



**UNIVERSIDADE ESTADUAL PAULISTA**  
**“JÚLIO DE MESQUITA FILHO”**  
Camus de São José do Rio Preto

**João Victor Piloto**

**Simulações moleculares da interação da região conservada de  
cisteínas do ectodomínio da proteína G do hRSV com o receptor  
celular CX3CR1 do hospedeiro**

São José do Rio Preto

2022

João Victor Piloto

**Simulações moleculares da interação da região conservada de cisteínas do ectodomínio da proteína G do hRSV com o receptor celular CX3CR1 do hospedeiro**

Dissertação apresentada como parte dos requisitos para a obtenção do título de Mestre em Biofísica Molecular, junto ao Programa de Pós-Graduação em Biofísica Molecular do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio Mesquita Filho”. Campus de São José do Rio Preto.

Financiadora: CAPES – Proc.:  
88887.502565/2020-00

Orientador: Prof. Dr. Ícaro Putinhon Caruso

Coorientador: Profa. Dra. Fátima Pereira de Souza

São José do Rio Preto  
2022

P643s      Piloto, João Victor

Simulações moleculares da interação da região conservada de cisteínas do ectodomínio da proteína G do hRSV com o receptor celular CX3CR1 do hospedeiro / João Victor Piloto. -- São José do Rio Preto, 2022    129 p.

Dissertação (mestrado) - Universidade Estadual Paulista (Unesp), Instituto de Biociências Letras e Ciências Exatas, São José do Rio Preto

Orientador: Ícaro Putinhon Caruso

Coorientadora: Fátima Pereira de Souza

1. Biofísica Molecular. 2. Virologia. 3. Dinâmica molecular. 4. Biomoléculas. I.

Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca do Instituto de Biociências Letras e Ciências Exatas, São José do Rio Preto. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

João Victor Piloto

**Simulações moleculares da interação da região conservada de  
cisteínas do ectodomínio da proteína G do hRSV com o receptor  
celular CX3CR1 do hospedeiro**

Dissertação apresentada como parte dos requisitos para a obtenção do título de Mestre em Biofísica Molecular, junto ao Programa de Pós-Graduação em Biofísica Molecular do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio Mesquita Filho”. Campus de São José do Rio Preto.

São José do Rio Preto, 10 de outubro de 2022

COMISSÃO EXAMINADORA

Prof. Dr. Ícaro Putinhon Caruso

Departamento de Física – UNESP- São José do Rio Preto

Prof. Dr. Francisco Gomes-Neto

Fundação Oswaldo Cruz – FIOCRUZ - Rio de Janeiro

Profa. Dra. Gabriela Campos de Araújo

Docente da União das Faculdades dos Grandes Lagos – UNILAGO – São José do Rio Preto

São José do Rio Preto  
10 de outubro de 2022

## **AGRADECIMENTOS**

O desenvolvimento deste trabalho de mestrado contou com a ajuda de diversas pessoas, dentre as quais eu agradeço:

Aos meus pais pela paciência que tiveram ao longo desses anos, ao meu irmão e minha namorada por compreenderem minha ausência em alguns momentos e a todos meus familiares.

Ao meu orientador Prof. Dr. Ícaro Putinhon Caruso pela ajuda, apoio, paciência que me concedia em todos os momentos, principalmente nos momentos mais difíceis.

A minha coorientadora Profa. Dra. Fátima Pereira de Souza pelo apoio e ajuda a quem me foi dado.

Aos professores orientadores, que durante todos esses meses me acompanharam pontualmente, dando auxílio necessário para a finalização deste projeto.

Aos professores da banca por aceitarem avaliar meu trabalho de mestrado.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. À CAPES pelo apoio financeiro neste projeto (88887.502565/2020-00), e as outras equipes de fomentos: FAPESP, CNPQ, FINEP, LNCC e FAPERJ.

A todos que participaram das pesquisas, pela colaboração e disposição de me ajudarem na realização, continuidade e finalização deste projeto.

**A todos meus sinceros agradecimentos!**

No fim tudo dá certo, e se não deu certo é  
porque ainda não chegou ao fim.

Fernando Sabino, 1998, p. 7

## RESUMO

O Vírus Sincicial Respiratório humano (hRSV) é um dos principais causadores de doenças respiratórias agudas como bronquiolite e pneumonia em crianças e idosos. Atualmente, as patologias causadas pelo hRSV não são bem entendidas e as vacinas formuladas não apresenta efetividade e segurança suficientes. A infetividade do vírus está relacionada com suas proteínas de membrana e dentre elas a glicoproteína G ou proteína G, que é responsável pela ligação do vírus à célula epiteliais do trato respiratório do hospedeiro e consequente instalação da infecção. Esta glicoproteína exerce um importante papel como antígeno de reconhecimento, sendo alvo para identificação do RSV através de anticorpos. Há evidências na literatura de que a proteína G interage com um receptor celular, conhecido como CX3CR1, porém não existe informações estruturais experimentais sobre essa interação. O objetivo do trabalho foi caracterizar a interação da região conservada de cisteína do ectodomínio da proteína G com o receptor celular CX3CR1 utilizando ferramentas computacionais. A partir da abordagem de modelagem molecular, modelos estruturais para as quatro isoformas do receptor celular foram calculados, os quais passaram por uma etapa de 300 ns de simulação de dinâmica molecular em bicamada lipídica de POPC para avaliação de suas estabilidades estruturais. Em seguida, cálculos de *docking* molecular buscaram pela conformação mais provável do peptídeo da região conservada de cisteína da G na interação com a região N-terminal do barril de alfa-hélices do CX3CR1. Por fim, novas simulações de 300 ns reportaram a estabilidade estrutural do modelo proposto para o complexo CX3CR1/peptídeo da G (CX3CR1/G), apontando os resíduos LYS-171, GLY-177, TYR-179, GLN-184 e ARG-272 da proteína como importante para a formação de pontes de hidrogênio com o peptídeo. Os cálculos de PCA e de RMSF mostraram uma comparação entre as formas livres e ligadas, mostrando um maior grau de flexibilidade das regiões N e C-terminais, assim como uma maior flexibilidade das alças entre as hélices  $\alpha$  das isoformas do CX3CR1. Os cálculos de energia livre de ligação teórica (MM/PBSA) mostraram que os resíduos ácidos interagem de forma favorável para a ligação e manutenção do peptídeo da G na região do *docking*, indicando que a isoforma 1 do CX3CR1 é aquela com energia de ligação mais favorável para a interação com o peptídeo da G. As informações

estruturais produzidas no presente trabalho podem trazer luz ao mecanismo de interação da proteína G do hRSV com o receptor celular CX3CR1, assim como proporcionar uma caracterização molecular do processo de adesão do vírus à célula do hospedeiro.

**Palavras chaves:** Hrsv. Proteína G. Receptor celular CX3CR1. Modelagem molecular. *Docking* molecular. Simulações de dinâmica molecular.

## ABSTRACT

Human Respiratory Syncytial Virus (hRSV) is one of the main causes of acute respiratory diseases such as bronchiolitis and pneumonia in children and the elderly. Currently, the pathologies caused by hRSV are not well understood and the formulated vaccines do not present sufficient effectiveness and safety. The infectivity of the virus is related to its membrane proteins and among them the G glycoprotein or G protein, which is responsible for binding the virus to the epithelial cells of the host's respiratory tract and consequent installation of the infection. This glycoprotein plays an important role as a recognition antigen, being a target for RSV identification through antibodies. There is evidence in the literature that the G protein interacts with a cellular receptor, known as CX3CR1, but there is no experimental structural information about this interaction. The objective of this work was to characterize the interaction of the conserved region of cysteine of the ectodomain of the G protein with the cellular receptor CX3CR1 using computational tools. From the molecular modeling approach, structural models for the four cellular receptor isoforms were calculated, which underwent a 300 ns step of molecular dynamics simulation in a POPC lipid bilayer to evaluate their structural stabilities. Then, molecular docking calculations looked for the most likely conformation of the peptide of the conserved region of cysteine of G in the interaction with the N-terminal region of the barrel of alpha-helices of CX3CR1. Finally, new 300 ns simulations reported the structural stability of the proposed model for the CX3CR1/G peptide complex (CX3CR1/G), pointing to residues LYS-171, GLY-177, TYR-179, GLN-184 and ARG- 272 of the protein as important for the formation of hydrogen bonds with the peptide. The PCA and RMSF calculations showed a comparison between the free and bound forms, showing a greater degree of flexibility of the N and C-terminal regions, as well as a greater flexibility of the loops between the  $\alpha$  helices of the CX3CR1 isoforms. Theoretical binding free energy (MM/PBSA) calculations showed that acid residues interact favorably for the binding and maintenance of the G peptide in the docking region, indicating that CX3CR1 isoform 1 is the one with the highest binding energy. Favorable for the interaction with the G peptide. The structural information produced in the present work may shed light on the mechanism of interaction of the hRSV G protein with the cell receptor CX3CR1, as well

as provide a molecular characterization of the virus adhesion process to the cell of the host.

**Keywords:** hRSV. Protein G. CX3CR1 cell receptor. Molecular modeling. Molecular docking. Molecular dynamics simulations.

## LISTA DE ILUSTRAÇÕES

- FIGURA 1:** Representação esquemática do RSV, sendo representadas as proteínas SH, F, G, M, P e N e o envelope da bicamada lipídica.....22
- FIGURA 2:** Representação esquemática dos domínios da Glicoproteína G do hRSV.....25
- FIGURA 3:** Sequências dos resíduos de aminoácidos das isoformas 1 (preto), 2 (vermelho), 3 (azul) e 4 (verde).....26
- FIGURA 4:** Modelos estruturais calculados pelos servidores de modelagem molelucar RosettaFOLD, Rosetta TR, I-Tasser e Phyre2. Modelos estruturais do receptor celular CX3CR1: isoforma 1 (A), 2 (B), 3 (C) e 4 (D). As proteínas são apresentadas com o modelo *cartoon* utilizando o programa de vizualização PyMOL. O barril de hélices alfa é denotado em cinza e as regiões N e C-terminal são destacadas em verde e azul, respectivamente.....45
- FIGURA 5:** Gráficos de Ramachandran para ângulos torcionais dos resíduos dos modelos estruturais para o receptor celular CX3CR1 ou isoforma 1 determinados pelos servidores Rosetta FOLD (A), Rosetta TR (B), Phyre2 (C) e I-Tasser (D). Em vermelho resíduos em regiões favoráveis, amarelo escuro para regiões adicionais permitidas, em amarelo claro para regiões generosamente adicionais e em branco para regiões não permitidas.....47
- FIGURA 6:** Gráficos de Ramachandran para ângulos torcionais dos resíduos dos modelos estruturais da isoforma 2 determinados pelos servidores Rosetta FOLD (A), Rosetta TR (B), Phyre2 (C) e I-Tasser (D). Em vermelho resíduos em regiões favoráveis, amarelo escuro para regiões adicionais permitidas, em amarelo claro para regiões generosamente adicionais e em branco para regiões não permitidas....49
- FIGURA 7:** Gráficos de Ramachandran para ângulos torcionais dos resíduos dos modelos estruturais da isoforma 3 determinados pelos servidores Rosetta FOLD (A), Rosetta TR (B), Phyre2 (C) e I-Tasser (D). Em vermelho resíduos em regiões favoráveis, amarelo escuro para regiões adicionais permitidas, em amarelo claro para regiões generosamente adicionais e em branco para regiões não permitidas....51
- FIGURA 8:** Gráficos de Ramachandran para ângulos torcionais dos resíduos dos modelos estruturais da isoforma 4 determinados pelos servidores Rosetta FOLD (A), Rosetta TR (B), Phyre2 (C) e I-Tasser (D). Em vermelho resíduos em regiões favoráveis, amarelo escuro para regiões adicionais permitidas, em amarelo claro para regiões generosamente adicionais e em branco para regiões não permitidas....53
- FIGURA 9:** Análise de RMSD dos átomos da cadeia principal (A), do número de contados < 0,6 nm por resíduos entre os átomos dos resíduos N-terminal e da fita beta GLU174–GLY177 (B), do ângulo de inclinação da proteína com relação a bicamada lipídica (C) e do ângulo do eixo principal com relação a normal da bicamada (D) ao longo das trajetórias de 300 ns das simulações de DM dos modelos estruturais da isoforma 1 do C3XCR1 determinados pelos servidores Rosetta FOLD (preto), Rosetta TR (vermelho), I-tasser (azul) e Phyre2 (ciano).....56
- FIGURA 10:** Estruturas representativas para os modelos estruturais da isoforma 1 do Rosetta FOLD (A), Rosetta TR (B), I-Tasser (C) e Phyre2 (D) em bicamada lipídica de POPC, as quais foram obtidas

pela análise de clusterização dos 300 ns de simulações de DM. A proteína é representada em modelo de *cartton* e os lipídeos como modelos de esferas e linhas. O barril de hélices alfa da proteína são denotados em cinza, a região N-terminal em verde e a C-terminal em azul. O átomo de fósforo da cabeça polar dos lipídeos é mostrado como esfera laranja e a cauda hidrofóbica como linhas cinzas. As imagens da direita mostram uma vista de cima de como a proteína encontra-se inserida na bicamada lipídica de POPC.....58

**FIGURA 11:** Perfil de densidade de massa para a bicamada de POPC solvatada para as quatro isoformas. A espessura da bicamada é determinada pelas distância P8-P8. A parte apolar da bicamada de POPC é denotada com uma linha preta, os fósforos P8 com linha vermelho, a proteína com linha verde e moléculas de água com linha azul.....60

**FIGURA 12:** Análise de RMSD dos átomos da cadeia principal (A) e do número de contados < 0,6 nm entre os átomos dos resíduos N-terminal e da fita beta GLU174–GLY177 (B) ao longo das trajetórias de 300 ns das simulações de DM dos modelos estruturais das isoformas 2, 3 e 4 do C3XCR1 determinados pelos servidores Rosetta FOLD.....61

**FIGURA 13:** Estruturas representativas para os modelos estruturais da isoforma 2 (A), isoforma 3 (B) e isoforma 4 (C) em bicamada lipídica de POPC, as quais foram obtidas pela análise de clusterização dos 300 ns de simulações de DM. A proteína é representada em modelo de *cartton* e os lipídeos como modelos de esferas e linhas. O barril de hélices alfa da proteína são denotados em cinza, a região N-terminal em verde e a C-terminal em azul. O átomo de fósforo da cabeça polar dos lipídeos é mostrado como esfera laranja e a cauda hidrofóbica como linhas cinzas. As imagens da direita mostram uma vista de cima de como a proteína encontra-se inserida na bicamada lipídica de POPC.....62

**FIGURA 14:** Perfil de densidade de massa para a bicamada de POPC solvatada com NaCl para as quatro isoformas. A espessura da bicamada é determinada pelas distância P8-P8. A parte apolar da bicamada de POPC é denotada com uma linha preta, os íons de NaCl são mostrados com linha amarela, os fósforos P8 com linha vermelho, a proteína com linha verde e moléculas de água com linha azul.....64

**FIGURA 15:** Análise de RMSD dos átomos da cadeia principal das isoformas (A) e do ângulo de inclinação da proteína com relação a bicamada lipídica de POPC com 150 mM de NaCl (B) ao longo das trajetórias de 300 ns das simulações de DM das isoformas do modelo estrutural de CX3CR1. Sendo a isoforma 1 representada em preto, a isoforma 2 em vermelho, a isoforma 3 em verde e a isoforma 4 em azul.....66

**FIGURA 16:** Comparação entre as estruturas de POPC e DPPC.....67

**FIGURA 17:** Perfil de densidade de massa para a bicamada de DPPC solvatada para as quatro isoformas. A espessura da bicamada é determinada pelas distância P8-P8. A parte apolar da bicamada de DPPC é denotada com uma linha preta, os fósforos P8 com linha vermelho, a proteína com linha verde e moléculas de água com linha azul.....68

**FIGURA 18:** Perfil de densidade de massa para a bicamada de DPPC solvatada com concentração de íons NaCl de 150mM para as quatro isoformas. A espessura da bicamada é determinada pelas

distância P8-P8. A parte apolar da bicamada de DPPC (preto), os íons de NaCl (amarelo), os fósforos P8 (vermelho), a parte da proteína (verde) e a água (azul).....69

**FIGURA 19:** Análise de RMSD dos átomos da cadeia principal das isoformas (A e B) e do ângulo de inclinação da proteína com relação a normal da bicamada lipídica de DPPC (C e D), sendo os resultados na presença de 150 mM de cloreto de sódio na coluna da direita e sem NaCl na coluna da esquerda, ao longo das trajetórias de 300 ns das simulações de DM das isoformas para os modelos estruturais do receptor celular CX3CR1. Sendo a isoforma 1 representada em preto, isoforma 2 em vermelho, isoforma 3 em verde e isoforma 4 em azul.....70

**FIGURA 20:** Estrutura de cristal (PDB 4XTI) do complexo formado entre a fractalcina CX3CL1 (magenta) com o receptor homólogo US28 (cinza). As proteínas foram representadas com o modelo *cartoon*, com os resíduos N- e C-terminais denotados em verde e azul, respectivamente. A correspondente fita beta do homólogo CX3CR1 (GLU174–GLY177, numeração referente a isoforma 1) é destacada em ciano na cavidade de ligação da proteína.....72

**FIGURA 21:** Resultados dos cálculos de *docking* molecular para o servidor ClusPro (A) e HADDOCK (B). A proteína é apresentada com o modelo de *cartoon*, sendo o barril de hélices alfa colorido em cinza e o N e C-terminal em verde e azul, respectivamente. A fita beta GLU174–GLY177 (numeração referente a isoforma 1) na cavidade de ligação da proteína é destacada em ciano. O peptídeo da região de cisteínas conservadas da proteína G do hRSV é mostrado como modelo de *cartoon* em magenta e suas pontes de dissulfeto são denotadas com o modelo de *sticks* em amarelo.....74

**FIGURA 22:** Valores de RMSD calculados para os átomos da cadeia principal do proteína e do peptídeo da G ao longo dos 300 ns de simulação de DM para os modelos estruturais dos complexos determinados pelo servidores ClusPro (esquerda) e HADDOCK (direita). Os resultados de cada isoforma é apresentado no seguinte esquema de cores: preto para a isoforma 1, vermelho para a 2, azul para a 3 e ciano para a 4. A linha grossa na parte B denota os valores de RMSD do peptídeo para o alinhamento estrutural realizado com relação a proteína, enquanto que a linha fina se refere a um alinhamento feito com relação ao próprio peptídeo.....76

**FIGURA 23:** Valores de variação/diferença de distância (distância num tempo qualquer subtraído pela distância na posição inicial  $t = 0$  ns) calculadas entre os centros de massa do barril de hélices alfa da proteína e do peptídeo da G ao longo de 300 ns de simulação para os modelos estruturais dos complexos determinados pelo ClusPro (esquerda) e HADDOCK (direita). Os resultados de cada isoforma é apresentado no seguinte esquema de cores: preto para a isoforma 1, vermelho para a 2, azul para a 3 e ciano para a 4.....77

**FIGURA 24:** Comparação entre as posições do peptídeo da G nos modelos estruturais dos cálculos de *docking* molecular, (A) ClusPRO e (B) HADDOCK, e nas estruturas representativas determinadas pela análise de clusterização das trajetórias de 300 ns de simulações dos complexos. A proteína é apresentada com o modelo de *cartoon*, sendo o barril de hélices alfa colorido em cinza e o N e C-terminal em verde e azul, respectivamente. A fita beta GLU174–GLY177 (numeração referente a isoforma 1) na cavidade de ligação da proteína é destacado em ciano. O peptídeo da região de cisteínas conservadas

da proteína G do hRSV é mostrado como modelo de *cartoon* em magenta para o modelo de *docking* e em amarelo para o estrutura representativa da clusterização.....79

**FIGURA 25:** Número de contatos menor que 0,6 nm determinados entre os átomos da proteína e do peptídeo da G ao longo dos 300 ns de simulação para os modelos estruturais dos complexos calculados pelo ClusPro (esquerda) e HADDOCK (direita). Os resultados de cada isoforma é apresentado no seguinte esquema de cores: preto para a isoforma 1, vermelho para a 2, azul para a 3 e ciano para a 4.....80

**FIGURA 26:** Número de ligações de hidrogênios determinados entre os átomos da proteína e do peptídeo da G ao longo dos 300 ns de simulação para os modelos estruturais dos complexos calculados pelo ClusPro (esquerda) e HADDOCK (direita). Os resultados de cada isoforma é apresentado no seguinte esquema de cores: preto para a isoforma 1, vermelho para a 2, azul para a 3 e ciano para a 4.....81

**FIGURA 27:** Representação esquemática em cartoon do receptor celular CX3CR1 e peptídeo G e em sticks os resíduos mais recorrentes da persistência das ligações de hidrogênios. Em cinza o receptor celular, com o N-terminal em verde, em ciano a fita  $\beta$  e os resíduos recorrentes na forma de sticks, carbonos em cinza, nitrogênio em azul e em vermelho oxigênio (A). Em magenta o peptídeo G, sendo as cisteínas representadas em amarelo na forma de sticks e os resíduos recorrentes em sticks, sendo carbono em bege, nitrogênio azul e oxigênio em vermelho (B).....88

**FIGURA 28:** Valores de RMSF em função do número de resíduos para as quatro isoformas, isoforma 1 (A), isoforma 2 (B), isoforma 3 (C) e isoforma 4 (D) do receptor celular CX3CR1, com a forma livre do receptor celular CX3CR1 (preto), sua estrutura na forma de complexo CX3CR1/G obtida pela DM do resultado do ClusPro (verde) e sua estrutura na forma de complexo CX3CR1/G obtida pela DM do resultado do HADDOCK (azul). A estrutura secundária está mostrada ao topo de cada isoforma, sendo as hélices  $\alpha$  do barril em amarelo, as folhas  $\beta$  em verde, e em cinza a estrutura secundária do N-terminal e C-terminal. Os cálculos de RMSF foram realizados para os 300 ns de simulação de dinâmica molecular.....91

**FIGURA 29:** Gráficos de dispersão PCA, PCA1 e PCA2 para CX3CR1 livre (pontos vermelhos) e para complexo CX3CR1/G da DM dado pelo *docking* do ClusPro(pontos pretos) da isoforma 1 (A). (B e C) Movções filtradas de os autovetores de PCA1 (B) e PCA2 (C) para a dinâmica dados de CX3CR1 em sua forma livre e na sua forma de complexo CX3CR1/G (D e E) . O movimento é indicada pela variação de cor de azul para vermelho. Para realizar os cálculos de PCA, os 300 ns de trajetórias foram concatenados para cada uma das moléculas sistemas (CX3CR1 livre ou CX3CR1/G).....92

**FIGURA 30:** Gráficos de dispersão PCA, PCA1 e PCA2 para CX3CR1 livre (pontos vermelhos) e para complexo CX3CR1/G da DM dado pelo *docking* do ClusPro(pontos pretos) da isoforma 2 (A). (B e C) Movções filtradas de os autovetores de PCA1 (B) e PCA2 (C) para a dinâmica dados de CX3CR1 em sua forma livre e na sua forma de complexo CX3CR1/G (D e E) . O movimento é indicada pela variação de cor de azul para vermelho. Para realizar os cálculos de PCA, os 300 ns de trajetórias foram concatenados para cada uma das moléculas sistemas (CX3CR1 livre ou CX3CR1/G).....94

**FIGURA 31:** Gráficos de dispersão PCA, PCA1 e PCA2 para CX3CR1 livre (pontos vermelhos) e para complexo CX3CR1/G da DM dado pelo *docking* do ClusPro(pontos pretos) da isoforma 3 (A). (B e C) Moções filtradas de os autovetores de PCA1 (B) e PCA2 (C) para a dinâmica dados de CX3CR1 em sua forma livre e na sua forma de complexo CX3CR1/G (D e E) . O movimento é indicada pela variação de cor de azul para vermelho. Para realizar os cálculos de PCA, os 300 ns de trajetórias foram concatenados para cada uma das moléculas sistemas (CX3CR1 livre ou CX3CR1/G).....95

**FIGURA 32:** Gráficos de dispersão PCA, PCA1 e PCA2 para CX3CR1 livre (pontos vermelhos) e para complexo CX3CR1/G da DM dado pelo *docking* do ClusPro(pontos pretos) da isoforma 4 (A). (B e C) Moções filtradas de os autovetores de PCA1 (B) e PCA2 (C) para a dinâmica dados de CX3CR1 em sua forma livre e na sua forma de complexo CX3CR1/G (D e E) . O movimento é indicada pela variação de cor de azul para vermelho. Para realizar os cálculos de PCA, os 300 ns de trajetórias foram concatenados para cada uma das moléculas sistemas (CX3CR1 livre ou CX3CR1/G).....97

**FIGURA 33:** Gráficos de dispersão PCA, PCA1 e PCA2 para CX3CR1 livre (pontos vermelhos) e para complexo CX3CR1/G da DM dado pelo *docking* do HADDOCK(pontos azuis) da isoforma 1 (A). (B e C) Moções filtradas de os autovetores de PCA1 (B) e PCA2 (C) para a dinâmica dados de CX3CR1 em sua forma livre e na sua forma de complexo CX3CR1/G (D e E) . O movimento é indicada pela variação de cor de azul para vermelho. Para realizar os cálculos de PCA, os 300 ns de trajetórias foram concatenados para cada uma das moléculas sistemas (CX3CR1 livre ou CX3CR1/G).....99

**FIGURA 34:** Gráficos de dispersão PCA, PCA1 e PCA2 para CX3CR1 livre (pontos vermelhos) e para complexo CX3CR1/G da DM dado pelo *docking* do HADDOCK(pontos azuis) da isoforma 2 (A). (B e C) Moções filtradas de os autovetores de PCA1 (B) e PCA2 (C) para a dinâmica dados de CX3CR1 em sua forma livre e na sua forma de complexo CX3CR1/G (D e E) . O movimento é indicada pela variação de cor de azul para vermelho. Para realizar os cálculos de PCA, os 300 ns de trajetórias foram concatenados para cada uma das moléculas sistemas (CX3CR1 livre ou CX3CR1/G).....100

**FIGURA 35:** Gráficos de dispersão PCA, PCA1 e PCA2 para CX3CR1 livre (pontos vermelhos) e para complexo CX3CR1/G da DM dado pelo *docking* do HADDOCK(pontos azuis) da isoforma 3 (A). (B e C) Moções filtradas de os autovetores de PCA1 (B) e PCA2 (C) para a dinâmica dados de CX3CR1 em sua forma livre e na sua forma de complexo CX3CR1/G (D e E) . O movimento é indicada pela variação de cor de azul para vermelho. Para realizar os cálculos de PCA, os 300 ns de trajetórias foram concatenados para cada uma das moléculas sistemas (CX3CR1 livre ou CX3CR1/G).....102

**FIGURA 36:** Gráficos de dispersão PCA, PCA1 e PCA2 para CX3CR1 livre (pontos vermelhos) e para complexo CX3CR1/G da DM dado pelo *docking* do HADDOCK(pontos azuis) da isoforma 4 (A). (B e C) Moções filtradas de os autovetores de PCA1 (B) e PCA2 (C) para a dinâmica dados de CX3CR1 em sua forma livre e na sua forma de complexo CX3CR1/G (D e E) . O movimento é indicada pela variação de cor de azul para vermelho. Para realizar os cálculos de PCA, os 300 ns de trajetórias foram concatenados para cada uma das moléculas sistemas (CX3CR1 livre ou CX3CR1/G).....103

**FIGURA 37:** Contribuição individual de cada resíduo de aminoácido na energia total do receptor celular CX3CR1 para os resultados de docking obtidos pelas quatro isoformas do server ClusPro. Sendo contribuição favorável (vermelho) e contribuição desfavorável (azul).....105

**FIGURA 38:** Contribuição individual de cada resíduo de aminoácido na energia total do receptor celular CX3CR1 para os resultados de docking obtidos pelas quatro isoformas do server HADDOCK. Sendo contribuição favorável (vermelho) e contribuição desfavorável (azul).....106

**FIGURA 39:** Estrutura representativa das isoformas do receptor celular CX3CR1 com os resíduos que apresentam uma contribuição favoráveis e desfavoráveis significativas (acima da média mais desvio padrão) para a energia livre de ligação total para as simulações de DM dos modelos estruturais dos complexos com o peptídeo da G determinados pelo servidor ClusPro. As contribuições favoráveis e desfavoráveis para as energias totais são representadas gradientes de cores indo vermelho para o azul, respectivamente. As barras de cores denotam o aspecto quantitativo de cada gradiente de cor. As isoformas 1 (A), 2 (B), 3 (C) e 4 (D) são apresentadas em modelos de *cartoon* em cinza e os resíduos de aminoácidos na forma de sticks.....111

**FIGURA 40:** Estrutura representativa das isoformas do receptor celular CX3CR1 com os resíduos que apresentam uma contribuição favoráveis e desfavoráveis significativas (acima da média mais desvio padrão) para a energia livre de ligação total para as simulações de DM dos modelos estruturais dos complexos com o peptídeo da G determinados pelo servidor HADDOCK. As contribuições favoráveis e desfavoráveis para as energias totais são representadas gradientes de cores indo vermelho para o azul, respectivamente. As barras de cores denotam o aspecto quantitativo de cada gradiente de cor. As isoformas 1 (A), 2 (B), 3 (C) e 4 (D) são apresentadas em modelos de *cartoon* em cinza e os resíduos de aminoácidos na forma de sticks.....116

## LISTA DE TABELAS

<b>TABELA 1:</b> Valores de área por lipídio em bicamada lipídica de POPC dos programas utilizados para modelagem molecular.....	40
<b>TABELA 2:</b> Valores de área por lipídio das isoformas utilizando Rosetta FOLD para os modelos de bicamada lipídica em POPC e DPPC.....	40
<b>TABELA 3:</b> Valores das porcentagens dos ângulos torcionais dos resíduos dos modelos estruturais do CX3CR1 ou isoforma 1 encontrados nas regiões descritas pelo gráfico de Ramachandran.....	47
<b>TABELA 4:</b> Valores das porcentagens dos ângulos torcionais dos resíduos dos modelos estruturais da isoforma 2 encontrados nas regiões descritas pelo gráfico de Ramachandran.....	49
<b>TABELA 5:</b> Valores das porcentagens dos ângulos torcionais dos resíduos dos modelos estruturais da isoforma 3 encontrados nas regiões descritas pelo gráfico de Ramachandran.....	51
<b>TABELA 6:</b> Valores das porcentagens dos ângulos torcionais dos resíduos dos modelos estruturais da isoforma 4 encontrados nas regiões descritas pelo gráfico de Ramachandran.....	53
<b>TABELA 7:</b> Valores de RMSD para a região do barril de hélices alfa calculados para os modelos estruturais determinados pelos servidores Rosetta FOLD, Rosetta TR, I-tasser e Phyre2. O modelo do Rosetta FOLD foi utilizado como referência para o alinhamento estrutural e cálculo de RMSD.....	54
<b>TABELA 8:</b> Porcentagens de persistência maior que 10% para as ligações de hidrogênio formadas entre o receptor celular CX3CR1 e o peptídeo da G ao longo dos 300 ns de simulações de DM para o modelo estrutural do complexo obtidos a partir do servidor ClusPro. Para diferenciar da proteína, os resíduos do peptídeo da G são denotados com um asterisco (*).....	83
<b>TABELA 9:</b> Porcentagens de persistência maior que 10% para as ligações de hidrogênio formadas entre a isoforma 2 do receptor celular CX3CR1 e o peptídeo da G ao longo dos 300 ns de simulações de DM para o modelo estrutural do complexo obtidos a partir do servidor ClusPro. Para diferenciar da proteína, os resíduos do peptídeo da G são denotados com um asterisco (*).....	83
<b>TABELA 10:</b> Porcentagens de persistência maior que 10% para as ligações de hidrogênio formadas entre a isoforma 3 do receptor celular CX3CR1 e o peptídeo da G ao longo dos 300 ns de simulações de DM para o modelo estrutural do complexo obtidos a partir do servidor ClusPro. Para diferenciar da proteína, os resíduos do peptídeo da G são denotados com um asterisco (*).....	84
<b>TABELA 11:</b> Porcentagens de persistência maior que 10% para as ligações de hidrogênio formadas entre a isoforma 4 do receptor celular CX3CR1 e o peptídeo da G ao longo dos 300 ns de simulações de DM para o modelo estrutural do complexo obtidos a partir do servidor ClusPro. Para diferenciar da proteína, os resíduos do peptídeo da G são denotados com um asterisco (*).....	84
<b>TABELA 12:</b> Porcentagens de persistência maior que 10% para as ligações de hidrogênio formadas entre o receptor celular CX3CR1 e o peptídeo da G ao longo dos 300 ns de simulações de DM para o modelo estrutural do complexo obtidos a partir do servidor HADDOCK. Para diferenciar da proteína, os resíduos do peptídeo da G são denotados com um asterisco (*).....	85
<b>TABELA 13:</b> Porcentagens de persistência maior que 10% para as ligações de hidrogênio formadas entre a isoforma 2 do receptor celular CX3CR1 e o peptídeo da G ao longo dos 300 ns de simulações	

de DM para o modelo estrutural do complexo obtidos a partir do servidor HADDOCK. Para diferenciar da proteína, os resíduos do peptídeo da G são denotados com um asterisco (*).....	86
<b>TABELA 14:</b> Porcentagens de persistência maior que 10% para as ligações de hidrogênio formadas entre a isoforma 3 do receptor celular CX3CR1 e o peptídeo da G ao longo dos 300 ns de simulações de DM para o modelo estrutural do complexo obtidos a partir do servidor HADDOCK. Para diferenciar da proteína, os resíduos do peptídeo da G são denotados com um asterisco (*).....	86
<b>TABELA 15:</b> Porcentagens de persistência maior que 10% para as ligações de hidrogênio formadas entre a isoforma 4 do receptor celular CX3CR1 e o peptídeo da G ao longo dos 300 ns de simulações de DM para o modelo estrutural do complexo obtidos a partir do servidor HADDOCK. Para diferenciar da proteína, os resíduos do peptídeo da G são denotados com um asterisco (*).....	87
<b>TABELA 16:</b> Resíduos de aminoácidos da isoforma 1 do CX3CR1 e seus respectivos valores das contribuições favoráveis e desfavoráveis (acima do valor de corte média mais desvio padrão) para a energia livre de ligação total para a trajetória de DM do modelo estrutural determinado pelo servidor ClusPro.....	108
<b>TABELA 17:</b> Resíduos de aminoácidos da isoforma 2 do CX3CR1 e seus respectivos valores das contribuições favoráveis e desfavoráveis (acima do valor de corte média mais desvio padrão) para a energia livre de ligação total para a trajetória de DM do modelo estrutural determinado pelo servidor ClusPro.....	108
<b>TABELA 18:</b> Resíduos de aminoácidos da isoforma 3 do CX3CR1 e seus respectivos valores das contribuições favoráveis e desfavoráveis (acima do valor de corte média mais desvio padrão) para a energia livre de ligação total para a trajetória de DM do modelo estrutural determinado pelo servidor ClusPro.....	109
<b>TABELA 19:</b> Resíduos de aminoácidos da isoforma 4 do CX3CR1 e seus respectivos valores das contribuições favoráveis e desfavoráveis (acima do valor de corte média mais desvio padrão) para a energia livre de ligação total para a trajetória de DM do modelo estrutural determinado pelo servidor ClusPro.....	110
<b>TABELA 20:</b> Resíduos de aminoácidos da isoforma 1 do CX3CR1 e seus respectivos valores das contribuições favoráveis e desfavoráveis (acima do valor de corte média mais desvio padrão) para a energia livre de ligação total para a trajetória de DM do modelo estrutural determinado pelo servidor HADDOCK.....	112
<b>TABELA 21:</b> Resíduos de aminoácidos da isoforma 2 do CX3CR1 e seus respectivos valores das contribuições favoráveis e desfavoráveis (acima do valor de corte média mais desvio padrão) para a energia livre de ligação total para a trajetória de DM do modelo estrutural determinado pelo servidor HADDOCK.....	113
<b>TABELA 22:</b> Resíduos de aminoácidos da isoforma 3 do CX3CR1 e seus respectivos valores das contribuições favoráveis e desfavoráveis (acima do valor de corte média mais desvio padrão) para a energia livre de ligação total para a trajetória de DM do modelo estrutural determinado pelo servidor HADDOCK.....	113

<b>TABELA 23:</b> Resíduos de aminoácidos da isoforma 4 do CX3CR1 e seus respectivos valores das contribuições favoráveis e desfavoráveis (acima do valor de corte média mais desvio padrão) para a energia livre de ligação total para a trajetória de DM do modelo estrutural determinado pelo servidor HADDOCK.....	114
<b>TABELA 24:</b> Valores das contribuições favoráveis e desfavoráveis para os resíduos de aminoácidos do peptídeo da G com as respectivas isoformas (1,2,3 e 4) obtidos para as simulações de DM dos modelos estruturais dos complexos determinados pelo servidor ClusPro.....	117
<b>TABELA 25</b> Valores das contribuições favoráveis e desfavoráveis para os resíduos de aminoácidos do peptídeo da G com as respectivas isoformas (1,2,3 e 4) obtidos para as simulações de DM dos modelos estruturais dos complexos determinados pelo servidor HADDOCK.....	118
<b>TABELA 26:</b> Valores de variação de energia livre de ligação total para a interação do peptídeo da G com as isoformas (1,2,3 e 4) calculados a partir das simulações de DM para os modelos estruturais dos complexos obtidos pelos servidores ClusPro e HADDOCK.....	119

## LISTA DE ABREVIATURAS E SIGLAS

CX3CR1 - Receptor Celular para a Adesão da Proteína G

DM - Dinâmica Molecular

DPPC - Dipalmitoilfosfatidilcolina

hRSV - *Respiratory Syncytial Virus human* (Vírus Sincicial Respiratório humano)

LNCC- Laboratório Nacional de Computação Científica

POPC - 1-palmitoil-2-oleoil-sn-glicero-3-fosfocolina.

PCA - *Principal Component Analysis* (Análise da Componente Principal)

RMSD - *Root Mean Square Deviation* (Desvio da Raiz Quadrática Média)

RMSF - *Root Mean Square Fluctuation* (flutuação da Raiz Quadrática Média)

GROMACS- Groningen Machine for Chemical Simulations

GROMOS –Groningen Molecular Simulations

PDB- Protein Data Bank

PyMOL- Sistema de Visualização Molecular

VMD- Visual Molecular Dynamics

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	200
1.1	Dinâmica Molecular	27
1.2	Docking Molecular	30
1.3	Cálculo de Energia Livre de Ligação	33
<b>2</b>	<b>OBJETIVOS GERAIS E ESPECÍFICOS</b>	35
<b>3</b>	<b>METODOLOGIA</b>	36
3.1	Modelagem Molecular	36
3.2	Dinâmica Molecular em Bicamada Lipídica POPC e DPPC	38
3.3	Docking Molecular	40
3.4	Dinâmica Molecular dos Complexos das Isoformas do CX3CR1 com o Peptídeo da Proteína G	42
3.5	Cálculo de Variação de Energia Livre de Ligação Teórica	44
<b>4</b>	<b>RESULTADOS e DISCUSSÕES</b>	44
4.1	Modelagem Molecular	44
4.2	Simulações de Dinâmica Molecular do Receptor e suas Isoformas	54
4.3	Simulações de Dinâmica Molecular de CX3CR1 e suas Isoformas em POPC na Presença de Força Iônica e em DPPC na Ausência e Presença de Força Iônica	63
4.4	Cálculos de Docking Molecular e Simulações de Dinâmica Molecular da Interação do CX3CR1 com o Peptídeo da Região Conservada de Cisteínas da Proteína G do hRSV	71
4.5	Cálculo de Energia Livre de Ligação	104
<b>5</b>	<b>CONCLUSÕES e PERSPECTIVAS</b>	120
	<b>REFERÊNCIAS</b>	122

## 1 INTRODUÇÃO

O Vírus Sincicial Respiratório humano (do inglês, *human Respiratory Syncytial Virus* - hRSV) foi isolado pela primeira vez em 1956, por Morris, Blount e Savage, em estudo para investigar o agente etiológico responsável pela doença respiratória aguda em chimpanzés do Instituto de Pesquisa Walter Reed Army, nos Estados Unidos da América, sendo descrito como “agente de coriza de chimpanzé” (MORRIS et al., 1956; COLLINS & GRAHAM, 2008). Em 1957, após estudos em cultura de tecido, o vírus foi renomeado para Vírus Sincicial Respiratório (RSV), devido ao seu padrão exclusivo de formar sincícios celulares multinucleados e, só em 1981, começou a caracterização detalhada do RSV com a clonagem molecular e sequenciamento do seu genoma (COLLINS et al., 2013; HOGAN et al., 2018).

Este vírus é um dos principais agentes causadores de doenças agudas do trato respiratório, tais como pneumonia e bronquiolite em crianças com idade igual e abaixo de cinco anos, em todo o mundo, através de epidemias sazonais decorrente do clima de cada região do planeta. Todavia, não são somente as crianças que são acometidas por esse vírus, idosos e imunocomprometidos também fazem parte do grupo de risco, aumentando a taxa de mortalidade nessas epidemias. O principal grupo de risco são crianças que apresentam prematuridade, imunodeficiência congênita e doenças cardiorrespiratórias (COLLINS; CROWE JR, 2007; HALL et al., 2009; OTHUMPANGAT et al., 2009; FALSEY et al., 2005; RICCETTO et al., 2009; YE, et al., 2019).

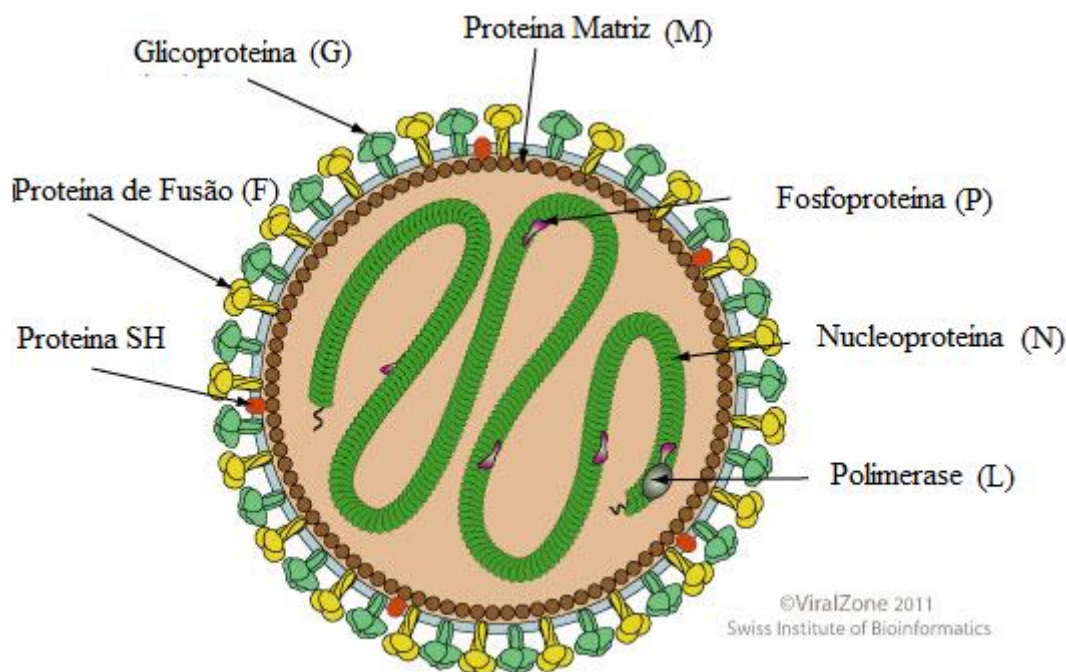
As infecções por hRSV podem causar desde hospitalizações com quadros clínicos graves e até a morte. Um levantamento epidemiológico global relatou que 33,1 milhões de episódios de infecções por hRSV ocorreram em 2015 no mundo todo, sendo 3,2 milhões de hospitalizações e 59.600 mil mortes de crianças menores de cinco anos (SHI et al., 2017).

As infecções pelo hRSV podem acometer pessoas de qualquer idade, apresentando-se como um quadro gripal ou um resfriado. Ainda não há um tratamento eficaz para os casos de hRSV graves, nem mesmo vacinas para atenuar e/ou evitar a infecção por este vírus (MORRIS et al., 2009; LIU et al., 2007; MARTÍNEZ et al., 2007). Estudos foram realizados para a busca de um medicamento ou até uma vacina

eficiente contra o hRSV, porém a variabilidade antigênica viral e as características do sistema imune do hospedeiro dificultam a produção destas vacinas (PARVEEN et al., 2006). A administração intravenosa de anticorpo monoclonal humanizado (Palivizumab), que age contra a atividade de fusão da proteína F, é a principal medida de prevenção utilizada atualmente. Todavia, devido ao alto custo dessa intervenção e, ela tem sido aplicada apenas em crianças prematuras com grande risco de desenvolver infecções graves, visto que o Palivizumab confere uma imunidade temporária, assim diminui o risco da criança prematura vir a óbito (MEISSNER et al., 2003). A única droga antiviral aprovada para tratamento do hRSV é a Ribavirina aerolizada (usada para tratamentos de outras doenças respiratórias), a qual é atribuída a atividade de inibição da síntese de ácido nucléico viral. A Ribavirina tem grande importância no tratamento de adultos imunocomprometidos, no entanto é responsável pela ocorrência de alguns efeitos colaterais. (NOKES e CANE, 2008).

