

Letícia Alves da Silva

**Estudo da interação do complexo
antiNS1/NS1-DENV2 usando ancoramento
molecular e métodos quânticos
semi-empíricos**

Uberlândia

2020

Letícia Alves da Silva

**Estudo da interação do complexo
antiNS1/NS1-DENV2 usando ancoramento molecular
e métodos quânticos semi-empíricos**

Dissertação apresentada ao Programa de Pós-Graduação em Química do Instituto de Química da Universidade Federal de Uberlândia, como requisito para obtenção do Título de Mestre em Química.

Área de concentração: Físico-Química

Orientador: Prof. Dr. Eduardo de Faria Franca

Uberlândia
2020

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

S586 Silva, Letícia Alves da, 1995-
2020 Estudo da interação do complexo antiNS1/NS1-DENV2 usando
ancoramento molecular e métodos quânticos semi-empíricos
[recurso eletrônico] / Letícia Alves da Silva. - 2020.

Orientador: Eduardo de Faria Franca.
Dissertação (Mestrado) - Universidade Federal de Uberlândia,
Pós-graduação em Química.
Modo de acesso: Internet.
Disponível em: <http://doi.org/10.14393/ufu.di.2020.322>
Inclui bibliografia.

1. Química. I. Franca, Eduardo de Faria,1981-, (Orient.). II.
Universidade Federal de Uberlândia. Pós-graduação em Química.
III. Título.

CDU: 54



UNIVERSIDADE FEDERAL DE UBERLÂNDIA

Coordenação do Programa de Pós-Graduação em Química

Av. João Naves de Ávila, 2121, Bloco 51 - 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:	Dissertação de Mestrado Acadêmico, 311, PPGQUI			
Data:	vinte e sete de fevereiro de dois mil e vinte	Hora de início:	14:00	Hora de encerramento:
Matrícula do Discente:	11812QMI011			
Nome do Discente:	Letícia Alves da Silva			
Título do Trabalho:	Estudo da interação do complexo antiNS1/NS1-DENV2 usando ancoramento molecular e métodos quânticos semi-empíricos			
Área de concentração:	Química			
Linha de pesquisa:	Química Teórica e Físico-Química de Macromoléculas e Coloides			
Projeto de Pesquisa de vinculação:	Controle molecular no desenvolvimento de nanobiossensores.			

Reuniu-se no Auditório Prof. Manuel Gonzalo Hernández Terrones , Campus Santa Mônica, da Universidade Federal de Uberlândia, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Química, assim composta: Professores Doutores: Luciano Tavares da Costa, da Universidade Federal Fluminense; Sarah Arvelos, da Universidade Federal de Uberlândia e Eduardo de Faria Franca, orientador da candidata.

Iniciando os trabalhos o(a) presidente da mesa, Dr. Eduardo de Faria Franca, apresentou a Comissão Examinadora e a candidata, agradeceu a presença do público, e concedeu à Discente a palavra para a exposição do seu trabalho. A duração da apresentação da 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(as) 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):

Aprovada.

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

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 27/02/2020, às 16:53, 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 **Sarah Arvelos, Professor(a) do Magistério Superior**, em 27/02/2020, às 16:54, 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 **LUCIANO TAVARES DA COSTA, Usuário Externo**, em 03/03/2020, às 15:32, 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 **1893132** e o código CRC **OFA9CA84**.

Referência: Processo nº 23117.013611/2020-68

SEI nº 1893132

Dedico este trabalho a mim mesma por ter me permitido ser livre pra
trilhar meu próprio destino e superar todos os obstáculos para no final
deste árduo trabalho me sentir **aprovada**.

Agradecimentos

Reservo este pequeno espaço para agradecer às pessoas que tanto me ajudaram até aqui. A começar por Aquele que mantém a minha fé e esperança e, que me dá razões pra me tornar uma pessoa melhor: Deus.

Agradeço aos meus pais Rosangela e Acriziomar que mesmo sem ter tido a oportunidade de trilhar o caminho dos estudos, me ajudaram a chegar até aqui, com muito amor, carinho, dedicação, fé e perseverança.

Agradeço ao meu noivo Renan Maciel pela linda caminhada juntos e por todo incentivo, o qual me fez transpor grandes barreiras para chegar onde estou. Serei eternamente grata por todos esses anos juntos.

Agradeço ao meu querido orientador, Eduardo Franca, que acreditou no meu potencial e me deu a oportunidade de trilhar este caminho, o qual me trouxe grandes conquistas e aprendizados.

Agradeço ao meu amigo Renan Faria, que me auxiliou durante toda execução deste projeto e que muito contribui com várias ideias e ensinamentos, além de ter me dado a oportunidade de trabalhar em um projeto tão desafiador.

Agradeço aos meus colegas de trabalho, que fazem parte do grupo de pesquisa do LCQC por terem aberto as portas para mim, por toda consideração, especialmente pelo carinho durante estes dois anos. Com certeza todo apoio vindo deles fizeram uma grande diferença na minha caminhada. Os levarei comigo para sempre.

Agradeço aos orgãos de fomento CAPES, CNPQ e FAPEMIG.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

"Taking a new step, uttering a new word, is what people fear most."
Fyodor Dostoyevsky, Crime e Castigo

"Pain and suffering are always inevitable for
a large intelligence and a deep heart."

Fyodor Dostoyevsky, Crime e Castigo

"The flower that blooms in adversity, is the most
rare and beautiful of all."
Imperador da China, Filme Mulan

Resumo

Flaviroses, como a Dengue e a Zika, são doenças causadas pelos vírus do gênero *Flavirirus*. Devido aos recorrentes quadros epidêmicos, essas doenças tornaram-se emergência de saúde pública de preocupação internacional. Países de clima tropical, como o Brasil, são propícios a tais epidemias, pois apresentam clima favorável para a proliferação do mosquito transmissor. A profilaxia restrita ao combate do agente transmissor e a falta de tratamento específico, tornam os estudos de desenvolvimento de diagnósticos mais precisos e precoces, uma alternativa promissora na mudança do cenário atual. Neste sentido, muitos estudos têm voltado a atenção para as proteínas codificadas pelo vírus dentro da célula hospedeira. Dentre elas, a proteína não estrutural 1 (NS1) recebeu destaque, devido ao importante papel desempenhado nos processos de replicação e evasão viral. Atualmente, o diagnóstico da Dengue e da Zika, é realizado por meio de ensaios imunoenzimáticos, como o ELISA, que utilizam a imunoglobulina IgG como marcador. No entanto, este diagnóstico não é específico à nenhuma flavirose e pode fornecer resultados falso-positivos, devido à reações cruzadas. Alternativamente, a proteína sNS1 já é utilizada como um marcador e pode fornecer diagnósticos precoces e sem reações cruzadas. No entanto, a capacidade de discriminação entre as flaviroses por meio da afinidade do complexo antígeno-anticorpo é desconhecida. Neste contexto, este trabalho tem como objetivo compreender as interações moleculares entre a proteína não estrutural da Dengue sorotipo 2 (NS1-DENV2) e um anticorpo monoclonal antiNS1 (22NS). Para isso, foi proposto um protocolo de análise *in silico* que combina ancoramento molecular e cálculos quânticos semi-empíricos para avaliar a região e a afinidade de interação do complexo antiNS1/NS1-DENV2. Foram utilizados dois softwares de ancoramento molecular (ClusPro e PatchDock), baseados em algoritmos de busca global distintos, para obtenção das configurações iniciais do complexo. Depois, o software Rosetta *online*, foi utilizado para refinamento dessas configurações. O modelo do complexo 22NS/DENV2 e a estrutura cristalográfica 22NS/WNV, tiveram a entalpia de ligação (ΔH_{lig}) calculada à partir das informações do calor de formação (ΔH_f). Ele foi calculado no MOPAC 2016, através do Hamiltoniano semi-empírico PM7, com a aproximação MOZYME e modelo de solvatação implícita. Os resultados, mostraram que a região de interação do complexo 22NS/NS1-DENV2 localiza-se nas porções do domínio *Wing*, em acordo com dados experimentais de mapeamento de epítópos. O valor de ΔH_{lig} foi de -460,07 kJ mol⁻¹ comparável ao ΔH_{lig} do complexo cristalográfico (-620,95 kJ mol⁻¹). O protocolo proposto mostrou-se promissor na discriminação de modelos precisos e imprecisos, fornecendo modelos em concordância com estudos experimentais e, portanto, é útil para a avaliação da região e da afinidade de interação em sistemas semelhantes.

Palavras-Chave: Ancoramento molecular, Semi-empírico, Dengue, NS1, Flaviroses

Abstract

Flaviviruses, such as Dengue and Zika, are diseases caused by viruses from the genus *Flavivirus*. Responsible for epidemic conditions, they have become a public health emergency of international concern. Tropical climate countries, such as Brazil, are prone to these epidemics because they have a favorable climate for the proliferation of the transmitting mosquito. The prophylaxis restricted to combating the transmitting agent and the lack of specific treatment have made studies focus on developing more accurate and early diagnoses, a promising alternative in changing the current scenario. In this sense, many researches have been studying the viral proteins encoded within the host cell. Among them, the non-structural protein 1 (NS1) has received focus due to its important role in replication processes and viral evasion. Currently, the diagnosis of Dengue and Zika diseases is performed by using enzyme immunoassays, like ELISA, which uses immunoglobulin G as a marker. However, this kind of diagnosis is non-specific and may provide false-positive results due to cross-reactions. Alternatively, sNS1 protein is already used as a marker, and it provides early diagnostics without cross-reactions. However, the capability to discriminate between flaviviruses through the antigen-antibody complex affinity is still unknown. In this context, this study aims to understand the molecular interactions between the non-structural protein of Dengue serotype 2 (NS1-DENV2) and a monoclonal antibody antiNS1. In order to evaluate the interacting affinity region through getting the antiNS1/NS1-DENV2 complex, we are proposing an *in silico* analysis protocol, which combines docking molecular and semiempirical approach. By using two molecular docking softwares (ClusPro and PatchDock), based on different global search algorithms, we got the initial configurations of the complex. Next, the Rosetta online server was used to refine those configurations. The bonding enthalpy (ΔH_{bind}) of the 22NS/DENV2 model and the 22NS/WNV crystallographic structure were calculated from their heat of formation, ΔH_f . All ΔH_f were calculated in MOPAC 2016 using the semiempirical Hamiltonian PM7, along with the MOZYME approach and implicit solvation model. The results showed that the 22NS/NS1-DENV2 complex interaction region is located on the Wing domain, in agreement with experimental epitopes mapping data. The value of ΔH_{lig} obtained from 22NS/DENV2 was $-460.07 \text{ kJ mol}^{-1}$, which is comparable to ΔH_{lig} the crystallographic complex ($-620.95 \text{ kJ mol}^{-1}$). According to this, the proposed protocol might be a promising tool for discriminating accurate and imprecise models, providing models in agreement with experimental data. Therefore, this protocol is useful to evaluate the interaction region and the affinity for similar systems.

Keywords: Molecular docking, Semiempirical, Dengue, NS1, Flaviviruses

Lista de Figuras

Figura 1 – Diagrama esquemático da estrutura comum dos aminoácidos	19
Figura 2 – Representação esquemática da classificação dos aminoácidos.	20
Figura 3 – Diagrama esquemático de uma cadeia peptídica evidenciando as ligações peptídicas e as extremidades N-terminal e C-terminal.	21
Figura 4 – Figura esquemática evidenciando ângulos de rotação ϕ e ψ	22
Figura 5 – Representação esquemática dos tipos de estrutura secundária: α -hélice, folha- β e dobra β	23
Figura 6 – Diagrama topológico da estrutura do monômero NS1.	25
Figura 7 – Representação estrutural da NS1 (A) dimérica e (B) hexamérica, evidenciando os domínios β -roll, wing e β -ladder.	26
Figura 8 – Representação da estrutura geral das imunoglobulinas.	27
Figura 9 – Fluxograma da metodologia de ancoramento proposta.	39
Figura 10 – Representações cartoon da (A) estrutura cristalográfica do sistema WNV/22NS e, das <i>poses</i> de menor energia obtidas por re-ancoramento (B) pelo ClusPro e (C) PatchDock.	41
Figura 11 – Representação em <i>cartoon</i> das <i>poses</i> de menor energia selecionadas do <i>webserver</i> Cluspro 2.0, ordenadas em forma crescente de energia A-E. .	42
Figura 12 – Representação em <i>cartoon</i> das <i>poses</i> de menor energia selecionadas do <i>webserver</i> PatchDock, ordenadas em forma crescente de energia A-C. .	42
Figura 13 – Quantidade de <i>poses</i> por modelo de acordo com o critério estabelecido pela CAPRI.	43
Figura 14 – Representação em malha da sobreposição da interface de interação da estrutura cristalográfica 22NS/WNV, em coloração amarela, com os modelos (A) WC01, em azul, e (B) WP01, em vermelho.	45
Figura 15 – Representação em malha da sobreposição da interface de interação da estrutura cristalográfica 22NS/WNV, em coloração amarela, com os modelos (A) WC01, em azul, com resíduos em esferas amarelas e (B) WP01, em vermelho, com resíduos em esferas vermelhas.	46
Figura 16 – Interfaces de interação dos modelos do complexo 22NS/DENV2 com destaque aos resíduos pertencentes aos epítopos E1, na coloração verde, E2, em magenta e E3, em laranja.	47

Lista de abreviaturas e siglas

NS	Proteína não estrutural
NS1	Proteína não estrutural 1
sNS1	Proteína não estrutural 1 solúvel hexamérica
mNS1	Proteína não estrutural 1 fixada na membrana dimérica
ELISA	Ensaio Imunienzimático de Imunoabsorção Enzimática
IgG	Imunoglobulina G
IgM	Imunoglobulina M
DENV	Vírus da Dengue
DENV2	Vírus da Dengue sorotipo 2
WNV	Vírus da Febre do Nilo do Oeste
22NS	Anticorpo monoclonal anti-NS1
ΔH_{lig}	Variação de Entalpia de Ligação
ΔH_f	Calor de Formação
PCR-RT	Reação da Transcriptase Reversa
Fc	Fragmento Cristalizável
Fab	Fragmento de Ligação ao Antígeno
V_L	Região Variável da Cadeia Leve
V_H	Região Variável da Cadeia Pesada
C_L	Região Constante da Cadeia Leve
C_H	Região Constante da Cadeia Pesada
CDR	Região de Determinação da Complementariedade
AM	Ancoramento Molecular
VDW	Van der Waals
PMF	Potencial de Força Médio

CAPRI	Avaliação Crítica de Interações Preditas
CASP	Avaliação Crítica das Técnicas de Previsão da Estrutura Proteica
RMSD	Desvio da Raiz Quadrada Média das Posições Atômicas
i-RMSD	RMSD dos Átomos da Interface
Fnat	Fração de Resíduos Nativos
PDB	Banco de Dados de Proteínas
OM	Orbital Molecular
OML	Orbital Molecular Localizado
SCF	Campo Auto-Consistente Iterativo
MNDO	<i>Modified Neglect of Differential Overlap</i>
AM1	<i>Austin Model 1</i>
PM3	<i>Parametric Method 3</i>
PM6	<i>Parametric Method 6</i>
PM7	<i>Parametric Method 7</i>
MOPAC	Pacote de Orbital Molecular
WP	Modelos Obtidos no PatchDock para a estrutura da NS1 do WNV
WC	Modelos Obtidos no ClusPro para a estrutura da NS1 do WNV
DP	Modelos Obtidos no PatchDock para a estrutura da NS1 do DENV2
DC	Modelos Obtidos no ClusPro para a estrutura da NS1 do DENV2
WNV/22NS	Complexo entre a NS1 do WNV e o anticorpo 22NS
DENV2/22NS	Complexo entre a NS1 do DENV2 e o anticorpo 22NS
NRPE	Número de Resíduos que Pertencem a Epítopo

Sumário

1	INTRODUÇÃO	16
2	ESTRUTURA PROTEICA	19
2.1	Aminoácidos e peptídeos	19
2.2	Conformação: níveis estruturais	21
3	SISTEMA ESTUDADO: ANTINS1/NS1-DENV2	25
3.1	Montagem estrutural da proteína NS1	25
3.2	Montagem estrutural das imunoglobulinas	26
3.3	Interação antígeno-anticorpo	28
4	FERRAMENTAS METODOLÓGICAS	29
4.1	Ancoramento molecular	29
4.2	Método quântico semi-empírico	33
5	METODOLOGIA	36
5.1	Ancoramento molecular	36
5.2	Cálculo da entalpia de ligação: MOPAC	39
6	RESULTADOS E DISCUSSÕES	41
6.1	Ancoramento molecular	41
6.2	Análise da viabilidade do procedimento metodológico	44
7	CONSIDERAÇÕES FINAIS	49
	REFERÊNCIAS	50
	APÊNDICE A – REVISÃO DOS MÉTODOS IMUNOENZIMÁTICOS PARA DETECÇÃO DA NS1.	59

1 Introdução

As flaviroses estão entre as doenças que mais causam mortes anualmente, sendo muitas delas devido ao diagnóstico incorreto ou tardio da infecção. Estas doenças são causadas pelos vírus do gênero *Flavivirus*, que podem ser transmitidos para o indivíduo pela picada de mosquito de espécies específicas, ou outras formas secundárias como transfusão de sangue e, em alguns casos, como do vírus da Zika, por fluidos corporais, como o sêmen. A Dengue e a Zika são as flaviroses mais recorrentes, acometendo principalmente a população de países subdesenvolvidos de clima tropical, onde a proliferação do mosquito transmissor, *Aedes aegypti*, é favorável. Outros exemplos de flaviroses de menor ocorrência são: a Chikungunya, a Febre Japonesa, a Febre Amarela e a Febre do Nilo do Oeste [1, 2, 3].

