

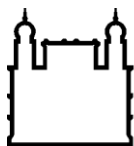
MINISTÉRIO DA SAÚDE
FUNDAÇÃO OSWALDO CRUZ
INSTITUTO OSWALDO CRUZ

Mestrado em Biologia Computacional e Sistemas

ESTUDO DA INTERAÇÃO MOLECULAR DO ANTICORPO 19CC6CG2
COM O ANTÍGENO HBSAG POR MÉTODOS COMPUTACIONAIS

ARTUR HERMANO SAMPAIO DIAS

Rio de Janeiro
Janeiro de 2019



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

Programa de Pós-Graduação em Biologia Computacional e Sistemas

ARTUR HERMANO SAMPAIO DIAS

Estudo da Interação Molecular do Anticorpo 19CC6CG2 com o Antígeno HBsAg por Métodos Computacionais

Dissertação apresentada ao Instituto Oswaldo Cruz como parte dos requisitos para obtenção do título de Mestre em Biologia Computacional e Sistemas

Orientadores: Dr. Ernesto Raul Caffarena
Dr. João Hermínio Martins da Silva

RIO DE JANEIRO

Janeiro de 2019

Dias, Artur Hermano Sampaio.

Estudo da interação molecular do anticorpo 19CC6CG2 com o antígeno HBsAg por métodos computacionais / Artur Hermano Sampaio Dias. - Rio de Janeiro, 2019.

xi, 79f f.; il.

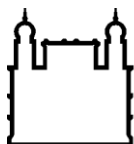
Dissertação (Mestrado) - Instituto Oswaldo Cruz, Pós-Graduação em Biologia Computacional e Sistemas, 2019.

Orientador: Ernesto Raul Caffarena.

Co-orientador: João Hermínio Martins da Silva.

Bibliografia: f. 68-79

1. Biofísica molecular. 2. Anticorpos. 3. Hepatite B. 4. Cálculo de energia livre. I. Título.



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

Programa de Pós-Graduação em Biologia Computacional e Sistemas

AUTOR: ARTUR HERMANO SAMPAIO DIAS

ESTUDO DA INTERAÇÃO MOLECULAR DO ANTICORPO 19CC6CG2 COM O ANTÍGENO HBSAG POR MÉTODOS COMPUTACIONAIS

**ORIENTADORES: Dr. Ernesto Raúl Caffarena
Dr. João Hermínio Martins da Silva**

Aprovada em: 13/02/2019

EXAMINADORES:

Dr. Floriano Paes Silva Junior – Presidente	(Instituto Oswaldo Cruz)
Dra. Livia Melo Villar	(Instituto Oswaldo Cruz)
Dr. Alessandro Silva Nascimento	(Instituto de Física de São Carlos)
Dra. Ana Carolina Ramos Guimarães	(Instituto Oswaldo Cruz)
Dra. Isabella Guedes	(Laboratório Nacional de Computação Científica)

Rio de Janeiro, 13 de fevereiro de 2019

AGRADECIMENTOS

Agradeço à minha mãe, Maria Cecília, e ao meu pai, José Dias, por terem sempre acreditado em mim e me incentivado a conquistar uma educação superior completa e de qualidade, mesmo que isso significasse passar muito tempo longe deles.

Agradeço ao meu orientador na Fiocruz do Rio de Janeiro, Dr. Ernesto Caffarena, que me acolheu com todo o suporte que eu precisei para desenvolver meu projeto de mestrado, me ensinou a fazer ciência com humildade e honestidade, me ajudou a enxergar e trabalhar meus defeitos como cientista e me ensinou a buscar a felicidade profissional.

Agradeço ao meu orientador na Fiocruz do Ceará, Dr. João Martins, quem primeiro me pôs em contato com a biofísica molecular e a biologia computacional. Obrigado por ter tido a paciência de me ensinar quando eu ainda não tinha qualquer experiência na área e por ter me dado suporte quando decidi ingressar na pós-graduação.

Agradeço à coordenadora da BCS, Dra. Carolina Guimarães, e a todos os amigos de Pós-graduação do Instituto Oswaldo Cruz – Lia Pinho, Rafael Ferreira, Alessandra Jordano, Deborah Antunes, Aline Rodrigues, Lucas Machado, Ronald Martins, Valdemir Junior, Vanessa Silva, Alexandra Almeida, Gisele Vieira, Christian Sagave, Pedro Ivo, Fernanda Cristina, Paula Cristina, Victor Barbosa, Liliane Contevile, Letícia Lima, Paulino Ribeiro, Shayenne Olsson e Esdras Matheus – os quais, de diferentes formas e em todos os momentos desses dois anos, me ajudaram e ensinaram muito. Obrigado por terem sido minha família no Rio de Janeiro.

Agradeço ao Dr. Alexandre Suman pela grande ajuda com o método de ABF ao longo do trabalho.

Agradeço ao Dr. Roberto Lins e aos seus alunos Madson de Luna, Marjorie Freire, Carlos Cruz e Danilo Coêlho pela recepção no Instituto Ageu Magalhães em Recife e pelas valiosas contribuições ao trabalho desenvolvidas na Fiocruz de Pernambuco.

Agradeço às secretárias da BCS, Rose Pani e Clarissa Lacerda, por seu trabalho sempre impecável no sentido de ajudar a mim e a todos os meus companheiros de pós-graduação em todos os assuntos acadêmicos.

Agradeço a todos os professores com os quais tive a chance de ter aulas nesta pós-graduação. Obrigado pela paciência e pelo rigor científico que me fizeram aprender e evoluir como aluno.

Agradeço à Dra. Márcia Arissawa e ao Laboratório de Tecnologia de Anticorpos Monoclonais pela concessão das sequências dos clones 19DD1AE3 e 19CC6CG2 para a realização deste projeto, iniciado em 2015 na Fiocruz Ceará.

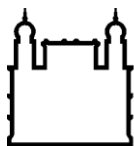
Agradeço a todos profissionais com os quais tive a oportunidade de conversar sobre meu trabalho neste período e que de alguma forma me ajudaram a compreender melhor diferentes aspectos do projeto. Agradeço especialmente ao Dr. Carlos Otávio, à Dra. Márcia Arissawa e ao Dr. Laurent Dardenne, com os quais tive a chance de aprender à época da minha qualificação discente.

Agradeço aos meus amigos, antigos e novos, que de forma direta ou indireta, me ajudaram a concluir o mestrado. Próximos ou distantes, a sua presença na minha vida me deu força e alegria para continuar, mesmo nos momentos mais difíceis.

Agradeço à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES – e também à Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro – FAPERJ – pelos auxílios financeiros concedidos.

“E você acha que dedicou todos os seus esforços à questão? Que exerceu toda a sua enorme inventividade? Que não deixou de explorar nenhuma possibilidade em sua busca [...]?”

J. K. Rowling



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

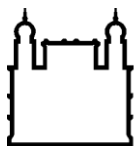
ESTUDO DA INTERAÇÃO MOLECULAR DO ANTICORPO 19CC6CG2 COM O ANTÍGENO HBSAG POR MÉTODOS COMPUTACIONAIS

RESUMO

DISSERTAÇÃO DE MESTRADO EM BIOLOGIA COMPUTACIONAL E SISTEMAS

Artur Hermano Sampaio Dias

A hepatite B é um problema de saúde global: só no ano de 2015 foram quase 900 mil óbitos decorrentes de complicações relacionadas à infecção com o vírus HBV. No Brasil, a situação também é grave: no ano de 2000 até o ano de 2015 o número de óbitos atingiu 13 mil. O diagnóstico da doença e o tratamento dos pacientes crônicos podem ser feitos através do uso de imunoglobulinas que tem afinidade pelo antígeno de superfície do vírus, HBsAg. Portanto, estratégias que busquem diminuir os custos de produção de imunoglobulinas anti-HBsAg ou aumentar sua afinidade frente a este antígeno são desejáveis em âmbitos nacional e global. O estudo de biomoléculas através de técnicas computacionais tem produzido bons resultados, capazes de orientar estudos experimentais, economizando tempo e recursos e, frequentemente, resolvendo problemas biológicos. Dentre estas técnicas computacionais, destaca-se o cálculo de energia livre. A aplicação do cálculo de energia livre a complexos anticorpo-antígeno pode fornecer informações detalhadas sobre a afinidade do anticorpo frente ao antígeno. Neste trabalho, estudamos a interação do anticorpo 19CC6CG2, desenvolvido no Laboratório de Tecnologia de Anticorpos Monoclonais de Bio-Manguinhos, com o HBsAg através de duas técnicas de cálculo de energia livre: MM-PBSA e *Adaptive Biasing Force*. Embora a primeira técnica tenha fornecido um valor de ΔG de ligação de -12 kcal/mol, a análise mais robusta através do segundo método mostrou um ΔG de dissociação de -7,6 kcal/mol. Adicionalmente, foram propostas mutações na estrutura do anticorpo visando ao aumento da sua afinidade pelo antígeno. O anticorpo mutante foi então modelado *in silico* e a sua afinidade frente ao HBsAg foi mensurada através da técnica de ABF. Resultados preliminares mostraram um valor de ΔG de ligação de -4,2 kcal/mol. As mutações na estrutura do anticorpo favoreceram a formação de ligações hidrogênio e pontes salinas intermoleculares mais estáveis nas simulações de dinâmica molecular. No entanto, simulações mais longas e/ou o aumento da dimensionalidade do espaço através de variáveis coletivas no cálculo de ABF, podem melhorar a convergência do método, tornando evidente se as mutações aqui propostas são favoráveis à afinidade do anticorpo 19CC6CG2 contra o HBsAg.



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

MOLECULAR INTERACTION BETWEEN THE 19CC6CG2 ANTIBODY AND HBSAG ANTIGEN STUDIED USING COMPUTATIONAL METHODS

ABSTRACT

MASTER DISSERTATION IN COMPUTATION AND SYSTEMS BIOLOGY

Artur Hermano Sampaio Dias

Hepatitis B is a global health problem: only in 2015 there were almost 900 thousand deaths caused by complications related to HBV infection. In Brazil, the situation is also serious: from year 2000 to 2015 the number of deaths reached 13 thousand. Diagnosis of the disease and treatment of chronic patients can be done using immunoglobulins that have affinity for the virus surface antigen, HBsAg. Therefore, strategies that seek either to decrease the production costs of anti-HBsAg immunoglobulins or to increase their affinity against this antigen are desirable at national and global levels. The study of biomolecules through computational techniques has produced good results, capable of guiding experimental studies, saving time and resources, and often solving biological problems. Among these computational techniques, the free energy calculation stands out. The application of the free energy calculation to antibody-antigen complexes can provide detailed information on the affinity of the antibody against the antigen. In this work, we studied the interaction of the 19CC6CG2 antibody, which was developed in the Bio-Manguinhos Monoclonal Antibody Technology Laboratory, against HBsAg through two free energy calculation techniques: MM-PBSA and Adaptive Biasing Force. Although the first technique provided a ΔG binding value of -12 kcal/mol, the more robust analysis through the second method showed a dissociation ΔG of -7,6 kcal/mol. In addition, mutations were proposed in the structure of the antibody aiming to increase its affinity for the antigen. The mutant antibody was then modeled *in silico* and its affinity against HBsAg was measured through the ABF technique. Preliminary results showed a binding ΔG value of -4,2 kcal/mol. Mutations in the structure of the antibody favored the formation of hydrogen bonds and intermolecular salt bridges which were more stable in molecular dynamics simulations. However, longer simulations and/or increased dimensionality of space through collective variables in the ABF calculation can improve the convergence of the method, making it evident whether the mutations proposed herein are favorable to the affinity of the 19CC6CG2 antibody against HBsAg.

RESUMO	VII
ABSTRACT	VIII
1 INTRODUÇÃO	1
1.1 Hepatite B.....	1
1.2 HBV – O Vírus da Hepatite B	2
1.3 Anticorpos.....	5
1.4 Dinâmica Molecular e Cálculo de Energia Livre	7
1.4.1 Dinâmica Molecular	7
1.4.2 Molecular Mechanics Poisson-Boltzmann Surface Area. 10	
1.4.3 Adaptive Biasing Force	12
1.5 Justificativa.....	14
2 OBJETIVOS	16
2.1 Objetivo Geral.....	16
2.2 Objetivos Específicos	16
3 MATERIAL E MÉTODOS	17
3.1 Modelagem e <i>Docking</i> Molecular do Anticorpo 19CC6CG2 e do Antígeno HBsAg	17
3.2 Dinâmica Molecular do Complexo Anticorpo-Antígeno.....	18
3.2.1 Cálculo de MM-PBSA	19
3.3 <i>Adaptive Biasing Force</i> dos Sistemas-Modelo 5A2J e 5JYL.....	20
3.4 <i>Adaptive Biasing Force</i> do Complexo Anticorpo (Nativo)-Antígeno.....	23
3.5 Engenharia do Anticorpo 19CC6CG2	24
3.6 <i>Adaptive Biasing Force</i> do Complexo Anticorpo (Mutante)-Antígeno.....	25
4 RESULTADOS E DISCUSSÃO	27
4.1 Modelagem e <i>Docking</i> Molecular do Anticorpo 19CC6CG2 e do Antígeno HBsAg	27
4.2 Dinâmica Molecular do Complexo Anticorpo-Antígeno e Cálculo de MM-PBSA	31
4.3 <i>Adaptive Biasing Force</i> dos Sistemas-Modelo 5A2J e 5JYL.....	33

4.3.1	Primeiro Sistema - 5A2J	33
4.3.2	Segundo Sistema – 5JYL	40
4.4	<i>Adaptive Biasing Force</i> do Complexo Anticorpo (Nativo)- Antígeno	48
4.5	Engenharia do Anticorpo 19CC6CG2	56
4.6	<i>Adaptive Biasing Force</i> do Complexo Anticorpo (Mutante)- Antígeno	59
5	PERSPECTIVAS	66
6	CONCLUSÕES	67
7	REFERÊNCIAS BIBLIOGRÁFICAS	68

LISTA DE SIGLAS E ABREVIATURAS

HBV	Vírus da Hepatite B
HBsAg	Antígeno de Superfície do Vírus da Hepatite B
IFN- α	Intérferon alfa
V _H	Região Variável da Cadeia Pesada do Anticorpo
V _L	Região Variável da Cadeia Leve do Anticorpo
CDR	Região Determinante de Complementaridade
MHR	Maior Região Hidrofílica
ns	Nanossegundo
ps	Picossegundo
fs	Fentossegundo
RMSD	Raiz do desvio quadrático médio
PMF	Potencial da Força Média
kcal	Quilocaloria
kJ	Quilojoule
HADDOCK	<i>High Ambiguity Driven Docking</i>
scFv	<i>single-chain Fragment variable</i>
MM-PBSA	<i>Molecular Mechanics Poisson Boltzmann Surface Area</i>
ABF	<i>Adaptive Biasing Force</i>
RMSE	Raiz do erro quadrático médio
PDB	<i>Protein Data Bank</i>
APBS	<i>Adaptive Poisson-Boltzmann Solver</i>

1 INTRODUÇÃO

1.1 Hepatite B

A hepatite B é uma infecção, causada pelo vírus HBV, que ataca o fígado e é transmitida através do contato com sangue ou espermatozoides de pessoas infectadas. Apesar de poder ser assintomática, o indivíduo infectado pelo HBV pode apresentar sintomas como febre, icterícia, dor no corpo, colúria, náuseas, vômito e perda de apetite (1).

Entre as hepatites, a do tipo B é a segunda mais letal no Brasil: de 2000 a 2015, foram mais de 13.200 óbitos decorrentes da infecção com o vírus (Figura 1.1). Desde 1999, quando se iniciou a notificação compulsória da hepatite B no Brasil, até 2016, foram notificados 212.031 casos dessa virose no País (2). O quadro é ainda mais crítico a nível global: só em 2015 cerca de 887.000 pessoas morreram devido a complicações relacionadas à hepatite B, e atualmente estima-se que 257 milhões de pessoas vivam com a infecção no mundo, segundo a Organização Mundial da Saúde (OMS) (3).

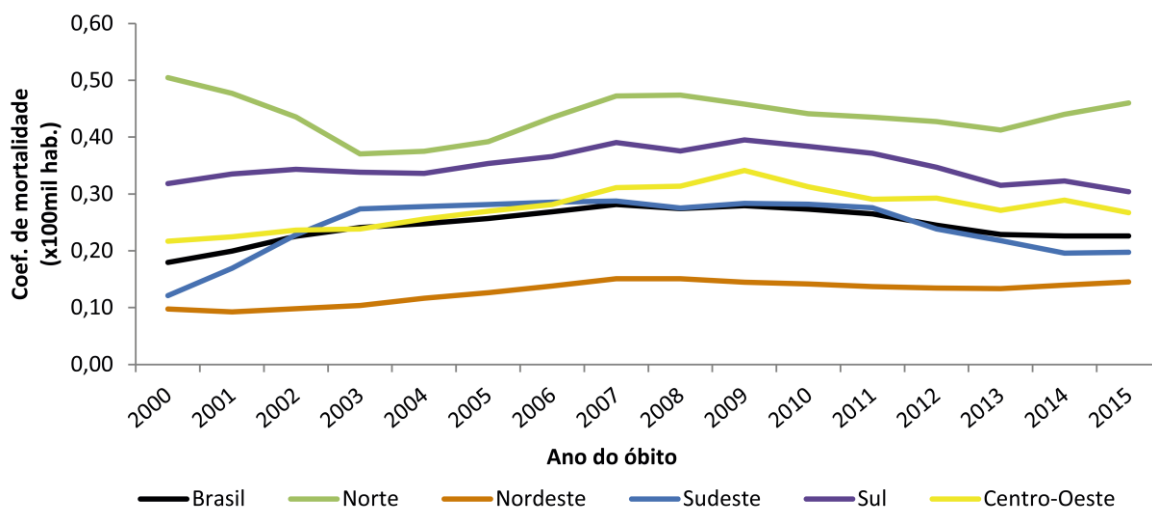


Figura 1.1 - Coeficiente de mortalidade por hepatite B segundo região de residência e ano do óbito no Brasil, de 2000 a 2015 (Fonte: Ministério da Saúde, 2016).

Cerca de 20% dos casos de hepatite B evoluem para um quadro crônico. A forma crônica dessa doença em um grande número de casos evolui para cirrose, o que aumenta a mortalidade e o risco de hepatocarcinoma. Quadros mais graves

podem levar à necessidade de um transplante hepático (4). Quando nesse estado, o paciente é tratado com imunossuppressores e, portanto, encontra-se sensível à recidiva do vírus, que ocorre em cerca de 80% dos pacientes transplantados (5). Então é necessária a administração de medicamentos que neutralizem as partículas virais ainda presentes em seu corpo (6), de forma a evitar nova proliferação do vírus.

Nesses casos, o tratamento usual é feito através da administração de interferon-alfa (IFN- α), IFN- α peguillado e análogos de nucleosídeos e de nucleotídeos (tenofovir, entecavir, lamivudina e adefovir, por exemplo). Contudo, o tratamento com interferons mostra-se eficaz em apenas um terço dos pacientes, com uma série de efeitos colaterais, tais como fadiga, anorexia, depressão, alopecia, leucopenia e plaquetopenia. Da mesma forma, os análogos de nucleotídeos e de nucleosídeos atuam apenas na fase tardia da multiplicação do HBV – a síntese de DNA –, portanto, apesar de reduzirem consideravelmente a viremia, eles negligenciam a transcrição de RNA viral e a tradução de proteínas virais, levando à baixa soroconversão e não-redução das taxas sanguíneas do antígeno de superfície do HBV - HBsAg (7,8).

Portanto, uma alternativa viável é o uso de imunoglobulinas que reconheçam e neutralizem as partículas virais ainda presentes no corpo do paciente. Imunoglobulinas que reconheçam e liguem-se às partículas virais mediante o reconhecimento do HBsAg e levem à formação de imunocomplexos e à sua subsequente remoção da corrente sanguínea são potenciais ferramentas para a diminuição dos títulos sanguíneos do HBsAg, seguida pela produção de anticorpos anti-HB, levando à soroconversão que é considerada como a recuperação da infecção viral (8).

1.2 HBV – O Vírus da Hepatite B

O HBV, também chamado de “partícula de Dane” é um vírus envelopado de aproximadamente 44 nm de diâmetro cujo capsídeo tem a forma de um icosaedro. No interior do seu capsídeo encontra-se o material genético: 3.200 pares de bases de DNA parcialmente dupla-fita em cuja extremidade 5' está ligada covalentemente a sua DNA polimerase (9).

O seu genoma contém regiões codificantes para as proteínas virais de superfície, a saber: regiões S, pré-S1 e pré-S2 (Figura 1.2). As três regiões juntas codificam a proteína de superfície L (do inglês “*large*”); mas quando apenas as regiões pré-S2 e S são transcritas e traduzidas, o resultado é a proteína M (do inglês

“medium”). Por sua vez, a região S sozinha codifica a proteína S (do inglês “small”), que é responsável por 95% das proteínas que compõem o que se chama HBsAg (10).

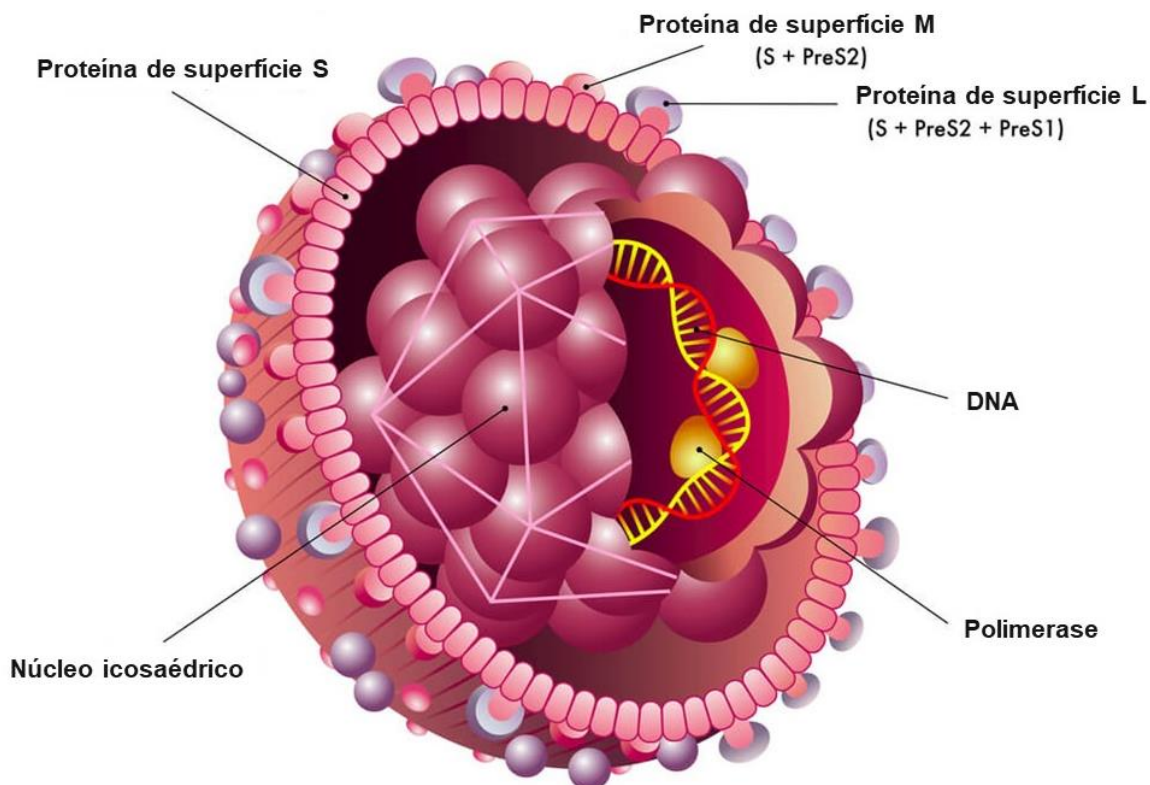


Figura 1.2 - Representação do HBV (Adaptado de <https://drug-dev.com>).

A proteína S do HBsAg é composta por quatro hélices transmembrana que estão envolvidas na integração da proteína com a membrana do retículo endoplasmático (Figura 1.3). A proteína S também é rica em voltas – regiões sem estruturação secundária – e estas regiões são responsáveis pela sua antigenicidade. Em sua estrutura, existe uma região exposta à matriz extracelular chamada MHR (do inglês “*Major Hydrophilic Region*”, Maior Região Hidrofílica), a qual compreende os resíduos 99 a 160 e é responsável pela ativação da resposta imune humoral. Nesta região se encontram os resíduos 121 a 147, que correspondem ao “determinante a”, uma região extremamente conservada do HBsAg à qual comumente se ligam anticorpos neutralizantes do HBV (11).

A proteína S media a primeira etapa do ciclo de infecção celular pelo HBV: o domínio S permite a adesão do vírion aos proteoglicanos de heparam sulfato da superfície celular, que são chamados de “receptores de baixa afinidade”. Esta etapa inicial expõe o domínio pré-S1 da proteína L para que interaja com o chamado “receptor de alta afinidade”, o polipeptídeo cotransportador de taurocolato de sódio. A

infecção celular então tem início com a penetração do vírus na célula iniciada por estas duas etapas. Portanto, anticorpos que tenham afinidade pelas regiões S ou pré-S1 podem ser utilizados para neutralizar a infecção pelo HBV (12–14).

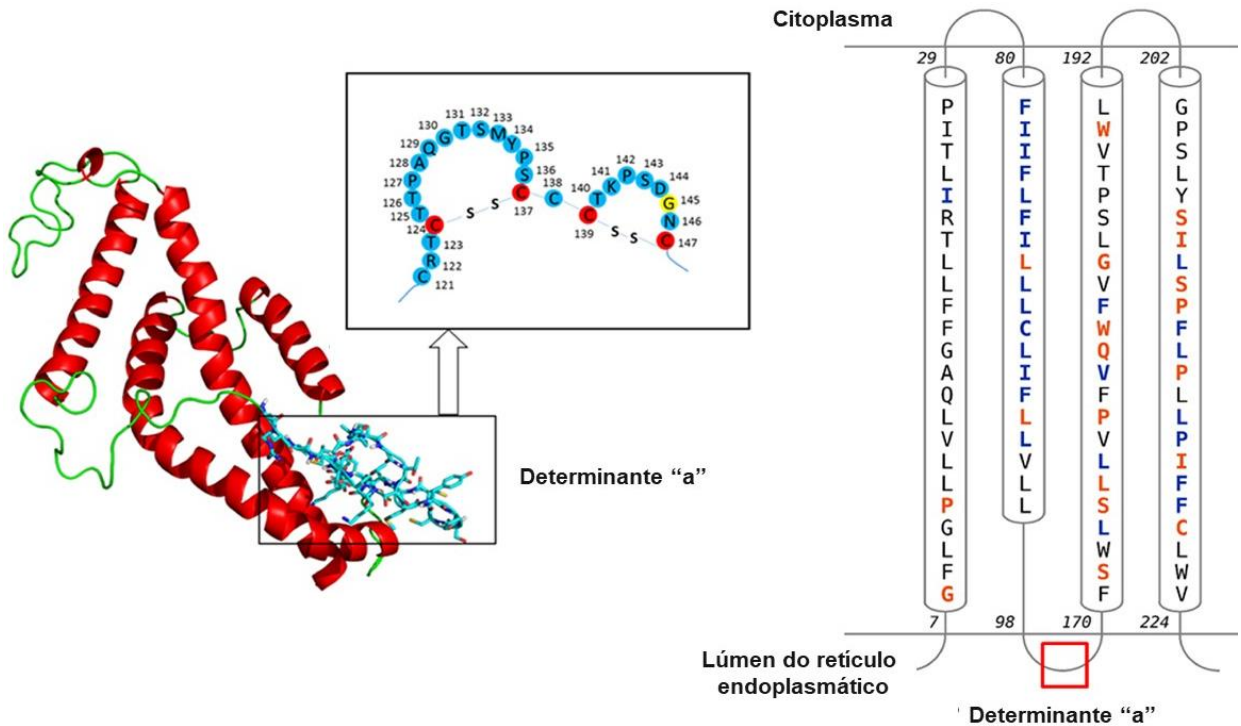


Figura 1.3 - Estrutura teórica da proteína S do HBsAg, com destaque para sua topologia transmembrana e o seu determinante "a". As regiões que atravessam a membrana do Retículo Endoplasmático (ER) estão mostradas como α -hélices vermelhas à esquerda e cilindros à direita (Fonte: adaptado de Rezaee e colaboradores, 2016).

Adicionalmente, Kil Lyong Kim e colaboradores (15) conduziram um estudo de mapeamento de epítomos de uma imunoglobulina anti-HBV comercialmente disponível na Coreia do Sul e chegaram à conclusão de que esse preparado de anticorpos exerce o efeito neutralizante através unicamente do reconhecimento da região S do HBsAg. Porém, os mesmos autores reconhecem que nos casos de falha da imunoglobulina anti-HBV (quando a administração do preparado de anticorpos comerciais não concede imunização passiva contra a infecção com HBV), pode ser necessário suplementar esse produto com anticorpos específicos contra as regiões pré-S1 e pré-S2 para aumentar a potência antiviral do tratamento.

1.3 Anticorpos

Anticorpos ou imunoglobulinas são proteínas circulantes produzidas por vertebrados em resposta à exposição a antígenos e apresentam grande diversidade e especificidade pelos seus alvos moleculares. São os principais mediadores da resposta imune humoral contra todas as classes de microrganismos. Todos os anticorpos apresentam a mesma estrutura geral básica, mas também apresentam enorme variabilidade na sua porção de ligação ao antígeno específico para o qual foram produzidos (16).

