

DEPARTAMENTO DE QUÍMICA UNIVERSIDADE FEDERAL DE MINAS GERAIS

COLEGIADO DOS CURSOS DE QUÍMICA E QUÍMICA TECNOLÓGICA

TRABALHO DE CONCLUSÃO DE CURSO DE BACHARELADO EM QUÍMICA

SAMUEL DE MACEDO LOURENÇO

**MODELAGEM DA INTERAÇÃO DO GLICOSAMINOGLICANO FONDAPARINUX
COM A PROTEÍNA E DA PARTÍCULA DO VÍRUS DA DENGUE**

BELO HORIZONTE

2026

SAMUEL DE MACEDO LOURENÇO

**MODELAGEM DA INTERAÇÃO DO GLICOSAMINOGLICANO FONDAPARINUX
COM A PROTEÍNA E DA PARTÍCULA DO VÍRUS DA DENGUE**

Trabalho de Conclusão de Curso (TCC)
apresentado ao Departamento de Química da
Universidade Federal de Minas Gerais, como
parte dos requisitos necessários à obtenção do
título de Bacharel em Química.

Orientador(a): Prof. Dr. Adolfo Henrique de
Moraes Silva

Co-orientador(a): Me. Philipe de Oliveira
Fernandes

BELO HORIZONTE

2026

UFMG

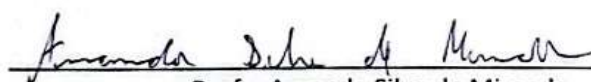
COLEGIADO DE QUÍMICA
Instituto de Ciências Exatas
Fone: (031) 3409-5799/6386
E-mail: colgradqui@icex.ufmg.br

Ata da sessão de avaliação do Trabalho de Conclusão de Curso de Bacharelado em Química da Universidade Federal de Minas Gerais, do aluno Samuel de Macedo Lourenço, nº de registro 2021038380.

O Trabalho de Conclusão de Curso do aluno Samuel de Macedo Lourenço, intitulado "Modelagem da interação do glicosaminoglicano fondaparinux com a proteína e da partícula do vírus da dengue", foi avaliado e aprovado pela Banca Examinadora constituída por Profa. Amanda Silva de Miranda (DQ - UFMG), Dr. Mozart Silvio Pereira (DQ - UFMG), Me. Philipe de Oliveira Fernandes (DQ - UFMG), coorientador do trabalho de conclusão de curso, e Prof. Adolfo Henrique de Moraes Silva (DQ - UFMG), orientador do trabalho de conclusão de curso, no dia 24 de junho de 2026.


Prof. Adolfo Henrique de Moraes Silva


Me. Philipe de Oliveira Fernandes


Profa. Amanda Silva de Miranda


Dr. Mozart Silvio Pereira


Prof. Dario Windmüller / Profa. Dalva Ester da Costa Ferreira
Coordenador dos Cursos de Graduação em Química e Química Tecnológica
ICEx – UFMG

AGRADECIMENTOS

Aos meus pais que eu amo incondicionalmente e apesar de todas as dificuldades, sempre me incentivaram a ser uma pessoa melhor, desde cedo ensinando-me o valor do estudo, esforço e dedicação, fundamentais para a construção da minha identidade e dos meus valores enquanto homem e filho.

À minha família, que sempre se fez presente e unida me oferecendo apoio constante, especialmente nos momentos de maior adversidade, nos alegres momentos juntos, nas reuniões em família, na coleção de bons momentos proporcionados em todos esses anos, fica a todos vocês o meu eterno amor e gratidão.

A UFMG por me permitir estudar e obter acesso a conhecimentos, experiências e vivências que eu não conseguiria de nenhuma outra forma. Pelas experiências cotidianas de contato com pessoas incríveis que marcaram a minha vida e eu serei eternamente grato.

Aos meus orientadores que desde o início do projeto se demonstraram muito prestativos e presentes me auxiliando no desenvolvimento de um projeto que finaliza a minha jornada acadêmica grandemente, agregando muitos dos conhecimentos que adquiri nos últimos anos com muitos aprendizados novos, terminando essa etapa da minha vida com muito mérito e orgulho desse trabalho desenvolvido.

A todos os meus amigos que moraram comigo e aos que não moraram, mas compartilharam a vivência universitária ao meu lado, seja em Belo Horizonte ou em Sete Lagoas, dividindo momentos de lazer e aprendizado.

“Saber de onde você veio não é menos importante do que saber para onde você está indo.”-

Neil deGrasse Tyson

Resumo

A dengue é uma arbovirose com alta incidência em regiões tropicais e subtropicais, para a qual não há terapias antivirais específicas aprovadas. A proteína E do vírus da dengue sorotipo 2 (DENV-2) é fundamental no processo de entrada viral, mediando o reconhecimento de GAGs nas células hospedeiras por meio de resíduos de aminoácidos localizados nas estruturas oligoméricas do vírion maduro. Estudos mostraram que GAGs análogos ao sulfato de heparano podem inibir a infecção viral. Apesar dos estudos bioquímicos, ainda não há, na literatura, um modelo estrutural da interação entre a proteína E da partícula viral e GAGs. Recentemente, nosso grupo identificou e caracterizou *druggable pockets* na estrutura trimérica da proteína E, descrevendo o hotspot DP-C16-T como um sítio de alta drogabilidade. Neste trabalho, 16 conformações representativas foram utilizadas como receptores em um protocolo de *ensemble docking* com o fondaparinux. Essa abordagem teve como objetivo investigar, por métodos computacionais, a capacidade do fondaparinux de interagir com o sítio de ligação a GAG da proteína E. Foram geradas 800 poses de docking, submetidas a uma análise do perfil de interações e a uma clusterização hierárquica por centro de massa. Foram detectadas 18.518 interações distribuídas predominantemente pela cadeia A do trímero. A clusterização identificou cinco clusters, dos quais o Cluster 3 coincidiu espacialmente com o DP-C16-T e concentrou 52,6% das poses, demonstrando convergência do fondaparinux para o sítio biologicamente relevante. A sub-clusterização por RMSD desse cluster revelou três modos de ligação geometricamente distintos, cada um representado por uma pose selecionada por critérios físico-químicos. As três poses representativas compartilham um núcleo de interação composto por Glu172, Thr180, Lys291 e Lys295 e se diferenciam por sub-regiões farmacofóricas específicas dentro do *pocket*. Os resultados demonstram que o fondaparinux reconhece o *pocket* através dos mesmos resíduos responsáveis pela infecção viral. As três poses representativas identificadas constituem estruturas de partida para simulações de DM e estudos futuros que permitirão o design de candidatos antivirais para o DENV-2.

Palavras-chave: Dengue, fondaparinux, modelagem molecular, docking, Proteína E

Abstract

Dengue is an arboviral disease with high incidence in tropical and subtropical regions, for which there are no approved specific antiviral therapies. The envelope (E) protein of dengue virus serotype 2 (DENV-2) is crucial for viral entry, mediating the recognition of GAGs on host cells via amino acid residues located on the oligomeric structures of the mature virion. Studies have shown that GAGs analogous to heparan sulfate can inhibit viral infection. Despite biochemical studies, a structural model of the interaction between the viral particle's E protein and GAGs is still lacking in the literature. Recently, our group identified and characterized druggable pockets within the trimeric structure of the E protein, describing the DP-C16-T hotspot as a site with high druggability. In this study, 16 representative conformations were used as receptors in an ensemble docking protocol with fondaparinux. This approach aimed to investigate, using computational methods, the ability of fondaparinux to interact with the E protein's GAG-binding site. A total of 800 docking poses were generated and subjected to interaction profile analysis and hierarchical clustering based on the center of mass. We detected 18,518 interactions, distributed predominantly across chain A of the trimer. Clustering identified five clusters; Cluster 3 spatially coincided with DP-C16-T and contained 52.6% of the poses, demonstrating the convergence of fondaparinux toward the biologically relevant site. RMSD-based sub-clustering of this cluster revealed three geometrically distinct binding modes, each represented by a pose selected according to physicochemical criteria. The three representative poses share an interaction core composed of Glu172, Thr180, Lys291, and Lys295, while differing in specific pharmacophoric sub-regions within the pocket. The results demonstrate that fondaparinux recognizes the pocket via the same residues responsible for viral infection. The three representative poses identified constitute starting structures for MD simulations and future studies that will enable the design of antiviral candidates for DENV-2.

Keywords: Dengue, fondaparinux, molecular modeling, docking, E protein

LISTA DE FIGURAS

Figura 1: Países com o maior número de casos notificados de dengue em 2024.....	11
Figura 2: Percentual de óbitos por dengue por país em 2024.....	12
Figura 3: Aumento do número de casos prováveis no Brasil.....	13
Figura 4: Ilustração da estrutura do DENV.....	14
Figura 5: Representação simplificada da inoculação de vírus e saliva de mosquito na pele.....	16
Figura 6: Componentes da superfície celular envolvidos no reconhecimento/ligação/anexação e internalização do DENV.....	17
Figura 7: Ciclo de replicação do vírus da dengue em hospedeiro humano, através de endocitose mediada por clatrina.....	18
Figura 8: Estrutura da glicoproteína E do DENV.....	20
Figura 9: Sulfato de Heparano.....	22
Figura 10: Fondaparinux.....	22
Figura 11: Esquema ilustrativo de um docking molecular pelo Autodock Vina.....	25
Figura 12: Output do Vina.....	30
Figura 13: Preparo das estruturas.....	34
Figura 14: Experimento realizado com a grid box em diferentes posições:.....	35
Figura 15: Resultados do docking com a caixa dentro e docking com a caixa fora e resultados do docking com a caixa dentro.....	36
Figura 16: Quantidade de interações por conformação.....	37
Figura 17: Distribuição das interações por cadeia.....	38
Figura 18: Boxplot de score x conformação.....	38
Figura 19: Visualização e dados gerais da clusterização.....	40
Figura 20: Visualização e dados gerais da sub-clusterização.....	42
Figura 21: Distribuição das poses por conformação dentro do cluster 3.....	42
Figura 22: Poses representativas de cada sub-cluster.....	44
Figura 23: Visualização das interações das poses representativas de cada sub-cluster de diferentes formas.....	45
Figura A-1: 20 resíduos que mais interagiram com o fondaparinux.....	55

Figura A-2: Top 15 resíduos que mais interagiram com o fondaparinux por cadeia do trímero.....	55
Figura B-1: Método do cotovelo “Elbow Method” da clusterização por centro de massa.....	56
Figura B-2: Método do cotovelo “Elbow Method” da clusterização por RMSD.....	57

LISTA DE TABELAS

Tabela 1: Categorias das funções de score.....	23
Tabela 2: Tipos de interação detectadas pelo PLIP.....	30
Tabela 3: Classificação dos clusters.....	37

LISTA DE ABREVIATURAS, SIGLAS, SÍMBOLOS E UNIDADES

CADD- *Computational-aided drug design*

Cryo-EM - Criomicroscopia eletrônica

DI - Domínio I

DII - Domínio II

DIII - Domínio III

DM - Dinâmica molecular

FS - Funções de *score*

GAGs - Glicosaminoglicanos

HCA - Análise Hierárquica de Clusters

IDE- Ambiente de desenvolvimento integrado

IUPAC- União Internacional de Química Pura e Aplicada

LBDD- *Ligand Based Drug Design*

MOPAC - Molecular Orbital PACkage

MRC- *Multiple Receptor Conformations*

PDB - *Protein Data Bank*

PLIP - *Protein-Ligand interaction profiler*

SBDD - *Structure-Based Drug Design*

SH - Sulfato de Heparano

SUMÁRIO

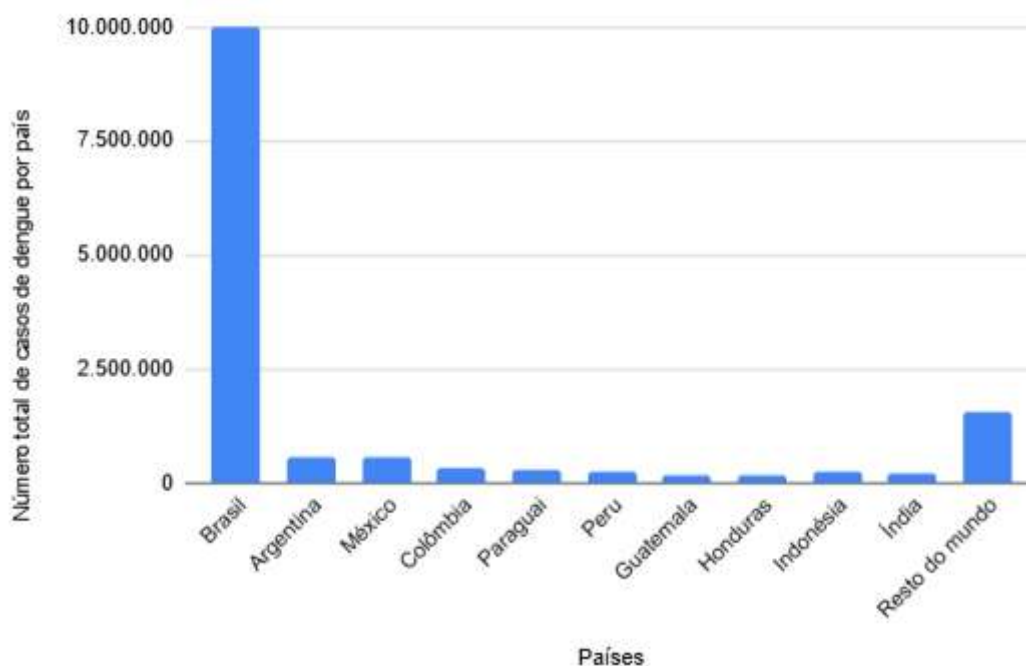
1. INTRODUÇÃO.....	11
2. REVISÃO DE LITERATURA.....	14
3. OBJETIVO.....	27
3.1 Objetivo geral.....	27
3.2 Objetivos específicos.....	27
4. MATERIAIS E MÉTODOS.....	28
4.1 Obtenção e preparação das estruturas.....	28
4.2 Docking molecular.....	28
4.3 PLIP.....	31
4.4 HCAs.....	32
5. RESULTADOS E DISCUSSÃO.....	34
5.1.Ensemble docking.....	34
5.2 Análise do ensemble docking.....	36
6. CONCLUSÃO.....	47
7. REFERÊNCIAS.....	49
8. APÊNDICE.....	55
Apêndice A: Resultados da Análise do PLIP.....	55
Apêndice B: Resultados da análise do HCA.....	56

1. INTRODUÇÃO

A dengue é uma arbovirose endêmica em regiões tropicais e subtropicais, com expansão global crescente nas últimas décadas. Em 2024 a OMS classificou a doença como endêmica em mais de 130 países, ameaçando a vida de mais de 4 bilhões de pessoas em todo o mundo (Organização das Nações Unidas, 2024). Embora mais prevalentes na América do Sul e no Sudeste Asiático, surtos também têm sido registrados em áreas temperadas, como a Europa e os Estados Unidos, incluindo episódios de transmissão local e óbitos (Sharp et al., 2021).

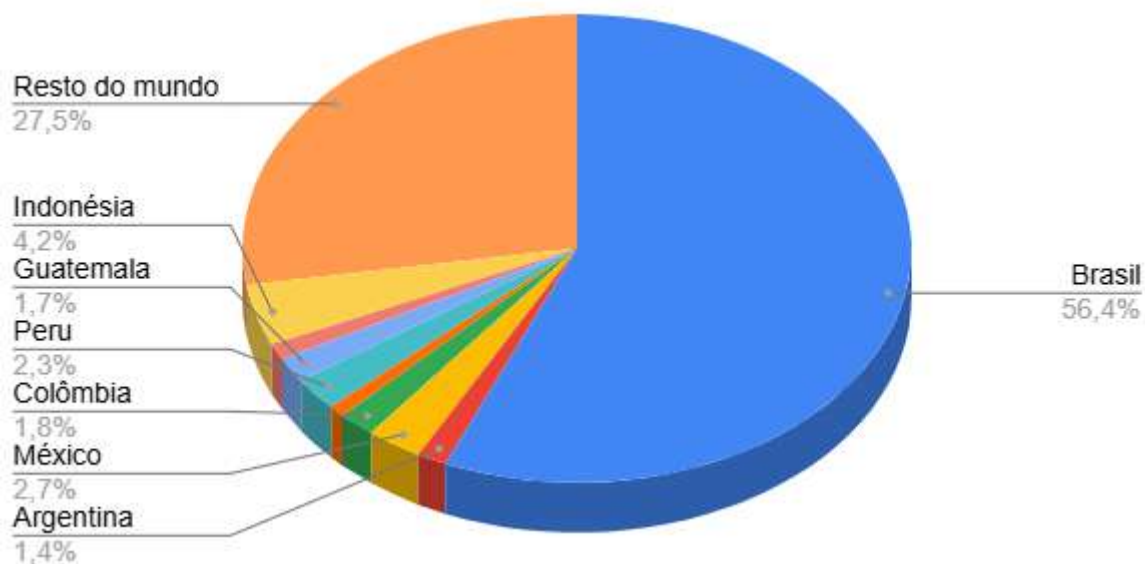
Os dados referentes a 2025 ainda não foram completamente consolidados em nível global pela Organização Mundial da Saúde (OMS). Em 2024, o Brasil enfrentou a maior epidemia de dengue já documentada, com mais de 6,5 milhões de casos prováveis e 6.321 óbitos confirmados (Brasil, 2026). Em 2024, foram notificados aproximadamente 14,5 milhões de casos, e apenas o Brasil notificou mais de 10 milhões (Figura 1). Além disso, foram notificadas mais de 11 mil mortes no mundo inteiro, com o Brasil representando 56% dos casos letais (Figura 2) (World Health Organization, 2024).

