

UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS

Mateus Sá Magalhães Serafim

**TRIAGEM VIRTUAL E BIOLÓGICA DE SUBSTÂNCIAS COM POTENCIAL
ATIVIDADE ANTIVIRAL CONTRA TRÊS FLAVIVÍRUS: dengue virus, yellow fever
virus e Zika virus**

Belo Horizonte

2020

Mateus Sá Magalhães Serafim

**TRIAGEM VIRTUAL E BIOLÓGICA DE SUBSTÂNCIAS COM POTENCIAL
ATIVIDADE ANTIVIRAL CONTRA TRÊS FLAVIVÍRUS: dengue virus, yellow fever
virus e Zika virus**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Microbiologia da Universidade Federal de Minas Gerais, como requisito parcial para obtenção do título de Mestre em Microbiologia.

Orientadora: Profa. Dra. Erna Geessien Kroon (ICB/UFMG)

Coorientadores: Prof. Dr. Bruno Eduardo Fernandes Mota (FAFAR/UFMG)

Prof. Dr. Vinícius Gonçalves Maltarollo (FAFAR/UFMG)

Colaborador: Dr. Thales Kronenberger (University Hospital of Tübingen/Alemanha)

Belo Horizonte

2020

043 Serafim, Mateus Sá Magalhães.
Triagem virtual e biológica de substâncias com potencial atividade antiviral
contra três flavivírus: dengue virus, yellow fever virus e Zika virus [manuscrito]
/ Mateus Sá Magalhães Serafim. – 2020.
236 f. : il. ; 29,5 cm.

Orientadora: Profa. Dra. Erna Geessien Kroon. Coorientadores: Prof. Dr.
Bruno Eduardo Fernandes Mota, Prof. Dr. Vinícius Gonçalves Maltarollo.
Colaborador: Dr. Thales Kronenberger.

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

1. Microbiologia. 2. Biologia molecular. 3. Antivirais. 4. Flavivirus. I. Kroon,
Erna Geessien. II. Mota, Bruno Eduardo Fernandes. III. Maltarollo, Vinícius
Gonçalves. IV. Kronenberger, Thales. V. Universidade Federal de Minas Gerais.
Instituto de Ciências Biológicas. VI. Título.

CDU: 579



UNIVERSIDADE FEDERAL DE MINAS GERAIS

INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM MICROBIOLOGIA

ATA DA DEFESA DE DISSERTAÇÃO DE Mateus Sá Magalhães Serafim
Nº REGISTRO: 2018685060

Às 13h30 do dia 18 de fevereiro de 2020, reuniu-se, no Instituto de Ciências Biológicas da UFMG, a Comissão Examinadora composta pelas Prof(a). Rafaela Salgado Ferreira, Jaqueline Maria Siqueira Ferreira e a Prof(a). Erna Geessien Kroon, Orientadora, além dos Prof. Bruno Eduardo Fernandes Mota e Vinícius Gonçalves Maltarollo, coorientadores, para julgar o trabalho final "TRIAGEM VIRTUAL E BIOLÓGICA DE SUBSTÂNCIAS COM POTENCIAL ATIVIDADE ANTIVIRAL CONTRA TRÊS FLAVIVÍRUS: dengue virus, yellow fever virus e Zika virus", do aluno Mateus Sá Magalhães Serafim, requisito final para a obtenção do Grau de **MESTRE EM CIÊNCIAS BIOLÓGICAS: MICROBIOLOGIA**. Abrindo a sessão, o Presidente da Comissão, Prof. Flávio Guimarães da Fonseca - Coordenador do Programa, após dar a conhecer aos presentes o teor das Normas Regulamentares do Trabalho Final, passou a palavra ao candidato, para a apresentação de seu trabalho. Seguiu-se a arguição pelos Examinadores, com a respectiva defesa do candidato. Logo após, a Comissão se reuniu, sem a presença do candidato e do público, para julgamento e expedição de resultado final. O candidato foi considerado APROVADO. O resultado final foi comunicado publicamente ao candidato pelo Presidente da Comissão. Nada mais havendo a tratar, o Presidente encerrou a reunião e lavrou a presente ata, que será assinada por todos os membros participantes da Comissão Examinadora. O candidato tem 60 (sessenta) dias, a partir desta data, para entregar a versão final da dissertação ao Programa de Pós-Graduação em Microbiologia da UFMG e requerer seu diploma.

Prof(a). Rafaela Salgado Ferreira Rafaela S. Ferreira

Prof(a). Jaqueline Maria Siqueira Ferreira J. Siqueira

Prof(a). Erna Geessien Kroon [Orientadora] E. Kroon

Prof(a). Bruno Eduardo Fernandes Mota [Coorientador] Bruno E. F. Mota

Prof(a). Vinícius Gonçalves Maltarollo [Coorientador] V. Maltarollo

F. Bitânia Paiva Durand

Prof. Flávio Guimarães da Fonseca
Coordenador do PPGM/UFMG

Agradecimentos

Reservei esse espaço para agradecimentos pessoais ao final da defesa.

RESUMO

Aproximadamente três bilhões de pessoas vivem em regiões com risco de infecção por *Dengue virus* (DENV), *Zika virus* (ZIKV) e *Yellow fever virus* (YFV). Além dos casos graves de dengue, Zika exibe complicações como síndromes congênitas em fetos. Ademais, surtos de febre amarela ocorreram no Brasil em 2017, mesmo havendo vacina contra este vírus. Até o momento não há fármacos antivirais aprovados para o tratamento dessas doenças, ressaltando a necessidade de estratégias para se obter novas moléculas com atividade antiviral. Portanto, o objetivo deste trabalho foi avaliar, empregando estratégias *in silico* e *in vitro*, a atividade antiviral de substâncias em DENV, ZIKV e YFV, tendo como alvo a protease viral (NS3_{PRO}). Estruturas tridimensionais de compostos peptidomiméticos inibidores de protease de DENV-3 foram empregadas para a geração de modelos da relação quantitativa entre a estrutura química e a atividade biológica por holograma (HQSAR) e de acoplamento molecular (*docking*). Paralelamente, foram preditos sítios de ligação e *hotspots* de ligação das proteases alvo (PDB: 3U1I e 5YOF) utilizando programas FTSite, FTMap e AutoGrid. Esses resultados foram utilizados na construção de um modelo farmacofórico para triagem virtual. Propriedades físico-químicas, presença de fragmentos estruturais e características de interações com o alvo molecular (farmacóforo) foram delimitadas na triagem. Foi realizada dinâmica molecular dos *hits* virtuais, bem como predições de toxicidade e atividade biológica, selecionando oito substâncias para aquisição. A validação da atividade biológica foi corroborada empregando o teste colorimétrico de brometo de 3-(4,5-dimetiltiazol-2-il)-2,5-difeniltetrazólio (MTT), determinando as concentrações citotóxicas (CC₅₀), concentrações efetivas (CE₅₀) e índices de seletividade (IS). A atividade antiviral também foi avaliada por ensaios de placas de lise, efeito de pré-tratamento e de atividade virucida das substâncias. Inicialmente, modelos de HQSAR foram construídos e validados, sendo um modelo ótimo ($q^2 = 0.675$ e $r^2 = 0.873$) selecionado utilizando mapas de contribuição dos fragmentos moleculares. Na estrutura da protease foi delimitada a tríade catalítica, ligantes, sítios e subsítios, sendo utilizados na construção de um modelo farmacofórico baseado no mapeamento, interpretação do modelo de HQSAR e interações observadas em estudos prévios de *docking*. Uma triagem virtual de aproximadamente 7.600.000 compostos foi conduzida em diversas etapas (farmacóforo, *docking* e dinâmica molecular), identificando oito potenciais inibidores da protease de DENV-3 (dois fármacos e seis moléculas sintéticas sob investigação). Na dinâmica molecular foram observadas interações estáveis no sítio alvo, complementadas por predições de atividade (5/8) e toxicidade favoráveis (8/8). Dentre as substâncias, uma foi ativa em ZIKV (CE₅₀ = 11,86 ± 0,43 µM; IS = 1,42), quatro em YFV (CE₅₀ = 4,79 ± 0,17 a 24,15 ± 1,15 µM; IS = 1,97 a 3,74), duas em DENV-2 (CE₅₀ = 4,21 ± 0,14 e 9,51 ± 0,12 µM; IS = 2,26 e 3,16) e uma em DENV-3 (CE₅₀ = 37,51 ± 0,8 µM; IS = 2,85). Nos ensaios de placa, as substâncias 136 e 128 reduziram cerca de 1 a 1,5 log₁₀ do título viral (UFP/mL) de ZIKV, YFV e DENV-2. As substâncias 136 e 140 também reduziram cerca de 1 log₁₀ do título viral de YFV nos ensaios de pré-tratamento. Posteriormente, ensaios de inibição enzimática serão realizados para confirmação do alvo planejado neste trabalho.

Palavras-chave: Atividade antiviral. Dinâmica molecular. Docking. *Flavivirus*. HQSAR. NS3_{PRO}.

ABSTRACT

Approximately three billion people live in regions at risk of infection by the dengue virus (DENV), Zika virus (ZIKV) and yellow fever virus (YFV). In addition to severe cases of dengue, Zika presents complications such as congenital syndromes in fetuses. Yellow fever outbreaks also happened in Brazil in 2017, even with a vaccine available against it. Currently, there are no licensed antivirals for the treatment of these diseases, emphasizing the need for strategies to obtain new bioactive molecules with antiviral activity. The objective of this work was to evaluate, through *in vitro* and *in silico* strategies, the antiviral activity of substances against DENV, ZIKV and YFV, targeting the viral protease (NS3_{PRO}). Three-dimensional structures of peptidomimetic compounds inhibiting DENV-3 protease were used to generate models of the quantitative relationship between chemical structure and biological activity by holograms (HQSAR) and molecular docking. Binding sites and hotspots of protein structures (PDB: 3U1I and 5YOF) were detected using the FTSite, FTMap and AutoGrid programs. These results were used to the construction of a pharmacophoric model for virtual screening. Physical and chemical properties, presence molecular fragments and interactions with the molecular target (pharmacophore) were defined in the screening. Eight compounds were selected after molecular dynamics, as well as predictions of toxicity and biological activity. Biological activity was confirmed by the colorimetric assay of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT), in order to obtain the cytotoxic (CC₅₀) and effective (EC₅₀) concentrations, as well as selectivity indexes (SI). The antiviral activity was also evaluated by plaque reduction, pre-treatment and virucide activity assays. Initially, HQSAR models were built and validated, with an optimal model ($q^2 = 0.675$ and $r^2 = 0.873$) selected using chemical contribution maps of molecular fragments. The catalytic triad, ligands and binding sites pointed at the structures were used to build a pharmacophoric model based on mapping, interpretation of the HQSAR model and interactions observed in previous fitting studies. A virtual screening of approximately 7,600,000 compounds was conducted in several stages (pharmacophore, docking and molecular dynamics), identifying eight potential inhibitors to the DENV-3 protease (two drugs and six synthetic molecules under investigation). Molecular dynamics simulations showed stable interactions at the target sites, and were complemented by favorable biological activity (5/8) and toxicity (8/8) predictions. Among the substances, one was active against ZIKV (EC₅₀ = 11.86 ± 0.43 μM; SI = 1.42), four against YFV (EC₅₀ = 4.79 ± 0.17 to 24.15 ± 1.15 μM; SI = 1.97 to 3.74), two against DENV-2 (EC₅₀ = 4.21 ± 0.14 and 9.51 ± 0.12 μM; 2.26 and 3.16) and one against DENV-3 (EC₅₀ = 37.51 ± 0.8 μM; SI = 2.85). In plaque reduction assays, substances 136 and 128 reduced about 1.0 to 1.5 log₁₀ of the viral titer (PFU/mL) of ZIKV, YFV and DENV-2. Substances 136 and 140 also reduced about 1.0 log₁₀ of YFV titer in pre-treatment assays. As a perspective, enzyme inhibition assays will be carried out in order to confirm the molecular target predicted in this work.

Keywords: Antiviral activity. Docking. *Flavivirus*. HQSAR. Molecular dynamics. NS3.

LISTA DE FIGURAS

Figura 1. Ciclo de multiplicação dos vírus do gênero <i>Flavivirus</i>.	20
Figura 2. Representação de partícula viral dos vírus do gênero <i>Flavivirus</i>.	21
Figura 3. Organização do genoma viral e proteínas traduzidas conforme modelo em vírus do gênero <i>Flavivirus</i>.	22
Figura 4. Estrutura tridimensional da protease (NS2B-NS3) de dengue virus.	26
Figura 5. Sequência em etapas do mecanismo de ação da protease NS2B-NS3 de dengue virus.	27
Figura 6. Ciclo de transmissão de dengue em humanos e mosquitos	29
Figura 7. Distribuição global de sequências de DENV até 2016.	30
Figura 8. Critérios para classificação da gravidade de dengue.	31
Figura 9. Surtos de Zika ao longo da última década (2007-2017).	33
Figura 10. Manifestações clínicas e consequências da infecção por ZIKV.	36
Figura 11. Marcadores moleculares e sorológicos na infecção por DENV: NS1, IgM, IgG, ELISA e RT-PCR.	40
Figura 12. Dengvaxia® (CYD-TDV) da <i>Sanofi Pasteur</i>: vacina recombinante tetravalente de DENV.	42
Figura 13. Representações estruturais da aprotinina.	46
Figura 14. Molécula 104 (peptideomimético) com atividade anti-DENV.	47
Figura 15. Exemplos de inibidores da protease de DENV descobertos.	47
Figura 16. Esquema de etapas da triagem virtual.	66
Figura 17. Estrutura geral das moléculas utilizadas na construção do conjunto de dados.	77
Figura 18. (A) Coeficiente de validação cruzada de Y.	85
Figura 18. (B) Aleatorização de Y (Y scramble).	85
Figura 19. Leave many out (LMO).	86
Figura 20. Gráfico de frequência de moléculas por bloco de distribuição do modelo selecionado de HQSAR.	87
Figura 21. Mapas de contribuição química das cinco moléculas mais ativas (menores valores de IC50) experimentalmente na protease DENV-3.	88
Figura 22. Mapas de contribuição química das cinco moléculas menos ativas (maiores valores de IC50) experimentalmente na protease DENV-3.	89
Figura 23. Mapas de contribuição química das moléculas mais ativas (62 e 65) e respectivos fragmentos de maior importância.	90
Figura 24. Visualização em modelo de cartoon da sobreposição das 15 estruturas da NS3 de DENV importadas do PDB.	91
Figura 25. Estrutura tridimensional do complexo da protease NS2B-NS3 do dengue virus depositada no PDB sob o código 3U1I.	92

Figura 26. (A) Protease (NS2B-NS3) de DENV-3.	93
Figura 26. (B) Protease (NS2B-NS3) de DENV-3 e grupamentos de moléculas-sonda.	93
Figura 27. Peptídeo inibidor NDL (ligante de origem) da protease NS2B-NS3 de DENV-3.	94
Figura 28. (A) Percentual de interações de hidrogênio do tipo acceptor (HBA) por aminoácidos (aa) no sítio de ligação da protease de DENV-3 (3U1I).	95
Figura 28. (B) Campo de interação molecular de hidrogênio do tipo acceptor (vermelho) próximo ao aminoácido (aa) Ser135.	95
Figura 29. (A) Percentual de interações de hidrogênio do tipo doador (HBD) por aa no sítio de ligação da protease de DENV-3 (3U1I).	96
Figura 29. (B) Campo de interação molecular de hidrogênio do tipo doador (azul) próximo aos aa Asp75 e Asp81.	96
Figura 30. (A) Percentual de interações eletrostáticas (iônicas) por aa no sítio de ligação da protease de DENV-3 (3U1I).	97
Figura 30. (B) Campo de interação molecular do tipo eletrostática (amarelo) próximo a Asp75.	97
Figura 31. (A) Percentual de interações de van der Waals (hidrofóbicas) no sítio de ligação da protease de DENV-3 (3U1I).	98
Figura 31. (B) Campo de interação molecular do tipo hidrofóbica (laranja) próximo a Val52.	98
Figura 32. Estrutura tridimensional do complexo da protease NS2B-NS3 de ZIKV depositada no PDB sob o código 5YOF.	100
Figura 33. (A) Protease NS2B-NS3 de ZIKV e suas cavidades representando sítios de ligação	101
Figura 33. (B) Protease NS2B-NS3 de ZIKV e grupamentos de moléculas-sonda.	101
Figura 34. Dipeptídeo inibidor da protease (5YOF) de ZIKV.	102
Figura 35. Percentual de interações de hidrogênio do tipo acceptor (HBA) por aa no sítio de ligação da protease de ZIKV (5YOF).	102
Figura 36. Percentual de interações de hidrogênio do tipo doador (HBD) por aa no sítio de ligação da protease de ZIKV (5YOF).	103
Figura 37. Percentual de interações eletrostáticas (iônicas) por aa no sítio de ligação da protease de ZIKV (5YOF).	103
Figura 38. Percentual de interações de van der Waals (hidrofóbicas) por aa no sítio de ligação da protease de ZIKV (5YOF).	104
Figura 39. Campos de interação de ligação de hidrogênio do tipo acceptor nas proteases de DENV-3 (3U1I) e ZIKV (5YOF).	105
Figura 40. Campos de interação do tipo eletrostático nas proteases de DENV-3 (3U1I) e ZIKV (5YOF).	105
Figura 41. Sobreposição das caixas tridimensionais utilizadas nos cálculos de campo de interação molecular (MIF) das proteases de DENV (3U1I) e ZIKV (5YOF).	106

Figura 42. Análise comparativa do docking da molécula 62 na protease de DENV-3 (3U1I). (A) Reprodução do acoplamento molecular utilizando programa Surflex Docking.	108
Figura 42. (B) Docking do composto 62 na protease (3U1I) de DENV-3.	108
Figura 43. Sequências de NS3 de diferente flavivírus: DENV-1 a 4, JEV, WNV e YFV.	110
Figura 44. (A) Esquema de contatos de átomos do peptídeo inibidor (ligante de origem) nos aa da protease de DENV-3 (3U1I).	112
Figura 44. (B) Esquema de contatos de átomos do peptídeo inibidor (ligante de origem) nos aa da protease DENV-3 (3U1I) (duplicata).	112
Figura 45. Frequência de interações do ligante (peptídeo) com aa da protease de DENV-3 (duplicatas A e B).	113
Figura 46. (A) Esquema de contatos de átomos do inibidor dipeptídico (ligante de origem) nos aa da protease de ZIKV (5YOF).	115
Figura 46. (B) Frequência de interações do ligante (dipeptídeo) com aa da protease de ZIKV.	115
Figura 47. (A) Esquema de contatos de átomos de 62 com aa da protease de DENV-3 (3U1I).	117
Figura 47. (B) Frequência de interações do ligante 62 com aa da protease de DENV-3.	117
Figura 48. (A) Esquema de contatos de átomos de 65 com aa da protease de DENV-3 (3U1I).	119
Figura 48. (B) Frequência de interações do ligante 65 com aa da protease de DENV-3.	119
Figura 49. (A) Esquema de contatos de átomos de 62 com aa da protease de ZIKV (5YOF).	120
Figura 49. (B) Frequência de interações do ligante 62 com aa da protease de ZIKV.	120
Figura 50. Modelo farmacofórico construído utilizando a estrutura do complexo entre a protease de DENV-3 e o ligante de origem.	124
Figura 51. Pose de docking de um potencial inibidor triado pelo farmacóforo.	126
Figura 52. Modelo do protomol ou envelope de interações.	127
Figura 53. (A) Exemplo de análise de interações do docking no sítio do farmacóforo	128
Figura 53. (B) Exemplo aumentado do modelo de pontuação das interações.	128
Figura 54. Comparação entre os resultados dos protocolos de docking nos programas Glide (P2) e Surflex (P1).	132
Figura 55. Avaliação da atividade antiviral das cinco substâncias ativas por ensaios de redução de placas de lise.	154
Figura 56. Avaliação do efeito de pré-tratamento das cinco substâncias ativas.	158
Figura 57. Avaliação da atividade virucida das cinco substâncias ativas.	160

LISTA DE TABELAS

Tabela 1. Espécies de arbovírus de maior importância clínica da família <i>Flaviviridae</i> e do gênero <i>Flavivirus</i>: manifestações clínicas comuns, hospedeiros e vetores.	17
Tabela 2. Características de vírus da família <i>Flaviviridae</i>.	18
Tabela 3. Proteínas não estruturais dos flavivírus.	25
Tabela 4. Sorotipos e respectivos genótipos de DENV.	30
Tabela 5. Dengue: fases, duração, sinais clínicos e testes laboratoriais.	32
Tabela 6. Sinais e sintomas clínicos de dengue e Zika.	35
Tabela 7. Descrição dos diferentes tipos de fragmentos utilizados no método de relação quantitativa entre a estrutura química e atividade biológica por holograma (HQSAR).	50
Tabela 8. Moléculas-sonda utilizadas pelos servidores FTMap e FTSite e suas respectivas características físico-químicas.	53
Tabela 9. Parâmetros definidos para o modelo de triagem virtual.	64
Tabela 10. Relação quantitativa entre a estrutura química e atividade biológica por holograma (HQSAR) do conjunto de dados.	79
Tabela 11. Variação do tamanho dos fragmentos dos quatro modelos de HQSAR. (5 ou 6) do tamanho dos fragmentos de cada.	81
Tabela 12. Validação externa dos modelos de HQSAR selecionados por q^2.	83
Tabela 13. Comparação dentre três modelos ótimos de HQSAR.	84
Tabela 14. Coordenadas dos centróides dos <i>grids</i> utilizados nos cálculos de campo de interação molecular (MIF) de DENV e ZIKV.	107
Tabela 15. Pontuações dos 33 <i>hits</i> (protocolo 1) no farmacóforo.	129
Tabela 16. Pontuação dos 18 <i>hits</i> (protocolo 2) no farmacóforo.	131
Tabela 17. Estrutura das moléculas selecionadas na análise de duplo filtro dos protocolos 1 e 2 (P1'' e P2'').	133
Tabela 18. Simulações das substâncias do banco MolPort selecionadas nas DM.	135
Tabela 19. Simulações das substâncias do banco DrugBank selecionadas nas DM.	137
Tabela 20. Percentual predito das substâncias triadas e selecionadas nos parâmetros do servidor Pred-hERG.	140
Tabela 21. Parâmetros farmacocinéticos preditos pelo servidor admetSAR para as oito substâncias selecionadas.	143
Tabela 22. Predição de atividade biológica das oito substâncias: kNNw, SVM polinomial e floresta aleatória (RF).	145
Tabela 23. Estrutura química das oito substâncias selecionadas na triagem para avaliação da atividade biológica <i>in vitro</i>: 134, 151, 154, 156, 136, 128, 140 e 158.	147

Tabela 24. Avaliação da atividade antiviral das oito substâncias selecionadas na triagem em três *Flavivirus*: *Zika virus* (ZIKV), *Yellow fever virus* (YFV) e *Dengue virus* (DENV).

149

LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS

Å	Angstrom
C	Proteína do capsídeo
°C	Grau <i>Celsius</i>
CC ₅₀	Concentração citotóxica 50%
CDC	<i>Centers for Disease Control and Prevention</i> (Centro de Controle e Prevenção de Doenças)
cDNA	DNA complementar
CE ₅₀	Concentração efetiva 50%
CMC	Carboximetilcelulose
CO ₂	Dióxido de carbono
DENV	Dengue virus
DMSO	Dimetilsulfóxido
DNA	<i>Deoxyribonucleic acid</i> (Ácido desoxirribonucleico)
E	Proteína E
EDTA	Ácido etilenodiaminotetracético
h	Hora(s)
IS	Índice de seletividade
JEV	Japanese encephalitis virus
MEM	<i>Minimum Essential Medium</i> (Meio mínimo essencial de <i>Eagle</i>)
mg	Miligrama(s)
mL	Mililitro(s)
MIF	<i>Molecular Interaction Field</i> (Campo de interação molecular)
MTT	Brometo de 3-(4,5-dimetiltiazol-2-il)-2,5-difeniltetrazólio
µL	Microlitro(s)
µg/mL	Micrograma(s) por mililitro
nm	Nanômetro
OMS	Organização Mundial de Saúde
OPAS	Organização Pan-Americana de Saúde
prM	Proteína M
RNA	<i>Ribonucleic acid</i> (Ácido ribonucleico)
rRNA	<i>Ribosomal RNA</i> (RNA ribossômico)

RT	<i>Reverse transcriptase</i> (Transcriptase reversa)
SLEV	Saint Louis encephalitis virus
TABV	Tamana bat virus
TBEV	Tick-borne encephalitis virus
tRNA	Transfer RNA (RNA transportador)
U	Unidade Internacional
UFP	Unidade formadora de placa
UFMG	Universidade Federal de Minas Gerais
v/v	Volume por volume
WHO	<i>World Health Organization</i> (Organização Mundial de Saúde)
WNV	West Nile virus
YFV	Yellow fever virus
ZIKV	Zika virus
%	Porcentagem
®	Marca registrada

SUMÁRIO

1 INTRODUÇÃO	17
1.1 Arboviroses.....	17
1.1.1 Família <i>Flaviviridae</i>	18
1.1.1.1 Taxonomia	18
1.1.2 O gênero <i>Flavivirus</i>	19
1.1.2.1 Ciclo de multiplicação	19
1.1.2.2 Estrutura da partícula e genoma dos flavivírus	21
1.1.2.2.1 Proteínas estruturais	23
1.1.2.2.2 Proteínas não estruturais	23
1.1.2.2.3 O complexo da protease NS2B-NS3.....	25
1.1.2.3 Dengue virus (DENV).....	27
1.1.2.4 Zika virus (ZIKV)	33
1.1.2.5 Yellow fever virus (YFV).....	36
1.1.3.1 Epidemiologia de febre amarela, dengue e Zika	38
1.1.3.2 Diagnóstico de febre amarela, dengue e Zika	39
1.1.3.3 Prevenção e tratamento de dengue, Zika e febre amarela.....	42
1.2 Estratégias no planejamento, triagem e desenvolvimento de fármacos antivirais	44
1.2.1 Inibidores da protease NS2B-NS3.....	45
1.2.2 Triagem virtual de moléculas bioativas	48
1.2.3 Métodos de aprendizado de máquina (MLT): relação quantitativa entre a estrutura química e atividade biológica por holograma (HQSAR)	49
1.2.4 Acoplamento molecular (<i>molecular docking</i>)	51
1.2.5 Dinâmica molecular	52
1.2.6 Predição e mapeamento de sítios de ligação em proteínas alvo	53
1.2.7 Predições de biodisponibilidade (farmacocinética) e toxicidade.....	55

2 JUSTIFICATIVA	57
3 OBJETIVOS	58
3.1 Objetivo Geral	58
3.2 Objetivos Específicos	58
4 MATERIAL E MÉTODOS	59
4.1 Análises <i>in silico</i>	59
4.1.1 Conjunto de dados para análise: seleção de moléculas bioativas	59
4.1.2 Relação quantitativa entre a estrutura química e atividade biológica por holograma (HQSAR) de compostos peptidomiméticos.....	59
4.1.3 Seleção e preparo das estruturas da NS2B-NS3	60
4.1.4 Predição dos sítios de ligação da protease e campos de interação molecular (MIF)	61
4.1.5 Acoplamento molecular (molecular docking): ligantes e peptidomiméticos....	62
4.1.6 Construção do modelo farmacofórico e triagem virtual (VS)	63
4.1.7 Dinâmica molecular	67
4.1.8 Predições de toxicidade	69
4.1.9 Modelos de aprendizado de máquina (MLT) para predição da atividade inibitória da protease de DENV	69
4.1.10 Validações e seleção dos modelos de aprendizado de máquina	70
4.2 Linhagens celulares	71
4.3 Vírus	72
4.3.1 Estoques.....	72
4.3.2 Multiplicação	72
4.3.3 Titulação	72
4.4 Ensaios biológicos	73
4.4.1 Avaliação da citotoxicidade: concentração citotóxica 50% (CC50)	73
4.4.2 Avaliação da atividade antiviral: concentração efetiva a 50% (CE50)	74
4.4.3 Avaliação da atividade virucida	75
4.4.4 Ensaio de redução do número de placas de lise.....	75

4.4.5 Ensaio de pré-tratamento.....	76
5 RESULTADOS E DISCUSSÃO	77
5.1 Análises <i>in silico</i>	77
5.1.1 Construção e preparação do conjunto de dados	77
5.1.2 Seleção do modelo de relação quantitativa entre a estrutura química e atividade biológica por holograma (HQSAR).....	78
5.1.3 Validações do modelo ótimo de HQSAR selecionado	84
5.1.4 Interpretação físico-química do modelo de HQSAR selecionado	87
5.1.5 Seleção da estrutura da protease (NS3pro) de DENV	90
5.1.6 Mapeamento da estrutura de DENV selecionada	92
5.1.7 Seleção e mapeamento da estrutura da protease de ZIKV	99
5.1.8 Comparação entre as estruturas de DENV e ZIKV	104
5.1.9 Análises de acoplamento molecular	107
5.1.10 Dinâmicas moleculares	110
5.1.11 Construção do modelo farmacofórico e triagem virtual	123
5.1.12 Triagem virtual: filtros e crossdocking.....	125
5.1.13 Análise e classificação das moléculas obtidas na triagem.....	127
5.1.14 Filtrando os protocolos de triagem	131
5.1.15 Simulações de DM dos <i>hits</i> e seleção dos candidatos	135
5.2 Predições das atividades biológicas.....	138
5.2.1 Cardiotoxicidade.....	138
5.2.2 Toxicidade oral	141
5.2.3 Características farmacocinéticas (biodisponibilidade)	142
5.2.4 Predição da atividade biológica	144
5.3 Ensaio biológicos	147
5.3.1 Avaliação da citotoxicidade das substâncias).....	147
5.3.2 Avaliação da concentração efetiva a 50% (CE ₅₀).....	148
5.3.3 Ensaio de redução do número de placas de lise	152

5.3.4 Ensaio de pré-tratamento.....	157
5.3.5 Avaliação da atividade virucida.....	159
6 CONCLUSÕES.....	162
7 PERSPECTIVAS.....	163
8 REFERÊNCIAS BIBLIOGRÁFICAS	164
9 APÊNDICES	181
10 ANEXOS	186

1. INTRODUÇÃO

1.1 Arboviroses

Diversas espécies de vírus apresentam importância ambiental, veterinária e humana de saúde pública no cenário mundial, em particular aqueles classificados como arbovírus (do inglês, *arthropod-borne viruses*), transmitidos por vetores artrópodes. A circulação destes vírus na natureza ocorre pela transmissão de artrópodes hematófagos, como carrapatos e principalmente mosquitos, a exemplo dos gêneros *Culex* e *Aedes*, com destaque para duas espécies, *A. aegypti* e *A. albopictus* (WEAVER; BARRETT, 2004; PATTERSON; SAMMON; GARG, 2016a).

O termo arboviroses é abrangente, compreendendo e classificando vírus de diferentes famílias, como *Flaviviridae*, *Togaviridae*, *Bunyaviridae*, *Reoviridae* e *Rhabdoviridae* (KUNO; CHANG, 2005). Dentre estas, na família *Flaviviridae* se destacam diversos gêneros de importância veterinária e humana, a exemplo do gênero *Flavivirus* (**Tabela 1**). Embora a maioria das infecções tenha caráter assintomático ou sintomatologia inespecífica (semelhante a uma gripe ou resfriado), muitas dessas espécies são patógenos que denotam relevância clínica, especialmente em humanos, podendo causar diversas doenças graves. Neste âmbito, a dengue, por exemplo, pode desenvolver casos como febres hemorrágicas e evoluir ao choque sistêmico, além de acometimentos neurológicos como as encefalites, podendo ocasionar a morte dos indivíduos (WEAVER; BARRETT, 2004; GROSSI-SOYSTER et al., 2017).

Tabela 1. Espécies de arbovírus de maior importância clínica da família *Flaviviridae* e do gênero *Flavivirus*: manifestações clínicas comuns, hospedeiros e vetores.

Vírus	Manifestações clínicas	Hospedeiros	Vetores
<i>Dengue virus</i>	Febre, síndrome hemorrágica	Primatas não humanos (PNH) e humanos	<i>Aedes aegypti</i> <i>Aedes albopictus</i>
<i>Japanese encephalitis virus</i>	Febre, encefalite	Aves, suínos e humanos	<i>Culex</i> spp.
<i>West Nile virus</i>	Febre, encefalite	Aves e humanos	<i>Culex quinquefasciatus</i>
<i>Saint Louis encephalitis virus</i>	Encefalite	Aves e humanos	<i>Culex</i> spp.
<i>Yellow Fever virus</i>	Febre, encefalite	PNH e humanos	<i>Aedes</i> , <i>Sabethes</i> e <i>Haemagogus</i> spp.
<i>Zika virus</i>	Febre, síndromes congênitas e Guillain-Barré	PNH e humanos	<i>Aedes aegypti</i> <i>Aedes albopictus</i>

Fonte: WEAVER; BARRET, 2004 e GROSSI-SOYSTER et al., 2017 (*adaptado*).

1.1.1 Família *Flaviviridae*

A família *Flaviviridae* compreende vírus envelopados com cerca de 40 a 60 nm de diâmetro, com um genoma de RNA fita simples e polaridade positiva, contendo aproximadamente 9.000 a 13.000 bases (**Tabela 2**). Até o presente momento, de acordo com o Comitê Internacional de Taxonomia de Vírus (do inglês, *International Committee on Taxonomy of Viruses*; ICTV), a família *Flaviviridae* compreende 89 espécies responsáveis por causar diversas doenças em aves e mamíferos, afetando especificamente os seres humanos, podendo ocasionar a morte desses indivíduos (SIMMONDS et al., 2017).

Tabela 2. Características de vírus da família *Flaviviridae*

Característica	Descrição
Protótipo	17D (X03700), espécie <i>Yellow fever virus</i> (YFV), gênero <i>Flavivirus</i> .
Vírus	Envelopados, 40–60 nm com um único nucleocapsídeo (exceto gênero <i>Pegivirus</i>) e duas ou três glicoproteínas de envelope.
Genoma	RNA fita simples de polaridade positiva (ssRNA+), com cerca de 9 a 13 kb.
Multiplicação	Citoplasmática, em vesículas de membranas do retículo endoplasmático (RE). Brotamento no lúmen do retículo e secreção transportada por vesículas.
Tradução	Diretamente do RNA genômico, via estrutura <i>cap</i> tipo 1 (<i>Flavivirus</i>) ou sítio de entrada ribossomal interno (IRES; <i>Hepacivirus</i> , <i>Pegivirus</i> e <i>Pestivirus</i>).
Hospedeiros	Mamíferos e aves. São classificados como arbovírus a maioria das espécies do gênero <i>Flavivirus</i> .
Taxonomia	Quatro gêneros compreendendo 89 espécies.

Fonte: ICTV, *Virus Taxonomy*, 10th Report, 2017, revised 2019 (*adaptado*).

1.1.1.1 Taxonomia

A família *Flaviviridae* é constituída por quatro gêneros: *Hepacivirus* (do grego *hepar* e/ou *hepatos*, “fígado”), da qual pertence o vírus da Hepatite C (HCV); *Pegivirus* (*Pe*, “persistente” e *g* “Hepatite G”), que infectam diversos mamíferos; *Pestivirus* (do latim *pestis*, “praga”), que infectam porcos e ruminantes, como os vírus causadores da febre suína e diarreia bovina; além do gênero *Flavivirus* (do latim *flavus*, “amarelo”), que abrange patógenos humanos de grande importância clínica, como *Dengue virus* (DENV), *Yellow fever virus* (YFV) e *Zika virus* (ZIKV) (MUKHOPADHYAY; KUHN; ROSSMANN, 2005; SIMMONDS et al., 2017).

Os membros dessa família se diferem quanto à sua antigenicidade, embora reações cruzadas na sorologia possam ocorrer entre espécies de um mesmo gênero. Nesse sentido, sequências de domínios conservados na RNA polimerase dependente de RNA (do inglês, *RNA-dependent RNA polymerase* – RdRp) são utilizadas para agrupar os membros da família dentre os quatro gêneros (SIMMONDS *et al.* 2017).

1.1.2 O gênero *Flavivirus*

O gênero *Flavivirus* compreende 53 espécies (embora existam algumas espécies relacionadas ao gênero, mas ainda não classificadas), sendo a maioria transmitida por artrópodes, nos quais os vírus se multiplicam de forma ativa. O gênero também pode ser classificado contendo três grandes grupos: (i) vírus transmitidos por mosquitos como DENV, *Japanese encephalitis virus* (JEV), *Saint Louis encephalitis virus* (SLEV), *West Nile virus* (WNV), YFV e ZIKV; (ii) vírus transmitidos por carrapatos, como o *Tick-borne encephalitis virus* (TBEV); e (iii) vírus sem vetor artrópode ainda conhecido, por exemplo, *Tamana bat virus* (TABV), transmitido entre roedores e morcegos. Usualmente, mamíferos e aves são os hospedeiros primários, acometidos por infecções assintomáticas até febres hemorrágicas e acometimentos neurológicos. Nesse cenário, podem acarretar em prejuízos econômicos e veterinários, em razão da infecção em animais domésticos e de criação (MUKHOPADHYAY; KUHN; ROSSMANN, 2005; SIMMONDS *et al.*, 2017).

1.1.2.1 Ciclo de multiplicação

Assim como outros vírus de RNA de polaridade positiva, o ciclo de multiplicação dos flavivírus ocorre no citoplasma das células do hospedeiro. As partículas virais se adsorvem a receptores na superfície celular por interação da proteína do envelope ao chegarem no tecido do hospedeiro, sendo internalizadas por endocitose mediada por clatrina. No pH ácido do endossomo ocorre um rearranjo conformacional do dímero da proteína E, fusionando o envelope viral com a membrana do endossomo, permitindo o desnudamento da partícula e liberação do RNA para o citoplasma. A replicação do genoma ocorre associada a membranas celulares, sendo traduzido a uma única poliproteína associada ao retículo endoplasmático. Próximas a sua porção N-terminal se encontram codificadas as três proteínas estruturais do capsídeo, membrana e envelope, enquanto as proteínas não estruturais se dispõem no restante do polipeptídeo, em que se localizam a protease, helicase e a RNA polimerase (MUKHOPADHYAY; KUHN; ROSSMANN, 2005; NEUFELDT *et al.*, 2018).

Logo após a clivagem da poliproteína por proteases virais e celulares (furinas e secretases) ocorre a montagem da partícula e maturação do vírus, com rearranjo e glicosilação da proteína do envelope e clivagem da proteína de membrana. A formação do envelope lipídico acontece por brotamento de membranas intracelulares e as partículas são transportadas por vesículas citoplasmáticas nas vias secretórias da célula antes de serem liberadas pelo processo de exocitose, na qual também é possível a liberação de RNA infeccioso via exossomos (**Figura 1**) (RAMAKRISHNAIAH et al., 2013; FERNÁNDEZ-SANLÉS et al., 2017).

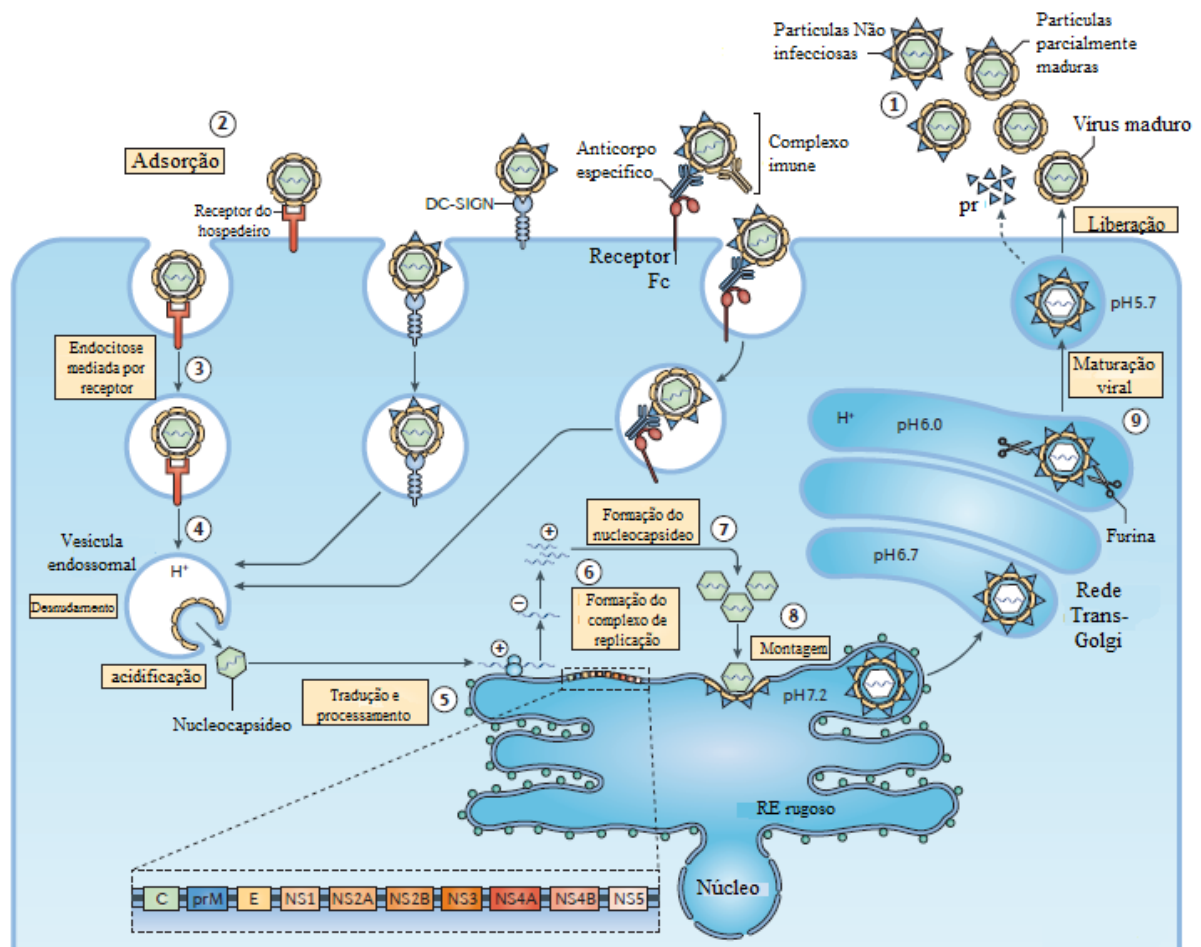


Figura 1. Ciclo de multiplicação do gênero *Flavivirus*. Partículas virais maduras e parcialmente maduras são infecciosas (1), podendo se ligar a um grupo de receptores celulares, aderindo à célula do hospedeiro (2). Ocorre a internalização dos vírus por endocitose mediada por clatrina (3), seguido da fusão da vesícula formada com o endossomo, engatilhada no baixo pH, promovendo o desnudamento da partícula e liberação do genoma para o citoplasma (4). O RNA é traduzido a uma poliproteína que é processada por proteases virais e celulares do hospedeiro (5). Proteínas não estruturais do vírus realizam a replicação do genoma (6) e, em seguida, o RNA recém-sintetizado é encapsulado pela proteína estrutural do capsídeo, formando um nucleocapsídeo (7). A montagem (8) da partícula ocorre na membrana do retículo endoplasmático (RE), sendo envolvida pela membrana do RE e glicoproteínas para formação de partículas imaturas. As partículas virais imaturas são transportadas

através da via secretora da célula e, no pH baixo da rede trans-Golgi (TGN) ocorre a clivagem de prM mediada por furina, liberando a partícula viral madura para o citoplasma (9). Fonte: (SCREATON et al., 2015) (*adaptado*).

1.1.2.2 Estrutura da partícula e genoma dos flavivírus

Os vírus do gênero *Flavivirus* apresentam estrutura esférica com aproximadamente 50 nm de diâmetro, de simetria icosaédrica e são distinguíveis entre a partícula madura, que apresenta as proteínas M e E associadas à membrana; e a imatura, que contém a proteína precursora prM, que é clivada durante a maturação (**Figura 2**). A proteína dimérica do envelope (E) por sua vez, é orientada paralelamente à membrana, não formando as projeções características (espículas) comuns a outros vírus (LI et al., 2008).

Dentre as três proteínas estruturais presentes no vírus (C de 11 kDa, E apresentando 50 kDa e M com 8 kDa na partícula madura, enquanto 26 kDa da prM nas partículas imaturas), a proteína E (envelope) é responsável pela interação da partícula viral ao receptor celular. Logo após penetração na célula via endocitose, a fusão de membranas no endossomo é dependente da acidificação nesse processo, permitindo a liberação do RNA viral e replicação do genoma (LINDENBACH; RICE, 2003; MURRAY; JONES; RICE, 2008).

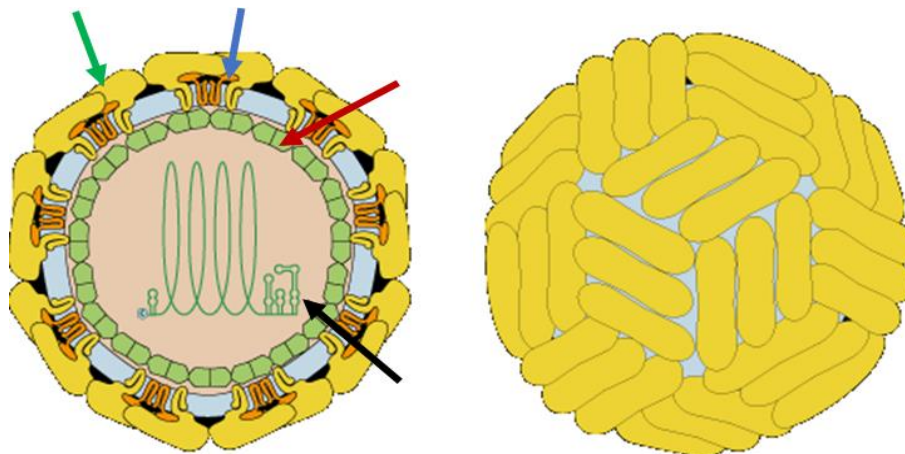


Figura 2. Representação de partícula viral do gênero *Flavivirus*. Modelo de reconstrução tridimensional da partícula de um isolado de DENV. No corte transversal estão indicados, por setas, o dímero da proteína de envelope (verde), a proteína de membrana (azul) e a proteína do capsídeo (vermelha) envolvendo o genoma viral (preto). Logo ao lado se representa a organização dos dímeros da proteína E na superfície do envelope viral. Fonte: ViralZone, 2016 (*adaptado*).

O genoma dos flavivírus é composto por uma fita simples de RNA (ssRNA) de polaridade positiva, com aproximadamente 9 a 11 kb. A porção terminal 5' do genoma ainda apresenta uma estrutura *cap* do tipo I (m-7GpppAmp), no qual as amidinas são seguidas por guaninas, enquanto a porção terminal 3' não possui uma cauda poli-A. O genoma dos vírus

maduros é constituído de uma fita simples de RNA de polaridade positiva, que funciona como molde para replicação do genoma e RNAm para a síntese da uma única poliproteína por meio de uma janela aberta de leitura (ORF), flanqueada por regiões não traduzidas (do inglês, *untranslated region* – UTR), que contém sequências conservadas do genoma viral (HEINZ; STIASNY, 2012; LEONEL et al., 2018).

Posteriormente, a poliproteína traduzida é clivada por proteases do hospedeiro ou do vírus (NS2B-NS3) em três proteínas estruturais (C, prM e E) e sete proteínas não estruturais, NS1 (46 kDa), NS2A (22 kDa), NS2B (14 kDa), NS3 (70 kDa), NS4A (16 kDa) e NS4B (27 kDa) e NS5 (103 kDa), sendo estas essenciais para a multiplicação viral (**Figura 3**) (LINDENBACH; RICE, 2003; ALESHIN et al., 2007; LEONEL et al., 2018). As proteínas estruturais são codificadas na porção N-terminal do genoma, enquanto as proteínas não estruturais estão diretamente envolvidas na replicação do RNA associadas às membranas intracelulares, não sendo detectáveis na partícula viral (MURRAY; JONES; RICE, 2008; TAKAHASHI et al., 2012).

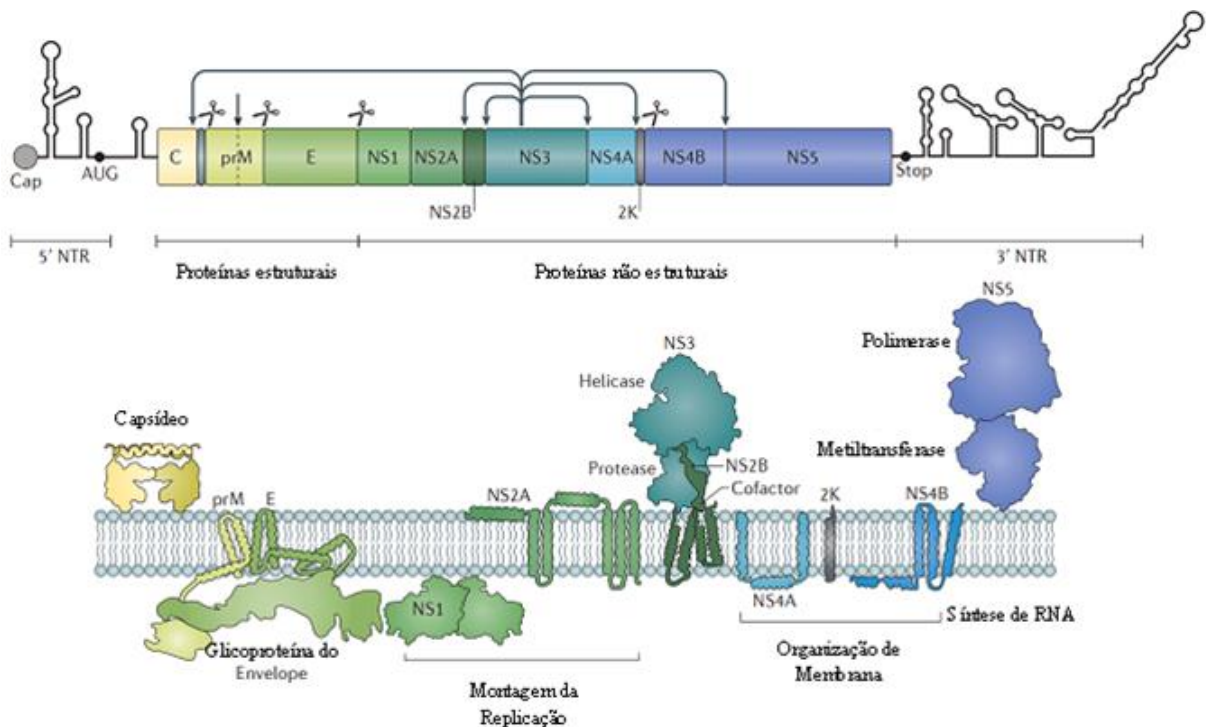


Figura 3. Organização do genoma viral e proteínas traduzidas conforme modelo em vírus do gênero *Flavivirus*. Acima: genoma de *Dengue virus* (DENV). As tesouras indicam a região de clivagem da poliproteína por proteases celulares, enquanto as setas indicam a clivagem pela protease viral e a seta vertical sinaliza a região de clivagem pela furina. Abaixo: estruturas das proteínas estruturais e não estruturais em contato com a bicamada lipídica após clivagem. Fonte: NEUFELDT, et al. 2018 (adaptado).

1.1.2.2.1 Proteínas estruturais

A proteína do capsídeo (C), com cerca de 11 kDa, é a primeira a ser traduzida. Esta possui uma cauda terminal hidrofóbica com aproximadamente 27% de aa de lisina e arginina, conferindo-a um caráter alcalino e permitindo sua interação com o RNA viral. Ela é responsável pela encapsidação do genoma, na qual múltiplas cópias da proteína interagem com o material genético, resultando na formação de um nucleocapsídeo de formato esférico e aproximadamente 30 nm de diâmetro (LINDENBACH; RICE, 2003). A proteína precursora da proteína de membrana (prM), por sua vez, é uma glicoproteína de aproximadamente 26 kDa traduzida após sinalização promovida pela cauda terminal da proteína C. Após sofrer uma clivagem proteolítica, a prM dará origem à proteína M, que atuará como uma chaperona da proteína do envelope (E), ou seja, evitará dobramentos incorretos na conformação desta proteína. Dessa forma, a proteína M é capaz de impedir rearranjos da proteína E durante a montagem e liberação dos vírions (LI et al., 2008; SIMMONDS et al., 2017).

Finalmente, a proteína do envelope (E), uma glicoproteína de superfície com cerca de 54 kDa participa de diversas funções importantes, incluindo a montagem da partícula e ligação ao receptor celular. Embora apresente cerca de 40% de identidade entre os flavivírus, a quantidade de aminoácidos glicosados em sua estrutura não é conservada entre as diferentes amostras de um mesmo vírus. Ainda, porções da proteína E na superfície da partícula são capazes de modular a especificidade ou interação ao receptor, fatores que permitem uma diferenciação antigênica entre espécies, além de subdividir os flavivírus em diversos complexos antigênicos. Por exemplo, os sorocomplexos de JEV, SLEV e WNV apresentam antígenos imunologicamente distintos de vírus como DENV ou YFV (MUKHOPADHYAY; KUHN; ROSSMANN, 2005; SIMMONDS et al., 2017). Além disso, a proteína E é o principal alvo de anticorpos neutralizantes, que atuam no bloqueio da infecção inibindo a interação ao receptor, o processo de endocitose ou fusão das membranas no endossomo (HEINZ; STIASNY, 2012).

