

UFRRJ
INSTITUTO DE QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA

DISSERTAÇÃO DE MESTRADO

Identificação de Ligantes Alostéricos da Polimerase do Vírus da Dengue Usando
Estratégias de SBDD e FBDD

Sarah Hessing Louzada

2025



UFRRJ
INSTITUTO DE QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA

**IDENTIFICAÇÃO DE LIGANTES ALOSTÉRICOS DA POLIMERASE
DO VÍRUS DA DENGUE USANDO ESTRATÉGIAS DE SBDD E FBDD**

SARAH HESSING LOUZADA

Sob a Orientação do Professor:

Carlos Mauricio R. Sant'Anna

Dissertação submetida como requisito
parcial para obtenção do grau de Mestra
em Química, no Programa de Pós
Graduação em Química, Área de
Concentração Química.

Seropédica, RJ

Setembro de 2025

Universidade Federal Rural do Rio de Janeiro
Biblioteca Central / Seção de Processamento Técnico

Ficha catalográfica elaborada
com os dados fornecidos pelo(a) autor(a)

L886i

Louzada, Sarah Hessing, 2001-
Identificação de ligantes alostéricos da polimerase
do vírus da dengue usando estratégias de SBDD e FBDD /
Sarah Hessing Louzada. - Seropédica, 2025.
140 f.: il.

Orientador: Carlos Mauricio Rabello Sant'Anna.
Dissertação(Mestrado). -- Universidade Federal Rural
do Rio de Janeiro, Programa de Pós-Graduação em Química,
2025.

1. Química Medicinal. 2. Modelagem Molecular. I.
Sant'Anna, Carlos Mauricio Rabello, 1965-, orient. II
Universidade Federal Rural do Rio de Janeiro.
Programa de Pós-Graduação em Química III. Título.

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

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível
Superior - Brasil (CAPES) - Finance Code 001.

UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
INSTITUTO DE QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA

SARAH HESSING LOUZADA

Dissertação submetida como requisito parcial para a obtenção do grau de Mestra em Química, no Programa de Pós-Graduação em Química, Área de concentração Química.

DISSERTAÇÃO APROVADA EM 16/09/2025

Carlos Mauricio Rabello de Sant'Anna Dr. UFRRJ
(Orientador)

Magaly Girão Albuquerque Dra. UFRJ

Renata Barbosa Lacerda Dra. UFRRJ



TERMO Nº 798/2025 - PPGQ (12.28.01.00.00.00.60)

(Nº do Protocolo: NÃO PROTOCOLADO)

(Assinado digitalmente em 17/09/2025 10:35)

CARLOS MAURICIO RABELLO DE SANT ANNA

PROFESSOR DO MAGISTERIO SUPERIOR

DQF (11.39.00.25)

Matrícula: ####204#4

(Assinado digitalmente em 17/09/2025 09:30)

RENATA BARBOSA LACERDA

PROFESSOR DO MAGISTERIO SUPERIOR

DCFar (12.28.01.00.00.00.47)

Matrícula: #

(Assinado digitalmente em 17/09/2025 10:19)

MAGALY GIRÃO ALBUQUERQUE

ASSINANTE EXTERNO

CPF: ####.###.847-##

Visualize o documento original em <https://sipac.ufrrj.br/documentos/> informando seu número: 798, ano: 2025, tipo:

TERMO, data de emissão: 17/09/2025 e o código de verificação: 1c4d451064

RESUMO

LOUZADA, Sarah Hessing. **IDENTIFICAÇÃO DE LIGANTES ALOSTÉRICOS DA POLIMERASE DO VÍRUS DA DENGUE USANDO ESTRATÉGIAS DE SBDD E FBDD**. 2025. 140p. Dissertação (Mestrado em Química, Química Medicinal e Biológica). Instituto de Química, Universidade Federal Rural do Rio de Janeiro (UFRRJ), Seropédica, RJ. 2025

A dengue é uma arbovirose endêmica em regiões tropicais e subtropicais, cuja incidência tem se intensificado nas últimas décadas, especialmente devido às mudanças climáticas, representando um sério problema de saúde pública, agravado pela ausência de antivirais específicos licenciados. A proteína não estrutural 5 (NS5), uma polimerase de RNA dependente de RNA (RdRp) essencial na replicação viral, é altamente conservada entre sorotipos e sem homólogos em humanos, o que a torna um alvo estratégico para o desenvolvimento de fármacos. Neste trabalho, aplicou-se uma abordagem de triagem virtual baseada em estrutura por docagem, visando à identificação de ligantes do sítio alostérico *N* da NS5. Foi utilizada uma quimioteca inédita, não comercial, composta por 657 compostos derivados de dissertações e teses do Programa de Pós-Graduação em Química da UFRRJ. A triagem foi conduzida com o programa GOLD 2023.2.0, função de pontuação ChemPLP, utilizando a técnica de docagem em conjunto. As pontuações obtidas foram normalizadas por massa molecular e área superficial, a fim de selecionar fragmentos (massa<300 Da) e por \sqrt{N} (número de átomos diferentes de hidrogênios), a fim de reduzir viés relacionado ao tamanho molecular. Foram selecionados 23 compostos, sendo 17 fragmentos e 5 moléculas maiores. O ácido cinâmico, melhor fragmento no critério massa, foi otimizado por técnicas de crescimento de fragmentos, gerando o ligante ácido (*E*)-3-(3-(2-carboxivinil)fenoxi)benzóico como o derivado mais promissor. Os compostos maiores destacaram-se por estabelecer interações relevantes com resíduos-chave do sítio, como Arg729, Thr794, Trp795 e His800, incluindo ligações de hidrogênio e interações cátion-π. As entalpias de interação foram estimadas com o método semiempírico PM7 (MOPAC2016), considerando solvente implícito, revelando valores de até -96 kcal/mol para o ácido (*E*)-3-(3-(2-carboxivinil)fenoxi)benzóico, o que supera o valor dos inibidores de referência. Entre os compostos maiores, dois atendem integralmente à regra dos cinco de Lipinski, sugerindo potencial para administração oral e previsão de toxicidade comparável ou melhor à dos inibidores de referência. Nesse contexto, destaca-se a série PPGQ2000701. A otimização dos fragmentos também levou a melhora do perfil farmacocinético, eliminando a previsão de interação com CYPs. Os resultados indicam que compostos da quimioteca do PPGQ-UFRRJ são bons candidatos a inibidores da NS5 do DENV e estruturas iniciais promissoras para o planejamento de fármacos. Além disso, as técnicas de design de fármacos baseada em

fragmentos (FBDD) se mostraram eficazes em melhorar os perfis preditos de interação e farmacocinético dos ligantes.

Palavras-chave: *Dengue, Polimerase, Triagem Virtual;*

ABSTRACT

LOUZADA, Sarah Hessing. **IDENTIFICATION OF ALLOSTERIC LIGANDS OF DENGUE VIRUS POLYMERASE USING SBDD AND FBDD STRATEGIES.** 2025. 140p. Dissertation (Master's in Chemistry, Medicinal and Biological Chemistry). Institute of Chemistry, Federal Rural University of Rio de Janeiro (UFRRJ), Seropédica, RJ. 2025

Dengue is an arbovirus endemic in tropical and subtropical regions. Its incidence has intensified in recent decades, especially due to climate change. It represents a serious public health problem, exacerbated by the lack of licensed specific antivirals. Nonstructural protein 5 (NS5), an RNA-dependent RNA polymerase (RdRp) essential for viral replication, is highly conserved across serotypes and has no homologues in humans, making it a strategic target for drug development. In this work, we applied a virtual screening approach based on structure docking to identify ligands for the N allosteric site of NS5. We used a novel, non-commercial chemical library of 657 compounds derived from dissertations and theses from the Graduate Program in Chemistry at UFRRJ. Screening was conducted with the GOLD 2023.2.0 program, ChemPLP scoring function, and the ensemble docking technique. The scores obtained were normalized by molecular mass and surface area to select fragments (mass <300 Da) and by \sqrt{N} (number of non-hydrogen atoms) to reduce bias related to molecular size. Twenty-three compounds were selected, 17 fragments and 5 larger molecules. Cinnamic acid, the best fragment in terms of mass, was optimized by fragment growth techniques, generating the ligand (E)-3-(3-(2-carboxyvinyl)phenoxy)benzoic acid as the most promising derivative. The larger compounds stood out for establishing relevant interactions with key residues of the site, such as Arg729, Thr794, Trp795, and His800, including hydrogen bonds and cation- π interactions. Interaction enthalpies were estimated using the PM7 semiempirical method (MOPAC2016), considering implicit solvent, revealing values up to -96 kcal/mol for (E)-3-(3-(2-carboxyvinyl)phenoxy)benzoic acid, exceeding the value of reference inhibitors. Among the larger compounds, two fully comply with Lipinski's rule of five, suggesting potential for oral administration and predicted toxicity comparable to or better than that of reference inhibitors. In this context, the PPGQ2000701 series stands out. Fragment optimization also led to improved pharmacokinetic profiles, eliminating the predicted interaction with CYPs. The results indicate that compounds from the PPGQ-UFRRJ library are good candidates for DENV NS5 inhibitors and promising starting structures for drug development. Furthermore, fragment-based drug design (FBDD) techniques proved effective in improving the predicted interaction and pharmacokinetic profiles of the ligands.

Keywords: Dengue, Polymerase, Virtual Screening;

LISTA DE FIGURAS

Figura 1: Estrutura do RNA viral de ortoflavivírus. O genoma completo inclui uma única frente de leitura aberta, que codifica dez proteínas virais. As extremidades 5' e 3' contam com regiões não codificantes (UTR). Fonte: Hodge et al, 2019. 21

Figura 2: Proteínas sintetizadas pela tradução do material genético dos ortoflavivírus. Fonte: Roney; Mohd Aluwi, 2024. 21

Figura 3: A estrutura do cap do RNA: o cap consiste em uma 7-metil-guanosina (azul) ligada ao nucleosídeo 5' da cadeia do RNAm por uma subunidade trifosfato 5'-5' (rosa). O grupo metila da guanosina na posição N-7 está circulado em verde. Os grupos metila 2'-O-metil no primeiro e segundo resíduo de nucleotídeo, formando as estruturas do CAP1 e CAP2, respectivamente, estão circuladas em laranja. Fonte: FERRON et al., 2012. 23

Figura 4: Estrutura da NS5 (DENV 3) em representação cartoon. Domínio MTase em amarelo. Substrato GTP e co-fator SAH representados em bastões. Íons metálicos representados por esferas. Subdomínios da RdRp representados: Dedos em verde, polegar em rosa e palma em azul. À direita, são destacadas por quadrados tracejados as áreas de interface entre os domínios. Fonte: ZHAO et al., 2015. 24

Figura 5: Inibidores da NS5 de DENV 1, DENV 2, DENV 3 e DENV 4 em ordem crescente de potência. Os valores de IC_{50} foram medidos por ensaio FAPA (ensaio de polímerase acoplado à fosfatase alcalina baseado em fluorescência) de iniciação "de novo" para a NS5. A numeração apresentada corresponde à do artigo original. Fonte: Lim et al., 2016. 30

Figura 6: Compostos com ação antiviral em célula propostos como inibidores da NS5 de DENV 1, DENV 2, DENV 3 e DENV 4 em ordem de potência. Os valores de EC_{50} foram medidos por ensaio de placa contra infecção viral por DENV (GenBank: EU081177.1) após 47h da infecção. A numeração apresentada corresponde à do artigo original. Fonte: Felicetti et al., 2023. 32

Figura 7: Antivirais contra DENV-2 que superam o composto 27 ($IC_{50} = 19,67 \mu M$ ($\pm 1,12 \mu M$)). Os valores foram medidos para DENV2-NGC por infecção de cultura de células. As culturas foram mantidas após a infecção a 37°C por 3 dias e os resultados foram coletados por leitura luminescente. A numeração apresentada corresponde à do artigo original. Fonte: Zong et al. 2025. 34

Figura 8: Os dois antivirais mais potentes contra DENV-2 apresentados por Wei e colaboradores. Os valores foram medidos para DENV2-NGC por infecção de cultura de células. As culturas foram mantidas após a infecção a 37°C por 3 dias e os resultados foram coletados

por leitura luminescente. A numeração apresentada corresponde à do artigo original. Fonte: Wei et al. 2025.	35
Figura 9: Estruturas de quatro medicamentos aprovados oriundos de FBDD. O PLX3397 (pexidartinibe) foi desenvolvido a partir do 7-azaindol (Zhang et al., 2013). O vemurafenibe foi construído a partir do mesmo fragmento (Tsai et al., 2008). As estruturas do venetoclax (um inibidor de Bcl2) (Souers et al., 2013) e do erdafitinibe e de seus respectivos fragmentos de origem (Murray C.W. et al., 2019) também são apresentadas. Fonte: Li, 2020	47
Figura 10: Estruturas dos ligantes selecionados, seus nomes e códigos na quimioteca do PPGQ-UFRJ.....	60
Figura 11: Composto 27 no sítio alostérico. (Estrutura 5K5M - DENV2). Ligantes com carbonos em rosa claro, resíduos de aminoácido com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6).....	65
Figura 12: Composto 29 no sítio alostérico. (Estrutura 5I3Q - DENV3). Ligantes com carbonos em rosa claro, resíduos de aminoácido com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6).....	66
Figura 13: Sobreposição do ácido cinâmico e inibidor 29 na estrutura 5I3Q otimizada com o método PM7. Inibidor com carbonos em rosa claro, ácido cinâmico com carbonos em verde, resíduos de aminoácido com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6).....	68
Figura 14: Previsão de pose do ácido <i>p</i> -cumárico, com carbonos em verde, resíduos de aminoácido com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6).....	69
Figura 15: Previsão de pose do ácido ferúlico, com carbonos em verde, resíduos de aminoácido com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6).....	70
Figura 16: Previsão de pose do ácido 2-metoxicinâmico, com carbonos em verde, resíduos de aminoácido com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6).....	71

Figura 17: Previsão de pose do ácido <i>o</i> -cumárico, com carbonos em verde, resíduos de aminoácido com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6).....	72
Figura 18: Previsão de pose do 3-octanol, com carbonos em verde, resíduos de aminoácido com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6).....	73
Figura 19: Previsão de pose do (<i>Z</i>)-ocimeno, com carbonos em verde, resíduos de aminoácido com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6).....	74
Figura 20: Previsão de pose do ácido <i>p</i> -nitrocinâmico, com carbonos em verde, resíduos de aminoácido com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6).....	75
Figura 21: Previsão de pose do ácido <i>p</i> -flúor-cinâmico, com carbonos em verde, resíduos de aminoácido com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6).....	76
Figura 22: Previsão de pose do ácido cafeico, com carbonos em verde, resíduos de aminoácidos com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6).....	77
Figura 23: Previsão de pose da drupanina, com carbonos em verde, resíduos de aminoácidos com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6).....	78
Figura 24: Previsão de pose do ácido 4-metoxicinâmico com carbonos em verde, resíduos de aminoácidos com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6).....	80
Figura 25: Previsão de pose do ácido <i>p</i> -metoxibenzoíco com carbonos em verde, resíduos de aminoácidos com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6).....	81
Figura 26: Previsão de pose do mirceno com carbonos em verde, resíduos de aminoácidos com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6).....	82

Figura 27: Previsão de pose do *p*-cimeno com carbonos em verde, resíduos de aminoácidos com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6).....83

Figura 28: Previsão de pose do ligante (*1E,4Z,6E*)-1-(4-((etiniloxi)metil)-3-metoxifenil)-5-hidroxi-7-(3-metoxi-4-(prop-2-in-1-iloxi)fenil)hepta-1,4,6-trien-3-ona com carbonos em verde, resíduos de aminoácidos com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6).....84

Figura 29: Previsão de pose do ligante [1,7- bis -(3, 4- dimetoxifenil)] -heptano - 3,5-diona com carbonos em verde, resíduos de aminoácidos com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6).....85

Figura 30: Previsão de pose do ligante (*1E,4Z,6E*)-5-hidroxi-1,7-bis(4-hidroxifenil)hepta-1,4,6-trien-3-ona com carbonos em verde, resíduos de aminoácidos com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6).....86

Figura 31: Previsão de pose do ligante (*E,E*)-N¹,N⁴-bis(4-clorobenzilideno)butano-1,4-diamina com carbonos em verde, resíduos de aminoácidos com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6).....87

Figura 32: Previsão de pose do ligante (*E,E*)-N¹,N⁶-dibenzilidenohexano-1,6-diamina com carbonos em verde, resíduos de aminoácidos com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6).....88

Figura 33: Previsão de pose do ligante (*E,E*)-N¹,N⁶-dibenzilidenohexano-1,6-diamina com carbonos em verde, resíduos de aminoácidos com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6).....88

Figura 34: Previsão de pose do ligante 2-((*1E,3E*)-4-(benzo[*d*][1,3]dioxol-5-il)buta-1,3-dien-1-il)-5-(dodeciltio)-1,3,4-oxadiazol com carbonos em verde, resíduos de aminoácidos com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6).....89

Figura 35: Previsão de pose do ligante (*E*)-3-((2*S,3S*)-2-(4-hidroxi-3-metoxifenil)-3-(hidroximetil)-7-metoxi-2,3-di-hidrobenzofuran-5-il)acrilaldeído com carbonos em verde,

resíduos de aminoácidos com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6)	90
.....
Figura 36: Sobreposições de fragmentos selecionados com o ácido cinâmico. Oxigênios em vermelho, nitrogênios em azul e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6) A: Todos os derivados do ácido cinâmico, exceto drupanina, sobrepostos (carbonos em cores variadas). B: Drupanina (carbonos em ciano) e ácido cinâmico (carbonos em amarelo) sobrepostos.	93
.....
Figura 37: Previsão de pose do ligante ácido m-fenil-etyl-cinâmico com carbonos em verde, resíduos de aminoácidos com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6)	96
.....
Figura 38: Estruturas 2D dos ligante propostos na segunda série de otimização do ácido cinâmico	98
.....
Figura 39: Previsão de pose do ligante ácido (E)-3-(3-(2-carboxivinil)fenoxy)benzóico com carbonos em verde, resíduos de aminoácidos com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6)	100
.....

LISTA DE ESQUEMAS

Esquema 1: Resumo das etapas de planejamento de fármacos, destacando onde o processo de FBDD pode ser inserido. Adaptado de Li, 2020.....46

Esquema 2: Fluxograma geral da metodologia utilizada.....55

LISTA DE QUADROS

Quadro 1: Limitações dos compostos 27 e 29 (Fontes: Lim et al., 2018; Lim et al., 2019)	31
.....
Quadro 2: Códigos PDB das estruturas analisadas e sua descrição.....	56
Quadro 3: Resíduos diferentes entre as sequências primárias de DENV-2 (5K5M) e DENV-3 (5I3Q)	66

LISTA DE TABELAS

Tabela 1: RMSD da pose melhor pontuada na redocagem dos inibidores cocristalizados nas estruturas selecionadas para o sítio N ¹	56
Tabela 2: Dados dos ligantes docados selecionados. Colocações de 1 a 10 foram ressaltadas em negrito.....	59
Tabela 3: Entalpias de interação dos ligantes selecionados pela triagem virtual.....	64
Tabela 4: Análise dos ligantes grandes pelas regras de Lipinski	91
Tabela 5: Previsão de inibição das CYPs pelos ligantes selecionados. Gerado pelo servidor Swiss-ADME.....	92
Tabela 6: Pontuações de docagem e entalpias de interação para novos derivados alquilados do ácido cinâmico.	95
Tabela 7: Pontuações de docagem e entalpias de interação para os novos derivados do ácido m-fenil-etil-cinâmico.	99
Tabela 8: Análise dos derivados selecionados pela regra dos cinco de Lipinski	101
Tabela 9: Previsão de inibição das CYPs pelos ligantes selecionados, gerado pelo servidor Swiss-ADME.....	101

SUMÁRIO

1.	Introdução e Justificativa.....	17
2.	Revisão da Literatura.....	20
2.1.	O gênero ortoflavivírus	20
2.2.	A replicação viral e a proteína não-estrutural 5 (NS5)	22
2.3.	Motivos da polimerase dos ortoflavivírus.....	27
2.4.	Inibidores da RdRp do vírus da dengue	29
2.5.	Técnicas de descoberta de fármacos baseadas em estrutura	36
2.5.1.	Docagem molecular.....	39
2.5.2.	Métodos quânticos semi-empíricos	43
2.5.3.	Fragment based drug design.....	46
3.	Objetivos.....	49
4.	Metodologia.....	50
4.1.	Obtenção e modelagem dos ligantes.....	50
4.2.	Obtenção e preparação dos receptores	51
4.3.	Escolha da função de pontuação e protocolo de triagem virtual	52
4.4.	Análise das interações dos ligantes no sítio de ligação e FBDD	53
4.5.	Otimização das geometrias previstas e cálculo da entalpia de interação	54
4.6.	Análises de absorção, distribuição, metabolismo, excreção e toxicidade:	55
5.	Resultados e discussão	56
5.1.	Escolha da função de docagem	56
5.2.	Ligantes selecionados por triagem virtual	58
5.3.	Cálculo da entalpia de interação dos complexos proteína-ligante	63
5.4.	Análise das interações de cada ligante selecionado	65
5.4.1.	Fragmentos selecionados da série MPPGQ2004R010	68
5.4.2.	Fragmentos selecionados da série PPGQ2010R01	73

5.4.3. Fragmentos selecionados da série PPGQ2018202	75
5.4.4. Fragmentos selecionados da série PPGQ2016R02	77
5.4.5. Fragmentos selecionados da série MPPGQ2004R03	80
5.4.6. Fragmentos selecionados da série PPGQ1998R01	82
5.4.7. Compostos selecionados da série PPGQ2000701	84
5.4.8. Compostos selecionados da série PPGQ2019201	87
5.4.9. Ligante PPGQ201430113	89
5.4.10. Ligante MPPGQ2003M0103	90
5.5. Análise das propriedades ADMET dos ligantes selecionados	91
5.6. Otimização dos fragmentos por estratégias de FBDD	93
5.6.1. Série 1: Substituintes alquila e alquil-arila	95
5.6.2. Série 2: Adição de grupos polares	98
5.6.3. Avaliação ADMET dos derivados propostos.....	101
6. Conclusões.....	102
7. Referências Bibliográficas.....	103
8. ANEXOS	116
ANEXO 1A: Interações previstas pelo PLIP (Adasme et al., 2021) para o fragmento MPPGQ2004R0105	116
ANEXO 1B: Interações previstas pelo PLIP (Adasme et al., 2021) para o fragmento MPPGQ2004R0106	116
ANEXO 1C: Interações previstas pelo PLIP (Adasme et al., 2021) para o fragmento MPPGQ2004R0108	117
ANEXO 1D: Interações previstas pelo PLIP (Adasme et al., 2021) para o fragmento MPPGQ2004R0109	117
ANEXO 1E: Interações previstas pelo PLIP (Adasme et al., 2021) para o fragmento MPPGQ2004R0110	118
ANEXO 2A: Interações previstas pelo PLIP (Adasme et al., 2021) para o fragmento PPGQ2010R0101.....	118

ANEXO 2B: Interações previstas pelo PLIP (Adasme et al., 2021) para o fragmento PPGQ2010R0102.....	119
ANEXO 3A: Interações previstas pelo PLIP (Adasme et al., 2021) para o fragmento PPGQ201820209	119
ANEXO 3B: Interações previstas pelo PLIP (Adasme et al., 2021) para o fragmento PPGQ201820210	120
ANEXO 4A: Interações previstas pelo PLIP (Adasme et al., 2021) para o fragmento PPGQ2016R0201.....	120
ANEXO 4B: Interações previstas pelo PLIP (Adasme et al., 2021) para o fragmento PPGQ2016R0211.....	121
ANEXO 5A: Interações previstas pelo PLIP (Adasme et al., 2021) para o fragmento MPPGQ2004R0302	121
ANEXO 5B: Interações previstas pelo PLIP (Adasme et al., 2021) para o fragmento MPPGQ2004R0303	122
ANEXO 6A: Interações previstas pelo PLIP (Adasme et al., 2021) para o fragmento PPGQ1998R0105.....	122
ANEXO 6B: Interações previstas pelo PLIP (Adasme et al., 2021) para o fragmento PPGQ1998R0108.....	122
ANEXO 7A: Interações previstas pelo PLIP (Adasme et al., 2021) para o ligante PPGQ200070110	123
ANEXO 7B: Interações previstas pelo PLIP (Adasme et al., 2021) para o ligante PPGQ200070111.....	123
ANEXO 7C: Interações previstas pelo PLIP (Adasme et al., 2021) para o ligante PPGQ200070130	124
ANEXO 8A: Interações previstas pelo PLIP (Adasme et al., 2021) para o ligante PPGQ201920105	124
ANEXO 8B: Interações previstas pelo PLIP (Adasme et al., 2021) para o fragmento PPGQ201920108	125
ANEXO 8C: Interações previstas pelo PLIP (Adasme et al., 2021) para o ligante PPGQ201920109	125

ANEXO 9: Interações previstas pelo PLIP (Adasme et al., 2021) para o ligante PPGQ201430113	126
ANEXO 10: Interações previstas pelo PLIP (Adasme et al., 2021) para o ligante MPPGQ2003M0103	126
ANEXO 11A: Interações previstas pelo PLIP (Adasme et al., 2021) para o ligante 2a	127
ANEXO 11B: Interações previstas pelo PLIP (Adasme et al., 2021) para o ligante 2b	127
ANEXO 12A: Previsão ADMET pelo servidor Swiss-ADME (Daina et al., 2017) para o ligante MPPGQ2003M0103	128
ANEXO 12B: Previsão ADMET pelo servidor Swiss-ADME (Daina et al., 2017) para o ligante PPGQ200070110.....	129
ANEXO 12C: Previsão ADMET pelo servidor Swiss-ADME (Daina et al., 2017) para o ligante PPGQ200070111.....	130
ANEXO 12D: Previsão ADMET pelo servidor Swiss-ADME (Daina et al., 2017) para o ligante PPGQ200070130	131
ANEXO 12E: Previsão ADMET pelo servidor Swiss-ADME (Daina et al., 2017) para o ligante PPGQ201430113.....	132
ANEXO 12F: Previsão ADMET pelo servidor Swiss-ADME (Daina et al., 2017) para o ligante PPGQ201920105	133
ANEXO 12G: Previsão ADMET pelo servidor Swiss-ADME (Daina et al., 2017) para o ligante PPGQ201920109	134
ANEXO 12H: Previsão ADMET pelo servidor Swiss-ADME (Daina et al., 2017) para o inibidor 27.....	135
ANEXO 12I: Previsão ADMET pelo servidor Swiss-ADME (Daina et al., 2017) para o inibidor 29.....	136
ANEXO 13A: Previsão ADMET pelo servidor Swiss-ADME (Daina et al., 2017) para o fragmento 1h	137
ANEXO 13B: Previsão ADMET pelo servidor Swiss-ADME (Daina et al., 2017) para o fragmento 11.....	138

ANEXO 13C: Previsão ADMET pelo servidor Swiss-ADME (Daina et al., 2017) para o fragmento 2A 139

ANEXO 13D: Previsão ADMET pelo servidor Swiss-ADME (Daina et al., 2017) para o fragmento 2b 140

1. Introdução e Justificativa

A dengue é a doença viral transmitida por artrópodes mais relevante no contexto mundial atual. A incidência anual da doença ao redor do mundo cresceu de cerca de 500 mil casos relatados à Organização Mundial da Saúde (OMS) em 2000 a mais de 5 milhões em 2019. Pela grande taxa de subnotificação, é estimado que metade da população mundial está em risco de contrair dengue e que os casos anuais podem chegar a 400 milhões (World Health Organization, 2024). Há 4 sorotipos da dengue que, entre si, compartilham aproximadamente 65% do genoma e causam síndromes quase idênticas (Guzman et al., 2010 *apud* Oliveira; Neto, 2024).

A maior prevalência de casos é em regiões tropicais e subtropicais, sendo a doença endêmica na maior parte dos países da América Latina, incluindo o Brasil (Prestes-Carneiro et al., 2024). Em 2023, foram registrados 1094 óbitos confirmados por dengue no país, sendo esse o maior número, até então, da série histórica, registrada desde 2000. Ainda considerando a série histórica de 2000 até 2023, o maior número anual de casos prováveis havia ocorrido em 2015, sendo 1.688.688 (Ministério da Saúde, 2024a). Entretanto, nas primeiras 11 semanas do ano de 2024, o recorde histórico foi ultrapassado. No dia 18 de março de 2024, foi atualizado o Painel de Arboviroses do Ministério da Saúde, contabilizando 1.899.206 casos prováveis e confirmados (Ministério da Saúde, 2024b).

Em 23 de julho de 2025, data da atualização mais recente do Painel, constavam, para o ano de 2024, 6.593.268 casos prováveis, 6297 óbitos confirmados por dengue e 378 óbitos em investigação. A letalidade dos casos graves chega em 5,92 a cada 100. Para o ano de 2025, já são contabilizados mais de 1,5 milhão de casos prováveis e quase 1500 mortes. As regiões mais afetadas são Sudeste e Sul, em ordem decrescente. Em 2025, o estado com maior coeficiente de incidência de casos a cada 100 mil habitantes é São Paulo, sendo seguido pelo Paraná (Ministério da Saúde, 2024b).

A urbanização e fatores como o aumento de depósitos de resíduos sólidos urbanos, o desmatamento e o aquecimento global têm correlação positiva com o aumento de casos de dengue e outras arboviroses, como demonstrado para a cidade de São Paulo segundo o trabalho de Prestes-Carneiro e colaboradores (2024).

Um fator também relevante é que, ao observarmos o início da série histórica brasileira de casos de dengue, a região Sul era a menos afetada, condição muito relacionada ao clima mais temperado. Todavia, atualmente, é a segunda região brasileira mais afetada (Prestes-Carneiro et al., 2024). A questão climática é apontada como um fator de influência para o surgimento, ressurgimento e endemicidade de diversas doenças tropicais, especialmente as que contam com

vetores artrópodes (Tidman; Abela-Ridder; De Castañeda, 2021). Todo o apresentado demonstra a necessidade de ações para combater a dengue.

O tratamento atual das pessoas com dengue é inespecífico, com foco em alívio dos sintomas (World Health Organization, 2024). A prevenção da doença é majoritariamente focada no controle vetorial, embora existam, atualmente, vacinas. A vacina Dengvaxia® (Pitisuttithum; Bouckenooghe, 2016) foi aprovada pela Agência Nacional de Vigilância Sanitária (ANVISA) em 2015 para uso em indivíduos previamente expostos ao vírus. A vacina Qdenga® (Patel et al., 2023) foi aprovada pela ANVISA em 2023, para uso geral em indivíduos de 4 a 60 anos. Ambas as vacinas são tetravalentes, ou seja, previnem os 4 sorotipos da dengue. A última foi incluída, em 2024, ao Programa Nacional de Imunização (PNI) (Secretaria de Comunicação Social, 2024).

Apesar das boas perspectivas quanto à eficácia das vacinas, alguns pontos devem ser observados. A região brasileira mais afetada atualmente concentra as maiores densidades de pessoas idosas (com mais de 60 anos), que ainda não estão incluídas no programa de vacinação pela necessidade de mais estudos. Também podemos citar que a vacina é contraindicada para lactantes ou gestantes e em quase todas as regiões brasileiras há predominância de mulheres cis com idades entre 20 e 59 anos (Oliveira; Neto, 2024). Por fim, outra ressalva é a falta de estudos em números estatísticos consideráveis para o sorotipo DENV-4, considerando que os quatro sorotipos da doença são atualmente circulantes no Brasil. O DENV-4 foi detectado em 2025 na forma de vírus selvagem em amostras dos estados de São Paulo e Minas Gerais (Informe Semanal n°17 – Ministério da Saúde, 2025).

No que tange ao tratamento de pessoas infectadas, não há nenhum fármaco licenciado para uso, embora candidatos sejam conhecidos. Um dos expoentes nesse contexto é o mosnodenvir (JNJ-1802), um antiviral proposto pela farmacêutica Janssen e seguro em testes clínicos (Kesteleyn et al., 2024). Apesar do bom desempenho nos testes de fase 1, a fase 2 dos testes clínicos foi descontinuada pela farmacêutica (Raritan, 2024). Há limitações para estudos *in vivo* uma vez que não há modelo animal adequado para mimetizar a doença. Ensaios *in vitro*, com modelos de células e ensaios clássicos da virologia são utilizados. Há diversos alvos possíveis, levando em conta todo o ciclo da doença, uma vez que o vírus infecta o hospedeiro (Altamish et al., 2022).

A otimização e descoberta de novos candidatos a fármacos pode utilizar estratégias de Planejamento de Fármacos Auxiliado por Computadores (CADD, do inglês *Computer Aided Drug Design*), que se dividem em dois grandes grupos: o Planejamento de Fármacos Baseado na Estrutura dos Ligantes (LBDD, do inglês *Ligand Based Drug Design*) e o Planejamento de

Fármacos Baseado na Estrutura do Receptor (SBDD, do inglês *Structure Based Drug Design*) (Barreiro et al., 2015). Como exemplo de antiviral modelado por técnicas de SBDD, entre outras, temos o boceprevir, um antiviral utilizado no tratamento de pessoas com hepatite C (Njoroge et al., 2008); recentemente, temos o exemplo do ensitrelvir, um fármaco para tratamento de pessoas com COVID-19, causada pelo SARS-CoV-2 aprovado para uso no Japão em novembro de 2022. A descoberta do ensitrelvir, o primeiro candidato clínico não peptídico para o tratamento de casos graves da COVID-19, iniciou com uma triagem virtual, seguida de triagem biológica e otimização dos *hits* por técnicas de SBDD (Unoh et al., 2022).

Diante da necessidade premente por novas terapias contra a dengue, sobretudo após a infecção viral, e considerando o potencial das técnicas computacionais na descoberta de fármacos, este trabalho propõe-se a identificar, avaliar e otimizar potenciais inibidores da porção polimerase da proteína NS5 do vírus da dengue. Utilizando técnicas de docagem molecular, cálculos quânticos, combinando estruturas cristalográficas da proteína-alvo (YOKOKAWA et al., 2016) e bancos de dados de compostos previamente isolados ou sintetizados, buscou-se, por meio de triagem virtual, novos candidatos a agentes antivirais capazes de interferir no ciclo de replicação viral.

2. Revisão da Literatura

2.1. O gênero ortoflavivírus

Em 2022, o Comitê Internacional de Taxonomia de Vírus (ICTV, do inglês *International Committee on Taxonomy of Viruses*) aprovou a alteração do nome do gênero Flavivírus para Ortoflavivírus, a fim de eliminar ambiguidades entre o nome da família e do gênero. Os conhecidos Flavivírus “*sensu stricto*” formam o maior gênero dentro da família *Flaviviridae* (International Committee on Taxonomy of Viruses, 2023; Lu; Gong, 2017). Atualmente, de acordo com o Comitê, há 89 espécies na família, distribuídas em quatro gêneros. Desses, 53 espécies são caracterizadas como ortoflavivírus. Há, ainda, espécies relacionadas, mas não definitivamente classificadas (International Committee on Taxonomy of Viruses, 2023).

A maior parte dos ortoflavivírus é transmitido por vetores artrópodes, sejam mosquitos ou carapatos. Alguns dos mais conhecidos são os vírus do Nilo Ocidental (WNV), da febre amarela (YF), da dengue (DENV) e da encefalite japonesa (JEV) (Noble et al., 2016). Outro vírus relevante do gênero, no contexto brasileiro, é o Zika vírus, responsável por uma grave epidemia em 2016, chegando a mais de 200.000 casos (Ministério da Saúde, 2023). Apesar de diferentes possíveis mecanismos de transmissão, a organização genômica dentro do gênero é conservada. O RNA do vírus completo e infeccioso (vírion) dos ortoflavivírus é um RNA de fita simples (ssRNA) de sentido positivo que varia de 9.200 a 11.000 bases nitrogenadas. A extremidade 5' do genoma possui um *cap* tipo I e a extremidade 3' não possui uma cauda poli(A), terminando com o dinucleotídeo conservado CU (International Committee on Taxonomy of Viruses, 2023).

Ambas as extremidades possuem regiões não codificantes - também conhecidas como não traduzidas, UTR (do inglês, *Untranslated Region*) – sendo a UTR 5' do comprimento de aproximadamente 100 nucleotídeos (nt) e a UTR 3' do comprimento aproximado de 500 nt (Hodge et al., 2019). Após o contágio, o material genético do ortoflavivírus (Figura 1) entra no citoplasma das células do hospedeiro, onde, liberado, funciona como um RNA mensageiro (RNAm) endógeno, sendo traduzido pelo retículo endoplasmático rugoso em uma poliproteína (Figura 2). Refere-se ao material genético desse gênero viral como contendo um único quadro de leitura aberta (ORF, do inglês *open reading frame*), já que sua tradução dá origem a uma única poliproteína (Howley et al., 2021)

A tradução pode ocorrer de forma canônica, em que o componente ribossomal 43S lê o códon de início, ou por um mecanismo alternativo, não canônico (Hodge et al., 2019). Em experimentos realizados na presença de inibidores da tradução canônica, a replicação viral

prosseguiu. Há indicativos que mesmo o mecanismo alternativo de tradução passa pelo reconhecimento do terminal 5', ressaltando a importância dessa região (Edgil; Polacek; Harris, 2006).

A região 3'UTR também é relevante por ser a região que se liga a proteínas do hospedeiro que estão envolvidas na tradução, como a proteína de ligação a poli(A) (PABP) e o fator de alongamento 1- α (Hodge et al., 2019). Além disso, diversos experimentos demonstraram que alterações na região terminal 3' afetam a produção de proteínas virais. O mecanismo, entretanto, não parece passar por leitura ribossomal da região (Chiu; Kinney; Dreher, 2005; Holden et al., 2006; Reid et al., 2018).

Por fim, uma conclusão interessante do trabalho de Reid e colaboradores (2018) é que a tradução do RNA viral é lenta quando comparada com a tradução de RNAm endógenos do hospedeiro. Os autores sugerem que a lentidão do processo pode ser uma estratégia evolutiva para escapar da resposta imune do hospedeiro. Caso a tradução fosse mais rápida, a homeostase do retículo endoplasmático seria mais seriamente afetada, desencadeando mecanismos de defesa.

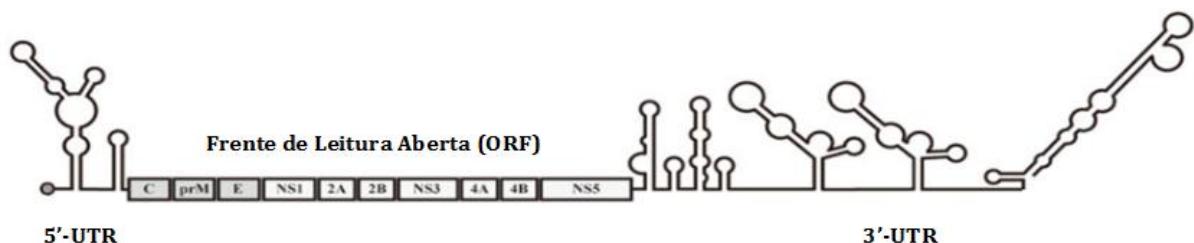


Figura 1: Estrutura do RNA viral de ortoflavivírus. O genoma completo inclui uma única frente de leitura aberta, que codifica dez proteínas virais. As extremidades 5' e 3' contêm regiões não codificantes (UTR). Fonte: Hodge et al, 2019.

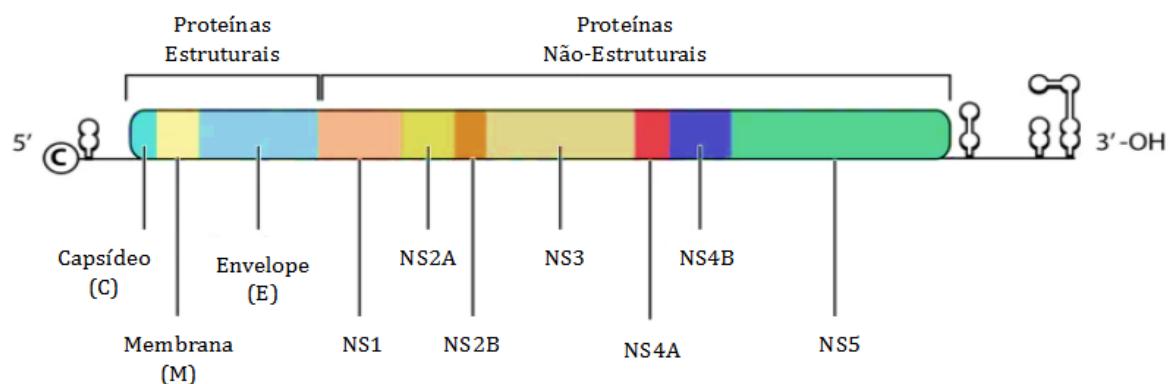


Figura 2: Proteínas sintetizadas pela tradução do material genético dos ortoflavivírus. Fonte: Roney; Mohd Aluwi, 2024.

2.2. A replicação viral e a proteína não-estrutural 5 (NS5)

Ao ser traduzido, o RNA viral origina uma poliproteína ligada ao retículo endoplasmático (RE), que é processada tanto por proteínas virais como do hospedeiro para gerar dez proteínas (Ashour et al., 2009). Três dessas têm funções estruturais: capsídeo (C), membrana (prM) e envelope (E). As demais sete são proteínas não estruturais: NS1, NS2A, NS2B, NS3, NS4A, NS4B e NS5 (Figura 2), que se agrupam formando, junto com co-fatores do hospedeiro, um complexo de replicação viral (Lu; Gong, 2017; Zhao et al., 2015).

A NS5 é a maior e mais conservada proteína entre os ortoflavivírus, consiste em uma metiltransferase (MTase) e uma polimerase de RNA dependente de RNA (RdRp), fundidas por um *linker* flexível de 10 nucleotídeos (Lu; Gong, 2017). O domínio *N*-terminal, correspondente à MTase, pertence à superfamília das MTases dependentes de *S*-adenosil-*L*-metionina (SAM), e tem como função mais marcante o capeamento do RNA. O processo de capeamento é de suma relevância no contexto do reconhecimento do RNA viral e seu processamento como RNAm. Muitas famílias virais utilizam uma via de capeamento muito semelhante à utilizada por células eucarióticas. A estrutura do *cap* (Figura 3) é adicionada à fita nascente 5' em uma série de reações, que passam por hidrólise da ligação 5'-trifosfato, adição do *cap* na forma de guanosina 5'-monofosfato (catálise por uma guanililtransferase, GTPase), metilação na posição N7 da guanina (catálise por MTase) e posteriores reações de metilação nas posições 2'-O das riboses 1 e 2 da fita nascente (também catalisadas por MTase) (Ferron et al., 2012; Issur et al., 2009; Zhao et al., 2015).

A adição do *cap* assegura a estabilidade do material transcrito a frente de uma série de enzimas que degradam RNA (as exonucleases celulares). Além disso, garante que o RNA viral ou RNAm seja lido pelo fator de iniciação da tradução eucariótica eIF4E e corretamente traduzido. Estruturas de RNA sem *cap* são comumente detectadas pelo sistema imune como invasoras. Sendo assim, visando escapar da resposta imunológica, o capeamento deve ser rápido e bem sucedido (Ferron et al., 2012). A NS5, nesse contexto, é uma enzima-chave. O domínio *N*-terminal tem a atividade de MTase, realizando todas as metilações citadas como necessárias no processo do capeamento (Zhao et al., 2015).

Em 2009, Issur e colaboradores forneceram evidências bioquímicas de que a proteína NS5 dos ortoflavivírus possui também atividade de RNA GTPase, uma função essencial para uma das etapas da formação do *cap* do RNA viral, como citado. O estudo demonstrou que o domínio *N*-terminal da NS5 é capaz de utilizar GTP como substrato para formar um intermediário covalente entre a guanosina monofosfato (GMP) e a enzima, por meio de uma ligação fosfoamida.

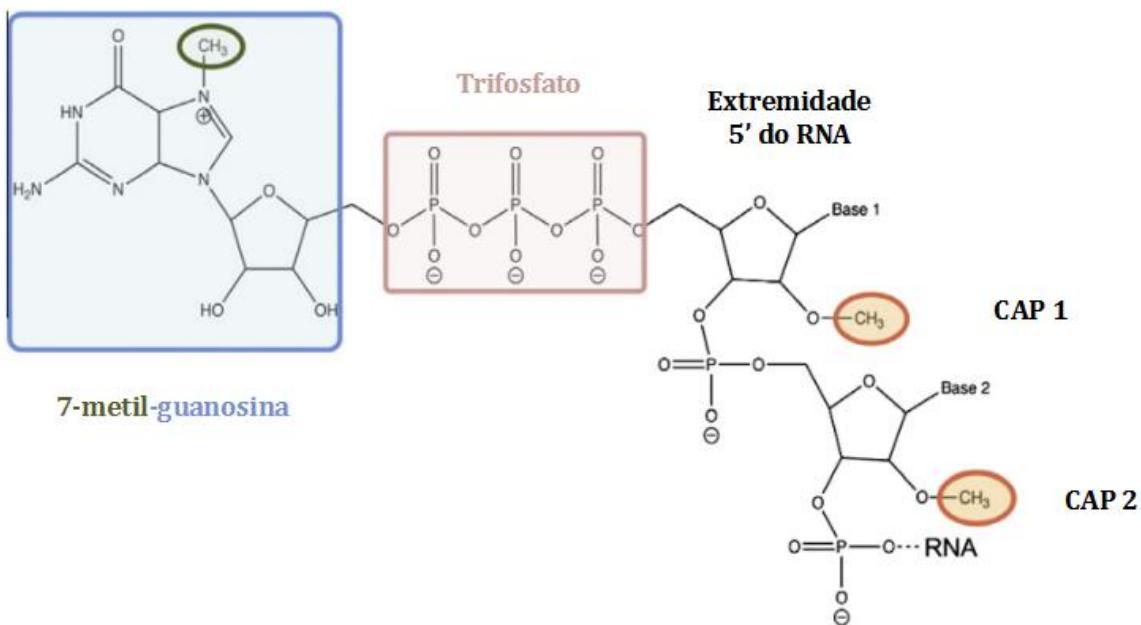


Figura 3: A estrutura do *cap* do RNA: o cap consiste em uma 7-metil-guanosina (azul) ligada ao nucleosídeo 5' da cadeia do RNAm por uma subunidade trifosfato 5'-5' (rosa). O grupo metila da guanosina na posição N-7 está circulado em verde. Os grupos metila 2'-O-metil no primeiro e segundo resíduo de nucleotídeo, formando as estruturas do CAP1 e CAP2, respectivamente, estão circuladas em laranja. Fonte: FERRON et al., 2012.