Figura 1: Quantidade de casos de dengue notificados entre os países com o maior número de casos de dengue no ano de 2024.



Fonte: OMS

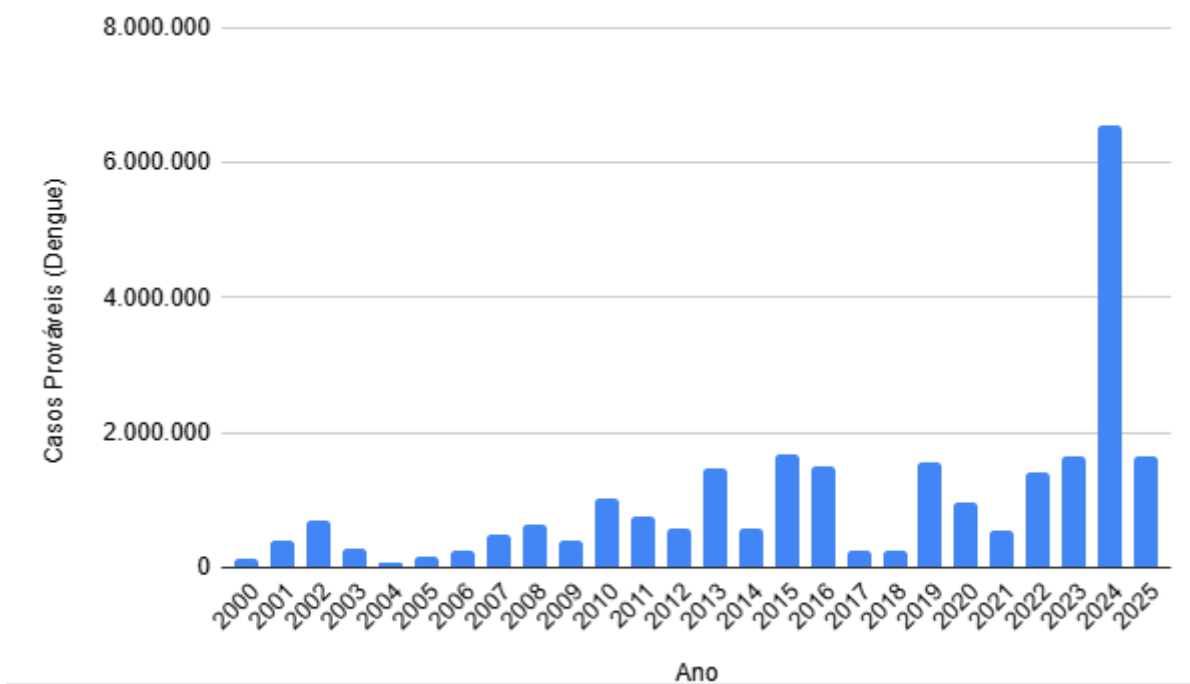
Figura 2: Países com o maior percentual de óbitos por dengue no mundo no ano de 2024.



Fonte: Painel de Monitoramento de Arboviroses

De acordo com o painel de monitoramento de arboviroses do Ministério da Saúde, no Brasil em 2025 observou-se uma redução no número de casos após o surto registrado em 2024. Os registros indicam mais de 1,6 milhão de casos no país, com cerca de 1.793 mortes, que representa uma diminuição significativa em comparação com o ano anterior. Apesar dessa redução, estudos indicam que surtos epidemiológicos de dengue tendem a ocorrer em ciclos de aproximadamente dois a três anos no Brasil (Figura 3) e em países do hemisfério sul (Jones, 2004; Vincenti-Gonzalez et al., 2018), o que sugere a possibilidade de novos aumentos no número de casos nos próximos anos.

Figura 3: Aumento do número de casos prováveis de dengue no Brasil entre os anos de 2000 e 2025.



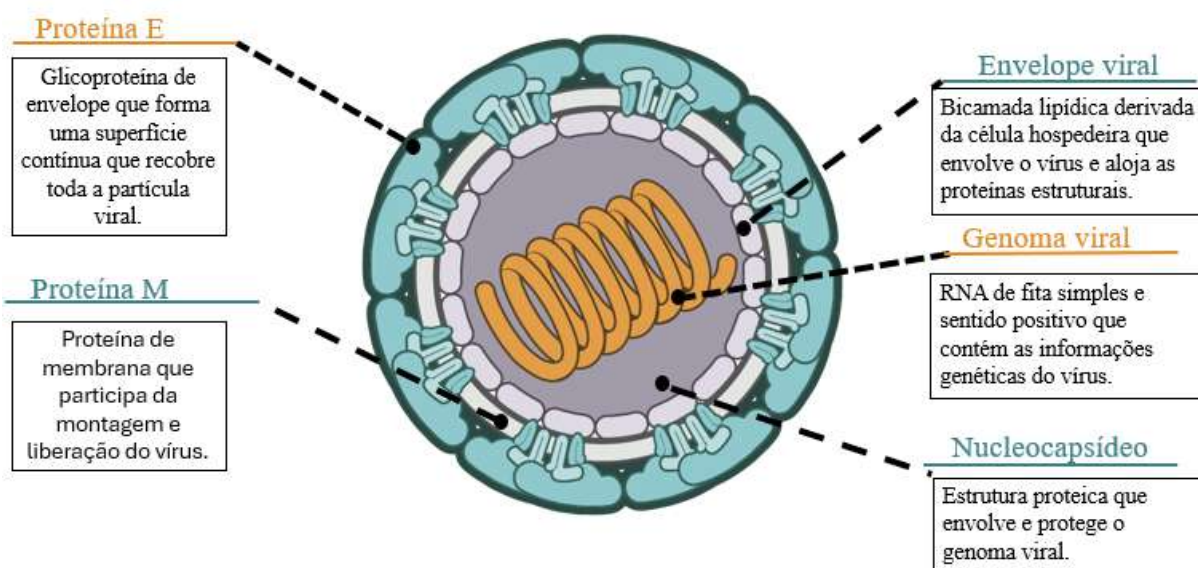
Adaptado de gov.br

A expansão geográfica da dengue é associada a fatores como urbanização acelerada e mal planejada, más práticas de saneamento e higiene, mobilidade humana (migrações e viagens internacionais) além dos fenômenos de mudanças climáticas, que favorecem a dispersão do vetor e a manutenção da transmissão, sendo essa uma tendência observada em outras doenças arbovirais (Organização das Nações Unidas, 2024; Fong, 2020). Embora a origem geográfica da dengue ainda seja motivo de debate (Gubler ,2012; Brady; Hay, 2020), o seu impacto na saúde pública é inegável e vem aumentando drasticamente nos últimos 25 anos.

2. REVISÃO DE LITERATURA

O vírus da dengue (DENV) é um vírus de RNA de fita simples e sentido positivo, com formato esférico e um nucleocapsídeo icosaédrico coberto por uma bicamada lipídica (Nanaware et al., 2021) (Figura 4). Os vírus da dengue (DENV) causam um amplo espectro de manifestações clínicas, incluindo condições potencialmente fatais, como a síndrome do choque hemorrágico e, menos frequentemente, hepatite aguda com insuficiência hepática e encefalopatia (Hilgard; Stockert, 2000). Além disso, o DENV é o membro mais prevalente do gênero *Flavivirus* ao redor do mundo, que também inclui vírus como a febre amarela (YFV), o vírus do Nilo Ocidental e o Zika, que também emergiram como ameaças significativas à saúde pública (Brady; Hay, 2020).

Figura 4: Ilustração da estrutura do DENV



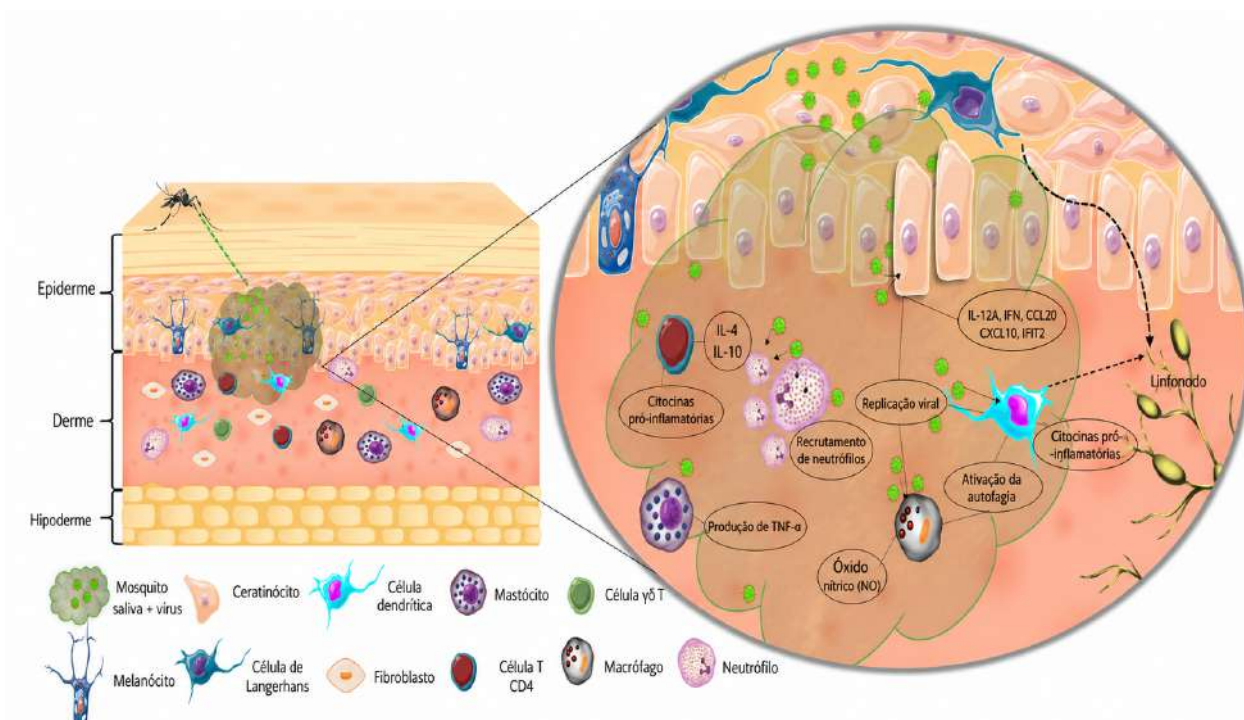
Fonte: Elaborado pelo autor

O DENV existe em pelo menos quatro clados geneticamente distintos, frequentemente denominados sorotipos (DENV 1-4) (Brady; Hay, 2020). Os quatro sorotipos do DENV diferem no nível dos aminoácidos entre 25–40% e podem ser ainda classificados em genótipos separados que variam até 3% (Tamura et al., 2022). Cada sorotipo de DENV é classificado em diferentes genótipos com base em variações de nucleotídeos.

O DENV-1 é subdividido em cinco genótipos, sendo o genótipo I asiático, enquanto os genótipos II e III pertencem à Tailândia e à Malásia, respectivamente. O genótipo IV é do Pacífico Sul e o genótipo V é o americano/africano. O DENV-2 foi classificado em seis genótipos: os asiáticos (I e II), o americano/asiático (III), o cosmopolita (IV) e o americano (V). O DENV-3 é ainda dividido em cinco genótipos (I-V) na linhagem humana. A maioria das infecções por DENV-3 é causada pelos genótipos I, II e III, que estão associados a epidemias nas regiões do Sudeste Asiático, Pacífico Sul, subcontinente indiano, África Oriental e Américas. Os genótipos I (Tailândia, Sri Lanka, Filipinas e Japão), II (Indonésia, Malásia, Taiti, Caribe e Américas), III (isolados tailandeses) e IV (cepas silvestres da Malásia) do DENV-4 foram analisados usando sequências completas do gene E (Agarwal et al., 2024).

A infecção é transmitida por fêmeas do mosquito *Aedes aegypti* durante o repasto sanguíneo, momento em que a saliva do inseto, rica em proteínas com capacidade de modular a homeostase e a resposta imunológica do hospedeiro, é inoculada nos tecidos subepidérmicos (Figura 5). Esses componentes salivares promovem efeitos como aumento da permeabilidade vascular, modulação da sinalização de citocinas da imunidade inata e recrutamento de células inflamatórias, para o sítio de inoculação, fazendo com que a saliva do mosquito não atue apenas como veículo, mas como um facilitador da infecção viral, modulando a resposta imune local e favorecendo a replicação do vírus (Guerrero; Cantaert; Missé, 2020). Nesse microambiente, a partícula viral interage com diferentes receptores presentes em diversos tipos celulares, como fibroblastos, monócitos, macrófagos, células epiteliais e células dendríticas, favorecendo o estabelecimento inicial da infecção (Urcuqui-Inchima et al., 2010). A partir dessas interações iniciais no sítio de inoculação, torna-se fundamental compreender os mecanismos moleculares que mediam o reconhecimento, a ligação e a internalização do vírus nas células hospedeiras.

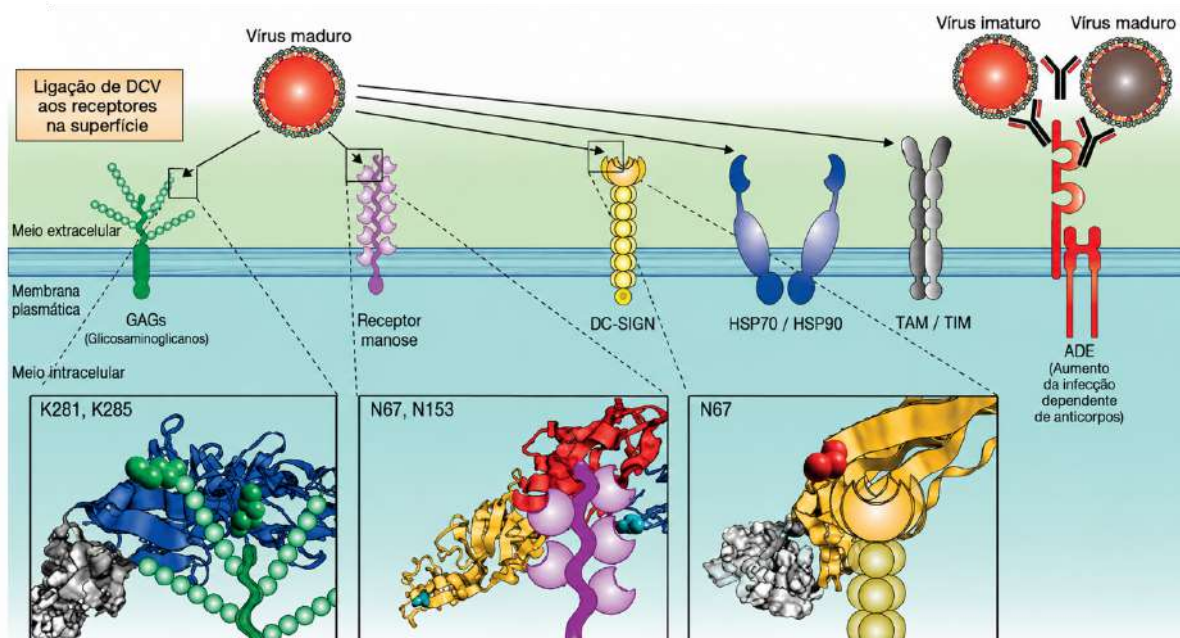
Figura 5: Representação simplificada da inoculação de vírus e saliva de mosquito na pele. Reconhecimento do vírus por células de Langerhans (LCs) e células dendríticas (DCs), e migração para o linfonodo. Efeito da saliva do mosquito nas células imunes residentes e infiltrantes da pele.



Adaptada de Guerrero D. et.al 2020.

Estudos sugerem que o DENV não utiliza um receptor único e específico para sua entrada na célula, mas reconhece e se liga a diversas moléculas (Cruz-Oliveira et al., 2015). Os principais receptores do hospedeiro associados à entrada do DENV incluem: o GAG sulfato de heparano (SH), receptor de manose (RM) o complexo lipopolissacarídeo (LPS)-CD14, a molécula de adesão de células dendríticas (DC-SIGN), proteínas induzidas por estresse, como HSP70 e HSP90 e a chaperonina do RE GRP78 (Figura 6) (Cruz-Oliveira et al., 2015). Um receptor que merece destaque é o sulfato de heparano, um copolímero linear e um glicosaminoglicano (GAG), carregado negativamente, de resíduos de carboidratos de ácido urônico e glucosamina com sulfatação variável amplamente expresso na superfície celular, que atua como importante fator de ligação inicial em diferentes linhagens epiteliais (Hilgard; Stockert, 2000) .

