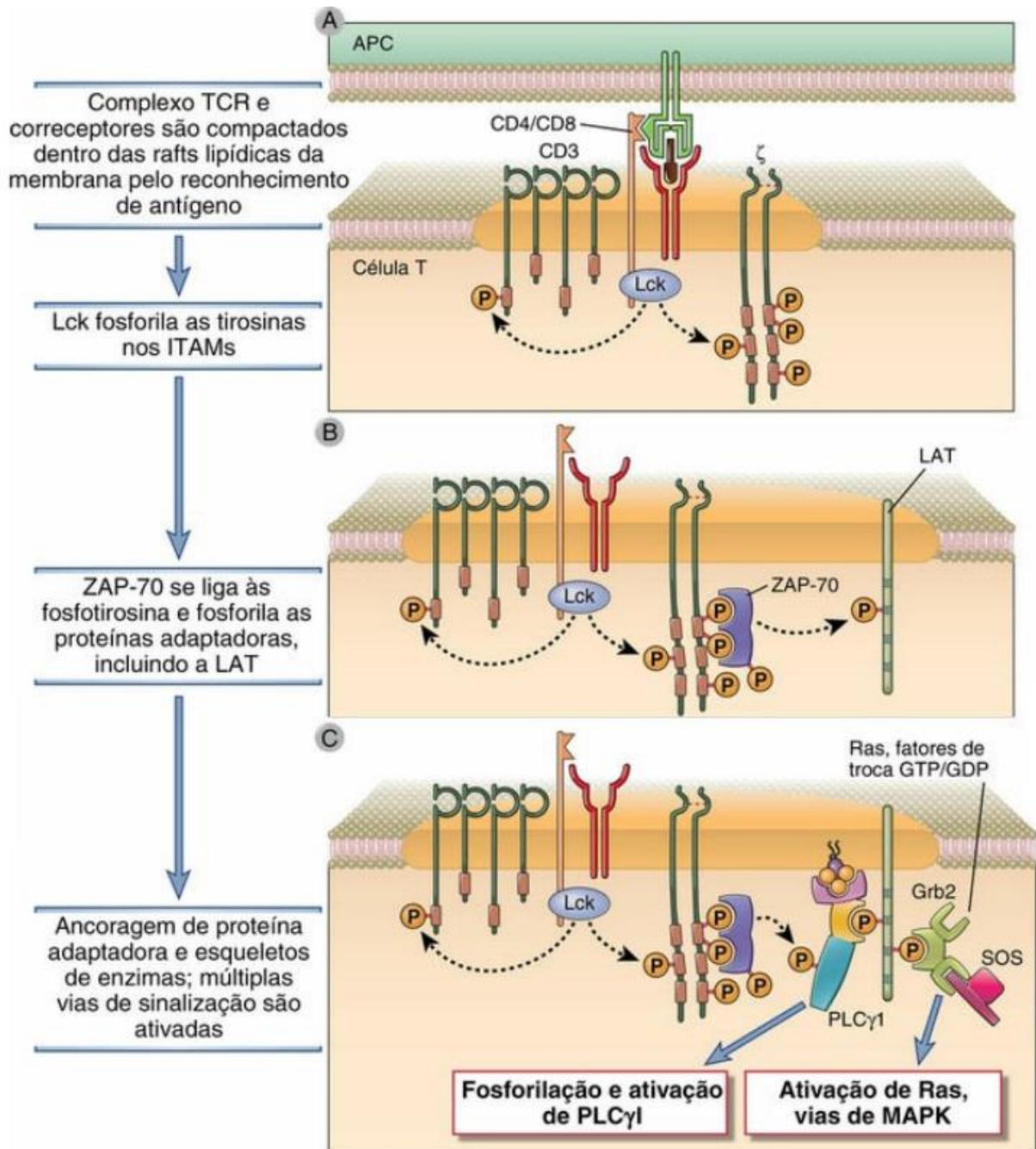
As proteínas γ CD3, δ e ϵ são homólogas entre si. As regiões N-terminais extracelulares de γ , δ , e cadeias de ϵ CD3, cada uma contém um único domínio do tipo Ig, e, por esta razão, estas três proteínas são membros da superfamília das Ig. Os segmentos transmembranares de todas as três cadeias CD3 contêm um resíduo de ácido aspártico, negativamente carregado, que interage com resíduos carregados positivamente nos domínios transmembranares das cadeias α e β do TCR. Cada complexo TCR é formado por um heterodímero TCR $\alpha\beta$ associado a um

heterodímero $\gamma\epsilon$ CD3, um heterodímero CD3 $\delta\epsilon$, e um homodímero ζ ligado por pontes dissulfeto. Os domínios citoplásmicos das proteínas γ CD3, δ e ϵ variam de 44-81 resíduos e cada um destes domínios contém um motivo de ativação à base de tirosina do imunorreceptor (ITAM), região de proteínas de sinalização que são fosforiladas em resíduos de tirosina e tornam-se locais de ligação para outras moléculas de sinalização. A cadeia ζ tem uma pequena região extracelular de nove aminoácidos, uma região transmembranar contendo um resíduo de ácido aspártico, carregado negativamente (semelhante às cadeias CD3), e uma longa região citoplasmática (113 aminoácidos) que contém três ITAMs (VALLEJO *et al.*, 2004).

A ligação do TCR aos complexos MHC-peptídeo resulta no agrupamento dos correceptores com o receptor de antígeno e a fosforilação de resíduos de tirosina de ITAM. A fosforilação de tirosinas ITAM inicia a transdução de sinal e a ativação de tirosina quinases subsequentes, que fosforilam resíduos de tirosina em outras proteínas adaptadoras. Os passos subsequentes na transdução de sinal são gerados pelo recrutamento específico de enzimas-chave que iniciam diferentes vias de sinalização cada uma. O reconhecimento de complexos MHC-peptídeo pode induzir uma alteração conformacional no TCR, fazendo com que os ITAMs associados aos CD3 ou cadeias ζ ligadas tornem-se disponíveis para a fosforilação da tirosina por quinases da família Src (ABBAS *et al.*, 2015).

A fosforilação de proteínas tem um papel central na transdução de sinal a partir do complexo TCR e correceptores. Mesmo antes da ativação do TCR, há fosforilação basal da tirosina de ITAM e recrutamento de ZAP-70 para estas ITAMs fosforiladas. Em poucos segundos após a ligação do TCR, a Lck é trazida para perto dos resíduos de tirosina no interior das ITAMs do CD3 e cadeias ζ , que são, portanto, mais extensivamente fosforiladas (Figura 4). Além da Lck associada ao correceptor, a outra quinase da família Src encontrada em associação física ao complexo de TCR é Fyn associada ao CD3, que pode contribuir para a fosforilação basal da tirosina de ITAM (ABBAS *et al.*, 2015).

Figura 4 - Eventos iniciais da fosforilação da tirosina durante a ativação de células T.



Fonte: ABBAS *et al.* (2015).

Os ITAMs com tirosinas fosforiladas na cadeia ζ são locais de ligação para a tirosina quinase da família Syk, chamada ZAP-70, proteína associada a ζ de 70 kDa. ZAP-70 contém dois domínios SH2 que podem se ligar às fosfotirosinas dos ITAM. Conforme discutido anteriormente, cada ITAM tem dois resíduos de tirosina, e ambos devem ser fosforilados para proporcionar um local de ancoragem para uma molécula de ZAP-70. A ZAP-70 acoplada torna-se um substrato para a Lck adjacente após o reconhecimento do antígeno por TCR, e a Lck

fosforila resíduos específicos de tirosina de ZAP-70 (Figura 4). Como resultado, a ZAP-70 adquire a sua atividade de tirosinaquinase e, em seguida, é capaz de fosforilar outras moléculas de sinalização citoplasmática. Um limiar crítico de atividade de ZAP-70 pode ser necessário antes que ocorram eventos de sinalização subsequentes, e este limite é alcançado pelo recrutamento de múltiplas moléculas de ZAP-70 para os ITAMs fosforilados nas cadeias ζ e nas caudas do CD3 (ABBAS *et al.*, 2015).

1.2 Receptores de antígenos quiméricos (CARs)

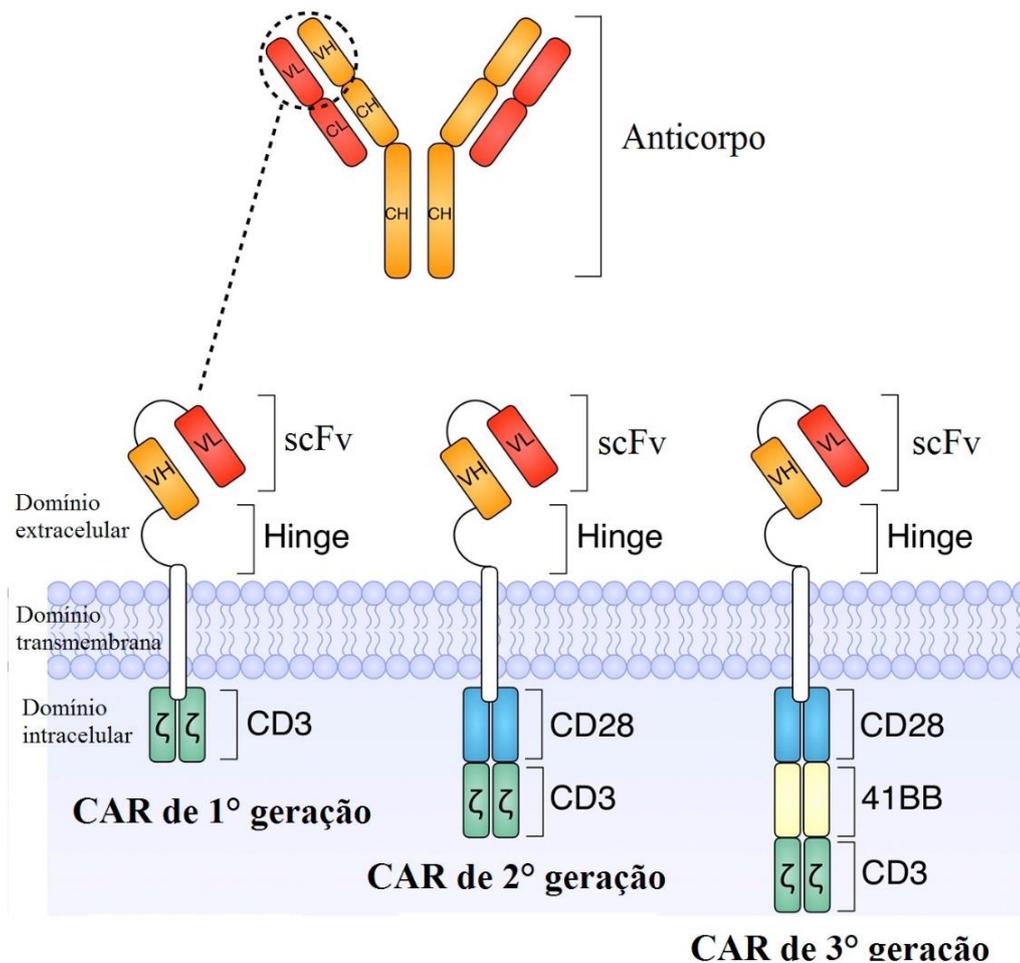
A transferência passiva de células T autólogas, que são cultivadas a partir de linfócitos tumorais infiltrantes, pode causar regressões de melanoma avançado em humanos. Como as células T reativas ao tumor não podem ser cultivadas de forma confiável, foram desenvolvidos métodos de manipulação de células T a partir da maioria dos tumores humanos para expressar genes que codificam receptores celulares específicos a antígenos tumorais (MILONE *et al.*, 2009; BRENTJENS *et al.*, 2010). A transferência passiva desse tipo de células geneticamente modificadas é uma abordagem promissora para a imunoterapia do câncer (MICHAEL C. JENSEN *et al.*, 2010). Outra abordagem terapêutica com células T passivas é projetá-las para expressar receptores de antígenos quiméricos (CARs).

Em ensaios clínicos recentes, infusões de células CAR T anti-CD19 causaram remissões completas prolongadas em pacientes com malignidades avançadas de células B (RAMOS; DOTTI, 2011; ANNESLEY *et al.*, 2017; WEI *et al.*, 2017). A transferência adotiva de células T manipuladas para expressar um receptor de antígeno quimérico (CAR) emerge então como uma poderosa imunoterapia direcionada, mostrando respostas impressionantes em pacientes com Leucemia Linfoblástica Aguda (LLA) altamente refratária (QIAN *et al.*, 2016; TOMULEASA *et al.*, 2018). No entanto, essa terapia ainda está em um estágio inicial de desenvolvimento, e muitos aspectos moleculares e estruturais necessitam de estudos mais aprofundados (ALABANZA *et al.*, 2017).

Os receptores imunes quiméricos foram desenvolvidos pela primeira vez em meados da década de 1980 e consistiam inicialmente das regiões variáveis de um anticorpo monoclonal da cadeia pesada e leve (VH e VL, respectivamente), responsáveis pelo reconhecimento do antígeno, e das regiões constantes das cadeias α e β do receptor de células T (TCR) (KUWANA *et al.*, 1987). Em 1993, Eshhar e colaboradores modificaram essa estrutura para a que é utilizada hoje, comercialmente (ESHAR *et al.*, 1993), denominada de receptor de antígeno quimérico

(CAR), que é constituído por três componentes: um domínio extracelular, um domínio transmembrana e um domínio intracelular (Figura 5). Avanços na engenharia genética combinados com um melhor entendimento do reconhecimento de células T levaram ao desenvolvimento dos CARs. Estes podem ser introduzidos em células T humanas para redirecionar a especificidade ao antígeno e melhorar a função na imunoterapia passiva, um tratamento que aumenta a capacidade do corpo de combater o câncer e atua fortalecendo as células T (KULEMZIN *et al.*, 2017).

Figura 5 - Modelo estrutural do CAR. Estrutura de CARs de primeira geração, segunda e terceira, exemplificando os três domínios que o compõem: domínio extracelular, transmembrana e intracelular.



Fonte: BORRIE; MALEKI VAREKI (2018, com adaptações).

O conceito básico relacionado à concepção de CARs é a conexão entre um domínio de reconhecimento de ligação extracelular, geralmente um *single chain fragment variable* (scFv), a um módulo de sinalização intracelular que inclui CD3-ζ, que possui a função de induzir a

ativação de células T após o reconhecimento do alvo na célula tumoral (KULEMZIN *et al.*, 2017). Os CARs são receptores projetados para serem expressos e conduzidos à superfície da membrana de células imunes efectoras (células T) e conferir especificidade a células tumorais. Assim, combinam a propriedade do anticorpo monoclonal de interagir com o antígeno presente nas células tumorais com a capacidade de lise celular e de auto renovação das células T, demonstrando algumas vantagens sobre as células T que não expressam os CARs (RAMOS; DOTTE, 2011). Uma das vantagens é que as células CAR T reconhecem e matam as células tumorais independentemente do complexo principal de histocompatibilidade (MHC). Portanto, o reconhecimento da célula alvo pelo CAR não é afetado pelos principais mecanismos que fazem com que os tumores evitem o reconhecimento de células T restritas ao MHC, como o processamento defeituoso do antígeno (CHMIELEWSKI *et al.*, 2013).

O domínio extracelular é o domínio responsável pelo reconhecimento de antígeno, que geralmente é formado por um *single chain fragment variable* (scFv); e um espaçador (*hinge*), que conecta o domínio extracelular do CAR ao domínio transmembrana, sendo os *hinges* de CD8 α ou de CD28 muito utilizados na construção de CARs (IMAI *et al.*, 2004; MILONE *et al.*, 2009). Os domínios de reconhecimento de antígeno baseados em fragmentos de anticorpos do tipo scFv são os mais utilizados nas construções de CARs, pois mantêm a especificidade do anticorpo monoclonal completo ao antígeno, possui maior penetração em tumores pouco vascularizados (YOKOTA *et al.*, 1992) e menor imunogenicidade em relação ao anticorpo completo (BEGENT *et al.*, 1996).

O desenvolvimento de novos CARs depende muito da construção de scFvs, a partir da sequência dos domínios variáveis da cadeia leve e pesada (VL e VH) de anticorpos monoclonais específicos a receptores presentes na membrana de células tumorais. A utilização de scFvs é limitada não só pelo pequeno repertório de anticorpos monoclonais seletivos ao tumor, mas também pela afinidade ao antígeno, que é significativamente maior que a do TCR, podendo interferir na sinalização. Além disso, sua especificidade é limitada a antígenos expressos na superfície das células tumorais, e não a antígenos intracelulares (SRIVASTAVA; RIDDELL, 2015).

A maioria dos CARs anti-CD19, testados clinicamente até o momento, contém scFvs derivados de anticorpos monoclonais murinos. As respostas imunes ao CAR direcionadas aos scFv murinos podem eliminar as células T que expressam o CAR e, conseqüentemente, anular a eficácia terapêutica. Já foram relatadas respostas imunes das células T direcionadas a componentes de CAR anti-CD19, sendo estas respostas problemáticas quando infusões são administradas a um paciente repetidas vezes. Com o objetivo de limitar as respostas imunes

anti-CAR, novos CARs anti-CD19 contendo regiões variáveis de um anticorpo totalmente humano vêm sendo construídos e estudados (ALABANZA *et al.*, 2017; SOMMERMEYER *et al.*, 2017).

A ligação ao antígeno realizada pelo scFv e a ativação subsequente feita pelo domínio intracelular podem ser moduladas pela introdução de uma sequência flexível de ligação no CAR, a região espaçadora (*hinge*), entre o scFv e o domínio transmembrana (ALABANZA *et al.*, 2017). O *hinge* pode provocar mudanças na função da célula CAR T ao alterar a quantidade e a flexibilidade dos resíduos que a formam (GUEST *et al.*, 2005; HUDECEK *et al.*, 2013). Hudecek e colaboradores demonstraram que *hinges* menores conferem uma maior atividade antitumoral ao CAR, por mediarem melhor a secreção de citocinas após o reconhecimento do tumor, conduzindo a um aumento significativo na citotoxicidade, proliferação e funções efetoras *in vitro* (HUDECEK *et al.*, 2013).

O *hinge* de CD8 α é utilizado no CAR anti-CD19 já comercializado pela empresa Novartis, Kimriah[®] (ZHENG *et al.*, 2018). Estudos confirmam a preferência por esse *hinge*, uma vez que já foi comprovado que CARs com *hinge* de CD8 α produzem menores níveis de citocina e menor morte celular induzida por ativação (AICD) quando comparado ao *hinge* de CD28 (ALABANZA *et al.*, 2017), o que confere uma maior persistência, *in vivo*, da célula CAR T.

Uma comparação entre as estruturas dos domínios extracelulares de CD28 e CD8 α revela diferenças entre a interface de homodimerização, devido a diferenças nos domínios espaçadores dessas duas moléculas. Os resíduos que abrangem a interface de dimerização do domínio espaçador do homodímero de CD28 desempenham um papel importante na formação de dímeros, pois formam uma área de interface não polar, contendo uma grande quantidade de átomos completamente enterrados, e também, devido à ligação dissulfeto entre os resíduos Cys123 de cada monômero (EVANS *et al.*, 2005). Já estudos realizados com a interface de dimerização do homodímero de CD8 α , revelam que a cauda C-terminal, contendo 27 resíduos, não é ordenada, o que sugere flexibilidade para essa região. Assim, é menos provável que o *hinge* de CD8 α esteja localizado em interface de dimerização, com os resíduos do C-terminal afastados da estrutura. Isso sugere que o *hinge* de CD8 α próximo ao domínio scFv não interage na configuração de um CAR dimérico (ALABANZA *et al.*, 2017). Além disso, a porção extracelular do CD8 α possui potenciais locais de O-glicosilação e resíduos carregados negativamente, o que também reduz a propensão à auto associação (LEAHY *et al.*, 1992).

O domínio transmembrana tem como principal função conectar os domínios extra e intracelular do CAR e, assim como o *hinge*, pode influenciar na função efetora da célula CAR

T. O domínio transmembrana consiste de uma α -hélice hidrofóbica que atravessa a membrana (GUEDAN *et al.*, 2019). Os domínios transmembrana derivados de CD4, CD28 ou CD8 α são os mais utilizados para CARs de segunda e terceira geração (IMAI *et al.*, 2004; KOCHENDERFER *et al.*, 2009). A ligação do domínio intracelular próxima ao domínio transmembrana pode permitir a sinalização adequada das células CAR T. O uso dos domínios transmembrana de CD8 α ou CD28, já amplamente utilizados, pode melhorar a expressão ou a estabilidade do CAR (GUEDAN *et al.*, 2019).

O domínio intracelular é a porção ativadora de células T, formada geralmente de CD3- ζ nos CARs de primeira geração (TASIAN; GARDNER, 2015; LIU *et al.*, 2017). No entanto, a sinalização de CD3- ζ pode não ser suficiente para produzir uma resposta imune durável. Sem um segundo sinal, ou coestímulo, as células T sofrem apoptose rapidamente após estimulação (LIEBOWITZ *et al.*, 1998; SALOMON; BLUESTONE, 2001). Esses domínios simulam a coestimulação que é proporcionada durante o reconhecimento do receptor de células T (TCR) por células apresentadoras de antígeno (APC) e requerida para ativação de células T fisiológicas completas (SRIVASTAVA; RIDDELL, 2015). Além disso, os domínios coestimuladores conferem às células CAR T uma maior persistência, *in vivo*, e maior eficácia antitumoral em relação às que possuem apenas CD3- ζ no seu domínio intracelular (FINNEY *et al.*, 1998). Portanto, os CARs de segunda e terceira geração, que apresentam respectivamente um ou dois domínios coestimuladores intracelulares, geralmente CD28 ou 4-1BB (Figura 5), ligados ao CD3- ζ , apresentam-se como alternativa aos CARs de primeira geração. O comprimento do domínio intracelular do CAR determina sua capacidade de interagir com moléculas de sinalização endógenas (RAMELLO *et al.*, 2019). O que pode explicar por que alguns CARs promovem sinalização tônica, independente de ligante, que tem sido associada a perda antecipada de função (LONG *et al.*, 2015), e as diferenças entre os CARs no início e na amplificação da ativação das células T.

CARs de segunda geração, específicos para CD19, contendo os coestimuladores CD28 ou 4-1BB, demonstraram atividade antitumoral, *in vivo*, em modelos pré-clínicos e ensaios clínicos para células B malignas (QIAN *et al.*, 2016; WEI *et al.*, 2017). Além disso, CARs de segunda e terceira geração mostram-se mais efetivos que os de primeira, pois promovem maior secreção de citocinas, expansão de células T após repetidas exposições ao antígeno e persistência de células T (MAHER *et al.*, 2002). A ativação das células T ocorre com dupla sinalização (sinal 1 e 2), onde participam três tipos diferentes de receptores, receptores de antígenos das células T (TCR), receptores de citocinas e receptores coestimuladores. O primeiro é o sinal desencadeado pelo TCR, que reconhece o complexo peptídico antigênico-MHC na

superfície das células apresentadoras de antígeno (APC). O segundo sinal é o sinal coestimulador, produzido por uma molécula coestimuladora, como CD28, que promove a síntese de interleucina 2 (IL-2) para completar a ativação das células T e evitar a apoptose. As células T naive não podem desempenhar seu papel normal se o sinal coestimulador estiver ausente, mesmo se as células T forem estimuladas pelo antígeno. Portanto, os CARs que incluem apenas a sequência CD3- ζ não podem ativar as células CAR T sem um sinal coestimulador. Consequentemente, os CARs de segunda geração possuem domínios de sinalização intracelular de vários receptores de proteínas coestimuladoras ligados à cauda citoplasmática dos CARs para fornecer sinais adicionais à célula T, que podem melhorar a proliferação, citotoxicidade e resposta sustentada, e prolongar a vida das células CAR-T *in vivo* (FINNEY *et al.*, 1998; PARK *et al.*, 2011).

O domínio 4-1BB pode manter o sinal de resposta das células T, que desempenha um papel fundamental na sobrevivência das células T e na memória das células T CD8⁺ citotóxicas (ACUTO; MICHEL, 2003; FINNEY *et al.*, 2004). O 4-1BB, membro da família de receptores de fator de necrose tumoral, foi utilizado no presente trabalho, pois evita a morte celular induzida por ativação de células T (BUKCYNSKI *et al.*, 2003) e reduz a exaustão de células T induzida pela sinalização tônica de CARs (LONG *et al.*, 2015). Além de induzir a expansão de células CD8⁺ e melhorar as respostas das células T CD8⁺ citotóxicas durante infecção viral e rejeição de aloenxertos (SHUFORD *et al.*, 1997). Há também evidências experimentais com modelos animais que apontam para um papel crucial da sinalização de 4-1BB para respostas antitumorais eficazes. Além disso, células CAR T anti-CD19 de segunda geração, que possuem 4-1BB como coestimulador, apresentam potente atividade antileucêmica (IMAI *et al.*, 2004).

A especificidade e segurança das células CAR T são determinadas pela escolha da molécula alvo, e a maioria dos alvos identificados até agora não são totalmente restritos a tumores em sua expressão. Por exemplo, CD19 e CD20 são expressos em células B malignas e saudáveis e, portanto, células T expressando CARs específicos para esses antígenos destroem células tumorais, mas também células B saudáveis, resultando em depleção de células B por muito tempo. As células CAR T funcionais persistem *in vivo*, assim a ativação de um gene suicida coexpresso condicional a células CAR T pode ser ativado para eliminá-las ou regular sua expressão, superando a imunodepleção com a reposição constante do conjunto de células B pelas células-tronco hematopoiéticas (SRIVASTAVA; RIDDELL, 2015).

Nem todos os pacientes que recebem células CAR T anti-CD19 alcançam a remissão total de suas malignidades, assim faz-se necessário a proposição de células CAR T mais eficientes contra células tumorais (ALABANZA *et al.*, 2017). As respostas antitumorais após

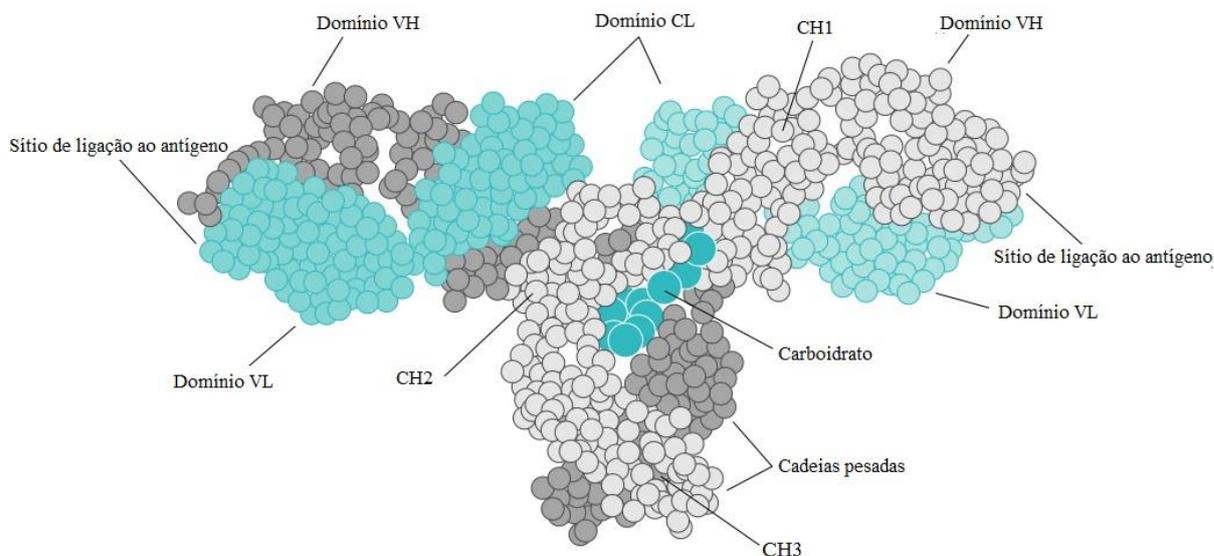
infusões de células CAR T têm sido consistentemente associadas a altos níveis de células T no sangue, portanto um objetivo geral é desenvolver abordagens que não só aumentem os níveis de expressão dessas células nos pacientes, mas também modulem a afinidade entre o scFv utilizado no CAR e o antígeno presente na célula alvo e a ação citotóxica, após a ativação das células CAR T.

1.3 Anticorpos

Imunoglobulinas humanas (anticorpos) são um grupo de glicoproteínas estrutural e funcionalmente similares que conferem imunidade a seres humanos. A estrutura geral de um anticorpo consiste de uma estrutura bifuncional (Figura 6), que compreende, de um lado, dois sítios de ligação ao antígeno (*antigen binding fragment*, Fab), e um sítio efetor (*crystallized fragment*, Fc) do outro lado (ABBAS, ABUL K.; LICHTMAN, ANDREW H.; PILAI, 2012). Todos os anticorpos possuem as mesmas características estruturais básicas, mas apresentam uma grande variedade nas regiões que se ligam aos antígenos.

Os anticorpos são constituídos de quatro cadeias, sendo duas cadeias leves idênticas (isotipos κ ou λ) e duas cadeias pesadas idênticas (isotipos γ , ϵ , δ , α , ou μ). A cadeia leve é formada por dois domínios, chamados de domínio variável (V_L) e constante (C_L), devido à variabilidade da sequência de aminoácidos. A cadeia pesada contém um domínio variável (V_H) e três ou quatro domínios constantes (C_{H1} , C_{H2} , C_{H3} , C_{H4}), dependendo do isotipo (ALBERTS *et al.*, 1989). Cada domínio é formado pelo empacotamento, face a face, de duas folhas β , unidas por pontes dissulfeto conservadas e pela interação entre as fitas dos *loops* (CHOTHIA; LESK, 1987). Pontes dissulfeto intracadeias estabilizam os domínios de ligações dissulfeto intercadeias dobráveis e também a interação entre as cadeias pesada e leve, e entre as pesadas.

Figura 6 - Estrutura tridimensional do anticorpo, em formato de Y. Cada esfera representa um resíduo da estrutura do anticorpo. Azul: cadeias leves e cinza: cadeias pesadas.

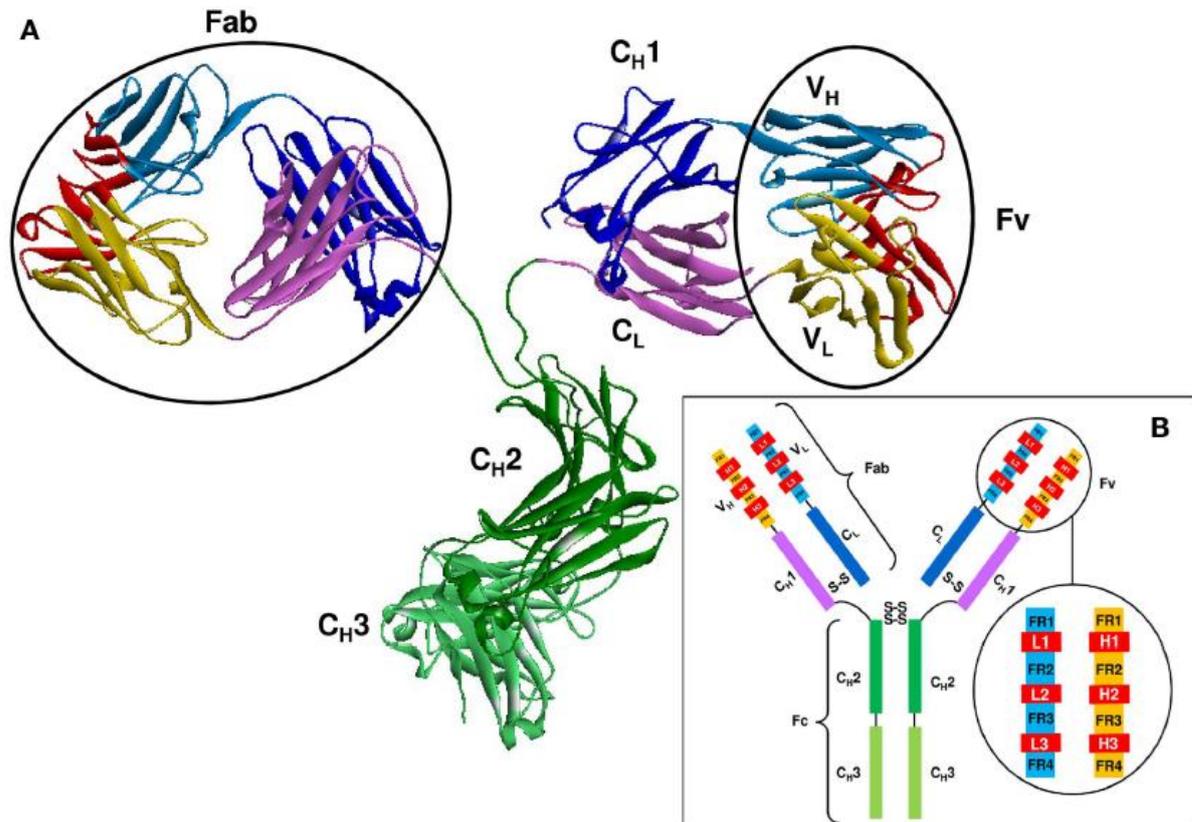


Fonte: ELGERT (1998, com adaptações).

Os sítios de ligação ao antígeno da maioria dos anticorpos são formados, primariamente, por seis *loops* ou regiões determinantes de complementariedade (CDRs), sendo três no domínio VL (CDR L1, L2 e L3) e três no domínio VH (CDR H1, H2 e H3), enquanto que as outras regiões dos domínios variáveis, que não compõem os *loops*, são denominadas “*framework*” (HOLLIGER; HUDSON, 2005; NELSON; REICHERT, 2009). Essas regiões são altamente conservadas, tanto em sequência quanto na conformação da cadeia principal, podendo, desta forma, ser modeladas com técnicas padrão de modelagem por homologia (MOREA *et al.*, 1998).

O scFv é formado pelos domínios VH e VL (Fv) de um anticorpo monoclonal (Figura 7) ligados por um polipeptídeo flexível (*linker*) (AHMAD *et al.*, 2012), que previne a dissociação dos domínios (MOHAMMADI, M.; NEJATOLLAHI, 2014). O scFv possui simultaneamente uma melhor capacidade farmacocinética de penetração em tecidos e demanda menor custo computacional durante a simulação de Dinâmica Molecular (DM) (KANDA; WATANABE, 1996; HOLLIGER; HUDSON, 2005; NELSON; REICHERT, 2009), em relação ao anticorpo completo. Além disso, esse fragmento mantém a região de ligação do anticorpo ao antígeno e possui a mesma especificidade e afinidade do anticorpo monoclonal (WANG; DUAN, 2011). A figura 7B evidencia a representação das regiões CDRs, responsáveis pelo reconhecimento do antígeno, e as regiões *framework* de um anticorpo monoclonal (SELA-CULANG *et al.*, 2013).

Figura 7 – Representação estrutural do anticorpo monoclonal. A) Estrutura tridimensional de um anticorpo monoclonal. Em evidência fragmentos Fab e Fv. B) Representação esquemática de um anticorpo monoclonal, destacando as regiões CDRs e *framework*.



Fonte: SELA-CULANG *et al.* (2013).

1.4 Receptor CD19

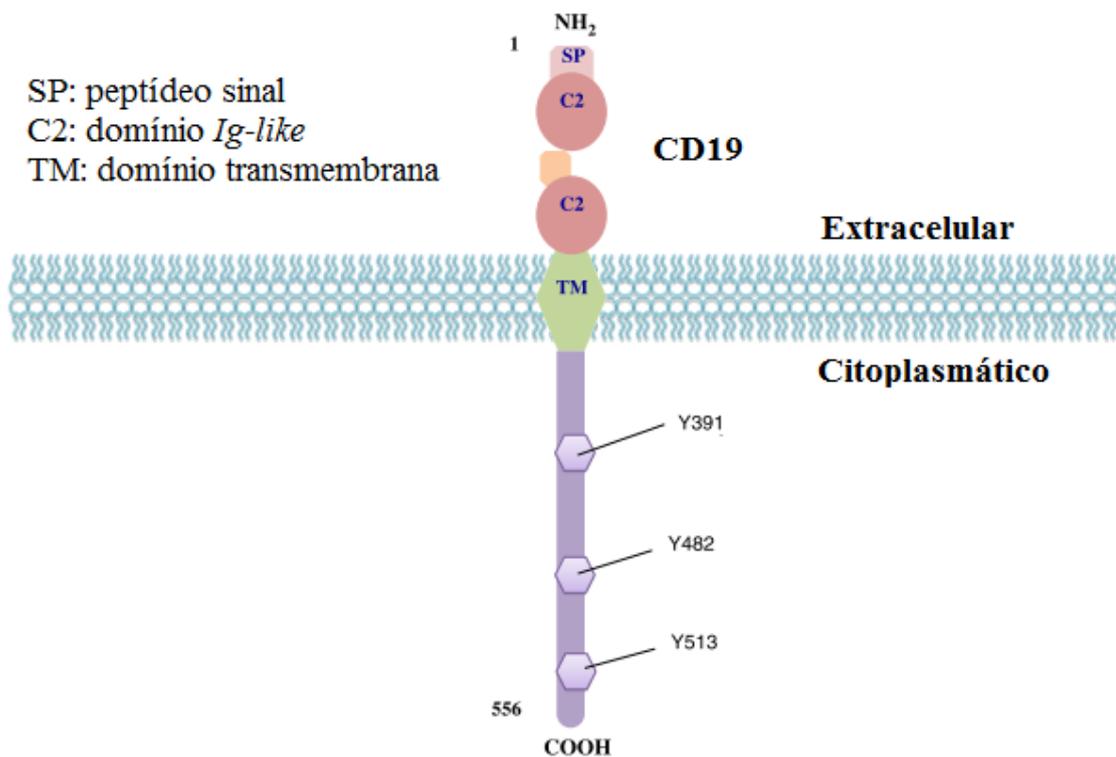
O antígeno CD19 humano é uma glicoproteína transmembrana de 95 kDa presente nas células B progenitoras, naive e de memória, bem como em células tumorais que acometem linfócitos B, mas não nas células hematopoiéticas nem nas células plasmáticas terminalmente diferenciadas (TEPLYAKOV *et al.*, 2018). O CD19 não é prontamente eliminado das células e é internalizado com o anticorpo ligado, permitindo a entrega de toxinas ligadas a anti-CD19 (UCKUN *et al.*, 1988).

O CD19 é um biomarcador para o desenvolvimento de células B, diagnóstico de linfoma e imunoterapia. Este receptor é classificado como uma proteína transmembrana do tipo I, com um único domínio transmembrana, um C-terminal citoplasmático e um N-terminal extracelular. Não existe homologia significativa entre o CD19 e outras proteínas conhecidas, sendo um alvo perfeito para imunoterapia mediada por CAR (TEPLYAKOV *et al.*, 2018).

A porção extracelular do CD19 contém dois domínios do tipo Ig (imunoglobulina), do tipo C2, divididos por um domínio não semelhante a Ig, ligado por ligação dissulfeto, bem

como locais de adição de carboidratos ligados a N (Figura 8). O domínio citoplasmático altamente conservado consiste em 242 aminoácidos com nove resíduos de tirosina próximos ao C-terminal (HAAS; TEDDER, 2005; TEDDER, 2009). Vários estudos sugerem que as funções biológicas do CD19 dependem de três resíduos citoplasmáticos da tirosina - Y391, Y482 e Y513, representados na Figura 8 (WANG *et al.*, 2012).

