



UNIVERSIDADE FEDERAL DE SERGIPE
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA

**AVALIAÇÃO DA DINÂMICA ESTRUTURAL DO
DOMÍNIO VARIÁVEL DE ANTICORPOS DE CADEIA
PESADA ÚNICA DERIVADOS DE CAMELÍDEOS (VHH)**

Dissertação de Mestrado

WENNY CAMILLA DOS SANTOS ADAN



São Cristóvão – Sergipe

2020

UNIVERSIDADE FEDERAL DE SERGIPE
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA

WENNY CAMILLA DOS SANTOS ADAN

**AVALIAÇÃO DA DINÂMICA ESTRUTURAL DO
DOMÍNIO VARIÁVEL DE ANTICORPOS DE CADEIA
PESADA ÚNICA DERIVADOS DE CAMELÍDEOS (VHH)**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Química da Universidade Federal de Sergipe como requisito para a obtenção do título de Mestre em Química.

Orientador(a): Dr. Nivan Bezerra da Costa Jr.
Coorientador(a): Dr. Roberto Dias Lins Neto

São Cristóvão – Sergipe

2020

**FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL
UNIVERSIDADE FEDERAL DE SERGIPE**

Adan, Wenny Camilla dos Santos
A221a Avaliação da dinâmica estrutural do domínio variável de anticorpos de cadeia pesada única derivados de camelídeos (VHH) / Wenny Camilla dos Santos Adan ; orientador Nivan Bezerra da Costa Junior - São Cristóvão, 2020.
74 f. : il.

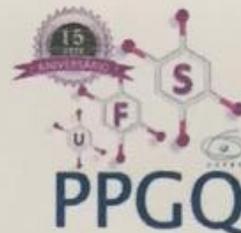
Dissertação (mestrado em Química) – Universidade Federal de Sergipe, 2020.

1. Química. 2. Autoanticorpos. 3. Dinâmica molecular. I. Costa Junior, Nivan Bezerra da orient. II. Título.

CDU 54



SERVIÇO PÚBLICO FEDERAL
MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE SERGIPE
Programa de Pós-Graduação em Química
PPGQ



FOLHA DE APROVAÇÃO

Membros da Comissão Julgadora da Dissertação de Mestrado de Wenny Camilla dos Santos Adan apresentada ao Programa de Pós-Graduação em Química da Universidade Federal de Sergipe em 28/02/2020.

Prof. Dr. Nivân Bezerra da Costa Junior
Departamento de Química - UFS

Prof. Dr. Luis Eduardo Almeida
Departamento de Ciência e Engenharia de Materiais - UFS

Prof. Dr. Tiago Branquinho Oliveira
Departamento de farmácia - UFS

Agradecimentos

“This most beautiful System of the Sun, Planets, and Comets, could only proceed from the counsel and dominion of an intelligent and powerful being.” (Isaac Newton, The Mathematical Principles of Natural Philosophy, trans. by Andrew Motte (London, 1729), pp. 387-393).

Iniciando por meio das dimensões filosóficas da ciência gostaria de agradecer ao Deus em que acredito por ter me sustentado até aqui. Agradeço ao meu orientador Dr. Nivan Bezerra que me introduziu à ciência e me impulsionou a dar passos maiores. Agradeço ao Dr. Roberto Lins pelas oportunidades e, principalmente, por ter recuperado o meu interesse pela ciência. Agradeço a todos os meus colegas do BIOMAT/FIOCRUZ-PE por todo o aprendizado e por terem me acolhido tão bem no Recife. Agradeço à minha mãe Ana Cristina Alves (*in memoriam*), mulher forte e determinada, que desde cedo enxergou a criança cientista que havia em mim, priorizando os meus estudos, nunca me deixando desviar do caminho científico ou desacreditar da minha capacidade. Agradeço à minha avó Lió Alves, que não me deixava faltar um dia de aula se quer, por toda criação e infância maravilhosa que pôde me proporcionar. Agradeço à minha irmã Ana Alves por toda a sua lealdade, amor e por acreditar que eu sempre irei solucionar qualquer problema. Agradeço aos meus tios Ezequiel e Kátia Alves, vocês são os pilares da nossa família. Agradeço ao meu pai Wallace Adan pelas conversas motivadoras, aos meus primos Caroline, Diego, Bruno, Júnior Alves e Átina Adan. Agradeço à Sabrina Porto por todo amor, companheirismo e por acreditar no meu potencial sempre com bom otimismo tomando as minhas causas como se fossem suas. Agradeço aos meus irmãos de famílias distintas Jéssica Barreto, Hector Santana e Julie Alencar. Agradeço aos amigos, os melhores, Raysa Ramos, Lorena Argôlo, Monalisa Montalvão e Matheus Ferraz. Agradeço à Camila Oliveira e toda a sua família por terem me acolhido tão bem no Recife. Por fim, agradeço a minha casa científica, Universidade Federal de Sergipe-UFS, por todo o suporte e por ter me proporcionado o contato com pessoas incríveis, destaco Isaac Matos. Diante da impossibilidade de citar todas as pessoas que se fizeram presente nesse período conturbado, deixo aqui o meu muito obrigada por fazerem parte da minha história.

Resumo

Anticorpos de cadeia pesada única derivados de camelídeos (HCAb) utilizam o domínio variável N-terminal para estabelecer a interação com o antígeno. Quando isolado, este domínio recebe o nome de VHH. Apesar de possuírem metade do tamanho da região de reconhecimento ao antígeno de anticorpos convencionais, os VHHs apresentam especificidade e afinidade equivalentes por antígenos. Para se manterem funcionais mesmo com o tamanho reduzidos, os VHHs sofreram adaptações evolutivas e as principais modificações se concentram nas regiões determinantes de complementariedade (CDR1-3), constituídas majoritariamente por alças, cuja função é o reconhecimento ao antígeno. A informação a respeito da dinâmica estrutural dessas regiões frente a estabilidade, flexibilidade e afinidade de VHHs ainda é elusiva e o entendimento das bases moleculares envolvidas na interação VHH-antígeno podem e contribuir para o desenho racional de anticorpos sintéticos. Através de simulações por dinâmica molecular, o presente estudo identificou e caracterizou o perfil dinâmico estrutural das alças CDR1-3 para 40 estruturas de VHHs. Diferentemente do que era esperado, a amplitude da diversidade conformacional e a flexibilidade não possuem correlação com o comprimento de alças CDR. Dados de RMSF referentes a alça CDR3 alcançaram uma correlação de $R^2 = 0.71$ com a energia livre de ligação (ΔG_{bind} experimental, indicando que um aumento na flexibilidade desta alça leva os VHHs a maiores afinidades de ligação. Estas informações são inéditas na literatura de VHHs e suplantam a hipótese de que a rigidificação ou flexibilização de alças é apenas um dos mecanismos biofísicos que os VHHs utilizam para aumentar a afinidade. Gráficos de Ramachandran apontaram um perfil bem estruturado para a CDR3, destacando a presença de uma região característica de α -hélices. De fato, alças muito longas possuem sabidamente uma tendência de modular a sua região C-terminal. Por fim, a construção do índice de acessibilidade ao solvente (IAS) permitiu a classificação do caráter químico das regiões CDR, apontando a identidade dos resíduos e a sua propensão de conferir especificidade ou assumir funções estruturais.

Palavras-chave: HCAb. VHH. Nanocorpos. Dinâmica Molecular.

Abstract

Camelid-derived Heavy-chain antibodies (HCAb) use the N-terminal variable domain to establish interaction with the antigen. When Isolated, this domain is referred as VHH. Despite having half the size of the antigen recognition region of conventional antibodies, VHHs have equivalent specificity and affinity for antigens. In order to remain functional even with a small size, the VHHs have undergone evolutionary adaptations and the main modifications are concentrated in the complementarity determining region (CDR1-3), constituted mainly by loops, whose function is the recognition of the antigen. The information regarding the structural dynamics of these regions in view of the stability, flexibility and affinity of VHHs is still elusive and the understanding of the molecular bases involved in the VHH-antigen interaction can contribute to the rational design of synthetic antibodies. Through molecular dynamics simulations, the present study identified and characterized the structural dynamic profile of the CDR1-3 loops for 40 VHH structures. Contrary to what was expected, the amplitude of conformational diversity and flexibility are not correlated with the length of CDR loops. RMSF data regarding the CDR3 loop achieved a correlation of $R^2 = 0.71$ with the experimental bind free energy (ΔG_{bind}), indicating that an increase in the flexibility of this loop leads the VHHs to higher binding affinities. This information is unprecedented in the literature of VHHs and supersedes the hypothesis that the rigidity or flexibility of loops is only one of the biophysical mechanisms that VHHs use to increase affinity. Ramachandran plots showed a well-structured profile for CDR3, highlighting the presence of a characteristic region of α -helices. In fact, very long loops are known to have a tendency to modulate their C-terminal region. Finally, the construction of the solvent accessibility index (SAI) allowed the classification of the chemical character of the CDR regions, pointing out the identity of the residues and their propensity to confer specificity or assume structural functions.

Keywords: VHH. Nanobody. Flexibility. Molecular Dynamics Simulation.

Lista de ilustrações

Figura 1 – Anticorpos exercendo a função efetora neutralizante frente a antígenos (Vírus, Bactérias, Parasitas, Fungos e Proteínas estranha	12
Figura 2 – Representação gráfica da arquitetura molecular dos anticorpos convencionais e anticorpos encontrado nas espécies de camelídeos (HCAbs).	13
Figura 3 – . Estrutura da IgG com destaque para a região VH-VL e as suas 6 alças CDR.	17
Figura 4 – . Fórmula estrutural de uma folha- composta por duas fitas- β	18
Figura 5 – . Esquema estrutural da IgG1, IgG2 e IgG3.	19
Figura 6 – . Esquema estrutural de um VHH com destaque para seu enovelamento da imunoglobulina.	20
Figura 7 – Representação do oscilador harmônico através da modelo massa mola.	21
Figura 8 – . - Ilustração do método de integração leap-frog.	28
Figura 9 – - Ilustração do método de integração leap-frog.	28
Figura 10 – - Ilustração do método de integração leap-frog.	29
Figura 11 – -- Ressonância apresentada pela ligação peptídica.	35
Figura 12 – -- Ressonância apresentada pela ligação peptídica.	35
Figura 13 – Representação gráfica da estrutura de VHH de código PDB 2X6M. As regiões de CDR estão destacadas da seguinte maneira: CDR1 (preto), CDR2 (Vermelho) e CDR3 (Azul).	37
Figura 14 – RMSD médio de cada região de CDR função do comprimento de CDR.	43
Figura 15 – A estrutura A é a representação tridimensional do VHH anti α -HCG; Estrutura B é a representação tridimensional do VHH anti <i>alpha</i> -lisozima. As alças CDR3 estão destacadas em vermelho.	45
Figura 16 – Regiões do gráfico de Ramachandran para estruturas secundárias.	46
Figura 17 – Gráfico de Ramachandran em 2D e 3D, a região contornada em vermelho é a mais favorável e a mais populosa; a amarela permitida e a azul aceitável. A legenda traz o número de combinações encontradas em determinada região (população).	47
Figura 18 – Gráfico do RMSF médio de cada região de CDR função do comprimento de CDR.	51
Figura 19 – Correlação entre a energia livre de ligação e o RMSF médio para o CDR1-3 de 9 estruturas de VHH.	52
Figura 20 – índice de acessibilidade ao solvente para o CDR1-3.	54
Figura 21 – Gráfico de Conservação de sequência; as regiões referentes aos CDR1, 2 e 3 estão demarcadas por bastões em preto, vermelho e azul.	55
Figura 22 – Caráter químico das regiões CDR1-3 de acordo com as contribuições para a área acessível ao solvente.	57

Figura 23 – Gráficos de RMSD em função do tempo de simulação referente ao Backbone (C, N, O) gerados por dinâmica molecular paraca cada uma das estruturas de VHH.	65
Figura 29 – Gráficos de RMSF em função do tempo de simulação referente ao Backbone (C, N, O) gerados por dinâmica molecular paraca cada uma das estruturas de VHH.	71

Lista de tabelas

Tabela 1 – Mapeamento de sequência e posição dos aminoácidos que compõem a CDR1-3 em cada uma das estruturas de VHH.	38
Tabela 1 – Mapeamento de sequência e posição dos aminoácidos que compõem a CDR1-3 em cada uma das estruturas de VHH.	39
Tabela 2 – Valores médios de RMSD e seus respectivos desvios padrão para cada uma das estruturas, acompanhadas da quantidade de resíduos que compõem cada CDR.	41
Tabela 2 – Valores médios de RMSD e seus respectivos desvios padrão para cada uma das estruturas, acompanhadas da quantidade de resíduos que compõem cada CDR.	42
Tabela 3 – Valores médios de RMSF e seus respectivos desvios padrão para cada uma das estruturas, acompanhadas da quantidade de resíduos que compõem cada CDR.	49
Tabela 3 – Valores médios de RMSF e seus respectivos desvios padrão para cada uma das estruturas, acompanhadas da quantidade de resíduos que compõem cada CDR.	50

Sumário

1	Introdução	12
2	Fundamentação Teórica	16
2.1	Anticorpos	16
2.2	VHH	18
2.3	Mecânica Molecular	21
2.3.1	Campo de Força	23
2.4	Dinâmica Molecular	25
2.4.1	Método das diferenças finitas	27
2.4.2	Condições Periódicas de Contorno	28
2.4.3	Mecânica Estatística: Espaço de fase e Ensemble	29
3	Objetivos	31
3.1	Objetivo Geral	31
3.1.1	Objetivos Específicos	31
4	Materiais e Métodos	32
4.1	Construção da biblioteca de estruturas de VHH e preparação das estruturas	32
4.2	Dinâmica Molecular	33
4.2.1	Cálculo do RMSD	33
4.2.2	Cálculo do RMSF	33
4.2.3	Cálculo do SASA	34
4.2.4	Identificação das CDRs nas estruturas de VHHs	34
4.2.5	Construção dos gráficos de RMSD e RMSF	34
4.2.6	Gráficos de Ramachandran	34
4.2.7	Alinhamento múltiplo de sequência e conservação de resíduo por posição	35
5	Resultados e Discussão	37
5.1	Identificação das regiões CDRs nas estruturas	37
5.2	Análise das simulações por dinâmica molecular	40
5.2.1	RMSD	40
5.2.2	Ramachandran	46
5.2.3	RMSF	48
5.2.4	Análise da flexibilidade em função da energia livre de ligação de VHHs	52
5.2.5	SASA e Conservação de Sequência	53
6	Conclusões	58

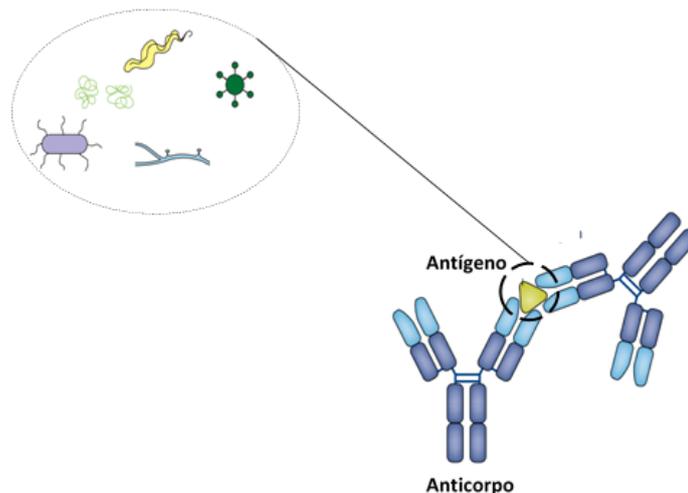
Referências	59
Anexos	64
ANEXO A RMSD	65
ANEXO B RMSF	71

1

Introdução

Os anticorpos, também designados como imunoglobulinas (IGs), são proteínas que desempenham um papel central no sistema imune de humanos e outros animais vertebrados. Essas moléculas, produzidas via linfócitos B¹, dispõem da capacidade de reconhecer e ligar-se com alta especificidade e afinidade a patógenos ou substâncias químicas (antígenos), tais como proteínas de vírus e bactérias, provocando ações diversas denominadas respostas imunes que atuarão no sentido de neutralizar o alvo [Figura 1](#) [1, 2]

Figura 1 – Anticorpos exercendo a função efetora neutralizante frente a antígenos (Vírus, Bactérias, Parasitas, Fungos e Proteínas estranha)

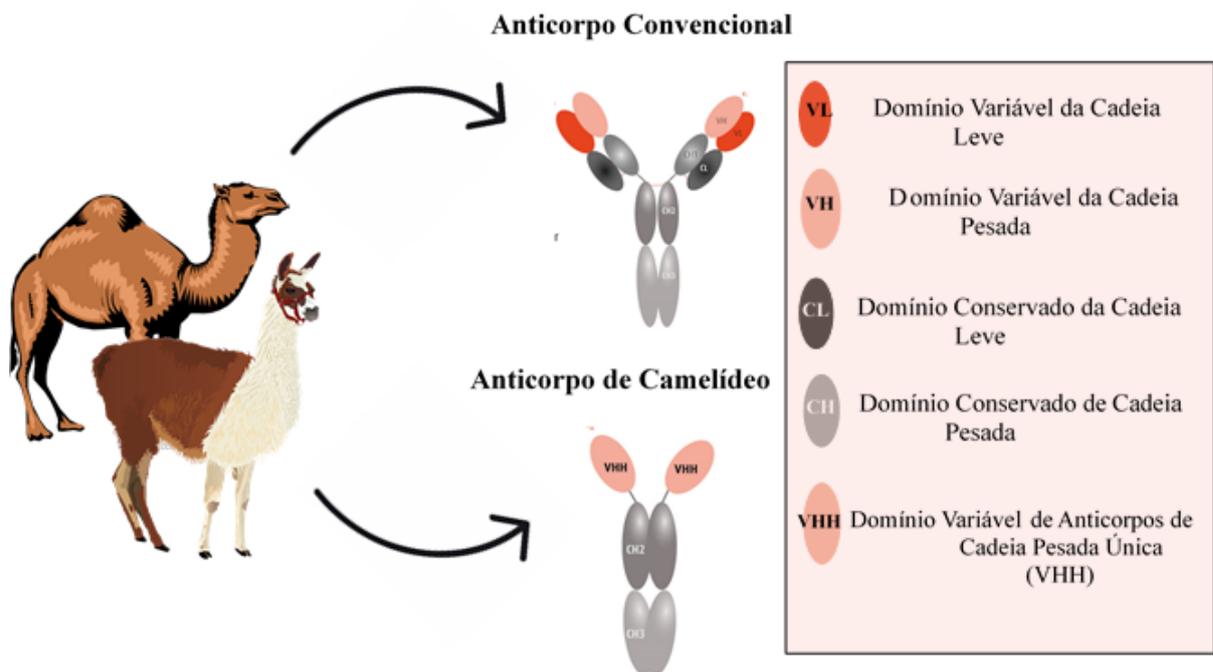


Fonte: Adaptado de ABBAS (2015)[1] e SCHIMD (2019)[3].

¹ Linfócito B é um tipo de leucócito presente no sistema imune e a sua principal função é a produção de anticorpos para o reconhecimento de antígenos.

As moléculas de anticorpo possuem uma unidade estrutural básica altamente conservada em mamíferos, formada por 2 cadeias leves idênticas (25 kDa) e 2 cadeias pesadas também idênticas (50 kDa) **Figura 2** (Figura 2) [6]. A região responsável pelo reconhecimento antigênico é composta pelo emparelhamento dos domínios variáveis da cadeia pesada (VH) e da cadeia leve (VL) [1,2,7,8]. A especificidade ao alvo é mediada por pequenos trechos, constituídos majoritariamente por alças, as regiões hipervariáveis ou regiões determinantes da complementaridade (CDRs, do inglês Complementarity-Determining Regions). Juntos, os domínios variáveis (VH-VL) possuem 6 alças CDR integrando o paratopo de ligação ao antígeno. Em torno das CDRs, atuando como um arcabouço de suporte às alças, estão posicionadas as regiões de *framework* cuja sequência e estrutura são conservadas.

Figura 2 – Representação gráfica da arquitetura molecular dos anticorpos convencionais e anticorpos encontrado nas espécies de camelídeos (HCABs).



Fonte: Autora.

Essa visão canônica acerca da arquitetura molecular dos anticorpos foi alterada drasticamente após a descoberta de que, além dos anticorpos convencionais, as espécies da família Camelidae possuíam um novo tipo de anticorpo, desprovido da cadeia leve e do primeiro domínio constante da cadeia pesada (CH1) **Figura 2** [9,10,3]. Com base em sua estrutura peculiar, esses anticorpos foram denominados de anticorpos de cadeia pesada única (HCABs, do inglês Heavy Chain-Only Antibodies). Posto que os HCABs são desprovidos de cadeia leve e, portanto, de um domínio VL, o reconhecimento antigênico se dá unicamente através do domínio variável N-terminal e, quando isolado, este domínio é denominado de VHH (do inglês, Variable Domain of the Heavy-Chain of Heavy-Chain Antibodies) **Figura 2**.

Devido ao tamanho reduzido, os VHHS possuem apenas 3 alças integrando a sua região determinante da complementaridade (CDR1, CDR2 e CDR3) [17]. Entretanto, quando comparados à região correspondente nos anticorpos convencionais (VH-VL), com aproximadamente o dobro de seu tamanho, os VHHs mantêm afinidade e especificidade equivalentes frente a antígenos [9,18]. Além das características usuais de anticorpos, VHHs possuem propriedades biofísicas, bioquímicas e farmacológicas exclusivas, como estabilidade térmica, resistência a pHs extremos e alta solubilidade [8,12]. Combinado a isso, a sua produção custa um quinto do valor de produção dos anticorpos convencionais, uma vez que não requer o uso de células de mamíferos [10]. Em face das suas características atrativas, os VHHs vêm sendo amplamente aplicados em diversos campos de pesquisa e na indústria biotecnológica. Alguns exemplos dessas aplicações incluem o direcionamento de VHHs (i) contra o coronavírus da síndrome respiratória aguda grave 2 (SARS-CoV-2); (ii) na imagiologia tumoral (iii) contra toxinas de venenos.; (iv) como terapia antiviral para HIV-1.

A estrutura de VHHs é similar àquela do domínio VH dos anticorpos convencionais. No entanto, para manter a sua capacidade de ligação a antígenos em um formato de domínio único, os VHHs adquiriram ao longo da evolução importantes modificações de sequência e estrutura. Uma das distinções destacadas entre VHH e VH ocorre na região framework-2 (FR2). Esta região no domínio VH é altamente conservada e acomoda aminoácidos hidrofóbicos que servem como um ponto de ancoragem para o domínio VL. Em contraste, na região correspondente em VHHs a substituição dessas espécies hidrofóbicas por aminoácidos hidrofílicos ocorre da seguinte maneira: V37F/Y, G44E/Q, L45R e W47G/F/L (com base na numeração de Kabat). A reconstrução da superfície do framework-2 (FR2) em VHHs explica tanto a sua falha ao se associar a um domínio VL quanto a sua alta solubilidade (18). Outras três características distintas estão situadas na região hipervariável dos VHHs. Primeiramente, a alça CDR1 é ampliada em direção a região N-terminal devido a mutagênese no códon dos resíduos 28 e 30, resultando em hotspots para hipermutação somática. Em segundo, os VHHs apresentam, em média, alças CDR3 mais longas. O comprimento médio da CDR3 em VHHs de dromedários, por exemplo, é de 18 resíduos comparado com 14 e 11 resíduos observados em VHs humanos e camundongos, respectivamente. Por fim, em adição às ligações dissulfeto intradomínio, a CDR3 dos VHHs comumente apresenta uma cisteína que estabelece uma ligação dissulfeto com uma outra cisteína funcionalizada na CDR1 ou FR2, a depender da espécie de camelídeos. A presença da ligação dissulfeto pode restringir a flexibilidade conformacional da alça CDR3, minimizando a penalidade entrópica ao ligar-se ao antígeno.

