

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA
DOUTORADO EM QUÍMICA

RENAN FARIA GUERRA

**MODELAGEM MOLECULAR DAS INTERAÇÕES DO COMPLEXO
NS1/ANTI-NS1 PARA APLICAÇÃO NO DIAGNÓSTICO DIFERENCIAL DE
FLAVIROSES**

UBERLÂNDIA
2021

RENAN FARIA GUERRA

**MODELAGEM MOLECULAR DAS INTERAÇÕES DO COMPLEXO
NS1/ANTI-NS1 PARA APLICAÇÃO NO DIAGNÓSTICO DIFERENCIAL DE
FLAVIROSES**

Tese de doutorado apresentada à Coordenação do
Curso de Pós-Graduação em Química, como
requisito parcial para a obtenção do Título de Doutor
em Química.

Área de Concentração: Físico-Química

Orientador: Prof. Dr. Eduardo de Faria Franca

UBERLÂNDIA
2021

Ficha Catalográfica Online do Sistema de Bibliotecas da UFU
com dados informados pelo(a) próprio(a) autor(a).

| | |
|--------------|---|
| G934 2021 | <p>Guerra, Renan Faria, 1992- Modelagem molecular das interações do complexo NS1/anti-NS1 para aplicação no diagnóstico diferencial de Flaviroses [recurso eletrônico] / Renan Faria Guerra. - 2021.</p> <p>Orientador: Eduardo de Faria Franca. Tese (Doutorado) - Universidade Federal de Uberlândia, Pós-graduação em Química. Modo de acesso: Internet. Disponível em: http://doi.org/10.14393/ufu.te.2021.124 Inclui bibliografia. Inclui ilustrações.</p> <p>1. Química. I. Franca, Eduardo de Faria, 1981-, (Orient.). II. Universidade Federal de Uberlândia. Pós- graduação em Química. III. Título.</p> |
|--------------|---|

CDU: 54

Bibliotecários responsáveis pela estrutura de acordo com o AACR2:

Gizele Cristine Nunes do Couto - CRB6/2091



UNIVERSIDADE FEDERAL DE UBERLÂNDIA

Coordenação do Programa de Pós-Graduação em Química
Av. João Naves de Ávila, 2121, Bloco 5I - Bairro Santa Mônica, Uberlândia-MG, CEP
38400-902
Telefone: (34) 3239-4385 - www.cpgquimica.iq.ufu.br - cpgquimica@ufu.br



ATA

| | | | | | |
|------------------------------------|--|-----------------|---------|-----------------------|---------|
| Programa de Pós-Graduação em: | Química | | | | |
| Defesa de: | Tese de Doutorado Acadêmico, 112, PPGQUI | | | | |
| Data: | vinte e quatro de fevereiro de dois mil e vinte e um | Hora de início: | [08:05] | Hora de encerramento: | [12:45] |
| Matrícula do Discente: | 11713QMI009 | | | | |
| Nome do Discente: | Renan Faria Guerra | | | | |
| Título do Trabalho: | Modelagem molecular das interações do complexo NS1/anti-NS1 para aplicação no diagnóstico diferencial de Flavivírus | | | | |
| Área de concentração: | Química | | | | |
| Linha de pesquisa: | Físico-Química de Macromoléculas e Coloides | | | | |
| Projeto de Pesquisa de vinculação: | Uso de modelagem molecular no desenvolvimento de biossensores para detecção de herbicidas e doenças causadas pelo mosquito Aedes Aegypti | | | | |

Reuniu-se por meio de webconferência, Plataforma zoom, link: <https://us04web.zoom.us/j/76711581263?pwd=SjR3OWVJaytyRWZyWHhlcckrazhOUT09>,

a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Química, assim composta: Professores Doutores: Odonório Abraão Júnior, da Universidade Federal do Triângulo Mineiro; Osmair Vital de Oliveira, do Instituto Federal de Educação, Ciência e Tecnologia de São Paulo; Tayana Mazin Tsubone e Moacyr Comar Júnior, da Universidade Federal de Uberlândia e Eduardo de Faria Franca, orientador(a) do(a) candidato(a).

Iniciando os trabalhos o(a) presidente da mesa, Dr(a). Eduardo de Faria Franca, apresentou a Comissão Examinadora e o candidato(a), agradeceu a presença do público, e concedeu ao Discente a palavra para a exposição do seu trabalho. A duração da apresentação do Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir o senhor(a) presidente concedeu a palavra, pela ordem sucessivamente, aos(às) examinadores(as), que passaram a arguir o(a) candidato(a). Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o(a) candidato(a):

Aprovado(a).

Esta defesa faz parte dos requisitos necessários à obtenção do título de Doutor.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.



Documento assinado eletronicamente por **Eduardo de Faria Franca, Professor(a) do Magistério Superior**, em 24/02/2021, às 12:54, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Moacyr Comar Junior, Professor(a) do Magistério Superior**, em 24/02/2021, às 12:56, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Tayana Mazin Tsubone, Professor(a) do Magistério Superior**, em 24/02/2021, às 12:56, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Odonório Abrahão Júnior, Usuário Externo**, em 24/02/2021, às 12:56, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Osmair Vital de Oliveira, Usuário Externo**, em 24/02/2021, às 12:57, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **2565581** e o código CRC **0C807167**.

Ao jovem e sonhador Renan de outrora.
Conseguimos.

“Herdamos de nossos antepassados um profundo desejo por um conhecimento unificado e abrangente. [...] Mas o alargamento nos singulares últimos cem anos das múltiplas ramificações do conhecimento, tanto em extensão como em profundidade, confrontou-nos com um difícil dilema. Sentimos claramente que só agora começamos a adquirir material confiável para reunir tudo o que se sabe em uma só totalidade. Mas, por outro lado, tornou-se quase impossível para uma só mente dominar por completo mais que uma pequena porção especializada desse conhecimento. Não vejo outra saída para esse dilema [...] que nos aventurarmos a embarcar, alguns de nós, numa síntese de fatos e de teorias, ainda que munidos de conhecimento incompleto e de segunda mão sobre alguns deles, e sob o risco de parecermos tolos”

Erwin Schrodinger

AGRADECIMENTOS

À Deus por me manter firme em seus desígnios e pela fé renovada a cada manhã.

À minha Mãe pelo amor incondicional, pelo exemplo de dedicação, por ser a responsável pela minha formação de vida e por acreditar em mim mais que eu mesmo, sempre me apoiando em minhas escolhas acreditando na minha felicidade e realização pessoal. Muito Obrigado por TUDO.

À minha avó Célida pelo exemplo de vida, pelo amor e ajuda a mim concedidos, mas, sobretudo por me apoiar em seguir o caminho escolhido mesmo nele estando o ônus da sua ausência diária.

Aos amigos do coração, Rafa e Luiz André, que tenho como irmãos e que levarei para o resto da vida.

Aos meus amigos do trabalho da UFU-Patos de Minas pelo companheirismo e pela compreensão com meus afastamentos para a realização de mais essa etapa: Vanessa, Lucas e Luciana.

À minha amiga Letícia por tornar possível a realização de parte desse trabalho e pelo companheirismo e trabalho nessa árdua jornada.

Ao Roger pela lealdade e pela compreensão que ajudaram a tornar menos árdua essa etapa.

Aos colegas do Laboratório de Cristalografia e Química Computacional (LCQC).

Aos demais amigos, familiares e profissionais que de alguma forma e em algum momento colaboraram para a realização desse trabalho.

À minha sobrinha Rebeca por me apresentar uma forma diferente de amar e por me estimular a buscar sempre a evolução para quem sabe deixar um pedaço desse mundo melhor pra ela.

Ao meu orientador Prof. Dr. Eduardo Franca pelos ensinamentos, paciência e constante disponibilidade nessa etapa da minha formação acadêmica, mas, em especial, pela

confiança em me permitir trabalhar em um projeto tão desafiador e novo dentro de suas linhas de pesquisa. Exemplo a ser seguido como profissional, ao exercer com mestria o papel de orientador, e ser humano, ao saber como lidar com o próximo.

Ao Prof^o Dr. Fábio Amaral pelo exemplo de excelência na prática docente que tanto admiro, pela confiança e pelo enorme aprendizado a mim oferecido nos fundamentos da pesquisa científica.

À professora Dra. Sarah Arvelos pela participação e contribuições na banca de qualificação desse trabalho.

À professora Dra. Tayane Mazin Tsubone e aos professores Dr. Moacyr Comar Júnior, Dr. Odonório Abrahão Júnior e Dr. Osmair Vital de Oliveira pela disponibilidade de participação e pelas contribuições enquanto comissão examinadora desse trabalho.

Ao Prof^o Dr. Matheus de Souza, coordenador do Laboratório de Bioinformática e Análises Moleculares (LBAM), por disponibilizar um computador e o espaço do laboratório para meu uso pessoal na realização desse trabalho.

Aos responsáveis pelo Laboratório de Computação Científica Aplicada e Tecnologia de Informação (LCCA) pela disponibilização do cluster Orion para realização de parte dos cálculos.

À Universidade Federal de Uberlândia por ter sido uma extensão da minha casa nos últimos anos. Uma escolha acertada que me ofereceu formação acadêmica de qualidade e, ainda, me recebeu para exercer os ofícios da carreira de químico.

Aos órgãos de fomento CAPES/CNPq/FAPEMIG/RQ-MG pelos recursos financeiros disponibilizados para a realização desse trabalho.

RESUMO

O desenvolvimento de métodos de diagnóstico diferencial das flavirose Dengue e Zika, que superem as limitações de reatividade cruzada, é um desafio que encontra nos imunossensores uma alternativa promissora pela sua maior seletividade e especificidade. Neste trabalho, métodos de modelagem molecular foram utilizados para compreender e elucidar, *in silico*, as interações moleculares específicas de complexos formados pela proteína não-estrutural 1 (NS1) e um anticorpo anti-NS1 requeridas por esses dispositivos. As simulações de dinâmica molecular sugeriram a estabilidade energética e estrutural do anticorpo anti-NS1 e das proteínas NS1, mesmo na ausência de glicosilações, além de possibilitarem a amostragem estatística de suas estruturas representativas em meio aquoso com boa qualidade estereoquímica. As estruturas das proteínas NS1 e do anticorpo foram submetidas ao protocolo de Docking Molecular para a predição da estrutura 3D de complexos com caráter nativo. O protocolo proposto combina Docking Molecular de corpo rígido com diferentes métodos de busca, critérios de avaliação CAPRI, busca por resíduos de aminoácidos pertencentes a epítomos descritos na literatura e cálculos quânticos semi-empíricos, e apresentou melhor desempenho para sistemas estudados com arquivos input cristalográficos do receptor e ligante. A partir da sua aplicação, os modelos mais representativos de cada um dos sistemas estudados foram preditos pelo *webservice* ClusPro e selecionados considerando o maior número de resíduos pertencentes a epítomos conservados nas PPIs caracterizadas por um maior número de ligações de hidrogênio, hidrofóbicas e interações específicas π - π intermoleculares. Portanto, esses foram descritos como mais representativos da interação específica entre antígeno e anticorpo. Os domínios distintos wing e β -ladder foram, respectivamente, as regiões de interação preferencial do anticorpo com as proteínas NS1-DENV2 e NS1-ZIKV. Ademais, valores de energia livre de ligação na faixa de -6,5 a -13,0 kcal mol⁻¹, característica de complexos dessa natureza, e constantes de dissociação (K_d) em ordens de grandeza diferentes (10⁻⁸ a 10⁻¹⁰ M) sugerem seletividade e especificidade para os sistemas estudados, com potencial aplicabilidade em dispositivos de sensoriamento, em particular os sensores do tipo AFM, no diagnóstico diferencial e precoce de Dengue e Zika.

Palavras-chave: Dengue, Zika, NS1, Docking Molecular Proteína-Proteína, Modelagem Molecular, Diagnóstico Diferencial.

ABSTRACT

The development of differential diagnosis methods between Dengue and Zika Flaviviruses, which overcome cross-reactivity, is a challenge that finds in immunosensors a promising alternative due to its greater selectivity and specificity. In this work, molecular modelling methods were used to understand and elucidate, *in silico*, specific molecular interactions of nonstructural 1 protein (NS1) and antibody anti-NS1 complexes, required by these devices. The MD simulations suggested energetic and structural stability of the anti-NS1 antibody and the NS1 protein, even in the absence of glycosylation, and allowed the statistical sampling of their representative structures in aqueous environment with good stereochemical quality. The NS1 proteins and antibody structures were submitted to the Molecular Docking protocol to predict 3D structure of near-native complexes. The proposed protocol combines rigid body Molecular Docking with different pose prediction methods, CAPRI evaluation criteria, search for epitope amino acids residues and semiempirical approach, and presented better performance for systems studied with crystallographic ligand and receptor input files. From its application, the most representative models of each studied system were predicted by ClusPro webserver and selected considering the higher number of conserved epitopes residues in the PPIs characterized by a higher number of hydrogen bonds, hydrophobic and π - π intermolecular specific interactions. Therefore, these models were described as the most representative models of antigen-antibody specific interaction. The wing and β -leader domains were, respectively, the preferential interaction region of the antibody with NS1-DENV2 and NS1-ZIKV proteins. Furthermore, binding free energy values ranges from -6.5 to -13.0 kcal mol⁻¹, which is usual for this kind of complexes, and dissociation constants (K_d) in different magnitude orders (10⁻⁸ to 10⁻¹⁰ M) suggested selectivity and specificity for the studied systems, with potential application in sensing devices, in particular AFM-type sensors, for differential and early diagnosis of Dengue and Zika.

Keywords: Dengue, Zika, NS1, Protein-Protein Molecular Docking, Molecular Modeling, Differential Diagnosis

LISTA DE FIGURAS

| | |
|--|----|
| Figura 1.1 – Número de casos total e confirmados em laboratório de Dengue (a) e Zika (b) no Brasil nos últimos anos. | 20 |
| Figura 1.2 – Distribuição Mundial das áreas de infecção e sobreposição para DENV e ZIKV..... | 26 |
| Figura 1.3 – Representação esquemática da organização estrutural e do capsídeo esférico da partícula viral de um flavivírus..... | 29 |
| Figura 1.4 – Representação da via de síntese e processamento da proteína NS1..... | 29 |
| Figura 1.5 – Biomarcadores da Dengue e Zika ao longo do tempo de evolução de suas fases agudas..... | 30 |
| Figura 1.6 – Representação esquemática dos domínios da proteína NS1..... | 31 |
| Figura 1.7 – Representação tridimensional das estruturas dimérica (a) e hexamérica (b) da proteína NS1..... | 31 |
| Figura 1.8 – Alinhamento das sequências FASTA das proteínas NS1 PDB ID – NWV: 4OII, DENV2: 4O6B e ZIKV: 5K6K utilizadas nesse trabalho..... | 32 |
| Figura 1.9 – Representação <i>new cartoon</i> das superfícies β -ladder (esquerda) e de loops (direita) dos dímeros da proteína NS1 da Dengue (a) e da Zika (b), com identificação <i>wireframe surface</i> dos epítomos descritos na literatura e representados nas cores: verde (E1), magenta (E2), alaranjado (E3) e azul (E4). | 35 |
| Figura 2.1 – Estrutura geral dos α -aminoácidos..... | 43 |
| Figura 2.2 – Representação esquemática, identificação e classificação das estruturas dos 20 aminoácidos mais comuns encontrados em proteínas..... | 44 |
| Figura 2.3 – Reação de condensação entre aminoácidos formando uma ligação peptídica (a) e suas rotações permitidas sobre as ligações N-C α e C α -C' definidas, respectivamente, pelos ângulos diedros ϕ e ψ (b)..... | 45 |
| Figura 2.4 – Exemplo de um diagrama de Ramachandran..... | 46 |
| Figura 2.5 – Representação esquemática dos tipos mais comuns de estrutura secundária: α -hélice (a) e folha β (b). | 47 |
| Figura 2.6 – Representação esquemática da estrutura geral das imunoglobulinas..... | 48 |
| Figura 2.7 – Sequência FASTA da região Fab do anticorpo 22-anti-NS1 (a) e sua representação <i>new cartoon</i> coloridas em branco (chain L) e bege (chain H) com identificação <i>wireframe surface</i> das CDRs preditas pelo <i>software</i> Paratome (b). | 49 |
| Figura 2.8 – Representação de uma molécula sob a ótica da Mecânica Molecular. | 54 |
| Figura 2.9 – Representação das condições periódicas de contorno em três dimensões. | 59 |
| Figura 2.10 – Representação do Raio de Corte esférico para um sistema periódico em duas dimensões..... | 60 |
| Figura 2.11 - Representação dos modelos de água do tipo: SPC e TIP3P (a), TIP4P (b) e TIP5P (c); I_1 e I_2 representam as distâncias entre o átomo hidrogênio e oxigênio; q_1 e q_2 representam as cargas atômicas do hidrogênio e oxigênio, respectivamente; θ e Φ são os ângulos de ligação e σ é a distância para o potencial zero. | 62 |
| Figura 2.12 – Energia Total dos sistemas das proteínas NS1 (a) e do anticorpo (b) simulados em meio aquoso..... | 66 |

| | |
|--|-----|
| Figura 2.13 – RMSD (a) e Raio de Giro (b) dos sistemas das proteínas NS1 e do anticorpo simulados em meio aquoso..... | 67 |
| Figura 2.14 – Representação <i>new cartoon</i> da sobreposição entre a estrutura do anticorpo no início (cinza) e ao final da simulação (preto). | 68 |
| Figura 2.15 – Representação <i>new cartoon</i> da sobreposição entre a estrutura das proteínas NS1-DENV2 (a) e NS1-ZIKV (b) no início (cinza) e ao final da simulação (azul/vermelho)..... | 69 |
| Figura 2.16 – RMSF para os resíduos de aminoácido das proteínas NS1-DENV2 e NS1-ZIKV e indicação das regiões de maior flutuação..... | 70 |
| Figura 2.17 – Área acessível ao solvente (SASA) e suas contribuições polar e apolar para as proteínas NS1 de ambas as flavirose no início a 0 ns (a) e final a 100 ns (b) de simulação. | 73 |
| Figura 2.18 – Representação <i>new cartoon</i> das estruturas representativas das proteínas a) NS1-DENV2, b) NS1-ZIKV e c) 22-anti-NS1 em solução aquosa. | 74 |
| Figura 2.19 – Diagramas de Ramachandran das estruturas representativas das proteínas a) NS1-DENV2, b) NS1-ZIKV e c) 22-anti-NS1..... | 75 |
| Figura 3.1 – Identificação dos fragmentos geométricos (azul: côncavo, vermelho: convexo e verde: plano) em uma proteína (a) conforme função de forma (b) aplicada pelo algoritmo implementado no <i>webservice</i> PatchDock..... | 87 |
| Figura 3.2 – Fluxograma da metodologia de Docking Molecular..... | 92 |
| Figura 3.3 – Representações <i>new cartoon</i> da estrutura cristalográfica do sistema WNV/22-anti-NS1 (a) e das poses melhor ranqueadas preditas pelo redocking realizado utilizando os <i>webservers</i> ClusPro (b) e PatchDock (c). | 93 |
| Figura 3.4 – Representação em malha do <i>fitting</i> entre a interface de interação da estrutura cristalográfica WNV/22-anti-NS1, em coloração amarela, com poses melhor ranqueadas preditas pelo redocking realizado utilizando os <i>webservers</i> ClusPro, em azul, e PatchDock, em vermelho..... | 94 |
| Figura 3.5 – Representação <i>new cartoon</i> das poses de menor energia selecionadas dos <i>webservers</i> PatchDock (a-c) e ClusPro (d-h), conforme critério de viabilidade biológica, ordenadas em forma crescente de energia para a NS1-DENV2 Cristalográfica (1). | 95 |
| Figura 3.6 – Representação <i>new cartoon</i> das poses de menor energia selecionadas dos <i>webservers</i> PatchDock (a-c) e ClusPro (d-j), conforme critério de viabilidade biológica, ordenadas em forma crescente de energia para a NS1-ZIKV Cristalográfica (2). | 96 |
| Figura 3.7 – Representação <i>new cartoon</i> das poses de menor energia selecionadas do <i>webservers</i> PatchDock (a-e) e ClusPro (f-k), conforme critério de viabilidade biológica, ordenadas em forma crescente de energia para a NS1-DENV2 MD (3)..... | 97 |
| Figura 3.8 – Representação <i>new cartoon</i> das poses de menor energia selecionadas do <i>webservers</i> PatchDock (a-c) e ClusPro (d-i), conforme critério de viabilidade biológica, ordenadas em forma crescente de energia para a NS1-ZIKV MD (4). | 98 |
| Figura 3.9 – Número de poses após reranking conforme critério estabelecido pela CAPRI para os sistemas: (a) NS1-DENV2 Cristalográfica, (b) NS1-ZIKV Cristalográfica, (c) NS1-DENV2 MD e (d) NS1-ZIKV MD. | 102 |

| | |
|---|-----|
| Figura 3.10 – Interfaces de interação dos modelos dos complexos NS1-DENV2/22-anti-NS1: DC09 (a) e DC07 (b), com os resíduos pertencentes aos epítomos E1, E2 e E3 destacados nas cores verde, magenta e alaranjado, respectivamente. | 104 |
| Figura 3.11 – Interfaces de interação dos modelos dos complexos NS1-ZIKV/22-anti-NS1: ZC02 (a), ZC04 (b) e ZP09 (c) referentes ao sistema 2 e ZC02 (d) e ZP02 (e) referentes ao sistema 4, com os resíduos pertencentes aos epítomos E1, E2, E3 e E4 destacados nas cores verde, magenta, alaranjado e azul, respectivamente..... | 106 |
| Figura 4.1 – Alinhamento das sequências das CDRs do anticorpo identificando os resíduos de aminoácidos pertencentes às ICs com as proteínas NS1 das Flaviviruses WNV (PDB ID 4OII), DENV2 (DC09) e ZIKV (ZC02)..... | 124 |

LISTA DE TABELAS

| | |
|---|-----|
| Tabela 1.1 – Levantamento bibliográfico de estudos teórico teóricos e/ou experimentais de predição de epítomos das proteínas NS1 considerados na análise das interfaces de interação preditas neste trabalho..... | 34 |
| Tabela 2.1 – Descrição dos termos de potenciais ligados e não-ligados do campo de força OPLS-AA utilizado nos cálculos de DM. | 56 |
| Tabela 2.2 – Comparação dos parâmetros de diferentes modelos de molécula de água.62 | |
| Tabela 2.3 – Descrição dos sistemas simulados..... | 63 |
| Tabela 2.4 – Número de ligações de hidrogênio entre moléculas de água e a proteína, e pontes salinas quantificadas durante as simulações por DM..... | 72 |
| Tabela 3.1 – Número de resíduos pertencentes aos epítomos e das interações presentes na interface dos modelos preditos. | 104 |
| Tabela 3.2 – Número de resíduos pertencentes ao epítomos e das interações presentes na interface dos modelos preditos. | 105 |
| Tabela 4.1 – Valores de ΔH_{lig} calculados pelo método semi-empírico PM7..... | 121 |
| Tabela 4.2 – Valores de ΔG_{lig} e constantes de dissociação (K_d) e associação (K_a), a 298,15 K, para os modelos preditos pelo protocolo proposto. | 123 |

LISTA DE ABREVIATURAS E SIGLAS

- a.a. – Aminoácido
- abd – anticorpo (*antibody*)
- ACC – Acetil coenzima A carboxilase
- ADARS – *assymetric Decoys as the Reference State*
- AFM – Microscopia de Força Atômica
- AHAS – Acetohidróxiácido Sintase
- AM1 – Modelo Austin 1
- CAPRI – Avaliação Crítica de Interações Preditas
- CDR – Região Determinante de Complementariedade
- CHARMM – *Chemistry at Harvard Macromolecular Mechanics*
- CHIKV – Chikungunya
- CPU – Unidade Central de Processamento
- DARS – *Decoys as the Reference State*
- DENV – Dengue Vírus
- DENV2 – Dengue Vírus Sorotipo 2
- DFT – Teoria do Funcional de Densidade
- DM – Dinâmica Molecular
- DTNs – Doenças Tropicais Negligenciadas
- Fab – Fragmento de ligação ao antígeno
- Fc – Fragmento cristalizável
- FFT – Transformada Rápida de Fourier
- fs – femtossegundos (10^{-15} s)
- $G_{predição}$ – Energia livre de predição
- GRACE – *Graphing Advanced Computation and Exploration of Data*
- HF – Hartree-Fock
- H_f – Entalpia de Formação
- H_{lig} – Entalpia de Ligação
- ICs – Contatos Interfaciais
- IgG – Imunoglobulina G
- IgM – Imunoglobulina M
- iRMSD – Raiz Quadrada do Desvio Quadrático Médio de Interface
- IUPAC – União Internacional de Química Pura e Aplicada

JEV – Febre Japonesa
kb – quilobase
mNS1 – NS1 de membrana (estrutura dimérica)
MOG – Glicoproteína Oligodentrócita de Mielina
NIS – Superfícies que não interagem
nm – nanômetro
NpT – *ensemble* isotérmico-isobárico
ns – nanossegundos (10^{-9} s)
NS1 – Proteína não-estrutural 1
OMS – Organização Mundial de Saúde
OPLS-AA – Parâmetros Otimizados para Simulação *all-atoms* de Líquidos
PCR – Reação em cadeia da polimerase
PCR-RT – Reação da transcriptase reversa seguida pela reação em cadeia da polimerase
PDB – Protein Data Bank
PM3 – Método Paramétrico 3
PM6 – Método Paramétrico 6
PM7 – Método Paramétrico 7
PME – Potencial Mesh Ewald
PPI – Interação Proteína Proteína
ps – picossegundos (10^{-12} s)
R_c – Raio de Corte
R_g – Raio de Giro
RM1 – Modelo Recife 1
RMN – Ressonância Magnética Nuclear
RMSD – Raiz Quadrada do Desvio Quadrático Médio
RMSF – Raiz Quadrada da Flutuação Quadrática Média
RNA – Ácido ribonucleico
ROSIE – *The Rosetta Online Server that Includes Everyone*
SASA – Superfície Acessível ao Solvente
sNS1 – NS1 secretada (estrutura hexamérica)
SPC – *Single Point Charge*
TIP – *Transferable Intermolecular Potentials*
WNV – Febre do Nilo Ocidental
YFV – Febre Amarela

ZDO – Zero Differential Overlap Approximation

ZIKV – Zika Virus

SUMÁRIO

| | |
|---|-----------|
| CAPÍTULO 1 - FLAVIROSES | 20 |
| 1.1 INTRODUÇÃO E JUSTIFICATIVA | 20 |
| 1.1.1 Dengue..... | 25 |
| 1.1.2 Zika..... | 27 |
| 1.1.3 A Proteína NS1..... | 28 |
| 1.2 OBJETIVOS..... | 36 |
| 1.2.1 Objetivo Geral | 36 |
| 1.2.2 Objetivos Específicos | 36 |
| REFERÊNCIAS | 37 |
| CAPÍTULO 2 - PROTEÍNAS: ANTÍGENO NS1 E ANTICORPO 22-ANTI-NS1 | 43 |
| 2.1 AMINOÁCIDOS E PROTEÍNAS | 43 |
| 2.1.1 Imunoglobulinas: anticorpos | 47 |
| 2.2 MODELAGEM MOLECULAR DE SISTEMAS BIOLÓGICOS | 50 |
| 2.2.1 Simulação por Dinâmica Molecular (DM)..... | 52 |
| 2.2.2 Algoritmo (Leap-Frog)..... | 57 |
| 2.2.3 <i>Ensemble</i> NpT | 57 |
| 2.2.4 Condições Periódicas de Contorno..... | 58 |
| 2.2.5 Raio de Corte..... | 59 |
| 2.2.6 Modelos de moléculas de água para solvatação explícita | 61 |
| 2.3 METODOLOGIA..... | 63 |
| 2.3.1 Dinâmica Molecular | 63 |
| 2.3.2 Análise da população (cluster) das simulações por Dinâmica Molecular..... | 65 |
| 2.4 RESULTADOS E DISCUSSÕES..... | 66 |
| 2.5 CONCLUSÕES PARCIAIS..... | 76 |
| REFERÊNCIAS | 76 |
| CAPÍTULO 3 - INTERAÇÃO ANTÍGENO-ANTICORPO: DOCKING MOLECULAR PROTEÍNA-PROTEÍNA | 82 |
| 3.1 DOCKING PROTEÍNA-PROTEÍNA..... | 82 |
| 3.1.1 ClusPro | 84 |
| 3.1.2 PatchDock..... | 87 |
| 3.1.3 Rosie | 89 |
| 3.2 METODOLOGIA..... | 90 |
| 3.3 RESULTADOS E DISCUSSÃO | 93 |
| 3.3.1 Redocking..... | 93 |

| | |
|---|------------|
| 3.3.2 Docking Molecular | 94 |
| 3.3.3 Reranking | 100 |
| 3.3.4 Análise das Interfaces de Interação | 103 |
| 3.4 CONCLUSÕES PARCIAIS | 108 |
| REFERÊNCIAS | 109 |
| CAPÍTULO 4 - CÁLCULO DE PROPRIEDADES TERMODINÂMICAS DOS COMPLEXOS NS1/ANTI-NS1 | 114 |
| 4.1 INTRODUÇÃO | 114 |
| 4.1.1 Mecânica Quântica – O Método Hartree-Fock | 114 |
| 4.1.2 Métodos Semi-Empíricos: PM7 | 117 |
| 4.2 METODOLOGIA | 119 |
| 4.2.1 Cálculo da Entalpia de Ligação | 119 |
| 4.2.2 Cálculo da Energia de Ligação | 120 |
| 4.3 RESULTADOS E DISCUSSÃO | 120 |
| 4.4 CONSIDERAÇÕES FINAIS | 125 |
| 4.5 TRABALHOS FUTUROS | 127 |
| REFERÊNCIAS | 128 |

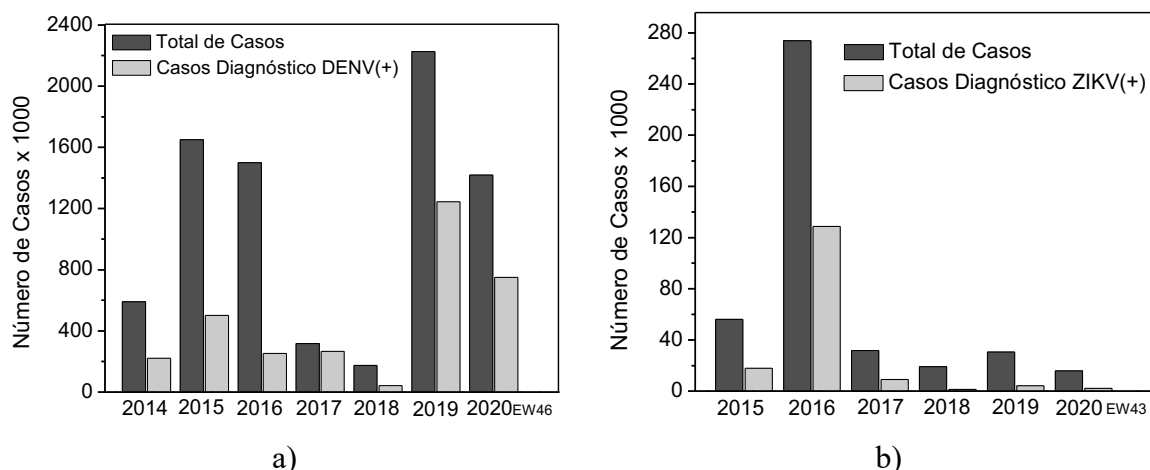
FLAVIROSES

1.1 INTRODUÇÃO E JUSTIFICATIVA

Segundo a Organização Mundial de Saúde (OMS), as Doenças Tropicais Negligenciadas (DTNs) são uma variedade de patologias causadas por agentes infecciosos e parasitários que acometem, em especial, a população dos países de clima tropical e subtropical. As DTNs, além de comprometer o estado de saúde dos habitantes, acabam impondo a eles um conjunto de ônus, pois afetam o desenvolvimento econômico e social dessas regiões em função da falta de: atenção à saúde pública, recursos para pesquisas, desenvolvimento e tratamentos específicos (HORSTICK; TOZAN; WILDER-SMITH, 2015). Dentro desse conjunto de doenças, destacam-se aquelas transmitidas por mosquitos, como a Malária, Febre Amarela, Dengue, Zika e Chikungunya, que correspondem, em nível global, a cerca de 17 % de todas as infecções (WHO, 2019).

Na última década, no Brasil, a expansão e o surgimento dos quadros epidêmicos de Dengue (DENV) e Zika (ZIKV), respectivamente, mereceram a atenção e os esforços dos órgãos governamentais de saúde e da população. Os gráficos apresentados na Figura 1.1 mostram a evolução temporal nos últimos anos do número total de casos e de diagnósticos laboratoriais positivos dessas arboviroses no Brasil.

Figura 1.1 – Número de casos total e confirmados em laboratório de Dengue (a) e Zika (b) no Brasil nos últimos anos.



Fonte: Elaborado pelo autor com base nos dados da OPS (2020)

Conforme observado nos gráficos apresentados na Figura 1.1, o ano de 2015 marcou um momento importante na saúde pública do Brasil. Isso porque, após os quadros epidêmicos do século passado e a expansão desde o início dos anos 2000, uma nova epidemia de Dengue foi declarada. Além disso, nesse mesmo ano, também foram registrados os primeiros casos de Zika e apontada sua associação com casos de microcefalia em recém-nascidos de gestantes diagnosticadas com essa patologia (BRADY et al., 2019; LOVER, 2016) e com outras desordens neurológicas, como a Síndrome de Guillain-Barré (BARBI et al., 2018).

Portanto, Dengue e Zika Vírus, além de serem as flavirozes de maior ocorrência no Brasil possuem em comum o mesmo vetor de transmissão (a fêmea do mosquito *Aedes aegypti*) e o quadro clínico característico de uma doença febril aguda autolimitada. Outros exemplos de flavirozes de menor ocorrência são: Febre Amarela (YFV), Febre do Nilo do Oeste (WNV) e Febre Japonesa (JEV) (GUARNER; HALE, 2019).

As medidas de prevenção e controle das doenças se restringem ao combate do agente transmissor, ou seja, evitar a proliferação do mosquito, eliminando possíveis criadouros. Os métodos profiláticos são limitados, porque não há ainda um método eficaz e seguro de imunização ativa. Dessa forma, tornam-se pertinentes e promissores estudos relacionados ao agente transmissor e, em especial, ao desenvolvimento de métodos de diagnóstico.

Segundo Ferreira (2015, p. 71), para flavirozes “os testes devem ser padronizados para o diagnóstico durante o período febril da doença. Nessa fase, os testes deverão ser muito sensíveis e capazes de diferenciar as infecções”. Esse diagnóstico preciso, precoce e diferencial entre essas patologias é essencial no manejo clínico dos pacientes e na orientação de ações epidemiológicas voltadas aos seus controles (MELO, 2012).

O diagnóstico dos casos clínicos de suspeita de Dengue e Zika envolve em um primeiro momento a avaliação, por um profissional médico, dos sintomas observados e relatados. No entanto, os quadros clínicos dessas patologias, em especial na fase aguda, apresentam sintomatologias caracterizadas por manifestações inespecíficas e variáveis conforme a resposta imunofisiológica do infectado. Isso pode ocasionar subnotificação ou notificação equivocada e, por consequência, um dimensionamento indevido da ocorrência dessas patologias (SILVA, 2019).

Nesse sentido, os métodos laboratoriais de diagnóstico da infecção primária dessas patologias são de quatro tipos: (1) microbiológico – isolamento do vírus, (2) biologia molecular

– reação em cadeia da polimerase (PCR) e PCR em tempo real (PCR-RT), (3) imunológicos – pesquisa de componentes antigênicos do vírus por meio de testes imunoenzimáticos como o ELISA e (4) teste de neutralização por redução em placa (PRNT) (MULLER; DEPELSENAIRE; YOUNG, 2017; PETERS; STEVENSON, 2019; RABE et al., 2016). Cada um dos testes supracitados, apesar de utilizados, apresentam, respectivamente, as seguintes limitações: (1) baixa especificidade, por conta da reatividade cruzada, (2) são dispendiosos e demandam laboratoristas altamente qualificados, (3) apresentam ocorrência de falsos resultados negativos em virtude da forma de detecção indireta inerente ao método e (4) trata-se de uma técnica complexa, de custo elevada realizada apenas em laboratórios de pesquisa e centros de referência por utilizar os vírus vivos.

É importante salientar ainda que em áreas onde há cocirculação dos vírus da Dengue e Zika, como no Brasil, a probabilidade de ocorrência de reação cruzada¹ em testes diagnósticos é elevada sendo sua principal limitação. Desse modo, a circulação simultânea desses vírus, que possuem semelhanças funcionais e estruturais significativas, acaba por dificultar a distinção laboratorial entre a viroses (ABRASCO, 2016). Isso porque a reatividade cruzada entre a flaviroses acaba por impossibilitar a identificação específica do vírus infectante ou de seus outros fatores de virulência, em especial, nos casos de reinfecção ou de vacina contra outros flavivírus.

Em resposta a esses desafios, os estudos mais recentes da literatura apontam como potencial a identificação da proteína não-estrutural 1 (NS1) como um marcador antigênico para Dengue e Zika. Isso porque trata-se de um fator de virulência encontrado em níveis elevados no soro de pacientes infectados, da ordem de $50 \mu\text{g mL}^{-1}$, desde a manifestação dos primeiros sintomas até cerca de nove dias do quadro febril. Portanto, essa proteína pode ser identificada por dispositivos analíticos ao mesmo tempo que o RNA viral, mas antes que seja necessária uma resposta imune com formação de anticorpos, o que ocorre apenas na fase secundária das doenças (FREIRE et al., 2017; MULLER; DEPELSENAIRE; YOUNG, 2017; XU et al., 2016).

Na busca por métodos mais específicos e sensíveis, diversos grupos de pesquisa têm dispendido esforços na proposição de biossensores para quantificação da proteína NS1 baseados no caráter imunogênico do complexo antígeno-anticorpo. No entanto, a exigência de diversas etapas laboratoriais e a ausência de escalonamento para detecção simultânea de diversos biomarcadores acabam por restringir a aplicabilidade desses métodos no diagnóstico

¹ Reatividade Cruzada refere-se à competição direta entre a molécula de interesse e outras moléculas por sítios de ligação do anticorpo devido às suas semelhanças estruturais. Desse modo, ocorre quando um anticorpo dirigido contra um antígeno específico é capaz de se ligar eficientemente com outro antígeno diferente.

Assim sendo, é desafiadora a integração das características analíticas requeridas por esses métodos imunoenzimáticos e o escalonamento desses dispositivos para identificar e discriminar biomarcadores sendo, desse modo, capazes de realizar um diagnóstico diferencial entre as flaviviruses.

Nesse contexto, o presente trabalho parte da premissa de que esse desafio tem uma alternativa promissora no uso da Microscopia de Força Atômica (AFM) como ferramenta de sensoriamento, altamente sensível e seletiva, das proteínas NS1 das flaviviruses. Isso porque o uso desse tipo de biossensor possibilita o mapeamento de grupamentos funcionais no substrato em escala micro e nanométrica. Desse modo, permite detectar interações intermoleculares específicas ou não-específicas em eventos de reconhecimento molecular como os estabelecidos entre antígeno e anticorpo (BUTT; CAPPELLA; KAPPL, 2005; DEDA et al., 2013; ITO; IBRAHIM; GRABOWSKA, 2010). Sendo assim, a precisão desse mapeamento, princípio dessa técnica, associada às características específicas das proteínas NS1 da Dengue e Zika, conferem *seletividade* ao biossensor. Contornando assim, as limitações de reatividade cruzada associadas aos métodos atuais de diagnóstico das flaviviruses.

Além disso, pelo seu uso é possível quantificar forças entre a molécula de reconhecimento biológico, imobilizada na ponteira do AFM, e a amostra. Logo, é capaz de mensurar forças de interação na faixa das encontradas em condições fisiológicas garantido assim a *sensibilidade* requerida para o dispositivo (JALILI; LAXMINARAYANA, 2004; LI et al., 2002).

Para tanto, o imunossensor em questão utilizaria um anticorpo anti-NS1 como elemento de reconhecimento biológico da proteína NS1 das flaviviruses Dengue e Zika. Isso pode ser realizado pela imobilização do anticorpo na ponteira (*tip*) do AFM, pelas técnicas de funcionalização bem descritas na literatura (CAMARA et al., 2013; KATZ, 1990), a fim de que esse estabeleça uma interação com a amostra contendo o analito de interesse. Como resultado da ligação biomolecular específica entre antígeno e anticorpo, uma mudança na energia livre é induzida (FRITZ, 2000) e, por sua vez, o cantilever, que atua como um transdutor de força, sofre uma deflexão. Essa deflexão, deve ser diretamente proporcional à concentração da proteína NS1 na amostra analisada (BOWEN; HILAL, 2009; FERREIRA; YAMANAKA, 2006).

Logo, para análise da viabilidade de aplicação desse sistema biológico em dispositivos analíticos dessa natureza torna-se fundamental a compreensão e distinção das interações moleculares entre a proteína NS1 de ambas flavirose e um anticorpo anti-NS1. Isso porque reconhecer e mensurar essas forças de interação, estabelecidas por esses complexos, é essencial na elucidação das bases moleculares dos processos de reconhecimento das proteínas NS1.

Portanto, para os objetivos do presente trabalho a Química Teórica e sua aplicação na área biológica/biomédica, que também pode ser conhecida como Biologia Computacional, se apresenta, como uma ferramenta potencial. Isso porque, de acordo com Kennely e Rodwell (2014, p. 99), essa área de estudo busca

desenvolver modelos de computador que apliquem os princípios físicos, químicos e biológicos para reproduzir o comportamento das moléculas e processos biológicos. (...) Ao realizar experiências e análises virtuais “*in silico*”, o que significa ser realizada em computador ou por simulação em computador, a biologia computacional oferece a promessa de acelerar a velocidade e a eficiência da descoberta científica. (...) A criação de algoritmos computadorizados que imitam com exatidão o comportamento das proteínas, enzimas, células, etc. promete aumentar a velocidade, eficiência e segurança da pesquisa biomédica.

