

UNIVERSIDADE FEDERAL DE MINAS GERAIS

Instituto de Ciências Biológicas

Programa de Pós-graduação em Bioquímica e Imunologia

Beatriz Murta Rezende Moraes Ribeiro

**TRIAGEM VIRTUAL PARA PROSPECÇÃO DE INIBIDORES CONTRA A
PROTEASE NS2B-NS3 DE ZIKA VÍRUS**

Belo Horizonte

2023

Beatriz Murta Rezende Moraes Ribeiro

**TRIAGEM VIRTUAL PARA PROSPECÇÃO DE INIBIDORES CONTRA A
PROTEASE NS2B-NS3 DE ZIKA VÍRUS**

Dissertação apresentada ao programa de pós-graduação em Bioquímica e Imunologia do Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais, como requisito parcial para obtenção do título de Mestra em Bioquímica e Imunologia.

Orientadora: Profa. Dra. Rafaela Salgado Ferreira

Belo Horizonte

2023

043 Ribeiro, Beatriz Murta Rezende Moraes.
Triagem virtual para prospecção de inibidores contra a protease NS2B-NS3 de zika vírus [manuscrito] / Beatriz Murta Rezende Moraes Ribeiro. – 2023.
91 f. : il. ; 29,5 cm.

Orientador: Rafaela Salgado Ferreira.

Dissertação (mestrado) – Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas. Programa de Pós-Graduação em Bioquímica e Imunologia.

1. Bioquímica e Imunologia. 2. Zika vírus. 3. Preparações Farmacêuticas. 4. Triagem. I. Ferreira, Rafaela Salgado. II. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. III. Título.

CDU: 577.1



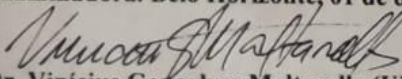
Universidade Federal de Minas Gerais
 Curso de Pós-Graduação em Bioquímica e Imunologia ICB/UFMG
 Av. Antônio Carlos, 6627 - Pampulha
 31270-901 - Belo Horizonte - MG
 e-mail: pg-biq@icb.ufmg.br (31)3409-2615

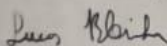


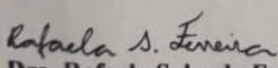
ATA DA DEFESA DA DISSERTAÇÃO DE MESTRADO DE BEATRIZ MURTA REZENDE MORAES RIBEIRO. Ao um dia do mês de dezembro de 2023 às 09:00 horas, reuniu-se no Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais, a Comissão Examinadora da dissertação de Mestrado, indicada *ad referendum* do Colegiado do Curso, para julgar, em exame final, o trabalho intitulado "Triagem virtual para prospecção de inibidores contra a protease NS2B-NS3 de Zika vírus", requisito final para a obtenção do grau de Mestre em Bioquímica e Imunologia, área de concentração: Bioquímica. Abrindo a sessão, a Presidente da Comissão, Profª. Rafaela Salgado Ferreira, da Universidade Federal de Minas Gerais, após dar a conhecer aos presentes o teor das Normas Regulamentares do Trabalho Final, passou a palavra à candidata para apresentação de seu trabalho. Seguiu-se a arguição pelos examinadores, com a respectiva defesa da candidata. Logo após a Comissão se reuniu, sem a presença da candidata e do público, para julgamento e expedição do resultado final. Foram atribuídas as seguintes indicações: Dr. Vinícius Gonçalves Maltarollo (Universidade Federal de Minas Gerais), aprovada; Dr. Lucas Bleicher (Universidade Federal de Minas Gerais), aprovada; Dra. Rafaela Salgado Ferreira - Orientadora (Universidade Federal de Minas Gerais), aprovada. Pelas indicações a candidata foi considerada:

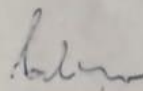
APROVADA
 REPROVADA

O resultado final foi comunicado publicamente à candidata pela Presidente da Comissão. Nada mais havendo a tratar, a Presidente da Comissão encerrou a reunião e lavrou a presente Ata que será assinada por todos os membros participantes da Comissão Examinadora. Belo Horizonte, 01 de dezembro de 2023.


 Dr. Vinícius Gonçalves Maltarollo (UFMG)


 Dr. Lucas Bleicher (UFMG)


 Dra. Rafaela Salgado Ferreira - Orientadora (UFMG)


 Dra. Fabíola Mara Ribeiro
 Coordenadora do Programa de
 Pós-graduação em Bioquímica e Imunologia

AGRADECIMENTOS

A minha família que sempre me apoiou em todas minhas decisões de vida e sempre estão comigo quando preciso.

A minha orientadora Rafaela Salgado Ferreira, pela oportunidade de fazer parte do seu laboratório desde a iniciação científica e agora no mestrado sempre me auxiliando em cada etapa.

A todos membros do laboratório LMMPF, pelo apoio, ajuda e incentivo

As instituições CNPQ e CAPES pelo suporte financeiro para o desenvolvimento do trabalho.

À OpenEye Scientific Software, pyMOL, Chimera, Discovery Studio, PDB-REDO pela disponibilização da licença acadêmica ou acesso aberto para utilização de suas ferramentas

RESUMO

A doença de Zika, causada pelo vírus de mesmo nome, geralmente se manifesta com sintomas leves, mas pode resultar em complicações graves, como a síndrome congênita de Zika em recém-nascidos e Guillain-Barré em adultos. Não existem tratamentos aprovados, sendo necessários estudos nesse sentido. A protease NS2B-NS3 realiza um papel essencial na replicação viral, tornando-se um alvo promissor para estudos de desenvolvimento de fármacos. Este trabalho integrou abordagens computacionais e experimentais para identificar inibidores desta protease. O protocolo de triagem virtual foi desenvolvido considerando as estruturas resolvidas experimentalmente da enzima, levando em consideração sua estrutura e conformação do sítio ativo. O programa utilizado foi o Dock6 e a validação do protocolo foi feita em relação a reprodução de poses cristalográficas (*cross-docking*) e a performance na triagem virtual (avaliação da AUC da curva ROC e fator de enriquecimento). O *cross-docking* não obteve sucesso na reprodução de poses em nenhuma estrutura tridimensional utilizada. Devido ao predomínio de ligantes macrocíclicos, foram realizados estudos sobre o impacto do uso de conformações de entrada mais similares as poses cristalográficas no sucesso do *cross-docking*, mas não observou-se melhora significativa no resultado. Foram construídas curvas ROC utilizando dois conjuntos testes, decoys gerados pelo DUD-E e compostos inativos reportados na literatura. A curva referente ao primeiro grupo apresentou baixa performance (AUC = 0,487), enquanto a do segundo apresentou um excelente desempenho (AUC = 0,962). No entanto, esse resultado pode estar enviesado devido as características dos ligantes inativos, que possuem predomínio de cargas 0 e 1, enquanto entre em relação aos ligantes ativos predominam ligantes com cargas 3. Além do sítio ativo ter característica negativa, a função de pontuação empregada não considera a energia de dessolvatação, favorecendo compostos com carga formal positiva. A partir da triagem virtual da biblioteca BraCoLi foram selecionados dezesseis compostos, dos quais 10 foram disponibilizados e submetidos a experimentos *in vitro*. Todos apresentaram baixa inibição quando avaliados a 100 µM. O ligante BR20151 (31% de inibição) foi submetido a caracterizações adicionais que indicaram comportamento agregador para a molécula.

Palavras-chaves: Zika vírus, planejamento de fármacos, triagem virtual

ABSTRACT

Zika disease is caused by the Zika virus and is known to cause mild symptoms but also result in severe complications such as Zika congenital syndrome in newborns and Guillain-Barré syndrome in adults. There are no approved treatments, thus further research in this direction is needed. The NS2B-NS3 protease plays a crucial role in viral replication, making it a promising target for drug development studies. This work integrated computational and experimental approaches to identify protease inhibitors. The virtual screening protocol was developed considering experimentally resolved enzyme structures, taking into account its overall structure and active site conformation. Dock6 was employed for docking, and the protocol was validated considering the reproduction of crystallographic poses (cross-docking) and its screening performance (based on evaluation of the ROC curve's AUC and enrichment factor). Cross-docking results did not succeed in any structure considered. Due to the prevalence of macrocyclic ligands, we evaluated the impact of using entry conformations more similar to the crystallographic poses on the success of cross-docking, but no significant improvement was observed. ROC curves were constructed using two test sets: decoys generated by DUD-E and inactive compounds reported in the literature taken from Abrams (2020). The curve for the first group showed low performance (AUC = 0.487), while the second exhibited excellent performance (AUC = 0.962). However, this result may be biased due to the characteristics of inactive ligands, which predominantly have charges of 0 and 1, compared to active ligands among which charges of 3 dominate. Besides, the active site has a negative electrostatic potential, and the scoring function employed does not consider the desolvation energy, favoring compounds with a positive formal charge. Virtual screening of the BraCoLi library resulted in a selection of sixteen compounds, of which 10 were made available and subjected to in vitro experiments. All showed low inhibition when evaluated at 100 μ M. The ligand BR20151 (31% inhibition) underwent additional characterizations indicating an aggregating behavior for the molecule.

Key words: Zika virus, drug design, virtual screening

LISTA DE FIGURAS

| | |
|---|-----------|
| Figura 1- Estrutura da poliproteína viral. os aminoácidos ao longo da cadeia. | 15 |
| Figura 2 - Estrutura da protease NS2B-NS3PRO de Zika vírus. | 17 |
| Figura 3 - Estrutura dos inibidores de NS2B-NS3pro ZIKV. | 19 |
| Figura 4 – Sequência proteica resultante do inserto construído por Lei e colaboradores (2016). | 28 |
| Figura 5 – Análise estrutural das 49 estruturas da NS2B-NS3pro ZIKV com ligantes disponíveis no PDB. | 40 |
| Figura 6 – Diferença estrutural entre a estrutura 7M1V e 5YOF. | 40 |
| Figura 7 – Comparação da cobertura da densidade eletrônica de ligantes utilizando a ferramenta <i>density fitness</i> do site RCSB. | 41 |
| Figura 8 - Análise do perfil das interações proteína-ligante entre as 49 estruturas cristalográficas estudadas. | 43 |
| Figura 9 - Sobreposição dos resíduos VAL155 e ALA132. | 46 |
| Figura 10 - Comparação da orientação dos resíduos ASP129 e SER135 nas estruturas do <i>cross-docking</i> | 47 |
| Figura 11 – Comparação do modo de ligação entre o ligante 5YOF cristalográfico da estrutura 6kk2 o ligante 5YOF correspondente à pose 1 do docking com a estrutura 6KK2. | 51 |
| Figura 12 - Comparação das conformações e poses do ligante L7PFY. | 58 |
| Figura 13 – Comparação das conformações e poses do ligante 5YOF. | 59 |
| Figura 14 – Clusterização dos ligantes ativos usados na construção das curvas ROC. | 61 |
| Figura 15 - Resultado das análises de enriquecimento. | 63 |
| Figura 16 – Distribuição de cargas nos diferentes grupos de ligantes. | 64 |
| Figura 17 - Caixa usada para o <i>docking</i> com a estrutura tridimensional 5YOF. Representação da caixa desenvolvida para a triagem virtual. | 66 |
| Figura 18 – Estrutura química dos compostos do grupo 1 e modo de ligação predito no sítio ativo da NS2B/NS3pro ZIKV com a estrutura tridimensional 5YOF | 67 |
| Figura 19 – Estrutura química dos compostos do grupo 2 e modo de ligação predito no sítio ativo da NS2B/NS3pro ZIKV 2 e modo de ligação com a estrutura tridimensional 5YOF. | 68 |

| | |
|---|----|
| Figura 20 – Estrutura química dos compostos do grupo 3 e modo de ligação predito no sítio ativo da NS2B/NS3pro ZIKV 3 e modo de ligação com a estrutura tridimensional 5YOF. | 69 |
| Figura 21 – Estrutura química dos compostos do grupo e modo de ligação predito no sítio ativo da NS2B/NS3pro ZIKV 4 e modo de ligação com a estrutura tridimensional 5YOF..... | 70 |
| Figura 22 – Estrutura química dos compostos do grupo 3 e modo de ligação predito no sítio ativo da NS2B/NS3pro ZIKV 3 e modo de ligação com a estrutura tridimensional 5YOF..... | 71 |
| Figura 23 – Gráfico de interações versus resíduos do sítio ativo da NS2B-NS3pro. Cada gráfico representa um conjunto distinto de ligantes..... | 72 |
| Figura 24 – Resíduos dos PDBs com maior frequência de interação com o ligante.. .. | 73 |
| Figura 25 – Estruturas dos compostos submetidos aos ensaios enzimáticos | 75 |
| Figura 26 – Curvas de IC ₅₀ do composto BR20151. | 77 |
| Figura 27 – Comportamento da inibição da enzima com variação da concentração do substrato.. .. | 79 |

LISTA DE TABELAS

| | |
|---|----|
| Tabela 1 - Dimensões das caixas utilizadas para o <i>docking</i> | 30 |
| Tabela 2 – Informações sobre as estruturas cristalográficas e inspeção visual de densidade eletrônica | 38 |
| Tabela 3 – RMSD dos resíduos do sítio ativo em relação a estrutura 5LC0 | 44 |
| Tabela 4 – Características dos ligantes cristalográficos e seus respectivos RMSD em relação a cada receptor | 49 |
| Tabela 5 – RMSD entre a conformação dos ligantes cristalográficos e a conformação gerada pelo OMEGA antes e após processamento pelo <i>docking</i> | 54 |
| Tabela 6 - RMSD entre a conformação dos ligantes cristalográficos e a melhor conformação gerada pelo OMEGA antes e após processamento pelo <i>docking</i> | 55 |
| Tabela 7 - RMSD entre a conformação dos ligantes cristalográficos e o resultado do <i>docking</i> realizado com a melhor conformação do OMEGA | 56 |
| Tabela 8 – Atividade inibitória experimentalmente determinada dos compostos selecionados na triagem virtual a 100 μM | 75 |
| Tabela 9 – Caracterização do composto BR20151 | 78 |

LISTA DE ABREVIATURAS

AUC Área sob a curva

BSA Albumina de Soro Bovino

CHAPS (3-((3-cholamidopropyl) dimethylammonio)-1-propanesulfonate)

DMSO Dimetilsulfóxido

FE Fator de enriquecimento

HIV Vírus da imunodeficiência humana

HTS High-throughput screening

IC₅₀ Concentração do inibidor que reduz a atividade de um processo biológico em 50%

NS (proteína) não estrutural

NS2B-NS3pro Protease NS2B-NS3

PDB *Protein data Bank*

pH Potencial de hidrogênio

RMSD Root-mean-square deviation - Desvio Quadrático Médio

ROC Receiver operator characteristic - Característica de Operação do receptor

ZIKV Vírus da Zika

SUMÁRIO

| | | |
|--------|--|----|
| 1 | Introdução..... | 14 |
| 1.1 | Histórico e epidemiologia..... | 14 |
| 1.2 | Estrutura viral..... | 15 |
| 1.2.1 | NS2B-NS3 protease..... | 16 |
| 1.3 | Inibidores da ZIKV NS2B-NS3pro..... | 18 |
| 1.4 | Desenvolvimento racional de fármacos..... | 20 |
| 1.4.1 | <i>Docking</i> molecular..... | 21 |
| 1.5 | Caracterização bioquímica de potenciais inibidores..... | 25 |
| 2 | Objetivos..... | 27 |
| 3 | Materiais e métodos..... | 28 |
| 3.1 | Análise das estruturas cristalográficas disponíveis da NS2B-NS3pro ZIKV..... | 28 |
| 3.2 | Seleção das estruturas tridimensionais..... | 29 |
| 3.3 | Preparação das estruturas tridimensionais..... | 30 |
| 3.4 | Preparação dos ligantes..... | 30 |
| 3.5 | Protocolo adotado para o <i>docking</i> molecular..... | 31 |
| 3.6 | <i>Cross-docking</i> | 31 |
| 3.6.1 | Avaliação dos resultados do <i>cross-docking</i> | 31 |
| 3.7 | Comparação das conformações estruturais dos ligantes..... | 32 |
| 3.8 | Avaliação do enriquecimento do protocolo..... | 32 |
| 3.9 | Triagem virtual..... | 34 |
| 3.10 | Ensaio <i>in vitro</i> para caracterização de inibidores..... | 35 |
| 3.10.1 | Teste de solubilidade..... | 35 |
| 3.10.2 | Ensaio para caracterização bioquímica..... | 35 |
| 4 | Resultados e discussão..... | 38 |

| | | |
|-------|---|----|
| 4.1 | Similaridade global das estruturas de NS2B-NS3pro de Zika | 38 |
| 4.2 | Diferenças estruturais entre os resíduos do sítio ativo entre as estruturas tridimensionais | 42 |
| 4.3 | Avaliação da reprodução da pose cristalográfica dos ligantes | 48 |
| 4.3.1 | Resultados do <i>cross-docking</i> utilizando a conformação do ligante de menor energia..... | 48 |
| 4.3.2 | Avaliação da reprodução da conformação dos ligantes cristalográficos | 52 |
| 4.3.3 | Resultados do <i>cross-docking</i> utilizando conformações dos ligantes com menor RMSD | 54 |
| 4.3.4 | Análise de enriquecimento | 60 |
| 4.4 | Triagem virtual..... | 66 |
| 4.5 | Caracterização dos potenciais inibidores de NSB2-NS3 ZIKV | 74 |
| 5 | Conclusão..... | 80 |
| 6 | Publicações e apresentações de trabalho | 81 |
| 7 | Publicações e manuscritos submetidos..... | 81 |
| 7.1 | Trabalhos publicados durante o mestrado | 81 |
| 7.2 | Apresentações em congressos | 81 |
| 8 | REFERÊNCIAS..... | 83 |

1 Introdução

1.1 Histórico e epidemiologia

Em 1946, a floresta de Zika, localizada na Península de Entebbe, Uganda, sediava estudos sentinela com macacos rhesus, originalmente destinados a pesquisas relacionadas a febre amarela. No entanto, em 1947, os pesquisadores isolaram um vírus até então desconhecido, o qual recebeu o nome de “Zika vírus” em homenagem a essa floresta. No ano seguinte, em 1948, esse vírus também foi isolado em *Aedes africanus*, indicando a possibilidade de transmissão por artrópodes (DICK; KITCHEN; HADDOW, 1952).

Os primeiros indícios de infecção humana causada pelo Zika vírus começaram a surgir na década de 1950, assim como sua disseminação por alguns países africanos e asiáticos (OLSON et al., 1981). A primeira epidemia registrada fora da África e Ásia ocorreu nas ilhas Yap, em 2007, nos Estados Federados da Micronésia, onde cerca de 73% da população foi acometida pelo vírus. Antes disso, apenas quatorze casos da doença haviam sido relatados. Os sintomas mais comuns, até o momento, incluíam febre baixa, rash cutâneo, artralgia e conjuntivite (DUFFY et al., 2009).

Em outubro de 2013, na Polinésia Francesa, eclodiu uma nova epidemia de Zika, sendo a maior registrada até aquele momento. Além do elevado número de contaminações com manifestações clínicas, observou-se um aumento no número de casos de síndrome de Guillain-Barré, associando-a à epidemia (MUSSO; NILLES; CAO-LORMEAU, 2014; OEHLER et al., 2014). Essa síndrome é um distúrbio autoimune que pode ser desencadeado após infecções virais, resultando em fraqueza muscular que pode causar paralisia temporária e, em casos mais graves, levar o paciente a óbito (KATZ; GILBURD; SHOVMAN, 2019).

No ano de 2015, um surto da doença se iniciou no estado da Bahia e se espalhou pelo Brasil (CAMPOS; BANDEIRA; SARDI, 2015; KINDHAUSER et al., 2016). Durante esse período, houve aumento nos casos de microcefalia em recém-nascidos, suspeitando-se de que essa malformação estivesse relacionada à doença. Estudos posteriores comprovaram essa suspeita, demonstrando a capacidade do vírus de atravessar a placenta e afetar o desenvolvimento fetal (DE ARAÚJO et al., 2018; WU et al., 2016). Um dos mecanismos descritos é a capacidade de induzir apoptose e autofagia de células progenitoras corticais prejudicando o neurodesenvolvimento do

feto (CUGOLA et al., 2016). Atualmente, a microcefalia está categorizada dentro do que chamamos de síndrome congênita do Zika, a qual engloba diversas anomalias no desenvolvimento fetal, incluindo problemas oculares e malformações cerebrais (MOORE et al., 2017).

Devido a todo esse histórico, especialmente devido às suas associações com síndromes congênitas, a doença de Zika foi designada como emergência global pela Organização Mundial de Saúde (OMS) em 2016 (GULLAND, 2016). Embora a classificação de emergência global tenha sido retirada, a doença ainda persiste em escala global. O boletim epidemiológico da OMS de fevereiro de 2022 indica que 89 países e territórios continuam registrando a transmissão da doença (WORLD HEALTH ORGANISATION, 2022). Até o momento, não foram aprovados medicamentos ou vacinas para o combate da doença, ressaltando a necessidade de aprofundar as pesquisas nessa área.

1.2 Estrutura viral

O Zika vírus pertence à família *Flaviridae* e ao gênero *Flavivirus*, assim como os vírus da dengue, encefalite japonesa, febre amarela e vírus do Nilo ocidental. Ele é classificado como um vírus de RNA de fita simples de sentido positivo, que utiliza o hospedeiro para sintetizar uma poliproteína com aproximadamente 3500 aminoácidos. Essa poliproteína é posteriormente processada em três proteínas estruturais denominadas capsídeo (C), precursor da membrana (prM) e envelope (Env) e sete proteínas não estruturais (NS1, NS2A, NS2B, NS3, NS4A, NS4B, NS5) (COX; STANTON; SCHINAZI, 2015).

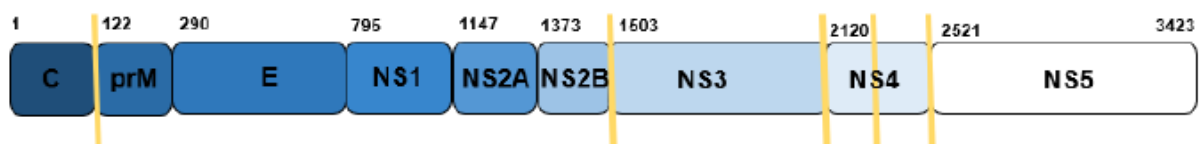


Figura 1- Estrutura da poliproteína viral. As três proteínas estruturais são capsídeo (C), precursor da membrana (prM), e glicoproteína do envelope (E). As proteínas não estruturais (NS) estão numeradas de 1 a 5. As linhas amarelas indicam os sítios de clivagem pela protease NS3, e os números indicam os aminoácidos ao longo da cadeia. Adaptada de (COX; STANTON; SCHINAZI, 2015).

A proteína NS1 desempenha um papel importante na evasão do sistema imunológico do hospedeiro durante a infecção viral. Embora essa proteína seja compartilhada com outros membros do gênero *Flavivirus*, suas características estruturais divergentes influenciam na especificidade da resposta imune. Além disso, no início da infecção,

essa proteína apresenta alta concentração, o que a torna um biomarcador em testes diagnósticos (BROWN et al., 2016; COX; STANTON; SCHINAZI, 2015; XU et al., 2016).

A NS2A é uma proteína que recruta o RNA viral, as poliproteínas não processadas C-prM-E e o complexo NS2B-NS3 para o local de montagem do vírion. Mutações nessa proteína tem a capacidade de interromper o processo de montagem, destacando sua importância. No entanto, devido ao seu tamanho reduzido, seu uso como alvo farmacológico pode ser dificultado (LEUNG et al., 2008; ZHANG et al., 2019).

As proteínas NS4A e NS4B atuam em conjunto na evasão do sistema imunológico do hospedeiro e têm sido associadas à indução de autofagia nas células do tronco neural, o que está relacionado com o surgimento de síndromes congênitas, como a microcefalia (MA et al., 2018). A NS4B também desempenha um papel na replicação viral, facilitando a formação do complexo de replicação viral e interagindo com outras proteínas não estruturais do vírus (WANG; XIE; SHI, 2022).

A proteína NS5 é conservada entre as diferentes linhagens de Zika vírus e desempenha papéis na replicação viral e na interação vírus-hospedeiro. Devido às suas funções e à existência de caracterização estrutural definida, ela é utilizada como alvo farmacológico em pesquisas para o desenvolvimento de inibidores (WANG et al., 2018).

Embora existam pesquisas para o desenvolvimento de inibidores com foco nas proteínas não estruturais NS4B e NS5, para esse trabalho, foi escolhido o complexo NS2B-NS3. A seguir, serão apresentados detalhes mais aprofundados sobre esse complexo e sua importância no contexto de replicação viral.

1.2.1 NS2B-NS3 protease

A proteína NS3 possui dois domínios funcionais distintos em sua estrutura. O primeiro é uma serino protease localizada na região N-terminal, responsável por realizar a clivagem da poliproteína viral em proteínas individuais, além de ter a capacidade de clivar algumas proteínas do hospedeiro. A segunda região, localizada na extremidade C-terminal, corresponde a uma helicase responsável pelo desenrolamento do RNA durante ao processo de replicação viral (GRUBA et al., 2016; TIAN et al., 2016; XU et al., 2019). A ativação da protease NS3 só ocorre na presença do cofator NS2B, pois

a ligação entre as proteínas gera uma conformação que forma o sítio de clivagem da poliproteína (XING et al., 2020).

A protease NS2B-NS3 do ZIKV (NS2B-NS3pro ZIKV) possui peso molecular de aproximadamente 26 kDa e apresenta um sítio ativo raso e exposto ao solvente. Ela é classificada como uma serino protease e contém a tríade catalítica composta pelos resíduos de aminoácidos His51, Asp75 e Ser135, que são conservados nessa família de proteases (LEI et al., 2016; PHOO et al., 2016) .

O sítio ativo da NS2B-NS3pro ZIKV possui em sua maioria densidade de carga negativa e é subdividido em cinco subsítios (figura 2). O subsítio S1 é composto principalmente pelos resíduos Tyr150, Tyr161, Asp129 e Ala132. O subsítio S1' inclui, além da histidina catalítica (His51), os resíduos Val52 e Thr53. O subsítio S2 contém o Asp75, componente da tríade catalítica, além dos aminoácidos Asp83 (exclusivo do Zika vírus, quando comparado a outras proteases relacionadas) e Ser81. O subsítio S3, embora menor, é formado, entre outros, pelos resíduos Gly128, Tyr127, Phe84 da NS2B. A região composta pelos resíduos Val154 e Val155 constitui o subsítio S4 (GRUBA et al., 2016; PHOO et al., 2016).

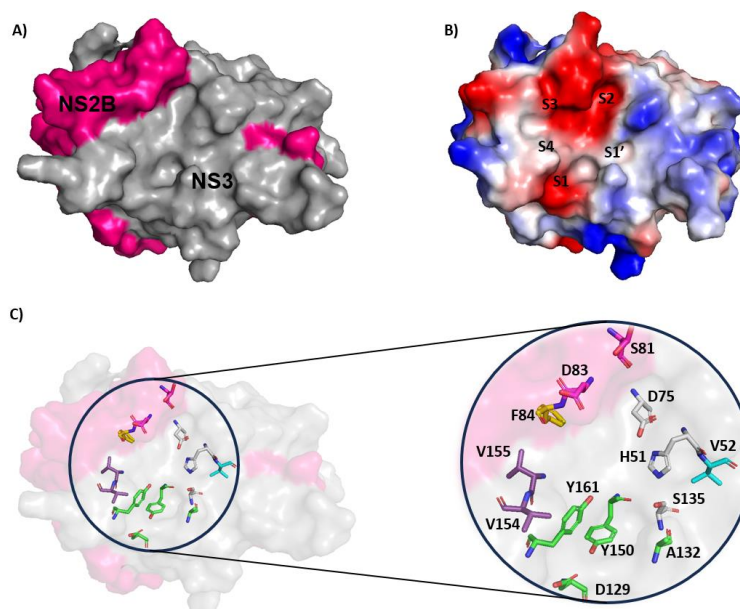


Figura 2 - Estrutura da protease NS2B-NS3pro de Zika vírus. A) Estrutura representada em superfície para a NS3 (cinza) e a NS2B (rosa); B) Superfície de Poisson-Boltzmann destacando regiões negativas (vermelho), positivas (azul) e neutras (branco). C) Sítio ativo da enzima destacando resíduos pertencentes aos subsítios S1 (verde), S1' (ciano), S2 (magenta), S3 (amarelo) e S4 (roxo) em sticks. A tríade catalítica está destacada em sticks da cor branca.

O planejamento de fármacos é um processo voltado para a identificação de novos compostos com atividade farmacológica, empregando tanto abordagens

experimentais quanto computacionais (ZHOU; ZHONG, 2017). Nesse contexto, o complexo NS2B-NS3pro se apresenta como um excelente alvo farmacológico não apenas devido à sua função essencial na replicação viral, mas também à sua conservação entre outros flavivírus (MAHAWATHTHA et al., 2017).

1.3 Inibidores da ZIKV NS2B-NS3pro

Proteases, de forma geral, representam alvos farmacológicos promissores. Exemplos clássicos incluem os inibidores da protease do vírus da imunodeficiência humana (HIV-1) (GHOSH; OSSWALD; PRATO, 2016) e os inibidores da protease da hepatite C (ZHANG, 2016). Dada a eficácia dessa classe de inibidores em outros vírus, é razoável considerar que a inibição das proteases dos flavivírus possa ser uma estratégia clínica bem-sucedida.

Um estudo recente de revisão abordou inibidores direcionados à NS2B-NS3pro dos vírus da Zika, dengue e do Nilo ocidental, que foram relatados na literatura entre 2015 e 2021 (SAMRAT et al., 2022). Dos 35 inibidores identificados, 22 demonstraram atividade inibitória contra a ZIKV NS2B-NS3pro, com valores médios de concentração inibitória (IC_{50}) variando entre 0,13 e 21,7 μ M. O artigo divide os inibidores entre competitivos e não competitivos.

Os inibidores competitivos interagem com a protease em seu sítio ativo, inibindo sua atividade enzimática. Dentro desse grupo, o artigo faz uma subdivisão entre inibidores reposicionados, inibidores sintéticos e compostos naturais. O primeiro grupo consiste em medicamentos já aprovados para outros fins, que são submetidos a testes para avaliar sua eficácia contra a ZIKV NS2B-NS3pro. Destaca-se nessa categoria, o novobiocin (**1**) (**figura 3**), que é um dos inibidores mais promissores contra o alvo em questão e o único dos compostos competitivos a ter passado por ensaios *in vitro* (YUAN et al., 2017).

A segunda subdivisão refere-se a compostos sintéticos. Inicialmente, essas pesquisas eram voltadas para o desenvolvimento de inibidores baseados em peptídeos, que poderiam ou não conter armadilha de serina, como a incorporação de ácido bórico para mimetizar o substrato da enzima. Porém, peptídeos frequentemente enfrentam desafios em termos de permeabilidade através das barreiras de membrana. Por esse motivo, houve uma mudança de foco em direção ao desenvolvimento de moléculas

pequenas como alternativa para inibir a protease (LEI et al., 2016; LI et al., 2018; YIN et al., 2006).

Dentro desse conjunto, podemos destacar alguns compostos potentes contra a ZIKV NS2B-NS3pro. Rassias et al. (2019) sintetizou uma série de compostos com base no grupo carbazol, resultando em inibidor (**2**) com uma IC_{50} de apenas $0,52 \mu M$ (**figura 3**) (RASSIAS et al., 2019). Outra abordagem empregada na descoberta de inibidores foi conduzida por Abrams et al. (2020) que utilizou triagens quantitativas de alto rendimento (do inglês *quantitative high throughput screenings – qHTS*) para identificar pequenas moléculas com baixo IC_{50} (**3**) (**figura 3**).

A última categoria de inibidores competitivos são os compostos naturais, que têm origem em fontes como plantas e microorganismos e frequentemente exibem propriedades medicinais. Um exemplo é o composto derivado do chá preto, theaflavin-3-3'-digallate (**4**), que demonstrou uma ação moderada contra a protease de Zika, com uma IC_{50} de $7,65 \mu M$ (**figura 3**) (CUI et al., 2020).