O Brasil encontra-se entre os países que apresentaram maior número de casos associado a flaviroses. A primeira epidemia de Dengue tem registro em 1986, e anos depois, em 2015, foi registrada a segunda epidemia. Além disso, no mesmo ano, o país sofreu com o surgimento de casos de microcefalia associados à Zika. Além disso, em 2018, foi isolado pela primeira vez o vírus da Febre do Nilo do Oeste, encontrado em cavalos no estado de Santa Catarina, mas há evidências de que essa flavirose já circulava pelo país por pelo menos oito anos. Infelizmente, tanto a Dengue quanto a Zika, se tornaram sérios problemas de saúde pública, sendo responsáveis por um grande número de casos em todo o país. No entanto, a principal medida profilática promovida pelos órgãos governamentais, tem sido a conscientização da população para eliminação de focos do mosquito transmissor [4, 5, 6, 7].

O *Flavivírus* causador dessas doenças tem como material genético uma fita simples de RNA positiva que codifica dentro da célula hospedeira três proteínas estruturais, o capsídio (C), a pré-membrana (prM) e o envelope (E) e, sete proteínas não estruturais (NS1, NS2A, NS2B, NS3, NS4A, NS4B, NS5). Essas proteínas não estruturais (NS), desempenham papéis essenciais à infecção viral, seja no processo de replicação ou na evasão viral [8, 9].

A falta de um método eficaz e seguro de imunização ativa é a maior dificuldade encontrada pelas equipes de saúde na promoção das medidas profiláticas. Diante disso, o estudo do agente causador e o desenvolvimento de métodos de diagnóstico das Flaviroses tornam-se uma alternativa promissora na mudança deste cenário. Esta ação pode ter grande impacto em flaviroses endêmicas, como a Dengue, uma vez que “o diagnóstico preciso e precoce da dengue é de fundamental importância para o manejo clínico dos pacientes, assim como para a orientação de ações epidemiológicas direcionadas para o controle da doença” (MELO, 2016, p.36) [10].

O quadro clínico dos indivíduos infectados com Dengue é genérico e variável, ou seja, os sintomas relatados são comuns a quase todas as flaviroses, mas variam muito diante das

particularidades do paciente infectado. Logo, torna-se inviável um diagnóstico confiável apenas pela avaliação clínica realizada por um profissional médico [11].

Desse modo, o diagnóstico atual da Dengue se baseia na observação dos sintomas apresentados em associação a exames laboratoriais. Dentre os quais cita-se: (i) isolamento viral (dosagem dos níveis de IgG e IgM no soro do paciente) [12], (ii) testes fotométricos de ELISA [13] e, principalmente, (iii) os testes virológicos de transcrição seguida da reação em cadeia da enzima polimerase em tempo real, mais conhecidos como PCR-RT [14]. No entanto, todos esses testes apresentam limitações que restringem sua ampla aplicação sendo, respectivamente, (i) a baixa especificidade em virtude da possibilidade de reatividade cruzada, (ii) ocorrência de falso-positivos associados à detecção indiretamente e (iii) o alto custo associado a necessidade de uma ampla estrutura laboratorial.

Nesse contexto, os métodos baseados na identificação de proteínas não-estruturais - em especial NS1, NS3 e NS5, visto que elas possuem papel importante na replicação viral e elevado grau de conservação estrutural entre os sorotipos [15] - tem emergido como uma alternativa para o diagnóstico da DENV ainda na fase aguda (infecção primária). Dentre as supracitadas, a NS1 participa simultaneamente dos processos de evasão e replicação viral, e tem sido alvo de diversos estudos com aplicações da área biomédica, devido a sua versatilidade dentro de tais processos. Mais especificamente, essa proteína é secretada para o meio extracelular na sua forma hexamérica (sNS1) em concentrações da ordem de até $50 \mu\text{g mL}^{-1}$ no soro de pacientes infectados pelo DENV, portanto, em níveis detectáveis por dispositivos analíticos [16, 17, 18, 19, 20]. Sendo assim, a NS1 é considerada o maior biomarcador antigênico de Flaviroses, como a Dengue [21].

Portanto, na literatura são encontrados diversos estudos direcionados ao desenvolvimento de métodos imunoenzimáticos que tem como alvo a proteína NS1 para o diagnóstico específico e precoce da dengue. As tabelas do Apêndice A, summarizam de modo sistemático os principais estudos e seus resultados encontrados durante a realização do levantamento bibliográfico desse trabalho.

Através da análise desses estudos, é possível perceber a existência de esforços direcionados ao desenvolvimento de métodos mais específicos, sensíveis e acessíveis na detecção da Dengue. No entanto, quase todos tratam-se de ensaios que requerem diversas etapas laboratoriais e que não visam um diagnóstico diferencial em relação as demais Flaviroses, ou mesmo a detecção simultânea de diversos biomarcadores [22]. Mas, vale ressaltar que, a maioria dos estudos dedicam seus esforços à identificação da proteína NS1, mais precisamente, por métodos baseados na especificidade da interação antígeno-anticorpo, ou seja, no estudo de sistema antiNS1/NS1.

Nesse contexto, o presente trabalho teve como principal objetivo a compreensão das interações moleculares entre a proteína não estrutural da Dengue sorotipo 2 (NS1-DENV2) e um anticorpo monoclonal antiNS1 (22NS).

Para o cumprimento do objetivo geral, os objetivos específicos deste trabalho consis-

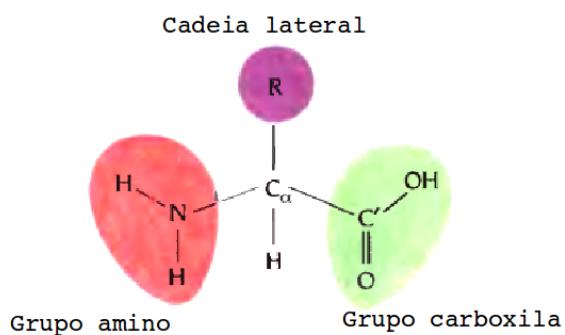
tiram em: **1.**identificar o local de interação entre o antígeno e o anticorpo por meio de cálculos de ancoramento molecular e **2.**calcular a entalpia de ligação entre o antígeno e o anticorpo utilizando métodos quânticos semi-empíricos.

2 Estrutura proteica

2.1 Aminoácidos e peptídeos

As proteínas constituem o grupo mais abundante de biomoléculas presentes no interior das células, e podem ser encontradas na forma de pequenos peptídeos ou de moléculas poliméricas. Elas derivam-se da combinação entre 20 diferentes subunidades, denominadas aminoácidos. Todo aminoácido apresenta uma estrutura comum formada por um átomo de carbono central (C_α), ao qual estão ligados: um átomo de hidrogênio (H), um grupo carboxílico (COOH), um grupo amino (NH_2) e um grupo R ou cadeia lateral, a qual é responsável pela diferenciação entre os aminoácidos [23]. Na Figura 1 é esquematizado a estrutura geral de um aminoácido.

Figura 1 – Diagrama esquemático da estrutura comum dos aminoácidos

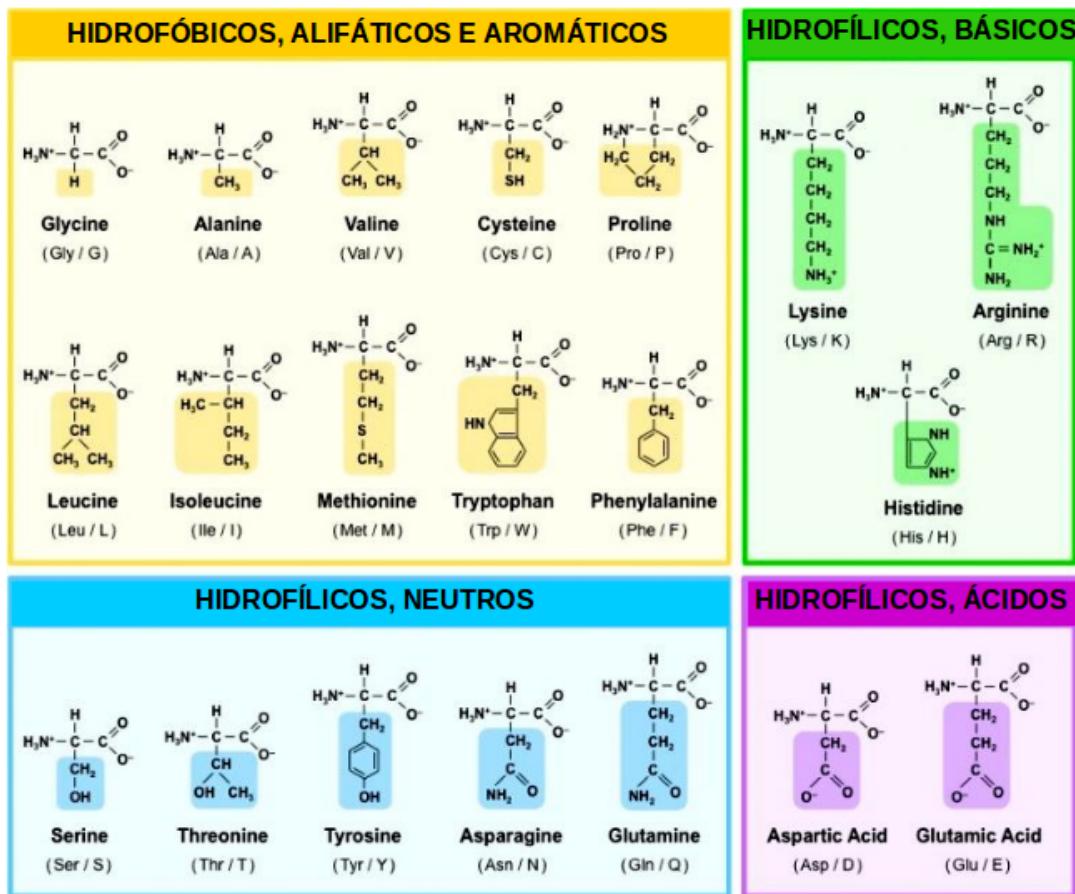


Fonte: Branden (1999) [23] - ADAPTADO

Os aminoácidos são classificados de acordo com as características da cadeia lateral, que inicialmente pode ser hidrofóbica (apolar) ou hidrofílica (polar). Dentre as cadeias laterais hidrofóbicas existem as alifáticas e as aromáticas. Enquanto que as hidrofílicas podem ser neutras (sem carga), básicas (carregadas positivamente) ou ácidas (carregadas negativamente). A classificação de cada um dos 20 aminoácidos pode ser vista na Figura 2^a. Como veremos em seções posteriores, a compreensão das diferentes características das cadeias laterais é importante para estudo da conformação estrutural das proteínas, porque as interações realizadas entre os átomos das cadeias laterais dos aminoácidos definem o dobramento (em inglês *folding*) proteico e também a afinidade de interação com outras moléculas [24].

^a Os nomes dos aminoácidos na Figura 2 estão em inglês para facilitar a compreensão da sigla que representa cada aminoácido. Além disso as letras de identificação da sequência FASTA (siglas de única letra de aminoácidos) são colocados para facilitar a identificações de resíduos de aminoácidos posteriormente. Exemplo – nome:Lysine, sigla: Lys, fasta:K

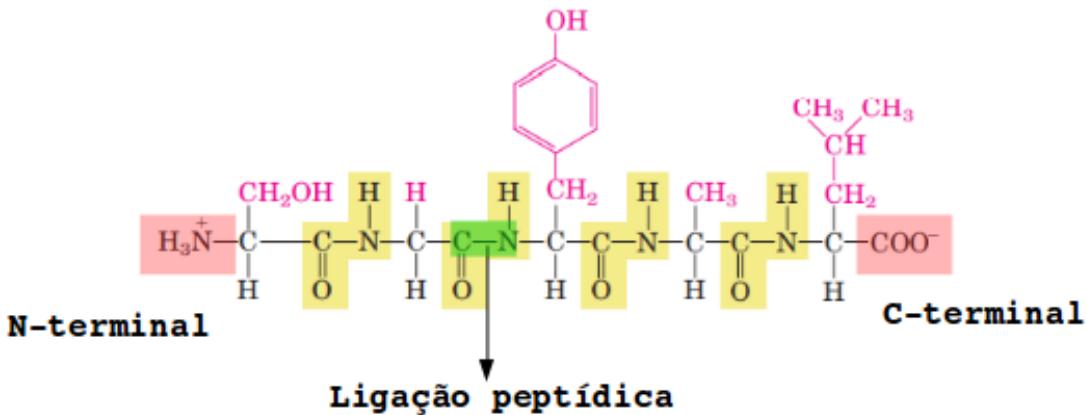
Figura 2 – Representação esquemática da classificação dos aminoácidos.



Fonte: Cornell (2016)[28] - ADAPTADO.

Todos os aminoácidos encontrados nas proteínas são L-aminácidos devido a sua forma quiral L, determinada pelos grupos amino, carboxílico e R no sentido horário com relação ao carbono central (C_α). Durante o processo de síntese das proteínas esses aminoácidos unem-se uns aos outros por ligações covalentes. Essas ligações são formadas por meio de uma reação de condensação entre o grupo amino de um com o grupo carboxílico do outro e, é denominada ligação peptídica (Figura 3). O grupo de aminoácido que se origina da perda de água, durante essa reação, é denominado resíduo. Devido ao padrão de ligação seguido durante a síntese, o grupo amino (α -amino) do primeiro aminoácido e o grupo carboxílico (α -carbonila) do último, ambos ligados a um carbono C_α , encontram-se intactos nas extremidades da cadeia peptídica, de modo que a cadeia estende-se da extremidade N-terminal à C-terminal, Figura 3 [24, 23].

Figura 3 – Diagrama esquemático de uma cadeia peptídica evidenciando as ligações peptídicas e as extremidades N-terminal e C-terminal.



Fonte: Lehninger (2005) [24] - ADAPTADO

2.2 Conformação: níveis estruturais

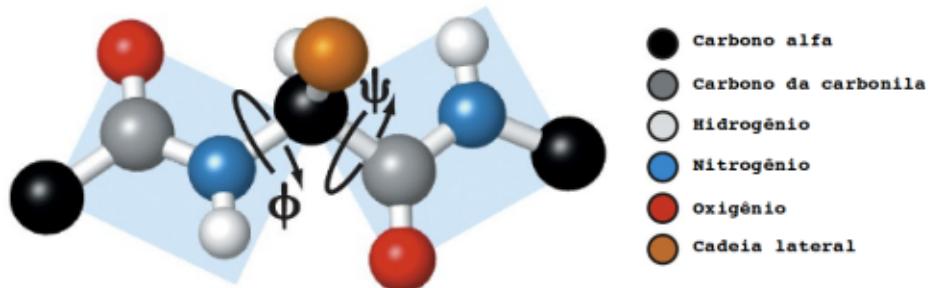
A união de vários aminoácidos dá origem a longas cadeias peptídicas, denominadas cadeias polipeptídicas e são essas cadeias que dão origem às estruturas proteicas. O arranjo espacial no qual todos átomos de uma proteína se encontra é denominado conformação, ou estrutura secundária. Ela é resultante das rotações entre as ligações simples dos resíduos que formam a cadeia polipeptídica. Devido ao grande número de rotações possíveis, é esperado que a estrutura proteica assuma uma infinidade de conformações. No entanto, nem todos os arranjos espaciais são concebíveis, eles são limitados pelas condições do meio, de maneira que a conformação que permanece é aquela que apresenta maior estabilidade termodinâmica, definida pelo menor valor de energia livre de Gibbs (ΔG). Dessa forma, quando as proteínas estão em sua conformação funcional mais estável, ela é denominada conformação nativa. A manutenção dessa conformação é feita por meio das ligações e interações estabelecidas entre os átomos das cadeias laterais [26].

Os carbonos α dos resíduos de aminoácidos adjacentes estão separados por três ligações covalentes em um arranjo que pode ser representado por $C_\alpha-C-N-C_\alpha$. A ligação C–N não apresenta rotação devido ao efeito de ressonância observado entre a ligação dupla do átomo de oxigênio da carbonila com o átomo de nitrogênio adjacente. Como consequência, os átomos de um mesmo aminoácido encontram-se em um mesmo plano e os átomos de oxigênio da carbonila e o átomo de hidrogênio da amina estão em posição *trans*, Figura 4. A livre rotação dentro da cadeia peptídica só é possível nas ligações N–C α e C α –C, o que limita o número de conformações que uma proteína pode assumir [26, 24].

Os ângulos de rotação das ligações N–C α e C α –C são denominados ϕ e ψ , respectivamente. Na teoria, esses ângulos podem assumir valores que vão de -180° a $+180^\circ$,

assumindo $\phi = \psi = 180^\circ$ em que todos os átomos do grupo peptídeos estão no mesmo plano. No entanto, existem algumas combinações que são inatingíveis devido a efeitos estéricos, Figura 4. Dessa forma, a análise mais detalhada das combinações possíveis dos valores de ϕ e ψ fornecem informações relevantes sobre a conformação nativa de um proteína [24].

Figura 4 – Figura esquemática evidenciando ângulos de rotação ϕ e ψ .



Fonte: Lehninger (2005) [24] - ADAPTADO

Dado a complexidade do arranjo espacial das proteínas, sua conformação é dividida em quatro níveis estruturais com ordem crescente de complexidade: estrutura primária, secundária, terciária e quaternária. A estrutura primária, mais simples de todas, trata-se da sequência linear dos resíduos de aminoácidos que formam a cadeia polipeptídica da proteína. As demais estruturas apresentam definições mais extensas, por isso serão discutidas separadamente.