As alças CDR1 e CDR2 adotam um conjunto restrito de conformações que desviam substancialmente das conformações canônicas de alças estabelecidas com base nos anticorpos convencionais. Embora a sequência dentro das alças seja hipervariável, a variação de comprimento e diversidade conformacional é limitada exceto para a alça CDR3. Dessa maneira, a alça CDR3 fornece a principal contribuição para a diversidade e especificidade de ligação e, por conseguinte, têm uma maior influência na energia de ligação. Essas configurações enriquecem o repertório

estrutural dos VHHs, assim, para alcançar alta especificidade e afinidade essas moléculas empregam uma ampla variedade de representações geométricas na região de contato com o antígeno, que irão depender do tamanho e tipo antígeno [3]. Dessa forma, observações particulares sobre o comportamento dos VHHs frente a antígenos específicos não deixam claro quais características podem ser validadas de maneira geral.

Embora apresentem inúmeras vantagens, VHHs são moléculas complexas para se executar mutações pois a variação de um único resíduo pode modificar completamente as suas características estruturais e funcionais, resultando em uma perda crítica de estabilidade, além de baixo rendimento devido a agregação. À vista disso, a aplicação da engenharia racional no desenho de novos VHHs detentores de alta afinidade por alvos específicos requer o entendimento das características estruturais das alças CDR e de que maneira estas impactam a sua funcionalidade.

Trabalhos recentes dedicaram-se ao estudo da estrutura das alças CDR de VHHs por meio de análises baseadas em sequência e em poses estáticas (estruturas cristalográficas) [21,22,23]. Entretanto, proteínas são sistemas moleculares que estão constantemente alternando entre diversos estados conformacionais, o que torna a análise desse dinamismo em VHHs essencial à elucidação da relação entre as alças CDR e as suas funções. A simulação por dinâmica molecular (DM) é uma abordagem computacional amplamente aplicada e bem desenvolvida na obtenção de informações a nível atômico com respeito à variabilidade conformacional de proteínas ao longo do tempo. O estudo do dinamismo estrutural de VHHs através de simulações por DM é uma área ainda pouco explorada, mas que detém o potencial de complementar as informações obtidas a partir de dados cristalográficos.

No presente trabalho, descritores moleculares extraídos a partir de simulações por dinâmica molecular foram capazes de caracterizar as bases moleculares das alças CDR em VHHs. O delineamento do perfil dinâmico-estrutural da CDR permitiu a exploração da relação entre o comprimento de suas alças, estabilidade, flexibilidade e afinidade de ligação. Até onde sabemos, este é o primeiro trabalho a compilar dados de dinâmica molecular para um grande conjunto de estruturas de VHHs buscando compreender de que maneira o dinamismo das alças CDR afeta afinidade de ligação de VHHs.

2

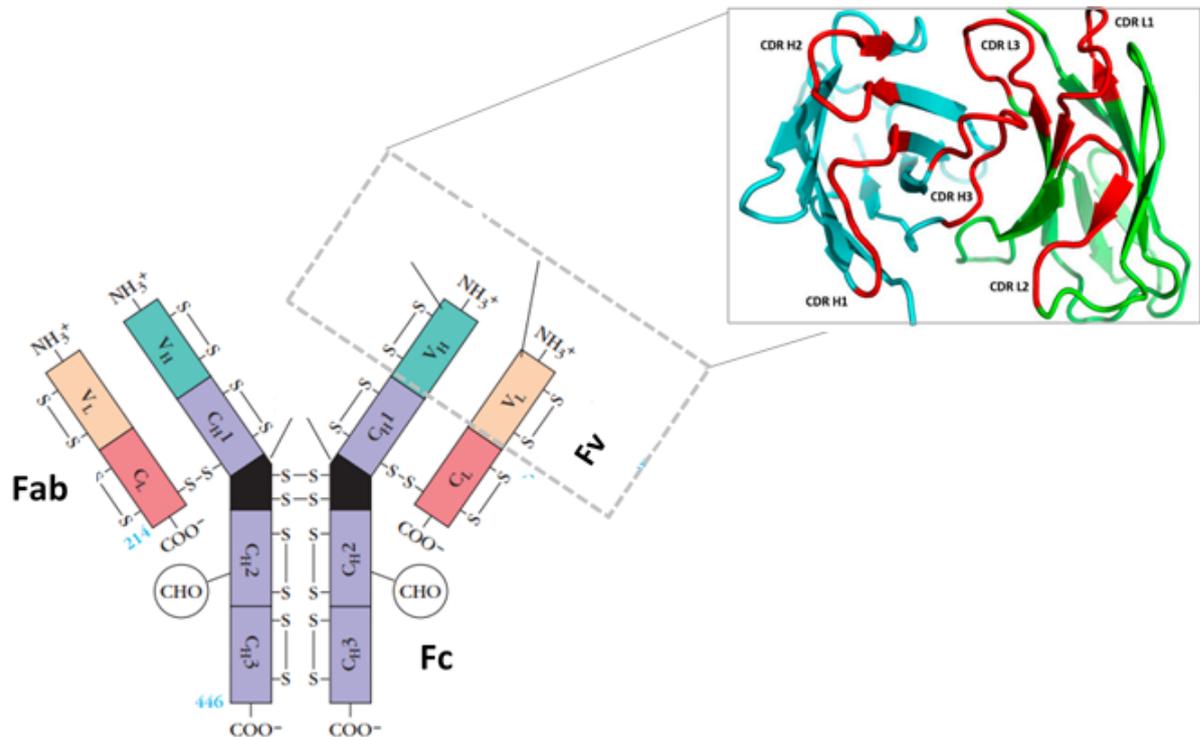
Fundamentação Teórica

2.1 Anticorpos

Através do processo de centrifugação o sangue pode ser separado em duas componentes: um fluido e uma fração celular. O fluido é denominado plasma e a fração celular contém as células vermelhas, leucócitos e plaquetas. Contido no plasma estão todas as moléculas pequenas e macromoléculas solúveis do sangue, incluindo as proteínas necessárias para a formação dos coágulos sanguíneos. Se o sangue ou o plasma coagularem, a fase fluida resultante é chamada soro [4]. A primeira evidência que os anticorpos residem no soro veio do clássico experimento de Tiselius e Kabat em 1939, no entanto o conhecimento a respeito estrutura tridimensional emergiu através estudos de cristalografia destas moléculas reportados no começo dos anos 70 [5, 4]. Os anticorpos secretados circulam no sangue onde agem como efetores da imunidade humoral identificando e neutralizando antígenos ou marcando-os para a eliminação [5].

Os anticorpos recebem a denominação de imunoglobulinas (IGs) pois são proteínas que conferem imunidade e possuem características de globulinas como a mobilidade eletroforetica lenta [1]. A arquitetura completa da molécula de anticorpo tem um peso molecular de aproximadamente 150 kDa [4] e é mostrada na [Figura 3](#). Estruturalmente o anticorpo pode ser dividido em três regiões de acordo com a sua funcionalidade: Duas regiões designadas fragmento de ligação ao antígeno (Fab, do inglês, *Fragment Antigen-Binding*) e o fragmento cristalino (Fc, do inglês, *Fragment crystalline*) [1, 6].

Figura 3 – . Estrutura da IgG com destaque para a região VH-VL e as suas 6 alças CDR.



Fonte: Adaptado de WOOF (2004) [7] e KINDT (2007) [8]

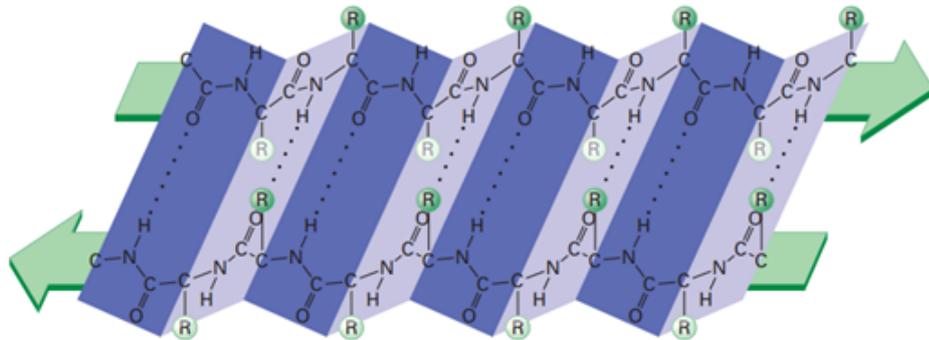
Cada Fab é idêntico e contém os domínios da cadeia leve VL e CL conectados aos domínios da cadeia pesada VH e CH1. O Fab retém a porção necessária para o reconhecimento ao antígeno (VH-VL), o fragmento variável (Fv). A região Fc define o perfil da função efetora que dita como o anticorpo irá interagir com outro componente dos sistemas imunes adaptativo e humoral. Entre as regiões Fc e Fab existe um porção flexível chamada de dobradiça. A dobradiça permite que as duas regiões Fab se movam independente uma da outra, permitindo-lhes ligar simultaneamente a epítopos de um antígeno que estão separados por distâncias variáveis [1, 5].

Dentre as interações que mantém as cadeias leves conectadas às cadeias pesadas estão: ligações dissulfeto, interações não-covalentes como interações de hidrogênio, interações hidrofóbicas e interações iônicas [4]. O dímero VH-VL possui 6 alças CDR das quais 3 localizam-se no domínio leve variável CDR L1, CDR L2 E CDR L3 e as outras 3 no domínio pesado variável CDR H1, CDR H2 e CDR H3. As regiões *framework* são estruturadas em formato de fitas antiparalelas distribuídas em folhas- e configuram-se, basicamente, como um núcleo que suporta as CDRs em sua superfície [4, 9].

As fitas dentro de uma folha são estabilizadas por ligações de hidrogênio que conectam os grupos $-NH$ em uma fita aos grupos carbonila de uma fita adjacente [Figura 4](#). As fitas são caracterizadas por aminoácidos hidrofóbicos e hidrofílicos alternados cujas cadeias laterais estão dispostas perpendicularmente ao plano da folha; os aminoácidos hidrofóbicos são orientados para o interior do sanduíche e os aminoácidos hidrofílicos voltados para fora. Esta estrutura

característica e compacta em que os domínios dos anticorpos enovelam-se é denominada de “enovelamento da imunoglobulina” [4].

Figura 4 – . Fórmula estrutural de uma folha- composta por duas fitas- β



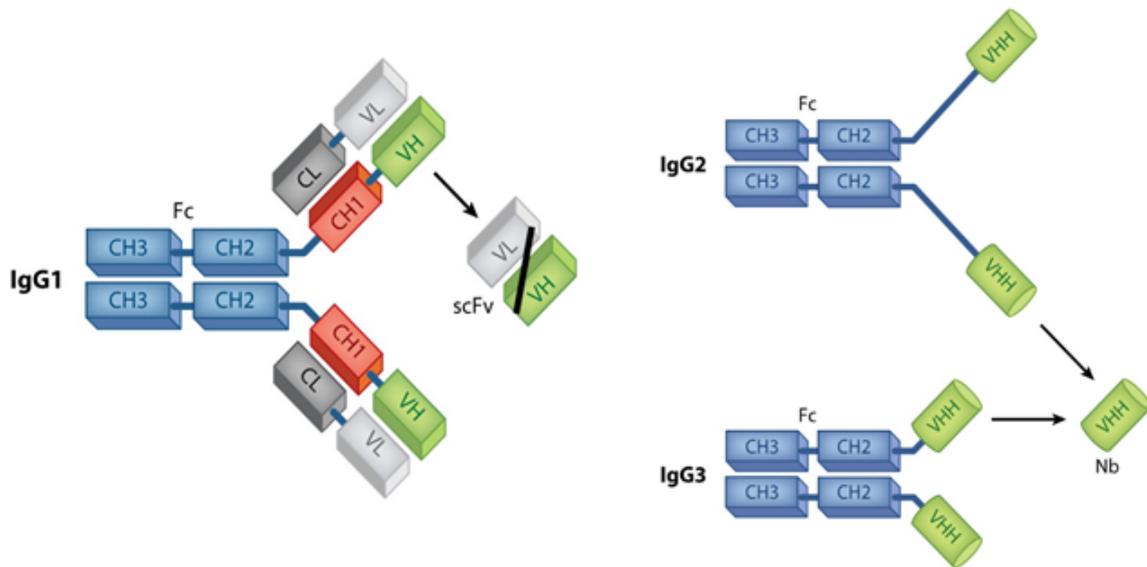
Fonte: KINDT (2007) [8]

2.2 VHH

Três frações contendo anticorpos do tipo IgG com pesos moleculares distintos podem ser isoladas do soro de dromedário através da adsorção diferencial à colunas de Proteína A e Proteína G na cromatografia por afinidade: IgG, IgG2 e IgG3. A fração IgG1 contém os anticorpos convencionais que compreendem dois domínios de cadeia leve e quatro domínios de cadeia pesada. As frações "IgG2" e "IgG3" diferenciam-se da IgG1 em razão de serem constituídas apenas por domínios formados por cadeias pesadas (HCAs) [Figura 5](#). As cadeias pesadas nesses anticorpos de cadeia pesada única são aproximadamente 10 a 12 kDa menores em comparação as cadeias pesadas de anticorpos convencionais e essa massa molecular reduzida é atribuída a ausência da cadeia leve e do primeiro domínio pesado conservado CH1, que funciona como uma âncora para a cadeia leve em anticorpos convencionais [2, 7].

A porcentagem de HCAs e IgG1 convencional nos soros de camelídeos é variável, nos camelos pode variar entre 50 e 80 % contrastando com as espécies sul-americanas de camelídeos em que essa quantidade varia entre 10 e 25%. Tendo em vista estas proporções significativas fica clara a importância dos HCAs no sistema imune dos camelídeos [2]. As IgG2 e IgG3 são tipos de anticorpos homodiméricos e funcionalmente dividem-se em duas regiões: região de reconhecimento ao antígeno, VHH, e fragmento cristalino (CH2 e CH3). A diferença entre estes anticorpos está na dobradiça da fração IgG2 que possui maior comprimento em comparação a do tipo IgG3. Na IgG1, O menor fragmento de ligação ao antígeno funcional intacto que pode ser gerado a consiste no par VH-VL ligado por um oligopeptídeo, e recebe a nomenclatura de scFv (do inglês, *Single Chain Fragment Variable*).

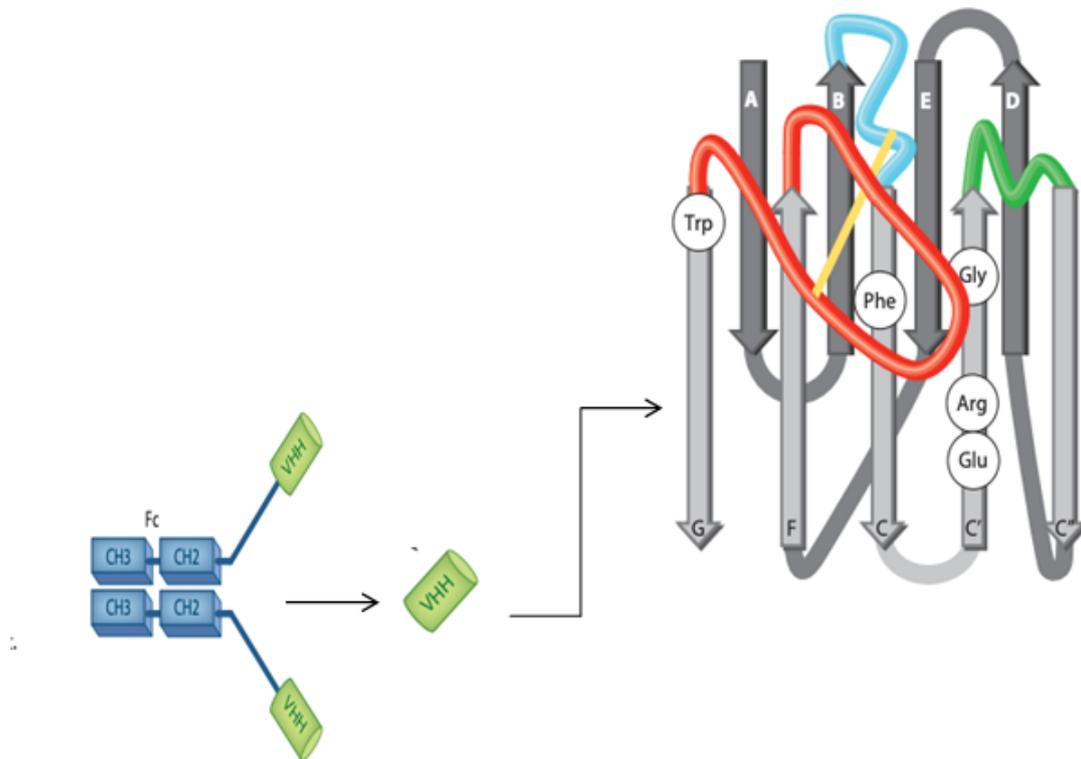
Figura 5 – . Esquema estrutural da IgG1, IgG2 e IgG3.



Fonte: MUYLDERMANS (2013) [2]

O menor fragmento intacto e funcional, equivalente ao scFV, nos anticorpos de cadeia pesada única é domínio pesado variável é o VHH. Os VHHs também adotam o enovelamento da imunoglobulina: 3 regiões de alças hipervariáveis, CDR1-3, rodeadas por regiões *Framework* organizadas em duas folhas- α a compostas por 4 e 5 fitas- β antiparalelas conectadas por meio de alças e uma ligação dissulfeto conservada entre Cys 23 e Cys 94 [Figura 6](#). Nesta arquitetura as regiões CDR1, CDR2 e CDR3 conectam as fitas B-C, C' -C'' e F-G, respectivamente [2, 10, 11]. Em suas sequências os VHHs apresentam a substituição de cinco aminoácidos que são conservados no domínio VH dos anticorpos convencionais em vertebrados [Figura 6](#). Essas substituições são uma marca da diferença VH-VHH e se dão da seguinte maneira: Leu12Ser, Val42Phe/Tyr, Gly49Glu, Leu50Arg/Cys, e Trp52Gly (sendo esta menos conservada). Em anticorpos convencionais os resíduos nas posições 42, 49, 50 e 52 do domínio VH estão localizados na superfície da interface com o domínio VL. Dado o caráter hidrofóbico da superfície do domínio VH, a substituição de aminoácidos apolares por aminoácidos polares observada nos VHHs explica tanto a falha para associar-se com o domínio VL quanto a alta solubilidade dessas moléculas [2, 10, 12].

Figura 6 – . Esquema estrutural de um VHH com destaque para seu enovelamento da imunoglobulina.



Fonte: MUYLDERMANS (2013) [2]

As substituições de aminoácidos apolares por polares aumentam a hidrofili­cidade da superfície do VHH. Este efeito é ainda melhorado pela rotação das cadeias laterais hidrofóbicas de resíduos adjacentes para expor as partes mais hidrofílicas ao solvente, sem deformar o C do *backbone*. Além disso, as substituições nas posições 42 e 52 causam um deslocamento de uma rede de grupos hidrofóbicos volumosos em direção ao centro das 5 fitas- β . A alça CDR3 geralmente dobra-se sobre esses resíduos e os tornam inacessíveis ao solvente.

Além das arquiteturas padrões apresentadas pelo seu parátipo ¹ como grooves, cavidades e superfícies planas, VHHs são capazes de arquitetar a sua longa alça CDR3 em um formato convexo comparado com a geometria de “uma bola de rugby”. Por exemplo, em um VHH que liga-se à lisozima parte do longo CDR3 se dobra sobre a região que seria ocupada pelo domínio VL em um anticorpo convencional, enquanto a outra parte remanescente do parátipo penetra o interior do ativo enzimático. Os VHHs também são capazes de formar uma cavidade utilizando os seus três CDRs para acomodar a ligação com um hapteno ². Todo este aparato molecular permite que os VHHs reconheçam a epítomos ³ geralmente não antigênicos para anticorpos convencionais como o sítio catalítico de enzimas e cânions em biomarcadores de doenças virais e infecciosas .

¹ Sítio que compreende a região do anticorpo que está em contato com o antígeno.

² Molécula não-proteica de baixo peso molecular.

³ Região do antígeno que se liga ao anticorpo.

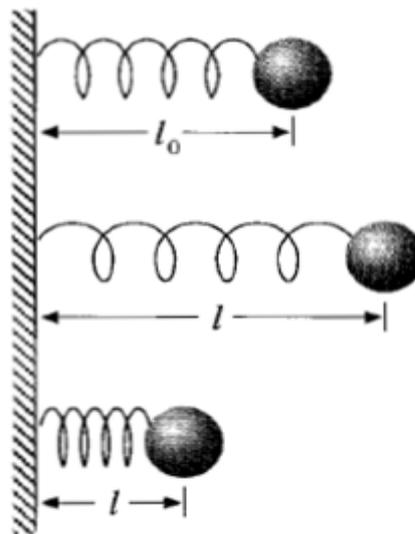
Portanto, se teoriza que HCAs tenham sido selecionados e mantidos nas espécies de camelídeos para exercer uma função complementar em sua resposta imune humoral [2, 10, 12].

2.3 Mecânica Molecular

A Mecânica Molecular (MM) é um método teórico que descreve sistemas moleculares através das leis da Mecânica Clássica [13]. Os modelos baseados na MM para a representação de uma molécula consistem em uma aproximação ao oscilador harmônico onde conjunto de esferas carregadas (átomos) são conectadas por molas (ligações) [13, 14, 15, 16]. O uso das coordenadas nucleares para representar átomos é justificado em termos da aproximação de Born-Oppenheimer, na qual assume-se que o movimento do núcleo pode ser desacoplado do movimento dos elétrons [17, 18, 19]. Notavelmente, a mecânica clássica não se aplica ao estudo do comportamento dos elétrons e estes não são considerados explicitamente [16, 20, 21]. À vista disto, uma vez que a movimentação do núcleo é livre e as interações nucleares possuem natureza aditiva a densidade eletrônica tende a ajustar-se a qualquer mudança geométrica dos núcleos [20].

No oscilador harmônico simples [Figura 7](#) se uma esfera de massa m , conectada a uma mola, sofre uma perturbação do seu estado de equilíbrio, estará sujeita a ação de uma força proporcional ao deslocamento [15, 22].

Figura 7 – Representação do oscilador harmônico através da modelo massa mola.