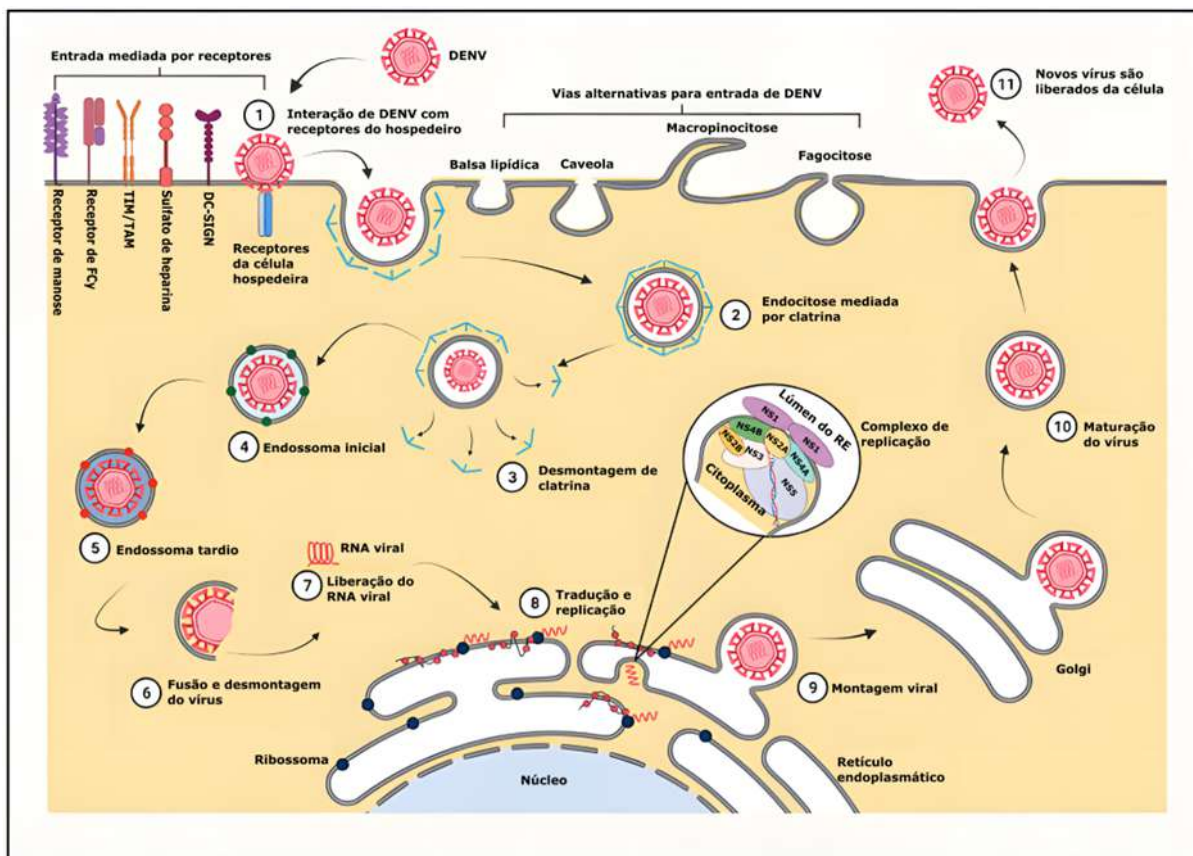
Figura 6: Componentes da superfície celular envolvidos no reconhecimento/ligação/anexação e internalização do DENV: GAGs, receptor de manose, DC-SIGN, HSP90/HSP70 e TIM/TAM, bem como os receptores Fc γ dos macrófagos, que participam da internalização dos imunocomplexos do DENV maduro ou imaturo, resultando no fenômeno do ADE.



Adaptado de Cruz-Oliveira et.al 2015

Após a interação do vírus com os receptores celulares do hospedeiro, diferentes vias podem mediar a entrada da partícula viral na célula, sendo a endocitose mediada por clatrina (Figura 7) a mais descrita na literatura. Nesse contexto, a porção da membrana celular recoberta por clatrina promove a endocitose em uma vesícula que, em seguida, é liberada, iniciando o processo endossomal. O endossoma, por sua vez, passa por mudanças conformacionais marcadas pela substituição de proteínas Rab5 por Rab7, juntamente com a redução do pH de 6,5 para 5. Como resultado, essas alterações favorecem a fusão entre o envelope viral e a membrana do hospedeiro, levando à liberação do genoma viral no citoplasma. A partir disso, o RNA viral serve de molde para os processos de tradução e replicação no retículo endoplasmático. Por fim, os genomas replicados e as proteínas sintetizadas são organizados em novas partículas virais, que passam por maturação na rede trans do Golgi e, posteriormente, são exocitadas da célula infectada (Nanaware et al., 2021).

Figura 7: Ciclo de replicação do vírus da dengue em hospedeiro humano, mediado por endocitose mediada por clatrina. A invaginação coberta por clatrina é o sítio de internalização (2), e, após a endocitose, a clatrina é liberada do endossoma (3), que passa por diferentes estágios de maturação (4) e (5), caracterizados por presença das proteínas Rab5 e Rab7 e valores de pH próximos de 6,5 e 5,5, respectivamente. Mudanças conformacionais resultam na fusão do envelope viral com a membrana do hospedeiro (6). O desnudamento da partícula viral libera seu genoma no citoplasma (7), dando início ao processo de tradução e replicação do RNA viral (8). As proteínas virais são produzidas e compõem o vírus imaturo (9), que passa por maturação mediada pela furina no complexo de Golgi (10), antes de ser liberado da célula infectada (11).



Adaptado de Nanaware et al., 2021.

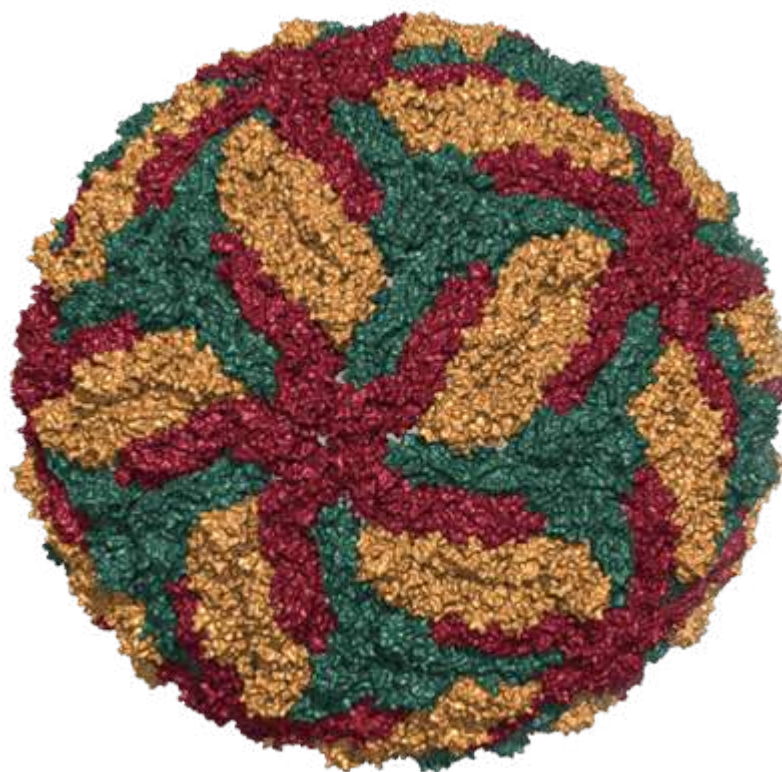
Como mencionado na introdução, há uma prevalência crescente e uma disseminação mundial e o tratamento permanece sintomático devido à ausência de terapias antivirais específicas para a doença. Embora o desenvolvimento de vacinas contra a dengue tenha apresentado melhorias, a vacina atualmente licenciada CYD-TDV (Dengvaxia) demonstrou eficácia parcial em crianças e adultos soronegativos, ressaltando a importância do status sorológico prévio como fator determinante para a segurança e eficácia da vacinação, o que impõe desafios adicionais para sua implementação em larga escala (Huang et al., 2021). Outros candidatos, como a vacina TAK-003, demonstraram resultados promissores em

ensaios de fase III para diferentes sorotipos, com uso recomendado para crianças entre 6 e 16 anos de idade (Leite, 2025).

Portanto, cresce o interesse no desenvolvimento de fármacos que atuem sobre alvos virais bem definidos, preferencialmente sem homólogos funcionais no hospedeiro. Alvos moleculares eficazes incluem proteínas envolvidas em etapas essenciais do ciclo viral, como entrada na célula, fusão membranar, replicação e a montagem (Kausar et al., 2021). Entre essas proteínas, destaca-se a glicoproteína de envelope E, responsável pela ligação ao receptor celular e pela fusão entre o envelope viral e a membrana endossomal (Villalaín, 2022).

A glicoproteína de envelope (proteína E), é a principal proteína estrutural presente na superfície dos vírions maduros do vírus da dengue, sendo classificada como uma proteína integral de membrana do tipo I e desempenhando papel central em etapas críticas do ciclo viral, como reconhecimento celular, entrada e fusão de membranas (Diamond; Pierson, 2015). Foi demonstrado que a proteína E do DENV forma homodímeros de orientação antiparalela (cabeça-cauda) (Rey et al., 1995). No virion maduro, a proteína E adota um arranjo mais complexo, formando uma superfície icosaédrica composta por 180 monômeros de proteína E organizados em conjuntos triméricos e pentaméricos, que seguem o arranjo do dímero (Figura 8) (Leite et al., 2026). Cada monômero possui três domínios distintos: domínio I (DI), o domínio N-terminal central; domínio II (DII), o domínio de dimerização; e domínio III (DIII), o domínio semelhante à imunoglobulina (Ig) (Rey et al., 1995). O DIII, composto por aproximadamente 100 aminoácidos (resíduos 296–395) da região C-terminal, tem sido amplamente proposto como o principal domínio de reconhecimento e de ligação aos GAGs que atuam como receptores celulares (Cruz-Oliveira et al., 2015). A dobra semelhante à Ig presente na proteína DIII é comumente associada a estruturas com função de adesão (Rey et al., 1995). Além disso, evidências experimentais indicam que o DIII contém determinantes estruturais diretamente envolvidos na interação com receptores celulares e na especificidade de ligação do vírus (Huerta et al., 2008). Mutações nesse domínio podem resultar em alterações na virulência viral ou favorecer o escape da neutralização imunológica, destacando sua relevância tanto na patogênese quanto na resposta imune (Chin; Chu; Ng, 2007).

Figura 8: Estrutura da partícula viral e do recobrimento da glicoproteína E do DENV, obtida por cryo-EM PDB3J27 . Os Pentâmeros da proteína E estão representados em vermelho, trímeros em verde e dímeros em amarelo



Fonte: Leite, 2025

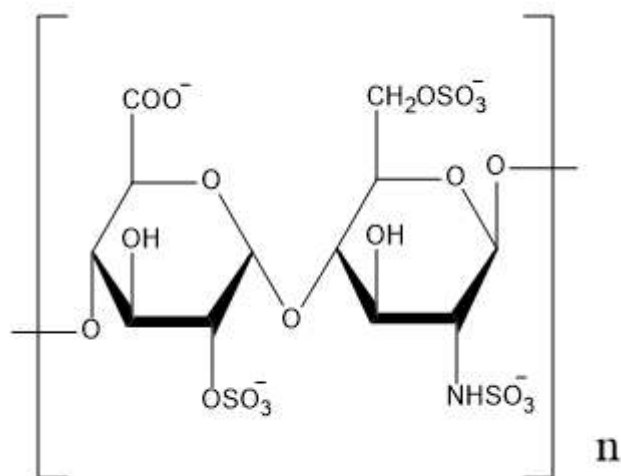
Já foi demonstrado em estudos prévios que esses arranjos triméricos e pentaméricos expõem *pockets* drogáveis, regiões específicas na estrutura de uma proteína onde uma pequena molécula pode se ligar e alterar a função da proteína, que são ausentes nas formas monoméricas e diméricas isoladas da proteína E (Leite et al., 2026). Alguns resíduos de aminoácidos presentes nesses arranjos estruturais são validados experimentalmente como essenciais para a ligação a glicosaminoglicanos na superfície de célula hospedeira (Watterson et al. 2012). No trímero, um hotspot drogável designado DP-C16-T localizado na interface entre os três monômeros é caracterizado pela presença de resíduos com papel funcional no reconhecimento de glicosaminoglicanos, incluindo Lys291, Lys295, Glu172, Glu174, Tyr178 e Thr180 (Leite et al., 2026). Investigações acerca do papel dos resíduos da glicoproteína E do DENV na ligação a glicosaminoglicanos da superfície celular, demonstraram que a Lys 291 e a Lys 295 exercem papel central nesse processo (Watterson et al. 2012). As cadeias laterais

carregadas positivamente dessas lisinas estabelecem interações eletrostáticas com os grupos negativamente carregados dos glicosaminoglicanos, presentes na superfície das células hospedeiras, mediando o contato inicial do vírion com a membrana celular, etapa que precede a endocitose mediada por clatrina e a fusão endossomal (Watterson et al. 2012). Além disso, uma análise de conservação de sequência entre flavivírus revelou que os resíduos do *pocket* trimétrico são bem conservados em múltiplos sorotipos do DENV e de vírus relacionados, sugerindo que inibidores direcionados a esse sítio poderiam apresentar atividade de amplo espectro (Leite et al., 2026).

O conhecimento acerca da inoculação e do ciclo de replicação do DENV (Figuras 5, 6 e 7), juntamente com informações sobre a proteína E e os seus receptores celulares podem ser muito importantes para o desenvolvimento de antivirais e têm imensa importância para a saúde pública. Estudos revelaram que a ligação do DENV à linhagem celular de hepatoma humano HuH-7 foi estritamente dependente do pH e substancialmente inibida pelo glicosaminoglicano heparina, com a análise por ligand-blot mostrando duas proteínas de ligação à superfície celular contendo sulfato de heparano (Hilgard; Stockert, 2000). Isso mostra que o papel dos GAGs, com destaque para o sulfato de heparano (Figura 9) utilizado pelo vírus como receptor durante a etapa inicial de adsorção celular, também é fundamental para o desenvolvimento de uma terapia antiviral. Esse conhecimento fundamenta o interesse pelo fondaparinux (Figura 10), um pentassacarídeo altamente sulfatado que mimetiza regiões específicas do SH.

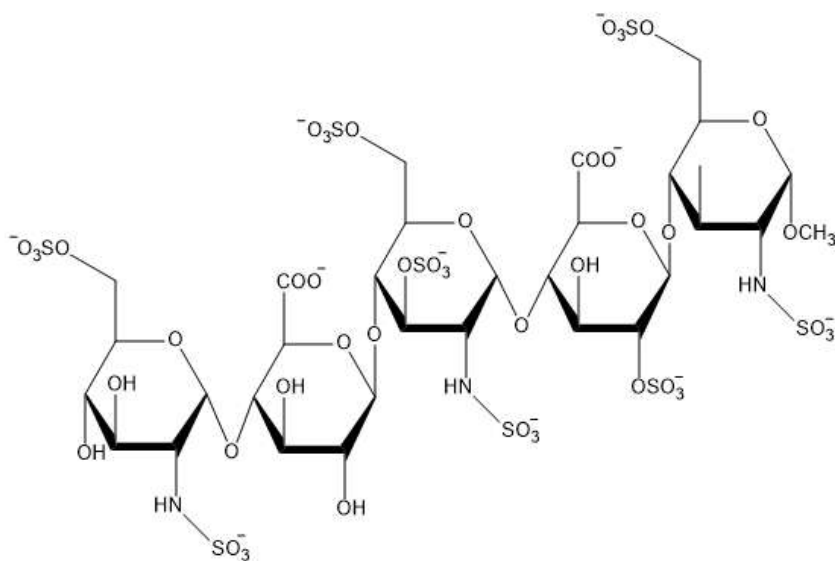
O fondaparinux é um anticoagulante sintético, estruturalmente semelhante aos glicosaminoglicanos endógenos. Além de sua conhecida atuação clínica como inibidor indireto do fator Xa (Chen et al., 2022). Sua elevada densidade de cargas negativas permite interações específicas com regiões altamente carregadas de proteínas virais, incluindo a proteína E do vírus da dengue, e têm sido relatada a capacidade do composto de interferir em etapas críticas da ligação e da replicação virais (Xie et al., 2017). Essa propriedade aliada a uma estrutura previsível e pequena, comparada a outros GAGs, somada a uma alta pureza sintética, tornam o fondaparinux um excelente candidato para estudos de interação molecular com proteínas virais, especialmente com a proteína E. Além disso, por se tratar de um fármaco já aprovado para uso clínico, o fondaparinux apresenta perfil de toxicidade e propriedades farmacológicas bem estabelecidas (European Medicines Agency, 2026), o que reforça seu potencial no desenvolvimento de uma terapia antiviral.

Figura 9: Estrutura química do sulfato de heparano.



Fonte: Elaborado pelo autor

Figura 10: Estrutura química do fondaparinux.



Fonte: Elaborado pelo autor

Entretanto, o desenvolvimento de fármacos e terapias antivirais não é tão simples, sendo uma tarefa que demanda muito capital e tempo. Uma maneira de planejar um fármaco é recorrer a métodos computacionais na fase pré-clínica. O planejamento de fármacos assistido por computador, do inglês *Computational-aided drug design* (CADD), pode ser definido como um conjunto de abordagens computacionais utilizadas para descobrir, desenvolver e

analisar fármacos e moléculas ativas a partir de suas propriedades bioquímicas (Sabe et al., 2021). O CADD se divide em duas abordagens: uma de planejamento de fármacos baseada na estrutura do alvo molecular, do inglês Structure-Based Drug Design (SBDD), e outra de planejamento de fármacos baseada na estrutura do ligante, do inglês Ligand-Based Drug Design (LBDD) (Piccirillo; Amaral, 2018). Modelagem por homologia, *docking* molecular, triagem virtual (VS), relação quantitativa estrutura-atividade (QSAR) e mapeamento farmacofórico tridimensional (3D) são, em geral, os principais componentes do CADD (Sabe et al., 2021).

A metodologia de *docking* molecular pode ser entre proteína-proteína, que explora a interação entre duas proteínas, e proteína-ligante, que explora a interação de pequenas moléculas no sítio de ligação de uma proteína alvo (Figura 10) (Forli et al., 2016; Li; Fu; Zhang, 2019). Os programas de *docking* molecular possuem basicamente dois componentes principais: um algoritmo de busca e uma função de *score*. Esses programas executam o algoritmo de busca no qual a conformação do ligante é avaliada recursivamente até que a convergência para a energia mínima seja alcançada e em seguida uma função de *score* é empregada para classificar as poses candidatas (Pagadala; Syed; Tuszynski, 2017). Estudos recentes classificam as funções de *score* (FS) em quatro categorias principais (Tabela 1), baseadas principalmente nos seus princípios individuais (Liu; Wang, 2015) .