Nesse escopo, as técnicas computacionais de modelagem molecular, como Dinâmica e Docking Molecular, são utilizadas para avaliar o efeito do solvente nas dinâmicas estruturais, os modos de ligação com caráter nativo e a energia de interação envolvida no processo de interação entre a proteína NS1 e o anticorpo. Portanto, fornecem uma chave para compreensão das conformações moleculares que estão relacionadas com os processos de reconhecimento da proteína em estudo e, por sua vez, podem ser identificadas com boa correlação teórica-experimental tal como obtido por Franca et al. (2011). Dessa forma, ao utilizar as ferramentas computacionais, os modelos desenvolvidos neste trabalho poderão subsidiar e direcionar de forma direta futuros experimentos de AFM para diagnóstico diferencial e precoce de Dengue e Zika reduzindo seus custos laboratoriais e tempo de bancada.

Estudos similares ao deste trabalho já foram proposto para doenças autoimunes e neurodegenerativas (IERICH et al., 2019a; OLIVEIRA et al., 2019) e analitos de interesse na área ambiental (AMARANTE et al., 2014; BUENO et al., 2014; DEDA et al., 2013; FERREIRA; FRANCA; LEITE, 2017; FRANCA et al., 2011; GUERRA et al., 2019; IERICH et al., 2015a; OLIVEIRA et al., 2013). Essa abordagem e aplicações têm sido objeto de estudo do presente grupo de pesquisa (Laboratório de Cristalografia e Química Computacional) em parceria com o Laboratório Multidisciplinar de Nanoneurobiofísica da Universidade Federal de

São Carlos, pioneiros na descrição do subsídio das técnicas computacionais para direcionamento do uso de biossensores de AFM.

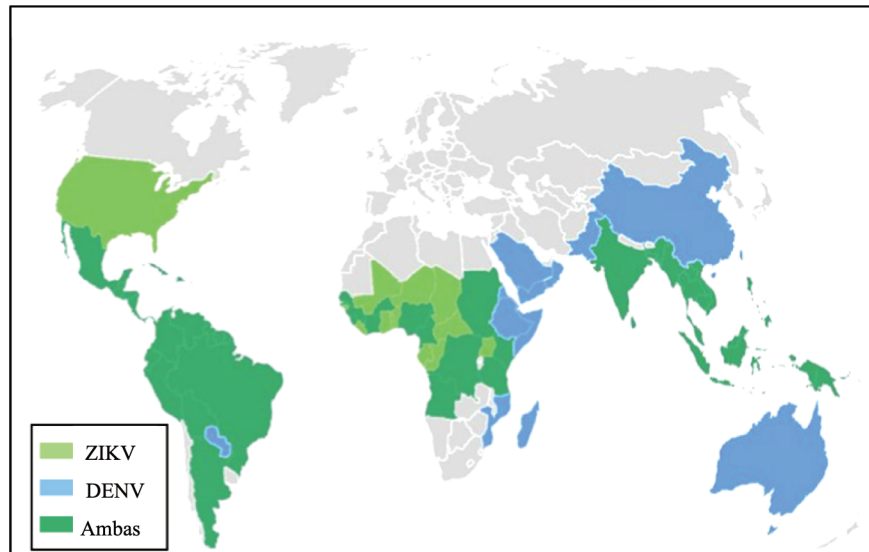
1.1.1 Dengue

A Dengue é uma doença viral não contagiosa causada por quatro sorotipos do dengue vírus (DENV 1-4) pertencente ao gênero *Flavivirus* e família Flaviviridae, e transmitido principalmente pela fêmea do mosquito *Aedes aegypti* e, de modo secundário, pelo *Aedes albopictus*. Trata-se da arbovirose de maior ocorrência em nível mundial, em especial em regiões de clima tropical e subtropical, ocasionando impactos consideráveis nas áreas social e econômica em função da extensão geográfica atingida, do número de casos e da gravidade da doença (BHATT et al., 2013; GUZMAN; HARRIS, 2015).

As primeiras referências a uma doença com manifestações clínicas similares às da Dengue datam do século III na China, mas somente após o século XVII as taxas de incidência em outras regiões cresceram em virtude do comércio marítimo (WEAVER; VASILAKIS, 2009). Em meados do século XVIII na Indonésia, Egito e Estados Unidos da América (EUA), e, ao longo do século XIX, no Caribe e na região sul dos EUA, foram registrados os primeiros quadros epidêmicos de Dengue (GUBLER, 1998). Ao longo do século XX, eventos importantes relacionados à doença ocorreram: (1) epidemias no sudoeste asiático nas décadas de 40 e 50 que, em função da Segunda Guerra Mundial acabaram contribuindo para a difusão do vírus, (2) na década de 40, pesquisadores japoneses e norte-americanos isolaram o vírus da Dengue em laboratório e, após testes de inoculação em animais, foram capazes de associar o vírus aos sintomas humanos, (3) na década de 50, na região da Tailândia e das Filipinas, foi reconhecido que o DENV era capaz de provocar uma forma mais grave da doença, a Dengue Hemorrágica, e (4) mosquitos *Aedes* foram vinculados ao ciclo evolutivo enquanto vetor de transmissão (HOTTA, 1952; PINHEIRO; CORBER, 1997; SNOW et al., 2014).

Até os anos 70, apenas nove países haviam registrado quadros epidêmicos de Dengue; no entanto, o vírus se propagou e a doença se difundiu de modo que, segundo dados da Organização Mundial de Saúde, o número de casos da doença cresceu 15 vezes nas últimas duas décadas. Desse modo, atualmente 129 países possuem áreas de risco de infecção, atingindo cerca 3,9 bilhões de pessoas, especialmente no sudeste da Ásia, Pacífico Ocidental, Américas, África e Mediterrâneo, conforme apresentado na Figura 1.2 (WHO, 2020).

Figura 1.2 – Distribuição Mundial das áreas de infecção e sobreposição para DENV e ZIKV.



Fonte: Adaptado de Collins e Metz (2017)

Segundo Guzman e Harris (2015), essa rápida disseminação da Dengue ao longo dos últimos 50-60 anos está intrinsicamente relacionada aos seguintes fatores: crescimento populacional sem precedentes, aumento nos movimentos populacionais, urbanização descontrolada, mudanças climáticas e falhas nos sistemas de saúde em relação à infraestrutura e aos programas de controle do vetor (GUZMAN; HARRIS, 2015).

De acordo com os sintomas, a Dengue pode ser dividida em três categorias segundo a OMS: dengue sem ou com sintomas de alerta e a dengue severa. Essa classificação, reformulada em 2017, está associada às manifestações clínicas da doença, que são inespecíficas e podem variar desde casos assintomáticos até quadros graves com choque hemorrágico ou não, ou ainda evoluir para óbito (GUARNER; HALE, 2019; MARTINEZ; GARZA; CUELLAR-BARBOZA, 2019).

Para cerca de 80 % dos casos, as infecções são assintomáticas ou caracterizadas por quadros febris leves. No entanto, quando a doença se instala, os sintomas começam a aparecer já nas primeiras 48 horas após a infecção (fase aguda), sendo os mais característicos: febre alta, dor de cabeça intensa, dores musculares e nas articulações e rash cutâneo (FERREIRA, 2015; MARTINEZ; GARZA; CUELLAR-BARBOZA, 2019). O aparecimento de petéquias (pequenas manchas avermelhadas) na epiderme, em número maior que 10 por cm², é uma manifestação específica da dengue (GRANDE et al., 2016). Na pequena parcela de pacientes em que a doença evolui para manifestações mais severas, são observados sintomas como

hepatomegalia, dor abdominal, vômitos persistentes e hemorragias, desencadeando no comprometimento do sistema hematológico com falha na circulação sanguínea. A perda de um volume crítico de plasma leva ao choque, que é caracterizado por alternância na frequência do pulso, diminuição da pressão de pulso, extremidades frias, demora no enchimento capilar, pele pegajosa e agitação (GUARNER; HALE, 2019; MINISTÉRIO DA SAÚDE, 2016).

Não há um tratamento específico para a Dengue, sendo recomendado o uso de paracetamol para controle da dor e da febre, além de evitar a administração de aspirina e outros anti-inflamatórios não-esteroides para evitar complicações hemorrágicas. Na observação de sinais de alerta, o paciente deve ser mantido em unidade hospitalar apropriada para avaliação clínica contínua (MARTINEZ; GARZA; CUELLAR-BARBOZA, 2019).

1.1.2 Zika

Similarmente à Dengue, a Zika pertence à categoria das flaviviruses que compartilham do mesmo vetor de transmissão: os mosquitos do gênero *Aedes*. No entanto, há evidências de que o Zika vírus (ZKV) também seja transmitido sexual e verticalmente (WANG et al., 2016). É uma patologia sobre a qual há poucos estudos, visto que surgiu como preocupação de saúde pública em nível mundial ao longo da última década. Isso porque foi relatada a sua associação com um aumento significativo de casos de defeitos congênitos e síndromes neurológicas como microcefalia (BRADY et al., 2019; LOVER, 2016) e síndrome de Guillain-Barré (BARBI, et al. 2018), respectivamente.

Historicamente, o ZIKV foi isolado pela primeira vez do sangue de um macaco em Uganda, no ano de 1947, durante um estudo sobre Febre Amarela. Nove meses depois, o vírus foi isolado de mosquitos *Aedes africanus* coletados na mesma região. Em 1952, os primeiros casos de Zika em humanos foram registrados em pacientes da Uganda e Tanzânia (MLACKER et al., 2016). Durante mais de cinco décadas, a doença permaneceu silenciosa, com registro de poucos casos leves e restritos a determinadas regiões dos continentes asiático e africano. No ano de 2007, um surto de infecção por ZIKV foi notificado na Micronésia, região do Pacífico (GOURINAT et al., 2015). Em seguida, durante os anos de 2013 e 2014, a Polinésia Francesa registrou uma das maiores epidemias de Zika, com a notificação de aproximadamente 270.000 casos. Esse quadro epidêmico foi superado em 2015, quando o ZIKV chegou ao continente

americano, com destaque para o grande número de casos registrados no Brasil, conforme dados apresentados na Figura 1.1 (MLACKER et al., 2016).

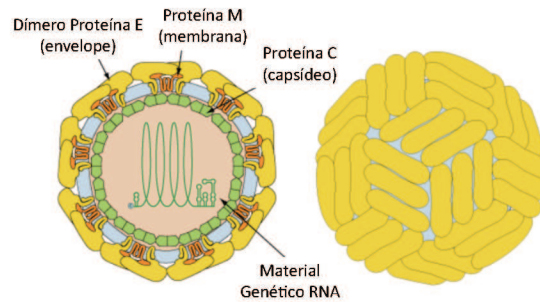
Nesse contexto, em 1º de Fevereiro de 2016, a OMS declarou a Zika como emergência de saúde pública em nível mundial (ROOS, 2016). Desde então, em função da circulação do vírus, a sua transmissão é relatada em 55 países e territórios, conforme panorama mostrado na Figura 1.2. Desde 2018, o número de infecções por ZIKV decresceu consideravelmente, fenômeno que, segundo especialistas, ocorreu em função de infecções primárias assintomáticas que conferiram imunização à parte da população. No entanto, as medidas preventivas devem ser preconizadas até que haja uma vacina para imunização eficaz e segura (OSUNA et al., 2016).

O período de incubação do ZIKV é de três a 12 dias e, em mais de 80 % dos casos, as manifestações da doença são assintomáticas ou subclínicas e provocam sintomas moderados (MARTINEZ; GARZA; CUELLAR-BARBOZA, 2019). Os sintomas mais característicos são: vermelhidão em todo o corpo, febre baixa (37,8 – 38,5 °C), conjuntivite, mialgia e dores de cabeça e nas articulações (YADAV; RAWAL; BAXI, 2016).

1.1.3 A Proteína NS1

O Flavivírus causador da Dengue e da Zika possui formato esférico com dimensão de 50-55 nm, e como material genético uma fita simples de RNA de cerca 11 kb com sentido positivo, responsável por codificar uma grande poliproteína (Figura 1.3) (LINDENBACH; RICE, 2003). Essa proteína, por sua vez, é clivada em 3 proteínas estruturais (Capsídeo – C, Membrana – M e Envelope – E) e 7 proteínas não-estruturais (NS1, NS2A, NS2B, NS3, NS4A, NS4B e NS5) (RICE et al., 1985). Na Figura 1.3 é representada a fita de RNA dentro do capsídeo, constituído por proteínas C e envolvido pelo envelope, constituído por proteínas M e E.

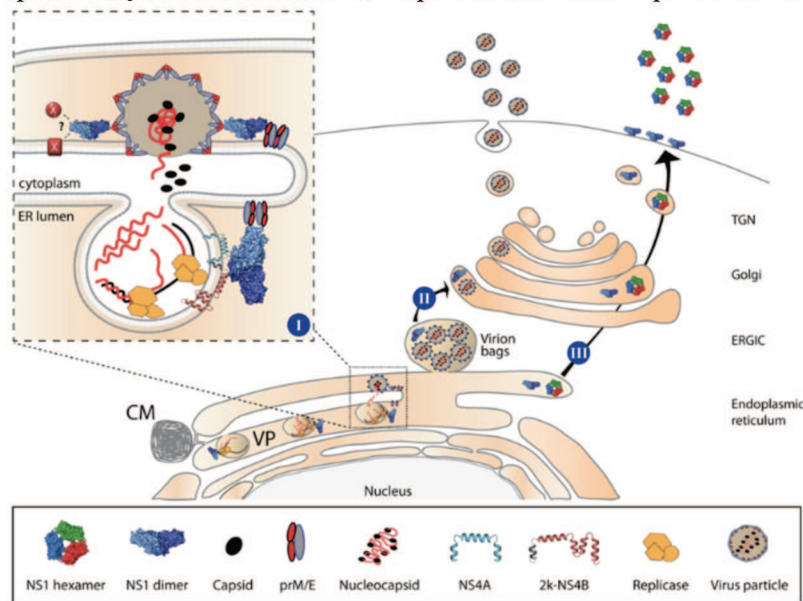
Figura 1.3 – Representação esquemática da organização estrutural e do capsídeo esférico da partícula viral de um flavivírus.



Fonte: Adaptado de Hulo et al. (2018)

Dentre as proteínas não-estruturais, a NS1 tem ganhado atenção no estudo dessas patologias por se tratar de um fator de virulência multifuncional, ou seja, por estar envolvida em diversas etapas do ciclo de vida do vírus como: replicação viral, escape do sistema imune e processos de modulação da célula do hospedeiro (AMORIM et al., 2014). A NS1 é uma glicoproteína que se apresenta em diferentes formas oligoméricas sendo elas: (1) o monômero, que é encontrado apenas no interior do compartimento das membranas do retículo endoplasmático; (2) o dímero (mNS1), que uma fração participa do processo de replicação viral dentro da membrana do retículo endoplasmático e outra encontra-se associado à membrana celular e (3) o hexâmero (sNS1), lipoproteico que trata-se da forma secretada para o meio extracelular (FLAMAND et al., 1999), conforme esquematizado na Figura 1.4.

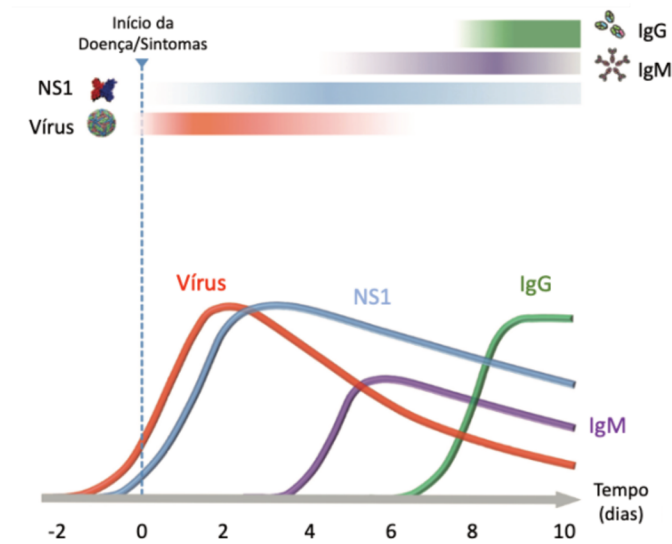
Figura 1.4 – Representação da via de síntese e processamento da proteína NS1.



Fonte: (SCATURRO et al., 2015)

A forma sNS1 é secretada e pode ser encontrada na circulação sanguínea de infectados, ainda na fase aguda das doenças, em concentrações da ordem de até $50 \mu\text{g mL}^{-1}$. Portanto, pode ser detectada em amostras de soro desde o início até o nono dia após o início do período febril (YOUNG et al., 2000). Os demais marcadores imunológicos IgM e IgG apenas aparecem, respectivamente, a partir do dia 3 ou ao final da fase aguda (VAUGHN et al., 1997), conforme apresentado na Figura 1.5. Portanto, a proteína sNS1 é considerada o maior marcador antigênico de flavirose.

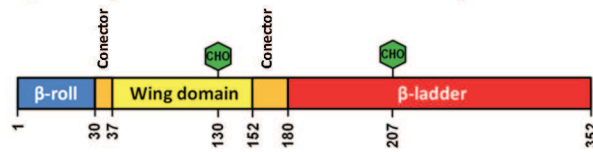
Figura 1.5 – Biomarcadores da Dengue e Zika ao longo do tempo de evolução de suas fases agudas.



Fonte: Adaptado de Muller; Depelsenaire e Young (2017)

A forma monomérica da proteína NS1, com peso molecular aproximado de 48 kDa, é constituída por 352 resíduos de aminoácidos, seis ligações dissulfeto intramoleculares e sítios de glicosilação nos aminoácidos Asn130 e Asn207 (BEESETTI; KHANNA; SWAMINATHAN, 2014; BLITVICH et al., 2001). Além disso, a proteína enovela-se em 3 domínios diferentes: β -roll (azul), wing (amarelo) e β -ladder (vermelho) (Figura 1.6) (AKEY et al., 2014a).

Figura 1.6 – Representação esquemática dos domínios da proteína NS1.

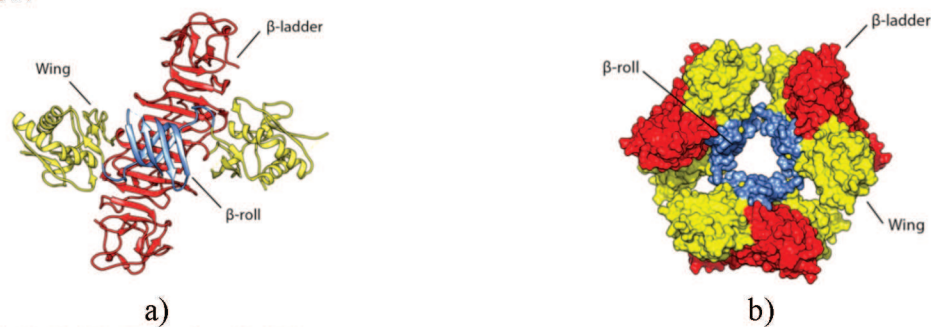


Fonte: Adaptado de Scaturro et al. (2015)

O domínio β -roll do monômero é formado pelos aminoácidos a.a. 1-29 com enovelamento de folhas β na porção N-terminal cujo entrelace entre dois monômeros formam o domínio β -roll do dímero. O domínio wing é formado pela sequência de aminoácidos a.a. 38-151, caracterizado pelas regiões de α -hélice e de loops, essas de elevada flexibilidade e flutuação estrutural compreendida entre os aminoácidos 110-130. Já o domínio β -ladder é o de maior extensão compreendendo a região formada pela sequência de aminoácidos a.a. 181-352, em suas porções C-terminal, caracterizada por 9 folhas β antiparalelas (AKEY et al., 2014a; BROWN et al., 2016).

A forma dimérica mNS1 (Figura 1.7a) apresenta caráter hidrofóbico, em virtude das propriedades do domínio β -roll, que confere a ela considerável importância biológica, pois é essa propriedade que (1) permite a associação da proteína com a membrana e (2) promove a estabilidade da conformação *open-barrel shell* da forma hexamérica (Figura 1.7b). Essa forma secretada sNS1 é constituída pela associação entre três dímeros, unidos por interações hidrofóbicas do domínio β -roll, que a conferem um formato de “barril” cujo centro é uma região inacessível ao solvente por onde são carregados *packs* de lipídeo da membrana celular (MULLER; YOUNG, 2013; SCATURRO et al., 2015). Durante o processo de polimerização da NS1, as enzimas glicosidases e glicosiltransferases removem os açúcares, tornando a forma secretada sNS1 isenta de glicosilações (RASTOGI; SHARMA; SINGH, 2016).

Figura 1.7 – Representação tridimensional das estruturas dimérica (a) e hexamérica (b) da proteína NS1.



Fonte: (SCATURRO et al., 2015)

Os cálculos de Docking Molecular desse trabalho, para os sistemas antígeno-anticorpo das flavirozes Dengue e Zika, são continuidade do estudo prévio apresentado na dissertação de mestrado: “Estudo da interação do complexo antiNS1/NS1-DENV2 usando ancoramento molecular e métodos quânticos semi-empíricos” (SILVA, 2020). Neste trabalho foi proposto um protocolo para seleção do modelo de complexo mais representativo para o sistema NS1-DENV2/22-anti-NS1 utilizando diferentes *softwares* de Docking Molecular, incluindo flexibilidade conformacional, suportados pela determinação das entalpias de ligação obtidas de cálculos quânticos semi-empíricos em convergência com regiões de epítomos determinados por estudos prévios da literatura. As considerações finais desse trabalho apontaram que diante da funcionalidade do protocolo proposto para o complexo da Dengue o mesmo poderia ser “utilizado para avaliar as interações da proteína NS1 de outras flavirozes, como a Zika” por possuírem características estruturais e funcionais semelhantes (SILVA, 2020, p. 46). Nesse tocante, o presente trabalho aplicou o protocolo proposto para a estrutura cristalográfica da NS1-ZIKV, bem como realizado por Silva (2020) para a NS1-DENV2 (Sistema 1). Além disso, foram realizadas simulações de dinâmica molecular para avaliar o efeito do solvente no comportamento das proteínas envolvidas na formação do complexo. Desse modo, estruturas representativas da NS1 de ambas flavirozes e do anticorpo também foram utilizadas como *input* dos cálculos de Docking Molecular.

O protocolo testado baseia a seleção da conformação mais representativa do complexo NS1- anti-NS1 (interação específica), predita por *websserver* de Docking Molecular proteína-proteína, na concordância entre os valores de entalpia de formação e resíduos localizados na interface de interação pertencentes a epítomos determinados por estudos prévios da literatura. Os estudos teóricos e/ou experimentais de predição de epítomos das proteínas NS1 considerados na análise das interfaces de interação preditas estão sumarizados na Tabela 1.1.

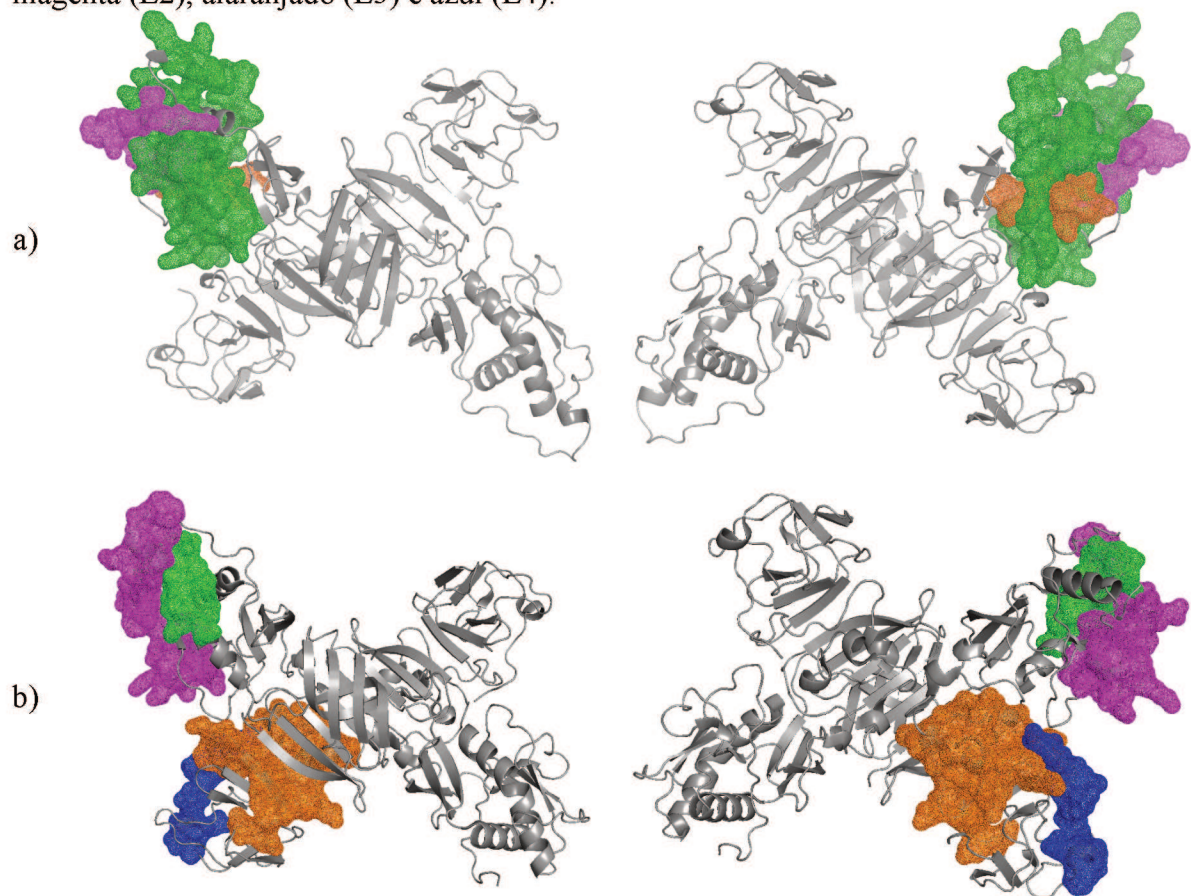
Tabela 1.1 – Levantamento bibliográfico de estudos teórico teóricos e/ou experimentais de predição de epítomos das proteínas NS1 considerados na análise das interfaces de interação preditas neste trabalho.

| Referência | Descrição |
|--------------------------------|---|
| (FALCONAR; YOUNG; MILES, 1994) | A partir de imunoenaios com 9 anticorpos monoclonais, os autores definiram a presença de epítomos lineares da NS1 do vírus da Dengue, identificados por a.a. 25-33; a.a. 61-69; a.a. 111-121; a.a. 299-309. |
| (JIANG et al., 2010) | Estudo teórico-experimental de seleção e identificação de epítomos na proteína NS1 da Dengue sorotipo 2 pela técnica de <i>Phage Display</i> utilizando anticorpos policlonais. Os epítomos confirmados pelo uso de <i>softwares online</i> de bioinformática compreendem as seguintes sequências de resíduos a.a. 32-40; a.a. 80-89; a.a. 103-112; a.a. 121-130; a.a. 187-196; a.a. 295-304 e a.a. 315-324. |
| (JONES et al., 2017) | Estudo <i>in silico</i> de identificação de epítomos potenciais da proteína NS1-DENV2 frente a um painel de anticorpos protetivos. O estudo indica a potencialidade do domínio wing nesse processo de reconhecimento NS1-abd, com relevância para os seguintes epítomos: a.a. 97-116; a.a. 125-134; a.a. 143-151. |
| (FREIRE et al., 2017) | Estudo de predição <i>in silico</i> de epítomos da NS1 da DENV2 e ZIKV, pela metodologia de consenso entre os resultados dos <i>webservers</i> ElliPro, Epitopia e DiscoTope, com reatividade confirmada experimentalmente pela técnica de microarranjo de peptídeos. |
| (WANG et al., 2017a) | Estudos cristalográfico que indicou similaridade de 13 dos 21 resíduos de a.a. nas interfaces de interação entre o domínio β -ladder das proteínas NS1 das flavirozes WNV e ZIKV com o anticorpo 22-anti-NS1. |
| (LEE et al., 2018) | Estudo experimental de identificação de epítomos da NS1-ZIKV e sua atividade de ligação a anticorpos em soro de pacientes infectados. Dos nove epítomos com 100 % de afinidade, seis encontram-se localizados no domínio β -ladder. Dos quais as sequências compreendidas por a.a. 236-250, a.a. 246-275, a.a. 266-280 e a.a. 341-352 estão localizadas em região com potencial eletrostático de superfície distinto em relação a NS1-DENV. |

Portanto, para a NS1-DENV2, conforme critérios adotados por Silva (2020), foram considerados apenas os epítomos localizados no domínio wing e, portanto, as seguintes sequências a.a. 50-116 (Epítomo 1 – E1); a.a. 122-134 (Epítomo 2 – E2) e a.a. 143-151 (Epítomo 3 – E3), conforme representado na Figura 1.9a.

Para a NS1-ZIKV foram considerados nessa análise tanto os epítomos localizados no domínio wing quanto no domínio β -ladder. Portanto, foram consideradas as seguintes sequências: a) Domínio wing: a.a. 70-84 (Epítomo 1 – E1) e a.a. 118-147 (Epítomo 2 – E2); b) Domínio β -ladder: a.a. 236-280 (Epítomo 3 – E3) e a.a. 341-352 (Epítomo 4 – E4), conforme representado na Figura 1.9b. (FREIRE et al., 2017; LEE et al., 2018; WANG et al., 2017).

Figura 1.9 – Representação *new cartoon* das superfícies β -ladder (esquerda) e de loops (direita) dos dímeros da proteína NS1 da Dengue (a) e da Zika (b), com identificação *wireframe surface* dos epítomos descritos na literatura e representados nas cores: verde (E1), magenta (E2), alaranjado (E3) e azul (E4).



Fonte: Autor

1.2 OBJETIVOS

1.2.1 Objetivo Geral

O presente trabalho tem como objetivo compreender e elucidar, *in silico*, as interações moleculares específicas de complexos antígeno-anticorpo de caráter nativo, formados pela proteína NS1, enquanto marcador antigênico de flavirozes, e um anticorpo monoclonal. Desse modo, fornecer as bases moleculares e verificar as potencialidades de aplicação do sistema em um imunossensor construído com AFM para diagnóstico precoce e diferencial entre a Dengue e Zika, visando contornar as limitações dos métodos atuais, em especial, a reatividade cruzada.

1.2.2 Objetivos Específicos

- Simular a dinâmica estrutural das proteínas envolvidas nos complexos estudados (NS1 DENV2, NS1 ZIKV e 22-anti-NS1) para analisar seus parâmetros estruturais e energéticos, a fim de selecionar a estrutura proteica mais representativa em solução salina (Capítulo 2);
- Identificar a região de interação e o modo de ligação mais favorável para os sistemas antígeno-anticorpo (NS1/22-anti-NS1) com caráter nativo, por meio de um protocolo proposto que utiliza metodologias de Docking Molecular, conhecimento prévio do sistema em estudo e avaliação dos critérios CAPRI (Capítulo 3);
- Calcular as entalpias de formação dos complexos antígeno-anticorpo (NS1/22-anti-NS1), utilizando métodos quânticos semi-empíricos e as energias livres de ligação para mensurar as forças de interação (adesão) desses complexos com a finalidade de verificar a viabilidade da aplicação no diagnóstico diferencial das flavirozes Dengue e Zika (Capítulo 4).

REFERÊNCIAS

- AKEY, D. L. et al. Flavivirus NS1 structures reveal surfaces for associations with membranes and the immune system. **Science**, v. 343, n. 6173, p. 881–885, 2014.
<https://doi.org/10.1126/science.1247749>
- AMARANTE, A. M. et al. Modeling the coverage of an AFM tip by enzymes and its application in nanobiosensors. **Journal of Molecular Graphics and Modelling**, v. 53, p. 100–104, 2014. <https://doi.org/10.1016/j.jmgm.2014.07.009>
- AMORIM, J. H. et al. The dengue virus non-structural 1 protein: Risks and benefits. **Virus Research**, v. 181, p. 53–60, 2014. <https://doi.org/10.1016/j.virusres.2014.01.001>
- ANTUNES, P. et al. Quantification of NS1 dengue biomarker in serum via optomagnetic nanocluster detection. **Scientific Reports**, v. 5, n. 16145, p. 1–10, 2015.
<https://doi.org/10.1038/srep16145>
- BARBI, L. et al. Prevalence of Guillain-Barré syndrome among Zika virus infected cases: a systematic review and meta-analysis. **Brazilian Journal of Infectious Diseases**, v. 22, n. 2, p. 137–141, 2018. <https://doi.org/10.1016/j.bjid.2018.02.005>
- BEESETTI, H.; KHANNA, N.; SWAMINATHAN, S. Drugs for dengue: a patent review (2010 – 2014). **Expert Opinion on Therapeutic Patents**, v. 24, n. 11, p. 1171–1184, 2014.
<https://doi.org/10.1517/13543776.2014.967212>
- BHATT, S. et al. The global distribution and burden of dengue. **Nature**, v. 496, n. 7446, p. 504–507, 2013. <https://doi.org/10.1038/nature12060>
- BLITVICH, B. J. et al. Determination of the intramolecular disulfide bond arrangement and biochemical identification of the glycosylation sites of the nonstructural protein NS1 of Murray Valley encephalitis virus. **Journal of General Virology**, v. 82, n. 9, p. 2251–2256, 2001. <https://doi.org/10.1099/0022-1317-82-9-2251>
- BOWEN, W. R.; HILAL, N. Future Prospects. In: **Atomic Force Microscopy in Process Engineering**. 1. ed. Oxford: Elsevier Inc., 2009. p. 275–279. <https://doi.org/10.1016/B978-1-85617-517-3.00010-9>
- BRADY, O. J. et al. The association between zika virus infection and microcephaly in brazil 2015–2017: An observational analysis of over 4 million births. **PLoS Medicine**, v. 16, n. 3, p. e1002755, 2019. <https://doi.org/10.1371/journal.pmed.1002755>
- BROWN, W. C. et al. Extended surface for membrane association in Zika virus NS1 structure. **Nature Structural and Molecular Biology**, v. 23, n. 9, p. 865–867, 2016.
<https://doi.org/10.1038/nsmb.3268>
- BUENO, C. C. et al. Nanobiosensor for diclofop detection based on chemically modified AFM probes. **IEEE Sensors Journal**, v. 14, n. 5, p. 1467–1475, 2014.
<https://doi.org/10.1109/JSEN.2014.2301997>
- BUTT, H. J.; CAPPELLA, B.; KAPPL, M. Force measurements with the atomic force microscope: Technique, interpretation and applications. **Surface Science Reports**, v. 59, n. 1–6, p. 1–152, 2005. <https://doi.org/10.1016/j.surfrep.2005.08.003>

CAMARA, A. R. et al. Dengue immunoassay with an LSPR fiber optic sensor. **Optics Express**, v. 21, n. 22, p. 27023, 2013. <https://doi.org/10.1364/OE.21.027023>

COLLINS, M. H.; METZ, S. W. Progress and Works in Progress: Update on Flavivirus Vaccine Development. **Clinical Therapeutics**, v. 39, n. 8, p. 1519–1536, 2017. <https://doi.org/10.1016/j.clinthera.2017.07.001>

COMISSÃO DE EPIDEMIOLOGIA DA ABRASCO. **Zika vírus: desafios da saúde pública no Brasil**. Revista Brasileira de Epidemiologia. [s.l.: s.n.]. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1415-790X2016000200225&lng=pt&tlng=pt>.

DEDA, D. K. et al. The use of functionalized AFM tips as molecular sensors in the detection of pesticides. **Materials Research**, v. 16, n. 3, p. 683–687, 2013. <https://doi.org/10.1590/S1516-14392013005000043>

EDELING, M. A.; DIAMOND, M. S.; FREMONT, D. H. Structural basis of flavivirus NS1 assembly and antibody recognition. **Proceedings of the National Academy of Sciences of the United States of America**, v. 111, n. 11, p. 4285–4290, 2014. <https://doi.org/10.1073/pnas.1322036111>

FALCONAR, A. K. I.; YOUNG, P. R.; MILES, M. A. Precise location of sequential dengue virus subcomplex and complex B cell epitopes on the nonstructural-1 glycoprotein. **Archives of Virology**, v. 137, n. 3–4, p. 315–326, 1994. <https://doi.org/10.1007/BF01309478>

FERREIRA, A. A. P.; YAMANAKA, H. Microscopia de força atômica aplicada em imunoensaios. **Química Nova**, v. 29, n. 1, p. 137–142, 2006. <https://doi.org/10.1590/S0100-40422006000100024>

FERREIRA, A. W. Dengue. In: **Diagnóstico laboratorial das principais doenças infecciosas e autoimunes: correlações clínico-laboratoriais**. 3. ed. Rio de Janeiro: Guanabara Koogan, 2015. p. 70–72.

FERREIRA, M. F.; FRANCA, E. F.; LEITE, F. L. Unbinding pathway energy of glyphosate from the EPSPs enzyme binding site characterized by Steered Molecular Dynamics and Potential of Mean Force. **Journal of Molecular Graphics and Modelling**, v. 72, p. 43–49, 2017. <https://doi.org/10.1016/j.jmgm.2016.11.010>

FLAMAND, M. et al. Dengue Virus Type 1 Nonstructural Glycoprotein NS1 Is Secreted from Mammalian Cells as a Soluble Hexamer in a Glycosylation-Dependent Fashion. **Journal of Virology**, v. 73, n. 7, p. 6104–6110, 1999. <https://doi.org/10.1128/JVI.73.7.6104-6110.1999>

FRANCA, E. F. et al. Designing an enzyme-based nanobiosensor using molecular modeling techniques. **Physical chemistry chemical physics : PCCP**, v. 13, n. 19, p. 8894–8899, 2011. <https://doi.org/10.1039/c1cp20393b>

FREIRE, M. C. L. C. et al. Mapping Putative B-Cell Zika Virus NS1 Epitopes Provides Molecular Basis for Anti-NS1 Antibody Discrimination between Zika and Dengue Viruses. **ACS Omega**, v. 2, p. 3913–3920, 2017. <https://doi.org/10.1021/acsomega.7b00608>

FRITZ, J. Translating Biomolecular Recognition into Nanomechanics. **Science**, v. 288, n. 5464, p. 316–318, 2000. <https://doi.org/10.1126/science.288.5464.316>

GOURINAT, A. C. et al. Detection of Zika Virus in Urine. **Emerging Infectious Diseases**, v. 21, n. 1, p. 84–86, 2015. <https://doi.org/10.3201/eid2101.140894>

GRANDE, A. J. et al. Tourniquet Test for Dengue Diagnosis: Systematic Review and Meta-analysis of Diagnostic Test Accuracy. **PLOS Neglected Tropical Diseases**, v. 10, n. 8, p. e0004888, 2016. <https://doi.org/10.1371/journal.pntd.0004888>

GUARNER, J.; HALE, G. L. Four human diseases with significant public health impact caused by mosquito-borne flaviviruses: West Nile, Zika, dengue and yellow fever. **Seminars in Diagnostic Pathology**, v. 36, n. 3, p. 170–176, 2019. <https://doi.org/10.1053/j.semdp.2019.04.009>

GUBLER, D. J. Dengue and Dengue Hemorrhagic Fever. **Clinical Microbiology Reviews**, v. 11, n. 3, p. 480–496, 1998. <https://doi.org/10.1128/CMR.11.3.480>

GUERRA, R. F. et al. Molecular modelling to understand AFM tip functionalisation and imazaquin-AHAS interactions in order to design a new nanobiosensor. **Molecular Simulation**, v. 45, n. 16, p. 1332–1340, 2019. <https://doi.org/10.1080/08927022.2019.1634266>

GUZMAN, M. G.; HARRIS, E. Dengue. **The Lancet**, v. 385, n. 9966, p. 453–465, 2015. [https://doi.org/10.1016/S0140-6736\(14\)60572-9](https://doi.org/10.1016/S0140-6736(14)60572-9)

HORSTICK, O.; TOZAN, Y.; WILDER-SMITH, A. Reviewing Dengue: Still a Neglected Tropical Disease? **PLoS Neglected Tropical Diseases**, v. 9, n. 4, p. e0003632, 2015. <https://doi.org/10.1371/journal.pntd.0003632>

HOTTA, S. Experimental Studies on Dengue: I. Isolation, Identification and Modification of the Virus. **Journal of Infectious Diseases**, v. 90, n. 1, p. 1–9, 1952. <https://doi.org/10.1093/infdis/90.1.1>

HULO, C. D. C. E. et al. **ViralZone - Flavivirus**. Disponível em: <<https://viralzone.expasy.org/24>>. Acesso em: 14 jun. 2020.

