

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA DE SÃO CARLOS

Ketllyn Irene Zagato de Oliveira

Biologia estrutural e busca por inibidores da proteína não-estrutural NS5
domínio RNA polimerase dependente de RNA do vírus Zika

São Carlos
2021

KETLLYN IRENE ZAGATO DE OLIVEIRA

Biologia estrutural e busca por inibidores da proteína não-estrutural NS5
domínio RNA polimerase dependente de RNA do vírus Zika

Tese apresentada ao Programa de Pós-Graduação em Física do Instituto de Física de São Carlos da Universidade de São Paulo, para obtenção do título de Doutora em Ciências.

Área de concentração: Física Biomolecular
Orientador: Prof. Dr. Glaucius Oliva

Versão Original

São Carlos
2021

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Oliveira, Ketllyn Irene Zagato de
Biologia estrutural e busca por inibidores da proteína não-estrutural NS5 domínio RNA polimerase dependente de RNA do vírus Zika / Ketllyn Irene Zagato de Oliveira; orientador Glaucius Oliva -- São Carlos, 2021.
99 p.

Tese (Doutorado - Programa de Pós-Graduação em Física Biomolecular) -- Instituto de Física de São Carlos, Universidade de São Paulo, 2021.

1. Flavíivirus . 2. Zika vírus. 3. RNA polimerase dependente de RNA. I. Oliva, Glaucius, orient. II. Título.

FOLHA DE APROVAÇÃO

Ketllyn Irene Zagato de Oliveira

Tese apresentada ao Instituto de Física de São Carlos da Universidade de São Paulo para obtenção do título de Doutora em Ciências. Área de Concentração: Física Biomolecular.

Aprovado(a) em: 07/04/2021

Comissão Julgadora

Dr(a). Glaucius Oliva

Instituição: (IFSC/USP)

Dr(a). Ana Carolina Gomes Jardim

Instituição: (UFU/Uberlândia)

Dr(a). Dulce Helena Ferreira de Souza

Instituição: (UFSCar/São Carlos)

Dr(a). Andrea Balan Fernandes

Instituição: (ICB/USP)

Dr(a). Marcio Vinícius Bertacine Dias

Instituição: (ICB/USP)

Dedico este trabalho à minha sogra Ivanilde e ao meu tio Gerson,
que me deixaram recentemente, mas que foram fundamentais,
para que eu chegasse até aqui.

AGRADECIMENTOS

A Deus pela oportunidade desta vida, pelas pessoas que colocou em meu caminho, e por ter sido o meu suporte sempre.

Ao meu companheiro de vida Wellington, pela paciência, incentivo e por estar ao meu lado, não me deixando desistir jamais, muitas vezes acreditando muito mais em mim do que eu mesma.

Aos meus pais, Rovilson e Francisca, que não mediram esforços para que eu tivesse acesso as melhores opções educacionais, o que me possibilitou melhores oportunidades. E a minha irmã, Kevllyn pelo apoio ao longo da vida. Sou grata por tudo que fizeram por mim.

Ao Prof. Dr. Glaucius, pela oportunidade e pela orientação. Durante esse processo, muito me ensinou, contribuindo para meu crescimento científico, intelectual e pessoal.

Aos meus amigos de laboratório (Guilherme, Juliana, Camila Z., Gabriela, Victor O., Camila T., Camila B. e Victor R.) que me acompanharam nos almoços de todos os dias, vocês me ajudaram e me alegraram durante esse doutorado. Nossos almoços foram regados a muitas risadas e discussões científicas que contribuíram para esse trabalho. E em especial a Naiara por todo incentivo, apoio e por ter sido uma luz durante meu doutorado.

Às Dras. Caroline e Nathalya, por todos os ensinamentos e paciência, com vocês aprendi muito sobre a ciência e a humildade.

Aos Drs. André, Rafaela, Renata, Amanda e Fernando Maluf e a Marjorie por todos os ensinamentos e ajuda durante esse processo.

Aos técnicos da Biofísica Andressa, Fernando, Rafael, em especial, à Bel, e aos técnicos da cristalografia Humberto e Suzana, pelo suporte, ensinamentos e cafés.

Às minhas amigas Thais, Larissa, Lhaís e Catharina por todo apoio e suporte durante esse processo.

Aos Profs. Drs. Marcio e Rafael, pelas oportunidades e discussões onde foi possível muito aprendizado.

Ao Instituto de Física de São Carlos, pela oportunidade de realização do curso de doutorado.

À Fundação de Amparo à Pesquisa do Estado de São Paulo, pela concessão da bolsa de Doutorado Direto (Processos 2016/17153-2 e 2018/19574-0) pelo apoio financeiro para a realização desta pesquisa.

A todos os professores que passaram pela minha vida, vocês foram essenciais para que eu chegasse até aqui, em especial à Dras. Cintia Matiucci Pereira e Andreia Pereira Matos, que me apresentaram às possibilidades acadêmicas.

E por fim a todos aqueles amigos e familiares que de forma direta ou indireta colaboraram para que esse sonho fosse possível.

*“Você nunca sabe que resultados virão da sua ação. Mas se
você não fizer nada, não existirão resultados.”*

Mahatma Gandhi

RESUMO

OLIVEIRA, K. I. Z. **Biologia estrutural e busca por inibidores da proteína não-estrutural NS5 domínio RNA polimerase dependente de RNA do vírus Zika.** 2021. 99 p. Tese (Doutorado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2021.

O vírus Zika (ZIKV) é um membro da família *Flaviviridae* que teve a primeira ocorrência relatada no Brasil em abril de 2015. O maior impacto do ZIKV ocorreu devido a sua associação com o aumento no número de casos de microcefalia em bebês e fetos de gestantes infectadas com o vírus durante a epidemia que ocorreu no Brasil entre 2015 e 2016. Desde então, mais de 240.000 casos de febre Zika foram notificados no país, sendo 10.768 casos registrados no ano de 2019 e 6.705 casos prováveis no ano de 2020. Diante desses fatos, muitos estudos encontram-se em andamento para a busca de candidatos a fármacos eficazes contra o ZIKV. Dentre as proteínas não estruturais, a NS5 desperta grande interesse como alvo farmacológico, devido a já existir um inibidor de NS5 domínio RNA polimerase dependente de RNA (RdRp) aprovado para uso clínico contra o vírus da Hepatite C. A RdRp é responsável pela síntese de RNA *de novo*, ou seja, sem a necessidade de um primer para o início da síntese. Promove a elongação do RNA viral, através da polimerização fazendo o uso de NTPs, garantindo assim a replicação viral. A inibição de qualquer uma dessas funções acarretaria o impedimento da replicação viral. Nesse contexto, o objetivo deste trabalho foi obter a proteína NS5 do ZIKV domínio RdRp a fim de determinar a sua estrutura tridimensional por cristalografia de raios X, seguida da descoberta de inibidores. Para isso, planejamos duas construções para a proteína NS5 RdRp: a RdRp_truncada e a RdRp_integral. Com a RdRp_truncada foi possível obter cristais que produziram difração de raios X, o que possibilitou a resolução da estrutura cristalográfica da NS5 domínio RdRp a 1.9 Å. Essa estrutura foi utilizada para realizar a triagem computacional de inibidores, assim como ensaios biofísicos para a busca de possíveis ligantes. Com a RdRp_integral, foi possível a padronização de um ensaio baseado no marcador fluorescente de ácido nucleico SYBR Green I, que monitora a atividade da proteína em tempo real, e que foi utilizado na busca por inibidores em larga escala. Foi possível triar a inibição de mais de 500 compostos, incluindo a biblioteca de compostos “drug-like” da Pandemic Response Box (Medicines for Malaria Venture - MMV) (PRB), e os compostos sintéticos e provenientes de produtos naturais que foram encontrados por triagem biofísica e virtual. Ao todo, foram identificados 17 potenciais inibidores da NS5 RdRp, sendo que 9 deles apresentaram inibição na faixa micromolar ($IC_{50} \leq 10 \mu M$) e 5 deles apresentaram inibição na faixa nanomolar ($IC_{50} \leq 1 \mu M$). Os resultados obtidos neste trabalho demonstram que, apesar da ausência de estruturas cristalográficas complexadas com ligantes, foi possível encontrar moléculas com alta inibição, que podem ter sua ação otimizada por ensaios computacionais e posteriormente testadas contra a RdRp. Dentre elas, os compostos provenientes de produtos naturais que se mostraram extremamente interessantes, em especial os derivados do anel flavonol, que se mostram potentes na inibição da atividade da RdRp.

Palavras-chave: Flavivírus. Zika vírus. RNA polimerase dependente de RNA.

ABSTRACT

OLIVEIRA, K. I. Z. **Structural biology and search for inhibitors of the non-structural protein NS5 RNA-dependent RNA polymerase domain of the Zika virus**. 2021. 99 p. Tese (Doutorado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2021.

The Zika virus (ZIKV) is a member of the *Flaviviridae* family that was first reported in Brazil in April 2015. The major impact of ZIKV was due to its association with the increase in the number of microcephaly cases reported in newborns of pregnant women infected with the virus during the epidemic that occurred in Brazil between 2015 and 2016. Since then, more than 240.000 cases of Zika fever have been reported in the country, with 10.768 cases registered in 2019 and 6.705 probable cases in 2020. Given these facts, many studies are underway to search for candidates for effective drugs against ZIKV. Among the ZIKV non-structural proteins, NS5 is of great interest as a pharmacological target, due to the existence of an inhibitor approved for clinical use against the Hepatitis C virus NS5 RNA-dependent RNA polymerase domain (RdRp). RdRp is responsible for the *de novo* synthesis of viral RNA, that is, without the need for a primer to start the synthesis. It promotes the elongation of viral RNA, through polymerization using NTPs, thus guaranteeing viral replication. Inhibition of any of these functions would prevent viral replication. In this context, the objective of this project was to obtain the NS5 RdRp domain from the ZIKV in order to determine its three-dimensional structure by X-ray crystallography, followed by the discovery of inhibitors. For this purpose, we designed two constructions for the NS5 RdRp protein: the RdRp_truncated and the RdRp_integral. With the RdRp_truncated, it was possible to obtain crystals that produced X-ray diffraction, which enabled the elucidation of the crystallographic structure of the NS5 RdRp domain at 1.9 Å resolution. This structure was used to perform a computational screening of inhibitors as well as biophysical assays to search for possible ligands. With RdRp_integral, it was possible to standardize an assay based on the fluorescent nucleic acid marker SYBR Green I, which monitors the activity of the protein in real time, and which was used in the extended search for inhibitors. It was possible to test the inhibition of more than 500 compounds, including the drug-like compound library *Pandemic Response Box* (PRB, Medicines for Malaria Venture - MMV), as well as synthetic compounds and natural products that were previously identified by biophysical methods and virtual screening. In all, 17 potential NS5 RdRp inhibitors were identified, with 9 of them showing inhibition in the micromolar range ($IC_{50} \leq 10 \mu M$) and 5 of them showing inhibition in the nanomolar range ($IC_{50} \leq 1 \mu M$). The results obtained in this work demonstrate that despite the absence of crystallographic structures complexed with ligands, it was possible to find molecules with high inhibition, which can have their action optimized by computational approaches and subsequently tested against the ZIKV NS5 RdRp. Among them, the natural products that have shown good RdRp inhibitory activity seem to be very promising for further studies, particularly those derived from the flavonol ring.

Keywords: Flavivirus. Zika vírus. NS5 RNA-dependent RNA polymerase domain.

LISTA DE FIGURAS

- Figura 1 Mapa da disseminação do ZIKV pelo mundo. Em vermelho está Uganda, país africano, onde houve a primeira ocorrência do ZIKV em 1947, em roxo todos os países onde houve relatos de infecção pelo ZIKV e em amarelo são os países onde não houve nenhum relato de infecção pelo ZIKV.....26
- Figura 2 A partícula viral do ZIKV possui um material genético constituído por uma fita simples de cadeia positiva de RNA, com aproximadamente 11 Kb. Esta fita é traduzida como uma única poliproteína e após maturação é clivada por proteases virais e do hospedeiro em 3 proteínas estruturais e 7 proteínas não estruturais as NSs.....29
- Figura 3 Esquema do ciclo de replicação do vírus da Zika. A proteína E é responsável pela ligação do vírus aos receptores de membrana do hospedeiro, levando o vírus a entrar na célula por endocitose. A redução no pH endossomal induz a mudança conformacional na proteína E que permite a fusão das membranas do vírus às do hospedeiro O RNA viral é então liberado no citoplasma. O RNA de fita simples (ssRNA+) é traduzido em uma poliproteína, que é então clivada em várias proteínas virais. Os genomas virais se replicam na superfície do retículo endoplasmático (RE). No complexo de Golgi, a proteína prM é clivada, induzindo a maturação dos virions. Por fim a partícula viral madura sai da célula por exocitose.....30
- Figura 4 O Sofosbuvir é um pró-fármaco, usado no tratamento de HCV, cujo alvo molecular é a RdRp. A metabolização intracelular produz sua molécula ativa, que atua no sítio de síntese de RNA, competindo com os substratos naturais e impedindo a síntese de RNA.34
- Figura 5 Malet e colaboradores propõem duas construções para o domínio RdRp de WNV, uma integral e ativa (2HFZ) que ao ser cristalizada apresenta média resolução (3,0 Å) – construção em azul; e uma construção truncada (2HCN), que possui as mesmas orientações dos aminoácidos do sítio ativo e conformações, porém inativa, que ao ser difratada apresenta mais alta resolução (1,8 Å) – construção laranja.....37
- Figura 6 Gel de agarose 1% da PCR das colônias, contendo a sequência da proteína da construção RdRp_integral (1800 bp), clonas nos vetores pETSumo-1a/LIC, pETTrx-1a/LIC; e pETM11/LIC. Marcador Molecular Mid Range DNA Ladder Cellco Biotec.49
- Figura 7 Géis SD-PAGE 10% do teste de expressão realizado com a construção RdRp_integral. FI é a fração insolúvel e FS a fração solúvel. A seta em vermelho representa a proteína de interesse. (A) RdRp_integral fusionada com a cauda de 6-His (75 kDa), (b) RdRp_integral fusionada com a Trx (86 kDa)50
- Figura 8 Purificação RdRp_truncada. (A) Gel SDS PAGE 10%, etapas de purificação da construção RdRp_truncada *1) Marcador Molecular; 2) Fração solúvel; 3) Fração eluída na primeira cromatografia de afinidade; 4) Fração após a troca de tampão; 5) Digestão com TEV

	protease; 6) Fração não-ligada na segunda cromatografia de afinidade; 7) Fração eluída de segunda cromatografia de afinidade, 8) Fração eluída na cromatografia de exclusão molecular, 65 kDa, como mostra a seta vermelha. (B) Cromatograma da etapa de exclusão molecular mostrando a eluição da proteína em um pico único e simétrico.....	51
Figura 9	Cristais obtidos para a construção RdRp_truncada, a partir do kit de cristalização Morpheus, da Molecular Dimensions. As condições em que foi possível observar a formação de cristais que apresentaram difração foram: A) C6; B) D2; C) E2; D) F2; E) F6; F) G10 e G) H6.	53
Figura 10	Estrutura cristalográfica da NS5-RdRp, obtida a partir da construção RdRp_truncada. (A) Estrutura cristalográfica NS5 RdRp proveniente da construção RdRp_truncada. Em ciano claro, o subdomínio “dedos”. Em rosa claro o subdomínio “Palma”. Em salmão, o subdomínio “polegar”. Em magenta, mostramos o “priming loop”. As esferas violeta representam a posição dos átomos de Zn ²⁺ , e nos zooms podem ser visualizados os aminoácidos que coordenam esses Zn ²⁺ . (B) Proteína NS5 com seus domínios Metiltransferase (verde) e RNA Polimerase dependente de RNA, seus subdomínios mantem as cores mostradas na estrutura cristalográfica da Figura (A), e os motivos A à G estão mostrados em seus respectivos locais.	56
Figura 11	Estruturas cristalográficas do domínio RdRp de DENV cristalizadas com ligantes no sítio N-pocket. No quadro é possível observar a estrutura dos ligantes, nomeados de acordo com as estruturas às quais são correspondentes.....	58
Figura 12	Estruturas de NS5 domínio RdRp de ZIKV. Estruturas cristalográficas de ZIKV complexadas com ligantes na região N-pocket sobrepostas com a Estrutura 5U04. Ligantes presentes na região N-pocket, denominados de acordo com sua respectiva estrutura cristalográfica.....	59
Figura 13	Priming Loop das estruturas cristalográficas de ZIKV, DENV2 e DENV3. (A) Sobreposição das regiões do priming loop com os aminoácidos que sofrem substituições em cada uma das estruturas. (B) Priming loop da estrutura cristalográfica de ZIKV (5U04) e os seus respectivos aminoácidos substituídos em comparação com as demais, (C) priming loop de DENV2 (5K5M) e os seus respectivos aminoácidos substituídos; (D) priming loop de DENV3 (5HMW) e os seus respectivos aminoácidos substituídos. (E) Alinhamento da região do priming loop das sequências das estruturas cristalográficas de DENV2, DENV3 e ZIKV.	60
Figura 14	Purificação da RdRp_integral. (A)Gel de SDS-PAGE 12,5% com as etapas de purificação da RdRp_integral.1) Marcador molecular Bio Rad; 2) Fração Insolúvel; 3) Fração Solúvel; 4) Fração não ligada na primeira cromatografia de afinidade; 5) Fração eluída na primeira cromatografia de afinidade; 6) Digestão com TEV protease; 7) Fração não-ligada na segunda cromatografia de afinidade; 8) Fração eluída na cromatografia de exclusão molecular, 74kDa, como mostra a seta vermelha. (B) Cromatograma da etapa de exclusão molecular mostrando a eluição da proteína em um pico único e simétrico.	61

- Figura 15 Espectro obtido após o experimento de CD, onde temos um espectro para cada construção em vermelho temos a construção RdRp_truncada e em azul a construção RdRp_integral. Observamos a definição das bandas em 208 e 220 nm nas bandas em azul, que mostram a estabilidade da proteína em solução.63
- Figura 16 Esquema do ensaio de atividade. Esse ensaio registra a síntese de dsRNA, realizado pela RdRp (em ciano), a partir de um molde de +ssRNA (vermelho) e na presença de ATP. O SYBR Green I irá se intercalar ao dsRNA resultante, a partir de uma interação eletrostática entre o N⁺ (círculo vermelho) e o fosfato do RNA, formando um complexo SYBR Green I-dsRNA que irá emitir fluorescência.64
- Figura 17 Otimização das concentrações dos reagentes para a detecção da atividade da proteína NS5 domínio RdRp. Em (A) variou-se a concentração de proteína frente aos demais reagentes, após a padronização na concentração da proteína, em (B) foi realizada uma variação na concentração do RNA, mantendo fixo a concentração dos demais reagente. Em (C) e (D) segue-se o mesmo protocolo porem variando as concentrações de ATP e SYBr Green I, respectivamente. ...66
- Figura 18 Gráfico da atividade relativa pela posição de cada poço nas placas da PRB, pontos pretos se referem a poços com compostos enquanto os pontos azuis a controles positivos, já os pontos em vermelho representam as moléculas que diminuem a fluorescência da leitura da atividade da proteína em mais de 80%, a linha laranja mostra a atividade da proteína em 20% ou menos.68
- Figura 19 Gráfico da % de inibição dos compostos descobertos, frente ao tampão e na ausência da proteína. Dos 24 compostos testados observamos que 17 ainda interagem com o tampão, isso mostra que eles não inibem a atividade da proteína.70
- Figura 20 Curvas concentração-resposta para os cinco compostos do PRB, que tiveram IF abaixo de 20%, com suas respectivas estruturas moleculares. Sendo (A) o composto da placa A, referente ao poco A5, código PRB MMV1578560. (B) o composto da placa E, referente ao poco B9, MMV1782220; (C) o composto da placa B, referente ao poco C3, MMV1593540; (D) o composto da placa B, referente ao poco B9, MMV1593541 e (E) o composto da placa B, referente ao poco B4, o mais potente entre os compostos PRB encontrado, a clofazamina, MMV687800.71
- Figura 21 Esquema simplificado da síntese do composto MMV1593540.72
- Figura 22 Disposição da biblioteca de fragmentos e compostos do professor Marcio Dias, ICB-USP. Cada bloco é composto por 80 compostos/fragmentos e 16 controles, foram utilizados 8 controles de proteína, sem a presença de DMSO e 8 controles de proteína na presença de DMSO na mesma quantidade do composto a ser utilizado.75
- Figura 23 Média dos T_{ms} da proteína NS5 domínio RdRp_truncada na presença de DMSO como controle em cada um dos blocos testados. Em cada bloco foram realizadas 8 amostras da proteína com DMSO. As barras

	em azul mostram o valor médio dessas medidas e a barra em vermelho o desvio padrão assumido.	76
Figura 24	Análises realizada por blocos, onde estão apresentados os T_m de cada bloco com seu respectivo desvio padrão, em y estão o número de hits e em x, o ΔT_m ($^{\circ}C$) para os hits.	77
Figura 25	Gráficos do ensaio de inibição dos compostos que diminuíram a estabilidade da proteína pelo DSF. (A) Resultado dos cento e quarenta e três compostos a 200 μM frente a proteína, dos quais, somente nove foram capazes de diminuir a atividade da proteína em mais de 50 %. (B) Esses nove compostos foram submetidos a uma segunda rodada de ensaio de inibição, dessa vez a concentração dos mesmos foi de 20 μM , porém somente 3 deles inibiram levemente a atividade da proteína nessa concentração.	79
Figura 26	Resultado da triagem computacional e do ensaio enzimático da colaboração CIBFar. (A) Tabela com o resultado da melhor simulação realizada para os compostos Mangiferina e Nubbe_285. (B) Gráfico de barras da atividade enzimática da RdRp e a inibição da atividade dos compostos Nubbe_285 e Magiferina. (C) Imagem gerada pelo Pymol, onde é possível observar os resíduos que interagem com o composto Mangiferina (ciano), (D) gráfico de dose-resposta do composto Mangiferina com o IC_{50} de $6 \pm 2 \mu M$ (E) Imagem gerada pelo Pymol, onde é possível observar os resíduos que interagem com o composto Nubbe_285 (salmão), (F) gráfico de dose-resposta do composto Nubbe_285 com o IC_{50} de $10 \pm 2 \mu M$	82
Figura 27	Estruturas químicas dos compostos LabMol, para os quais foi possível determinar a atividade inibitória.	85
Figura 28	Resultado da triagem computacional e do ensaio enzimático da colaboração OpenZika, para os três compostos com $IC_{50} \geq 1 \mu M$. (A) Imagem gerada pelo Pymol, onde é possível observar os resíduos que interagem com o composto LabMol-301 (rosa) e (B) sua respectiva curva de concentração-resposta, com o IC_{50} de $1 \pm 0.1 \mu M$ (C) Imagem gerada pelo Pymol, onde é possível observar os resíduos que interagem com o composto LabMol-309 (ciano) e (D) sua respectiva curva de concentração-resposta, com o IC_{50} de $0.7 \pm 0.1 \mu M$ (E) Imagem gerada pelo Pymol, onde é possível observar os resíduos que interagem com o composto LabMol-343 (salmão) e (F) sua respectiva curva de concentração-resposta, com o IC_{50} de $0.48 \pm 0.09 \mu M$	87
Figura 29	Curvas de concentração-resposta dos compostos Nubbe_285 (A), LabMol-344 (B) e LabMol-343 (C). As estruturas químicas dos compostos são análogas e derivadas do anel flavonol.	88

LISTA DE TABELAS

Tabela 1	Pares senso-direto (Fw) e anti-senso (Rv) (5' → 3') dos oligonucleotídeos iniciadores e seus respectivos genes de interesse utilizados nas ampliações das ORF correspondentes as construções RdRp_truncada e RdRp_integral	39
Tabela 2	Tampões utilizados para a purificação de cada uma das construções de NS5 RdRp	42
Tabela 3	Condições do kit Morpheus onde foram possíveis o crescimento e a visualização de cristais da RdRp_truncada	52
Tabela 4	Parâmetros estatísticos relativos à coleta de dados e refinamento da NS5 RdRp de Zika vírus	55
Tabela 5	Propriedades físico-químicas das moléculas encontradas na triagem de fragmentos e compostos do LBEA.	80
Tabela 6	Resultado da colaboração OZ onde estão apresentados os valores de inibição IC ₅₀ (µM) e o valor da pontuação de docagem (kcal/mol).	85

LISTA DE ABREVIATURAS E SIGLAS

ATP	Adenosina trifosfato
6-His	Cauda de hexa-histidina
C	Capsídeo
CD	Dicroísmo Circular
DENV	Vírus da Dengue
DENV2	Vírus da Dengue sorotipo 2
DENV3	Vírus da Dengue sorotipo 3
DMSO	Dimetilsulfóxido
DNA	Ácido desoxirribonucleico
DO _{600 nm}	Densidade ótica em 600 nm
DSF	Fluorimetria Diferencial de Varredura
dsRNA	RNA de fita dupla
DTT	Ditiotreitol
E	Envelope
FBDD	Planejamento de Fármacos Baseado em Fragmentos
FZ	Febre Zika
HCV	Vírus da hepatite C
HTS	<i>High Throughput Screening</i>
IC ₅₀	Concentração inibitória média
IF	Inibição de Fluorescência
IFSC	Instituto de Física de São Carlos
IMAC	Cromatografia de afinidade por íon metálico imobilizado
IPTG	Isopropil β-D-1-tiogalactopiranosídeo
LabMol	Laboratório de Planejamento de Fármacos e Modelagem Molecular
LB	<i>Lisogeny broth</i>
LBDD	Planejamento de Fármacos Baseado na Estrutura dos Ligantes
LBEA	Laboratório de Biologia Estrutural Aplicada
LIC	Clonagem Independente de Ligase
LNLS	Laboratório Nacional de Luz Síncrotron
M	Membrana
MMV	<i>Medicines for Malaria Venture</i>

Mtase	Metiltransferase
NIs	Inibidores nucleosídeos
NLS	Sequência de localização nuclear funcional
NNIs	Inibidores não-nucleosídeos
NS	Não-estrutural
OZ	Projeto OpenZika
PAINS	<i>Pan Assay Interference Compounds</i>
PCR	Reação em cadeia da polimerase
PDB	<i>Protein Data Bank</i>
PEG	Polietilenoglicol
PMSF	Fluoreto de fenilmetilsulfonil
PRB	<i>Pandemic Response Box</i>
prM	Pré-membrana
RE	Reticulo endoplasmático
RdRp	RNA polimerase dependente de RNA
RNA	Ácido ribonucleico
RT-PCR	Reação da Transcriptase Reversa Seguida de Reação em Cadeia da Polimerase
SAM	S-adenosil-L-metionina
SAH	S-adenosil-L-hemocisteina
SBDD	Planejamento de Fármacos Baseado na Estrutura do Receptor
SCZ	Síndrome Congênita Zika
SDS-PAGE	Eletroforese em gel de poliacrilamida com dodecil-sulfato de sódio
SGB	Síndrome de Guillian-Barré
ssRNA	RNA de fita simples
TCEP	Tris 2-carboxietil-fosfina
TEV	<i>Tobacco Etch Virus</i>
Tm	Temperatura de ponto médio
TRX	Tiorredoxina
WNV	<i>West Nile Virus</i>
ZIKV	Vírus Zika

Sumário

1	Introdução	25
1.1	O vírus Zika	27
1.2	A Proteína Não Estrutural 5 (NS5).....	31
1.3	Busca por um tratamento anti-ZIKV.....	32
1.4	Busca por inibidores da atividade da NS5 RdRp do ZIKV	33
2	Objetivos	35
3	Materiais e Métodos	37
3.1	Planejamento das construções proteicas de RdRp	37
3.2	Clonagem do domínio RdRp em diferentes construções.....	38
3.3	Testes de expressão da NS5 domínio RdRp das construções RdRp_truncada e RdRp_integral	39
3.4	Expressão em larga escala da NS5 domínio RdRp construções RdRp_truncada e RdRp_integral	40
3.5	Purificação da NS5 domínio RdRp nas construções RdRp_truncada e RdRp_integral	40
3.6	Cristalização da NS5 RdRp de ZIKV e coleta de dados de difração de raios X.....	42
3.7	Processamento dos Dados Cristalográficos e Determinação da Estrutura 3D	43
3.8	Fluorimetria Diferencial de Varredura (DSF)	43
3.9	Dicroísmo Circular (do inglês Circular Dichroism - CD)	44
3.10	Ensaio de atividade da construção RdRp_integral	44
3.11	Ensaio de inibição da atividade RdRp	45
3.12	Triagem virtual	46
4	Resultados	49
4.1	Clonagem	49
4.2	Expressão.....	49
4.3	Purificação, cristalização e estrutura cristalográfica	51
4.3.1	Purificação, cristalização e determinação da estrutura cristalográfica da RdRp_truncada	51

4.3.2	Purificação e tentativas de cristalização da RdRp_integral	61
4.3.3	Análise da estabilidade da construção RdRp_integral em solução por Dicroísmo Circular (CD).....	62
4.3.4	Análise da atividade da construção RdRp_integral	63
4.4	Busca por ligantes.....	67
4.4.1	Ensaio de inibição em larga escala (<i>HTS – High Throughput screening</i>) ...	67
4.4.2	Triagem de ligantes a partir de técnicas biofísicas	73
4.4.3	Triagem de ligantes a partir de técnicas computacionais	80
4.4.3.1	Colaboração CIBFar	81
4.4.3.2	Colaboração OpenZika.....	83
5	Conclusão	89
	Referências	91

1 Introdução

O vírus Zika (ZIKV) é um flavivírus, pertencente à família *Flaviridae*, que possui mais de 70 espécies de vírus transmitidas por mosquitos, incluindo os vírus da febre amarela, dengue, encefalite japonesa e do Nilo Ocidental.¹ Os sintomas clínicos iniciais decorrentes da infecção pelo ZIKV são febre moderada, erupções cutâneas, artralgias, dor de cabeça e conjuntivite.²

O primeiro relato do ZIKV foi feito em 1947, na floresta Zika, em Uganda, África.³ O vírus foi encontrado no sangue de um macaco Rhesus e, até sua disseminação pelas ilhas do Pacífico, poucos casos de febre Zika foram documentados, sendo esses raros e geograficamente isolados.⁴

Após a identificação inicial do vírus, o primeiro grande surto ocorreu somente em 2007, na ilha de Yap, na Micronésia, onde 73% da população foi infectada, embora sem relatos de mortes ou hospitalizações. Porém em 2013/2014, esse quadro mudou quando, durante um surto na Polinésia Francesa, o ZIKV infectou mais de 30.000 pessoas, levando algumas dessas a serem hospitalizadas.⁵⁻⁶ Durante essa epidemia, foram relatados 42 casos da Síndrome de *Guillian-Barré* (SGB), uma doença que acomete os nervos periféricos dos membros superiores e inferiores, levando à fraqueza progressiva, dificuldade na movimentação dos membros, e, potencialmente, à paralisia das pernas. Estudos sugeriram que esses casos poderiam estar associados ao ZIKV.⁷⁻⁹

O ZIKV é transmitido por mosquitos do gênero *Aedes*, em especial o *Aedes aegypti*, razão pela qual é classificado como um arbovírus (***arthropod-bourne virus***). Além da transmissão por mosquitos, já foram confirmadas transmissões de ZIKV por relação sexual (uma vez que o ZIKV foi encontrado no sêmen assim como na urina), por transfusão sanguínea, doação de órgãos e de mães para prole durante a gestação.¹⁰

O Brasil foi o primeiro país da América do Sul a detectar a presença de infecções ocasionadas por ZIKV, em abril de 2015,¹¹ fazendo uso da técnica de Reação da Transcriptase Reversa Seguida de Reação em Cadeia da Polimerase (RT-PCR) em amostras sorológicas de pacientes do estado da Bahia.¹² Além dos sintomas normais do ZIKV, o número de casos de SGB aumentou 20 vezes.¹³ Um mapa com a disseminação global da doença pode ser observado na Figura 1.

conjunto de sintomas a OMS chamou de Síndrome Congênita Zika (SCZ). Em alguns desses casos, o dano cerebral apresentou-se tão significativo que o crânio foi preenchido com fluído em sua maior parte, diminuindo assim o tecido cerebral.