Inibidores não competitivos modulam a atividade enzimática ao se ligarem a locais distintos do sítio ativo da NS2B-NS3pro. Neste grupo, destacam-se dois inibidores (**5**, **6**) (**figura 3**), que não apenas exibem baixos valores de IC_{50} , mas também demonstraram efeitos protetores em organoides 3D de mini cérebros quando expostos ao Zika (LI et al., 2020a, 2020b).

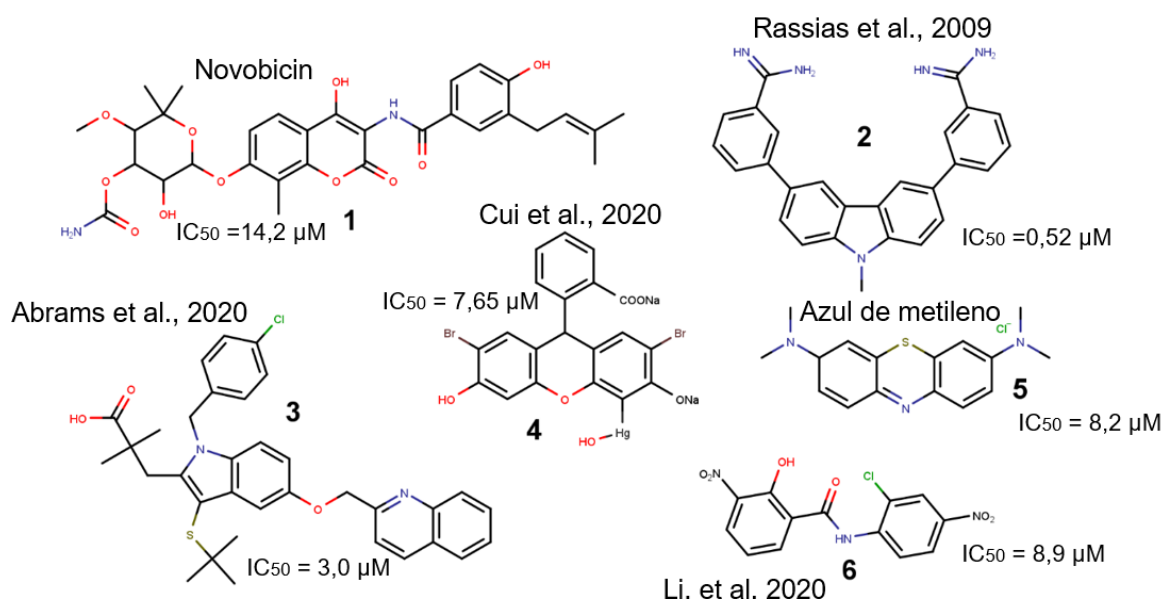


Figura 3 - Estrutura dos inibidores de NS2B-NS3pro ZIKV.

Além dos inibidores citados na revisão, que incluem peptídeos e pequenas moléculas, vale ressaltar outra categoria de inibidores: os macrociclos. Esses compostos são caracterizados por apresentar um sistema de anéis contendo 12 ou mais átomos em sua composição. Exemplos desse grupo incluem fármacos como a ciclosporina A e a rapamicina (ALLEN; DOKHOLYAN; BOWERS, 2016). A ciclização de peptídeos representa uma estratégia empregada para restringir a flexibilidade conformacional da molécula, resultando em melhorias em sua potência, seletividade e estabilidade. Esse aprimoramento é exemplificado no estudo de Braun *et al* (2020), no qual a ciclização de um peptídeo resultou em um aumento de 1000 vezes na capacidade de inibir a ZIKV NS2B-NS3pro em comparação com a estrutura acíclica correspondente (BRAUN *et al.*, 2020).

1.4 Desenvolvimento racional de fármacos

O desenvolvimento de novos medicamentos com alta eficácia e baixa toxicidade é uma demanda atual, no entanto, essa tarefa enfrenta diversos desafios significativos. O processo é notavelmente dispendioso, com custos que podem alcançar bilhões de dólares, e pode estender-se por mais de 15 anos, desde a descoberta de um protótipo até a aprovação final no mercado (CHANG *et al.*, 2022). Além disso, é importante destacar que aproximadamente 90% dos compostos iniciais não conseguem chegar ao mercado como medicamentos aprovados (CHANG *et al.*, 2022; EKINS; MESTRES; TESTA, 2007).

O processo de desenvolvimento de um medicamento é dividido em duas etapas principais. A primeira fase envolve os ensaios pré-clínicos que compreendem os testes realizados *in vitro* e *in vivo* para assegurar segurança e eficácia dos compostos testados. A segunda fase, conhecida como fase clínica, consiste em testes em seres humanos para definição de dose, comprovação de eficácia e segurança tendo como desfecho a aprovação, ou não, do medicamento no mercado (BRODNIEWICZ; GRYNKIEWICZ, 2010; CHANG *et al.*, 2022).

Entretanto, para iniciar a fase pré-clínica, é necessário identificar e otimizar um composto líder que será submetido aos testes de segurança e eficácia *in vitro* e *in vivo*. Nesse ponto, existem diversas abordagens disponíveis, que podem ser experimentais, como a triagem de alto rendimento (HTS), bem como técnicas computacionais (CHANG *et al.*, 2022; EKINS; MESTRES; TESTA, 2007; WALTERS;

STAHL; MURCKO, 1998). A utilização de ferramentas *in silico* tem como objetivo reduzir custos e otimizar o tempo de triagem.

Ao empregarmos ferramentas computacionais, nos deparamos com duas abordagens principais para identificação de compostos que possam ter efeito farmacológico: o desenvolvimento de fármacos baseados em ligantes e o desenvolvimento de fármacos baseados na estrutura do alvo molecular. Na primeira abordagem, temos um ligante conhecido, que será usado como ponto de partida para a busca de novos compostos. As técnicas podem ser empregadas independentemente de conhecimento prévio de um alvo farmacológico.. Portanto, algumas técnicas, como a busca por similaridade estrutural com outros ligantes já caracterizados, são empregadas, partindo do pressuposto de que se a estrutura for semelhante, há uma alta probabilidade de que a atividade farmacológica também o seja. Outra técnica é a relação quantitativa estrutura-atividade (QSAR), que calcula a correlação entre as propriedades moleculares de um grupo de compostos e sua atividade experimentalmente determinada (CHANG et al., 2022; EKINS; MESTRES; TESTA, 2007).

No caso do design de fármacos baseados na estrutura do alvo molecular, utilizada nesse trabalho, esta estrutura precisa ser conhecida. Nesse contexto, o *Protein Data Bank* (PDB) (BERMAN, 2000), é um banco de dados relevante pois contém milhares de estruturas de macromoléculas e seus complexos resolvidos experimentalmente. Quando não se tem determinação experimental de sua estrutura, ela pode ser obtida pela modelagem por homologia, que prevê a estrutura do alvo a partir da comparação de sua sequência proteica com proteínas que servirão de modelo para sua construção (CHANG et al., 2022; EKINS; MESTRES; TESTA, 2007). A partir da identificação do alvo de interesse, utilizamos a triagem virtual para identificação de compostos promissores.

Independentemente da forma como se obtêm os dados estruturais do alvo molecular, seja experimentalmente ou por meio da modelagem por homologia, uma ferramenta usada para realização da triagem virtual é o *docking* molecular (CHANG et al., 2022).

1.4.1 *Docking* molecular

O *docking* molecular é uma ferramenta computacional utilizada para prever a conformação de um complexo receptor-ligante, sendo o receptor geralmente uma

proteína ou molécula de ácido nucleico (RNA e DNA), enquanto o ligante pode ser uma molécula pequena ou outra proteína (BROOIJMANS; KUNTZ, 2003; DIAS; DE AZEVEDO JR., 2008). Essa abordagem desempenha um papel essencial na triagem de grandes bancos de dados de ligantes em um curto período de tempo, e com redução de custos, uma vez que não é necessário a síntese das estruturas avaliadas computacionalmente. Os pontos-chaves para um resultado de *docking* bem-sucedido incluem a velocidade e a acurácia do programa utilizado, seja para descobrir novos compostos líderes por meio de triagem virtual ou para reproduzir modos de ligação validados por dados experimentais (BROOIJMANS; KUNTZ, 2003).

A determinação do modo de ligação de um composto com seu alvo envolve algoritmos de busca conformacional adotados pelos programas de *docking*. Os algoritmos podem ser classificados em três tipos: determinísticos, estocásticos e sistemáticos. No algoritmo determinístico o estado inicial da molécula irá determinar a geração do próximo estado (BROOIJMANS; KUNTZ, 2003). A dinâmica molecular e a minimização de energia são exemplos de métodos determinísticos (BROOIJMANS; KUNTZ, 2003; KITCHEN et al., 2004). A dinâmica molecular simula o comportamento molecular ao longo do tempo para obter uma representação próxima da realidade biológica, considerando características como flexibilidade molecular e temperatura (BRUNO CÉSAR FELTES, 2015). Por sua vez, a minimização de energia busca a energia mínima local e geralmente é usada como complemento em outros algoritmos (KITCHEN et al., 2004).

Algoritmos estocásticos operam por meio de alterações aleatórias, frequentemente mudando os graus de liberdade do sistema a cada alteração. Um exemplo dessa categoria é o algoritmo genético, que utiliza conceitos de genética evolucionária para buscar a conformação de energia mínima do ligante. Nesse contexto, a posição espacial e os ângulos diedros de uma molécula são modelados em um vetor que é denominado cromossomo. O algoritmo inicia com uma população composta por um grupo pai de cromossomos, a qual é avaliada em relação a sua energia de ligação. Populações filhas são criadas a partir do cruzamento entre os componentes da população inicial, ocorrendo também mutações pontuais. Esse processo ocorre diversas vezes até obter uma conformação com a melhor pontuação (DIAS; DE AZEVEDO JR., 2008; KITCHEN et al., 2004).

Algoritmos sistemáticos baseiam-se em valores pré-estabelecidos para cada grau de liberdade, e cada um desses valores é explorado durante a busca conformacional. O algoritmo de construção incremental ou ancoragem e crescimento (do inglês *anchor-and-grow*) é um exemplo dessa classe (BROOIJMANS; KUNTZ, 2003; DIAS; DE AZEVEDO JR., 2008; KITCHEN et al., 2004). Nessa abordagem, um ligante flexível é tratado como uma coleção de segmentos rígidos. Cada segmento consiste em átomos adjacentes separados por ligações não rotacionáveis, e eles são separados pela presença de uma ligação rotacionável. Caso o ligante contenha um anel em sua estrutura, ele será tratado como um fragmento rígido. A construção incremental começa fixando o segmento rígido de maior tamanho em uma região da molécula alvo. Em seguida, os demais segmentos são adicionados a essa porção inicial até que o ligante seja reconstruído na cavidade. Entre cada adição ocorre um processo de minimização da estrutura (LANG et al., 2023).

Além da capacidade de identificar corretamente o modo de ligação entre o ligante e o alvo molecular, o *docking* deve ser capaz de ranquear eficazmente as melhores poses geradas. Para isso, são empregadas diferentes classes de funções de pontuação, sendo as três principais: físico-químicas (ou baseadas em campo de força), empíricas e baseadas em conhecimento (DIAS; DE AZEVEDO JR., 2008). As funções físico-químicas usam termos não-ligados típicos de campo de força para calcular a energia de ligação de cada pose gerada. As energias normalmente consideradas para esses cálculos são a van der Waals e eletrostática. Pontuações calculadas a partir de campo de força padrão possuem limitações por não considerarem energias de solvatação e termos entrópicos (KITCHEN et al., 2004). As funções empíricas utilizam diferentes termos de interações intermoleculares, cujo peso é ponderado por coeficientes obtidos através de análises de regressão que utilizam dados experimentais para construir a curva. Uma desvantagem desse método é a dependência de conjuntos de dados experimentais usados para realizar as análises de regressão. Por último, as funções baseadas em conhecimento (do inglês, *knowledge-based*) utilizam informações extraídas de estruturas experimentalmente resolvidas, identificando padrões de distâncias entre tipos atômicos que serão usados para prever o modo de ligação de outras moléculas (DIAS; DE AZEVEDO JR., 2008; KITCHEN et al., 2004).

Várias combinações de algoritmos de busca e função de pontuação são usadas, gerando uma variedade de programas de *docking* molecular. Nesse trabalho

utilizamos o Dock6 que utiliza o algoritmo de construção incremental em associação com a função de pontuação físico-química (LANG et al., 2023) representada na equação 1:

$$E_{vdW} + E_{electrostatic} = \sum_{prot} \sum_{lig} \left[\left(\frac{A_{ij}}{d_{ij}^a} + \frac{B_{ij}}{d_{ij}^b} \right) + 332.0 \frac{q_i q_j}{\epsilon(d_{ij}) d_{ij}} \right] \quad (1)$$

Onde a energia total é dada pela soma das energias de van der Waals (E_{vdW}) e energia eletrostática ($E_{electrostatic}$). Os termos "i" e "j" representam dois átomos quaisquer, A_{ij} e B_{ij} são os parâmetros de van der Waals para determinados tipos atômicos, d_{ij} é a distância interatômica, q_i e q_j são cargas parciais atômicas e $\epsilon(d_{ij})$ é uma função dielétrica dependente da distância (KITCHEN et al., 2004; LANG et al., 2023).

Independente da escolha do programa e protocolo empregado para triagem virtual, é necessário validar sua capacidade em identificar corretamente o modo de ligação do composto e classificar corretamente as poses geradas. Ao trabalhar com alvos moleculares resolvidos experimentalmente, pode-se utilizar as estruturas cristalográficas que contenham ligantes para testar a capacidade do protocolo em reproduzir a pose do ligante cristalográfico. Essa técnica pode ser denominada *re-docking* ou *cross-docking*. No *re-docking*, o ligante é separado de sua estrutura cristalográfica e o protocolo é utilizado para testar se houve reprodução do modo de ligação. No caso do *cross-docking*, o processo inicial é semelhante, separa-se a estrutura tridimensional do ligante, mas, em vez de usar a estrutura tridimensional nativa, o teste de reprodução do modo de ligação será feito em outra estrutura tridimensional da mesma proteína (MUKHERJEE; BALIUS; RIZZO, 2010; TEMML; KUTIL, 2021). Em ambos casos, o sucesso é determinado através do cálculo do RMSD (do inglês *Root Mean Square Deviation* – RMSD) entre o resultado de *docking* e o modo de ligação observado experimentalmente. O RMSD permite estimar a distância quadrada média entre átomos equivalentes de duas estruturas sobrepostas (ARMOUGOM et al., 2006). Seu cálculo é feito a partir da diferença entre as coordenadas x, y e z de um par de moléculas 1 e 2. Essas diferenças são elevadas ao quadrado, somadas e divididas pelo número total de átomos (n), sendo posteriormente tirada a raiz quadrada.

$$RMSD = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_{i1} - x_{i2})^2 + (y_{i1} - y_{i2})^2 + (z_{i1} - z_{i2})^2} \quad (2)$$

A outra validação está relacionada ao enriquecimento da triagem virtual, que seria a capacidade do protocolo em identificar ligantes ativos em comparação com os demais (BROZELL et al., 2012). Uma maneira de realizar essa avaliação é a por meio da análise das curvas ROC, que permitem a visualização da sensibilidade versus especificidade do protocolo adotado. Para construir uma curva ROC, é necessário ter um conjunto de ligantes ativos e um grupo de ligantes presumidamente, ou experimentalmente, inativos. Em uma triagem perfeita, espera-se que todos compostos ativos estejam classificados no topo, resultando em uma área sob a curva (AUC) igual a 1. Portanto, a expectativa é que um bom protocolo se aproxime desse valor. Por outro lado, quando há classificação aleatória dos ligantes ativos, ocorrerá a formação de uma linha cuja AUC = 0,5, resultado indesejado para um protocolo de triagem virtual (ALLEN et al., 2015; TRIBALLEAU et al., 2005). Além da AUC, o fator de enriquecimento pode ser usado para essa avaliação. Ele permite avaliar a proporção de ligantes encontrados pelo protocolo em uma determinada faixa da base triada, comparando-se com a proporção esperada caso os compostos da base de dados fossem ordenados de forma aleatória. O cálculo do fator de enriquecimento para uma porcentagem x da base triada é dada por:

$$Fator\ de\ enriquecimento_x = \frac{\left(\frac{A_x}{N_x}\right)}{\frac{A_{total}}{N_{total}}} \quad (3)$$

Onde A_x é o número de ligantes ativos presentes na faixa desejada (x%), A_{total} é o número total de compostos ativos, N_x é o número total de compostos em x% e N_{total} é o número de total de ligantes da base. O esperado é que um bom protocolo gere valores altos do fator de enriquecimento em baixas porcentagens da base triada (BROZELL et al., 2012).

1.5 Caracterização bioquímica de potenciais inibidores

Após a seleção de compostos de interesse, a validação de sua atividade biológica para NS2B-NS3pro pode ser feita por meio do monitoramento da clivagem de peptídeos contendo um grupo fluorescente. Esses peptídeos são projetados de forma a serem específicos para a protease e possuem em sua composição um grupo fluorescente, como a 7-Amino-4-Metilcumarina (AMC) (GRUBA

et al., 2016; KNIGHT, 1995). Quando ocorre a clivagem do peptídeo, há liberação do grupo AMC resultando no aumento da fluorescência observada no ensaio (KNIGHT, 1995). Portanto, para avaliar o efeito do inibidor, medimos a diminuição da atividade enzimática provocada por compostos inibidores. Na porção inicial das curvas geradas nesses experimentos, observamos um comportamento linear onde a velocidade é constante, permitindo a mensuração da taxa de reação. A velocidade medida nessas condições é denominada velocidade inicial. Os ensaios são conduzidos de forma a medir a velocidade inicial por ser a melhor medida da taxa de reação enzimática e é a fração da curva mais sensível à presença de inibição, tornando mais efetiva a detecção de inibidores durante a triagem (COPELAND, 2003).

Os ensaios fluorimétricos proporcionam a determinação de diferentes parâmetros cinéticos, constantes de inibição e valores de IC_{50} . A IC_{50} representa a concentração do inibidor que reduz a velocidade de reação enzimática em 50% em determinada condição de ensaio. Embora seja uma medida útil para classificar a potência dos compostos triados, é importante ressaltar que não se trata de um valor absoluto, uma vez que alterações nas condições do ensaio podem influenciá-la (COPELAND, 2003). Além de determinar a capacidade inibitória e potência de um composto, são necessários experimentos confirmatórios para descartar a possibilidade de que a inibição encontrada ocorra por mecanismos inespecíficos (SHOICHET, 2006).

2 Objetivos

O presente trabalho visa a identificação de potenciais inibidores contra a protease NS2B-NS3pro de Zika vírus a partir de técnicas *in silico* e *in vitro*, através dos seguintes objetivos específicos:

- Desenvolver e validar um protocolo de *docking* para triagem de pequenas moléculas contra a NS2B-NS3pro ZIKV;
- Selecionar moléculas da biblioteca BraCoLi a serem testadas experimentalmente contra a NS2B-NS3pro ZIKV a partir de triagem virtual;
- Avaliar experimentalmente a atividade biológica das moléculas selecionadas contra a NS2B-NS3pro de Zika.

3 Materiais e métodos

3.1 Análise das estruturas cristalográficas disponíveis da NS2B-NS3pro ZIKV

Com o intuito de analisar as estruturas cristalográficas previamente disponíveis da protease NS2B-NS3 do Zika vírus (ZIKV NS2B-NS3pro) foi realizada uma pesquisa no site RCSB, em 15/05/2023 (<https://www.rcsb.org/>), utilizando a sequência do inserto construída por Lei e colaboradores (LEI et al., 2016) como ponto de partida (figura 4).

```

MGSSHHHHHSSGLVPRGSHMVDMYIERAGDITWE
KDAEVTGNSPRLDVALDESGDFSLVEDDGPPMAGG
GGSGGGGSGALWDVPAPKEVKKGETTDGVYRVMT
RLLGSTQVGVGVMQEGVFHTMWHVTKGSALRSG
EGRLDPYWGDVKQDLVSYCGPWKLDAAWDGHSEV
QLLAVPPGERARNIQTLPGIFKTKDGDIGAVALDYPAG
TSGSPILDKCGRVIGLYGNGVVIKNGSYVSAITQGRR

```

Figura 4 – Sequência proteica resultante do inserto construído por Lei e colaboradores (2016). O destaque em amarelo corresponde a NS2B e o destaque em azul a NS3.

Essa busca resultou na identificação de um total de 59 estruturas. Além disso, foram coletadas manualmente informações adicionais sobre cada estrutura, sendo elas resolução global, ano de depósito, artigo de referência e classe à qual o ligante pertence.

A ferramenta *electron density*, fornecida pelo próprio site, foi utilizada para realizar uma análise visual da distribuição da densidade eletrônica dos ligantes a partir do mapa $2F_o - F_c$ σ 1,5. Cada ligante foi classificado em uma escala de 0 a 10, de acordo com a cobertura da densidade eletrônica observada em uma análise visual, sendo 0 indicando nenhuma cobertura e 10 uma cobertura total.

Outra ferramenta usada para avaliação da densidade eletrônica do ligante foi a *density fitness* disponibilizada pela PDB-REDO. Essa ferramenta utiliza o mapa de densidade eletrônica padrão $2F_o - F_c$ para calcular o ajuste de átomos individuais ou de um conjunto de átomos na densidade eletrônica, sendo esse cálculo denominado EDIA. Os valores de EDIA variam entre 0 e 1,2 e sua classificação é dividida em faixas:

- 0 – 0,4: átomos com inconsistências substanciais em relação ao ajuste da densidade eletrônica
- 0,4 – 0,8: átomos com inconsistências pequenas em relação ao ajuste da densidade eletrônica
- 0,8 – 1,2: átomos bem cobertos pela densidade eletrônica (MEYDER et al., 2017).

Para avaliar a similaridade global entre as estruturas selecionadas, foi realizada uma clusterização de todas as estruturas em relação umas às outras utilizando o pacote Bio3D do programa R (B.J et al., 2006). O RMSD foi utilizado como medida de similaridade.

Outro aspecto avaliado foi a análise da frequência e tipos de interações entre os ligantes cristalográficos e os resíduos do sítio de ligação. Para realizá-la, empregamos a ferramenta LUNA (FASSIO et al., 2022), para gerar gráficos que destacam os resíduos frequentemente envolvidos em interações com o alvo molecular, suas respectivas frequências e qual o tipo de interação realizada.

Com base nas informações sobre os resíduos que apresentaram um maior número de interações, foi conduzida uma nova clusterização a fim de investigar possíveis diferenças estruturais no interior do sítio. Foram considerados apenas dez resíduos com maior frequência de interação. Para realizar essa análise, uma estrutura foi selecionada como referência e o RMSD foi calculado em relação a todos os resíduos comparados a ela.

3.2 Seleção das estruturas tridimensionais

Além da análise dos resíduos frequentemente envolvidos em interações com ligantes, descrito na sessão 3.1, as estruturas tridimensionais foram separadas em cinco grupos com base na classificação de seus ligantes: acíclico, macrociclo, peptídeos, peptídeo-mimético e fragmento. Posteriormente, realizamos a comparação entre os ligantes dentro de cada grupo, utilizando a ferramenta PyMOL (SCHRÖDINGER, 2015) para uma análise visual. Durante essa avaliação, direcionamos nosso foco para os dez resíduos previamente identificados pelo LUNA como tendo uma frequência maior de interações. As estruturas foram alinhadas e sobrepostas, permitindo uma avaliação das semelhanças e diferenças entre os ligantes de cada categoria.

Após essa etapa, selecionamos pelo menos uma estrutura representante de cada categoria. Em seguida uma última comparação foi realizada entre as estruturas selecionadas, com o objetivo de identificar aquelas que apresentassem diferenças na conformação dos resíduos em relação às demais estruturas. Em situações de empate, nos quais não foram observadas diferenças conformacionais significativas nos resíduos dos sítios entre as diferentes estruturas tridimensionais, recorreremos aos critérios adicionais de resolução global da proteína e densidade eletrônica do ligante.

3.3 Preparação das estruturas tridimensionais proteicas

As estruturas tridimensionais proteicas foram preparadas com a ferramenta “*dockPrep*” presente no Chimera (PETTERSEN et al., 2004) seguido das etapas necessárias para criação das esferas, caixa e grid necessárias para o funcionamento do Dock6 (ALLEN et al., 2015). Durante o processo de preparação, o estado de protonação da histidina 51 foi alterado para HID, uma vez que estudos prévios demonstram que este é o estado tautomérico preferencial para esta proteína (SANTOS; CAFFARENA; FERREIRA, 2022). A caixa foi criada a com base nas esferas centralizadas nos resíduos His51, Ser135 e Tyr161. Essa escolha foi feita levando em consideração a observação de cobertura dos resíduos previamente classificados pelo LUNA como mais frequentes para ligações intermoleculares. As dimensões das caixas podem ser verificadas na **tabela 1**.

Tabela 1 - Dimensões das caixas utilizadas para o *docking*

| Dimensões Código (Å) | X | Y | Z |
|-------------------------|------|------|------|
| 5YOF | 22,7 | 30,6 | 35,8 |
| 5ZMS | 33,2 | 19,5 | 36,7 |
| 6L4Z | 20,6 | 29,0 | 33,8 |
| 7ZLC | 28,4 | 21,2 | 35,9 |
| 6KK2 | 22,1 | 28,3 | 34,2 |

3.4 Preparação dos ligantes

Todos ligantes utilizados nesse trabalho passaram por um processo de preparação semelhante, diferindo apenas na obtenção dos SMILES, formato usado para início da preparação. Quando esses não estavam disponíveis, foi necessário desenhar a

estrutura do ligante no Marvin Sketch (MARVIN 23.2, 2023), baseado na estereoquímica fornecida nos artigos correspondentes, para obtê-los. A protonação dos ligantes foi feita utilizando a ferramenta FixpKa, que retorna o estado de ionização mais abundante no pH 7,4. A obtenção da conformação tridimensional foi realizada pelo OMEGA, que calculou um conformero de menor energia. Vale ressaltar que foram empregadas duas funções específicas, para macrociclos foi utilizada a função -*macrocycle* e no caso das demais utilizou-se a função -*classic*. Por último, para calcular as cargas atômicas parciais AM1BCC, foi utilizada a ferramenta MolCharge. Todas essas ferramentas disponibilizadas pela OpenEye Scientific Software (HAWKINS et al., 2010; OPENEYE, 2019).

3.5 Protocolo adotado para o *docking* molecular

O programa de *docking* escolhido foi o Dock6 (ALLEN et al., 2015), do qual adotamos o protocolo para ligantes flexíveis com estrutura proteica rígida e pontuação baseada em grid (equação 1).

Os segmentos rígidos dos ligantes, contendo no mínimo 20 átomos, foram considerados como âncoras, limitados a um número máximo de 100 orientações. Para cada ligante, foram geradas até 500 conformações, e 10 poses foram classificadas para avaliação. A pose inicial, ou seja, a melhor ranqueada, foi empregada na maior parte das análises.

3.6 *Cross-docking*

Para realização do *cross-docking* foram selecionadas cinco estruturas cristalográficas, seguindo critérios previamente estabelecidos (tópicos 3.1 e 3.2). Os ligantes utilizados foram selecionados dentre os ligantes cristalográficos disponíveis. Foram utilizados os ligantes cuja densidade eletrônica foi considerada igual ou superior a 6 a partir da análise de inspeção visual, com exceção do ligante correspondente à estrutura 6KK2 que apresentou uma densidade mais baixa. Outro critério adotado foi a existência de artigo para a estrutura. No total, obtivemos 25 ligantes que foram desenhados utilizando a ferramenta Marvin Sketch (MARVIN 23.2, 2023).

3.6.1 Avaliação dos resultados do *cross-docking*

Para avaliar a eficácia do *cross-docking* em reproduzir as poses cristalográficas, conduzimos uma análise de RMSD usando um *script* fornecido pelo Schrodinger (SCHRÖDINGER, 2023). Realizamos dois tipos de análise. O primeiro foi direcionado

para a identificação da reprodução da pose cristalográfica. Nessa análise, a posição dos ligantes no interior do alvo molecular é levado em consideração. Antes de realizar o cálculo de RMSD, a estrutura tridimensional do ligante cristalográfico (referência) analisado foi alinhado com a estrutura tridimensional do *cross-docking* utilizando o PyMOL. Após esse alinhamento, o cálculo de RMSD foi feito utilizando o *script* abaixo:

```
rmsd.py -use_neutral_scaffold referencia.mae pose.mae
```

A opção “-use_neutral_scaffold” indica que o programa reduzirá a estrutura ao seu esqueleto, contendo apenas átomos pesados neutros conectados por ligações simples, facilitando a comparação das estruturas.

3.7 Comparação das conformações estruturais dos ligantes

Visando comparar a reprodução da conformação cristalográfica do ligante, realizamos outro tipo de análise utilizando o RMSD. Nesse caso, não consideramos a posição dentro do sítio ativo, apenas a conformação dos ligantes. Portanto, nessa análise, a sobreposição das estruturas foi feita diretamente pelo script descrito na seção 3.6.2 adicionando a flag “-m” que realiza essa sobreposição. Aqui avaliamos as conformações tanto da estrutura usada como arquivo de entrada para o *docking*, quando a conformação das poses resultantes após o processamento do *docking*.

3.8 Avaliação do enriquecimento do protocolo

Para avaliar a capacidade do protocolo em distinguir e priorizar ligantes experimentalmente ativos (verdadeiros positivos) de outros ligantes (falsos positivos), construímos três conjuntos de ligantes que foram avaliados através da construção de curvas ROC.

O conjunto de ligantes ativos é formado pelas estruturas usadas no *cross-docking*, além de outras mencionadas em seus artigos de origem (BRAUN et al., 2020; HAMMERSCHMIDT et al., 2023; HUBER et al., 2022; LI et al., 2017, 2018; NITSCHKE et al., 2019; PATIL et al., 2021; PHOO et al., 2018; QUEK et al., 2020; XIONG et al., 2022). Foram considerados ligantes ativos aqueles que apresentaram valores de K_i ou IC_{50} inferiores a 50 μ M.

Os ligantes ativos foram avaliados quanto sua similaridade utilizando duas métricas: coeficiente de Tanimoto e clusterização utilizando fingerprints de Morgan. Ambos foram calculados a partir de um script que utilizam ferramentas do RDKit. O fingerprint

de Morgan utilizado avalia, para cada átomo, um raio de 3 Å, se determinado grupo químico é encontrado ele é representado por 1 e, caso esteja ausente, por 0. Após classificar os grupos químicos de todos ligantes, eles são divididos em clusters de acordo com sua similaridade. O coeficiente de Tanimoto, por sua vez, é uma representação gráfica da similaridade entre dois ligantes, portanto ele irá utilizar os resultados gerados pelo fingerprint de Morgan para gerar o heatmap (BAJUSZ; RÁCZ; HÉBERGER, 2015; ZHONG; GUAN, 2023).

O primeiro grupo de teste é composto por decoys gerados através do DUD-E, ferramenta que explora o banco de dados ZINC20 em busca de estruturas com características físico-químicas semelhantes aos ligantes ativos: massa molecular, logP, número de ligações rotacionáveis e número de doadores e aceptores de ligação de hidrogênio. Para que os decoys funcionem como controle negativo, eles não devem ser capazes de se ligar ao sítio de interesse. Portanto, o programa utiliza fingerprints de similaridade 2D para reduzir a similaridade topológica entre os ligantes e decoys. Os decoys computacionalmente gerados foram moléculas com propriedades físico-químicas semelhantes, porém estruturalmente distintas dos ligantes, portanto espera-se que eles não sejam ativos contra ZIKV NS2B-NS3pro. O objetivo é gerar decoys desafiadores para o *docking*, mantendo características físicas semelhantes aos ligantes, ao mesmo tempo em que são topologicamente distintos, diminuindo a probabilidade de interação com o alvo. O programa busca gerar cinquenta decoys para cada ligante ativo utilizado (MYSINGER et al., 2012).