1. Funções de *score* baseadas em física: calculam a afinidade de ligação somando as energias de interação entre o ligante e a proteína;
2. Funções de *score* empíricas: estimam a afinidade de ligação somando termos de energia individuais ponderados, calibrados com base em dados experimentais
3. Funções de *score* de potenciais baseados em conhecimento: derivam parâmetros de *score* analisando as preferências estatísticas de pares de átomos proteína-ligante em estruturas conhecidas, em vez de calcular a energia diretamente;
4. Funções de *score* baseadas em aprendizado de máquina: métodos baseados em aprendizado de máquina que frequentemente utilizam modelos não lineares complexos

Tabela 1: Categorias das funções de score, exemplos de cada grupo e as fórmulas que usam para calcular a energia de ligação

Categoria	Exemplo	Fórmula
Funções de <i>score</i> baseadas em física	DOCK	$E_{lig} = \sum_{i=1}^L \sum_{j=1}^R \left(\frac{A_{ij}}{r_{ij}^{12}} \frac{B_{ij}}{r_{ij}^6} + \frac{q_i q_j}{\epsilon(r_{ij}) r_{ji}} \right)$
Funções de <i>score</i> empíricas	X-Score	$E_{lig} = w_0 + w_1 \Delta G_{vdw} + w_2 \Delta G_{ligH} + w_3 \Delta G_{rot} + w_4 \Delta G_{hidro}$
Funções de <i>score</i> de potenciais baseados em conhecimento	PMF	$E_{lig} = \sum_{i=1}^L \sum_{j=1}^R -k_B T \ln[g(r)]$
Funções de <i>score</i> baseada em aprendizado de máquina	RF-Score	$E_{lig} = f_{RF}(xm)$ com $x_{ij} = \sum_{i=1}^L \sum_{j=1}^R \theta[d_{romp} - d_{ij}]$

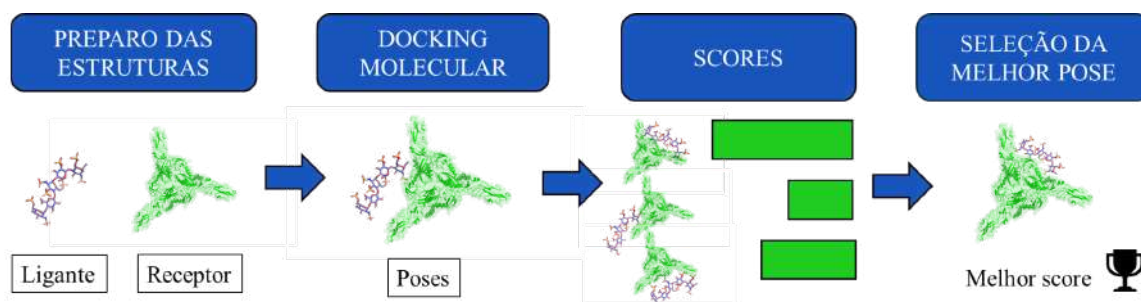
Adaptado de Li et al. 2019

As três primeiras funções de *score* (consideradas clássicas) são classificadas de acordo com os tipos de descritores utilizados e empregam principalmente o método de regressão linear, enquanto o quarto tipo de função de *score* incorpora o método de aprendizado de máquina de regressão não linear (Li; Fu; Zhang, 2019).

AutoDock é um dos conjuntos de softwares mais populares, gratuito e de código aberto para a modelagem molecular e a triagem virtual de pequenas moléculas a receptores macromoleculares. O conjunto inclui atualmente diversas ferramentas complementares como o AutoDock Vina, um programa de *docking* molecular baseado em uma função de score simples e em uma busca conformacional rápida por otimização de gradiente, e o AutoDock Tools, uma interface gráfica de usuário interativa para preparação de coordenadas, acoplamento e análise (Forli et al., 2016). O AutoDock Vina utiliza uma função de *score* empírica e não linear (Trott; Olson, 2010), mais comumente utilizadas em softwares de

docking proteína-ligante (Li; Fu; Zhang, 2019), inspirada na função X-score, baseada principalmente em termos gaussianos de interação, repulsão estérica, hidrofobicidade, ligações de hidrogênio e penalidade torsional (Trott; Olson, 2010) a partir da qual classifica as melhores poses geradas (Figura 11).

Figura 11: Esquema ilustrativo das macroetapas de uma simulação de docking molecular pelo Autodock Vina.



Fonte: Elaborado pelo autor

As funções de *score* empíricas, como a do Autodock Vina, estimam a afinidade de ligação de um complexo proteína-ligante por meio de uma soma ponderada de termos energéticos, cujos coeficientes de ponderação são determinados pelo ajuste (*fitting*) dos dados de afinidade de um conjunto de treinamento composto por complexos proteína-ligante com estruturas tridimensionais conhecidas (Ain et al., 2015). Devido aos seus termos de energia simples, essas FS são boas para prever a afinidade de ligação, a pose do ligante e a triagem virtual com baixo custo computacional, especialmente quando a flexibilidade do alvo não desempenha papel significativo (De Ávila; Pereira; Dardenne, 2018), mas são pouco adequadas para descrever a relação entre a afinidade de ligação e as estruturas cristalinas (Li; Fu; Zhang, 2019). A correta predição da energia livre de ligação associada ao deslocamento de moléculas de água pelo ligante e a integração adequada a flexibilidade proteica às FS empíricas continuam sendo uma limitação importante para estimativas precisas de afinidade (De Ávila; Pereira; Dardenne, 2018).

A natureza paramétrica dessas FS também introduz uma série de erros sistemáticos: a aplicabilidade geral dessas funções pode ser comprometida pelo conjunto de treinamento utilizado, dado que foram ajustadas a um conjunto de dados limitado (Ain et al., 2015). Além disso, modelos não lineares, ao fazerem uso de grande quantidade de descritores sem significado físico, oferecem o risco de produzir índices de desempenho excelentes em razão de *overfitting* e/ou viés na construção do conjunto de treinamento, capturando regras adotadas

durante a seleção de compostos ativos e *decoys* ao invés de interações fisicamente relevantes (De Ávila; Pereira; Dardenne, 2018).

Considerando tudo o que foi mencionado, este trabalho pretende dar continuidade ao estudo de Leite et al. 2026 sobre o arranjo trimétrico da partícula viral do DENV-2 e empregar metodologias computacionais, utilizando as 16 conformações representativas geradas pelas simulações de dinâmica molecular dos autores como receptores para um experimento de *docking* do fondaparinux. As análises desses sítios de interação por meio de um *docking* de múltiplas estruturas vai contribuir para a construção de um modelo e a melhor compreensão dos mecanismos de interação entre a proteína E e os GAGs, além de oferecer subsídios para o desenvolvimento de fármacos e de estratégias terapêuticas inovadoras contra a dengue.

Vale ressaltar que o fondaparinux é um anticoagulante sintético, enquanto a principal causa de mortalidade associada à infecção pelo DENV é a dengue severa, forma grave da doença caracterizada por aumento da permeabilidade vascular, trombocitopenia e manifestações hemorrágicas espontâneas. A administração de um anticoagulante nesse cenário clínico poderia potencializar o risco hemorrágico representando um risco para o paciente. Dessa forma, o fondaparinux deve ser compreendido neste trabalho não como candidato terapêutico, mas como modelo estrutural que valida computacionalmente o sítio drogável e o farmacóforo necessário para o reconhecimento desse sítio.

3. OBJETIVO

3.1 Objetivo geral

O projeto tem como objetivo geral realizar a modelagem molecular do complexo entre o fondaparinux e as estruturas da proteína E do DENV-2, desenhando um modelo da interação entre o trímero e os GAGs que atuam como receptores celulares.

3.2 Objetivos específicos

- I. Realizar o *docking* molecular do fondaparinux com o sítio mapeado do DIII, (Leite et al., 2026) dentro das 16 conformações do trímero;
- II. Analisar as melhores poses e *scores* gerados a partir desses *dockings*;
- III. Analisar o perfil das interações predominantes entre o ligante e o receptor em cada uma das poses e discutir sobre as interações mais relevantes, em termos de recorrência e estabilidade;
- IV. Agrupar essas poses em clusters para averiguar se há um padrão de comportamento entre a distribuição dessas poses que permite visualizar, caracterizar e distinguir esses clusters;
- V. Selecionar e caracterizar as poses que melhor representam as 800 poses geradas e a interação do sistema como um todo;

4. MATERIAIS E MÉTODOS

4.1 Obtenção e preparação das estruturas

As estruturas empregadas neste estudo foram obtidas a partir do agrupamento de trajetórias de dinâmica molecular (DM) do trímero, realizado em trabalhos prévios do grupo de pesquisa (Leite et al., 2026). As DMs utilizaram como base a estrutura da proteína E do DENV2 depositada no *Protein Data Bank* sob o código 3J27, determinada por criomicroscopia eletrônica (Cryo-EM). Com base nos resultados do HCA, foram definidas 16 estruturas representativas do trímero (Leite et al., 2026). Para esse trabalho, cada uma das estruturas representativas foi utilizada e foram rotuladas de c1 à c16 de forma arbitrária.

Para a visualização das estruturas, utilizou-se o PyMOL, um software que permite a representação tridimensional e o alinhamento estrutural. Como as estruturas iniciais não se encontravam previamente orientadas no eixo Z, a gridbox nas simulações de *docking* molecular estava desalinhada, o que fazia com que o espaço de busca do software não fosse homogêneo entre as cadeias do trímero. Para resolver esse problema, a primeira conformação foi alinhada ao longo do eixo Z, e as demais foram alinhadas com base no alinhamento da primeira.

Após essa etapa, procedeu-se à preparação dos sistemas para o *docking* molecular. Para isso, utilizou-se o AutoDock Tools. Como as estruturas foram obtidas a partir de estudos prévios, não foi necessário realizar etapas adicionais de preparo estrutural, uma vez que já se encontravam tratadas para representar condições fisiológicas (pH 7,4). Dessa forma, o único procedimento realizado para as proteínas foi a atribuição de cargas de Gasteiger e a conversão dos arquivos para o formato pdbqt.

Por outro lado, o ligante fondaparinux exigiu etapas adicionais de preparação. Inicialmente, verificou-se que sua estrutura não apresentava estados de protonação compatíveis com pH fisiológico, sendo necessária a correção. Além disso, as ligações rotacionáveis não estavam devidamente definidas, exigindo sua definição através do AutoDock Tools para garantir uma amostragem conformacional adequada durante as simulações de *docking*. Por fim, o ligante foi convertido para o formato pdbqt.

4.2 *Docking* molecular

Em virtude das limitações relacionadas ao hardware disponível, ao sistema operacional e à organização da estrutura de arquivos necessária à execução dos cálculos, optou-se pelo uso do Google Colaboratory (Colab) como ambiente de desenvolvimento para a realização dos experimentos de *docking*. Essa escolha foi motivada, principalmente pela sua natureza em nuvem, que permite a execução de notebooks Jupyter diretamente no navegador, facilitando o acesso a recursos computacionais escaláveis e a padronização do ambiente de execução. Além disso, o Colab possibilita a integração eficiente com bibliotecas e com o Google Drive, proporcionando acesso a diversas ferramentas e atalhos que são fundamentais para a obtenção dos resultados propostos e a reprodutibilidade dos experimentos. Foi utilizada a versão mais recente disponível do Python (3.12.13), bem como as versões atualizadas do AutoDock Vina (1.2.7) e do AutoDock Tools (1.5.7), assegurando compatibilidade e acesso às funcionalidades mais recentes dessas ferramentas.

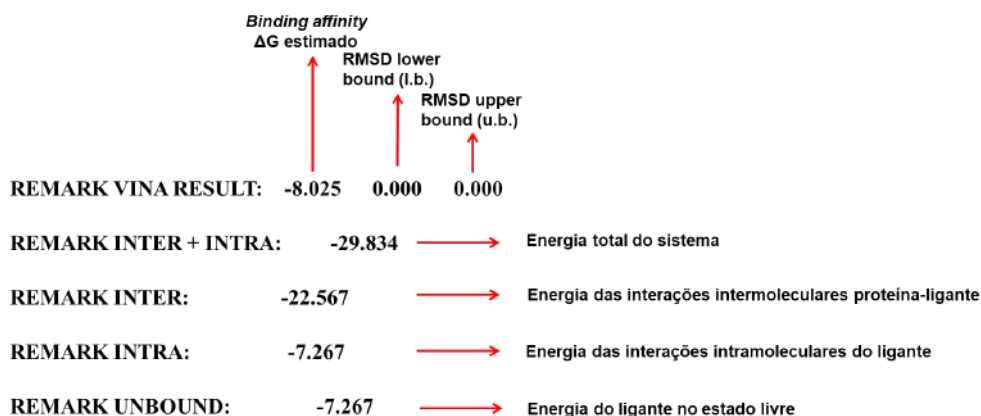
A metodologia para esse trabalho não consiste em um *docking* tradicional, mas de uma metodologia similar a de um *ensemble docking*. O *ensemble docking* representa um método inovador em relação ao *docking* convencional por incorporar a flexibilidade conformacional do receptor na análise dos resultados (Maltarollo *et al.*, 2025). As proteínas não são entidades estáticas, mas exibem movimentos intrínsecos e não aleatórios que são cruciais para suas interações com outras moléculas, conferindo ao *ensemble docking* maior similaridade à realidade biológica. Entretanto, esse trabalho difere de um *ensemble docking*, pois o objetivo final do *ensemble docking* é somente selecionar a pose de melhor *score* para estudo, enquanto aqui serão selecionadas algumas poses por critérios que irão incluir o *score* para caracterizar o sistema, sendo adotado para esse trabalho o nome de *docking* de múltiplas estruturas.

Ao se utilizar nos *dockings* uma *gridbox* centralizada na estrutura, as interações proteína-ligante talvez não se dirigissem para o sítio previamente mapeado, pois a maioria dos resíduos de interesse previamente mapeados (Leite *et al.*, 2026) estavam localizados na superfície da proteína. Isso seria um problema caso acontecesse, pois não simularia as interações da proteína com os receptores celulares visto que o interesse recai sobre as interações da superfície externa do capsídeo. Para solucionar esse problema, primeiramente foi realizado um experimento de *docking* molecular em uma conformação do trímico, selecionada aleatoriamente, para definir a posição ideal da caixa, mais próxima da superfície ou mais centralizada nas estruturas. Para o *docking* na superfície foram utilizadas as seguintes

coordenadas de centro da caixa $x=122$, $y=60$ e $z=80$, enquanto para o *docking* no centro da caixa foram estabelecidas as coordenadas $x=122$, $y=60$ e $z=60$. Ambos com a dimensão da caixa em 60, 60, 60 e o espaçamento padrão de $0,375\text{\AA}$, recobrimdo boa parte da proteína, desde o centro até parte das hélices. Esses parâmetros foram definidos utilizando como referência os estudos prévios de localização do DP-C16-T, além da grande quantidade de resíduos de aminoácidos carregados positivamente localizados naquela região, gerando um estudo mais direcionado. Os resultados obtidos nessa etapa preliminar permitiram identificar a região mais favorável à interação ligante-proteína, servindo como base para a definição dos parâmetros empregados nas simulações subsequentes (*exhaustiveness* de 64 e 50 poses por simulação).

O fondaparinux, com 32 torções ativas e aproximadamente 1728 Da, representa um desafio computacional para o Vina. O *exhaustiveness* padrão de 8 foi projetado para moléculas menores com 5-10 torções. Usar 64 representa 8 vezes mais esforço amostral para garantir que o espaço conformacional seja adequadamente explorado. Enquanto as 50 poses por conformação garantem densidade estatística para a análise de clustering.

Após o experimento inicial, foram realizados 16 *dockings*, o que corresponde a uma simulação para cada conformação da proteína, com tempo médio entre 8 e 11 horas por *docking*, gerando outputs analisáveis (Figura 12). As 50 poses geradas por conformação foram numeradas de 1 a 50 por ordem crescente de score, a pose de melhor score foi a pose 1 enquanto a pose de pior score foi a pose 50. Obteve-se um total de 800 poses ao final do experimento, o que possibilitou uma análise mais abrangente do comportamento conformacional e das interações moleculares envolvidas.

Figura 12: Output do Vina.

Fonte: Elaborado pelo autor

4.3 PLIP

Foi realizada uma análise das interações intermoleculares entre o fondaparinux e as conformações da proteína, utilizando o software PLIP (Protein–Ligand Interaction Profiler), que permite identificar e caracterizar as interações predominantes (tipos de ligação, resíduos de aminoácidos e átomos envolvidos, e distâncias interatômicas).

Tabela 2: Tipos de interação detectadas pelo PLIP

Tipo de interação	Distância máxima detectada pelo PLIP (em Å)	Justificativa dessa distância	Ângulo máximo detectado pelo PLIP
Ligações de hidrogênio	4,1	Distância máxima entre doador e acceptor de ligação de hidrogênio (Hubbard & Haider, 2001) + 0,6 Å	Mínimo 100°
Interações hidrofóbicas	4,0	Limite de distância para detecção de contatos hidrofóbicos	-
Interações π - π (empilhamento aromático)	5,5	Máx. distância para empilhamento paralelo ou deslocado (McGaughey, 1998)	Desvio 30°
Interações cátion- π	6,0	Distância máxima entre o átomo carregado e o centro do anel aromático (Gallivan e Dougherty, 1999)	-
Pontes salinas	5,5	Distância máxima entre os centros de carga para pontes salinas (Barlow e Thornton, 1983) + 1,5	-
Interações envolvendo halogênios	4,0	Distância máxima entre os centros de carga para pontes salinas (Barlow e Thornton, 1983) + 1,5 e halogênio (Ligações de halogênio em moléculas biológicas, Auffinger) + 0,5	Aceptor 120° Doador 165° Desvio 30°

Adaptado documentação do PLIP

As interações identificadas foram organizadas em planilhas, contendo as informações de cada ligação associadas aos sistemas e às poses correspondentes. Essa organização permitiu a realização de análises quantitativas, incluindo a frequência de interações por tipo, identificação de resíduos recorrentes e a comparação entre diferentes sistemas estruturais.