Fonte: MCQUARRIE [15]

Esta força é dada pela [Equação 2.1](#), denominada lei de Hooke .

$$f = -k(l - l_0) = -kx \quad (2.1)$$

Na [Equação 2.1](#) o deslocamento é definido por $x = (l - l_0)$ sendo l o comprimento da mola distorcida, l_0 o comprimento de equilíbrio e k a sua constante de força. O sinal negativo indica que a força direciona-se para a direita quando a mola é comprimida e para a esquerda quando a mola é estirada $l > l_0$ [15]. Este modelo foi originalmente aplicado no entendimento do espectro vibracional, à medida que dados espectroscópicos tornavam-se disponíveis no início do século 20 [16, 18]. Particularmente no âmbito da química orgânica, inúmeras moléculas que possuíam mesmo padrão de ligações pudera ser caracterizadas evidenciando que o comprimento de ligação de equilíbrio e a constante de força, o mesmo termo utilizado na mecânica molecular, eram estendidos de uma molécula para a outra [18]. A este fenômeno é concedido o nome de transferibilidade, um dos principais pilares para a eficácia global da mecânica molecular [14, 16, 18]

A energia potencial $V(l)$ associada à deformação da geometria de equilíbrio de uma molécula $l < l_0$ pode ser escrita como uma expansão em séries de Taylor apresentada na [Equação 2.2](#) [18, 20]:

$$\begin{aligned}
 V(l) = V(l_0) + \left(\frac{dV}{dl} \right)_{l=l_0} (l - l_0) + \\
 \frac{1}{2!} \left(\frac{d^2V}{dl^2} \right)_{l=l_0} (l - l_0)^2 + \\
 \frac{1}{3!} \left(\frac{d^3V}{dl^3} \right)_{l=l_0} (l - l_0)^3 + \dots
 \end{aligned}
 \tag{2.2}$$

Na [Equação 2.2](#) por conveniência, o primeiro termo $V(l_0)$ é designado como o zero de energia e é, portanto, igualado a zero. O segundo termo $\left(\frac{dV}{dl} \right)_{l=l_0}$ desaparece em razão do ponto $l - l_0$ ser correspondente ao mínimo da curva de energia potencial. Essencialmente, esta atribuição reflete que a força agindo entre os núcleos é zero neste ponto. Truncando a série após o termo quadrático, o primeiro diferente de zero, têm-se a expressão mais simplória para a energia potencial vibracional como mostra a [Equação 2.3](#) [15, 16, 20].

$$V(l) = \frac{1}{2} k (l - l_0)^2
 \tag{2.3}$$

Sendo $\left(\frac{d^2V}{dl^2} \right)_{l=l_0}$ denotado por k , a constante de força. Naturalmente uma expansão de Taylor truncada terá o seu melhor funcionamento próximo ao ponto de referência, nesse caso l_0 [18]. Para desvios severos do equilíbrio este tratamento não seria válido e termos superiores na equação 3 teriam que ser considerados, os denominados efeitos anarmônicos [15, 20]. De forma funcional a MM envolve a construção de energias potenciais, uma função de posições atômicas, partindo de um vasto corpo de dados moleculares (estruturas cristalográficas, espectroscopia vibracional, etc). Frente às forças intramoleculares e intermoleculares as moléculas estiram-se,

dobram e rotacionam ao redor das ligações [14]. De acordo com o princípio da aditividade a energia total ou estérica destas moléculas é expressa pela soma dos potenciais derivados destas forças. O somatório destas funções de energia potencial é denominado campo de força [14, 23, 24].

2.3.1 Campo de Força

A maioria dos campos de força de modelagem molecular, em uso atualmente para sistemas moleculares, pode ser interpretado como a soma das contribuições dos seguintes tipos de termos: comprimento de ligação, deformação do ângulo de ligação, torsão, interações eletrostáticas e de Van der Waals [14, 24]. A Equação 2.4 traz a representação matemática das contribuições supracitadas.

$$\begin{aligned}
 V(r^N) = & \sum_{\text{Ligação}} \frac{k_i}{2} (l_i - l_{i,0})^2 + \sum_{\text{Ângulos}} \frac{k_i}{2} (\theta_i - \theta_{i,0})^2 + \\
 & \sum_{\text{Torsões}} \frac{V_n}{2} (1 + \cos(n\omega - \gamma)) + \\
 & \sum_{i=1}^N \sum_{j=i+1}^N \left(4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} + \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right)
 \end{aligned} \tag{2.4}$$

Na Equação 2.4 $V(r^N)$ denota a energia potencial total, que é uma função das posições (l) dos N átomos. O primeiro termo modela a interação entre os pares de átomos ligados por meio do potencial harmônico, dessa forma, a energia eleva-se quando a ligação é afastada do comprimento de referência $l_{i,0}$. O segundo termo corresponde a soma de todos os ângulos de valência da molécula e é, também, modelado pelo potencial harmônico. O terceiro potencial é o torcional que modela o quanto a energia modifica quando uma ligação rotacional em torno de uma ligação simples. A quarta contribuição abrange as interações entre pares (i, j) de átomos não ligados modeladas pelo potencial de Coulomb e potencial de Lennard-Jones, correspondentes às interações eletrostáticas e de Van der Waals, respectivamente. Campos de força mais sofisticados podem dispor de outros termos no entanto, invariavelmente, irão conter os componentes da Equação 2.4 [24].

A utilização dos métodos de campo de força na otimização de geometrias traz consigo vantagens consideráveis como velocidade de execução e precisão satisfatória, podendo igualar-se ou, até mesmo, superar a confiabilidade de resultados obtidos experimentalmente. Logicamente, para alcançar tais vantagens é necessário que o campo de força seja implementado às mesmas classes de moléculas que foram aplicadas em sua parametrização [30]. Ao longo do tempo, diversos campos de força para simulação de sistemas biomoleculares foram desenvolvidos e apesar de possuírem formas semelhantes de função de interação diferem, substancialmente, em sua filosofia de parametrização e valores dos parâmetros [24].

Para as simulações desenvolvidas no presente estudo, fora utilizado o campo de força GROMOS 53A6 (Groningen Molecular Simulation - GROMOS) [25]. O pacote de simulação computacional GROMOS tem sido desenvolvido em associação às funções de interação interatômica para simulações de dinâmica molecular de sistemas biomoleculares. A filosofia básica implícita no campo de força GROMOS propõe alcançar o equilíbrio entre uma descrição precisa da energia de interação como uma função da conformação, porém, mantendo-se uma forma funcional relativamente simples. Partindo de uma forma funcional simples é possível minimizar o número de parâmetros necessários, facilitando a transferência de parâmetros entre unidades similares, e limitando o custo computacional para avaliar a energia potencial.

O GROMOS 53A6 inclui em sua parametrização as entalpias livres de hidratação e solvatação para uma diversidade de moléculas apolares e, além disso, possui parâmetros aprimorados para reproduzir propriedades termodinâmicas de líquidos puros de uma gama de pequenas moléculas polares. Vale ressaltar, assim, que as cargas parciais foram então ajustadas para reproduzir as entalpias livres de hidratação. As considerações dos efeitos de solvatação são imprescindíveis pois desempenham papel fundamental na maioria dos processos biológicos de interesse, como o enovelamento de proteínas, reconhecimento e associação de biomoléculas, entre outros.

A forma funcional do campo de força GROMOS 53A6 é construída pelos seguintes termos:

$$V_{\text{Ligação}}(r; s) = V_{\text{Ligação}}(r; K_b, b_0) = \sum_{n=1}^N \frac{1}{4} K_b [b_n^2 - b_{0n}^2]^2 \quad (2.5)$$

$$V_{\text{Ângulos}}(r; s) = V_{\text{Ângulos}}(r; K_{\Theta}, \Theta_0) = \sum_{n=1}^{N_{\Theta}} \frac{1}{2} K_{\Theta_n} [\cos \Theta_n - \cos \Theta_{0n}]^2 \quad (2.6)$$

$$V_{\text{Impróprios}}(r; s) = V_{\text{Impróprios}}(r; K_{\xi}, \xi_0) = \sum_{n=1}^{N_{\xi}} \frac{1}{2} K_{\xi_n} [\xi_n - \xi_{0n}]^2 \quad (2.7)$$

$$V_{\text{Diedros}}(r; s) = V_{\text{Diedros}}(r; K_{\phi}, \delta, m) = \sum_{n=1}^{N_{\phi}} K_{\phi_n} [1 + \cos(\phi_n) \cos(m_n \phi_n)] \quad (2.8)$$

$$V_{\text{van der Waals}}(r; s) = V_{\text{van der Waals}}(r; C12, C6) = \sum_{\text{pares } i,j} \left(\frac{C12_{ij}}{r_{ij}^{12}} - \frac{C6_{ij}}{r_{ij}^6} \right) \quad (2.9)$$

$$V^C(r; s) = V^C(r; q) = \sum_{\text{pares } i,j} \frac{q_i q_j}{4\pi\epsilon_0\epsilon_1} - \frac{1}{r_{ij}} \quad (2.10)$$

$$V^{CR}(r; s) = V^{CR}(r; q) = \sum_{\text{pares } i,j} \frac{q_i q_j}{4\pi\epsilon_0\epsilon_1} \frac{-\frac{1}{2}C_{cr}r_{cr}^2}{R_{cr}^3} \quad (2.11)$$

$$V^{Rcr}(s) = V^{CR}(q) = \sum_{\text{pares } i,j} \frac{q_i q_j}{4\pi\epsilon_0\epsilon_1} \frac{-(1 - \frac{1}{2}C_{cr})}{R_{cr}} \quad (2.12)$$

As primeiras equações (2.5 a 2.8) descrevem os quatro tipos de interações por ligações covalente, mencionadas anteriormente, onde K equivale às constantes de força, b_n é o comprimento de ligação, b_{0n} é o comprimento de ligação ideal, Θ_n é o ângulo de ligação, Θ_{0n} é o ângulo de ligação ideal, ϕ_n é o ângulo diedro, ϕ_{0n} é o ângulo-diedro de deslocamento de fase, ϵ_n é o ângulo diedro impróprio, ϵ_{0n} é o ângulo diedro impróprio de referência. A [Equação 2.6](#) especifica as interações não-ligadas de Van der Waals, onde r_{ij} é a distância entre os átomos i e j , C_{12ij} e C_{6ij} são os parâmetros de Lennard-Jones.

As interações eletrostáticas entre pares que interagem são descritas através de três contribuições. A [Equação 2.10](#) apresenta-se como a soma sobre todos os pares que interagem por meio da interação de Coulomb V^C , os parâmetros q_i são definidos como as cargas parciais q_j sobre os átomos, ϵ_0 como a permissividade dielétrica do vácuo e ϵ_1 equivale à permissividade relativa do meio em que os átomos estão incluídos, cujo valor é padronizado como 1. Em adição, a [Equação 2.11](#) refere-se a uma contribuição do campo de reação V^{CR} às interações eletrostáticas e é calculada por meio da representação da interação de um átomo i com o campo induzido de um meio contínuo dielétrico fora de uma distância de um raio de corte, R_{cr} , devido a presença do átomo j .

Por fim, a última contribuição é exposta, comumente, como um termo de campo de reação independente da distância V^{Rcr} . Esta é uma contribuição constante, ou seja, independe da configuração, para a energia de cada par que é levado em conta. Por conta da independência das distâncias interatômicas r_{ij} não há contribuição para as forças, o que garante que a energia eletrostática é zero para átomos que estão na distância do raio de corte e, assim, reduz ruídos de raio de corte na energia. Além disso, o somatório neste termo é capaz de excluir pares, como o autotermo *i.i.*, proporcionando que a contribuição total para o $V^{Rcr}(q)$ seja zero para sistemas neutros.

2.4 Dinâmica Molecular

A dinâmica molecular clássica (DM) como abordagem computacional baseia-se na simulação do movimento de um sistema atômico, sob a influência de um campo de força

específico, ao longo do tempo de acordo com a equação de movimento Newtoniana [14]. Como resultado deste processo configurações sucessivas do sistema são geradas, construindo uma trajetória, que especifica como as posições e velocidades das partículas do sistema variam com o tempo [17, 24]. Em contrapartida aos métodos experimentais, as simulações de DM fornecem uma estimativa da dinâmica estrutural e propriedades de equilíbrio e de sistemas complexos, as quais não podem ser obtidas analiticamente. Desta forma, a DM estabelece-se como uma interface entre teoria e experimento, ocupando uma posição de destaque por estar na interseção entre a matemática biologia, química, física e ciência da computação [14].

Para um átomo, i , com massa m_i , e posição definida pelo vetor tridimensional r_i , o relacionamento entre a velocidade e momento, p_i , é dado por [19]:

$$\frac{dr_i}{dt} = \frac{dp_i}{m_i} \quad (2.13)$$

Implementado o campo força, as forças atuantes sobre cada átomo do sistema são calculadas pela derivada negativa da energia potencial, obtida do campo de força escolhido, com relação a posição desses átomos, como a mostra a equação:

$$F_i(t) = -\frac{dV}{dr_i} \quad (2.14)$$

A equação de movimento Newtoniana para o átomo i é dada por:

$$\frac{dp_i}{dt} = F_i \quad (2.15)$$

Considerando a posição com respeito a uma única componente do vetor r_i , a componente x , em um tempo específico, t , a mudança de posição após um curto e finito intervalo de tempo, Δt , é dada pela seguinte uma série de Taylor:

$$x(t + \Delta t) = x(t) + \frac{dx(t)}{dt} \Delta t + \frac{d^2x(t)}{dt^2} \frac{\Delta t^2}{2} + \dots \quad (2.16)$$

A posição $x(t)$, a velocidade $\frac{dx(t)}{dt}$ e a aceleração $\frac{d^2x(t)}{dt^2}$ são suficientes para uma solução numérica da equação de movimento. Para uma única dimensão a aceleração é descrita pela segunda lei de Newton:

$$\frac{d^2x(t)}{dt^2} = \frac{F_x}{m} \quad (2.17)$$

Onde, F_x é a força que age sobre o átomo na mesma direção de x . Sob a influência de um potencial contínuo as movimentações de todas as partículas estão acopladas dando origem a um sistema de muitos corpos. Dentro desta circunstância, há a necessidade de submeter as equações de movimento à integração utilizando algoritmos baseados no método das diferenças finitas [24].

2.4.1 Método das diferenças finitas

A essência dos algoritmos baseados no método das diferenças finitas, está em particionar a integração em pequenos passos, separadas no tempo por um intervalo fixo, Δt . A força total em cada partícula, para uma configuração, no tempo t , é calculada como o vetor soma de suas interações com as demais partículas. Partindo da força é possível determinar as acelerações das partículas que, por sua vez, são combinadas às posições e velocidades no tempo t para calcular as posições e velocidade no tempo $t + \Delta t$. Durante o intervalo de tempo assume-se que a força é constante, dessa forma, as forças sobre as partículas em novas posições são determinadas, levando à novas posições e velocidades no tempo $t + 2\Delta t$, e assim por diante [24].

Existe uma variedade de algoritmos baseados nesta metodologia, funcionalizados na integração das equações de movimento, no entanto, dentro do âmbito da dinâmica molecular o algoritmo de Verlet é o mais utilizado pois requer uma quantidade pequena de memória computacional e tempo de CPU [26, 27]. O algoritmo de Verlet faz uso das posições atômicas e acelerações no tempo t e posições do passo anterior $x(t - \Delta t)$, para determinar as novas posições em $t + \Delta t$ [19, 24, 27]:

$$x(t + \Delta t) = 2x(t) - x(t - \Delta t) + \frac{d^2x(t)}{dt^2} \Delta t^2 \quad (2.18)$$

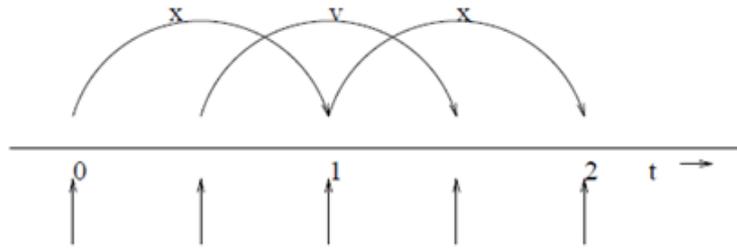
O algoritmo leap-frog é construído através de uma pequena modificação no algoritmo de Verlet. A implementação do leap-frog consiste no uso das posições r no tempo t e velocidade no instante $t - \left(\frac{\Delta t}{2}\right)$, a atualização das velocidades v e posições r são executadas através do cálculo das forças $F(t)$ determinadas pelas posições no tempo t , como mostra as equações [26]:

$$v\left(t + \frac{1}{2}\Delta t\right) = v\left(t - \frac{1}{2}\Delta t\right) + \frac{\Delta t}{m} \mathbf{F}(t) \quad (2.19)$$

$$r(t + \Delta t) = r(t) + \Delta t v\left(t + \frac{1}{2}\Delta t\right) \quad (2.20)$$

O Algoritmo é denominado leap-frog porque r e v estão “saltando” um após o outro (**Figura 8**).

Figura 8 – - Ilustração do método de integração leap-frog.

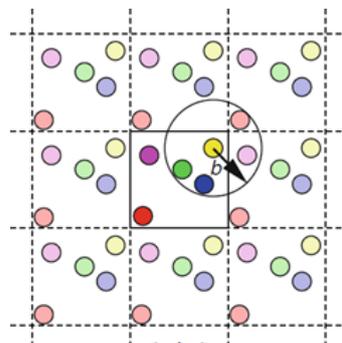


2.4.2 Condições Periódicas de Contorno

As simulações de DM são capazes de mimetizar diferentes condições experimentais [19]. Os sistemas submetidos a estas simulações possuem em geral entre 10^2 e 10^6 partículas, no entanto, este número torna-se insignificante quando comparado à quantidade de átomos contidos nos sistemas macroscópicos, aproximadamente 10^{23} partículas. Como consequência do tamanho reduzido, o número relativo de átomos na superfície do sistema é muito maior do que para um sistema macroscópico, causando efeitos de superfície que podem dominar as propriedades do sistema de forma indesejável [18, 20]. Uma maneira clássica de minimizar efeitos de superfície em sistemas finitos é através da aplicação das condições periódicas de contorno (PBC, *Periodic Boundary Conditions*) [28]. Sob a PBC o sistema modelado (biomoléculas e moléculas do solvente) é construído dentro de uma caixa de simulação cúbica considerada como uma célula unitária que é replicada em todas as direções de forma periódica [14, 18].

Na Figura 9 a caixa é mostrada em uma representação bidimensional, neste exemplo cada caixa está rodeada por 8 vizinhos. As partículas experimentam ação das forças como se estivessem no interior do sistema, dessa forma, o sistema modelado é colocado na caixa e esta é replicada em todas as dimensões [24].

Figura 9 – - Ilustração do método de integração leap-frog.



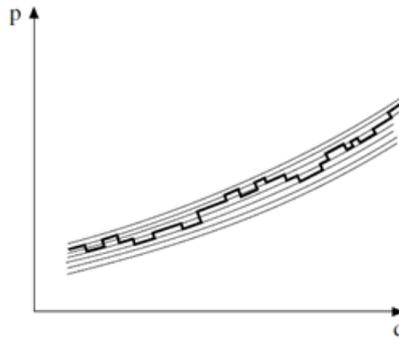
SCHLICK (2010) [14]

A convenção mínima de imagem garante que se um átomo da caixa central cruza o limite da célula unitária, um átomo idêntico entrará pela face oposta e o substituirá garantindo que interações duplicadas não sejam levadas em consideração e mantendo o número de moléculas constante [13, 19]. Um outro aspecto chave da definição da PBC é que nenhuma molécula dentro da célula de simulação vê "vácuo" dentro da faixa de seus pontos de corte de interação e, portanto, os artefatos de superfície são evitados [18].

2.4.3 Mecânica Estatística: Espaço de fase e Ensemble

O espaço de Fase é um conceito imprescindível no âmbito das simulações computacionais [12]. Para um sistema contendo N átomos há uma demanda de $6N$ coordenadas, 3 coordenadas relativas à posição atômica q_1, \dots, q_{3N} e 3 componentes do momento p_1, \dots, p_{3N} , para que o estado do sistema possa ser definido. Cada ponto no espaço de fase, construído através das combinações entre as $3N$ posições e $3N$ momento, representa um microestado e, a coleção de todos os microestados é denominada ensemble. Assim, todos os microestados possíveis contidos no ensemble representam um mesmo macroestado, o qual é caracterizado pelas variáveis macroscópicas (E, V, P, \dots) [21, 24, 29, 30].

Figura 10 – - Ilustração do método de integração leap-frog.



SCHWAB (2006) [30]

Um ponto individual no espaço de fase, corresponde à uma geometria em particular de todas as moléculas de um sistema. Existem inúmeros pontos no espaço que, de fato, nunca ocorreriam em um sistema real. Dessa forma, para descrever de o sistema forma realista é necessário determinar a probabilidade de ocorrência de determinada configuração [29]. Os microestados são distribuídos de acordo com a função de probabilidade de Boltzmann [27]:

$$P(q, p) = Q^{-1} e^{-E(q,p)/k_B T} \quad (2.21)$$

A dinâmica molecular gera uma sequência de pontos no espaço de fase em função do tempo (trajetória) [24]. Obtendo a trajetória é possível calcular a média temporal de uma

propriedade termodinâmica ao tirar a média dessa propriedade com respeito ao tempo ao longo da trajetória [28]. Esse método baseia-se no princípio da ergodicidade que assume que a média de tempo para um sistema isolado é equivalente a média sobre os microestados que compõe o ensemble [21]. Para um número finito de pontos M , é possível calcular o valor médio de uma dada propriedade X de acordo com a seguinte equação:

$$\langle X \rangle_M = \frac{1}{M} \sum_{i=1}^M X(q_i, p_i) \quad (2.22)$$

3

Objetivos

3.1 Objetivo Geral

Caracterizar o comportamento da dinâmica estrutural de VHHs para a elucidação do comportamento das bases moleculares responsáveis pela interação VHH-antígeno, assim bem como, fornecer informações que poderão contribuir para a otimização do desenho racional de novos VHHs.

3.1.1 Objetivos Específicos

- Construir uma biblioteca de estruturas tridimensionais de VHHs;
- Realizar simulações de dinâmica molecular para 40 estruturas de VHHs em sua forma livre e em solução aquosa;
- Caracterizar a amplitude da diversidade conformacional e a flexibilidade das regiões CDR1-3 por meio das métricas RMSD e RMSF;
- Construir gráficos de de Ramachandran para as regiões CDR1-3;
- Avaliar o comportamento da flexibilidade das regiões CDR1-3 em função da energia livre de ligação experimental (ΔG_{bind});
- Construir um gráfico de conservação de sequência para as 40 estruturas;
- Calcular a área da superfície acessível ao solvente (SASA) por resíduo de aminoácido que compõem a CDR1-3;

4

Materiais e Métodos

4.1 Construção da biblioteca de estruturas de VHH e preparação das estruturas

O conjunto de estruturas tridimensionais de VHH foram obtidas através do banco de proteínas PDB (do inglês, Protein Data Bank) (<https://www.rcsb.org/>), utilizando os seguintes filtros: A palavra-chave nanobody e o limite de resolução de 3 angstroms (Å).

A busca resultou em 130 estruturas tridimensionais, contudo foram rejeitadas aquelas estruturas cujos resíduos faltosos excederam o total de 5 aminoácidos em sequência a fim de preservar confiabilidade das coordenadas atômicas. O conjunto de dados final resultou em 40 estruturas tridimensionais de VHHs. Isto decorre pois as 40 estruturas tiveram a sua estrutura tridimensional determinada por meio de difração de raios X e desde que essa técnica recai na obtenção de cristais com várias proteínas em posições idênticas, regiões muito flexíveis geram problemas porque geralmente não são observadas e então as suas coordenadas não são incluídas no arquivo PDB.

Àqueles VHHs complexados a antígeno foram desacoplados e utilizados em sua forma livre. As estruturas que apresentaram ausência de resíduos (>5) tiveram as suas coordenadas atômicas modeladas através do pacote de modelagem molecular MODELLER 9.20 (<https://salilab.org/modeller/>) utilizando o protocolo loopmodel DOPE-based [31]. Os algoritmos do MODELLER são baseados na linguagem de programação Python e os scripts aplicados na modelagem das nossas estruturas estão disponíveis no Anexo I.

Em seguida, todas as estruturas foram submetidas a uma minimização utilizando o algoritmo MinMover implementado pacote computacional Rosetta versão 3.9 (<https://www.rosettacommons.org/>) selecionando como o limite um RMSD de 1.5 Å em relação à estrutura inicial [32].

4.2 Dinâmica Molecular

Todas as simulações moleculares das estruturas completas dos VHHs em sua forma livre foram realizadas através do pacote de simulação GROMACS v. 4.6.7. utilizando o o campo de força GROMOS e o conjunto d parâmetros 53A6 [26, 25]. As proteínas foram inseridas em uma caixa cúbica de dimensões 12x12x12 nm e explicitamente solvatadas com o modelo de água SPC (*Single point-charge*) sob condições periódicas de contorno nas direções x, y e z [33]. Contra-íons foram adicionados ao sistema para garantir a neutralidade de cargas e uma concentração fisiológica de 0.15 mM. Uma minimização de geometria de 5.000 ciclos foi realizada utilizando o algoritmo steepest descent. As interações eletrostáticas de longo alcance foram levadas em consideração através do campo de reação com $\epsilon=66$ [34]. Um raio de corte de 1,4 nm foi usado para interações de curto alcance . O algoritmo LINCS foi aplicado para restringir todos os comprimentos de ligações covalentes [??]. A produção de dinâmica molecular foi de 200 ns utilizando um ensemble isotérmico-isobárico (NPT). As temperaturas dos solutos e dos solventes foram acopladas separadamente e a temperatura foi mantida a 300K utilizando o termostato v-rescale com um tempo de acoplamento de 0,4 fs [35]. A pressão foi mantida a 1 bar através do barostato Berendsen com um tempo de relaxamento de 0,4 ps e compressibilidade de $4,5 \times 10^{-5} \text{ bar}^{-1}$ [36]. O arquivo de entrada utilizado na produção da dinâmica molecular está disposto no Anexo I.