A estrutura geral de um anticorpo (Figura 1.4) é composta por duas cadeias leves idênticas e duas cadeias pesadas idênticas que se associam através de pontes de dissulfeto. Cada uma dessas cadeias apresenta uma composição de aminoácidos bastante diversa em sua porção N-terminal, chamada de região variável. A junção das regiões variáveis das cadeias leve e pesada (V_L e V_H , respectivamente) cria o sítio de ligação do anticorpo ao antígeno, sítio esse que é definido por resíduos de aminoácidos específicos em Regiões Determinantes de Complementaridade (CDRs, do inglês “*Complementarity Determining Region*”). Já as porções C-terminais de cada cadeia pesada interagem com outras moléculas e células do sistema imune através da porção Fc do anticorpo, mediando assim a função efetora dessas biomoléculas (16).

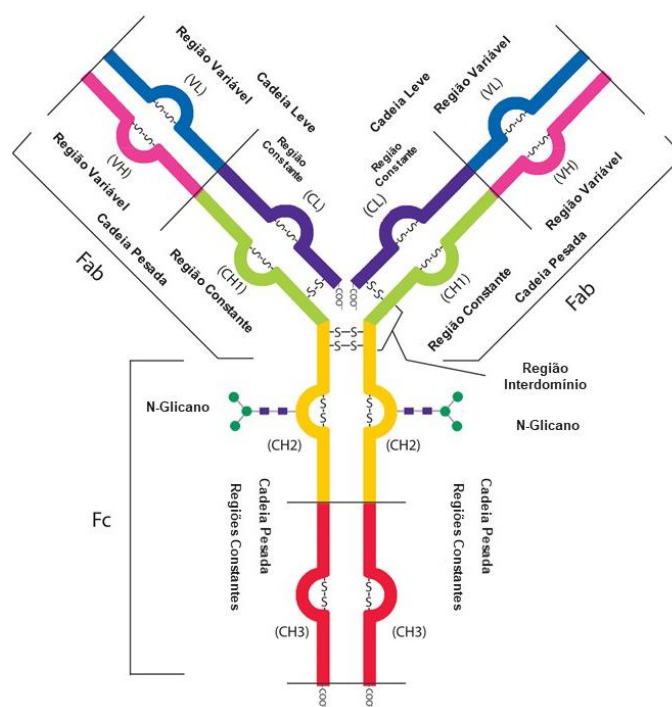


Figura 1.4 - Estrutura básica de um anticorpo (Fonte: adaptado de Ferreira, 2017).

A alta especificidade e afinidade de ligação dos anticorpos os qualifica como insumos biotecnológicos e terapêuticos que podem ser usados no tratamento de diversas doenças (17). De fato, as imunoglobulinas são a classe de agentes terapêuticos que mais vêm crescendo nas últimas décadas (18). A estrutura básica dos anticorpos é conservada na natureza, e a sua especificidade e afinidade são definidas principalmente por seis voltas em suas cadeias variáveis da sua porção Fab – onde localizam-se os CDRs –, portanto é possível engenheirar essas regiões dos anticorpos sem, no entanto, desestabilizar a sua estrutura (19).

Desde a descoberta da tecnologia de produção de anticorpos monoclonais em 1975 através do uso de hibridomas (20), várias estratégias têm sido utilizadas para engenheirar anticorpos, de forma a melhorar propriedades desejáveis dessas moléculas ou neutralizar propriedades desvantajosas (21). Assim, tem sido comum a engenharia de anticorpos com o intuito de aumentar a sua afinidade pelo antígeno, diminuir a sua resposta imunogênica e aumentar sua estabilidade, por exemplo (22). Segundo estudos recentes, 40% dos anticorpos monoclonais em testes clínicos são ou estão sendo engenheirados (23).

Uma das áreas de maior interesse dentro da engenharia de anticorpos é o controle fino da afinidade dessas moléculas: através de modificações (mutações) em resíduos de aminoácidos em sua estrutura – comumente nos CDRs – é possível aumentar ou reduzir sua afinidade pelo antígeno-alvo. O aumento da afinidade de um anticorpo usado como biofármaco melhora o seu efeito terapêutico e pode levar à redução da dose efetiva de administração, conseqüentemente reduzindo efeitos colaterais e o custo do tratamento (24).

Atualmente a profilaxia pós-exposição com imunoglobulinas anti-HBV usa a HBIg - Imunoglobulina Humana Específica Anti-Hepatite B - um anticorpo que se liga ao HBsAg e é comercialmente conhecido pelos nomes Hepagam B, HyperHEP S/D ou Nabi-HB (4). Porém seu uso ainda levanta questões sobre a sua baixa atividade específica e baixa disponibilidade (25). Além disso, a HBIg tem alto custo, pois é derivada do sangue de indivíduos hiperimunizados com a vacina anti-HBV, o que a torna inacessível para países e regiões com poucos recursos financeiros; além disso, seu uso levanta questões de biossegurança, por se tratar de um produto originado de sangue humano que pode ser contaminado com patógenos humanos. Uma alternativa viável à HBIg seria o uso de anticorpos monoclonais anti-HBV geneticamente

engenheirados com alta afinidade pelo determinante “a” capazes de neutralizar as partículas virais no corpo do paciente (8,26).

O Laboratório de Tecnologia de Anticorpos Monoclonais da Fundação Bio-Manguinhos desenvolveu diversos clones de anticorpos monoclonais contra o HBsAg através da imunização subcutânea de camundongos BALB/C com partículas do antígeno, obtidas através do plasma de pacientes infectados com o vírus. Entre eles, um dos que apresentou maiores níveis de reatividade em testes de imunoadsorção enzimática (ELISA em inglês, para *Ezyme-Linked Immunosorbent Assay*) contra o HBsAg nativo, alquilado ou reduzido e alquilado foi o clone 19CC6CG2 (27). Portanto, esse clone seria um candidato interessante à engenharia de anticorpos visando à criação de uma alternativa terapêutica à HBIg.

Atualmente diversas técnicas de engenharia de anticorpos com o propósito de aumentar sua afinidade estão disponíveis, incluindo métodos *in vivo* que usam o sistema imune de mamíferos e também manipulação gênica (28,29). Contudo, essas técnicas muitas vezes são custosas e demandam muito tempo (30). Portanto, métodos que sejam capazes de gerar resultados acurados com baixo custo e em tempos menores são desejáveis, e neste contexto se encaixam as técnicas *in silico* de simulação da interação entre biomoléculas.

Neste contexto, a crescente produção científica acerca das interações anticorpos-antígenos, assim como o progressivo entendimento sobre a estrutura e dinâmica dessas proteínas pode ajudar no planejamento e desenvolvimento de anticorpos engenheirados. Adicionalmente, o entendimento das propriedades físico-químicas dos resíduos de aminoácidos da interface anticorpo-antígeno e o uso de cálculos computacionais que monitorem a formação e a extinção de interações intermoleculares mediante o reconhecimento do antígeno pelo anticorpo são essenciais para a engenharia de anticorpos moderna (31).

1.4 Dinâmica Molecular e Cálculo de Energia Livre

1.4.1 Dinâmica Molecular

Segundo a União Internacional de Química Pura e Aplicada (IUPAC, do inglês *International Union of Pure and Applied Chemistry*), a Dinâmica Molecular é “um procedimento de simulação que consiste na computação do movimento dos átomos

em uma molécula ou de átomos individuais ou moléculas em sólidos, líquidos e gases, de acordo com as leis de movimento de Newton”(32). Sabe-se que tanto a sequência de aminoácidos de uma proteína quanto sua estrutura tridimensional são diretamente relacionados à sua função; então, sendo a estrutura proteica flexível em menor ou maior grau, pode-se dizer que a compreensão do movimento dos átomos de uma proteína ao longo do tempo é fundamental para entender sua função. Portanto, a Dinâmica Molecular é uma técnica computacional robusta usada para o cálculo e previsão do comportamento temporal de moléculas em um meio (33).

Para que uma simulação computacional represente de forma acurada o comportamento biofísico de um sistema, é necessário trabalhar com funções matemáticas que descrevam a sua energia potencial de forma a reproduzir os fenômenos e interações intra e intermoleculares desse sistema. O conjunto de funções de energia potencial (Figura 1.5) e parâmetros calibrados a partir de dados experimentais e cálculos quânticos chama-se campo de força (34).

$$E_{total} = \underbrace{\sum_{\text{ligações}} K_r (r - r_{eq})^2 + \sum_{\text{ângulos}} K_\theta (\theta - \theta_{eq})^2 + \sum_{\text{diedros}} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)]}_{\text{Ligado}} + \underbrace{\sum_{i < j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{\epsilon R_{ij}} \right]}_{\text{Não-Ligado}}$$

Figura 1.5 – Representação matemática de um Campo de Força Clássico
(Fonte: Adaptado de Durrant e McCammon, 2011).

O primeiro termo do lado direito da equação da Figura 1.5 é um potencial ligado que trata do comprimento das ligações químicas da molécula simulada como um sistema massa-mola que obedece à Lei de Hooke. De forma semelhante, o segundo termo trata os ângulos de ligação da molécula, também através da Lei de Hooke. O terceiro termo também é um potencial ligado, mas trata da rotação de ligações químicas que separam dois planos: esse ângulo torcional é computado através de um potencial descrito por funções harmônicas. Os dois últimos termos tratam de potenciais não-ligados, ou seja, potenciais referentes a interações de átomos que não estabelecem entre si ligações químicas. O quarto termo é equivalente ao potencial de

Lennard-Jones, que descreve as interações de Van der Waals – interações atrativas a curto alcance e repulsivas a curtíssimo alcance. Finalmente, o último termo faz referência ao potencial de Coulomb, que descreve as interações eletrostáticas, de longo alcance (35).

$$F_i = \frac{d^2 x_i}{dt^2} m_i \cong \frac{\Delta v_i}{\Delta t} m_i = \langle a_i \rangle m_i \quad (1)$$

Quando se define o campo de força a ser usado em uma simulação, é possível calcular as forças que atuam sobre cada átomo “*i*” do sistema, calculando-se o gradiente da energia potencial, fornecida pelo campo de força, em relação às posições desses átomos. De posse da força e da massa de cada átomo, é possível obter diretamente a sua aceleração. A partir desta, integrando-se as equações de movimento, pode-se obter as velocidades, cuja integral, por sua vez, proporciona as novas posições dos átomos. Com as novas posições e velocidades de cada átomo, obtêm-se as energias potencial e cinética do sistema. Assim, de posse da massa, *m*, e dos vetores posição e aceleração – **x** e **a**, respectivamente – da partícula *i*, é possível, através da resolução numérica da Segunda Lei de Newton (Equação 1) sucessivas vezes ao longo do tempo de simulação, obter um perfil da energia do sistema e descrever a trajetória de todos os seus átomos (35).

A engenharia de proteínas usa métodos computacionais como a dinâmica molecular para estudar detalhes atomísticos de sistemas proteicos, como fármaco-receptor ou antígeno-anticorpo. Diversos métodos computacionais têm sido empregados com êxito no estudo molecular de problemas biológicos (36–43). Essa estratégia tem a vantagem de fornecer resultados bons, com baixo custo operacional e frequentemente de forma mais rápida que métodos experimentais. Dentre os métodos computacionais, destaca-se o cálculo de energia livre (ΔG) de ligação, que tem sido empregado para modulação da afinidade de biomoléculas, como demonstrado por Paloni e Chong, por exemplo (43,44).

A energia livre de ligação é a variável termodinâmica que reflete a tendência de duas moléculas se associarem espontaneamente (45). Ao longo dos anos, diversos métodos computacionais para calcular o ΔG de ligação foram propostos, incluindo Perturbação de Energia Livre (FEP) (46), Dinâmica Molecular Guiada (47) e *Umbrella Sampling* (48). No entanto, o cálculo do ΔG da interação entre duas proteínas continua sendo uma tarefa árdua (49) devido principalmente à complexidade inerente da

associação proteica advinda do grande número de interações possíveis entre cadeias laterais na interface das duas proteínas (45). Uma das alternativas computacionalmente simples para o cálculo da energia livre de ligação é o método de *Molecular Mechanics Poisson-Boltzmann Surface Area* (50).

1.4.2 *Molecular Mechanics Poisson-Boltzmann Surface Area*

Um dos métodos computacionais mais frequentemente empregados no cálculo da energia livre de ligação – e conseqüentemente da afinidade – entre duas moléculas é o método de *Molecular Mechanics Poisson-Boltzmann Surface Area* (MM-PBSA) (50). Esse método já foi aplicado em dezenas de estudos de sistemas biológicos, com bons resultados para a avaliação de modos de ligação de fármacos aos seus receptores, para a determinação da estabilidade estrutural de macromoléculas e para a predição de *hot-spots* no estudo da afinidade entre moléculas (51–57).

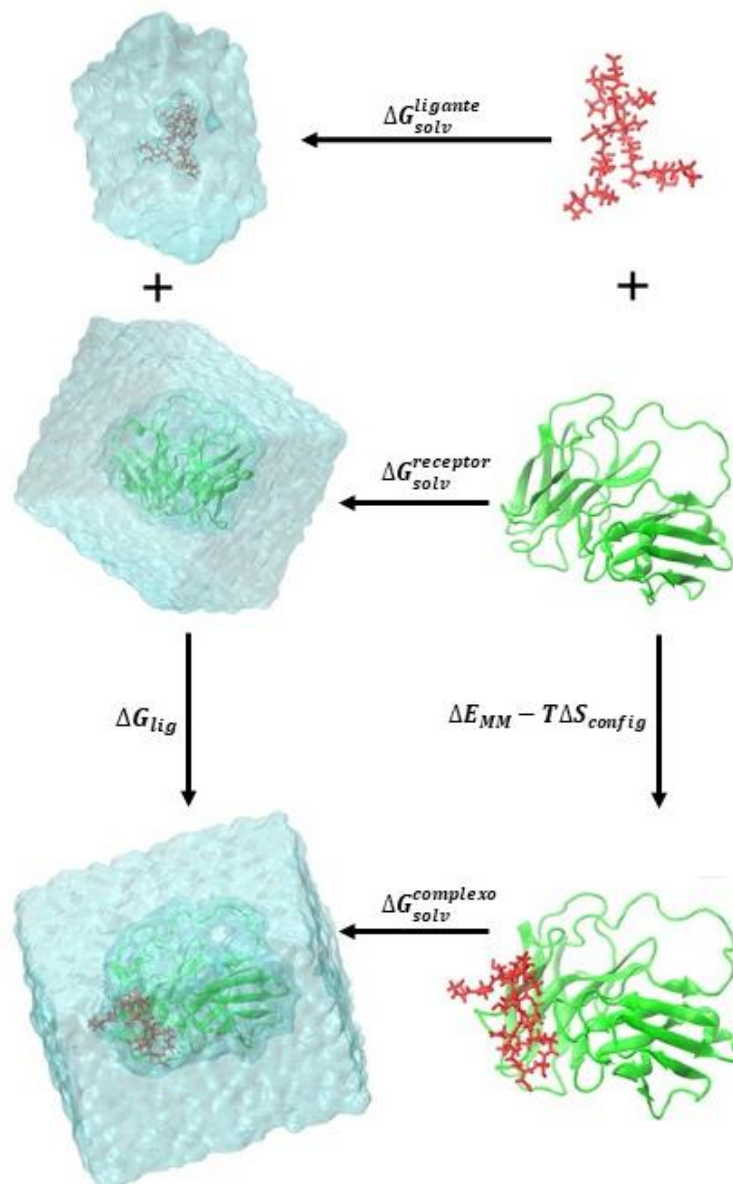


Figura 1.6 - Ciclo termodinâmico para a associação entre uma proteína e uma molécula pequena.

Como a energia livre de Gibbs é uma função de estado do sistema, é possível descrever sua variação através de um ciclo termodinâmico (Figura 1.6), cujo $\Delta G_{total} = 0$. Assim, através da decomposição da energia livre e de uma aritmética simples entre as “contribuições energéticas” individuais, é possível estimar a energia livre de ligação entre duas moléculas, sejam elas um fármaco e seu receptor proteico ou um antígeno e um anticorpo. O valor de ΔG de ligação em solução então assume o valor:

$$\Delta G_{lig} = \Delta E_{MM} + (\Delta G_{solv}^{complexo} - \Delta G_{solv}^{receptor} - \Delta G_{solv}^{ligante}) - T\Delta S_{config} \quad (2)$$

A componente ΔE_{MM} da Equação 2 faz referência à contribuição energética da associação das moléculas na fase gasosa e é obtida através do campo de força usado na simulação da dinâmica molecular. Nela estão inclusas as energias referentes aos comprimentos, ângulos e torções de ligações, assim como as interações eletrostáticas e de Van der Waals. Como os processos biológicos devem acontecer em um solvente, deve haver um termo que contabilize a energia de solvatação do sistema – ΔG_{solv} –, o qual pode ser dividido em suas contribuições polares e apolares. Sua componente polar é calculada como a energia de transferência de uma partícula carregada entre dois meios – vácuo e solvente – com constantes dielétricas distintas, sendo o solvente modelado de forma implícita. Esse cálculo é feito via solução das diferenças finitas da equação de Poisson-Boltzmann (Equação 3), na qual ∇ representa o operador gradiente, ε representa a permissividade do meio, ϕ é o potencial eletrostático do meio, k é o fator de Debye-Hückel e ρ_f é soma algébrica das cargas fixas do meio.

$$-\nabla \cdot \varepsilon(x) \nabla \phi(x) + \bar{k}^2(x) \phi(x) = 4\pi \rho_f(x) \quad (3)$$

Já a componente apolar do termo energético de solvatação advém da formação de uma cavidade no solvente que possa comportar o sistema (soluto). Esta componente é calculada como a soma entre energias “desfavoráveis” pela ruptura de interações solvente-solvente e energias “favoráveis” pela formação de interações soluto-solvente. Por fim, a componente $-T\Delta S_{config}$ mede o gasto entrópico do sistema mediante associação entre as duas moléculas e pode ser calculada através de análises de modos normais ou quase-harmônicas (51,52,54).

O método MM-PBSA é versátil, fácil de ser aplicado e muito pouco custoso computacionalmente quando comparado com técnicas mais robustas de cálculo de energia livre, como Perturbação de Energia Livre ou Integração Termodinâmica. Diversos estudos de biologia estrutural e planejamento de fármacos obtiveram sucesso com a aplicação do método de MM-PBSA (53,58–63). Portanto, esse método mostra-se adequado à análise da afinidade e da interação molecular entre o anticorpo 19CC6CG2 e o HBsAg. Contudo, existem métodos mais robustos capazes de gerar resultados mais acurados às expensas de um maior custo computacional. Um desses métodos é o *Adaptive Biasing Force* (64).

1.4.3 *Adaptive Biasing Force*

Em simulações de dinâmica molecular, frequentemente é necessário reduzir os graus de liberdade de um sistema físico, simplificando-os em um parâmetro cuja distribuição estatística possa ser analisada individualmente ou usada para definir potenciais úteis para controlar a dinâmica do sistema. Este parâmetro - denotado por ξ - é chamado de coordenada de reação, parâmetro de reação ou variável coletiva (65).

Este conceito é importante porque, segundo a Teoria da Integração Termodinâmica, é possível obter o perfil de energia livre de um processo físico a partir da força média que age ao longo da coordenada de reação (66). Tratando-se do processo de reconhecimento e associação entre duas proteínas, durante o qual as duas moléculas percorrem uma distância em meio ao solvente para então se ligarem, uma variável coletiva apropriada para estudar este processo seria a distância entre os centros de massa das proteínas, por exemplo (Figura 1.7).

Uma característica comum dos métodos modernos de cálculo de energia livre é a dependência da amostragem da variável em estudo. Contudo, simulações computacionais de processos biofísicos em que o sistema passa por grandes mudanças conformacionais frequentemente atingem mínimos locais de energia potencial, os quais agem como “armadilhas cinéticas”, pois mantêm o sistema “preso” em configurações não-relevantes e dificultam a exploração do processo, reduzindo a amostragem da variável (67).

Assim, foram criados métodos para o cálculo de energia livre que contornam esse problema, de forma a uniformizar a amostragem, aumentando a exploração estocástica da coordenada de reação. Métodos como *replica-exchange*, *metadinâmica*, *simulated annealing* e *adaptive biasing force* baseiam-se em coletar amostras da variável de estudo ao longo da coordenada de reação até mesmo em eventos raros – configurações de alta energia potencial – do espaço configuracional, possibilitando assim o cálculo de ΔG .

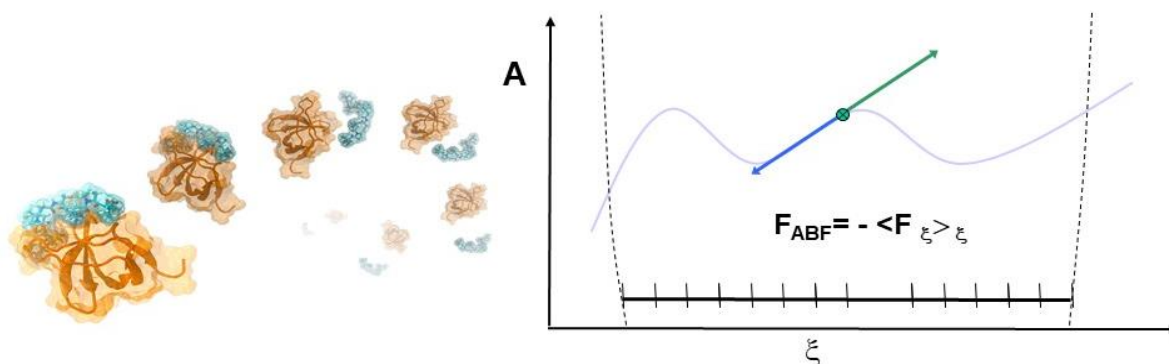


Figura 1.7 - Representação da evolução da coordenada de reação mediante a aplicação da técnica de ABF (Fonte: adaptado de Gumbart e colaboradores (2017) e Cheng (2009)).

O método *Adaptive Biasing Force* (ABF) tem se mostrado robusto e eficiente para o cálculo de energia livre e, desde a sua criação, tem sido empregado na resolução de problemas de interesse químico e biológico, tais como proteínas mecânicas e de transporte, reconhecimento e associação em sistemas proteína-ligante ou proteína-proteína (68). A técnica ABF utiliza uma versão de viés adaptativo da amostragem aumentada, na qual uma força, \mathbf{F}_{ABF} , de valor igual à força média, $\langle \mathbf{F}_\xi \rangle$, mas de sentido contrário, é aplicada ao longo da coordenada da reação e é ajustada continuamente a cada $\delta\xi$, de forma a anular a força resultante sobre o sistema (Figura 1.7), que então evolui por meio de efeitos de difusão sobre uma superfície de energia plana ao longo da coordenada de reação (58).

A robustez, versatilidade e precisão do método ABF sinalizaram como ferramenta para o cálculo do ΔG de ligação neste projeto de análise da afinidade e interação entre o anticorpo 19CC6CG2 e o antígeno HBsAg.

1.5 Justificativa

Os métodos experimentais para medição do ΔG de ligação (Ressonância Plasmônica de Superfície, por exemplo) são caros, demorados e laboriosos; portanto é importante que sejam estudadas técnicas rápidas, menos custosas e capazes de fornecer resultados tão precisos quanto o fazem as técnicas experimentais. Nesse contexto as técnicas computacionais de cálculo de energia livre, iniciadas desde o final do século XX e aprimoradas desde então, se apresentam como uma alternativa viável e atraente. Além disso, as técnicas computacionais que fazem uso da dinâmica

molecular permitem obter detalhamento a nível atômico sobre os processos que ocorrem mediante o reconhecimento de um antígeno pelo anticorpo.

Neste trabalho, os resultados obtidos com cálculo computacional de energia livre permitirão entender em detalhes a interação anticorpo-HBsAg. Com esse conhecimento, a engenharia de anticorpos, baseada em mutações sítio-dirigidas, poderá ser acurada, visando aumentar a afinidade do anticorpo 19CC6CG2 pelo antígeno de superfície do vírus da hepatite B.

2 OBJETIVOS

2.1 Objetivo Geral

Explorar o perfil de energia livre de ligação do anticorpo 19CC6CG2 ao antígeno HBsAg e propor mutações sítio-dirigidas sobre este anticorpo que aumentem a sua afinidade pelo antígeno.

2.2 Objetivos Específicos

1. Propor um possível modo de ligação do anticorpo 19CC6CG2 ao antígeno HBsAg;
2. Simular o comportamento dinâmico do complexo anticorpo-antígeno para explorar a superfície de energia livre de ligação do anticorpo ao antígeno;
3. Propor mutações na estrutura do anticorpo que aumentem sua afinidade pelo antígeno;
4. Comparar os perfis de energia livre de ligação dos anticorpos selvagem e mutante frente ao HBsAg.

3 MATERIAL E MÉTODOS

3.1 Modelagem e *Docking* Molecular do Anticorpo 19CC6CG2 e do Antígeno HBsAg

O anticorpo 19CC6CG2 neste trabalho foi estudado no formato de fragmento variável de cadeia única (scFv, do inglês *single chain Fragment variable*), com um *linker* de 8 resíduos – GGGSGGGG – entre as cadeias variáveis leve e pesada. O uso de scFv é vantajoso porque retém a afinidade e a especificidade do anticorpo parental (69–74) e, em um estudo computacional, o número reduzido de átomos, comparado à estrutura completa do anticorpo parental, torna os cálculos computacionais muito mais rápidos sem, no entanto, comprometer a acurácia da simulação.

O HBsAg neste trabalho foi estudado inicialmente sob duas formas: a primeira consistiu da volta que compreende o segmento desde o resíduo 135 até o resíduo 151, correspondente à região do determinante *a*, à qual comumente se ligam anticorpos neutralizantes do HBV; já a segunda forma corresponde à proteína S completa do antígeno.

Ambas estruturas, do anticorpo 19CC6CG2-scFv e do antígeno HBsAg, foram previamente obtidas a partir do grupo de Modelagem, Simulação e Evolução *in silico* de Biomoléculas da Fiocruz-CE.

Os *dockings* anticorpo-antígeno foram feitos através do servidor HADDOCK versão 2.2 (75,76), o qual recebe como entrada as duas estruturas proteicas e um conjunto de dados sobre os resíduos ativos – os resíduos que sabidamente fazem contato direto dentro do complexo – e passivos – os que potencialmente fazem contato direto – de cada molécula. Tais informações são convertidas pelo servidor em Restrições Ambíguas de Interação (RAI), que direcionam todo o processo de *docking*: para cada resíduo ativo, uma RAI é definida entre este resíduo e todos os resíduos ativos e passivos da outra proteína, e todos os cálculos do processo são desenvolvidos a partir de então (75).

Os resíduos ativos do scFv foram selecionados com base na análise estrutural dos CDRs: os resíduos de aminoácidos que estavam mais expostos em cada volta destas regiões foram selecionados como ativos. Para as duas formas do antígeno, um raciocínio semelhante foi aplicado: os resíduos de aminoácidos do determinante *a* mais expostos ao solvente foram escolhidos como ativos. Em cada molécula, foram

definidos como “passivos” todos os resíduos que estivessem dentro de um raio de corte de 6,5 Å de qualquer resíduo ativo (75).

O protocolo de *docking* seguiu três passos básicos: a) randomização das orientações e minimização de corpo rígido das duas proteínas; b) três etapas de *simulated annealing* semirrígido; c) e um refinamento final com solvente explícito (77).

Ao final do protocolo de *docking*, todas as estruturas são agrupadas pela comparação par-a-par dos valores das raízes de seus desvios quadráticos médios (RMSD, do inglês *Root Mean Square Deviation*) em relação aos esqueletos peptídicos na interface, com raio de corte de 1 Å. Os grupos são então analisados e classificados de acordo com o seu “HADDOCK score”, que é uma soma ponderada entre as energias de Van der Waals, eletrostáticas, energias de desolvatação e de violação de restrições junto com uma média da área superficial não-exposta (77).

A avaliação de cada complexo gerado nos *dockings* contra as duas formas do HBsAg foi feita com base nos seguintes parâmetros:

- i. HADDOCK score (descrito acima);
- ii. *Cluster size* – o número de estruturas finais que compõem cada grupo;
- iii. RMSD em relação à estrutura de menor energia;
- iv. Posição do CDR H3 em relação ao antígeno (dado que os CDRs H3 são os que mais contribuem energeticamente para o reconhecimento do antígeno, portanto espera-se uma posição favorável desta volta na interface com o HBsAg).

3.2 Dinâmica Molecular do Complexo Anticorpo-Antígeno

O *docking* gerou diversos complexos com diferentes modos de ligação ao HBsAg. Destes, o complexo de maior qualidade (de acordo com os parâmetros descritos acima, na seção 3.1) foi usado como entrada para uma simulação de dinâmica molecular.

A simulação foi realizada através do pacote de programas Gromacs (78) versão 5.1.2, e as ligações e interações químicas foram tratadas conforme o campo de força CHARMM22/CMAP (79). O algoritmo *Particle Mesh Ewald* (PME) foi aplicado no tratamento de todas as interações eletrostáticas (80). As interações de curto alcance foram descritas mediante a aplicação do potencial de Lennard-Jones. As simulações foram realizadas no ensemble estatístico isotérmico-isobárico (NPT) (81). O complexo scFv-HBsAg foi simulado no centro de uma caixa cúbica de aresta 70 Å preenchida

com moléculas do modelo de água explícito TIP3P (82) e com dois íons cloreto de forma a neutralizar a carga líquida do sistema. A simulação foi feita em três estágios:

- minimização de energia;
- aquecimento;
- e equilíbrio, seguida pela dinâmica de produção.

No primeiro estágio, a minimização de energia do sistema foi feita gradualmente em 100.000 passos do algoritmo Gradiente Conjugado seguidos de mais 100.000 passos do algoritmo *Steepest Descent*, ambos com gradiente de tolerância de 0,01 kJ.mol⁻¹.nm⁻¹. No segundo estágio, o sistema foi aquecido de 0 a 310 K ao longo de 1000 picossegundos. Nesta fase, cada átomo pesado do sistema teve sobre si uma restrição de posição com constante de força de 1.000 kJ mol⁻¹.nm⁻². Durante a fase de equilíbrio, que durou 4 nanossegundos (ns), a força harmônica de restrição sobre os átomos pesados diminuiu progressivamente até zero, para então dar início à dinâmica molecular de produção, que durou 170 ns.