1.1.2.2.2 Proteínas não estruturais

A proteína não estrutural 1 (do inglês, *non-structural 1* – NS1) contém aproximadamente 46 kDa e pode ser encontrada em três formas: disposta na superfície da célula, em associação com a célula do hospedeiro ou na matriz extracelular. A forma intracelular é localizada em sítios de síntese do RNA, e tem uma participação importante na replicação do genoma (LINDENBACH; RICE, 2003; MURRAY; JONES; RICE, 2008). Esta proteína é secretada por células de mamíferos infectadas, se acumulando em altos níveis no

sangue e no tecido do hospedeiro, sendo utilizada, por exemplo, no diagnóstico de alguns flavivírus, como dengue e Zika. Além disso, a NS1 também induz uma resposta imune no hospedeiro, estimulando a produção de anticorpos com atividade reativa de fixação de complemento, acarretando na lise de células infectadas. (MURRAY; JONES; RICE, 2008; HUNSPERGER et al., 2014a; NEUFELDT et al., 2018).

A NS2, por sua vez, é clivada em duas subunidades: NS2A e NS2B. A primeira (22 kDa) apresenta caráter hidrofóbico e é necessária na coordenação da replicação do RNA e auxiliando no empacotamento do genoma. Já a proteína NS2B (14 kDa), a menor delas, está associada à membrana e influencia diretamente na atividade proteolítica, formando um complexo com a protease NS3, cofator essencial à atividade desta serina-protease (MURRAY; JONES; RICE, 2008).

Dentre as proteínas não estruturais, a NS3 (70 kDa), altamente conservada entre os flavivírus, possui uma característica multifuncional com três porções distintas e essenciais na multiplicação viral: (i) serina-protease (associada a seu cofator NS2B), atuando na clivagem de junções proteicas entre NS2A–NS2B, NS2B–NS3, NS3–NS4A e NS4B–NS5, permitindo a liberação da RNA polimerase (NS5); (ii) helicase, desfazendo a forma intermediária da cadeia dupla de RNA durante a replicação, liberando as cópias do genoma; e (iii) RNA trifosfatase, necessária para adicionar a estrutura *cap* ao RNA viral replicado, necessária para o reconhecimento ribossomal na etapa de tradução. Estas funções refletem sua importância na multiplicação desses vírus, se tornando um potencial e promissor alvo molecular para o desenvolvimento de inibidores antivirais, devido a sua seletividade e especificidade aos flavivírus, além da possibilidade de interações alostéricas ou no sítio ativo (TAKAHASHI et al., 2012; OLIVEIRA et al., 2014).

A proteína NS4 é clivada em NS4A (16 kDa) e NS4B (27 kDa). Embora sejam pouco elucidadas as funções e interações às demais proteínas, é conhecido que a NS4B inibe a resposta de interferon do tipo I (IFN-1) em células do hospedeiro, além de modular a multiplicação viral ao interagir com a proteína NS3 (ZOU et al., 2015).

A proteína NS5, maior (105 kDa) e mais conservada, é a última proteína não estrutural codificada na ORF. Esta proteína apresenta uma função de RNA polimerase dependente de RNA (RdRp), bem como função de metiltransferase, essencial no *capping* do RNA. Além disso, é encontrada de forma abundante no núcleo de células infectadas, supostamente suprimindo a transcrição de genes importantes na resposta do hospedeiro, como interleucinas (PRYOR et al.,

2007; CALY; WAGSTAFF; JANS, 2012; SELISKO et al., 2014). Sua atividade enzimática também é indispensável à replicação dos flavivírus, bem como alvo de estudos para o desenvolvimento de fármacos antivirais (PASTORINO et al., 2010; OLIVEIRA et al., 2014; LIM; NOBLE; SHI, 2015). Na **Tabela 3** se dispõe um resumo das funções dessas proteínas.

Tabela 3. Proteínas não estruturais dos flavivírus

Proteína	Tamanho (kDa)	Característica	Funções principais
NS1	46	Ancorada ao RE, associada à membrana plasmática ou secretada (sNS1).	Envolvida na replicação inicial do genoma; NS1 ativa o sistema imune promovendo extravasamento vascular.
NS2A	22	Proteína de membrana hidrofóbica.	Envolvida na replicação do RNA.
NS2B	14	Proteína hidrofóbica.	Cofator da NS3 (complexo NS2B-NS3).
NS3	70	Proteína multifuncional de diversos domínios.	Protease, helicase e trifosfatase na síntese do RNA.
NS4A	16	Proteína de membrana hidrofóbica.	Associada à formação de vesículas de replicação.
NS4B	27	Proteína hidrofóbica.	Promove supressão da sinalização de IFN-1
NS5	105	Maior e mais conservada proteína não estrutural.	Polimerase dependente de RNA e metiltransferase viral.

Fonte: GUZMAN *et al.* 2016 (*adaptado*).

1.1.2.2.3 O complexo da protease NS2B-NS3

A enzima proteolítica NS2B-NS3 é classificada como uma serina-protease, em razão da presença de um aminoácido (aa) serina (Ser135) na tríade catalítica do sítio ativo da porção NS3 da protease. Esta porção se apresenta dividida nas subunidades N-terminal (NS3proN) e C-terminal (NS3proC), em que se encontram os aa da tríade, histidina51, aspartato75 e serina135 (**Figura 4**). A NS2B, por sua vez, é menor, apresenta características hidrofóbicas e atua como cofator essencial à atividade protease (TIMIRI; SINHA; JAYAPRAKASH, 2016).

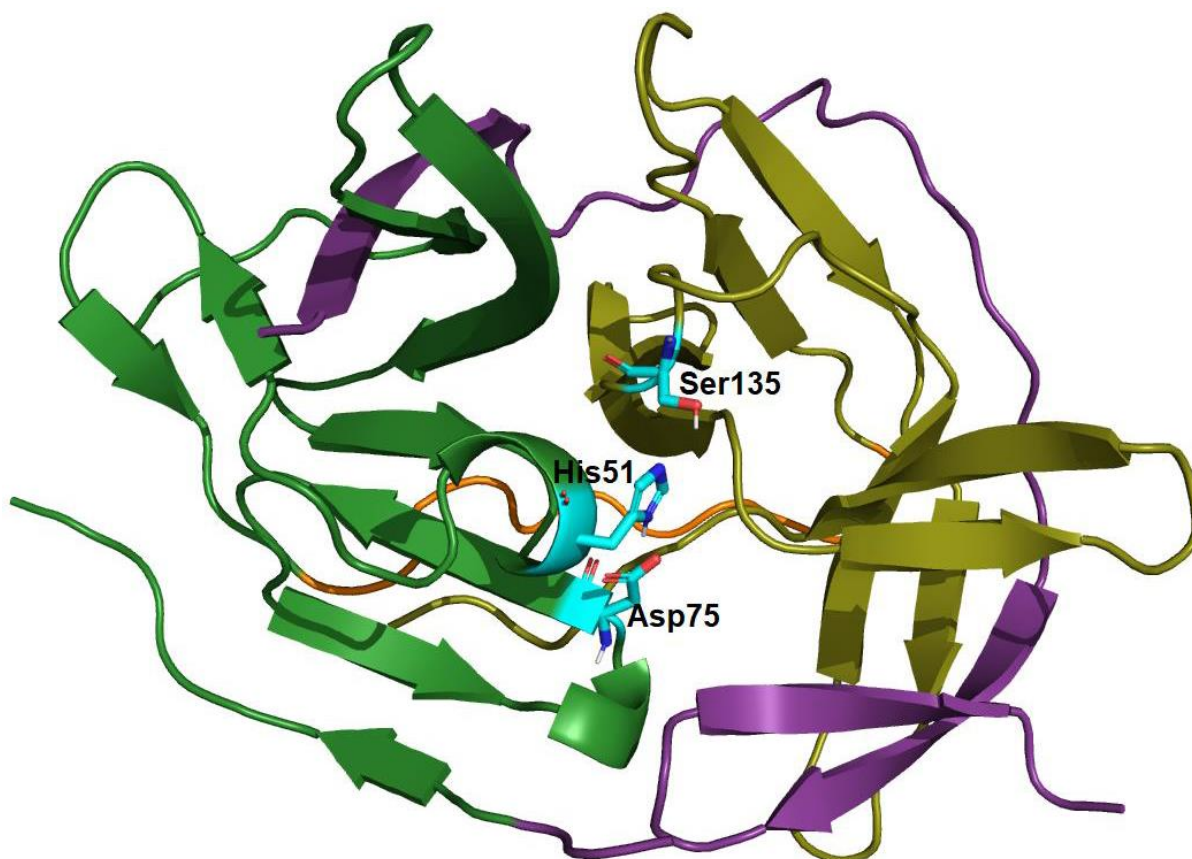


Figura 4. Estrutura tridimensional da protease (NS2B-NS3) de dengue virus. Estrutura representativa em modelo de *cartoon* da protease de DENV-3 depositada no banco de dados de proteínas (do inglês, *Protein Data Bank* – PDB) sob código **3U1I**. Porção NS3proN (verde), porção NS3proC (oliva), porção NS2B (roxo), alça ligadora (laranja) e tríade catalítica composta pelos aminoácidos His51, Asp 75 e Ser135 (ciano).

Dentre os três aminoácidos (aa) que constituem a tríade catalítica, a serina possui um papel essencial na catálise. A transferência de um próton entre a histidina e serina gera um oxianion altamente reativo, realizando um ataque nucleofílico no carbono carbonílico da ligação peptídica do substrato. Assim, ocorre a formação de um intermediário tetraédrico que sofre rearranjo, liberando o fragmento amino-terminal. Logo após, por um processo de hidrólise, a serina libera o fragmento carboxílico e a tríade catalítica é regenerada. Sua estrutura é conservada dentre os flavivírus e outros membros da família *Flaviviridae* (exceto os hepacivírus e pestivírus). Portanto, moléculas capazes de interagir com uma ou mais dessas regiões podem ser potenciais inibidores do ciclo de multiplicação dos vírus (ALESHIN et al., 2007; TIMIRI; SINHA; JAYAPRAKASH, 2016).

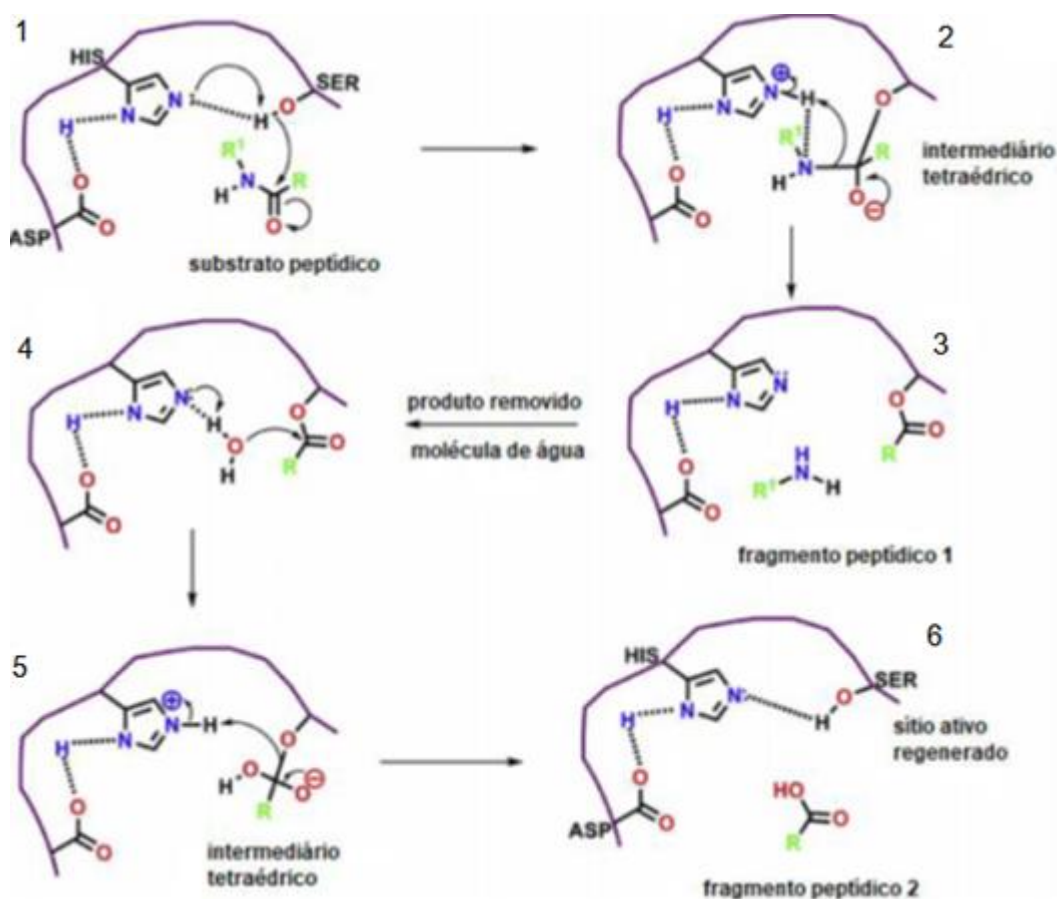


Figura 5. Sequência em etapas do mecanismo de ação da protease NS2B-NS3. Inicialmente, na triade catalítica, o oxoânion reativo da serina interage com uma carbonila do substrato, gerando um intermediário tetraédrico e fragmentando, por exemplo, um peptídeo (1 a 3). Por seguinte, o produto removido interage com moléculas de água no meio, novamente produzindo um intermediário tetraédrico, liberando o segundo fragmento peptídico (4 a 6). Fonte: TIMIRI; SINHA; JAYAPRAKASH, 2016 (*adaptado*).

Neste trabalho serão abordados três arbovírus do gênero *Flavivirus*: dengue virus (DENV), yellow fever virus (YFV) e Zika virus (ZIKV). Epidemias, surtos e novas manifestações clínicas exaltam a relevância dessas espécies, especialmente no Brasil, com os casos de Zika em 2015 e 2016, surtos de febre amarela em 2017 e 2018, além da persistente endemia de dengue ao longo das duas últimas décadas (VASCONCELOS; VASCONCELOS, 2017a; ADACHI et al., 2018; NEUFELDT et al., 2018).

1.1.2.3 Dengue virus (DENV)

O vírus da dengue é causador de uma das mais importantes doenças tropicais no mundo, com um aumento abrupto de sua incidência nas últimas décadas, associado à expansão do mosquito vetor. A ocorrência abrange regiões do mediterrâneo, Américas, sudeste da Ásia, regiões da África e Europa, com as epidemias promovendo um alto custo aos serviços de saúde

e sistemas econômicos (GUZMAN et al., 2016; SHEPARD et al., 2016). A dengue é considerada pela Organização Mundial de Saúde (OMS; do inglês, *World Health Organization* – WHO) a arbovirose mais prevalente e de mais rápida transmissão e disseminação mundial, além de ser endêmica em mais de 128 países, exceto apenas a algumas regiões da Europa, norte da América e Antártica, as quais não há presença dos mosquitos vetores, especialmente em razão das condições climáticas (WORLD HEALTH ORGANIZATION, 2009; BRADY et al., 2012; NEUFELDT et al., 2018).

Estima-se que mais de três bilhões de pessoas estão em áreas de risco de infecção por DENV, com uma ocorrência aproximada de 390 milhões (Intervalo de Confiança de 95%: 284-528 milhões) de infecções anuais, sendo 96 milhões (IC 95%: 67-136 milhões) de casos sintomáticos ocorrendo anualmente. Dentre esses, cerca de 2 milhões são casos graves da doença, com mais de 20.000 mortes por ano, números denotam um aumento significativo dos 50 milhões de casos anuais estimados em 2009, além de reforçar a importância da doença no cenário mundial (BRADY et al., 2012; BHATT et al., 2013; STANAWAY et al., 2016).

Esse cenário também se deve à disseminação dos mosquitos vetores ao redor do mundo, bem como de sua adaptação ao ciclo urbano, além da adaptação do vírus ao ser humano. Mosquitos do gênero *Aedes*, especialmente o principal vetor *A. aegypti*, emergiram há séculos de ciclos selváticos, competindo com demais espécies do gênero em florestas da África e Ásia (GUZMAN et al., 2016). Esta espécie foi introduzida no continente americano nos anos 1600, juntamente ao tráfico de escravos, se disseminando pelo mundo durante o processo de desenvolvimento industrial e urbano, interligado ao crescimento populacional (SOUTHWOOD et al., 1972).

A disseminação mais evidente do *A. aegypti* teve início no sudeste da Ásia, em razão do crescimento da região após a segunda Guerra Mundial, nas décadas de 1970 e 80, especialmente em razão das viagens aéreas, que começavam a se popularizar no mundo. Ainda, para uma população difundida mundialmente, alguns fatores como o transporte moderno, comércio internacional, crescimento populacional em áreas urbanas e, não menos importante, da existência de quatro sorotipos de DENV, favoreceram a disseminação de vetor e do vírus (HALSTEAD, 1992).

A fêmea do mosquito *A. aegypti*, bem adaptada às regiões urbanas, vive intimamente com seres humanos, em um tempo curto de uma a pouco mais de duas semanas de vida e se alimentando do sangue do hospedeiro durante o repasto sanguíneo. Esse período é suficiente

para se alimentar de indivíduos infectados, se infectar, permitir a multiplicação do vírus em seu tecido, realizar a oviposição e então transmitir os diferentes sorotipos de DENV por nova picada ou pela transmissão direta para os ovos (**Figura 6**) (GUZMAN et al., 2016).

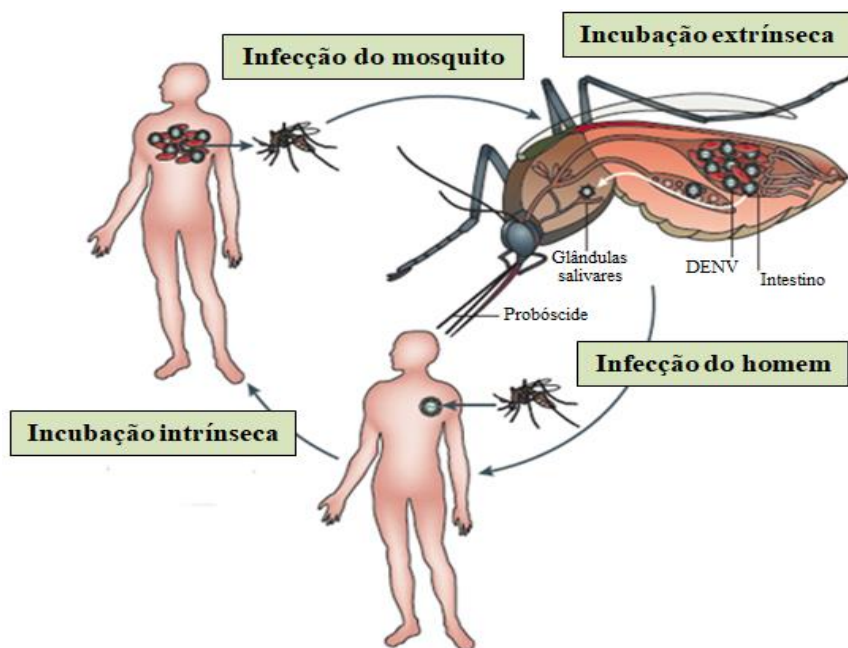


Figura 6. Ciclo de transmissão de dengue em humanos e mosquitos. A fêmea do mosquito se infecta após realizar o repasto sanguíneo de um indivíduo (ser humano) em estado virêmico. O vírus infecta primeiramente o intestino e outras células do mosquito antes de se disseminar para as glândulas salivares (incubação extrínseca). Um único mosquito pode então infectar diversos indivíduos. O indivíduo infectado apresenta sinais após 4 a 7 dias (incubação intrínseca) e, caso ainda em viremia, é capaz de transmitir o vírus a outros mosquitos, reiniciando o ciclo. Fonte: GUZMAN *et al.* 2016 (*adaptado*).

Nesse aspecto, algumas características como a resposta imune do hospedeiro, suscetibilidade e permissividade do mosquito à infecção pelo vírus, bem como sua capacidade de transmissão, além dos fatores associados entre sorotipos e genótipos, como gravidade da doença, potencial epidemiológico e transmissão pelo vetor, também são responsáveis pela incidência de infecções (RICO-HESSE, 2003). Uma relação antigênica entre esses vírus os permite serem classificados em quatro sorotipos (DENV-1 a DENV-4) baseado na similaridade da proteína do envelope (E), além de diferentes genótipos, de acordo com diferenças nas sequências nucleotídicas (**Tabela 4**). As diferenças de nucleotídeos no genoma dentre os quatro sorotipos podem ser de até cerca de 30% e, entre genótipos de um mesmo sorotipo, cerca de 6% (YAMASHITA et al., 2013). Os genótipos possuem diferentes características de transmissão pelo vetor, infecção no hospedeiro humano e de patogênese nestes indivíduos, além de apresentarem diferentes distribuições geográficas (**Figura 7**) (RICO-HESSE, 2003).

Tabela 4. Sorotipos e respectivos genótipos de DENV

Sorotipo	Genótipos
DENV-1	I, II, III, IV e V (5).
DENV-2	Americano, Asiático I, Asiático II, Asiático–Americano, Cosmopolita e Selvático (6).
DENV-3	I, II, III, IV e V (5).
DENV-4	I, IIA, IIB, III e Selvático (5).

Fonte: YAMASHITA *et al.* 2013 (*adaptado*).

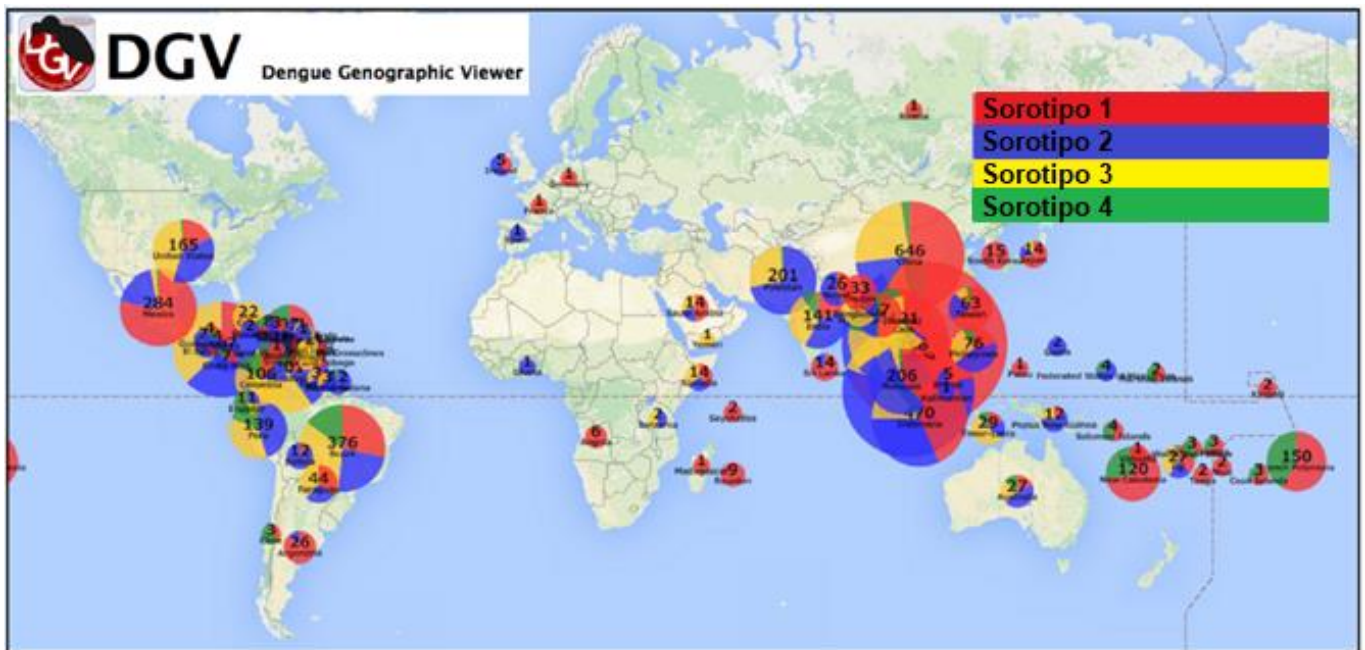


Figura 7. Distribuição global de sequências de DENV até 2016. Figura representativa da distribuição mundial de sequências de dengue virus (DENV) reportadas até o ano de 2016 (*Dengue Geographic Viewer*[®]). Sorotipos de DENV-1 (vermelho), DENV-2 (azul), DENV-3 (amarelo) e DENV-4 (verde) representados em gráficos circulares por país. Fonte: YAMASHITA *et al.* 2016 (*adaptado*).

Essas diferenças de sequências entre genótipos e sorotipos interferem na multiplicação e na gravidade da doença (sinais e sintomatologia clínicos) em uma infecção. Neste aspecto, de acordo com a OMS, a dengue pode ser classificada entre: dengue, dengue com sinais de alarme e dengue grave (**Figura 8**) (WORLD HEALTH ORGANIZATION, 2009; GUZMAN *et al.*, 2016).

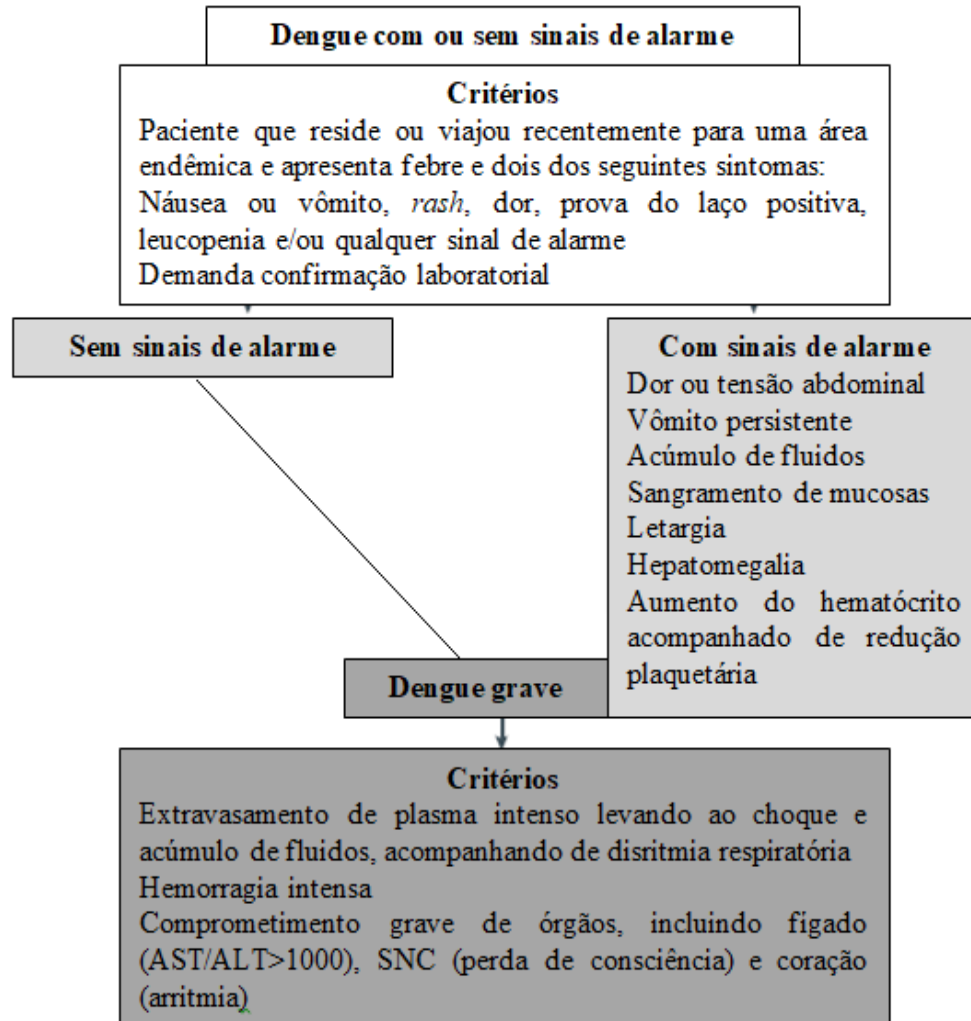


Figura 8. Critérios para classificação da gravidade de dengue. Critérios da Organização Mundial de Saúde descrevendo sinais e sintomas clínicos para classificação e confirmação de dengue sem sinais de alarme, dengue com sinais de alarme e dengue grave. Fonte: WHO, 2009; revisado por GUZMAN *et al.* 2016 (*adaptado*).

Esse padrão de triagem almeja prever e direcionar as condutas clínicas para essas três classificações possíveis da doença. Além disso, a dengue ainda pode ser dividida em três diferentes fases após o período de incubação: aguda ou febril, crítica (extravasamento de plasma) e de convalescência, caracterizando seus sinais e manifestações clínicas (**Tabela 5**) (WORLD HEALTH ORGANIZATION, 2009; PATTERSON; SAMMON; GARG, 2016a).

Tabela 5. Dengue: fases, duração, sinais clínicos e testes laboratoriais

Fase	Incubação	Fase aguda/febril					Fase crítica				Convalescência
Tempo	3-14 dias	3-7 dias					1-2 dias				3-5 dias
Sintomas	Ausentes	Febre presente					Resolução da febre				<ul style="list-style-type: none"> ▪ Reabsorção de fluidos ▪ Diurese
		<ul style="list-style-type: none"> ▪ Mialgia ▪ Rash cutâneo ▪ Petéquias ▪ Leucopenia ▪ Sangramentos 	Sinais de alarme podem ocorrer				<ul style="list-style-type: none"> ▪ Extravasamento capilar ▪ Choque ▪ Hemorragia severa ▪ Comprometimento de órgãos 				
Testes	-						DENV IgM				
		DENV NS1									
		DENV PCR									
Período da doença (dias)	0	1	2	3	4	5	6+	7	8	9	10+

Fonte: PATTERSON J; SAMMON M e GARG M, 2016 (*adaptado*).

As infecções por DENV estabelecem uma imunidade eficiente e duradoura àquele sorotipo para toda vida do hospedeiro, prevenindo uma nova infecção por este mesmo vírus (homotípica). Entretanto, ainda que ocorra uma reação cruzada na sorologia para os diferentes sorotipos de DENV, uma exposição e resposta a um sorotipo não irá conferir proteção contra outro sorotipo (heterotípica). Os casos de dengue grave, por exemplo, podem estar mais associados à uma segunda exposição heterotípica, sendo sugerida a hipótese do fenômeno denominado exacerbação dependente de anticorpos (do inglês, *Antibody Dependent Enhancement* – ADE). Nesta hipótese, um segundo sorotipo de dengue poderia utilizar anticorpos anti-DENV para facilitar sua penetração em macrófagos, acelerando o processo de multiplicação viral e, possivelmente, culminando em uma exacerbação inflamatória e do quadro clínico (VAN DER SCHAAR; WILSCHUT; SMIT, 2009; MIDGLEY et al., 2011).

Discussões acerca dessa hipótese também atribuem casos de dengue grave a uma resposta heterotípica linfocitária (T CD8+), durante uma infecção secundária por DENV. Essa resposta mediada por células T é considerada insuficiente, diferente de respostas primárias, em que ocorre uma maior produção de citocinas e atividade citotóxica. Curiosamente, esse fenômeno ainda não foi observado em outros vírus do gênero, ainda que co-circulantes, como ZIKV (ROTHMAN, 2011; RASMUSSEN et al., 2018).

1.1.2.4 Zika virus (ZIKV)

O vírus Zika foi isolado pela primeira vez em 1947, a partir de um macaco sentinela da reserva florestal de Zika em Uganda, no leste africano. O primeiro caso notificado de uma infecção em humanos ocorreu em 1954 na Nigéria, mas a primeira confirmação de uma infecção pelo vírus Zika foi descrita em Uganda, entre 1962 e 1963 (WIKAN; SMITH, 2016). Embora houvesse conhecimento da infecção no ser humano, a doença teve pouca investigação ao longo dos anos, possivelmente tendo seu diagnóstico confundido com dengue e outras arboviroses, em razão de seus sinais clínicos semelhantes e reação cruzada em testes sorológicos. Somado à ausência de relatos de epidemias e surtos, bem como de subnotificações da doença e alta taxa de casos assintomáticos, uma transmissão silenciosa e esporádica pôde disseminar o vírus ao longo dos continentes africano e asiático, com menos de 20 infecções confirmadas ao longo de 60 anos da descoberta da doença (BAUD et al., 2017).

No ano de 2007, com um surto na ilha de Yap, seguido de pequenos surtos no sul do pacífico até 2013, com um grande surto Polinésia Francesa, mudariam o curso histórico da Zika (DUFFY et al., 2009; MUSSO; NILLES; CAO-LORMEAU, 2014). No Brasil, uma linhagem Asiática do vírus foi introduzida entre os anos de 2013 e 2015, se disseminando rapidamente e gerando uma epidemia no final de 2015, até, posteriormente, relatos confirmarem a presença da doença na América Latina e América do Norte, além de novos surtos em Cabo Verde na África e Cingapura na Ásia (**Figura 9**) (PETTERSSON et al., 2016; MUSSO; LANTERI, 2017).

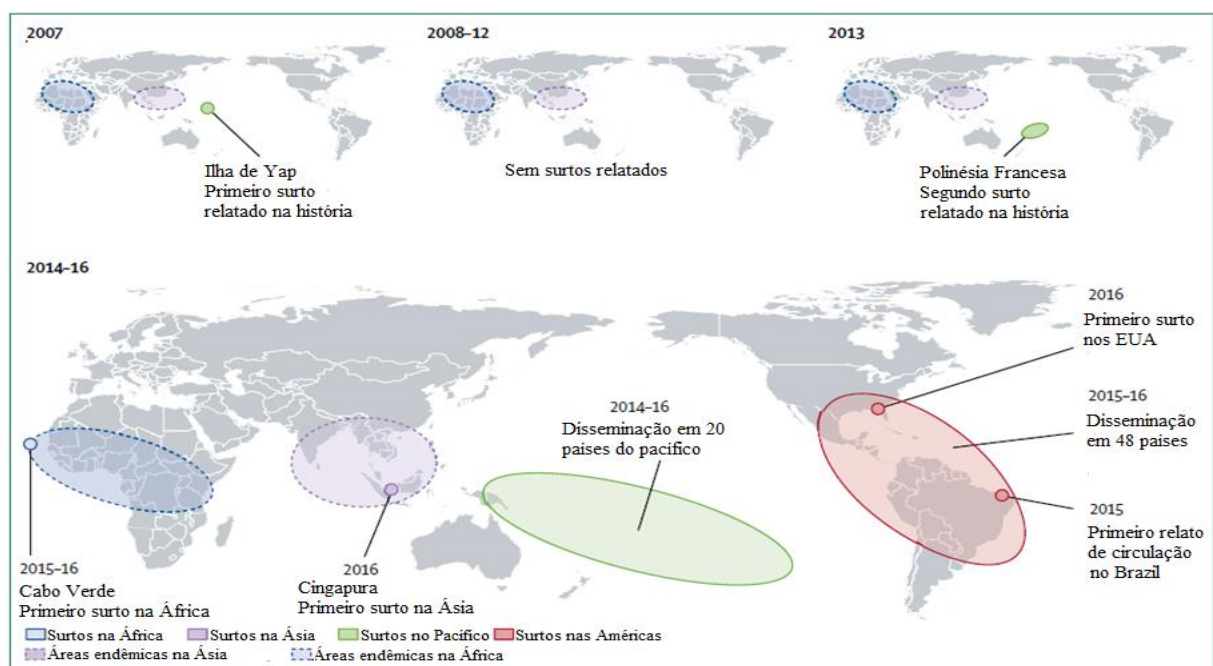


Figura 9. Surtos de Zika ao longo da última década (2007-2017). Primeiro surto na ilha de Yap em 2007. Ausência de relatos até o segundo surto na Polinésia Francesa em 2013. Disseminação na América em 2015 e 2016, com os primeiros surtos no Brasil (2015) e Estados Unidos (2016). Em paralelo, no mesmo período ocorrem os primeiros surtos na África (Cabo Verde) e Ásia (Cingapura). Fonte: BAUD *et al.* 2017 (*adaptado*).

A semelhança genética e epidemiológica entre Zika e dengue, bem como os mesmos fatores de adaptação do vetor e condições climáticas, além do ambiente urbano e densidade populacional, reforçam a provável hipótese que justifica a disseminação de Zika no mundo: a disseminação geográfica de isolados do vírus e suas mudanças genéticas, ambos associados aos fluxos migratórios, podem ter originado linhagens mais virulentas e melhor adaptadas para infecções em seres humanos (YOKOYAMA; STARMER, 2017).

Os isolados de ZIKV podem ser classificados em duas ou três principais linhagens: Africana, Asiática (incluindo as linhagens americanas) e a designada Africana II, uma linhagem negligenciada circulante na África, que consiste de um grupo irmão das linhagens Africana e Asiática previamente identificadas (SAIZ *et al.*, 2017; ZHANG *et al.*, 2017). Embora diferentes linhagens existam, suas relações antigênicas se apoiam na existência de um único sorotipo, o que pode ser crucial para o desenvolvimento futuro de uma vacina ou favorecer o desenvolvimento de fármacos anti-panflavivírus (DOWD *et al.*, 2016).

Enquanto uma vacina seria uma solução agente-específico, especialmente nos casos em gestantes abordados a seguir, uma estratégia panflavivírus também se mostra relevante, haja visto que mais de 50% das espécies de flavivírus estão associadas por causarem doenças no ser humano. Dentre as mais importantes, como encefalite japonesa e encefalite do oeste do Nilo, febre amarela, dengue e Zika também denotam relevância, especialmente por estarem disseminadas mundialmente (PATTERSON; SAMMON; GARG, 2016b; FERNÁNDEZ-SANLÉS *et al.*, 2017).

Assim como descrito para outros flavivírus, após infectar um hospedeiro, o ZIKV se dissemina para os nódulos linfáticos e sangue, alcançando vários órgãos. Cerca de 80% dos casos de infecção em humanos são assintomáticos (SHILEY, NATARO, 2017). Os pacientes sintomáticos apresentam sintomatologia clínica semelhante a dengue e outras arboviroses, caracterizada por febre, artralgias, mialgias, exantema cutâneo, conjuntivite, febre hemorrágica e acometimentos do sistema nervoso central (SNC), como meningoencefalites. Alguns sinais e sintomas (**Tabela 6**) são específicos a cada uma dessas duas doenças, auxiliando no diagnóstico clínico diferencial entre elas e demais arboviroses.

Tabela 6. Sinais e sintomas clínicos de dengue e Zika

SINAIS / SINTOMAS	Dengue	Zika
Febre (duração)	Acima de 38 °C (4 a 7 dias)	Ausente/subfebril (1 a 2 dias)
Manchas na pele (frequência)	4º dia (30-50% dos casos)	1º ou 2º dia (90% dos casos)
Mialgias (frequência)	+++	++/+++
Dores articulares	+ /+++	++/+++
Intensidade de dores articulares	Leve	Leve/moderada
Edemas articulares	Raro	Frequente e intensidade leve
Conjuntivite	Raro	50-90%
Dores de cabeça (frequência)	+++	++
Coceira	Leve/moderada	Moderada/intensa
Hipertrofias ganglionares (frequência)	Leve	Moderada/intensa
Discrasia hemorrágica (frequência)	Moderada	Ausente
Acometimentos neurológicos	Raro	Mais frequente que Dengue

Fonte: Hospital das Clínicas da Faculdade de Medicina de Botucatu (UNESP), 2016 (*adaptado*). Barras (+/+++) indicam faixa de intensidade/frequência.

Algumas manifestações clínicas apresentaram relevância para infecções por ZIKV, como a síndrome congênita em fetos e recém-nascidos, ressaltando, por exemplo, os casos de microcefalia. Além disso, a síndrome de Guillain-Barré (SGB) em adultos, bem como meningoencefalites em infecções por Zika em adultos e crianças (**Figura 10**), também remeteram à relevância clínica dos casos, em razão do atraso no desenvolvimento e morbidade associada (EILLANNY; MARINHO, 2015; SAIZ et al., 2017; SILVA MARINHO; KROON, 2019).

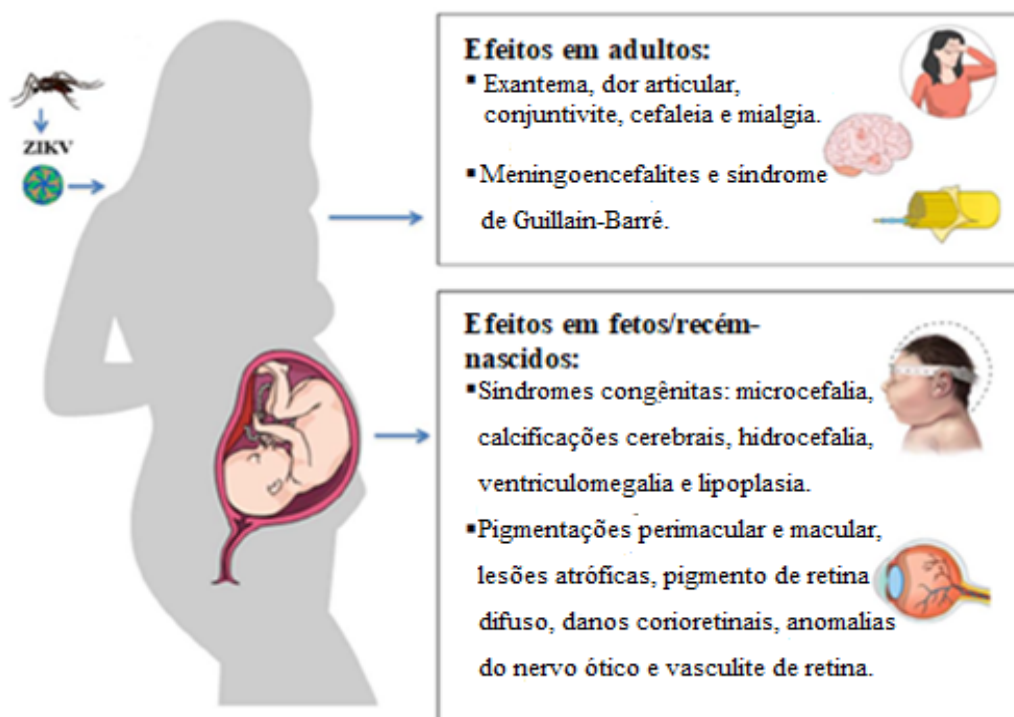


Figura 10. Manifestações clínicas e consequências da infecção por ZIKV. Ilustração esquemática dos efeitos da infecção por *Zika virus* (ZIKV) em adultos e fetos e/ou recém-nascidos. Fonte: SAIZ *et al.* 2017 (*adaptado*).

O primeiro relato de ZIKV associado com complicações neurológicas foi durante o surto na Polinésia Francesa em 2013, com um aumento dos relatos dos casos de SGB (MUSSO; NILLES; CAO-LORMEAU, 2014). A polineuropatia periférica aguda (ou síndrome de Guillain-Barré) pode ser desencadeada por fatores não infecciosos e infecciosos, como vírus e bactérias (CAUCHEMEZ *et al.*, 2016). O desenvolvimento dessa síndrome também já havia sido anteriormente descrito após infecções por DENV, mas o aumento de notificações na epidemia no Brasil desde 2015 a tornou mais associada a infecções por ZIKV (SANTOS *et al.*, 2004; FRAGOSO *et al.*, 2016; MISHRA; BEHERA, 2016).

Os casos de microcefalia, no entanto, ainda não haviam sido associados a outros flavivírus (LAZEAR; DIAMOND, 2016). Desde o início da epidemia de ZIKV no Brasil, um aumento no número de casos de microcefalia no país foi observado principalmente na região Nordeste, indicando uma relação entre a infecção por ZIKV e transmissão vertical em gestantes (DE ARAÚJO *et al.*, 2016; MLAKAR *et al.*, 2016).

1.1.1.5.1 Yellow fever virus (YFV)

O vírus da febre amarela é originário da África, trazido às Américas durante o tráfico de escravos no século XVII. Já o primeiro surto da doença foi relatado em 1648, em Yucatán, no

México (BRYAN; MOSS; KAHN, 2004). Durante os dois séculos seguintes, diversos surtos foram relatados na porção tropical do continente americano, cidades na América do Norte e na Europa. Apenas no século XIX, Carlos Finlay sugeriu uma transmissão da doença via vetor artrópode (mosquito), confirmada, posteriormente, por Walter e Reed, refutando uma hipótese de transmissão via contato direto (GARDNER; RYMAN, 2010).

O primeiro isolado do YFV ocorreu décadas depois, em 1927, dando origem à primeira linhagem do vírus, Asibi, em homenagem ao nome do próprio paciente, até que Max Theiler e colaboradores viriam a produzir a primeira linhagem vacinal em 1930, denominada 17D, resultante de sucessivas passagens da amostra Asibi. Alguns testes demonstraram a atenuação do vírus, ainda mantendo seu potencial imunogênico e a capacidade de estimular resposta imune em homens e macacos. Posteriormente, essa seria a linhagem utilizada como vacina contra febre amarela (STAPLES; MONATH, 2008; MONATH; VASCONCELOS, 2015).

A vacina é a melhor medida profilática contra a febre amarela, uma estratégia que pôde, no Brasil, eliminar a circulação do vírus no ciclo urbano em 1940. Contudo, no ano de 2017 ocorreu novo surto da doença, em razão da falta nos estoques da vacina da FA e baixa cobertura vacinal em algumas regiões do país, com estimativas chegando até cerca de 13% da população em 2016 (CAMPI-AZEVEDO et al., 2016; SHEARER et al., 2017; VASCONCELOS; VASCONCELOS, 2017b; POSSAS et al., 2018). Nos serviços públicos brasileiros, a vacina é produzida pelo Instituto de Tecnologia em Imunobiológicos Bio-Manguinhos, a partir da amostra 17DD atenuada, cultivada em ovos embrionados de galinha e livre de outros agentes patogênicos, como preconizado pela OMS (MINISTÉRIO DA SAÚDE, 2019).

Os seres humanos são infectados esporadicamente após o repasto sanguíneo de mosquitos silvestres dos gêneros *Aedes*, *Haemagogus* e *Sabethes*, previamente infectados ao picar primatas não-humanos (PNH) em estado de viremia (febre amarela silvestre). Espécies desses gêneros, especialmente *A. aegypti* também podem atuar como reservatórios na transmissão entre humanos, na febre amarela urbana. A infecção pode progredir para uma doença grave, caracterizada por febre, náuseas e vômitos, dor epigástrica, hepatite, falência renal, hemorragias, síndrome do choque sistêmico e morte em até 60% dos casos. A febre amarela é tipicamente uma febre hemorrágica viral, com sinais e sintomas semelhantes a outras arboviroses, como a dengue, embora apresente uma gravidade maior associada à lesão do tecido hepático (BARNETT, 2007; MONATH; VASCONCELOS, 2015).

Curiosamente, o continente africano apresenta uma grande disseminação da doença e possui, de acordo com a literatura, uma mortalidade inferior a América do Sul. Isso pode indicar que o efeito de fatores genéticos associados à evolução e seleção na região, bem como a resposta dos hospedeiros, determinam características da doença, como a letalidade e até a produção de anticorpos neutralizantes contra o vírus após a vacinação (MONATH et al., 2002; TUBOI et al., 2007).

1.1.3.1 Epidemiologia de febre amarela, dengue e Zika

Os quatro sorotipos de DENV se encontram circulantes no mundo há pelo menos 800 anos, desde então se mantendo pouco associados a casos de surtos e disseminação, além de estarem restritos geograficamente à região africana e asiática. Logo após os desdobramentos da Segunda Guerra Mundial e consequente transporte do mosquito vetor pelo mundo, os primeiros casos, especialmente de dengue grave, ganharam relevância, especialmente nas primeiras epidemias na década de 1950 nas Filipinas e Tailândia, bem como dos primeiros casos no Caribe e América Latina no início da década de 1980. Desde então, a dengue tem atingindo uma abrangência geográfica em quase todos os continentes, além de ser considerada uma das doenças tropicais mais importantes (GUZMAN; HARRIS, 2015).

Já para o vírus Zika, desde 2008 a disponibilidade de informações tem aumentado, incluindo dados sobre epidemiologia, manifestações clínicas e até a possível transmissão sexual (MUSSO; LANTERI, 2017). Este caso em específico foi primeiramente sugerido após evidências clínicas e sorológicas em que um cientista norte-americano havia contraído ZIKV enquanto trabalhava no Senegal em 2008 e acabou transmitindo-o a sua esposa ao retornar para casa (FOY et al., 2011). Uma evidência adicional foi encontrada durante um surto de ZIKV na Polinésia Francesa em 2013, em que o vírus foi isolado do sêmen de um paciente no Taiti que procurou tratamento para um caso de hematospémia (presença de sangue no esperma) (D'ORTENZIO et al., 2016).

Além dessa descoberta, pesquisas retrospectivas identificaram um aglomerado heterogêneo e incomum de malformações cerebrais congênitas e disfunções do tronco cerebral em fetos e recém-nascidos após a epidemia de ZIKV na Polinésia Francesa, sugerindo uma associação causal entre a infecção e esta manifestação clínica (CAUCHEMEZ et al., 2016; DE ARAÚJO et al., 2016). Logo após sua aparição inicial no Brasil, em maio de 2015, ZIKV se disseminou rapidamente pelas Américas em 2016, ano em que 48 países e territórios das Américas informaram mais de 500.000 casos suspeitos de Zika, sendo 175.063 casos

confirmados. Além disso, 22 países e territórios relataram 2.439 casos de síndrome congênita associada ao Zika, além de casos via transmissão sexual em mais cinco deles (IKEJEZIE et al., 2017; WHO, 2017).

A febre amarela, por sua vez, mostrou importância especialmente entre os séculos XVIII e XX, em que epidemias disseminaram os vírus por zonas costeiras em regiões da América do Norte e Caribe. Desde a identificação de vetores como *A. aegypti*, além da promoção de estratégias de controle, ocorreu uma redução dos casos de febre amarela em regiões não-endêmicas. Posteriormente, o desenvolvimento da vacina de febre amarela na década de 1930 e 40 reduziu significativamente os casos da doença, ainda que ocorressem surtos e epidemias em regiões sem programas de imunizações ou sem cobertura vacinal eficaz, a exemplos das epidemias no Brasil desde 2016 (MONATH; VASCONCELOS, 2015; VASCONCELOS, 2017).

No Brasil, dados do Ministério da Saúde remetem à gravidade das infecções por DENV, ZIKV e YFV nos últimos anos, associados aos diversos casos de dengue, febre amarela e Zika reportados. Somente nos anos de 2015 e 2016, foram estimados cerca de 1,5 milhões de casos para dengue, enquanto Zika apresentou uma estimativa de mais de 200.000 casos em 2016, com cerca de 3.000 casos de má formação do SNC em fetos e neonatos. Já em 2017, cerca de 155.000 e 14.000 casos foram estimados para dengue e Zika, respectivamente. Uma das possíveis causas para essa queda se deve às prováveis subnotificações concomitante ao surto de febre amarela ainda naquele mesmo ano, no qual se confirmaram 1.376 casos e 483 óbitos no período de julho de 2017 a junho de 2018. No mesmo período de 2016, foram confirmados 632 casos e 201 óbitos. (VASCONCELOS; VASCONCELOS, 2017b; GARCIA, 2018; MINISTÉRIO DA SAÚDE, 2019). No ano de 2019 (semana epidemiológica 1 a 34), cerca de 1.439.471 casos prováveis de dengue foram relatados no Brasil, comparado aos 205.791 em 2018. Zika, por sua vez, apresentou cerca de 9.813 casos prováveis em 2019, comparado aos 6.669 em 2018. Ainda em 2019, dentre os 1.281 casos notificados de febre amarela, 82 destes foram confirmados com 14 óbitos (MINISTÉRIO DA SAÚDE, 2019).

1.1.3.2 Diagnóstico de febre amarela, dengue e Zika

Todos os flavivírus se relacionam sorologicamente, o que é demonstrado em técnicas de detecção e diagnóstico aplicados a esses vírus, como ensaios ELISA (do inglês, *Enzyme-linked Immunosorbent Assay*) e o método de inibição da hemaglutinação. Ainda sim, ensaios de

neutralização são indicados para discriminar os sorotipos e são utilizados para identificar sorocomplexos mais relacionado dentre os flavivírus, ainda que não seja totalmente específico dentre as espécies do gênero. A proteína E é o principal alvo para anticorpos neutralizantes, embora também induza reações cruzadas com anticorpos não-neutralizantes e, por tal razão, sítios antigênicos alternativos de domínios da proteína E, prM e NS1 podem induzir anticorpos detectáveis no diagnóstico (MURRELL; WU; BUTLER, 2011).

O diagnóstico de DENV, YFV e ZIKV consiste na detecção do RNA genômico ou de proteínas virais, além do isolamento viral, bem como de testes sorológicos. Dentre eles, podem ser citados a reação em cadeia da polimerase com transcriptase reversa (RT-PCR), a inoculação de amostra em culturas celulares e a detecção de anticorpos anti-DENV, anti-YFV e anti-ZIKV (PATTERSON; SAMMON; GARG, 2016b).

Dentre algumas das abordagens diagnósticas, a detecção de viremia para DENV é possível entre 24 e 48 horas antes do início da febre, persistindo por 5 a 6 dias (**Figura 11**). Durante esse período, o vírus, o RNA genômico e a proteína NS1 podem ser detectados no sangue, soro e plasma do paciente e também nos tecidos de casos fatais. Métodos virológicos, moleculares e sorológicos são usados para confirmar a infecção por DENV para vigilância epidemiológica e diagnóstico clínico (WORLD HEALTH ORGANIZATION, 2009; HUNSPERGER et al., 2014b; PATTERSON; SAMMON; GARG, 2016b).