2.2.1 Estrutura secundária

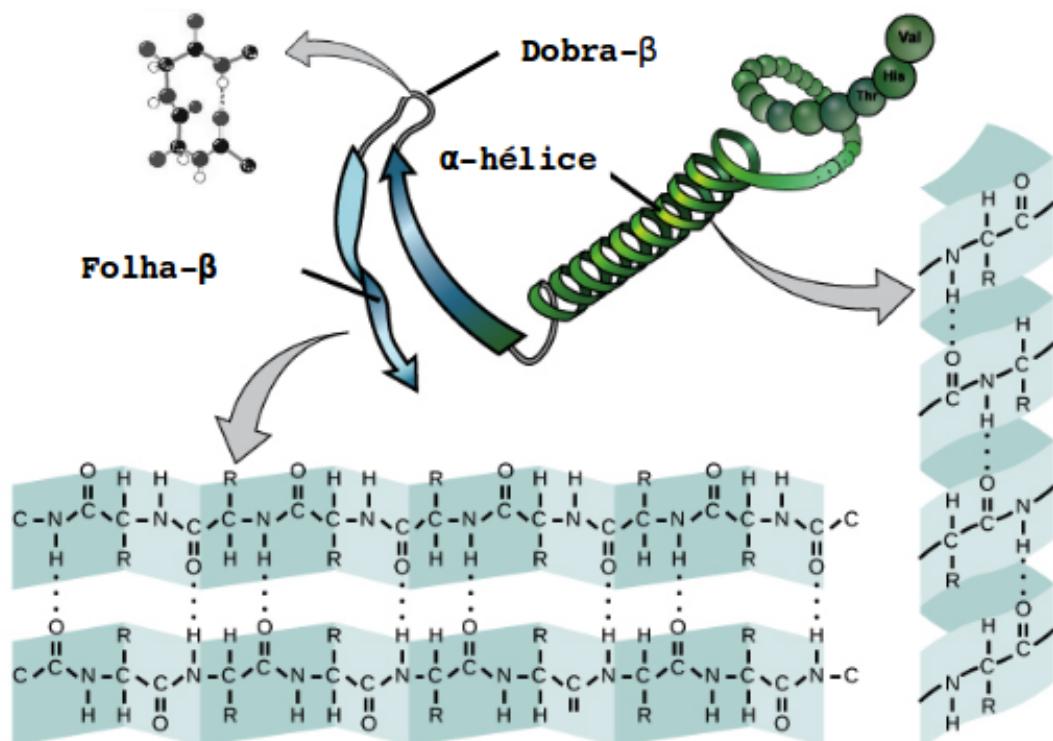
A estrutura secundária consiste no arranjo tridimensional da cadeia polipeptídica adquirido quando os átomos das cadeias laterais adjacentes interagem entre si, dando origem a padrões estruturais recorrentes denominados conformações α -hélice (em inglês, α -helix) e folha β (em inglês, β -sheet) [24, 27, 23].

A conformação α -hélice possui uma estrutura helicoidal formada pela torção (para direita ou para esquerda) da cadeia polipeptídica em torno de um eixo longitudinal que passa pelo centro da hélice. Nesse arranjo as cadeias laterais projetam-se para a face externa da hélice, como mostrado na Figura 5 [24, 27, 23].

A conformação folha β , Figura 5, é formada pela disposição das cadeias polipeptídicas lado a lado (paralela ou antiparalela), em zigue-zague, dando origem a uma estrutura que se assemelha a uma série de pregas. Ela é mantida por meio das ligações de hidrogênio formadas entre os seguimentos adjacentes das cadeias polipeptídicas. As cadeias laterais dos resíduos de aminoácidos adjacentes projetam-se em direções opostas a partir da estrutura em zigue-zague, criando um padrão alternado [24, 27, 23].

Além das conformações α -hélice e folha β existem as dobras β (em inglês, *beta turn*), também conhecidas como alças (em inglês, *loops*). Essa conformação é um elemento de conexão, formada a partir dos resíduos que conectam uma estrutura secundária à outra, sendo constituída por apenas quatro resíduos de aminoácidos. Sua estrutura é mantida fixa graças a ligação de hidrogênio formada entre o átomo de oxigênio do grupo carbonila do primeiro resíduo com o átomo de hidrogênio da amina do quarto resíduo, Figura 5 [24, 27, 23].

Figura 5 – Representação esquemática dos tipos de estrutura secundária: α -hélice, folha β e dobra β .



Fonte: Lehninger (2005) [24] - ADAPTADO

2.2.2 Estruturas terciárias e quaternárias

A maneira como as estruturas secundárias estão dispostas no espaço é denominada estrutura terciária. Esse nível estrutural é alcançado pelo arranjo global de todos os átomos da proteína que pertencem a diferentes estruturas secundárias. A conformação adquirida se mantém por meio das interações, fracas ou covalentes, entre os resíduos, sejam eles adjacentes ou não [24, 27, 23].

Todas as proteínas apresentam estrutura primária, secundária e terciária, no entanto, um quarto nível estrutural é observado apenas em proteínas que apresentam mais de uma cadeia polipeptídica, sejam elas iguais ou diferentes, formando multímeros. A es-

trutura quaternária é, portanto, a conformação obtida por meia das interações entre as subunidades que formam o grupo proteico [24, 27, 23].

De acordo com a complexidade dos níveis estruturais de uma proteína, ela pode ser classificada em fibrosa ou globular. Proteínas denominadas fibrosas geralmente são formados por apenas um tipo de estrutura secundária e suas cadeias se encontram dispostas em fios ou folhas longas; enquanto que as globulares apresentam uma mistura de estruturas secundárias e suas cadeias polipeptídicas se encontram enoveladas em forma esférica ou globular. O foco será dado apenas a proteínas globulares, pois as estruturas proteicas trabalhadas aqui pertencem a essa classificação [24, 27, 23].

A estrutura tridimensional de uma proteína globular é formada pela combinação de pequenos grupos de estruturas secundárias organizadas em arranjos geométricos específicos, denominados motivos (em inglês, *motifs*) conhecidos também como estruturas supersecundárias. Em geral, essas conformações estão associadas à características funcionais da proteína e por isso são arranjos recorrentes. Quando diferentes motivos se combinam originando uma estrutura globular compacta, esta conformação é conhecida como domínio (em inglês, *domain*). Uma estrutura proteica, seja ela formada por apenas uma ou várias cadeias polipeptídicas, se enovelá (em inglês, *folds*) em diferentes domínios, os quais tem seu ordenamento espacial de acordo com a função biológica exibida pela proteína [24, 27, 23].

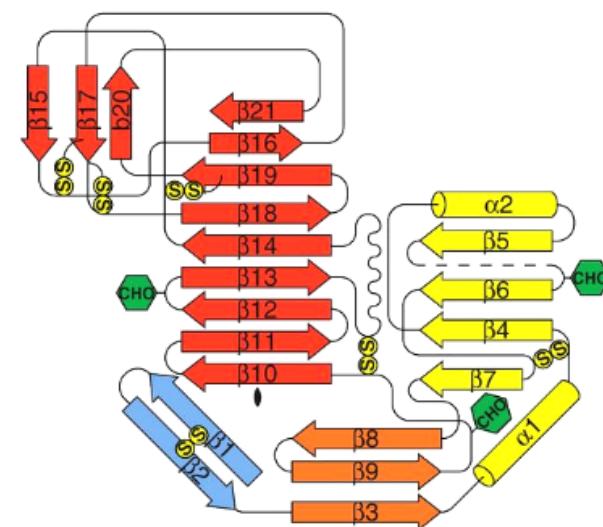
3 Sistema estudado: antiNS1/NS1-DENV2

3.1 Montagem estrutural da proteína NS1

A proteína não estrutural 1 (NS1) é uma glicoproteína^b e pode ser encontrada na forma monomérica, dimérica (associada à membrana, mNS1) e hexamérica (forma secretada, sNS1). O monômero, Figura 6, de 48 kDa apresenta dois sítios de glicosilação (verde) e seis ligações de dissulfeto (amarelo) intramoleculares. Esta proteína enovelava-se formando três diferentes domínios, a saber *β-roll* (azul), *wing* (amarelo e laranja) e *β-ladder* (vermelho) [29, 30].

O domínio *β-roll* é formado pela sequência de aminoácidos a.a. 1-29 em um enovelamento conhecido como *β-harpin* e está localizado na porção N-terminal da cadeia polipeptídica. O domínio *wing* encontra-se entre as extremidades da cadeia polipeptídica e é constituído pela sequência 38-151 em uma combinação de folhas β e α -hélices. E por fim, o domínio *β-ladder* (a.a.181-352) é composto por folhas β anti-paralelas que são conectadas por vários segmentos de conexão formando uma região em alças, conhecida como superfície de alças (em inglês, *loop surface*). A composição de estruturas secundárias de cada domínio pode ser melhor visualizada no diagrama topológico abaixo [29, 30].

Figura 6 – Diagrama topológico da estrutura do monômero NS1.



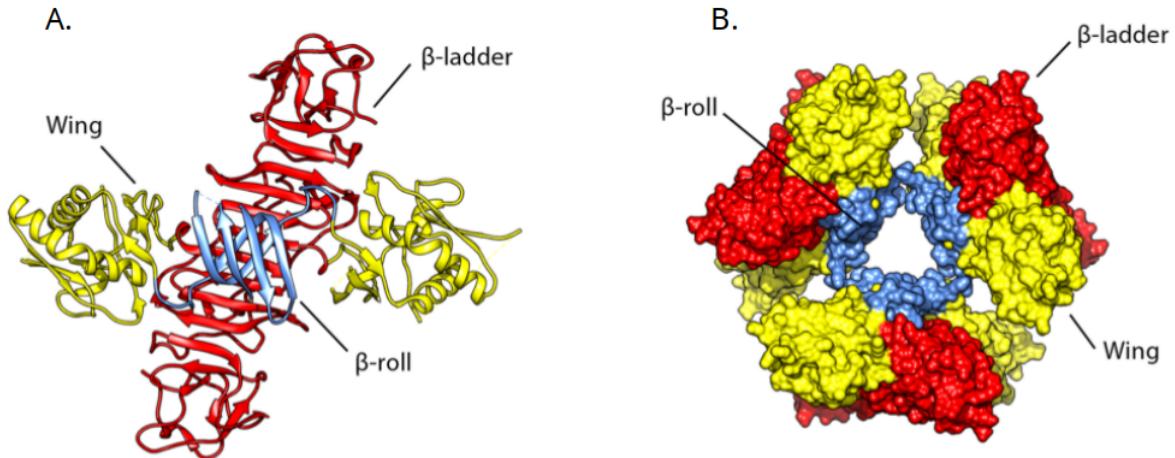
Fonte: Akey et al. (2014) [29]

^b Glicoproteína é o nome dado à proteínas que tem porções de carboidrato ligado em sua estrutura, geralmente nos resíduos de Cisteína.

A dimerização da NS1 ocorre rapidamente dentro do retículo endoplasmático (RE) por meio do enlaçamento dos domínios β -roll de dois monômeros, dando origem à forma dimérica (mNS1), Figura 7-A. O dímero formado sofre processo de glicosilação com carboidratos de alto teor de manose e, logo depois, é transportado para seus principais locais de atuação: retículo endoplasmático (replicação viral), membrana plasmática e compartimento extracelular (evasão viral) [9, 8].

O processo de evasão viral é caracterizado pelo transporte extracelular da NS1 em sua forma hexamérica ou forma secretada (sNS1), Figura 7-B. Ela é uma lipoproteína composta por três estruturas diméricas que se mantêm unidas por meio de interações hidrofóbicas entre os domínios β -roll, formando um núcleo lipídico central inacessível ao solvente. Esse processo ocorre dentro do complexo de Golgi no retículo endoplasmático da célula hospedeira e, a NS1 é transportada para o meio extracelular através de vesículas transportadoras. Vale ressaltar que, as glicosilações presentes da forma dimérica (mNS1) não permanecem na forma secretada (sNS1), elas são removidas pelas enzimas glicosidases e glicosiltransferase durante o processo de polimerização da NS1 [8].

Figura 7 – Representação estrutural da NS1 (A) dimérica e (B) hexamérica, evidenciando os domínios β -roll, wing e β -ladder.



Fonte: Scaturro (2015) [9]

3.2 Montagem estrutural das imunoglobulinas

Uma resposta imunológica surge como resultado da exposição a um estímulo estranho, o qual é desencadeado por compostos denominados抗ígenos. Dentre as respostas imunológicas existentes a de maior relevância para este trabalho é a adaptativa humoral. Ela é responsável pela proteção do organismo contra infecções e vírus encontrados na porção

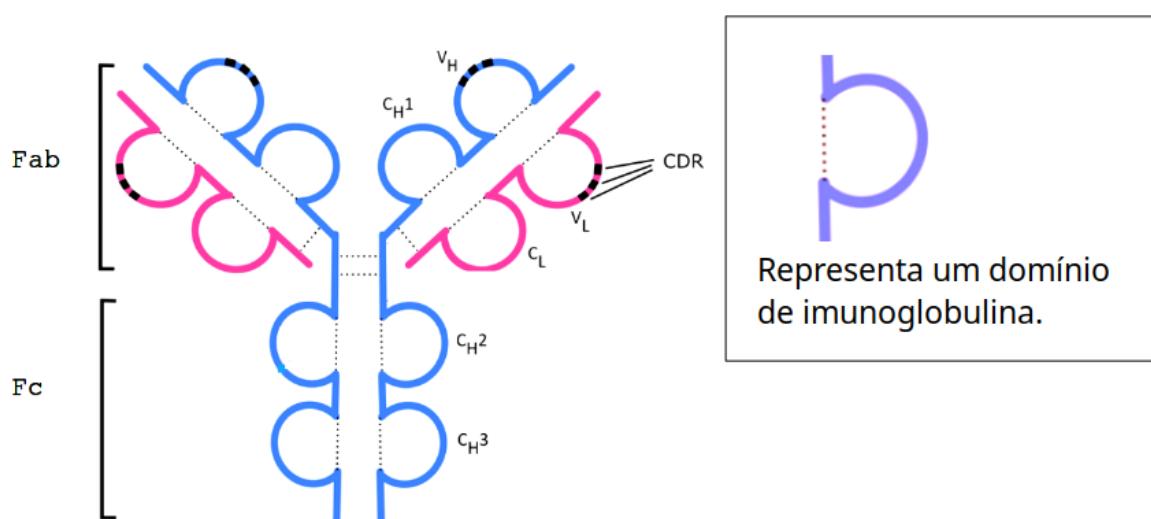
extracelular. As principais proteínas participantes desse processo são os anticorpos ou imunoglobulinas, produzidos e secretados pelas células B ou linfócitos [26].

A estrutura molecular das imunoglobulinas é composta por quatro unidades básicas: duas cadeias leves (2.200 Da) e duas cadeias pesadas (5.300 Da), que se unem por meio de ligações dissulfeto dando origem a uma estrutura em formato de Y, Figura 8. Elas são enoveladas seguindo um padrão estrutural, denominado de domínio das imunoglobulinas. Este domínio é formado por folhas β antiparalelas conectadas por ligações de dissulfeto intra-cadeias em uma estrutura de dobra compacta. Cada cadeia leve se enovelá de forma a obter dois domínios e, cada cadeia pesada é formada por quatro desses domínios [26, 31].

As cadeias leve e pesada estão distribuídas em dois tipos de fragmentos, o fragmento de união ao antígeno (Fab do inglês *fragment antigen-binding*) e o fragmento cristalizável (Fc), Figura 8. Cada Fab é composto por dois domínios de cadeia leve e dois de cadeia pesada e o Fc é formado pelos demais domínios da cadeia pesada. Dentro dessa divisão, os resíduos das cadeias que constituem os domínios do Fab apresentam uma variabilidade de resíduos. Esta região é comumente representada pela letra V de variável, sendo V_H para região variável da cadeia pesada e V_L para cadeia leve. Portanto, as demais são representadas pela letra C, de constante, seguindo a mesma lógica [32].

A região responsável pelo reconhecimento molecular está localizada na extremidade da região variável dos Fabs e é denominada região de determinação da complementariedade (CDR, em inglês *Complementary-determining Region*), Figura 8. Cada cadeia possui três CDRs, as quais são comumente formadas pelas seguintes sequências de aminoácidos: CDR1 (a.a.31-34), CDR2 (a.a.50-56) e CDR3 (a.a.93-99) [33, 34].

Figura 8 – Representação da estrutura geral das imunoglobulinas.



Fonte: Abbas (2008) [26]- ADAPTADO.

3.3 Interação antígeno-anticorpo

Tanto a NS1 secretada (sNS1) quanto a associada à membrana celular (mNS1) são altamente imunogênicas, ou seja, são susceptíveis ao reconhecimento pelas proteínas do sistema imunológico: os anticorpos. Os contatos interatômicos estabelecidos entre o anticorpo e o antígeno são altamente específicos e resultantes da complementariedade química entre os resíduos de aminoácido que formam o epítopo^c do antígeno e os da CDR do anticorpo. A complementariedade é obtida por meio de mudanças conformacionais da interface de interação e, diz respeito à forma e localização dos resíduos polares, apolares e de ligações de hidrogênio. Sendo, portanto, o fator condutor para formação das interações reversíveis responsáveis pela manutenção do complexo antígeno-anticorpo [35, 32, 24].

O mecanismo de reconhecimento das moléculas de anticorpo tem sido utilizado na detecção específica de moléculas em diferentes tipos de amostras [36, 37, 38]. Atualmente, este princípio tem sido aplicado a amostras biológicas, como o soro, para o diagnóstico de doenças. Tal procedimento é realizado por meio da detecção de moléculas marcadoras, como proteínas antigênicas [39, 40].

Desta forma, o estudo das interações de complexos antígeno-anticorpo podem fornecer bases moleculares para compreensão do processo de reconhecimento. Dessa maneira, a longo prazo, doenças que compartilham marcadores similares, como a Zika e a Dengue, podem ser diferenciadas por diagnósticos, que utilizam o princípio de afinidade específica dos complexos antígeno-anticorpo.