A região C-terminal, cerca de 2/3 da NS5 em número de nucleotídeos, é um domínio catalítico de RdRp. Os sete motivos de sequência de aminoácidos, conservados em todas RdRps virais, formam a estrutura semelhante a uma mão direita, com subdomínios palma, polegar e dedos (Figura 4). O mecanismo catalítico para a incorporação de nucleotídeos também é comum, utilizando dois íons zinco coordenados a resíduos de aspartato conservados. O sítio ativo tem formato circular, cercado pelos elementos conhecidos como “pontas dos dedos”, que ligam os subdomínios dedos ao “polegar”. O subdomínio da palma é o mais conservado dentro das polimerases, sendo onde o sítio catalítico se aloca. A estrutura cristalográfica da RdRp do vírus da dengue também revelou que os dois íons de zinco contribuem para a estabilidade estrutural (Yap et al., 2007; Zhao et al., 2015).

Dentre os resíduos da polimerase, algumas sequências interagem com outras proteínas virais e humanas. Algumas dessas interações são fatores de virulência relevantes, por exemplo, pode-se citar um dos mecanismos pelo qual a dengue escapa da resposta imune antiviral do hospedeiro, que passa pela antagonização da via de sinalização do interferon (IFN) (Ashour et al., 2009; Zhao et al., 2015).

Em suma, essa via age retardando a replicação viral e estimulando a ativação de células efetoras antivirais. A cascata de sinalização é complexa, se iniciando pelo reconhecimento de padrões virais por receptores semelhantes ao gene I induzível por ácido retinoico (RIG-I) e do

tipo Tool (Ashour et al., 2009). Há ativação do fator de regulação do IFN 3, fator nuclear kappa B (NF- κ B) e proteína ativadora 1 (AP1). O receptor de IFN tipo I (IFNAR) atua como um sítio de ancoragem para a proteína transdutora de sinal e ativadora da transcrição 2 (STAT2), que, ativada, recruta STAT1, heterodimerizando e formando um complexo juntamente com outros componentes, chamado fator de gene estimulado por interferon 3 (ISGF3). A ativação do ISGF3 resulta em níveis de produção aumentados de mais de 100 proteínas, promovendo uma situação antiviral na célula, capaz de inibir a replicação viral (Ashour et al., 2009). A relevância da NS5 nessa via pode ser sucintamente explicada. Durante a infecção por DENV, há diminuição da expressão de STAT2, pois a polimerase viral se liga a esse transdutor e, consequentemente, inibe a sinalização dependente do IFN (Jones et al., 2005).

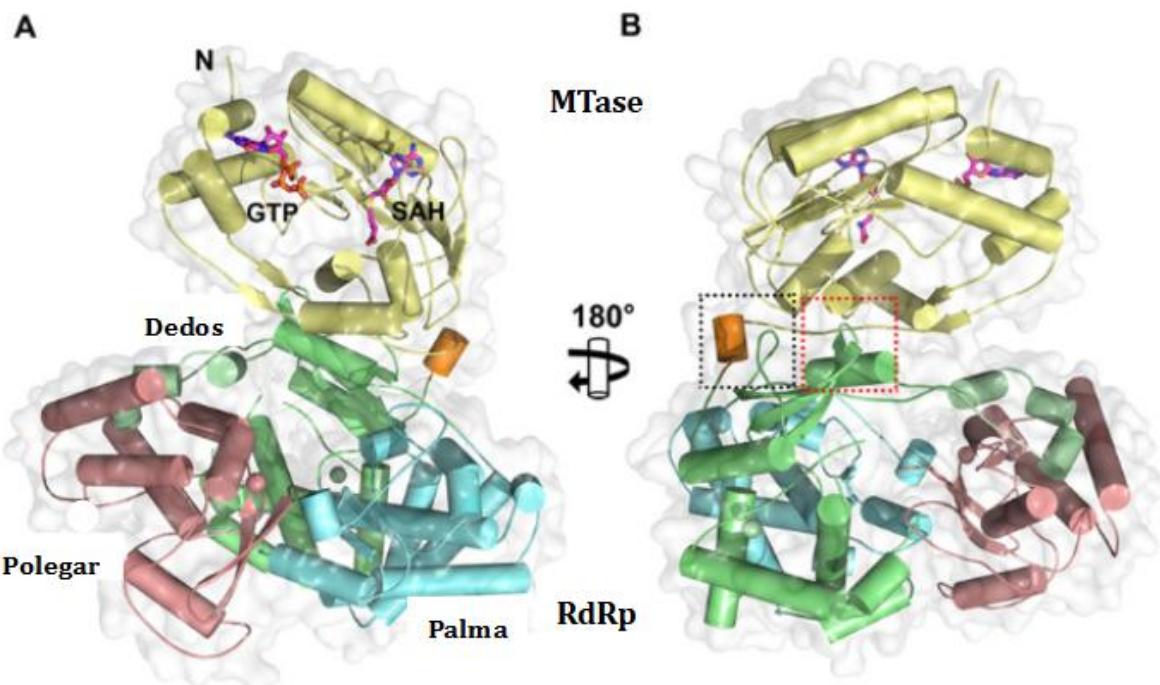


Figura 4: Estrutura da NS5 (DENV 3) em representação cartoon. Domínio MTase em amarelo. Substrato GTP e co-fator SAH representados em bastões. Íons metálicos representados por esferas. Subdomínios da RdRp representados: Dedos em verde, polegar em rosa e palma em azul. À direita, são destacadas por quadrados tracejados as áreas de interface entre os domínios. Fonte: ZHAO et al., 2015.

É interessante que as proteínas NS5 de outros ortoflavivírus, como JEV, vírus Langat (LGT), WNV e o vírus da encefalite transmitida por carrapatos (TBEV) inibem a via do interferon tipo 1 impedindo a ativação do STAT1, enquanto a NS5 do DENV se liga especificamente ao STAT2, promovendo a sua degradação. É também intrigante que a interação da NS5 do DENV2 com o STAT2 é espécie-específica. A mesma proteína não se liga ao STAT de murinos (Hannemann et al., 2013).

Retornando à ideia inicial, como exemplo essencial de interação entre proteínas virais

para a replicação é a interação NS3-NS5. A NS3 conta com uma região helicase, responsável por desenovelar o dsRNA (do inglês, *double-stranded RNA*) e permitir a síntese de novas fitas. Uma sequência peptídica específica da RdRp da NS5 de cerca de 20 resíduos interage com a helicase, permitindo o equilíbrio na produção de fitas tanto positivas como negativas de RNA. A localização das enzimas em células infectadas também é a mesma, demonstrando a necessária interação (Tay et al., 2015). Um resíduo específico da NS5, Lys-330, localizado no sítio catalítico, quando mutado por alanina inibe a interação entre as proteínas, apesar de não afetar a atividade RdRp *in vitro*. Há, portanto, indicativos de que esse trecho da proteína pode apresentar uma possibilidade pertinente para o planejamento de antivirais (Zou et al., 2011).

Vale ressaltar que, diferente do imaginado inicialmente, a proteína NS5 não tem duas regiões catalíticas isoladas. O *linker* entre as duas regiões catalíticas não tem sequência altamente conservada, apenas seu comprimento é preservado entre os ortoflavivírus. A estrutura cristalográfica estudada por Zhao e colaboradores (2015) apresenta uma região *linker* bem organizada, com interface majoritariamente polar entre os domínios. Estudos de mutagênese e replicação viral indicaram que há interações críticas inter-domínios. O avanço no estudo de estruturas cristalográficas e replicação do RNA, demonstra que a proteína inteira tem maior capacidade catalítica (ações de alongamento de fita e iniciação *de novo*) que o domínio RdRp isolado. É sugerido que o domínio MTase influencia dinamicamente na regulação da síntese de RNA, demonstrando que o ciclo viral é dependente de muitas interações e interregulações (Zhao et al., 2015).

Tratando especificamente da atividade catalítica, a RdRp sintetiza, inicialmente, a fita negativa (-) complementar à fita positiva (+) do RNA viral. Essa fita negativa apenas é encontrada associada à positiva. Uma vez associadas, formam a fita dupla de RNA viral (dsRNA). O dsRNA atua como modelo para a síntese de excesso de fitas positivas, que podem ser direcionadas a três principais objetivos: tradução para síntese de novos complexos de replicação, síntese de novas fitas negativas complementares ou empacotamento para invasão de novas células (Hodge et al., 2019; Yap et al., 2007).

Para que essa síntese ocorra corretamente, o primeiro ato da NS5 é diferenciar RNAm do hospedeiro de RNA viral +. No trabalho de Reid e colaboradores (2018) sugere-se que o processo de replicação ocorre em vesículas remodeladas, derivadas do retículo endoplasmático, que se localizam no citoplasma das células infectadas. Outra questão importante é a necessidade de mudança para uma conformação específica, comumente chamada de ciclização do RNA, embora se refira mais à complementaridade 5'-3' do que a um formato circular. Fato é que há necessidade de emparelhamento 5'-3' e modificações nas UTRs prejudicam a síntese de RNA

viral - (Hahn et al., 1987; You; Padmanabhan, 1999). Além da complementaridade das UTRs, diversas sequências foram achadas fundamentalmente conservadas para que a replicação ocorra, entretanto, não serão aprofundadas nesse trabalho.

Embora a replicação ocorra no citoplasma, estruturas de NS5 de ortoflavivirus como JEV, YF e TBEV, além do DENV, foram encontradas no núcleo das células infectadas. O transporte de proteínas grandes, de mais de 45kDa, como é o caso da NS5, é ativo, mediado por proteínas da família das beta carioferinas, conhecidas como importinas e exportinas. Para que a NS5 entre no núcleo celular, uma importina deve se ligar a uma sequência específica dessa proteína. As sequências que permitem essa ligação são chamadas sequências de localização nuclear (NLS, do inglês *nuclear localization sequences*). A importina que reconhece a proteína a entrar no núcleo pode reconhecer dois tipos de NLS. O primeiro, monopartido, é uma sequência única de resíduos de aminoácidos básicos. O segundo, bipartido, são duas sequências de resíduos de aminoácidos básicos ligados por um *linker* de 10 a 12 aminoácidos (Hannemann et al., 2013).

Um NLS bipartido (denominado a/bNLS) foi encontrado na NS5 do DENV-2 em seu domínio polimerase. Experimentos de mutagênese que eliminaram o a/bNLS do DENV-2 demonstraram que a replicação viral diminuiu e os níveis de interleucina-8 (IL-8, uma citocina) aumentaram (Pryor et al., 2007). Foi percebido, então, que a polimerase também contém uma sequência funcional para a exportação nuclear, chamada bNLS, que é reconhecida por uma exportina e carrega a proteína para o citoplasma novamente (Rawlinson et al., 2009). Essas observações sugerem que o transporte núcleo-citoplasma é regulado de maneira a controlar os níveis de IL-8. A sinalização de infecção com variação dos níveis de IL-8 ocorre na contaminação por todos os sorotipos de dengue, entretanto, apenas o sorotipo DENV-2 apresenta o a/bNLS, de forma que os demais sorotipos controlam o escape imunológico por diferentes mecanismos. Provavelmente, a localização da NS5 no núcleo ou no citoplasma não é estritamente necessária para a replicação viral, mas sim um fator de regulação do ciclo celular para sorotipos específicos. Considerando o dito, é necessário avaliar as diferenças e semelhanças entre os diferentes sorotipos no planejamento de fármacos (Hannemann et al., 2013).

A importância da NS5 do vírus da dengue, sua conservação entre sorotipos e a participação na resposta imune a tornam um excelente alvo para o planejamento racional de fármacos antivirais. Casos de inibidores da NS5 de flavivírus comprovados serão relatados nos tópicos seguintes, mas, antes, é necessário falar mais dos motivos conservados entre sorotipos.

2.3. Motivos da polimerase dos ortoflavivírus

Similarmente a todas polimerases de RNA virais, a RdRp da NS5 adota uma estrutura semelhante a uma mão direita formando um círculo com as pontas dos dedos e polegar. Os subdomínios dos dedos podem ser divididos nos subdomínios indicador, médio, anelar e mínimo. A ponta do dedo indicador interage com a ponta do polegar, formando uma estrutura circular, sendo essa característica exclusiva de polimerases de RNA virais (Lu; Gong, 2017).

Considerando apenas as RdRps virais, podemos subclassificá-las de acordo com o mecanismo de iniciação da síntese de RNA. Polimerases dependentes de *primer* apresentam um polegar fino e um canal espaçoso entre o mínimo e o polegar. As polimerases das famílias que seguem o mecanismo de iniciação *de novo* requerem um elemento estrutural de *priming*, como é o exemplo da família Flaviviridae. Esse elemento estabelece uma plataforma que compreende uma inserção no polegar e um trecho C-terminal, tornando o polegar mais volumoso. Essa plataforma estabiliza a fita molde e os nucleotídeos iniciadores. Especificamente nos ortoflavivírus, o elemento de *priming* é chamado “*priming loop*”, que conecta duas hélices do polegar e compreende os resíduos 790 a 812 da NS5 de JEV. A sequência do *priming loop* da NS5 de JEV, comparada ao DENV, apresenta 15 dos 22 resíduos idênticos e em mesma posição (Lu; Gong, 2017). O *priming loop* é essencial no mecanismo da síntese de RNA das enzimas características *de novo* (Butcher et al., 2001).

Todas as polimerases de RNA virais compartilham sete motivos, listados de A até G (Bruenn, 2003). Sua identificação se dá tanto por características estruturais como por conservação de sequência. Os motivos A, B, C e F contém resíduos conservados em todas as famílias virais, enquanto os motivos D, E e G diferem entre as polimerases que seguem mecanismo dependente de *primer* e as que seguem mecanismo *de novo*, mas seus arranjos espaciais em torno do núcleo tornam possível a sua identificação (Te Velthuis, 2014). Os motivos A a E se localizam no domínio palma, o mais conservado. Os motivos F e G estão nos dedos (Lu; Gong, 2017).

O motivo A contém um resíduo de ácido aspártico conservado universalmente em polimerases de ácidos nucleicos que é responsável pela coordenação dos íons zinco necessários para o mecanismo (Beese; Steitz, 1991).

O motivo B inclui uma sequência altamente conservada de resíduos de serinas e glicinas. A serina desse motivo é característica de RdRps, apesar de em algumas cepas resistentes ser modificada para uma treonina (Te Velthuis, 2014). Essa serina sofre mudanças conformacionais relevantes para o mecanismo catalítico de polimerases do tipo *de novo*, como a do DENV e do HCV. A serina sofre uma mudança conformacional de forma a interagir com o ácido aspártico

conservado do motivo A, de forma que o ácido aspártico se move para distante do substrato (nucleotídeo) (Appleby et al., 2015). Esse par Ser-Asp reconhece NTPs que serão direcionados para a síntese, de forma que pode representar um ponto essencial no reconhecimento de inibidores análogos de nucleosídeos. Uma das glicinas do motivo B (que contém poucos resíduos) é conservada tanto em polimerases como em transcriptases reversas (RT). Acredita-se que essa glicina permite flexibilidade do esqueleto polipeptídico, necessária para as movimentações da serina adjacente. O motivo B é relativamente dinâmico e acompanha a transferência de grupo fosforila (Lu; Gong, 2017).

O motivo C contém o segundo ácido aspártico que coordena zinco metálico essencial (Beese; Steitz, 1991). Além disso, contém a sequência de assinatura SGDD de RdRps virais, que interage com a fita produto de RNA durante a síntese.

O motivo D forma uma folha β antiparalela com o motivo A. Durante o fechamento do sítio, esse motivo também se move em direção ao motivo C (Appleby et al., 2015). O motivo D não contém resíduos conservados em todas famílias virais, sendo específico dos ortoflavivírus (Lu; Gong, 2017).

O motivo E tem formato de grampo, se dobrando com o motivo C e interagindo com o esqueleto da fita produto de RNA durante a síntese (Lu; Gong, 2017).

O motivo F também é compartilhado entre RdRps e RTs. Sua principal função é se ligar aos NTPs, formando o “telhado” do túnel de entrada de nucleotídeos (Bruenn, 2003; Butcher et al., 2001). Há 4 resíduos altamente conservados nesse motivo. A Lys459 e o Glu461 interagem com a fita nascente, a Arg474 interage com a região do trifosfato e a Leu ou Ile476 interage com a fita molde (Appleby et al., 2015).

O motivo G é específico de RdRps (Te Velthuis, 2014) e se localiza aproximadamente paralelo à fita molde, porém, não tem interações conservadas com ela. Há dois resíduos Ala ou equivalentes de cadeia pequena conservados que parecem ser relevantes para o processo de translocação de fosfato (Wu et al., 2015).

Com base nesses motivos altamente conservados entre os ortoflavivírus, torna-se possível explorar regiões específicas da polimerase como alvos para inibição seletiva, o que nos leva a discutir os inibidores da NS5.

2.4. Inibidores da RdRp do vírus da dengue

O presente tópico tem por objetivo listar algumas substâncias com ação antiviral contra o vírus da dengue comprovada, com enfoque no domínio polimerase da NS5 viral, relatando resumidamente seus mecanismos de ação. Por ser o alvo escolhido para o presente trabalho, apenas inibidores com ação focada na polimerase foram selecionados.

O trabalho mais relevante até o momento relativo à inibição da polimerase do vírus da dengue é o de Lim e colaboradores (2016; 2018; 2019). Sendo esses pesquisadores do *Novartis Institute of Tropical Diseases*, publicam artigos relacionados à busca de candidatos a fármacos contra a dengue. Em 2016, o grupo publicou um artigo descrevendo uma série de inibidores alostéricos da polimerase do DENV. Até o momento, não havia sido registrada a existência de um sítio alostérico na NS5. O trabalho intitula esse sítio como “*N-pocket*”, que aqui traduziremos como sítio N. A metodologia utilizada foi uma triagem de fragmentos acoplada à cristalografia de raios-X, que permitiu a descoberta de uma forma “fechada” da RdRp, conformação que ocorre antes do início da síntese de uma fita complementar negativa. Nessa conformação, há uma alça de preparação (*priming loop*) que se estende do subdomínio polegar para o túnel de saída do dsRNA, próximo ao sítio ativo com resíduos de aminoácidos conservados (sequência GDD) e que foi identificado como sítio alostérico. A proposta de mecanismo de ação é que a ligação dos compostos nesse sítio impede a mudança conformacional para o estado “aberto”, necessária para o início da síntese de RNA. Os compostos apresentaram valores de IC₅₀ na faixa nanomolar contra todos os quatro sorotipos de dengue em ensaios de cultura de células. A metodologia do trabalho aperfeiçoou seus *hits* a partir de técnicas de SBDD, sendo capaz de partir de valores de IC₅₀ > 700 μM para a faixa nanomolar citada.

Para facilitar a comparação com a literatura original, a numeração dos compostos dos artigos foi mantida. Os compostos planejados foram apresentados (Figura 5) em ordem crescente de potência, sendo o composto **3** o originalmente triado. A partir do composto 15, os valores de IC₅₀ para os 4 sorotipos já eram da faixa do baixo micromolar. Entretanto, em ensaios baseados em células, a baixa permeabilidade impedia a inibição desejada. Então, os compostos continuaram a ser otimizados a fim de aumentar ou manter a potência juntamente com um perfil de permeabilidade celular adequado. Os compostos de melhor performance foram **27** e **29**, e suas estruturas cristalográficas tanto na NS5 inteira como apenas na porção RdRp de DENV-3 estão disponíveis no *Protein Data Bank* (PDB) sob códigos 5I3P, 5I3Q, 5JJS e 5JJR (Lim et al., 2016).

Algumas limitações existem para os compostos **27** e **29**, apesar de seu notável

desempenho. Essas limitações estão justificadas no Quadro 1, sendo válido ressaltar que ainda não foram inteiramente superadas por nenhum outro inibidor de acordo com a literatura encontrada até a escrita do presente documento. Lim, em 2018, declarou que o trabalho foi colocado "em espera" e publicado para que outros pesquisadores tivessem acesso aos avanços até então obtidos (Lim, 2019; Lim et al., 2018).

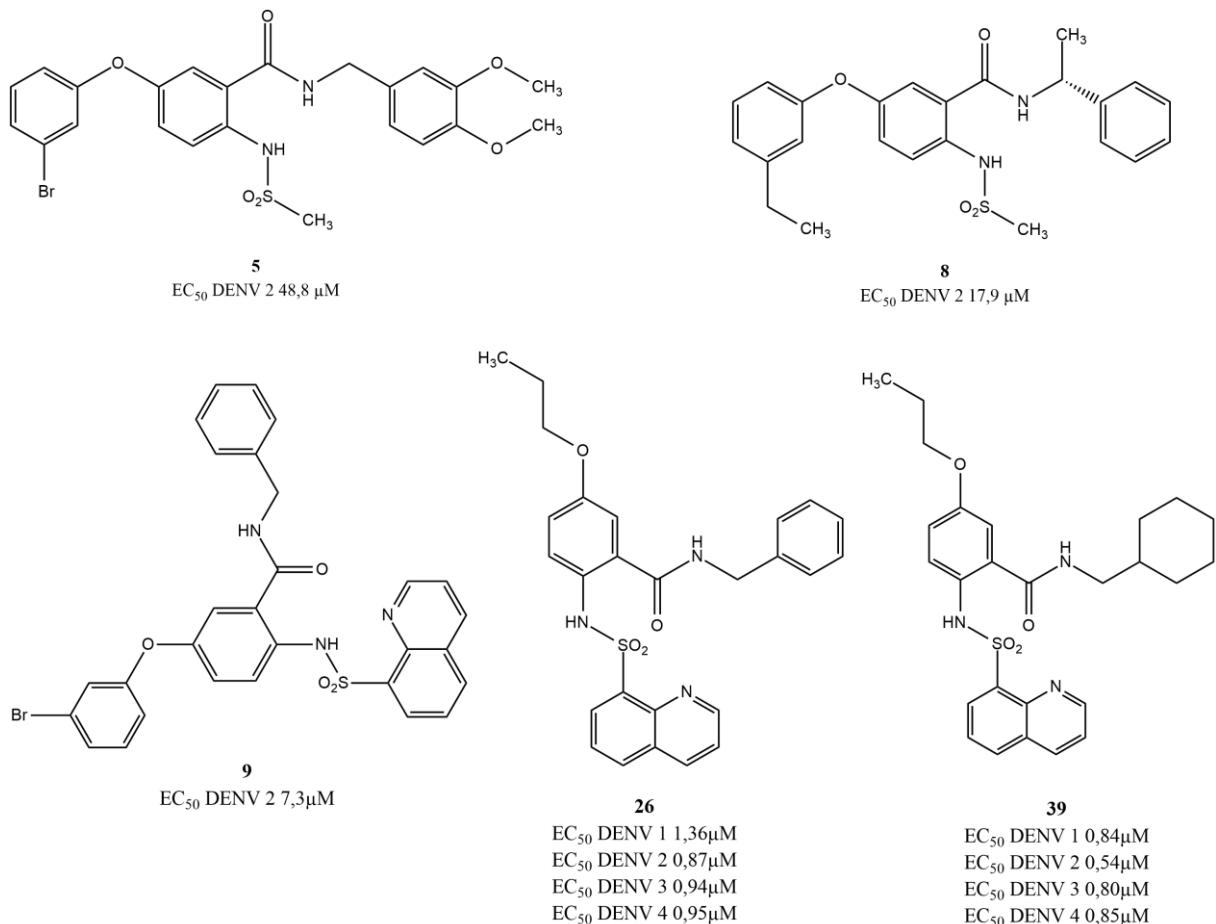


Figura 5: Inibidores da NS5 de DENV 1, DENV 2, DENV 3 e DENV 4 em ordem crescente de potência. Os valores de IC₅₀ foram medidos por ensaio FAPA (ensaio de polimerase acoplado à fosfatase alcalina baseado em fluorescência) de iniciação "de novo" para a NS5. A numeração apresentada corresponde à do artigo original. Fonte: Lim et al., 2016.

Em 2023, Felicetti e colaboradores divulgaram seus trabalhos utilizando derivados do ácido sulfonil-antranílico para a inibição alostérica da polimerase do vírus da dengue, a partir do arcabouço 2,1-benzotiazina 2,2-dióxido, previamente utilizado pelo grupo e capaz de promover inibição enzimática, porém, sem ação em ensaios celulares. Os derivados iniciais mostravam IC₅₀ na faixa micromolar, com perfil de inibição mista, coerente com inibição enzimática alostérica, sendo a inabilidade em ensaios celulares justificada provavelmente por baixa permeabilidade celular, grupo éster lável frente a esterases celulares e propensão à retenção dos compostos nas membranas celulares, pela alta lipofilicidade.

Visando a melhora do perfil farmacocinético de seus compostos, o mesmo grupo propôs modificações estruturais chegando a compostos derivados com melhor perfil metabólico e retenção da capacidade inibitória, seguindo para a síntese apenas após confirmar que a sobreposição dos núcleos base era adequada. O primeiro conjunto de derivados apresentou resultados em ensaios celulares, porém, com alta citotoxicidade e, portanto, baixo índice de seletividade. A maior seletividade foi observada nos compostos **5** e **8** (Figura 6) (Felicetti et al., 2023).

Quadro 1: Limitações dos compostos **27** e **29** (Fontes: Lim et al., 2018; Lim et al., 2019)

Limitação	Descrição
1. Permeabilidade celular deficiente	Compostos iniciais com grupos ácidos fortemente ionizados, dificultando a entrada nas células, mesmo após otimização, a permeabilidade ainda não é ideal.
2. Potência celular ainda insuficiente	Apesar de melhores IC_{50}/EC_{50} , compostos ainda não atingiram a potência submicromolar desejada pelos pesquisadores em ensaios baseados em células.
3. Problemas de bioestabilidade e farmacocinética	Alta depuração <i>in vitro</i> , inviabilizando estudos de eficácia em modelo animal.
4. Alterações conformacionais no alvo durante elongação	Mudanças na conformação da NS5 durante a elongação reduzem a afinidade de ligação dos inibidores. Além disso, não há estrutura cristalográfica disponível da RdRp com RNA, o que dificulta o estudo da forma aberta da enzima.
5. Resistência viral	Surgimento de mutações (L511V e E802D) que reduzem drasticamente a potência dos compostos. Essas mutações foram geradas em ensaios com o composto 27.
6. Falta de ancoragem estrutural comparável a outros flavivírus, como o HCV	No vírus da dengue, o bolso N tem apenas uma "parede" estrutural, o que dificulta o encaixe e a estabilidade dos inibidores. Há menos pontos de ancoragem quando comparado com sítios alostéricos de outros vírus, como o da hepatite C (HCV).

Diante da elevada citotoxicidade observada, os autores buscaram modificações estruturais que mantivessem ou aumentassem atividade inibitória da polimerase e reduzissem os efeitos tóxicos. (Felicetti et al., 2023).

Para isso, investigaram alterações em diversas regiões do composto inicial, aplicando estudos de correlação estrutura-atividade. Foram investigados: o grupo sulfonamida, a adição

de um nitrogênio quinolínico, a variação da extensão da aromaticidade nos substituintes e, por fim, o grupo amida. Do segundo conjunto de ligantes sintetizados, o melhor desempenho foi de **9**, no terceiro conjunto destacou-se **26** e, no quarto e último conjunto, **39**. Pode-se observar como relevantes duas regiões aromáticas, ligados ao anel aromático central por ligações amida e sulfonamida. Por fim, uma pequena cadeia alquila ligada por um grupo éter ao anel aromático central parece ser responsável por um grande ganho de potência (Felicetti et al., 2023).

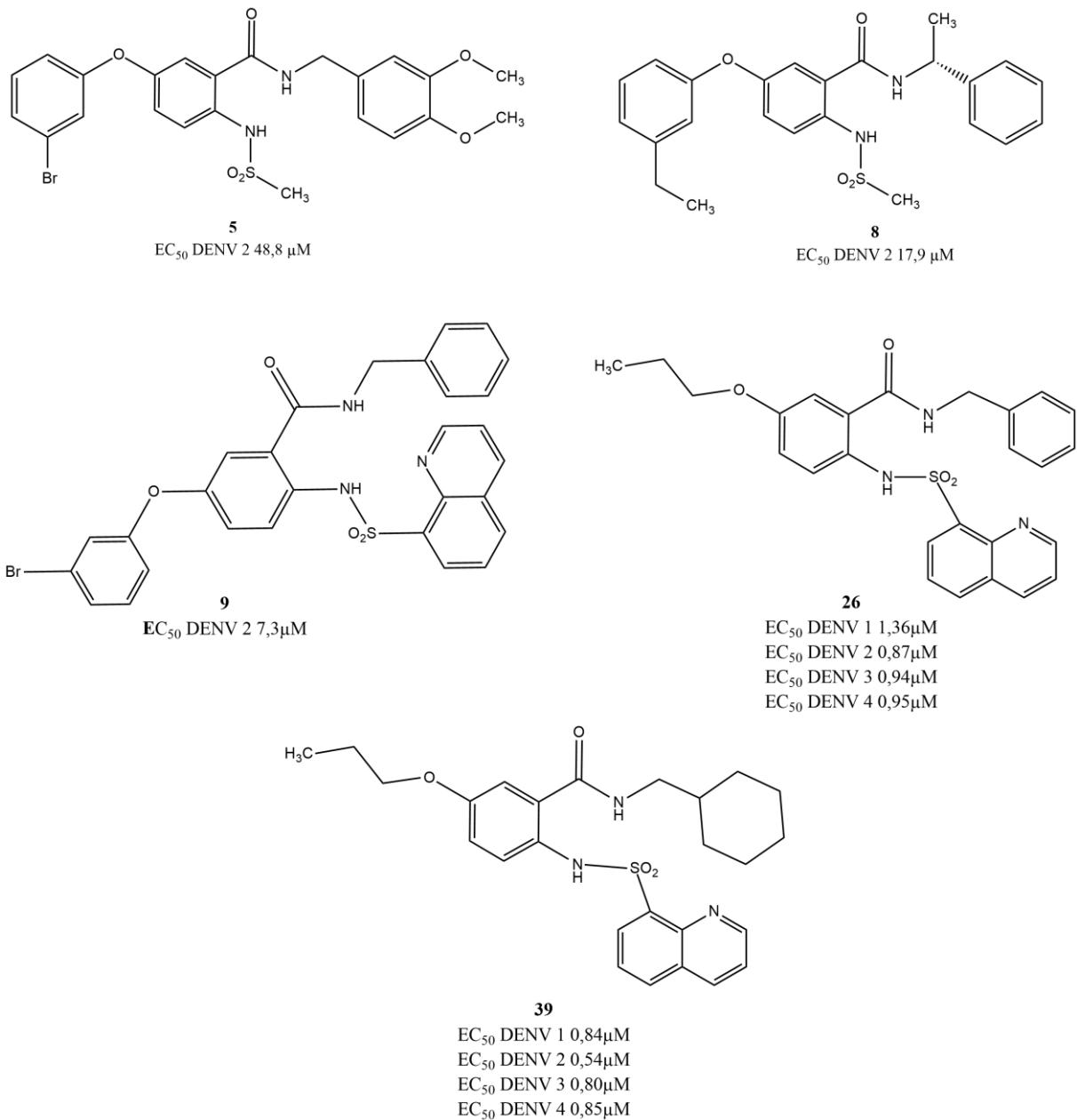


Figura 6: Compostos com ação antiviral em célula propostos como inibidores da NS5 de DENV 1, DENV 2, DENV 3 e DENV 4 em ordem de potência. Os valores de EC_{50} foram medidos por ensaio de placa contra infecção viral por DENV (GenBank: EU081177.1) após 47h da infecção. A numeração apresentada corresponde à do artigo original. Fonte: Felicetti et al., 2023.

Os autores ainda realizaram ensaios baseados em células para os 4 sorotipos, sendo a atividade dos compostos **26** e **39** da faixa submicromolar (0,54 – 1,36 μM). Concluindo, foram

assim propostos dois antivirais com ação pan-sorotipo. Curiosamente, ensaios *in vitro* de inibição da NS5 e NS3 foram realizados com **26**, sem inibição demonstrada. Utilizando diferentes linhagens celulares, também foi observado que o inibidor é mais potente para linhagens primatas e humana, com EC₅₀ mais alta para células de mosquito e hamster. Nenhuma linhagem mutante resistente foi obtida para as células primatas mesmo após 30 passagens do vírus, indicando uma especificidade do inibidor para o sistema celular do hospedeiro. O mecanismo de ação de **26** e sua relação com o sistema de defesa do hospedeiro ainda é estudado (Felicetti et al., 2023).

Em 2025, foi publicado o trabalho de Zong e colaboradores, que, por meio de uma abordagem mista (computacional e experimental), propõem novos inibidores de DENV-2, que superariam o composto **27** do trabalho de Lim e colaboradores. A trajetória do trabalho se inicia em uma triagem virtual de 1,6 milhão de compostos oriundos de dois bancos de moléculas combinados. 27 compostos são selecionados, sendo o objetivo do trabalho a descoberta de moléculas pequenas, que, de acordo com os autores, sejam de síntese e administração facilitada. No artigo são tabeladas as estruturas, IC₅₀, CC₅₀ e SI para 23 desses compostos. Os testes foram realizados para DENV-NGC, uma linhagem de DENV 2. Testes *in vitro* específicos para a NS5 não são relatados.

Três dos compostos avaliados apresentam IC₅₀ menor do que o composto **27** (Lim et al., 2016). A estabilidade dos complexos, analisando força de interação e flexibilidade frente a mudanças conformacionais, além da ocupação do sítio alostérico, são avaliados por estudos computacionais de dinâmica molecular a partir dos complexos obtidos pela docagem molecular. Ao fim de todas as análises, dois compostos são apresentados como mais potentes que o composto **27** (IC₅₀ = 19,67 μ M (\pm 1,12 μ M) segundo Zong et al., 2025): **D1** e **D8** (Figura 7). Os próprios autores concluem que o mecanismo de ação e sua relação com vias de sinalização imune dos hospedeiros precisam ser mais explorados (Zong et al., 2025). Ademais, o trabalho ainda não foi expandido para os demais sorotipos, nem mesmo computacionalmente.

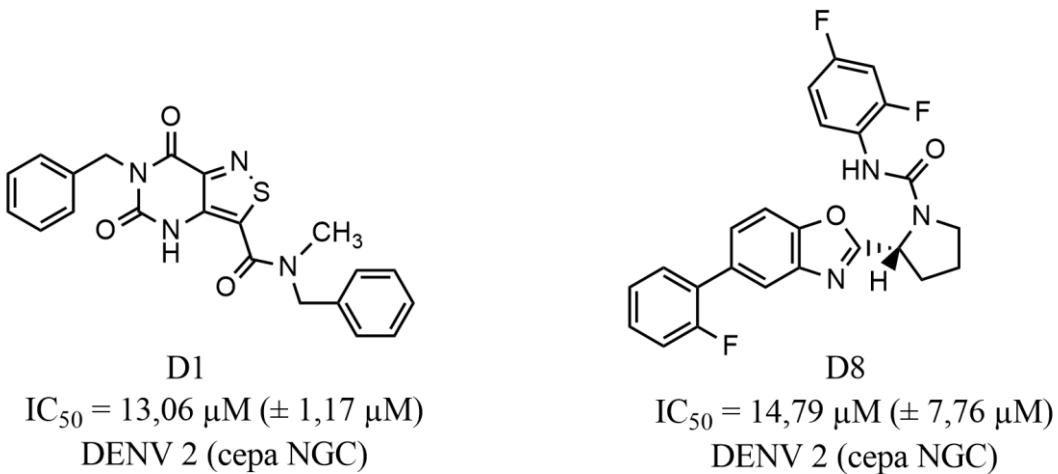


Figura 7: Antivirais contra DENV-2 que superam o composto 27 ($IC_{50} = 19,67 \mu M (\pm 1,12 \mu M)$). Os valores foram medidos para DENV2-NGC por infecção de cultura de células. As culturas foram mantidas após a infecção a 37°C por 3 dias e os resultados foram coletados por leitura luminescente. A numeração apresentada corresponde à do artigo original. Fonte: Zong *et al.* 2025.

O mesmo grupo de pesquisa publicou, também em 2025, outro artigo de metodologia semelhante. Utilizando outros dois bancos de moléculas, somando um total de 1,1 milhão de candidatos iniciais, a triagem virtual baseada em docagem foi executada tendo como alvo a estrutura de código PDB 5K5M (RdRp – DENV2). As moléculas foram filtradas por propriedades farmacocinéticas favoráveis e viabilidade sintética. A previsão de energia livre da interação foi calculada usando modelo de solvatação VSGB (do inglês, *variable solvation model based on the generalized Born approach*) considerando solvente aquoso. Diferente do trabalho anterior, os 47 compostos melhor classificados pela triagem foram testados não só pela sua capacidade antiviral, mas também especificamente para a RdRp. Além disso, estudos de dinâmica molecular avaliaram a estabilidade dos complexos em função do tempo (Wei *et al.*, 2025).

O composto **R37** (Figura 8) foi o antiviral mais potente, mas apresentou índice de seletividade (SI) baixo, indicando alta citotoxicidade. Além disso, sua estrutura é capaz de interagir bem em ambientes hidrofóbicos, podendo levar a interações fora do alvo. O composto **R29** (Figura 8) apresentou melhor equilíbrio entre potência antiviral e citotoxicidade, sendo um candidato considerado mais seguro. Apenas **R37** foi selecionado para o ensaio de inibição da polimerase viral, realizado por fluorescência, onde obteve o resultado de 10,86 μM . O composto **27**, usado como padrão positivo, não foi testado nas mesmas condições (Wei *et al.*, 2025).

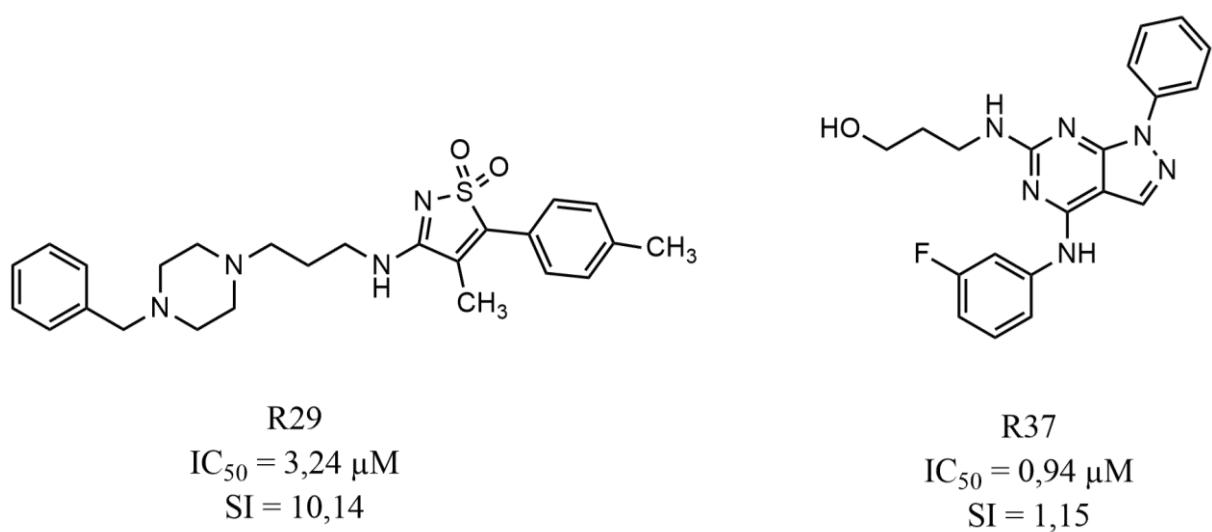


Figura 8: Os dois antivirais mais potentes contra DENV-2 apresentados por Wei e colaboradores. Os valores foram medidos para DENV2-NGC por infecção de cultura de células. As culturas foram mantidas após a infecção a 37°C por 3 dias e os resultados foram coletados por leitura luminescente. A numeração apresentada corresponde à do artigo original. Fonte: Wei *et al.* 2025.

Apesar das diversas abordagens descritas na literatura, a busca por inibidores efetivos e clinicamente viáveis da NS5 continua desafiadora, o que justifica o uso de metodologias computacionais como as baseadas na estrutura da proteína, tratadas a seguir.

2.5. Técnicas de descoberta de fármacos baseadas em estrutura

A modelagem molecular tem sido uma abordagem de grande importância na compreensão de sistemas biológicos complexos, permitindo, por exemplo, a visualização e predição de interações moleculares que seriam difíceis de acessar experimentalmente, seja por limitações metodológicas ou financeiras (Rocha; Sant'Anna, 2024). O estudo computacional baseado em modelos teóricos tem custos consideravelmente menores e grande potencial para maior rapidez que métodos experimentais (Duarte et al., 2025).

Um dos conceitos fundamentais na Química Medicinal é a relação entre a estrutura molecular e a atividade biológica. Desde a proposta do modelo chave-fechadura por Fischer, em 1894, entende-se que a interação entre uma biomolécula e seu ligante depende da complementaridade estrutural entre eles. Esse modelo, que evoluiu ao longo do tempo, como pela inclusão de aspectos conformatacionais na interação, é um princípio que norteia grande parte das estratégias modernas de descoberta de fármacos, que se beneficiam do avanço de métodos computacionais para prever e otimizar essas interações (Rocha; Sant'Anna, 2024).

Com o desenvolvimento do planejamento de fármacos auxiliado por computador (CADD, do inglês *computer aided drug design*), tornou-se possível utilizar abordagens computacionais para guiar a identificação e otimização de novos compostos bioativos. Como dito antes, o CADD pode ser dividido em duas grandes categorias: O planejamento baseado no ligante (LBDD, do inglês *ligand based drug design*) e o planejamento baseado na estrutura do alvo (SBDD, do inglês *structure based drug design*).

No LBDD, a busca por novos compostos é baseada em semelhanças estruturais ou estatísticas com ligantes conhecidos, que podem ser exploradas por meio de modelos de relações quantitativas entre a estrutura e a atividade (QSAR, do inglês *Quantitative Structure-Activity Relationship*), mapas farmacofóricos e métodos de similaridade molecular (Rocha; Sant'Anna, 2024). O modelo de QSAR tem sua origem nos estudos da relação entre a estrutura e as propriedades (QSPR, do inglês *Quantitative Structure-Property Relationships*). Nesse caso, usam-se dados de origem empírica ou de cálculos computacionais como descritores que podem ser relacionados com as propriedades de determinado composto. O primeiro exemplo desses estudos foi feito por Hammet, que buscou quantificar o efeito eletrônico de substituintes em anéis aromáticos sobre a reatividade de compostos orgânicos, a relacionando a constantes termodinâmicas e cinéticas (Duarte et al., 2025).

No SBDD, a estrutura tridimensional do alvo biológico, geralmente uma proteína, é

utilizada diretamente para prever e otimizar interações com potenciais inibidores, tornando esse método especialmente útil quando há informações detalhadas sobre o sítio ativo ou sítios alostéricos da biomolécula. Essa estrutura também pode ser um modelo tridimensional do alvo, ou seja, construído por técnicas de modelagem de proteínas. As técnicas incluem a modelagem comparativa (ou por homologia), métodos de aprendizado de máquina (ML, do inglês *Machine Learning*) e métodos *ab initio* (Rocha; Sant'Anna, 2024).

A modelagem comparativa utiliza a estrutura de um molde tridimensional determinada experimentalmente e a estrutura primária da proteína que se deseja modelar. Para a modelagem comparativa ser possível, deve haver entre o molde e a proteína uma identidade de sequência de peptídeos de ao menos 30% (Rocha; Sant'Anna, 2024).

No campo de ML, há duas classes de métodos: o aprendizado supervisionado e o não-supervisionado. O aprendizado supervisionado consiste em treinar um modelo computacional a partir de dados rotulados, isto é, entradas associadas a saídas conhecidas, com o objetivo de aprender uma função que relate essas variáveis e permita fazer previsões sobre novos dados. O aprendizado não supervisionado opera sobre conjuntos de dados não rotulados, buscando identificar estruturas latentes, agrupamentos naturais ou padrões subjacentes que descrevam a organização dos dados no espaço de características. Enquanto o primeiro é mais voltado à previsão, o segundo é exploratório por natureza (Duarte et al., 2025).

O método de ML utilizado para a modelagem de proteínas é o aprendizado por reforço, um derivado híbrido das classes citadas. O processo de enovelamento proteico é simulado, de forma que um agente (modelo computacional) interage com um ambiente (representação estrutural da proteína) realizando ações (modificações geométricas) em um sistema inicialmente desnaturado. A cada passo, o agente recebe recompensas ou penalidades com base em critérios definidos, como a proximidade entre a conformação atual e a de estruturas nativas (enoveladas). O objetivo é aprender, por tentativa e erro, a sequência de transformações que leva à estrutura energeticamente mais estável (Duarte et al., 2025).

Por fim, a modelagem *ab initio* (do latim, desde o início), busca construir a estrutura tridimensional da proteína a partir de sua sequência, ou seja, estrutura primária. É o método de modelagem que busca definir como uma proteína adota uma estrutura tridimensional específica a partir de princípios físico-químicos. A previsão é feita por um algoritmo computacional, que faz a amostragem conformacional, ou seja, gera um grande número de possíveis conformações para uma dada sequência, guiado por uma função de energia. Sendo assim, a eficiência do

método depende da eficiência da função (Lee; Freddolino; Zhang, 2017).

As funções para esse tipo de modelagem podem ser baseadas em princípios físicos ou em conhecimento. As primeiras podem usar campos de força, como AMBER, CHARMM, OPLS, UNRES e QUARK, entre outros. As últimas utilizam dados estatísticos das estruturas determinadas experimentalmente e depositadas no PDB. A amostragem conformacional, por sua vez, pode empregar métodos como simulação por dinâmica molecular, algoritmos genéticos ou simulação de Monte Carlo, com o objetivo de explorar o espaço conformacional de estruturas possíveis e identificar aquelas com menor energia (Lee; Freddolino; Zhang, 2017).

Retomando a ideia geral do SBDD, a estrutura do alvo deve guiar o planejamento de ligantes. Esse processo pode usar técnicas como a docagem molecular, que prevê modos de ligação e afinidades de interação entre ligantes e proteínas, e dinâmica molecular, que permite avaliar a estabilidade desses complexos ao longo do tempo. Com a crescente disponibilidade de estruturas de proteínas em bases de dados como o PDB e o aprimoramento de algoritmos de simulação, o SBDD se tornou uma abordagem poderosa para acelerar e tornar mais eficiente o processo de descoberta de novos fármacos (Rocha; Sant'Anna, 2024).

2.5.1. Docagem molecular

A docagem molecular depende de dois elementos: o algoritmo de busca e a função de pontuação. O algoritmo de busca será responsável por propor diferentes conformações e orientações do ligante no sítio de ligação da proteína, o que definimos por “pose” (Torres et al., 2019). A função de pontuação avalia as poses obtidas pelo algoritmo e as classifica, gerando uma pontuação. Quanto maior a pontuação, maior é a afinidade prevista do ligante pelo sítio de ligação, na pose específica explicitada (Da Silva Rocha et al., 2019).

Uma limitação importante da docagem molecular é a impossibilidade de prever com precisão se o ligante testado será capaz de induzir modificações conformacionais significativas no sítio de ligação da proteína. No ambiente biológico, é comum que enzimas apresentem alterações estruturais relevantes ao interagirem com cofatores, passando da forma apo (sem cofator) para a forma holo (com cofator). Como essa transição conformacional induzida pelo cofator não é reproduzida, em geral, nas simulações de docagem, torna-se necessário selecionar previamente qual conformação da proteína será utilizada — uma decisão que deve considerar tanto a disponibilidade de estruturas cristalográficas quanto os objetivos específicos do estudo (Kirsch et al., 2019).