IERICH, J. C. M. et al. A Computational Protein Structure Refinement of the Yeast Acetohydroxyacid Synthase. **Journal of the Brazilian Chemical Society**, v. 26, n. 8, p. 1702–1709, 2015. <https://doi.org/10.5935/0103-5053.20150144>

IERICH, J. C. M. et al. Antibody-mediated biorecognition of myelin oligodendrocyte glycoprotein: computational evidence of demyelination-related epitopes. **Scientific Reports**, v. 9, n. 1, p. 1–10, 2019. <https://doi.org/10.1038/s41598-018-36578-8>

ITO, T.; IBRAHIM, S.; GRABOWSKA, I. Chemical-force microscopy for materials characterization. **TrAC Trends in Analytical Chemistry**, v. 29, n. 3, p. 225–233, 2010. <https://doi.org/10.1016/j.trac.2009.12.008>

JALILI, N.; LAXMINARAYANA, K. A review of atomic force microscopy imaging systems: application to molecular metrology and biological sciences. **Mechatronics**, v. 14, n. 8, p. 907–945, 2004. <https://doi.org/10.1016/j.mechatronics.2004.04.005>

JIANG, L. et al. Selection and identification of B-cell epitope on NS1 protein of dengue virus type 2. **Virus Research**, v. 150, n. 1–2, p. 49–55, 2010. <https://doi.org/10.1016/j.virusres.2010.02.012>

- JONES, M. et al. Computational Identification of Antibody Epitopes on the Dengue Virus NS1 Protein. **Molecules**, v. 22, n. 4, p. 607, 2017. <https://doi.org/10.3390/molecules22040607>
- KATZ, E. Y. A chemically modified electrode capable of a spontaneous immobilization of amino compounds due to its functionalization with succinimidyl groups. **Journal of Electroanalytical Chemistry and Interfacial Electrochemistry**, v. 291, n. 1–2, p. 257–260, 1990. [https://doi.org/10.1016/0022-0728\(90\)87193-N](https://doi.org/10.1016/0022-0728(90)87193-N)
- KENNELLY, P. J.; RODWELL, V. W. Bioinformática e Biologia Computacional. In: MURRAY, R. K. et al. (Eds.). **Bioquímica Ilustrada de Harper**. 29. ed. Porto Alegre: AMGH, 2014. p. 818.
- LARKIN, M. A. et al. Clustal W and Clustal X version 2.0. **Bioinformatics**, v. 23, n. 21, p. 2947–2948, 2007. <https://doi.org/10.1093/bioinformatics/btm404>
- LEE, H. J. et al. Identification of peptide based B-cell epitopes in Zika virus NS1. **Biochemical and Biophysical Research Communications**, v. 505, n. 4, p. 1010–1014, 2018. <https://doi.org/10.1016/j.bbrc.2018.10.024>
- LI, L. et al. In Situ Single-Molecule Detection of Antibody–Antigen Binding by Tapping-Mode Atomic Force Microscopy. **Analytical Chemistry**, v. 74, n. 23, p. 6017–6022, 2002. <https://doi.org/10.1021/ac0258148>
- LINDENBACH, B. D.; RICE, C. M. Molecular biology of flaviviruses. In: **Advances in Virus Research**. 59. ed. Nova Iorque: Elsevier Inc., 2003. p. 23–61. [https://doi.org/10.1016/S0065-3527\(03\)59002-9](https://doi.org/10.1016/S0065-3527(03)59002-9)
- LOVER, A. A. Zika virus and microcephaly. **The Lancet Infectious Diseases**, v. 16, n. 12, p. 1331–1332, 2016. [https://doi.org/10.1016/S1473-3099\(16\)30462-5](https://doi.org/10.1016/S1473-3099(16)30462-5)
- MARTINEZ, J. D.; GARZA, J. A. C. L. A.; CUELLAR-BARBOZA, A. Going Viral 2019. **Dermatologic Clinics**, v. 37, n. 1, p. 95–105, 2019. <https://doi.org/10.1016/j.det.2018.07.008>
- MELO, K. M. S. **Desenvolvimento e Validação de insumos para diagnóstico de infecções pelo vírus da Dengue**. Dissertação (Mestrado em Saúde Pública), Departamento de Saúde Coletiva, Fundação Oswaldo Cruz. Recife, 178 p., 2012.
- MINISTÉRIO DA SAÚDE. **Dengue - Informações Técnicas**. Disponível em: <<https://www.saude.gov.br/saude-de-a-z/dengue/informacoes-tecnicas>>. Acesso em: 24 maio. 2020.
- MLACKER, S. et al. Origin of the Zika virus revealed: a historical journey across the world. **International Journal of Dermatology**, v. 55, n. 12, p. 1369–1372, 2016. <https://doi.org/10.1111/ijd.13399>
- MULLER, D. A.; DEPELSENAIRE, A. C. I.; YOUNG, P. R. Clinical and laboratory diagnosis of dengue virus infection. **Journal of Infectious Diseases**, v. 215, p. S89–S95, 2017. <https://doi.org/10.1093/infdis/jiw649>
- MULLER, D. A.; YOUNG, P. R. The flavivirus NS1 protein: Molecular and structural biology, immunology, role in pathogenesis and application as a diagnostic biomarker. **Antiviral Research**, v. 98, n. 2, p. 192–208, 2013. <https://doi.org/10.1016/j.antiviral.2013.03.008>

OLIVEIRA, G. S. et al. Molecular Modeling of Enzyme Attachment on AFM Probes. **Journal of Molecular Graphics & Modelling**, v. 45, p. 128–136, 2013. <https://doi.org/10.1016/j.jmgm.2013.08.007>

OLIVEIRA, G. S. et al. Immobilization and unbinding investigation of the antigen-antibody complex using theoretical and experimental techniques. **Journal of Molecular Graphics and Modelling**, v. 86, p. 219–227, 2019. <https://doi.org/10.1016/j.jmgm.2018.10.012>

OPS. **Temas de Saúde**. Disponível em: <<https://www.paho.org/data/index.php/es/>>. Acesso em: 26 dez. 2020.

OSUNA, C. E. et al. Zika viral dynamics and shedding in rhesus and cynomolgus macaques. **Nature Medicine**, v. 22, n. 12, p. 1448–1455, 2016. <https://doi.org/10.1038/nm.4206>

PETERS, R.; STEVENSON, M. Zika virus diagnosis: challenges and solutions. **Clinical Microbiology and Infection**, v. 25, n. 2, p. 142–146, 2019. <https://doi.org/10.1016/j.cmi.2018.12.002>

PINHEIRO, F. P.; CORBER, S. J. Global situation of dengue and dengue haemorrhagic fever, and its emergence in the Americas. **World Health Statistics Quarterly**, v. 50, p. 161–169, 1997.

RABE, I. B. et al. Interim Guidance for Interpretation of Zika Virus Antibody Test Results. **MMWR. Morbidity and Mortality Weekly Report**, v. 65, n. 21, p. 543–546, 2016. <https://doi.org/10.15585/mmwr.mm6521e1>

RASTOGI, M.; SHARMA, N.; SINGH, S. K. Flavivirus NS1: a multifaceted enigmatic viral protein. **Virology Journal**, v. 13, n. 1, p. 131, 2016. <https://doi.org/10.1186/s12985-016-0590-7>

RICE, C. et al. Nucleotide sequence of yellow fever virus: implications for flavivirus gene expression and evolution. **Science**, v. 229, n. 4715, p. 726–733, 1985. <https://doi.org/10.1126/science.4023707>

ROOS, R. P. Zika Virus—A Public Health Emergency of International Concern. **JAMA Neurology**, v. 73, n. 12, p. 1395–1396, 2016. <https://doi.org/10.1001/jamaneurol.2016.3677>

SCATURRO, P. et al. Dengue Virus Non-structural Protein 1 Modulates Infectious Particle Production via Interaction with the Structural Proteins. **PLOS Pathogens**, v. 11, n. 11, p. e1005277, 2015. <https://doi.org/10.1371/journal.ppat.1005277>

SILVA, V. L. O. DA. **Desenvolvimento de RT-PCR em tempo real para diagnóstico diferencial de chikungunya, dengue e Zika**. Dissertação (Mestrado em Ciências), Programa de Pós-Graduação em Ciências e Biotecnologia, Universidade de São Paulo. Ribeirão Preto, 103 p., 2019.

SILVA, L. A. **Estudo da interação do complexo antiNS1 / NS1-DENV2 usando ancoramento molecular e métodos quânticos**. Dissertação (Mestrado em Química), Programa de Pós-Graduação em Química, Universidade Federal de Uberlândia. Uberlândia, 58 p., 2020.

SNOW, G. E. et al. Research on Dengue During World War II Revisited. **The American Journal of Tropical Medicine and Hygiene**, v. 91, n. 6, p. 1203–1217, 2014.

<https://doi.org/10.4269/ajtmh.14-0132>

VAUGHN, D. W. et al. Dengue in the Early Febrile Phase: Viremia and Antibody Responses. **The Journal of Infectious Diseases**, v. 176, n. 2, p. 322–330, 1997.

<https://doi.org/10.1086/514048>

WANG, D. et al. A mutation identified in neonatal microcephaly destabilizes zika virus NS1 assembly in vitro. **Scientific Reports**, v. 7, n. 42580, p. 1–10, 2017.

<https://doi.org/10.1038/srep42580>

WANG, L. et al. From Mosquitos to Humans: Genetic Evolution of Zika Virus. **Cell Host & Microbe**, v. 19, n. 5, p. 561–565, 2016. <https://doi.org/10.1016/j.chom.2016.04.006>

WEAVER, S. C.; VASILAKIS, N. Molecular evolution of dengue viruses: Contributions of phylogenetics to understanding the history and epidemiology of the preeminent arboviral disease. **Infection, Genetics and Evolution**, v. 9, n. 4, p. 523–540, 2009.

<https://doi.org/10.1016/j.meegid.2009.02.003>

WHO. **Mosquito sterilization offers new opportunity to control dengue, Zika and chikungunya**. Disponível em: <<https://www.who.int/tdr/news/2019/mosquito-sterilization-new-opportunities-to-control-diseases/en/>>. Acesso em: 8 abr. 2020.

WHO. **Dengue and severe dengue**. Disponível em: <<https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue>>. Acesso em: 23 mai. 2020.

XU, X. et al. Contribution of intertwined loop to membrane association revealed by Zika virus full-length NS 1 structure . **The EMBO Journal**, v. 35, n. 20, p. 2170–2178, 2016.

<https://doi.org/10.15252/embj.201695290>

YADAV, S.; RAWAL, G.; BAXI, M. Zika virus: A pandemic in progress. **Journal of Translational Internal Medicine**, v. 4, n. 1, p. 42–45, 2016. <https://doi.org/10.1515/jtim-2016-0009>

YOUNG, P. R. et al. An antigen capture enzyme-linked immunosorbent assay reveals high levels of the dengue virus protein NS1 in the sera of infected patients. **Journal of Clinical Microbiology**, v. 38, n. 3, p. 1053–1057, 2000. <https://doi.org/10.1128/JCM.38.3.1053-1057.2000>

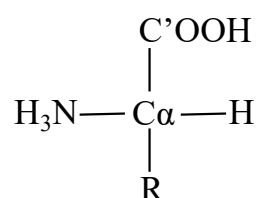
PROTEÍNAS: ANTÍGENO NS1 E ANTICORPO 22-ANTI-NS1

2.1 AMINOÁCIDOS E PROTEÍNAS

Dentre as macromoléculas biológicas, as proteínas são as mais abundantes, pois estão no centro da ação dos processos biológicos, sendo encontradas em todas as células. O termo proteína vem do grego, *proteios*, e significa primário. Tal terminologia foi adotada por G. J. Mulder (1983) ao denominar uma classe de compostos naturais contendo o elemento químico nitrogênio e que são essenciais à nutrição animal (VOET; VOET; PRATT, 2008).

As proteínas podem ser encontradas na forma de peptídeos ou de moléculas poliméricas resultantes da combinação das suas unidades estruturais, os aminoácidos. Estruturalmente, todos eles possuem em comum um átomo de carbono central ($C\alpha$) ligado a um átomo de hidrogênio e aos grupamentos amino (NH_3) e carboxil ($C'OOH$). No entanto, se distinguem pela cadeia lateral (grupo R) ligada ao $C\alpha$ através de sua quarta valência, conforme apresentado na Figura 2.1 (NELSON; COX, 2011).

Figura 2.1 – Estrutura geral dos α -aminoácidos.



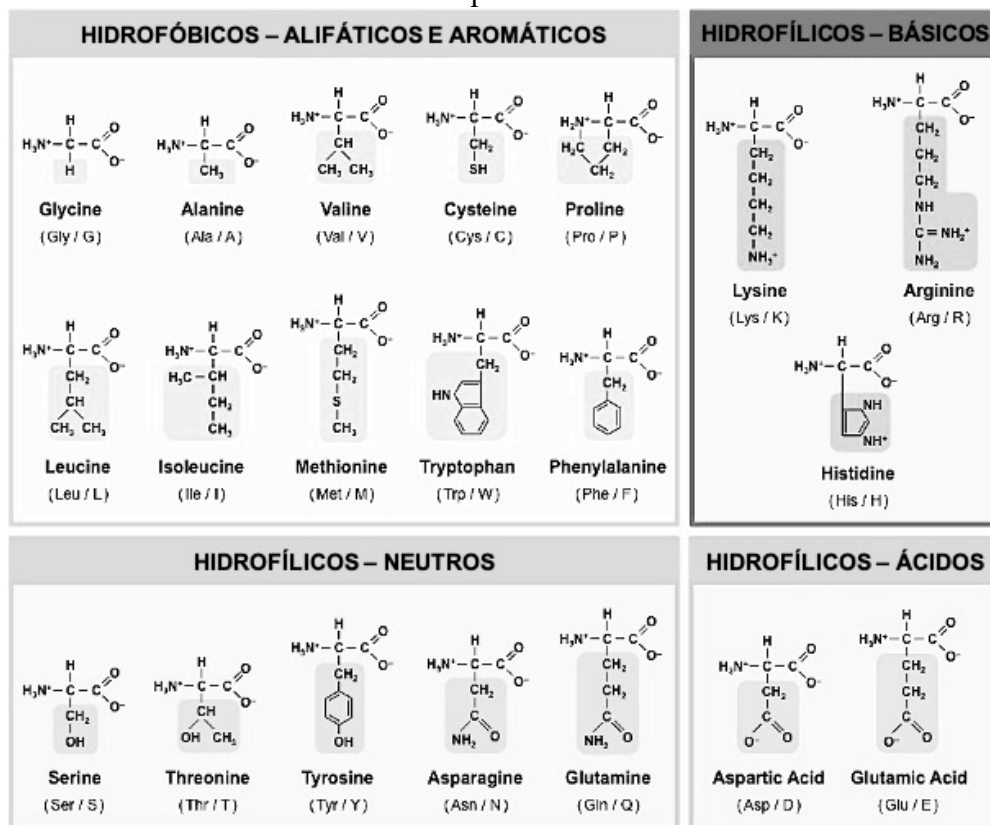
Fonte: Autor

Apesar da ocorrência natural de mais de 300 aminoácidos, apenas 20 deles constituem unidades fundamentais usadas pelas células para formarem as proteínas. Estes são os chamados α -aminoácidos. Com exceção da glicina, que possui um átomo de hidrogênio como cadeia

lateral, a quiralidade é uma característica química de todos os aminoácidos, visto que o $C\alpha$ está covalentemente ligado a quatro grupos distintos. Embora haja, para cada uma dessas estruturas, dois estereoisômeros (L- e D-), todas compartilham a configuração absoluta do L-glutaraldeído sendo, portanto, denominadas, L- α -aminoácidos (MURRAY et al., 2014).

O critério utilizado na classificação dos aminoácidos é a natureza química da cadeia lateral (grupamento R), ou seja, de acordo com sua polaridade. Portanto, os aminoácidos são divididos em dois grupos: hidrofóbicos (alifáticos e aromáticos) e hidrofílicos (ácidos, básicos e neutros). Os grupamentos R são, portanto, também os responsáveis por determinar as funções bioquímicas próprias dos aminoácidos. Na Figura 2.2 estão apresentadas as estruturas dos aminoácidos, classificadas de acordo com as características fundamentais da cadeia lateral (porções sombreadas) (NELSON; COX, 2011).

Figura 2.2 – Representação esquemática, identificação e classificação das estruturas dos 20 aminoácidos mais comuns encontrados em proteínas.

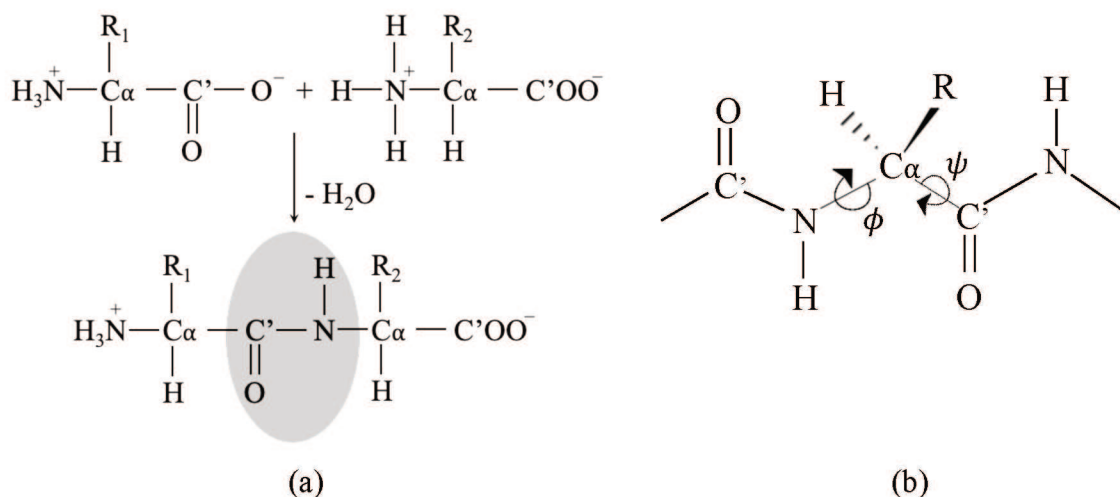


Fonte: Adaptado de Cornell (2016).

As estruturas apresentadas na Figura 2.2 estão em seu estado de ionização predominante em pH neutro, ou seja, nas suas respectivas formas iônicas em que os grupamentos amino e carboxil estão carregados positiva e negativamente, respectivamente. Isso se deve a uma particularidade do comportamento químico dessas estruturas: o caráter zwitteriônico. Um zwitterion em solução aquosa pode se comportar como ácido (doador de prótons) ou como base (receptor de prótons) (SOLOMONS, 2005).

Quando muitos desses aminoácidos se ligam, formam-se cadeias polipeptídicas que, na grande maioria dos casos, isoladas ou aos pares, resultam nas proteínas. Os aminoácidos ligam-se uns aos outros por meio de uma reação de condensação, em que o agrupamento α -amino de um se liga ao agrupamento α -carboxil do outro resultando na eliminação de uma molécula de água (Figura 2.3a). A ligação resultante é denominada ligação peptídica, caracterizada por sua rigidez e planaridade. Isso porque, embora representada por uma ligação covalente simples, a ligação peptídica possui caráter de dupla ligação parcial em função do efeito ressonante estabilizador. Desse modo, a liberdade de rotação em torno dessa ligação é restringida, tornando os átomos O, C', N e H coplanares. Portanto, o modo como os peptídeos se arranjam para gerar estruturas de ordem superior é direcionada unicamente pelas rotações sobre as ligações N-C α e C α -C'. Os ângulos definidos por essas rotações são, por convenção, denominados de phi (ϕ) e psi (Ψ), respectivamente, conforme apresentado na Figura 2.3b (BRANDEN; TOOZE, 2009; MURRAY et al., 2014).

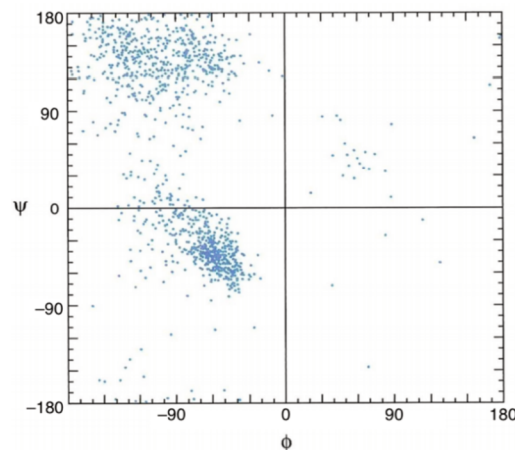
Figura 2.3 – Reação de condensação entre aminoácidos formando uma ligação peptídica (a) e suas rotações permitidas sobre as ligações N-C α e C α -C' definidas, respectivamente, pelos ângulos diedros ϕ e Ψ (b).



Fonte: Autor

A maioria das combinações entre ângulos ϕ e Ψ não são permitidas em função do efeito estérico entre as cadeias lateral e principal, seja no mesmo peptídeo ou entre peptídeos diferentes. Os cálculos de regiões estericamente permitidas foram realizados pelo biofísico N. G. Ramachandran que apresentou, em 1963, o diagrama de Ramachandran, que contém os pares desses ângulos plotados na forma ϕ versus Ψ (RAMACHANDRAN; RAMAKRISHNAN; SASISEKHARAN, 1963). A Figura 2.4 apresenta o diagrama de Ramachandran dos ângulos ϕ e Ψ observados para todos os tipos de resíduos de aminoácidos, exceto glicina, de proteínas cujas estruturas foram solucionadas em alta resolução por meio de cristalografia de raios-x (RICHARDSON, 1981).

Figura 2.4 – Exemplo de um diagrama de Ramachandran.



Fonte: Adaptado de Murray et al. (2014)

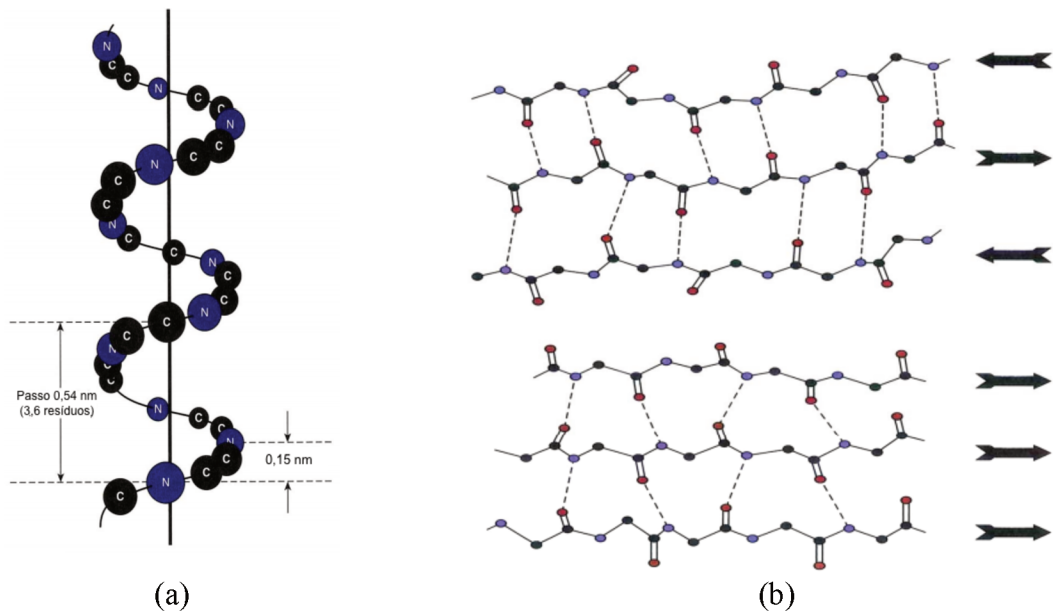
Na Figura 2.4 as regiões definidas pelos pontos azuis e os espaços brancos representam combinações permitidas e proibidas, respectivamente. As regiões permitidas correspondem às conformações adquiridas pelos peptídeos na estrutura secundária, em especial, os tipos mais comuns: α -hélice e folha β . Os ângulos que as definem caem, respectivamente, dentro dos quadrantes esquerdos superior e inferior (MURRAY et al., 2014).

A conformação α -hélice possui uma estrutura helicoidal torcida por uma quantidade igual de cada $C\alpha$ com ϕ e Ψ aproximados de -60° e -50° , respectivamente. Possui a cada volta cerca de 3,6 resíduos com os grupamentos R voltados para fora, de modo que ligações de

hidrogênio sejam formadas entre o átomo de oxigênio ligado ao C' e o hidrogênio ligado ao N do 4º resíduo adiante (Figura 2.5a) (BRANDEN; TOOZE, 2009).

Nas folhas β , os resíduos formam um padrão zigue-zague no qual os resíduos apontam em direções opostas. Sua estabilidade advém das ligações de hidrogênios estabelecidas entre segmentos adjacentes da estrutura. Podem se apresentar nas formas paralela (segmentos na mesma direção) e anti-paralela (segmentos em direções opostas) (Figura 2.5b) (BRANDEN; TOOZE, 2009).

Figura 2.5 – Representação esquemática dos tipos mais comuns de estrutura secundária: α -hélice (a) e folha β (b).



Fonte: Adaptado de Murray et al. (2014)

2.1.1 Imunoglobulinas: anticorpos

As globulinas, assim denominadas por sua estrutura globular, são uma classe de proteínas que exercem um papel importante no sistema imunológico dos vertebrados. Mais conhecidas como anticorpos, as imunoglobulinas (Ig) são “proteínas solúveis que circulam livremente e exibem propriedades que contribuem especificamente para a imunidade e proteção contra material estranho” (COICO; SUNSHINE, 2010, p. 41).

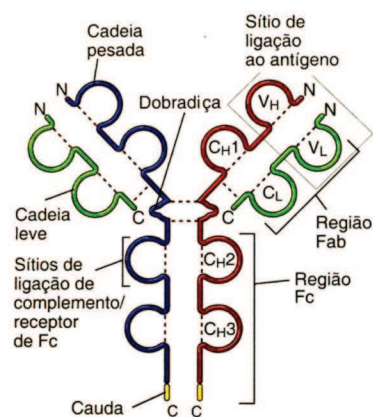
Os anticorpos são, portanto, os mediadores da imunidade humana contra todas as classes de microorganismos, sendo suas principais características a atividade biológica e a especificidade, intrinsicamente ligadas à sua estrutura molecular.

De modo geral, todas as moléculas de anticorpo apresentam características estruturais básicas idênticas, sendo distintas entre si pela elevada variabilidade existente na região de ligação aos antígenos.

A estrutura dos anticorpos (Figura 2.6), além do formato característico em Y, é composta por 3 fragmentos de tamanhos similares: dois fragmentos Fab e um fragmento Fc. O fragmento Fc é a região cristalizável, responsável pelas funções efetoras e propriedades físico-químicas do anticorpo após sua ligação ao antígeno por meio dos fragmentos Fab (do inglês, *fragment antigen-binding*) que, portanto, correspondem à duas regiões idênticas com capacidade de se ligarem ao antígeno específico (ABBAS; LICHTMAN; PILLAI, 2011; COICO; SUNSHINE, 2010).

Esses domínios, por sua vez, são compostos por duas cadeias leves idênticas (Cadeia L – verde) e duas cadeias pesadas (Cadeia H – azul e vermelho) mantidas juntas por várias pontes dissulfeto e compostas por domínios constantes (C_L , C_{H1} , C_{H2} e C_{H3}) e variáveis (V_L e V_H). As regiões variáveis são assim denominadas, pois correspondem a áreas de variabilidade na sequência de aminoácidos que diferem os anticorpos e os conferem especificidades diferentes. Portanto, pelo fato de juntas formarem o sítio de ligação ao antígeno são as de maior interesse em estudos de interação antígeno-anticorpo (ABBAS; LICHTMAN; PILLAI, 2011; COICO; SUNSHINE, 2010).

Figura 2.6 – Representação esquemática da estrutura geral das imunoglobulinas.

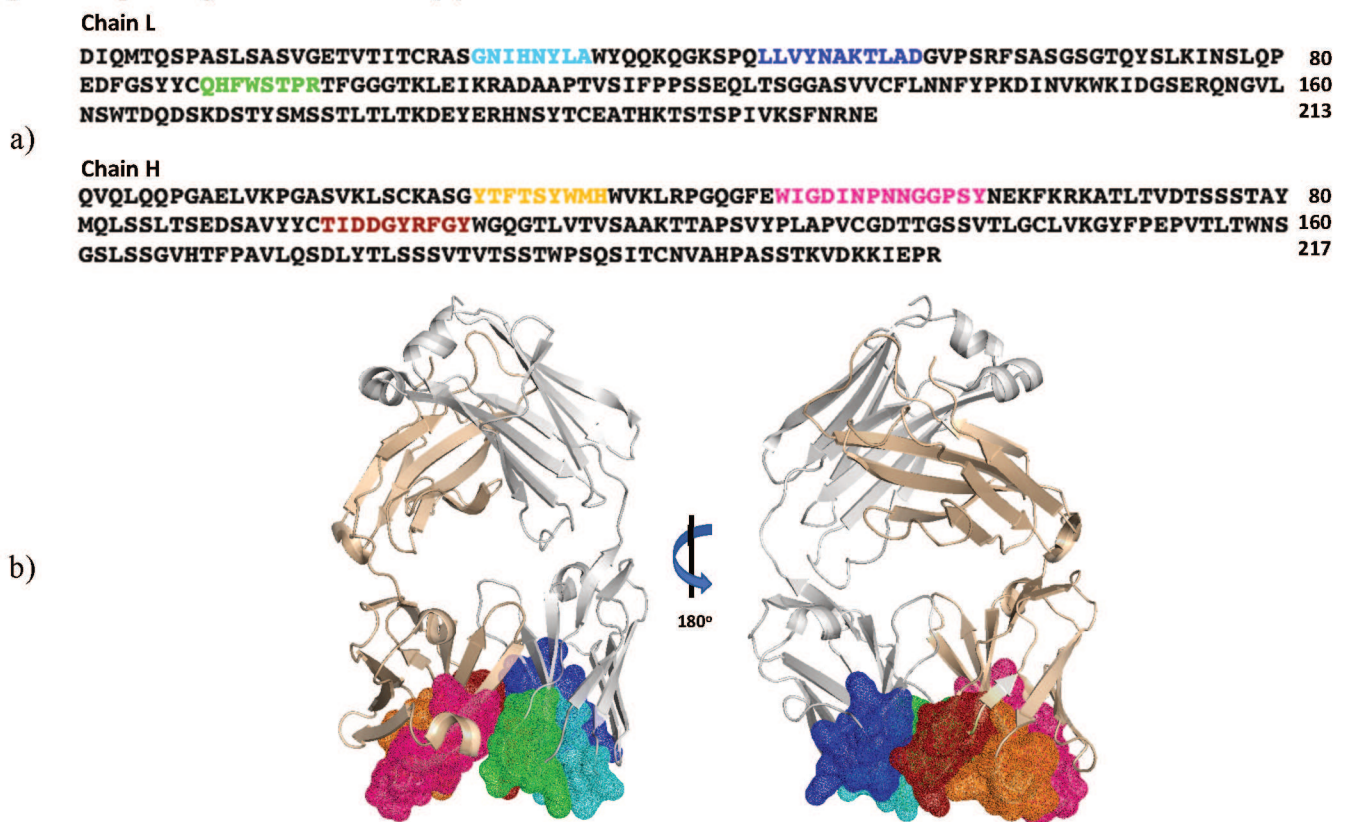


Fonte: (ABBAS; LICHTMAN; PILLAI, 2011)

As porções variáveis das cadeias leve e pesada são ainda compostas por três regiões de hipervariabilidade, denominadas regiões de determinação da complementaridade (CDR). São os aminoácidos presentes nas CDRs – CDR1 (a.a. 31-34), CDR2 (a.a. 50-56) e CDR3 (a.a. 93-99) – que efetivamente participam da ligação ao antígeno (SELA-CULANG; KUNIK; OFRAN, 2013).

O anticorpo anti-NS1 desse estudo é monoclonal do isotipo IgG (gene do domínio j) obtido no estudo de Chung e colaboradores (2006), que testou a imunoafinidade de 22 anticorpos distintos em relação à proteína NS1 da flavirose Vírus do Nilo Ocidental (WNV). A estrutura da sua região variável (Figura 2.7b) foi cristalizada e resolvida na forma de um complexo 2:2 (cadeias A,B:L,H e A,B:M,I) com o domínio β -ladder da WNV-NS1₁₇₂₋₃₅₂, com resolução de 3 Å e um epítipo estrutural composto por 16 resíduos de contato na interface de loop do domínio em questão (EDELING; DIAMOND; FREEMONT, 2014).

Figura 2.7 – Sequência FASTA da região Fab do anticorpo 22-anti-NS1 (a) e sua representação *new cartoon* coloridas em branco (chain L) e bege (chain H) com identificação *wireframe surface* das CDRs preditas pelo *software* Paratome (b).



Fonte: Autor e (KUNIK; ASHKENAZI; OFRAN, 2012; KUNIK; PETERS; OFRAN, 2012)

Conforme sequência FASTA da estrutura Fab do anticorpo (Figura 2.7a), nas CDRs é possível identificar a relevância dos resíduos de aminoácidos polares e não carregados, Ser (S) e Tyr (Y), que correspondem a cerca de 17 % dos resíduos e estão presentes em ao menos um dos paratopos. Esses resíduos estão presentes na CDR para estabelecer um contato maior com o antígeno e, portanto, são intrinsecamente responsáveis pelo reconhecimento molecular, em especial, pelo papel específico de estabelecer ligações de hidrogênio diretas ou mediadas (FELLOUSE et al., 2005, 2006). Além disso, os resíduos de Tyr são responsáveis por estabelecer interações do tipo π - π e CH- π (OSAJIMA et al., 2014).

O comportamento dessas regiões e a forma como ocorre o seu ajuste aos epítomos (região de um antígeno que está diretamente envolvida na interação antígeno-receptor) das proteínas de interesse são objetos de estudo do presente trabalho. Isso porque a afinidade de ligação (medida da força de ligação) do anticorpo com os epítomos serão diferentes em decorrência das diferenças nos números e na natureza das forças de ligações disponíveis, em sua maioria interações não-covalentes fracas, como interações iônicas, ligações de hidrogênio, forças de van der Waals e interações hidrofóbicas. Nesse contexto, o estudo das interações de complexos antígeno-anticorpo pode fornecer bases moleculares para compreensão do processo de reconhecimento das proteínas NS1 das Flaviviruses, Dengue e Zika, a ser aplicado no diagnóstico diferencial e precoce.

2.2 MODELAGEM MOLECULAR DE SISTEMAS BIOLÓGICOS

A compreensão de fenômenos/processos nos quais os sistemas biológicos estão envolvidos passa pelo conhecimento dos seus detalhes a nível atômico e molecular. Pela sua natureza, a maior parte desses sistemas é composta por macromoléculas como as proteínas. Desse modo, os fenômenos biológicos estão intrinsecamente ligados à flexibilidade dessas biomoléculas.

Sendo assim, é relevante o entendimento dos detalhes da dinâmica destes sistemas e, em particular, da análise da conformação geral das estruturas proteicas componentes, determinada diretamente pelas interações estabelecidas pela sua estrutura secundária. Portanto, é necessário que seja avaliada sua estabilidade conformacional em condições fisiológicas para que a disposição espacial garanta a funcionalidade e regulação da proteína. Isso porque a maior parte

dos fenômenos biológicos, bem como os químicos, ocorrem em meio condensado, mais especificamente, em meio aquoso.

Nesse contexto, a modelagem molecular possibilita a representação e estudo desses sistemas biomoleculares em escala atômica com base na premissa de que a compreensão sobre sua função biológica requer conhecimentos acerca da estrutura e da dinâmica e dos efeitos da presença do solvente.

De acordo com a IUPAC, modelagem molecular “é a investigação de estruturas moleculares e propriedades usando química computacional e técnicas de visualização gráfica, a fim de fornecer uma representação tridimensional plausível sob um dado conjunto de circunstâncias”. Resumidamente, os objetivos da modelagem molecular são o uso de modelos teóricos, a nível atômico, que permitam o manuseio de estruturas moleculares e a avaliação do comportamento químico visando a descrever ou interpretar as propriedades macroscópicas, valendo-se de uma série de técnicas computacionais que possibilitam a construção, a visualização, a manipulação e a estocagem de modelos moleculares tridimensionais (LEACH, 2001).

Dentre os métodos de Modelagem Molecular, aqueles de mecânica estatística clássica têm a vantagem de permitir rapidez na avaliação de sistemas maiores e mais complexos, sendo adequados na simulação de processos que não envolvam a transferência eletrônica e quebra e/ou formação de ligação química, visto que os átomos são tratados como as unidades fundamentais e não os elétrons ou prótons. Sendo assim, permitem o entendimento das interações intermoleculares, do comportamento das moléculas em solução e a previsão das propriedades de líquidos, soluções e sólidos (ALLEN; TILDESLEY, 1993).

As simulações de dinâmica molecular são uma metodologia de vasta aplicação e que fornecem a adequada conexão entre a estrutura tridimensional, obtida com o auxílio de técnicas como cristalografia de raios-x e RMN, e a dinâmica em nível molecular, que é mais difícil de ser estudada experimentalmente (KARPLUS; KURIYAN, 2005; PATODIA; BAGARIA; CHOPRA, 2014). Segundo Karplus e Kuriyan (2005, p. 1, nossa tradução)

as simulações de dinâmica molecular fornecem ferramentas poderosas para a exploração do panorama de energia conformacional acessível a essas moléculas, e o rápido aumento no poder computacional associado a melhorias na metodologia torna este um momento promissor para a aplicação da simulação na biologia estrutural.

Em 1979, foi reportada na literatura a primeira simulação de dinâmica molecular de proteína. A lisozima foi simulada, no vácuo, e obtida uma trajetória de 9,2 ps (ARTYMIUK et al., 1979). Anos depois, a mesma proteína foi simulada em meio aquoso numa trajetória de 210 ps (LEVITT; SHARON, 1988). Desde então, o avanço computacional nos aspectos relacionados ao aprimoramento no *hardware* e *software*, em especial, o aumento na velocidade dos cálculos, o uso de memórias maiores e mais eficientes e a obtenção de resultados mais precisos viabilizaram a realização de dinâmicas moleculares de proteínas de pesos moleculares maiores, com elucidação dos efeitos de solvente e concentrações salinas e em tempos da ordem de dezenas e centenas de nanossegundos. Portanto, a DM de sistemas biológicos em meio aquoso permite, que mesmo diante da considerável complexidade das estruturas, as dinâmicas sejam estudadas e propriedades termodinâmicas calculadas com base em estruturas representativas, que correspondem às médias de um extenso conjunto conformacional.

Nos tópicos a seguir, serão abordados com maiores detalhes os princípios da Mecânica Molecular e de sua aplicação através da Dinâmica Molecular, pois se trata da metodologia principal utilizada neste capítulo para estudo das proteínas em questão.

2.2.1 Simulação por Dinâmica Molecular (DM)

Segundo a IUPAC, 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”. Dessa forma, a descrição de sistemas moleculares pelos pacotes de Dinâmica Molecular se baseia na Aproximação de Born-Oppenheimer, segundo a qual, por conta da diferença de massa, a movimentação dos elétrons é consideravelmente mais rápida que a dos núcleos atômicos. Portanto, o movimento dos elétrons pode ser negligenciado e a energia do sistema calculada somente em função das posições nucleares de modo que as leis de Newton sejam aplicáveis (LEVINE, 1991).

Portanto, a descrição conformacional oferecida para os sistemas baseia-se na resolução da 2ª Lei de Newton, que fornece a relação temporal da força F_{x_i} aplicada ao átomo i na posição x , conforme Equação 2.1:

$$\frac{d^2 x_i}{dt^2} = \frac{F_{x_i}}{m_i} \quad (2.1)$$

As equações de Newton na forma diferencial são numericamente integradas simultânea e sucessivamente para cada um dos N átomos do sistema. Esse processo é conhecido como *passo* e a apresentação cronológica de n passos gera o que conhecemos como trajetória. Dessa forma, a cada novo passo, as equações são novamente integradas com adição de determinado Δt , o que acarreta a alteração da velocidade e da posição atômica em função do tempo (CUNHA, 2013). Logo, as propriedades macroscópicas são obtidas a partir das médias temporais do conjunto de trajetórias em equilíbrio (VAN DER SPOEL et al., 2005). A escolha de um tempo adequado de integração tem relação direta com o tamanho da simulação e, conseqüentemente, o custo computacional. Mas vale considerar que esses pequenos intervalos de tempo devem ser da ordem de femtossegundos ($1 \text{ fs} = 10^{-15}$ segundos) ou menores (LEACH, 2001) para que sejam capazes de modelar com precisão suficiente eventos moleculares (rotações e vibrações) que, caracteristicamente, ocorrem em frequências da ordem de 10^{11} - 10^{14} s^{-1} , referentes a comprimentos de onda na faixa de 3 - 3300 cm^{-1} . Além disso, se o tamanho do passo temporal for muito grande, pode acarretar a supressão de eventos moleculares importantes e ocasionar erros significativos nos cálculos das forças e da energia potencial total do sistema.

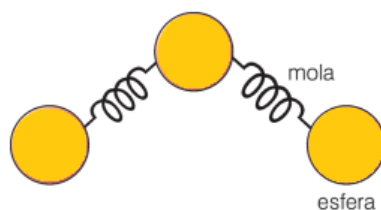
Assim, apesar de cada uma das partículas atômicas do sistema possuírem suas massas atômicas conhecidas, elas estão submetidas a diferentes forças F_i , por estarem ligadas a outros átomos formando moléculas, genericamente descritas conforme representado na Equação 2.2:

$$F_i = - \frac{\partial V}{\partial r_i} \quad (2.2)$$

em que, r_i é o raio atômico e V é o potencial descrito por uma função matemática empírica que considera as forças intra e interatômicas, conforme descritas pelo campo de força utilizado.

Antes de compreender como as interações físicas entre os átomos são descritas, é importante considerar que os modelos, utilizados nos métodos de Mecânica Molecular Clássica, são representados por sistemas massa-mola. Nesses sistemas, as moléculas são tratadas como coleção de partículas (esferas maciças e indivisíveis) unidas por forças harmônicas ou elásticas (molas devidamente calibradas) conforme representado na Figura 2.8 (VAN DER SPOEL et al., 2005).

Figura 2.8 – Representação de uma molécula sob a ótica da Mecânica Molecular.



Fonte: Adaptado de Rodrigues (2001).

Portanto, as características dos sistemas massa-mola estão contidas no campo de força (GUIZADO, 2008). De acordo com a IUPAC, campo de força pode ser descrito como “um conjunto de funções e parametrização usadas em cálculos de mecânica molecular” ao fornecer, em detalhes, informações a respeito dos átomos e suas interações em uma molécula, considere-se aqui os termos relacionados às interações entre átomos ligados e não-ligados entre si (VERLI et al., 2014).