4.2.1 Cálculo do RMSD

O desvio quadrático médio (do inglês, *Root Mean Square Deviation* - RMSD) dos átomos de uma molécula com respeito a uma estrutura de referência, aqui a estrutura cristalográfica, foi calculado por meio do programa *g_rms* implementado no pacote computacional GROMACS v 4.6.7 [26, 37]. As diferentes conformações obtidas durante a simulação de DM são alinhadas uma a uma à estrutura de referência e o RMSD é computado como mostra a [Equação 4.1](#):

$$\text{RMSD} = \sqrt{\frac{\sum_{i=1}^N (r_{i,j} - r_i)^2}{N}} \quad (4.1)$$

Onde N é o número de átomos da macromolécula, $r_{(i,j)}$ é a posição do átomo i na j-ésima conformação e r_i é a posição do átomo i na estrutura de referência. Maiores valores de RMSD indicam que determinada estrutura se afasta da conformação da estrutura de referência (cristalográfica).

4.2.2 Cálculo do RMSF

A flutuação quadrática média (do inglês, *Root Mean Square Fluctuation*- RMSF) foi calculada por meio do programa *g_rmsf* implementado no pacote computacional GROMACS

v4.6.7 [26]. O RMSF é uma medida da flexibilidade macromolecular e é dada pela seguinte equação:

$$\text{RMSF} = \frac{1}{N} \sum_{j=1}^N \sqrt{(r_i(j) - \bar{r}_i)^2} \quad (4.2)$$

Onde, N é o número de estruturas da trajetória (conformações ao longo do tempo), $r_i(j)$ são as coordenadas cartesianas do átomo i na posição j da trajetória e \bar{r}_i é a média estrutural dos átomos na simulação.

4.2.3 Cálculo do SASA

O cálculo da área da superfície acessível ao solvente (do inglês, *Solvent Accessible Surface Area-SASA*) para os resíduos de aminoácidos presentes no CDR1-3 de cada estrutura de VHH foi obtido através do programa *g_sasa*, que computa superfícies hidrofóbicas, hidrofílicas e totais acessíveis ao solvente, implementado no pacote computacional GROMACS v4.6.7 [26].

4.2.4 Identificação das CDRs nas estruturas de VHHs

As regiões referentes ao CDR 1-3 foram determinadas de maneira automatizada pelo servidor online e banco de dados PyIgClassify disponível em <http://dunbrack2.fccc.edu/pyigclassify/> [38].

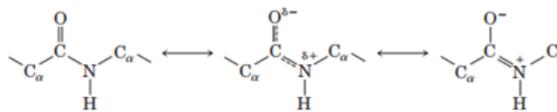
4.2.5 Construção dos gráficos de RMSD e RMSF

As funções utilizadas na construção dos gráficos de RMSD e RMSF foram providas pelo pacote MDPlot implementado no RStudio, uma interface gráfica para o programa estatístico R (<https://rstudio.com/>) e o códigos utilizados estão disposto nos Apêndice I [39].

4.2.6 Gráficos de Ramachandran

Átomos de carbono de resíduos de aminoácidos adjacentes são separados por três ligações covalentes, arranjadas da seguinte forma: C-C-N-C. A ligação C-N em um peptídeo é mais curta do que a ligação C-N em uma amina e os átomos associados na ligação peptídica são coplanares. Isto indica que existe uma ressonância ou um compartilhamento de cargas parciais entre os dois pares de elétrons entre o oxigênio do grupo carboxila e o nitrogênio do grupo amida [Figura 11](#). [40].

Figura 11 – -- Ressonância apresentada pela ligação peptídica.

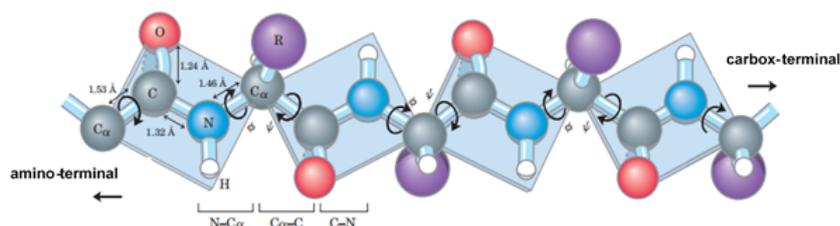


LEHNINGHER (2005) [40]

O backbone uma cadeia polipeptídica pode, assim, ser representada como uma série de planos rígidos com planos consecutivos compartilhando um ponto de rotação comum no C (**Figura 12**). As ligações peptídicas rígidas limitam a faixa de conformações que podem ser assumidas para uma determinada cadeia polipeptídica [40]

Por convenção, os ângulos de rotação resultam da combinação de rotações no C rotuladas ϕ (phi) para a ligação N - C α e ψ (psi) para a ligação **Figura 12**. Em princípio $\phi\psi$ e pode assumir qualquer valor entre -180° e $+180^\circ$, mas muitos valores são proibitivos devido a impedimentos estéricos entre os átomos do backbone de um polipeptídeo e a cadeia lateral de aminoácidos [40].

Figura 12 – -- Ressonância apresentada pela ligação peptídica.



LEHNINGHER (2005) [40]

O gráfico de Ramachandran demonstra teoricamente quais as combinações dos ângulos diedricos $\phi\psi$ são possíveis e pequenas alterações de combinações destes ângulos levam à mudanças de conformação. A combinação desses ângulos em função do tempo de simulação para cada região CDR, em cada uma das estruturas, foi obtida por meio do programa *g_rama* implementado no pacote computacional GROMACS v 4.6.7 [26]. Posteriormente, os valores obtidos para uma mesma região (CDR1, CDR2 e CDR3), nas diferentes estruturas, foram compilados e a matriz de frequência de combinações de ângulos para foi gerada também por meio do pacote MDPlot [39].

4.2.7 Alinhamento múltiplo de sequência e conservação de resíduo por posição

Realizou-se um alinhamento múltiplo de sequências utilizando o servidor da web Clusta Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) [41] Para tanto, foram removidas das

sequências fasta obtidas no PDB todos os resíduos de metionina N-terminal, visto a ocorrência da excisão da metionina N-terminal [42]. Para representar graficamente a conservação dos resíduos empregou-se o servidor Seq2Logo 2.0 (<http://nbc-222.ucsd.edu/pdb2pqr> 2.0.0/) utilizando em conjunto com o tipo P-Weighted Kullback-Leibler. Para diminuir a redundância foi utilizado o método heuristic clustering. E o valor do prior foi de 200. Os parâmetros da exibição gráfica do layout foram definidos como o padrão do servidor.

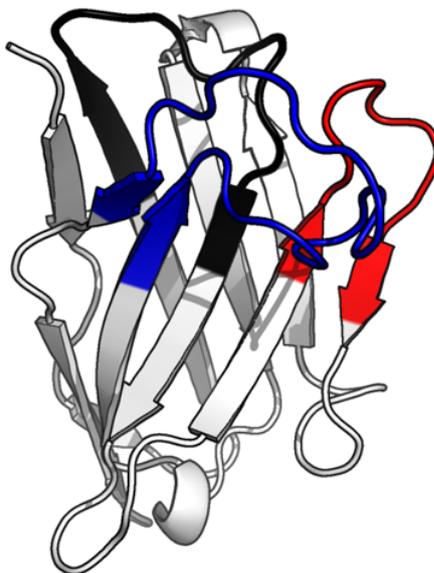
5

Resultados e Discussão

5.1 Identificação das regiões CDRs nas estruturas

A [Tabela 1](#) traz o mapeamento das sequências e posições dos aminoácidos que compõem a CDR1, CDR2 e CDR3 de cada uma das 40 estruturas tridimensionais de VHH aplicadas no presente estudo. A região da proteína onde se situa as CDRs é ilustrada na [Figura 13](#)

Figura 13 – Representação gráfica da estrutura de VHH de código PDB 2X6M. As regiões de CDR estão destacadas da seguinte maneira: CDR1 (preto), CDR2 (Vermelho) e CDR3 (Azul).



Fonte: A Autora.

Tabela 1 – Mapeamento de sequência e posição dos aminoácidos que compõem a CDR1-3 em cada umas das estruturas de VHH.

Código PDB	CDR1		CDR2		CDR3	
	Sequência	Posição	Sequência	Posição	Sequência	Posição
2X89	AASGYTDSRYCMA	23 - 35	RINSGRDITY	50-59	ATDIPLRCRDIVAKGGDGFY	97-117
4ZG1	AASGSTFGIRTMG	23 -35	IISGGSTD	50 – 58	NARVGITMLAH	96 – 106
4JVP	TASGFTDDYYAIG	23 – 35	CITNFDGGTY	50 – 59	AADKGLCSWLRAGGKVTFGS	97 – 116
4QLR	AASGRTFSDYALG	23-35	AISWSGGSTY	50-59	AADKYYTGPGGESVYDY	97-113
4S11	AASGRTFSSFVMG	23-35	SISRSGSVTR	50-59	AADLHRPYGPGSQRTDDYDT	97-116
5BOP	AASGGTFSTYGMG	23 – 35	ASSWTGANTY	50 – 59	AARRWLGGSYFDPGNYDF	97 – 114
5DA4	AASGRTFSSDVMG	26 – 38	AVTRSGGKSY	53 – 62	AAGDTAITSWYGYDY	100 – 114
5E0Q	AASGRSFSNSVMA	25 – 37	VLNWSSGRTS	52 – 61	AASNRGSLYTLDNQNRVED	99 – 117
5VL2	AASGRTHTPYAMG	25 – 37	GIGGVAATT	52 – 61	ATRAMAVSTSPRVYPI	99 – 115
2X10	AAAGRNLRLMYRMG	23 – 35	TMVWSSDTIY	50 – 59	AAGAGWAGTMTDYN	97 – 111
5TP3	AASGFTLDYYYIG	23 – 35	CISGSSGSTY	50 – 58	ATIRSSSWGVCVHYGMDY	93 – 102
5M2W	SITSGGTTN	23 – 35	KAFFRRDYVGYDY	50 – 58	AASGSISGIVVMA	96 – 108
5IP4	TASVNIIGGNHWA	23 – 35	SLSRYNAN	50 – 57	ALENY	95 – 100
5IVN	TASGFTLDHYDIG	23 – 35	CINNSDDDTY	50 – 59	AEARGCKRGRYEYDF	97– 111
5H8D	AASGFTFSRYAMS	25 – 37	RISGGGTIY	52 – 61	YVGGF	99 – 103
5OCL	AASGIIFRELID	25 – 37	SIASAGMTN	52 – 60	HTLPRFSLH	98 – 106
5M7Q	AVSENTGRMG	23 – 32	IITRLGGYTS	47 – 56	AADSRPIYSGTWRY	94 – 107
3STB	AASGRTLSSYAMG	23-35	AINRSGSTF	50-58	AADRFSPVVPPIPVNTVDS	96-115
5E7B	TASGFTFDDSDMG	25 – 37	AIFSDGSTY	52 – 60	AAATTVASPPVRHVCNGY	98 – 116
3EAK	AASGGSEYSYSTFSLG	23 - 38	AIASMGGLTY	53-62	AAVRGYFMRLPSSHNFY	100-117
5HGG	AASGFTLDSYAIG	23-35	CISASGGSTN	50-59	AADHPGLCTSESGRRRYLEV	97-116

Tabela 1 – Mapeamento de sequência e posição dos aminoácidos que compõem a CDR1-3 em cada uma das estruturas de VHH.

	CDR1		CDR2		CDR3	
4WEM	EASGNVDRIDAMG	823-835	YISEGGILN	850-858	AASHWGTLLIKIEH	896-910
4LHQ	SASGGSFSRNAMG	23-35	AINWSASSTY	50-59	AGSSVYAEMPYADSVKATSYNY	97-118
5LZ0	AASGRTFSSYSMG	23 - 35	AISRSDNSTY	50-59	AATPYGSRYYLRELREYDY	97-114
5IML	AASGRTFSSYGMG	22-34	AIRWNGGSTY	49-58	AAGRWDKYGSSFQDEYDY	96-113
5U65	VVSGLTISNYCMR	23-35	SINSAGTTY	50-58	ASSTRVWGGYCGGLDDATNND	96-116
4W6X	TASGYTYRKYCMG	23-35	CINSGGGTSY	50-59	ALSSNSVCPGGHVAWYND	97-114
4EIG	KASGIIFSVYKMT	23-35	LITNNNTM	50-58	NANRGLAGPAY	96-106
4N9O	AASGRTFSSYNMG	23-35	SITSSGDKSD	50-59	ARGLGIYIIRARGGYDH	97-113
3K74	AASGFTFNYYWY	182-194	MINPGGIITK	209-218	AKDWATGLAK	256-265
4AQ1	AASGRTSSAYAMG	23-35	GISSKGGSTY	50-59	AASDKYNFDTSHAGYGY	97-113
2XT1	AASGSFFMSNVMA	23-35	AIRGGDMSTV	50-58	KASGSS	93-102
3G9A	AASGDTFSSYSMA	23-35	NILRDGTTT	50-58	AADSGTQLGYVGAVGLSCLDYVMDY	96-120
5IMK	AASGRTFSSYGMG	22-34	AIRWNGGSTY	49-58	AAGRWDKYGSSFQDEYDY	96-113
4GFT	SAAPERAFSNYAMG	24-37	GITGSGRSQY	52-61	AARVVPVFSdstkgyvy	99-115
4I13	AASERLTVDYAIG	23-35	AISWGGGLTV	50-59	AASRISYRVWNTIPYNKLT	97-116
4LGR	AASGSIASIYRTC	23-35	AITSGGNTY	50-58	NADEAGIGGFNDY	96-108
6EZW	AFSGRTFSMYTMG	24-36	ANRGRGLSPD	51-60	AADLQYGSSWPQRSSAEYDY	98-117
5M13	VASGDIKYISYLG	23-35	ALYTSTGRTY	50-59	AAAEWGSQSPLTQWFYRY	97-114
5MY6	AASGYIFNSCGMG	23-35	RISGDGDTW	50-58	AVCYNLETY	96-104

5.2 Análise das simulações por dinâmica molecular

5.2.1 RMSD

As simulações foram monitoradas através do gráfico do desvio médio quadrático (RMSD) para os átomos do *backbone* (C, N, O), de cada uma das estruturas, ao longo do tempo e tendo como referência as posições atômicas da estrutura cristalográfica inicial. Toda a trajetória foi considerada a fim de preservar a amostragem conformacional. Por meio do RMSD é possível obter uma estimativa da estabilidade estrutural do sistema assim bem como da amplitude da diversidade conformacional.

Os valores de RMSD, calculados para cada uma das 40 estruturas individualmente, apresentam uma faixa de variação entre 0.19 e 0.34 nm e os seus respectivos gráficos da variação de RMSD em função do tempo de simulação estão dispostos no Apêndice II. Para este conjunto de estruturas a média de RMSD foi de 0.20 nm (± 0.04 nm). Em comparação com valores de RMSD reportados na literatura para proteínas o nosso conjunto de VHHs apresenta um perfil de RMSD válido, indicando estabilidade estrutural e a ausência de desvios abruptos de conformação em relação a estrutura inicial [43, 44, 45]. Levando em consideração que algumas estruturas de VHH inseridas em nossa biblioteca encontravam-se complexadas a outras moléculas e as simulações foram executadas apenas para VHHs em sua forma livre e em solução aquosa, flutuações nos gráficos de RMSD durante a produção de DM são aceitáveis.

A literatura dispõe de diversos estudos sobre a dinâmica estrutural de anticorpos convencionais e seus fragmentos, em contrapartida a disponibilidade destas informações a respeito de VHHs ainda é bastante limitada. Dados de simulação por DM para um número pequeno de VHHs, reportados na literatura, exibiram valores de RMSD abaixo de 0.2 nm [46, 47, 48]. Outros dados de simulações por DM, também retirados da literatura, referentes a scFv e Fab de anticorpos convencionais exibiram valores de RMSD na faixa de 0.07 a 0.33 nm e 0.1 a 0.16 nm, respectivamente [49, 50]. Análises de DM para as cadeias pesadas e cadeias leves do Fab, simuladas separadamente, apresentaram os seguintes valores de RMSD médio: 0.23 nm (± 0.04 nm) e 0.17 nm (± 0.03 nm), respectivamente [51].

Essa faixa de valores é semelhante aos valores de RMSD apresentados pelos VHHs que integram o nosso conjunto de dados, apontando para um perfil de mobilidade global semelhante entre estas moléculas. Uma vez que estamos interessados no comportamento das regiões CDR 1-3, os valores médios de RMSD e os seus respectivos desvios padrão foram computados para cada alça CDR em cada uma das estruturas de VHH individualmente e os valores encontram-se no [Tabela 2](#). O comprimento de cada CDR, ou seja o número de resíduos que o compõem, também é mostrado na [Tabela 2](#).

Tabela 2 – Valores médios de RMSD e seus respectivos desvios padrão para cada uma das estruturas, acompanhadas da quantidade de resíduos que compõem cada CDR.

Código PDB	Comp. ^a	CDR1		Comp. ^a	CDR2		Comp. ^a	CDR3	
		RMSD Médio	DP ^b		RMSD Médio	DP ^b		RMSD Médio	DP ^b
2X89	13	0.198535	0.0372402	10	0.05528428	0.01542516	21	0.358625	0.0734304
4GZ1	13	0.147523	0.0341796	8	0.1145642	0.02358114	11	0.136037	0.01937609
4JVP	13	0.1203125	0.03342595	10	0.1314587	0.03202579	20	0.3229403	0.08350516
4QRL	13	0.1457718	0.02181246	10	0.08729875	0.0175684	17	0.219496	0.03903221
4S11	13	0.1581969	0.04705528	10	0.07745051	0.0165915	20	0.08543301	0.02371821
5BOP	13	0.1978736	0.04001426	10	0.1395942	0.01592163	18	0.1786378	0.03604988
5DA4	13	0.1324927	0.03477959	10	0.13149	0.0125253	15	0.3366076	0.0525911
5E0Q	13	0.1540898	0.02644915	10	0.1238716	0.05686578	19	0.2876271	0.165752
5VL2	13	0.1485467	0.02033043	10	0.1184668	0.03563635	17	0.07240001	0.01650517
2X1O	13	0.1937915	0.02147979	10	0.08326433	0.01947944	15	0.1835341	0.03326327
5TP3	13	0.1204921	0.02880867	10	0.09715567	0.04980613	10	0.2468627	0.04373953
5M2W	13	0.2581384	0.03556023	9	0.1110708	0.02132352	13	0.0965056	0.01889537
5IP4	13	0.1827404	0.0437286	8	0.08395986	0.02732435	6	0.05987062	0.01203567
5IVN	13	0.1253076	0.02358382	10	0.1014021	0.0135605	15	0.1509028	0.01875304
5H8D	13	0.1770073	0.05541751	10	0.1774538	0.06942184	5	0.05702809	0.01032236
5OCL	13	0.1169595	0.02384481	9	0.09851026	0.03022394	9	0.07068828	0.01378121
5M7Q	10	0.1443448	0.02107775	10	0.1466733	0.02569157	14	0.1557839	0.0139067
3STB	13	0.198471	0.05200875	9	0.07901419	0.01479679	20	0.2980742	0.1003069
5E7B	13	0.2386534	0.03567436	9	0.06412484	0.03392434	19	0.3487867	0.1260592
3EAK	16	0.1944556	0.0223888	10	0.06499018	0.01681974	18	0.1066424	0.01955968

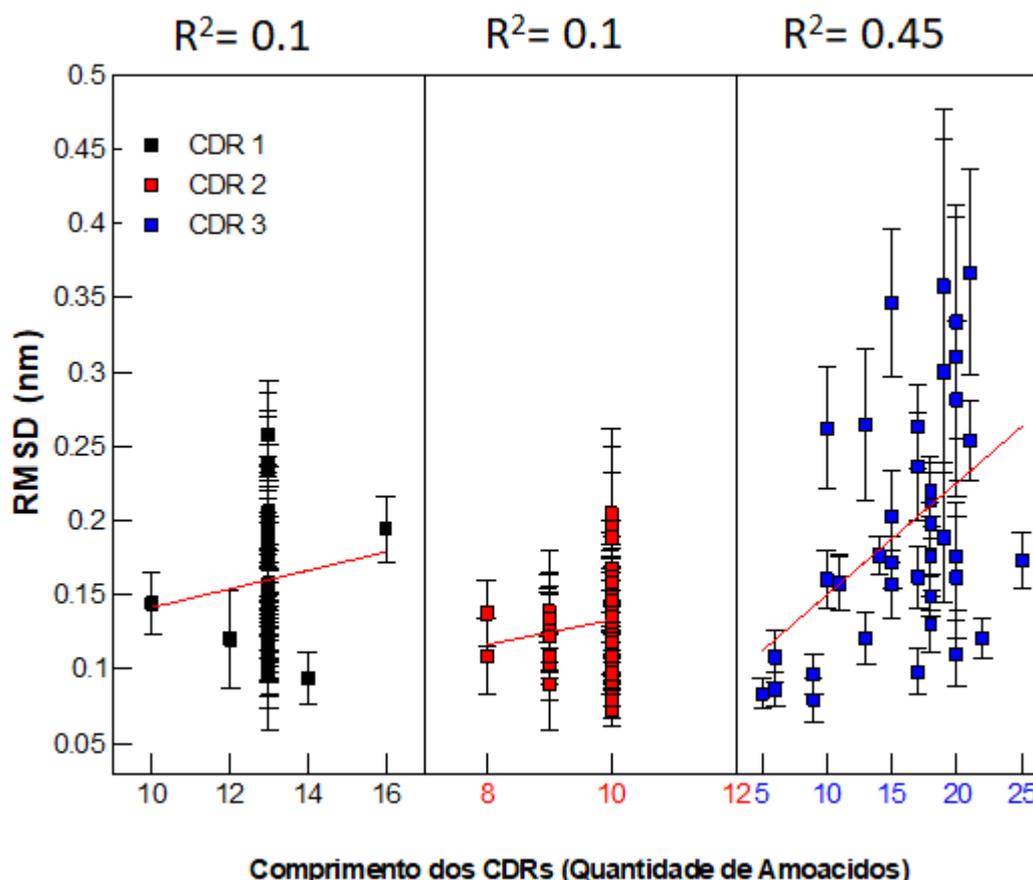
Tabela 2 – Valores médios de RMSD e seus respectivos desvios padrão para cada uma das estruturas, acompanhadas da quantidade de resíduos que compõem cada CDR.