A relevância da escolha da conformação da proteína utilizada na docagem é evidenciada por Zhang e colaboradores (2022), que propuseram uma abordagem robusta para refinar estruturas apo e gerar conformações holo-similares por meio do refinamento do sítio de ligação com base em moldes holo de baixa homologia, mas alta similaridade topológica. Utilizando simulações de dinâmica molecular com restrições derivadas dessas estruturas modelo, os autores aproximaram a conformação apo da conformação holo. Seus resultados demonstraram melhorias substanciais tanto na acurácia da predição do modo de ligação dos ligantes quanto no desempenho em triagens virtuais. Essa estratégia mostra-se especialmente útil na ausência de estruturas holo disponíveis, sendo uma alternativa eficiente para melhorar os resultados de docagem, sobretudo quando há mudanças conformacionais relevantes no sítio de ligação (Zhang et al., 2022).

Deve ser destacado que há algoritmos de docagem que permitem conferir liberdade conformacional parcial para cadeias laterais e para o esqueleto polipeptídico da proteína, o que permite algum nível de ajuste induzido pela interação com o ligante durante a docagem, como, por exemplo, o do programa *Genetic Optimisation for Ligand Docking* (GOLD), que será apresentado mais a frente (Cambridge Crystallographic Data Centre, 2022).

Os algoritmos de busca podem ser classificados em três grandes grupos: sistemáticos, estocásticos e determinísticos. Os algoritmos sistemáticos realizam varreduras incrementais nos graus de liberdade do ligante, podendo ser subdivididos em abordagens exaustivas, construção incremental e conjuntos conformacionais. Os algoritmos estocásticos realizam modificações aleatórias nos graus de liberdade e, por meio de processos iterativos, buscam superar o risco de aprisionamento em mínimos locais; exemplos comuns incluem os métodos de Monte Carlo, algoritmos genéticos, busca tabu e otimização por enxame de partículas, sendo amplamente utilizados. Por fim, algoritmos determinísticos, como os baseados em minimização de energia ou dinâmica molecular, atualizam a pose do ligante com base direta na iteração anterior, mas apresentam maior custo computacional e risco de convergência para mínimos locais (Torres et al., 2019).

De modo geral, o desempenho desses algoritmos pode ser comprometido quando se trabalha com ligantes de alta flexibilidade, em razão do aumento exponencial no espaço conformacional. Nesse cenário, a utilização de fragmentos (conforme discutido no item 2.5.3) se mostra vantajosa, pois reduz os graus de liberdade e permite uma exploração mais eficiente e acurada das interações relevantes com o sítio de ligação (Torres et al., 2019).

As funções de pontuação podem ser classificadas em três categorias principais: baseadas em campos de força, empíricas e baseadas em conhecimento. Mais recentemente, foram lançadas as funções baseadas em aprendizado de máquina, que aprendem padrões a partir de um conjunto de treinamento, sendo capazes de generalizar o conhecimento e estimar a afinidade de ligação entre proteínas e moléculas (Scantlebury et al., 2023).

As funções de campo de força somam termos energéticos interatônicos (relacionados às ligações) e intermoleculares (relacionados às interações), frequentemente negligenciando efeitos do solvente e contribuição entrópica. Já as funções empíricas seguem a lógica QSAR, empregando dados experimentais para estimar a energia de ligação. Por fim, as funções baseadas em conhecimento, utilizam dados estatísticos extraídos de estruturas cristalográficas para inferir a probabilidade de ocorrência de interações favoráveis. Embora cada abordagem tenha limitações, a combinação de diferentes funções por meio de estratégias de pontuação por consenso tem sido amplamente utilizada para aumentar a confiabilidade dos resultados. Para esse fim, pode ser usado o chamado “rescore”, que funciona como uma repontuação da pose com uma função diferente daquela usada na sua busca (Torres et al., 2019).

O GOLD (CCDC), sigla para *Genetic Optimisation for Ligand Docking*, é um programa

largamente utilizado para a docagem molecular, especialmente no Brasil, pela parceria com a CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) por meio da BDEC (Base de Estruturas Cristalinas). Pela BDEC, pesquisadores de áreas correlatas de ciências e engenharias vinculados a instituições de ensino superior conveniadas se registram e recebem acesso a uma plataforma com conteúdo voltado para a disseminação de informações técnicas e científicas de estruturas cristalinas, bem como suas formas, propriedades, aplicações e funções. Na plataforma, está disponível o pacote da *Cambridge Structural Database* (CSD), que inclui o GOLD (“BdEC - Apresentação, Características, Diferenciais.”, [S.d.]).

O GOLD utiliza um algoritmo genético para encontrar a melhor pose do ligante na cavidade de docagem. Os algoritmos genéticos são métodos computacionais que simulam o processo de evolução natural para resolver problemas complexos. Eles pertencem à categoria de métodos heurísticos, que são estratégias práticas desenvolvidas para encontrar boas soluções quando não é viável testar todas as possibilidades. Em vez de garantirem a resposta perfeita, heurísticas buscam respostas “boas o suficiente” de forma mais rápida e eficiente. No caso dos algoritmos genéticos, uma população inicial de soluções é criada e, a cada ciclo (ou geração), as melhores soluções são selecionadas para gerar novas, por meio dos chamados operadores genéticos. Com o passar das gerações, essas soluções vão se aperfeiçoando gradualmente, como se estivessem “evoluindo” em direção à melhor resposta possível (Lee; Freddolino; Zhang, 2017).

Os operadores genéticos utilizados no programa são mutação, cruzamento e migração. A mutação é o emprego de pequenas alterações na posição ou conformação do ligante; o cruzamento é a combinação de elementos de diferentes poses para gerar novas soluções; e a migração é a seleção de algumas soluções para serem movidas entre subpopulações (ilhas) para evitar convergência prematura (Cambridge Crystallographic Data Centre, 2022).

Convém aos propósitos do trabalho destacar duas funções de pontuação: ChemPLP e ChemScore.

A função de pontuação ChemPLP (*Piecewise Linear Potential*) é a função padrão utilizada pelo programa GOLD. Essa função pertence à classe das funções empíricas e é baseada em um potencial linear por partes, ou seja, uma função matemática que descreve a energia potencial de interação entre pares de átomos em função da distância, modelando de forma diferenciada regiões de atração e repulsão. O objetivo é representar, de forma eficiente, a complementaridade entre o ligante e o sítio de ligação da proteína (Cambridge

Crystallographic Data Centre, 2022).

No modelo ChemPLP, todos os átomos pesados (diferentes de hidrogênio) da proteína e do ligante são classificados com base no seu potencial de interação química como doadores ou acetores de ligação de hidrogênio ou doador/aceptor ou, ainda, não polares. Além disso, íons metálicos presentes no sítio de ligação recebem uma classificação específica, permitindo a modelagem de interações específicas com cofatores metálicos, quando presentes (Korb; Stützle; Exner, 2009).

A função ChemPLP incorpora termos adicionais de outra função, ChemScore, incluindo contribuições empíricas para ligações de hidrogênio e interações com metais, além de penalidades associadas à geometria do ligante, como energia torcional e colisões (Korb; Stützle; Exner, 2009). Devido ao seu desempenho robusto na predição de poses de ligação, a ChemPLP é amplamente recomendada para triagens virtuais e tarefas de docagem baseadas na estrutura, oferecendo um bom equilíbrio entre eficiência computacional e acurácia preditiva (Cambridge Crystallographic Data Centre, 2022).

A função ChemScore foi desenvolvida a partir de dados empíricos de 82 complexos proteína-ligante com afinidades de ligação conhecidas, sendo treinada por regressão contra dados medidos. Ela estima a mudança total de energia livre que ocorre na interação do ligante com seu receptor. Inclui penalidades para colisões e conformações internas desfavoráveis. É recomendada para estimativas de afinidade por meio da predição de energia livre da ligação (Cambridge Crystallographic Data Centre, 2022; Eldridge et al., 1997a).

2.5.2. Métodos quânticos semi-empíricos

Funções utilizadas para cálculos de energia de estruturas moleculares podem recorrer a aproximações clássicas ou quânticas. Fatores como a precisão desejada e a capacidade computacional disponível devem ser considerados na escolha do método a ser utilizado (Jensen, 2017).

Nessa perspectiva, dentro das aproximações clássicas, existem os métodos de mecânica molecular (MM), que tratam os átomos como esferas, sem considerar que são formados por núcleo e elétrons, e as ligações entre os átomos como molas. Um campo de força é utilizado, definindo um conjunto de funções que determinam como a variação de parâmetros estruturais afetará a energia da estrutura (Lipkowitz; Boyd, 1991). Essas funções contêm constantes características geralmente oriundas de dados experimentais, sendo, por consequência, o método chamado de empírico (Barreiro; Fraga, 2014). Essa abordagem é bastante eficiente, mas apresenta algumas limitações por não considerar os elétrons explicitamente.

Métodos quanto-mecânicos consideram que a estrutura molecular é formada por núcleos e elétrons, dependentes, a princípio, apenas das constantes fundamentais da natureza (Jensen, 2017). O maior complicador nessa consideração é a natureza ondulatória dessas partículas, que foi considerada por E. Schrödinger na equação para a determinação de energia do sistema, a qual leva seu nome. Na equação de Schrödinger, temos o operador Hamiltoniano e a função de onda. O primeiro corresponde à soma de operadores de energia cinética e potencial, enquanto a segunda é dependente das coordenadas dos núcleos e elétrons (Barreiro; Fraga, 2014). Por ser uma equação diferencial, a solução da equação de Schrödinger requer integração.

Simplificações na equação são necessárias para sua resolução, uma primeira é a aproximação de Born-Oppenheimer, que considera os núcleos como partículas "congeladas" e calcula somente a energia eletrônica no campo gerado por esses núcleos pela resolução da equação de Schrödinger. Posteriormente, é adicionado um termo de repulsão entre núcleos, chegando à previsão de energia total. Por se considerar a natureza ondulatória dos elétrons, são seguidas as definições da mecânica quântica, resultando valores quantizados para a energia eletrônica (Barreiro; Fraga, 2014).

Para compreender os métodos quânticos SE é preciso conhecer a definição de um campo autoconsistente (SCF, do inglês *self-consistent field*). A ideia do SCF vem do modelo de Hartree-Fock (HF), em que cada elétron é considerado inserido em dois campos de força: um dos núcleos "congelados" e outro dos demais elétrons do sistema. Para isso, são geradas funções

de onda moleculares para cada elétron, aproximadas como combinações lineares de funções de onda atômicas (aproximação LCAO, do inglês *Linear Combination of Atomic Orbitals*), ponderadas por coeficientes que adequam a participação de cada orbital atômico no orbital molecular. Esses coeficientes são variados objetivando a menor energia eletrônica possível. Como o campo sobre um determinado elétron depende apenas das funções de onda dos demais e essas ainda precisam ser determinadas, inicia-se o processo com funções de onda tentativas para os demais elétrons, que vão sendo melhoradas em um processo cíclico. O processo ocorre até que de um ciclo para outro não haja mudança significativa na energia. Nessa situação, é dito que o campo gerado pelos elétrons se tornou autoconsistente (Barreiro; Fraga, 2014; Custodio et al., 2002).

Assim, é possível diferenciar os métodos *ab initio* dos métodos SE. Para os métodos *ab initio*, não são feitas mais aproximações além das citadas na resolução da equação de Schrödinger (Jensen, 2017). Para os métodos SE, parâmetros experimentais ou previamente calculados são inseridos para reduzir o número de integrais a serem resolvidas na resolução da equação de Schrödinger, levando a um menor custo computacional (Stewart; Seiler, 1990).

Michael J. S. Dewar e colaboradores desenvolveram alguns dos primeiros métodos semiempíricos (SE), estabelecendo as bases para uma abordagem que buscava um equilíbrio entre precisão e eficiência computacional. Além disso, objetivavam que os programas para realização dos cálculos deveriam ser acessíveis aos não especialistas (Barreiro; Fraga, 2014). Nos métodos SE desenvolvidos por eles, foram incluídos na resolução da equação de Schrödinger apenas os elétrons de valência, considerando as camadas eletrônicas mais internas como um caroço ou “core” (“Semiempirical Calculations”, 2004). Dados experimentais foram utilizados para que houvesse previsão correta de dados como momentos dipolo, energias de ionização, calores de formação e geometrias de equilíbrio (Barreiro; Fraga, 2014).

A abordagem de Dewar levou à criação inicial do CNDO (*Complete Neglect of Differential Overlap*), que posteriormente foi aprimorado ao método conhecido como MNDO (*Modified Neglect of Differential Overlap*), que foi largamente utilizado e tornou obsoletos os demais métodos da época. Entretanto, esse ainda não apresentava resultados com qualidade suficiente para maior aceitação. O modelo AM1 (*Austin Model 1*) trouxe grandes evoluções, incluindo parametrizações para 20 elementos (“Semiempirical Calculations”, 2006).

James J. P. Stewart aprimorou a abordagem SE ao desenvolver o PM3 (*Parametric Method 3*), introduzindo um novo esquema de parametrização, mais rápido. O método foi

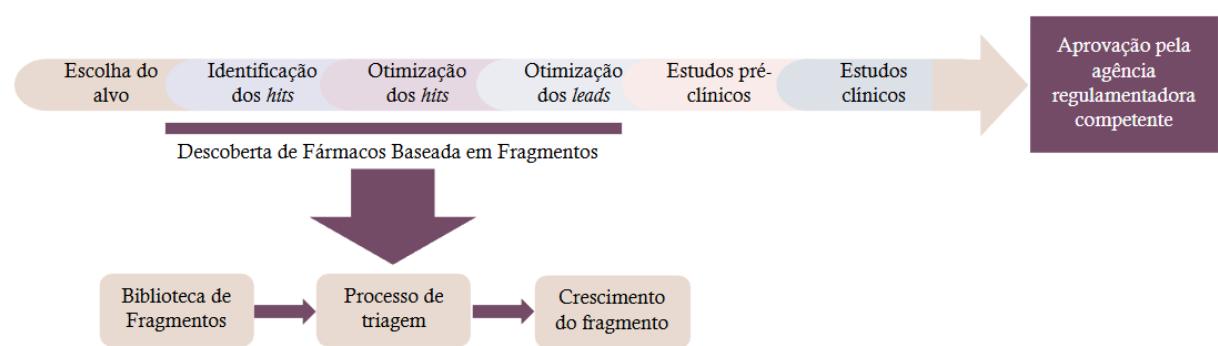
parametrizado para um amplo conjunto de elementos, sendo comprovadamente mais preciso para moléculas contendo H, C, N, O, F, Cl, Br e I. Apesar do avanço, houve uma divergência conceitual entre os pesquisadores da área. Dewar acreditava que o PM3 era apenas minimamente mais preciso que seu método mais recente, o AM1, de forma que, para ele, um novo método SE só deveria ser aceito se tornasse o anterior obsoleto (“Semiempirical Calculations”, 2006). Cartas aos editores foram submetidas à época (anos 1980 - 1990) e publicadas pelo *Journal of Computational Chemistry*. Evidentemente, a concorrência entre os programas semi-empíricos contribuiu, à época, para a rixa. O programa AMPAC foi inicialmente proposto por Dewar em 1985, ao passo que o MOPAC foi inicialmente proposto por Stewart em 1989. Atualmente, o AMPAC também conta com a colaboração de Stewart (Dewar et al., 1990; Stewart, 1990a, 1990b).

A evolução dos métodos semiempíricos acompanhou a necessidade de maior precisão na modelagem de biomoléculas e interações intermoleculares. O PM6, desenvolvido em 2007, apresentou um avanço significativo pelo refinamento de parâmetros empíricos e melhoramento da descrição de interações núcleo-núcleo, além da inclusão de nova parametrização para 70 elementos. Posteriormente, o PM7, introduzido em 2013, incorporou termos adicionais para modelar efeitos de dispersão e ligações de hidrogênio com maior precisão, tornando-o mais adequado para sistemas biológicos (Stewart, 2017).

Ainda mais recentemente, em 2023, foi lançado o PM6-ORG, um método especificamente parametrizado para modelar fenômenos de interesse bioquímico. A modificação principal foi no conjunto de treinamento para a parametrização; elementos comumente encontrados em enzimas tiveram sua parametrização otimizada, focando na minimização de colisões e diminuição dos erros na previsão de interações intermoleculares e entalpias de formação. A diferença no erro absoluto médio das entalpias de formação calculas pelo PM7 (4,22kcal/mol) e pelo PM6-ORG (3,88kcal/mol) é baixa, de menos de 1kcal.mol⁻¹. O erro médio de RMS na geometria do esqueleto polipeptídico também tem pouca diferença, de forma que a utilização do PM7 não prejudica a análise de sistemas bioquímicos (Stewart, 2023).

2.5.3. Fragment based drug design

FBDD é uma técnica de desenvolvimento de moléculas bioativas que parte de moléculas pequenas e simples em sua estrutura química, os chamados fragmentos. A literatura mais recente restringe a massa de um fragmento em 300 Da. Esses fragmentos são selecionados pela sua interação com um alvo, podendo esse ser uma enzima, um receptor, um canal iônico ou qualquer outra macroestrutura de interesse biológico. A interação dos fragmentos com os alvos costuma ser fraca, já que pensando em complementaridade das estruturas fragmento-alvo, o fragmento é muito simples e pequeno. A partir de uma biblioteca de fragmentos, os de melhor desempenho são selecionados e otimizados, visando uma molécula com propriedades fármaco-similares. O Esquema 1 ilustra as etapas do processo (Kirsch et al., 2019; Li, 2020).



Esquema 1: Resumo das etapas de planejamento de fármacos, destacando onde o processo de FBDD pode ser inserido. Adaptado de Li, 2020.

O FBDD não tem como objetivo substituir completamente técnicas bem estruturadas como a triagem de alto desempenho (HTS, do inglês *high-throughput screening*), que é, por definição, a testagem robotizada de mais de 10.000 diferentes compostos por dia, em placa com poços, por meio de um ensaio biológico em comum, desenhado para o estudo do alvo de interesse (Wildey et al., 2017). Entretanto, oferece uma alternativa de custo reduzido capaz de explorar uma diversidade químico-estrutural abrangente, com muitas possibilidades de otimização.

Os *hits* no FBDD precisam ser identificados por métodos de alta sensibilidade, uma vez que as interações não são comumente fortes e os fragmentos são pequenos. Os métodos biofísicos mais comuns são a cristalografia de raios-X e ressonância magnética nuclear (RMN) (Toungé; Parker, 2011). Um exemplo de fármaco descoberto por técnicas de FBDD é o vemurafenibe, um antineoplásico utilizado no tratamento de melanoma em estágio avançado (Li, 2020). Outros exemplos e seus fragmentos iniciais podem ser observados na Figura 9. É

interessante notar que, devido ao menor tamanho, os fragmentos também podem ser mais adequados a um sítio alostérico, caso em que normalmente a cavidade disponível é de menor volume que o sítio ortostérico. Em suma, para sítios essencialmente pequenos, a FBDD é muito promissora (Kirsch et al., 2019).

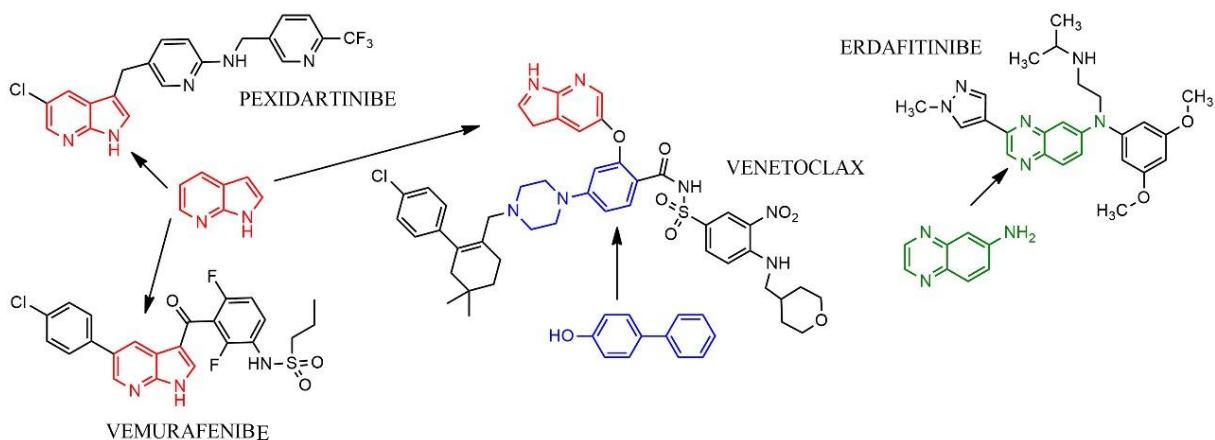


Figura 9: Estruturas de quatro medicamentos aprovados oriundos de FBDD. O PLX3397 (penidartinibe) foi desenvolvido a partir do 7-azaindol (Zhang et al., 2013). O vemurafenibe foi construído a partir do mesmo fragmento (Tsai et al., 2008). As estruturas do venetoclax (um inibidor de Bcl2) (Souers et al., 2013) e do erdafitinibe e de seus respectivos fragmentos de origem (Murray C.W. et al., 2019) também são apresentadas. Fonte: Li, 2020

Diferentemente da conhecida regra dos 5 de Lipinski (Lipinski et al., 2001) para moléculas *drug-like*, para fragmentos utiliza-se a regra dos 3: Massa molecular menor que 300, $\log P < 3$ e menos do que três doadores ou aceptores de ligações de hidrogênio (Jhoti et al., 2013). Entretanto, como os compostos são facilmente modificáveis, a regra pode ser adotada como uma sugestão e não uma limitação. Outras características que podem ser benéficas são área superficial polar menor que 60 \AA^2 e número de ligações rotacionáveis menor que 3. É comum que os fragmentos tenham menos que 20 átomos diferentes de hidrogênio. Como a otimização dos *hits* permite combinar diversos fragmentos, a biblioteca inicial não precisa ter um número tão alto de estruturas, como comumente é necessário em outras técnicas de descoberta de fármacos (Kirsch et al., 2019; Li, 2020).

Para triagem virtual de fragmentos, as mesmas ideias já citadas no tópico SBDD são utilizadas. A determinação estrutural do alvo é essencial e estratégias como docagem e dinâmica molecular são aplicadas. Posteriormente à seleção, há ferramentas para provar as interações entre proteína e fragmento, como ensaios celulares e espectrometria de massas no caso de ligações covalentes entre alvo e *hit* (Li, 2020).

Uma vez identificados e comprovados, a etapa de otimização se inicia. As três estratégias mais comumente utilizadas são: crescimento de fragmentos (do inglês, *fragment*

growing), fusão de fragmentos (do inglês *fragment hopping* ou *merging*) e união de fragmentos (do inglês *fragment linking*) (Kirsch et al., 2019).

A técnica de crescimento de fragmentos é a mais utilizada para aumentar massa e potência e foi a técnica utilizada para o desenvolvimento dos quatro fármacos da Figura 9. A técnica consiste em, a partir do fragmento (bloco inicial), adicionar substituintes químicos que aumentem sua potência. Apesar da técnica ser otimizada na presença de informações estruturais, essas não são essenciais para a sua aplicação (Li, 2020). É necessário o estudo dos vetores de crescimento, que vão indicar a direção adequada na qual o fragmento deve ser aumentado. Além disso, é essencial observar a complementaridade geométrica, flexibilidade e tamanho do substituinte (Kirsch et al., 2019).

A técnica de fusão dos fragmentos prevê a combinação das características de diferentes *hits* que interagem diferentemente no mesmo sítio alvo, se sobrepondo. Dessa forma, é como se ocorresse uma “fusão” dos fragmentos, aproveitando grupos em comum e agregando os demais em uma mesma molécula. Para o uso dessa estratégia, é necessária a visualização das interações *hit*-proteína, podendo essa ser obtida por cristalografia, RMN ou docagem (Li, 2020).

Por fim, a estratégia de união de fragmentos é considerada a mais poderosa. A ideia é ligar dois ou mais fragmentos que interajam em bolsões diferentes, mas adjacentes, os potencializando como candidatos a um composto líder. O desafio é que o conector entre os fragmentos, que podemos entender como um espaçador, não afete sua atividade e os mantenha na geometria necessária para a soma das interações. De acordo com cálculos de energia livre, ligar dois fragmentos de atividade na faixa mM pode resultar em um composto na faixa nM. Para essa técnica, é importante que o sítio seja grande o suficiente para acomodar um composto maior. Informações estruturais auxiliam no planejamento da ligação entre os *hits* (Harner; Frank; Fesik, 2013; Li, 2020).

Na perspectiva do atual trabalho, em que as estruturas do alvo e dos inibidores são conhecidos e boa parte dos compostos da quimioteca tem massa molecular < 300 Da, a FBDD surge como uma valiosa ferramenta, como veremos nos resultados.

3. Objetivos

O objetivo geral do presente trabalho é realizar uma triagem virtual por docagem molecular de compostos da quimioteca do PPGQ-UFRRJ visando identificar candidatos a ligantes do sítio alostérico da NS5 do DENV. A triagem será feita em duas vertentes: a de fragmentos (moléculas com até 300 Da) e a de ligantes grandes (massa até 550 Da).

A seguir, serão feitos (i) a otimização dos ligantes-fragmentos, utilizando crescimento de fragmentos; (ii) o cálculo das entalpias de interação para os ligantes originais e otimizados e (iii) previsão de propriedades ADMET para os ligantes melhor classificados.

4. Metodologia

4.1. Obtenção e modelagem dos ligantes

Os compostos usados na triagem foram obtidos da quimioteca em construção do Laboratório de Modelagem Molecular do Instituto de Química da Universidade Federal Rural do Rio de Janeiro (LabMol – IQ/UFRRJ). Os componentes da quimioteca virtual são oriundos do banco de teses e dissertações do Programa de Pós Graduação em Química (PPGQ) da UFRRJ. Os alunos de graduação e pós-graduação responsáveis pela criação, expansão e curadoria da quimioteca, além da presente autora, são: Ana Carolina J. Oliveira; André Ávila L. Gonçalves; Bruno Sena de Oliveira; Layla D. Araújo, Lidiane D. Medeiros, Vinícius N. da Rocha, Vitor Fontes Biancardi e Yuri Terra Maia, todos sob orientação do Prof. Dr. Carlos Maurício R. Sant’Anna.

A modelagem dos ligantes foi realizada com o programa *Spartan’24* (Wavefunction, Inc.). A obtenção das geometrias de equilíbrio foi feita pelo método semi-empírico PM6, excetuando-se ligantes com presença de bromo, em que o método semiempírico PM3 foi utilizado. A quimioteca oferece, ao tempo do presente trabalho, arquivos de 657 compostos gravados no formato Spartan (*.spartan*), SYBYL (*.mol2*) e PDB (*.pdb*), modelados assumindo-se pH fisiológico (7,4).

4.2. Obtenção e preparação dos receptores

Foram coletados do Protein Data Bank (PDB) (BERMAN, 2000; RCSB.ORG) 9 estruturas cristalográficas da NS5 do vírus da dengue (5JJS, 5JJR, 5I3P, 5I3Q, 5K5M, 6XD0, 6XD1, 6IZZ e 6IZX). A escolha das citadas se dá por todas terem ligantes co-cristalizados em sítios alostéricos e apresentarem resoluções adequadas ao trabalho (até 2,5 Å). Posteriormente, observou-se quais estruturas foram exploradas por um mesmo grupo de pesquisa, com diferentes ligantes co-cristalizados e boa ocupação do sítio, necessária para o estudo computacional mais aprofundado. Nesse sentido, destacaram-se as estruturas 5JJS, 5JJR, 5I3P, 5I3Q e 5K5M. Dentre essas, 5JJR e 5JJS, assim como 5I3Q e 5I3P, são correspondentes, mudando apenas o ligante, portanto, selecionamos, em cada par, a estrutura de melhor resolução. Além disso, testes de redocagem foram conduzidos (detalhados no tópico 4.3; resultados no tópico 5.1), a fim de avaliar a capacidade da predição das geometrias de interação das funções de pontuação de docagem.

Os códigos das estruturas escolhidas foram 5JJS, 5I3Q e 5K5M. Todas são referenciadas por Lim e colaboradores (2016). As estruturas coletadas foram refinadas no programa *Swiss-PDBViewer 4.1.0* (Guex; Peitsch, 1997) em que quaisquer incompletudes da cristalografia foram reparadas. Todas as moléculas de água foram retiradas das estruturas para o estudo de triagem.

Utilizando o programa *GOLD 2022.3.0* (Jones et al., 1997), os átomos de H foram incluídos de acordo com os critérios padrão do programa e os ligantes co-cristalizados em cada estrutura foram removidos computacionalmente, abrindo o espaço do sítio para as posteriores docagens.

4.3. Escolha da função de pontuação e protocolo de triagem virtual

Após a retirada computacional dos ligantes originais, co-cristalizados nas estruturas, os mesmos foram redocados em suas estruturas de origem e tiveram a pontuação e raiz do desvio médio da posição (RMSD, do inglês, *root mean square deviation*) em relação à estrutura cristalográfica registrados. O procedimento, denominado redocagem, foi realizado com o programa *GOLD 2022.3.0*. Foi escolhida a função que deu maiores pontuações para as poses de menor RMSD. Algumas estruturas não obtiveram bons RMSD em nenhuma função de pontuação, de forma que foram consideradas inadequadas para o prosseguimento do estudo.

Ainda no *GOLD 2022.3.0*, iniciou-se a triagem virtual com a docagem em conjunto, isto é, as estruturas selecionadas da proteína foram sobrepostas e transformadas em um conjunto. Os 657 ligantes, em formato SYBYL, foram selecionados para a docagem, definindo-se o sítio de docagem com um raio de 6 Å a partir da posição do ligante original da proteína 5I3Q. A opção de detectar cavidade foi escolhida para restringir a região de interesse a partes côncavas da superfície do sítio de interação (opção “**Detect cavity – restrict atom selection to solvent-accessible surface**”); ademais, ativou-se a opção para gerar um arquivo do sítio de interação selecionado.

A função de pontuação escolhida foi ChemPLP (Korb; Stützle; Exner, 2009b), com repontuação em ChemScore (Eldridge et al., 1997b), eficiência de busca de 100% e baixa velocidade, com maior acurácia. Foi autorizado o encerramento da docagem com menos de 10 resultados se as 3 melhores soluções apresentassem um RMSD dentro de uma faixa de 1,5 Å entre si, a fim de tornar o processo mais rápido, gerando menos poses repetidas.

As pontuações na função ChemPLP foram tabeladas e normalizadas por 3 critérios – dividindo-se pela massa molecular, área superficial e \sqrt{N} , onde N é o número de átomos que não são hidrogênios - para cada ligante para a pose melhor pontuada. Os ligantes que foram melhor pontuados em cada critério foram selecionados, 10 para cada critério. Nessa lógica, seriam esperados 30 compostos selecionados. No entanto, houve repetição entre os compostos selecionados: vários apareceram em mais de um critério, e um deles foi selecionado pelos três. Como resultado, o total final foi de 23 compostos. Desses, 16 são fragmentos e 7 são ligantes grandes.

Com o protocolo de docagem estabelecido, passamos à análise das interações dos ligantes com o sítio de ligação e ao uso das estratégias FBDD para o refinamento estrutural.

4.4. Análise das interações dos ligantes no sítio de ligação e FBDD

Os 17 fragmentos tiveram suas interações com o sítio avaliadas utilizando o servidor online PLIP (*Protein-Ligand Interaction Profiler*) (Adasme et al., 2021), sendo tabeladas suas interações hidrofóbicas (*van-der-Waals*) e ligações de hidrogênio. Para isso, a melhor solução da docagem de interesse foi exportada e salva como arquivo *.pdb*. A razoabilidade química também foi considerada no mapeamento das interações, observando individualmente cada arquivo gerado.

Foram procurados padrões de interação e os resíduos de aminoácidos mais importantes para o funcionamento da proteína foram estudados. Técnicas de FBDD foram empregadas para se planejar melhores ligantes. A modificação dos fragmentos foi desenhada e otimizada como citado em 4.1 e sua docagem foi feita como descrito em 4.3 apenas na estrutura 5I3Q, por ser apenas o domínio RdRp da NS5, diminuindo-se a exigência computacional. As séries de fragmentos modificados foram testados sempre nas mesmas condições de cálculo, a fim de fornecer comparação válida. Além do critério de pontuação, as interações adicionais decorrentes das modificações estruturais foram analisadas no servidor PLIP e no programa *Discovery Studio Visualizer* (Biovia, 2024).

O planejamento das modificações foi realizado em algumas etapas: primeiro, os resíduos de aminoácidos em torno do ligante foram identificados e analisados, pretendendo entender o perfil do bolsão (hidrofóbico ou polar); sabendo o perfil, resíduos chave já detectados por estudos de mutação (Lim et al., 2016) foram o foco do planejamento, objetivando as interações mais fortes possíveis com esses resíduos e seu consequente comprometimento. Séries homólogas foram geradas para compreender como as substituições afetavam a pose e interação dos ligantes. Por fim, a partir das poses melhor pontuadas, mais modificações foram propostas, visando mais interações ou interações mais fortes.

Para validar e aprofundar a análise das interações previstas, realizamos ainda a otimização das geometrias e o cálculo da entalpia de interação, conforme descrito a seguir.

4.5. Otimização das geometrias previstas e cálculo da entalpia de interação

A entalpia de interação (ΔH_{int}) entre o ligante e a proteína foi estimada pela diferença entre a entalpia de formação do complexo otimizado e a soma das entalpias de formação das estruturas isoladas, calculadas com o programa Mopac2016 (Equação 1).

$$\Delta H_{int} = \Delta H_f(\text{complexo}) - [\Delta H_f(\text{ligante}) + \Delta H_f(\text{proteína})] \quad (\text{Equação 1})$$

Concomitantemente ao descrito em 4.4, iniciamos os cálculos da entalpia de interação entre os ligantes selecionados e a enzima.

Para o cálculo no Mopac2016, o estado de ionização foi atribuído assumindo-se estados ionizados aos grupos das extremidades N e C terminais e às cadeias laterais dos resíduos de Asp, Glu, Arg e Lys, e garantindo-se que as ligações dissulfeto estivessem corretamente atribuídas, assim como as cargas dos íons metálicos das estruturas e dos resíduos que os coordenam. Para todas essas verificações, a palavra-chave “**charges**” foi utilizada. Ajustes necessários foram realizados no GOLD ou por edição do arquivo em um editor de textos.

Acertada essa etapa, a otimização foi conduzida em etapa única. Ao redor do ligante, em um raio de 6 Å, todos os átomos foram liberados para otimização de coordenadas usando-se o hamiltoniano PM7. O ligante foi otimizado inteiramente também utilizando-se o PM7. A escolha de otimizar as coordenadas dos átomos em um raio ao redor do ligante se deu pela exigência computacional dos cálculos. Também para reduzir o custo computacional durante a otimização dos complexos e proteínas, foi utilizada a palavra-chave “*mozyme*”, que substitui o procedimento padrão de cálculo SCF por um método baseado em orbitais moleculares localizados (LMOs). Esse método foi desenvolvido para possibilitar o tratamento eficiente de sistemas volumosos, permitindo que o tempo de cálculo só aumente de forma aproximadamente linear com o tamanho do sistema.

Para simular parcialmente os efeitos do solvente, foi utilizado o modelo de contínuo COSMO (*Conductor-like Screening Model*) no MOPAC. Esse modelo representa o solvente como um meio contínuo com uma constante dielétrica adequada, que afeta a distribuição eletrônica das moléculas inseridas.

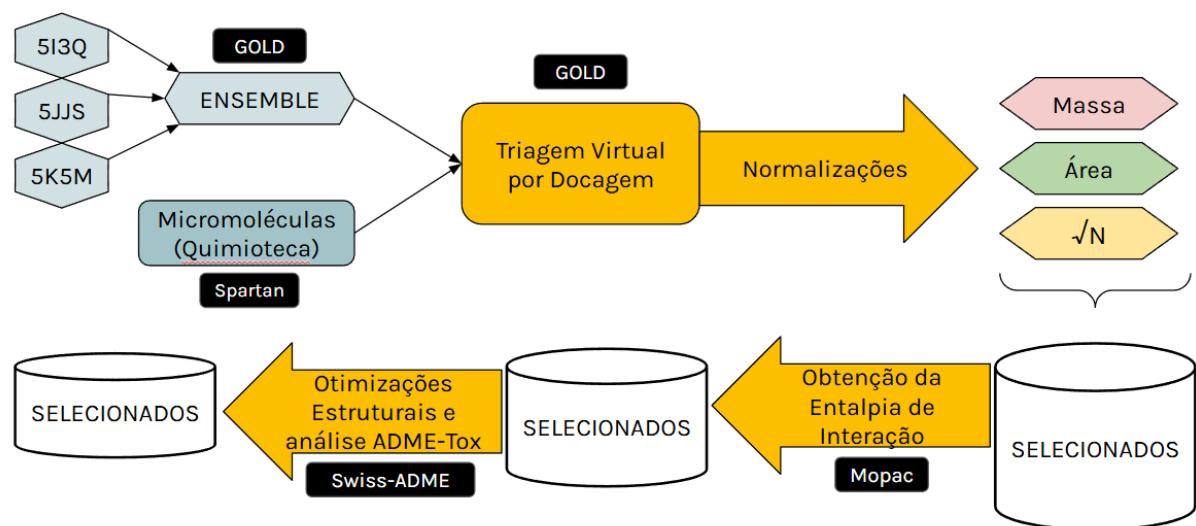
A constante dielétrica da água à temperatura ambiente (298K) é 78,4, valor adotado nos cálculos usando a palavra-chave “**EPS=78.4**”. Isso permite incluir parcialmente o efeito do solvente sobre a energia e estabilidade do sistema, uma vez que a polaridade do meio pode influenciar a interação entre o ligante e a macromolécula.

Concluídas as etapas de triagem e análise estrutural, os dados obtidos permitiram a seleção de compostos promissores, apresentados na seção de resultados.

4.6. Análises de absorção, distribuição, metabolismo, excreção e toxicidade:

Todos os compostos selecionados após as normalizações das pontuações obtidas pela triagem virtual foram submetidos à predição de propriedades ADME-Tox pelo servidor online Swiss-ADME (Daina et al., 2017).

Um fluxograma geral resumindo todas as etapas da metodologia é apresentado no Esquema 2.



Esquema 2: Fluxograma geral da metodologia utilizada.

5. Resultados e discussão

5.1. Escolha da função de docagem

No processo de redocagem, cada inibidor cocristalizado foi excluído computacionalmente da sua estrutura de origem e teve sua pose então definida pelo programa de docagem molecular. A Tabela 1 sumariza os resultados obtidos para as 5 estruturas cristalográficas oriundas do trabalho de Noble e colaboradores, 2016.

Tabela 1: RMSD da pose melhor pontuada na redocagem dos inibidores cocristalizados nas estruturas selecionadas para o sítio N¹.

PDB (Resolução, Å)	ASP	ChemPLP	RMSD (Å)	
			Função de pontuação	ChemScore
<u>5I3P (2,45)</u>	2,1476*	1,6209	1,6020*	1,4746*
<u>5JJR (1,99)</u>	2,5824*	2,2498*	2,4890*	1,6639*
5I3Q (1,88)	1,7190*	0,4633*	2,4563*	2,1352*
5K5M (2,04)	3,1685*	1,7915*	2,4415	2,8867*
5JJS (1,65)	0,3425	0,3886	0,7971	0,4305*
MÉDIA	1,9920	1,3028	1,9572	1,7182

¹Resultados em que o melhor RMSD não foi o da pose melhor pontuada são marcados com *. Estruturas excluídas por não terem bons resultados estão sublinhadas. Valores considerados bons estão em negrito.

O primeiro processo a partir dos resultados foi a exclusão de duas estruturas (5I3P e 5JJR) porque as funções de pontuação avaliadas tiveram pior desempenho nos experimentos de redocagem com elas, de forma que prosseguir com seu uso poderia levar à imprecisão nos resultados independentemente da função de pontuação utilizada. A exclusão das duas estruturas não prejudica o planejamento do trabalho, fato que pode ser justificado analisando as informações do Quadro 2. As estruturas dos ligantes foram apresentadas no tópico 2.4.

Quadro 2: Códigos PDB das estruturas analisadas e sua descrição.

Código PDB	Descrição
5I3P	Polimerase de RNA dependente de RNA de DENV-3 ligada ao composto 27.
5JJR	Proteína não-estrutural 5 de DENV-3 ligada ao composto 29.
5I3Q	Polimerase de RNA dependente de RNA de DENV-3 ligada ao composto 29.
5K5M	Polimerase de RNA dependente de RNA de DENV-2 ligada ao composto 27.
5JJS	Proteína não-estrutural 5 de DENV-3 ligada ao composto 27.

A estrutura 5I3P é equivalente à estrutura 5I3Q, sendo alterados apenas os ligantes cocristalizados. Dessa forma, excluir 5I3P da seleção não é prejudicial, pois 5I3Q teve resultados

melhores nos experimentos, talvez devido à sua melhor resolução cristalográfica (1,88 Å). Quando observamos o caso da estrutura 5JJR, ela é equivalente à 5JJS, mas esta tem melhor resolução cristalográfica (1,65 Å). O mesmo cenário se repete e excluímos com segurança duas estruturas para as quais as funções de pontuação apresentaram pior desempenho.

Observando os resultados, a exclusão da estrutura 5K5M também seria esperada. Entretanto, a exclusão da única estrutura relativa ao sorotipo 2 restringiria as análises apenas ao sorotipo 3, reduzindo a cobertura teórica. Além disso, a supracitada é a única de linhagem não-mutante, sendo essa mais uma razão para não a excluir. A fim de otimizar o trabalho e explorar ao mesmo tempo os modos de interação com as 3 estruturas remanescentes, o procedimento de *ensemble docking* (docagem em conjunto) foi utilizado.

Nesse procedimento, todos os compostos são testados em um dado conjunto de estruturas disponíveis do alvo. A compatibilidade máxima é avaliada e o resultado fornecido é a maior pontuação possível dentre as estruturas selecionadas do alvo. As principais razões para a escolha dessa metodologia são: *i*) a possibilidade de avaliar diferentes conformações do mesmo sítio alostérico (ex.: diferenças entre 5JJS e 5I3Q); *ii*) intencionalmente distinguir as diferenças estruturais entre os sítios alostéricos das polimerases de dois sorotipos e suas consequências para as interações com os candidatos a ligantes; *iii*) obter em procedimento único as pontuações nas três estruturas, sendo destacada a melhor combinação.

A escolha da função ChemPLP se deu pelo seu melhor desempenho da redocagem, apresentando o menor valor médio de RMSD de todas as redocagens, sendo seguida por ChemScore, que foi escolhida para o procedimento de rescore.

Além disso, ChemPLP é recomendada para uso em triagens virtuais e docagens diversas em que a determinação da pose do ligante é o foco (Cambridge Crystallographic Data Centre, 2022). Já a função ChemScore contém componentes empíricos eficazes em demonstrar a intensidade de interações. Dessa forma, a pose foi determinada pela ChemPLP e teve a pontuação recalculada pela ChemScore.

5.2. Ligantes selecionados por triagem virtual

Cabe ressaltar, inicialmente, alguns dados sobre a diversidade da quimioteca empregada no trabalho, ainda em construção. Nem todos os descritores de espaço químico foram determinados, entretanto, algumas informações relevantes são: dos 655 ligantes bem sucedidos na docagem, temos moléculas de 130 Da até 902 Da. As áreas superficiais variam de 151,93 Å² até 869,60 Å². Os trabalhos da pós-graduação em Química dos quais as moléculas foram coletadas, datam de 1974 até 2022. 12 professores-orientadores do PPGQ são citados, sendo eles: Dr. Anselmo Alpande Morais; Dr. Arthur E. Kümmerle; Dra. Áurea Echevarria Aznar Neves Lima; Dra. Ceres Maria Andrade da Mata; Dra. Cristiane Martins Cardoso De Salles; Dr. João Batista Neves da Costa; Dra. Leila Vilela Alegrio; Dr. Marco Edilson Freire De Lima; Dr. Mário Geraldo de Carvalho; Dr. Otto Richard Gottlieb; Dr. Raimundo Braz Filho e Dra. Rosane Nora Castro. As linhas de pesquisa dos supracitados pesquisadores incluem produtos naturais, semissintéticos e sintéticos.

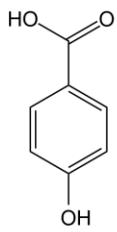
Para a seleção dos ligantes melhor pontuados, três critérios foram utilizados. Uma das limitações da triagem virtual por docagem é a pontuação enviesada – quanto maior o ligante, mais interações faz com o sítio, maior a pontuação. Isso é um problema do ponto de vista da Química Medicinal, porque há um limite máximo para a massa molecular associado à biodisponibilidade oral dos candidatos a fármacos; considerando a regra dos 5 de Lipinski, por exemplo, esse limite seria 500 Da. Para reduzir esse viés, utilizamos normalizações dividindo as pontuações por quantidades adequadas. Na literatura, há referência à normalização por divisão por \sqrt{N} , em que N é o número de átomos do ligante diferentes de hidrogênios (Ibrahim; Bauer; Boeckler, 2015). A normalização por divisão por massa ou área superficial pode potencialmente criar outro viés – de pontuar melhor moléculas menores. Cientes das limitações de cada normalização, decidiu-se utilizar as três. Por massa e área foram selecionados fragmentos (moléculas de até 300 Da), que foram otimizados por estratégias de FBDD. Por \sqrt{N} algumas moléculas de maior massa molecular foram selecionadas, na perspectiva do SBDD. A Tabela 2 apresenta os ligantes selecionados, suas referências e pontuação em cada critério de normalização.

Os resultados da triagem virtual foram tabelados pelos seguintes parâmetros: pontuação (ChemPLP), proteína em que atingiu essa pontuação, fórmula molecular do ligante, massa molecular, área superficial e número de átomos diferentes de hidrogênio. As estruturas dos ligantes selecionados são mostradas na Figura 9.

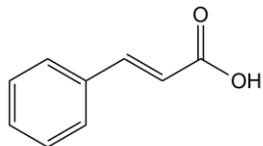
Tabela 2: Dados dos ligantes docados selecionados. Colocações de 1 a 10 foram ressaltadas em negrito.