O outro grupo de ligantes que foi testado são ligantes experimentalmente inativos. Para obtê-los foi realizada uma busca no bando de dados PubChem resultando em diversos artigos que continham informações sobre moléculas ativas e inativas contra NS2B-NS3pro ZIKV. O artigo de triagem experimental de alto rendimento (*High-throughput screen* -HTS) (ABRAMS et al., 2020) foi selecionado por conter uma ampla base de dados com milhares de compostos caracterizados. A partir desse artigo, foram selecionadas as moléculas inativas.

No total, foram obtidos 73 ligantes ativos, 307 decoys e 3617 inativos, preparados de acordo com o descrito na seção preparação dos ligantes (3.4). Os ligantes ativos foram desenhados no Marvin Sketch. Foram realizados *docking* para ambos conjuntos teste seguindo o protocolo descrito em no item 3.5. Por fim a curva ROC e a áreas

sob a curva (AUC) foram calculadas utilizando a biblioteca “ROCR” do R (SING et al., 2005), para os grupos decoys e compostos inativos.

3.9 Triagem virtual

O *docking* foi realizado de acordo com o protocolo estabelecido nesse trabalho, utilizando a biblioteca BraCoLi, base de dados desenvolvida por professores da faculdade de farmácia da UFMG e coordenada pelo professor Vinícius Maltarollo, que é composta por 1.076 compostos (VERÍSSIMO et al., 2022).

Os resultados do *docking* foram avaliados a partir do ranqueamento realizado pelo Dock6, e os 150 primeiros ligantes foram submetidos a uma inspeção visual. Essa inspeção foi feita no programa Chimera e levou em consideração diversos aspectos. A complementariedade estérica foi a primeira etapa, onde se analisou o encaixe do ligante na cavidade do sítio ativo da proteína. A presença de ligações de hidrogênio também foi considerada e ligantes que não apresentavam nenhuma foram descartados. A última parte da análise foi a observação da interação entre os ligantes e os resíduos identificados pelo LUNA como mais frequentes em interações intermoleculares nos complexos cristalográficos entre alvo molecular e ligante.

Para complementar a inspeção visual, o software Discovery Studio (SYSTÈMES, 2020) foi utilizado para realizar o mapeamento 2D de interações entre o ligante e o alvo molecular. Além de ligações de hidrogênio, o programa consegue identificar outros tipos de ligações favoráveis, como interações de van der Waals, pi-stacking, hidrofóbicas, empilhamentos aromáticos, iônicas, cátion-pi, cátion-nucleófilo e multipolares. Além disso, ele também permite a detecção de possíveis impedimentos estéricos (*clashes*) entre o ligante e o alvo molecular.

Com base nos resultados da inspeção visual e da análise do Discovery Studio, foram selecionados os compostos que apresentaram características promissoras e interações favoráveis com o alvo molecular. Esses compostos foram escolhidos para serem submetidos a ensaios enzimáticos, visando avaliar sua atividade potencial como inibidores da NS2B-NS3pro ZIKV.

3.10 Ensaios *in vitro* para caracterização de inibidores

3.10.1 Teste de solubilidade

Os ensaios enzimáticos foram conduzidos de acordo com os protocolos estabelecidos pelo Laboratório de Modelagem Molecular e Planejamento de Fármacos (LMMPF) do Departamento de Bioquímica e Imunologia da Universidade Federal de Minas Gerais.

Para a correta determinação da atividade inibitória dos ligantes, é necessário que eles sejam solúveis nas condições experimentais. Realizamos um teste de solubilidade dos compostos em um tampão composto por 10 mM de Tris, 5% de glicerina, 1 mM de CHAPS, pH ajustado para 8,5 usando HCl (1mol/L), diluído em água destilada.

Os compostos foram previamente diluídos em DMSO para obter uma solução estoque de 10 mM. Os experimentos de solubilidade foram realizados de duas maneiras diferentes:

1. Teste em placa de 96 poços: Foram adicionados 100 μ L do tampão e 1 μ L de solução estoque do composto. A mistura foi homogeneizada com pipeta, seguida por homogeneização manual.
2. Teste em microtubos de 1,5 mL: Foram adicionados 200 μ L do tampão e 2 μ L de solução estoque do composto. A mistura foi homogeneizada com pipeta e centrifugada por aproximadamente quinze segundos.

Em ambos testes, a concentração final dos compostos foi de 100 μ M. Após a realização dos experimentos, observou-se se houve formação de algum tipo de precipitado na solução.

3.10.2 Ensaios para caracterização bioquímica

A caracterização bioquímica foi realizada utilizando o fluorímetro Biotek Synergy 2, pertencente ao laboratório de multiusuários do departamento de bioquímica e imunologia da UFMG. Os ensaios fluorimétricos baseiam-se na mensuração da liberação de produtos fluorescentes resultantes da clivagem do substrato Bz-Nle-Lys-Lys-Arg-AMC, com excitação em 340 nm e emissão em comprimentos de onda de 440 nm. Os experimentos foram realizados em placas pretas de 96 poços e com fundo em U, em triplicata e em dois experimentos independentes com uma leitura de dois minutos. Uma solução tampão contendo 10 mM de Tris-HCl, 20% de glicerina e 1 mM

de CHAPS foi preparada com pH ajustado para 8,5. As concentrações finais de substrato e enzima nos ensaios foram respectivamente 44 μM e 0,2 nM.

3.10.2.1 Triagem inicial

A triagem inicial dos compostos para detecção de inibição foi realizada com a concentração final dos compostos de 100 μM e em duas condições distintas. A primeira foi o ensaio com pré-incubação onde 1 μL do composto, juntamente com 50 μL da solução de enzima, foi adicionado ao poço da placa. A placa foi incubada por 10 minutos a 37 °C. Após esse tempo, 50 μL da solução de substrato foi adicionada para realizar a leitura. Compostos que apresentaram inibição nessas condições foram submetidos ao segundo experimento para avaliar se a inibição era tempo dependente. Para isso foi realizado o ensaio sem incubação onde 1 μL da solução estoque do composto, juntamente com 50 μL da solução de substrato foi adicionado a um poço de uma. Em seguida, foi adicionado 50 μL da solução da enzima e a leitura foi realizada.

3.10.2.2 Determinação de valores de IC_{50}

Os ensaios de IC_{50} foram realizados em sete concentrações diferentes de inibidor, variando de 400 μM a 96 nM. Ele foi realizado em duplicata e em dois experimentos independentes, os demais parâmetros foram os mesmos descritos em 3.9.2. A ordem de adição dos elementos, e suas quantidades, foram a mesma utilizada para os ensaios de triagem inicial com incubação (seção 3.9.2.1). Antes desse ensaio, foi realizado um teste de solubilidade com o tampão (seção 3.9.1) para a concentração de 400 μM . O ajuste da curva foi realizado com o programa Graph Pad Prism 6.0 (GRAPHPAD SOFTWARE, 2023), usando a equação de quatro parâmetros para IC_{50} :

$$Y = \text{Base} + \frac{(\text{topo} - \text{base})}{1 + 10^{\log(\text{IC}_{50} - x) * s}} \quad (4)$$

Onde x é a concentração do inibidor, s é a inclinação da curva concentração-resposta (*Hill slope*), e o topo e a base são, respectivamente, os platôs superior e inferior.

3.10.2.3 Avaliação da inibição na presença de BSA

Para esse ensaio foi empregado o tampão sem a presença de CHAPS, contendo apenas Tris e glicerina nas mesmas concentrações descritas anteriormente (3.9.2). O composto, na concentração próxima à IC_{50} , foi incubado por 10 minutos com 25 μL

BSA (concentração de 4 mg/mL, resultando em concentração final de 1 mg/mL no ensaio), seguido pela adição de 25 μ L enzima. Uma segunda incubação de 10 minutos foi realizada antes da adição de 50 μ L do substrato iniciando a reação enzimática. Em ambos ensaios as concentrações enzimáticas e do substrato permaneceram iguais as empregadas na triagem inicial.

3.10.2.4 Ensaio de variação de concentração enzimática

Com exceção da concentração final de enzima, esse ensaio foi conduzido mantendo todos os parâmetros descritos em 3.9.2. Foram avaliadas duas concentrações enzimáticas, 0,2 nM (concentração padrão) e 2 nM. Ele foi realizado com pré-incubação com a enzima, seguindo as mesmas quantidades e ordem de adição das soluções descritas em 3.10.2.1 e a concentração dos compostos usados foi próxima à IC_{50} .

3.10.2.5 Ensaio de variação da concentração do substrato

Com exceção da concentração final do substrato e o número de réplicas empregado, esse ensaio foi conduzido mantendo todos os parâmetros descritos em 3.9.2. Foram avaliadas cinco concentrações substrato (88 μ M, 44 μ M, 22 μ M, 11 μ M e 5,5 μ M), em duplicata. O ensaio foi realizado com pré-incubação com a enzima e a concentração dos compostos usados foi próxima à IC_{50} .

4 Resultados e discussão

4.1 Similaridade global das estruturas de NS2B-NS3pro de Zika

O primeiro passo para realização de um *docking* é a seleção das estruturas tridimensionais a serem utilizadas. Essa escolha foi baseada em 49 estruturas disponíveis no PDB da protease NS2B-NS3 de Zika, todas obtidas por difração de raios-X e contendo ligantes. Para cada uma dessas estruturas, foram obtidas informações suplementares, como resolução, artigo de referência, quando disponível, e a classe do ligante. A resolução cristalográfica das estruturas pode ser dividida em três faixas: menor ou igual a 1,51 Å, entre 1,51 e 2,00 Å e maior que 2,00 Å. Apenas 18% das estruturas apresentam resolução pior que 2,00 Å, a maior parte (59%) está concentrada na faixa intermediária e 22% apresentam resoluções melhores que 1,51 Å (tabela 2).

Tabela 2 – Informações sobre as estruturas cristalográficas e inspeção visual de densidade eletrônica

| Código PDB | Resolução (Å) | Classe do ligante | Densidade eletrônica do ligante | |
|------------|---------------|-------------------|--|------------------------------|
| | | | Inspeção visual ¹ (0-10) | EDIA ² (0-1,2) |
| 7ZLC | 1,75 | acíclico | 8,0 | 0,65 |
| 7ZNO | 1,70 | acíclico | 7,0 | 0,39 |
| 5H4I | 2,00 | fragmento | 5,0 | 0,17 |
| 5YOD | 1,90 | fragmento | 9,5 | 0,71 |
| 6L4Z | 1,90 | fragmento | 9,5 | 0,77 |
| 6L50 | 1,95 | fragmento | 9,0 | 0,65 |
| 7VXX | 1,90 | fragmento | 8,0 | 0,55 |
| 6KK2 | 2,02 | macrociclo | 4,0 | 0,13 |
| 6KK3 | 2,05 | macrociclo | 7,0 | 0,57 |
| 6KK4 | 1,74 | macrociclo | 8,0 | 0,98 |
| 6KK5 | 2,03 | macrociclo | 9,0 | 0,68 |
| 6KK6 | 1,74 | macrociclo | 9,5 | 0,98 |
| 6KPQ | 2,62 | macrociclo | 7,0 | 0,27 |
| 6Y3B | 1,59 | macrociclo | 9,5 | 1,00 |
| 7O2M | 1,90 | macrociclo | 8,0 | 0,58 |
| 7O55 | 1,95 | macrociclo | 6,0 | 0,67 |
| 7OBV | 1,30 | macrociclo | 6,0 | 0,98 |
| 7OC2 | 1,50 | macrociclo | 6,0 | 0,95 |
| 7PFQ | 1,45 | macrociclo | 10,0 | 0,97 |
| 7PFY | 1,38 | macrociclo | 9,5 | 0,95 |
| 7PFZ | 1,45 | macrociclo | 10,0 | 0,77 |
| 7PG1 | 1,95 | macrociclo | 7,0 | 0,71 |

| | | | | |
|------|------|-------------------|------|------|
| 7PGC | 1,55 | macrociclo | 5,0 | 0,88 |
| 7VLG | 1,77 | macrociclo | 8,0 | 0,93 |
| 7VLH | 2,62 | macrociclo | 6,0 | 0,24 |
| 7VLI | 2,38 | macrociclo | 5,0 | 0,45 |
| 7ZLD | 1,61 | macrociclo | 9,0 | 0,83 |
| 7ZMI | 2,15 | macrociclo | 5,0 | 0,15 |
| 7ZPD | 1,40 | macrociclo | 10,0 | 0,95 |
| 7ZQ1 | 1,52 | macrociclo | 10,0 | 0,98 |
| 7ZQF | 1,68 | macrociclo | 10,0 | 0,98 |
| 7ZTM | 1,45 | macrociclo | 10,0 | 1,00 |
| 7ZUM | 1,75 | macrociclo | 9,5 | 0,76 |
| 7ZV4 | 1,69 | macrociclo | 9,5 | 0,47 |
| 7ZWK | 2,00 | macrociclo | 8,0 | 0,81 |
| 7ZYS | 1,26 | macrociclo | 9,5 | 0,75 |
| 8AQA | 1,35 | macrociclo | 10,0 | 0,97 |
| 8AQB | 1,28 | macrociclo | 10,0 | 0,99 |
| 8AQK | 1,20 | macrociclo | 10,0 | 1,00 |
| 5GPI | 1,58 | peptídeo | 10,0 | ND |
| 5H6V | 2,42 | peptídeo | 8,0 | 0,47 |
| 5YOF | 1,51 | peptídeo | 9,5 | 0,77 |
| 6JPW | 1,95 | peptídeo | ND | 0,44 |
| 7DOC | 1,90 | peptídeo | ND | 0,32 |
| 7VXY | 1,90 | peptídeo | ND | 0,30 |
| 5LC0 | 2,70 | peptídeo mimético | 9,0 | 0,36 |
| 5ZMQ | 1,99 | Peptídeo mimético | ND | 0,75 |
| 5ZMS | 1,80 | Peptídeo mimético | 7,0 | 0,83 |
| 5ZOB | 2,00 | Peptídeo mimético | 5,0 | 0,63 |

¹ Inspeção visual realizada usando a ferramenta “*eletron density*” disponibilizada pelo site RCSB. ² Cálculo feito utilizando o programa “*density fitness*”. ND: Ligantes que não foram identificados pela ferramenta *eletron density* presente no site RCSB, ou que não foram processadas pelo *density fitness*

A fim de identificar possíveis diferenças entre as estruturas cristalográficas, realizamos uma análise de similaridade global entre elas, utilizando o cálculo do RMSD. A **Figura 5** ilustra os resultados dessa análise, revelando que a maioria das estruturas apresenta um RMSD próximo de 0,5 Å, indicando uma boa sobreposição entre elas. No entanto, é importante ressaltar que a escala do RMSD foi limitada a, no máximo, 1 Å com o objetivo de facilitar a visualização das diferenças estruturais entre as estruturas tridimensionais. Essa limitação se deve à estrutura 7MV1, que apresentou um valor médio de RMSD de 4,5 Å em relação as demais estruturas. Essa diferença pode ser explicada pelo fato de que essa estrutura corresponde a uma conformação super aberta, enquanto as demais correspondem à conformação fechada (**figura 6**)

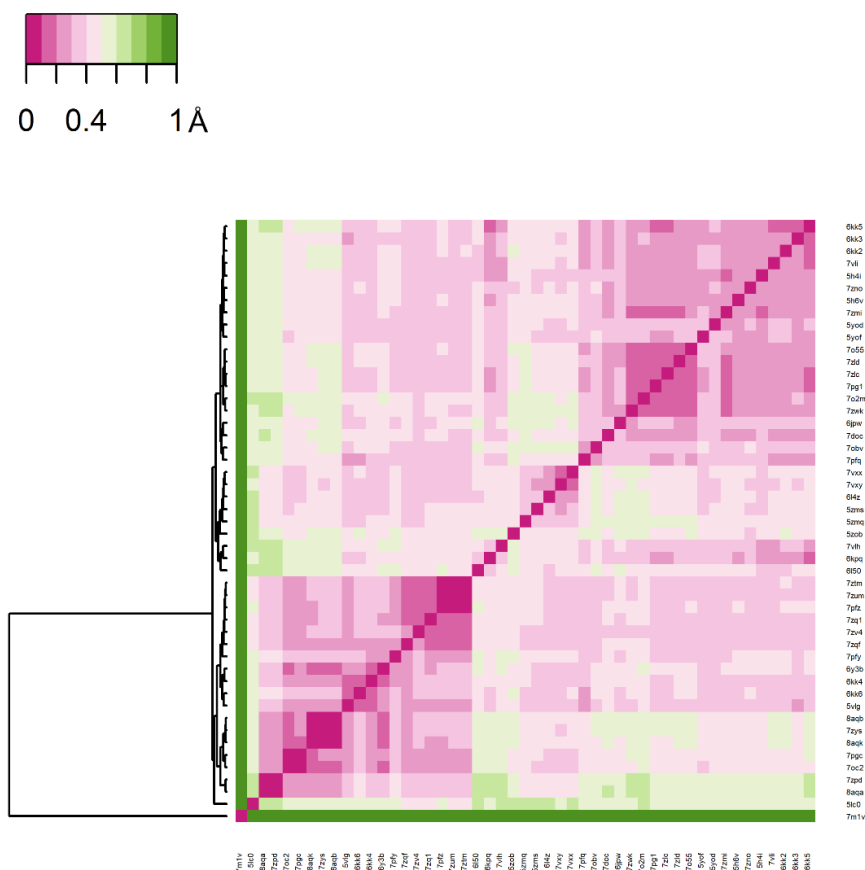


Figura 5 – Análise estrutural das 49 estruturas da NS2B-NS3pro ZIKV com ligantes disponíveis no PDB. O heatmap representa o desvio quadrático médio (RMSD) dos carbonos alfa (C α) entre as estruturas. As variações de RMSD são codificadas por cores, indo de 0 (rosa escuro) a 1,0 Å (verde), destacando-se os maiores valores associados à presença da estrutura 7M1V, cuja conformação é super aberta.

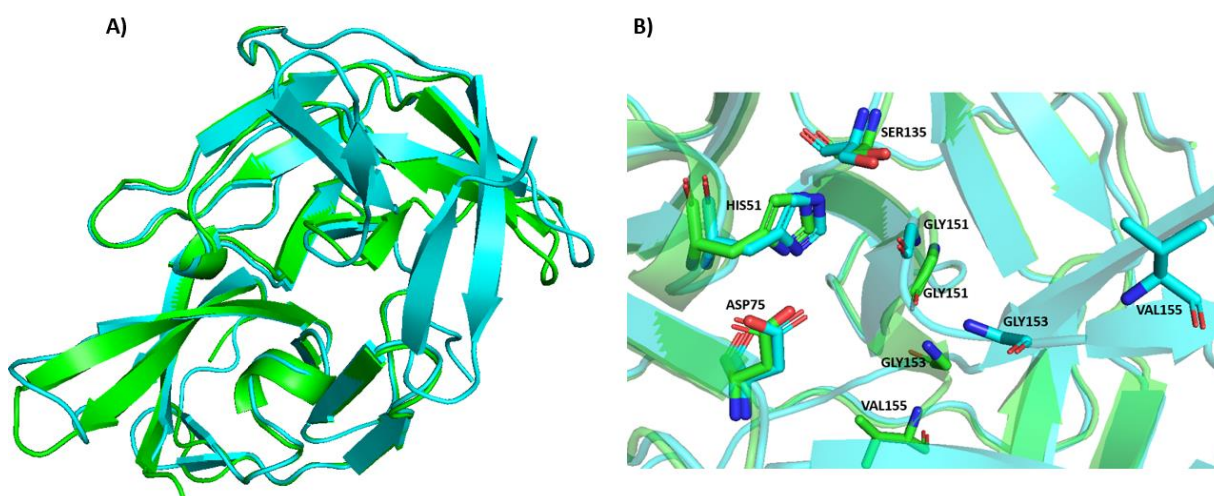
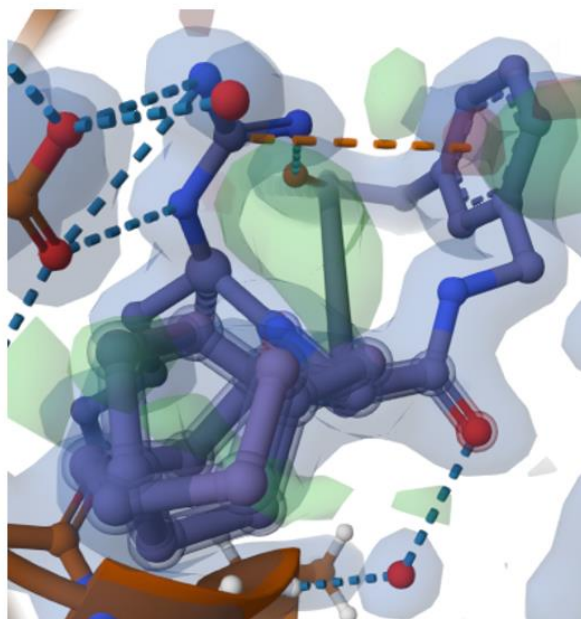


Figura 6 – Diferença estrutural entre a estrutura 7M1V e 5YOF. Alinhamento das duas estruturas, sendo 7M1V em rosa e 5YOF em azul. A) Representação em cartoon da sobreposição de ambas proteínas. B) Sobreposição dos resíduos do sítio ativo, podemos observar que os resíduos da tríade catalítica estão localizados na mesma região, enquanto há uma grande diferença da posição de resíduos como GLY151, GLY153 e VAL155, destacados na imagem.

Outro ponto importante a ser avaliado é a qualidade dos ligantes. O PDB possui uma métrica de qualidade para avaliar o ajuste da estrutura do ligante aos dados experimentais denominada “wwPDB Validation”. No entanto, essa métrica não pôde ser aplicada como parâmetro em nossas análises devido à ausência dessa classificação na maioria das estruturas tridimensionais analisadas, o que comprometeria sua precisão como método de avaliação. Dessa forma, a avaliação dos ligantes foi realizada mediante inspeção visual da densidade eletrônica, utilizando a ferramenta *electron density* disponível no site do PDB e pela ferramenta *density fitness* fornecida pelo PDB-REDO (**tabela 2**). A avaliação feita utilizando a *electron density* foi subjetiva com atribuição de notas que variavam de 0, indicando nenhuma cobertura do ligante, até 10, indicando cobertura total do ligante. A menor cobertura observada foi 4, enquanto a maior cobertura foi 10 (**figura 7**). Cerca de 80% das estruturas foram classificadas como tendo cobertura igual ou superior a 6. É importante ressaltar que para as estruturas 5ZMQ, 6JPW, 7DOC e 7VXY, a maioria pertencente à classe dos peptídeos, a ferramenta não conseguiu identificar a localização do ligante, impossibilitando a avaliação de sua densidade eletrônica.

7PFZ – Inspeção visual = 10,0



6KK2 – Inspeção visual = 4,0

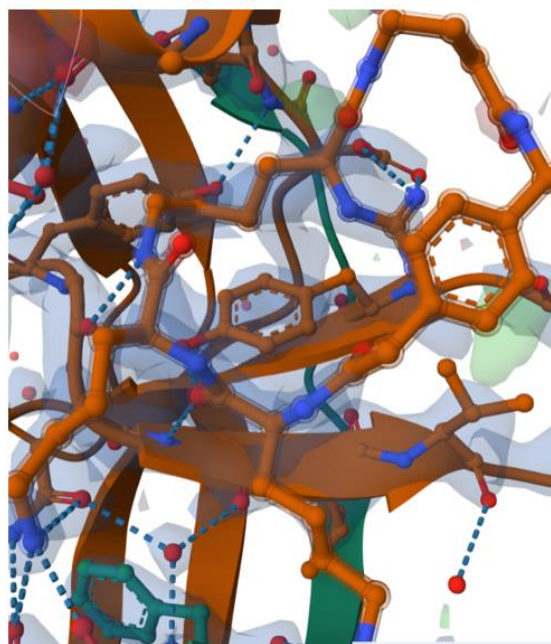


Figura 7 – Comparação da cobertura da densidade eletrônica de ligantes utilizando a ferramenta *density fitness* do site RCSB. Mapa de densidade $2Fo-Fc$ σ 1,5, onde a cobertura do ligante está sendo representada em azul. Observa-se que na estrutura 7PFZ, a densidade eletrônica cobre integralmente o ligante, enquanto na estrutura 6KK2 diversas partes do ligante encontram-se fora da área de cobertura.

A avaliação pela *density fitness* é feita de forma automatizada através do cálculo de EDIA que é o ajuste de múltiplos átomos na densidade eletrônica do mapa 2Fo-Fc fornecido para a ferramenta. O EDIA varia de 0 a 1,2 e a avaliação é dividida em três faixas que avaliam a cobertura da densidade eletrônica, sendo 0-0,4 correspondente a um baixo ajuste dos átomos em relação a densidade eletrônica, 0,4-0,8 indicando pequenas inconsistências em relação ao ajuste da densidade e 0,8-1,2 indicando uma boa cobertura pela densidade eletrônica (MEYDER et al., 2017). Considerando essa divisão temos cerca de 35% dos ligantes com boa cobertura.

Ao comparar os resultados obtidos pela análise visual com os da ferramenta automatizada, observam-se algumas discrepâncias na avaliação da cobertura dos ligantes com uma superestimação da qualidade da cobertura. Um exemplo é o ligante da estrutura 5LC0, que na inspeção visual, foi considerado como 9, enquanto, ao ser processado pela *density fitness*, seu valor foi de 0,36, se enquadrando-se na faixa mais baixa do programa. Tais diferenças destacam a importância de métodos sistematizados para avaliação de diferentes parâmetros.

4.2 Diferenças estruturais entre os resíduos do sítio ativo entre as estruturas tridimensionais

Uma vez que a similaridade global das estruturas cristalográficas foi evidenciada, indicando alta similaridade global das estruturas terciárias, realizamos uma análise mais específica do sítio ativo para comparar as conformações dos resíduos. Para identificar quais resíduos de interesse, utilizamos o LUNA, que proporcionou um panorama abrangente das interações estabelecidas pelos ligantes cristalográficos com a protease, conforme observado na **figura 8**. Essa análise permitiu identificar os resíduos mais frequentes envolvidos nessas interações, com suas respectivas frequências de interação, sendo eles: Gly151 (91,7%), Tyr130 (89,6%), Ser135 (89,6%), Tyr161 (89,6%), His51 (87,5%), Asp129 (85,4%), Tyr150 (77,1%), Ala132 (75%), Val155 (70,8%) e Gly153 (56,3%). Entende-se por frequência a porcentagem de ligantes que interagem com um resíduo específico, calculada pela divisão do número total de ligantes que interagem com esse resíduo pelo número total de ligantes analisados.

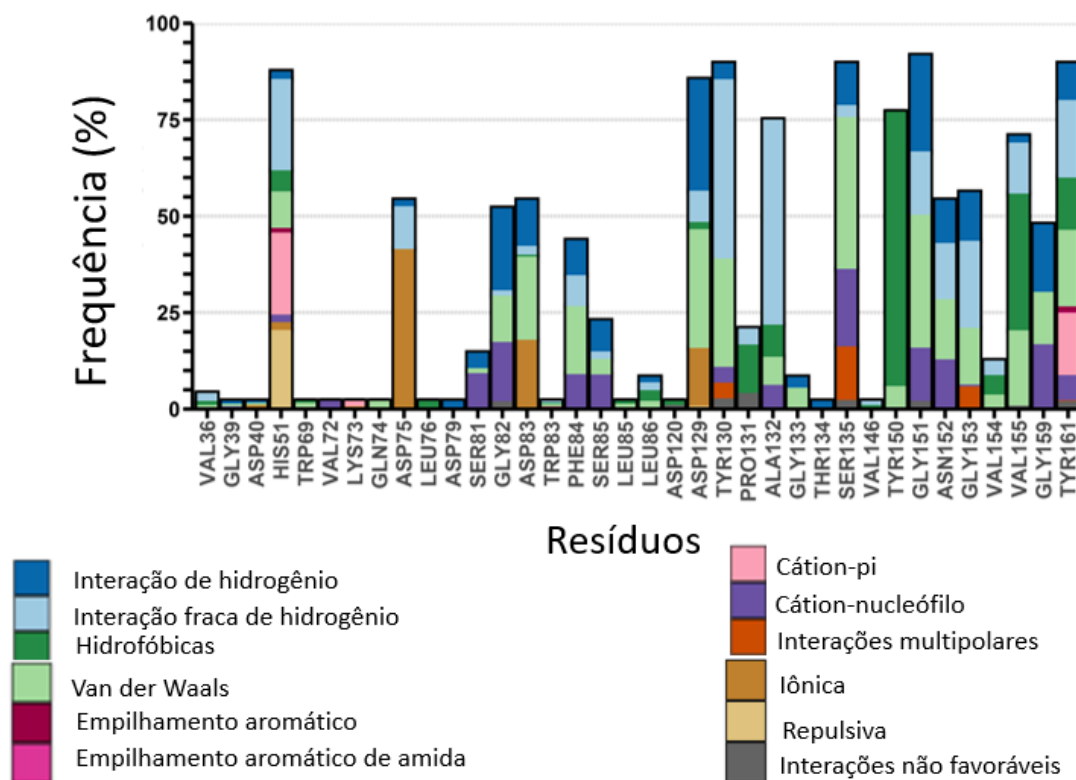


Figura 8 - Análise do perfil das interações proteína-ligante entre as 49 estruturas cristalográficas estudadas. A frequência de ligação se refere à porcentagem de ligantes cristalográficos que interagem com determinado resíduo. As cores dos gráficos indicam o tipo de interação realizada, cujo significado encontra-se na legenda.

Em relação aos tipos de interações observadas, identificamos um predomínio de interações de hidrogênio fortes, quando há interação do hidrogênio com átomos de nitrogênio ou oxigênio, bem como interações de hidrogênio fracas que ocorrem com outros heteroátomos. Além disso, as interações de van der Waals e hidrofóbicas também se mostraram amplamente distribuídas entre os ligantes analisados. Outros tipos de interação foram observadas em menor proporção. Um destaque é para o Asp75, onde se observou um predomínio de interações iônicas.

Os resíduos Asp129, Tyr130, Ala132, Tyr150 e Tyr161 compõem o subsítio S1, que é documentado na literatura como um dos principais sítios de ligação para inibidores. Além disso, essa região é conservada nos demais vírus da família, como o vírus da dengue, sendo interessante a busca de inibidores que ocupem essa cavidade (LI et al., 2018; PHOO et al., 2018). A Ser135 e His51 fazem parte da tríade catalítica. É interessante notar que, embora o Asp75 também faça parte da tríade, ele ocupa a décima primeira posição com 54,2% de interação com os ligantes. Adicionalmente, a

Val155 pertence ao subsítio S4, enquanto a Gly151 e Gly153, embora não estejam associadas a um subsítio específico, fazem parte da estrutura da proteína NS3 (LI et al., 2017; PHOO et al., 2016; ZHANG et al., 2016).

Com base na informação sobre a frequência de interação dos resíduos com os ligantes, procedemos ao alinhamento dos dez resíduos melhores classificados, utilizando a estrutura tridimensional 5LC0 como referência (**tabela 3**). Os resíduos com RMSD inferior a 1 Å estão destacado em verde, aqueles com RMSD entre 1 e 2 Å em amarelo, e acima de 2 Å em vermelho. A maioria dos resíduos, em todas as estruturas tridimensionais, apresentou RMSD inferior a 2 Å, tendo como maior exceção a estrutura 7M1V, o que era esperado devido à sua conformação super aberta.