Código PDB	Comp. ^a	CDR1		Comp. . ^a	CDR2		Comp. ^a	CDR3	
		RMSD Médio	DP ^b		RMSD Médio	DP ^b		RMSD Médio	DP ^b
5HGG	13	0.1320518	0.02619536	10	0.1860103	0.04849846	20	0.2671636	0.05743637
4WEM	13	0.149312	0.02205825	9	0.1148474	0.01507324	15	0.1351627	0.02379144
4LHQ	13	0.185926	0.03396678	10	0.0455781	0.01150975	22	0.09653122	0.01390359
5LZ0	13	0.1965426	0.02995731	10	0.1687489	0.04592344	19	0.1687489	0.04592344
5IML	13	0.1927941	0.03415979	10	0.1233535	0.03818648	18	0.1570311	0.02636615
5U65	13	0.1031913	0.02893218	9	0.1158437	0.04307312	21	0.2385486	0.02859407
4W6X	13	0.1030699	0.02077182	10	0.1348466	0.03415562	18	0.1263148	0.01468721
4EIG	13	0.1487153	0.02884758	9	0.1036022	0.02946879	11	0.1360753	0.01989291
4N9O	13	0.2351382	0.0512528	10	0.05225461	0.01233863	18	0.1952966	0.02733214
3K74	13	0.09813245	0.03937275	10	0.1405876	0.01899357	10	0.1386842	0.02103011
4AQ1	13	0.2058599	0.06373135	10	0.09394384	0.03742601	17	0.2481767	0.03035223
2XT1	13	0.1744391	0.02374836	10	0.1047752	0.03292098	6	0.08309417	0.01855072
3G9A	13	0.134585	0.01225168	9	0.08411558	0.01605085	25	0.1526757	0.02031082
5IMK	13	0.1538599	0.04887579	10	0.107756	0.03612757	18	0.2017458	0.02513632
4GFT	14	0.09420854	0.01740922	10	0.1126223	0.03357252	17	0.1404988	0.0222445
4I13	13	0.1261177	0.01913298	10	0.07474372	0.02605077	20	0.1550716	0.03848692
4LGR	13	0.2048378	0.03805738	9	0.1110146	0.03166587	13	0.2492933	0.05470699
6EZW	13	0.1086061	0.02504498	10	0.1362389	0.01258092	20	0.1402113	0.04315279
5M13	13	0.1196115	0.0254131	10	0.07182669	0.01253286	18	0.1548991	0.03878046
5MY6	13	0.1876223	0.0420441	9	0.098002	0.04636095	9	0.05265397	0.0154366

A fim de caracterizar a amplitude da diversidade conformacional das alças CDRs e avaliar a existência de uma correlação com a sua extensão, os valores médios de RMSD e desvios padrão foram plotados em função da quantidade de resíduos presentes no CDR 1, 2 e 3 [Figura 14](#). A região CDR1, 2 e 3 de cada estrutura foi analisada individualmente e cada ponto no gráfico corresponde a região CDR para uma determinada estrutura.

<https://www.overleaf.com/project/5e9099ea79b6da0001154f75>

Figura 14 – RMSD médio de cada região de CDR função do comprimento de CDR.



Cada posição da alça CDR possui a probabilidade de assumir a identidade de qualquer um dos 20 aminoácidos existentes e o tamanho médio de alça CDR 1, CDR 2 e CDR 3, para o nosso conjunto de estruturas, é de 13, 10 e 16 resíduos, gerando 20^{13} , 20^{10} e 20^{16} possíveis combinações de seqüências, respectivamente.

A primeira observação que chama atenção para a CDR1 é que 90% das estruturas analisadas possuem o mesmo comprimento de alça, 13 resíduos. Da mesma maneira, na CDR2 70% das estruturas possuem alças compostas por 10 resíduos. Apesar de muitos pontos estarem sobrepostos, existe uma variação nos valores de RMSD para a CDR1 (0.09 nm a 0.25 nm) e para a CDR2 (0.09 nm a 0.2 nm), o que sugere diversidade conformacional dentro de cada grupo, porém uma diversidade limitada pelo comprimento de alça. A semelhança no comportamento dos pontos na CDR1 e CDR2 corrobora com a literatura no que diz respeito a existência de

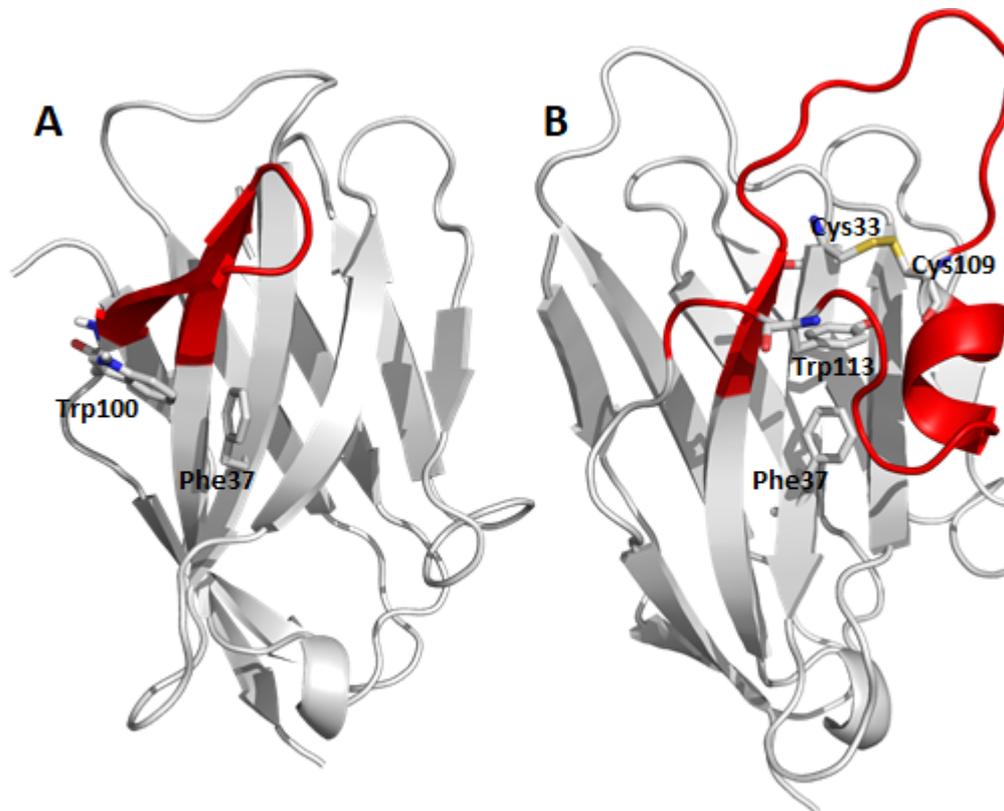
um comprimento de alça ótimo no qual aminoácidos situados em posições chaves levariam à mudança conformacional dessa região [2].

A CDR3 é sabidamente uma região com alta diversidade de sequência e estrutural. Através da inspeção visual do gráfico (**Figura 14**) é possível observar o espaçamento entre os pontos e as grandes variações de RMSD. O comprimento de alças CDR3 é bastante diverso e na faixa que vai de 15 a 20 resíduos observa-se maior quantidade de estruturas de mesmo comprimento de alças. Nesta região a variação de RMSD dentro de grupos que apresentam mesmo comprimento de alças é mais acentuada. Tomando como exemplo, uma estrutura composta por uma alça de 20 resíduos apresenta um RMSD de aproximadamente 0.3 nm, dentro deste mesmo grupo de comprimento de alças existe uma estrutura apresentando um RMSD de 0.1, o equivalente a amplitude conformacional apresentada uma alça composta por 6 resíduos.

Este é um dado surpreendente pois à medida que o número de resíduos de aminoácidos que compõe determinada alça CDR aumenta, seria esperado o aumento do número de graus de liberdade e, conseqüentemente, uma elevação da amplitude da diversidade conformacional, resultando em maiores valores de RMSD. Contudo, não é o que observamos no gráfico da figura 14, nenhuma correlação significativa entre a amplitude da diversidade conformacional das CDRs e o seu comprimento foi encontrada. O estudo de Mitchell et al, menciona que o número de conformações estruturais que alças CDR3 podem adotar para uma dada sequência aumenta em função do comprimento da alça [52]. Em nossos dados, alguns pontos até seguem este comportamento, no entanto no padrão geral isto não é observado e a amplitude conformacional não aumenta linearmente com o comprimento de alças.

A alça CDR3 além de ser determinante para a especificidade de ligação, contribui para a estabilidade da estrutura de VHH [53]. Em algumas estruturas de VHH, resíduos aromáticos posicionados na CDR3 dobram-se em direção a região framework, na qual, em anticorpos convencionais, estaria conectada à VL. Basicamente a perda da VL é compensada pelas interações entre o CDR3 e o framework em VHHs. Os tipos de interações empregadas nas diferentes arquiteturas de VHH variam consideravelmente, entretanto possuem o mesmo objetivo geral que é conferir estabilidade à estrutura. Estas diferenças estruturais podem ser melhor exemplificadas através da comparação entre um VHH anti-gonadotrofina coriônica humana (α -HCG) derivado de lhama, e um VHH anti-lisozima (α -lisozima) derivado de camelo (**Figura 15**) [53].

Figura 15 – A estrutura A é a representação tridimensional do VHH anti α -HCG; Estrutura B é a representação tridimensional do VHH anti *alpha*-lisozima. As alças CDR3 estão destacadas em vermelho.



Fonte: Adaptadp de BOND (2003) [53]

A principal diferença entre as estruturas A e B é o comprimento da alça CDR3, 10 resíduos e 26 resíduos, respectivamente. Na estrutura A um Trp na posição 100 da alça CDR3, dobra-se em direção ao resíduo Phe na posição 37 e uma interação do tipo T-shape é estabelecida. Nesta configuração de alça CDR3 a Phe é blindada do contato com a água, formando um núcleo aromático com Trp100 e Trp103, levando a estabilidade. Para a estrutura B, observa-se uma alça CDR3 mais estruturada, formando uma α -hélice no final C-terminal, que ao dobrar-se em direção a Phe 37 é mantida no lugar por uma ligação dissulfeto entre a Cys109 situada no CDR3 e a Cys33 situada CDR1 [53].

Os tipos de interações específicas entre a CDR3 e o framework são, provavelmente, dependentes do comprimento e da sequência da CDR3. Dessa forma, uma longa alça CDR3 irá necessitar de interações adicionais, que podem restringir a sua mobilidade conformacional, para se manter posicionada sobre o framework conferindo estabilidade à estrutura global [53]. E este fato pode explicar o porquê alças mais compridas nem sempre refletem em maiores valores de RMSD.

5.2.2 Ramachandran

Os valores de phi e psi permitidos para várias estruturas secundárias estão dispostas na [Figura 18](#). As áreas sombreadas em azul escuro refletem conformações que não envolvem impedimento estérico e, portanto, são totalmente favoráveis, em azul médio estão indicadas as conformações permitidas nos limites extremos para contatos atômicos desfavoráveis e em azul claro estão refletidas as conformações que são permitidas se houver um pouco de flexibilidade nos ângulos de ligação. A glicina (Gly) é o resíduo menos impedido estericamente já que sua cadeia lateral tem apenas um átomo de hidrogênio e, por isso, possui maior flexibilidade de combinações de ângulos phi e psi exibindo uma faixa muito mais ampla de conformações permitidas [40].

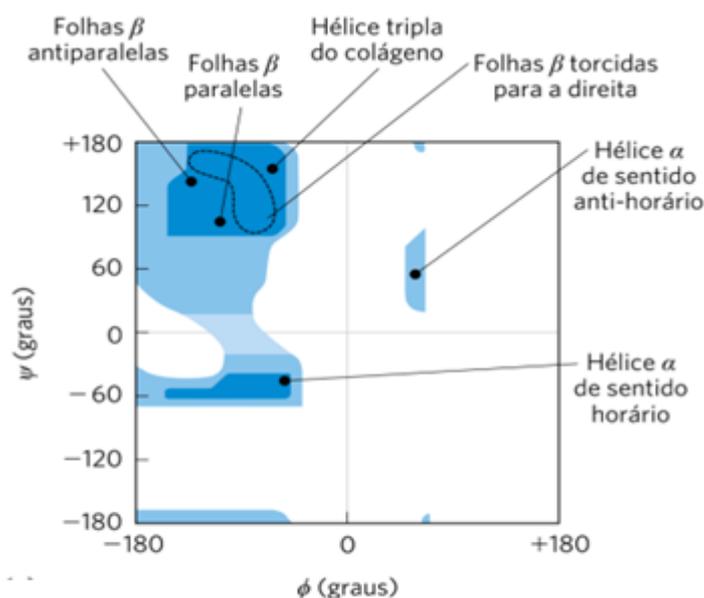
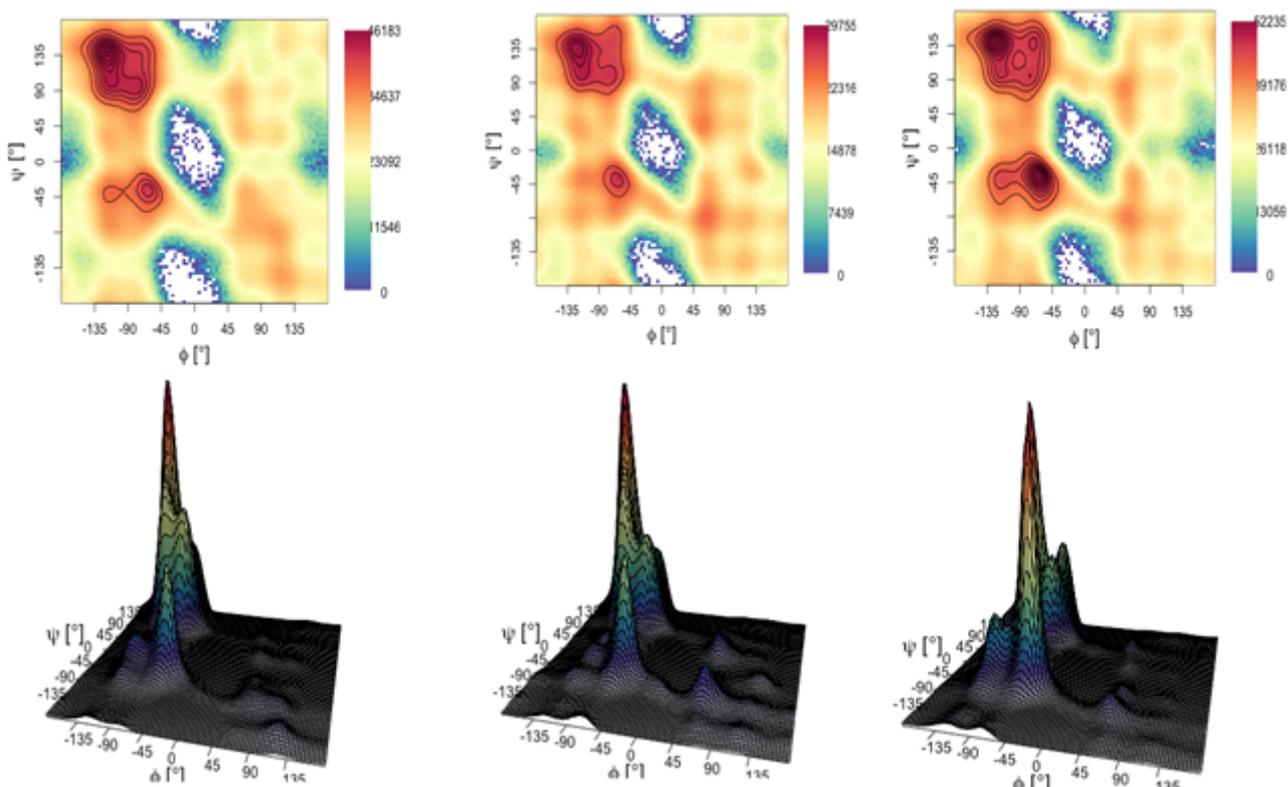


Figura 16 – Regiões do gráfico de Ramachandran para estruturas secundárias.

A amostragem da combinação de ângulos diédricos ϕ/ψ do **backbone** das regiões CDR1-3 é mostrado pelos gráficos de Ramachandran na [Figura 17](#). Este tipo de gráfico é frequentemente utilizado para atribuir propensões de elementos de estrutura secundária à proteína.

Figura 17 – Gráfico de Ramachandran em 2D e 3D, a região contornada em vermelho é a mais favorável e a mais populosa; a amarela permitida e a azul aceitável. A legenda traz o número de combinações encontradas em determinada região (população).



No gráfico de Ramachandran as populações são agrupadas por meio da frequência de combinações ϕ/ψ . Regiões de alça que possuem alta diversidade conformacional podem adotar combinações de ângulos ϕ/ψ em todos os quatro quadrantes no gráfico de Ramachandran. Em contraste aos outros elementos de estrutura secundária mostrados na [Figura 17](#), que ocupam regiões limitadas. Uma vez que alças não possuem região exata para cair no gráfico de Ramachandran, a construção do perfil de distribuição de frequências de combinações de ângulo ϕ/ψ para as alças CDR dos VHH pode ser um dado útil do ponto de vista da engenharia de novos VHHs. Este é o primeiro dado de mapeamento da torção dos ângulos diédricos ao longo do tempo para alças de um grande conjunto de estruturas de VHHs.

O perfil de amostragem de conformações para a CDR1 e CDR2 é bastante símil diferindo nas regiões típicas de glicina, que são mais expressivas para a CDR2, e podem ser melhor visualizadas no gráfico em 3D. A informação mais importante que podemos extrair dos gráficos de Ramachandran é a respeito da CDR3, que apesar da sua alta variabilidade de sequência e estrutural apresenta um perfil bem estruturado, ou seja, apresentando-se com picos concentrados. Além disso, a CDR3 é a única a amostrar na região destinada a -hélice, indicando que a formação deste tipo de estrutura em uma região de alça pode ser bastante comum em VHHs de alças longas.

5.2.3 RMSF

A flutuação quadrática média (RMSF) foi computada para os átomos do Backbone (C, N, O) de cada resíduo de aminoácido em função do tempo durante as simulações de dinâmica molecular, e podem ser verificadas através dos gráficos do Apêndice III. O RMSF fornece uma estimativa da flexibilidade de uma proteína, assim, quanto maiores forem os seus valores, mais elevada será a flexibilidade de uma dada região. Novamente, como as regiões de interesse compreendem aos CDR1, 2 e 3 os seus valores médios de RMSF por resíduo acompanhados dos seus desvios foram calculados [Tabela 3](#). O valor médio de RMSF para as estruturas de VHHs que compõem o nosso conjunto de dados é de 0.12 nm, indicando um perfil de baixa mobilidade global. Dados reportados na literatura apresentam valores de RMSF para VHHs equivalentes aos apresentados aqui (aproximadamente 0.1 nm) [47, 48]. Regiões que possuem seu núcleo baseado em folhas- β flutuam relativamente pouco, de tal forma que, valores altos de RMSF para a estrutura global não eram esperados. Porém, cabe ressaltar que 0.12 nm é um valor baixo de RMSF em proteínas.

Tabela 3 – Valores médios de RMSF e seus respectivos desvios padrão para cada uma das estruturas, acompanhadas da quantidade de resíduos que compõem cada CDR.

CódigoPDB	Comp.	CDR1		Comp.	CDR2		Comp.	CDR3	
		RMSF Médio	DP		RMSF Médio	DP		RMSF Médio	DP
2X89	13	0.198535	0.0372402	10	0.04145	0.00614297	21	0.197795	0.051995
4GZ1	13	0.147523	0.0341796	8	0.0808444	0.0157641	11	0.0635273	0.0141056
4JVP	13	0.1203125	0.03342595	10	0.08316	0.0291577	20	0.163505	0.0645271
4QRL	13	0.1457718	0.02181246	10	0.04879	0.0143248	17	0.115871	0.0430464
4S11	13	0.1581969	0.04705528	10	0.05149	0.020301	20	0.0611273	0.0203554
5BOP	13	0.1978736	0.04001426	10	0.05856	0.0160551	18	0.0806778	0.0310907
5DA4	13	0.1324927	0.03477959	10	0.0376556	0.00772255	15	0.15178	0.0395734
5E0Q	13	0.1540898	0.02644915	10	0.10558	0.023468	19	0.224595	0.0898233
5VL2	13	0.1485467	0.02033043	10	0.07621	0.0186019	17	0.0480941	0.00767777
2X1O	13	0.1937915	0.02147979	10	0.05143	0.00855636	15	0.120287	0.0386899
5TP3	13	0.1204921	0.02880867	10	0.08388	0.0268589	10	0.124933	0.0387674
5M2W	13	0.2581384	0.03556023	9	0.0829333	0.0177867	13	0.0646	0.014747
5IP4	13	0.1827404	0.0437286	8	0.0576125	0.00853387	6	0.0364333	0.00493261
5IVN	13	0.1253076	0.02358382	10	0.03747	0.00887206	15	0.0724933	0.019347
5H8D	13	0.1770073	0.05541751	10	0.13891	0.0529439	5	0.0271	0.00427843
5OCL	13	0.1169595	0.02384481	9	0.0762111	0.0152323	9	0.0395889	0.00646886
5M7Q	10	0.1443448	0.02107775	10	0.08556	0.0269215	14	0.0422357	0.0107596
3STB	13	0.198471	0.05200875	9	0.0487444	0.0105694	20	0.19916	0.0720653
5E7B	13	0.2386534	0.03567436	9	0.0597667	0.0168667	19	0.239011	0.0616162
3EAK	16	0.1944556	0.0223888	10	0.05001	0.00995082	18	0.07515	0.0127362

Tabela 3 – Valores médios de RMSF e seus respectivos desvios padrão para cada uma das estruturas, acompanhadas da quantidade de resíduos que compõem cada CDR.

CódigoPDB	Comp.	CDR1		Comp.	CDR2		Comp.	CDR3	
		RMSF Médio	DP		RMSF Médio	DP		RMSF Médio	DP
5HGG	13	0.1320518	0.02619536	10	0.12185	0.0159275	20	0.11836	0.0360556
4WEM	13	0.149312	0.02205825	9	0.0482778	0.012679	15	0.06562	0.0191274
4LHQ	13	0.185926	0.03396678	10	0.03978	0.00837029	22	0.05515	0.0132792
5LZ0	13	0.1965426	0.02995731	10	0.09169	0.0313977	19	0.1419	0.059432
5IML	13	0.1927941	0.03415979	10	0.06586	0.0198723	18	0.0851056	0.0359794
5U65	13	0.1031913	0.02893218	9	0.0955111	0.0323645	21	0.121919	0.037491
4W6X	13	0.1030699	0.02077182	10	0.08434	0.0292034	18	0.07688899	0.022261
4EIG	13	0.1487153	0.02884758	9	0.0712556	0.0192802	11	0.0765364	0.0208634
4N9O	13	0.2351382	0.0512528	10	0.04034	0.00579045	18	0.122388	0.0353857
3K74	13	0.09813245	0.03937275	10	0.07218	0.0157197	10	0.05408	0.00967813
4AQ1	13	0.2058599	0.06373135	10	0.07041	0.0259456	17	0.134794	0.0292714
2XT1	13	0.1744391	0.02374836	10	0.08023	0.0156437	6	0.04239	0.00881589
3G9A	13	0.134585	0.01225168	9	0.0539111	0.017966	25	0.086036	0.0336967
5IMK	13	0.1538599	0.04887579	10	0.07085	0.026613	18	0.0561111	0.0169592
4GFT	14	0.09420854	0.01740922	10	0.05485	0.024869	17	0.0612176	0.0108869
4I13	13	0.1261177	0.01913298	10	0.06099	0.014916	20	0.107005	0.021426
4LGR	13	0.2048378	0.03805738	9	0.08782222	0.0207315	13	0.146538	0.0453755
6EZW	13	0.1086061	0.02504498	10	0.04574	0.00443777	20	0.10587	0.032483
5M13	13	0.1196115	0.0254131	10	0.04113	0.00756557	18	0.04113	0.00756557
5MY6	13	0.1876223	0.0420441	9	0.0927333	0.0180997	9	0.0468333	0.00731693

Com o aumento do comprimento de alças e por conseguinte o aumento do número de graus de liberdade, esperávamos observar uma tendência crescente na flexibilidade estrutural das alças CDR. A elevação da flexibilidade é observada através do aumento nos valores de RMSF. A fim de avaliar a influência do comprimento das CDRs na flexibilidade desta região, os valores médios de RMSF e seus respectivos desvios padrão foram plotados em função do comprimento de alça (**Figura 18**).