Médicos no Panamá e na Colômbia relataram alguns casos de mães que apresentaram os sintomas do ZIKV durante a gestação e os bebês nasceram com problemas cardíacos, no trato gastrointestinal e nos órgãos genitais.¹⁷⁻¹⁸

Assim, em 1 de fevereiro de 2016, devido ao surto do ZIKV e a sua associação à complicações neurológicas e à microcefalia no Brasil (e em vários outros países, principalmente da América do Sul), a OMS declarou a FZ como emergência de saúde pública de preocupação internacional.¹⁹

Dados atuais do Ministério da Saúde, mostram que mais de 240.000 casos de FZ foram notificados no país desde 2015, sendo 10.768 casos prováveis registrados no ano de 2019 e 7.119 casos prováveis no ano de 2020.²⁰⁻²¹ Além disso, em relação ao número de casos suspeitos de SCZ de 2015 a 2020, foram notificados 19.492, dos quais 3.563 foram confirmados e 2.913 permanecem em investigação.²²

Apesar da diminuição do número de casos desde a descoberta da FZ, uma nova linhagem do ZIKV foi recentemente detectada em circulação no Brasil.²³ Esses achados mostram que o vírus continua circulando no país e está sofrendo mudanças para adaptação aos seus hospedeiros. Assim, apesar da gravidade causada pela infecção por ZIKV, nenhum tratamento profilático e / ou terapêutico específico está disponível.

1.1 O vírus Zika

Estudos de análises filogenéticas envolvendo sequências de genoma completo e genes específicos identificaram três linhagens virais: asiática, africana oriental e africana ocidental.^{11,24-25} As análises mostraram que a linhagem africana continua predominantemente em seu ciclo silvestre, envolvendo principalmente macacos e mosquitos. Já a linhagem asiática espalhou-se pelas ilhas do Pacífico e mais recentemente na América do Sul, Central, México e Porto Rico, infectando humanos.^{13,26}

Estudos com base nos genomas do ZIKV circulante no Brasil no início da epidemia causada pelo ZIKV, demonstraram que as cepas virais brasileiras

possuem a mesma origem genética que a linhagem asiática, pois compartilham mais de 99,7% de identidade sequencial de RNA e 99,9% de similaridade sequencial de aminoácidos com o ZIKV relatado na Polinésia Francesa em 2013.²⁷

Em 2019, um grupo da Fiocruz identificou a circulação de linhagens africanas, até então inexistentes no Brasil, em duas regiões brasileiras – no Sul e no Sudeste – além de indícios de que essas linhagens já estão circulantes no Brasil há algum tempo, devido à distância geográfica e à diferença entre os hospedeiros (uma foi encontrada em mosquito e a outra, em uma espécie de macaco).^{23,28-29}

O ZIKV é formado por partículas esféricas com capsídeo icosaédrico e envelopadas, com aproximadamente 50 nm de diâmetro.³⁰⁻³¹ Possui um material genético constituído de RNA de fita simples de cadeia positiva, com aproximadamente 11 kb, que codifica por uma única poliproteína que é clivada em 10 proteínas virais: três proteínas estruturais - capsídeo (C) , envelope (E) e a pré-membrana (prM ou M) - e sete proteínas não estruturais (NS), NS1, NS2A, NS2B, NS3, NS4A, NS4B e NS5.³²⁻³⁴ A Figura 2 mostra um esquema da partícula viral e do genoma do ZIKV.

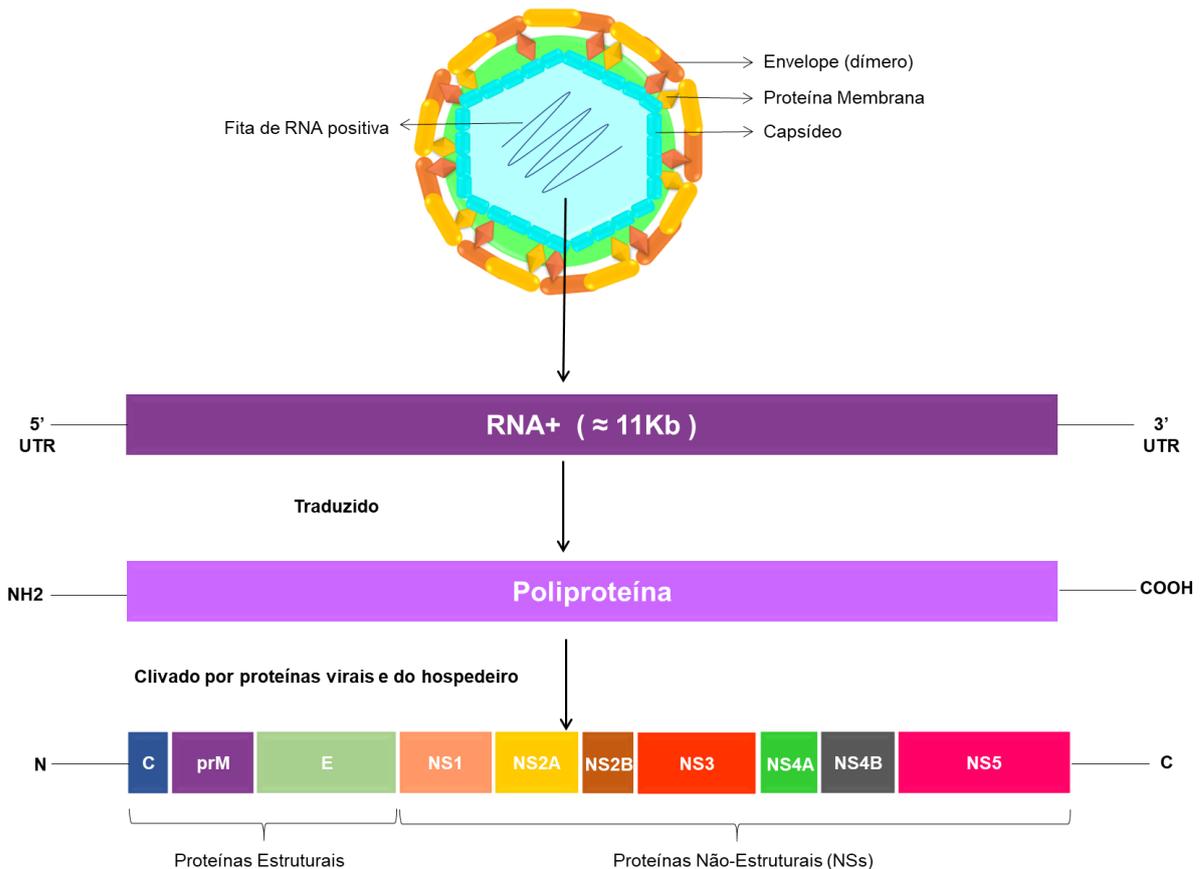


Figura 2 A partícula viral do ZIKV possui um material genético constituído por uma fita simples de cadeia positiva de RNA, com aproximadamente 11 Kb. Esta fita é traduzida como uma única poliproteína e após maturação é clivada por proteases virais e do hospedeiro em 3 proteínas estruturais e 7 proteínas não estruturais as NSs.

Fonte: Elaborada pela autora; Adaptada de SINGH, R. K. *et al.* ³⁵

O ciclo de replicação dos flavivírus ocorre nas células hospedeiras por meio de endocitose mediada por receptores. Após a endocitose, a proteína E sofre mudanças conformacionais, que são induzidas pela diminuição do pH vesicular, resultando na fusão da membrana viral com a membrana do hospedeiro, o que leva à liberação do genoma que está ligado ao capsídeo. O RNA viral de sentido positivo (ssRNA+) é liberado no citoplasma e transportado para a superfície do retículo endoplasmático (RE). Esse ssRNA+ é traduzido nos ribossomos para formar a poliproteína viral, a qual depois de clivada, permite a formação do complexo de replicação composto pelas proteínas não-estruturais, NS3, NS5, NS4B, NS4A, NS2B, NS2A. Este complexo de replicação se liga ao ssRNA+ para, numa primeira passagem, polimerizar a fita negativa complementar ao genoma viral, assim resultando em uma fita dupla de RNA (dsRNA), a forma replicativa. Esta molécula de dsRNA passa a ser o molde com o qual o complexo de replicação passa a produzir muitas cópias do genoma viral de ssRNA+.

Estas novas cópias do genoma viral repetem amplamente o ciclo de tradução ribossômica, acumulando em grande quantidade dentro da célula infectada as proteínas virais estruturais e não estruturais. Após sucessivos ciclos de transcrição e a tradução, as múltiplas cópias do ssRNA+ viral são empacotadas juntamente com as proteínas estruturais C, E e prM em virions imaturos. No trajeto das partículas virais imaturas do seu locus de formação no complexo de Golgi até a membrana externa, a proteína E passa por mudanças conformacionais e a prM passará por clivagem, mediada por proteases endógenas da célula, formando assim a proteína M madura e a partícula viral pronta para eclosão. As partículas virais maduras serão liberadas da célula hospedeira por exocitose.^{36–39} Um esquema do ciclo de replicação viral no hospedeiro pode ser observado na Figura 3.

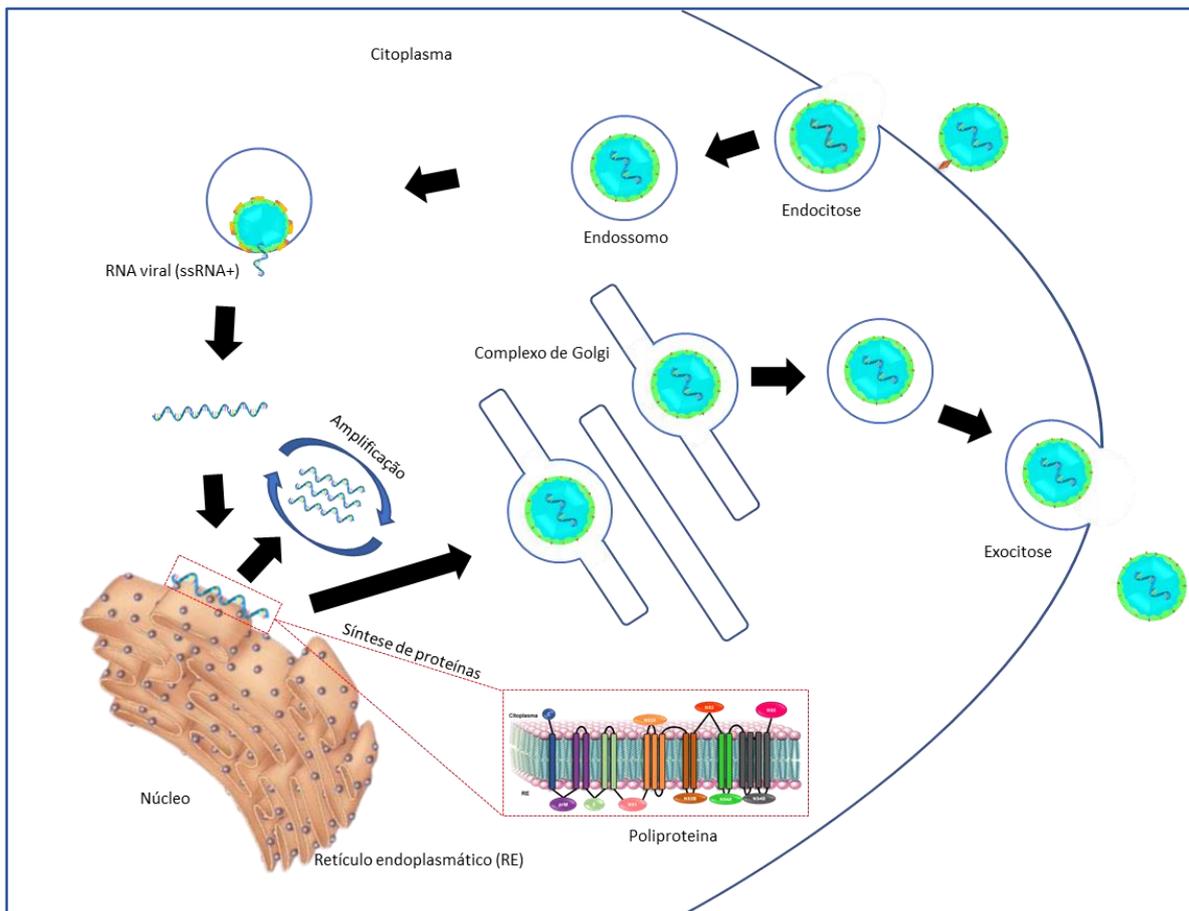


Figura 3 Esquema do ciclo de replicação do vírus da Zika. A proteína E é responsável pela ligação do vírus aos receptores de membrana do hospedeiro, levando o vírus a entrar na célula por endocitose. A redução no pH endossomal induz a mudança conformacional na proteína E que permite a fusão das membranas do vírus às do hospedeiro. O RNA viral é então liberado no citoplasma. O RNA de fita simples (ssRNA+) é traduzido em uma poliproteína, que é então clivada em várias proteínas virais. Os genomas virais se replicam na superfície do retículo endoplasmático (RE). No complexo de Golgi, a proteína prM é clivada, induzindo a maturação dos virions. Por fim a partícula viral madura sai da célula por exocitose.

Fonte: Elaborada pela autora; Adaptada de KATO, F *et al.*³⁸

Como foi possível observar, as proteínas NSs associam-se para formar o complexo de replicação, que se ancora nas estruturas membranosas do retículo endoplasmático da célula hospedeira, sendo o responsável pela replicação do genoma viral. As NSs também contribuem para a montagem da partícula viral e da evasão da imunidade inata da célula hospedeira, mostrando-se, portanto, como alvos para estudos de desenvolvimento de fármacos antivirais.⁴⁰

Dentre as proteínas não estruturais, existem duas com funções enzimáticas e que não estão diretamente inseridas na membrana do retículo endoplasmático ou às regiões transmembranares e despertam especial interesse: a NS3 e a NS5.⁴¹

A NS3 possui um domínio N-terminal de função protease, que tem como cofator a NS2B, formando o complexo NS2B-NS3 e um domínio C-terminal, de função helicase.⁴²⁻⁴³

1.2 A Proteína Não Estrutural 5 (NS5)

A proteína NS5 do ZIKV é a maior (900 kDa) e a mais conservada dentre as proteínas não estruturais do ZIKV.⁴⁴⁻⁴⁵ Ela é constituída por dois domínios proteicos: um domínio N-terminal Metiltransferase (Mtase) e um domínio C-terminal com atividade RNA polimerase dependente de RNA (RdRp).⁴⁶

O domínio Mtase é responsável por gerar um *cap* na extremidade 5' do RNA genômico viral recém sintetizado, que promove a tradução, além de proteger o RNA viral da degradação pela resposta imune inata celular.⁴⁷

O Mtase possui um sítio de ligação específico à ligação de S-adenosil-L-metionina (SAM), cofator essencial para a catalise da metilação da guanina na posição N-7, bem como da ribose na posição 2'-O. Após o processo de metilação, a S-adenosil-L-hemocisteína (SAH) é gerada como um subproduto da reação.^{39,48-49}

O domínio RdRp apresenta-se com uma arquitetura que mostra sua estrutura em forma classicamente representada como uma mão direita, com três principais subdomínios que a constituem, que são: os dedos, a palma e o polegar. Esses três subdomínios formam três canais que possibilitam a entrada do RNA molde, a entrada de ribonucleotídeos trifosfatos (NTPs) e a saída do RNA recém-sintetizado.⁵⁰⁻⁵²

Assim, o domínio RdRp é responsável pela transcrição do genoma viral, e síntese de novas fitas de RNA. Esse processo ocorre em três etapas principais, que são: a iniciação *de novo*, que começa a síntese de uma nova fita de RNA na ausência de um *primer*; a transcrição da nova fita de RNA com a incorporação de nucleotídeos complementares ao molde sintetizado, prosseguindo na fase de alongamento da polimerização.⁵³⁻⁵⁴

Além de exercer sua função de polimerase, o domínio RdRp da NS5 pode interagir com componentes de RNA e com outras proteínas do complexo de replicação responsáveis pelas várias etapas necessárias à replicação do RNA viral.⁴⁶

1.3 Busca por um tratamento anti-ZIKV

Tendo em vista o grande impacto que a epidemia do ZIKV gerou no Brasil e no mundo, especialmente em mulheres grávidas, há um grande esforço da comunidade científica na busca por recursos terapêuticos para a profilaxia e tratamento da infecção viral.

No que tange à prevenção, o objetivo maior é o desenvolvimento de uma vacina eficaz e segura.⁵⁵⁻⁵⁶ Entretanto, esse desenvolvimento tem encontrado algumas dificuldades, visto que a população-alvo durante o surto são mulheres grávidas ou em idade fértil e homens com possibilidade de transmissão sexual do vírus. Além disso, o surto da doença foi breve, o que dificulta a testagem e avaliação das vacinas em estudos clínicos controlados de larga escala. Assim, não é possível prever a eficácia e a durabilidade da imunidade induzida pela vacina, bem como o número de doses necessárias para induzir a imunização, além do impacto dos anticorpos produzidos com reatividade cruzada para outros flavivírus.⁵⁷⁻⁶⁰

Dessa forma, estudos paralelos na busca por antivirais se fazem necessários, visto que, até o momento, não há nada que possa prevenir os danos fetais causados pela infecção de ZIKV durante a gravidez. Assim, um medicamento não teratogênico, de baixo ou nenhum risco para o feto seria um recurso fundamental para o tratamento de mulheres grávidas eventualmente infectadas.⁴

Diferentes abordagens e metodologias vêm sendo estudadas e utilizadas, tais como: o reposicionamento de medicamentos já aprovados pela *Food and Drug*

Administration (FDA),⁶¹⁻⁶⁶ produtos naturais, ⁶⁷⁻⁷⁰ triagem de diferentes bibliotecas comerciais de moléculas bioativas, triagem fenotípica e triagem computacional para descoberta de novos candidatos a fármacos, ⁷¹⁻⁷⁵ porém, até o momento, não existem estudos clínicos sendo realizados sobre a segurança e eficácia de compostos para o tratamento de infecções pelo vírus Zika.

No que se refere à busca por antivirais, as estratégias baseiam-se na busca por inibidores do vírus ou inibidores que atuam em proteínas do hospedeiro que são cooptadas pelo vírus durante a infecção celular, implicando assim nas diferentes etapas do ciclo de vida viral, desde a ligação, entrada e fusão, até a formação do complexo de replicação, reprodução do genoma e das proteínas estruturais do vírus e o subsequente encapsulamento, maturação e eclosão viral.⁷⁶⁻⁷⁷ Compostos que agem contra proteínas específicas do vírus, como a proteína de envelope viral ou as proteínas não-estruturais, são denominadas de Antivirais de Ação Direta (DAA, do inglês *Direct-Acting Antivirals*).^{37,77}

1.4 Busca por inibidores da atividade da NS5 RdRp do ZIKV

A NS5 apresenta-se altamente conservada entre os flavivírus quando se analisa o sítio catalítico. O domínio RdRp apresenta-se como um alvo atraente para o desenvolvimento de fármacos antivirais porque, dentre outros fatores, é ausente de similares em células humanas, o que contribui para a busca por um antiviral para essa proteína que seja específico e portanto mais seguro.^{62,78}

Foram descritas duas classes de inibidores virais de RdRp: os inibidores nucleosídeos (NIs) e os inibidores não-nucleosídeos (NNIs). Os NIs são intracelularmente convertidos em uma forma trifosfato e podem competir com os substratos naturais de nucleosídeos trifosfatados (NTP), impedindo que o alongamento da cadeia de RNA seja concluído. Já os NNIs se ligam ao sítio ativo ou a sítios alostéricos da RdRp, inativando a enzima ou evitando a ligação de substratos ou alterações conformacionais necessárias para que ocorra a síntese de RNA.⁷⁹

Estudos relacionados à busca de inibidores para a NS5 mostraram-se promissores, visto que já existe uma classe de inibidores de RNA polimerase descritos na literatura, como por exemplo o medicamento de uso clínico Sofosbuvir

(Sofaldi[®]), cujo o alvo é o domínio RdRp da proteína NS5 do vírus da hepatite C (HCV).^{33,71}

O Sofosbuvir é um inibidor do tipo NI, análogo de uridina, tratando-se de um pró-fármaco metabolizado dentro do hepatócito por enzimas endógenas, que levam à formação do composto ativo fosforilado, o 2'-C-metil-2'-fluoro-uridina-5'-trifosfato, (Figura 4). Este composto se liga ao sítio ativo da RdRp, competindo com os substratos naturais para a síntese de RNA,⁸⁰⁻⁸¹ Estudos computacionais e celulares mostram que há possibilidades desse medicamento se ligar ao domínio RdRp de outros flavivírus.^{62,82}



Figura 4 O Sofosbuvir é um pró-fármaco, usado no tratamento de HCV, cujo alvo molecular é a RdRp. A metabolização intracelular produz sua molécula ativa, que atua no sítio de síntese de RNA, competindo com os substratos naturais e impedindo a síntese de RNA.

Fonte: Elaborada pela autora.

Esses estudos demonstram que a RdRp é um alvo válido e apontam para a importância da busca por novos candidatos a fármacos anti-ZIKV a partir de potenciais inibidores da atividade da RdRp.

2 Objetivos

O objetivo central deste trabalho foi buscar por inibidores específicos contra a proteína NS5 RNA polimerase dependente de RNA do vírus Zika. Para tal, foi necessária a obtenção e a elucidação da estrutura cristalográfica do domínio RdRp. A estrutura da proteína nos auxilia na busca por possíveis ligantes e inibidores a partir de técnicas bioquímicas, biofísicas e computacionais. Para atingir esse objetivo central, os objetivos específicos do presente projeto são:

- Clonagem, expressão e purificação da proteína NS5 domínio RdRp de ZIKV em diferentes construções, utilizando sistemas de expressão na bactéria *Escherichia coli*;
- Cristalização e determinação da estrutura cristalográfica da proteína NS5 domínio RdRp de ZIKV;
- Estabelecimento de um ensaio de atividade enzimática, para avaliar a inibição de potenciais compostos inibidores em triagens de larga escala;
- Triagem de compostos por técnicas bioquímicas, biofísicas e computacionais;

3 Materiais e Métodos

3.1 Planejamento das construções proteicas de RdRp

Baseado no artigo de Malet *et al.*⁸³ foram planejadas duas construções proteicas para a RdRp de ZIKV. Estes autores desenharam e cristalizaram duas construções proteicas para o domínio RdRp do flavivírus *West Nile Virus* (WNV), como mostra a Figura 5. Uma é a forma ativa, a qual foi denominada POL1 (resíduos 273-905; 3,0 Å) e a outra é forma inativa, denominada POL2 (resíduos 317-905; 2,3 Å). Após a cristalização o autor compara as estruturas e conclui que ambas possuem as mesmas orientações de resíduos dentro do sítio ativo, porém a resolução estrutural obtida para POL2 é mais alta.

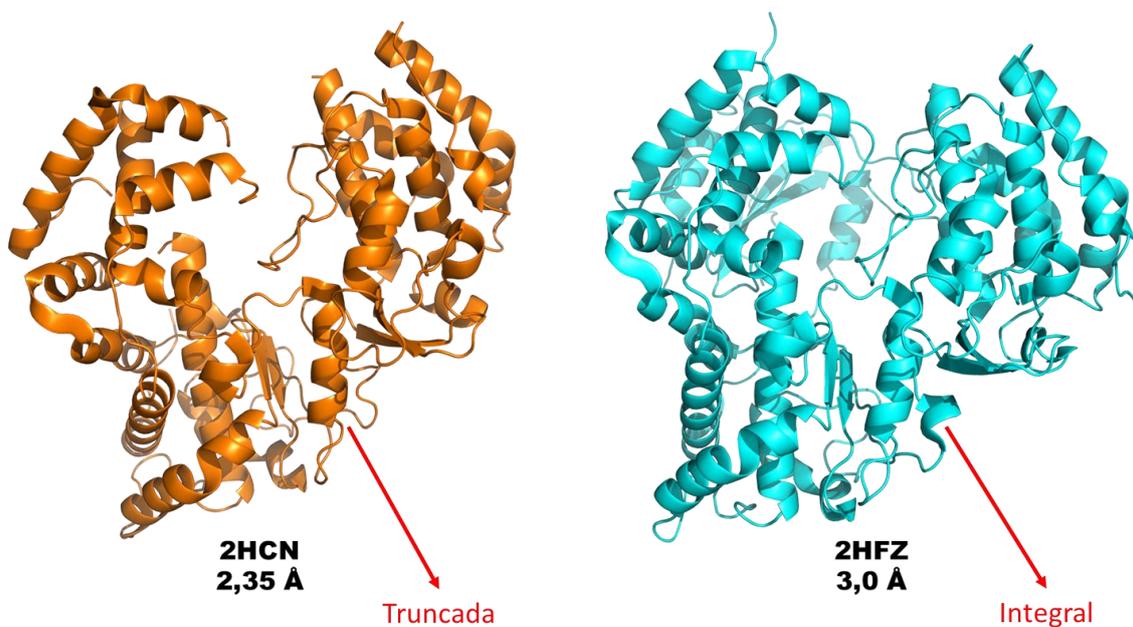


Figura 5 Malet e colaboradores propõem duas construções para o domínio RdRp de WNV, uma integral e ativa (2HFZ) que ao ser cristalizada apresenta média resolução (3,0 Å) – construção em azul; e uma construção truncada (2HCN), que possui as mesmas orientações dos aminoácidos do sítio ativo e conformações, porém inativa, que ao ser difratada apresenta mais alta resolução (1,8 Å) – construção laranja.

Fonte: Elaborada pela autora.

Devido a esse indicativo da literatura, as duas construções planejadas para a proteína NS5 domínio RdRp de ZIKV foram:

- RdRp_truncada, baseada na POL2, que foi truncada em seus resíduos, sendo esta a forma inativa, resíduos 306 a 904.
- RdRp_integral, sendo esta a forma ativa, resíduos 276 a 888.

3.2 Clonagem do domínio RdRp em diferentes construções

A clonagem das sequências codificantes das proteínas de Zika vírus iniciou-se com a aquisição de genes sintéticos.

As sequências do DNA que codificam a proteína NS5 foram obtidas no banco de dados GenBank (AY632535.2, proveniente da cepa ancestral do macaco *Rhesus* isolada em Uganda na África em 1947 e ALU33341.1, proveniente de uma cepa humana, isolada no Brasil em março de 2015). As sequências de DNA codificante, produzidos por encomenda pela empresa GenScript, foram desenhados com códons que otimizam a expressão em *E.coli*. Baseado nessas sequências, os respectivos pares de oligonucleotídeos entre as bases 7655-10363 foram planejados para a amplificação e sintetizados pela empresa Exxtend (<https://www.exxtend.com.br/>).

A partir das sequências amplificadas foram obtidas duas construções: a construção proveniente da cepa ancestral do macaco *Rhesus*, que é a construção MR766_RdRp⁽³⁰⁶⁻⁹⁰⁴⁾, neste trabalho denominada como RdRp_truncada. E a construção proveniente da cepa humana, que é a construção Alu33341_RdRp⁽²⁷⁶⁻⁸⁹⁸⁾, denominada neste trabalho como RdRp_integral.

A clonagem da RdRp_truncada foi realizada em colaboração com a empresa Cellco Biotec, uma startup à época incubada em nosso laboratório, fazendo o uso dos plasmídeos pETM-11/LIC e pETTrx-1a/LIC. Já a clonagem da RdRp_integral foi realizada integralmente pela autora deste trabalho, utilizando os plasmídeos pETM11/LIC e pETTrx-1a/LIC.