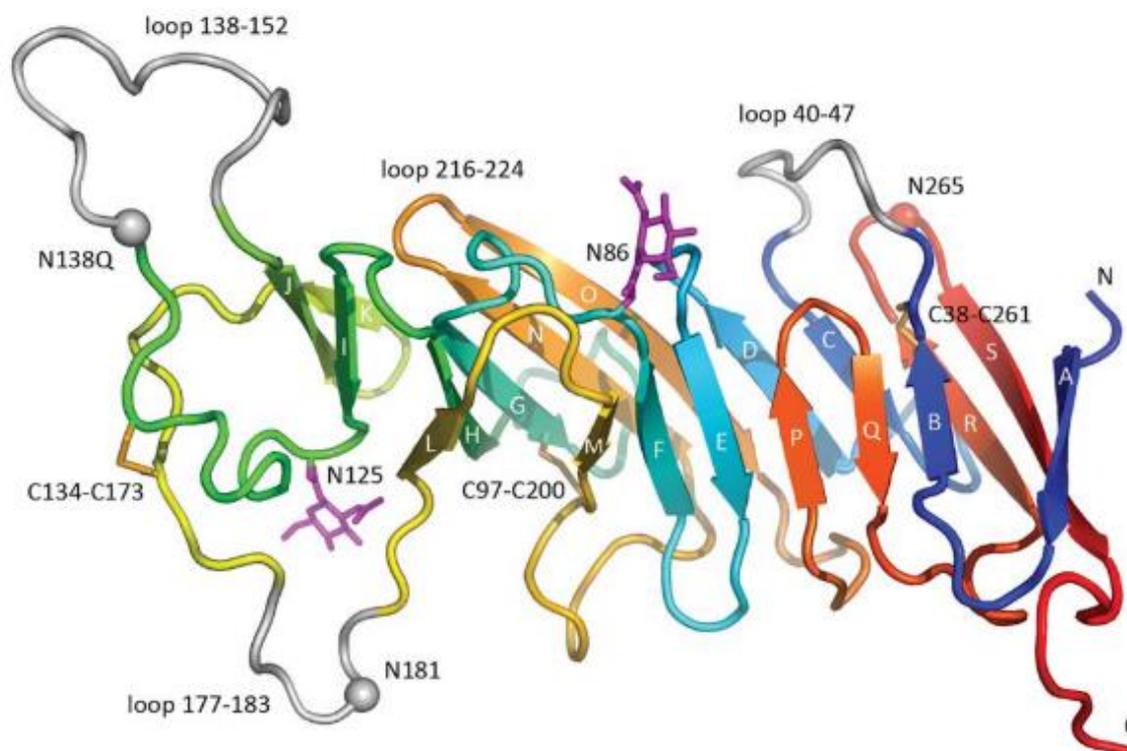
Figura 8 – Representação da estrutura do CD19. Os dois domínios extracelulares C2 semelhantes a Ig (*Ig-like*) são separados por um pequeno domínio helicoidal não semelhante a Ig com possíveis ligações dissulfeto. O domínio citoplasmático de 242 aminoácidos, altamente conservado, possui vários resíduos de tirosina, sendo três desses resíduos chave para as funções biológicas do CD19, os que estão destacados na figura.



Fonte: adaptado de WANG *et al.* (2012).

A estrutura cristalográfica do domínio extracelular do CD19 humano (resíduos 21–277 código UniProtKB CD19_HUMAN) cocristalizada com o fragmento Fab do anticorpo murino B43 (PDB 6AL5) já foi obtida a uma resolução de 3,0 Å. Esse complexo permitiu identificar o epítipo do CD19 e sugeriu que a atuação deste anticorpo seja por um mecanismo chave-fechadura de reconhecimento de antígeno (TEPLYAKOV *et al.*, 2018). O CD19 extracelular possui uma estrutura secundária de folhas- β , formada por seis folhas- β de um lado e 10 folhas- β do outro como mostrado na Figura 9, o que faz dele parte da superfamília das imunoglobulinas.

Figura 9 - Representação em *ribbon* da porção extracelular do CD19. Cores: azul (N-terminal) ao vermelho (C-terminal). Os *loops* que não foram definidos na densidade eletrônica da estrutura cristalográfica (PDB 6AL5) estão mostrados em cinza. Os carboidratos observados nos locais de N-glicosilação estão mostrados como *sticks* em magenta; outros potenciais locais de N-glicosilação mostrados como esferas. As pontes dissulfeto estão representadas em *sticks* amarelas.



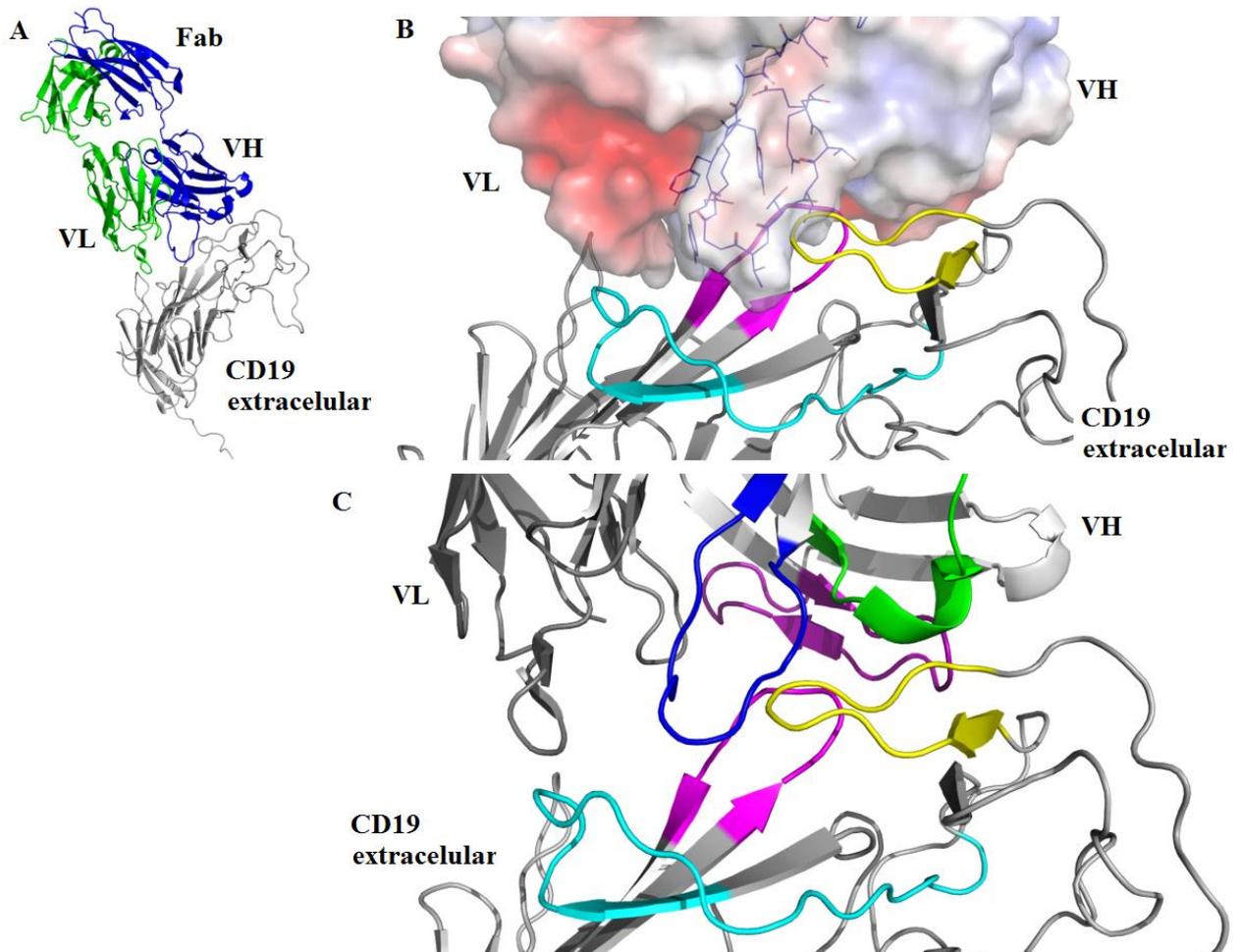
Fonte: TEPLYAKOV *et al.* (2018).

O anticorpo monoclonal B43 é uma imunotoxina anti-CD19 de alta afinidade capaz de ocasionar apoptose em células B leucêmicas, com um fenótipo resistente a drogas (MEANY *et al.*, 2015). O anticorpo Blinatumomab, uma construção de anticorpo biespecífico no formato BiTE (dispositivo biespecífico de células T), formado pelo B43 que reconhece o CD19 com uma constante de dissociação de 1.6×10^{-9} M (LEE, K. J. *et al.*, 2016), foi aprovado para uso em pacientes com precursores de Leucemia Linfoblástica Aguda (LLA) de células B refratária ou recidiva (KANTARJIAN *et al.*, 2017).

Esse anticorpo interage com o CD19 na porção central do domínio extracelular do CD19 (Figura 10A) e reconhece o epítipo, formado por 16 resíduos, compreendido de três alças (*loops*) formadas pelos resíduos 77–97 (*loop 1*), 135–146 (*loop 2*) e 196–204 (*loop 3*) (Figura 10B). O *loop 3* ocupa o bolso central na região CDR, enquanto os outros dois interagem com a CDR H3 no exterior do Fab, de modo que a CDR H3 se encaixa na cavidade entre os *loops 2* e *3*, como mostrado na Figura 10B. O paratopo do anticorpo B43 (resíduos do anticorpo em

contato direto com o CD19, usando um raio de corte de 4 Å, inclui 15 resíduos, todos da cadeia pesada. A cadeia leve não está em contato com o CD19 (Figura 10C). As três CDRs da cadeia pesada contribuem para a ligação (Figura 10C) (TEPLYAKOV *et al.*, 2018). Interações eletrostáticas desempenham um papel importante na interação Fab-CD19. A carga positiva das duas lisinas e duas histidinas do *loop* 3 do CD19 e a carga global negativa de B43, devido a oito resíduos carregados negativamente nas CDRs, promovem atração eletrostática de longo alcance (coulômbica) Fab-CD19.

Figura 10 – Estruturas do Fab do anticorpo B43 e do CD19 cocristalizados. A) Representação em *cartoon* do complexo Fab B43-CD19. CD19 está representado em cinza, cadeia pesada do Fab em azul e cadeia leve do Fab em verde. B) Interface de interação entre os domínios variáveis da cadeia leve (VL) e pesada (VH) do anticorpo B43 (superfície) e o CD19 extracelular, evidenciando os *loops* do epítipo, *loop* 1 (azul claro), *loop* 2 (amarelo) e *loop* 3 (lilás) e a região CDR H3 em linhas azuis encaixada entre os *loops* do epítipo. C) Interface de interação VL-VH e CD19, evidenciando as CDRs do domínio VH, CDR H1 (verde), H2 (roxo) e H3 (azul), e as alças do epítipo do CD19.



Fonte: elaborada pelo autor.

O CD19 é expresso na maioria das malignidades que acometem células B, incluindo Leucemia Linfoblástica Aguda, Leucemia Linfocítica Crônica e linfoma não-Hodgkin. Além disso, está ausente nas células progenitoras da medula óssea (UCKUN *et al.*, 1988) e é expresso ao longo do desenvolvimento das células B, assim, desempenha um papel crucial na manutenção do equilíbrio entre a resposta imune humoral, antígeno-induzida e na indução de tolerância. Portanto, a imunoterapia baseada em células CAR T, que possui como alvo o CD19, emerge como uma abordagem terapêutica promissora para doenças autoimunes e malignidades de células B.

1.5 Simulação de Dinâmica Molecular (DM)

A simulação por Dinâmica Molecular (DM) é uma técnica essencial para o entendimento da relação entre a estrutura e a função de moléculas biológicas. Essa técnica considera uma molécula como um modelo dinâmico em que as movimentações internas e as mudanças conformacionais de um conjunto de átomos são essenciais para a determinação de sua função (KARPLUS; MCCAMMON, 2002). A partir dos detalhes atomísticos fornecidos pela simulação de DM, é possível acessar informações detalhadas sobre a estrutura tridimensional de uma proteína em equilíbrio em um solvente ou uma enzima interagindo com um substrato, por exemplo (HANSSON *et al.*, 2002).

As simulações de DM podem facilitar a interpretação de dados experimentais, por meio do teste de modelos propostos para descrever fenômenos naturais, resultando em dados que são difíceis ou impossíveis de se obter experimentalmente (FRENKEL; SMIT, 2002; ALLEN; TILDESLEY, 2017). Além disso, a DM acessa detalhes atomísticos e descreve movimentos moleculares que as técnicas de cristalografia de raios-X, ressonância magnética nuclear, espalhamento de nêutrons e espectroscopia de fluorescência não alcançam (VAN GUNSTEREN *et al.*, 2008).

As interações entre átomos e moléculas são empiricamente descritas como uma função do potencial, da qual derivam as forças que atuam sobre cada partícula. Com o conhecimento dessas forças, é possível calcular a dinâmica dos átomos no sistema usando as equações de movimento de Newton.

A função do potencial consiste de um conjunto de equações que descrevem empiricamente as interações entre átomos ligados e não ligados (ABRAHAM *et al.*, 2013). Esta função, juntamente com o conjunto de seus parâmetros empíricos, é referida como campo de força. Este possibilita o cálculo do potencial total, V (Equação 1), do sistema a partir de

coordenadas atômicas espaciais e leva em consideração a interação entre átomos ligados, incluindo ângulos de ligação, torsionais e diedrais, e a interação entre átomos não ligados ($V_{\text{não-ligados}}$) (FRENKEL; SMIT, 2002; ALLEN; TILDESLEY, 2017).

$$V = \frac{1}{2} \sum_{\text{ligações}} K_b (b_{eq} - b)^2 + \frac{1}{2} \sum_{\text{ângulos}} K_\theta (\theta_{eq} - \theta)^2 + \frac{1}{2} \sum_{\text{torsões}} K_\varphi [1 + \cos(n\varphi)] + \sum V_{\text{não-ligados}} \quad (1)$$

Em um determinado momento (t), b é o comprimento de ligação (instantâneo) que representa o desvio em relação ao comprimento de equilíbrio b_{eq} da ligação; θ é o ângulo instantâneo entre duas ligações químicas consecutivas que descreve o desvio em relação ao ângulo de equilíbrio θ_{eq} entre as duas ligações químicas; K_b é a constante de Hooke associada à deformação da ligação química específica, e K_θ é a constante de Hooke para a restituição do ângulo de equilíbrio (θ_{eq}) entre as duas ligações químicas, visto que as distâncias de ligações e as amplitudes dos ângulos podem oscilar, considerando um modelo de oscilador harmônico, entre as posições de deformação máxima e a posição de equilíbrio; φ representa o ângulo de torção, e n é a multiplicidade do ângulo quando este passa por 360° (FRENKEL; SMIT, 2002; ALLEN; TILDESLEY, 2017).

O potencial entre átomos não ligados (Equação 2) é representado pelo somatório do potencial de Lennard-Jones (V_{L-J}) (Equação 3), que modela a interação de *van der Waals* (vdW), responsável por interações de curto alcance, com o potencial de Coulomb (Equação 4), que modela as interações eletrostáticas (V_{elet}), consideradas de longo alcance.

$$V_{\text{não-ligados}} = V_{L-J} + V_{elet} \quad (2)$$

$$V_{L-J} = \sum 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] \quad (3)$$

$$V_{elet} = \sum \frac{q_i q_j}{4\pi\epsilon_0 \epsilon_r r_{ij}} \quad (4)$$

Na Equação 3, σ_{ij} é a distância finita entre os átomos i e j , distância em que o potencial entre átomos é zero, sendo que ϵ_{ij} é a profundidade do potencial entre a barreira repulsiva e

atrativa entre os átomos i e j . Esses são parâmetros ajustados por cálculos teóricos ou experimentais. Na Equação 4, V_{elet} é referente às interações eletrostáticas, r_{ij} corresponde à distância entre os átomos i e j com as cargas, q_i e q_j pontuais, ϵ_r é a constante dielétrica relativa do meio e ϵ_0 a permissividade do espaço livre.

Na Equação 5, é possível encontrar as forças (F) que atuam sobre cada átomo i , calculando-se a derivada do potencial (V), em relação à posição (r) de cada átomo.

$$\vec{F}_i = -\frac{dV(r_1, r_2, r_3, \dots, r_N)}{dr_i} \quad (5)$$

As posições e a aceleração dos átomos no tempo t são utilizadas para calcular a posição no tempo anterior ($t - dt$) e a nova posição ($t + dt$). Dada uma configuração inicial do sistema (posições nos eixos x , y e z), as forças exercidas para cada átomo podem ser resolvidas pela derivada primeira do potencial (V), considerando o campo de força (Equação 1), em relação às posições desses átomos (Equação 5). Ao obtermos as forças que atuam sobre cada átomo, os valores de aceleração são obtidos segundo a equação de Newton (princípio fundamental da dinâmica), que é resolvida para cada um dos átomos, seguindo a Equação 6:

$$\vec{F}_i = m_i \vec{a}_i = m_i \frac{d^2 \vec{r}_i}{dt^2} \quad (6)$$

Onde F_i é a força exercida em cada átomo i , a_i é a aceleração para o átomo i que possui massa m_i e dt é o intervalo de tempo de para evoluir o sistema temporalmente.

Na DM, após determinada a aceleração, integra-se as equações de movimento de Newton por algoritmos como o de Verlet (VERLET, 1967), para obter a nova posição do átomo. As posições atômicas são gravadas em arquivos de trajetória para posterior análise. Portanto, aplicando a integração das equações de movimento sucessivamente às novas posições (Equação 7), obtidas em sucessivos intervalos de tempo dt , é possível atingir o intervalo de tempo total e consequentemente obter a trajetória atômico-molecular, ao longo do tempo. A partir da trajetória obtida da simulação, podem ser realizadas análises que permitem o acesso a informações sobre a interação entre moléculas e as energias envolvidas nessa interação.

$$r(t + dt) = 2r(t) - r(t - dt) + a(t)dt^2 \quad (7)$$

Onde as posições e acelerações dos átomos no tempo t são utilizadas para calcular $r(t - dt)$, posição no tempo anterior e $(t + dt)$, a nova posição, sendo dt o intervalo temporal para evoluir o sistema.

As etapas de simulação de DM são essenciais para se obter a trajetória das moléculas na caixa de simulação e realizar análises posteriores. Primeiramente, é obtida a configuração inicial das moléculas que irão compor a caixa de simulação. Esta configuração pode ser obtida a partir de dados cristalográficos do *Protein Data Bank* (PDB) (BERMAN *et al.*, 2000). Depois da obtenção da configuração inicial, é realizado o cálculo das forças exercidas sobre cada átomo do sistema, utilizando as informações de parâmetros do campo de força e posições atômicas. Em seguida, é necessário realizar a minimização de energia da interação entre as moléculas do sistema para reacomodá-las na caixa de simulação e evitar choques entre os átomos. Após se obter um mínimo de energia, a estrutura e a dinâmica do sistema precisam ser ajustadas para se obter o estado termodinâmico de interesse por meio da termalização ou equilíbrio do sistema. Essa etapa é controlada em *ensemble* canônico (NVT), representado por número de moléculas e volume fixos e temperatura controlada ou em *ensemble* isotérmico-isobárico (NPT) (ALLEN; TILDESLEY, 1987), no qual o número de moléculas é fixo e a pressão e temperatura são controladas para convergir com os parâmetros de referência (BERENDSEN, H J C *et al.*, 1984; HUANG, 1987; BERENDSEN *et al.*, 1995). Após a termalização, o sistema está pronto para ser simulado em *ensemble* NVT ou NPT. Por fim, são realizadas análises energéticas e estruturais usando as trajetórias obtidas.

1.6 Importância e justificativa

O Ministério da Saúde estima cerca de 10.000 novos casos de Leucemia e também cerca de 10.000 casos de Linfoma não-Hodgkin para cada ano de 2018 e 2019 no Brasil (INCA, 2018). Os pacientes são convencionalmente tratados com quimioterapia e anticorpos monoclonais (mAbs) como adjuvantes, sendo que alguns deles podem alcançar sobrevida prolongada livre da doença após o transplante de células-tronco alogênicas. Entretanto, o transplante é limitado pela significativa mortalidade relacionada ao procedimento e pela falta de doadores adequados. Em pacientes com neoplasias de células B que recidivam após o transplante alogênico de células-tronco, a infusão de linfócitos do doador alogênico pode induzir remissões. A eficácia dessas infusões de linfócitos fornece uma justificativa para tentativas de desenvolver outras imunoterapias celulares para malignidades de células B (KOCHENDERFER *et al.*, 2010).

O CD19 é um alvo promissor para terapias com células T específicas ao antígeno, pois é expresso significativamente em células B malignas, assim como em células B saudáveis e células dendríticas foliculares, porém ausente em células-tronco hematopoiéticas e em células pró-B (KOCHENDERFER *et al.*, 2010).

A imunoterapia utilizando células CAR-T anti-CD19 é um tratamento efetivo para leucemias e linfomas que acometem células B (MAKITA *et al.*, 2017). Com o objetivo de limitar as respostas imunes anti-CAR, novos CARs anti-CD19 vêm sendo construídos e estudados (ALABANZA *et al.*, 2017; SOMMERMEYER *et al.*, 2017). As sequências dos domínios variáveis (VL e VH) para compor o scFv escolhido para o CAR nesse estudo, foram retiradas do anticorpo monoclonal (mAb) murino B43 (PDB 6AL5) que já possui estrutura cristalográfica, facilitando a realização dos estudos de interação entre scFv-CD19 e entre CAR-CD19 por simulação de Dinâmica Molecular.

O *hinge* de CD8 α , o domínio transmembrana de CD8 α e intracelular, formado pelo coestimulador 4-1BB ligado ao domínio de sinalização CD3- ζ , foram utilizados para a montagem do CAR anti-CD19 nesse trabalho, pois já são utilizados no CAR comercial Kymriah[®] da empresa Novartis (SALMIKANGAS *et al.*, 2018), já utilizado para tratamento de Leucemia Linfoblástica Aguda.

Não há estudos descritos na literatura que explicam molecularmente o mecanismo de ativação de células CAR T pela transdução de sinal ou predizem a estrutura de CAR anti-CD19 completo. Assim, os componentes do CAR: scFv, *hinge*, domínio transmembrana e domínio intracelular foram modelados, a fim de propor um modelo para a estrutura completa do CAR em membrana com constituição semelhante à de células T. Além de obter informações estruturais, através de simulação de Dinâmica Molecular, relacionadas ao mecanismo de sinalização do CAR modelado, interagindo com CD19, marcador de células cancerosas.

2 OBJETIVOS

2.1 Objetivos gerais

Obter informações estruturais, através de simulação de Dinâmica Molecular, relacionadas ao mecanismo de sinalização de um CAR em membrana de células T modelado, interagindo com a proteína de membrana CD19, marcador de células cancerosas.

2.2 Objetivos específicos

- a) Construir o scFv anti-CD19 com base nas informações cristalográficas do fragmento Fab do anticorpo B43 (PDB 6AL5) e submeter o scFv obtido à simulação de DM;
- b) Construir modelos de membrana de tamanhos diferentes no programa CHARMM-GUI, suficientes para conter a porção transmembrana, o CAR e o complexo CAR-CD19;
- c) Modelar por homologia o *hinge* de CD8 α e submeter o modelo obtido à simulação de DM para, posteriormente, unir ao scFv construído, finalizando a porção extracelular do CAR;
- d) Modelar por homologia a porção transmembrana de CD8 α em modelo de membrana gerado no programa CHARMM-GUI, submeter o modelo obtido à simulação de DM e, posteriormente, unir ao scFv-*hinge*;
- e) Modelar por montagem (*threading*) a porção intracelular (4-1BB-CD3- ζ) e submeter o modelo obtido à simulação de DM. Em seguida, unir o modelo obtido à porção transmembrana do CAR;
- f) Modelar o CAR em interface membrana/água e submetê-lo à DM;
- g) Realizar análises estruturais e energéticas, utilizando as trajetórias de DM, para verificar a estabilidade dos modelos construídos em água;
- h) Modelar o sistema CAR-CD19 em interface membrana/água e submetê-lo à DM;
- i) Realizar análises estruturais e energéticas da interação CAR-CD19 a partir das trajetórias de DM obtidas;
- j) Realizar Análise de Componentes Principais (PCA) para inferir a transdução de sinal do CAR, comparando-se o CAR livre ao CAR interagindo com CD19.

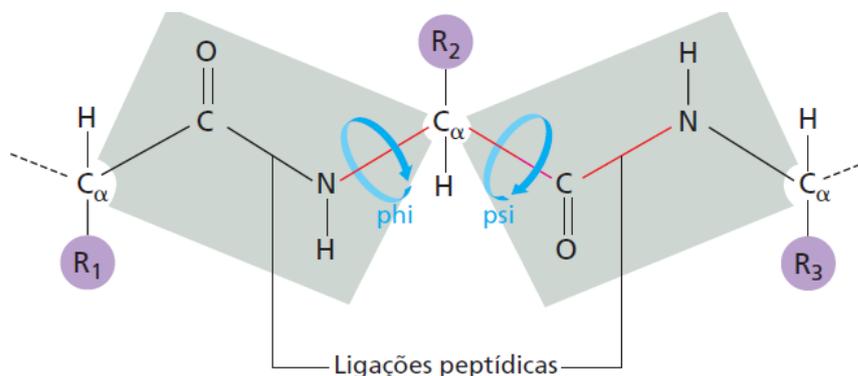
3 METODOLOGIA

3.1 Modelagem Molecular

A estrutura tridimensional (3D) de uma proteína pode ser determinada experimentalmente por Difração de raios X e Ressonância Magnética Nuclear (RMN). Essas técnicas possuem limitações, por exemplo, a obtenção de amostras em quantidade suficiente para os ensaios necessários ou a dificuldade de se conseguir cristais de qualidade para difração (FILHO; ALENCASTRO, 2003). Apesar da grande evolução nesse campo, a obtenção de estruturas tridimensionais por esses métodos ainda é um processo demorado e sem sucesso garantido. Com o aumento do número de projetos de sequenciamento genômico e tais limitações da predição experimental de estruturas proteicas, surgiu a necessidade de aumentar o uso de técnicas teóricas de predição (WATERHOUSE *et al.*, 2018). Sabe-se, ainda, que um gene pode codificar várias proteínas diferentes e, desta forma, obter a estrutura tridimensional de todas elas por meio de métodos experimentais, torna-se impossível. Porém, conhecendo-se a sequência de aminoácidos de uma proteína de interesse, e de posse de uma estrutura já resolvida, que possua certa identidade estrutural (molde), pode-se prever a estrutura 3D de uma proteína ainda não resolvida. A essa técnica dá-se o nome de modelagem comparativa ou modelagem por homologia. Essa técnica utiliza parâmetros observados durante a evolução, como a semelhança estrutural e funcional observada quando há alta similaridade entre sequências, a presença de regiões conservadas em proteínas homólogas (principalmente estrutura secundária) e os “*loops*” ou regiões de alça em que se ocasionam as principais diferenças estruturais entre proteínas homólogas (FILHO; ALENCASTRO, 2003).

A avaliação da qualidade dos modelos é uma importante etapa no processo de modelagem por homologia. Quando um modelo é predito através dessa abordagem, é necessário avaliar diferentes propriedades de organização estrutural, como a estereoquímica, a confiabilidade do enovelamento conforme os ambientes químicos em que os resíduos se encontram e o empacotamento de resíduos (FILHO; ALENCASTRO, 2003; SILVA; SILVA, 2007). A qualidade estereoquímica dos modelos é verificada através de parâmetros como comprimento das ligações, ângulos entre ligações, ângulos torsionais (Φ , Ψ e Ψ), representados na Figura 11. e são avaliados ao assumirem conformações permitidas no diagrama de Ramachandran (RAMACHANDRAN *et al.*, 1963). São verificados também os maus contatos (impedimentos estéricos), a quiralidade dos aminoácidos, desvios em relação à geometria tetraédrica dos carbonos α ($C\alpha$), dentre outros.

Figura 11 - Representação da ligação peptídica e dos ângulos diédricos de rotação Phi e Psi.



Fonte: ALBERTS *et al.* (2017).

Uma vez que não há estrutura tridimensional já descrita do CAR, foram construídos modelos baseados em estruturas tridimensionais já resolvidas e depositadas em banco de dados (BERMAN *et al.*, 2000) para cada um de seus domínios, através das técnicas de modelagem comparativa e modelagem por *threading*. Os programas utilizados para a modelagem foram MODELLER (WEBB; SALI, 2016) e I-TASSER (YANG *et al.*, 2015).

O MODELLER é um programa disponível para uso livre que necessita da interação do usuário, através do uso de *scripts*. Ele utiliza um método de modelagem comparativa pela satisfação de restrições espaciais com base no alinhamento da sequência a ser modelada com sequências de estruturas já conhecidas (ŠALI; BLUNDELL, 1993). O algoritmo de programação dinâmica local é utilizado para identificar sequências relacionadas (SMITH; WATERMAN, 1981). O melhor modelo é selecionado através do valor da função objetiva MODELLER, DOPE *score* (SHEN; SALI, 2006) ou SOAP (DONG *et al.*, 2013). Então, informações geométricas e de restrições são extraídas do molde e as restrições estereoquímicas são obtidas do campo de força de mecânica molecular CHARMM-22 (MACKERELL *et al.*, 1998). A seleção das melhores informações acerca da geometria da sequência de referência é baseada em análises estatísticas de um conjunto representativo de todas as estruturas de proteínas conhecidas (WEBB; SALI, 2016). O modelo construído é otimizado, para minimização das violações das restrições espaciais, utilizando o algoritmo *conjugate gradient* (HESTENES; STIEFEL, 1952) e Dinâmica Molecular.

O I-TASSER é um programa, alocado em um servidor online, de predição de estrutura e função de proteínas. Ele gera um modelo completo de proteína através da remoção de fragmentos contínuos encontrados após os alinhamentos com as estruturas molde (molde) e

remontando-os, utilizando simulações de Monte Carlo *replica-exchanged*. Para cada estrutura que se quer modelar, as simulações do I-TASSER geram um grande conjunto de conformações estruturais, denominadas de *decoys*, que são geradas durante a simulação e agrupadas pelo programa SPICKER em *clusters* com base na similaridade da estrutura. Os cinco principais *clusters* de configurações são selecionados para gerar os modelos atômicos completos. A densidade do *cluster* é definida como o número de *decoys* da estrutura em uma unidade de espaço no respectivo *cluster*. Uma densidade de *cluster* mais alta significa que a estrutura ocorre com mais frequência na trajetória da simulação e, portanto, significa um modelo de melhor qualidade (ZHANG, 2008; ROY *et al.*, 2010; YANG *et al.*, 2015).

A confiabilidade de cada modelo é medida quantitativamente pelo *C-score*, que é calculado com base no alinhamento às estruturas molde e nos parâmetros de convergência das simulações de montagem (*threading*) da estrutura. O *C-score* é obtido na faixa de [-5, 2], onde um *C-score* de maior valor significa um modelo com maior confiabilidade. Os parâmetros de *TM-score* e RMSD, que medem a similaridade entre duas estruturas, são estimados com base no *C-score* e são geralmente usados para medir a precisão da modelagem da estrutura. Um $TM > 0.5$ indica um modelo de topologia correta e uma pontuação de $TM < 0.17$; significa uma similaridade aleatória.

Existem vários programas que analisam os melhores modelos escolhidos, mas os modelos obtidos nesse trabalho foram submetidos nos programas PROCHECK (LASKOWSKI *et al.*, 1993) para validação da estrutura no gráfico de Ramachandram, determinando se seus resíduos se encontram em regiões favoráveis ou permitidas. No gráfico, estas áreas são representadas pelas cores mais escuras. Um modelo de boa qualidade deve apresentar mais de 90% de seus resíduos nessas regiões. Os modelos também foram submetidos no programa *MolProbity* (DAVIS *et al.*, 2007; CHEN *et al.*, 2010), que proporciona uma análise detalhada de contato de todos os átomos em qualquer impedimento estérico. Os valores são fornecidos em contagem bruta e em percentil. O *clash score* é o número de sobreposições estéricas graves por 1000 átomos. Esse valor é dado em percentil, que é uma medida utilizada para dividir uma amostra de valores, ordenados de forma crescente, em cem partes, assim, o centésimo percentil é o melhor entre estruturas de resolução comparável e o 0 percentil é o pior. O *MolProbity score* combina as avaliações de *clashscore*, rotâmeros e *Ramachandran* em uma única pontuação, normalizada para estar na mesma escala da resolução de raios-X. Um modelo de boa qualidade deve apresentar uma combinação de parâmetros, como *clash score* e *MolProbity score* mais próximo de 100° percentil, menor número de rotâmeros ruins e menos de 0.05% de resíduos em regiões não permitidas, por exemplo.

Em seguida, os modelos foram submetidos à minimização e termalização para se obter as estruturas de menor energia através de Mecânica Molecular (MM) e Dinâmica Molecular (DM).

3.2 Simulação de Dinâmica Molecular e análises

As simulações de DM seguiram os protocolos incluindo etapas de minimização de energia, termalização e fase de aquisição da trajetória que foram realizadas utilizando o pacote GROMACS versão 2018, (BERENDSEN *et al.*, 1995; LINDAHL; HESS, 2001; SPOEL, VAN DER *et al.*, 2005; HESS *et al.*, 2008; PRONK *et al.*, 2013; ABRAHAM *et al.*, 2015; PALL *et al.*, 2015). O campo de força utilizado para descrever as interações atômicas foi o CHARMM36 (HUANG; MACKERELL, 2013). As coordenadas atômicas das estruturas do scFv, do *hinge*, dos domínios transmembrana e intracelular, e do CD19 foram submetidas no servidor H++ (GORDON *et al.*, 2005; MYERS *et al.*, 2006; ANANDAKRISHNAN *et al.*, 2012) para determinar a protonação dos resíduos de histidina a um pH de 7,4 (pH fisiológico). Os terminais de todas as proteínas foram determinados como carregados.

A etapa de minimização de energia do sistema usou o algoritmo *steepest descent* (MORSE; FESHBACH, 1953; ARFKEN; WEBER, 1985), visando melhorar a geometria da proteína e reduzir a energia potencial total da mesma, e seguiu-se com simulações curtas.

Na DM, os comprimentos de ligação envolvendo os átomos de hidrogênio das proteínas foram restringidos utilizando o algoritmo LINCS (HESS *et al.*, 1997; HESS, 2008) e as ligações das moléculas de água, utilizando o algoritmo SHAKE (RYCKAERT *et al.*, 1977). O modelo de água escolhido para todos os sistemas foi o de água explícitas no modelo TIP3P (JORGENSEN *et al.*, 1983). As interações entre átomos não ligados, de van der Waals e Coulomb, foram avaliadas com um raio de corte de 1.3 nm, sendo que as correções para interações eletrostáticas foram calculadas utilizando o método *Particle Mesh Ewald* (PME) (DARDEN *et al.*, 1993).

Foram usadas para termalização dos sistemas DMs curtas e sucessivas nos *ensembles* NVT (BERENDSEN, H. J. C. *et al.*, 1984; TAYLOR *et al.*, 1984; HOOVER, 1985; BERENDSEN *et al.*, 1995) e NPT (PARRINELLO; RAHMAN, 1981; BERENDSEN, H. J. C. *et al.*, 1984; HUANG, 1987; BERENDSEN *et al.*, 1995b; TAYLOR *et al.*, 2013) a 310 K. A temperatura foi controlada pelo termostato *V-rescale* com a proteína, membrana, solvente e íons acoplados separadamente (BUSSI *et al.*, 2007), com uma constante de acoplamento de 1 ps. Da mesma forma, a pressão do sistema foi controlada pelo barostato de Berendsen (BERENDSEN,

H J C *et al.*, 1984; BERENDSEN *et al.*, 1995), nas etapas de termalização, ou Parrinello-Rahman (PARRINELLO; RAHMAN, 1980. 1981), nas etapas de produção e aquisição de trajetórias, com uma constante de 5 ps a 1 bar. A concentração salina de cada sistema simulado foi a fisiológica, de 0.150 M. A aquisição das trajetórias foi realizada em DM no *ensemble* NPT e $dt = 2$ fs. As coordenadas da DM foram gravadas a cada 100 ps. Durante a etapa de termalização, as posições dos átomos dos componentes do CAR, do CD19 e da membrana foram restringidas com *posre*, aplicando-se uma força decrescente ao longo da termalização. No sistema CAR-CD19, as posições dos átomos do CD19 e de resíduos do scFv do CAR, localizados na interface de interação com o CD19, foram mantidos restritos também na etapa de produção para observar a movimentação do *hinge*, dos domínios transmembrana e intracelular.

A trajetória foi analisada utilizando programas do pacote GROMACS 2018 (BERENDSEN *et al.*, 1995; LINDAHL; HESS, 2001; SPOEL, VAN DER *et al.*, 2005; HESS *et al.*, 2008; PRONK *et al.*, 2013; ABRAHAM *et al.*, 2015; PALL *et al.*, 2015).

As variações na estrutura das biomoléculas em solução foram avaliadas pelo desvio quadrático médio (RMSD – *Root Mean Square Deviation*) dos átomos, pelo programa *gmx rms*. O RMSD em função do tempo foi obtido após a sobreposição dos carbonos alfa ($C\alpha$), de cada estrutura ao longo do tempo da DM (t_2), em relação à estrutura no tempo inicial de simulação (referência) (t_1). O cálculo do RMSD provém da média do desvio das distâncias entre as posições dos átomos dos $C\alpha$ das duas estruturas em tempos distintos. Cada uma das estruturas configuracionais amostradas ao longo da DM são comparadas com a de referência, resultando em RMSDs ao longo do tempo, seguindo a Equação 8:

$$RMSD(t_1, t_2) = \left[\frac{1}{M} \sum_{i=1}^N m_i \|r_i(t_1) - r_i(t_2)\|^2 \right]^{\frac{1}{2}} \quad (8)$$

Sendo $M = \sum_{i=1}^N m_i$, m_i corresponde à massa do átomo i , $r_i(t)$ é a posição do átomo i no tempo t de referência, t_1 , e nos tempos posteriores, t_2 . Os RMSDs médios e os desvios foram calculados a partir do momento em que se considerou que a estrutura alcança o equilíbrio (*teq*).

O Raio de giro ou Raio de giração (Rg) foi calculado pelo programa *gmx gyrate*, segundo a Equação 9, para medir possíveis alterações estruturais na estrutura do *hinge* nos sistemas CAR e CAR-CD19. A curva de Rg em equilíbrio revela o alto grau de compacidade

proteica, reflexo do enovelamento estável da proteína.

$$Rg = \left[\frac{\sum_i \|r_i\|^2 m_i}{\sum_i m_i} \right]^{\frac{1}{2}} \quad (9)$$

O programa *gmx energy* foi utilizado para obtenção dos perfis de energia configuracional, que incluem o potencial eletrostático (Equação 4) e de Lennard-Jones (Equação 3) entre átomos do scFv e CD19, dentro de um raio de corte de 1.3 nm. O potencial de interação intermolecular (PII) médio por resíduos entre os resíduos do scFv e os do CD19 dentro de um raio de corte de 0.5 nm, foi obtido com os respectivos desvios, utilizando um programa *homemade* que lê os dados de potencial gerados pelo pacote GROMACS.

A análise de *cluster* (DAURA *et al.*, 1999) que considera a representatividade das estruturas durante a simulação foi calculada pelo programa *gmx cluster* para as estruturas dos componentes do CAR, com objetivo de determinar as estruturas utilizadas na montagem do CAR e do complexo CAR-CD19. O programa *gmx cluster* pode agrupar estruturas com base nas distâncias entre essas estruturas, determinadas a partir da trajetória de simulação. O RMSD após a sobreposição das estruturas ou o RMSD das distâncias entre pares de átomos podem ser usados para definir a distância entre as estruturas. Foram utilizados os algoritmos *single linkage*, que adiciona uma estrutura a um cluster quando sua distância a qualquer elemento do cluster for menor que o raio de corte definido; e *gromos*, que contabiliza o número de vizinhos usando o raio de corte. A estrutura com o maior número de vizinhos e seus vizinhos, presentes dentro do raio de corte, são adicionados a um cluster.

O programa *gmx distance* foi utilizado para calcular: a) a distância entre o centro de massa do scFv e o C α do resíduo Trp da extremidade extracelular da α -hélice do domínio transmembrana (Distância A); b) a distância em z do centro de massa do scFv e do C α do resíduo Trp da extremidade extracelular da α -hélice do domínio transmembrana (Distância B); c) a distância em função do tempo entre os C α dos resíduos Trp e Leu das extremidades extra e intracelular da α -hélice do domínio transmembrana (Distância C); d) a distância em z dos C α dos mesmos resíduos (Distância D). O ângulo formado entre dois vetores em função do tempo de simulação foi calculado utilizando o programa *gmx gangle*. O ângulo β formado entre o vetor sobre a posição do centro de massa do scFv e do C α do resíduo Trp da α -hélice do domínio transmembrana (vetor O) e o vetor sobre o eixo z (vetor M). O ângulo α formado entre o vetor

definido pela posição dos C α dos resíduos Trp e Leu das extremidades extra e intracelular da α -hélice do domínio transmembrana (Vetor N) e o outro vetor pelo eixo z (Vetor M), possibilitando inferir a inclinação da α -hélice em relação ao eixo z.