Código	Colocação			Referência de origem	Nome
	Massa	Área	\sqrt{N}		
MPPGQ2004R0105	3°	4°	14°		Ácido <i>p</i> -cumárico
MPPGQ2004R0106	24°	10°	27°		Ácido ferúlico
MPPGQ2004R0108	8°	6°	49°	Silva, da; Ricardo Figueira, 2004	Ácido <i>o</i> -cumárico
MPPGQ2004R0109	1°	3°	12°		Ácido cinâmico
MPPGQ2004R0110	26°	1°	87°		Ácido 2-metoxi cinâmico
PPGQ2010R0101	4°	71°	117°	Souza, de; Marco André Alves, 2010	3-octanol
PPGQ2010R0102	2°	43°	62°		(<i>Z</i>)-ocimeno
PPGQ201820209	22°	7°	23°	Drumond, Natália Lopes, 2018.	Ácido <i>p</i> -nitro-cinâmico
PPGQ201820210	5°	5°	18°		Ácido <i>p</i> -flúor-cinâmico
PPGQ2016R0201	6°	2°	8°	Salgueiro, Fernanda Barbosa, 2016	Ácido cafeico
PPGQ2016R0211	27°	16°	4°		Drupanina
MPPGQ2004R0302	7°	8°	79°	Liandra, Regina Lúcia Pelachim, 2004.	Ácido 4-metoxicinâmico
MPPGQ2004R0303	23°	9°	316°		Ácido <i>p</i> -hidroxibenzóico
MPPGQ1998R0105	10°	133°	249°	Alvez, Luciana Vignólio, 1998.	Mirceno
MPPGQ1998R0108	9°	45°	244°		<i>p</i> -cimeno
PPGQ200070110	116°	187°	1°		<i>(1E,4Z,6E)-1-(4-((etiniloxi)metil)-3-metoxifenil)-5-hidroxi-7-(3-metoxi-4-(prop-2-in-1-iloxi)fenil)hepta-1,4,6-trien-3-ona</i>
PPGQ200070111	134°	168°	6°	Castro, de; Denise, 2000.	[1,7- bis -(3, 4- dimetoxifenil)] -heptano - 3,5- diona
PPGQ200070130	101°	119°	9°		<i>(1E,4Z,6E)-5-hidroxi-1,7-bis(4-hidroxifenil)hepta-1,4,6-trien-3-ona</i>
PPGQ201920105	104°	111°	3°		<i>(E,E)-N1,N4-bis(4-clorobenzilideno)butano-1,4-diamina</i>
PPGQ201920108	44°	100°	5°	Rodrigues, Arthur Valbon, 2019.	<i>(E,E)-N1,N6-dibenzilidenohexano-1,6-diamina</i>
PPGQ201920109	52°	148°	2°		<i>(E,E)-N1,N8-dibenzilidenoctano-1,8-diamina</i>
PPGQ201430113	212°	395°	7°	Ferreira, Welisson da Silva, 2014.	<i>2-((1E,3E)-4-(benzo[d][1,3]dioxol-5-il)buta-1,3-dien-1-il)-5-(dodeciltio)-1,3,4-oxadiazol</i>
MPPGQ2003M0103	100°	67°	10°	Nascimento, do; Ildomar Alves, 2003.	<i>(E)-3-((2S,3S)-2-(4-hidroxi-3-metoxifenil)-3-(hidroximetil)-7-metoxi-2,3-di-hidrobenzofuran-5-il)acrilaldeído</i>

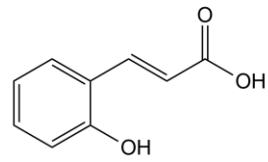
ÁCIDO *PARA*-HIDROXI-BENZÓICO
MPPGQ2004R0303



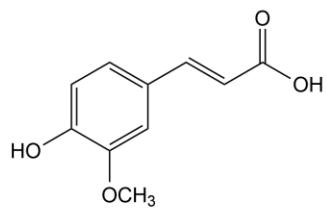
ÁCIDO CINÂMICO
MPPGQ2004R0109



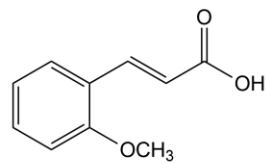
ÁCIDO *ORTO*-CUMÁRICO
MPPGQ2004R0108



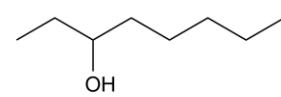
ÁCIDO FERÚLICO
MPPGQ2004R0106



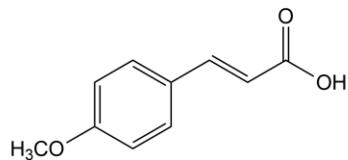
ÁCIDO 2-METOXI-CINÂMICO
MPPGQ2004R0110



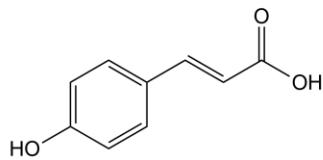
3-OCTANOL
PPGQ2010R0101



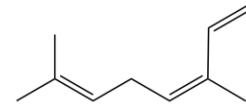
ÁCIDO 4-METOXI-CINÂMICO
MPPGQ2004R0302



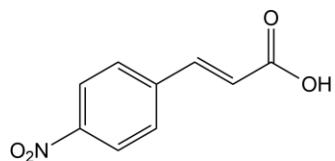
ÁCIDO *PARA*-CUMÁRICO
MPPGQ2004R0105



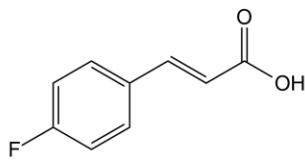
(*Z*)-OCIMENO
PPGQ2010R0102



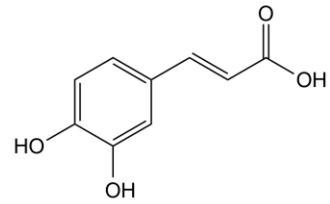
ÁCIDO *PARA*-NITRO-CINÂMICO
PPGQ20182029



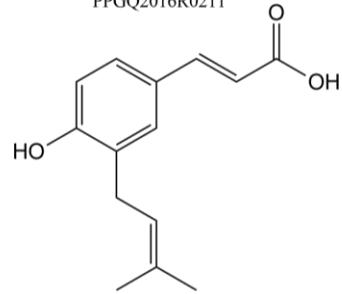
ÁCIDO *PARA*-FLUOR-CINÂMICO
PPGQ201820210



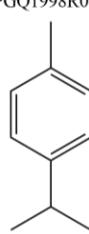
ÁCIDO CAFEICO
PPGQ2016R0201



DRUPANINA
PPGQ2016R0211



PARA-CIMENO
PPGQ1998R0108



MIRCENO
PPGQ1998R0105

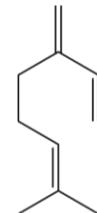
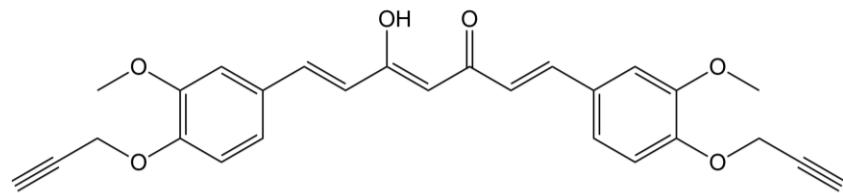
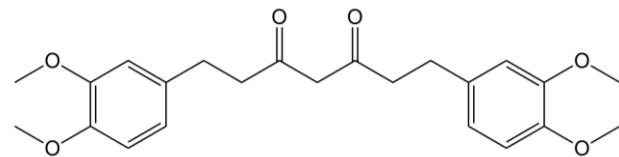


Figura 10: Estruturas dos ligantes selecionados, seus nomes e códigos na quimioteca do PPGQ-UFRRJ.

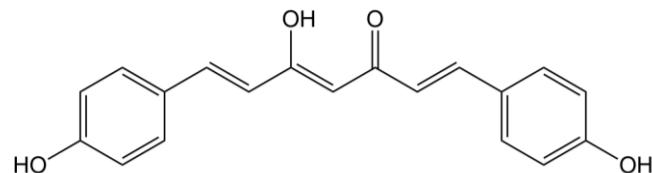
(1E,4Z,6E)-1-(4-((etiniloxi)methyl)-3-metoxifenil)-5-hidroxi-7-(3-metoxi-4-(prop-2-in-1-oxi)fenil)hepta-1,4,6-trien-3-ona
PPGQ200070110



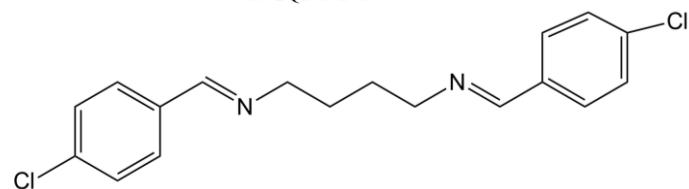
1,7-bis(3,4-dimetoxifenil)heptano-3,5-diona
PPGQ200070111



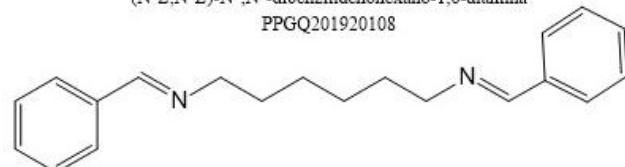
(1E,4Z,6E)-5-hidroxi-1,7-bis(4-hidroxifenil)hepta-1,4,6-trien-3-ona
PPGQ200070130



(N¹E,N⁴E)-N¹,N⁴-bis(4-chlorobenzildeno)butano-1,4-diamina
PPGQ201920105



(N¹Z,N⁶E)-N¹,N⁶-dibenzildidenohexano-1,6-diamina
PPGQ201920108



(N¹Z,N⁸E)-N¹,N⁸-dibenzildidenooctano-1,8-diamina
PPGQ201920109

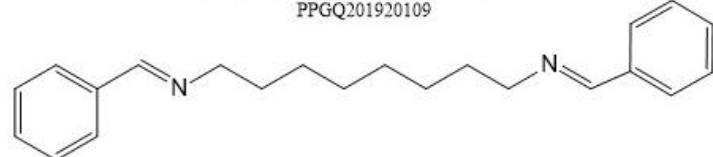


Figura 10 - continuação: Estruturas dos ligantes selecionados, seus nomes e códigos na quimioteca do PPGQ-UFRRJ.

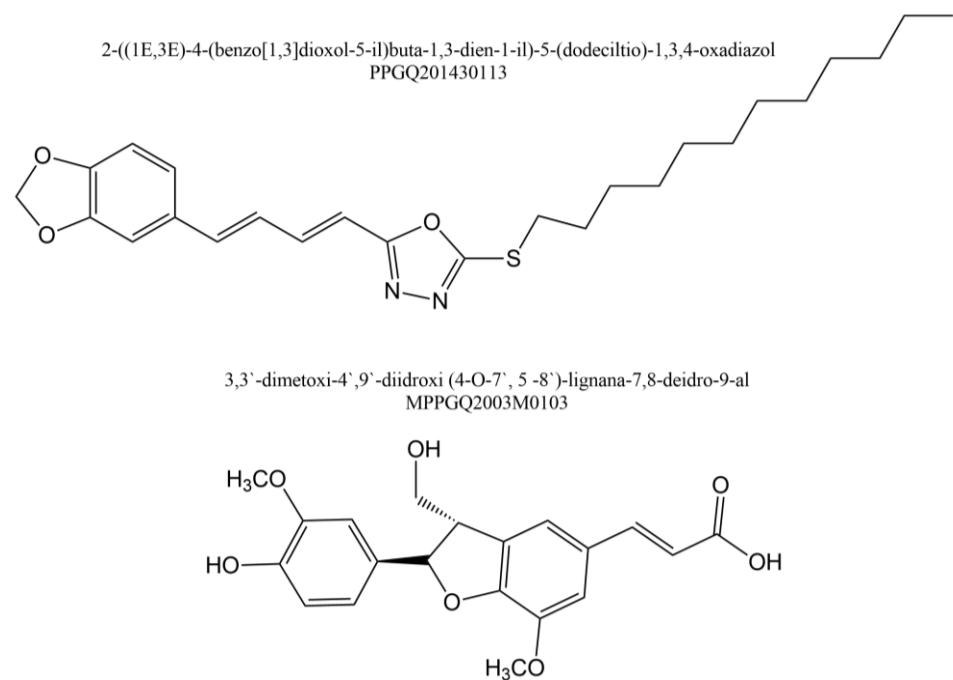


Figura 10 - continuação: Estruturas dos ligantes selecionados, seus nomes e códigos na quimioteca do PPGQ-UFRRJ.

5.3. Cálculo da entalpia de interação dos complexos proteína-ligante

A metodologia utilizada, em que a otimização das coordenadas atômicas foi permitida a partir de 6 Å dos átomos dos ligantes, levou a diferentes raios de otimização completa, já que os ligantes têm diferentes tamanhos. Para calcular a entalpia de interação é necessário calcular a diferença entre a entalpia de formação do complexo (que foi otimizado como descrito anteriormente) e a soma das entalpias de formação do ligante e proteína separados. Aqui nos deparamos com um ponto de reflexão: o ligante, tanto inserido no complexo como otimizado separadamente foi totalmente otimizado, mas a proteína está sendo apenas parcialmente otimizada.

No entanto, é esperado que o efeito conformacional causado pelo ligante na proteína seja maior nas proximidades do sítio de interação, havendo pouca ou nenhuma mudança conformacional nas maiores distâncias. Apesar disso, trabalhar com a estrutura completa da proteína traz efeitos significativos na estabilidade das interações proteína-ligante e, portanto, na entalpia de interação. Isso se dá pela redistribuição de carga entre os átomos do sistema após a minimização de energia eletrônica pelo modelo quântico semi-empírico.

Para simplificar o cálculo, decidiu-se determinar a entalpia de interação considerando apenas a inserção do ligante no interior da cavidade de interação em todos os casos. Ou seja, para a proteína com sítio vazio manteve-se a mesma conformação resultante da inserção do ligante e minimização da energia do complexo. Assim, foi feito o cálculo da entalpia de formação da proteína vazia a partir do complexo otimizado, retirando-se o ligante e fazendo o cálculo por *single point*. Nessa otimização, calcula-se a energia do sistema com todas as coordenadas atômicas fixas, portanto idênticas às coordenadas do arquivo de entrada.

O inibidor **27** teve melhor entalpia de interação, mostrando que nenhum dos compostos selecionados foi capaz de formar interações mais fortes que ele. Em segundo lugar, temos a drupanina (PPGQ2016R0211) e em terceiro lugar o inibidor **29**. Esse resultado demonstra que a drupanina tem grande potencial de estabelecer interação com o sítio alostérico da polimerase do DENV-3.

Observando tendências gerais, com piores entalpias de interação estão hidrocarbonetos, o que era esperado, já que não são capazes de estabelecer interações fortes. A entalpia, da maneira que foi calculada, prevê que os compostos em seu estado inicial estão imersos em um dielétrico com propriedades semelhantes à da água e, para a interação, devem ser dessolvatados para interagir com a proteína.

Dentre os derivados do ácido cinâmico, a pior entalpia de interação é a do ligante PPGQ2004R0108, correspondente ao ácido *o*-cumárico, o que pode indicar que substituintes polares na posição *ortho* não são tão favoráveis quanto em outras posições, como *para*.

Considerando os compostos com melhor entalpia de interação, após a drupanina, está o ligante grande MPPGQ2003M0103, que foi capaz de estabelecer maior número de interações fortes no sítio.

Tabela 3: Entalpias de interação dos ligantes selecionados pela triagem virtual

Código do Ligante	Entalpia de Interação (kcal/mol)
MPPGQ1998R0105	-27,54
MPPGQ1998R0108	-23,32
MPPGQ2004R0105	-54,36
MPPGQ2004R0106	-56,37
MPPGQ2004R0108	-43,03
MPPGQ2004R0109	-56,30
MPPGQ2004R0110	-54,47
MPPGQ2004R0302	-49,34
MPPGQ2004R0303	-25,87
PPGQ2010R0101	-26,16
PPGQ2010R0102	-25,15
PPGQ2016R0201	-59,61
PPGQ2016R0211	-79,32
PPGQ201820209	-56,28
PPGQ201820210	-54,74
PPGQ201920105	-54,01
PPGQ201920108	-52,23
PPGQ201920109	-57,11
MPPGQ2003M0103	-70,98
PPGQ200070110	-61,18
PPGQ200070111	-49,52
PPGQ200070130	-60,20
PPGQ201430113	-69,33
INIBIDOR 29	-74,51
INIBIDOR 27	-81,38

5.4. Análise das interações de cada ligante selecionado

As geometrias de interação das poses identificadas pela docagem molecular foram otimizadas por minimização de energia com o método PM7, conforme descrito na Metodologia. Nesse item, serão analisadas as principais interações observadas após essa otimização. Tomando como referência os compostos **27** e **29** (Lim et al., 2016), nossa análise será comparativa com as interações desses, que são comprovadamente inibidores pan-sorotípico. As estruturas 2D desses inibidores são apresentadas na Figura 5. As suas poses no sítio alostérico da polimerase são apresentadas nas Figuras 11 e 12, respectivamente.

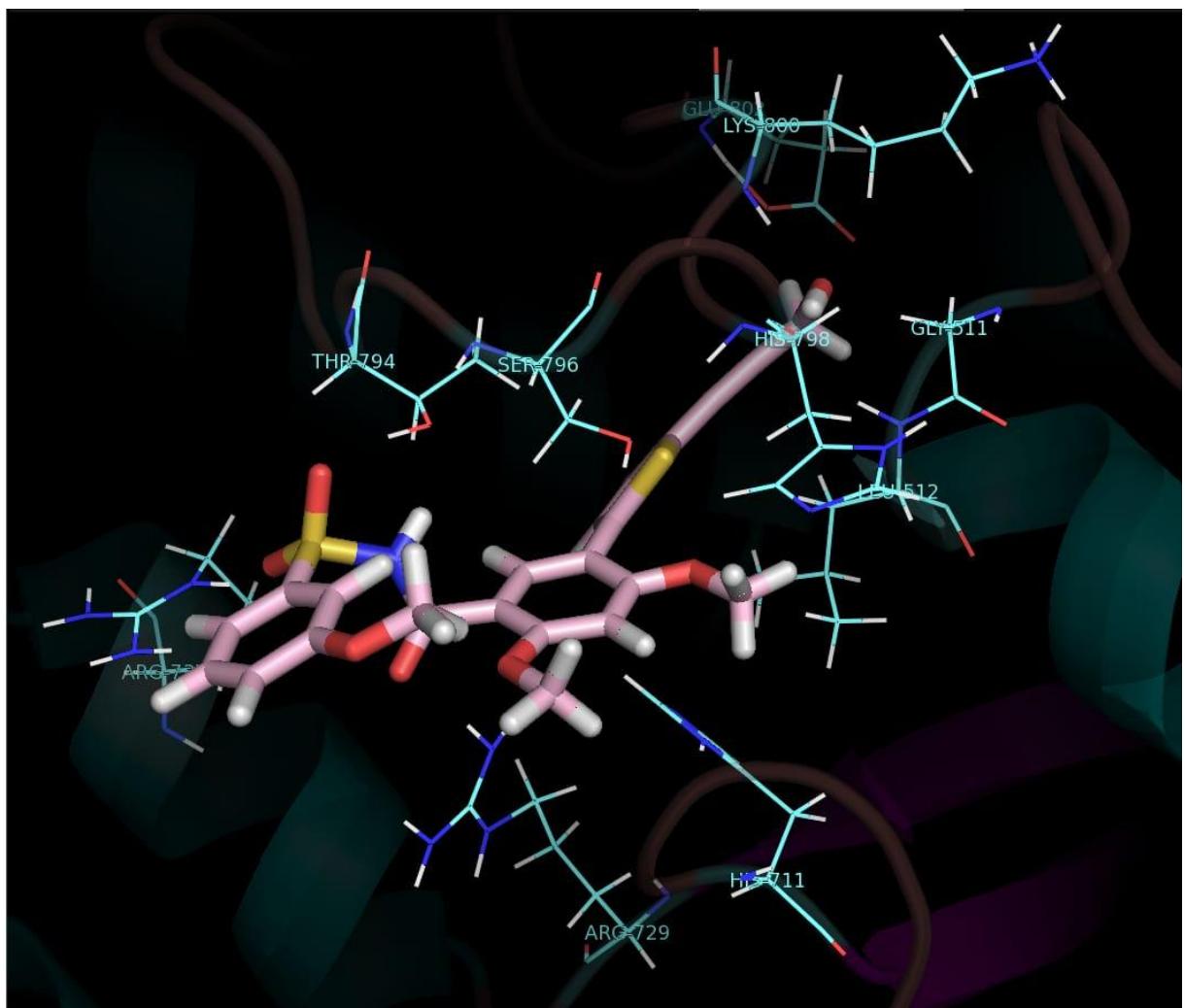


Figura 11: Composto 27 no sítio alostérico. (Estrutura 5K5M - DENV2). Ligantes com carbonos em rosa claro, resíduos de aminoácido com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6)

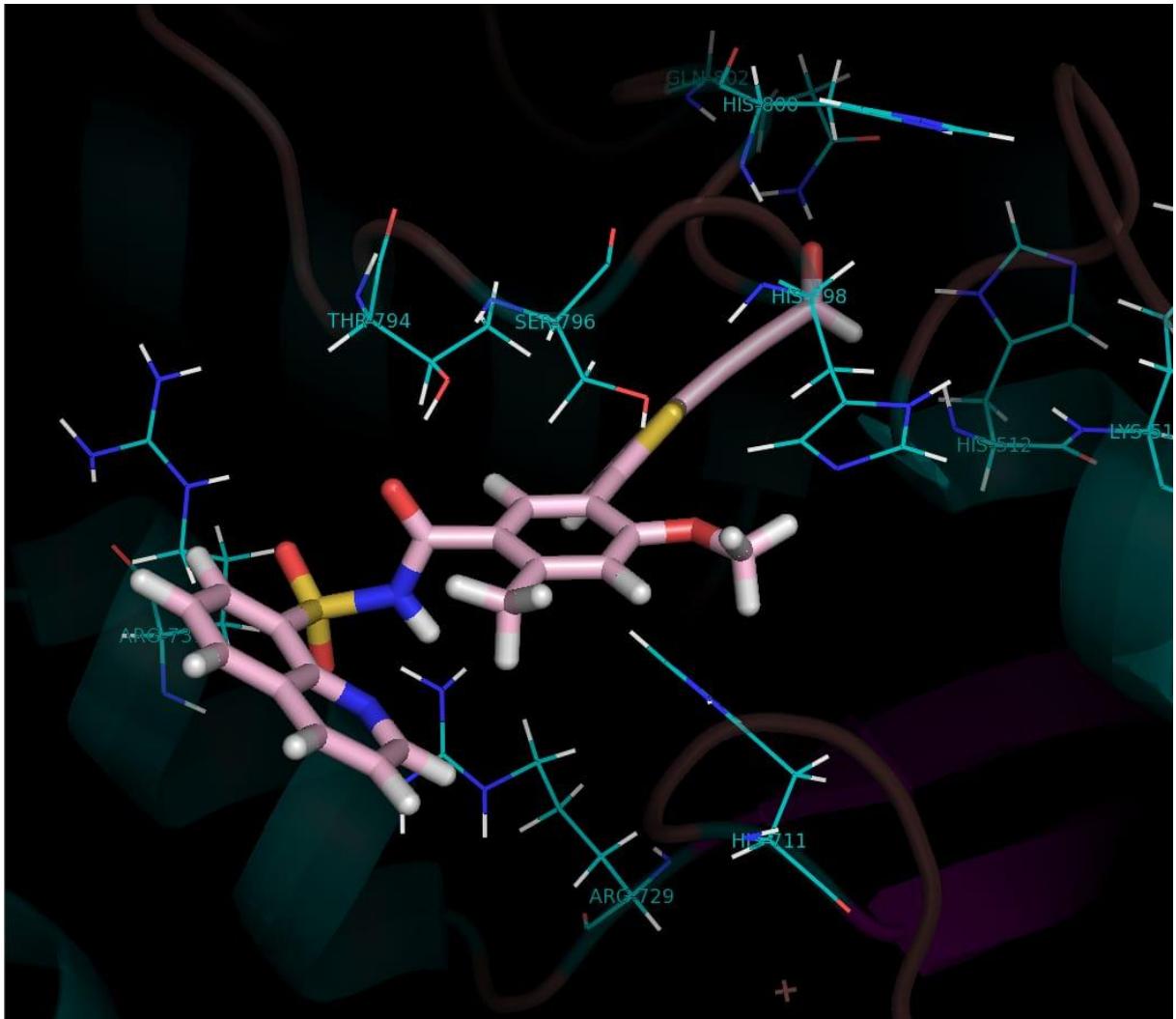


Figura 12: Composto 29 no sítio alostérico. (Estrutura 5I3Q - DENV3). Ligantes com carbonos em rosa claro, resíduos de aminoácido com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6)

Entre os sítios alostéricos de DENV-2 e DENV-3 poucas mudanças relevantes acontecem na sequência primária de aminoácidos. O comparativo dos aminoácidos do sítio que sofrem modificação de uma estrutura para a outra é apresentado no quadro 3.

Quadro 3: Resíduos diferentes entre as sequências primárias de DENV-2 (5K5M) e DENV-3 (5I3Q)

DENV-2	GLY 511; LEU 512; LYS 800; GLU 802;
DENV-3	LEU 511; HIS 512; HIS 800; GLN 802;

Essas modificações interferem em como um mesmo ligante pode interagir com os diferentes sítios. Por exemplo, o álcool propargílico, presente em ambos os ligantes, desempenha papel fundamental estabelecendo duas ligações de hidrogênio com resíduos do sítio (Lim et al., 2016). Enquanto na estrutura de DENV-2 essas ligações são com o esqueleto

polipeptídico da Lys800 e cadeia lateral da Gln802, na estrutura de DENV-3 ocorrem as mesmas ligações, porém, com o esqueleto polipeptídico da His800 e a cadeia lateral do Glu802. Um ponto de discussão interessante é que em pH=7,4 o ácido glutâmico está na forma de glutamato, permitindo uma interação íon-dipolo, enquanto na glutamina o mesmo não pode ocorrer.

Tendo conhecimento do sítio e disposição dos inibidores de referência, pode-se analisar cada ligante selecionado pela docagem.

5.4.1. Fragmentos selecionados da série MPPGQ2004R010

Todos os fragmentos selecionados oriundos da série MPPGQ2004R010 (Silva; Da, 2004) têm em comum a conservação do esqueleto do ácido cinâmico. Os ácidos *p*-cumárico, ferúlico, cinâmico e 2-metoxi-cinâmico foram selecionados para DENV-3, enquanto o ácido *o*-cumárico foi selecionado para DENV-2.

O carboxilato do ácido cinâmico, de acordo com a previsão da docagem, aponta para a mesma direção que o álcool propargílico do inibidor (Figura 13). Provavelmente, o fato do ácido cinâmico ser plano, assim como o grupo 2-propargil-tiofenila do composto **29**, além da presença em comum de oxigênios, assegura essa possibilidade. Considerando contatos polares, enquanto o inibidor **29** está restrito a duas ligações de hidrogênio, por só ter um oxigênio nesse grupo, o ácido cinâmico pode estar envolvido em até 4 ligações de hidrogênio. As ligações de hidrogênio previstas pelo servidor PLIP são com os resíduos Leu511, His512, His800 e Gln802. As distâncias de ligação podem ser observadas nos anexos 1A, 1B, 1C, 1D e 1E.



Figura 13: Sobreposição do ácido cinâmico e inibidor **29** na estrutura 5I3Q otimizada com o método PM7. Inibidor com carbonos em rosa claro, ácido cinâmico com carbonos em verde, resíduos de aminoácido com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6)

O ácido *p*-cumárico (Figura 14) também tem ligações de hidrogênio previstas com Leu511, His512, His800 e Gln802, exatamente como o ácido cinâmico. A diferença está na presença de uma ligação de hidrogênio extra com o resíduo Thr794, pela hidroxila na posição 4 do anel. É oportuno avaliar que, pela grande proximidade entre o carboxilato do ligante e o grupo amino do esqueleto da His512, o programa PyMol considera que esses grupos estão ligados.

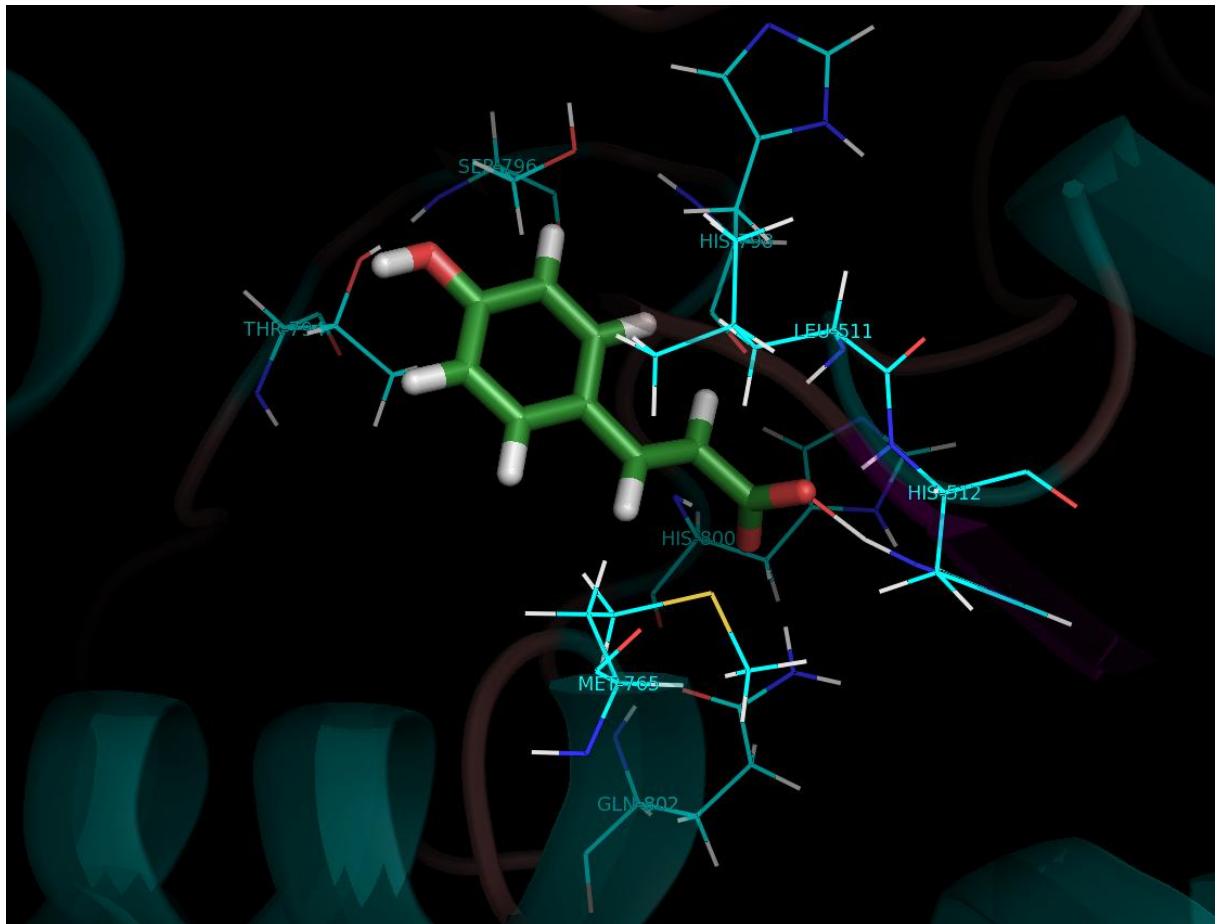


Figura 14: Previsão de pose do ácido *p*-cumárico, com carbonos em verde, resíduos de aminoácido com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6)

No caso do ácido Ferúlico (figura 15), são previstas pelo PLIP as mesmas ligações de hidrogênio a partir do carboxilato e a mesma ligação de hidrogênio pela hidroxila na posição 4. Com a adição do grupo metoxila na posição 3, é prevista uma ligação de hidrogênio com a Ser796.

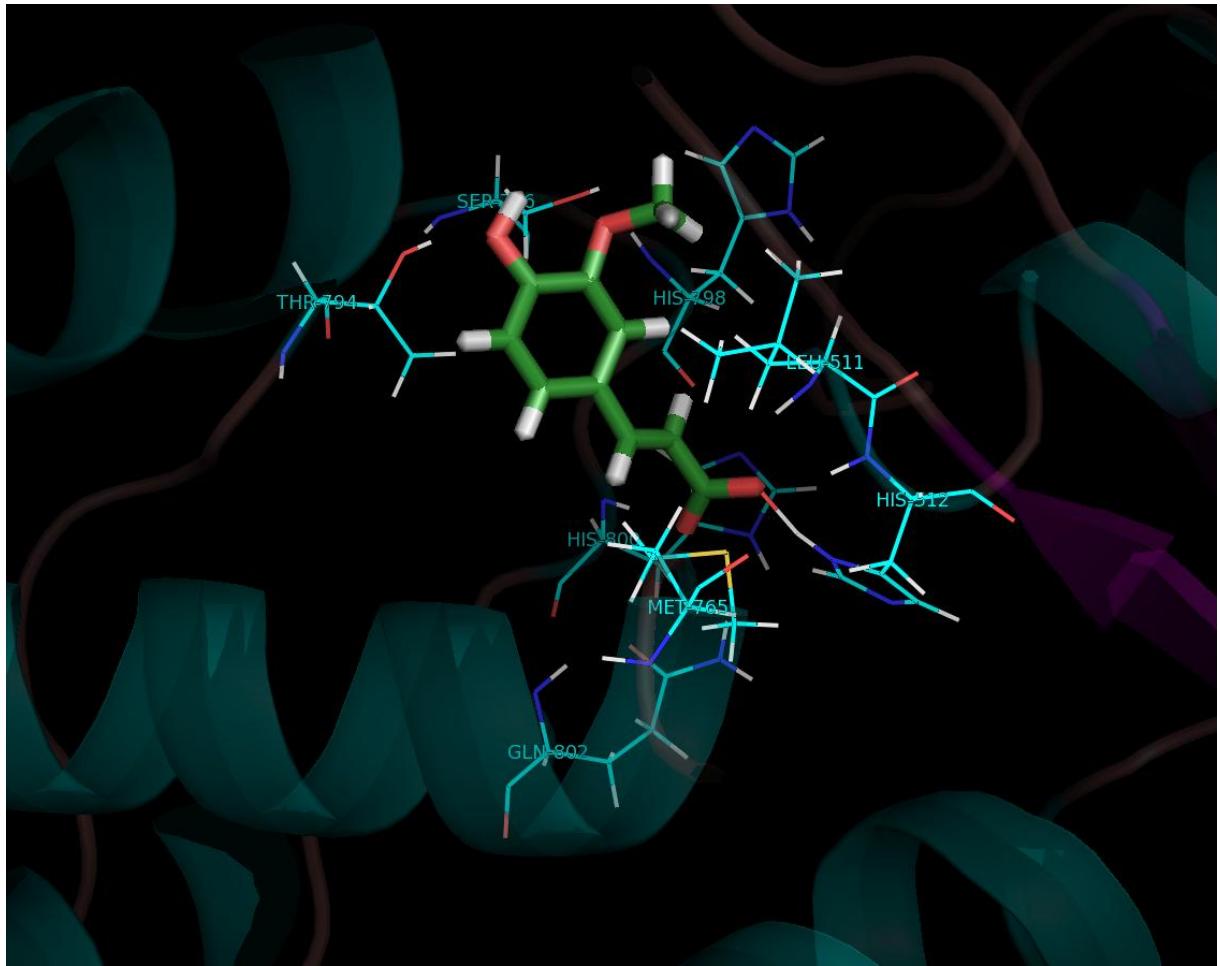


Figura 15: Previsão de pose do ácido ferúlico, com carbonos em verde, resíduos de aminoácido com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6)

Para o ácido 2-metoxi-cinâmico (Figura 16) são previstas as mesmas interações que para o ácido cinâmico, sendo todas com o grupo carboxilato do ligante. A substituição com o grupo metoxila na posição 2 não leva a interações fortes no DENV-3, diferentemente do que será discutido a seguir para o ácido *o*-cumárico, selecionado para DENV-2. O resíduo de Ser796 foi ressaltado, apesar disso, para comparar com a Figura 17.

Na solução para o ácido *o*-cumárico (Figura 17), selecionado para DENV-2, há uma observação a ser destacada: todos os ligantes derivados do ácido cinâmico até então apresentados conservam interação prevista com o resíduo de aminoácido Gln802, cuja cadeia lateral é uma amida. Ao observar o quadro de resíduos diferentes entre as estruturas, percebe-se que na estrutura 5K5M esse resíduo é um ácido glutâmico, que, em pH fisiológico, está na forma ionizada, ou seja, como glutamato.

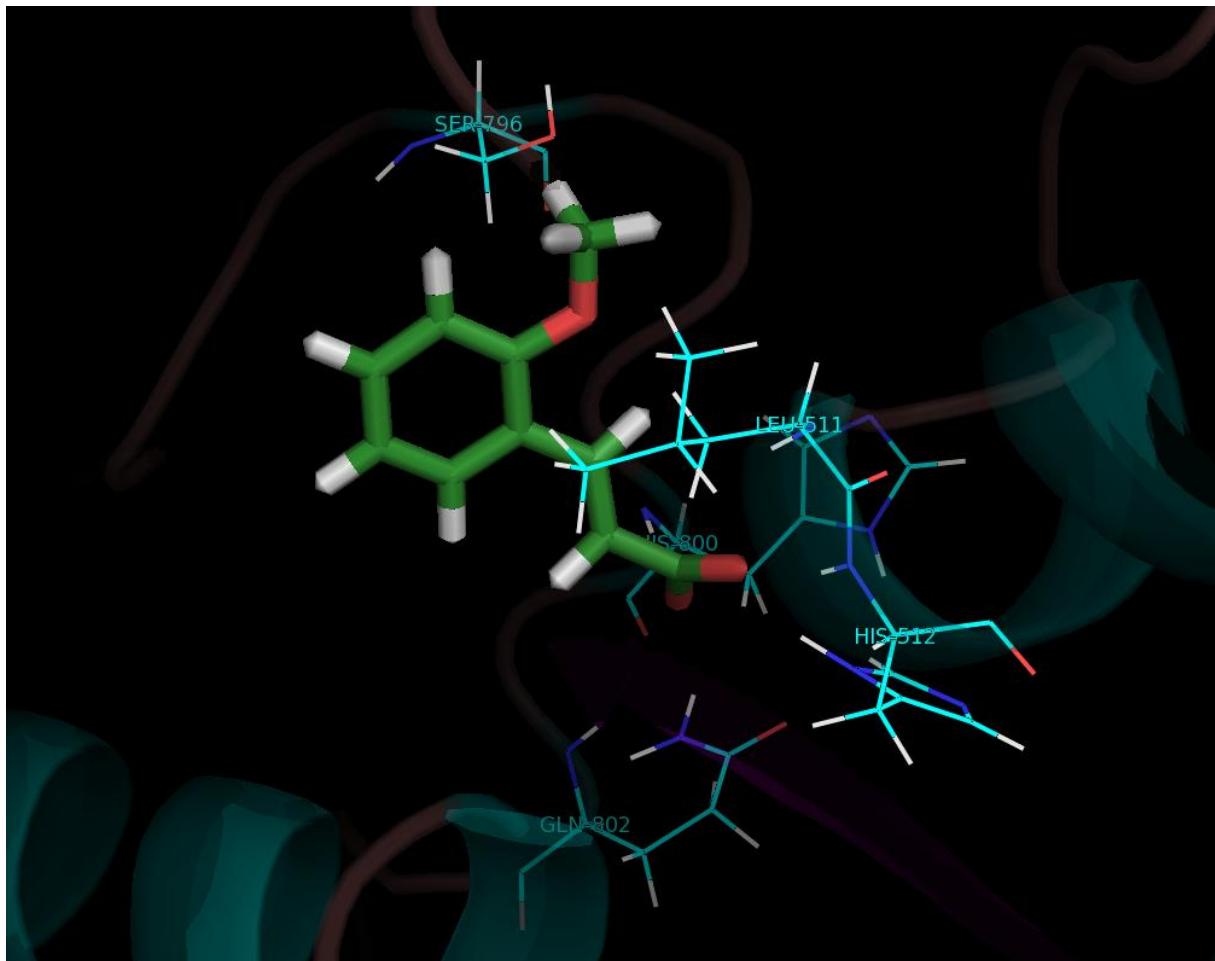


Figura 16: Previsão de pose do ácido 2-metoxicinâmico, com carbonos em verde, resíduos de aminoácido com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6)

O glutamato, por ser um ânion, não participa das mesmas interações que a glutamina. Além disso, pela carga negativa, é esperado que o ânion carboxilato do ligante seja repelido. Ao analisar a previsão de pose após otimização com o PM7, de fato há um deslocamento, entretanto, novas interações favoráveis são apresentadas. Para o ácido *o*-cumárico são previstas ligações de hidrogênio a partir do carboxilato com os resíduos Leu512, His513, His798 e Lys800. A mudança de His para Lys no caso do resíduo 800 não impacta a interação porque o grupo que participa da interação é o da cadeia principal e não da cadeia lateral.

O deslocamento causado pela repulsão resulta em ligações de hidrogênio envolvendo a hidroxila na posição 2 do anel com a Ser796. No caso do ácido 2-metoxy-cinâmico, não há interações fortes possíveis, pois a distância não é favorável. Sendo assim, os derivados do ácido cinâmico mostram bom ancoramento tanto na NS5 de DENV-2 como na NS5 de DENV-3, apesar das interações diferentes em cada sorotipo.



Figura 17: Previsão de pose do ácido *o*-cumárico, com carbonos em verde, resíduos de aminoácido com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6)

5.4.2. Fragmentos selecionados da série PPGQ2010R01

Dois compostos foram selecionados da série PPGQ2010R01 (Souza, 2010), não tendo relação estrutural entre si. Um desses compostos foi o 3-octanol, provavelmente selecionado por ter uma grande região hidrofóbica e um grupo polar adequados à interação com o sítio. As análises possíveis a partir do complexo são das áreas majoritariamente apolares do sítio alostérico.

Na figura 18, é possível visualizar a presença dos resíduos majoritariamente apolares Leu511, Ala799 e Trp803. O 3-octanol interagiu por forças de dispersão com esses resíduos e fez uma ligação de hidrogênio com o resíduo de Thr794. A percepção dos subsítios apolares é importante para o planejamento de posteriores otimizações estruturais. As distâncias de ligação podem ser observadas nos anexos 2A e 2B.

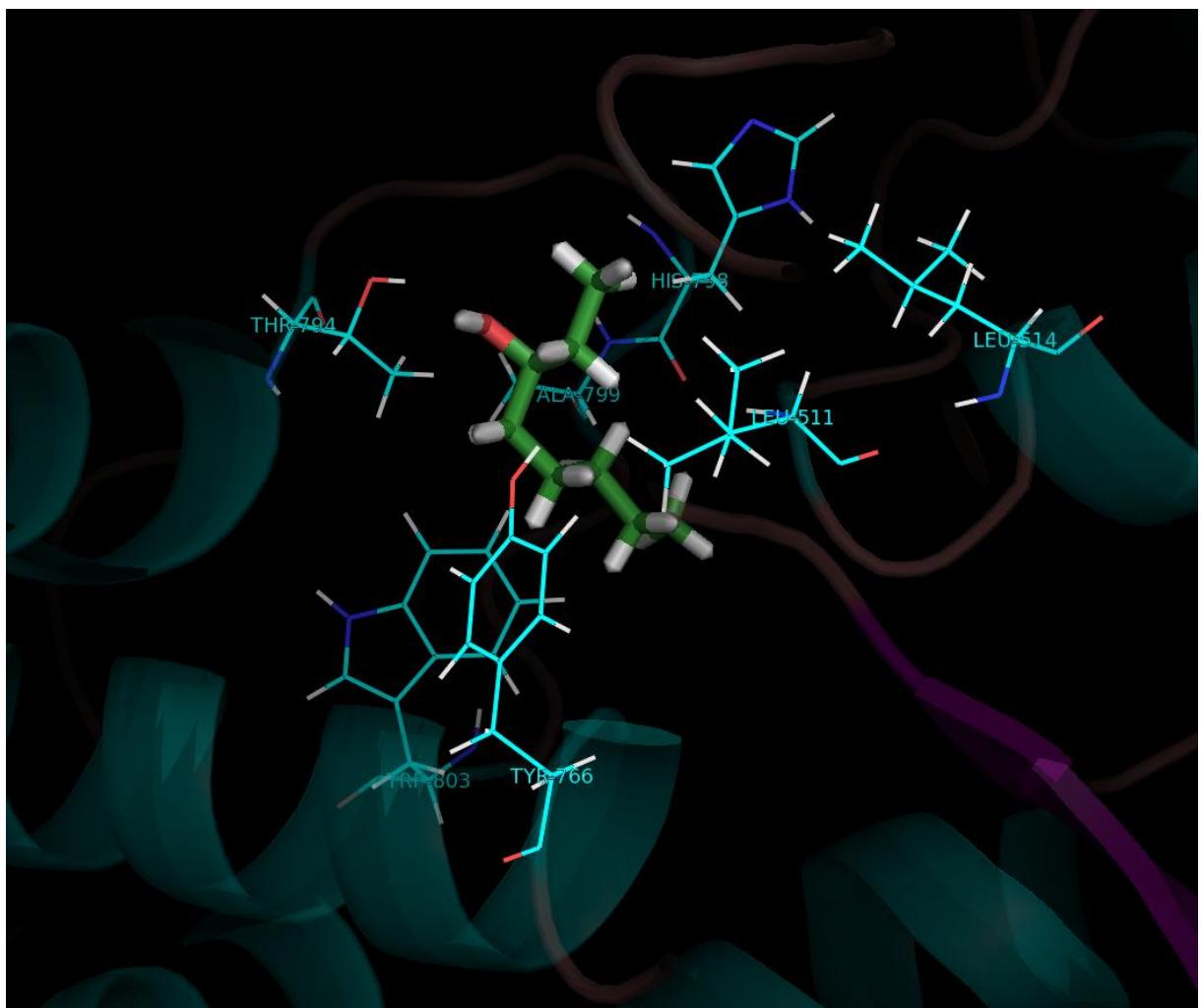


Figura 18: Previsão de pose do 3-octanol, com carbonos em verde, resíduos de aminoácido com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6)

Partindo para a análise do (*Z*)-ocimeno (Figura 19), um hidrocarboneto, vemos ancoramento também por forças de dispersão com os resíduos Leu511, Leu514, Tyr766, His798, Ala799 e Trp803.