Os métodos de mecânica molecular se fundamentam, dessa forma, no grau de transferibilidade dos parâmetros geométricos de uma molécula à outra, ou seja, na capacidade de representar as moléculas em termos das estruturas de valência. Isso porque o comprimento das ligações e os ângulos podem ser transferidos de uma molécula para outra resultando na diminuição da necessidade de determinação de parâmetros para compostos diferentes, respeitando a semelhança entre suas estruturas químicas e a dependência previsível dos parâmetros da hibridação atômica (HEHRE, 2003).

Diferentes equações (número e natureza dos termos) e parametrizações podem ser utilizadas originando diferentes campos de força (VERLI et al., 2014). Dentre os campos de força mais utilizados atualmente, cita-se: AMBER (*Assisted Model Building with Energy Refinement*), GROMOS (*Groningen Molecular Simulation packages*), OPLS-AA (*Optimized Parameters for Liquid Simulation – all atom*), CHARMM (*Chemistry at Harvard Macromolecular Mechanics*) e Martini. A escolha por um determinado campo de força não é trivial, ou seja, não existe o melhor campo de força. A melhor opção de modelo e campo de força depende do tipo de sistema molecular e o tipo de propriedade na qual se tem interesse (VAN GUNSTEREN; BERENDSEN, 1990). Portanto, para o desenvolvimento deste estudo foi escolhido o campo de força OPLS-AA, visto que seus parâmetros foram obtidos e otimizados para a descrição adequada de sistemas proteicos, além de sua compatibilidade com

o *software* gratuito GROMACS (JORGENSEN; MAXWELL; TIRADO-RIVES, 1996; JORGENSEN; MCDONALD, 1998).

A forma funcional da energia potencial total desse campo de força é dividida entre potenciais ligados e não-ligados. Os potenciais ligados são o potencial harmônico de ligação ou linear (V_d – comprimento de ligação), potencial harmônico angular (V_θ – ângulo de ligação), potencial torsional ou diedral (ângulo diedro) próprio ($V_{\varphi\text{prop}}$) e impróprio ($V_{\varphi\text{improp}}$). Os potenciais não-ligados são potencial de van der Waals ou de Lennard-Jones (V_{LJ}) e potencial de Coulomb (V_C – eletrostáticas).

$$V_{\text{total}} = V_d + V_\theta + V_{\varphi\text{prop}} + V_{\varphi\text{improp}} + V_{LJ} + V_C \quad (2.3)$$

O significado e as expressões de cada um dos termos da Equação 2.3 para campo de força OPLS-AA estão descritos na Tabela 2.1.

Tabela 2.1 – Descrição dos termos de potenciais ligados e não-ligados do campo de força OPLS-AA utilizado nos cálculos de DM.

| Potencial | | Descrição |
|---|--|---|
| Ligados | Harmônico Linear | Também conhecido como potencial de ligação química e descreve a energia associada ao estiramento entre dois átomos ligados como o desvio a uma distância de equilíbrio, d_0 (Equação 2.4). |
| | | $V_d = k_d (d - d_0)^2$ (2.4) |
| | em que: d é o comprimento da ligação entre os dois átomos, d_0 é o comprimento de equilíbrio e k_d é a constante de força da ligação. Sendo d_0 e k_d os parâmetros envolvidos no campo de força. | |
| | Harmônico Angular | Descreve o desvio de ângulos, formados entre três átomos ligados por ligações covalentes como o desvio em relação ao seu valor de equilíbrio θ_0 , (Equação 2.5). |
| | | $V_\theta = k_\theta (\theta - \theta_0)^2$ (2.5) |
| Onde: θ é o ângulo definido pela ligação, θ_0 é o ângulo de equilíbrio e k_θ é a constante elástica da força. | | |
| Torsional/Diedral Próprio | Considera-se um conjunto de quatro átomos ligados (i, j, k e l) onde o ângulo diedro φ é definido como o ângulo entre os planos formados pelos átomos ijk e jkl , em que o valor zero corresponde à conformação cis, na qual os átomos i e j estão do mesmo lado. Dessa forma, V_χ é a representação explícita da barreira energética atribuído à rotação dos átomos i e l ao redor da ligação $j - k$ (Equação 2.6). | |
| | $V_\chi = \sum_{\chi} \left\{ \frac{V_n}{2} [1 + \cos(\chi + f_1)] + \frac{V_2}{2} [1 - \cos(\chi + f_2)] + \frac{V_3}{2} [1 + \cos(\chi + f_3)] \right\}$ (2.6) | |
| em que: V_n é a constante que define a altura da barreira de rotação (coeficientes da Série de Fourier) e f_n são as fases dos ângulos correspondentes ao número de mínimos para a torção de uma ligação química específica. | | |
| Diedral Impróprio | Está associado com deformações dos ângulos de torção impróprios, ou seja, aqueles associados a átomos de hibridização sp^2 , que geram deformações fora do plano (Equação 2.7). | |
| | $V_{\varphi_{\text{improp}}} = \frac{1}{2} k_{\varphi_{\text{improp}}} (\varphi - \varphi_0)^2$ (2.7) | |
| em que: $k_{\varphi_{\text{improp}}}$ é a constante de força do diedro impróprio, φ é o ângulo diedral impróprio e φ_0 é ângulo diedral de equilíbrio. | | |
| Não-Ligados | Lennard-Jonnes | O potencial de Lenard-Jones V_{LJ} é utilizado para descrever interações intra e intermoleculares do tipo van der Waals entre átomos não ligados. No caso de átomos em uma mesma molécula apenas são consideradas interações entre átomos separados por três ou mais ligações covalentes (Equação 2.8). |
| | | $V_{LJ} = 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right]$ (2.8) |
| em que: r_{ij} é a distância entre os átomos i e j , ε é a profundidade do poço de energia potencial e σ é o diâmetro de Lennard-Jones que é dependente da natureza dos átomos envolvidos. O primeiro termo corresponde à repulsão entre dois átomos e o segundo, à atração. | | |
| Coulomb | Coulomb | O potencial de Coulomb V_c é utilizado para descrever as interações puramente eletrostáticas (de cargas) entre dois átomos i e j (Equação 2.9). |
| | | $V_c = q_i q_j / 4\pi \varepsilon_0 r_{ij}$ (2.9) |
| em que: r_{ij} é a distância entre os átomos i e j , q_i e q_j são as cargas dos átomos e ε_0 a permissividade no vácuo. | | |

Fonte: Dados compilados de Jorgensen; Maxwell e Tirado-Rives (1996) e Jorgensen; Tirado-Rives (1988)

2.2.2 Algoritmo (Leap-Frog)

Para realizar a integração das equações de movimento dos átomos, são utilizados algoritmos que se baseiam nos métodos das diferenças finitas, ou seja, que permitem que a integração aconteça em intervalos pequenos de tempo (passo de integração) de modo a simular os movimentos mais frequentes no sistema, as vibrações dos átomos (NAMBA; DA SILVA; DA SILVA, 2008). Apesar da ampla variedade de algoritmos disponíveis, dentre os integradores capazes de efetuar a integração numérica das equações acima mencionadas, os mais confiáveis e utilizados, em função do menor consumo de memória RAM e processamento, são: *Verlet* (VERLET, 1967), *Beeman* (BEEMAN, 1976) e *Leap-frog* (HOCKNEY, 1970). O GROMACS, *software* de distribuição livre que foi utilizado na condução dos cálculos de DM, utiliza o algoritmo de *Leap-frog*. Este consiste em um método simplificado para a integração das equações de movimento que inclui explicitamente a velocidade em seus cálculos - composto por dois passos:

- 1) Cálculo das velocidades na metade do intervalo de integração (Equação 2.10)

$$v\left(t + \frac{1}{2}\delta t\right) = v\left(t - \frac{1}{2}\delta t\right) + \delta t a(t) \quad (2.10)$$

- 2) Cálculo das novas posições utilizando as velocidades calculadas (Equação 2.11).

$$r(t + \delta t) = r(t) + v\left(t + \frac{1}{2}\delta t\right) + \delta t v\left(t + \frac{1}{2}\delta t\right) \quad (2.11)$$

nas quais v é a velocidade, a aceleração e r posição dos átomos no tempo t e no intervalo do tempo de integração δt .

A limitação em relação ao uso desse algoritmo é a não-simultaneidade na determinação das posições e velocidades pelo fato de eles estarem em descompasso por $\frac{1}{2}\delta t$.

2.2.3 Ensemble NpT

Após a realização da simulação, para estudar o comportamento médio do sistema é necessário calcular a média temporal dos valores que as propriedades físicas de interesse assumem nos diversos microestados acessíveis. Para que o cálculo das propriedades

macroscópicas do sistema seja mais preciso, algumas propriedades físicas podem ser reproduzidas sendo mantidas constantes outras propriedades ao longo da simulação, dentre as quais cita-se: número de partículas (N), energia total (E), pressão (p), temperatura (T), volume (V) e potencial químico (μ). Formalmente, uma das alternativas para tanto é a combinação e reprodução do conjunto de alguns desses parâmetros, definindo uma equação de estado para o sistema, caracterizando o que chamamos de *ensemble*. Desse modo, garante que em um determinado instante um grande número de sistemas seja caracterizado pelo mesmo macroestado mesmo podendo estar em qualquer um dos microestados compatíveis, ou seja, a média temporal de uma grandeza se iguala à média do *ensemble*.

Resumidamente, o termo *ensemble* tem o seu significado associado ao vocábulo “conjunto” e refere-se, portanto, ao conjunto imaginário de réplicas da caixa de simulação com grandezas físicas idênticas (ATKINS; PAULA, 2009). Dessa forma, possibilita a ligação entre os microestados possíveis e as propriedades físicas macroscópicas do sistema por meio dos princípios fundamentais da Mecânica Estatística (SOUZA, 2009).

Nas simulações realizadas no presente trabalho, foi utilizado um *ensemble* isotérmico-isobárico (NpT) caracterizado por manter ao longo de toda a simulação a mesma composição, pressão e temperatura.

Para a manutenção do equilíbrio térmico e da pressão, é necessário o acoplamento do sistema molecular a um banho de aquecimento externo e de pressão, respectivamente (BERENDSEN et al., 1984). Uma das vantagens desse tipo de *ensemble* é o fato de as partículas não ficarem restritas a entrar ou sair do sistema de acordo com as condições periódicas de contorno, descritas na seção a seguir.

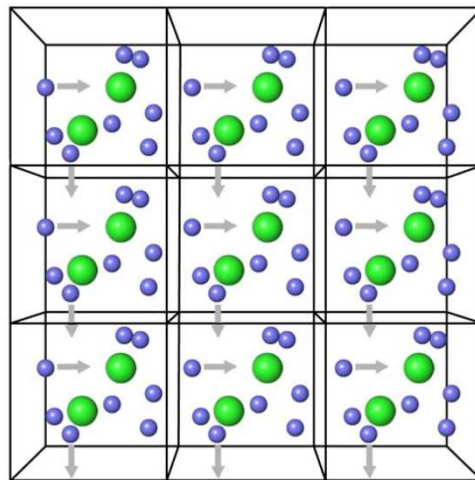
2.2.4 Condições Periódicas de Contorno

Em simulações de MD, os sistemas, compostos pela molécula de interesse (biomolécula) e as moléculas do solvente (explícito), são introduzidos em um espaço tridimensional (caixa de simulação). As formas de caixa de simulação empregadas com maior frequência são cúbicas, retangulares, octaédricas e dodecaédricas (VERLI et al., 2014). No entanto, essa forma de descrever condições experimentais ocasiona dois problemas a serem contornados, visto que afetam significativamente os resultados:

- Efeito de superfície: as moléculas da margem da caixa de simulação estão sujeitas a forças diferentes daquelas do centro (HINCHLIFFE, 2003);
- Número de átomos limitado: o quantitativo de partículas dentro da caixa é da ordem de 10^5 - 10^6 , valor pequeno quando comparado aos sistemas macroscópicos que apresentam número de átomos da ordem de Avogadro, que é da ordem de 10^{23} (VAN GUNSTEREN et al., 2006).

Para minimizar esses efeitos, aplicam-se as *condições periódicas de contorno* que consistem na realização de cópias da caixa de simulação em todas as direções, conforme representação tridimensional apresentada na Figura 2.9. Trata-se de uma abstração em que uma molécula, ao sair de um lado da caixa original, tem sua imagem entrando na face oposta da caixa imagem com mesma velocidade ao longo de toda a simulação mantendo, dessa forma, constante o número de átomos na caixa central e no sistema como um todo.

Figura 2.9 – Representação das condições periódicas de contorno em três dimensões.



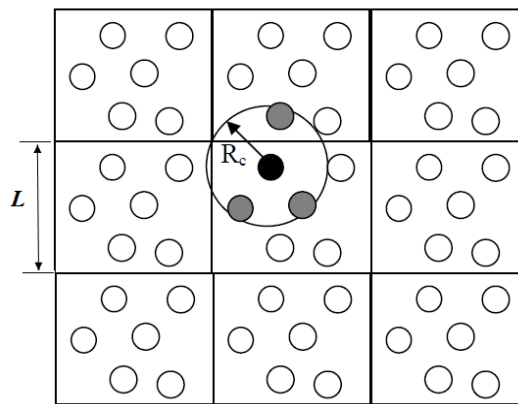
Fonte: (VERLI et al., 2014)

2.2.5 Raio de Corte

Em relação ao tamanho da caixa de simulação, é necessário que não seja definido um tamanho excessivamente pequeno, de modo a evitar que a molécula interaja com suas imagens geradas conforme definido pelas condições periódicas de contorno. Para tanto, é necessário

restringir a interação desses átomos com suas imagens através de um *cut-off*, ou seja, estabelecer um raio de corte (R_c) que deve ser da ordem de $L/2$ ($R_c \leq L/2$, na qual L é o comprimento da caixa de simulação), procedimento conhecido como convenção da imagem mínima (Figura 2.10). Outra vantagem do uso do raio de corte é a economia de memória computacional, visto que, em metodologias de DM, o cálculo das interações não-ligadas (Coulomb e Lennard-Jonnes) é a parte mais dispendiosa computacionalmente.

Figura 2.10 – Representação do Raio de Corte esférico para um sistema periódico em duas dimensões.



Fonte: Adaptado de Franca (2009).

Enquanto o potencial de Lennard-Jonnes decai rapidamente com o aumento da distância entre as partículas na proporção de r^{-6} (Equação 2.8), as interações de longo alcance (interação coulômbicas) não podem ser tratadas da mesma forma, já que o potencial V_C decai com o inverso da distância entre as partículas. Dessa forma, a eliminação da avaliação dessas interações de longa distância pelo estabelecimento de um corte pode ocasionar instabilidade e, por conseguinte, erros significativos na amostragem de simulações, em especial, aquelas que envolvam espécies carregadas (SCHLICK, 2010). Sendo assim, para descrever adequadamente essas forças de longo alcance, os métodos mais utilizados são: a soma de Ewald, sua variação PME (do inglês, *Particle Mesh Ewald*) e a aproximação do campo de reação (*reaction-field*).

O método soma de Ewald é mais utilizado na descrição de sistemas altamente carregados ou polares e, neste método, uma dada partícula interage com todas as demais partículas da caixa de simulação e, também, com todas as suas imagens num arranjo infinito de células periódicas, de modo que, no limite, esse adquire formato esférico (EWALD, 1921). Já a aproximação do campo de reação (*reaction field*) é utilizada para tratar as interações

Coulômbicas em sistemas em fase condensada como, por exemplo, macromoléculas em solução, visto que considera o ambiente fora do raio de corte como um contínuo dielétrico com constante dielétrica ϵ referente ao solvente utilizado e permite que esse ambiente apresente também átomos e moléculas de soluto (LINDAHL; HESS; VAN DER SPOEL, 2001).

2.2.6 Modelos de moléculas de água para solvatação explícita

O solvente de um sistema, em particular a água, exerce influência considerável sobre a estrutura do soluto e a termodinâmica do sistema. As enzimas, por exemplo, requerem um elevado grau de hidratação para que possam exercer sua função biológica. Por essa razão, é importante, em simulações computacionais de sistemas biomoleculares solvatados, recriar o ambiente aquoso da forma mais precisa possível, de modo a incluir os efeitos do solvente, seja pelos métodos implícitos ou explícitos (MAKAROV; PETTITT; FEIG, 2002; NUTT; SMITH, 2007).

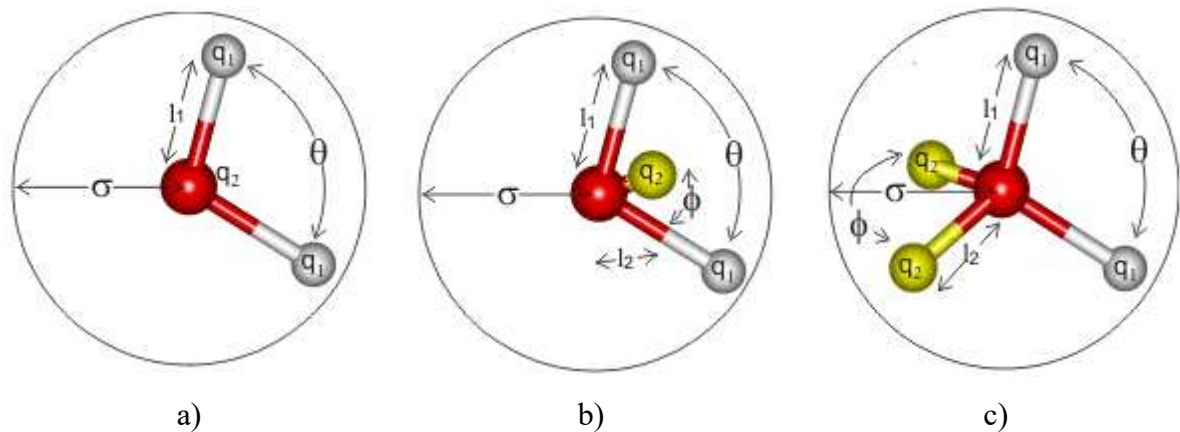
O modelo de solvatação implícita não considera os graus de liberdade do solvente, mas o trata como um meio contínuo que contém as propriedades médias do solvente real. Nesse tratamento, a influência do solvente é considerada através da estimativa direta da energia livre de solvatação, que se trata de uma medida da afinidade do soluto com o solvente. Conceitualmente, essa propriedade pode ser definida com o trabalho reversível requerido para a transferência do soluto em uma configuração fixa do vácuo para a solução (NAMBA; DA SILVA; DA SILVA, 2008).

Na solvatação explícita, o solvente é tratado como parte do sistema, ou seja, como um elemento adicional. Dessa forma, é necessária a parametrização do potencial para a reprodução das propriedades macroscópicas e, portanto, um modelo que descreva apropriadamente a molécula de água.

Dentre os vários modelos de água disponíveis mais utilizados na solvatação de sistemas simulados por dinâmica molecular, tem-se o SPC (do inglês, *Simple Point Charge*) e o conjunto dos potenciais TIP (do inglês, *Transferable Intermolecular Potentials*). Assim, no modelo SPC, as moléculas de água são tratadas como cargas pontuais com três sítios de interação para as interações eletrostáticas e as cargas parciais dos átomos de hidrogênio balanceadas apropriadamente pela carga negativa do átomo de oxigênio (BERENDSEN; GRIGERA; STRAATSMA, 1987). Os modelos rígidos do tipo TIP n P são simples e se baseiam na

interações não-ligadas, já que as interações ligadas são tratadas implicitamente. As moléculas de água do tipo TIP3P apresentam três sítios de interação de modo análogo ao SPC, enquanto as do tipo TIP4P e TIP5P utilizam, respectivamente, 4 e 5 sítios de interação, conforme mostrado na Figura 2.11. A comparação descritiva entre as características desses modelos pode ser observada na Tabela 2.2.

Figura 2.11 - Representação dos modelos de água do tipo: SPC e TIP3P (a), TIP4P (b) e TIP5P (c); I_1 e I_2 representam as distâncias entre o átomo hidrogênio e oxigênio; q_1 e q_2 representam as cargas atômicas do hidrogênio e oxigênio, respectivamente; θ e Φ são os ângulos de ligação e σ é a distância para o potencial zero.



Fonte: (CHAPLIN, 2000)

Tabela 2.2 – Comparação dos parâmetros de diferentes modelos de molécula de água.

| Parâmetros | SPC | TIP3P | TIP4P | TIP5P |
|-------------------------|----------|----------|----------|----------|
| $I_1 / \text{Å}$ | 1,0000 | 0,9572 | 0,9572 | 0,9572 |
| $I_2 / \text{Å}$ | --- | --- | 0,15 | 0,70 |
| q_1 / e | + 0,4238 | + 0,4170 | + 0,5200 | + 0,2410 |
| q_2 / e | - 0,8476 | - 0,8340 | - 1,0400 | - 0,2410 |
| $\sigma / \text{Å}$ | 3,16600 | 3,15061 | 3,15365 | 3,12000 |
| θ / graus | 109,47 | 104,52 | 104,52 | 104,52 |
| Φ / graus | --- | --- | 52,26 | 109,47 |
| μ / D | 2,35 | 2,35 | 2,18 | 2,29 |

Fonte: Dados compilados de (BERENDSEN; GRIGERA; STRAATSMA, 1987; JORGENSEN et al., 1983; MARK; NILSSON, 2001; NERIA; FISCHER; KARPLUS, 1996).

2.3 METODOLOGIA

2.3.1 Dinâmica Molecular

Os sistemas estudados nessa etapa inicial do trabalho foram as proteínas NS1 do flavírus da Dengue Sorotipo 2 (DENV2) e da Zikavirus (ZIKV) e o anticorpo monoclonal 22-anti-NS1. As estruturas tridimensionais da NS1-DENV2, NS1-ZIKV e do 22-anti-NS1 foram retiradas do banco de dados de proteínas RCSC-PDB registradas, respectivamente, sob os códigos 4O6B (AKEY et al., 2014b), 5K6K (BROWN et al., 2016) e 4OII (EDELING; DIAMOND; FREMONT, 2014b). As regiões de loop ausentes na estrutura 4O6B foram modeladas conforme protocolo apresentado por Freire et al. (2017) e os átomos faltantes na estrutura 5K6K (átomos CG, CD1, CD2, CE1, CE2, CZ e OH do resíduo TYR1) foram modelados utilizando o *software Modeller* implementado na interface gráfica UCSF *Chimera* (YANG et al., 2012). A estrutura final da modelagem da NS1-ZIKV foi escolhida com base nos parâmetros GA341 (MELO; SÁNCHEZ; SALI, 2002), zDOPE (SHEN; SALI, 2006), RMSD e *Overlap* estimados (ERAMIAN et al., 2008). As simulações foram conduzidas para as estruturas diméricas isentas de glicosilações com a finalidade de avaliar a influência da ausência dos resíduos de N-acetilglicosamina na manutenção estrutural das proteínas. Os dímeros foram escolhidos para esse estudo considerando a boa qualidade dessas estruturas disponíveis nas bases de dados, a existência de estudos comparativos e o fato de que a forma hexamérica seria computacionalmente muito dispendiosa, tanto para os cálculos de DM quanto para as etapas seguintes de Docking Molecular e cálculos semi-empíricos. Os sistemas simulados estão descritos na Tabela 2.3 apresentada a seguir.

Tabela 2.3 – Descrição dos sistemas simulados.

| Sistema | Cadeias da Proteína | Nº átomos Proteína | Nº total de átomos |
|-------------|---------------------|--------------------|--------------------|
| NS1-DENV2 | A e B | 10996 | 469742 |
| NS1-ZIKV | A e B | 11181 | 469967 |
| 22-anti-NS1 | L e H | 6477 | 163517 |

Fonte: Autor

As simulações de dinâmica molecular de cada uma das proteínas descritas na Tabela 2.3 foram, separadamente, conduzidas no pacote computacional GROMACS 5.1.2 (ABRAHAM et al., 2015) utilizando o campo de força OPLS-AA (JORGENSEN; TIRADORIVES, 1988).

Inicialmente, as proteínas foram inseridas a 1,5 nm de distância das arestas da caixa de simulação cúbica, em seguida, preenchida com moléculas de água do tipo TIP4P. A força iônica dos sistemas foi ajustada para 150 mM a fim de mimetizar as condições do soro sanguíneo (WANG et al., 2017a). Em seguida, a energia foi minimizada com o algoritmo *steepest descent* (QIAN, 1999) utilizando-se 1000 passos de integração numérica. Este algoritmo utiliza a derivada primeira para determinar a direção do mínimo de energia do *ensemble* simulado. Os sistemas foram termalizados por 100 ps, com aumento gradual e sucessivo da temperatura, a 50 K, 150 K, e finalmente a 310 K, sendo mantida constante a esta temperatura, com o acoplamento soluto-solvente realizado a cada 0,1 ps. Nesta etapa, as coordenadas dos átomos mais pesados, com exceção do hidrogênio, foram fixadas com uma força constante de $1,0 \cdot 10^3 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ (FREIRE et al., 2017).

As simulações por dinâmica molecular foram realizadas utilizando o intervalo de tempo de integração (Δt) de 2 fs, totalizando 100 ns em um *ensemble* isotérmico-isobárico (NPT), utilizando o algoritmo *leap-frog* (HOCKNEY, 1970) para integração das equações de movimento de Newton. Além disso, o barostato de Parrinello-Rahman (PARRINELLO; RAHMAN, 1981) e o termostato *v-rescale* (BUSSI; DONADIO; PARRINELLO, 2007) foram utilizados para manter constantes a pressão e a temperatura, respectivamente, em 1,0 bar (tempo de relaxação de 2,0 ps e constante de compressibilidade de $4,5 \times 10^{-5}$ bar) e 310 K.

O movimento de estiramento linear e angular dos átomos de hidrogênio e o comprimento das ligações foram restringidos usando-se o algoritmo LINCS (do inglês, *Linear Constraint Solver*) (HESS et al., 1997). Visando à caracterização das interações eletrostáticas e de van der Waals de curto alcance, foi utilizado um raio de corte de 1,5 nm e as contribuições eletrostáticas de longo alcance foram tratadas pelo método PME.

As trajetórias obtidas foram analisadas utilizando o *software* VMD (do inglês, *Visual Molecular Dynamics*) (HUMPHREY; DALKE; SCHULTEN, 1996) e os gráficos e análises estatísticas realizadas nos *softwares* GRACE (do inglês, *Graphing Advanced Computation and Exploration of Data*) (STAMBULCHIK, 2008) e Origin 8.0. Os valores médios e seus

respectivos desvios-padrão para as propriedades calculadas apresentaram nível de confiança de 95 %.

2.3.2 Análise da população (cluster) das simulações por Dinâmica Molecular

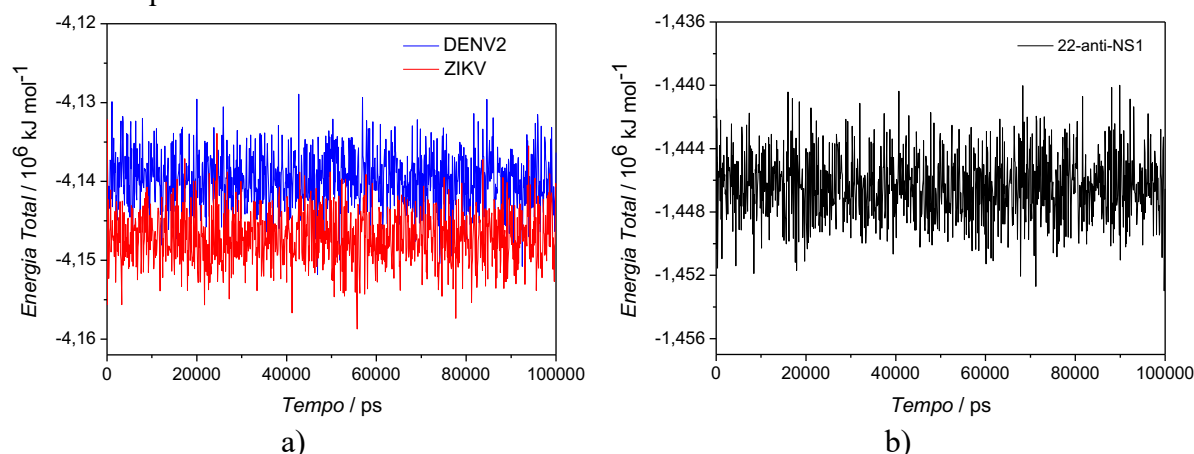
Os últimos 20 ns (convergência estrutural e energética) das dinâmicas moleculares de todos os sistemas foram clusterizados para obtenção de uma estrutura representativa da proteína em solução, ou seja, considerando o efeito do solvente. Para tanto, foi utilizado o método *single linkage* do módulo *g_cluster* disponível no programa GROMACS 5.1.2 (ABRAHAM et al., 2015).

2.4 RESULTADOS E DISCUSSÕES

As simulações por DM de biomoléculas, como as proteínas dos sistemas em estudo, são realizadas para observar seu comportamento temporal em meio aquoso, ou seja, para estudar a proteína buscando mimetizar seu ambiente nativo. Para tanto, é importante que exista a manutenção das suas estruturas secundária e terciária, ou seja, que não ocorra desnaturação. Dessa forma, foi necessária a avaliação das suas respectivas estabilidades energética e estrutural.

Considerando o parâmetro energético, a Figura 2.12 apresenta a evolução temporal da energia total para os sistemas: NS1-DENV2 (azul), NS1-ZIKV (vermelho) e 22-anti-NS1 (preto), sendo possível observar um patamar estável dos valores de energia para todos os sistemas. Os níveis energéticos encontram-se estabilizados e com pequenas flutuações (baixos valores de desvio padrão), respectivamente, em $-4,137 \times 10^6 (\pm 0,003)$, $-4,147 \times 10^6 (\pm 0,003)$ e $-1,44 \times 10^6 (\pm 0,002)$ kJ mol⁻¹. Portanto, o equilíbrio energético pode ser sugerido para os sistemas em estudo sendo possível observar alta similaridade nos parâmetros energéticos para os sistemas das proteínas NS1 de ambas flavivíroses.

Figura 2.12 – Energia Total dos sistemas das proteínas NS1 (a) e do anticorpo (b) simulados em meio aquoso.



Fonte: Autor

Verificada essa propriedade para os sistemas, torna-se viável analisar seus parâmetros estruturais ao longo das simulações. Portanto, foi verificada a evolução temporal da raiz quadrada do desvio quadrático médio (do inglês, *root mean square deviation*, RMSD) e do Raio de Giro (R_g) que avaliam, respectivamente, o distanciamento das posições atômicas da proteína

em relação à estrutura nativa e uma aproximação da medida da compactação de sua estrutura. Essas propriedades são calculadas pelas equações que se seguem e os gráficos para os sistemas em estudo estão apresentados na Figura 2.13:

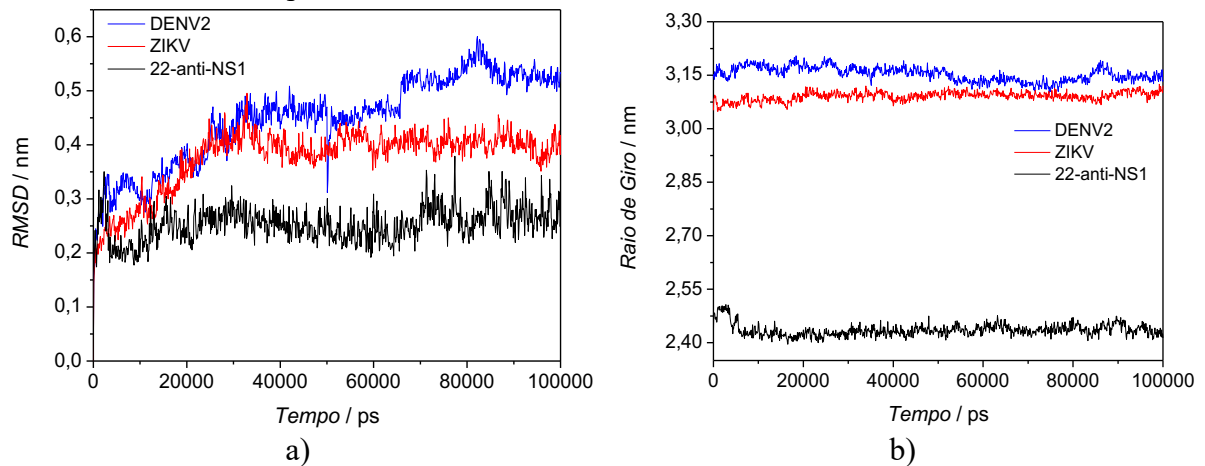
$$RMSD(t_1, t_2) = \left[\frac{1}{M} \sum_{i=1}^N m_i |r_i(t_1) - r_i(t_2)|^2 \right]^{\frac{1}{2}} \quad (2.12)$$

em que, $M = \sum_{i=1}^N m_i$, corresponde à somatória da massa de todos os N átomos da molécula e $r_i(t)$ é a posição do átomo i no tempo t .

$$R_g = \left[\frac{\sum_{i=1}^N |r_i|^2 m_i^{1/2}}{\sum_{i=1}^N m_i} \right] \quad (2.13)$$

em que, m_i é a massa do átomo i e r_i é sua posição em relação ao centro de massa.

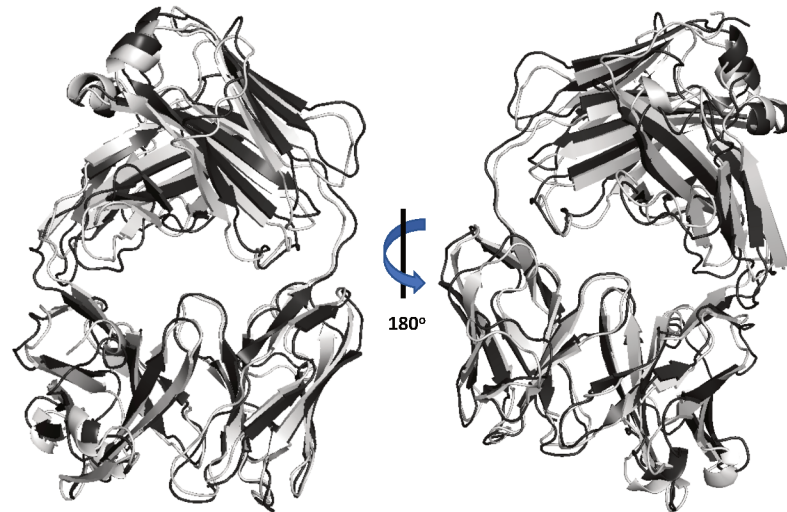
Figura 2.13 – RMSD (a) e Raio de Giro (b) dos sistemas das proteínas NS1 e do anticorpo simulados em meio aquoso.



Fonte: Autor

Para o anticorpo, os valores médios de RMSD e Raio de Giro (R_g) foram de, respectivamente, $0,253 (\pm 0,032)$ e $2,44 (\pm 0,02)$ nm associados à pequena mobilidade estrutural em solução aquosa salina, ou seja, uma leve e sistemática flutuação conforme observado na Figura 2.14.

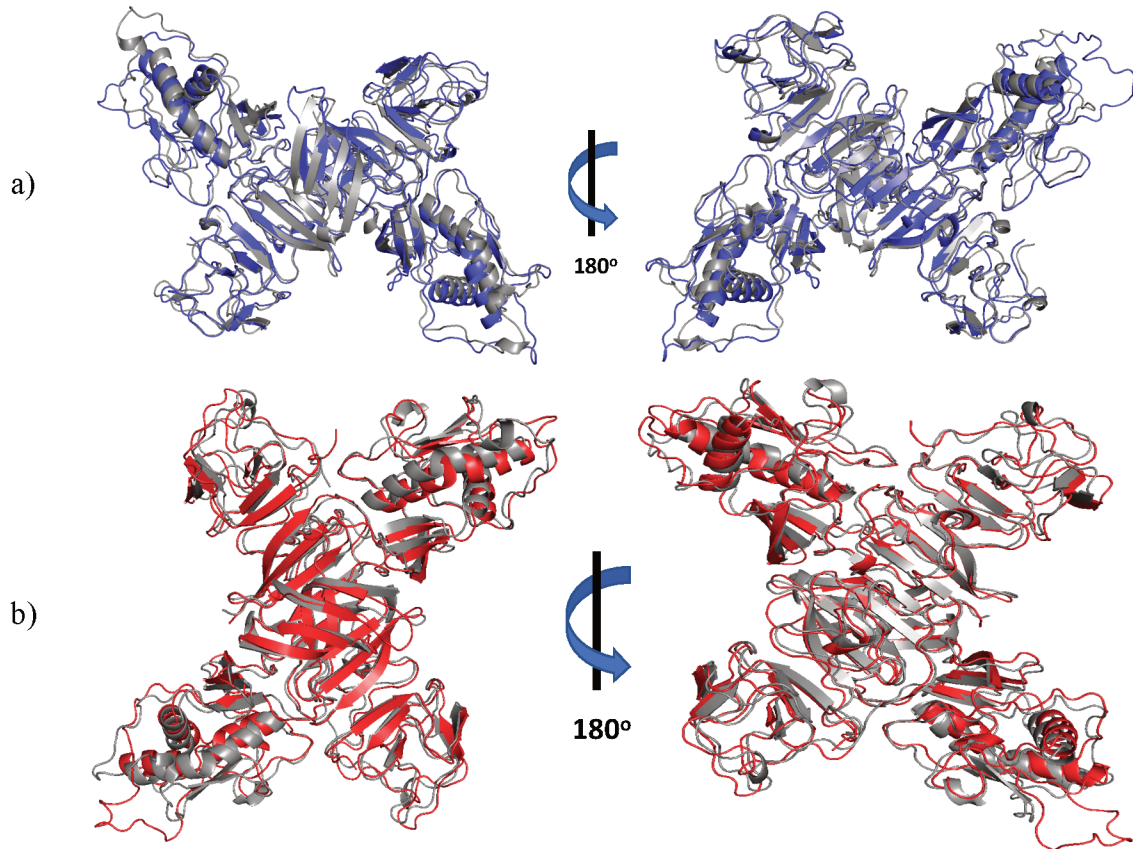
Figura 2.14 – Representação *new cartoon* da sobreposição entre a estrutura do anticorpo no início (cinza) e ao final da simulação (preto).



Fonte: Autor

Para as proteínas NS1-DENV2 e NS1-ZIKV, os valores médios de RMSD e Raio de Giro (R_g) foram, respectivamente, 0,449 ($\pm 0,082$) e 3,15 ($\pm 0,02$) nm; 0,376 ($\pm 0,057$) e 3,09 ($\pm 0,01$) nm. Esses valores e os gráficos indicam maior flutuação da estrutura da NS1-DENV2 quando comparada com a NS1-ZIKV, em particular, nos últimos nanossegundos de simulação associados aos movimentos mais acentuados no domínio β -roll. Isso pode ser observado na Figura 2.15 que apresenta a sobreposição das estruturas das proteínas no início da simulação e a estrutura representativa das proteínas em meio aquoso.

Figura 2.15 – Representação *new cartoon* da sobreposição entre a estrutura das proteínas NS1-DENV2 (a) e NS1-ZIKV (b) no início (cinza) e ao final da simulação (azul/vermelho).



Fonte: Autor

No entanto, dentro dos intervalos de desvio-padrão dos valores médios, foi observada similaridade nas propriedades estruturais das proteínas NS1 da Dengue e da Zika e, portanto, ausência de diferenças significativas no comportamento e na dinâmica geral dos dímeros. Esse fato está em concordância com o apresentado por Brown et al. (2016) e Song et al. (2016), segundo os quais as dobras e o arranjo dos domínios são praticamente idênticos entre as proteínas.

Vale considerar ainda que os sistemas simulados correspondem à forma não glicosilada das proteínas NS1. Isso porque a forma secretada sNS1 e, portanto, detectável em um dispositivo de sensoriamento, está isenta de glicosilações. Comparativamente aos estudos teóricos de Freire et al. (2017) e Gonçalves (2017), nos quais essas proteínas foram simuladas nas suas formas glicosiladas, os parâmetros estruturais são semelhantes indicando que a ausência dos resíduos de açúcar parece não afetar na manutenção estrutural das proteínas.

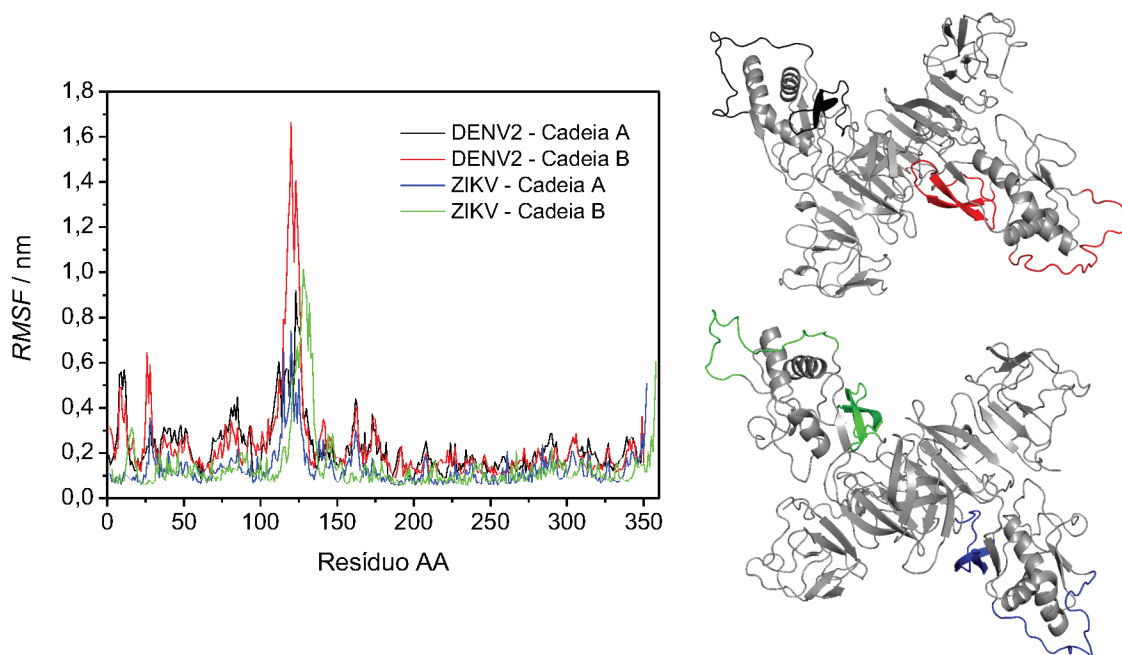
Portanto, as simulações em questão não apresentam eventos moleculares significativos que indiquem a perda da estrutura ou desnaturação das proteínas em solução aquosa salina.