Durante a dinâmica molecular, o algoritmo de Verlet foi utilizado para a integração das equações de movimento, que utiliza as posições e acelerações dos átomos no instante t e as suas posições no tempo $t - \delta t$ para calcular suas posições no tempo $t + \delta t$ (83). O passo de integração usado foi de 2 fs. As interações químicas ligadas envolvendo ligações covalentes foram tratadas através do algoritmo LINCS (84). A temperatura do sistema foi mantida constante a 310 K pela aplicação do termostato V-rescale, e a pressão do sistema foi mantida constante a 1 atm pelo acoplamento do barostato de Berendsen à caixa de simulação. A simulação de dinâmica molecular foi feita usando condições periódicas de contorno.

3.2.1 Cálculo de MM-PBSA

O cálculo de energia livre foi inicialmente feito através da técnica de MM-PBSA, seguindo a “abordagem de trajetória única” (51). Portanto, as contribuições energéticas individuais de cada membro do sistema (scFv isolado, HBsAg isolado e complexo scFv-HBsAg) foram extraídas a partir de uma única trajetória de dinâmica molecular.

A trajetória de produção da dinâmica molecular foi monitorada em relação aos valores de RMSD dos carbonos-alfa. A partir do momento em que se observou menos variação deste parâmetro, pode-se extrair configurações do sistema durante a trajetória. Essas configurações foram agrupadas em 54 *clusters* com um raio de corte

de RMSD = 1 Å. De cada um desses grupos foi extraída a configuração mais representativa, portanto 54 configurações da simulação foram analisadas quanto às contribuições energéticas (ΔE_{MM} e ΔG_{solv}) de cada membro do sistema – scFv, HBsAg e complexo scFv-HBsAg.

Os valores das componentes ligadas e não-ligadas (ΔE_{MM}) foram extraídos do cálculo do próprio campo de força CHARMM22/CMAP usado na dinâmica molecular. Os termos polar e apolar de solvatação (ΔG_{solv}) foram calculados através do programa APBS (85), considerando uma constante de proporcionalidade de tensão superficial igual a 0,0001 kcal.mol⁻¹.Å⁻² e constantes dielétricas dos solutos, do solvente e do vácuo iguais a 1,5, 80 e 1, respectivamente.

Finalmente, a cada membro do sistema foi designado um valor médio, “G”, representativo de sua energia: a média da soma de ΔE_{MM} e ΔG_{solv} em todas as configurações do componente do sistema. Como a variação de energia livre é calculada como a diferença entre os estágios final e inicial de um processo, o ΔG de ligação foi então estimado segundo a equação 4.

$$\Delta G_{lig} = G_{complexo} - (G_{anticorpo} + G_{antígeno}) \quad (4)$$

3.3 Adaptive Biasing Force dos Sistemas-Modelo 5A2J e 5JYL

A aplicação de uma técnica intrinsecamente mais robusta que MM-PBSA sugeriu um estudo inicial com sistemas semelhantes ao complexo scFv-HBsAg que tivessem valores de ΔG de ligação determinados experimentalmente. Portanto, inicialmente, a partir do *Brookhaven Protein Data Bank* (86) foram obtidas:

a) a estrutura depositada sobre o PDB ID 5A2J (87), referente ao fragmento de anticorpo SM3 complexado com o peptídeo $_{APDTRP}$ (Figura 3.1);

b) e a estrutura depositada sobre o PDB ID 5JYL (88), referente ao fragmento de anticorpo TSP7 complexado com a P-caderina humana MEC1 (Figura 3.1).

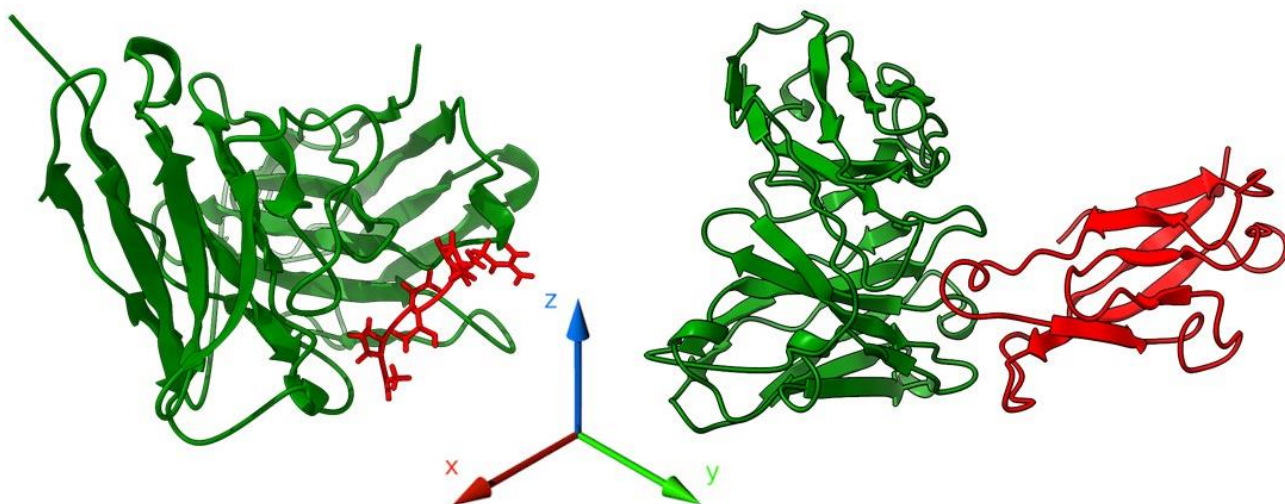


Figura 3.1 - Anticorpo SM3 complexado com peptídeo APDTRP (esquerda) e anticorpo TSP7 complexado com P-caderina MEC1 (direita).

Em seguida, ambos os complexos foram manipulados individualmente com a ferramenta *editconf* do pacote de programas Gromacs (89) de forma que os eixos de inércia das duas moléculas fossem alinhados com relação aos eixos cartesianos (Figura 3.1). Moléculas de água TIP3P (82) foram usadas para solvatar ambos os sistemas, 5A2J e 5JYL, em caixas de dimensões $78 \times 56 \times 70 \text{ \AA}^3$ e $127 \times 103 \times 93 \text{ \AA}^3$, respectivamente. Adicionalmente, ambos os sistemas tiveram suas cargas elétricas líquidas neutralizadas mediante a adição de íons (2 Na^+ à caixa de 5A2J e 6 Na^+ à caixa de 5JYL). A solvatação e a neutralização foram feitas através do programa VMD versão 1.9.2 (90).

Todas as simulações foram realizadas utilizando o programa NAMD versão 2.9 (91). Novamente, o algoritmo de integração das equações de movimento utilizado foi o de Verlet (83), com um passo de integração de 2 fs, em conjunto com o campo de força CHARMM36 (92,93). As interações ligadas envolvendo ligações covalentes foram ajustadas pelo algoritmo SHAKE (94), e as interações eletrostáticas não-ligadas foram tratadas através do método PME (80). As interações de curto alcance foram descritas mediante a aplicação do potencial de Lennard-Jones. As simulações foram realizadas no *ensemble* estatístico NPT (81). A temperatura de cada sistema foi mantida constante em 300 K mediante aplicação do termostato de Langevin. Da mesma forma, a pressão dos sistemas foi mantida constante a 1 atm através do pistão de Langevin acoplado a cada caixa de simulação (95). Foram aplicadas condições periódicas de contorno a ambos os sistemas. O processo se deu em três etapas:

- i. Minimização;
- ii. Equilibração;

iii. ABF.

A energia potencial de cada sistema foi previamente minimizada através de 2.500 passos do algoritmo Gradiente Conjugado.

Como preparação para dinâmica de equilíbrio, foi aplicado um potencial harmônico de restrição de constante de força igual a $1 \text{ kcal.mol}^{-1}.\text{Å}^{-2}$ sobre os átomos da cadeia principal de cada proteína em ambos os sistemas de forma a restringir seus movimentos e evitar deformações estruturais da proteína. Em seguida, os sistemas evoluíram à temperatura de 300 K por 5 ns.

Após a etapa de equilíbrio, as restrições adicionadas sobre os átomos da cadeia principal do peptídeo APDTRP e da P-caderina MEC1 foram removidas. Em seguida, foi aplicado o método ABF a ambos os sistemas, de forma que durante 5 ns as duas proteínas em cada complexo progressivamente afastaram-se uma da outra cerca de 16 Å dentro da caixa de simulação, a 300 K e 1 atm. Neste trabalho, o cálculo dos potenciais da força média foi feito seguindo uma estratégia de estratificação. Portanto, essa simulação inicial foi utilizada para determinar as configurações iniciais de cada etapa (ou janela) de estratificação.

Assim, ambas as trajetórias foram estratificadas em 8 janelas de 2 Å ; e em cada janela os complexos proteicos evoluíram sequencialmente até atingir 15 ns. O monitoramento do comportamento do potencial de força média foi feito a cada 5 ns. Cada dinâmica de produção de 5 ns, nas mesmas condições supracitadas, foi realizada com o intuito de aumentar a amostragem e verificar a convergência do método. A estratégia de estratificação baseia-se no fato de que o tempo para alcançar convergência cresce com o quadrado da tamanho da coordenada de reação (68). Portanto, trajetórias em intervalos espaciais menores requerem menos tempo de simulação para atingir convergência.

Ao longo da coordenada de reação, definida pela variável coletiva 'd' - a distância entre as duas proteínas -, a força média foi calculada em intervalos $\delta\xi = 0,1 \text{ Å}$. Os valores da força média dos intervalos de cada janela foram então concatenados e integrados subsequentemente para gerar o perfil total do potencial da força média ao longo da coordenada de reação para cada tempo de simulação (5 ns, 10 ns e 15 ns por janela). Adicionalmente, foi feita a análise das energias de interação entre as duas proteínas de cada complexo ao longo da dissociação, além do cálculo do RMSD do antígeno e da amostragem ao longo da simulação mais longa em cada sistema.

O erro quadrático médio (RMSE) por janela foi o parâmetro usado para averiguar convergência, considerando-se atingida se o RMSE entre duas simulações

consecutivas fosse menor ou igual a 1 kcal/mol. O RMSE por janela é calculado de acordo com a equação 4 abaixo.

$$RMSE = \sqrt{\frac{\sum_{m=1}^n (a_m - b_m)^2}{n}} \quad (4)$$

Nesta equação, a_m e b_m são os valores do potencial da força média para um $\delta\xi$ específico (m) entre duas simulações consecutivas (5 ns e 10 ns, por exemplo) e n é o número de pontos na janela específica (o número de $\delta\xi$ contidos em uma janela).

Finalmente, com o auxílio do programa VMD versão 1.9.2 (90) as interações não-ligadas do sistema 5JYL foram analisadas ao longo da coordenada de reação, com destaque para a formação, extinção e ocupância de ligações hidrogênio e pontes salinas (96).

3.4 *Adaptive Biasing Force* do Complexo Anticorpo (Nativo)-Antígeno

De volta ao sistema formado pelo anticorpo 19CC6CG2 ligado ao HBsAg obtido ao fim do item 3.1, a energia livre de ligação desse complexo foi adicionalmente explorada através de um segundo método – ABF. O complexo scFv-HBsAg (Figura 3.2) foi submetido exatamente ao mesmo protocolo descrito no item 3.3, com algumas diferenças, a saber:

- i. A caixa de simulação teve dimensões de $125 \times 112 \times 123 \text{ \AA}^3$;
- ii. O RMSD do HBsAg e as energias de interação intermoleculares foram calculados na simulação inicial de 5 ns;
- iii. A força iônica neste sistema foi modelada de forma que a concentração de $\text{NaCl}_{(aq)}$ fosse igual a 0,15 M, com carga total líquida nula – portanto foram adicionados 153 íons Na^+ e 153 íons cloreto à caixa de simulação;
- iv. Durante a terceira fase, ABF, a força média sobre a coordenada de reação foi calculada a cada $\delta\xi = 0,05 \text{ \AA}$.

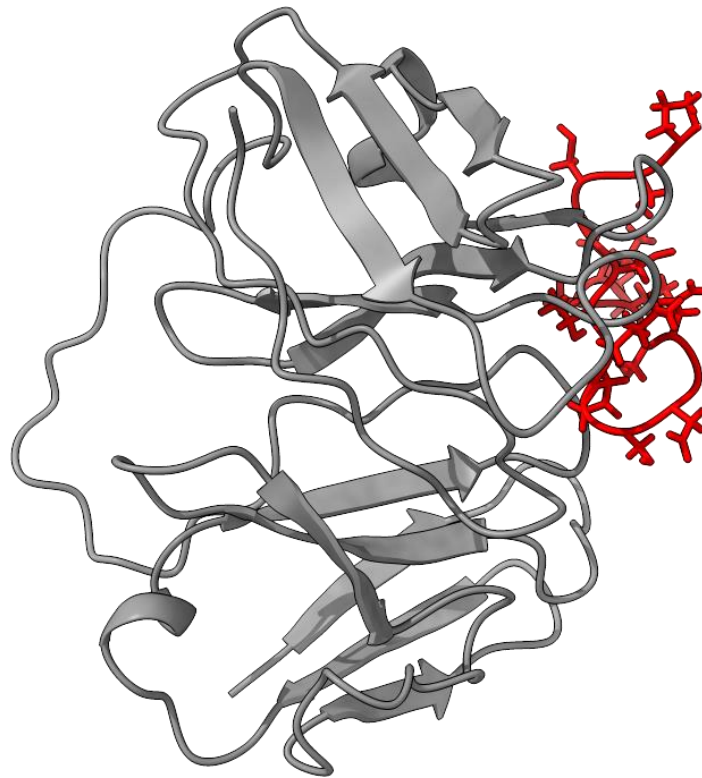


Figura 3.2 - Anticorpo 19CC6CG2 (cinza) complexado com parte do HBsAg (vermelho).

3.5 Engenharia do Anticorpo 19CC6CG2

A partir da dinâmica molecular descrita no item 3.2, foi possível realizar mutações na estrutura do scFv que aumentariam a sua afinidade pelo HBsAg. Essa engenharia do anticorpo foi feita em RosettaScripts, uma linguagem de programação que provê acesso às funcionalidades de modelagem molecular da plataforma Rosetta (97).

Uma vez que os valores de RMSD da simulação do item 3.2 se estabilizaram e as configurações dessa dinâmica molecular foram agrupadas conforme o item 3.2.1, foram obtidos 54 *clusters*. Então foi extraída a configuração mais representativa do maior *cluster* da dinâmica molecular, e, após ter as moléculas de água e íons removidos do sistema, esta pose do complexo scFv-HBsAg foi usada como entrada no protocolo de engenharia de anticorpos do RosettaScripts.

O protocolo realizou 100.000 ciclos de modelagem do scFv e *docking* com o antígeno. Cada ciclo foi independente dos demais e pode ser descrito como um processo estocástico no qual são feitas mutações nos resíduos da interface com o HBsAg, e o mutante então obtido é dockado com o antígeno. Finalmente, o complexo

resultante é avaliado quanto a dois principais parâmetros – complementaridade de interface e energia de ligação entre as duas proteínas (98) – estimados segundo a função de escore Talaris2014 (99).

Todos os resíduos de aminoácidos do scFv dentro de um raio de corte de 8 Å a partir do HBsAg foram suscetíveis a mutações aleatórias, exceto aqueles resíduos que mantiveram ligações hidrogênio com o antígeno com ocupância superior a 50% ou se mantiveram a 4 Å de distância ou menos em 50% das configurações da simulação. Além disso, cada ciclo do protocolo permitiu flexibilidade ao esqueleto do scFv mediante as mudanças de cadeias laterais resultantes das mutações – o chamado movimento de *backrub* (100,101).

Adicionalmente, para comparação do scFv nativo com os mutantes gerados, foi realizado um único ciclo do protocolo com as funções de modelagem e *docking* desativadas. Assim, este ciclo forneceu apenas a avaliação de energia de ligação e complementaridade de interface do anticorpo nativo segundo a função de escore Talaris2014.

3.6 Adaptive Biasing Force do Complexo Anticorpo (Mutante)-Antígeno

O ΔG de dissociação entre o anticorpo mutante criado no item 3.5 e o HBsAg foi calculado através do método ABF. O complexo scFv_(mut)-HBsAg (Figura 3.3) foi submetido exatamente ao mesmo protocolo descrito nos itens 3.3 e 3.4, com as particularidades a seguir:

- i. A caixa de simulação desse sistema teve dimensões de 122x119x118 Å³;
- ii. O RMSD do HBsAg e as energias de interação intermoleculares foram calculados na simulação inicial de 5 ns;
- iii. Novamente a força iônica neste sistema foi simulada com $[\text{NaCl}_{(\text{aq})}] = 0,15 \text{ M}$, com carga total líquida neutra – portanto foram adicionados 160 íons Na⁺ e 161 íons cloreto à caixa de simulação;
- iv. Igualmente ao item 3.4, durante a simulação de ABF, a força média sobre a coordenada de reação foi calculada a cada $\delta\xi = 0,05 \text{ Å}$.

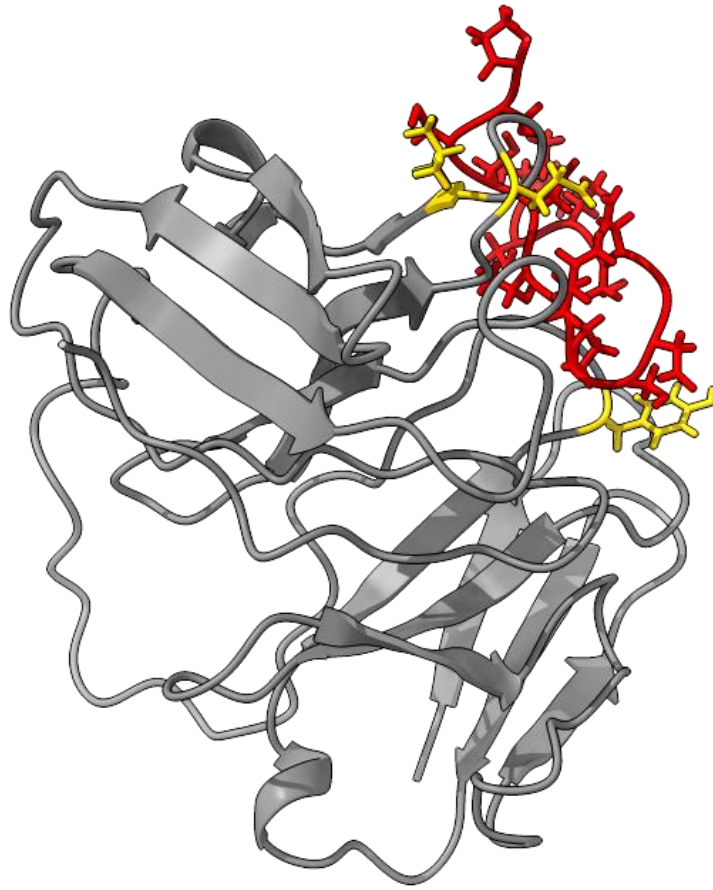


Figura 3.3 - Anticorpo mutante (cinza) em complexo com parte do HBsAg (vermelho). Os resíduos mutados do anticorpo estão destacados em amarelo.

4 RESULTADOS E DISCUSSÃO

4.1 Modelagem e *Docking* Molecular do Anticorpo 19CC6CG2 e do Antígeno HBsAg

As estruturas do anticorpo 19CC6CG2 e do HBsAg (Figuras 4.1 e 4.2) foram previamente obtidas a partir do grupo de Modelagem, Simulação e Evolução *in silico* de Biomoléculas da Fiocruz-CE .

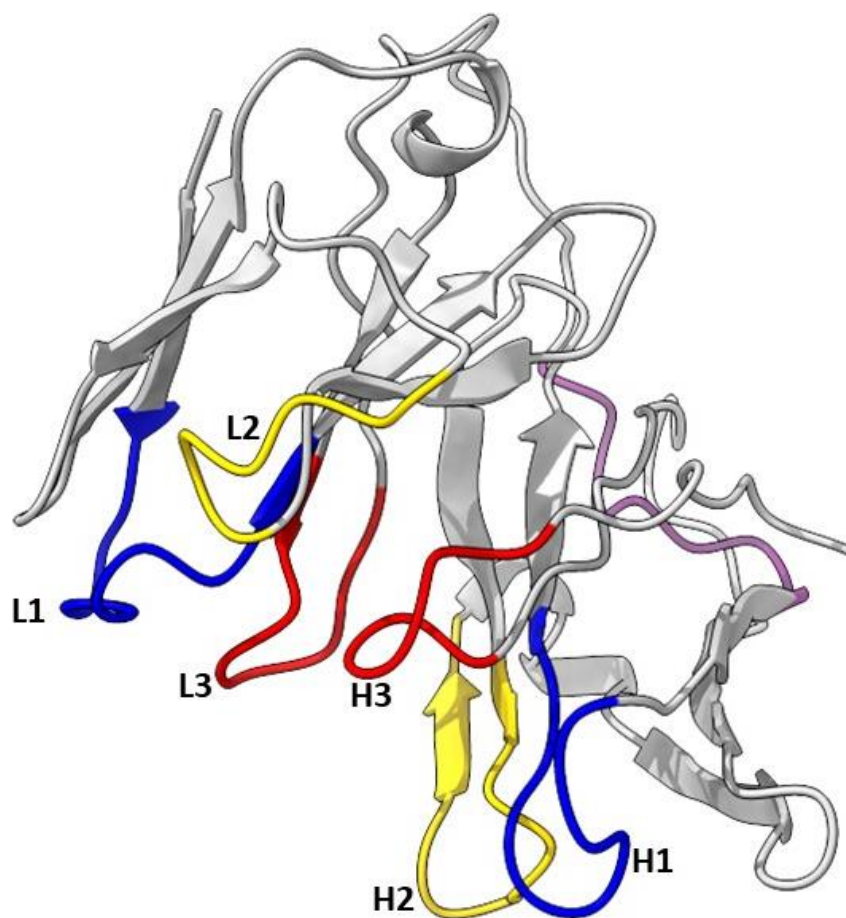


Figura 4.1 - Modelo do scFv 19CC6CG2 com os CDRs destacados em suas cores padrão. Em violeta está destacado o peptídeo que conecta as cadeias V_H e V_L . L1, L2, L3 e H1, H2, H3 são, respectivamente, os CDRs das regiões V_L e V_H .

O HBsAg neste trabalho foi inicialmente avaliado sob duas formas: a volta que compreende o segmento desde o resíduo 135 até o resíduo 151 do antígeno e a proteína S completa (Figura 4.2).

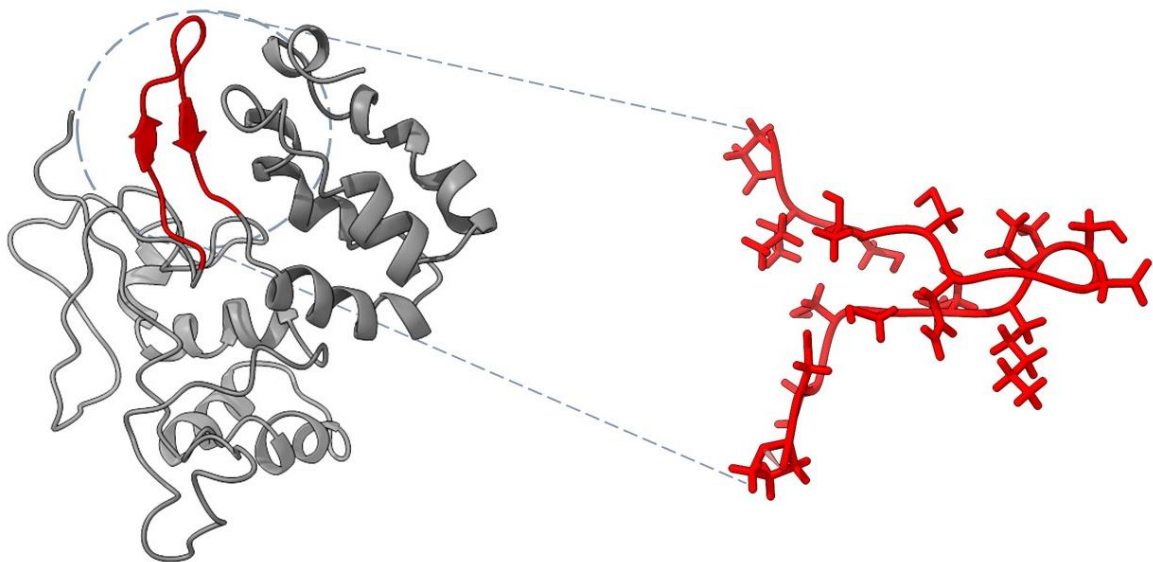


Figura 4.2 - Modelos do HBsAg utilizados neste trabalho: a proteína S completa está à esquerda, e a volta que compreende os resíduos 135 a 151 está destacada à direita.

Para o *docking* no servidor HADDOCK, foram considerados como ativos os resíduos 139 a 147 das duas formas do HBsAg, pois é esta a região na qual 90% dos anticorpos neutralizantes do HBV se ligam para exercer sua atividade (11). Apesar desta estatística vantajosa, é pertinente notar que os pesquisadores que criaram o clone 19CC6CG2 não deixaram claro em qual região específica da proteína S o anticorpo se liga, apenas sugeriram que poderia ser em uma região glicosilada do antígeno (dados não-publicados).

Portanto, é razoável supor que a escolha dos resíduos ativos do HBsAg feita neste trabalho seja apenas uma predição com acurácia incerta. Caso o anticorpo de fato ligue-se ao determinante *a*, então as análises desenvolvidas nesta dissertação podem ser consideradas precisas; no entanto, se o anticorpo se liga a outras regiões do antígeno, como as regiões dos determinantes *d/y* ou *w/r*, então novos estudos deverão ser feitos partindo-se de novas estruturas. Shin e colaboradores (2007) exemplificaram como o reconhecimento do HBsAg por um anticorpo neutralizante depende fortemente da sua especificidade por um ou mais determinantes, sendo os anticorpos direcionados ao determinante *a* capazes de reconhecer todas as formas do antígeno (25).

Para o *docking*, os resíduos mais expostos ao solvente em cada CDR do scFv foram considerados como ativos, exceto para o CDR H3: todos os resíduos deste CDR foram considerados ativos. Esta escolha particular quanto ao CDR H3 leva em

consideração que este é, usualmente, a região mais variável, menos conservada e mais flexível entre todas as seis regiões determinantes de complementaridade (102,103). O protocolo de *docking*, portanto, deve levar em consideração cada um dos resíduos deste CDR para direcionar o cálculo do modo de ligação entre as duas proteínas.

O HADDOCK agrupou as soluções de cada *docking* em “clusters” e os classificou segundo sua função interna de pontuação, o “HADDOCK score”, que quantifica as energias de Van der Waals, eletrostáticas, energias de desolvatação e de violação de restrições junto com uma média da área superficial não-exposta (77). Os melhores “clusters” resultantes de cada rodada estão detalhados de acordo com a forma do HBsAg usada em cada *docking*.

Tabela 1 - Avaliação dos melhores resultados dos dockings scFv-HBsAg

Parâmetro	Volta 135-151 Proteína S	
HADDOCK score	-73,9 ± 1,4	-128,1 ± 1,8
Cluster size	14 (de 200)	47 (de 200)
RMSD em relação a estrutura de menor energia (Å)	1,8 ± 0,1	0,9 ± 0,1
Menor distância do CDR H3 ao HBsAg (Å)	3,9	5,9
Número de contatos do CDR H3 com o HBsAg	31	16

O algoritmo de *docking* do HADDOCK mostrou ser acurado na resolução de diversos problemas biológicos (104–108), inclusive na análise do modo de ligação entre fragmentos de anticorpos e seus antígenos (109,110). Contudo, já mostrou soluções inacuradas em experimentos de *re-docking*, portanto nem sempre a sua função de pontuação é confiável, sendo necessário avaliar outros aspectos da solução proposta, como o número de contatos com o CDR H3, por exemplo (103).

O RMSD entre as melhores soluções de cada *docking* foi de 0,809 Å, e, considerando apenas as duas estruturas do HBsAg, o RMSD entre elas foi de 1,119

Â; portanto, entende-se que as duas soluções não divergiram estruturalmente. Este é um resultado positivo e surpreendente, visto que o HADDOCK leva em consideração a flexibilidade do esqueleto peptídico em seu cálculo (103), poderiam ser esperadas soluções com conformações bem distintas nos dois *dockings*.

Apesar da pontuação do *docking* com a proteína S completa ter sido melhor, a solução do *docking* com a volta 135-151 exibe uma proximidade maior entre o HBsAg e o CDR H3, além de ter praticamente o dobro de contatos entre estas regiões. O algoritmo de *docking* do HADDOCK tem sido reportado como muito acurado para a predição do modo de ligação de V_H (onde localiza-se o CDR H3), mas ineficaz na predição da pose de V_L (103). Portanto é razoável que neste trabalho se dê prioridade à pose obtida no *docking* com a volta 135-151 (Figura 4.3), na qual os resíduos do HBsAg fazem mais contatos com o CDR H3; e, além da pose favorável desta região, os CDRs H2 e L3 nesta pose, reportados por frequentemente manterem, respectivamente, pontes salinas e interações polares com o antígeno (19), também estão bem orientados com relação ao HBsAg.