4.4 HCAs

Para cada pose obtida no *docking*, foram extraídas as coordenadas tridimensionais dos átomos, o valor de score predito pelo algoritmo de *docking* e um identificador de origem, correspondente ao sistema estrutural analisado. A partir das coordenadas atômicas, foi calculado o centróide de massa de cada pose, utilizando átomos pesados, definido como o ponto único no espaço que representa a posição média ponderada de toda a massa do sistema (Halliday; Resnick; Walker, 2017).

Para a análise dos resultados do *docking* molecular, empregamos a Análise Hierárquica de Clusters (HCA), visando distribuir as poses obtidas em clusters homogêneos. Foram realizados dois agrupamentos, o primeiro baseou-se em parâmetros de centro de massa por permitir uma representação mais simples das moléculas e conseguir fornecer uma boa noção do posicionamento espacial delas no espaço e o segundo baseou-se no desvio quadrático médio (RMSD) por permitir um entendimento geométrico acerca das poses clusterizadas. O número ótimo de clusters foi determinado pela regra do “cotovelo” (*elbow method*), uma abordagem visual para encontrar o número de clusters (k), baseada na análise do *within-cluster sum of squares* (WCSS), em português soma dos quadrados intra-cluster conhecida como Inércia, (Michilles, 2026) evitando a escolha arbitrária de parâmetros. A matriz de distâncias foi construída com a função *pdist*, e o agrupamento foi realizado por meio do método de ligação (*linkage*), o que permitiu a construção de um dendrograma representativo da relação hierárquica entre as conformações.

A partir da definição dos clusters, cada pose foi rotulada conforme o grupo ao qual pertence, o que permitiu identificar regiões preferenciais de ligação, reduzir a redundância estrutural e organizar as poses em conjuntos estruturalmente semelhantes. Essa clusterização exemplifica a dispersão das poses na superfície da proteína.

O agrupamento mais populoso foi então submetido a uma nova análise de HCA, gerando sub-clusters baseados em RMSD, de modo a avaliar o comportamento predominante das poses geradas no *docking*. A visualização desses clusters será realizada em 3D, tanto a clusterização inicial por centro de massa quanto a clusterização específica por RMSD, possibilitando o mapeamento preciso de sua distribuição na superfície da macromolécula.

Após a obtenção desses clusters, foi realizada uma análise detalhada para elucidar as propriedades químicas e biológicas do agrupamento predominante. Foram avaliados o perfil de interações, as conformações mais favoráveis à interação ligante-receptor, as poses de maior score e a viabilidade de utilizar estruturas representativas para caracterizar cada cluster.

Para cada cluster identificado, foram selecionadas conformações representativas com base no valor de afinidade, priorizando poses com melhores *scores* dentro de cada agrupamento e poses que as coordenadas do centro de massa possuem o menor distanciamento entre as poses do cluster. Essas poses foram então caracterizadas quanto às ligações estabelecidas por meio de dois softwares distintos, o PLIP e o Discovery Studio. As interações detectadas pelo PLIP foram então visualizadas em 3D no Pymol enquanto o Discovery Studio gerou uma visualização 2D dessas interações.

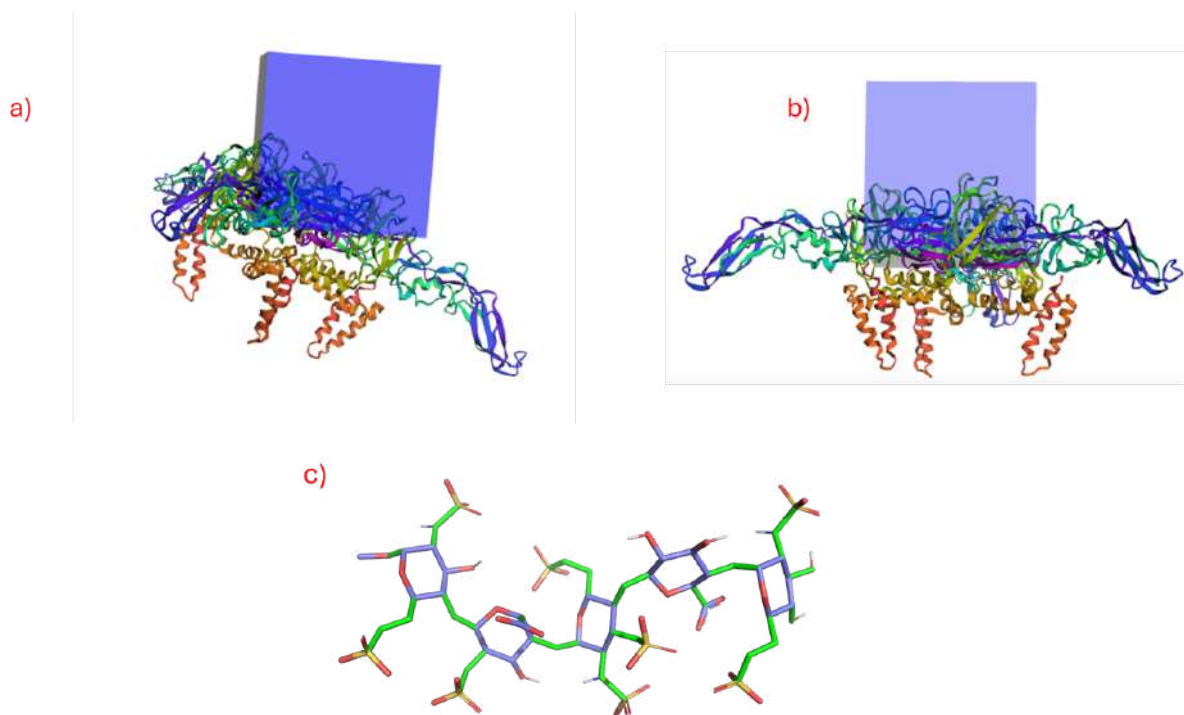
Uma análise integrada entre os resultados da clusterização estrutural e da análise do PLIP permite uma interpretação mais abrangente dos dados, possibilitando a identificação de regiões de ligação recorrentes e a avaliação da consistência das interações entre diferentes poses.

5. RESULTADOS E DISCUSSÃO

5.1. Docking de múltiplas estruturas

As 16 conformações representativas definidas por Leite et al.2026 foram alinhadas ao longo do eixo Z, de modo que a caixa ficasse homogênea sobre a superfície do trímico, de modo que o *docking* ocorresse apenas sobre a superfície do trímico, sem a entrada do fondaparinux devido a uma área ter maior profundidade que a outra. Já o ligante fondaparinux passou por uma correção para o pH fisiológico e pela definição de ligações rotacionáveis (Figura 13).

Figura 13: Preparo das estruturas a) trímico desalinhado; b) trímico alinhado; e c) Ligante Fondaparinux preparado para docking com as ligações rotacionáveis em verde



Fonte: Elaborado pelo autor

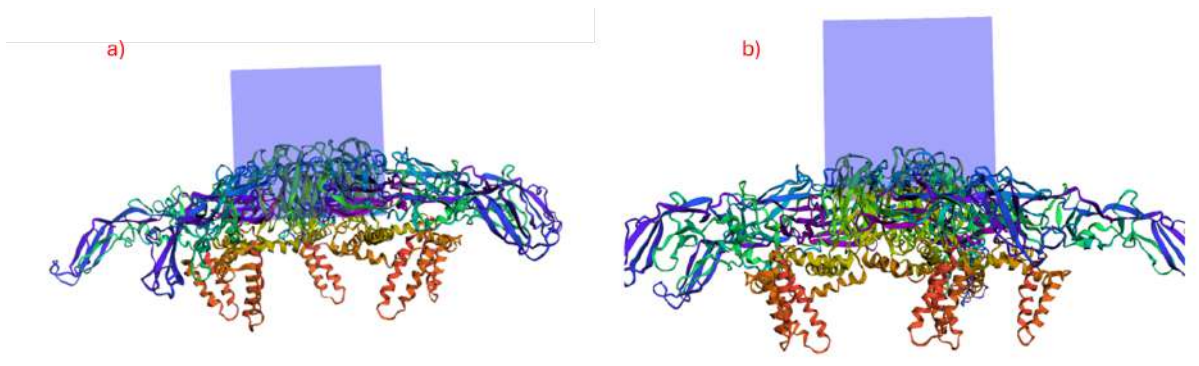
Vale ressaltar que a proteína E do DENV-2 é um sistema dinâmico, que em solução explora continuamente um conjunto de estados conformacionais. A estrutura de *cryo*-EM 3J27 representa apenas um *snapshot* de baixa resolução de um sistema que, biologicamente, existe como um conjunto. Como demonstrado em LEITE et al., 2026, a resolução limitada da *cryo*-EM implica incerteza significativa nas posições das cadeias laterais, especialmente em

resíduos carregados e flexíveis como a Lys 291 e a Lys 295. Isso significa que um único receptor de *docking* representaria não apenas uma conformação, mas também uma conformação com geometria local potencialmente incorreta.

A estratégia de *docking* de múltiplas estruturas a partir de 16 conformações representativas de clusters de dinâmica molecular resolve dois problemas simultaneamente. Primeiro, captura a diversidade conformacional do sistema e segundo, permite identificar quais conformações são mais receptivas ao fondaparinux, ambas questões que são fundamentais na construção de um modelo das interações da proteína com os GAGs por abranger tanto.

O experimento inicial de definição da melhor posição da caixa (Figura 14), foi importante para garantir que o experimento simulasse o meio biológico de forma que a parte externa da proteína do vírus interagisse com o receptor e para averiguar qual a tendência de comportamento do fondaparinux com toda a área de proteína disponível.

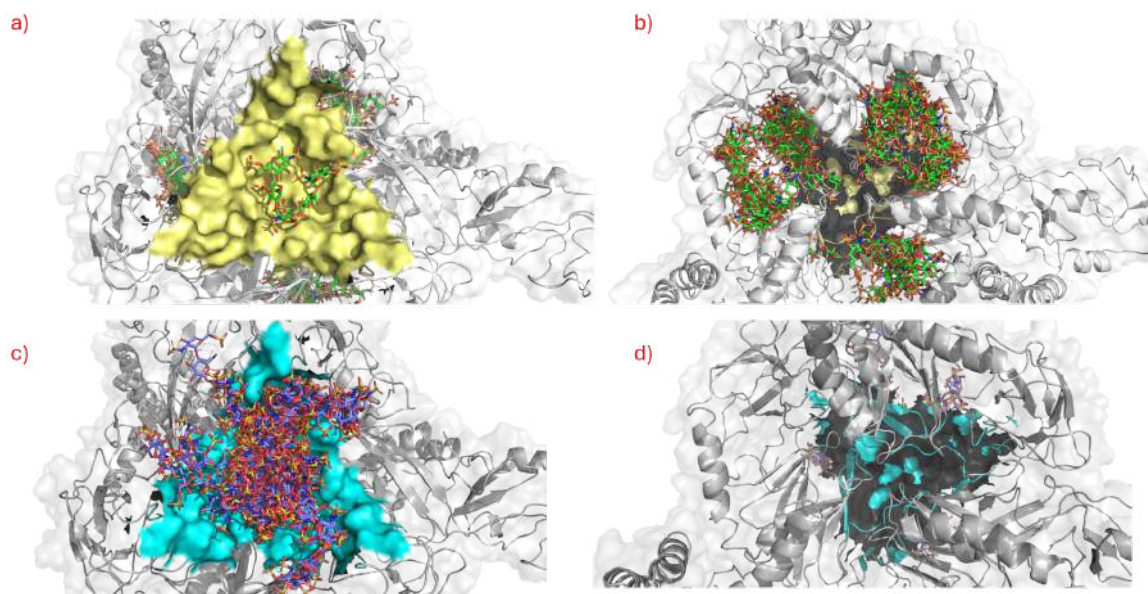
Figura 14: Experimento realizado com a grid box em diferentes posições: a) grid box centralizada na estrutura; e b) grid box na superfície



Fonte: Elaborado pelo autor

Com a realização desse experimento foi possível perceber que o experimento com a caixa dentro teve uma quantidade muito pequena de poses na superfície (Figura 15), o que não gera as interações desejadas, visto que o objetivo inicial era simular a interação dos receptores com a superfície externa do vírus. Sendo, portanto estabelecido como critério para o restante dos experimentos a utilização da caixa na superfície.

Figura 15: Resultados do docking com a caixa dentro e com a caixa fora da estrutura da proteína a) superfície da proteína com a caixa dentro da estrutura; b) parte interna da proteína com a caixa dentro; c) superfície da proteína com a caixa fora; d) parte interna da proteína com a caixa fora



Fonte: Elaborado pelo autor

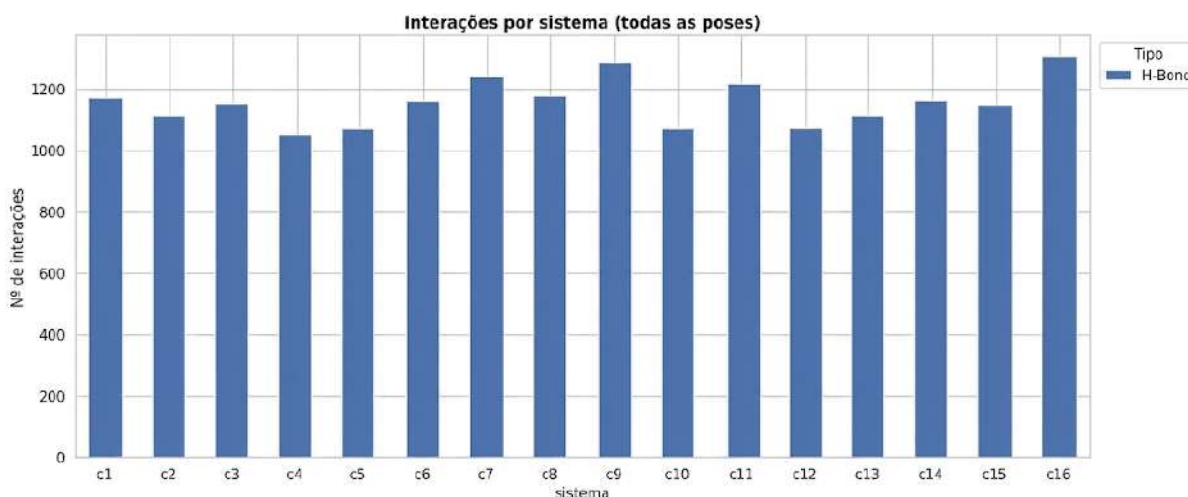
Com a definição dos parâmetros do *docking* (posição e tamanho da caixa, número de poses e exaustiveness) e a realização dos experimentos iniciais, seguiu-se para a etapa do *docking* de múltiplas estruturas, com cada conformação recebendo um nome de forma arbitrária de c1 à c16 e a obtenção de 800 poses do ligante fondaparinux ao final dessa etapa.

5.2 Análise do *docking* de múltiplas estruturas

Foram detectadas um total de 18.518 interações entre as 800 poses do fondaparinux e as conformações da proteína, com o software PLIP classificando todas como interações de hidrogênio. Possivelmente interações iônicas entre os grupos sulfato do fondaparinux e as lisinas foram capturadas como interações de hidrogênio pela geometria doador-aceptor satisfeita pelos hidrogênios do amônio dos resíduos de aminoácidos interagindo com os oxigênios do sulfato do fondaparinux. Entretanto, isso não invalida o resultado, apenas demonstra que essas interações possuem um caráter misto eletrostático/ligação de hidrogênio. A ausência de interações hidrofóbicas, π -cátion ou π -stacking é consistente com a natureza do fondaparinux: um ligante polar, carregado negativamente, sem anéis aromáticos ou regiões hidrofóbicas extensas.

O número de interações de hidrogênio por sistema varia de 1050 (c4) a 1300 (c16) (Figura 16). As conformações com mais interações não são necessariamente as que tiveram os melhores valores de score, confirmando que a quantidade de interações não garante a estabilidade do sistema, e que a força, a distância e a geometria das interações importam mais do que a quantidade em si.