Para avaliar com maior rigor a flexibilidade dos sistemas proteicos, foram obtidos os valores de raiz quadrada da flutuação média (do inglês, *root mean square fluctuation*, RMSF). O RMSF é calculado como a média temporal para cada um dos i átomos de aminoácido da proteína sem levar em consideração seu peso molecular, conforme equação a seguir:

$$RMSF_i = (|r_i - r_{i\text{médio}}|)^2 \quad (2.14)$$

em que $r_i(t)$ é a posição do átomo i .

Figura 2.16 – RMSF para os resíduos de aminoácido das proteínas NS1-DENV2 e NS1-ZIKV e indicação das regiões de maior flutuação.



Fonte: Autor

O gráfico da Figura 2.16 indica amplitudes de RMSF maiores para a quase totalidade dos resíduos de aminoácido do dímero da DENV2 em relação aos da ZIKV. No entanto, há conformidade nas faixas de resíduos de maior flutuação, corroborando com o observado nas análises de RMSD e Raio de Giro. Essas regiões de maior flutuação estão relacionadas a regiões de importância funcional para a proteína.

Na Figura 2.16 estão identificadas na estrutura secundária das proteínas duas regiões importantes: (1) a.a. 108-130: região pertencente ao domínio wing e composta por loops não

estruturados, caracterizados por sua maior flexibilidade. A flutuação considerável, da ordem de até 1,7 nm, tem relação com a ausência dos resíduos de açúcar estruturalmente ligados à Asp130. Portanto, a ausência de glicosilação, apesar de não afetar a manutenção estrutural, influencia nas flutuações da proteína, em especial, nas regiões próximas; e (2) a.a. 152-180: região conectora entre os domínios wing e β -ladder, portanto, composta por alças flexíveis que ligam duas folhas β importantes no reconhecimento imunogênico da proteína.

Uma terceira região de flutuação considerável pode ser apontada: a.a. 1-25 (DENV) e 10-30 (ZIKV): região que compreende o domínio β -roll e, portanto, as porções N-terminal. Os valores maiores para a DENV2 devem-se em especial a uma maior desorganização das folhas- β ao final da simulação visualmente anotado (Figura 2.15). No entanto, vale considerar que para a finalidade do presente trabalho a função dessa flutuação é secundária, visto que, na forma sNS1 esse domínio encontra-se na parte interna hidrofóbica da proteína não acessível para detecção.

Portanto, mesmo que os dímeros de NS1 e o anticorpo apresentem regiões instáveis e flexíveis, evidencia-se a manutenção estrutural ao final da simulação e a importância biológica dessas regiões nas prováveis interações antígeno-anticorpo.

A estrutura terciária das proteínas, ou seja, o arranjo tridimensional das estruturas secundárias, é estabilizada majoritariamente por ligações de caráter não-covalente como, por exemplo, as ligações de hidrogênio, as pontes salinas e interações hidrofóbicas (KESSEL; BEN-TAL, 2010). Nesse sentido, o número de ligações de hidrogênio e de pontes salinas foi adotado como parâmetro para avaliação dos efeitos de hidratação, manutenção da estrutura tridimensional e estabilidade estrutural conforme proposto por Franca et al. (2011). Os critérios adotados para as análises de ligações de hidrogênio entre as proteínas e as moléculas de água foram 3,5 Å para a distância em relação ao raio de corte e 30° para o ângulo entre os centros de referência escolhido entre os átomos doador e aceptor (TORSHIN; WEBER; HARRISON, 2002). Nas análises de interação eletrostática entre resíduos de aminoácidos de cargas opostas (pontes salinas), foi adotada uma distância de 3,2 Å entre os átomos de oxigênio de resíduos ácidos e átomos de nitrogênio de resíduos básicos (TRABUCO; VILLA, 2011). A Tabela 2.4 sumariza o número de ligações de hidrogênio e pontes salinas monitorado ao longo das simulações por DM, em intervalos de 20 ns.

Tabela 2.4 – Número de ligações de hidrogênio entre moléculas de água e a proteína, e pontes salinas quantificadas durante as simulações por DM.

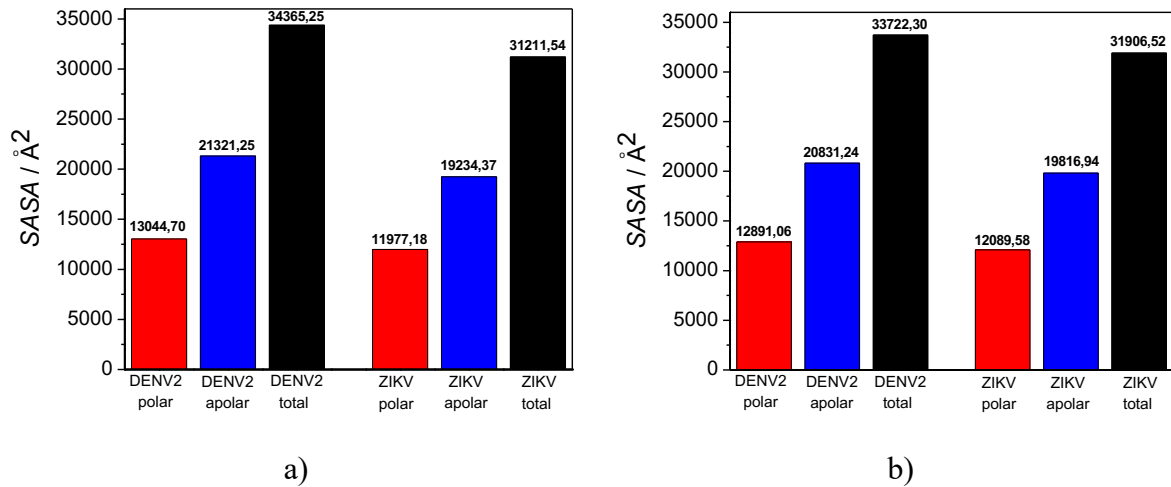
| Sistema | Interação | 0 ns | 20 ns | 40 ns | 60 ns | 80 ns | 100 ns |
|-------------|-----------------------|------|-------|-------|-------|-------|--------|
| NS1-DENV2 | Ligação de Hidrogênio | 1511 | 1504 | 1479 | 1457 | 1473 | 1477 |
| | Ponte Salina | 49 | 54 | 52 | 49 | 49 | 53 |
| NS1-ZIKV | Ligação de Hidrogênio | 1468 | 1522 | 1497 | 1515 | 1438 | 1484 |
| | Ponte Salina | 63 | 66 | 71 | 69 | 72 | 75 |
| 22-anti-NS1 | Ligação de Hidrogênio | 859 | 860 | 874 | 888 | 865 | 896 |
| | Ponte Salina | 35 | 32 | 31 | 30 | 26 | 21 |

Fonte: Autor

Para o anticorpo, observou-se uma tendência inversa entre o número de ligações de hidrogênio e de pontes salinas. Esse resultado indica que a diminuição no número de pontes salinas ao longo da simulação está associada às interações estabelecidas entre os resíduos de aminoácidos carregados (ARG, ASP, GLU e LYS) e as moléculas de água. Portanto, a solvatação do sistema leva à formação de novas ligações de hidrogênio e, conseqüentemente, à manutenção da estabilidade estrutural do sistema. Essa tendência está em consonância com os resultados obtidos utilizando esse protocolo ao avaliar a estabilidade de proteínas como, por exemplo, acetil coenzima A carboxilase (ACC) (FRANCA et al., 2011), acetohidroxiácido sintase (AHAS) (IERICH et al., 2015b) e complexo da glicoproteína MOG e a região Fab do anticorpo (MOG e anti-MOG Fab) (IERICH et al., 2019b).

No entanto, essa mesma tendência não é observada para as proteínas NS1 de ambas as flavivírus. Para a NS1-DENV2, a tendência foi inversa: com o aumento do número de pontes salinas, houve a diminuição do número de ligações de hidrogênio; para a NS1-ZIKV: os dois parâmetros aumentaram ao longo dos 100 ns de simulação. Nesses casos, a análise deve ser relativizada em relação à contribuição das porções polar e apolar de cada uma das proteínas. Para tanto, a Área Acessível ao Solvente (SASA, do inglês, *Solvent Accessible Surface Area*) foi calculada utilizando o *software online* Getarea (FRACZKIEWICZ; BRAUN, 1998) com raio de 1,4 Å e os resultados estão apresentados na Figura 2.17.

Figura 2.17 – Área acessível ao solvente (SASA) e suas contribuições polar e apolar para as proteínas NS1 de ambas as flavivíroses no início a 0 ns (a) e final a 100 ns (b) de simulação.



Fonte: Autor

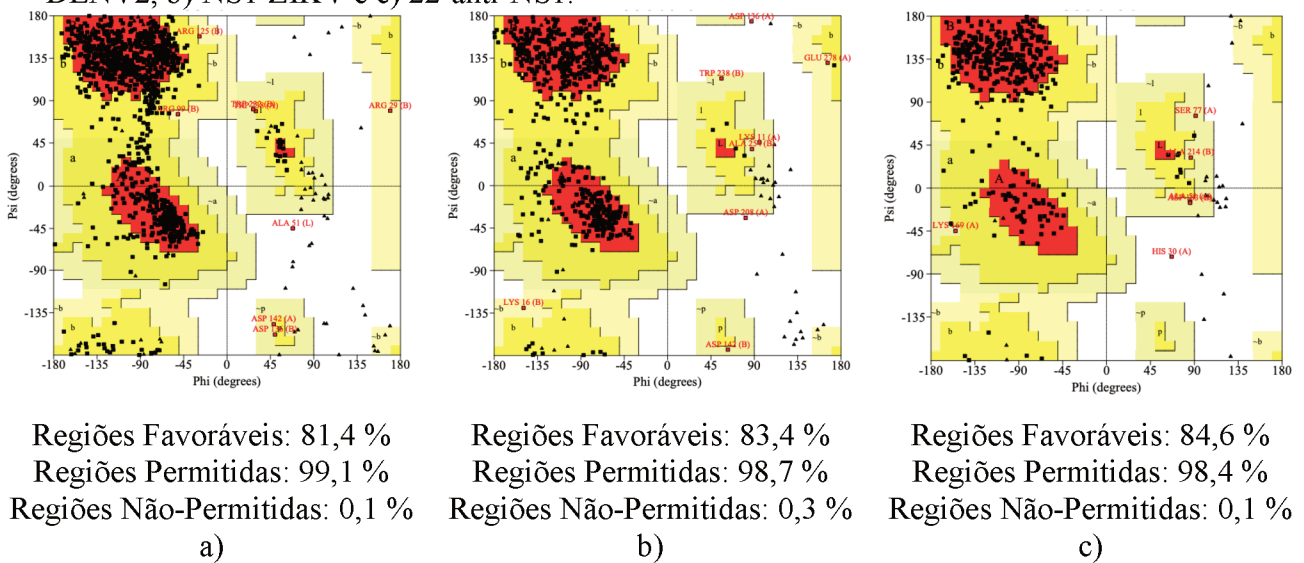
De modo geral, houve diminuição e aumento na área acessível ao solvente, respectivamente, para NS1-DENV2 e NS1-ZIKV. No entanto, é preciso levar em consideração que a contribuição da região polar dessas proteínas é mais significativa em implicações de resposta imune, visto que se trata da parte externa (hidrofílica). Isso porque a parte apolar é responsável pela estabilização do hexâmero por meio de ligações hidrofóbicas entre os dímeros, que ficam inacessíveis nessa forma detectável da proteína.

Para tanto, é preciso compreender ainda que, segundo Freire e colaboradores (2017), a região polar acessível ao solvente, essencialmente a região de loop (domínio wing) apresenta diferença significativa em termos de superfície de energia potencial. Sendo a NS1-DENV2 carregada positivamente e a NS1-ZIKV, negativamente.

Dessa forma, a $SASA_{NS1-DENV2}$ da região polar diminuiu ($\Delta_{SASA} = -153,64 \text{ \AA}^2$), porque apesar da carga positiva dessa região da proteína, cerca de 43 % dos resíduos de aminoácido carga positiva (Arg, His e Lys) estão na região interna da proteína (*buried residues*), ou seja, inacessíveis ao solvente. Portanto, o efeito da solvatação não foi efetivo, ocasionando a diminuição na formação de ligações de hidrogênio intermoleculares entre os resíduos de aminoácido e as moléculas de água e, conseqüentemente, possibilitando que os resíduos carregados pudessem estabelecer interações do tipo ponte salina cujo número aumentou. Vale considerar ainda que, mesmo com a diminuição da SASA após a DM para esse dímero, em termos absolutos, essa é superior à da ZIKV, sendo sua estabilidade adquirida pelo aumento da entropia do solvente.

A qualidade estereoquímica desses modelos foi avaliada pela distribuição dos ângulos torcionais ϕ e Ψ da cadeia principal e apresentados na forma dos diagramas de Ramachandran obtidos utilizando o *webservice* PROCHECK (LASKOWSKI et al., 1993; MORRIS et al., 1992) e apresentados na Figura 2.19. As áreas interiores às marcações em vermelho (*A*, *B* e *L*), amarelo claro (*a*, *b*, *l* e *p*) e branco indicam, respectivamente, regiões favoráveis, permitidas e não permitidas de distribuição dos ângulos torcionais ϕ e Ψ .

Figura 2.19 – Diagramas de Ramachandran das estruturas representativas das proteínas a) NS1-DENV2, b) NS1-ZIKV e c) 22-anti-NS1.



Fonte: Autor

A maior parcela dos resíduos se encontra em regiões favoráveis ou permitidas sugerindo que os aminoácidos das proteínas tendem a ocupar regiões bem definidas e que, portanto, os comprimentos e ângulos das ligações e dos diedros dos modelos são razoáveis. Sendo assim, a análise dos diagramas indica boa qualidade estereoquímica para os 3 modelos obtidos. Portanto, essas estruturas serão utilizadas como arquivos de entrada (*input*) nos cálculos de docking molecular.

2.5 CONCLUSÕES PARCIAIS

Os parâmetros estruturais e energéticos, obtidos a partir das simulações por dinâmica molecular das proteínas NS1 das flavirozes DENV2 e ZIKV e do anticorpo anti-NS1, evidenciaram a manutenção estrutural dessas proteínas. No entanto, as regiões a.a. 108-130 e a.a. 152-180, além de importância biológica nas prováveis interações antígeno-anticorpo, apresentaram maior flutuação.

As ligações de hidrogênio e pontes salinas possuem papel importante na avaliação dessa estabilidade estrutural, em especial, em relação ao efeito do solvente que é mais significativo para a NS1-ZIKV. Por outro lado, a ausência de glicosilações, apesar de contribuir para maior flutuação no domínio wing nas estruturas diméricas simuladas, não afeta na manutenção estrutural delas apontando para a representatividade dessas em relação aos sistemas biológicos reais, que envolvem a detecção da forma hexamérica.

Estruturas de boa qualidade estequiométrica foram estatisticamente amostradas, a partir das simulações por DM, para cada uma das biomoléculas em estudo e serão utilizadas como arquivos de entrada no protocolo de Docking Molecular.

REFERÊNCIAS

ABBAS, A. K.; LICHTMAN, A. H.; PILLAI, S. Anticorpos e Antígenos. In: **Imunologia Celular e Molecular**. 7. ed. Rio de Janeiro: Elsevier, 2011. p. 89–108.

ABRAHAM, M. J. et al. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. **SoftwareX**, v. 1, n. 2, p. 19–25, 2015. <https://doi.org/10.1016/j.softx.2015.06.001>

AKEY, D. L. et al. Flavivirus NS1 Structures Reveal Surfaces for Associations with Membranes and the Immune System. **Science**, v. 343, n. 6173, p. 881–885, 2014. <https://doi.org/10.1126/science.1247749>

ALLEN, M. P.; TILDESLEY, D. J. **Computer Simulation in Chemical Physics**. Dordrecht: Springer Netherlands, 1993. <https://doi.org/10.1007/978-94-011-1679-4>

ARTYMIUK, P. J. et al. Crystallographic studies of the dynamic properties of lysozyme. **Nature**, v. 280, n. 5723, p. 563–568, 1979. <https://doi.org/10.1038/280563a0>

ATKINS, P.; PAULA, J. DE. **Físico-Química - Volume 2**. 9. ed. Rio de Janeiro: LTC, 2009.

- BEEMAN, D. Some multistep methods for use in molecular dynamics calculations. **Journal of Computational Physics**, v. 20, n. 2, p. 130–139, 1976. [https://doi.org/10.1016/0021-9991\(76\)90059-0](https://doi.org/10.1016/0021-9991(76)90059-0)
- BERENDSEN, H. J. C. et al. Molecular dynamics with coupling to an external bath. **The Journal of Chemical Physics**, v. 81, n. 8, p. 3684–3690, 1984. <https://doi.org/10.1063/1.448118>
- BERENDSEN, H. J. C.; GRIGERA, J. R.; STRAATSMA, T. P. The Missing Term in Effective Pair Potentials. **Journal of Physical Chemistry**, v. 91, n. 24, p. 6269–6271, 1987. <https://doi.org/10.1021/j100308a038>
- BRANDEN, C.; TOOZE, J. **Introduction to Protein Structure**. 2. ed. New York: Garland Science, Taylor & Francis Group, 2009.
- BROWN, W. C. et al. Extended surface for membrane association in Zika virus NS1 structure. **Nature Structural and Molecular Biology**, v. 23, n. 9, p. 865–867, 2016. <https://doi.org/10.1038/nsmb.3268>
- BUSSI, G.; DONADIO, D.; PARRINELLO, M. Canonical sampling through velocity rescaling. **Journal of Chemical Physics**, v. 126, n. 014101, p. 1–7, 2007. <https://doi.org/10.1063/1.2408420>
- CHUNG, K. M. et al. Antibodies against West Nile Virus Nonstructural Protein NS1 Prevent Lethal Infection through Fc- γ Receptor-Dependent and -Independent Mechanisms. **Journal of Virology**, v. 80, n. 3, p. 1340–1351, 2006.
- COICO, R.; SUNSHINE, G. **Imunologia**. 6. ed. Rio de Janeiro: Koogan, Guanabara, 2010.
- CORNELL, B. **Amino Acids**. Disponível em: <<http://ib.bioninja.com.au>>. Acesso em: 7 jun. 2020.
- CUNHA, R. A. **Caracterização molecular de mecanismos de interação de quitosana com bicamadas lipídicas compostas de dipalmitoilfosfatidilcolina**. Dissertação (Mestrado em Química) - Instituto de Química, Universidade Federal de Uberlândia. Uberlândia, 87 p., 2013
- EDELING, M. A.; DIAMOND, M. S.; FREMONT, D. H. Structural basis of flavivirus NS1 assembly and antibody recognition. **Proceedings of the National Academy of Sciences of the United States of America**, v. 111, n. 11, p. 4285–4290, 2014. <https://doi.org/10.1073/pnas.1322036111>
- ERAMIAN, D. et al. How well can the accuracy of comparative protein structure models be predicted? **Protein Science**, v. 17, n. 11, p. 1881–1893, 2008. <https://doi.org/10.1110/ps.036061.108>
- FELLOUSE, F. A. et al. Molecular Recognition by a Binary Code. **Journal of Molecular Biology**, v. 348, n. 5, p. 1153–1162, 2005. <https://doi.org/10.1016/j.jmb.2005.03.041>
- FELLOUSE, F. A. et al. Tyrosine Plays a Dominant Functional Role in the Paratope of a Synthetic Antibody Derived from a Four Amino Acid Code. **Journal of Molecular Biology**, v. 357, n. 1, p. 100–114, 2006. <https://doi.org/10.1016/j.jmb.2005.11.092>
- FRACZKIEWICZ, R.; BRAUN, W. Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. **Journal of Computational**

Chemistry, v. 19, n. 3, p. 319–333, 1998. [https://doi.org/10.1002/\(SICI\)1096-987X\(199802\)19:3<319::AID-JCC6>3.0.CO;2-W](https://doi.org/10.1002/(SICI)1096-987X(199802)19:3<319::AID-JCC6>3.0.CO;2-W)

FRANCA, E. F. et al. Designing an enzyme-based nanobiosensor using molecular modeling techniques. **Physical chemistry chemical physics : PCCP**, v. 13, n. 19, p. 8894–8899, 2011. <https://doi.org/10.1039/c1cp20393b>

FREIRE, M. C. L. C. et al. Mapping Putative B-Cell Zika Virus NS1 Epitopes Provides Molecular Basis for Anti-NS1 Antibody Discrimination between Zika and Dengue Viruses. **ACS Omega**, v. 2, p. 3913–3920, 2017. <https://doi.org/10.1021/acsomega.7b00608>

GONÇALVES, R. L. **Mecanismos da interação entre monômeros da NS1 dos vírus Zika e Dengue como alvo do design racional de fármaco**. Dissertação (Mestrado em Ciências Aplicadas), Programa de Pós-Graduação em Ciências Aplicadas, Universidade Federal de Goiás. Jataí, 65 p., 2017.

GUIZADO, T. R. C. **Estudos Computacionais da Interação de Porfirinas e seus Complexos de Ferro com Albumina Sérica Humana**. Dissertação (Mestrado em Física), Departamento de Física, Pontifícia Universidade Católica. Rio de Janeiro, 138 p., 2008.

HEHRE, W. J. **A Guide to Molecular Mechanics and Quantum Chemical Calculations**. Irvine: WaveFunction, 2003.

HESS, B. et al. LINCS: A linear constraint solver for molecular simulations. **Journal of Computational Chemistry**, v. 18, n. 12, p. 1463–1472, 1997. [https://doi.org/10.1002/\(SICI\)1096-987X\(199709\)18:12<1463::AID-JCC4>3.0.CO;2-H](https://doi.org/10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H)

HOCKNEY, R. W. The potential calculation and some applications (Potential calculation from given source distribution, including direct and iterative methods, error analysis. **Meth. Compt. Phys**, v. 9, p. 136–211, 1970.

HUMPHREY, W.; DALKE, A.; SCHULTEN, K. VMD: Visual molecular dynamics. **Journal of Molecular Graphics**, v. 14, n. 1, p. 33–38, 1996. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5)

IERICH, J. C. M. et al. A Computational Protein Structure Refinement of the Yeast Acetohydroxyacid Synthase. **Journal of the Brazilian Chemical Society**, v. 26, n. 8, p. 1702–1709, 2015. <https://doi.org/10.5935/0103-5053.20150144>

IERICH, J. C. M. et al. Antibody-mediated biorecognition of myelin oligodendrocyte glycoprotein: computational evidence of demyelination-related epitopes. **Scientific Reports**, v. 9, n. 1, p. 2033, 2019. <https://doi.org/10.1038/s41598-018-36578-8>

JORGENSEN, W. L. et al. Comparison of simple potential functions for simulating liquid water. **The Journal of Chemical Physics**, v. 79, n. 2, p. 926, 1983. <https://doi.org/10.1063/1.445869>

JORGENSEN, W. L.; MAXWELL, D. S.; TIRADO-RIVES, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. **Journal of the American Chemical Society**, v. 118, n. 45, p. 11225–11236, 1996. <https://doi.org/10.1021/ja9621760>

JORGENSEN, W. L.; MCDONALD, N. A. Development of an all-atom force field for

heterocycles. Properties of liquid pyridine and diazenes. **Journal of Molecular Structure: THEOCHEM**, v. 424, n. 1–2, p. 145–155, 1998. [https://doi.org/10.1016/S0166-1280\(97\)00237-6](https://doi.org/10.1016/S0166-1280(97)00237-6)

JORGENSEN, W. L.; TIRADO-RIVES, J. The OPLS Potential Functions for Proteins. Energy Minimizations for Crystals of Cyclic Peptides and Crambin. **Journal of the American Chemical Society**, v. 110, n. 6, p. 1657–1666, 1988. <https://doi.org/10.1021/ja00214a001>

KARPLUS, M.; KURIYAN, J. Molecular dynamics and protein function. **Proceedings of the National Academy of Sciences**, v. 102, n. 19, p. 6679–6685, 2005. <https://doi.org/10.1073/pnas.0408930102>

KESSEL, A.; BEN-TAL, N. **Introduction to Proteins: Structure, Function, and Motion**. Londres: CRC Press, 2010. <https://doi.org/10.1201/b10456>

KUNIK, V.; ASHKENAZI, S.; OFRAN, Y. Paratome: an online tool for systematic identification of antigen-binding regions in antibodies based on sequence or structure. **Nucleic Acids Research**, v. 40, n. W1, p. W521–W524, 2012. <https://doi.org/10.1093/nar/gks480>

KUNIK, V.; PETERS, B.; OFRAN, Y. Structural Consensus among Antibodies Defines the Antigen Binding Site. **PLoS Computational Biology**, v. 8, n. 2, p. e1002388, 2012. <https://doi.org/10.1371/journal.pcbi.1002388>

LASKOWSKI, R. A. et al. PROCHECK: a program to check the stereochemical quality of protein structures. **Journal of Applied Crystallography**, v. 26, n. 2, p. 283–291, 1993. <https://doi.org/10.1107/S0021889892009944>

LEACH, A. R. **Molecular Modelling: Principles and Applications**. 2. ed. New York: Pearson Prentice Hall, 2001.

LEVINE, I. N. **Quantum Chemistry**. 4. ed. New York: Pearson Prentice Hall, 1991.

LEVITT, M.; SHARON, R. Accurate simulation of protein dynamics in solution. **Proceedings of the National Academy of Sciences**, v. 85, n. 20, p. 7557–7561, 1988. <https://doi.org/10.1073/pnas.85.20.7557>

LINDAHL, E.; HESS, B.; VAN DER SPOEL, D. GROMACS 3.0: A package for molecular simulation and trajectory analysis. **Journal of Molecular Modeling**, v. 7, n. 8, p. 306–317, 2001. <https://doi.org/10.1007/s008940100045>

MAKAROV, V.; PETTITT, B. M.; FEIG, M. Solvation and Hydration of Proteins and Nucleic Acids: A Theoretical View of Simulation and Experiment. **Accounts of Chemical Research**, v. 35, n. 6, p. 376–384, 2002. <https://doi.org/10.1021/ar0100273>

MARK, P.; NILSSON, L. Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. **The Journal of Physical Chemistry A**, v. 105, n. 43, p. 9954–9960, 2001. <https://doi.org/10.1021/jp003020w>

MELO, F.; SÁNCHEZ, R.; SALI, A. Statistical potentials for fold assessment. **Protein Science**, v. 11, n. 2, p. 430–448, 2002. <https://doi.org/10.1002/pro.110430>

MORRIS, A. L. et al. Stereochemical quality of protein structure coordinates. **Proteins: Structure, Function, and Genetics**, v. 12, n. 4, p. 345–364, 1992.

<https://doi.org/10.1002/prot.340120407>

MURRAY, R. K. et al. **Bioquímica Ilustrada de Harper**. 29. ed. Porto Alegre: AMGH, 2014.

NAMBA, A. M.; DA SILVA, V. B.; DA SILVA, C. H. T. P. Dinâmica molecular: Teoria e aplicações em planejamento de fármacos. **Ecletica Química**, v. 33, n. 4, p. 13–24, 2008. <https://doi.org/10.1590/S0100-46702008000400002>

NELSON, D. L.; COX, M. M. **Princípios de Bioquímica de Lehninger**. 6. ed. Porto Alegre: Artmed, 2011.

NERIA, E.; FISCHER, S.; KARPLUS, M. Simulation of activation free energies in molecular systems. **The Journal of Chemical Physics**, v. 105, n. 5, p. 1902–1921, 1996. <https://doi.org/10.1063/1.472061>

NUTT, D. R.; SMITH, J. C. Molecular Dynamics Simulations of Proteins: Can the Explicit Water Model Be Varied? **Journal of Chemical Theory and Computation**, v. 3, n. 4, p. 1550–1560, 2007. <https://doi.org/10.1021/ct700053u>

OSAJIMA, T. et al. Computational and statistical study on the molecular interaction between antigen and antibody. **Journal of Molecular Graphics and Modelling**, v. 53, p. 128–139, 2014. <https://doi.org/10.1016/j.jmglm.2014.07.005>

PARRINELLO, M.; RAHMAN, A. Polymorphic transitions in single crystals: A new molecular dynamics method. **Journal of Applied Physics**, v. 52, n. 12, p. 7182–790, 1981. <https://doi.org/10.1063/1.328693>

PATODIA, S.; BAGARIA, A.; CHOPRA, D. Molecular Dynamics Simulation of Proteins: A Brief Overview. **Journal of Physical Chemistry & Biophysics**, v. 4, n. 6, p. 1–4, 2014. <https://doi.org/10.4172/2161-0398.1000166>

QIAN, N. On the momentum term in gradient descent learning algorithms. **Neural Networks**, v. 12, n. 1, p. 145–151, 1999. [https://doi.org/10.1016/S0893-6080\(98\)00116-6](https://doi.org/10.1016/S0893-6080(98)00116-6)

RAMACHANDRAN, G. N.; RAMAKRISHNAN, C.; SASISEKHARAN, V. Stereochemistry of polypeptide chain configurations. **Journal of Molecular Biology**, v. 7, n. 1, p. 95–99, 1963. [https://doi.org/10.1016/S0022-2836\(63\)80023-6](https://doi.org/10.1016/S0022-2836(63)80023-6)

RICHARDSON, J. S. The Anatomy and Taxonomy of Protein Structure. In: **Advances in Protein Chemistry**. [s.l: s.n.]. p. 167–339. [https://doi.org/10.1016/S0065-3233\(08\)60520-3](https://doi.org/10.1016/S0065-3233(08)60520-3)

RODRIGUES, C. R. Processos Modernos no Desenvolvimento de Fármacos: Modelagem Molecular. **Cadernos Temáticos de Química Nova na Escola**, v. 3, p. 7, 2001.

SCHLICK, T. **Molecular Modeling and Simulation: An Interdisciplinary Guide**. 2. ed. V. 21. Nova Iorque: Springer, 2010. <https://doi.org/10.1007/978-1-4419-6351-2>

SELA-CULANG, I.; KUNIK, V.; OFRAN, Y. The Structural Basis of Antibody-Antigen Recognition. **Frontiers in Immunology**, v. 4, 2013. <https://doi.org/10.3389/fimmu.2013.00302>

SHEN, M.; SALI, A. Statistical potential for assessment and prediction of protein structures. **Protein Science**, v. 15, n. 11, p. 2507–2524, 2006. <https://doi.org/10.1110/ps.062416606>

SOLOMONS, T. W. G. **Química Orgânica - Volume 1**. 5. ed. Rio de Janeiro: Editora LTC, 2005.

SONG, H. et al. Zika virus NS1 structure reveals diversity of electrostatic surfaces among flaviviruses. **Nature Structural & Molecular Biology**, v. 23, n. 5, p. 456–458, 2016. <https://doi.org/10.1038/nsmb.3213>

SOUZA, P. C. T. **Simulações de Dinâmica Molecular de Mutantes do Receptor do Hormônio Tireoideano**. Dissertação (Mestrado em Química). Programa de Pós-Graduação em Química – Departamento de Físico-Química, Universidade Estadual de Campinas, Campinas, 107 p. 2009.

STAMBULCHIK, E. **Grace 5.1.22**. Cambridge, USA Grace Team, 2008. Disponível em: <<http://plasma-gate.weizmann.ac.il/Grace/>>. Acesso em: 12 abr. 2020.

TORSHIN, I. Y.; WEBER, I. T.; HARRISON, R. W. Geometric criteria of hydrogen bonds in proteins and identification of “bifurcated” hydrogen bonds. **Protein Engineering**, v. 15, n. 5, p. 359–363, 2002. <https://doi.org/10.1093/protein/15.5.359>

TRABUCO, L.; VILLA, E. **Salt Bridges Plugin, Version 1.1**. Disponível em: <<http://www.ks.uiuc.edu/Research/vmd/plugins/saltbr/>>. Acesso em: 10 ago. 2020.

VAN DER SPOEL, D. et al. GROMACS: Fast, flexible, and free. **Journal of Computational Chemistry**, v. 26, n. 16, p. 1701–1718, 2005. <https://doi.org/10.1002/jcc.20291>

VAN GUNSTEREN, W. F. et al. Biomolecular modeling: Goals, problems, perspectives. **Angewandte Chemie - International Edition**, v. 45, n. 25, p. 4064–4092, 2006. <https://doi.org/10.1002/anie.200502655>

VAN GUNSTEREN, W. F.; BERENDSEN, H. J. C. Computer Simulation of Molecular Dynamics: Methodology, Applications, and Perspectives in Chemistry. **Angewandte Chemie International Edition in English**, v. 29, n. 9, p. 992–1023, 1990. <https://doi.org/10.1002/anie.199009921>

VERLET, L. Computer “experiments” on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. **Physical Review**, v. 159, n. 1, p. 98–103, 1967. <https://doi.org/10.1103/PhysRev.159.98>

VERLI, H. et al. **Bioinformática : da Biologia à Flexibilidade Molecular**. 1. ed. São Paulo: Sociedade Brasileira de Bioquímica e Biologia Molecular (SBBq), 2014.

VOET, D.; VOET, J. G.; PRATT, C. W. **Fundamentos de Bioquímica – A vida em nível molecular**. 2. ed. Porto Alegre: Artmed, 2008.

WANG, D. et al. A mutation identified in neonatal microcephaly destabilizes zika virus NS1 assembly in vitro. **Scientific Reports**, v. 7, n. 42580, p. 1–10, 2017. <https://doi.org/10.1038/srep42580>

YANG, Z. et al. UCSF Chimera, MODELLER, and IMP: An integrated modeling system. **Journal of Structural Biology**, v. 179, n. 3, p. 269–278, 2012. <https://doi.org/10.1016/j.jsb.2011.09.006>

INTERAÇÃO ANTÍGENO-ANTICORPO: DOCKING MOLECULAR PROTEÍNA- PROTEÍNA

3.1 DOCKING PROTEÍNA-PROTEÍNA

De modo geral, as interações das proteínas com outras moléculas e, em particular, com outras proteínas, formando complexos de caráter permanente ou transiente, influenciam diretamente nas funções por elas desempenhadas. Portanto, as interações proteína-proteína (PPI, do inglês, *protein-protein interaction*) constituem processos-chave de muitos eventos biológicos básicos essenciais para a manutenção da vida. Nesse sentido, compreender os modos de reconhecimento molecular entre as superfícies de interação de duas proteínas torna-se uma tarefa relevante para diversas aplicações envolvendo sistemas biológicos, como os complexos antígeno-anticorpo deste estudo.

A obtenção da estrutura tridimensional desses complexos é o modo mais adequado de compreender as interações atômicas presentes nas interfaces. Essas interfaces de interação, em nível atômico, podem ser caracterizadas por métodos experimentais como Cristalografia de Raios-X e Ressonância Magnética Nuclear (RMN) que, além de demandarem grande quantidade de recursos, tem suas aplicações limitadas a questões de ordem técnica como, respectivamente, a obtenção de cristais de boa qualidade para difração e o tamanho da proteína em estudo (DESTA et al., 2020; HIGA, 2009). Nesse contexto, na ausência de dados estruturais do complexo formado entre proteínas, torna-se promissor o uso de abordagens computacionais, em especial, o Docking Molecular.

O Docking Molecular proteína-proteína pode ser definido como uma técnica computacional rápida e acessível de predição da conformação do complexo (estrutura 3D) a partir de suas estruturas proteicas componentes como base para o reconhecimento, afinidade e especificidade (LYSKOV; GRAY, 2008; PEDOTTI et al., 2011).

Segundo Higa (2009, p. 3), “basicamente, os algoritmos de molecular docking executam um processo de busca para encontrar a orientação espacial preferencial de uma molécula com

relação a uma segunda, de acordo com uma função de mérito” também denominada função de *scoring* ou função de pontuação característica de cada *software* utilizado.

Apesar de se tratar de uma técnica recente e em constante desenvolvimento capaz de obter estruturas de complexos multi-moleculares, com qualidade similar a de técnicas experimentais, o Docking Molecular possui uma das tarefas mais desafiadoras e interessantes da bioinformática estrutural. Isso se deve a algumas razões como: (1) elevado número de combinações possíveis para ajuste de superfícies; (2) as estruturas das moléculas não são exatas, possuindo erros experimentais; (3) as biomoléculas sofrem uma mudança na conformação de cargas quando associadas, chamada de ajuste induzido (DUHOVNY; NUSSINOV; WOLFSON, 2002; KOSHLAND, 1958).

Portanto, os algoritmos de Docking Molecular proteína-proteína devem considerar esses fatores na busca do melhor modelo para a descrição da interação desses complexos. Logo, é necessário que as funções de *scoring* ponderem, conforme abordagem aplicada, os parâmetros estruturais e energéticos, relacionados às propriedades físico-químicas, envolvidos nas PPI (GUEDES; MAGALHÃES; DARDENNE, 2014).

Fundamentalmente, é necessário destacar a classificação das técnicas de Docking Molecular proteína-proteína conforme as estruturas individuais das moléculas componentes: (1) ligado (*Bound*), estrutura do complexo experimentalmente determinada; (2) não-ligado (*Unbound*), estruturas das proteínas experimentalmente determinadas de modo individual, sejam elas livres ou ligadas a outro par; (3) modelada (*Modeled*), estruturas previstas por modelagem por homologia e/ou outras técnicas computacionais como predições *ab initio* ou dinâmica molecular (PEDOTTI et al., 2011).

No presente trabalho, para a busca global, foram utilizadas estruturas não-ligadas como conformações de partida, em especial, por se tratar de um complexo antígeno-anticorpo. Isso porque, a maioria das estruturas de anticorpo, disponíveis no banco de dados, estão na forma de complexos, visto que, estruturas de anticorpos livres não são úteis já que mudam drasticamente sua conformação após a ligação (RINI; SCHULZE-GAHMEN; WILSON, 1992).

Dentre os principais servidores de docking disponíveis na internet para interações proteína-proteína temos: ClusPro (COMEAU et al., 2004), GRAMM-X (TOVCHIGRECHKO; VAKSER, 2006) e ZDOCK (CHEN; LI; WENG, 2003) com métodos de busca baseados em

Transformada Rápida de Fourier (FFT, do inglês, *fast-Fourier Transform*), PatchDock e SymmDock (SCHNEIDMAN-DUHOVNY et al., 2005) baseados nos princípios de complementariedade de forma e restrições simétricas (*hashing* geométrico) e Hex (RITCHIE; KEMP, 2000) baseado na representação de esferas harmônicas.

Foram escolhidos para uso no presente trabalho os *softwares* ClusPro e PatchDock pelas seguintes razões: (1) caracterizados por serem métodos sistemáticos de docking de busca global e de corpo rígido, nos quais a flexibilidade da proteína é mínima, sendo permitidas apenas algumas sobreposições – os métodos de corpo rígido possuem a vantagem de tolerar diferenças moderadas entre as estruturas ligada e não-ligada (DESTA et al., 2020); (2) possuem versão *on-line* disponível necessitando, portanto, de um menor tempo computacional para prever modelos *all-atoms* e dispensando o uso de clusters; (3) dispõem do modo anticorpo (*antibody mode*), que restringe a região de busca às CDRs do anticorpo; (4) a qualidade dos modelos preditos é periodicamente avaliada, desde 2002, pelos experimentos internacionais em andamento de Avaliação Crítica de Interações Preditas (CAPRI, do inglês, *Critical Assessment of PRedicted Interactions*) (JANIN, 2002); (5) possuem busca global baseada em métodos distintos, ampliando a amostragem de estruturas de complexos para avaliação. Nos tópicos a seguir serão descritos com maiores detalhes os métodos de busca e as funções de *scoring* adotadas por ambos os *softwares*.

3.1.1 ClusPro

O ClusPro trata-se de um *webservice* de docking de corpo rígido entre proteínas que utiliza o PIPER, um algoritmo baseado na abordagem da correlação de Transformada Rápida de Fourier (FFT) – introduzida, em 1992, por Katchalski-Katzir e colaboradores – e que usa potencial de pares como parte da função de *scoring*.

Neste método, uma das proteínas (receptor) é colocada na origem do sistema de coordenadas em um grid (grade tridimensional) fixo e a outra proteína (ligante) é inserida em um grid móvel com passos rotacional e translacional randômicos de 5° e 1 Å, respectivamente. A energia de interação é então escrita na forma de uma função de correlação. A eficiência numérica do método parte do princípio de que as funções de energia podem ser eficientemente calculadas usando FFT e resulta na habilidade de amostrar exaustivamente bilhões de

conformações (10^9 – 10^{10} conformações) de duas proteínas interagindo com a energia sendo avaliada a cada ponto (KOZAKOV et al., 2017).

Esses métodos atuais de FFT para docking introduzem funções de *scoring* mais complexas e precisas incluindo, simultaneamente, termos que representam as interações eletrostáticas e de desolvatação. A função de *scoring* do servidor possui a seguinte forma (BRENKE et al., 2012; COMEAU et al., 2004; KOZAKOV et al., 2017; VAJDA et al., 2017):

$$E = E_{vdw} + w_3 E_{elec} + w_4 E_{pair} \quad (3.1)$$

na qual, o termo E_{vdw} corresponde à implementação da energia de van der Waals com os componentes repulsivo (E_{rep}) e atrativo (E_{att}) (Equação 3.2), E_{elec} a energia eletrostática coulômbica (Equação 3.3) e E_{pair} denota o potencial de pares.

$$E_{vdw} = w_1 E_{rep} + w_2 E_{att} \quad (3.2)$$

$$E_{elec} = \sum_{i=1}^{N_r} \sum_{j=1}^{N_l} \frac{q_i q_j}{\left[r_{ij}^2 + D^2 \exp\left(-r_{ij}/4D^2\right) \right]^{1/2}} \quad (3.3)$$

na qual, q_i e q_j são as cargas dos átomos i e j , r_{ij} a distância entre esses átomos na proteína e D o raio de van der Waals para cada átomo.