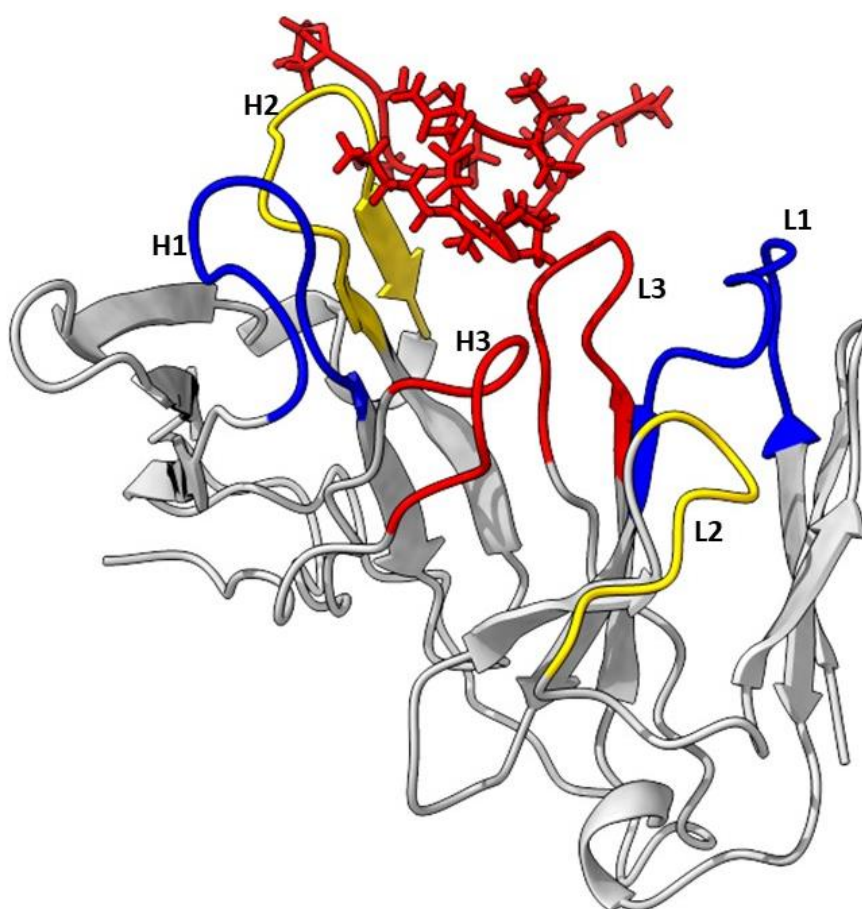


Figura 4.3 - Modo de ligação entre o scFv e o HBsAg proposto pelo HADDOCK.

4.2 Dinâmica Molecular do Complexo Anticorpo-Antígeno e Cálculo de MM-PBSA

A dinâmica molecular do complexo scFv-HBsAg foi feita a partir da estrutura gerada no *docking* (Figura 4.3) e objetivando o cálculo de energia livre de ligação do anticorpo ao antígeno (Figura 1.6), descrito amplamente na literatura científica (31,51,52,54,59,61,112–114), inclusive com aplicações a sistemas anticorpo-antígeno (43,114).

Os valores de RMSD dos carbonos-alfa do complexo foram monitorados durante a dinâmica de produção (Figura 4.4). A partir de 70.000 ps, quando este parâmetro apresentou menos variação, todas as configurações do complexo foram então reunidas em 54 grupos cujo raio de corte foi $\text{RMSD} = 1 \text{ \AA}$, e de cada grupo foi extraída a configuração mais representativa do complexo para o cálculo de MM-PBSA. Este é um procedimento importante para que o cálculo de energia livre seja feito sobre conformações de um sistema termalizado (51).

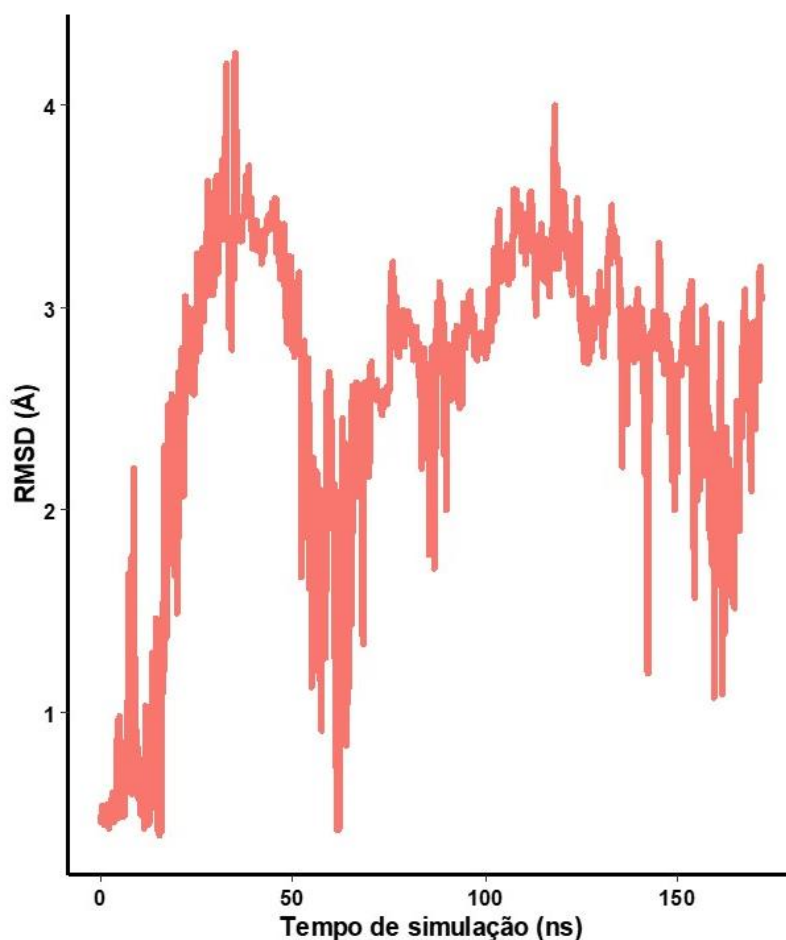


Figura 4.4 - RMSD do complexo scFv-HBsAg ao longo da dinâmica molecular.

A “abordagem de trajetória única” é vantajosa por sua execução simples e por usualmente ter resultados acurados e precisos em comparação com a “abordagem de trajetória tripla” (52), na qual são analisadas trajetórias independentes do complexo, do antígeno e do anticorpo. A principal razão para esta vantagem é o fato de que, apesar de negligenciar as mudanças conformacionais das proteínas em seu estado não-ligado, o cálculo da abordagem de trajetória única conta com o cancelamento das contribuições energéticas intramoleculares – os termos ligados do campo de força. Esta abordagem é comum na análise de sistemas em que não se espera grandes mudanças conformacionais nos componentes do sistema (51).

Finalmente, o valor obtido para a energia livre de ligação entre as duas proteínas foi $\Delta G_{LIG} = -12,39$ kcal/mol. Este valor está em consonância com valores obtidos experimentalmente para a afinidade de anticorpos e fragmentos de anticorpos (13,43,54,74,88,114). Este resultado, portanto, sugere que a associação entre o anticorpo 19CC6CG2 e o HBsAg é entalpicamente dirigida, visto que o termo entrópico “ $-T\Delta S_{config}$ ” foi negligenciado no cálculo e, ainda assim, o resultado obtido tem a mesma ordem de grandeza de resultados experimentais com sistemas semelhantes (Tabela 2).

Tabela 2 - Exemplos de ΔG de ligação para complexos proteicos.

Complexo anticorpo-antígeno	ΔG_{LIG} experimental	Referência
<i>Fragmento Fab – Proteína L</i>	-7 kcal/mol	(43)
<i>Complexo Ras-Raf</i>	-9,6 kcal/mol	(54)
<i>scFv – 11-deoxycortisol</i>	-13,78 kcal/mol*	(74)
<i>scFv TSP7 – P-caderina humana MEC1</i>	-10,9 kcal/mol*	(88)
<i>scFv TSP11 - P-caderina humana MEC1</i>	-12,06 kcal/mol*	(88)
<i>scFv TSP5 - P-caderina humana MEC1</i>	-11,94 kcal/mol*	(88)
<i>mAb 17.1.41 - HBsAg</i>	-12,42 kcal/mol*	(13)
<i>mAb 19.79.5 - HBsAg</i>	-12,67 kcal/mol*	(13)

*Valores deduzidos a partir do parâmetro K_d ou K_a obtidos experimentalmente.

Caso a contribuição energética da entropia fosse estimada, o seria através de análise de modos normais, a qual frequentemente gera erros sistemáticos por não contabilizar contribuições não-harmônicas (53,61) e muitas vezes superestimar a

variação entrópica de um processo biofísico (51). Embora a opção de computar o termo entrópico através de análise quase-harmônica seja atraente por contabilizar implicitamente movimentos não-harmônicos (53), esta frequentemente requer tempos de simulação muito grandes para atingir convergência, mesmo em simulações de sistemas pequenos (54). No entanto, métodos computacionais mais robustos como Integração Termodinâmica, por exemplo, podem ser capazes de fazer estimativas ainda mais acuradas em análises biofísicas (115), levando em consideração as mudanças entrópicas advindas da interação entre biomoléculas, como é demonstrado nas seções seguintes.

4.3 Adaptive Biasing Force dos Sistemas-Modelo 5A2J e 5JYL

Uma vez obtidas as estruturas 5A2J e 5JYL, foi necessário ter valores de referência para as energias livres de ligação de ambos, a fim de comparar os resultados fornecidos por ABF com resultados obtidos através de outras metodologias, experimentais ou não. A aplicação de uma metodologia a sistemas simples antes de usá-la para obter resultados que demandem mais confiabilidade é usual, serve como calibração do método computacional, reduz a possibilidade de erros, além de tornar o usuário familiarizado com a técnica e, portanto, capacitado a interpretar melhor os resultados obtidos.

As energias livres de dissociação para esses dois complexos foram obtidas de formas distintas: para o complexo 5A2J, o ΔG de ligação entre o fragmento de anticorpo SM3 e o peptídeo APDTRP não havia sido estudado usando técnicas experimentais, então para este trabalho o ΔG de ligação entre as duas moléculas foi estimado através do servidor PISA (116). O valor apresentado pelo servidor foi de -0,9 kcal/mol, apesar de este não ser um valor acurado. Para o complexo 5JYL, uma pesquisa bibliográfica mostrou que o ΔG de ligação entre o fragmento de anticorpo TSP7 e a P-caderina MEC1, obtido experimentalmente através da técnica de Ressonância Plasmônica de Superfície, é de -10,9 kcal/mol (88).

4.3.1 Primeiro Sistema - 5A2J

O potencial da força média ao longo da coordenada de reação, ξ , que representa a distância, em Å, entre os centros de massa do fragmento de anticorpo

SM3 e do peptídeo $_{APDTRP}$ evoluiu durante três diferentes tempos de simulação (Figura 4.5). Quando o sistema evoluiu por 5 ns, o ΔG de ligação obtido foi de -4,4 kcal/mol. Esta primeira medida serve como referência para estimar a convergência do método. Logo, permitindo-se que o sistema evoluísse até 10 ns, foi observado que este valor mudou para -2,8 kcal/mol com RMSE associado de 3,169 kcal/mol. E, finalmente, com 15 ns de simulação por janela, obtém-se o valor de -1,5 kcal/mol, com RMSE de 1,954 kcal/mol (Figura 4.6).

Este resultado revela não apenas a evolução em direção a maior acurácia (tendo como referência o valor obtido pelo PISA), como também ilustra o refinamento da exploração da coordenada de reação conforme aumenta a amostragem: diversos picos e vales no caminho de reação são revelados à medida que o tempo de simulação por janela aumenta. Por exemplo, o intervalo de ξ (18 Å, 19 Å) contém um declive de energia que só é revelado com amostragem suficientemente grande, a partir de 10 ns de simulação. Este mesmo intervalo corresponde a um aumento da atração entre as duas moléculas, descrito pela queda nos valores de energia de interação e de RMSD do peptídeo (Figuras 4.7 e 4.8).

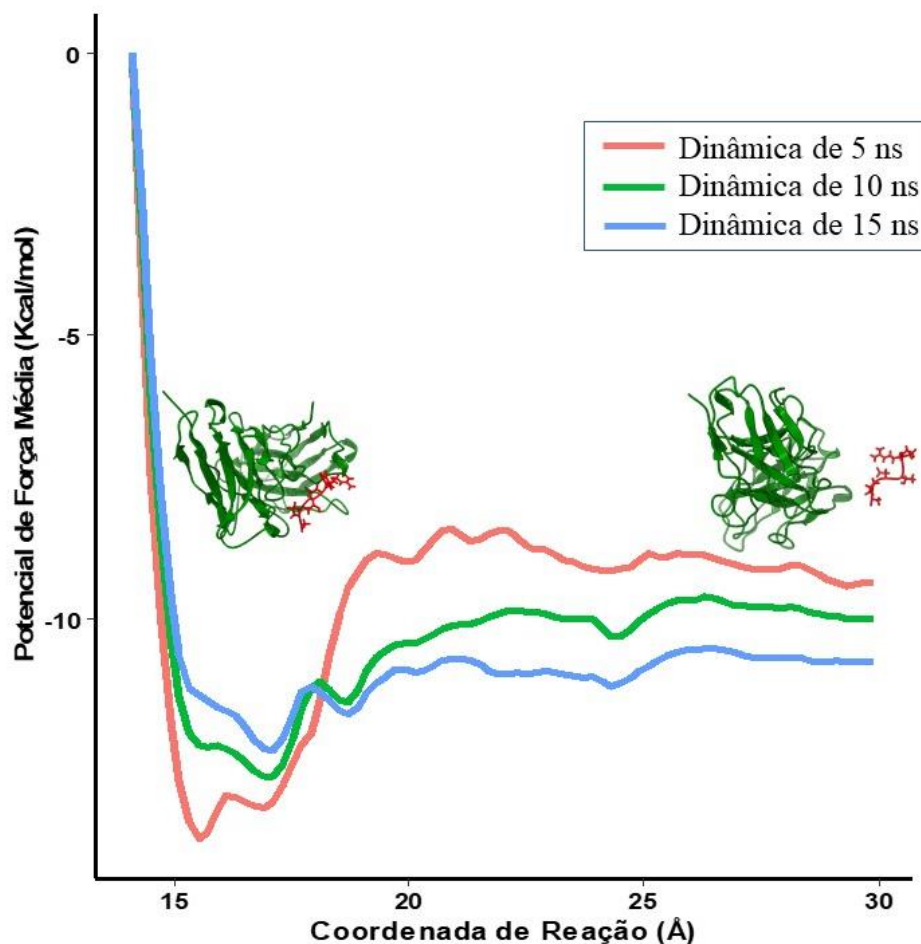


Figura 4.5 - Perfil do potencial da força média em função da coordenada de reação de dissociação entre o anticorpo SM3 e o peptídeo APDTRP.

Com 15 ns de simulação, estratificados em 8 janelas, obteve-se um valor de ΔG de ligação acima do previsto pelo servidor PISA (erro de 33% sobre o valor esperado). Embora os dois valores não sejam próximos, é preciso dizer que esse resultado não põe em xeque a acurácia do método ABF, dada a incerteza associada ao valor atribuído pelo servidor PISA. Esse valor é calculado apenas com base em estimativas sobre um conjunto pequeno de complexos proteicos depositados no PDB, e tem divergências já reportadas em relação a cálculos acurados feitos usando dinâmica molecular (116).

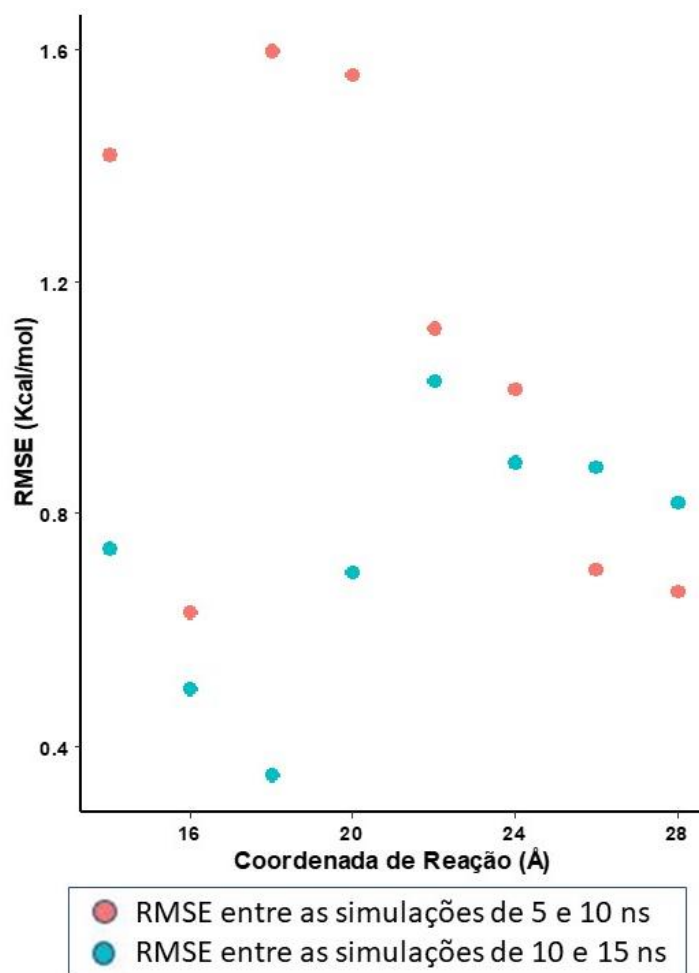


Figura 4.6 - Erro quadrático médio entre os valores de energia calculados para as simulações com 5A2J.

A dissociação entre o fragmento de anticorpo SM3 e o peptídeo $_{APDTRP}$ permite traçar a evolução das energias de interação entre as duas moléculas de acordo com a coordenada de reação (Figura 4.7) durante a simulação em que foi alcançada convergência (15 ns, com estratificação em 8 janelas de 2 Å).

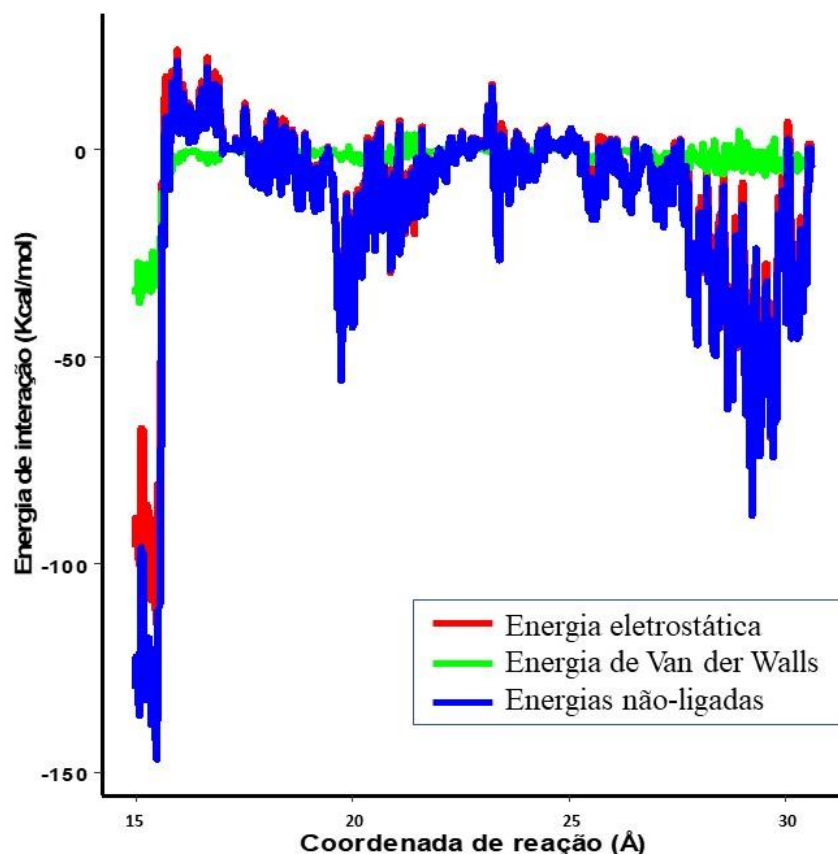


Figura 4.7 - Energias de interação entre o fragmento anticorpo SM3 e o peptídeo APDTRP em função da coordenada de reação.

As energias advindas de ligações de curto alcance rapidamente caíram para valores próximos a zero no início do processo de afastamento; enquanto que as energias advindas do potencial de Coulomb sofrem uma queda inicial, mas mostram diversos aumentos ao longo de várias janelas durante a simulação. Esses aumentos na energia eletrostática devem-se a artefatos da dinâmica molecular, decorrentes do formato da caixa de simulação e das condições periódicas de contorno. Por acontecer dentro de uma caixa pequena e aproximadamente cúbica, a translação do peptídeo fez com que o fragmento SM3 interagisse com uma ou mais imagens do peptídeo APDTRP.

Uma das possibilidades da análise computacional é a possibilidade de acompanhar a variação da configuração tridimensional das moléculas envolvidas. Esta análise é feita com base no RMSD de uma molécula específica (Figura 4.8). Durante a simulação em que foi alcançada convergência (15 ns, com estratificação em 8 janelas de 2 Å), inicialmente o peptídeo mudou drasticamente sua configuração estrutural – o que é ilustrado pela rápida elevação inicial no gráfico da Figura 4.8 –,

pois desliga-se do fragmento de anticorpo e então passa a mover-se livremente pela caixa de simulação, quando seus valores de RMSD estabilizam-se ao redor de 2 Å.

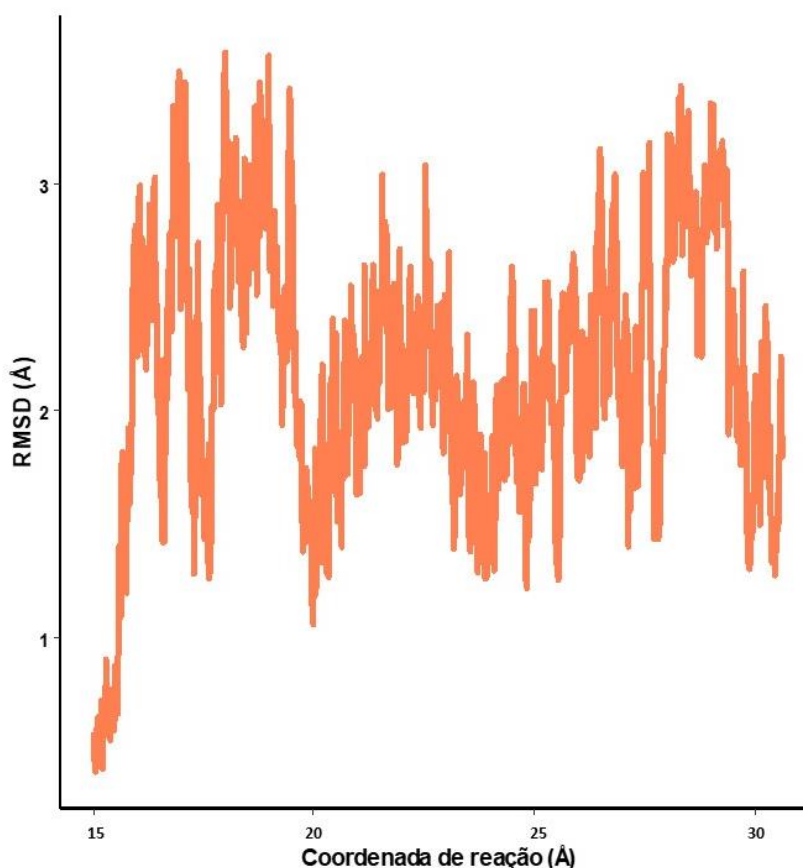


Figura 4.8 - Desvio Quadrático Médio do peptídeo APDTRP em função da coordenada de reação.

De acordo com o Teorema Central do Limite, quanto maior o tamanho da amostra de uma variável, mais a distribuição amostral de sua média aproxima-se de uma distribuição normal, e menor torna-se o seu desvio padrão (117). De forma análoga, técnicas computacionais que usam modelos matemáticos para descrever sistemas reais sustentam-se no tamanho da amostra da variável coletada para produzir resultados acurados. Assim, a amostragem de uma variável é fundamental para avaliar a acurácia de uma simulação. Utilizando o algoritmo de ABF, a variável coletada são as forças que agem sobre a coordenada de reação. Para este sistema, a força média agindo ao longo da variável coletiva foi calculada a cada $\delta\xi = 0,1 \text{ \AA}$. Este cálculo interno do algoritmo é feito através da média de todas as forças que agem sobre ξ em cada $\delta\xi$ específico através da equação 5.

$$\bar{F}_{\xi}(N_{step}, k) = \frac{1}{N_{step}^k} \sum_{\mu=1}^{N_{step}^k} F_{\mu}^k \quad (5)$$

Na qual k é o $\delta\xi$ em questão, N_{step} é o número de passos de dinâmica molecular dentro de $\delta\xi$ e μ representa o número ordinal do passo de dinâmica molecular no qual a força média, $\langle \mathbf{F}_{\xi} \rangle$, é calculada dentro de $\delta\xi$. Apenas quando o sistema evoluiu por um mínimo de $\mu = 1.000$ em cada $\delta\xi$, então a força média foi calculada, para que então a força de viés adaptativo, \mathbf{F}_{ABF} , fosse aplicada (68), levando o sistema ao $\delta\xi$ seguinte e armazenando o valor de $\langle \mathbf{F}_{\xi} \rangle$.

A amostragem, definida como o número de passos de dinâmica molecular usados para estimar a força média $\langle \mathbf{F}_{\xi} \rangle$ em cada $\delta\xi$, varia conforme a dissociação progride e pode refletir a precisão do cálculo da força média, conforme o Teorema do Limite Central. Assim, os valores de amostragem cresceram conforme a exploração da coordenada de reação aumentou, indicando, portanto, evolução em direção à convergência no cálculo de $\langle \mathbf{F}_{\xi} \rangle$ (Figura 4.9). Apenas a partir de 22 Å a amostragem tornou-se uniforme, e este é o ponto aproximado da coordenada de reação a partir do qual o potencial da força média atinge um platô (Figura 4.5).

Comer e colaboradores (68) sugerem que essa dificuldade em alcançar amostragem uniforme antes do equilíbrio pode ser devido à presença de barreiras ocultas ao longo de graus de liberdade ortogonais à coordenada de reação, que podem levar à descontinuidade da força de viés adaptativo entre janelas adjacentes. Contudo, os mesmos autores assumem que seja normal e esperado o comportamento observado neste sistema, no qual a amostragem torna-se uniforme, e o erro estatístico entre simulações diferentes diminui com o tempo (Figuras 4.5 e 4.6 acima).

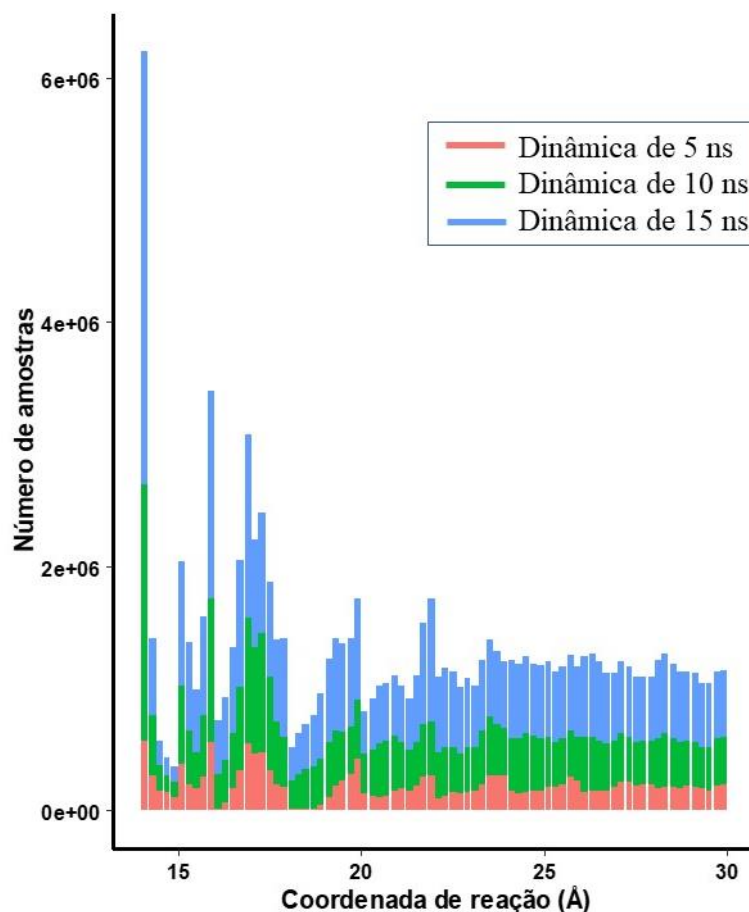


Figura 4.9 - Amostragem (μ máximo) sobre ξ coletada ao longo da dissociação entre o fragmento de anticorpo SM3 e o peptídeo $_{APDTRP}$.

A primeira parte deste projeto foi concluída com o estudo do perfil de energia livre de ligação do fragmento de anticorpo SM3 ao peptídeo $_{APDTRP}$. O valor de ΔG de ligação obtido e a referência medida pelo servidor PISA, apesar de não serem próximos, têm a mesma ordem de grandeza e o mesmo sinal, o que exemplifica a congruência entre diferentes métodos computacionais para resolver um problema biofísico. Contudo, o método de ABF é robusto o suficiente para fornecer valores semelhantes àqueles obtidos por métodos experimentais e para fazer previsões quanto à associação entre biomoléculas, conforme será mostrado na discussão sobre o sistema 5JYL a seguir.