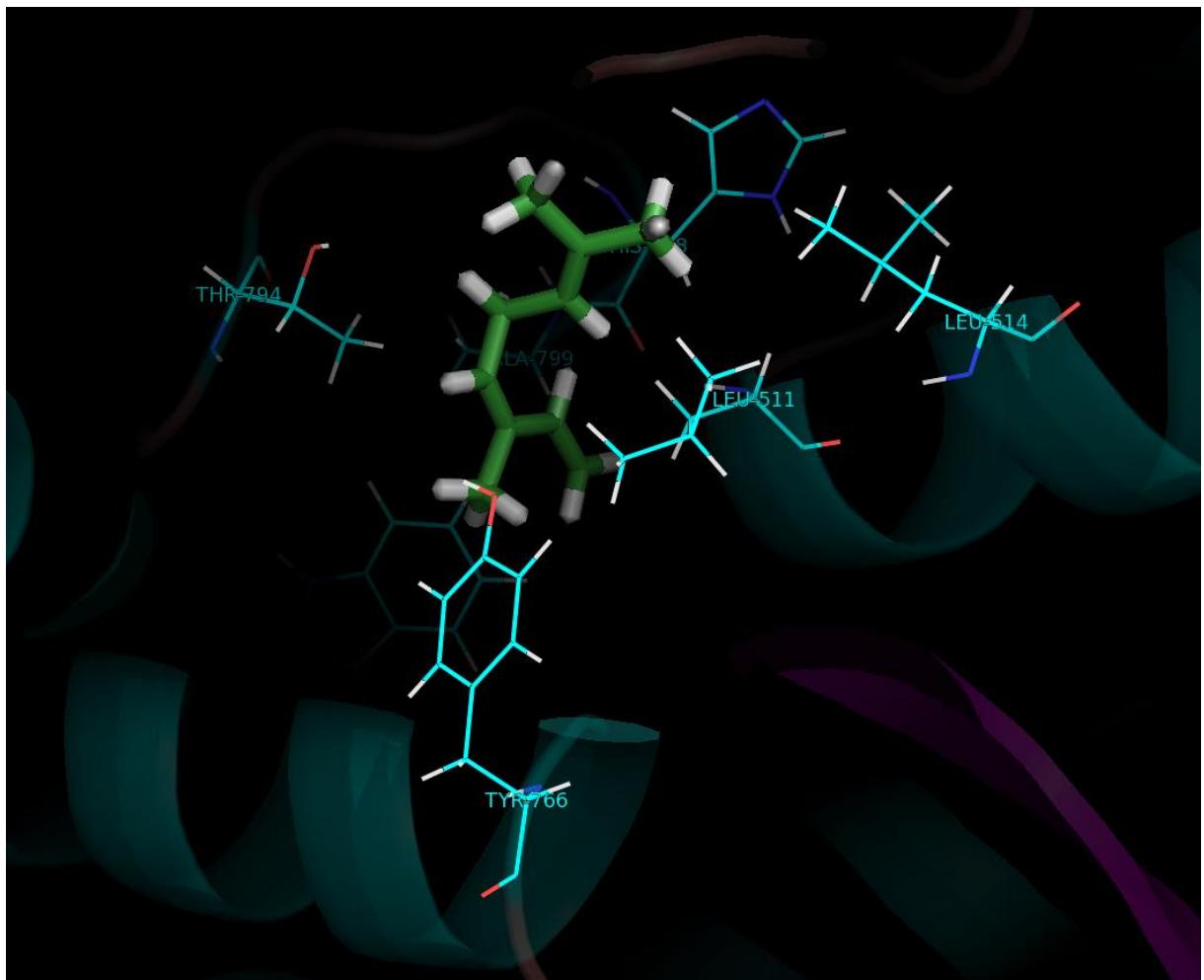


Figura 19: Previsão de pose do (*Z*)-ocimeno, com carbonos em verde, resíduos de aminoácido com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6)

À título de registro, ambos ligantes da série foram selecionados para DENV-3.

5.4.3. Fragmentos selecionados da série PPGQ2018202

Nessa série (Lopes, 2018), dois ligantes foram selecionados, ambos derivados do ácido cinâmico e selecionados para DENV-3.

Iniciando pelo ácido *p*-nitro-cinâmico (Figura 20), assim como para os ligantes da série MPPGQ2004R010, as ligações de hidrogênio em que o carboxilato se envolve são sustentadas pelos resíduos de aminoácidos Leu511, His512, His800 e Gln802. Além de estabilizações por interações fracas, não são registradas pelo servidor PLIP outras interações. O grupo nitro não demonstra estar realizando nenhuma interação com resíduos vizinhos. As distâncias de ligação podem ser observadas nos anexos 3A e 3B.

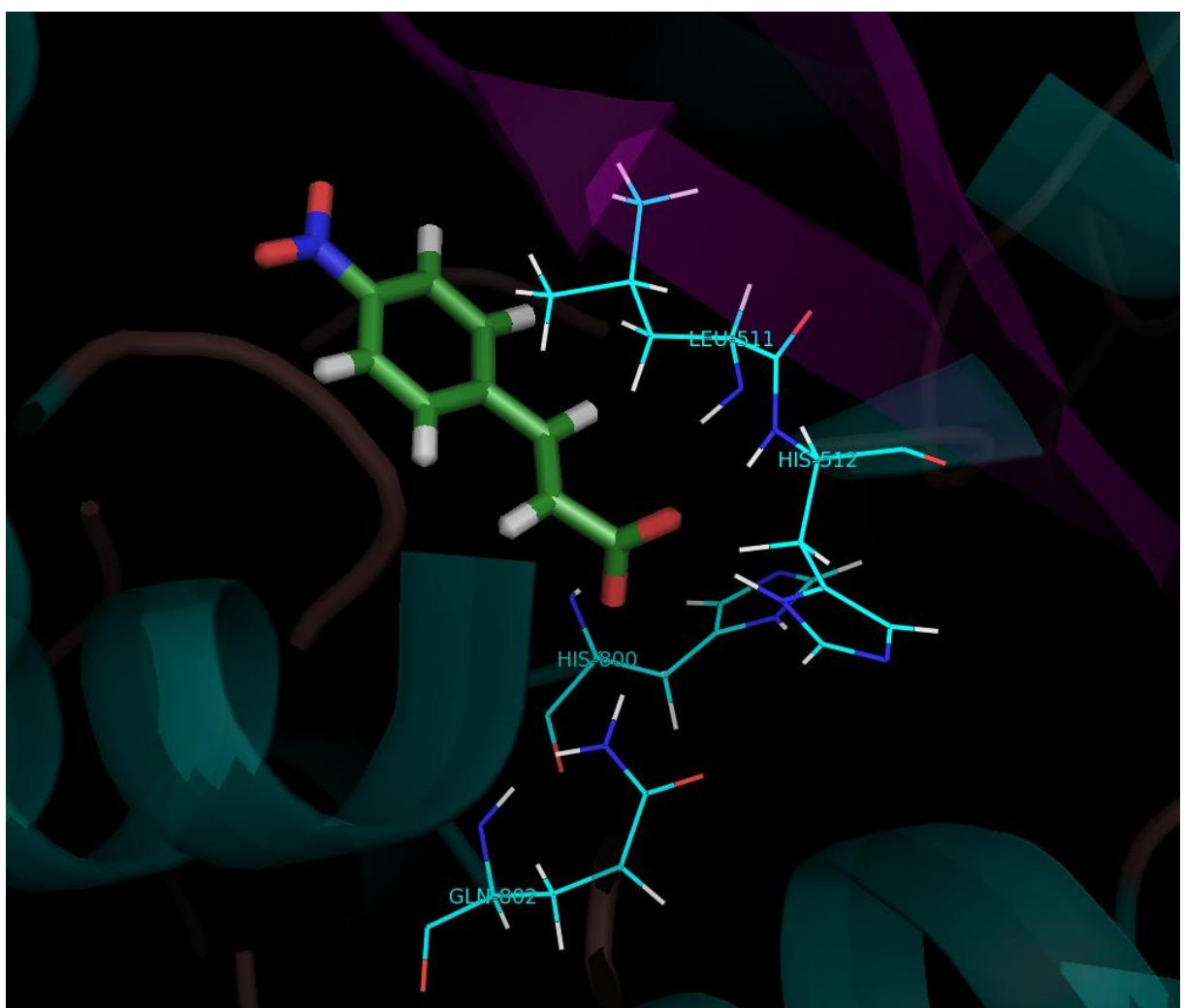


Figura 20: Previsão de pose do ácido *p*-nitrocinnâmico, com carbonos em verde, resíduos de aminoácido com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6)

Como a cavidade do sítio alostérico é bastante ampla, é possível que qualquer substituinte no lugar do grupo nitro também não realizasse interações complementares porque não há resíduos de aminoácidos próximos nessa região.

Uma discussão possível é se, para DENV-2, onde há o ânion glutamato, o grupo nitro poderia ser menos repelido que o carboxilato do ácido cinâmico, uma vez que o perfil polar se manteria, mas a carga total seria nula. Pelas soluções obtidas e pela cristalografia dos inibidores conhecidos, é necessário um espaçador hidrofóbico e planar entre a região mais volumosa do ligante e o grupo polar que irá interagir com os resíduos His800, Gln802 e vizinhos. Sendo assim, a substituição no anel não parece atender aos requisitos para alocação no pequeno túnel disponível. Uma proposta a ser futuramente testada é a substituição do grupo carboxilato do ácido cinâmico por um grupo nitro, verificando se é possível as ligações de hidrogênio se manterem, com a maximização do perfil de interação para DENV-2.

Observando o ácido *p*-flúor-cinâmico (Figura 21) vemos basicamente o mesmo perfil de interações. Não são reconhecidas interações intermoleculares envolvendo o flúor, sendo, aparentemente, o mesmo caso do substituinte nitro.

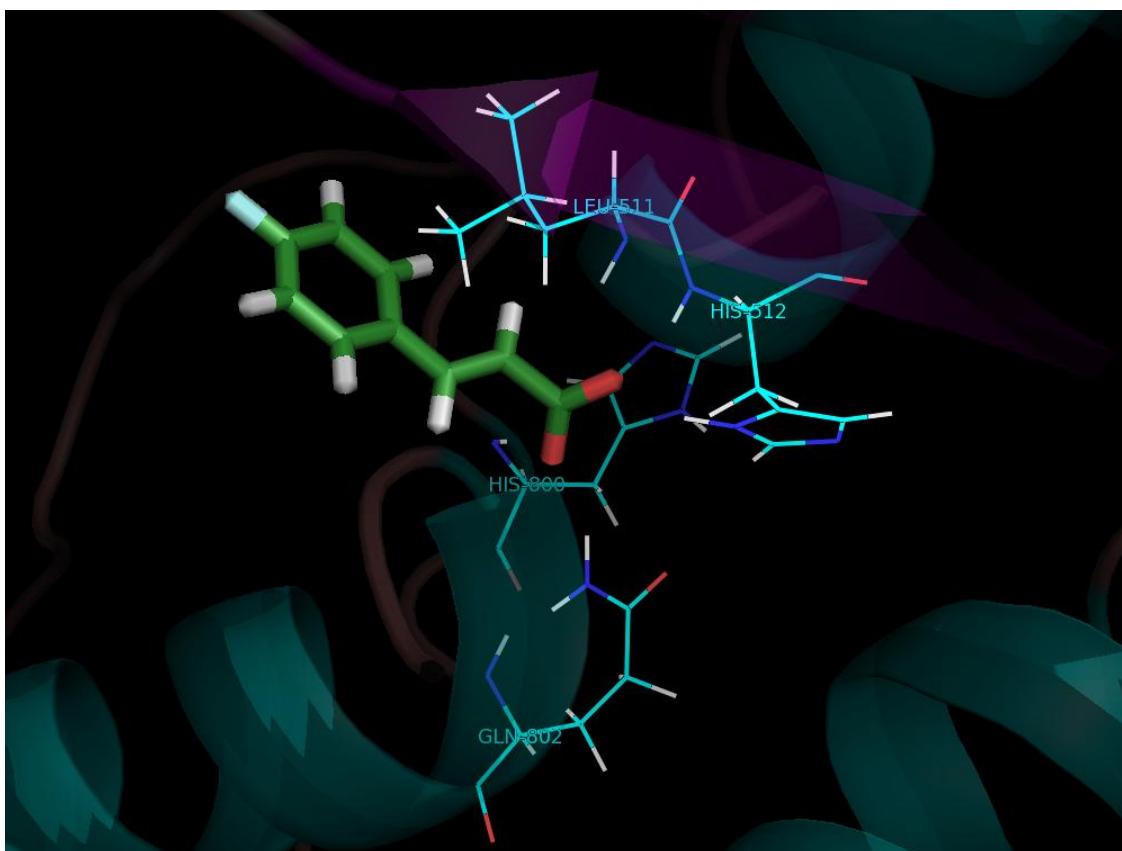


Figura 21: Previsão de pose do ácido *p*-flúor-cinâmico, com carbonos em verde, resíduos de aminoácido com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6)

5.4.4. Fragmentos selecionados da série PPGQ2016R02

Um dos resultados mais interessantes da triagem foi a seleção do ácido cafeico (Figura 22) pelos três critérios de normalização utilizados. Apesar da grande sobreposição entre os compostos selecionados por massa e área superficial, poucos foram os selecionados por \sqrt{N} com boas colocações nos outros critérios. Ambos os compostos selecionados dessa série (Salgueiro, 2016) tiveram sua melhor correspondência em DENV-3. As distâncias de ligação podem ser observadas nos anexos 4A e 4B.

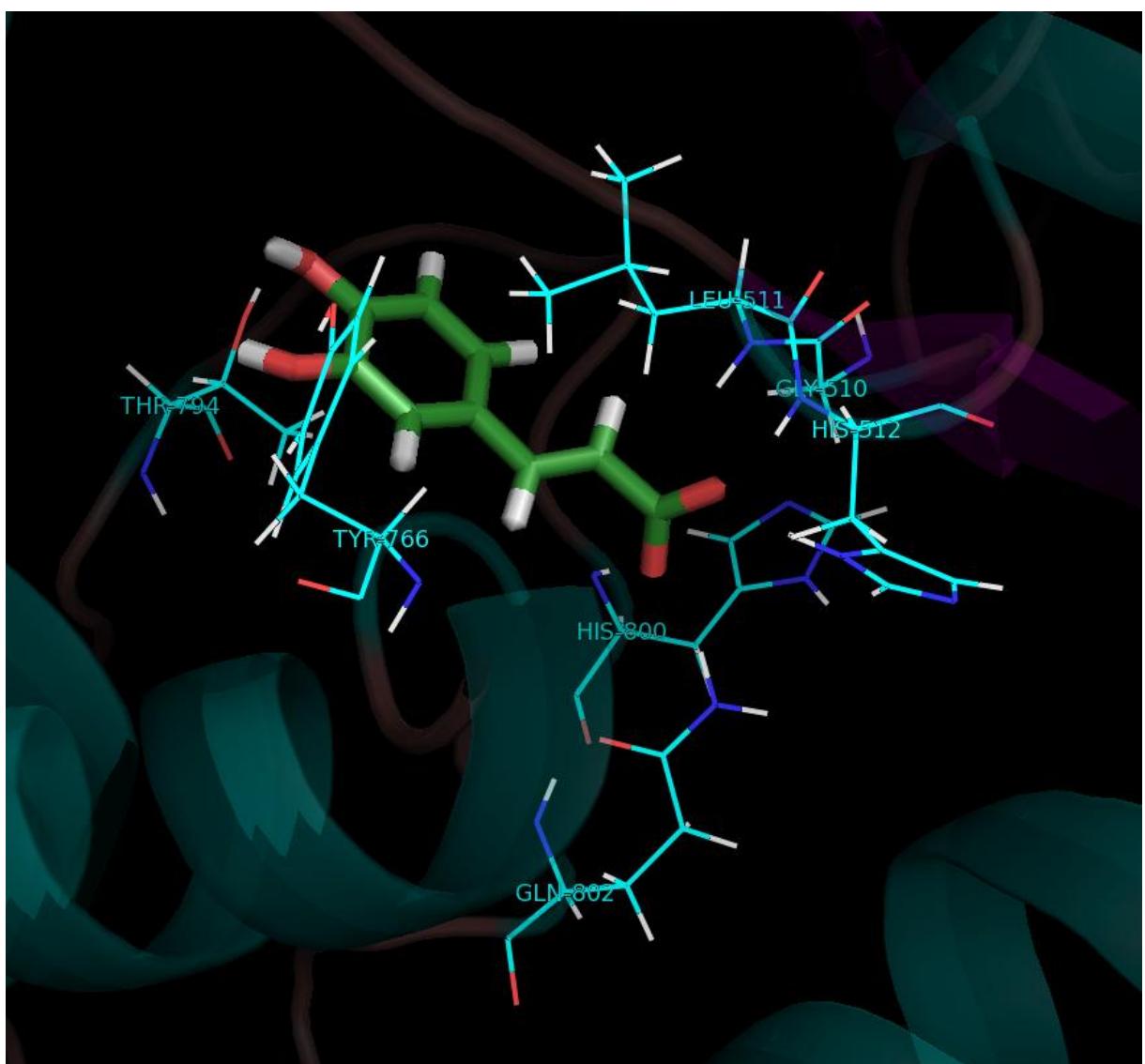


Figura 22: Previsão de pose do ácido cafeico, com carbonos em verde, resíduos de aminoácidos com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6)

Previamente, foi mostrado o perfil de interações do ácido *p*-cumárico, que difere do ácido cafeico apenas porque o último tem uma hidroxila a mais na posição *meta*. As interações

do ácido cafeico com o sítio são as mesmas do ácido *p*-cumárico. Devido a presença da hidroxila na posição 3 há uma ligação de hidrogênio com a cadeia lateral da Tyr766.

A drupanina (Figura 23), por sua vez, contém na posição *meta* uma porção alquílica que não faz muitas interações complementares. Próximo aos carbonos da ramificação há os resíduos Leu514 e His798. O subsítio hidrofóbico para o qual a porção alquílica se direciona é o mesmo visto anteriormente na série PPGQ2010R01, entretanto, as interações por ligações de hidrogênio tiveram maior influência na definição da pose, levando a uma orientação que não maximiza a potencialidade das interações por forças de dispersão. Mesmo as interações que ocorrem estão a uma distância de mais de 3Å.

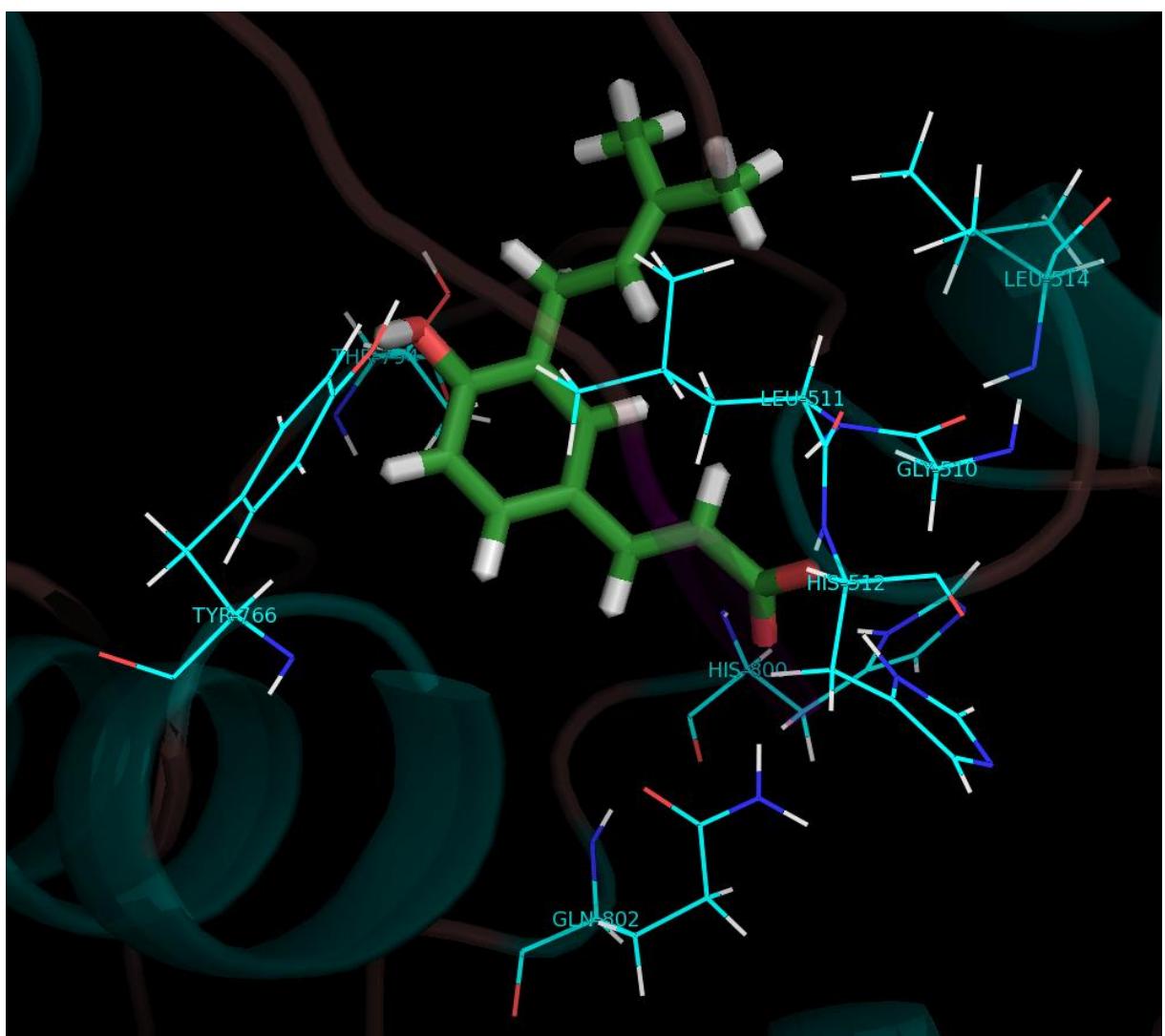


Figura 23: Previsão de pose da drupanina, com carbonos em verde, resíduos de aminoácidos com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6)

As demais interações são conservadas com o já apresentado. O carboxilato faz ligações de hidrogênio com os resíduos Leu511, His512, His800 e Gln802. A hidroxila em *para* faz ligação de hidrogênio com a Tyr766.

5.4.5. Fragmentos selecionados da série MPPGQ2004R03

Os compostos selecionados da série incluem o ácido 4-metoxi-cinâmico e o ácido *p*-hidroxibenzóico (Lianda, 2004). O primeiro foi selecionado para DENV-2 e o segundo para DENV-3.

Iniciando a análise pelo ácido 4-metoxi-cinâmico (Figura 24), o grupo carboxilato realiza interações com os resíduos Lys800, His513 e Leu 512. Devido ao carboxilato do ácido glutâmico, há deslocamento da posição que outros derivados cinâmicos ocupam em DENV-2. A distância entre o centro dos dois carboxilatos é de 5,54 Å. De modo relativamente semelhante, o oxigênio do grupo éter interage com a cadeia lateral da Thr794. As distâncias de ligação podem ser observadas nos anexos 5A e 5B.

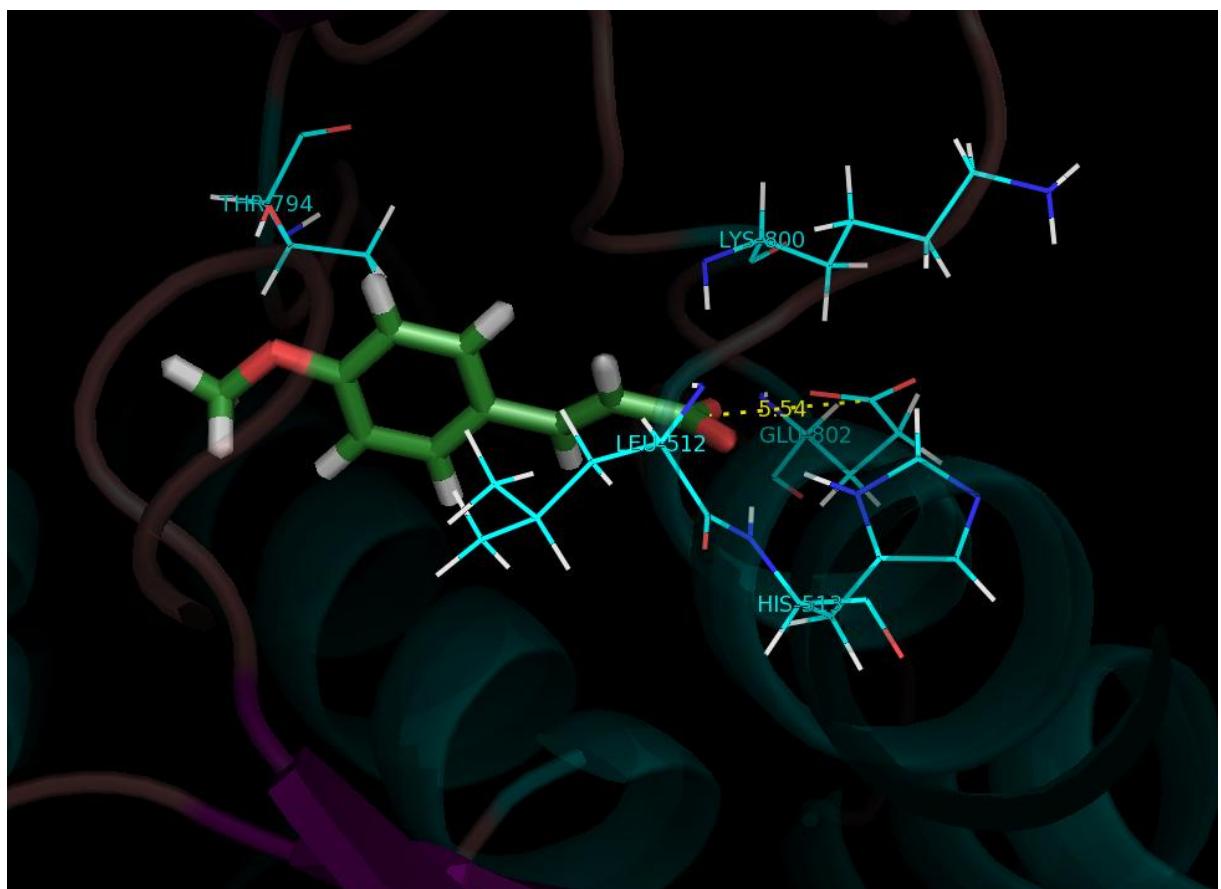


Figura 24: Previsão de pose do ácido 4-metoxicinâmico com carbonos em verde, resíduos de aminoácidos com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6)

O ácido *p*-metoxibenzóico (Figura 25) também contém um grupo carboxilato, mas não há espaçador entre o anel e o carboxilato como nos derivados cinâmicos. A análise desse caso fornece a possibilidade de discutir a relevância desses dois carbonos que funcionam como espaçador plano.

É interessante perceber que o anel permanece no mesmo centro que nos demais ligantes, mas há a mudança na orientação do carboxilato, que passa a interagir com Thr794 e Tyr766. Os resíduos His512 e Gln802 estão mais distantes e há interação da hidroxila com His 798, His800 e Leu511. Outra informação que se pode retirar dessa solução é que substituintes polares podem ser bem estabilizados em *para*, pelo menos, desde que estejam ligados diretamente ao anel aromático. Essa informação pode ser relevante na proposição de modificações estruturais dos fragmentos selecionados.

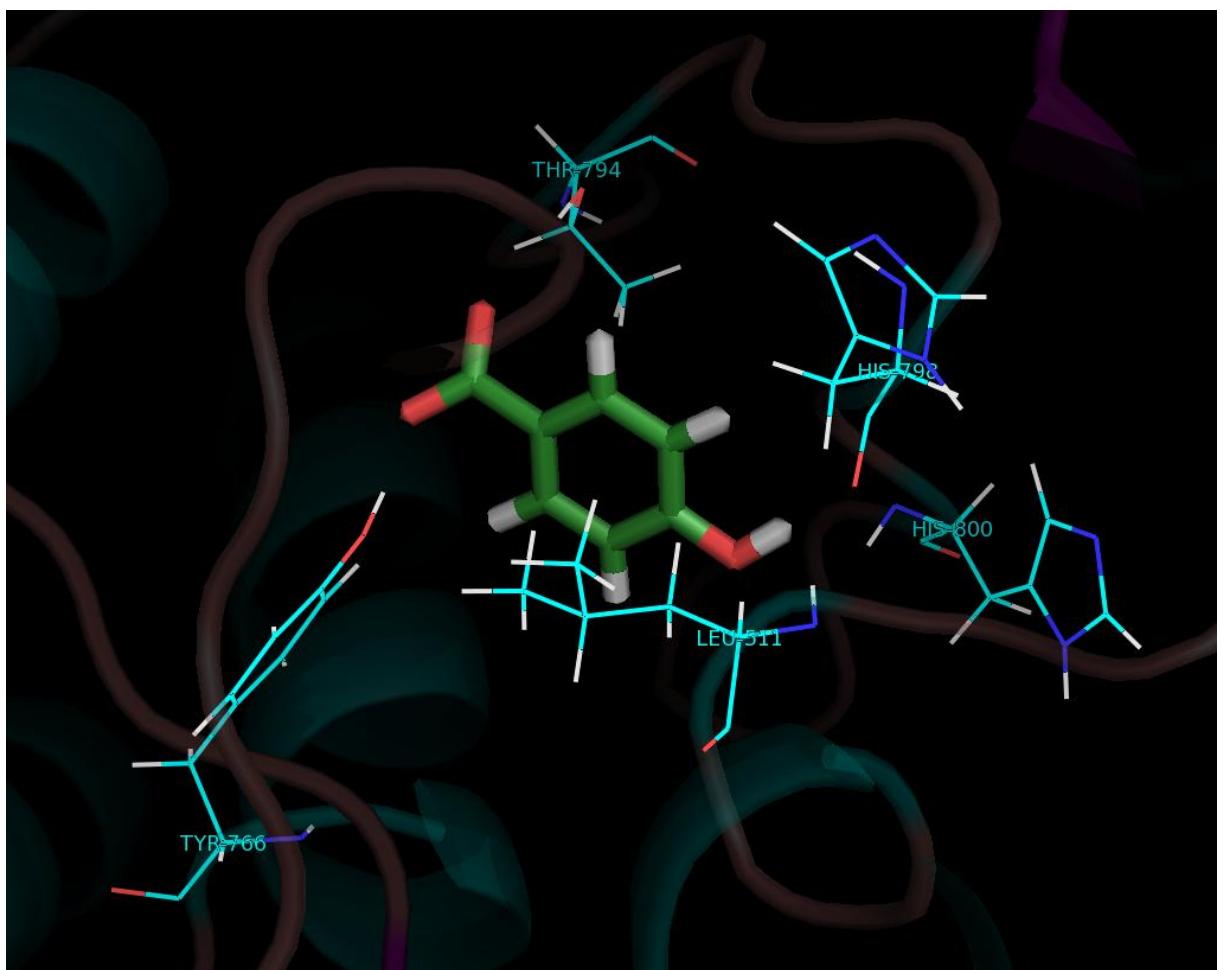


Figura 25: Previsão de pose do ácido p-metoxibenzóico com carbonos em verde, resíduos de aminoácidos com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6)

5.4.6. Fragmentos selecionados da série PPGQ1998R01

Em mais uma série de hidrocarbonetos selecionada (Alvez, 1998), temos os produtos naturais mirceno e *p*-cimeno. O primeiro foi selecionado para DENV-3 e o segundo para DENV-2. Novamente, sendo os selecionados hidrocarbonetos, utilizamos suas soluções como fonte de informações estruturais do sítio.

O mirceno (Figura 26) faz interações complementares por apenas uma face em um subsítio majoritariamente apolar que foi registrado anteriormente para o ocimeno. Entretanto, o mirceno não tem geometria adequada para alcançar outros resíduos, como o hidrocarboneto apresentado anteriormente foi capaz. As distâncias das ligações de hidrogênio, entre outras interações, podem ser observadas nos anexos 6A e 6B.

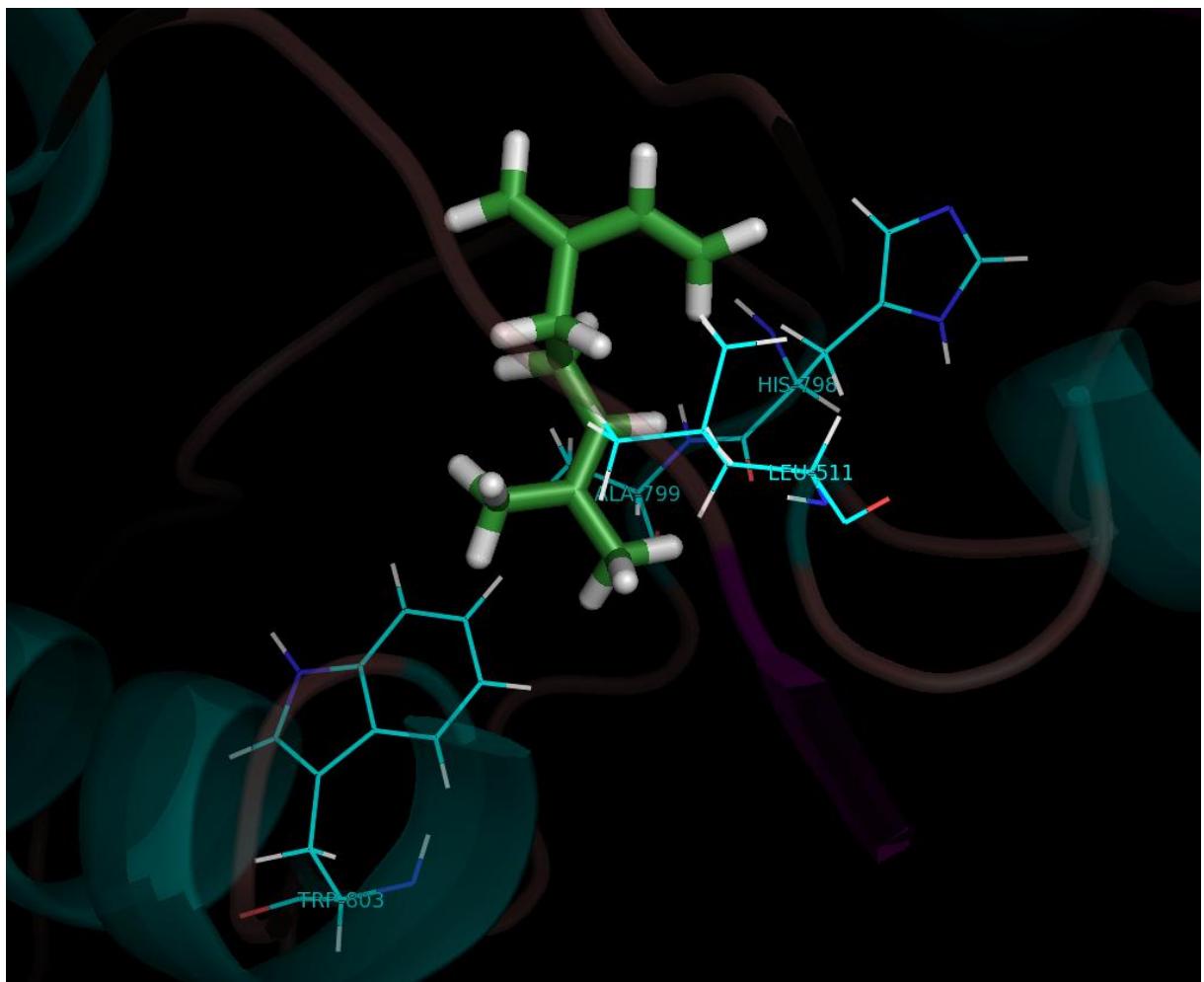


Figura 26: Previsão de pose do mirceno com carbonos em verde, resíduos de aminoácidos com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6)

Quanto ao *p*-cimeno, são registradas pouquíssimas interações que se relacionam com o mesmo subsítio que todos os hidrocarbonetos até então analisados. É relevante notar que a

seleção dessas moléculas que interagem fracamente é uma ocorrência esperada da seleção dos compostos normalizados por massa. A divisão da pontuação pela massa é uma super-correção, como mencionado, tendendo a selecionar moléculas pequenas.

O *p*-cimeno (Figura 27), segundo o servidor PLIP, faz interações complementares apenas com Leu512, Ala799 e Trp803. Na Figura 26 mais resíduos foram incluídos para servir de referência da posição do hidrocarboneto. Ao comparar a Figura 27 com a Figura 24, que mostra o ácido 4-metoxicinâmico, também em DENV-2, conseguimos perceber que a posição geral do anel segue a mesma. Essas posições semelhantes auxiliam o planejamento do que deve ou não ser modificado em otimizações estruturais.

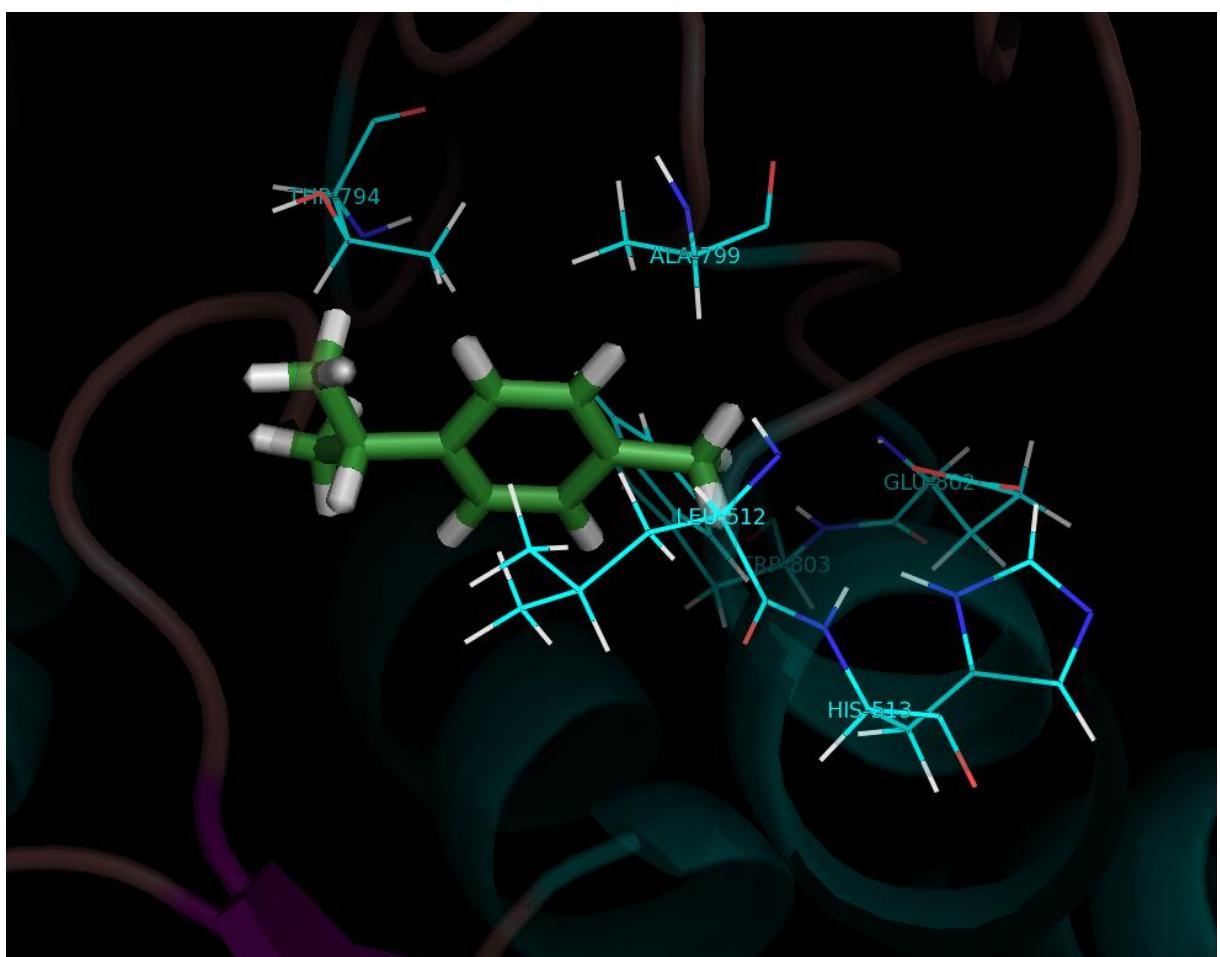


Figura 27: Previsão de pose do *p*-cimeno com carbonos em verde, resíduos de aminoácidos com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6)

5.4.7. Compostos selecionados da série PPGQ2000701

Desse subtópico em diante serão analisadas as séries que incluem moléculas de massa maior que 300Da selecionadas na triagem virtual. Todas foram selecionadas pelo critério de normalização \sqrt{N} e serão analisadas no viés de possíveis candidatos a fármacos.

O ligante PPGQ200070110 (Ferreira, 2000), selecionado para DENV-2, tem ótima ocupação do sítio, tendo tamanho e flexibilidade adequadas (Figura 28). Apesar disso, a maior parte de suas interações complementares são fracas. Apenas duas ligações de hidrogênio são registradas. A primeira entre uma hidroxila do ligante e a cadeia lateral da Thr794 e a segunda entre um dos oxigênios do grupo éter e o N da cadeia principal da Leu512. Essa segunda é razoavelmente distante, perdendo força devido aos 3,26 Å que separam os átomos. As distâncias de ligação podem ser observadas nos anexos 7A a 7C.

Diferente dos fragmentos, temos interação com a Arg737, que é comprovadamente essencial para a função polimerase da enzima (Lim et al., 2016). Essa interação é fraca e, portanto, compromete pouco a arginina, mas a presença dessa interação em um dos ligante que contém anel aromático dá ideia das substituições necessárias nos fragmentos para alcançar o resíduo Arg737.

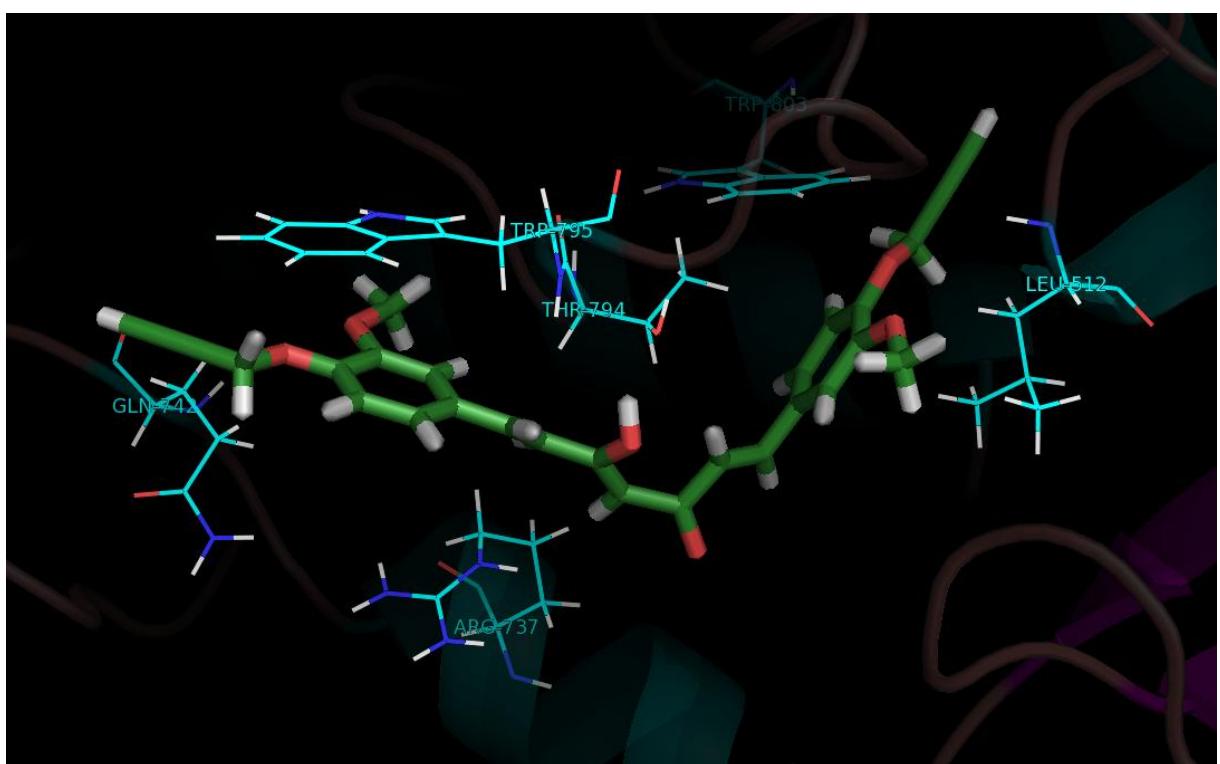


Figura 28: Previsão de pose do *ligante (1E,4Z,6E)-1-(4-((etiniloxi)metyl)-3-metoxifenil)-5-hidroxi-7-(3-metoxi-4-(prop-2-in-1-iloxi)fenil)hepta-1,4,6-trien-3-ona* com carbonos em verde, resíduos de aminoácidos com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6)

Por fim, há uma interação do tipo empilhamento- π entre um dos anéis aromáticos do ligante e a cadeia lateral do resíduo Trp795.

O ligante PPGQ20070111 (Castro, 2000), selecionado para DENV-2, foi capaz de realizar 4 ligações de hidrogênio, com os resíduos Arg729, Thr794, Trp795 e Ser796 (Figura 29). O resíduo Arg729 também é essencial para a função polimerase (Lim et al., 2016), sendo essa uma interação muito desejável. Novamente, é registrada uma interação fraca com a Arg737 e é sugerida uma interação do tipo empilhamento- π com o Trp795.

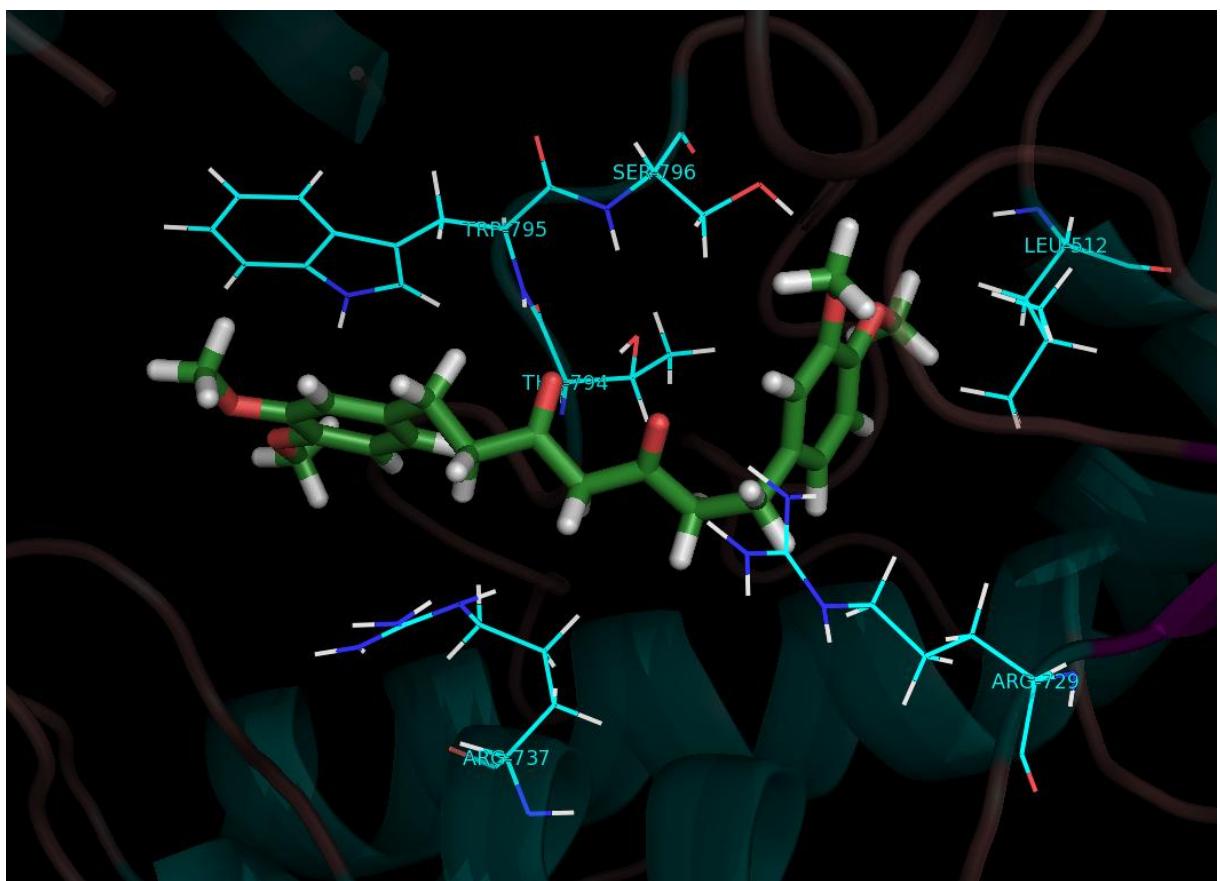


Figura 29: Previsão de pose do ligante [1,7- bis -(3, 4- dimetóxifenil)] -heptano - 3,5- diona com carbonos em verde, resíduos de aminoácidos com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6)

O ligante PPGQ200070130 foi selecionado para DENV-3 e apresenta os resultados mais promissores. São registradas 6 ligações de hidrogênio, com os resíduos Leu511, Arg729, Tyr766, Trp795, His798 e His800 e uma interação cátion- π com a Arg737 (Figura 30). A previsão de um mesmo ligante ser capaz de realizar interações com ambos os resíduos de arginina essenciais para a catálise é vantajosa quando se planeja a inibição da NS5. A distância da ligação de hidrogênio envolvendo a Arg729 e o oxigênio da cetona do ligante é de 1,87 \AA ,

sendo bastante forte. A distância da interação cátion-π, que envolve a outra arginina, é de 3,55Å, em média.