A maneira mais eficiente de estudar interações atômicas entre duas moléculas, é através da estrutura tridimensional do complexo formado por elas. Dentre as técnicas de maior relevância para o estudo de estruturas proteicas está a cristalografia de raio X. Esta técnica permite obter estruturas tridimensionais com resoluções consideravelmente boas, mas, o procedimento experimental envolvido é demorado, difícil e caro. Diante disso, ferramentas computacionais, como o ancoramento molecular (em inglês, *docking molecular*), permitem o estudo de interações entre proteínas por meio da modelagem computacional. Dessa forma, os complexos proteicos podem ser avaliados em um tempo muito menor com boa concordância a resultados experimentais [41].

^c Epítopo é a parte de um antígeno que está diretamente envolvida na interação antígeno-receptor (pequena parte do antígeno) é chamada de determinante antigênico ou epítopo. Um antígeno pode ter numerosos epítopos que podem ser iguais (epítopos repetidos)

4 Ferramentas metodológicas

4.1 Ancoramento molecular

Compreender a natureza do reconhecimento macromolecular significa saber como e onde ocorre a interação proteína-proteína, e isto é um grande desafio. Entretanto, a partir da estrutura tridimensional do complexo proteico, é possível analisar a região e os resíduos envolvidos na interface de interação do complexo. Isto permite compreender a afinidade ou não afinidade entre duas proteínas [42].

Nas últimas duas décadas, uma grande variedade de algoritmos teóricos foram desenvolvidos com o objetivo de prever estruturas de complexos proteína-proteína. Estes algoritmos constituem o método computacional denominado ancoramento molecular (AM). O desenvolvimento desta nova metodologia despertou grande interesse da comunidade científica acerca do ancoramento molecular de proteínas, o que permitiu a aplicação de abordagens interdisciplinares à modelagem, predição e compreensão deste tipo de interação. No entanto, a predição correta de complexos proteicos permanece um grande desafio na biologia estrutural, especialmente quando trata-se de sistemas do tipo antígeno-anticorpo [43, 44].

O primeiro algoritmo de ancoramento molecular destinado à predição de complexos proteicos, foi desenvolvido por Janin e Wodak em 1978 [45], dando início à criação de uma grande variedade de algoritmos de ancoramento molecular. De modo geral, o objetivo principal de todos esses algoritmos é,

“[...] prever o arranjo tridimensional de um complexo proteína-proteína a partir das coordenadas de suas moléculas componentes, sendo uma previsão precisa a que apontará a maioria dos contatos resíduos-resíduos envolvidos na interação do alvo.” (MOREIRA; FERNANDES; RAMOS, 2010, p.318, tradução nossa)[46].

Para cumprir este objetivo, cada metodologia de AM é composta por dois procedimentos principais: o algoritmo de busca e a função de pontuação.

O algoritmo de busca consiste na pesquisa de todas as possíveis orientações e conformações relativas entre o receptor e o ligante. Este procedimento envolve todas as possibilidades de movimentos no espaço 6D (3 graus de liberdade translacional e 3 graus de liberdade rotacional) do ligante ao redor do receptor (no caso mais comum, onde o ancoramento é do tipo rígido). A busca no espaço 6D, torna tal procedimento um tanto dispendioso, por isso a maioria dos *softwares* de ancoramento molecular utilizam métodos, tais como Monte Carlo, *hashing* geométrico ou transformada rápida de Fourier, para acelerar a busca configuracional [47, 48].

Uma função de pontuação é uma função matemática simplificada usada para predizer as forças de interação entre duas moléculas em uma determinada configuração. Elas são tradicionalmente divididas em três categorias: baseadas em campo de força, baseadas em conhecimento de pares e empírica. Dentro da metodologia de ancoramento molecular, essas funções têm como objetivo estimar a energia de ligação entre duas moléculas. Para isso, diferentes termos relacionados às propriedades físico-químicas das interações proteína-ligante são considerados e o cálculo da energia de ligação é feito de acordo com a abordagem aplicada [49].

Funções de pontuação baseadas em campo de força calculam a energia de ligação através da soma dos termos de interação individual, os quais são descritos por meio de um conjunto de parâmetros derivados de algum campo de força. Dentre os termos mais comumente utilizados estão: energias de Van de Waals (VDW), energias eletrostática, energias de interação ligada (comprimento, rotação, translação), etc [49, 47]. Um exemplo desse tipo de função é dado por

$$E_{\text{ligação}} = \sum_i \sum_j \left(\frac{A_{ij}}{d_{ij}^{12}} - \frac{B_{ij}}{d_{ij}^6} + \frac{q_i q_j}{\varepsilon(d_{ij}) d_{ij}} \right), \quad (4.1)$$

em que, d_{ij} representa a distância entre o átomo i da proteína e o átomo j do ligante, A_{ij} e B_{ij} são os parâmetros de VDW, q_i e q_j são as cargas atômicas e $\varepsilon(d_{ij})$ é a constante dielétrica dependente da distância.

Funções de pontuação baseadas em conhecimento de pares utilizam a relação inversa da distribuição de Boltzmann, como uma análise estatística para relacionar os pares de resíduos presentes na interação de complexos cristalográficos proteína-ligante. Dessa forma, a energia de ligação do complexo é calculada a partir dos dados de pares emparelhados. Para isso, esses dados são convertidos em um pseudo-potencial, conhecido como potencial de força médio (PMF), apresentado na Eq. 4.2. A soma destes potenciais, $w(r)$, fornece então a energia de ligação do complexo por meio da preferência geométrica dos pares de resíduos emparelhados [50, 51].

$$w(r) = -k_B T \ln[\rho(r)/\rho^*(r)], \quad (4.2)$$

em que, T é a temperatura absoluta, k_B é a constante de Boltzmann, $\rho(r)$ é a densidade numérica do par proteína-ligante à distância r no conjunto de treinamento, $\rho^*(r)$ é a densidade do par em um estado de referência onde as interações interatômicas são zero.

Funções de pontuação empírica baseiam-se no princípio de que a energia livre de ligação pode ser calculada por meio da soma de termos não correlacionados. Dessa forma, cada termo tem sua contribuição relativa dada por um coeficiente que é parametrizado de maneira a reproduzir os dados experimentais de afinidade de ligação de complexos

proteína-ligante existentes. A maneira mais simples de representar esse tipo de função é

$$\Delta G = \sum W_i \cdot \Delta G_i, \quad (4.3)$$

em que, W_i é o coeficiente de cada termo ΔG_i relacionado a alguma propriedade físico-química considerada, por exemplo, energias livres de solvatação ($\Delta G_{solv.}$), de interações de Van der Waals (ΔG_{vdw}), de ligações de hidrogênio ΔG_{ligH} [49].

As metodologias de AM são implementadas em diversos *softwares*, os quais diferem-se entre si de acordo com o algoritmo de busca e a função pontuação utilizados. Eles passam por um processo de teste e validação, o qual é realizado por meio da avaliação da qualidade dos modelos preditos. Este processo compreende os experimentos internacionais de Avaliação Crítica de Interações Preditas (CAPRI, do inglês *Critical Assessment of PRedicted Interactions*), os quais foram desenvolvidos de acordo com o modelo de Avaliação Crítica das Técnicas de Previsão da Estrutura Proteica (CASP, do inglês *Critical Assessment of Techniques for Protein Structure Prediction*). Nesses experimentos, os modelos gerados são avaliados de acordo com os critérios de L-RMSD (RMSD global), i-RMSD (RMSD de interface) e Fnat (fracão de resíduos nativos) estabelecidos pela CAPRI [46].

A maioria destes *softwares* foram desenvolvidos para o estudo de interações de pequenas moléculas em sítios de proteínas alvo. No entanto, avanços computacionais e científicos permitiram a ampliação para sistemas maiores, como complexos proteína-proteína. Entretanto, apesar do considerável progresso desta área, alguns problemas ainda permanecem relutantes, dado a complexidade estrutural das proteínas [43, 52].

Durante o processo de formação do complexo proteico é necessário que as interfaces de ambas moléculas estejam dispostas de maneira favorável, ou seja, em uma orientação e conformação altamente específicas. Isto significa que a estrutura de ambas moléculas estão inicialmente em um conformação não ligada (*unbound*) e, à medida que elas aproximam-se uma da outra ocorre uma mudança conformacional das cadeias principal e laterais. Ao final, as moléculas que formam o complexo estão em uma conformação diferente daquela apresentada antes da interação; este estado é denominado conformação ligada (*bound*). Este modelo de interação é denominado modelo de ajuste induzido e, foi proposto por Daniel Koshland em 1958 [53]. Esse modelo evidencia que ocorre uma adaptação mútua das duas proteínas durante a formação do complexo, sendo os resíduos de aminoácido da interface de interação os que sofrem maior perturbação.

Neste sentido, para que uma predição seja considerada perfeita, é preciso que a metodologia de ancoramento molecular forneça um modelo de complexo correspondente à conformação ligada de ambas proteínas; tendo como configuração de partida as estruturas nas conformações não ligadas. No entanto, dentre as milhares estruturas de proteínas disponíveis no banco de dados de proteínas (PDB, do inglês *Protein Data Bank*), apenas algumas centenas apresentam estruturas não ligadas com alta resolução. Em especial, para moléculas de anticorpo a maioria das estruturas disponíveis estão na conformação

ligada, pois complexos do tipo antígeno-anticorpo são complexos transientes, ou seja, são complexos que permanecem unidos durante um curto espaço de tempo [46].

Dessa forma, o primeiro desafio enfrentado pelo ancoramento molecular é a ausência de flexibilidade estrutural durante o procedimento de busca conformacional, realizada através do algoritmo de busca. Essa privação exclui conformações mais energeticamente favoráveis, pois torna a superfície de energia pouco representativa [54]. Além disso, é necessário que o complexo proteico de mais baixa energia livre seja encontrado. Entretanto, devido à complexidade envolvida nos fatores energéticos que regem as interações, esta tarefa não é nada fácil. É então que surge o segundo maior desafio do ancoramento molecular, a ineficiência das funções de pontuação em combinar todos os fatores energéticos a um reduzido custo computacional [55].

De modo geral, o impasse enfrentado pelo AM está na limitada capacidade de discriminação entre modelos precisos e imprecisos. Isto significa que não se pode confiar que o modelo de maior pontuação, ou seja, de energia mais negativa, será o modelo que mais se aproxima da conformação nativa. Este fato é ainda pior quando uma predição às cegas é realizada, ou seja, quando não se tem o conhecimento do complexo alvo [56].

No entanto, as falhas envolvidas no processo de discriminação podem ser superadas quando os modelos obtidos em um ancoramento rígido (ausência de flexibilidade) são submetidos a um pós-tratamento. Esse pós-tratamento tem como objetivo refinar ou repontuar os modelos de modo a realizar uma discriminação entre as configurações representativas e pouco representativas. Diversas abordagens são relatadas na literatura e, dentre elas está o refinamento a partir da inclusão de flexibilidade conformacional [57, 58]. Para isso, existem alguns *softwares*, tais como o Rosetta *online*, que utilizam algoritmos que permitem uma leve flexibilidade das cadeias principal e laterais através de uma busca local [59]. Além disso, informações experimentais podem auxiliar tanto na orientação quanto na validação do desempenho do ancoramento molecular. Essa abordagem já foi utilizada em alguns estudos de interação antígeno-anticorpo envolvendo a proteína de envelope (E) dos vírus da Dengue e Zika e, foi observado sucesso na discriminação de modelos mais representativos [60, 56, 61].

Neste sentido, no presente trabalho foi proposto um protocolo com o objetivo de selecionar o modelo de complexo mais representativo para o sistema estudado (22NS/DENV2). Esse protocolo é composto por uma série de procedimentos que englobam diferentes *softwares* de ancoramento molecular, incluindo flexibilidade conformacional e, informações de entalpia de ligação ($\Delta H_{lig.}$), calculada a partir do calor de formação (ΔH_f), o qual foi obtido por métodos quânticos semi-empíricos. Vale ressaltar que neste estudo, informações experimentais só foram utilizadas para avaliar a utilidade do protocolo sugerido. Detalhes sobre a sequência metodológica utilizada na construção desse protocolo estão apresentados no Capítulo 5.

4.2 Método quântico semi-empírico

Na mecânica quântica, a energia total de um sistema é dada pela soma das energias potencial e cinética entre núcleos e elétrons. A partir da aproximação de Born-Oppenheimer, o termo de energia eletrônica pode ser calculado separado do termo de energia nuclear. Dessa forma, a equação de Schrödinger pode ser descrita apenas para a parte eletrônica do átomo

$$\hat{H}_e \psi_e = E_e \psi_e, \quad (4.4)$$

em que, \hat{H}_e é o Hamiltoniano eletrônico, ψ_e é a função de onda mono-eletrônica e E_e é a energia eletrônica [62, 63, 64]. A solução da Eq. 4.4 pode ser encontrada a partir da combinação linear dos orbitais atômicos (LCAO do inglês, *Linear Combination of Atomic Orbitals*).

Em 1951, Roothaan [65] propôs que um orbital atômico (ψ_A) pode ser descrito como a combinação linear de um conjunto de funções de base (X_i),

$$\psi_A = \sum_i c_i X_i, \quad (4.5)$$

em que, X_i formam um conjunto completo de funções de base, c_i é o coeficiente de expansão e ψ_A é o orbital atômico de interesse. Ao aplicar esta ideia para obter a função de onda Ψ_M que descreve um orbital molecular, encontramos

$$\Psi_M = \sum_i C_i \psi_{A_i}, \quad (4.6)$$

sendo Ψ_M a soma dos coeficientes C_i e das funções de onda atômicas ψ_{A_i} que contribuem para formação do orbital molecular. Aqui, o coeficiente C_i representa a contribuição individual de cada função atômica ψ_{A_i} . A função de onda ψ_{A_i} é uma função mono-eletrônica obtida a partir da aproximação de Hartree-Fock [66].

Os orbitais moleculares (OM) dados a partir da função da Eq. 4.6 são obtidos por meio da aproximação de Hartree-Fock via método de campo auto-consistente iterativo (SCF, do inglês *Self-Consistent Field*). Esse método soluciona de maneira iterativa a equação de autovalores $F\Psi_M = E\Psi_M$ por meio da aproximação da matriz de Fock. Os elementos que compõem esta matriz são obtidas via integrais de troca que correspondem as diversas interações entre os elétrons que estão presentes nos átomos [66, 67, 68].

Para sistemas com um grande número de átomos, a solução dessas integrais torna os cálculos computacionalmente custoso. Para superar este problema vários pesquisadores, dentre eles Michael J. S. Dewar [69], desenvolveram novas técnicas com o objetivo de tornar a aproximação Hartree-Fock computacionalmente menos intensa, mas sem sacrificar drasticamente a precisão do método. A maneira mais eficiente encontrada foi negligenciar algumas integrais e utilizar aproximações através do emprego de funções paramétricas que

ajustam-se a dados experimentais. Este novo conjunto de técnicas constitui os métodos quânticos semi-empíricos.

Durante os anos, o formalismo e os parâmetros que constituem os métodos quânticos semi-empíricos foram aperfeiçoados e deram origem a diversos métodos padrões, tais como MNDO [69], AM1 [70], PM3 [71, 72], PM6 [73] e PM7 [74]. Atualmente, esses métodos encontram-se implementados em pacotes de *software* de domínio público, tais como o MOPAC (*Molecular Orbital PACKage*) [75], e são usados no cálculo da energia eletrônica para obter orbitais moleculares e calor de formação de moléculas em uma determinada geometria molecular, além de outras informações como espectro vibracional, propriedades termodinâmicas, dentre outras informações moleculares [76].

Dentre os métodos semi-empíricos atuais, o PM7, destaca-se devido às melhorias na descrição de algumas propriedades, como calor de formação, e a inclusão de uma parametrização melhorada para descrição de interações não covalentes [77].

De acordo com Nikita e colaboradores (2004) [78], as interações não covalentes estabelecidas em um complexo proteína-ligante estão diretamente relacionadas com a sua entalpia de ligação (ΔH_{lig}), ou seja, uma ΔH_{lig} favorável pode indicar que as interações formadas são relevantes para manutenção do complexo. Dessa forma, as interações não covalentes, envolvidas na formação do complexo, podem ser melhor caracterizadas por meio da entalpia de ligação.

Para obter a entalpia de ligação (ΔH_{lig}) é necessário calcular o calor de formação (ΔH_f) dos componentes do complexo. No *software* MOPAC [79], o ΔH_f é dado pela combinação da energia eletrônica (E_{elec}) e a da energia total de repulsão núcleo-núcleo (E_{nucl}) com a soma das energias envolvidas na atomização ($E_{isol}(A)$) e ionização ($\sum E_{atom}(A)$) das moléculas num estado gasoso, resultando na seguinte equação

$$\Delta H_f = (E_{elec} + E_{nucl}) - \sum E_{isol}(A) + \sum E_{atom}(A), \quad (4.7)$$

em que,

$$E_{elec} = \frac{1}{2} \sum_{\mu} \sum_{\nu} P_{\mu\nu} (H_{\mu\nu} + F_{\mu\nu}), \quad (4.8)$$

$$E_{nucl} = \sum_A \sum_{A>B} Z'_A Z'_B \langle s_A s_A | s_B s_B \rangle [1 + F(A) + F(B)]. \quad (4.9)$$

Na equação Eq. 4.8, os termos P, F e H são as matrizes de densidade, de Fock e de um elétron, respectivamente, e os subscritos μ e ν indicam que os orbitais atômicos são centrados em um átomo. As matrizes de densidade e de Fock são geradas pelo método SCF.

A Eq. 4.9 descreve a energia total de repulsão entre o núcleo A e o núcleo B a uma distância R_{AB} , com cargas efetivas Z'_A e Z'_B , em que $\langle s_A s_A | s_B s_B \rangle$ é a integral de repul-

são elétron-elétron e $F(A)$ e $F(B)$ são funções que dependem da distância internuclear, R_{AB} .