Tabela 3 – RMSD dos resíduos do sítio ativo em relação a estrutura 5LC0

| Estrutura tridimensional/resíduo | GLY151 | TYR130 | SER135 | TYR161 | HIS51 | ASP129 | TYR150 | ALA132 | VAL155 | GLY153 |
|----------------------------------|--------|--------|--------|--------|-------|--------|--------|--------|--------|--------|
| 5LC0 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 5H41 | 0,35 | 0,59 | 0,10 | 0,16 | 0,95 | 0,50 | 1,69 | 2,07 | 1,94 | 0,23 |
| 5H6V | 0,29 | 0,36 | 0,22 | 0,20 | 0,83 | 0,64 | 0,45 | 1,77 | 0,72 | 0,31 |
| 5VLG | 0,47 | 0,45 | 0,31 | 0,66 | 1,09 | 0,39 | 1,66 | 1,86 | 1,97 | 0,35 |
| 5YOD | 0,33 | 0,44 | 1,20 | 0,31 | 0,71 | 1,26 | 0,48 | 1,85 | 2,06 | 0,18 |
| 5YOF | 0,34 | 0,42 | 0,29 | 0,23 | 0,92 | 0,45 | 0,36 | 1,27 | 2,04 | 0,44 |
| 5ZMQ | 0,32 | 0,39 | 1,71 | 0,31 | 0,65 | 0,25 | 0,39 | 1,11 | 2,23 | 0,45 |
| 5ZMS | 0,20 | 0,27 | 1,75 | 0,46 | 0,68 | 0,34 | 1,64 | 1,32 | 2,03 | 0,41 |
| 5ZOB | 0,35 | 1,77 | 0,74 | 0,48 | 0,73 | 1,08 | 0,54 | 0,39 | 0,99 | 0,38 |
| 6JPW | 0,40 | 0,41 | 0,59 | 0,59 | 0,64 | 0,93 | 1,66 | 0,36 | 1,93 | 0,36 |
| 6KK2 | 0,40 | 0,42 | 0,08 | 0,24 | 0,96 | 1,65 | 1,66 | 1,62 | 2,05 | 0,36 |
| 6KK3 | 0,38 | 0,44 | 0,33 | 0,42 | 0,92 | 0,57 | 1,66 | 1,54 | 1,84 | 0,38 |
| 6KK4 | 0,42 | 0,39 | 0,37 | 0,56 | 1,02 | 0,49 | 1,66 | 1,77 | 1,92 | 0,35 |
| 6KK5 | 0,34 | 0,40 | 0,32 | 0,38 | 1,02 | 0,72 | 1,66 | 1,57 | 1,91 | 0,34 |
| 6KK6 | 0,42 | 0,37 | 0,41 | 0,51 | 1,01 | 0,52 | 1,66 | 1,68 | 1,93 | 0,39 |
| 6KPQ | 0,31 | 0,36 | 0,55 | 0,27 | 0,83 | 0,73 | 1,66 | 1,35 | 2,00 | 0,19 |
| 6L4Z | 0,21 | 0,48 | 0,10 | 0,14 | 0,87 | 0,21 | 0,34 | 2,33 | 2,17 | 0,21 |
| 6L50 | 0,22 | 0,41 | 0,25 | 0,15 | 0,94 | 0,40 | 1,66 | 2,16 | 2,03 | 0,21 |
| 6Y3B | 0,39 | 0,35 | 0,38 | 0,55 | 1,03 | 0,51 | 1,66 | 1,66 | 1,93 | 0,37 |
| 7DOC | 0,22 | 1,80 | 0,53 | 0,44 | 0,77 | 0,63 | 1,65 | 0,96 | 1,75 | 0,31 |
| 7M1V | 3,75 | 2,22 | 1,65 | NA* | 2,06 | 8,21 | 3,67 | 7,14 | 9,39 | 3,51 |
| 7O2M | 0,43 | 0,48 | 0,38 | 0,57 | 1,04 | 0,82 | 1,65 | 1,71 | 1,99 | 0,48 |
| 7O55 | 0,40 | 0,37 | 0,44 | 0,51 | 1,04 | 0,72 | 1,64 | 1,58 | 1,77 | 0,36 |
| 7OBV | 0,42 | 0,37 | 0,53 | 0,64 | 1,12 | 0,65 | 1,66 | 1,42 | 1,99 | 0,50 |
| 7OC2 | 0,41 | 0,42 | 0,34 | 0,66 | 1,07 | 0,45 | 1,66 | 1,72 | 1,91 | 0,35 |
| 7PFQ | 0,39 | 0,35 | 0,55 | 0,55 | 1,08 | 0,56 | 1,66 | 1,40 | 1,94 | 0,46 |
| 7PFY | 0,42 | 0,41 | 0,35 | 0,47 | 1,09 | 0,69 | 1,65 | 1,63 | 0,23 | 0,32 |
| 7PFZ | 0,43 | 0,47 | 0,38 | 0,61 | 1,07 | 0,50 | 1,66 | 1,67 | 1,89 | 0,33 |

| Estrutura tridimensional/ resíduo | GLY151 | TYR130 | SER135 | TYR161 | HIS51 | ASP129 | TYR150 | ALA132 | VAL155 | GLY153 |
|--------------------------------------|--------|--------|--------|--------|-------|--------|--------|--------|--------|--------|
| 7PG1 | 0,41 | 0,47 | 0,36 | 0,46 | 1,01 | 0,78 | 1,66 | 1,66 | 1,82 | 0,35 |
| 7PGC | 0,44 | 0,46 | 0,42 | 0,66 | 1,05 | 0,45 | 1,66 | 1,70 | 1,89 | 0,34 |
| 7VLH | 0,22 | 0,42 | 0,41 | 0,16 | 0,95 | 0,59 | 1,67 | 1,37 | 1,97 | 0,20 |
| 7VLI | 0,46 | 0,31 | 0,34 | 0,31 | 1,01 | 0,77 | 1,66 | 1,29 | 1,81 | 0,38 |
| 7VXX | 0,30 | 0,41 | 0,21 | 0,23 | 0,86 | 0,19 | 0,32 | 2,20 | 2,17 | 0,33 |
| 7VXY | 0,34 | 1,73 | 0,64 | 0,26 | 0,85 | 0,35 | 0,33 | 1,29 | 2,07 | 0,28 |
| 7ZLC | 0,33 | 0,36 | 0,34 | 0,45 | 0,98 | 0,53 | 1,65 | 1,64 | 0,81 | 0,31 |
| 7LZD | 0,38 | 0,38 | 0,47 | 0,50 | 1,02 | 0,75 | 1,66 | 1,61 | 1,90 | 0,37 |
| 7ZMI | 0,34 | 0,44 | 0,20 | 0,27 | 0,99 | 0,59 | 1,67 | 1,78 | 1,83 | 0,29 |
| 7ZNO | 0,34 | 0,44 | 0,32 | 0,30 | 0,72 | 2,01 | 1,67 | 0,40 | 1,96 | 0,31 |
| 7ZPD | 0,38 | 0,40 | 0,42 | 0,59 | 1,03 | 0,55 | 1,65 | 1,64 | 1,91 | 0,36 |
| 7ZQ1 | 0,42 | 0,41 | 0,45 | 0,47 | 1,04 | 0,59 | 1,65 | 1,69 | 1,96 | 0,33 |
| 7ZQF | 0,39 | 0,37 | 0,46 | 0,39 | 1,03 | 0,60 | 1,66 | 1,65 | 1,89 | 0,36 |
| 7ZTM | 0,40 | 0,48 | 0,39 | 0,53 | 1,08 | 0,46 | 1,66 | 1,67 | 1,87 | 0,35 |
| 7ZUM | 0,43 | 0,50 | 0,38 | 0,55 | 1,06 | 0,47 | 1,67 | 1,60 | 1,89 | 0,30 |
| 7ZV4 | 0,40 | 0,43 | 0,36 | 0,65 | 1,03 | 0,38 | 1,65 | 1,60 | 1,91 | 0,33 |
| 7ZWK | 0,40 | 0,50 | 0,51 | 0,39 | 1,03 | 0,97 | 1,67 | 1,51 | 1,93 | 0,40 |
| 7ZYS | 0,40 | 0,42 | 0,45 | 0,57 | 1,01 | 0,52 | 1,65 | 1,64 | 1,93 | 0,33 |
| 8AQA | 0,43 | 0,41 | 0,47 | 0,58 | 1,03 | 0,60 | 1,65 | 1,67 | 1,90 | 0,33 |
| 8AQB | 0,38 | 0,43 | 0,45 | 0,55 | 1,07 | 0,57 | 1,65 | 1,56 | 1,94 | 0,38 |
| 8AQK | 0,37 | 0,43 | 0,42 | 0,51 | 1,04 | 0,56 | 1,65 | 1,76 | 1,94 | 0,45 |

*O resíduo não está presente na estrutura cristalográfica. Verde: RMSD < 1 Å, amarelo: 1 > RMSD < 2 Å, vermelho: RMSD > 2 Å.

Em seguida, realizamos uma análise visual, sobrepondo as estruturas tridimensionais que mostraram maior RMSD para determinado resíduo, buscando identificar as possíveis alterações conformacionais. Contudo, essa análise não nos permitiu identificar alterações significativas, pois os resíduos que apresentaram maior diferença de RMSD em relação à estrutura tridimensional 5LC0 mostraram similaridade conformacional entre eles, conforme observado na **figura 9**.

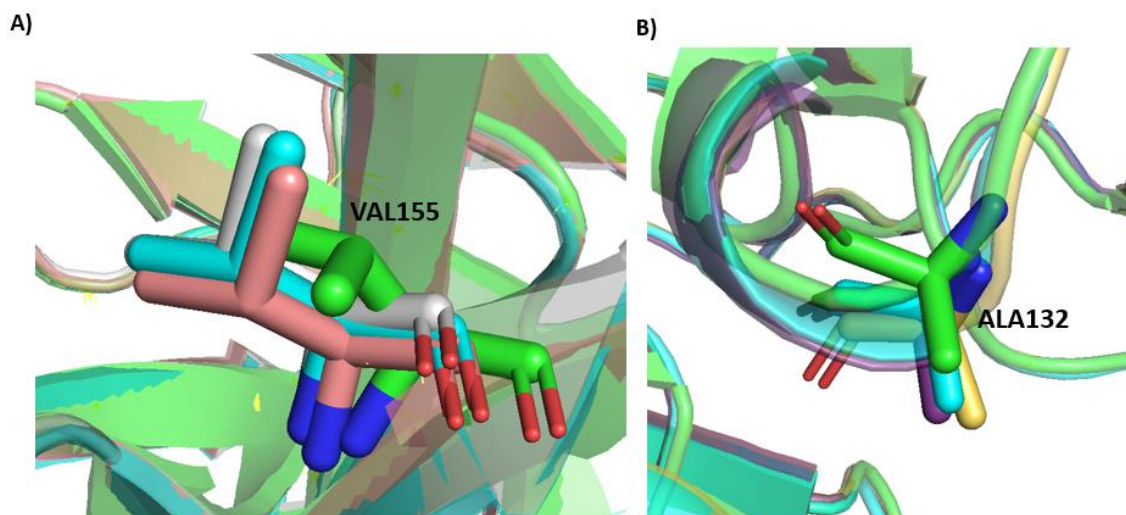


Figura 9 - Sobreposição dos resíduos VAL155 e ALA132. A) Sobreposição dos resíduos VAL155 das estruturas tridimensionais 5LC0 (verde), 5ZMQ (branco), 7VXX (azul) e 6L4Z (rosa). B) Sobreposição dos resíduos ALA132 das estruturas tridimensionais 5LC0 (verde), 5H4I (azul), 6L4Z (amarelo), 6L50 (roxo).

Devido ao resultado da análise anterior, optamos por uma nova abordagem para identificar as diferenças conformacionais dos resíduos no sítio ativo. Inicialmente, dividimos as estruturas tridimensionais conforme o tipo de ligante associado: fragmento, macrociclo, peptídeo, peptídeo-mimético e acíclico. Realizamos uma análise visual comparativa, sobrepondo as estruturas tridimensionais dentro de cada grupo individualmente. Após a pré-seleção em cada grupo, conduzimos uma segunda inspeção visual, sobrepondo essas estruturas escolhidas para selecionar uma estrutura tridimensional representativa de cada categoria inicial.

No grupo de fragmentos não foi observada diferença significativa nas conformações dos resíduos em destaque. Portanto, a escolha da estrutura para o *cross-docking* foi feita considerando a resolução global da proteína, densidade eletrônica do ligante e dados experimentais disponíveis. A estrutura 5H4I foi descartada por ter a pior densidade eletrônica do ligante (5) e pior resolução (2 Å), critérios que também se aplicam para a 6L50 que apresenta densidade igual a 9 e resolução de 1,95 Å. A estrutura 7VXX foi descartada por apresentar densidade eletrônica do ligante inferior as demais (8). Além disso, a molécula 5YOD se liga de forma covalente a serina catalítica (LI et al., 2018). Por estarmos usando uma metodologia voltada para ligantes não-covalentes esse aspecto poderia interferir nos resultados. Desta forma, a escolha final para esse grupo foi a estrutura 6L4Z.

A estrutura 6KK2, pertencente ao grupo dos macrociclos, foi selecionada devido ao valor de RMSD de 1,65 Å do resíduo Asp129 em relação à estrutura 5LC0. No entanto, após análises visuais dos resultados de docking, identificamos que essa diferença era um artefato metodológico causado pela rotação do grupo carboxílico. Essa rotação não teria impacto no *docking* devido à equivalência química dos oxigênios presentes neste grupo funcional (**figura 10**).

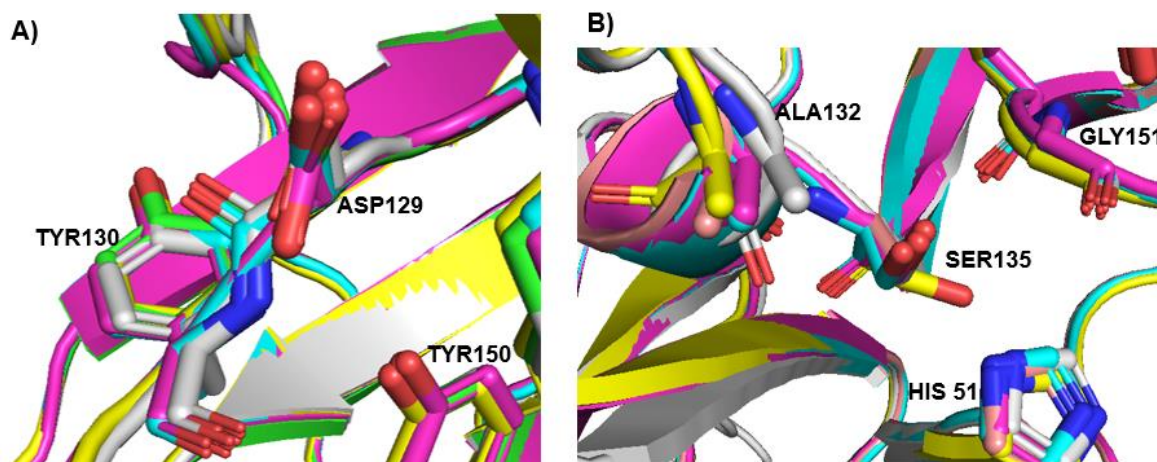


Figura 10 - Comparação da orientação dos resíduos ASP129 e SER135 nas estruturas do cross-docking. Ambas figuras apresentam as seguintes estruturas tridimensionais: 7ZLC (ciano), 6KK2 (magenta), 5ZMS (amarelo), 5YOF (verde) e 6L4Z (branco). A) Rotação do ácido carboxílico do ASP129 na estrutura 6KK2 em relação as demais. B) Rotação do grupo hidroxila da SER135 da estrutura 5ZMS.

A estrutura do 5ZMS, um peptídeomimético, apresentou uma pequena diferença conformacional na Ser135, tendo uma diferença na rotação do grupo hidroxila (**figura 10**), sendo, portanto, escolhida.

No grupo dos peptídeos, inicialmente, foram observadas diferenças na conformação do Asp129 nos compostos 5YOF e 6JPW. No entanto, ao compará-los com os demais, essa conformação se repetiu em pelo menos uma das outras estruturas. Como não foi identificada nenhuma diferença estrutural significativa, a escolha da estrutura deste grupo foi baseada em outros critérios. A estrutura 5H6V foi descartada devido a sua pior resolução (2,42 Å). As estruturas 6JPW, 7DOC, 7VXY foram descartadas porque os ligantes correspondentes não foram identificados pelo programa de densidade eletrônica e possuíam resolução igual ou pior que 1,90 Å. A estrutura escolhida foi a 5YOF por apresentar, em relação a outra restante (5GPI), melhor resolução (1,51 Å).

Por último, as estruturas no grupo dos acíclicos não apresentaram diferença significativa na conformação dos resíduos entre si. Optou-se pela estrutura 7ZLC porque o ligante da 7ZNO realiza ligação covalente com a cadeia lateral da serina

(BRAUN et al., 2023) e o *docking* realizado nesse trabalho é voltado para ligantes que realizam ligações não covalentes.

Para a execução do *cross-docking* foram selecionadas as estruturas tridimensionais depositadas sob os seguintes PDB IDs: 5YOF (ligante peptídico), 5ZMS (ligante peptídeo-mimético), 6KK2 (ligante macrociclo), 6L4Z (ligante fragmento) e 7ZLC (ligante acíclico). O conjunto de ligantes usado nessa análise compreende 25 estruturas, incluindo os ligantes das cinco estruturas tridimensionais mencionadas, bem como outros ligantes cristalográficos com densidade eletrônica superior à 6.

4.3 Avaliação da reprodução da pose cristalográfica dos ligantes

4.3.1 Resultados do *cross-docking* utilizando a conformação do ligante de menor energia

O *cross-docking* é um procedimento empregado para validação de um protocolo de *docking* quanto a capacidade de reproduzir modo de ligação de ligantes cristalográficos. Nesse processo, ligantes cujo modo de ligação já foram experimentalmente determinados são usados para realizar simulações computacionais de *docking* com estruturas tridimensionais proteicas diferentes das originais (MUKHERJEE; BALIUS; RIZZO, 2010; TEMML; KUTIL, 2021). Já o *re-docking* avalia a capacidade do protocolo de *docking* reproduzir a pose cristalográfica do ligante na estrutura tridimensional proteica de origem (TEMML; KUTIL, 2021).

As estruturas tridimensionais utilizadas para o *cross-docking* foram preparados da mesma maneira conforme descrito anteriormente (3.2.2), assim como os vinte e cinco ligantes (3.2.3). O protocolo utilizado para o *docking* foi o mesmo para os cinco grupos analisados, utilizando o programa Dock6, cuja amostragem é realizada por meio da construção incremental e a pontuação é baseada em grid, conforme descrita na equação 1. Essa análise permitiu obter resultados de *cross-docking* e *re-docking*.

A avaliação do sucesso tanto do *cross-docking* quanto do *re-docking* foi realizada através do cálculo do RMSD, utilizando a pose cristalográfica do ligante como referência. Na literatura, considera-se sucesso quando o valor de RMSD é inferior a 2 Å (CASTRO-ALVAREZ; COSTA; VILARRASA, 2017; TEMML; KUTIL, 2021), indicando que a pose prevista pelo protocolo de *docking* é similar à pose observada na estrutura cristalográfica. Os valores obtidos para cada um dos ligantes

crystalográficos, representados por “L + código PDB da estrutura tridimensional”, estão descritos na **tabela 4**. A avaliação do sucesso tanto do *cross-docking* quanto do *re-docking* foi realizada através do cálculo do RMSD, utilizando a pose cristalográfica do ligante como referência. Na literatura, considera-se sucesso quando o valor de RMSD é inferior a 2 Å (CASTRO-ALVAREZ; COSTA; VILARRASA, 2017; TEMML; KUTIL, 2021), indicando que a pose prevista pelo protocolo de *docking* é similar à pose observada na estrutura cristalográfica. Os valores obtidos para cada um dos ligantes cristalográficos, representados por “L + código PDB da estrutura tridimensional”, estão descritos na **tabela 4**.

Tabela 4 – Características dos ligantes cristalográficos e seus respectivos RMSD em relação a cada estrutura tridimensional

| Ligantes | | | Estruturas tridimensionais | | | | |
|----------|------------------|----------------------------------|----------------------------|-------|-------|-------|-------|
| Código | Tipo | Número de ligações rotacionáveis | RMSD (Å) | | | | |
| | | | 5YOF | 5ZMS | 6KK2 | 6L4Z | 7ZLC |
| L5YOF | Peptídeo | 12 | 5,11 | 7,78 | 2,18 | 8,00 | 3,30 |
| L5ZMQ | Peptídeomimético | 20 | 7,21 | 8,81 | 9,01 | 7,91 | 8,58 |
| L5ZMS | Peptídeomimético | 22 | 11,08 | 8,95 | 10,87 | 9,46 | 6,76 |
| L6JPW | Peptídeo | 28 | 13,32 | 14,79 | 14,83 | 12,91 | 14,08 |
| L6KK2 | Macrocielo | 16 | 9,20 | 8,40 | 8,41 | NA* | 9,56 |
| L6KK3 | Macrocielo | 16 | 7,19 | 11,79 | 9,09 | 9,81 | 8,07 |
| L6KK4 | Macrocielo | 16 | 9,94 | 9,65 | 8,09 | 8,98 | 9,90 |
| L6KK5 | Macrocielo | 15 | 6,92 | 8,50 | 7,65 | 10,13 | 7,30 |
| L6KK6 | Macrocielo | 15 | 7,84 | 9,43 | 7,85 | 9,96 | 7,85 |
| L6KPK | Macrocielo | 17 | 8,54 | 7,29 | 8,40 | 9,33 | 7,48 |
| L6L4Z | Fragmento | 2 | 11,19 | 12,87 | 11,64 | 13,06 | 12,42 |
| L6L50 | Fragmento | 0 | 14,45 | 3,12 | 12,40 | 13,50 | 12,47 |
| L6Y3B | Macrocielo | 15 | 8,45 | 21,00 | 6,99 | 9,80 | 8,39 |
| L7DOC | Peptídeo | 26 | 12,20 | 9,85 | 7,77 | 9,56 | 12,04 |
| L7O2M | Macrocielo | 25 | 8,34 | 8,25 | 8,03 | 6,58 | 7,78 |
| L7O55 | Macrocielo | 22 | 8,50 | 7,83 | 7,39 | 9,26 | 8,98 |
| L7OBV | Macrocielo | 24 | 8,97 | 8,62 | 9,08 | 8,39 | 8,84 |
| L7OC2 | Macrocielo | 17 | 10,49 | 8,69 | 7,55 | 10,88 | 10,34 |
| L7PFQ | Macrocielo | 25 | 7,67 | 6,15 | 6,12 | 6,62 | 7,63 |
| L7PFY | Macrocielo | 23 | 6,70 | 8,37 | 7,73 | 9,03 | 7,41 |
| L7PG1 | Macrocielo | 22 | 8,34 | 8,61 | 8,61 | 8,60 | 8,25 |

| | | | | | | | |
|-------|------------|----|------|------|------|------|------|
| L7VLH | Macrociclo | 21 | 8,48 | 7,83 | 6,79 | 9,86 | 6,64 |
| L7VXY | Peptídeo | 22 | 9,93 | 7,96 | 7,31 | 8,51 | 7,29 |
| L7ZLC | Acíclico | 19 | 8,36 | 8,56 | 8,34 | 8,94 | 8,34 |
| L7ZLD | Macrociclo | 22 | 7,57 | 9,20 | 9,18 | 9,80 | 9,27 |

*NA: O programa não conseguiu realizar o *docking* para este par alvo molecular-ligante

Nenhum dos ligantes usados obteve sucesso no *cross-docking* e *re-docking*. As falhas na reprodução da pose cristalográfica podem ser categorizadas em dois tipos distintos. O primeiro é a falha do ranqueamento, onde o *docking* teve êxito em reproduzir a pose cristalográfica, mas essa pose não foi classificada como a de melhor pontuação. O segundo tipo de falha ocorre na amostragem, indicando que o *docking* não conseguiu explorar e identificar a pose correta (ALLEN et al., 2015; MUKHERJEE; BALIUS; RIZZO, 2010).

Uma hipótese levantada para a falha do *cross-docking* está relacionada com falha da amostragem devido às classes de ligantes empregadas. Os ligantes do grupo dos fragmentos são moléculas muito pequenas com capacidade de se ligarem a diferentes regiões do sítio de ligação, o que pode justificar a diferença nos valores de RMSD entre a pose referência e a do *docking*. Considerando que apenas a pose melhor ranqueada de cada ligante foi analisada, não permitindo a análise da falha por erro de ranqueamento, é possível que outras poses tenham se ligado próximas ao sítio observado na estrutura cristalográfica e reproduzido sua pose, representando uma possível falha da pontuação do *docking* em identificar a pose mais próxima a de referência como a melhor.

Ligantes peptídeos e peptídeo-miméticos são compostos com uma alta flexibilidade estrutural, decorrente do elevado número de ligações rotacionáveis em sua configuração (CIEMNY et al., 2018; SANTOS et al., 2020). A maioria dos ligantes dessas classes, utilizada nesse trabalho, possuem mais de 15 ligações rotacionáveis, conforme evidenciado na **tabela 2**. A reprodução de poses cristalográficas de compostos com essa complexidade estrutural representa um desafio para o dock6, conforme discutido por Allen (2015). Em suas análises, observaram-se taxas significativas de falha para esse tipo de ligantes, especialmente quando comparadas a grupos de moléculas com menos de oito ligações rotacionáveis. Apesar de não identificado sucesso na reprodução do modo de ligação do ligante cristalográfico levando em consideração o corte como 2 Å, observamos que o ligante 5YOF dentro

da estrutura 6KK2 apresentou um RMSD de 2,18 Å. Como ilustrado na **figura 11**, o ligante cristalográfico (azul, B) e a pose 1 do docking (rosa, C) compartilham a maioria das interações realizadas. Portanto, mesmo não estando dentro da faixa considerada como sucesso em termos de RMSD, é possível considerá-la como uma reprodução adequada da maior parte do modo de ligação do ligante cristalográfico.

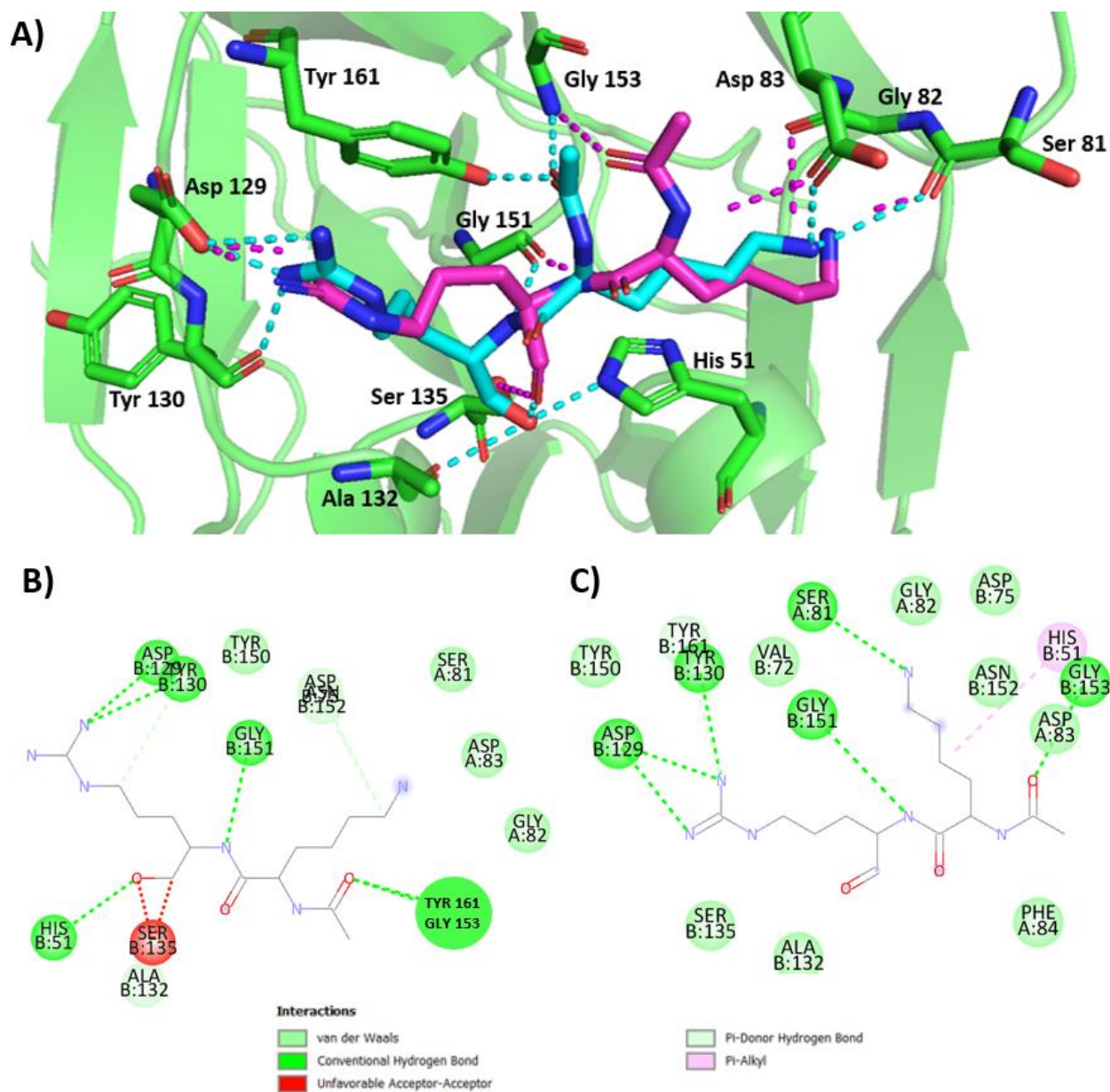


Figura 11 – Comparação do modo de ligação entre o ligante 5YOF cristalográfico da estrutura 6kk2 o ligante 5YOF correspondente à pose 1 do docking com a estrutura 6KK2. A) O ligante 5YOF cristalográfico está representado em azul, enquanto o ligante 5YOF da pose 1 do docking está em rosa. As linhas pontilhadas representam

interações polares, e sua coloração está de acordo com a estrutura que representam. B) interações do ligante 5YOF cristalográfico com a estrutura 6KK2. C) Interações da pose 1 do docking com a estrutura 6KK2.

Os ligantes macrociclos apresentam maior complexidade em termos de flexibilidade estrutural, pois, além de possuírem pelo menos quinze ligações rotacionáveis (**tabela 2**), sua natureza cíclica implica na necessidade de rotação de múltiplos diédricos para transitar entre diferentes conformações (ALLEN; DOKHOLYAN; BOWERS, 2016). Em uma revisão na literatura, encontramos um artigo que realiza uma comparação entre a capacidade de diversos programas em reproduzir poses cristalográficas de ligantes macrocíclicos (CASTRO-ALVAREZ; COSTA; VILARRASA, 2017). O Dock6.5, uma versão anterior do programa adotado nesse trabalho, foi um dos avaliados no artigo e demonstrou sucesso na reprodução de poses para quase todos os ligantes. Esse resultado nos levou a comparar a metodologia utilizada na preparação dos ligantes a qual utilizou programas distintos em relação ao nosso na obtenção das conformações tridimensionais, além de terem realizado uma preparação adicional com busca da melhor conformação para o macrociclo. Essa discrepância suscitou a reflexão sobre a possibilidade de a geração das conformações dos ligantes, realizadas apenas pelas ferramentas da OpenEye, pode influenciar nos resultados do *cross-docking*. Esse questionamento se baseia na compreensão de que, para obter sucesso na obtenção de poses em ligantes macrocíclicos, o ideal é que a conformação usada para o docking seja próxima ao do ligante cristalográfico, principalmente em relação à conformação do anel, pois a construção incremental empregada na amostragem do Dock6 considera qualquer tipo de anel como segmento rígido para formação da estrutura da molécula (ALLEN; DOKHOLYAN; BOWERS, 2016; LANG et al., 2023).