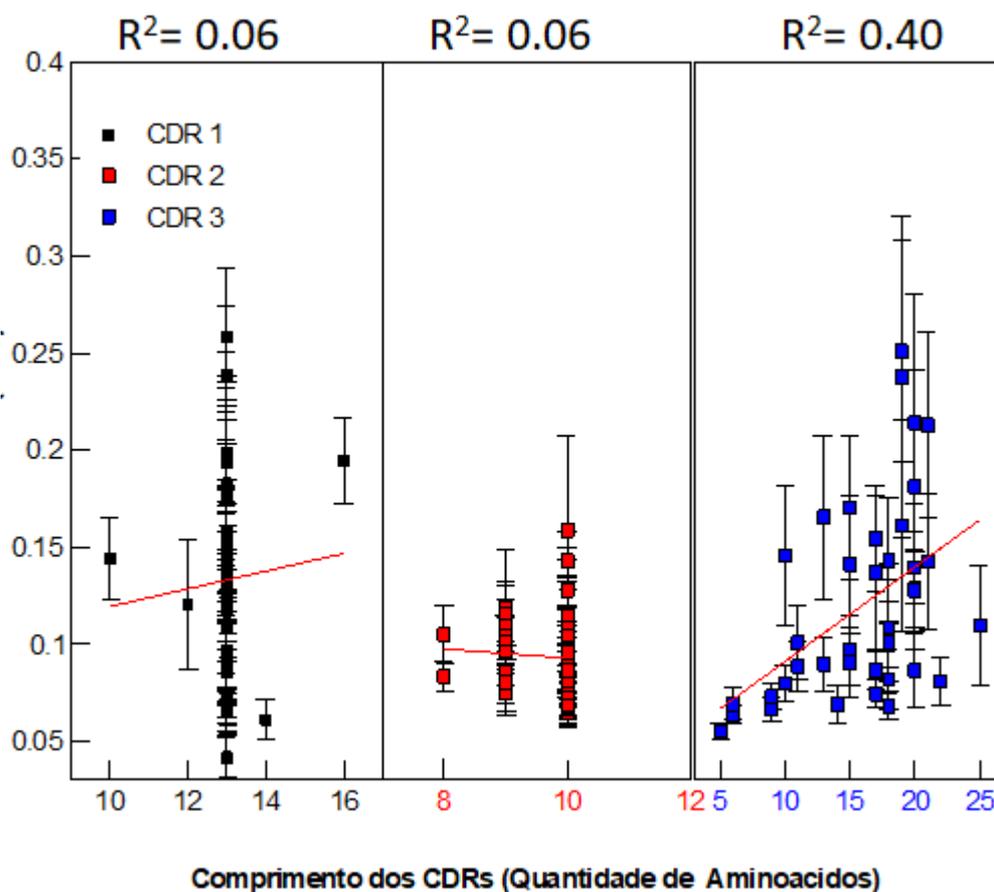


Figura 18 – Gráfico do RMSF médio de cada região de CDR função do comprimento de CDR.

Como mostram os valores de R^2 , nenhuma correlação significativa foi encontrada entre o comprimento de alças CDR e a sua flexibilidade. Os valores médios de RMSF para a CDR1, CDR2 e CDR3 foram de 0.09 nm, 0.09 nm e 0.05 nm, respectivamente. Uma vez que as CDRs são regiões compostas por alças existe naturalmente uma maior liberdade para se mover. Dessa forma, era esperado que os valores de RMSF médios fossem mais elevados em comparação à estrutura global, principalmente para a CDR3, em contrapartida o que estes resultados indicam é que as CDRs são regiões surpreendentemente rígidas.

Como o CDR1 e CDR2 não possuem diversidade de comprimento de alça, novamente observa-se o agrupamento de pontos. A CDR2 é a região que possui as alças mais curtas, além disso não possui alta diversidade de sequência em comparação as demais CDRs dos VHs e isto nos dá uma clareza do porquê esta região apresentou a menor flexibilidade. Já para a

CDR1 o que chama a atenção é a expressiva variação nos valores de RMSF dentro o grupo que apresenta mesmo comprimento de alça, que pode ser justificada: **I)** Pela variabilidade de sequência conferida pelos hotspots mutacionais nas posições 28 e 30; **II)** [10]. Comumente, algumas estruturas de VHHs que possuem uma ligação dissulfeto extra situada entre as alças CDR1 e CDR3, podendo levar a uma redução da mobilidade de ambas as regiões [2, 12]. A explicação plausível para a baixa flexibilidade das longas alças CDR3 também poderia decair sobre a sua funcionalidade estrutural na manutenção da estabilidade global da estrutura de VHH. Para alguns pontos, até é possível observar uma tendência crescente entre flexibilidade e comprimento de alça, porém este não é o perfil geral apresentado pelos nossos dados.

5.2.4 Análise da flexibilidade em função da energia livre de ligação de VHHs

Para 9 estruturas de VHH do nosso conjunto valores de energia livre de ligação, medidos experimentalmente, frente a antígenos foram reportados na literatura [54]. A flexibilidade de alças é crítica para funções biológicas, exercendo um importante papel no reconhecimento molecular [55]. Se faz necessário que haja um balanço de flexibilidade, pois extremos são desfavoráveis entropicamente à interação intermolecular frente a antígenos.

As alças CDRs de VHHs são as regiões responsáveis pelo reconhecimento ao antígeno e os nossos dados até o momento indicaram que a flexibilidade seria um reflexo direto não do comprimento de alças, mas da identidade de resíduos que as compõem. Para avaliar a influência da flexibilidade das regiões CDR1-3 na energia livre de ligação a antígenos, valores de RMSF médios por resíduo para cada CDR1-3, para cada uma das 9 estruturas, foram plotados versus ΔG_{bind} Figura 19.

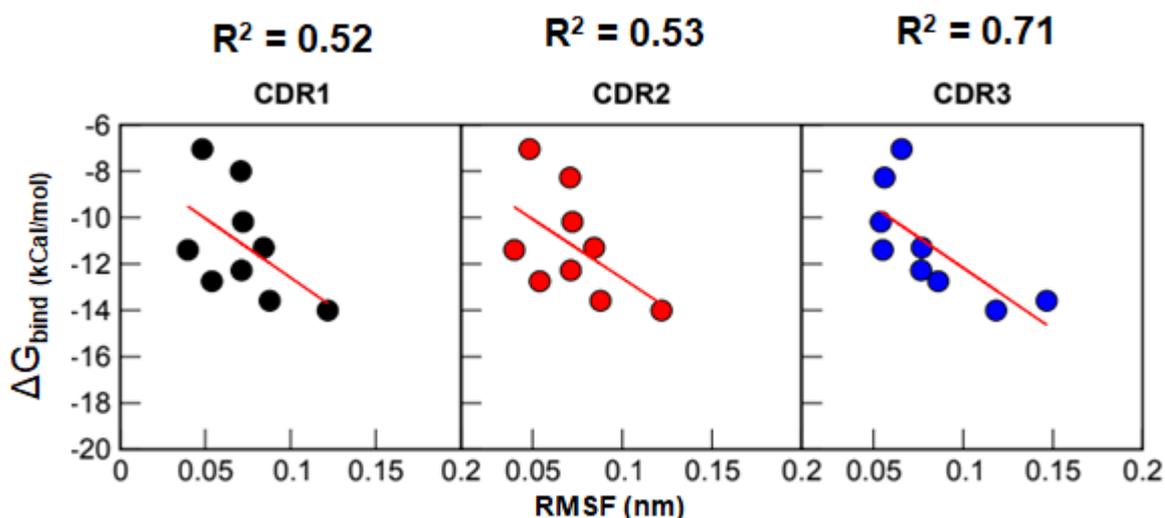


Figura 19 – Correlação entre a energia livre de ligação e o RMSF médio para o CDR1-3 de 9 estruturas de VHH.

Correlações baixas foram encontradas para a CDR1 e CDR2, 0.52 e 0.53, respectivamente. Já para o CDR3 pudemos obter uma correlação moderada de 0.71. Dentro do nosso conjunto de dados todos os CDRs apresentam-se rígidos, mas para o CDR3, dentro dessa faixa, quanto maior a flexibilidade melhor a energia de ligação. E este é um forte indicio da ocorrência de um ajuste induzido para esta região ao interagir com o antígeno. Utilizando como referência os dados obtidos no presente estudo este valor ótimo de flexibilidade deve ter o seu limite em 0.2 nm. Adicionalmente, uma variável que não está sendo avaliada leva em conta quais os tipos de antígenos a que estes VHHs estão ligados, pois VHHs podem interagir com um amplo espectro de moléculas através de diferentes mecanismos. Diferentes mecanismos de interação implicarão em diferentes comprimentos de alças e a identidade de resíduos que as compõem.

A flexibilidade de alças em anticorpos convencionais têm sido tema de alguns estudos e duas tendências podem ser delineadas a este respeito. A primeira, indica que um aumento da rigidez da alça CDR3 do domínio VH de anticorpos convencionais ocorre após a maturação de afinidade e, dessa maneira, a flexibilidade conformacional é redistribuída para o domínio VL [56]. A segunda, indica que não há grandes diferença de flexibilidade entre anticorpos antes de depois da maturação de afinidade. Sugerindo que a rigidificação é apenas um dos mecanismos biofísicos para aumentar a afinidade [57].

Dentro do nosso conjunto de dados de uma maneira geral as alças CDR apresentaram baixa flexibilidade, todavia para uma parcela das estruturas foi demonstrado que um aumento controlado de flexibilidade reflete em aumento da afinidade de ligação. Partindo desse pressuposto e estendendo as observações a respeito de rigidez/flexibilidade de anticorpos convencionais para os VHHs, os nossos dados ajustam-se à sugestão de que de que diferentes antígenos exigirão, de fato, diferentes mecanismos biofísicos para aumentar a afinidade.

5.2.5 SASA e Conservação de Sequência

Para identificar a exposição dos resíduos nos CDRs, avaliou-se a superfície de acessibilidade ao solvente (SASA). Por meio do SASA é possível avaliar a propensão de um determinado resíduo interagir com o antígeno, baseado na sua área exposta ao solvente. O SASA foi calculado para cada resíduo que compõem a CDR1, 2 e 3 individualmente em cada uma das estruturas de VHHs. Os valores de SASA foram compilados para uma mesma região (CDR1, 2 e 3) nas diferentes estruturas e aplicados na construção de um índice que caracterizou estatisticamente os resíduos que mais contribuem dentro de cada área de CDR1, CDR2 e CDR3 para os VHHs do nosso conjunto de dados. O índice de acessibilidade ao solvente (IAS) foi cálculo por meio da seguinte equação:

$$IAS = \frac{\sum SASA_{RES} f_{CDR}}{N} \quad (5.1)$$

Na qual, o somatório dos valores de SASA para um dado resíduo foi multiplicado pela

frequência com que esse resíduo está presente naquela CDR dividido pelo número total de estruturas do conjunto (N). Os valores de contribuição para cada resíduo em cada CDR é mostrado nas Figura 20.

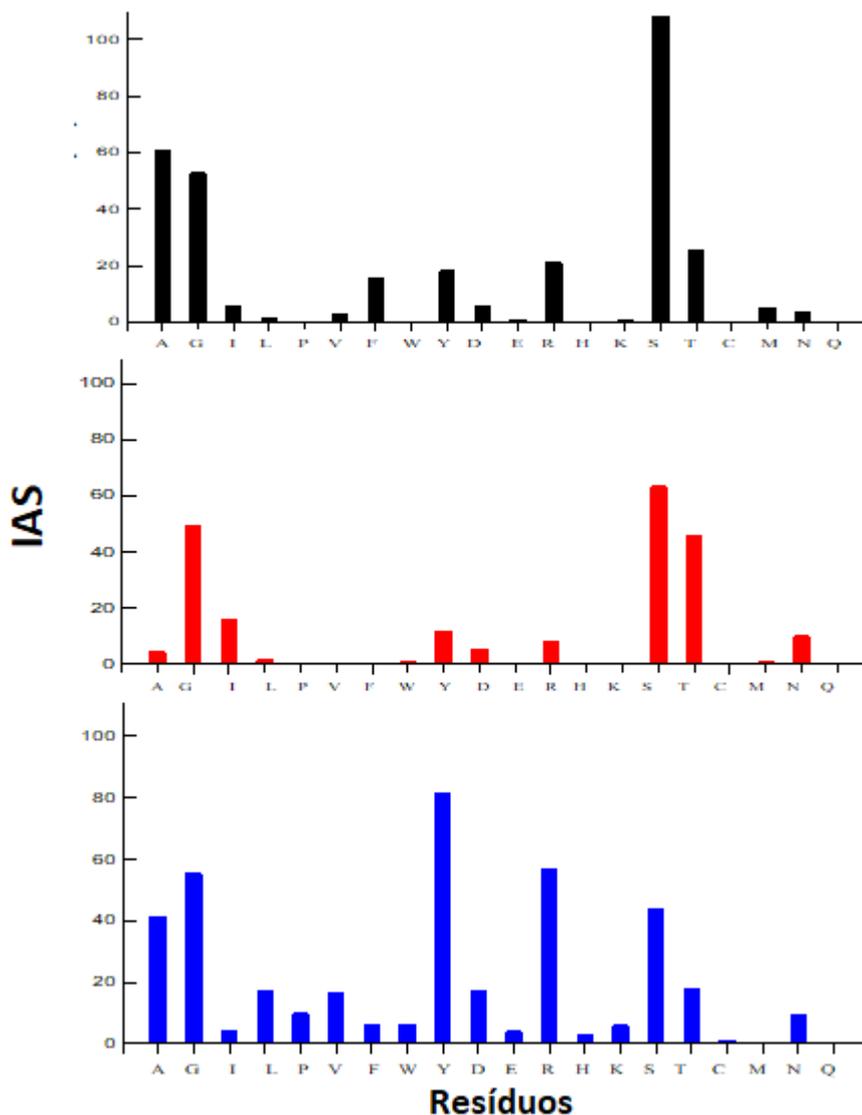
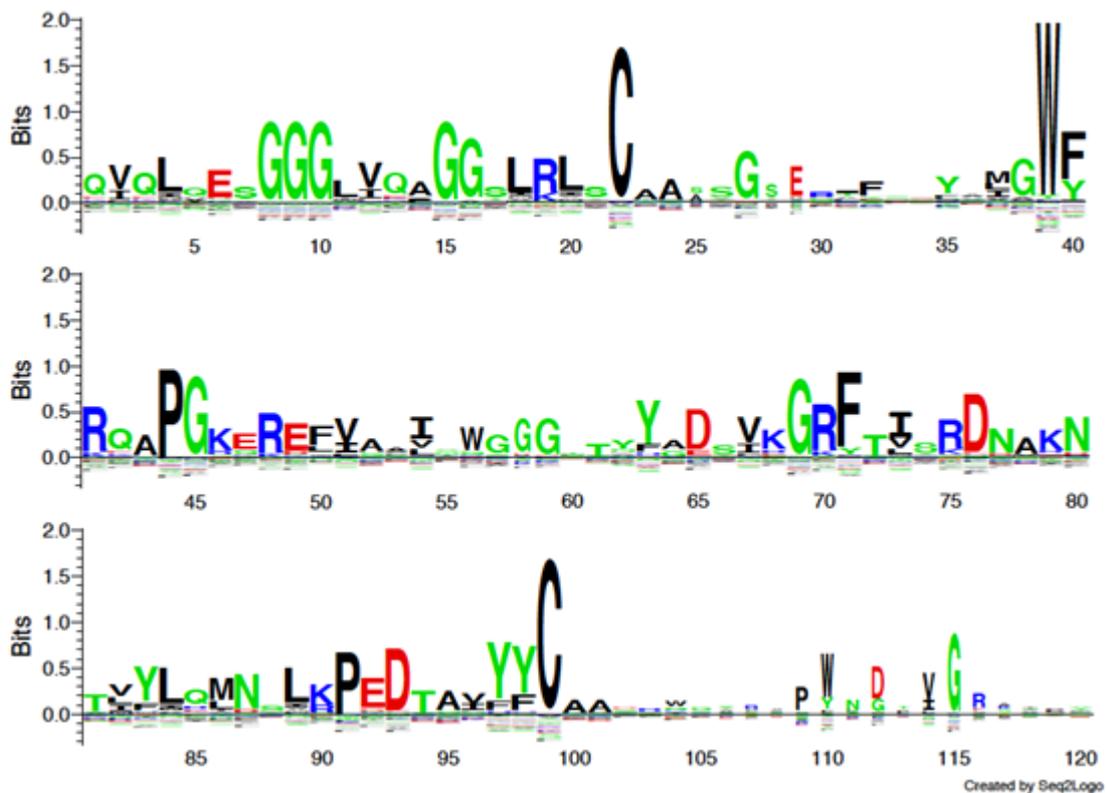


Figura 20 – índice de acessibilidade ao solvente para o CDR1-3.

A informação fornecida por esses índices para as CDRs acoplada ao gráfico de conservação de sequência referente as 40 estruturas do nosso conjunto de dados Figura 21 sugerem a possibilidade que um determinado resíduo possui de exercer uma função estrutural ou de ser determinante para a especificidade de ligação. No gráfico de conservação de sequência, quanto maior a escala de representação do código de uma letra do aminoácido, maior a frequência desse aminoácido nessa posição.

Figura 21 – Gráfico de Conservação de sequência; as regiões referentes aos CDR1, 2 e 3 estão demarcadas por bastões em preto, vermelho e azul.



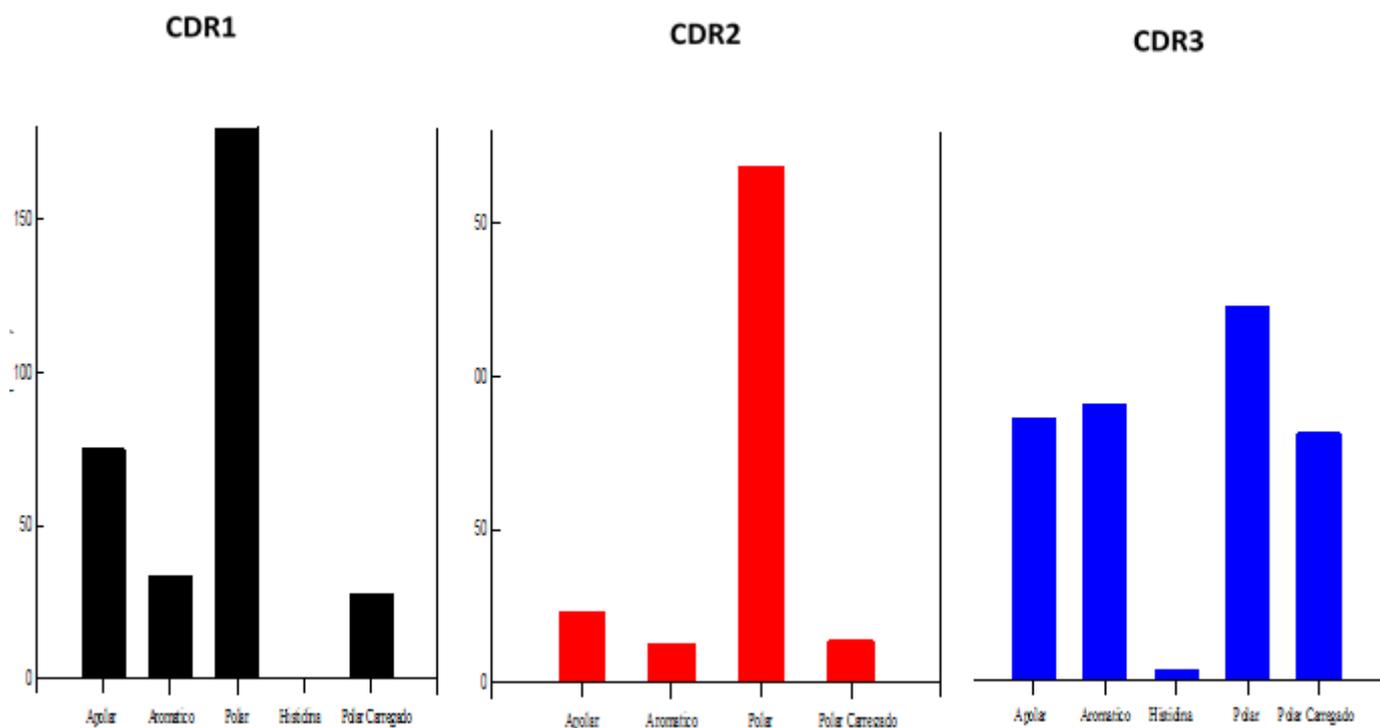
Se um determinado resíduo é conservado em qualquer posição de uma determinada CDR e, além disso, apresenta baixo índice de acessibilidade ao solvente dentro daquela CDR, sugere-se que este resíduo tem alta propensão de exercer uma função estrutural. Em uma outra situação, para o caso de um resíduo que não é conservado para nenhuma posição de uma determinada CDR, mas apresenta alto índice de acessibilidade ao solvente dentro daquela CDR, sugere-se que este resíduo tem alta propensão de conferir especificidade frente a antígenos. Ressalta-se que o fato de um resíduo ser apontado como uma peça na manutenção estrutural de determinada alça não exclui a possibilidade do mesmo estar envolvido em uma interação com o antígeno. Em média, a região CDR1 situa-se entre as posições 23-35 e os resíduos que mais contribuem para a sua área exposta ao solvente são a alanina (A), glicina (G), treonina (T) e serina (S). Dentre estes resíduos a A, S e G são conservadas no CDR1, sendo que a A e S ocupam mais de uma posição. Apesar do ácido glutâmico (E) ocupar uma posição conservada no CDR1 o seu valor de IAS é baixo, sugerindo que provavelmente apenas a sua região polar carregada está disponível na superfície, enquanto que o resto de sua cadeia está posicionada em direção ao núcleo proteico. Esse tipo de configuração confere solubilidade à estrutura de VHH. A arginina (R) também está numa posição conservada mas apresenta maior contribuição para a superfície da CDR1, podendo estar envolvida em tipos de interações diferentes da última exemplificada. Outros resíduos polares (ácido aspártico (D), lisina (K) e asparagina (N)) e apolares (valina (V),

isoleucina (I), e leucina (L), assumem contribuições diversas para a área da CDR1, sugerindo que possuem a possibilidade de conferir especificidade. Dentre os aromáticos conservados e que contribuem para a área do CDR1 estão a tirosina (Y) e a fenilalanina (F).

A CDR2 em média situa-se entre as posições 59 e 60 para o nosso conjunto de dados. Em comparação as demais CDRs, a CDR2 é região a que possui a área da superfície menos exposta ao solvente. Os resíduos que possuem maior contribuição para a área exposta são a serina (S), treonina (T) e glicina (G). Dentre estes, apenas a G é conservada e, de maneira interessante ocupa três posições consecutivas na sequência. Uma vez que a CDR2 é a alça mais curta e menos diversa sequencialmente, estas três G devem conferir maior flexibilidade a região. Os resíduos polares arginina (R), ácido aspártico (D) e asparagina (N) e o aromática tirosina (Y) não são conservados e estão expostos ao solvente, sugerindo que podem conferir especificidade ao CDR2. Isoleucina (I), alanina (A) e valina (V) são conservados e é muito provável que tenham funcionalidade estrutural. A CDR3 é bastante variada e para no nosso conjunto de dados assume uma faixa que vai da posição 94 a 122. A variabilidade de sequência desta região fica bem evidente no gráfico de conservação de sequência. Os resíduos que mais contribuem para a área exposta ao solvente são alanina (A), glicina (G), arginina (R), serina (S) e tirosina (Y). Dentre estes, a A, Y, G e R são conservados sugerindo função estrutural. A prolina (P) e o triptofano (W) são sabidamente resíduos que tem função estrutural e, como mostram os nossos dados, não exibiram altos valores de exposição ao solvente e são conservados. O ácido glutâmico também aparece em posição conservada e possui um valor de exposição ao solvente razoável. Todos os demais resíduos que possuem a possibilidade de conferir especificidade a essa região. A CDR1 não possui nenhuma contribuição da glutamina(Q), já a CDR2 não possui nenhuma contribuição da histidina. Todos os resíduos de aminoácidos foram observados para a CDR3.