A partícula viral do hRSV é constituída das seguintes proteínas: proteína hidrofóbica pequena (SH), proteína de fusão (F), proteína matriz (M), proteína M2-1, proteína L, fosfoproteína (P), nucleoproteína (N) e glicoproteína (G) (Figura 1). Além dessas proteínas presentes no vírus maduro, as proteínas não estruturais NS1 e NS2 são produzidas durante o processo de infecção viral na célula hospedeira, assim como a proteína M2-2, para a qual há escassa informação na literatura sobre suas funções.

**FIGURA 1:** Representação esquemática do hRSV, sendo representadas as proteínas SH, F, G, M, P e N e o envelope da bicamada lipídica.



Fonte: Adaptada Swiss Institute of Bioinformatics. ViralZone® [database]. Suíça: SIB; 2009. Disponível em: [http://www.expasy.org/viralzone/all\\_by\\_species/90.html](http://www.expasy.org/viralzone/all_by_species/90.html)

O nucleocapsídeo presente nas partículas virais maduras é formado pelo RNA genômico complexado com as proteínas N, P, L e M2-1 (COLLINS et al. 2001). A proteína M2-1 é necessária para uma eficiente transcrição do genoma pela polimerase viral (FEARNS E COLLINS, 1999). A proteína N está associada ao RNA genômico e é a principal proteína estrutural do nucleocapsídeo (COLLINS et al., 2001). O RNA genômico do hRSV encontra-se sempre complexado com a proteína N e este complexo é usado como molde para a síntese de RNA. Este complexo proteína N-RNA previne a formação de estruturas secundárias no RNA, protegendo-o de degradação por nucleases e reduzindo a formação de RNAs dupla fita, podendo levar a resposta celular antiviral (COWTON et al., 2006).

As proteínas não estruturais NS1 e NS2 interferem na indução e sinalização do interferon (COLLINS & CROWE JR, 2007), o qual representa o principal mecanismo da imunidade inata no combate a infecções virais. O hRSV, através das proteínas NS1

e NS2, pode agir inibindo ou neutralizando várias etapas da via dos interferons, o que favorece o maior sucesso da progênie viral (SWEDAN et al., 2009; LO et al., 2005).

A proteína P é altamente fosforilada e desempenha função importante na transcrição e replicação do hRSV (COLLINS et al., 2001). A proteína P constitui uma subunidade da RNA-polimerase viral (formada pelas proteínas P e L) e parece mediar o contato entre a proteína L e o conjunto formado pela proteína N e o RNA (TRAN et al., 2007).

A proteína L, a maior proteína do hRSV, contém seis segmentos altamente conservados sendo a proteína estrutural de menor abundância na partícula viral (COLLINS et al., 2001). Atribui-se a essa proteína as atividades de RNA-polimerase e de poliadenilação e para formação de estrutura de *cap* do RNAm viral (MASON et al., 2004).

A proteína M2-1 é interna e não glicosilada e interage com as proteínas N e P, atuando no processo de transcrição (COLLINS et al., 2001). Essa proteína tem função na replicação do hRSV e alterações em sua sequência podem diminuir a infectividade do vírus (TANG et al., 2001). A proteína M2-2 está relacionada à regulação da transcrição e replicação do RNA viral (ATREYA et al., 1998). A M2-2, quando expressa em altos níveis, inibe completamente a replicação do hRSV, mostrando que se essa proteína estiver em grande quantidade, pode ser um fator crítico para a replicação viral (CHENG et al., 2005).

A proteína M encontra-se na face interna do envelope viral, apresentando um papel-chave na montagem e liberação do vírus e associando o complexo ribonucleoproteico com as glicoproteínas do envelope viral (GHILDYAL et al., 2002).

A glicoproteína de superfície SH é uma molécula pequena, hidrofóbica e transmembrânica cuja função ainda não está completamente elucidada, mas possivelmente sua atuação esteja relacionada ao Complexo de Golgi da célula hospedeira, no qual facilitaria o transporte das proteínas virais pela via de secreção (RIXON et al., 2004). Também se associa a SH um importante fator de virulência, uma vez que essa proteína altera a permeabilidade da membrana celular (Perez et al., 1997).

A proteína F é uma glicoproteína responsável pela fusão do envelope viral com a membrana citoplasmática da célula hospedeira, fazendo com que o nucleocapsídeo

seja liberado no interior do citoplasma (WALSH e HRUSKA, 1983). Essa proteína também proporciona a fusão de células infectadas formando sincícios (COLLINS et al., 2001).

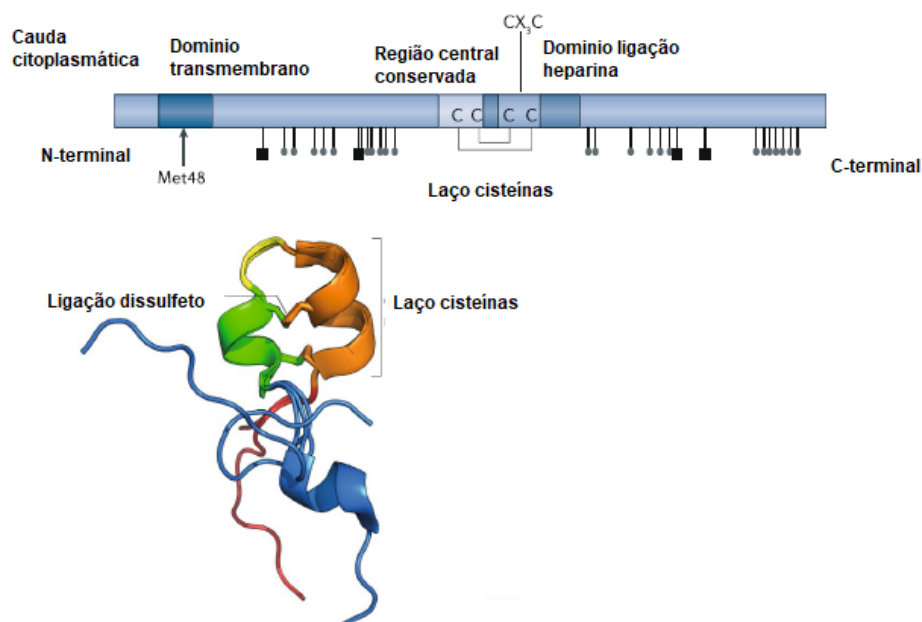
A proteína G do hRSV é uma proteína de superfície, tem de 289 a 319 resíduos de aminoácidos dependendo da amostra viral (COLINS et al., 2002). Ela foi identificada como a responsável pela adesão do vírus à célula hospedeira, a partir da observação de que anticorpos específicos contra a mesma inibem esse processo (COLLINS et al., 2001, LEVINE et al., 1987; BOTOSSO et al., 2009). A proteína G tem sido alvo de estudos sobre a variabilidade genética do hRSV. Estudos indicam que ela é pouco conservada entre os grupos A e B, mostrando até 50% de divergência na sequência de aminoácidos (ARBIZA et al., 2006). Alguns estudos mostraram que vários genótipos diferentes cocirculam e alguns predominam em uma comunidade a cada ano (OSTLUND et al., 2008; PERET et al., 1998).

A glicoproteína G é uma proteína transmembrânica integral do Tipo II (ou seja, está ancorada à membrana próximo à sua porção amino terminal por um domínio hidrofóbico que se estende do resíduo 38 ao 66), altamente glicosilada, com uma região localizada no citoplasma (N-terminal), uma região transmembrana e um grande ectodomínio C-terminal (GARCIA, O. et al., 1994). Esta proteína é sintetizada como um precursor de massa molecular de 36 kDa, sendo posteriormente glicosilada originando um complexo intermediário de 45 kDa. Esses açúcares são convertidos para tipos complexos produzindo uma forma madura de 90 kDa.

Sua variabilidade e glicosilação estão localizadas principalmente no ectodomínio, que contém dois segmentos hipervariáveis com alto teor de serina e treonina, entre os aminoácidos 69-164 e 207-298, com aproximadamente 56% de divergência entre os tipos A e B (JOHNSON *et al.*, 1987), separados por uma região altamente conservada (Figura 2) entre os aminoácidos 164 e 199, e quatro resíduos de cisteínas (173, 176, 182, 186). Esta sequência *motif* de 35 resíduos de aminoácidos é considerada como provável sítio de ligação ao receptor de quimiocina CX3CR1 da célula hospedeira (JOHNSON *et al.*, 1987; CANE *et al.*, 1991; SULLENDER *et al.*, 2000). A proteína G interage com o receptor celular CX3CR1, modulando assim a resposta imunológica do hospedeiro e facilitando a infecção (COLLINS *et al.*, 2008). Essa interação do receptor CX3CR1 com a proteína G é um contribuidor importante

para a infecção do hRSV em células epiteliais das vias aéreas dos humanos (CHIRKOVA, 2015).

**FIGURA 2:** Representação esquemática dos domínios da Glicoproteína G do hRSV.



Fonte: Adaptada de BATTLES, 2019.

O receptor celular CX3CR1 é encontrado em 4 formas parecidas em suas sequências, denominadas de isoformas, pois essas apresentam sequências de resíduos de aminoácidos semelhantes. A isoforma 1 é constituída de 355 resíduos de aminoácidos, enquanto as isoformas 2 e 4 são constituídas de 387 resíduos, visto que a isoforma 4 apresenta uma fenilalanina no resíduo 8 enquanto a isoforma 2 tem uma leucina na mesma posição, e a isoforma 3 é composta de 362 resíduos. A principal diferença encontrada nas isoformas está na constituição dos resíduos de aminoácidos na região N-terminal, enquanto a parte central e a região C-terminal é idêntica para todas as isoformas (Figura 3). A diferença na região do N-terminal é a extensão da sequência primária que para as isoformas 2 e 4 têm 32 resíduos a mais com relação a isoforma 1, sendo que a diferença entre a isoforma 2 e 4 é a mudança de um resíduo de lisina para fenilalanina para a isoforma 4, e a isoforma 3 têm 7 resíduos a mais do que a isoforma 1. Em um estudo realizado por Burg e colaboradores (2015) mostrou que uma parte da região N-terminal do US28, um homólogo do receptor celular

CX3CR1 (38% de identidade), interage com a fractalcina (CX3CL1), o ligante nativo do CX3CR1 que apresenta semelhante estrutural com a região conservadas de cisteínas do ectodomínio da proteína G. Esse resultado indica que a região N-terminal das isoformas pode ter um papel importante para a interação com a região conservada de cisteínas do ectodomínio da proteína responsável pelo processo de adesão do hRSV.

**FIGURA 3:** Sequências dos resíduos de aminoácidos das isoformas 1 (preto), 2 (vermelho), 3 (azul) e 4 (verde).

```

isoforma1      -----MDQFPESVTENFEYDDLAEACYIGDIVV
isoforma2      MREPLEALKLADLDFRKSSLASGWRMASGAFTMDQFPESVTENFEYDDLAEACYIGDIVV
isoforma3      -----MASGAFTMDQFPESVTENFEYDDLAEACYIGDIVV
isoforma4      MREPLEAFKLADLDFRKSSLASGWRMASGAFTMDQFPESVTENFEYDDLAEACYIGDIVV
                *****

Isoforma1      FGTVFLSIFYSVIFAIGLVGNLLVVFALTNSKKPKSVTDIYLLNLALSDDLFFVATLPFWT
Isoforma2      FGTVFLSIFYSVIFAIGLVGNLLVVFALTNSKKPKSVTDIYLLNLALSDDLFFVATLPFWT
Isoforma3      FGTVFLSIFYSVIFAIGLVGNLLVVFALTNSKKPKSVTDIYLLNLALSDDLFFVATLPFWT
Isoforma4      FGTVFLSIFYSVIFAIGLVGNLLVVFALTNSKKPKSVTDIYLLNLALSDDLFFVATLPFWT
                *****

Isoforma1      HYLINEKGLHNAMCKFTTAAFFIGFFGSIFFITVISIDRYLAIVLAANSMMNRTVQHGV
Isoforma2      HYLINEKGLHNAMCKFTTAAFFIGFFGSIFFITVISIDRYLAIVLAANSMMNRTVQHGV
Isoforma3      HYLINEKGLHNAMCKFTTAAFFIGFFGSIFFITVISIDRYLAIVLAANSMMNRTVQHGV
Isoforma4      HYLINEKGLHNAMCKFTTAAFFIGFFGSIFFITVISIDRYLAIVLAANSMMNRTVQHGV
                *****

Isoforma1      ISLGVWAAAILVAAPQFMFTKQKENECLGDYPEVLQEIWPVLRNVEITNFGFLLPLLIMS
Isoforma2      ISLGVWAAAILVAAPQFMFTKQKENECLGDYPEVLQEIWPVLRNVEITNFGFLLPLLIMS
Isoforma3      ISLGVWAAAILVAAPQFMFTKQKENECLGDYPEVLQEIWPVLRNVEITNFGFLLPLLIMS
Isoforma4      ISLGVWAAAILVAAPQFMFTKQKENECLGDYPEVLQEIWPVLRNVEITNFGFLLPLLIMS
                *****

Isoforma1      YCYFRIIQTLFSCKNHKKAKAIKLILLVVIVFFLFWTPYVNMIFLETCLKLYDFFPSCDMR
Isoforma2      YCYFRIIQTLFSCKNHKKAKAIKLILLVVIVFFLFWTPYVNMIFLETCLKLYDFFPSCDMR
Isoforma3      YCYFRIIQTLFSCKNHKKAKAIKLILLVVIVFFLFWTPYVNMIFLETCLKLYDFFPSCDMR
Isoforma4      YCYFRIIQTLFSCKNHKKAKAIKLILLVVIVFFLFWTPYVNMIFLETCLKLYDFFPSCDMR
                *****

Isoforma1      KDLRLALSVTETVAFSHCLNPLIYAFAGEKFRRYLYHLYGKCLAVLCGRSVHVDSSSE
Isoforma2      KDLRLALSVTETVAFSHCLNPLIYAFAGEKFRRYLYHLYGKCLAVLCGRSVHVDSSSE
Isoforma3      KDLRLALSVTETVAFSHCLNPLIYAFAGEKFRRYLYHLYGKCLAVLCGRSVHVDSSSE
Isoforma4      KDLRLALSVTETVAFSHCLNPLIYAFAGEKFRRYLYHLYGKCLAVLCGRSVHVDSSSE
                *****

Isoforma1      SQRSRHGSVLSSNFTYHTSDGDALLLL
Isoforma2      SQRSRHGSVLSSNFTYHTSDGDALLLL
Isoforma3      SQRSRHGSVLSSNFTYHTSDGDALLLL
Isoforma4      SQRSRHGSVLSSNFTYHTSDGDALLLL
                *****

```

Fonte: Elaborada pelo Autor.

Com relação ao contexto apresentado, o presente trabalho faz uso de ferramentas computacionais para buscar compreender aspectos moleculares da

interação entre as isoformas do receptor celular CX3CR1 com a proteína G do hRSV, responsável pela adesão viral à célula hospedeira e aqui mimetizada por sua região de cisteínas conservadas, sendo nomeada neste trabalho como peptídeo da G. Vale destacar que não existe na literatura qualquer estudo experimental que detalhe molecularmente essa interação CX3CR1/G (peptídeo da G). Portanto, o presente estudo proporciona, além de importantes informações de biologia estrutural computacional para o entendimento do processo de interação vírus/célula hospedeira, uma base molecular útil para nortear um futuro desenvolvimento de novos compostos com ação antiviral que tenham como alvo o bloqueio da interação CX3CR1/G fundamental para o processo de infecção viral pelo hRSV.

### 1.1 Dinâmica Molecular

A simulação de dinâmica molecular (DM) é uma das principais ferramentas computacionais para estudo de sistemas atômico-moleculares. É uma técnica computacional em que se determinam os movimentos das partículas de qualquer sistema, do qual se conhecem o potencial de interação entre essas partículas e as equações que regem seu movimento. A DM permite estudar a evolução temporal das configurações dos constituintes do sistema, e a partir das sequências de posições geradas, determinar as propriedades macroscópicas do sistema. Neste método, partículas interagentes, inicialmente dispostas em uma determinada configuração, movimentam-se sob a influência de potenciais intermoleculares (COUTINHO *et al.* 2007).

Com base na mecânica molecular, as moléculas são tratadas como uma coleção de átomos que pode ser descrita por forças newtonianas, ou seja, são tratadas como uma coleção de partículas mantidas unidas por forças harmônicas ou elásticas. Um conjunto completo dos potenciais de interação entre as partículas é referido como “campo de força” (BROOKS *et al.*, 1998. GUNSTEREN *et al.*, 1992).

Para utilizar este método é preciso utilizar um determinado campo de força. Para o presente trabalho, o campo de força utilizado foi o GROMOS54A7 do pacote de programas GROMACS (GUNSTEREN *et al.*, 1987), que para o estudo de

macromoléculas por dinâmica molecular se mostrou satisfatório (COUTINHO *et al.*2007).

A escolha dos potenciais de interação intramolecular ( $V_{intra}$ ) e intermolecular ( $V_{inter}$ ) constitui uma etapa essencial para a descrição correta do sistema em estudo, já que são esses potenciais que determinam as forças que atuam em cada partícula, e, conseqüentemente, determinarão como o sistema irá evoluir no tempo. Para descrever as interações em sistemas moleculares, assume-se que o potencial entre duas moléculas possa ser descrito como uma soma de interações entre cada par de átomos ou sítios de interação, e que estas interações dependem apenas da distância que separa esses sítios. Desse modo,

$$V_{total} = \sum V_{inter} + \sum V_{intra} \quad (1)$$

Um modelo de potencial intermolecular entre moléculas  $i$  e  $j$  consiste na soma das interações entre cada par de átomos dessas moléculas, as quais são dadas por um potencial de Lennard-Jones (LJ) acrescido do termo correspondente às interações eletrostáticas entre cada par de sítio  $a$  e  $b$  pertencentes às moléculas  $i$  e  $j$ , respectivamente,

$$V_{ij} = \sum_{a \in i} \sum_{b \in j} V_{ab}(r_{ab}) \quad (2)$$

$$V_{ab}(r_{ab}) = 4\varepsilon_{ab} \left[ \left( \frac{\sigma_{ab}}{r_{ab}} \right)^{12} - \left( \frac{\sigma_{ab}}{r_{ab}} \right)^6 \right] + \frac{q_a q_b}{4\pi\epsilon_0 r_{ab}} \quad (3)$$

sendo que  $q_a$  e  $q_b$  correspondem às cargas parciais,  $\varepsilon_{ab}$  e  $\sigma_{ab}$  são os parâmetros de energia e diâmetro de Lennard-Jones do sítio  $a$ , e  $r_{ab}$  é a separação entre os sítios  $a$  e  $b$  de moléculas distintas. No potencial de (3), o primeiro termo  $(r_{ab})^{-12}$  descreve a repulsão a curtas distâncias e o segundo termo corresponde à contribuição dispersiva de London, e o terceiro termo às interações eletrostáticas.

As interações intramoleculares, também conhecidos por potenciais ligados, descrevem as deformações moleculares. Dessa forma,

$$V_{intra} = \sum V_{ligação} + \sum V_{angular} + \sum V_{torção} \quad (4)$$

em que  $V_{ligação}$  corresponde aos estiramentos de uma ligação química e  $V_{angular}$  e  $V_{torção}$  correspondem, respectivamente, às descrições das deformações angulares e das deformações dos ângulos diedros. Os termos de estiramento e deformação angular são descritos por potenciais harmônicos como:

$$V_{ligação}(r) = \frac{1}{2}K_r(r - r_0)^2 \quad V_{angular}(\theta) = \frac{1}{2}K_\theta(\theta - \theta_0)^2 \quad (5)$$

no qual  $r_0$  e  $\theta_0$  correspondem ao comprimento e ângulo de equilíbrio da ligação, respectivamente, e  $K_r$  e  $K_\theta$  as constantes de força. O potencial de torção é dado por:

$$V_{torção}(\phi) = \sum_n \frac{1}{2}V_n [1 + \cos(n\phi - \delta)] \quad (6)$$

sendo que  $\phi$  é o valor do ângulo diedro,  $V_n$  é a amplitude de cada função e  $\delta$  é a fase.

Assim, o potencial total anteriormente apresentado na Equação 1 pode ser descrito como.

$$V_{total} = V_{ligado} + V_{não-ligado} \quad (7)$$

Partindo dos funcionais apresentados, a expressão para a determinação do potencial total sobre cada uma das moléculas do sistema em estudo, assume a seguinte (COUTINHO, 2007):

$$V_{ligado} = \sum_{ligações} \frac{1}{2}K_r(r - r_0)^2 + \sum_{ângulos} \frac{1}{2}K_\theta(\theta - \theta_0)^2 + \sum_{n,diedros} \frac{1}{2}V_n [1 + \cos(n\phi - \delta)] \quad (8)$$

$$V_{não-ligado} = \sum_{i,j} 4\varepsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \quad (9)$$

$$V_{total} = \sum_{ligações} \frac{1}{2}K_r(r - r_0)^2 + \sum_{ângulos} \frac{1}{2}K_\theta(\theta - \theta_0)^2 + \sum_{n,diedros} \frac{1}{2}V_n [1 + \cos(n\phi - \delta)] + \sum_{i,j} 4\varepsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \quad (10)$$

## 1.2 Docking Molecular

O *docking* molecular é um método de modelagem molecular que prevê a orientação preferencial de um ligante, normalmente uma pequena molécula, na estrutura de uma macromolécula que na maioria das vezes se trata de uma proteína. O conhecimento da orientação preferencial de uma pequena molécula no sítio de ligação em uma proteína pode ser utilizado para prever as interações não covalente envolvidas na associação entre essas duas moléculas.

A utilização dos cálculos de *docking* molecular desempenha um papel importante na ciência atualmente, sendo utilizado por exemplo para a triagem virtual de grandes bibliotecas de compostos contra proteína de interesse e para propor hipóteses estruturais de como um ligante se liga ao alvo proteico. Esses cálculos aumentam a probabilidade de encontrar novos inibidores de proteínas-chave de patógenos, prevendo quão bom um composto pode ligar-se ao alvo, onde ele pode interagir e que tipos de interações não covalentes poderia formar no local de ligação (MORRIS et al., 1998; MORGON et al., 2007).

Para o presente trabalho foi escolhido servidores *on-line* de *docking*, entre eles destacamos ClusPro e HADDOCK, os quais realizam *docking* proteína-proteína, visto que o sistema estudado neste trabalho é o receptor celular CX3CR1 e um peptídeo que corresponde a região de cisteínas conservadas da proteína G do hRSV. Dessa forma, empregar os programas de *docking* proteína-proteína se torna relevante para o presente trabalho.

O ClusPro é um servidor *on-line* para modelagem molecular do encaixe direto de duas proteínas que interagem. Ele foi criado em 2004, mas desde então foi substancialmente modificado e expandido. O servidor executa três etapas computacionais da seguinte forma: (1) ancoragem de corpo rígido por amostragem de bilhões de conformações; (2) agrupamento baseado em desvio quadrático médio (RMSD) das 1.000 estruturas de menor energia geradas para encontrar os maiores clusters que representarão os modelos mais prováveis do complexo; e (3) refinamento das estruturas selecionadas usando minimização de energia. A etapa de ancoragem de corpo rígido usa o PIPER, um programa de ancoragem baseado na abordagem de correlação *Fast Fourier Transform* (FFT). A abordagem FFT, introduzida por

Katchalski-Katzir e colaboradores em 1992, levou a um grande progresso no *docking* proteína-proteína de corpo rígido. Neste método, uma das proteínas (nomeada receptor) é colocada na origem do sistema de coordenadas em uma grade fixa, a segunda proteína (nomeada ligante) é colocada em uma grade móvel e a energia de interação é escrita na forma de uma função de correlação (ou como uma soma de algumas funções de correlação) (KOZAKOV, 2017).

A eficiência numérica do método decorre do fato de que tais funções de energia podem ser calculadas eficientemente usando FFT, e resulta na capacidade de amostragem exaustiva de bilhões de conformações das duas proteínas que interagem, avaliando as energias em cada ponto da grade. Assim, o algoritmo baseado em FFT permite o encaixe de proteínas sem qualquer informação a priori sobre a estrutura do complexo. Katchalski-Katzir e coautores (1992) usaram uma função de pontuação simples que levava em conta apenas a complementaridade da forma. No entanto, métodos subsequentes baseados na abordagem de correlação FFT para ancoragem introduziram funções de pontuação mais complexas e mais precisas que também incluíam termos representando interações eletrostáticas, ou tanto eletrostáticas quanto contribuições de solvatação. Uma das características distintivas do PIPER, o programa de *docking* usado na versão atual do ClusPro, é que essa implementação do método de correlação FFT emprega uma função de pontuação, incluindo um termo de interação pareado baseado em estrutura e a combinação com os outros termos na função de energia, que aumenta substancialmente a precisão do encaixe e resulta em estruturas mais próximas da nativa (KOZAKOV, 2006).

Como mencionado anteriormente, o servidor ClusPro é baseado no PIPER que realiza a amostragem. O centro de massa do receptor é fixado na origem do sistema de coordenadas e as possíveis posições rotacionais e translacionais do ligante são avaliadas em um dado nível de discretização. O espaço rotacional é amostrado em uma grade baseada em esferas que define uma subdivisão de uma superfície esférica na qual cada pixel cobre a mesma área de superfície que todos os outros pixels. As 70.000 rotações que se considera correspondem a cerca de 5 graus em termos dos ângulos de Euler. O tamanho do passo da grade translacional é 1 Å (YERSHOVA, 2010).

O PIPER representa a interação de energia entre duas proteínas usando uma expressão da forma

$$E = w_1 E_{rep} + w_2 E_{attr} + w_3 E_{elec} + w_4 E_{DARS} \quad (11)$$

sendo que  $E_{rep}$  e  $E_{attr}$  representam as contribuições de energia repulsiva e atrativa, respectivamente, para as interações de energia de van der Waals e  $E_{elec}$  é a contribuição de energia eletrostática. O  $E_{DARS}$  é um potencial baseado em estrutura pareado construído pela abordagem *Decoys* como o Estado de Referência (DARS) e representa principalmente contribuições de solvatação, ou seja, a mudança de energia livre devido à remoção das moléculas de água da interface. Os coeficientes  $w_1$ ,  $w_2$ ,  $w_3$  e  $w_4$  definem os pesos dos termos correspondentes e são selecionados de forma otimizada para diferentes tipos de problemas de *docking* (ZOARKOV, 2017).

O servidor HADDOCK faz uso de dados experimentais bioquímicos e biofísicos na forma de informações sobre a região de interface entre os componentes moleculares e/ou suas orientações relativas (DOMINGUEZ, 2003). Essas informações podem ser derivadas, por exemplo, de mutagênese, espectrometria de massa ou uma variedade de experimentos de Ressonância Magnética Nuclear (RMN) (perturbação de deslocamento químico (CSP), acoplamentos dipolares residuais (RDCs), troca de hidrogênio/deutério ou, se disponível, restrições de distância). Quando a informação experimental é escassa ou ausentes, as previsões de interface bioinformática também podem ser utilizadas (de VRIES, 2005,2006,2008).

Para orientar os processos de *docking* são introduzidas restrições de interação ambígua (*Ambiguous Interaction Restraints – AIRs*), os AIRs são definidos pelo HADDOCK com base nos resíduos ativos e passivos definidos pelo usuário, para cada resíduo ativo, uma única restrição de AIR é definida entre esse resíduo e todos os resíduos ativos e passivos na molécula parceira, dessa forma para as interações ambíguas o usuário determina os resíduos ativos e passivos, assim o *server* tem um guia para realizar o *docking* molecular.

Dessa forma, após o cálculo, as estruturas são classificadas de acordo com sua energia intermolecular, que envolvem energia eletrostática ( $E_{elec}$ ), energia de van der Waals ( $E_{VDW}$ ), energia de solvatação ( $E_{solv}$ ) e as AIRs ( $E_{AIRs}$ ), além do valor para

área de superfície aterrada ( $E_{BSA}$ ). Portanto, os valores energéticos são determinados através de uma combinação linear que forma o HADDOCK score (HS) (12).

$$HS = 0,01E_{VDW} + 0,1E_{elec} + 1,0E_{solv} - 0,01E_{BSA} + 0,01E_{AIRs} \quad (12)$$

As AIRs, são provenientes de qualquer informação experimental existente sobre resíduos que estão envolvidos na interação intermolecular. Permitindo que o HADDOCK pesquise todas as conformações possíveis em torno do local de interação definido pelos dados e encontrar o par mais favorável de aminoácidos interagindo entre os resíduos ativos e passivos (DOMINGUEZ et al, 2003).

O protocolo do HADDOCK requer os arquivos PDB das proteínas individuais e as AIRs, sendo dividida em três partes: (1) randomização de orientações e minimização de energia; (2) alinhamento semirrígido simulado no espaço de ângulo e torção; e (3) refinamento final com o solvente explícito. Por fim, as estruturas são agrupadas em *clusters* usando RMSD para a interface de interação e, em seguida, são analisadas essas estruturas são classificadas de acordo com as suas energias média de interação e sua área de superfície média de interação (DOMINGUEZ et al, 2003).

### 1.3 Cálculo de Energia Livre de Ligação

O processo de reconhecimento molecular entre proteína e ligante é dirigido através de efeitos entálpicos e entrópicos. Esses efeitos podem ser estimados através da energia livre de ligação de Gibbs  $\Delta G_{lig}$ , que está relacionada diretamente com a constante de ligação  $K_{lig}$ . Desta forma, a energia livre de ligação pode se calculada da seguinte forma.

$$\Delta G_{lig} = \Delta H - T\Delta S = -RT \ln K_{lig} \quad (13)$$

sendo que  $\Delta H$  é a variação de entalpia,  $T$  é a temperatura absoluta,  $\Delta S$  é a variação de entropia e  $R$  é a constante universal dos gases perfeitos (COUTINHO, 2014).

A força de interação biomolecular, como a envolvida em reconhecimento ou catalíse, pode ser quantificada em termos de sua energia livre de ligação. Existem abordagens computacionais que podem ser utilizada para estimar as energias livre de ligação e uma delas é o método MM-PBSA, que se trata de dinâmica molecular combinada com uma análise de área superficial via Poisson-Boltzmann. O MM-PBSA é um método para estimar energias livres de ligação entre proteínas e ligantes (pequenas moléculas, peptídeos e ácidos nucleicos), que tem sido cada vez mais utilizado no estudo de interações biomoleculares. Em termos gerais, a energia livre de ligação é caracterizada como uma soma das seguintes energias livre (KUMARI *et al*, 2014)

$$\Delta G_{\text{ligação}} = G_{\text{complexo}} - (G_{\text{proteína}} + G_{\text{ligante}}) \quad (14)$$

sendo que  $G_{\text{complexo}}$  é a energia livre total do complexo proteína-ligante e  $G_{\text{proteína}}$  e  $G_{\text{ligante}}$  são energias livres totais do proteína isolada e ligante em solvente, respectivamente.

O método MM-PBSA combina três termos energéticos para explicar a energia livre de ligação. O primeiro termo corresponde a uma mudança na energia potencial no vácuo, o qual inclui termos energéticos de pares ligados como estiramento, ângulo e diedro de ligação, bem como termos de pares não ligados como potencial de van der Waals e interações eletrostáticas. O segundo termo é responsável pela dissolução das diferentes espécies do sistema molecular, o qual é quantificado pela soma de dois termos energéticos, um polar e outro apolar de energia de solvatação usando um modelo implícito (GILSON *et al*, 1988). Por fim, o terceiro termo é responsável pela entropia configuracional associada à formação de complexo na fase gasosa (KUMARI *et al*, 2014).

A abordagem MM-PBSA cresceu e se tornou um dos métodos mais amplamente usados para calcular energias de interação e muitas vezes é empregada para estudar complexos biomoleculares (HOMEYER *et al*, 2012). Este método é comumente combinado com simulações de DM, incorporando flutuações conformacionais e contribuições entrópicas para a energia de ligação.

## 2 OBJETIVOS GERAIS E ESPECÍFICOS

O objetivo principal do presente trabalho foi caracterizar através de simulações computacionais, tais como modelagem molecular, *docking* molecular, dinâmica molecular e cálculos de energia livre de ligação teórica, a interação da região conservada de cisteínas do ectodomínio da proteína G do hRSV com as isoformas do receptor celular CX3CR1 do hospedeiro. A caracterização computacional apresentada neste trabalho representa o detalhamento molecular da interação da proteína G com seu receptor celular e, portanto, pode proporcionar informações importantes para o desenvolvimento de inibidores moleculares da ligação CX3CR1/G, dificultando a adesão do vírus a célula hospedeira. Com relação ao contexto apresentado, os seguintes objetivos específicos foram delineados:

- 1) Realizar modelagem molecular da estrutura tridimensional das isoformas 1, 2, 3 e 4 do receptor celular CX3CR1 através de diferentes programas de modelagem estrutural de proteína (I-Tasser, Phyre2, Rosetta Fold e RosettaTR);
- 2) Realizar simulações de dinâmica molecular em bicamada lipídica de POPC e DPPC das estruturas modeladas para as isoformas do CX3CR1;
- 3) Realizar cálculos de *docking* molecular do peptídeo da região conservada de cisteínas da proteína G do hRSV (peptídeo da G) com as isoformas do receptor CX3CR1;
- 4) Realizar as simulações de dinâmica molecular dos modelos estruturais dos complexos formados entre o peptídeo da G e as isoformas do CX3CR1 em bicamada lipídica de POPC;
- 5) Realizar cálculos de energia livre de ligação entre o peptídeo da G e as isoformas do CX3CR1, analisando os resíduos de aminoácidos que mais contribuem para a formação dos complexos.

### 3 METODOLOGIA

#### 3.1 Modelagem Molecular

Há um grande interesse nos estudos estruturais de proteínas, pois com detalhamento atômico, pode-se ter uma maior compreensão sobre seus sítios ativos ou sítios de ligação, facilitando o desenvolvimento de fármacos. Atualmente, técnicas experimentais como a difração de Raios-X e a espectroscopia de Ressonância Magnética Nuclear (RMN) são ferramentas utilizadas pelos pesquisadores para determinar a estrutura tridimensional de proteínas. Contudo, essas técnicas experimentais demandam alto custos tanto de recursos quanto de tempo de trabalho. Dessa forma, os cálculos de modelagem molecular podem ser empregados como uma abordagem alternativa para a obtenção de resultados mais rápidos e de menor custo na determinação das estruturas tridimensionais de proteínas.

No presente trabalho, os cálculos de modelagem molecular foram realizados para determinar computacionalmente a estrutura tridimensional do receptor celular CX3CR1, uma vez que ainda não há nenhum dado estrutural experimental. Para isso, a sequência de resíduos de aminoácidos das quatro isoformas do CX3CR1 foram obtidas do banco de dados de UNIPROT sob o código de acesso P49238. Em seguida, essas sequências de resíduos foram utilizadas como dado de entrada nos servidores de modelagem molecular I-Tasser (Zhang *et al.*, 2015), Phyre2 (Kelley *et al.*, 2015), Rosetta FOLD (Baek *et al.*, 2021) e Rosetta TR (Yang *et al.*, 2020).

O servidor I-Tasser (Zhang *et al.*, 2015) utiliza as informações da estrutura como referência, de forma semelhante ao *threading* este método ajusta a conformação da cadeia principal da sequência alvo à uma coleção de proteínas de referência. Assim é gerado um valor de ajuste, que mede a qualidade do modelo dado e isso é calculado para cada modelo. Este método de *threading* é utilizado pois acredita-se que na natureza há um número finitos de enovelamento, então as proteínas são ajustadas dependendo dessa biblioteca de enovelamentos que já tenha sido formada e mostrada por meio do uso de resultados de difração de Raios X e RMN. Depois de ter as estruturas por meio deste método de *threading*, as estruturas se formam em grupos maiores passam pelo método de Monte Carlo para terem um maior refinamento e dessa forma obter uma estrutura que pode ser mais apropriada (Sabbag, 2012).

O método usado pelo Phyre2 (Kelley *et al*, 2015) constrói os modelos através de metodologias de predição de estruturas utilizando o enovelamento. Há uma construção do perfil do modelo alvo, esse perfil é composto através de alinhamento de outras proteínas de estruturas já determinadas e conhecidas. Logo após este feito, o perfil é submetido a um modelo de Markov Oculto - *Hidden Markov Model* (HMM) e este modelo é comparado contra um banco de dados de HMM de proteínas com estruturas conhecidas através do algoritmo de alinhamento *HHsearch* (Söding, 2005). Dessa forma, é realizada a construção dos modelos com base nos fragmentos do molde selecionado e ao final é inserido as cadeias laterais (Kelley *et al*, 2015).

O método Rosetta TR ou denominado Rosetta Restrito por Transformação utiliza como componentes uma rede convulacional residual profunda (Yang *et al*, 2020). Essa rede é caracterizada por um algoritmo de aprendizagem profunda que capta uma imagem de entrada e atribui importâncias a ela, além de diferenciar alguns aspectos entre elas. Assim o método Rosetta TR utiliza informações como distância relativas e orientações de todos os pares de resíduos da proteína, levando em consideração outras distâncias como  $C_{\beta} - C_{\beta}$  e a procura por orientações entre os resíduos. Após obter essas informações, o método passa para uma nova etapa na qual utiliza restrições do tipo Rosetta (*ab initio*) e minimização diretas de energias, usando também restrições de distância entre resíduos e restrições de orientações (Yang *et al*, 2020). Dessa forma, é um algoritmo que insere restrições e faz buscas no PDB para se determinar a melhor estrutura para a modelagem molecular.

O programa Rosetta é um programa de predição de estruturas de proteínas totalmente automatizado usando o método de inserção de fragmentos para a modelagem proteica. A biblioteca de fragmentos é composta por estruturas determinadas experimentalmente. Assim, o programa funciona para orientar a busca do espaço conformacional e as previsões das estruturas. Para o Rosetta FOLD é utilizado um método de aprendizagem oculta ou aprendizagem profunda, muito parecido com o método de aprendizagem de máquina, no qual a máquina aprende a fazer uma busca conformacional para se determinar a melhor estrutura. O método de aprendizagem profunda trata-se de Redes Neurais Artificiais, uma área que busca simular computacionalmente o cérebro enquanto máquina, ou seja, inteligência artificial. O Rosetta FOLD utiliza este princípio usando uma rede neural de três vias,

essas vias são divididas em (i) leitura dos resíduos de aminoácidos da proteína (ii) determinação entre as distâncias de pares de resíduos entre carbonos betas ( $C_{\beta} - C_{\beta}$ ), a qual obtém uma rede de probabilidade entre os resíduos e (iii) otimização dos dados através de cálculos de RMSD. Dessa forma, obtém uma otimização mais acurada (BAEK *et al*, 2020), calculando, portanto, uma estrutura mais otimizada. No último passo da modelagem molecular, as estruturas são levadas para um algoritmo de Monte Carlo e as melhores estruturas são selecionadas.

### 3.2 Dinâmica Molecular em bicamada lipídica POPC e DPPC

Para a simulação de dinâmica molecular (DM) em bicamada lipídica de POPC e DPPC, foi utilizado o campo de força GROMOS 54A7 (SCHMID *et al*, 2011) com modificação para inclusão dos parâmetros de topologia para lipídeos (BERGER *et al*, 1997) e o modelo de água TIP3P (JORGENSEN *et al*, 1983). Todas as simulações foram realizadas no Supercomputador Santos Dumont do Laboratório Nacional de Computação Científica (LNCC) no Rio de Janeiro. Utilizando o servidor CHARMM-GUI (<https://www.charmm-gui.org>) foi construído a bicamada lipídica de POPC em uma caixa retangular contendo a proteína alinhada com o eixo Z da bicamada lipídica, sendo adicionada uma camada de moléculas de água de 22,5 Å com 50 moléculas de água por lipídio. A bicamada lipídica é constituída em média com 200 lipídios na parte superior e 200 lipídios na parte inferior, contabilizando um total de 400 lipídios. Duas condições salinas foram empregadas: na primeira, íons de sódio foram adicionados para promover a neutralização do sistema na caixa retangular; e na segunda, íons de sódio e cloro foram adicionados ao sistema de modo a conferir uma concentração de 150 mM de NaCl. O sistema foi devidamente minimizado e o arquivo de PDB resultante do servidor CHARMM-GUI foi utilizado para dar prosseguimento aos cálculos de DM com o pacote de programas GROMACS. O estado de protonação dos resíduos ionizáveis da proteína foram obtidas pelo programa PROPKA (OLSSON *et al*, 2011), considerando um pH de 7,0. Condições periódicas de contorno foram usadas para o *ensemble* NPT mantendo a temperatura constante em 298 K e a pressão em 1,0 bar. Para o potencial de Lennard-Jones foi usado um *cut-off* de 12 Å. Foi usado o mesmo *cut-off* para o potencial eletrostático, no qual usou o algoritmo

PME (Partículas de malhas Ewald). As simulações foram realizadas usando um passo de tempo de 2,0 fs salvando a trajetória a cada 20 ps e todas as ligações covalentes foram limitados à sua distância de equilíbrio. Um algoritmo de minimização de gradiente conjugado foi utilizado para relaxar a superposição de átomos gerados no processo de construção da caixa. Vinte e cinco milhões passos (50 ns) foram realizados para cada processo de equilibrção NPT do sistema, aplicando uma força constante de  $1000 \text{ kJ.mol}^{-1}.\text{nm}^{-2}$  para todos os átomos pesados do modelo estrutural da proteína em bicamada lipídica. O *script shrink.pl* foi empregado para calcular o valor de área por lipídios (Tabela 1 e 2). Para a bicamada lipídica de POPC, todos os passos anteriormente descritos foram aplicados para os quatro modelos estruturais da isoforma 1 do CX3CR1 determinados pelo Rosetta FOLD, Rosetta TR, I-tasser e Phyre2, assim como para os modelos das isoformas 2, 3 e 4 obtidos a partir do servidor Rosetta FOLD. Já para DPPC, todos os passos anteriores foram empregados para os quatro modelos estruturais das isoformas do CX3CR1 determinados pelo Rosetta FOLD. Ao final das etapas de preparação, simulações de DM de 300 ns foram realizadas dos modelos estruturais em bicamada lipídica de POPC e DPPC. Para avaliação das simulações, as trajetórias foram alinhadas e analisadas de acordo com: (i) desvio quadrático médio (RMSD) dos átomos da cadeia principal da proteína; (ii) ângulo de inclinação para do barril de hélices alfa da proteína usando o comando *bundle* do GROMACS; (iii) ângulo do eixo principal (comando *principal* do GROMACS) do barril de hélices alfa com relação ao eixo principal de partida no início da simulação; (iv) número de contatos menor do que 0,6 nm entre os átomos da fita beta GLU174–GLY177 (numeração referente a isoforma 1) na cavidade de ligação da proteína e os resíduos da região N-terminal de cada isoforma (1: resíduos 1-22, 2: resíduos 1-51, 3: resíduos 1-31, e 4: resíduos 1-51), sendo que a cavidade é importante para este trabalho pois é nessa região que o peptídeo da G se liga no receptor celular CX3CR1.

**TABELA 1:** Valores de área por lipídio em bicamada lipídica de POPC dos programas utilizados para modelagem molecular.

	Área por lipídio em x (Å <sup>2</sup> )	Área por lipídio em y (Å <sup>2</sup> )	Área por lipídio em z (Å <sup>2</sup> )
<i>RosettaFold</i>	63,6	65,6	65,5
<i>RosettaTR</i>	63,8	64,1	65,1
<i>I-Tasser</i>	63,4	63,7	65,0
<i>Phyre2</i>	63,5	64,1	64,2

Fonte: Elaborado pelo autor.

**TABELA 2:** Valores de área por lipídio das isoformas utilizando Rosetta FOLD para os modelos de bicamada lipídica em POPC e DPPC.

	POPC			DPPC		
	Área por lipídio em x (Å <sup>2</sup> )	Área por lipídio em y (Å <sup>2</sup> )	Área por lipídio em z (Å <sup>2</sup> )	Área por lipídio em x (Å <sup>2</sup> )	Área por lipídio em y (Å <sup>2</sup> )	Área por lipídio em z (Å <sup>2</sup> )
<i>Isoforma1</i>	63,6	65,6	65,5	59,9	61,5	62,0
<i>Isoforma2</i>	62,7	63,0	64,4	61,6	61,9	63,5
<i>Isoforma3</i>	61,8	63,8	63,5	60,3	61,9	61,5
<i>Isoforma4</i>	61,6	64,4	63,6	60,4	62,9	62,7

Fonte: Elaborado pelo autor.