As sequências de DNA codificante utilizadas para cada uma das construções estão representadas na Tabela 1. Ao planejar os oligonucleotídeos iniciadores (*primers*) para a reação, incluímos a sequência 5'-CAGGGCGCCATG-3' na extremidade 5' de cada *primer* senso-direto e a sequência 5'-GACCCGACGCGGTTA-3' na extremidade 5' de cada *primer* anti-senso, como representado no esquema abaixo:

- Senso-direto (Fw) (5' → 3') CAGGGCGCCATG-gene de interesse
- Anti-senso (Rv) (5' → 3') GACCCGACGCGGTTA-gene de interesse

Tabela 1 Pares senso-direto (Fw) e anti-senso (Rv) (5' → 3') dos oligonucleotídeos iniciadores e seus respectivos genes de interesse utilizados nas ampliações das ORF correspondentes as construções RdRp_truncada e RdRp_integral

Construção proteica	Oligonucleotídeos
RdRp_truncada	CAGGGCGCCATGTACCATGGGAGCTACGAAGC (Fw) GACCCGACGCGGTTACAACACTCCGGGTGTGG (Rv)
RdRp_integral	CAGGGCGCCATGAAAATTATTGGCAATCGTATTGAG (Fw) GACCCGACGCGGTTAACTACCTTCTTCGCCAGATAG (Rv)

Fonte: Elaborada pela autora.

Todas as clonagens foram realizadas fazendo o uso do sistema LIC - “Clonagem Independente de Ligase” (LIC, do inglês, *Ligase Independent Cloning*), que baseia-se na atividade exonuclease da enzima T4 DNA polimerase (5' → 3'), que gera extremidades coesivas nos vetores e no inserto. Essas extremidades serão complementares, assim os fragmentos irão se unir em uma reação de hibridização, sem a necessidade de uma Ligase, conforme descrito por Aslanidis e Jong.⁸⁴

3.3 Testes de expressão da NS5 domínio RdRp das construções RdRp_truncada e RdRp_integral

Protocolos distintos foram empregados para analisar as condições de expressão mais apropriadas. Em ambos os protocolos o pré-inóculo foi obtido a partir de uma mesma colônia de *Escherichia coli Rosetta 2* (DE3), contendo os plasmídeos de expressão clonados, após a clonagem e crescimento em placa, em meio LB (*Lisogeny broth*) suplementado com 50 µg/ml de Canamicina e 34 µg/ml de Cloranfenicol por 16 h a 37 °C, à 150 rpm de agitação.

Protocolo 1 - A cultura foi inoculada em meio LB até atingir a densidade ótica em 600 nm (DO₆₀₀) entre 0,6-0,8. Em seguida, a expressão proteica foi induzida pela adição de 1 mM de Isopropil β-D-1-tiogalactopirosídeo (IPTG), diminui-se a temperatura para 18 °C e a cultura bacteriana foi mantida a 150 rpm de agitação por 20 h.

Protocolo 2 - A cultura foi inoculada em meio auto-indutor ZYM-505227,⁸⁵ que foi mantido sob agitação constante de 175 rpm a 37 °C e a DO₆₀₀ foi acompanhada até atingir 0,6 e 0,8. A expressão proteica foi induzida pela diminuição da temperatura para 18 °C e a cultura foi mantida a 150 rpm de agitação por 20 h.

Após concluído o período de expressão de 12 h, amostras de cultura foram retiradas a cada 2 h. As células foram centrifugadas e lisadas (como descrito na seção 3.5). Para avaliar a expressão da NS5 RdRp, o lisado obtido foi centrifugado e as frações solúvel e insolúvel foram submetidas à análise por eletroforese em gel de poliacrilamida (SDS-PAGE 10%), corado com *Comassie Brilliant Blue*.

3.4 Expressão em larga escala da NS5 domínio RdRp construções RdRp_truncada e RdRp_integral

A condição de expressão otimizada foi: a partir de uma amostra de um estoque glicerinado de cultura mantido a -80 °C, contendo células de expressão transformadas com os vetores de expressão citados na seção 3.1, respectivamente. Um pré-inóculo foi cultivado em meio LB, suplementado com 50 µg/ml de Canamicina e 34 µg/ml de Cloranfenicol por 16 h a 37 °C. Para cada 1 L de meio auto-indutor ZYM-5052, foram utilizados 10 mL de pré-inóculo. A cultura inoculada foi mantida sob agitação constante de 175 rpm a 37 °C até atingir a DO₆₀₀ entre 0,2 e 0,6. Após atingida a DO₆₀₀ necessária, a temperatura foi reduzida a 18 °C e mantida sob agitação constante por 20 h.

Ao final da expressão, as células foram coletadas por centrifugação (4000g, 40 min, 4 °C), ressuspendidas em 40 mL de tampão de lise (50 mM Tris pH 8,5; 100 mM NaCl; 20 mM Imidazol e 10% Glicerol), congeladas e estocadas a -20 °C.

3.5 Purificação da NS5 domínio RdRp nas construções RdRp_truncada e RdRp_integral

O mesmo protocolo foi utilizado para todas as construções, havendo modificação apenas dos tampões C. Dessa forma, um procedimento geral é descrito a seguir e os tampões utilizados estão listados na Tabela 2.

Para a lise da membrana celular, o pellet de bactérias foi mantido em banho de gelo durante 30 min, na presença de 0,5% de Triton X-100; 0,5 mg/mL de lisozima, 10 U/mL de Turbo Nuclease (Cellco Biotec), 4 mM de Ditioneitol (DTT) e 1 mM fluoreto de fenilmetilsulfonil (PMSF).

Após esse período, as células foram submetidas ao processo de sonicação nas condições 40 Hz, 15 min, 30 s *on* e 1 min *off*. A suspensão resultante foi

submetida à centrifugação (12900 rpm, 40 min, 4 °C), para a separação das frações solúvel (extrato proteico) e insolúvel (restos celulares).

Para isolar a proteína-alvo do extrato proteico, a fração solúvel foi submetida à primeira etapa de cromatografia de afinidade por íon metálico imobilizado (do inglês, IMAC), em uma coluna HisTrap HP 5 mL com resina Ni Sepharose (GE Healthcare), pré-equilibrada com Tampão A. Após as etapas de lavagem da coluna, a proteína fusionada à proteína de fusão 6-His-Trx foi eluída utilizando Tampão B. A clivagem da proteína de fusão e a troca de tampão (Tampão B para Tampão A) foram realizadas de forma simultânea, a fim de permitir a ação da enzima proteolítica de *Tobacco Etch Virus* (TEV protease). Essa troca foi realizada por meio de diálise por 16 h a 4 ° C na presença de TEV protease (proporção 1 mg de TEV protease para 20 mg de proteína de interesse) e 4 mM de DTT. Nessas condições, a TEV protease reconhece seu sítio de clivagem e separa a proteína de interesse da proteína de fusão ligada à cauda de histidina (6-His-Trx). O isolamento da proteína de interesse das proteínas remanescentes foi efetuado utilizando-se um segundo passo de IMAC na coluna HisTrap HP, com resina Ni Sepharose 5 mL, equilibrada em Tampão A.

Para garantir a homogeneidade do estado oligomérico e conformacional, realizou-se um passo final de cromatografia de exclusão molecular em uma coluna *Hiload Superdex 75 pg 16/60*. Nesta etapa foi utilizado o Tampão C.

A pureza da amostra e seu progresso durante as etapas de purificação foram acompanhados por eletroforese em gel de poliacrilamida 10% (SDS-Page), como descrito por Laemmli.⁸⁶

Tabela 2 Tampões utilizados para a purificação de cada uma das construções de NS5 RdRp

Construção	Tampão A	Tampão B	Tampão C
RdRp_truncada	50 mM de Tris pH 9,0; 500 mM de NaCl; 20 mM de Imidazol; 10% de glicerol	50 mM de Tris pH 9,0; 500 mM de NaCl; 500 mM de Imidazol; 10% de glicerol	20 mM de HEPES pH 7,5; 200 mM NaCl; 5% de glicerol
RdRp_integral	50 mM de Tris pH 9,0; 500 mM de NaCl; 20 mM de Imidazol; 10% de glicerol	50 mM de Tris pH 9,0; 500 mM de NaCl; 500 mM de Imidazol; 10% de glicerol	20 mM de Tris pH 7,5; 200 mM NaCl, 5% de glicerol; 0,5 mM de TCEP

Fonte: Elaborada pela autora.

3.6 Cristalização da NS5 RdRp de ZIKV e coleta de dados de difração de raios X

Os ensaios de cristalização da proteína NS5-RdRp foram realizados com proteína nas concentrações de 5, 10 e 20 mg/mL. Uma triagem inicial das condições de cristalização foi realizada com auxílio do robô de cristalização *Crystal Gryphon* (*Art Robbins Instruments*), por meio do método de difusão de vapor em gota sentada (do inglês, *sitting drop*), utilizando kits comerciais das seguintes empresas *Molecular Dimensions* (Morpheus, PACT), *Hampton Research* (Index HT, Crystal Screen), *Qiagen* (PEGs I, PEGs II, MPD Suite e JCSG Core Suite I, II, III e IV). Cada gota era constituída de uma solução 1:1 de 200 nL de proteína e solução do kit, equilibrada por difusão de vapor com 60 µL de solução do poço (kit). As caixinhas com os experimentos de cristalização foram mantidas nas temperaturas de 4 e 18 °C. O experimento foi acompanhado por observação em microscópio em dias alternados, até que cristais fossem observados.

Os cristais obtidos durante os experimentos de triagem de condições foram retirados de suas gotas utilizando uma alça de nylon (*Hampton Research*) e transferidos para uma solução crio-protetora contendo o tampão de cristalização da respectiva gota de cristalização e 10% de um agente crio-protetor (por exemplo, etilenoglicol, glicerol, PEG 200, entre outros).

A coleta de dados cristalográficos foi realizada na linha MX-2 do Laboratório Nacional de Luz Síncrotron (LNLS) – Campinas/SP, onde cristais foram submetidos à difração de raios X, utilizando um detector PILATUS 2M.

3.7 Processamento dos Dados Cristalográficos e Determinação da Estrutura 3D

Os dados de difração foram processados, indexados e integrados por meio do pacote XDS,⁸⁷ que converte as imagens bidimensionais obtidas durante a coleta de dados em um arquivo de intensidades (e.g., XDS_ASCII.HKL) e reduzidos com o auxílio do programa *Aimless*, disponível no pacote de programas CCP4.⁸⁸

A determinação das fases iniciais foi realizada com a utilização da estrutura da proteína homóloga 2HCN do vírus de *West Nile*,⁸³ depositada no banco de estruturas proteicas PDB (do inglês, *Protein Data Bank*), pois essa possui uma sequência com identidade de 72%. Para tanto, foi utilizado o método de substituição molecular com o programa *Phaser*.⁸⁹

O refinamento da estrutura no espaço recíproco foi realizado com as rotinas do pacote de programas *Phenix*,^{90–92} enquanto o refinamento no espaço real e a construção dos modelos 3D foram realizados com o programa *Coot*.⁹³

A validação dos modelos finais ocorreu pela avaliação dos parâmetros estatísticos desde o processamento até o refinamento, bem como pelos parâmetros estereoquímicos e geométricos com o programa *MolProbity*.⁹⁴

3.8 Fluorimetria Diferencial de Varredura (DSF)

A investigação da estabilidade térmica da construção RdRp_truncada foi realizada com a técnica de DSF. A mesma técnica foi utilizada para a triagem da biblioteca de fragmentos e compostos, disponível no Laboratório de Biologia Estrutural Aplicada (LBEA) do Instituto de Ciências Biomédicas da USP (ICB-USP), gentilmente cedida pelo Prof. Dr. Marcio Dias.

O experimento foi composto por uma solução de 25 μL , contendo 35 μM da construção RdRp_truncada, sonda fluorescente Sypro® Orange (Sigma- Aldrich) 5x concentrada e 2 μL de cada composto a ser analisado, que se encontravam na concentração de 5-8 mM diluídos em DMSO. A amostra controle consistiu em solução de ensaio e 2 μL de DMSO 100%, que não ultrapassou a concentração final de 8%.

Os ensaios foram realizados em um termociclador de tempo real (Bio-Rad CFX connect). A fluorescência do Sypro® Orange ($\lambda_{\text{excitação}} = 490 \text{ nm}$ e $\lambda_{\text{emissão}} = 575$

nm) foi monitorada, enquanto as amostras foram aquecidas de 25 a 95 °C, com uma taxa de incremento de 0,5°C/min e 10 s de repouso. Os valores de T_m foram determinados pelo programa Bio-Rad CFX Manager usando a primeira derivada de fluorescência vs. temperatura. A temperatura do desenovelamento da proteína na ausência de qualquer ligante (T_{mRef}) foi subtraída dos valores na presença de ligantes (T_{mLig}), com o propósito de obter os ΔT_m . A fórmula usada para essa análise foi:

$$\Delta T_m = T_{mLig} - T_{mRef}$$

Para a seleção dos fragmentos, foram considerados como *hits* somente aqueles que incrementaram o ΔT_m em pelo menos três vezes o valor do desvio padrão em relação à T_{mRef} . Esse ponto de corte foi selecionado por representar o mínimo erro possível na seleção dos ligantes.

3.9 Dicroísmo Circular (do inglês Circular Dichroism - CD)

Para avaliarmos o enovelamento de estruturas secundárias das construções RdRp_truncada e RdRp_integral em solução, experimentos de CD foram realizados em um espectropolarímetro *Jasco J-815 CD Spectrometer*.

As medidas foram registradas em um intervalo de comprimento de onda de 195 nm a 280 nm, utilizando-se cubetas de quartzo de 1 mm de caminho óptico. Para os experimentos, 2 μ M de cada construção em tampão C foram utilizados. O espectro de CD resultante foi gerado por uma média de 4 leituras consecutivas. Os dados foram analisados pelo programa CD tool⁹⁵ e representados graficamente com o programa GraphPad Prisma 6.0.

3.10 Ensaio de atividade da construção RdRp_integral

Os ensaios de atividade foram baseados no marcador fluorescente de ácido nucleico SYBR Green I (Thermo Fischer) e a técnica foi adaptada do ensaio descrito por Sáez-Álvarez *et al.*⁹⁶

Inicialmente o RNA foi incubado a 55 °C por 5 min em um termociclador para PCR T100 (BioRad) para garantir a formação do grampo (*hairpin*) de RNA. Posteriormente, em um tampão de ensaio contendo 50 mM Tris pH 7,0; 2,5 mM de

MnCl₂ e 0,01% Triton X-100, foram preparadas duas soluções, a solução 1 contendo 250 nM de RdRp (Proteína 2x) e a solução 2 contendo 400 nM RNA, 1 mM ATP, 1x SYBR Green I (RNA+ATP+ SYBR Green I 2x). Para iniciar a reação, volumes iguais da solução 1 e da solução 2 foram homogeneizados em uma placa de PCR de 96 poços a um volume final de 50 µL.

Os controles foram preparados homogeneizando a solução 2 e tampão de ensaio, sem a presença da proteína. Em seguida, a reação foi incubada a 30 °C em um termociclador Stratagene MX3005P qPCR Agilent Technologies®. As intensidades de fluorescência do SYBR Green I foram avaliadas a cada 30 s com o uso do filtro FAM por 60 min. Os comprimentos de onda utilizados para excitação e emissão foram, respectivamente, 494 nm e 521 nm. Todos os experimentos foram realizados em duplicatas ou triplicatas. Os resultados foram analisados e plotados usando o software OriginPro 9.0.

Foi realizada a padronização do ensaio, na qual as seguintes condições foram variadas: concentração de proteína (1,0 - 0,062 µM), concentração de RNA (5,0 - 0,078 µM), concentração de ATP (15,62 - 1,0 mM) e concentração de SYBR Green I (5,0 - 0,156 x).

Todas as soluções foram preparadas com os mesmos tampões e RNA do ensaio anterior. A fim de evitar contaminação de RNAses todo o ensaio foi realizado dentro do fluxo laminar, com ponteiras com filtro, tubos eppendorf e placas de PCR livres de RNase.

3.11 Ensaio de inibição da atividade RdRp

Para o ensaio de inibição, utilizou-se o protocolo descrito na seção 3.10, com a diferença de que 0,5 µL de ligante a uma concentração desejada, diluído em DMSO, foi adicionado em 24,5 µL de solução 1 e incubado por 15 min à temperatura ambiente. Após o tempo de incubação, 25 µL de solução 2 foram adicionados e a reação foi homogeneizada em uma placa de PCR para dar início à reação. O controle positivo foi preparado adicionando 0,5 µL de DMSO em 24,5 µL da solução 1, iniciando a reação com a adição de 25 µL de solução 2, enquanto, para o branco, 24,5 µL de tampão de ensaio foi utilizado, ao invés de Proteína 2x. Como anteriormente, a reação e os controles foram incubados a 30 °C em um

termociclador Stratagene MX3005P qPCR Agilent Technologies®, as intensidades da fluorescência do SYBR Green I foram avaliadas a cada 30 s com o uso do filtro FAM por 60 min.

A porcentagem de inibição é obtida pela leitura da fluorescência da reação, do branco (tampão na ausência da proteína) e do controle positivo (reação na presença de 1% de DMSO), definida pela seguinte expressão:

$$\% \text{ Inibição} = 1 - \text{Atividade relativa} = 1 - \frac{\text{Reação} - \text{Branco}}{\text{Controle Positivo} - \text{Branco}}$$

As curvas de dose-resposta foram geradas a partir da % de inibição pela concentração de ligante. A melhor curva foi determinada a partir da regressão linear dos dados utilizando o modelo de Hill 1 no software Origin Pro 9, e a concentração inibitória média (IC₅₀) foi obtida a partir da interpolação dos dados da curva gerada.

3.12 Triagem virtual

A triagem virtual foi realizada utilizando o software Autodock Vina 1.1.2. O programa Autodock Vina é um programa livre, comumente utilizado em triagens virtuais de larga escala, que utiliza um algoritmo genético como estratégia de busca das poses de menor energia. Além disso, o programa utiliza uma função para classificar as poses, que considera a somatória das interações mais favoráveis entre o ligante flexível e o *grid* da proteína.⁹⁷ A estrutura cristalográfica de ZIKV utilizada foi a obtida em nosso laboratório e depositada no Protein Data Bank PDB (ID 5U04)⁹⁸ e os átomos de hidrogênio foram adicionados usando o servidor MolProbity.⁹⁴ O programa AutoDockTools 1.5.6 foi utilizado para formatar os tipos de átomos, calcular cargas de Gasteiger-Marsili e mesclar hidrogênios não polares em seus respectivos átomos pesados, por meio do protocolo de preparação padrão do software. As estruturas do ligante foram preparadas no AutoDockTools 1.5.6, seguindo o protocolo de preparação padrão para ligantes, permitindo total flexibilidade desses.⁹⁹ A grade foi centrada nos sítios ativo e alostérico da RdRp.

Esses experimentos foram realizados em colaboração com o professor Dr. Rafael Guido (IFSC-USP) para a triagem virtual de moléculas da biblioteca do CIBFar e com a professora Dra. Carolina Horta Andrade, do LabMol – Laboratório de

Planejamento de Fármacos e Modelagem Molecular (UFG) para a triagem de compostos oriundos do projeto OpenZika (OZ).¹⁰⁰

4 Resultados

4.1 Clonagem

Os resultados da clonagem da construção RdRp_integral estão apresentados na Figura 6. Os oligonucleotídeos iniciadores foram planejados em vetores pET-LIC em *E. coli* utilizando 3 vetores: cauda 6x-His (pETM11), cauda 6x-His + tiorredoxina (pETTrx-1a) e cauda 6x-His + SUMO protease (pETSUMO-1a). As proteínas de fusão tiorredoxina e SUMO protease foram adicionadas com o objetivo de verificar se, na presença dessas, há um aumento na solubilidade da proteína de interesse.

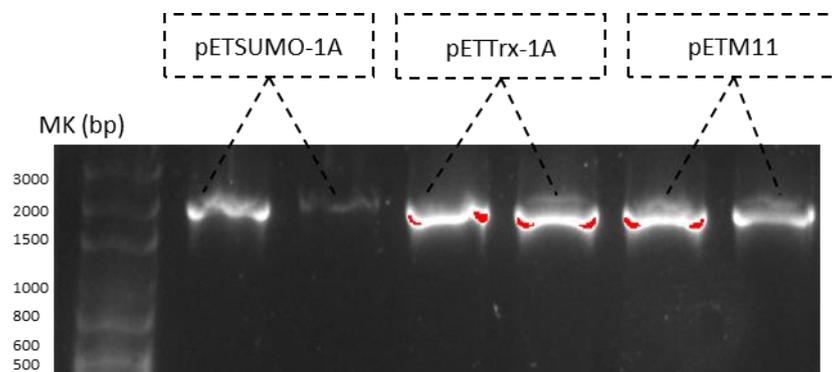


Figura 6 Gel de agarose 1% da PCR das colônias, contendo a sequência da proteína da construção RdRp_integral (1800 bp), clonadas nos vetores pETSUMO-1a/LIC, pETTrx-1a/LIC; e pETM11/LIC. Marcador Molecular *Mid Range DNA Ladder* Cellco Biotec.

Fonte: Elaborada pela autora.

A integridade da sequência de DNA codificante da proteína, de 1800 pb, foi confirmada também por sequenciamento gênico.

A construção RdRp_integral clonada no vetor pETSUMO-1a/LIC, apresentou mutações em suas sequências, e por isso, não foi utilizada nos estudos posteriores.

4.2 Expressão

Os testes de expressão foram realizados conforme descrito na seção 3.3, tendo sido analisados por gel de SDS-PAGE 10 %.

A proteína de interesse foi observada a partir da análise dos géis de SDS-PAGE 10%, na fração solúvel nos plasmídeos pETM11/LIC e pETTrx-1a/LIC para ambas as construções, em meio ZYM-5052 e LB.

Os testes para a construção RdRp_truncada foram realizados em colaboração com a empresa Cellco Biotec (Dra. N.U.Torres) e não são apresentados

neste trabalho, já os testes com a construção RdRp_integral podem ser visualizados na Figura 7.

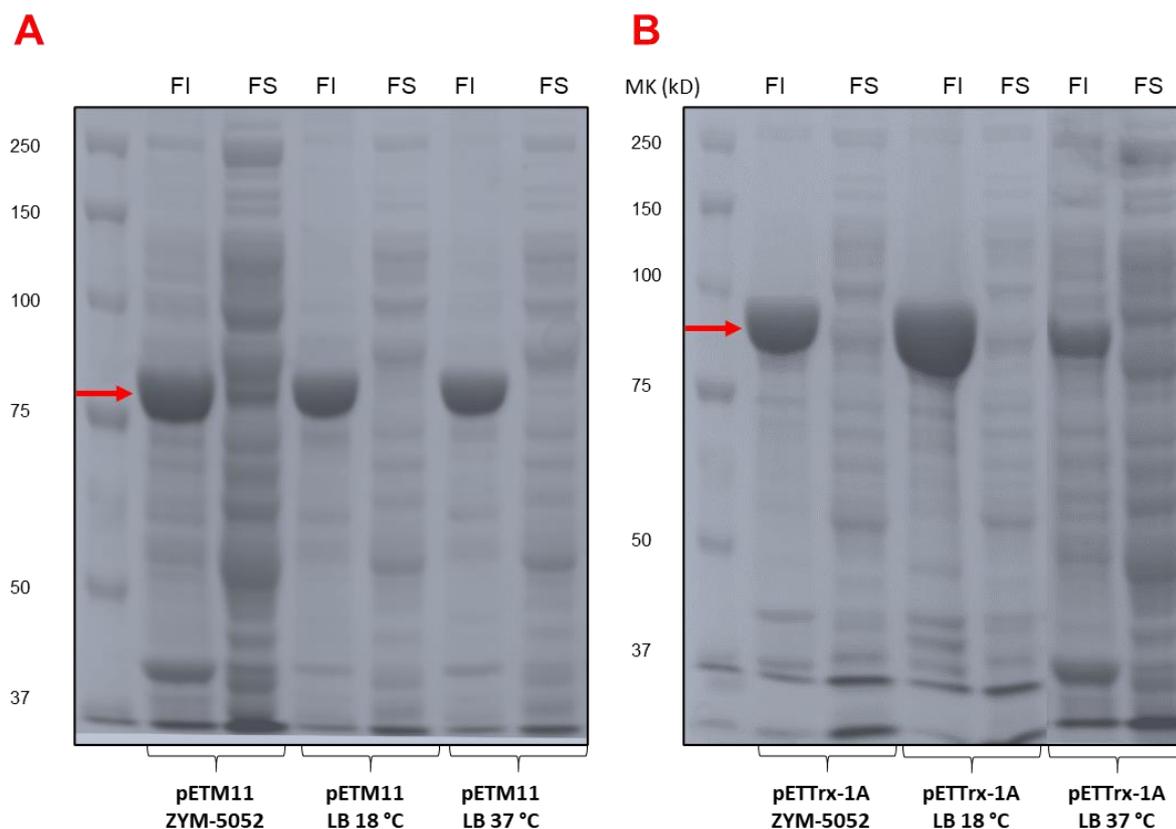


Figura 7 Géis SD-PAGE 10% do teste de expressão realizado com a construção RdRp_integral. FI é a fração insolúvel e FS a fração solúvel. A seta em vermelho representa a proteína de interesse. (A) RdRp_integral fusionada com a cauda de 6-His (75 kDa), (b) RdRp_integral fusionada com a Trx (86 kDa)

Fonte: Elaborada pela autora.

Como pode ser visualizado nos géis das Figuras referentes à construção RdRp_integral, ambos os sistemas de expressão pETM11/LIC e pETTrx-1a/LIC mostraram ser eficientes em meio auto indutor ZYM-5052 e em meio LB, apresentando uma banda larga de 77 kDa para a RdRp_integral fusionada com a 6-His (A) e 86 kDa para a RdRp_integral fusionada com a 6-His e a Trx (B) na fração insolúvel, foi possível observar banda na fração solúvel em ambos os vetores no meio auto indutor ZYM-5052. Entretanto, em meio LB, não foi possível detectar expressão proteica na fração solúvel em nenhum dos tempos analisados.

Protocolos de purificação foram testados para ambas as construções, porém, para o vetor pETM11/LIC, não foi possível obter a proteína pura no final das etapas de purificação. Por isso, os dados apresentados são para a purificação da construção no vetor pETTrx-1a/LIC.

4.3 Purificação, cristalização e estrutura cristalográfica

Quando este trabalho foi iniciado, em 2016, não existia nenhuma estrutura cristalográfica da NS5 domínio RdRp de ZIKV, a qual seria necessária para o desenvolvimento deste e de demais estudos. Em virtude disso, iniciamos o trabalho com foco na obtenção e resolução da estrutura cristalográfica de alta resolução, a começar pela construção RdRp_truncada, que segundo Malet *et al.*⁸³, facilita o processo de cristalização, por ser uma construção sem as regiões mais flexíveis da proteína.

4.3.1 Purificação, cristalização e determinação da estrutura cristalográfica da RdRp_truncada

Após clonagem e padronização do protocolo de expressão da construção RdRp_truncada, utilizamos a metodologia de purificação descrita no item 3.4, garantindo um rendimento de 5mg de proteína por litro de expressão e um alto grau de pureza, como podemos observar pelo gel de SDS-PAGE 12,5% apresentado na Figura 8A.

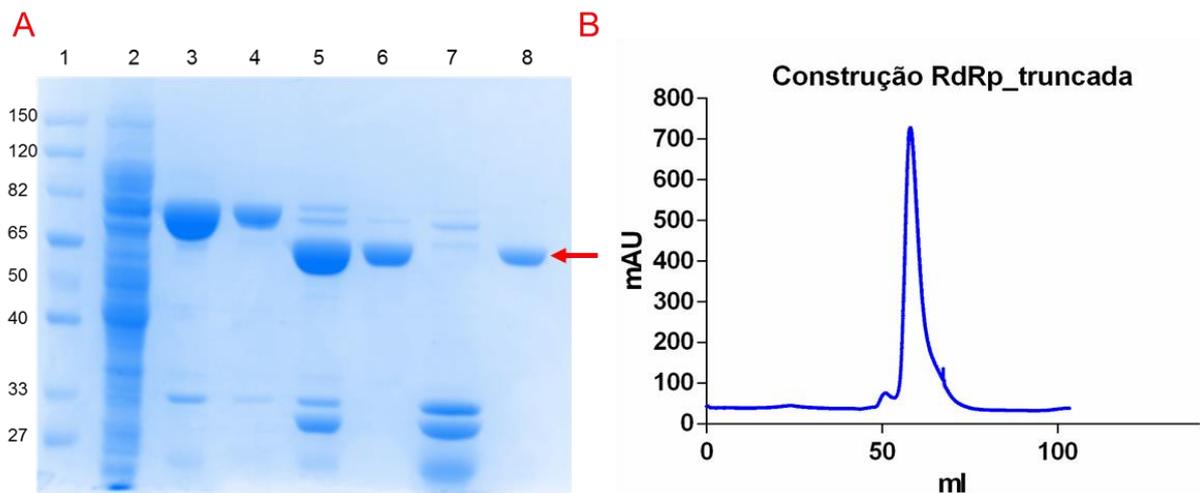


Figura 8 Purificação RdRp_truncada. (A) Gel SDS PAGE 10%, etapas de purificação da construção RdRp_truncada *1) Marcador Molecular; 2) Fração solúvel; 3) Fração eluída na primeira cromatografia de afinidade; 4) Fração após a troca de tampão; 5) Digestão com TEV protease; 6) Fração não-ligada na segunda cromatografia de afinidade; 7) Fração eluída de segunda cromatografia de afinidade, 8) Fração eluída na cromatografia de exclusão molecular, 65 kDa, como mostra a seta vermelha. (B) Cromatograma da etapa de exclusão molecular mostrando a eluição da proteína em um pico único e simétrico.