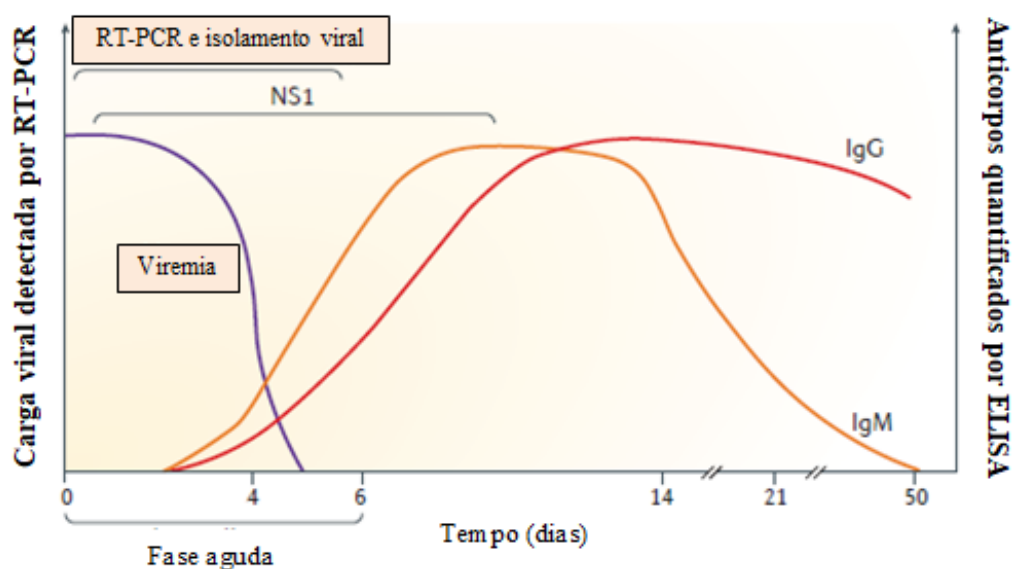


Figura 11. Marcadores moleculares e sorológicos na infecção por DENV: NS1, IgM, IgG, ELISA e RT-PCR. Variação da viremia, concentração do antígeno NS1 e de anticorpos ao longo do tempo. Diferentes testes serão adequados dependendo do estágio de infecção. ELISA – Ensaio imunoadsorvente ligado à enzima; RT – Transcriptase reversa. Fonte: GUZMAN *et al.* 2016 (*adaptado*).

A detecção de anticorpos IgM anti-DENV é o teste mais utilizado, podendo ser mensurados qualitativamente ou quantitativamente nas diluições dos soros de pacientes na fase aguda da doença, a fim de se confirmar a infecção aguda, enquanto pacientes na fase de convalescência podem ser identificados por soroconversão de IgM e IgG (HUNSPERGER et al., 2014b). Para os pacientes com suspeita da doença, um diagnóstico presuntivo pode ser feito pela detecção anti-IgM em amostras coletadas próximo ao sexto dia de sintomas agudos, havendo *kits* comerciais para detecção de IgM ou IgG em ensaios imunoenzimáticos (ELISA), além dos formatos de teste rápido, menos sensíveis (HUNSPERGER et al., 2009). Ademais, uma infecção secundária por DENV é marcada pela ausência da identificação de anticorpos IgM, apenas de altos títulos de IgG, o que denota um baixo valor de diagnóstico de IgM por si só (CUCUNAWANGSIH; LUGITO; KURNIAWAN, 2015).

Embora o isolamento e a identificação do vírus sejam altamente específicos, sua sensibilidade é relativamente baixa, porém, em contraste, a detecção de RNA de DENV fornece um método rápido, sensível e específico para confirmação do diagnóstico. A detecção da proteína NS1 também fornece uma janela de oportunidade para um diagnóstico etiológico precoce, com sensibilidade e especificidade variável de acordo com o sorotipo, tempo de coleta da amostra, tipo de infecção (primária ou secundária) e modelo do teste (GUZMAN et al., 2010; HUNSPERGER et al., 2014b).

A introdução de um “novo” arbovírus em 2015 no Brasil, o ZIKV, impôs novos desafios para o diagnóstico de ambos os vírus e de outras arboviroses, uma vez que o diagnóstico é baseado na detecção de RNA e/ou detecção de IgM específico. No entanto, a duração da viremia é diferente entre as infecções e uma reação cruzada entre DENV e ZIKV pode ser observada nos testes sorológicos (FARIA et al., 2016; ADACHI et al., 2018). A detecção de IgM por ELISA, por exemplo, apresenta reação cruzada com outros flavivírus, enquanto o teste de ELISA competitivo para NS1, por sua vez, apresenta ainda uma baixa especificidade dentre os flavivírus, mesmo apresentando alta sensibilidade (STAPLES et al., 2016; BALMASEDA et al., 2017; KADKHODA; GRETCHEN; RACANO, 2017).

Pesquisas de campo também já demonstraram associação na soroneutralização de ZIKV em concomitância à DENV e YFV, corroborando as reações cruzadas. Nesse aspecto, DENV foi até mesmo considerado inicialmente como um dos causadores do surto de Zika na ilha de Yap, parte em razão de resultados falso-positivos de um teste rápido de DENV (WAGGONER; PINSKY, 2016).

Respondendo a essa problemática, a Organização Pan-Americana da Saúde (OPAS), recomendou um novo protocolo de diagnóstico para os dois vírus. Na primeira etapa, o RNA viral de DENV ou ZIKV é detectado por PCR em tempo real em amostras de casos agudos das doenças. Caso disponível, a sorologia para IgM em amostras de soro coletadas de indivíduos com suspeita clínica de dengue ou Zika deve ser testada por ELISA específico para os dois vírus e, se ambos os casos forem positivos, uma infecção secundária por flavivírus deve ser considerada (OPAS, 2017).

1.1.3.3 Prevenção e tratamento de dengue, Zika e febre amarela

Dentre as formas de prevenção atuais, duas estratégias para o controle do mosquito vetor podem ser aplicadas: evitar áreas endêmicas e a implementação de programa governamentais e não-governamentais para vigilância do vetor, a nível local e global. Alguns métodos têm apresentado propostas promissoras, como o emprego de insetos estéreis, a introdução de linhagens da bactéria *Wolbachia*, capaz de bloquear a transmissão vertical do vírus da dengue em mosquitos, além das estratégias convencionais de controle (ALPHEY et al., 2010; BEISEL; BOËTE, 2013; JEFFRIES; WALKER, 2016).

A única vacina aprovada até o momento para o uso em populações endêmicas de dengue consiste em uma forma tetravalente quimérica e recombinante, estruturada em YFV atenuado, Dengvaxia® (CYD-TDV) (Figura 12). Esta é composta de quatro vírus vacinais 17D recombinados com genes da prM e proteína E de cada um dos quatro sorotipos de dengue (TORRESI; EBERT; PELLEGRINI, 2017b).

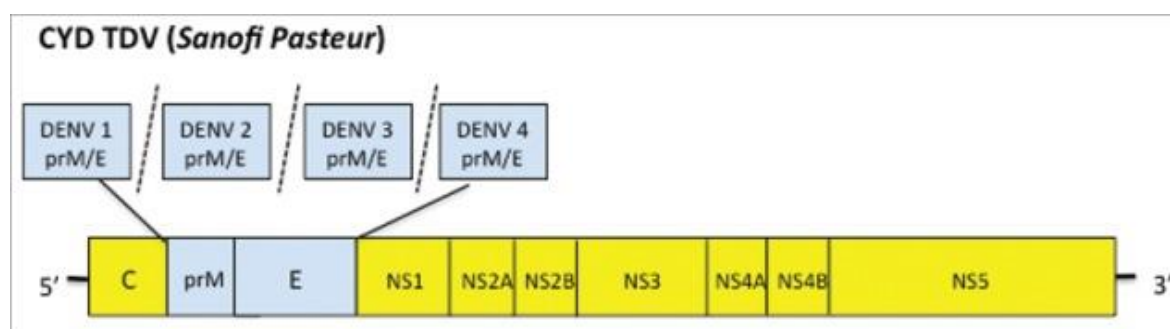


Figura 12. Dengvaxia® (CYD-TDV) da Sanofi Pasteur: vacina recombinante tetravalente de DENV. Dengvaxia® (CYD-TDV), vacina de estrutura quimérica tetravalente construída utilizando recombinação de genes da prM e proteína E dos quatro sorotipos de DENV, inseridos em sequências dos vírus vacinais (17D) de febre amarela. Fonte: TORRESI; EBERT e PELLEGRINI, 2017 (adaptado).

Estudos pré-clínicos da vacina demonstraram a estabilidade genética e fenotípica dos vírus vacinais recombinados, com a eficácia da CYD-TVD testada em um estudo de prova de conceito de fase II na Tailândia (CYD-23), além dois estudos de fase III no sul da Ásia e América Latina (CYD-15 e CYD-17), em um esquema de três doses (0, 6 e 12 meses) (TORRESI; EBERT; PELLEGRINI, 2017b). Embora tenha sido rapidamente submetida às fases clínicas e disponibilizada à população, os resultados não foram promissores, com variações na eficácia da vacina: foi demonstrado para CYD-23 que, dentre os pacientes confirmados e sintomáticos para dengue, uma eficácia entre 3,5% para DENV-2 até 90% para DENV-4; CYD-15, um estudo com crianças, teve uma eficácia média de 56,5%; e CYD-17, também em crianças, eficácia de 60,8% (SABCHAREON et al., 2012; CAPEDEING et al., 2014; VILLAR et al., 2015).

Embora essas variações coexistissem dentre os diferentes estudos de fase clínica, a eficácia na proteção de dengue grave, nos casos de febre hemorrágica e na redução da hospitalização nos primeiros anos se mostraram mais robustos, fatores que também favoreceram a liberação e comercialização da vacina no final de 2015 (TORRESI; EBERT; PELLEGRINI, 2017a; WHO, 2017). Contudo, resultados de até três anos de acompanhamento pós-vacinal demonstraram uma maior probabilidade de desenvolvimento de casos de dengue grave em indivíduos na exposição ao vírus após vacinação, estimando cerca de cinco casos de hospitalização e dois casos da forma grave da doença a mais a cada 1000 indivíduos vacinados. A OMS recomenda a vacinação somente em casos potencialmente graves, especialmente em indivíduos que já foram expostos ao vírus e/ou após uma avaliação clínico-epidemiológica prévia criteriosa (WHO, 2017).

Paralelamente, outras vacinas como a DENVax, desenvolvida pela Takeda Vaccines Inc., também se encontram em desenvolvimento e avaliações em estudos clínicos. Esta especificamente consiste em uma linhagem atenuada de DENV-2 (TDV-2), construída quimERICAMENTE contendo os genes da prM e proteína E de DENV-1, 3 e 4 no genoma de DENV-2 (OSORIO et al., 2015; SIRIVICHAYAKUL et al., 2016). Outro exemplo é o desenvolvimento e fabricação de uma vacina tetravalente liofilizada de vírus atenuado contendo os quatro sorotipos de DENV, desenvolvida pelo instituto Butantan, no Brasil (PRECIOSO et al., 2015).

Os métodos de prevenção para Zika, além das estratégias de controle do vetor, consistem em evitar o contato sexual com indivíduos expostos à doença ou oriundos de regiões com surtos

confirmados, além de um controle mais rigoroso de transfusões sanguíneas, embora apenas um caso tenha sido relatado até o momento (MOTTA et al., 2016). Sabendo das diversas vacinas produzidas para outros flavivírus como YFV e TBE, da inexistência de fármacos aprovados para o tratamento de Zika, dengue e febre amarela, somado às reações cruzadas no diagnóstico sorológico, o desenvolvimento de fármacos antivirais se torna uma opção única e viável contra esses três vírus (TAKAHASHI et al., 2012; DOWD et al., 2016; SAIZ et al., 2017).

1.2 Estratégias no planejamento, triagem e desenvolvimento de fármacos antivirais

Até o presente momento, não há fármacos antivirais aprovados para nenhum flavivírus, sendo o tratamento baseado no alívio de sintomas clínicos. Uma importante estratégia se volta para o desenvolvimento de fármacos antivirais, empregando análises *in silico*, como análises de similaridade estrutural, modelos de relações quantitativas entre a estrutura química e atividade biológica (do inglês, *Quantitative Structure-Activity Relationship* – QSAR), acoplamento molecular (*molecular docking*), dinâmicas moleculares e outros métodos de aprendizado de máquina (do inglês, *machine learning techniques* – MLT). Estas análises, aliadas a testes *in vitro* e triagens de substâncias sintéticas, semi-sintéticas e de produtos naturais, bem como a tentativa de reposicionamento de fármacos, podem ser estratégias favoráveis ao desenvolvimento de novos fármacos (BARROWS et al., 2016; LEONEL et al., 2018).

A triagem de bibliotecas de substâncias, como do banco de fármacos disponíveis e aprovados pelo FDA (do inglês, *Food and Drug Administration*), já demonstraram substâncias com atividade antiviral, especialmente de compostos semi-peptídicos e outras substâncias sintetizadas contra DENV, embora não tenham, até o momento, ultrapassando as primeiras fases de estudos clínicos. Takagi et al. (2017), desenvolveu e sintetizou peptídeos cíclicos mimetizando substratos da protease de DENV-2, confirmando a atividade inibitória em ensaios enzimáticos, com o melhor candidato apresentando valores de IC_{50} de 0,95 μ M (TAKAGI et al., 2017; LEONEL et al., 2018).

Dois alvos apresentam perfis promissores no desenvolvimento desses fármacos, a protease NS3 e a polimerase dependente de RNA (RdRp) NS5, ambas essenciais no processo de multiplicação destes vírus. A NS5 é altamente conservada, com aproximadamente 67 a 82% de identidade dentre os quatro sorotipos de DENV e essencial para a replicação do genoma do vírus, ressaltando sua importância, uma vez que não se encontra em células do hospedeiro, tornando-a um potencial alvo para o desenho de inibidores específicos. Contudo, especialmente

o complexo da protease NS2B-NS3 (NS3pro) como alvo de estudo pode contribuir para novas opções terapêuticas contra a dengue, Zika e demais doenças causadas por flavivírus, uma vez que essa proteína não estrutural também apresenta uma sequência altamente conservada dentro dos membros do gênero *Flavivirus*, tornando esse complexo um alvo multi-viral (FRECER; MIERTUS, 2010; TAKAHASHI et al., 2012; TAKAGI et al., 2017; LEONEL et al., 2018).

Além disso, os multidomínios do complexo NS2B-NS3 também são atrativos no desenvolvimento de ligantes com atividade antiviral e potenciais fármacos. A existência da protease, helicase e RNA trifosfatase, capacidade de modulação imune interagindo com NS4 e interações com NS4B e a própria NS5, corroboram a versatilidade e possibilidade das opções e estratégias no desenvolvimento de inibidores (LUO; VASUDEVAN; LESCAR, 2015; TAY et al., 2015; ZOU et al., 2015; YUAN et al., 2017).

1.2.1 Inibidores da protease NS2B-NS3

A protease de DENV apresenta uma afinidade por substratos com sequências dibásicas ou polibásicas, o que torna necessário a presença de fragmentos alcalinos ou polares na estrutura de ligantes, como os inibidores. Contudo, essas características moleculares resultam em baixa permeabilidade passiva das moléculas por membranas celulares, justificando uma baixa eficácia de inibidores peptídicos, conjugados peptídicos e inibidores não-peptídicos em ensaios *in vitro* (LI et al., 2005).

Dentre essas classes de inibidores, os peptídicos se baseiam nos substratos enzimáticos (ligantes naturais da protease), tendo o local de clivagem (sítio ativo da enzima) como foco principal. Estudos comparativos com inibidores naturais e avaliação de perfis cinéticos demonstraram que inibidores com menor número de aa apresentavam maior potência. Neste âmbito, inibidores menores possuindo grupos funcionais eletrofilicos como aldeído e ácido borônico apresentaram melhores resultados de inibição na protease de DENV (CHANPRAPAPH et al., 2005; YIN et al., 2006).

Peptídeos de cadeias lineares, por sua vez, se mostram instáveis ao metabolismo do hospedeiro e com baixa biodisponibilidade, diferentes de peptídeos cíclicos, mais estáveis e com maior permeabilidade celular, como a aprotinina (**Figura 13**) um inibidor natural da tripsina pancreática bovina. Estas moléculas, encontradas em animais e plantas, são utilizadas como protótipo para análogos sintéticos, podendo conferir maior especificidade a outras

proteases, como a NS2B-NS3 (FRITZ; WUNDERER, 1983; HORTON; BOURNE; SMYTHE, 2002; GAO et al., 2010).

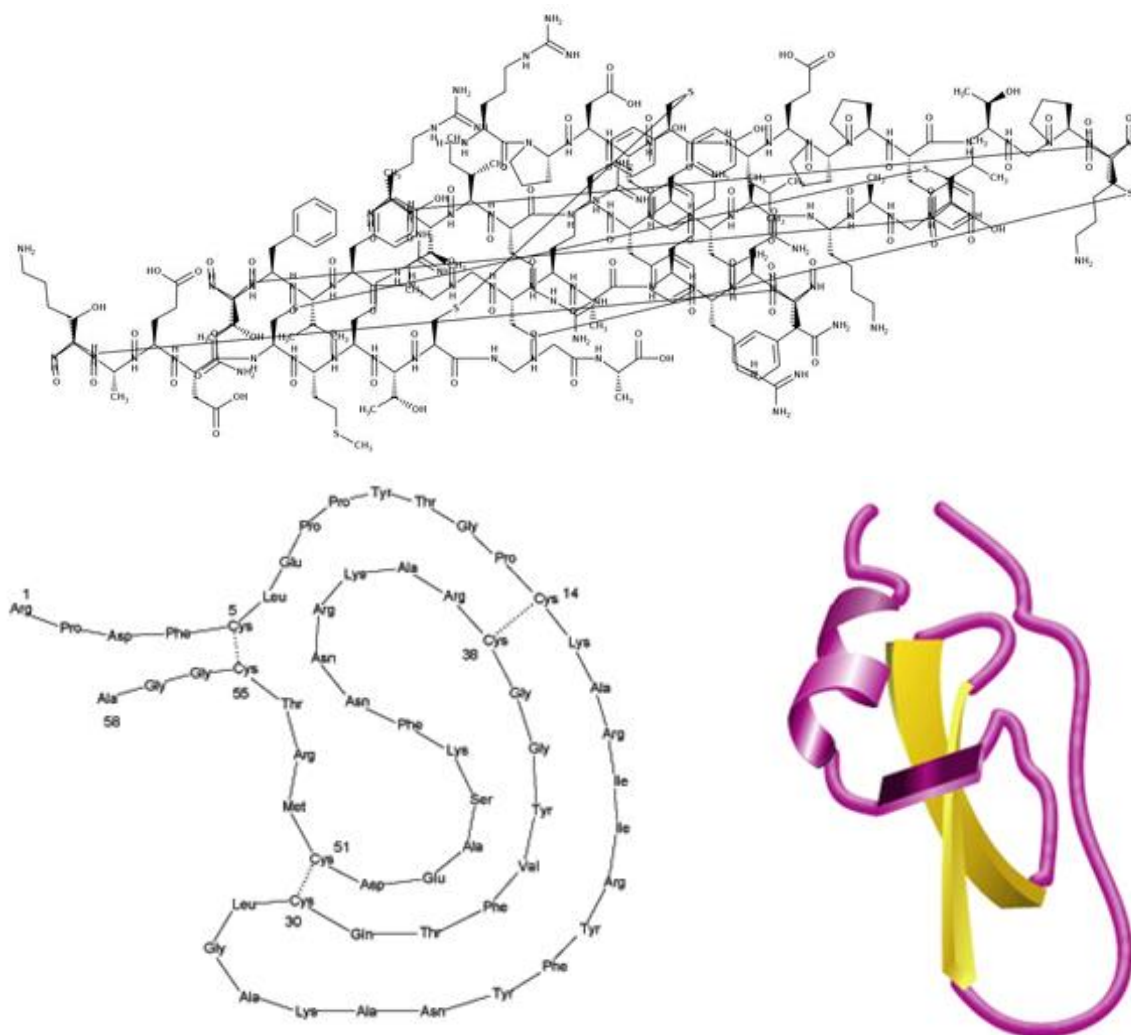


Figura 13. Representações estruturais da aprotinina. Acima: estrutura primária; abaixo à esquerda: secundária; abaixo à direita: terciária (amarelo e magenta). Fonte: Merck Index (FRITZ; WUNDERER, 1983; PUBCHEM, 2016) (*adaptado*)

Além desses inibidores, os conjugados peptídicos, ou seja, pequenas moléculas orgânicas ligadas a peptídeos (híbridos), podem interferir na atividade da protease. Contudo, esses híbridos não demonstraram a afinidade necessária devido às propriedades alcalinas dos substratos. Neste âmbito, inibidores não-peptídicos (peptideomiméticos), que apresentam frações eletrofílicas, podem ser opções mais viáveis, uma vez que podem mimetizar as funções peptídicas enquanto apresentam características farmacocinéticas mais viáveis (TIMIRI; SINHA; JAYAPRAKASH, 2016). Behnam et al. (2015), identificou diversos peptideomiméticos com atividade antiviral em DENV, confirmando a atividade antiviral em ensaios de redução de placas de lise, a exemplo da molécula na **Figura 14**.

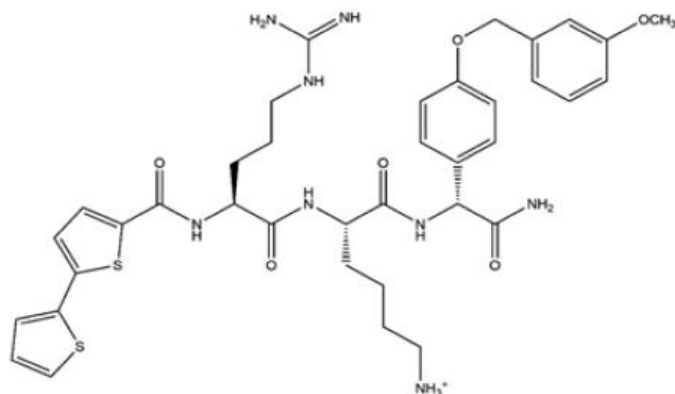


Figura 14. Molécula 104 (composto peptideomimético) com atividade anti-DENV. Fonte: Behnam et al. 2015.

Essas e outras moléculas bioativas podem ser identificadas por diversas estratégias, como triagens *in vitro* de quimiotecas e triagens virtuais de grandes bancos de dados. WU et al. (2015a), identificaram policresulen como um inibidor da protease de DENV, enquanto WU et al. (2015b) também observaram outras duas moléculas, denominadas “7” e “8”, com essa atividade inibitória na protease e em ensaios antivirais. Pambudi et al. (2013), também analisaram a substância SK-12, confirmando a inibição da protease e atividade antiviral em DENV (**Figura 15**).

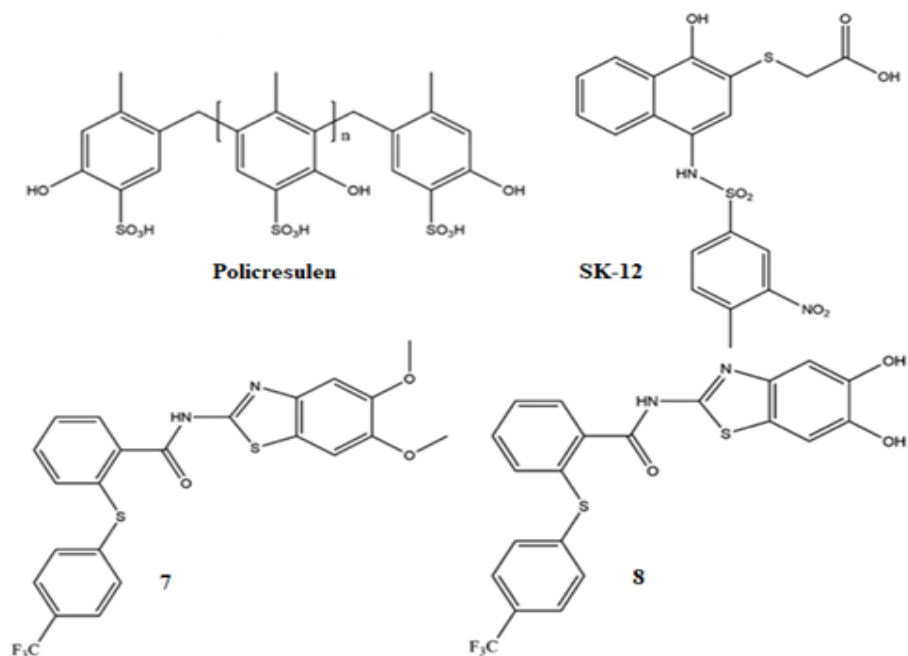


Figura 15. Exemplos de inibidores da protease de DENV descobertos. Polycresulen, SK-12, moléculas 7 e 8. Fonte: Tian et al. 2018 (*adaptado*).

1.2.2 Triagem virtual de moléculas bioativas

Dentre as diversas estratégias de planejamento de fármacos, a triagem virtual (do inglês, *virtual screening* – VS) é uma abordagem *in silico* que visa filtrar grandes bancos de dados de substâncias de modo a identificar ligantes com potencial atividade biológica precedendo, geralmente, ensaios *in vitro* confirmatórios. Baseada na modelagem molecular da estrutura de um alvo (do inglês, *structure based drug design* – SBDD) ou na estrutura do ligante (do inglês, *ligand based drug design* – LBDD), uma triagem virtual é capaz de avaliar dezenas a centenas de milhões de substâncias, o que implica em versatilidade e alto desempenho em uma etapa de desenvolvimento inicial de moléculas bioativas (GUIDO; ANDRICOPULO; OLIVA, 2010; FERREIRA; GLAUCIUS; ANDRICOPULO, 2011).

Métodos de SBDD se baseiam no arranjo de alvos biológicos, como proteínas, utilizando informações de uma macromolécula obtida por cristalografia de raios-X, espectroscopia por ressonância magnética nuclear (RMN) ou construção por homologia, para predição da interação (acoplamento molecular) de um ligante (molécula) no sítio de ligação. Os métodos de LBDD, por sua vez, avaliam comparativamente a conformação e propriedades de um ou diversos ligantes de um alvo molecular, buscando se obter informações acerca de suas características físico-químicas e/ou estruturais. Essas estratégias já se mostraram promissoras e capazes de identificar novas moléculas bioativas, como inibidores seletivos com atividade inibitória da protease de DENV, a partir da triagem virtual de milhões de substâncias baseado na estrutura do alvo (NS3), otimizando etapas iniciais na busca e desenvolvimento de inibidores de protease (VISWANATHAN et al., 2014; MIRZA et al., 2016, 2018).

Uma estratégia amplamente empregada é a caracterização e limitação do espaço químico utilizado nas triagens. Os estudos de Lipinski e Hopkins especificam que estes espaços podem ser descritos como espaços virtuais que descrevem moléculas com características e comportamentos específicos daquele espaço físico-químico. No contexto do planejamento de fármacos, funcionam como um filtro de moléculas que possuem as características físico-químicas e/ou estruturais que possibilitariam a seleção de melhores candidatos a fármaco. Nesse cenário, modelos empíricos capazes de criar espaços químicos foram desenvolvidos especialmente para caracterização do caráter *druglike* ou fármaco-similar. Contendo fármacos administrados oralmente com propriedades físico-químicas bem definidas, poderia se delimitar um perfil de parâmetros favoráveis ao desenvolvimento de um novo fármaco (LIPINSKI; HOPKINS, 2004; SCHNEIDER; FECHNER, 2005).

Lipinski e colaboradores (1997) propuseram a “Regra do 5 de Lipinski” (do inglês *Rule of Five* – Ro5), em que um *hit* teria grandes chances de ter uma boa absorção no trato gastrointestinal. A Ro5 estima que para um composto apresentar uma maior disponibilidade oral, ele deve apresentar massa molecular menor que 500 Da, não mais que cinco grupos doadores e 10 grupos aceptores de ligação de hidrogênio, além de log P (logaritmo do coeficiente de partição *n*-octanol/água) menor que cinco, sendo permitido violar uma das condições (LIPINSKI et al., 1997). Outro modelo, desenvolvido por Veber e colaboradores (2002), as “regras de Veber”, estimam que a polaridade e o grau de liberdade conformacional da molécula são importantes para a disponibilidade oral, em que o fármaco deve apresentar uma superfície polar menor do que 140 Å² e não mais que 10 ligações rotacionáveis (VEBER et al., 2002). Ambos os parâmetros auxiliam na otimização de triagens virtuais, filtrando e selecionando moléculas como potenciais inibidores e candidatos à fármacos, com maiores chances de sucesso em etapas avançadas do desenvolvimento de fármacos (LIPINSKI; HOPKINS, 2004).

1.2.3 Métodos de aprendizado de máquina (MLT): relação quantitativa entre a estrutura química e atividade biológica por holograma (HQSAR)

Os métodos ou técnicas de aprendizado de máquina (do inglês, *machine learning techniques* – MLT) compreendem algoritmos utilizados para classificação, regressão e agrupamento de tarefas ou funções para predição de uma propriedade ou atividade biológica, utilizando uma informação física, química ou ambas como descritor (XU; HAGLER, 2002; LO et al., 2018). Uma vantagem das MLT comparada a outras estratégias computacionais se baseia na capacidade de identificar padrões em grandes bancos de dados armazenados, os quais poderiam ser ignorados ou desconsiderados caso não ocorresse uma avaliação e caracterização não-linear, normalmente superando métodos de classificação e regressão clássicos na química medicinal (GERTRUDES et al., 2012; MALTAROLLO et al., 2015).

Dentre essas técnicas, os métodos de QSAR, embora sejam ferramentas importantes na descoberta e desenvolvimento de potenciais fármacos, apresentam algumas complexidades como o cálculo de descritores, geração de confôrmeros e alinhamentos de estruturas, bem como a dificuldade em se analisar grandes bancos de dados. No QSAR-3D, por exemplo, o alinhamento das estruturas geradas demanda tempo, além de ser necessário se alinhar mutuamente as moléculas no conjunto de dados, com base em alguma regra ou estratégia consistente (HERITAGE; LOWIS, 1999).

Nesse cenário, o método da relação quantitativa entre a estrutura química e atividade biológica por holograma (do inglês, *Hologram Quantitative Structure Activity Relationship* – HQSAR), permite agilizar análises estruturais e biológicas de uma substância, uma vez que estabelece a correlação entre a atividade biológica e um holograma construído de uma molécula, dispensando também o alinhamento das estruturas. Esse método é embasado no uso fragmentos moleculares como átomos (A), ligações (B), conectividade (C), átomos de hidrogênio (H), quiralidade (Q) e átomos doadores ou aceptores (DA), (**Tabela 7**), ambos gerados utilizando a estrutura bidimensional (2D) da substância (HERITAGE; LOWIS, 1999; KRONENBERGER et al., 2017).

Tabela 7. Descrição dos diferentes tipos de fragmentos utilizados no método de relação quantitativa entre a estrutura química e atividade biológica por holograma (HQSAR)

Distinção do fragmento	Definição
(A) Átomos	Tipos de átomos utilizados para construção do fragmento
(B) Ligações	Tipos de ligações: simples, duplas, triplas e aromáticas
(C) Conectividade	Hibridização dos átomos contidos nos fragmentos
(H) Hidrogênio	Presença e número de átomos de hidrogênio
(Q) Quiralidade	Presença de átomos quirais
(DA) Doadores e Aceptores	Presença de átomos aceptores ou doadores de ligações de hidrogênio

Fonte: GAD, 2008 (*adaptado*).

Os fragmentos moleculares gerados são traduzidos em um código binário (bins), sendo 0 o número que representa a ausência de determinado fragmento e 1 o valor que representa sua presença, de forma análoga à geração de *fingerprints* moleculares. (BROWN e MARTIN, 1996). A combinação desses fragmentos moleculares permite a criação dos hologramas, com sua extensão disposta de um tamanho mínimo e máximo (50 a 500 bins). Estes valores são organizados por números primos, com o objetivo de evitar que dois fragmentos parecidos ou iguais recebam uma mesma identificação (DODDAREDEY et al., 2004).

O HQSAR também é capaz de corroborar variações observadas nas atividades biológicas pela quantificação dos hologramas calculados, permitindo a construção de mapas de contribuição de cada fragmento das moléculas avaliadas (SALUM; ANDRICOPULO, 2009).

Os hologramas calculados consistem nas variáveis independentes do modelo matemático gerado. Na construção de um modelo, o método de regressão pelos mínimos quadrados parciais (do inglês, *Partial Least Squares*) juntamente com uma validação cruzada são realizados com o intuito de correlacionar os dados dos hologramas e os dados de atividade biológica (variável dependente), gerando os modelos matemáticos utilizados para posterior predição das atividades biológicas das moléculas (YE e DAWSON, 2009).

Adicionalmente, por não depender de alinhamentos das moléculas, a diversidade de substâncias e substituintes em um grupo teste não interfere no modelo (LU et al., 2012). Assim, a análise desses aspectos permite um planejamento de novas modificações moleculares, visando o desenvolvimento de moléculas bioativas mais potentes, possibilitando empregar modelos na predição das moléculas selecionadas nas etapas finais de um *virtual screening* (CRAMER, PATTERSON e BUNCE, 1988).

A utilização de um modelo de HQSAR como predição da atividade *in vitro* de determinadas substâncias já se mostrou eficaz na triagem e seleção de inibidores do tipo nardaprevir contra *Hepatitis C virus* (HCV), obtendo novos análogos com potencial atividade antiviral (ZHU et al., 2012). Essa estratégia também já se mostrou favorável na pesquisa e desenvolvimento de potenciais substâncias antitumorais, novos agentes terapêuticos para síndromes metabólicas e de compostos antimaláricos (EFFERTH et al., 2008; GARCIA; HONÓRIO, 2011b; LIU et al., 2016). Estes estudos, por sua vez, também demonstram a geração de modelos estatisticamente robustos (elevados valores de q^2) e preditivos (maiores valores de r^2 pred.).

Lima et al. (2018) desenvolveram modelos de QSAR 2D e 3D baseados em inibidores da trifosfatase de *Plasmodium falciparum*. Os melhores modelos nessa abordagem permitiram determinar um consenso para triagem virtual de moléculas bioativas, obtendo *hits* com atividade em diferentes estágios do parasito e estrutura química distinta de antimaláricos disponíveis, validando uma estratégia promissora no desenvolvimento de fármacos. Esta estratégia computacional ainda não foi avaliada em estudos envolvendo espécies do gênero *Flavivirus*, o que reforça a relevância deste estudo.

1.2.4 Acoplamento molecular (*molecular docking*)

A técnica de acoplamento molecular pode ser dividida basicamente em duas partes: o posicionamento correto do confôrmero de um ligante em um determinado sítio de ligação e o

sucesso de seu reconhecimento pontuado por uma função *score*. Essa abordagem simula um estado conformacional do complexo proteína-ligante, permitindo analisar e solucionar possíveis problemas que normalmente resultariam em baixa afinidade e/ou atividade de um ligante e/ou inibidor (WARREN et al., 2006).

A técnica de acoplamento molecular também já demonstrou amplamente a capacidade em prever atividades biológicas no planejamento e triagem de substâncias contra DENV, como policresulen, um inibidor recém determinado do complexo NS2B-NS3 (WU et al., 2015a), além de outros inibidores sintéticos também contra a protease deste vírus (PAMBUDI et al., 2013; WU et al., 2015b). Além destes, compostos sintéticos com atividade contra JEV, *Bovine viral diarrhea virus* (BVDV) e *Herpes simplex virus 2* (HSV-2) também já foram demonstrados utilizando esta técnica e confirmando seus resultados com ensaios *in vitro* (FANG et al., 2016; SORAIRE SANTACRUZ et al., 2017; ARUNKUMAR; RAJARAJAN, 2018).

1.2.5 Dinâmica molecular

A dinâmica molecular (DM) consiste em um sistema de simulações que permite modelar os movimentos de moléculas e macromoléculas a nível atômico, tornando observáveis alguns fenômenos difíceis de se observar experimentalmente. Diferentemente do acoplamento molecular onde apenas um estado conformacional do complexo proteína-ligante é simulado, a DM calcula as variações conformacionais deste complexo em função do tempo como uma resolução das equações baseadas na segunda lei de Newton: a força resultante que age sobre um corpo deve ser igual ao produto da massa do corpo por sua aceleração (DURRANT; MCCAMMON, 2011).

Além disso, as interações interatômicas são descritas por campos de força, unindo diversas funções dependentes das posições relativas de cada átomo e parâmetros numéricos que representam aproximações e correções do sistema. Tais simulações se mostram promissoras em áreas de estudo como a bioquímica, biologia molecular, bioinformática e química farmacêutica, permitindo observações a nível funcional de proteínas, ácidos nucleicos, membranas, receptores e outras estruturas (BOWERS et al., 2006; DURRANT; MCCAMMON, 2011).

Algumas funções biológicas ainda se encontram fora da alçada de uma DM por limitações computacionais e de tempo, bem como do interesse farmacêutico. Ainda assim, a capacidade de se observar as simulações de alterações estruturais proteicas, bem como de suas

interações com, por exemplo, um potencial candidato a fármaco, em escalas de mili e microssegundos resalta a importância do método. Essa simulação que mimetiza uma realidade biológica pode envolver centenas de milhares de átomos, representativos de macromoléculas em um ambiente que considera solvente, como água, além dos íons das moléculas e estruturas solvatadas (DUAN; KOLLMAN, 1998; BOWERS et al., 2006).

Desta forma, as simulações de DM são úteis no refinamento de resultados, por exemplo, de acoplamento molecular, permitindo estudar a estabilidade do complexo proteína-ligante e suas interações intermoleculares preditas. Essa estratégia já foi utilizada na busca de sítios de ligação adjacentes ao sítio ativo da integrase de HIV-1, um achado que anos depois auxiliaria no desenvolvimento de inibidores de integrase deste vírus (SCHAMES et al., 2004; DURRANT; MCCAMMON, 2011). Utilizando DM, também já foi demonstrada uma cavidade universal no sítio de adesão da neuraminidase de vírus influenza, auxiliando, por exemplo, no desenvolvimento de fármacos contra este vírus (AMARO et al., 2011).

1.2.6 Predição e mapeamento de sítios de ligação em proteínas alvo

Diversas técnicas são utilizadas para estudar os possíveis sítios de ligação de proteínas de interesse no planejamento de fármacos e para identificar regiões favoráveis à interação com grupos funcionais específicos. Dentre elas, os *hot spots* são definidos como pequenas regiões ou pontos em sítios de ligação que apresentam alguma relevância energética em interações moleculares. A capacidade destas regiões se ligarem a uma variedade de moléculas-sonda pode ser predita por servidores como FTMap e FTSite, desenvolvidos para estudos de sítios de ligação (KOZAKOV et al., 2015). Ambos podem ser considerados análogos computacionais de triagens experimentais de macromoléculas (cristalografia de raios-X e RMN) que identificam e caracterizam sítios de ligação e quais suas possíveis ligações. Ambos realizam o mapeamento de macromoléculas simulando a interação de pequenas moléculas orgânicas (moléculas-sonda) com a proteína, DNA ou RNA alvo. Essas sondas variam em forma, tamanho e características físico-químicas, como descrito na **Tabela 8** (KOZAKOV et al., 2015).

Tabela 8. Moléculas-sonda utilizadas pelos servidores FTMap e FTSite e suas respectivas características físico-químicas.

Molécula-sonda	Características físico-químicas
acetamida (ACD)	polar, doadora e aceptora de ligação de hidrogênio
acetonitrila (ACN)	polar, aceptora de ligação de hidrogênio
acetona (ACT)	polar, aceptora de ligação de hidrogênio
acetaldeído (ADY)	polar, aceptora de ligação de hidrogênio
metanamina (AMN)	polar, positiva, doadora e aceptora de ligação de hidrogênio
benzaldeído (BDY)	polar, aromática, aceptora de ligação de hidrogênio
benzeno (BEN)	hidrofóbica, aromática
tert-butanol (BUT)	hidrofóbica, aromática
ciclohexano (CHX)	polar, doadora e aceptora de ligação de hidrogênio
<i>N,N</i> -dimetilformamida (DFO)	polar, aceptora de ligação de hidrogênio
dimetil éter (DME)	polar, aceptora de ligação de hidrogênio
etanol (EOL)	polar, doadora e aceptora de ligação de hidrogênio
etano (ETH)	hidrofóbica
fenol (PHN)	polar, aromática, doadora e aceptora de ligação de hidrogênio
isopropanol (THS)	polar, aceptora de ligação de hidrogênio
ureia (URE)	polar, positiva, doadora e aceptora de ligação de hidrogênio

Fonte: KOZAKOV et al., 2015 (*adaptado*).

Expressões energéticas como a interação eletrostática, energia de dessolvatação, atração e repulsão de interações de van der Waals, além do potencial de interação pareado baseado na estrutura das macromoléculas são utilizados como amostragem no mapeamento do alvo. A partir dessa amostragem, o FTMap busca as conformações de menor energia dessas sondas, agrupando-as de acordo com a energia média entre elas. Baseado nesses grupos (*clusters*) é possível definir e caracterizar as interações entre as sondas e as regiões da proteína onde estão situadas. No FTSite, por sua vez, regiões para possíveis ligações de inibidores são pesquisadas de acordo com a localização das sondas em uma caixa tridimensional ao redor da proteína, também agrupadas de acordo com a energia média do grupo, sendo as regiões com maior quantidade de sobreposições de sondas consideradas como os sítios de ligação (MOTTIN et al., 2017; PANTALEÃO et al., 2018). Desse modo, o FTSite identifica os sítios de ligação enquanto FTMap caracteriza tais sítios, fornecendo informações relevantes ao planejamento de potenciais inibidores e candidatos a fármacos.

Uma outra estratégia para identificação de regiões favoráveis de interação entre ligantes e alvos moleculares é o cálculo dos campos de interação molecular (do inglês, *Molecular interaction fields* – MIF). Esta ferramenta computacional permite a avaliação da complementaridade e capacidade de interação molecular entre um ligante e receptor, baseado

nas propriedades estéricas e eletrostáticas dos átomos (sondas) das moléculas envolvidas (ROCHA; FREITAS; MONTANARI, 2010). Esse tipo de cálculo é feito nos programas AutoGrid, implementado no pacote do AutoDockTools (MORRIS et al., 2009) e GRID (KASTENHOLZ et al., 2000).

No AutoGrid, as sondas são movidas por todos os pontos de uma caixa tridimensional construída em volta de uma região de estudo da macromolécula, como o sítio ativo da protease. Cada ponto de interação da sonda com os aa é calculado, sendo MIF resultante de pontos favoráveis em uma determinada energia, representando a força de atração entre o átomo (sonda) e a região delimitada na macromolécula. Os resultados do GRID distinguem regiões por características de interações intermoleculares específicas de uma área delimitada no alvo, permitindo, por exemplo, a construção de modelos farmacofóricos (KASTENHOLZ et al., 2000; GOODFORD, 2006).

1.2.7 Predições de biodisponibilidade (farmacocinética) e toxicidade

Os estudos *in silico*, além de buscar potenciais moléculas bioativas, também são empregados no intuito de eliminar substâncias com potencial toxicidade e perfil farmacocinético inadequado em etapas mais avançadas do desenvolvimento de fármacos. Assim, não só moléculas com pouca ou nenhuma atividade durante os ensaios *in vitro* e *in vivo* são dispensadas, mas também aquelas que poderiam apresentar perfis farmacocinéticos inadequados ou toxicidade elevada (PATRICK, 2005; PEREIRA, 2007; RUIZ-GARCIA et al., 2008; WERMUTH, 2008).

Neste âmbito, modelos computacionais foram desenvolvidos para predição de parâmetros envolvidos na farmacocinética ou toxicidade de moléculas bioativas (*hits*). Alguns, como os já citados QSAR e outras técnicas de aprendizado de máquina (MLT), envolvem modelos matemáticos que correlacionam a estrutura química e propriedades físico-químicas e moleculares calculadas com um parâmetro estudado, como a atividade biológica, toxicidade ou perfil farmacocinético de uma molécula (BARREIRO, FRAGA, 2015; PATRICK, 2005; WERMUTH, 2008; YANG et al., 2018). Dentre os programas e servidores utilizados para predição das propriedades das moléculas referentes às características farmacocinéticas (biodisponibilidade) e toxicológicas, são exemplos os servidores admetSAR (YANG et al., 2018; <http://lmmd.ecust.edu.cn/admetSar2/>), ProTox-II (BANERJEE et al., 2018; http://tox.charite.de/prottox_II/) e Pred-hERG 4.2 (BRAGA, 2014; <http://predherg.labmol.com.br/>).

O servidor admetSAR permite uma predição qualitativa de parâmetros farmacocinéticos como a absorção intestinal e capacidade de atravessar a barreira hematoencefálica. No ProTox-II, por sua vez, a predição dos potenciais mutagênicos, hepatotóxicos, carcinogênicos, imunotóxicos, citotóxicos e o valor teórico de dose letal mediana (DL_{50}) podem ser calculados e avaliados quanto a sua acurácia de predição. Por último, Pred-hERG 4.2 avalia o potencial de bloqueio dos canais para potássio em células cardíacas, também demonstrando uma acurácia de predição (BRAGA, 2014; BANERJEE et al., 2018; YANG et al., 2018).

2 JUSTIFICATIVA

As arboviroses, com destaque às infecções por DENV, YFV e ZIKV são ainda um grande problema de Saúde Pública mundial, com aproximadamente três bilhões de pessoas vivendo em regiões de risco para a dengue e mais de 100 milhões de casos dessa doença anualmente. Adicionalmente, ZIKV apresenta manifestações clínicas de importância, como a síndrome de Guillain-Barré e síndrome congênitas como má-formações do SNC em fetos, somando cerca de 3.000 casos relatados em 2016 no Brasil. A febre amarela, ainda que prevenível por vacinação desde a década de 1940, apresentou surtos desde 2017, somando quase 1.000 óbitos confirmados até agora. No país, nos anos de 2015 e 2016 foram estimados mais de 1,5 milhões de casos de dengue e 200.000 casos de Zika, enquanto cerca de 1.439.471 casos prováveis de dengue e 9.813 de Zika foram relatados em 2019.

Até o momento, não há fármacos antivirais aprovados para o uso em caso de infecções por esses três vírus. Não existem vacinas específicas para ZIKV e, embora uma vacina contra dengue tenha sido aprovada, além de outras em fases de testes clínicos, sua taxa de proteção ainda permanece baixa. Ainda que exista também uma vacina para febre amarela, os surtos em 2017 e 2018 ressaltam a necessidade de uma opção terapêutica disponível. Além disso, a dificuldade no diagnóstico sorológico destes vírus, em razão de reações cruzadas pela semelhança entre as espécies, deixa os testes confirmatórios às técnicas de biologia molecular, difíceis de ser aplicadas em ampla escala. Sendo assim, a busca por novas moléculas bioativas com potente atividade contra estes vírus denota relevância e urgência.

Devido a sua atividade enzimática indispensável à multiplicação destes vírus, a proteína NS3 se apresenta como alvo favorável para o desenvolvimento de novos fármacos antivirais, sendo a NS3pro excepcionalmente importante como alvo multiviral ou panflavivírus, uma vez que apresenta característica multifuncional (protease, helicase e trifosfatase), essencial em diferentes etapas no processo de multiplicação viral. Altamente conservada dentre diferentes espécies de *Flavivirus*, a NS3 apresenta capacidade de modulação imune (NS4), interage com NS4B e NS5, além de apresentar multidomínios diversos do complexo NS2B-NS3 e, conseqüentemente, mais opções de sítios alvo. Portanto, a presente proposta objetiva, empregando estratégias *in silico* e *in vitro* de triagem e planejamento de moléculas bioativas, descobrir e/ou desenvolver novas substâncias antivirais contra DENV, YFV e ZIKV, estabelecendo como alvo molecular a proteína NS3pro.

3 OBJETIVOS

3.1 Objetivo Geral

Identificar substâncias com potencial atividade antiviral, planejadas para atuarem contra a protease viral NS3 de DENV, YFV e ZIKV, utilizando estratégias de planejamento de moléculas bioativas *in silico* e validação de sua atividade *in vitro*.

3.2 Objetivos Específicos

- 3.2.1 Construir modelos quantitativos da relação quantitativa entre a estrutura química e atividade biológica por holograma (HQSAR) e similaridade estrutural de compostos peptídeomiméticos com atividade anti-DENV pré-selecionados da literatura;
- 3.2.2 Realizar o mapeamento e predição dos sítios de ligação nas proteases de DENV e ZIKV;
- 3.2.3 Construir um modelo farmacofórico utilizando as informações dos modelos de HQSAR, mapeamento das estruturas, acoplamento molecular (*molecular docking*) e dinâmica molecular;
- 3.2.4 Selecionar potenciais inibidores de multiplicação de flavivírus por triagem virtual, utilizando filtros de propriedades e farmacóforo, em quimiotecas como ZINC, FDA, NuBBE e de substâncias sintetizadas no Laboratório de Química Farmacêutica Medicinal e Modelagem Molecular da UFMG;
- 3.2.5 Selecionar potenciais inibidores de protease dos vírus (NS3pro) utilizando as técnicas de *docking* e dinâmica molecular;
- 3.2.6 Avaliar propriedades de biodisponibilidade e toxicidade das substâncias avaliadas *in silico* para desconsiderar aquelas indesejáveis na aquisição;
- 3.2.7 Determinar a concentração citotóxica 50% (CC₅₀) e concentração efetiva 50% (CE₅₀) para 50% das células, além dos respectivos índices de seletividade (IS) das substâncias selecionadas na triagem virtual pelo teste colorimétrico de MTT e ensaios de redução do número de placas de lise;
- 3.2.8 Avaliar a atividade virucida e determinar a concentração virucida (CV₅₀) das substâncias que apresentarem atividade antiviral *in vitro*;
- 3.2.9 Avaliar o efeito de pré-tratamento das substâncias no ciclo de multiplicação dos vírus.

4. MATERIAL E MÉTODOS

4.1 Análises *in silico*

4.1.1 Conjunto de dados para análise: seleção de moléculas bioativas

Inicialmente, foi realizada uma busca na literatura por compostos com atividade de inibição enzimática e/ou antiviral contra DENV e ZIKV. Dentre os critérios, foram incluídos os termos “NS3”, “*antiviral*” e “*synthetic compounds*”, delimitando um período de 2013 a 2018. Foram encontrados cerca de 70 trabalhos, priorizando aqueles que continham grupos de pelo menos 20 moléculas, a fim de se obter um número mínimo experimental de substâncias que abrangesse maior quantidade de estruturas com grupos funcionais e/ou substituintes distintos. Os modelos estruturais tridimensionais das moléculas selecionadas foram construídos utilizando o programa Discovery Studio (BIOVIA, USA, 2016).

Ao final da construção os compostos foram submetidos ao pacote computacional KNIME® (BERTHOLD et al. 2007), para criação do banco de conformêros necessários. Foi gerado o conformêro de menor energia de cada composto dentre um número máximo de conformações reunidas e ordenadas de acordo com a energia, ou seja, a unidade conformacional de maior estabilidade. Todas as estruturas tiveram seus estados de ionização corrigidos de acordo com pH fisiológico, também utilizando pacote computacional KNIME®.

Finalmente, foi realizada uma separação das moléculas em grupos ou conjuntos treinamento e teste (80% e 20% das moléculas, respectivamente), para elaboração do modelo de HQSAR. Essa separação obedeceu dois critérios: (i) exclusão de moléculas inativas ou não validadas experimentalmente dos grupos treinamento e teste; e (ii) exclusão da(s) molécula(s) mais ativa(s) da(s) série(s) do grupo teste.

4.1.2 Relação quantitativa entre a estrutura química e atividade biológica por holograma (HQSAR) de compostos peptidomiméticos

O princípio do método de HQSAR foi embasado na combinação dos descritores: átomos (A), ligações (B), conectividade (C), átomos de hidrogênio (H), quiralidade (Q) e átomos doadores ou aceptores (DA) para geração dos modelos, utilizando a estrutura 2D das moléculas. O tamanho do fragmento foi fixado em relação ao número de átomos (4 a 7), além de um tamanho mínimo e máximo de extensão dos hologramas (50 a 500 *bins*), organizados em números primos.

Os quatro melhores modelos (maiores valores de q^2) obtidos nessa fase de otimização foram submetidos a variações dos tamanhos dos fragmentos, em intervalos de 2, 3 e 4 átomos, mantendo sempre os descritores de cada modelo. Dessa forma, foram selecionados 20 modelos (cinco melhores de cada), obtendo uma otimização dos resultados na primeira fase, buscando construir um modelo estatisticamente mais representativo, que correlacionasse o holograma com a atividade biológica experimental (variável dependente) das moléculas.

Os 20 modelos selecionados foram submetidos a uma validação externa no pacote computacional KNIME®, em que diversas métricas podem ser avaliadas na busca da capacidade preditiva mais robusta de um modelo, por exemplo, maiores médias nos testes de q^2 e r^2 aplicados, bem como menores média de erro absoluto (do inglês, *Mean Absolute Error* – MAE) e erro quadrático médio (do inglês, *Root Mean Squarred Error* – RMSE). MAE corresponde a uma pontuação linear, distribuindo igualmente as diferenças individuais de um grupo analisado. RMSE, por sua vez, compara a média da magnitude do erro em um grupo (CHIRICO; GRAMATICA, 2011; CHAI; DRAXLER, 2014; GRAMATICA; SANGION, 2016).

Finalmente, uma regressão pelos mínimos quadrados parciais e uma validação cruzada também foram empregados para seleção de um modelo ótimo de HQSAR. Devidamente validado pelos testes de validações externas e cruzadas, foi possível obter mapas de contribuições dos fragmentos moleculares, sendo esses positivos ou negativos para atividade biológica das moléculas utilizadas no modelo, o que pode indicar regiões fundamentais das moléculas avaliadas na inibição da protease de DENV.