### 3.3 Docking Molecular

Para os cálculos de *docking* molecular foram utilizados os seguintes servidores: ClusPro (<https://cluspro.org/>), HADDOCK (<https://wenmr.science.uu.nl/>), pyDock (<https://life.bsc.es/pid/pydockweb>), Zdock (<https://zdock.umassmed.edu/>), HDock (<http://hdock.phys.hust.edu.cn/>), GalaxyWeb (<http://galaxy.seoklab.org/>), HpepDock(<http://huanglab.phys.hust.edu.cn/>), MDockPeP (<https://zougrouptoolkit.missouri.edu/mdockpep/>) e FlexPepDock (<http://flexpepdock.furmanlab.cs.huji.ac.il/>). As estruturas representativas das isoformas do CX3CR1 utilizadas nos cálculos de *docking* foram obtidas a partir do

primeiro *cluster* de uma análise de clusterização realizada para a trajetória de 300 ns das proteínas em bicamada lipídica de POPC. A estrutura do peptídeo da região conservada de cisteínas da proteína G do hRSV foi adquirida a partir do banco de dados PDB (<https://www.rcsb.org>) sob o código de acesso 5WN9 (FEDECHKIN *et al*, 2017). Quando permitido pelo servidor, os resíduos CYS-176, SER-177, ASN-178, ASN-179 e PRO-180 do peptídeo da G e os da fita beta GLU-174–GLY-177 (numeração referente a isoforma 1) no CX3CR1 foram selecionados como pontos de atração para direcionar os cálculos de *docking* na cavidade de ligação da proteína. Para as isoformas 2 e 4 os resíduos da fita beta são GLU-206–GLY209 e para a 3 são GLU-181–GLY-184.

Os servidores ClusPro e HADDOCK apresentaram resultados mais apropriados tendo como referência uma comparação com a estrutura de cristal da fractalcina (CX3CL1) interagindo com receptor US28 homólogo de CX3CR1 (PDB 4XT1). O HADDOCK (*High Ambiguity Driven Protein-Protein Docking-docking* molecular orientado por alta ambiguidade) é um servidor que utiliza *docking* flexível orientado por informações para a modelagem de complexos biomoleculares. Este método diferencia dos métodos de *docking ab initio* pelo fato de codificar informações de interfaces de proteínas identificadas ou previstas de interações ambíguas e permite interações do tipo não ambíguas (resíduos ativos e passivos, respectivamente), tudo isso para conduzir o processo de ancoragem molecular (HONORATO *et al*, 2021 & ZUNDERT *et al*, 2016). Este método realmente é diferente dos demais, pois além de ser um método computacional, ele faz uso de informações experimentais obtidas por difração de Raios-X, restrições de interface definidas por RMN e dados de espalhamento de Raios-X a baixo ângulo. No final, o HADDOCK utiliza todos esses parâmetros para fazer o melhor encaixe, e seus resultados são colocados em *clusters* contendo as estruturas na ordem do menor valor de sua função de *score* (HADDOCK *score* (12)) para o maior, como o HAADDOCK utiliza uma área de superfície significativa para os cálculos de energia, então este *server* analisa os resultados, entregando o melhor modelo selecionado dado pela menor energia livre de ligação, ou seja, a conformação que apresentou maior afinidade com a proteína.

O método utilizado para o ClusPro é um método de *docking* rígido, utilizando a Transformada Rápida de Fourier (FFT) para a avaliação de energia. Nos métodos

FFT, uma das proteínas é colocada na origem do sistema de coordenadas em uma grade fixa e a segunda proteína em uma grade móvel. A energia de interação é escrita como uma soma de algumas funções de correlação (11). A eficiência mínima dos métodos decorre do fato de que tais funções de energia podem ser avaliadas simultaneamente para todas as traduções usadas das Transformadas Rápidas de Fourier e apenas as rotações precisam ser consideradas explicitamente. Isso resulta na capacidade de amostrar bilhões das conformações das duas proteínas em interação, obtendo valores de energia em cada parte da grade (DESTA *et al*, 2020). Obtendo esse número de conformações, o servidor utiliza um algoritmo em que se escolhe os melhores resultados pela energia de ligação colocando estes resultados em *clusters*. Assim o algoritmo baseado em FFT permite encaixe global sem qualquer tipo de informação *a priori* sobre a estrutura do complexo.

#### **3.4 Dinâmica Molecular dos complexo das isoformas do CX3CR1 com o peptídeo da proteína G**

Para a simulação de DM em bicamada lipídica de POPC dos complexos isoformas/peptídeo, foi utilizado o campo de força GROMOS 54A7 (SCHMID *et al*, 2011) com modificação para inclusão dos parâmetros de topologia para lipídeos (BERGER *et al*, 1997) e o modelo de água TIP3P (JORGENSEN *et al*, 1983). Dessa forma, retirou a proteína dos lipídios (*cluster*) e com, somente a proteína foi realizado os *dockings* descritos em 3.3, assim com o resultado dos *dockings* voltou o complexo CX3CR1/peptídeo da G para a bicamada lipídica e realizando alinhamento das estruturas, dessa forma correu as dinâmicas moleculares. Todas as simulações foram realizadas no Supercomputador Santos Dumont do Laboratório Nacional de Computação Científica (LNCC) no Rio de Janeiro. A bicamada lipídica é constituída em média com 200 lipídios na parte superior e 200 lipídios na parte inferior, contabilizando um total de 400 lipídios. Íons de sódio foram adicionados para promover a neutralização do sistema na caixa retangular. O estado de protonação dos resíduos ionizáveis da proteína foram obtidas pelo programa PROPKA (OLSSON *et al*, 2011), considerando um pH de 7,0. Condições periódicas de contorno foram usadas para o *ensemble* NPT mantendo a temperatura constante em 298 K e a pressão em 1,0 bar. Para o potencial de Lennard-Jones foi usado um *cut-off* de 12 Å.

Foi usado o mesmo *cut-off* para o potencial eletrostático, no qual usou o algoritmo PME (Partículas de malhas Ewald). As simulações foram realizadas usando um passo de tempo de 2,0 fs salvando a trajetória a cada 20 ps e todas as ligações covalentes foram limitados à sua distância de equilíbrio. Um algoritmo de minimização de gradiente conjugado foi utilizado para relaxar a superposição de átomos gerados no processo de construção da caixa. Vinte e cinco milhões passos (50 ns) foram realizados para cada processo de equilíbrio NPT do sistema, aplicando uma força constante de  $1000 \text{ kJ.mol}^{-1}.\text{nm}^{-2}$  para todos os átomos pesados modelo estrutural da proteína em bicamada lipídica. Todos os passos anteriormente descritos foram aplicados para os complexos das isoformas 1, 2, 3 e 4 do CX3CR1 com o peptídeo da G calculados pelos servidores de *docking* do ClusPro e do HADDOCK. Ao final das etapa de preparação, simulações de DM de 300 ns foram realizadas para os modelos estruturais dos complexos isoformas/peptídeo da G em bicamada lipídica de POPC. Para avaliação das simulações, as trajetórias foram alinhadas e analisadas de acordo com: (i) desvio quadrático médio (RMSD) dos átomos da cadeia principal das isoformas; (ii) RMSD da estrutura do peptídeo em relação ao alinhamento da estrutura da proteína ; (iii) RMSD da estrutura do peptídeo em relação ao alinhamento da estrutura do peptídeo; (iv) número de contatos menor do que 0,6 nm entre os átomos das isoformas do receptor celular e os átomos do peptídeo da G; (v) cálculos de distância entre o centro de massa do peptídeo da G em relação ao centro de massa das isoformas do receptor, para determinar se o peptídeo se distanciou ou não do sítio de ligação; (vi) quantificação das ligações de hidrogênio entre o receptor e o peptídeo, além do cálculo das porcentagens de persistência que essas ligações apresentaram ao longo dos 300 ns de simulação computacional; (vii) RMSF dos átomos da cadeia principal das isoformas do receptor livre (*backbone*) e ligado ao peptídeo (ClusPro e HADDOCK); e (viii) análise de componente principal (PCA) para os átomos da cadeia principal das isoformas do receptor livre (*backbone*) e ligado ao peptídeo (ClusPro e HADDOCK). Os gráficos de dispersão de PCA (PC1 & PC2) foram gerados para as isoformas livres e ligadas ao peptídeo, assim como movimentos conformacionais (30 *frames*) obtidos a partir dos autovetores da primeira e segunda componente principal (PC1 e PC2, respectivamente)

### 3.5 Cálculos de Variação de Energia Livre de Ligação Teórica

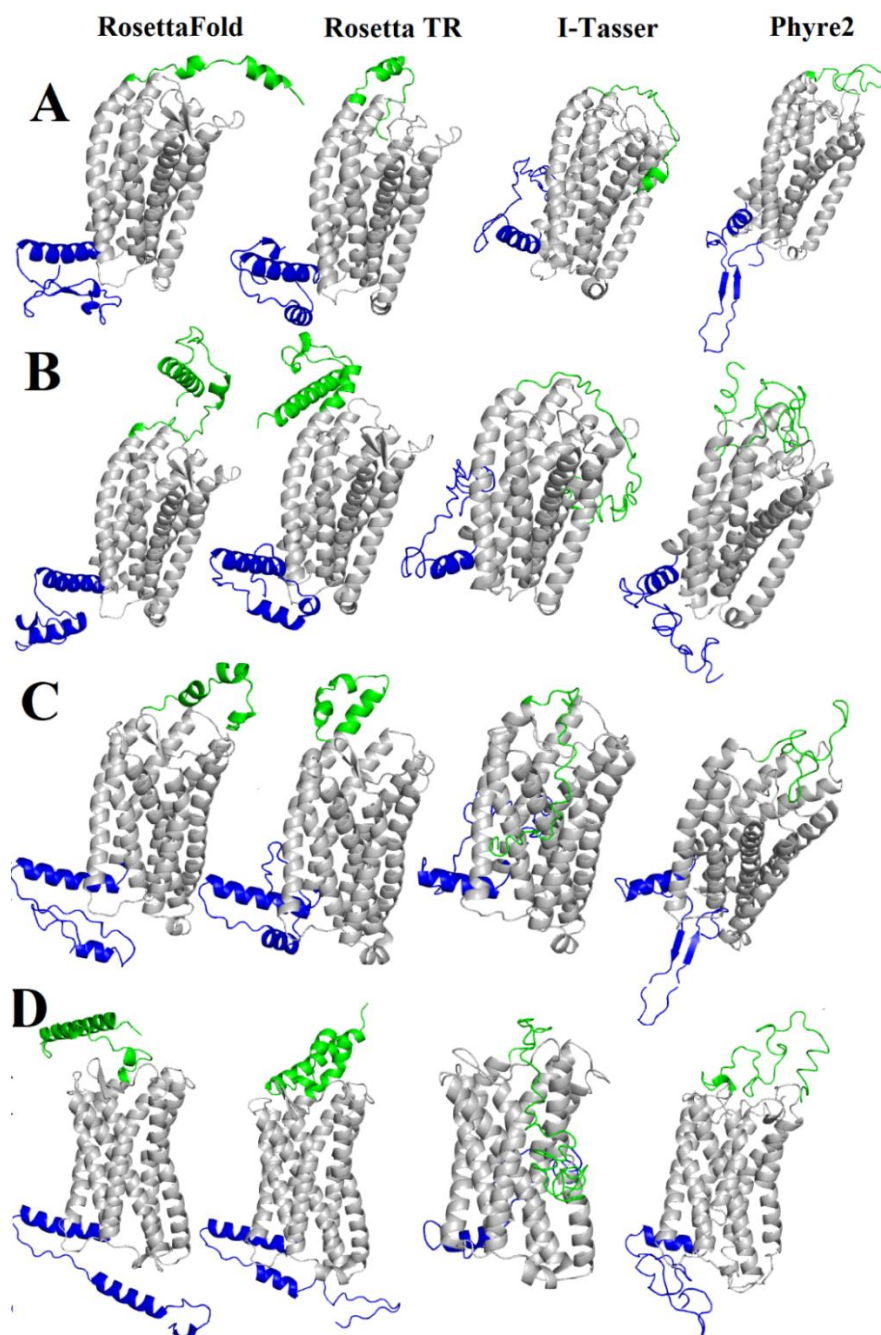
A variação de energia livre de ligação teórica ( $\Delta G_b$ ) entre as isoformas do receptor celular CX3CR1 e o peptídeo da G foi calculada a partir da trajetória de DM para o complexo em bicamada lipídica de POPC dos modelos estruturais obtidos pelo ClusPro e HADDOCK usando o método de Mecânica Molecular de Poisson-Boltzmann baseado na Área de Superfície (*Molecular Mechanic Poisson-Boltzmann Surface Area* – MM-PBSA) implementado no programa *g\_mmpbsa* para o pacote GROMACS, junto com os *scripts* *MmPbSaStat.py* e *MmPbSaDecomp.py* (Baker et al., 2001; Kumari et al., 2014). Uma análise *bootstrap* ( $n = 5000$ ) foi realizada obter erros padrão para as energias determinadas pela análise MM-PBSA. Foi utilizado um valor de 2 para a constante dielétrica do soluto (proteína), enquanto para a constante dielétrica do solvente foi empregado um valor de 80. Uma concentração de NaCl de 150 mM foi utilizada com uma temperatura de 25 °C (298 K). Para determinar os resíduos de aminoácidos que significativamente contribuem de forma favorável ou desfavorável para a energia livre de ligação total, um ponto de corte foi determinado como sendo a média mais um desvio padrão calculado a partir dos valores de energia de cada um dos resíduos.

## 4 RESULTADOS e DISCUSSÕES

### 4.1 Modelagem Molecular

A modelagem molecular foi feita utilizando os servidores I-Tasser (Zhang *et al.*, 2015), Phyre2 (Kelley LA *et al.*, 2015), Rosetta FOLD (Baek *et al.* 2021) e Rosetta TR (Yang *et al.*, 2020), inserindo os dados de sequência primária das isoformas 1, 2, 3 e 4 do CX3CR1 (<https://www.uniprot.org/uniprot/P49238>). A Figura 4 mostra os modelos estruturais gerados pelos servidores para o receptor celular e suas isoformas. É possível observar que uma configuração de barril de hélices alfa se mantém para todas as estruturas tridimensionais das isoformas preditas pelos servidores, diferenciando entre elas a conformação e posicionamento das regiões N e C terminais, as quais estão destacadas em verde e azul na Figura 4, respectivamente.

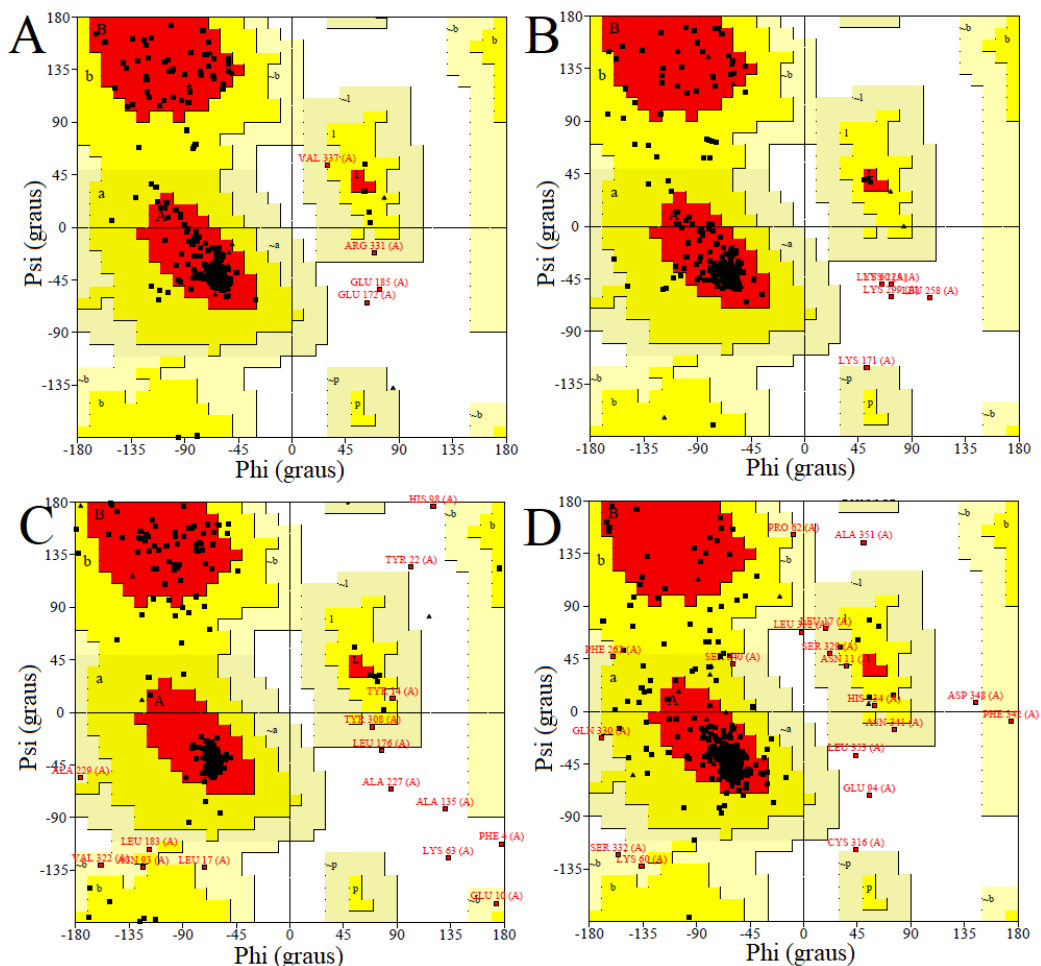
**FIGURA 4:** Modelos estruturais calculados pelos servidores de modelagem moleculuar RosettaFOLD, Rosetta TR, I-Tasser e Phyre2. Modelos estruturais do receptor celular CX3CR1: isoforma 1 (A), 2 (B), 3 (C) e 4 (D). As proteínas são apresentadas com o modelo *cartoon* utilizando o programa de vizualização PyMOL. O barril de hélices alfa é denotado em cinza e as regiões N e C-terminal são destacadas em verde e azul, respectivamente.



Fonte: Elaborado pelo autor.

Os gráficos de Ramachandran foram empregados para verificar a qualidade dos modelos estruturais gerados para CX3CR1 e suas isoformas 1, 2, 3 e 4. Esse tipo de gráfico apresenta uma relação entre os ângulos torcionais ( $\varphi$  – Phi e  $\Phi$  – Psi) da cadeia principal da proteína, os quais foram analisados para cada modelo estrutural calculado pelos servidores. Para o receptor celular CX3CR1 original que se refere a isoforma 1 (Figura 5) foi verificado que o modelo estrutural determinado pelo servidor Rosetta FOLD (Figura 5-A e Tabela 3) se mostrou mais apropriado apresentando valores de 92,7 % dos ângulos torcionais nas regiões favoráveis (vermelho), 6,1% nas regiões adicionais permitidas (amarelo escuro), 0,6% nas regiões generosamente adicionais (amarelo claro) e 0,6% nas regiões não permitidas (branco). Os valores referentes aos ângulos torcionais calculados pelos Rosetta TR, I-Tasser e Phyre2 são mostrados na Tabela 3. É possível observar que a porcentagem de ângulos torcionais em regiões não permitidas é maior para os demais servidores, o que pode caracterizar problemas com a determinação de seus respectivos modelos estruturais.

**FIGURA 5:** Gráficos de Ramachandran para ângulos torcionais dos resíduos dos modelos estruturais para o receptor celular CX3CR1 ou isoforma 1 determinados pelos servidores Rosetta FOLD (A), Rosetta TR (B), Phyre2 (C) e I-Tasser (D). Em vermelho resíduos em regiões favoráveis, amarelo escuro para regiões adicionais permitidas, em amarelo claro para regiões generosamente adicionais e em branco para regiões não permitidas.



Fonte: Elaborado pelo autor.

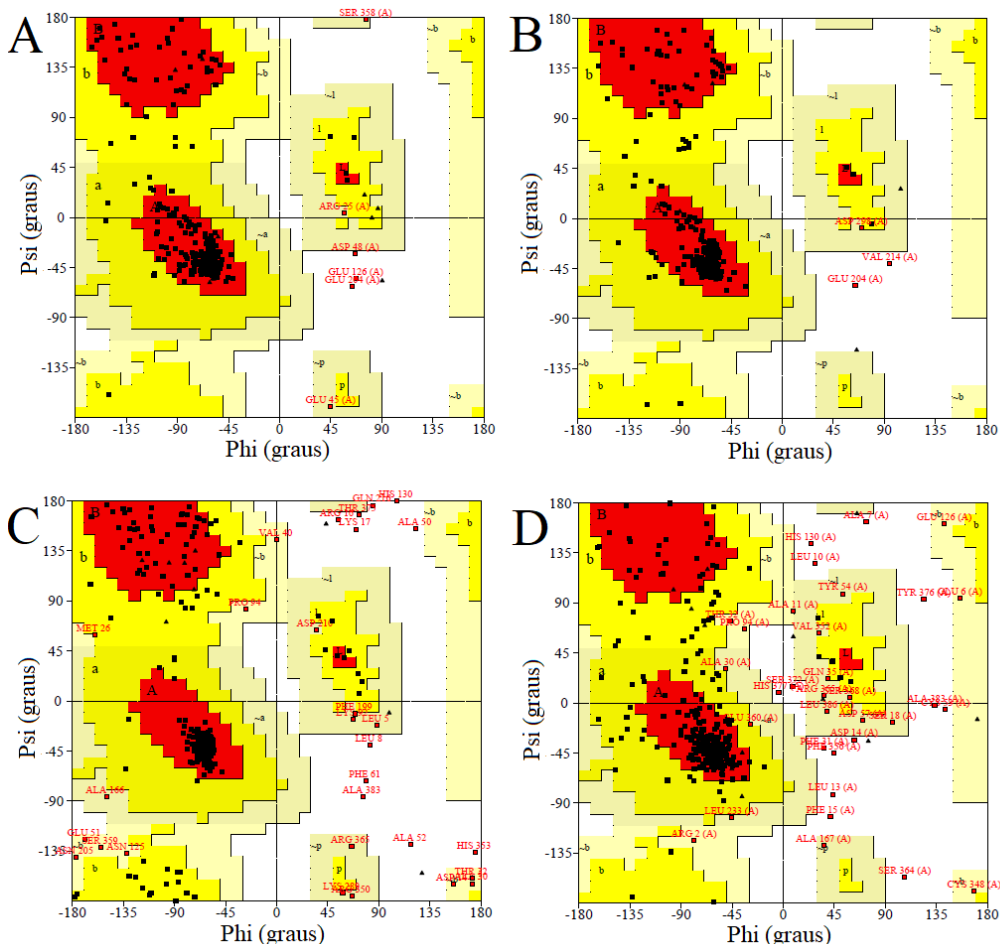
**TABELA 3:** Valores das porcentagens dos ângulos torcionais dos resíduos dos modelos estruturais do CX3CR1 ou isoforma 1 encontrados nas regiões descritas pelo gráfico de Ramachandran.

	RosettaFold	RosettaTR	Phyre2	I-Tasser
<i>Resíduos em regiões mais favoráveis (A, B e L)</i>	92,7%	90,8%	85,3%	77,1%
<i>Resíduos em regiões adicionais permitidas</i>	6,1%	7,6%	10,1%	17,7%
<i>Resíduos em regiões generosamente adicionais</i>	0,6%	0,3%	2,4%	3,4%
<i>Resíduos em regiões adicionais não permitidas</i>	0,6%	1,2%	2,1%	1,8%

Fonte: Elaborado pelo autor.

O mesmo processo de análise pelo gráfico de Ramachandran foi realizado para os modelos estruturais determinados para a isoforma 2 do receptor celular CX3CR1. A Figura 6 e Tabela 4 mostram que os servidores Rosetta FOLD e TR determinaram modelos para a isoforma 2 com as menores porcentagens de ângulos torcionais dos resíduos nas regiões generosamente adicionais (amarelo claro) e não permitidas (branco) do gráfico de Ramachandran, todos abaixo de 1%. Vale notar que o Rosetta FOLD apresentou a maior porcentagem de ângulos torcionais (93%) dos resíduos em regiões favoráveis (vermelho). Assim como para a isoforma 1, as porcentagens determinadas para os servidores I-Tasser e Phyre2 não foram comparativamente satisfatórias, pois observou-se valores significativos nas regiões generosamente adicionais e não permitidas dos gráficos de Ramachandran (Figura 6-C-D e Tabela 4).

**FIGURA 6:** Gráficos de Ramachandran para ângulos torcionais dos resíduos dos modelos estruturais da isoforma 2 determinados pelos servidores Rosetta FOLD (A), Rosetta TR (B), Phyre2 (C) e I-Tasser (D). Em vermelho resíduos em regiões favoráveis, amarelo escuro para regiões adicionais permitidas, em amarelo claro para regiões generosamente adicionais e em branco para regiões não permitidas.



Fonte: Elaborado pelo autor.

**TABELA 4:** Valores das porcentagens dos ângulos torcionais dos resíduos dos modelos estruturais da isoforma 2 encontrados nas regiões descritas pelo gráfico de Ramachandran.

	RosettaFold	RosettaTR	Phyre2	I-Tasser
<i>Resíduos em regiões mais favoráveis (A, B e L)</i>	93,0%	91,9%	81,2%	73,3%
<i>Resíduos em regiões adicionais permitidas</i>	5,3%	7,3%	11,0%	17,7%
<i>Resíduos em regiões generosamente adicionais</i>	0,8%	0,3%	4,8%	5,1%
<i>Resíduos em regiões adicionais não permitidas</i>	0,8%	0,6%	3,1%	3,9%

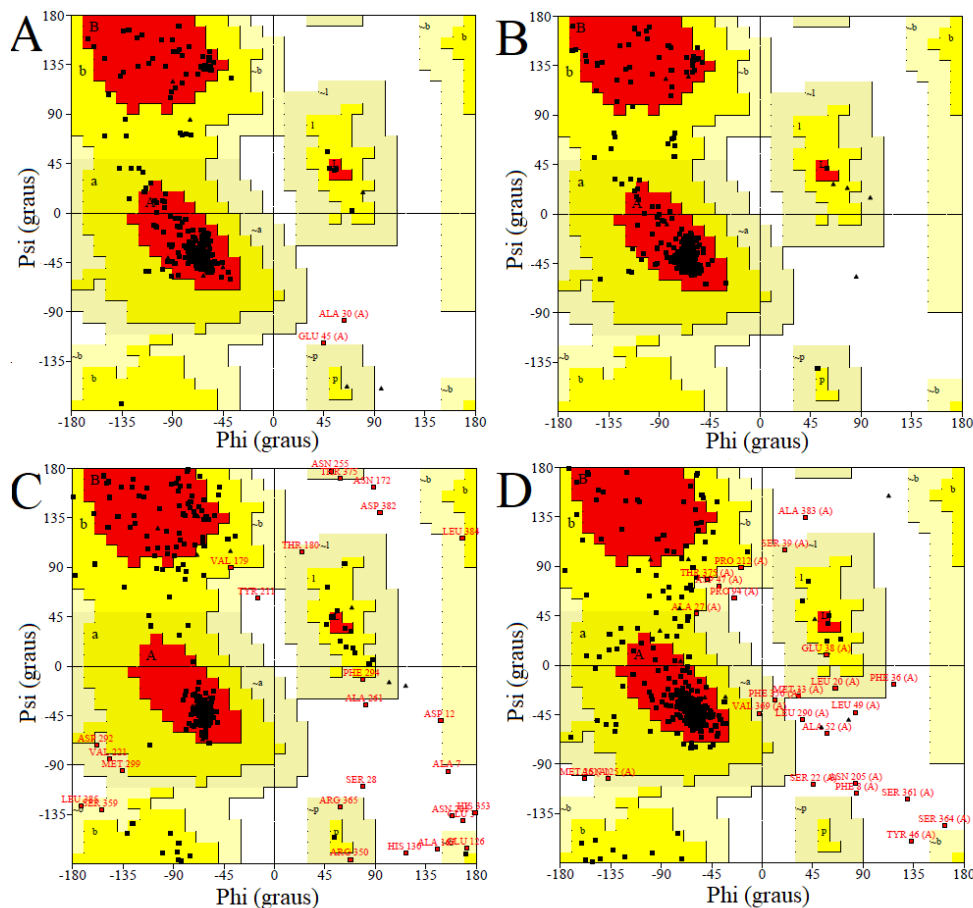
Fonte: Elaborado pelo autor.

A análise do gráfico de Ramachandran realizada para os modelos estruturais calculados para a isoforma 3 (Figura 7 e Tabela 5) revela nenhum resíduo com ângulos torcionais em regiões generosamente adicionais (amarelo claro) e não permitidas (branca) para a estrutura determinada pelo servidor Rosetta TR. Para o Rosetta FOLD, 0,6% e 1,2% dos resíduos do modelo estrutural apresentam ângulos torcionais em regiões generosamente adicionais e não permitidas, respectivamente (Tabela 5). Com relação aos resíduos ASP-32, ALA-142, GLU-179, VAL-189, ASN-230 e LYS-232 identificados nas regiões não permitidas (Figura 7-A), vale mencionar que esses são encontrados em regiões de loop no modelo estrutural da isoforma 3. Assim como para as isoformas 1 e 2, os resultados obtidos para os servidores I-Tasser e Phyre2 são menos apropriados quando comparados àqueles obtidos pelo Rosetta FOLD e TR.



Para a isoforma 4 do receptor celular CX3CR1, os melhores resultados das análises dos modelos estruturais gerados foram observados para os servidores Rosetta FOLD e TR, apresentando gráficos de Ramachandran semelhantes (Figura 8 – A e B e Tabela 6). As porcentagens de resíduos com ângulos torcionais nas regiões favoráveis (vermelho) são 93,3% e 93,5%, nas regiões adicionais permitidas (amarelo escuro) foram de 6,2% e 6,5%, nas regiões generosamente adicionais (amarelo claro) foram de 0,0% e 0,0% e nas regiões não permitidas (branco) foram de 0,6% e 0,0% para análise dos resultados determinados pelo Rosetta FOLD e TR, respectivamente (Tabela 6). Vale mencionar que o valor de 0,6% para ângulos torcionais em regiões não permitidas verificado para o Rosetta FOLD corresponde a dois resíduos, ALA-30 e GLU-45, os quais encontram-se na porção N-terminal da proteína. A análise dos resultados dos servidores I-Tasser e Phyre2 mostraram-se inferiores quando comparados com aqueles obtidos pelo Rosetta (FOLD e TR), da mesma forma como foi visto para as demais isoformas do CX3CR1 (Tabela 6).

**FIGURA 8:** Gráficos de Ramachandran para ângulos torcionais dos resíduos dos modelos estruturais da isoforma 4 determinados pelos servidores Rosetta FOLD (A), Rosetta TR (B), Phyre2 (C) e I-Tasser (D). Em vermelho resíduos em regiões favoráveis, amarelo escuro para regiões adicionais permitidas, em amarelo claro para regiões generosamente adicionais e em branco para regiões não permitidas.



Fonte: Elaborado pelo autor.

**TABELA 6:** Valores das porcentagens dos ângulos torcionais dos resíduos dos modelos estruturais da isoforma 4 encontrados nas regiões descritas pelo gráfico de Ramachandran.

	RosettaFold	RosettaTR	Phyre2	I-Tasser
Resíduos em regiões mais favoráveis (A, B e L)	93,3%	93,5%	80,3%	73,0%
Resíduos em regiões adicionais permitidas	6,2%	6,5%	12,4%	20,8%
Resíduos em regiões generosamente adicionais	0,0%	0,0%	3,9%	2,8%
Resíduos em regiões adicionais não permitidas	0,6%	0,0%	3,4%	3,4%

Fonte: Elaborado pelo autor.

Um alinhamento estrutural e cálculos de RMSD foram realizados para se comparar os modelos estruturais da CX3CR1 e suas isoformas determinados pelos servidores Rosetta FOLD, Rosetta TR, I-Tasser e Phyre2. Para isso utilizou-se como referência o modelo estrutural obtido pelo Rosetta FOLD e a região das proteínas empregadas no alinhamento e cálculo de RMSD foi parte do barril de hélices alfa, descartando o N- e C-terminais. A Tabela 7 mostra valores de RMSD acima de 1 Å para os modelos estruturais determinados pelos programas I-Tasser e Phyre2 quando comparados com o Rosetta FOLD e valores menores do que 1 Å para as estruturais calculadas pelo servidor Rosetta TR.

**TABELA 7:** Valores de RMSD para a região do barril de hélices alfa calculados para os modelos estruturais determinados pelos servidores Rosetta FOLD, Rosetta TR, I-tasser e Phyre2. O modelo do Rosetta FOLD foi utilizado como referência para o alinhamento estrutural e cálculo de RMSD.

	Isoforma 1 (Å)	Isoforma 2 (Å)	Isoforma 3 (Å)	Isoforma 4 (Å)
RosettaFold	0	0	0	0
RosettaTR	0,635	0,670	0,560	0,713
I-Tasser	1,005	1,447	1,063	1,441
Phyre2	1,319	1,244	1,158	1,676

Fonte: Elaborado pelo autor.

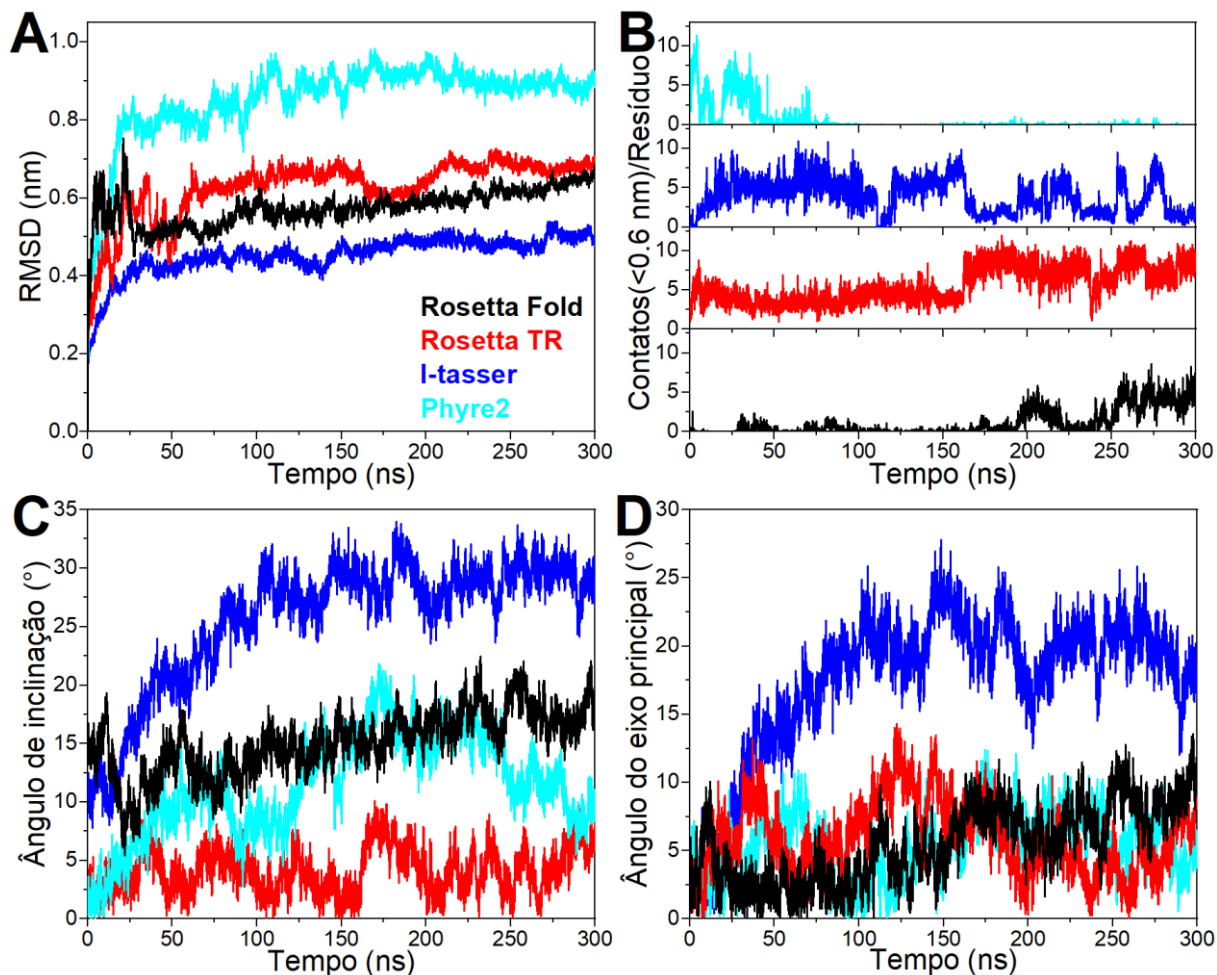
## 4.2 Simulações de Dinâmica Molecular do Receptor e suas Isoformas

Simulações de Dinâmica Molecular (DM) de 300 ns em bicamada lipídica de POPC foram realizadas para as estruturas tridimensionais do receptor celular CX3CR1, com a finalidade de avaliar a estabilidade estrutural dos modelos calculados pelo Rosetta FOLD, Rosetta TR, I-Tasser e Phyre2. Para otimizar o tempo computacional, foram realizadas simulações de DM somente para a isoforma 1 do receptor celular CX3CR1 determinados pelos quatro servidores utilizados neste trabalho. A Figura 9-A mostra os valores de RMSD calculados para os átomos da cadeia principal da isoforma 1 ao longo dos 300 ns de simulação, empregando o primeiro *frame* (modelo estrutural) como a estrutura de referência para o alinhamento estrutural. É possível observar que após 50 ns de simulação os valores de RMSD alcançam patamares de

estabilidade para todos os modelos estruturais. Nota-se que para o modelo determinado pelo Phyre2 foram observados os maiores valores de RMSD, em torno de 0,9 nm (~9 Å), enquanto que para o modelo calculado pelo I-Tasser e Rosetta FOLD foram observados os menores valores ao longo dos 300 ns, mantendo entre 0,4 (~4) e 0,5 nm (~5 Å), respectivamente.

Com a finalidade de verificar a exposição da cavidade de ligação do receptor celular CX3CR1 (isoforma 1) com relação a região N-terminal da proteína, o número de contatos < 0,6 nm foram calculados entre os átomos dos resíduos N-terminais (MET1–ILE23) e os da fita β (GLU174–GLY177) na entrada da cavidade de ligação ao longo dos 300 ns de simulação. A partir da Figura 9-B observa-se um número maior de contatos para os modelos estruturais determinados pelo Rosetta TR e I-Tasser, o que significa que a cavidade de ligação na proteína está ocluída pelos resíduos do N-terminal na maior parte do tempo de simulação. Por outro lado, para os modelos do Phyre2 e Rosetta FOLD, é possível observar valores reduzidos de número de contatos e com uma frequência menor de ocorrência quando comparados com o I-Tasser e Rosetta FOLD, indicando que a cavidade de ligação encontra-se exposta na maior parte do tempo de simulação. Uma análise de clusterização das trajetórias dos 300 ns de simulação dos modelos estruturais da isoforma 1 corroboram com a contagem do número de contatos < 0,6 nm, revelando que as cavidades de ligação na porção N-terminal da proteína é exposta nas estruturas representativas para os modelos do Rosetta FOLD (Figura 10-A) e Phyre2 (Figura 10-D), enquanto que ela é ocluída pelos resíduos N-terminais (verde) para as estruturas determinadas pelo Rosetta TR (Figura 10-B) e I-Tasser (Figura 10-C).

**FIGURA 9:** Análise de RMSD dos átomos da cadeia principal (A), do número de contatos < 0,6 nm por resíduos entre os átomos dos resíduos N-terminal e da fita beta GLU174–GLY177 (B), do ângulo de inclinação da proteína com relação a bicamada lipídica (C) e do ângulo do eixo principal com relação a normal da bicamada (D) ao longo das trajetórias de 300 ns das simulações de DM dos modelos estruturais da isoforma 1 do C3XCR1 determinados pelos servidores Rosetta FOLD (preto), Rosetta TR (vermelho), I-tasser (azul) e Phyre2 (ciano).



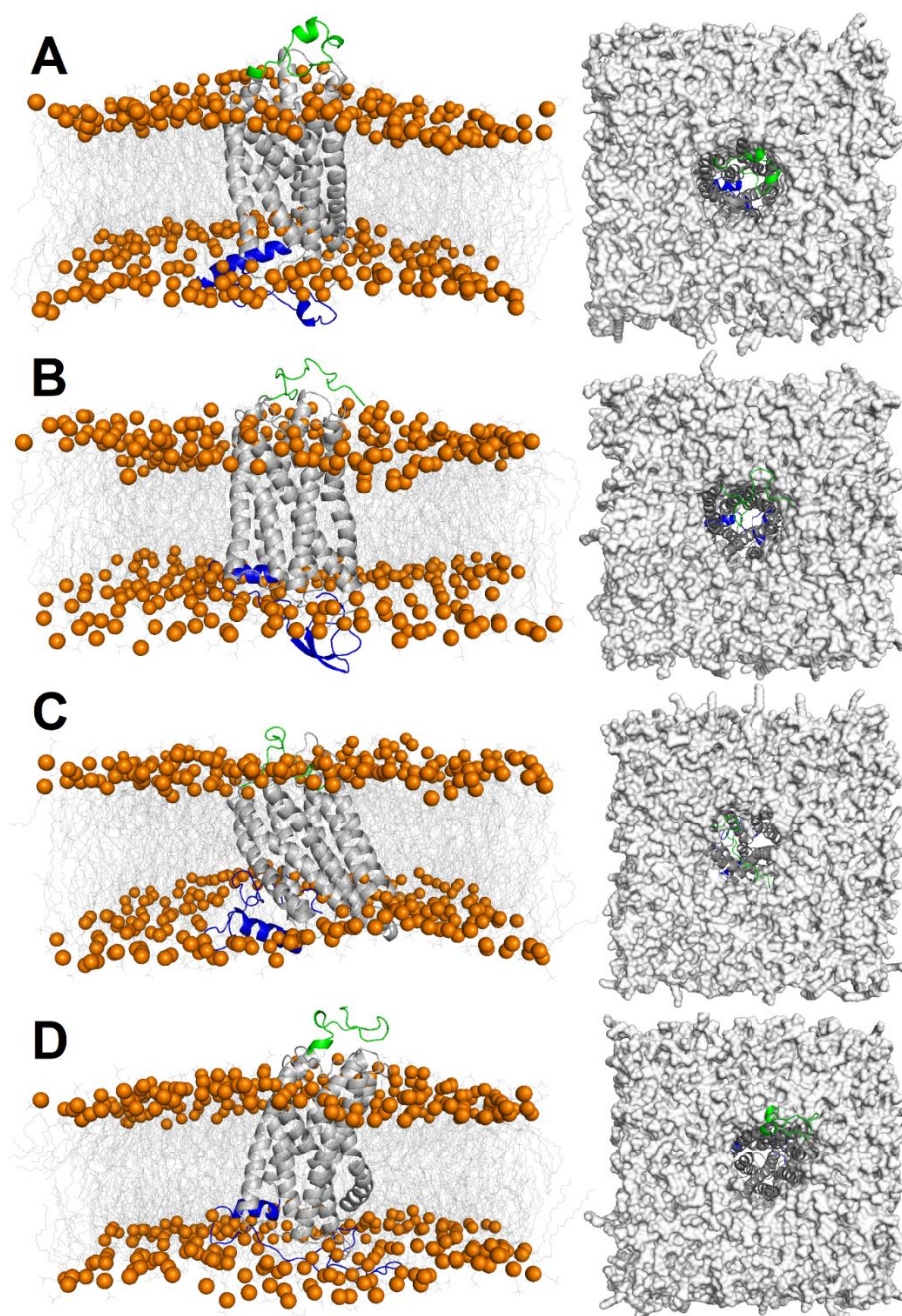
Fonte: Elaborado pelo autor.