Os gráficos apresentados foram gerados usando o programa OriginPro 8.0 (ORIGINLAB CORPORATION). As representações de estruturas dos sistemas analisados foram geradas usando o programa Pymol 2.2.0 (DeLano Scientific) (DELANO, 2002).

3.3 Dinâmica essencial de proteínas

A dinâmica essencial (DE) é uma técnica que utiliza a análise de componentes principais, que identifica a natureza e a importância relativa dos modos de deformação essenciais de uma macromolécula, a partir de amostras de trajetórias de DM (MEYER *et al.*, 2006). A matriz de covariância cartesiana original, que contém as posições dos átomos nos 3 eixos de coordenadas sobre a estrutura média, é diagonalizada para se obter um conjunto de autovetores e autovalores. Os autovetores descrevem a natureza dos movimentos de deformação no espaço cartesiano, enquanto os autovalores representam a quantidade de variância explicada por cada movimento. Os autovetores definem um conjunto de bases completas e ortogonal e, portanto, qualquer instante na trajetória pode ser reproduzido exatamente neste novo conjunto de bases 3N-6, em que N é o número de átomos do sistema (Equação 10).

$$\{R\}_{x, y, z} \rightarrow \{P\}_v \quad (10)$$

Onde R representa as coordenadas cartesianas originais (x, y, z) e P representa as projeções nos autovetores 3N-6 (v), que são definidos para maximizar a quantidade de variância explicada em uma ordem decrescente.

Para proteínas e ácidos nucleicos, o número de autovetores necessários para explicar 95 a 99% da variância total é muito menor que o 3N-6. Assim, se apenas um conjunto de autovetores, M, descrevem 95% da variância total, as projeções das coordenadas cartesianas originais ao longo do conjunto de autovetores importantes contêm quase todas as informações originais e de uma maneira muito mais comprimida (MEYER *et al.*, 2006).

A Análise de Componentes Principais (PCA) realiza uma transformação linear de dados e pode ser usada para reduzir dados multidimensionais a duas ou três dimensões, facilitando a análise da trajetória de simulação. Muitos conjuntos de grande quantidade de dados contêm

correlações entre as dimensões, assim uma parte dos dados é redundante. Assim, o método de PCA transforma os dados para que o máximo de variação possível seja compactado em menos dimensões possíveis. Isso permite que os dados sejam comprimidos, ignorando informações não relevantes (WALLISCH *et al.*, 2014). Assim, o objetivo dessa técnica é reduzir a dimensionalidade do conjunto de dados, agrupando variáveis correlacionadas e substituindo-as por novos descritores, chamados de componentes principais (PCs), nas quais os dados são projetados. As PCs são variáveis completamente não correlacionadas, obtidas por combinação linear simples dos dados originais. Além disso, as PCs contêm a maior parte da variabilidade dos dados originais em um espaço dimensional reduzido (FERREIRA *et al.*, 1999).

Recentes avanços em algoritmos, campos de força e poder computacional aumentaram a utilização de simulações de Dinâmica Molecular (DM) com objetivo de se ter uma melhor compreensão do comportamento estrutural e dinâmico das biomoléculas. Embora a DM esteja se tornando uma ferramenta padrão e as trajetórias estejam sendo coletadas para sistemas cada vez maiores e para tempos de simulação mais longos, os arquivos de trajetória gerados também são cada vez maiores, podendo sobrecarregar os sistemas de armazenamento em disco. Assim, técnicas de redução de dimensionalidade, como PCA, estão sendo cada vez mais aplicadas à análise de dados de simulação molecular (AMADEI *et al.*, 1993; WLODEK *et al.*, 1997; PAPAEO *et al.*, 2009; NG *et al.*, 2013), uma vez que permite observar a convergência do comportamento dinâmico de um sistema entre réplicas e a conservação ou perturbação desse sistema pela interação com um ligante (ROY; LAUGHTON, 2010). A técnica de PCA permite que os modos dominantes de flexibilidade molecular sejam identificados, de maneira rigorosa, e apresentados na forma de variações nos valores de um pequeno número de coordenadas coletivas (PC1 e PC2), em vez das coordenadas cartesianas independentes (3N) de cada átomo, facilitando a interpretação e a visualização dos resultados de DM.

A análise de PCA foi realizada utilizando o pacote de programa pyPcazip, formado por um conjunto de ferramentas baseado na compactação e análise de dados de simulação molecular, sem perda significativa de informação (SHKURTI *et al.*, 2015). Os dados compactados podem ser analisados por uma variedade de métodos que produzem métricas objetivas, quantitativas e comparativas relacionadas à convergência e amostragem, bem como métricas sobre a similaridade entre trajetórias de simulação molecular.

A trajetória da DM foi compactada, utilizando PCA, seguindo-se os seguintes passos: a) cálculo da estrutura média e sobreposição de cada configuração obtida ao longo da trajetória a essa estrutura média; b) cálculo da matriz de covariância de coordenadas cartesianas dos dados sobrepostos; c) diagonalização dessa matriz de covariância para obter um conjunto de N

autovetores e autovalores; d) seleção dos M autovetores principais que capturam uma porcentagem escolhida da variância total (por exemplo, 90%); e) cálculo da projeção correspondente no subespaço M-dimensional e registro de autovetores, autovalores e das projeções selecionadas (PC1 e PC2) para cada configuração da trajetória no arquivo de saída (SHKURTI *et al.*, 2015).

As entradas do programa são compatíveis com os formatos de arquivo de trajetória gerados pela maioria dos pacotes de simulação de DM atuais, como AMBER, CHARMM, GROMACS e NAMD, e é paralelizado para permitir o processamento eficiente de grande quantidade de dados.

4 PROCEDIMENTOS EXPERIMENTAIS

4.1 Modelagem dos componentes do CAR

4.1.1 Modelagem do scFv anti-CD19

A estrutura do scFv é constituída pelos domínios variáveis das cadeias leve e pesada (VL e VH, respectivamente) do anticorpo conectadas por uma cadeia polipeptídica (*linker*). As sequências de VL e VH (Tabela 1) foram obtidas a partir do Fab do anticorpo B43, depositado com o código 6AL5 no *Protein Data Bank* (Figura 10A). A estrutura do *linker*, denominado de *Whitlow linker*, constituído de 18 aminoácidos (Tabela 1), foi modelada no servidor I-TASSER (YANG *et al.*, 2015), onde foi também conectada aos domínios variáveis. O *linker* iniciou no C-terminal do domínio VL e finalizou no N-terminal do VH. Foram gerados 5 modelos de scFv a partir de 10 moldes.

Tabela 1 - Sequências de aminoácidos dos domínios variáveis do anticorpo B43 e do *linker* que compõem o scFv com a respectiva numeração dos resíduos e dos átomos de acordo com o arquivo do sistema em formato PDB submetido à simulação de DM.

	Resíduos	Átomos	Sequência
Domínio VL*	1-111	1-1663	DIQLTQSPSFLSASVGDRVTITCKASQSVDYSG DSYLNWYQQKPGKAPKLLIYDASNLVSGVPSR FSGSGSGTEFTLTISLQPEDFATYYCQQSTENP WTFGGGTKLEIK
<i>linker</i>	112-129	1664-1868	GSTSGSGKPGSGEGSTKG
Domínio VH*	130-252	1869-3737	EVQLVQSGAEVKKPGSSVKVSKASGYAFSSY WMNWRQAPGQGLEWMGQIWPGDSDTNYA QKFQGRVTITADESTSTAYMELSSLRSEDVAVY YCARRETTTVGRYYYAMDYWGQGTITVTVSS

*Sequências obtidas do *Protein Data Bank* (PDB).

Fonte: elaborada pelo autor.

4.1.2 Modelagem do hinge de CD8 α

A sequência do *hinge* de CD8 α está apresentada na Tabela 2. Como não há estrutura para o *hinge* de CD8 α depositada no banco de dados do PDB, seu modelo foi obtido por modelagem comparativa *multi*-molde pelo programa *Modeller 9.18* (WEBB; SALI, 2016). Foram selecionados 5 moldes, 3DMM, 3VU2, 3AML, 1QB4 e 2R49, que apresentaram identidade acima de 30 %, 80; 72; 64.2; 53 e 48.3%, respectivamente, a partir da busca realizada no programa *blastp*, utilizando o banco de dados PDB. Foram gerados 200 modelos.

4.1.3 Modelagem do domínio transmembrana de CD8 α

Também não há estrutura depositada no banco de dados do PDB para o domínio transmembrana de CD8 α utilizado para o CAR em estudo, assim foi realizada uma busca no *blastp*, usando o banco de dados PDB, para encontrar moldes com identidade suficiente para a realização da modelagem comparativa. Desse modo, o domínio transmembrana de CD8 α (Tabela 2) foi modelado por homologia pelo programa *Modeller 9.18* (WEBB; SALI, 2016), usando a abordagem de *multi-templates*. Foram utilizados 5 moldes, 1BD9, 3V2W, 6IWD, 6BWI e 5U1D, que apresentaram identidade acima de 30 %, 85.7; 81.8; 64.3; 50 e 47.6%, respectivamente. Foram gerados 200 modelos.

Após escolha e validação do melhor modelo, ele foi submetido no programa CHARMM-GUI para ser inserido em modelo de membrana com composição de lipídeos semelhante à membrana de linfócitos T humanos, utilizando-se os parâmetros do próprio programa. A composição da membrana de linfócitos T humanos foi utilizada como referência para construir o modelo de membrana, determinando-se as porcentagens de cada tipo de ácido graxo que compõe os fosfolipídeos de membrana. Sendo a composição dos ácidos graxos da membrana de 43 % de fosfatidilcolina (PC); 28 % de fosfatidiletanolamina (PE), 9 % de fosfatidilserina (PS) e 8 % de fosfatidilinositol (PI) (BOUGNOUX *et al.*, 1985). Foi determinada também a relação entre ácidos graxos insaturados e saturados; sendo 0.64 para PC e 1.72 para PE. As contribuições de PS e PI foram ignoradas, pois além de pequenas comparadas a PC e PE, possuem carga negativa, o que pode gerar artefatos na simulação e maior custo computacional.

O sistema contendo o domínio transmembrana em membrana e água foi submetido à minimização e termalização para se obter a estrutura proteica de menor energia em interface

membrana/água através de DM.

4.1.4 Modelagem do domínio intracelular formado por 4-1BB e CD3- ζ

O domínio intracelular do CAR, formado pelo domínio coestimulador 4-1BB e o domínio de sinalização CD3- ζ , também não possui estruturas depositadas no PDB. Então foi feita a busca no *blastp* por *templates* que pudessem ser utilizados na modelagem comparativa, mas não foram encontrados moldes com identidade acima de 30 %. Assim, foi utilizado o programa I-TASSER (YANG *et al.*, 2015) que realiza modelagem por montagem (*threading-múltiplo*) para modelar esse domínio a partir de partes de estruturas já depositadas em banco de dados próprio. Foram gerados 5 modelos com base em 10 moldes. A sequência dos domínios 4-1BB e CD3- ζ submetida no programa está presente na Tabela 2.

Tabela 2 – Sequência de resíduos dos componentes do CAR comercial Kymriah[®] usados na modelagem dos domínios espaçador (*hinge*), transmembrana e intracelular do CAR anti-CD19 em estudo.

Domínio	Sequência
Hinge (CD8α)	TTTPAPRPPTPAPTIASQPLSLRPEACRPAAGGAVHTRGLDFACD
Transmembrana (CD8α)	IYIWAPLAGTCGVLLLSLVITLYC
Intracelular (4-1BB e CD3-ζ)	KRGRKLLLYIFKQPFMRPVQTTQEEDGCSCRFPEEEEGGCEL RVKFSRSADAPAYQQGQNQLYNELNLGRREEYDVLDKRRGRD PEMGGKPRRKNPQEGLYNELQKDKMAEAYSEIGMKGERRRGK GHDGLYQGLSTATKDTYDALHMQUALPPR

Fonte: elaborada pelo autor.

4.1.5 Validação de cada modelo obtido

Cada modelo obtido foi validado nos programas MolProbity (Duke Biochemistry) e PROCHECK. No MolProbity, foram observados parâmetros como os valores de *clash score* em percentil, de MolProbity *score* de rotâmeros ruins e de ligações e ângulos em conformações não desejadas e o número de resíduos fora de regiões permitidas (*outliers*) no gráfico de Ramachandran. Já no programa PROCHECK, foi determinado se os resíduos de cada modelo avaliado se encontram em regiões favoráveis ou permitidas. No gráfico, estas áreas são representadas pelas cores mais escuras.

4.2 Simulação de Dinâmica Molecular e análises

4.2.1 Dinâmica Molecular do scFv

O sistema foi formado pelo scFv, 25205 moléculas de água, 74 íons Na^+ e 73 íons Cl^- , em uma caixa cúbica de 9.4 nm de aresta. O seguinte protocolo foi utilizado para termalização do sistema e preparação para a fase de simulação de DM: 10000 passos de minimização de energia para otimização da geometria da proteína e redução de maus contatos. Em seguida, foi realizada a termalização para relaxar a estrutura do scFv e termalizar as moléculas de água adicionadas ao sistema, consistindo em 12.5 ps, com o scFv restrito e a água livre em *ensemble* NVT; 5 ps, com o scFv livre, também em NVT; e 80 ps em *ensemble* NPT. A etapa de aquisição da trajetória seguiu por 300 ns em *ensemble* NPT.

O *linker* foi desconsiderado para os cálculos de RMSD, pois flutuações consideráveis já são esperadas devido a sua característica flexível. A influência da oscilação na estrutura de cada domínio (VL e VH) foi avaliada por cálculos de RMSD após a sobreposição individual desses domínios.

A análise de *cluster* foi realizada no programa *gmx cluster*. Para a estrutura do scFv, foi utilizado um raio de corte de 0.18 nm, em que foram consideradas estruturas que possuem uma distância de até 0.18 nm em um único *cluster*.

4.2.2 Dinâmica Molecular do scFv-CD19

A partir das estruturas do scFv e do CD19 equilibradas em água, avaliou-se a interação entre elas, com base nas coordenadas cristalográficas depositadas no PDB, código 6AL5. A estrutura final do scFv foi sobreposta à estrutura cristalográfica dos domínios VL e VH para o melhor posicionamento com o CD19, respeitando a conformação do sítio de interação e as regiões de CDR descritas na Figura 13.

A caixa de simulação, com dimensões de 12.8, 11.8 e 13.8 nm (x y z), abrigou o sistema composto pelo scFv, o CD19, 68449 moléculas de água explícitas, 198 íons Na^+ e 197 íons Cl^- , adicionados para manter a neutralidade de cargas.

O seguinte protocolo foi utilizado para termalizar e preparar o sistema para a fase de produção das simulações de DM: 5000 passos de minimização de energia para otimização da geometria do complexo scFv-CD19 e redução dos maus contatos, utilizando *steep*; seguindo-se de 7.5 ps NVT, mantendo scFv e CD19 fixos; 5 ps NVT com as cadeias principais do scFv e CD19 fixas; e finalmente 15 ps NPT com o sistema livre de restrições de posição. A simulação

de DM foi executada em NPT durante 600 ns.

4.2.3 Dinâmica Molecular do hinge

Após a modelagem do *hinge* de CD8 α , o modelo escolhido e validado foi inserido em uma caixa cúbica de 7.64 nm de aresta que foi preenchida com 13848 moléculas de água, 40 íons NA⁺ e 41 íons Cl⁻. O sistema foi minimizado por 10000 passos. Em seguida, foram realizadas simulações curtas de DM para termalizar a estrutura do *hinge* modelado em água. Inicialmente, o sistema foi simulado em NVT por 12.5 ps com as coordenadas da proteína congeladas. Em seguida, em NPT por 15 ps com a proteína livre. Após a termalização, o sistema foi submetido à etapa de produção por DM durante 400 ns.

4.2.4 Dinâmica Molecular do domínio transmembrana

Foram construídas três modelos de membranas, de tamanhos diferentes, seguindo-se a mesma composição de fosfolipídios. Cada um deles recebe um conjunto de elementos do CAR para serem simulados, com a região transmembrana inserida nesses modelos de membrana. O modelo transmembrana9-9 abriga o domínio transmembrana e o *hinge* em um sistema, e o domínio transmembrana e o intracelular em outro, e possui dimensão de 9 x 9 nm. O modelo transmembrana12-12 possui dimensão de 12 x 12 nm e a região transmembrana inserida ancora as porções extracelulares scFv-*hinge* (scFv-*hinge*-transmembrana) e o CAR (scFv-*hinge*-transmembrana-intracelular). No caso do CAR em interação com o CD19, foi usado o modelo de membrana de dimensão 14 x 14 nm, denominado de transmembrana14-14.

O modelo do domínio transmembrana selecionado e validado foi inserido em cada uma das membranas montadas no próprio programa CHARMM-GUI (JO *et al.*, 2009; WU *et al.*, 2014; LEE, J. *et al.*, 2016) para, em seguida, minimizar e termalizar a estrutura em membrana e água. A faixa de água que preenche o sistema é determinada também no próprio programa, tendo sido escolhido 1.5 nm para cima e 1.5 nm para baixo de cada membrana construída. Após a montagem do sistema transmembrana-membrana-água, foram realizadas duas etapas de minimização utilizando o algoritmo *steepest descent* e *conjugate gradient* por 20000 passos cada etapa, mantendo-se a posição da proteína e da membrana restritas. Em seguida, foram realizadas simulações curtas em NVT durante 50 ps, ainda com a proteína e a membrana restritas, depois 225 ps em NPT, reduzindo a força de restrição na proteína e membrana, e, por fim, 200 ps em NPT com a proteína e membrana livres. Após as etapas de simulações curtas,

cada membrana foi simulada durante 200 ns para equilibrar em água.

Após a equilibração das membranas e dos domínios transmembrana, as estruturas finais da simulação foram utilizadas para a montagem dos sistemas e para as simulações seguintes.

4.2.5 Modelagem e Dinâmica Molecular do *hinge*-transmembrana

Após a equilibração do *hinge* em água e do transmembrana em membrana e água as duas estruturas foram conectadas através da ligação entre o resíduo ácido aspártico no C-terminal do *hinge* e o resíduo isoleucina no N-terminal do transmembrana no programa *Pymol* 2.2.0 (DeLano Scientific), cuja distância de ligação e ângulos foram estabelecidos usando a etapa de minimização de energia com algoritmo *steepest descent*. A caixa com dimensões 9 x 9 x 12 nm contendo os domínios *hinge*-transmembrana em membrana foi preenchido com 17019 moléculas de água, 77 íons Na⁺ e 78 íons Cl⁻. Inicialmente, foi minimizado por 20000 passos com as posições da membrana restritas. Em seguida, foi termalizado por 7.5 ps em NVT e por 106 ps em NPT com a membrana ainda restrita. Após a termalização, o sistema foi simulado por DM durante 200 ns.

4.2.6 Modelagem e Dinâmica Molecular do *scFv*-*hinge*-transmembrana

Após simular o *scFv* em água e os domínios *hinge*-transmembrana em membrana e água, o C-terminal do *scFv* foi aproximado ao N-terminal do *hinge* no programa *Pymol* 2.2.0 e, em seguida, conectados na minimização de energia, finalizando o domínio extracelular do CAR. O *scFv* foi selecionado com base na análise de *cluster*, em que a estrutura final presente no *cluster* mais representativo durante a simulação do *scFv* em água foi escolhida. A configuração *hinge*-transmembrana foi determinada com base na posição do *hinge*, sendo escolhida a que o *hinge* estava em uma posição mais esticada para facilitar a conexão com o *scFv*, como já falado anteriormente. O sistema com dimensões de 12 x 12 x 15 nm contendo *scFv*-*hinge*-transmembrana e membrana foi preenchido com 56686 moléculas de água, 153 íons Na⁺ e 153 íons Cl⁻. Inicialmente, foi minimizado por 10000 passos com as posições da proteína e membrana restritas. Em seguida, foi termalizado por 50 ps em NVT com proteína e membrana ainda restritas. Finalizando a termalização com 225 ps em NPT reduzindo-se a força de restrição na proteína e na membrana e 100 ps também em NPT com a proteína e membrana livres. Após a termalização, o sistema foi simulado por DM durante 500 ns.

4.2.7 Dinâmica Molecular transmembrana-intracelular

O modelo do domínio intracelular obtido do programa I-TASSER foi unido no *pymol* ao domínio transmembrana já inserido em membrana e simulado por 200 ns. O sistema com dimensões 9 x 9 x 15 nm foi preenchido com 24976 moléculas de água, 110 íons Na⁺ e 115 íons Cl⁻. Inicialmente, foi minimizado por 20000 passos com as posições da proteína e membrana restritas. Em seguida, foi termalizado por 7.5 ps em NVT com proteína e membrana ainda restritas. Finalizando a termalização com 46 ps em NPT com restrição na proteína e na membrana e 40 ps também em NPT com a proteína e membrana livres. Após a termalização, o sistema foi simulado por DM durante 500 ns.

4.2.8 Modelagem e Dinâmica Molecular do CAR

Para a montagem do CAR, foi escolhida a configuração após 200 ns do sistema scFv-*hinge*-transmembrana em membrana e água, uma vez que era a configuração que o *hinge* se encontrava mais linear. A configuração do domínio intracelular escolhida foi a do final da simulação do sistema transmembrana-intracelular. A membrana utilizada foi a de 12 x 12 nm já com o domínio transmembrana inserido. O domínio transmembrana do scFv-*hinge*-transmembrana foi sobreposto ao domínio transmembrana que já estava inserido na membrana para orientar a posição do CAR. Foi feita a mesma sobreposição para o transmembrana-intracelular, orientando o domínio intracelular do CAR. A caixa com dimensões 12 x 12 x 21 nm contendo o CAR (scFv-*hinge*-transmembrana-intracelular) e a membrana foi preenchida com 67763 moléculas de água, 186 íons Na⁺ e 191 íons Cl⁻. Inicialmente, o sistema foi minimizado por 20000 passos com as posições da membrana restritas e a proteína livre. Em seguida, foi termalizado por 2.5 ps em NVT com membrana ainda restrita. Finalizando a termalização com 131 ps em NPT com a membrana ainda restrita. Após a termalização, o sistema foi simulado por DM durante 1.2 μ s.

4.2.9 Dinâmica Molecular do CAR-CD19

Após a simulação do CAR em água e do estudo da interação apenas do scFv com o CD19 (sistema scFv-CD19), foi montado o sistema CAR-CD19 em interface membrana-água. A membrana utilizada foi a de 14 x 14 nm já equilibrada por 200 ns. A estrutura inicial do CAR foi aquela proveniente do final da equilibração do sistema CAR. A estrutura do CD19 escolhida

para a montagem do complexo foi a do final da simulação scFv-CD19, mantendo-se a interface de interação.

O domínio transmembrana do CAR foi sobreposto ao domínio transmembrana já inserido na membrana¹⁴⁻¹⁴, usando o programa *Pymol 2.2.0*. para manter a posição inserida em membrana. A estrutura do scFv do final da simulação do complexo scFv-CD19 foi sobreposta ao scFv do CAR para se manter a interface de interação scFv-CD19 com base na estrutura cristalográfica (PDB 6AL5). Depois de montar o CAR-CD19, foram adicionadas 122.612 moléculas de água TIP3P e íons 425 Na⁺ e 429 Cl⁻ para neutralizar o sistema a 0.150 M. As dimensões da caixa de simulação foram de 14 x 14 x 23 nm, suficiente para conter o complexo CAR-CD19 e a membrana.

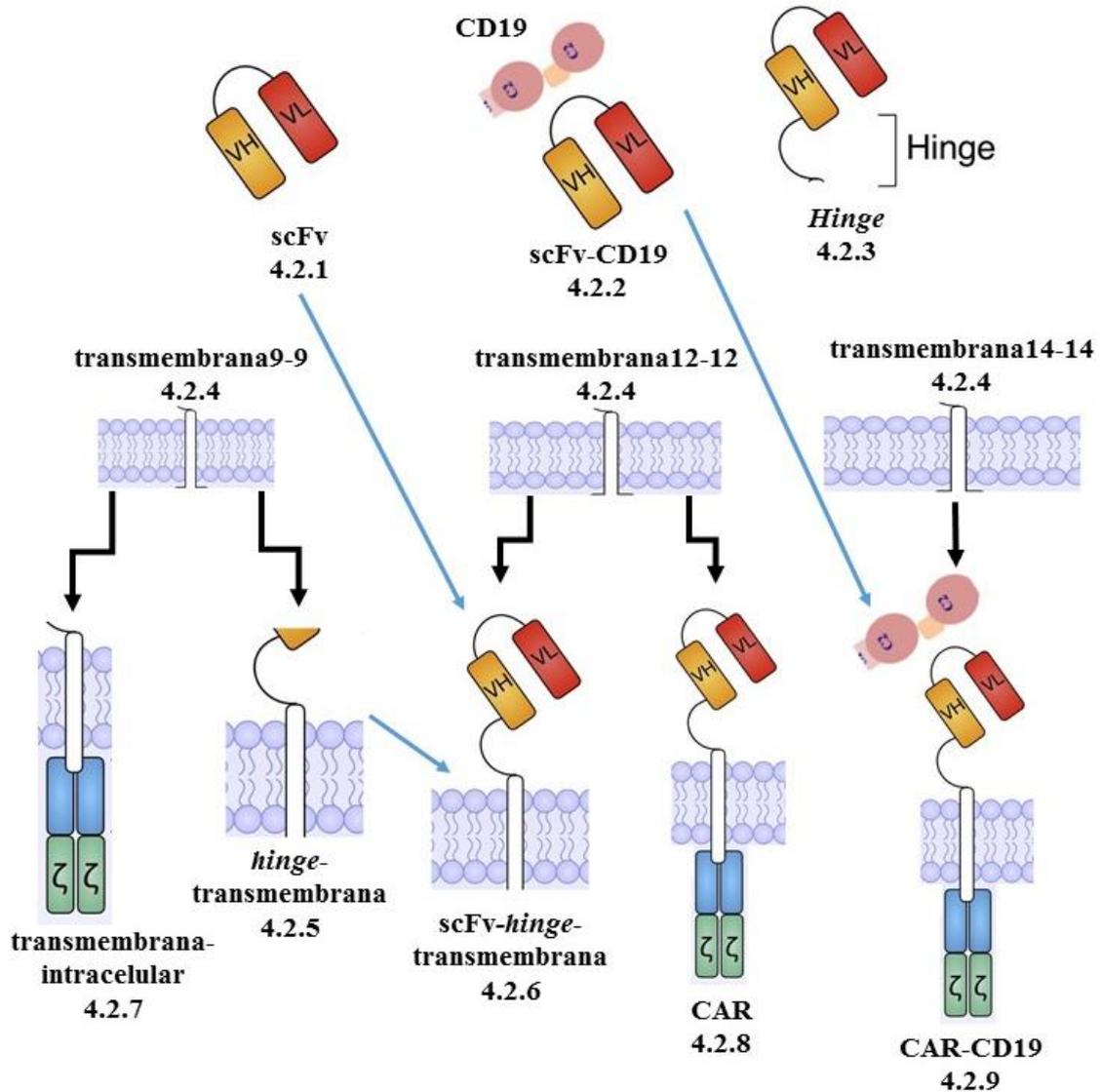
Inicialmente, o sistema foi minimizado por 5000 passos com as posições da membrana e do CD19 restritas e do CAR livre. Em seguida, foi termalizado por 50 ps em NVT com membrana e CD19 ainda restritos. Finalizando a termalização com 225 ps em NPT com a membrana e o CD19 ainda restritos e com 100 ps também em NPT com a membrana livre e apenas o CD19 restrito. Após a termalização, o sistema foi simulado por DM durante 1 μ s.

O CD19 e alguns resíduos do scFv próximos a interface de interação foram mantidos restritos durante toda a simulação para se avaliar a movimentação do *hinge* e dos domínios transmembrana e intracelular, possibilitando inferir a transdução de sinal na célula T.

4.2.10 Definição dos sistemas simulados

Para facilitar o entendimento dos procedimentos experimentais já descritos e dos resultados que serão descritos a seguir, foi construído o fluxograma esquemático apresentado na Figura 12, contendo a definição dos sistemas simulados nesse trabalho. Abaixo do esquema de cada sistema está a nomenclatura usada e a sessão onde se encontra nos procedimentos experimentais. As setas pretas indicam que a respectiva membrana e domínio transmembrana foram utilizados para a montagem dos sistemas apresentados logo abaixo. As setas azuis indicam que as estruturas do scFv e do *hinge*-transmembrana já equilibradas foram utilizadas para a montagem do sistema scFv-*hinge*-transmembrana, e o CD19 do sistema scFv-CD19 foi utilizado para a montagem do sistema CAR-CD19, mantendo-se a interface de interação do final da DM.

Figura 12 – Fluxograma esquemático de definição dos sistemas simulados. Abaixo de cada esquema está sua identificação e a sessão onde está descrito nos procedimentos experimentais.



Fonte: elaborada pelo autor.

4.3 Análise de Componentes Principais (PCA)

A Análise de Componentes Principais (PCA) foi realizada para os sistemas CAR e CAR-CD19 no programa pyPcazip. Os resultados de PCA foram obtidos individualmente para cada componente do CAR nos dois sistemas. As análises de PCA em função do tempo, do CAR e dos componentes nos sistemas CAR e CAR-CD19, possibilitam inferir o papel das mudanças estruturais no CAR quanto em interação, ou não, com o CD19.

A trajetória de DM para os componentes do CAR: scFv, hinge, domínio transmembrana e intracelular, e para o CD19 foram comprimidas e processadas utilizando as ferramentas do programa pyPcazip. A análise foi feita com base na cadeia principal de cada uma das estruturas,

átomos N, C α , C, O. Foram utilizados autovetores suficientes para capturar 95% da variância dos dados.

Para visualizar os principais processos de movimentação das estruturas dos componentes do CAR, as trajetórias foram examinadas quando projetadas no subespaço bidimensional, definido pelos dois principais componentes principais (PC1 e PC2). PC1 e PC2 representam cerca de metade dos movimentos gerais. Portanto, eles são considerados uma representação adequada da dinâmica geral do sistema.

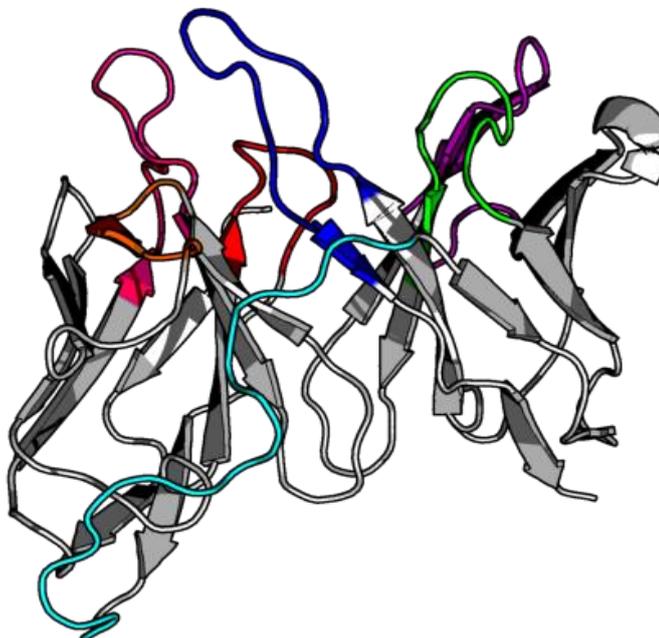
5 RESULTADOS

5.1 Modelagem dos componentes do CAR

5.1.1 Modelagem do scFv

A Figura 13 representa o modelo de scFv do anticorpo B43 obtido a partir da junção dos domínios VH e VL, provenientes do fragmento do tipo Fab do anticorpo B43 (PDB 6AL5), ao *linker* modelado no I-TASSER. O *linker* e as regiões determinantes de complementariedade (CDRs), que serão foco das análises energéticas e estruturais posteriores na interação com o CD19, estão destacados na figura. A sequência dos resíduos das regiões CDRs (CHOTHIA; LESK, 1987), está presente na Tabela 3 com as cores relacionadas à Figura 13.

Figura 13 – Representação estrutural do scfv do anticorpo B43. Representação em *cartoon* da estrutura do scFv, destacando as regiões CDRs. L1: rosa; L2: laranja; L3: vermelho; *linker*: azul claro; H1: verde; H2: roxo e H3: azul.



Fonte: elaborada pelo autor.

Tabela 3- Numeração dos resíduos e sequência de aminoácidos das CDRs do anticorpo B43 de acordo com o arquivo PDB 6AL5 submetido à simulação de DM.

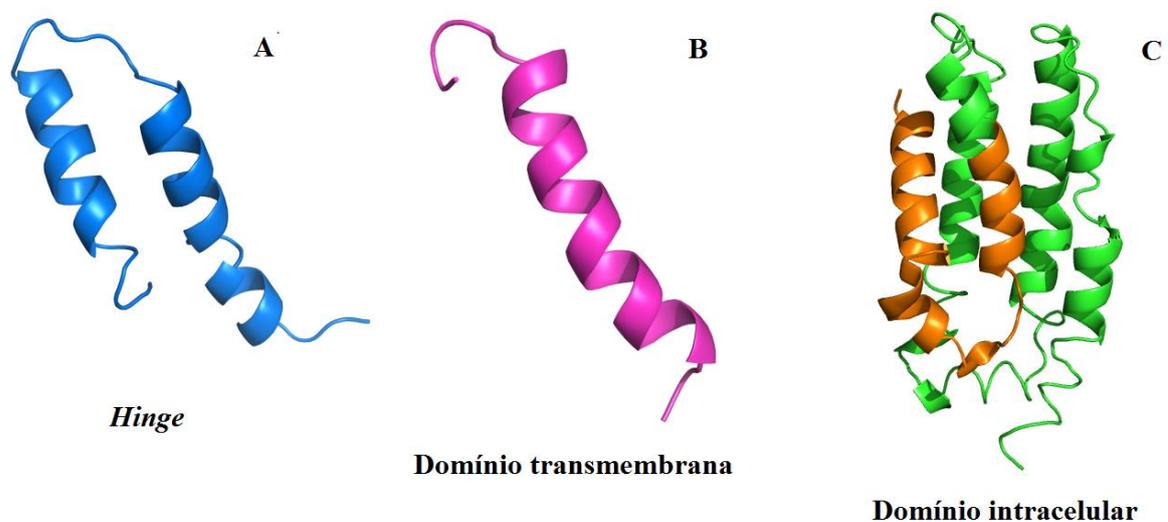
CDRs B43	Sequência	Resíduos
L1	KASQSVDYSGDSYLN	24-38
L2	DASNLVS	54-60
L3	QQSTENPWT	92-101
H1	GYAFSSYWMN	154-164
H2	QIWPGDSDTNYAQKFQG	178-194
H3	RETTTVGRYYYYAMDY	227-241

Fonte: elaborada pelo autor.

5.1.2 Modelagem do hinge, do domínio transmembrana e intracelular

Os melhores modelos selecionados do *hinge* e do domínio transmembrana de CD8 α e do domínio intracelular estão representados na Figura 14. Esses modelos foram, em seguida, validados, minimizados para só depois seguirem para a montagem do CAR.

Figura 14 – Representação estrutural dos modelos obtidos utilizando os programas Modeller e I-TASSER. A) Modelo de *hinge* e B) do domínio transmembrana de CD8 α obtidos por Modelagem Comparativa utilizando o programa Modeller. C) Modelo do domínio intracelular, formado por 4-1BB (laranja) e CD3 ζ (verde), obtido da modelagem por *threading* no programa I-TASSER.



Fonte: elaborada pelo autor.

5.1.3 Validação dos modelos dos componentes do CAR obtidos

Na Tabela 4, estão representados os resultados de validação obtidos no programa MolProbity para cada modelo, como os valores de *clash score* em percentil, de MolProbity *score*, de rotâmeros ruins, de ligações e ângulos em conformações não desejadas (Anexo 1) e o número de resíduos fora de regiões permitidas (*outliers*) no gráfico de Ramachandran (Anexo 2).

Tabela 4 – Resultados obtidos no programa MolProbity para cada modelo obtido de scFv, *hinge*, transmembrana e intracelular.

Modelos	<i>Clash score</i> (percentil)	MolProbity <i>score</i>	Rotâmeros ruins	Ângulos ruins	Ligações ruins	<i>Outliers</i> resíduos
scFv	100°	1.41	9	19	0	4
<i>Hinge</i>	100°	0.82	0	6	0	1
Transmembrana	100°	0.50	0	2	0	0
Intracelular	100°	1.57	4	76	23	7

Fonte: elaborada pelo autor.

Os dados de validação de cada modelo no programa PROCHECK estão representados no Anexo 3. O número de resíduos presentes em regiões mais favoráveis, em regiões permitidas e em regiões não permitidas para cada modelo está apresentado na Tabela 5. Assim, cada modelo dos componentes do CAR mostra-se de boa qualidade por ter acima de 90 % dos aminoácidos em regiões mais favoráveis ou permitidas.

Tabela 5 – Resultado de validação obtido no programa PROCHECK para cada modelo obtido de scFv, *hinge*, transmembrana e intracelular, indicando o número de resíduos em regiões mais favoráveis, em regiões permitidas e em regiões não permitidas.

Modelos	Regiões mais favoráveis (%)	Regiões permitidas (%)	Regiões não permitidas (%)	Resíduos em regiões não permitidas
scFv	87,7	11,8	0,5	Ala55
<i>Hinge</i>	90,3	9,7	0	-
Transmembrana	94,7	5,3	0	-
Intracelular	81,1	18,1	0,8	Gln20

Fonte: elaborada pelo autor.

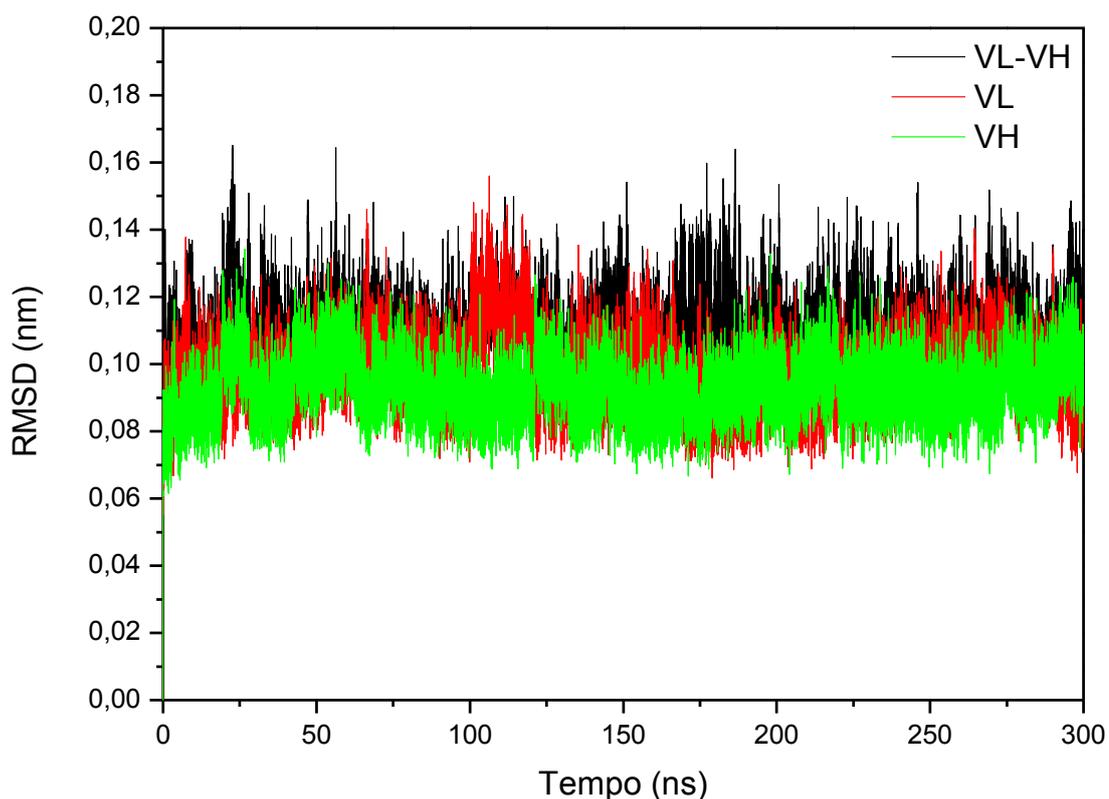
5.2 Simulação de Dinâmica Molecular (DM) e análises

5.2.1 Dinâmica Molecular do scFv

O equilíbrio estrutural dos domínios variáveis do scFv foi determinado pela análise de

RMSD. Os perfis de RMSD para os domínios VL-VH, VL e VH são semelhantes, como mostrado na Figura 15, com RMSDs médios de 0.10 ± 0.01 ; 0.09 ± 0.01 e 0.09 ± 0.01 nm (Tabela 6), respectivamente. Com base nos perfis de RMSD, pode-se inferir que a estrutura dos domínios VL e VH alcançam equilíbrio (*teq*) após 25 ns de simulação e se mantêm até o final.