4.1.3 Seleção e preparo das estruturas da NS2B-NS3

A fim de selecionar a estrutura da protease com qualidade mais adequada aos estudos computacionais, foi realizada uma busca no banco de dados de proteínas (do inglês, *Protein Data Bank* – PDB; (BERMAN et al., 2000), um repositório de arquivos contendo a estrutura tridimensional de proteínas, obtidas por métodos experimentais como difração de raios-X por cristalografia, espectroscopia de RMN e microscopia crio-eletrônica.

Foram feitas buscas utilizando os seguintes termos como palavras-chave: “NS3”, “DENV”, “ZIKV”, “YFV” e “protease”. Quatro critérios para eleição da estrutura mais apropriada para o estudo, foram estabelecidos, descritos do mais relevante para o menos relevante: (i) valor da resolução <2.0 Å, indicando estruturas de alta resolução, ou seja, maior

confiabilidade da posição dos átomos dos aa da proteína alvo; (ii) presença de inibidores interagindo com a proteína, a fim de compará-los com ligantes planejados, bem como de obter um estado conformacional apropriado para a ligação; (iii) ausência de mutações, buscando sequências mais similares às daquelas dos modelos experimentais; e (iv) cadeias completas, evitando a necessidade de construção por homologia. O programa PyMOL 2.0.4 (Schrödinger, Inc., DeLano Scientific, San Carlos, CA, USA, version 0.99c, 2017) foi empregado para as análises visuais e confecção das figuras apresentadas como alguns dos resultados deste trabalho.

Além das moléculas selecionadas, as estruturas das proteases de DENV e ZIKV também foram selecionadas no PDB e visualizadas no programa PyMOL. Os arquivos das estruturas selecionadas foram obtidos no PDB, sob os códigos de identificação **3UII** (DENV-3) e **5YOF** (ZIKV) (NOBLE et al., 2012a; PHOO et al., 2016). Ambos foram importados ao programa e submetidos a coloração em suas frações NS3 C' e N' terminais, NS2B, alça *linker*, tríade catalítica, ligante e helicase (quando houvesse). Não haviam estruturas cristalográficas da protease de YFV disponíveis no PDB até o início do projeto e, considerando a similaridade e identidade entre as proteases, após as análises estruturais conduzidas, apenas a protease de DENV-3 foi selecionada para os estudos posteriores de acoplamento molecular, permitindo sua comparação com dados obtidos dos compostos peptidomiméticos de Behnam et al. (2015).

4.1.4 Predição dos sítios de ligação da protease e campos de interação molecular (MIF)

As estruturas selecionadas da protease de DENV e ZIKV, no formato PDB, foram submetidas aos servidores FTSite e FTMap (“The FTMap family of web servers for determining and characterizing ligand-binding hot spots of proteins. - PubMed - NCBI”,), buscando prever cavidades da protease susceptíveis a atuarem como sítios de ligação e quais tipos de ligações seriam propícios a ocorrerem nestes sítios, respectivamente. O servidor FTSite resultou nos potenciais sítios de ligação e aa de interesse da protease, enquanto o servidor FTMap resultou no agrupamento de moléculas-sonda que interagem com os aa da protease, divididos por cada sítio de ligação.

Durante a caracterização do perfil de afinidade dessas regiões, foi utilizado o programa Discovery Studio (BIOVIA, USA, 2016), em que foram observadas interações ocorrendo entre cada grupo funcional das sondas moleculares e aa dos sítios de ligação da enzima. Dessa maneira, foi possível estimar quais aa fazem mais ligações, ou seja, interagem mais com alguma molécula-sonda que esteja próxima, bem como quais sondas são mais importantes estarem presentes em um inibidor e em qual cavidade da proteína poderiam se ligar, além dos tipos de

ligação que ocorreriam. Na análise da interação destas sondas com os aa da protease, foram identificados quais aa interagem com determinada molécula-sonda e a quantidade de ligações presentes pela sonda, divididas em acceptor de ligação de hidrogênio (HBA), doador de ligação de hidrogênio (HBD), interação eletrostática (iônica) e interações de van der Waals (hidrofóbicas).

O programa AutoGrid (AutoDockTools) foi utilizado para caracterização dos campos de interação moleculares. Previamente, a proteína foi preparada de acordo com o tutorial “*Using AutoDock 4 and AutoDock Vina with AutoDockTools: A Tutorial*” (HUEY ET AL., 2012), seguindo os seguintes passos: i) foram retiradas as moléculas de água presentes na estrutura; ii) foram deletados os dados de dupla ocupância, ou seja, duas probabilidades de posição de átomos, removendo aqueles que apresentavam menor valor; iii) foram adicionados átomos de hidrogênios; iv) foram calculadas as cargas atômicas parciais computadas para todo o ligante; v) o arquivo preparado foi salvo utilizando a extensão .pdbqt.

A extensão da caixa tridimensional foi definida em 20x20x20 Å, com espaçamento padrão de 1 Å entre os pontos do *grid*, sendo o centróide definido como o centro de massa do ligante presente na estrutura. Nos cálculos, foram selecionadas as seguintes sondas: hidrofóbica, representada pelo carbono alifático (C); doadora de ligação de hidrogênio, representada por um hidrogênio doador (HD); acceptora de ligação de hidrogênio, representada por um oxigênio acceptor (OA); e eletrostática (E). Os resultados também foram visualizados no programa PyMOL.

4.1.5 Acoplamento molecular (*molecular docking*): ligantes e peptideomiméticos

Os estudos de interação entre os ligantes selecionados e o alvo molecular (NS3pro) foram realizados utilizando o método de acoplamento molecular (*molecular docking*). Somente o modelo da proteína de DENV (PDB: **3U1I**) foi selecionado para estes estudos e o sítio de ligação foi definido ao redor do ligante co-cristalizado da protease, utilizando os programas Surflex Docking e GOLD 5.1 (JAIN, 2003; VERDONK et al., 2003).

Os parâmetros intrínsecos dos programas, como função de pontuação, flexibilidade de aa e átomos, além do número de poses foram testados e avaliados previamente de forma a melhor reproduzir o modo de ligação experimental de ligantes conhecidos (VALLONE A, *et al.* 2018). Nesse caso, foi feita a reprodução do acoplamento molecular da molécula mais ativa (62) do trabalho de Behnam et al. (2015), fornecendo dois resultados: uma pontuação (baseada

no algoritmo do programa) referente à afinidade com a proteína estudada (ranqueamento das substâncias) e o modo de ligação e suas interações intermoleculares (avaliação de interações necessárias das substâncias para inibir a enzima em estudo).

Ademais, interações de grupos funcionais com aa da protease de inibidores conhecidos e utilizados neste trabalho foram comparados aos resultados dos mapas dos sítios de ligação, além dos mapas de contribuição química do HQSAR. Assim, o conjunto de dados permitiu avaliar, junto ao *docking*, informações relevantes de cada estratégia na construção do modelo farmacofórico.

4.1.6 Construção do modelo farmacofórico e triagem virtual (VS)

Utilizando as informações dos modelos de HQSAR construídos e validados, além do mapeamento das estruturas dos alvos, acoplamento molecular (*molecular docking*) e simulações de dinâmica molecular, um modelo farmacofórico foi construído no programa UNITY, implementado na plataforma computacional SYBYL-X 2.1. O modelo representa um conjunto de características que definem um ou mais grupos funcionais ou subunidades estruturais, necessários para o melhor reconhecimento molecular de um ligante a um determinado alvo, sugerindo maior afinidade (YANG, 2010). Esse compilado de variáveis e informações nos resultados desse trabalho (HQSAR, FTSite, FTMap, AutoGrid, *docking* e DM), possibilitou selecionar aa de interesse e definir centroides de interação molecular no modelo, permitindo filtrar moléculas de interesse, como potenciais inibidores da protease de DENV utilizada (LEONEL et al., 2018).

As bibliotecas empregadas na triagem virtual foram: i) quimioteca de substâncias sintetizadas pelo Laboratório de Química Farmacêutica e Medicinal (QF) da Faculdade de Farmácia (FAFAR) da UFMG, que contempla mais de 1.200 substâncias catalogadas; ii) um subconjunto do banco de dados ZINC (<https://zinc.docking.org/>), com 7.586.000 moléculas, previamente selecionadas baseado em propriedades “*druglike*” e/ou disponibilidade de aquisição via MolPort (<https://www.molport.com/shop/index>); iii) NuBBE, Núcleo de Bioensaios, Biossíntese e Ecofisiologia da UNESP, contendo 2.712 produtos naturais; iv) DrugBank, banco contendo 9.821 fármacos aprovados pelo FDA ou em fase experimental e disponíveis comercialmente. O número de candidatos corresponde ao total de moléculas disponíveis em cada banco de dados até a execução do protocolo de triagem virtual, em março de 2019.

Na etapa inicial do protocolo de triagem, um filtro foi estabelecido baseado em modificações na regra dos cinco de Lipinski (LIPINSKI, 1997) e nas regras de Veber (VEBER, 2002), além de alguns critérios específicos (centros quirais, carga total, solubilidade e elementos permitidos) em razão dos métodos aplicados no estudo (HQSAR, *docking* e dinâmica molecular), buscando selecionar candidatos obedecendo a certos parâmetros de afinidade de fármaco, a fim de obter moléculas que apresentassem, de uma maneira geral, maior biodisponibilidade. Os parâmetros foram listados na **Tabela 9**.

Tabela 9. Parâmetros definidos para o modelo de triagem virtual

Parâmetros	Condições
Elementos permitidos	C, O, N, H, P, S, F, Cl, Br, I
Doador de hidrogênio (HBD)	3 a 5
Aceptor de hidrogênio (HBA)	0 a 10
Log P	-2,0 a 5,5
Massa molecular (MM)	150 a 550
Carga total	0 a +2
Centro quiral	≤ 2
Ligações rotacionáveis	≤ 10
Área de superfície polar (PSA)	0 a 300
Solubilidade	Baixa (limitado a DMSO)

Algumas dessas variáveis foram modificadas buscando aprimorar o filtro de triagem inicial. Uma margem de tolerância de 10% foi ajustada para os parâmetros de massa molecular (550 Da) e de log P (5,5), comparado com a regra de Lipinski, de até 500 Da e 5,0, respectivamente. Respeitando ao menos 3 dos 4 parâmetros da regra, a margem foi definida em razão das características de tamanho e peso molecular dos protótipos – peptidomiméticos (LIPINSKI, 1997). Igualmente, esse ajuste considerou o número de doadores e aceptores de hidrogênio limitados pelo farmacóforo, buscando evitar excluir potenciais candidatos.

Já comparado às regras de Veber, pouco mais do dobro da superfície polar foi considerado (300 Å²), em razão do alto caráter polar dos compostos peptidomiméticos utilizados (VEBER, 2002). Os átomos permitidos, por sua vez, foram limitados a carbono, oxigênio, nitrogênio, hidrogênio, fósforo, enxofre, flúor, cloro, bromo e iodo, em razão do campo de força do ligante no *docking* e considerando maximizar a viabilidade sintética dos *hits* obtidos neste trabalho. Neste caso, o acoplamento realizado no programa Surflex Docking se

limita a ausência de metais no cálculo da função *score*, sendo necessário adicioná-los manualmente para as predições.

Dois centros quirais também foram delimitados, buscando reduzir a possibilidade de enantiômeros e facilitar uma eventual viabilidade de síntese de moléculas. A carga total (0 a +2), por sua vez, se baseia em uma maior disponibilidade de cargas positivas para interagir com átomos no alvo, permitindo, por exemplo, interações iônicas e ligações de hidrogênio. Contudo, o número não pode ser elevado o bastante, haja visto que o estado de ionização influencia diretamente na solubilidade e, conseqüentemente, na permeabilidade em membranas celulares (TEN BRINK; EXNER, 2009; H. BROOKS; C. GUIDA; G. DANIEL, 2011).

Por último, um filtro de solubilidade das moléculas foi delimitado em dimetilsulfóxido (DMSO), em razão da capacidade deste solvente em dissolver uma ampla gama de propriedades físico-químicas, apresentar baixa volatilidade, miscibilidade com água e outros solventes. Além disso, DMSO apresenta relativa baixa toxicidade em culturas celulares e possibilidade de congelamento abaixo de -196 °C em nitrogênio líquido, sendo opção universal de solvente mais favorável, especialmente à manutenção da estabilidade das moléculas em solução (BALAKIN, 2003; BALAKIN et al., 2004; WAYBRIGHT; BRITT; MCCLOUD, 2009).

Foi novamente gerado o confôrmero de menor energia das cerca de 7.600.000 moléculas e corrigido seu estado de ionização de acordo com pH fisiológico (ambos no pacote computacional KNIME®). Dessa forma, o protocolo consistiu em sete etapas apresentadas no esquema abaixo (**Figura 16**)

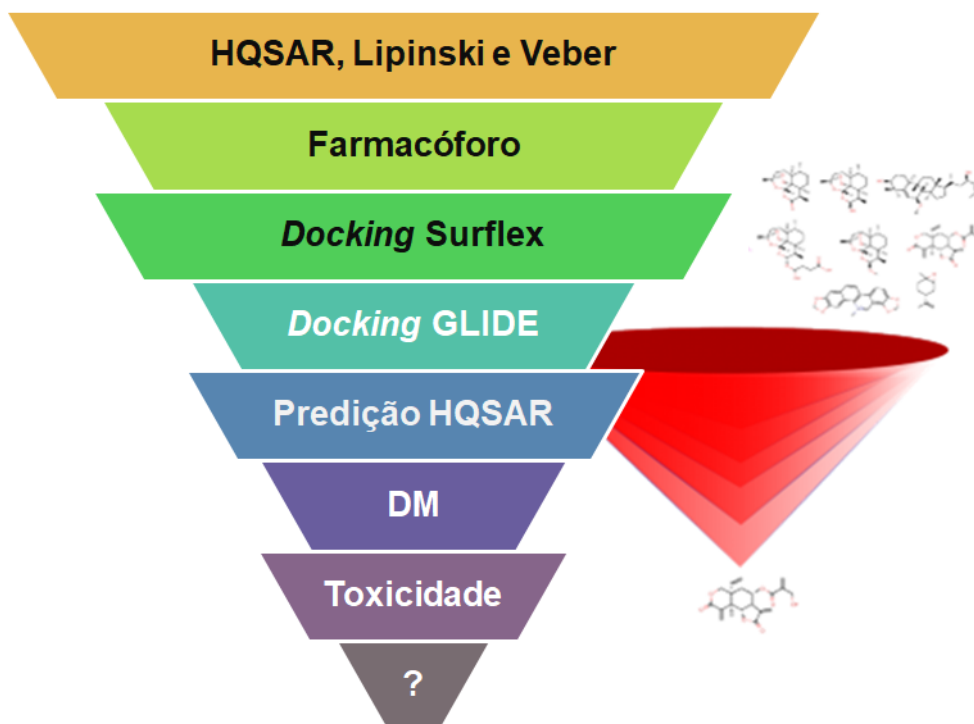


Figura 16. Esquema de etapas da triagem virtual. Paralelamente, os quatro bancos de dados foram filtrados nas condições dos parâmetros das propriedades moleculares de interesse (fragmentos arginina, lisina e modificações das regras de Lipinski e Veber), utilizando pacote computacional KNIME® (1). As moléculas restantes foram filtradas pelo modelo farmacofórico na plataforma UNITY (2). Estudos de acoplamento molecular para selecionar as moléculas com maior afinidade com a enzima foram realizados utilizando o programa Surflex (3). O *docking* também utilizando o programa GLIDE foi realizado (4). Seguidamente, a predição das moléculas com o modelo de HQSAR selecionado foi realizada (5). Simulações de dinâmica molecular foram empregadas para avaliar a afinidade e estabilidade de interação das moléculas no sítio alvo (farmacóforo) (6). Finalmente, predições de biodisponibilidade (admetSAR) e toxicidade (Pred-hERG e ProTox-II) foram realizadas a fim de se desconsiderar moléculas indesejáveis na aquisição (7).

A ordem dos programas na abordagem Surflex > GLIDE > HQSAR também foi alterada para GLIDE > Surflex > HQSAR e HQSAR > Surflex > GLIDE, empregando assim uma abordagem complementar dentre os dois programas nas análises de acoplamento molecular e o modelo de HQSAR, gerando maior diversidade de resultados. No final das quatro etapas, os *hits* putativos identificados foram adquiridos e/ou sintetizados para validação e confirmação da atividade biológica nos testes *in vitro*. Posteriormente, complementando os resultados dessas etapas e das moléculas obtidas outros MLT, como vizinhos mais próximos de “k” (do inglês, como *k Nearest Neighbours* – kNN), floresta aleatória (do inglês, *random forest* – RF) e máquina de vetor suporte polinomial (do inglês, *Support Vector Machine* – SVM – *polynomial*) foram empregados para predição da atividade biológica.

Na construção do *protomol* ou cavidade filtro do modelo farmacofórico, dois parâmetros foram considerados: (i) *threshold* (0,01-0,99), um limiar de volume máximo à cavidade passível de encaixe de moléculas na protease; (ii) *bloat* (0-10,0), tamanho (Å) do espaço permissível do *protomol*, que teoricamente, quanto menor, mais estável. Além dos critérios de orientação e volume selecionados, alguns parâmetros foram analisados comparativamente e selecionados para padronização de um modelo ótimo baseado nas interações do ligante (SYBYL and related Tripos modules ©, 2010,).

Os parâmetros analisados foram: (i) Surflex Docking (Screen, Geom e GeomX): variantes do modelo de *docking*. Screen é considerado o mais rápido (densidade de *spin* eletrônico = 3, ou seja, probabilidade de encontro de um elétron), Geom o intermediário (*spin* = 6) e GeomX o mais exaustivo ou de maior acurácia (*spin* = 9); (ii) mobilidade de hidrogênios (sim ou não): permite uma pré-simulação de interferência ou interação de ligações de hidrogênio externas à cavidade ou dos átomos na cavidade; (iii) conformações adicionais iniciais (0-16): permite flexibilidade, com mais conformações de grandes moléculas, cadeias ou aa, sendo no mínimo seis quando presentes argininas ou lisinas na molécula; (iv) expansão de busca (Å): limita a distância de interação entre cada átomo do ligante com o sítio/subsítios (sendo 6Å suficiente para as três ligações possíveis: hidrogênio, iônica e hidrofóbica); (v) conformações máximas por fragmento (0-100): delimita a quantidade de confôrmeros que o ligante poderá apresentar ao tentar interagir com a protease; (vi) ligações rotacionáveis máximas (0-100): flexibilidade rotacional de cada átomo ou grupo funcional ao interagir com átomos da protease; (vii) *spins* por alinhamento de moléculas (12-36): controle de rotações dos vetores de ligação, ou seja, o direcionamento da ligação e do momento polo-dipolo; (viii) poses máximas por ligantes (3-100): número de posições possíveis das moléculas na cavidade de interação. Um número menor favorece posições de flexibilidade controlada, reduzindo a possibilidade de interações menos prováveis ou dificultadas; (ix) densidade de busca (3, 6 ou 9Å): limita a uma triagem simples (3Å), de maior acurácia (6Å) ou exaustiva (9Å); (x) RMSD mínimo (0,01-0,99): valor da média da distância das posições atômicas entre estruturas ou ligações (AN, 2003, 2007).

4.1.7 Dinâmica molecular

As dinâmicas moleculares (DM) na protease de DENV, junto ao inibidor presente na estrutura selecionada (PDB: 3U11) ou moléculas obtidas na triagem foram preparadas ajustando estados de ionização de aa e corrigindo átomos ausentes de cadeias laterais (PrepWiz, Maestro

v2019.1). Segmentos ausentes de NS3 foram modelados utilizando o programa Prime, removendo sequências incompletas e íons sulfato. Moléculas de água com três ou mais ligações de hidrogênio foram mantidas e foi delimitado Ser135 para simulações de interações covalentes. Posteriormente, o sistema teve seu estado de protonação ajustado utilizando o programa Epik, especificamente na tríade catalítica, além da minimização da energia do sistema (JACOBSON et al., 2004; SHELLEY et al., 2007; MAGILL et al., 2014).

As DM foram realizadas utilizando programa Desmond com o campo de força OPLS3e (BOWERS et al., 2006). O sistema simulado abrangeu o complexo ligante-alvo, um modelo de água predefinido (TIP3P) como solvente explícito e contra-íons (Na^+ ou Cl^-) ajustados para neutralizar a carga geral do sistema (em torno de 4 a 5 átomos de Na^+). O sistema foi tratado em uma caixa cúbica com condições de contorno periódicas, especificando a forma e o tamanho da caixa a uma distância de 13 Å das bordas da caixa a qualquer átomo da proteína (cerca de 40.000 átomos entre proteína, ligante, solvente e íons). Um intervalo de tempo de 1 fs foi utilizado, tratando as interações coulômbicas de curto alcance com um valor de corte de 9 Å (método de curto alcance), enquanto o método de malha suave de partículas de Ewald (do inglês, *Particle Mesh Ewald* – PME) foi utilizado para interações coulômbicas de longo alcance (DARDEN; YORK; PEDERSEN, 1993).

Inicialmente, o relaxamento do sistema foi realizado usando os algoritmos de fase estacionário (*Steepest Descent*) e de memória limitada (Broyden-Fletcher-Goldfarb-Shanno) de maneira híbrida. Esta simulação foi realizada sob o conjunto por 5 ns, implementando os métodos de termostato e barostato de Berendsen. Uma temperatura constante de 310 K (Kelvin), equivalente a 36,85 °C, foi mantida durante toda a simulação usando o algoritmo do termostato Nose-Hoover e o algoritmo Martyna-Tobias-Klein Barostat para manter 1 atm de pressão. Finalizada a minimização e relaxamento do sistema, foi realizada etapa de produção de 1 μs . Todas as simulações de DM foram realizadas pelo menos em três execuções independentes com sementes geradas aleatoriamente. Nas simulações, as estruturas representativas da proteína foram selecionadas com base na inspeção visual das interações conservadas e na variação do RMSD. As interações e distâncias foram determinadas usando o pipeline de análise de eventos de simulação implementado no programa Maestro (Maestro v2019.1).

Os critérios geométricos atuais para a ligação de hidrogênio (H) do ligante de proteína foram: (i) distância de 2,5 Å entre os átomos doador (D) e aceptor (A) (D-H...A); (ii) ângulo

doador $\geq 120^\circ$ entre os átomos doador-acceptor de hidrogênio (D-H...A); e (iii) ângulo acceptor $\geq 90^\circ$ entre os átomos ligados ao acceptor de hidrogênio (H...A-X). Quanto à ligação de hidrogênio da proteína-água ou do ligante água: (i) distância de 2,8 Å entre os átomos doador e acceptor (D-H...A); (ii) ângulo doador $\geq 110^\circ$ entre os átomos doador-acceptor de hidrogênio (D-H...A); e (iii) ângulo acceptor $\geq 90^\circ$ entre os átomos ligados ao acceptor de hidrogênio (H...A-X). As interações hidrofóbicas inespecíficas são definidas pela cadeia lateral hidrofóbica a 3,6 Å dos carbonos aromáticos ou alifáticos de um ligante e as interações π - π requerem dois grupos aromáticos empilhados frente a frente, a 4,5 Å de distância. As trajetórias da dinâmica molecular foram visualizadas e as figuras foram produzidas utilizando programa PyMOL.

4.1.8 Predições de toxicidade

As propriedades das moléculas referentes às características farmacocinéticas (biodisponibilidade) e toxicológicas foram preditas empregando os servidores admetSAR (YANG et al., 2018; <http://lmm.d.ecust.edu.cn/admetSAR2/>), ProTox-II (BANERJEE et al., 2018; http://tox.charite.de/prottox_II/) e Pred-hERG 4.2 (BRAGA, 2014; <http://predherg.labmol.com.br/>). Todas as moléculas oriundas da triagem virtual foram submetidas em formato sdf (do inglês, *spatial data file*) aos servidores, retornando dados de predição, a partir de moléculas já descritas com uma ou mais propriedades associadas à farmacocinética ou toxicidade.

4.1.9 Modelos de aprendizado de máquina (MLT) para predição da atividade inibitória da protease de DENV

Diversos modelos computacionais foram desenvolvidos para predição de parâmetros envolvidos na farmacocinética, toxicidade de moléculas bioativas e da própria atividade biológica. As MLT por exemplo, envolvem modelos matemáticos que correlacionam uma estrutura química e suas propriedades físico-químicas com determinados parâmetros em estudo (BARREIRO, FRAGA, 2015; YANG et al., 2018).

As propriedades físico-químicas das substâncias foram calculadas utilizando o pacote computacional KNIME® (BEISKEN et al., 2013). Os dados foram analisados e comparados com os parâmetros definidos pelas regras de Lipinski e de Veber, a fim de descrever o caráter *druglike* das moléculas obtidas na triagem. Foram utilizados para avaliar as propriedades moleculares, seus respectivos cálculos preditivos, bem como sua classificação quanto às regras empíricas.

Inicialmente, os mesmos compostos peptidomiméticos separados nos grupos treinamento foram empregados para treinar MLT como árvores de decisão (do inglês, *decision trees* – DT), RF, multicamadas *perceptron* (do inglês, *Multilayer Perceptron* – MLP), kNN, *Naive Bayes* (NB), do teorema de Bayes e máquinas de vetores suporte, ambos utilizados modelos de classificação. Já os compostos dos conjuntos teste foram empregados na validação externa.

As propriedades citadas anteriormente (peso molecular, Log P, aceptores e doadores de hidrogênio, ligações rotacionáveis e área de superfície polar), bem como doze conjuntos de impressões digitais, foram calculados com o descritor PaDEL: *fingerprint*, *fingerprint* estendido, *EState*, gráficos, MACCS, PubChem, subestruturas, contagem de subestruturas, Kleklota Roth, contagem de Kleklota Roth, *Atom Pair 2D* e contagem de *Atom Pair 2D* (YAP, 2011). Todos os conjuntos de impressões digitais foram usados para gerar modelos baseados em combinação com uma variação de parâmetros intrínsecos de cada técnica.

Os modelos de DT foram construídos variando o número mínimo de registros por nó (2 a 100). Os parâmetros de RF foram rastreados heurísticamente por número variável de níveis (10 a 50, em etapas de 10 níveis) e vários modelos (10.000 a 50.000 em etapas de 10.000). Os parâmetros de MLP também foram rastreados heurísticamente variando um número de camadas (1 a 5, nas etapas de uma camada) e um número de neurônios ocultos por camada (20 a 100, nas etapas de 20 neurônios). Os modelos kNN foram gerados variando o número de vizinhos mais próximos de “k” entre 1 a 27, em números ímpares. Modelos de kNN com vizinhos ponderados por distância também foram gerados usando mesma variação de parâmetro k. O parâmetro de probabilidade de NB foi triado de 0,1 a 5 em etapas de 0,1.

Finalmente, os modelos de SVM foram treinados usando três núcleos diferentes: função de base radial (do inglês, *Radial Basis Function* – RBF), hiper tangente (do inglês, *Hyper Tangent* – HT) e polinomial. Cada modelo conta com um conjunto de parâmetros rastreado e variável: sigma de 0,5 a 300 (RBF); kappa e delta de 1 a 50; viés e gama de 1 a 5 (polinomial). Uma sobreposição de penalidade (do inglês, *C parameter*) também variou de 1 a 1500 para todos os três modelos de SVM (MALTAROLLO et al., 2019).

4.1.10 Validações e seleção dos modelos de aprendizado de máquina

Todos os modelos gerados foram validados pelo coeficiente de correlação de Mathew (do inglês, *Matthew's Coefficient Correlation* – MCC), *score* F1, área sob a curva (do inglês,

area under curve – AUC) de uma curva característica operacional do receptor (ROC) e taxa de verdadeiros positivos (do inglês, *true positive rate* – TPR). Esses quatro parâmetros de validação selecionados foram obtidos para os compostos do grupo teste (métricas de validação externa) e em uma validação cruzada (CV) de cinco vezes (métricas de validação interna) (ROY; KAR; DAS, 2015).

Os modelos foram classificados de acordo com o MCC de validação externa (extMCC) calculado com o conjunto de testes, seguido por *score* F1, AUC e TPR também de validações externas. As métricas calculadas cinco vezes na CV foram empregadas para verificar a consistência dos modelos. Nesse caso, quando um modelo mais preditivo (valor extMCC mais alto) apresentou um cvMCC baixo, também foi selecionado um modelo consistente ao selecionar as médias mais altas das métricas somadas (cvMCC + extMCC) e com a menor diferença (Δ MCC) entre si.

Finalizados os treinamentos e validações dos modelos, apenas os três que apresentaram valores mais robustos de predição (ROC mais próximo de 1,0) e nas métricas de validação (superiores a 0,6), foram selecionados. Nestes modelos, as moléculas triadas foram classificadas como ativas ou inativas a partir de um valor de corte de IC_{50} equivalente a 10 μ M.

4.2 Linhagens Celulares

A produção dos estoques virais e titulação de ZIKV e YFV, bem como a avaliação da citotoxicidade (CC_{50}) e determinação da atividade antiviral (CE_{50}) das substâncias, foi feita em uma linhagem celular aderente de epitélio renal de macaco-verde africano (*Cercopithecus aethiops*), Vero (ATCC® CCL-81™). Já para os vírus DENV-2 e DENV-3, foi utilizada uma linhagem celular aderente do epitélio renal de hamster-sírio (*Mesocricetus auratus*), BHK-21 (ATCC® CCL-10™).

As células foram cultivadas em meio mínimo essencial de Eagle (do inglês, *Eagles's Minimum Essential Medium* – MEM) (Cultilab, Brasil). O meio foi suplementado com 5% de soro fetal bovino (SFB) (Cultilab, Brasil), além de 100 UI/mL de penicilina (Cellofarm, Brasil), 100 μ g/mL estreptomicina (Merck, Alemanha) e 0,25 μ g/mL anfotericina B (Cultilab, Brasil). As células permaneceram sob cultivo a 37 °C em atmosfera de 5% de CO_2 . Posteriormente, as células foram lavadas com salina tamponada com tampão fosfato (do inglês, *phosphate-buffered saline* – PBS), homogeneizadas em tripsina acrescida de ácido etilenodiaminoacético (EDTA) (Sinapse Biotecnologia, Brasil) e contadas para realização dos ensaios.

4.3 Vírus

4.3.1 Estoques

Vírus foram obtidos da coleção de vírus do Laboratório de Vírus, do Departamento de Microbiologia do Instituto de Ciências Biológicas (ICB) da UFMG. ZIKV (PE243) [LOTE: PE243/2015] foi isolado em Recife, Pernambuco, Brasil em 2015, na Fundação Oswaldo Cruz (FIOCRUZ). O paciente, do sexo masculino, apresentava sintomas clássicos de Zika: exantemas na face e membros, artralgia das mãos, pulsos e tornozelos, além de edema nas mãos e pulsos, sem sintomas neurológicos, tendo o genoma desta linhagem Asiática (Americana) sido sequenciado (DONALD et al., 2016). DENV-2 (PI59), linhagem II [L.: DENV2/PI59/2006] e DENV-3 (PI76) [L.: DENV2/PI76/2006], foram identificados e isolados de pacientes do sexo masculino durante um surto no Piauí, Brasil em 2006, como descrito anteriormente (FIGUEIREDO, 2014). A vacina contendo YFV (17DD) foi gentilmente cedida por Bio-Manguinhos, Brasil, 2008 (FIOCRUZ N. 980FB066Z).

4.3.2 Multiplicação

Células Vero foram cultivadas em meio MEM suplementado com 5% de SFB e incubadas a 37 °C em atmosfera de 5% de CO₂, em frascos de cultura de 75 cm². Logo após 24h de incubação, as monocamadas celulares com aproximadamente 60 a 80% de confluência foram lavadas com PBS para remoção de debris celulares e aminoácidos de soro, sendo infectadas com uma multiplicidade de infecção (MOI) de 0,01, ou seja, uma partícula viral para cada 100 células em cultivo (1:100), dos estoques virais (*pool* semente) de ZIKV e YFV em 2,0 mL de meio MEM sem adição de SFB. A adsorção foi feita com 1h de incubação em estufa a 37 °C e 5% de CO₂, homogeneizando delicadamente os frascos a cada 10 minutos. Logo após o período de adsorção, foram adicionados 12 mL de meio MEM com 2% de SFB aos frascos, incubados sob as mesmas condições, sendo observados diariamente em microscópio óptico para visualização da formação de efeitos citopáticos. Depois de visualizados efeitos em aproximadamente 70 a 80% das células, o sobrenadante foi removido e centrifugado a 3.000 rpm em centrífuga RT6000B (Sorvall, Thermo Scientific, EUA) por 10 minutos a 4°C para remoção de debris celulares, sendo aliquoteado em microtubos tipo Eppendorf de 1,5 mL, congelado a -70°C e, posteriormente, titulado.

4.3.3 Titulação

As células Vero e BHK-21 foram cultivadas em placas de 24 poços contendo $8,0 \times 10^5$ células/poço e incubadas a 37 °C e 5% de CO₂ por 24h. Logo após, as placas foram infectadas com 100 µL de diluições seriadas do vírus em proporção 1:10 (10^{-1} a 10^{-5}) e, após 1h de adsorção a 37 °C e 5% de CO₂, com homogeneização delicada das placas a cada 10 minutos, o meio foi desprezado, adicionando a cada poço 1,0 mL de meio M199 (do inglês, *Medium 199* – M199) (Cultilab, Brasil) suplementado com antibióticos e 2% de SFB, além de 1% de carboximetilcelulose (CMC) (Synth, Brasil). As placas foram incubadas por cinco (ZIKV e YFV) ou sete (DENV-2 e DENV-3) dias em estufa a 37 °C e 5% de CO₂ e, em seguida, foram fixadas com formalina 10% *overnight*, lavadas e coradas com cristal violeta 1% por 20 minutos. As placas foram novamente lavadas e colocadas para secar, contando a olho nu as placas de lise formadas (COELHO et al., 2017). Poços contendo cerca de 10 a 100 placas de lise foram considerados para cálculo do título viral em unidades formadoras de placa (UFP/mL), corrigindo o resultado pelo fator de correção do volume de adsorção (10^1) e o inverso da diluição seriada (10^1 a 10^6) do(s) poço(s) em que as placas foram contadas.

4.4 Ensaios biológicos

4.4.1 Avaliação da citotoxicidade: concentração citotóxica 50% (CC₅₀)

Monocamadas celulares foram cultivadas em placas de 96 poços contendo $4,0 \times 10^4$ células/poço em meio MEM com 1% de SFB, incubadas a 37 °C em atmosfera com 5% de CO₂ por 24h. Logo após a incubação, o meio foi descartado e foram adicionados 200 µL de meio MEM com 1% de SFB contendo diluições seriadas das substâncias (100 a 1,56 µM). Poços contendo uma diluição seriada de DMSO foram utilizados como controle do veículo, contendo uma solução 10% de DMSO como controle de inibição, contendo apenas meio como controle de viabilidade celular e aqueles sem células como controle de esterilidade e branco.

Logo após 72h de incubação sob as mesmas condições, a viabilidade celular foi avaliada empregando o teste colorimétrico de brometo de 3-(4,5-dimetiltiazol-2-il)-2,5-difeniltetrazólio (MTT), que permite avaliar indiretamente a viabilidade celular (MOSMANN, 1983). Neste teste, as mitocôndrias celulares reduzem o MTT em cristais de formazana, o que permite avaliar a concentração de cristais por densidade óptica (D.O.) após dissolver as células presente com DMSO. Assim, quanto maior a leitura da concentração, maior é a viabilidade (%) das células em comparação aos respectivos controles.

O meio foi removido e 100 µL de uma solução de MTT (ThermoFischer Scientific, Estados Unidos) em meio com 1% de SFB (0,5 mg/mL) foram adicionados em cada poço. As placas foram incubadas por 3h sob as mesmas condições, seguido da remoção do meio e adição de 100 µL de DMSO a cada poço, para dissolução dos cristais de formazana. Logo após, as placas foram agitadas por 20 minutos e a leitura da absorbância foi realizada em aparelho espectrofotômetro VersaMax (Molecular Devices, Estados Unidos), a uma D.O. de 570 nm. Os percentuais de inibição da viabilidade celular foram calculados utilizando o controle de veículo, conforme **Fórmula 1**, determinando os valores da concentração citotóxica 50% (CC₅₀) de cada molécula.

$$\text{Inibição de crescimento} = \left(\frac{A_{\text{amostra}} - A_{\text{branco}}}{A_{\text{veículo}} - A_{\text{branco}}} \right) \times 100 \quad (1)$$

Regressão linear foi utilizada para as análises, considerando válidos os resultados com valores de r² superiores a 0,9. Todas as condições foram testadas em triplicatas e em dois ensaios independentes.

4.4.2 Avaliação da atividade antiviral: concentração efetiva a 50% (CE₅₀)

Monocamadas celulares foram cultivadas em placas de 96 poços contendo 4,0 x 10⁴ células/poço em meio MEM com 1% de SFB, incubadas a 37 °C em atmosfera com 5% de CO₂ por 24h. Poços não contendo células foram utilizados como controle de esterilidade e/ou branco. Logo após a incubação, o meio foi descartado e foram adicionados 100 µL das diluições das substâncias (diluídas a partir dos valores de CC₅₀) juntamente a 100 µL da suspensão viral de ZIKV, YFV, DENV-2 ou DENV-3, com uma MOI de 0,1, ambos preparados em meio MEM com 1% de SFB. Poços contendo uma diluição seriada de DMSO e vírus foram utilizados como controle do veículo, contendo diluições de ribavirina ou ácido micofenólico e vírus como controles de inibição, contendo apenas meio e células como controle de viabilidade celular e aqueles sem células como controle de esterilidade ou branco.

Brevemente, após 72h de incubação sob as mesmas condições, a viabilidade celular foi avaliada empregando o teste colorimétrico de brometo de 3-(4,5-dimetiltiazol-2-il)-2,5-difeniltetrazólio (MTT), seguindo as mesmas etapas em **4.4.1** (TAKEUCHI; BABA; SHIGETA, 1991). Nesse ensaio, a concentração efetiva a 50% (CE₅₀) é considerada como um efeito de proteção de uma substância nas células infectadas, sendo calculada de acordo com a **Fórmula 2**. Os índices de seletividade (IS) foram obtidos calculando a razão entre a CC₅₀ e os valores de CE₅₀.

$$\text{Inibição de multiplicação} = \left(\frac{A_{vs} - A_v}{A_c - A_v} \right) \times 100 \quad (2)$$

A_{vs} : absorvância de células infectadas com vírus na presença das substâncias;

A_v : absorvância de células infectadas com vírus e não tratadas; e

A_c : absorvância de células não infectadas e não tratadas.

Regressão linear for utilizada para as análises, considerando válidos os resultados com valores de r^2 superiores a 0,9. Todas as condições foram testadas em triplicatas e em dois ensaios independentes. Ensaio de redução de placas de lise serão realizados para corroborar a atividade antiviral avaliada por MTT, considerando uma redução mínima de 50% de UFP/mL.

4.4.3 Avaliação da atividade virucida

Monocamadas celulares foram cultivadas em placas de 24 poços contendo $8,0 \times 10^4$ células/poço e incubadas a 37 °C em atmosfera 5% de CO₂ por 24h. As diluições das substâncias nos valores de EC₅₀ foram incubadas juntamente às suspensões virais a contendo 8.000 UFP/mL por 1h a 37 °C e 5% de CO₂. Vírus incubado apenas com meio e DMSO, na proporção das diluições das substâncias, foram considerados como controles. Logo após as 24h, o meio das placas foi removido e a monocamada celular lavada com PBS. As suspensões virais, pré-incubadas com as diluições das substâncias, foram diluídas em proporção 1:10 (10^{-1} a 10^{-5}) e 100 µL plaqueados por poço. A adsorção foi realizada por 1h com incubação em estufa a 37 °C e 5% de CO₂, homogeneizando delicadamente as placas a cada 10 minutos. Logo após a adsorção o meio foi desprezado, adicionando a cada poço 1,0 mL de meio M199 2% de SFB e 1% de CMC. As placas foram incubadas por cinco (ZIKV e YFV) ou sete dias (DENV-2 e DENV-3) a 37 °C em atmosfera 5% de CO₂ por 24h e, posteriormente, fixadas com formalina 10% *overnight*, seguida por coloração com cristal violeta 1% por 20 minutos. As placas foram lavadas e colocadas para secar, contando a olho nu as placas de lise formadas. O título foi determinado como descrito no item 4.4.1. Foram consideradas ativas as moléculas que apresentassem uma redução mínima de 50% de UFP/mL.

4.4.4 Ensaio de redução do número de placas de lise

Monocamadas celulares foram cultivadas em placas de 24 poços contendo $8,0 \times 10^4$ células/poço e incubadas a 37 °C em atmosfera 5% de CO₂ por 24h. Diluições das substâncias, sendo em uma e duas vezes os valores de EC₅₀, foram incubadas sob as mesmas condições por 48h, juntamente às suspensões virais em uma MOI de 0,1. Vírus incubado apenas com meio e

DMSO, na proporção das diluições das substâncias, foram considerados como controles. Logo após, os sobrenadantes foram coletados e titulados segundo item 4.4.1. Poços contendo cerca de 10 a 100 placas de lise foram considerados para cálculo do título viral em unidades formadoras de placa (UFP/mL), corrigindo o resultado pelo fator de diluição do volume de adsorção (1:10) e o inverso da diluição seriada do(s) poço(s) em que as placas foram contadas.

4.4.5 Ensaios de pré-tratamento

Monocamadas celulares foram cultivadas em placas de 24 poços contendo $8,0 \times 10^4$ células/poço e incubadas a 37 °C em atmosfera 5% de CO₂ por 24h. O meio foi removido, a monocamada celular lavada com PBS e diluições das moléculas (nos valores obtidos de EC₅₀) foram adicionadas e incubadas sob as mesmas condições por 1h. Logo após este período de incubação, o meio foi removido e a monocamada celular novamente lavada com PBS, adicionando 100 µL das suspensões virais a uma MOI de 0,1, diluídas em proporção 1:10 (10^{-1} a 10^{-5}). Vírus incubado apenas com meio e DMSO, na proporção das diluições das substâncias, foram considerados como controles. A adsorção foi realizada conforme descrito acima. Logo após a adsorção o meio foi desprezado, adicionando a cada poço 1,0 mL de meio M199 2% de SFB e 1% de CMC. As placas foram incubadas por cinco (ZIKV e YFV) ou sete dias (DENV-2 e DENV-3) a 37 °C em atmosfera 5% de CO₂ por 24h e, posteriormente, fixadas com formalina e coradas com cristal violeta conforme descrito. O título foi determinado como no item 4.4.1.

5. RESULTADOS E DISCUSSÃO

5.1 Análises *in silico*

5.1.1 Construção e preparação do conjunto de dados

Foram selecionadas, para construção dos modelos de HQSAR, 111 moléculas da literatura derivadas de peptídeos sintéticos (peptideomiméticos), com atividade inibitória determinada contra a protease de DENV-3 e confirmadas por ensaios de redução de placas de lise (BEHNAM et al., 2015; WEIGEL et al., 2015). Alguns fatores favoreceram a seleção destas moléculas para as análises de HQSAR, como a estrutura geral conservada e variedade de substituintes (**Figura 17**), atividade inibitória (IC_{50}) descrita contra a protease de DENV-3 e atividade antiviral (EC_{50}) confirmada contra o vírus. O conjunto das estruturas utilizadas estão dispostos no **Apêndice A**. Destes, 86 (numeradas de 1 a 86) compostos pertencem ao trabalho de Behnam e 25 (numeradas de 87 a 111) aos estudos de Weigel.

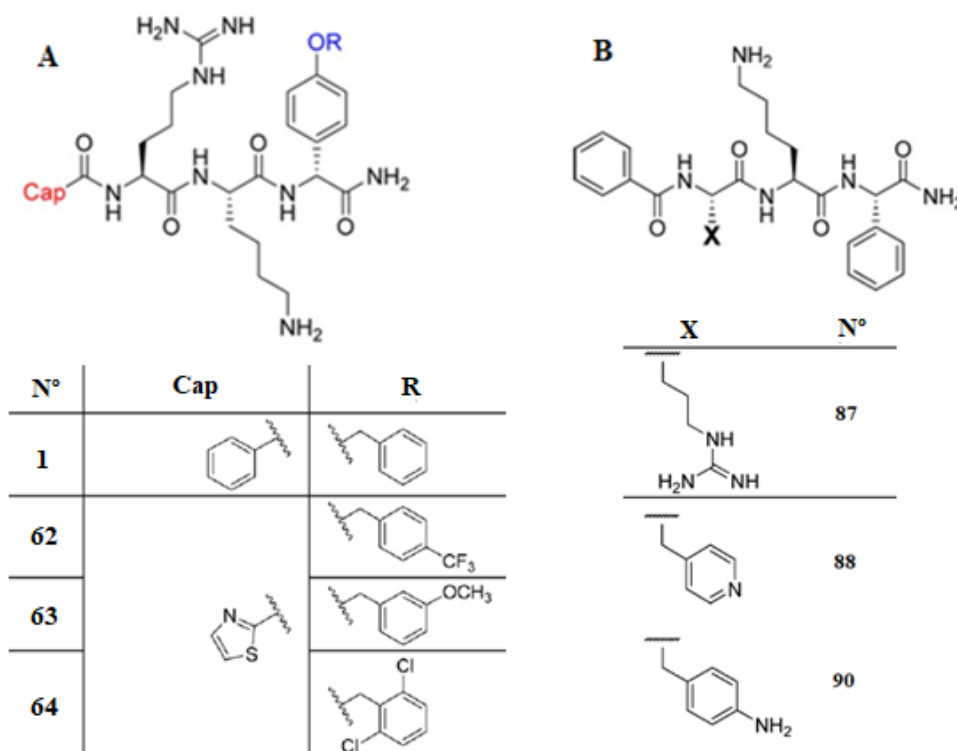


Figura 17. Estrutura geral das moléculas utilizadas na construção do conjunto de dados. (A). Estrutura referente aos 86 compostos de BEHNAM et al., 2015 e exemplos de substituintes Cap (vermelho) e R (azul). (B). Estrutura dos 25 compostos de WEIGEL et al., 2015 e exemplos do substituinte X.

Cada uma das 111 moléculas construída foi individualmente conferida quanto à sua estereoquímica. Essa etapa denota uma importância para o modelo de HQSAR, uma vez que a estereoquímica pode estar diretamente relacionada à atividade biológica de uma molécula.

Além disso, a configuração do estereocentro, ou seja, a orientação R (sentido horário) ou S (sentido anti-horário) dos grupos substituintes ligados molécula, quando incorreta, poder induzir um modelo a erros de predição da atividade e/ou de contribuição química da estrutura (WINKLER, 2002).

Logo após a construção da estrutura tridimensional (3D), um banco de confôrmeros foi criado, obtendo a conformação de maior estabilidade de cada molécula. Essa etapa também é importante em análises computacionais como HQSAR, uma vez que, quanto maior o número de moléculas e confôrmeros em um banco de dados, maior são, por exemplo, as variações de coordenadas atômicas de hidrogênios nas moléculas (BORODINA et al., 2007).

Finalmente, cada molécula também teve seu estado de ionização corrigido de acordo com pH fisiológico. Obtendo o estado iônico mais favorável, um melhor perfil das interações possíveis e maior confiabilidade de um modelo é obtido, favorecendo, por exemplo, análises posteriores de acoplamento molecular (BOLCATO et al., 2019). Dentre as 111 moléculas, 14 delas (6 de Behnam e 8 de Wiegel) foram retiradas do conjunto de dados, uma vez que não apresentaram atividade inibitória na protease de DENV ou não foram avaliadas experimentalmente. Dessa forma, a separação dos grupos no método de HQSAR foi realizada com as 97 moléculas restantes.

5.1.2 Seleção do modelo de relação quantitativa entre a estrutura química e atividade biológica por holograma (HQSAR)

As 97 moléculas restantes foram divididas aleatoriamente em dois conjuntos, treinamento ($n = 77$) e teste ($n = 20$), removendo do conjunto teste as moléculas mais ativas de cada série de compostos nos dois estudos, 62 (Behnam et al. 2015; $IC_{50} = 0,018 \mu M$) e 106 (Weigel et al. 2015; $IC_{50} = 0,21 \mu M$). Os modelos de HQSAR foram criados fixando o tamanho do fragmento em relação ao número de átomos (4 a 7) (**Tabela 10**), sendo selecionados os quatro modelos com maior valor de q^2 . Essa métrica é considerada uma prova de robustez de um modelo selecionado em técnicas de QSAR, quando apresentados altos valores estatísticos e obedecendo a alguns critérios de validação, como $q^2 > 0,5$ (GOLBRAIKH; TROPSHA, 2002). Nesse aspecto, também se faz necessária à validação da capacidade de estimar a atividade predita do método, com valores mínimos de $r^2 > 0,6$ e, idealmente, superiores 0,65 e não muito distantes de q^2 (TROPSHA; GRAMATICA; GOMBAR, 2003; GRAMATICA; SANGION, 2016). Nesse sentido, os quatro modelos superaram os valores mínimos de q^2 (0,629; 0,61; 0,609 e 0,584) e ideias de r^2 (0,929; 0,735; 0,736 e 0,915).

Tabela 10. Relação quantitativa entre a estrutura química e atividade biológica por holograma (HQSAR) do conjunto de dados. A: átomos; B: ligações; C: conectividade; HA: hidrogênios; Ch: quiralidade; DA: aceptores e doadores. Tamanho de fragmentos 4 a 7 átomos. q^2 : coeficiente de validação; r^2 : coeficiente de correlação não-validado; SEE: erro padrão da estimativa; SEV: erro padrão da validação; BL: extensão do holograma; COMP: componentes principais. Quatro modelos (valores de q^2) selecionados em negrito.

4 a 7 átomos	q^2	r^2	SEV	SEE	BL	COMP
A, C	0,629	0,929	0,38	0,166	307	13
A, Ch	0,610	0,735	0,365	0,301	401	4
A	0,609	0,736	0,365	0,3	401	4
A,C,Ch	0,584	0,915	0,397	0,18	307	11
A,B,Ch	0,572	0,754	0,385	0,292	61	5
A,C,HA,Ch	0,566	0,824	0,393	0,25	307	7
A, B	0,560	0,797	0,393	0,267	353	6
A,C,HA	0,553	0,806	0,402	0,265	71	8
A,B,HA	0,552	0,918	0,425	0,182	71	15
A,B,C,Ch	0,550	0,784	0,392	0,271	353	4
A,B,C	0,549	0,796	0,395	0,266	353	5
A,HA,Ch	0,546	0,782	0,399	0,277	151	6
A,B,HA,DA	0,543	0,887	0,419	0,208	83	12
A, DA	0,540	0,834	0,408	71	8	3
A, HA	0,537	0,786	0,403	0,274	151	6
A,B,C,HA	0,532	0,938	0,434	0,158	353	15
A,HA,DA	0,525	0,927	0,437	0,171	257	15
A,B,C,HA,Ch	0,524	0,747	0,406	0,296	401	5
A,B,Ch,DA	0,522	0,752	0,404	0,291	83	4
A,B,HA,Ch	0,522	0,705	0,407	0,32	199	5
A,B,DA	0,519	0,789	0,411	0,272	59	6
A,C,HA,DA	0,519	0,94	0,44	0,155	307	15
A,Ch,DA	0,517	0,819	0,418	0,256	71	8
A,C,DA	0,486	0,921	0,441	0,173	307	11
A,B,C,HA,DA	0,482	0,944	0,457	0,15	353	15
A,B,C,DA	0,480	0,949	0,458	0,144	353	15
A,C,Ch,DA	0,464	0,933	0,457	0,162	307	13
A,HA,Ch,DA	0,462	0,848	0,441	0,234	307	8
A,B,HA,Ch,DA	0,460	0,755	0,435	0,293	53	6
A,C,HA,Ch,DA	0,454	0,885	0,451	0,207	307	10
A,B,C,Ch,DA	0,432	0,729	0,44	0,304	151	4
A,B,C,HA,Ch,DA	0,428	0,54	0,439	0,394	83	3

A análise de HQSAR com tamanho fixo dos fragmentos (4 a 7 átomos) apresentou átomos e conectividade (A, C) como o modelo mais robusto, com maior valor da métrica de q^2 (0,629). Esse resultado indica que a conectividade, ou seja, o estado de hibridização dos átomos na molécula, são suficientes para prever a atividade biológica. Além disso, apresentar somente uma distinção além de átomos torna este um modelo mais simples para previsões. Já o segundo modelo mais robusto envolve átomos e quiralidade (A, Ch; $q^2 = 0,610$), indicando que a estereoquímica da molécula (orientação R ou S) também é importante para a previsão da atividade biológica. Curiosamente, o quarto melhor modelo une ambas as distinções dos dois melhores modelos (A, C e Ch; $q^2 = 0,584$), o que pode sugerir uma correlação importante entre a presença da conectividade dos átomos e a estereoquímica da molécula na previsão da atividade.

Pensando nisso, os quatro modelos de maiores valores de q^2 também foram submetidos a variações dos tamanhos dos fragmentos, em intervalos de 2, 3 e 4 átomos (1 a 3 até 11 a 13; 1 a 4 até 10 a 13; e 1 a 5 até 9 a 13), mantendo sempre os descritores de cada modelo. Dentre todos estes modelos, 20 foram selecionados (5 de cada), considerando, novamente, os maiores valores de q^2 (**Tabela 11**). Essa abordagem permite avaliar todos os fragmentos possíveis no holograma, incluindo sobreposições e a contagem de cada vez que um fragmento ocorre. Dessa forma, com todas as possibilidades de fragmentos com tamanhos entre átomos cobertas no *fingerprint*, os hologramas codificam as informações químicas das moléculas, incluindo sua topologia, representando dados mais robustos de correlação de um modelo de previsão (HERITAGE; LOWIS, 1999; GARCIA; HONÓRIO, 2011a).