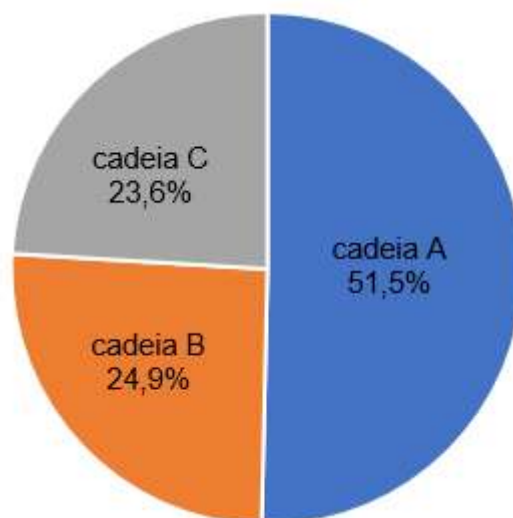
Figura 16: Quantidade de interações com o fondaparinux por conformação do trímero



Fonte: Elaborado pelo autor

Essas interações não se distribuíram de maneira igual entre as três cadeias monoméricas que compõem o trímero. A cadeia A concentrou isoladamente mais de 50% do total de interações (Figura 17), revelando uma assimetria pronunciada na forma como o fondaparinux interage com as três subunidades que compõem o arranjo trimétrico. Essa distribuição assimétrica é consistente com a localização do *druggable pocket* DP-C16-T, cuja posição espacial está majoritariamente associada à cadeia A. A correspondência entre a dominância da cadeia A nas interações globais e sua posição central na definição do *pocket* constitui uma validação estrutural interna dos resultados, pois o fondaparinux interage preferencialmente com a cadeia que abriga o sítio biologicamente relevante. Essa interpretação é reforçada pela análise dos resíduos com maior frequência de interação: entre os 20 resíduos com o maior número acumulado de interações ao longo de todo o experimento, a grande maioria corresponde a resíduos integrantes ou adjacentes ao *pocket* DP-C16-T (Figura A-1).

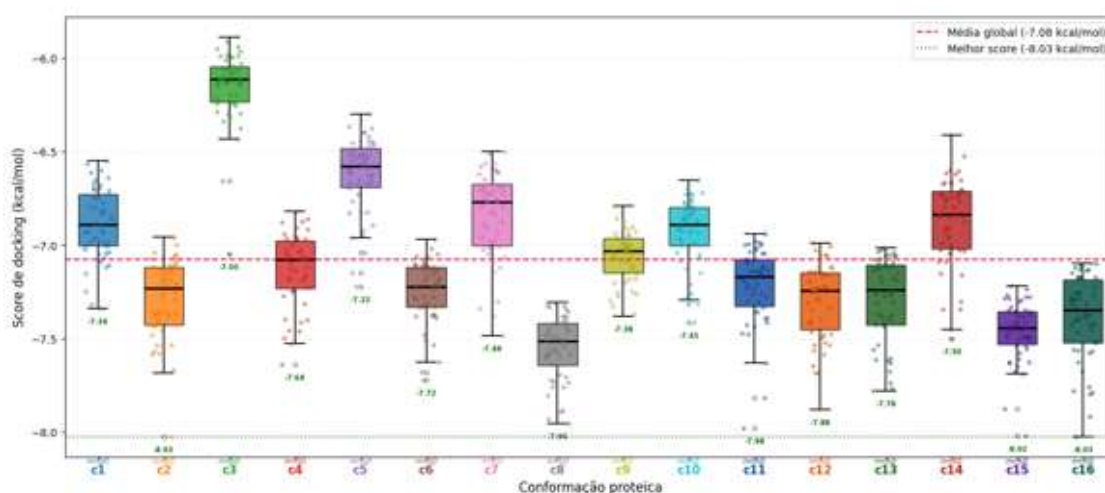
Figura 17: Distribuição das interações do fondaparinux com o trímico por cadeia monomérica



Fonte: Elaborado pelo autor

Antes de realizar a clusterização, foi realizada uma análise dos scores gerados por cada conformação. Ao analisar a Figura 18 é possível perceber que existem conformações que recebem de forma mais favorável o ligante que outras. A c3 se mostra como desfavorável no aspecto energético quando comparada a todas as outras e a c8 e c16 se mostram mais favoráveis a esse tipo de interação.

Figura 18: Boxplot de distribuição dos scores pelas 16 conformações



Fonte: Elaborado pelo autor

As 800 poses foram então clusterizadas pelo centro de massa, com o número de clusters definido pelo método *Elbow* e a clusterização realizada com um *linkage Ward*. A

curva de inércia do elbow plot mostra uma redução acentuada até $k=5$, com ganho marginal a partir desse ponto (Figura B-1). Como a clusterização visava manter o número estabelecido pelo método de *Elbow* foi estabelecido para esse HCA um *linkage Ward* que tende a gerar uma clusterização mais robusta pois se baseia no Método de Ward que calcula a soma dos quadrados dentro dos clusters, usando a distância euclidiana ao quadrado. Foram gerados 5 clusters para as 800 poses e eles se dividiram da seguinte forma:

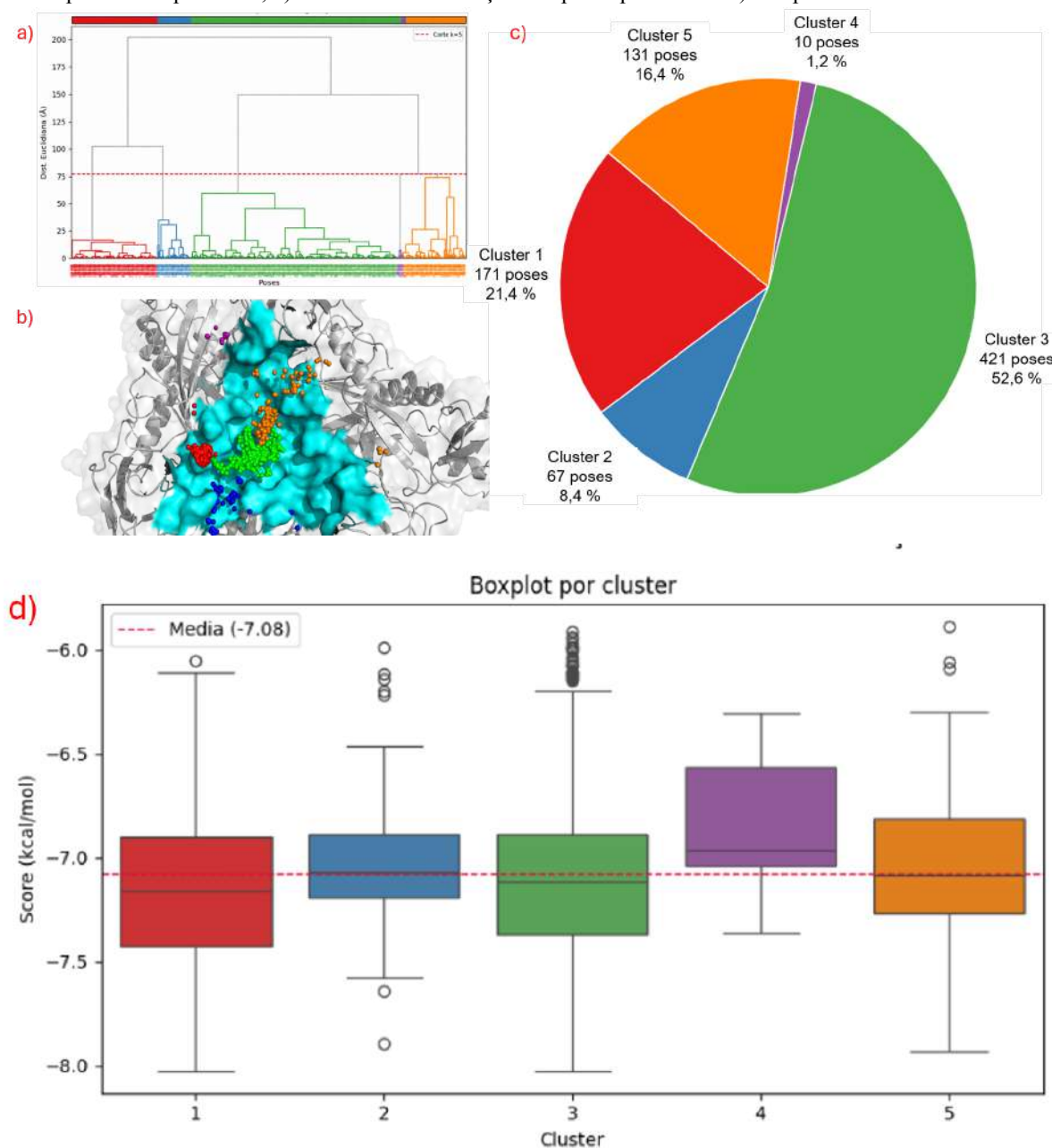
Tabela 3: Classificação dos clusters

Cluster	Cor	Nº de poses	Score médio (kcal/mol)
1	Vermelho	171	-7.14
2	Azul	67	-7.03
3	Verde	421	-7.08
4	Roxo	10	-6.84
5	Laranja	131	-7.04

Fonte: Elaborado pelo autor

O dendrograma gerado a partir dessa clusterização (Figura 18a) confirma essa separação, juntamente com a visualização espacial dessas poses que foi gerada utilizando o Pymol (Figura 18b).

Figura 19: Visualização e dados gerais da clusterização a) Dendograma; b) Centros de massa das poses sobre a superfície da proteínas; c) Gráfico de distribuição das poses por cluster d) Boxplot de score x cluster



Fonte: Elaborado pelo autor

Os clusters de maior destaque são o Cluster 1 que tem a maior variância, o Cluster 3 tem o maior número de poses e o maior número de outliers e o Cluster 4 que tem o menor número de poses e a menor variância. Os outros clusters tiveram padrões semelhantes, visto a sua quantidade de poses e comportamento do sistema.

O boxplot (Figura 18d) revela que os cinco clusters têm distribuições de score notavelmente similares, sendo praticamente degenerados com médias entre -6.84 e -7.14 kcal/mol e média global de -7.08 kcal/mol. A degenerescência dos clusters difere do que foi visto ao analisar as conformações (Figura 18) que variam de forma mais considerável nos scores, o que demonstra que as conformações estão bem distribuídas entre os clusters. Por um lado, indica que o fondaparinux não discrimina fortemente entre regiões da superfície protéica em termos de score bruto, característica esperada para um ligante de superfície altamente carregado interagindo com uma proteína de superfície também carregada. Por outro lado, a ausência de separação clara por score reforça a importância da geometria como critério de classificação complementar ao *score*. O que diferencia biologicamente os clusters não é a energia de ligação, mas a posição que eles ocupam na superfície da proteína, isso significa que o fondaparinux adota poses de alta qualidade em sub-regiões específicas do cluster, não uniformemente distribuídas.

O Cluster 3 se destaca não apenas pela quantidade de poses, mas também pela localização espacial, pois está na região do DIII e do *druggable pocket* DP-C16-T (LEITE et al., 2026). Isso indica que mesmo que o fondaparinux interaja de formas distintas com a proteína, há favorecimento das interações de GAGs com essa região e que as heparinas são interessantes no desenvolvimento de terapias que atuem nesse *pocket*.

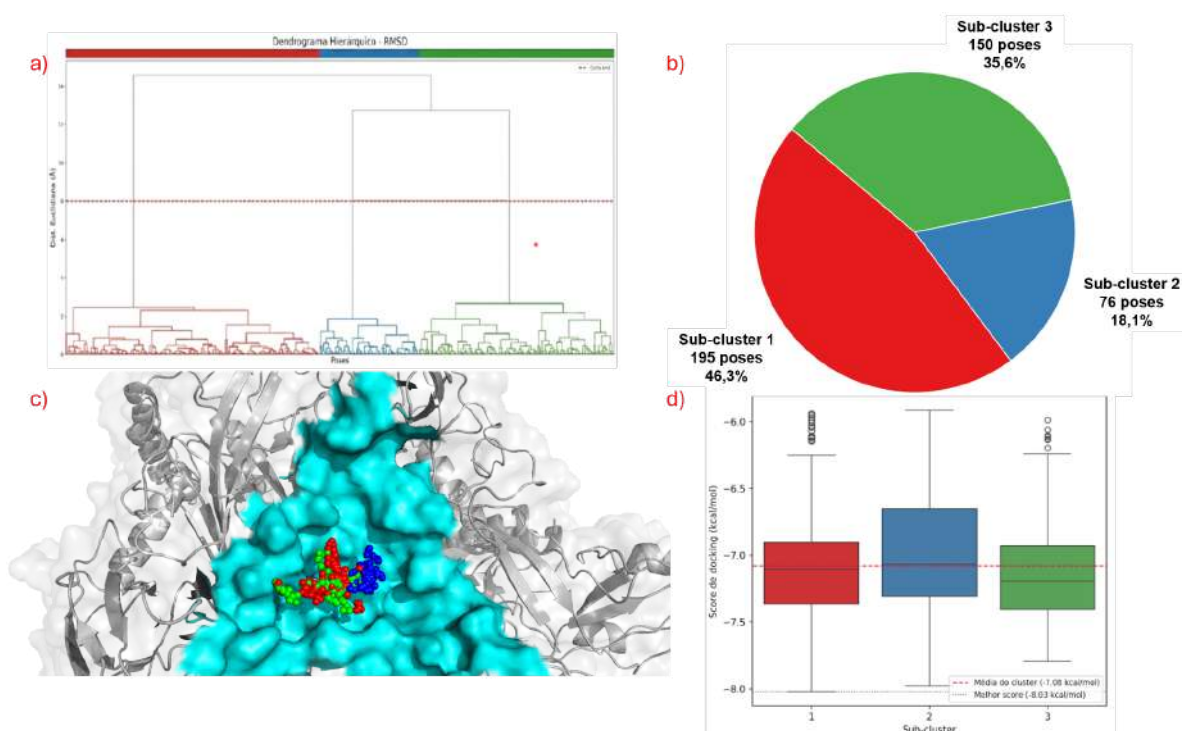
A clusterização inicial por centro de massa foi utilizada como etapa de triagem espacial para identificar a dispersão do ligante na superfície da proteína. Para o cluster de interesse biológico (Cluster 3), uma segunda clusterização por RMSD foi aplicada às suas 421 poses, capturando a posição, a orientação e a conformação interna do fondaparinux, resultando em três modos de ligação geometricamente distintos.

A sub-clusterização por RMSD difere metodologicamente da clusterização inicial por centro de massa. Enquanto o centro de massa agrupa as poses com base na posição espacial do ligante na superfície do sítio mapeado na proteína E, o RMSD agrupa as poses com base na semelhança conformacional entre si. Poses que com centroide similar podem apresentar grande variação de RMSD se o ligante estiver rotacionado ou apresentar conformações internas diferentes.

O elbow plot da sub-clusterização do Cluster 3 indica $k=3$ como número ótimo de clusters (Figura B-2). O dendograma (Figura 20a) e o gráfico (Figura 20b) ilustram como

essas poses foram agrupadas. A projeção (Figura 20c) mostra que os três sub-clusters se sobrepõem parcialmente no espaço do centroide, mas diferem na geometria interna das poses adotadas pelo fondaparinux. O boxplot (Figura 20d), também demonstra uma degenerescência entre esses sub-clusters. Esses resultados têm implicações diretas para o design de inibidores: existem três modos conformacionais distintos pelos quais um mimetizador de GAG pode se acomodar no pocket trimétrico, e cada um deles pode ser explorado como ponto de partida para otimização estrutural.

Figura 20: Visualização e dados gerais da sub-clusterização a) Dendograma da sub-clusterização; b) Distribuição das poses por sub-cluster; c) centros de massa do cluster 3 após a sub-clusterização por RMSD; e d) Boxplot de score x sub-clusters

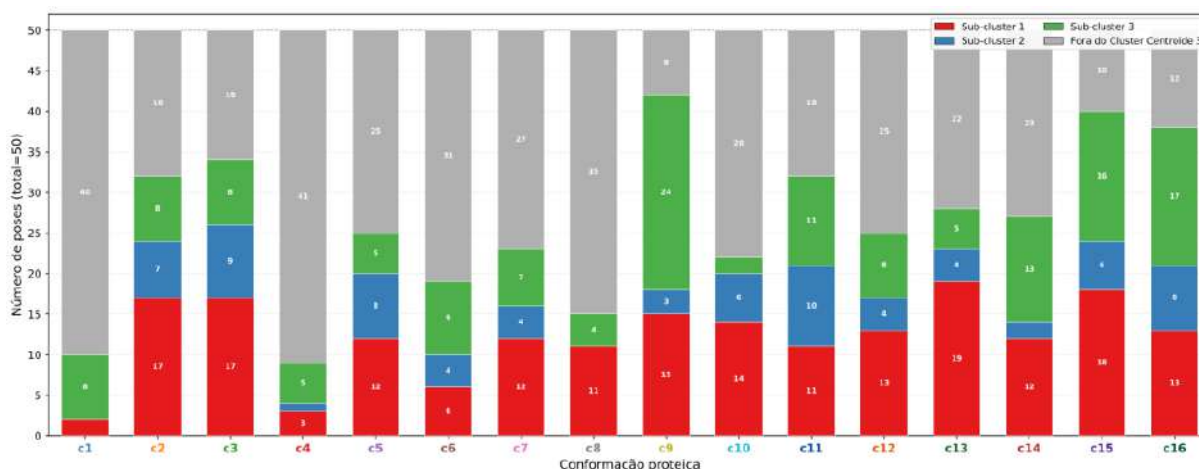


Fonte: Elaborado pelo autor

A distribuição dos sub-clusters no Cluster 3 (Sub-cluster 1: 46.3%, Sub-cluster 3: 35.6%, Sub-cluster 2: 18.1%) indica que o espaço conformacional do pocket trimétrico é ocupado de forma desigual; o modo 1 é claramente dominante, mas o modo 3 apresenta contribuição significativa e é energeticamente mais favorável.

A Figura 21 integra toda a informação de clusterização em uma única visualização. Cada barra representa as 50 poses de cada conformação, divididas entre os três sub-clusters (coloridos) e as poses fora do Cluster 3 (cinza).