Figura 15 – Perfil de RMSD para os domínios variáveis do scFv. Curva de RMSD entre as configurações estruturais do VL-VH (sobreposição e cálculo de RMSD de VL-VH), VL (sobreposição e cálculo de RMSD de VL) e VH (sobreposição e cálculo de RMSD de VH), ao longo do tempo de simulação, comparadas com a estrutura de referência inicial.



Fonte: elaborada pelo autor.

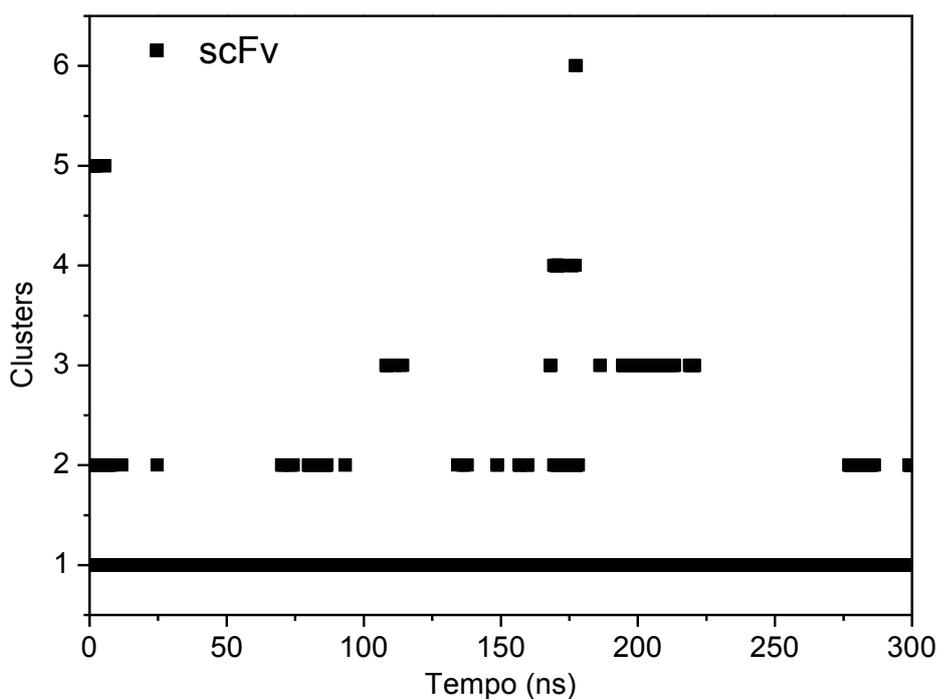
Tabela 6 – Médias e desvios de RMSD calculados para VL-VH, VL e VH a partir de *teq* = 25 ns.

Estrutura	RMSD (nm)
VL-VH	0.10 ± 0.01
VL	0.09 ± 0.01
VH	0.09 ± 0.01

Fonte: elaborada pelo autor.

A análise de *cluster* para o VL-VH mostra que o *cluster* 1 é o mais representativo desde o início da simulação (Figura 16), reforçando que a estrutura do scFv não varia muito em água. O que pode ser observado também na Figura 17, em que ao longo da simulação os domínios VL e VH se mantêm, variando apenas a estrutura do *linker*. Essa análise foi feita com base na estrutura do scFv, utilizando um raio de corte de 0.18 nm.

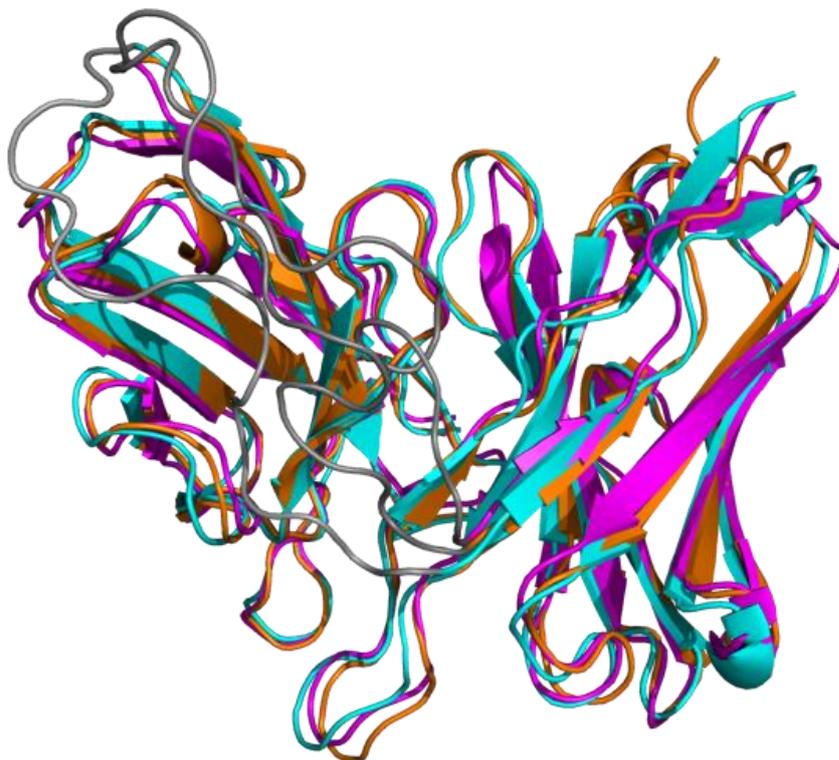
Figura 16 – Resultado da análise de *cluster* para o scFv. *Clusters* de configurações de scFv ao longo da simulação, utilizando um raio de corte de 0.18 nm.



Fonte: elaborada pelo autor.

Com base na análise de *cluster*, a estrutura do scFv escolhida para a montagem do CAR foi a configuração do final da simulação, uma vez que não houve mudança conformacional significativa para o scFv que agrupasse as estruturas obtidas durante a simulação em diferentes *clusters* representativos. O RMSD após a sobreposição dos domínios VL das estruturas de 100 e 200 ns apresentaram RMSD de 0.07 nm e entre as estruturas de 300 e 100 ns, o RMSD foi de 0.08 nm, ambos dentro do raio de corte para a análise de *cluster*.

Figura 17 – Representação estrutural do scFv após a sobreposição do domínio VL. A) após 100 ns (laranja); B) 200 ns (rosa) e C) 300 ns (azul claro) de simulação. O *linker* está destacado em cinza.

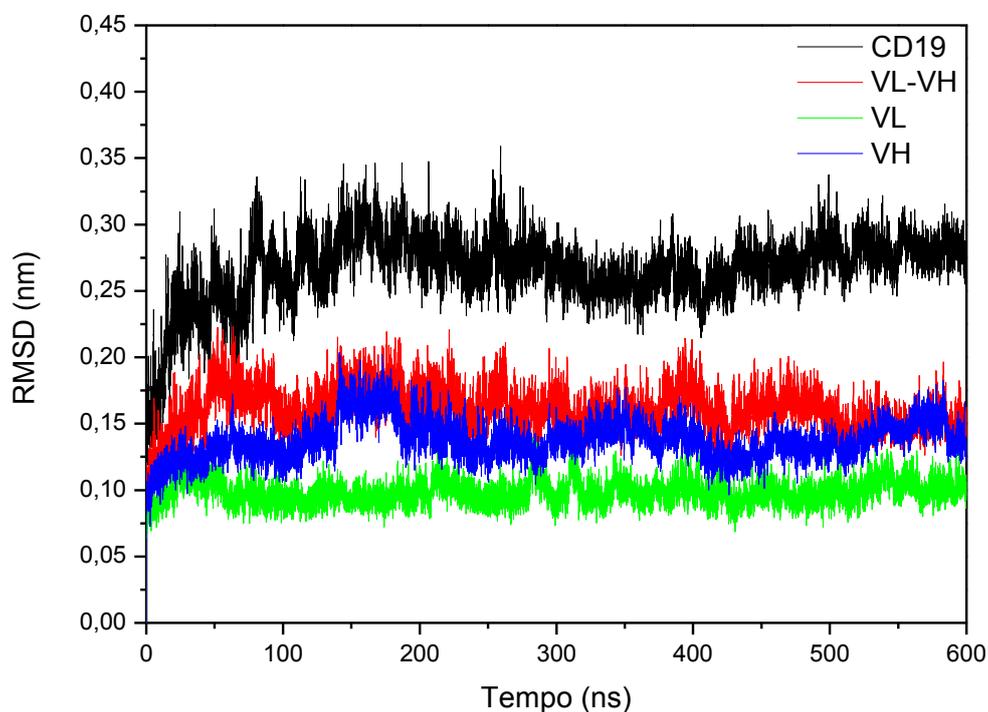


Fonte: elaborada pelo autor.

5.2.2 Dinâmica Molecular do scFv-CD19

O RMSD foi calculado para os domínios variáveis do scFv e para o CD19 no sistema scFv-CD19 com objetivo de observar o equilíbrio dessas estruturas em solução. O perfil de RMSD para a estrutura dos domínios variáveis VL-VH (Figura 18), sem levar em conta o *linker*, apresenta menor RMSD médio de 0.15 ± 0.01 nm (Tabela 7) a partir de 200 ns de simulação. Já os domínios VL e VH separados possuem RMSDs médios ainda menores, de 0.09 ± 0.01 e 0.13 ± 0.01 nm (Tabela 7), respectivamente. O RMSD médio para o scFv, calculado a partir de $t_{eq} = 300$ ns, foi de 0.26 ± 0.01 nm (Tabela 7). O que indica que as estruturas dos domínios variáveis do scFv e o CD19 estão em equilíbrio em solução.

Figura 18 – Perfis de RMSD para o CD19 e os domínios variáveis do scFv. Curva de RMSD entre as configurações estruturais dos componentes do scFv (sobreposição e cálculo de RMSD do VL-VH; do VL e do VH) e do CD19 (sobreposição e cálculo de RMSD do CD19), ao longo do tempo de simulação, comparadas com a estrutura inicial como referência.



Fonte: elaborada pelo autor.

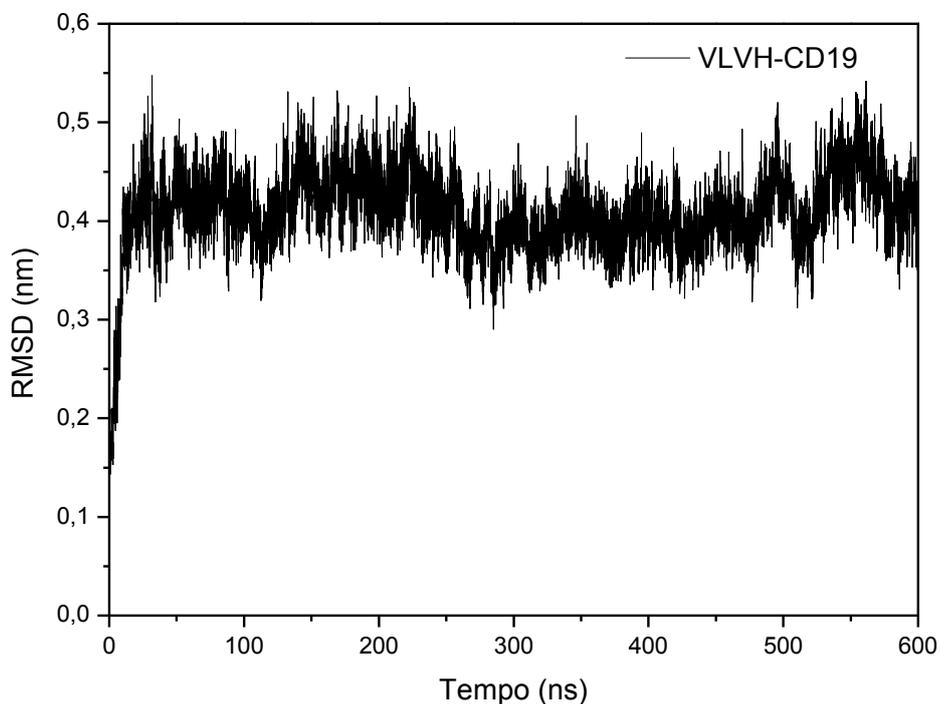
Tabela 7 - Médias e desvios de RMSD calculados para os domínios VL-VH, VL e VH e também para o CD19 em solução a partir do *teq*.

Estrutura	RMSD (nm)	<i>teq</i> (ns)
VL-VH	0.15 ± 0.01	200
VL	0.09 ± 0.01	200
VH	0.13 ± 0.01	200
CD19	0.26 ± 0.01	300

Fonte: elaborada pelo autor.

A Figura 19 mostra o perfil de RMSD para o complexo VLVH-CD19, desconsiderando a influência do *linker* na estrutura do scFv. A redução no perfil de RMSD, em 250 ns, indica uma mudança na interface scFv-CD19, seguida da estabilização da interface a partir de 300 ns até o final da simulação com RMSD médio de 0.40 ± 0.03 nm.

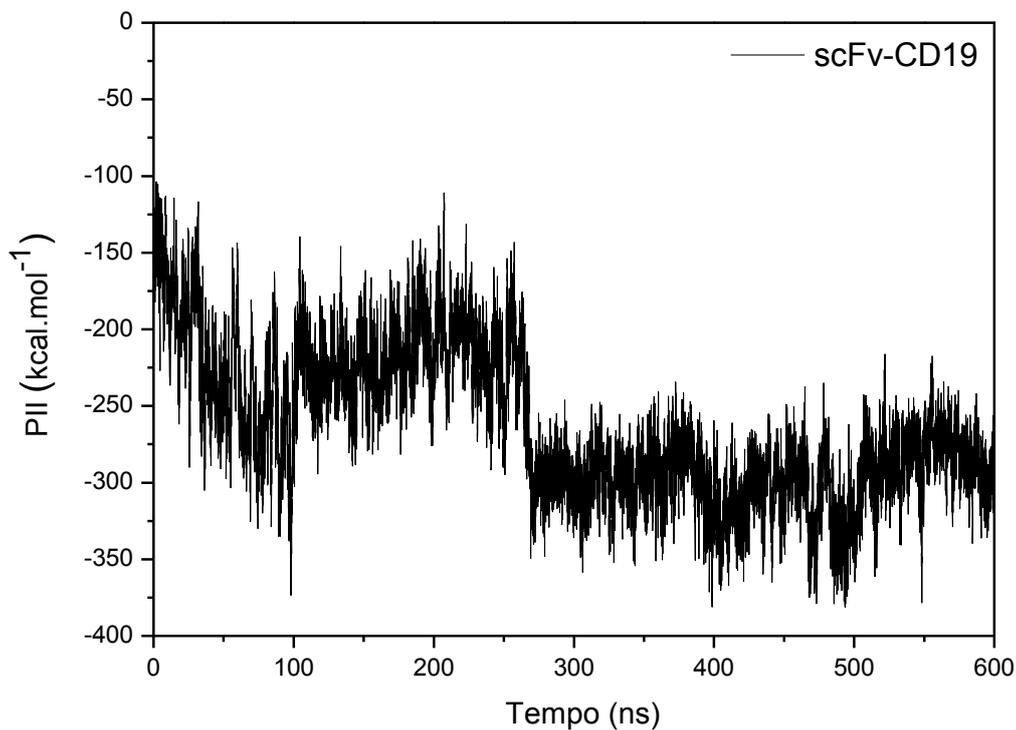
Figura 19 – Perfil de RMSD do complexo VLVH-CD19. Curva de RMSD das configurações estruturais do VLVH-CD19 ao longo do tempo de simulação (sobreposição e cálculo de RMSD do VLVH-CD19), comparadas com a estrutura inicial como referência.



Fonte: elaborada pelo autor.

O Potencial de Interação Intermolecular (PII) entre o scFv e o CD19 foi calculado para se inferir a natureza da interação scFv-CD19, se é atrativa ou repulsiva. A Figura 20 mostra o perfil de PII scFv-CD19. Há uma redução do PII em 250 ns, o que realça a interação atrativa entre o scFv e o CD19, no mesmo instante em que houve mudança no perfil de RMSD VLVH-CD19 (Figura 19). O PII médio para essa interação foi de $-297,0 \pm 24,7$ kcal.mol⁻¹ calculado a partir de 300 ns de simulação (Figura 20), o que indica interação atrativa entre o scFv e o CD19, realçada após mudança na interface de interação scFv-CD19 (Figura 19).

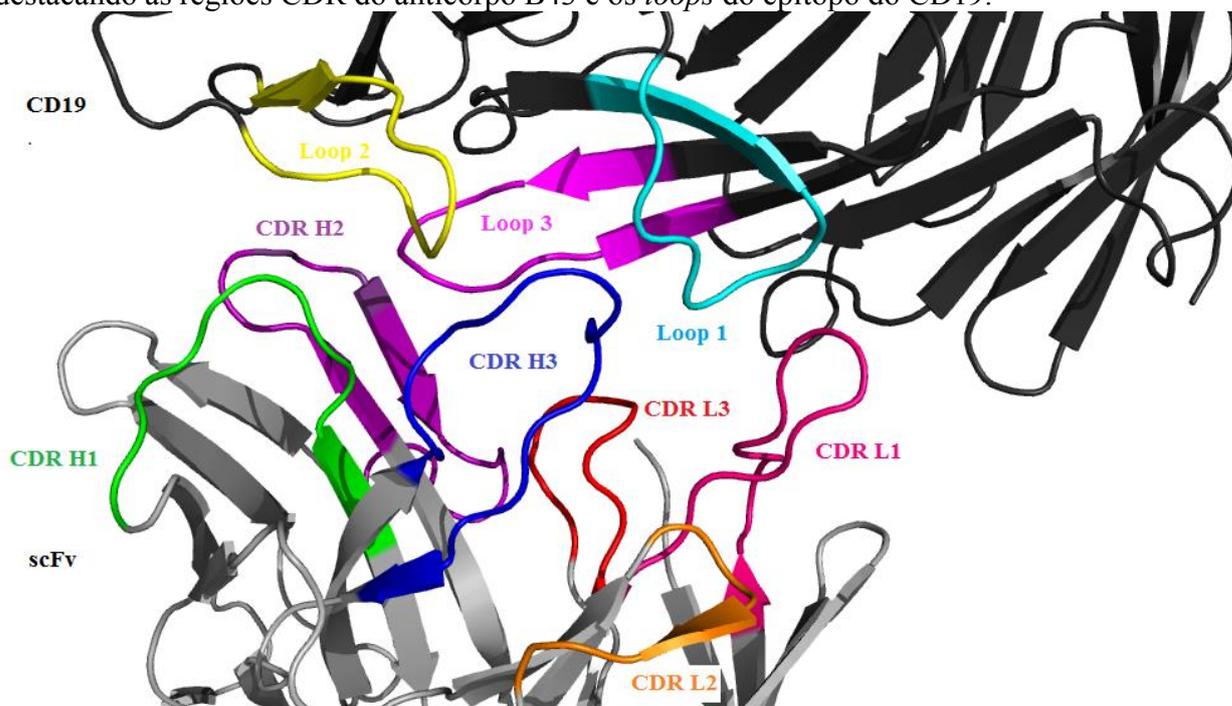
Figura 20 – Potencial de interação intermolecular (PII) scFv-CD19. PII entre o scFv e o CD19 calculado durante 600 ns, utilizando um raio de corte de 1.3 nm.



Fonte: elaborada pelo autor.

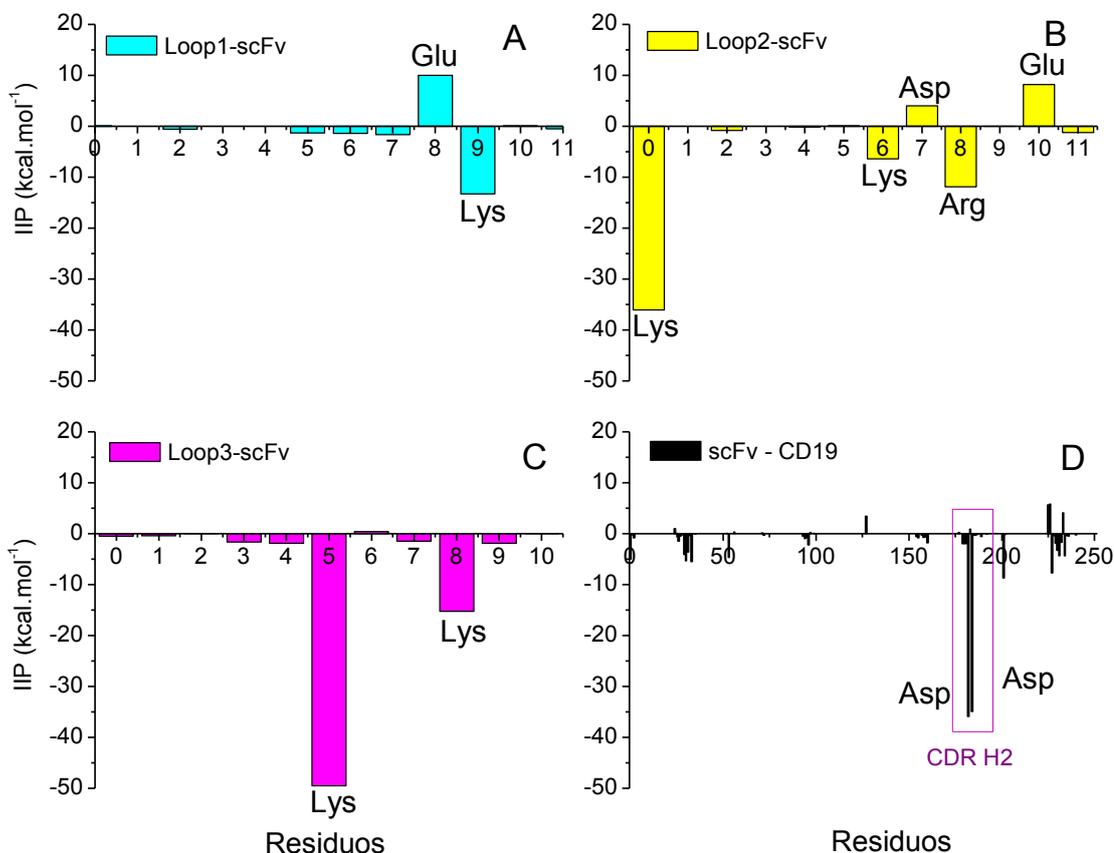
Na Figura 21 está representada a interface de interação entre as regiões CDR do anticorpo B43, de acordo com as cores já definidas na Figura 13, e os *loops* do epítipo do CD19. Sendo as CDRs da cadeia pesada, as que mais contribuem para a interação com o CD19, como já descrito por Teplyakov e colaboradores (TEPLYAKOV *et al.*, 2018). As cores dos *loops* do epítipo destacadas nessa figura se relacionam com as cores dos gráficos da Figura 22.

Figura 21 – Representação da interface de interação scFv-CD19. Interface scFv-CD19, destacando as regiões CDR do anticorpo B43 e os *loops* do epítipo do CD19.



Para determinar a contribuição de cada resíduo do scFv e dos *loops* do epítipo do CD19 na interação scFv-CD19 foi realizada a análise de PII por resíduo entre os resíduos do scFv e o CD19. Os resultados dessa análise estão apresentados na Figura 22, em que se pode observar que resíduos Asp da CDR H2 contribuem para uma interação atrativa com o CD19, uma vez que possuem PII negativos (Figura 22D). O que corrobora o trabalho de Teplyakov e colaboradores (TEPLYAKOV *et al.*, 2018), em que foi destacado que as CDRs da cadeia pesada interagem com o CD19 na estrutura cristalográfica (PDB 6AL5). Já na análise inversa em que foi calculado o PII entre os resíduos dos *loops* do epítipo do CD19 e o scFv, pode-se observar que os resíduos carregados positivamente como lisina (Lys) e arginina (Arg) (Figura 22A, B e C) contribuem para a interação atrativa com o scFv, já que apresentam PII negativos. Os resíduos de lisina dos *loops* 2 e 3 são os que contribuem mais atrativamente.

Figura 22 – Potencial de interação intermolecular (PII) por resíduo. PII por resíduo calculado por programa *homemade* entre os resíduos dos *loops* do epítipo e o scFv (A, B e C) e entre os resíduos do scFv e o CD19 (D).

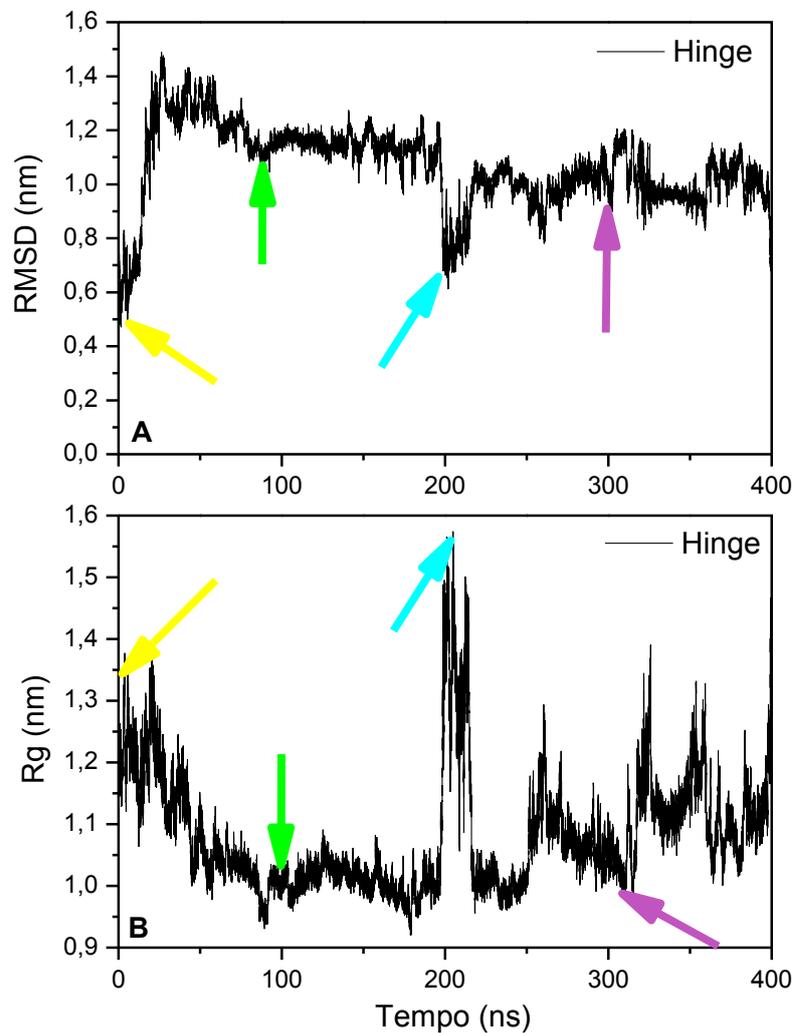


Fonte: elaborada pelo autor.

5.2.3 Dinâmica Molecular do hinge de CD8 α

O perfil de RMSD e o Raio de giro (Rg) para a estrutura do *hinge* indica oscilações na estrutura, acompanhadas por oscilações bruscas (Figura 23). Na figura 24, estão representadas quatro configurações da estrutura do *hinge*, indicadas na Figura 23 por setas com as respectivas cores relacionadas à Figura 24, ilustrando mudança estrutural em relação ao modelo do *hinge* do início da DM. Em 100 ns, o *hinge* apresenta-se como uma estrutura mais compacta com Rg pequeno (Figura 23B) e, em 200 ns, em forma mais linear, com redução no RMSD (Figura 23A) e aumento do Rg, indicando descompactação. Em 300 ns, a estrutura do *hinge* já se apresenta mais compacta novamente, o que confere oscilações na estrutura do *hinge*.

Figura 23 – Perfil de RMSD e Raio de giro (Rg) para o *hinge*. A) Curva de RMSD entre as configurações estruturais do *hinge*, em função do tempo de simulação, comparadas com a estrutura de referência inicial. B) Rg calculado para a estrutura do *hinge*. Cada seta indica uma configuração durante a simulação. Amarelo: modelo; verde: em 100 ns; azul claro: em 200 ns e roxo: em 300 ns.



Fonte: elaborada pelo autor.

A estrutura do *hinge* escolhida para a montagem do CAR foi a estrutura em 200 ns (Figura 24), em que ele se encontra mais linear com a extremidade N-terminal mais distendida, para se conectar mais facilmente ao C-terminal do scFv, e a extremidade C-terminal para baixo para se conectar ao N-terminal do domínio transmembrana do CAR.

Figura 24 – Representação estrutural do *hinge*. Estruturas do *hinge* após sobreposição de cada estrutura ao modelo. Amarelo: configuração do modelo; verde: em 100 ns; azul claro: em 200 ns e roxo: em 300 ns de simulação.

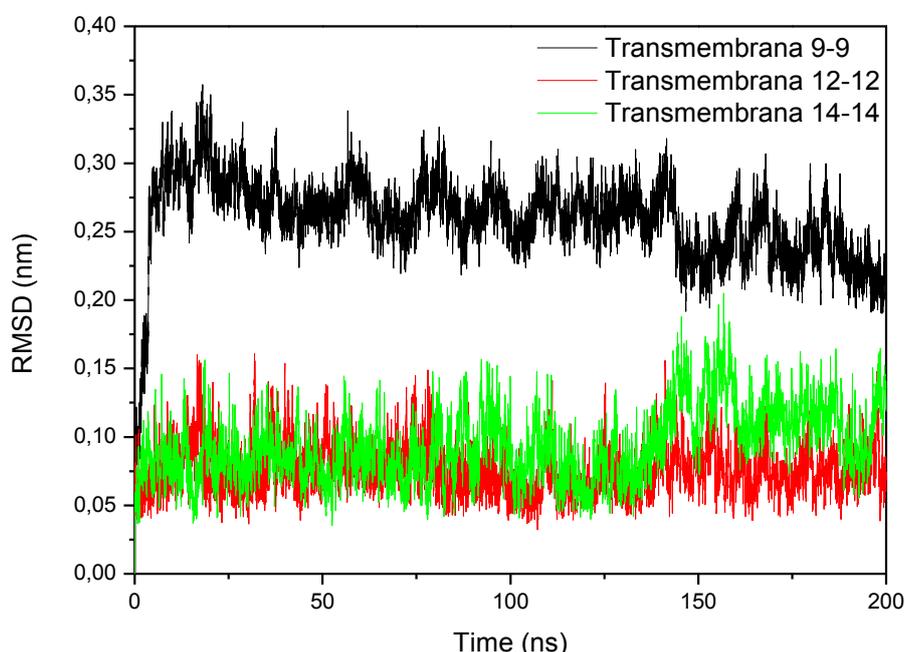


Fonte: elaborada pelo autor.

5.2.4 Dinâmica Molecular do domínio transmembrana

Os perfis de RMSD do domínio transmembrana, nos três sistemas, transmembrana9-9, transmembrana12-12 e transmembrana14-14 estão apresentados na Figura 25. O perfil de RMSD do domínio na região transmembrana9-9 mostra que após $teq = 50$ ns a estrutura apresenta RMSD médio de 0.25 ± 0.02 nm (Tabela 8). Já nos sistemas transmembrana12-12 e transmembrana14-14, $teq = 10$ ns e RMSD médio de 0.07 ± 0.01 e 0.09 ± 0.02 nm, respectivamente (Tabela 8). Os resultados de RMSD para o domínio transmembrana nos três sistemas inferem que ele entra em equilíbrio rapidamente e pode ser utilizado para a montagem dos sistemas seguintes.

Figura 25 – Perfis de RMSD para o domínio transmembrana. Curva de RMSD para as configurações estruturais do domínio transmembrana por tempo de simulação, após sobreposição com a estrutura de referência inicial nos sistemas transmembrana9-9; 12-12 e 14-14. (sobreposição e cálculo de RMSD do domínio transmembrana).



Fonte: elaborada pelo autor.

Tabela 8 - Médias e desvios de RMSD da estrutura do domínio transmembrana calculados a partir do momento em que se atinge o equilíbrio (*teq*) nos sistemas transmembrana9-9; 12-12 e 14-14.

Transmembrana	RMSD (nm)	<i>teq</i>
9-9	0.25 ± 0.02	50 ns
12-12	0.07 ± 0.01	10 ns
14-14	0.09 ± 0.02	10 ns

Fonte: elaborada pelo autor.

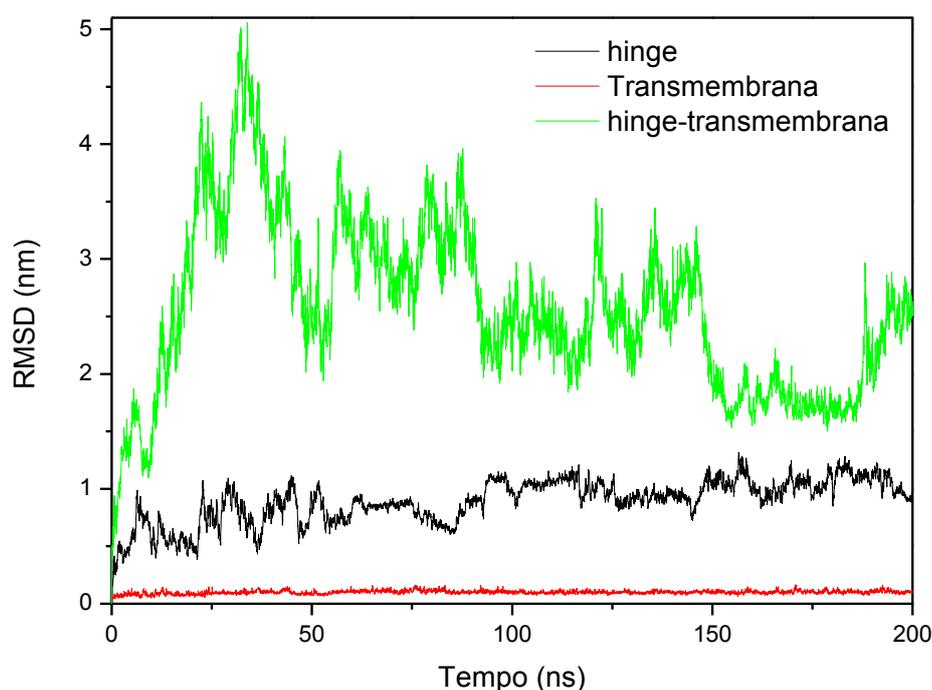
As estruturas finais de cada simulação são representativas, visto os baixos RMSD médios e desvios, ou seja, não foram observadas mudanças estruturais e movimentação acima do esperado para movimentação térmica. Assim, as estruturas finais da simulação foram utilizadas para a montagem dos sistemas e para as simulações seguintes.

5.2.5 Modelagem e Dinâmica Molecular do hinge-transmembrana

O perfil de RMSD do *hinge* calculado a partir de 100 ns, com oscilações de grande

amplitude, e RMSD médio de 1.01 ± 0.09 nm, apresentando grande desvio, indicam que a estrutura não atinge equilíbrio em 200 ns (Figura 26). O que era esperado já que essa estrutura possui poucos resíduos e possui a extremidade livre, sendo bem flexível. Além disso, o *hinge* tem como característica ser uma estrutura flexível para garantir flexibilidade ao domínio extracelular do CAR. Já a estrutura do domínio transmembrana, como já estava equilibrada na membrana, manteve-se em equilíbrio, com RMSD médio de 0.10 ± 0.01 nm, a partir de $t_{eq} = 10$ ns. O cálculo de RMSD do *hinge*, após a sobreposição do domínio transmembrana (*hinge-transmembrana*), com objetivo de observar a movimentação do *hinge*, está mostrado na Figura 26. O que infere oscilação de grande amplitude da estrutura do *hinge* em relação ao domínio transmembrana. O perfil de oscilações caracteriza diferentes configurações ao longo da DM como se pode observar na Figura 27 pela presença de vários *clusters*. Três configurações ao longo da DM estão representadas na Figura 28, ilustrando a variação estrutural do *hinge*, após sobreposição do domínio transmembrana das três estruturas.

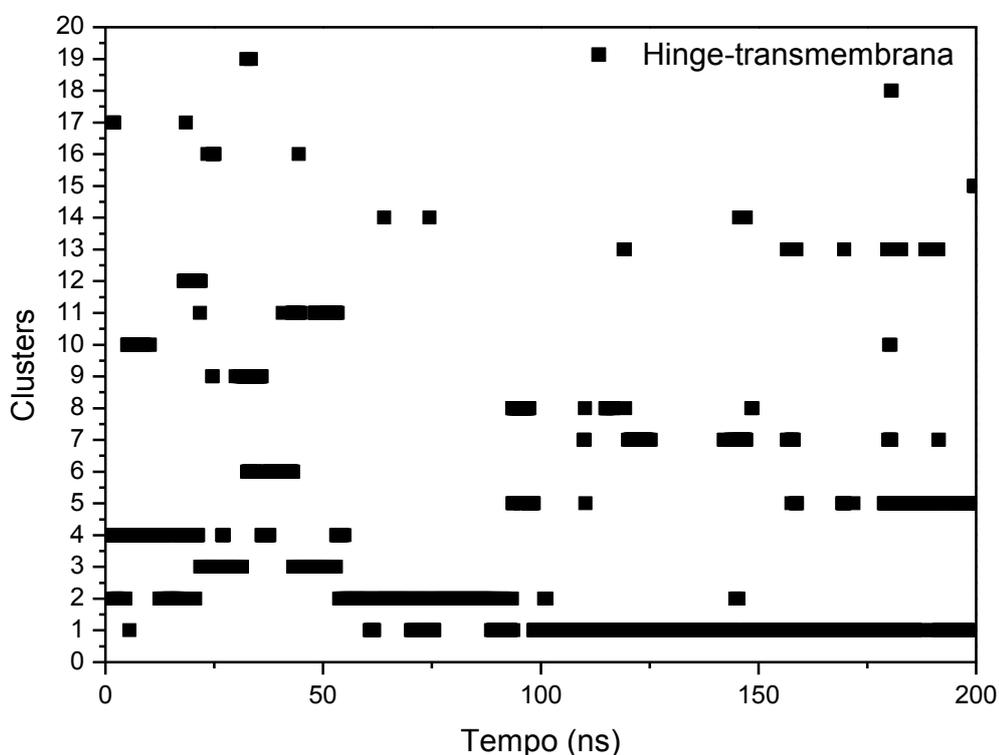
Figura 26 – Perfis de RMSD para o *hinge-transmembrana*. Curva de RMSD entre as configurações estruturais do *hinge* (sobreposição e cálculo de RMSD do *hinge*), transmembrana (sobreposição e cálculo de RMSD do transmembrana) e *hinge-transmembrana* (sobreposição do domínio transmembrana e cálculo de RMSD do *hinge*), ao longo do tempo de simulação, comparadas com a estrutura de referência inicial.



Fonte: elaborada pelo autor.