Tabela 11. Variação do tamanho dos fragmentos dos quatro modelos de HQSAR. (5 ou 6) do tamanho dos fragmentos de cada. A: átomos; B: ligações; C: conectividade; HA: hidrogênios; Ch: quiralidade; DA: aceptores e doadores. Tamanho de fragmentos 4 a 7 átomos. q^2 : coeficiente de validação; r^2 : coeficiente de correlação não-validado; SEE: erro padrão da estimativa; SEV: erro padrão da validação; BL: extensão do holograma; COMP: componentes principais.

Distinção	Tamanho	q^2	r^2	SEV	SEE	BL	COMP
A	11_13	0,672	0,809	0,334	0,255	59	4
A	10_12	0,668	0,851	0,34	0,231	53	7
A	7_11	0,664	0,869	0,348	0,218	83	8
A	8_10	0,664	0,905	0,351	0,186	353	9
A	8_11	0,663	0,908	0,352	0,183	353	9
A,C	8_10	0,677	0,879	0,337	0,206	257	6
A,C	9_12	0,677	0,873	0,339	0,213	59	7
A,C	8_11	0,673	0,853	0,336	0,226	353	5
A,C	7_11	0,671	0,87	0,34	0,214	353	6
A,C	7_10	0,669	0,861	0,341	0,221	199	6
A, Ch	9_13	0,691	0,918	0,342	0,176	83	11
A, Ch	7_10	0,684	0,914	0,343	0,179	257	10
A, Ch	6_10	0,681	0,914	0,345	0,179	257	10
A, Ch	9_12	0,674	0,91	0,349	0,183	97	10
A, Ch	8_10	0,668	0,915	0,352	0,178	257	10
A, C, Ch	8_11	0,687	0,878	0,332	0,207	199	6
A, C, Ch	10_12	0,676	0,876	0,337	0,209	199	6
A, C, Ch	7_11	0,675	0,873	0,338	0,211	199	6
A, C, Ch	7_9	0,674	0,867	0,338	0,216	199	6
A, C, Ch	10_13	0,669	0,868	0,341	0,215	151	6

A variação do tamanho dos fragmentos resultou em modelos mais robustos ($q^2 = 0,691$) comparados àqueles com tamanho fixo de 4 a 7 átomos ($q^2 = 0,629$). Similarmente aos resultados anteriores, o modelo mais robusto ao realizar a variação de tamanho dos fragmentos, envolve quiralidade (A, Ch de 9 a 13 átomos; $q^2 = 0,691$), resultado que também corrobora à importância da estereoquímica de uma molécula nos modelos de predição. Nessa análise também, a presença da conectividade é evidenciada no segundo modelo mais robusto (A, C, Ch de 8 a 11 átomos; $q^2 = 0,687$), denotando também a relevância dessa distinção nos modelos de HQSAR, novamente sugerindo uma correlação com a quiralidade, igualmente aos resultados da análise anterior.

Ainda que as variações de tamanho dos fragmentos moleculares dentre as quatro distinções obtidas apresentem um maior valor geral de q^2 , esta é uma métrica que pode subestimar a qualidade preditiva de um modelo ao avaliar apenas as moléculas que foram utilizadas na construção do modelo, necessitando de demais validações para se obter um modelo confiável (GOLBRAIKH, 2002). Nesse caso, uma validação externa e outros métodos de validações cruzadas são indispensáveis para se obter os melhores parâmetros de caracterização e validação da preditividade dos modelos de QSAR (GRAMATICA; SANGION, 2016).

Os 20 modelos foram submetidos a uma validação externa no pacote computacional KNIME®, em que se obteve um modelo ótimo $q^2 = 0,675$ e $r^2 = 0,873$ (A, C, Ch de 7 a 11 átomos) (**Tabela 12**). Embora não apresentasse os maiores valores de q^2 e r^2 , outras métricas são usadas em conjunto para avaliar o modelo, como menores valores de RMSE e MAE, que distribuem igualmente as diferenças individuais de um grupo analisado e a média da magnitude do erro em um grupo, respectivamente, indicando um modelo que apresenta o menor erro preditivo (CHAI; DRAXLER, 2014). Esses dois parâmetros acabam corroborando as métricas de correlação por regressão obtidas nos testes, a exemplo dos testes de coeficientes de validação de q^2 (coeficiente de correlação preditiva quadrático) e médias de r^2 , ambos importantes individualmente e em concordância, testificando a robustez de um modelo nos métodos de QSAR (KIRALJ; FERREIRA, 2009; CHIRICO; GRAMATICA, 2011; CHAI; DRAXLER, 2014).

Tabela 12. Validação externa dos modelos de HQSAR selecionados por q^2 . Modelo selecionado (A, C, Ch de 7 a 11 átomos) assinalado em negrito. F: coeficiente de correlação preditiva quadrático. CCC: coeficiente de correlação de concordância. AVG r^2m : teste de Roy (médias de parâmetros r^2m e r^2m'). Δ : diferença entre as médias r^2m e r^2m' (erro). RMSE: erro quadrático médio. MAE: média do erro absoluto.

Modelo	q^2	r^2	BL	q^2 F1	q^2 F2	q^2 F3	CCC	r^2m	r^2m'	AVG r^2m	Δ r^2m	RMSE	MAE
A/C/Ch 7_11	0,675	0,873	199	0,866	0,859	0,877	0,932	0,837	0,782	0,809	0,055	0,203	0,161
A/C/Ch 8_11	0,687	0,878	199	0,856	0,848	0,868	0,926	0,821	0,771	0,796	0,050	0,210	0,168
A/C/Ch 7_9	0,674	0,867	199	0,840	0,832	0,854	0,916	0,764	0,787	0,775	0,023	0,221	0,172
A/C/Ch 10_12	0,676	0,876	199	0,830	0,821	0,844	0,901	0,658	0,815	0,736	0,157	0,228	0,173
A 8_11	0,663	0,908	353	0,829	0,820	0,844	0,911	0,768	0,760	0,764	0,008	0,229	0,170
A/Ch 8_10	0,668	0,915	257	0,826	0,817	0,841	0,918	0,834	0,694	0,764	0,140	0,231	0,188
A/Ch 7_10	0,684	0,914	257	0,806	0,796	0,823	0,910	0,814	0,663	0,738	0,151	0,244	0,198
A/Ch 6_10	0,681	0,914	257	0,790	0,779	0,808	0,905	0,793	0,635	0,714	0,158	0,254	0,212
A 7_11	0,664	0,869	83	0,789	0,778	0,807	0,891	0,735	0,700	0,717	0,036	0,254	0,187
A/Ch 9_12	0,674	0,91	97	0,789	0,778	0,807	0,892	0,732	0,689	0,711	0,044	0,254	0,179
A/C 8_10	0,677	0,879	257	0,766	0,754	0,786	0,866	0,586	0,725	0,655	0,139	0,268	0,204
A/C 7_10	0,669	0,861	199	0,761	0,748	0,781	0,878	0,698	0,659	0,679	0,039	0,271	0,214
A/C 8_11	0,673	0,853	353	0,752	0,739	0,773	0,851	0,523	0,740	0,632	0,217	0,275	0,197
A/C 7_11	0,671	0,87	353	0,744	0,730	0,765	0,843	0,494	0,722	0,608	0,228	0,280	0,202
A 11_13	0,672	0,809	59	0,727	0,712	0,750	0,869	0,724	0,604	0,664	0,119	0,289	0,225
A/C/Ch 10_13	0,669	0,868	151	0,684	0,667	0,710	0,810	0,448	0,636	0,542	0,187	0,311	0,252
A 8_10	0,664	0,905	353	0,677	0,660	0,704	0,849	0,700	0,540	0,620	0,160	0,315	0,261
A/C 9_12	0,677	0,873	59	0,660	0,642	0,689	0,803	0,454	0,604	0,529	0,150	0,323	0,236
A/Ch 9_13	0,691	0,918	83	0,394	0,361	0,445	0,721	0,447	0,320	0,384	0,127	0,431	0,324
A 10_12	0,668	0,851	53	0,182	0,138	0,251	0,618	0,287	0,235	0,261	0,052	0,501	0,289

Embora, comparativamente, o modelo ótimo não tivesse os maiores valores de correlação (q^2) e regressão (r^2) dentre os modelos validados, as métricas dos testes e coeficientes de validação, bem como testes dos erros, o colocam como modelo ideal dentre aqueles obtidos na etapa anterior, demonstrando, comparativamente, uma maior preditividade do modelo ótimo selecionado (Tabela 13).

Tabela 13. Comparação dentre três modelos ótimos de HQSAR. Comparativo determinado pelas métricas do método predito e da validação externa aplicada (q^2 , r^2 , SEE, SEV, BL e COMP comparado a q^2 F1, q^2 F2, q^2 F3, CCC, AVG r^2 m, Δr^2 m, RMSE e MAE).

Técnica	HQSAR		
Distinção do Fragmento	A/C/Ch	A/Ch	A/C
Tamanho do Fragmento	7 a 11	9 a 13	8 a 10
q²	0,675	0,691	0,677
r²	0,873	0,918	0,879
SEV	0,338	0,342	0,337
SEE	0,211	0,176	0,206
BL	199	83	257
COMP	6	11	6
q² F1	0,866	0,394	0,766
q² F2	0,859	0,361	0,754
q² F3	0,877	0,445	0,786
CCC	0,932	0,721	0,918
AVG r²m	0,809	0,384	0,655
Δr^2m	0,055	0,127	0,139
RMSE	0,203	0,431	0,268
MAE	0,161	0,324	0,204

5.1.3 Validações do modelo ótimo de HQSAR selecionado

O modelo ótimo selecionado também foi submetido a testes de validação independentes, sendo: (i) um coeficiente de validação cruzada (do inglês, *cross validation* – CV) (**Figura 18A**), em que foi excluída uma molécula do conjunto, construindo um novo modelo, assim predizendo o valor desta variável dependente do modelo original, ou seja, uma correlação entre a qualidade do modelo original e dos modelos aleatórios; (ii) aleatorização de Y (*Y scramble* ou *y-randomization*) (**Figura 18B**), em que foi avaliada a robustez (q^2) e o erro de predição (SEV) na probabilidade de ocorrer uma correlação dependente de cada variável no modelo, ou seja, parâmetros estatisticamente correlacionados a Y, mas sem uma relação de causa e efeito direta à atividade biológica; e (iii) *leave many out* (LMO) ou *leave-N-out* (**Figura 19**), no qual as amostras (moléculas) do conjunto treinamento foram divididas em subgrupos múltiplos de 5, avaliando a robustez do modelo em função de diferentes tamanhos de subgrupos de validação (KIRALJ; FERREIRA, 2009).

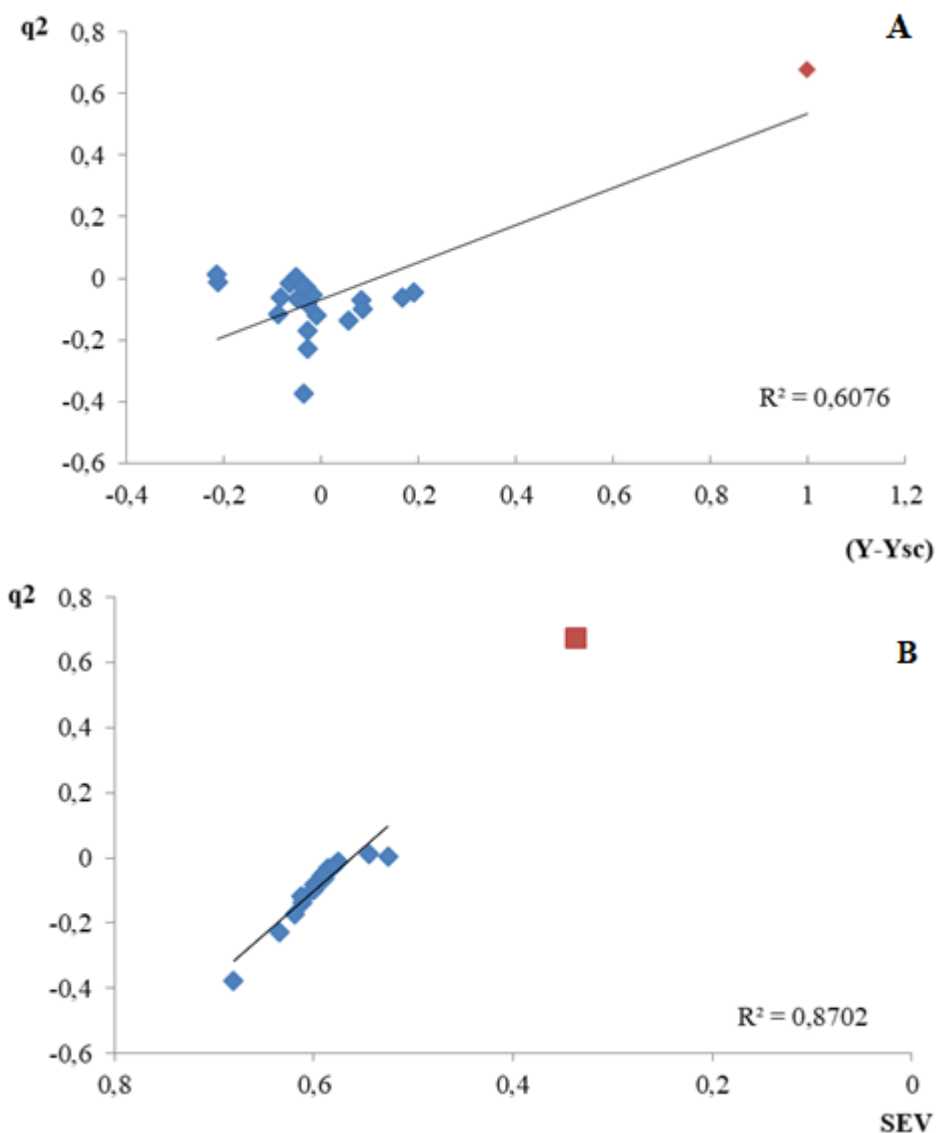


Figura 18. (A) Coeficiente de validação cruzada de Y. Teste de validação por exclusão de moléculas do grupo teste do modelo. Cada novo modelo (azul) prediz o efeito da remoção de uma variável (molécula excluída) comparado ao valor do modelo original (vermelho). Valor de correlação entre modelos: $r^2 = 0,608$. **(B) Aleatorização de Y (Y scramble).** Gráfico da probabilidade da correlação de variáveis (azul) no modelo selecionado, indicando a aleatoriedade do conjunto de moléculas no grupo teste em relação à sua atividade biológica, comparado ao valor do modelo original (vermelho). Valor da correlação entre modelos: $r^2 = 0,8702$.

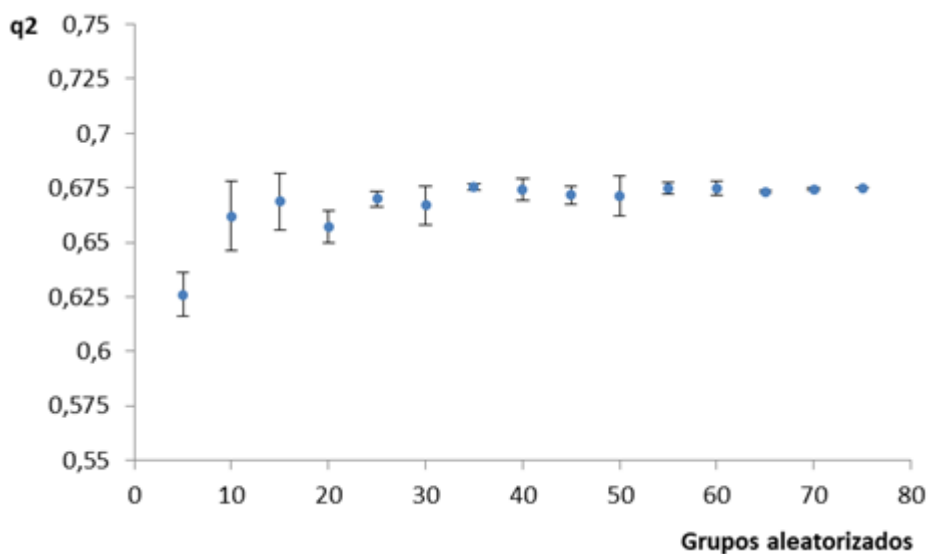


Figura 19. Leave many out (LMO). Correlação estatística da aleatorização das moléculas do grupo treinamento em grupos (múltiplos) de 5. Modelo realizado em triplicata, representando desvios-padrão das análises. Média geral das 15 aleatorizações = $0,6676 \pm 0,0057$.

Reunidos, os testes superam os requisitos mínimos ($q^2 > 0,5$ e $r^2 > 0,6$) propostos e permitem validar a correlação da qualidade do modelo selecionado (A, C, Ch de 7 a 11 átomos) com os modelos aleatórios, bem como sua confiabilidade de predição. Uma vez validado, o modelo pode ser testado, correlacionando, por exemplo, dados obtidos experimentalmente (resultados da atividade biológica), com dados preditos computacionalmente (valores calculados pelo grupo teste).

Foi observada uma correlação com perfil semelhante entre as frequências de distribuição do grupo treinamento comparado ao grupo teste (**Figura 20**). Esses resultados corroboram a similaridade entre os grupos treinamento e teste, mesmo aleatorizados, também testificando o modelo de HQSAR selecionado. Nesse histograma, a maior quantidade de erros de predição (eixo X) está próxima de zero, o que indica uma boa qualidade do modelo.

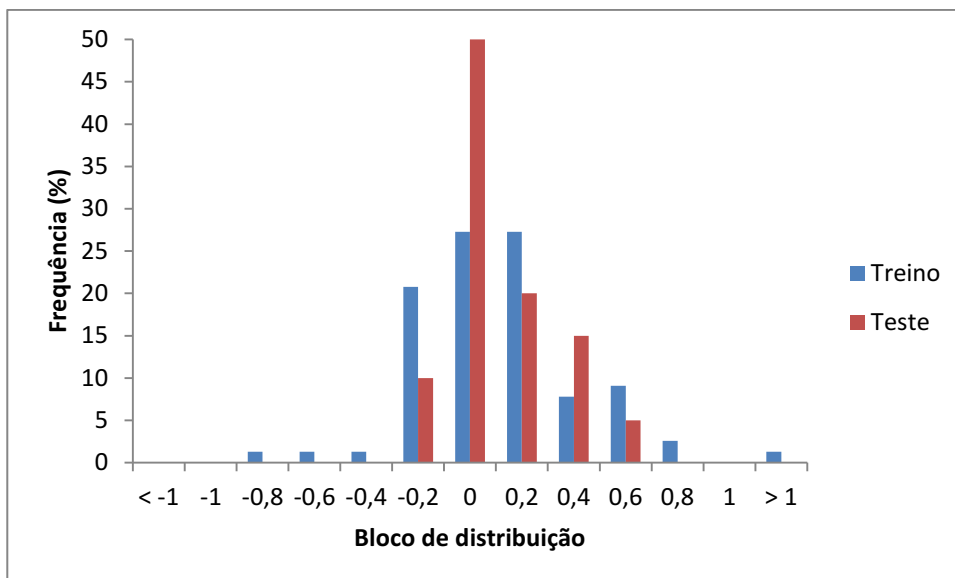


Figura 20. Gráfico de frequência de compostos por bloco de distribuição do modelo selecionado de HQSAR. Modelo comparativo da similaridade de distribuição das amostras dos grupos treinamento e teste.

5.1.4 Interpretação físico-química do modelo de HQSAR selecionado

Foram gerados os mapas de contribuição química das moléculas mais ativas da série no modelo ótimo de HQSAR: 62, 65, 63, 64 e 66 (**Figura 21**); e menos ativas, 101, 3, 105, 93 e 109 (**Figura 22**), juntamente aos respectivos valores calculados de potência da atividade (pIC_{50}) das moléculas avaliadas experimentalmente na protease de DENV-3, segundo Behnam et al. (2015) e Wiegel et al. (2015).

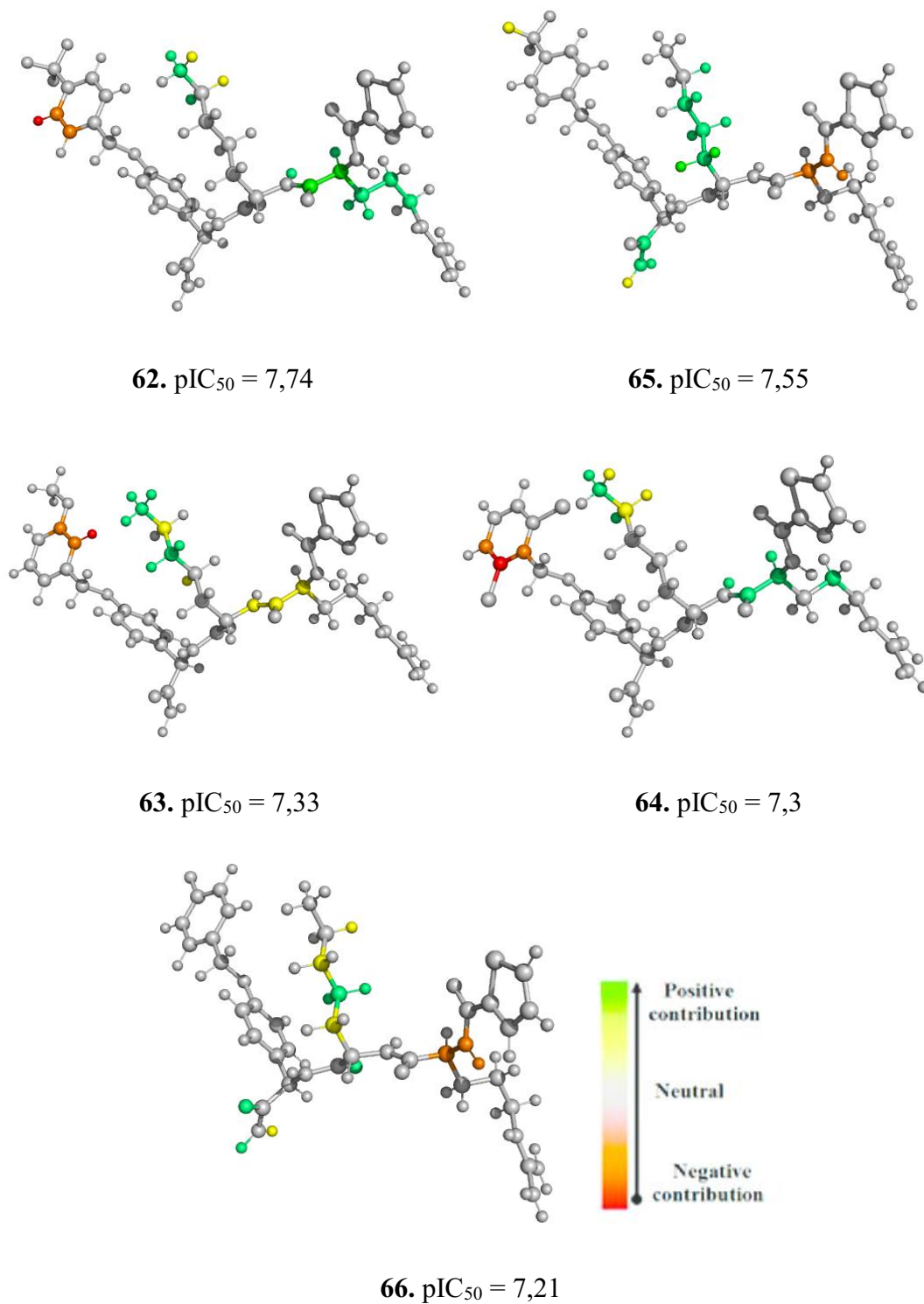
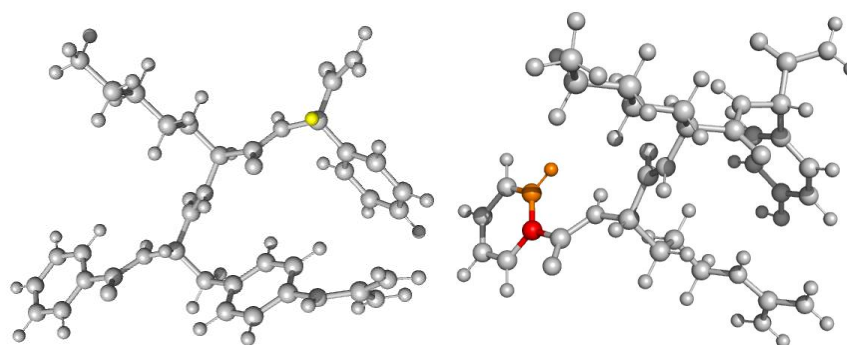
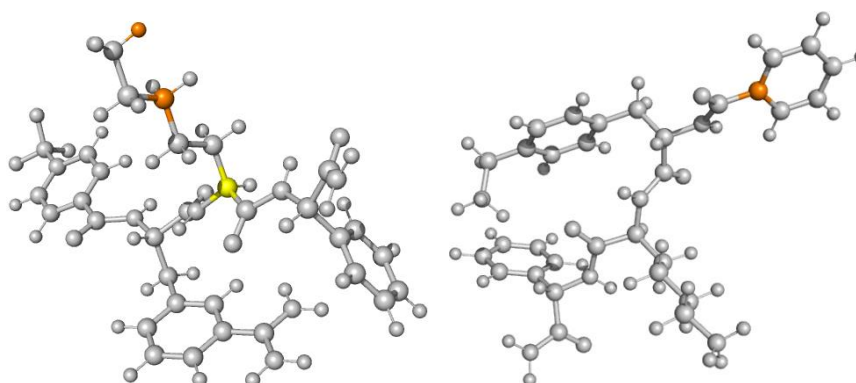


Figura 21. Mapas de contribuição química das cinco moléculas mais ativas (menores valores de IC_{50}) experimentalmente na protease DENV-3. Foram identificados fragmentos específicos de maior contribuição positiva nos compostos peptidomiméticos do conjunto de dados. Valores de IC_{50} calculados a pIC_{50} . Formatação condicional das estruturas em escala de cor: verde a vermelho (contribuição positiva a negativa, respectivamente). Estruturas em branco representam uma contribuição neutra.



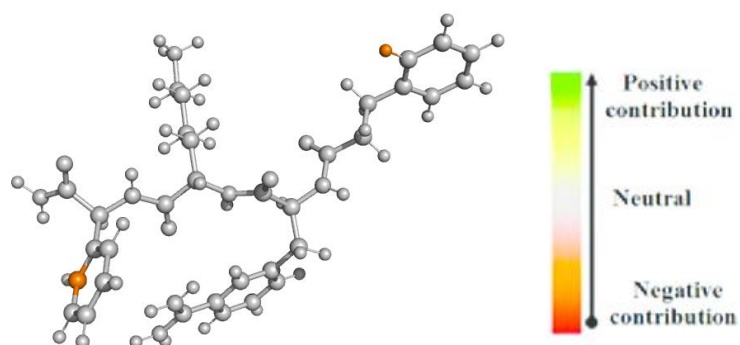
101. $pIC_{50} = 4,9$

3. $pIC_{50} = 5,01$



105. $pIC_{50} = 5,23$

109. $pIC_{50} = 5,39$



93. $pIC_{50} = 5,36$

Figura 22. Mapas de contribuição química das cinco moléculas menos ativas (maiores valores de IC_{50}) experimentalmente na protease DENV-3. Não foram identificados fragmentos de maior contribuição negativa nos compostos peptidomiméticos do conjunto de dados. Valores de IC_{50} calculados a pIC_{50} . Formatação condicional das estruturas em escala de cor: verde a vermelho (contribuição positiva a negativa, respectivamente). Estruturas em branco representam uma contribuição neutra.

De uma maneira geral, dentre os mapas de contribuição, pode se observar a importância de dois fragmentos na estrutura das moléculas: as cadeias laterais dos aa arginina e lisina (**Figura 23**). Estas informações estruturais serão utilizadas para o filtro na triagem virtual e construção do modelo de um farmacóforo, após triagem baseada em modificações dos parâmetros de biodisponibilidade oral na “regra dos 5” de Lipinski (Ro5) (Lipinski et al. 2001; Lipinski and Hopkins, 2004) e as propriedades das “regras de Veber” (Veber et al. 2002).

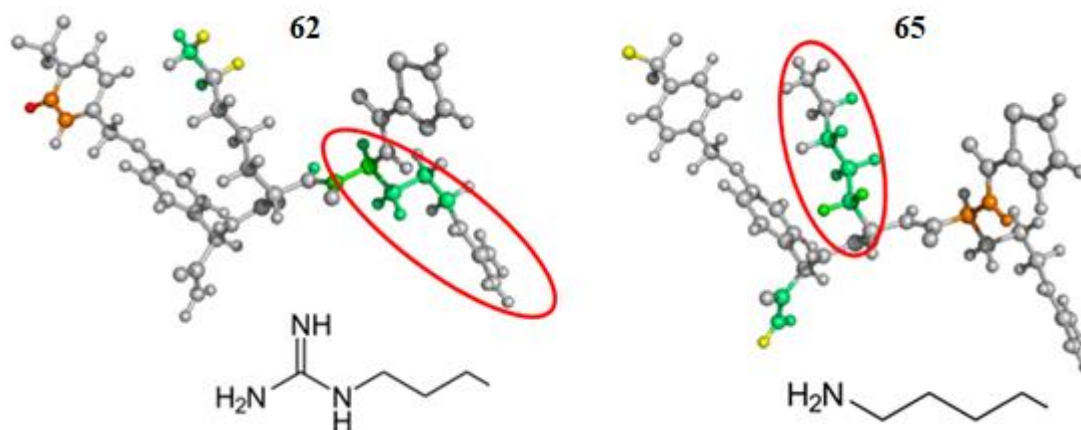


Figura 23. Mapas de contribuição química das moléculas mais ativas (62 e 65) e respectivos fragmentos moleculares de maior importância. Fragmentos de arginina (62) e lisina (65) assinalados em vermelho e representados em suas estruturas químicas.

5.1.5 Seleção da estrutura da protease (NS3pro) de DENV

Inicialmente, para DENV, quinze estruturas (2FOM, 2M9P, 2M9Q, 2VBC, 2WHX, 2WZQ, 3L6P, 3LKW, 3U1I, 3U1J, 4M9I, 4M9F, 4M9K, 4M9M e 4M9T) foram importadas e submetidas ao programa PyMOL, removendo moléculas de água e sulfatos. A sobreposição das estruturas permitiu evidenciar a similaridade conformacional dentre elas, o que também corrobora à conservação estrutural dentre as diferentes estruturas (**Figura 24**).

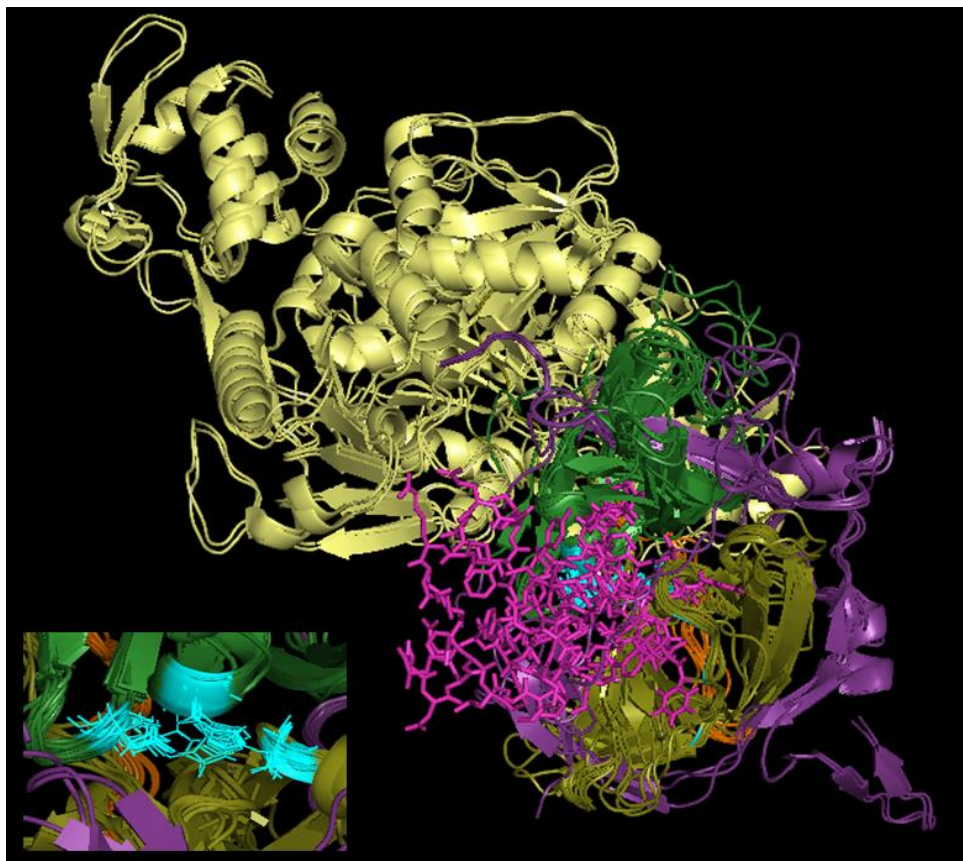


Figura 24. Visualização em modelo de *cartoon* da sobreposição das 15 estruturas da NS3 de DENV importadas do PDB. Porções NS3 C-terminal (verde); NS3 N-terminal (cobre); NS2B (roxo); alças *linker* (laranja); tríades (his51, asp75 e ser135) catalíticas (ciano); ligantes em modelo de *sticks* (magenta); e frações da helicase (pastel). No canto inferior esquerdo se encontra um inserto mostrando, em maior aumento, a sobreposição das regiões do sítio ativo (tríades catalíticas).

Posterior à análise das estruturas que atendessem os quatro critérios de eleição delimitados (**item 4.1.3**), duas estruturas da protease de DENV-3 foram selecionadas, sob os códigos do PDB 3U1I e 3U1J. Embora 3U1J possuísse a aprotinina como ligante, um inibidor conhecido de serina-proteases, a estrutura 3U1I (**Figura 25**) foi selecionada, uma vez que este cristal também foi utilizado como modelo nas análises computacionais nos trabalhos de Behnam (2015) e Weigel (2015) com os 111 compostos peptidomiméticos.

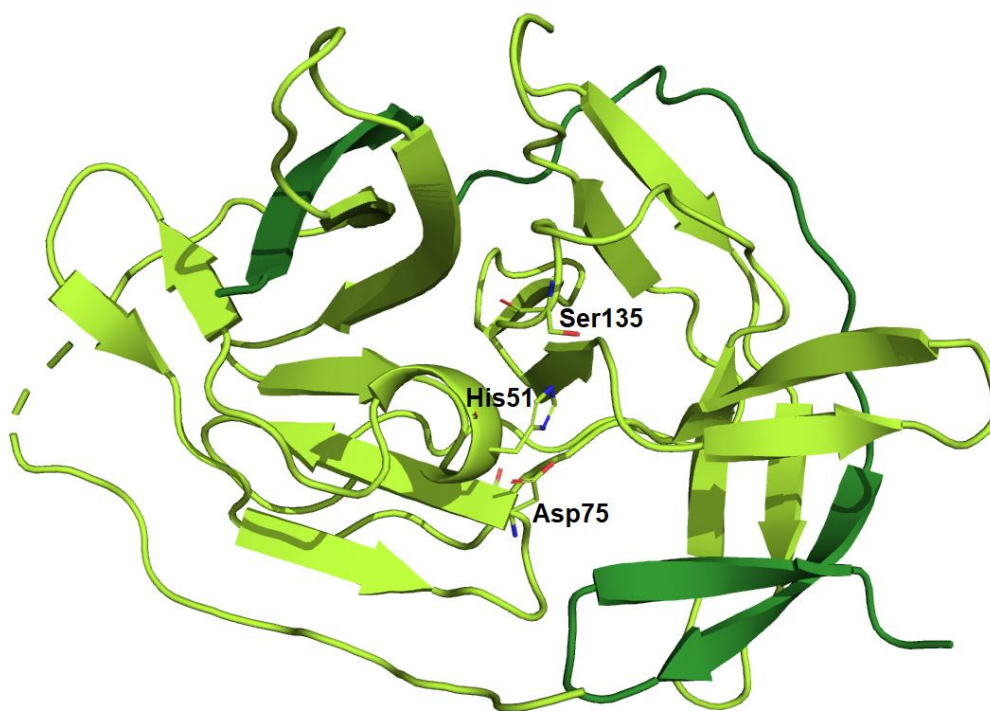


Figura 25. Estrutura tridimensional do complexo da protease NS2B-NS3 do dengue virus depositada no PDB sob o código 3U1I. NS2B (verde escuro), NS3 (verde claro) e alça *linker*. O sítio catalítico (tríade) da enzima é mostrado: aa histidina 51 (His51), aspartato 75 (Asp75) e serina 135 (Ser135).

5.1.6 Mapeamento da estrutura de DENV selecionada

Seguidamente, a predição e caracterização dos sítios de ligação da protease de DENV selecionada (3U1I) foi realizado nos servidores FTSite e FTMap. O servidor FTSite fez a predição de três possíveis sítios de ligação da protease (**Figura 26A**). Neste resultado é ressaltada a importância de observar as cavidades próximas ao sítio ativo da protease (tríade catalítica), em que um ligante poderia interagir para atuar como um efetivo inibidor, representadas pelos sítios 1 (rosa), 2 (laranja) e 3 (azul). O arquivo resultante do processo no FTMap, por sua vez, apresenta agrupamentos de moléculas-sonda divididos por sítios de ligação (**Figura 26B**), as quais interagem com os aa daquele sítio da protease delimitados no FTSite. Foram selecionados para análise somente os aa localizados nos sítios preditos pelo FTsite.

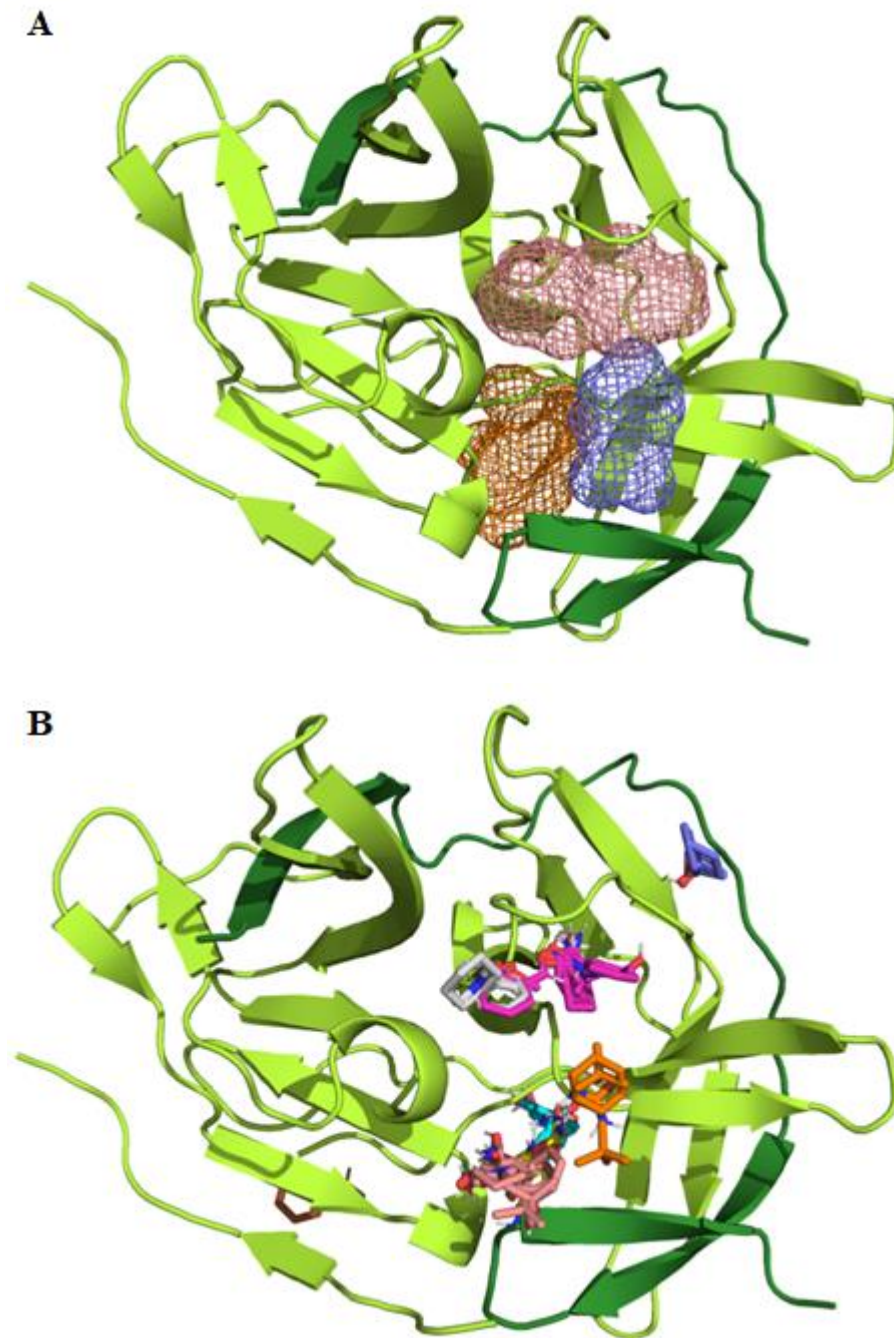


Figura 26. (A) Protease (NS2B-NS3) de DENV-3 e suas cavidades representando sítios de ligação. Predição dos sítios calculada pelo servidor FTSite. Sítio 1 ou S1 (rosa), sítio 2 ou S2 (laranja) e sítio 3 ou S3 (azul) representados em redes tridimensionais. **(B) Protease (NS2B-NS3) de DENV-3 e agrupamentos de moléculas-sonda.** Agrupamentos de moléculas calculado pelo servidor FTMap. Grupamentos 00 (ciano), 01 (magenta), 02 (amarelo), 03 (rosa), 04 (branco), 05 (roxo), 06 (laranja) e 07 (marrom).

Os grupamentos 01, 03, 04 e 06 estão localizados próximos aos sítios de ligação S1 e S3 identificados no resultado do FTSite. Juntamente aos resultados dos campos de interação molecular, os pontos de complementaridade molecular entre a protease e o peptídeo inibidor NDL (ligante de origem) (**Figura 27**) foram observados e apresentados por tipo de interação (**Figuras 28 a 31**).

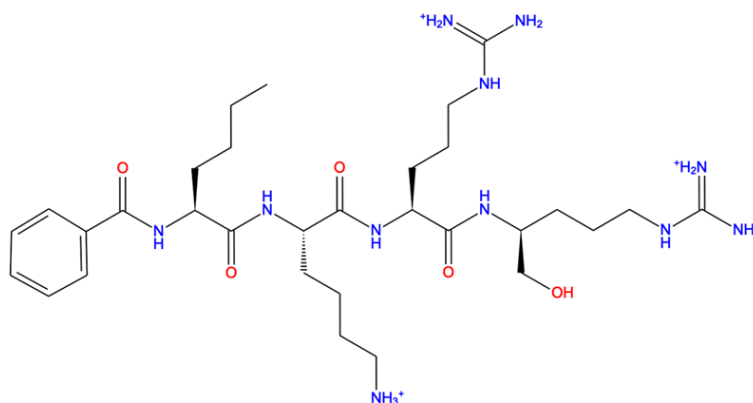


Figura 27. Peptídeo inibidor NDL (ligante de origem) da protease NS2B-NS3 de DENV-3. Estrutura molecular de N-benzoil-L-norleucil-6-amônio-L-norleucil-N-5-[amino(imino)metil]-N-[(2S)-5-carbamidamido-1-hidroxipentan-2-il]-L-oritinamida. Fonte: NOBLE et al. 2012

Os dados obtidos no FTSite para predição dos sítios de ligação para interaçõesceptoras de hidrogênio indicaram seis aa: um para porção NS2B, Met84; e cinco para porção NS3, Gly133, Thr134, Ser135, Gly153 e Tyr161; sendo Gly133 e Ser135 com maior frequência (34 e 37%) nas interações das moléculas-sonda (**Figura 28A**). Os resultados nos campos de interação molecular corroboram a essas observações, indicando uma interação do tipo acceptor (vermelho) próximo ao aa Ser135 (**Figura 28B**).

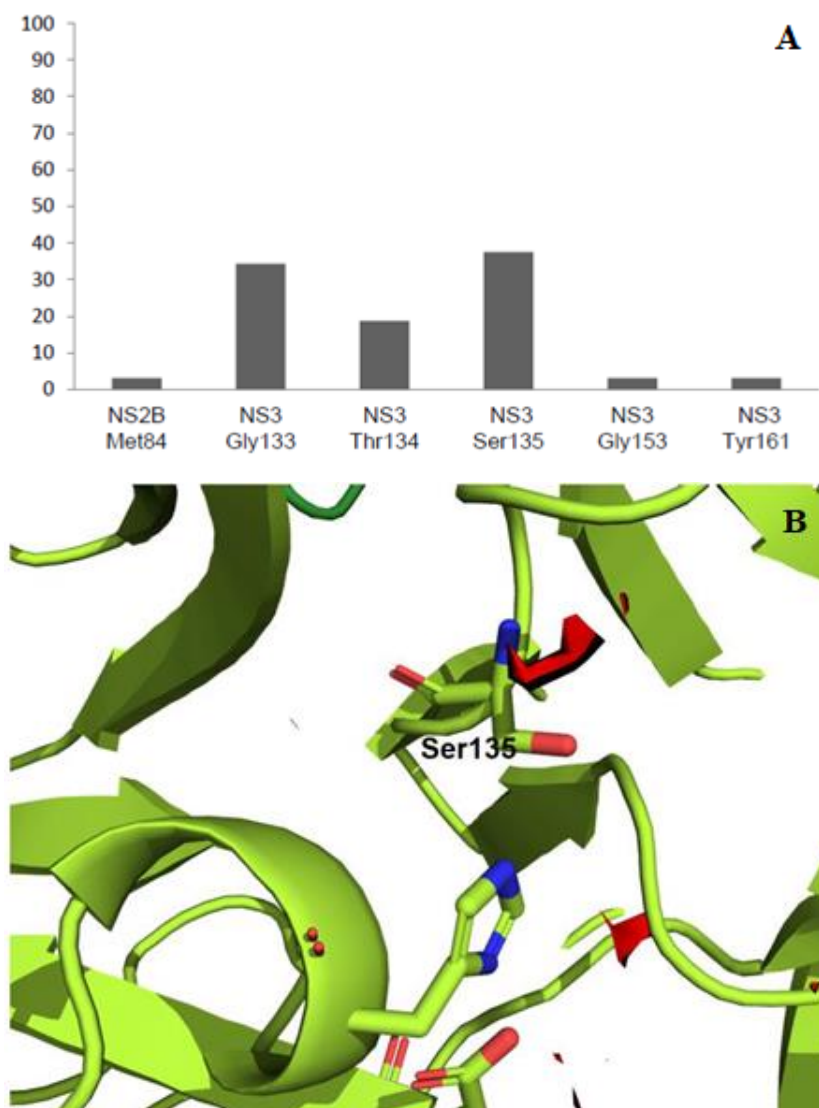


Figura 28. (A) Percentual de interações de hidrogênio (HBA) por aminoácidos (aa) no sítio de ligação da protease de DENV-3 (3U11). A estrutura da protease de DENV-3, obtida sob código de acesso 3U11 no PDB, foi submetida ao programa FTMap para predição dos sítios de ligação e as estruturas visualizadas utilizando o programa DiscoveryStudio. O gráfico apresenta Gly133 e Ser135 preditos como aa interagindo mais frequentemente com as moléculas-sonda. **(B) Campo de interação molecular do tipo aceptor de hidrogênio (vermelho) próximo ao aminoácido (aa) Ser135.** Os campos de interação do tipo aceptor de hidrogênio corroboram os dados de interações das sondas com o aa Ser135.

As predições para interações doadoras de hidrogênio, por sua vez, indicaram oito aa: um para porção NS2B, Asp81; e sete para porção NS3, Val36, Val72, Phe130, Lys131, Gly151, Asn152 e Gly153; sendo Asp81 com maior frequência (41%) nas interações das moléculas-sonda (**Figura 29A**). Os resultados nos campos de interação molecular também corroboram essas observações, indicando uma interação do tipo doador (azul) próximo ao aa Asp81 (**Figura 29B**).

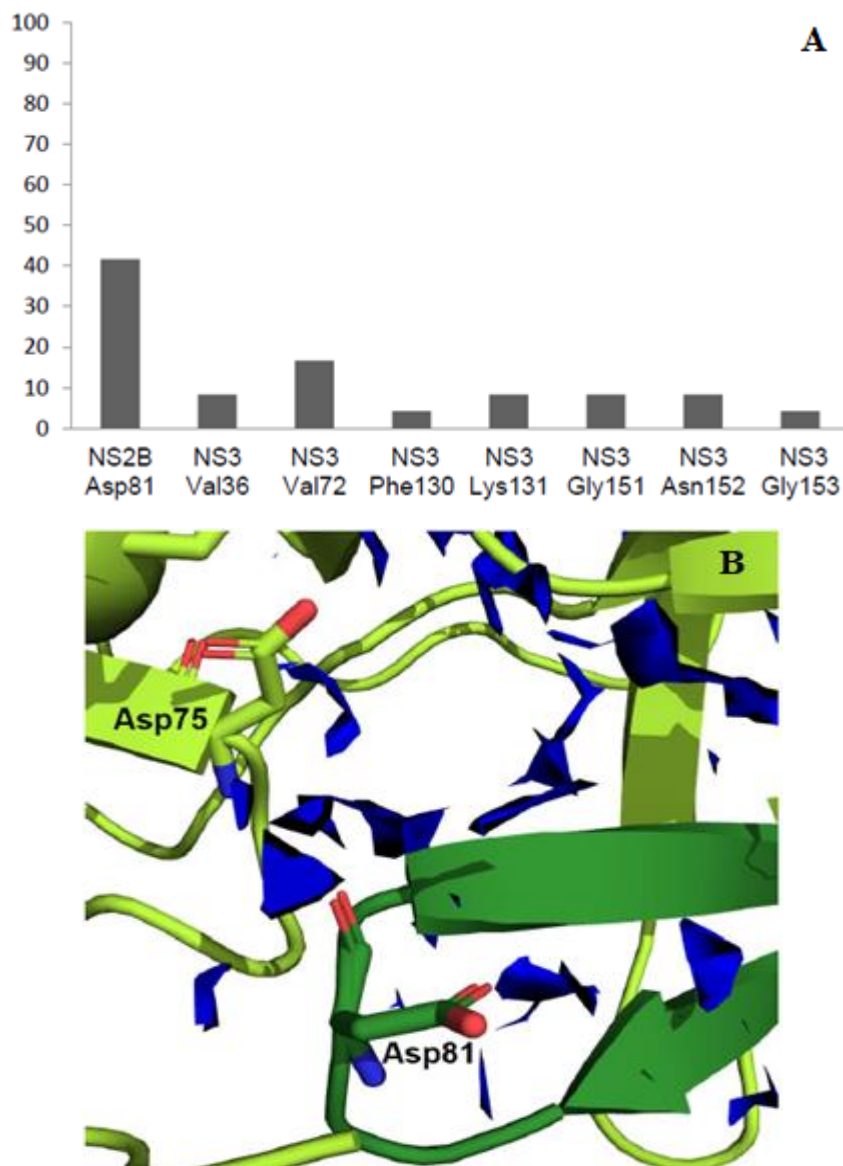


Figura 29. (A) Percentual de interações doadoras de hidrogênio (HBD) por aa no sítio de ligação da protease de DENV-3 (3U1I). A estrutura da protease de DENV-3, obtida sob código de acesso 3U1I no PBD, foi submetida ao programa FTMap para predição dos sítios de ligação e as estruturas visualizadas utilizando o programa DiscoveryStudio. O gráfico apresenta Asp81 predito como aa interagindo mais frequentemente com as moléculas-sonda. **(B) Campo de interação molecular do tipo doador de hidrogênio (azul) próximo aos aa Asp75 e Asp81.** Os campos de interação do tipo doador de hidrogênio corroboram os dados de interações das sondas com o aa Asp81. Também se encontra assinalado o aa Asp75.

As predições para interações eletrostáticas, por outro lado, indicaram somente Asp75 em interação com as moléculas-sonda na NS3 (**Figura 30A**). Os resultados nos campos de interação molecular também corroboram essa observação, indicando uma interação do tipo iônica (amarelo) próximo ao aa Asp75 (**Figura 30B**).

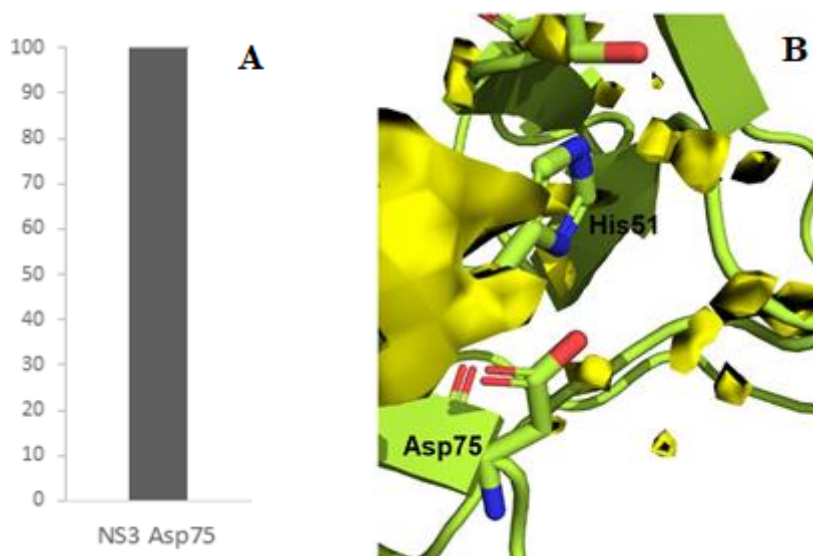


Figura 30. (A) Percentual de interações eletrostáticas (iônicas) por aa no sítio de ligação da protease de DENV-3 (3U1I). A estrutura da protease de DENV-3, obtida sob código de acesso 3U1I no PBD, foi submetida ao programa FTMap para predição dos sítios de ligação e as estruturas visualizadas utilizando o programa DiscoveryStudio. O gráfico apresenta Asp75 predito como único aa interagindo com as moléculas-sonda. **(B) Campo de interação molecular do tipo eletrostática (amarelo) próximo a Asp75.** Os campos de interação do tipo eletrostática corroboram os dados de interações das sondas com o aa Asp75. Também se encontra assinalado outro aa da tríade catalítica, His51.