Figura 21: Distribuição das poses por conformação do docking(c1-c16) dentro do cluster 3, cluster por centro de massa que foi definido na figura 19 e caracterizado como biologicamente relevante, sendo dividido em sub-clusters por RMSD na figura 20 para definir os modos de ligação do sistema fondaparinux-trímero



Fonte : Elaborado pelo autor

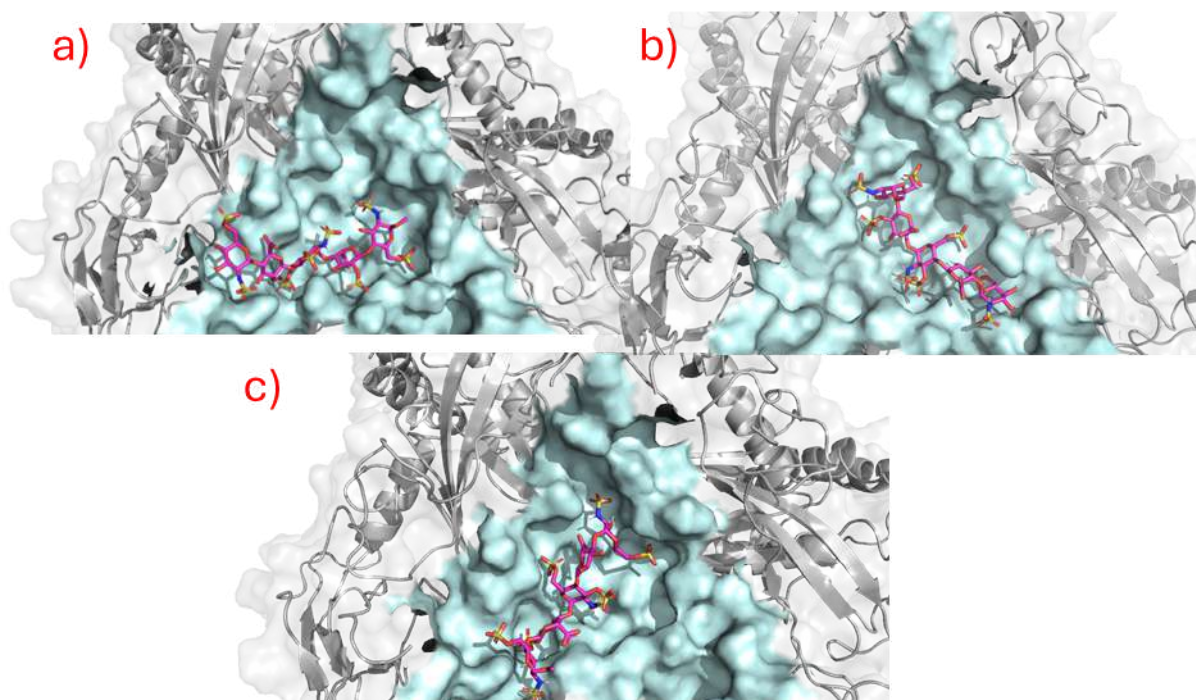
As conformações c1 e c4 são os casos mais extremos: com aproximadamente 20% das poses no sítio de interesse, distribuídas em sua maioria no Sub-cluster 3 (região representada em vermelho). Isso indica uma maior compactação dessas conformações, possivelmente os *pockets* estão fechados demais para acomodar favoravelmente o fondaparinux, gerando uma preferência por outros sítios da proteína. Caso esse indício se comprove, será possível estabelecer uma relação entre a compactação da interface e a baixa ocupação do sítio de *docking*, quanto maior a compactação da interface trimétrica o volume do *pocket* é reduzido, limitando o acesso do fondaparinux.

A conformação c9 é o oposto: apenas 8 poses fora do Cluster 3 (cinza), bem distribuídas entre os sub-clusters e com distribuição expressiva no Sub-cluster 3 (24 poses). Seu *pocket* possivelmente está numa geometria de abertura ótima para o Sub-cluster 3 (verde), mostrando que c9 é uma boa conformação para caracterizar a interação do trímero com GAGs. Além de c9, c15 e c16 também têm alta ocupação do sítio, com distribuição relevante pelos três sub-clusters, respectivamente de 80% e 76% das posições dessas conformações no cluster 3.

A partir de toda a construção do HCA foram selecionadas as poses que melhor representavam esses sub-clusters, para serem usadas em estudos futuros de dinâmica molecular. Para selecionar essas poses, os critérios adotados foram as poses de melhor score

que estavam nas conformações mais populosas de cada sub-cluster, pois a estabilidade é proporcional à população do estado. As poses selecionadas foram para o sub-cluster 1: c13 pose 1, sub-cluster 2: c11 pose 1, sub-cluster 3: c9 pose 2.

Figura 22: Poses representativas de cada sub-cluster a)Sub-cluster 1: c13 pose 1 score = -7,779 kcal/mol
b)Sub-cluster 2: c11 pose 1 score = -7,978 kcal/mol c)Sub-cluster 3: c9 pose 2 score = -7,368 kcal/mol

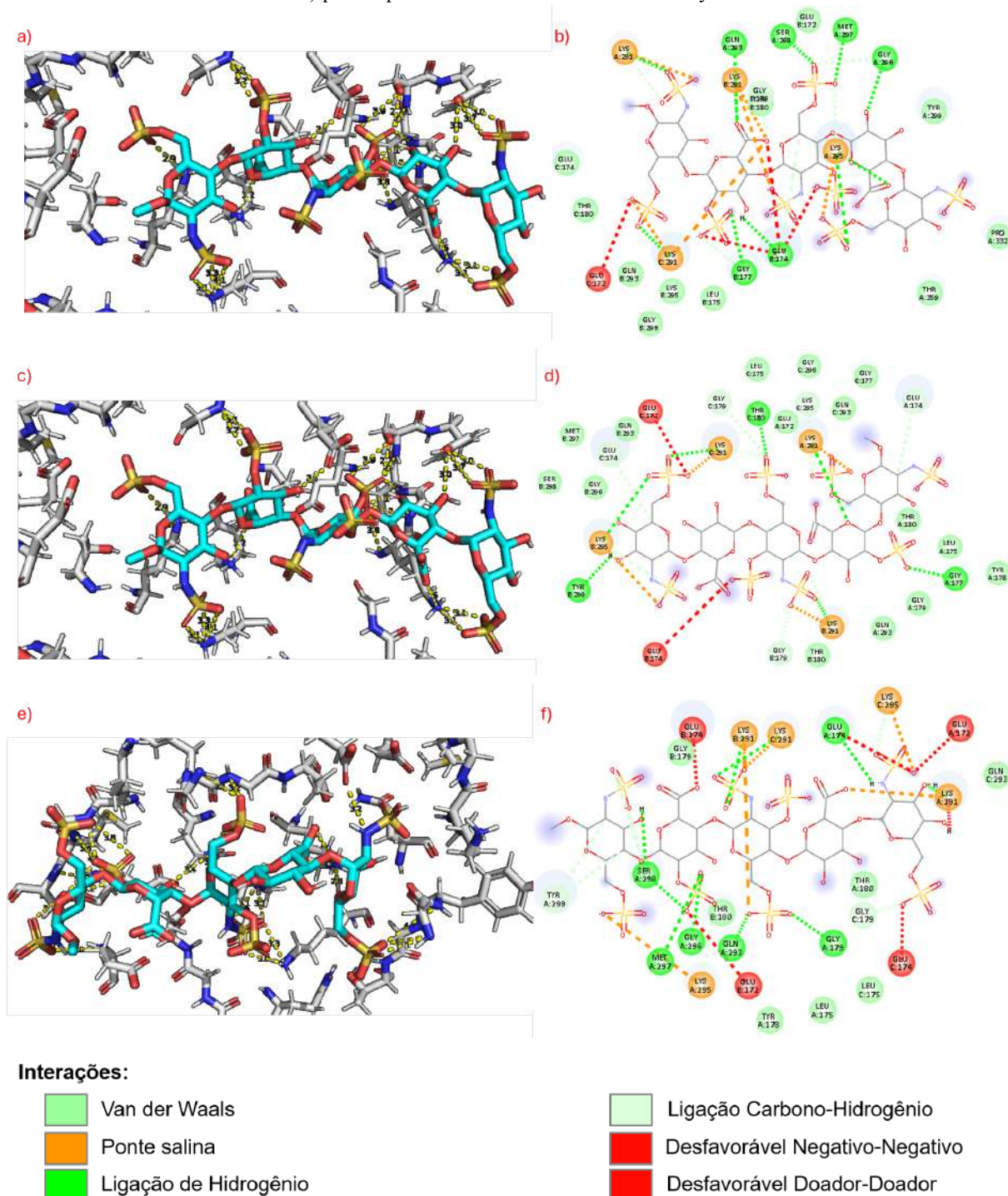


Fonte: Elaborado pelo autor

Essas poses foram selecionadas por critérios físico-químicos e populacionais, entretanto, é possível reparar que elas retratam muito bem a simetria do sistema, pois estão orientadas seguindo cada uma das cadeias monoméricas que compõem a estrutura trimétrica.

Com um entendimento de como essas poses estão localizadas no espaço, foi realizada uma etapa de caracterização das interações realizadas por essas poses utilizando os resultados do PLIP que foram convertidos em uma visualização 3D pelo Pymol e uma visualização 2D pelo Discovery Studio (Figura 23).

Figura 23: Visualização das interações das poses representativas de cada sub-cluster de diferentes formas, a) pose representativa sub-cluster 1 Pymol; b) pose representativa sub-cluster 1 Discovery; c) pose representativa sub-cluster 2 Pymol; d) pose representativa sub-cluster 2 Discovery; e) pose representativa sub-cluster 3 Pymol; f) pose representativa sub-cluster 3 Discovery



Fonte: Elaborado pelo autor

As três poses representativas selecionadas convergem para uma região de interações que valida computacionalmente o *pocket* DP-C16-T como sítio preferencial de ligação do

fondaparinux na superfície trimétrica da proteína E do DENV-2. Os resíduos Glu172, Glu174, Gly177, Thr180, Lys291 e Lys295, interagem nas três poses sem exceção, constituindo uma região farmacofórica necessária para o reconhecimento do sítio por um mimetizador de GAG. Esse núcleo de poses selecionadas por critérios independentes, *score* de *docking* e representatividade populacional dos sub-clusters, possui uma correspondência direta com os resíduos identificados por *solvent mapping* (Leite et al. 2026) representam uma validação cruzada entre metodologias completamente distintas.

Os resíduos de maior destaque dentre os que são comuns às três poses são: Lys 291, Lys 295, Glu 172 e Thr 180. As lisinas 291 e 295 emergem como os principais mediadores de interação eletrostática com os grupos sulfatos do fondaparinux, devido à natureza eletrostática e de ligação de hidrogênio de suas interações. O glutamato 172 é o resíduo que realiza o maior número de interações com o fondaparinux, demonstrando que está em uma região ótima de interações. Já a treonina 180 contribui com contatos anfipáticos que estabilizam os anéis de açúcar do pentassacarídeo no *pocket*.

Cada pose revela extensões específicas do modo de ligação que enriquecem o mapa farmacofórico do sítio. A coexistência desses três modos de ligação geometricamente distintos dentro do mesmo sítio reflete a natureza flexível do fondaparinux como pentassacarídeo, capaz de adotar múltiplas orientações complementares ao *pocket* trimétrico, todas biologicamente plausíveis. A escolha dessas 3 poses mostra que todas são bons representantes para esse sistema biológico e podem servir de base em estudos futuros de dinâmica molecular que caracterizem melhor a biologia do sistema e no desenvolvimento de uma terapia antiviral.

6. CONCLUSÃO

Os resultados obtidos neste trabalho estabelecem uma base estrutural consistente para a investigação do fondaparinux como modelo estrutural de um inibidor de entrada do DENV-2, mas demonstram uma necessidade de estudos mais aprofundados.

A variação de score entre as 16 conformações, demonstra que o *docking* de múltiplas estruturas é fundamental para a caracterização de sistemas biológicos. Uma única conformação teria produzido resultado parcial ou enganoso. A c3 como outlier negativo e c1/c4 com baixa ocupação do cluster 3 estabelecem que conformações mais compactas do *pocket* são menos receptivas ao fondaparinux, enquanto conformações com geometria de abertura intermediária como c16 maximizam os contatos com o ligante.

Das 800 poses geradas, 421 (52.6%) convergiram para o Cluster 3, que espacialmente está no DIII e acima das coordenadas do *pocket* DP-C16-T. O fato de mais da metade das poses se dirigirem para aquele sítio específico é uma evidência computacional de que o *pocket* é genuinamente atrativo para o fondaparinux e para GAGs como heparina e SH. Biologicamente, isso faz sentido, pois o DP-C16-T está localizado na interface entre os três monômeros do trîmero, exatamente na região onde Lys291 e Lys295 estão expostas para o exterior do virion.

Embora 52.6% das poses convirjam para o DP-C16-T, as demais 47.4% se distribuem pelos outros clusters, indicando que múltiplas regiões da superfície protéica são energeticamente compatíveis com o ligante. Essa diversidade de ligação decorre principalmente de dois fatores estruturais interdependentes. O primeiro é a extensão da superfície acessível do trîmero que expõe uma área de contato extensa com muitos sítios carregados positivamente. O segundo fator é a flexibilidade conformacional do fondaparinux: como pentassacarídeo com 32 torsões ativas, ele possui liberdade rotacional suficiente para adaptar sua geometria a diferentes microambientes de superfície. A combinação de um superfície extensa e positivamente carregada com um ligante grande e flexível negativamente carregado configura um sistema de baixa especificidade geométrica, onde a seletividade biológica observada pelo sítio DP-C16-T é determinada primariamente pela densidade e organização espacial dos resíduos de GAG *binding*. Esse resultado é consistente com o comportamento conhecido de GAGs endógenos, que também interagem de forma diversa com

múltiplas regiões de proteínas de superfície viral, sendo sua especificidade funcional determinada por contexto biológico e não por geometria molecular restrita.

A sub-clusterização por RMSD revelou que o fondaparinux não adota um único modo de ligação no pocket trimétrico, mas três possíveis orientações distintas com diferentes perfis de interação. Isso tem implicação direta para o design de inibidores: cada sub-modo representa uma estratégia distinta de otimização e pode ser usado para representar de forma significativa a biologia do sistema em estudos de dinâmica molecular.

As três poses representativas foram selecionadas precisamente por representarem os modos de ligação mais populados e energeticamente favoráveis do *docking*, reunindo características necessárias para servir como estruturas de partida em simulações de dinâmica molecular. A progressão metodológica de *docking* estático para *docking* de múltiplas estruturas e finalmente para dinâmica molecular do complexo representa o fluxo de trabalho natural do design estrutural de antivirais, e as poses aqui identificadas fornecem os pontos de entrada necessários para que essa investigação prossiga sobre fundamentos computacionais sólidos e biologicamente justificados.

7. REFERÊNCIAS

AGARWAL, Ankita; GANVIR, Ruchi; KALE, Dipesh; CHAURASIA, Deepti; KAPOOR, Garima. Continued dominance of dengue virus serotype 2 during the recent central India outbreaks (2019–2021) with evidence of genetic divergence. **Pathogens and Global Health**, v. 118, n. 2, p. 109–119, 2024. DOI: <https://doi.org/10.1080/20477724.2023.2246712>.

AIN, Q. U.; ALEKSANDROVA, A.; ROESSLER, F. D.; BALLESTER, P. J. Machine-learning scoring functions to improve structure-based binding affinity prediction and virtual screening. **WIREs Computational Molecular Science**, v. 5, n. 6, p. 405–424, 2015. DOI: <https://doi.org/10.1002/wcms.1225>.

BioArt. Disponível em: <<https://bioart.niaid.nih.gov/bioart/604>>. Acesso em: 2 jun. 2026.

BRADY, Oliver J.; HAY, Simon I. The global expansion of dengue: how *Aedes aegypti* mosquitoes enabled the first pandemic arbovirus. **Annual Review of Entomology**, v. 65, n. 1, p. 191–208, 2020. DOI: <https://doi.org/10.1146/annurev-ento-011019-024918>.

BRASIL. Ministério da Saúde. **Painel de monitoramento das arboviroses**. Brasília, DF: Ministério da Saúde, 2026. Disponível em: <https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/a/aedes-aegypti/monitoramento-das-arboviroses>. Acesso em: 8 maio 2026.

CHEN, Li-Yu; KHAN, Nida; LINDENBAUER, Annerose; NGUYEN, Thi-Huong. When will fondaparinux induce thrombocytopenia? **Bioconjugate Chemistry**, v. 33, n. 8, p. 1574–1583, 2022. DOI: <https://doi.org/10.1021/acs.bioconjchem.2c00316>.

CHIN, J. F. L.; CHU, J. J. H.; NG, M. L. The envelope glycoprotein domain III of dengue virus serotypes 1 and 2 inhibit virus entry. **Microbes and Infection**, v. 9, n. 1, p. 1–6, 2007. DOI: <https://doi.org/10.1016/j.micinf.2006.09.009>.

CRAMER, C. J. **Essentials of computational chemistry: theories and models**. 2. ed. Chichester: Wiley, 2004.

CRUZ-OLIVEIRA, Christine; FREIRE, João Miguel; CONCEIÇÃO, Thaís M.; HIGA, Luiza M.; CASTANHO, Miguel A. R. B.; DA POIAN, Andrea T. Receptors and routes of dengue virus entry into the host cells. **FEMS Microbiology Reviews**, v. 39, n. 2, p. 155–170, 2015. DOI: <https://doi.org/10.1093/femsre/fuu004>.

DASSAULT SYSTÈMES. **BIOVIA Discovery Studio Modeling Environment**, Release 2026. San Diego: Dassault Systèmes, 2026.

DE ÁVILA, M. B.; PEREIRA, F. S. S.; DARDENNE, L. E. Empirical scoring functions for structure-based virtual screening: applications, critical aspects, and challenges. **Frontiers in Pharmacology**, v. 9, p. 1089, 2018. DOI: <https://doi.org/10.3389/fphar.2018.01089>.