4.3.2 Avaliação da reprodução da conformação dos ligantes cristalográficos

Na preparação dos ligantes usados no *cross-docking*, feita com as ferramentas da OpenEye, optamos por utilizar apenas uma conformação tridimensional considerada pelo programa OMEGA como a de menor energia. Na preparação dos ligantes usados no *cross-docking*, feita com as ferramentas da OpenEye, optamos por utilizar apenas uma conformação tridimensional considerada pelo programa OMEGA como a de menor energia. Diante a suposição de que essa conformação poderia não representar a conformação cristalográfica, especialmente para os anéis macrociclos, realizamos uma análise de RMSD comparando a conformação cristalográfica de cada ligante em relação à conformação gerada pelo OMEGA. Para avaliar o acerto dessa reprodução,

foi utilizado as seguintes métricas: acerto na reprodução da conformação cristalográfica quando o RMSD for inferior a 2 Å (destacado em verde), acerto parcial da reprodução da conformação quando o RMSD ficar entre 2 e 3 Å (destacado em amarelo) e erro na reprodução da conformação cristalográfica quando o RMSD for superior a 3 Å (em branco). Como pode ser observado na **tabela 5** tivemos apenas quatro acertos, sendo que dois pertencem ao grupo dos fragmentos (L64Z e L6L50), por terem um baixo número de ligações rotacionáveis, era esperado que a conformação fosse reproduzida. Os outros dois ligantes pertencem a classe dos macrociclos (L6KK4 e L6KK6). Obtivemos também três acertos parciais para outros macrociclos (L7O2M, L7O55 e L7PFQ).

Outra análise realizada foi a comparação entre a conformação do ligante cristalográfico com a conformação das poses geradas pelo *docking*. Essa abordagem foi adotada devido à capacidade do algoritmo de construção incremental de modificar a conformação das regiões flexíveis do ligante durante a geração da pose. O programa divide a molécula em dois tipos de cadeias (rígidas e flexíveis). As cadeias rígidas podem ser segmentos separados por ligações não rotacionáveis ou anéis de qualquer tipo. Cadeias flexíveis são aquelas que possuem ligações rotacionáveis em sua composição. A construção do ligante na cavidade do sítio ativo é realizada fixando uma porção rígida e, em seguida, adicionando as porções flexíveis. Entre cada uma dessas adições ocorre a minimização de energia da pose (LANG et al., 2023)

Nosso objetivo era avaliar se, após o processamento, haveria melhor reprodução da conformação cristalográfica do ligante em relação a conformação de entrada do OMEGA, previamente discutida. Uma vez que nosso foco era analisar exclusivamente a conformação dos ligantes processados (estrutura) e não sua posição no sítio ativo (pose), antes de calcular o RMSD realizamos um alinhamento da conformação do ligante cristalográfico e a conformação resultante das poses do *docking*. Os dados referentes a essa análise também estão descritos na **tabela 5**. Nós avaliamos a conformação referente às primeiras dez poses geradas utilizando a estrutura tridimensional 5YOF. Para ser classificado como acerto, era necessário que pelo menos uma das dez reproduções apresentasse RMSD inferior a 2 Å. Podemos observar que após o *docking*, o número de acertos aumentou de quatro para sete. Os ligantes L5YOF (peptídeo), L7O55 e L7PFQ (macrociclo) foram incluídos no grupo dos

acertos, indicando uma melhoria na reprodução da conformação cristalográfica após o processamento.

Tabela 5 – RMSD entre a conformação dos ligantes cristalográficos e a conformação gerada pelo OMEGA antes e após processamento pelo *docking*

| Ligantes | RMSD (Å) entre ligante cristalográfico e conformações analisadas | | | | | | | | | | |
|----------|--|---|------|------|------|------|------|------|------|------|------|
| | Conformação de entrada | Classificação das poses geradas após <i>docking</i> | | | | | | | | | |
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| L5YOF | 3,58 | 2,31 | 2,16 | 2,85 | 2,88 | 2,86 | 2,32 | 2,88 | 2,74 | 2,09 | 1,99 |
| L5ZMQ | 5,20 | 4,72 | 3,50 | 2,93 | 3,05 | 4,72 | 2,90 | 3,06 | 2,95 | 2,97 | 3,08 |
| L5ZMS | 5,01 | 4,10 | 4,57 | 4,55 | 4,70 | 4,72 | 4,12 | 4,34 | 4,36 | 4,77 | 4,14 |
| L6JPW | 4,37 | 4,53 | 4,51 | 4,16 | 4,13 | 4,14 | 4,86 | 4,23 | 5,01 | 5,66 | 5,77 |
| L6KK2 | 3,03 | 2,65 | 2,64 | 2,48 | 2,80 | 2,63 | 2,76 | 2,77 | 2,94 | 2,65 | 2,80 |
| L6KK3 | 3,46 | 2,84 | 2,80 | 2,92 | 2,92 | 2,91 | 2,87 | 2,94 | 2,92 | 3,03 | 2,96 |
| L6KK4 | 1,96 | 2,12 | 2,00 | 2,20 | 2,07 | 2,13 | 2,26 | 2,22 | 2,14 | 2,20 | 2,22 |
| L6KK5 | 3,90 | 3,47 | 3,54 | 3,60 | 3,56 | 3,49 | 3,26 | 3,41 | 3,80 | 3,65 | 3,45 |
| L6KK6 | 1,58 | 2,02 | 1,79 | 1,89 | 1,96 | 1,85 | 1,93 | 1,88 | 2,08 | 1,83 | 1,84 |
| L6KPQ | 4,92 | 4,72 | 4,74 | 4,73 | 4,76 | 4,70 | 4,82 | 4,99 | 4,77 | 4,91 | 4,82 |
| L6L4Z | 0,42 | 0,55 | 0,55 | 0,57 | 0,51 | 0,54 | 0,50 | 0,39 | 0,66 | 0,29 | 0,47 |
| L6L50 | 0,05 | 0,05 | 0,05 | 0,05 | 0,05 | 0,05 | 0,05 | 0,05 | 0,05 | 0,05 | 0,05 |
| L6Y3B | 3,92 | 3,92 | 3,93 | 3,88 | 3,99 | 3,97 | 4,14 | 3,91 | 4,17 | 3,91 | 4,06 |
| L7DOC | 3,83 | 4,76 | 4,71 | 4,74 | 4,63 | 4,65 | 4,71 | 4,69 | 4,37 | 4,63 | 4,61 |
| L7O2M | 2,42 | 2,93 | 2,80 | 2,82 | 2,90 | 2,75 | 2,97 | 2,80 | 3,06 | 2,84 | 2,95 |
| L7O55 | 2,11 | 2,05 | 1,97 | 2,01 | 1,95 | 2,03 | 2,01 | 1,95 | 2,11 | 2,18 | 2,05 |
| L7OBV | 5,15 | 5,18 | 5,18 | 5,19 | 5,14 | 5,20 | 5,15 | 5,19 | 5,13 | 5,18 | 5,13 |
| L7OC2 | 3,40 | 3,43 | 3,29 | 3,35 | 3,50 | 3,42 | 3,39 | 3,25 | 3,50 | 3,28 | 3,29 |
| L7PFQ | 2,29 | 2,44 | 1,87 | 2,20 | 1,99 | 2,49 | 2,01 | 1,99 | 1,95 | 2,05 | 2,44 |
| L7PFY | 4,78 | 4,60 | 4,66 | 4,60 | 4,69 | 4,64 | 4,63 | 4,38 | 4,51 | 4,77 | 4,47 |
| L7PG1 | 3,06 | 2,76 | 2,80 | 2,82 | 2,75 | 2,87 | 2,82 | 2,84 | 2,79 | 2,78 | 2,78 |
| L7VLH | 4,81 | 4,65 | 4,83 | 4,83 | 4,90 | 4,88 | 4,85 | 4,89 | 4,80 | 4,78 | 4,78 |
| L7VXY | 4,20 | 3,71 | 4,19 | 4,04 | 4,04 | 3,98 | 3,98 | 4,35 | 4,30 | 4,30 | 4,27 |
| L7ZLC | 4,83 | 3,97 | 3,96 | 3,45 | 4,07 | 3,95 | 3,37 | 3,37 | 4,07 | 4,03 | 4,18 |
| L7ZLD | 6,17 | 5,67 | 5,71 | 5,71 | 5,79 | 5,73 | 5,65 | 5,78 | 5,71 | 5,73 | 5,70 |

4.3.3 Resultados do *cross-docking* utilizando conformações dos ligantes com menor RMSD

Devido à baixa reprodução das conformações dos ligantes cristalográficos, tanto antes quanto após o processamento do *docking*, decidimos investigar se o uso de conformações próximas às cristalográficas (RMSD < 2 Å) como arquivos de entrada para o *docking* poderia diminuir a falha relacionada a amostragem. Para explorar essa hipótese, geramos novas conformações através do OMEGA, gerando um máximo de 400 conformações para cada ligante. Em seguida, alinhemos essas novas

conformações com os ligantes cristalográficos e calculamos seus RMSD. As conformações com melhor RMSD (**tabela 6**) foram selecionadas e foram submetidas a um novo *docking*.

Com exceção de quatro ligantes (L5ZMS, L6JPW, L7DOC, L7VXY), todas as novas conformações usadas como entrada tiveram acerto ou acerto parciais (apenas a L5ZMQ) na reprodução da conformação cristalográfica. Além disso, observamos melhora na taxa de acertos das conformações após o *docking*, com dez ligantes tendo sido classificados como acerto. Entretanto, é importante ressaltar que ter uma boa reprodução da conformação cristalográfica como arquivo de entrada não garantiu a reprodução desse acerto após o *docking*, como evidenciado em alguns casos, como o L7O2M, L7OC2 e 7PFQ, cujas conformações iniciais possuem RMSD inferior a 2 Å, mas todas as conformações derivadas das poses do *docking* foram classificadas como erro.

Tabela 6 - RMSD entre a conformação dos ligantes cristalográficos e a melhor conformação gerada pelo OMEGA antes e após processamento pelo *docking*

| Ligantes | RMSD (Å) entre ligante cristalográfico e conformações analisadas | | | | | | | | | | |
|----------|--|---|------|------|-------|------|------|------|------|------|------|
| | Conformação de entrada de menor RMSD | Conformação pós processamento do <i>docking</i> | | | | | | | | | |
| | | Número das poses cuja conformação foi analisada | | | | | | | | | |
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| L5YOF | 1,49 | 3,69 | 2,47 | 2,51 | 2,55 | 2,19 | 2,30 | 2,70 | 2,39 | 2,18 | 2,56 |
| L5ZMQ | 2,06 | 4,30 | 4,30 | 2,62 | 2,62 | 2,64 | 2,64 | 2,31 | 2,54 | 2,54 | 2,49 |
| L5ZMS | 3,47 | 4,55 | 4,42 | 4,55 | 4,35 | 4,48 | 4,37 | 4,38 | 4,59 | 4,61 | 4,60 |
| L6JPW | 3,74 | 4,49 | 4,49 | 4,48 | 4,37 | 4,37 | 4,36 | 4,45 | 4,53 | 4,35 | 4,54 |
| L6KK2 | 1,35 | 1,80 | 1,38 | 1,42 | 1,47 | 1,73 | 1,84 | 1,85 | 1,42 | 1,72 | 1,67 |
| L6KK3 | 1,86 | 4,26 | 4,27 | 4,29 | 4,229 | 4,27 | 4,34 | 4,38 | 4,37 | 4,29 | 4,22 |
| L6KK4 | 1,43 | 1,75 | 1,74 | 1,72 | 1,54 | 1,71 | 1,58 | 1,56 | 1,54 | 1,56 | 1,58 |
| L6KK5 | 1,47 | 3,68 | 3,73 | 3,64 | 3,77 | 3,74 | 3,79 | 3,74 | 3,68 | 3,79 | 3,70 |
| L6KK6 | 1,51 | 1,92 | 1,84 | 2,05 | 1,91 | 1,79 | 1,91 | 1,8 | 1,97 | 2,03 | 2,06 |
| L6KPQ | 1,80 | 1,98 | 2,05 | 2,04 | 2,00 | 2,03 | 1,94 | 2,00 | 1,96 | 1,97 | 2,13 |
| L6L4Z | 0,21 | 0,55 | 0,50 | 0,53 | 0,49 | 0,62 | 0,50 | 0,18 | 0,57 | 0,75 | 0,26 |
| L6L50 | 0,05 | 0,05 | 0,05 | 0,05 | 0,05 | 0,05 | 0,05 | 0,05 | 0,05 | 0,05 | 0,05 |
| L6Y3B | 1,72 | 1,70 | 1,66 | 1,66 | 1,61 | 1,68 | 1,72 | 1,78 | 1,74 | 1,81 | 1,89 |
| L7DOC | 3,05 | 3,42 | 3,42 | 3,35 | 3,35 | 4,61 | 4,63 | 4,66 | 3,51 | 3,45 | 3,06 |
| L7O2M | 1,84 | 4,63 | 4,65 | 4,65 | 4,43 | 4,61 | 5,52 | 4,67 | 4,50 | 4,78 | 4,50 |
| L7O55 | 1,70 | 1,93 | 1,88 | 1,94 | 1,85 | 1,95 | 1,86 | 1,89 | 1,91 | 1,88 | 1,81 |
| L7OBV | 1,73 | 2,15 | 2,15 | 2,16 | 2,10 | 2,15 | 2,09 | 2,14 | 2,23 | 2,24 | 2,29 |
| L7OC2 | 1,31 | 4,09 | 4,10 | 4,11 | 4,14 | 4,18 | 4,22 | 4,38 | 4,19 | 4,33 | 4,37 |

| | | | | | | | | | | | |
|-------|------|------|------|------|------|------|------|------|------|------|------|
| L7PFQ | 1,66 | 4,49 | 4,45 | 4,48 | 4,78 | 4,47 | 4,78 | 4,75 | 4,46 | 4,82 | 4,76 |
| L7PFY | 1,54 | 1,69 | 1,70 | 1,52 | 1,69 | 1,66 | 1,67 | 1,73 | 1,53 | 1,73 | 1,70 |
| L7PG1 | 1,62 | 4,71 | 4,98 | 4,64 | 4,76 | 4,93 | 5,01 | 4,94 | 4,93 | 5,19 | 4,88 |
| L7VLH | 1,56 | 1,82 | 1,78 | 1,88 | 1,93 | 2,01 | 1,80 | 1,98 | 1,8 | 1,75 | 1,88 |
| L7VXY | 4,33 | 3,89 | 4,06 | 3,59 | 3,90 | 3,98 | 4,04 | 4,05 | 4,05 | 3,84 | 3,71 |
| L7ZLD | 1,74 | 3,29 | 3,21 | 3,23 | 3,30 | 3,22 | 3,25 | 3,25 | 3,29 | 3,31 | 3,27 |

Na etapa final dessa análise, realizamos o cálculo de RMSD para verificar se o uso da conformação próxima a do ligante cristalográfico como arquivo de entrada impactou nos resultados em relação à reprodução da pose cristalográfica, objetivo central do *cross-docking*. Utilizamos o resultado do *docking* realizado com os ligantes com a melhor conformação geradas pelo OMEGA. Nesse ponto, a avaliação não se restringe apenas à estrutura dos ligantes, e sim considera sua orientação dentro do alvo molecular (pose). Para realizar essa avaliação, adotamos os seguintes critérios: sucesso quando o RMSD foi inferior a 2 Å (destacado em verde) e sucesso parcial quando o RMSD ficou entre 2 e 3 Å (destacado em amarelo) (RAMÍREZ; CABALLERO, 2018; THOMAS; PARRILL; BAKER, 2022) e falha quando o RMSD foi superior a 3 (branco). Observamos que, mesmo empregando conformações similares à do ligante cristalográfico, não conseguimos obter sucesso no *cross-docking* (tabela 7). Contudo, em contraste com o resultado do primeiro *cross-docking* (tabela 4) que não apresentou nenhum tipo de sucesso, obtivemos três sucessos parciais (L6KK4, L7O55, 7PFY) nesta abordagem. Os demais ligantes continuaram apresentando falha na amostragem. Vale destacar que o L7O55, apesar de ter obtido um sucesso parcial na identificação da pose, apresentou falha no ranqueamento pois a pose relacionada ao sucesso parcial não foi a melhor classificada.

Tabela 7 - RMSD entre a conformação dos ligantes cristalográficos e o resultado do *docking* realizado com a melhor conformação do OMEGA

| Ligantes | RMSD (Å) entre os ligantes cristalográficos e as poses resultantes do <i>docking</i> com 5YOF | | | | | | | | | |
|----------|---|--------|--------|--------|--------|--------|--------|--------|--------|---------|
| | Pose 1 | Pose 2 | Pose 3 | Pose 4 | Pose 5 | Pose 6 | Pose 7 | Pose 8 | Pose 9 | Pose 10 |
| L5YOF | 6,7 | 7,29 | 7,28 | 7,33 | 6,51 | 6,95 | 7,36 | 7,05 | 6,58 | 12,34 |
| L5ZMQ | 8,07 | 8,07 | 7,55 | 7,55 | 7,64 | 7,64 | 7,44 | 7,61 | 7,61 | 7,52 |
| L5ZMS | 11,07 | 10,93 | 11,06 | 10,82 | 10,88 | 10,81 | 10,95 | 11,06 | 11,06 | 10,97 |
| L6JPW | 12,81 | 12,8 | 12,81 | 12,93 | 12,93 | 12,97 | 12,77 | 12,85 | 12,96 | 12,85 |
| L6KK2 | 7,15 | 7,1 | 7,17 | 7,02 | 10,61 | 7,36 | 7,22 | 7,06 | 10,88 | 10,86 |
| L6KK3 | 9,25 | 9,25 | 9,29 | 9,08 | 9,25 | 9,27 | 9,02 | 9,02 | 9,35 | 8,68 |
| L6KK4 | 2,98 | 10,05 | 10,22 | 9,65 | 9,75 | 10 | 9,93 | 9,69 | 9,6 | 2,75 |

| | | | | | | | | | | |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| L6KK5 | 6,5 | 6,48 | 6,37 | 6,50 | 6,68 | 6,69 | 6,44 | 6,52 | 6,63 | 6,49 |
| L6KK6 | 7,11 | 7,24 | 7,2 | 7,31 | 7,48 | 6,97 | 7,57 | 6,92 | 7,1 | 7,03 |
| L6KPQ | 7,09 | 10,07 | 7,05 | 10,07 | 9,85 | 10,86 | 10,02 | 10,64 | 10,13 | 10,18 |
| L6L4Z | 11,17 | 11,11 | 11,04 | 11,05 | 11,09 | 11,07 | 11,36 | 11,24 | 11,02 | 12,83 |
| L6L50 | 14,45 | 13,58 | 13,25 | 15,5 | 13,58 | 13,8 | 16 | 14,44 | 14,76 | 10,77 |
| L6Y3B | 8,48 | 8,75 | 8,81 | 8,85 | 8,74 | 8,53 | 9,3 | 8,81 | 8,73 | 8,68 |
| L7DOC | 10,32 | 10,32 | 10,36 | 10,36 | 12,63 | 12,64 | 12,61 | 10,33 | 10,35 | 11,42 |
| L7O2M | 7,72 | 7,66 | 7,65 | 9,47 | 10,69 | 10,6 | 10,67 | 10,63 | 10,8 | 10,61 |
| L7O55 | 3,05 | 2,51 | 2,81 | 7,65 | 3,06 | 2,92 | 7,61 | 7,59 | 3 | 2,78 |
| L7OBV | 7,76 | 7,75 | 7,76 | 7,78 | 7,79 | 7,78 | 7,76 | 7,73 | 7,76 | 7,80 |
| L7OC2 | 6,98 | 7,04 | 6,99 | 9,45 | 9,49 | 9,91 | 10,17 | 9,58 | 10,1 | 10,33 |
| L7PFQ | 11,68 | 11,74 | 11,56 | 8,74 | 11,77 | 8,87 | 8,7 | 11,48 | 8,8 | 8,73 |
| L7PFY | 2,22 | 2,35 | 2,15 | 2,25 | 2,41 | 2,34 | 2,37 | 2,3 | 2,47 | 2,28 |
| L7PG1 | 5,4 | 11,77 | 6,28 | 6,51 | 11,65 | 11,62 | 7,43 | 11,69 | 10,57 | 11,66 |
| L7VLH | 6,98 | 6,89 | 6,97 | 6,98 | 6,78 | 6,98 | 6,78 | 6,93 | 6,8 | 6,72 |
| L7VXY | 7,3 | 7,31 | 7,41 | 7,32 | 7,32 | 7,34 | 4,28 | 7,28 | 7,25 | 7,31 |
| L7ZLD | 7,5 | 7,48 | 7,44 | 7,49 | 7,54 | 7,44 | 7,43 | 7,78 | 7,54 | 7,45 |

Com base nessas análises, podemos concluir, que a conformação utilizada para executar o *docking* pode influenciar no resultado final, diminuindo falhas na amostragem, como o observado para o ligante L7PFY (**figura 12**) em que a inserção de uma conformação mais próxima à cristalográfica resultou em poses e interações similares ao observado no ligante cristalográfico. No entanto, apenas usar uma conformação próxima à cristalográfica não solucionou o sucesso na reprodução das poses, como visto na **figura 13** referente ao ligante 5YOF, em que, apesar de haver uma melhora na conformação inicial utilizada, o resultado do *docking* foi pior do que o obtido no primeiro *cross-docking*. A modificação de outros parâmetros no *docking*, tais como redução do número de átomos presentes na âncora utilizada e o aumento na geração de posições da âncora e nas conformações dos ligantes dentro do sítio, também podem influenciar em uma melhora no resultado do *cross-docking*.

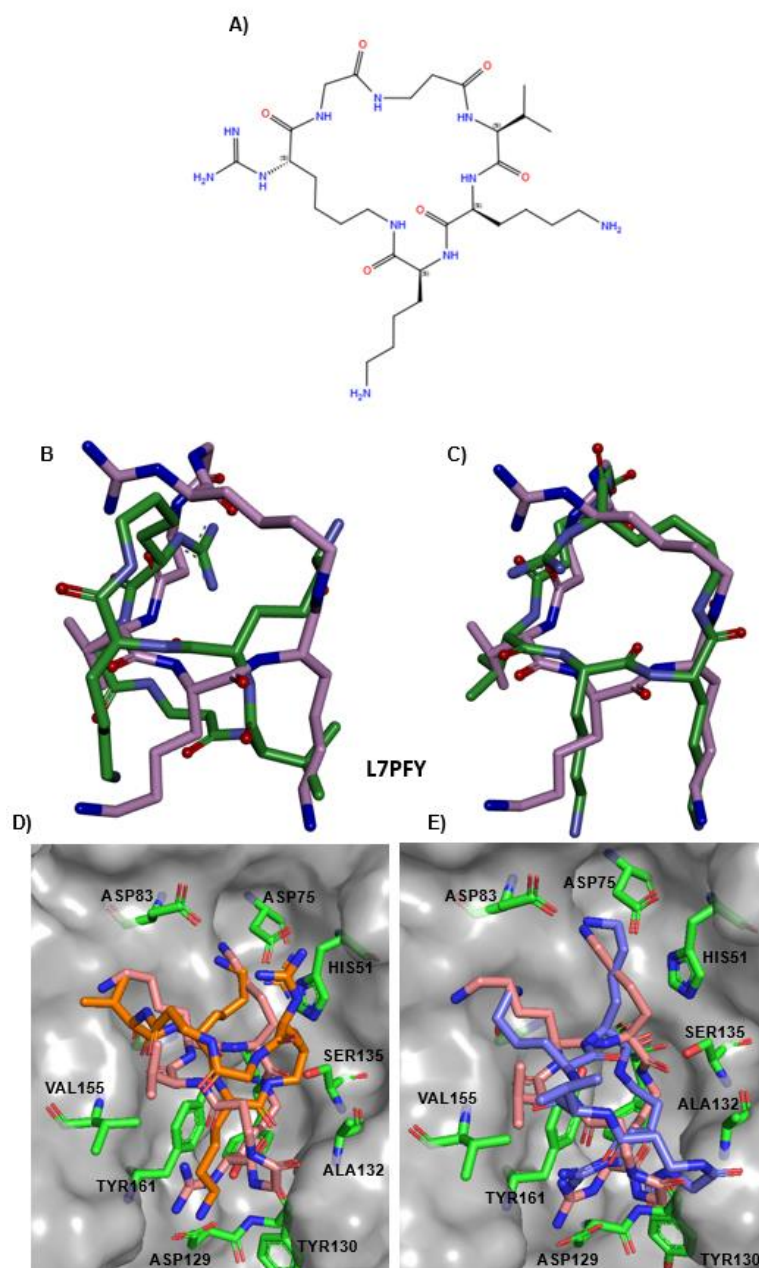


Figura 12 - Comparação das conformações e poses do ligante L7PFY. A) estrutura 2D do ligante L7PFY. B) Conformação cristalográfica do ligante 7PFY em cinza, com a conformação de menor energia do OMEGA em verde (RMSD: 4,78 Å). C) Conformação cristalográfica do ligante 7PFY em cinza, com a conformação de menor RMSD em verde (RMSD: 1,54 Å). D) Pose cristalográfica em rosa e a pose 1 *docking* antes da análise conformacional do ligante L7PFY em laranja (6,70 Å). E) Pose cristalográfica em rosa e a pose 1 do *docking* após a análise conformacional do ligante L7PFY em roxo (RMSD: 2,22 Å).

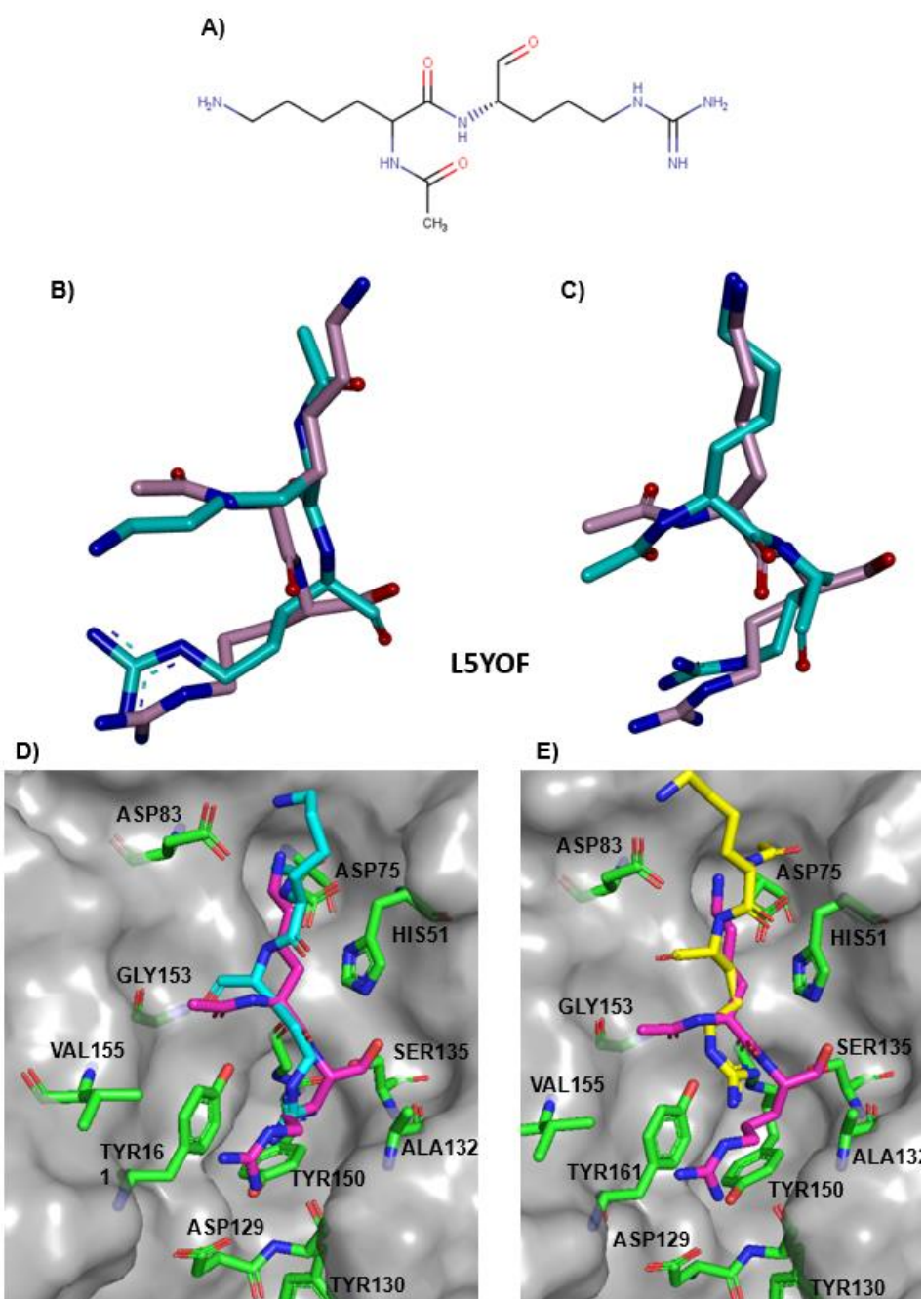


Figura 13 – Comparação das conformações e poses do ligante 5YOF. A) Estrutura 2D do ligante L5YOF. B) Conformação cristalográfica do ligante 5YOF em rosa, com a conformação de menor energia do OMEGA em azul (RMSD: 3,58 Å) C) Conformação cristalográfica do ligante 5YOF em rosa, com a conformação de menor energia do OMEGA em azul (RMSD: 1,49 Å). D) Pose cristalográfica em rosa e a pose 1 do *docking* antes da análise conformacional do ligante 5YOF em azul (RMSD: 5,11 Å) E) Pose cristalográfica em rosa e a pose 1 do *docking* após a análise conformacional do ligante 5YOF em amarelo (RMSD: 6,70 Å).

4.3.4 Análise de enriquecimento

Outra análise realizada para avaliar o desempenho do protocolo de *docking* é sua performance na triagem virtual (BROZELL et al., 2012). Isso envolve a análise do enriquecimento, que compara a habilidade do protocolo em priorizar a classificação de ligantes experimentalmente ativos em detrimento de ligantes que são presumidamente (decoys) ou comprovadamente inativos. Uma métrica utilizada para medir o enriquecimento é a área sob a curva de uma curva ROC. Essa representação gráfica permite analisar visualmente a sensibilidade versus a especificidade do protocolo, proporcionando uma visão do enriquecimento global do protocolo (ALLEN et al., 2015; TRIBALLEAU et al., 2005).

Para a construção das curvas ROC, compilamos um conjunto de moléculas ativas, incluindo os ligantes usados no *cross-docking*, bem como outros comprovadamente ativos que foram identificados na literatura. Além disso, utilizamos dois grupos de ligantes provenientes de duas fontes diferentes: decoys gerados pelo DUD-E (MYSINGER et al., 2012) e um conjunto de ligantes inativos descrito no estudo de triagem em larga escala realizado por Abrans et al. (2020).

Os ligantes que compuseram o grupo dos ativos foram retirados de dez artigos relacionados com as estruturas cristalográficas (BRAUN et al., 2020; HAMMERSCHMIDT et al., 2023; HUBER et al., 2022; LI et al., 2017, 2018; PATIL et al., 2021; PHOO et al., 2018; QUEK et al., 2020 ; XIONG et al., 2022), totalizando 73 ligantes. A fim de melhor caracterização desse grupo foram realizadas duas análises: o cálculo do coeficiente de Tanimoto para cada par de compostos do conjunto e a clusterização dos ligantes com base em suas estruturas.

Ao analisar o coeficiente de Tanimoto (**figura 14**) observamos a distinção de dois grupos principais: um grupo maior relacionado a moléculas macrocíclicas e um grupo menor composto por moléculas acíclicas. Além disso, ao realizar a clusterização (**figura 14**), os ligantes foram agrupados em três clusters distintos, devido à divisão dos ligantes macrociclos em dois subgrupos. O cluster 0 (37% dos ligantes) é composto por ligantes macrociclos que possuem um anel aromático em sua cadeia principal. O cluster 1 (20% dos ligantes) compreende os ligantes acíclicos, caracterizados por cadeias moleculares abertas. Já o cluster 2 (43% dos ligantes) é

composto por ligantes macrociclos sem a presença de anéis aromáticos em sua estrutura principal.