Por último, a predição das interações de van der Waals indicaram seis aa em interação com as moléculas-sonda, todos na porção da protease: Val36, His51, Val52, Pro132, Tyr150 e Tyr161 **(Figura 31A)**. Os resultados nos campos de interação molecular também corroboram essa observação, indicando uma interação do tipo hidrofóbica (laranja) próximo ao aa Val52 **(Figura 31B)**.

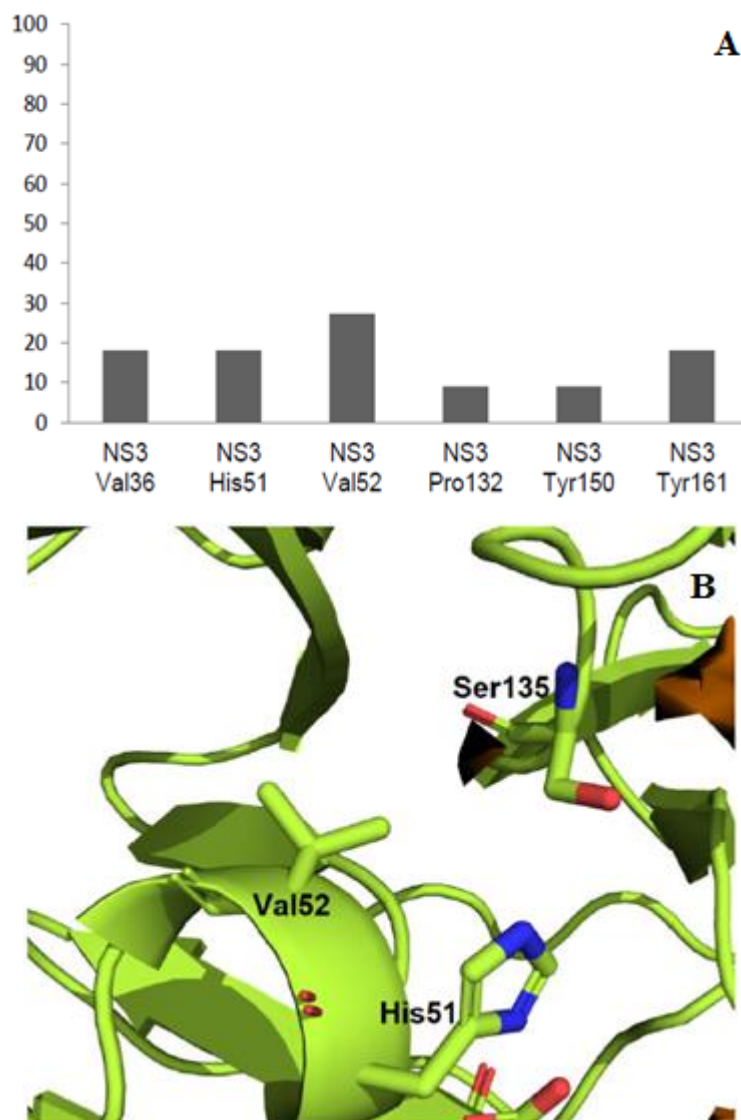


Figura 31. (A) Percentual de interações de van der Waals (hidrofóbicas) por aa no sítio de ligação da protease de DENV-3 (3U1I). O gráfico apresenta Val52 predito como aa interagindo mais frequentemente com as moléculas-sonda. **(B) Campo de interação molecular do tipo hidrofóbica (laranja) próximo a Val52.** Os campos de interação do tipo hidrofóbico corroboram os dados de orientação das interações das sondas com o aa Val52. Assinalados também se encontram os resíduos da tríade Ser135 e His51.

Os dados de interação de cada sonda com os aa, bem como os campos de interação apresentam sítios e aa potencialmente importantes para as interações entre uma molécula (ligante) e a protease (alvo). Aguilera-Pesantes et al. (2017), por exemplo, demonstraram os aa His51, Val52, Asp75 e Ser135 interagindo com diversas moléculas, entre os sítios S1 e S3 da protease. Nesse trabalho também foram compilados diversos aa de interesse já identificados, da literatura, em interações com ligantes da protease de DENV. Dentre os 22 aa identificados, a maioria das moléculas em interação também foram preditas em associação aos aa da tríade catalítica, His51, Asp75 ou Ser135 (AGUILERA-PESANTES et al., 2017).

Streuer et al. (2011), identificaram compostos cetoamida especialmente interagindo ao bolso S1 da protease de DENV. Nesse estudo, as moléculas interagem especialmente com o oxigênio de Ser135, formando diferentes confôrmeros quirais na interação. Dentre os melhores ranqueados, foi observada uma ligação covalente com Ser135 em todas as conformações, além de interações frequentes com His51 e Tyr161, os quais também foram identificados interagindo com sonda nos resultados apresentados nas interações de van der Waals (**Figura 31**). Também em concordância com o trabalho de Aguilera-Pesantes et al. (2017), não foram observadas interações no sítio S2 da protease (STEUER et al., 2011; AGUILERA-PESANTES et al., 2017).

Esses dados, em conjunto, permitem obter informações importantes acerca de regiões de ligação nos sítios da protease, bem como dos aa de maior relevância ou frequência de interações com ligantes como inibidores ou potenciais moléculas com atividade na protease. Juntamente aos resultados do HQSAR, essas informações foram utilizadas na construção do modelo farmacofórico para triagem virtual de moléculas com potencial atividade em DENV.

5.1.7 Seleção e mapeamento da estrutura da protease de ZIKV

Posterior à análise da estrutura da NS3 de DENV-3 (3U1I), o mesmo foi realizado para protease de ZIKV (PDB: 5YOF), buscando comparar os dados em DENV e aprimorar a construção do farmacóforo. A estrutura da protease de ZIKV, juntamente com a posição dos aa da tríade catalítica (His51, Asp75 ou Ser135) é mostrada na figura abaixo (**Figura 32**).

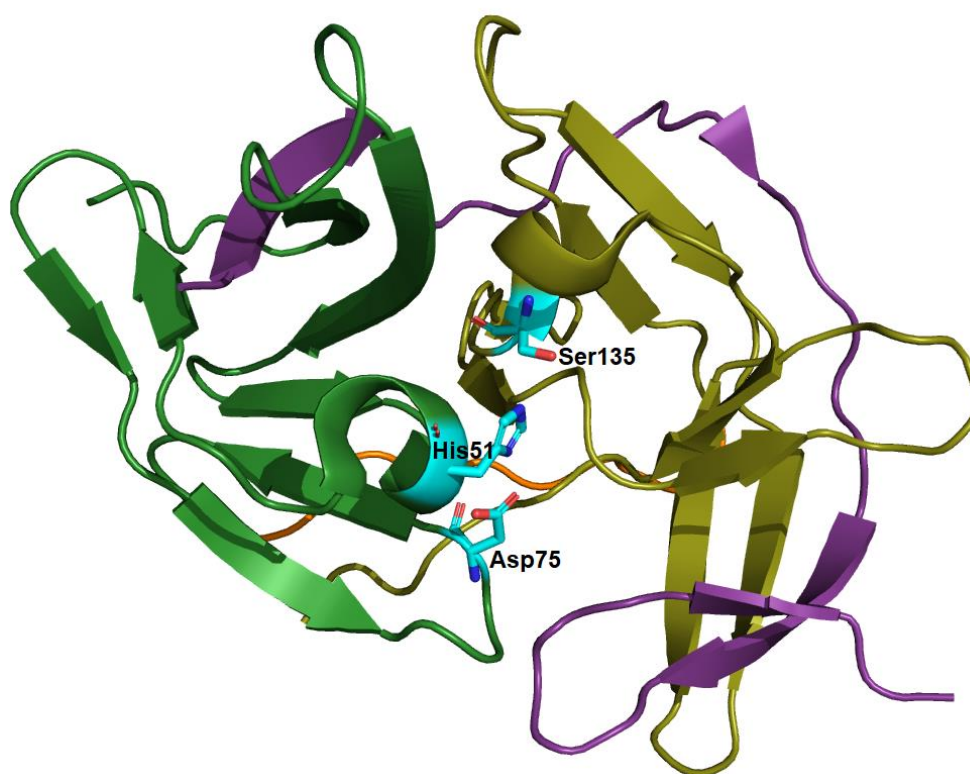


Figura 32. Estrutura tridimensional do complexo da protease NS2B-NS3 de ZIKV depositada no PDB sob o código 5YOF. NS2B (roxo), NS3pro N-terminal (verde), NS3pro C-terminal (oliva), alça *linker* (laranja) e tríade catalítica (ciano) composta pelos aa His51, Asp 75 e Ser135.

O servidor FTSite fez a predição de três possíveis sítios de ligação da protease (**Figura 33A**). Assim como em DENV, se ressalta a importância em observar quais cavidades se encontram próximas ao sítio ativo da protease (tríade catalítica), em que um ligante poderia interagir para atuar como um efetivo inibidor, agora representados pelos sítios 1 (rosa), 2 (verde) e 3 (azul). O arquivo resultante do FTMap, por sua vez, apresenta agrupamentos de moléculas-sonda divididos por sítios de ligação (**Figura 33B**), as quais interagem com os aa daquele sítio da protease delimitados no FTSite. Foram selecionados para análise somente os aa localizados nos sítios preditos pelo FTsite.

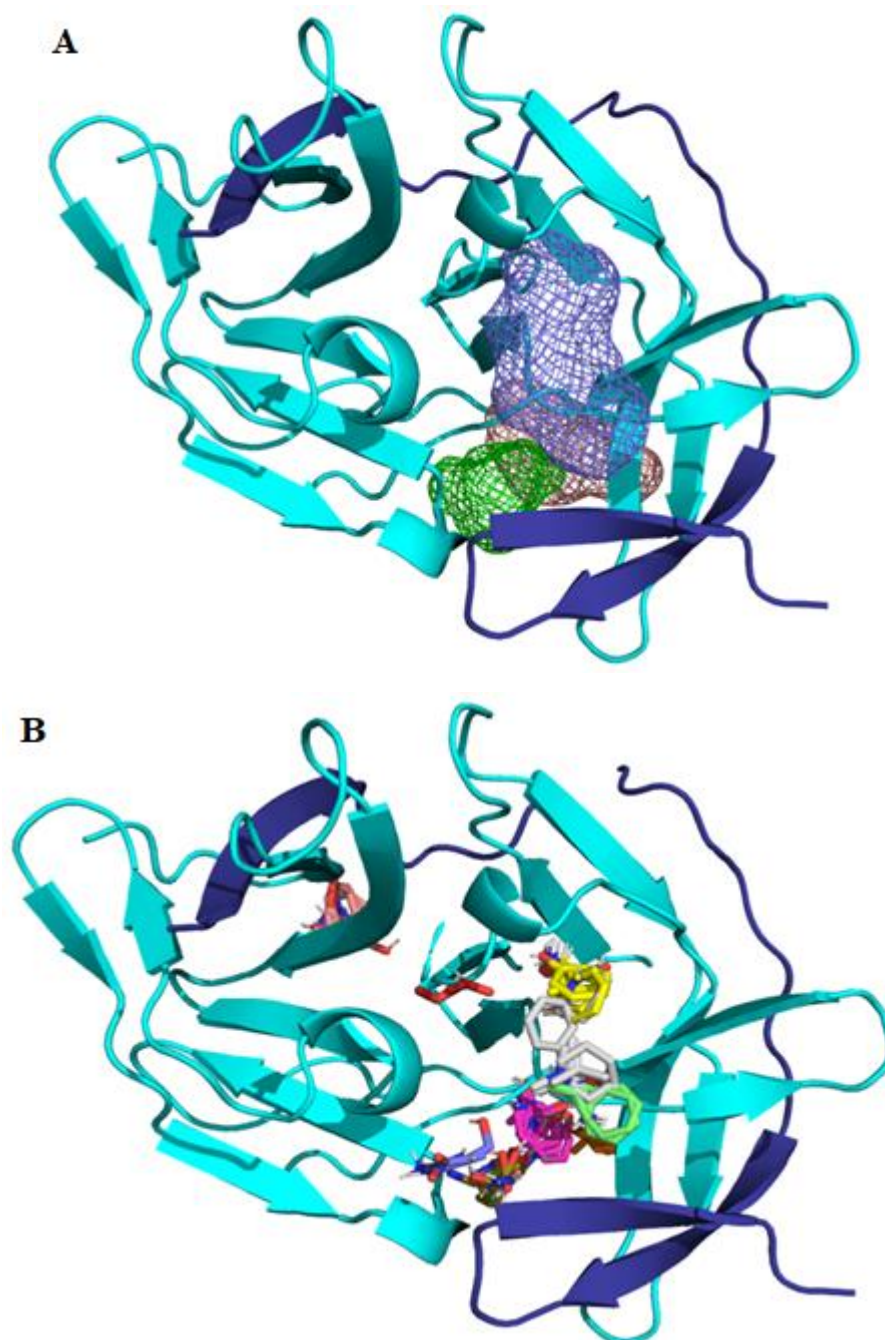


Figura 33. (A) Protease NS2B-NS3 de ZIKV e suas cavidades representando sítios de ligação. Predição dos sítios calculada pelo servidor FTSite. Sítio 1 ou S1 (rosa), sítio 2 ou S2 (verde) e sítio 3 ou S3 (azul) representados em redes tridimensionais. **(B) Protease NS2B-NS3 de ZIKV e grupamentos de moléculas-sonda.** Agrupamentos de moléculas calculado pelo servidor FTMap. Grupamentos 00 (oliva), 01 (magenta), 02 (amarelo), 03 (rosa), 04 (branco), 05 (verde), 06 (laranja), 07 (lilás) e 08 (vermelho).

Os grupamentos 01, 02, 04 e 05, se localizam próximos aos sítios de ligação S2 e S3, identificados no resultado do FTSite. Na análise da interação destas sondas com os aa da protease, também foram identificados quais aa interagem com determinada molécula-sonda e a quantidade de ligações presentes pela sonda, divididas em acceptor de ligação de hidrogênio

(HBA), doador de ligação de hidrogênio (HBD), interação eletrostática (iônica) e interação de van der Waals (hidrofóbica). Juntamente aos resultados dos campos de interação molecular, os pontos de complementaridade molecular entre a protease e o ligante de origem (**Figura 34**) também foram observados e apresentados por tipo de interação (**Figuras 35 a 38**).

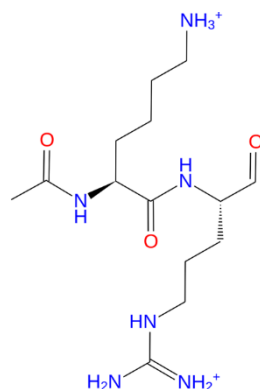


Figura 34. Dipeptídeo inibidor da protease (5YOF) de ZIKV. Estrutura molecular de (S)-2-acetamido-6-amino-N-((S)-5-guanidino-1-oxopentan-2-il)hexanamida. Fonte: LI et al. 2018.

Os dados obtidos no FTSite para predição dos sítios de ligação para interações aceptoras de hidrogênio indicaram nove aa na protease: Gln74, Thr134, Ser135, Tyr150, Asn152, Gly153, Tyr161, Thr166 e Gln167; sendo Tyr161 com maior frequência (40%) nas interações das moléculas-sonda (**Figura 35**). Dentre eles, quatro dos seis aa preditos em DENV foram identificados em ZIKV (Thr134, Ser135, Gly153 e Tyr161).

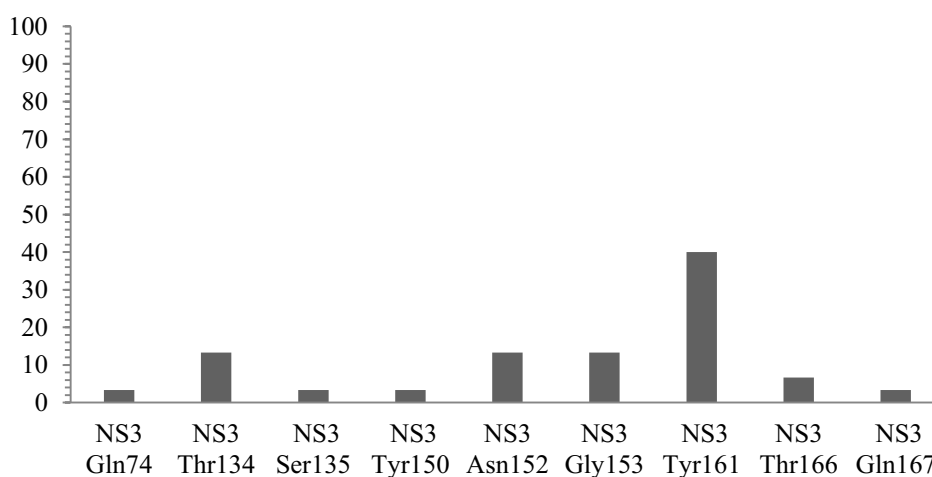


Figura 35. Percentual de interações de hidrogênio do tipo aceptor (HBA) por aa no sítio de ligação da protease de ZIKV (5YOF). O gráfico apresenta Tyr161 predito interagindo mais frequentemente com as moléculas-sonda.

As predições para interações doadoras de hidrogênio, por sua vez, indicaram doze aa na

porção da protease: Val72, Lys73, Gln74, Asp75, Pro131, Thr134, Gly151, Asn152, Gly153, Ala164, Ile165 e Gln167; sendo Pro131 e Gly151 com maiores frequências (18 e 20%) nas interações das moléculas-sonda (**Figura 36**). Quatro destes aa foram os mesmos preditos na protease de DENV (Val72, Gly151, Asn152 e Gly153), além de uma quinta posição idêntica (131), porém com aa diferente (Gly>Pro).

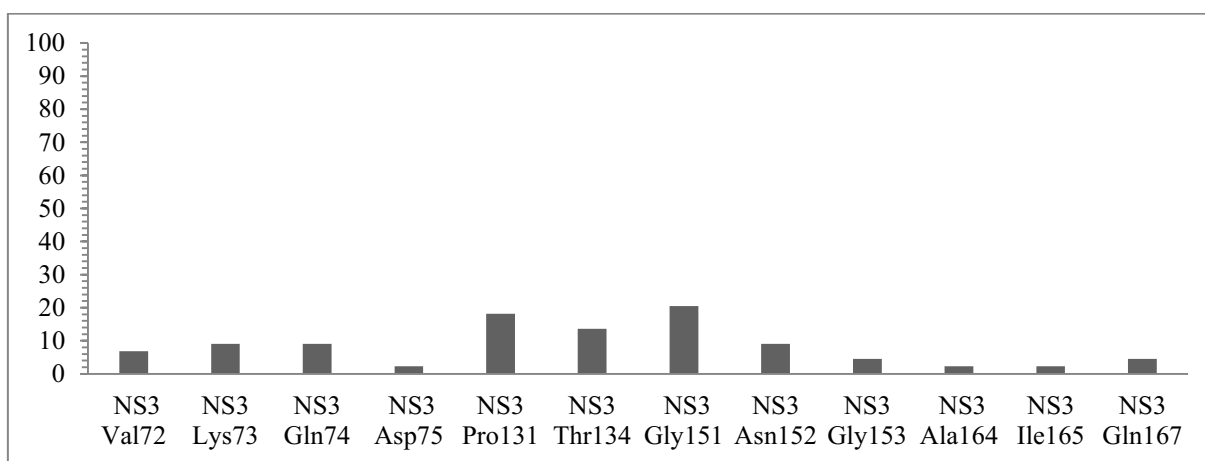


Figura 36. Percentual de interações de hidrogênio do tipo doador (HBD) por aa no sítio de ligação da protease de ZIKV (5YOF). O gráfico apresenta Pro131 e Gly151 preditos interagindo mais frequentemente com as moléculas-sonda.

As predições para interações eletrostáticas em ZIKV, por outro lado, indicaram cinco aa interagindo com as moléculas-sonda: um para porção NS2B, Asp83; e quatro para NS3pro, His51, Asp75, Pro131 e Tyr161; sendo Asp75 com maior frequência (33%) de interações (**Figura 37**). Destes, somente Asp75, único aa com interação eletrostática em DENV, foi igualmente predito em ZIKV.

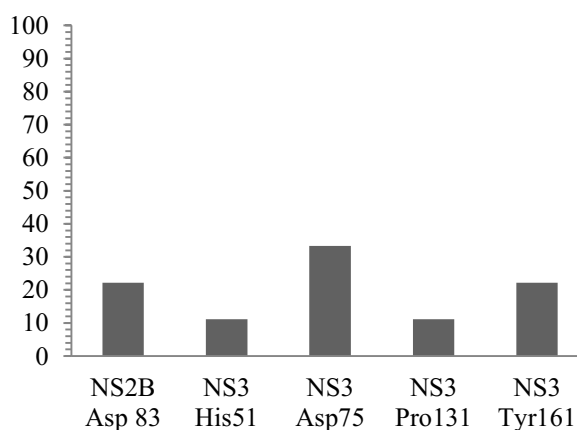


Figura 37. Percentual de interações eletrostáticas (iônicas) por aa no sítio de ligação da protease de ZIKV (5YOF). O gráfico apresenta Asp75 predito interagindo mais frequentemente com as moléculas-sonda.

Por último, a predição das interações de van der Waals indicaram nove aa em interação com as moléculas-sonda: dois para NS2B, Leu78 e Phe84; e sete para NS3pro, His51, Ile123, Pro131, Ala132, Tyr150, Tyr161 e Ala164; sendo Ile126 e Tyr161 com maior frequência (28 e 31%) de interações (**Figura 38**). Destes, His51, Tyr150 e Tyr161 foram igualmente preditos em DENV comparado a ZIKV.

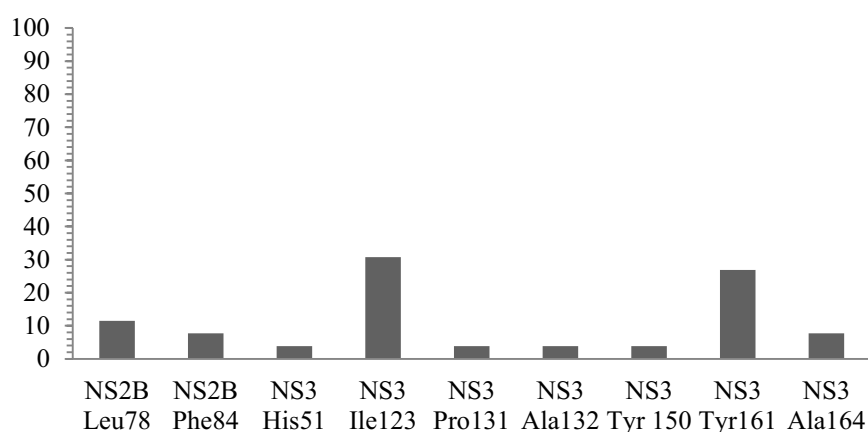


Figura 38. Percentual de interações de van der Waals (hidrofóbicas) por aa no sítio de ligação da protease de ZIKV (5YOF). O gráfico apresenta Ile123 e Tyr161 preditos como aa principais interagindo com as moléculas-sonda.

5.1.8 Comparação entre as estruturas de DENV e ZIKV

Comparativamente, Ser135 (DENV) e Tyr161 (ZIKV) apresentaram maior porcentagem de interação como ligação de hidrogênio do tipo acceptor, identificadas com a presença do campo de interação molecular das sondas. Embora distintos, a região de interação em ambos os aa é similar, o que configura, estruturalmente, um sítio de ligação de hidrogênio equivalente em ambas as proteases (**Figura 39**). Isso representa uma característica importante, por exemplo, para a construção de um modelo farmacofórico único para ambos os alvos em uma triagem virtual (YANG, 2010).

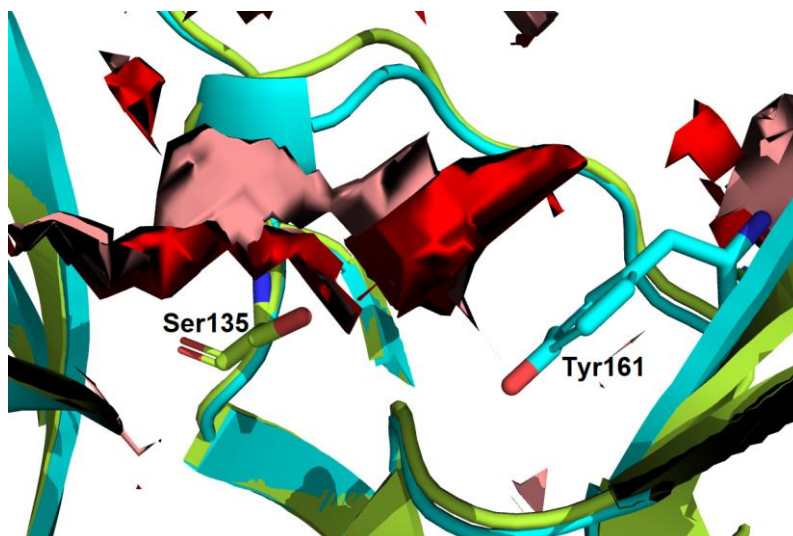


Figura 39. Campos de interação de ligação de hidrogênio do tipo aceptor nas proteases de DENV-3 (3U1I) e ZIKV (5YOF). Sítio de ligação predito da protease de ZIKV (ciano) e DENV (verde) sobrepostas. Campos de interação de ligação de hidrogênio com os aa Ser135 (rosa) e Tyr161 (vermelho).

Uma mesma observação pode ser feita para Asp75 em DENV e ZIKV. Este aa apresentou maior porcentagem de interação iônica predita com as sondas nas duas proteases, visualizado na presença do campo de interação eletrostática localizado próximo a ambos os aa nas duas estruturas (**Figura 40**). Igualmente, durante o planejamento de um farmacóforo único, estes sítios também podem ser considerados como *hot spots* ou pontos de acesso de ligação e, neste caso, equivalentes em ambas as estruturas (BARILLARI; MARCOU; ROGNAN, 2008).

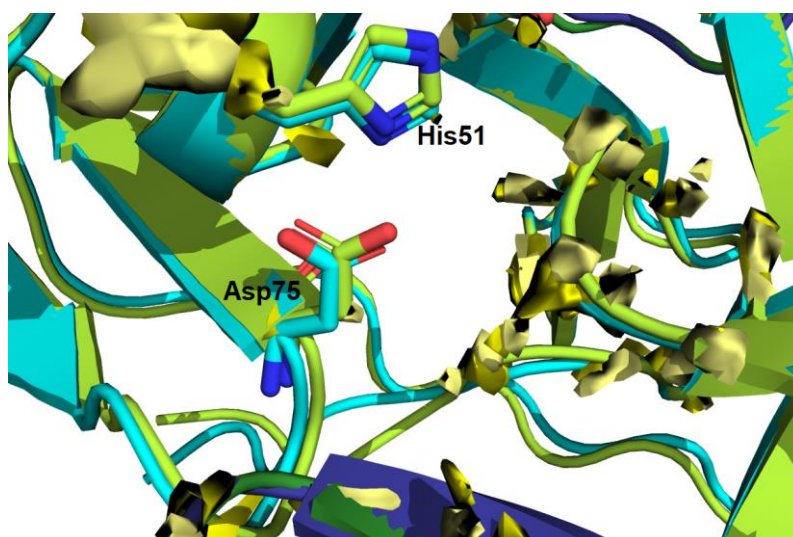


Figura 40. Campos de interação do tipo eletrostático nas proteases de DENV-3 (3U1I) e ZIKV (5YOF). Estrutura do sítio de ligação predito da protease NS2B-NS3 de ZIKV (ciano) e estrutura da mesma região do sítio de ligação da protease NS2B-NS3 de DENV (verde). Campo de interação com o aa Asp75 da protease de DENV (amarelo claro) e campo de interação com o aa Asp75 da protease de ZIKV (amarelo escuro).

Adicionalmente, esses dados reforçam a escolha das proteases de DENV e ZIKV como alvos moleculares de interesse, haja vista sua similaridade entre os flavivírus e papel essencial na multiplicação destes vírus (NEUFELDT et al., 2018). Além disso, corroborando os dados das análises descritas, também foi realizada uma comparação das características estruturais das proteases de DENV e ZIKV, empregando análises das coordenadas por cálculos de MIF.

As análises das coordenadas dos centros de cada uma das caixas (*grids*), utilizada para os cálculos de MIF, indicam, além de uma semelhança estrutural, uma correspondência dos sítios de ligação das proteases. Foram observadas regiões favoráveis em comum, haja vista a sobreposição das estruturas (**Figura 41**), corroboradas pela proximidade das coordenadas nos eixos X, Y e Z (**Tabela 14**). Esses resultados corroboram também para os pontos de acesso (*hot spots*) equivalentes em ambas proteases, tornando possível a construção de um único modelo farmacofórico e conveniente o planejamento de um único inibidor para ambas proteases.

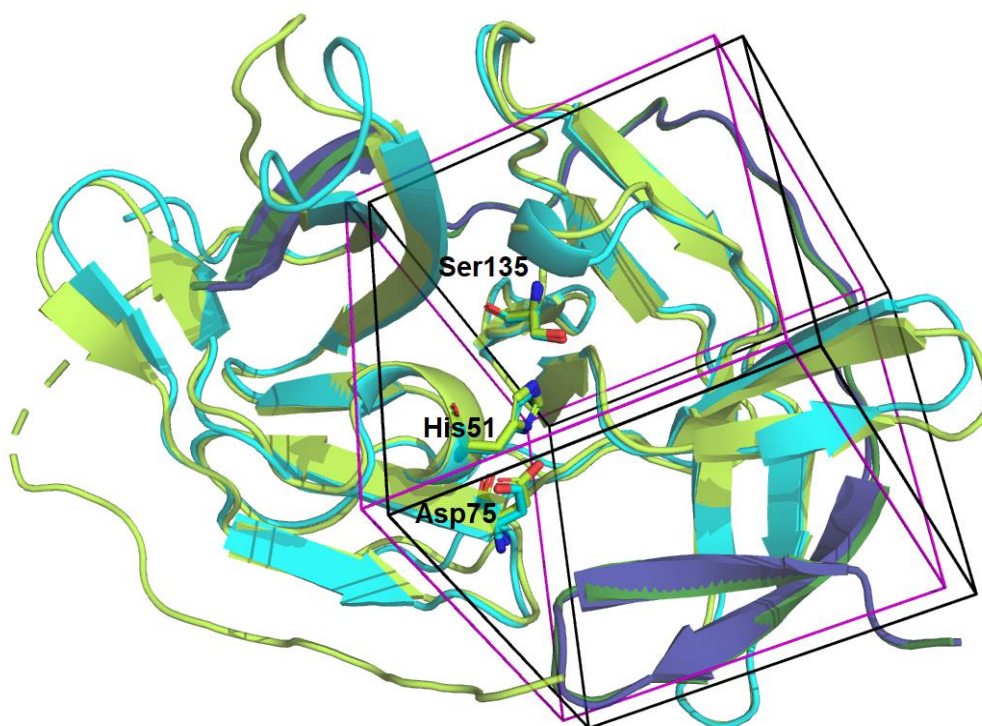


Figura 41. Sobreposição das caixas tridimensionais utilizadas nos cálculos de campo de interação molecular (MIF) das proteases de DENV (3U1I) e ZIKV (5YOF). Protease de DENV (verde-lima); protease de ZIKV (ciano); *grid* referente à protease de DENV (preto); e *grid* referente à protease de ZIKV (magenta).

Tabela 14. Coordenadas dos centróides dos *grids* utilizados nos cálculos de campo de interação molecular (MIF) de DENV e ZIKV.

	DENV	ZIKV
Eixo X	23,186	23,200
Eixo Y	-20,512	-19,206
Eixo Z	8,114	9,057

5.1.9 Análises de acoplamento molecular

Juntamente aos dados obtidos dos modelos de HQSAR e do mapeamento dos sítios de ligação das estruturas, o acoplamento molecular (*docking*) do ligante de origem (NDL) da protease em estudo (3U1I), bem como dos compostos mais ativos do conjunto de dados utilizado (62-66), além do *docking* do trabalho Behnam et al. (2015), foram realizados como estratégia para evidenciar quais aa apresentavam interação mais frequente na protease alvo. Posteriormente, estes ligantes também foram submetidos à dinâmica molecular, buscando complementar as análises de *docking*.

O trabalho de Behnam et al. (2015), além de base para o conjunto de dados deste trabalho, forneceu informações importantes em relação às interações dos compostos peptidomiméticos na protease de DENV (3U1I). A estratégia empregada foi baseada nos fragmentos de estruturas, direcionando a síntese e modificações de compostos semi-peptídicos, a fim de se obter moléculas com atividade inibitória nas proteases de DENV e WNV. Essa estratégia também já obteve sucesso em outros estudos de inibidores de proteases virais, como no desenvolvimento de fármacos inibidores das proteases de HCV, telaprevir e boceprevir, além da protease de HIV, como ritonavir e lopinavir (DE CLERCQ, 2009; AU; POCKROS, 2014).

Empregando a síntese e avaliações biológicas de potenciais inibidores, éteres benzílicos de 4-hidroxifenilglicina foram usados como esqueletos não-peptídicos para uma síntese em três etapas: (i) extensão de um fragmento C-terminal de éter-benziloxi; (ii) otimização C- e N-terminal; e (iii) mistura ou fusão de fragmentos com diferentes grupos funcionais substituintes (**Apêndice A**). Esta gama de modificações estruturais dos compostos permitiu obter, por exemplo, moléculas com atividade na faixa de nanomolar (moléculas 62, $IC_{50} = 18$ nM e 65, $IC_{50} = 28$ nM) (BEHNAM, 2015).

Sabendo disso, neste trabalho foi realizada uma reprodução e comparação do *docking* da molécula 62 no programa Surflex Docking (**Figura 42**), com a estrutura da protease (3U1I)

importada do PDB. A reprodução do acoplamento molecular da pose de maior *score* da molécula 62 resultaram na identificação de nove aa interagindo com essa conformação do ligante avaliado: Val36, His51, Val52, Asp75, Asp129, Phe130, Gly133, Tyr150 e Gly151.

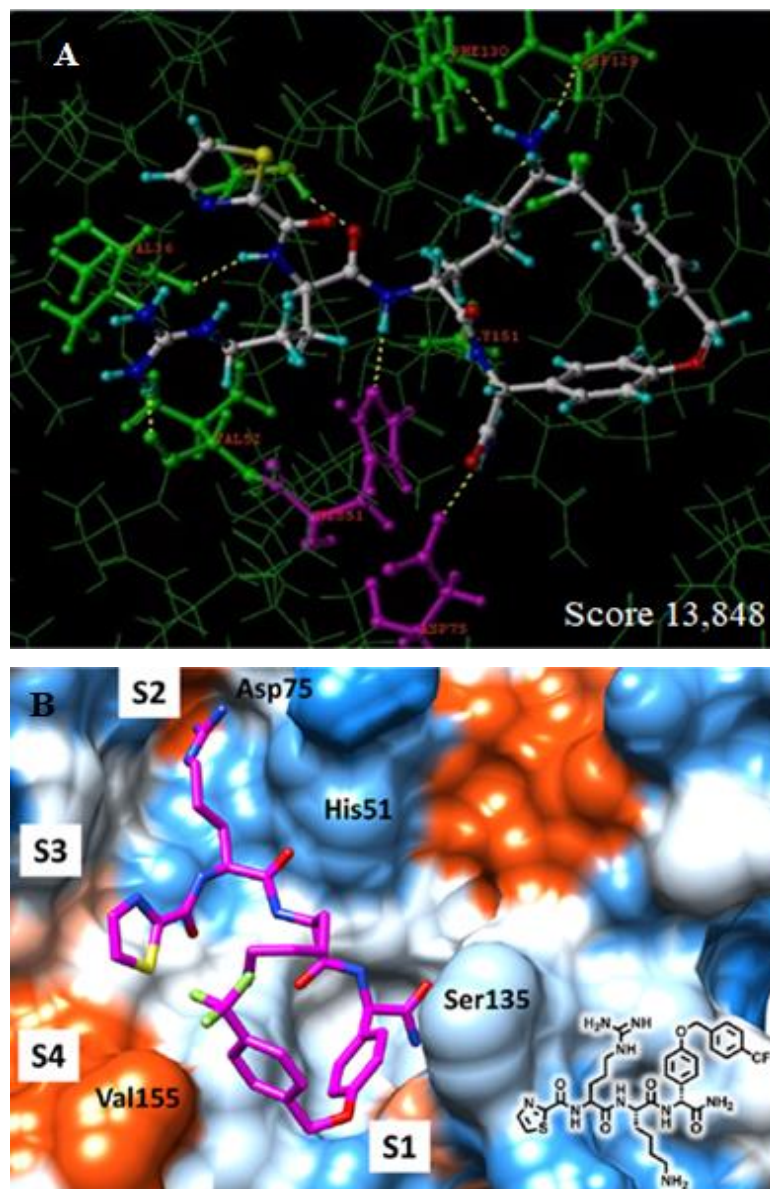


Figura 42. (A) Análise comparativa do *docking* da molécula 62 na protease de DENV-3 (3U1I). (A) Reprodução do acoplamento molecular utilizando programa Surflex Docking. Molécula 62 interagindo com aa Val36, His51, Val52, Asp75, Asp129, Phe130, Gly133, Tyr150 e Gly151. Estrutura da protease em modelo se *sticks* (verde), estrutura do ligante: átomos de carbono (branco), nitrogênio (azul), oxigênio (vermelho), flúor (verde-lima), enxofre (amarelo) e tracejados de interações moleculares (amarelo). Pontuação (*score*) = 13,848. (B) *Docking* do composto 62 na protease (3U1I) de DENV-3. Acoplamento molecular utilizando programa AutoDock Vina. Superfície da proteína delimitada em regiões hidrofílicas (azul) e hidrofóbicas (laranja). Ligante em representação de *sticks* (magenta). Sítios de ligação 1 a 4 (S1 a 4) e aa relevantes na interação com o composto (His51, Asp75, Ser135 e Val155) foram assinalados. Fonte: BEHNAM et al., 2015. (*adaptado*)

Dentre os aa identificados na reprodução do *docking*, somente His51 e Asp75 foram concomitantes às interações na pose do trabalho de Behnam et al. (2015). Por outro lado, na reprodução somente Asp129 não foi identificado como aa em potencial para interações com este ligante, quando comparado aos resultados do FTSite. A pose do ligante no alvo no acoplamento de Behnam et al. (2015), demonstra o anel de fenilglicina ocupando o bolso S1, com a arginina em S2 e a porção *cap* em S3, potencialmente se estendendo a S4. A cadeia de lisina, por sua vez, ocupa espaços entre os bolsos, denotando uma liberdade de interação com o alvo ou ligante, como já observado em outras simulações de acoplamento de peptideomiméticos na protease de DENV (BEHNAM et al., 2015). Ademais, mesmo com as inúmeras modificações e substituintes das moléculas utilizadas, ainda se preserva uma similaridade estrutural do esqueleto dos compostos peptideomiméticos, o que implica em poucas alterações nas poses ou interações dessas moléculas com o alvo.

Além dos aa da tríade catalítica serem essenciais à atividade da protease, há uma preferência por fragmentos moleculares de caráter básico nos sítios S1 e S2, como arginina e lisina, dados que reforçam aqueles observados por Behnam, bem como os resultados obtidos previamente nos modelos de HQSAR deste trabalho (CREGAR-HERNANDEZ et al. 2011). Estes e outros grupamentos de caráter polar, usualmente apresentam baixa permeabilidade celular e perfil farmacocinético desfavorável ao desenvolvimento de inibidores, ainda que moléculas com boa atividade *in vitro*. Dados similares já foram demonstrados na protease de WNV e, paralelamente às diversas classes de inibidores peptídicos e não-peptídicos, favoreceram o desenvolvimento de moléculas bioativas como inibidores de protease, por exemplo, inibidores sintéticos em DENV-2 (MUELLER et al., 2007; WU et al., 2015b).

O aa valina (Val155) identificado nos estudos de Behnam et al. (2015), interage com o anel benzil do ligante. A ausência desse aa está associada, por exemplo, a uma redução da afinidade da aprotinina na protease de WNV quando comparado à de DENV. Ademais, esse aa é conservado nos quatro sorotipos de DENV, bem como nas proteases de ZIKV, WNV e JEV, além de apresentar homologia em YFV (isoleucina; I155), além de outros flavivírus, como exemplificado na **Figura 43** (ALESHIN et al., 2007; NOBLE et al., 2012b). Por estes motivos, esse aa também se mostra importante como um potencial sítio de ligação a ser considerado, por exemplo, na construção de um farmacóforo para triagem virtual.



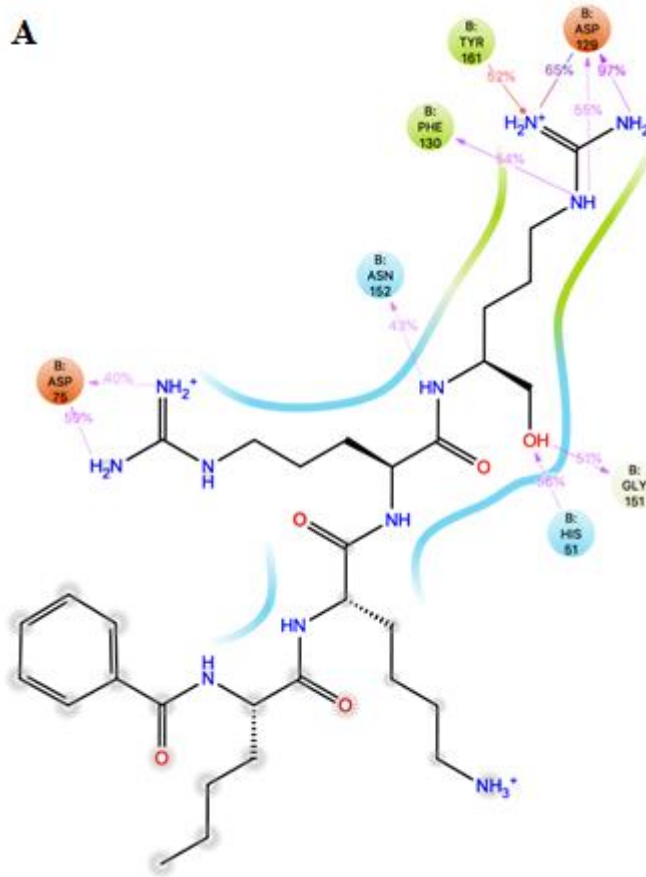
Figura 43. Sequências de NS3 de diferente flavivírus: DENV-1 a 4, JEV, WNV e YFV. Sequências de estruturas depositadas no banco de dados PDB. Identidade (caixas vermelhas) e similaridade (letras vermelhas) dos aa das sequências. Asterisco representa o aa Ser135. Fonte: ALESHIN *et al.* 2007. (*adaptado*)

5.1.10 Dinâmicas moleculares

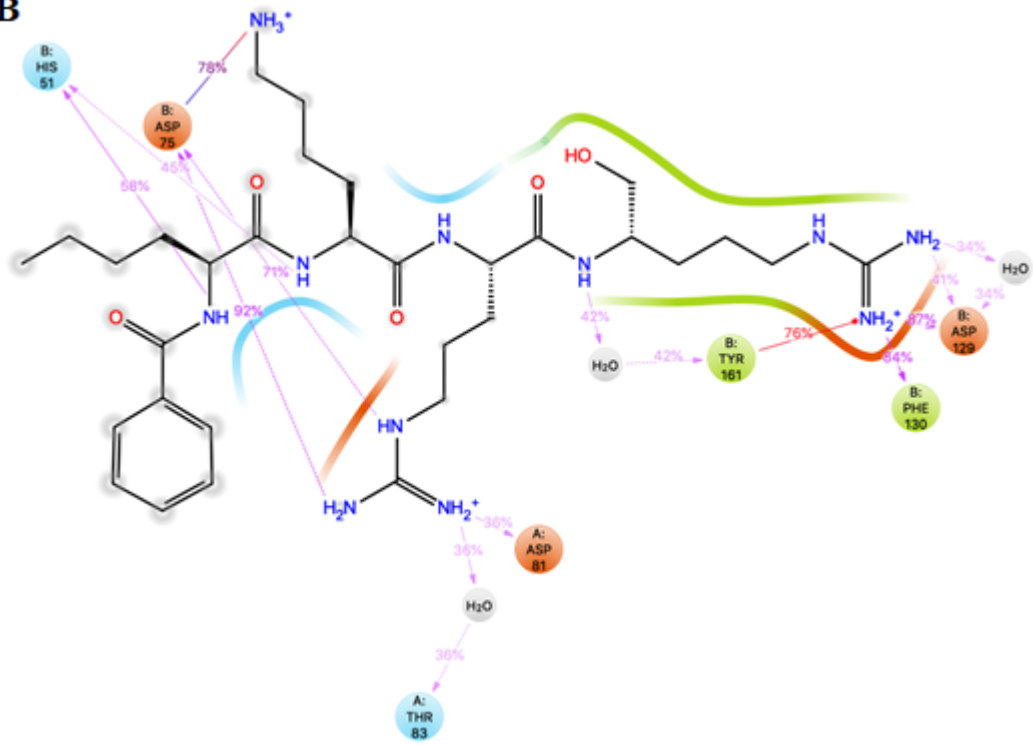
Complementando os dados sugeridos pelos estudos de *docking*, foi realizada dinâmica molecular (DM) dos ligantes originais de ambas proteases, bem como da molécula mais ativa (62), a fim de se simular os movimentos de átomos a nível molecular, analisando possíveis alterações estruturais proteicas (alvo), conformacionais (ligantes) e interações entre os complexos ligante-alvo. Dessa forma foi possível confirmar pontos de interação preditos anteriormente (FTSite e *docking*), fragmentos com maiores interações (HQSAR) e ainda determinar o percentual de interação de cada aa da protease com as moléculas avaliadas.

Inicialmente, uma simulação de DM de 1 μ s do ligante (Bz-Nle-Lys-Arg-Arg-H) original da protease de DENV-3 (3U1I) foi realizada em duplicata, resultando em mapas dos contatos (interações) do ligante na proteína (**Figura 44**). Na primeira replicata (**Figura 44A**) foram preditos contatos de diferentes grupamentos químicos em sete aa da protease: His51 (56%), Asp75 (40 e 59%), Asp129 (55, 65 e 97%), Phe130 (54%), Gly151 (51%), Asn152 (43%) e Tyr161 (52%). O aa Asp75 interage com um fragmento de arginina, assim como Asp129, Phe130 e Tyr161 a outro fragmento de arginina. Aminoácidos podem apresentar interações superiores a 100%, em razão de contatos múltiplos de um mesmo tipo com átomos do ligante, por exemplo, Asp129. Na segunda replicata (**Figura 44B**), resultados similares foram preditos, com contatos em His51 (45 e 58%), Asp75 (71, 78 e 92%), Asp129 (41 e 87%), Phe130 (84%), Tyr161 (76%) e Asp81 (36%), além de Thr83 (36%), este interagindo por contato com uma molécula de água. Novamente, Asp129, Phe130 e Tyr161 interagem com um fragmento de arginina, além de Asp81 e Thr83 a outra arginina. Nesta replicata, Asp75 interage com a amina da lisina, diferentemente da simulação anterior. Abaixo os histogramas representando graficamente as frações de interação.

A



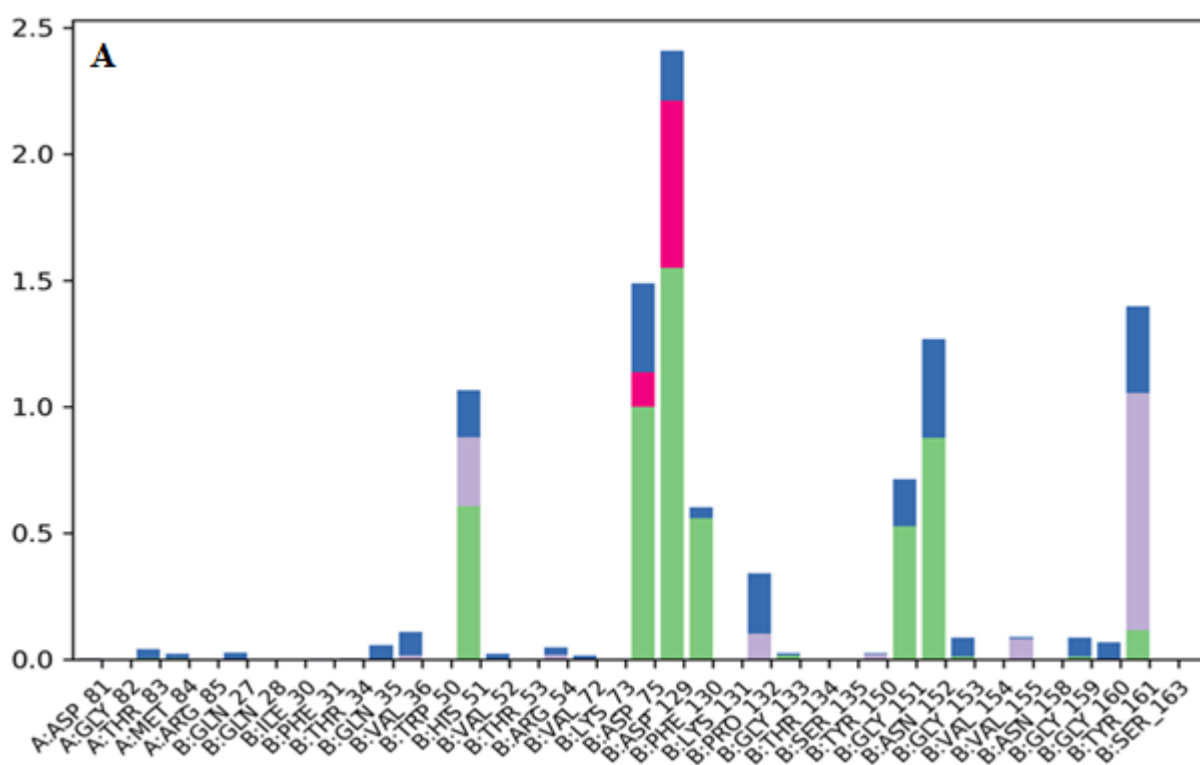
B



- Charged (negative)
- Hydrophobic
- Polar
- Water
- Pi-cation
- Salt bridge
- Solvent exposure

Figura 44. (A) Esquema de contatos de átomos do peptídeo inibidor (ligante de origem) nos aa da protease de DENV-3 (3U1I). Simulações de 1 μ s. Interações que ocorrem em tempo superior a 30% da simulação são representadas. Linhas de tendência de interação: carga ou negativa (vermelho), polar (azul) e hidrofóbica (verde). **(B) Esquema de contatos de átomos do peptídeo inibidor (ligante de origem) nos aa da protease DENV-3 (3U1I) (duplicata).** Simulações de 1 μ s. Interações que ocorrem em tempo superior a 30% da simulação são representadas. Linhas de tendência de interação: ligação de hidrogênio ou polar (azul), iônica ou carga negativa (vermelho) e hidrofóbica (verde).

Essas interações do ligante com a proteína também podem ser monitoradas durante a simulação, o que permite caracterizar cada tipo de contato, sendo categorizados em quatro tipos: ligações de hidrogênio ou de água e interações hidrofóbicas ou iônicas. Dessa forma, foi possível obter um histograma da frequência de interações dos aa do alvo com a molécula (**Figura 45**). Nesta análise, Asp81, His51, Asp75, Asp129, Gly151 e Tyr161 foram identificados com frações de interação superiores a 1 (IF>1.0), resultados que podem indicar um aa realizando múltiplos contatos de um mesmo subtipo com o ligante avaliado, indicando, por exemplo, uma possível inibição e maior estabilidade da interação da molécula no alvo (BOWERS, 2006).