DIAMOND, M. S.; PIERSON, T. C. Molecular insight into dengue virus pathogenesis and its implications for disease control. **Cell**, v. 162, n. 3, p. 488–492, 2015. DOI: <https://doi.org/10.1016/j.cell.2015.07.005>.

EUROPEAN MEDICINES AGENCY. **Arixtra**. Amsterdam: EMA, [2026]. Disponível em: <https://www.ema.europa.eu/en/medicines/human/EPAR/arixtra>. Acesso em: 8 maio 2026.

FONG, I. W. Climate change: impact on health and infectious diseases globally. In: FONG, I. W. (org.). **Current trends and concerns in infectious diseases**. Cham: Springer International Publishing, 2020. p. 165–190.

FORLI, Stefano; HUEY, Ruth; PIQUE, Michael E.; SANNER, Michel F.; GOODSSELL, David S.; OLSON, Arthur J. Computational protein–ligand docking and virtual drug screening with the AutoDock suite. **Nature Protocols**, v. 11, n. 5, p. 905–919, 2016. DOI: <https://doi.org/10.1038/nprot.2016.051>.

GUBLER, D. J. The economic burden of dengue. **The American Journal of Tropical Medicine and Hygiene**, v. 86, n. 5, p. 743–744, 2012. DOI: <https://doi.org/10.4269/ajtmh.2012.12-0157>.

GUERRERO, D.; CANTAERT, T.; MISSÉ, D. Aedes mosquito salivary components and their effect on the immune response to arboviruses. **Frontiers in Cellular and Infection Microbiology**, v. 10, 2020. DOI: <https://doi.org/10.3389/fcimb.2020.00407>.

HALLIDAY, D.; RESNICK, R.; WALKER, J. **Fundamentals of physics**. Volume 1. Hoboken: John Wiley & Sons, 2017.

HILGARD, Philip; STOCKERT, Richard. Heparan sulfate proteoglycans initiate dengue virus infection of hepatocytes. **Hepatology**, v. 32, n. 5, p. 1069–1077, 2000. DOI: <https://doi.org/10.1053/jhep.2000.18713>.

HUANG, C.-H.; TSAI, Y.-T.; WANG, S.-F.; WANG, W.-H.; CHEN, Y.-H. Dengue vaccine: an update. **Expert Review of Anti-Infective Therapy**, v. 19, n. 12, p. 1495–1502, 2021.

HUERTA, Vivian; CHINEA, Glay; FLEITAS, Noralvis et al. Characterization of the interaction of domain III of the envelope protein of dengue virus with putative receptors from CHO cells. **Virus Research**, v. 137, n. 2, p. 225–234, 2008. DOI: <https://doi.org/10.1016/j.virusres.2008.07.022>.

JONES, Susan. Periodicity of dengue. **Nature Reviews Microbiology**, v. 2, n. 3, p. 182, 2004. DOI: <https://doi.org/10.1038/nrmicro851>.

KAUSAR, S. et al. A review: mechanism of action of antiviral drugs. **International Journal of Immunopathology and Pharmacology**, v. 35, p. 20587384211002621, 2021.

KLAMT, A.; SCHÜÜRMAN, 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**, n. 5, p. 799–805, 1993.

LEITE, P. **Mapeamento de bolsos susceptíveis à interação de substâncias bioativas na estrutura da proteína E na partícula viral da dengue**. 2025. 44 p. Trabalho acadêmico.

LEITE, Pedro T.T. F.; FERNANDES, Philipe O.; LACERDA, Pedro S.; *et al.* Reading between the Chains: Surface Mapping and Druggable Pockets on the Biological Assemblies of DENV-2's Protein E. **ACS Omega**, v. 11, n. 5, p. 7518–7531, 2026.

LI, Jin; FU, Ailing; ZHANG, Le. An overview of scoring functions used for protein–ligand interactions in molecular docking. **Interdisciplinary Sciences: Computational Life Sciences**, v. 11, n. 2, p. 320–328, 2019. DOI: <https://doi.org/10.1007/s12539-019-00327-w>.

LIU, J.; WANG, R. Classification of current scoring functions. **Journal of Chemical Information and Modeling**, v. 55, n. 3, p. 475–482, 2015.

MALTAROLLO, Vinícius Gonçalves (Org.). **Computer-Aided and Machine Learning-Driven Drug Design: From Theory to Applications**. Cham: Springer Nature Switzerland, 2024. (Computer-Aided Drug Discovery and Design). Disponível em: <<https://link.springer.com/10.1007/978-3-031-76718-0>>. Acesso em: 9 jun. 2026.

MICHILLES, R. **Elbow method e silhouette score: como identificar o número ideal de clusters**. [S.l.]: Asimov Academy, [2026]. Disponível em: <https://hub.asimov.academy/blog/elbow-method-e-silhouette-score-numero-cluster/>. Acesso em: 8 maio 2026.

NANAWARE, Nikita; BANERJEE, Anwesha; BAGCHI, Satarupa Mullick; BAGCHI, Parikshit; MUKHERJEE, Anupam. Dengue virus infection: a tale of viral exploitations and host responses. **Viruses**, v. 13, n. 10, p. 1967, 2021. DOI: <https://doi.org/10.3390/v13101967>.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS. **Dengue ameaça 4 bilhões de pessoas no mundo**. [S.l.]: ONU News, 2024. Disponível em: <https://news.un.org/pt/story/2024/10/1838466>. Acesso em: 8 maio 2026.

PAGADALA, Nataraj S.; SYED, Khajamohiddin; TUSZYNSKI, Jack. Software for molecular docking: a review. **Biophysical Reviews**, v. 9, n. 2, p. 91–102, 2017. DOI: <https://doi.org/10.1007/s12551-016-0247-1>.

PICCIRILLO, Erika; AMARAL, Antonia. Busca virtual de compostos bioativos: conceitos e aplicações. **Química Nova**, publicação prévia em linha, 2018. DOI: <https://doi.org/10.21577/0100-4042.20170210>.

PLIP – Protein-Ligand Interaction Profiler. **PLIP - Welcome**. [S.l.]: TU Dresden, [2026]. Disponível em: <https://plip-tool.biotec.tu-dresden.de/plip-web/plip/index>. Acesso em: 8 maio 2026.

REY, F. A.; HEINZ, F. X.; MANDL, C.; KUNZ, C.; HARRISON, S. C. The envelope glycoprotein from tick-borne encephalitis virus at 2 Å resolution. **Nature**, v. 375, p. 291–298, 1995.

SABE, Victor T.; NTOMBELA, Thandokuhle; JHAMBA, Lindiwe A. et al. Current trends in computer aided drug design and a highlight of drugs discovered via computational techniques: a review. **European Journal of Medicinal Chemistry**, v. 224, p. 113705, 2021. DOI: <https://doi.org/10.1016/j.ejmech.2021.113705>.

SHARP, T. M.; MORRIS, S.; MORRISON, A.; DE LIMA CORVINO, D.; SANTIAGO, G. A.; SHIEH, W.-J.; RICO, E.; KOPP, E.; MUÑOZ-JORDÁN, J. L.; MARTTOS, A. et al. Fatal dengue acquired in Florida. **New England Journal of Medicine**, v. 384, n. 23, p. 2257–2259, 2021.

STEWART, J. J. 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.

SULIMOV, V. B.; KUTOV, D. C.; SULIMOV, A. V. Combined docking with classical force field and quantum chemical semiempirical method PM7. **Advances in Bioinformatics**, v. 2017, p. 7167691, 2017. DOI: <https://doi.org/10.1155/2017/7167691>.

SULIMOV, V. B. et al. Supercomputing, docking and quantum mechanics in quest for inhibitors of papain-like protease of SARS-CoV-2. **Lobachevskii Journal of Mathematics**, v. 42, n. 8, p. 2043–2050, 2021. DOI: <https://doi.org/10.1134/S1995080221070222>.

TAMURA, Tomokazu; ZHANG, Jiayu; MADAN, Vrinda et al. Generation and characterization of genetically and antigenically diverse infectious clones of dengue virus serotypes 1–4. **Emerging Microbes & Infections**, v. 11, n. 1, p. 227–239, 2022. DOI: <https://doi.org/10.1080/22221751.2021.2021808>.

TROTT, Oleg; OLSON, Arthur J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. **Journal of Computational Chemistry**, v. 31, n. 2, p. 455–461, 2010. DOI: <https://doi.org/10.1002/jcc.21334>.

URCUQUI-INCHIMA, S. et al. Recent developments in understanding dengue virus replication. **Advances in Virus Research**, p. 1–39, 2010.

VILLALAIN, J. Envelope E protein of dengue virus and phospholipid binding to the late endosomal membrane. **Biochimica et Biophysica Acta (BBA) – Biomembranes**, v. 1864, n. 5, p. 183889, 2022.

VINCENTI-GONZALEZ, M. F.; TAMI, A.; LIZARAZO, E. F.; GRILLET, M. E. ENSO-driven climate variability promotes periodic major outbreaks of dengue in Venezuela. **Scientific Reports**, v. 8, n. 1, p. 5727, 2018. DOI: <https://doi.org/10.1038/s41598-018-24003-z>.

WATTERSON, D.; KOBE, B.; YOUNG, P. R. Residues in domain III of the dengue virus envelope glycoprotein involved in cell-surface glycosaminoglycan binding. **Journal of General Virology**, v. 93, n. 1, p. 72–82, 2012.

WORLD HEALTH ORGANIZATION. **Dengue: global situation, surveillance and progress – 2024 update**. Geneva: WHO, 2024. Disponível em: <https://www.who.int/publications/i/item/who-wer10052-665-678>.

XIE, Qing; SPEAR, John M.; NOBLE, Alex J. et al. The 2.8 Å electron microscopy structure of adeno-associated virus-DJ bound by a heparinoid pentasaccharide. **Molecular Therapy – Methods & Clinical Development**, v. 5, p. 1–12, 2017. DOI: <https://doi.org/10.1016/j.omtm.2017.02.004>.

8. APÊNDICE

Apêndice A: Resultados da Análise do PLIP

Figura A-1: 20 resíduos que mais interagiram com o fondaparinux

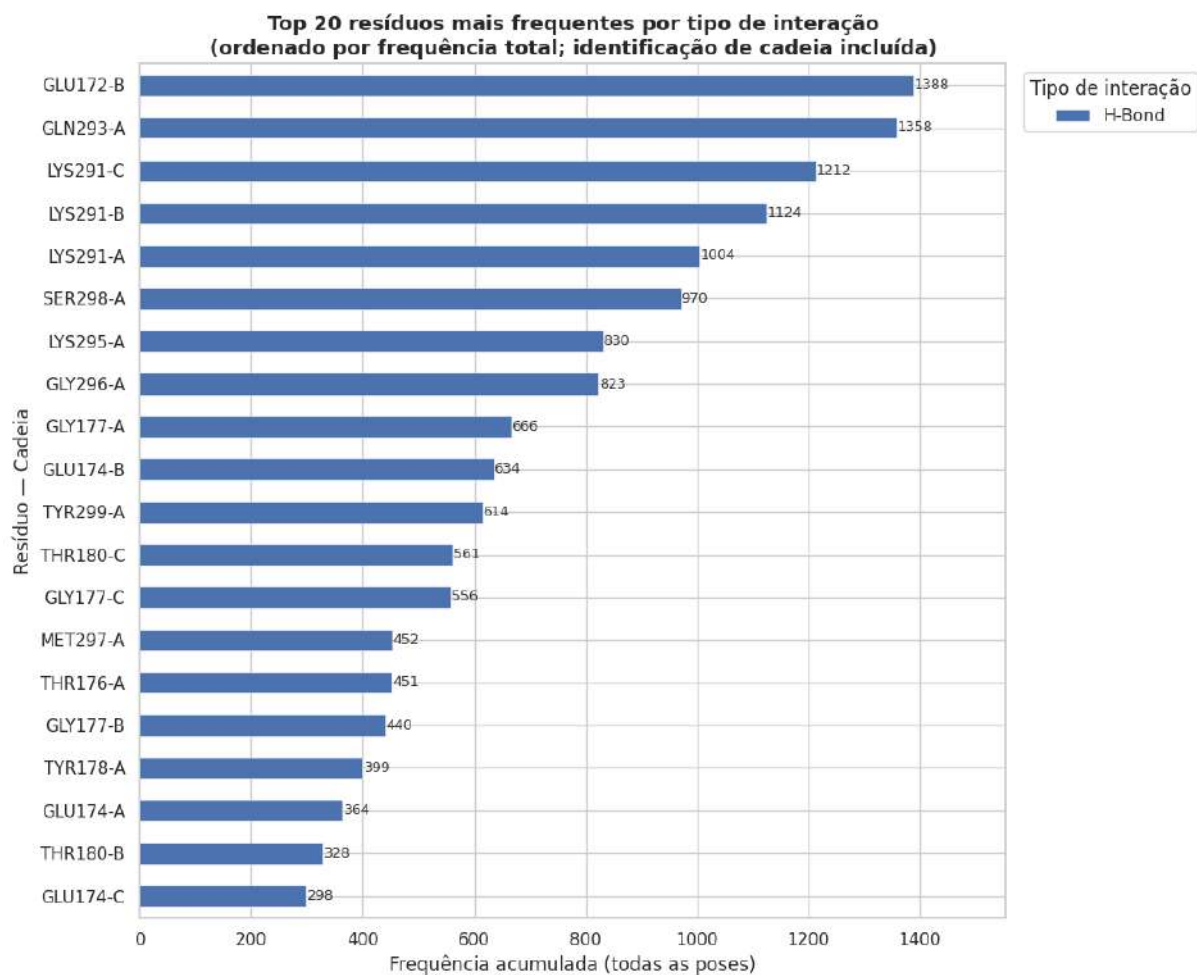
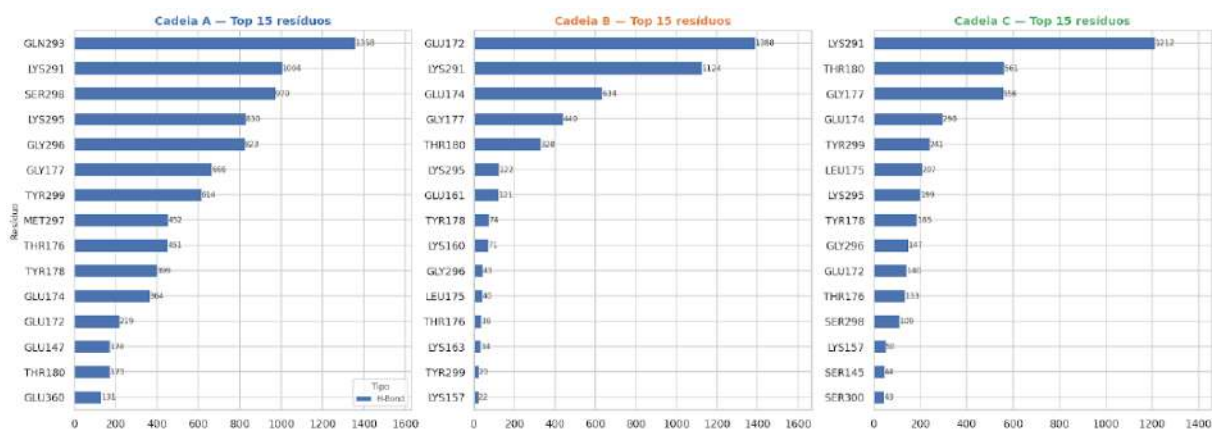


Figura A-2: Top 15 resíduos que mais interagiram com o fondaparinux por cadeia do trímico

Apêndice B: Resultados da análise do HCA

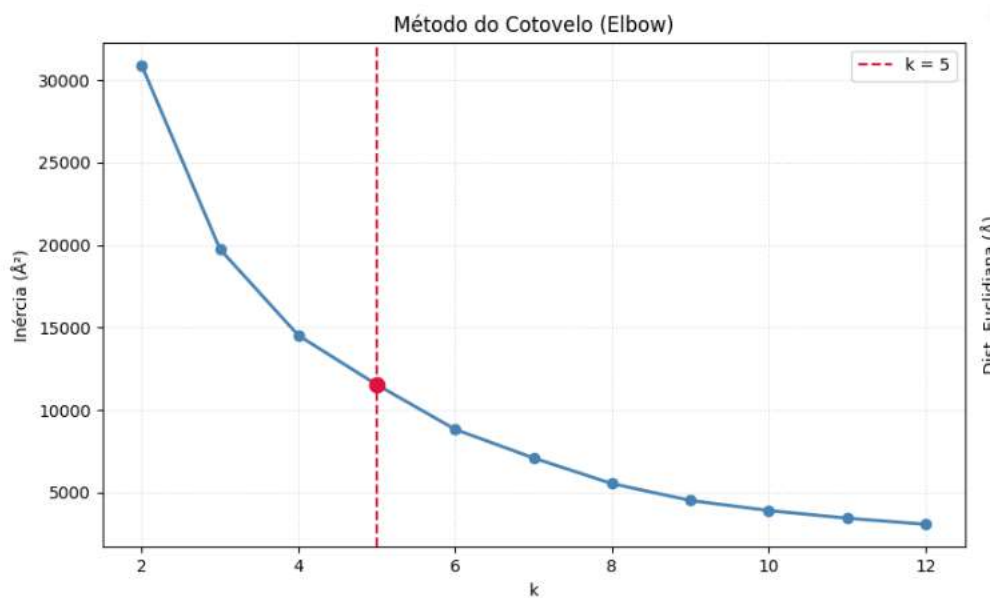
Figura B-1: Método do cotovelo “Elbow Method” da clusterização por centro de massa

Figura B-2: Método do cotovelo “Elbow Method” da clusterização por RMSD