Uma análise da orientação dos modelos estruturais da isoforma 1 do CX3CR1 com relação a bicamada lipídica foi realizada ao longo das trajetórias de 300 ns de simulação. Essa análise foi realizada de duas maneiras calculando o ângulo de inclinação (Figura 9-C) e o ângulo do eixo principal (Figura 9-D) do barril de hélices alfa da proteína utilizando as ferramentas *bundle* e *principal* do pacote GROMACS. Dentre os modelos estruturais simulados, a estrutura referente ao servidor I-Tasser apresentou as maiores variações dos ângulos de inclinação e do eixo principal (Figura

9-C e 9-D), indicando que a proteína se inclina em direção ao núcleo hidrofóbico (cauda dos lipídeos) da bicamada lipídica. Na Figura 10-C é possível observar uma inclinação na estrutura representativa do modelo do I-Tasser obtida para a análise de clusterização feita para os 300 ns de simulação. Para as simulações dos demais modelos estruturais, há apenas flutuações nos valores dos ângulos de inclinação e do eixo principal (Figura 9-C e 10-D), não apresentando mudanças significativas que resultem em uma inclinações da proteína em direção as caudas hidrofóbicas da bicamada lipídica nas estruturas representativas da clusterização (Figura 10-A, 10-B e 10-D).

Com base nos resultados dos gráficos de Ramachandran e análises das trajetórias de DM para os modelos estruturais da isoforma 1 do CX3CR1, decidiu-se prosseguir com a modelagem do Rosetta FOLD para as demais isoformas 2, 3 e 4 devido aos resultados satisfatórios dos ângulos torcionais, estabilidade estrutural (valores de RMSD e ângulos de orientação com relação a bicamada) ao longo dos 300 ns de simulação e exposição da cavidade de ligação na porção N-terminal do barril de hélices alfa na estrutura representativa obtida pela clusterização.

**FIGURA 10:** Estruturas representativas para os modelos estruturais da isoforma 1 do Rosetta FOLD (A), Rosetta TR (B), I-Tasser (C) e Phyre2 (D) em bicamada lipídica de POPC, as quais foram obtidas pela análise de clusterização dos 300 ns de simulações de DM. A proteína é representada em modelo de *cartton* e os lipídeos como modelos de esferas e linhas. O barril de hélices alfa da proteína são denotados em cinza, a região N-terminal em verde e a C-terminal em azul. O átomo de fósforo da cabeça polar dos lipídeos é mostrado como esfera laranja e a cauda hidrofóbica como linhas cinzas. As imagens da direita mostram uma vista de cima de como a proteína encontra-se inserida na bicamada lipídica de POPC.

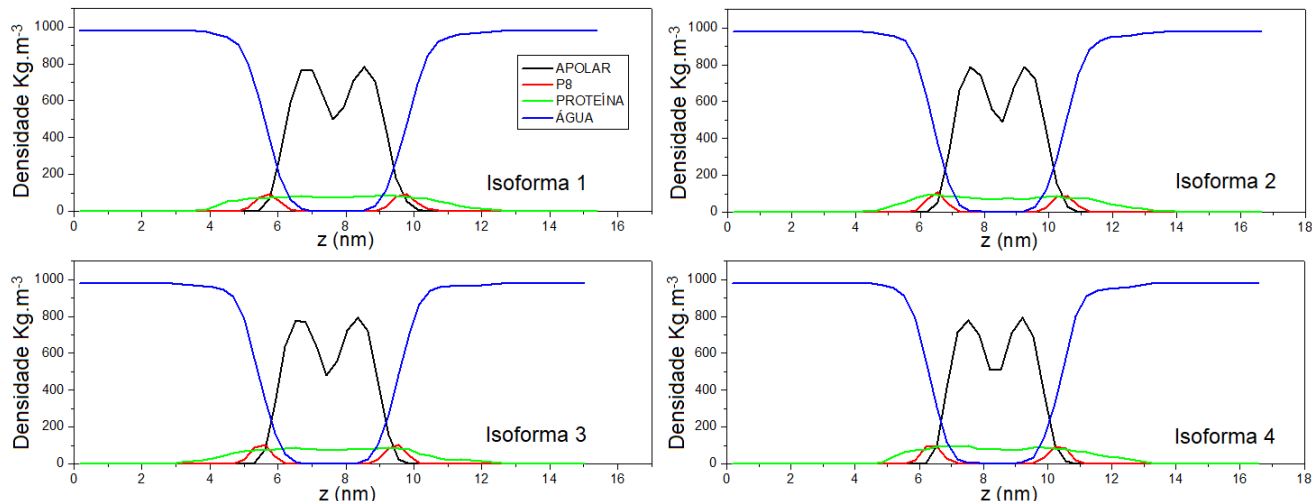


Fonte: Elaborado pelo autor.

Ao final das simulações de DM, uma análise pode ser feita para investigar a estabilidade da bicama lipídica de POPC, uma vez que essa estabilidade é de fundamental importância para prosseguir com os próximos cálculos. Dessa forma, curvas de densidade de massa ( $\text{Kg}\cdot\text{m}^{-3}$ ) para diferentes grupos que compõem o sistema foram calculadas (Figura 11). Através de uma varredura ao longo do eixo Z (eixo normal à bicamada) é quantificada a massa de cada molécula presente em cada ponto que faz parte do sistema molecular. Este mapeamento gera um perfil de densidade de massa ( $\text{Kg}\cdot\text{m}^{-3}$ ). Desta forma, pode-se determinar a região ocupada pelo solvente, pela membrana e pelos resíduos da proteína CX3CR1/Isoformas.

As mesmas análises de RMSD e contatos realizadas para as trajetórias da isoforma 1 do CX3CR1 também foram feitas para as simulações de DM dos modelos estruturais da isoforma 2, 3 e 4 determinados pelo servidor Rosetta FOLD (Figura 12). Assim como para a isoforma 1, os valores de RMSD (Figura 12-A) para as isoformas 2, 3 e 4 alcançaram patamares de estabilidade após os 50 ns simulações de DM. As isoformas 2 e 4 apresentaram mudanças mais significativas nos valores de RMSD, alcançando patamares por volta de 0,9 ( $\sim 9$ ) e 0,8 nm ( $\sim 8$  Å), respectivamente, enquanto que a isoforma 3 mostrou uma menor variação ( $\sim 5$  Å) com relação a estrutura de início (o *frame* de referência). Uma explicação plausível para esse resultado é que as isoformas 2 e 4 apresentam uma região N-terminal mais longa, com 32 resíduos de aminoácidos a mais em relação a isoforma 1.

**FIGURA 11:** Perfil de densidade de massa para a bicamada de POPC solvatada para as quatro isoformas. A espessura da bicamada é determinada pela distância P8-P8. A parte apolar da bicamada de POPC é denotada com uma linha preta, os fósforos P8 com linha vermelha, a proteína com linha verde e moléculas de água com linha azul.



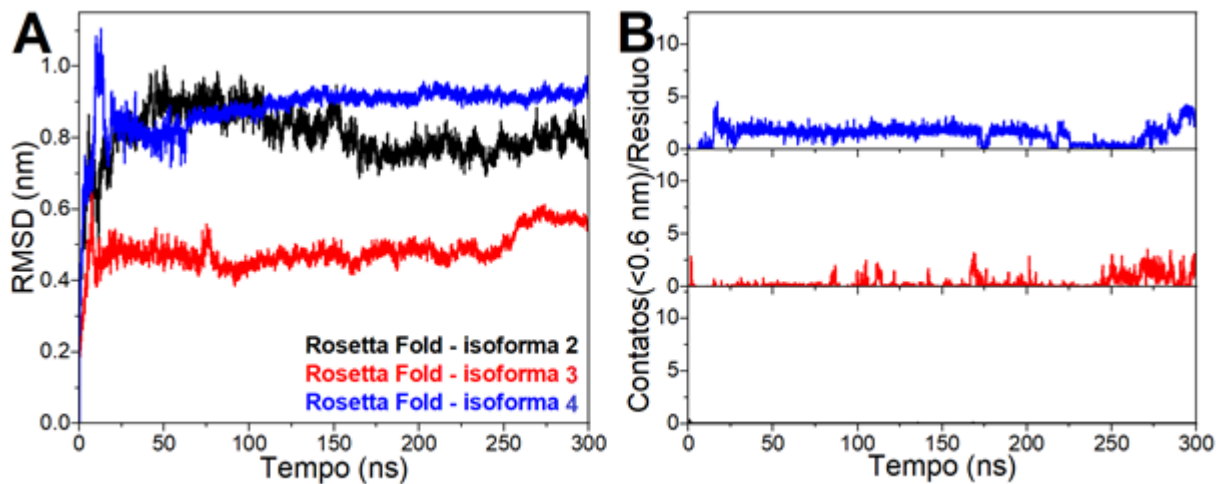
Fonte: Elaborado pelo autor.

Pelos perfis de densidade de massa observados na Figura 11, é possível observar que a bicama lipídica de POPC é estável ao longo dos 300 ns de simulação, uma vez que: a cadeia apolar dos lipídeos ocupa a região central do eixo de coordenada Z (perpendicular a bicamada lipídica); as moléculas de águas ocupam as extremidades do eixo de coordenada Z (maior densidade) e não ocupando o centro da bicamada; a proteína ocupa porções interiorizadas na bicamada lipídica e apresenta regiões (terminais) expostas ao solvente; e os átomos de fósforo ocupam a região de limiar entre as moléculas de água e os átomos apolares dos lipídeos.

A Figura 12-B mostra que o número de contatos < 0,6 nm entre os átomos dos resíduos N-terminais e a fita beta GLU174–GLY177 (numeração referente à isoforma 1) na entrada da cavidade de ligação apresenta uma contagem significativamente menor do que aquelas observadas nas verificações feitas para a isoforma 1. Praticamente nenhum contato é observado para a isoforma 2, caracterizando que a cavidade de ligação é exposta, o que corrobora com a estrutura representativa na Figura 10-A obtida pela análise de clusterização. Para a isoformas 3, são observados contatos menores que 0,6 nm, mas esses não são frequentes ao longo dos 300 ns de simulação, resultando em uma estrutura representativa sem oclusão da cavidade

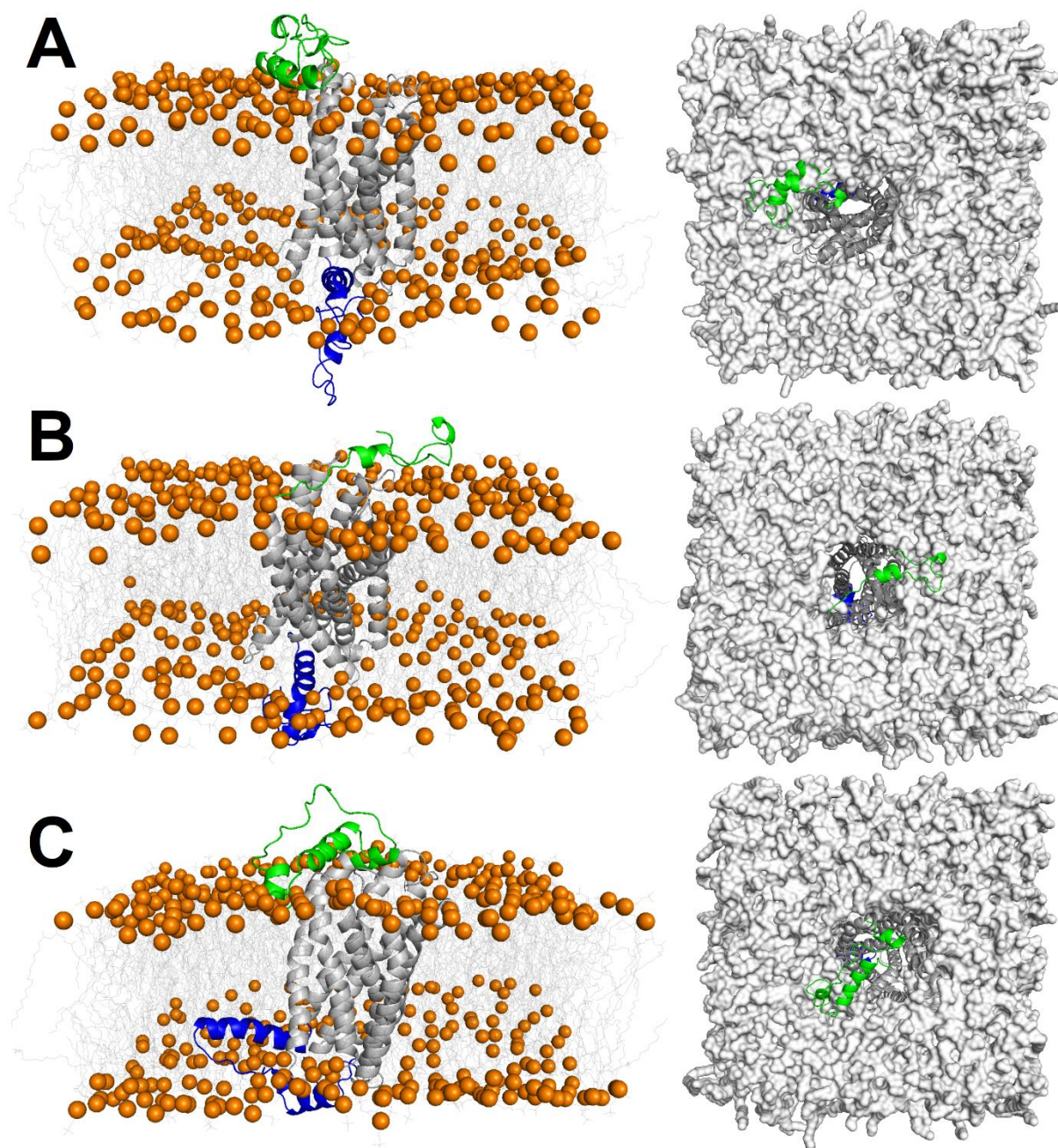
de ligação na região N-terminal da proteína (Figura 13-B). No caso da isoforma 4, a contagem de contatos é relativamente constante na maior parte do tempo de simulação, mas esses envolvem os resíduos MET-1 e ARG-2 que são próximos da fita beta GLU206-GLY208 (para a isoforma 1 são GLU174-GLY177), sem promover oclusão da cavidade de ligação na região N-terminal da proteína (estrutura representativa na Figura 17-C).

**FIGURA 12:** Análise de RMSD dos átomos da cadeia principal (A) e do número de contatos < 0,6 nm entre os átomos dos resíduos N-terminal e da fita beta GLU174-GLY177 (B) ao longo das trajetórias de 300 ns das simulações de DM dos modelos estruturais das isoformas 2, 3 e 4 do C3XCR1 determinados pelos servidores Rosetta FOLD.



Fonte: Elaborado pelo autor.

**FIGURA 13:** Estruturas representativas para os modelos estruturais da isoforma 2 (A), isoforma 3 (B) e isoforma 4 (C) em bicamada lipídica de POPC, as quais foram obtidas pela análise de clusterização dos 300 ns de simulações de DM. A proteína é representada em modelo de *cartoon* e os lipídeos como modelos de esferas e linhas. O barril de hélices alfa da proteína são denotados em cinza, a região N-terminal em verde e a C-terminal em azul. O átomo de fósforo da cabeça polar dos lipídeos é mostrado como esfera laranja e a cauda hidrofóbica como linhas cinzas. As imagens da direita mostram uma vista de cima de como a proteína encontra-se inserida na bicamada lipídica de POPC.



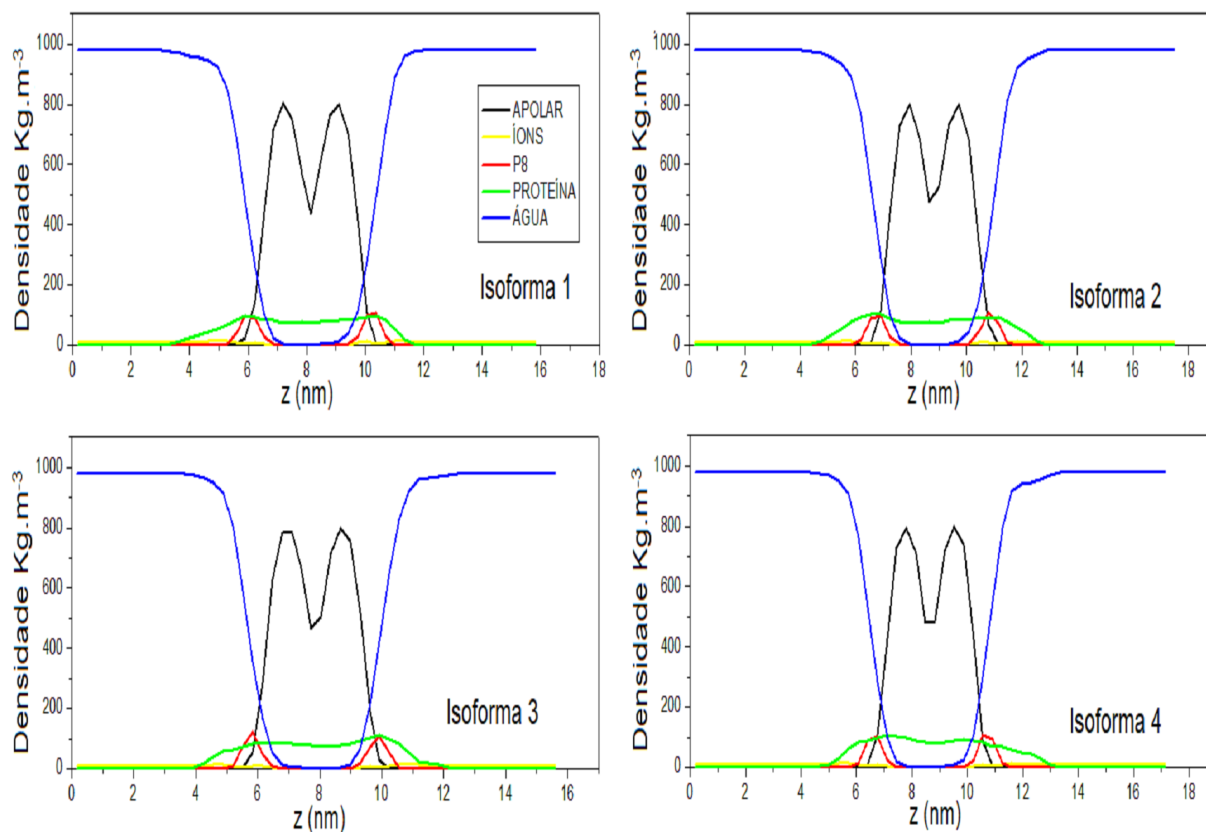
Fonte: Elaborado pelo autor.

### **4.3 Simulações de Dinâmica Molecular de CX3CR1 e suas Isoformas em POPC na presença de Força Iônica e em DPPC na ausência e presença de Força Iônica**

Neste trabalho foi realizado simulações de MD para analisar o efeito da força iônica nos modelos computacionais obtidos pelo programa Rosetta FOLD das isoformas do receptor celular CX3CR1, para isso foi adicionado uma concentração de 150 mM de NaCl (Cloreto de sódio) para a bicama lipídica de POPC e, em seguida, realizou-se a simulação de DM de 300 ns do receptor e suas isoformas.

Para os perfis de densidade de massa para o sistema molecular em bicamada lipídica de POPC com adição de 150 Mm de NaCl é estável ao longo dos 300 ns de simulação, uma vez que: a cadeia apolar dos lipídeos ocupa a região central do eixo de coordena Z (perpendicular a bicamada lipídica); as moléculas de águas ocupam as extremidades do eixo de coordenada Z (maior densidade) e não ocupando o centro da bicamada; a proteína ocupa porções interiorizadas na bicamada lipídica e apresenta regiões (terminais) expostas ao solvente; e os átomos de fósforo ocupam a região de limiar entre as moléculas de água e os átomos apolares dos lipídeos (Figura 14).

**FIGURA 14:** Perfil de densidade de massa para a bicamada de POPC solvatada com NaCl para as quatro isoformas. A espessura da bicamada é determinada pelas distância P8-P8. A parte apolar da bicamada de POPC é denotada com uma linha preta, os íons de NaCl são mostrados com linha amarela, os fósforos P8 com linha vermelha, a proteína com linha verde e moléculas de água com linha azul.



Fonte: Elaborado pelo autor.

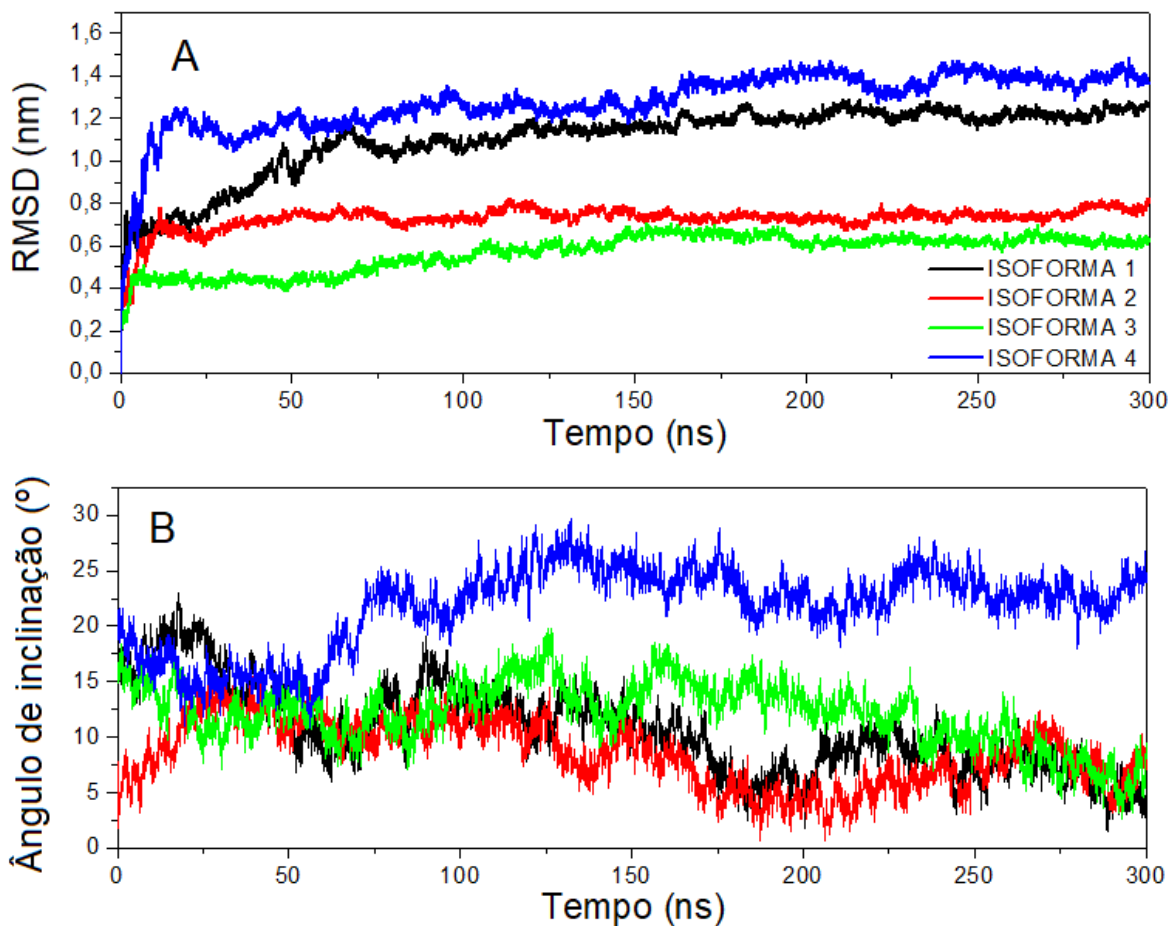
As mesmas análises de RMSD realizadas para as trajetórias das isoformas do CX3CR1 em POPC sem adição de NaCl também foram feitas para as simulações de DM das isoformas em POPC/NaCl. Os valores de RMSD (Figura 15-A) para a isoformas alcançaram patamares de estabilidade após 25 ns simulações de DM. As isoformas 2 e 3 estabilizaram nos menores tempos de simulação, no qual apresentaram valores de RMSD por volta de 0,6 (6 Å) e 0,4 nm (4 Å), respectivamente, enquanto a isoforma 1 alcançou um patamar de estabilidade de RMSD em torno 1 nm (10 Å) após 50 ns de simulação. A isoforma 4 foi a que apresentou mais variação significativa para os valores de RMSD, os quais foram em torno de 1,2 nm (12 Å) com relação a estrutura de início (o *frame* de referência). Uma explicação plausível para

esse resultado é que a isoformas 4 apresenta uma região N-terminal mais longo, com 32 resíduos de aminoácidos.

As análises do ângulo de inclinação da proteína em relação a bicamada lipídica foram realizadas (Figura 15-B), mostrando que não houve variação significativa nos ângulos de inclinação, ou seja, as isoformas se mantiveram em um posicionamento vertical em relação a bicamada lipídica ao longo dos 300 ns de simulação.

Uma vez que as simulações das isoformas em POPC com 150 mM de NaCl são parecidas com as simulações em POPC sem NaCl, pode-se concluir que a utilização dos sistemas moleculares com bicamada lipídica com íons apenas para a neutralização é suficiente para prosseguir com as simulações de DM dos modelos estruturais dos complexos determinados pelos cálculos de *docking* molecular do peptídeo da G.

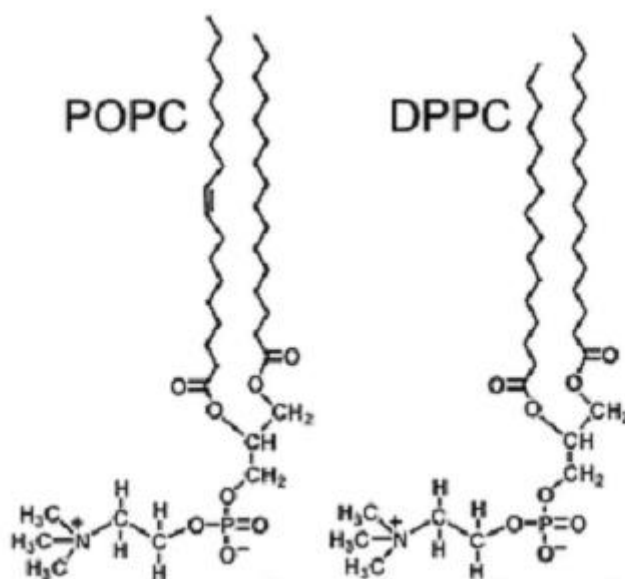
**FIGURA 15:** Análise de RMSD dos átomos da cadeia principal das isoformas (A) e do ângulo de inclinação da proteína com relação a bicamada lipídica de POPC com 150 mM de NaCl (B) ao longo das trajetórias de 300 ns das simulações de DM das isoformas do modelo estrutural de CX3CR1. Sendo a isoforma 1 representada em preto, a isoforma 2 em vermelho, a isoforma 3 em verde e a isoforma 4 em azul.



Fonte: Elaborado pelo autor.

O DPPC é um fosfolípido nomeado dipalmitoilfosfatidilcolina que consiste em dois grupos de ácido palmítico C16 ligados a um grupo de cabeça polar de fosfatidilcolina. A molécula de DPPC se diferencia do POPC basicamente pelo tamanho de sua cauda apolar, mantendo os grupos de cabeça polar iguais (Figura 16). Enquanto o POPC tem fórmula química como  $C_{42}H_{82}NO_8P$ , o DPPC tem a fórmula  $C_{40}H_{80}NO_8P$ , o que diferencia em quantidade de dois átomos de carbono na cauda apolar do lipídeo, e outra diferença a ser notada é uma instauração ou dupla ligação no POPC e que não há no DPPC (Figura 20).

**FIGURA 16:** Comparação entre as estruturas de POPC e DPPC.



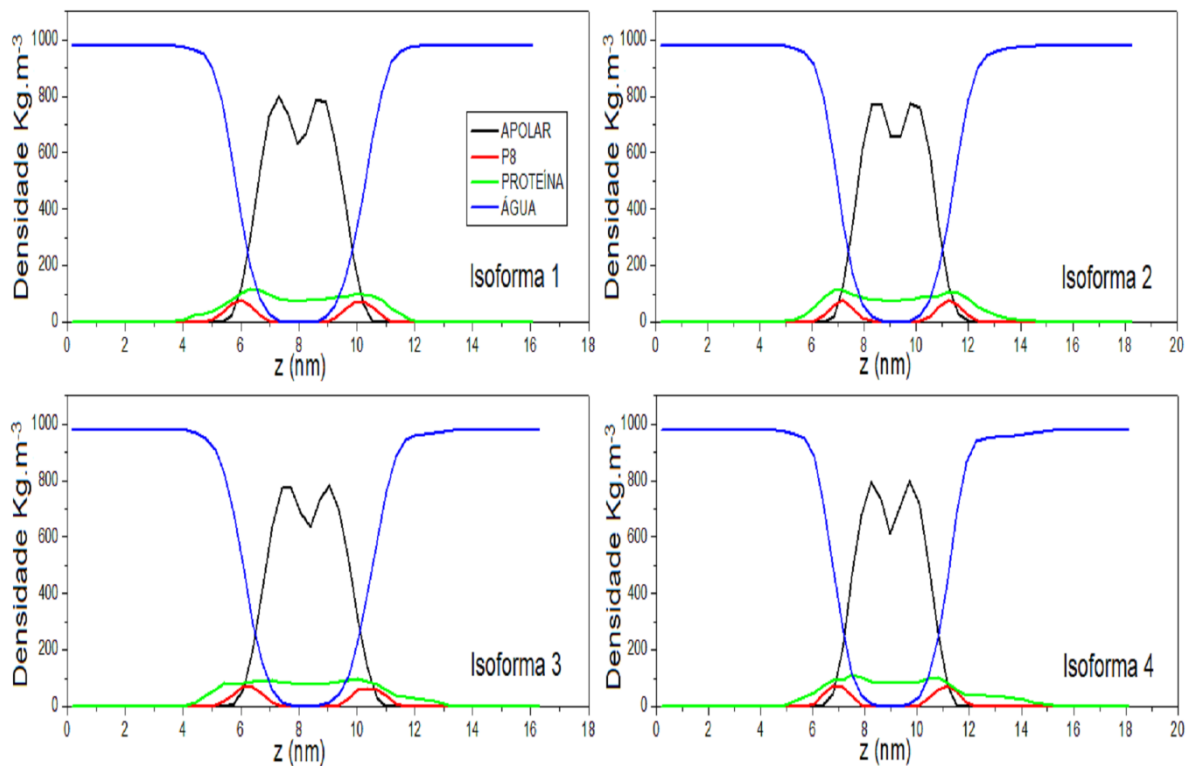
Fonte: Adaptada de DILBECK, 2013.

Dessa forma, as simulações de DM de 300 ns foram realizadas para as isoformas do CX3CR1 em bicamada lipídica de DPPC obtidas pelos cálculos de modelagem molecular pelo servidor Rosetta FOLD, a fim de obter dados comparativos entre as duas bicamadas lipídicas. Perfis de densidade de massa também foram realizadas para a bicamada lipídica de DPPC neutralizada (Figura 17) e solvatada com 150 mM de NaCl (Figura 18).

Perfis de densidade de massa foram calculadas, tanto para a forma com os íons de cloreto de sódio (Figura 18) quanto para a forma neutralizada da caixa de simulação (Figura 17). Para os perfis de densidade de massa em bicamada lipídica de DPPC, apresentaram perfis similares a bicamada lipídica de POPC. Dessa forma, curvas de densidade de massa ( $\text{Kg}\cdot\text{m}^{-3}$ ) para diferentes grupos que compõem o sistema foram calculadas (Figura 17 e 18). Através de uma varredura ao longo do eixo Z (eixo normal à bicamada) é quantificada a massa de cada molécula presente em cada ponto que faz parte do sistema molecular. Este mapeamento gera um perfil de densidade de massa ( $\text{Kg}\cdot\text{m}^{-3}$ ). Desta forma, pode-se determinar a região ocupada pelo solvente, pela membrana e pelos resíduos da proteína CX3CR1/Isoformas. Pelos

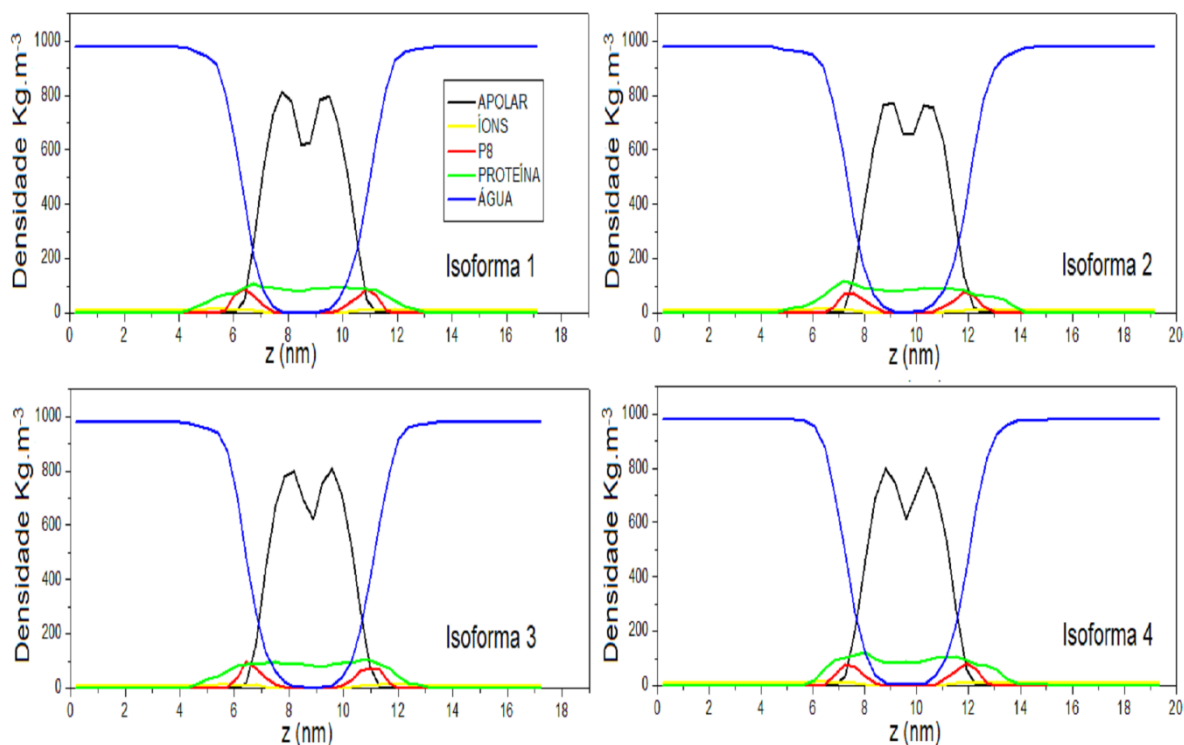
perfis de densidade de massa observados na Figura 17 e 18, é possível observar que a bicama lipídica de DPPC é estável ao longo dos 300 ns de simulação, uma vez que: a cadeia apolar dos lipídeos ocupa a região central do eixo de coordena Z (perpendicular a bicamada lipídica); as moléculas de águas ocupam as extremidades do eixo de coordenada Z (maior densidade) não ocupando o centro da bicamada; a proteína ocupa porções interiorizadas na bicamada lipídica e apresenta regiões (terminais) expostas ao solvente; e os átomos de fósforo ocupam a região de limiar entre as moléculas de água e os átomos apolares dos lipídeos (Figura 17 e 18); os íons de NaCl ocupam as extremidades do eixo de coordenada Z, semelhante a água, porém com baixa densidade (Figura 18-linha amarela). Observando os graficos abaixo, estima-se que o comprimento da camada de DPPC seja em torno de 15 Å menor em relação ao de POPC que é em torno de 20 Å.

**FIGURA 17:** Perfil de densidade de massa para a bicamada de DPPC solvatada para as quatro isoformas. A espessura da bicamada é determinada pela distância P8-P8. A parte apolar da bicamada de DPPC é denotada com uma linha preta, os fósforos P8 com linha vermelha, a proteína com linha verde e moléculas de água com linha azul.



Fonte: Elaborado pelo autor.

**FIGURA 18:** Perfil de densidade de massa para a bicamada de DPPC solvatada com concentração de íons NaCl de 150mM para as quatro isoformas. A espessura da bicamada é determinada pela distância P8-P8. A parte apolar da bicamada de DPPC (preto), os íons de NaCl (amarelo), os fósforos P8 (vermelho), a parte da proteína (verde) e a água (azul).

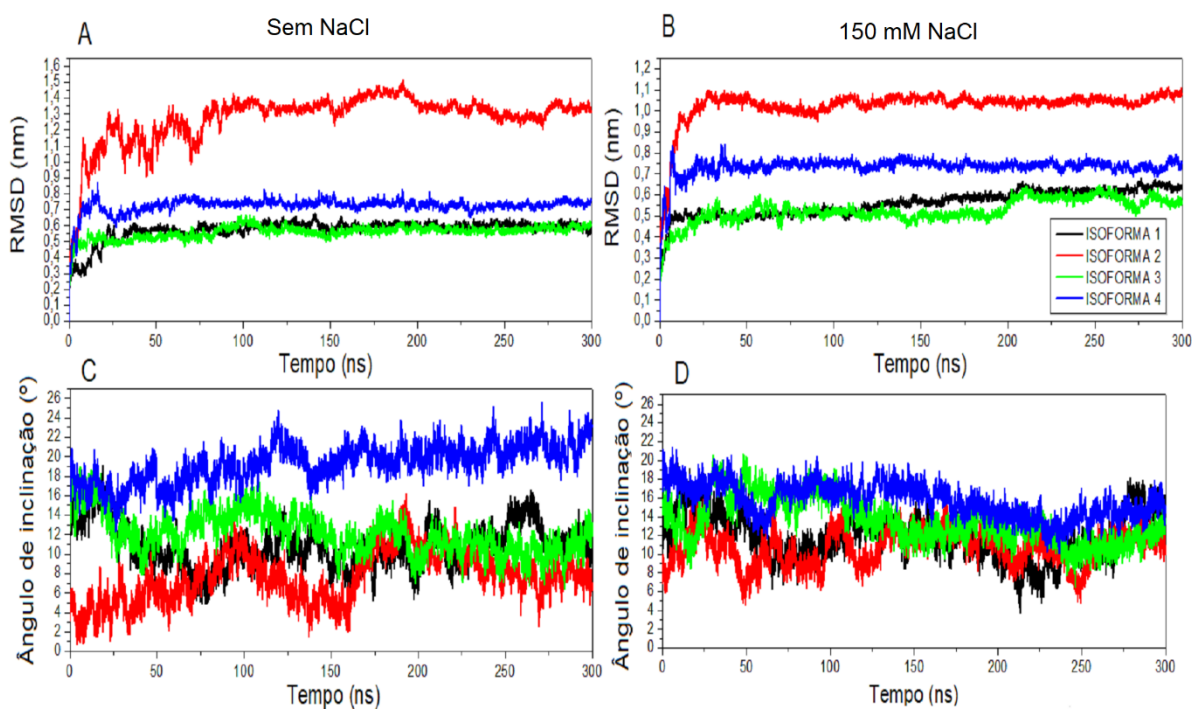


Fonte: Elaborado pelo autor.

As mesmas análises de RMSD realizadas para as trajetórias das isoformas do CX3CR1 em POPC também foram feitas para as simulações de DM das isoformas em DPPC e DPPC com NaCl. Os valores de RMSD (Figura 19-A e B) para as isoformas alcançarem patamares de estabilidade após os 50 ns de simulações de DM para isoformas 1 e 3 foram por volta de 0,5 nm (5 Å), e esse mesmo valor foi obtido para a simulação de DM com 150 mM de NaCl. Enquanto a isoforma 4 apresentou valores de RMSD em torno 0,7 nm (7 Å), a isoforma 2 foi a que mais apresentou variação significativa nos valores de RMSD em torno de 1,3 nm (13 Å) para a amostra sem adição de NaCl e ~1 nm (10 Å) para o sistema com adição de 150 mM de NaCl. Para explicar esse resultado é concluído que as isoformas 2 e 4 apresentam regiões N-terminais mais longas, com 32 resíduos de aminoácidos, dessa forma um maior alongamento do N-terminal dita um maior valor de RMSD. As análises do ângulo de

inclinação da proteína em relação ao eixo normal da bicamada lipídica foram realizadas (Figura 19-C e D) e mostraram que não houve variação significativa desses ângulos de inclinação, ou seja, as isoformas se mantiveram em uma orientação praticamente vertical com relação a normal da bicamada lipídica de DPPC com ou sem a adição de íons de cloreto de sódio durante os 300 ns de simulação. Dessa forma a proteína se manteve estável na sua posição vertical não afetando nenhum limite da bicamada lipídica.

**FIGURA 19:** Análise de RMSD dos átomos da cadeia principal das isoformas (A e B) e do ângulo de inclinação da proteína com relação a normal da bicamada lipídica de DPPC (C e D), sendo os resultados na presença de 150 mM de cloreto de sódio na coluna da direita e sem NaCl na coluna da esquerda, ao longo das trajetórias de 300 ns das simulações de DM das isoformas para os modelos estruturais do receptor celular CX3CR1. Sendo a isoforma 1 representada em preto, isoforma 2 em vermelho, isoforma 3 em verde e isoforma 4 em azul.



Fonte: Elaborado pelo autor.

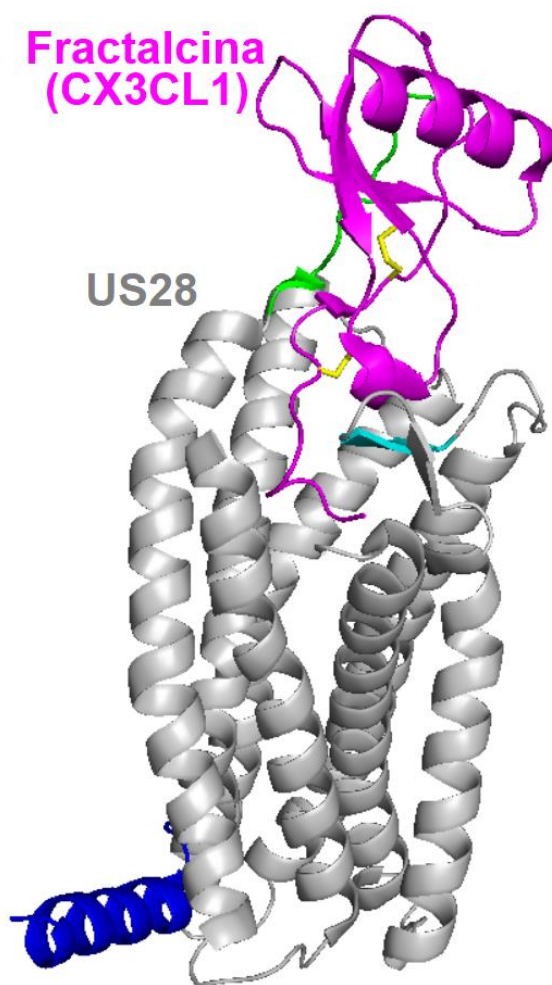
Com os resultados obtidos no excerto acima, é possível afirmar que o emprego de 150 mM de NaCl nos sistemas moleculares estudados não afetou o processo de estabilidade da bicamada lipídica com as isoformas como evidenciado pelos perfis das análises de densidade de massa, assim como também não afetou os tempos de

estabilidade estrutural da proteína (por volta de 50 ns). Da mesma forma, como não houveram variações também para os resultados usando DPPC ao invés de POPC, e como o POPC é o lipídio que melhor mimetiza o sistema biológico em questão, para o restante deste trabalho os próximos passos são utilizados o lipídio POPC com íons de cloreto de sódio empregados apenas para a neutralização do sistema.

#### **4.4 Cálculos de *Docking* Molecular e Simulações de Dinâmica Molecular da Interação do CX3CR1 com o Peptídeo da Região Conservada de Cisteínas da Proteína G do hRSV**

Os cálculos de *docking* molecular foram realizados com a estruturas representativas obtidas da análise de clusterização feitas para as simulações de 300 ns dos modelos estruturais das isoformas determinadas pelo servidor Rosetta FOLD. Para esses cálculos foram empregados diferentes servidores como ZDOCK, HDOCK, ClusPRO, HADDOCK, pyDock, GalaxyWeb, HpepDock e MDockPeP (seção 3.3). Entretanto, apenas ClusPro e HADDOCK apresentaram resultados satisfatórios, com a região de volta entre as duas hélices alfa do peptídeo (Figura 2 – região conservada do ectodomínio da proteína G) voltada para a cavidade de ligação na porção N-terminal das isoformas do CX3CR1. A busca por essa orientação dentre os resultados dos diferentes *servers* teve como base a interação da fractalcina (CX3CL1), o ligante biológico do receptor celular CX3CR1, com um receptor homólogo US28 que apresenta 38% de identidade de sequência (BURG *et al*, 2015) (Figura 20). Dessa forma, os modelos estruturais determinados pelos servidores ClusPro e HADDOCK para o complexo formado entre as isoformas do CX3CR1 e o peptídeo da região conservada de cisteínas da proteína G do hRSV prosseguiram para as simulações de DM de 300 ns.