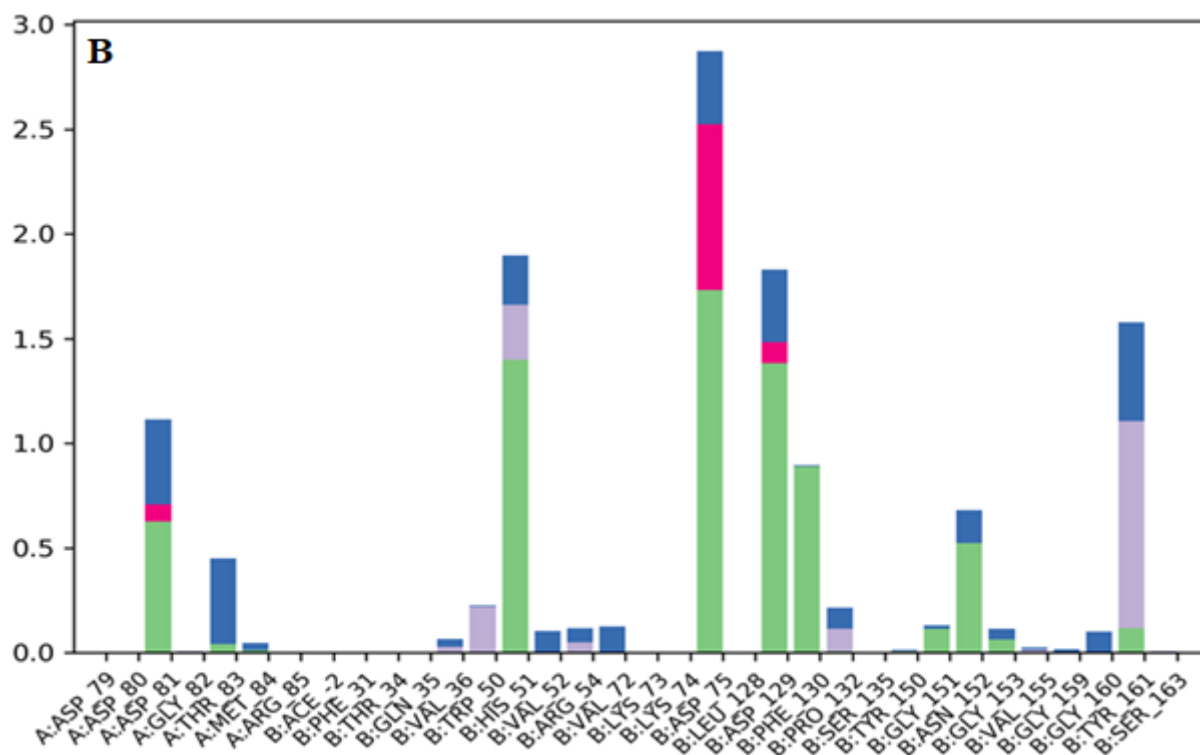


Figura 45. Frequência de interações do ligante (peptídeo) com aa da protease de DENV-3 (duplicatas A e B). Barras representam as frequências de interação (IF) por tipo de cada aa: ligação de hidrogênio (azul), iônica (magenta), hidrofóbica (verde) e ligação de água (cinza). Valores normalizados de acordo com as trajetórias de interação, ou seja, 1.0 sugere 100% de uma interação específica sendo mantida ao longo do tempo da simulação. A: cadeia da NS2B; B: cadeia da NS3.

As DM na protease de ZIKV também é similar às de DENV, sendo preditos no mapa de contatos os aa Asp83 (93%) interagindo com fragmento de lisina, Phe84 (37%) e Gly151 (88%), além de Tyr161 (32%), Asp129 (97%) e Tyr130 (96%), ambos interagindo com fragmento de arginina (**Figura 46A**). A tirosina (Tyr130), por sua vez, possui mesmo caráter químico de fenilalanina (Phe130), mantendo um perfil similar de interação com os ligantes entre as proteases de ZIKV e DENV. Na análise de frequência dos contatos (**Figura 46B**), Asp83, Asp129, Tyr130, Gly153 e Tyr161 foram identificados com frações de interação superiores a 1 (IF>1.0), resultados que podem indicar um aa realizando múltiplos contatos de um mesmo subtipo com o ligante avaliado. Curiosamente, embora não evidenciado na representação de contatos, o aa Asp75 apresenta IF>0.8, ressaltando, também em ZIKV, seu importante papel nas interações com ligantes.

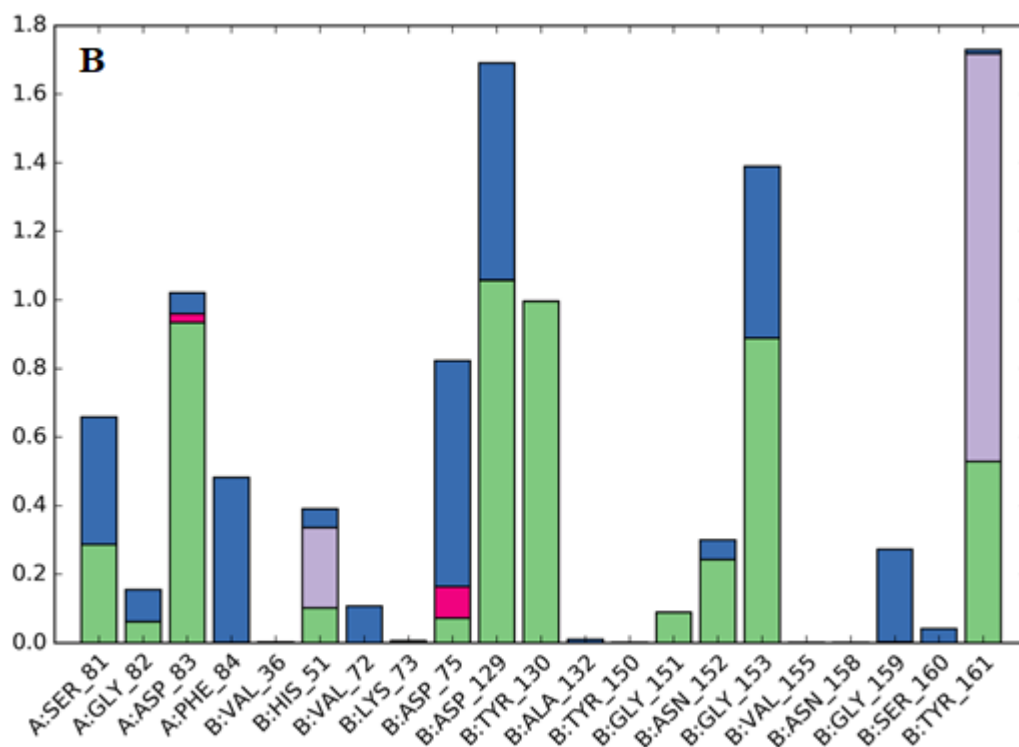
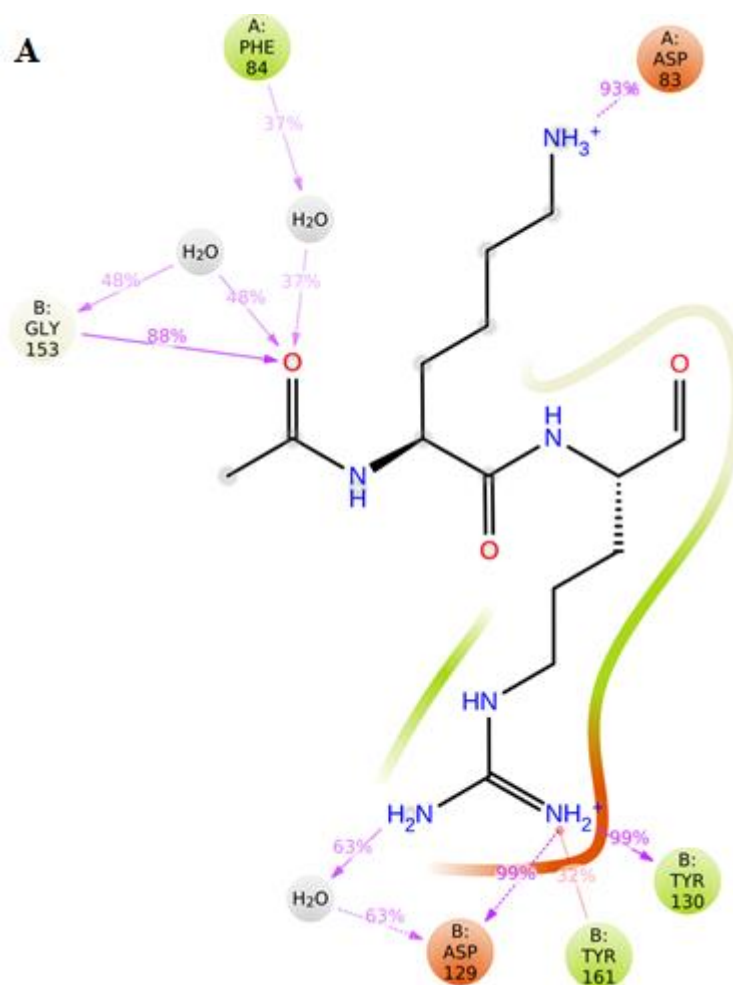


Figura 46. (A) Esquema de contatos de átomos do inibidor dipeptídico (ligante de origem) nos aa da protease de ZIKV (5YOF). Simulações de 1 μ s. Interações que ocorrem em tempo superior a 30% da simulação são representadas. Linhas de tendência de interação: ligação de hidrogênio ou polar (azul), iônica ou carga negativa (vermelho) e hidrofóbica (verde). **(B) Frequência de interações do ligante (dipeptídeo) com aa da protease de ZIKV.** Barras representam as frequências de interação (IF) por tipo de cada aa: ligação de hidrogênio (azul), iônica (magenta), hidrofóbica (verde) e ligação de água (cinza). Valores normalizados de acordo com as trajetórias de interação, ou seja, 1.0 sugere 100% de uma interação específica sendo mantida ao longo do tempo da simulação. A: cadeia da NS2B; B: cadeia da NS3.

Por último, dinâmicas moleculares de 1 μ s das duas moléculas mais ativas da série, 62 e 65 (ambas em DENV e 62 em ZIKV), foram realizadas a fim de se comparar as interações entre os ligantes originais dos cristais das proteases com os compostos peptideomiméticos. Nos mapas de contatos (**Figura 47A**), curiosamente foram preditos apenas três aa da protease: Asp75 (51 e 56%), Asp79 (42%) e Asn152 (41%). O aa Asp75 interage com um fragmento de arginina, enquanto Asp79 a um fragmento de lisina. Novamente, um aa apresentou interações superiores a 100% (Asp75), em razão de contatos múltiplos de um mesmo tipo com átomos do ligante. Já para a frequência de interações (**Figura 47B**), Asp75 e Tyr161 foram identificados com frações de interação superiores a 1 (IF>1.0), resultados que também podem indicar um aa realizando múltiplos contatos de um mesmo subtipo com o ligante avaliado.

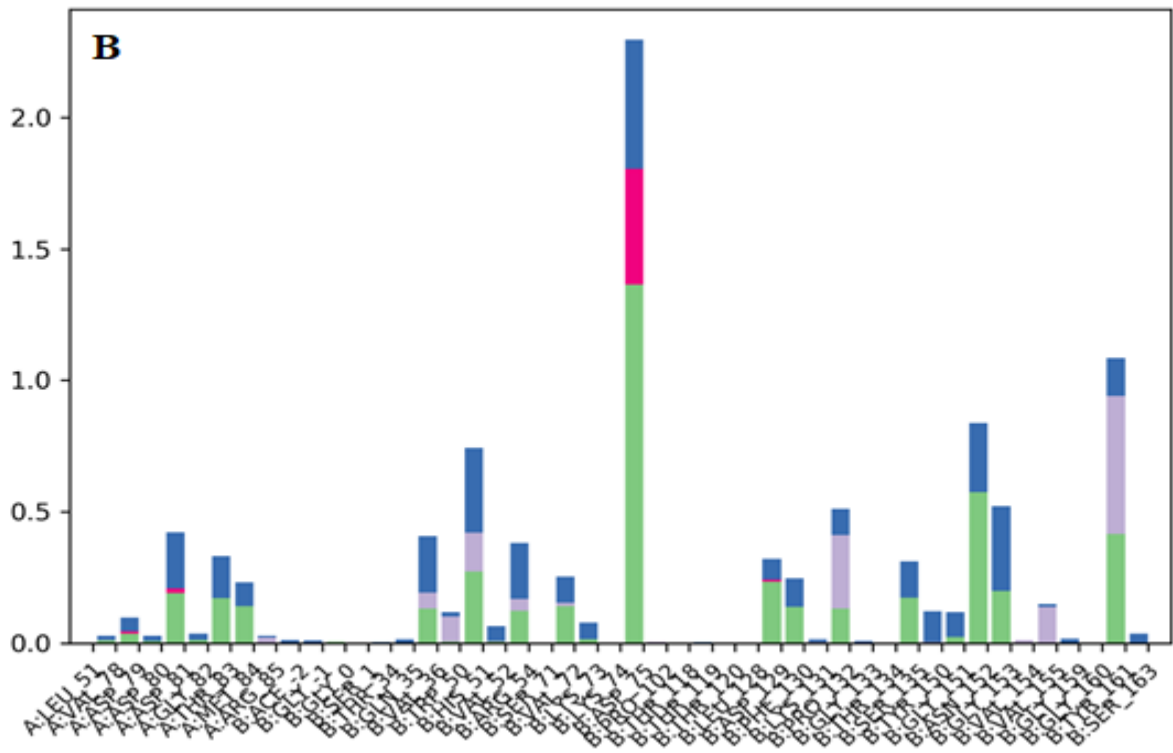
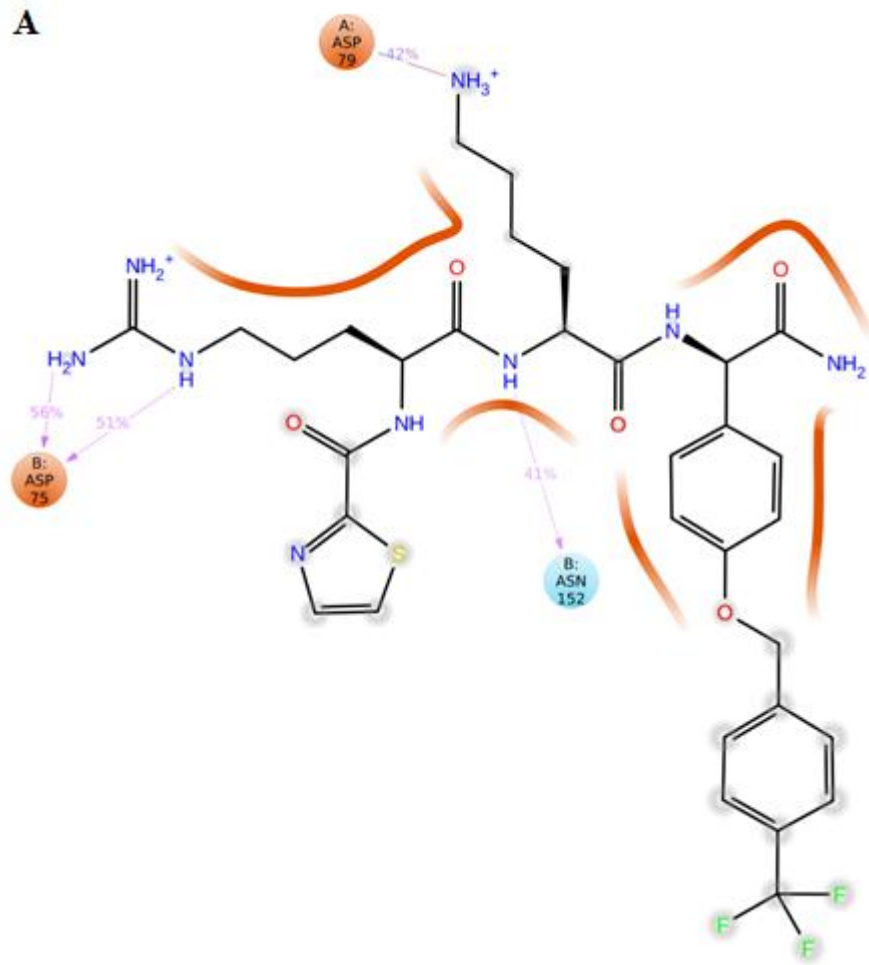


Figura 47. (A) Esquema de contatos de átomos de 62 nos aa da protease de DENV-3 (3U1I). Simulações de 1 μ s. Interações que ocorrem em tempo superior a 30% da simulação são representadas. Linhas de tendência de interação: ligação de hidrogênio ou polar (azul), iônica ou carga negativa (vermelho) e hidrofóbica (verde). **(B) Frequência de interações do ligante 62 com aa da protease de DENV-3.** Barras representam as frequências de interação (IF) por tipo de cada aa: ligação de hidrogênio (azul), iônica (magenta), hidrofóbica (verde) e ligação de água (cinza). Valores normalizados de acordo com as trajetórias de interação, ou seja, 1.0 sugere 100% de uma interação específica sendo mantida ao longo do tempo da simulação. A: cadeia da NS2B; B: cadeia da NS3.

Nas análises de contatos com a molécula 65 (**Figura 48A**), por sua vez, foram preditos contatos em sete aa da protease: Asp75 (96 e 97%), Phe130 (57%), Tyr150 (54%), Gly151 (76 e 84%), Asn152 (54%), Gly153 (74%) e Tyr161 (45 e 73%). O aa Asp75 mais uma vez interage com um fragmento de arginina e, novamente, aa (Asp75, Gly151 e Tyr161) apresentam interações superiores a 100%, em razão de contatos múltiplos de um mesmo tipo com átomos do ligante. Já para a frequência de interações (**Figura 48B**), novamente Asp75 e Tyr161, além de Gly151 e Gly153, foram identificados com frações de interação superiores a 1 (IF>1.0), sugerindo múltiplos contatos de um mesmo subtipo destes aa com o ligante avaliado.

A

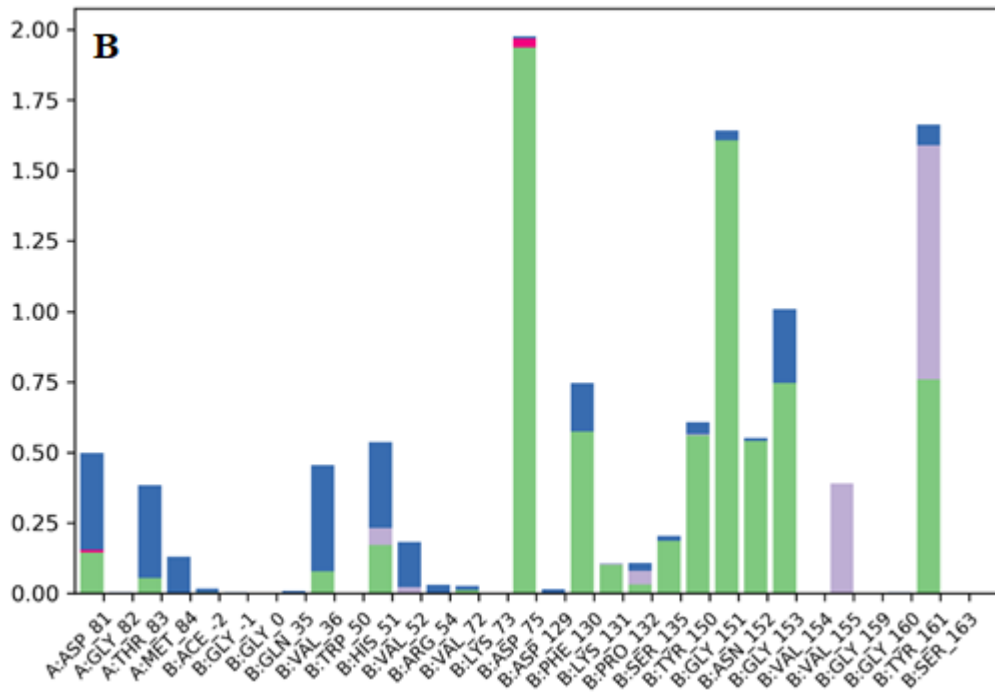
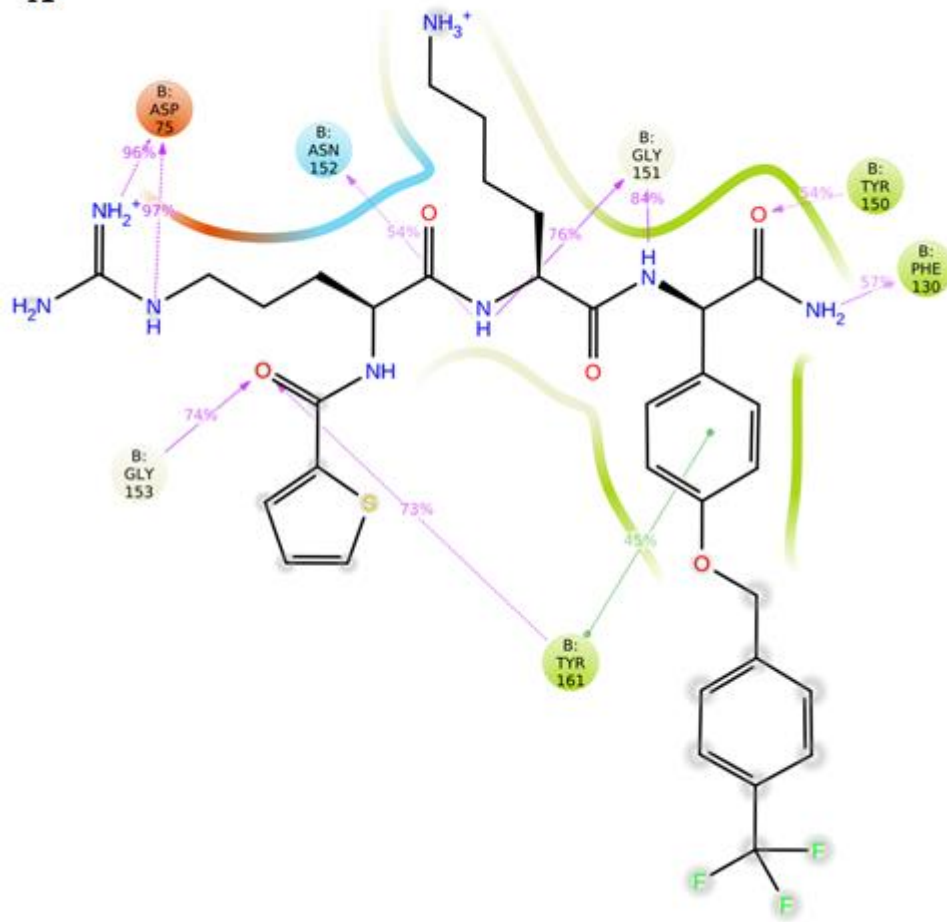


Figura 48. (A) Esquema de contatos de átomos de 65 nos aa da protease de DENV-3 (3U1I). Simulações de 1 μ s. Interações que ocorrem em tempo superior a 30% da simulação são representadas. Linhas de tendência de interação: ligação de hidrogênio ou polar (azul), iônica ou carga negativa (vermelho) e hidrofóbica (verde). **(B) Frequência de interações do ligante 65 com aa da protease de DENV-3.** Barras representam as frequências de interação (IF) por tipo de cada aa: ligação de hidrogênio (azul), iônica (magenta), hidrofóbica (verde) e ligação de água (cinza). Valores normalizados de acordo com as trajetórias de interação, ou seja, 1.0 sugere 100% de uma interação específica sendo mantida ao longo do tempo da simulação. A: cadeia da NS2B; B: cadeia da NS3.

A DM da molécula 62 em ZIKV, na análise de contatos (**Figura 49A**), foram preditos os aa Val36 (54%), Tyr161 (50%), His51 (36%) e, pela primeira vez dentre as simulações, Ser135 (35%). Os aa Ser81 (61%), Asp83 (38, 51 e 53%), Asp75 (56%) e Asn152 (56%) se mostraram interagindo com um fragmento de arginina. Novamente, um aa apresentou interações superiores a 100% (Asp83), em razão de contatos múltiplos de um mesmo tipo com átomos do ligante. Já para a frequência de interações (**Figura 49B**), Asp83, His51 e Asp75 foram identificados com frações de interação superiores a 1 (IF>1.0). Estes resultados, assim como os anteriores, podem indicar um aa realizando múltiplos contatos de um mesmo subtipo com o ligante avaliado.

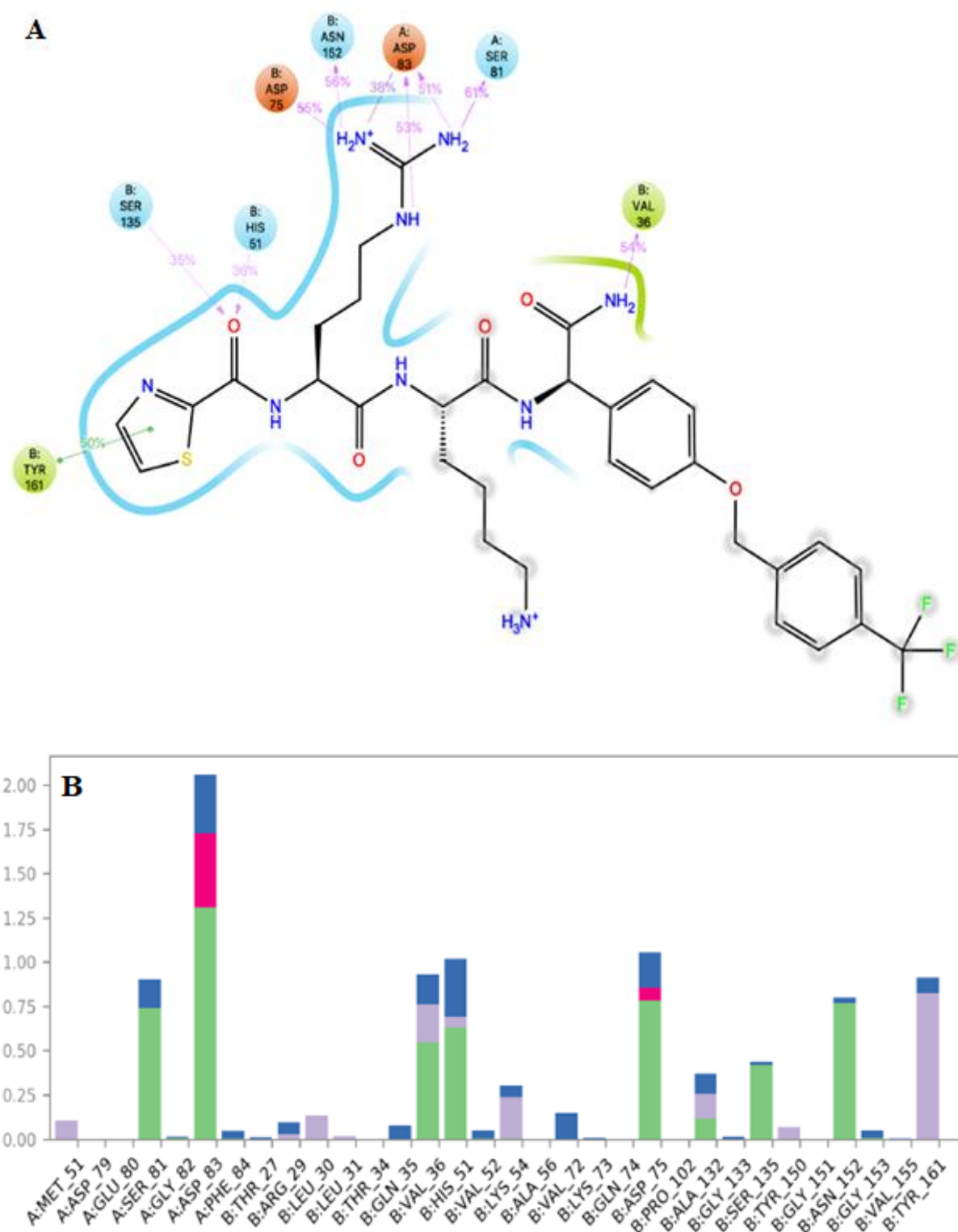


Figura 49. (A) Esquema de contatos de átomos de 62 nos aa da protease de ZIKV (5YOF). Simulações de 1 μ s. Interações que ocorrem em tempo superior a 30% da simulação são representadas. Linhas de tendência de interação: ligação de hidrogênio ou polar (azul), iônica ou carga negativa (vermelho) e hidrofóbica (verde).

(B) Frequência de interações do ligante 62 com aa da protease de ZIKV. Barras representam as frequências de interação (IF) por tipo de cada aa: ligação de hidrogênio (azul), iônica (magenta), hidrofóbica (verde) e ligação de água (cinza). Valores normalizados de acordo com as trajetórias de interação, ou seja, 1.0 sugere 100% de uma interação específica sendo mantida ao longo do tempo da simulação. A: cadeia da NS2B; B: cadeia da NS3.

Essa abordagem por simulações permite, como já mencionado, avaliar não somente a estabilidade de um ligante no alvo, mas a frequência de interações com aa potencialmente importantes na atividade inibitória de uma molécula (BOWERS, 2006). Ademais, a combinação de estratégias computacionais distintas permite reforçar a robustez de métodos almejando obter inibidores seletivos ou corroborar dados preditos, por exemplo, em um acoplamento molecular.

Ainda sim, a potência de um sistema computacional pode limitar o uso de uma DM como estratégia majoritária no planejamento de fármacos, especificamente na capacidade de minimizações, cálculos (especialmente energéticos), predição de interações e, acima destes, a redução do tempo gasto nos ensaios computacionais, permitindo longas simulações de eventos bioquímicos. Contudo, esses fatores, da mesma forma, favorecem o uso dessas simulações como uma importante ferramenta na descoberta de novas moléculas bioativas e no desenvolvimento de potenciais fármacos (CASE, 2005; BOWERS, 2006).

Diversas abordagens que combinaram as DM com outras estratégias computacionais já obtiveram sucesso na busca de moléculas bioativas, por exemplo, com atividade antibacteriana. Rehnem et al. (2019), por exemplo, utilizaram dinâmicas moleculares para análises comparativas de inibidores oriundos de uma triagem virtual tendo NDM-1, uma beta-lactamase, como alvo. As DM permitem avaliar a estabilidade das moléculas testadas no complexo da proteína, o que permitiu, nesse estudo, indicar interações no sítio ativo da enzima, corroborando ligações de hidrogênio, interações eletrostáticas (iônicas) e hidrofóbicas, resultados que foram validados confirmando a inibição enzimática em ensaios *in vitro* determinando a IC₅₀ dos *hits* triados. Uma abordagem semelhante também demonstrou o potencial de simulações de dinâmica molecular na validação de um potencial inibidor de bomba de efluxo de *Acinetobacter baumannii* predito em análises de *docking* (VERMA; TIWARI; TIWARI, 2018).

Ajmi et al. (2018) também demonstraram a vantagem de aplicações da dinâmica molecular na identificação de produtos naturais como inibidores de uma quinase (PLK-1) associada a diversos tipos de câncer. Novamente, as DM permitiram validar, a partir de uma triagem virtual, a estabilidade de moléculas bioativas no sítio ativo do receptor, bem como

corroborar os aa associados à interação dos ligantes. Similarmente, essa estratégia já havia auxiliado no planejamento de potenciais inibidores dessa mesma classe de enzimas, analisando semelhanças estruturais dos receptores e das moléculas (JANI; DALAFAVE, 2012).

Na virologia, por outro lado, estudos com dinâmicas moleculares usualmente apresentavam enfoque nas interações vírus-hospedeiro, buscando elucidar, por exemplo, o papel de proteínas no ciclo de multiplicação e/ou respostas imunológicas, como já demonstrado na construção de um modelo estrutural de CCR5, um co-receptor para HIV-1 (MAEDA et al., 2008; ODE et al., 2012). Paralelamente, estratégias com simulações na descoberta e desenvolvimento de fármacos antivirais já obtiveram sucesso, podendo permitir simulações mais longas de sistemas mais complexos com o avanço dos recursos disponíveis e velocidade de processamento dos computadores no futuro (DURRANT; MCCAMMON, 2011; BORHANI; SHAW, 2012).

Por exemplo, Schames et al., ainda em 2004, aplicaram essa técnica em análises da integrase de HIV, identificando uma região de interação na enzima ainda não identificada em estruturas de cristais disponíveis na época. Lee et al. (2006), indicaram a influência desse sítio adjacente com estudos de mutagênese induzida nessa região, permitindo auxiliar, posteriormente, no desenvolvimento de inibidores de integrase, como o fármaco Raltegravir, pela Merck, o primeiro inibidor de integrase aprovado pelo FDA no tratamento de HIV (HAZUDA et al., 2004; SCHAMES et al., 2004; LEE; ROBINSON, 2006).

No mesmo âmbito, DM já foram capazes de demonstrar uma cavidade universal no sítio de adesão da neuraminidase do vírus influenza, como H1N1, H5N1 e H2N2, auxiliando, por exemplo, no desenvolvimento de fármacos contra estes vírus (AMARO et al., 2011; DE LIMA NETO et al., 2018). Voltado para os flavivírus, interesse desse trabalho, estudos como os de Lima Neto et al. (2018) já haviam proposto dinâmicas moleculares para o refinamento de análises de acoplamento molecular de moléculas comercialmente disponíveis e com atividade descrita em DENV, sendo reposicionadas como potenciais inibidores da NS3, NS5 ou proteína E de ZIKV.

Considerando a importância das dinâmicas moleculares, suas associações com estratégias de triagem, as informações obtidas nesse trabalho, além daquelas apresentadas acerca das sequências da NS3, estrutura da protease e interações com ligantes permitem, por exemplo, considerar a construção de um farmacóforo para triagem virtual de moléculas com potencial atividade antiviral em DENV, YFV e ZIKV. Além disso: (i) a ausência de uma

estrutura de cristalografia da protease de DENV-2 em complexo com um inibidor, (ii) a alta taxa de conservação da protease entre os demais sorotipos de DENV e espécies do gênero *Flavivirus*; (iii) a qualidade do cristal de DENV-3 (3U1I); (iv) os estudos de Behnam; (v) a similaridade dos sítios de ligação entre a protease de ZIKV e DENV-3; e (vi) a ausência de uma estrutura de cristalografia de YFV, reforçam a seleção da protease de DENV-3 (3U1I) como alvo único nessa abordagem.

5.1.11 Construção do modelo farmacofórico e triagem virtual

Uma estratégia de triagem virtual na busca de potenciais inibidores ou moléculas bioativas, permite, por exemplo, reduzir tempo e custos com novas rotas de síntese e tentativas de modificações estruturais, identificando diretamente moléculas de rápida ou fácil obtenção, ao se aplicar diferentes filtros de triagem. Nesse caso, a construção de um modelo farmacofórico permite filtrar compostos de interesse que possam interagir em regiões específicas de um alvo, como a protease de DENV em estudo (LEONEL, 2018).

Tratando dessa estratégia, com o compilado de variáveis e informações nos resultados desse trabalho (HQSAR, FTSite, FTMap, AutoGrid, análises de *docking* e DM), um modelo farmacofórico da protease de DENV-3 foi construído para aplicação em uma triagem virtual, sendo selecionados aa de interesse e definidos centroides de interação molecular. O modelo para triagem dos potenciais inibidores (ligantes) foi definido e construído via plataforma UNITY no programa SYBYL-X 2.1, consistindo nos três aa da tríade (His51, Asp75 e Ser135), além de Asp79, Asn152 e Val155, além de quatro centroides de interação molecular, sendo um hidrofóbico (**HYDROPHOBIC2**), um doador de hidrogênio (**DONOR_ATOM1**) e dois positivos (**POSITIVE_N1** e **POSITIVE_N2**), sendo o segundo em condição de *partial match*, ou seja, correspondência parcial de interações (**Figura 50**). Os aa Asp129, Phe130 e Tyr161 foram desconsiderados, haja visto uma certa inconsistência da presença de interações apresentadas em um tipo de ligante ou protease nos diferentes métodos avaliados. Além disso, uma substituição presente na posição 130 entre as proteases de ZIKV e DENV e a ocupação próxima dos outros dois aa (aspartato e tirosina), reforçaram essa escolha.



Figura 50. Modelo farmacofórico construído utilizando a estrutura do complexo entre a protease de DENV-3 e o ligante de origem. O modelo apresenta quatro centroides de interação sendo: um hidrofóbico (ciano), um doador (azul) e dois positivos (vermelho). Aa assinalados (*sticks*): His51, Asp75, Asp79, Ser135, Asn152 e Val155 (azul). Estrutura da protease em modelo de *cartoon* (PDB: 3U1I; em verde).

Construído o modelo farmacofórico, filtros foram aplicados aos quatro conjuntos de moléculas, buscando otimizar a triagem virtual a ser executada. Inicialmente, modificações dos parâmetros baseados na Ro5 de Lipinski (LIPINSKI et al., 1997) e regras de Veber (VEBER et al., 2002) foram aplicados, como já apresentados e discutidos anteriormente.

Além dessas variáveis iniciais, foi estabelecido um segundo filtro baseado nos resultados preditos no modelo de HQSAR, especificamente nos dois fragmentos moleculares de interesse, arginina e lisina. Dessa forma, dentre as 7.600.000 moléculas dos quatro conjuntos de dados (MolPort/ZINC, DrugBank/FDA, NuBBE e QF/UFMG), foram selecionadas apenas aquelas que apresentassem fragmentos similares ou do tipo arginina e/ou lisina de acordo com o filtro estrutural delimitado na triagem pelo KNIME®, haja visto sua importância na atividade inibitória das proteases observada nos resultados anteriores e em estudos anteriores já citados (CREGAR-HERNANDEZ, 2011; WU, 2014; BEHANM, 2015).

Os quatro bancos de dados utilizados foram submetidos ao mesmo tratamento dos compostos peptidomiméticos realizados no pacote computacional KNIME®, para criação de quatro respectivos bancos de moléculas contendo o confômero de menor energia. Posteriormente, cada molécula também teve seu estado de protonação corrigido de acordo com pH fisiológico, também utilizando o pacote computacional KNIME®.

As aplicações desses comandos permitiram, para todas as moléculas e potenciais ligantes a serem triadas no alvo, a obtenção de uma unidade conformacional de maior estabilidade e estado iônico mais favorável, conseqüentemente, com um melhor perfil energético e de possíveis interações. Seguidamente, ambos os filtros também foram executados no pacote computacional KNIME®, descrevendo cada variável dos parâmetros definidos e propriedades físico-químicas das moléculas a serem triadas, filtrando, etapa a etapa, os quatro bancos de dados paralelamente.

5.1.12 Triagem virtual: filtros e *crossdocking*

Na execução do primeiro filtro, avaliando os parâmetros de biodisponibilidade (regras de Lipinski e Veber), dentre as 7.600.000 moléculas oriundas do MolPort, DrugBank, NuBBE e QF, cerca de 820.000 moléculas restaram. Logo após, com o segundo filtro, considerando os fragmentos arginina e lisina, aproximadamente 220.000 substâncias resultaram da triagem. Finalmente, com os candidatos restantes, a triagem pelo farmacóforo foi realizada (**Figura 51**), resultando em 2.775 moléculas, oriundas somente dos bancos MolPort (2.763) e DrugBank (12). Nenhuma molécula dos bancos NuBBE e QF foi obtida ao final da triagem pelo farmacóforo.



Figura 51. Pose de *docking* de um potencial inibidor triado pelo farmacóforo. Modelo farmacofórico construído na estrutura da protease de DENV-3 (3U1I) (verde). O modelo apresenta quatro centroides de interação sendo: um hidrofóbico (ciano), um doador (azul) e dois positivos (vermelho). Ligante (*sticks*) representativo de uma simulação de interações da molécula no farmacóforo da protease durante a triagem.

A fim de reduzir o número de moléculas triadas, uma estratégia de *docking* foi implementada em dois protocolos distintos (P1 e P2), uma abordagem que permite selecionar apenas os potenciais inibidores com melhor *fit* ou interações na protease, mais especificamente em uma cavidade específica do alvo. Essa abordagem também permite contornar algumas variações que podem ocorrer em diferentes algoritmos ou programas, como polarizações incorretas, dessolvatação aquosa, influências iônicas, pouca entropia entre a proteína e o alvo, bem como concentrações insuficientes de ambos, além de estados de ionização e protonação desfavoráveis, presunções incorretas da rigidez de uma proteína e deficiências na pontuação de moléculas acopladas (MARS DEN et al., 2004; HAWKINS; SKILLMAN; NICHOLLS, 2007).

O primeiro protocolo (P1) foi realizado no programa Surflex (modelo Geom), em que um *protomol* ou cavidade foi então padronizado e delimitado baseado na ocupância do ligante de origem (peptídeo NDL) da protease de DENV. Assim, foram definidos um limiar de 0,5 e um volume igual a 0 (espaço mínimo de ocupância), para delimitação do envelope de interações (**Figura 52**). Além destas características, foram testados 66 modelos distintos com o ligante de origem (NDL), variando diversos parâmetros, até a definição de um modelo ótimo para o *crossdocking*: limiar (0,5); tamanho (0); movimento de hidrogênios (não); movimento de hidrogênios de átomos (não); *CScore* (sim); poses otimizadas (não); número de conformações (8); expansão de busca (6Å); máximo de conformações (60); máximo de ligações rotacionáveis (100); densidade de busca (6); número de spins possíveis (24); Poses (20); e RMSD (0,5).

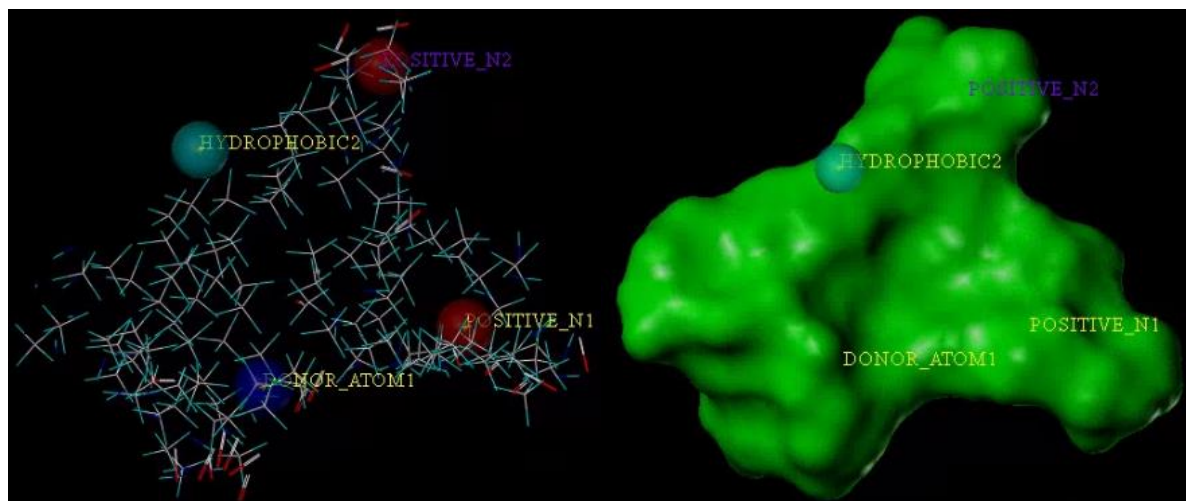


Figura 52. Modelo do *protomol* ou envelope de interações. O envelope inicial é composto de átomos de C, O, N e H, delimitando pontos possíveis de interação, considerando os centroides do farmacóforo: hidrofóbico (ciano), um doador (azul) e dois positivos (vermelho). O modelo é então renderizado com o tamanho da ocupância e limiar do volume máximo do *protomol* (verde).

5.1.13 Análise e classificação das moléculas obtidas na triagem

Selecionado o modelo, o *crossdocking* das 2.775 moléculas foi realizado (P1), selecionando os 100 melhores *hits* do banco MolPort de acordo com os valores de *score* (>8,7), além das 12 moléculas triadas no DrugBank, independentemente de *score* (6,788 a 8,968). Destas 112 moléculas resultantes, foi feita uma inspeção visual dos resultados do *crossdocking*, uma a uma (**Figura 53A**), estabelecendo um ranqueamento a partir de uma pontuação arbitrária de quatro critérios: (i) tipo de interação com aa da protease; (ii) interação ocorrendo na região dos centroides; (iii) interações com aa do farmacóforo, incluindo a tríade catalítica; e (iv) interações com demais aa da protease (**Figura 53B**). Segundo esses critérios, a pontuação arbitrária recebeu pesos de acordo com a força da ligação estabelecida ou aa específico de interação do farmacóforo (maior importância), sendo: iônica = 2, hidrogênio = 1, hidrofóbica = 0,5, além do dobro de pontos quando houvesse posicionamento exato em um centroide ou uma interação com His51, Asp75 e Ser135, Asp79, Asn152 ou Val155.

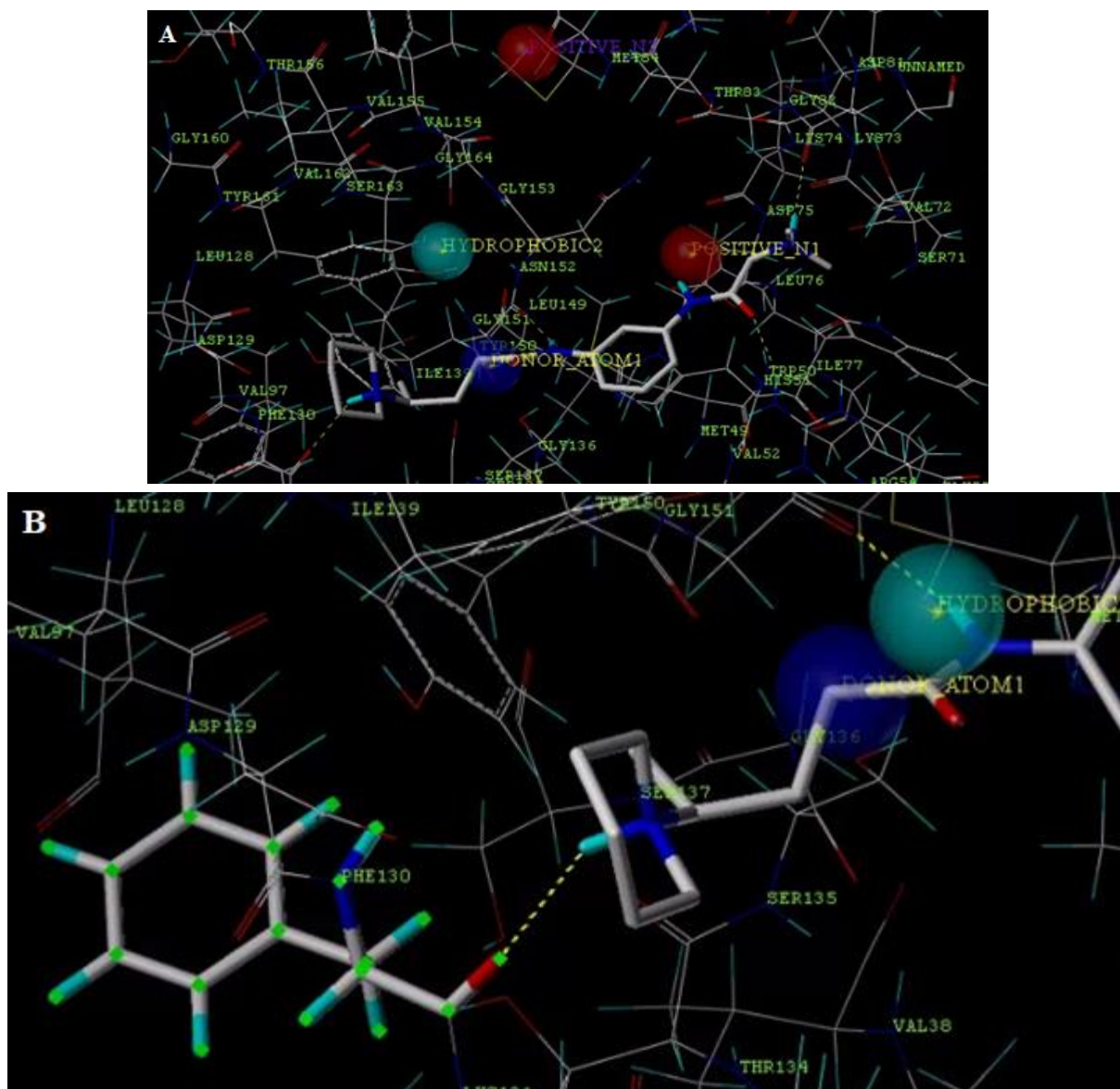


Figura 53. (A) Exemplo de análise de interações do *docking* no sítio do farmacóforo. Molécula interagindo com aa da protease. Centroides do farmacóforo: hidrofóbico (ciano), um doador (azul) e dois positivos (vermelho). **(B) Exemplo aumentado do modelo de pontuação das interações.** Aminoácido Phe130 realizando uma interação iônica (íon amônio). Paralelamente, outra interação ocorre no centroide hidrofóbico. Considerando os quatro critérios de pontuação: uma interação iônica (2) + uma interação no centroide hidrofóbico (2) = 4.

Dentre as 112 substâncias selecionadas (100 oriundas do MolPort e 12 do DrugBank), ao menos 20% do conjunto ou aquelas que apresentassem soma total mínima maior que 10 pontos (pontuação padrão baseada nos resultados do ligante de origem da protease), foram selecionadas. No primeiro protocolo (P1), 33 moléculas foram selecionadas, sendo 25 moléculas oriundas do MolPort e 6 do DrugBank, como demonstrado na **Tabela 15**. Dessas, apenas 31 apresentaram soma superior a 10, sendo selecionadas às etapas seguintes.

Tabela 15. Pontuações dos 33 hits (protocolo 1) no farmacóforo. A tabela apresenta a identificação de cada molécula, junto ao valor de *Score* resultante do *crossdocking*. Ocupação nos centróides positivo (P'), doador (D), hidrofóbico (H) e positivo *partial match* (P''). Pontos de interações possíveis com aa do farmacóforo (His51, Asp75, Ser135 e Asn152). Pontos de interações possíveis com aa relevantes no completo ligante-alvo. ST: soma total das pontuações (demais aa não apresentados). Valores em vermelho correspondem a mais de uma interação ocorrendo em um mesmo aa. Linhas destacadas (cinza) correspondem a moléculas desconsideradas desta etapa.

Moléculas Código	score	Centroides				Farmacóforo				Aminoácidos relevantes					ST	
		P'	D	H	P''	His51	Asp75	Ser135	Asn152	Asp129	Phe130	Gly151	Gly153	Tyr161		
112	9,140					2			4				1			10
113	8,895							4	2							12
114	8,714							2		6	1					13
115	8,776					2		6			1		1	1		14
116	8,800							2			2		1	1		10
117	8,748	1					4		4			2	1	1		12
118	9,690		1					8					1	1		11
119	9,466							4		2		2				10
120	8,795							4		3						10
121	9,008											2				11
122	9,004							2		5	2	1				11
123	9,613				1				4				1			10
124	10,458							4		4	2					13
125	9,320								4				1			10
126	11,174		1					4		2	2		1	1		11
127	8,765		1			2		4		2	2					11
128	8,975		2					4	4		2		1	1		12
129	8,941	1							4				2			10
130	8,810		2					2		4		1	2	2		11
131	8,975							4			2	2	1	1		10
132	8,708	1							6				1			10
133	10,572	1		1				4	4			2				11
134	10,512			1				4	4			2				11
135	8,914	1						4		2	2		1	1		12
136	8,807	1						4					1			10
137	8,799		1	1				4	4	4		2				14
138	9,232		2								2	2				10
139	8,664							2				1	1			9
140	7,567	1				2		4								9
141	8,317							4	4			2				11
142	8,585							4	4	1	2	2				15
143	8,968							4	4	1	2	2				14
144	8,293							4	2	4		2				14

Curiosamente, não houveram pontos de interações possíveis nos aa Asp79 e Val155 do farmacóforo. No entanto, os aa Asp129, Phe130, Gly151, Gly153 e Tyr161 apresentaram inúmeros pontos de interações possíveis dentre as moléculas, dados que corroboram os resultados observados previamente nas análises das predições de sítios, *docking* e dinâmica molecular. No geral, dentre os demais, Ser135 apresentou o maior número pontos somados de interações totais e o maior número de pontos de interações em um único aa. Esse resultado pode sugerir o potencial inibitório dessas moléculas, uma vez que o aa serina da tríade catalítica é essencial para função da enzima, atuando como o gatilho que desencadeia a atividade proteolítica (TIMIRI; SINHA; JAYAPRAKASH, 2016).

Posteriormente, um segundo protocolo (P2) de *crossdocking* das 2.775 moléculas foi realizado no programa Glide, utilizando somente o farmacóforo, sem outras condições limitantes de poses ou conformações (parâmetros de P1), bem como quaisquer delimitações no alvo (*protomol*), considerando todos os tautômeros preditos, estados de ionização das moléculas e presença de moléculas de água na estrutura do alvo molecular. Dessa forma, isômeros em equilíbrio dinâmico interagindo na protease, além da presença ou ausência de seus respectivos estados ionizados e moléculas de água no alvo foram analisados, aumentando as variáveis possíveis de cada composto interagindo no sítio (cerca de 7000 possibilidades relativas às 2775 moléculas do protocolo anterior). O protocolo 2 resultou em 18 *hits*, submetidos à mesma inspeção visual de interações, restando 8 moléculas selecionadas com valores superiores a 10 **(Tabela 16)**.

Tabela 16. Pontuação dos 18 hits (protocolo 2) no farmacóforo. A tabela apresenta a identificação de cada molécula, junto ao valor de *score* resultante do *crossdocking*. Ocupação nos centróides positivo (P'), doador (D), hidrofóbico (H) e positivo *partial match* (P''). Pontos de interações possíveis com aa do farmacóforo (His51, Asp75, Ser135 e Asn152). Pontos de interações possíveis com aa relevantes no completo ligante-alvo. Soma total das pontuações (demais aa não apresentados). Valores em vermelho correspondem a mais de uma interação ocorrendo em um mesmo aa. Linhas destacadas (cinza) correspondem a moléculas desconsideradas desta etapa.

Moléculas Código (VS)	score	Centroides				Farmacóforo				Aminoácidos relevantes					ST
		P'	D	H	P''	His51	Asp75	Ser135	Asn152	Asp129	Phe130	Gly151	Gly153	Tyr161	
145	6,425							4							6
146	9,3502						2					2			8
147	6,9365					4		4			2				11
148	7,9251					2					1				4
149	7,7360					4		4			2				11
150	8,4199					4		4			2				13
151	7,9467	1				4		4	1		2				12
152	9,4513							4	2	4					15
153	5,8769								2	2		2		1	9
154	7,6906								7	1		1		1	11
155	8,2129								7	1					10
156	7,3833						2		4		1	2		2	11
157	7,4902	1							2	2	2				7

Nos resultados de P2, se evidencia uma tendência de interação nos aa considerados relevantes, aquém daqueles do farmacóforo, em que não ocorreram interações, por exemplo, com His51, além de menos interações com Ser135. Esse perfil pode sugerir um efeito limitante dos parâmetros utilizados em P1, condicionando interações específicas, ou uma tendência inespecífica de interações de um modelo potencialmente menos seletivo de triagem (P2).