Fonte: Elaborada pela autora.

Observou-se, pelo Gel de SDS-PAGE 10%, uma proteína pura e com 65 kDa, conforme o esperado e, pelo cromatograma de exclusão molecular por tamanho,

que a eluição da proteína no volume de retenção de 58 ml é condizente com monômero em um pico único e simétrico, indicando que a amostra está homogênea, como apresentado na Figura 8B. Esse dado foi observado quando comparamos o gráfico de calibração da coluna (não apresentado).

Após a obtenção de proteína pura, diversas triagens de cristalização foram realizadas como descrito na seção 3.6. Após o monitoramento do crescimento dos cristais foi possível observar a formação destes cristais em diversas condições do kit Morpheus®. ¹⁰¹ Esses cristais foram submetidos à difração de raios X, no difratômetro disponível no Instituto de Física de São Carlos (IFSC-USP) e paralelamente na linha de luz MX2, no Laboratório Nacional de Luz Síncrotron (LNLS - Campinas), onde foi possível coletar diversos conjuntos. Os cristais que difrataram encontravam-se nas condições C6, D2, D6, E2, E6, F2, F6, G10 e H6. A composição de cada uma dessas condições pode ser observada na Tabela 3.

Tabela 3 Condições do kit Morpheus onde foram possíveis o crescimento e a visualização de cristais da RdRp_truncada

Condição	Mix de precipitantes	Mix de aditivos	Tampão
C6	10% PEG 8000, 20% Etileno glicol	0,03 M de cada NPS	0,1 M MOPS/ Hepes-Na pH 7,5
D2	10% PEG 8000, 20% Etileno glicol	0,02 M de cada álcool	0,1 M MES/Imidazol pH 6,5
D6	10% PEG 8000, 20% Etileno glicol	0,02 M de cada álcool	0,1 M MOPS/ Hepes-Na pH 7,5
E2	10% PEG 8000, 20% Etileno glicol	0,03 M de cada etileno glicol	0,1 M MES/Imidazol pH 6,5
E6	10% PEG 8000, 20% Etileno glicol	0,03 M de cada etileno glicol	0,1 M MOPS/ Hepes-Na pH 7,5
F2	10% PEG 8000, 20% Etileno glicol	0,02 M de cada monossacarídeo	0,1 M MES/Imidazol pH 6,5
F6	10% PEG 8000, 20% Etileno glicol	0,02 M de cada monossacarídeo	0,1 M MOPS/ Hepes-Na pH 7,5
G10	10% PEG 8000, 20% Etileno glicol	0,02 M de cada ácido carboxílico	0,1 M Bicina/ Trizma base pH 8,5
H6	10% PEG 8000, 20% Etileno glicol	0,02 M de cada aminoácido	0,1 M MOPS/ Hepes-Na pH 7,5

Fonte: Elaborada pela autora.

Observamos que a composição dos precipitantes que ocasionam a formação de cristais é a mesma para todas as condições, porém, o tampão e a os aditivos variam. Os melhores padrões de difração foram dos cristais obtidos no tampão 0,1 M

MOPS/ HEPES-Na pH 7,5, correspondentes às condições C6 (1,9 Å), F6 (2,1 Å) e H6 (2,4 Å)

A morfologia dos cristais obtidos e que apresentaram difração pode ser visualizada na Figura 9.

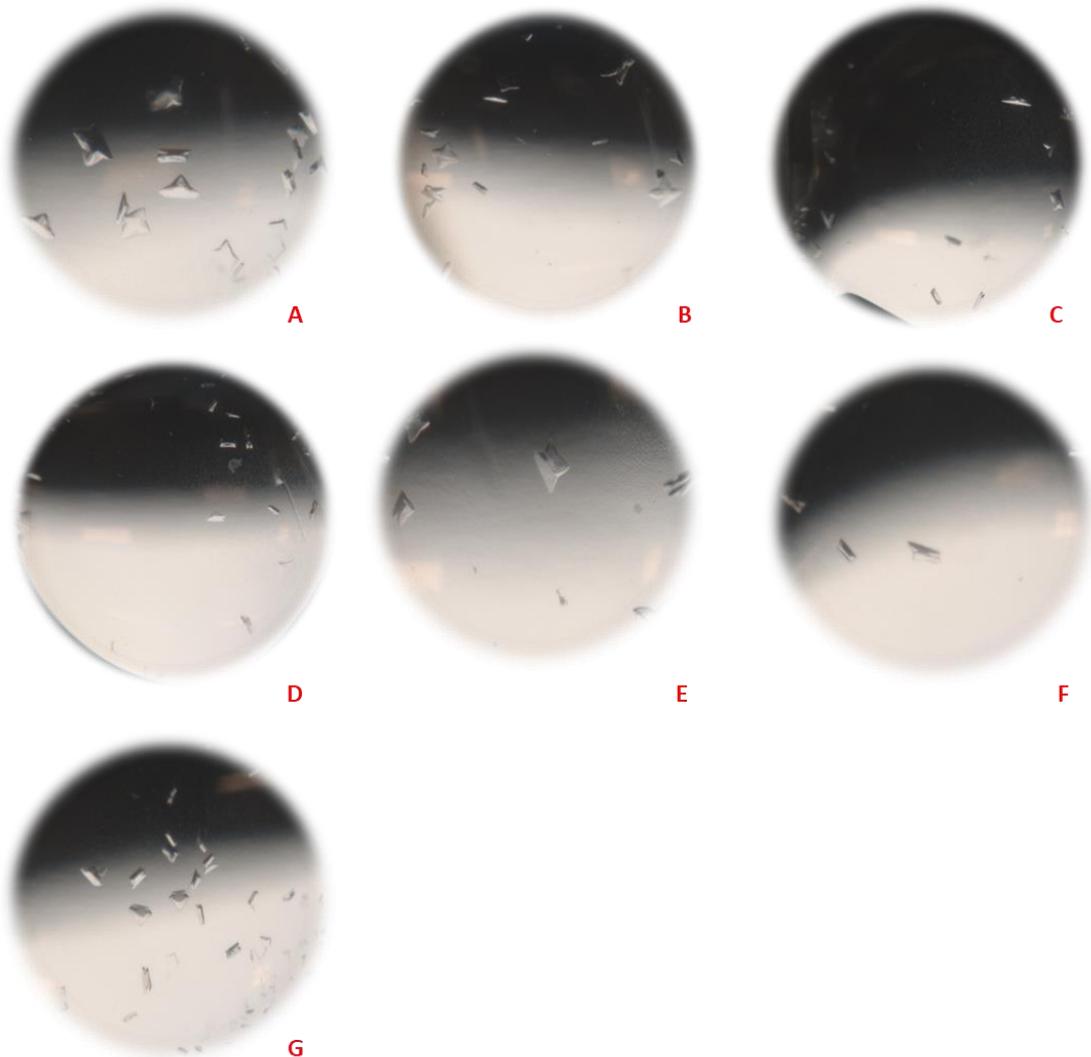


Figura 9 Cristais obtidos para a construção RdRp_truncada, a partir do kit de cristalização Morpheus, da Molecular Dimensions. As condições em que foi possível observar a formação de cristais que apresentaram difração foram: A) C6; B) D2; C) E2; D) F2; E) F6; F) G10 e G) H6.

Fonte: Elaborada pela autora.

Os cristais submetidos à difração de raios X, tiveram conjuntos de dados coletados em sua forma apo, ou seja, sem ligantes específicos adicionados. Simultaneamente, os cristais capazes de ocasionar difração foram submetidos a testes de *soaking* com possíveis ligantes. Ainda não foi encontrado nenhum ligante cristalizado com a proteína.

Um conjunto completo dos dados de difração de raios X à resolução de 1,9 Å foi obtido a partir de um cristal da condição C6. Esse conjunto de dados permitiu a primeira resolução da estrutura do domínio RdRp da NS5 de ZIKV por substituição molecular, utilizando como modelo a proteína homóloga 2HCN do *West Nile Virus*,⁸³ sendo que a sequência de identidade entre ambas é de 72%. A determinação da estrutura a partir deste conjunto de dados foi conduzida em colaboração com o Dr. Andre Schutzer de Godoy, pós-doutorando no nosso laboratório e colaborador deste trabalho.

Os resultados dessa estrutura foram publicados por nosso grupo de pesquisa na revista *Nature Communications*, com o seguinte título “*Crystal structure of Zika virus NS5 RNA-dependent RNA polymerase*”¹⁰², o qual já recebeu 55 citações computadas na base Web of Science na data de 12/02/2021.

Um resumo dos dados de difração e da resolução da estrutura cristalográfica de RdRp da NS5 de ZIKV estão listados na Tabela 4, na qual podemos verificar um conjunto com 99,7% de completeza, com alta multiplicidade e alta resolução. O refinamento da estrutura foi realizado até a obtenção dos valores estatísticos de $R_{work} = 17\%$ e $R_{free} = 20\%$.¹⁰²

Tabela 4 Parâmetros estatísticos relativos à coleta de dados e refinamento da NS5 RdRp de Zika vírus.

Coleta de dados	
Grupo espacial	$P4_32_12$
Célula unitária	
a, b, c (Å)	78,9; 78,9; 210,02
α, β, γ (°)	90,0; 90,0; 90,0
Resolução (Å)	29,66 - 1,9 (1,95 - 1,9)
R _{p.i.m.}	0,074 (0,85)
I/ σ I	9,9 (1,8)
Completeza (%)	99,7 (96,1)
Multiplicidade	24,3 (11,1)
CC _{1/2}	0,99 (0,72)
Refinamento	
Número de reflexões	53499
R _{work} /R _{free}	0,17/0,20
Número de átomos (não-H)	
Proteína	3878
Ligantes	7
Água	663
B-fatores	
Proteína	31,7
Ligantes	73
Água	44
RMSD	
Tamanho da ligação (Å)	0,004
Ângulos da ligação (°)	0,58

Fonte: Adaptada de GODOY, A. S. *et al.* ¹⁰²

A NS5-RdRp (código PDB 5U04), apresenta-se com uma arquitetura estrutural clássica das polimerases de flavivírus, como uma mão direita, com seus três principais subdomínios: dedos (resíduos 321-488 e 542-608), palma (resíduos 489-541 e 609-714) e polegar (resíduos 715-903). Os subdomínios dedos e polegar interagem formando um bolsão, que cerca o subdomínio polegar, onde estão contidos os sítios essenciais para a atividade da proteína. O subdomínio polegar apresenta uma protusão denominada *priming loop* (resíduos 785-810), que se localiza no centro do bolsão formado pelos três subdomínios. Além disso, foi possível verificar a presença de dois íons de Zn²⁺, um localizado entre os dedos e outro no polegar, como representado na Figura 10.⁵⁰⁻⁵²

Dentre os subdomínios, existem 7 motivos estruturais, que são altamente conservados em diferentes polimerases e que desempenham diferentes papéis importantes para a atividade enzimática, tais como: coordenação dos metais divalentes que catalisam a reação (motivos A e C), estabilização e direcionamento

da fita molde e produto (F, G e E), reconhecimento e estabilização dos NTPs (B e F), e liberação e neutralização do pirofosfato (D) (Figura 10 B).^{50–52}

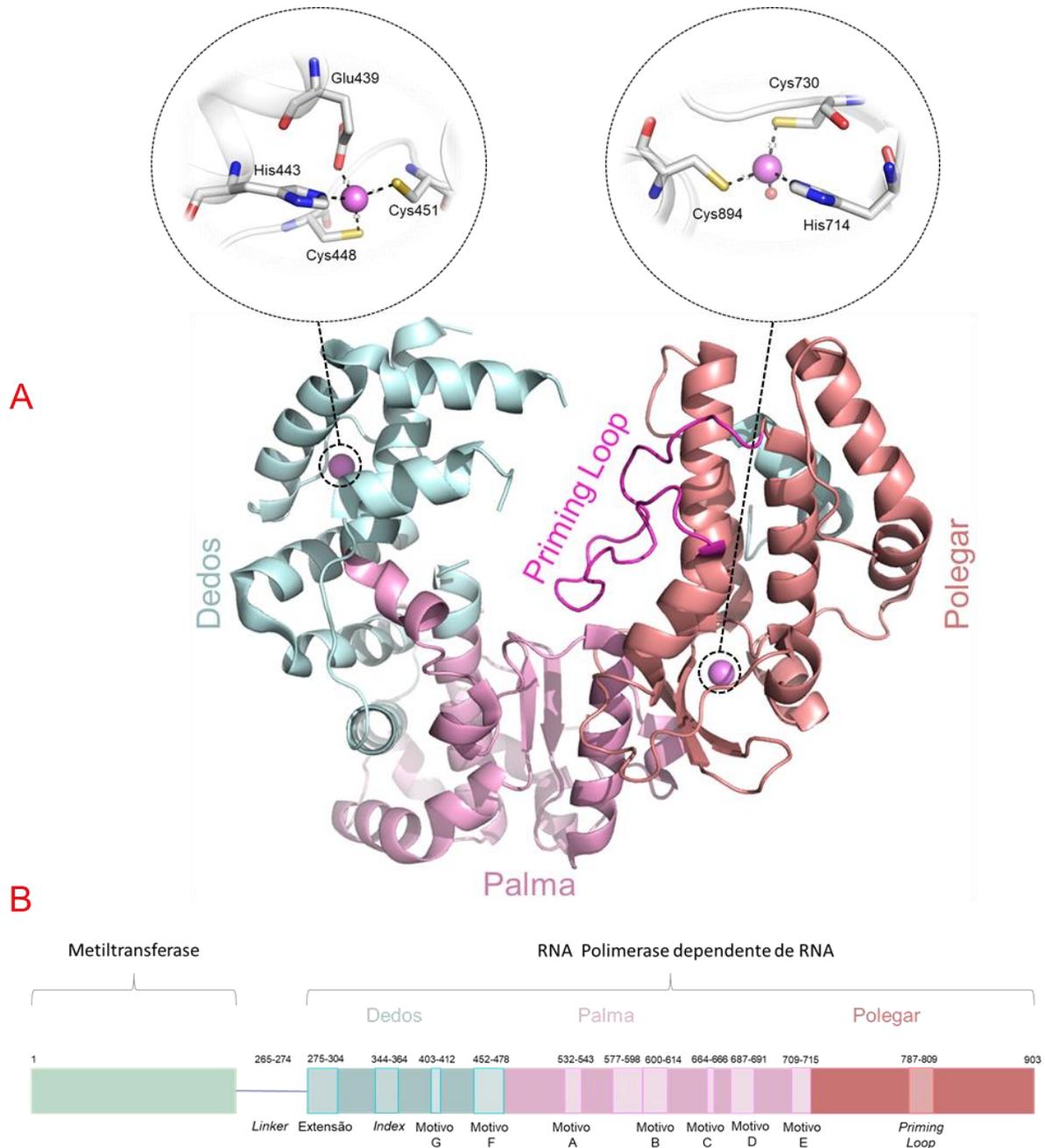


Figura 10 Estrutura cristalográfica da NS5-RdRp, obtida a partir da construção RdRp_truncada. (A) Estrutura cristalográfica NS5 RdRp proveniente da construção RdRp_truncada. Em ciano claro, o subdomínio “dedos”. Em rosa claro o subdomínio “Palma”. Em salmão, o subdomínio “polegar”. Em magenta, mostramos o “priming loop”. As esferas violeta representam a posição dos átomos de Zn^{2+} , e nos zooms podem ser visualizados os aminoácidos que coordenam esses Zn^{2+} . (B) Proteína NS5 com seus domínios Metiltransferase (verde) e RNA Polimerase dependente de RNA, seus subdomínios mantem as cores mostradas na estrutura cristalográfica da Figura (A), e os motivos A à G estão mostrados em seus respectivos locais.

Fonte: Elaborada pela autora.

A estrutura cristalográfica de ZIKV mostrou-se conservada, quando comparada às estruturas de outros flavivírus, assim como quando comparada às outras estruturas já descritas de ZIKV, possuindo 7 motivos, denominados de A a G.⁵⁰ A construção RdRp_truncada não contém a porção N-terminal composta pelos aminoácidos 275 a 305, sendo essa a região que contém a sequência de localização nuclear funcional (NLS) considerada importante para interações com outras proteínas virais e hospedeiras. Na estrutura cristalográfica 5U04 não foi possível identificar algumas porções de estruturas secundárias que interagem com a região de extensão NLS, tais como as regiões dos aminoácidos 268 a 320, 341 a 361, 407 a 420, uma vez que não foi observada densidade eletrônica associada a essas regiões.¹⁰²

Dois íons de zinco (Zn^{2+}) foram encontrados na estrutura, co-purificados naturalmente durante a expressão, pois as condições de purificação e cristalização não continham adição destes íons. Assim, como em outras estruturas de RdRp de flavivírus, um deles encontra-se no subdomínio dedos, que faz interação com os aminoácidos E439, H443, C448 e C451 e o outro no polegar, interagindo com os aminoácidos H714, C730 e C849.¹⁰³ Essas coordenações são mostradas na esfera da figura 10A. Estes íons zinco estão distantes do centro de reações catalíticas da enzima e possivelmente contribuem para a estabilização da estrutura durante o enovelamento e na catálise.

A catálise de entrada de nucleotídeos é realizada em cooperação com os aspartatos catalíticos D535, D665 e D666, localizados no subdomínio palma, com D535 localizado no motivo A e D665 e D666 localizados no motivo C. Esses aspartatos interagem com íons metálicos (*in vitro* Mg^{2+} ou Mn^{2+} ; *in vivo* ainda é uma questão aberta), que coordenam as reações de transferência de prótons dos nucleotídeos.^{52,103}

Como em outras estruturas da NS5 RdRp de flavivirus, o *priming loop*, localizado no centro do bolsão formado pelos três subdomínios, é responsável pelo posicionamento de nucleotídeos para a polimerização, fechando o sítio durante o início da polimerização *de novo* do RNA viral.^{104,50}

Estruturas cristalográficas de DENV2 e do DENV3 complexadas com ligantes na região do *priming loop*, mostram que esses ligantes se encontram entre a alça de iniciação e os subdomínios do polegar e da palma, o que pode acarretar numa

mudança conformacional na proteína durante o processo de alongamento. Esse bolsão na região do *priming loop* foi denominado de *N-pocket*. Na Figura 11, podemos observar as estruturas dos complexos cristalográficos de DENV sobrepostas e os ligantes encontrados na região *N-pocket*.^{105–107}

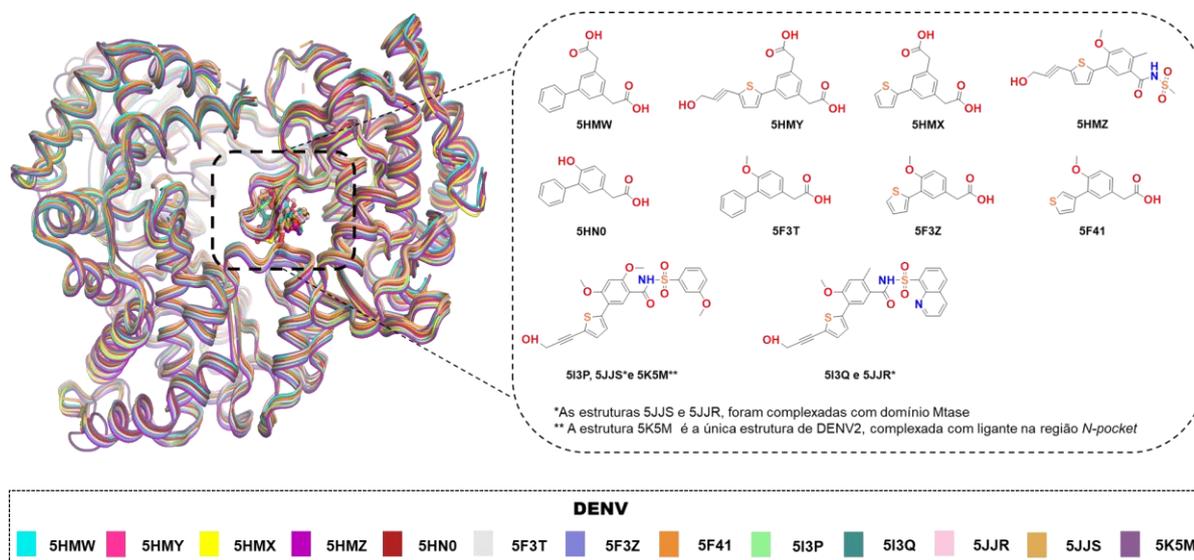


Figura 11 Estruturas cristalográficas do domínio RdRp de DENV cristalizadas com ligantes no sítio *N-pocket*. No quadro é possível observar a estrutura dos ligantes, nomeados de acordo com as estruturas às quais são correspondentes.

Fonte: Elaborada pela autora.

Esses mesmos autores também realizaram ensaios enzimáticos e puderam comprovar que, além de ligantes, algumas dessas moléculas inibem a atividade enzimática da proteína, mostrando que o bolsão é um sítio de atividade alostérica na proteína e comprovando a funcionalidade e a importância dessa região na busca por moléculas contra a atividade RdRp.^{105–107}

Baseando-se nesses artigos, Gharbi-Ayachi e colaboradores utilizaram inibidores já descritos para DENV e obtiveram estruturas cristalográficas com moléculas complexadas na região *N-pocket* para ZIKV. Eles também conseguiram determinar a inibição de algumas dessas moléculas e realizar modificações químicas nas mesmas para que seu potencial de inibição fosse aumentado, mostrando que essas moléculas se ligam ao sítio alostérico da proteína, sendo NNIs e *hits* promissores, conforme podemos observar na Figura 12.¹⁰⁸

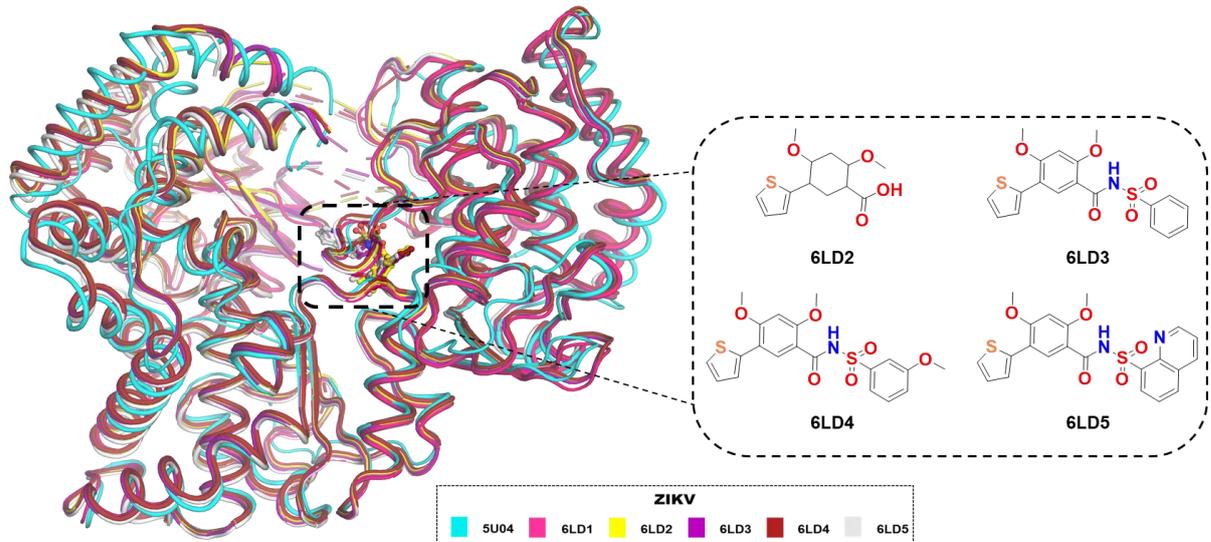


Figura 12 Estruturas de NS5 domínio RdRp de ZIKV. Estruturas cristalográficas de ZIKV complexadas com ligantes na região N-pocket sobrepostas com a Estrutura 5U04. Ligantes presentes na região *N-pocket*, denominados de acordo com sua respectiva estrutura cristalográfica.

Fonte: Elaborada pela autora.

Podemos observar que as estruturas de ZIKV cristalizadas na presença de ligantes possuem 100% de identidade com a estrutura resolvida, depositada e publicada pelo nosso grupo. Já quando comparadas às estruturas de DENV2 e DENV3, as identidades de sequência de ZIKV são de 76 e 81%, respectivamente.

As diferenças encontradas são observadas nos resíduos no início do *priming loop*. Em ZIKV, temos: V787, D788, G793; já em DENV2: S785, H786, S791 e em DENV3: H786, S791. Com relação aos resíduos expostos ao solvente e que estão dispostos no final do loop, as principais diferenças são em ZIKV: G801, G803, enquanto em DENV2 e DENV3, esses resíduos são substituídos por A799, H801, respectivamente e finalmente em ZIKV: K802, E804 e DENV2: K800, E802, esses resíduos são conservados, enquanto em DENV3 são substituídos por H800 e Q802. Todas essas substituições podem ser visualizadas na Figura 13.¹⁰²

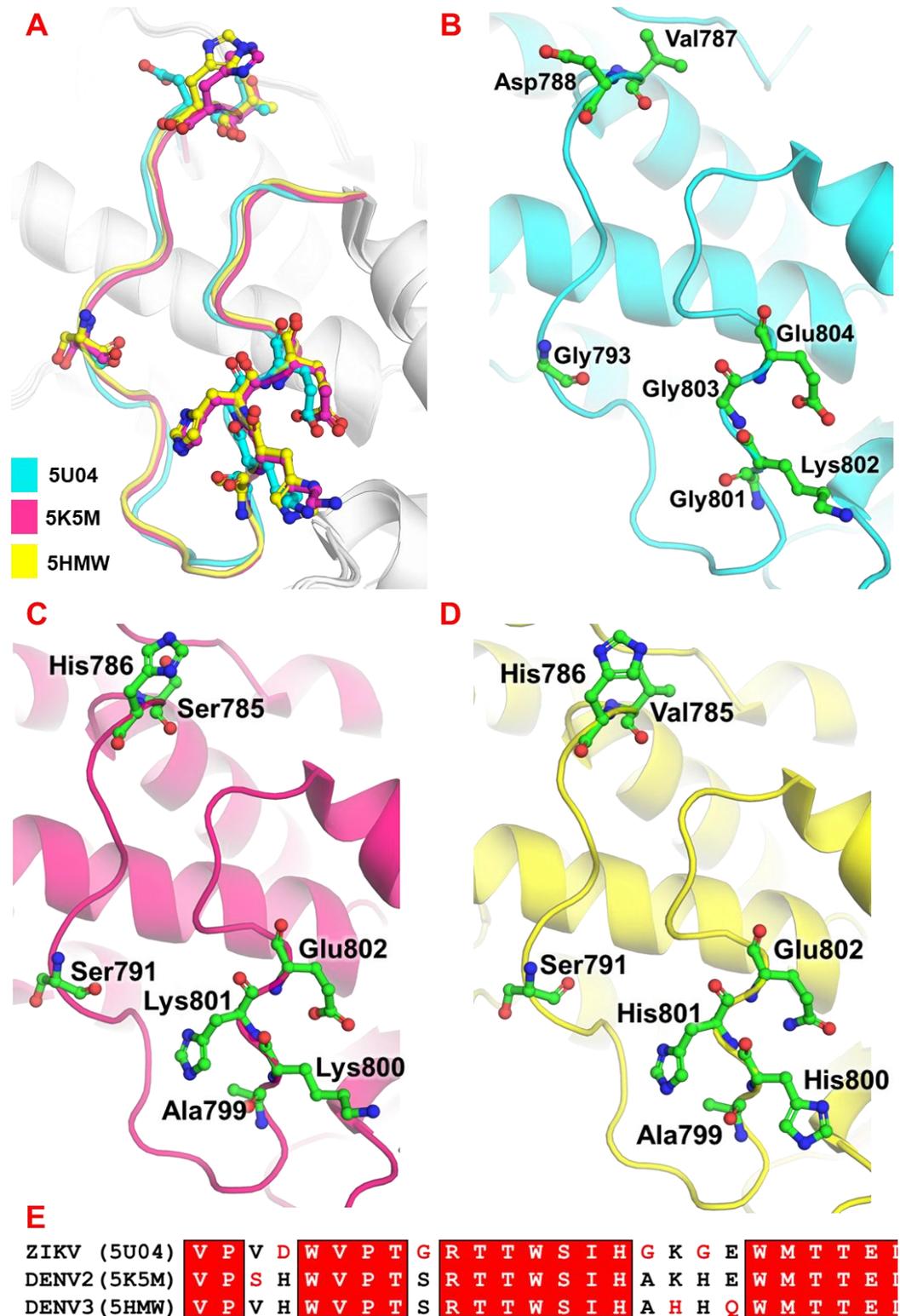


Figura 13 *Priming Loop* das estruturas cristalográficas de ZIKV, DENV2 e DENV3. (A) Sobreposição das regiões do *priming loop* com os aminoácidos que sofrem substituições em cada uma das estruturas. (B) *Priming loop* da estrutura cristalográfica de ZIKV (5U04) e os seus respectivos aminoácidos substituídos em comparação com as demais, (C) *priming loop* de DENV2 (5K5M) e os seus respectivos aminoácidos substituídos; (D) *priming loop* de DENV3 (5HMW) e os seus respectivos aminoácidos substituídos. (E) Alinhamento da região do *priming loop* das sequências das estruturas cristalográficas de DENV2, DENV3 e ZIKV.

Fonte: Elaborada pela autora.

Observamos que não há grandes alterações nessa região e isso nos permite demonstrar a importância do *priming loop* para a atividade proteica e mesmo que a estrutura analisada seja truncada, mantém-se conservada na região *N-pocket*, reforçando que pode ser usada na busca por ligantes do tipo NNIs.