A fim de caracterizar o perfil químico geral das regiões de CDR, agrupamos os resíduos encontrados em cada CDR em cinco grupos de contribuição: apolares, aromáticos, polares histidina e carregados [Figura 22](#).

Figura 22 – Caráter químico das regiões CDR1-3 de acordo com as contribuições para a área acessível ao solvente.



Para a CDR1 e CDR2 a propensão de interações é majoritariamente polar. A CDR2 é a região com menor contribuição de resíduos aromáticos. Na CDR3 o que se observa é um balanço entre resíduos de caráter hidrofóbico e polar.

6

Conclusões

Através de simulações por dinâmica molecular o presente estudo identificou e caracterizou o perfil dinâmico estrutural das alças CDR1-3 para 40 estruturas de VHHs. Diferentemente do que era esperado, a amplitude da diversidade conformacional e a flexibilidade não possuem correlação com o comprimento das alças CDR. Tanto a diversidade de conformacional quanto a flexibilidade são um reflexo não do comprimento de alças, mas da identidade de resíduos que compõem essas alças. Excepcionalmente, ao elevar a flexibilidade da CDR3 há um aumento na afinidade de ligação. Estes dados são inéditos na literatura de VHHs até o momento e suplantam a hipótese de que a rigidificação ou flexibilização de alças é apenas um dos mecanismos biofísicos para aumentar a afinidade. A combinação de variados comprimentos de alças e diversidade de sequência elevam o repertório das interações que o CDR3 utiliza para exercer as suas duas funções: contribuir com estabilidade global da estrutura de VHH e a interação com o antígeno. Apesar da diversidade de sequência a CDR3 apresentou um perfil de alça bem estruturada, com picos concentrados, reforçando dessa alça na manutenção da estabilidade global da estrutura de VHH. O índice de acessibilidade ao solvente (IAS) forneceu informações valiosas a respeito do caráter químico das CDRs e da identidade dos resíduos que a compõem, atribuindo pesos à propensão que estes possuem de atuarem como uma peça estrutural ou conferindo especificidade frente a antígenos. Todos esses achados poderão, potencialmente, racionalizar o desenho racional de novos VHHs.

Referências

- 1 Abul K. Abbas, Andrew H. Lichtman, S. P.; LICHTMAN, A. H.; PILLAI, S. *Basic Immunology: Functions and Disorders*. [S.l.: s.n.], 2016. 232–233, 243–244 p. ISBN 978-0-323-39082-8. Citado 3 vezes nas páginas 12, 16 e 17.
- 2 MUYLDERMANS, S. Nanobodies: Natural Single-Domain Antibodies. *Annual Review of Biochemistry*, Annual Reviews, v. 82, n. 1, p. 775–797, jun 2013. ISSN 0066-4154. Disponível em: <<http://www.annualreviews.org/doi/10.1146/annurev-biochem-063011-092449>>. Citado 7 vezes nas páginas 12, 18, 19, 20, 21, 44 e 52.
- 3 SCHMID, A. S.; NERI, D. *Advances in antibody engineering for rheumatic diseases*. Nature Publishing Group, 2019. 197–207 p. Disponível em: <<http://www.nature.com/articles/s41584-019-0188-8>>. Citado na página 12.
- 4 GOLDSBY, R. A. et al. Osborne. *Immunology, Fifth Edition*. WH Freeman, p. 228–291, 2002. Citado 3 vezes nas páginas 16, 17 e 18.
- 5 CHIU, M. L. et al. Antibody structure and function: The basis for engineering therapeutics. *Antibodies*, Multidisciplinary Digital Publishing Institute, v. 8, n. 4, p. 55, 2019. Citado 2 vezes nas páginas 16 e 17.
- 6 MAHAJAN, S. P.; VELEZ-VEGA, C.; ESCOBEDO, F. A. Tilting the Balance between Canonical and Noncanonical Conformations for the H1 Hypervariable Loop of a Llama VHH through Point Mutations. *The Journal of Physical Chemistry B*, American Chemical Society, v. 117, n. 1, p. 13–24, jan 2013. ISSN 1520-6106. Disponível em: <<https://pubs.acs.org/doi/10.1021/jp3075496>>. Citado na página 16.
- 7 WOOF, J. M.; BURTON, D. R. *Human antibody-Fc receptor interactions illuminated by crystal structures*. Nature Publishing Group, 2004. 89–99 p. Disponível em: <<http://www.nature.com/articles/nri1266>>. Citado 2 vezes nas páginas 17 e 18.
- 8 KINDT, T. J. et al. *Kuby immunology*. [S.l.]: Macmillan, 2007. Citado 2 vezes nas páginas 17 e 18.
- 9 KARADAG, M. et al. Physicochemical determinants of antibody-protein interactions. In: . [s.n.], 2019. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S1876162319300653>>. Citado na página 17.
- 10 VINCKE, C.; MUYLDERMANS, S. Introduction to heavy chain antibodies and derived nanobodies. In: *Single Domain Antibodies*. [S.l.]: Springer, 2012. p. 15–26. Citado 3 vezes nas páginas 19, 21 e 52.
- 11 MITCHELL, L. S.; COLWELL, L. J. Analysis of nanobody paratopes reveals greater diversity than classical antibodies. *Protein Engineering, Design and Selection*, Oxford University Press, v. 31, n. 7-8, p. 267–275, jul 2018. ISSN 1741-0126. Disponível em: <<https://academic.oup.com/peds/article/31/7-8/267/5058991>>. Citado na página 19.
- 12 FLAJNIK, M. F.; DESCHACHT, N.; MUYLDERMANS, S. A case of convergence: Why did a simple alternative to canonical antibodies arise in Sharks and Camels? *PLoS Biology*,

Public Library of Science, v. 9, n. 8, aug 2011. ISSN 15449173. Citado 3 vezes nas páginas 19, 21 e 52.

13 NAMBA, A. M.; SILVA, V. B. da; SILVA, C. D. Dinâmica molecular: teoria e aplicações em planejamento de fármacos. *Eclética Química*, SciELO Brasil, v. 33, n. 4, p. 13–24, 2008. Citado 2 vezes nas páginas 21 e 29.

14 SCHLICK, T. *Molecular modeling and simulation: an interdisciplinary guide: an interdisciplinary guide*. [S.l.]: Springer Science & Business Media, 2010. v. 21. Citado 5 vezes nas páginas 21, 22, 23, 26 e 28.

15 MCQUARRIE, D. A.; SIMON, J. D. *Physical chemistry: a molecular approach*. [S.l.]: University science books Sausalito, CA, 1997. v. 1. Citado 2 vezes nas páginas 21 e 22.

16 ALLINGER, N. L. *Molecular structure: understanding steric and electronic effects from molecular mechanics*. [S.l.]: John Wiley & Sons, 2010. Citado 2 vezes nas páginas 21 e 22.

17 SALMASO, V.; MORO, S. Bridging molecular docking to molecular dynamics in exploring ligand-protein recognition process: An overview. *Frontiers in pharmacology*, Frontiers, v. 9, p. 923, 2018. Citado 2 vezes nas páginas 21 e 26.

18 CHATFIELD, D. Christopher j. cramer: *Essentials of computational chemistry: Theories and models wiley, new york, 2002. 562pp (isbn 0-471-48551-9) 110.00. Theoretical Chemistry Accounts*, v. 6, n. 108, p. 367–368, 2002. Citado 4 vezes nas páginas 21, 22, 28 e 29.

19 ADCOCK, S. A.; MCCAMMON, J. A. Molecular dynamics: survey of methods for simulating the activity of proteins. *Chemical reviews*, ACS Publications, v. 106, n. 5, p. 1589–1615, 2006. Citado 5 vezes nas páginas 21, 26, 27, 28 e 29.

20 COELHO, L. W. et al. Aplicação de mecânica molecular em química inorgânica. *Química Nova*, SciELO Brasil, v. 22, n. 3, p. 396–404, 1999. Citado 3 vezes nas páginas 21, 22 e 28.

21 JENSEN, F. *Introduction to computational chemistry*. [S.l.]: John wiley & sons, 2017. Citado 3 vezes nas páginas 21, 29 e 30.

22 BASSALO, J. M. F.; CATTANI, M. S. D. *Osciladores Harmônicos Clássicos e Quânticos*. [S.l.]: Editora Livraria da Física, 2009. Citado na página 21.

23 HÖLTJE, H.; FOLKERS, G. Small molecules. *Molecular Modeling: Basic Principles and Applications; Mannhold R., Kubinyi H., Timmerman H., Eds*, p. 9–63, 1996. Citado na página 23.

24 LEACH, A. R.; LEACH, A. R. *Molecular modelling: principles and applications*. [S.l.]: Pearson education, 2001. Citado 5 vezes nas páginas 23, 26, 27, 28 e 29.

25 OOSTENBRINK, C. et al. A biomolecular force field based on the free enthalpy of hydration and solvation: the gromos force-field parameter sets 53a5 and 53a6. *Journal of computational chemistry*, Wiley Online Library, v. 25, n. 13, p. 1656–1676, 2004. Citado 2 vezes nas páginas 24 e 33.

26 HESS, B. et al. Gromacs 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *Journal of chemical theory and computation*, ACS Publications, v. 4, n. 3, p. 435–447, 2008. Citado 4 vezes nas páginas 27, 33, 34 e 35.

- 27 VERLET, L. Computer” experiments” on classical fluids. i. thermodynamical properties of lennard-jones, 159, 98–103. *Phys. Rev.*, 1967. Citado 2 vezes nas páginas 27 e 29.
- 28 PAQUET, E.; VIKTOR, H. L. Molecular dynamics, monte carlo simulations, and langevin dynamics: a computational review. *BioMed research international*, Hindawi, v. 2015, 2015. Citado 2 vezes nas páginas 28 e 30.
- 29 YOUNG, D. C. et al. Computational chemistry: A practical guide for applying techniques to real world problems. 2001. Citeseer, 1999. Citado na página 29.
- 30 SCHWABL, F.; BREWER, W. *Statistical Mechanics*. Springer Berlin Heidelberg, 2002. (Advanced Texts in Physics). ISBN 9783540431633. Disponível em: <<https://books.google.com.br/books?id=o-HyHvRZ4VcC>>. Citado na página 29.
- 31 FISER, A.; SALI, A. ModLoop: automated modeling of loops in protein structures. *Bioinformatics*, Oxford Academic, v. 19, n. 18, p. 2500–2501, dec 2003. ISSN 1367-4803. Disponível em: <<https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btg362>>. Citado na página 32.
- 32 LEAVER-FAY, A. et al. Rosetta3: An object-oriented software suite for the simulation and design of macromolecules. *Methods in Enzymology*, v. 487, n. C, p. 545–574, 2011. ISSN 00766879. Citado na página 32.
- 33 BERENDSEN, H. J. C. et al. Interaction Models for Water in Relation to Protein Hydration. In: . [S.l.]: Springer, Dordrecht, 1981. p. 331–342. Citado na página 33.
- 34 TIRONI, I. G. et al. A generalized reaction field method for molecular dynamics simulations. *The Journal of Chemical Physics*, American Institute of Physics, v. 102, n. 13, p. 5451–5459, apr 1995. ISSN 00219606. Disponível em: <<http://aip.scitation.org/doi/10.1063/1.469273>>. Citado na página 33.
- 35 BUSSI, G.; DONADIO, D.; PARRINELLO, M. Canonical sampling through velocity rescaling. *Journal of Chemical Physics*, American Institute of Physics, v. 126, n. 1, p. 014101, jan 2007. ISSN 00219606. Disponível em: <<http://aip.scitation.org/doi/10.1063/1.2408420>>. Citado na página 33.
- 36 BERENDSEN, H. J. et al. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, American Institute of Physics, v. 81, n. 8, p. 3684–3690, oct 1984. ISSN 00219606. Disponível em: <<http://aip.scitation.org/doi/10.1063/1.448118>>. Citado na página 33.
- 37 MAIOROV, V. N.; CRIPPEN, G. M. Size-independent comparison of protein three-dimensional structures. *Proteins: Structure, Function, and Genetics*, John Wiley & Sons, Ltd, v. 22, n. 3, p. 273–283, jul 1995. ISSN 0887-3585. Disponível em: <<http://doi.wiley.com/10.1002/prot.340220308>>. Citado na página 33.
- 38 ADOLF-BRYFOGLE, J. et al. Pyigclassify: a database of antibody cdr structural classifications. *Nucleic acids research*, Oxford University Press, v. 43, n. D1, p. D432–D438, 2015. Citado na página 34.
- 39 MARGREITTER, C.; OOSTENBRINK, C. MDplot: Visualise molecular dynamics. *R Journal*, v. 9, n. 1, p. 164–186, 2017. ISSN 20734859. Citado 2 vezes nas páginas 34 e 35.

- 40 LEHNINGER, A. L.; NELSON, D. L.; COX, M. M. *Princípios de bioquímica*. [S.l.]: Ediciones Omega, 2005. Citado 3 vezes nas páginas 34, 35 e 46.
- 41 THOMPSON, G. A.; SCHULZ, A. *Macromolecular trafficking in the phloem*. [S.l.]: Elsevier Current Trends, 1999. 354–360 p. Citado na página 35.
- 42 GIGLIONE, C.; BOULAROT, A.; MEINNEL, T. *Protein N-terminal methionine excision*. [S.l.]: Birkhauser Verlag Basel, 2004. 1455–1474 p. Citado na página 36.
- 43 FREIRE, M. C. et al. Mapping Putative B-Cell Zika Virus NS1 Epitopes Provides Molecular Basis for Anti-NS1 Antibody Discrimination between Zika and Dengue Viruses. *ACS Omega*, American Chemical Society, v. 2, n. 7, p. 3913–3920, jul 2017. ISSN 24701343. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/30023708><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC6044859>>. Citado na página 40.
- 44 COELHO, D. F. et al. The influence of biotinylation on the ability of a computer designed protein to detect B-cells producing anti-HIV-1 2F5 antibodies. *Journal of Molecular Graphics and Modelling*, Elsevier Inc., v. 93, p. 107442, dec 2019. ISSN 18734243. Citado na página 40.
- 45 TRZESNIAK, D.; LINS, R. D.; GUNSTEREN, W. F. van. Protein under pressure: Molecular dynamics simulation of the arc repressor. *Proteins: Structure, Function, and Bioinformatics*, John Wiley & Sons, Ltd, v. 65, n. 1, p. 136–144, aug 2006. ISSN 08873585. Disponível em: <<http://doi.wiley.com/10.1002/prot.21034>>. Citado na página 40.
- 46 SOLER, M. A.; De Marco, A.; FORTUNA, S. Molecular dynamics simulations and docking enable to explore the biophysical factors controlling the yields of engineered nanobodies. *Scientific Reports*, Nature Publishing Group, v. 6, n. 1, p. 1–11, oct 2016. ISSN 20452322. Citado na página 40.
- 47 VOORDIJK, S. et al. Molecular dynamics simulations highlight mobile regions in proteins: A novel suggestion for converting a murine V(H) domain into a more tractable species. *Journal of Molecular Biology*, Academic Press, v. 300, n. 4, p. 963–973, jul 2000. ISSN 00222836. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0022283600938907>>. Citado 2 vezes nas páginas 40 e 48.
- 48 SOLER, M. A. et al. Binding affinity prediction of nanobody-protein complexes by scoring of molecular dynamics trajectories. *Physical Chemistry Chemical Physics*, Royal Society of Chemistry, v. 20, n. 5, p. 3438–3444, feb 2018. ISSN 14639076. Citado 2 vezes nas páginas 40 e 48.
- 49 OSAJIMA, T.; HOSHINO, T. Roles of the respective loops at complementarity determining region on the antigen-antibody recognition. *Computational Biology and Chemistry*, Elsevier Ltd, v. 64, p. 368–383, oct 2016. ISSN 14769271. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S1476927116301773>>. Citado na página 40.
- 50 MOHAMMADI, M. et al. Insilico analysis of three different tag polypeptides with dual roles in scFv antibodies. *Journal of Theoretical Biology*, Academic Press, v. 402, p. 100–106, aug 2016. ISSN 10958541. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0022519316300443>>. Citado na página 40.
- 51 STAVRAKOUDIS, A. Computational modelling and molecular dynamics simulations of a cyclic peptide mimotope of the CD52 antigen complexed with CAMPATH-1H antibody. *Molecular Simulation*, Taylor & Francis Group, v. 36, n. 2, p. 127–137, feb 2010. ISSN 0892-7022.

Disponível em: <<http://www.tandfonline.com/doi/abs/10.1080/08927020903124593>>. Citado na página 40.

52 MITCHELL, L. S.; COLWELL, L. J. Comparative analysis of nanobody sequence and structure data. *Proteins: Structure, Function, and Bioinformatics*, John Wiley and Sons Inc., v. 86, n. 7, p. 697–706, jul 2018. ISSN 08873585. Disponível em: <<http://doi.wiley.com/10.1002/prot.25497>>. Citado na página 44.

53 BOND, C. J.; MARSTERS, J. C.; SIDHU, S. S. Contributions of CDR3 to VHH domain stability and the design of monobody scaffolds for naive antibody libraries. *Journal of Molecular Biology*, Academic Press, v. 332, n. 3, p. 643–655, sep 2003. ISSN 00222836. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0022283603009677>>. Citado 2 vezes nas páginas 44 e 45.

54 ZAVRTANIK, U.; HADŽI, S. A non-redundant data set of nanobody-antigen crystal structures. *Data in Brief*, v. 24, p. 0–3, 2019. ISSN 23523409. Citado na página 52.

55 VERLI, H. Bioinformática: da biologia à flexibilidade molecular. Sociedade Brasileira de Bioquímica e Biologia Molecular, 2014. Citado na página 52.

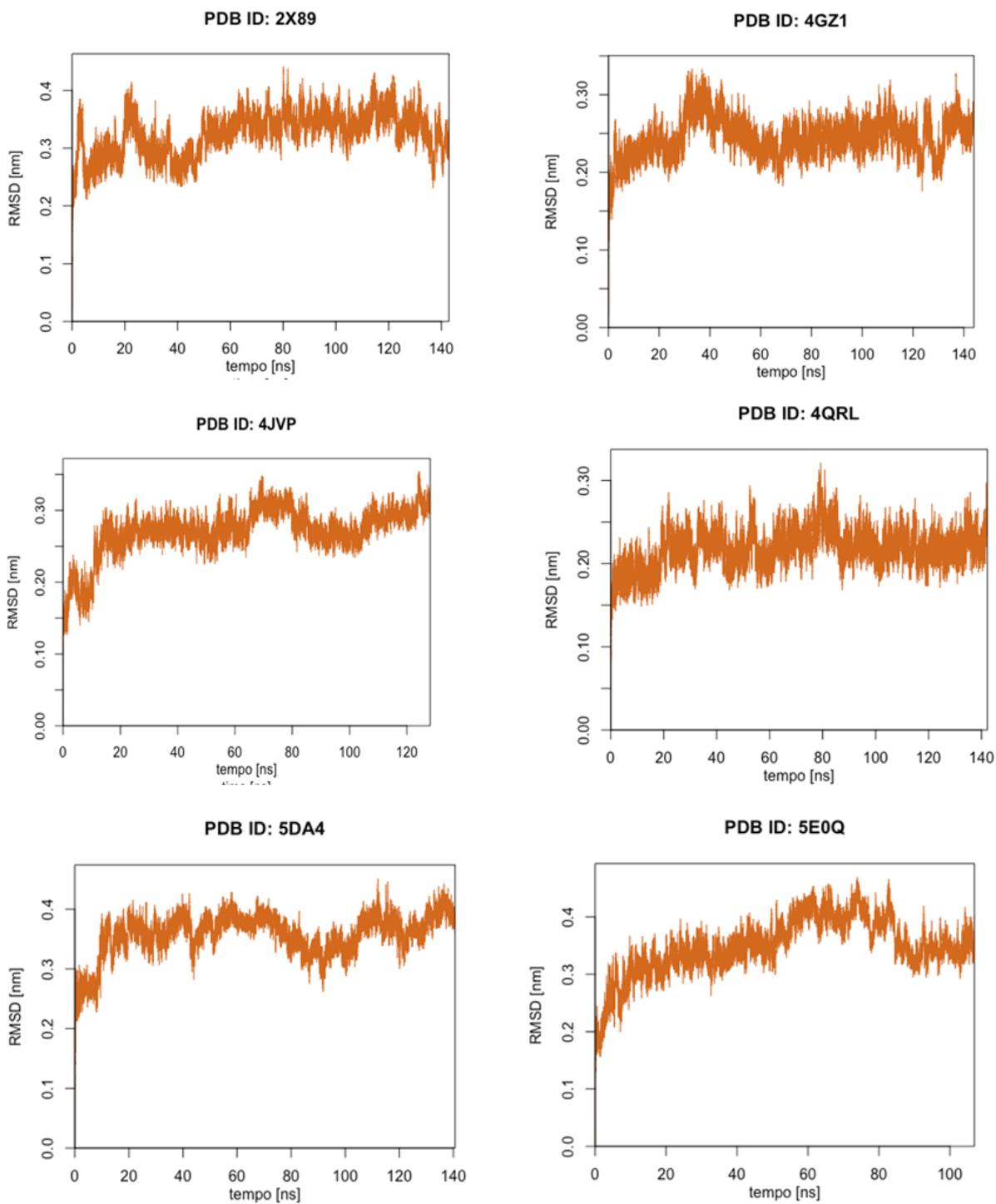
56 LI, T. et al. Rigidity emerges during antibody evolution in three distinct antibody systems: evidence from qsfr analysis of fab fragments. *PLoS computational biology*, Public Library of Science, v. 11, n. 7, 2015. Citado na página 53.

57 JELIAZKOV, J. R. et al. Repertoire analysis of antibody cdr-h3 loops suggests affinity maturation does not typically result in rigidification. *Frontiers in immunology*, Frontiers, v. 9, p. 413, 2018. Citado na página 53.

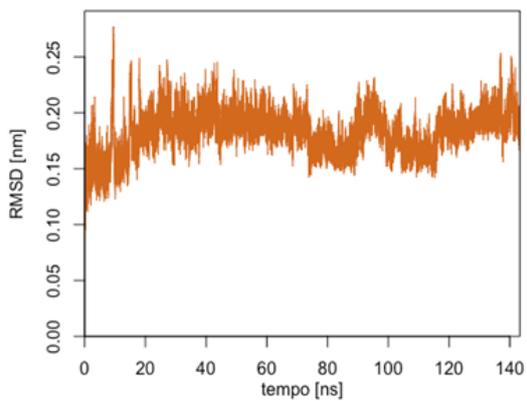
Anexos

ANEXO A – RMSD

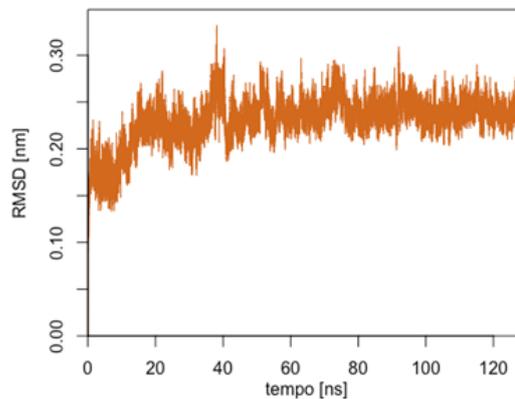
Figura 23 – Gráficos de RMSD em função do tempo de simulação referente ao Backbone (C, N, O) gerados por dinâmica molecular paraca cada uma das estruturas de VHH.