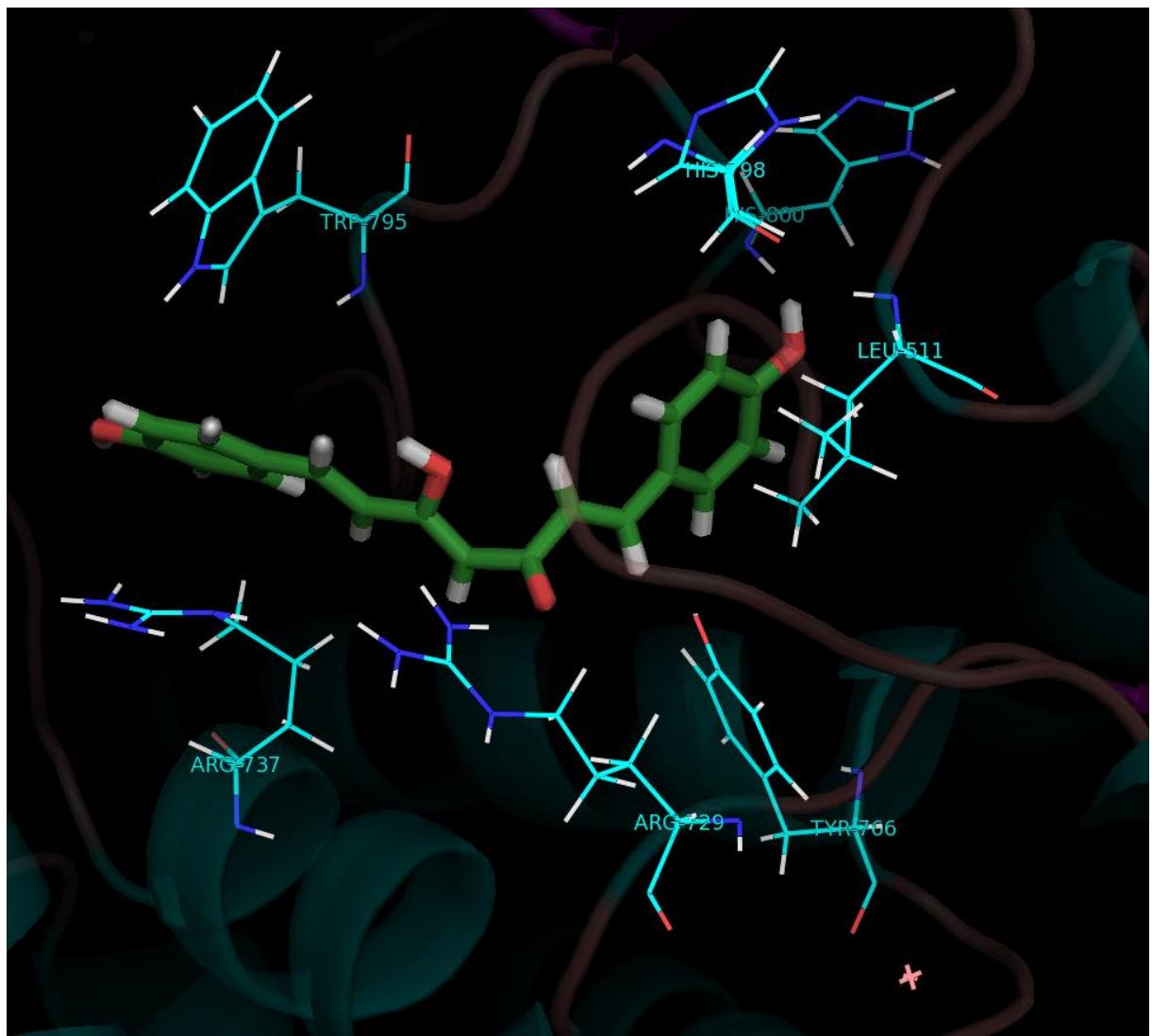


Figura 30: Previsão de pose do ligante $(1E,4Z,6E)$ -5-hidroxi-1,7-bis(4-hidroxifenil)hepta-1,4,6-trien-3-ona com carbonos em verde, resíduos de aminoácidos com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6)

5.4.8. Compostos selecionados da série PPGQ2019201

O ligante PPGQ201920105 (Rodrigues, 2020), selecionado para DENV-2, realiza duas ligações de hidrogênio, sendo essas com os resíduos Tyr766 e Thr794 (Figura 31). O ligante é bem ancorado por interações do tipo dispersões de London, com os resíduos de aminoácidos Leu511, Ala799 e Trp803. Além disso, ocorre uma interação de empilhamento- π em forma de T com o resíduo Trp795. As distâncias de ligação podem ser observadas nos anexos 8A a 8C.

O ligante PPGQ201920108 realiza apenas 3 ligações de hidrogênio com o sítio alostérico, mas é ancorado por interações fracas com diversos resíduos próximos por ambas as faces (Figura 32). As ligações de hidrogênio envolvem os resíduos Thr794 e Trp795 e um nitrogênio de imina e entre Tyr766 e outro nitrogênio de imina. O resíduo Arg737 está próximo do ligante, mas interage apenas por interações mediadas por forças de London.

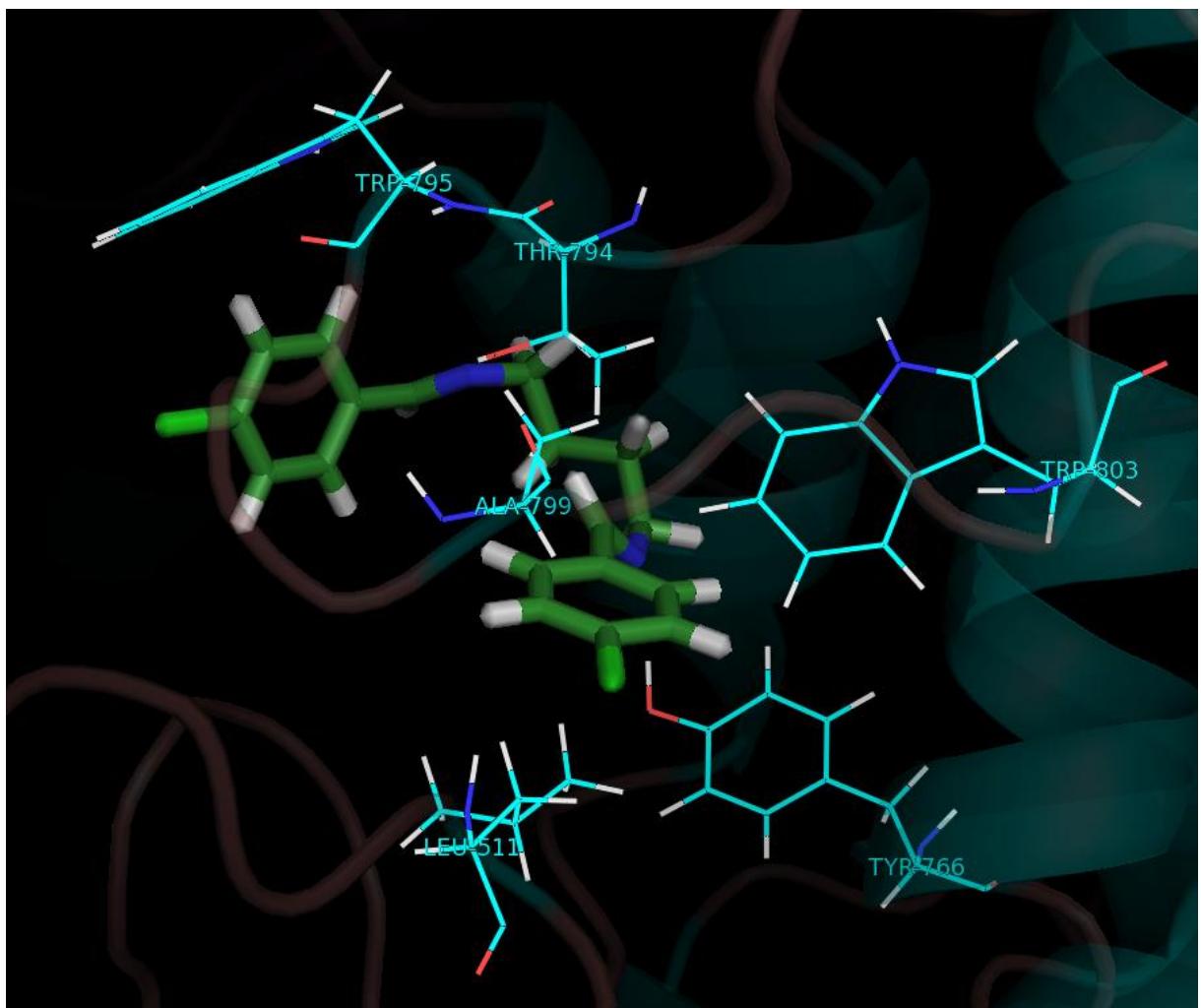


Figura 31: Previsão de pose do ligante (E,E) -N¹,N⁴-bis(4-clorobenzilideno)butano-1,4-diamina com carbonos em verde, resíduos de aminoácidos com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6)

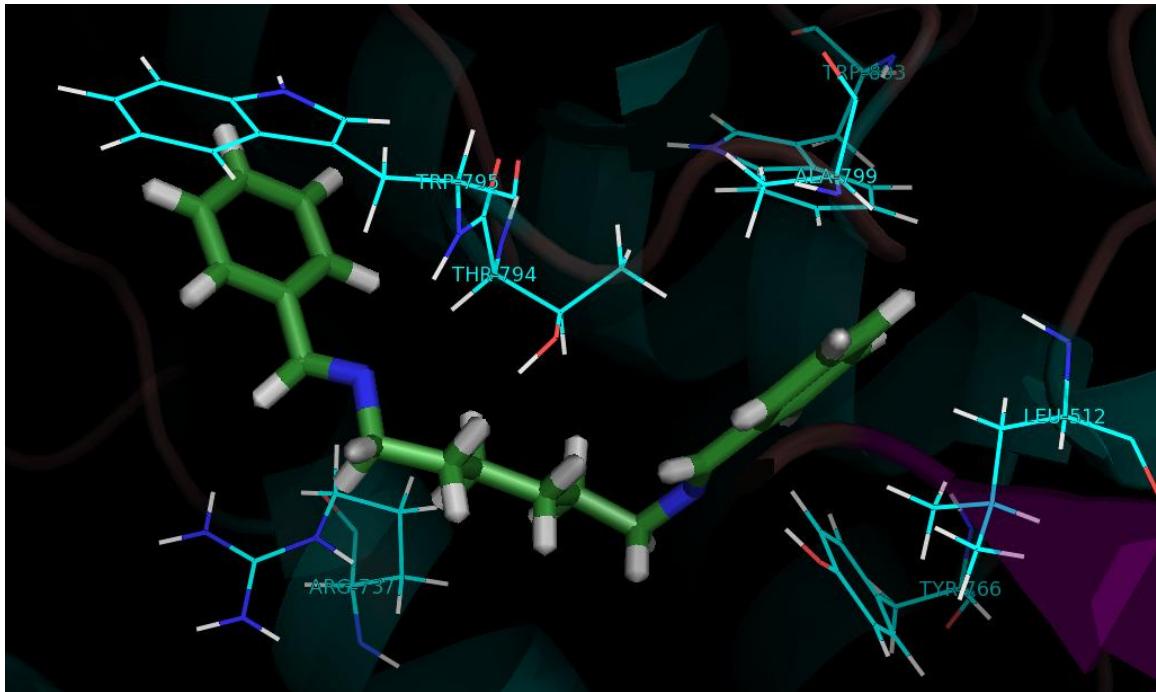


Figura 32: Previsão de pose do ligante (E,E) -N¹,N⁶-dibenzilidenohexano-1,6-diamina com carbonos em verde, resíduos de aminoácidos com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6)

O ligante PPGQ201920109 interage por apenas duas ligações de hidrogênio, sendo essas com os resíduos Tyr766 e Trp795 e um empilhamento- π envolvendo o resíduo Trp795. Além dessas, são registradas 9 interações mediadas por forças de dispersão com resíduos próximos, conforme pode ser observado na Figura 33.

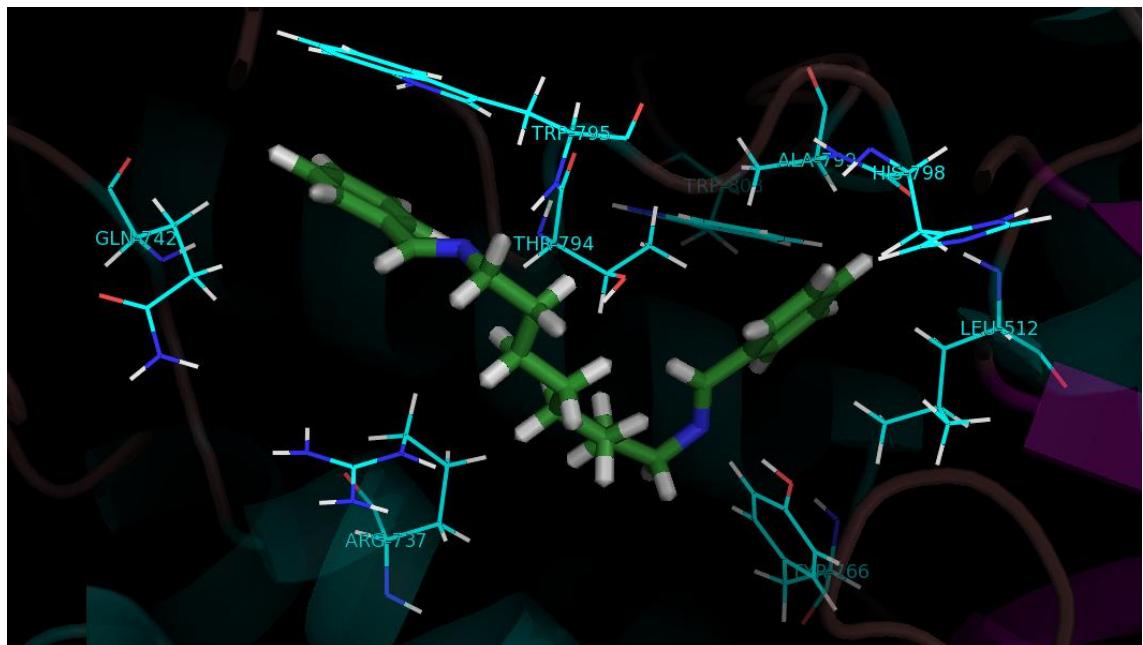


Figura 33: Previsão de pose do ligante (E,E) -N¹,N⁶-dibenzilidenohexano-1,6-diamina com carbonos em verde, resíduos de aminoácidos com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6)

5.4.9. Ligante PPGQ201430113

A seleção desse ligante (Ferreira, 2014) muito provavelmente foi enviesada pelo seu tamanho apesar do processo de normalização proposto. A presença de uma cadeia dodecila permite interações por forças de dispersão de London como interações complementares e influencia na leitura do ligante como bem pontuado (Figura 34). O único resíduo de aminoácido com o qual o ligante interage por ligações de hidrogênio é Arg729, que é essencial para a catálise (Lim et al., 2016). As distâncias de ligação podem ser observadas no Anexo 9.

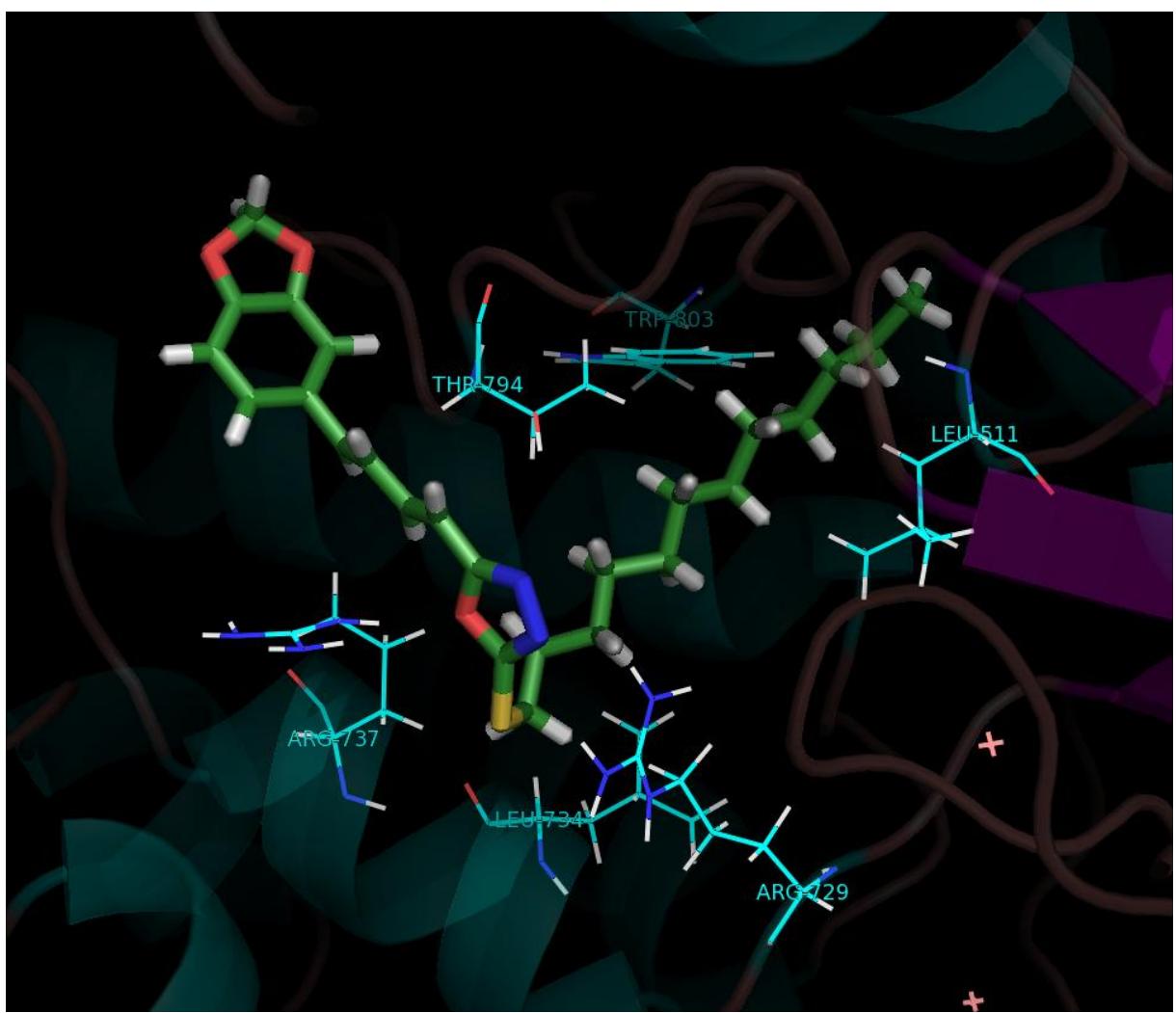


Figura 34: Previsão de pose do ligante 2-((1*E*,3*E*)-4-(benzo[*d*][1,3]dioxol-5-il)buta-1,3-dien-1-il)-5-(dodeciltio)-1,3,4-oxadiazol com carbonos em verde, resíduos de aminoácidos com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6)

5.4.10. Ligante MPPGQ2003M0103

O composto (*E*)-3-((2*S*,3*S*)-2-(4-hidroxi-3-metoxifenil)-3-(hidroximetil)-7-metoxi-2,3-di-hidrobenzofuran-5-il)acrilaldeído é um produto natural extraído da madeira de pau-brasil. O aldeído é ancorado pelas duas faces, realizando 9 ligações de hidrogênio com resíduos de aminoácidos do subsítio N. A Arg729 é um desses resíduos e está envolvida em uma ligação de hidrogênio com 2,27Å de distância entre uma hidroxila do ligante e o grupo guanidínio da cadeia lateral da arginina. As distâncias de ligação podem ser observadas no Anexo 10.



Figura 35: Previsão de pose do ligante (*E*)-3-((2*S*,3*S*)-2-(4-hidroxi-3-metoxifenil)-3-(hidroximetil)-7-metoxi-2,3-di-hidrobenzofuran-5-il)acrilaldeído com carbonos em verde, resíduos de aminoácidos com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6)

Além do grande número de ligações de hidrogênio, quando comparamos com o inibidor **29**, há 5 ligações de hidrogênio em comum. Isso significa que o ligante é capaz de ocupar o sítio e interagir de maneira similar a de um inibidor comprovado.

5.5. Análise das propriedades ADMET dos ligantes selecionados

Nesse item, analisamos os ligantes grandes (massa > 300Da) pelas regras de Lipinski, criando um filtro de quais compostos selecionados são mais promissores do ponto de vista da previsão da biodisponibilidade oral.

A regra de Lipinski ou regra dos 5 (Ro5) fornece fundamentos para a análise da biodisponibilidade oral de um candidato a fármaco (Baell et al., 2013), sendo constituída por 4 critérios:

- i. o número de doadores de ligações de hidrogênio deve ser menor que 5;
- ii. o número de aceptores de ligações de hidrogênio deve ser menor que 10;
- iii. a massa molecular deve ser menor que 500 Da;
- iv. o coeficiente de partição octanol-água (ou LogP) deve ser até 5.

Os resultados completos da análise ADMET pelo servidor Swiss-ADME estão disponíveis nos anexos 12A até 12I.

Tabela 4: Análise dos ligantes grandes pelas regras de Lipinski

Código	Critério I	Critério II	Critério III	Critério IV
MPPGQ2003M0103	2 doadores	6 aceptores	356,37 Da	LogP: 2,70
PPGQ200070110	1 doador	6 aceptores	450,52 Da	LogP: 4,22
PPGQ200070111	0 doadores	6 aceptores	400,46 Da	LogP: 3,73
PPGQ200070130	3 doadores	4 aceptores	314,38 Da	LogP: 2,86
PPGQ201430113	0 doadores	5 aceptores	446,65 Da	LogP: 7,11
PPGQ201920105	1 doador	2 doadores	335,27 Da	LogP: 4,13
PPGQ201920109	1 doador	2 aceptores	322,49 Da	LogP: 4,31
INIBIDOR 27	2 doadores	7 aceptores	487,55 Da	LogP: 3,05
INIBIDOR 29	2 doadores	6 aceptores	492,57 Da	LogP: 3,69

Para o valor de LogP (Djaković Sekulić et al., 2018), o Swiss-ADME apresenta diversas previsões por métodos diferentes. Aqui, o consenso entre as previsões é apresentado. Em primeira análise, pela baixíssima solubilidade aquosa, que se relaciona com o alto valor de LogP, eliminamos o ligante PPGQ201430113 como candidato.

A seguir, foi feita a previsão da capacidade dos compostos em inibir enzimas do citocromo P450, cuja inibição pode levar a efeitos colaterais indesejados (Souza; Oliveira, 2022).

Tabela 5: Previsão de inibição das CYPs pelos ligantes selecionados. Gerado pelo servidor Swiss-ADME.

Código	CYP1A2	CYP2C19	CYP2C9	CYP2D6	CYP3A4
MPPGQ2003M0103	Não	Não	Sim	Sim	Sim
PPGQ200070110	Não	Não	Não	Sim	Sim
PPGQ200070111	Não	Não	Não	Sim	Sim
PPGQ200070130	Não	Não	Não	Sim	Sim
PPGQ201920105	Sim	Sim	Não	Sim	Sim
PPGQ201920109	Sim	Sim	Não	Sim	Sim
INIBIDOR 27	Não	Sim	Sim	Não	Sim
INIBIDOR 29	Não	Sim	Sim	Não	Sim

Analizando a Tabela 5, percebemos que todos os ligantes são potencialmente inibidores de pelo menos duas CYPs, apresentando, portanto, potencial toxicidade. Dentre eles, entretanto, os que interagem com menor quantidade de CYPs são os da série PPGQ2000701. Comparando com os inibidores de referência **27** e **29**, a série supracitada é, potencialmente, menos tóxica.

5.6. Otimização dos fragmentos por estratégias de FBDD

Inicialmente, as interações do ácido cinâmico e seus derivados no sítio alostérico foram analisadas. Podemos observar grande conservação em algumas interações nas poses que os ligantes assumem no sítio onde foram docados. Os carboxilatos em geral fazem ligações de hidrogênio com os mesmos resíduos de aminoácidos, como apresentado. Outro dado que comprova a conservação das interações complementares, apesar das substituições nos ácidos cinâmicos, é a sobreposição dos complexos, como pode ser visto na Figura 36A. Percebe-se que todos os carboxilatos apontam para a mesma direção, ainda que com pequenas mudanças na conformação. As interações também são conservadas. Esse dado trouxe a possibilidade de explorar modificações na estrutura do ácido cinâmico que maximizassem interações em outros subsítios, porém, sem perder as interações conservadas dos carboxilatos.

Dentre as técnicas clássicas de otimização de fragmentos, as estratégias de fusão e junção não são aplicáveis aos resultados da pesquisa. Como mostrado, as posições são muito conservadas. Decidiu-se, portanto, otimizá-los por crescimento. Para guiar a estratégia, utilizou-se a sobreposição do ácido cinâmico ao inibidor 29 (Figura 13). Além disso, observou-se a pose da drupanina no sítio, tendo sido esse o ligante melhor pontuado na docagem e com entalpia de interação mais favorável. A sobreposição da drupanina e do ácido cinâmico também foi analisada para guiar o planejamento racional dos candidatos (Figura 36B).

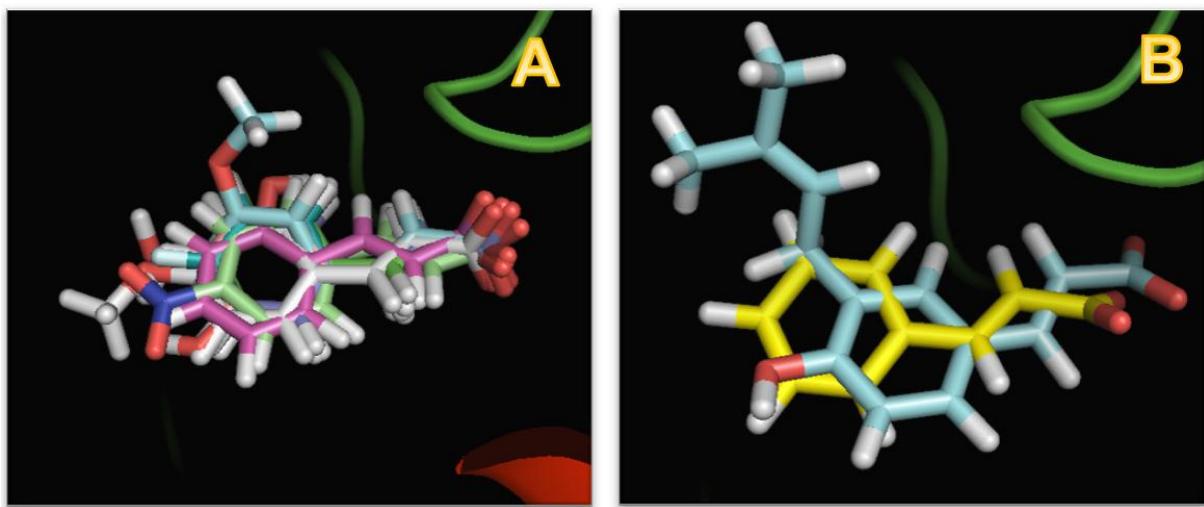


Figura 36: Sobreposições de fragmentos selecionados com o ácido cinâmico. Oxigênios em vermelho, nitrogênios em azul e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6) A: Todos os derivados do ácido cinâmico, exceto drupanina, sobrepostos (carbonos em cores variadas). B: Drupanina (carbonos em ciano) e ácido cinâmico (carbonos em amarelo) sobrepostos.

Entre a sobreposição dos fragmentos com o ácido cinâmico, exceto drupanina e da drupanina com ácido cinâmico, percebe-se deslocamento da pose antes conservada entre os

carboxilatos dos fragmentos. Ainda assim, na comparação do ácido cinâmico com a drupanina, as ligações de hidrogênio, em geral, são iguais. A drupanina é substituída em *para* (hidroxila) e em *meta* (3,3-dimetil-propenila). A substituição em *meta* levou a interações com um subsídio razoavelmente hidrofóbico.

Diante da possibilidade de maximizar as interações complementares do ácido cinâmico pela adição de grupos alquila e alquil-arila na posição *meta* do anel aromático, propôs-se a primeira série de derivados e esses foram submetidos à docagem na estrutura 5I3Q.

5.6.1. Série 1: Substituintes alquila e alquil-arila

Foi realizada a docagem dos fragmentos modificados e análise da pose melhor pontuada. Além disso, foram realizados cálculos SE para estimar a entalpia de interação entre os derivados alquilados e o sítio alostérico da RdRp do DENV-3. Na Tabela 7, são apresentados os valores de entalpia de interação absolutos e relativos, tendo o ácido cinâmico como referência, para que se possa observar o quanto benéficas foram as otimizações em relação ao fragmento base, na perspectiva da entalpia de interação. Os resultados da docagem também estão apresentados na forma de pontuações relativas, ou seja, a diferença entre a pontuação do ligante modificado e a do ácido cinâmico, base para o crescimento.

Tabela 6: Pontuações de docagem e entalpias de interação para novos derivados alquilados do ácido cinâmico.

Ácido Cinâmico	Código	Pontuação Relativa (ChemPLP)	ΔH interação	ΔH interação relativo
			(kcal/mol)	(kcal/mol)
Substituição em <i>meta</i>				
Metila	1a	2,07	-53,60	2,78
Etila	1b	1,35	-51,44	4,94
Propila	1c	8,24	-67,72	-11,34
Isopropila	1d	3,38	-63,98	-7,60
Butila	1e	14,2	-65,30	-8,92
Isobutila	1f	11,06	-64,27	-7,89
Terc-butila	1g	-3,79	-54,43	1,95
Pentila	1h	9,44	-95,40	-39,01
Isopentila	1i	13,75	-66,43	-10,05
Neopentila	1j	-0,16	-61,10	-4,71
2-fenil-etila	1l	21,34	-81,79	-25,41
3-fenil-propila	1m	15,09	-66,91	-10,53

A primeira série (1a – 1m) pretendia testar como os complexos de docagem variam mediante a adição de grupos alquila na posição *meta*. Sendo assim, foram propostas substituições com grupos alquila variados, inspiradas nos fragmentos apolares selecionados e, por fim, a inserção de anéis aromáticos com espaçadores, considerando como objetivo que o fragmento otimizado pudesse realizar interação cátion-π com o resíduo de aminoácido Arg737, como ocorreu, por exemplo, para o ligante PPGQ200070130.

Iniciando a discussão apenas pelas interações conhecidas e o que se obteve da docagem, substituintes muito volumosos perto do anel aromático principal não trazem boas interações complementares. O sítio parece responder melhor a substituintes mais compridos e flexíveis. Por exemplo, olhando as pontuações da docagem, o substituinte butila é previsto como melhor acomodado que um substituinte isobutila. A partir da adição de dois metilenos como espaçadores e então um grupo de maior volume, houve melhores resultados, como no caso do

substituinte isopentila, quando comparado ao isobutila. Entretanto, o grande ganho de pontuação ocorre na adição do substituinte 2-fenetila. Para compreender esse resultado, procedemos os cálculos de entalpia de interação com o PM7, obtendo as entalpias de interação mostradas na Tabela 6.

Em segunda colocação está o ácido *m*-fenil-etyl-cinâmico (Figura 38), que apresenta, além das interações conservadas do ácido cinâmico, interações por forças de dispersão de London com átomos da fenila e uma interação cátion-π com a cadeia lateral do resíduo de Arg737.

As ligações de hidrogênio conservadas são com Leu511, His512, His800 e Gln802. As três primeiras são com o esqueleto polipeptídico da proteína e o carboxilato do ligante. A última envolve a cadeia lateral do resíduo de aminoácido Gln802. Interações envolvendo os resíduos de aminoácido Leu511 e Trp803 estabilizam o anel do ácido cinâmico por forças de dispersão. A modificação estrutural causa novas interações, que são as destacadas na figura 38. Leu734, Arg737 e Tyr758 interagem com átomos do substituinte fenila e, como interação mais forte, temos a citada cátion-π, com distância média de 3,69Å.

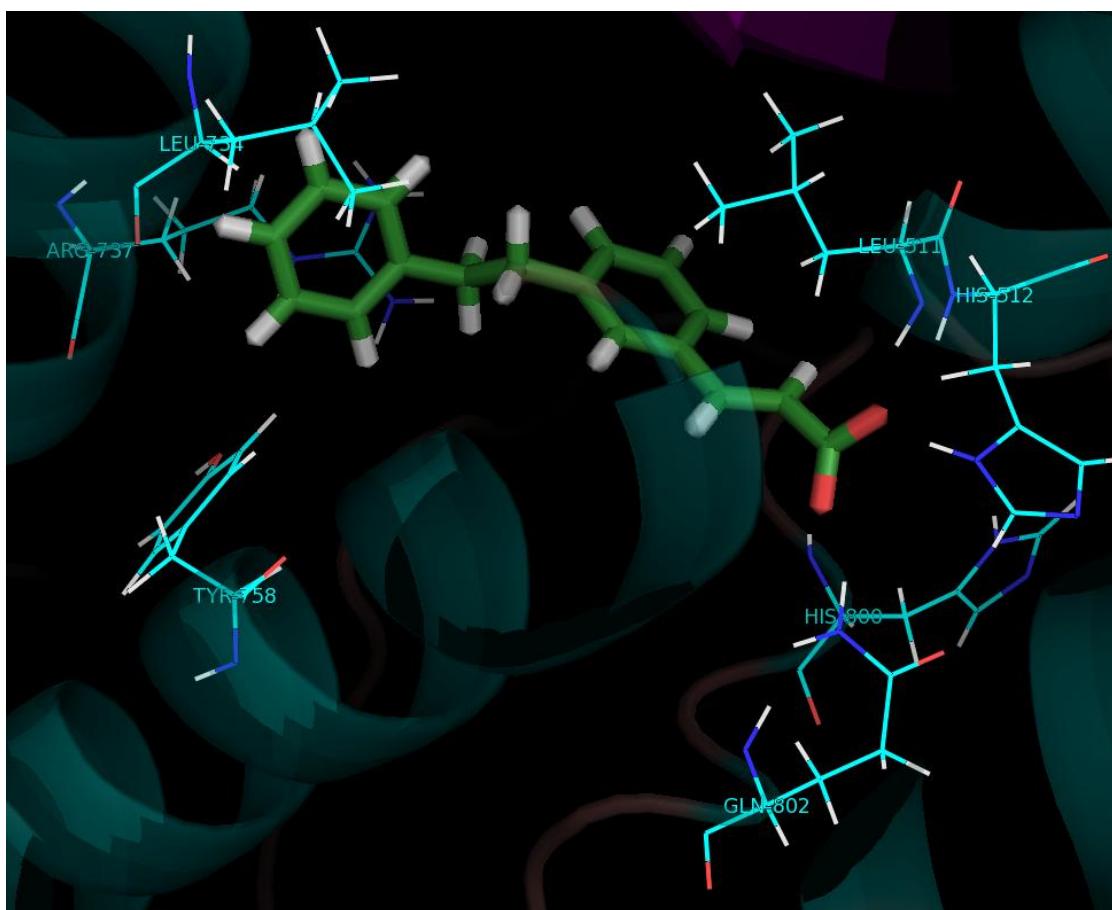


Figura 37: Previsão de pose do ligante ácido *m*-fenil-etyl-cinâmico com carbonos em verde, resíduos de aminoácidos com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6)

Em primeira colocação está **1h**, com substituinte pentila. Ao analisar a entalpia de interação sempre se segue a Equação 1 (citada no tópico 4.5).

Como os cálculos foram realizados em contínuo dielétrico que simula a água, quanto menos estável o ligante for nesse meio, menos negativo é o valor de ΔH_f (ligante) e, por consequência, o resultado de ΔH_{int} é mais negativo. No caso do ligante **1h**, esse fator deve ser preponderante. De toda a série 1, ele é o que apresenta menor estabilidade em simulação de solvente aquoso e, dessa forma, resulta, pelo cálculo, no valor mais favorável de entalpia de interação. Entretanto, em quantidades termodinâmicas, a relação entre ambos os lados da equação deve ser considerada. Se o ligante fosse igualmente pouco estável no sítio da proteína, o valor da entalpia de interação não seria tão favorável.

Não são registradas interações intermoleculares com o substituinte pentila, entretanto, o ligante, de caráter mais apolar, deve ser mais estável em meio proteico apolar do que em água, um solvente fortemente polar.

Dando continuidade, cientes das limitações das substituições de caráter exclusivamente apolar e ambicionando o estabelecimento de interações mais fortes, além da melhora do perfil farmacocinético (a ser discutido no tópico 5.6.3) e, possivelmente, facilitação da rota sintética, uma segunda série de otimizações estruturais foi proposta a partir do arcabouço inicial do ácido *m*-etil-fenil-cinâmico.

5.6.2. Série 2: Adição de grupos polares

A segunda série (2a a 2d) (Figura 38) traz duas modificações no arcabouço inicial: a primeira é a adição de um grupo carboxilato na posição *meta* do anel fenila. Essa substituição é comum aos 4 ligantes propostos na série e objetivava, inicialmente, o estabelecimento de uma ponte salina entre o carboxilato e o grupo guanidínio do resíduo de aminoácido Arg737 ou Arg729. A segunda modificação, variável entre os ligantes propostos, é um grupo espaçador entre o ácido cinâmico e o anel fenila.

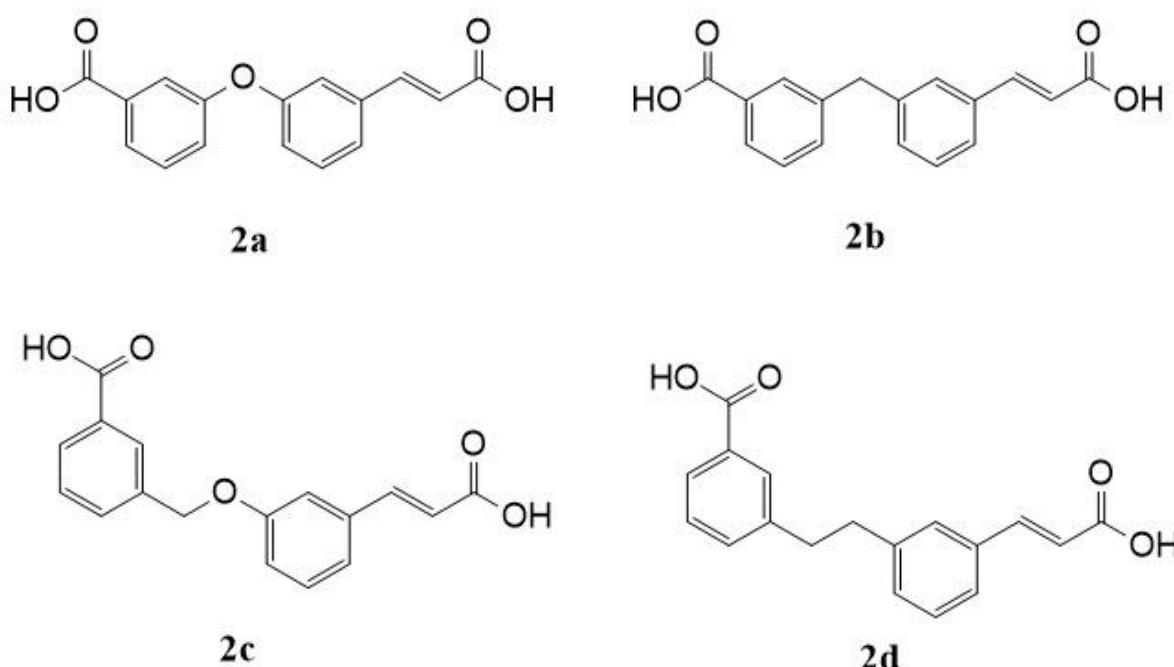


Figura 38: Estruturas 2D dos ligante propostos na segunda série de otimização do ácido cinâmico

A variação dos espaçadores tem duas justificativas: primeiro, a substituição do carbono ligado ao ácido cinâmico por um oxigênio facilita a rota sintética; segundo, para que a posição geométrica do carboxilato seja ideal para interagir com o grupo guanidínio, possivelmente, o espaçador deveria ser diminuído.

Os resultados da segunda rodada de otimizações são apresentados na Tabela 7.

Analizando, inicialmente, as pontuações da docagem, não há grandes variações. O pequeno aumento no caso do ligante **2d** pode ser devido ao aumento do tamanho do ligante, fazendo, teoricamente, mais interações. Esse é um viés comentado anteriormente. Além disso, a previsão de que para realizar interações mais favoráveis seria necessário diminuir o tamanho do espaçador se confirma ao observar as entalpias de interação relativas dos ligantes **2c** e **2d**.

Tabela 7: Pontuações de docagem e entalpias de interação para os novos derivados do ácido *m*-fenil-etyl-cinâmico.

	Pontuação (ChemPLP)	ΔH interação (kcal/mol)	ΔH interação relativo (kcal/mol)
Ácido <i>m</i>-fenil-etyl-cinâmico	75,22	-82	0
Fragmentos Otimizados			
2a	73,29	-94	-12
2b	68,16	-87	-5
2c	73,23	-63	19
2d	75,31	-69	13

Observando as interações estabelecidas para o ligante que teve melhor entalpia de interação, o ligante **2a** manteve as mesmas interações na região do ácido cinâmico (Figura 40) O oxigênio do grupo éter do ligante não estabelece interações intermoleculares com nenhum resíduo de aminoácido, o que leva à análise de que o caráter do espaçador não é importante no ponto de vista de interações intermoleculares, mas sim seu tamanho. É interessante, porém, que a entalpia de interação melhore 7 kcal/mol ao trocar o metileno pelo oxigênio, talvez a explicação esteja no efeito de ressonância, de forma que o aumento da densidade eletrônica nos carboxilatos pela ação do oxigênio reforça as ligações de hidrogênio.

O carboxilato ligado à fenila está, de acordo com a previsão da geometria otimizada pelo método PM7, envolvido em ponte salina com a Arg729, essencial para a catálise da RdRp, sendo esse resultado muito promissor. Ademais, ocorre uma ligação de hidrogênio com a cadeia lateral do resíduo de aminoácido Thr794. Para o ligante **2b**, as mesmas interações se repetem. As tabelas das interações e distâncias previstas pelo PLIP para os ligantes **2a** e **2b** estão nos anexos 11A e 11B.

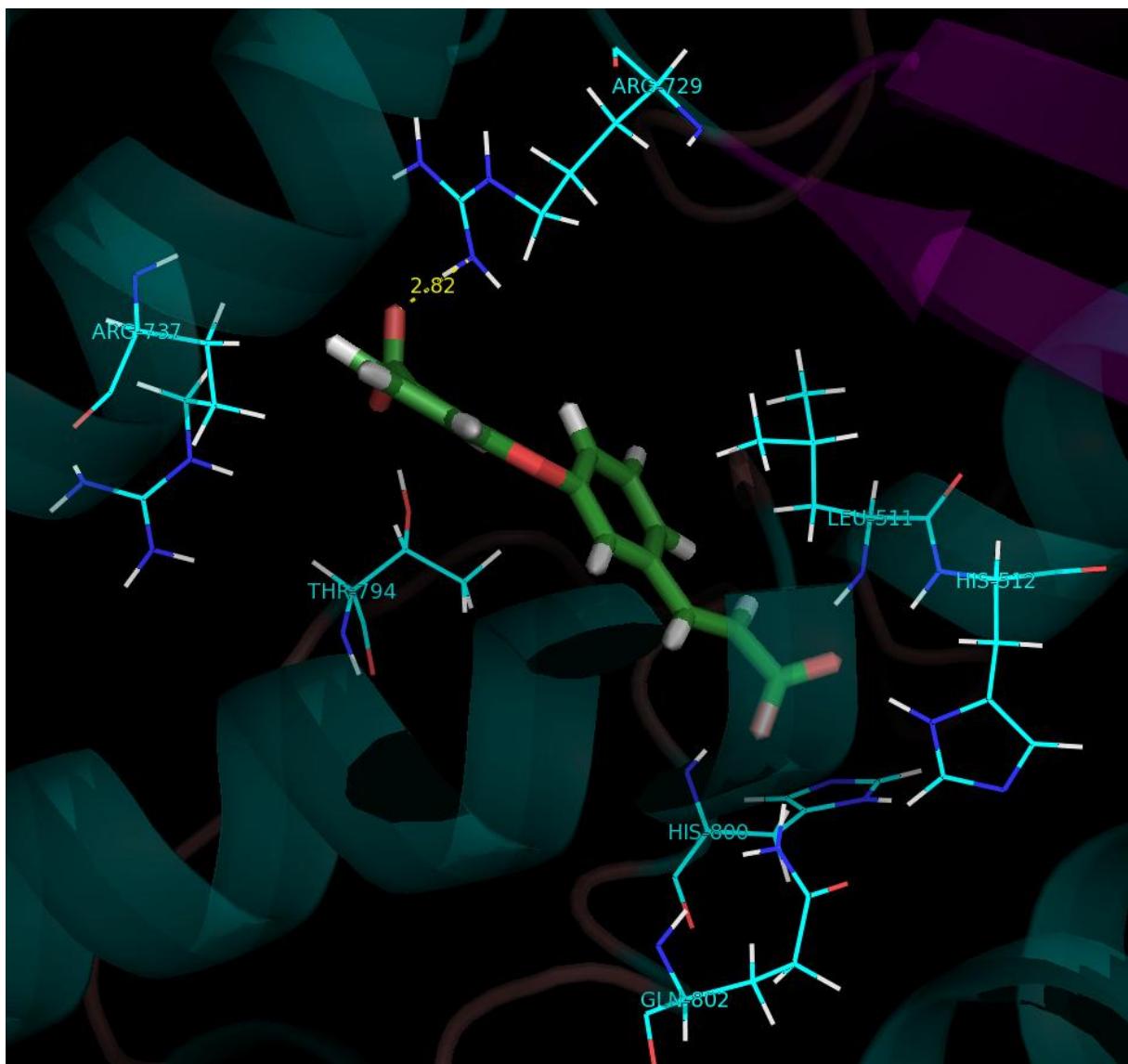


Figura 39: Previsão de pose do ligante ácido (*E*)-3-(3-(2-carboxivinil)fenoxi)benzóico com carbonos em verde, resíduos de aminoácidos com carbonos em ciano. Nitrogênios em azul, oxigênios em vermelho, enxofres em amarelo e hidrogênios em branco. (Figura gerada com o programa PyMol 0.99rc6)

5.6.3. Avaliação ADMET dos derivados propostos

Um dos objetivos das modificações estruturais era obter propriedades farmacocinéticas mais adequadas. Os melhores colocados das séries 1 (**1h** e **1l**) e 2 (**2a** e **2b**) foram analisados. Na tabela 8 pode-se perceber que nenhum critério de Lipinski é violado.