4.3.2 Purificação e tentativas de cristalização da RdRp_integral

O desenho da construção RdRp_integral foi planejado de acordo com o artigo publicado por Duan *et al.*,¹⁰⁹ que depositou, em 2017, uma estrutura cristalográfica em alta resolução (1,8 Å) do domínio RdRp de ZIKV em sua forma integral.

Após clonagem e padronização do protocolo de expressão da construção RdRp_integral, utilizamos a metodologia de purificação descrita no item 3.4, garantindo um rendimento de 2,5 mg de proteína por litro de expressão e um alto grau de pureza. Como podemos observar pelo gel de SDS-PAGE 12,5% apresentado na Figura 14A.

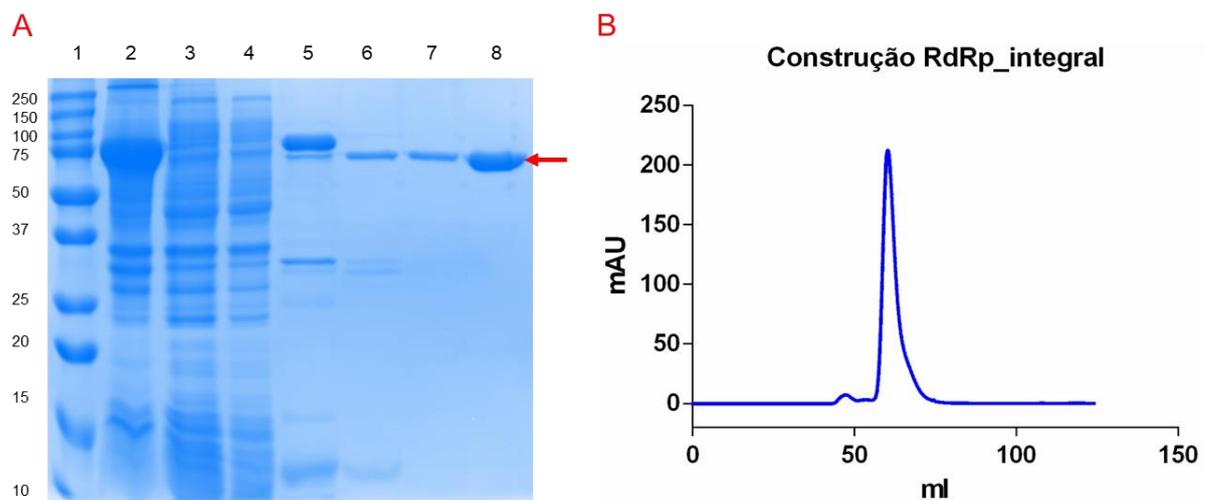


Figura 14 Purificação da RdRp_integral. (A) Gel de SDS-PAGE 12,5% com as etapas de purificação da RdRp_integral. 1) Marcador molecular Bio Rad; 2) Fração Insolúvel; 3) Fração Solúvel; 4) Fração não ligada na primeira cromatografia de afinidade; 5) Fração eluída na primeira cromatografia de afinidade; 6) Digestão com TEV protease; 7) Fração não-ligada na segunda cromatografia de afinidade; 8) Fração eluída na cromatografia de exclusão molecular, 74kDa, como mostra a seta vermelha. (B) Cromatograma da etapa de exclusão molecular mostrando a eluição da proteína em um pico único e simétrico.

Fonte: Elaborada pela autora.

Observou-se, pelo Gel de SDS-PAGE 12,5%, uma proteína pura e com 74 kDa, conforme o esperado e, pelo cromatograma de exclusão molecular por tamanho, que a eluição da proteína no volume de retenção de 61 ml é condizente com monômero em um pico único e simétrico, como observado quando comparado

ao gráfico de calibração (não apresentado), indicando que a amostra está homogênea

A construção RdRp_integral foi concentrada a 10 e 5 mg/mL. Uma triagem de diferentes condições de cristalização em diferentes concentrações de proteína foi realizada, porém não foi possível observar a formação de cristais.

4.3.3 Análise da estabilidade da construção RdRp_integral em solução por Dicroísmo Circular (CD)

A técnica de CD se baseia na interação da proteína, que é composta por aminoácidos quirais, com a luz circularmente polarizada. A diferença na absorção da luz circularmente polarizada à direita e à esquerda, resultando numa luz elipticamente polarizada, assim gerando um espectro em função do comprimento de onda que pode ser interpretado qualitativa e quantitativamente.¹¹⁰⁻¹¹¹

Uma análise qualitativa de estruturas secundárias foi realizada, a fim de avaliar o enovelamento e a estabilidade de construção RdRp_integral em solução, comparando-a com a construção RdRp_truncada, que foi cristalizada e mostrou-se conservada e estável. Na Figura 15 é possível observar os espectros obtidos, após a normalização por concentração de proteína, realizada pelo programa CDtool.

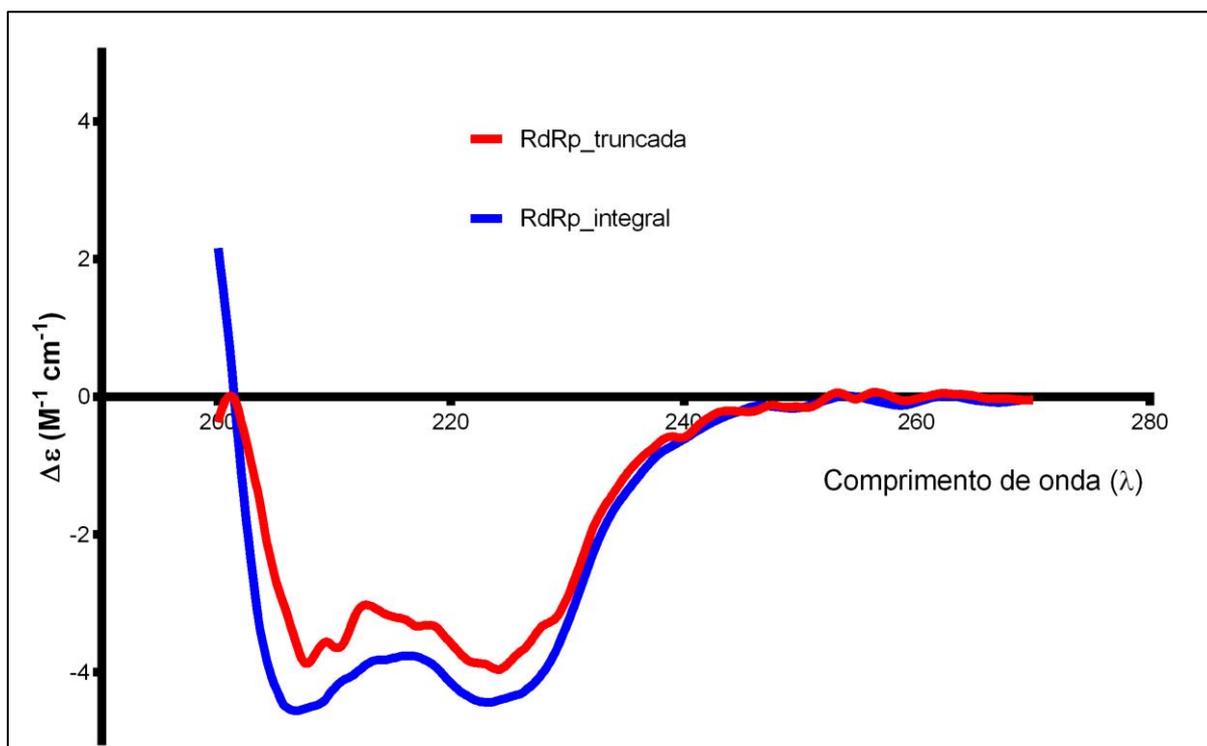


Figura 15 Espectro obtido após o experimento de CD, onde temos um espectro para cada construção em vermelho temos a construção RdRp_truncada e em azul a construção RdRp_integral. Observamos a definição das bandas em 208 e 220 nm nas bandas em azul, que mostram a estabilidade da proteína em solução.

Fonte: Elaborada pela autora.

Podemos observar que a construção RdRp_integral se apresenta estável em solução nas suas estruturas secundárias, pois foi observado que o espectro azul mostra as bandas de 208 e 220 nm (Figura 15), características bem definidas de estruturas secundárias do tipo α -hélices. A ausência de alguns resíduos na construção RdRp_truncada não permite a formação de alguns motivos e essas regiões podem ficar parcialmente desenoveladas, deixando o espectro vermelho ligeiramente mais indefinido na região de 208 nm. Porém, isso não impediu a ocorrência da cristalização, visto que essa região faltante apresenta grande mobilidade e sua ausência facilita o processo, como já discutido por Malet.⁸³

4.3.4 Análise da atividade da construção RdRp_integral

Após a obtenção e a confirmação de que a RdRp_integral é estável em solução, partimos para os ensaios que pudessem confirmar sua atividade polimerase.

Foi possível padronizar um ensaio fluorescente que monitora a atividade da proteína em tempo real, conforme descrito por Sáez-Álvarez *et al.*⁹⁶

O uso de sondas análogas ao SYBR Green para análise de formação e quantificação de dupla fita de RNA (dsRNA) é bem descrito na literatura e mais recentemente, demonstrou-se que elas podem avaliar a atividade de RNA polimerases de forma eficiente, em diferentes ensaios de atividades virais.^{96,112-113} Um esquema do ensaio está representado na Figura 16.

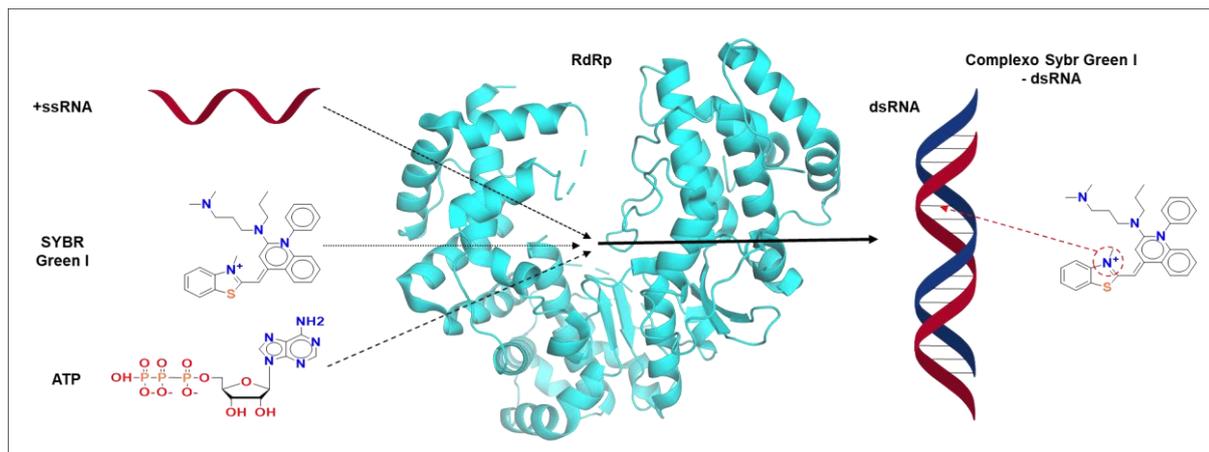


Figura 16 Esquema do ensaio de atividade. Esse ensaio registra a síntese de dsRNA, realizado pela RdRp (em ciano), a partir de um molde de +ssRNA (vermelho) e na presença de ATP. O SYBR Green I irá se intercalar ao dsRNA resultante, a partir de uma interação eletrostática entre o N^+ (círculo vermelho) e o fosfato do RNA, formando um complexo SYBR Green I-dsRNA que irá emitir fluorescência.

Fonte: Elaborada pela autora.

O ensaio fluorescente é acompanhado em tempo real em um equipamento de RT-PCR, pois o complexo formado SYBR Green I – dsRNA absorve a luz azul em 497 nm e emite luz verde em 520 nm. São utilizados como substrato o ATP, que é doador de nucleotídeos e uma fita molde de RNA poli U. O SYBR Green I é capaz de intercalar em duplas fitas de RNA com mais intensidade do que fitas simples, sem inibir a atividade enzimática. Dessa forma, garantimos um ensaio que mede a formação de dupla fita de RNA em tempo real.

Para otimizar a detecção da síntese de RNA e garantir que o ensaio é reprodutível e confiável, fizemos variações nas concentrações dos reagentes e analisamos o quanto essas afetaram a atividade enzimática. Os reagentes foram variados um a cada vez e foi possível observar as concentrações ótimas para a visualização da atividade, conforme observado na Figura 16.

Para o ensaio de otimização da concentração da proteína, a concentração de RNA foi fixada em 200 nM, a de ATP em 20 μ M, a concentração de SYBR Green I em 1x e realizou-se uma diluição seriada da concentração de RdRp 250-30 nM, na presença de 2,5 mM de $MnCl_2$, com monitoramento da reação por 60 minutos. A Figura 16A mostra todas as curvas obtidas, sendo possível observar um aumento na fluorescência com o tempo. Também não se observa decaimento do sinal, o que indica que a proteína permanece estável e ativa durante todo o ensaio. Observando

a fluorescência inicial do ensaio em relação ao seu crescimento, optou-se por trabalhar com 125 nM da enzima RdRp para os demais ensaios.

Para o ensaio da variação na concentração de RNA, fixou-se a concentração de RdRp em 125 nM e manteve-se a concentração de MnCl₂, ATP e SYBR Green I idênticas ao ensaio anterior, sendo que nesta etapa foi realizada a diluição seriada na concentração do RNA 5 – 0,08 µM. Vimos que em altas concentrações de RNA há um aumento na fluorescência inicial do ensaio, o que pode indicar a ligação do SYBR Green I ao substrato ssRNA. No entanto, à medida que o ensaio acontece, há um aumento da fluorescência ao longo do tempo com um padrão típico de cinética enzimática, o que indica que o SYBR Green I é capaz de intercalar dsRNA. Como observado na Figura 16B, optou-se por trabalhar com 200 µM de RNA, pois nesta concentração não se observa uma alta leitura na fluorescência inicial.

Na Figura 16C, observa-se que a presença do ATP pode causar interferência na reação. Para otimizar este componente, fixou-se a concentração de RdRp, MnCl₂ e SYBR Green I, como nos ensaios anteriores e foi mantida a concentração de RNA em 200 µM. Realizou-se, também, a diluição seriada da concentração e ATP de 1,0 - 0,01 mM. Observamos que o comportamento linear se manteve para diversas concentrações de ATP com inclinações proporcionais às concentrações do substrato, como esperado em um comportamento enzimático típico. Devido à linearidade e à intensidade da fluorescência observada, optou-se por utilizar ATP na concentração de 0,5 mM, idêntica à adotada por Sáez-Álvarez *et al.*⁹⁶

Por último, avaliou-se a variação na concentração de SYBR Green I em solução (Figura 16D). Após definir demais concentrações dos reagentes, realizou-se uma diluição seriada de 5,0-0,16 x de SYBR Green I, para avaliação da concentração ideal. É possível observar que altas concentrações de SYBR Green I inibem a atividade da RdRp, achado também relatado por Sáez-Álvarez *et al.*⁹⁶ Dessa forma, optou-se por utilizar a concentração do SYBR Green I em 0,5x, que foi a menor concentração sem interferência na fluorescência inicial ou na atividade enzimática.

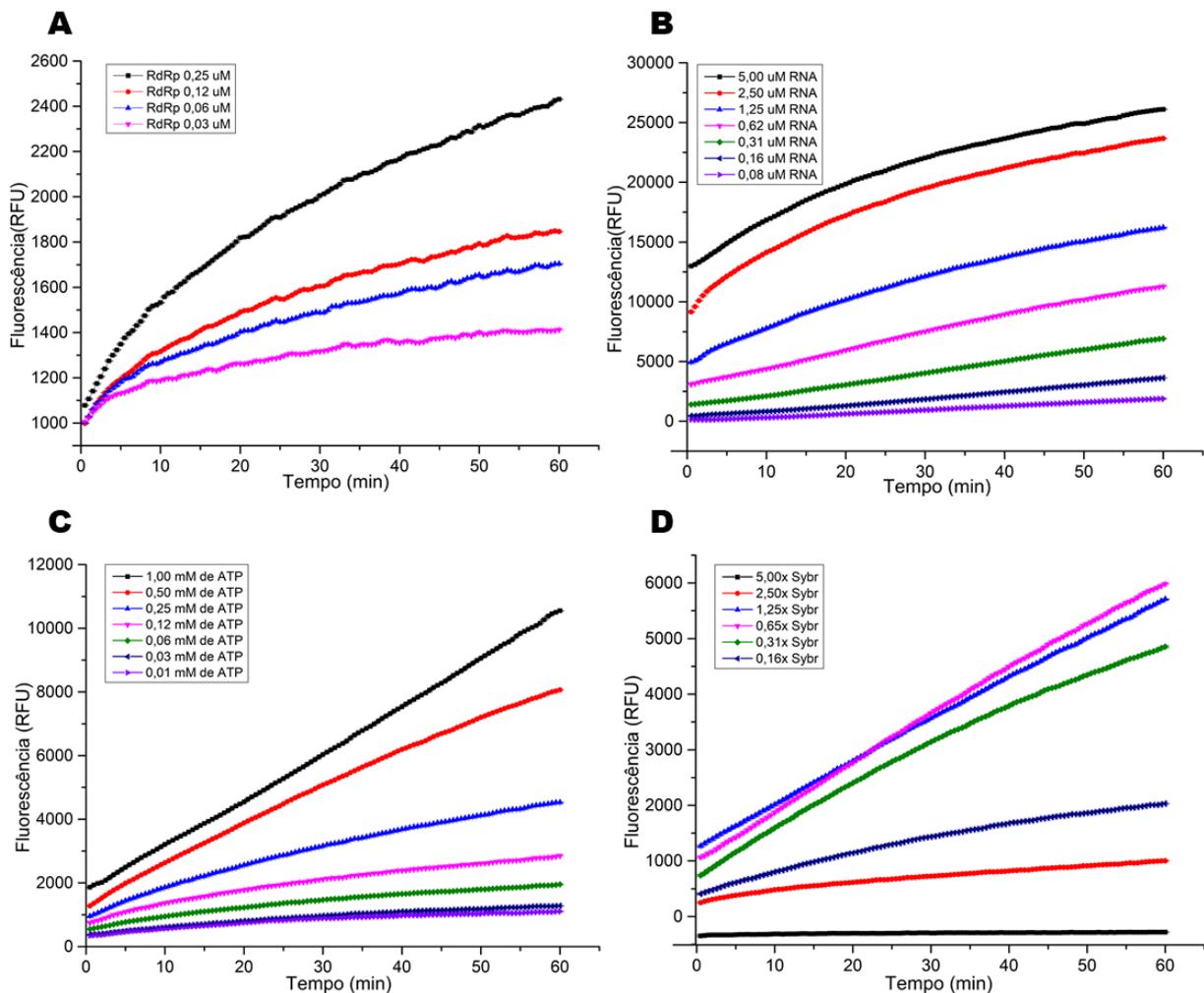


Figura 17 Otimização das concentrações dos reagentes para a detecção da atividade da proteína NS5 domínio RdRp. Em (A) variou-se a concentração de proteína frente aos demais reagentes, após a padronização da concentração de proteína, em (B) foi realizada uma variação na concentração do RNA, mantendo fixo a concentração dos demais reagentes. Em (C) e (D) segue-se o mesmo protocolo porém variando as concentrações de ATP e SYBR Green I, respectivamente.

Fonte: Elaborada pela autora.

A sonda fluorescente SYBR Green I é capaz de intercalar em duplas fitas de RNA com maior intensidade do que fitas simples sem inibir a atividade enzimática em concentrações mais baixas, garantindo um ensaio que mede a formação de dupla fita de RNA em tempo real. Além da vantagem de ser um ensaio em tempo real, os valores de intensidade observados no ensaio foram significativos, garantindo uma leitura ampla e contribuindo para a precisão dos ensaios de inibição. Uma característica importante deste ensaio de atividade é a sua realização em placas de 96 poços de fácil montagem, o que possibilita o seu uso em triagens de larga escala.

Os resultados obtidos para a triagem mostraram que a proteína manteve uma atividade estável mesmo na presença de 1% de DMSO. A qualidade do ensaio pode

ser medida a partir do fator-Z médio para cada placa, que é calculado a partir das médias e do desvio padrão, dos controles positivos e negativos do ensaio, para as 5 placas o valor obtido foi de $0,65 \pm 0,13$. Considerando que cada placa foi montada em dias diferentes e que, em duas delas lotes diferentes de proteína foram usados, o fator-Z é excelente.¹¹⁴ O valor de fator-Z obtido mostra que o ensaio padronizado pode ser utilizado para a triagem em larga escala e que os resultados obtidos são válidos.

4.4 Busca por ligantes

4.4.1 Ensaios de inibição em larga escala (*HTS – High Throughput screening*)

Para avaliar a efetividade do ensaio de atividade em larga escala, testou-se a atividade proteica por meio da biblioteca de compostos *Pandemic Response Box* (PRB - <https://www.mmv.org/mmv-open/pandemic-response-box>). A biblioteca, elaborada em conjunto pelas organizações *Medicines for Malaria Venture* (MMV) e *Drugs for Neglected Diseases initiative* (DNDi), possui 400 compostos com características “*drug-like*”, ativos contra bactérias, vírus ou fungos, e é dividida em 5 placas nomeadas de A a E. Cada placa possui 8 poços para controle positivo, 8 para controle negativo e 80 poços com diferentes compostos. A placa A consiste majoritariamente de compostos antifúngicos, as placas B e C de compostos antibacterianos e as placas D e E de compostos antivirais.

Para cada composto, realizou-se um ensaio de ponto único com cada uma das placas. Nesses ensaios, a atividade da proteína foi avaliada na presença de 20 μM de cada composto. O resultado está resumido na Figura 17. Cerca de 24 compostos diminuíram a leitura da fluorescência da atividade da proteína em até 80 % ou mais. Foram eles: 4 compostos da Placa A, 8 Placa B, 4 Placa C, 4 Placa D e 4 Placa E.

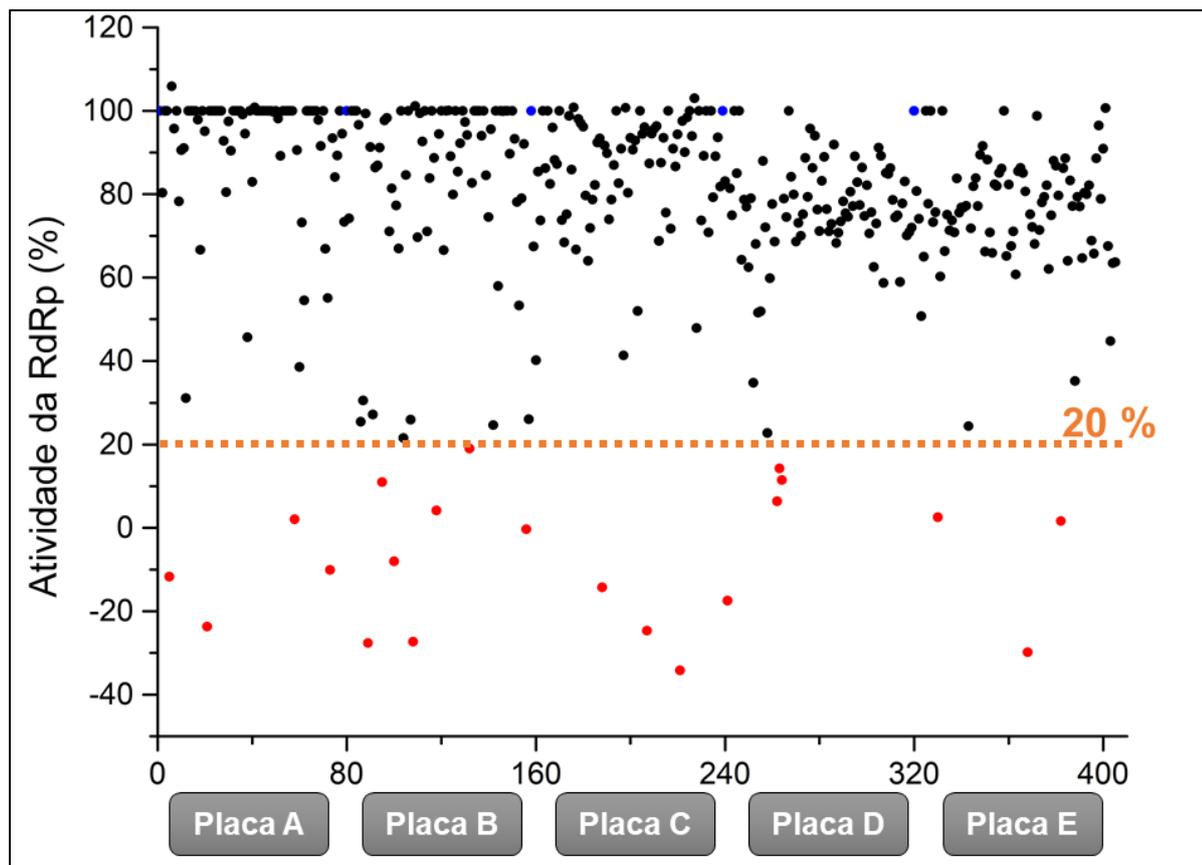


Figura 18 Gráfico da atividade relativa pela posição de cada poço nas placas da PRB, pontos pretos se referem a poços com compostos enquanto os pontos azuis a controles positivos, já os pontos em vermelho representam as moléculas que diminuem a fluorescência da leitura da atividade da proteína em mais de 80%, a linha laranja mostra a atividade da proteína em 20% ou menos.

Fonte: Elaborada pela autora.

O número elevado de compostos que diminuíram a atividade da proteína chamou a atenção, uma vez que, usualmente, a chance de acerto em triagens de larga escala é tipicamente da ordem de 0,4-0,5%. Assim, os compostos foram avaliados com chance de serem falsos positivos. Por se tratar de um ensaio em que se utiliza a fluorescência como medida, compostos falsos positivos podem ocorrer devido a sua interferência com a sonda. Medidas para predizer falsos positivos são importantes em triagens em larga escala, já que, ao evitar que esses sejam levados à etapas posteriores, tempo e dinheiro podem ser salvos.¹¹⁵

No ensaio com SYBR Green I podemos citar dois exemplos de compostos falsos positivos: aqueles que interagem de alguma maneira com o RNA, impedindo a intercalação da sonda (o que poderia dar um sinal aparente de inibição, mas na realidade é apenas o composto interagindo com o RNA); ou compostos que

interferem diretamente com a fluorescência da sonda, seja por “*quenching*” ou por fluorescência intrínseca.

A solução encontrada para medir a interferência do composto com o ensaio, foi utilizar a capacidade do SYBR Green I de emitir um sinal de fluorescência mesmo em fitas simples de RNA. Por meio da incubação do composto com o RNA, foi possível estabelecer uma % de Inibição de Fluorescência (IF) para cada composto e medir a relação entre a reação na ausência de proteína e contendo somente o RNA e ligante, subtraída pelo branco (que é a reação sem RNA e ligante), que por sua vez é dividida pela reação na presença de RNA e sem ligante subtraída pelo branco, como podemos observar na equação abaixo:

$$\% \text{ Inibição da Fluorescência} = 1 - \frac{(RNA + Ligante) - (Branco)}{RNA - Branco}$$

Os valores que se mantiveram com a fluorescência acima de 80% indicam que o composto não interfere com a sonda por meio de *quenching*, enquanto valores abaixo de 80% podem indicar que houve tanto *quench* quanto interferência com RNA, como mostra o gráfico de barras da Figura 19. Desse modo, foi possível determinar os compostos que interagiram com o RNA, ou com a sonda, e os que de fato inibiram a atividade da proteína.

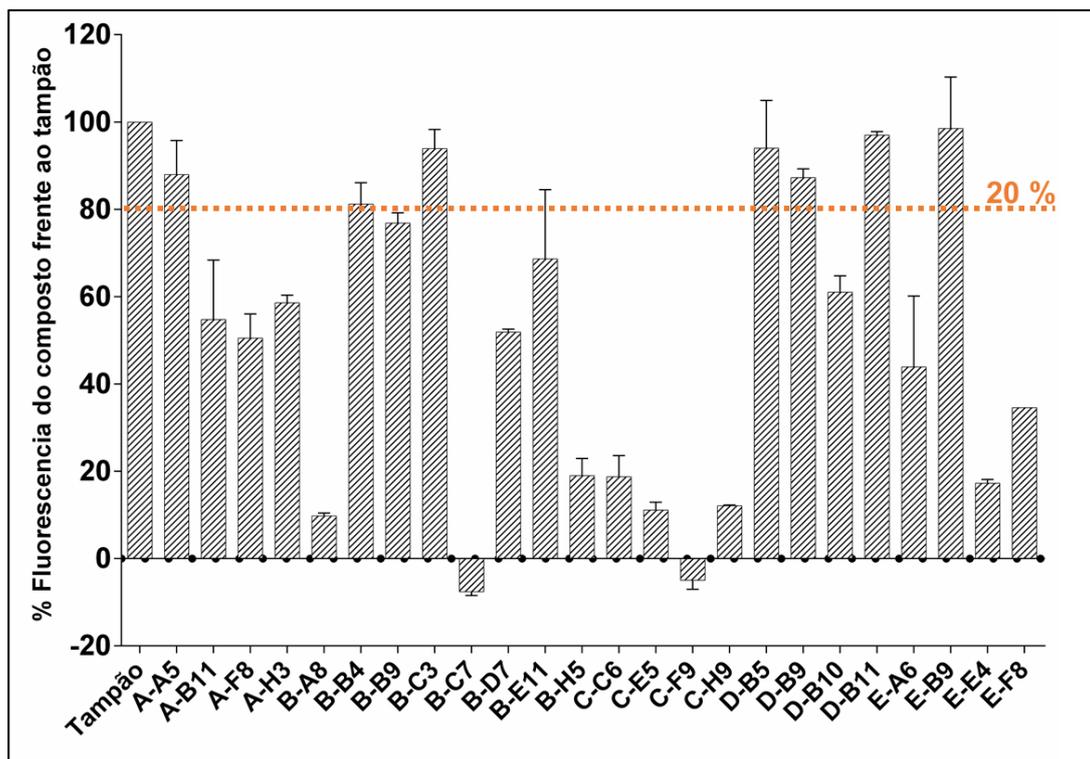


Figura 19 Gráfico da % de inibição dos compostos descobertos, frente ao tampão e na ausência da proteína. Dos 24 compostos testados observamos que 17 ainda interagem com o tampão, isso mostra que eles não inibem a atividade da proteína.