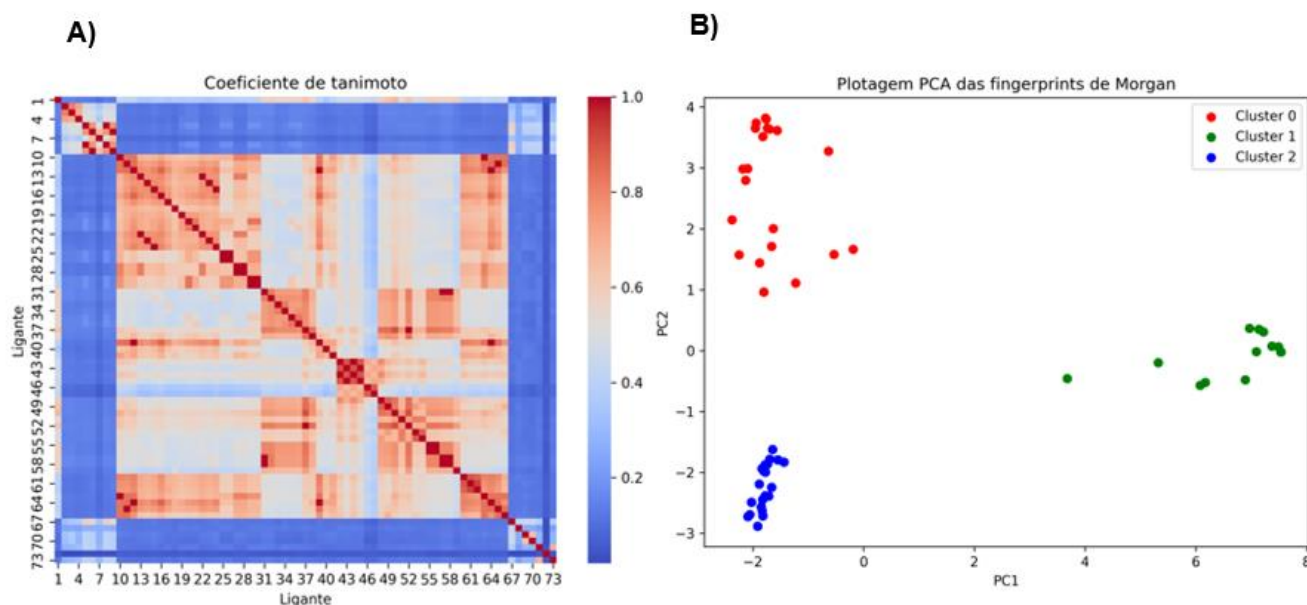


Figura 14 – Clusterização dos ligantes ativos usados na construção das curvas ROC. A) Heatmap do coeficiente de Tanimoto. Dois grupos distintos são observados, um maior central e outro menor na região superior esquerda do mapa. A variação da similaridade entre os ligantes é representada por tons de azul (0) a vermelho (1). B) Clusterização dos ligantes utilizando fingerprints de Morgan. Três grupos distintos de ligantes são identificados, onde o cluster 0 (vermelho) é composto por ligantes acíclicos, e os clusters 1 e 2 (verde e azul, respectivamente) são compostos por ligantes macrociclos.

O conjunto de decoys foi gerado através do DUD-E (MYSINGER et al., 2012). Esse programa utiliza as informações físico-químicas dos ligantes fornecidos para fazer uma busca no banco de dados ZINC20 por estruturas com características semelhantes, ao mesmo tempo que aplica filtros relacionados à estrutura topológica para que as estruturas encontradas sejam estruturalmente dissimilares em relação as ativas (MYSINGER et al., 2012). Vale ressaltar que a versão atualizada dessa plataforma, DUDE-z (STEIN et al., 2021), não pode ser utilizada devido a uma restrição quanto à massa molar dos ligantes. Diversos ligantes ativos utilizados em nosso estudo possuem massa molecular que ultrapassa o limite estabelecido pela plataforma (600 g/mol), o que inviabilizou seu uso.

A primeira curva construída utilizou os decoys gerados pelo DUD-E, após serem filtrados para evitar duplicidades, resultando em 307 moléculas distintas. No entanto os resultados obtidos não foram satisfatórios, como pode ser visto na **figura 15**. A curva cruzou a linha de aleatoriedade, e o valor da AUC foi baixo (0,487).

Nas triagens virtuais, além do valor de AUC, também se utiliza a métrica de enriquecimento em várias porcentagens. Essa métrica visa avaliar o número de ligantes ativos recuperados em diferentes faixas da base triada, como 0,1%, 1% e 2%, comparando este valor com o que seria esperado pela ordenação aleatória dos compostos (BROZELL et al., 2012). Contudo, devido ao número reduzido de ligantes empregados na construção dessa curva (307 decoys e 73 ativos) o uso dessas faixas seria inviável, pois o número de ligantes pertencentes a esses intervalos seria pequeno. Portanto para análise da curva ROC, optou-se por considerar os primeiros 10% da curva, onde foram recuperados apenas cinco ligantes, correspondendo a um fator de enriquecimento (EF) de apenas 0,68. Apesar de considerar uma porcentagem maior do que as faixas mencionadas anteriormente, esse resultado é considerado insatisfatório pois representa uma baixa taxa de recuperação dos ligantes de interesse, abaixo da curva de um resultado esperado aleatoriamente.

É importante ressaltar que o desejado na construção de uma curva empregando decoys é que tenhamos entre 20 a 50 decoys para cada ligante ativo a fim de garantir uma maior diversidade no banco de dados (MYSINGER et al., 2012; STEIN et al., 2021). Entretanto, nesse caso, não foi possível obter um número suficiente de decoys, resultando em uma proporção de aproximadamente 4 decoys para cada ligante ativo. Essa limitação pode ter impactado negativamente na curva ROC, uma vez que o banco de dados triados foi relativamente pequeno.

O conjunto de ligantes inativos foi obtido a partir do artigo de Abrans et al. (2020) que conduziu uma triagem de alto rendimento visando a identificação de inibidores da ZIKV NS2B-NS3pro realizando análise bioquímica de milhares de compostos. Portanto foi possível recuperar um grande número de ligantes inativos experimentalmente testados para serem utilizado na construção da curva ROC. Para esta segunda curva, a proporção de decoys para ligantes foi de 50 para 1, o que é considerado satisfatória para esse tipo de análise. No total, a base de dados considerada continha 3674 compostos (3600 inativos e 73 ativos). Como pode ser observado na **Figura 15**, o protocolo utilizado apresentou boa performance na recuperação de verdadeiros positivos, com valor da AUC de 0,962. Além disso, o fator de enriquecimento (FE) foi avaliado em três faixas distintas: 10%, onde o FE foi de 9,5; 5% com FE de 17,5 e 1% da base triada, com FE de 18,9. Esses resultados contrastam com o que foi observado na curva ROC dos decoys gerados pelo DUD-E.

Além disso, trata-se de um resultado inesperado considerando-se a baixa performance do protocolo de *docking* nas análises de *cross-docking* e *re-docking*.

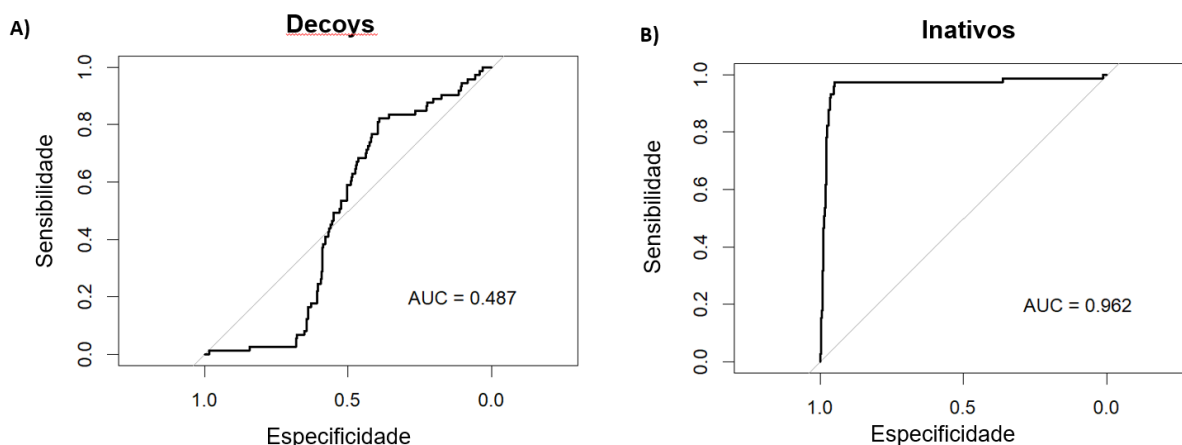


Figura 15 - Resultado das análises de enriquecimento. A) Curva ROC para o conjunto de dados dos decoys (AUC = 0,487). B) Curva ROC para o conjunto de dados dos ligantes inativos, (AUC = 0,962).

Um possível viés que pode explicar esse resultado está relacionado à natureza do sítio ativo da NS2B-NS3pro ZIKV, que é predominantemente carregado negativamente (ZHANG et al., 2016). A hipótese subjacente é que os compostos com uma carga total positiva mais alta seriam melhor classificados em relação aqueles com cargas neutras e negativas. Para investigar essa possibilidade, dividimos cada grupo de ligantes de acordo com sua faixa de carga total, conforme mostrado na **figura 16**. A maioria dos ligantes ativos possui carga igual a 3, enquanto os ligantes inativos tendem a ter cargas predominantemente 0 e 1. Os decoys, embora mais uniformemente distribuídos, também apresentam representantes na faixa de carga inferior a 3. Porém, eles também estão distribuídos em faixas de carga mais elevadas, como 7 e 8, que contém apenas decoys. Além disso, nas faixas de cargas 4 e 5, o número de decoys supera o de ligantes ativos. Esse desbalanceamento de cargas na geração de decoys, em relação aos ligantes ativos, pelo DUD-E foi destacado como uma correção realizada na criação do DUDE-Z (STEIN et al., 2021). Além das características do sítio, a função de pontuação do Dock6, conforme descrito na equação 1, leva em consideração apenas as energias de van der Waals e eletrostática, não incorporando o custo de dessolvatação do ligante. Isso implica que os ligantes carregados não irão sofrer esse tipo de penalidade, podendo ter melhor classificação nesse sistema. Portanto, os resultados obtidos nas curvas ROC se mostram coerentes, uma vez que no caso dos decoys, eles possuem maior carga e

no caso da curva dos inativos, os ligantes ativos são os mais positivamente carregados.

Além das características do sítio, a função de pontuação do Dock6, conforme descrito na equação 1, leva em consideração apenas as energias de van der Waals e eletrostática, não incorporando o custo de dessolvatação do sistema. Isso implica que os ligantes carregados não irão sofrer esse tipo de penalidade, podendo ter melhor classificação devido às interações eletrostáticas favoráveis. Portanto, os resultados obtidos nas curvas ROC se mostram coerentes, uma vez que no caso dos decoys computacionais, eles possuem maior carga e no caso da curva dos inativos, os ligantes ativos são os mais positivamente carregados.

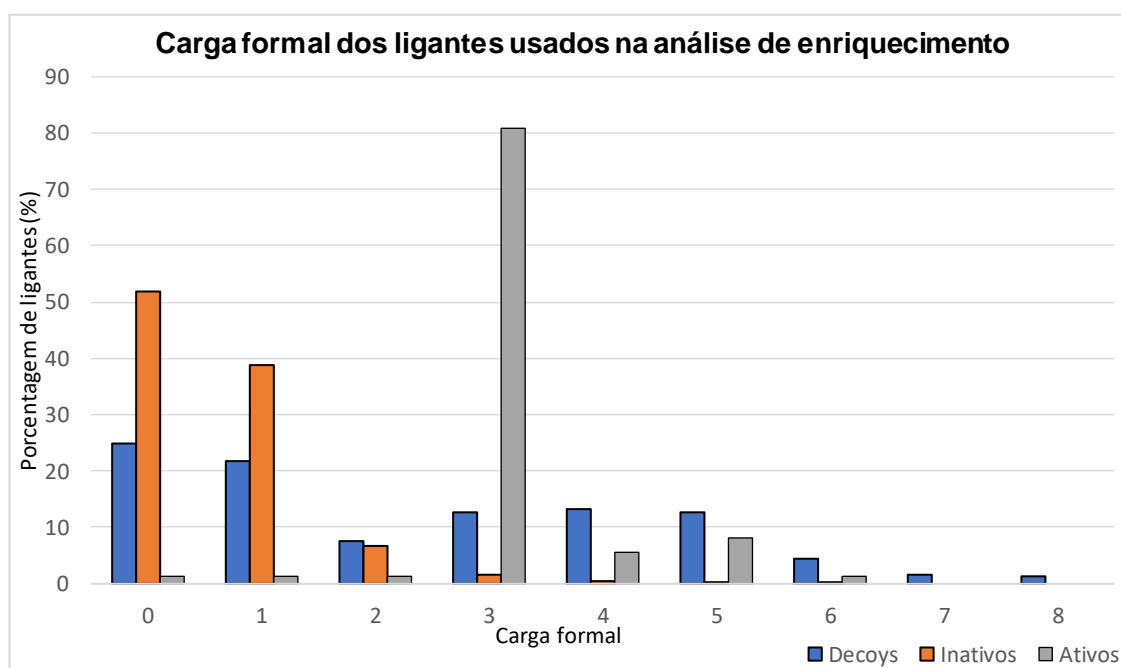


Figura 16 – Distribuição de cargas nos diferentes grupos de ligantes. Os ligantes inativos (barras cinzas) exibem distribuição da carga formal variando de 0 a 3, com maior concentração nas cargas 0 e 1. Os decoys (representados por barras azuis) estão distribuídos em todas as cargas formais, com concentração predominante nas cargas 0 e 1, sendo também os únicos a apresentarem carga 7 e 8. Os ligantes ativos (representados por barras laranjas) mostram-se distribuídos entre 0 e 6, com a maioria concentrada na carga formal 3.

A validação retrospectiva do protocolo desenvolvido revelou um desempenho insatisfatório na reprodução da pose cristalográfica e identificação de verdadeiros positivos em detrimento de falsos positivos. No entanto, é importante observar que diversos cuidados tomados durante a definição do protocolo estão de acordo com o

recomendado para o estabelecimento do mesmo. A caixa do grid criada (**figura 17**) abrange todo o sítio ativo, bem como os resíduos preditos como os mais frequentes para interações entre o alvo molecular e ligante. Os ligantes foram protonados utilizando a ferramenta Fixpka que gera o estado de protonação mais abundante para o pH 7,4. Quanto ao alvo molecular, sua preparação foi realizada no Chimera, a partir *do dock prep*, abordagem recomendada pelos desenvolvedores do Dock6. Durante essa preparação, foi considerada a protonação dos resíduos ionizáveis no sítio, mantendo a His com o hidrogênio no carbono delta (HID). Essa configuração já foi descrita em estudos de dinâmica como mais estável (SANTOS; CAFFARENA; FERREIRA, 2022). A protonação dos ligantes foi realizada em pH 8,5, o qual foi selecionado com base nas condições experimentais e é considerado ótimo para a atividade da enzima. Os parâmetros usados para o *docking*, para amostragem das poses, descritos na metodologia (3.2.4), são compatíveis com o descrito na literatura para o uso do programa (ALLEN et al., 2015). Além disso, os compostos usados na validação, como já mencionado, são desafiadores devido à sua estrutura flexível, frequentemente com mais de quinze ligações rotacionáveis. Essa complexidade estrutural não é refletida nos compostos da biblioteca utilizada na triagem virtual, que será discutida no próximo tópico, apresentando compostos menores e menos flexíveis. Por fim, vale ressaltar que a seleção de compostos para os testes bioquímicos não se baseia apenas na classificação gerada pelo programa, mas também em uma inspeção visual, onde são consideradas características como complementariedade e interações do ligante com o alvo molecular. Portanto, mesmo com os resultados da validação, o protocolo desenvolvido foi utilizado para a realização da triagem virtual.

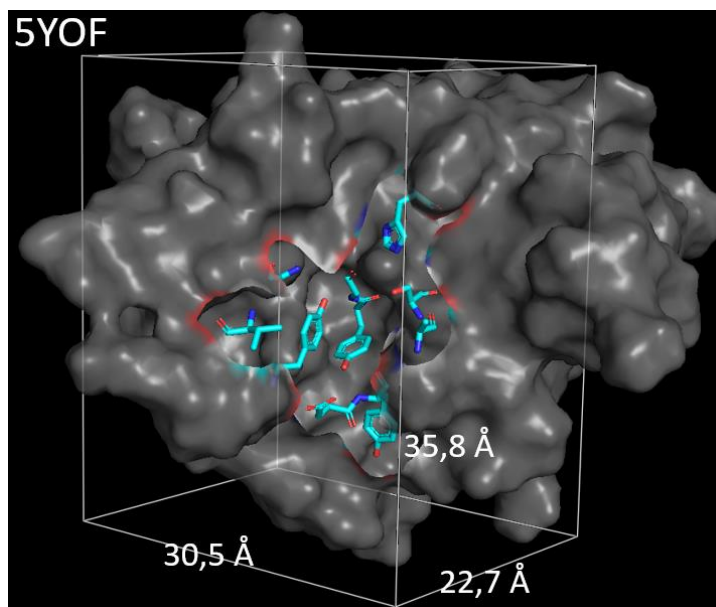


Figura 17 - Caixa usada para o docking com a estrutura tridimensional 5YOF. Representação da caixa desenvolvida para a triagem virtual. Os sticks em azul destacam alguns resíduos pertencentes ao sítio ativo (Dimensões: X: 22,7 Å , Y: 30,5 Å Z: 35,8 Å)

4.4 Triagem virtual

A triagem virtual foi conduzida utilizando a biblioteca BraCoLi que contém um total de 1076 compostos com diversidade estrutural. Seguindo o protocolo desenvolvido neste trabalho e utilizando a estrutura tridimensional 5YOF, 99,35% dos ligantes foram docados. Para realizar uma análise mais detalhada dos resultados, os 150 compostos melhores ranqueados foram submetidos a uma inspeção visual. A primeira etapa consistiu em uma análise de encaixe na superfície da ZIKV NS2B-NS3pro utilizando o Chimera. Durante essa etapa, os compostos que não apresentavam uma ocupação satisfatória do sítio ativo foram descartados. Na próxima etapa, ainda utilizando o Chimera, foram destacados os dez aminoácidos relevantes indicados pelo LUNA, bem como o Asp75, para verificar interações intermoleculares e possíveis *clashes*. Compostos que não apresentavam pelo menos uma ligação de hidrogênio, e/ou não realizavam interações com o subsítio S1 foram descartados durante essa etapa. Por fim, os compostos remanescentes foram submetidos ao Discovery Studio, permitindo uma análise mais detalhada e clara das interações diversas no sítio ativo.

Como resultado, 16 moléculas foram selecionadas para os ensaios enzimáticos. Para simplificar a explicação das escolhas, os compostos foram agrupados com base em suas características químicas. No grupo 1, temos os compostos BR10186 e BR10265, ambos contendo uma glicose ligada a um grupo tiazol. Como observado na **figura**

18, o carboidrato realiza interações de hidrogênio com a Ser81 e Asp83, na região do subsítio S2, além de um empilhamento aromático com Tyr161 no subsítio S1.

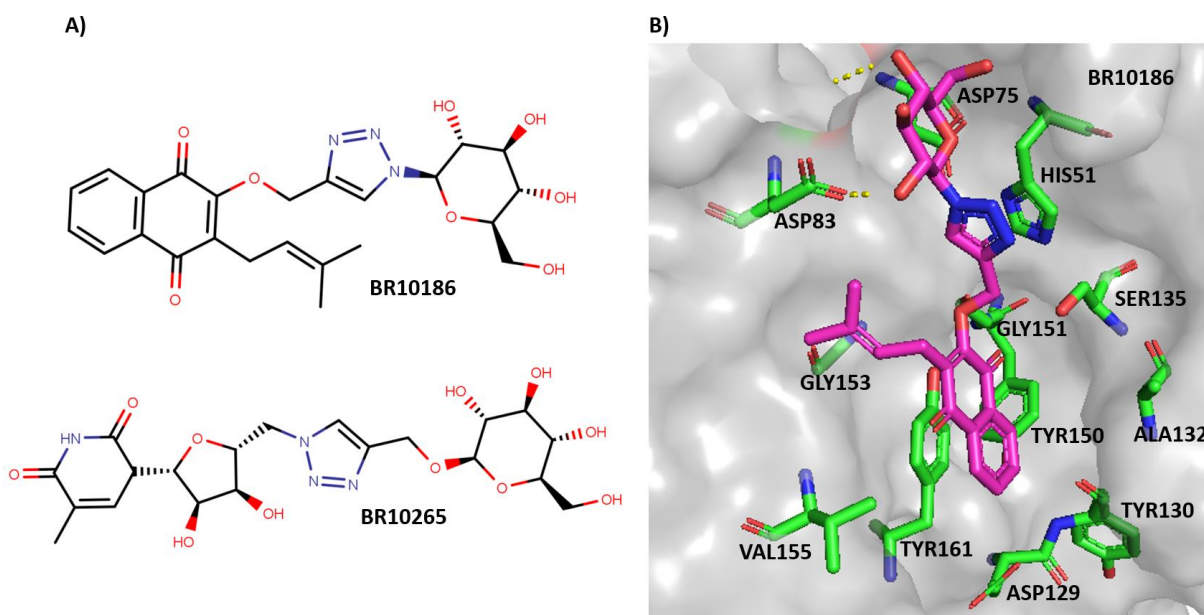


Figura 18 – Estrutura química dos compostos do grupo 1 e modo de ligação predito no sítio ativo da NS2B/NS3pro ZIKV, com a estrutura tridimensional 5YOF. A) estrutura química dos ligantes BR10186 e BR10265. B) Resultado do *docking* da molécula BR10186 na estrutura de PDB 5YOF. Ligações de hidrogênio estão destacadas em amarelo.

O grupo 2 inclui os compostos BR10153, BR10668, BR10713 e BR10402, caracterizados pela presença de uma glicose ligada a um anel aromático. Um representante desse grupo, mostrado na **figura 19**, exibe ligações de hidrogênio com

as Tyr130 e 161 do subsítio S1, além de interações pi com a His51 da tríade catalítica.

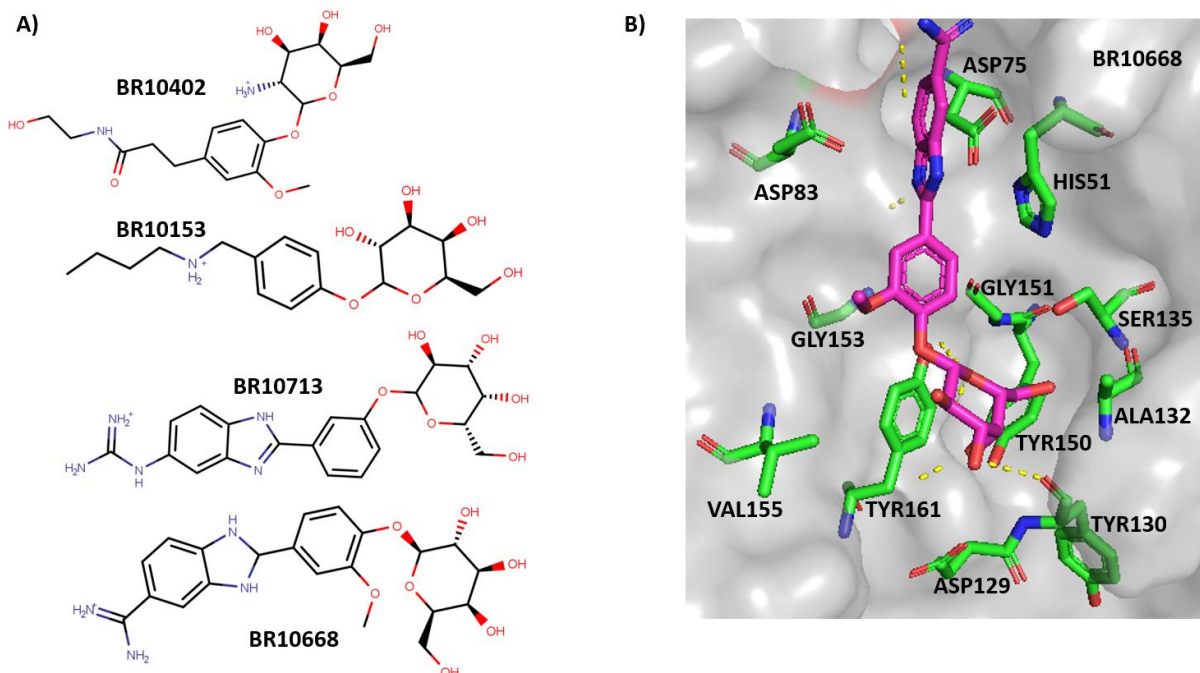


Figura 19 – Estrutura química dos compostos do grupo 2 e modo de ligação predito no sítio ativo da NS2B/NS3pro ZIKV e modo de ligação com a estrutura tridimensional 5YOF. A) estrutura bidimensional dos ligantes BR10402, BR10153, BR10713 e BR10668. B) Resultado do *docking* da molécula BR10668 na estrutura de PDB 5YOF. Ligações de hidrogênio estão destacadas em amarelo.

O grupo 3 consiste nos compostos BR10617, BR10620, BR10636. Diferentemente dos demais, esses compostos não ocupam o sítio S1, mas foram selecionados devido às suas interações com outros subsítios e uma boa ocupação geral do sítio ativo. Na **figura 20**, observamos interações de empilhamento aromático com His51 e ligações de hidrogênio com Asp83.

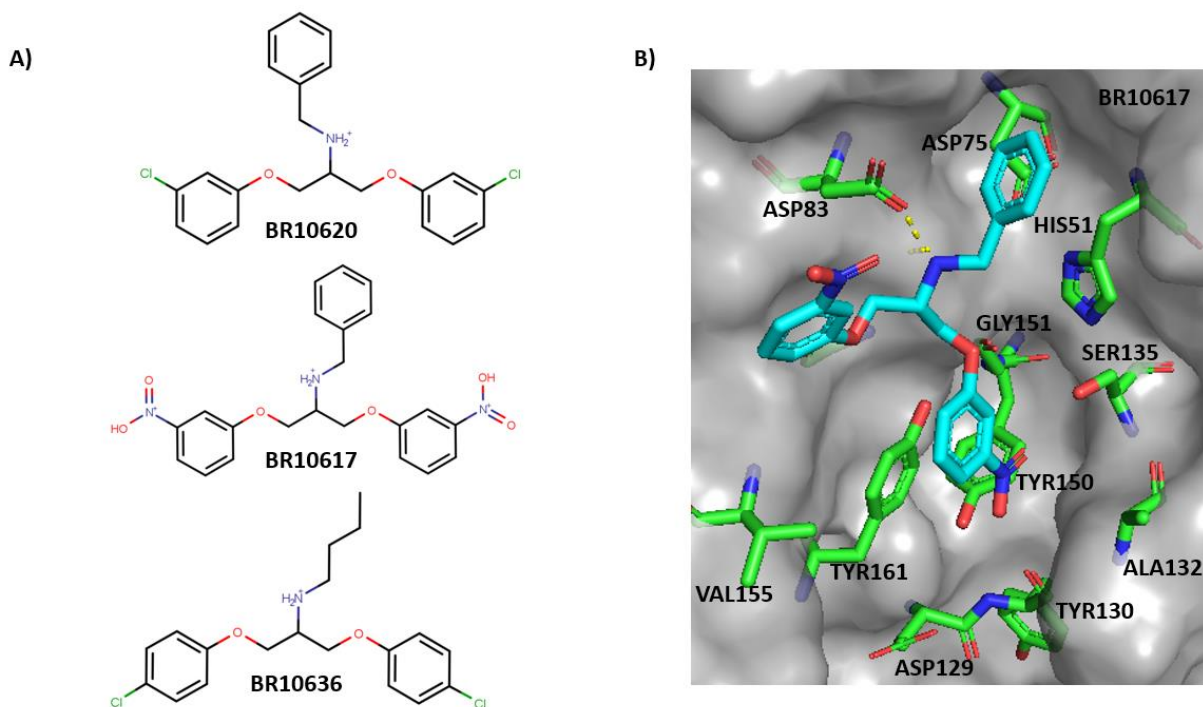


Figura 20 – Estrutura química dos compostos do grupo 3 e modo de ligação predito no sítio ativo da NS2B/NS3pro ZIKV e modo de ligação com a estrutura tridimensional 5YOF A) estrutura bidimensional dos ligantes BR10620, BR10617 e BR10636. B) Resultado do *docking* da molécula BR10617 na estrutura de PDB 5YOF. Ligações de hidrogênio estão destacadas em amarelo.

O grupo 4 é composto por BR20427, BR20442, BR20441, BR20426, caracterizados por conter um anel morfolino ligado a uma cadeia linear de três carbonos seguido de uma amina secundária. Tomando o BR20442 como exemplo (**figura 21**), vemos uma boa ocupação do subsítio S1, com interações com a Tyr161, também há ligações de hidrogênio com Gly151 e Ans152.

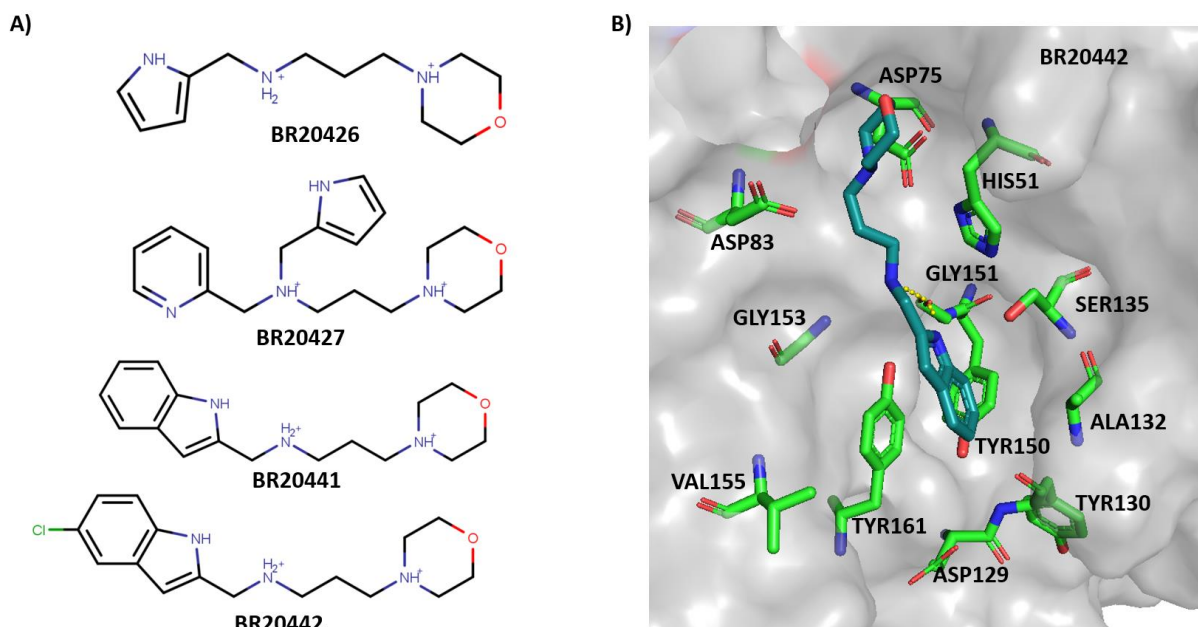


Figura 21 – Estrutura química dos compostos do grupo e modo de ligação predito no sítio ativo da NS2B/NS3pro ZIKV e modo de ligação com a estrutura tridimensional 5YOF. A) estrutura bidimensional dos ligantes BR20426, BR20427, BR20441 e BR20442. B) Resultado do *docking* da molécula BR20442 na estrutura de PDB 5YOF. Ligações de hidrogênio estão destacadas em amarelo.