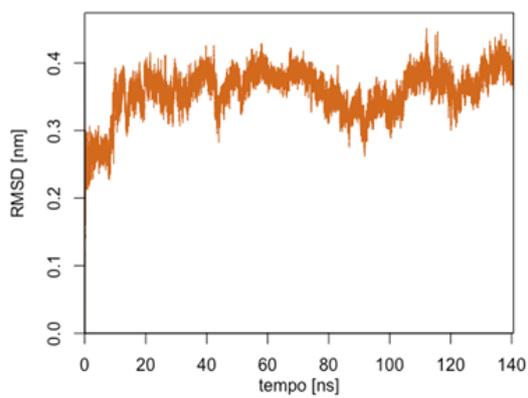
PDB ID: 4S11



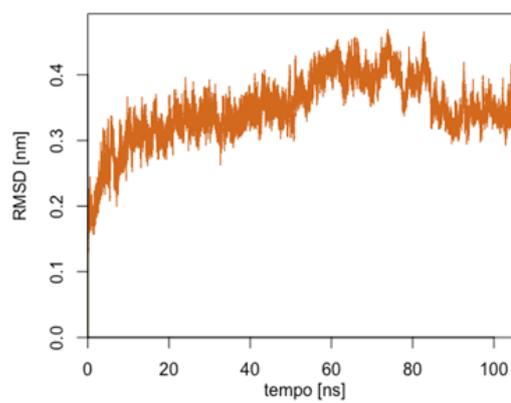
PDB ID: 5BOP



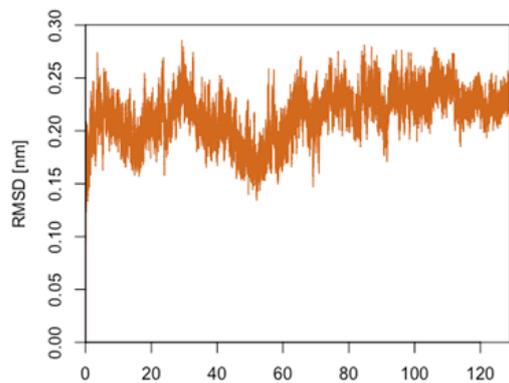
PDB ID: 5DA4



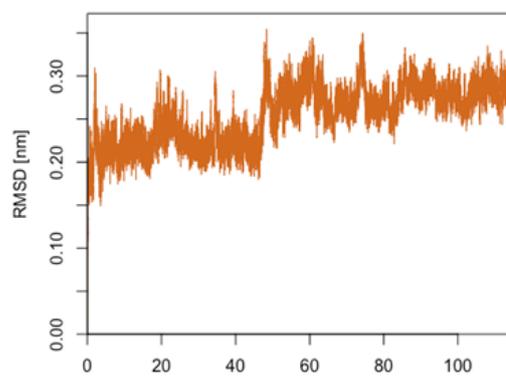
PDB ID: 5E0Q



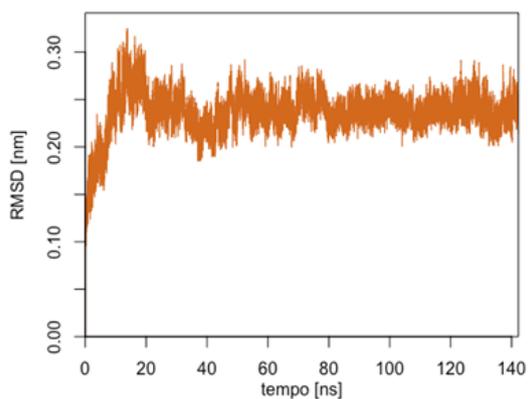
PDB ID: 5VL2



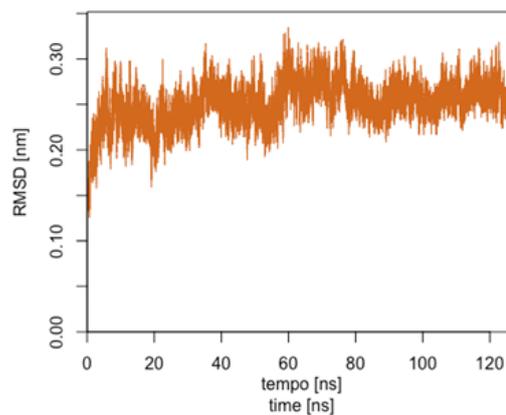
PDB ID: 2X1O



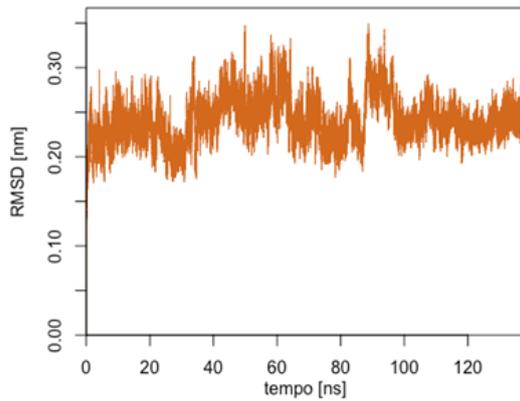
PDB ID: 5TP3



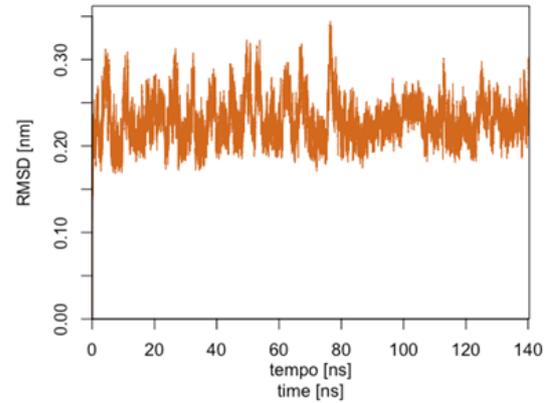
PDB ID: 5M2W



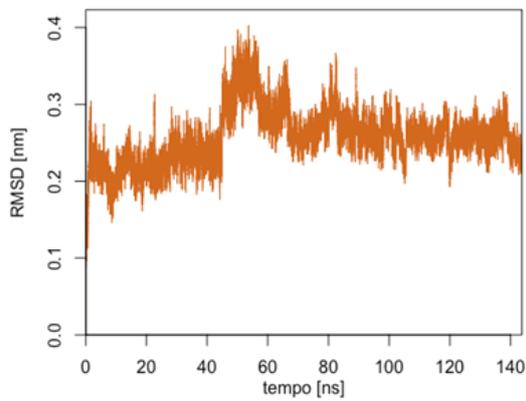
PDB ID: 5IP4



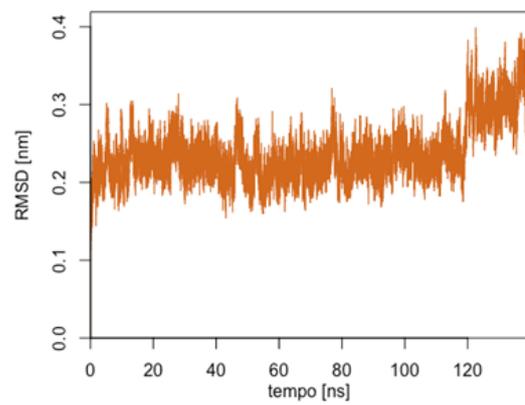
PDB ID: 5IVN



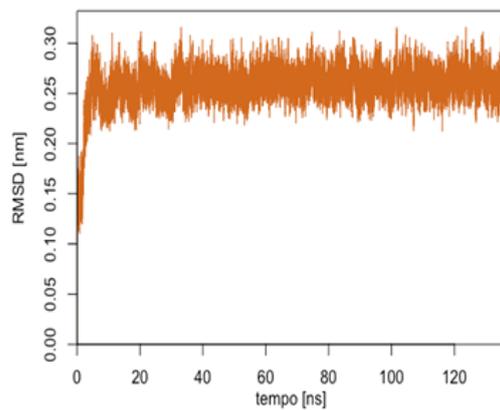
PDB ID: 5H8D



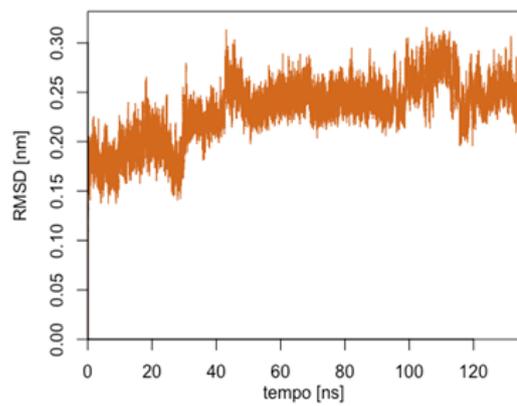
PDB ID: 5OCL



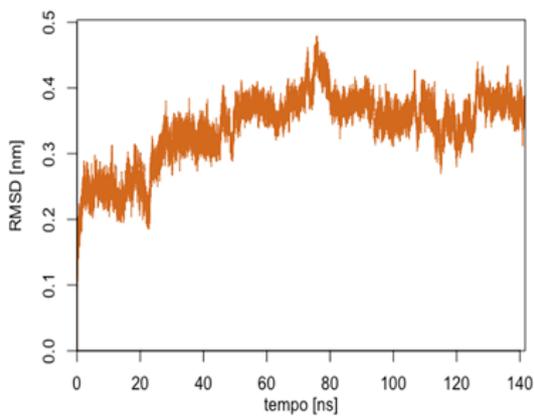
PDB ID: 5M7Q



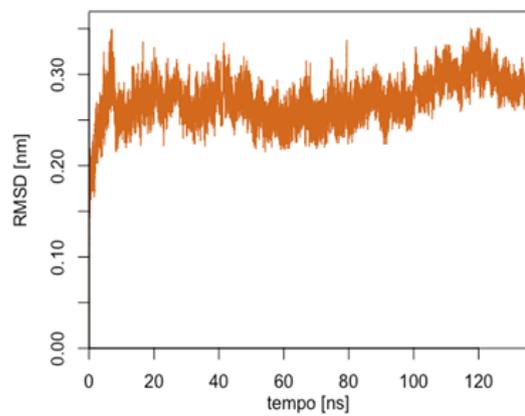
PDB ID: 3STB



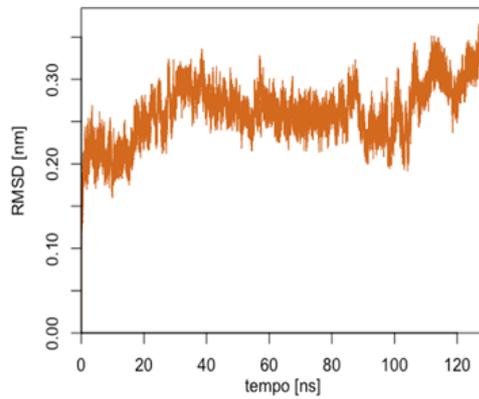
PDB ID: 5E7B



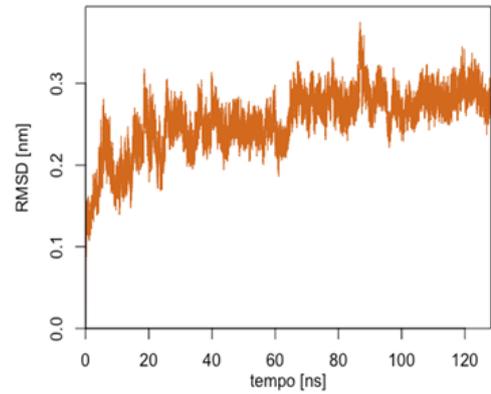
PDB ID: 3EAK



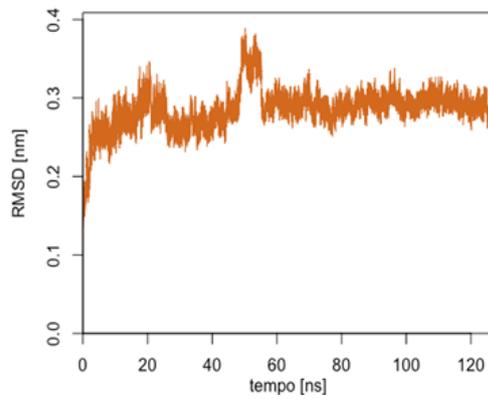
PDB ID: 5HGG



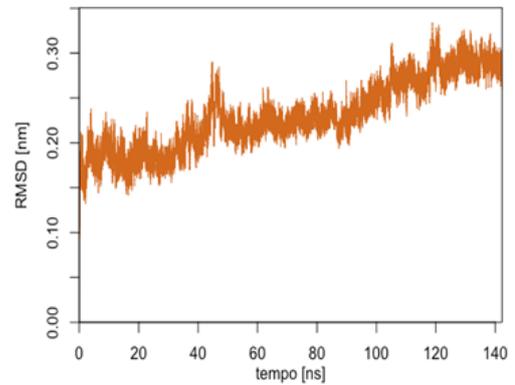
PDB ID: 4WEM



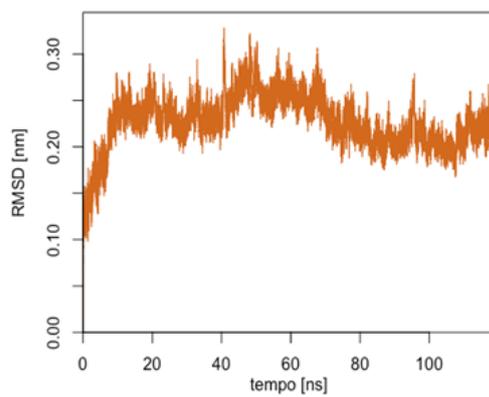
PDB ID: 4LHQ



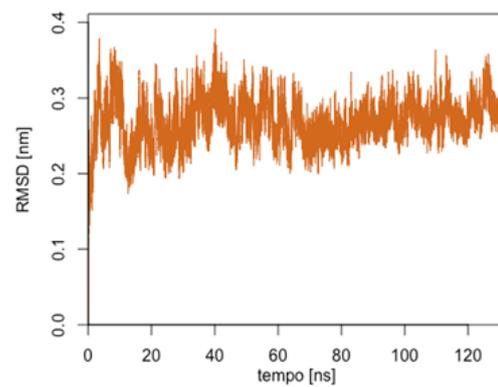
PDB ID: 5LZ0



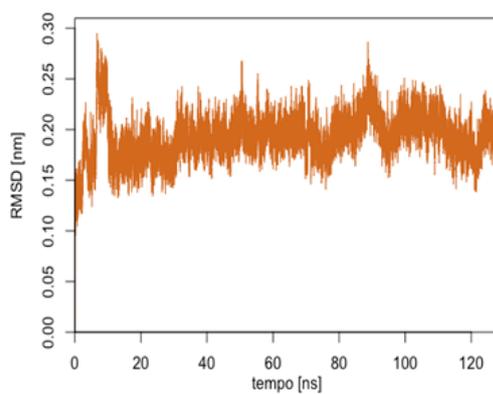
PDB ID: 5IML



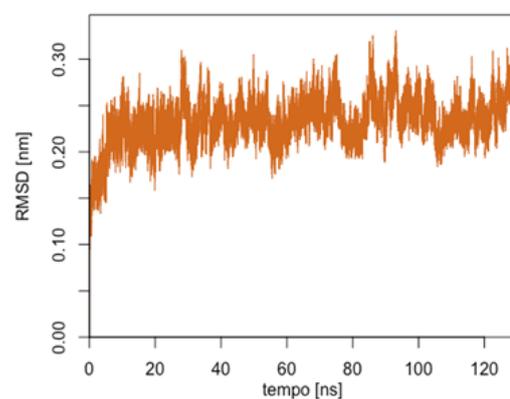
PDB ID: 5U65



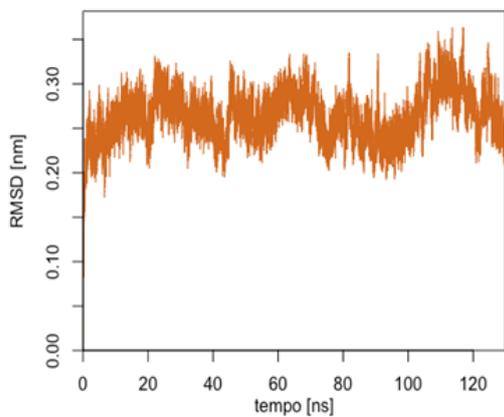
PDB ID: 4W6X



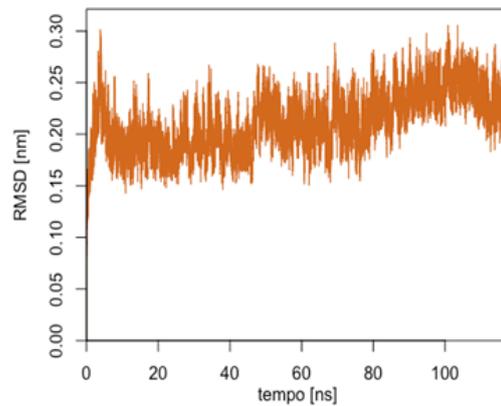
PDB ID: 4EIG



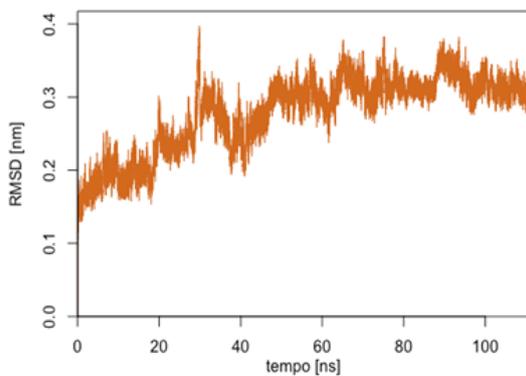
PDB ID: 4N9O



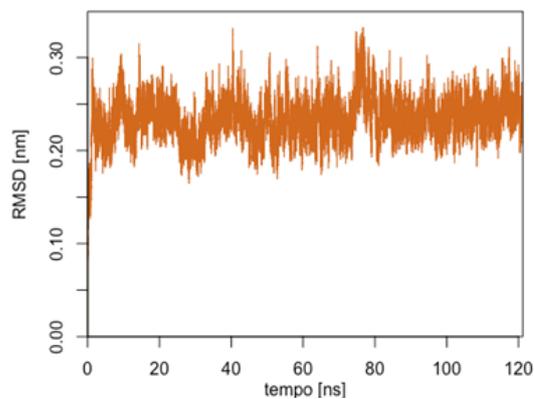
PDB ID: 3K74



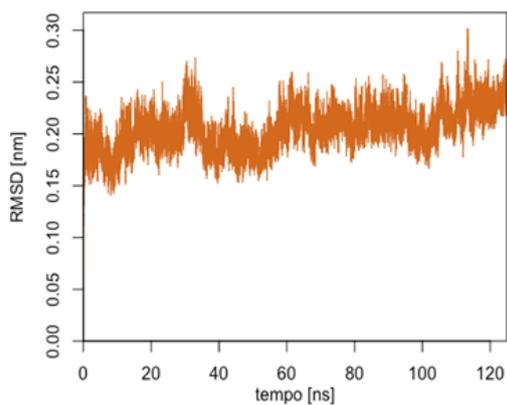
PDB ID: 4AQ1



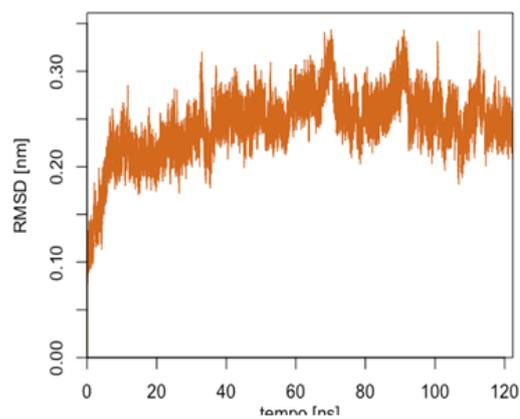
PDB ID: 2XT1



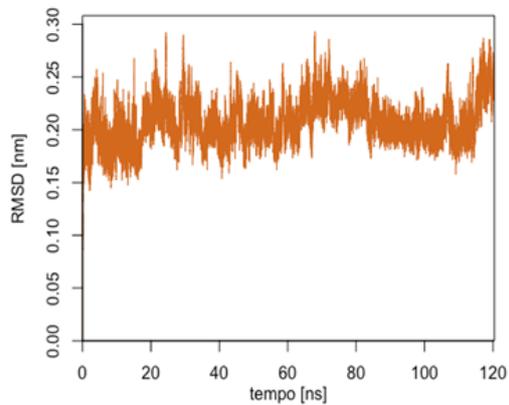
PDB ID: 3G9A



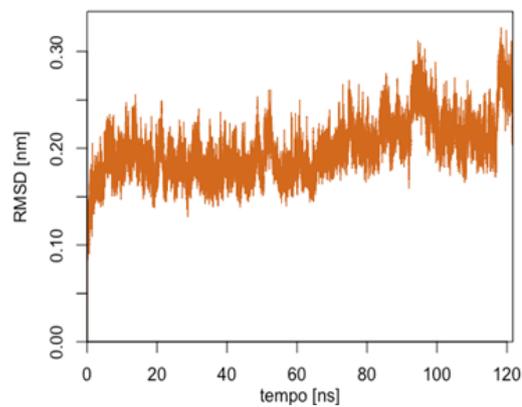
PDB ID: 5IMK



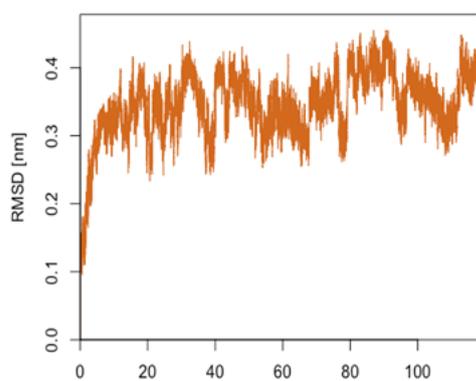
PDB ID: 4GFT



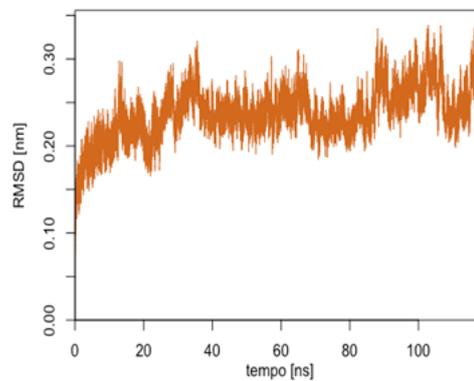
PDB ID: 4I13



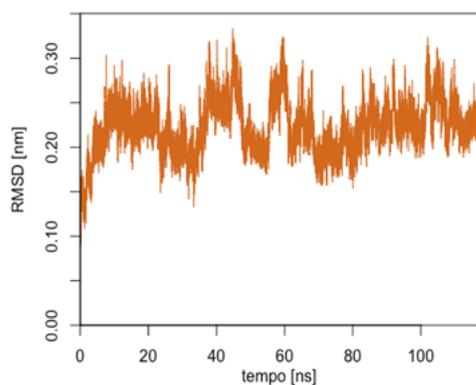
PDB ID: 4LGR



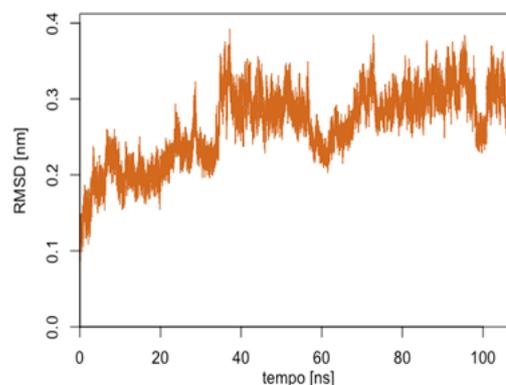
PDB ID: 6EZW



PDB ID: 5M13



PDB ID: 5MY6



ANEXO B – RMSF

Figura 29 – Gráficos de RMSF em função do tempo de simulação referente ao Backbone (C, N, O) gerados por dinâmica molecular paraca cada uma das estruturas de VHH.