Fonte: Elaborada pela autora.

Podemos identificar que 17 compostos provocaram diminuição na fluorescência, indicando que houve interferência de fluorescência com componentes do tampão, devendo, portanto, ser descartados. Um dos motivos para o número de interferências se deve à região de excitação e emissão da sonda. Bibliotecas de compostos tendem a ter uma interferência maior na região azul-verde do espectro de emissão, exatamente a região onde atua o SYBR Green I. ¹¹⁵

Sete desses compostos testados não inibiram a fluorescência do ensaio, assim, mostraram não interagir com a sonda ou com o RNA e foram submetidos a ensaios de concentração-resposta, que possibilitaram a determinação do IC₅₀ para cinco deles (Figura 20).

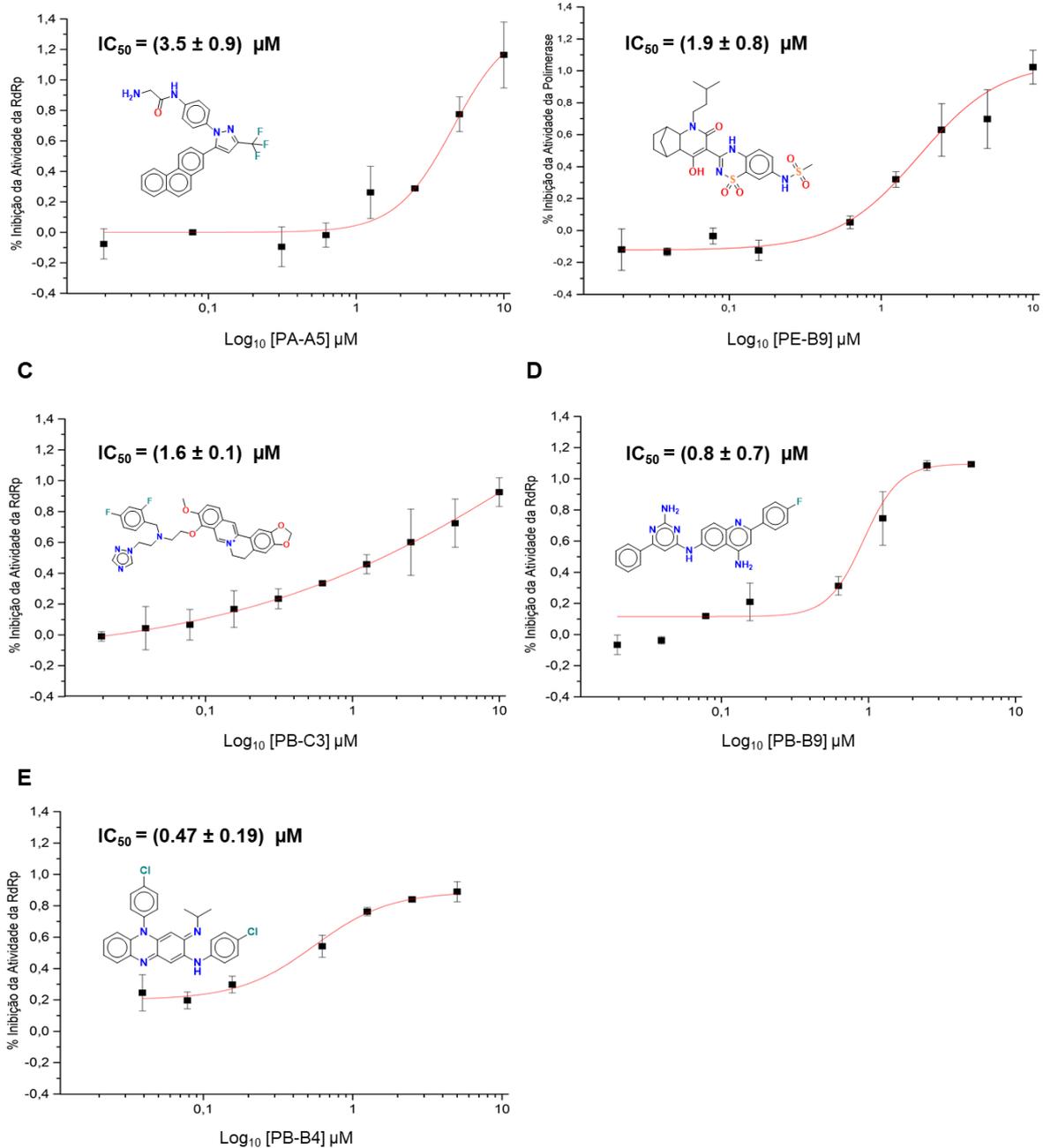


Figura 20 Curvas concentração-resposta para os cinco compostos do PRB, que tiveram IF abaixo de 20%, com suas respectivas estruturas moleculares. Sendo (A) o composto da placa A, referente ao poço A5, código PRB MMV1578560. (B) o composto da placa E, referente ao poço B9, MMV1782220; (C) o composto da placa B, referente ao poço C3, MMV1593540; (D) o composto da placa B, referente ao poço B9, MMV1593541 e (E) o composto da placa B, referente ao poço B4, o mais potente entre os compostos PRB encontrado, a clofazamina, MMV687800.

Fonte: Elaborada pela autora.

O composto da placa A, referente ao poço A5, é relatado pelo MMV com o código MMV1578560 e apresentou um IC_{50} de $3.5 \pm 0.9 \mu M$. Esse composto é descrito na literatura como um derivado do antitumoral e antimicrobiano Celecoxib, porém, ao contrário de seu precursor, não inibe COX-2, estudado inicialmente por

sua atividade anticâncer (inibidor de PDK1). Esse composto foi combinado com outros fármacos e apresentou atividade antifúngica contra *S. cerevisiae*, *C. albicans* e *C. neoformans*,¹¹⁶ e antiviral contra vírus que causam febre hemorrágica, como Lassa, Marburg e Ebola.¹¹⁷

O composto da placa E, poço B9, código MMV1782220, descrito na literatura como um composto NNI, demonstrou atividade antiviral contra RdRp de HCV, embora outros estudos mostrem que análogos dessa molécula foram mais potentes e com maior biodisponibilidade contra o replicon de HCV.¹¹⁸ Nossos resultados mostraram um IC₅₀ de $1.9 \pm 0.8 \mu\text{M}$ contra a RdRp de Zika.

O composto da placa B, poço C3, código MMV1593540, é um derivado da berberina, um alcaloide isoquinolínico natural conhecido por possuir atividade farmacológicas diversas, entre elas antimicrobiana^{119,120} e antiviral.⁷⁰ A proposta do composto MMV1593540 foi sintetizar a berberina combinada com o fluconazol, um antifúngico, como mostrado na Figura 20, gerando análogos de fluconazol, a fim de aumentar seus índices terapêuticos e ampliar o espectro antimicrobiano.^{119,120} Contra a RdRp de Zika, esse composto apresentou um IC₅₀ de $1.6 \pm 0.1 \mu\text{M}$.

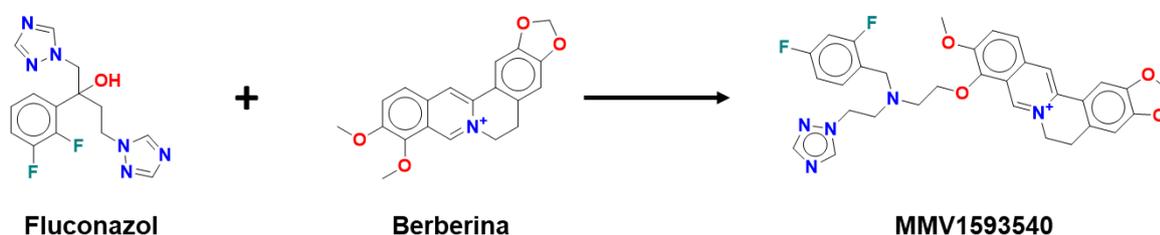


Figura 21 Esquema simplificado da síntese do composto MMV1593540.
Fonte: Elaborada pela autora.

O composto da placa B, poço B9, código MMV1593541 é um dos derivados de uma classe de compostos sintetizados por pesquisadores da AstraZeneca, com intuito de encontrar um potente agente antituberculose. Esses compostos possuem atividade antibacteriana e IC₅₀ na casa de baixo micromolar e alguns compostos apresentaram IC₅₀ na faixa de nanomolar contra a NADH-2 de *Mycobacterium tuberculosis*.¹²¹ Contra a RdRp de ZIKV o mesmo apresentou um IC₅₀ de $0.8 \pm 0.7 \mu\text{M}$.

O último composto em que foi possível determinar o valor de IC₅₀, era pertencente à placa B, poço B4, código MMV687800, que possui atividade

antibacteriana e apresentou um IC_{50} de $0.47 \pm 0.19 \mu M$, contra RdRp de ZIKV. Esse composto é relatado na literatura como a Clofazamina, um fármaco de amplo espectro usado contra o *Mycobacterium leprae*, descrito em 1957 por Barry e colaboradores atividade antituberculose,¹²² porém este mostrou-se ineficaz contra a tuberculose e atualmente é usado clinicamente para o tratamento da Hanseníase.¹²³ Ademais mostrou-se eficaz contra o vírus Sars-Cov-2.¹²⁴

A maioria dos compostos não possuem atividade relatada em polimerases de flavivírus. O único com atividade relatada em RdRp de HCV é o composto MMV1782220,¹¹⁸ portanto, tais resultados apresentam novas estruturas químicas que podem contribuir para a descoberta de antivirais contra o ZIKV.

Esse ensaio demonstrou sua adequação para a busca por ligantes em bibliotecas amplas de compostos, e dessa forma, demos início aos ensaios de inibidores com as moléculas provenientes dos nossos colaboradores.

4.4.2 Triagem de ligantes a partir de técnicas biofísicas

Após a obtenção da estrutura cristalizada proveniente da construção RdRp_truncada, iniciamos a busca por possíveis ligantes, usando a abordagem de Planejamento Baseado em Fragmentos (FBDD, *Fragment-Based Drug Discovery*), a partir de técnicas biofísicas e, posteriormente, de análise da inibição dos compostos encontrados.

O FBDD utiliza pequenas moléculas com peso molecular entre 150-300 Da ou 8-18 átomos pesados (não-H), de diferentes classes e funcionalidades químicas. Assim, um número maior de diferentes grupos químicos pode ser triado, aumentando a possibilidade de que vários desses fragmentos liguem-se à proteína, uma vez que a probabilidade de compostos pequenos ligarem-se ao sítio ativo ou a sítios alostéricos da proteína é maior, ainda que com baixa afinidade.¹²⁵⁻¹²⁶

No âmbito da triagem de fragmentos, existem diversas abordagens metodológicas, sendo uma delas a utilização de técnicas biofísicas de caracterização da proteína em solução.¹²⁷⁻¹²⁸

Para isso, utilizamos um ensaio simples e de baixo custo, a Fluorimetria Diferencial de Varredura (DSF, *Differential Scanning Fluorimetry*), que é uma técnica que permite investigar a estabilidade térmica de uma proteína em solução. Quando a

proteína é exposta a diferentes compostos, é possível medir o aumento da fluorescência ocasionado pela desnaturação térmica por meio da inserção de uma sonda fluorescente que se liga a regiões hidrofóbicas da proteína e ficam expostas ao solvente devido ao desenovelamento proteico. Podemos, assim, diferenciar dois estágios de enovelamento de uma proteína: enovelada (baixa fluorescência) e desenovelada (alta fluorescência). Tendo os dois estágios bem definidos é possível realizar um ajuste através de uma Função de Boltzmann- utilizando ferramentas gráficas – e determinar a Temperatura de *midpoint* (T_m), ou seja, a temperatura em que a metade das proteínas em solução se encontram enoveladas e assim, inferir uma maior ou menor estabilidade térmica da proteína.¹²⁷⁻¹²⁸

A técnica foi utilizada frente a *uma* biblioteca com 514 fragmentos e compostos, disponível no Laboratório de Biologia Estrutural Aplicada (LBEA) do Instituto de Ciências Biomédicas da USP (ICB-USP), cedida pelo Prof. Dr. Marcio Dias, utilizando a construção RdRp_truncada, uma vez que esta é a única construção com protocolo estabelecido para a obtenção de cristais, visando a obtenção de complexos que possam auxiliar na elucidação dos mecanismos de ligação desses compostos frente a proteína. Os ensaios foram realizados como descrito na seção 3.6.

A biblioteca de compostos do Prof. Dr. Marcio Dias é organizada em 7 blocos, denominados de A a G, sendo composta ao todo por 514 compostos/fragmentos. Tais compostos são denominados pelo código do laboratório (LBEA) e seu número correspondente, e 112 possibilidades para controle positivo e negativo. Um esquema da biblioteca é apresentado na Figura 22.

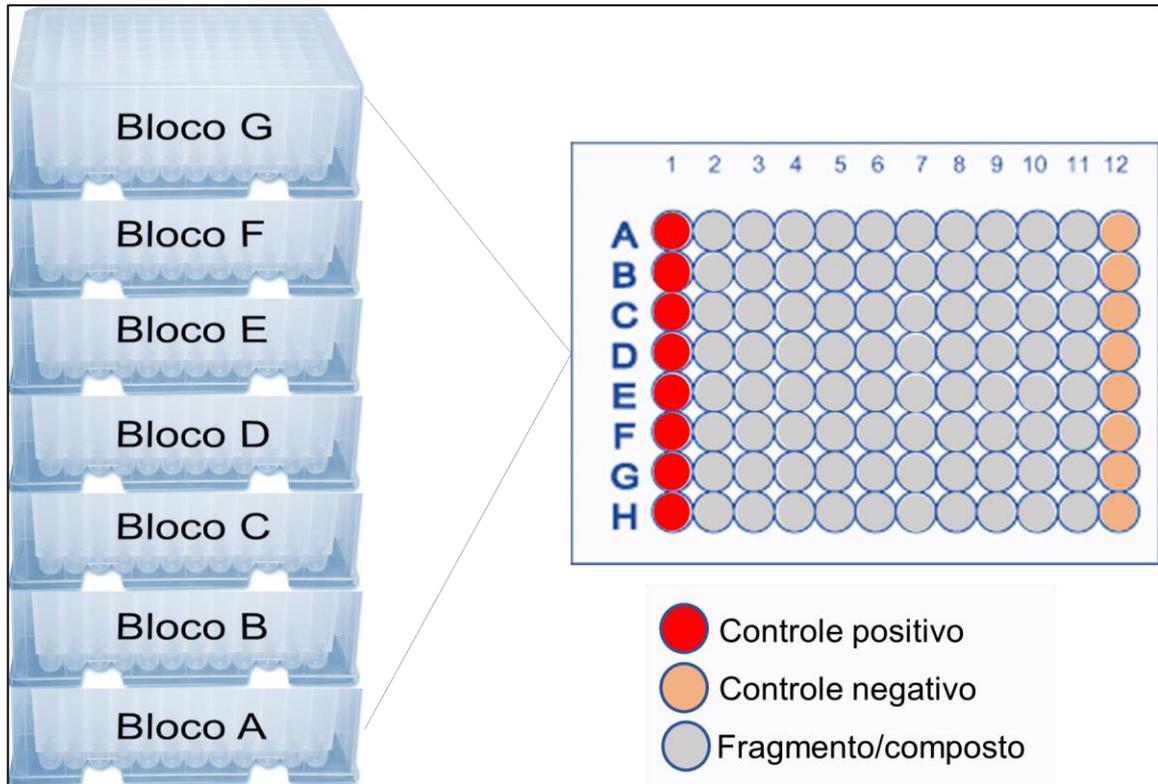


Figura 22 Disposição da biblioteca de fragmentos e compostos do professor Marcio Dias, ICB-USP. Cada bloco é composto por 80 compostos/fragmentos e 16 controles, foram utilizados 8 controles de proteína, sem a presença de DMSO e 8 controles de proteína na presença de DMSO na mesma quantidade do composto a ser utilizado.

Fonte: Elaborada pela autora.

No gráfico apresentado na Figura 23 exibe-se a média dos valores do controle (proteína com DMSO) assim como o desvio padrão, em cada um dos blocos que constituíam a biblioteca a ser testada, mostrando-se a contundência dos valores de T_m obtidos em cada placa avaliada.

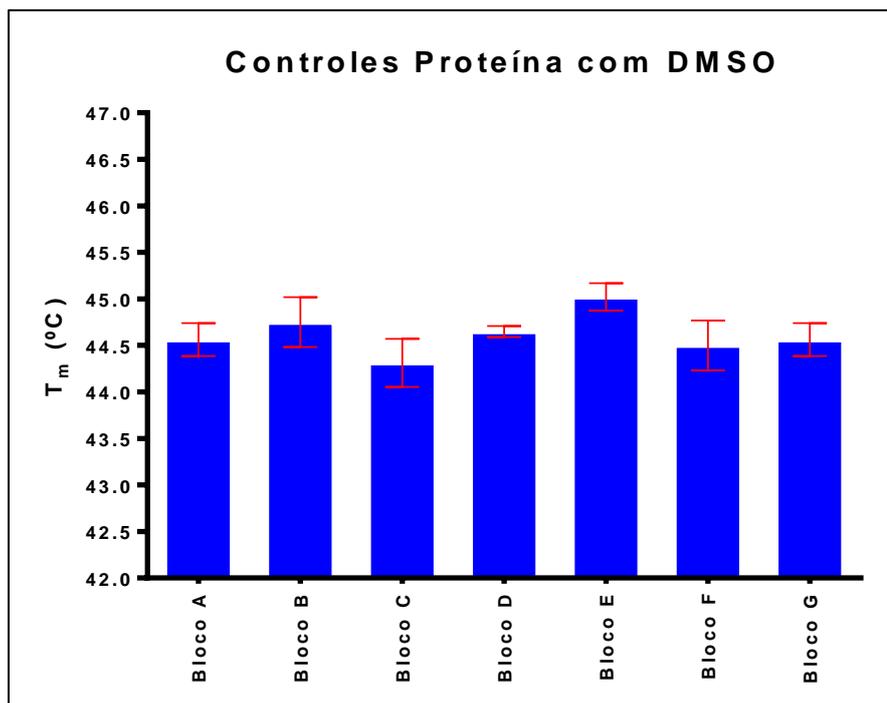


Figura 23 Média dos T_{ms} da proteína NS5 domínio RdRp_truncada na presença de DMSO como controle em cada um dos blocos testados. Em cada bloco foram realizadas 8 amostras da proteína com DMSO. As barras em azul mostram o valor médio dessas medidas e a barra em vermelho o desvio padrão assumido.

Fonte: Elaborada pela autora.

Pode-se, assim, calcular a média da temperatura de midpoint (T_m) da proteína, sobre todos os blocos de experimentos realizados, que resulta em $T_m = 44,6 \text{ }^\circ\text{C}$, $\pm 0,2$ (média dos valores médios de cada placa de experimento).

Seguiu-se, então, com a análise dos compostos triados bloco a bloco. Adotando como critério de seleção de *hits* aqueles compostos capazes de aumentar o T_m em pelo menos três vezes o valor do desvio padrão, estes dados estão dispostos na Figura 24.

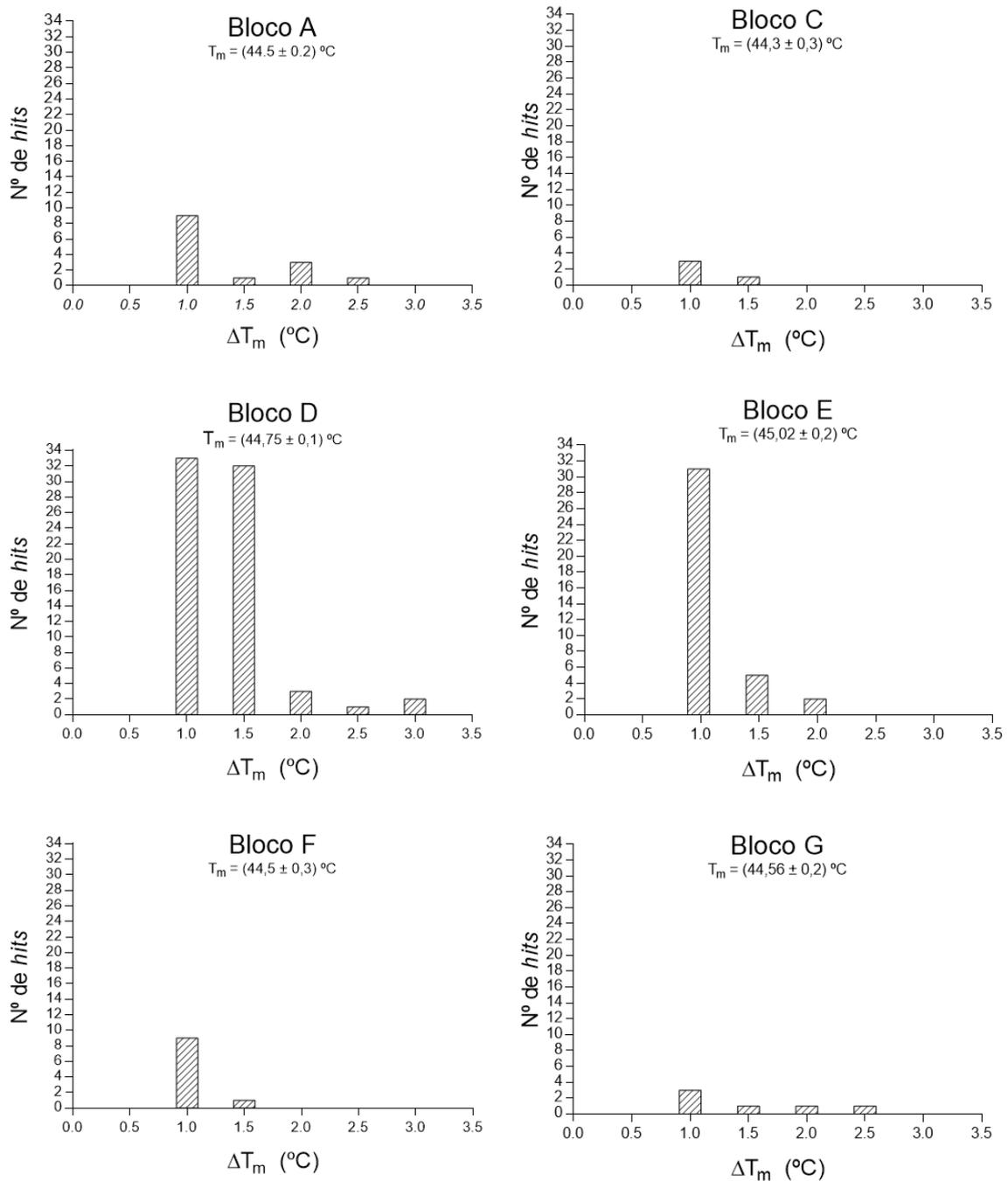


Figura 24 Análises realizada por blocos, onde estão apresentados os T_m de cada bloco com seu respectivo desvio padrão, em y estão o número de *hits* e em x, o ΔT_m (°C) para os *hits*.
 Fonte: Elaborada pela autora.

Ao todo, foram obtidos 143 *hits* que ocasionaram alteração positiva do T_m da proteína de 44,5°C para: 45,5 °C (21 fragmentos e 36 compostos, 57 no total); 46 °C (14 fragmentos e 53 compostos, 67 no total); 46,5 °C (6 fragmentos e 6 compostos, 12 no total); 47 °C (2 fragmentos e 2 compostos, 4 no total) e 47,5 °C (3 compostos).

Não houve nenhum *hit* no bloco B que alterou o T_m da proteína significativamente. Como o intuito desta colaboração é a obtenção de estruturas cristalográficas com ligantes, selecionou somente os compostos que aumentaram a T_m da proteína, deixando-a mais estável. Os compostos que diminuem essa a T_m , desestabilizam-na,¹²⁹⁻¹³⁰ o que dificultaria sua cristalização e, conseqüentemente, a obtenção de complexos cristalográficos, e portanto foram excluídos.

Todos os compostos que deslocaram o T_m da proteína foram submetidos ao ensaio de atividade enzimática, para analisar se podem inibir a atividade da proteína. Ressaltamos que para esse ensaio fizemos o uso da construção RdRp_integral.

Por serem fragmentos e pequenas moléculas, tipicamente de menor afinidade, esses compostos foram testados frente a proteína em uma concentração de composto (200 μ M), maior do que a usada para compostos *drug-like*, visto que em baixas concentrações, dificilmente serão capazes de inibir a atividade proteica.

O resultado do ensaio de inibição pode ser observado na Figura 25A, na qual estão dispostos os resultados de inibição dos 143 compostos que alteraram a estabilidade térmica da proteína. Ao final, 9 compostos inibiram mais de 50% da atividade e foram submetidos a um segundo ensaio de ponto único a uma concentração de 20 μ M (concentração usada para analisar a inibição dos demais inibidores encontrados neste trabalho). Desses, somente 3 mostraram algum tipo de inibição na proteína, como apresentado na Figura 25B.

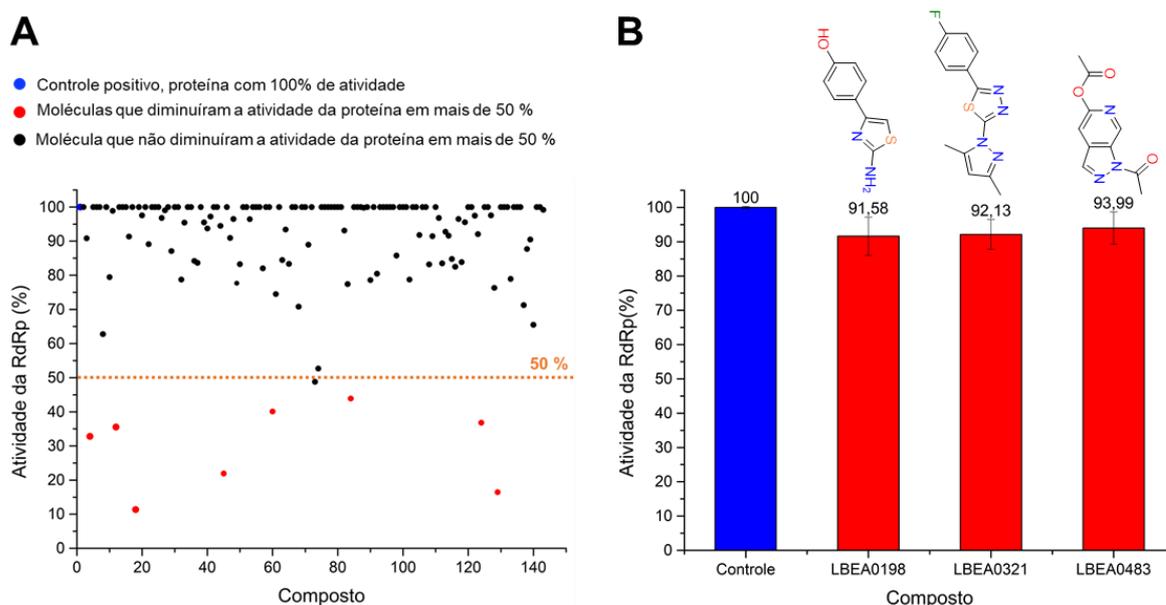


Figura 25 Gráficos do ensaio de inibição dos compostos que diminuíram a estabilidade da proteína pelo DSF. (A) Resultado dos cento e quarenta e três compostos a 200 μM frente a proteína, dos quais, somente nove foram capazes de diminuir a atividade da proteína em mais de 50 %. (B) Esses nove compostos foram submetidos a uma segunda rodada de ensaio de inibição, dessa vez a concentração dos mesmos foi de 20 μM , porém somente 3 deles inibiram levemente a atividade da proteína nessa concentração.

Fonte: Elaborada pela autora.

Comparando ambos os experimentos, podemos observar que apesar de inibirem a atividade da proteína em uma concentração alta de composto (200 μM), quando essa concentração é diminuída, não apresentam uma inibição significativa na atividade da proteica.

Das nove moléculas somente três, LBEA0198 inibiu 8,42%, LBEA0321 inibiu 7,87% e o LBEA0483 inibiu 6% a atividade enzimática, como podemos observar na Figura 25B, na qual um gráfico de barra demonstra a molécula e a inibição que apresenta na atividade relativa da proteína.

Essas três moléculas satisfazem as cinco regras de Lipinski ($\log P$ (lipofilicidade) <5 – que descreve P como o coeficiente de partição entre octanol e água –, Massa Molecular (MM) <500 Da, número de ligações de hidrogênio ≤ 5 , número de doadores e aceptores de hidrogênio ≤ 10), pois segundo Lipinski, moléculas que se encontram fora dessa faixa possuem uma menor probabilidade de serem absorvidas por via oral.¹³¹ Um outro critério indicador de bons fragmentos no desenvolvimento de fármacos por FBDD é a “regra dos três” (RO3),¹³² que sugere o uso de fragmentos com MM < 300 Da, $\log P < 3$, número de átomos doadores e

aceptores de hidrogênio ≤ 3 , número de ligações rotacionáveis ≤ 3).¹³³ Estes parâmetros físico-químicos foram calculados para os compostos que identificamos por DSF neste trabalho, para tanto utilizando a plataforma *SwissADME*¹³⁴. Os resultados destes cálculos estão apresentados na Tabela 5.