Enquanto a forma matemática desses dois termos é idêntica independente de qual seja a natureza do complexo proteico, o potencial de pares (E_{pair}) possui para complexos antígeno-anticorpo uma forma específica, adaptada da forma inicial do potencial baseada em estrutura de pares desenvolvida pelo *Decoys as the Reference State* (E_{DARS}). O E_{ADARS} (*assymmetric DARS*) foi adequadamente adaptado para complexos dessa natureza considerando que eles são, em grande parte, planos, menos hidrofóbicos e possuem assimetria inerente entre os dois lados da interface. Portanto, esse termo é responsável por otimizar a localização da interface

permitindo a sua obtenção com conformação mais próxima à nativa. Isso porque as estatísticas de interação de complexos desse tipo foram obtidas com base na premissa de que um átomo de determinado tipo no anticorpo (pertencente ao paratopo) difere das preferências de um átomo do mesmo tipo no antígeno (pertencente ao epítipo) (BRENKE et al., 2012).

Após a obtenção dos modelos dos complexos, o programa opera uma segunda etapa, a organização de clusters das 1000 estruturas melhor ranqueadas dentro do parâmetro energético. Portanto, são selecionados os centros (estrutura média) dos 10 clusters mais populosos de estruturas de baixa energia, em contraponto aos demais *softwares* que usualmente selecionam as conformações de mais baixa energia. Para tanto, são calculados os valores de *i*RMSD (RMSD de interface) para cada uma das 1000 estruturas e, por sua vez, localizada aquela com maior número de estruturas vizinhas dentro do raio de 9 Å formando, assim, o primeiro cluster. O processo é então sucessivamente repetido com as estruturas restantes até a obtenção de todos os clusters. Como as estruturas de cada cluster são selecionadas em uma faixa energética de baixa amplitude, mesmo com os erros associados a essas medidas, é razoável considerar que a energia da estrutura média representa as demais dentro do cluster. Além disso, o ranqueamento dos cluster é baseado no fato de que a probabilidade de ocorrência de cada cluster (P_k) é proporcional ao tamanho da sua população (N_k) conforme equação 3.4:

$$P_k = \exp(-E/RT) \times N_k / Z \quad (3.4)$$

em que, E é a energia da estrutura média do cluster de ordem k e Z a função de partição do espaço conformacional amostral considerado como um *ensemble* canônico (KOZAKOV et al., 2005, 2013; SHORTLE; SIMONS; BAKER, 1998).

Por fim, o *software* realiza uma etapa final de minimização da energia, utilizando o campo de força CHARMM, das estruturas preditas pelo docking, responsável por remover potenciais efeitos estéricos, mas sem mudanças significativas na conformação dos complexos (BROOKS et al., 1983; KOZAKOV et al., 2013).

Além disso, cabe salientar que o *webservice* tem cerca de 15.000 usuários registrados e realizou 98.300 cálculos de docking, em 2019, tendo um impacto relevante na literatura ao ser citado por, aproximadamente, 600 publicações (DESTA et al., 2020).

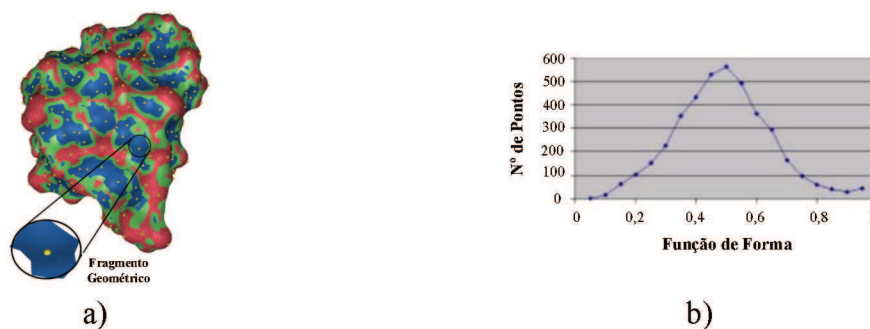
3.1.2 PatchDock

O PatchDock é um *webservice* de docking de corpo rígido entre proteínas que utiliza um algoritmo baseado nos princípios de complementariedade molecular de forma introduzidos por Connolly (1986) e Kuntz et al. (1982). Portanto, trata-se de um modelo baseado na geometria que tem o objetivo de encontrar transformações que induzem amplas áreas de interface e número reduzido de efeitos estéricos garantindo boa complementariedade molecular de forma (SCHNEIDMAN-DUHOVNY et al., 2005).

O software foi testado e validado tanto para complexos proteína-ligante quanto antígeno-anticorpo, com boas taxas de acerto para estruturas próximas às nativas. Para tanto, se utiliza basicamente de três etapas descritas a seguir (CHEN et al., 2003).

Após o upload dos arquivos, é realizada a etapa de representação da forma molecular na qual é calculada a superfície molecular das proteínas pelo programa MS (CONNOLLY, 1983). Essa superfície é então segmentada em três porções geométricas de áreas similares (côncava, convexa ou plana – Figura 3.1a) conforme cálculo da função de forma. A função de forma em cada ponto pode ser definida como a fração não-acessível ao solvente, de uma esfera de raio 6 Å, colocada sobre esse ponto. Portanto, para cada ponto é anotado um valor entre 0 e 1, no qual, valores menores que $\frac{1}{3}$ correspondem a regiões convexas (*knobs*), maiores que $\frac{2}{3}$ a regiões côncavas (*hole*) e no intervalo entre esses limites a regiões planas (*flat*) conforme histograma apresentado na Figura 3.1b (DUHOVNY; NUSSINOV; WOLFSON, 2002).

Figura 3.1 – Identificação dos fragmentos geométricos (azul: côncavo, vermelho: convexo e verde: plano) em uma proteína (a) conforme função de forma (b) aplicada pelo algoritmo implementado no *webservice* PatchDock.



Fonte: Adaptado de Axenopoulos et al. (2011); Duhovny; Nussinov e Wolfson (2002)

Essas porções são divididas conforme o número de centros de gravidade de cada face (1, 2 ou 3 átomos) que, por sua vez, são computados como um centroide e projetados na direção normal da superfície, sendo chamados de pontos críticos, representados em amarelo na Figura 3.1a. Esses pontos são o input para a segunda etapa responsável por realizar a correspondência entre os fragmentos. Vale considerar que ao operar no modo anticorpo, a busca por esses pontos críticos se restringe aos fragmentos geométricos que se intersectam nas CDRs do anticorpo.

Portanto, dados os fragmentos do receptor e do ligante então são computadas as transformações hipotéticas de encaixe baseados na complementariedade geométrica local a partir do Hashing Geométrico (WOLFSON; RIGOUTSOS, 1997). A ideia central é de que os fragmentos côncavos se conectem a fragmentos convexos e fragmentos planos a qualquer outro tipo. Poses com valores de RMSD menor que 4,0 Å (parâmetro definido na interface gráfica de *upload* dos arquivos das estruturas *input*) são então clusterizadas (SCHNEIDMAN-DUHOVNY et al., 2005; STOCKMAN, 1987).

Por fim, as poses candidatas, selecionadas na etapa anterior, são avaliadas em relação a existência de impedimento estérico entre os átomos do receptor e do ligante. São descartadas poses cuja distância entre o grid de transformadas do ligante e as coordenadas de cada ponto da superfície do receptor sejam maiores que o limite de penetração (5.0 Å). Por fim, é realizada a pontuação das poses restantes de acordo com uma função de complementariedade de forma (DUHOVNY; NUSSINOV; WOLFSON, 2002).

Para refinamento dos resultado é sugerido o uso do FireDock que otimiza as conformações da cadeia lateral na interface das poses ranqueadas (MASHIACH et al., 2010).

No entanto, mesmo após a realização de uma busca global a discriminação entre modelos precisos e imprecisos, ou seja, correspondentes à estrutura de conformação nativa, ainda é uma tarefa complexa. Nesse sentido, é sugerido um reranking dessas estruturas, ou seja, repontuar os modelos de modo a realizar uma discriminação entre configurações representativas e pouco representativas. Uma das alternativas é conferir graus de liberdade aos complexos (cadeias laterais e regiões de loop) por meio de um software de Docking Molecular baseado em métodos de busca local, ou seja, em regiões pré-selecionadas (PIERCE; WENG, 2008). Nesse caso, foi utilizado o software Rosetta (LYSKOV; GRAY, 2008) que possui uma ampla aplicabilidade em estudos computacionais envolvendo anticorpos, além de também possuir uma versão on-line disponível. Abordagens similares de reranking já foram aplicadas com sucesso

em estudos de sistemas antígeno-anticorpo que vislumbravam a proteína de envelope (E) das flavirose Dengue e Zika como marcador antigênico (SIMONELLI et al., 2010, 2013; WANG et al., 2017b).

3.1.3 Rosie

O ROSIE (do inglês, *The Rosetta Online Server that Includes Everyone*) é a versão online e totalmente automatizada (LYSKOV et al., 2013) do amplamente utilizado software Rosetta, que de acordo com Moretti et al. (2018, p. 260, nossa tradução)

é um pacote de modelagem molecular que fornece uma ampla gama de ferramentas para predição e design de macromoléculas biológicas. Rosetta tem sido utilizado em um número de publicações diversas, desde a previsão de estrutura de proteína, estrutura de ácido ribonucleico (RNA), interações proteína-proteína, interações proteína-peptídeo e interações proteína-pequenas moléculas, para o uso de ressonância magnética nuclear e informação de densidade de elétrons na predição e refinamento de estrutura e para o design de nova proteínas.

Para aplicações no estudo de PPI, o RosettaDock é uma ferramenta de docking molecular de corpo rígido e busca local que se utiliza de um algoritmo de busca de Monte Carlo requerendo para tanto uma conformação de partida entre as proteínas indicando a região de busca. O protocolo utilizado (Local Docking) realiza uma amostragem ativa da orientação dos pares de docking com passos de translação e rotação de 5° e 0,7 Å, respectivamente (perturbação de baixa resolução). Em seguida, é realizado um refinamento de alta resolução que inclui a otimização e minimização das cadeias laterais das estruturas iniciais e uma pequena melhoria do corpo rígido dos pares. Portanto, os autores indicam o uso do software para o refinamento e reranking de poses candidatas obtidas a partir de qualquer software de docking de busca global (GRAY et al., 2003; LYSKOV; GRAY, 2008).

Dados de Moretti e colaboradores (2018) apontam que o *webserver* tinha cerca de 5.000 usuários registrados e realizou mais de 31.000 cálculos desde sua implementação, em 2012, com uma média mensal de 875 cálculos. Em termos de demanda computacional, foram necessárias mais de 3,5 milhões de horas de CPU ou o equivalente a aproximadamente 100

CPUs em uso contínuo nos últimos quatro anos fornecidas pelo cluster Stampede (Texas - EUA).

3.2 METODOLOGIA

Primeiramente, a avaliação da capacidade de predição dos *softwares* ClusPro e PatchDock foi realizada por meio do redocking da estrutura cristalográfica do complexo formado pelo anticorpo 22-anti-NS1 e da NS1 do Vírus do Nilo do Oeste (NS1-WNV), PDB ID 4OII. Os resultados de Silva (2020) serão apresentados na Seção 3.3, visto que se trata de uma estrutura nativa cujas características e propriedades são relevantes na discussão dos resultados.

Para identificar a região de interação e o modo de ligação mais favorável para os sistemas antígeno-anticorpo (NS1/22-anti-NS1) da Dengue e Zika, foi aplicado o protocolo proposto por Silva (2020) composto por quatro etapas descritas a seguir:

Etapa 1: Busca Global - Docking Molecular Proteína-Proteína

Os modelos de complexos para os sistemas estudados foram obtidos utilizando os softwares online de Docking Molecular Proteína-Proteína ClusPro (BRENKE et al., 2012) e PatchDock (SCHNEIDMAN-DUHOVNY et al., 2005) cujos algoritmos de busca global se baseiam, respectivamente, nos métodos de transformada de Fourier e hashing geométrico.

Os arquivos de entrada, no formato PDB, para ambos os softwares eram constituídos das coordenadas estruturais das proteínas, de modo que o anticorpo 22-anti-NS1 foi definido com o receptor e a proteína NS1 como ligante. Todos os cálculos foram realizados no modo antígeno-anticorpo, que restringe a região de busca às CDRs do anticorpo, e em quintuplicata, para evitar possíveis predições geradas equivocadamente por erros de software.

Ambos *softwares* são de uso livre em suas plataformas disponíveis nos seguintes sítios eletrônicos: www.bioinfo3d.cs.tau.ac.il/PatchDock/ e www.cluspro.bu.edu.

Etapa 2: Top 10 - Seleção dos 10 modelos mais favoráveis segundo critério energético

Cada software gerou 1000 modelos ranqueados de acordo com a função de scoring própria para cada um dos sistemas. A energia de ligação mínima foi o critério utilizado para a seleção das 10 melhores configurações de ambos os softwares (Top 10). Em particular, para os modelos obtidos no PatchDock, o ranqueamento das 1000 configurações foi realizado pelo software FireDock (MASHIACH et al., 2008).

Etapa 3: Viabilidade Biológica – Exclusão por Inacessibilidade

Considerando o fato de que a SNS1 é a forma detectável dessa proteína, formada por interações hidrofóbicas entre os domínios β -roll de três unidades da forma mNS1, foi realizada a exclusão das configurações, pertencentes ao Top 10, que apresentavam interação nessa região ou nas proximidades. A visualização que possibilitou a exclusão dos complexos, por não apresentarem susceptibilidade de interação em virtude da inacessibilidade do domínio β -roll do antígeno, foi realizada no *software* de visualização gráfica VMD (HUMPHREY; DALKE; SCHULTEN, 1996).

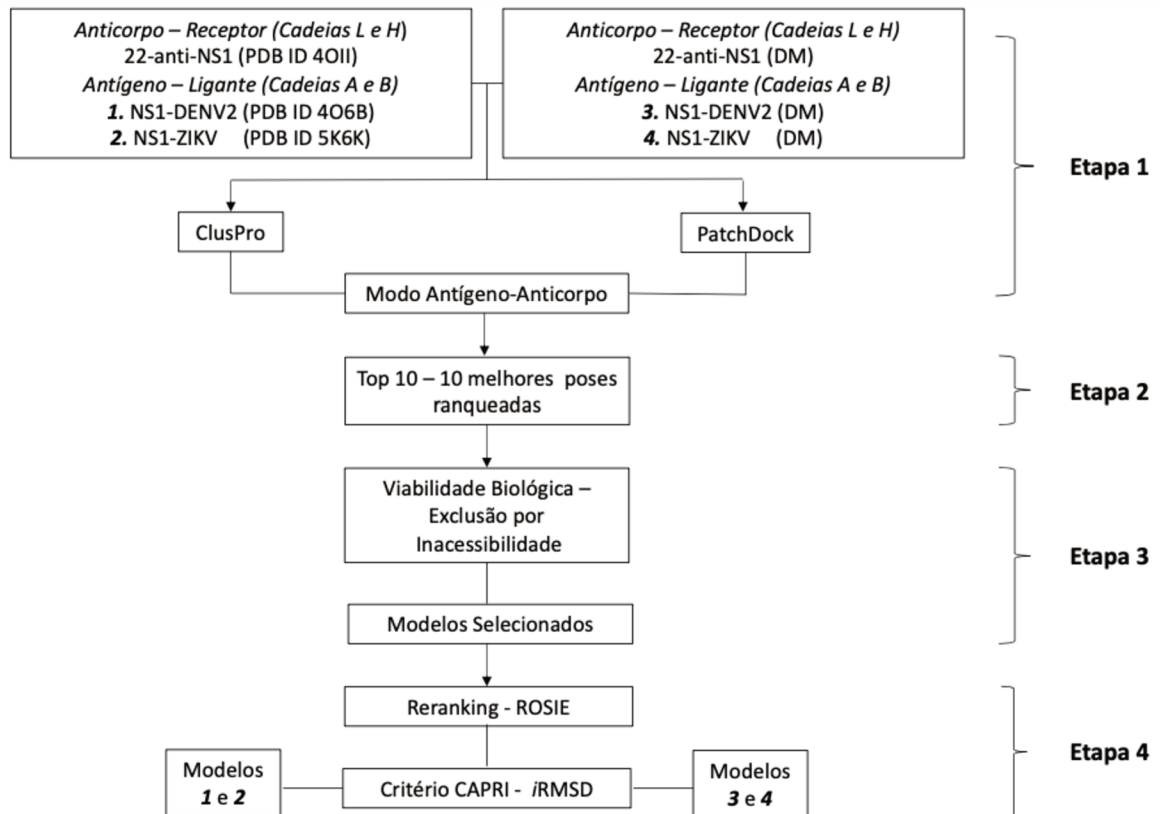
Etapa 4: Busca Local e Reranking – Aplicação dos critérios CAPRI

Por fim, foram realizados cálculos de docking com busca local com objetivo de repontuar todas as estruturas dos complexos, selecionados na etapa anterior, com base no critério de comportamento de estrutura nativa. O reranqueamento das estruturas foi realizado utilizando o ROSIE (LYSKOV et al., 2013). O módulo Docking 2 do *webserver* foi utilizado com o protocolo Local Docking e a sequência LH_AB como pares de docking. Os arquivos PDB input combinados dos complexos foram preparados conforme critérios estabelecidos por (LYSKOV; GRAY, 2008). Os gráficos de score *versus* RMSD de interface (*i*RMSD) foram avaliados para aplicação dos critérios CAPRI.

Esse *software* também é de uso livre em sua plataforma disponível no sítio eletrônico: <https://rosie.rosettacommons.org/>.

O fluxograma apresentado na Figura 3.2 sintetiza a metodologia de docking molecular empregada no protocolo proposto.

Figura 3.2 – Fluxograma da metodologia de Docking Molecular.



Fonte: Adaptado de Silva (2020)

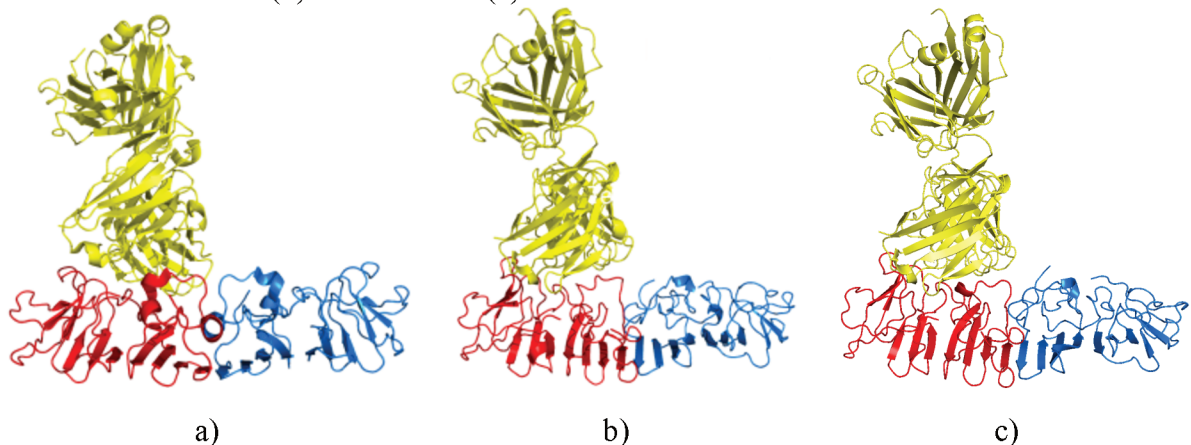
Para realizar a análise das interfaces de interação dos complexos foi utilizado o *software* AppA (NGUYEN et al., 2017; NGUYEN; VERMA; ZHONG, 2019), disponível no sítio eletrônico <http://mspc.bii.a-star.edu.sg/minhn/appa.html>. As interfaces de interação foram obtidas considerando os resíduos de contato que são identificados como sendo aqueles cuja área acessível ao solvente (ASA) na estrutura do anticorpo ou do antígeno muda com a formação do complexo e estão a 6,0 Å do epítipo ou paratopo na forma complexada (NGUYEN et al., 2017). As interfaces obtidas foram analisadas e as imagens confeccionadas utilizando o *software* Pymol (DELANO, 2020). As interações do tipo π - π foram computadas utilizando o *webserver* RING 2.0 (PIOVESAN; MINERVINI; TOSATTO, 2016).

3.3 RESULTADOS E DISCUSSÃO

3.3.1 Redocking

Os resultados de redocking, previamente apresentados por Silva (2020), indicam a capacidade de ambos os *softwares* de prever corretamente a região de interação para o complexo antígeno-anticorpo. Essa conclusão é corroborada pela análise das superfícies de interação da pose mais bem ranqueada (top 1) predita por cada um dos *webservers*, conforme pode ser observado na Figura 3.3.

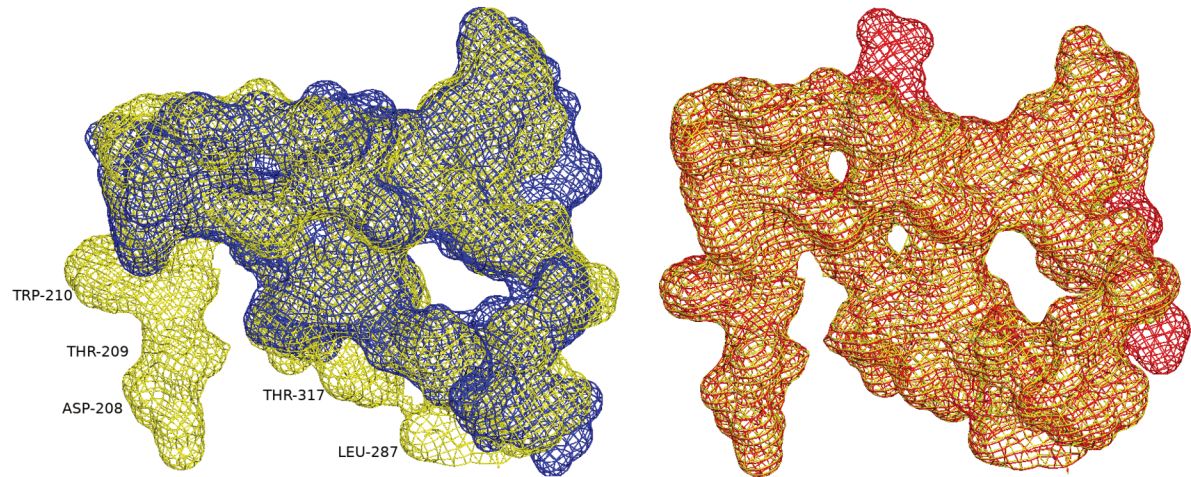
Figura 3.3 – Representações *new cartoon* da estrutura cristalográfica do sistema WNV/22-anti-NS1 (a) e das poses melhor ranqueadas preditas pelo redocking realizado utilizando os *webservers* ClusPro (b) e PatchDock (c).



Fonte: Autor e (SILVA, 2020)

Na estrutura cristalográfica (PDB ID 4OII) são anotados 31 e 29 resíduos de contato do anticorpo e da proteína NS1, respectivamente. Para a pose 1 predita pelo ClusPro 29 e 23 desses resíduos idênticos foram identificados indicando *fitting* de 0,94 (RMSD: 0,0 e Z-Score: 34,80) e 0,82 (RMSD: 0,7 nm e Z-Score: 12,31) entre as interfaces das estruturas do anticorpo e da proteína, respectivamente (NGUYEN et al., 2019). Já a pose mais bem ranqueada pelo PatchDock (top 1) a predição foi de elevada precisão com *fitting* de 1,00 (RMSD: 0,02 nm e Z-Score: 40,80).

Figura 3.4 – Representação em malha do *fitting* entre a interface de interação da estrutura cristalográfica WNV/22-anti-NS1, em coloração amarela, com poses melhor ranqueadas previstas pelo redocking realizado utilizando os *webservers* ClusPro, em azul, e PatchDock, em vermelho.



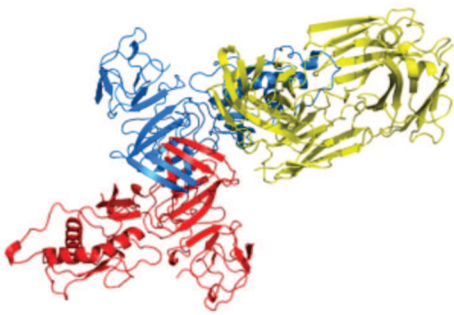
Fonte: Autor

Nesse sentido, o redocking indica a capacidade de prever com precisão considerável a superfície de interação das poses mais bem ranqueadas em ambos *webservers* em comparação com o complexo cristalizado sendo, portanto, utilizados nos cálculos de docking molecular realizados a seguir.

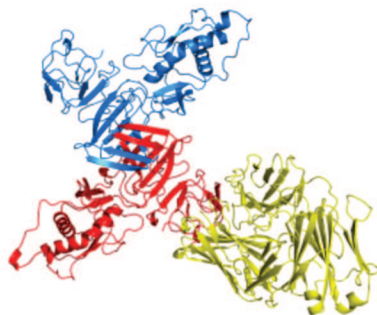
3.3.2 Docking Molecular

As configurações dos sistemas, após avaliação do critério de viabilidade biológica, realizada na etapa 3, estão apresentadas nas Figuras 3.5-3.8, nas quais, a estrutura do Fab do anticorpo, composta pelas cadeias L (leve) e H (pesada), está representada na cor amarela, e as cadeias A e B da estrutura do antígeno (NS1) de ambas flavirose, em vermelho e azul, respectivamente. Vale considerar que a nomenclatura DC e DP foi adotada para indicar poses previstas para os sistemas envolvendo a proteína NS1-DENV2 (D), respectivamente, pelos *webservers* ClusPro (C) e PatchDock (P). Analogamente, ZC e ZP são as nomenclaturas adotadas para as poses previstas para os sistemas envolvendo a proteína NS1-ZIKV. Os números cardinais que seguem essas notações indicam a posição de cada uma das poses dentro das 10 melhor ranqueadas (*top 10*) na predição dos *webservers*.

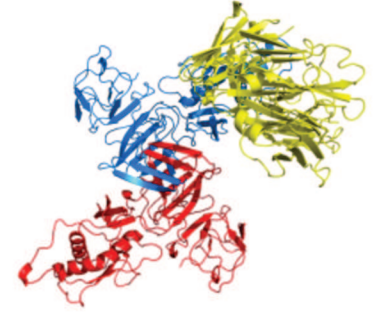
Figura 3.5 – Representação *new cartoon* das poses de menor energia selecionadas dos *webservers* PatchDock (a-c) e ClusPro (d-h), conforme critério de viabilidade biológica, ordenadas em forma crescente de energia para a NS1-DENV2 Cristalográfica (I).



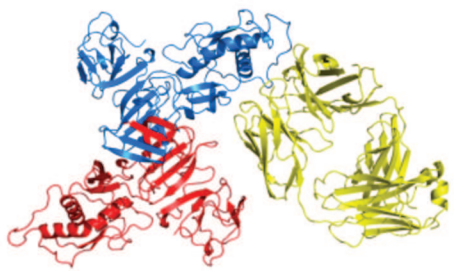
a) DP07



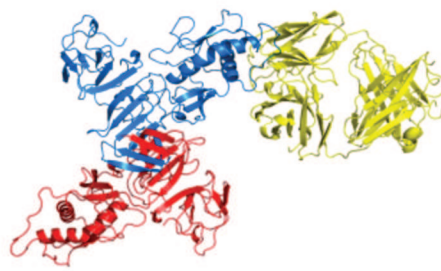
b) DP08



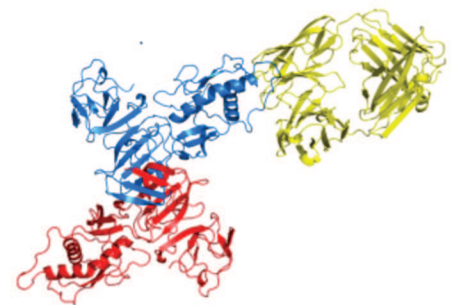
c) DP09



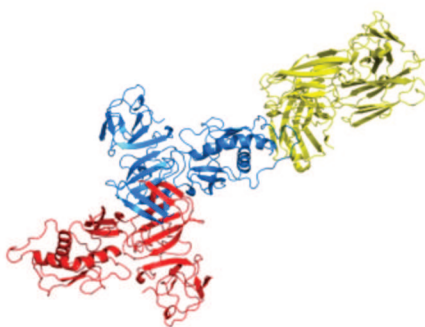
d) DC01



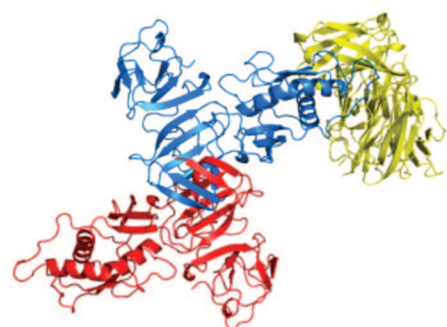
e) DC02



f) DC05



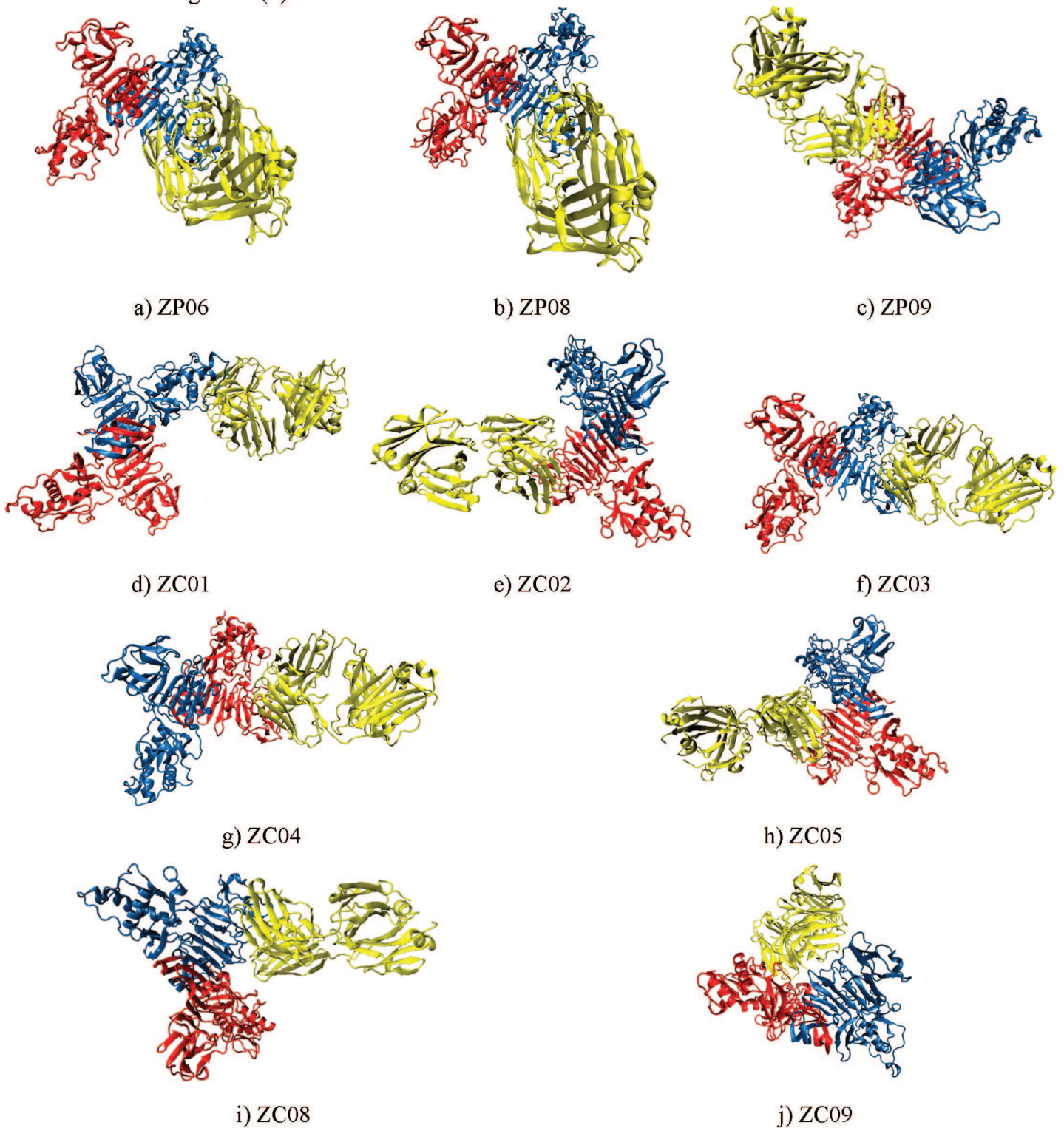
g) DC07



h) DC09

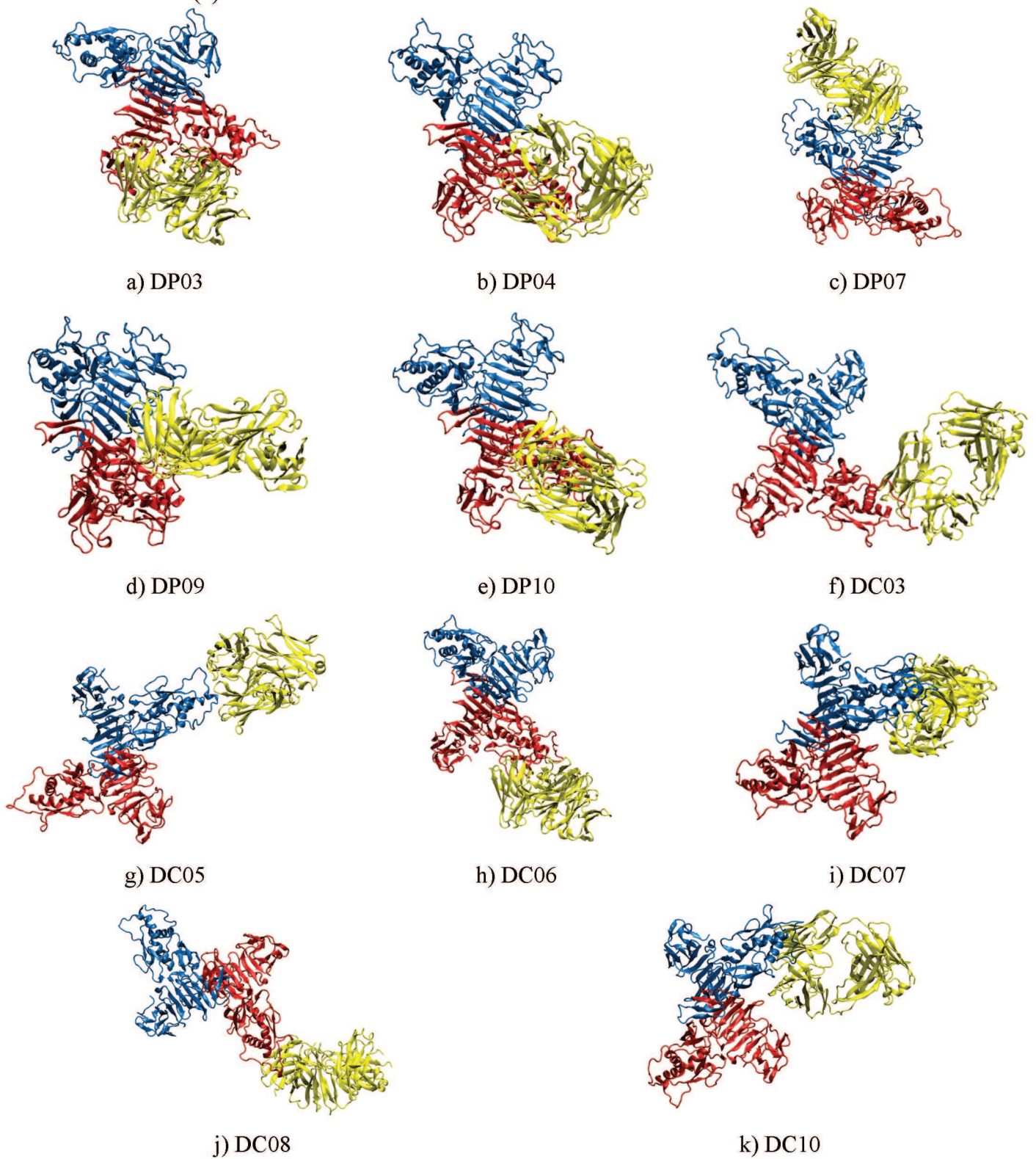
Fonte: (SILVA, 2020)

Figura 3.6 – Representação *new cartoon* das poses de menor energia selecionadas dos *webservers* PatchDock (a-c) e ClusPro (d-j), conforme critério de viabilidade biológica, ordenadas em forma crescente de energia para a NS1-ZIKV Cristalográfica (2).



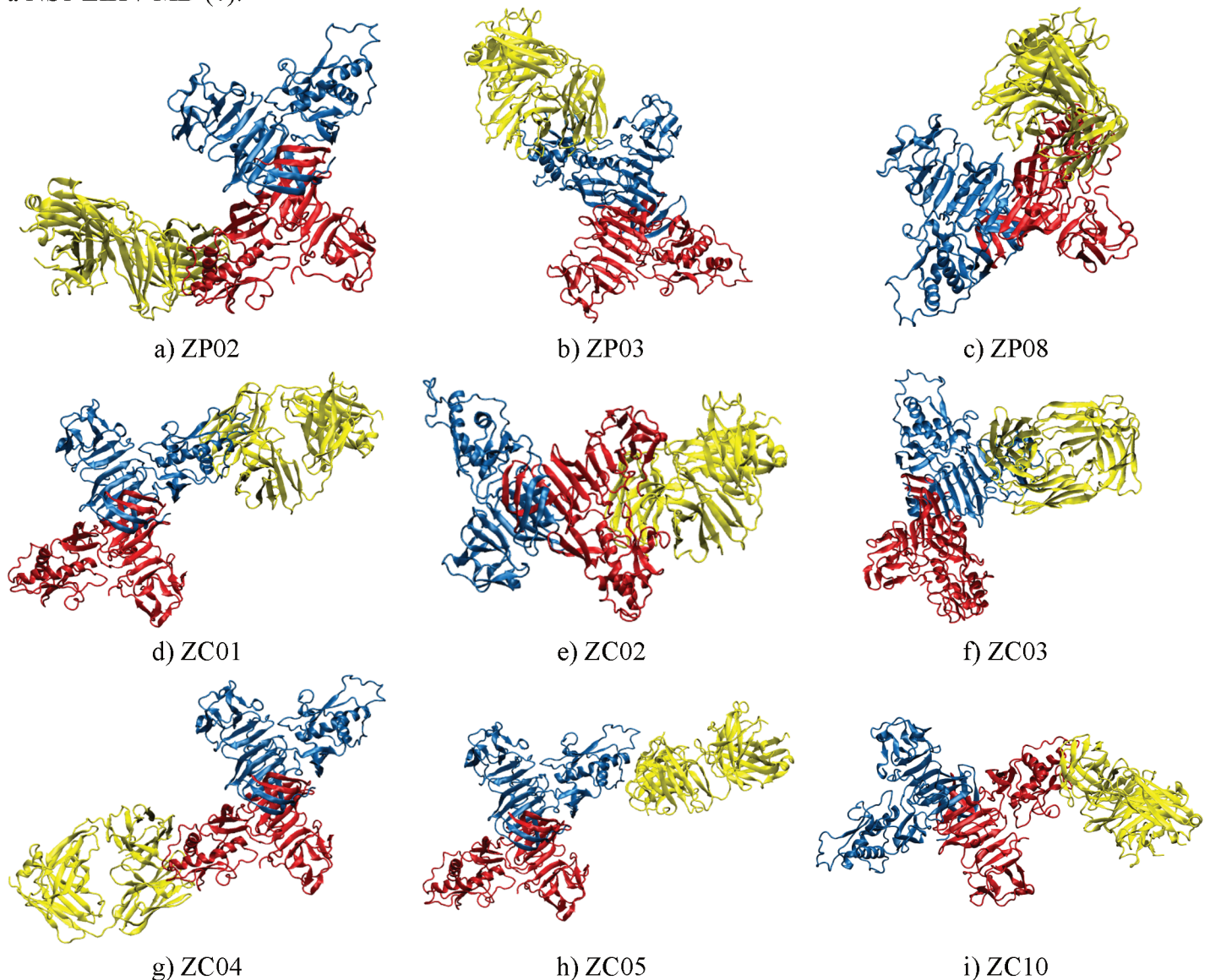
Fonte: Autor

Figura 3.7 – Representação *new cartoon* das poses de menor energia selecionadas do *webservers* PatchDock (a-e) e ClusPro (f-k), conforme critério de viabilidade biológica, ordenadas em forma crescente de energia para a NS1-DENV2 MD (3).



Fonte: Autor

Figura 3.8 – Representação *new cartoon* das poses de menor energia selecionadas do *webservers* PatchDock (a-c) e ClusPro (d-i), conforme critério de viabilidade biológica, ordenadas em forma crescente de energia para a NS1-ZIKV MD (4).