**FIGURA 20:** Estrutura de cristal (PDB 4XTI) do complexo formado entre a fractalcina CX3CL1 (magenta) com o receptor homólogo US28 (cinza). As proteínas foram representadas com o modelo *cartoon*, com os resíduos N- e C-terminais denotados em verde e azul, respectivamente. A correspondente fita beta do homólogo CX3CR1 (GLU174–GLY177, numeração referente a isoforma 1) é destacada em ciano na cavidade de ligação da proteína.

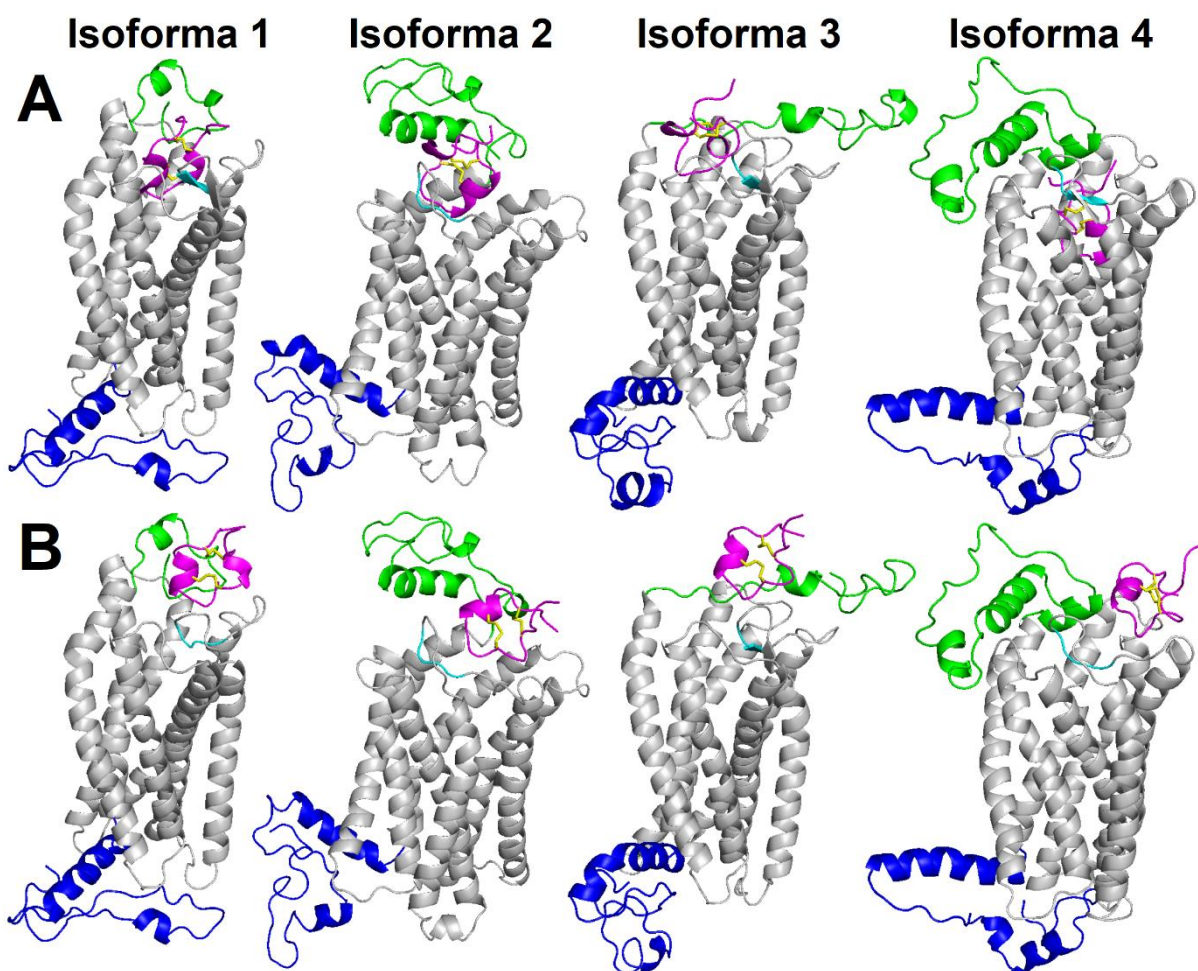


Fonte: Elaborado pelo autor.

O servidor ClusPro apresenta seus resultados em forma de conjuntos, ranqueado-os de 0 à 10. Neste trabalho, os resultados obtidos pelo ClusPro para a modelagem do complexo formado entre o CX3CR1 e o peptídeo da região conservada de cisteínas da proteína G não são todos provenientes do conjunto 0, pois buscou-se por orientações do peptídeo que era análogas aquela discutida para a fractalcina

(CX3CL1, Figura 20). Dessa forma, foi utilizado como resultado final do *docking* o primeiro modelo estrutural do complexo que apresentasse o peptídeo da G direcionado a fita beta GLU174–GLY177 (numeração referente a isoforma 1) na cavidade de ligação na porção N-terminal do receptor celular CX3CR1 (Figura 21-A). Por outro lado, os modelos estruturais do complexo CX3CR1/peptídeo da G dos cálculos de *docking* molecular do HADDOCK foram obtidos a partir do melhor resultado julgado pelo servidor com base na pontuação de sua função *score* (HADDOCK *score*). Nesse caso, todos os resultados do HADDOCK apresentaram a conformação desejada do peptídeo da G orientando o *loop* entre as hélices alfas em direção a fita beta GLU174–GLY177 (numeração referente a isoforma 1) na cavidade de ligação na região N-terminal do CX3CR1 (Figura 21-B). Comparando os resultados de *docking* molecular da Figura 21, é possível notar que, para o modelo estrutural do ClusPro, o peptídeo da região conservada de cisteínas da proteína G encontra-se mais enterrado na cavidade de ligação da porção N-terminal do barril de hélices alfa do CX3CR1, enquanto que para o modelo do HADDOCK, o mesmo encontra-se na superfície da entrada da cavidade. Mesmo com essa diferença de localização na cavidade de ligação da CX3CR1, os modelos estruturais determinados pelo ClusPro e HADDOCK prosseguiram para as simulações de DM em bicamada lipídica de POPC, uma vez que como não há informação estrutural experimental sobre a interação CX3CR1/peptídeo da G, a diversidade conformacional mostradas pelos dois servidores pode ser importante para uma amostragem do espaço de fase na interação do peptídeo com a proteína.

**FIGURA 21:** Resultados dos cálculos de *docking* molecular para o servidor ClusPro (A) e HADDOCK (B). A proteína é apresentada com o modelo de *cartoon*, sendo o barril de hélices alfa colorido em cinza e o N e C-terminal em verde e azul, respectivamente. A fita beta GLU174–GLY177 (numeração referente a isoforma 1) na cavidade de ligação da proteína é destacada em ciano. O peptídeo da região de cisteínas conservadas da proteína G do hRSV é mostrado como modelo de *cartoon* em magenta e suas pontes de dissulfeto são denotadas com o modelo de *sticks* em amarelo.

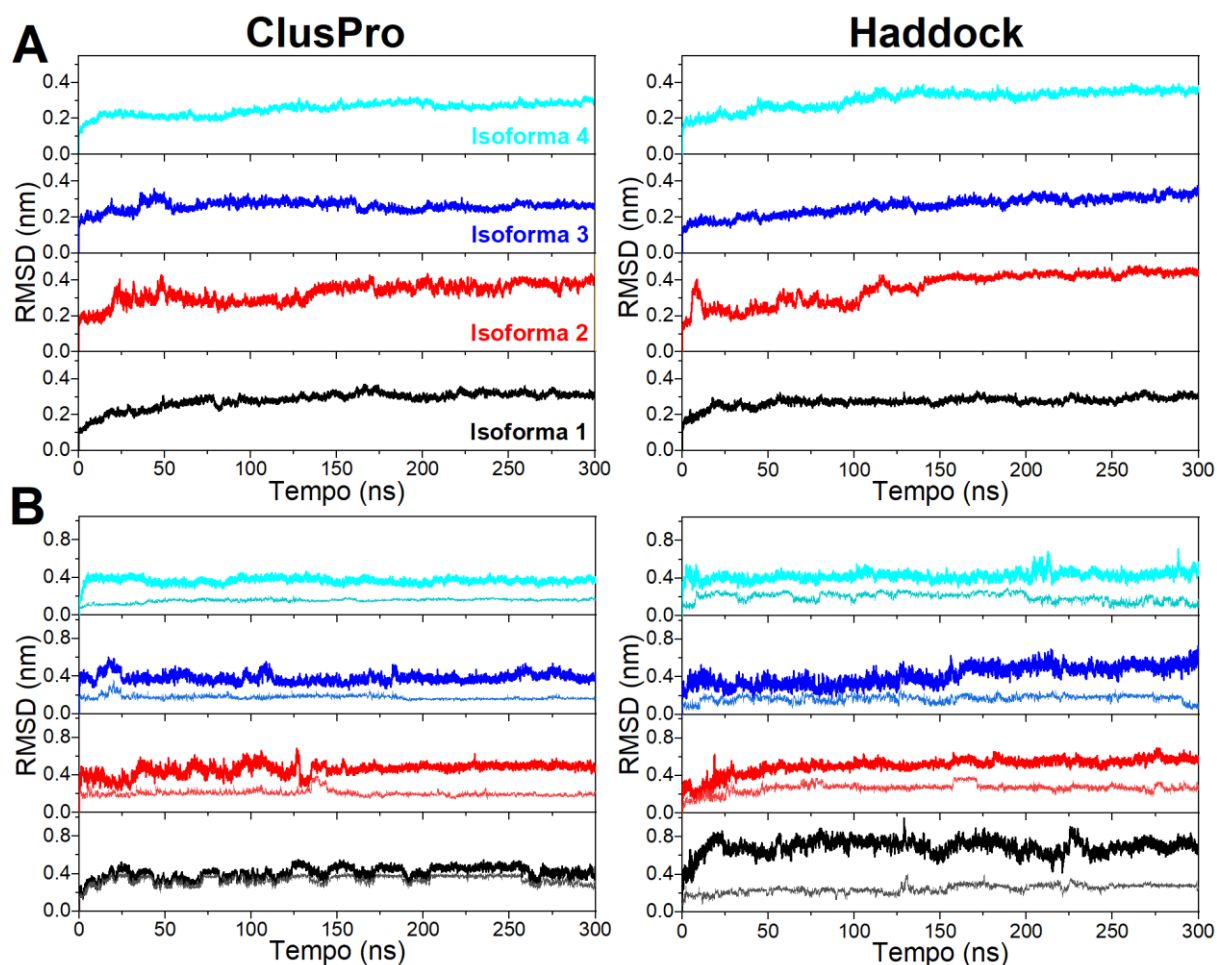


Fonte: Elaborado pelo autor.

Após a realização das simulações de DM de 300 ns dos modelos estruturais dos complexo determinados pelo ClusPro e HADDOCK, as trajetórias foram analisadas com relação aos valores de RMSD da cadeia principal da proteína e do peptídeo da G, distância entre o centro de massa entre a proteína e o peptídeo, número de contatos  $< 0,6$  nm entre qualquer átomo da proteína e do peptídeo, número de ligações de hidrogênio entre o receptor e o peptídeo e porcentagem de persistência das ligações de hidrogênio ao longo de todo o tempo de simulação.

Para as análises de RMSD da Figura 22, foram considerados os átomos da cadeia principal da proteína com alinhamento estrutural realizado para os mesmos, assim como os átomos da cadeia principal do peptídeo da G com alinhamento feito para os mesmos e também para cadeia principal da proteína. Os valores de RMSD para a cadeia principal da proteína (Figura 22-A) mostram que todas as isoformas do CX3CR1 não apresentaram variações conformacionais significativas ( $< 3 \text{ \AA}$ ) ao longo dos 300 ns de simulações, tanto para modelo do complexo determinado pelo ClusPro como para o HADDOCK. Com relação aos valores de RMSD do peptídeo da G, quando o alinhamento estrutural é realizado em sua própria cadeia principal (Figura 22-B, linha fina), pode-se observar para todas as isoformas que os RMSDs estabilizam por volta  $2 \text{ \AA}$ , com exceção da isoforma 1 do ClusPro ( $\sim 4 \text{ \AA}$ ), o que significa que o peptídeo não sofre significativas mudanças conformacionais ao longo do tempo de simulação e preserva sua estrutura inicial. Já com relação aos valores de RMSD para quando o alinhamento é feito tomando como referência a cadeia principal da proteína (Figura 22-B, linha grossa), observa-se variações maiores que alcançam estabilidade por volta de  $4 \text{ \AA}$  para todas as isoformas, com exceção da isoforma 1 do HADDOCK ( $\sim 8 \text{ \AA}$ ), indicando que o peptídeo da G pode ter passado por variações em seu posicionamento com relação a posição inicial na cavidade de ligação na proteína (modelo de *docking*), seja aproximando, afastando e/ou rotacionando.

**FIGURA 22:** Valores de RMSD calculados para os átomos da cadeia principal do proteína e do peptídeo da G ao longo dos 300 ns de simulação de DM para os modelos estruturais dos complexos determinados pelo servidores ClusPro (esquerda) e HADDOCK (direita). Os resultados de cada isoforma é apresentado no seguinte esquema de cores: preto para a isoforma 1, vermelho para a 2, azul para a 3 e ciano para a 4. A linha grossa na parte B denota os valores de RMSD do peptídeo para o alinhamento estrutural realizado com relação a proteína, enquanto que a linha fina se refere a um alinhamento feito com relação ao próprio peptídeo.

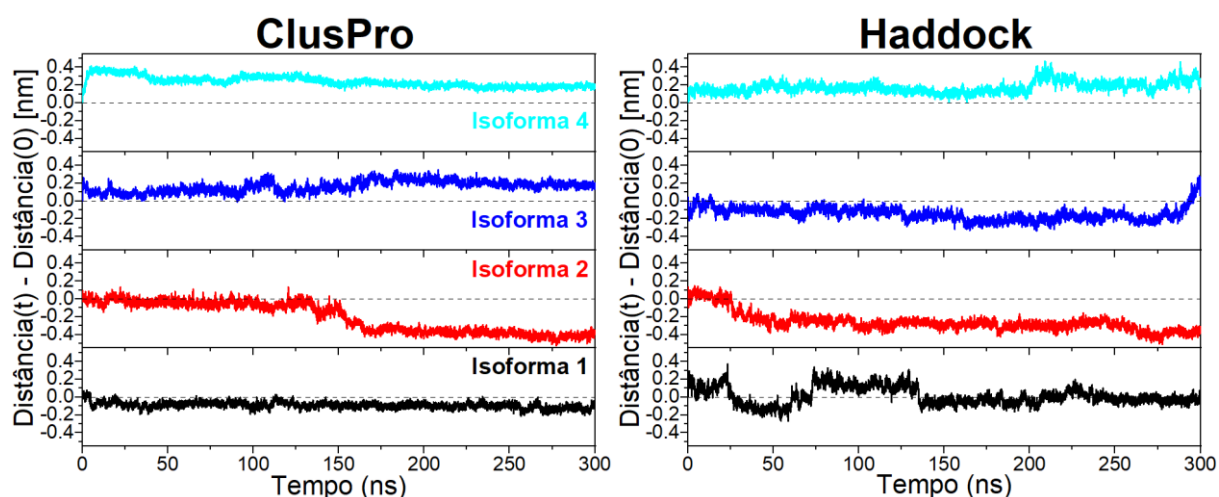


Fonte: Elaborado pelo autor.

Para se determinar o quão o peptídeo da região conservada de cisteínas da G se movimentou em relação a sua posição inicial (modelo de *docking*), calculou-se a distância entre os centro de massas do barril de hélices alfa da proteína e do peptídeo ao longo dos 300 ns, tomando como ponto de referência o ponto inicial da simulação ( $t = 0$  ns). A Figura 23 mostra valores de variação/diferença de distância com relação ao ponto de referência  $t = 0$  ns calculados para os modelos estruturais dos complexos determinados pelo ClusPro e HADDOCK. Para os modelos estruturais determinados

pelo ClusPro (Figura 23, esquerda), observa-se que: (i) para as isoformas 3 e 4 há uma variação positiva de distância de até 2 Å (0,2 nm), indicando que a proteína e o peptídeo se afastaram levemente um do outro; (ii) para a isoforma 1 não há mudanças significativas ( $< 1 \text{ \AA}$ ); e (iii) para a isoforma 2 ocorrem variações negativas de distância após 150 ns de simulação, as quais sugerem que a proteína e o peptídeo se aproximam com distâncias de até 4 Å menores do que aquela da posição inicial ( $t = 0 \text{ ns}$ ). Para os modelos de complexos do HADDOCK (Figura 23, direita), os cálculos de variação/diferença de distância revelam que: (i) para a isoforma 1 há flutuações iniciais que se estabilizam ao final da simulação, mantendo os centros de massa da proteína e do peptídeo com a mesma distância assim como no ponto de referência; (ii) para as isoformas 2 e 3 ocorrem variações negativas de até 2 e 4 Å, respectivamente, o que caracterizam uma mútua aproximação da proteína e do peptídeo; e (iii) para a isoforma 4 há variações positivas de distância de até 3 Å que indicam um leve afastamento entre o receptor celular CX3CR1 e o peptídeo da G.

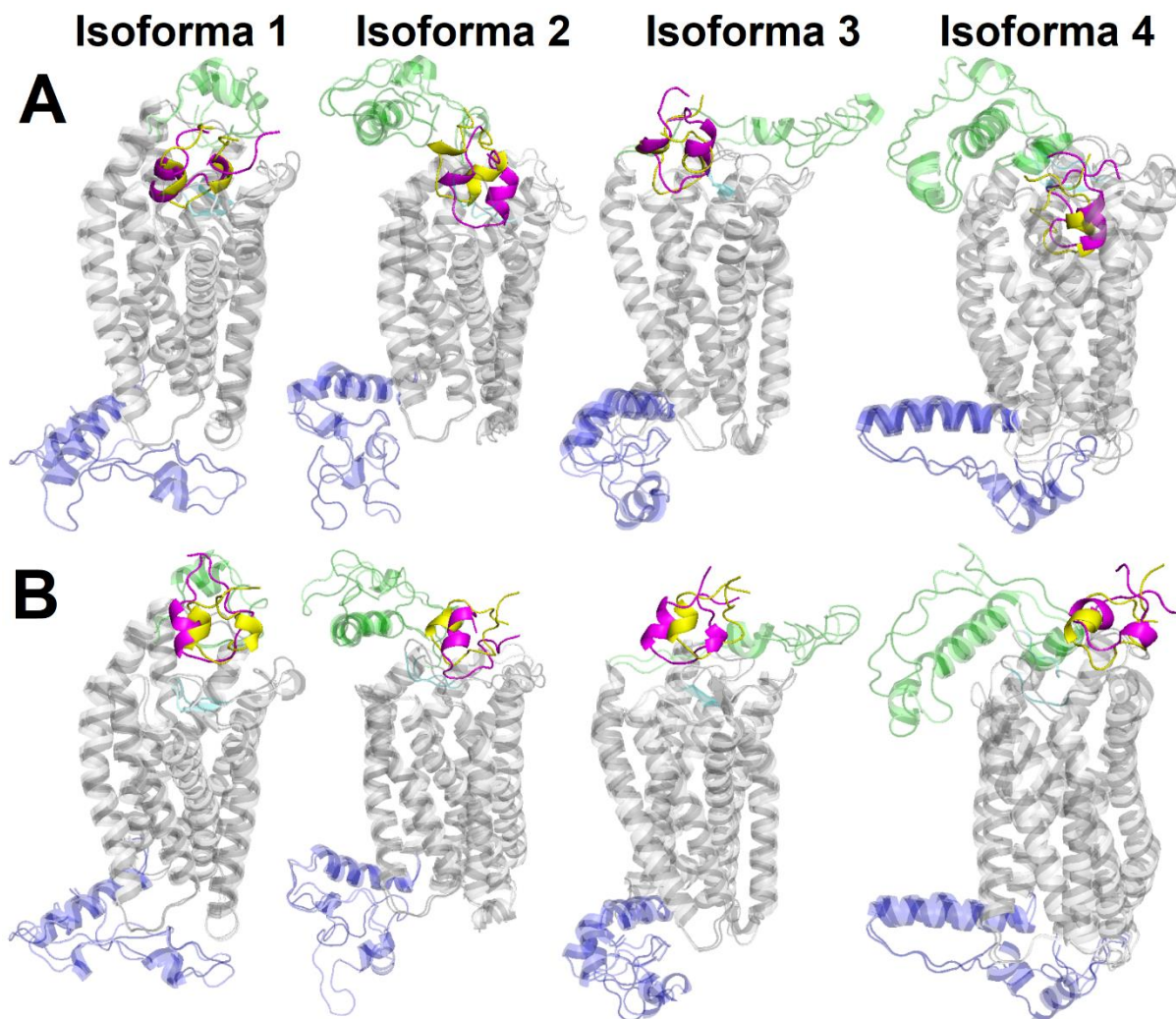
**FIGURA 23:** Valores de variação/diferença de distância (distância num tempo qualquer subtraído pela distância na posição inicial  $t = 0 \text{ ns}$ ) calculadas entre os centros de massa do barril de hélices alfa da proteína e do peptídeo da G ao longo de 300 ns de simulação para os modelos estruturais dos complexos determinados pelo ClusPro (esquerda) e HADDOCK (direita). Os resultados de cada isoforma é apresentado no seguinte esquema de cores: preto para a isoforma 1, vermelho para a 2, azul para a 3 e ciano para a 4.



Fonte: Elaborado pelo autor.

Pode-se notar pelo resultado da Figura 23 e as análises feitas no parágrafo anterior que as variações/diferenças de distância entre o CX3CR1 e o peptídeo da G não são maiores do que 4 Å ao longo dos 300 ns de simulações. Esses resultados indicam que o posicionamento do peptídeo não muda significativamente com relação a posição inicial vinda dos cálculos de *docking* molecular feitos pelos ClusPro e HADDOCK, possivelmente apresenta um afastamento, aproximação e/ou rotação marginal. Para confirmar essa análise e também proporcionar uma inspeção visual da estabilidade estrutural do complexo, foi realizada cálculos de clusterização dos 300 ns da trajetória dos modelos. A partir da Figura 24, é possível observar que comparativamente as posições dos peptídeos nas estruturas representativas (em amarelo) não mudam significativamente com relação a posição inicial determinadas pelos cálculos de *docking* molecular (em magenta), sugerindo que a proposta de interação do peptídeo da G, em uma orientação como a observada para a fractalcina (CX3CL1, Figura 20) na cavidade de ligação na porção N-terminal do barril de hélices alfa das isoformas do CX3CR1 é estável estruturalmente. Isso tanto para uma posição mais enterrada do peptídeo como a predita pelo ClusPro (Figura 24-A), como uma mais na superfície da cavidade de ligação determinada pelo HADDOCK (Figura 24-B).

**FIGURA 24:** Comparação entre as posições do peptídeo da G nos modelos estruturais dos cálculos de *docking* molecular, (A) ClusPRO e (B) HADDOCK, e nas estruturas representativas determinadas pela análise de clusterização das trajetórias de 300 ns de simulações dos complexos. A proteína é apresentada com o modelo de *cartoon*, sendo o barril de hélices alfa colorido em cinza e o N e C-terminal em verde e azul, respectivamente. A fita beta GLU174–GLY177 (numeração referente a isoforma 1) na cavidade de ligação da proteína é destacado em ciano. O peptídeo da região de cisteínas conservadas da proteína G do hRSV é mostrado como modelo de *cartoon* em magenta para o modelo de *docking* e em amarelo para o estrutura representativa da clusterização.

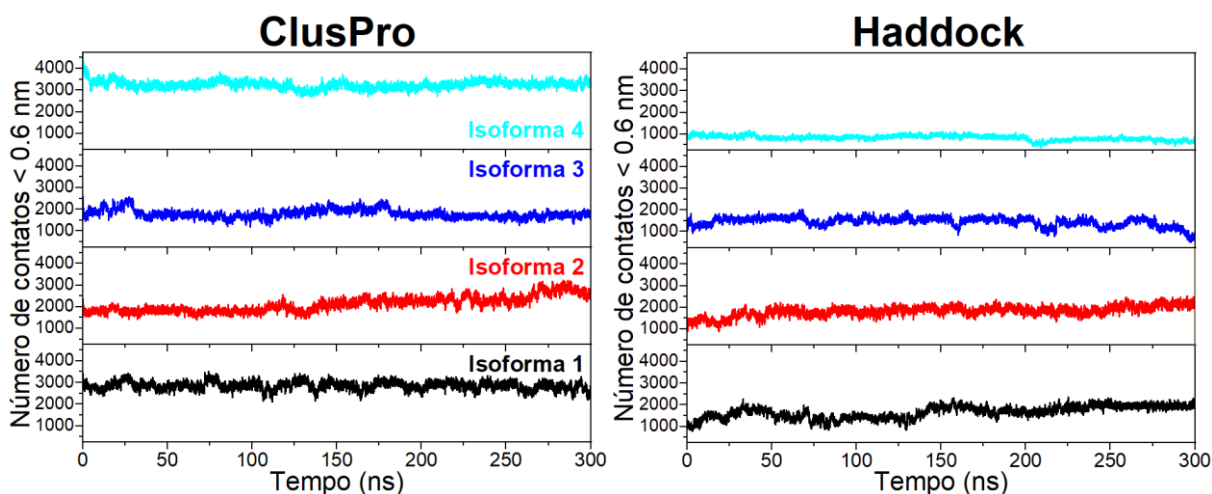


Fonte: Elaborado pelo autor.

A estabilidade estrutural dos modelos dos complexos (ClusPro e HADDOCK) também foi avaliada de acordo com o número de contatos < 0,6 nm calculados ao longo dos 300 ns de simulações de DM. Pode-se observar a partir da Figura 25 que todos os modelos estruturais dos complexos apresentam patamares estáveis de contagem de contatos ao longo de toda a simulação, indicando uma estabilidade

estrutural dos modelos uma vez que em nenhum momento o receptor CX3CR1 e o peptídeo perdem contato. O que também pode ser observado pelos gráficos do número de contatos é uma maior contagem para os resultados com os modelos estruturais dos complexos vindos do ClusPro (esquerda) quando comparado com os do HADDOCK (direita), destacando as isoformas 1 e 4. Esse resultado reforça as afirmações de maior enterramento do peptídeo da G na cavidade de ligação do barril de hélices alfa do CX3CR1 para os modelos estruturais determinados pelo servidor ClusPro.

**FIGURA 25:** Número de contatos menor que 0,6 nm determinados entre os átomos da proteína e do peptídeo da G ao longo dos 300 ns de simulação para os modelos estruturais dos complexos calculados pelo ClusPro (esquerda) e HADDOCK (direita). Os resultados de cada isoforma é apresentado no seguinte esquema de cores: preto para a isoforma 1, vermelho para a 2, azul para a 3 e ciano para a 4.

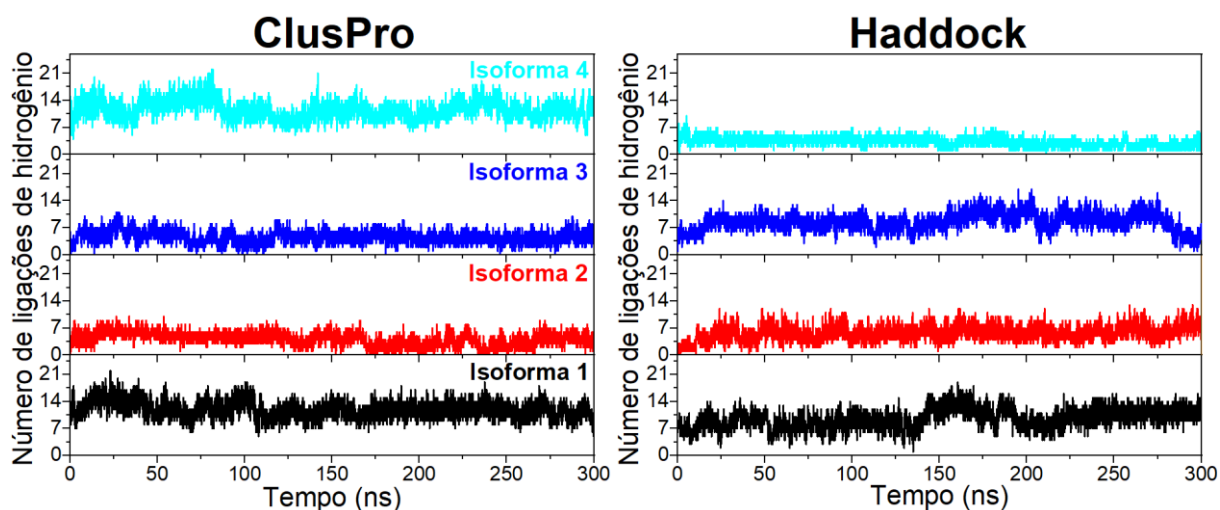


Fonte: Elaborado pelo autor.

A avaliação do número de ligações de hidrogênio formadas entre as isoformas do CX3CR1 e o peptídeo da região conservada de cisteínas da proteína G (Figura 26) revela que esse tipo de interação não covalente representa um aspecto importante na estabilidade da interação entre a proteína e o peptídeo ao longo de todos os 300 ns de simulação de DM, tanto para os modelos estruturais determinados pelo ClusPro como pelo HADDOCK. É possível observar a partir da Figura 26 que as isoformas 1 e 4 do ClusPro apresentam contagens de ligações de hidrogênio maior do que as calculada para o modelo de HADDOCK, o que nesse caso se deve provavelmente ao

maior nível de enterramento do peptídeo na cavidade de ligação na proteína. Por outro lado, para as isoformas 2 e 3 não se pode fazer a mesmo tipo de raciocínio uma vez que os modelos dos complexos CX3CR1/peptídeo da G determinados pelo HADDOCK foram os quais mostraram um maior número de ligações de hidrogênio ao longo dos 300 ns de simulação.

**FIGURA 26:** Número de ligações de hidrogênios determinados entre os átomos da proteína e do peptídeo da G ao longo dos 300 ns de simulação para os modelos estruturais dos complexos calculados pelo ClusPro (esquerda) e HADDOCK (direita). Os resultados de cada isoforma é apresentado no seguinte esquema de cores: preto para a isoforma 1, vermelho para a 2, azul para a 3 e ciano para a 4.



Fonte: Elaborado pelo autor.

Uma outra forma de analisar as ligações de hidrogênio é determinar a porcentagem de persistência delas ao longo dos 300 ns de simulação de DM. As Tabelas de 8 a 15 mostram os resíduos do receptor CX3CR1 e do peptídeo da G que formam ligações de hidrogênio com porcentagens de persistência acima de 10%. Esse limite de porcentagem de persistência de 10% foi usado recentemente com sucesso para análises de complexos proteína-ligantes realizadas pelo nosso grupo de pesquisa em trabalhos como De SÁ *et al.*, 2020 e CARUSO *et al.*, 2021. Vale mencionar que as sequências primárias das isoformas do receptor celular CX3CR1 diferenciam com relação aos resíduos N-terminais, a isoforma 2 e 4 têm 32 resíduos a mais que a isoforma 1, enquanto que a isoforma 3 têm 7 a mais. Essa informação ajuda na interpretação e identificação dos resíduos nas Tabelas 8–15.

Uma análise dos resíduos das isoformas do receptor celular com porcentagens de persistências >10% para os modelos estruturais dos complexos determinados pelo ClusPro revela que a ARG-272 na isoforma 1 (Tabela 8) é recorrente na 3 como ARG-279 (Tabela 10) e na 4 como ARG-304 (Tabela 11). Para a isoforma 2, a ARG-304 apresenta porcentagem de persistência menor do que 10% e portanto não é listada na Tabela 9. Para os modelos vindos do servidor HADDOCK, a GLY-177 na isoforma 1 (Tabela 12) é recorrente na 2 como GLY-209 (Tabela 13) e na 4 como GLY-209 (Tabela 15), sendo que nessa última apresenta uma porcentagem de persistência >96%. Outros dois resíduos recorrentes no receptor CX3CR1 foram TYR-179 e GLN-184 na isoforma 1 (Tabela 12), sendo a TYR-211 e GLN-216 nas isoformas 2 (Tabela 13) e 4 (Tabela 15). Na isoforma 3, somente TYR-186 foi identificada nos dados originais com porcentagens de persistência <10%.

Para os resíduos do peptídeo G marcados com asterisco (\*) observa-se que ARG-188\*, ASN-178\* e SER-174\* é frequente para todas as isoformas do resultado obtido pelo ClusPro apresentando uma persistência maior que 10% (Tabelas 8, 9, 10 e 11), exceto o resíduo SER-174\* para a isoforma 3 não apresentado na tabela pois apresenta uma persistência de 9,2% da ligação de hidrogênio, todavia há essa ligação em algum tempo de simulação só não se mantendo com frequência. Também observa-se que os resíduos SER-177\* e ASN-179\* são recorrentes para as isoformas 1, 2 e 3, apresentando uma porcentagem de persistência maior que 10% (Tabela 12,13 e 14), tanto para os modelos estruturais do ClusPro como HADDOCK. Contudo, para a isoforma 4, a SER-177\* não é mostrada nas Tabelas 11 e 15 pois apresenta uma porcentagem de persistência pouco maior do que 4%, enquanto que para a ASN-179\* não foi detectada qualquer ligação de hidrogênio. Vale destacar que a ASN-178\* do peptídeo da G foi encontrado em todos os resultados tanto para o modelo estrutural vindo do ClusPro quanto do HADDOCK, com porcentagens de persistência que ultrapassam os 65% (Tabelas 10 e 15).

**TABELA 8:** Porcentagens de persistência maior que 10% para as ligações de hidrogênio formadas entre o receptor celular CX3CR1 e o peptídeo da G ao longo dos 300 ns de simulações de DM para o modelo estrutural do complexo obtidos a partir do servidor ClusPro. Para diferenciar da proteína, os resíduos do peptídeo da G são denotados com um asterisco (\*).

<b>Doador</b>	<b>Átomo</b>	<b>Aceitador</b>	<b>Átomo</b>	<b>%Persistência</b>
CYS173*	N	GLU6	OE1	70.484
CYS173*	N	GLU6	OE2	66.418
SER174*	N	GLU6	OE1	59.872
SER174*	N	GLU6	OE2	33.516
ILE175*	N	GLU6	OE1	30.689
ILE175*	N	GLU6	OE2	58.419
SER177*	OG	GLU254	OE1	49.047
SER177*	OG	GLU254	OE2	48.660
ASN178*	ND2	GLU254	OE1	43.114
ASN178*	ND2	GLU254	OE2	43.674
ASN178*	ND2	GLU279	OE1	15.051
ASN178*	ND2	GLU279	OE2	16.291
ASN179*	N	SER276	OG	49.847
ASN179*	ND2	GLU279	OE1	39.181
ASN179*	ND2	GLU279	OE2	43.674
ARG188*	N	GLN3	OE1	23.610
ARG188*	NH1	CYS21	O	28.316
ARG188*	NH2	GLN3	OE1	14.558
ARG188*	NH2	CYS21	O	25.757
ILE189*	O	GLU6	OE1	63.991
ILE189*	O	GLU6	OE2	34.902
GLN3	NE2	SER174*	OG	22.517
GLN3	NE2	CYS186*	O	19.104
GLN184	NE2	PHE170*	O	38.475
ARG191	NH2	CYS176*	O	10.439
ARG191	NH2	SER177*	OG	15.065
ARG191	NH2	ASN178*	OD1	12.985
ARG272	NH1	SER177*	OG	31.809

Fonte: Elaborado pelo autor.

**TABELA 9:** Porcentagens de persistência maior que 10% para as ligações de hidrogênio formadas entre a isoforma 2 do receptor celular CX3CR1 e o peptídeo da G ao longo dos 300 ns de simulações de DM para o modelo estrutural do complexo obtidos a partir do servidor ClusPro. Para diferenciar da proteína, os resíduos do peptídeo da G são denotados com um asterisco (\*).

<b>Doador</b>	<b>Átomo</b>	<b>Aceitador</b>	<b>Átomo</b>	<b>%Persistência</b>
ASN169*	ND2	LEU49	O	16.238
ASN169*	ND2	GLU51	O	27.836
ASN169*	ND2	CYS53	O	44.314
PHE170*	N	LEU49	O	37.182
ASN178*	ND2	TYR122	O	24.623
ASN178*	ND2	CYS207	O	10.092
THR181*	OG1	TYR54	OH	40.288
ARG188*	NH1	TYR211	OH	10.932
TYR122	OH	SER174*	O	20.677

Fonte: Elaborado pelo autor.

**TABELA 10:** Porcentagens de persistência maior que 10% para as ligações de hidrogênio formadas entre a isoforma 3 do receptor celular CX3CR1 e o peptídeo da G ao longo dos 300 ns de simulações de DM para o modelo estrutural do complexo obtidos a partir do servidor ClusPro. Para diferenciar da proteína, os resíduos do peptídeo da G são denotados com um asterisco (\*).

<b>Doador</b>	<b>Átomo</b>	<b>Aceitador</b>	<b>Átomo</b>	<b>%Persistência</b>
SER177*	OG	LEU98	O	81.802
ASN178*	N	TYR97	O	66.378
ASN179*	ND2	ARG279	O	22.357
ASN179*	ND2	SER283	OG	20.837
ARG188*	NE	GLU181	OE1	11.918
ARG188*	NE	GLU181	OE2	10.012
ARG188*	NH1	GLU101	O	11.998
ARG188*	NH2	GLU101	O	11.079
ILE189*	O	GLU101	OE1	12.945
ILE189*	O	GLU101	OE2	17.598
TYR29	OH	PRO172*	O	13.811
ARG279	NE	ASN178*	OD1	24.170
ARG279	NH2	ASN178*	OD1	11.159
SER283	OG	ASN178*	OD1	11.452

Fonte: Elaborado pelo autor.

**TABELA 11:** Porcentagens de persistência maior que 10% para as ligações de hidrogênio formadas entre a isoforma 4 do receptor celular CX3CR1 e o peptídeo da G ao longo dos 300 ns de simulações de DM para o modelo estrutural do complexo obtidos a partir do servidor ClusPro. Para diferenciar da proteína, os resíduos do peptídeo da G são denotados com um asterisco (\*).

<b>Doador</b>	<b>Átomo</b>	<b>Aceitador</b>	<b>Átomo</b>	<b>%Persistência</b>
ASN169*	N	THR287	OG1	13.478
ASN169*	N	ASP292	OD1	11.385
ASN169*	ND2	TYR211	OH	19.051
ASN169*	ND2	ASP292	OD1	18.491
ASN169*	ND2	ASP292	OD2	21.810
SER177*	OG	GLU311	OE1	39.821
SER177*	OG	GLU311	OE2	14.825
ASN178*	N	GLU311	OE1	26.410
ASN178*	N	GLU311	OE2	23.197
ASN178*	ND2	TYR70	OH	11.159
ASN178*	ND2	THR310	O	24.023
ASN178*	ND2	GLU311	OE1	15.665
ASN178*	ND2	GLU311	OE2	15.278
ASN179*	ND2	THR227	OG1	46.434
ASN179*	ND2	TYR279	OH	75.497
THR181*	OG1	ASN224	OD1	10.999
THR181*	OG1	GLU226	OE2	45.661
THR181*	OG1	THR227	OG1	20.184
LYS187*	NZ	GLY209	O	45.421
LYS187*	NZ	ASP210	OD1	13.545
LYS187*	NZ	ASP210	OD2	14.371
LYS187*	NZ	TYR211	OH	17.651
ARG188*	N	CYS207	O	96.121
ILE189*	N	GLU206	OE1	32.742
ILE189*	N	GLU206	OE2	65.298
GLY209	N	CYS186*	O	88.748

TYR211	OH	ASN169*	OD1	21.064
ASN228	ND2	ASN179*	OD1	13.012
ARG300	NH1	ASN169*	OD1	21.370
ARG300	NH2	ASN169*	OD1	15.118
ARG304	NE	SER174*	OG	27.076

Fonte: Elaborado pelo autor.

**TABELA 12:** Porcentagens de persistência maior que 10% para as ligações de hidrogênio formadas entre o receptor celular CX3CR1 e o peptídeo da G ao longo dos 300 ns de simulações de DM para o modelo estrutural do complexo obtidos a partir do servidor HADDOCK. Para diferenciar da proteína, os resíduos do peptídeo da G são denotados com um asterisco (\*).

Doador	Átomo	Aceitador	Átomo	%Persistência
CYS173*	N	GLU10	OE1	41.861
CYS173*	N	GLU10	OE2	41.381
SER174*	N	GLU10	OE1	45.181
SER174*	N	GLU10	OE2	43.141
SER174*	OG	GLU10	OE1	33.062
SER174*	OG	GLU10	OE2	33.622
SER177*	OG	GLU6	OE1	12.718
SER177*	OG	GLU6	OE2	11.678
ASN178*	N	GLU6	O	21.837
ASN178*	ND2	GLU6	OE1	14.531
ASN178*	ND2	GLU6	OE2	14.865
ASN178*	ND2	GLY177	O	23.850
ASN178*	ND2	TYR179	OH	44.394
ASN179*	ND2	LYS171	O	11.812
ASN179*	ND2	GLU172	OE1	17.051
ASN179*	ND2	GLU172	OE2	16.318
ASN179*	ND2	GLU174	OE1	23.863
ASN179*	ND2	GLU174	OE2	31.276
THR181*	N	GLU174	OE1	44.967
THR181*	N	GLU174	OE2	18.371
THR181*	OG1	GLU172	OE1	15.985
THR181*	OG1	GLU172	OE2	16.984
THR181*	OG1	GLU174	OE1	49.740
THR181*	OG1	GLU174	OE2	45.301
TRP183*	NE1	ASP2	O	18.011
MET1	N	TRP183*	O	11.972
MET1	N	CYS186*	O	18.691
GLN3	NE2	PRO180*	O	20.517
LYS171	NZ	SER177*	OG	10.985
LYS171	NZ	SER177*	O	12.265
LYS171	NZ	ASN178*	OD1	20.677
LYS171	NZ	ASN178*	O	23.664
TYR179	N	ASN178*	OD1	12.185
GLN184	NE2	SER174*	O	11.812

Fonte: Elaborado pelo autor.

**TABELA 13:** Porcentagens de persistência maior que 10% para as ligações de hidrogênio formadas entre a isoforma 2 do receptor celular CX3CR1 e o peptídeo da G ao longo dos 300 ns de simulações de DM para o modelo estrutural do complexo obtidos a partir do servidor HADDOCK. Para diferenciar da proteína, os resíduos do peptídeo da G são denotados com um asterisco (\*).

<b>Doador</b>	<b>Átomo</b>	<b>Aceitador</b>	<b>Átomo</b>	<b>%Persistência</b>
SER177*	N	TYR211	OH	73.350
SER177*	OG	TYR211	OH	26.010
ASN179*	ND2	GLY209	O	76.496
ASN179*	ND2	ASP210	OD1	26.690
ASN179*	ND2	ASP210	OD2	25.117
ILE189*	N	LEU49	O	44.514
LYS203	NZ	THR181*	OG1	12.425
GLN216	NE2	ASN169*	O	19.784
ARG300	NH2	CYS176*	O	21.904
ARG304	NH1	ASN178*	OD1	22.064
ARG304	NH2	ASN178*	OD1	18.264

Fonte: Elaborado pelo autor.

**TABELA 14:** Porcentagens de persistência maior que 10% para as ligações de hidrogênio formadas entre a isoforma 3 do receptor celular CX3CR1 e o peptídeo da G ao longo dos 300 ns de simulações de DM para o modelo estrutural do complexo obtidos a partir do servidor HADDOCK. Para diferenciar da proteína, os resíduos do peptídeo da G são denotados com um asterisco (\*).

<b>Doador</b>	<b>Átomo</b>	<b>Aceitador</b>	<b>Átomo</b>	<b>%Persistência</b>
SER174*	N	ASP23	OD1	18.144
SER174*	N	ASP23	OD2	15.651
SER174*	OG	ASP23	OD1	16.918
SER174*	OG	ASP23	OD2	15.905
SER177*	N	GLU181	OE1	54.086
SER177*	N	GLU181	OE2	26.530
SER177*	OG	ASP22	OD1	14.371
SER177*	OG	ASP22	OD2	12.452
ASN178*	N	GLU181	OE1	59.472
ASN178*	N	GLU181	OE2	38.235
ASN178*	ND2	GLU181	OE1	19.584
ASN178*	ND2	GLU181	OE2	23.104
ASN179*	N	GLU181	OE1	55.073
ASN179*	N	GLU181	OE2	25.983
ASN179*	ND2	GLU101	OE1	40.968
ASN179*	ND2	GLU101	OE2	43.381
THR181*	OG1	GLU101	OE1	23.264
THR181*	OG1	GLU101	OE2	26.810
ARG188*	NH1	ASP23	O	18.011
ILE189*	O	ASP23	OD1	20.117
ILE189*	O	ASP23	OD2	22.424
CYS28	N	ASN178*	OD1	11.398
CYS28	N	ASN178*	O	35.155
TYR29	N	ASN178*	O	14.611
LYS178	NZ	SER174*	O	11.425
LYS178	NZ	SER177*	OG	17.638
ASN180	ND2	ILE175*	O	10.972
ASN180	ND2	ASN179*	OD1	25.130

Fonte: Elaborado pelo autor.