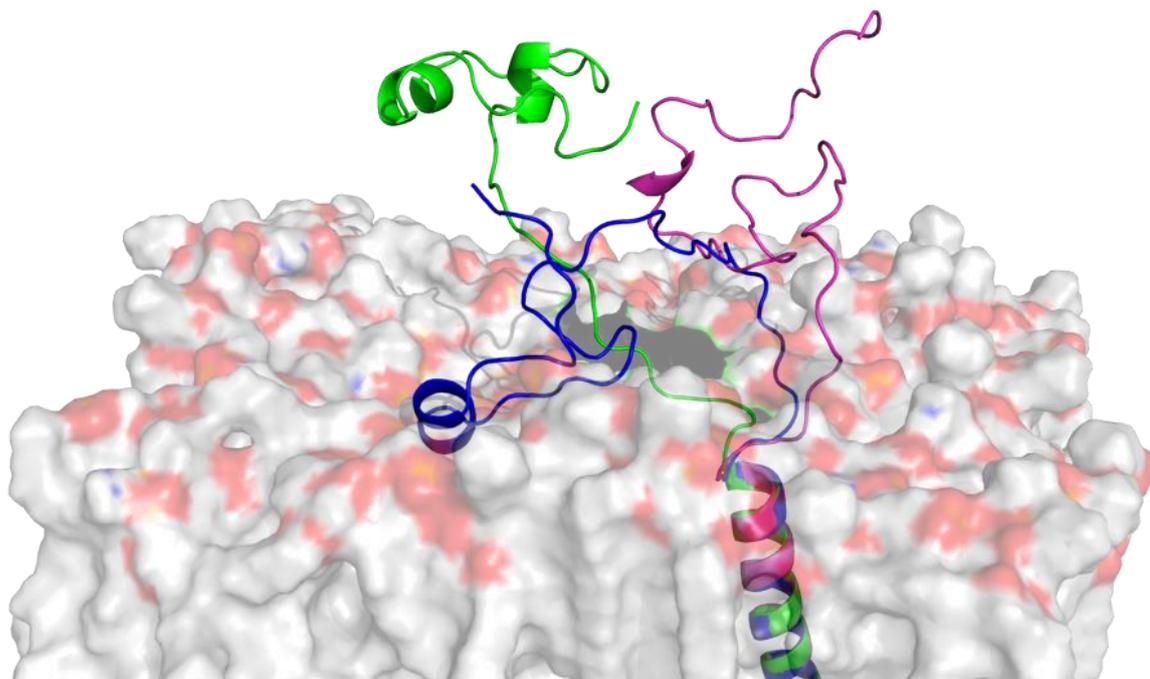
A Figura 27 mostra o resultado da análise de *cluster* para o *hinge*-transmembrana, mostrando os diferentes *clusters* de configurações obtidas na DM, utilizando raio de corte de 0.6 nm, com base na oscilação do perfil de RMSD do *hinge*. Na figura 28, estão representadas diferentes configurações de *hinge*-transmembrana com base no resultado da análise de *cluster*, em 0. 78 e 200 ns, exemplificando as diferentes conformações do *hinge* nos *clusters* 4, 2 e 1 respectivamente. A configuração escolhida para a montagem do CAR foi a de 78 ns (Figura 28) em que o N-terminal do *hinge* está mais linear, que corresponde ao *cluster* 2 (Figura 27), apesar de não ser a configuração do *cluster* mais representativa em solução.

Figura 27 – Resultado da análise de *cluster* para o *hinge*-transmembrana. *Clusters* de configurações de *hinge*-transmembrana ao longo da simulação, utilizando um raio de corte de 0.6 nm.



Fonte: elaborada pelo autor.

Figura 28 - Representação estrutural do *hinge*-transmembrana após a sobreposição do domínio transmembrana. A) no início (verde); B) em 78 ns (roxo) e C) em 200 ns (azul) de simulação.

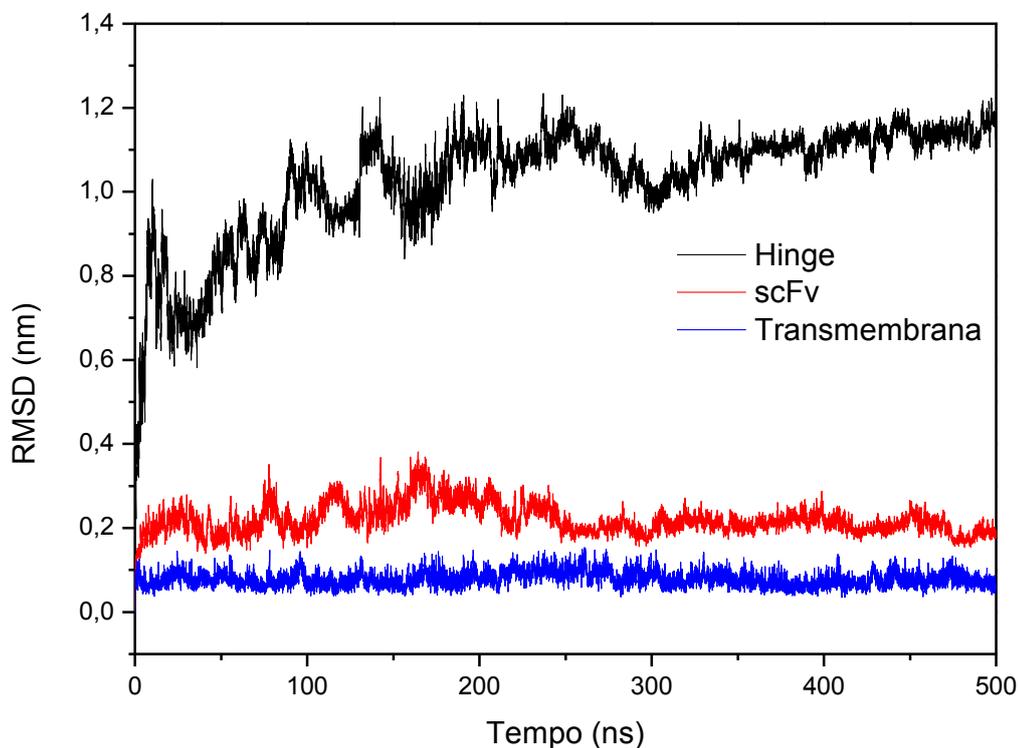


Fonte: elaborada pelo autor.

5.2.6 Modelagem e Dinâmica Molecular do *scFv*-*hinge*-transmembrana

O RMSD foi calculado para o *scFv*, *hinge* e transmembrana, após a sobreposição de cada componente. Os resultados de RMSD estão apresentados na Figura 29. O domínio transmembrana entra em equilíbrio em $teq \sim 20$ ns de simulação com RMSD médio de 0.07 ± 0.01 nm (Tabela 9). Já o *scFv* entra em equilíbrio com $teq \sim 250$ ns, com RMSD médio de 0.20 ± 0.01 nm. A estrutura do *hinge* apresenta RMSD crescente, reduzindo-se em 250 ns de simulação (Figura 29). Em seguida, o perfil de RMSD do *hinge* se mantém em equilíbrio no $teq = 350$ ns até o final da simulação, com RMSD médio de 1.12 ± 0.02 nm (Tabela 9).

Figura 29 – Perfis de RMSD para o scFv, *hinge* e transmembrana. Curva de RMSD das configurações estruturais do scFv (sobreposição e cálculo de RMSD do scFv), *hinge* (sobreposição e cálculo de RMSD do *hinge*) e transmembrana (sobreposição e cálculo de RMSD do transmembrana) por tempo de simulação, comparadas com a estrutura de referência inicial.



Fonte: elaborada pelo autor.

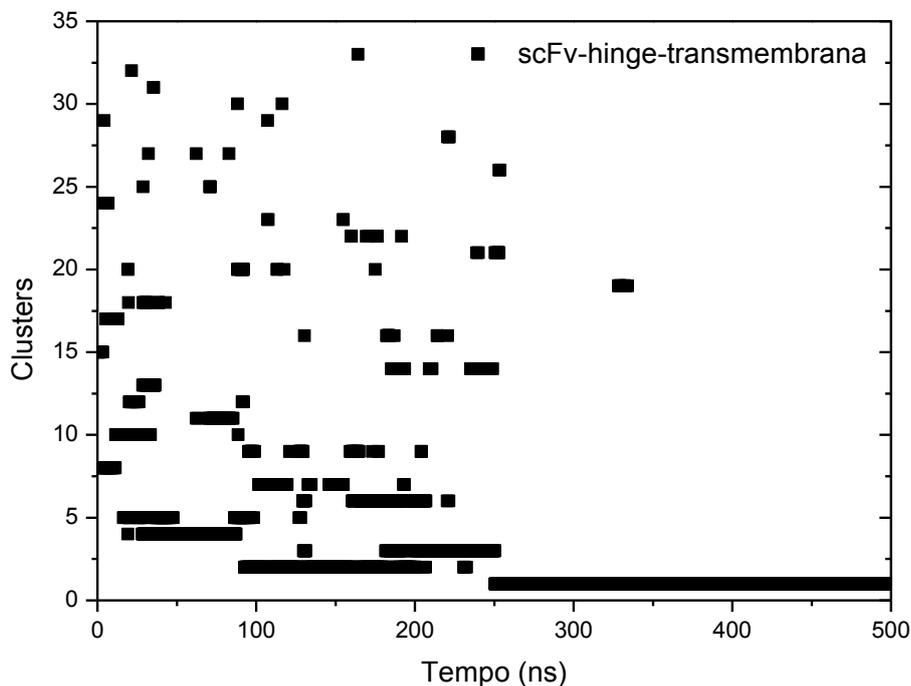
Tabela 9 - Médias e desvios de RMSD calculados para o scFv, *hinge* e transmembrana, a partir do tempo em que se atinge o equilíbrio (*teq*).

Estrutura	RMSD (nm)	<i>teq</i> (ns)
scFv	0.20 ± 0.01	250
<i>Hinge</i>	1.12 ± 0.02	350
Transmembrana	0.07 ± 0.01	20

Fonte: elaborada pelo autor.

A análise de *cluster* (Figura 30) mostra a presença de 32 *clusters*, utilizando um raio de corte de 0.6 nm com base no perfil de RMSD do *hinge*. A presença de muitos *clusters* no início da simulação pode ocorrer devido à flexibilidade do *hinge* que movimenta também o scFv até 250 ns (Figura 31B). Em 250 ns, o scFv começa a se aproximar da membrana e interagir atrativamente, como visto no PII utilizando raio de corte 1.3 nm (Figura 31A).

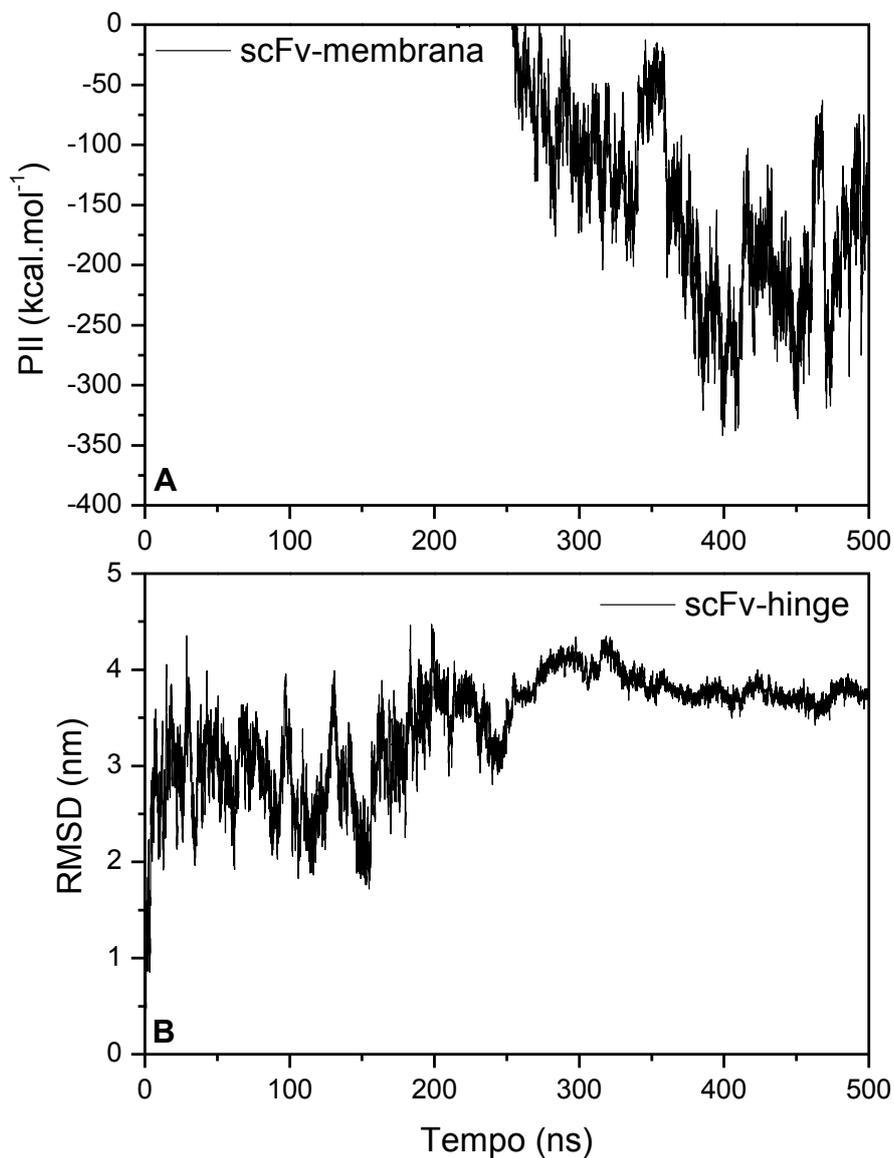
Figura 30 - Resultado da análise de *cluster* para o scFv-*hinge*-transmembrana. *Clusters* de configurações de scFv-*hinge*-transmembrana ao longo da simulação, utilizando um raio de corte de 0.6 nm.



Fonte: elaborada pelo autor.

A Figura 31A mostra o resultado do PII entre o scFv e a membrana. O PII scFv-membrana reduz a partir de 250 ns, mostrando que a atratividade é realçada entre o scFv e a membrana, com a aproximação do scFv da membrana. Em seguida, a estrutura do *hinge* se mantém, ~ 350 ns, como mostrado na Figura 29. A Figura 31B mostra o perfil de RMSD do *hinge* calculado após a sobreposição do scFv (scFv-*hinge*), inferindo oscilação estrutural do *hinge*, em relação ao scFv, até ~ 250 ns. Quando o scFv se aproxima da membrana, há redução da oscilação da estrutura do *hinge*.

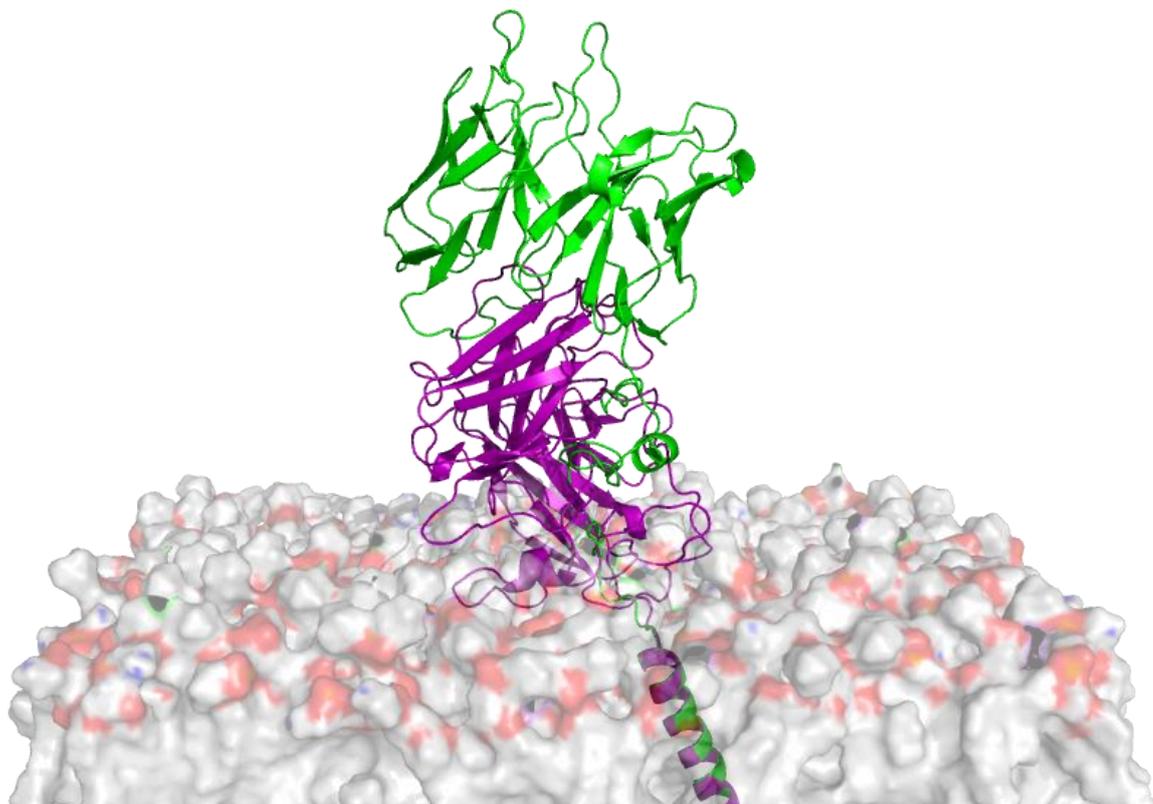
Figura 31 - Potencial de interação intermolecular (PII) entre o scFv e a membrana (scFv-membrana) e RMSD do *hinge* em relação ao scFv. A) PII entre o scFv e a membrana calculado durante 500 ns, utilizando um raio de corte de 1.3 nm. B) RMSD do *hinge* em relação ao scFv (scFv-*hinge*).



Fonte: elaborada pelo autor.

A configuração escolhida para a montagem do CAR foi a que a orientação scFv-*hinge* se apresentava mais linear, de 200 ns (Figura 32), representada no *cluster 2* (Figura 30). Apesar de não ser a configuração mais representativa da simulação, foi escolhida por facilitar o estudo posterior de análise da variação da estrutura do CAR em membrana e também da interação CAR-CD19.

Figura 32 – Representação estrutural do scFv-*hinge*-transmembrana após sobreposição do domínio transmembrana. Conformações do scFv-*hinge*-transmembrana inserido em membrana após 200 (verde) e 500 ns (roxo).

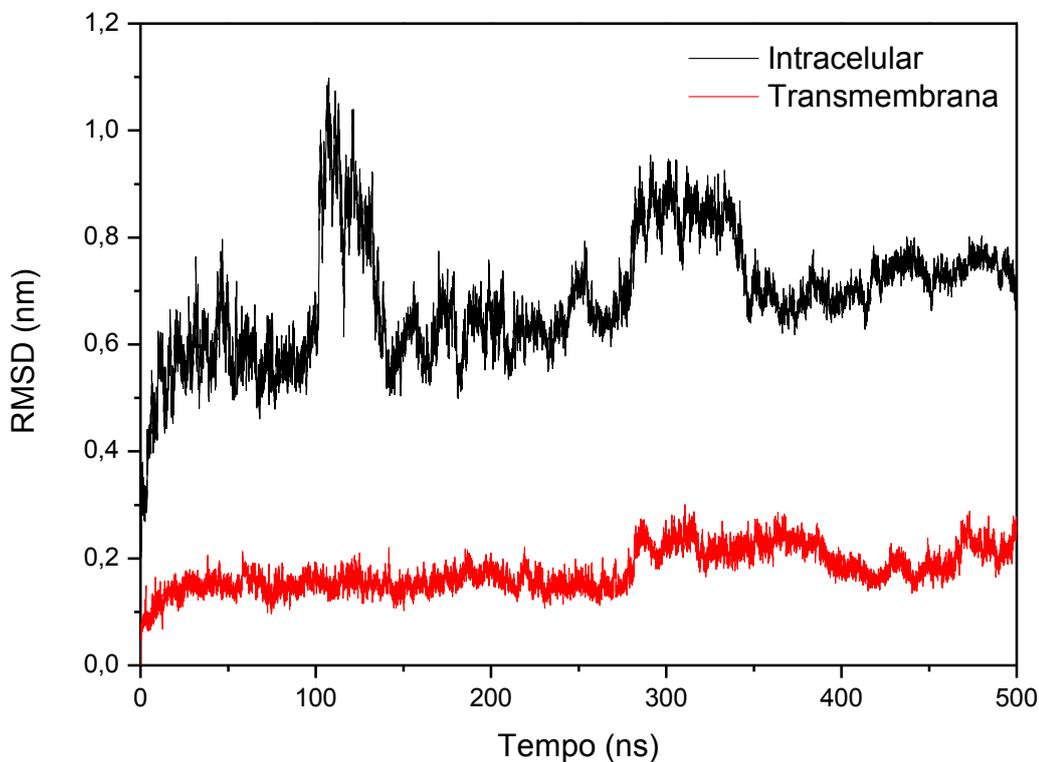


Fonte: elaborada pelo autor.

5.2.7 *Dinâmica Molecular do transmembrana-intracelular*

A estrutura do domínio intracelular apresenta um perfil de RMSD bastante variável, com dois picos de RMSD em 100 e 300 ns (Figura 33), indicando variações estruturais/conformacionais nesse domínio. O RMSD médio para o domínio transmembrana foi calculado a partir de 300 ns, sendo 0.20 ± 0.02 nm. Já o domínio intracelular apresenta RMSD médio mais alto de 0.71 ± 0.03 nm, calculado a partir de 350 ns. A última configuração do domínio intracelular foi escolhida para a montagem do CAR.

Figura 33 – Perfis de RMSD para os domínios transmembrana e intracelular. Curva de RMSD das configurações estruturais dos domínios intracelular (sobreposição e cálculo de RMSD do intracelular) e transmembrana (sobreposição e cálculo de RMSD do transmembrana), ao longo do tempo de simulação, comparadas com a estrutura de referência inicial.

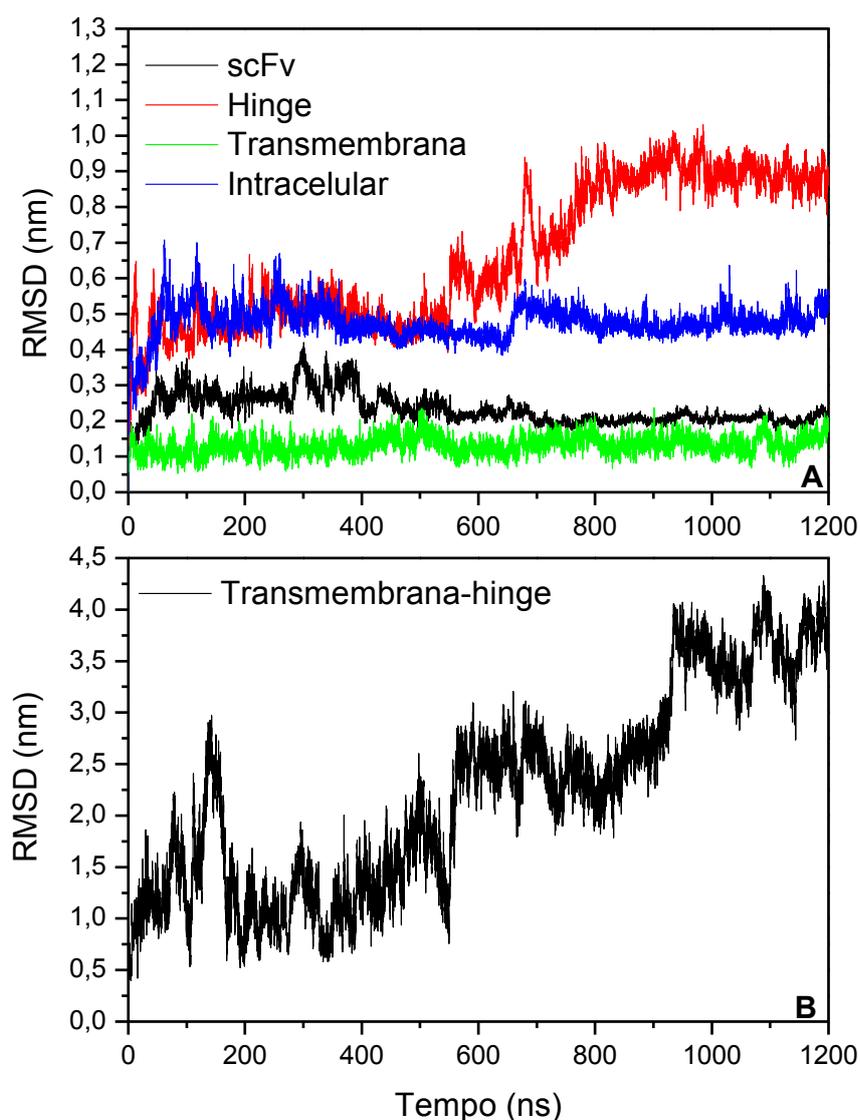


Fonte: elaborada pelo autor.

5.2.8 Modelagem e Dinâmica Molecular do CAR

O RMSD foi calculado para os componentes do CAR com o intuito de inferir a variação de cada uma dessas estruturas no sistema. O perfil de RMSD apresentado na Figura 34A infere que o scFv e o domínio transmembrana alcançam equilíbrio estrutural em $teq = 400$ ns. O perfil de RMSD do *hinge* em relação ao domínio transmembrana (Figura 34B) sugere que há mudança conformacional do *hinge* no sistema, devido a oscilações de grande amplitude. O equilíbrio estrutural do domínio intracelular ocorre em $teq = 700$ ns com RMSD médio de 0.48 ± 0.02 nm. No caso do *hinge*, em $teq = 800$ ns de simulação com RMSD médio de 0.89 ± 0.03 nm (Tabela 10), devido à sua flexibilidade, o que confere maior variação estrutural ao CAR. A partir de 800 ns, o scFv se mantém posicionado na membrana e não há movimentos que indicam afastamento até 1200 ns (Figura 36).

Figura 34 – Perfis de RMSD para os componentes do CAR: scFv, *hinge*, transmembrana e intracelular. A) Curva de RMSD das configurações estruturais do scFv (sobreposição e cálculo de RMSD do scFv), *hinge* (sobreposição e cálculo de RMSD do *hinge*), transmembrana (sobreposição e cálculo de RMSD do transmembrana) e intracelular (sobreposição e cálculo de RMSD do intracelular); B) Curva de RMSD das configurações estruturais do *hinge* após a sobreposição do domínio transmembrana (Transmembrana-*hinge*) ao longo do tempo de simulação, comparadas com a estrutura de referência inicial.



Fonte: elaborada pelo autor.

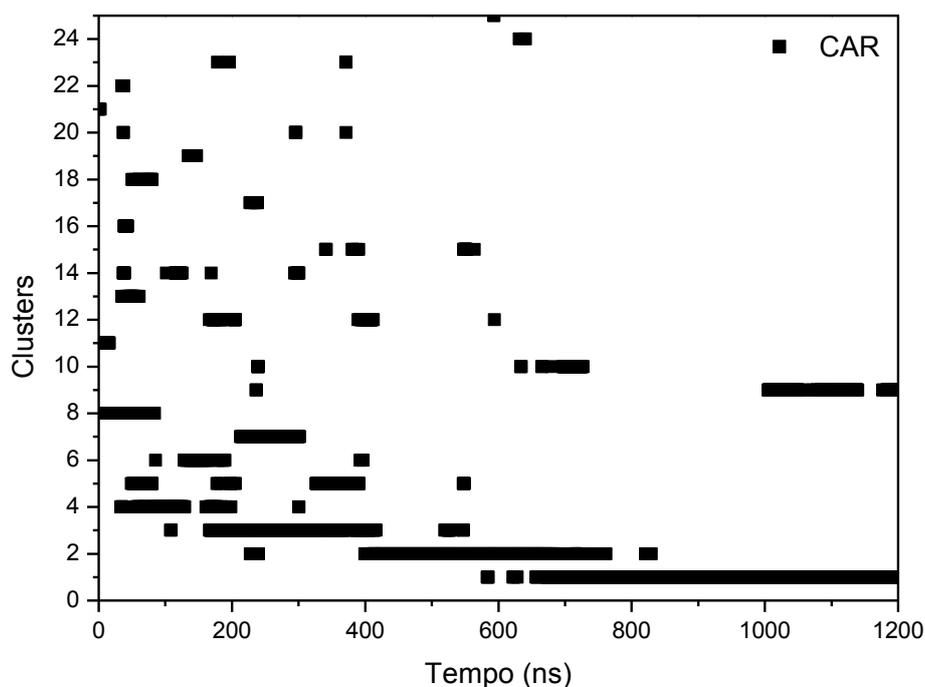
Tabela 10 - Médias e desvios de RMSD calculados para os componentes do CAR em solução calculados a partir do *teq*.

Estrutura	RMSD (nm)	<i>teq</i> (ns)
scFv	0.21 ± 0.02	400
<i>Hinge</i>	0.89 ± 0.03	800
Transmembrana	0.13 ± 0.02	400
Intracelular	0.48 ± 0.02	700

Fonte: elaborada pelo autor.

A análise de *cluster* (Figura 35) mostra a presença de 25 *clusters*, sendo o *cluster* 1 o mais representativo. A presença de muitos *clusters* no início da simulação ocorre devido à flexibilidade do *hinge* que movimenta também a estrutura do CAR. Por volta de 600 ns, a conformação da estrutura do scFv se aproxima da membrana (Figura 37) e interage atrativamente com a mesma (Figura 36C).

Figura 35 – Resultado da análise de *cluster* para o CAR. *Clusters* de configurações do CAR ao longo da simulação, utilizando um raio de corte de 0.6 nm.

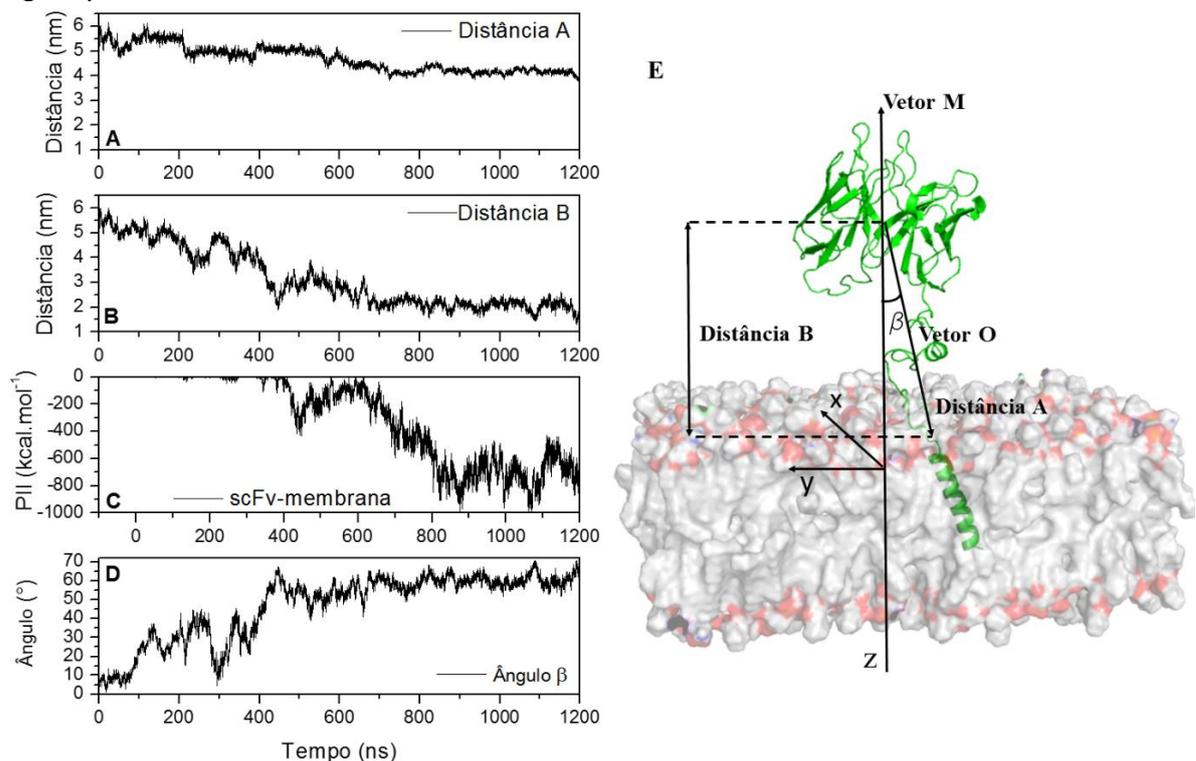


Fonte: elaborada pelo autor.

A distância entre o centro de massa do scFv e o carbono alpha ($C\alpha$) do Trp301 da α -hélice do domínio transmembrana (Figura 36E), localizado na extremidade extracelular da membrana (Distância A), foi calculada com intuito de observar a aproximação entre o scFv e a

membrana. Na figura 36A, está representada a Distância A, em que se pode observar a redução da distância entre o scFv e a membrana de 6 para $\sim 3,5$ nm. A distância em z do centro de massa do scFv e do Trp301 (Distância B) varia como mostrado na Figura 36B, reduzindo de 6 para ~ 1 nm. A Distância B reduz a partir de 200 ns, mesmo instante em que o PII entre o scFv e a membrana também reduz, evidenciando uma interação mais atrativa entre o scFv e a membrana (scFv-membrana) até o final da DM (Figura 36C). Esse comportamento é confirmado pelo monitoramento do ângulo β formado entre o vetor O, formado sobre a posição do centro de massa do scFv e do C α do resíduo Trp301 da α -hélice do domínio transmembrana, e o vetor projetado no eixo z (vetor M). A curva de ângulo β em função do tempo de DM, na Figura 36D, mostra que o ângulo é de 70° em ~ 400 ns, ou seja, no mesmo tempo em que foi observada a menor distância projetada em z entre o scFv e a membrana (Figura 36B). A inclinação do scFv se mantém a partir de 800 ns, com ângulos variando entre 50 e 70° , ao mesmo tempo em que a Distância B se mantém, com o scFv interagindo atrativamente com a membrana (Figura 36C).

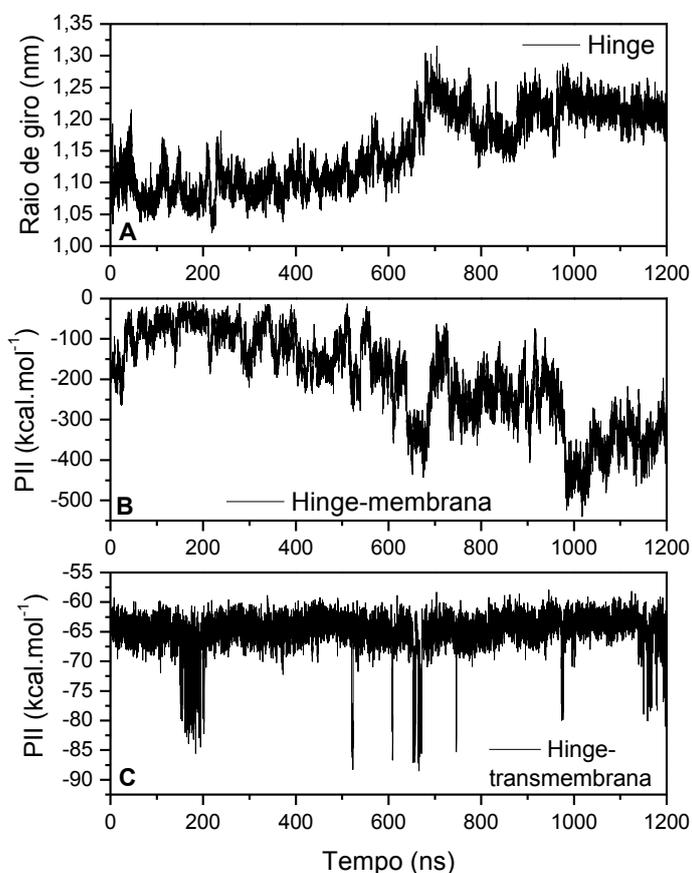
Figura 36 – Distância e PII entre o scFv e a membrana. A) Distância entre o centro de massa do scFv e o carbono alpha (C α) do resíduo Trp301 da α -hélice do domínio transmembrana (Distância A). B) Distância projetada em z do centro de massa do scFv e o C α do resíduo Trp301 do domínio transmembrana (Distância B). C) PII scFv-membrana. D) Ângulo β formado entre os vetores M e O. E) Representação das Distâncias A e B e dos vetores M e O, formando o ângulo β .



Fonte: elaborada pelo autor.

Para monitorar a compacidade do *hinge* no CAR ao longo da DM foi calculado o Raio de giro (R_g) do *hinge* e o PII entre o *hinge* e a membrana (*hinge*-membrana) e entre o *hinge* e o domínio transmembrana (*hinge*-transmembrana) (Figura 37). O Raio de giro (R_g) do *hinge* aumenta até 1.05 a 1.3 nm logo após o scFv se aproximar da membrana, em 700 ns (Figura 37A), o que indica que não há mudança na compacidade do *hinge*. Em 900 ns, o R_g se mantém entre 1.15 e 1.25 nm. O PII *hinge*-membrana é mais atrativo em ~ 1000 ns, de -550 kcal.mol $^{-1}$ tempo em que o scFv também interage mais atrativamente com a membrana (Figura 36C). Em 640 ns, o PII *hinge*-transmembrana é mais atrativo, -90 kcal.mol $^{-1}$ (Figura 37C), quando o scFv se aproxima da membrana.

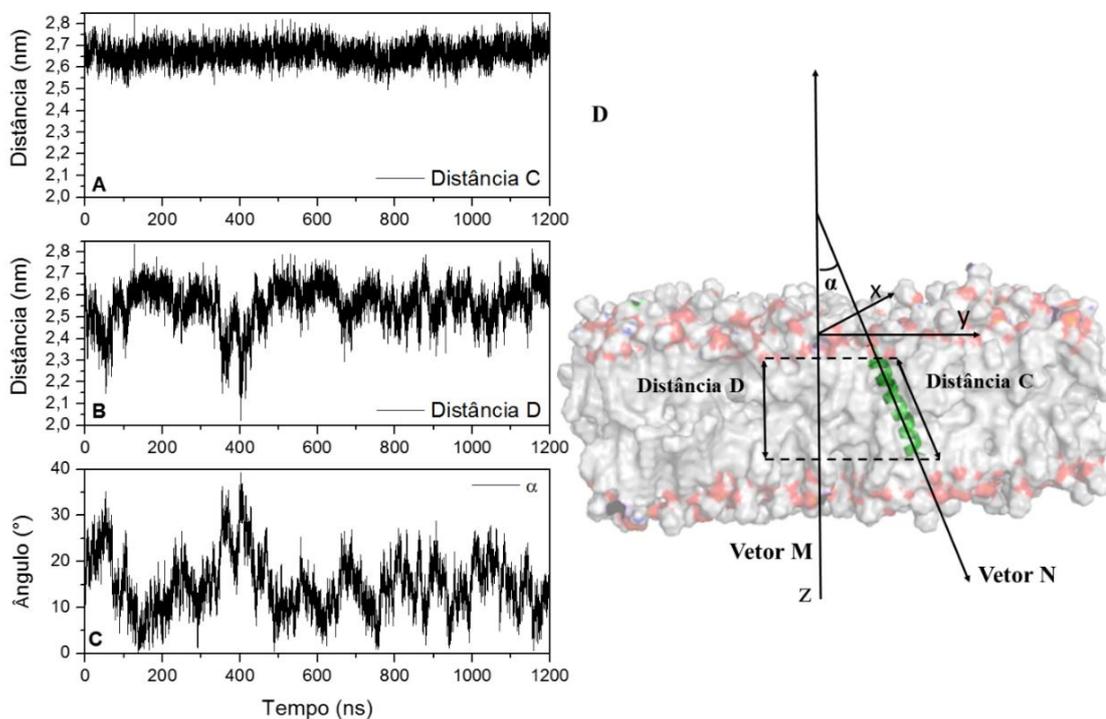
Figura 37 – Raio de giro (R_g) do *hinge* e PII entre o *hinge* e a membrana (*hinge*-membrana) e, o domínio transmembrana (*hinge*-transmembrana). A) R_g do *hinge* ao longo do tempo. B) PII *hinge*-membrana. C) PII *hinge*-transmembrana.



Fonte: elaborada pelo autor.

Com intuito de monitorar o comportamento da região transmembrana durante a aproximação do scFv à membrana, foi calculada a distância entre os carbonos alpha ($C\alpha$) dos resíduos Trp301 e Leu319 (Distância C) (Figura 38D), resíduos das extremidades extra e intracelular da α -hélice do domínio transmembrana. A Distância C varia entre 2,5 e 2,8 nm (Figura 38A), o que infere manutenção da estrutura da α -hélice do domínio transmembrana. No entanto, observa-se uma maior variação da distância entre as coordenadas em z dos átomos $C\alpha$ dos resíduos Trp560 e Leu578 (Distância D), descrita na Figura 38B. Sendo que em 400 ns, a Distância D reduz de $\sim 2,6$ para 2 nm e, em seguida, aumenta até 2,8 nm. Uma vez que a distância entre os $C\alpha$ reduz bem menos que a distância em z e que a estrutura em α -hélice do domínio transmembrana é mantida, conclui-se que a α -hélice (domínio transmembrana) sofre uma inclinação em relação ao eixo z. Isso é confirmado pelo monitoramento do ângulo α entre o vetor N, formado pela posição dos $C\alpha$ dos resíduos Trp301 e Leu319, e o vetor sobre o eixo z (vetor M), lançado em gráfico na Figura 38C. Em 400 ns, quando a Distância D é menor, o ângulo α alcança 40° , o maior ângulo observado. Em seguida, quando a Distância D aumenta, há uma redução do ângulo α , que permanece oscilando entre 0 e 20° .