Como já dito, a construção da matriz de Fock demanda a resolução de várias integrais e, na metodologia semi-empírica o número e complexidade destas integrais diminuem. Mas, os cálculos SCF utilizam métodos de álgebra matricial na sua resolução. Isto implica que o número de operações aumenta à terceira potência do tamanho do sistema, ou seja, será necessário realizar N^3 operações para um sistema com N número de átomos. Desta forma, o tratamento de grandes sistemas, como é o caso das proteínas, seria inviável pelos métodos quânticos semi-empíricos. Uma maneira para solucionar isso, foi proposta por Stewart no desenvolvimento da abordagem MOZYME [80], disponível no *software* MOPAC.

Resumidamente, a abordagem MOZYME faz com que o procedimento padrão para a resolução das equações do método SCF seja substituído por um método de Orbitais Moleculares Localizados. Dessa maneira, as funções de base usadas para descrever os orbitais atômicos são substituídas por funções de Orbital Molecular Localizado (OML). Isto faz com que a construção da matriz de Fock seja realizada de maneira mais rápida, pois à medida que aumenta-se o número de átomos também é aumentado o número de OMLs e, em sistemas macromoleculares a maioria dos OMLs virtuais e ocupados ainda não têm átomos em comum, o que significa nesse caso que, a interação é automaticamente nula, diminuindo o número de termos a serem calculado [81, 80].

De modo geral, a abordagem semi-empírica, a partir do formalismo adotado pelos métodos *ab initio*, conserva a física essencial à descrição das propriedades de um sistema. Ao passo que as aproximações aplicadas contribuem para a rapidez do método e, boas parametrizações corrigem os erros introduzidos por essas aproximações. Dessa forma, a descrição de um sistema utilizando métodos semi-empíricos, tem visado principalmente grandes biomoléculas [82].

Neste sentido, a grande vantagem de utilizar os métodos quânticos semi-empíricos é que eles permitem o estudo de sistemas com um grande número de átomos, o que não é viável computacionalmente pelos métodos de primeiros princípios e, por considerar a estrutura eletrônica, estes métodos tornam-se muito mais versáteis do que os mecânicos moleculares no cálculo de propriedades termoquímicas [83, 68].

5 Metodologia

5.1 Ancoramento molecular

Na metodologia de ancoramento molecular, o objetivo principal foi obter o modelo de complexo mais representativo para o sistema 22NS/DENV2. Para atingir esse objetivo foi desenvolvido um protocolo composto por cinco etapas. Na primeira etapa, foi avaliado a capacidade de predição dos *softwares* escolhidos por meio de um re-ancoramento da estrutura cristalográfica 22NS/WNV (PDB.ID:4OII). A segunda etapa, foi destinada à obtenção de modelos de complexo para o sistema estudado, utilizando *softwares* online, baseados em diferentes algoritmos de busca global.

Na terceira etapa, foi realizada a seleção dos modelos de menor energia, uma vez que cada *software* gerou 1000 modelos; portanto, todas essas conformações foram classificadas de acordo com a energia de ligação, sendo a energia de ligação mínima o critério para seleção das 10 melhores configurações de cada *software*. A quarta etapa foi destinada à exclusão das configurações, dentre as selecionadas na terceira etapa, que não apresentaram susceptibilidade de interação devido à inacessibilidade do domínio β -roll do antígeno. Dessa forma, todas as configurações que apresentaram interação nessa região foram excluídas.

Por fim, na quinta etapa foi realizado o procedimento de pós-tratamento, e teve como objetivo avaliar o comportamento de estrutura nativa. Para avaliar esse comportamento foi realizado a análise do RMSD de interface (i-RMSD) dos modelos gerados a partir de cada configuração oriunda da quarta etapa. Todas essas etapas estão summarizadas na Tabela 1.

Neste estudo foram utilizados três *softwares* de ancoramento molecular: dois para obtenção de configurações e um para o refinamento dos modelos no pós-tratamento. O critério para a escolha destes programas foi, essencialmente, o método utilizado no algoritmo de busca. Os *softwares* ClusPro [84] e PatchDock [85] realizam busca global por meio dos métodos de transformada de Fourier e *hashing* geométrico, respectivamente. Enquanto, o Rosetta [59] *webserver* (ROSIE) realiza uma busca local, utilizando método Monte Carlo, e emprega um algoritmo que permite uma leve flexibilidade das cadeias principais e laterais da proteína. Por esses motivos, é que ele foi utilizado para o refinamento das configurações obtidas pelos *softwares* anteriores.

Todos os *softwares* de ancoramento molecular utilizados são acessíveis online nos seguintes sítios eletrônicos de domínio público: www.bioinfo3d.cs.tau.ac.il/PatchDock/, www.cluspro.bu.edu e www.rosie.graylab.jhu.edu.

Tabela 1 – Resumo das cinco etapas do protocolo proposto neste estudo.

Etapa	Softwares	Sistema	Objetivo
1 - Re-ancoramento	ClusPro PatchDock Rosetta	22NS/WNV	Avaliar capacidade de predição dos softwares para o sistema de interesse.
2 - Busca Global	ClusPro PatchDock	22NS/DENV2	Obter modelos de complexo para o sistema estudado.
3 - Modelos Top10	-	22NS/DENV2	Selecionar os 10 modelos de menor energia de cada software.
4 - Critério de Acessibilidade	-	22NS/DENV2	Excluir dentre dos 10 modelos de menor energia aqueles que apresentaram interação no domínio β -roll.
5 - Pós-tratamento	Rosetta	22NS/DENV2	Avaliar comportamento de estrutura nativa a partir da análise do RMSD de interface (i-RMSD).

Fonte: A autora.

5.1.1 Re-ancoramento e modelos iniciais: Etapas 1-4

A estrutura inicial da proteína não estrutural 1 do vírus da Dengue sorotipo 2 (NS1-DENV2) e do fragmento de união ao antígeno do anticorpo monoclonal antiNS1 (22NS) estudadas neste trabalho, foram obtidas do banco de dados de proteínas RCSB-PDB (do inglês, *Research Collaboratory for Structural Bioinformatics — Protein Data Bank*) identificadas pelo código do modelo cristalográfico 4O6B e 4OII, respectivamente.

No cálculo de ancoramento molecular, realizado no ClusPro e PatchDock, as configurações iniciais submetidas foram as estruturas cristalográficas do anticorpo 22NS (receptor) e do antígeno NS1-DENV2 (ligante). Em ambos, o algoritmo de busca global foi configurado para o modo anticorpo-antígeno, de modo a limitar a região de busca às CDRs do anticorpo.

De acordo com a etapa 3 do protocolo, em cada programa foram selecionadas como melhores *poses* as 10 configurações de menor energia de ligação, conforme padrão adotado em cálculos de ancoramento molecular. Em particular, para os modelos obtidos no PatchDock, o ordenamento das 1000 configurações foi realizado pelo *software* FireDock [86].

As etapas descritas acima foram empregadas, da mesma forma, à etapa 1 de re-ancoramento da estrutura 22NS/WNV. Neste caso, apenas a *pose* de menor energia de cada programa: WC01, do ClusPro e WP01, do PatchDock, foi selecionada.

Após a seleção realizada na etapa 3, os modelos foram submetidos à etapa 4 para avaliar a localização da interação. Isto foi feito de acordo com a área acessível para interação da estrutura da proteína solúvel (sNS1), a qual apresenta seus domínios β -roll

internalizados na forma hexamérica. Dessa forma, todos os modelos que apresentaram interação no domínio β -roll foram descartados. Ao final, dentre as 10 *poses* resultaram cinco provenientes do ClusPro, nomeados por DC00, DC01, DC04, DC06 e DC08, e três do PatchDock, nomeadas por DP07, DP08 e DP09.

5.1.2 Pós-tratamento: Etapa 5

O objetivo da etapa 5 foi realizar o pós-tratamento dos modelos obtidos de um ancoramento molecular rígido. Como discutido na Seção 4.1, a realização de um pós-tratamento faz-se necessária para que as falhas de discriminação entre modelos precisos e imprecisos sejam superadas. Portanto, neste estudo foi proposto a avaliação do comportamento de estrutura nativa a partir da análise do i-RMSD. Para isso, os modelos resultantes das etapas anteriores foram utilizados como conformação inicial para os cálculos de ancoramento molecular no servidor online Rosetta. Este procedimento foi utilizado para refinar os modelos selecionados (DC00, DC01, DC04, DC06, DC08, DP07, DP08 e DP09) por meio de uma nova avaliação.

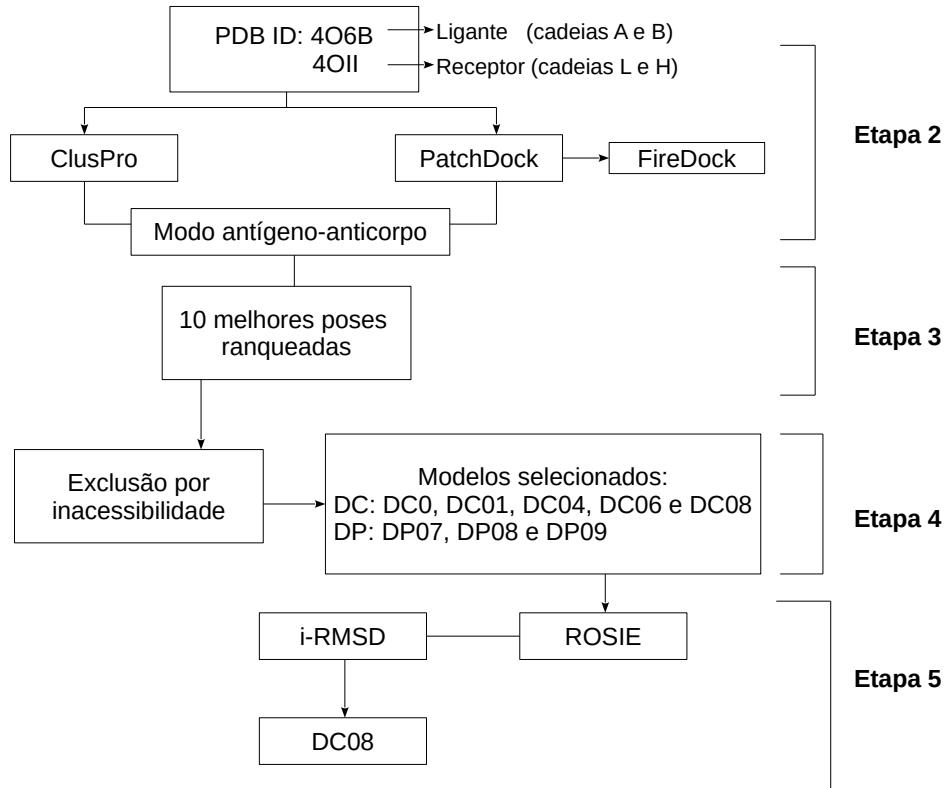
Para tanto, todos os modelos e a estrutura cristalográfica do complexo 22NS/WNV, foram submetidos ao protocolo de busca local, realizado a partir de uma amostragem translacional e rotacional da orientação das estruturas iniciais, seguido de um refinamento de alta resolução. Este refinamento conta com a atomização e minimização das cadeias laterais das estruturas e permite leve flexibilidade da cadeia principal.

Para realizar a nova avaliação proposta, inicialmente, foi utilizado o re-ancoramento da estrutura cristalográfica do complexo 22NS/WNV, para obter um perfil de comportamento de conformação nativa. Este perfil foi traçado de acordo com a quantidade de poses que se adequaram aos critérios de i-RMSD, estabelecidos pela CAPRI. Esse critério avalia o RMSD da cadeia principal dos resíduos da interface do complexo com relação à configuração inicial [87]. Os intervalos de valores de i-RMSD e suas categorias são: $0 < \text{i-RMSD} \leq 1$: EXCELENTE; $1 < \text{i-RMSD} \leq 2$: MODERADO e $2 < \text{i-RMSD} \leq 4$: ACEITÁVEL.

Sendo assim, os modelos gerados no Rosetta foram classificados como excelentes, moderados ou aceitos. E a quantidade desses modelos possibilitou a construção de um perfil de conformação nativa.

O fluxograma abaixo resume a metodologia de ancoramento molecular empregada no protocolo proposto.

Figura 9 – Fluxograma da metodologia de ancoramento proposta.



Fonte: A autora.

5.2 Cálculo da entalpia de ligação: MOPAC

As configurações obtidas como arquivo de saída nos programas de ancoramento molecular e as estruturas cristalográficas do PDB, não possuem hidrogênios em sua estrutura. Portanto, os modelos DC08, WC01, WP01 e a estrutura cristalográfica 22NS/WNV, foram submetidos a um pré-processamento para adição de hidrogênios. Esta tarefa foi realizada utilizando-se o módulo *pdb2gmx* implementado no programa GROMACS 2016.4 [88], considerando a forma de protonação dos resíduos de histidina (HIS) como N101.

Adicionalmente, para evitar que pequenos erros nas coordenadas dos átomos da proteína contribuam com energias não realistas no cálculo de energia, as estruturas foram otimizadas. Dessa forma, foi realizado apenas a otimização dos átomos de hidrogênios (OPT-H), utilizando o método PM7, juntamente com a aproximação MOZYME, usando o modelo COSMO [89] como um modelo de solvente implícito e mantendo os átomos pesados fixos nas suas posições cristalográficas (NOOPT). Considerando a solvatação em água, a constante dielétrica utilizada foi EPS=78,4 à T= 25°C e, o critério para uma geometria otimizada foi configurado em GNORM=5.

Utilizando os modelos previamente otimizados, o cálculo do ΔH_f , foi realizado pelo o método PM7, juntamente com a abordagem MOZYME e solvatação implícita, da mesma maneira como foi para a otimização. Adicionalmente, nesta etapa foi solicitado a rea-

lização de apenas um cálculo de campo auto-consistente, 1SCF, para corrigir possíveis erros no calor de formação. Estes erros podem surgir no uso da abordagem MOZYME e referem-se à perda de ortogonalidade dos orbitais moleculares localizados. Ao realizar apenas um cálculo SCF para a estrutura otimizada esses erros são então superados [75].

Por fim, a entalpia de ligação (ΔH_{lig}) de cada complexo foi calculada de acordo com a equação proposta por Nikitta e colaboradores (2004) [78], expressa por

$$\Delta H_{lig} = \Delta H_{f(complexo)} - (\Delta H_{f(anticorpo)} + \Delta H_{f(antigeno)}) , \quad (5.1)$$

em que cada um dos valores de ΔH_f para o complexo, anticorpo e antígeno foram obtidos a partir dos cálculos semi-empíricos realizados no *software* MOPAC 2016.

6 Resultados e Discussões

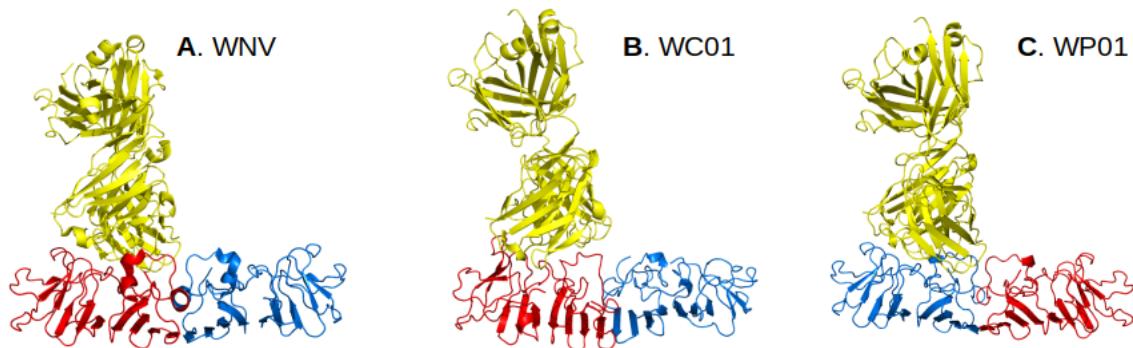
6.1 Ancoramento molecular

6.1.1 Re-ancoramento e modelos iniciais: Etapas 1-4

Re-ancoramento: WNV/22NS

As *poses* de menor energia do sistema WNV/22NS obtidos no re-ancoramento estão expostas na Figura 10, em que a estrutura do Fab do anticorpo, composta pelas cadeias L (leve) e H (pesada), estão na coloração amarela e, as cadeias A e B da estrutura do antígeno (WNV), em vermelho e azul, respectivamente.

Figura 10 – Representações cartoon da (A) estrutura cristalográfica do sistema WNV/22NS e, das *poses* de menor energia obtidas por re-ancoramento (B) pelo ClusPro e (C) PatchDock.



Fonte: A autora, Pymol [90].

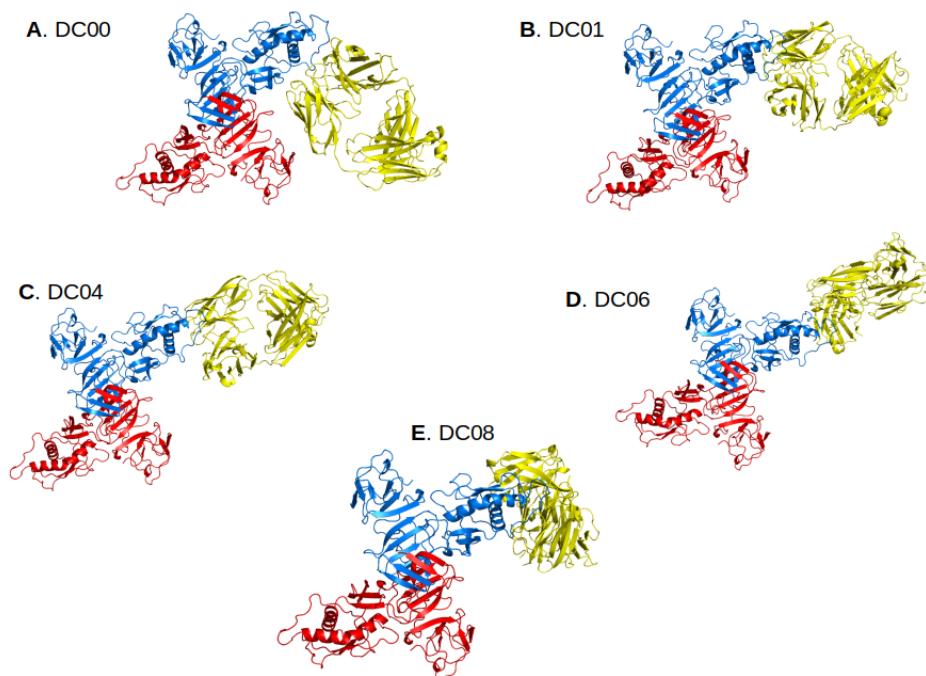
Visualmente, foi possível verificar que ambos programas retornaram configurações de complexos que apresentam orientação semelhante à encontrada na estrutura cristalográfica. A princípio, essa semelhança pode ser atribuída ao fato de que, ambos *softwares* previram corretamente a região de interação. Ela está localizada na porção do domínio β -ladder, mais especificamente, na região oposta àquela formada pelas estruturas de folhas β , conhecida como superfície de alça.