Fonte: Autor

Acerca do apresentado nas Figuras 3.5–3.8, é possível observar que em relação à predição dos dois *webservers*, o ClusPro foi capaz de prever maior número de poses biologicamente viáveis no *top* 10 para as proteínas de ambas flaviviruses, tanto em sua forma cristalográfica quanto sob efeito de solvatação. Isto está intrinsicamente relacionado ao método de busca empregado pelo algoritmo de cada um dos *softwares*. Do ponto de vista exclusivamente estrutural e da complementariedade (método de busca do PatchDock), pode-se sugerir que, os sítios de ligação dos anticorpos são superfícies amplas capazes de acomodar

epítomos conformacionais de antígenos extensos a fim de assegurar que haja uma força de ligação adequada para garantir a estabilidade do complexo (ABBAS; LICHTMAN; PILLAI, 2011). Nos resultados do docking molecular com a estrutura dimérica das proteínas NS1, os padrões de folhas- β dos domínios β -roll e β -ladder encontram-se disponíveis na superfície β -ladder. Na face oposta (superfície de loop), o domínio β -ladder é caracterizado por regiões de alças. Portanto, sugere-se que pode ter havido uma indicação preferencial desse *software* pela região de folhas- β do domínio β -roll, por seu caráter similar ao das cadeias leve e pesadas do anticorpo. No entanto, essas não estão acessíveis na forma hexamérica e, portanto, não atendem ao critério de viabilidade biológica. Além disso, segundo Duhovny; Nussinov e Wolfson (2002) a complementariedade de forma é muito mais proeminente em interfaces côncavo/convexo, como é o caso de complexos enzima-inibidor, em detrimento de interfaces planas, caso dos complexos antígeno-anticorpo. Isso indica uma limitação do princípio de busca do *webservice* PatchDock em relação a sistemas dessa natureza como é o caso do presente trabalho.

Ainda em relação às Figuras 3.5–3.8, no tocante aos domínios de interação preferenciais para cada um dos sistemas temos: (1) todas as poses apresentaram o domínio wing como região de interação; (2) todas as poses preditas apresentaram o domínio β -ladder como região de interação, com exceção da configuração ZC01; (3) todas as poses preditas pelo ClusPro apresentaram interfaces de interação no domínio wing, bem como as poses DP04 e DP10. Já as poses DP03 e DP09 indicam interação na região de interface entre os dois domínios e, apenas a pose DP07 apresentou interação exclusivamente no domínio β -ladder; (4) todas as poses apresentaram o domínio wing como região de interação, com exceção, das configurações ZC02 e ZC03 que apresentaram interação, respectivamente, na interface entre os domínios e apenas no domínio β -ladder. Esses resultados podem ser adequadamente compreendidos sob a perspectiva dos estudos teórico-experimentais sintetizados na Tabela 1.1:

- No estudo de Jiang e colaboradores (2010), ao utilizarem anticorpos policlonais, foi identificada uma série de epítomos lineares para a proteínas NS1-DENV2, sendo eles os correspondentes às seguintes sequências: a.a. 32-40; a.a. 80-89; a.a. 103-112; a.a. 121-130; a.a. 187-196; a.a. 295-304 e a.a. 315-324. No entanto, na forma hexamérica, detectável como analito, nem todos estão acessíveis. Posteriormente, Jones e colaboradores (2017) indicaram que os resíduos que de fato estão acessíveis para interação com anticorpo correspondem a sequências pertencentes ao domínio wing: a.a. 97-116; a.a. 125-134; a.a. 143-151 e a.a. 170-178. Esses dados são corroborados pelo

estudo teórico-experimental de Freire et al. (2017) que apontou a localização de três epítomos nesse domínio: a.a. 50-65; a.a. 74-88 e a.a. 108-129. Portanto, os epítomos da proteína NS1-DENV2 encontram-se preferencialmente localizados no domínio wing de acordo com a maioria das configurações de docking obtidas pelos dois *webservers* para os sistemas 1 e 3;

- A divergência de região de interação preferencial nos sistemas envolvendo a NS1-ZIKV (sistemas 2 e 4) tem relação com o fato de essa apresentar uma interface, caracterizada por uma superfície variada e potencial eletrostático divergente, que pode afetar as propriedades de ligação à anticorpos protetivos (XU et al., 2016);
- Nesse sentido, o trabalho de Wang e colaboradores (2017) aponta que a NS1-ZIKV apresenta maior similaridade estrutural em relação a NS1-WNV em detrimento da NS1-DENV2. Dessa forma, ao interagir com o anticorpo 22-anti-NS1 a interface entre as proteínas ocorre na região de loops do domínio β -ladder, com conservação de 13 dos 21 epítomos identificados na NS1-WNV. Portanto, é justificável a orientação preferencial da maioria das poses do sistema 2;
- No entanto, no estudo de Lee et al. (2018) foram identificados oito epítomos para a proteína NS1-ZIKV sendo dois deles (a.a. 70-84 e a.a. 118-147), localizados no domínio wing, apontados como de alta especificidade para esse antígeno. Nesse sentido, os autores apontaram que essa região do domínio wing na superfície de loop, apresenta diferenças significativas entre as proteínas, inclusive em termos de potencial eletrostático, indicando sua importância para a ligação a anticorpos protetivos. A relevância desses epítomos justifica as configurações preferencialmente obtidas para o sistema 4. Todavia, não se sobrepõe a importância de as interfaces de interação terem sido identificadas em regiões diferentes das proteínas, no sentido de viabilizar o diagnóstico diferencial entre as flavirozes.

3.3.3 Reranking

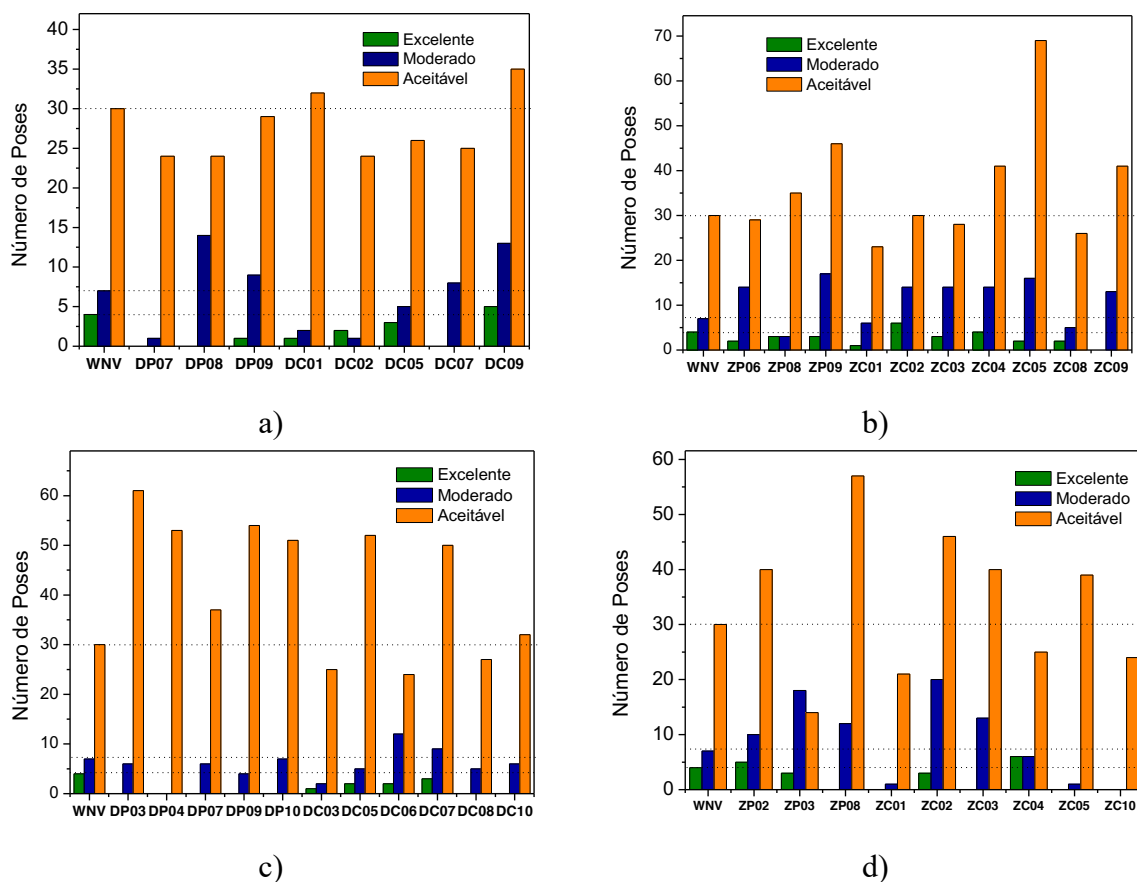
O reranking das estruturas selecionadas, conforme critério de viabilidade biológica, permitiu a obtenção de 1.000 novas configurações que foram avaliadas de acordo com sua qualidade, conforme definido pelo CAPRI (JANIN *et al.*, 2003). De acordo com esse critério

as estruturas são classificadas da seguinte forma e representadas com a notação que se segue: $iRMSD \leq 1$: Excelente (***) , $1 < iRMSD \leq 2$: Moderado (**) e $2 < iRMSD \leq 4$: Aceitável (*) (LENSINK; MÉNDEZ; WODAK, 2007). Portanto, a notação adotada, seguindo o modelo utilizado pelo CAPRI, é a seguinte (x^{***}, y^{**}, z^*) , na qual x , y e z são números naturais e indicam, respectivamente, o número de estruturas de qualidade excelente (***), moderada (**) e aceitável (*).

Como não há conhecimento da estrutura nativa do complexo para as flaviroses em estudo, uma abordagem aplicável é o comparativo dos resultados das poses candidatas com os de uma estrutura nativa de características e/ou naturezas semelhantes. Caso os resultados apresentem um comportamento igual ou similar aos observados na estrutura nativa há indicativos de que a pose escolhida possui caráter nativo (JANIN, 2005). Isso porque a estrutura de um complexo na sua conformação nativa encontra-se no seu estado de maior estabilidade termodinâmica, associada à sua energia livre de ligação. Dessa forma, se uma pose candidata apresenta valor de $iRMSD > 4$, foram necessários ajustes consideráveis, durante o refinamento/otimização realizado pelo ROSIE, para que o complexo atingisse energia de ligação favorável. Portanto, mudanças conformacionais significativas sugerem que esta configuração não se encontra em sua forma nativa.

Nesse sentido, é esperado que a configuração mais representativa da forma nativa do complexo seja aquela que apresenta um número de poses com qualidades semelhantes ao do sistema definido com referência. O perfil utilizado como referência foi o da estrutura cristalográfica do complexo NS1–WNV/22-anti-NS1 representada nos gráficos da Figura 3.9 pelo conjunto das três primeiras barras ($4^{***}+7^{**}+30^*$). Desse modo, esse complexo após refinamento possui quatro estruturas de excelente qualidade (barra verde nos gráficos), sete estruturas de qualidade moderada (barra azul nos gráficos) e 30 estruturas aceitáveis (barra alaranjada nos gráficos).

Figura 3.9 – Número de poses após reranking conforme critério estabelecido pela CAPRI para os sistemas: (a) NS1-DENV2 Cristalográfica, (b) NS1-ZIKV Cristalográfica, (c) NS1-DENV2 MD e (d) NS1-ZIKV MD.



Fonte: Autor

Conforme percebido, pelo auxílio das linhas tracejadas, as poses que possuem comportamento mais similar ao da estrutura de referência conforme número de modelos excelentes, moderados e aceitáveis são, para os sistemas: (1) NS1-DENV2: DC09 (5***+13**+35*); (2) NS1-ZIKV: ZC02 (6***+14**+30*), ZC04 (4***+14**+41*) e ZP09 (3***+17**+46*); (3) NS1-DENV2 MD: DC07 (3***+9**+50*) e (4) NS1-ZIKV MD: ZC02 (3***+20**+46*) e ZP02 (5***+10**+40*).

Desse modo, o reranking possibilita a discriminação entre os modelos preditos com base nas suas precisões. Além disso, é possível observar que houve um melhor desempenho, em termos de qualidade, dos modelos preditos pelo *webserver* ClusPro. Esse resultado está em consonância com o observado nas últimas rodadas do CAPRI (2009, 2013 e 2016) que apontam o ClusPro como o servidor automatizado de Docking com melhor desempenho, quando

comparado a outros servidores, e com desempenho similar ou superior ao de grupos humanos de pesquisa em predição, para alvos de dificuldade moderada (VAJDA et al., 2017). O bom desempenho se refere aos complexos do tipo proteína-inibidor, proteína-proteína e antígeno-anticorpo. Os autores sugerem que a etapa de ranqueamento do algoritmo, baseada no tamanho do cluster, torna confiável a predição de configurações de elevada precisão (KOZAKOV et al., 2013; KOZAKOV; SCHUELER-FURMAN; VAJDA, 2008).

No entanto, vale ainda colocar que bons resultados, nas avaliações do CAPRI, têm sido obtidos por outros *softwares* de docking que realizam busca global, como PatchDock, ZDOCK e GRAMM-X (LENSINK; WODAK, 2010, 2013).

3.3.4 Análise das Interfaces de Interação

A análise das interfaces de interação foi realizada para avaliação das poses preditas e classificadas acima como candidatas a configurações de conformação nativa. Nessa etapa, serão avaliados separadamente os sistemas referentes às flavirose Dengue (1 e 3) e Zika (2 e 4).

a) Dengue

Considerando-se as conclusões obtidas na seção anterior e os estudos *in silico* e *in vitro* apresentados na Tabela 1.1 foram considerados nessa análise apenas os epítomos localizados no domínio wing. Portanto, conforme critérios adotados por Silva (2020), foram consideradas as seguintes sequências a.a. 50-116; a.a. 122-134 e a.a. 143-151, nomeadas por Epítomo 1 (E1), Epítomo 2 (E2) e Epítomo 3 (E3), respectivamente.

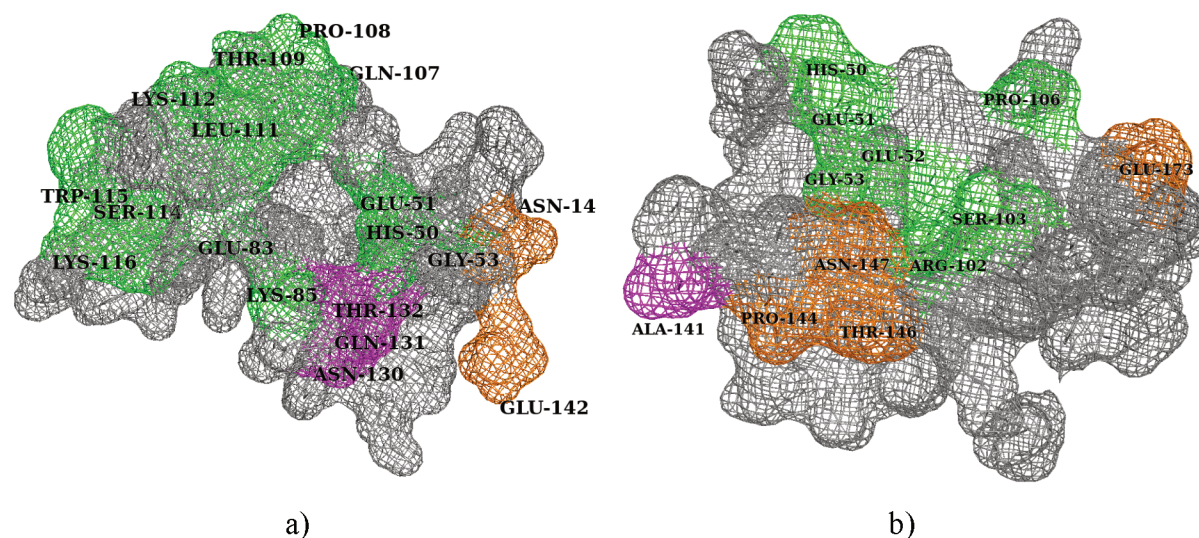
Para o sistema (1), o modelo DC09 já foi definido como a estrutura mais representativa do sistema com 18 resíduos de aminoácidos na superfície pertencentes aos epítomos. Em relação ao sistema (3), a estrutura DC07 foi a única que apresentou parâmetros de *i*RMSD similares a da estrutura nativa com 12 resíduos de aminoácidos na superfície pertencentes a epítomos. Além disso, essa similaridade se verifica ainda nas interações presentes nas interfaces de ambas as configurações (Figura 3.10), sumarizadas na Tabela 3.1.

Tabela 3.1 – Número de resíduos pertencentes aos epítomos e das interações presentes na interface dos modelos preditos.

| Sistema | Modelo | Número de Resíduos | Ligações de Hidrogênio | Interações de van der Waals | Interações hidrofóbicas | Interações Iônicas |
|----------|--------|--------------------|------------------------|-----------------------------|-------------------------|--------------------|
| <i>1</i> | DC09 | 18 | 14 | 105 | 39 | 8 |
| <i>3</i> | DC07 | 12 | 18 | 110 | 28 | 7 |

Fonte: Autor

Figura 3.10 – Interfaces de interação dos modelos dos complexos NS1-DENV2/22-anti-NS1: DC09 (a) e DC07 (b), com os resíduos pertencentes aos epítomos E1, E2 e E3 destacados nas cores verde, magenta e alaranjado, respectivamente.



Fonte: Autor

Os dados apresentados na Tabela 3.1 estão de acordo com o estudo de Dalkas et al. (2014), segundo o qual as interações do tipo ligação de hidrogênio e as de van der Waals/hidrofóbicas são as mais frequentes e, portanto, contribuem mais significativamente para a estabilidade de complexos antígeno-anticorpo. Esse estudo destaca ainda que interações do tipo π - π (interações hidrofóbicas envolvendo dois anéis aromáticos) e π -cátion, apesar de presentes em menor quantidade, favorecem a estabilidade das interfaces de interação envolvendo antígeno e anticorpo. Portanto, tratam-se de interações relevantes no reconhecimento específico desses complexos.

Os resíduos His-50, Glu-51 e Gly-53 encontram-se presentes em ambas interfaces de interação estabelecendo com as CDRs do anticorpo – DC09, sistema 1, e o DC07, para o sistema 3 –, respectivamente, três e cinco ligações de hidrogênio, cinco e nenhuma interação iônica e 18 e 21 interações vdW. Nesse sentido, indica-se a importância, em especial, das ligações de hidrogênio dentre as PPIs estabelecidas por esses resíduos para a formação das superfícies de interação, em especial, pelo caráter hidrofílico dos resíduos de Histidina e Ácido Glutâmico.

Para o modelo DC09 (sistema 1), as interações do tipo π - π intermoleculares ocorrem entre os resíduos Tyr-32(H) --- His-50(B) e Trp-33(H) --- His-129(B) indicando que esses resíduos nos seus respectivos epítomos contribuem para o reconhecimento específico.

b) Zika

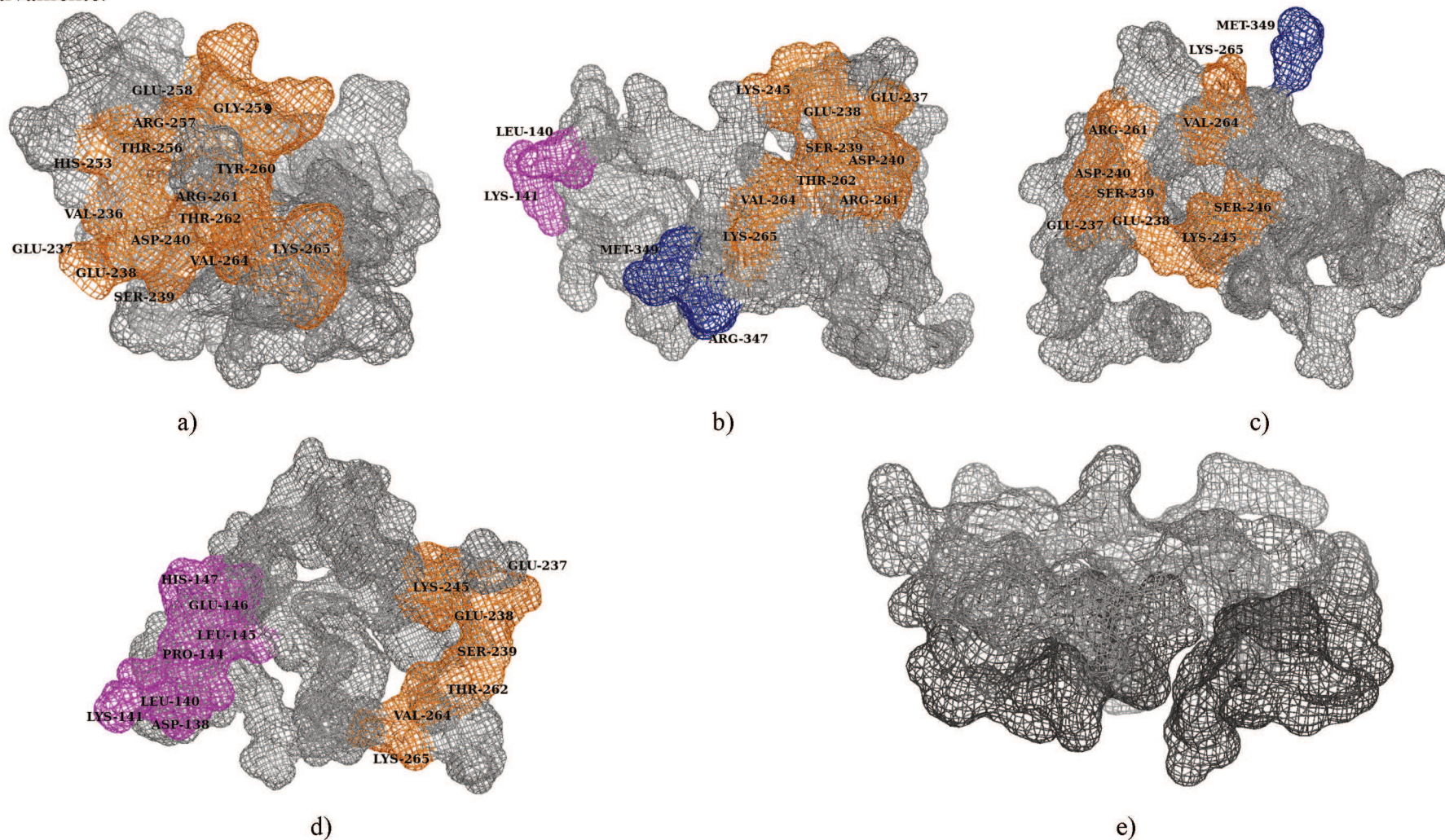
Considerando-se as conclusões obtidas na seção anterior e os estudos *in silico* e *in vitro* realizados, foram considerados nessa análise tanto os epítomos localizados no domínio wing quanto no domínio β -ladder. Portanto, foram consideradas as seguintes sequências: a) Domínio wing: a.a. 70-84 (Epítomo 1 – E1) e a.a. 118-147 (Epítomo 2 – E2); b) Domínio β -ladder: a.a. 236-280 (Epítomo 3 – E3) e a.a. 341-355 (Epítomo 4 – E4) (FREIRE et al., 2017; LEE et al., 2018; WANG et al., 2017). As análises do número de resíduos pertencentes à epítomos e das interações identificadas nas interfaces de cada configuração bem como suas representações estão apresentadas, respectivamente, na Tabela 3.2 e na Figura 3.11.

Tabela 3.2 – Número de resíduos pertencentes ao epítomos e das interações presentes na interface dos modelos preditos.

| Sistema | Modelo | Número de Resíduos | Ligações de Hidrogênio | Interações de van der Waals | Interações hidrofóbicas | Interações Iônicas |
|---------|--------|--------------------|------------------------|-----------------------------|-------------------------|--------------------|
| 2 | ZC02 | 15 | 23 | 150 | 40 | 25 |
| | ZC04 | 13 | 27 | 171 | 31 | 15 |
| | ZP09 | 10 | 10 | 218 | 42 | 11 |
| 4 | ZC02 | 14 | 13 | 100 | 43 | 13 |
| | ZP02 | 0 | 4 | 176 | 37 | 0 |

Fonte: Autor

Figura 3.11 – Interfaces de interação dos modelos dos complexos NS1-ZIKV/22-anti-NS1: ZC02 (a), ZC04 (b) e ZP09 (c) referentes ao sistema 2 e ZC02 (d) e ZP02 (e) referentes ao sistema 4, com os resíduos pertencentes aos epítotos E1, E2, E3 e E4 destacados nas cores verde, magenta, alaranjado e azul, respectivamente.



Fonte: Autor

Analisando as imagens da Figura 3.11 e os valores na Tabela 3.2, é possível verificar que ZC02, para o sistema 2, e ZC02, para o sistema 4, são os modelos que apresentam a maior quantidade de resíduos pertencentes a epítomos. Nesse sentido, é relevante anotar a importância, em especial, do epítomo E2 e também do E3 nos modelos preditos, indicando que tanto o domínio β -ladder quanto o wing contribuem para uma configuração nativa nos complexos formados por proteínas NS1-ZIKV. Por outro lado, o epítomo E1, localizado no domínio wing, apesar do caráter imunogênico não possui contribuição para os modelos preditos.

Vale destacar que os resíduos Glu-237, Glu-238, Asp-240, Arg-261, Val-264, Lys-265 estão conservados nas superfícies de interação dentre todos esses modelos preditos. Dentre os modelos selecionados para os sistemas 2 e 4 estão descritas a seguir a contribuição, em termos de interações, dos principais dentre os resíduos supracitados:

ZC02 (sistema 2): Glu-237 e Glu-238 (3 h-bond, 4 interações iônicas e 20 vdW), e Lys-265 (9 h-bond e 2 interações iônicas).

ZC02 (sistema 4): Glu-237 e Glu-238 (4 h-bond, 4 interações iônicas e 19 vdW) e Lys-265 (4 h-bond e 2 interações iônicas).

Nesse sentido, destaca-se a relevância das ligações de hidrogênio e das interações iônicas, dentre as PPIs estabelecidas por esses resíduos para a formação das superfícies de interação, bem como a similaridade no número e natureza dessas interações.

Para o modelo ZC02 (sistema 2), ocorre uma interação do tipo π - π Tyr-92(L) --- His-253(A) e uma π -cátion Tyr-98(H) --- Arg-261(A). Já para o modelo ZC02 (sistema 4) as interações do tipo π - π intermoleculares ocorrem entre os resíduos Tyr-98(H) --- Trp-98(A) e Trp-92(L) --- Tyr-175(A). Essas interações reforçam a importância dos resíduos identificados nas interfaces no reconhecimento molecular específico entre a proteína NS1-ZIKV e o anticorpo.

Dessa forma, essas interfaces preditas apresentam resíduos realmente significativos para uma interação do tipo antígeno-anticorpo, com caráter de nativa. Numa tendência geral, esses modelos têm, em comparação com os demais, número maior de ligações de hidrogênio e interações hidrofóbicas e iônicas, e menor número de interações do tipo van der Waals. Segundo Abbas; Lichtman e Pillai (2011, p. 102) “a relativa importância de cada uma destas interações depende das estruturas do sítio de ligação de cada anticorpo e do determinante antigênico” não sendo possível estabelecer uma relação *genérica* entre o número dessas interações não-covalentes e o caráter nativo e/ou afinidade dos complexos. No entanto, ainda assim nesse caso,

observa-se similaridade entre a natureza e o número de interações estabelecidas na PPIs para os modelos selecionados, principalmente em relação aos resíduos conservados nas interfaces de interação.

3.4 CONCLUSÕES PARCIAIS

Nesse capítulo foi apresentado e aplicado um protocolo para o uso de *webservers* de docking molecular em complexos das flavirose Dengue e Zika. O procedimento utiliza o modelo de corpo rígido totalmente automatizado com diferentes métodos de busca, critérios CAPRI e conhecimento prévio dos sistemas em estudo. O protocolo mostrou-se promissor na obtenção de modelos representativos, com boa precisão e concordantes com estudos experimentais, para sistemas antígeno-anticorpo de natureza semelhante ao das flavirose, superando assim as principais limitações da técnica de Docking Molecular Proteína-Proteína. Nesse sentido, o conhecimento prévio do sistema em estudo é primordial para a aplicação do protocolo proposto.

Nesse sentido, o método de busca global baseado em correlação de Transformada Rápida de Fourier (FFT), implementado no *webservice* ClusPro, foi mais eficiente na predição das configurações selecionadas. Portanto, os modelos selecionados – DC09 (sistema 1), ZC02 (sistema 2), DC07 (sistema 3) e ZC02 (sistema 4) – apontaram convergência das regiões de interação entre os sistemas envolvendo as proteínas em sua forma cristalográfica e sob efeito de solvatação bem como a conservação de resíduos pertencentes a epítomos entre os modelos analisados. Os domínios wing e β -ladder são, respectivamente, as regiões de interação preferencial do anticorpo com as proteínas NS1-DENV2 e NS1-ZIKV. Essa diferença entre as regiões de interação para os complexos das flavirose Dengue e Zika, em um primeiro momento, indica a possibilidade de aplicação dos complexos estudados em diagnóstico diferencial. A predominância das interações do tipo ligação de hidrogênio e hidrofóbicas em maiores quantidades, bem como interações específicas do tipo π - π e π -cátion nas interfaces de interação reforçam o caráter nativo dos complexos selecionados.

No próximo capítulo, parâmetros físico-químicos serão utilizados para complementar a escolha dos modelos, realizada com base no protocolo proposto neste capítulo, e definir a magnitude da afinidade de ligação em cada um deles.

REFERÊNCIAS

- ABBAS, A. K.; LICHTMAN, A. H.; PILLAI, S. Anticorpos e Antígenos. In: **Imunologia Celular e Molecular**. 7. ed. Rio de Janeiro: Elsevier, 2011. p. 89–108.
- AXENOPOULOS, A. et al. A Shape Descriptor for Fast Complementarity Matching in Molecular Docking. **IEEE/ACM Transactions on Computational Biology and Bioinformatics**, v. 8, n. 6, p. 1441–1457, 2011. <https://doi.org/10.1109/TCBB.2011.72>
- BRENKE, R. et al. Application of asymmetric statistical potentials to antibody-protein docking. **Bioinformatics**, v. 28, n. 20, p. 2608–2614, 2012. <https://doi.org/10.1093/bioinformatics/bts493>
- BROOKS, B. R. et al. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. **Journal of Computational Chemistry**, v. 4, n. 2, p. 187–217, 1983. <https://doi.org/10.1002/jcc.540040211>
- CHEN, R. et al. A protein-protein docking benchmark. **Proteins: Structure, Function, and Genetics**, v. 52, n. 1, p. 88–91, 2003. <https://doi.org/10.1002/prot.10390>
- CHEN, R.; LI, L.; WENG, Z. ZDOCK: An initial-stage protein-docking algorithm. **Proteins: Structure, Function, and Genetics**, v. 52, n. 1, p. 80–87, 2003. <https://doi.org/10.1002/prot.10389>
- COMEAU, S. R. et al. ClusPro: an automated docking and discrimination method for the prediction of protein complexes. **Bioinformatics**, v. 20, n. 1, p. 45–50, 2004. <https://doi.org/10.1093/bioinformatics/btg371>
- CONNOLLY, M. L. Analytical molecular surface calculation. **Journal of Applied Crystallography**, v. 16, n. 5, p. 548–558, 1983. <https://doi.org/10.1002/bip.360250705>
- CONNOLLY, M. L. Shape complementarity at the hemoglobin $\alpha 1\beta 1$ subunit interface. **Biopolymers**, v. 25, n. 7, p. 1229–1247, 1986.
- DALKAS, G. A. et al. Cation- π , amino- π , π - π , and H-bond interactions stabilize antigen-antibody interfaces. **Proteins: Structure, Function, and Bioinformatics**, v. 82, n. 9, p. 1734–1746, 2014.
- DELANO, W. L. **The PyMOL Molecular Graphics System, Version 2.3**. Schrödinger LLC, 2020.
- DESTA, I. T. et al. Performance and Its Limits in Rigid Body Protein-Protein Docking. **Structure**, v. 28, n. 9, p. 1071–1081.e3, 2020. <https://doi.org/10.1016/j.str.2020.06.006>
- DUHOVNY, D.; NUSSINOV, R.; WOLFSON, H. J. Efficient Unbound Docking of Rigid Molecules. In: **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**. [s.l.: s.n.]. p. 185–200. https://doi.org/10.1007/3-540-45784-4_14
- FREIRE, M. C. L. C. et al. Mapping Putative B-Cell Zika Virus NS1 Epitopes Provides Molecular Basis for Anti-NS1 Antibody Discrimination between Zika and Dengue Viruses. **ACS Omega**, v. 2, p. 3913–3920, 2017. <https://doi.org/10.1021/acsomega.7b00608>

- GRAY, J. J. et al. Protein–Protein Docking with Simultaneous Optimization of Rigid-body Displacement and Side-chain Conformations. **Journal of Molecular Biology**, v. 331, n. 1, p. 281–299, 2003. [https://doi.org/10.1016/S0022-2836\(03\)00670-3](https://doi.org/10.1016/S0022-2836(03)00670-3)
- GROMIHA, M. M.; YUGANDHAR, K.; JEMIMAH, S. Protein–protein interactions: scoring schemes and binding affinity. **Current Opinion in Structural Biology**, v. 44, p. 31–38, 2017. <https://doi.org/10.1016/j.sbi.2016.10.016>
- GUEDES, I. A.; MAGALHÃES, C. S. DE; DARDENNE, L. E. Atracamento Molecular. In: VERLI, H. (Ed.). **Bioinformática: da biologia à flexibilidade molecular**. 1. ed. São Paulo: Sociedade Brasileira de Bioquímica e Biologia Molecular - SBBq, 2014. p. 189–208.
- HIGA, R. H. **Predição de regiões de interface proteína-proteína baseada em informações estruturais**. Tese (Doutorado em Engenharia Elétrica), Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas. Campinas, 146 p., 2009.
- HUMPHREY, W.; DALKE, A.; SCHULTEN, K. VMD: Visual molecular dynamics. **Journal of Molecular Graphics**, v. 14, n. 1, p. 33–38, 1996. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5)
- JANIN, J. Welcome to CAPRI: A Critical Assessment of PRedicted Interactions. **Proteins: Structure, Function, and Genetics**, v. 47, n. 3, p. 257–257, 2002. <https://doi.org/10.1002/prot.10111>
- JANIN, J. et al. CAPRI: A Critical Assessment of PRedicted Interactions. **Proteins: Structure, Function, and Genetics**, v. 52, n. 1, p. 2–9, 2003. <https://doi.org/10.1002/prot.10381>
- JANIN, J. Assessing predictions of protein-protein interaction: The CAPRI experiment. **Protein Science**, v. 14, n. 2, p. 278–283, 2005. <https://doi.org/10.1110/ps.041081905>
- JIANG, L. et al. Selection and identification of B-cell epitope on NS1 protein of dengue virus type 2. **Virus Research**, v. 150, n. 1–2, p. 49–55, 2010. <https://doi.org/10.1016/j.virusres.2010.02.012>
- JONES, M. et al. Computational Identification of Antibody Epitopes on the Dengue Virus NS1 Protein. **Molecules**, v. 22, n. 4, p. 607, 2017. <https://doi.org/10.3390/molecules22040607>
- KATCHALSKI-KATZIR, E. et al. Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques. **Proceedings of the National Academy of Sciences**, v. 89, n. 6, p. 2195–2199, 1992. <https://doi.org/10.1073/pnas.89.6.2195>
- KOSHLAND, D. E. Application of a Theory of Enzyme Specificity to Protein Synthesis. **Proceedings of the National Academy of Sciences**, v. 44, n. 2, p. 98–104, 1958. <https://doi.org/10.1073/pnas.44.2.98>
- KOZAKOV, D. et al. Optimal Clustering for Detecting Near-Native Conformations in Protein Docking. **Biophysical Journal**, v. 89, n. 2, p. 867–875, 2005. <https://doi.org/10.1529/biophysj.104.058768>
- KOZAKOV, D. et al. How good is automated protein docking? **Proteins: Structure, Function, and Bioinformatics**, v. 81, n. 12, p. 2159–2166, 2013.

<https://doi.org/10.1002/prot.24403>

KOZAKOV, D. et al. The ClusPro web server for protein–protein docking. **Nature Protocols**, v. 12, n. 2, p. 255–278, 2017. <https://doi.org/10.1038/nprot.2016.169>

KOZAKOV, D.; SCHUELER-FURMAN, O.; VAJDA, S. Discrimination of near-native structures in protein-protein docking by testing the stability of local minima. **Proteins: Structure, Function, and Bioinformatics**, v. 72, n. 3, p. 993–1004, 2008. <https://doi.org/10.1002/prot.21997>

KUNTZ, I. D. et al. A geometric approach to macromolecule–ligand interactions. **Journal of Molecular Biology**, v. 161, n. 2, p. 269–288, 1982. [https://doi.org/10.1016/0022-2836\(82\)90153-X](https://doi.org/10.1016/0022-2836(82)90153-X)

LEE, H.-J. et al. Identification of peptide based B-cell epitopes in Zika virus NS1. **Biochemical and Biophysical Research Communications**, v. 505, n. 4, p. 1010–1014, 2018. <https://doi.org/10.1016/j.bbrc.2018.10.024>

LENSINK, M. F.; MÉNDEZ, R.; WODAK, S. J. Docking and scoring protein complexes: CAPRI 3rd Edition. **Proteins: Structure, Function, and Bioinformatics**, v. 69, n. 4, p. 704–718, 2007. <https://doi.org/10.1002/prot.21804>

LENSINK, M. F.; WODAK, S. J. Docking and scoring protein interactions: CAPRI 2009. **Proteins: Structure, Function, and Bioinformatics**, v. 78, n. 15, p. 3073–3084, 2010. <https://doi.org/10.1002/prot.22818>

LENSINK, M. F.; WODAK, S. J. Docking, scoring, and affinity prediction in CAPRI. **Proteins: Structure, Function, and Bioinformatics**, v. 81, n. 12, p. 2082–2095, 2013. <https://doi.org/10.1002/prot.24428>

LYSKOV, S. et al. Serverification of Molecular Modeling Applications: The Rosetta Online Server That Includes Everyone (ROSIE). **PLoS ONE**, v. 8, n. 5, p. e93906, 2013. <https://doi.org/10.1371/journal.pone.0063906>

LYSKOV, S.; GRAY, J. J. The RosettaDock server for local protein-protein docking. **Nucleic Acids Research**, v. 36, n. Web Server, p. W233–W238, 2008. <https://doi.org/10.1093/nar/gkn216>

MASHIACH, E. et al. FireDock: a web server for fast interaction refinement in molecular docking. **Nucleic acids research**, v. 36, n. suppl_2, p. W229–W232, 2008. <https://doi.org/10.1093/nar/gkn186>

MASHIACH, E. et al. An integrated suite of fast docking algorithms. **Proteins: Structure, Function, and Bioinformatics**, v. 78, n. 15, p. 3197–3204, 2010. <https://doi.org/10.1002/prot.22790>

MORETTI, R. et al. Web-accessible molecular modeling with Rosetta: The Rosetta Online Server that Includes Everyone (ROSIE). **Protein Science**, v. 27, n. 1, p. 259–268, 2018. <https://doi.org/10.1002/pro.3313>

NGUYEN, M. N. et al. The interfacial character of antibody paratopes: Analysis of antibody-antigen structures. **Bioinformatics**, v. 33, n. 19, p. 2971–2976, 2017. <https://doi.org/10.1093/bioinformatics/btx389>

- NGUYEN, M. N. et al. Discovering Putative Protein Targets of Small Molecules: A Study of the p53 Activator Nutlin. **Journal of Chemical Information and Modeling**, v. 59, n. 4, p. 1529–1546, 2019. <https://doi.org/10.1021/acs.jcim.8b00762>
- NGUYEN, M. N.; VERMA, C. S.; ZHONG, P. AppA: a web server for analysis, comparison, and visualization of contact residues and interfacial waters of antibody-antigen structures and models. **Nucleic acids research**, v. 47, n. W1, p. W482–W489, 2019. <https://doi.org/10.1093/nar/gkz358>
- PEDOTTI, M. et al. Computational Docking of Antibody-Antigen Complexes, Opportunities and Pitfalls Illustrated by Influenza Hemagglutinin. **International Journal of Molecular Sciences**, v. 12, n. 1, p. 226–251, 2011. <https://doi.org/10.3390/ijms12010226>
- PIERCE, B.; WENG, Z. A combination of rescoring and refinement significantly improves protein docking performance. **Proteins: Structure, Function, and Bioinformatics**, v. 72, n. 1, p. 270–279, 2008. <https://doi.org/10.1002/prot.21920>
- PIOVESAN, D.; MINERVINI, G.; TOSATTO, S. C. E. The RING 2.0 web server for high quality residue interaction networks. **Nucleic Acids Research**, v. 44, n. W1, p. W367–W374, 2016. <https://doi.org/10.1093/nar/gkw315>
- RINI, J.; SCHULZE-GAHMEN, U.; WILSON, I. Structural evidence for induced fit as a mechanism for antibody-antigen recognition. **Science**, v. 255, n. 5047, p. 959–965, 1992. <https://doi.org/10.1126/science.1546293>
- RITCHIE, D. W.; KEMP, G. J. L. Protein docking using spherical polar Fourier correlations. **Proteins: Structure, Function, and Genetics**, v. 39, n. 2, p. 178–194, 2000. [https://doi.org/10.1002/\(SICI\)1097-0134\(20000501\)39:2<178::AID-PROT8>3.0.CO;2-6](https://doi.org/10.1002/(SICI)1097-0134(20000501)39:2<178::AID-PROT8>3.0.CO;2-6)
- SCHNEIDMAN-DUHOVNY, D. et al. PatchDock and SymmDock: Servers for rigid and symmetric docking. **Nucleic Acids Research**, v. 33, n. suppl_2, p. W363–W367, 2005. <https://doi.org/10.1093/nar/gki481>
- SHORTLE, D.; SIMONS, K. T.; BAKER, D. Clustering of low-energy conformations near the native structures of small proteins. **Proceedings of the National Academy of Sciences**, v. 95, n. 19, p. 11158–11162, 1998. <https://doi.org/10.1073/pnas.95.19.11158>
- SILVA, L. A. **Estudo da interação do complexo antiNS1 / NS1-DENV2 usando ancoramento molecular e métodos quânticos**. Dissertação (Mestrado em Química), Programa de Pós-Graduação em Química, Universidade Federal de Uberlândia. Uberlândia, 58 p., 2020.
- SIMONELLI, L. et al. Rapid Structural Characterization of Human Antibody–Antigen Complexes through Experimentally Validated Computational Docking. **Journal of Molecular Biology**, v. 396, n. 5, p. 1491–1507, 2010. <https://doi.org/10.1016/j.jmb.2009.12.053>
- SIMONELLI, L. et al. Rational Engineering of a Human Anti-Dengue Antibody through Experimentally Validated Computational Docking. **PLoS ONE**, v. 8, n. 2, p. e55561, 2013. <https://doi.org/10.1371/journal.pone.0055561>
- STOCKMAN, G. Object recognition and localization via pose clustering. **Computer Vision, Graphics, and Image Processing**, v. 40, n. 3, p. 361–387, 1987.