Tabela 5 Propriedades físico-químicas das moléculas encontradas na triagem de fragmentos e compostos do LBEA.

Molécula	Massa molecular	log <i>P</i>	Nº de átomos doadores H	Nº de átomos aceptores H
LBEA0198	192,24	1.8	2	2
LBEA0321	374,32	3.39	0	4
LBEA0483	219,20	0.95	0	5

Fonte: Elaborada pela autora, baseada nos dados gerados pelo *SwissADME*.

O IC₅₀ dessas moléculas se encontra entre 200 e 20 μ M, o que é típico para fragmentos e compostos de baixo peso molecular, pois fragmentos geralmente possuem menos átomos pesados (não-H), assim não são capazes de formar múltiplas interações atrativas com a superfície da molécula, ao contrário das moléculas maiores.^{133,135}

Apesar dos nossos esforços, não foi possível obter estruturas cristalográficas desses compostos complexadas com a RdRp, seja pelo método de “*soaking*” ou por co-cristalização. A obtenção dessas estruturas cristalográficas permitiria a identificação do modo de ligação, possibilitando, assim, a proposição de modificações baseadas na estrutura do ligante a fim de potencializar o potencial de inibição frente a proteína.

4.4.3 Triagem de ligantes a partir de técnicas computacionais

Neste trabalho também estabelecemos colaborações com outros grupos que se dispuseram a prover coleções de compostos para serem testados, como a coleção CIBFar (todos os compostos já obtidos ou sintetizados nos laboratórios participantes do Centro de Pesquisa e Inovação em Biodiversidade e Fármacos – CIBFar/CEPID) e a iniciativa OpenZika (openzika.ufg.br),¹⁰⁰ coordenada pela Profa. Dra. Carolina Horta Andrade da Universidade Federal de Goiás. A estratégia, neste caso, foi pré-selecionar compostos utilizando as técnicas de triagem virtual e posteriormente verificar a inibição a partir do ensaio de atividade enzimática.

4.4.3.1 Colaboração CIBFar

A primeira colaboração foi com o professor Dr. Rafael Guido (IFSC-USP) integrante do CIBFar, que é um CEPID apoiado pela FAPESP envolvendo físicos, químicos de produtos naturais, químicos sintéticos, farmacêuticos e biólogos, entre outros. O CIBFar possui um extenso banco de dados de moléculas, que consiste em aproximadamente 4000 moléculas de origem sintética ou isoladas de produtos naturais.

Neste sentido, todas as moléculas já identificadas ou sintetizadas pelos grupos integrantes do CIBFar foram integradas em um mesmo banco de dados, selecionando previamente as moléculas que se encaixam nas cinco regras de Lipinski ¹³¹. Este banco de dados foi utilizado em uma campanha de triagem virtual, utilizando a nossa estrutura cristalográfica obtida com a construção RdRp_truncada (código PDB 5U04)¹⁰² como alvo, com busca específica para o sítio ativo, *grid box* na região dos resíduos Asp535, Asp665 e Asp666 e para o sítio alostérico na região *N-pocket*, *grid box para a região do priming loop* (resíduos Thr796 a Lys802). A metodologia utilizada nesta busca está descrita no item 3.12 acima.

A partir dessa triagem dois compostos mais promissores foram priorizados, os pesquisadores do CIBFar foram contactados e nos enviaram as moléculas disponíveis na quantidade (~1mg) e pureza (~95%) minimamente necessárias, para que pudéssemos testar sua inibição frente ao ensaio de atividade enzimática da RdRp_integral.

Foi possível encontrar dois compostos que inibiram significativamente a atividade relativa da proteína e que foram submetidos a ensaio de dose resposta para determinação do IC₅₀. Esses compostos inibiram mais de 80% da atividade relativa da proteína no ensaio de inibição enzimática de ponto único e são provenientes do Núcleo de Bioensaios, Biossíntese e Ecofisiologia de Produtos Naturais – NuBBE, Instituto de Química de Araraquara da UNESP. Os resultados destes compostos podem ser visualizados na Figura 26.

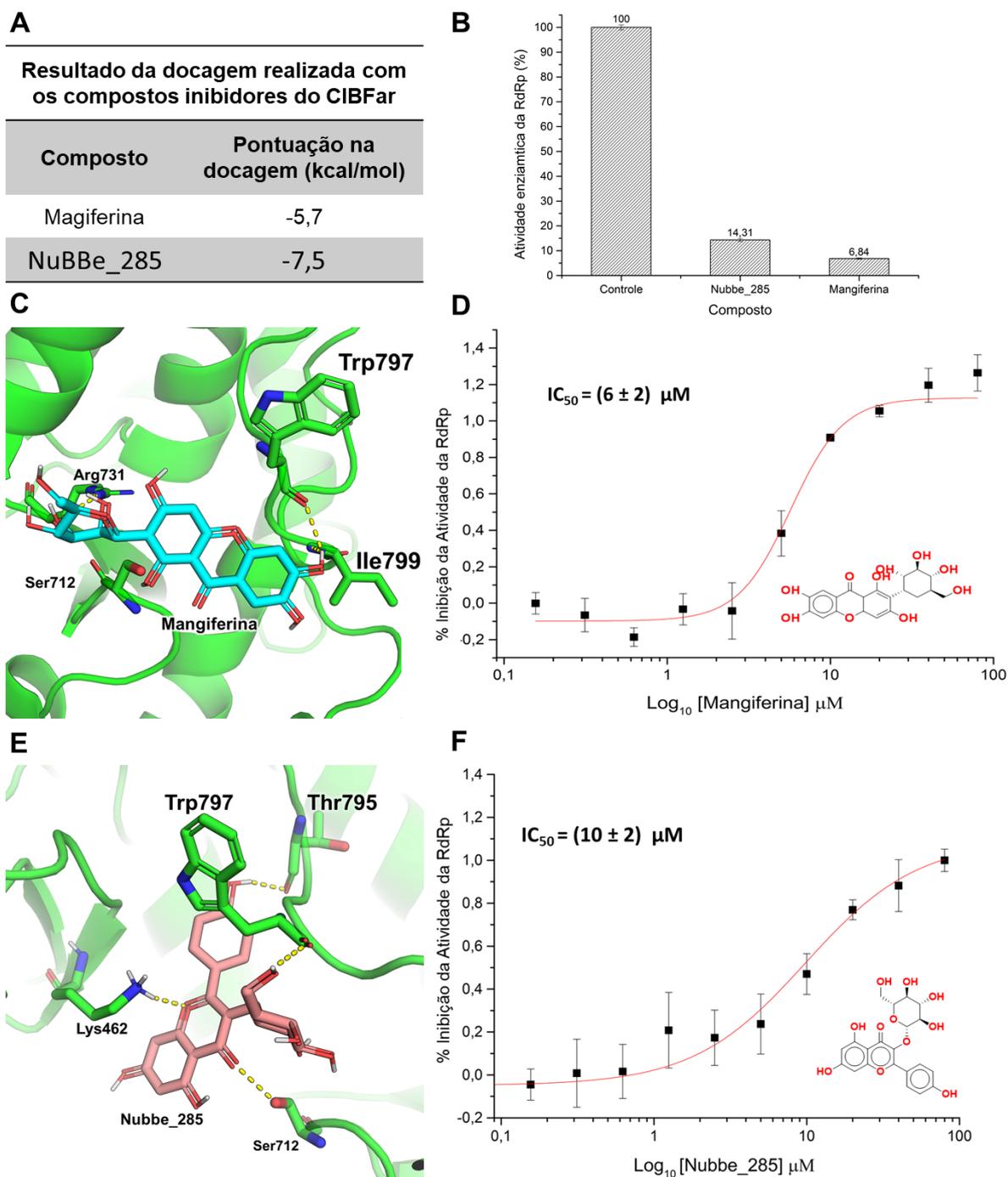


Figura 26 Resultado da triagem computacional e do ensaio enzimático da colaboração CIBFar. (A) Tabela com o resultado da melhor simulação realizada para os compostos Mangiferina e Nubbe_285. (B) Gráfico de barras da atividade enzimática da RdRp e a inibição da atividade dos compostos Nubbe_285 e Magiferina. (C) Imagem gerada pelo Pymol, onde é possível observar os resíduos que interagem com o composto Mangiferina (ciano), (D) gráfico de dose-resposta do composto Mangiferina com o IC_{50} de $6 \pm 2 \mu M$ (E) Imagem gerada pelo Pymol, onde é possível observar os resíduos que interagem com o composto Nubbe_285 (salmão), (F) gráfico de dose-resposta do composto Nubbe_285 com o IC_{50} de $10 \pm 2 \mu M$.

Fonte: Elaborada pela autora.

Dois compostos dessa colaboração inibiram a atividade da proteína abaixo de 10 μM , foram eles a Mangiferina IC_{50} de $6 \pm 2 \mu\text{M}$ (Figura 26D) e o Nubbe_285 IC_{50} de $10 \pm 2 \mu\text{M}$ (Figura 26F).

A mangiferina é um bioativo natural, isolado inicialmente da *Mangifera indica* L, a mangueira, e foi a primeira xantona a ser farmacologicamente investigada.¹³⁶ É usada como um fitomedicamento anti-inflamatório e tem atividades biológicas amplamente estudadas contra diversas classes de doenças,^{137–139} entre elas antialérgica, antitumoral, antioxidante, antiviral, antimicrobiana, entre outras. Apesar de sua baixa biodisponibilidade oral, seu potencial terapêutico e sua baixa toxicidade tem despertado grande interesse.^{140–143}

O composto Nubbe_285, também conhecido como trifolin ou kaempferol 3-galactoside, é um flavonoide da classe dos flavonóis, isolado em 1977,¹⁴⁴ derivado do kaempferol, que é um flavonol amplamente estudado como um antitumoral.¹⁴⁵ O trifolin é relatado como um antifúngico e vem sendo estudado como antitumoral de baixa toxicidade.¹⁴⁶⁻¹⁴⁷

A modelagem por triagem virtual indica que ambos os compostos podem interagir com a região da RdRp conhecida como *Priming Loop*, no resíduo triptofano (Trp) 797. A Mangiferina ainda interage com o resíduo Isoleucina (Ile) 799 e o Nubbe_285 com o resíduo Treonina (Thr) 795 do *Priming Loop*. Essa simulação pode ser usada para que novos compostos possam vir a ser desenhados a fim de que a inibição dessas moléculas frente a proteína seja potencializada, visto que produtos naturais tem grande potencial para o desenvolvimento de agentes antivirais por serem moléculas com baixa toxicidade e amplamente estudadas para diversas finalidades.¹⁴⁸⁻¹⁴⁹

4.4.3.2 Colaboração OpenZika

A colaboração foi realizada com o projeto OpenZika (OZ, ver <http://openzika.ufg.br>), coordenado pela professora Dra. Carolina Horta Andrade, do LabMol – Laboratório de Planejamento de Fármacos e Modelagem Molecular – UFG. O objetivo principal do projeto é realizar a triagem virtual de milhares de compostos disponíveis comercialmente em bases de dados nacionais e internacionais, de acesso aberto, com modelos estruturais das proteínas do vírus Zika determinados experimentalmente ou obtidos com modelagem por homologia.¹⁰⁰

A proteína alvo, os sítios selecionados para a docagem molecular, os programas e os parâmetros computacionais utilizados foram os mesmos para ambas as colaborações computacionais, para que obtivéssemos resultados comparáveis. Os bancos de dados utilizados nesta triagem computacional foram: a Biblioteca *Chembridge*, composto por 1 milhões de moléculas que estão disponíveis comercialmente e a biblioteca de produtos naturais do professor Dr. Luis Octávio Regasini (IBILCE-UNESP São José do Rio Preto), composta por 150 moléculas naturais e semissintéticas de *Pterogyne nitens*.

Devido ao tamanho da biblioteca de *Chembridge* alguns filtros foram utilizados para garantir que somente os compostos que apresentassem as propriedades semelhantes a fármacos (drug-like) fossem selecionados após a triagem. Para isso foi utilizado o programa FILTER versão 1.9, esses filtros satisfazem as regras de Lipinsk, removem grupos funcionais considerados tóxicos e reativos e os compostos PAINS (do inglês *Pan Assay Interference Compounds*), que podem interferir nos resultados em ensaios experimentais, como falso-positivos. ¹⁵⁰⁻

151

Os melhores compostos foram adquiridos pelo grupo da Profa. Carolina H. Andrade e nos foram enviados para os ensaios de inibição enzimática. Foram feitos inicialmente os ensaios de ponto único, com uma concentração de composto fixada em 20 μM . Para os compostos que inibiram pelo menos 80% da atividade enzimática foram feitos os ensaios de concentração-resposta para determinar o IC_{50} de cada composto.

O resultado dessa colaboração pode ser visualizado na Tabela 6, na qual estão dispostos os compostos com seus respectivos IC_{50} (μM) e valor da pontuação de docagem (kcal/mol).

Tabela 6 Resultado da colaboração OZ onde estão apresentados os valores de inibição IC_{50} (μM) e o valor da pontuação de docagem (kcal/mol).

Composto	IC_{50} (μM)	Pontuação na docagem (kcal/mol)
LabMol-346	16 ± 2	-5.2
LabMol-202	15 ± 4	-5.7
LabMol-204	12 ± 3	-6.8
LabMol-340	6.9 ± 0.9	-9.0
LabMol-317	5 ± 1	-7.9
LabMol-344	5 ± 1	-7.5
LabMol-319	1.8 ± 0.5	-7.4
LabMol-301	1 ± 0.1	-7.2
LabMol-309	0.7 ± 0.1	-7.6
LabMol-343	0.48 ± 0.09	-6.5

Fonte: Elaborada pela autora.

Para todos os compostos, foi analisado se algum deles provocaram interferência com alguns dos componentes do tampão ou com o fluoróforo SYBR Green I, usando a fórmula de IF, como relatado na padronização do ensaio para triagem em larga escala. Foi possível determinar a atividade inibitória de 10 compostos, conforme observado na Tabela 6 e suas estruturas químicas podem ser visualizadas na Figura 27.

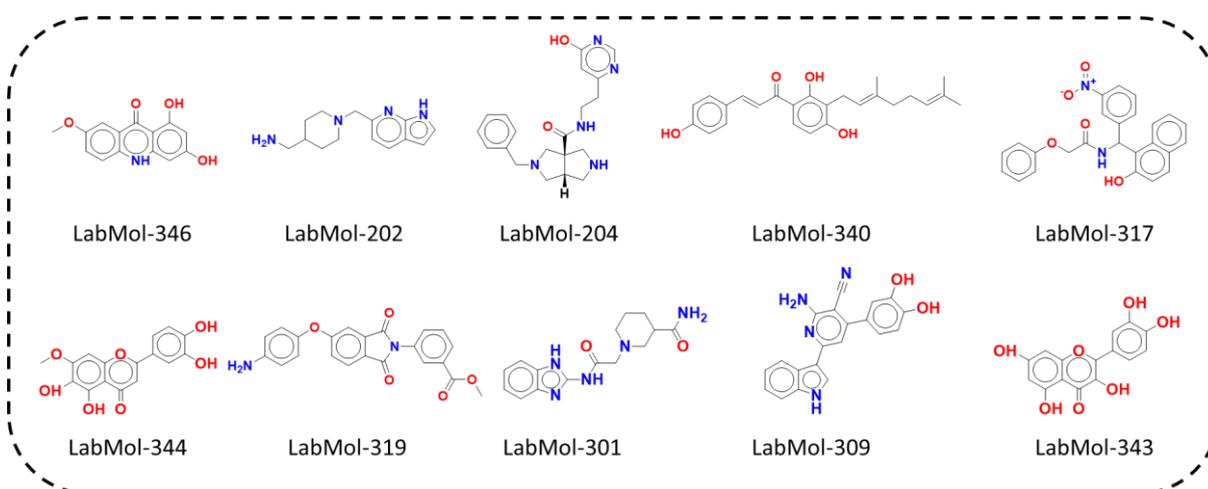


Figura 27 Estruturas químicas dos compostos LabMol, para os quais foi possível determinar a atividade inibitória.

Fonte: Elaborada pela autora.

Vemos que dos dez compostos encontrados, sete estão com inibição abaixo de 10 μM e três possuem uma inibição abaixo de 1 μM . As curvas de concentração-resposta para esses três compostos mais potentes podem ser visualizadas na Figura 28.

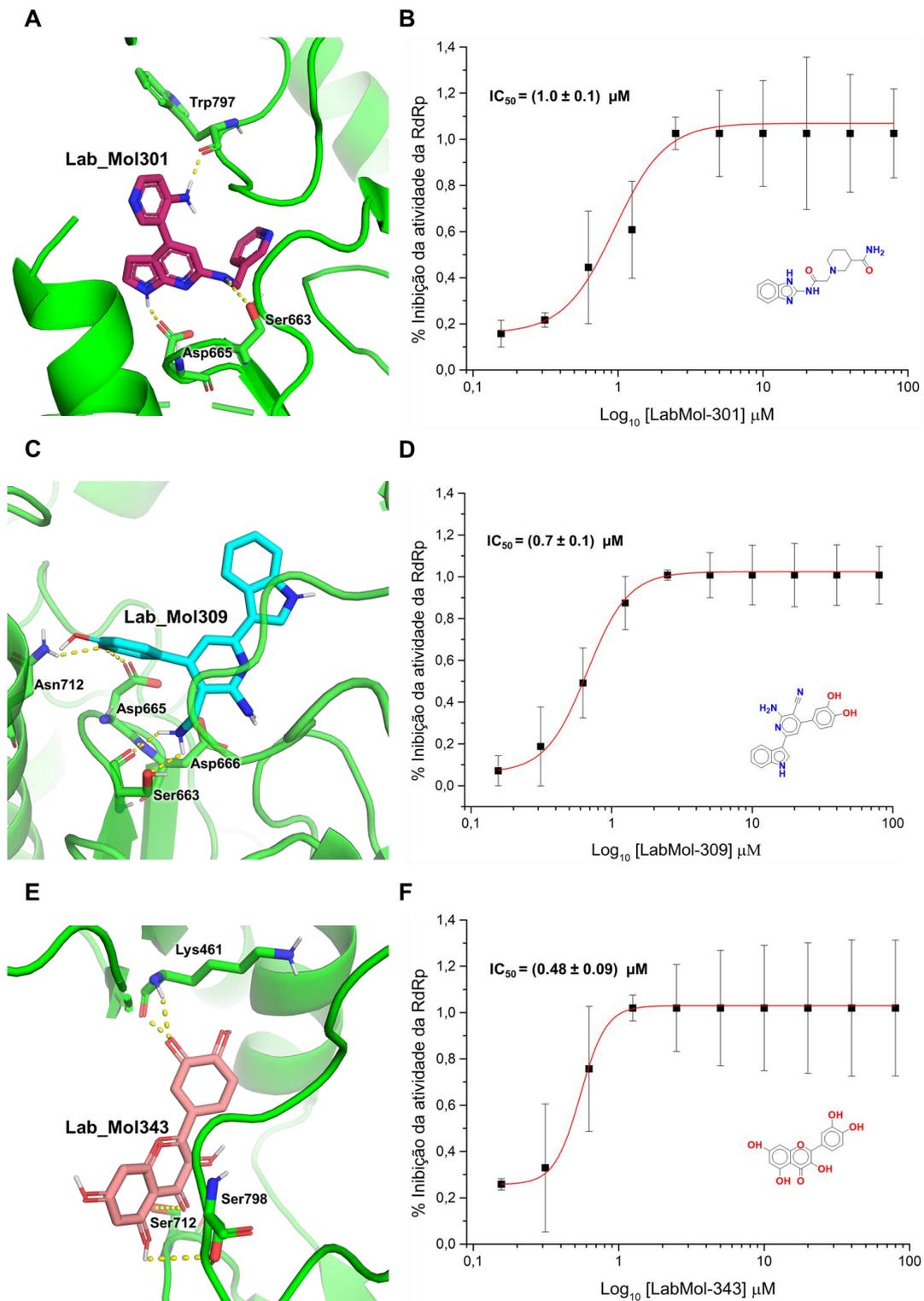


Figura 28 Resultado da triagem computacional e do ensaio enzimático da colaboração OpenZika, para os três compostos com $IC_{50} \geq 1 \mu M$. (A) Imagem gerada pelo Pymol, onde é possível observar os resíduos que interagem com o composto LabMol-301 (rosa) e (B) sua respectiva curva de concentração-resposta, com o IC_{50} de $1 \pm 0.1 \mu M$ (C) Imagem gerada pelo Pymol, onde é possível observar os resíduos que interagem com o composto LabMol-309 (ciano) e (D) sua respectiva curva de concentração-resposta, com o IC_{50} de $0.7 \pm 0.1 \mu M$ (E) Imagem gerada pelo Pymol, onde é possível observar os resíduos que interagem com o composto LabMol-343 (salmão) e (F) sua respectiva curva de concentração-resposta, com o IC_{50} de $0.48 \pm 0.09 \mu M$.

Fonte: Elaborada pela autora.

Pela predição da triagem, os compostos LabMol-301 e LabMol-309 virtual interagem diretamente com os aspartatos catalíticos (D665 e D666). Como observação interessante, o composto LabMol-309 parece ter maior potencial de inibição pois a modelagem prevê que ele interage com ambos os aspartatos catalíticos.

O composto LabMol-343, assim como o composto Nubbe_285 e o LabMol-344, é um flavonol, conhecido como quercetina, um antioxidante amplamente estudado com diversas atividades biológicas relacionadas, tais como anti-hipertensiva, antiarrítmica, antineoplásica, antimicrobiana, anti-inflamatória e antiviral.^{152–155}

Uma análise das curvas de concentração-resposta dos compostos análogos e derivadas do anel flavonol, LabMol-343, Nubbe_285 e LabMol-344 assim como suas respectivas estruturas químicas podem ser visualizadas na Figura 29.

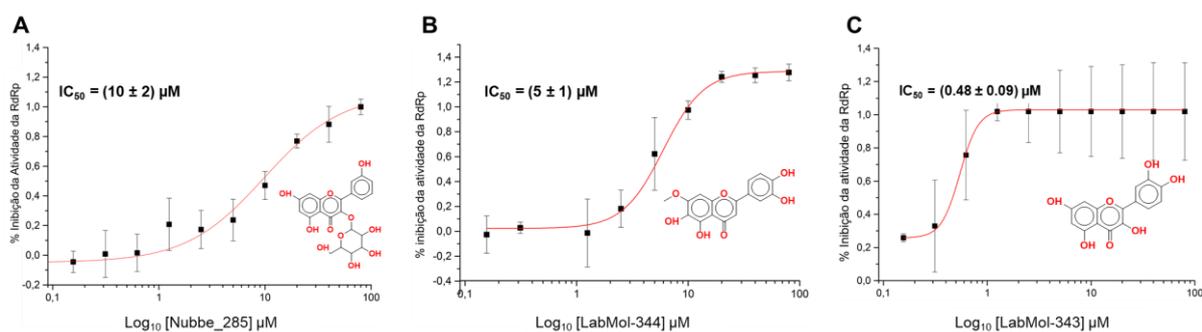


Figura 29 Curvas de concentração-resposta dos compostos Nubbe_285 (A), LabMol-344 (B) e LabMol-343 (C). As estruturas químicas dos compostos são análogas e derivadas do anel flavonol.

Fonte: Elaborada pela autora.

Correlacionando esses compostos, mostramos que modificações baseadas no anel flavonol auxiliaram-nos na busca por um potente inibidor para a RdRp e de baixa toxicidade, visto que eles apresentam baixa toxicidade nos ensaios relatados na literatura.

Para todos esses compostos ativos contra a RdRp, seria de extremo interesse a obtenção de cristais com a presença do ligante, o que, apesar dos nossos esforços, até o presente momento não foi alcançado.

Como perspectivas deste projeto, todas as moléculas deverão ser submetidas a ensaios de toxicidade pelos colaboradores, assim como passar por ensaios biofísicos e cristalográficos para a validação de sua ligação à proteína.

5 Conclusão

Com a realização deste trabalho, foi possível estabelecer protocolos de expressão e purificação para ambas as construções propostas para a proteína RdRp de ZIKV.

Para a construção RdRp_truncada foi possível obter cristais que ocasionaram difração, possibilitando a resolução da estrutura cristalográfica a 1,9 Å. Os resultados obtidos com essa construção foram publicados na revista *Nature Communications*. Além disso, essa estrutura cristalográfica mostrou-se promissora para a busca por novos ligantes no sítio alostérico *N-pocket*, a partir de ensaios virtuais e biofísicos.

Ademais, foi possível desenvolver e padronizar um ensaio baseado no marcador fluorescente de ácido nucleico SYBR Green I, para a construção RdRp_integral, que monitora a atividade da proteína em tempo real, o qual foi utilizado na busca por inibidores em larga escala, utilizando moléculas provenientes de diversas bibliotecas, como a do Pandemic Response Box (Medicines for Malaria Venture - MMV) (PRB), moléculas sintéticas e provenientes de produtos naturais.

Com a estrutura cristalográfica em alta resolução e com o ensaio de atividade padronizado, foi possível estabelecer colaborações para a busca por possíveis inibidores da RdRp de ZIKV a partir de técnicas computacionais.

Ao todo, analisou-se a inibição de mais de 500 moléculas contra a RdRp de ZIKV, das quais 17 se mostraram inibidoras da atividade enzimática. Dessas, 5 são provenientes da biblioteca PRB, 2 da colaboração CIBFar, provenientes de produtos naturais isolados pelos grupos Nubbe e 10 da colaboração OpenZika, proveniente de diversas bibliotecas sintéticas e de produtos naturais. Algumas dessas moléculas apresentaram inibição na faixa de micromolar, sendo, três entre 20 e 10 μM , dez delas entre 10 e 1 μM e cinco (2 do PRB e 2 OZ) na faixa de nanomolar, ou seja, com potencial de inibição $< 1 \mu\text{M}$.

Dentre esses compostos encontramos 3 que possuem como anel principal um flavonol, derivados de produtos naturais, o Nubbe_285 com IC_{50} de $10 \pm 2 \mu\text{M}$, o LabMol-344 com IC_{50} de $5 \pm 1 \mu\text{M}$ e o LabMol-343 com IC_{50} de $0,48 \pm 0,09 \mu\text{M}$. Modificações baseadas na estrutura do ligante, mantendo o anel flavonol, podem ser úteis para propor uma nova classe de inibidores de RdRp.

Além disso, uma molécula altamente potente contra a RdRp foi encontrada na biblioteca PRB, o MMV687800, com IC_{50} de $0,47 \pm 0,17 \mu M$. Essa molécula, a Clofazamina, pode ser aplicada à técnica de reposicionamento de fármacos, por já ser utilizada no tratamento da Hanseníase, assim que estudos celulares e em animais comprovarem sua efetividade. Se demonstrada sua atividade *in vitro* e *in vivo* (modelos animais) contra o ZIKV, este composto poderia ser diretamente utilizado em ensaios clínicos de eficácia, visto que seu perfil de segurança já é conhecido. Isso diminuiria os custos referentes às pesquisas, assim como o tempo de aprovação pelos órgãos regulatórios.

Portanto, analisando-se todos os resultados obtidos neste período, concluímos que os mesmos são potencialmente importantes para o desenvolvimento de fármacos para o tratamento das infecções com o vírus Zika. A plataforma obtida para avaliação da inibição da enzima RdRp mostrou ser uma ferramenta confiável e eficaz para triagem de compostos. Com estes ensaios, aliados às informações estruturais da RdRp, os compostos obtidos mostram-se promissores para desenvolvimento futuro, aliados a ensaios celulares e em modelos animais.

Referências

- 1 HEINZ, F. X.; STIASNY, K. Flaviviruses and their antigenic structure. *Journal of Clinical Virology*, v. 55, n. 4, p. 289-295, 2012.
- 2 ROSSI, S. L. *et al.* Characterization of a novel murine model to study Zika virus. *The American Journal of Tropical Medicine and Hygiene*, v. 94, n. 6, p. 1362-1369, 2016.
- 3 FAYE, O. *et al.* Molecular evolution of Zika virus during its emergence in the 20 th century. *PLoS Neglected Tropical Diseases*, v. 8, n. 1, p. e2636, 2014..
- 4 GORSHKOV, K. *et al.* Zika virus: origins, pathological action, and treatment strategies. *Frontiers in Microbiology*, v. 9, p. 3252, 2019..
- 5 ENSERINK, M. An obscure mosquito-borne disease goes global. *Science*, v. 350, n. 6264, p. 1012-1013, 2015.
- 6 MLAKAR, J. *et al.* Zika virus associated with microcephaly. *New England Journal of Medicine*, v. 374, n. 10, p. 951-958, 2016.
- 7 CAO-LORMEAU, V. *et al.* Guillain-Barré Syndrome outbreak associated with Zika virus infection in French Polynesia: a case-control study. *The Lancet*, v. 387, n. 10027, p. 1531-1539, 2016.
- 8 OEHLER, E. *et al.* Zika virus infection complicated by Guillain-Barre syndrome—case report, French Polynesia, December 2013. *Eurosurveillance*, v. 19, n. 9, p. 20720, 2014.
- 9 IOOS, S. *et al.* Current Zika virus epidemiology and recent epidemics. *Medicine et Maladies Infectieuses*, v. 44, n. 7, p. 302-307, 2014.
- 10 SHARMA, A.; LAL, SUNIL K. Zika virus: transmission, detection, control, and prevention. *Frontiers in Microbiology*, v. 8, p. 110, 2017.
- 11 FARIA, N. R. *et al.* Zika virus in the Americas: early epidemiological and genetic findings. *Science*, v. 352, n. 6283, p. 345-349, 2016.
- 12 BRASIL, P. *et al.* Zika virus infection in pregnant women in Rio de Janeiro. *New England Journal of Medicine*, v. 375, n. 24, p. 2321-2334, 2016.
- 13 DE MELO FREIRE, C. C. *et al.* Spread of the pandemic Zika virus lineage is associated with NS1 codon usage adaptation in humans. *BioRxiv*, p. 032839, 2015. DOI 10.1101/032839.
- 14 CENTERS OF DISEASE CONTROL AND PREVENTION. *Areas with risk of dengue*.2019. Disponível em: <https://www.cdc.gov/dengue/areaswithrisk/index.html>. Acessado em:14/11/2019
- 15 VENTURA, C. V. *et al.* Zika virus in Brazil and macular atrophy in a child with microcephaly. *The Lancet*, v. 387, n. 10015, p. 228, 2016.
- 16 CUGOLA, F. R. *et al.* The Brazilian Zika virus strain causes birth defects in experimental models. *Nature*, v. 534, n. 7606, p. 267-271, 2016.
- 17 VOGEL, G. Experts fear Zika's effects may be even worse than thought. *Science*, v. 352, n. 6292, p. 1375-1376, 2016.
- 18 PACHECO, O. *et al.* Zika virus disease in Colombia—preliminary report. *New England Journal of Medicine*, v. 383, n. 6, p. e44, 2020.
- 19 COHEN, J. The race for a Zika vaccine is on. *Science*, v. 351, n. 6273, p. 543-544, 2016.
- 20 BRASIL. Ministerio da Saude.Secretaria de Vigilância em Saúde. (SVS/MS). Monitoramento dos

casos de Arboviroses urbanas transmitidas pelo Aedes (dengue, chikungunya e Zika), Semanas Epidemiológicas 01 a 52. *Boletim Epidemiológico Arboviroses*, v.51,n.24, p.1–13 , 2019.