Considerando-se apenas o domínio de interação, os resultados do re-ancoramento mostraram a capacidade, de ambos *softwares*, de predizer corretamente a região de interação para o complexo antígeno-anticorpo. No entanto, uma discussão mais detalhada da interface de interação está apresentada na Seção 6.1.2.

Sistema estudado: DENV2/22NS

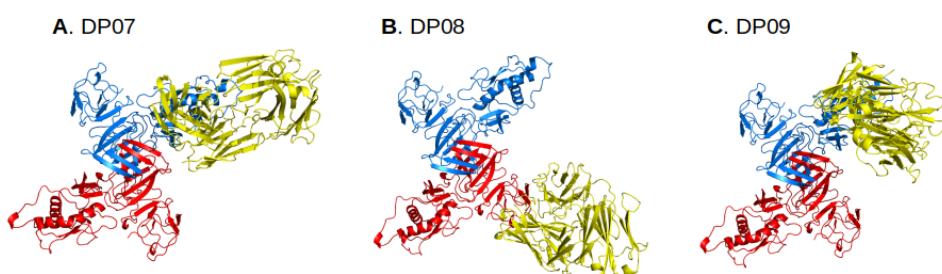
As configurações do sistema DENV2/22NS após a seleção realizada na etapa 4, dos softwares ClusPro e PatchDock, estão expostas nas Figuras 11 e 12, respectivamente. Em ambas figuras, a estrutura do Fab do anticorpo, composta pelas cadeias L (leve) e H (pesada), estão na coloração amarela e, as cadeias A e B da estrutura do antígeno (DENV2), em vermelho e azul, respectivamente.

Figura 11 – Representação em *cartoon* das *poses* de menor energia selecionadas do *web-server* Cluspro 2.0, ordenadas em forma crescente de energia A-E.



Fonte: A autora, Pymol [90].

Figura 12 – Representação em *cartoon* das *poses* de menor energia selecionadas do *web-server* PatchDock, ordenadas em forma crescente de energia A-C.



Fonte: A autora, Pymol [90].

Todas as *poses* apresentaram como local de interação a porção do domínio *Wing*. Este domínio é constituído pelos resíduos a.a. 38-151. Nesta região já foram identificados três

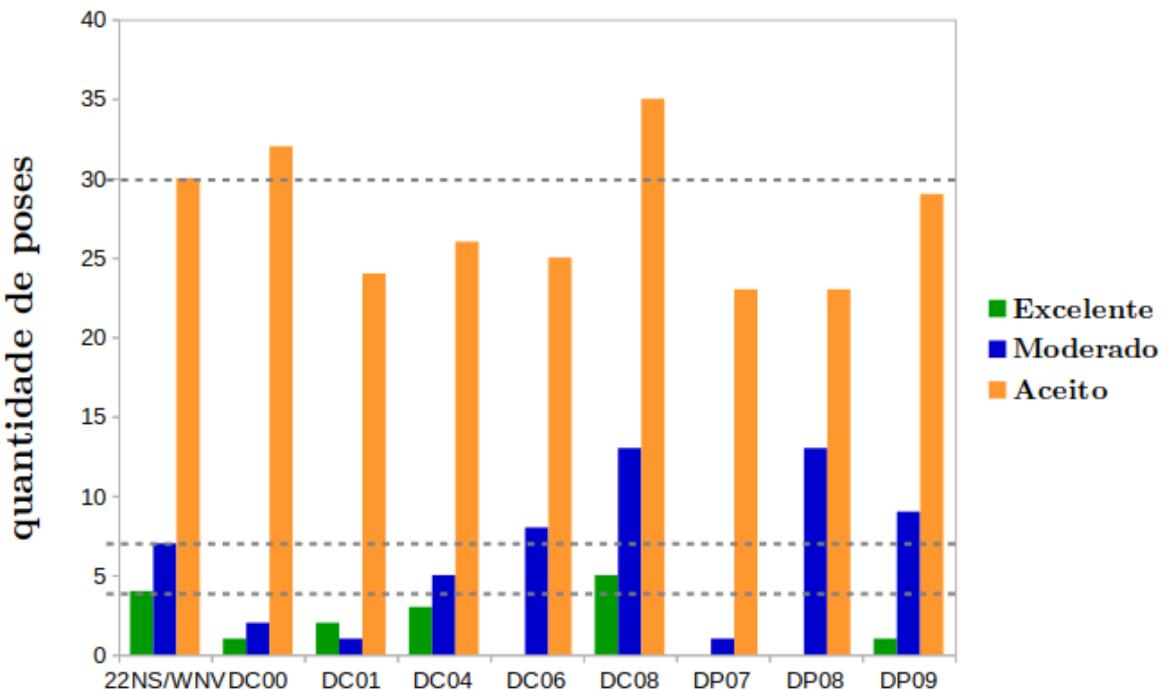
possíveis epítOTOS, evidenciando que a orientação predita, por ambos *softwares*, trata-se de uma possível região de interação conforme estudos in silico e in vitro encontrados na literatura [21, 91, 92, 93].

6.1.2 Pós-tratamento: Etapa 5

De posse das mil configurações geradas a partir do Rosetta Webserver e de seus respectivos valores de i-RMSD um perfil de comportamento de conformação nativa foi obtido.

Na Figura 13 é mostrada a relação entre a quantidade de *poses*, para cada modelo, que adequaram-se aos critérios de i-RMSD (excelente, moderado e aceito). O perfil utilizado como referência foi traçado de acordo com os resultados de re-ancoramento da estrutura cristalográfica WNV/22NS. Este perfil é dado pelo conjunto das três primeiras barras do gráfico (22NS/WNV) na Figura 13.

Figura 13 – Quantidade de *poses* por modelo de acordo com o critério estabelecido pela CAPRI.



Fonte: A autora.

Como dito na Seção 2.2, um complexo está em sua conformação nativa quando a estrutura tridimensional estabelecida por ele encontra-se em um estado de maior estabilidade termodinâmica. Esta estabilidade está diretamente relacionada com a energia livre de ligação do complexo, calculada através da função de pontuação implementada nos *softwares* de ancoramento molecular.

Desse modo, a análise da etapa 5 foi realizada baseada na seguinte premissa: se uma pose apresenta $i\text{-RMSD} \geq 4$ foi necessária mudança conformacional significativa para que o complexo atingisse energia de ligação favorável. Considerando que a estrutura inicial é uma estrutura nativa, uma elevada mudança na conformação da interface sugere que, esta configuração não encontra-se em sua conformação nativa.

Partindo desse pressuposto e sabendo que a estrutura cristalográfica WNV/22NS está em sua conformação nativa, a análise da quantidade de modelos que adequaram-se ao critério de $i\text{-RMSD}$ da CAPRI pode fornecer um perfil de comportamento de uma conformação nativa. Dessa forma, é esperado que a configuração mais representativa seja aquela que apresenta quantidades de *poses* semelhantes ao sistema de comparação [87]. Dentre todos os modelos analisados, o DC08 foi o que melhor ajustou-se a esse critério. Como pode ser percebido, com a ajuda das linhas tracejadas, a quantidade de modelos excelentes, moderados e aceitos, é semelhante àquela obtida no perfil do modelo cristalográfico. Em contrapartida, o modelo DP07 foi o que mais distanciou-se ao apresentar quantidades abaixo da linha de referência.

6.2 Análise da viabilidade do procedimento metodológico

A análise de comportamento de estrutura nativa foi validada de duas maneiras: (a) a partir dos valores de entalpia de ligação, obtidos através de cálculos quânticos semi-empíricos e (b) pela análise da interface de interação predita.

(a) Entalpia de ligação

Os valores de ΔH_{lig} para o modelo cristalográfico WNV/22NS e para os modelos WC01, WP01 E DC08 estão apresentados na Tabela 2.

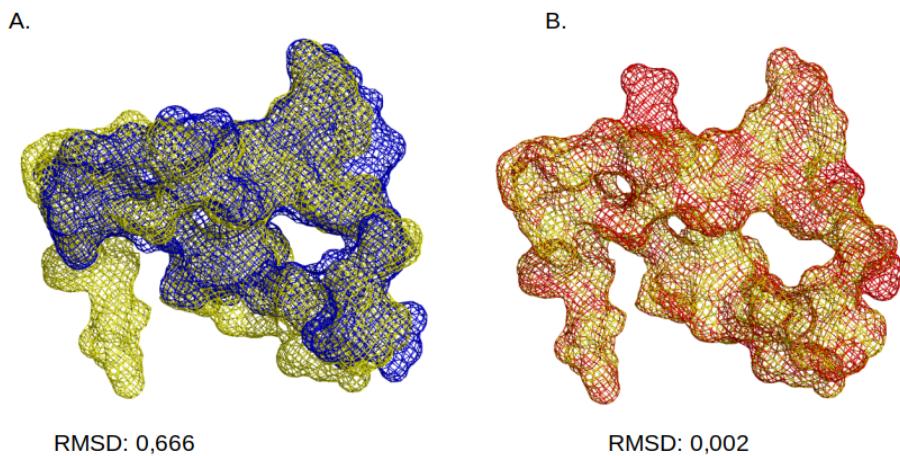
Tabela 2 – Valores de ΔH_{lig} , calculados a partir dos valores de ΔH_f obtidos no programa MOPAC 2016.

Complexo	$\Delta H_{lig}(kJ\ mol^{-1})$
WNV/22NS	-620,95
WC01	-614,71
WP01	-44,81
DC08	-460,07

Fonte: A autora.

Para melhor compreensão dos valores de ΔH_{lig} , foi realizada a comparação das interfaces de interação dos complexos, apresentada na Figura 14.

Figura 14 – Representação em malha da sobreposição da interface de interação da estrutura cristalográfica 22NS/WNV, em coloração amarela, com os modelos (A) WC01, em azul, e (B) WP01, em vermelho.

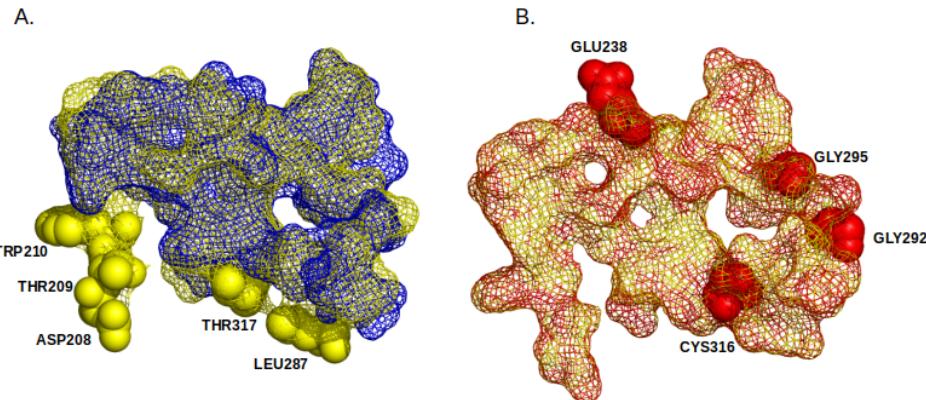


Fonte: A autora, Appa [94] e Pymol [90].

Analizando-se os valores ΔH_{lig} , Tabela 2 e os valores de RMSD, notou-se uma controvérsia. Para o complexo WC01, de maior RMSD, o valor ΔH_{lig} está na mesma ordem de grandeza do valor para estrutura cristalográfica 22NS/WNV. Enquanto que, para o WP01, de menor RMSD, o valor de ΔH_{lig} distanciou-se daquele obtido para a estrutura cristalográfica. Tal evento pode sugerir que, os resíduos preditos no modelo WC01 são mais significativos para a formação do complexo e que, alguns resíduos preditos no WP01 contribuem negativamente para a formação do complexo.

Dante do exposto, foi realizada uma análise dos resíduos ausentes ou presentes na interface de interação, como mostrado na Figura 15. Na figura, esses resíduos estão identificados pelos código e o identificador (ex: CYS316 - resíduo de cisteína localizado na posição 316 da sequência FASTA da proteína).

Figura 15 – Representação em malha da sobreposição da interface de interação da estrutura cristalográfica 22NS/WNV, em coloração amarela, com os modelos (A) WC01, em azul, com resíduos em esferas amarelas e (B) WP01, em vermelho, com resíduos em esferas vermelhas.



Fonte: A autora, Appa [94] e Pymol [90].

Foi observado que os resíduos CYS316, GLY292, GLY295 E GLU238 (Figura.15-B) estão presentes apenas na superfície de interação do modelo WP01. Isso pode indicar que, a baixa estabilidade deste complexo pode estar associada à presença desses resíduos na interface de interação. Por outro lado, os resíduos THR317, LEU287, TRP210, THR209 e ASP208, presentes apenas na superfície de interação do complexo WNV/22NS, sugere que tais resíduos pouco contribuem para a estabilidade do complexo.

Neste sentido, o *software* ClusPro mostrou melhor desempenho na predição de resíduos que interagem mais significativamente com o anticorpo 22NS. Além disso, foi possível verificar que o valor de ΔH_{lig} obtido para o modelo DC08 é comparável ao valor obtido para a estrutura experimental 22NS/WNV. Isto indica que o DC08 é um candidato em potencial para o complexo 22NS/DENV2 em sua conformação nativa.

(b) Interface de interação

Inicialmente, foi utilizado o AppA [94] (<http://mspc.bii.a-star.edu.sg/minhn/appa.html>) para obtenção das interfaces^d. Estas interfaces estão apresentadas na Figura 16 e o Número de Resíduos que Pertencem a Epítopo (NRPE) está exposto na Tabela 3.

Como dito no Capítulo 3, a estrutura dimérica da NS1 é formada por dois monômeros da NS1, que enovelam-se em três diferentes domínios: *β-roll*, *wing* e *β-ladder*. O domínio *β-roll* é formado por a.a. 1-29; o *wing* por a.a. 38-151 e *β-roll* por a.a. 181-352. Os domínios de uma proteína apresentam ordenamento espacial de acordo com a função

^d A interface de interação é obtida considerando-se os resíduos de contato. Os resíduos de contato são identificados como resíduos de uma estrutura de anticorpo (antígeno) cujas áreas de superfície acessível a solvente (ASA) mudam com a formação de seu complexo anticorpo-antígeno e estão a 6 Å do epítopo complexado (paratopo). [95]

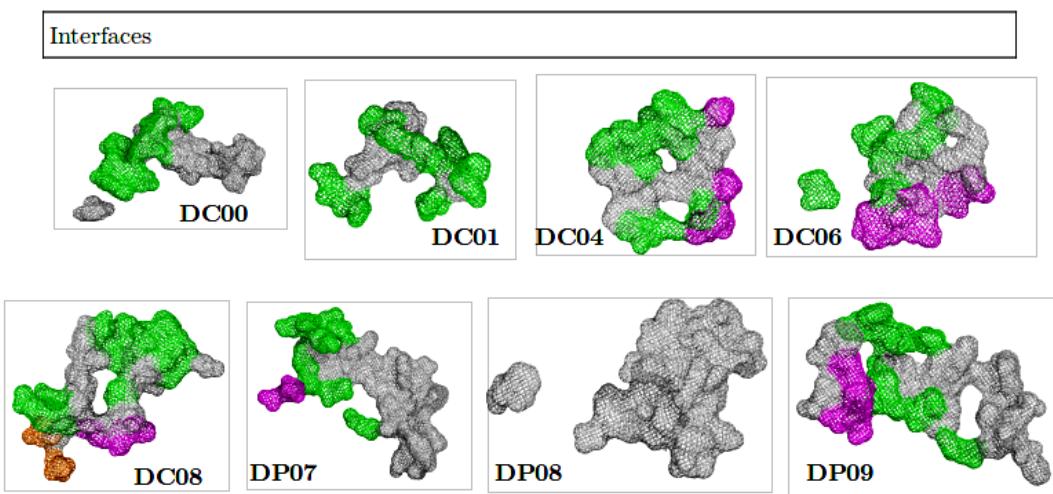
biológica exibida pela proteína. Portanto, no caso de moléculas antigênicas, os domínios são regiões susceptíveis à presença de epítópos.

Estudos experimentais realizados por Falconar [93] e colaboradores em 1994, com uso de anticorpos monoclonais, relevaram a presença de epítópos 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. Posteriormente Jian [92] e colaboradores (2010), utilizando anticorpos policlonais, identificaram uma variedade de epítópos lineares, os quais estão compreendidos nas 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.