4.3.2 Segundo Sistema – 5JYL

Para a estrutura 5JYL, o potencial da força média ao longo da variável coletiva (distância entre os centros de massa do fragmento de anticorpo TSP7 e da P-caderina MEC1) também foi calculado (Figura 4.10). O valor de referência obtido inicialmente, com 5 ns de simulação, foi de $\Delta G = -19,3$ kcal/mol. Quando o sistema desenvolveu 5

ns adicionais, até 10 ns, o valor de energia livre mudou para $\Delta G = -9,5$ kcal/mol, ao que se associa um RMSE de 1,379 kcal/mol em relação ao valor anterior. Finalmente, quando o sistema completou 15 ns de evolução, o valor de ΔG atingiu $-8,5$ kcal/mol, e este por sua vez tem um RMSE de 0,524 kcal/mol, comparado ao valor alcançado em 10 ns. Portanto, a convergência ocorreu com 15 ns de simulação, considerando um raio de corte de 1 kcal/mol (Figura 4.11).

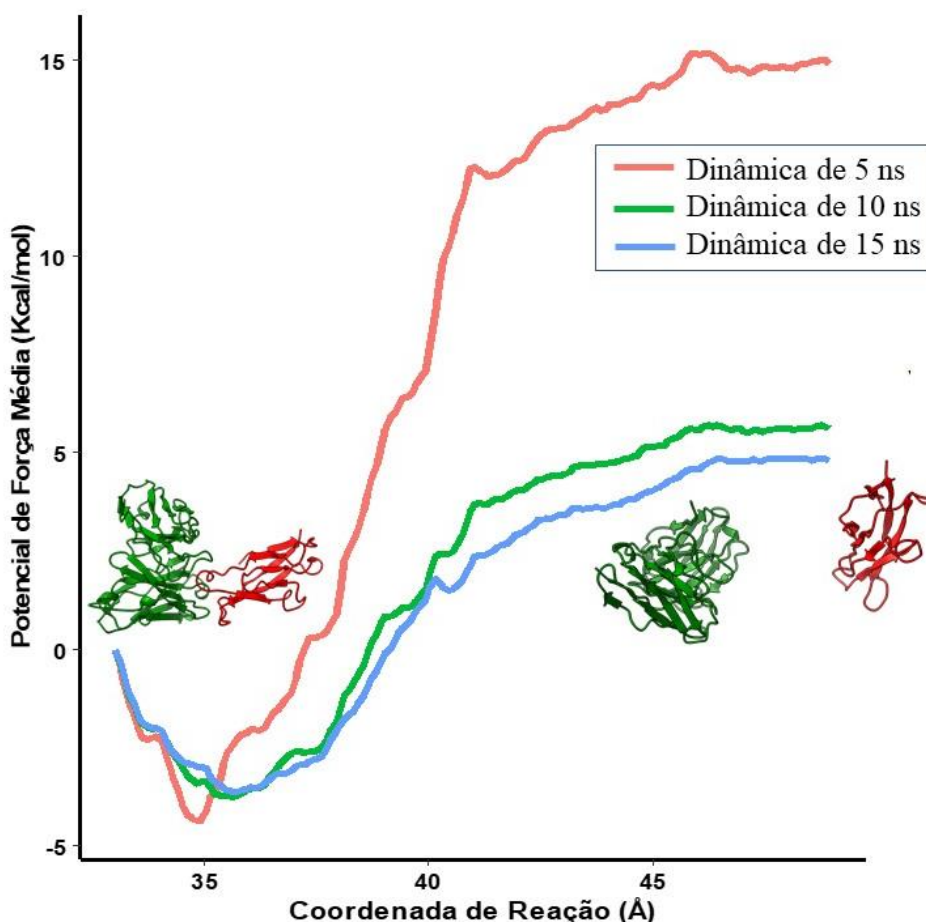


Figura 4.10 - Perfil do potencial da força média em função da coordenada de reação de dissociação entre o fragmento de anticorpo TSP7 e a P-caderina MEC1.

Trabalhos como Malmqvist (118) e Fägerstam e colaboradores (118) exemplificam como a técnica de ressonância plasmônica de superfície tem sido usada há décadas para o estudo da afinidade de anticorpos e da sua cinética de reação com antígenos. Kudo e colaboradores (88) aplicaram essa técnica para medir a constante de dissociação do complexo 5JYL à temperatura ambiente, obtendo o valor de 8,7 nM, que equivale a um ΔG de ligação igual a $-10,9$ kcal/mol. Com 15 ns de simulação, estratificados em 8 janelas, obteve-se um valor de ΔG de ligação, em módulo, levemente menor do que o obtido experimentalmente através da técnica de

ressonância plasmônica de superfície (erro de 12% sobre o valor experimental). Estudos semelhantes que usaram complexos proteína-proteína encontraram o mesmo erro estatístico ao comparar resultados experimentais com os valores obtidos por ABF (45). Este resultado reforça a acurácia, precisão e convergência do método para o estudo da termodinâmica de associação entre proteínas.

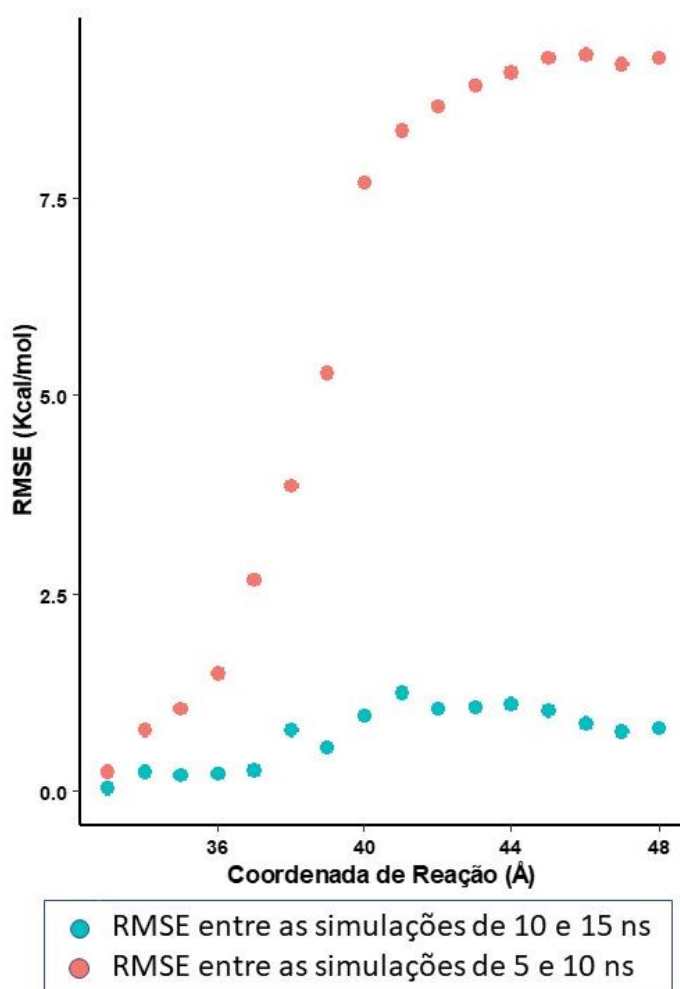


Figura 4.11 - Erro quadrático médio entre os valores de energia calculados para as simulações com 5JYL.

Assim como ocorreu no primeiro sistema, a evolução dos potenciais de força média foi explorada mais detalhadamente quanto melhor amostrada: a partir de 10 ns de simulação, um pico de potencial fica evidente no gráfico da Figura 4.10 quando o sistema passa por $\xi = 40$ Å. Esse pico de potencial é concomitante a um aumento da repulsão entre as duas proteínas, mostrado pelo aumento das energias não-ligadas na mesma abscissa do gráfico da Figura 4.12 abaixo. Além disso, o vale subsequente de potencial, em $\xi \cong 41$ Å, é acompanhado por um aumento da atração entre as duas moléculas, evidenciado pela queda dos valores de energia de interação (Figura 4.12).

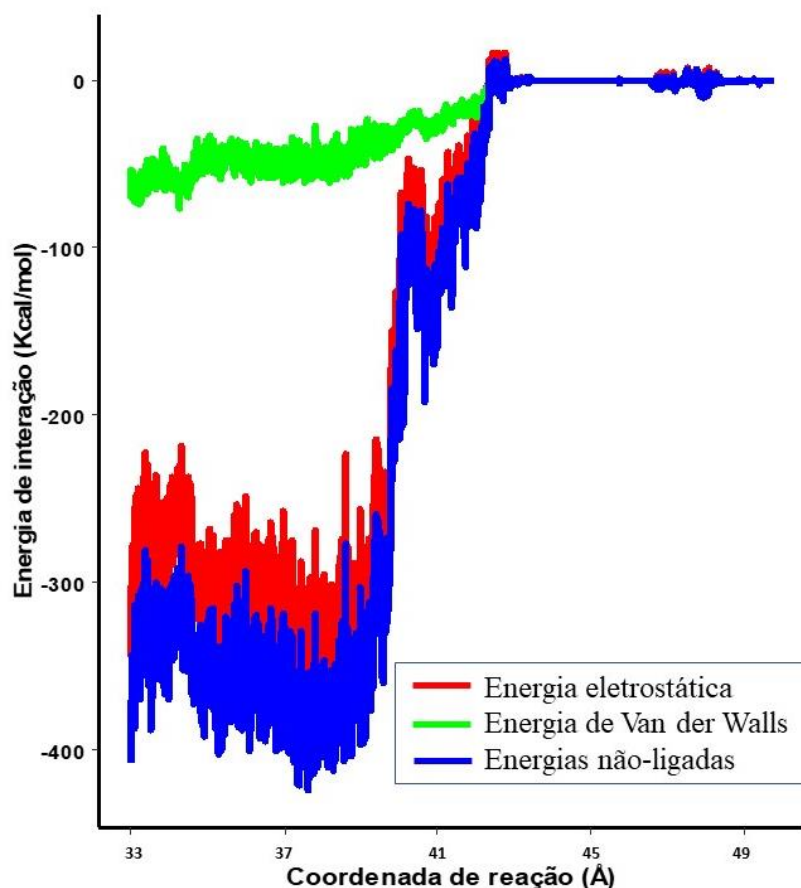


Figura 4.12 - Energias de interação entre o fragmento de anticorpo TSP7 e a P-caderina MEC1 em função da coordenada de reação.

Os valores de energia de ligação neste complexo decrescem de forma mais suave que os mesmos valores de 5A2J (Figura 4.7). Por envolver um fragmento maior de anticorpo e um antígeno também muito maior que o peptídeo de 5A2J, os resíduos de ambas proteínas permanecem em contato até metade da trajetória, quando os potenciais de Coulomb e de Lennard-Jones caem para valores próximos de zero. Portanto o decréscimo das energias de interação não-ligadas foi concomitante à evolução da coordenada de reação, com menos ruídos e artefatos devido às condições periódicas de contorno que aconteceram durante a simulação do primeiro sistema, provavelmente devido às diferenças do formato entre as duas caixas de simulação e ao tamanho do complexo antígeno-anticorpo.

Durante a simulação de 15 ns, a estrutura tridimensional da P-caderina MEC1 permaneceu praticamente inalterada ao longo de toda a trajetória, pois seu RMSD flutuou em valores pequenos, próximos de 2 Å (Figura 4.13).

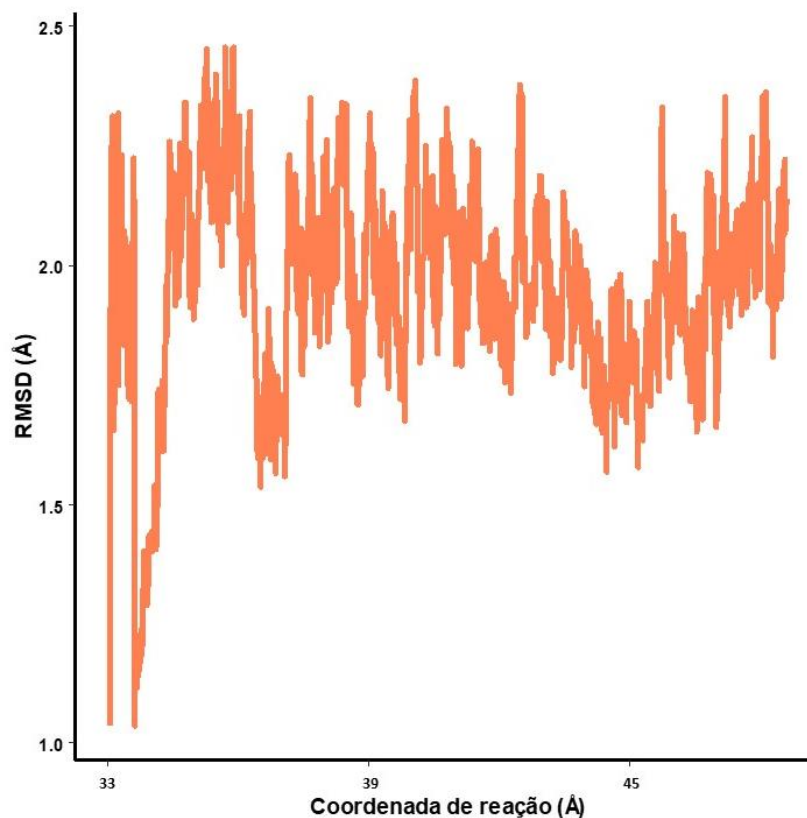


Figura 4.13 - Desvio Quadrático Médio da P-caderina MEC1 em função da coordenada de reação.

É interessante notar que pouco depois da metade da trajetória, por volta de $\xi = 44$ Å, o RMSD sofre um aumento considerável, que é simultâneo à anulação dos potenciais de Lennard-Jones e de Coulomb (Figura 4.12). Este era um resultado esperado, visto que a energia de interação com o fragmento TSP7 é uma das interações que mantém a P-caderina na sua configuração de co-cristalização com o fragmento de anticorpo (Figura 3.1).

Assim como no primeiro sistema, a amostragem ao longo da coordenada de reação foi quantificada para 5JYL, de acordo com o tempo de simulação por janela (Figura 4.14). O mesmo problema do primeiro sistema foi novamente observado aqui, com amostragem não-uniforme nos primeiros passos da variável coletiva. De fato, a amostragem uniforme só foi alcançada depois que o potencial de força média atingiu um platô (Figura 4.10). Além disso, o intervalo com menor amostragem em todo o caminho de reação foi $\xi = 41$ Å, que corresponde à região em que a repulsão foi mais intensa entre as duas proteínas, o que é mostrado pelo rápido e intenso aumento das energias de interação não-ligadas em $\xi = 41$ Å na Figura 4.8 acima.

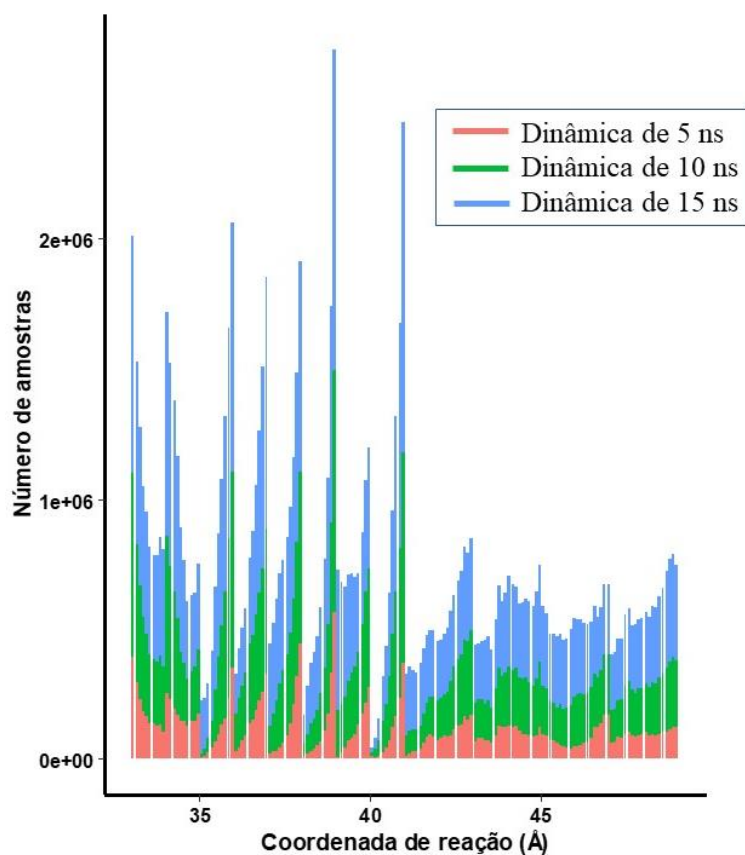


Figura 4.14 - Amostragem (μ máximo) sobre ξ coletada ao longo da dissociação entre o fragmento de anticorpo TSP7 e a P-caderina MEC1.

Contudo, como foi destacado a respeito da amostragem do primeiro sistema, este fenômeno de amostragem não-uniforme pode acontecer como consequência da estratégia de estratificação, mas não apresenta um problema para a interpretação dos resultados obtidos, uma vez que a convergência tenha sido alcançada em cada janela individual.

Adicionalmente, as interações proteína-proteína ao longo da dinâmica também foram analisadas. Foram contabilizadas as ligações hidrogênio e as pontes salinas entre o anticorpo e a caderina, seguindo os critérios internos do VMD (90).

A análise das ligações hidrogênio foi feita em função de suas ocupâncias em cada janela da coordenada de reação (Figura 4.15). A ocupância aqui é definida como a porcentagem de configurações da janela em que a ligação específica esteve presente. Foram selecionadas para compor o grupo da Figura 4.15 apenas as ligações hidrogênio que apresentaram ocupância de 50% ou mais em qualquer janela da trajetória (119). Os resíduos que estabelecem cada ligação hidrogênio estão mostrados abaixo, em cores de destaque.

A maior parte dos resíduos da P-caderina que estabelecem ligações hidrogênio com o anticorpo estão concentrados em uma única volta, com exceção do Asp87, que estabelece a interação com o resíduo Tyr151 do fragmento TSP7. Apesar de estarem distantes comparados aos outros resíduos que formam ligações hidrogênio, estes dois resíduos interagem durante praticamente toda a primeira metade do processo de dissociação, formando a segunda ligação hidrogênio de maior ocupância.

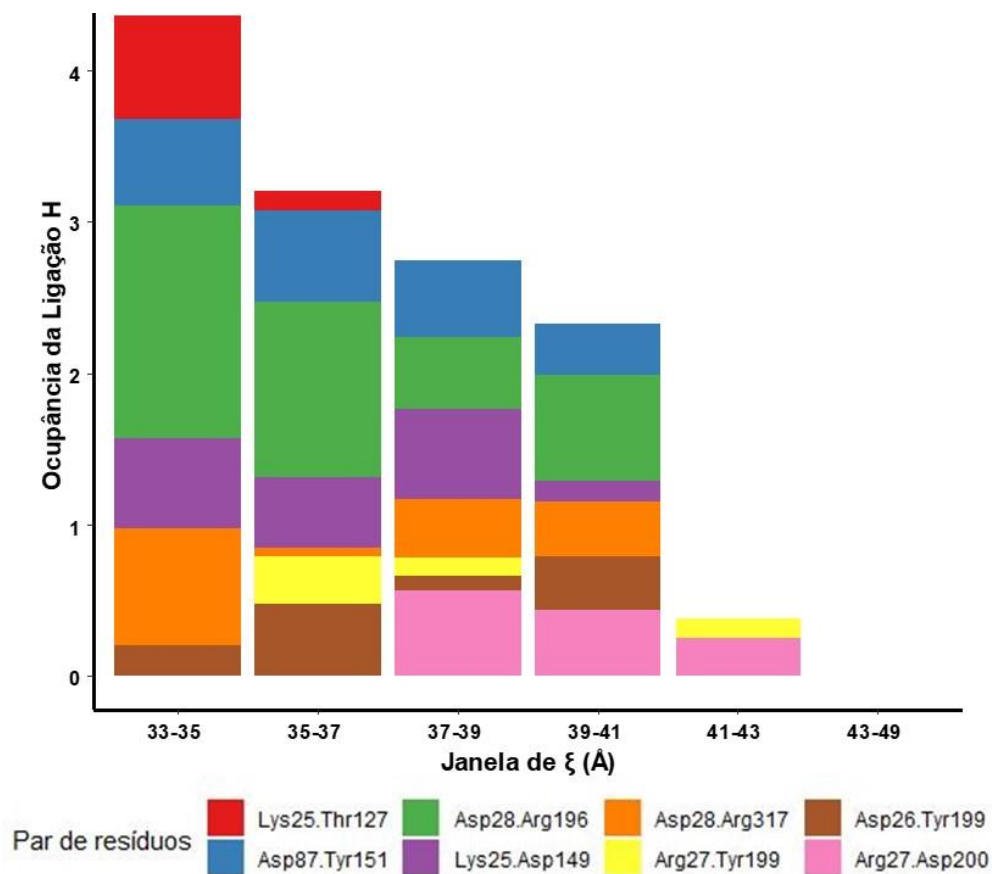


Figura 4.15 - Ligações Hidrogênio e sua ocupância de acordo com a coordenada de reação durante a simulação de 15 ns.

As ligações hidrogênio Lys25-Thr127, Asp87-Tyr151, Asp28-Arg196 e Lys25-Asp149 haviam sido anteriormente previstas por Kudo e colaboradores. (88) através do servidor PISA. Dentre as quatro novas destacadas na Figura 4.15, as duas estabelecidas pela Arg27 (Arg27-Tyr199 e Arg27-Asp200) merecem destaque, pois formam-se apenas nas últimas janelas em que ainda há contato entre o fragmento TSP7 e a P-caderina. De forma análoga, no processo reverso, de associação entre antígeno e anticorpo, estas duas ligações hidrogênio seriam cruciais para o reconhecimento inicial entre as duas proteínas. Estratégias de maturação da afinidade do fragmento de anticorpo TSP7 pela P-caderina MEC1, portanto, poderiam se basear

neste dado para realizar mutações sítio-dirigidas, de forma a promover associação mais forte entre as duas moléculas.

As pontes salinas intermoleculares e seus respectivos comprimentos também foram analisados durante a trajetória de separação entre as duas proteínas (Figura 4.16). Assim como com as ligações hidrogênio, as pontes salinas são extintas à medida que a simulação evolui e as duas proteínas se afastam. Dentre as seis pontes encontradas ao longo da simulação, Kudo e colaboradores (88) encontraram apenas duas usando o servidor PISA: Asp149-Lys25 e Arg196-Asp28. Essas duas são visivelmente as ligações mais estáveis do complexo (ocupância 43,86% e 37,5%, respectivamente considerando todo o caminho de reação), pois são mantidas durante toda a primeira metade da dissociação, quando as duas proteínas ainda estão em contato. As demais pontes salinas são estabelecidas apenas durante pequenos trechos da coordenada de reação. Contudo, as pontes salinas estabelecidas entre Asp200-Arg27 e Arg317-Asp28 (ocupâncias 5,73% e 12,43%, respectivamente) merecem destaque, pois são as últimas a serem extintas, e, portanto, seriam as primeiras a se formarem durante a associação antígeno-anticorpo. Este resultado sugere que essas duas ligações são responsáveis por direcionar a formação das demais interações e por orientar a associação entre as duas proteínas.

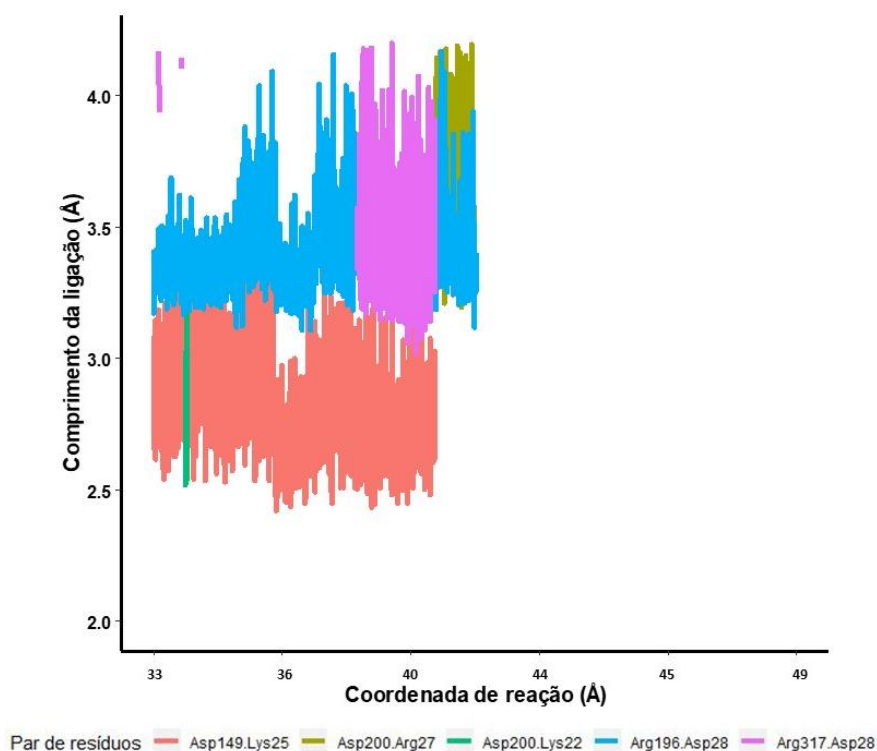


Figura 4.16 - Comprimento das pontes salinas intermoleculares de acordo com a coordenada de reação durante a simulação de 15 ns.

4.4 Adaptive Biasing Force do Complexo Anticorpo (Nativo)-Antígeno

A difícil tarefa de medir a energia livre de ligação entre proteínas através de técnicas computacionais mostrou-se viável, com certa acurácia, para o sistema 5JYL destacado na seção 4.3.2. No entanto, para o anticorpo 19CC6CG2, nunca foi feita nenhuma medida experimental de sua afinidade contra o antígeno de superfície do vírus da hepatite B, apenas a estimativa entálpica oriunda do cálculo de MM-PBSA (seções 3.2.1 e 4.2). Contudo, a negligência do termo entrópico nessa estimativa de ΔG_{LIG} e a inerente dificuldade do método de computar termos energéticos importantes, como a contribuição das água estruturais, a solvatação de grupos polares e de grupos não-expostos ao solvente (51), torna necessária uma nova abordagem através de um cálculo mais robusto de ΔG_{LIG} , daí o uso de ABF.

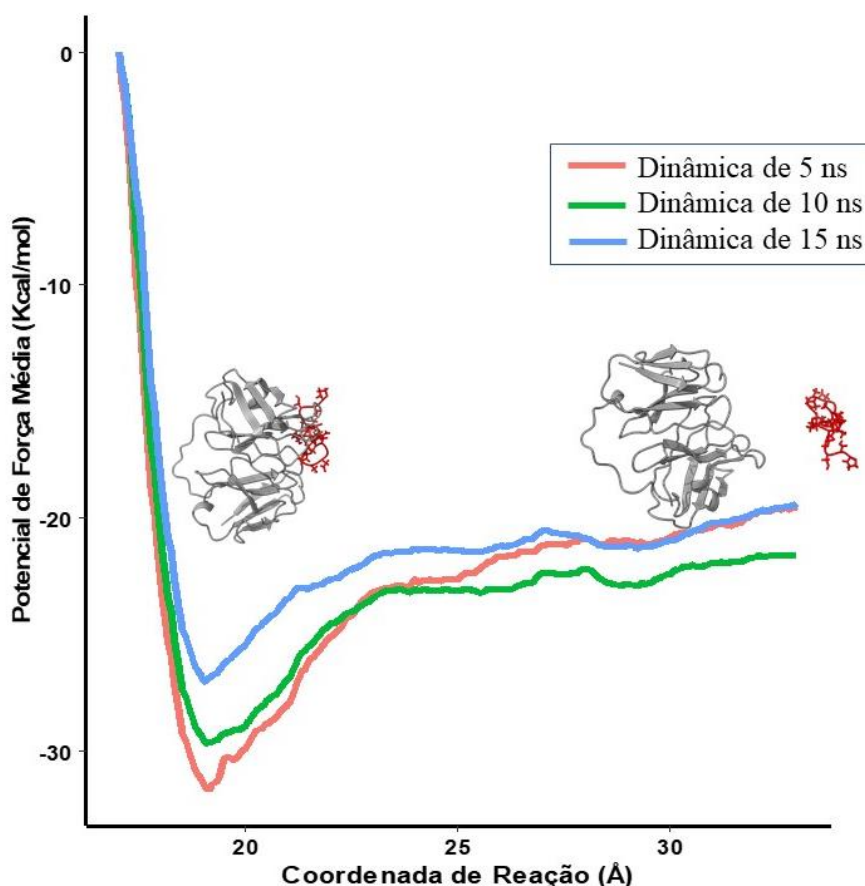


Figura 4.17 - Perfil do potencial da força média em função da coordenada de reação de dissociação entre o anticorpo 19CC6CG2 e o HBsAg.

Quando a técnica foi aplicada sobre o sistema durante a dinâmica molecular de 5 ns, o valor de ΔG_{LIG} obtido foi de -12,09 kcal/mol, mas quando a amostragem do sistema foi aumentada por mais 5 ns, o valor fornecido foi de -8,03 kcal/mol. Entre

estes dois valores se associa um RMSE de 1,23 kcal/mol. Finalmente, quando o sistema evolui por 15 ns, o ΔG_{LIG} assume o valor de -7,62 kcal/mol, com um RMSE de 1,09 kcal/mol em relação à simulação anterior (Figuras 4.17 e 4.18).