21 BRASIL. Ministerio da Saude.Secretaria de Vigilância em Saúde (SVS/MS). Monitoramento dos casos de Arboviroses urbanas transmitidas pelo Aedes (dengue, chikungunya e Zika), Semanas Epidemiológicas 01 a 50. *Boletim Epidemiológico Arboviroses*, v.51, p.1–13 , 2020.

22 BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. Situação epidemiológica da síndrome congênita associada à infecção pelo vírus Zika em 2020: até a SE 25. *Boletim Epidemiológico* v.51, p.23–28 , 2020.

23 KASPRZYKOWSKI, J. I. *et al.* A recursive sub-typing screening surveillance system detects the appearance of the ZIKV African lineage in Brazil: Is there a risk of a new epidemic? *International Journal of Infectious Diseases*, v. 96, p. 579-581, 2020.

24 SHEN, S. *et al.* Phylogenetic analysis revealed the central roles of two African countries in the evolution and worldwide spread of Zika virus. *Virologica Sinica*, v. 31, n. 2, p. 118-130, 2016.

25 WANG, L. *et al.* From mosquitos to humans: genetic evolution of Zika virus. *Cell Host & Microbe*, v. 19, n. 5, p. 561-565, 2016.

26 KAMELIAN, K. *et al.* Phylogenetic surveillance of travel-related Zika virus infections through whole-genome sequencing methods. *Scientific Reports*, v. 9, n. 1, p. 1-10, 2019.

27 HAN, J. F. *et al.* Homologous recombination of Zika viruses in the Americas. *Journal of Infection*, v. 73, n. 1, p. 87-88, 2016.

28 ALENCAR, J. *et al.* Natural infection and vertical transmission of two flaviviruses (Yellow fever and Zika) in mosquitoes in primary forests in the Brazilian state of Rio de Janeiro (Diptera: Culicidae). *BioRxiv*, p. 688713, 2019. DOI:10.1101/688713.

29 DE ALMEIDA, P. R. *et al.* Detection of a novel African-lineage-like Zika virus naturally infecting free-living neotropical primates in Southern Brazil. *BioRxiv*, p. 828871, 2019. DOI:10.1101/828871.

30 SIROHI, D. *et al.* The 3.8 Å resolution cryo-EM structure of Zika virus. *Science*, v. 352, n. 6284, p. 467-470, 2016.

31 SIROHI, D.; KUHN, R. J. Zika virus structure, maturation, and receptors. *The Journal of Infectious Diseases*, v. 216, suppl.10, p. S935-S944, 2017.

32 KAUFMANN, B.; ROSSMANN, M. G. Molecular mechanisms involved in the early steps of flavivirus cell entry. *Microbes and Infection*, v. 13, n. 1, p. 1-9, 2011.

33 BOLDESCU, V. *et al.* Broad-spectrum agents for flaviviral infections: dengue, Zika and beyond. *Nature Reviews Drug Discovery*, v. 16, n. 8, p. 565-586, 2017.

34 NANTHINI, D. P.; CHITRALEKHA, S. Zika Virus : an emerging vector borne disease. *Research & Reviews: Journal of Medical and Health Sciences*, v.7, n.1, p.10–30 , 2018.

35 SINGH, R. K. *et al.* Zika virus—emergence, evolution, pathology, diagnosis, and control: current global scenario and future perspectives—a comprehensive review. *Veterinary Quarterly*, v. 36, n. 3, p. 150-175, 2016.

36 BARROWS, N. J. *et al.* Biochemistry and molecular biology of flaviviruses. *Chemical Reviews*, v. 118, n. 8, p. 4448-4482, 2018.

37 DIAMOND, M. S.; PIERSON, T. C. Molecular insight into dengue virus pathogenesis and its implications for disease control. *Cell*, v. 162, n. 3, p. 488-492, 2015.

38 KATO, F.; HISHIKI, T. Dengue virus reporter replicon is a valuable tool for antiviral drug discovery

and analysis of virus replication mechanisms. *Viruses*, v. 8, n. 5, p. 122, 2016.

39 GANJI, L. V.; KANYALKAR, M. A. Non-structural proteases as a target of dengue virus. *Journal of Antivirals & Antiretrovirals*. v.11, n.4, p.1–15 , 2019.

40 SHAN, C. *et al.* An infectious cDNA clone of Zika virus to study viral virulence, mosquito transmission, and antiviral inhibitors. *Cell Host & Microbe*, v. 19, n. 6, p. 891-900, 2016.

41 OLIVEIRA, A. S. *et al.* NS3 and NS5 proteins: important targets for anti-dengue drug design. *Journal of the Brazilian Chemical Society*, v. 25, n. 10, p. 1759-1769, 2014.

42 BUKREJEWSKA, M. *et al.* Crystal structures of the methyltransferase and helicase from the ZIKA 1947 MR766 Uganda strain. *Acta Crystallographica Section D: structural biology*, v. 73, n. 9, p. 767-774, 2017.

43 ZHANG, Z. *et al.* Crystal structure of unlinked NS2B-NS3 protease from Zika virus. *Science*, v. 354, n. 6319, p. 1597-1600, 2016.

44 YOKOKAWA, F. *et al.* Discovery of potent non-nucleoside inhibitors of dengue viral RNA-dependent RNA polymerase from a fragment hit using structure-based drug design. *Journal of Medicinal Chemistry*, v. 59, n. 8, p. 3935-3952, 2016. DOI:10.1021/acs.jmedchem.6b00143.

45 LI, C. *et al.* Cellular DDX3 regulates Japanese encephalitis virus replication by interacting with viral un-translated regions. *Virology*, v. 449, p. 70-81, 2014.

46 KLEMA, V. J.; PADMANABHAN, R.; CHOI, K. H. Flaviviral replication complex: coordination between RNA synthesis and 5'-RNA capping. *Viruses*, v. 7, n. 8, p. 4640-4656, 2015.

47 FAJARDO JUNIOR, T. *et al.* The flavivirus polymerase NS5 regulates translation of viral genomic RNA. *Nucleic Acids Research*, v. 48, n. 9, p. 5081-5093, 2020.

48 BARRAL, K. *et al.* Development of specific dengue virus 2'-O-and N7-methyltransferase assays for antiviral drug screening. *Antiviral Research*, v. 99, n. 3, p. 292-300, 2013.

49 POTISOPON, S. *et al.* The methyltransferase domain of dengue virus protein NS5 ensures efficient RNA synthesis initiation and elongation by the polymerase domain. *Nucleic Acids Research*, v. 42, n. 18, p. 11642-11656, 2014.

50 ZHAO, B. *et al.* Structure and function of the Zika virus full-length NS5 protein. *Nature Communications*, v. 8, n. 1, p. 1-9, 2017.

51 WANG, B. *et al.* Structure and function of Zika virus NS5 protein: perspectives for drug design. *Cellular and Molecular Life Sciences*, v. 75, n. 10, p. 1723-1736, 2018. DOI:10.1007/s00018-018-2751-x.

52 NG, K. K.-S.; ARNOLD, J. J.; CAMERON, C. E. Structure-function relationships among RNA-dependent RNA polymerases. *In: PADDISON, P. G.; VOGT, P. K. RNA interference*. Berlin: Springer, 2008. p. 137-156.

53 LIM, S. P.; NOBLE, C. G.; SHI, P. Y. The dengue virus NS5 protein as a target for drug discovery. *Antiviral Research*, v. 119, p. 57-67, 2015.

54 VENKATARAMAN, S.; PRASAD, B. V.; SELVARAJAN, R. RNA dependent RNA polymerases: insights from structure, function and evolution. *Viruses*, v. 10, n. 2, p. 76, 2018. DOI:10.3390/v10020076.

55 GLOBAL RESEARCH COLLABORATION FOR INFECTIOUS DISEASE PREPAREDNESS. *Ongoing and upcoming research on the Zika virus*. 2016. Disponível em: <https://www.glopid-r.org/wp-content/uploads/2016/05/findings-of-zika-research-working-groups-on-vaccines-aetiology-and-diagnostics.pdf>. Acesso em: 19.02.21.

- 56 LAROCCA, R. A. *et al.* Vaccine protection against Zika virus from Brazil. *Nature*, v. 536, n. 7617, p. 474-478, 2016.
- 57 KENNEDY, R. B. *et al.* Zika vaccines: current state. In: POLAND, G. A. (ed.) *Vaccinations*. New York:Elsevier, 2019. p. 75-88. DOI:10.1016/b978-0-323-55435-0.00005-7.
- 58 BAROUCH, D. H.; THOMAS, S. J.; MICHAEL, N. L. Prospects for a Zika virus vaccine. *Immunity*, v. 46, n. 2, p. 176-182, 2017.
- 59 POLAND, G. A.; OVSYANNIKOVA, I. G.; KENNEDY, R. B. Zika vaccine development: current status. *Mayo Clinic Proceedings*. Elsevier, v.94, n.12, p. 2572-2586,2019.
- 60 POLAND, G. A. *et al.* Development of vaccines against Zika virus. *The Lancet Infectious Diseases*, v. 18, n. 7, p. e211-e219, 2018.
- 61 BARROWS, N. J. *et al.* A screen of FDA-approved drugs for inhibitors of Zika virus infection. *Cell Host & Microbe*, v. 20, n. 2, p. 259-270, 2016.
- 62 BULLARD-FEIBELMAN, K. M. *et al.* The FDA-approved drug sofosbuvir inhibits Zika virus infection. *Antiviral Research*, v. 137, p. 134-140, 2017.
- 63 SHIRYAEV, S. A. *et al.* Repurposing of the anti-malaria drug chloroquine for Zika Virus treatment and prophylaxis. *Scientific Reports*, v. 7, n. 1, p. 1-9, 2017.
- 64 KAMIYAMA, N. *et al.* Ribavirin inhibits Zika virus (ZIKV) replication in vitro and suppresses viremia in ZIKV-infected STAT1-deficient mice. *Antiviral Research*, v. 146, p. 1-11, 2017.
- 65 DE SOUZA, A. A. A. *et al.* Inhibition of Brazilian ZIKV strain replication in primary human placental chorionic cells and cervical cells treated with nitazoxanide. *Brazilian Journal of Infectious Diseases*, v. 24, n. 6, p. 505-516, 2020.
- 66 CAO, R. Y. *et al.* Pediatric drug nitazoxanide: a potential choice for control of Zika. *Open Forum Infectious Diseases*. v.3, n.4,p.1, 2017.
- 67 CLAIN, E. *et al.* Extract from *Aphloia theiformis*, an edible indigenous plant from Reunion Island, impairs Zika virus attachment to the host cell surface. *Scientific Reports*, v. 8, n. 1, p. 1-12, 2018.
- 68 HADDAD, J. G. *et al.* *Doratoxylon apetalum*, an indigenous medicinal plant from Mascarene Islands, is a potent inhibitor of Zika and dengue virus infection in human cells. *International Journal of Molecular Sciences*, v. 20, n. 10, p. 2382, 2019.
- 69 CLAIN, E. *et al.* The polyphenol-rich extract from *psiloxylon mauritianum*, an endemic medicinal plant from Reunion Island, inhibits the early stages of dengue and Zika virus infection. *International Journal of Molecular Sciences*, v. 20, n. 8, p. 1860, 2019.
- 70 BATISTA, M. N. *et al.* Natural products isolated from oriental medicinal herbs inactivate Zika virus. *Viruses*, v. 11, n. 1, p. 49, 2019.
- 71 MOTTIN, M. *et al.* The A–Z of Zika drug discovery. *Drug Discovery Today*, v. 23, n. 11, p. 1833-1847, 2018.
- 72 HAN, Y.; MESPLÈDE, T. Investigational drugs for the treatment of Zika virus infection: a preclinical and clinical update. *Expert Opinion on Investigational Drugs*, v. 27, n. 12, p. 951-962, 2018.
- 73 FINK, S. L. *et al.* The antiviral drug arbidol inhibits Zika virus. *Scientific Reports*, v. 8, n. 1, p. 1-9, 2018.
- 74 GODOY, A. S. *et al.* Structural and mechanistic insight from antiviral and antiparasitic enzyme drug targets for tropical infectious diseases. *Current Opinion in Structural Biology*, v. 59, p. 65-72, 2019.

- 75 GARCÍA, L. L.; PADILLA, L.; CASTAÑO, J. C. Inhibitors compounds of the flavivirus replication process. *Virology Journal*, v. 14, n. 1, p. 1-12, 2017.
- 76 SAIZ, J. C.; MARTÍN-ACEBES, M.I A. The race to find antivirals for Zika virus. *Antimicrobial Agents and Chemotherapy*, v. 61, n. 6, 2017.
- 77 BAZ, M.; BOIVIN, G. Antiviral agents in development for Zika virus infections. *Pharmaceuticals*, v. 12, n. 3, p. 101, 2019.
- 78 YUAN, J. *et al.* Antibiotic fidaxomicin is an RdRp inhibitor as a potential new therapeutic agent against Zika virus. *BMC Medicine*, v. 18, n. 1, p. 1-16, 2020.
- 79 LIN, Y. *et al.* Identification and characterization of Zika virus NS5 RNA-dependent RNA polymerase inhibitors. *International Journal of Antimicrobial Agents*, v. 54, n. 4, p. 502-506, 2019.
- 80 CADA, D. J.; CONG, J.; BAKER, D. E. Sofosbuvir. *Hospital Pharmacy*, v. 49, n. 5, p. 466-478, 2014.
- 81 APPLEBY, T. C. *et al.* Structural basis for RNA replication by the hepatitis C virus polymerase. *Science*, v. 347, n. 6223, p. 771-775, 2015.
- 82 SACRAMENTO, C. Q. *et al.* The clinically approved antiviral drug sofosbuvir inhibits Zika virus replication. *Scientific Reports*, v. 7, n. 1, p. 1-12, 2017. DOI:10.1038/srep40920.
- 83 MALET, H. *et al.* Crystal structure of the RNA polymerase domain of the West Nile virus non-structural protein 5. *Journal of Biological Chemistry*, v. 282, n. 14, p. 10678-10689, 2007.
- 84 ASLANIDIS, C.; DE JONG, P. J. Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Research*, v. 18, n. 20, p. 6069-6074, 1990.
- 85 STUDIÉ, F. W. Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification*, v. 41, n. 1, p. 207-234, 2005.
- 86 LAEMMLI, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, v. 227, n. 5259, p. 680-685, 1970.
- 87 KABSCH, W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallographica Section D: biological crystallography*, v. 66, n. 2, p. 133-144, 2010.
- 88 WINN, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D: biological crystallography*, v. 67, n. 4, p. 235-242, 2011.
- 89 MCCOY, A. J. *et al.* Phaser crystallographic software. *Journal of Applied Crystallography*, v. 40, n. 4, p. 658-674, 2007.
- 90 TERWILLIGER, T. C. *et al.* Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallographica Section D: biological crystallography*, v. 64, n. 1, p. 61-69, 2008.
- 91 ADAMS, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: biological crystallography*, v. 66, n. 2, p. 213-221, 2010.
- 92 AFONINE, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallographica Section D: biological crystallography*, v. 68, n. 4, p. 352-367, 2012.
- 93 EMSLEY, P. *et al.* Features and development of Coot. *Acta Crystallographica Section D: biological crystallography*, v. 66, n. 4, p. 486-501, 2010.
- 94 CHEN, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D: biological crystallography*, v. 66, n. 1, p. 12-21,

2010.

95 MILES, A. J.; WALLACE, B. A. CDtoolX, a downloadable software package for processing and analyses of circular dichroism spectroscopic data. *Protein Science*, v. 27, n. 9, p. 1717-1722, 2018. DOI:10.1002/pro.3474

96 SÁEZ-ÁLVAREZ, Y. *et al.* Development of a fluorescence-based method for the rapid determination of Zika virus polymerase activity and the screening of antiviral drugs. *Scientific Reports*, v. 9, n. 1, p. 1-11, 2019.

97 TROTT, O.; OLSON, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, v. 31, n. 2, p. 455-461, 2010.

98 BERMAN, H. M. *et al.* The protein data bank. *Acta Crystallographica Section D: biological crystallography*, v. 58, n. 6, p. 899-907, 2002.

99 MORRIS, G. M. *et al.* AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *Journal of Computational Chemistry*, v. 30, n. 16, p. 2785-2791, 2009.

100 EKINS, S.; PERRYMAN, A. L.; HORTA ANDRADE, C. OpenZika: an IBM world community grid project to accelerate Zika virus drug discovery. *PLoS Neglected Tropical Diseases*, v. 10, n. 10, p. e0005023, 2016.

101 GORREC, F. The MORPHEUS protein crystallization screen. *Journal of Applied Crystallography*, v. 42, n. 6, p. 1035-1042, 2009.

102 GODOY, A. S. *et al.* Crystal structure of Zika virus NS5 RNA-dependent RNA polymerase. *Nature Communications*, v. 8, n. 1, p. 1-6, 2017.

103 SELISKO, B. *et al.* Regulation of Flavivirus RNA synthesis and replication. *Current Opinion in Virology*, v. 9, p. 74-83, 2014.

104 SELISKO, B. *et al.* Molecular basis for nucleotide conservation at the ends of the dengue virus genome. *PLoS Pathogens*, v. 8, n. 9, p. e1002912, 2012.

105 LIM, S. P. *et al.* Discovery of potent non-nucleoside inhibitors of dengue viral RNA-dependent RNA polymerase from fragment screening and structure-guided design. *Advances Experimental Medicine and Biology*, v.2018, n.1062, p. 187-198, 2018.

106 NOBLE, C. G. *et al.* A conserved pocket in the dengue virus polymerase identified through fragment-based screening. *Journal of Biological Chemistry*, v. 291, n. 16, p. 8541-8548, 2016.

107 LIM, S. P. *et al.* Potent allosteric dengue virus NS5 polymerase inhibitors: mechanism of action and resistance profiling. *PLoS Pathogens*, v. 12, n. 8, p. e1005737, 2016.

108 GHARBI-AYACHI, A. *et al.* Non-Nucleoside Inhibitors of Zika virus RNA-dependent RNA polymerase. *Journal of Virology*, v. 94, n. 21, 2020. DOI:10.1128/jvi.00794-20.

109 DUAN, W. *et al.* The crystal structure of Zika virus NS 5 reveals conserved drug targets. *The EMBO Journal*, v. 36, n. 7, p. 919-933, 2017.

110 KELLY, S. M.; JESS, T. J.; PRICE, N. C. How to study proteins by circular dichroism. *Biochimica et Biophysica Acta (BBA): proteins and proteomics*, v. 1751, n. 2, p. 119-139, 2005.

111 RANJBAR, B.; GILL, P. Circular dichroism techniques: biomolecular and nanostructural analyses-a review. *Chemical Biology & Drug Design*, v. 74, n. 2, p. 101-120, 2009. DOI:10.1111/j.1747-0285.2009.00847.x

112 KOCABAS, F.; TURAN, R. D.; ASLAN, Galip S. Fluorometric RdRp assay with self-priming

RNA. *Virus Genes*, v. 50, n. 3, p. 498-504, 2015.

113 TARANTINO, D. *et al.* Targeting flavivirus RNA dependent RNA polymerase through a pyridobenzothiazole inhibitor. *Antiviral Research*, v. 134, p. 226-235, 2016.

114 IVERSEN, P. W. *et al.* HTS assay validation. *In: MARKOSSIAN, D. et al. (eds.) Assay guidance manual*. Bethesda: Eli Lilly & Company, 2012.

115 SIMEONOV, A.; DAVIS, M. I. Interference with fluorescence and absorbance. *In: MARKOSSIAN, D. et al. (eds.) Assay guidance manual*. Bethesda: Eli Lilly & Company, 2018.

116 KOSELNY, K. *et al.* Antitumor/antifungal celecoxib derivative AR-12 is a non-nucleoside inhibitor of the ANL-family adenylyating enzyme acetyl CoA synthetase. *ACS Infectious Diseases*, v. 2, n. 4, p. 268-280, 2016.

117 MOHR, E. L. *et al.* Inhibitors of cellular kinases with broad-spectrum antiviral activity for hemorrhagic fever viruses. *Antiviral Research*, v. 120, p. 40-47, 2015.

118 RUEBSAM, F. *et al.* Discovery of tricyclic 5, 6-dihydro-1H-pyridin-2-ones as novel, potent, and orally bioavailable inhibitors of HCV NS5B polymerase. *Bioorganic & Medicinal Chemistry Letters*, v. 19, n. 22, p. 6404-6412, 2009.

119 ZHANG, S. L. *et al.* Berberine azoles as antimicrobial agents: synthesis, biological evaluation and their interactions with human serum albumin. *MedChemComm*, v. 4, n. 5, p. 839-846, 2013.

120 ZHANG, S. L. *et al.* Novel berberine triazoles: synthesis, antimicrobial evaluation and competitive interactions with metal ions to human serum albumin. *Bioorganic & Medicinal Chemistry Letters*, v. 23, n. 4, p. 1008-1012, 2013.

121 SHIRUDE, P. S. *et al.* Quinoliny pyrimidines: potent inhibitors of NDH-2 as a novel class of anti-TB agents. *ACS Medicinal Chemistry Letters*, v. 3, n. 9, p. 736-740, 2012.

122 BARRY, V. C. *et al.* A new series of phenazines (rimino-compounds) with high antituberculosis activity. *Nature*, v. 179, p. 1013-15, 1957.

123. CHOLO, M. C.; STEEL, H. C.; FOURIE, P. B.; GERMISHUIZEN, W. A.; ANDERSON, R. Clofazimine: current status and future prospects. *Journal Antimicrobial Chemotherapy*, v. 67, p. 290-298, 2012.

124 RIVA, L. *et al.* Discovery of SARS-CoV-2 antiviral drugs through large-scale compound repurposing. *Nature*, v. 586, n. 7827, p. 113-119, 2020.

125 REES, D. C. *et al.* Fragment-based lead discovery. *Nature Reviews Drug Discovery*, v. 3, n. 8, p. 660-672, 2004.

126 MURRAY, C. W.; REES, D. C. The rise of fragment-based drug discovery. *Nature Chemistry*, v. 1, n. 3, p. 187-192, 2009.

127 NIESEN, F. H.; BERGLUND, H.; VEDADI, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature Protocols*, v. 2, n. 9, p. 2212, 2007.

128 PANTOLIANO, M. W. *et al.* High-density miniaturized thermal shift assays as a general strategy for drug discovery. *Journal of Biomolecular Screening*, v. 6, n. 6, p. 429-440, 2001.

129 ASKIN, S. *et al.* Selective protein unfolding: a universal mechanism of action for the development of irreversible inhibitors. *Chemical Communications*, v. 54, n. 14, p. 1738-1741, 2018.

130 MCDONNELL, P. A. *et al.* Assessing compound binding to the Eg5 motor domain using a thermal shift assay. *Analytical Biochemistry*, v. 392, n. 1, p. 59-69, 2009.

- 131 LIPINSKI, C. A. *et al.* Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*, v. 23, n. 1-3, p. 3-25, 1997.
- 132 REES, D. C. *et al.* Fragment-based lead discovery. *Nature Reviews Drug Discovery*, v. 3, n. 8, p. 660-672, 2004.
- 133 JHOTI, H. *et al.* The 'rule of three' for fragment-based drug discovery: where are we now? *Nature Reviews Drug Discovery*, v. 12, n. 8, p. 644-644, 2013.
- 134 DAINA, A.; MICHIELIN, O.; ZOETE, V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific Reports*, v. 7, n. 1, p. 1-13, 2017.
- 135 KIRSCH, P. *et al.* Concepts and core principles of fragment-based drug design. *Molecules*, v. 24, n. 23, p. 4309, 2019.
- 136 FINNEGAN, R. A. *et al.* Occurrence of mangiferin in *Hiptage madablota* geartn. *Journal of Pharmaceutical Sciences*, v. 57, n. 6, p. 1039-1040, 1968.
- 137 PARDO-ANDREU, G. L. *et al.* Vimang (*Mangifera indica* L. extract) induces permeability transition in isolated mitochondria, closely reproducing the effect of mangiferin, Vimang's main component. *Chemico-Biological Interactions*, v. 159, n. 2, p. 141-148, 2006.
- 138 RIVERA, D. G. *et al.* Anti-allergic properties of *Mangifera indica* L. extract (Vimang) and contribution of its glucosylxanthone mangiferin. *Journal of Pharmacy and Pharmacology*, v. 58, n. 3, p. 385-392, 2006.
- 139 GARCÍA-RIVERA, D. *et al.* Gallic acid indanone and mangiferin xanthone are strong determinants of immunosuppressive anti-tumour effects of *Mangifera indica* L. bark in MDA-MB231 breast cancer cells. *Cancer Letters*, v. 305, n. 1, p. 21-31, 2011.
- 140 DU, S. *et al.* Mangiferin: an effective therapeutic agent against several disorders. *Molecular Medicine Reports*, v. 18, n. 6, p. 4775-4786, 2018.
- 141 LI, M. *et al.* Mangiferin inhibition of proliferation and induction of apoptosis in human prostate cancer cells is correlated with downregulation of B-cell lymphoma-2 and upregulation of microRNA-182. *Oncology Letters*, v. 11, n. 1, p. 817-822, 2016.
- 142 MIRZA, R. H.; CHI, N.; CHI, Y. Therapeutic potential of the natural product mangiferin in metabolic syndrome. *Journal of Nutritional Therapeutics*, v. 2, n. 2, p. 74-79, 2013. DOI:10.6000/1929-5634.2013.02.02.2.
- 143 WALIA, V.; CHAUDHARY, S. K.; SETHIYA, N. K. Therapeutic potential of mangiferin in the treatment of various neuropsychiatric and neurodegenerative disorders. *Neurochemistry International*, v.143, p. 104939, 2020.
- 144 ROSHCHIN, Y. V. Trifolin from *Euphorbia condylocarpa*. *Chemistry of Natural Compounds*. v.13, n.4, p.481-482, 1977.
- 145 WANG, X. *et al.* The mechanism of anticancer action and potential clinical use of kaempferol in the treatment of breast cancer. *Biomedicine & Pharmacotherapy*, v. 117, p. 109086, 2019.
- 146 LI, S. *et al.* Antifungal activity of camptothecin, trifolin, and hyperoside isolated from *Camptotheca acuminata*. *Journal of Agricultural and Food Chemistry*, v. 53, n. 1, p. 32-37, 2005.
- 147 TORRES, F. *et al.* Trifolin acetate-induced cell death in human leukemia cells is dependent on caspase-6 and activates the MAPK pathway. *Apoptosis*, v. 13, n. 5, p. 716-728, 2008.
- 148 OLIVEIRA, A. F. C. S. *et al.* Potential antivirals: natural products targeting replication enzymes of

dengue and chikungunya viruses. *Molecules*, v. 22, n. 3, p. 505, 2017.

149 MARTINEZ, J. P. *et al.* Antiviral drug discovery: broad-spectrum drugs from nature. *Natural Product Reports*, v. 32, n. 1, p. 29-48, 2015.

150 HAWKINS, P. C. *et al.* Conformer generation with OMEGA: algorithm and validation using high quality structures from the Protein Databank and Cambridge Structural Database. *Journal of Chemical Information and Modeling*, v. 50, n. 4, p. 572-584, 2010.

151 BAELELL, J. B.; HOLLOWAY, G. A. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *Journal of Medicinal Chemistry*, v. 53, n. 7, p. 2719-2740, 2010.

152 CALDERÓN-OLIVER, M.; PONCE-ALQUICIRA, E. Fruits: a source of polyphenols and health benefits. *In: CALDERÓN-OLIVER, M.; PONCE-ALQUICIRA, E. Natural and artificial flavoring agents and food dyes.* New York:Academic Press,2018 p.189-228.DOI:10.1016/b978-0-12-811518-3.00007-7.

153 DAVID, A. V. A.; ARULMOLI, R.; PARASURAMAN, S. Overviews of biological importance of quercetin: a bioactive flavonoid. *Pharmacognosy Reviews*, v. 10, n. 20, p. 84, 2016.

154 RAMOS, S. Effects of dietary flavonoids on apoptotic pathways related to cancer chemoprevention. *The Journal of Nutritional Biochemistry*, v. 18, n. 7, p. 427-442, 2007.

155 FORMICA, J. V.; REGELSON, W. Review of the biology of quercetin and related bioflavonoids. *Food and Chemical Toxicology*, v. 33, n. 12, p. 1061-1080, 1995.