Os descritores ADMET calculados para os ligantes otimizados (Tabela 8) com o servidor Swiss-ADME, estão disponíveis nos anexos 13A a 13D.

Tabela 8: Análise dos derivados selecionados pela regra dos cinco de Lipinski

Código	Critério I	Critério II	Critério III	Critério IV
1h	1 doador	2 aceptores	218,29 g/mol	LogP: 3,48
1l	1 doador	2 aceptores	252,31 g/mol	LogP: 4,22
2a	2 doadores	5 aceptores	284,26 g/mol	LogP: 2,71
2b	2 doadores	4 aceptores	282,29 g/mol	LogP: 2,93

Das moléculas grandes anteriormente selecionadas e avaliadas (tópico 5.5.1), a maior preocupação era a previsão de inibição das CYPs. Nos fragmentos otimizados, o perfil de toxicidade também é aprimorado, de forma que, como pode ser observado na Tabela 9, com os ligantes **2a** e **2b** chega-se em moléculas sem previsão de inibição de nenhuma CYP, o que, inclusive, mostra aprimoramento em relação ao perfil dos próprios inibidores de referência, **27** e **29**.

Tabela 9: Previsão de inibição das CYPs pelos ligantes selecionados, gerado pelo servidor Swiss-ADME.

Código	CYP1A2	CYP2C19	CYP2C9	CYP2D6	CYP3A4
1h	Não	Sim	Não	Não	Não
1l	Não	Não	Não	Sim	Não
2a	Não	Não	Não	Não	Não
2b	Não	Não	Não	Não	Não

6. Conclusões

Neste trabalho, aplicaram-se estratégias de SBDD e FBDD combinadas, utilizando triagem virtual por docagem, otimização de fragmentos e cálculos semiempíricos para a identificação de potenciais inibidores da NS5, proteína-chave do vírus da dengue. A triagem inicial de 657 candidatos, oriundos de uma quimioteca inédita, não explorada por outros grupos e com grande quantidade de produtos naturais, levou à seleção de 17 fragmentos com massa inferior a 300 Da e 6 moléculas maiores.

Os fragmentos foram otimizados por estratégias de crescimento e reavaliação por docagem. Os cálculos de entalpia de interação via MOPAC, utilizando o hamiltoniano PM7, revelaram que diversos derivados apresentaram valores mais favoráveis do que seus fragmentos de origem, indicando ganhos na interação proteína-ligante. Destacaram-se compostos que exibiram entalpias de interação significativamente negativas, como o ácido (*E*)-3-(3-(2-carboxivinil)fenoxi)benzóico, com -96 kcal/mol, derivado do ácido cinâmico, produto natural abundante. O resultado sugere potencial competitivo em relação a inibidores descritos na literatura, de acordo com os fatores que aqui foram analisados.

Quanto aos ligantes grandes, com suas estruturas completas, quando submetidos à etapa de avaliação *in silico* de propriedades ADMET demonstraram, em grande parte, perfis compatíveis com a biodisponibilidade oral, cumprindo critérios de Lipinski. Alguns candidatos também apresentaram potencial de toxicidade predita comparável a de inibidores conhecidos, reforçando sua viabilidade como pontos de partida para posterior otimização. Pela entalpia de interação, destaca-se o ligante (*E*)-3-((2S,3S)-2-(4-hidroxi-3-metoxifenil)-3-(hidroximetil)-7-metoxi-2,3-di-hidrobenzofuran-5-il)acrilaldeído, com -71 kcal/mol. Pelo perfil ADMET, destaca-se a série PPGQ2000701.

As limitações inerentes às abordagens computacionais são reconhecidas, de forma que os compostos identificados fornecem um ponto de partida importante para síntese e ensaios biológicos, podendo servir como base para futuras otimizações guiadas por dados experimentais.

Portanto, este estudo contribui para o avanço da busca por antivirais contra a dengue, reforçando o potencial das técnicas de SBDD e FBDD como uma abordagem ágil, de baixo custo e de alto impacto no desenvolvimento inicial de fármacos.

7. Referências Bibliográficas

ADASME, Melissa F. *et al.* PLIP 2021: expanding the scope of the protein–ligand interaction profiler to DNA and RNA. **Nucleic Acids Research**, v. 49, n. W1, p. W530–W534, 2 jul. 2021.

ALTAMISH, Mohammad *et al.* Therapeutic Potential of Medicinal Plants against Dengue Infection: A Mechanistic Viewpoint. **ACS Omega**, v. 7, n. 28, p. 24048–24065, 19 jul. 2022.

ALVEZ, Luciana Vignólio. **Análise dos constituintes químicos do óleo essencial de Eugenia speciosa Camb. (Myrtaceae). [S.l.: S.n.]**.

APPLEBY, Todd C. *et al.* Structural basis for RNA replication by the hepatitis C virus polymerase. **Science**, v. 347, n. 6223, p. 771–775, 13 fev. 2015.

ASHOUR, Joseph *et al.* NS5 of Dengue Virus Mediates STAT2 Binding and Degradation. **Journal of Virology**, v. 83, n. 11, p. 5408–5418, jun. 2009.

BAELL, Jonathan *et al.* Ask the experts: past, present and future of the rule of five. **Future medicinal chemistry**, v. 5, n. 7, p. 745–752, maio 2013.

BARREIRO, Eliezer J.; FRAGA, Carlos Alberto Manssour. **Química Medicinal - As Bases Moleculares da Ação dos Fármacos**. 3. ed. [S.l.]: Artmed editora, 2014.

BdEC - Apresentação, Características, Diferenciais. Disponível em: <<https://bdec.dotlib.com.br/ajuda/caracteristicas-diferenciais.pdf>>. Acesso em: 25 fev. 2025.

BEESE, Lorena S.; STEITZ, Thomas A. Structural basis for the 3'-5' exonuclease activity of Escherichia coli DNA polymerase I: a two metal ion mechanism. **The EMBO Journal**, v.10, n. 1, p. 245–253, 1 jan. 1991.

BERMAN, H. M. The Protein Data Bank. **Nucleic Acids Research**, v. 28, n. 1, p. 235–242, 1 jan. 2000.

BRUENN, Jeremy A. A structural and primary sequence comparison of the viral RNA-dependent RNA polymerases. **Nucleic Acids Research**, v. 31, n. 7, p. 1821–1829, 1 abr. 2003.

BUTCHER, Sarah J. *et al.* A mechanism for initiating RNA-dependent RNA polymerization. **Nature** 2001 **410:6825**, v. 410, n. 6825, p. 235–240, 8 mar. 2001.

CAMBRIDGE CRYSTALLOGRAPHIC DATA CENTRE. **GOLD User Guide** , 2022.

CHIU, Wei-Wei; KINNEY, Richard M.; DREHER, Theo W. Control of Translation by the 5-and 3-Terminal Regions of the Dengue Virus Genome. **Journal of Virology**, v. 79, n. 13, p. 8303–8315, 2005.

CUSTODIO, Rogério *et al.* Quatro alternativas para resolver a equação de Schrödinger para o átomo de hidrogênio. **Química Nova**, v. 25, n. 1, p. 159–170, fev. 2002.

DAINA, A.; MICHELIN, O.; ZOETE, V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. **Scientific Reports**, v. 7, p. 42717, mar. 2017.

DA SILVA ROCHA, Sheisi F. L. *et al.* Virtual Screening Techniques in Drug Discovery: Review and Recent Applications. **Current Topics in Medicinal Chemistry**, v. 19, n. 19, p. 1751–1767, 16 ago. 2019.

DEWAR, Michael J. S. *et al.* Comments on a comparison of AM1 with the recently developed PM3 method. **Journal of Computational Chemistry**, v. 11, n. 4, p. 541–542, 1990.

DJAKOVIĆ SEKULIĆ, Tatjana *et al.* Chromatographic and in silico assessment of logP measures for new spirohydantoin derivatives with anticancer activity. **Journal of Chemometrics**, v. 32, n. 4, p. e2991, 1 abr. 2018.

DUARTE, Julio Cesar *et al.* A Survey of Basic Concepts and Applications of Machine Learning to Chemistry. **Journal of the Brazilian Chemical Society**, v. 36, n. 8, p. 1-20, 2025.

EDGIL, Dianna; POLACEK, Charlotta; HARRIS, Eva. Dengue Virus Utilizes a Novel Strategy for Translation Initiation When Cap-Dependent Translation Is Inhibited. **Journal of Virology**, v. 80, n. 6, p. 2976–2986, 15 mar. 2006.

ELDRIDGE, Matthew D. *et al.* Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. **Journal of Computer-Aided Molecular Design**, v. 11, n. 5, p. 425–445, 1997a.

ELDRIDGE, Matthew D. *et al.* Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. **Journal of Computer-Aided Molecular Design**, v. 11, n. 5, p. 425–445, 1997b.

FELICETTI, Tommaso *et al.* Functionalized sulfonyl anthranilic acid derivatives inhibit replication of all the four dengue serotypes. **European Journal of Medicinal Chemistry**, v. 252, p. 115283, 5 abr. 2023.

FERREIRA, Denise de Castro Gomes. Síntese e modificações estruturais de derivados curcuminóides e avaliação de suas atividades anti-leishmania. 2000. 188 f. Tese (Doutorado em Química) - Instituto de Ciências Exatas, Universidade Federal Rural do Rio de Janeiro, Seropédica-RJ, 2000.

Ferreira, Welisson da Silva. Planejamento, síntese e avaliação das atividades antifúngica e tripanocida da piperina e derivados do núcleo 1,3,4-Oxadiazol-2-Tiona. 2014. [251 f]. Tese (Programa de Pós-Graduação em Química) - Universidade Federal Rural do Rio de Janeiro, [Seropédica-RJ] .

FERRON, François *et al.* The viral RNA capping machinery as a target for antiviral drugs. **Antiviral Research**, v. 96, n. 1, p. 21–31, 1 out. 2012.

GUEX, Nicolas; PEITSCH, Manuel C. SWISS-MODEL and the Swiss-Pdb Viewer: An environment for comparative protein modeling. **Electrophoresis**, v. 18, n. 15, p. 2714–2723, 1 jan. 1997.

HAHN, Chang S. *et al.* Conserved elements in the 3' untranslated region of flavivirus RNAs and potential cyclization sequences. **Journal of Molecular Biology**, v. 198, n. 1, p. 33–41, 5 nov. 1987.

HANNEMANN, Holger *et al.* Serotype-specific Differences in Dengue Virus Non-structural Protein 5 Nuclear Localization. **The Journal of Biological Chemistry**, v. 288, n. 31, p. 22621, 2 ago. 2013.

HARNER, Mary J.; FRANK, Andreas O.; FESIK, Stephen W. Fragment-based drug discovery using NMR spectroscopy. **Journal of Biomolecular NMR**, v. 56, n. 2, p. 65–75, 18 jun. 2013.

HODGE, Kenneth *et al.* **Flavors of Flaviviral RNA Structure: towards an Integrated View of RNA Function from Translation through Encapsidation.** **BioEssays** John Wiley and Sons Inc., v. 41, p.1900003, 1 ago. 2019.

HOLDEN, Katherine Lynn *et al.* Inhibition of dengue virus translation and RNA synthesis by a morpholino oligomer targeted to the top of the terminal 3' stem-loop structure. **Virology**, v. 344, n. 2, p. 439–452, 20 jan. 2006.

HOWLEY *et al.* **Fields Virology: Emerging Viruses.** 7. ed. [S.l.]: Wolters Kluwer, 2021.

IBRAHIM, Tamer M.; BAUER, Matthias R.; BOECKLER, Frank M. Applying DEKOIS 2.0 in structure-based virtual screening to probe the impact of preparation procedures and score normalization. **Journal of Cheminformatics**, v. 7, n. 1, p. 21, 4 dez. 2015.

INTERNATIONAL COMMITTEE ON TAXONOMY OF VIRUSES. **Current ICTV Taxonomy Release | ICTV.** Disponível em: <<https://ictv.global/taxonomy>>. Acesso em: 10 set. 2024.

ISSUR, Moheshwarnath *et al.* The flavivirus NS5 protein is a true RNA guanylyltransferase that catalyzes a two-step reaction to form the RNA cap structure. **RNA**, v. 15, n. 12, p. 2340–2350, 1 dez. 2009.

JENSEN, Frank. **Introduction to Computational Chemistry**. 3. ed. [S.l.]: Wiley, 2017.

JHOTI, Harren *et al.* The “rule of three” for fragment-based drug discovery: Where are we now? **Nature Reviews Drug Discovery**, v. 12, n. 8, p. 644, 12 ago. 2013.

JONES, Gareth *et al.* Development and validation of a genetic algorithm for flexible docking. **Journal of Molecular Biology**, v. 267, n. 3, p. 727–748, 4 abr. 1997.

JONES, Meleri *et al.* Dengue Virus Inhibits Alpha Interferon Signaling by Reducing STAT2 Expression. **Journal of Virology**, v. 79, n. 9, p. 5414–5420, maio 2005.

KIRSCH, Philine *et al.* Concepts and Core Principles of Fragment-Based Drug Design. **Molecules 2019, Vol. 24, Page 4309**, v. 24, n. 23, p. 4309, 26 nov. 2019.

KORB, Oliver; STÜTZLE, Thomas; EXNER, Thomas E. Empirical Scoring Functions for Advanced Protein–Ligand Docking with PLANTS. **Journal of Chemical Information and Modeling**, v. 49, n. 1, p. 84–96, 26 jan. 2009.

LEE, Jooyoung; FREDDOLINO, Peter L.; ZHANG, Yang. Ab Initio Protein Structure Prediction. In: **From Protein Structure to Function with Bioinformatics**. Dordrecht: Springer Netherlands, 2017. p. 3–35.

LI, Qingxin. Application of Fragment-Based Drug Discovery to Versatile Targets. **Frontiers in Molecular Biosciences**, v. 7, p. 552427, 5 ago. 2020.

Lianda, Regina Lucia Pelachim. Caracterização de mel de *Apis mellifera* pelo seu perfil em substâncias fenólicas por cromatografia líquida de alta eficiência e avaliação da atividade biológica. 2004. 142 f. Dissertação (Programa de Pós-Graduação em Química) - Universidade Federal Rural do Rio de Janeiro, Seropédica-RJ.

LIM, Siew Pheng *et al.* Potent Allosteric Dengue Virus NS5 Polymerase Inhibitors: Mechanism of Action and Resistance Profiling. **PLOS Pathogens**, v. 12, n. 8, p. e1005737, 1 ago. 2016.

LIM, Siew Pheng *et al.* Discovery of Potent Non-nucleoside Inhibitors of Dengue Viral RNA-Dependent RNA Polymerase from Fragment Screening and Structure-Guided Design. In: HILGENFELD, R.; VASUDEVAN, S. (Orgs.). **Advances in Experimental Medicine and Biology**. [S.l.]: Springer, Singapore, 2018. v. 1062 p. 187–198.

LIM, Siew Pheng. Dengue drug discovery: Progress, challenges and outlook. **Antiviral Research**, v. 163, p. 156–178, mar. 2019.

LIPINSKI, Christopher A. *et al.* Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. **Advanced drug delivery reviews**, v. 46, n. 1–3, p. 3–26, 1 mar. 2001.

LIPKOWITZ, Kenny B.; BOYD, Donald B. (ORG.). **Reviews in Computational Chemistry**. [S.l.]: Wiley, 1991. v. 2

LOPES, Natália Drumond. **Síntese e avaliação da atividade biológica de compostos mesoiônicos da classe dos 1,3,4-tiadiazólio-2-arylaminidas**. [S.l.: S.n.].

LU, Guoliang; GONG, Peng. A structural view of the RNA-dependent RNA polymerases from the Flavivirus genus. *Virus Research* Elsevier B.V., 15 abr. 2017.

MINISTÉRIO DA SAÚDE. **Situação Epidemiológica.** Disponível em: <<https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/a/arboviroses/informe-semanal/2025/informe-semanal-no-17.pdf/view>>. Acesso em 24 jul. 2025.

MINISTÉRIO DA SAÚDE. **Atualização de Casos de Arboviroses.** Disponível em: <<https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/a/aedes-aegypti/monitoramento-das-arboviroses>>. Acesso em 21 out. 2024.

NJOROGE, F. George *et al.* Challenges in Modern Drug Discovery: A Case Study of Boceprevir, an HCV Protease Inhibitor for the Treatment of Hepatitis C Virus Infection. **Accounts of Chemical Research**, v. 41, n. 1, p. 50–59, 1 jan. 2008.

NOBLE, Christian G. *et al.* A Conserved Pocket in the Dengue Virus Polymerase Identified through Fragment-based Screening. **The Journal of biological chemistry**, v. 291, n. 16, p. 8541–8548, 15 abr. 2016.

OLIVEIRA, Cintia Crysiane da Silva de; NETO, Paulo de Oliveira Paes de Lira. Vacina da dengue x sorotipo circulante: uma discussão da cobertura vacinal de acordo com a epidemiologia das regiões do Brasil. **Revista JRG de Estudos Acadêmicos**, v. 7, n. 14, p. e14951–e14951, 5 mar. 2024.

PATEL, Sanjay S. *et al.* Clinical Safety Experience of TAK-003 for Dengue Fever: A New Tetravalent Live Attenuated Vaccine Candidate. **Clinical Infectious Diseases**, v. 76, n. 3, p. e1350–e1359, 8 fev. 2023.

PITISUTTITHUM, Punnee; BOUCKENOOGHE, Alain. The first licensed dengue vaccine: an important tool for integrated preventive strategies against dengue virus infection. **Expert review of vaccines**, v. 15, n. 7, p. 795–798, 2 jul. 2016.

PRESTES-CARNEIRO, Luiz Euribel *et al.* Dengue outbreaks in a city with recent transmission in São Paulo state, Brazil. **International Journal of Environmental Health Research**, v. 34, n. 3, p. 1370–1383, 3 mar. 2024.

PRYOR, Melinda J. *et al.* Nuclear Localization of Dengue Virus Nonstructural Protein 5 Through Its Importin α/β -Recognized Nuclear Localization Sequences is Integral to Viral Infection. **Traffic**, v. 8, n. 7, p. 795–807, 1 jul. 2007.

RARITAN, N. J. **Johnson & Johnson to Discontinue Phase 2 Field Study Evaluating Investigational Antiviral for the Prevention of Dengue**. Disponível em: <<https://www.jnj.com/media-center/press-releases/johnson-johnson-to-discontinue-phase-2-field-study-evaluating-investigational-antiviral-for-the-prevention-of-dengue>>. Acesso em: 28 set. 2025.

RAWLINSON, Stephen M. *et al.* CRM1-mediated nuclear export of dengue virus RNA polymerase NS5 modulates interleukin-8 induction and virus production. **Journal of Biological Chemistry**, v. 284, n. 23, p. 15589–15597, 5 jun. 2009.

RCSB.ORG. **Protein Data Bank**. Disponível em: <rscb.org>. Acesso em 30 set. 2023.

REID, David W. *et al.* Dengue Virus Selectively Annexes Endoplasmic Reticulum-Associated Translation Machinery as a Strategy for Co-opting Host Cell Protein Synthesis. **Journal of virology**, v. 92, n. 7, abr. 2018.

ROCHA, Vinícius N.; SANT'ANNA, Carlos M. R. **From Origin to Current Methods: An Overview of Molecular Modeling Applied to Medicinal Chemistry in the Last 30 Years.** **Journal of the Brazilian Chemical Society** Sociedade Brasileira de Química, v. 35, n. 10, p. e-20240103, 2024.

RODRIGUES, Arthur Valbon. Síntese de bases de Schiff, polivinilpirrolidona e obtenção de extrato vegetal como inibidores verdes de corrosão do aço carbono em meio ácido. 2020. 155 f. Tese (Mestrado em Química) - Instituto de Química, Universidade Federal Rural do Rio de Janeiro, Seropédica, 2020.

RONEY, Miah; MOHD ALUWI, Mohd Fadhlizil Fasihi. Computational studies demonstrating dithymoquinone of *Nigella sativa* as a potential anti-dengue agent: Short review. **Intelligent Pharmacy**, v. 2, n. 3, p. 335–338, 1 jun. 2024.

SALGUEIRO, Fernanda Barbosa. **Caracterização da própolis verde brasileira: substâncias fenólicas, atividade biológica e análise quimiométrica.** [S.l.: S.n.].

SCANTLEBURY, Jack *et al.* A Small Step Toward Generalizability: Training a Machine Learning Scoring Function for Structure-Based Virtual Screening. **J. Chem. Inf. Model**, v. 63, p. 2960–2974, 2023.

SECRETARIA DE COMUNICAÇÃO SOCIAL. **Vacina contra dengue já está incorporada ao Programa Nacional de Imunização.** Disponível em: <<https://www.gov.br/secom/pt-br/fatos/brasil-contra-fake/noticias/2024/vacina-contra-dengue-ja-esta-incorporada-ao-programa-nacional-de-imunizacao>>. Acesso em 08 jan. 2024.

Semiempirical Calculations. *In: Computational Chemistry.* [S.l.]: Springer, Boston, MA, 2004. p. 339–379.

Silva, Ricardo Figueira da. Investigação de ácidos fenólicos em amostras de mel por cromatografia líquida de alta eficiência e sua aplicação na caracterização da origem floral. 2004. 82 f. Dissertação (Programa de Pós-Graduação em Química) - Universidade Federal Rural do Rio de Janeiro, Seropédica-RJ .

SOUZA, Marco Andre Alves de. Expressão gênica relacionada à produção de óleo essencial e avaliação do metabolismo de *Mentha arvensis* L. sob diferentes condições de cultivo. 2010. 106 f. Tese (Programa de Pós-Graduação em Química) - Universidade Federal Rural do Rio de Janeiro, Seropédica-RJ .

SOUZA, Rafael A. C.; OLIVEIRA, Carolina G. Investigação teórica por Docking molecular e parâmetros farmacocinéticos do potencial antibacteriano de complexos de manganês(II) com tiossemicarbazonas. **Open Science Research IIN**, v. 2, p. 577-593, 2022.

STEWART, James J. P. Reply to “Comments on a comparison of AM1 with the recently developed PM3 method”. **Journal of Computational Chemistry**, v. 11, n. 4, p. 543–544, 1990a.

STEWART, James J. P. MOPAC: a semiempirical molecular orbital program. **Journal of computer-aided molecular design**, v. 4, n. 1, p. 1–103, mar. 1990b.

STEWART, James J. P.; SEILER, Frank J. MOPAC: A semiempirical molecular orbital programJournal of Computer-Aided Molecular Design. **Journal of Computer-Aided Molecular Design**, v. 4, p. 1-105, 1990.

TAY, Moon Y. F. *et al.* The C-terminal 50 amino acid residues of dengue NS3 protein are important for NS3-NS5 interaction and viral replication. **Journal of Biological Chemistry**, v. 290, n. 4, p. 2379–2394, 23 jan. 2015.

TE VELTHUIS, Aartjan J. W. Common and unique features of viral RNA-dependent polymerases. **Cellular and Molecular Life Sciences**, v. 71, n. 22, p. 4403–4420, 24 out. 2014.

TIDMAN, Rachel; ABELA-RIDDER, Bernadette; DE CASTAÑEDA, Rafael Ruiz. The impact of climate change on neglected tropical diseases: a systematic review. **Transactions of The Royal Society of Tropical Medicine and Hygiene**, v. 115, n. 2, p. 147–168, 28 jan. 2021.

TORRES, Pedro H. M. *et al.* **Key topics in molecular docking for drug design**. International Journal of Molecular Sciences, v. 20, p. 4574, 2 set. 2019.

TOUNGE, Brett A.; PARKER, Michael H. Designing a Diverse High-Quality Library for Crystallography-Based FBDD Screening. **Methods of Enzymology**, v. 493, p. 3–20, 2011.

UNOH, Yuto *et al.* Discovery of S-217622, a Noncovalent Oral SARS-CoV-2 3CL Protease Inhibitor Clinical Candidate for Treating COVID-19. **Journal of Medicinal Chemistry**, v. 65, n. 9, p. 6499–6512, 2022.

WEI, Chaochun *et al.* Discovery of novel inhibitors of dengue viral RNA-dependent RNA polymerase by molecular docking, *in vitro* assay, DFT, and MD simulations. **International Journal of Biological Macromolecules**, v. 305, p. 141328, 1 maio 2025.

WILDEY, Mary Jo *et al.* High-Throughput Screening. **Annual Reports in Medicinal Chemistry**, v. 50, p. 149–195, 2017.

WORLD HEALTH ORGANIZATION. **Dengue and Severe Dengue**. Disponível em: <<https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue>>. Acesso em 8 out. 2023.

WU, Jiqin *et al.* A Structural Overview of RNA-Dependent RNA Polymerases from the Flaviviridae Family. **International Journal of Molecular Sciences**, v. 16, n. 6, p. 12943, 8 jun. 2015.

YAP, Thai Leong *et al.* Crystal Structure of the Dengue Virus RNA-Dependent RNA Polymerase Catalytic Domain at 1.85-Angstrom Resolution. **Journal of Virology**, v. 81, n. 9, p. 4753–4765, maio 2007.

YOU, Shihyun; PADMANABHAN, R. A novel in vitro replication system for Dengue virus. Initiation of RNA synthesis at the 3'-end of exogenous viral RNA templates requires 5'- and 3'-terminal complementary sequence motifs of the viral RNA. **The Journal of biological chemistry**, v. 274, n. 47, p. 33714–33722, 19 nov. 1999.

ZHANG, Jinze *et al.* Holo Protein Conformation Generation from Apo Structures by Ligand Binding Site Refinement. **Journal of Chemical Information and Modeling**, v. 62, n. 22, p. 5806–5820, 28 nov. 2022.

ZHAO, Yongqian *et al.* A Crystal Structure of the Dengue Virus NS5 Protein Reveals a Novel Inter-domain Interface Essential for Protein Flexibility and Virus Replication. **PLOS Pathogens**, v. 11, n. 3, p. e1004682, 1 mar. 2015.

ZONG, Keli *et al.* Identification of novel inhibitors of dengue viral NS5 RNA-dependent RNA polymerase through molecular docking, biological activity evaluation and molecular dynamics simulations. **Journal of Enzyme Inhibition and Medicinal Chemistry**, v. 40, n. 1, p. 2463006, 2025.

ZOU, Gang *et al.* Functional analysis of two cavities in flavivirus NS5 polymerase. **Journal of Biological Chemistry**, v. 286, n. 16, p. 14362–14372, 22 abr. 2011.

8. ANEXOS

ANEXO 1A: Interações previstas pelo PLIP (Adasme et al., 2021) para o fragmento MPPGQ2004R0105

▼ Hydrophobic Interactions

Index	Residue	AA	Distance	Ligand Atom	Protein Atom
1	511A	LEU	3.72	13691	7703
2	511A	LEU	3.35	13705	7700
3	794A	THR	3.96	13695	12220
4	799A	ALA	3.72	13693	12303

▼ Hydrogen Bonds —

Index	Residue	AA	Distance H-A	Distance D-A	Donor Angle	Protein donor?	Side chain	Donor Atom	Acceptor Atom
1	511A	LEU	2.12	2.85	126.02	✓	✗	7696 [Nam]	13708 [O.co2]
2	512A	HIS	1.93	2.94	168.71	✓	✗	7715 [Nam]	13708 [O.co2]
3	794A	THR	2.21	3.16	160.86	✓	✓	12219 [O3]	13701 [O2]
4	800A	HIS	2.09	3.10	164.43	✓	✗	12309 [Nam]	13709 [O.co2]
5	802A	GLN	1.86	2.80	151.30	✓	✓	12351 [Nam]	13709 [O.co2]

▼ Salt Bridges

Index	Residue	AA	Distance	Protein positive?	Ligand Group	Ligand Atoms
1	512A	HIS	3.93	✓	Carboxylate	13709, 13708
2	800A	HIS	5.03	✓	Carboxylate	13709, 13708

ANEXO 1B: Interações previstas pelo PLIP (Adasme et al., 2021) para o fragmento MPPGQ2004R0106

▼ Hydrophobic Interactions

Index	Residue	AA	Distance	Ligand Atom	Protein Atom
1	511A	LEU	3.81	13691	7703
2	511A	LEU	3.29	13709	7700
3	794A	THR	3.46	13695	12220
4	799A	ALA	3.94	13691	12303

▼ Hydrogen Bonds —

Index	Residue	AA	Distance H-A	Distance D-A	Donor Angle	Protein donor?	Side chain	Donor Atom	Acceptor Atom
1	511A	LEU	1.87	2.75	140.66	✓	✗	7696 [Nam]	13713 [O.co2]
2	512A	HIS	1.99	2.99	166.60	✓	✗	7715 [Nam]	13713 [O.co2]
3	794A	THR	2.11	3.08	166.82	✓	✓	12219 [O3]	13700 [O2]
4	796A	SER	2.14	2.97	140.42	✓	✓	12257 [O3]	13702 [O2]
5	800A	HIS	2.11	3.14	177.81	✓	✗	12309 [Nam]	13712 [O.co2]
6	802A	GLN	2.53	3.41	145.62	✓	✓	12351 [Nam]	13712 [O.co2]

▼ Salt Bridges

Index	Residue	AA	Distance	Protein positive?	Ligand Group	Ligand Atoms
1	512A	HIS	3.88	✓	Carboxylate	13712, 13713
2	800A	HIS	5.17	✓	Carboxylate	13712, 13713

ANEXO 1C: Interações previstas pelo PLIP (Adasme et al., 2021) para o fragmento MPPGQ2004R0108

▼ Hydrophobic Interactions

Index	Residue	AA	Distance	Ligand Atom	Protein Atom
1	512A	LEU	3.64	8955	3218
2	512A	LEU	3.23	8964	3215
3	794A	THR	3.89	8957	7379
4	799A	ALA	3.99	8964	7462
5	803A	TRP	3.82	8966	7535

▼ Hydrogen Bonds —

Index	Residue	AA	Distance H-A	Distance D-A	Donor Angle	Protein donor?	Side chain	Donor Atom	Acceptor Atom
1	512A	LEU	2.06	2.70	117.66	✓	✗	3211 [Nam]	8969 [O.co2]
2	513A	HIS	1.99	2.98	161.34	✓	✗	3230 [Nam]	8969 [O.co2]
3	796A	SER	2.38	3.06	124.73	✓	✓	7416 [O3]	8962 [O2]
4	798A	HIS	2.94	3.71	135.71	✗	✗	8962 [O2]	7444 [O2]
5	800A	LYS	1.76	2.80	167.45	✓	✗	7468 [Nam]	8970 [O.co2]

▼ Salt Bridges

Index	Residue	AA	Distance	Protein positive?	Ligand Group	Ligand Atoms
1	513A	HIS	3.89	✓	Carboxylate	8970, 8969

ANEXO 1D: Interações previstas pelo PLIP (Adasme et al., 2021) para o fragmento MPPGQ2004R0109

▼ Hydrophobic Interactions

Index	Residue	AA	Distance	Ligand Atom	Protein Atom
1	511A	LEU	3.68	9446	3454
2	511A	LEU	3.23	9456	3451
3	799A	ALA	3.69	9444	8069
4	803A	TRP	3.43	9454	8139

▼ Hydrogen Bonds —

Index	Residue	AA	Distance H-A	Distance D-A	Donor Angle	Protein donor?	Side chain	Donor Atom	Acceptor Atom
1	511A	LEU	2.04	2.90	138.94	✓	✗	3447 [Nam]	9460 [O.co2]
2	512A	HIS	2.17	3.18	167.23	✓	✗	3466 [Nam]	9460 [O.co2]
3	800A	HIS	2.08	3.09	166.45	✓	✗	8075 [Nam]	9459 [O.co2]
4	802A	GLN	1.78	2.76	156.06	✓	✓	8117 [Nam]	9459 [O.co2]

▼ Salt Bridges

Index	Residue	AA	Distance	Protein positive?	Ligand Group	Ligand Atoms
1	512A	HIS	4.21	✓	Carboxylate	9459, 9460
2	800A	HIS	4.96	✓	Carboxylate	9459, 9460

ANEXO 1E: Interações previstas pelo PLIP (Adasme et al., 2021) para o fragmento MPPGQ2004R0110

▼ Hydrophobic Interactions

Index	Residue	AA	Distance	Ligand Atom	Protein Atom
1	511A	LEU	3.66	9446	3454
2	511A	LEU	3.16	9453	3451
3	794A	THR	3.99	9447	7986

▼ Hydrogen Bonds —

Index	Residue	AA	Distance H-A	Distance D-A	Donor Angle	Protein donor?	Side chain	Donor Atom	Acceptor Atom
1	511A	LEU	2.16	2.91	128.03	✓	✗	3447 [Nam]	9458 [O.co2]
2	512A	HIS	2.17	3.16	162.44	✓	✗	3466 [Nam]	9458 [O.co2]
3	800A	HIS	1.93	2.95	167.52	✓	✗	8075 [Nam]	9459 [O.co2]
4	802A	GLN	1.94	2.91	154.52	✓	✓	8117 [Nam]	9459 [O.co2]

▼ Salt Bridges

Index	Residue	AA	Distance	Protein positive?	Ligand Group	Ligand Atoms
1	512A	HIS	3.96	✓	Carboxylate	9459, 9458
2	800A	HIS	5.13	✓	Carboxylate	9459, 9458

ANEXO 2A: Interações previstas pelo PLIP (Adasme et al., 2021) para o fragmento PPGQ2010R0101

▼ Hydrophobic Interactions

Index	Residue	AA	Distance	Ligand Atom	Protein Atom
1	511A	LEU	3.68	13709	7700
2	794A	THR	3.80	13700	12220
3	799A	ALA	3.94	13706	12303
4	803A	TRP	3.69	13709	12372
5	803A	TRP	3.63	13706	12373

▼ Hydrogen Bonds —

Index	Residue	AA	Distance H-A	Distance D-A	Donor Angle	Protein donor?	Side chain	Donor Atom	Acceptor Atom
1	794A	THR	2.18	3.08	151.61	✓	✓	12219 [O3]	13716 [O3]

ANEXO 2B: Interações previstas pelo PLIP (Adasme et al., 2021) para o fragmento PPGQ2010R0102

▼ Hydrophobic Interactions

Index	Residue	AA	Distance	Ligand Atom	Protein Atom
1	511A	LEU	3.52	9446	3453
2	511A	LEU	3.60	9459	3454
3	511A	LEU	3.43	9460	3451
4	514A	LEU	3.77	9446	3511
5	766A	TYR	3.97	9465	7533
6	798A	HIS	3.53	9446	8052
7	799A	ALA	3.81	9460	8069
8	803A	TRP	3.57	9462	8139

ANEXO 3A: Interações previstas pelo PLIP (Adasme et al., 2021) para o fragmento PPGQ201820209

▼ Hydrophobic Interactions

Index	Residue	AA	Distance	Ligand Atom	Protein Atom
1	511A	LEU	3.59	9446	3454
2	511A	LEU	3.14	9456	3451
3	766A	TYR	3.85	9447	7533
4	803A	TRP	3.65	9458	8139

▼ Hydrogen Bonds —

Index	Residue	AA	Distance H-A	Distance D-A	Donor Angle	Protein donor?	Side chain	Donor Atom	Acceptor Atom
1	511A	LEU	2.20	2.95	128.68	✓	✗	3447 [Nam]	9462 [O.co2]
2	512A	HIS	2.13	3.13	164.09	✓	✗	3466 [Nam]	9462 [O.co2]
3	800A	HIS	1.91	2.92	164.25	✓	✗	8075 [Nam]	9461 [O.co2]
4	802A	GLN	1.94	2.87	150.53	✓	✓	8117 [Nam]	9461 [O.co2]

▼ Salt Bridges

Index	Residue	AA	Distance	Protein positive?	Ligand Group	Ligand Atoms
1	512A	HIS	4.08	✓	Carboxylate	9461, 9462
2	800A	HIS	5.14	✓	Carboxylate	9461, 9462

ANEXO 3B: Interações previstas pelo PLIP (Adasme et al., 2021) para o fragmento PPGQ201820210

▼ Hydrophobic Interactions

Index	Residue	AA	Distance	Ligand Atom	Protein Atom
1	511A	LEU	3.66	9443	3454
2	511A	LEU	3.25	9454	3451
3	799A	ALA	3.70	9448	8069
4	803A	TRP	3.43	9452	8139

▼ Hydrogen Bonds —

Index	Residue	AA	Distance H-A	Distance D-A	Donor Angle	Protein donor?	Side chain	Donor Atom	Acceptor Atom
1	511A	LEU	2.06	2.90	136.85	✓	✗	3447 [Nam]	9458 [O.co2]
2	512A	HIS	2.13	3.14	165.78	✓	✗	3466 [Nam]	9458 [O.co2]
3	800A	HIS	2.06	3.07	165.79	✓	✗	8075 [Nam]	9457 [O.co2]
4	802A	GLN	1.80	2.77	155.14	✓	✓	8117 [Nam]	9457 [O.co2]

▼ Salt Bridges

Index	Residue	AA	Distance	Protein positive?	Ligand Group	Ligand Atoms
1	512A	HIS	4.20	✓	Carboxylate	9457, 9458
2	800A	HIS	4.96	✓	Carboxylate	9457, 9458

ANEXO 4A: Interações previstas pelo PLIP (Adasme et al., 2021) para o fragmento PPGQ2016R0201

▼ Hydrophobic Interactions

Index	Residue	AA	Distance	Ligand Atom	Protein Atom
1	511A	LEU	3.88	13694	7703
2	511A	LEU	3.29	13706	7700
3	799A	ALA	3.56	13697	12303
4	803A	TRP	3.61	13704	12373

▼ Hydrogen Bonds —

Index	Residue	AA	Distance H-A	Distance D-A	Donor Angle	Protein donor?	Side chain	Donor Atom	Acceptor Atom
1	511A	LEU	1.99	2.80	133.18	✓	✗	7696 [Nam]	13709 [O.co2]
2	512A	HIS	2.02	3.03	168.29	✓	✗	7715 [Nam]	13709 [O.co2]
3	766A	TYR	2.00	2.98	171.75	✓	✓	11769 [O3]	13702 [O2]
4	794A	THR	2.31	3.25	157.73	✓	✓	12219 [O3]	13700 [O2]
5	800A	HIS	2.03	3.07	173.95	✓	✗	12309 [Nam]	13710 [O.co2]
6	802A	GLN	3.10	4.07	158.62	✓	✓	12351 [Nam]	13709 [O.co2]

▼ Salt Bridges

Index	Residue	AA	Distance	Protein positive?	Ligand Group	Ligand Atoms
1	512A	HIS	4.11	✓	Carboxylate	13710, 13709
2	800A	HIS	5.08	✓	Carboxylate	13710, 13709

ANEXO 4B: Interações previstas pelo PLIP (Adasme et al., 2021) para o fragmento PPGQ2016R0211

▼ Hydrophobic Interactions

Index	Residue	AA	Distance	Ligand Atom	Protein Atom
1	511A	LEU	3.47	13719	7702
2	514A	LEU	3.59	13719	7760
3	766A	TYR	3.81	13693	11767
4	798A	HIS	3.39	13719	12286
5	799A	ALA	3.69	13692	12303
6	803A	TRP	3.32	13697	12373

▼ Hydrogen Bonds —

Index	Residue	AA	Distance H-A	Distance D-A	Donor Angle	Protein donor?	Side chain	Donor Atom	Acceptor Atom
1	510A	GLY	3.49	4.04	120.05	✗	✗	13705 [O.co2]	7692 [O2]
2	511A	LEU	2.05	2.72	120.30	✓	✗	7696 [Nam]	13705 [O.co2]
3	512A	HIS	2.01	3.02	167.54	✓	✗	7715 [Nam]	13705 [O.co2]
4	766A	TYR	1.75	2.69	154.67	✗	✓	13707 [O2]	11769 [O3]
5	800A	HIS	1.99	3.00	163.49	✓	✗	12309 [Nam]	13706 [O.co2]
6	802A	GLN	2.04	2.96	150.66	✓	✓	12351 [Nam]	13706 [O.co2]

▼ Salt Bridges

Index	Residue	AA	Distance	Protein positive?	Ligand Group	Ligand Atoms
1	512A	HIS	4.08	✓	Carboxylate	13706, 13705
2	800A	HIS	4.20	✓	Carboxylate	13706, 13705

ANEXO 5A: Interações previstas pelo PLIP (Adasme et al., 2021) para o fragmento MPPGQ2004R0302

▼ Hydrophobic Interactions

Index	Residue	AA	Distance	Ligand Atom	Protein Atom
1	512A	LEU	3.93	8954	3218
2	512A	LEU	3.23	8967	3215
3	794A	THR	3.87	8953	7379
4	799A	ALA	3.65	8957	7462
5	803A	TRP	3.55	8965	7535

▼ Hydrogen Bonds —

Index	Residue	AA	Distance H-A	Distance D-A	Donor Angle	Protein donor?	Side chain	Donor Atom	Acceptor Atom
1	512A	LEU	1.86	2.69	132.86	✓	✗	3211 [Nam]	8970 [O.co2]
2	513A	HIS	2.04	3.04	163.92	✓	✗	3230 [Nam]	8970 [O.co2]
3	794A	THR	2.51	3.39	149.35	✓	✓	7378 [O3]	8960 [O2]
4	800A	LYS	2.11	3.15	176.71	✓	✗	7468 [Nam]	8971 [O.co2]

▼ π-Cation Interactions

Index	Residue	AA	Distance	Offset	Protein charged?	Ligand Group	Ligand Atoms
1	711A	HIS	5.58	1.97	✓	Aromatic	8952, 8953, 8954, 8955, 8956, 8957

▼ Salt Bridges

Index	Residue	AA	Distance	Protein positive?	Ligand Group	Ligand Atoms
1	513A	HIS	4.09	✓	Carboxylate	8971, 8970

ANEXO 5B: Interações previstas pelo PLIP (Adasme et al., 2021) para o fragmento MPPGQ2004R0303

▼ Hydrophobic Interactions

Index	Residue	AA	Distance	Ligand Atom	Protein Atom
1	511A	LEU	3.67	13695	7703
2	511A	LEU	3.62	13696	7700
3	799A	ALA	3.63	13696	12303
4	803A	TRP	3.49	13695	12373

▼ Hydrogen Bonds —

Index	Residue	AA	Distance H-A	Distance D-A	Donor Angle	Protein donor?	Side chain	Donor Atom	Acceptor Atom
1	511A	LEU	2.65	3.50	141.35	✓	✗	7696 [Nam]	13701 [O2]
2	766A	TYR	1.67	2.68	168.33	✓	✓	11769 [O3]	13704 [O.co2]
3	794A	THR	2.22	3.10	146.71	✓	✓	12219 [O3]	13705 [O.co2]
4	794A	THR	2.41	3.10	130.21	✗	✓	13705 [O.co2]	12219 [O3]
5	798A	HIS	1.67	2.68	174.67	✗	✗	13701 [O2]	12285 [O2]
6	800A	HIS	3.22	4.02	137.03	✓	✗	12309 [Nam]	13701 [O2]

▼ Salt Bridges

Index	Residue	AA	Distance	Protein positive?	Ligand Group	Ligand Atoms
1	711A	HIS	5.23	✓	Carboxylate	13704, 13705

ANEXO 6A: Interações previstas pelo PLIP (Adasme et al., 2021) para o fragmento PPGQ1998R0105

▼ Hydrophobic Interactions

Index	Residue	AA	Distance	Ligand Atom	Protein Atom
1	511A	LEU	3.65	13702	7705
2	511A	LEU	3.55	13701	7704
3	511A	LEU	3.53	13707	7702
4	798A	HIS	3.86	13701	12289
5	799A	ALA	3.62	13704	12306
6	803A	TRP	3.55	13706	12376

ANEXO 6B: Interações previstas pelo PLIP (Adasme et al., 2021) para o fragmento PPGQ1998R0108

▼ Hydrophobic Interactions

Index	Residue	AA	Distance	Ligand Atom	Protein Atom
1	512A	LEU	3.74	8956	3218
2	512A	LEU	3.43	8954	3215
3	799A	ALA	3.83	8954	7462
4	803A	TRP	3.47	8955	7535

ANEXO 7A: Interações previstas pelo PLIP (Adasme et al., 2021) para o ligante PPGQ200070110

▼ Hydrophobic Interactions ----

Index	Residue	AA	Distance	Ligand Atom	Protein Atom
1	512A	LEU	3.70	9006	3215
2	737A	ARG	3.75	8972	6476
3	742A	GLN	3.46	9002	6563
4	795A	TRP	3.93	8982	7394
5	795A	TRP	3.76	9000	7398
6	803A	TRP	3.98	8954	7535

▼ Hydrogen Bonds ——

Index	Residue	AA	Distance H-A	Distance D-A	Donor Angle	Protein donor?	Side chain	Donor Atom	Acceptor Atom
1	512A	LEU	3.18	4.05	143.38	✓	✗	3211 [Nam]	8966 [O2]
2	729A	ARG	2.07	2.68	116.02	✓	✓	6353 [Ng+]	8987 [O2]
3	794A	THR	1.87	2.69	138.17	✗	✓	8995 [O2]	7378 [O3]
4	795A	TRP	2.97	3.40	106.93	✓	✓	7395 [Nar]	8989 [O2]

ANEXO 7B: Interações previstas pelo PLIP (Adasme et al., 2021) para o ligante PPGQ200070111

▼ Hydrophobic Interactions ----

Index	Residue	AA	Distance	Ligand Atom	Protein Atom
1	512A	LEU	3.83	8971	3218
2	737A	ARG	3.73	8988	6476