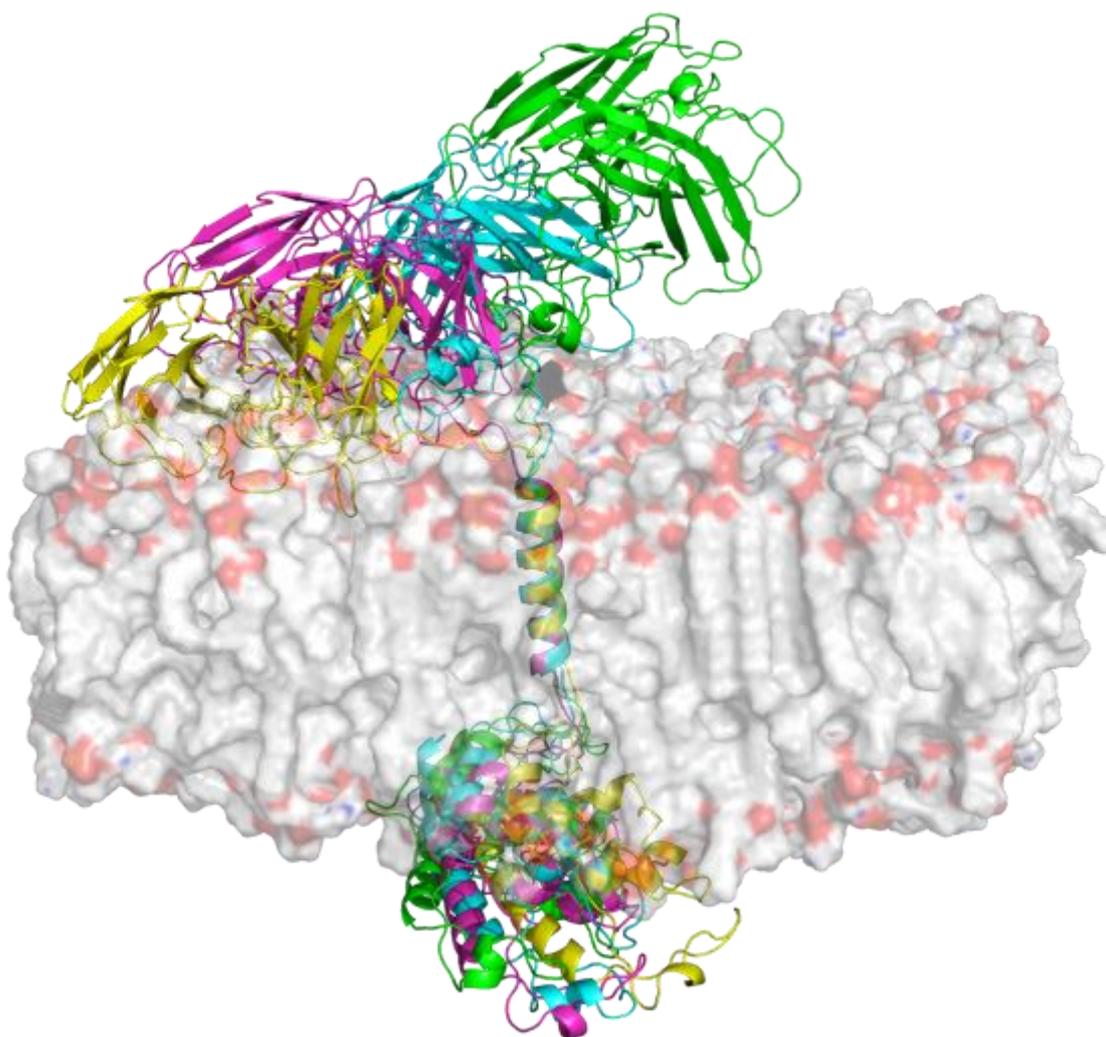
Figura 38 – Distância C e D e ângulo α , entre os vetores M e N. A) Distância entre os $C\alpha$ dos resíduos Trp301 e Leu319 do domínio transmembrana (Distância C). B) Distância em z entre as Trp301 e Leu319 do domínio transmembrana projetada no eixo z dos resíduos (Distância D). C) Ângulo α formado entre os vetores M e N. D) Representação das Distâncias C e D e dos vetores M e N, formando o ângulo α .



Fonte: elaborada pelo autor.

A configuração escolhida para montar o complexo CAR-CD19 foi a do início da simulação, antes do scFv se aproximar da membrana, para comparação entre a estrutura do CAR livre com o CAR interagindo com o CD19. Como se pode ver na Figura 39, a estrutura do scFv se aproxima da membrana durante a simulação, permanecendo interagindo com a mesma até o final, como mostrado no gráfico de PII scFv-membrana (Figura 36C).

Figura 39 – Representação estrutural do CAR em membrana. Conformações do CAR inserido em membrana após 200 (verde), 600 (azul claro), 1000 (roxo) e 1200 ns (amarelo), sobrepondo-se a estrutura do domínio transmembrana.

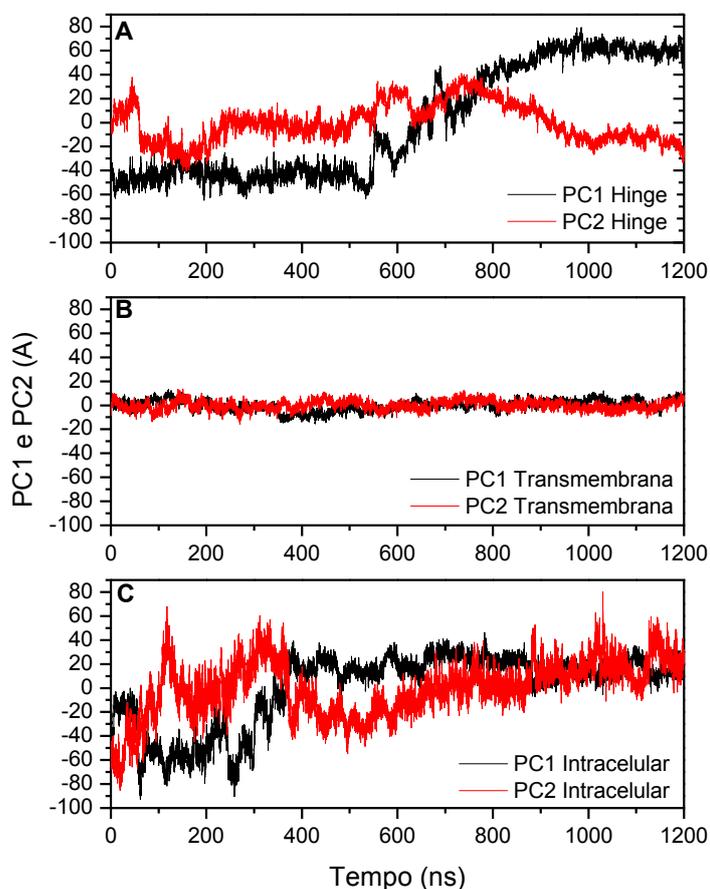


Fonte: elaborada pelo autor.

A Análise de Componentes Principais (PCA) em função do tempo de DM para cada uma das projeções PC1 e PC2 foi lançada em gráfico para observar a variação dos movimentos do *hinge*, dos domínios transmembrana e intracelular no CAR. Pode-se observar na Figura 40 a movimentação das projeções dos resíduos do *hinge*, dos domínios transmembrana e

intracelular. A estrutura que se movimenta mais durante a simulação é o *hinge*, uma vez que suas projeções oscilam mais (Figura 40A) durante a DM. A PC1 do *hinge* não varia consideravelmente até por volta de 600 ns de simulação, quando começa a aumentar de -60 a ~70 Å de 600 a 900 ns, mantendo-se até o final da simulação, quando o *hinge* já não se movimenta tanto, pois o scFv se aproximou e está bem próximo da membrana. Já a PC2, oscila até 900 ns entre -40 e 40 Å, e mantém-se por volta de -20 Å até o final. As projeções do domínio transmembrana oscilam entre -20 e 20 Å (Figura 40B), o que indica uma movimentação bem pequena dessa estrutura comparado ao *hinge*. As projeções da estrutura do domínio intracelular se movimentam entre -80 e 60 Å até 400 ns (Figura 40C), mantendo-se oscilando por volta de 20 Å até o final da DM.

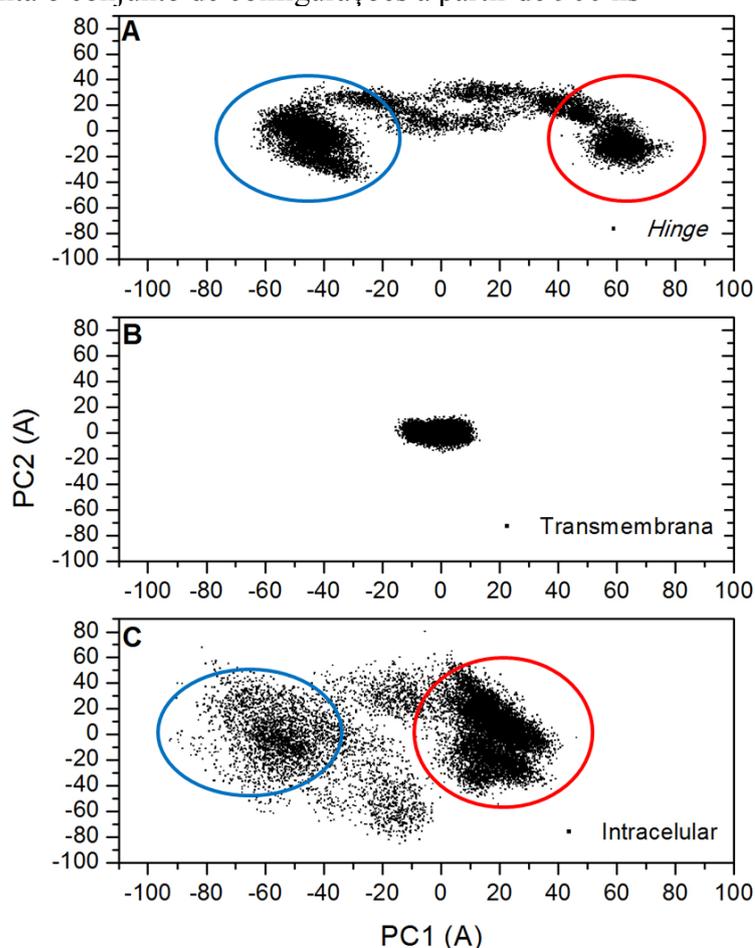
Figura 40 - Projeções dos componentes do CAR (PC1 e PC2) por tempo de simulação. PC1 e PC2 da estrutura do A) *hinge*; B) do domínio transmembrana; C) do domínio intracelular por tempo de simulação.



Fonte: elaborada pelo autor.

A Figura 41 mostra as projeções comuns, PC1 e PC2, do *hinge*, dos domínios transmembrana e intracelular. Pode-se observar uma maior variação da estrutura do *hinge* em relação aos dois domínios, indicada pela maior oscilação de suas projeções, principalmente em PC1 (Figura 41A). As projeções da estrutura do domínio transmembrana se mantêm próxima a inicial entre 20 e -20 Å (Figura 41B). Já o domínio intracelular varia menos que o *hinge* (Figura 41C), com manutenção da estrutura no final da DM.

Figura 41 - Projeções da simulação do CAR nos subespaços comuns definidos por PC1/PC2. PC1 e PC2 da estrutura do A) *hinge*; B) do domínio transmembrana; C) do domínio intracelular. O círculo azul representa o conjunto de configurações até 600 ns e o círculo vermelho representa o conjunto de configurações a partir de 900 ns



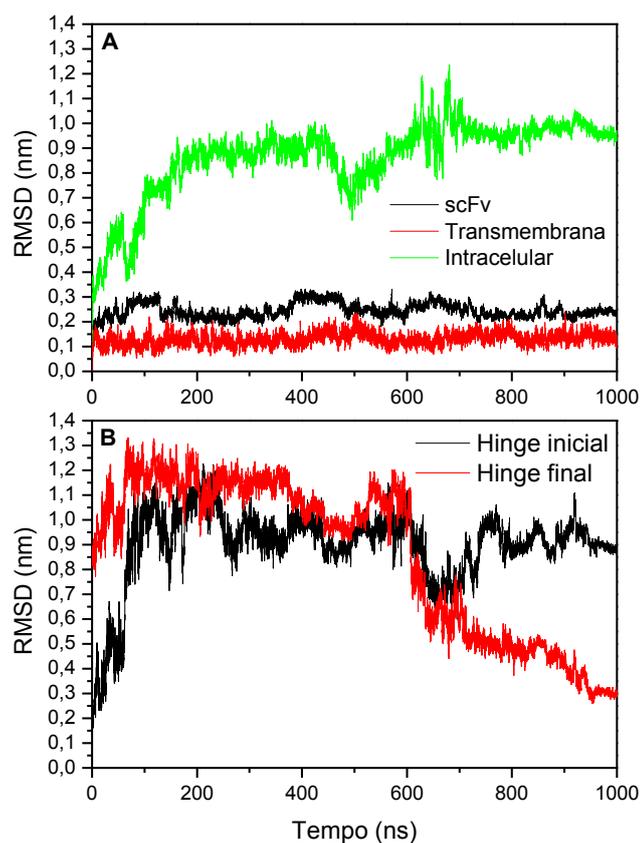
Fonte: elaborada pelo autor.

5.2.9 Dinâmica Molecular do CAR-CD19

O RMSD foi calculado para os componentes do CAR com o intuito de inferir a variação de cada uma dessas estruturas enquanto o CAR está mantido em interface com o CD19. O perfil

de RMSD apresentado na Figura 42 indica que a estrutura do scFv se mantém em equilíbrio estrutural após $teq \sim 400$ ns, com RMSD médio de 0.13 ± 0.02 nm (Tabela 11). A estrutura do domínio transmembrana entra em equilíbrio em 100 ns com RMSD médio de 0.27 ± 0.01 nm (Tabela 11). O perfil de RMSD para a estrutura do domínio intracelular oscila com grande amplitude até ~ 800 ns, quando entra em equilíbrio com RMSD médio de 0.97 ± 0.03 nm (Tabela 11). Os RMSDs após sobreposição das estruturas inicial (*hinge* inicial) ou a final da DM (configuração do *hinge* final) estão apresentados na Figura 42B. Em 600 ns, há uma mudança estrutural no *hinge* que permite a aproximação do *hinge* à membrana, como mostrado na Figura 44B com redução do PII *hinge*-membrana.

Figura 42 – Perfis de RMSD para os componentes do CAR: scFv, *hinge*, transmembrana e intracelular. A) Curva de RMSD das configurações estruturais do scFv (sobreposição e cálculo de RMSD do scFv), transmembrana (sobreposição e cálculo de RMSD do transmembrana) e intracelular (sobreposição e cálculo de RMSD do intracelular ao longo do tempo de simulação, comparadas com a estrutura de referência inicial. B) RMSD em função do tempo de simulação para o *hinge*. A curva preta mostra o RMSD utilizando como referência para sobreposição a estrutura inicial (*hinge* inicial) e a curva vermelha, tendo a estrutura final como referência (*hinge* final).



Fonte: elaborada pelo autor.

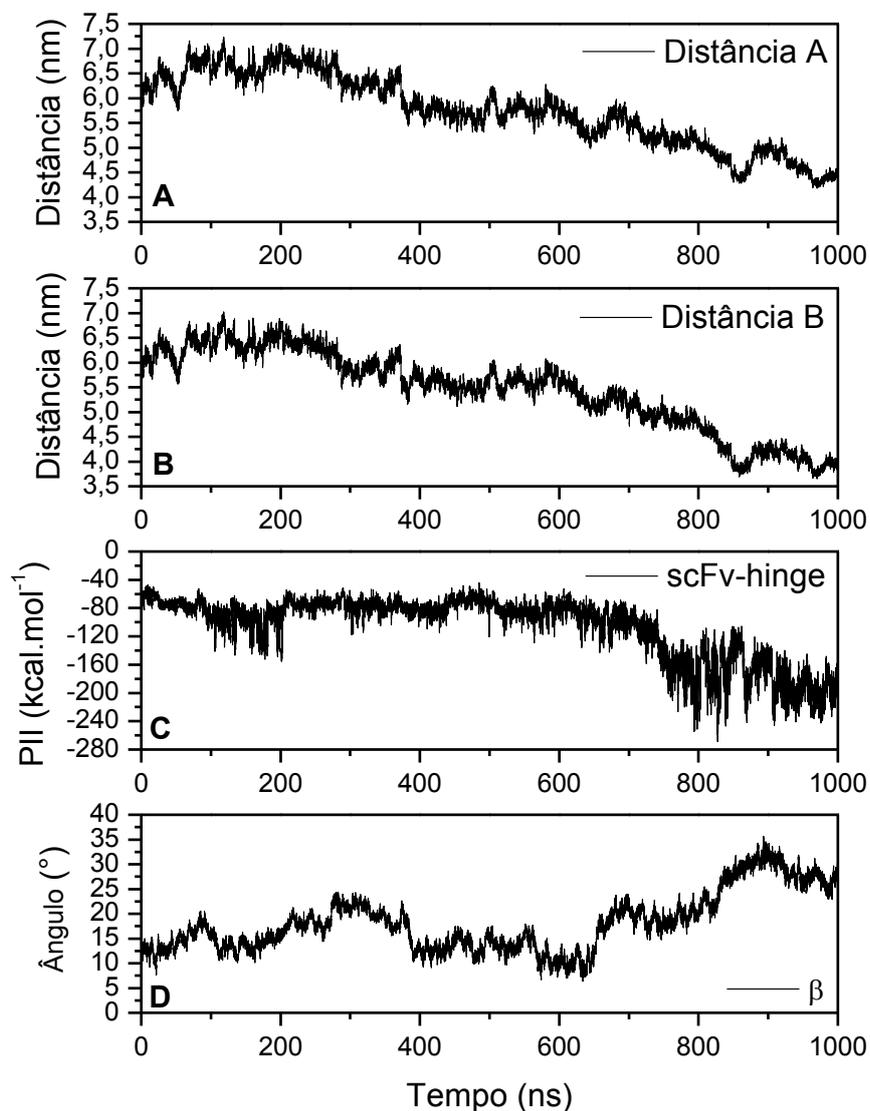
Tabela 11- Médias e desvios de RMSD calculados para os componentes do CAR em solução calculados a partir do *teq*.

Estrutura	RMSD (nm)	<i>teq</i> (ns)
scFv	0.24 ± 0.02	400
<i>Hinge</i>	0.87 ± 0.02	800
Transmembrana	0.13 ± 0.02	100
Intracelular	0.97 ± 0.03	800

Fonte: elaborada pelo autor.

A distância entre o centro de massa do scFv e o carbono alpha ($C\alpha$) do Trp560 da α -hélice do domínio transmembrana (Distância A), localizado na porção extracelular da membrana foi monitorada, como mostra a Figura 43A. Há um aumento da Distância A no início da DM de 6 para ~ 7 nm, mas após 200 ns, essa distância reduz a ~ 4.5 nm em ~ 860 ns, como mostrado na Figura 43A. A distância entre o centro de massa do scFv e o carbono alpha ($C\alpha$) do Trp560, projetada no eixo z (Distância B), varia de 6 a ~ 7 nm até 200 ns e reduz a próximo de 4 nm nos últimos 400 ns (Figura 43B). A redução da Distância A ocorre devido à uma mudança estrutural no *hinge* em 600 ns, observada nas curvas de RMSDs na Figura 42B. Observa-se, também, que a aproximação do scFv da membrana é acompanhada por uma redução do Potencial de Interação (PII) scFv-*hinge* (Figura 43C), que é evidenciada a partir de 640 ns. Os eventos sucessivos indicam que o *hinge* sofre uma mudança conformacional importante, que é indicada pelos RMSDs (Figura 42B), em $t = 600$ ns e que possibilita o scFv se aproximar da membrana, como visto na sua maior aproximação em $t = 890$ ns, que é acompanhada por um PII atrativo entre o *hinge* e o scFv. Esse comportamento é confirmado pelo monitoramento do ângulo β formado entre o vetor O, formado sobre a posição do centro de massa do scFv e do $C\alpha$ do resíduo Trp560 da α -hélice do domínio transmembrana, e o vetor projetado no eixo z (vetor M). A curva de ângulo em função do tempo de DM, na Figura 43D, mostra que o ângulo é de 35° em $t = 890$ ns, ou seja, no mesmo tempo em que foi observada a menor distância projetada em z entre o scFv e a membrana (Figura 43B). A inclinação do scFv reduz, após 900 ns, a 27° no final da DM (Figura 43D).

Figura 43 - Distância entre o scFv e a membrana e PII scFv-*hinge*. A) Distância entre o centro de massa do scFv e carbono alpha ($C\alpha$) do Trp560 da α -hélice do domínio transmembrana (Distância A). B) Distância em z do centro de massa do scFv e do $C\alpha$ do Trp560 do domínio transmembrana (Distância B). C) PII entre o scFv e o *hinge* (scFv-*hinge*). D) Ângulo β formado entre os vetores M e O.

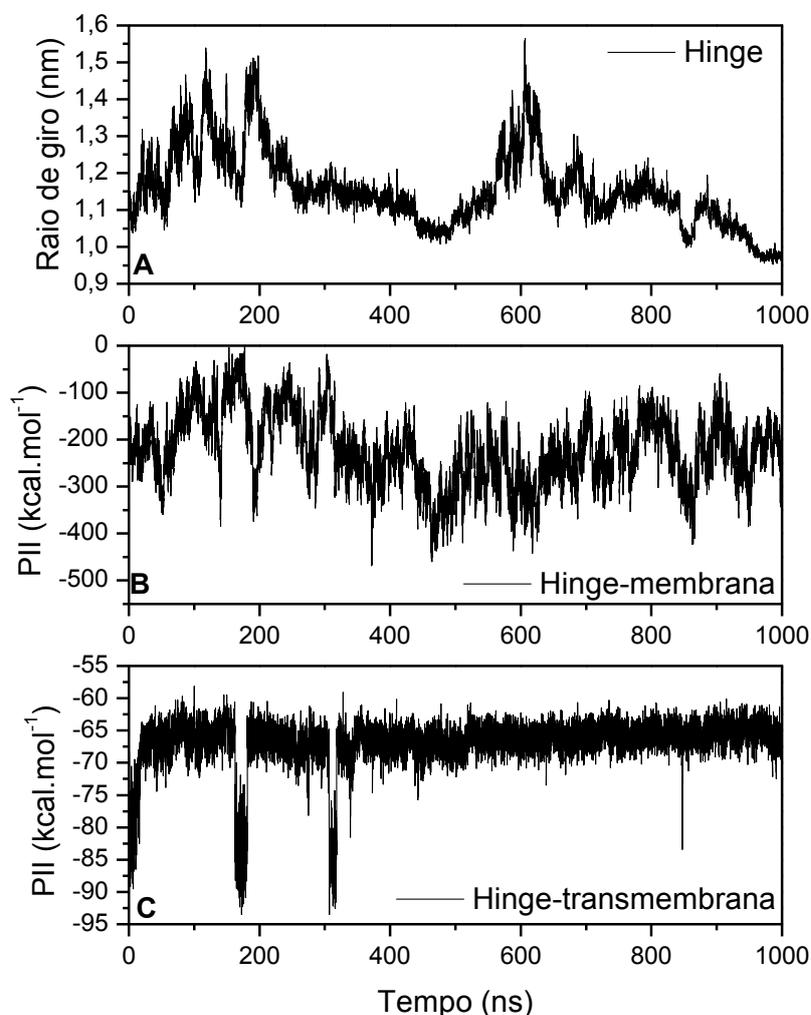


Fonte: elaborada pelo autor.

Para monitorar a compacidade do *hinge* no CAR interagindo com o CD19 ao longo da DM foi calculado o Raio de giro (R_g) do *hinge* e o PII entre o *hinge* e a membrana (*hinge*-membrana) e entre o *hinge* e o domínio transmembrana (Figura 44). O Raio de giro (R_g) do *hinge* aumenta oscila entre 1.1 e 1.5 nm até 200 ns. Em seguida, reduz a 1 em 500 ns, e aumenta

com um pico de 1.55 nm em 600 ns (Figura 44A), quando há aproximação *scFv-hinge* com redução do PII (Figura 43C). Essa variação no R_g , em 200 ns e em 600 ns, indica uma descompactação estrutural importante do *hinge*, devendo-se melhor investigar o papel dessa variação estrutural na sinalização da célula CAR T. O PII *hinge*-membrana oscila entre 0 e -475 kcal.mol⁻¹ até ~ 500 ns (Figura 44B), indicando interação mais atrativa. Em seguida, o PII se mantém entre -450 e -200 kcal.mol⁻¹ até o final da DM (Figura 44B). O PII *hinge*-transmembrana é mais atrativo em 175 ns, e, em ~ 300 ns, sendo -90 kcal.mol⁻¹ (Figura 44C).

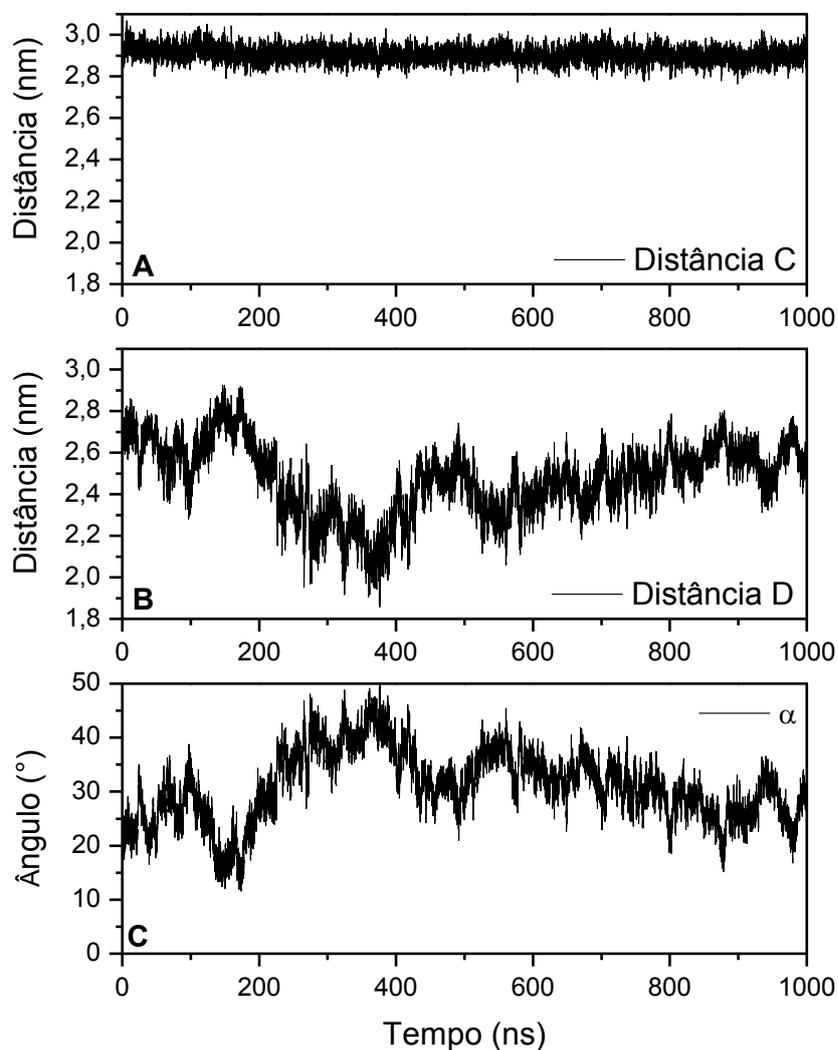
Figura 44 - Raio de giro (R_g) do *hinge* e PII entre o *hinge* e a membrana (*hinge*-membrana) e, o domínio transmembrana (*hinge*-transmembrana). A) R_g do *hinge* ao longo do tempo. B) PII *hinge*-membrana. C) PII *hinge*-transmembrana.



Fonte: elaborada pelo autor.

Com intuito de monitorar o comportamento da região transmembrana durante a aproximação do *hinge* ao scFv, foi calculada a distância entre os carbonos alpha ($C\alpha$) dos resíduos Trp560 e Leu578 (Distância C) (Figura 45A), resíduos das extremidades extra e intracelular da α -hélice do domínio transmembrana. A Distância C varia entre 2,8 e 3 nm, sendo bastante conservada e oscilando com alta frequência, como esperado para uma estrutura secundária em α -hélice. O que mostra que a estrutura é bastante estável na membrana ao longo da DM. No entanto, a distância projetada no eixo z dos $C\alpha$ dos resíduos Trp560 e Leu578 (Distância D), descrita na Figura 45B, reduz de 2,8 para 1.9 nm em ~ 370 ns e, em seguida, aumenta até 3,2 nm. O RMSD indica que a estrutura transmembrana é conservada durante a DM (Figura 42A), também, a variação da distância entre os $C\alpha$ é mantida em 2,9 nm (Figura 45A). Portanto, a redução da Distância D ao longo da DM revela que a α -hélice do domínio transmembrana inclina em relação ao eixo z, perpendicularmente a membrana. Esse comportamento é ratificado pelo monitoramento do ângulo α formado entre o vetor N, formado sobre a posição dos carbonos alpha ($C\alpha$) dos resíduos Trp560 e Leu578 da α -hélice do domínio transmembrana, e o vetor projetado no eixo z (vetor M). A curva do ângulo α em função do tempo de DM, na Figura 45C, mostra que o ângulo é de 50° em 370 ns, ou seja, no mesmo tempo onde foi observado a menor Distância D. Em seguida, há uma redução da inclinação da α -hélice, com ângulos variando entre 40 e 20° , ao mesmo tempo a Distância D aumenta.

Figura 45 – Distância C e D e ângulo α , entre os vetores M e N. A) Distância entre os $C\alpha$ dos resíduos Trp560 e Leu578 do domínio transmembrana (Distância C). B) Distância em z entre as Trp301 e Leu319 do domínio transmembrana projetada no eixo z dos resíduos (Distância D). C) Ângulo α formado entre os vetores M e N.

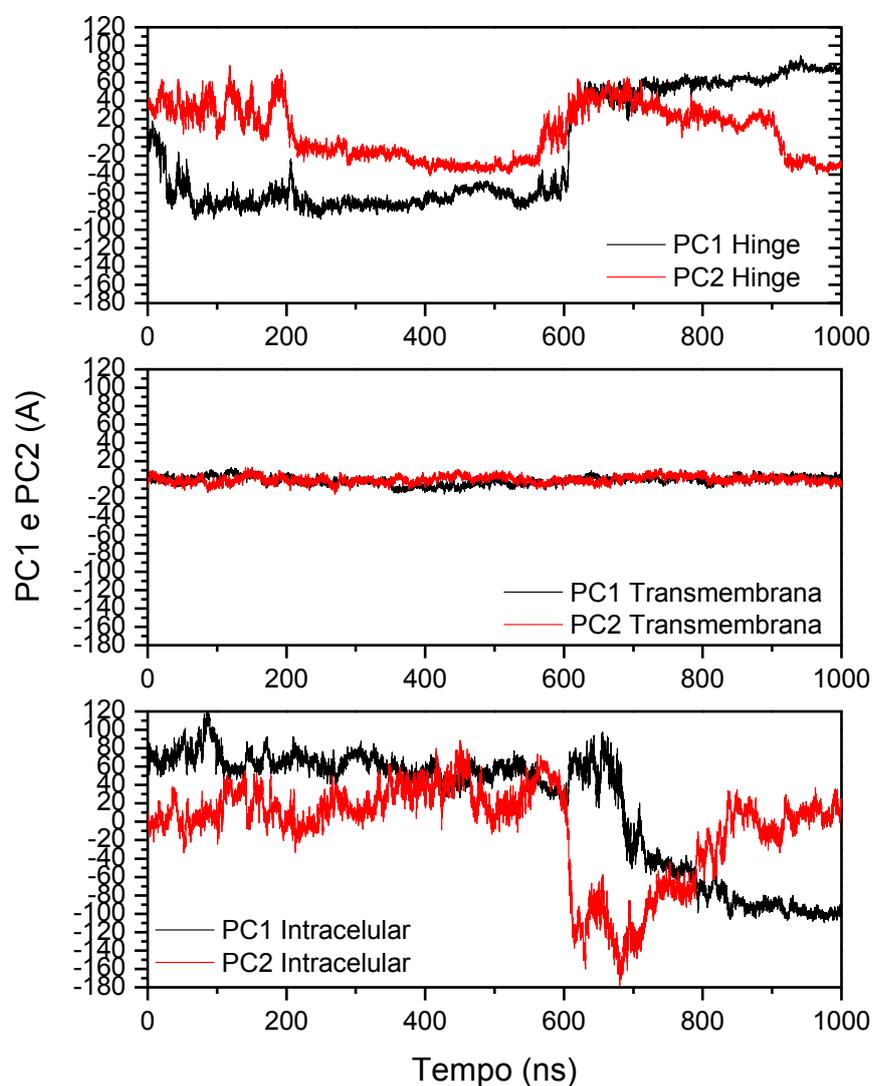


Fonte: elaborada pelo autor.

A Análise de Componentes Principais (PCA), em função do tempo, para cada uma das projeções PC1 e PC2 foi realizada para observar a variação de movimentos no *hinge*, nos domínios transmembrana e intracelular no CAR interagindo com o CD19. Pode-se observar na Figura 46 a movimentação das projeções dos resíduos do *hinge*, dos domínios transmembrana e intracelular. A estrutura que se movimenta mais durante a simulação é o domínio intracelular, uma vez que são observadas grandes amplitudes de PC1 e PC2 ao longo da DM (Figura 46C).

As curvas de PC1 e PC2 do *hinge* variam entre intervalos de tempo, antes de ~ 200 ns foi observado movimentos mais intensos, e entre ~ 200 ns e 600 ns os movimentos são menores e voltam a se intensificar após ~ 600 ns, chegando a 80 \AA no caso de PC1. As projeções do domínio transmembrana oscilam entre -15 e 15 \AA (Figura 46B) até 1000 ns, o que indica uma movimentação bem pequena dessa estrutura comparada ao *hinge*. As projeções da estrutura do domínio intracelular oscilam até ~ 600 ns de simulação, PC1 entre 40 e 120 \AA e PC2 entre -20 e 80 \AA (Figura 46C). Após 680 ns, há uma redução da PC1 a -40 \AA , que se mantém reduzindo até $\sim 100 \text{ \AA}$, indicando uma mudança na movimentação desse domínio. No caso da PC2 ocorre uma forte transição de movimentação, que era inicialmente nula e em 600 ns vai a -180 \AA e volta a ser nula em 850 ns. As mudanças nas movimentações corroboram as fases temporais e subsequentes de mudanças na estrutura do *hinge*, com concomitante aproximação entre scFv e membrana e consequente aproximação scFv/*hinge*, com uma formação de interface e realce no PII atrativo entre scFv-*hinge*. Nesse sentido, o PC2 para *hinge* é próximo de nulo e o PC1 se mantém alto, mas com mesmo padrão de movimentação (Figura 46A).

Figura 46 - Projeções dos componentes do CAR (PC1 e PC2) por tempo de simulação. PC1 e PC2 da estrutura A) do *hinge*; B) do domínio transmembrana; C) do domínio intracelular por tempo de simulação.

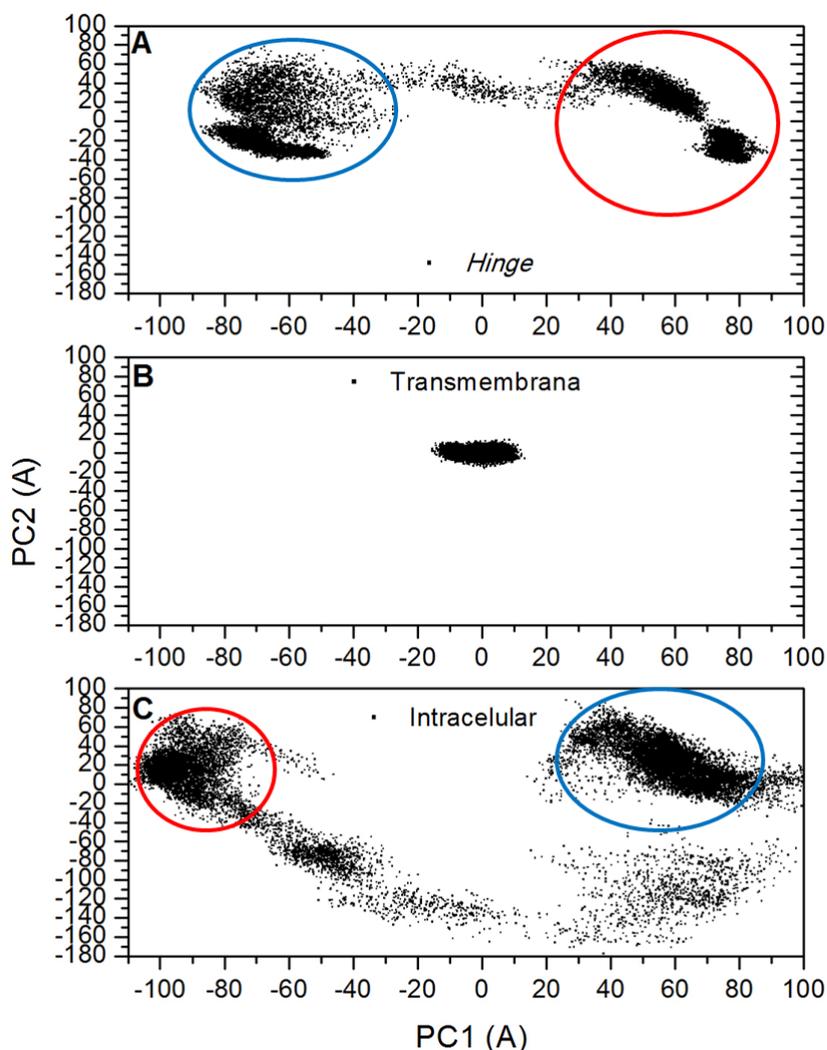


Fonte: elaborada pelo autor.

A Figura 47 mostra as projeções comuns, PC1 e PC2, do *hinge*, dos domínios transmembrana e intracelular. Pode-se observar uma maior variação da estrutura do intracelular em relação ao *hinge* e transmembrana, indicada pela maior oscilação de suas projeções (Figura 47C). As projeções da estrutura do domínio transmembrana se mantêm próxima à inicial (Figura 47B), como já observado na Figura 46B, oscilando entre 15 e -15 Å. As PCs do domínio intracelular oscilam bastante (Figura 47C), o que caracteriza uma maior variação estrutural do

domínio intracelular com um maior conjunto de conformações diferentes durante a DM.

Figura 47 - Projeções da simulação do CAR interagindo com o CD19 nos subespaços comuns definidos por PC1/PC2. PC1 e PC2 da estrutura A) do *hinge*; B) do domínio transmembrana. C) do domínio intracelular. O círculo azul representa o conjunto de configurações até 200 ns e o círculo vermelho representa o conjunto de configurações a partir de 600 ns.



Fonte: elaborada pelo autor.

6 DISCUSSÃO

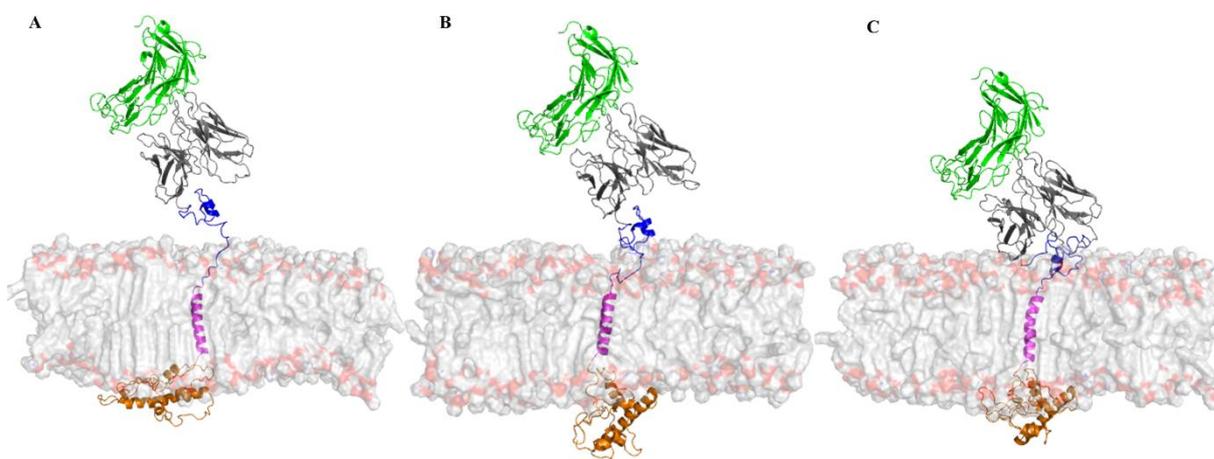
Os modelos do scFv, do *hinge*, do domínio transmembrana e intracelular foram obtidos e validados. As estruturas dos componentes do CAR foram submetidas a DM e equilibradas, em solução, e assim foi possível montar o CAR e analisar sua movimentação em interface membrana/água e em interação com o CD19. Além disso, foi possível também estudar a interação scFv-CD19, determinando quais regiões CDR mais contribuem para uma interação

atrativa com o CD19 e, também, quais resíduos dos *loops* do epítopo do CD19 mais contribuem para uma interação atrativa ao scFv. Sendo os resíduos dos *loops* carregados positivamente do CD19 que possuem PII mais negativos, portanto, mais atrativos ao scFv.