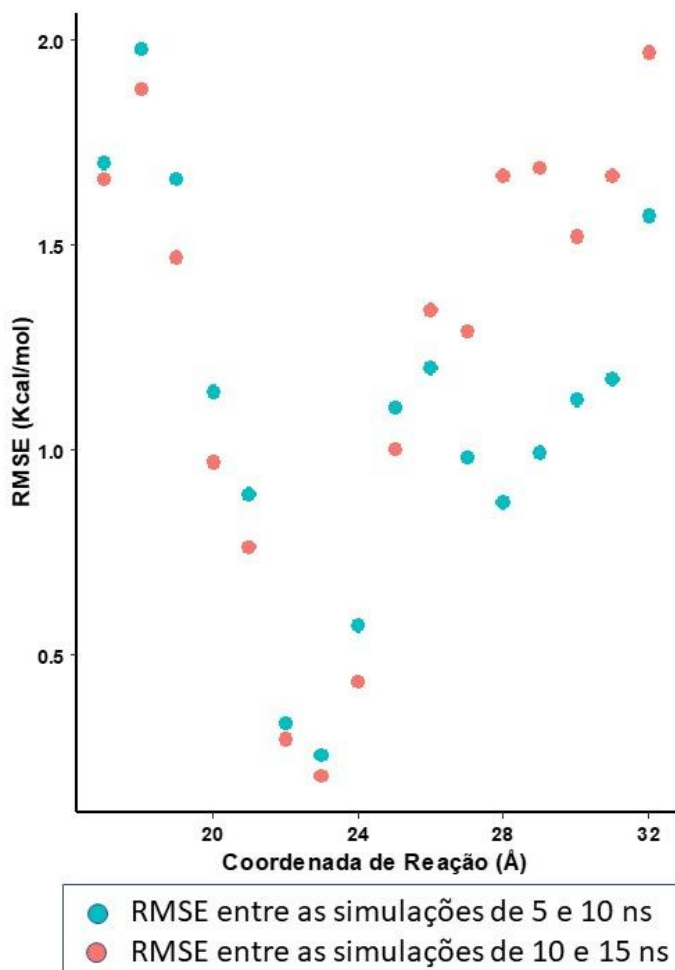


Figura 4.18 - Erro quadrático médio entre os valores de energia calculados para as simulações com o complexo anticorpo 19CC6CG2-HBsAg.

Considerando-se o valor entálpico obtido através do cálculo de MM-PBSA e o valor de ΔG_{LIG} obtido ao fim da simulação de 15 ns de ABF, pode-se estimar a perda entrópica do processo através da equação 6.

$$\Delta G = \Delta H - T\Delta S \quad (6)$$

O valor obtido é de, aproximadamente, $-16 \text{ cal}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$, consistente com o aumento da ordem do sistema esperado mediante associação entre antígenos e anticorpos em meio aquoso (120,121). Mesmo bastante aproximada, esta estimativa da entropia de ligação, portanto, corrobora com a confiabilidade dos resultados de ΔG_{LIG} obtidos neste trabalho.

Ao se analisar a evolução das energias de interação ao longo da trajetória inicial de 5 ns (Figura 4.19), nota-se que, conforme o antígeno se distancia do anticorpo, as energias de interação não-ligadas vão a zero, mas voltam a assumir valores altamente negativos ao fim da simulação. Esta é uma observação interessante, pois sugere que sob efeitos difusivos, resultantes do método de ABF, o sistema evolui até uma configuração de energia de interação duas vezes mais forte do que havia sido proposto pelos resultados do *docking*.

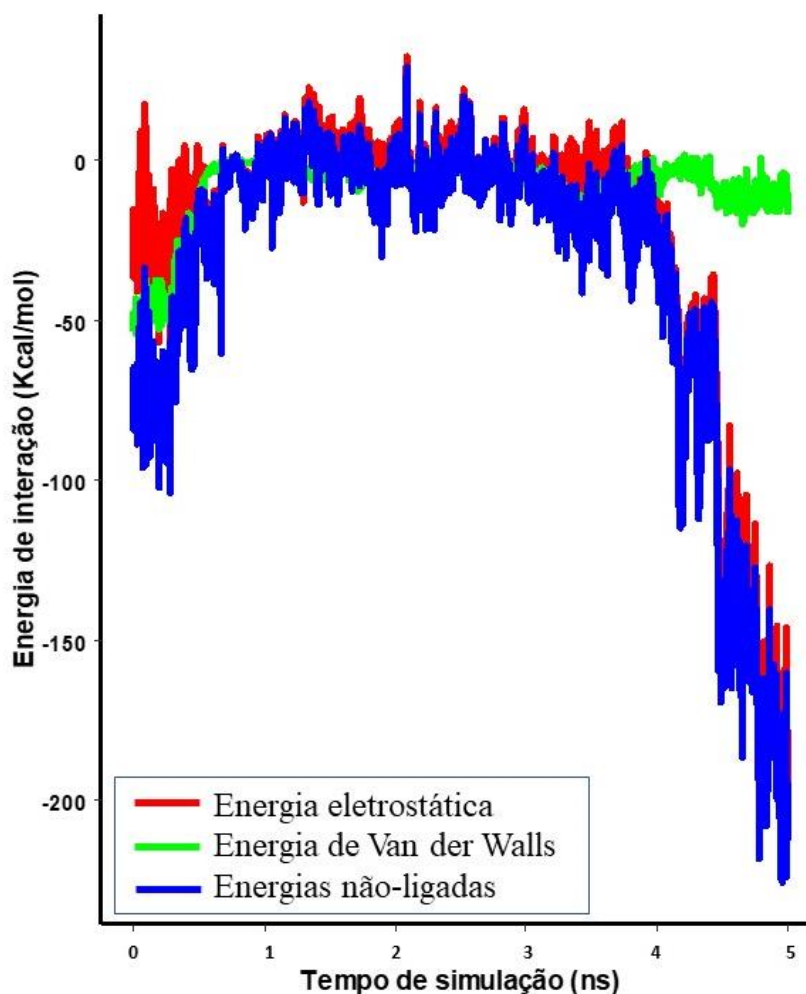


Figura 4.19 - Energias de interação entre o fragmento de anticorpo 19CC6CG2 e o HBsAg em função do tempo de simulação.

Em uma ampla revisão literária sobre algoritmos de *docking* aplicados a sistemas anticorpo-antígeno, Pedotti e colaboradores destacaram os pontos fortes e fracos do HADDOCK quando aplicado a sistemas como este: a flexibilidade do esqueleto peptídico inclusa no cálculo do HADDOCK o torna capaz de identificar corretamente a pose de V_H , mas falha quanto à pose de V_L , possivelmente numa tentativa de evitar choques estéricos entre as duas cadeias variáveis (103). Essa conclusão, portanto, é corroborada com o resultado encontrado nesta simulação

inicial: o modo de ligação do anticorpo 19CC6CG2 ao HBsAg pode ser diferente daquele previsto pelo HADDOCK, com uma participação mais expressiva de V_L , mais especificamente do CDR L1. De fato, quando são analisadas as últimas 10 configurações do complexo nesta dinâmica, constata-se que o HBsAg sofreu um desvio quadrático médio de cerca de 15 Å até chegar à sua nova posição (Figura 4.20), onde mantém três novas interações com o anticorpo:

- i. Uma ligação hidrogênio entre Pro217 e Lys152 do scFv, de ocupância 27,27%;
- ii. Uma ponte salina entre Lys207 e Glu9 (CDR L1) do scFv, de ocupância 45,45%;
- iii. Uma ligação hidrogênio entre Ser202 e Thr145 (CDR H2) do scFv, de ocupância 27,27%.

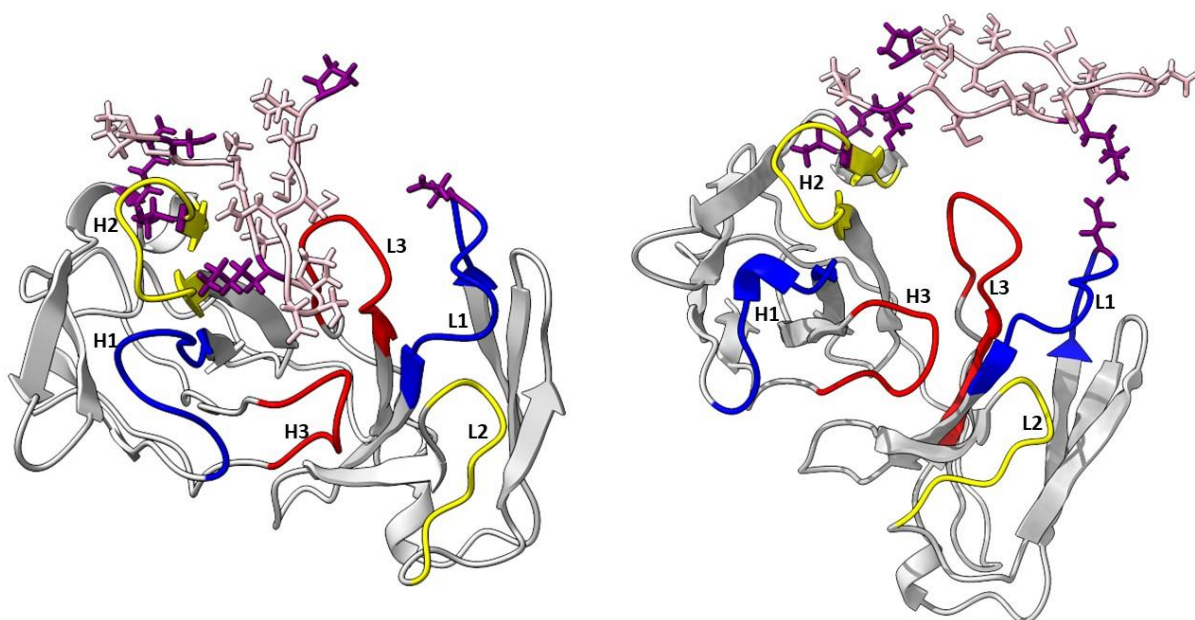


Figura 4.20 - À esquerda o complexo scFv-HBsAg antes da dinâmica com ABF, e à direita o mesmo complexo exibindo um provável novo modo de ligação. Os resíduos envolvidos nas novas ligações hidrogênio estão destacados em roxo.

A possibilidade de explorar o espaço-fase de forma mais uniforme graças ao algoritmo do método ABF permitiu que o sistema evoluísse sobre uma superfície plana de energia (68) e encontrasse um novo mínimo de energia de interação, provavelmente inacessível através de simulações não-enviesadas de dinâmica molecular como aquelas usadas em algoritmos de *docking* proteína-proteína. É possível, portanto, que, seguindo este raciocínio, com mais análises e mais tempo de

simulação seja possível propor modos de ligação entre proteínas através do método ABF.

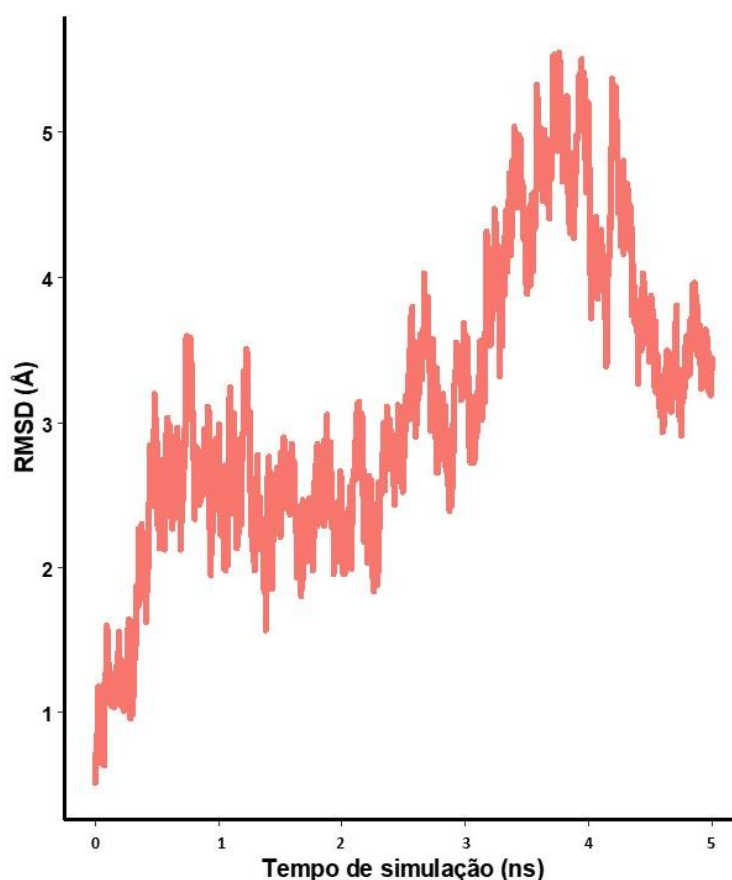


Figura 4.21 - Desvio Quadrático Médio do HBsAg em função do tempo de simulação.

O RMSD do HBsAg (alinhado com sua configuração inicial) também foi monitorado durante a simulação inicial (Figura 4.21). O antígeno manteve sua conformação, com pequenos desvios de até 5 Å neste tempo. De forma semelhante ao que ocorreu nos sistemas 5A2J e 5JYL, o aumento do RMSD foi concomitante ao enfraquecimento das interações intermoleculares (Figura 4.19 acima), e vice-versa: quando o complexo começou a formar interações que o levaram a assumir o novo “modo de ligação” (Figura 4.20), o seu RMSD diminuiu cerca de 2 Å.

A amostragem também foi quantificada ao longo das simulações de ABF com este sistema, e, como esperado, esse parâmetro aumentou conforme o crescimento do tempo de simulação (Figura 4.22). Apesar disso, houve trechos da coordenada de reação que tiveram amostragem muito baixa, como em $\xi = 21$ Å, e que podem ser essenciais para entender como o potencial de força média evolui nos últimos momentos em que o antígeno ainda faz contato com o anticorpo. O refinamento deste

intervalo de ξ , com a adição de mais tempo de simulação por $\delta\xi$, poderia ajudar a uniformizar a amostragem da coordenada de reação.

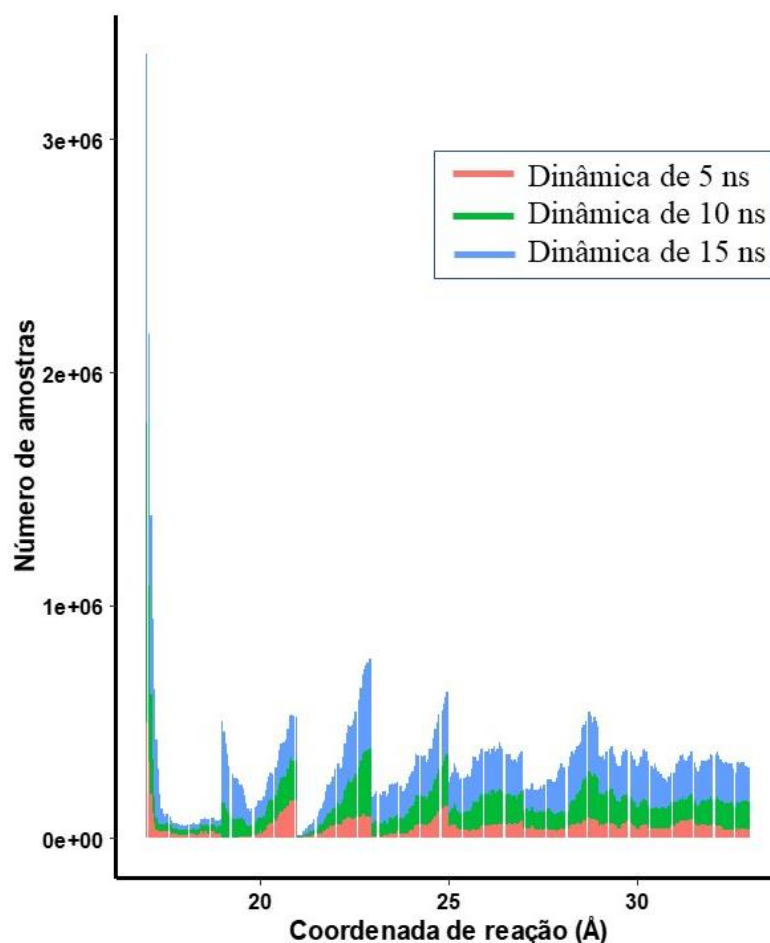


Figura 4.22 - Amostragem (μ máximo) sobre ξ coletada ao longo da dissociação entre o anticorpo 19CC6CG2 e o HBsAg.

Ao se analisar a formação e extinção de ligações hidrogênio e pontes salinas ao longo da dissociação entre as duas proteínas (Figuras 4.23 e 4.24), observa-se outro fato curioso: nenhuma das janelas de ξ está livre de interações intermoleculares, mesmo aquelas nas quais o antígeno se encontra a mais de 30 Å de distância do anticorpo. Isso acontece por dois motivos:

- A variável coletiva utilizada para dirigir o processo de ABF é a distância euclidiana entre os centros de massa das duas proteínas; portanto, quando $\xi = 30$ Å, por exemplo, existe uma gama diversa de posições (semelhante à superfície de uma esfera) ao redor do centro de massa do anticorpo que o centro de massa do antígeno pode explorar, e algumas dessas posições permitem que o HBsAg faça contatos polares com regiões distintas do anticorpo;
- A molécula do antígeno é muito pequena – apenas 17 resíduos de aminoácidos –, portanto a dinâmica molecular, mesmo em curtos intervalos de tempo, permite

que este explore diversas conformações estruturais ao longo da evolução de ξ , levando à formação de interações com diversas regiões do anticorpo.

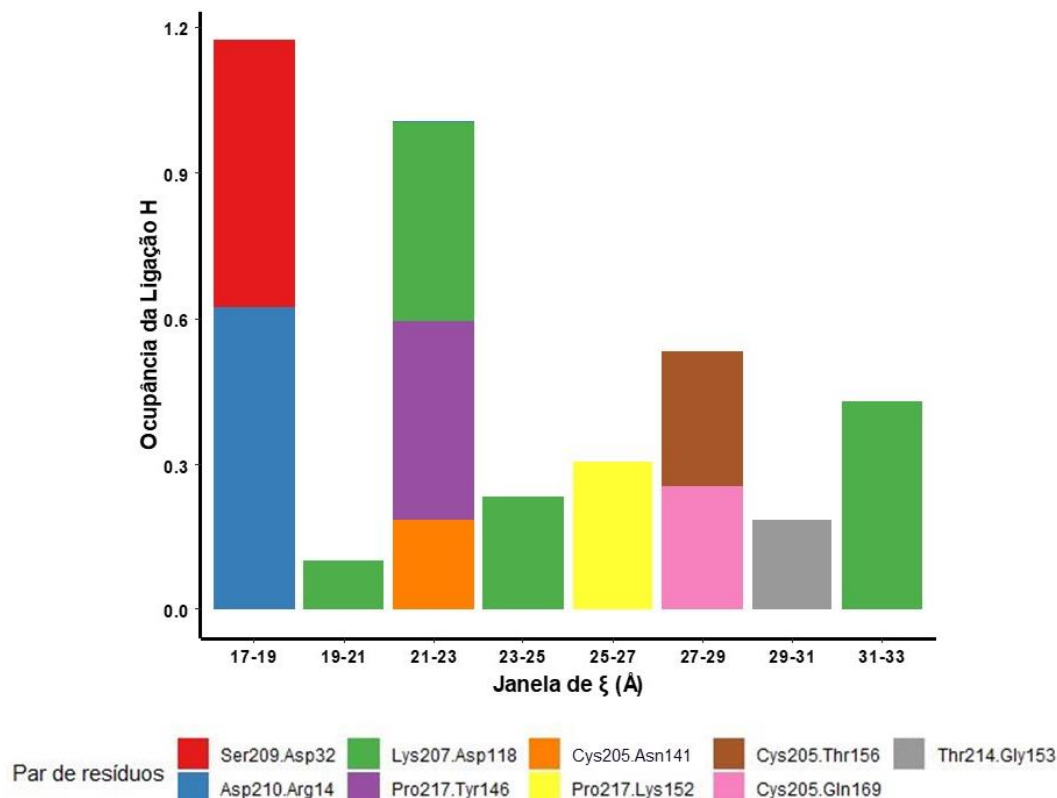


Figura 4.23 - Ligações Hidrogênio e sua ocupância de acordo com a coordenada de reação durante a simulação de 15 ns. Os resíduos do antígeno estão listados primeiro, seguidos do resíduo do anticorpo com o qual fazem contato.

Este problema com as interações polares é um efeito colateral do uso de apenas uma variável coletiva para descrever um processo biofísico: neste caso, para melhor representar a associação/dissociação entre as duas proteínas seria necessário a adição de uma segunda variável coletiva que permitisse a amostragem mais completa das suas entropias translacional, rotacional e conformacional. Seguindo esse raciocínio, Haohao e colaboradores aplicaram diversas restrições de graus de liberdade – ângulos polares, ângulos de Euler e RMSD, além da distância euclidiana – sobre sistemas proteína-ligante e obtiveram valores altamente acurados de ΔG_{LIG} (erro de 0,2 kcal/mol em comparação com os valores experimentais) (122).

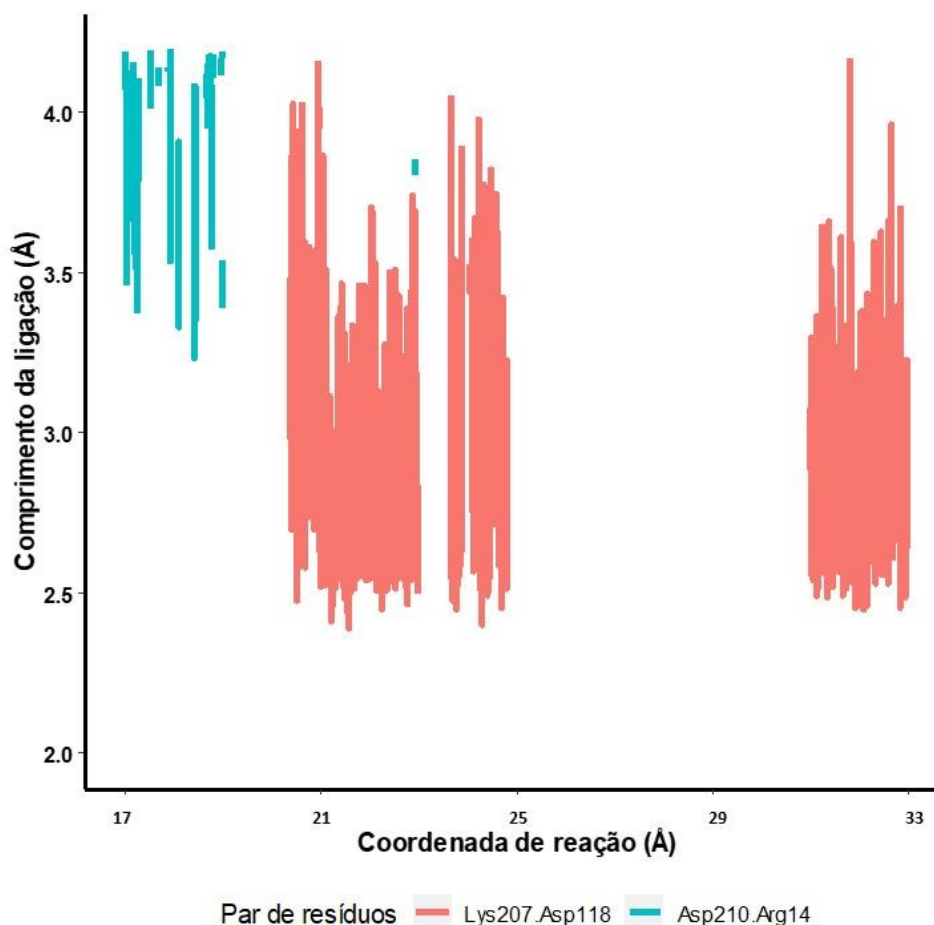


Figura 4.24 - Comprimento das pontes salinas intermoleculares de acordo com a coordenada de reação durante a simulação de 15 ns. Os resíduos do antígeno estão listados primeiro, seguidos do resíduo do anticorpo com o qual fazem contato.

Apesar da óbvia simplicidade do trabalho conduzido com o uso de apenas uma variável coletiva, foi possível estimar com certa confiabilidade a afinidade do anticorpo 19CC6CG2 contra o antígeno de superfície do vírus da hepatite B. Outros trabalhos também mostram o uso de uma coordenada de reação unidimensional no estudo da associação entre proteínas com acurácia desde que a amostragem seja uniforme ao longo de ξ (42,123,124). Portanto, para uma análise mais acurada das interações intermoleculares que dirigem a associação entre o anticorpo 19CC6CG2 e o HBsAg, a escolha de uma variável coletiva adicional pode ser vantajosa, mas o cálculo da sua energia livre de ligação pode ser estimado de forma mais simples, com uma variável coletiva e suficiente tempo de simulação para atingir uma amostragem uniforme ao longo de ξ .

4.5 Engenharia do Anticorpo 19CC6CG2

A dinâmica molecular executada no item 3.2 mostrou que os resíduos Asp118, Ser144, Tyr119, Asn139 e Ile76 mantiveram ligações hidrogênio com o antígeno com ocupância superior a 50% ou se mantiveram a 4 Å ou menos de distância do HBsAg em 50% da simulação, portanto esses resíduos de aminoácidos foram “protegidos” durante os ciclos de engenharia do RosettaScripts – não foi permitido que sofressem mutações.

A criação de mutantes a partir do anticorpo 19CC6CG2 gerou 100.000 estruturas com mutações e modos de ligação diversos, os quais foram avaliados segundo a energia de interface e o $\Delta\Delta G$, uma estimativa da diferença de energia livre advinda da(s) mutação(ões) realizada(s) e do modo de ligação proposto (Figura 4.25). Os valores destes parâmetros estão descritos em uma unidade arbitrária – *Rosetta Energy Unit* – a qual não é equivalente a nenhuma unidade de energia real, pois o seu cálculo varia entre os protocolos da plataforma Rosetta (125).

Na Figura 4.25 abaixo, a estrutura nativa do anticorpo 19CC6CG2 está destacada contra os mutantes gerados pelo protocolo do RosettaScripts. Apesar de não ter unidades físicas reais, os parâmetros energéticos fornecidos pelo RosettaScripts são tão melhores quanto menores forem seus valores (126). Portanto é possível inferir que todas as estruturas à direita e acima da estrutura nativa têm qualidade inferior a esta, enquanto que todas as estruturas abaixo e à esquerda desta são mutantes promissores, que supostamente têm maior afinidade pelo HBsAg.

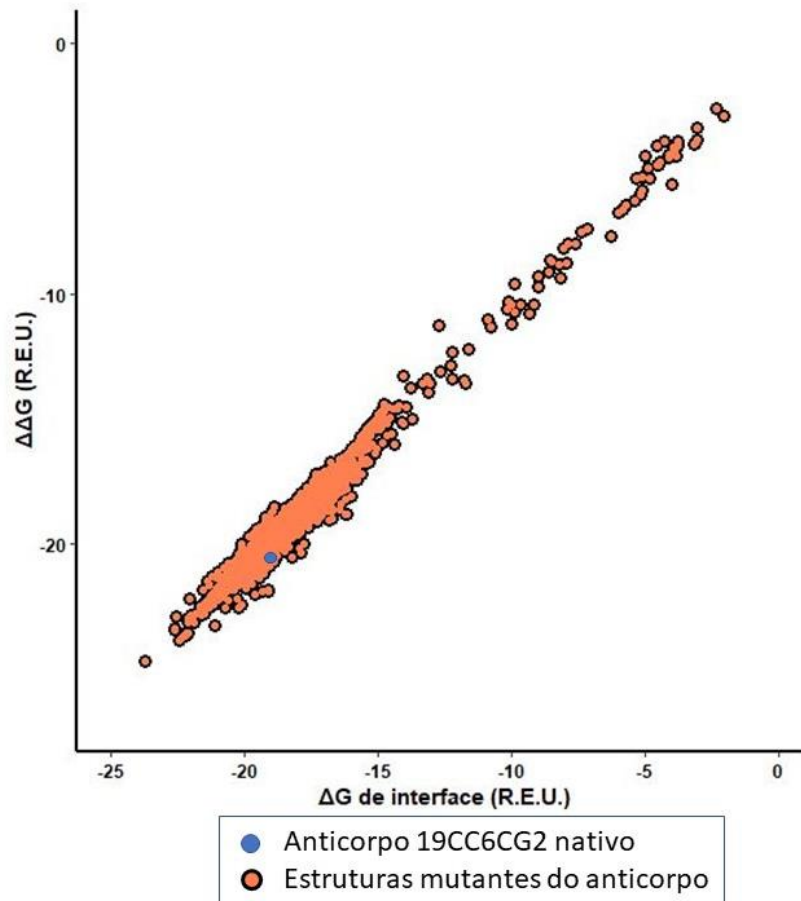


Figura 4.25 – Distribuição dos mutantes gerados pelo RosettaScripts em função de seus respectivos valores de $\Delta\Delta G$ e de energia de interface.

Assim, é evidente que a melhor solução do RosettaScripts para a engenharia do anticorpo 19CC6CG2 se encontra no canto inferior esquerdo da Figura 4.25: seus valores de $\Delta\Delta G$ e ΔG de interface são, respectivamente, -24,709 R.E.U. e -23,698 R.E.U., os menores entre todas as 100.000 estruturas geradas. Estes parâmetros para o anticorpo nativo têm valores de -20,385 R.E.U. e -19,311 R.E.U., respectivamente.

O anticorpo engenheirado conta com as seguintes mutações:

- i. W74Y (CDR L3);
- ii. N141E (CDR H2);
- iii. T145L (CDR H2).

A troca de um resíduo de triptofano por tirosina na posição 74 não afetou apreciavelmente o modo de ligação do anticorpo ao antígeno nem causou impacto considerável na conformação dos resíduos vizinhos à posição 74. Por sua vez, a substituição de um resíduo de asparagina por glutamato na posição 141 favoreceu contatos polares com os resíduos de asparagina 212 e lisina 207 do HBsAg: estes dois resíduos aproximaram-se cerca de 3 Å do glutamato 141 quando comparados às

suas posições no anticorpo nativo. Por fim, a troca de um resíduo polar – treonina – por um apolar e hidrofóbico – leucina – na posição 145 favoreceu contatos hidrofóbicos entre a vizinhança desta posição, composta majoritariamente por resíduos hidrofóbicos – glicina e tirosina, e a região predominantemente apolar que compreende os resíduos 201 a 205 do HBsAg. Apesar das mutações drásticas feitas pelo protocolo do RosettaScripts, o RMSD entre a estrutura nativa e o mutante é de apenas 0,095 Å (Figura 4.26).