No entanto, alguns desses resíduos encontram-se em regiões inacessíveis quando a NS1 está na forma hexamérica. Estudos teórico e experimental, realizados por Jones [91] e colaboradores (2017) avaliaram que os resíduos que realmente estão acessíveis para interação são: a.a. 97-116; a.a. 125-134; a.a. 143-151 e a.a. 170-178. Dentre eles, as sequências que englobam o domínio *wing* são as três primeiras. Estas sequências estão em concordância com outro estudo, por Freire [21] e colaboradores (2017), que identificaram as seguintes sequências de epítópos pertencentes ao domínio *wing*: a.a. 50-65; a.a. 74-88 e a.a. 108-129.

Considerando-se apenas os epítópos localizados no domínio *wing* e, englobando o máximo de resíduos pertencentes a eles, foram consideradas as seguintes sequências a.a. 50-116; a.a. 122-134 e a.a. 143-151, nomeadas por Epítopo 1 (E1), Epítopo 2 (E2) e Epítopo 3 (E3), respectivamente. Na análise das interfaces, apenas os resíduos classificados como expostos excluindo-se os intermediários e enterrados.

Figura 16 – Interfaces de interação dos modelos do complexo 22NS/DENV2 com destaque aos resíduos pertencentes aos epítópos E1, na coloração verde, E2, em magenta e E3, em laranja.



Fonte: A autora, Appa [94] e Pymol [90]

Tabela 3 – Tabela com o Número de Resíduos Pertencentes a Epítocos (NRPE), em cada modelo gerado para o complexo 22NS/DENV2.

Modelo	NRPE	Modelo	NRPE
DC00	9	DC08	17
DC01	8	DP07	10
DC04	13	DP08	0
DC06	13	DP09	11

Fonte: A autora.

Analisando as imagens da Figura 16 e os valores na Tabela 3, é possível verificar que o DC08 é o modelo que apresenta a maior quantidade de resíduos pertencentes a epítocos. Ou seja, a interface predita apresenta resíduos realmente significativos para uma interação do tipo antígeno-anticorpo. Enquanto que, o modelo DP08 não foi capaz de prever nenhum resíduo.

De acordo com os resultados da análise do gráfico da Figura 13, o DC08 foi o modelo que melhor se ajustou ao perfil traçado para a estrutura cristalográfica. Enquanto que, o modelo DP07 foi o que mais distanciou-se dos critérios. Disso pode-se inferir que, o procedimento metodológico, proposto neste trabalho, foi capaz de selecionar o modelo mais representativo, dentre todos. Apesar do julgamento do modelo menos representativo ter sido menos preciso e, portanto, ter necessitado de mais metodologias complementares, este protocolo mostrou-se promissor na obtenção de modelos representativos para sistemas antígeno-anticorpo semelhantes.

7 Considerações finais

O presente trabalho permitiu obter um modelo representativo do complexo entre a proteína não estrutural da Dengue sorotipo 2 (NS1-DENV2) e o anticorpo monoclonal antiNS1 (22NS). A partir dele, foi possível verificar que o local de interação encontra-se na região do domínio *wing* e comprehende três tipos de epítopos.

Além disso, o protocolo proposto mostrou-se promissor na seleção de modelos precisos, fornecendo modelos em concordância com estudos experimentais. Dessa forma, tal protocolo mostrou ser útil para a avaliação da região e da magnitude da afinidade de interação em sistemas antígeno-anticorpo com características semelhantes. Sendo, portanto, um estudo prévio acerca da afinidade de interação do anticorpo antiNS1 e com a proteína NS1 das flaviroses.

Como perspectivas futuras, o mesmo protocolo será utilizado para avaliar as interações da proteína NS1 de outras flaviroses, como a Zika, com o objetivo de avaliar a capacidade de discriminação entre flaviroses por meio dos princípios analíticos dos biossensores, em particular, os baseados na técnica de Microscopia de Força Atômica (AFM).

Referências

- 1 AZEREDO, E. L. et al. Clinical and laboratory profile of Zika and Dengue infected patients: lessons learned from the co-circulation of Dengue, Zika and Chikungunya in Brazil. **PLoS currents**, Public Library of Science, v. 10, 2018. Disponível em: <<https://doi.org/10.1371/currents.outbreaks.0bf6aeb4d30824de63c4d5d745b217f5>>.
- 2 WANG, D. et al. A mutation identified in neonatal microcephaly destabilizes Zika virus NS1 assembly *in vitro*. **Scientific reports**, Nature Publishing Group, v. 7, p. 425–480, 2017. Disponível em: <<https://doi.org/10.1038/srep42580>>.
- 3 MANSUY, J. M. et al. Zika virus: high infectious viral load in semen, a new sexually transmitted pathogen? **The Lancet Infectious Diseases**, Elsevier, v. 16, n. 4, p. 405, 2016. Disponível em: <[https://doi.org/10.1016/S1473-3099\(16\)00138-9](https://doi.org/10.1016/S1473-3099(16)00138-9)>.
- 4 SAMARASEKERA, U.; TRIUNFOL, M. Concern over Zika virus grips the world. **The Lancet**, Elsevier, v. 387, n. 10018, p. 521–524, 2016. Disponível em: <[https://doi.org/10.1016/S0140-6736\(16\)00257-9](https://doi.org/10.1016/S0140-6736(16)00257-9)>.
- 5 FARES, R. C. et al. Epidemiological scenario of Dengue in Brazil. **BioMed research international**, Hindawi, v. 2015, 2015. Disponível em: <<https://doi.org/10.1155/2015/321873>>.
- 6 MARTINS, L. C. et al. First isolation of west nile virus in brazil. **Mem Inst Oswaldo Cruz, Rio de Janeiro**, v. 114, p. e180332, 2019.
- 7 NERY, L. A. S. S. et al. 6. análise histórica de casos de dengue no brasil. **Revista Científica UMC**, v. 3, n. 2, p. 1–13, 2018.
- 8 RASTOGI, M.; SHARMA, N.; SINGH, S. K. Flavivirus NS1: a multifaceted enigmatic viral protein. **Virology journal**, Springer, v. 13, n. 1, p. 131, 2016. Disponível em: <<https://doi.org/10.1186/s12985-016-0590-7>>.
- 9 SCATURRO, P. et al. Dengue virus non-structural protein 1 modulates infectious particle production via interaction with the structural proteins. **PLoS pathogens**, Public Library of Science, v. 11, n. 11, p. e1005277, 2015. Disponível em: <<https://doi.org/10.1371/journal.ppat.1005277>>.
- 10 MELO, K. M. S. d. et al. **Desenvolvimento e validação de insumos para diagnóstico de infecções pelo vírus da dengue**. Tese (Doutorado) — Centro de Pesquisas Aggeu Magalhães, 2012.
- 11 GUBLER, D. J. Dengue and dengue hemorrhagic fever. In: ELSEVIER. **Seminars in Pediatric Infectious Diseases**. 1997. v. 8, n. 1, p. 3–9. Disponível em: <[https://doi.org/10.1016/S1045-1870\(97\)80003-9](https://doi.org/10.1016/S1045-1870(97)80003-9)>.
- 12 SAMUEL, P. P.; TYAGI, B. Diagnostic methods for detection & isolation of dengue viruses from vector mosquitoes. **Indian journal of medical research**, Indian Council of Medical Research, v. 123, n. 5, p. 615, 2006.

- 13 INNIS, B. et al. An enzyme-linked immunosorbent assay to characterize dengue infections where dengue and japanese encephalitis co-circulate. **The American journal of tropical medicine and hygiene**, ASTMH, v. 40, n. 4, p. 418–427, 1989. Disponível em: <<https://doi.org/10.4269/ajtmh.1989.40.418>>.
- 14 HUHTAMO, E. et al. Novel flaviviruses from mosquitoes: Mosquito-specific evolutionary lineages within the phylogenetic group of mosquito-borne flaviviruses. **Virology**, Elsevier, v. 464, p. 320–329, 2014. Disponível em: <<https://doi.org/10.1016/j.virol.2014.07.015>>.
- 15 CHAMBERS, T. J. et al. Flavivirus genome organization, expression, and replication. **Annual review of microbiology**, Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0139, USA, v. 44, n. 1, p. 649–688, 1990.
- 16 ALCON, S. et al. Enzyme-linked immunosorbent assay specific to dengue virus type 1 nonstructural protein ns1 reveals circulation of the antigen in the blood during the acute phase of disease in patients experiencing primary or secondary infections. **Journal of clinical microbiology**, Am Soc Microbiol, v. 40, n. 2, p. 376–381, 2002. Disponível em: <<https://doi.org/10.1128/JCM.40.02.376-381.2002>>.
- 17 FELIX, A. C. et al. Low sensitivity of ns1 protein tests evidenced during a dengue type 2 virus outbreak in santos, brazil, in 2010. **Clin. Vaccine Immunol.**, Am Soc Microbiol, v. 19, n. 12, p. 1972–1976, 2012. Disponível em: <<https://doi.org/10.1128/CVI.00535-12>>.
- 18 LIBRATY, D. H. et al. High circulating levels of the dengue virus nonstructural protein ns1 early in dengue illness correlate with the development of dengue hemorrhagic fever. **The Journal of infectious diseases**, The University of Chicago Press, v. 186, n. 8, p. 1165–1168, 2002. Disponível em: <<https://doi.org/10.1086/343813>>.
- 19 SHEN, W.-F. et al. Improving dengue viral antigens detection in dengue patient serum specimens using a low ph glycine buffer treatment. **Journal of Microbiology, Immunology and Infection**, Elsevier, v. 50, n. 2, p. 167–174, 2017. Disponível em: <<https://doi.org/10.1016/j.jmii.2015.05.008>>.
- 20 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**, Am Soc Microbiol, v. 38, n. 3, p. 1053–1057, 2000. Disponível em: <<https://doi.org/10.1128/JCM.38.3.1053-1057.2000>>.
- 21 FREIRE, M. C. et al. Mapping putative B-cell Zika virus NS1 epitopes provides molecular basis for anti-NS1 antibody discrimination between Zika and Dengue viruses. **Acs Omega**, ACS Publications, v. 2, n. 7, p. 3913–3920, 2017. Disponível em: <<https://doi.org/10.1021/acsomega.7b00608>>.
- 22 ANTUNES, P. et al. Quantification of ns1 dengue biomarker in serum via optomagnetic nanocluster detection. **Scientific Reports**, Nature Publishing Group, v. 5, n. 16145, p. 1–10, 2015. Disponível em: <<https://doi.org/10.1038/srep16145>>.
- 23 BRANDEN, C. I.; TOOZE, J. **Introduction to protein structure**. Garland Science, 2012. Disponível em: <<https://doi.org/10.1201/9781136969898>>.
- 24 LEHNINGER, A. L. et al. **Principles of biochemistry**. [S.l.]: Macmillan, 2005.

- 25 NUSSINOV, R.; SCHREIBER, G. **Computational protein-protein interactions.** CRC Press, 2009. Disponível em: <<https://doi.org/10.1201/9781420070071>>.
- 26 ABBAS, A. K.; LICHTMAN, A. H.; PILLAI, S. **Imunologia celular e molecular.** [S.l.]: Elsevier Brasil, 2008.
- 27 RANGWALA, H.; KARYPIS, G. **Introduction to protein structure prediction: methods and algorithms.** John Wiley & Sons, 2011. v. 18. Disponível em: <<https://doi.org/10.1002/9780470882207>>.
- 28 CORNELL, B. **Overview: aminoacids.** 2016. Disponível em: <<http://ib.bioninja.com.au>>. Acesso em: 5 de Outubro de 2019.
- 29 AKEY, D. L. et al. Flavivirus NS1 crystal structures reveal a surface for membrane association and regions of interaction with the immune system. **Science (New York, NY)**, NIH Public Access, v. 343, n. 6173, p. 881, 2014. Disponível em: <<https://doi.org/10.1126/science.1247749>>.
- 30 BROWN, W. C. et al. Extended surface for membrane association in Zika virus NS1 structure. **Nature structural & molecular biology**, Nature Publishing Group, v. 23, n. 9, p. 865, 2016. Disponível em: <<https://doi.org/10.1038/nsmb.3268>>.
- 31 HUBER, R. Structural basis for antigen-antibody recognition. **Science**, American Association for the Advancement of Science, v. 233, p. 702–704, 1986. Disponível em: <<https://doi.org/10.1126/science.2426777>>.
- 32 NISONOFF, A.; HOPPER, J. E.; SPRING, S. B. **The antibody molecule.** [S.l.]: Academic Press, 2014.
- 33 BARCLAY, A. N. Ig-like domains: evolution from simple interaction molecules to sophisticated antigen recognition. **Proceedings of the National Academy of Sciences**, National Acad Sciences, v. 96, n. 26, p. 14672–14674, 1999. Disponível em: <<https://doi.org/10.1073/pnas.96.26.14672>>.
- 34 SELA-CULANG, I.; KUNIK, V.; OFRAN, Y. The structural basis of antibody-antigen recognition. **Frontiers in immunology**, Frontiers, v. 4, p. 302, 2013. Disponível em: <<https://doi.org/10.3389/fimmu.2013.00302>>.
- 35 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**, Elsevier, v. 98, n. 2, p. 192–208, 2013. Disponível em: <<https://doi.org/10.1016/j.antiviral.2013.03.008>>.
- 36 CRIVIANU-GAITA, V.; THOMPSON, M. Aptamers, antibody scfv, and antibody Fab'fragments: An overview and comparison of three of the most versatile biosensor biorecognition elements. **Biosensors and Bioelectronics**, Elsevier, v. 85, p. 32–45, 2016. Disponível em: <<https://doi.org/10.1016/j.bios.2016.04.091>>.
- 37 TRILLING, A. K.; BEEKWILDER, J.; ZUILHOF, H. Antibody orientation on biosensor surfaces: a minireview. **Analyst**, Royal Society of Chemistry, v. 138, n. 6, p. 1619–1627, 2013. Disponível em: <<https://doi.org/10.1039/c2an36787d>>.

- 38 SIMÃO, E. P. et al. Biosensor based on cysteine monolayer and monoclonal antibody for specific detection of aflatoxin B1 in rice. **Journal of the Brazilian Chemical Society**, SciELO Brasil, v. 27, n. 6, p. 1040–1047, 2016. Disponível em: <<https://doi.org/10.5935/0103-5053.20150361>>.
- 39 GRANGER, D. et al. Serodiagnosis of dengue virus infection using commercially available antibody and NS1 antigen ELISAs. **Diagnostic microbiology and infectious disease**, Elsevier, v. 88, n. 2, p. 120–124, 2017. Disponível em: <<https://doi.org/10.1016/j.diagmicrobio.2017.03.015>>.
- 40 WU, G. et al. Bioassay of prostate-specific antigen (PSA) using microcantilevers. **Nature biotechnology**, Nature Publishing Group, v. 19, n. 9, p. 856, 2001. Disponível em: <<https://doi.org/10.1038/nbt0901-856>>.
- 41 PEDOTTI, M. et al. Computational docking of antibody-antigen complexes, opportunities and pitfalls illustrated by influenza hemagglutinin. **International journal of molecular sciences**, Molecular Diversity Preservation International, v. 12, n. 1, p. 226–251, 2011. Disponível em: <<https://doi.org/10.3390/ijms12010226>>.
- 42 RODRIGUES, J. P.; BONVIN, A. M. Integrative computational modeling of protein interactions. **The FEBS journal**, Wiley Online Library, v. 281, n. 8, p. 1988–2003, 2014. Disponível em: <<https://doi.org/10.1111/febs.12771>>.
- 43 EZKURDIA, I. et al. Progress and challenges in predicting protein–protein interaction sites. **Briefings in bioinformatics**, Oxford University Press, v. 10, n. 3, p. 233–246, 2009. Disponível em: <<https://doi.org/10.1093/bib/bbp021>>.
- 44 PONS, C. et al. Present and future challenges and limitations in protein–protein docking. **Proteins: Structure, Function, and Bioinformatics**, Wiley Online Library, v. 78, n. 1, p. 95–108, 2010. Disponível em: <<https://doi.org/10.1002/prot.22564>>.
- 45 JANIN, J. et al. Conformation of amino acid side-chains in proteins. **Journal of molecular biology**, v. 125, n. 3, p. 357–386, 1978. Disponível em: <[https://doi.org/10.1016/0022-2836\(78\)90408-4](https://doi.org/10.1016/0022-2836(78)90408-4)>.
- 46 MOREIRA, I. S.; FERNANDES, P. A.; RAMOS, M. J. Protein–protein docking dealing with the unknown. **Journal of computational chemistry**, Wiley Online Library, v. 31, n. 2, p. 317–342, 2010.
- 47 HUANG, S.-Y.; ZOU, X. Advances and challenges in protein-ligand docking. **International journal of molecular sciences**, Molecular Diversity Preservation International, v. 11, n. 8, p. 3016–3034, 2010. Disponível em: <<https://doi.org/10.3390/ijms11083016>>.
- 48 SMITH, G. R.; STERNBERG, M. J. Prediction of protein–protein interactions by docking methods. **Current opinion in structural biology**, Elsevier, v. 12, n. 1, p. 28–35, 2002. Disponível em: <[https://doi.org/10.1016/S0959-440X\(02\)00285-3](https://doi.org/10.1016/S0959-440X(02)00285-3)>.
- 49 VERLI, H. Bioinformática: da biologia à flexibilidade molecular. Sociedade Brasileira de Bioquímica e Biologia Molecular, 2014.