O *hinge* localizado entre a região de reconhecimento (scFv) e o domínio transmembrana é importante para garantir o correto posicionamento do sítio de ligação durante a interação CAR-antígeno (SHIRASU; KUROKI, 2012). No caso de CARs anti-CD19, já foi demonstrado que o reconhecimento do CD19, *in vitro*, é possível utilizando *hinges* de diferentes tamanhos, de 12 a 229 resíduos. Embora as células tumorais CD19+ sejam reconhecidas, observa-se maior produção e proliferação de citocinas, utilizando-se células CAR T com *hinges* curtos comparado àquelas com *hinges longos* (HUDECEK *et al.*, 2013a, 2015), o que pode ocorrer devido à estrutura e/ou localização do CD19 ou a alta densidade do CD19 nas células tumorais (DU *et al.*, 2008).

Comparando-se a distância entre o centro de massa do scFv e o C α do resíduo Trp da extremidade extracelular da α -hélice do domínio transmembrana (Distância A) nos sistemas CAR e CAR-CD19, observa-se que o scFv se aproxima da membrana no sistema CAR (Figuras 36A e 43A). Já a distância entre a projeção em z do centro de massa do scFv e do C α do resíduo Trp do domínio transmembrana (Distância B) varia com maior amplitude no sistema CAR (Figura 36B), comparada a variação no sistema CAR-CD19 (Figura 43B), devido à maior aproximação do scFv do CAR à membrana, realçando a interação atrativa scFv-membrana, como visto na Figura 36C. O ângulo β formado entre os vetores M e O foi maior no sistema CAR (Figura 36D), no instante em que o scFv se aproxima da membrana e o PII começa a reduzir (Figura 36B e C). Isso implica que, durante a aproximação do scFv da membrana no sistema CAR, ocorreu uma reorientação entre o domínio transmembrana e o scFv, que foi mais evidente em 860 ns, tempo em que a distância entre esses componentes foi menor e a interação scFv-membrana foi mais atrativa (Figura 36C). No caso em que o CAR interage com o CD19, a redução da distância em z (Figura 43B) implica no aumento do ângulo β , e na orientação do scFv se voltando para o plano da membrana após 500 ns, como mostrado na Figura 48. Nesse mesmo instante, há redução do PII scFv-*hinge* (Figura 43C), indicando interação mais atrativa scFv-*hinge* ao reduzir a distância scFv-transmembrana. O que evidencia o papel do *hinge* após a interação CAR-CD19 de manter a flexibilidade e o correto posicionamento scFv-CD19.

Figura 48 – Representação estrutural do complexo CAR-CD19. Conformações CAR-CD19 em A) 200; B) 600 e C) 1000 ns.



Fonte: elaborada pelo autor.

Embora a principal função do domínio transmembrana seja ancorar o CAR na membrana da célula T, algumas evidências sugerem que esse domínio pode ser relevante para a função da célula CAR T. O domínio transmembrana pode estar envolvido na sinalização de células T (MORIN *et al.*, 2015; GUEDAN *et al.*, 2019). A ligação do domínio intracelular próximo ao domínio transmembrana pode influenciar na sinalização adequada das células CAR T, enquanto o uso dos domínios transmembrana de CD8 α ou CD28, já amplamente utilizados, pode favorecer a expressão ou estabilidade do CAR (IMAI *et al.*, 2004; KOCHENDERFER *et al.*, 2009).

Os resultados das distâncias entre os C α dos resíduos Trp e Leu das extremidades extra e intracelular da α -hélice do domínio transmembrana (Distância C) e a projeção em z dessa distância (Distância D) nos sistemas CAR e CAR-CD19 permitem inferir a movimentação e a inclinação em relação ao eixo z do domínio transmembrana. No sistema CAR, a Distância C oscila entre 2,5 e 2,8 nm (Figura 38A), o que não indica mudança conformacional na estrutura secundária, em forma de α -hélice, do domínio transmembrana. No sistema CAR-CD19, pode-se observar uma maior Distância C, entre 2,8 e 3,1 nm (Figura 45A), mas também indicando estabilidade na estrutura da α -hélice. Já a Distância D aumenta a 2,8 nm e reduz a 2 nm em 400 ns no sistema CAR (Figura 38B), instante em que o ângulo α é máximo, 40° (Figura 38C). Após 400 ns, a Distância D aumenta e o ângulo α reduz, mantendo-se entre 0 e 20°. Já no sistema CAR-CD19, a redução da Distância D é ainda maior, a 1,8 nm (Figura 45B), e o ângulo α também é maior em 400 ns. Portanto, a redução da Distância D ao longo da

DM revela que a α -hélice do domínio transmembrana tem uma maior inclinação em relação ao eixo z, perpendicularmente à membrana, quando comparado ao sistema CAR. Esse comportamento é ratificado pelo monitoramento do ângulo α em função do tempo de DM, na Figura 45C, em que o ângulo é máximo, 50° , em ~ 400 ns, no mesmo tempo onde foi observado a menor Distância D. Em seguida, há uma redução da inclinação da α -hélice, que se mantém oscilando entre 20 e 40° , ao mesmo tempo a Distância D aumenta. É observado, no sistema CAR-CD19, uma mudança conformacional importante no *hinge*, que leva a mudança na inclinação da região transmembrana que pode influenciar na conformação do domínio intracelular.

No sistema CAR-CD19, há uma variação estrutural no *hinge* em 600 ns (Figura 44A), que permite uma maior aproximação e interação mais atrativa entre o *hinge* e o scFv. Após a redução da Distância D (Figura 45B) e o aumento da inclinação desse domínio, que pode ser observada pelo aumento do ângulo α entre os vetores M e N (Figura 45C), há uma aproximação do *hinge* ao scFv (Figura 43C). Em 600 ns, há uma variação na estrutura do domínio intracelular provocada pela inclinação da α -hélice do domínio transmembrana. O que sugere que para ocorrer o mecanismo de transdução de sinal no interior da célula T, o *hinge* se aproxima do scFv, em seguida, a α -hélice do domínio transmembrana inclina, ocasionando uma variação do domínio intracelular, expondo os domínios ITAM do CD3- ζ às proteínas quinases, que fosforilam esse domínio e ativam as células T.

Com base na Análise de Componentes Principais (PCA) por tempo de simulação para o *hinge*, os domínios transmembrana e intracelular do CAR nos sistemas CAR e CAR-CD19, é possível comparar a movimentação estrutural de cada uma dessas estruturas, quando o CAR interage ou não com o CD19. Com relação ao *hinge*, pode-se inferir uma maior movimentação dessa estrutura no sistema CAR-CD19 (Figuras 40A e 46A), com PC1 e PC2 máximos em 200 ns, indicando uma mudança conformacional no *hinge*. Após essa mudança, há inclinação da α -hélice do domínio transmembrana que resulta em redução brusca nas PC1 e PC2 do domínio intracelular em 600 ns (Figura 46C), indicando uma modificação conformacional desse domínio. O que realça o papel do *hinge* na ativação e ação antitumoral das células CAR T (GUEST *et al.*, 2005; HUDECEK *et al.*, 2015). Esse comportamento não é observado no sistema CAR, indicando que, quando o CAR não interage com o CD19, a estrutura do domínio intracelular não movimenta consideravelmente (Figura 40C).

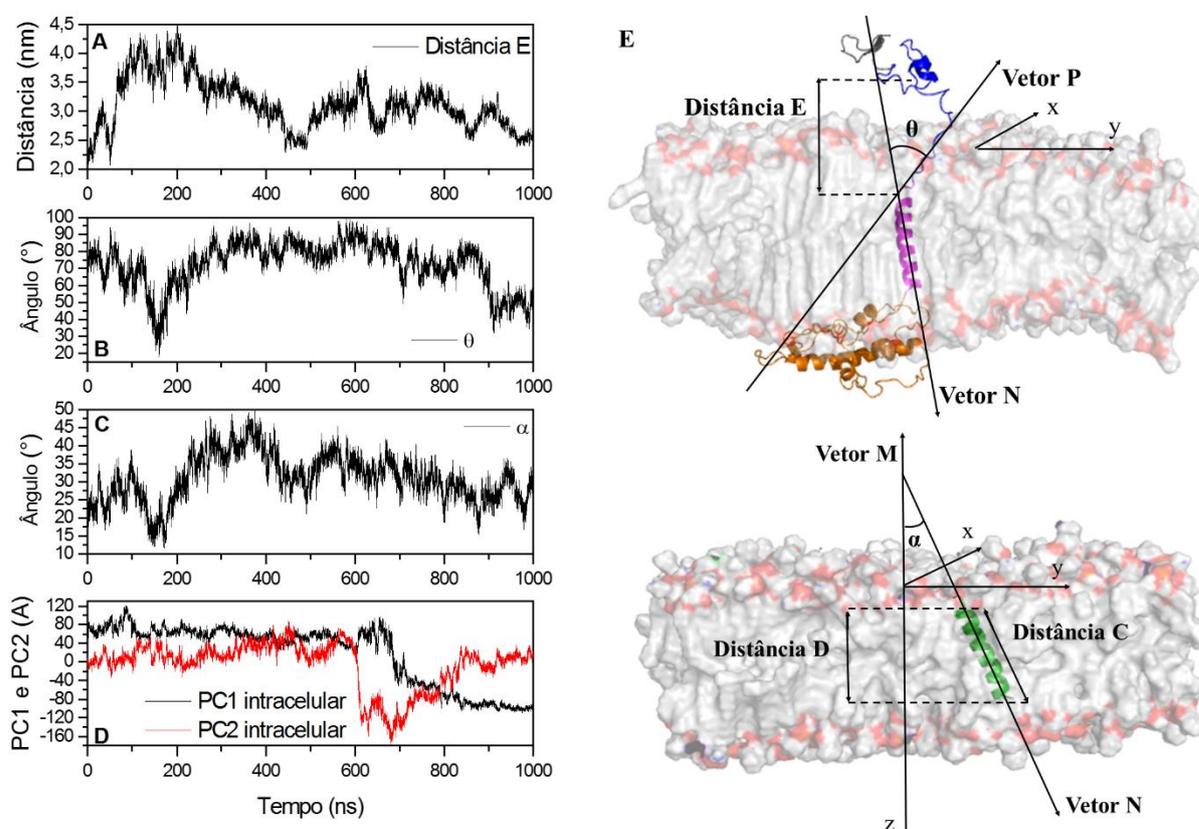
Pelas projeções PC1 e PC2 nos dois sistemas, pode-se inferir a movimentação do *hinge* e dos domínios transmembrana e intracelular do CAR. A movimentação do *hinge* é maior quando o CAR interage com o CD19 (Figuras 41A e 47A). Já para a estrutura do domínio

transmembrana, pode-se observar que ela se mantém nos dois sistemas (Figura 41B e 47B). A estrutura do domínio intracelular movimenta bem mais no sistema CAR-CD19, podendo assim inferir que a interação com o CD19 possibilita uma maior movimentação desse domínio, que possui mais conformações durante a DM (Figura 47C) comparado ao domínio transmembrana no sistema CAR (Figura 41C). Esse movimento descrito corrobora o que já é descrito na literatura em relação a disponibilização dos sítios do domínio intracelular, no sentido que mudanças conformacionais expõem os sítios de ativação de tirosina. Os domínios citoplásmicos da proteína CD3- ζ presente no domínio intracelular do CAR contém três motivos de ativação à base de tirosina do imunorreceptor (ITAM), região de proteínas de sinalização que são fosforiladas em resíduos de tirosina e tornam-se locais de ligação para outras moléculas de sinalização (VALLEJO *et al.*, 2004). A ativação da sinalização baseada em ITAM é capaz de dotar as células T com a capacidade de lisar as células tumorais alvo e liberar citocinas em resposta à ligação ao antígeno (BEZBRADICA; MEDZHITOV, 2012).

No sistema CAR-CD19, foi calculada a distância entre o centro de massa do *hinge* e o Ca do resíduo Trp560, localizado na extremidade extracelular da α -hélice do domínio transmembrana (Distância E), na superfície da membrana. Na Figura 49A, abaixo, é apresentada a curva de Distância E durante a DM. Observa-se que a distância máxima entre o *hinge* e o domínio transmembrana é de 4.5 nm em 200 ns, definindo que o *hinge* está com uma estrutura mais estirada. Do início a 200 ns, o ângulo θ , formado entre o vetor N e o vetor P (Figura 49E), definido pelas posições dos Ca dos resíduos Thr548 e Ile559 (resíduos da região menos flexível do *hinge* e os três primeiros resíduos do domínio transmembrana não pertencentes à α -hélice), atinge valor mínimo, de 20° (ver definição dos ângulos na Figura 49B), corroborando que essa região do *hinge* está estirada em relação à α -hélice. Isso ocorreu em decorrência de uma acomodação da estrutura do *hinge* na formação da interface scFv-*hinge*, como observado na redução do PII entre eles de -60 a -160 kcal.mol⁻¹ visto na Figura 43C. Após 200 ns, o ângulo θ aumenta, aproximando-se do máximo, 90°, em 370 ns, indicando que a região de ligação do *hinge*-transmembrana está mais frouxa, o que permite que a estrutura em α -hélice, bem conservada na DM, incline na bicamada, o que é ratificado pelo aumento do ângulo α , entre o vetor M e o vetor N (Figura 49C). Essa inclinação máxima é alcançada concomitante a maior frouxidão da região polipeptídica entre o *hinge* e o domínio transmembrana, em 370 ns e ângulo de 50°. Portanto, fica evidente que a inclinação da α -hélice ocorre após a formação de *hinge*-scFv e aproximação desse conjunto da membrana para reduzir a tensão na região de ligação *hinge*-transmembrana, o que possibilitou a inclinação da α -hélice. Essa inclinação perdurou de 370 ns a ~ 600 ns, momento em que foi observada mudança

conformacional no domínio intracelular (Figura 48A e B), ratificado pela transição rápida na curva de PC2 em 600 ns e PC1 em 680 ns, como mostrado na Figura 49D. Essa sequência de eventos descrita propõe um mecanismo de transdução de sinal na célula CAR T, sendo dependente desde a interação do CD19 com scFv, possibilitando formar uma interface com o *hinge* que tensiona e, posteriormente, afrouxa a região linear de ligação entre *hinge* e domínio transmembrana, possibilitando inclinar a α -hélice e promover a posterior mudança conformacional do domínio intracelular, que disponibiliza regiões para fosforilação e transmissão do sinal interno na célula CAR T.

Figura 49 – Mudanças conformacionais na região de ligação do *hinge* ao domínio transmembrana, α -hélice do domínio transmembrana e no domínio intracelular. A) Distância E entre o *hinge* e a α -hélice do domínio transmembrana. B) Ângulo θ , formado entre os vetores P e N. C) Ângulo α , formado entre os vetores M e N. D) PC1 e PC2 da estrutura do domínio intracelular por tempo de simulação. E) Representação dos vetores M, N e P, e dos ângulos α e θ .



Fonte: elaborada pelo autor.

A fosforilação de tirosinas nos motivos ITAM inicia a transdução de sinal e a ativação de tirosina quinases subsequentes, que fosforilam resíduos de tirosina em outras proteínas adaptadoras. Os passos subsequentes na transdução de sinal são gerados pelo recrutamento

específico de enzimas-chave que iniciam diferentes vias de sinalização cada uma. Portanto, a modificação estrutural do *hinge* em 200 ns, seguida da inclinação do domínio transmembrana em 370 ns observado nos resultados de DM possibilitaram mudanças conformacionais no domínio intracelular para expor os sítios e possibilitar a fosforilação, como mostrado na Figura 48. O reconhecimento de complexos MHC-peptídeo podem induzir uma alteração conformacional no TCR, fazendo com que os ITAMs associados aos CD3 ou cadeias ζ ligadas tornem-se disponíveis para a fosforilação da tirosina por quinases da família Src (ABBAS *et al.*, 2015). Da mesma forma, o reconhecimento do CD19 na célula alvo pode induzir alteração conformacional no CAR, tornando os domínios ITAM disponíveis para a fosforilação, com o mecanismo sugerido na DM. A fosforilação do CD3- ζ , presente no domínio intracelular, tem um papel central na ativação de células T (BENMEBAREK *et al.*, 2019). O sinal baseado em ITAM sozinho induz anergia da célula T, levando à secreção ineficiente de citocinas, à baixa proliferação e à morte celular induzida por ativação (AICD) das próprias células CAR T. Assim, sendo necessário o domínio coestimulador em adição ao sinal de CD3- ζ para a ativação completa da célula CAR T e prevenção de apoptose (GONG *et al.*, 1999).

O mecanismo de transdução de sinal da célula T modificada com CAR é semelhante ao mecanismo de ativação da célula T clássica pelo Receptor de células T (TCR). A proteína CD3- ζ presente no domínio intracelular do TCR e da célula CAR T possui papel na sinalização (CLEVERS *et al.*, 1988). Uma das vantagens da célula CAR T em relação à célula T convencional é que ela é capaz de fazer o reconhecimento do antígeno pelo scFv do CAR e, também, tem a capacidade de matar as células tumorais independentemente do complexo principal de histocompatibilidade (MHC) (CHMIELEWSKI *et al.*, 2013). Os linfócitos T citotóxicos convencionais (CTLs) matam e eliminam rapidamente a célula alvo com notável especificidade, devido à formação de uma sinapse imunológica (SI) distinta após o reconhecimento do TCR (MONKS *et al.*, 1998). Sabe-se que as células CAR T utilizam, pelo menos, em parte, o mecanismo convencional de sinalização do TCR (MAHER *et al.*, 2002). Assim, pode-se supor que a célula alvo da célula CAR T é comparável a da célula T clássica. Portanto, o reconhecimento da célula alvo pelo CAR não é afetado pelos principais mecanismos que fazem com que os tumores evitem o reconhecimento de células T restritas ao MHC, como o processamento defeituoso do antígeno (CHMIELEWSKI *et al.*, 2013). Além disso, o CAR combina a propriedade do anticorpo monoclonal de interagir com o antígeno presente na superfície das células tumorais, de forma específica, com a capacidade de lise celular e de auto renovação das células T (RAMOS; DOTTI, 2011).

O mecanismo de transdução de sinal proposto evidencia a importância dos componentes

do CAR para a sua correta atuação no organismo de pacientes com câncer, sendo sua composição e estrutura cruciais para a lise efetiva de células tumorais. A confirmação do mecanismo ainda necessita de mais estudos experimentais que corroborem os resultados encontrados neste trabalho. Até mesmo mais estudos teóricos realizando, por exemplo, triplicatas dessas simulações apresentadas aqui, que requerem um maior poder computacional e mais tempo para isso.

7 CONCLUSÃO

Os modelos do scFv, do *hinge*, do domínio transmembrana e intracelular foram utilizados na montagem do CAR e na montagem do complexo de interação CAR-CD19.

As estruturas do domínio intracelular e do *hinge* no CAR interagindo com o CD19 apresentam-se mais flexíveis em relação às mesmas estruturas do CAR que não interage com o CD19, assim a interação com o CD19 promove maiores mudanças conformacionais nos componentes do CAR.

As análises de distância, ângulo e PCA possibilitaram inferir um mecanismo de transdução de sinal no sistema CAR-CD19, por meio da variação estrutural do *hinge*, com posterior inclinação da α -hélice do domínio transmembrana, seguida de mudança conformacional do domínio intracelular, movimentos que não acontecem no sistema em que o CAR não interage com o CD19. A alteração nas conformações do domínio intracelular observadas na DM são essenciais para expor os sítios de fosforilação e corrobora a função do CD3- ζ que atua na sinalização das células T (CLEVERS *et al.*, 1988).

REFERÊNCIAS

ABBAS, ABUL K.; LICHTMAN, ANDREW H.; PILAI, S. **Imunologia Celular e Molecular**. 2012.

ABBAS, A. K.; LICHTMAN, A. H.; PILLAI, S. *Imunologia Celular e Molecular - Abbas 8ed-2*. Elsevier Editora Ltda, , n. 8, p. 631–667, 2015.

ABRAHAM, M.; HESS, B.; SPOEL, D. VAN DER; LINDAHL, E. **GROMACS USER MANUAL**. 2013.

ABRAHAM, M. J.; MURTOLA, T.; SCHULZ, R.; *et al.* Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. **SoftwareX**, v. 1–2, p. 19–25, 2015.

ACUTO, O.; MICHEL, F. CD28-mediated co-stimulation: A quantitative support for TCR signalling. **Nature Reviews Immunology**, v. 3, n. 12, p. 939–951. 2003.

AHMAD, Z. A.; YEAP, S. K.; ALI, A. M.; *et al.* scFv Antibody: Principles and Clinical Application. **Clinical and Developmental Immunology**, v. 2012, p. 1–15, 2012.

ALABANZA, L.; PEGUES, M.; GELDRES, C.; *et al.* Function of Novel Anti-CD19 Chimeric Antigen Receptors with Human Variable Regions Is Affected by *Hinge* and Transmembrane Domains. **Molecular Therapy**, v. 25, n. 11. p. 2452–2465, 2017. Elsevier Inc..

ALBERTS, B.; JOHNSON, A.; LEWIS, J.; *et al.* **Biologia Molecular da Célula 6º ed.** 2017.

ALLEN, M. P.; TILDESLEY, D. J. **Computer simulation of liquids**. 1987.

ALLEN, M. P.; TILDESLEY, D. J. **Computer simulation of liquids**. 2 ed. ed. Oxford University Press, 2017.

AMADEI, A.; LINSSSEN, A. B. M.; BERENDSEN, H. J. C. Essential dynamics of proteins. **Proteins: Structure, Function, and Genetics**, v. 17, n. 4, p. 412–425, 1993.

ANANDAKRISHNAN, R.; AGUILAR, B.; ONUFRIEV, A. V. H + + 3 . 0 : automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. **Nucleic acids research**, v. 40. p. 537–541. 2012.

ANNESLEY, C. E.; SUMMERS, C.; CEPPI, F.; GARDNER, R. A. The Evolution and Future of CAR T Cells for B-Cell Acute Lymphoblastic Leukemia. **Clinical Pharmacology & Therapeutics**, v. 00. n. 00. 2017.

ARFKEN, G. B.; WEBER, H. J. **Mathematical Methods for Physicists**. Orlando: FL.: Academic Press, 1985.

BEGENT, R. H. J.; VERHAAR, M. J.; CHESTER, K. A.; *et al.* Clinical evidence of efficient tumor targeting based on single-chain Fv antibody selected from a combinatorial library. **Nature Medicine**, v. 2, n. 9, p. 979–984, 1996.

- BENMEBAREK, M. R.; KARCHES, C. H.; CADILHA, B. L.; *et al.* Killing mechanisms of chimeric antigen receptor (CAR) T cells. **International Journal of Molecular Sciences**, v. 20. n. 6, p. 1–21. 2019.
- BERENDSEN, H J C; POSTMA, J. P. M.; GUNSTEREN, W. F. VAN; DINOLA, A.; HAAK, J. R. Molecular dynamics with coupling to an external bath. **The Journal of Chemical Physics**, v. 81. n. 8, p. 3684–3690. 1984.
- BERENDSEN, H. J. C.; POSTMA, J. P. M.; GUNSTEREN, W. F. VAN; DINOLA, A.; HAAK, J. R. Molecular dynamics with coupling to an external bath. **The Journal of Chemical Physics**, v. 81. n. 8, p. 3684–3690. 1984.
- BERENDSEN, H. J. C.; SPOEL, D. VAN DER; DRUNEN, R. VAN. GROMACS: A message-passing parallel molecular dynamics implementation. **Computer Physics Communications**, v. 91. n. 1–3, p. 43–56, 1995a.
- BERENDSEN, H. J. C.; SPOEL, D. VAN DER; DRUNEN, R. VAN. GROMACS: A message-passing parallel molecular dynamics implementation. **Computer Physics Communications**, v. 91. p. 43–56, 1995b.
- BERMAN, H. M.; WESTBROOK, J.; FENG, Z.; *et al.* The protein data bank. **Nucleic acids research**, v. 28, n. 1. p. 235–242, 2000.
- BEZBRADICA, J. S.; MEDZHITOV, R. Role of ITAM signaling module in signal integration. **Current Opinion in Immunology**, v. 24, p. 58–66, 2012.
- BORRIE, A. E.; MALEKI VAREKI, S. **T Lymphocyte–Based Cancer Immunotherapeutics**. Elsevier Ltd, 2018.
- BOUGNOUX, P.; SALEM, N.; LYONS, C.; HOFFMAN, T. Alteration in the membrane fatty acid composition of human lymphocytes and cultured transformed cells induced by interferon. **Molecular Immunology**, v. 22, n. 9, p. 1107–1113, 1985.
- BRENTJENS, R.; YEH, R.; BERNAL, Y.; RIVIERE, I.; SADELAIN, M. Treatment of chronic lymphocytic leukemia with genetically targeted autologous t cells: Case report of an unforeseen adverse event in a phase i clinical trial. **Molecular Therapy**, v. 18, n. 4, p. 666–668, 2010.
- BUKCZYNSKI, J.; WEN, T.; WATTS, T. H. Costimulation of human CD28- T cells by 4-1BB ligand. **European Journal of Immunology**, v. 33, p. 446–454, 2003.
- BUSSI, G.; DONADIO, D.; PARRINELLO, M. Canonical sampling through velocity rescaling. **Journal of Chemical Physics**, v. 126, n. 1. 2007.
- CHEN, V. B.; ARENDALL, W. B.; HEADD, J. J.; *et al.* MolProbity: All-atom structure validation for macromolecular crystallography. **Acta Crystallographica Section D: Biological Crystallography**, v. 66, n. 1. p. 12–21. 2010. International Union of Crystallography.
- CHMIELEWSKI, M.; HOMBACH, A. A.; ABKEN, H. Antigen-specific T-cell activation independently of the MHC: Chimeric antigen receptor-redirected T cells. **Frontiers in Immunology**, v. 4, n. NOV, p. 1–7, 2013.

- CHOTHIA, C. ; LESK, A. M. Canonical Structures for the Hypervariable Regions of Immunoglobulins. **Journal of Molecular Biology**, v. 196, p. 901–917, 1987.
- CLEVERS, H.; ALARCON, B.; WILEMAN, T.; TERHORST, C. The T cell receptor/CD3 complex: A dynamic protein ensemble. **Annual Review of Immunology**, v. 6, n. February, p. 629–662, 1988.
- DARDEN, T.; YORK, D.; PEDERSEN, L. Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems. **The Journal of Chemical Physics**, v. 98, n. 12, p. 10089, 1993.
- DAURA, X.; GADEMANN, K.; JAUN, B.; GUNSTEREN, W. F. VAN; MARK, A. E. Peptide Folding : When Simulation Meets Experiment. **Angew. Chem. Int. Ed**, v. 38, p. 236–240. 1999.
- DAVIS, I. W.; LEAVER-FAY, A.; CHEN, V. B.; *et al.* MolProbity: All-atom contacts and structure validation for proteins and nucleic acids. **Nucleic Acids Research**, v. 35, n. SUPPL.2, p. 375–383, 2007.
- DONG, G. Q.; FAN, H.; SCHNEIDMAN-DUHOVNY, D.; WEBB, B.; SALI, A. Optimized atomic statistical potentials: Assessment of protein interfaces and loops. **Bioinformatics**, v. 29, n. 24, p. 3158–3166, 2013.
- DU, X.; BEERS, R.; FITZGERALD, D. J.; PASTAN, I. Differential Cellular Internalization of Anti-CD19 and CD22 Immunotoxins Results in Different Cytotoxic Activity. **Cancer Res.**, v. 68, n. 15, p. 6300–6305, 2008.
- ELGERT, K. D. Antibody structure and function. **Immunology: Understanding the Immune System**. p.58–78, 1998.
- ESHAR, Z.; WAKS, T.; GROSS, G.; SCHINDLER, D. G. Specific activation and targeting of cytotoxic lymphocytes through chimeric single chains consisting of antibody-binding domains and the γ or ζ subunits of the immunoglobulin and T-cell receptors. **Proceedings of the National Academy of Sciences of the United States of America**, v. 90. n. 2, p. 720–724, 1993.
- EVANS, E. J.; ESNOUF, R. M.; MANSO-SANCHO, R.; *et al.* Crystal structure of a soluble CD28-Fab complex. , v. 6, n. 3, p. 271–279, 2005.
- FERREIRA, M. M. C.; ANTUNES, A. M.; MELGO, M. S.; VOLPE, P. L. O. Quimiometria i: Calibração multivariada, um tutorial. **Química Nova**, v. 22, n. 5, p. 724–731. 1999.
- FILHO, O. A. S.; ALENCASTRO, R. B. DE. Modelagem de proteínas por homologia. **Química Nova**, v. 26, n. 2, p. 253–259, 2003.
- FINNEY, H. M.; AKBAR, A. N.; LAWSON, A. D. G. Activation of Resting Human Primary T Cells with Chimeric Receptors: Costimulation from CD28, Inducible Costimulator, CD134, and CD137 in Series with Signals from the TCR ζ Chain. **The Journal of Immunology**, v. 172, n. 1. p. 104–113, 2004.
- FINNEY, H. M.; LAWSON, A. D.; BEBBINGTON, C. R.; WEIR, A. N. Chimeric receptors providing both primary and costimulatory signaling in T cells from a single gene product. **Journal of immunology (Baltimore, Md. : 1950)**, v. 161. n. 6, p. 2791–7, 1998.

FRENKEL, D.; SMIT, B. **Understanding Molecular Simulation - From Algorithms to Application**. 2 ed. ed. Academic Press, 2002.

GONG, M. C.; LATOUCHE, J.; KRAUSE, A.; HESTON, W. D. W.; BANDER, N. H. Cancer Patient T Cells Genetically Targeted to Prostate-Specific Membrane Antigen Specifically Lyse Prostate Cancer Cells and Release Cytokines in Response to Prostate-Specific Membrane Antigen. , v. 1. n. 2, p. 123–127, 1999.

GORDON, J. C.; MYERS, J. B.; FOLTA, T.; *et al.* H⁺⁺: A server for estimating pK_as and adding missing hydrogens to macromolecules. **Nucleic Acids Research**, v. 33, n. SUPPL. 2, p. 368–371. 2005.

GUEDAN, S.; CALDERON, H.; POSEY, A. D.; MAUS, M. V. Engineering and Design of Chimeric Antigen Receptors. **Molecular Therapy: Methods & Clinical Development**, v. 12, n. March, p. 145–156, 2019. Elsevier Ltd.

GUEST, R. D.; HAWKINS, R. E.; KIRILLOVA, N.; *et al.* The Role of Extracellular Spacer Regions in the Optimal Design of Chimeric Immune Receptors. **Journal of Immunotherapy**, v. 28, n. 3, p. 203–211. 2005.

GUNSTEREN, W. F. VAN; DOLENC, J.; MARK, A. E. Molecular simulation as an aid to experimentalists. **Current Opinion in Structural Biology**, v. 18, n. 2, p. 149–153, 2008.

HAAS, K. M.; TEDDER, T. F. Role of the CD19 and CD21/35 receptor complex in innate immunity, host defense and autoimmunity. **Advances in Experimental Medicine and Biology**, v. 560. p. 125–139, 2005.

HANSSON, T.; OOSTENBRINK, C.; GUNSTEREN, W. F. VAN. Molecular dynamics simulations. **Curr. Opin. Struct. Biol.**, v. 12, p. 190. 2002.

HESS, B. P-LINCS: A parallel linear constraint solver for molecular simulation. **Journal of Chemical Theory and Computation**, v. 4, n. 1. p. 116–122, 2008.

HESS, B.; BEKKER, H.; BERENDSEN, H. J. C.; FRAAIJE, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. **Journal of Computational Chemistry**, v. 18, n. 12, p. 1463–1472, 1997.

HESS, B.; KUTZNER, C.; SPOEL, D. VAN DER; LINDAHL, E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. **Journal of Chemical Theory and Computation**, v. 4, n. 3, p. 435–447, 2008.

HESTENES, M. R.; STIEFEL, E. Methods of conjugate gradients for solving linear systems. **Journal of Research of the National Bureau of Standards**, v. 49, n. 6, p. 409–436, 1952.

HOLLIGER, P.; HUDSON, P. J. Engineered antibody fragments and the rise of single domains. **Nature biotechnology**, v. 23, n. 9, p. 1126–1136, 2005.

HOOVER, W. G. Canonical Dynamics: Equilibrium phase-space distributions. **Physical Review A**, v. 31. n. 3, p. 1695–1697, 1985.

HUANG, J.; MACKERELL, A. D. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. **J Comput Chem.**, v. 34, n. 25, p. 2135–2145,

2013.

HUANG, K. **Statistical Mechanics, 2nd Edition**. 2nd ed. John Wiley & Sons, 1987.

HUDECEK, M.; KOSASIH, P. L.; JENSEN, M. C.; RADER, C.; RIDDELL, S. R. Receptor affinity and extracellular domain modifications affect tumor recognition by ROR1-specific chimeric antigen receptor T- cells. **Clin Cancer Res.**, v. 19, n. 12, p. 3153–3164, 2013.

HUDECEK, M.; SOMMERMEYER, D.; KOSASIH, P. L.; *et al.* The non-signaling extracellular spacer domain of chimeric antigen receptors is decisive for in vivo antitumor activity. **Cancer Immunol Res.**, v. 3, n. 2, p. 125–135, 2015.

IMAI, C.; MIHARA, K.; ANDREANSKY, M.; *et al.* Chimeric receptors with 4-1BB signaling capacity provoke potent cytotoxicity against acute lymphoblastic leukemia. **Leukemia**, v. 18, n. 4, p. 676–684, 2004.

INCA, 2018. **Estimativa 2018-Incidência de câncer no Brasil**. 2018.

JO, S.; LIM, J. B.; KLAUDA, J. B.; IM, W. CHARMM-GUI membrane builder for mixed bilayers and its application to yeast membranes. **Biophysical Journal**, v. 97, n. 1. p. 50–58, 2009. Biophysical Society.

JORGENSEN, W. L.; CHANDRASEKHAR, J.; MADURA, J. D.; IMPEY, R. W.; KLEIN, M. L. Comparison of simple potential functions for simulating liquid water. **The Journal of Chemical Physics**, v. 79, n. 2, p. 926–935, 1983.

KANDA, H.; WATANABE, T. Engineered antibodies. **Biotherapy**, v. 10. n. 2, p. 106–113, 1996.

KANTARJIAN, H.; STEIN, A.; GÖKBUGET, N.; *et al.* Blinatumomab versus Chemotherapy for Advanced Acute Lymphoblastic Leukemia. **N Engl J Med.**, v. 376, n. 9, p. 836–847, 2017.

KARPLUS, M.; MCCAMMON, J. A. Molecular dynamics simulations of biomolecules. **Nature Structural Biology**, v. 9, n. 9, p. 646–652, 2002.

KOCHENDERFER, J. N.; FELDMAN, S. A.; ZHAO, Y.; *et al.* Construction and Pre-clinical Evaluation of an Anti-CD19 Chimeric Antigen Receptor. **J Immunother**, v. 32, n. 7, p. 689–702, 2009.

KOCHENDERFER, J. N.; WILSON, W. H.; JANIK, J. E.; *et al.* Eradication of B-lineage cells and regression of lymphoma in a patient treated with autologous T cells genetically engineered to recognize Brief report Eradication of B-lineage cells and regression of lymphoma in a patient treated with autologous T cells. , v. 116, n. 20. p. 4099–4102, 2010.

KULEMZIN, S. V.; KUZNETSOVA, V. V.; MAMONKIN, M.; TARANIN, A. V.; GORCHAKOV, A. A. Engineering chimeric antigen receptors. **Acta Naturae**, v. 9, n. 1. p. 6–14, 2017.

KUWANA, Y.; ASAKURA, Y.; UTSUNOMIYA, N.; *et al.* Expression of chimeric receptor composed of immunoglobulin-derived V regions and T-cell receptor-derived C regions. **BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS**, v. 149, n. 3, p. 960–968, 1987.

- LASKOWSKI, R. A.; MACARTHUR, M. W.; MOSS, D. S.; THORNTON, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. **Journal of Applied Crystallography**, v. 26, n. 2, p. 283–291, 1993. International Union of Crystallography.
- LEAHY, D. J.; AXEL, R.; HENDRICKSON, W. A. Crystal Structure of a Soluble Form of the Human T Cell Coreceptor CD8 at 2.6 Å Resolution. **Cell**, v. 68, p. 1145–1162, 1992.
- LEE, J.; CHENG, X.; SWAILS, J. M.; *et al.* CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. **Journal of Chemical Theory and Computation**, v. 12, n. 1, p. 405–413, 2016.
- LEE, K. J.; CHOW, V.; WEISSMAN, A.; *et al.* Clinical use of blinatumomab for B-cell acute lymphoblastic leukemia in adults. **Therapeutics and Clinical Risk Management**, v. 12, p. 1301–1310, 2016.
- LIEBOWITZ, D. N.; LEE, K. R.; JUNE, C. H. Costimulatory approaches to adoptive immunotherapy. **Current Opinion in Oncology**, 1998.
- LINDAHL, E.; HESS, B. GROMACS 3.0: a package for molecular simulation and trajectory analysis. **J Mol Model**, v. 7, n. 1, p. 306–317, 2001.
- LIU, J.; ZHONG, J. F.; ZHANG, X.; ZHANG, C. Allogeneic CD19-CAR-T cell infusion after allogeneic hematopoietic stem cell transplantation in B cell malignancies. **Journal of Hematology & Oncology**, v. 10, n. 1, p. 35, 2017. *Journal of Hematology & Oncology*.
- LONG, A. H.; HASO, W. M.; SHERN, J. F.; *et al.* 4-1BB costimulation ameliorates T cell exhaustion induced by tonic signaling of chimeric antigen receptors. **Nature Medicine**, v. 21, n. 6, p. 581–590, 2015. Nature Publishing Group.
- MACKERELL, A. D.; BASHFORD, D.; BELLOTT, M.; *et al.* All-atom empirical potential for molecular modeling and dynamics studies of proteins. **Journal of Physical Chemistry B**, v. 102, p. 3586–3616, 1998.
- MAHER, J.; BRENTJENS, R. J.; GUNSET, G.; RIVIÈRE, I.; SADELAIN, M. Human T-lymphocyte cytotoxicity and proliferation directed by a single chimeric TCR ζ /CD28 receptor. **Nature Biotechnology**, v. 20, n. 1, p. 70–75, 2002.
- MAKITA, S.; YOSHIMURA, K.; TOBINAI, K. Clinical development of anti-CD19 chimeric antigen receptor T-cell therapy for B-cell non-Hodgkin lymphoma. **Cancer Science**, v. 108, n. 6, p. 1109–1118, 2017.
- MEANY, H. J.; SEIBEL, N. L.; KRAILO, M.; *et al.* Feasibility study of a novel Experimental induction protocol combining B43-PAP (anti-CD19) immunotoxin with standard induction chemotherapy in children and adolescents with relapsed B-lineage ALL: A report from the children's oncology group. **Journal of Immunotherapy**, v. 38, n. 7, p. 299–305, 2015.
- MEYER, T.; FERRER-COSTA, C.; PÉREZ, A.; *et al.* Essential dynamics: A tool for efficient trajectory compression and management. **Journal of Chemical Theory and Computation**, v. 2, n. 2, p. 251–258, 2006.