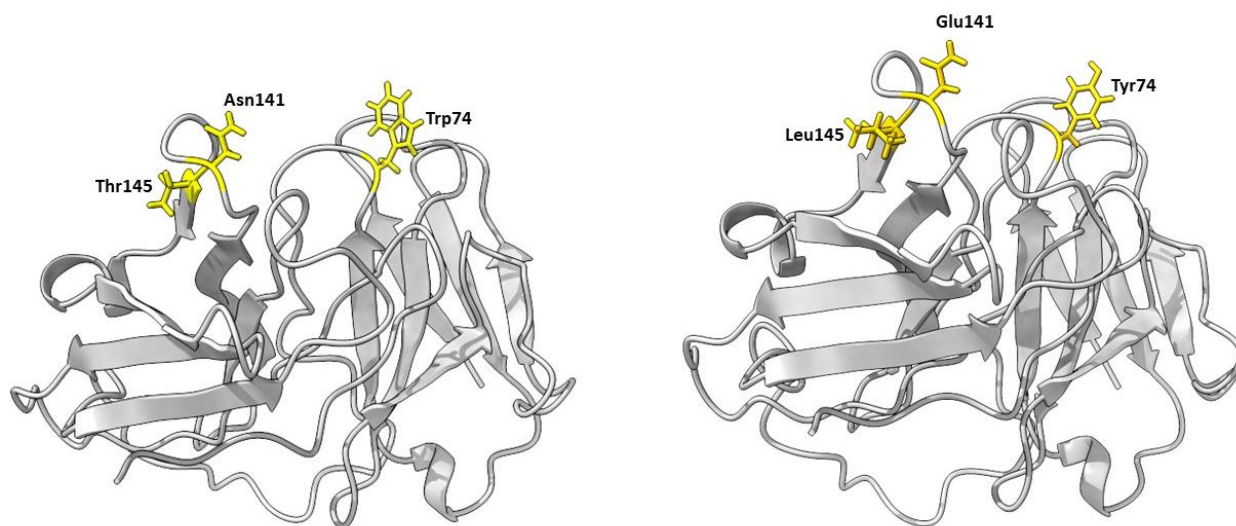


Figura 4.26 - O anticorpo 19CC6CG2 nativo, à esquerda, e o mutante, à direita.

As funcionalidades de modelagem molecular do Rosetta são extensivamente exploradas com sucesso na área de engenharia de proteínas e planejamento de interfaces proteicas, incluindo interfaces antígeno-anticorpo (127–133). O método desenvolvido nesta etapa de engenharia do anticorpo se assemelha à abordagem “*dock-and-design*” descrita por Procko e colaboradores (134). Estes autores deixam claro que a abordagem foi falha para o planejamento de um inibidor proteico da enzima lisozima, cujo sítio ativo é fortemente polar. Em seu trabalho com o RosettaScripts, os autores constataram que as mutações sugeridas para o inibidor não afetaram sua afinidade pelo sítio ativo da enzima; e, além disso, que mutações obtidas através de Reação em Cadeia da Polimerase Propensa a erros que aumentaram a afinidade deste inibidor pela enzima estavam todas fora das posições propostas pelo protocolo de “*dock-and-design*”.

A esta falha os autores atribuem a alta relevância dada pelo Rosetta aos contatos hidrofóbicos de interface, em contraste à baixa relevância dos contatos polares, e elaboram uma justificativa para este viés hidrofóbico das ferramentas do Rosetta: a amostragem de doadores e aceptores de ligações hidrogênio expostos ou

não na superfície é muito mais difícil do que o simples empacotamento de cadeias laterais apolares (134).

Neste trabalho, o efeito das mutações realizadas no anticorpo foi avaliado por meio da medição do ΔG_{LIG} do anticorpo nativo e do mutante e comparação destes valores. Para tal, lançamos mão novamente do método ABF.

4.6 Adaptive Biasing Force do Complexo Anticorpo (Mutante)-Antígeno

O anticorpo mutante descrito na seção anterior supostamente teria um valor de afinidade maior que o anticorpo nativo, contudo os resultados até este ponto são inconclusivos, pois a o cálculo de ABF ainda não atingiu convergência (Figura 4.27).

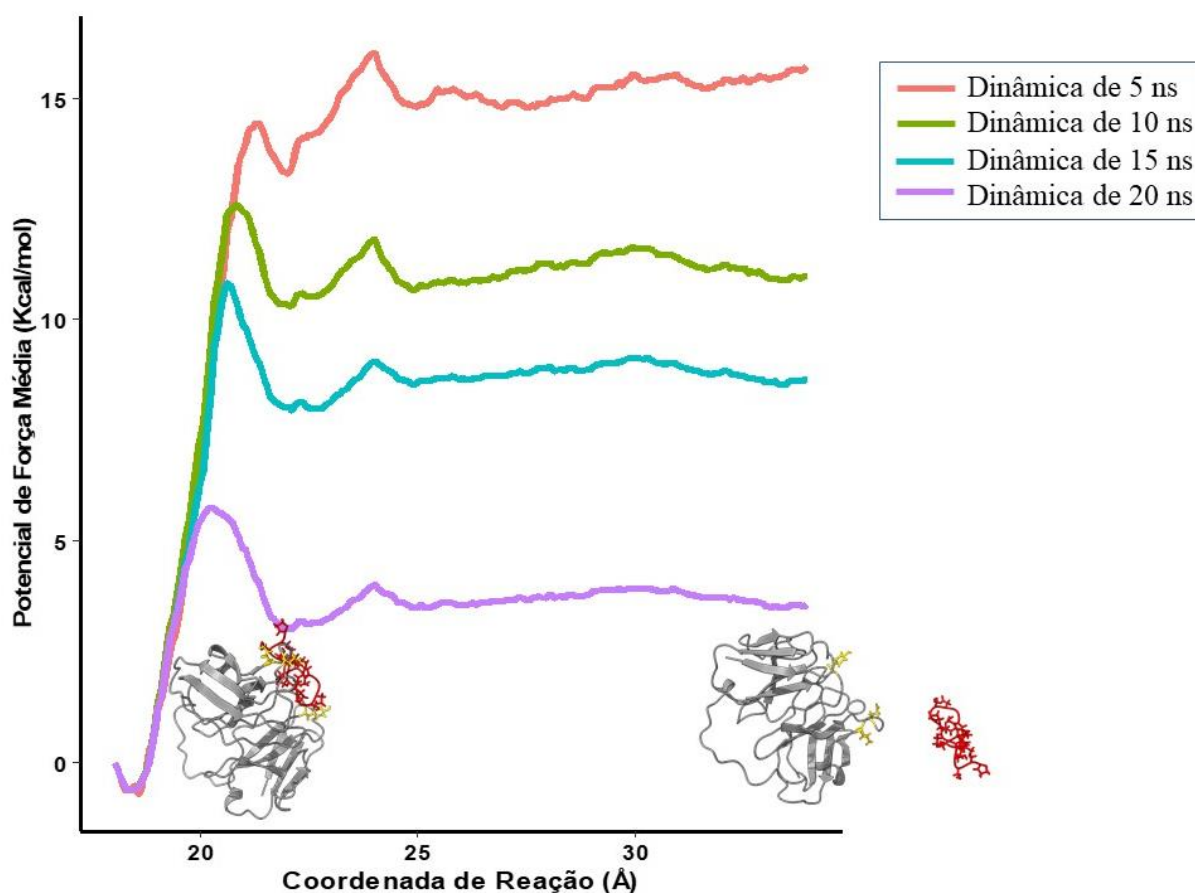


Figura 4.27 - Perfil do potencial da força média em função da coordenada de reação de dissociação entre o anticorpo mutante e o HBsAg.

Durante a dinâmica molecular de 5 ns, o valor de energia livre de ligação foi de -16,42 kcal/mol. Quando o sistema evoluiu até 10 ns, o valor obtido foi de -11,59 kcal/mol. A diferença entre estes dois valores é quantificada por um RMSE de 3,28

kcal/mol. Quando a amostragem do sistema evolui por 5 ns adicionais, o ΔG_{LIG} obtido então é de -9,26 kcal/mol, ao que se associa um RMSE de 2,07 kcal/mol em relação à simulação de 10 ns. Finalmente, quando o sistema evoluiu até 20 ns, o ΔG_{LIG} (inacurado) fornecido pelo método foi de -4,17 kcal/mol, com um RMSE de 4,39 kcal/mol em relação à simulação de 15 ns (Figuras 4.27 e 4.28).

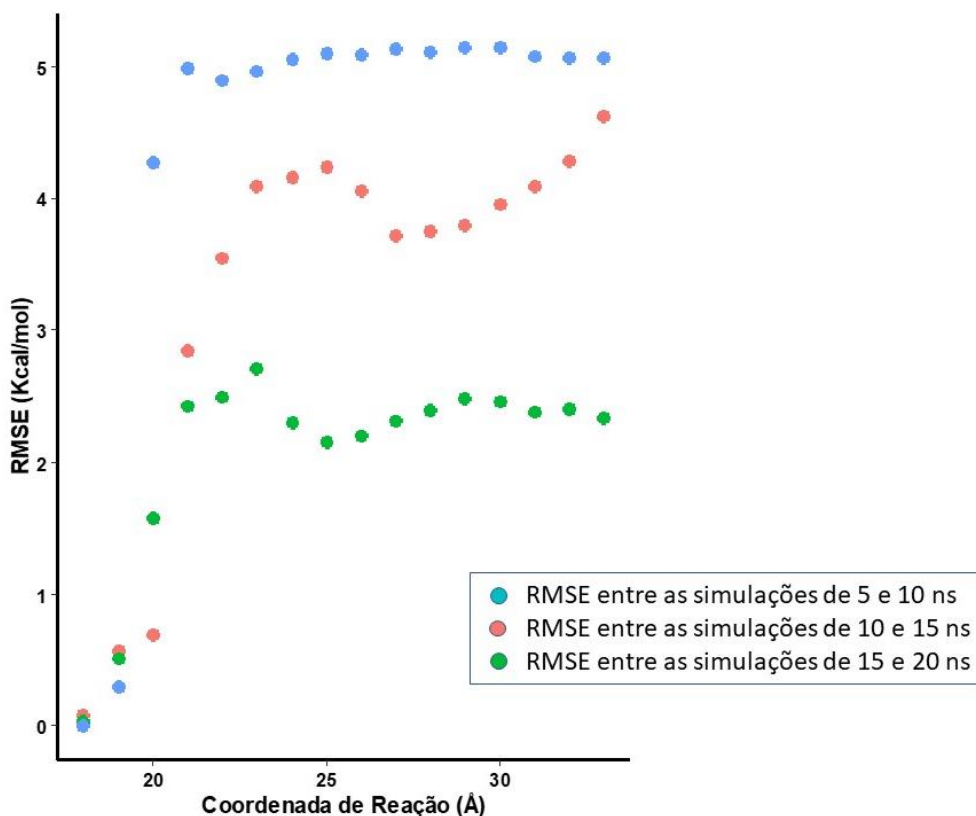


Figura 4.28 - Erro quadrático médio entre os valores de energia calculados para as simulações com o complexo anticorpo mutante-HBsAg.

As possíveis razões para esta não-convergência no cálculo de energia livre incluem:

- A necessidade de representar o processo usando mais de uma variável coletiva: conforme foi brevemente explicado nas seções 4.3.1 e 4.4, a escolha incorreta e/ou um número insuficiente de variáveis coletivas para estudar um processo físico pode resultar em uma pobre estimativa de $\langle F_{\xi} \rangle$, gerando imprecisões nos resultados, devido a graus de liberdade ortogonais a ξ que não são explorados, mas contribuem para a descrição termodinâmica do processo em estudo (68);
- A necessidade de aumentar a ergodicidade do sistema, ou seja, uniformizar a amostragem ao longo de ξ – o que pode ser feito através da restrição de graus de liberdade pela inclusão de uma (ou mais) variável(is) coletiva(s), ou através

do aumento do tempo de simulação por janela de ξ , ou através de outras estratégias, como a de réplicas múltiplas (68);

- As mutações e/ou o modo de ligação proposto pelo RosettaScripts foram falhos, assim como no caso reportado por Procko e colaboradores (134) – o resultado fornecido superestimou a importância de superfícies hidrofóbicas na interface anticorpo-antígeno, gerando um mutante com afinidade menor que o anticorpo nativo.

Embora as três possíveis fontes de erros sejam razoáveis para a explicação da não-convergência do método, o fato de que os três sistemas anteriores mostraram convergência com ABF unidimensional, dentro da escala de tempo de até 15 ns por janela de ξ , pode sugerir que as mutações propostas pelo algoritmo do RosettaScripts foram desvantajosas para a afinidade do anticorpo 19CC6CG2; contudo, são necessárias novas simulações para tirar conclusões mais precisas.

As energias de interação e o RMSD do HBsAg também foram computadas ao longo da simulação inicial (Figuras 4.29 e 4.30).

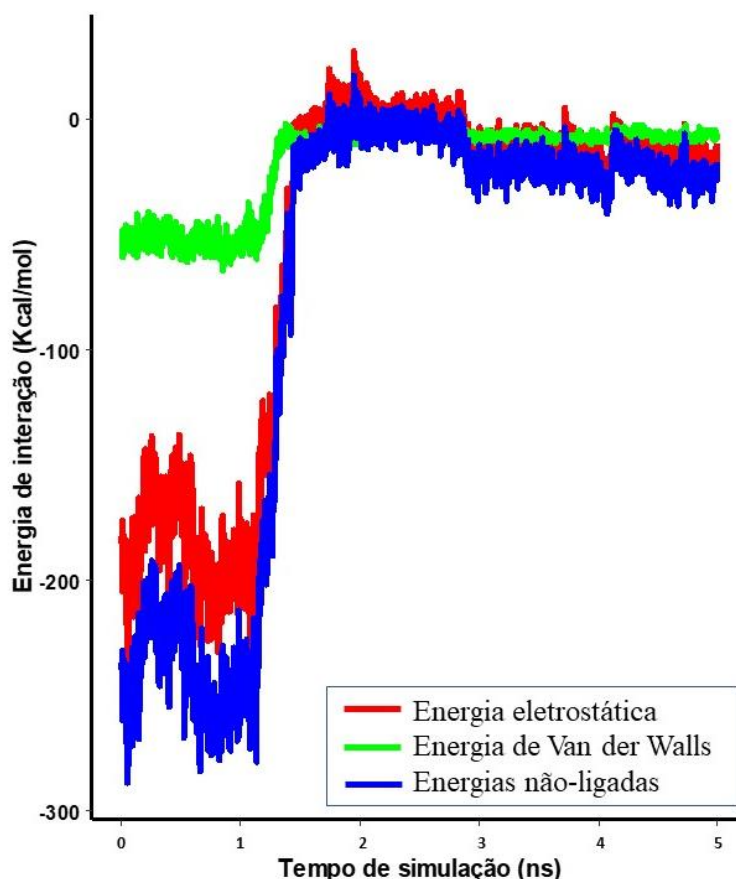


Figura 4.29 - Energias de interação entre o fragmento de anticorpo mutante e o HBsAg em função do tempo de simulação.

Apesar de haver indícios de que o RosettaScripts criou mutações desfavoráveis à interação com o HBsAg, os próximos resultados questionam esta hipótese, pois as energias de interação do mutante com o antígeno atingiram valores maiores (em módulo) do que aqueles obtidos para o complexo formado pelo HBsAg e o anticorpo nativo. Adicionalmente, as energias de interação entre o anticorpo mutante e o HBsAg levaram aproximadamente 500 ps a mais para chegar a zero em comparação com o anticorpo nativo, mediante a aplicação de $\langle F_{ABF} \rangle$.

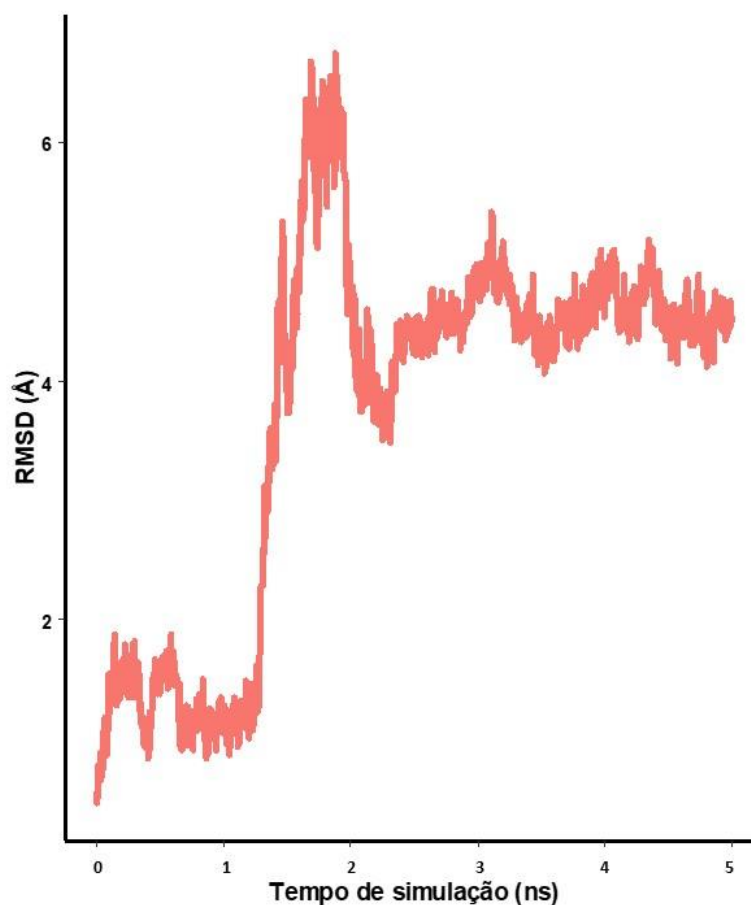


Figura 4.30 - Desvio Quadrático Médio do HBsAg em função da coordenada de reação.

O perfil de RMSD do antígeno mostra a sua evolução congruente à evolução das energias de interação, pois a depleção dessas energias de interação acontece no mesmo momento em que há um rápido aumento do RMSD, durante o segundo nanossegundo de simulação. De forma semelhante ao que foi observado nos três sistemas anteriores, as energias de interação desempenham papel importante na manutenção do HBsAg em sua conformação original, portanto quando estas energias chegam a zero, também o antígeno estabiliza sua conformação, em cerca de 5 Å.

A amostragem para este sistema também aumentou com o tempo de simulação (Figura 4.31). Mesmo com 20 ns de exploração do espaço-fase, o sistema ainda demonstrou sinais de comportamento não-ergódico: nos intervalos entre as duas primeiras janelas de ξ a amostragem está insuficiente. Com efeito, esta é uma das causas da não-convergência deste cálculo de energia livre. Embora o refinamento da simulação através do acréscimo de janelas extras nestes pontos de baixa amostragem possa ser pertinente para resolver este problema, Comer e colaboradores alertam contra a estratificação excessiva da coordenada de reação, indicando que isso pode dificultar ainda mais a exploração de ξ , principalmente nos limites entre as janelas (68).

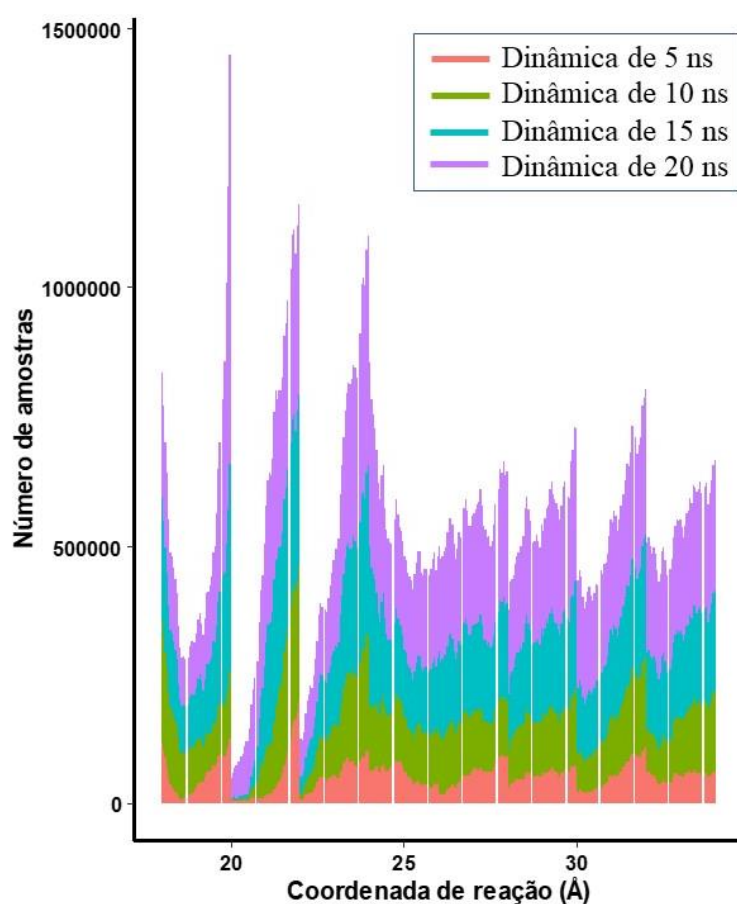


Figura 4.31 - Amostragem (μ máximo) sobre ξ coletada ao longo da dissociação entre o anticorpo mutante e o HBsAg.

Resultados adicionais que corroboram com a hipótese de que o anticorpo mutante de fato tem uma afinidade maior pelo antígeno são as análises de formação e extinção de ligações hidrogênio e pontes salinas ao longo da coordenada de reação (Figuras 4.32 e 4.33). Embora o novo modo de ligação atribuído pelo RosettaScripts ao anticorpo mutante tenha causado a perda do seu contato intermolecular mais

importante – entre o Asp118 do anticorpo e a Lys207 do antígeno –, o novo complexo anticorpo mutante-antígeno apresenta interações moleculares que se estendem por várias janelas de ξ , por motivos semelhantes àqueles descritos para as ligações hidrogênio do anticorpo nativo (seção 4.4).

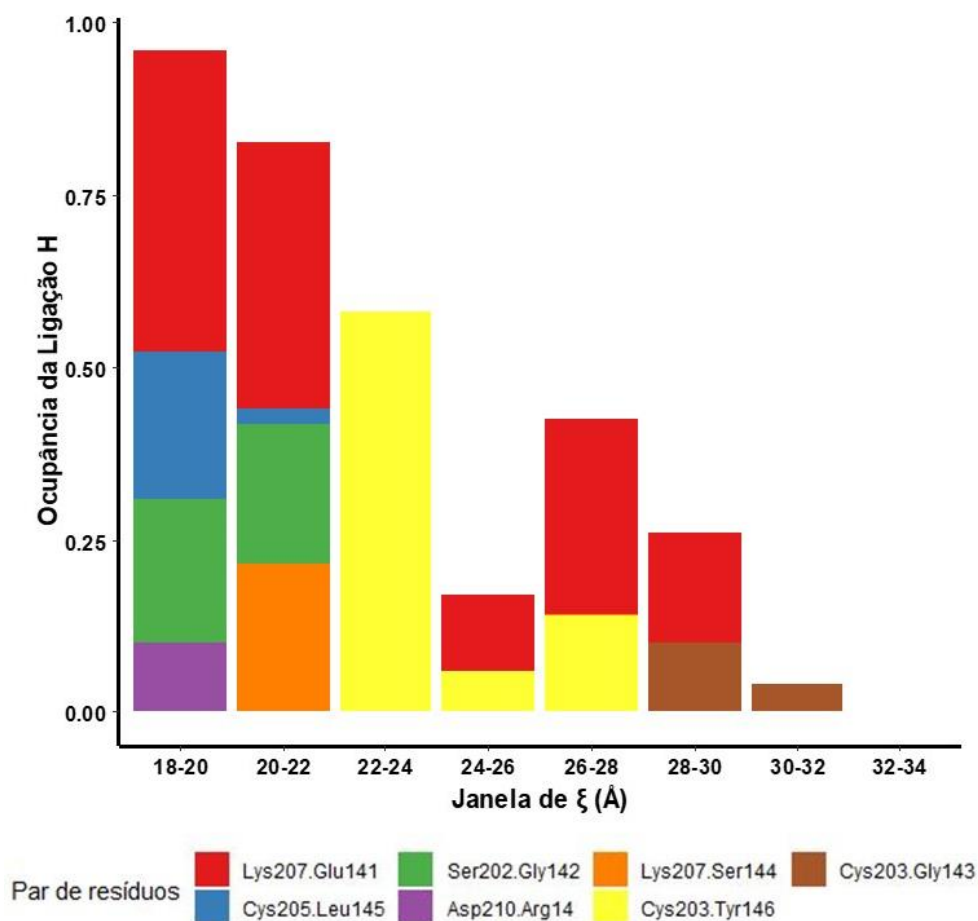


Figura 4.32 - Ligações Hidrogênio e sua ocupância de acordo com a coordenada de reação durante a simulação de 20 ns. Os resíduos do antígeno estão listados primeiro, seguidos do resíduo do anticorpo com o qual fazem contato.

A ligação hidrogênio entre Lys207 e Glu141, oriundo da mutação N141E, é mantida, mesmo com diminuta ocupância em quase todas as janelas da coordenada de reação. A ligação hidrogênio entre essa lisina do antígeno e Asp118 no anticorpo nativo também tem ocupância considerável (máximo de 41,57%) durante quatro janelas da dissociação entre as duas proteínas, mas é a única interação que permanece em mais de uma janela deste processo. Gly142 e Tyr146, vizinhas de resíduos que foram mutados, agora fazem ligações hidrogênio com alta ocupância que também permanecem em mais de uma janela de ξ .

Já a análise das pontes salinas ao longo de ξ dos anticorpos nativo e mutante mostra padrões semelhantes, nos quais Lys207 do antígeno perfaz uma ponte salina com um resíduo ácido do anticorpo – Asp118 no anticorpo nativo e Glu141 no anticorpo mutante, os quais distam 8.2 Å entre si. As pontes salinas em cada caso têm ocupâncias semelhantes distribuídas pela coordenada de reação.

As mutações e o *docking* do RosettaScripts fizeram o nitrogênio da Lys207 ajustar-se cerca de 1,9 Å, de forma a criar a nova ponte salina com o resíduo de glutamato 141. Esse deslocamento da cadeia lateral ocorreu devido ao empacotamento hidrofóbico das cadeias laterais entre os resíduos 203-206 (uma trinca de cisteínas seguida de uma treonina) do HBsAg com os resíduos 142-147 do anticorpo mutante. A provável causa deste empacotamento foi a substituição de um aminoácido polar por um apolar (mutação T145L) dentro dessa região majoritariamente hidrofóbica.

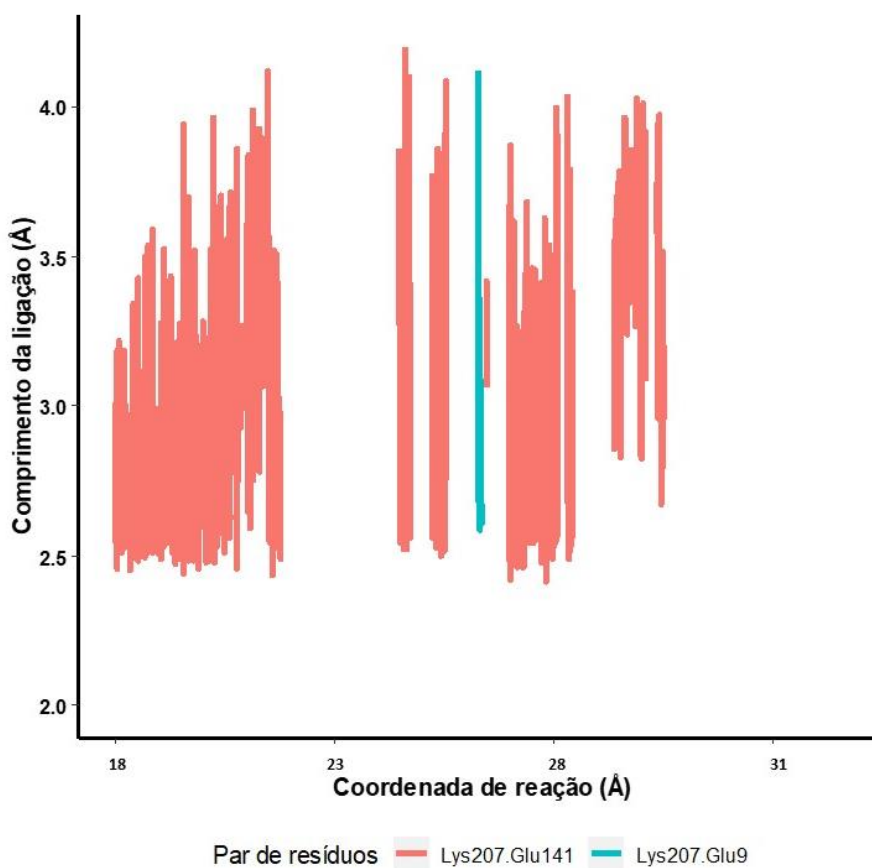


Figura 4.33 - Comprimento das pontes salinas intermoleculares de acordo com a coordenada de reação durante a simulação de 20 ns. Os resíduos do antígeno estão listados primeiro, seguidos do resíduo do anticorpo com o qual fazem contato.