[https://doi.org/10.1016/S0734-189X\(87\)80147-0](https://doi.org/10.1016/S0734-189X(87)80147-0)

TOVCHIGRECHKO, A.; VAKSER, I. A. GRAMM-X public web server for protein-protein docking. **Nucleic Acids Research**, v. 34, n. Web Server, p. W310–W314, 2006.

<https://doi.org/10.1093/nar/gkl206>

VAJDA, S. et al. New additions to the ClusPro server motivated by CAPRI. **Proteins: Structure, Function, and Bioinformatics**, v. 85, n. 3, p. 435–444, 2017.

<https://doi.org/10.1002/prot.25219>

WANG, D. et al. A mutation identified in neonatal microcephaly destabilizes zika virus NS1 assembly in vitro. **Scientific Reports**, v. 7, n. 42580, p. 1–10, 2017a.

<https://doi.org/10.1038/srep42580>

WANG, J. et al. A Human Bi-specific Antibody against Zika Virus with High Therapeutic Potential. **Cell**, v. 171, n. 1, p. 229- 241.e15, 2017b. <https://doi.org/10.1016/j.cell.2017.09.002>

WOLFSON, H. J.; RIGOUTSOS, I. Geometric hashing: an overview. **IEEE Computational Science and Engineering**, v. 4, n. 4, p. 10–21, 1997. <https://doi.org/10.1109/99.641604>

XU, X. et al. Contribution of intertwined loop to membrane association revealed by Zika virus full-length NS 1 structure . **The EMBO Journal**, v. 35, n. 20, p. 2170–2178, 2016.

<https://doi.org/10.15252/embj.201695290>

CÁLCULO DE PROPRIEDADES TERMODINÂMICAS DOS COMPLEXOS NS1/ANTI-NS1

4.1 INTRODUÇÃO

Os fundamentos da mecânica estatística são oriundos de uma série de experimentos realizados no final do século XIX, mas que se consolidaram no século XX, sendo a mecânica quântica e seus métodos considerados um dos grandes avanços da ciência nos últimos tempos. Suas teorias fornecem uma descrição matemática da estrutura molecular de um sistema em termos das interações entre os núcleos atômicos e a distribuição eletrônica ao seu redor, levando em consideração dados sobre a estrutura eletrônica, baseando-se no formalismo de orbitais moleculares com diferentes abordagens (LEVINE, 1991).

Dentre as vantagens dos métodos de mecânica quântica destaca-se o rigor metodológico e, por consequência, a precisão dos resultados na descrição de propriedades químicas e físicas dos sistemas em estudo. Para tanto, tais métodos demandam um maior tempo e capacidade de memória computacional.

4.1.1 Mecânica Quântica – O Método Hartree-Fock

Fundamentalmente, para modelar um sistema, os métodos da Mecânica Quântica envolvem a resolução da equação de autovalores do operador hamiltoniano do sistema composto por núcleos e elétrons descritos, respectivamente, pelos vetores de posição R_A e r_i , conhecida como Equação de Schrödinger:

$$H\Psi = E\Psi \quad (4.1)$$

em que E é a energia eletrônica em unidades atômicas, Ψ é a função de onda ou, ainda, chamada de função de coordenadas dos elétrons e H é o operador hamiltoniano, uma expressão matemática que relaciona a energia cinética dos elétrons e a energia potencial do sistema. O operador hamiltoniano genérico para um sistema multieletrônico tem suas unidades atômicas dadas por:

$$H = -\frac{1}{2} \sum_i^{\text{elétrons}} \nabla_i^2 - \frac{1}{2} \sum_A^{\text{núcleo}} \frac{1}{M_A} \nabla_A^2 - \sum_i^{\text{elétrons}} \sum_A^{\text{núcleo}} \frac{Z_A}{r_{iA}} + \sum_{i<j}^{\text{elétrons}} \frac{1}{r_{ij}} + \sum_{A<B}^{\text{núcleo}} \frac{Z_A Z_B}{R_{AB}} \quad (4.2)$$

em que, os termos correspondem, na sequência em que são descritos, às energias cinética eletrônica, cinética nuclear, potencial elétron-núcleo, potencial entre elétrons e potencial entre núcleos. Na equação, Z é a carga nuclear, M_A é a razão entre a massa do núcleo A e a massa do elétron, R_{AB} é a distância entre o núcleo A e o núcleo B, r_{ij} é a distância entre os elétrons i e j e r_{iA} é a distância entre o elétron i e o núcleo A.

Dessa forma, ao solucionar a equação de Schrödinger obtém-se uma função que possui todas as informações referentes ao sistema podendo, portanto, serem calculadas inúmeras propriedades da molécula, inclusive as eletrônicas (HEHRE, 2003; LEVINE, 2000).

A equação 4.1 não possui solução analítica, para sistemas de mais de um elétron, sendo necessária a aplicação de aproximações e simplificações para a descrição de sistemas moleculares. A primeira delas, chamada aproximação de Born-Oppenheimer considera que o núcleo possui massa muito superior aos elétrons, da ordem de 1000 vezes, e, por esse motivo, esses se movimentam mais rápida e independentemente, próximos da velocidade da luz. Além disso, em virtude de suas cargas opostas há uma força de atração mútua entre núcleo e elétrons. Como a magnitude da aceleração é inversamente proporcional à massa (Eq. 2.1), a aceleração dos elétrons é grande e dos núcleos atômicos pequena. Por consequência, os elétrons se movem e respondem às forças muito rapidamente em detrimento dos núcleos. Nesse sentido, uma aproximação razoável é descrever os estados eletrônicos de uma molécula considerando os núcleos estacionários. Nesse sentido, é possível separar o hamiltoniano nuclear do hamiltoniano eletrônico (Eq. 4.4) para resolver a Equação de Schrodinger eletrônica (Eq. 4.3) (DEMTRÖDER, 2010; LEVINE, 2000):

$$H_{\text{eletrônico}} \Psi_{\text{eletrônica}} = E_{\text{eletrônica}} \Psi_{\text{eletrônica}} \quad (4.3)$$

$$H_{eletr\hat{on}ico} = -\frac{1}{2} \sum_i^{\text{el\hat{e}trons}} \nabla_i^2 - \sum_i^{\text{el\hat{e}trons}} \sum_A^{\text{n\hat{u}cleo}} \frac{Z_A}{r_{iA}} + \sum_{i<j}^{\text{el\hat{e}trons}} \frac{1}{r_{ij}} \quad (4.4)$$

nas quais a $\Psi_{eletr\hat{on}ica}$ deve ser assim\hat{e}trica e dependente do conjunto de coordenadas eletr\hat{on}icas (3 eletr\hat{on}icas e 1 de spin) e $H_{eletr\hat{on}ico}$ descreve o movimento dos el\hat{e}trons do sistema definidos em uma regi\hat{a}o com N part\hat{ic}ulas fixas e carregadas eletricamente. Nesse sentido, os el\hat{e}trons podem ser representados por uma fun\c{c}o de onda, denominada orbital molecular (MO). Essa fun\c{c}o de onda \hat{e} assumida levando-se em conta que o movimento dos el\hat{e}trons \hat{e} independente entre si e que esses est\hat{a}o dentro de um campo m\hat{e}dio (v^{HF}) e resultante da presen\c{c}a dos demais (TRSIC; PINTO, 2009).

De acordo com essa aproxima\c{c}o o modo mais simples de se escrever a fun\c{c}o de onda eletr\hat{on}ica antiassim\hat{e}trica, garantindo assim o princ\hat{ip}io de exclus\hat{a}o de Pauli, \hat{e} utilizando o determinante de Slater (Eq. 4.5):

$$\psi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_i(x_1) & \chi_j(x_1) & \cdots & \chi_k(x_1) \\ \chi_i(x_2) & \chi_j(x_2) & \cdots & \chi_k(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_i(x_N) & \chi_j(x_N) & \cdots & \chi_k(x_N) \end{vmatrix} \quad (4.5)$$

na qual $\frac{1}{\sqrt{N!}}$ \hat{e} a constante de normaliza\c{c}o dependente dos N el\hat{e}trons do sistema e χ s\hat{a}o as fun\c{c}oes ortonormais que representam os spin-orbitais para cada um dos el\hat{e}trons (ZERNER, 1991).

Nesse ponto, \hat{e} necess\hat{a}rio que sejam determinadas as fun\c{c}oes χ do determinante de Slater, de modo que, esse forne\c{c}a a aproxima\c{c}o mais adequada para o estado fundamental do sistema. Isso porque o melhor conjunto de spin-orbitais \hat{e} o respons\hat{a}vel pela minimiza\c{c}o da energia, conforme preconiza o Princ\hat{ip}io Variacional. Esse conjunto \hat{e} determinado pelo m\hat{e}todo conhecido como campo auto-consistente (SCF, do ingl\hat{e}s, *self-consistent field*), sendo este o princ\hat{ip}io do m\hat{e}todo de Hartree-Fock. Esse m\hat{e}todo soluciona a equa\c{c}o de autovalores de Hartree-Fock (Eq. 4.6) de modo iterativo por meio da aproxima\c{c}o da matriz de Fock (F) cujos elementos s\hat{a}o obtidos via integrais de troca que correspondem as diversas intera\c{c}oes entre os el\hat{e}trons que est\hat{a}o presentes nos \hat{a}tomos (ZERNER, 1991).

$$F \chi(x_i) = \varepsilon \chi(x_i) \quad (4.6)$$

Essa teoria é o ponto de partida dos métodos *ab initio* que buscam resolver problemas quânticos diretamente dos postulados da mecânica quântica, sem aproximações e sem correlação a dados experimentais (HOLLAUER, 2008). Portanto, o seu uso envolve a avaliação de uma grande quantidade dessas integrais cujo número cresce exponencialmente com o número de funções de bases chegando à ordem de milhões para pequenas moléculas, demandando tempo e memória computacional consideráveis.

Portanto, é necessário que se estabeleça um objetivo considerando a linha tênue existente entre a precisão do método (rigor teórico) e o custo computacional, em especial, em se tratando de sistemas envolvendo proteínas como é o caso do presente trabalho. Nesse contexto, os métodos semi-empíricos se apresentam como uma alternativa viável no estudo de sistemas grandes (de centenas até milhares de átomos) em função de sua elevada velocidade quando comparada aos cálculos DFT ou *ab initio*, demasiadamente custosos computacionalmente (LASCHUK, 2005; ZERNER, 1991).

4.1.2 Métodos Semi-Empíricos: PM7

Os métodos semi-empíricos introduziram uma série de parametrizações e aproximações para deduzir os elementos da matriz de Fock juntando parâmetros experimentais ao formalismo do SCF para os elétrons de valência, o que permite negligenciar algumas integrais de troca para os elétrons de caroço (DEWAR; THIEL, 1977; LEACH, 2001). Dentro desse conjunto de métodos os mais utilizados são aqueles baseados na aproximação ZDO (do inglês, *Zero Differential Overlap Approximation*) (POPLE; BEVERIDGE, 1970): Modelo Austin 1 (AM1) (DEWAR et al., 1985), Modelo Recife 1 (RM1) (ROCHA et al., 2006), Método Paramétrico 3 (PM3) (STEWART, 1989), 6 (PM6) (STEWART, 2007) e 7 (PM7) (STEWART, 2013). No entanto, vale considerar que não há unanimidade em relação às suas aplicações visto que

alguns métodos são parametrizados para calores de formação, geometrias moleculares e momentos de dipolo; outros são parametrizados para reproduzir os espectros de transições eletrônicas. Não existem métodos semi-empíricos de aplicabilidade geral capazes de reproduzir resultados úteis para uma gama ampla de propriedades (LASCHUK, 2005, p. 6).

Nesse sentido, o método PM7 foi selecionado para aplicação nesse trabalho pelos seguintes motivos: (1) teve seus parâmetros ajustados e testados para uma correta otimização geométrica e precisão no cálculo da entalpia de formação de proteínas (HOSTAŠ; ŘEZÁČ; HOBZA, 2013; STEWART, 2013); (2) os dados de calor de formação, obtidos por sua aplicação, já foram utilizados com sucesso para complementar a função de scoring no reconhecimento de estruturas de proteínas de caráter nativo, conforme proposto para o sistema em estudo (URQUIZA-CARVALHO; FRAGOSO; ROCHA, 2016) e (3) encontra-se adequadamente implementado no software de distribuição livre MOPAC 2016 (STEWART, 2016).

De modo geral, esse método foi parametrizado considerando que as grandezas energia de ligação total (E_L) é numericamente equivalente às entalpias de atomização à 298 K, respectivamente, expressas pelas equações 4.7 e 4.8:

$$E_L = E_{TOT} - \sum_n E_n \quad (4.7)$$

$$\Delta H_{atom.}^{298} = \Delta H_f^o - \sum_n \Delta H_f^{298}(n) \quad (4.8)$$

em que E_{TOT} é a energia molecular total, E_n as energias de cada um dos n átomos constituintes e $\Delta H_f^{298}(n)$ o calor de formação do átomo n a 298 K.

Combinando-se as equações 4.7 e 4.8 temos que o valor de formação é dado por:

$$\Delta H_f^o = E_{TOT} - \sum_n (E_n - \Delta H_f^{298}(n)) \quad (4.9)$$

Vale considerar ainda que o uso da abordagem de escalonamento linear MOZYME, ao introduzir a aplicação dos Orbitais Moleculares Localizados (OMLs, do inglês, *Localized*

Molecular Orbitals), auxilia na diminuição da demanda computacional para os cálculos das propriedades de interesse de grandes biomoléculas (STEWART, 1996).

4.2 METODOLOGIA

4.2.1 Cálculo da Entalpia de Ligação

Após a execução do protocolo proposto (Fig. 3.2) para a escolha dos modelos mais representativos, conforme apresentado na Seção 3.3, suas informações estruturais foram preparadas e então convertidas em um formato compatível com o programa MOPAC 2016 (STEWART, 2016) utilizado na realização dos cálculos de Mecânica Quântica para obtenção das médias das propriedades termodinâmicas dos complexos.

O pré-processamento dos arquivos de coordenadas das estruturas dos modelos em estudo consistiu na adição de hidrogênios realizada com a utilização do módulo *pdb2gmx* implementado no *software* GROMACS 2016.4, considerando a forma de protonação dos resíduos de aminoácidos no pH sanguíneo fisiológico: 7,4 (SOARES; FERREIRA, 2004).

De acordo com Stewart; D'Elia e Attwood (2014) uma etapa inicial de otimização foi realizada mantendo fixas as posições de todos os átomos (NOOPT), com exceção dos átomos de hidrogênio (OPT-H). Os cálculos foram realizados utilizando o modelo COSMO (KLAMT; SCHUURMANN, 1993) de solvatação implícita, considerando a constante dielétrica da água: 78,4. Todos os sistemas foram otimizados considerando o método PM7, a aproximação MOZYME e a norma do gradiente menor que 5.0 (GNORM = 5.0). Valor considerado rigoroso e, portanto, satisfatório para sistemas moleculares grandes como os proteicos.

Em seguida, para as estruturas tridimensionais obtidas na etapa anterior (PDBOUT) foram realizados cálculos HF para obtenção da Entalpia de Formação (H_f) considerando o método PM7, a aproximação MOZYME, solvatação implícita (EPS 78,4) e a norma do gradiente menor que 5.0. Adicionalmente, um cálculo de campo auto-consistente (1SCF) foi realizado para correção de erros no cálculo dessa propriedade termodinâmica, decorrentes da perda de ortogonalidade dos orbitais moleculares localizados, ocasionada pelo uso da abordagem Mozyme.

A entalpia de ligação dos complexos foi calculada conforme metodologia proposta por Nikitina e colaboradores (2004):

$$H_{lig} = H_{f(\text{complexo})} - (H_{f \text{ antígeno}} + H_{f \text{ anticorpo}}) \quad (4.10)$$

em que, H_f é a entalpia de formação de cada um dos componentes do complexo (antígeno ou anticorpo) ou desse como um todo.

4.2.2 Cálculo da Energia de Ligação

O cálculo de predição das energias de ligação dos modelos de complexos antígeno-anticorpo previamente selecionados foi realizado utilizando o software PRODIGY no módulo *protein-protein* (VANGONE; BONVIN, 2015; XUE et al., 2016). Esse software estabelece uma correlação entre força de interação e o número de contatos interfaciais (ICs) no complexo. De modo que, por definição, dois resíduos estão em contato de algum de seus átomos, com exceção dos hidrogênios, estão em uma distância de até 5,5 Å. O modelo de predição implementado é dado pela seguinte equação de regressão linear simples entre as ICs e algumas propriedades das superfícies que não interagem (NIS, do inglês, *non-interacting surfaces*) (KASTRITIS et al., 2014):

$$\begin{aligned} \Delta G_{\text{predição}} = & -0,09459I \frac{CS_{\text{carga}}}{\text{carga}} - 0,10007 ICs_{\frac{\text{carga}}{\text{apolar}}} + 0,19577 ICs_{\frac{\text{polar}}{\text{polar}}} \\ & - 0,22671 ICs_{\frac{\text{polar}}{\text{apolar}}} + 0,18681 \%NIS_{\text{apolar}} + 0,3810 \%NIS_{\text{carga}} - 15,9433 \end{aligned} \quad (4.11)$$

4.3 RESULTADOS E DISCUSSÃO

Na Tabela 4.1 estão apresentados os valores de entalpia de ligação de cada um dos modelos dos sistemas previamente selecionados e do modelo do redocking.

Tabela 4.1 – Valores de ΔH_{lig} calculados pelo método semi-empírico PM7.

| Sistema | Complexo | $\Delta H_{lig} (kJ mol^{-1})$ |
|----------------------------------|-------------|--------------------------------|
| Redocking –NS1-WNV/22-anti-NS1* | PDB ID 4OII | - 620,95 |
| 1 – NS1-DENV2/22-anti-NS1 (PDB)* | DC09 | - 460,07 |
| 2 – NS1- ZIKV /22-anti-NS1 (PDB) | ZC02 | - 739,24 |
| 2 – NS1- ZIKV /22-anti-NS1 (PDB) | ZC04 | - 716,20 |
| 2 – NS1- ZIKV /22-anti-NS1(PDB) | ZP09 | + 1.063,79 |
| 3 – NS1-DENV2/22-anti-NS1 (MD) | DC07 | - 707,27 |
| 4 – NS1- ZIKV /22-anti-NS1(MD) | ZC02 | - 655,84 |

Fonte: Dados do autor e (SILVA, 2020)*

Os dados de Silva (2020) mostraram concordância entre as análises da superfície de interação para os modelos preditos pelos *softwares* de Docking Molecular e, posteriormente, selecionados conforme critérios adotados no protocolo de reranking para o sistema 1 (NS1-DENV2/22-anti-NS1). De modo similar, para o sistema 2 (NS1-ZIKV/22-anti-NS1) é possível anotar que os resultados indicam a tendência de proporcionalidade entre maior número de ICs pertencentes a epítopos (Tabela 3.2) e estabilidade termodinâmica, sob o ponto de vista entálpico, confirmando as conclusões parciais previamente apresentadas. Portanto, os modelos DC07 e ZC02 foram selecionados, respectivamente, como representativos para os sistemas 3 e 4 seguindo o protocolo proposto.

Além disso, foi possível verificar que os valores de ΔH_{lig} obtidos para os modelos DC09 (sistema 1), ZC02 (sistema 2), DC07 (sistema 3) e ZC02 (sistema 4) são comparáveis, ou seja, da mesma ordem de grandeza do valor obtido para a estrutura cristalográfica PDB ID 4OII. Isso indica que esses modelos são candidatos em potencial para, respectivamente, os complexos NS1-DENV2/22-anti-NS1 e NS1-ZIKV/22-anti-NS1 em sua conformação nativa.

Portanto, os modelos cristalográficos DC09 e ZC02 indicam interação das proteínas NS1 da Dengue e Zika, única e respectivamente, nos domínios wing e β -ladder. Bem como os modelos obtidos da amostragem da MD, DC07 e ZC02, indicam, respectivamente, o domínio wing e a interface entre os domínios wing e β -ladder como regiões de interação das proteínas NS1 com as CDRs do anticorpo. Essas diferenças corroboram os distintos valores de ΔH_{lig} . Portanto, os resultados apontam que as proteínas interagem de modo distinto com as

CDRs do anticorpo vislumbrando a viabilidade de aplicação dos sistemas no diagnóstico diferencial das flavirozes.

Considerando a aplicação proposta nesse trabalho, um biossensor de afinidade envolvendo complexos antígeno-anticorpo (FURTADO et al., 2008), a avaliação da energia livre de ligação (ΔG_{lig}) permite complementar as análises anteriores e mensurar a força de ligação entre antígeno e anticorpo no complexo. Isso porque essas forças de interação/adesão específicas regem o processo de reconhecimento do analito pelo elemento biológico sensível em um dispositivo de AFM enquanto biossensor (FRANCA; AMARANTE; LEITE, 2010). Segundo (ABBAS; LICHTMAN; PILLAI, 2011, p. 102)

a força da ligação entre um único sítio de combinação de um anticorpo e um epítipo do antígeno é chamada afinidade do anticorpo. A afinidade é comumente representada pela constante de dissociação (K_d), que indica a facilidade de separação do complexo antígeno-anticorpo em seus constituintes. Uma K_d baixa indica maior afinidade de interação, uma vez que uma menor concentração de antígeno ou anticorpo é necessária à formação do complexo.

A relação matemática entre as grandezas ΔG_{lig} e K_d pode ser adequadamente descrita pela Equação 4.12 (GILSON et al., 1997; WOO; ROUX, 2005):

$$\Delta G_{lig} = RT \ln K_d \quad (4.12)$$

em que R é a constante ideal dos gases ($kcal K^{-1} mol^{-1}$) e T a temperatura (K).

A constante de associação (K_a) também é uma medida da afinidade da interação do complexo antígeno-anticorpo no equilíbrio e está intrinsecamente relacionada com a constante de dissociação pois é medida pelo inverso dessa grandeza (OLIVEIRA; RIBEIRO; VIZZONI, 2013).

Tabela 4.2 – Valores de ΔG_{lig} e constantes de dissociação (K_d) e associação (K_a), a 298,15 K, para os modelos preditos pelo protocolo proposto.

| Sistema | Complexo | $\Delta G_{lig} (kcal\ mol^{-1})$ | $K_d (M)$ | $K_a (M^{-1})$ |
|----------------------------------|-------------|-----------------------------------|-----------------------|-------------------|
| Redocking NS1-WNV/22-anti-NS1 | PDB ID 4OII | - 12,7 | $4,9 \times 10^{-10}$ | $2,0 \times 10^9$ |
| 1 – NS1-DENV2/22-anti-NS1 (PDB) | DC09 | - 11,6 | $3,1 \times 10^{-9}$ | $3,2 \times 10^8$ |
| 2 – NS1- ZIKV /22-anti-NS1 (PDB) | ZC02 | - 13,5 | $1,3 \times 10^{-10}$ | $7,7 \times 10^9$ |
| 3 – NS1-DENV2/22-anti-NS1 (MD) | DC07 | - 12,4 | $8,7 \times 10^{-10}$ | $1,1 \times 10^9$ |
| 4 – NS1- ZIKV /22-anti-NS1 (MD) | ZC02 | - 9,7 | $7,4 \times 10^{-8}$ | $1,4 \times 10^7$ |

Fonte: Autor

Os valores de ΔG_{lig} obtidos para os complexos estão dentro ou próximos da faixa geral para complexos antígeno-anticorpo que varia entre -6,5 e -13,0 kcal mol⁻¹ (CHUANG et al., 2008). Esses valores quando convertidos em K_d , como medida de afinidade dos complexos, fornecem resultados da ordem de 10⁻⁸ a 10⁻¹⁰. De acordo com Kastiris et al. (2011), a afinidade para complexos antígeno-anticorpo segue a seguinte escala: $K_d < 10^{-10}$ M (alta afinidade), $10^{-6} < K_d < 10^{-10}$ M (média afinidade) e $K_d > 10^{-6}$ M (baixa afinidade). Essas diferenças podem ser mais adequadamente mensuráveis quando analisados os valores de K_a cujos valores maiores indicam maior afinidade do complexo em se tratando da ligação. Portanto, os sistemas em estudo apresentam afinidade alta/média indicando especificidade, ou seja, mesmo a baixas concentrações de antígeno esse é reconhecido pelo anticorpo. Vale salientar que os estudos de predição teórica das interações entre complexos, sejam enzima-inibidor e antígeno-anticorpo, conduzidos pelo presente grupo de estudo e citados no Capítulo 1, mostram boa concordância com resultados experimentais de AFM. Esses indicam que interações intermoleculares específicas ou não-específicas são distinguíveis entre si em eventos de reconhecimento molecular de um analito por seu inibidor. Desse modo, essas diferenças na magnitude das afinidades observadas para os sistemas dão indicativos de especificidade dos sistemas em questão.

Considerando os sistemas 1 e 2, DC09 e ZC02, é possível observar que há uma diferença significativa (aproximadamente 25 vezes) na afinidade das proteínas NS1 da DENV2 e da ZIKV pelo anticorpo em questão. Portanto, o reconhecimento das proteínas NS1 pelo anticorpo ocorre

em escalas diferentes sugerindo forças de adesão distinguíveis entre si. A diferença entre as afinidades observadas para esses sistemas se deve a dois fatores:

- i) Número de contatos interfaciais entre resíduos carregados é maior entre a proteína NS1-ZIKV e o anticorpo em comparação a NS1-DENV2, que por sua vez possui um maior número de contatos entre resíduos polares na superfície de contato. Logo, as interações coulômbicas são mais fortes e significativas na formação do complexo em detrimento das interações do tipo van der Waals;
- ii) Similaridade significativa entre as ICs da proteína NS1-ZKV e NS1-WNV e o anticorpo 22-anti-NS1 conforme alinhamento observado na Figura 4.1. Apenas o resíduo polar Asn27 da Cadeia Leve 1 (CL1) diferencia a contribuição das CDRs na formação dos complexos.

Figura 4.1 – Alinhamento das sequências das CDRs do anticorpo identificando os resíduos de aminoácidos pertencem às ICs com as proteínas NS1 das Flaviviruses WNV (PDB ID 4OII), DENV2 (DC09) e ZIKV (ZC02).

| | CL1 | CL2 | CL3 | CH1 | CH2 | CH3 |
|-------|----------|-------------|----------|-----------|----------------|-------------|
| WNV | GNIHNYLA | LLVYNAKTLAD | QHFWSTPR | YTFTSYWMH | WIGDINPNNGGPSY | TIDDDGYRFGY |
| DENV2 | GNIHNYLA | LLVYNAKTLAD | QHFWSTPR | YTFTSYWMH | WIGDINPNNGGPSY | TIDDDGYRFGY |
| ZIKV | GNIHNYLA | LLVYNAKTLAD | QHFWSTPR | YTFTSYWMH | WIGDINPNNGGPSY | TIDDDGYRFGY |

Fonte: Autor

Para os sistemas 3 e 4, observa-se diferença ainda mais significativa (aproximadamente 85 vezes) na afinidade dos complexos DC07 e ZC02. No entanto, uma tendência inversa é observada em relação ao anotado para os complexos cristalográficos DC09 e ZC02, visto que nesse caso, a proteína NS1-DENV2 apresentou maior afinidade com o anticorpo. Apesar de ainda sugerir forças de adesão distinguíveis entre si, esse resultado reflete uma das limitações de cálculos e/ou protocolos de Docking Molecular de corpo rígido. Movimentos locais e significativos da cadeia principal acontecem na interface de interação e os complexos antígeno-anticorpo, em geral, exibem grandes movimentos e/ou transições desordenadas (KASTRITIS et al., 2011).

Nesse caso, essa divergência deve estar associada ao aumento na área acessível ao solvente da proteína NS1-ZIKV que ocasionou mudanças estruturais mais significativas, no domínio de interação β -ladder, em relação à estrutura cristalográfica conforme resultados do Capítulo 2.

Desse modo, considera-se que o protocolo de Docking Molecular proposto, ao utilizar os *webservers* ClusPro e PatchDock, minimiza a contribuição dos efeitos entrópicos na busca global. Isso porque, apesar de suportarem diferenças moderadas entre as formas ligada e não-ligada das estruturas envolvidas nos complexos, não foram capazes de reproduzir com elevada precisão as mudanças estruturais das proteínas sob efeito de solvatação na formação dos complexos. Assim sendo, sugere-se que o protocolo proposto tem resultados mais satisfatórios ao utilizar arquivos de entrada (*input*) de natureza cristalográfica, em função da maior organização estrutural das cadeias principal e lateral das proteínas receptora e ligante.

Além disso, de acordo com Kastiris et al. (2011) os valores de K_d , por serem válidos apenas dentro de uma faixa pequena de temperatura, força iônica e pH, não devem ser tomados como absolutos, mas sim para auxiliar a compreender a ordem de magnitude da interação. Nesse sentido, cálculos teóricos de outra natureza, como Potencial de Força Média obtido a partir da amostragem estatística de cálculos *umbrella sampling*, devem ser aplicados para emular experimentos de AFM e prever a força de ligação propriamente dita dessas interações específicas.

Logo, mediante a verificação da seletividade e da especificidade dos sistemas em estudo sugere-se a viabilidade de um dispositivo de sensoriamento do tipo AFM com aplicabilidade no diagnóstico diferencial entre as Flavirose Dengue e Zika, utilizando a proteína NS1 como marcador antigênico.

4.4 CONSIDERAÇÕES FINAIS

A entalpia de ligação (ΔH_{lig}), foi utilizada para confirmar a escolha do modelo predito pelos servidores de Docking Molecular Proteína-Proteína, mais representativo, para os sistemas envolvendo as proteínas NS1 das flavirose Dengue e Zika. As propriedades termodinâmicas corroboram as análises prévias das superfícies de interação dos complexos, indicando que os

modelos DC09 (sistema 1), ZC02 (sistema 2), DC07 (sistema 3) e ZC02 (sistema 4) possuem caráter nativo. Nesse sentido, o protocolo proposto possui ineditismo ao mostrar que a associação de técnicas computacionais de diferentes abordagens aplicadas no estudo de sistemas, dos quais se possui um conhecimento preliminar, é capaz de fornecer para este uma descrição, a nível molecular, mais precisa.

Os modelos selecionados indicam que a interação do anticorpo com as proteínas NS1 de ambas flavirozes ocorrem em domínios diferentes e, portanto, a interação é seletiva e essa característica reforça a hipótese do uso do sistema para um diagnóstico diferencial.

As propriedades termodinâmicas para o complexo ZC02 (sistema 2) apresentam maior similaridade com as do complexo nativo de WNV. O resíduo de a.a. Asn27 da CL1 desse modelo, é o responsável pelas diferenças nas propriedades anotadas entre esse complexo e o DC09. Portanto, pode ser um indicativo no *design* racional de anticorpos mais específicos, para aplicação no desenvolvimento de vacina.

As constantes de dissociação (K_d), para os modelos DC09 (sistema 1) e ZC02 (sistema 2) estão em ordens de grandeza diferentes (10^{-9} e 10^{-10} M) sugerindo forças de adesão distinguíveis entre si, com maior afinidade da NS1-ZIKV pelo anticorpo em detrimento da NS1-DENV2.

O protocolo aqui detalhado e utilizado na modelagem molecular desses sistemas apresentou desempenho mais adequado para arquivos *input* cristalográficos do receptor e ligante. Isso em virtude das limitações de ponderação das mudanças conformacionais nas interfaces de interação em cálculos de docking rígido. Ainda assim, indica potencialidades descritivas para serem aplicadas em sistemas de caráter similar e nos mecanismos envolvidos em tais doenças.

Portanto, verificados os critérios de seletividade e especificidade, há indicativo da viabilidade de uso dos sistemas em estudo no direcionamento do desenvolvimento de imunossensores para diagnóstico diferencial e precoce de flavirozes.

Nesse sentido, os cálculos computacionais utilizados ao longo do estudo desses sistemas mostraram-se estar em consonância com informações da literatura obtidas de estudos *in silico* e *in vitro* sobre as proteínas NS1 das viroses Dengue e Zika. Portanto, a modelagem e simulação podem subsidiar consideravelmente pesquisas futuras nesta área, em particular, na redução dos custos e uso de modelos animais.

4.5 TRABALHOS FUTUROS

- Realizar cálculos de Dinâmica Molecular Direcional (SMD) para calcular o potencial de força média para os complexos selecionados nesse trabalho;
- Aprimorar o protocolo proposto (análise do *top* 30) para que os efeitos entrópicos sejam, em associação com os fatores entálpicos, mais adequadamente ponderados na seleção de complexos de caráter nativo visualizando a aplicação em diferentes sistemas;
- Aplicar protocolos de mutagênese computacional a fim de descrever estruturalmente anticorpos mais específicos para as flaviroses.

REFERÊNCIAS

- ABBAS, A. K.; LICHTMAN, A. H.; PILLAI, S. Anticorpos e Antígenos. In: **Imunologia Celular e Molecular**. 7. ed. Rio de Janeiro: Elsevier, 2011. p. 89–108.
- CHUANG, G.-Y. et al. DARS (Decoys As the Reference State) Potentials for Protein-Protein Docking. **Biophysical Journal**, v. 95, n. 9, p. 4217–4227, 2008.
<https://doi.org/10.1529/biophysj.108.135814>
- CUSTODIO, R. et al. Quatro alternativas para resolver a equação de Schrödinger para o átomo de hidrogênio. **Química Nova**, v. 25, n. 1, p. 159–170, 2002.
<https://doi.org/10.1590/S0100-40422002000100025>
- DEMTRÖDER, W. Basic Concepts of Quantum Mechanics. In: **Atoms, Molecules and Photons. Graduate Texts in Physics**. Berlin: Springer, 2010. p. 129–157.
https://doi.org/10.1007/978-3-642-10298-1_4
- DEWAR, M. J. S. et al. Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model. **Journal of the American Chemical Society**, v. 107, n. 13, p. 3902–3909, 1985.
<https://doi.org/10.1021/ja00299a024>
- DEWAR, M. J. S.; THIEL, W. Ground states of molecules. 38. The MNDO method. Approximations and parameters. **Journal of the American Chemical Society**, v. 99, n. 15, p. 4899–4907, 1977. <https://doi.org/10.1021/ja00457a004>
- FRANCA, E. F.; AMARANTE, A. M.; LEITE, F. L. Introduction to Atomic Force Microscopy Simulation. In: MÉNDEZ-VILAS, A.; DÍAZ, J. (Eds.). **Microscopy: Science, Technology, Applications and Education**. 4. ed. Badajoz: Formatex Research Center, 2010. p. 1338–1349.
- FURTADO, R. F. et al. **Aplicações de biossensores na análise da qualidade de alimentos**. 1. ed. Fortaleza: Embrapa Agroindústria Tropical, 2008.
- GILSON, M. K. et al. The statistical-thermodynamic basis for computation of binding affinities: a critical review. **Biophysical Journal**, v. 72, n. 3, p. 1047–1069, 1997.
[https://doi.org/10.1016/S0006-3495\(97\)78756-3](https://doi.org/10.1016/S0006-3495(97)78756-3)
- HEHRE, W. J. **A Guide to Molecular Mechanics and Quantum Chemical Calculations**. Irvine: WaveFunction, 2003.
- HOLLAUER, E. **Química Quântica**. Rio de Janeiro: LTC, 2008.
- HOSTAŠ, J.; ŘEZÁČ, J.; HOBZA, P. On the performance of the semiempirical quantum mechanical PM6 and PM7 methods for noncovalent interactions. **Chemical Physics Letters**, v. 568–569, p. 161–166, 2013. <https://doi.org/10.1016/j.cplett.2013.02.069>
- KASTRITIS, P. L. et al. A structure-based benchmark for protein-protein binding affinity. **Protein Science**, v. 20, n. 3, p. 482–491, 2011.
- KASTRITIS, P. L. et al. Proteins Feel More Than They See: Fine-Tuning of Binding Affinity by Properties of the Non-Interacting Surface. **Journal of Molecular Biology**, v. 426, n. 14, p. 2632–2652, 2014. <https://doi.org/10.1016/j.jmb.2014.04.017>
- KLAMT, A.; SCHUURMANN, G. COSMO : A New Approach to Dielectric Screening in

Solvents with Explicit. **Journal of Chemical Society Perkin Transactions 2**, n. 5, p. 799–805, 1993. <https://doi.org/10.1039/P29930000799>

LASCHUK, E. F. **Novo formalismo semi-empírico para cálculos químico-quânticos**. Tese (Doutorado em Química) – Programa de Pós-Graduação em Química, Universidade Federal do Rio Grande do Sul. Porto Alegre, 143 p., 2005.

LEACH, A. R. **Molecular Modelling: Principles and Applications**. 2. ed. New York: Pearson Prentice Hall, 2001.

LEVINE, I. N. **Quantum Chemistry**. 4. ed. New York: Pearson Prentice Hall, 1991.

LEVINE, I. N. **Quantum Chemistry**. 5. ed. New York: Pearson Prentice Hall, 2000.

NIKITINA, E. et al. Semiempirical Calculations of Binding Enthalpy for Protein-Ligand Complexes. **International Journal of Quantum Chemistry**, v. 97, n. 2, p. 747–763, 2004. <https://doi.org/10.1002/qua.10778>

OLIVEIRA, M. B. S. de C.; RIBEIRO, F. C.; VIZZONI, A. G (Org.). **Conceitos Básicos e Plificados em imuno-hematologia**. Rio de Janeiro: EPSJV, 2013.

POPLE, J.; BEVERIDGE, D. L. **Métodos semiempíricos de química quântica no ensino da entalpia das reações químicas**. New York: McGraw-Hill, 1970.

SILVA, L. A. **Estudo da interação do complexo antiNS1 / NS1-DENV2 usando ancoramento molecular e métodos quânticos**. Dissertação (Mestrado em Química), Universidade Federal de Uberlândia. Uberlândia, 58 p., 2020.

SOARES, T. A.; FERREIRA, R. Aplicação da equação de Poisson-Boltzmann ao cálculo de propriedades dependentes do pH em proteínas. **Química Nova**, v. 27, n. 4, p. 640–647, 2004. <https://doi.org/10.1590/S0100-40422004000400019>

STEWART, J. J. P. Optimization of parameters for semiempirical methods I. Method. **Journal of Computational Chemistry**, v. 10, n. 2, p. 209–220, 1989. <https://doi.org/10.1002/jcc.540100208>

STEWART, J. J. P. Application of localized molecular orbitals to the solution of semiempirical self-consistent field equations. **International Journal of Quantum Chemistry**, v. 58, n. 2, p. 133–146, 1996. [https://doi.org/10.1002/\(SICI\)1097-461X\(1996\)58:2<133::AID-QUA2>3.0.CO;2-Z](https://doi.org/10.1002/(SICI)1097-461X(1996)58:2<133::AID-QUA2>3.0.CO;2-Z)

STEWART, J. J. P. Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements. **Journal of Molecular Modeling**, v. 13, n. 12, p. 1173–1213, 2007. <https://doi.org/10.1007/s00894-007-0233-4>

STEWART, J. J. P. Optimization of parameters for semiempirical methods VI: More modifications to the NDDO approximations and re-optimization of parameters. **Journal of Molecular Modeling**, v. 19, n. 1, p. 1–32, 2013. <https://doi.org/10.1007/s00894-012-1667-x>

STEWART, J. J. P. **Stewart computational chemistry MOPAC2016**. Colorado, Colorado Springs, 2016. Disponível em: <<http://openmopac.net/>>.

STEWART, J. J. P.; D'ELIA, D.; ATTWOOD, T. **A quick guide to MOPAC**. Disponível em: <https://www.embnet.org/shared/quickguides/04-MOPAC-en_UK.pdf>. Acesso em: 2

abr. 2020.

TRSIC, M.; PINTO, M. F. S. **Química Quântica - Fundamentos e Aplicações**. Barueri: Manole, 2009.

URQUIZA-CARVALHO, G. A.; FRAGOSO, W. D.; ROCHA, G. B. Assessment of semiempirical enthalpy of formation in solution as an effective energy function to discriminate native-like structures in protein decoy sets. **Journal of Computational Chemistry**, v. 37, n. 21, p. 1962–1972, 2016. <https://doi.org/10.1002/jcc.24415>

VANGONE, A.; BONVIN, A. M. J. J. Contacts-based prediction of binding affinity in protein–protein complexes. **eLife**, v. 4, 2015. <https://doi.org/10.7554/eLife.07454>

WOO, H.-J.; ROUX, B. Calculation of absolute protein-ligand binding free energy from computer simulations. **Proceedings of the National Academy of Sciences**, v. 102, n. 19, p. 6825–6830, 2005. <https://doi.org/10.1073/pnas.0409005102>

XUE, L. C. et al. PRODIGY: a web server for predicting the binding affinity of protein–protein complexes. **Bioinformatics**, p. btw514, 2016. <https://doi.org/10.1093/bioinformatics/btw514>

ZERNER, M. C. Semiempirical Molecular Orbital Methods. In: LIPKOWITZ, K. B.; BOYD, D. B. (Eds.). **Reviews in Computational Chemistry**. 1. ed. Nova Jersey: John Wiley & Sons, Ltda, 1991. p. 45–81.