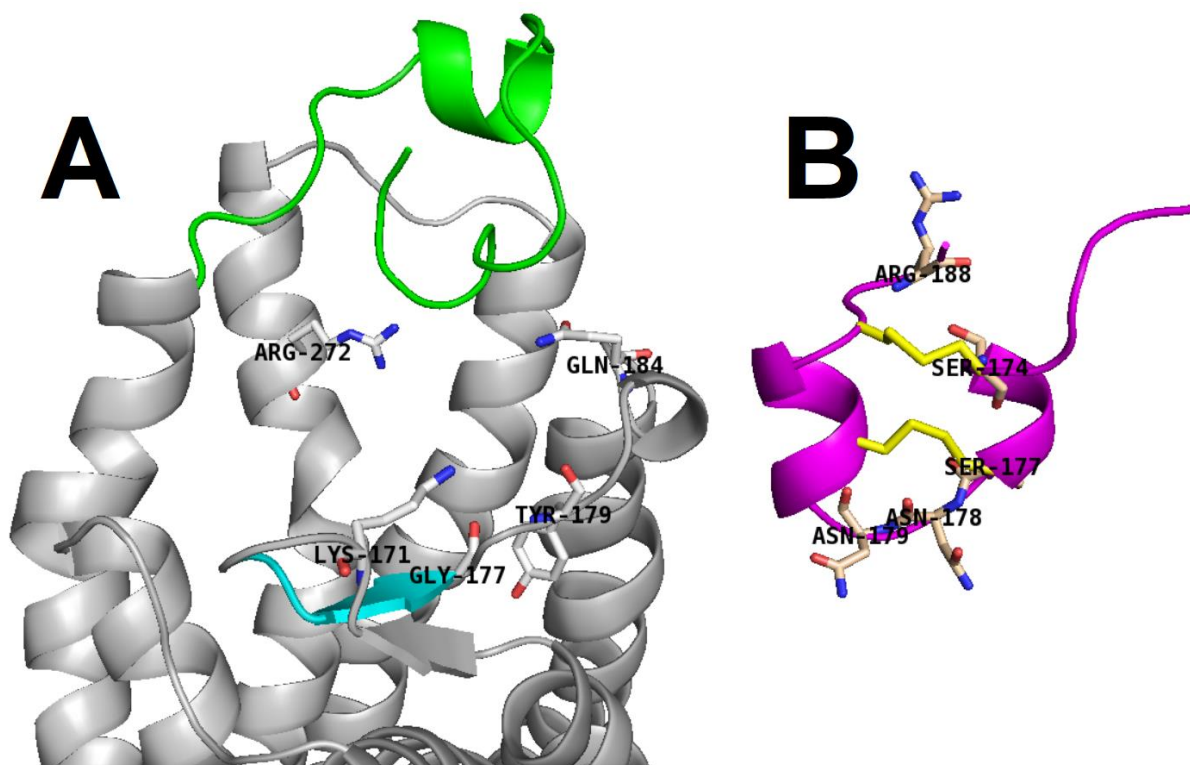
**TABELA 15:** Porcentagens de persistência maior que 10% para as ligações de hidrogênio formadas entre a isoforma 4 do receptor celular CX3CR1 e o peptídeo da G ao longo dos 300 ns de simulações de DM para o modelo estrutural do complexo obtidos a partir do servidor HADDOCK. Para diferenciar da proteína, os resíduos do peptídeo da G são denotados com um asterisco (\*).

<b>Doador</b>	<b>Átomo</b>	<b>Aceitador</b>	<b>Átomo</b>	<b>%Persistência</b>
SER174*	OG	GLU217	OE1	33.982
SER174*	OG	GLU217	OE2	31.916
ASN178*	N	GLN216	OE1	21.117
ASN178*	ND2	GLY209	O	96.947
LYS203	NZ	ASN178*	OD1	34.782
TYR211	N	ASN178*	OD1	68.911

Fonte: Elaborado pelo autor.

Os resíduos do receptor celular CX3CR1 que apareceram recorrentes nas isoformas encontram-se em sua maioria (LYS-171, GLY-177, TYR-179 e GLN-184, numeração referente a isoforma 1) do lado oposto a região N-terminal da proteína e próximos a fita beta (GLU-174–GLY-177, numeração referente a isoforma 1), com exceção a ARG-272 (na isoforma 1) que encontra-se mais internamente na cavidade de ligação do barril de hélices alfa da proteína, indicando um papel importante para o aterramento do peptídeo nesta região (Figura 27-A). Para o peptídeo da G, os resíduos ARG-188\*, SER-174\*, SER177\*, ASN-178\* e ASN-179\* (Figura 27-B) apresentaram significativa recorrência para a análise da porcentagem de persistência de ligações de hidrogênio (Tabelas 8–15). Notadamente, esses resíduos encontram-se na metade N-terminal do peptídeo da G, com exceção da ARG-188\* que está no C-terminal. Esse resultado sugere que a hélice alfa N-terminal do peptídeo apresenta um papel fundamental para a interação estável com a proteína via ligações de hidrogênio.

**FIGURA 27:** Representação esquemática em cartoon do receptor celular CX3CR1 e peptídeo G e em sticks os resíduos mais recorrentes da persistência das ligações de hidrogênios. Em cinza o receptor celular, com o N-terminal em verde, em ciano a fita  $\beta$  e os resíduos recorrentes na forma de sticks, carbonos em cinza, nitrogênio em azul e em vermelho oxigênio (A). Em magenta o peptídeo G, sendo as cisteínas representadas em amarelo na forma de sticks e os resíduos recorrentes em *sticks*, sendo carbono em bege, nitrogênio azul e oxigênio em vermelho (B).



Fonte: Elaborado pelo autor.

Para uma melhor compreensão de como a ligação do peptídeo da G altera a dinâmica das isoformas do receptor celular CX3CR1 no complexo CX3CR1/peptídeo da G, análises de RMSF e PCA foram realizadas das trajetórias dos 300 ns da simulação de DM. A Figura 28 mostra que CX3CR1 livre exibiu valores parecidos de RMSF para resíduos nas regiões N- e C-terminais, bem como a alça entre as folhas  $\beta_1$  e  $\beta_2$  e as alças entre as hélices  $\alpha$ , sugerindo grande flexibilidade conformacional dessas regiões. Além disso, para essas regiões, as porções N- e C-terminais são especialmente notáveis no contexto dos complexos proteína/peptídeo formados, visto que essas regiões tiveram um aumento da flexibilidade, exceto para a região N-terminal da simulação de DM das isoformas 3 e 4 (Figura 28). Para a isoforma 1, o

modelo estrutural do ClusPro apresentou uma maior flexibilidade nas regiões terminais como também nas regiões das folhas  $\beta_1$  e  $\beta_2$ , enquanto os valores de RMSF para o modelo do HADDOCK se mantiveram mais semelhante à forma livre. É interessante notar que a região de alça entre as hélices  $\alpha_6$  e  $\alpha_7$  e a porção N-terminal da hélice  $\alpha_7$  de ambos os complexos modelados pelo ClusPro e HADDOCK apresentam diminuição da flexibilidade em relação a CX3CR1 livre (Figura 28-A).

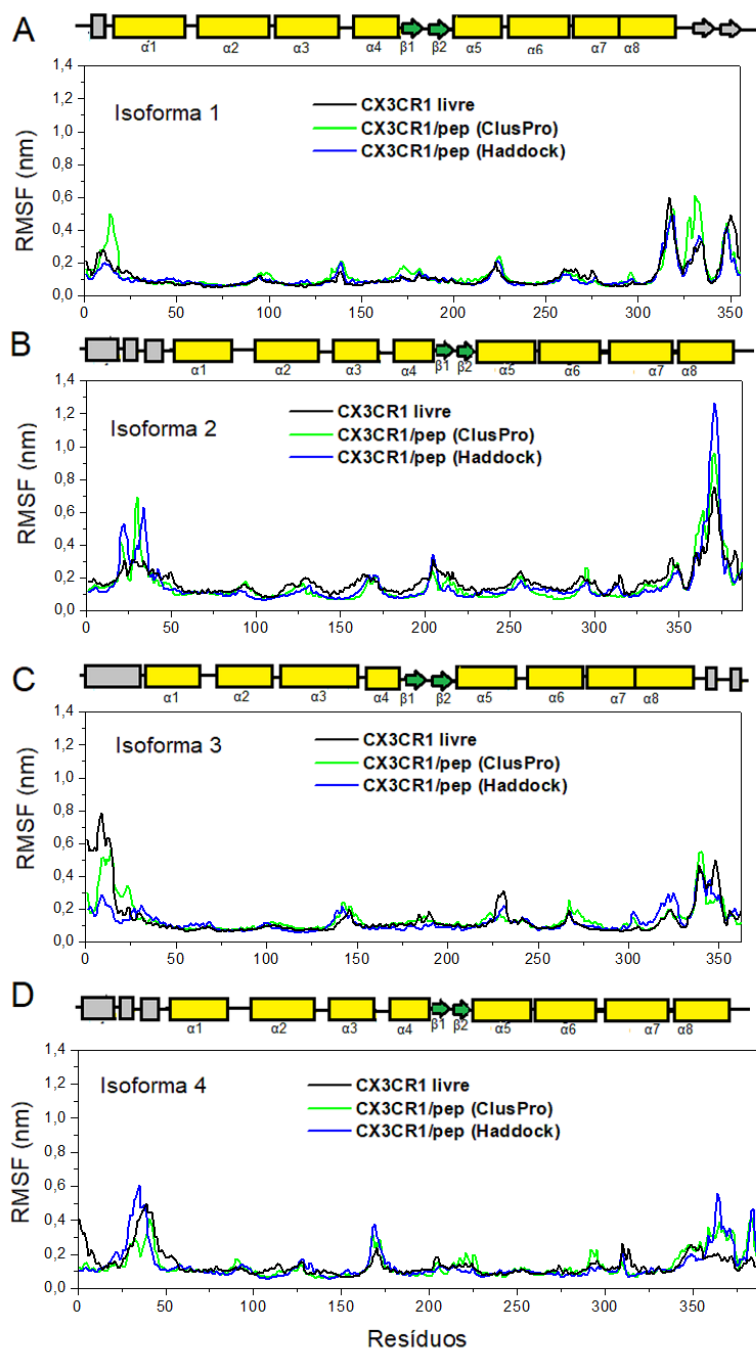
A isoforma 2 se comportou de maneira diferente da isoforma 1, pois a forma livre apresentou uma maior flexibilidade em relação as formas ligadas tanto do ClusPro quanto do HADDOCK, exceto para as regiões terminais, as quais mostraram maior grau de liberdade para as formas complexadas com o peptídeo da G (Figura 28-B). A isoforma 3 apresentou maior flexibilidade para a região N-terminal para a forma livre comparando com as formas complexadas vindas do ClusPro e HADDOCK, enquanto o C-terminal e as demais regiões da proteína se mantiveram parecidas entre as formas livre e ligadas (Figura 28-C). A isoforma 4 apresentou um comportamento semelhante de flexibilidade para o N-terminal e demais regiões da proteína livre e ligada ao peptídeo da G, enquanto a porção dos resíduos C-terminal mostraram uma maior flexibilidade para os modelos estruturais do ClusPro e HADDOCK em relação a forma livre. Vale destacar uma exceção de aumento de flexibilidade conformacional para a região entre alça entre as hélices  $\alpha_3$  e  $\alpha_4$  dos modelos estruturais do ClusPro e HADDOCK quando comparados com a proteína livre (Figura 28-D).

No geral, observa-se uma persistência na flexibilidade das alça entre as hélices  $\alpha$  das isoformas do CX3CR1 tanto para a forma livre como a ligada modelada pelos programas ClusPro e HAADOCK; entretanto, as regiões N e C terminais quando ligadas ao peptídeo da G mostraram uma maior variabilidade na flexibilidade conformacional. Nesse sentido, para observar um pouco mais desses movimentos internos do receptor celular, uma análise de componentes principais (PCA) foram realizadas com a finalidade de comparação entre as isoformas livres e seus complexos com o peptídeo da G modelados pelo ClusPro e HAADOCK.

O gráfico de dispersão PC1 & PC2 para a análise de PCA gerado para a isoforma 1 do CX3CR1 livre e na sua forma complexada com o peptídeo da G modelada pelo ClusPro revelou que a forma livre tem um espaço conformacional mais compacto com relação a forma ligada (Figura 29-A). Pode ser observada a partir dos

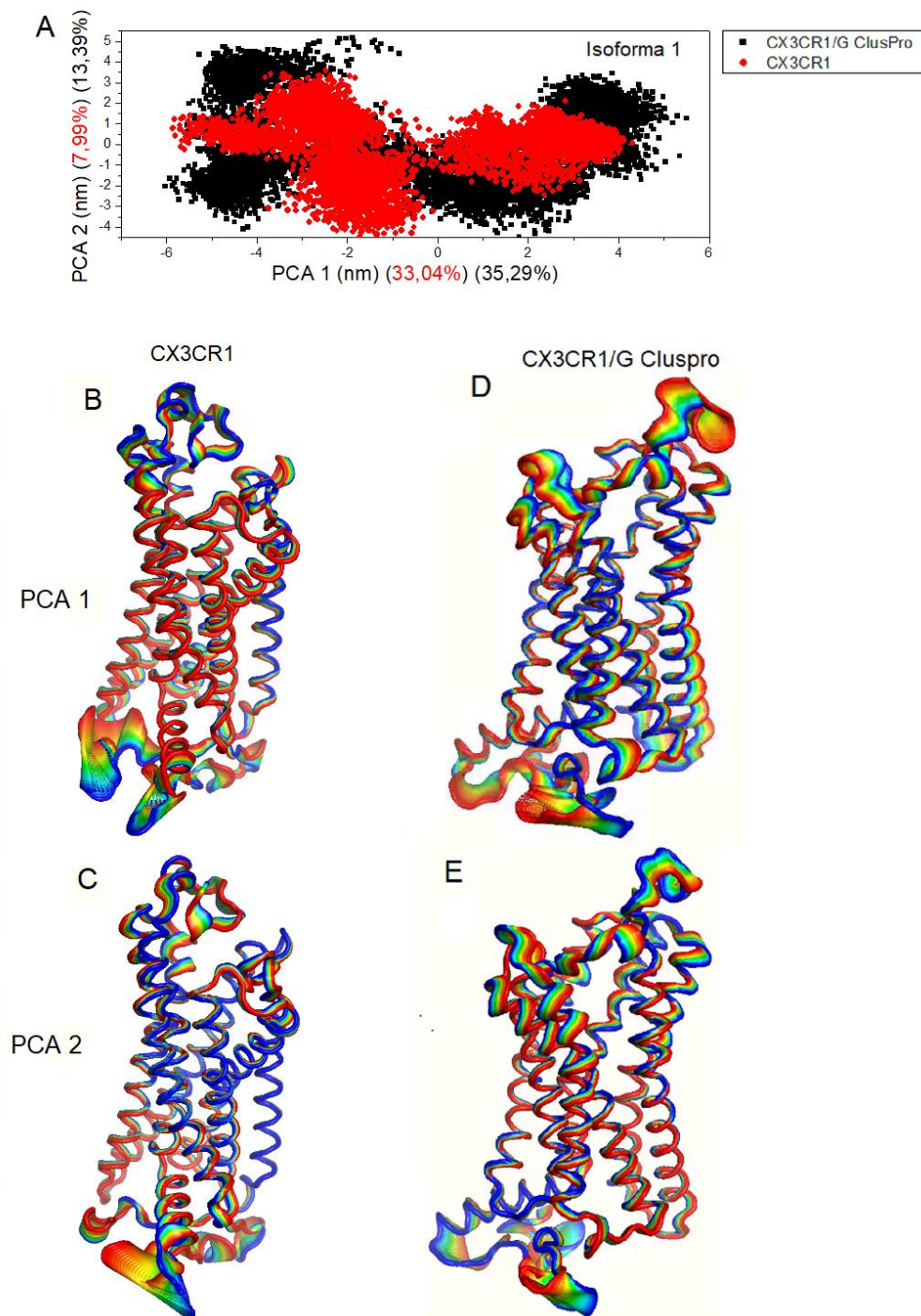
movimentos filtrados para o conjunto de autovetores de PC1 e PC2, para os quais um gradiente de cores do vermelho para azul (ou vice-versa) denotam um movimento de uma determinada região da proteína. Desse modo, o ganho de espaço conformacional da isoforma 1 complexada com o peptídeo da G é devido a um aumento de flexibilidade nos resíduos N- e C-terminais quando comparados com a forma livre da proteína (Figura 29-B e D) tanto para a PC1 como para PC2 (Figura 29-C e E). Esse resultado está em acordo com análises de RMSF. É possível notar também que a região N-terminal do barril de hélices  $\alpha$  envolvidos na interação com o peptídeo da G também ganham maior movimento quando comparados com a isoforma 1 no seu estado livre. Também é notável a semelhança dos movimentos das hélices  $\alpha$  do barril para a forma livre e ligada da isoforma 1 (Figura 29-B, C, D e E).

**FIGURA 28:** Valores de RMSF em função do número de resíduos para as quatro isoformas, isoforma 1 (A), isoforma 2 (B), isoforma 3 (C) e isoforma 4 (D) do receptor celular CX3CR1, com a forma livre do receptor celular CX3CR1 (preto), sua estrutura na forma de complexo CX3CR1/G obtida pela DM do resultado do ClusPro (verde) e sua estrutura na forma de complexo CX3CR1/G obtida pela DM do resultado do HADDOCK (azul). A estrutura secundária está mostrada ao topo de cada isoforma, sendo as hélices  $\alpha$  do barril em amarelo, as folhas  $\beta$  em verde, e em cinza a estrutura secundária do N-terminal e C-terminal. Os cálculos de RMSF foram realizados para os 300 ns de simulação de dinâmica molecular.



Fonte: Elaborado pelo autor.

**FIGURA 29:** Gráficos de dispersão PCA, PCA1 e PCA2 para CX3CR1 livre (pontos vermelhos) e para complexo CX3CR1/G da DM dado pelo *docking* do ClusPro (pontos pretos) da isoforma 1 (A). (B e C) Moções filtradas de os autovetores de PCA1 (B) e PCA2 (C) para a dinâmica dados de CX3CR1 em sua forma livre e na sua forma de complexo CX3CR1/G (D e E) . O movimento é indicada pela variação de cor de azul para vermelho. Para realizar os cálculos de PCA, os 300 ns de trajetórias foram concatenados para cada uma das moléculas sistemas (CX3CR1 livre ou CX3CR1/G).

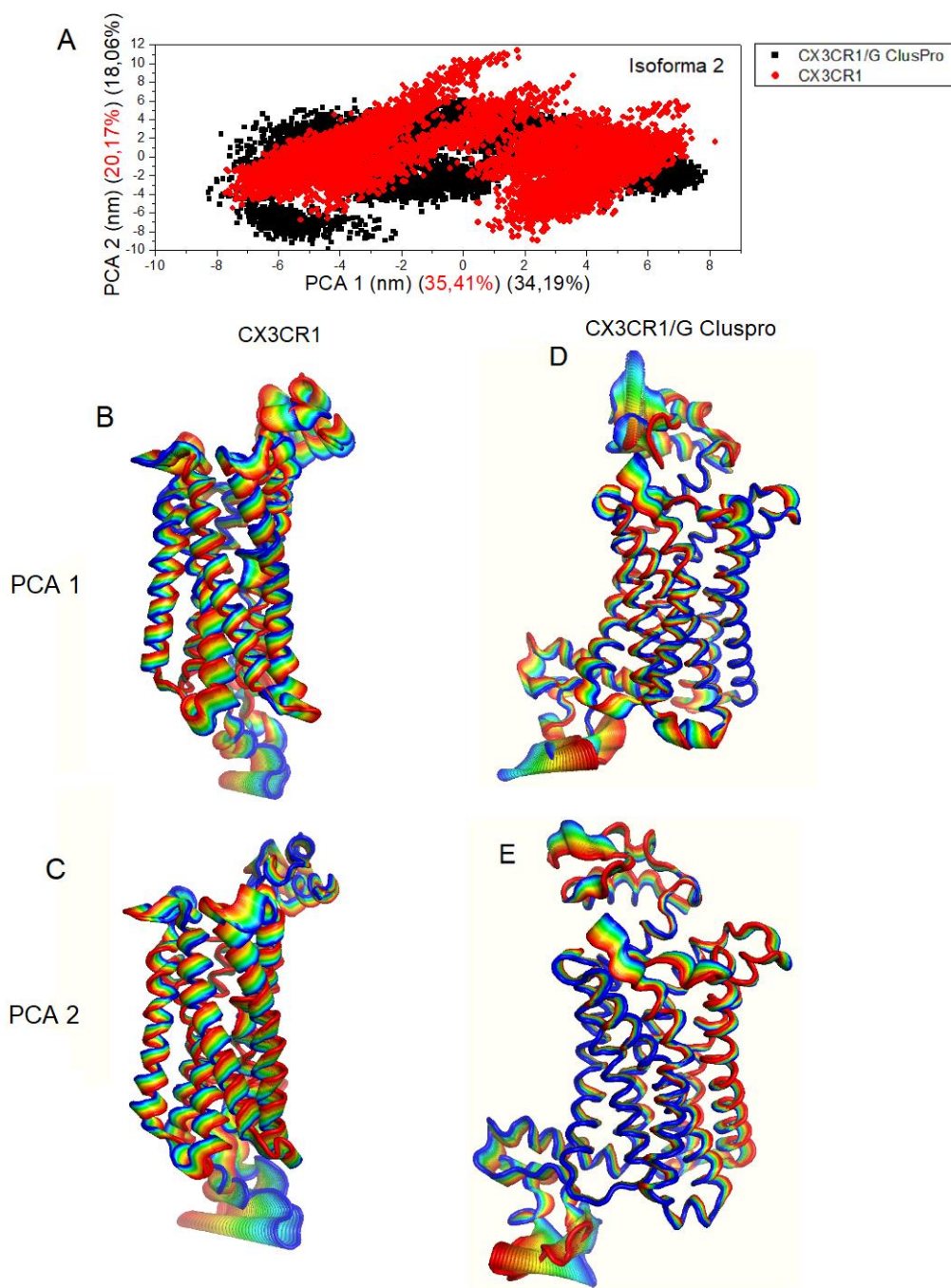


Fonte: Elaborado pelo autor.

O gráfico de dispersão PC1 & PC2 para a análise de PCA gerado para a isoforma 2 do CX3CR1 livre e na sua forma complexada com o peptídeo da G modelada pelo ClusPro revelou que a forma complexada tem um espaço conformacional mais abrangente com relação a forma livre (Figura 30-A). Pode ser observada a partir dos movimentos filtrados para o conjunto de autovetores, o ganho de espaço conformacional da isoforma 2 complexada com o peptídeo da G é devido a um aumento de flexibilidade nos resíduos N- e C-terminais quando comparados com a forma livre da proteína (Figura 30-B e D) tanto para a PC1 como para PC2 (Figura 30-C e E). Esse resultado está em acordo com análises de RMSF. É possível notar também que a região N-terminal do barril de hélices  $\alpha$  envolvidos na interação com o peptídeo da G também ganham maior movimento quando comparados com a isoforma 2 no seu estado livre. Também é notável uma diminuição dos movimentos das hélices  $\alpha$  do barril para a forma complexada em relação a forma livre da isoforma 2 (Figura 30-B, C, D e E), e isto está de acordo como gráfico de RMSF (Figura 28-B).

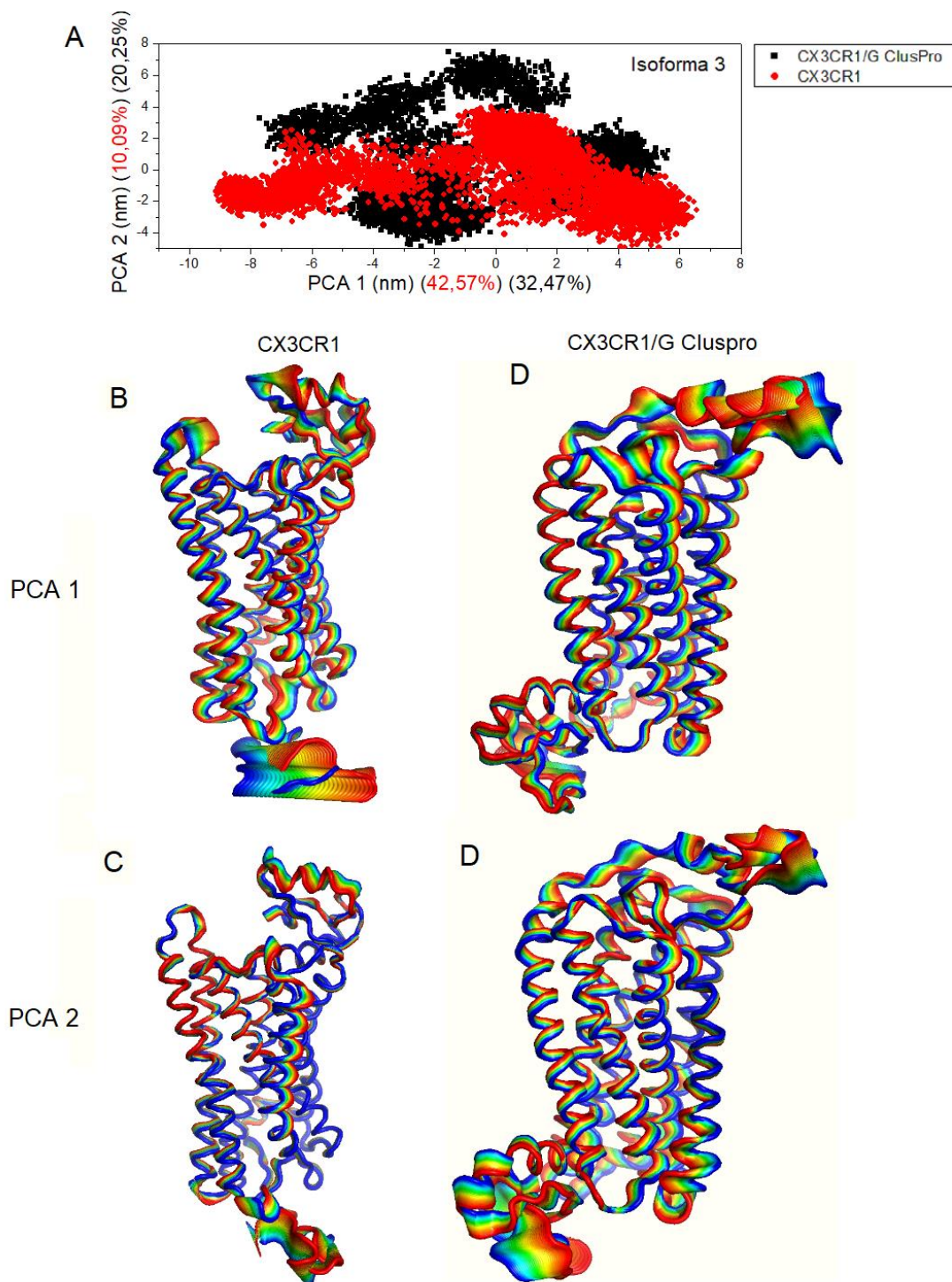
O gráfico gerado para a isoforma 3 do CX3CR1 livre e na sua forma complexada com o peptídeo da G modelada pelo ClusPro revelou que a forma livre tem um espaço conformacional mais compacto com relação a forma ligada (Figura 31-A). Pode ser observada a partir dos movimentos filtrados para o conjunto de autovetores, o ganho de espaço conformacional da isoforma 3 complexada com o peptídeo da G é devido a um aumento de flexibilidade nos resíduos N-terminais quando comparados com a forma livre da proteína (Figura 31-B e D) tanto para a PC1 como para PC2 em relação ao C-terminal mostrou uma forma mais compacta para PC1 e ocupando mais espaço conformacional para PC2 (Figura 31-C e E). É possível notar também que a região N-terminal do barril de hélices  $\alpha$  envolvidos na interação com o peptídeo da G também ganham maior movimento quando comparados com a isoforma 3 no seu estado livre. Também é notável um aumento dos movimentos das hélices  $\alpha$  do barril para a forma complexada em relação a forma livre da isoforma 3 (Figura 31-B, C, D e E), e isto está de acordo como gráfico de RMSF (Figura 28-C).

**FIGURA 30:** Gráficos de dispersão PCA, PCA1 e PCA2 para CX3CR1 livre (pontos vermelhos) e para complexo CX3CR1/G da DM dado pelo *docking* do ClusPro (pontos pretos) da isoforma 2 (A). (B e C) Moções filtradas de os autovetores de PCA1 (B) e PCA2 (C) para a dinâmica dados de CX3CR1 em sua forma livre e na sua forma de complexo CX3CR1/G (D e E) . O movimento é indicada pela variação de cor de azul para vermelho. Para realizar os cálculos de PCA, os 300 ns de trajetórias foram concatenados para cada uma das moléculas sistemas (CX3CR1 livre ou CX3CR1/G).



Fonte: Elaborado pelo autor.

**FIGURA 31:** Gráficos de dispersão PCA, PCA1 e PCA2 para CX3CR1 livre (pontos vermelhos) e para complexo CX3CR1/G da DM dado pelo *docking* do ClusPro (pontos pretos) da isoforma 3 (A). (B e C) Moções filtradas de os autovetores de PCA1 (B) e PCA2 (C) para a dinâmica dados de CX3CR1 em sua forma livre e na sua forma de complexo CX3CR1/G (D e E). O movimento é indicada pela variação de cor de azul para vermelho. Para realizar os cálculos de PCA, os 300 ns de trajetórias foram concatenados para cada uma das moléculas sistemas (CX3CR1 livre ou CX3CR1/G).

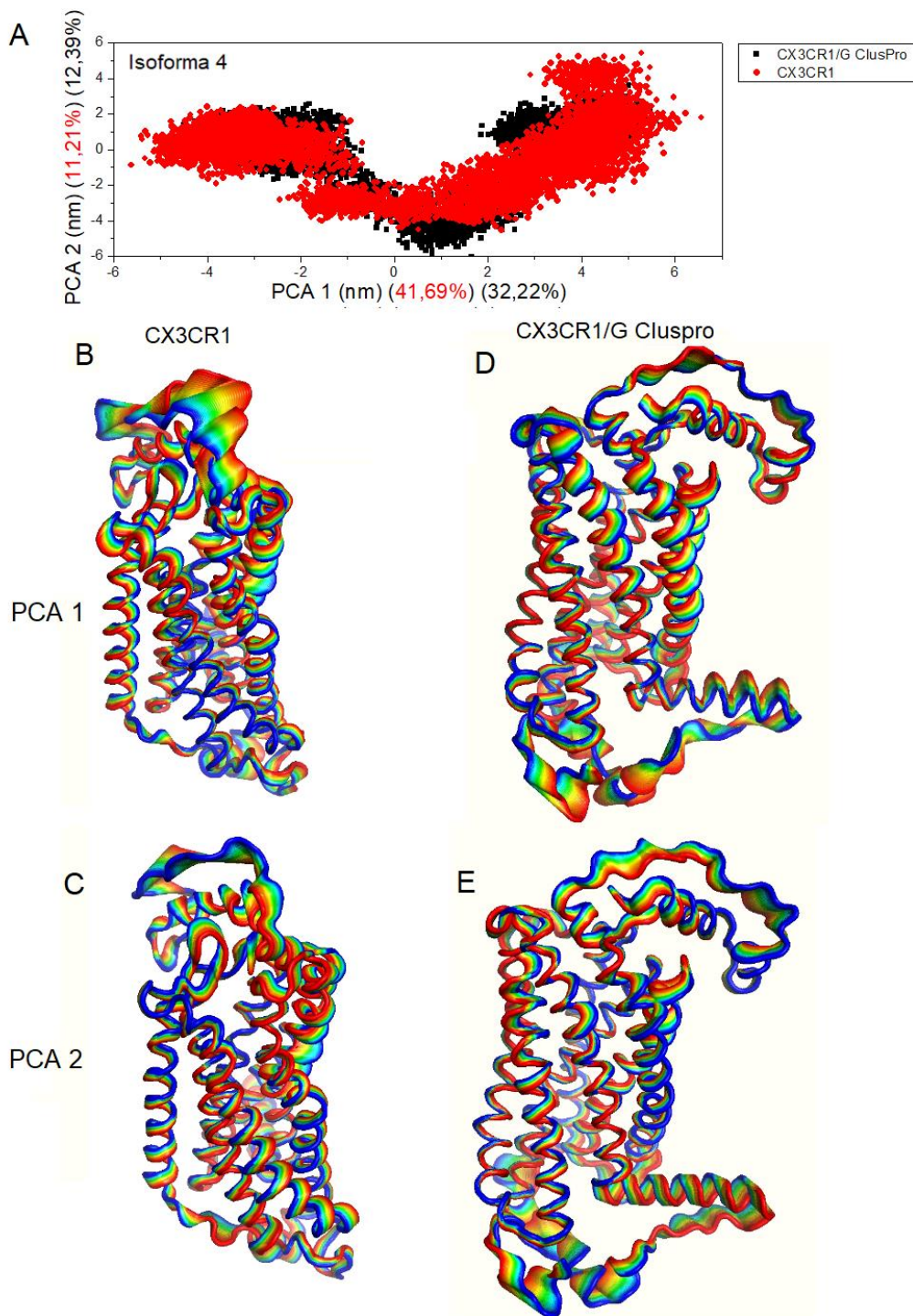


Fonte: Elaborado pelo autor.

O gráfico de dispersão PC1 & PC2 para a análise de PCA gerado para a isoforma 4 do CX3CR1 livre e na sua forma complexada com o peptídeo da G modelada pelo ClusPro revelou que a forma complexada tem um espaço conformacional mais compacto com relação a forma livre (Figura 32-A). Pode ser observada a partir dos movimentos filtrados para o conjunto de autovetores, a perda de espaço conformacional da isoforma 4 complexada com o peptídeo da G quando comparados com a forma livre da proteína (Figura 32-B e D) tanto para a PC1 como para PC2 (Figura 32-C e E). Esse resultado está em acordo com análises de RMSF. É possível notar também que a região N-terminal do barril de hélices  $\alpha$  envolvidos na interação com o peptídeo da G também perdem movimento quando comparados com a isoforma 4 no seu estado livre. Também é notável uma diminuição dos movimentos das hélices  $\alpha$  do barril para a forma complexada em relação a forma livre da isoforma 4 (Figura 32-B, C, D e E), e isto está de acordo como gráfico de RMSF (Figura 28-D).

Para estes resultados obtidos pela DM do *docking* do ClusPro notamos que as isoformas que mais tiveram seus movimentos aumentados foram as isoformas 1 e 3, pois elas apresentam uma região N-terminal menor em relação as isoformas 2 e 4, apresentando maior flexibilidade da forma livre. Outro fator observável é o aterramento do peptídeo pois quanto mais aterrado ele estiver no barril de hélices menor será a variação de movimento do complexo, e também devemos observar que quando o peptídeo da G é inserido na região N-terminal a região do C-terminal do complexo sofre maiores variações de movimento em relação ao CX3CR1 livre.

**FIGURA 32:** Gráficos de dispersão PCA, PCA1 e PCA2 para CX3CR1 livre (pontos vermelhos) e para complexo CX3CR1/G da DM dado pelo *docking* do ClusPro (pontos pretos) da isoforma 4 (A). (B e C) Moções filtradas de os autovetores de PCA1 (B) e PCA2 (C) para a dinâmica dados de CX3CR1 em sua forma livre e na sua forma de complexo CX3CR1/G (D e E) . O movimento é indicada pela variação de cor de azul para vermelho. Para realizar os cálculos de PCA, os 300 ns de trajetórias foram concatenados para cada uma das moléculas sistemas (CX3CR1 livre ou CX3CR1/G).

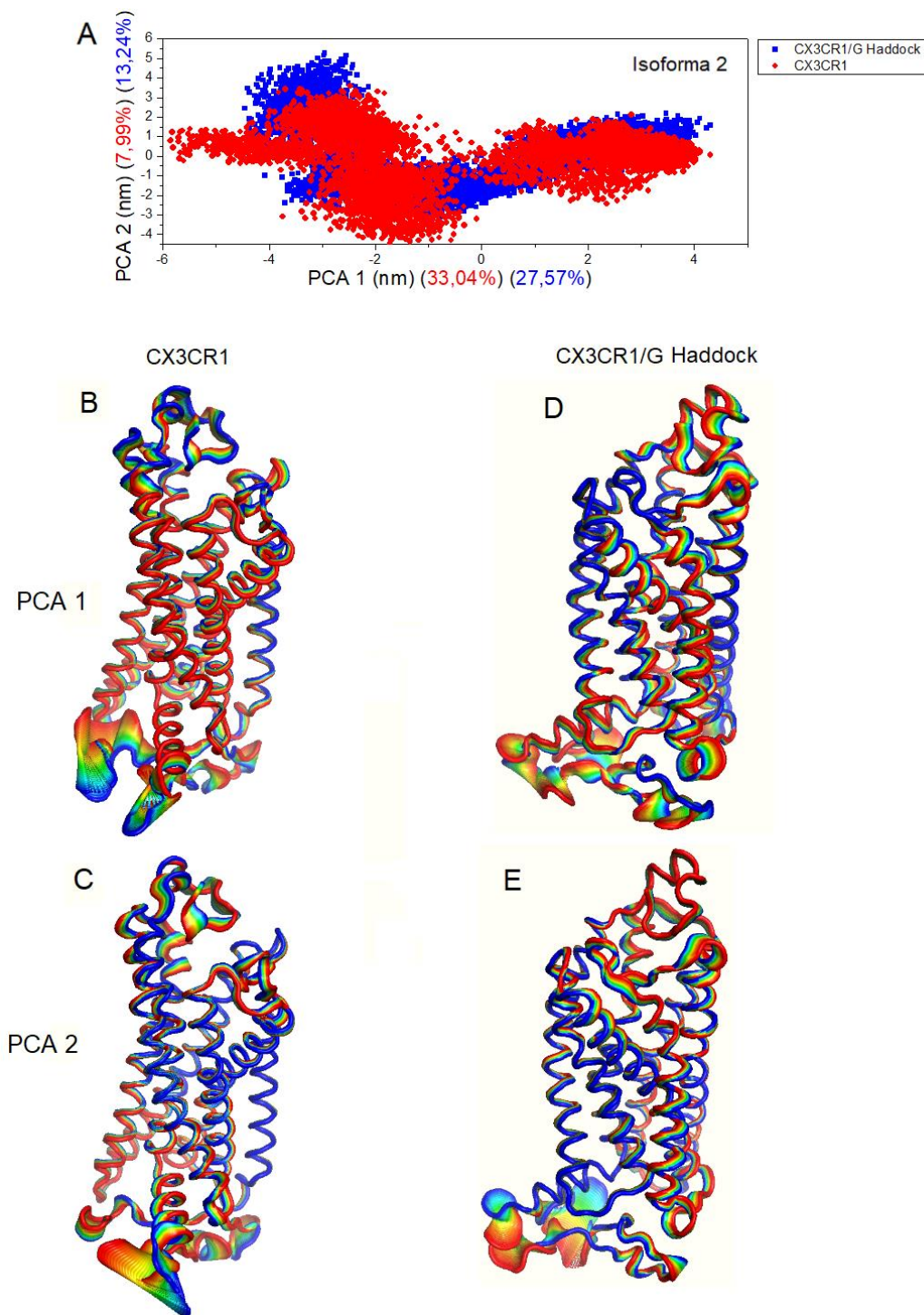


Fonte: Elaborado pelo autor.

O gráfico gerado para a isoforma 1 do CX3CR1 livre e na sua forma complexada com o peptídeo da G modelada pelo HADDOCK revelou que a forma livre tem um pequeno espaço conformacional mais compacto com relação a forma ligada (Figura 33-A). Pode ser observada a partir dos movimentos filtrados para o conjunto de autovetores, um leve ganho de espaço conformacional da isoforma 1 complexada quando comparados com a forma livre da proteína (Figura 33-B e D) tanto para a PC1 como para PC2 (Figura 33-C e E). Esse resultado está em acordo com análises de RMSF. É notável também que a região N-terminal do barril de hélices  $\alpha$  envolvidos na interação com o peptídeo da G manteve o mesmo movimento quando comparados com a isoforma 1 no seu estado livre. Também é notável uma manutenção dos movimentos das hélices  $\alpha$  do barril para a forma complexada em relação a forma livre da isoforma 1 (Figura 33-B, C, D e E), e isto está de acordo como gráfico de RMSF (Figura 28-A).

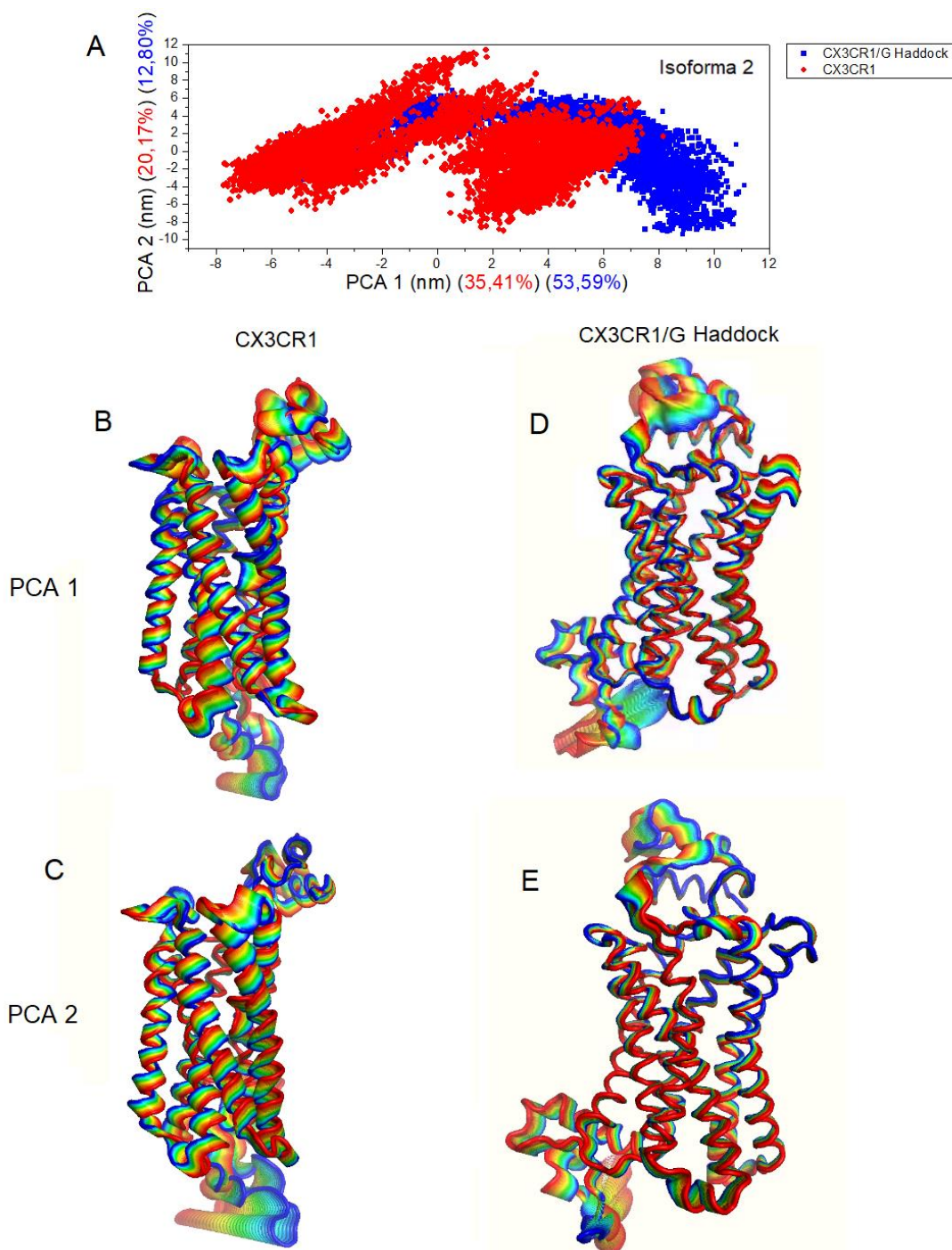
O gráfico gerado para a isoforma 2 do CX3CR1 livre e na sua forma complexada com o peptídeo da G modelada pelo HADDOCK revelou que a forma complexada tem um espaço conformacional mais abrangente com relação a forma livre (Figura 34-A). Pode ser observada a partir dos movimentos filtrados para o conjunto de autovetores, o ganho de espaço conformacional da isoforma 2 complexada com o peptídeo da G quando comparados com a forma livre da proteína (Figura 34-B e D) tanto para a PC1 como para PC2, porém esse ganho é apenas para N- e C-terminais (Figura 34-C e E) e está de acordo para a dispersão RMSF. É possível notar também que a região N-terminal do barril de hélices  $\alpha$  envolvidos na interação com o peptídeo da G juntamente com o barril de hélices  $\alpha$  obtiveram uma diminuição dos movimentos das hélices  $\alpha$  do barril para a forma complexada em relação a forma livre da isoforma 2 (Figura 34-B, C, D e E), e isto está de acordo como gráfico de RMSF (Figura 28-B).