5 PERSPECTIVAS

Os resultados encontrados neste trabalho destacam diversos aspectos da interação entre o anticorpo 19CC6CG2 e o antígeno de superfície do vírus da hepatite B; contudo simulações adicionais precisam ser realizadas para investigar com mais precisão os eventos que acontecem mediante o reconhecimento antígeno-anticorpo. O aumento da dimensionalidade das coordenadas de reação em um cálculo de ABF bi ou multidimensional, apesar de ser computacionalmente mais custoso e cujos resultados são mais complexos, poderia aumentar a acurácia dos valores de ΔG_{LIG} descritos neste trabalho, além de fornecer uma perspectiva mais sólida sobre as interações intermoleculares que dirigem a associação entre as duas proteínas. Uma estratégia mais simples que também poderia ser feita é o refinamento das janelas de ξ com baixos valores de amostragem de forma a aumentar a ergodicidade do sistema.

6 CONCLUSÕES

- O método MM-PBSA pode ser usado para estimar a energia livre da associação entre anticorpos e antígeno com acurácia razoável, dependendo da natureza do sistema e das mudanças conformacionais que ocorrem mediante o reconhecimento.
- O método ABF apresenta maior acurácia para a descrição da energia livre de ligação entre anticorpos e antígenos que o método MM-PBSA.
- As mutações W74Y, N141E e T145L podem ter influenciado a perda de afinidade do anticorpo 19CC6CG2, entretanto são necessárias novas simulações para determinar ao certo a sua contribuição para a afinidade frente ao HBsAg.
- A amostragem uniforme e a escolha de uma ou mais variável(is) coletiva(s) são essenciais para a descrição precisa de um processo biofísico como a associação entre um anticorpo e um antígeno.
- Os principais pontos fortes da metodologia explorada neste trabalho são a sua simplicidade e acurácia comparada a outros métodos computacionais de cálculo de ΔG .
- Entre as limitações e possíveis ameaças ao método, destacam-se associações anticorpo-antígeno que envolvem grandes modificações conformacionais e necessitam, portanto, de mais de uma variável coletiva ou de longos tempos de simulação para descrever e estudar o processo.
- A maior oportunidade destacada pela abordagem descrita neste trabalho é a oportunidade de correlacionar a extinção e a formação de interações intermoleculares com os eventos que ocorrem mediante a dissociação anticorpo-antígeno e com as variações de PMF do processo.

7 REFERÊNCIAS BIBLIOGRÁFICAS

1. Ministério da Saúde. Hepatites Virais HEPATITES VIRAIS Características clínicas e epidemiológicas Descrição. 15(Hepatite C):409–816. Available from: http://bvsmms.saude.gov.br/bvs/publicacoes/07_0044_M2.pdf
2. Secretaria de Vigilância em Saúde. Ministério da Saúde. Hepatites virais 2017. Bol Epidemiológico [Internet]. 2017;48(24):65. Available from: http://www.aids.gov.br/sites/default/files/anexos/publicacao/2016/59121/boletim_hepatites_05_08_2016_pdf_96185.pdf
3. World Health Organization. Global Hepatitis Report, 2017 [Internet]. World Health Organization. 2017. 62 p. Available from: <http://www.who.int/hepatitis>
4. Maiwall R, Kumar M. Review Article Prevention and Treatment of Recurrent Hepatitis B after Liver Transplantation. 2016;4:54–65.
5. Ku W, Wang U, Nguyen MH. Efficacy and effectiveness of anti-HBV therapy with early withdrawal of HBIG prophylaxis to prevent HBV recurrence following liver transplantation. Expert Opin Biol Ther [Internet]. 2015;15(5):665–77. Available from: <http://www.tandfonline.com/doi/full/10.1517/14712598.2015.1025045>
6. Shouval D, Samuel D. Hepatitis B immune globulin to prevent hepatitis B virus graft reinfection following liver transplantation: a concise review. Hepatology. 2000;32(1c):1189–95.
7. Ferreira MS, Borges AS. Avanços no tratamento da hepatite pelo vírus B. Rev Soc Bras Med Trop. 2007;40(4):451–62.
8. Cerino A, Bremer CM, Glebe D, Mondelli MU. A Human Monoclonal Antibody against Hepatitis B Surface Antigen with Potent Neutralizing Activity. PLoS One. 2015;10(4):e0125704.
9. Gilbert RJ, Beales L, Blond D, Simon MN, Lin BY, Chisari F V, et al. Hepatitis B small surface antigen particles are octahedral. Proc Natl Acad Sci U S A [Internet]. 2005;102(41):14783–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16203986>
10. Peterson DL. The Structure of Hepatitis B Surface Antigen and Its Antigenic Sites. 1987;6(6):258–62.
11. Cassidy A, Mossman S, Olivieri A, Ridder M De, Leroux-Roels G. Hepatitis B vaccine effectiveness in the face of global HBV genotype diversity. Expert Review of Vaccines. 2011.
12. Robert Neurath A, Seto B, Strick N. Antibodies to synthetic peptides from the

- preS1 region of the hepatitis B virus (HBV) envelope (env) protein are virus-neutralizing and protective. *Vaccine*. 1989;
13. Eren R, Ilan E, Nussbaum O, Lubin I, Terkieltaub D, Arazi Y, et al. Preclinical evaluation of two human anti-hepatitis B virus (HBV) monoclonal antibodies in the HBV-trimera mouse model and in HBV chronic carrier chimpanzees. *Hepatology*. 2000;
 14. Glebe D, Aliakbari M, Krass P, Knoop E V., Valerius KP, Gerlich WH. Pre-S1 Antigen-Dependent Infection of Tupaia Hepatocyte Cultures with Human Hepatitis B Virus. *J Virol*. 2003;
 15. Surface-antigen HB V. Neutralizing B Virus (HBV) Immunoglobulins by Epitope Mapping with Recombinant. *Microbiol Immunol*. 2000;44(8):703–10.
 16. Abbas, Abul; Abbas A. *Imunologia - Abbas 7^a Ed.pdf*. 2012. p. 560.
 17. Nelson AL, Dhimolea E, Reichert JM. Development trends for human monoclonal antibody therapeutics. *Nature Reviews Drug Discovery*. 2010.
 18. Carter PJ, Lazar GA. Next generation antibody drugs: Pursuit of the “high-hanging fruit.” *Nature Reviews Drug Discovery*. 2018.
 19. Fischman S, Ofra Y. *A Antibody Modeling, Engineering, and Design*. 2018; Available from: https://doi.org/10.1007/978-3-642-35943-9_10083-1
 20. Cambrosio A, Keating P. “Going Monoclonal”: Art, Science, and Magic in the Day-to-Day Use of Hybridoma Technology. *Soc Probl*. 1988;
 21. Buss NAPS, Henderson SJ, McFarlane M, Shenton JM, De Haan L. Monoclonal antibody therapeutics: History and future. *Current Opinion in Pharmacology*. 2012.
 22. Filpula D. Antibody engineering and modification technologies. *Biomolecular Engineering*. 2007.
 23. Ecker DM, Jones SD, Levine HL. The therapeutic monoclonal antibody market. *mAbs*. 2015.
 24. Porstmann T, Kiessig ST. Enzyme immunoassay techniques an overview. *J Immunol Methods*. 1992;
 25. Shin YW, Ryoo KH, Hong KW, Chang KH, Choi JS, So M, et al. Human monoclonal antibody against Hepatitis B virus surface antigen (HBsAg). *Antiviral Res*. 2007;75(2):113–20.
 26. Golsaz-Shirazi F, Amiri MM, Farid S, Bahadori M, Bohne F, Altstetter S, et al. Construction of a hepatitis B virus neutralizing chimeric monoclonal antibody recognizing escape mutants of the viral surface antigen (HBsAg). *Antiviral Res*

- [Internet]. Elsevier B.V.; 2017;144:153–63. Available from: <http://dx.doi.org/10.1016/j.antiviral.2017.06.013>
27. Alves Vianna CO, da Silva e Mouta Júnior S, da Glória Teixeira Martins M, Batoreu NM, Queiroz JL, Gomes SA, et al. Evaluation of murine monoclonal antibodies targeting different epitopes of the hepatitis B virus surface antigen by using immunological as well as molecular biology and biochemical approaches. *J Immunol Methods*. 2006;313(1–2):38–47.
 28. Wark KL, Hudson PJ. Latest technologies for the enhancement of antibody affinity. *Advanced Drug Delivery Reviews*. 2006.
 29. Carlsten M, Levy E, Karambelkar A, Li L, Reger R, Berg M, et al. Efficient mRNA-based genetic engineering of human NK cells with high-affinity CD16 and CCR7 augments rituximab-induced ADCC against lymphoma and targets NK cell migration toward the lymph node-associated chemokine CCL19. *Front Immunol*. 2016;
 30. Beck A, Goetsch L, Dumontet C, Corvaia N. Strategies and challenges for the next generation of antibody-drug conjugates. *Nature Reviews Drug Discovery*. 2017.
 31. Ebrahimi Z, Asgari S, Ahangari Cohan R, Hosseinzadeh R, Hosseinzadeh G, Arezumand R. Rational affinity enhancement of fragmented antibody by ligand-based affinity improvement approach. *Biochem Biophys Res Commun* [Internet]. 2018;(xxxx). Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0006291X18322939>
 32. IUPAC. Gold Book IUPAC. IUPAC. 2014.
 33. Verli H. *Bioinformática da Biologia à Flexibilidade Molecular*. 2014. 282 p.
 34. Allen MP, Tildesley DJ. *Computer simulation of liquids: Second edition*. Computer Simulation of Liquids: Second Edition. 2017.
 35. Namba AM, Silva VB, Silva CHTP. *Dinâmica molecular : teoria e aplicações em planejamento de fármacos*. 2008;33:13–23.
 36. Hu JP, Gong XQ, Su JG, Chen WZ, Wang CX. Study on the molecular mechanism of inhibiting HIV-1 integrase by EBR28 peptide via molecular modeling approach. *Biophys Chem*. 2008;
 37. Lavecchia A, Cosconati S, Novellino E, Calleri E, Temporini C, Massolini G, et al. Exploring the molecular basis of the enantioselective binding of penicillin G acylase towards a series of 2-aryloxyalkanoic acids: A docking and molecular dynamics study. *J Mol Graph Model*. 2007;

38. Capriles PVSZ, Baptista LPR, Guedes IA, Guimarães ACR, Custódio FL, Alves-Ferreira M, et al. Structural modeling and docking studies of ribose 5-phosphate isomerase from *Leishmania major* and *Homo sapiens*: A comparative analysis for Leishmaniasis treatment. *J Mol Graph Model*. 2015;
39. English NJ. Calculation of binding affinities of HIV-1 RT and ??-secretase inhibitors using the linear interaction energy method with explicit and continuum solvation approaches. *J Mol Model*. 2007;
40. Joti Y, Nakagawa H, Kataoka M, Kitao A. Hydration-dependent protein dynamics revealed by molecular dynamics simulation of crystalline staphylococcal nuclease. *J Phys Chem B*. 2008;
41. Díaz-Vergara N, Piñeiro Á. Molecular dynamics study of triosephosphate isomerase from *Trypanosoma cruzi* in water/decane mixture. *J Phys Chem B*. 2008;
42. Hénin J, Pohorille A, Chipot C. Insights into the recognition and association of transmembrane α -helices. The free energy of α -helix dimerization in glycophorin A. *J Am Chem Soc*. 2005;127(23):8478–84.
43. Paloni M, Cavallotti C. Molecular modeling of the interaction of protein L with antibodies. *ACS Omega*. 2017;2(10):6464–72.
44. Chong LT, Duan Y, Wang L, Massova I, Kollman PA. Molecular dynamics and free-energy calculations applied to affinity maturation in antibody 48G7. *Proc Natl Acad Sci [Internet]*. 1999;96(25):14330–5. Available from: <http://www.pnas.org/cgi/doi/10.1073/pnas.96.25.14330>
45. Gumbart JC, Chipot C. Efficient Determination of Protein – Protein Standard Binding Free Energies from First Principles. *J Chem Theory Comput*. 2013;xxxx(xxxx):xxxx.
46. Zwanzig RW. High† $\frac{1}{2}$ Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *J Chem Phys*. 1954;22(8):1420–6.
47. Park S, Khalili-Araghi F, Tajkhorshid E, Schulten K. Free energy calculation from steered molecular dynamics simulations using Jarzynski’s equality. *J Chem Phys*. 2003;119(6):3559–66.
48. Torrie GM, Valleau JP. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J Comput Phys*. 1977;23(2):187–99.
49. Lapelosa M. Free Energy of Binding and Mechanism of Interaction for the MEEVD-TPR2A Peptide-Protein Complex. *J Chem Theory Comput*. 2017;13(9):4514–23.

50. Srinivasan J, Cheatham TE, Cieplak P, Kollman PA, Case DA. Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate-DNA helices. *J Am Chem Soc.* 1998;
51. Homeyer N, Gohlke H. Free energy calculations by the Molecular Mechanics Poisson-Boltzmann Surface Area method. *Mol Inform.* 2012;31(2):114–22.
52. Genheden S, Ryde U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin Drug Discov [Internet].* 2015;0441(October):1–13. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25835573> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4487606&tool=pmcentrez&rendertype=abstract>
53. Wang J, Morin P, Wang W, Kollman PA. Use of MM-PBSA in reproducing the binding free energies to HIV-1 RT of TIBO derivatives and predicting the binding mode to HIV-1 RT of efavirenz by docking and MM-PBSA. *J Am Chem Soc.* 2001;
54. Gohlke H, Case D a, Biology M, Scripps T, Rd NTP. Converging Free Energy Estimates: MM-PB (GB) SA Studies on the Protein – Protein Complex Ras – Raf. *J Comput Chem.* 2003;25:238–50.
55. Thompson DC, Humblet C, Joseph-McCarthy D. Investigation of MM-PBSA rescoring of docking poses. *J Chem Inf Model.* 2008;
56. Steinbrecher T, Case DA, Labahn A. A multistep approach to structure-based drug design: Studying ligand binding at the human neutrophil elastase. *J Med Chem.* 2006;
57. Lindström A, Edvinsson L, Johansson A, Andersson CD, Andersson IE, Raubacher F, et al. Postprocessing of Docked Protein-Ligand Complexes Using Implicit Solvation Models. *J Chem Inf Model.* 2011;
58. Darve E, Pohorille A. Calculating free energies using average force. *J Chem Phys.* 2001;
59. Huo S, Wang J, Cieplak P, Kollman PA, Kuntz ID. Molecular dynamics and free energy analyses of cathepsin D-inhibitor interactions: Insight into structure-based ligand design. *J Med Chem.* 2002;
60. Kuhn B, Kollman PA. Binding of a diverse set of ligands to avidin and streptavidin: An accurate quantitative prediction of their relative affinities by a combination of molecular mechanics and continuum solvent models. *J Med Chem.* 2000;

61. Hou T, Wang J, Li Y, Wang W. Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J Chem Inf Model*. 2011;
62. Yang T, Wu JC, Yan C, Wang Y, Luo R, Gonzales MB, et al. Virtual screening using molecular simulations. *Proteins Struct Funct Bioinforma*. 2011;
63. Lee MR, Tsai J, Baker D, Kollman PA. Molecular Dynamics in the Endgame of Protein Structure Prediction. *J Mol Biol*. 2001;
64. Chipot C, Pohorille A. Free Energy Calculations: Theory and Applications in Chemistry and Biology (Springer Series in Chemical Physics). *J Am Chem Soc* 2007;299:1-100. ;
65. Git--- V, Bernardi R, Bhandarkar M, Bhatele A, Bohm E, Brunner R, et al. NAMD User ' s Guide. 2017;
66. Henin J, Fiorin G, Chipot C, Klein M. {E}xploring {M}ultidimensional {F}ree {E}nergy {L}andscapes {U}sing {T}ime{-D}ependent {B}iases on {C}ollective {V}ariables. *J Chem Theor Comp*. 2010;6:35–47.
67. Bernardi RC, Melo MCR, Schulten K. Enhanced sampling techniques in molecular dynamics simulations of biological systems. *Biochim Biophys Acta - Gen Subj* [Internet]. Elsevier B.V.; 2015;1850(5):872–7. Available from: <http://dx.doi.org/10.1016/j.bbagen.2014.10.019>
68. Comer J, Gumbart JC, Hénin J, Lelièvre T, Pohorille A, Chipot C. The {Adaptive} {Biasing} {Force} {Method}: {Everything} {You} {Always} {Wanted} {To} {Know} but {Were} {Afraid} {To} {Ask}. *J Phys Chem B* [Internet]. 2015;119(3):1129–51. Available from: <http://dx.doi.org/10.1021/jp506633n>
69. Shu L, Qi CF, Schlom J, Kashmiri S V. Secretion of a single-gene-encoded immunoglobulin from myeloma cells. *Proc Natl Acad Sci U S A*. 1993;
70. McCafferty J, Griffiths AD, Winter G, Chiswell DJ. Phage antibodies: filamentous phage displaying antibody variable domains. *Nature*. 1990. p. 552–4.
71. Chowdhury PS, Wu H. Tailor-made antibody therapeutics. *Methods*. 2005;36(1):11–24.
72. Bird RE, Hardman KD, Jacobson JW, Johnson S, Kaufman BM, Lee SM, et al. Single-chain antigen-binding proteins. *Science* (80-). 1988;
73. Levin AM, Weiss GA. Optimizing the affinity and specificity of proteins with molecular display. *Mol Biosyst*. 2006;2(1):49–57.
74. Kobayashi N, Shibahara K, Ikegashira K, Shibusawa K, Goto J. Single-chain Fv fragments derived from an anti-11-deoxycortisol antibody: Affinity, specificity,

- and idiotype analysis. *Steroids*. 2002;67(8):733–42.
75. Vries SJ De, Dijk M Van, Bonvin AMJJ. The HADDOCK web server for data-driven biomolecular docking. 2010;5(5):883–97.
 76. Van Zundert GCP, Rodrigues JPGLM, Trellet M, Schmitz C, Kastiris PL, Karaca E, et al. The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J Mol Biol*. 2016;
 77. Dominguez C, Boelens R, Bonvin AMJJ. HADDOCK: A Protein - Protein Docking Approach Based on Biochemical or Biophysical Information. 2003;(2):1731–7.
 78. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*. 2015;
 79. Mackerell AD, Feig M, Brooks CL. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulation. *J Comput Chem*. 2004;
 80. Tom Darden, Darrin York and LP. Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems. *J Chem Phys*. 1993;98(12):10089–92.
 81. Huang - *Statistical Mechanics - 2nd Edition - John Wiley & Sons.pdf*.
 82. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J Chem Phys*. 1983;
 83. Olivei A. Surface imaging by scanning-laser-induced voltage. *J Phys D Appl Phys*. 1976;9(2):183–95.
 84. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. LINCS: A Linear Constraint Solver for molecular simulations. *J Comput Chem*. 1997;
 85. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A* [Internet]. 2001 Aug 28 [cited 2016 Sep 16];98(18):10037–41. Available from: <http://www.pnas.org/cgi/content/short/98/18/10037>
 86. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. 2000;28(1):235–42.
 87. Martínez-Sáez N, Castro-López J, Valero-González J, Madariaga D, Compañón I, Somovilla VJ, et al. Deciphering the Non-Equivalence of Serine and Threonine O-Glycosylation Points: Implications for Molecular Recognition of the Tn Antigen by an anti-MUC1 Antibody. *Angew Chemie - Int Ed*. 2015;54(34):9830–4.

88. Kudo S, Caaveiro JMM, Nagatoishi S, Miyafusa T, Matsuura T, Sudou Y, et al. Disruption of cell adhesion by an antibody targeting the cell-adhesive intermediate (X-dimer) of human P-cadherin. *Sci Rep* [Internet]. Nature Publishing Group; 2017;7(July 2016):1–15. Available from: <http://dx.doi.org/10.1038/srep39518>
89. Spoel DVANDER, Lindahl E, Hess B, Groenhof G. GROMACS: Fast, Flexible, and Free. 2005;
90. William Humphrey, Andrew Dalke KS. VMD: Visual Molecular Dynamics. *J Mol Graph*. 1996;(14):33–8.
91. Road T. Molecular Modeling and Affinity Determination of scFv Antibody : Proper Linker Peptide Enhances Its Activity. 2009;
92. Huang J, Mackerell AD. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J Comput Chem*. 2013;
93. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. *Articulo_Gracia_Lopez*. 1983;4(2):187–217.
94. Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys*. 1977;23(3):327–41.
95. Fu H, Shao X, Chipot C, Cai W. Extended Adaptive Biasing Force Algorithm. An On-the-Fly Implementation for Accurate Free-Energy Calculations. *J Chem Theory Comput*. 2016;12(8):3506–13.
96. Mihailescu M, Gilson MK. On the theory of noncovalent binding. *Biophys J* [Internet]. Elsevier; 2004;87(1):23–36. Available from: <http://dx.doi.org/10.1529/biophysj.103.031682>
97. Fleishman SJ, Leaver-Fay A, Corn JE, Strauch EM, Khare SD, Koga N, et al. Rosettascripts: A scripting language interface to the Rosetta Macromolecular modeling suite. *PLoS One*. 2011;6(6):1–10.
98. Lewis S, Stranges P Ben. Rosetta 3.4. 2018;8–10.
99. O’Meara MJ, Leaver-Fay A, Tyka MD, Stein A, Houlihan K, Dimaio F, et al. Combined covalent-electrostatic model of hydrogen bonding improves structure prediction with Rosetta. *J Chem Theory Comput*. 2015;
100. Smith CA, Kortemme T. Backrub-Like Backbone Simulation Recapitulates Natural Protein Conformational Variability and Improves Mutant Side-Chain Prediction. *J Mol Biol*. 2008;380(4):742–56.
101. Davis IW, Arendall WB, Richardson DC, Richardson JS. The backrub motion:

- How protein backbone shrugs when a sidechain dances. *Structure*. 2006;14(2):265–74.
102. Regep C, Georges G, Shi J, Popovic B, Deane CM. The H3 loop of antibodies shows unique structural characteristics. *Proteins Struct Funct Bioinforma*. 2017;
103. Pedotti M, Simonelli L, Livoti E, Varani L. Computational docking of antibody-antigen complexes, opportunities and pitfalls illustrated by influenza hemagglutinin. *Int J Mol Sci*. 2011;12(1):226–51.
104. Schütz AK, Soragni A, Hornemann S, Aguzzi A, Ernst M, Böckmann A, et al. The amyloid-congo red interface at atomic resolution. *Angew Chemie - Int Ed*. 2011;50(26):5956–60.
105. van der Poel CE, Spaapen RM, van de Winkel JGJ, Leusen JHW. Functional Characteristics of the High Affinity IgG Receptor, Fc RI. *J Immunol [Internet]*. 2011;186(5):2699–704. Available from: <http://www.jimmunol.org/cgi/doi/10.4049/jimmunol.1003526>
106. Bochkov YA, Watters K, Ashraf S, Griggs TF, Devries MK, Jackson DJ, et al. Cadherin-related family member 3, a childhood asthma susceptibility gene product, mediates rhinovirus C binding and replication. *Proc Natl Acad Sci [Internet]*. 2015;112(17):5485–90. Available from: <http://www.pnas.org/lookup/doi/10.1073/pnas.1421178112>
107. Pruneda JN, Littlefield PJ, Soss SE, Nordquist KA, Chazin WJ, Brzovic PS, et al. Structure of an E3:E2~Ub Complex Reveals an Allosteric Mechanism Shared among RING/U-box Ligases. *Mol Cell [Internet]*. Elsevier; 2012;47(6):933–42. Available from: <http://dx.doi.org/10.1016/j.molcel.2012.07.001>
108. Hausmann J, Kamtekar S, Christodoulou E, Day JE, Wu T, Fulkerson Z, et al. Structural basis of substrate discrimination and integrin binding by autotaxin. *Nat Struct Mol Biol [Internet]*. Nature Publishing Group; 2011;18(2):198–205. Available from: <http://dx.doi.org/10.1038/nsmb.1980>
109. Attallah C, Aguilar MF, Garay AS, Herrera FE, Etcheverrigaray M, Oggero M, et al. An unusual cysteine VL87 affects the antibody fragment conformations without interfering with the disulfide bond formation. *Mol Immunol [Internet]*. Elsevier; 2017;90(April):143–9. Available from: <http://dx.doi.org/10.1016/j.molimm.2017.07.008>
110. Payandeh Z, Rajabibazl M, Mortazavi Y, Rahimpour A, Taramchi AH. Ofatumumab monoclonal antibody affinity maturation through in silico modeling. *Iran Biomed J*. 2018;22(3):180–92.

111. Hou T, Wang J, Li Y, Wang W. Assessing the performance of the molecular mechanics/Poisson Boltzmann surface area and molecular mechanics/generalized Born surface area methods. II. the accuracy of ranking poses generated from docking. *J Comput Chem*. 2011;
112. Srivastava HK, Sastry GN. Molecular dynamics investigation on a series of HIV protease inhibitors: Assessing the performance of MM-PBSA and MM-GBSA approaches. *J Chem Inf Model*. 2012;
113. Cappel D, Hall ML, Lenselink EB, Beuming T, Qi J, Bradner J, et al. Relative Binding Free Energy Calculations Applied to Protein Homology Models. *J Chem Inf Model*. 2016;
114. Lafont V, Schaefer M, Stote RH, Altschuh D, Dejaegere A. Protein–Protein Recognition and Interaction Hot Spots in an Antigen–Antibody Complex: Free Energy Decomposition Identifies “Efficient Amino Acids.” *PROTEINS Struct Funct Bioinforma*. 2007;5(67):418–34.
115. Martins SA, Perez MAS, Moreira IS, Sousa SF, Ramos MJ, Fernandes PA. Computational alanine scanning mutagenesis: MM-PBSA vs TI. *J Chem Theory Comput*. 2013;9(3):1311–9.
116. Krissinel E, Henrick K. Inference of Macromolecular Assemblies from Crystalline State. *J Mol Biol*. 2007;372(3):774–97.
117. Estatística Básica - Morettin, Bussab - 5ª edição.pdf.
118. Malmqvist M. Surface plasmon resonance for detection and measurement of antibody-antigen affinity and kinetics. *Curr Opin Immunol*. 1993;5(2):282–6.
119. Dhanik A, McMurray JS, Kaviraki LE. Binding Modes of Peptidomimetics Designed to Inhibit STAT3. *PLoS One*. 2012;7(12):34–9.
120. Zeder-Lutz G, Zuber E, Witz J, Van Regenmortel MHV. Thermodynamic analysis of antigen-antibody binding using biosensor measurements at different temperatures. *Anal Biochem*. 1997;
121. Swaminathan CP, Nandi A, Visweswariah SS, Surolia A. Thermodynamic analyses reveal role of water release in epitope recognition by a monoclonal antibody against the human guanylyl cyclase C receptor. *J Biol Chem*. 1999;
122. Fu H, Cai W, Hénin J, Roux B, Chipot C. New Coarse Variables for the Accurate Determination of Standard Binding Free Energies. *J Chem Theory Comput*. 2017;
123. Xu J, Crowley MF, Smith JC. Building a foundation for structure-based cellulosome design for cellulosic ethanol: Insight into cohesin-dockerin

- complexation from computer simulation. *Protein Sci.* 2009;18(5):949–59.
124. Souza PCT, Puhl AC, Martínez L, Aparício R, Nascimento AS, Figueira ACM, et al. Identification of a New Hormone-Binding Site on the Surface of Thyroid Hormone Receptor. *Mol Endocrinol [Internet]*. 2014;28(4):534–45. Available from: <https://academic.oup.com/mend/article-lookup/doi/10.1210/me.2013-1359>
 125. Kellogg EH, Leaver-Fay A, Baker D. Role of conformational sampling in computing mutation-induced changes in protein structure and stability. *Proteins Struct Funct Bioinforma.* 2011;
 126. Units in Rosetta [Internet]. [cited 2019 Jan 10]. Available from: https://www.rosettacommons.org/docs/latest/rosetta_basics/Units-in-Rosetta
 127. King C. Computational Design of Protein Therapeutics with Reduced Immunogenicity through Structural Modeling of Protein Interactions. 2014;
 128. Jardine J, Julien J-P, Menis S, Ota T, Kalyuzhniy O, McGuire A, et al. Rational HIV Immunogen Design to Target Specific Germline B Cell Receptors. *Science (80-)*. 2013;340(2):6.
 129. Lewis SM, Wu X, Pustilnik A, Sereno A, Huang F, Rick HL, et al. Generation of bispecific IgG antibodies by structure-based design of an orthogonal Fab interface. *Nat Biotechnol [Internet]*. Nature Publishing Group; 2014;32(2):191–8. Available from: <http://dx.doi.org/10.1038/nbt.2797>
 130. Whitehead TA, Chevalier A, Song Y, Dreyfus C, Fleishman SJ, De Mattos C, et al. Optimization of affinity, specificity and function of designed influenza inhibitors using deep sequencing. *Nat Biotechnol [Internet]*. Nature Publishing Group; 2012;30(6):543–8. Available from: <http://dx.doi.org/10.1038/nbt.2214>
 131. Procko E, Berguig GY, Shen BW, Song Y, Frayo S, Convertine AJ, et al. A computationally designed inhibitor of an Epstein-Barr viral Bcl-2 protein induces apoptosis in infected cells. *Cell [Internet]*. Elsevier Inc.; 2014;157(7):1644–56. Available from: <http://dx.doi.org/10.1016/j.cell.2014.04.034>
 132. Jha RK, Gaiotto T, Bradbury ARM, Strauss CEM. An improved Protein G with higher affinity for human/rabbit IgG Fc domains exploiting a computationally designed polar network. *Protein Eng Des Sel.* 2014;27(4):127–34.
 133. Tinberg CE, Khare SD, Dou J, Doyle L, Nelson JW, Schena A, et al. Computational design of ligand-binding proteins with high affinity and selectivity. *Nature [Internet]*. Nature Publishing Group; 2013;501(7466):212–6. Available from: <http://dx.doi.org/10.1038/nature12443>
 134. Procko E, Hedman R, Hamilton K, Seetharaman J, Fleishman SJ, Su M, et al.

Computational design of a protein-based enzyme inhibitor. J Mol Biol [Internet]. Elsevier B.V.; 2013;425(18):3563–75. Available from: <http://dx.doi.org/10.1016/j.jmb.2013.06.035>