▼ Hydrogen Bonds ——

Index	Residue	AA	Distance H-A	Distance D-A	Donor Angle	Protein donor?	Side chain	Donor Atom	Acceptor Atom
1	794A	THR	1.81	2.68	143.09	✓	✓	7378 [O3]	8999 [O2]
2	795A	TRP	2.21	3.12	148.45	✓	✗	7387 [Nam]	8999 [O2]
3	795A	TRP	3.00	3.77	134.74	✓	✓	7395 [Nar]	8966 [O2]
4	796A	SER	2.04	2.96	154.49	✓	✓	7416 [O3]	8977 [O2]

▼ π-Stacking ⚡ ⚡ ⚡

Index	Residue	AA	Distance	Angle	Offset	Stacking Type	Ligand Atoms
1	795A	TRP	3.74	18.26	0.68	P	8953, 8954, 8955, 8956, 8957, 8958

ANEXO 7C: Interações previstas pelo PLIP (Adasme et al., 2021) para o ligante PPGQ200070130

▼ Hydrophobic Interactions

Index	Residue	AA	Distance	Ligand Atom	Protein Atom
1	511A	LEU	3.71	13694	7703
2	511A	LEU	3.37	13695	7700
3	737A	ARG	3.82	13710	11327
4	799A	ALA	3.83	13695	12303
5	803A	TRP	3.29	13693	12373

▼ Hydrogen Bonds —

Index	Residue	AA	Distance H-A	Distance D-A	Donor Angle	Protein donor?	Side chain	Donor Atom	Acceptor Atom
1	511A	LEU	2.73	3.57	139.24	✓	✗	7696 [Nam]	13700 [O2]
2	729A	ARG	1.87	2.85	160.05	✓	✓	11204 [Ng<+]	13722 [O2]
3	794A	THR	2.13	3.09	164.52	✓	✓	12219 [O3]	13726 [O2]
4	798A	HIS	1.66	2.67	172.44	✗	✗	13700 [O2]	12285 [O2]
5	800A	HIS	2.89	3.76	143.46	✓	✗	12309 [Nam]	13700 [O2]

▼ π-Cation Interactions

Index	Residue	AA	Distance	Offset	Protein charged?	Ligand Group	Ligand Atoms
1	737A	ARG	4.15	1.18	✓	Aromatic	13713, 13714, 13715, 13716, 13717, 13718

ANEXO 8A: Interações previstas pelo PLIP (Adasme et al., 2021) para o ligante PPGQ201920105

▼ Hydrophobic Interactions

Index	Residue	AA	Distance	Ligand Atom	Protein Atom
1	511A	LEU	3.98	13707	7703
2	511A	LEU	3.56	13705	7700
3	799A	ALA	3.60	13705	12303
4	803A	TRP	3.28	13707	12373

▼ Hydrogen Bonds —

Index	Residue	AA	Distance H-A	Distance D-A	Donor Angle	Protein donor?	Side chain	Donor Atom	Acceptor Atom
1	766A	TYR	2.06	3.04	165.78	✓	✓	11769 [O3]	13701 [N2]
2	794A	THR	1.89	2.88	167.00	✓	✓	12219 [O3]	13728 [N2]

▼ π-Stacking ⚡ ⚡ ⚡

Index	Residue	AA	Distance	Angle	Offset	Stacking Type	Ligand Atoms
1	795A	TRP	4.59	69.00	0.90	T	13691, 13692, 13693, 13694, 13695, 13696

ANEXO 8B: Interações previstas pelo PLIP (Adasme et al., 2021) para o fragmento PPGQ201920108

▼ Hydrophobic Interactions ----

Index	Residue	AA	Distance	Ligand Atom	Protein Atom
1	512A	LEU	3.65	8954	3218
2	512A	LEU	3.43	8953	3215
3	737A	ARG	3.92	8961	6476
4	795A	TRP	3.55	8970	7398
5	799A	ALA	3.61	8952	7462
6	803A	TRP	3.45	8954	7535

▼ Hydrogen Bonds —

Index	Residue	AA	Distance H-A	Distance D-A	Donor Angle	Protein donor?	Side chain	Donor Atom	Acceptor Atom
1	766A	TYR	2.21	3.15	156.56	✓	✓	6934 [O3]	8995 [N2]
2	794A	THR	3.26	4.02	133.76	✓	✓	7378 [O3]	8964 [N2]
3	795A	TRP	2.89	3.72	138.35	✓	✗	7387 [Nam]	8964 [N2]

ANEXO 8C: Interações previstas pelo PLIP (Adasme et al., 2021) para o ligante PPGQ201920109

▼ Hydrophobic Interactions ----

Index	Residue	AA	Distance	Ligand Atom	Protein Atom
1	512A	LEU	3.61	8966	3218
2	512A	LEU	3.32	8965	3215
3	737A	ARG	3.91	8956	6476
4	742A	GLN	3.90	8997	6563
5	794A	THR	3.79	8963	7379
6	795A	TRP	3.54	8997	7398
7	798A	HIS	3.97	8964	7445
8	799A	ALA	3.66	8964	7462
9	803A	TRP	3.55	8966	7535

▼ Hydrogen Bonds —

Index	Residue	AA	Distance H-A	Distance D-A	Donor Angle	Protein donor?	Side chain	Donor Atom	Acceptor Atom
1	766A	TYR	2.01	2.99	165.99	✓	✓	6934 [O3]	8960 [N2]
2	795A	TRP	2.07	3.06	159.85	✓	✗	7387 [Nam]	8990 [N2]

▼ π-Stacking ··· ···

Index	Residue	AA	Distance	Angle	Offset	Stacking Type	Ligand Atoms
1	795A	TRP	3.74	16.23	1.11	P	8993, 8994, 8995, 8996, 8997, 8998

ANEXO 9: Interações previstas pelo PLIP (Adasme et al., 2021) para o ligante PPGQ201430113

▼ Hydrophobic Interactions

Index	Residue	AA	Distance	Ligand Atom	Protein Atom
1	511A	LEU	3.87	13737	7703
2	511A	LEU	3.60	13746	7700
3	737A	ARG	3.68	13722	11327
4	742A	GLN	3.93	13693	11415
5	794A	THR	3.61	13734	12220
6	800A	HIS	3.54	13752	12313
7	803A	TRP	3.63	13743	12373

▼ Hydrogen Bonds —

Index	Residue	AA	Distance H-A	Distance D-A	Donor Angle	Protein donor?	Side chain	Donor Atom	Acceptor Atom
1	729A	ARG	2.26	3.07	136.24	✓	✓	11204 [Ng+]	13717 [Nar]
2	729A	ARG	2.78	3.51	128.80	✓	✓	11205 [Ng+]	13717 [Nar]
3	792A	ARG	3.23	3.98	131.80	✓	✗	12176 [Nam]	13700 [O2]

ANEXO 10: Interações previstas pelo PLIP (Adasme et al., 2021) para o ligante MPPGQ2003M0103

▼ Hydrophobic Interactions

Index	Residue	AA	Distance	Ligand Atom	Protein Atom
1	511A	LEU	3.67	13697	7703
2	511A	LEU	3.27	13716	7700
3	737A	ARG	3.89	13726	11327
4	803A	TRP	3.67	13714	12373

▼ Hydrogen Bonds —

Index	Residue	AA	Distance H-A	Distance D-A	Donor Angle	Protein donor?	Side chain	Donor Atom	Acceptor Atom
1	511A	LEU	2.66	3.15	109.36	✓	✗	7696 [Nam]	13720 [O2]
2	512A	HIS	2.49	3.47	164.71	✓	✗	7715 [Nam]	13720 [O2]
3	512A	HIS	1.92	2.90	166.13	✓	✓	7721 [Nar]	13720 [O2]
4	729A	ARG	2.26	3.27	175.45	✓	✓	11204 [Ng+]	13712 [O3]
5	766A	TYR	1.75	2.69	158.97	✗	✓	13712 [O3]	11769 [O3]
6	794A	THR	3.01	3.99	175.91	✓	✓	12219 [O3]	13704 [O2]
7	795A	TRP	2.38	2.94	114.39	✓	✓	12236 [Nar]	13730 [O2]
8	795A	TRP	2.99	3.95	157.06	✓	✗	12228 [Nam]	13732 [O2]
9	796A	SER	1.75	2.70	160.05	✓	✓	12257 [O3]	13699 [O2]

▼ π-Cation Interactions

Index	Residue	AA	Distance	Offset	Protein charged?	Ligand Group	Ligand Atoms
1	711A	HIS	4.94	1.91	✓	Aromatic	13695, 13696, 13704, 13705, 13707

ANEXO 11A: Interações previstas pelo PLIP (Adasme et al., 2021) para o ligante 2a

▼ Hydrophobic Interactions

Index	Residue	AA	Distance	Ligand Atom	Protein Atom
1	511A	LEU	3.81	9446	3454
2	511A	LEU	3.28	9455	3451
3	734A	LEU	3.49	9464	7035
4	737A	ARG	3.67	9462	7078
5	803A	TRP	3.48	9453	8139

▼ Hydrogen Bonds —

Index	Residue	AA	Distance H-A	Distance D-A	Donor Angle	Protein donor?	Side chain	Donor Atom	Acceptor Atom
1	511A	LEU	2.00	2.87	140.67	✓	✗	3447 [Nam]	9459 [O.co2]
2	512A	HIS	2.12	3.10	161.48	✓	✗	3466 [Nam]	9459 [O.co2]
3	794A	THR	1.67	2.68	176.35	✓	✓	7985 [O3]	9473 [O.co2]
4	800A	HIS	1.97	2.97	161.93	✓	✗	8075 [Nam]	9458 [O.co2]
5	802A	GLN	1.92	2.87	154.28	✓	✓	8117 [Nam]	9458 [O.co2]

▼ Salt Bridges

Index	Residue	AA	Distance	Protein positive?	Ligand Group	Ligand Atoms
1	512A	HIS	3.97	✓	Carboxylate	9458, 9459
2	729A	ARG	4.75	✓	Carboxylate	9473, 9472
3	800A	HIS	5.07	✓	Carboxylate	9458, 9459

ANEXO 11B: Interações previstas pelo PLIP (Adasme et al., 2021) para o ligante 2b

▼ Hydrophobic Interactions

Index	Residue	AA	Distance	Ligand Atom	Protein Atom
1	511A	LEU	3.71	9445	3454
2	511A	LEU	3.28	9456	3451
3	737A	ARG	3.73	9463	7078
4	799A	ALA	3.95	9450	8069
5	803A	TRP	3.50	9454	8139

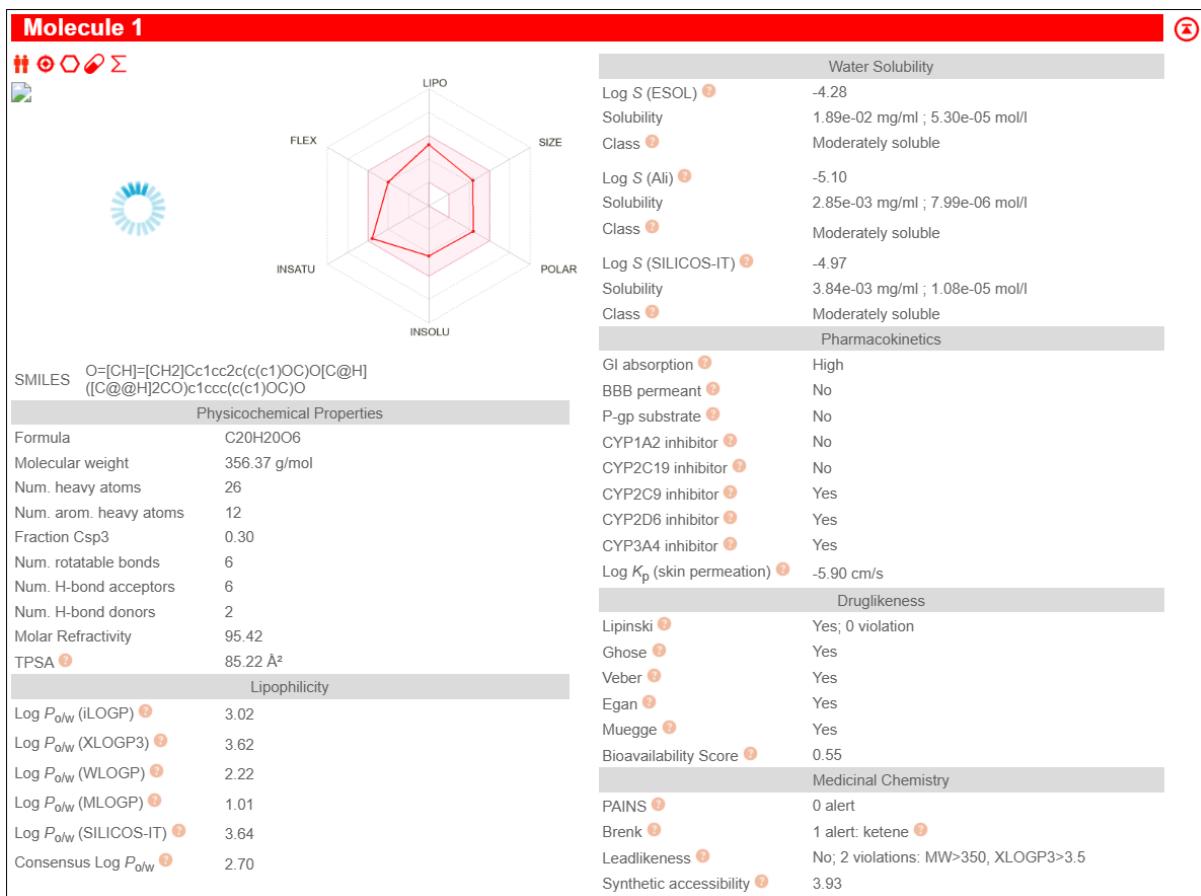
▼ Hydrogen Bonds —

Index	Residue	AA	Distance H-A	Distance D-A	Donor Angle	Protein donor?	Side chain	Donor Atom	Acceptor Atom
1	511A	LEU	1.99	2.88	142.66	✓	✗	3447 [Nam]	9460 [O.co2]
2	512A	HIS	2.14	3.12	161.61	✓	✗	3466 [Nam]	9460 [O.co2]
3	794A	THR	1.67	2.68	171.92	✓	✓	7985 [O3]	9473 [O.co2]
4	800A	HIS	2.03	3.03	163.42	✓	✗	8075 [Nam]	9459 [O.co2]
5	802A	GLN	1.94	2.90	153.33	✓	✓	8117 [Nam]	9459 [O.co2]

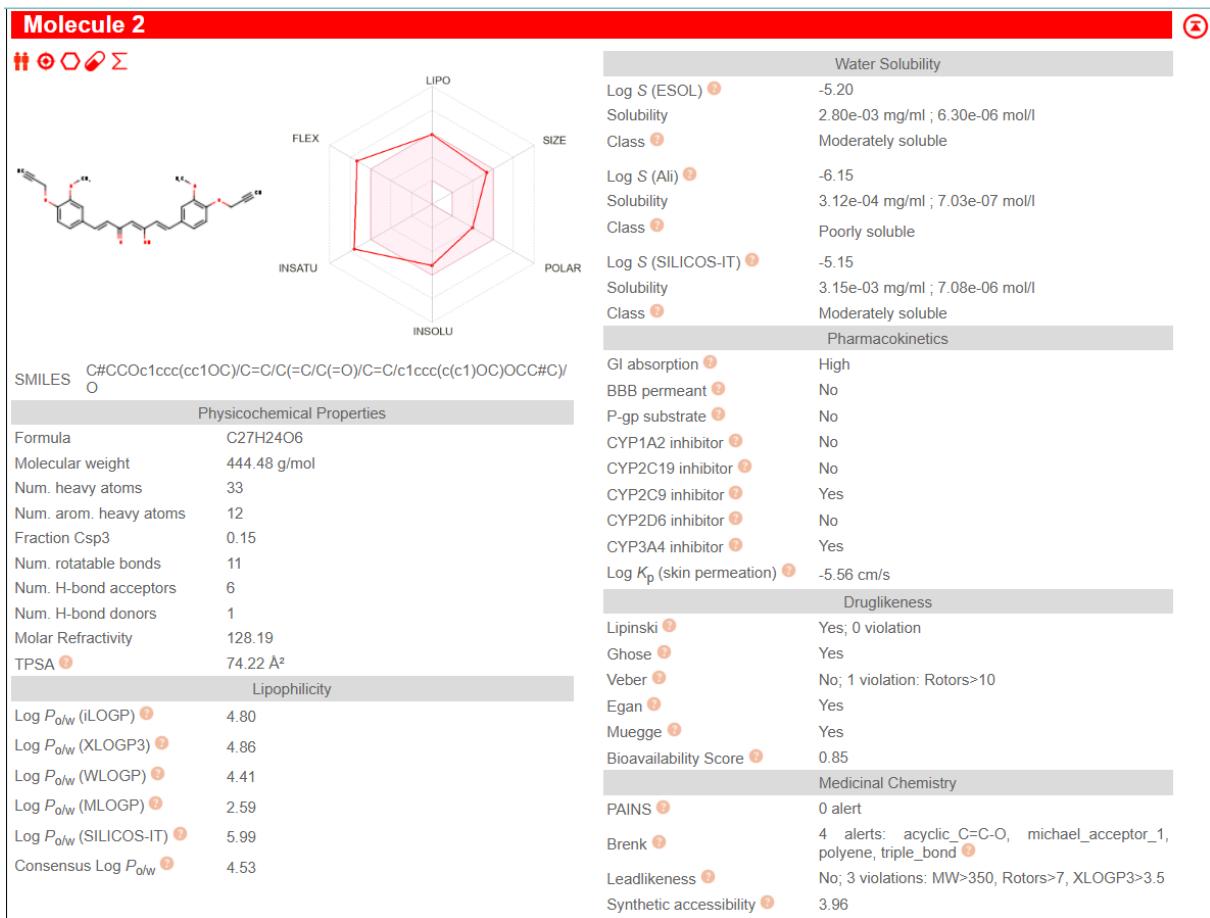
▼ Salt Bridges

Index	Residue	AA	Distance	Protein positive?	Ligand Group	Ligand Atoms
1	512A	HIS	4.00	✓	Carboxylate	9459, 9460
2	729A	ARG	4.89	✓	Carboxylate	9473, 9474
3	800A	HIS	5.32	✓	Carboxylate	9459, 9460

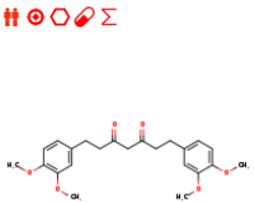
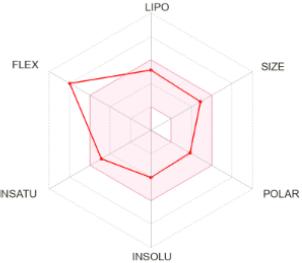
ANEXO 12A: Previsão ADMET pelo servidor Swiss-ADME (Daina et al., 2017) para o ligante MPPGQ2003M0103



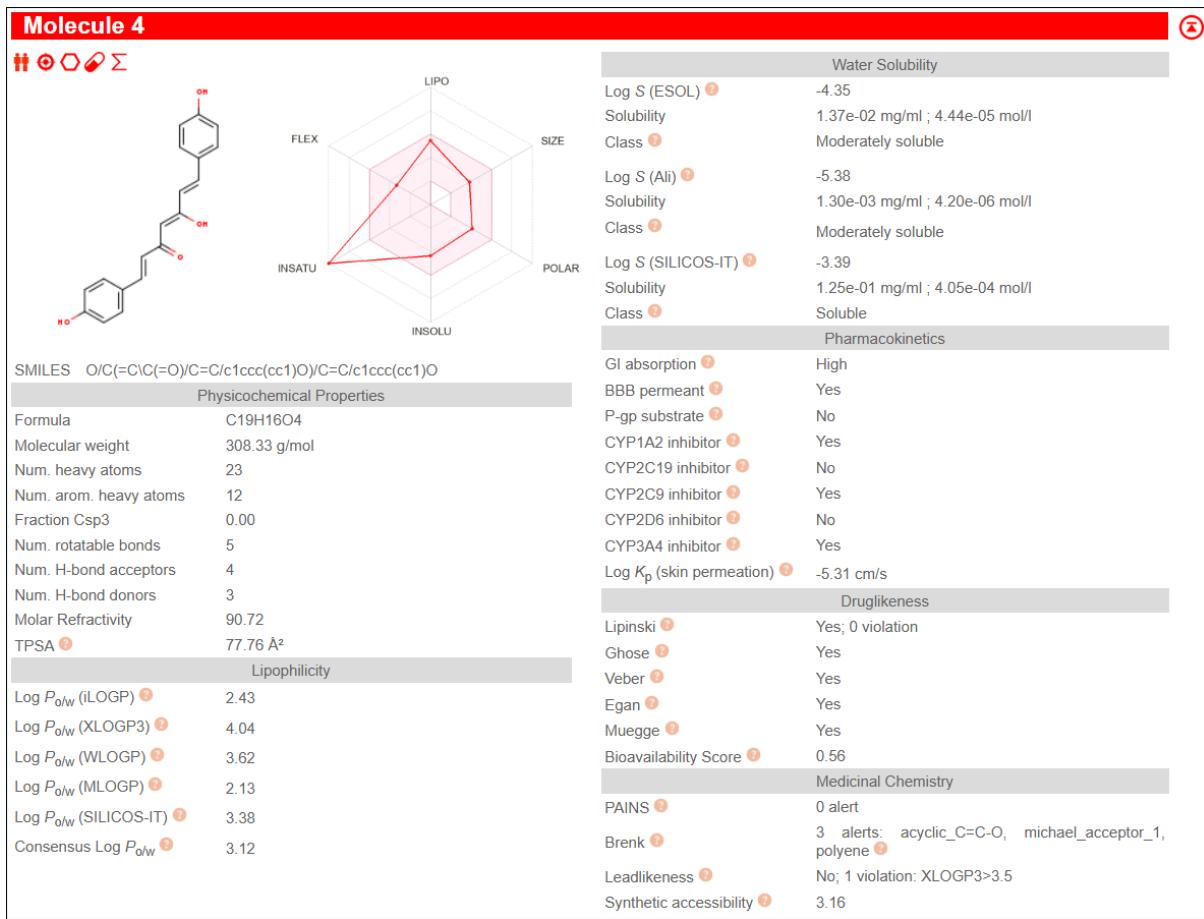
ANEXO 12B: Previsão ADMET pelo servidor Swiss-ADME (Daina et al., 2017) para o ligante PPGQ200070110



ANEXO 12C: Previsão ADMET pelo servidor Swiss-ADME (Daina et al., 2017) para o ligante PPGQ200070111

Molecule 3	
	
SMILES	CCc1cc(CCC(=O)CC(=O)CCc2ccc(c(c2)OC)OC)ccc1OC
Physicochemical Properties	
Formula	C23H28O6
Molecular weight	400.46 g/mol
Num. heavy atoms	29
Num. arom. heavy atoms	12
Fraction Csp3	0.39
Num. rotatable bonds	12
Num. H-bond acceptors	6
Num. H-bond donors	0
Molar Refractivity	111.10
TPSA	71.06 Å ²
Lipophilicity	
Log P _{ow} (iLOGP)	3.79
Log P _{ow} (XLOGP3)	3.47
Log P _{ow} (WLOGP)	3.81
Log P _{ow} (MLOGP)	2.06
Log P _{ow} (SILICOS-IT)	5.51
Consensus Log P _{ow}	3.73
Water Solubility	
Log S (ESOL)	-4.02
Solubility	3.80e-02 mg/ml ; 9.48e-05 mol/l
Class	Moderately soluble
Log S (Ali)	-4.64
Solubility	9.08e-03 mg/ml ; 2.27e-05 mol/l
Class	Moderately soluble
Log S (SILICOS-IT)	-7.27
Solubility	2.16e-05 mg/ml ; 5.40e-08 mol/l
Class	Poorly soluble
Pharmacokinetics	
GI absorption	High
BBB permeant	Yes
P-gp substrate	No
CYP1A2 inhibitor	No
CYP2C19 inhibitor	No
CYP2C9 inhibitor	No
CYP2D6 inhibitor	Yes
CYP3A4 inhibitor	Yes
Log K _p (skin permeation)	-6.28 cm/s
Druglikeness	
Lipinski	Yes; 0 violation
Ghose	Yes
Veber	No; 1 violation: Rotors>10
Egan	Yes
Muegge	Yes
Bioavailability Score	0.55
Medicinal Chemistry	
PAINS	0 alert
Brenk	1 alert: beta_keto_anhydride
Leadlikeness	No; 2 violations: MW>350, Rotors>7
Synthetic accessibility	2.71

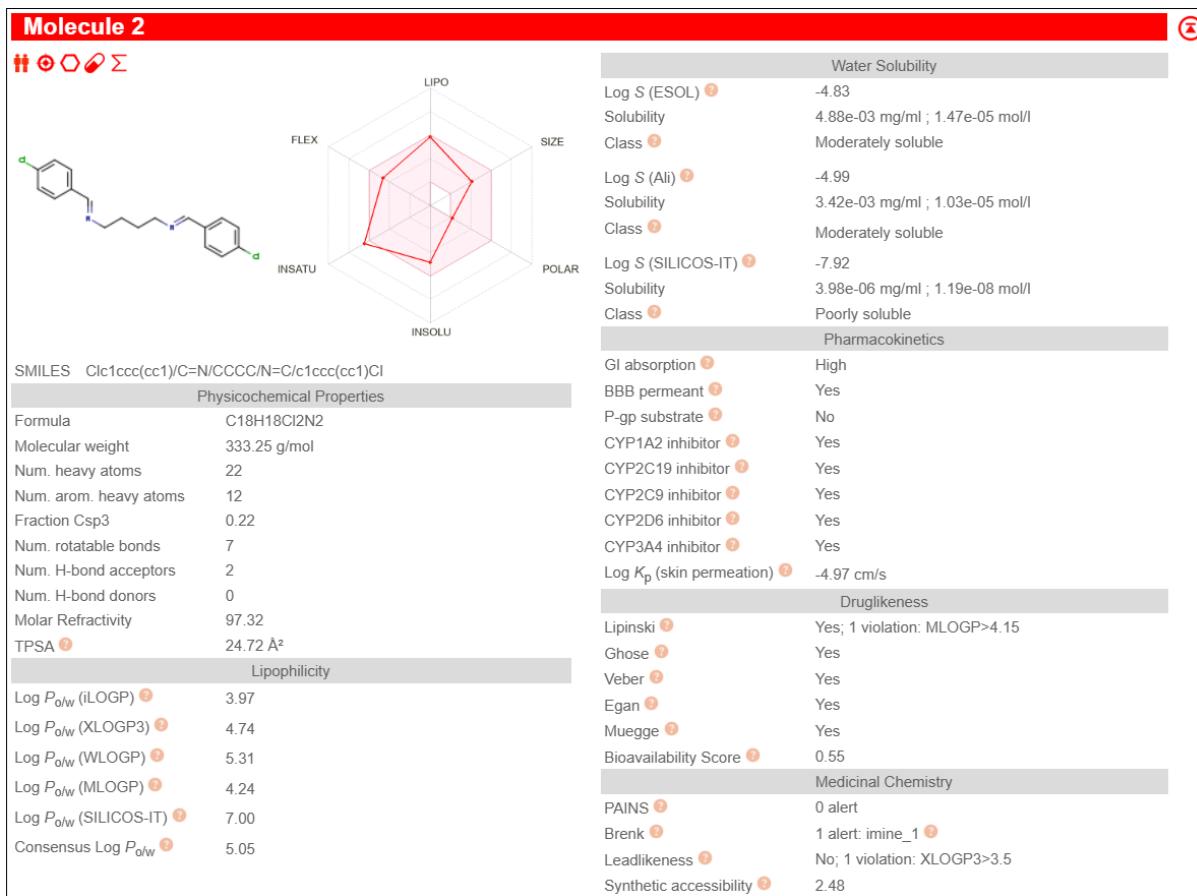
ANEXO 12D: Previsão ADMET pelo servidor Swiss-ADME (Daina et al., 2017) para o ligante PPGQ200070130



ANEXO 12E: Previsão ADMET pelo servidor Swiss-ADME (Daina et al., 2017) para o ligante PPGQ201430113



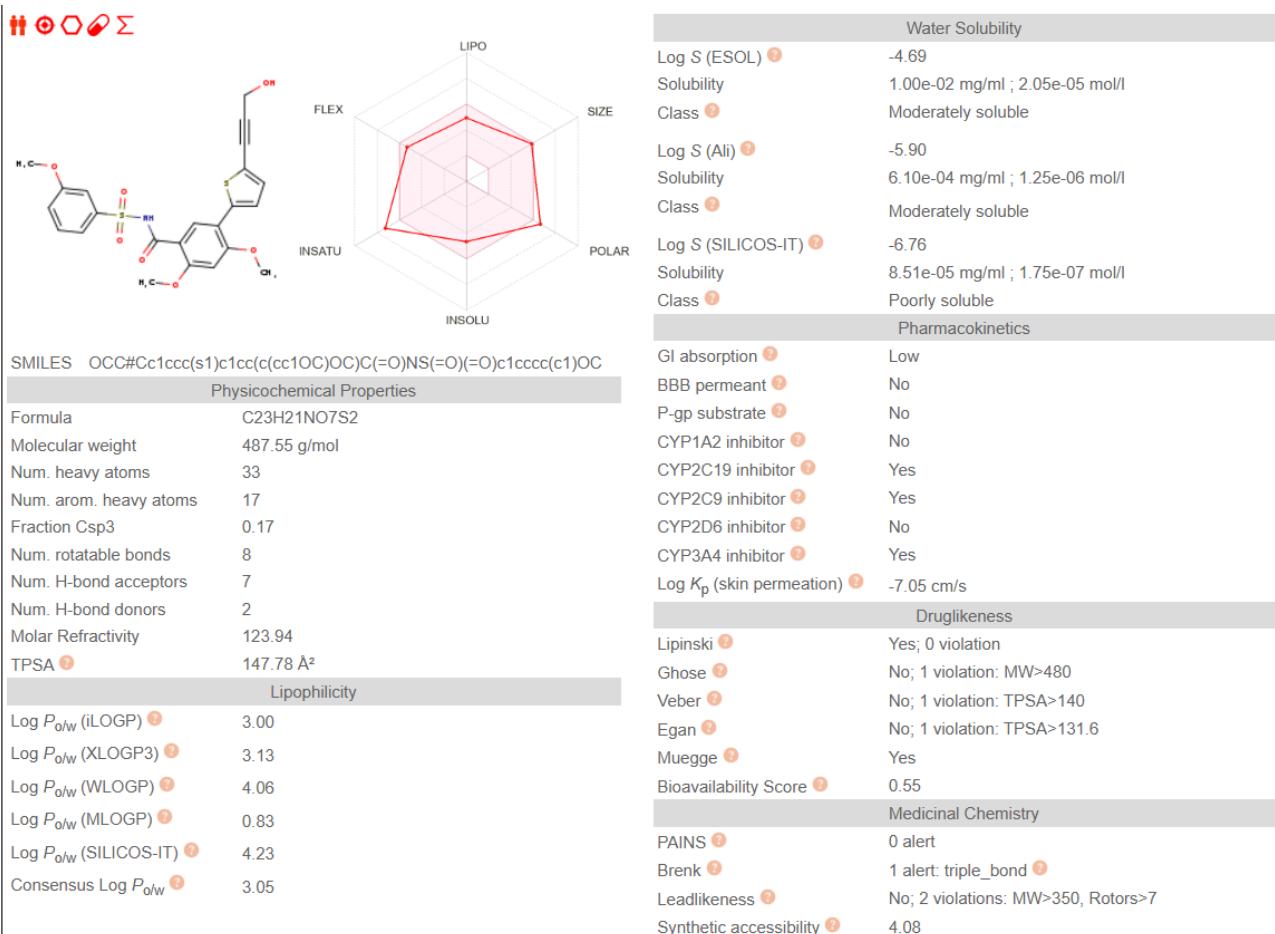
ANEXO 12F: Previsão ADMET pelo servidor Swiss-ADME (Daina et al., 2017) para o ligante PPGQ201920105



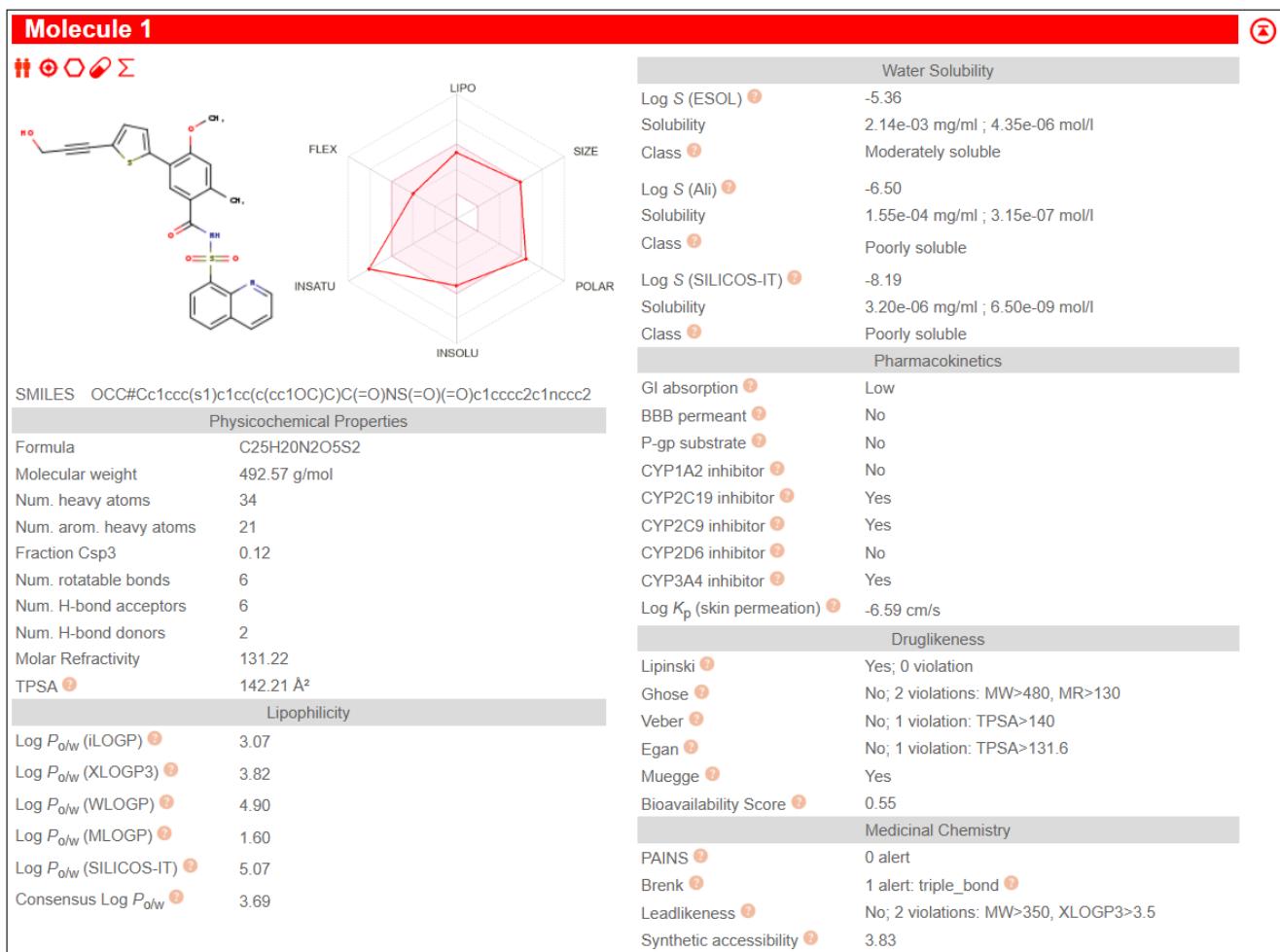
ANEXO 12G: Previsão ADMET pelo servidor Swiss-ADME (Daina et al., 2017) para o ligante PPGQ201920109

Molecule 3	
SMARTS	Water Solubility
<chem>CC(C)N=C/c1ccccc1)CCCC/N=C/c1ccccc1</chem>	Log S (ESOL) -4.80 Solubility 5.11e-03 mg/ml ; 1.59e-05 mol/l Class 3 Moderately soluble
Chemical Structure	Log S (Ali) -5.55 Solubility 9.05e-04 mg/ml ; 2.82e-06 mol/l Class 3 Moderately soluble
	Log S (SILICOS-IT) -8.32 Solubility 1.54e-06 mg/ml ; 4.82e-09 mol/l Class 3 Poorly soluble
3D QSAR	Pharmacokinetics
	GI absorption High BBB permeant Yes P-gp substrate No CYP1A2 inhibitor Yes CYP2C19 inhibitor Yes CYP2C9 inhibitor No CYP2D6 inhibitor Yes CYP3A4 inhibitor Yes Log Kp (skin permeation) -4.51 cm/s
Physicochemical Properties	Druglikeness
SMILES <chem>CC(C)N=C/c1ccccc1)CCCC/N=C/c1ccccc1</chem>	Lipinski Yes; 0 violation Ghose Yes Veber No; 1 violation: Rotors>10 Egan Yes Muegge No; 1 violation: XLOGP3>5 Bioavailability Score 0.55
Formula C22H28N2	Medicinal Chemistry
Molecular weight 320.47 g/mol	PAINS 0 alert Brenk 1 alert: imine_1 Leadlikeness No; 2 violations: Rotors>7, XLOGP3>3.5 Synthetic accessibility 2.76
Num. heavy atoms 24	
Num. arom. heavy atoms 12	
Fraction Csp3 0.36	
Num. rotatable bonds 11	
Num. H-bond acceptors 2	
Num. H-bond donors 0	
Molar Refractivity 106.53	
TPSA 24.72 Å²	
Lipophilicity	
Log P _{o/w} (iLOGP) 4.39	
Log P _{o/w} (XLOGP3) 5.28	
Log P _{o/w} (WLOGP) 5.57	
Log P _{o/w} (MLLOGP) 4.14	
Log P _{o/w} (SILICOS-IT) 7.30	
Consensus Log P _{o/w} 5.34	

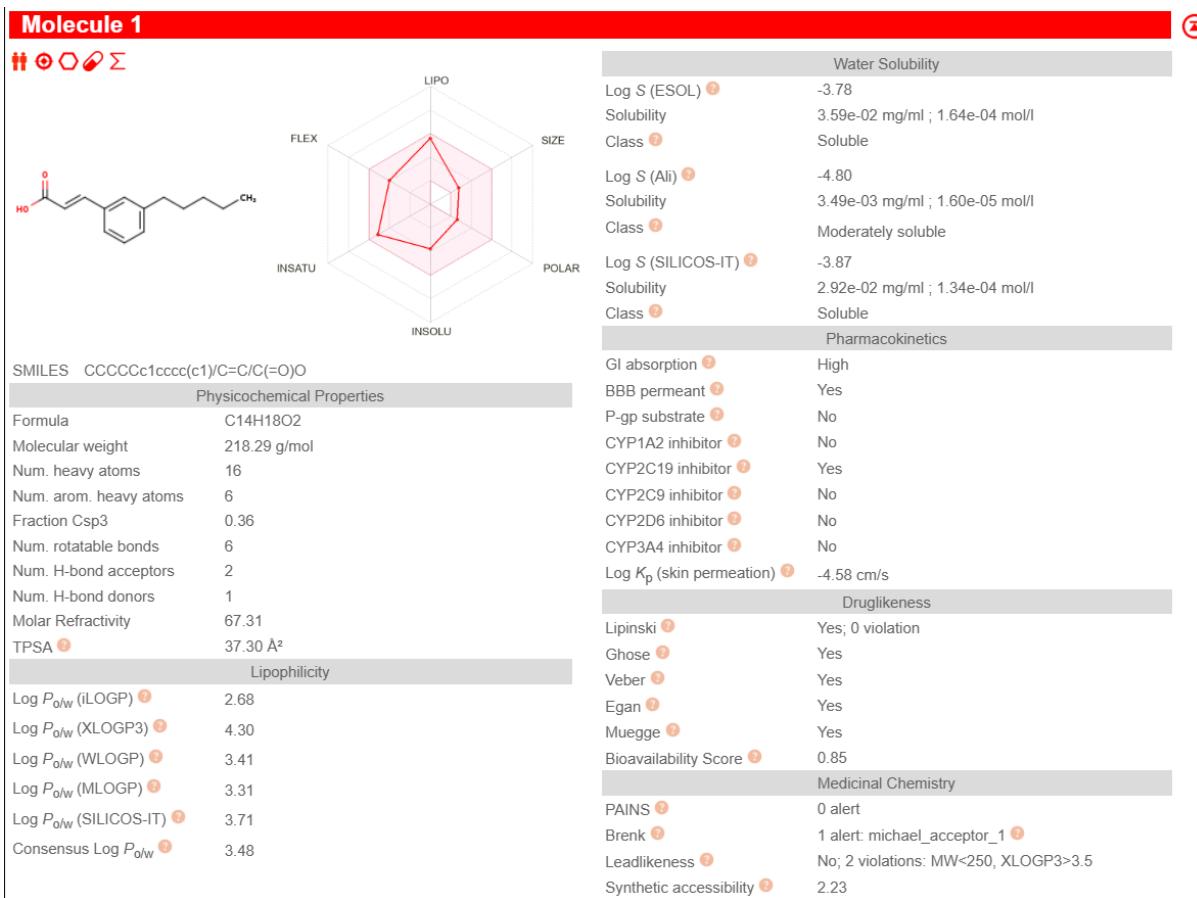
ANEXO 12H: Previsão ADMET pelo servidor Swiss-ADME (Daina et al., 2017) para o inibidor 27



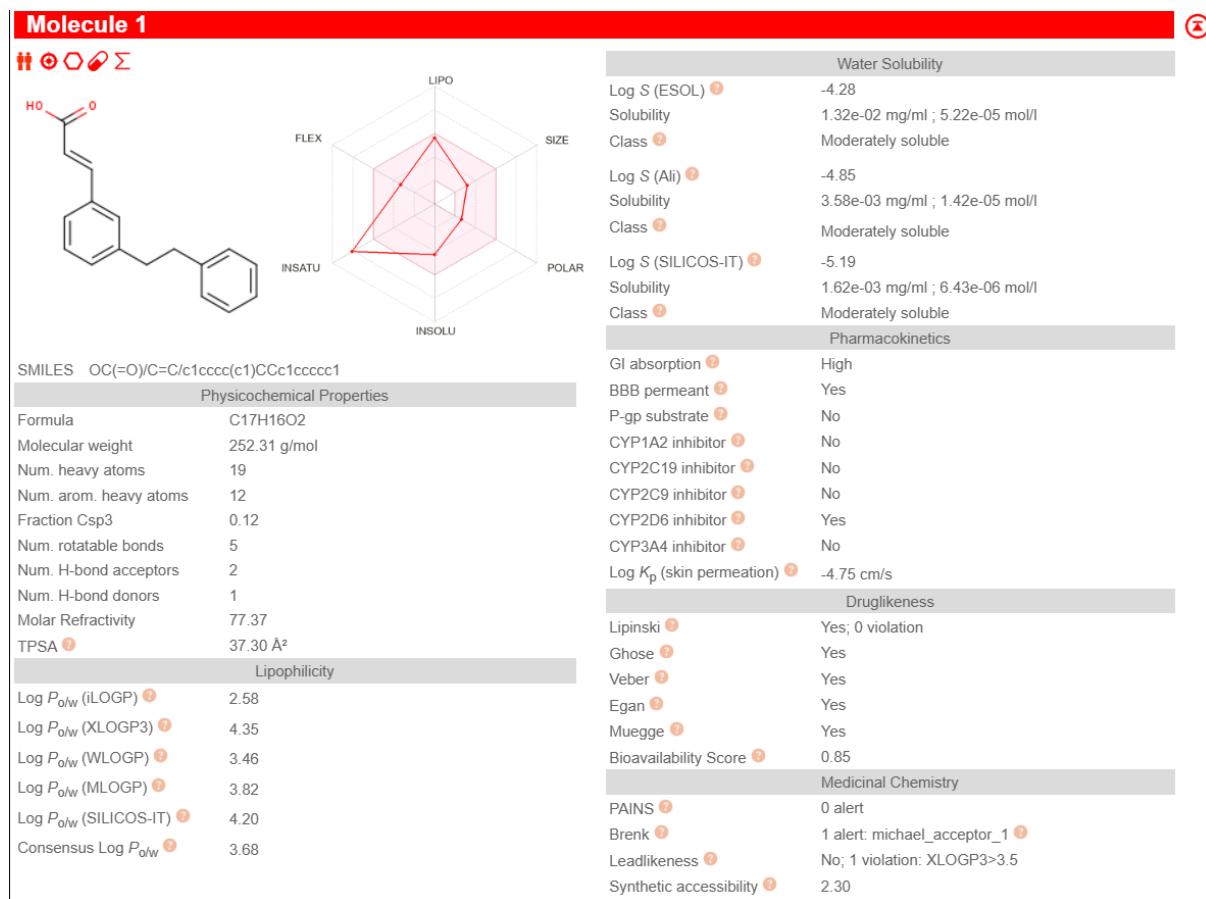
ANEXO 12I: Previsão ADMET pelo servidor Swiss-ADME (Daina et al., 2017) para o inibidor 29



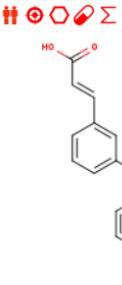
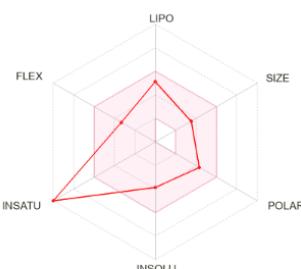
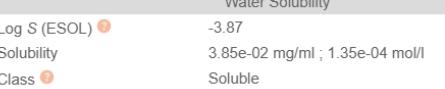
ANEXO 13A: Previsão ADMET pelo servidor Swiss-ADME (Daina et al., 2017) para o fragmento 1h



ANEXO 13B: Previsão ADMET pelo servidor Swiss-ADME (Daina et al., 2017) para o fragmento 11



ANEXO 13C: Previsão ADMET pelo servidor Swiss-ADME (Daina et al., 2017) para o fragmento 2A

Molecule 1	
 	     
SMILES	OC(=O)/C=C/c1cccc(c1)Oc1cccc(c1)C(=O)O
Physicochemical Properties	
Formula	C16H12O5
Molecular weight	284.26 g/mol
Num. heavy atoms	21
Num. arom. heavy atoms	12
Fraction Csp3	0.00
Num. rotatable bonds	5
Num. H-bond acceptors	5
Num. H-bond donors	2
Molar Refractivity	76.59
TPSA	83.83 Å²
Lipophilicity	
Log P _{ow} (iLOGP)	1.98
Log P _{ow} (XLOGP3)	3.45
Log P _{ow} (WLOGP)	3.17
Log P _{ow} (MLOGP)	2.55
Log P _{ow} (SILICOS-IT)	2.42
Consensus Log P _{ow}	2.71

ANEXO 13D: Previsão ADMET pelo servidor Swiss-ADME (Daina et al., 2017) para o fragmento 2b