- 50 GUEDES, I. A.; PEREIRA, F. S.; DARDENNE, L. E. Empirical scoring functions for structure-based virtual screening: applications, critical aspects, and challenges. **Frontiers in Pharmacology**, Frontiers, v. 9, p. 1089–1097, 2018. Disponível em: <<https://doi.org/10.3389/fphar.2018.01089>>.
- 51 GOHLKE, H.; HENDLICH, M.; KLEBE, G. Knowledge-based scoring function to predict protein-ligand interactions. **Journal of molecular biology**, Elsevier, v. 295, n. 2, p. 337–356, 2000. Disponível em: <<https://doi.org/10.1006/jmbi.1999.3371>>.
- 52 VANGONE, A. et al. Prediction of biomolecular complexes. In: **From Protein Structure to Function with Bioinformatics**. Springer, 2017. p. 265–292. Disponível em: <https://doi.org/10.1007/978-94-024-1069-3_8>.
- 53 KOSHLAND, D. E. Application of a theory of enzyme specificity to protein synthesis. **Proceedings of the National Academy of Sciences**, National Acad Sciences, v. 44, n. 2, p. 98–104, 1958. Disponível em: <<https://doi.org/10.1073/pnas.44.2.98>>.
- 54 WANG, C.; SCHUELER-FURMAN, O.; BAKER, D. Improved side-chain modeling for protein–protein docking. **Protein Science**, Wiley Online Library, v. 14, n. 5, p. 1328–1339, 2005. Disponível em: <<https://doi.org/10.1110/ps.041222905>>.
- 55 HALPERIN, I. et al. Principles of docking: An overview of search algorithms and a guide to scoring functions. **Proteins: Structure, Function, and Bioinformatics**, Wiley Online Library, v. 47, n. 4, p. 409–443, 2002. Disponível em: <<https://doi.org/10.1002/prot.10115>>.
- 56 SIMONELLI, L. et al. Rapid structural characterization of human antibody–antigen complexes through experimentally validated computational docking. **Journal of molecular biology**, Elsevier, v. 396, n. 5, p. 1491–1507, 2010. Disponível em: <<https://doi.org/10.1016/j.jmb.2009.12.053>>.
- 57 SMITH, G. R. et al. Incorporation of flexibility into rigid-body docking: Applications in rounds 3–5 of capri. **Proteins: Structure, Function, and Bioinformatics**, Wiley Online Library, v. 60, n. 2, p. 263–268, 2005. Disponível em: <<https://doi.org/10.1002/prot.20568>>.
- 58 ANDRUSIER, N. et al. Principles of flexible protein–protein docking. **Proteins: Structure, Function, and Bioinformatics**, Wiley Online Library, v. 73, n. 2, p. 271–289, 2008. Disponível em: <<https://doi.org/10.1002/prot.22170>>.
- 59 LYSKOV, S.; GRAY, J. J. The RosettaDock server for local protein–protein docking. **Nucleic acids research**, Oxford University Press, v. 36, n. suppl_2, p. W233–W238, 2008. Disponível em: <<https://doi.org/10.1093/nar/gkn216>>.
- 60 SIMONELLI, L. et al. Rational engineering of a human anti-dengue antibody through experimentally validated computational docking. **PLoS ONE**, Public Library of Science, v. 8, n. 2, p. e55561–e55561, 2013. Disponível em: <<https://doi.org/10.1371/journal.pone.0055561>>.
- 61 WANG, J. et al. A human bi-specific antibody against zika virus with high therapeutic potential. **Cell**, Elsevier, v. 171, n. 1, p. 229–241, 2017. Disponível em: <<https://doi.org/10.1016/j.cell.2017.09.002>>.

- 62 WEBSTER, B. C. **Chemical bonding theory**. [S.l.]: Blackwell scientific publications, 1990.
- 63 LEVINE, I. N.; BUSCH, D. H.; SHULL, H. **Quantum chemistry**. [S.l.]: Pearson Prentice Hall Upper Saddle River, NJ, 2009. v. 6.
- 64 DEMTRÖDER, W. **Atoms, molecules and photons**. Springer, 2010. v. 3. Disponível em: <<https://doi.org/10.1007/978-3-642-10298-1>>.
- 65 ROOTHAAN, C. C. J. New developments in molecular orbital theory. **Reviews of modern physics**, APS, v. 23, n. 2, p. 69–89, 1951.
- 66 CRAMER, C. J. **Essentials of computational chemistry: theories and models**. [S.l.]: John Wiley & Sons, 2013.
- 67 WARSHEL, A. Semiempirical methods of electronic structure calculation. **SGA, editor, Modern Theoretical Chemistry**, v. 7, p. 133, 1977. Disponível em: <https://doi.org/10.1007/978-1-4684-2556-7_5>.
- 68 STEWART, J. J. P. **Semiempirical Molecular Orbital Methods**. John Wiley & Sons, Ltd, 1990. v. 1. 45–81 p. Disponível em: <<https://doi.org/10.1002/9780470125786.ch2>>.
- 69 DEWAR, M. J.; THIEL, W. Ground states of molecules. 38. the mndo method. approximations and parameters. **Journal of the American Chemical Society**, ACS Publications, v. 99, n. 15, p. 4899–4907, 1977. Disponível em: <<https://doi.org/10.1021/ja00457a004>>.
- 70 DEWAR, M. J. 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**, ACS Publications, v. 107, n. 13, p. 3902–3909, 1985. Disponível em: <<https://doi.org/10.1021/ja00299a024>>.
- 71 STEWART, J. J. Optimization of parameters for semiempirical methods ii. applications. **Journal of computational chemistry**, Wiley Online Library, v. 10, n. 2, p. 221–264, 1989. Disponível em: <<https://doi.org/10.1002/jcc.540100209>>.
- 72 STEWART, J. Optimization of parameters for semiempirical methods. i: Method. **Journal of computational chemistry**, Wiley, v. 10, n. 2, p. 209–220, 1989. Disponível em: <<https://doi.org/10.1002/jcc.540100208>>.
- 73 STEWART, J. J. Optimization of parameters for semiempirical methods v: modification of nddo approximations and application to 70 elements. **Journal of Molecular modeling**, Springer, v. 13, n. 12, p. 1173–1213, 2007. Disponível em: <<https://doi.org/10.1007/s00894-007-0233-4>>.
- 74 STEWART, J. J. Optimization of parameters for semiempirical methods VI: more modifications to the NDDO approximations and re-optimization of parameters. **Journal of molecular modeling**, Springer, v. 19, n. 1, p. 1–32, 2013. Disponível em: <<https://doi.org/10.1007/s00894-012-1667-x>>.
- 75 STEWART, J. J. MOPAC2016; Stewart Computational Chemistry: Colorado Springs, CO, USA, 2016. **HTTP://OpenMOPAC.net (2016)**.

- 76 STEWART, J. J. Calculation of the geometry of a small protein using semiempirical methods. **Journal of Molecular Structure: THEOCHEM**, Elsevier, v. 401, n. 3, p. 195–205, 1997. Disponível em: <[https://doi.org/10.1016/S0166-1280\(97\)00050-X](https://doi.org/10.1016/S0166-1280(97)00050-X)>.
- 77 HOSTAŠ, J.; ŘEZÁČ, J.; HOBZA, P. On the performance of the semiempirical quantum mechanical pm6 and pm7 methods for noncovalent interactions. **Chemical Physics Letters**, Elsevier, v. 568, p. 161–166, 2013. Disponível em: <<https://doi.org/10.1016/j.cplett.2013.02.069>>.
- 78 NIKITINA, E. et al. Semiempirical calculations of binding enthalpy for protein–ligand complexes. **International journal of quantum chemistry**, Wiley Online Library, v. 97, n. 2, p. 747–763, 2004. Disponível em: <<https://doi.org/10.1002/qua.10778>>.
- 79 STEWART, J. J. Mopac: a semiempirical molecular orbital program. **Journal of computer-aided molecular design**, Springer, v. 4, n. 1, p. 1–103, 1990. Disponível em: <<https://doi.org/10.1007/BF00128336>>.
- 80 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. Disponível em: <[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)>.
- 81 STEWART, J. J. Application of localized molecular orbitals to the solution of semiempirical self-consistent field equations. **International Journal of Quantum Chemistry**, Wiley Online Library, v. 58, n. 2, p. 133–146, 1996.
- 82 THIEL, W. Semiempirical quantum–chemical methods. **Wiley Interdisciplinary Reviews: Computational Molecular Science**, Wiley Online Library, v. 4, n. 2, p. 145–157, 2014. Disponível em: <<https://doi.org/10.1002/wcms.1161>>.
- 83 DANIEL, C. R. A. et al. **Sequestro de CO₂ utilizando MOF-74-I: um estudo semiempírico**. Tese (Doutorado) — Universidade Federal de Sergipe, 2016.
- 84 BRENKE, R. et al. Application of asymmetric statistical potentials to antibody–protein docking. **Bioinformatics**, Oxford University Press, v. 28, n. 20, p. 2608–2614, 2012. Disponível em: <<https://doi.org/10.1093/bioinformatics/bts493>>.
- 85 SCHNEIDMAN-DUHOVNY, D. et al. Patchdock and Symmdock: servers for rigid and symmetric docking. **Nucleic acids research**, Oxford University Press, v. 33, n. suppl_2, p. W363–W367, 2005. Disponível em: <<https://doi.org/10.1093/nar/gki481>>.
- 86 MASHIACH, E. et al. Firedock: a web server for fast interaction refinement in molecular docking. **Nucleic acids research**, Oxford University Press, v. 36, n. suppl_2, p. W229–W232, 2008. Disponível em: <<https://doi.org/10.1093/nar/gkn186>>.
- 87 JANIN, J. Assessing predictions of protein–protein interaction: the CAPRI experiment. **Protein science**, Wiley Online Library, v. 14, n. 2, p. 278–283, 2005. Disponível em: <<https://doi.org/10.1110/ps.041081905>>.
- 88 ABRAHAM, M. J. et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. **SoftwareX**, Elsevier, v. 1, p. 19–25, 2015. Disponível em: <<https://doi.org/10.1016/j.softx.2015.06.001>>.

- 89 KLAMT, A.; SCHÜÜRMANN, G. Cosmo: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. **Journal of the Chemical Society, Perkin Transactions 2**, Royal Society of Chemistry, n. 5, p. 799–805, 1993. Disponível em: <<https://doi.org/10.1039/P29930000799>>.
- 90 DELANO, W. L. The pymol molecular graphics system. <http://www.pymol.org>, DeLano scientific, 2002.
- 91 JONES, M. L. et al. Computational identification of antibody epitopes on the dengue virus ns1 protein. **Molecules**, Multidisciplinary Digital Publishing Institute, v. 22, n. 4, p. 607–628, 2017. Disponível em: <<https://doi.org/10.3390/molecules22040607>>.
- 92 JIANG, L. et al. Selection and identification of b-cell epitope on ns1 protein of dengue virus type 2. **Virus research**, Elsevier, v. 150, n. 1-2, p. 49–55, 2010. Disponível em: <<https://doi.org/10.1016/j.virusres.2010.02.012>>.
- 93 FALCONAR, A.; YOUNG, P.; MILES, M. Precise location of sequential dengue virus subcomplex and complex b cell epitopes on the nonstructural-1 glycoprotein. **Archives of virology**, Springer, v. 137, n. 3-4, p. 315–326, 1994. Disponível em: <<https://doi.org/10.1007/BF01309478>>.
- 94 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**, 2019. Disponível em: <<https://doi.org/10.1093/nar/gkz358>>.
- 95 NGUYEN, M. N. et al. The interfacial character of antibody paratopes: analysis of antibody–antigen structures. **Bioinformatics**, Oxford University Press, v. 33, n. 19, p. 2971–2976, 2017. Disponível em: <<https://doi.org/10.1093/bioinformatics/btx389>>.
- 96 PEELING, R. W. et al. Evaluation of diagnostic tests: dengue. **Nature Reviews Microbiology**, Nature Publishing Group, v. 8, n. 12, p. S30–S37, 2010.
- 97 FRY, S. R. et al. The diagnostic sensitivity of dengue rapid test assays is significantly enhanced by using a combined antigen and antibody testing approach. **PLoS neglected tropical diseases**, Public Library of Science, v. 5, n. 6, 2011. Disponível em: <<https://doi.org/10.1371/journal.pntd.0001199>>.
- 98 LINARES, E. M. et al. Immunospot assay based on fluorescent nanoparticles for dengue fever detection. **Biosensors and Bioelectronics**, Elsevier, v. 41, p. 180–185, 2013. Disponível em: <<https://doi.org/10.1016/j.bios.2012.08.005>>.
- 99 HU, D. et al. Comparison of surface plasmon resonance, resonant waveguide grating biosensing and enzyme linked immunosorbent assay (elisa) in the evaluation of a dengue virus immunoassay. **Biosensors**, Multidisciplinary Digital Publishing Institute, v. 3, n. 3, p. 297–311, 2013. Disponível em: <<https://doi.org/10.3390/bios3030297>>.
- 100 SILVA, M. M. et al. Electrochemical detection of dengue virus ns1 protein with a poly (allylamine)/carbon nanotube layered immunoelectrode. **Journal of Chemical Technology & Biotechnology**, Wiley Online Library, v. 90, n. 1, p. 194–200, 2015. Disponível em: <<https://doi.org/10.1002/jctb.4305>>.

- 101 FIGUEIREDO, A. et al. Electrical detection of dengue biomarker using egg yolk immunoglobulin as the biological recognition element. **Scientific reports**, Nature Publishing Group, v. 5, p. 7865–7870, 2015. Disponível em: <<https://doi.org/10.1038/srep07865>>.

APÊNDICE A – Revisão dos métodos imunoenzimáticos para detecção da NS1.

Tabela 4 – Ensaios imunoenzimáticos de investigação do antígeno NS1.

Título do Trabalho	Principais Resultados	Referência
Evaluation of diagnostic tests: dengue.	Avaliação geral e comparativa sobre os métodos de diagnóstico da dengue. Os autores destacam que as principais características requeridas para um método diagnóstico da Dengue são: seletividade para diferenciar a dengue das demais patologias que causam quadros clínicos semelhantes, elevada sensibilidade durante a fase aguda da infecção, resultados rápidos, barato, fácil uso e estáveis a temperaturas acima de 30 °C para os casos de uso em campo.	(PEELING et al., 2010) [96]
Low sensitivity of NS1 protein tests evidenced during a dengue type 2 virus outbreak in Santos, Brazil, in 2010.	O trabalho realizou um estudo comparativo entre as performances dos kits de detecção da NS1 com os de detecção de outros marcadores como anticorpos IgM e de detecção viral em 379 casos de Dengue previamente evidenciados, pelo método de PCR em tempo real, na cidade de Santos-SP. Evidenciou-se que a detecção da proteína NS1 varia consideravelmente em função das condições epidemiológicas, portanto, a ocorrência de resultados falso-negativos precisa ser melhor compreendida e descrita.	(FELIX et al., 2012) [17]
The diagnostic sensitivity of dengue rapid test assays is significantly enhanced by using a combined antigen and antibody testing approach.	Estudo da performance do PanbioH®, um teste rápido de detecção precoce da Dengue pela investigação da NS1, e da possibilidade desse aumentar a sensibilidade do diagnóstico quando associado ao teste comercial de investigação de IgM/IgG. A teoria dos autores foi confirmada que ainda apontaram para o uso de diagnósticos rápidos como uma alternativa promissora em crescimento, principalmente nos países desenvolvidos.	(FRY et al., 2011) [97]
Improving dengue viral antigens detection in dengue patient serum specimens using a low pH glycine buffer treatment.	Os autores propuseram um tratamento prévio das amostras, envolvendo variação do pH, para investigação indireta de sNS1 no soro de pacientes infectados. Os níveis de detecção foram da ordem 3,840 ng ml-1 indicando sensibilidade do método, no entanto, os resultados de melhora na detecção foram mais significativos para a investigação da proteína E em detrimento da NS1.	(SHEN et al., 2017) [19]

Tabela 5 – Ensaios imunoenzimáticos de quantificação do antígeno NS1 (Biossensores).

Título do Trabalho	Principais Resultados	Referência
Immunospot assay based on fluorescent nanoparticles for Dengue fever detection.	Os autores propuseram como teste de detecção da Dengue o uso da técnica de Immunospot® que combina as vantagens do método de ELISA, já amplamente utilizado para esse fim, com nanopartículas fluorescentes. O limite de detecção do teste proposto é comparável ao método de ELISA convencional e com maior sensibilidade apresentando-se como uma alternativa de método rápido, específico e sensível para diagnóstico da dengue.	(LINARES et al., 2013) [98]
Comparison of surface plasmon resonance, resonant waveguide grating biosensing and enzyme-linked immunosorbent assay (ELISA) in the evaluation of a dengue virus immunoassay.	Dois métodos óticos de detecção, SPR e RWG, foram aplicados na detecção da proteína NS1 vislumbrando um diagnóstico mais sensível e específico alternativamente aos tradicionais. Os autores avaliaram as cinéticas de interação antígeno-anticorpo e concluíram que as técnicas podem ser promissoras quando utilizadas em associação ao método de ELISA.	(HU et al., 2013) [99]
Electrochemical detection of dengue virus NS1 protein with poly(allylamine)/ carbon nanotube layered immunoelectron.	Desenvolvimento de um imunossensor eletroquímico baseado num compósito de polímero depositados numa superfície de Nanotubos de carbono (NTCs) para quantificação da proteína NS1 do vírus da dengue. O sensor apresentou faixa de linearidade entre 0,1 e 2,5 µg mL ⁻¹ para detecções na fase aguda da virose o que diminuiu significativamente para detecções na fase secundária. Essa plataforma é, portanto, aplicável no diagnóstico precoce da dengue.	(SILVA et al., 2015) [100]
Electrical detection of dengue biomarker using egg yolk immunoglobulin as the biological recognition element.	Os autores propuseram um sensor elétrico para detecção da NS1 DENV sorotipo 2. O dispositivo analítico baseia-se na leitura de variações de Potencial de Circuito Aberto de um eletrodo de ouro, contendo anticorpos anti-NS1 imobilizados, e um eletrodo de referência. A faixa de linearidade para amostras padrão de NS1 foi de 0,1 e 10 µg mL ⁻¹ e 0,09 µg mL ⁻¹ de limite de detecção, indicando a possibilidade de aplicação do sistema em amostras reais de soro de pacientes DENV2 positivos haja visto que as concentrações apresentados são da ordem das observadas em casos clínicos reais.	(FIGUEIREDO, 2015) [101]