**FIGURA 33:** Gráficos de dispersão PCA, PCA1 e PCA2 para CX3CR1 livre (pontos vermelhos) e para complexo CX3CR1/G da DM dado pelo *docking* do HADDOCK (pontos azuis) da isoforma 1 (A). (B e C) Moções filtradas de os autovetores de PCA1 (B) e PCA2 (C) para a dinâmica dados de CX3CR1 em sua forma livre e na sua forma de complexo CX3CR1/G (D e E) . O movimento é indicada pela variação de cor de azul para vermelho. Para realizar os cálculos de PCA, os 300 ns de trajetórias foram concatenados para cada uma das moléculas sistemas (CX3CR1 livre ou CX3CR1/G).



Fonte: Elaborado pelo autor.

**FIGURA 34:** Gráficos de dispersão PCA, PCA1 e PCA2 para CX3CR1 livre (pontos vermelhos) e para complexo CX3CR1/G da DM dado pelo *docking* do HADDOCK (pontos azuis) da isoforma 2 (A). (B e C) Moções filtradas de os autovetores de PCA1 (B) e PCA2 (C) para a dinâmica dados de CX3CR1 em sua forma livre e na sua forma de complexo CX3CR1/G (D e E) . O movimento é indicada pela variação de cor de azul para vermelho. Para realizar os cálculos de PCA, os 300 ns de trajetórias foram concatenados para cada uma das moléculas sistemas (CX3CR1 livre ou CX3CR1/G).

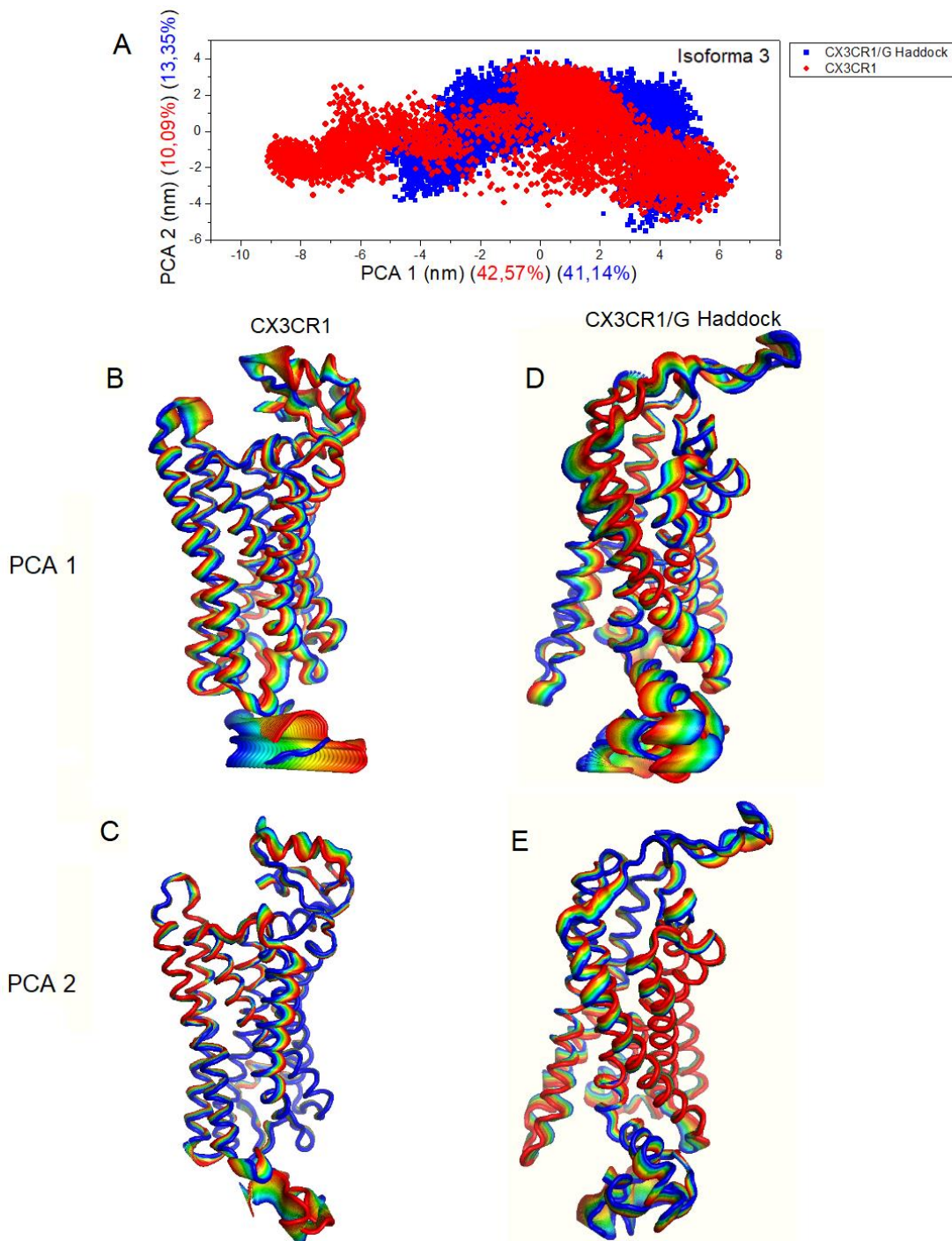


Fonte: Elaborado pelo autor.

O gráfico gerado para a isoforma 3 do CX3CR1 livre e na sua forma complexada com o peptídeo da G modelada pelo HADDOCK revelou que a forma ligada tem um pequeno espaço conformacional mais compacto com relação a forma livre (Figura 35-A). Pode ser observada a partir dos movimentos filtrados para o conjunto de autovetores, um ganho de espaço conformacional da isoforma 3 complexada quando comparados com a forma livre da proteína apenas para a região C-terminal (Figura 35-B e D) tanto para a PC1 como para PC2 (Figura 35-C e E). Entretanto para a região N-terminal há apenas uma manutenção dos movimentos para a forma ligada (Figura 35-B a E). Esse resultado está em acordo com análises de RMSF. É notável também que a região N-terminal do barril de hélices  $\alpha$  envolvidos na interação com o peptídeo da G obteve um aumento dos movimentos quando comparados com a isoforma 3 no seu estado livre. Também é notável um acréscimo dos movimentos das hélices  $\alpha$  do barril para a forma complexada em relação a forma livre da isoforma 3 (Figura 35-B, C, D e E), e isto está de acordo como gráfico de RMSF notando as hélices  $\alpha_1$ ,  $\alpha_7$  e  $\alpha_8$  (Figura 28-C).

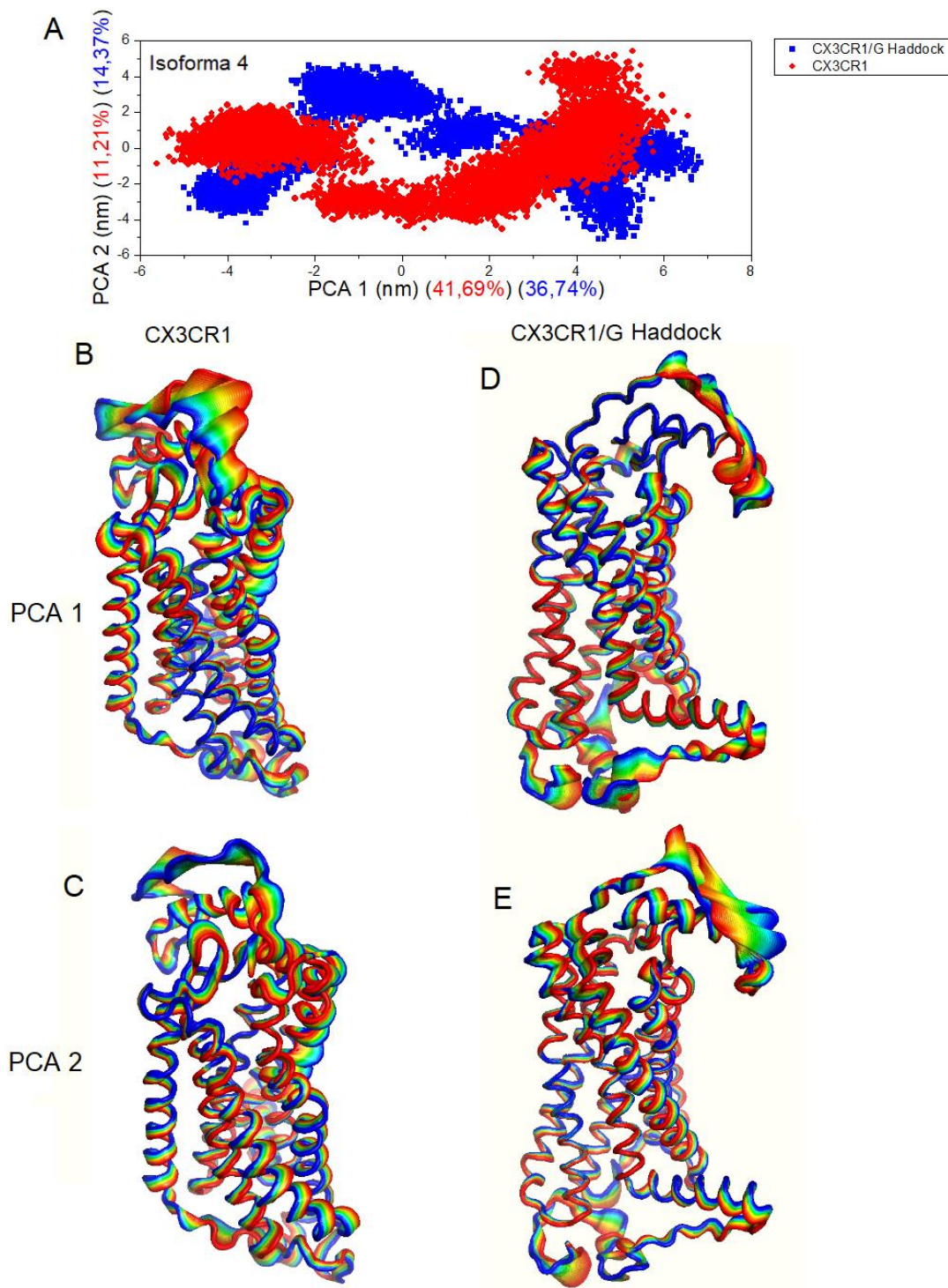
Para a isoforma 4 revelou uma distinção entre livre e a complexada, principalmente se formos observar na escala de PCA 2, todavia a forma livre apresenta concavidade para cima enquanto a forma ligada concavidade para baixo, (Figura 36-A). Pode ser observada a partir dos movimentos filtrados para o conjunto de autovetores que para PC1 o N-terminal da forma livre em relação ao N-terminal do complexo houve uma queda significativa de movimento enquanto C-terminal manteve parecido, todavia é mais notado no N-terminal (Figuras 36 B e D). Para as moções gerados pelos autovetores do PC2 a região N-terminal mostrou maior grau de movimento para a forma ligada (Figuras 36 C e E). Também é notável um decréscimo dos movimentos das hélices  $\alpha$  do barril para a forma complexada em relação a forma livre da isoforma 4 (Figura 36-B, C, D e E). E isto está de acordo como gráfico de RMSF (Figura 28-D).

**FIGURA 35:** Gráficos de dispersão PCA, PCA1 e PCA2 para CX3CR1 livre (pontos vermelhos) e para complexo CX3CR1/G da DM dado pelo *docking* do HADDOCK (pontos azuis) da isoforma 3 (A). (B e C) Moções filtradas de os autovetores de PCA1 (B) e PCA2 (C) para a dinâmica dados de CX3CR1 em sua forma livre e na sua forma de complexo CX3CR1/G (D e E) . O movimento é indicada pela variação de cor de azul para vermelho. Para realizar os cálculos de PCA, os 300 ns de trajetórias foram concatenados para cada uma das moléculas sistemas (CX3CR1 livre ou CX3CR1/G).



Fonte: Elaborado pelo autor.

**FIGURA 36:** Gráficos de dispersão PCA, PCA1 e PCA2 para CX3CR1 livre (pontos vermelhos) e para complexo CX3CR1/G da DM dado pelo *docking* do HADDOCK (pontos azuis) da isoforma 4 (A). (B e C) Moções filtradas de os autovetores de PCA1 (B) e PCA2 (C) para a dinâmica dados de CX3CR1 em sua forma livre e na sua forma de complexo CX3CR1/G (D e E) . O movimento é indicada pela variação de cor de azul para vermelho. Para realizar os cálculos de PCA, os 300 ns de trajetórias foram concatenados para cada uma das moléculas sistemas (CX3CR1 livre ou CX3CR1/G)..



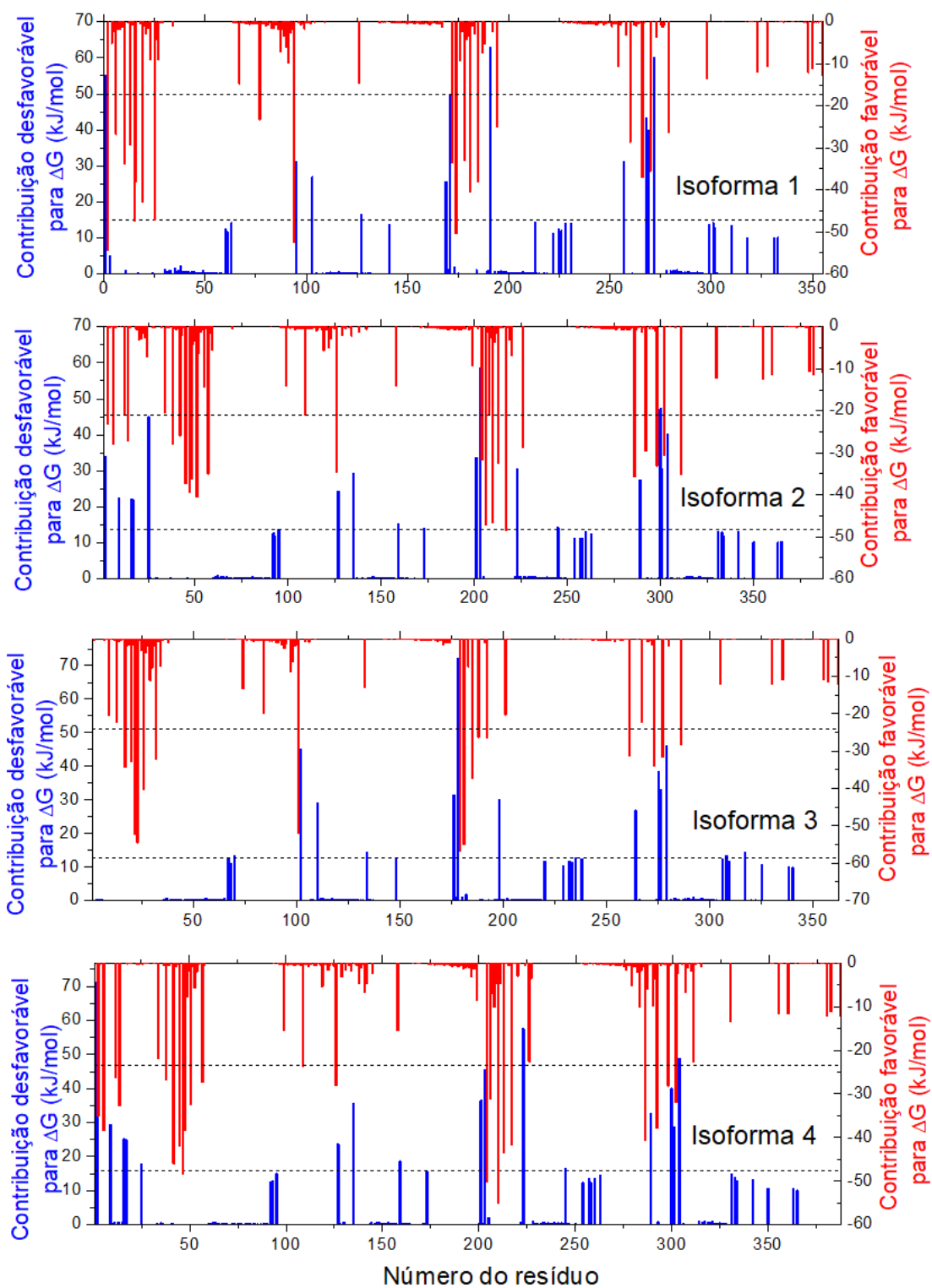
Fonte: Elaborado pelo autor.

Analisando os dados de RMSF e PCA das DM dos *dockings* dado pelos *servers* ClusPro e HADDOCK, podemos considerar que o peptídeo na região N-terminal influencia o movimento do receptor celular, principalmente nos N- e C-terminais e também nas alças entre as hélices  $\alpha$  do barril, ocasionando um maior movimento para a adaptação do peptídeo da G no receptor celular CX3CR1 (CX3CR1/peptídeo da G). Outro detalhe a ser observado é o movimento que propaga para o receptor celular no qual o peptídeo está ligado, influenciando no movimento das hélices  $\alpha$  e, conseqüentemente, o movimento da região C-terminal do receptor celular.

#### **4.5 Cálculos de Energia Livre de Ligação**

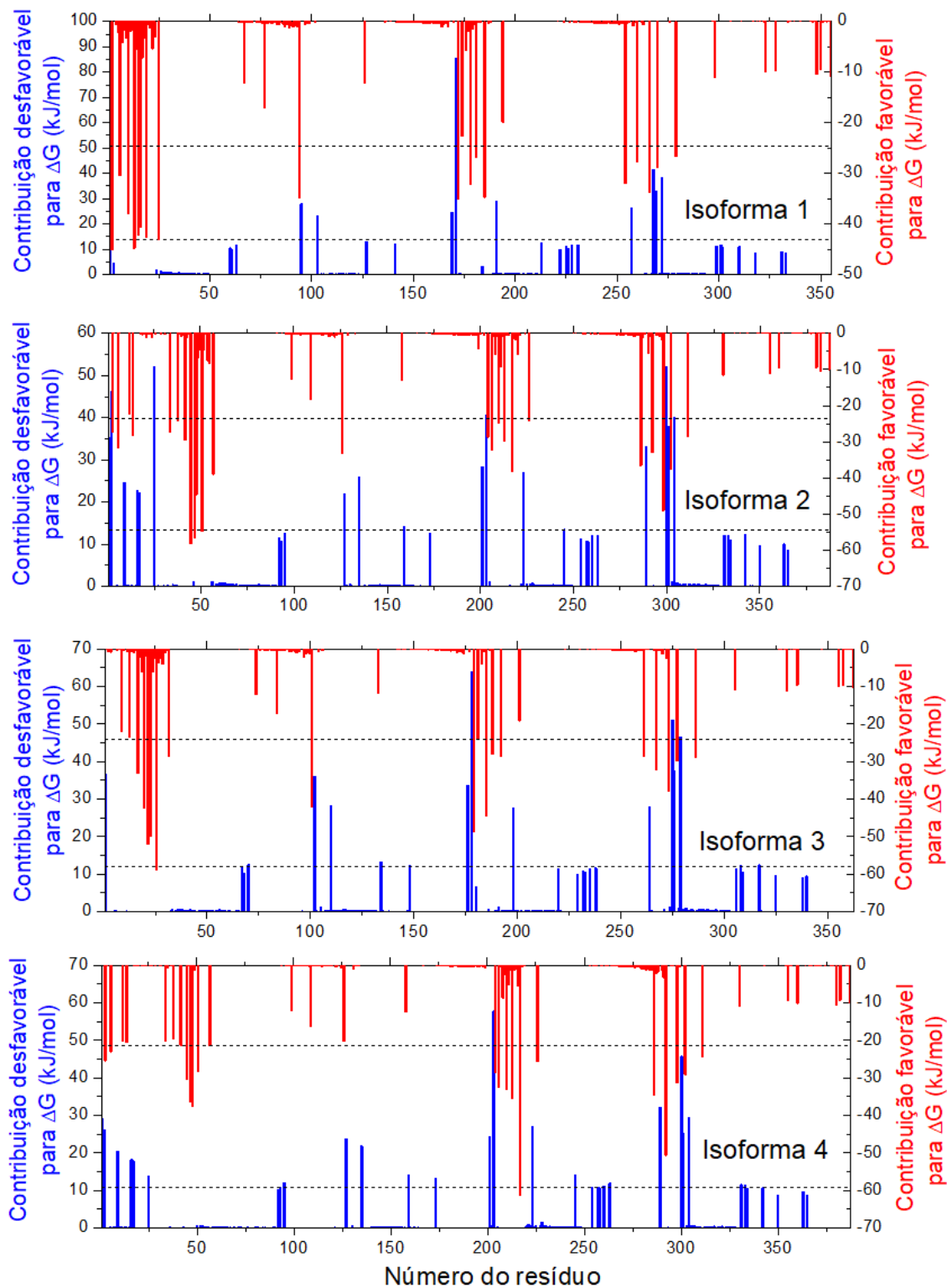
Os cálculos de energia livre de ligação teórica são importantes para uma estimativa da afinidade envolvida na associação entre moléculas, como no caso as isoformas do CX3CR1 e o peptídeo da G. Essa energia foi calculada para as trajetórias das DM dos modelos estruturais determinados pelos programas ClusPro (Figura 37) e HADDOCK (Figura 38). Nas Figuras 37 e 38, a energia livre de ligação favorável é representada pela cor vermelha enquanto a energia desfavorável é representada pela cor azul. Como as isoformas do receptor celular CX3CR1 tem um número significativo de resíduos de aminoácidos, um valor médio mais seu desvio padrão da energia de ligação foi utilizado como ponto de corte (linha tracejada nas Figuras 37 e 38) para caracterizar uma contribuição significativa de um determinado resíduo da proteína para energia livre de ligação total.

**FIGURA 37:** Contribuição individual de cada resíduo de aminoácido na energia total do receptor celular CX3CR1 para os resultados de docking obtidos pelas quatro isoformas do server ClusPro. Sendo contribuição favorável (vermelho) e contribuição desfavorável (azul).



Fonte: Elaborado pelo autor.

**FIGURA 38:** Contribuição individual de cada resíduo de aminoácido na energia total do receptor celular CX3CR1 para os resultados de docking obtidos pelas quatro isoformas do server HADDOCK. Sendo contribuição favorável (vermelho) e contribuição desfavorável (azul).



Fonte: Elaborado pelo autor.

Os resíduos das isoformas do CX3CR1 que apresentam valores de energia livre de ligação além do valor de corte (média mais desvio padrão) são apresentados nas Tabelas 16–19 para os modelos estruturais determinados pelo ClusPro e nas Tabelas 20–23 para os modelos calculados pelo HAADOCK. Vale mencionar que cada isoforma tem uma quantidade diferente de resíduos devido a diferença do alongamento do N-terminal, sendo a isoforma 1 com 355 resíduos, a isoforma 2 e 4 têm 387 resíduos e a isoforma 3 tem 362 resíduos (Figura 3), portanto numerações diferentes nas Tabelas 19–23 podem corresponder a um mesmo resíduo de aminoácido nas diferentes isoformas do CX3CR1 na interação como o peptídeo da G.

De modo em geral, a partir das Tabelas 19–23 é possível notar que todos os resíduos de aminoácidos com contribuição significativa para a energia livre ligação total são carregados polares, sendo ácidos aspárticos e glutâmicos aqueles com contribuição favorável para interação com o peptídeo da G e resíduos de lisina e arginina para contribuição desfavorável. Por apresentar uma carga líquida negativa, a metionina N-terminal sempre apresenta uma contribuição favorável para energia livre de ligação total. Esse aspecto de interação principalmente com resíduos polares carregados revela um papel fundamental das interações eletrostáticas na interação da porção N-terminal do barril de hélices  $\alpha$  das isoformas do CX3CR1 com o peptídeo da região de cisteínas conservadas da proteína G, o qual apresenta uma região de carga líquida positiva (LYS-187 e ARG-188).

Nas Tabelas 16-19 são apresentados os resíduos das isoformas do CX3CR1 das simulações de DM para os modelos estruturais do complexo determinado pelo programa ClusPro. Pode-se observar que os seguintes resíduos ácidos se repetem com relação a contribuição favorável: ASP-15, ASP-16, GLU-19, ASP-25, GLU-94, GLU-172, GLU-174, ASP-178, GLU-181, GLU-185, ASP-266 e ASP-270 (numeração da isoforma 1), os quais aparecem para as quatro isoformas, Vale destacar que os resíduos GLU-172 e GLU-174 das folhas  $\beta_1$  e  $\beta_2$  que interage com o peptídeo da G contribuem favoravelmente para energia livre de ligação total. Por outro lado, os resíduos MET-1, ARG-191, LYS-257, ARG-268, LYS-269 e ARG-272 (numeração da isoforma 1) contribuem desfavoravelmente para energia livre de ligação total na interação para as quatro isoformas do CX3CR1.

**TABELA 16:** Resíduos de aminoácidos da isoforma 1 do CX3CR1 e seus respectivos valores das contribuições favoráveis e desfavoráveis (acima do valor de corte média mais desvio padrão) para a energia livre de ligação total para a trajetória de DM do modelo estrutural determinado pelo servidor ClusPro.

<b>Resíduo</b>	<b>Contribuição favorável para <math>\Delta G_b^t</math> (kJ.mol<sup>-1</sup>)</b>	<b>Resíduo</b>	<b>Contribuição desfavorável para <math>\Delta G_b^t</math> (kJ.mol<sup>-1</sup>)</b>
ASP-2	-54.3742 ± 0.4488	MET-1	55.1252 ± 0.3829
GLU-6	-26.7860 ± 0.8622	ARG-191	63.0520 ± 0.7383
GLU-10	-33.8363 ± 0.3519	LYS-257	31.0818 ± 0.1331
GLU-13	-29.4089 ± 0.2291	ARG-268	43.2607 ± 0.2207
ASP-15	-47.4130 ± 0.5249	LYS-269	39.9330 ± 0.2151
ASP-16	-38.2271 ± 0.4115	ARG-272	60.1343 ± 0.8727
GLU-19	-42.9093 ± 0.4722		
ASP-25	-47.1302 ± 0.3040		
ASP-77	-23.3065 ± 0.0549		
GLU-94	-52.4474 ± 0.8221		
GLU-172	-33.5709 ± 0.4471		
GLU-174	-50.4883 ± 0.7043		
ASP-178	-32.9644 ± 0.4432		
GLU-181	-40.4746 ± 0.6105		
GLU-185	-38.1158 ± 0.4964		
GLU-194	-25.1414 ± 0.1572		
ASP-260	-28.7039 ± 0.1452		
ASP-266	-37.0811 ± 0.2522		
ASP-270	-35.5579 ± 0.1633		
GLU-279	-26.4540 ± 0.4929		

Fonte: Elaborado pelo autor.

**TABELA 17:** Resíduos de aminoácidos da isoforma 2 do CX3CR1 e seus respectivos valores das contribuições favoráveis e desfavoráveis (acima do valor de corte média mais desvio padrão) para a energia livre de ligação total para a trajetória de DM do modelo estrutural determinado pelo servidor ClusPro.

<b>Resíduo</b>	<b>Contribuição favorável para <math>\Delta G_b^t</math> (kJ.mol<sup>-1</sup>)</b>	<b>Resíduo</b>	<b>Contribuição desfavorável para <math>\Delta G_b^t</math> (kJ.mol<sup>-1</sup>)</b>
GLU-3	-23.1481 ± 0.1045	MET-1	27.6772 ± 0.1717
GLU-6	-27.8622 ± 0.1484	ARG-2	33.8554 ± 0.2193
ASP-14	-27.0833 ± 0.2071	LYS-9	22.4017 ± 0.1033
GLU-38	-27.9341 ± 0.3917	ARG-16	22.1936 ± 0.1524
GLU-42	-25.9748 ± 0.1941	LYS-17	21.7989 ± 0.1684
GLU-45	-37.2836 ± 0.3294	ARG-25	45.0447 ± 0.6947
ASP-47	-39.3430 ± 0.4317	LYS-127	24.4213 ± 0.2261
ASP-48	-36.2218 ± 0.3789	LYS-135	29.3349 ± 0.1355
GLU-51	-40.5842 ± 0.5162	ARG-159	15.3274 ± 0.0697
ASP-57	-35.0708 ± 0.1988	LYS-201	33.7078 ± 0.2129

GLU-126	-34.6957 ± 0.5053	LYS-203	58.6710 ± 0.9684
GLU-204	-31.7103 ± 0.2786	ARG-223	30.4671 ± 0.2304
GLU-206	-47.2354 ± 0.8033	ARG-245	14.2962 ± 0.0460
ASP-210	-46.6206 ± 0.5299	LYS-289	27.5063 ± 0.1768
GLU-213	-32.4105 ± 0.4028	ARG-300	47.3080 ± 0.7572
GLU-217	-48.5981 ± 0.6807	LYS-301	30.5810 ± 0.1868
GLU-226	-28.8294 ± 0.1697	ARG-304	40.2329 ± 0.4248
GLU-286	-35.6935 ± 0.6884		
ASP-292	-29.5460 ± 0.3864		
ASP-298	-33.1788 ± 0.3684		
ASP-302	-30.4932 ± 0.2005		
GLU-311	-37.2528 ± 0.4591		

Fonte: Elaborado pelo autor.

**TABELA 18:** Resíduos de aminoácidos da isoforma 3 do CX3CR1 e seus respectivos valores das contribuições favoráveis e desfavoráveis (acima do valor de corte média mais desvio padrão) para a energia livre de ligação total para a trajetória de DM do modelo estrutural determinado pelo servidor ClusPro.

<b>Resíduo</b>	<b>Contribuição favorável para <math>\Delta G_b^t</math> (kJ.mol<sup>-1</sup>)</b>	<b>Resíduo</b>	<b>Contribuição desfavorável para <math>\Delta G_b^t</math> (kJ.mol<sup>-1</sup>)</b>
GLU-17	-34.2432 ± 0.5610	MET-1	30.9913 ± 0.3824
GLU-20	-32.6866 ± 0.4204	LYS-70	13.4567 ± 0.0966
ASP-22	-52.4255 ± 0.9967	LYS-102	45.0240 ± 0.6288
ASP-23	-54.6023 ± 1.4741	LYS-110	29.2078 ± 0.1901
GLU-26	-40.3749 ± 0.4935	ARG-134	14.4692 ± 0.0710
GLU-101	-52.1270 ± 0.9762	LYS-176	31.3808 ± 0.2233
GLU-179	-56.7484 ± 0.8276	LYS-178	72.1759 ± 1.8854
GLU-181	-55.0986 ± 0.8947	ARG-198	30.1945 ± 0.1869
ASP-185	-37.3538 ± 0.3416	LYS-264	26.9555 ± 0.2072
GLU-188	-26.3911 ± 0.2428	ARG-275	38.5101 ± 0.5429
GLU-192	-26.5303 ± 0.1995	LYS-276	33.0066 ± 0.2570
GLU-261	-31.2973 ± 0.1970	ARG-279	46.1070 ± 1.1839
ASP-273	-33.9750 ± 0.4112	ARG-308	13.3624 ± 0.0614
ASP-277	-31.5745 ± 0.1978	LYS-317	14.2123 ± 0.1085
GLU-286	-28.2520 ± 0.7070		

Fonte: Elaborado pelo autor.

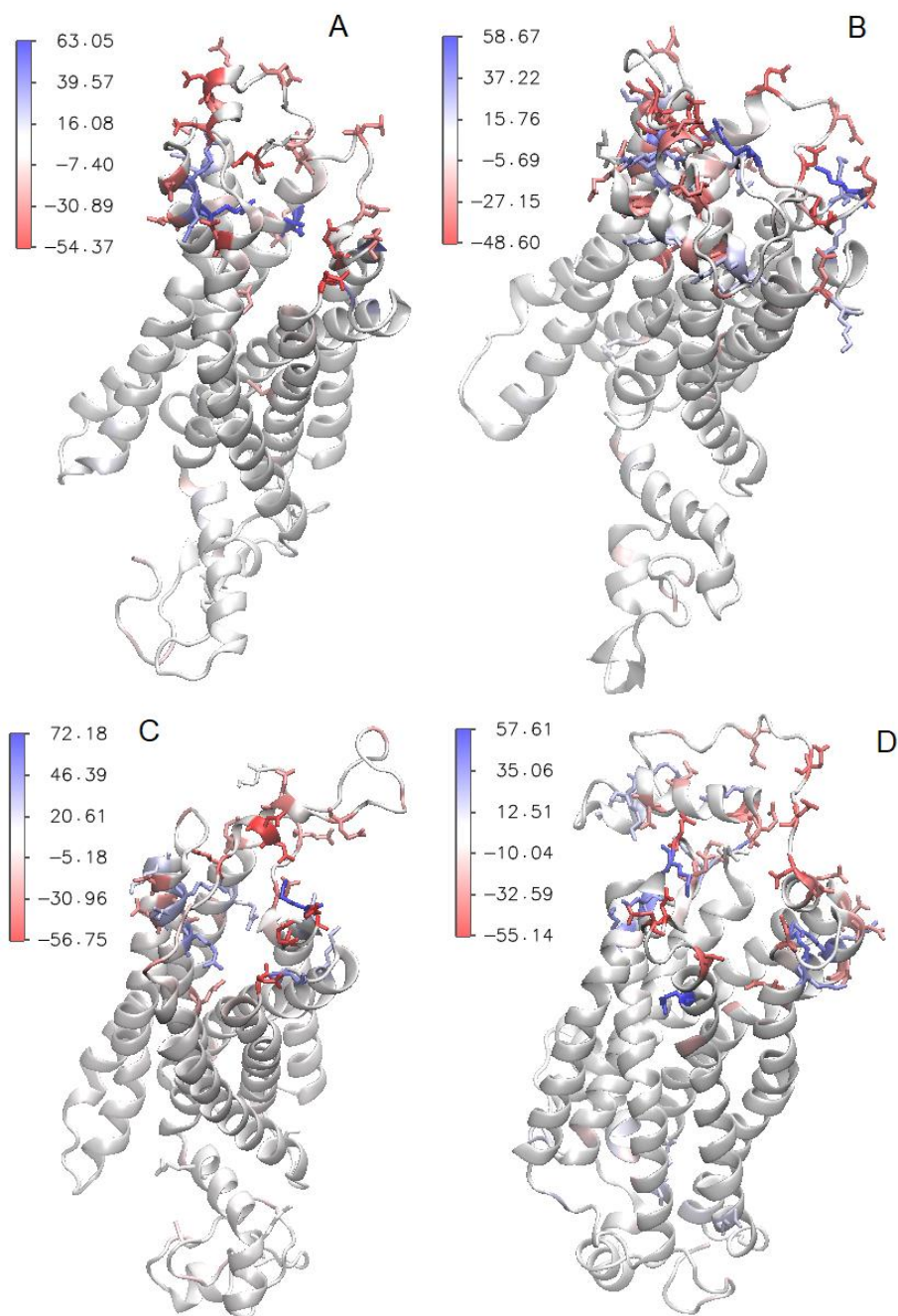
**TABELA 19:** Resíduos de aminoácidos da isoforma 4 do CX3CR1 e seus respectivos valores das contribuições favoráveis e desfavoráveis (acima do valor de corte média mais desvio padrão) para a energia livre de ligação total para a trajetória de DM do modelo estrutural determinado pelo servidor ClusPro.

<b>Resíduo</b>	<b>Contribuição favorável para <math>\Delta G_b^t</math> (kJ.mol<sup>-1</sup>)</b>	<b>Resíduo</b>	<b>Contribuição desfavorável para <math>\Delta G_b^t</math> (kJ.mol<sup>-1</sup>)</b>
GLU-3	-35.2479 ± 0.1372	MET-1	71.1883 ± 0.6204
GLU-6	-38.4306 ± 0.2190	LYS-9	29.3693 ± 0.1596
ASP-12	-26.4038 ± 0.1247	ARG-16	25.2481 ± 0.1219
ASP-14	-32.7465 ± 0.1965	LYS-17	24.8968 ± 0.1076
GLU-38	-26.7929 ± 0.2272	ARG-25	17.6975 ± 0.0569
GLU-42	-46.0251 ± 1.1194	LYS-127	23.7165 ± 0.1057
GLU-45	-41.8955 ± 0.8587	LYS-135	35.7651 ± 0.1417
ASP-47	-48.5117 ± 0.4523	ARG-159	18.6720 ± 0.0795
ASP-48	-38.4739 ± 0.3748	LYS-201	36.5710 ± 0.1688
GLU-51	-32.4130 ± 0.1150	LYS-203	45.4695 ± 0.2760
ASP-57	-27.3605 ± 0.0806	ARG-223	57.6072 ± 0.7286
GLU-126	-28.2314 ± 0.1341	ARG-245	16.3875 ± 0.0421
GLU-204	-50.3079 ± 0.8404	LYS-289	32.5713 ± 0.1402
GLU-206	-31.2928 ± 0.7592	ARG-300	40.0522 ± 0.4564
ASP-210	-55.1360 ± 0.6515	LYS-301	28.8323 ± 0.1671
GLU-213	-43.5746 ± 0.2834	ARG-304	48.9390 ± 0.4704
GLU-217	-41.6641 ± 0.3981		
GLU-286	-40.7208 ± 0.6169		
ASP-292	-37.9858 ± 0.3827		
ASP-298	-28.2532 ± 0.1146		
ASP-302	-32.0371 ± 0.1218		

Fonte: Elaborado pelo autor.

Para ilustrar a localização destes resíduos apresentados nas Tabelas 16–19, a Figura 39 foi feita denotando os resíduos com contribuições favoráveis e desfavoráveis por um gradiente de cores do vermelho para o azul. Para isso, utilizou-se uma estrutura representativa da simulação de DM para cada modelo estrutural do complexo isoforma/peptídeo da G obtida da análise de clusterização (Figura 39). É possível observar que a maioria dos resíduos de contribuição significativa, a energia livre de ligação pertencem as regiões de alças entre as hélices  $\alpha$  do barril das isoformas do CX3CR1, observando também que o N-terminal da proteína apresenta um maior número de resíduos com contribuição favorável para a energia total, como para as isoformas 2 e 4 (Figura 39-B e Figura 39-D).

**FIGURA 39:** Estrutura representativa das isoformas do receptor celular CX3CR1 com os resíduos que apresentam uma contribuição favoráveis e desfavoráveis significativas (acima da média mais desvio padrão) para a energia livre de ligação total para as simulações de DM dos modelos estruturais dos complexos com o peptídeo da G determinados pelo servidor ClusPro. As contribuições favoráveis e desfavoráveis para as energias totais são representadas gradientes de cores indo vermelho para o azul, respectivamente. As barras de cores denotam o aspecto quantitativo de cada gradiente de cor. As isoformas 1 (A), 2 (B), 3 (C) e 4 (D) são apresentadas em modelos de *cartoon* em cinza e os resíduos de aminoácidos na forma de sticks.



Fonte: Elaborado pelo autor.

Nas Tabelas 20-23 são apresentados os resíduos das isoformas do CX3CR1 das simulações de DM para os modelos estruturais do complexo determinado pelo programa HADDOCK. Pode-se observar que os seguintes resíduos ácidos se repetem com relação a contribuição favorável: ASP-15, ASP-16, GLU-19, GLU-172, ASP-178, GLU-181, GLU-185, GLU-254, ASP-260, ASP-266, ASP-270 e GLU-279 (numeração isoforma 1), os quais aparecem para as quatro isoformas, Vale destacar que o resíduo GLU-172 das folhas  $\beta_1$  e  $\beta_2$  que interage com o peptídeo da G contribui favoravelmente para energia livre de ligação total. Por outro lado, os resíduos MET-1, LYS-95, LYS-103, LYS-169, LYS171, ARG-191, LYS-257, ARG-268 e LYS-269 (numeração da isoforma 1) contribuem desfavoravelmente para energia livre de ligação total na interação para as quatro isoformas do CX3CR1.

**TABELA 20:** Resíduos de aminoácidos da isoforma 1 do CX3CR1 e seus respectivos valores das contribuições favoráveis e desfavoráveis (acima do valor de corte média mais desvio padrão) para a energia livre de ligação total para a trajetória de DM do modelo estrutural determinado pelo servidor HADDOCK.

Resíduo	Contribuição favorável para $\Delta G_b^t$ (kJ.mol <sup>-1</sup> )	Resíduo	Contribuição desfavorável para $\Delta G_b^t$ (kJ.mol <sup>-1</sup> )
ASP-2	-45.1132 ± 0.3017	MET-1	74.6410 ± 0.9001
GLU-6	-30.4376 ± 0.7164	LYS-95	27.9588 ± 0.2303
GLU-10	-37.9291 ± 0.7992	LYS-103	23.1837 ± 0.1276
GLU-13	-44.8379 ± 0.6296	LYS-169	24.6057 ± 0.1993
ASP-15	-42.2673 ± 0.5539	LYS-171	85.5512 ± 1.8107
ASP-16	-40.7070 ± 0.2930	ARG-191	28.8447 ± 0.2606
GLU-19	-42.5483 ± 0.5340	LYS-257	26.3803 ± 0.2163
ASP-25	-43.0511 ± 0.5403	ARG-268	41.4720 ± 0.2569
GLU-94	-34.8611 ± 0.5744	LYS-269	33.1205 ± 0.2692
GLU-172	-35.0441 ± 0.9902		
ASP-178	-32.1494 ± 0.3768		
GLU-181	-26.8148 ± 0.3490		
GLU-185	-34.7771 ± 0.3746		
GLU-254	-31.9821 ± 0.2222		
ASP-260	-27.7178 ± 0.1948		
ASP-266	-33.6628 ± 0.3019		
ASP-270	-28.8433 ± 0.1828		
GLU-279	-26.6952 ± 0.2198		

Fonte: Elaborado pelo autor.

**TABELA 21:** Resíduos de aminoácidos da isoforma 2 do CX3CR1 e seus respectivos valores das contribuições favoráveis e desfavoráveis (acima do valor de corte média mais desvio padrão) para a energia livre de ligação total para a trajetória de DM do modelo estrutural determinado pelo servidor HADDOCK.

<b>Resíduo</b>	<b>Contribuição favorável para <math>\Delta G_b^t</math> (kJ.mol<sup>-1</sup>)</b>	<b>Resíduo</b>	<b>Contribuição desfavorável para <math>\Delta G_b^t</math> (kJ.mol<sup>-1</sup>)</b>
GLU-3	-27.3764 ± 0.1408	MET-1	35.2221 ± 0.2716
GLU-6	-31.6573 ± 0.2269	ARG-2	46.1604 ± 0.3141
ASP-14	-28.1699 ± 0.2236	LYS-9	24.5728 ± 0.1338
ASP-34	-27.3113 ± 0.5230	ARG-16	22.7386 ± 0.1516
GLU-38	-24.3192 ± 0.1684	LYS-17	22.2053 ± 0.1371
GLU-42	-29.6621 ± 0.2595	ARG-25	52.1026 ± 0.8881
GLU-45	-58.1685 ± 0.6442	LYS-127	21.8343 ± 0.1476
ASP-47	-56.6717 ± 0.6135	LYS-135	25.8860 ± 0.1327
ASP-48	-44.6457 ± 0.3546	ARG-159	14.1607 ± 0.0385
GLU-51	-54.7472 ± 0.8876	LYS-201	28.3341 ± 0.1735
ASP-57	-39.0113 ± 0.2975	LYS-203	40.4701 ± 1.1425
GLU-126	-33.0888 ± 0.3452	ARG-223	26.9118 ± 0.1850
GLU-204	-28.7770 ± 0.5048	ARG-245	13.4818 ± 0.0405
GLU-206	-32.4170 ± 0.4499	LYS-289	33.1750 ± 0.2701
ASP-210	-24.8404 ± 0.8855	ARG-300	52.2045 ± 0.7087
GLU-213	-29.9395 ± 0.2396	LYS-301	37.8537 ± 0.2677
GLU-217	-38.1828 ± 0.5107	ARG-304	39.9469 ± 0.5120
GLU-226	-24.1259 ± 0.1279		
GLU-286	-36.6292 ± 0.3979		
ASP-292	-33.0820 ± 0.3449		
ASP-298	-49.0239 ± 0.5945		
ASP-302	-37.5973 ± 0.2518		
GLU-311	-28.6352 ± 0.2366		

Fonte: Elaborado pelo autor.