- MICHAEL C. JENSEN; LESLIE POPPLEWELL; COOPER, L. J.; *et al.* Anti-Transgene Rejection Responses Contribute to Attenuated Persistence of Adoptively Transferred CD20/CD19-Specific Chimeric Antigen Receptor Re-directed T Cells in Humans. **Biol Blood Marrow Transplant**, v. 16, n. 9, p. 1245–1256, 2010..
- MILONE, M. C.; FISH, J. D.; CARPENITO, C.; *et al.* Chimeric receptors containing CD137 signal transduction domains mediate enhanced survival of T cells and increased antileukemic efficacy in vivo. **Molecular Therapy**, v. 17, n. 8, p. 1453–1464, 2009. Nature Publishing Group.
- MOHAMMADI, M.; NEJATOLLAHI, F. . 3D STRUCTURAL MODELING OF NEUTRALIZING SCFV AGAINST GLYCOPROTEIN-D OF HSV-1 AND EVALUATION OF ANTIGEN-ANTIBODY INTERACTIONS BY BIOINFORMATIC METHODS. **International Journal of Pharma and Bio Sciences**, v. 5, n. 4, p. 835–847, 2014.
- MONKS, C. R. F.; FREIBERG, B. A.; KUPFER, H.; SCIAKY, N.; KUPFER, A. Three-dimensional segregation of supramolecular activation *clusters* in T cells. **Nature**, v. 395, p. 82–86, 1998.
- MOREA, V.; TRAMONTANO, A.; RUSTICI, M.; CHOTHIA, C.; LESK, A. M. Conformations of the third hypervariable region in the VH domain of immunoglobulins. **Journal of Molecular Biology**, v. 275, n. 2, p. 269–294, 1998.
- MORIN, S. O.; GIROUX, V.; FAVRE, C.; *et al.* In the absence of its cytosolic domain, the CD28 molecule still contributes to T cell activation. **Cellular and Molecular Life Sciences**, v. 72, n. 14, p. 2739–2748, 2015.
- MORSE, P. P.; FESHBACH, H. **Methods of Theoretical Physics**. New York: McGrawHill, 1953.
- MYERS, J.; GROTHAUS, G.; NARAYANAN, S.; ONUFRIEV, A. A Simple *Clustering* Algorithm Can Be Accurate Enough for Use in Calculations of pKs in Macromolecules. **PROTEINS: Structure, Function, and Bioinformatics**, v. 63, p. 928–938, 2006.
- NELSON, A. L.; REICHERT, J. M. Development trends for therapeutic antibody fragments. **Nature Biotechnology**, v. 27, n. 4, p. 331–337, 2009.
- NG, H. W.; LAUGHTON, C. A.; DOUGHTY, S. W. Molecular dynamics simulations of the adenosine A2a receptor: Structural stability, sampling, and convergence. **Journal of Chemical Information and Modeling**, v. 53, n. 5, p. 1168–1178, 2013.
- PALL, S.; ABRAHAM, M.; KUTZNER, C.; HESS, B.; LINDAHL, E. Tackling Exascale Software Challenges in Molecular Dynamics Simulations with GROMACS. **LNCS**, v. 8759, p. 3–27, 2015.
- PAPALEO, E.; MEREGHETTI, P.; FANTUCCI, P.; GRANDORI, R.; GIOIA, L. DE. Free-energy landscape, principal component analysis, and structural *clustering* to identify representative conformations from molecular dynamics simulations: The myoglobin case. **Journal of Molecular Graphics and Modelling**, v. 27, n. 8, p. 889–899, 2009.
- PARK, T. S.; ROSENBERG, S. A.; MORGAN, R. A. Treating Cancer with Genetically Engineered T Cells. **Trends Biotechnol.**, v. 29, n. 11, p. 550–557, 2011.

PARRINELLO, M.; RAHMAN, A. Crystal Structure and Pair Potentials: A Molecular-Dynamics Study. **The American Physical Society**, v. 45, n. 14, p. 1196–1199, 1980.

PARRINELLO, M.; RAHMAN, A. Polymorphic transitions in single crystals: A new molecular dynamics method. **Journal of Applied Physics**, v. 52, n. 12, p. 7182–7190, 1981.

PRONK, S.; PÁLL, S.; SCHULZ, R.; *et al.* GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. **Bioinformatics**, v. 29, n. 7, p. 845–854, 2013.

QIAN, L.; LI, D.; MA, L.; *et al.* The novel anti-CD19 chimeric antigen receptors with humanized scFv (single-chain variable fragment) trigger leukemia cell killing. **Cellular Immunology**, v. 304–305, p. 49–54, 2016. Elsevier Inc.

RAMACHANDRAN, G. N.; RAMAKRISHNAN, C.; SASISEKHARAN, V. Stereochemistry of polypeptide chain configurations. **Journal of Molecular Biology**, v. 7, n. 1, p. 95–99, 1963. Academic Press Inc. (London) Ltd.

RAMELLO, M. C.; BENZAÏD, I.; KUENZI, B. M.; *et al.* An immunoproteomic approach to characterize the CAR interactome and signalosome. **Science Signaling**, v. 12, n. 568, p. 1–16, 2019.

RAMOS, C. A.; DOTTI, G. Chimeric antigen receptor (CAR)-engineered lymphocytes for cancer therapy. **Expert Opinion on Biological Therapy**, v. 11, n. 7, p. 855–873, 2011.

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

ROY, J.; LAUGHTON, C. A. Long-timescale molecular-dynamics simulations of the major urinary protein provide atomistic interpretations of the unusual thermodynamics of ligand binding. **Biophysical Journal**, v. 99, n. 1, p. 218–226, 2010. Biophysical Society.

RYCKAERT, J. P.; CICCOTTI, G.; BERENDSEN, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. **Journal of Computational Physics**, v. 23, n. 3, p. 327–341, 1977.

ŠALI, A.; BLUNDELL, T. L. Comparative protein modelling by satisfaction of spatial restraints. **Journal of Molecular Biology**, v. 234, p. 779–815, 1993.

SALMIKANGAS, P.; KINSELLA, N.; CHAMBERLAIN, P. Chimeric Antigen Receptor T-Cells (CAR T-Cells) for Cancer Immunotherapy – Moving Target for Industry? **Pharmaceutical Research**, v. 35, n. 152, p. 1–8, 2018. Pharmaceutical Research.

SALOMON, B.; BLUESTONE, J. A. COMPLEXITIES OF CD28/B7: CTLA-4 COSTIMULATORY PATHWAYS IN AUTOIMMUNITY AND TRANSPLANTATION. **Annual Review of Immunology**, v. 19, n. 1, p. 225–252, 2001.

SELA-CULANG, I.; KUNIK, V.; OFRAN, Y. The structural basis of antibody-antigen recognition. **Frontiers in Immunology**, v. 4, n. OCT, p. 1–13, 2013.

SHEN, M.-Y.; SALI, A. Statistical potential for assessment and prediction of protein structures. **Protein Science**, p. 2507–2524, 2006.

- SHIRASU, N.; KUROKI, M. Functional design of chimeric T-cell antigen receptors for adoptive immunotherapy of cancer: Architecture and outcomes. **Anticancer Research**, v. 32, n. 6, p. 2377–2383, 2012.
- SHKURTI, A.; GONI, R.; ANDRIO, P.; *et al.* pyPcazip: A PCA-based toolkit for compression and analysis of molecular simulation data. **SoftwareX**, v. 5, p. 44–50. 2015. Elsevier B.V.
- SHUFORD, W. W.; KLUSSMAN, K.; TRITCHLER, D. D.; *et al.* 4-1BB costimulatory signals preferentially induce CD8+ T cell proliferation and lead to the amplification in vivo of cytotoxic T cell responses. **Journal of Experimental Medicine**, v. 186, n. 1. p. 47–55, 1997.
- SILVA, V. B. DA; SILVA, C. H. T. . Modelagem Molecular De Proteínas-Alvo Por Homologia Estrutural. **Revista Eletrônica de Farmácia**, 2007.
- SMITH, T. F.; WATERMAN, M. S. Identification of Common Molecular Subsequences. **Journal of Molecular Biology**, v. 147, p. 195–197, 1981.
- SOMMERMEYER, D.; HILL, T.; SHAMAH, S. M.; *et al.* Fully human CD19-specific chimeric antigen receptors for T-cell therapy. **Leukemia**, v. 31. n. 10. p. 2191–2199, 2017.
- SPOEL, D. VAN DER; LINDAHL, E.; HESS, B.; *et al.* GROMACS: Fast, flexible, and free. **Journal of Computational Chemistry**, v. 26, n. 16, p. 1701–1718, 2005.
- TASIAN, S. K.; GARDNER, R. A. CD19-redirected chimeric antigen receptor-modified T cells: A promising immunotherapy for children and adults with B-cell acute lymphoblastic leukemia (ALL). **Therapeutic Advances in Hematology**, v. 6, n. 5, p. 228–241. 2015.
- TAYLOR, P.; AHLSTRÖM, P.; JÖNSSON, B.; WALLQVIST, A.; ENGSTRÖM, S. A molecular dynamics method for simulations in the canonical ensemble. **Molecular Physics**, v. 52, n. 2, p. 255–268, 1984.
- TAYLOR, P.; NOSÉ, S.; KLEIN, M. L. Molecular Physics : An International Journal at the Interface Between Chemistry and Physics Constant pressure molecular dynamics for molecular systems. , , n. December, p. 37–41. 2013.
- TEDDER, T. F. CD19: A promising B cell target for rheumatoid arthritis. **Nature Reviews Rheumatology**, v. 5, n. 10. p. 572–577, 2009. Nature Publishing Group.
- TEPLYAKOV, A.; OBMOLOVA, G.; LUO, J.; GILLILAND, G. L. Crystal structure of B-cell co-receptor CD19 in complex with antibody B43 reveals an unexpected fold. **Proteins**, v. 86, n. 5, p. 495–500. 2018a.
- TEPLYAKOV, A.; OBMOLOVA, G.; LUO, J.; GILLILAND, G. L. Crystal structure of B-cell co-receptor CD19 in complex with antibody B43 reveals an unexpected fold. **Proteins**, v. 86, p. 495–500. 2018b.
- TOMULEASA, C.; FUJI, S.; BERCE, C.; *et al.* Chimeric antigen receptor T-cells for the treatment of B-cell acute lymphoblastic leukemia. **Frontiers in Immunology**, v. 9, n. FEB, p. 1–14, 2018.
- UCKUN, F.; JASZCZ, W.; AMBRUS, J.; *et al.* Detailed studies on expression and function of CD19 surface determinant by using B43 monoclonal antibody and the clinical potential of

anti-CD19 immunotoxins. **Blood**, v. 71. n. 1. p. 13–29, 1988.

VALLEJO, A. N.; DAVILA, E.; WEYAND, C. M.; GORONZY, J. J. Biology of T lymphocytes. **Rheumatic Disease Clinics of North America**, v. 30. n. 1. p. 135–157, 2004.

VERLET, L. Computer “Experiments” on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. **Physical Review**, v. 159, n. 1. p. 98–103, 1967.

WALLISCH, P.; LUSIGNAN, M. E.; BENAYOUN, M. D.; *et al.* Principal Components Analysis. In: Elsevier (Ed.); **MATLAB for Neuroscientists An Introduction to Scientific Computing in MATLAB**. 2nd ed., p.305–315, 2014.

WANG, K.; WEI, G.; LIU, D. CD19 : a biomarker for B cell development , lymphoma diagnosis and therapy. , v. 1. n. 1. p. 1. 2012.

WANG, T.; DUAN, Y. Probing the stability-limiting regions of an antibody single-chain variable fragment: A molecular dynamics simulation study. **Protein Engineering, Design and Selection**, v. 24, n. 9, p. 649–657, 2011.

WATERHOUSE, A.; BERTONI, M.; BIENERT, S.; *et al.* SWISS-MODEL: Homology modelling of protein structures and complexes. **Nucleic Acids Research**, v. 46, n. Web Server, p. 296–303, 2018. Oxford University Press.

WEBB, B.; SALI, A. Comparative Protein Structure Modeling Using MODELLER. **Current Protocols in Bioinformatics**, 2016.

WEI, G.; WANG, J.; HUANG, H.; ZHAO, Y. Novel immunotherapies for adult patients with B-lineage acute lymphoblastic leukemia. **Journal of Hematology and Oncology**, v. 10. n. 1. p. 1–13, 2017. Journal of Hematology & Oncology.

WLODEK, S. T.; CLARK, T. W.; SCOTT, L. R.; MCCAMMON, J. A. Molecular dynamics of acetylcholinesterase dimer complexed with tacrine. **Journal of the American Chemical Society**, v. 119, n. 40. p. 9513–9522, 1997.

WU, E. L.; CHENG, X.; JO, S.; *et al.* CHARMM-GUI Membrane Builder Toward Realistic Biological Membrane Simulations. **J Comput Chem.**, v. 35, n. 27, p. 1997–2004, 2014.

YANG, J.; YAN, R.; ROY, A.; *et al.* The I-TASSER Suite: protein structure and function prediction. **Nature Methods**, v. 12, n. 1. p. 1–5, 2015.

YOKOTA, T.; MILENIC, D. E.; WHITLOW, M.; SCHLOM, J. Rapid Tumor Penetration of a Single-Chain Fv and Comparison with Other Immunoglobulin Forms. **Cancer Research**, v. 52, n. 12, p. 3402–3408, 1992.

ZHANG, Y. I-TASSER server for protein 3D structure prediction. **BMC Bioinformatics**, v. 9, n. 40. p. 1–8, 2008.

ZHENG, P. P.; KROS, J. M.; LI, J. Approved CAR T cell therapies: ice bucket challenges on glaring safety risks and long-term impacts. **Drug Discovery Today**, v. 23, n. 6, p. 1175–1182, 2018. Elsevier Ltd.

ANEXO A- ANÁLISE DA QUALIDADE GEOMÉTRICA DOS MODELOS GERADOS NO PROGRAMA MOLPROBITY

a) scFv

All-Atom Contacts	Clashscore, all atoms:	0		100 th percentile* (N=1784, all resolutions)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.			
Protein Geometry	Poor rotamers	9	4.46%	Goal: <0.3%
	Favored rotamers	179	88.61%	Goal: >98%
	Ramachandran outliers	4	1.60%	Goal: <0.05%
	Ramachandran favored	234	93.60%	Goal: >98%
	MolProbity score [^]	1.41		97 th percentile* (N=27675, 0Å - 99Å)
	Cβ deviations >0.25Å	2	0.89%	Goal: 0
	Bad bonds:	0 / 1967	0.00%	Goal: 0%
	Bad angles:	19 / 2672	0.71%	Goal: <0.1%
Peptide Omegas	Cis Prolines:	2 / 10	20.00%	Expected: ≤1 per chain, or ≤5%
	Cis nonProlines:	1 / 241	0.41%	Goal: <0.05%
	Twisted Peptides:	3 / 251	1.20%	Goal: 0

b) Hinge

All-Atom Contacts	Clashscore, all atoms:	0		100 th percentile* (N=1784, all resolutions)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.			
Protein Geometry	Poor rotamers	0	0.00%	Goal: <0.3%
	Favored rotamers	31	93.94%	Goal: >98%
	Ramachandran outliers	1	2.33%	Goal: <0.05%
	Ramachandran favored	41	95.35%	Goal: >98%
	MolProbity score [^]	0.82		100 th percentile* (N=27675, 0Å - 99Å)
	Cβ deviations >0.25Å	0	0.00%	Goal: 0
	Bad bonds:	0 / 330	0.00%	Goal: 0%
	Bad angles:	6 / 457	1.31%	Goal: <0.1%
Peptide Omegas	Cis Prolines:	0 / 9	0.00%	Expected: ≤1 per chain, or ≤5%

c) Domínio transmembrana

All-Atom Contacts	Clashscore, all atoms:	0		100 th percentile* (N=1784, all resolutions)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.			
Protein Geometry	Poor rotamers	0	0.00%	Goal: <0.3%
	Favored rotamers	16	94.12%	Goal: >98%
	Ramachandran outliers	0	0.00%	Goal: <0.05%
	Ramachandran favored	22	100.00%	Goal: >98%
	MolProbity score [^]	0.50		100 th percentile* (N=27675, 0Å - 99Å)
	Cβ deviations >0.25Å	0	0.00%	Goal: 0
	Bad bonds:	0 / 186	0.00%	Goal: 0%
	Bad angles:	2 / 257	0.78%	Goal: <0.1%
Peptide Omegas	Cis Prolines:	0 / 1	0.00%	Expected: ≤1 per chain, or ≤5%

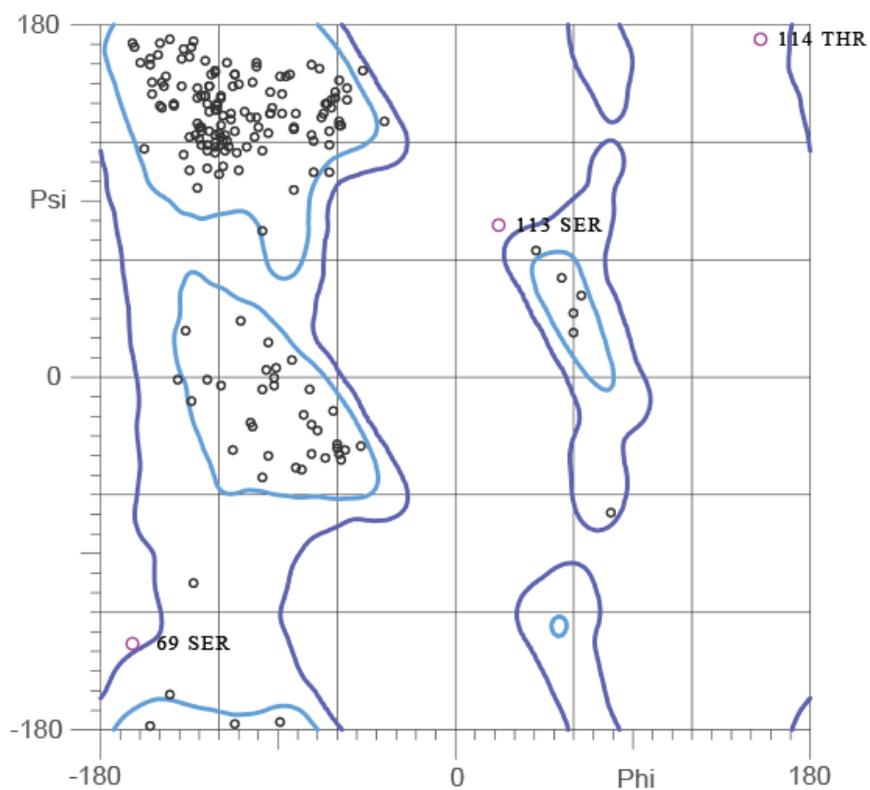
d) Domínio intracelular

All-Atom Contacts	Clashscore, all atoms:	0		100 th percentile* (N=1784, all resolutions)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.			
Protein Geometry	Poor rotamers	4	3.12%	Goal: <0.3%
	Favored rotamers	114	89.06%	Goal: >98%
	Ramachandran outliers	7	4.61%	Goal: <0.05%
	Ramachandran favored	126	82.89%	Goal: >98%
	MolProbity score [^]	1.57		93 rd percentile* (N=27675, 0Å - 99Å)
	Cβ deviations >0.25Å	18	13.04%	Goal: 0
	Bad bonds:	23 / 1273	1.81%	Goal: 0%
	Bad angles:	76 / 1699	4.47%	Goal: <0.1%
Peptide Omegas	Cis Prolines:	2 / 9	22.22%	Expected: ≤1 per chain, or ≤5%
	Twisted Peptides:	1 / 153	0.65%	Goal: 0

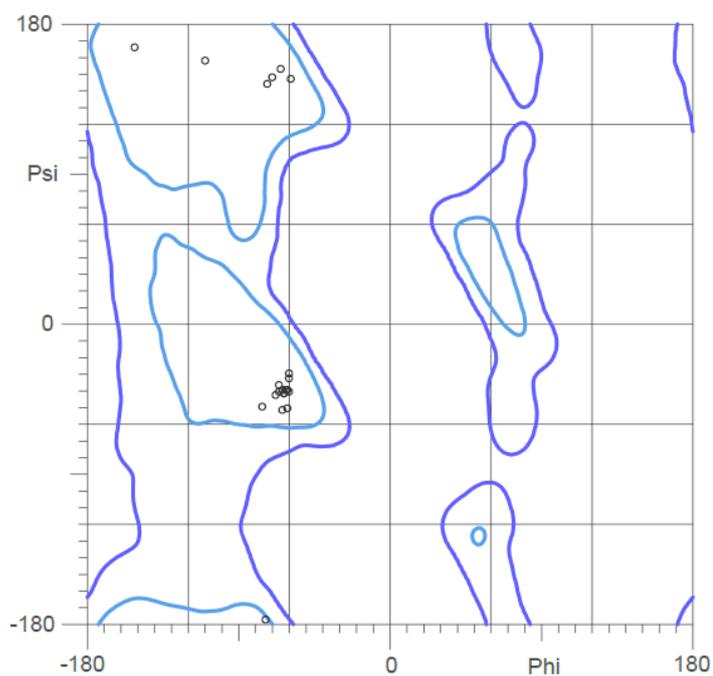
Fonte: MolProbity.

**ANEXO B- GRÁFICO RAMACHANDRAN GERADO PELO PROGRAMA
MOLPROBITY PARA CADA MODELO OBTIDO**

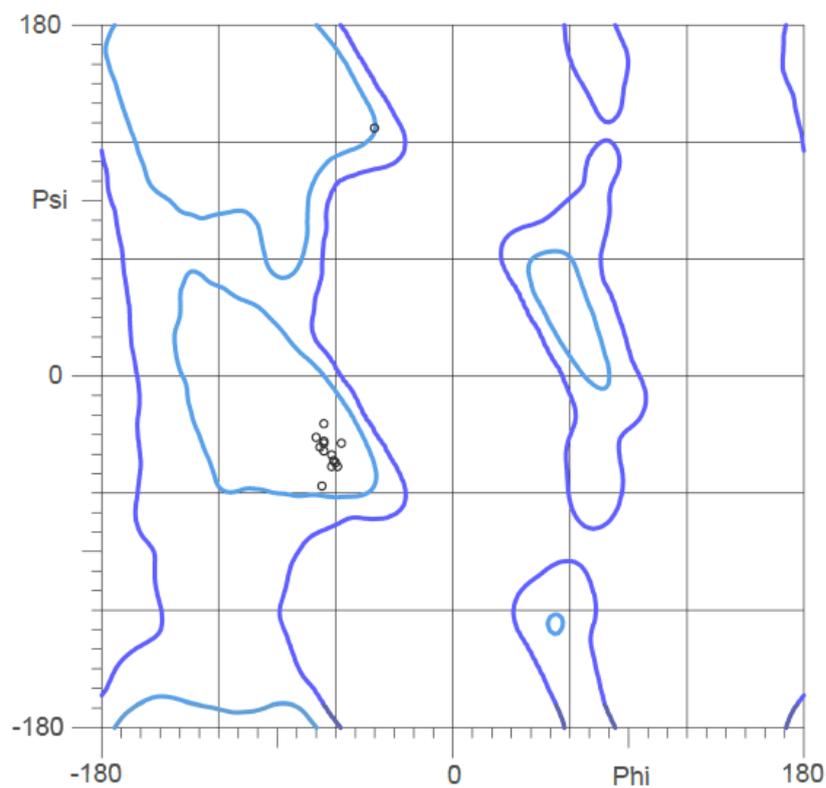
a) scFv



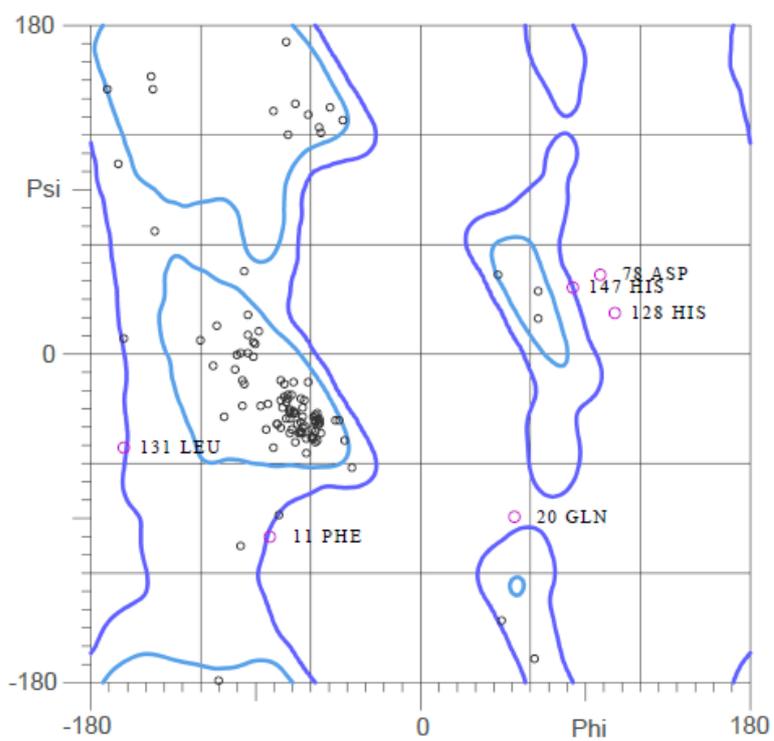
b) Hinge



c) Domínio transmembrana



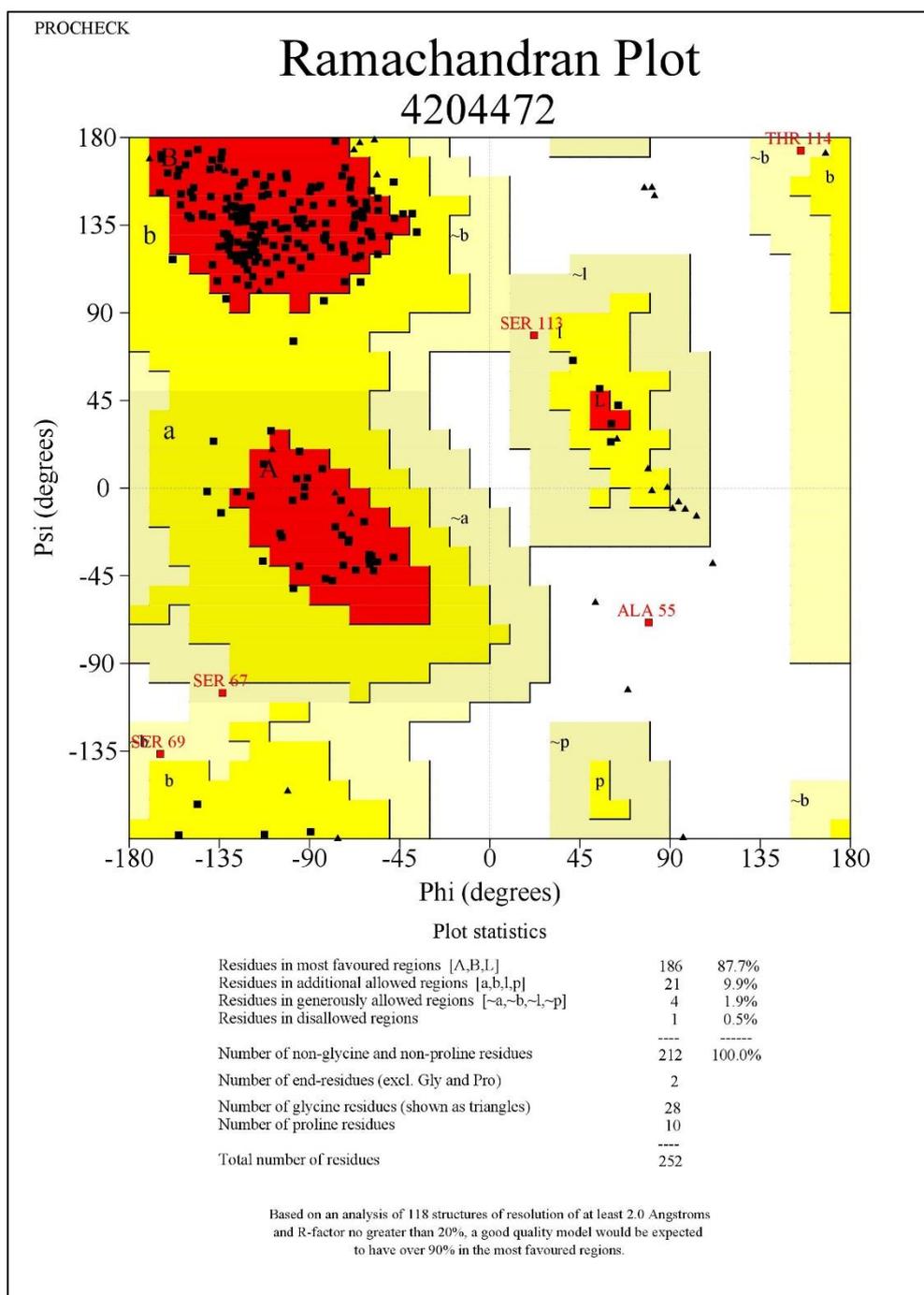
d) Domínio intracelular



Fonte: MolProbity.

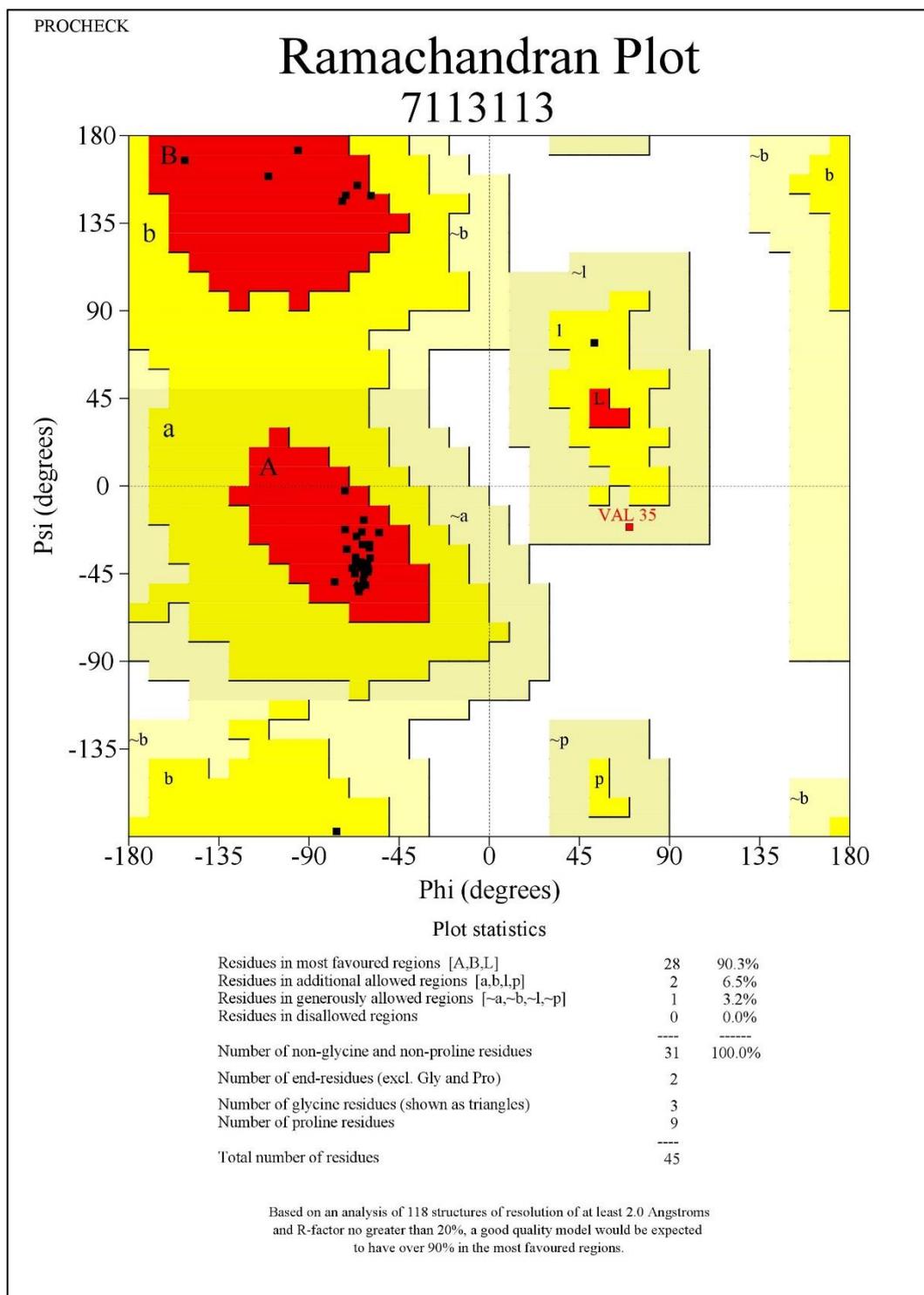
**ANEXO C - GRÁFICO RAMACHANDRAN GERADO PELO PROGRAMA
PROCHECK PARA CADA MODELO OBTIDO**

a) scFv

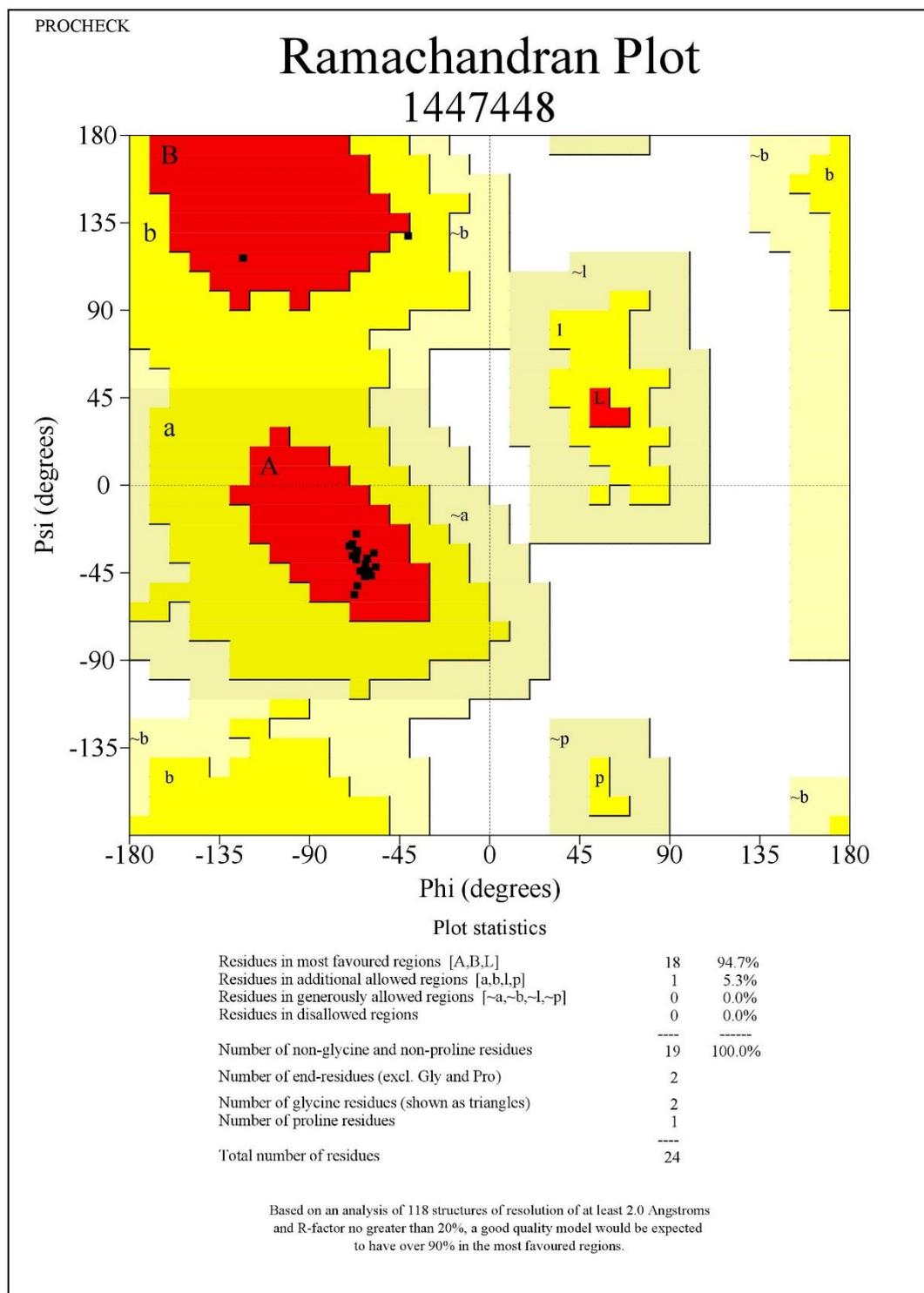


4204472_01.ps

b) Hinge

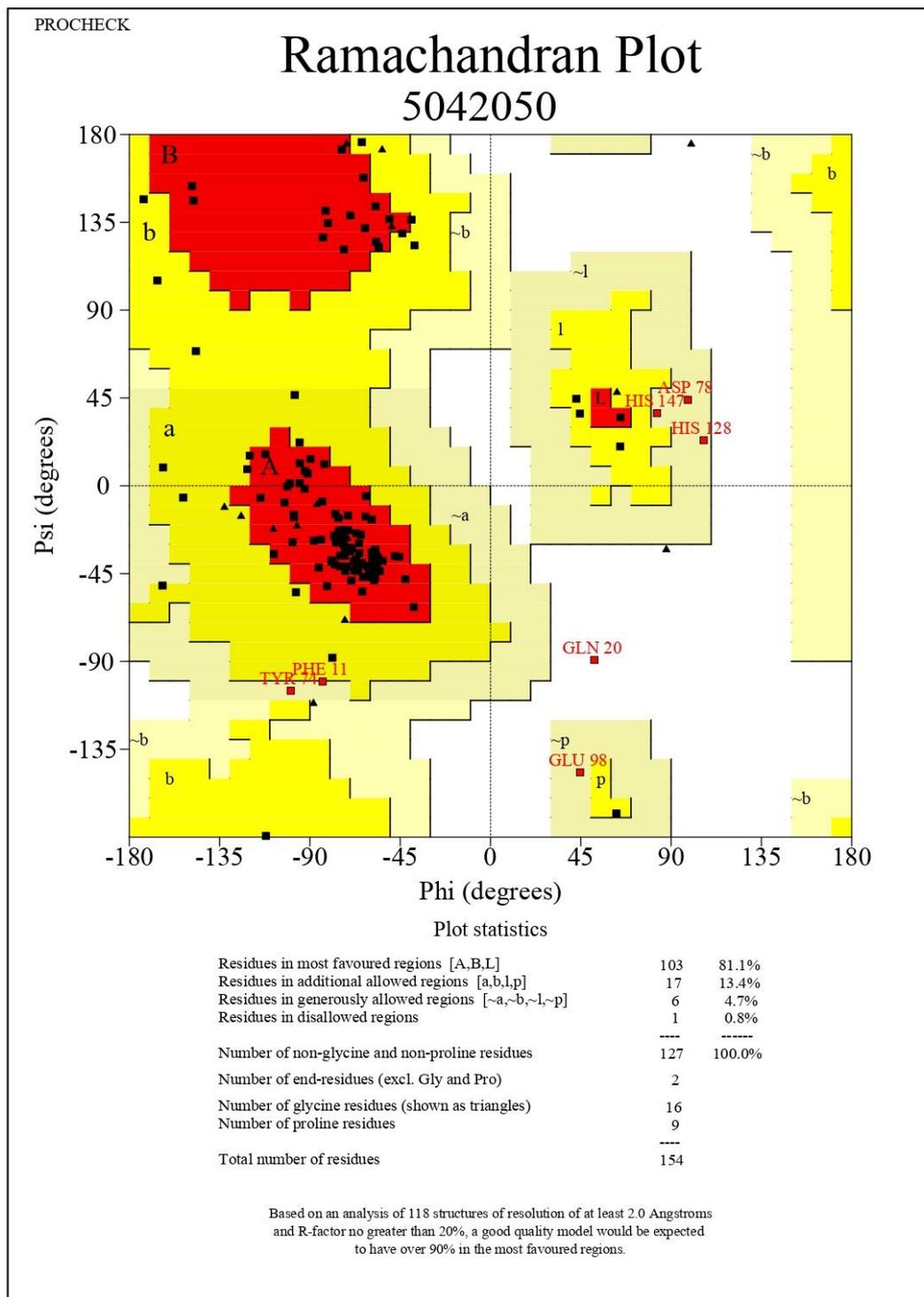


c) Domínio transmembrana



1447448_01.ps

d) Domínio intracelular



Fonte: PROCHECK.