O grupo 5 engloba compostos sem uma estrutura comum, todos destacados na **figura 22**. O BR20151, ocupa o subsítio S1 e realiza interações de empilhamento aromático com a Tyr161 e His51, ligações de hidrogênio com Tyr130, Asp83 e Gly153 além de interações alquil. O BR10718 interage por meio de ligações de hidrogênio com Gly82, Ans152, Asp83, além de interações alquil com Ala132, Tyr161 e Val72. Por fim, o BR10151, no sítio S1, estabelece ligações de hidrogênio com Asp129 e Tyr130, e empilhamento aromático com a Tyr161.

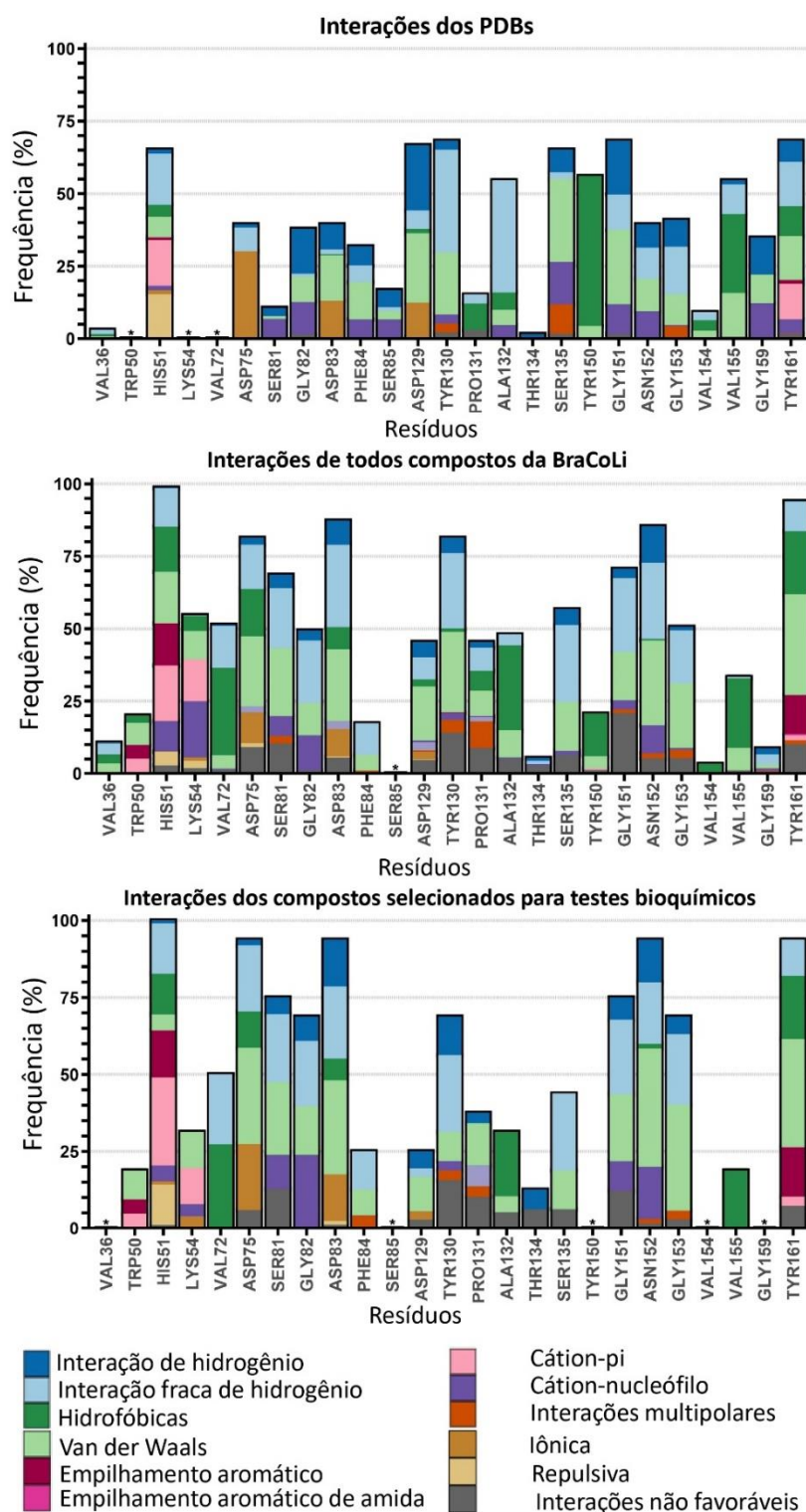


Figura 23 – Gráfico de interações versus resíduos do sítio ativo da NS2B-NS3pro. Cada gráfico representa um conjunto distinto de ligantes. “PDBs” refere-se aos 49 ligantes cristalográficos recuperados. “BraCoLi” refere-se aos 1076 ligantes após o *docking* molecular. “Compostos selecionados” são os 16 ligantes selecionados a partir dos ligantes da BraCoLi. O símbolo “*” indica a ausência de interações com o resíduo em cada um dos grupos.

Na comparação entre os resultados da BraCoLi e os ligantes cristalográficos (**figura 24**), observamos um enriquecimento na frequência de ligação de alguns resíduos na

BraCoLi, como His51, Asp75, Asp83, Asn152, Tyr161 e Pro131. Houve também uma mudança no perfil de interações, indicando uma maior diversidade na BraCoLi em comparação com os ligantes cristalográficos. Por outro lado, os resíduos Asp129, Tyr150 e Val155 apresentaram uma redução em sua frequência de interação, sendo que no caso do Asp129 houve modificação no tipo de interação, com o surgimento de interações ânion-eletrofilicas na biblioteca.

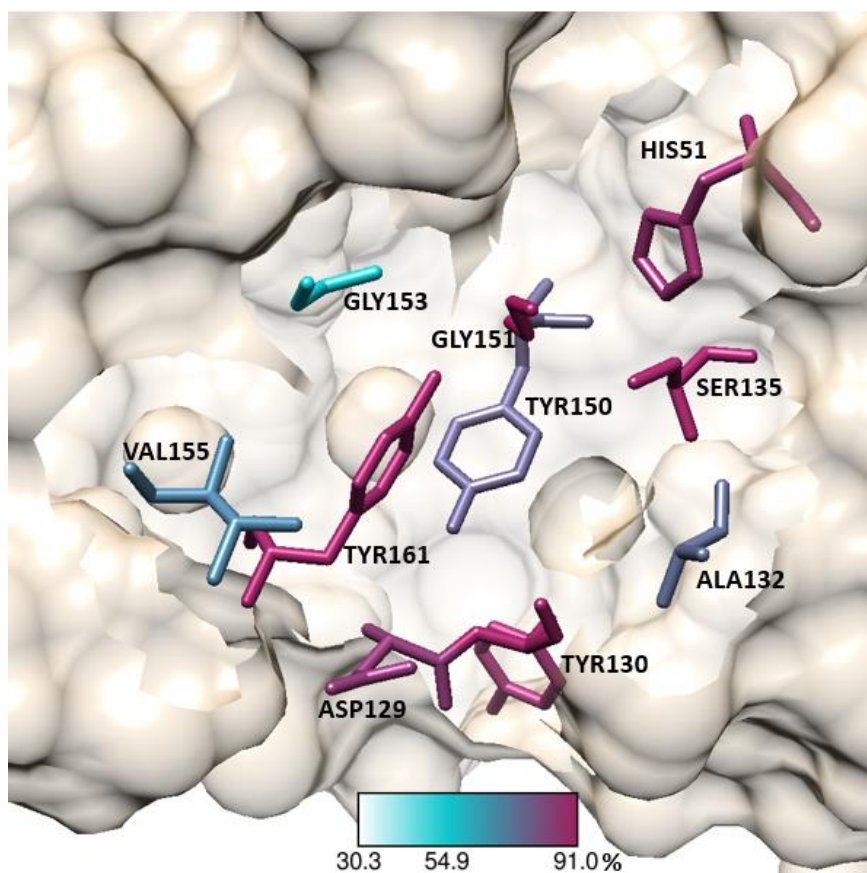


Figura 24 – Resíduos dos PDBs com maior frequência de interação com o ligante. Os resíduos estão coloridos de acordo com sua frequência de interação com os ligantes cristalográficos. Quanto mais próximo do azul menor a frequência e quanto mais próximo ao rosa maior a frequência encontrada.

O aumento da frequência de interação do Asp75, Asp83 e Asn152 está relacionado com o aumento da ocupação do sítio S2, considerada uma região crítica para o reconhecimento do substrato e atividade catalítica (CUI et al., 2020). Portanto, é interessante que ligantes tenham afinidade por esse subsítio. Ao analisarmos os resíduos que compõe o subsítio S1, notamos que houve tanto diminuição (Tyr150, Asp129) quanto aumento (Tyr161) na frequência das interações, demonstrando uma mudança no perfil de ligação com priorização de resíduos localizados próximo ao

subsítio S2. Também houve valorização do subsítio S1' devido ao aumento na frequência de interação com a histidina catalítica.

Ao analisar as interações de ligantes da BraCoLi, observamos três resíduos que interagem que não estavam presentes nos PDBs: Trp50, Lys54 e Val72. Os três estão localizados nos subsítios S1' e S2, próximos à histidina e ao aspartato catalíticos, corroborando com a possibilidade de maior interação nesses dois subsítios. Em contrapartida, nos ligantes cristalográficos há presença de interações com a Ser85, localizada na região S4, que não está presente na BraCoLi. A ausência dessa interação pode estar relacionada ao menor tamanho dos ligantes da biblioteca, quando comparado aos cristalográficos.

Comparamos também os três gráficos, para entender o perfil dos ligantes escolhidos para os testes enzimáticos. O perfil observado nos ligantes selecionados assemelha-se ao da biblioteca completa, no entanto, ao compará-las ainda é possível distinguir algumas diferenças em relação as frequências de interações de alguns resíduos. Essas pequenas alterações são relatadas a seguir. Houve aumento na frequência de ligação do Asp83, His51 e Asp75, juntamente com uma diminuição da frequência dos resíduos Asp129 e Val155, mantendo-se os tipos de interações observadas. Em relação a tríade catalítica, o aspartato e a histidina apresentaram aumento da frequência de interação. No entanto, a Ser135 não seguiu essa tendência; sua interação nas moléculas escolhidas é aproximadamente de 50%, enquanto nos ligantes cristalográficos era próximo de 75%. Houve uma ênfase na interação com a Gly82, localizada na região do subsítio S2, com um aumento de cerca de 25% em relação à biblioteca completa.

Em suma, podemos observar que os dezesseis ligantes selecionados para realização dos testes bioquímicos apresentam um perfil diferente dos ligantes cristalográficos. A seleção está concentrada principalmente no subsítio S2, mantendo interações com alguns resíduos do subsítio S1. Além disso, observou-se um aumento na interação com os resíduos catalíticos HIS51 e ASP75.

4.5 Caracterização dos potenciais inibidores de NSB2-NS3 ZIKV

Os compostos selecionados na triagem virtual foram fornecidos pelos professores Renata Barbosa de Oliveira e Ricardo Alves José, ambos da Faculdade de Farmácia da UFMG. Nem todos estavam disponíveis no momento em que os experimentos

foram realizados, portanto foram submetidos aos ensaios de caracterização dez compostos, sendo dois deles análogos do BR10153 cujo código é BR10155 e do BR10668 cujo código é BR10696 (**figura 25**).

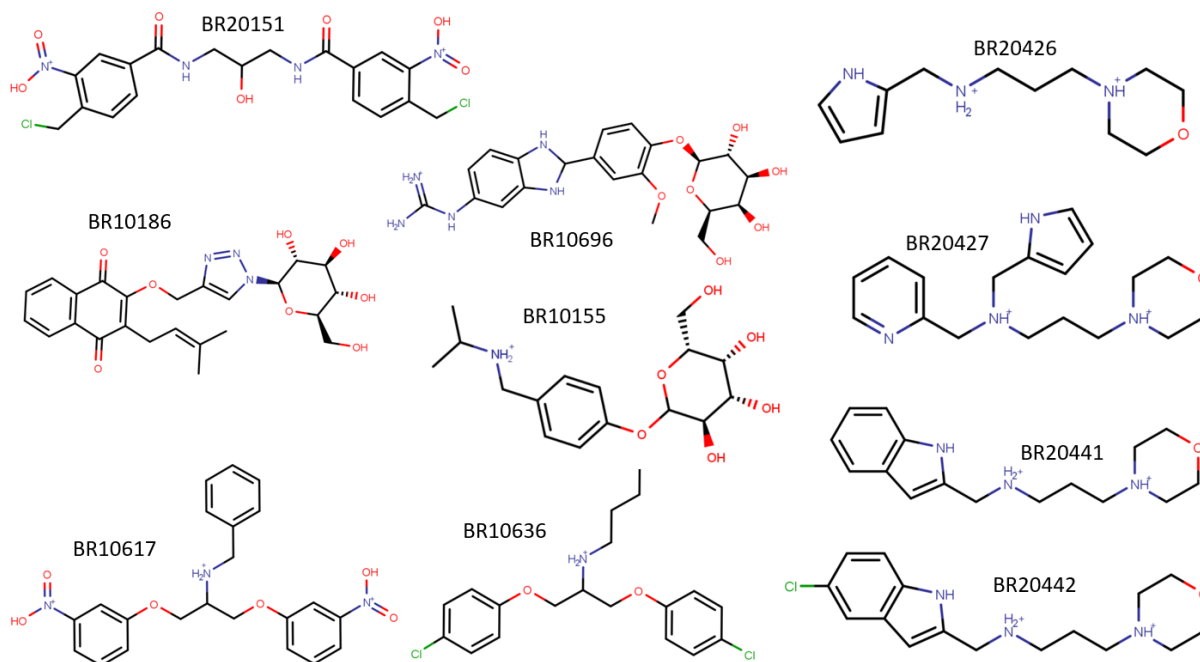


Figura 25 – Estruturas dos compostos submetidos aos ensaios enzimáticos

Os compostos BR10186 e BR20427 não foram completamente solúveis em DMSO durante a preparação da solução estoque com concentração de 10 mM, não sendo possível prosseguir com os ensaios para ambos. Os demais compostos não apresentaram problemas, tanto na diluição em DMSO quanto nos testes de solubilidade nas condições do ensaio. Diante disso, os oito compostos foram avaliados quanto à atividade inibitória na concentração de 100 μ M na presença de 0,2 nM da enzima e 44 μ M do substrato fluorescente Bz-Nle-Lys-Lys-Arg-AMC. Realizamos a triagem com incubação de dez minutos com a enzima. A inibição máxima encontrada foi de 31% (**tabela 8**).

Tabela 8 – Atividade inibitória experimentalmente determinada dos compostos selecionados na triagem virtual a 100 μ M

| Composto | Atividade inibitória com incubação a 100 μ M (%)* |
|----------|---|
| BR10155 | 4 \pm 6 |
| BR10617 | 26 \pm 6 |
| BR10636 | 14 \pm 5 |

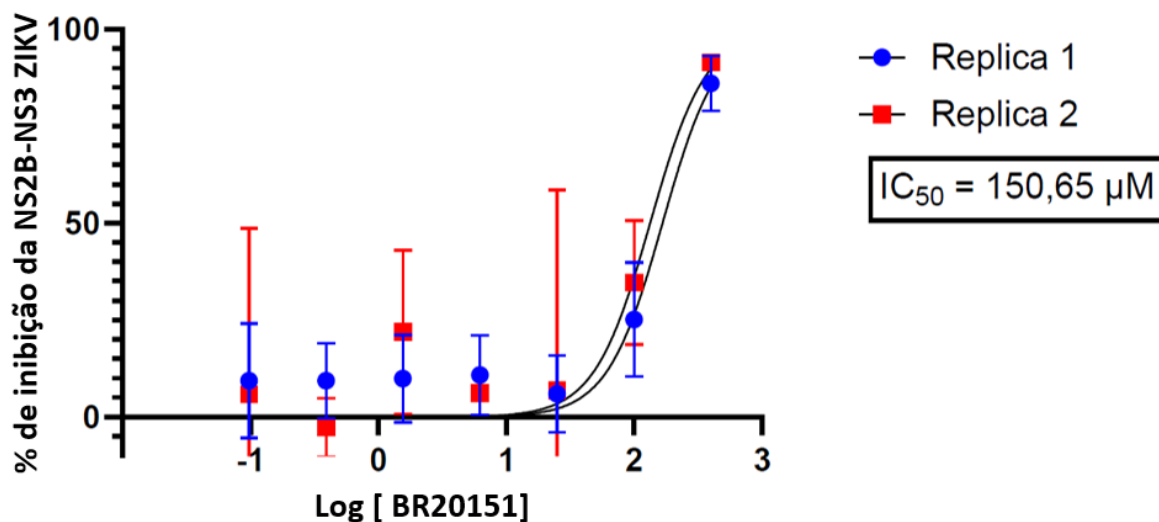
| | |
|---------|---------|
| BR10696 | 20 ± 5 |
| BR20151 | 31 ± 14 |
| BR20426 | 8 ± 7 |
| BR20441 | 21 ± 7 |
| BR20442 | 23 ± 10 |

*Os experimentos foram realizados em triplicatas e dois experimentos independentes. O valor da tabela representa a média dos experimentos e seus respectivos desvios padrões.

Os experimentos foram realizados em triplicatas e dois experimentos independentes. O valor da tabela representa a média dos experimentos e seus respectivos desvios padrões.

Apesar da baixa inibição inicial, optamos por prosseguir com a caracterização do composto BR20151. Primeiro avaliamos se a inibição da enzima era tempo dependente. Para isso, conduzimos o ensaio de inibição a 100 μM sem incubação prévia com a enzima, resultando em uma inibição de $37 \pm 18 \%$. Dado que essa inibição é comparável com àquela nas condições de incubação ($31 \pm 14\%$), podemos inferir que a inibição não é tempo-dependente. Devido à sua solubilidade em concentrações mais elevadas (400 μM), realizamos ensaios de IC_{50} com pré-incubação de 10 minutos com a enzima. Avaliamos a atividade da enzima em sete concentrações do composto, variando de 400 μM a 97 nM, em duplicata e em dois experimentos independentes. O composto apresentou IC_{50} média de 150,65 μM (**figura 26**) e coeficiente de Hill próximo a 2. O coeficiente de Hill é a medida do grau de cooperatividade entre ligantes ao interagirem com o alvo. Um coeficiente igual a 1 indica uma interação não cooperativa, enquanto valores maiores que 1 indicam cooperatividade positiva (NELSON; COX; NELSON, 2013). Valores elevados também estão associados a uma maior probabilidade de o composto ser agregador, no entanto esta não é uma evidência conclusiva para tal comportamento (DOUGLAS S. AULD; JAMES INGLESE; JAYME L. DAHLIN, 2017).

BR20151



| Parâmetros | Experimento 1 | Experimento 2 |
|---------------------------|---------------|---------------|
| IC ₅₀ (µM) | 167,4 | 133,5 |
| Log IC ₅₀ (µM) | 2,224 | 2,127 |
| Topo (%) | = 100 | = 100 |
| Base (%) | = 0 | = 0 |
| Coeficiente de Hill | 2,005 | 1,986 |

Figura 26 – Curvas de IC₅₀ do composto BR20151. Cada curva corresponde a um experimento independente determinada com sete concentrações do composto, em duplicata.

O teste utilizando diferentes concentrações de Triton X-100 é consolidado na literatura como um experimento para detecção de inibição por agregação. Em situações onde há presença de inibição não específica, a formação de agregados coloidais é uma das principais fontes desse comportamento (SHOICHET, 2006). A utilização do detergente iônico Triton X-100 visa desassociar os agregados formados, o que impactaria negativamente na inibição da enzima pelo composto testado (FENG; SHOICHET, 2006). Ao analisar a **tabela 13** observamos que não houve alteração na inibição da enzima ao alterar as concentrações de Triton X-100, sugerindo um comportamento não agregador. No entanto, ao compararmos com a inibição nas condições padrão, (representada pela coluna com concentração enzimática de 0,2 nM, na ausência de Triton X-100), notamos uma queda acentuada na inibição, passando de 70% para 25%. Devido a inconclusividade desse experimento, optamos por adotar outras abordagens para detectar comportamento agregador.

Compostos com comportamento promíscuo tendem a se ligar a diferentes tipos de proteínas. Incubamos o composto com uma alta concentração (1 mg/mL) de BSA

antes da adição da NS2B-NS3pro. Caso o composto inibisse a enzima de forma específica, esperávamos que a incubação com o BSA não alterasse a inibição da enzima de interesse (FENG; SHOICHET, 2006; SHOICHET, 2006). Como observado na **tabela 13**, a adição de BSA, assim como ocorreu com o Triton X-100, diminuiu aproximadamente 50% da inibição observada se comparada com a condição padrão, sendo um indicativo de comportamento agregador. Realizamos também dois experimentos onde variamos apenas a concentração da enzima ou do substrato. Aumentar a concentração enzimática do ensaio em dez vezes deveria resultar na manutenção da inibição se o composto fosse um inibidor específico. Isso ocorre porque, mesmo com o aumento da concentração enzimática, a concentração do composto permaneceria significante superior, podendo manter a taxa de inibição (FENG; SHOICHET, 2006; SHOICHET, 2006). No caso de inibição por agregados coloidais, a inibição diminuiria, pois a inibição enzimática ocorrendo devido a uma inibição altamente influenciada pela estequiometria, devido a possibilidade de saturação dos agregados. O BR20151 apresentou o segundo comportamento ao diminuir a inibição de 70% para zero (**tabela 9**).

Tabela 9 – Inibição da NS2B-NS3pro pelo composto BR20151 em diferentes condições experimentais

| Comp. | % de inibição da ZIKV NS2B-NS3pro ¹ | | | | | | | | |
|---------|--|----------------------------|------------------|-------------------|------------------|----------------------------------|---------------------------------|--------------------------------|-----------------|
| | 100 µM | | IC ₅₀ | 200 µM | | | | | |
| | Com incubação | Sem incubação ² | | [enzima] = 0,2 nM | BSA ³ | 0,001% ⁴ Triton X-100 | 0,01% ⁴ Triton X-100 | 0,1% ⁴ Triton X-100 | [enzima] = 2 nM |
| BR20151 | 31 ± 14 | 37 ± 18 | 151 ± 24 | 70 ± 9 | 35 ± 12 | 23 ± 10 | 23 ± 8 | 28 ± 8 | 0 ± 12 |

¹tampão base 10 mM Tris-HCl pH 8,5; 20% glicerol e 1 mM CHAPS ² O teste foi realizado sem incubar o composto com a enzima por 10 minutos, todos os demais sofreram esse processo. ³ O tampão utilizado não possui nenhum detergente. ⁴ Substituição do detergente CHAPS pelo Triton no tampão empregado. Os experimentos foram empregados em triplicata, com exceção do teste de IC₅₀ que foi conduzido em duplicata.

A última análise realizada buscou avaliar a existência de um comportamento inibitório competitivo do composto. Para tanto, avaliamos a inibição enzimática em condições nas quais houve variação da concentração do substrato utilizado, mantendo a concentração enzimática (SAMRAT et al., 2022). A expectativa para um inibidor competitivo seria um aumento na inibição quando a concentração do substrato é menor, já que haveria menos competição pelo sítio de ligação. No entanto, ao observar

a **figura 27**, notamos uma variação aleatória nas inibições, indicando que o comportamento desse ligante não é competitivo. Esse comportamento é coerente com inibição por agregação (FENG; SHOICHET, 2006; SHOICHET, 2006), e não seria esperado para inibidores que se ligassem de forma competitiva com o substrato, no sítio ativo da enzima, como proposto a partir do docking. Desta forma, os compostos selecionados a partir da triagem virtual da BraCoLi exibiram baixa capacidade de inibição da protease. Adicionalmente, a caracterização do composto BR20151 revelou baixa potência, evidenciada pela IC_{50} , e vários indícios de inibição relacionados à de formação de agregados.

Inibição da NS2B-NS3pro em relação a variação da concentração de substrato

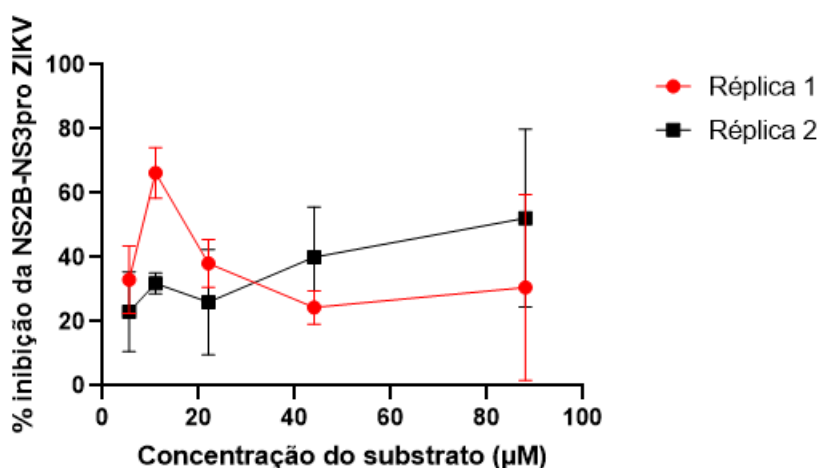


Figura 27 – Comportamento da inibição da enzima com variação da concentração do substrato. Experimento realizado em duplicatas e dois experimentos distintos. As concentrações de substrato usados foram 5,5 μM ; 11 μM ; 22 μM ; 44 μM e 88 μM .

5 Conclusão

Esse estudo nos mostrou uma grande similaridade entre as estruturas experimentalmente resolvidas da NS2B-NS3pro ZIKV, tanto em termos globais quanto no sítio ativo. A consistência dessas estruturas sugere que possuímos um modelo confiável da proteína em relação à conformação do sítio ativo na presença de um ligante que pode ser usado para futuras prospecções *in silico*. Contudo, o protocolo de triagem desenvolvido enfrentou desafios significativos durante as validações retrospectivas de reprodução de conformação de ligantes cristalográficos e performance na triagem. Essas dificuldades podem ser atribuídas a falhas tanto na preparação dos ligantes quanto a limitações do algoritmo de busca do Dock6 em lidar com ligantes muito flexíveis e que contenham anéis com muitos átomos, como os macrocíclicos. Além disso, a pontuação padrão do programa não considera o custo de dessolvatação em seu cálculo o que, junto com a característica negativa do sítio ativo, favorece ligantes com cargas positivas e não necessariamente os que teriam melhor encaixe. Uma abordagem sugerida para estudos futuros é utilizar outras funções de pontuação disponíveis no Dock6 para reavaliação das poses.

A utilização do LUNA se revelou uma ferramenta valiosa em diferentes fases do estudo. No preparo da caixa para realização do docking, permitiu identificar os resíduos essenciais a serem englobados por ela, a partir da frequência de interação destes com ligantes presentes em complexos cristalográficos.. Além disso, a análise de interação dos ligantes através do LUNA contribuiu para elucidação de seu perfil e compreensão das características do conjunto estudado.

A triagem da biblioteca BraCoLi, embora tenha resultado na seleção de 16 compostos para triagem experimental, resultou em compostos com baixa inibição enzimática. Mesmo no caso do BR20151, que foi submetido a uma caracterização mais detalhada, observamos indícios de que sua ação pode ser atribuída à agregação. O refinamento dos parâmetros utilizados para triagem virtual e a busca em bibliotecas mais extensas no futuro podem promover melhores resultados na identificação de potenciais inibidores.

6 Publicações e apresentações de trabalho

7 Publicações e manuscritos submetidos

7.1 Trabalhos publicados durante o mestrado

Lucianna H. Santos, Rafael E. O. Rocha, Diego L. Dias, Beatriz M. R. M. Ribeiro, Mateus Sá M. Serafim, Jônatas S. Abrahão, Rafaela S. Ferreira. Evaluating Know Zika Virus NS2B-NS3 Protease Inhibitor Scaffolds via In Silico Screening and Biochemical Assays. *Pharmaceuticals* 2023, 16, 1319. <https://doi.org/10.3390/ph16091319>

Trabalho realizado para identificação de inibidores contra NS2B-NS3pro ZIKV. Foram empregadas técnicas computacionais baseadas em ligantes e baseadas em receptor para identificação de possíveis inibidores. Os compostos selecionados foram testados contra a protease e em ensaios antivirais. Fui responsável por ensaios complementares sobre agregação dos compostos contra a NS2B-NS3pro ZIKV.

Renieidy F. C. Dias*, Beatriz M., R. M. Ribeiro*, Natasha M. Cassani*, Danilo N. Farago, Giovanna A. Antonucci, Rafael E. O. Rocha, Felipe O. Souza, Eduardo J. Pilau, Ana C. G. Jardim, Rafaela S. Ferreira, Celso O. Rezende Jr. Discovery and structural optimization of a new series of *N*-acyl-2-aminobenzothiazole as inhibitors of Zika virus. *Bioorganic & Medicinal Chemistry*, 2023, 95, 117488. [10.1016/j.bmc.2023.117488](https://doi.org/10.1016/j.bmc.2023.117488)

* co-primeira autoria

*Trabalho realizado para identificação de inibidores contra NS2B-NS3pro ZIKV. Foi desenvolvida e sintetizada uma série de compostos *N*-acil-2-aminobenzotiazol que foram avaliados quanto a inibição da replicação viral e na inibição da protease NS2B-NS3. Fui responsável pela caracterização bioquímica dos compostos testados contra NS2B-NS3pro ZIKV.*

7.2 Apresentações em congressos

34º Encontro regional da sociedade brasileira de química – 09-12/12/22022

Beatriz M. R. M. Ribeiro, Renieidy F. C. Dias; Celso O. Rezende Jr., Rafaela S. Ferreira. Detecção de falso-positivo durante triagem enzimática contra NS2B-NS3 Zika vírus

47th annual meeting of the Brazilian Biophysical Society – 19-22/10/2023

Beatriz M. R. M. Ribeiro, Lucianna H. S. Santos, Rafaela S. Ferreira. Development of a docking protocol Against Zika vírus NS2B-NS3pro

Brazmed Chem 2023 – 24-27/10/2023

Beatriz M. R. M. Ribeiro, Renieidy F. C. Dias; Celso O. Rezende Jr., Rafaela S. Ferreira. Biochemical characterization of potential Zika vírus NS2B-NS3pro inhibitors

8 REFERÊNCIAS

- ABRAMS, Rachel P. M. et al. Therapeutic candidates for the Zika virus identified by a high-throughput screen for Zika protease inhibitors. **Proceedings of the National Academy of Sciences**, [S. l.], v. 117, n. 49, p. 31365–31375, 2020. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.2005463117.
- ALLEN, Scott E.; DOKHOLYAN, Nikolay V.; BOWERS, Albert A. Dynamic Docking of Conformationally Constrained Macrocycles: Methods and Applications. **ACS Chemical Biology**, [S. l.], v. 11, n. 1, p. 10–24, 2016. ISSN: 1554-8929, 1554-8937. DOI: 10.1021/acscchembio.5b00663.
- ALLEN, William J.; BALIUS, Trent E.; MUKHERJEE, Sudipto; BROZELL, Scott R.; MOUSTAKAS, Demetri T.; LANG, P. Therese; CASE, David A.; KUNTZ, Irwin D.; RIZZO, Robert C. DOCK 6: Impact of new features and current docking performance. **Journal of Computational Chemistry**, [S. l.], v. 36, n. 15, p. 1132–1156, 2015. ISSN: 01928651. DOI: 10.1002/jcc.23905.
- ARMOUGOM, Fabrice; MORETTI, Sébastien; KEDUAS, Vladimir; NOTREDAME, Cedric. The iRMSD: a local measure of sequence alignment accuracy using structural information. **Bioinformatics**, [S. l.], v. 22, n. 14, p. e35–e39, 2006. ISSN: 1367-4811, 1367-4803. DOI: 10.1093/bioinformatics/btl218.
- BAJUSZ, Dávid; RÁCZ, Anita; HÉBERGER, Károly. Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations? **Journal of Cheminformatics**, [S. l.], v. 7, n. 1, p. 20, 2015. ISSN: 1758-2946. DOI: 10.1186/s13321-015-0069-3.
- BERMAN, H. M. Et al. The Protein Data Bank. **Nucleic Acids Research**, [S. l.], p. 235–242, 2000.
- B.J, Grant; A.P.C, Rodrigues; K.M, ElSawy; J.A, McCammon; L.S.D, Caves. Bio3D: An R package for the comparative analysis of protein structures. **Bioinformatics**, [S. l.], v. 22, p. 2695–2696, 2006.
- BRAUN, Niklas J. et al. Structure-Based Macrocyclization of Substrate Analogue NS2B-NS3 Protease Inhibitors of Zika, West Nile and Dengue viruses. **ChemMedChem**, [S. l.], v. 15, n. 15, p. 1439–1452, 2020. ISSN: 1860-7179, 1860-7187. DOI: 10.1002/cmdc.202000237.
- BRAUN, Niklas J.; HUBER, Simon; SCHMACKE, Luna C.; HEINE, Andreas; STEINMETZER, Torsten. Boroleucine-Derived Covalent Inhibitors of the ZIKV Protease. **ChemMedChem**, Germany, v. 18, n. 3, p. e202200336, 2023. ISSN: 1860-7187 1860-7179. DOI: 10.1002/cmdc.202200336.
- BRODNIEWICZ, Teresa; GRYNKIEWICZ, Grzegorz. Preclinical drug development. **Acta poloniae pharmaceutica**, Poland, v. 67, n. 6, p. 578–585, 2010. ISSN: 0001-6837.
- BROOIJMANS, Natasja; KUNTZ, Irwin D. Molecular Recognition and Docking Algorithms. **Annual Review of Biophysics and Biomolecular Structure**, [S. l.], v. 32,

n. 1, p. 335–373, 2003. ISSN: 1056-8700, 1545-4266. DOI: 10.1146/annurev.biophys.32.110601.142532.