**TABELA 22:** Resíduos de aminoácidos da isoforma 3 do CX3CR1 e seus respectivos valores das contribuições favoráveis e desfavoráveis (acima do valor de corte média mais desvio padrão) para a energia livre de ligação total para a trajetória de DM do modelo estrutural determinado pelo servidor HADDOCK.

<b>Resíduo</b>	<b>Contribuição favorável para <math>\Delta G_b^t</math> (kJ.mol<sup>-1</sup>)</b>	<b>Resíduo</b>	<b>Contribuição desfavorável para <math>\Delta G_b^t</math> (kJ.mol<sup>-1</sup>)</b>
GLU-17	-33.0951 ± 0.2840	MET-1	36.5784 ± 0.3569
GLU-20	-42.4794 ± 0.5082	LYS-70	12.5560 ± 0.0967
ASP-22	-51.9710 ± 0.5381	LYS-102	36.0491 ± 0.3847
ASP-23	-49.8257 ± 1.4122	LYS-110	28.2268 ± 0.1795
GLU-26	-58.7711 ± 0.5779	ARG-134	13.3326 ± 0.0705
ASP-32	-28.5530 ± 0.2397	LYS-176	33.6477 ± 0.2974

GLU-101	-42.0360 ± 0.6018	LYS-178	63.7654 ± 2.1584
GLU-179	-48.8122 ± 0.5621	ARG-198	27.6184 ± 0.2787
ASP-185	-44.3705 ± 0.5792	LYS-264	27.9215 ± 0.2452
GLU-188	-28.0738 ± 0.2113	ARG-275	51.0601 ± 0.7689
GLU-192	-28.7281 ± 0.2975	LYS-276	37.3575 ± 0.3908
GLU-261	-28.5006 ± 0.3133	ARG-279	46.5650 ± 0.8192
ASP-267	-32.2506 ± 0.4436	LYS-317	12.5437 ± 0.0899
ASP-273	-37.8855 ± 0.4307	MET-1	36.5784 ± 0.3569
ASP-277	-29.9251 ± 0.2586	LYS-70	12.5560 ± 0.0967
GLU-286	-28.8682 ± 0.2970		

Fonte: Elaborado pelo autor.

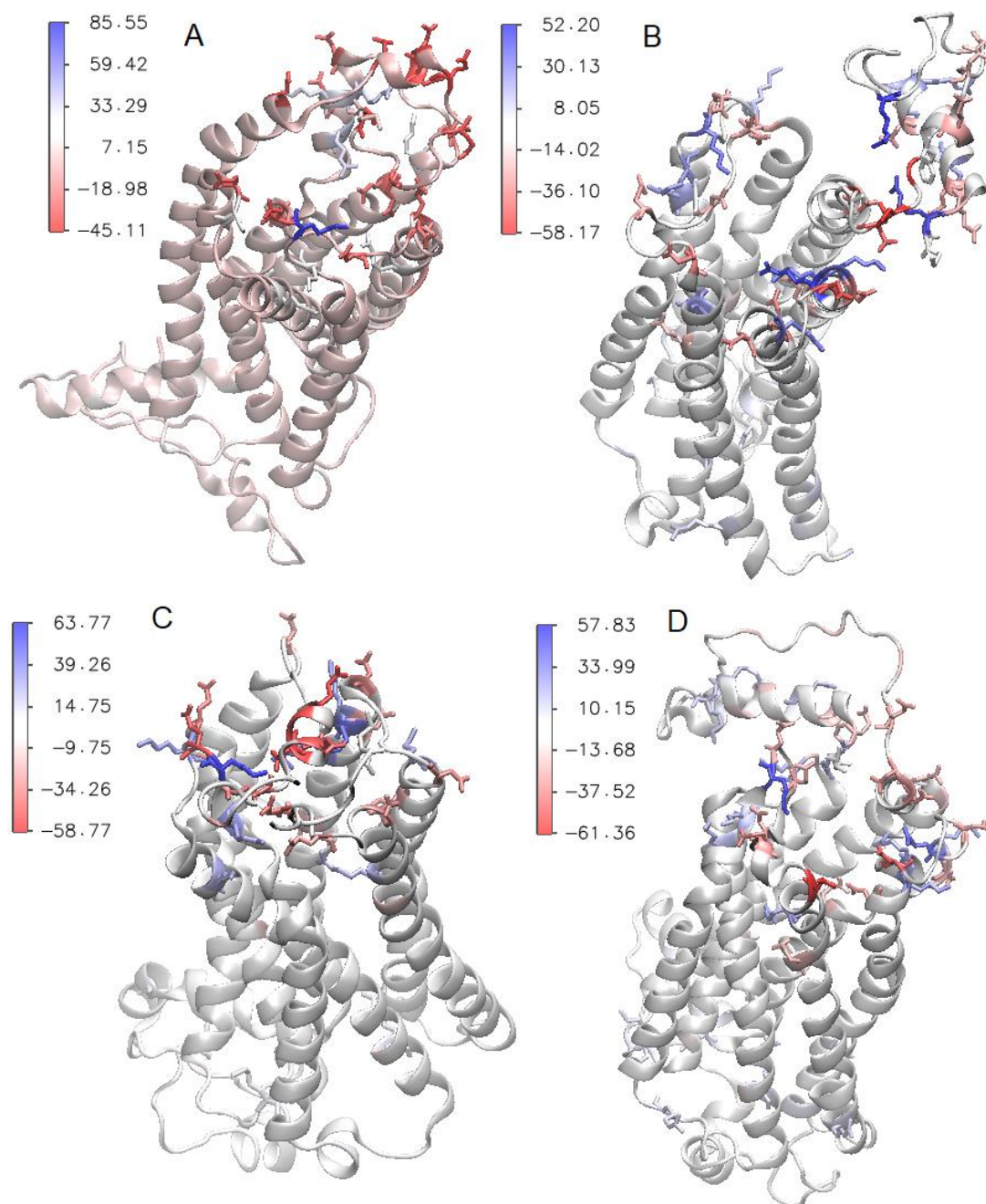
**TABELA 23:** Resíduos de aminoácidos da isoforma 4 do CX3CR1 e seus respectivos valores das contribuições favoráveis e desfavoráveis (acima do valor de corte média mais desvio padrão) para a energia livre de ligação total para a trajetória de DM do modelo estrutural determinado pelo servidor HADDOCK.

Resíduo	Contribuição favorável para $\Delta G_b^t$ (kJ.mol <sup>-1</sup> )	Resíduo	Contribuição desfavorável para $\Delta G_b^t$ (kJ.mol <sup>-1</sup> )
GLU-3	-25.5342 ± 0.2017	MET-1	29.0405 ± 0.4206
GLU-6	-23.0968 ± 0.2197	ARG-2	26.0239 ± 0.2595
GLU-45	-30.4233 ± 0.4161	LYS-9	20.4209 ± 0.1630
ASP-47	-36.5317 ± 0.3852	ARG-16	18.2530 ± 0.1167
ASP-48	-37.6053 ± 0.6651	LYS-17	17.9161 ± 0.1272
GLU-51	-28.3413 ± 0.3210	ARG-25	13.8358 ± 0.0744
GLU-204	-28.5697 ± 0.4203	LYS-95	12.0206 ± 0.1088
GLU-206	-32.5687 ± 0.3264	LYS-127	23.8010 ± 0.1733
ASP-210	-33.1682 ± 0.2923	LYS-135	21.8847 ± 0.1581
GLU-213	-35.4585 ± 0.5485	ARG-159	14.0413 ± 0.0875
GLU-217	-61.3591 ± 1.3769	ARG-173	13.3294 ± 0.0854
GLU-226	-25.6075 ± 0.2454	LYS-201	24.1631 ± 0.1795
GLU-286	-34.5830 ± 0.4608	LYS-203	57.8321 ± 0.7474
ASP-292	-50.6778 ± 0.9824	ARG-223	27.0191 ± 0.2537
ASP-298	-31.3157 ± 0.3250	ARG-245	14.2410 ± 0.0746
ASP-302	-29.1629 ± 0.2274	LYS-263	11.9060 ± 0.1074
GLU-311	-24.5189 ± 0.2210	LYS-289	32.1656 ± 0.2453
		ARG-300	45.8034 ± 0.5335
		LYS-301	25.2244 ± 0.1890
		ARG-304	29.3816 ± 0.3213
		LYS-331	11.6243 ± 0.0918
		ARG-334	10.4295 ± 0.0545

Fonte: Elaborado pelo autor.

Para demonstrar a localização destes resíduos apresentados nas Tabelas 20–23, foi feita a Figura 40 denotando os resíduos com contribuições favoráveis e desfavoráveis. Para isso, utilizou-se uma estrutura representativa da simulação de DM para cada modelo estrutural do complexo isoforma/peptídeo da G obtida da análise de clusterização (Figura 40). É mostrado que a maioria dos resíduos de contribuição significativa a energia livre de ligação pertencem as regiões de alças entre as hélices  $\alpha$  do barril de hélices das isoformas do CX3CR1, observando também que o N-terminal da proteína (região usada como sítio de *docking* do peptídeo da G) apresenta um maior número de resíduos com contribuição favorável para a energia total, como para as isoformas 2 e 4 (Figura 40-B e Figura 40-D).

**FIGURA 40:** Estrutura representativa das isoformas do receptor celular CX3CR1 com os resíduos que apresentam uma contribuição favoráveis e desfavoráveis significativas (acima da média mais desvio padrão) para a energia livre de ligação total para as simulações de DM dos modelos estruturais dos complexos com o peptídeo da G determinados pelo servidor HADDOCK. As contribuições favoráveis e desfavoráveis para as energias totais são representadas gradientes de cores indo vermelho para o azul, respectivamente. As barras de cores denotam o aspecto quantitativo de cada gradiente de cor. As isoformas 1 (A), 2 (B), 3 (C) e 4 (D) são apresentadas em modelos de *cartoon* em cinza e os resíduos de aminoácidos na forma de sticks.



Fonte: Elaborado pelo autor.

Uma comparação entre os resultados dos cálculos de energia livre de ligação para as trajetórias de DM dos modelos estruturais determinados pelo ClusPro (Tabela 16–19 e Figura 39) e pelo HADDOCK (Tabelas 20–23 e Figura 40) mostra que os resíduos ASP-15, ASP-16, GLU-19, GLU-172, ASP-178, GLU-181, GLU-185, ASP-266 e ASP-270 (numeração da isoforma 1) são responsáveis por uma contribuição favorável comum aos dois conjuntos, enquanto os resíduos MET-1, ARG-191, LYS-257, ARG-268 e LYS-269 (numeração da isoforma 1) contribuem de maneira desfavorável. Como mencionado antes, esses resíduos fazem parte do N-terminal das isoformas e também da região N-terminal do barril de hélices  $\alpha$ . Por conta do modelo dado pelo HADDOCK obter uma pose conformacional do peptídeo da G mais superficial, portanto obteve resíduos da parte mais externa do receptor CX3CR1. Uma diferença interessante é a contribuição de energia livre de ligação dos resíduos das folhas  $\beta$  utilizados como referência para a realização dos cálculos de *docking*. Para os resultados das trajetórias de DM dos modelos estruturais determinados pelo ClusPro, os resíduos GLU-172 e GLU-174 apresentam uma contribuição significativa para a energia total de todas as isoformas do CX3CR1, enquanto para os cálculos com modelo estrutural do HADDOCK apenas GLU-172 contribui significativamente para todos as isoformas. Esse resultado pode ter explicação no maior aterramento do peptídeo da G na ligação à porção N-terminal do barril de hélices  $\alpha$  do CX3CR1, o que favorece a interação com os resíduos das folhas  $\beta$ .

Os dados apresentados nas Tabelas 24 e 25 mostram as contribuições dos resíduos do peptídeo da G no modelo estrutural do ClusPro e do HADDOCK para a energia livre de ligação total, respectivamente.

**TABELA 24:** Valores das contribuições favoráveis e desfavoráveis para os resíduos de aminoácidos do peptídeo da G com as respectivas isoformas (1,2,3 e 4) obtidos para as simulações de DM dos modelos estruturais dos complexos determinados pelo servidor ClusPro.

Resíduo	Contribuição para $\Delta G_b^t$ (kJ.mol <sup>-1</sup> )			
	Isoforma 1	Isoforma 2	Isoforma 3	Isoforma 4
ASN-169	-1.0295 ± 0.5325	4.0033 ± 0.5416	-3.1637 ± 0.5913	6.0292 ± 0.8583
PHE-170	-9.1983 ± 0.3210	-11.7364 ± 0.5048	-2.8370 ± 0.1818	-9.3580 ± 0.3853
VAL-171	-10.0386 ± 0.2982	-11.4635 ± 0.2063	-4.6166 ± 0.1549	-9.3924 ± 0.2349
PRO-172	-22.4347 ± 0.3086	-5.2652 ± 0.2580	-2.4839 ± 0.1415	-12.3470 ± 0.1325
CYS-173	-17.9139 ± 0.2202	-3.7635 ± 0.1744	-5.3997 ± 0.1721	-11.4891 ± 0.1184
SER-174	-9.8494 ± 0.3557	-3.9553 ± 0.2713	-1.5464 ± 0.2068	-4.8913 ± 0.2646

ILE-175	-28.3479 ± 0.2968	-13.6331 ± 0.3402	-7.1646 ± 0.3244	-18.0592 ± 0.2054
CYS-176	-8.8155 ± 0.1989	-6.3404 ± 0.2453	-6.4663 ± 0.1903	-6.7289 ± 0.1931
SER-177	-6.7232 ± 0.3320	-2.3330 ± 0.2637	0.2445 ± 0.3698	-4.9567 ± 0.3141
ASN-178	-4.0945 ± 0.4353	8.2793 ± 0.7848	7.5908 ± 0.3901	2.7988 ± 0.5599
ASN-179	6.9847 ± 0.3574	-2.1421 ± 0.3652	5.2973 ± 0.3346	3.9440 ± 0.3514
PRO-180	-12.1393 ± 0.1978	-13.5741 ± 0.2822	-9.0748 ± 0.1565	-16.2681 ± 0.1905
THR-181	-5.2739 ± 0.1906	-4.6474 ± 0.2459	-4.2819 ± 0.2751	-8.5396 ± 0.3680
CYS-182	-1.5223 ± 0.0777	-5.0159 ± 0.1542	-4.3935 ± 0.2326	-7.5957 ± 0.1299
TRP-183	-18.1639 ± 0.2841	-17.7936 ± 0.4328	-12.6558 ± 0.2976	-12.5864 ± 0.2632
ALA-184	-4.1247 ± 0.1424	-1.9879 ± 0.1839	-1.3868 ± 0.2529	-3.7001 ± 0.1987
ILE-185	-12.9922 ± 0.2706	-11.1680 ± 0.3185	-11.2393 ± 0.2301	-18.4653 ± 0.2623
CYS-186	-3.1247 ± 0.1485	-3.1142 ± 0.0919	0.7537 ± 0.1328	-6.8925 ± 0.1721
LYS-187	-133.6001 ± 1.2906	-112.9848 ± 1.0048	-130.0099 ± 2.4406	-86.5694 ± 1.6229
ARG-188	-148.5841 ± 1.1378	-102.4266 ± 0.8962	-123.6633 ± 3.0624	-159.8975 ± 1.5243
ILE-189	-7.2126 ± 0.3510	-1.2960 ± 0.3010	0.2822 ± 0.2394	-22.5795 ± 0.4093

Fonte: Elaborado pelo autor.

**TABELA 25** Valores das contribuições favoráveis e desfavoráveis para os resíduos de aminoácidos do peptídeo da G com as respectivas isoformas (1,2,3 e 4) obtidos para as simulações de DM dos modelos estruturais dos complexos determinados pelo servidor HADDOCK.

Resíduo	Contribuição para $\Delta G_b^t$ (kJ.mol <sup>-1</sup> )			
	Isoforma 1	Isoforma 2	Isoforma 3	Isoforma 4
ASN-169	1.8701 ± 0.2801	1.7649 ± 0.4070	0.9715 ± 0.3121	-2.5258 ± 0.3799
PHE-170	-0.2897 ± 0.2553	-7.2147 ± 0.3735	1.9111 ± 0.0987	-1.1376 ± 0.0890
VAL-171	0.9447 ± 0.1159	-10.2543 ± 0.2120	-0.3949 ± 0.0887	0.9631 ± 0.0538
PRO-172	-5.7419 ± 0.1762	-4.0441 ± 0.0999	-3.6055 ± 0.2526	-0.6182 ± 0.0565
CYS-173	-8.1603 ± 0.2673	-3.4207 ± 0.2438	-4.6500 ± 0.3328	-1.5093 ± 0.1083
SER-174	-4.0923 ± 0.4056	-5.7096 ± 0.2058	-4.5199 ± 0.3670	-0.6804 ± 0.1696
ILE-175	-2.0256 ± 0.1201	-16.8984 ± 0.2085	-6.6689 ± 0.2916	-0.0353 ± 0.0882
CYS-176	-0.6824 ± 0.1725	-5.7819 ± 0.2812	-5.5699 ± 0.2027	-0.6270 ± 0.0966
SER-177	-4.1234 ± 0.2983	0.4673 ± 0.3103	-2.8450 ± 0.4695	0.4449 ± 0.2325
ASN-178	0.6674 ± 0.5490	-1.4640 ± 0.4288	-2.2376 ± 0.4576	2.8983 ± 0.4072
ASN-179	1.9182 ± 0.5178	4.8710 ± 0.3562	0.7820 ± 0.3871	-0.4731 ± 0.1567
PRO-180	-12.1521 ± 0.3029	-7.1975 ± 0.3002	-7.2135 ± 0.3648	-8.5086 ± 0.1753
THR-181	-8.2010 ± 0.3595	-0.4593 ± 0.2453	-1.0103 ± 0.1700	-0.8133 ± 0.1408
CYS-182	-2.2557 ± 0.0967	-1.2181 ± 0.0637	-2.7764 ± 0.1937	0.1855 ± 0.0727
TRP-183	-15.2818 ± 0.4819	-6.5006 ± 0.3463	-11.0354 ± 0.4415	-11.0858 ± 0.2938
ALA-184	-3.5873 ± 0.1999	0.0794 ± 0.0823	-0.2626 ± 0.1174	-0.9192 ± 0.1658
ILE-185	-2.8991 ± 0.0924	-0.6969 ± 0.0643	-1.3479 ± 0.0815	-1.4730 ± 0.0693
CYS-186	-3.4457 ± 0.1741	-0.4510 ± 0.0838	-1.0937 ± 0.0621	0.1292 ± 0.1914
LYS-187	-131.2469 ± 1.1040	-123.4543 ± 1.0765	-110.6768 ± 1.1504	-102.8973 ± 1.8372
ARG-188	-148.4333 ± 1.2962	-134.5803 ± 0.7859	-139.9564 ± 0.7784	-119.2138 ± 0.9984
ILE-189	-9.3482 ± 0.5784	-15.2215 ± 0.4525	-4.0811 ± 0.6104	-1.2784 ± 0.2141

Fonte: Elaborado pelo autor.

É possível observar que os resíduos LYS-187 e ARG-188 se destacam por um contribuição significativa para a energia livre de ligação total. Vale lembrar que a ARG-188 foi observada anteriormente por apresentar um alta porcentagem de persistência de ligação de hidrogênio. Como os resíduos de arginina e lisina são polares carregados positivamente, isso caracteriza uma interação de caráter eletrostático com os resíduos ácidos (ácido aspártico e glutâmico) da proteína, o que é possível graças ao posicionamento do peptídeo da G com suas extremidades interagindo com o N-terminal do barril de hélices  $\alpha$  de receptor celular.

**TABELA 26:** Valores de variação de energia livre de ligação total para a interação do peptídeo da G com as isoformas (1,2,3 e 4) calculados a partir das simulações de DM para os modelos estruturais dos complexos obtidos pelos servidores ClusPro e HADDOCK.

	Contribuição para $\Delta G_b^t$ (kJ.mol <sup>-1</sup> )	Contribuição para $\Delta G_b^t$ (kJ.mol <sup>-1</sup> )
	ClusPro	HADDOCK
Isoforma 1	-753,239 ± 4,763	-571,433 ± 4,099
Isoforma 2	-642,646 ± 5,437	-629,914 ± 3,702
Isoforma 3	-604,597 ± 7,896	-541,814 ± 5,352
Isoforma 4	-733,115 ± 6,432	-466,979 ± 5,813

Fonte: Elaborado pelo autor.

A Tabela 26 mostra a variação de energia livre de ligação total para a interação das isoformas do CX3CR1 com o peptídeo da G para as simulações de DM dos modelos estruturais determinados pelo ClusPro e HADDOCK. A partir da Tabela 26 é possível notar que, para o modelo do ClusPro, a isoforma 1 apresenta uma maior afinidade de ligação ao peptídeo comparada com as demais isoformas do CX3CR1, enquanto para o HADDOCK foi a isoforma 2. Por outro lado, a isoforma 3 para o ClusPro e a isoforma 4 para o HADDOCK apresentaram os menores valores da variação da energia livre de ligação total. Em um contexto em geral, levando em conta os modelos estruturais dos dois programas ClusPro e HADDOCK, as isoformas 1 e 2 apresentaram uma ligação mais favorável com o peptídeo da G (Tabela 26). Vale salientar que como os modelos estruturais determinados pelo ClusPro apresentaram variações de energia livre de ligação total mais favoráveis em comparação com os

valores de energia determinados para os modelos do HADDOCK. Esse resultado pode ser explicado devido a um maior enterramento do peptídeo no barril de hélices  $\alpha$  do receptor celular determinado pelo ClusPro.

## 5 CONCLUSÕES e PERSPECTIVAS

Os resultados obtidos ao longo deste trabalho com a utilização de simulações computacionais contribuem para o entendimento da interação das isoformas do receptor celular CX3CR1 com a região conservadas de cisteínas do ectodomínio da proteína G do Vírus Sincicial Respiratório humano (hRSV) ajudando na compreensão de detalhes moleculares do processo de adesão vírus/célula e também na busca de candidatos antivirais que poderiam atuar como potenciais inibidores dessa interação. As predições de estrutura tridimensional das isoformas do CX3CR1 calculadas no presente trabalho se mostraram estáveis ao longo de 300 ns de simulações de DM, indicando que a conformação em barril de hélices alfas, como a estrutura de cristal do receptor homólogo US28 (PDB 4XT1), é a mais provável para esse receptor celular. O modelo estrutural de interação da região conservada de cisteínas da proteína G com as isoformas do CX3CR1 também se mostrou estável em todas as simulações, sugerindo que a proposta de orientação do *loop* entre as hélices alfas do peptídeo em direção a fita beta GLU-174–GLY-177 (numeração referente a isoforma 1) na cavidade de ligação da porção N-terminal da proteína é aceitável e provável de ocorrer experimentalmente. Nesse contexto, os resíduos LYS-171, GLY-177, TYR-179, GLN-184 e ARG-272 da CX3CR1 (numeração referente a isoforma 1) destacam por sua importância para a estabilização da interação com o peptídeo, uma vez que participam de ligações de hidrogênio, outra característica observada são os resíduos do peptídeo da G, LYS-187 e ARG-188, os quais obtiveram a menor energia livre de ligação dos resíduos do peptídeo da G. Outro fator importante é as características traduzidas pelo RMSF e PCA, demonstrando como a ligação do peptídeo no N-terminal do receptor celular influencia o seu movimento, obtendo uma maior flexibilidade das regiões terminais para as formas ligadas, indicando que o peptídeo da G altera essas regiões. Os cálculos de energia livre de ligação destacaram um caráter de carga positiva para

o peptídeo, mostrando que os resíduos que favorecem a interação do peptídeo no receptor celular são ácidos, outro fator é a isoforma 1 tem maior favorecimento para o peptídeo da G, apresentando uma menor energia livre de ligação em torno de -700 KJ/mol. Vale mencionar que a configuração de interação do peptídeo da G na cavidade de ligação das isoformas é análoga a observada para a fractalcina (CX3CL1), o parceiro biológico do CX3CR1, com o receptor homólogo US28 (BURG *et al*, 2015).

Os passos realizados neste trabalho servem como uma locomoção para que pesquisas sejam feitas em busca de fármacos que reduzam ou inibam a adesão do Vírus Sincicial Respiratório humano (hRSV) na célula do hospedeiro, ocasionando um menor número de enfermos e conseqüentemente um menor número de hospitalizações e de óbitos.

## 6 REFERÊNCIAS

ARBIZA, J. et al. Genetic characterization of respiratory syncytial viruses isolated from consecutive acute respiratory infections in a HIV infected child. *Journal of Clinical Virology*, Amsterdam, v. 35, p. 41-45, 2006.

BAEK, M *et al.*, “Accurate prediction of protein structures and interactions using a three-track neural network,” *Science*, vol. 373, no. 6557, pp. 871–876, Aug. 2021, doi: 10.1126/SCIENCE.ABJ8754.

BAKER, N. A., SEPT, D., JOSEPH, S., HOLST, M. J., & McCAMMON, J. A. (2001). Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences of the United States of America*, 98(18), 10037–10041. <https://doi.org/10.1073/pnas.181342398>

BERGER, OLIVER, O. EDHOLM and Fritz JAHNING. 1997. “Molecular Dynamics Simulations of a Fluid Bilayer of Dipalmitoylphosphatidylcholine at Full Hydration, Constant Pressure, and Constant Temperature.” *Biophysical Journal* 72 (5): 2002–13. [https://doi.org/10.1016/S0006-3495\(97\)78845-3](https://doi.org/10.1016/S0006-3495(97)78845-3).

BOTOSSO, V. F.; ZANOTTO, P. M. A.; UEDA, M. et al. Positive selection results in frequent reversible aminoacid replacements in the G protein gene on human respiratory syncytial virus. *PLoS Pathogens*, San Francisco, v. 5, n. 1, p. faltam as páginas, 2009.

BURG, J.S., INGRAM, J., VENKATAKRISHNAN, A.J., JUDE, K.M., DUKKIPATI, A., *et al* *Structural basis for chemokine recognition and activation of a viral G protein-coupled receptor*. *SCIENCE* sciencemag.org 6 MARCH 2015 • VOL 347 ISSUE 6226 (2015)

C. L. BROOKS III, M. KARPLUS, B. M. PETTITT, *Proteins: A theoretical perspective of dynamics structure and Thermodynamics*, Advances in chemical physics, John Wiley & Sons, New York, 1988

CANE, P. A.; MATTHEWS, D. A.; PRINGLE, C. R. Identification of variable domains of the attachment (G) protein of subgroup A respiratory syncytial viruses. *Journal of General Virology*, London, v. 72, n. 9, p. 2091–2096, Sept. 1991.

CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput.* 12:405-413

CHIRKOVA, T et al. CX3CR1 is an important surface molecule for respiratory syncytial virus infection in human airway epithelial cells. *Journal of General Virology* (2015), 96, 2543–2556

COLLINS, P.L & GRAHAM, B.S. Viral and Host Factors in Human Respiratory Syncytial Virus Pathogenesis. *Journal of virology*, v. 82, n. 5, p. 2040–2055. 2008.

DELANO WL. 2002. *The PyMOL Molecular Graphics System*. DeLano Scientific, San Carlos, California, USA; <http://www.pymol.org>.

DE SÁ. J.M., PILOTO, J.V., CILLI, E.M., TASIC, L., FOSSEY, M.A., ALMEIDA, F.C.L., SOUZA, F.P., CARUSO, I.P. (2020): Hesperetin targets the hydrophobic pocket of the nucleoprotein/phosphoprotein binding site of human respiratory syncytial virus, *Journal of Biomolecular Structure and Dynamics*, DOI: 10.1080/07391102.2020.1835717

DESTA IT, PORTER KA, XIA B, KOZAKOV D, VAJDA S. Performance and Its Limits in Rigid Body Protein-Protein Docking. *Structure*. 2020 Sep; 28 (9):1071-1081

DILBECK, C.W and FINLAYSON-PITTS, B. J. s . (2013). Hydroxyl radical oxidation of phospholipid-coated NaCl particles. *Phys. Chem. Chem. Phys.*, 2013,15, 9833-9844 [doi.org/10.1039/C3CP51237A](https://doi.org/10.1039/C3CP51237A).

DOMINGUEZ, C., BOELEN, R. & BONVIN, A.M.J.J. HADDOCK: A proteinprotein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.* 125, 1731–1737 (2003).

ESWAR, N., JOHN, B., MIRKOVIC, N., FISER, A., ILYVIN, V. A., PIEPER, U., et al. (2003). Tools for comparative protein structure modeling and analysis. *Nucleic Acids Research*, 31(13), 3375-3380

FEDECHKIN, S.O., GEORGE, N.L., WOLFF, J.T., KAUVAR, L.M., DUBOIS, R.M. Structure of antibody 2D10 bound to the central conserved region of RSV G (2017) DOI: 10.2210/pdb5wn9/pdb

FERREIRA, L.G. *et al.* Molecular docking and structure-based drug design strategies. **Molecules** v. 20, pp. 13384-13421, 2015. DOI: 10.3390/molecules200713384.

FRISCH, M.J.; TRUCKS, G.W.; SCHLEGEL, H.B.; SCUSERIA, G.E.; ROBB, M.A.; et al.; Gaussian 09, Revision A.02. Gaussian Inc. Wallingford CT. Cheeseman, J.R, 2009.

GALLOUX, M.; TARUS, B.; BLAZEVIC, I.; FIX, J.; DUQUERROY, S.; ELÉOUËT, J.F. Characterization of a Viral Phosphoprotein Binding Site on the Surface of the Respiratory Syncytial Nucleoprotein. *Journal of Virology*, v. 86, p. 8375–8387, 2012.

GARCIA, O., MARTIN, M.; DOPAZO, ARBIZA, J.; FRABASILE, S.; RUSSI, J.; HORTAL, M; PEREZ-BREÑA, P.; MARTINEZ, I.; GARCIA-BARRENO, B. J. Evolutionary pattern of human respiratory syncytial virus (subgroup A): cocirculating lineages and correlation of genetic and antigenic changes in the G glycoprotein. *Journal of Virology*, Washington, v. 68, n. 9, p. 5448–5459, Sept. 1994.

GILSON, M. K.; HONIG, B., Calculation of the total electrostatic energy of a macromolecular system: solvation energies, binding energies, and conformational analysis. *Proteins* 1988, 4, 7-18.

GROTEWOLD, E. *The Science of Flavonoids*, New York: Springer, 2006.

HALL, C. B.; WEINBERG, G. A.; IWANE, M. K.; BLUMKIN, A. K.; EDWARDS, K. M.; STAAT, M. A.; AUINGER, P.; GRIFFIN, M. R.; POHLING, K. A.; ERDMAN, D.; GRIJALVA, C. G.; ZHU Y.; SZILAGYI P. The burden of respiratory syncytial virus infection in young children. *N. Engl. J. Med.*, v. 360, p. 588-598, 2009.

HOGAN, C.A., CAYA, C., PAPENBURG, J. Rapid and simple molecular tests for the detection of respiratory syncytial virus: a review. *Expert Review of Molecular Diagnostics*, 18(7), pp. 617–629. 2018. <http://www.rcsb.org/pdb/>

HOMEYER, N.; GOHLKE, H., Free Energy Calculations by the Molecular Mechanics Poisson–Boltzmann Surface Area Method. *Mol. Inform.* 2012, 31, 114-122.

JORGENSEN, WILLIAN L.; CHANDRASEKHAR, J; MADURA, JEFFRY D.; IMPERY, ROGER W.; KLEIN, M. L. *Journal of Chemical Physics* (1983), 79 (2), 926-35CODEN: JCPSA6; ISSN:0021-9606.

JOHNSON, P. R.; SPRIGGS, M. K.; OLMSTED, R. A.; COLLINS, P. L. The G glycoprotein of human respiratory syncytial viruses of subgroups A and B: extensive sequence divergence between antigenically related proteins. *Proceedings of the National Academy of Sciences of the United States of America*, Washington, v. 84, n. 16, p.5625–5629, Aug. 1987.

JO, S., KIM, T. IYER, V.G., and W. IM (2008) CHARMM-GUI: A Web-based Graphical User Interface for CHARMM. *J. Comput. Chem.* 29:1859-1865

KOZAKOV D, HALL DR, Xia B, PORTER KA, PADHORN D, YUEH C, BEGLOV D, VAJDA S. The ClusPro web server for protein-protein docking. *Nature Protocols*. 2017 Feb;12(2):255-278.

KOZAKOV D, BRENKE R, COMEAU SR, VAJDA S. PIPER: an FFT-based protein docking program with pairwise potentials. *Proteins*. 2006; 65:392–406. [PubMed: 16933295]

KUMARI, R.; KUMAR, R.; OSSD Consortium.; Lynn, A. g\_mmpbsa - A GROMACS tool for highthroughput MM-PBSA Calculations; *J. Chem. Inf. Model*, 2014

HOMEYER, N.; GOHLKE, H., Free Energy Calculations by the Molecular Mechanics Poisson–Boltzmann Surface Area Method. *Mol. Inform.* 2012, 31, 114-122.

HONORATO, R.V., KOUKOS, P.I., JIMENEZ-GARCIA, B., TSAREGORODTSEV, A., VERLATO, M., GIACHETTI, A., A. ROSATO and A.M.J.J. BONVIN (2021). "Structural biology in the clouds: The WeNMR-EOSC Ecosystem." *Frontiers Mol. Biosci.*, 8, fmolb.2021.729513.

KATCHALSKI-KATZIR E, et al. Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques. *Proc Natl Acad Sci U S A*. 1992; 89:2195–2199. [PubMed: 1549581]

KELLEY LA et al. The Phyre2 web portal for protein modeling, prediction and analysis *Nature Protocols* 10, 845-858 (2015)

KOZAKOV D, HALL DR, XIA B, PORTER KA, PADHORN D, YUEH C, BEGLOV D, VAJDA S. The ClusPro web server for protein-protein docking. *Nature Protocols*. 2017 Feb;12(2):255-278.

KOZAKOV D, BEGLOV D, BOHNUUD T, MOTARELLA S, XIA B, HALL DR, VAJDA, S. How good is automated protein docking? *Proteins: Structure, Function, and Bioinformatics*. 2013 Dec; 81(12):2159-66.

KUMARI, R., KUMAR, R., & LYNN, A, (2014). G-mmpbsa -A GROMACS tool for high-throughput MM-PBSA calculations. *Journal of Chemical Information and Modeling*, 54(7), 1951–1962

KUMARI, R.; KUMAR, R.; OSSD Consotium.; Lynn, A. g\_mmpbsa - A GROMACS tool for highthroughput MM-PBSA Calculations;J. Chem. Inf.model, 2014

LEVINE S.; KLAIBER-FRANCO, R.; PARADISO, P. R. Demonstration that glycoprotein G is the attachment protein of respiratory syncytial virus. *Journal of General Virology*, London, v. 68, n. 9, p. 2521-2524, Sept. 1987.

LIU C.; DAY, N. D.; BRANIGAN, P.; GUTSHALL, L. L.; SARISKY, R. T.; DEL VECCHIO, A. M. Relationship between the loss of neutralizing antibody binding and fusion activity of the F protein of human respiratory syncytial virus. *Viol. J.*, v. 4, 2007.

MALDE, A. K., ZUO, L., BREEZE, M., STROET, M., POGER, D., NAIR, P. C., OOSTENBRINK, C., & MARK, A. E. (2011). An automated force field topology builder (ATB) and repository: Version 1.0. *J Chem Theory Comput*, 7(12), 4026–4037. <https://doi.org/10.1021/ct200196m>

MENG, X.; ZHANG, H.; MEZEI, M.; CUI, M. Molecular Docking: A powerful approach for structure-based drug discovery, *Curr Comput Aided Drug Des.* **2011**; 7(2): 146-157

MORRIS, G. M.; LIM-WILBY, M. Molecular Docking, *Methods Mol Biol.* **2008**; 443: 365-382, Springer. 10.1007/978-1-59745-177-2\_19

MORRIS, J.A., BLOUNT JR, R.E SAVAGE Recovery of cytopathogenic agente from chimpanzees with goryza. *Proceedings of the Society for Experimental Biology and Medicine*, 92(3), pp. 544-549. 1956.

MEISSNER, H. C. et al. Revised indications for the use of palivizumab and respiratory syncytial virus immune globulin intravenous for the prevention of respiratory syncytial virus infections. *Pediatrics*, v. 112, p. 1447-1452, Dec. 2003.

MOHAMED.OUIZOUGUN OUBARIETAL; A Druggable Pocketatthe Nucleocapsid/Phosphoprotein Interaction Site of Human Respiratory Syncytial Virus November 2015 Volume 89 Number 21.

NOKES, J. D.; CANE, P. A. New strategies for control of respiratory syncytial virus infection: review. *Current Opinion of Infectious Diseases*, Philadelphia, v. 21, n. 6, p. 639- 643, Dec. 2008.

SCHMID, N., EICHENBERGER, A. P., CHOUTKO, A., RINIKER, S., WINGER, M., MARK, A. E., & Van GUNSTEREN, W. F. (2011). Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *European Biophysics Journal: Ebj*, 40(7), 843–856. <https://doi.org/10.1007/s00249-011-0700-9>

OLSSON, M. H. M., SØNDERGARD, C. R., ROSTKOWSK, M., & JENSEN, J. H. (2011).

PEREZ, M., GARCIA-BARRENO, B., MELERO, J.A., CARRASCO, L., and GUINEA, R. (1997). Membrane Permeability Changes Induced in *Escherichia coli* by the SH Protein of Human Respiratory Syncytial Virus. *Virology* 235, 342–351.

PROPKA3: Consistent treatment of internal and surface residues in empirical pKa predictions. *Journal of Chemical Theory and Computation*, 7(2), 525–537.

ÖLUND, M. R.; LINDELL, A. T. ; STENLER, S.; RIEDEL, H. M; WIRGART, B. Z.; GRILLNER. L. Molecular epidemiology and genetic variability of respiratory syncytial virus (RSV) in Stockholm, 2002–2003. *Journal of Medical Virology*, New York, v. 80, n. 1, p. 159–167, Jan. 2008.

OTHUMPANGAT, S.; GIBSON, L. F.; SAMSELL, L.; PIEDIMONTE, G. NGF is essential survival factor for bronchial epithelial cells during respiratory syncytial virus infection. PLoS ONE, v. 04, p. 1-13, 2009.

RAMACHANDRAN, G. N.; SASISEKHARAN, V.; Conformation of polypeptides and proteins. Advances in Protein Chemistry, New York, v. 23, p. 283-487, 1968.

REED, J.L., WELLIVER, T.P., SIMS, G.P., MCKINNEY, L., VELOZO, L., AVENDANO, L., HINTZ, K., LUMA, J., COYLE, A.J., and WELLIVER, R.C., Sr. (2009). Innate immune signals modulate antiviral and polyreactive antibody responses during severe respiratory syncytial virus infection. J. Infect. Dis. 199, 1128–1138.

RICCETTO, A. G. L.; SILVA, L. H. A.; SPILKI, F. R.; MORCILLO, A. M.; ARNS, C. W.; BARACAT, E. C. E. Genotypes and clinical data of respiratory syncytial virus and metapneumovirus in brazilian infants: A new perspective. Braz. J. Infect. Dis., v. 13, n. 1, p 35-39, 2009.

SABINO, F. No fim dá Certo, 1998, p. 7.

SENIOR, A.W., EVANS, R., JUMPER, J. *et al.* Improved protein structure prediction using potentials from deep learning. *Nature* **577**, 706–710 (2020). <https://doi.org/10.1038/s41586-019-1923-7>

SIT, A. Solving distance geometry problems for protein structure determination. Tese (Doutorado) — Iowa State University, 2010.

SHI, T., MCALLISTER, D. A., O'BRIEN, K. L., SIMOES, E. A. F., MADHI, S. A., GESSNER, B. D., et al. (2017). Global, regional, and national disease burden estimates of acute lower respiratory infections due to respiratory syncytial virus in young children in 2015: a systematic review and modelling study. *The Lancet*, 390(10098), 946–958, 2017.

SÖDING, J. (2005). Protein homology detection by HMM-HMM comparison. *Bioinformatics*, 21(7), 951-960.

SULLENDER, W. M. Respiratory syncytial virus genetic and antigenic diversity. *Clinical Microbiology Review*, Washington, v. 13, n. 1, p. 1-15, Jan. 2000.

TRAN, T. L., CASTAGNE, N., BHELLA, D., VARELA, P. F., BERNARD, J., CHILMONCZYK, S.,

WALLACE, A. C., LASKOWSK, R. A., & THORNTHON, J. M. (1995). LIGPLOT: A Program to Generate Schematic Diagrams of Protein-Ligand Interactions. *Protein Engineering*, 8(2), 127–134.

VAJDA S, YUEH C, BEGLOV D, BOHNUUD T, MOTARELLA SE, XIA B, HALL DR, KOZAKOV D. New additions to the ClusPro server motivated by CAPRI. *Proteins: Structure, Function, and Bioinformatics*. 2017 Mar; 85(3):435-444.

de VRIES, S.J., van DIJK, A.D.J. & BONVIN, A.M.J.J. WHISCY: What information does surface conservation yield? Application to data-driven docking. *Proteins* 63, 479–489 (2006).

de VRIES, S.J. & BONVIN, A.M. How proteins get in touch: interface prediction in the study of biomolecular complexes. *Curr. Protein Pept. Sci.* 9, 394–406 (2008).

de VRIES, S.J. et al. HADDOCK versus HADDOCK: new features and performance of HADDOCK2.0 on the CAPRI targets. *Proteins* (2007).

G.C.P VAN ZUNDERT, J.P.G.L.M. RODRIGUES, M. TRELLET, C. SCHMITZ, P.L. KASTRITIS, E. KARACA, A.S.J. MELQUIOND, M. van DIJK, S.J. de VRIES and A.M.J.J. BONVIN (2016). "The HADDOCK2.2 webserver: User-friendly integrative modeling of biomolecular complexes." *J. Mol. Biol.*, 428, 720-725 (2015).

WELLIVER, T.P., GAROFALO, R.P., HOSAKOTE, Y., HINTZ, K.H., AVENDANO, L., SANCHEZ, K., VELOSO, L., JAFRI, H., CHAVEZ-BUENO, S., OGRA, P.L., et al.

Severe human lower respiratory tract illness caused by respiratory syncytial virus and influenza virus is characterized by the absence of pulmonary cytotoxic lymphocyte responses. *J. Infect. Dis.* 195, 1126–1136, 2007

W. F. van GUNSTEREN, H. J. C. BERENDSEN, *Angew. Chem. Int. Ed. Engl.*, 29 (1990) 992.

W. F. van GUNSTEREN, H. J. C. BERENDSEN, *Groningen Molecular Simulation (GROMOS) Library Manual*, Biomos, Groningen, The Netherlands, 1987.

YE, X; IWUCHUKWU, O.P.; AVADHANULA, V.; AIDEYAN, L.O.; MCBRIDE, T.J.; FERLIC-STARK, L.L.; PATEL, K.D.; PIEDRA, F.A.; SHAH, D.P.; CHEMALY, R.F.; PIEDRA, P.A. Antigenic Site-Specific Competitive Antibody Responses to the Fusion Protein of Respiratory Syncytial Virus Were Associated With Viral Clearance in Hematopoietic Cell Transplantation Adults. *Front Immunol*, v. Mar 29, p-10:706, 2019.

YANG, J., YAN, R., ROY, A., XU, D., POISSON, J., ZHANG, Y. The I-TASSER Suite: Protein structure and function prediction. *Nature Methods*, 12: 7-8 (2015)

YANG, J., ANISHCHENKO, I., PARK, H., PENG, Z., OVCHINNIKOV, S., BAKER, D. Improved protein structure prediction using predicted interresidue orientations. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 1496–1503 (2020).

YERSHOVA A, JAIN S, LAVALLE SM, MITCHELL JC. Generating Uniform Incremental Grids on SO(3) Using the Hopf Fibration. *Int J Robot Res.* 2010; 29:801–812.

YUNTA, M.J.R. Docking and ligand binding affinity: uses and pitfalls. *American Journal of Modeling and Optimization* v. 4, n.3, pp. 74-114, 2016. DOI: 10.12691/ajmo-4-3-2 .

ZHANG, W., CHOI, Y., HAYNES, L.M., HARCOUT, J.L., ANDERSON, L.J., JONES, L.P and TRIPP, R.A (2010). Vaccination To Induce Antibodies Blocking the CX3C-CX3CR1 Interaction of Respiratory Syncytial Virus G Protein Reduces Pulmonary

Inflammation and Virus Replication in Mice *Journal of Virology*, p. 1148–1157 Vol. 84, No. 2 doi:10.1128/JVI.01755-09

ZHIVAKI, D., LEMOINE, S., LIM, A., MORVA, A., *et al* (2017). Respiratory Syncytial Virus Infects Regulatory B Cells in Human Neonates via Chemokine Receptor CX3CR1 and Promotes Lung Disease Severity. *Immunity* 46, 301–314. <http://dx.doi.org/10.1016/j.immuni.2017.01.010>