BROWN, W. Clay; AKEY, David L.; KONWERSKI, Jamie R.; TARRASCH, Jeffrey T.; SKINIOTIS, Georgios; KUHN, Richard J.; SMITH, Janet L. Extended surface for membrane association in Zika virus NS1 structure. **Nature Structural & Molecular Biology**, [S. l.], v. 23, n. 9, p. 865–867, 2016. ISSN: 1545-9993, 1545-9985. DOI: 10.1038/nsmb.3268.

BROZELL, Scott R.; MUKHERJEE, Sudipto; BALIUS, Trent E.; ROE, Daniel R.; CASE, David A.; RIZZO, Robert C. Evaluation of DOCK 6 as a pose generation and database enrichment tool. **Journal of Computer-Aided Molecular Design**, [S. l.], v. 26, n. 6, p. 749–773, 2012. ISSN: 0920-654X, 1573-4951. DOI: 10.1007/s10822-012-9565-y.

BRUNO CÉSAR FELTES, Et Al. **Bioinformática: da biologia à flexibilidade molecular**. [s.l.] : SBBq, 2015. ISBN: 978-85-69288-00-8.

CAMPOS, Gubio S.; BANDEIRA, Antonio C.; SARDI, Silvia I. Zika Virus Outbreak, Bahia, Brazil. **Emerging Infectious Diseases**, [S. l.], v. 21, n. 10, p. 1885–1886, 2015. ISSN: 1080-6040, 1080-6059. DOI: 10.3201/eid2110.150847.

CASTRO-ALVAREZ, Alejandro; COSTA, Anna; VILARRASA, Jaume. The Performance of Several Docking Programs at Reproducing Protein–Macrolide-Like Crystal Structures. **Molecules**, [S. l.], v. 22, n. 1, p. 136, 2017. ISSN: 1420-3049. DOI: 10.3390/molecules22010136.

CHANG, Yiqun; HAWKINS, Bryson A.; DU, Jonathan J.; GROUNDWATER, Paul W.; HIBBS, David E.; LAI, Felcia. A Guide to In Silico Drug Design. **Pharmaceutics**, [S. l.], v. 15, n. 1, p. 49, 2022. ISSN: 1999-4923. DOI: 10.3390/pharmaceutics15010049.

CIEMNY, Maciej; KURCINSKI, Mateusz; KAMEL, Karol; KOLINSKI, Andrzej; ALAM, Nawsad; SCHUELER-FURMAN, Ora; KMIECIK, Sebastian. Protein–peptide docking: opportunities and challenges. **Drug Discovery Today**, [S. l.], v. 23, n. 8, p. 1530–1537, 2018. ISSN: 13596446. DOI: 10.1016/j.drudis.2018.05.006.

COPELAND, Robert A. Mechanistic considerations in high-throughput screening. **Analytical Biochemistry**, [S. l.], v. 320, n. 1, p. 1–12, 2003. ISSN: 00032697. DOI: 10.1016/S0003-2697(03)00346-4.

COX, Bryan D.; STANTON, Richard A.; SCHINAZI, Raymond F. Predicting Zika virus structural biology: Challenges and opportunities for intervention. **Antiviral Chemistry and Chemotherapy**, [S. l.], v. 24, n. 3–4, p. 118–126, 2015. ISSN: 2040-2066, 2040-2066. DOI: 10.1177/2040206616653873.

CUGOLA, Fernanda R. et al. The Brazilian Zika virus strain causes birth defects in experimental models. **Nature**, [S. l.], v. 534, n. 7606, p. 267–271, 2016. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature18296.

CUI, Xiangling et al. Identification of Theaflavin-3,3'-Digallate as a Novel Zika Virus Protease Inhibitor. **Frontiers in Pharmacology**, [S. l.], v. 11, p. 514313, 2020. ISSN: 1663-9812. DOI: 10.3389/fphar.2020.514313.

DE ARAÚJO, Thalia Velho Barreto et al. Association between microcephaly, Zika virus infection, and other risk factors in Brazil: final report of a case-control study. **The Lancet Infectious Diseases**, [S. l.], v. 18, n. 3, p. 328–336, 2018. ISSN: 14733099. DOI: 10.1016/S1473-3099(17)30727-2.

DIAS, Raquel; DE AZEVEDO JR., Walter. Molecular Docking Algorithms. **Current Drug Targets**, [S. l.], v. 9, n. 12, p. 1040–1047, 2008. ISSN: 13894501. DOI: 10.2174/138945008786949432.

DICK, G. W. A.; KITCHEN, S. F.; HADDOW, A. J. Zika Virus (I). Isolations and serological specificity. **Transactions of the Royal Society of Tropical Medicine and Hygiene**, [S. l.], v. 46, n. 5, p. 509–520, 1952. ISSN: 00359203. DOI: 10.1016/0035-9203(52)90042-4.

DOUGLAS S. AULD, Ph. D.; JAMES INGLESE, Ph. D.; JAYME L. DAHLIN, Ph. D., M. D. **Assay Guidance Manual**. [s.l.]: Eli Lilly & Company and the National Center for Advancing Translational Sciences, 2017. Disponível em: <https://www.ncbi.nlm.nih.gov/books/NBK442297/>.

DUFFY, Mark R. et al. Zika Virus Outbreak on Yap Island, Federated States of Micronesia. **New England Journal of Medicine**, [S. l.], v. 360, n. 24, p. 2536–2543, 2009. ISSN: 0028-4793, 1533-4406. DOI: 10.1056/NEJMoa0805715.

EKINS, S.; MESTRES, J.; TESTA, B. *In silico* pharmacology for drug discovery: methods for virtual ligand screening and profiling: *In silico* pharmacology for drug discovery. **British Journal of Pharmacology**, [S. l.], v. 152, n. 1, p. 9–20, 2007. ISSN: 00071188. DOI: 10.1038/sj.bjp.0707305.

FASSIO, Alexandre V.; SHUB, Laura; PONZONI, Luca; MCKINLEY, Jessica; O'MEARA, Matthew J.; FERREIRA, Rafaela S.; KEISER, Michael J.; DE MELO MINARDI, Raquel C. Prioritizing Virtual Screening with Interpretable Interaction Fingerprints. **Journal of Chemical Information and Modeling**, [S. l.], v. 62, n. 18, p. 4300–4318, 2022. ISSN: 1549-9596, 1549-960X. DOI: 10.1021/acs.jcim.2c00695.

FENG, Brian Y.; SHOICHET, Brian K. A detergent-based assay for the detection of promiscuous inhibitors. **Nature Protocols**, [S. l.], v. 1, n. 2, p. 550–553, 2006. ISSN: 1754-2189, 1750-2799. DOI: 10.1038/nprot.2006.77.

GHOSH, Arun K.; OSSWALD, Heather L.; PRATO, Gary. Recent Progress in the Development of HIV-1 Protease Inhibitors for the Treatment of HIV/AIDS. **Journal of Medicinal Chemistry**, [S. l.], v. 59, n. 11, p. 5172–5208, 2016. ISSN: 0022-2623, 1520-4804. DOI: 10.1021/acs.jmedchem.5b01697.

GRAPHPAD SOFTWARE. **GraphPad Prism**. 2023.

GRUBA, Natalia et al. Substrate profiling of Zika virus NS2B-NS3 protease. **FEBS Letters**, [S. l.], v. 590, n. 20, p. 3459–3468, 2016. ISSN: 00145793. DOI: 10.1002/1873-3468.12443.

GULLAND, Anne. Zika virus is a global public health emergency, declares WHO. **BMJ**, [S. l.], p. i657, 2016. ISSN: 1756-1833. DOI: 10.1136/bmj.i657.

HAMMERSCHMIDT, Stefan J.; HUBER, Simon; BRAUN, Niklas J.; LANDER, Marc; STEINMETZER, Torsten; KERSTEN, Christian. Thermodynamic characterization of a macrocyclic Zika virus NS2B/NS3 protease inhibitor and its acyclic analogs. **Archiv der Pharmazie**, [S. l.], v. 356, n. 4, p. 2200518, 2023. ISSN: 0365-6233, 1521-4184. DOI: 10.1002/ardp.202200518.

HAWKINS, Paul C. D.; SKILLMAN, A. Geoffrey; WARREN, Gregory L.; ELLINGSON, Benjamin A.; STAHL, Matthew T. Conformer Generation with OMEGA: Algorithm and Validation Using High Quality Structures from the Protein Databank and the Cambridge Structural Database. **Journal of Chemical Information and Modeling**, [S. l.], v. 50, n. 4, p. 572–584, 2010.

HUBER, Simon et al. Structure-Based Optimization and Characterization of Macrocyclic Zika Virus NS2B-NS3 Protease Inhibitors. **Journal of Medicinal Chemistry**, [S. l.], v. 65, n. 9, p. 6555–6572, 2022. ISSN: 0022-2623, 1520-4804. DOI: 10.1021/acs.jmedchem.1c01860.

KATZ, Itai; GILBURD, Boris; SHOVMAN, Ora. Zika autoimmunity and Guillain–Barré syndrome. **Current Opinion in Rheumatology**, [S. l.], v. 31, n. 5, p. 484–487, 2019. ISSN: 1040-8711, 1531-6963. DOI: 10.1097/BOR.0000000000000629.

KINDHAUSER, Mary Kay; ALLEN, Tomas; FRANK, Veronika; SANTHANA, Ravi Shankar; DYE, Christopher. Zika: the origin and spread of a mosquito-borne virus. **Bulletin of the World Health Organization**, [S. l.], v. 94, n. 9, p. 675- 686C, 2016. ISSN: 0042-9686. DOI: 10.2471/BLT.16.171082.

KITCHEN, Douglas B.; DECORNEZ, Hélène; FURR, John R.; BAJORATH, Jürgen. Docking and scoring in virtual screening for drug discovery: methods and applications. **Nature Reviews Drug Discovery**, [S. l.], v. 3, n. 11, p. 935–949, 2004. ISSN: 1474-1776, 1474-1784. DOI: 10.1038/nrd1549.

KNIGHT, Graham C. Fluorimetric assays of proteolytic enzymes. **Methods in Enzymology**, [S. l.], v. 248, p. 18–34, 1995. DOI: 10.1016/0076-6879(95)48004-8.

LANG, PT et al. **DOCK 6.6 User's Manual**. [s.l.: s.n.].

LEI, Jian; HANSEN, Guido; NITSCHKE, Christoph; KLEIN, Christian D.; ZHANG, Linlin; HILGENFELD, Rolf. Crystal structure of Zika virus NS2B-NS3 protease in complex with a boronate inhibitor. **Science**, [S. l.], v. 353, n. 6298, p. 503–505, 2016. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.aag2419.

LEUNG, Jason Y.; PIJLMAN, Gorben P.; KONDRATIEVA, Natasha; HYDE, Jennifer; MACKENZIE, Jason M.; KHROMYKH, Alexander A. Role of Nonstructural Protein NS2A in Flavivirus Assembly. **Journal of Virology**, [S. l.], v. 82, n. 10, p. 4731–4741, 2008. ISSN: 0022-538X, 1098-5514. DOI: 10.1128/JVI.00002-08.

LI, Yan et al. Structural Dynamics of Zika Virus NS2B-NS3 Protease Binding to Dipeptide Inhibitors. **Structure**, [S. l.], v. 25, n. 8, p. 1242- 1250.e3, 2017. ISSN: 09692126. DOI: 10.1016/j.str.2017.06.006.

LI, Yan et al. Structural Insights into the Inhibition of Zika Virus NS2B-NS3 Protease by a Small-Molecule Inhibitor. **Structure**, [S. l.], v. 26, n. 4, p. 555- 564.e3, 2018. ISSN: 09692126. DOI: 10.1016/j.str.2018.02.005.

LI, Zhong et al. Methylene blue is a potent and broad-spectrum inhibitor against Zika virus *in vitro* and *in vivo*. **Emerging Microbes & Infections**, [S. l.], v. 9, n. 1, p. 2404–2416, 2020 a. ISSN: 2222-1751. DOI: 10.1080/22221751.2020.1838954.

LI, Zhong et al. JMX0207, a Niclosamide Derivative with Improved Pharmacokinetics, Suppresses Zika Virus Infection Both *In Vitro* and *In Vivo*. **ACS Infectious Diseases**, [S. l.], v. 6, n. 10, p. 2616–2628, 2020 b. ISSN: 2373-8227, 2373-8227. DOI: 10.1021/acsinfecdis.0c00217.

MA, Jinzhu et al. Zika Virus Non-structural Protein 4A Blocks the RLR-MAVS Signaling. **Frontiers in Microbiology**, [S. l.], v. 9, p. 1350, 2018. ISSN: 1664-302X. DOI: 10.3389/fmicb.2018.01350.

MAHAWATHTHA, Mithun C.; PEARCE, Benjamin J. G.; SZABO, Monika; GRAHAM, Bim; KLEIN, Christian D.; NITSCHKE, Christoph; OTTING, Gottfried. Solution conformations of a linked construct of the Zika virus NS2B-NS3 protease. **Antiviral Research**, [S. l.], v. 142, p. 141–147, 2017. ISSN: 01663542. DOI: 10.1016/j.antiviral.2017.03.011.

MARVIN 23.2. **Marvin was used for drawing, displaying and characterizing chemical structures, substructures and reactions**. ChemAxon, 2023. Disponível em: <http://www.chemaxon.com>.

MEYDER, Agnes; NITTINGER, Eva; LANGE, Gudrun; KLEIN, Robert; RAREY, Matthias. Estimating Electron Density Support for Individual Atoms and Molecular Fragments in X-ray Structures. **Journal of Chemical Information and Modeling**, [S. l.], v. 57, n. 10, p. 2437–2447, 2017. DOI: 10.1021/acs.jcim.7b00391.

MOORE, Cynthia A. et al. Characterizing the Pattern of Anomalies in Congenital Zika Syndrome for Pediatric Clinicians. **JAMA Pediatrics**, [S. l.], v. 171, n. 3, p. 288, 2017. ISSN: 2168-6203. DOI: 10.1001/jamapediatrics.2016.3982.

MUKHERJEE, Sudipto; BALIUS, Trent E.; RIZZO, Robert C. Docking Validation Resources: Protein Family and Ligand Flexibility Experiments. **Journal of Chemical Information and Modeling**, [S. l.], v. 50, n. 11, p. 1986–2000, 2010. ISSN: 1549-9596, 1549-960X. DOI: 10.1021/ci1001982.

MUSSO, D.; NILLES, E. J.; CAO-LORMEAU, V. M. Rapid spread of emerging Zika virus in the Pacific area. **Clinical Microbiology and Infection**, [S. l.], v. 20, n. 10, p. O595–O596, 2014. ISSN: 1198743X. DOI: 10.1111/1469-0691.12707.

MYSINGER, Michael M.; CARCHIA, Marco; IRWIN, John J.; SHOICHET, Brian K. DUD-E: The Drug-Set Enrichment Analysis database. **Journal of Medicinal Chemistry**, [S. l.], v. 55, n. 14, p. 6582–6594, 2012. DOI: 10.1021/jm300687e.

NELSON, David L.; COX, Michael M.; NELSON, David L. **Lehninger principles of biochemistry**. Sixth edition ed., Basingstoke: Macmillan Higher Education, 2013. ISBN: 978-1-4292-3414-6.

NITSCHKE, Christoph; ONAGI, Hideki; QUEK, Jun-Ping; OTTING, Gottfried; LUO, Dahai; HUBER, Thomas. Biocompatible Macrocyclization between Cysteine and 2-Cyanopyridine Generates Stable Peptide Inhibitors. **Organic Letters**, [S. l.], v. 21, n. 12, p. 4709–4712, 2019. ISSN: 1523-7060, 1523-7052. DOI: 10.1021/acs.orglett.9b01545.

OEHLER, E. et al. Zika virus infection complicated by Guillain-Barré syndrome – case report, French Polynesia, December 2013. **Eurosurveillance**, [S. l.], v. 19, n. 9, 2014. ISSN: 1560-7917. DOI: 10.2807/1560-7917.ES2014.19.9.20720. Disponível em: <https://www.eurosurveillance.org/content/10.2807/1560-7917.ES2014.19.9.20720>. Acesso em: 11 ago. 2023.

OLSON, J. G.; KSIAZEK, T. G.; SUHANDIMAN; TRIWIBOWO. Zika virus, a cause of fever in Central Java, Indonesia. **Transactions of the Royal Society of Tropical Medicine and Hygiene**, [S. l.], v. 75, n. 3, p. 389–393, 1981. ISSN: 00359203. DOI: 10.1016/0035-9203(81)90100-0.

OPENEYE, CADENCE MOLECULAR SCIENCES. **QUACPAC 2.2.2.0**. Santa Fe, NM, 2019. Disponível em: <http://www.eyesopen.com>.

PATIL, Nitin A.; QUEK, Jun-Ping; SCHROEDER, Barbara; MOREWOOD, Richard; RADEMANN, Jörg; LUO, Dahai; NITSCHKE, Christoph. 2-Cyanoisonicotinamide Conjugation: A Facile Approach to Generate Potent Peptide Inhibitors of the Zika Virus Protease. **ACS Medicinal Chemistry Letters**, [S. l.], v. 12, n. 5, p. 732–737, 2021. ISSN: 1948-5875, 1948-5875. DOI: 10.1021/acsmedchemlett.0c00657.

PETTERSEN, Eric F.; GODDARD, Thomas D.; HUANG, Conrad C.; COUCH, Gregory S.; GREENBLATT, Daniel M.; MENG, Elaine C.; FERRIN, Thomas E. UCSF Chimera-a visualization system for exploratory research and analysis. **Journal of computational chemistry**, United States, v. 25, n. 13, p. 1605–1612, 2004. ISSN: 0192-8651. DOI: 10.1002/jcc.20084.

PHOO, Wint Wint et al. Structure of the NS2B-NS3 protease from Zika virus after self-cleavage. **Nature Communications**, [S. l.], v. 7, n. 1, p. 13410, 2016. ISSN: 2041-1723. DOI: 10.1038/ncomms13410.

PHOO, Wint Wint; ZHANG, Zhenzhen; WIRAWAN, Melissa; CHEW, Edwin Jun Chen; CHEW, Alvin Bing Liang; KOURETOVA, Jenny; STEINMETZER, Torsten; LUO, Dahai. Structures of Zika virus NS2B-NS3 protease in complex with peptidomimetic inhibitors. **Antiviral Research**, [S. l.], v. 160, p. 17–24, 2018. ISSN: 01663542. DOI: 10.1016/j.antiviral.2018.10.006.

QUEK, Jun Ping; LIU, Shuang; ZHANG, Zhenzhen; LI, Yan; NG, Elizabeth Yihui; LOH, Ying Ru; HUNG, Alvin W.; LUO, Dahai; KANG, CongBao. Identification and structural characterization of small molecule fragments targeting Zika virus NS2B-NS3 protease. **Antiviral Research**, [S. l.], v. 175, p. 104707, 2020. ISSN: 01663542. DOI: 10.1016/j.antiviral.2020.104707.

RAMÍREZ, David; CABALLERO, Julio. Is It Reliable to Take the Molecular Docking Top Scoring Position as the Best Solution without Considering Available Structural Data?

Molecules, [S. l.], v. 23, n. 5, p. 1038, 2018. ISSN: 1420-3049. DOI: 10.3390/molecules23051038.

RASSIAS, Gerasimos et al. Cell-active carbazole derivatives as inhibitors of the zika virus protease. **European Journal of Medicinal Chemistry**, [S. l.], v. 180, p. 536–545, 2019. ISSN: 02235234. DOI: 10.1016/j.ejmech.2019.07.007.

SAMRAT, Subodh K.; XU, Jimin; LI, Zhong; ZHOU, Jia; LI, Hongmin. Antiviral Agents against Flavivirus Protease: Prospect and Future Direction. **Pathogens**, [S. l.], v. 11, n. 3, p. 293, 2022. ISSN: 2076-0817. DOI: 10.3390/pathogens11030293.

SANTOS, Karina B.; GUEDES, Isabella A.; KARL, Ana L. M.; DARDENNE, Laurent E. Highly Flexible Ligand Docking: Benchmarking of the DockThor Program on the LEADS-PEP Protein–Peptide Data Set. **Journal of Chemical Information and Modeling**, [S. l.], v. 60, n. 2, p. 667–683, 2020. ISSN: 1549-9596, 1549-960X. DOI: 10.1021/acs.jcim.9b00905.

SANTOS, Lucianna H.; CAFFARENA, Ernesto R.; FERREIRA, Rafaela S. pH and non-covalent ligand binding modulate Zika virus NS2B/NS3 protease binding site residues: Discoveries from MD and constant pH MD simulations. **Journal of Biomolecular Structure and Dynamics**, [S. l.], v. 40, n. 20, p. 10359–10372, 2022. ISSN: 0739-1102. DOI: 10.1080/07391102.2021.1943528.

SCHRÖDINGER, LLC. **The PyMOL Molecular Graphics System, Version 1.8**. [S. l.], 2015.

SCHRÖDINGER, LLC. **Schrödinger Release 2023-2: Maestro**. New York, NY, 2023.

SHOICHET, Brian K. Screening in a spirit haunted world. **Drug Discovery Today**, [S. l.], v. 11, n. 13–14, p. 607–615, 2006. ISSN: 13596446. DOI: 10.1016/j.drudis.2006.05.014.

SING, Tobias; SANDER, Oliver; BEERENWINKEL, Niko; LENGAUER, Thomas. ROCr: visualizing classifier performance in R. **Bioinformatics (Oxford, England)**, England, v. 21, n. 20, p. 3940–3941, 2005. ISSN: 1367-4803. DOI: 10.1093/bioinformatics/bti623.

STEIN, Reed M.; YANG, Ying; BALIUS, Trent E.; O'MEARA, Matt J.; LYU, Jiankun; YOUNG, Jennifer; TANG, Khanh; SHOICHET, Brian K.; IRWIN, John J. Property-Unmatched Decoys in Docking Benchmarks. **Journal of Chemical Information and Modeling**, [S. l.], v. 61, n. 2, p. 699–714, 2021. ISSN: 1549-9596. DOI: 10.1021/acs.jcim.0c00598.

SYSTÈMES, Dassault. **Discovery Studio Visualizer**. San Diego Dassault Systèmes, 2020.

TEMML, Veronika; KUTIL, Zsofia. Structure-based molecular modeling in SAR analysis and lead optimization. **Computational and Structural Biotechnology Journal**, [S. l.], v. 19, p. 1431–1444, 2021. ISSN: 20010370. DOI: 10.1016/j.csbj.2021.02.018.

THOMAS, Brittany N.; PARRILL, Abby L.; BAKER, Daniel L. Self-docking and cross-docking simulations of G protein-coupled receptor-ligand complexes: Impact of ligand type and receptor activation state. **Journal of Molecular Graphics and Modelling**, [S. l.], v. 112, p. 108119, 2022. ISSN: 10933263. DOI: 10.1016/j.jmgm.2021.108119.

TIAN, Hongliang et al. Structural basis of Zika virus helicase in recognizing its substrates. **Protein & Cell**, [S. l.], v. 7, n. 8, p. 562–570, 2016. ISSN: 1674-800X, 1674-8018. DOI: 10.1007/s13238-016-0293-2.

TRIBALLEAU, Nicolas; ACHER, Francine; BRABET, Isabelle; PIN, Jean-Philippe; BERTRAND, Hugues-Olivier. Virtual Screening Workflow Development Guided by the “Receiver Operating Characteristic” Curve Approach. Application to High-Throughput Docking on Metabotropic Glutamate Receptor Subtype 4. **Journal of Medicinal Chemistry**, [S. l.], v. 48, n. 7, p. 2534–2547, 2005. ISSN: 0022-2623, 1520-4804. DOI: 10.1021/jm049092j.

VERÍSSIMO, Gabriel Corrêa; DOS SANTOS JÚNIOR, Valtair Severino; DE ALMEIDA, Ingrid Ariela do Rosário; RUAS, Marina Sant’Anna Mitraud; COUTINHO, Lukas Galuppo; DE OLIVEIRA, Renata Barbosa; ALVES, Ricardo José; MALTAROLLO, Vinícius Gonçalves. The Brazilian compound library (BraCoLi) database: a repository of chemical and biological information for drug design. **Molecular diversity**, Netherlands, v. 26, n. 6, p. 3387–3397, 2022. ISSN: 1573-501X 1381-1991. DOI: 10.1007/s11030-022-10386-9.

WALTERS, W. Patrick; STAHL, Matthew T.; MURCKO, Mark A. Virtual screening—an overview. **Drug Discovery Today**, [S. l.], v. 3, n. 4, p. 160–178, 1998. ISSN: 13596446. DOI: 10.1016/S1359-6446(97)01163-X.

WANG, Boxiao; THURMOND, Stephanie; HAI, Rong; SONG, Jikui. Structure and function of Zika virus NS5 protein: perspectives for drug design. **Cellular and Molecular Life Sciences**, [S. l.], v. 75, n. 10, p. 1723–1736, 2018. ISSN: 1420-682X, 1420-9071. DOI: 10.1007/s00018-018-2751-x.

WANG, Yan; XIE, Xuping; SHI, Pei-Yong. Flavivirus NS4B protein: Structure, function, and antiviral discovery. **Antiviral Research**, [S. l.], v. 207, p. 105423, 2022. ISSN: 01663542. DOI: 10.1016/j.antiviral.2022.105423.

WORLD HEALTH ORGANISATION. **Zika Epidemiology Update - February 2022**. 2022. Disponível em: <https://www.who.int/publications/m/item/zika-epidemiology-update---february-2022>. Acesso em: 14 ago. 2023.

WU, Kong-Yan; ZUO, Guo-Long; LI, Xiao-Feng; YE, Qing; DENG, Yong-Qiang; HUANG, Xing-Yao; CAO, Wu-Chun; QIN, Cheng-Feng; LUO, Zhen-Ge. Vertical transmission of Zika virus targeting the radial glial cells affects cortex development of offspring mice. **Cell Research**, [S. l.], v. 26, n. 6, p. 645–654, 2016. ISSN: 1001-0602, 1748-7838. DOI: 10.1038/cr.2016.58.

XING, HuaiPeng; XU, Shan; JIA, Fangfei; YANG, Yang; XU, Caimin; QIN, Chengfeng; SHI, Lei. Zika NS2B is a crucial factor recruiting NS3 to the ER and activating its protease activity. **Virus Research**, [S. l.], v. 275, p. 197793, 2020. ISSN: 01681702. DOI: 10.1016/j.virusres.2019.197793.

XIONG, Yanchao; CHENG, Fei; ZHANG, Junyi; SU, Haixia; HU, Hangchen; ZOU, Yi; LI, Minjun; XU, Yechun. Structure-based design of a novel inhibitor of the ZIKA virus NS2B/NS3 protease. **Bioorganic Chemistry**, [S. l.], v. 128, p. 106109, 2022. ISSN: 00452068. DOI: 10.1016/j.bioorg.2022.106109.

XU, Shan; CI, Yali; WANG, Leijie; YANG, Yang; ZHANG, Leiliang; XU, Caimin; QIN, Chengfeng; SHI, Lei. Zika virus NS3 is a canonical RNA helicase stimulated by NS5 RNA polymerase. **Nucleic Acids Research**, [S. l.], v. 47, n. 16, p. 8693–8707, 2019. ISSN: 0305-1048, 1362-4962. DOI: 10.1093/nar/gkz650.

XU, Xiaoying; SONG, Hao; QI, Jianxun; LIU, Yuqian; WANG, Haiyuan; SU, Chao; SHI, Yi; GAO, George F. Contribution of intertwined loop to membrane association revealed by Zika virus full-length NS 1 structure. **The EMBO Journal**, [S. l.], v. 35, n. 20, p. 2170–2178, 2016. ISSN: 0261-4189, 1460-2075. DOI: 10.15252/embj.201695290.

YIN, Zheng et al. Peptide inhibitors of dengue virus NS3 protease. Part 1: Warhead. **Bioorganic & Medicinal Chemistry Letters**, [S. l.], v. 16, n. 1, p. 36–39, 2006. ISSN: 0960894X. DOI: 10.1016/j.bmcl.2005.09.062.

YUAN, Shuofeng et al. Structure-based discovery of clinically approved drugs as Zika virus NS2B-NS3 protease inhibitors that potently inhibit Zika virus infection in vitro and in vivo. **Antiviral Research**, [S. l.], v. 145, p. 33–43, 2017. ISSN: 01663542. DOI: 10.1016/j.antiviral.2017.07.007.

ZHANG, Xianwen; XIE, Xuping; XIA, Hongjie; ZOU, Jing; HUANG, Linfen; POPOV, Vsevolod L.; CHEN, Xinwen; SHI, Pei-Yong. Zika Virus NS2A-Mediated Virion Assembly. **mBio**, [S. l.], v. 10, n. 5, p. e02375-19, 2019. ISSN: 2161-2129, 2150-7511. DOI: 10.1128/mBio.02375-19.

ZHANG, Xingquan. Direct anti-HCV agents. **Acta Pharmaceutica Sinica B**, [S. l.], v. 6, n. 1, p. 26–31, 2016. ISSN: 22113835. DOI: 10.1016/j.apsb.2015.09.008.

ZHANG, Zhenzhen; LI, Yan; LOH, Ying Ru; PHOO, Wint Wint; HUNG, Alvin W.; KANG, CongBao; LUO, Dahai. Crystal structure of unlinked NS2B-NS3 protease from Zika virus. **Science**, [S. l.], v. 354, n. 6319, p. 1597–1600, 2016. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.aai9309.

ZHONG, Shifa; GUAN, Xiaohong. Count-Based Morgan Fingerprint: A More Efficient and Interpretable Molecular Representation in Developing Machine Learning-Based Predictive Regression Models for Water Contaminants' Activities and Properties. **Environmental Science & Technology**, [S. l.], v. 57, n. 46, p. 18193–18202, 2023. DOI: 10.1021/acs.est.3c02198.

ZHOU, Shu-Feng; ZHONG, Wei-Zhu. Drug Design and Discovery: Principles and Applications. **Molecules**, [S. l.], v. 22, n. 2, p. 279, 2017. ISSN: 1420-3049. DOI: 10.3390/molecules22020279.