Curiosamente também, nenhuma molécula foi obtida concomitantemente em ambos protocolos, o que valoriza a estratégia executada com modelos em paralelo. Pensando nisso, um segundo filtro foi aplicado, realizando um *crossdocking* reverso dos resultados de P1 e P2 entre os dois programas (Surflex e Glide), buscando evidenciar possíveis moléculas de características mútuas em ambos os protocolos.

5.1.14 Filtrando os protocolos de triagem

Assim, o *docking* dos hits de cada protocolo foi realizado no segundo programa, ou seja, os compostos oriundos do *docking* no Surflex foram submetidos ao Glide (P1'') e aqueles

oridundos do *docking* no Glide foram submetidos ao Surflex (P2''). Assim, um duplo filtro com as variáveis de dois programas distintos foi estabelecido, considerando agora, para ambos: (i) presença e ausência de tautômeros, estados de ionização e moléculas de água; (ii) presença e ausência de todos os parâmetros de orientação e interação dos ligantes (poses, *spins*, RMSD, expansão e densidade de busca, etc.); (iii) ranqueamento por inspeção visual individual de cada molécula.

Os resultados dos *hits* no Surflex (31) e Glide (8) foram analisados comparativamente dentre as poses similares e interações da molécula com aa equivalentes. Evitando desconsiderar moléculas triadas por apenas um dos programas, estas foram comparadas de acordo com os resultados recuperados dos acoplamentos (*crossdocking*) do protocolo oposto. As análises, também por inspeções visuais, dos dois protocolos foram realizadas no programa DiscoveryStudio, obtendo um total de 18 substâncias, sendo 10 oriundas do protocolo Surflex>Glide (P1') e 8 do Glide>Surflex (P2') (Figura 54).

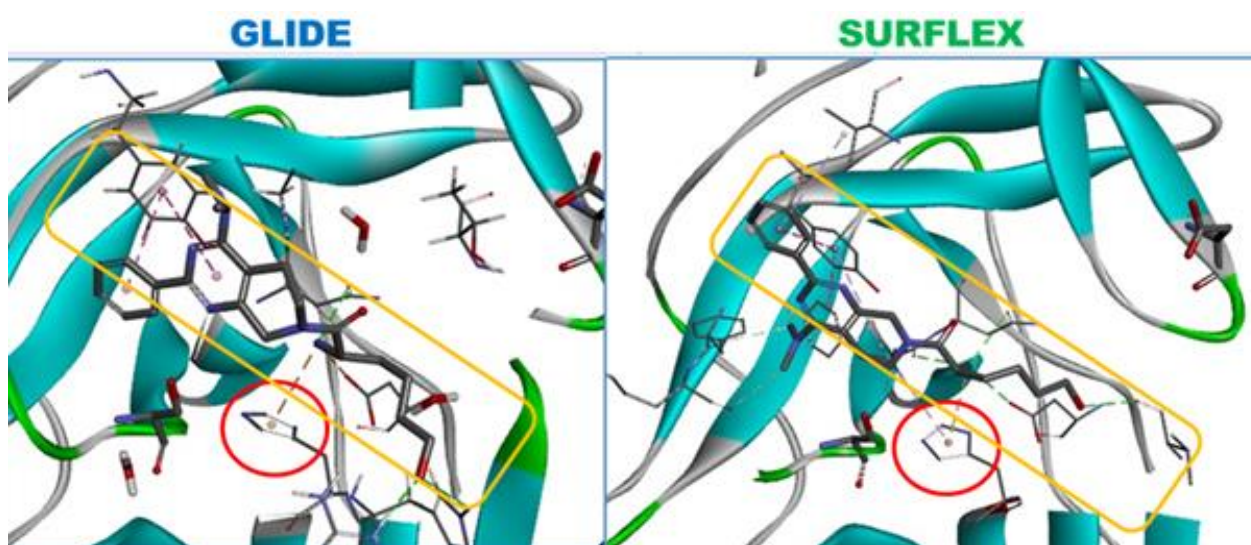


Figura 54. Comparação entre os resultados dos protocolos de *docking* nos programas Glide (P2) e Surflex (P1). Exemplo de análise entre ambos os protocolos. A conformação e orientação das poses devem ser respeitadas (amarelo), bem como as interações com os aa identificados (vermelho). Estrutura da protease de DENV em modelo de *cartoon* (ciano).

Contrário à diferença de perfil de interações observada entre os protocolos P1 e P2, na qual P1 demonstrou mais interações com variados aa, especialmente da tríade catalítica, o duplo filtro P2'' acabou se mostrando mais conservador mesmo após acoplamento no Glide seguido de Surflex. O protocolo denominado P1'', por outro lado, foi mais seletivo quando realizado Surflex>Glide, uma vez que apenas 10 moléculas restaram das 31 selecionadas em P1. Assim

se reforça a importância de uma abordagem mais extensa de *docking*, que aborde dois ou mais programas e algoritmos distintos nas análises de acoplamento molecular. As moléculas selecionadas nos dois protocolos estão dispostas na **Tabela 17**.

Tabela 17. Estrutura das moléculas selecionadas na análise de duplo filtro dos protocolos 1 e 2 (P1'' e P2''). Protocolos distinguidos por coloração em sombreamento : P1'' (verde) e P2'' (ciano).

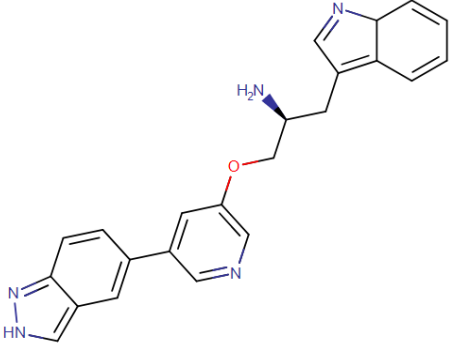
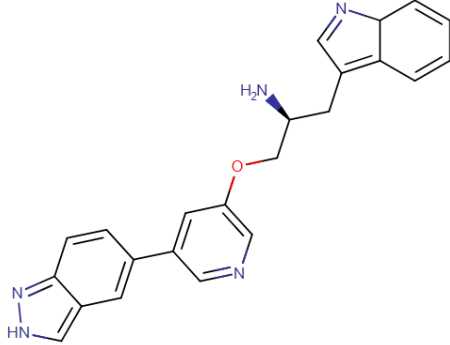
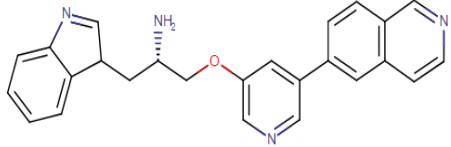
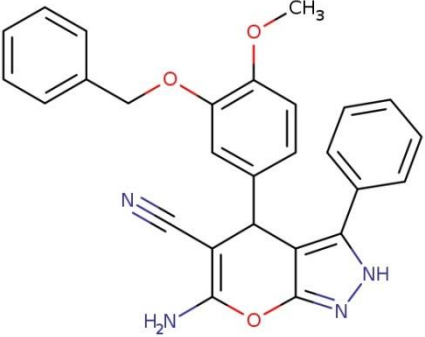
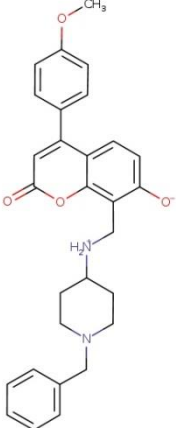
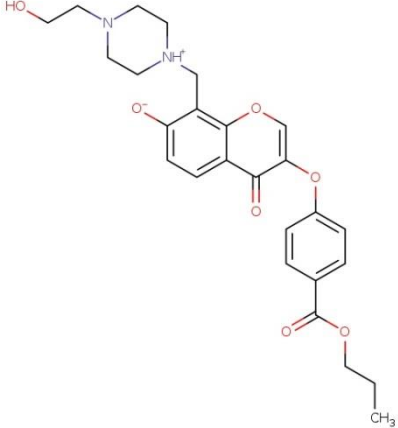
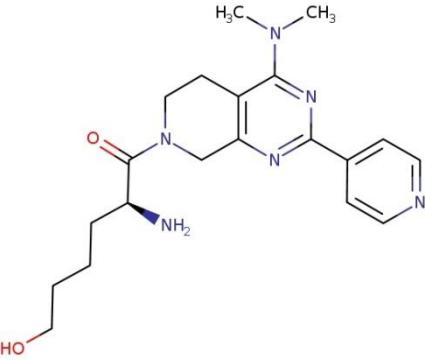
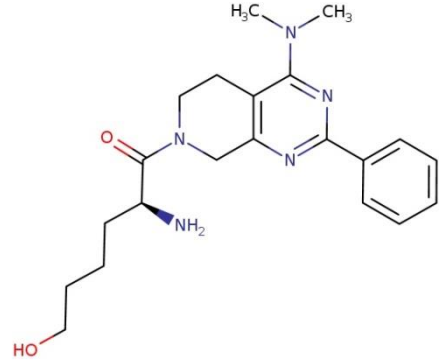
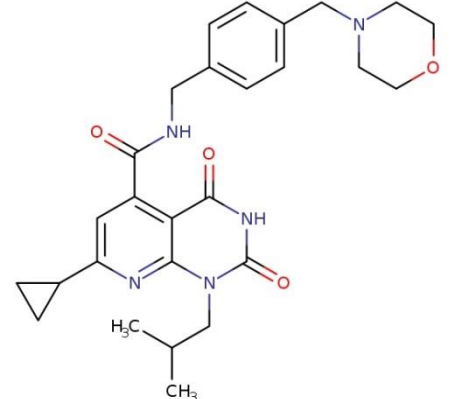
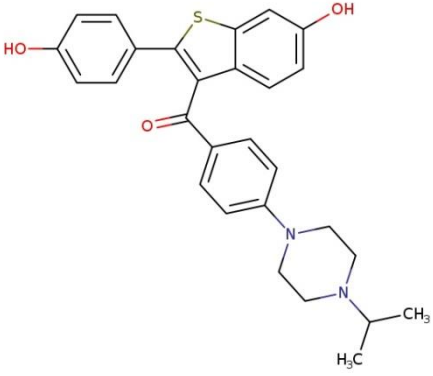
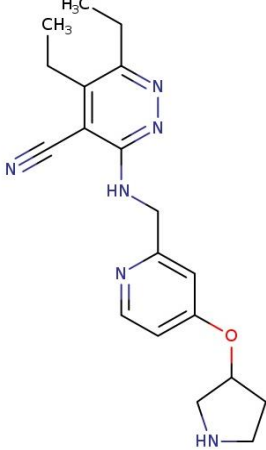
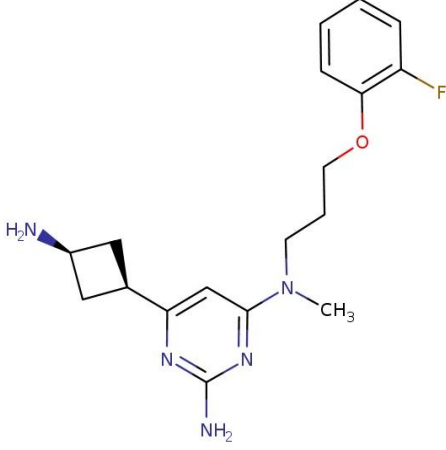
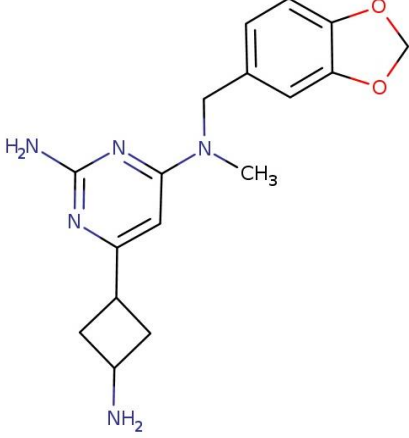
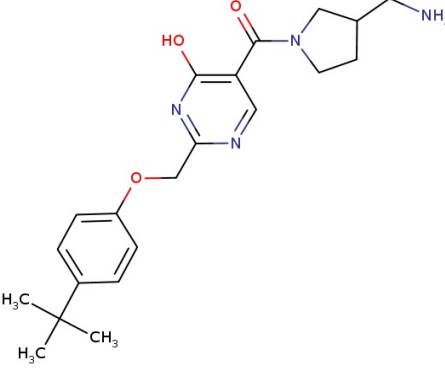
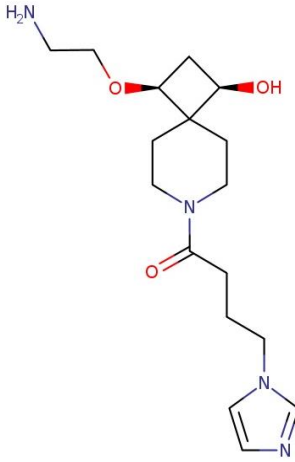
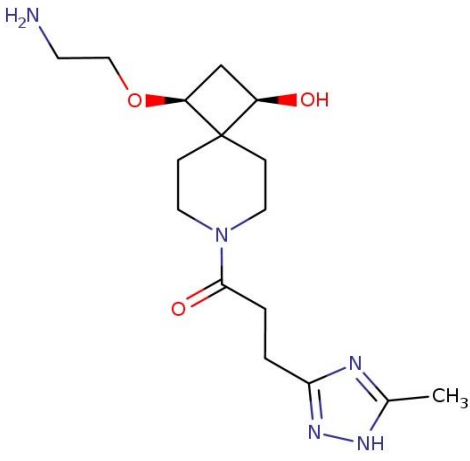
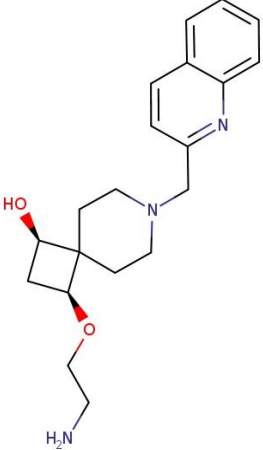
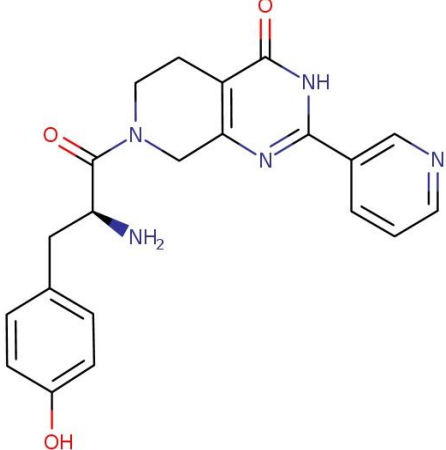
Molécula 142 (P1'')	Molécula 143 (P1'')	Molécula 141 (P1'')
 <p>Chemical structure of Molécula 142 (P1''). It features a central pyridine ring connected to a benzimidazole system on the left and a 2-aminoethyl chain on the right, which is further linked to another benzimidazole system.</p>	 <p>Chemical structure of Molécula 143 (P1''). It is identical to Molécula 142 (P1'').</p>	 <p>Chemical structure of Molécula 141 (P1''). It consists of a benzimidazole system connected to a 2-aminoethyl chain, which is further linked to a pyridine ring, which is in turn connected to a quinoline system.</p>
Molécula 115 (P1'')	Molécula 128 (P1'')	Molécula 103 (P1'')
 <p>Chemical structure of Molécula 115 (P1''). It features a central benzimidazole system with a nitrile group, an amino group, and a methoxy group, and is connected to a benzene ring and a benzimidazole system.</p>	 <p>Chemical structure of Molécula 128 (P1''). It features a central benzimidazole system with a methoxy group, a hydroxyl group, and a piperazine ring, and is connected to a benzene ring.</p>	 <p>Chemical structure of Molécula 103 (P1''). It features a central benzimidazole system with a hydroxyl group, a methoxy group, and a piperazine ring, and is connected to a benzene ring and a piperazine ring.</p>
Molécula 133 (P1'')	Molécula 134 (P1'')	Molécula 130 (P1'')
 <p>Chemical structure of Molécula 133 (P1''). It features a central benzimidazole system with a dimethylamino group, a hydroxyl group, and a piperazine ring, and is connected to a benzene ring and a piperazine ring.</p>	 <p>Chemical structure of Molécula 134 (P1''). It features a central benzimidazole system with a dimethylamino group, a hydroxyl group, and a piperazine ring, and is connected to a benzene ring and a piperazine ring.</p>	 <p>Chemical structure of Molécula 130 (P1''). It features a central benzimidazole system with a dimethylamino group, a hydroxyl group, and a piperazine ring, and is connected to a benzene ring and a piperazine ring.</p>

Tabela 17. Estrutura das moléculas selecionadas na análise de duplo filtro dos protocolos 1 e 2 (P1'' e P2''). Protocolos distinguidos por coloração em sombreamento : P1'' (verde) e P2'' (ciano).

(continuação)

Molécula 136 (P1'')	Molécula 156 (P2'')	Molécula 155 (P2'')
 <p>Chemical structure of Molécula 136 (P1''). It features a central benzothiazole ring system. One benzene ring of the thiazole is substituted with a 4-hydroxyphenyl group. The other benzene ring is substituted with a hydroxyl group and a carbonyl group. This carbonyl group is further substituted with a 4-(1-(2-isopropylpiperidin-4-yl)phenyl)phenyl group.</p>	 <p>Chemical structure of Molécula 156 (P2''). It consists of a 1,2,4-triazole ring substituted with a 2-(dimethylamino)ethyl group, a cyano group, and a methylene group. This methylene group is connected to a 4-(2-(4-(4-aminopyrrolidin-2-yl)phenoxy)pyridin-2-yl)pyridin-2-yl group.</p>	 <p>Chemical structure of Molécula 155 (P2''). It features a 1,2,4-triazole ring substituted with an amino group and a 2-(4-aminocyclohexyl)ethyl group. The other substituent is a 2-(4-(4-(4-fluorophenoxy)butyl)amino)ethyl group.</p>
Molécula 154 (P2'')	Molécula 152 (P2'')	Molécula 151 (P1'')
 <p>Chemical structure of Molécula 154 (P2''). It features a pyridine ring substituted with an amino group, a methyl group, and a 4-(4-(4-aminocyclohexyl)phenyl)phenyl group.</p>	 <p>Chemical structure of Molécula 152 (P2''). It consists of a pyridine ring substituted with a hydroxyl group, a methyl group, and a 2-(4-(4-(4-aminocyclohexyl)phenyl)phenoxy)ethyl group.</p>	 <p>Chemical structure of Molécula 151 (P1''). It features a piperidine ring substituted with a 2-(4-(4-(4-aminocyclohexyl)phenyl)phenoxy)ethyl group and a 2-(4-(4-(4-aminocyclohexyl)phenyl)phenoxy)ethyl group.</p>
Molécula 150 (P2'')	Molécula 149 (P2'')	Molécula 147 (P1'')
 <p>Chemical structure of Molécula 150 (P2''). It features a piperidine ring substituted with a 2-(4-(4-(4-aminocyclohexyl)phenyl)phenoxy)ethyl group and a 2-(4-(4-(4-aminocyclohexyl)phenyl)phenoxy)ethyl group.</p>	 <p>Chemical structure of Molécula 149 (P2''). It consists of a piperidine ring substituted with a 2-(4-(4-(4-aminocyclohexyl)phenyl)phenoxy)ethyl group and a 2-(4-(4-(4-aminocyclohexyl)phenyl)phenoxy)ethyl group.</p>	 <p>Chemical structure of Molécula 147 (P1''). It features a pyridine ring substituted with a 2-(4-(4-(4-aminocyclohexyl)phenyl)phenoxy)ethyl group and a 2-(4-(4-(4-aminocyclohexyl)phenyl)phenoxy)ethyl group.</p>

Finalmente, as 18 substâncias selecionadas foram submetidas a dinâmicas moleculares inicialmente de 100 ns e posteriormente de 500 ns, a fim de se avaliar a estabilidade de interação na protease, sob o mesmo protocolo utilizado anteriormente com o ligante de origem (NDL). Essas diferenças nos tempos de simulações são importantes para avaliar eventos de dobramento e desdobramentos proteicos, bem como o quão divergente são as interações de um complexo, quando comparados ao ligante de origem e moléculas (potenciais inibidores) de interesse em um alvo. Assim, é possível observar tais variações em diferentes condições experimentais, podendo ou não refletir estruturalmente no alvo e, conseqüentemente, na estabilidade de interações do complexo (KRESTEN LINDORFF, 2011).

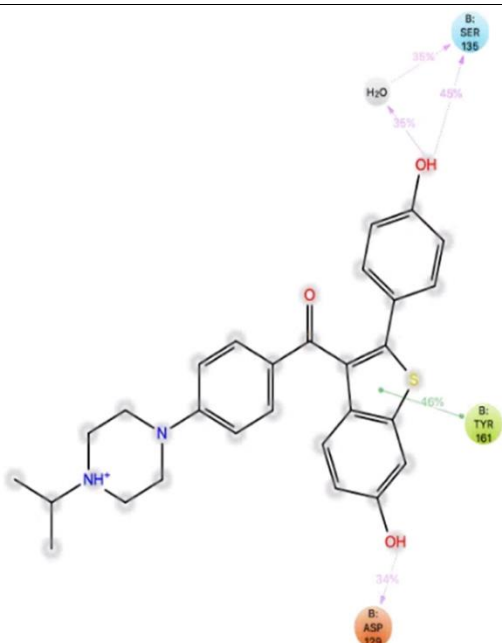
5.1.15 Simulações de DM dos *hits* e seleção dos candidatos

Os 18 potenciais ligantes foram analisados com a simulação dos movimentos de átomos a nível molecular, analisando possíveis alterações estruturais proteicas no alvo e/ou conformacionais dos ligantes, bem como as interações entre os complexos ligante-alvo. Dessa forma, foi possível corroborar os pontos de interação observados nos protocolos de *crossdocking* (P1, P2, P1'' e P2'') e ainda determinar o percentual de interação de cada aa da protease com as substâncias avaliadas nas DM.

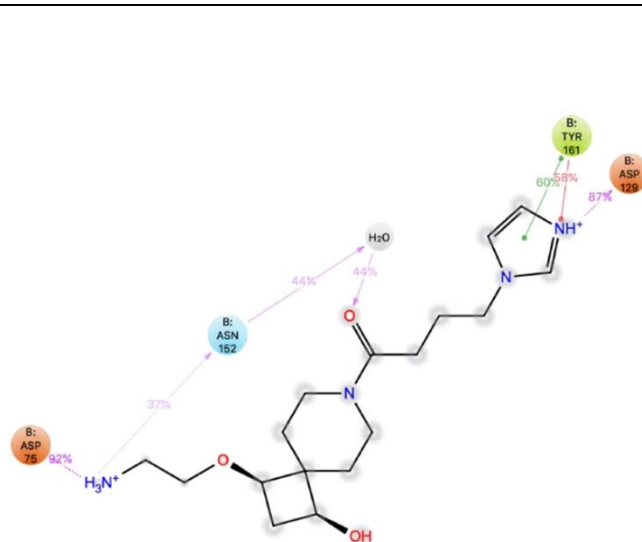
Os aa Val36, His51, Lys73, Asp75, Met84, Asp129, Phe130, Ser135, Gly151, Gly 153, Asn152, Tyr161 foram evidenciados com (IF>1,0). Exceto Val36, Lys73 e Met84, omitidos anteriormente na **Tabela 17**, todas as interações corroboram com os resultados observados no *crossdocking*, bem como àquelas interações preditas nas DM do ligante de origem (peptídeo) da protease de DENV-3, além daquelas respectivas às interações dos compostos peptidomiméticos 62 e 65. Na primeira dinâmica (100 ns), com objetivo de filtrar ligantes menos estáveis, dentre as 18 substâncias, apenas nove, sendo seis do MolPort (**Tabela 18**) e três do DrugBank (**Tabela 19**), se mostraram com alta estabilidade de interação ao longo de toda simulação. Submetidas à segunda DM (500 ns), as seis substâncias oriundas do MolPort e apenas uma do DrugBank mantiveram estabilidade da interação com a protease e foram selecionadas para as etapas seguintes (**Tabela 18**).

Tabela 18. Simulações das substâncias do banco MolPort selecionadas nas DM. Simulações de 100 ns. As figuras representam os esquemas de contatos de átomos de cada substância nos aa da protease de DENV-3 (3U1I). Interações que ocorrem em tempo superior a 30% da simulação são representadas. Linhas de tendência de interação: ligação de hidrogênio ou polar (azul), iônica ou carga negativa (vermelho) e hidrofóbica (verde).

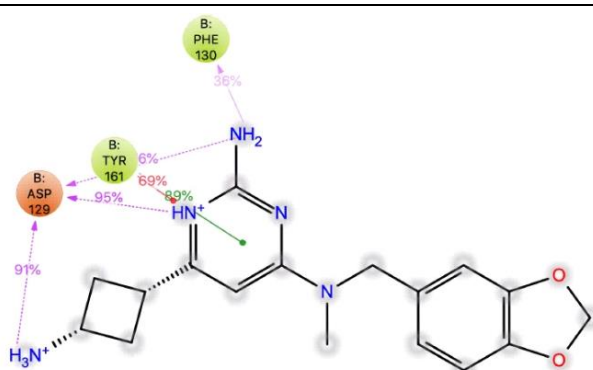
Substância 136



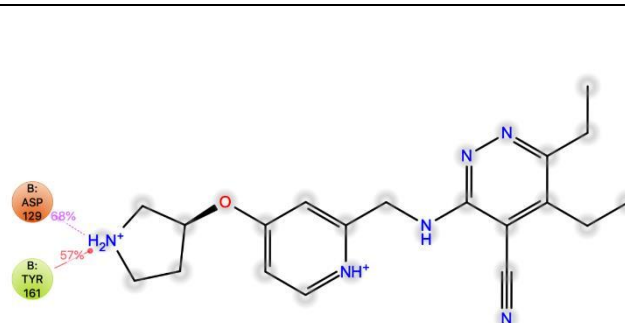
Substância 151



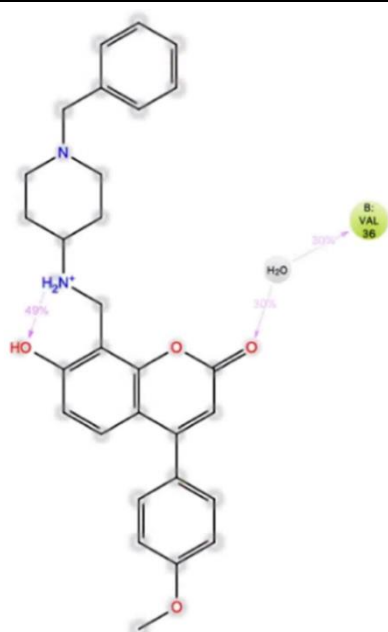
Substância 154



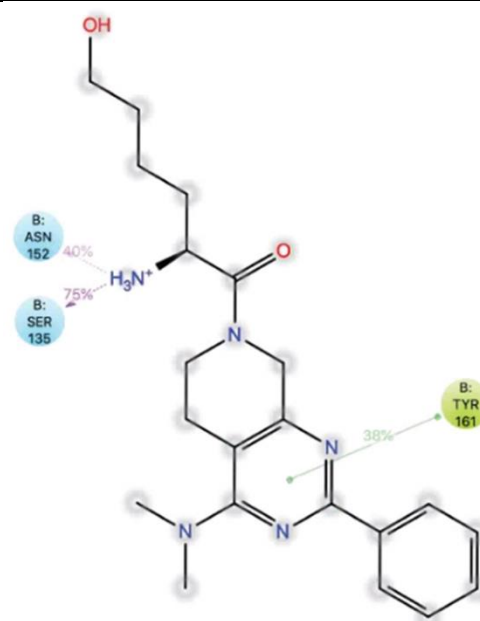
Substância 156



Substância 128 (P1')

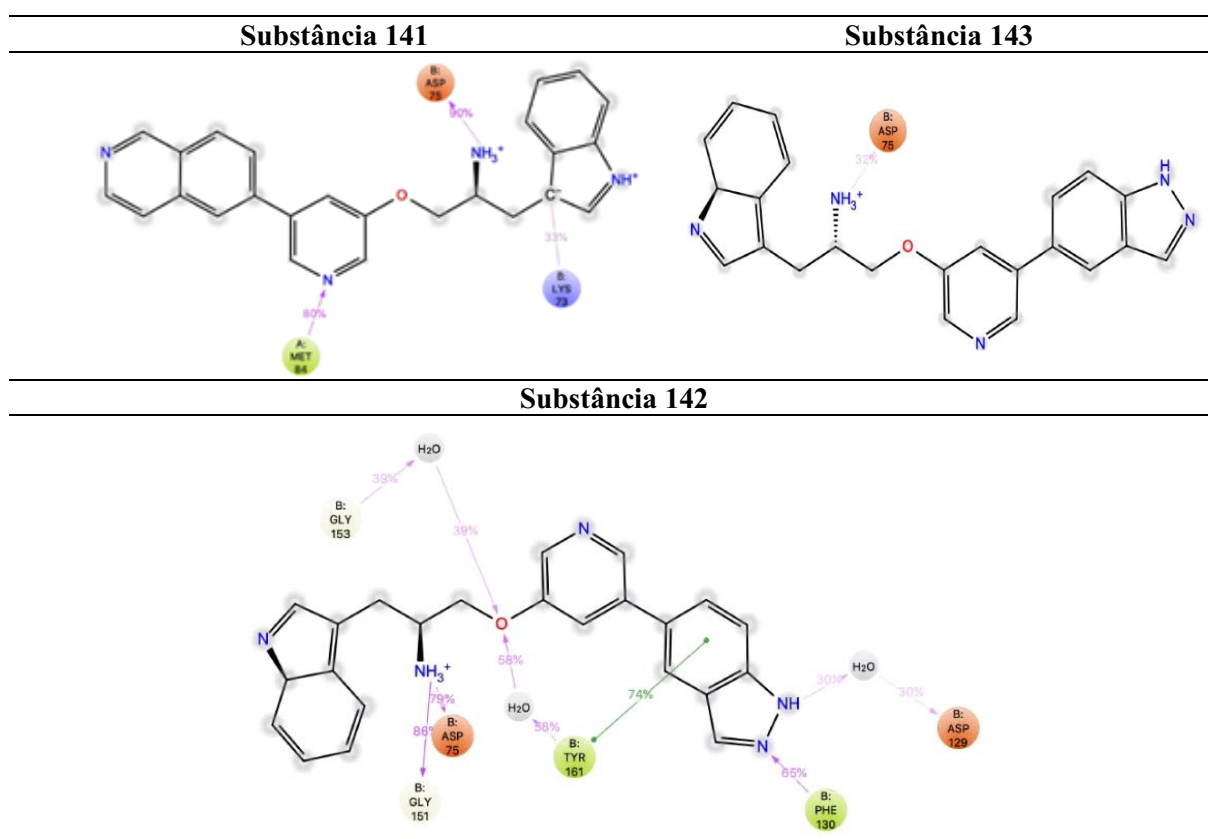


Substância 134 (P1')



Embora as três substâncias do DrugBank (**Tabela 19**) tivessem apresentado boa estabilidade na primeira simulação de 100 ns, ao serem replicadas a 500 ns, as substâncias 662 e 663 se deslocavam consideravelmente do sítio. Esses dados observados reforçam a importância de variações das condições de simulações nas previsões de interações no complexo ligante-alvo, como já discutido previamente (LINDORFF-LARSEN et al., 2011). A substância 653, por sua vez, embora apresentasse resultados estáveis em ambas dinâmicas, viria a ser reprovada nas previsões de toxicidade (imunotoxicidade).

Tabela 19. Simulações das substâncias do banco DrugBank selecionadas nas DM. Simulações de 100 ns. As figuras representam os esquemas de contatos de átomos de cada substância nos aa da protease de DENV-3 (3U1I). Interações que ocorrem em tempo superior a 30% da simulação são representadas. Linhas de tendência de interação: ligação de hidrogênio ou polar (azul), iônica ou carga negativa (vermelho) e hidrofóbica (verde).



Encerrada a triagem virtual e análises complementares dos *hits* obtidos, somente seis substâncias, dentre mais de sete milhões foram selecionadas para compra. Deste ponto, cada substância recebeu um código de identificação, sendo: **128; 134; 136; 151; 154; e 156.**

Todas as substâncias são sintéticas e, dentre elas, **151**, **156**, **134**, **154** e **128** não apresentam, até o momento, quaisquer atividades descritas na literatura. A substância **136**, no entanto, já foi descrita e validada como modulador de receptor de estrogênio (YANG; SHEN, 2005). Dentre as 12 substâncias triadas do DrugBank, dois fármacos apresentavam um mesmo alvo molecular que a substância **136**. As substâncias apresentadas anteriormente (**Tabela 15**) sob os códigos **158** e **140**, correspondem a raloxifeno e bazedoxifeno, respectivamente. Ambos são fármacos comercializados como medicamentos na prevenção e tratamento de osteoporose em mulheres na menopausa, atuando como moduladores seletivos dos receptores de estrogênio (do inglês, *selective estrogen receptor modulator* – SERM) (ARUN; ANTHONY; DUNN, 2002; YANG; SHEN, 2005).

Devido a sua facilidade de obtenção e por já terem ultrapassado etapas de estudos clínicos, ambos foram incluídos no estudo a fim de se comparar potenciais atividades inibitórias junto a substância **136** contra os três flavivírus de estudo (DENV, YFV e ZIKV). Ademais, raloxifeno apresentou atividade antiviral contra DENV-1 e 3, além de YFV e HCV em uma triagem fenotípica, sugerindo inibições em processos de fusão durante o ciclo de multiplicação destes vírus (MAZZON et al., 2019). Curiosamente, neste estudo de Mazzon et al. (2019), somente para ZIKV a atividade deste fármaco não foi determinada, deixando uma lacuna de interesse a ser investigada.

Portanto, incluindo raloxifeno (**158**) e bazedoxifeno (**140**), oito substâncias foram selecionadas para os modelos de predição de biodisponibilidade, toxicidade e atividade inibitória, bem como validações experimentais quanto às suas atividades biológicas *in vitro*. Paralelamente, as outras 12 substâncias desconsideradas também foram avaliadas nestas predições durante a execução das dinâmicas, sendo incluídas a seguir.

5.2 Predições das atividades biológicas

5.2.1 Cardiotoxicidade

Paralelamente às análises de dinâmica, as 18 substâncias triadas nos protocolos P1'' e P2'' foram submetidas a predições de toxicidade e biodisponibilidade, visando obter maiores informações preditivas dos *hits* e potencialmente excluir candidatos desfavoráveis, como perfis farmacocinéticos inadequados ou toxicidade elevada (PATRICK, 2005; WERMUTH, 2008). No primeiro software utilizado, Pred-hERG 4.2, um modelo de QSAR-2D convencional de cada substância é realizado para delimitar mapas de contribuição de átomos ou fragmentos

possivelmente capazes de bloquear canais para potássio em células cardíacas e, conseqüentemente, avaliem a possibilidade de toxicidade seletiva neste tecido.

Esses canais iônicos são importantes na transmissão de sinais elétricos que controlam contrações do miocárdio. Substâncias que inibem a proteína codificada por hERG são potenciais cardiotoxícos, podendo interferir na repolarização do potencial elétrico das células e transmissão desse sinal, causando problemas, por exemplo, de arritmia cardíaca (SANGUINETTI; TRISTANI-FIROUZI, 2006). O FDA, por exemplo, exige ensaios de inibição das proteínas oriundas da hERG para potenciais fármacos a serem registrados (FDA, 2005).

Os parâmetros utilizados nessa técnica apresentam uma contribuição positiva no modelo, em que cada nova substância submetida ao cálculo promove um aprendizado no servidor, contribuindo para um banco de dados agrupando cada contribuição de cada nova substância avaliada, aprimorando a sensibilidade da predição. Dessa forma, serão apresentados dois resultados possíveis: (i) binário, que resulta no percentual do mapa de contribuição individual da substância avaliada como um potencial bloqueador; e (ii) multi-classe, gerando um percentual comparativo da substância às substâncias já submetidas ao servidor, avaliando-a como potencial bloqueador ou não bloqueador.

Estes modelos são comparados a um banco de substâncias bioativas associadas com alguma atividade inibitória (bloqueio) ou interações conhecidas em canais para potássio, especialmente hERG, sendo 11.958 substâncias orgânicas somadas a fármacos, produtos naturais isolados e estruturas sob investigação, atualizadas periodicamente (BRAGA et al., 2014, 2015). No período das análises (maio/2019), constavam 16.932 substâncias no banco de dados.

O modelo prevê que quaisquer substâncias que superem 90% em um ou ambos os parâmetros (binário ou multi-classe), devem ser desconsideradas do um grupo de estudo, devido à possibilidade de toxicidade cardíaca. Dentre as 18 substâncias triadas nos protocolos P1'' e P2'', nenhuma foi predita como potencial bloqueador (>90%) na predição binária, enquanto todas foram preditas como não bloqueador na predição multi-classe (**Tabela 20**).

Tabela 20. Percentual predito das substâncias triadas e selecionadas nos parâmetros do servidor Pred-hERG. Código das substâncias triadas: três dígitos finais (MolPort) ou letras (*hits* selecionados). Predição binária e multi-classe em percentuais (%) de predição. Valores superiores a 90% sugerem potencial bloqueador de hERG.

MolPort	Binário (bloqueador)	Multi-classe (não bloqueador)	8 hits	Binário (bloqueador)	Multi-classe (não bloqueador)
113	70%	70%	134	70%	70%
155	70%	60%	151	70%	60%
152	70%	70%	154	70%	70%
130	70%	70%	156	70%	70%
115	70%	70%	136	60%	70%
149	70%	70%	128	70%	60%
150	70%	70%	140	80%	60%
133	70%	70%	158	70%	70%
147	70%	70%			
142	70%	70%			
143	70%	70%			
141	70%	70%			

Essa ferramenta já foi utilizada de forma complementar em outros estudos que envolviam estratégias de descoberta e desenvolvimento de substâncias bioativas, como o trabalho de Lima et al. (2018), que identificou antimaláricos empregando uma estratégia baseada em QSAR e análises de *docking*, além da validação experimental (LIMA et al., 2018). Paixão et al. (2019), por sua vez, também utilizaram Pred-hERG na identificação e avaliação produtos naturais como inibidores da redutase de tripanotiona de *Trypanosoma cruzi*, mais uma vez permitindo avaliar o potencial cardiotoxico de substâncias bioativas, prévio aos modelos *in vivo* (DA PAIXÃO; PITA, 2019).

Apesar dos avanços nas predições de modelos de aprendizado de máquina e QSAR, como Pred-hERG, não há uma garantia de que novas substâncias bioativas e potenciais inibidores não induzam alguma toxicidade celular específica. Nesse aspecto, o contínuo desenvolvimento e aprimoramento de métodos alternativos, especialmente *in silico*, são importantes nos processos de descoberta ou desenvolvimento de fármacos. Da mesma maneira, a expansão de dados de toxicidade disponíveis em bancos de dados computacionais e diferentes estratégias utilizadas em conjunto, também favorecem triagens e predições de substâncias (ALVES et al., 2018).

Além de predições específicas como hERG, também foi realizado um modelo de predição de toxicidade oral, submetendo as substâncias às análises no servidor ProTox-II. Neste

modelo, parâmetros como potenciais de carcinogenicidade, citotoxicidade, hepatotoxicidade, imunotoxicidade, mutagenicidade e valores teóricos de dose letal mediana (do inglês, *lethal dose* – LD50) são preditos baseados em modelos de aprendizagem de bibliotecas de dados, como substâncias que interagem em vias de sinalização e receptores, fármacos e compostos testados *in vitro* e *in vivo*, painéis de segurança de toxicidade como FDA, NIH e Novartis. São preditas classes (ativa ou inativa) para determinado alvo (toxicidade a um tecido ou receptores específicos), com uma probabilidade calculada entre similaridade da substância com o banco de dados, determinando também a acurácia de predição do modelo (BANERJEE et al., 2018).

5.2.2 Toxicidade oral

Nas predições feitas no ProTox-II, nenhuma das 18 substâncias apresentou potencial mutagênico, carcinogênico, citotóxico ou hepatotóxico. Esta baixa probabilidade de efeitos citotóxicos e hepatotóxicos são promissores, especialmente sabendo do tropismo hepático nas infecções por flavivírus (COUVELARD et al., 1999; REYES-DEL VALLE et al., 2014). Nenhuma das substâncias também apresentou valores de DL50 de classe V ou inferior. O servidor classifica esse parâmetro de acordo com o grau de toxicidade previsto no Sistema Globalmente Harmonizado de Classificação e Rotulagem de Produtos Químicos (do inglês, *Globally Harmonized System* – GHS), de I a VI, em que classe I indica que uma substância é fatal se ingerida (DL50 < 5,0 mg/kg), até VI, na qual o composto é considerado não tóxico (DL50 > 5.000 mg/kg) (COUVELARD et al., 1999).

A maioria das substâncias apresentou baixo grau de periculosidade (classes IV, V e VI), o que sugere uma baixa toxicidade oral (DRWAL et al., 2014). Contudo, dentre as 18 substâncias triadas, além de bazedoxifeno e raloxifeno, a substância 141 (DrugBank) apresentou uma predição ativa de imunotoxicidade, com acurácia de 97%. Além dessa, foi predita toxicidade para o receptor de hidrocarboneto arílico (RHA), com 66% de acurácia.

Os resultados, quando preditos como ativos para determinados alvos, também são representados em um modelo de radar, demonstrando graficamente a probabilidade de toxicidade de uma substância, a exemplo da substância 141. Havendo uma probabilidade de toxicidade superior àquela da média geral do painel do banco de dados, a substância deve ser rejeitada ou cautelosamente analisada nos ensaios *in vitro* e *in vivo* (BANERJEE et al., 2018).

A imunotoxicidade de uma substância ou composto químico condiz em uma redução da resistência imune de um indivíduo, especialmente frente a infecções ou exposição à antígenos

alergênicos. Esse contato com uma substância potencialmente imunotóxica, por exemplo, um xenobiótico, pode ocasionar efeitos histopatológicos diversos no sistema imune e até danos teciduais, por exemplo, alteração na maturação e função de células imunocompetentes, variação na população de linfócitos, reações imunes inespecíficas e falha nas respostas imunológicas (GERMOLEC et al., 2017).

Paralelamente, o receptor de hidrocarboneto arílico (RHA) atua como fator de transcrição que associam em um indivíduo, mediante interações por ligantes, os efeitos do ambiente, dieta, microbiota e metabolismo. Neste contexto, as interações de substâncias com AHR influenciam diretamente no organismo, como no desenvolvimento e progressão de doenças, além de perturbar o sistema imune, podendo estar associado a doenças auto-imunes, neoplasias e acometimentos neurológicos, além de disfunções imunológicas e de controle do trato digestório. Essas substâncias usualmente são xenobióticos, como contaminantes do ambiente, pesticidas, metabólitos, antígenos microbianos e alergênicos, além de fármacos e produtos naturais (LAMAS; NATIVIDAD; SOKOL, 2018; ROTHHAMMER; QUINTANA, 2019). Nesse sentido, também se evidencia por essa predição ativa o risco de toxicidade celular e tecidual, corroborando a predição de imunotoxicidade e, conseqüentemente, reforçando a desconsideração da substância 653.

Essa abordagem com o programa ProTox-II também já foi utilizada em estudos *in silico* por Supandi et al. (2018), em que derivados de aminoquinazolinás foram avaliados para predição de toxicidade. Esses derivados já haviam sido confirmados como inibidores do fator de crescimento de fibroblastos em células tumorais de leucemia, câncer gástrico, renal e pulmonar. As predições permitiram avaliar comparativamente e de forma complementar com outros servidores e programas, aprimorando uma triagem computacional da toxicidade dessas substâncias (LAZAR; PREDICTOR SUPANDI; MERDEKAWATI, 2018).

Karim et al. (2019), por outro lado, compararam diversos servidores de predição de toxicidade junto a modelos de aprendizado de máquina, como árvores de decisão, para um mesmo grupo de moléculas. Nas análises, ProTox-II apresentou um desempenho mais robusto, com valores superiores a 0,9 nas curvas ROC (Característica de Operação do Receptor) de validação e a quarta maior média geral de todas as curvas ROC (0,85) comparada dentre os 21 modelos (0,593 a 0,862) avaliados nas predições (KARIM et al., 2019).

5.2.3 Características farmacocinéticas (biodisponibilidade)

Além da toxicidade ser um fator de relevância na seleção de *hits* selecionados, suas propriedades físico-químicas em relação à biodisponibilidade, especialmente via oral, também são importantes para garantia do sucesso no desenvolvimento de um fármaco. Pensando nessa proposta, as propriedades referentes às características farmacocinéticas (biodisponibilidade) das oito substâncias selecionadas na triagem foram preditas empregando um modelo de QSAR no servidor admetSAR (<http://lmmd.ecust.edu.cn/admetSAR2/>), correlacionando cada propriedade dos *hits* selecionados a um banco de milhares de substâncias (YANG et al. 2018).

No modelo avaliado, quatro parâmetros foram preditos, sendo: (i) absorção intestinal; (ii) permeabilidade na B.B.B.; (iii) biodisponibilidade via oral; e (iv), localização intracelular. Nos resultados, todas as substâncias foram preditas positivamente para possibilidade de serem absorvidas no trato intestinal (86,52 a 99,3%) e em atravessarem a barreira hematoencefálica (96,97 a 98,12%). Quanto à biodisponibilidade oral, somente as substâncias **134**, **136** e **128** foram preditas positivamente, com valores de probabilidade entre 51,43 e 62,86%. Os dados de predição por admetSAR estão dispostos na **Tabela 21**.

Tabela 21. Parâmetros farmacocinéticos preditos pelo servidor admetSAR para as oito substâncias selecionadas. B.B.B. Barreira hemato-encefálica; + predição positiva; – predição negativa.

Substâncias	admetSAR [Condição/Probabilidade (%)]					
	Absorção intestinal	B.B.B.	Biodisponibilidade oral			
128	+	86.52	+	98.00	+	58.57
134	+	97.18	+	96.97	-	61.43
136	+	99.28	+	98.12	+	51.43
151	+	95.86	+	96.98	+	62.86
154	+	97.41	+	97.71	-	64.29
156	+	99.09	+	98.05	-	51.43
140	+	99.07	+	96.37	-	76.87
158	+	99.30	+	97.44	-	88.57

Nesses resultados, dentro os parâmetros avaliados, a predição de absorção intestinal e biodisponibilidade oral, de forma qualitativa ou quantitativa, por meio de modelos de análise estrutural ou por QSAR são estratégias favoráveis na descoberta e desenvolvimento de fármacos (YANG et al., 2018). A predição de permeabilidade na barreira hematoencefálica (B.B.B.), por sua vez, também pode ter implicações importantes nessas etapas.

A B.B.B., uma estrutura de capilares sanguíneos permeável presente SNC e responsável pela proteção contra agentes de diversas etiologias, como substâncias neurotóxicas ou

patógenos, avaliar a capacidade de permeabilidade nesta barreira pode auxiliar na avaliação de possíveis fármacos atuantes no SNC ou prever uma provável neurotoxicidade das substâncias (DANEMAN; PRAT, 2015). Todas as substâncias avaliadas apresentaram predições positivas com alta probabilidade de atravessarem a barreira hematoencefálica (> 96,97%), o que pode ser interessante no desenvolvimento de antivirais contra os flavivírus em estudo, devido à possibilidade da presença destes vírus no sistema nervoso central em infecções, como nos casos de meningoencefalites (SILVA MARINHO; KROON, 2019).

Já o modelo para predição de absorção no tecido intestinal humano é baseado em valores experimentais de substâncias avaliadas *in vivo*, comparando descritores moleculares utilizados no modelo de QSAR, de acordo com os dados de excreção de vias comuns (bile, fezes e urina). Os valores de predição indicam uma baixa (0 a 20%), média (20 a 70%) e alta probabilidade de absorção (70 a 100%) (ZHAO et al., 2001). Todas as substâncias avaliadas apresentaram alta probabilidade de serem absorvidas no trato intestinal (> 86,52%). Esses resultados corroboram a lipofilia das estruturas e reforçam a discussão acerca das regras de Lipinski, em que substâncias mais apolares são capazes de atravessar a membrana celular por transporte passivo, facilitando sua absorção (LIPINSKI, HOPKINS, 2004).

Esses resultados também podem reforçar aqueles observados para disponibilidade oral, uma vez que uma alta lipofilia resultaria em baixa solubilidade em água e, por consequência, baixa biodisponibilidade nesta via. O fato do filtro para predição dessa propriedade farmacocinética em substâncias bioativas se limitar somente a essa via de administração, acaba excluindo substâncias que podem ser administradas em outras vias, como intraperitoneal e intramuscular. Alguns fármacos baseados em produtos naturais já aprovados pelo FDA não seguiam, por exemplo, essa regra proposta por Lipinski, o que não excluiu tais candidatos e/ou impediu seu desenvolvimento a fármacos (NEWMAN; CRAGG, 2016).

Veber et al. (2002), também mencionam, por outro lado, a importância da área de superfície polar e o número de ligações rotacionáveis de substâncias, associados à permeabilidade de substâncias em membranas celulares. Valores maiores que 10 ligações rotacionáveis estipulados empiricamente podem estar associados a uma dificuldade na absorção e, conseqüentemente, na biodisponibilidade oral de substâncias.

5.2.4 Predição da atividade biológica

Finalmente, finalizadas as predições de toxicidade e biodisponibilidade, uma predição de atividade biológica (inibição enzimática da protease) somente das oito substâncias finais foi realizada empregando diferentes modelos de aprendizado de máquina no pacote computacional KNIME®, buscando avaliar, computacionalmente, se as substâncias preditas no alvo teriam potencial atividade inibitória *in vitro*. Dentre todos os modelos avaliados e parâmetros utilizados, apenas três superaram os valores das métricas de validação (maior que 0,6): kNN considerando massas moleculares (kNNw; cvMCC = 0,725 e extMCC = 0,825), RF (cvMCC = 0,638 e extMCC = 0,734) e SVM polinomial (cvMCC = 0,734 e extMCC = 0,631). Estes três modelos foram selecionados para as predições de atividade inibitória para separação entre moléculas preditas como ativas e inativas.

As curvas ROC dos três MLT selecionados superam 0,8, tendo o modelo de floresta aleatória (RF) apresentado o maior valor de curva ROC (0,976). Contudo, somente a predição empregando SVM polinomial apresentou cinco substâncias classificadas como ativas (classe 1): **136** (51,6%), **158** (59%), **140** (59,6%), **151** (59,9%) e **154** (76,2%). Novamente, os três moduladores de receptor de estrogênio (**136**, **140** e **158**) foram preditos para uma mesma característica, neste caso, em uma mesma classe (ativos) para atividade biológica. A substância **156**, por sua vez, não resultou em quaisquer predições no modelo de SVM empregado. Já os modelos kNNw e RF predizeram todas as substâncias como inativas (classe 0), com valores variando entre 69,4 e 80% e entre 60,8 a 97,3%, respectivamente (**Tabela 22**).

Tabela 22. Predição de atividade biológica das oito substâncias: kNNw, SVM polinomial e floresta aleatória (RF). As substâncias são preditas em duas classes distintas em cada modelo: inativo (0) e ativo (1). Percentuais de predição de cada classe são representados em valores decimais.

Substâncias (VS)	kNNw			SVM polinomial			RF		
	Classe	0	1	Classe	0	1	Classe	0	1
128	0	0,790	0,210	0	0,837	0,163	0	0,972	0,028
134	0	0,771	0,229	0	0,579	0,421	0	0,963	0,037
136	0	0,694	0,306	1	0,484	0,516	0	0,608	0,392
151	0	0,709	0,291	1	0,401	0,599	0	0,963	0,037
154	0	0,764	0,236	1	0,238	0,762	0	0,973	0,027
156	0	0,800	0,200	-	-	-	0	0,956	0,044
140	0	0,794	0,206	1	0,404	0,596	0	0,934	0,066
158	0	0,694	0,306	1	0,410	0,590	0	0,658	0,342

Modelos como o de floresta aleatória (RF), por sua vez, realizam combinações dos parâmetros aplicando algoritmos de aprendizado aos dados ou repetindo os mesmos algoritmos variando os parâmetros. Essa combinação de múltiplas decisões individuais, acerca das características avaliadas, pode ocorrer de diferentes maneiras, como por valores médios ou regras de ranqueamento por votação (POLIKAR, 2006). As máquinas de vetores suporte, por sua vez, consistem em gerar um hiperplano otimizado, ou seja, uma superfície de decisão em que se garante a melhor separação entre objetos (substâncias) de duas classes diferentes. Neste modelo, SVM projetam dados a um espaço e , em seguida, utilizam funções lineares para discriminação dos dados em classes distintas. Assim, há uma probabilidade de um dado (substância) novo ser diretamente alocado ao lado correspondente ao seu perfil ou característica principal, por exemplo, ativo e inativo (MALTAROLLO et al., 2019).

Maltarollo et al. (2019), já descreveram a aplicação de diversos modelos de aprendizado de máquina para inibidores e não-inibidores de uma enzima do metabolismo de ácidos graxos de *Staphylococcus aureus* (FabI). Dentre esses MLT, SVM e RF foram comparadas a outras técnicas de predições de classe, apresentando valores do coeficiente de correlação de Matthews (do inglês, *Matthews correlation coefficient* – MCC) superiores aos demais. MCC são calculados por matrizes de confusão, analisando taxas de verdadeiros e falsos positivos e negativos. Na ocasião, o modelo de RF também pode corroborar análises de QSAR que sugeriram grupos funcionais e sítios de interação com aa da enzima.

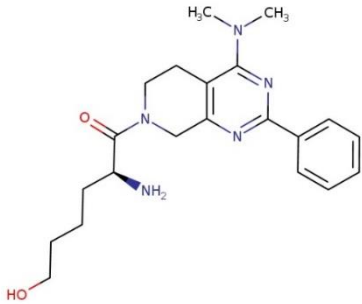
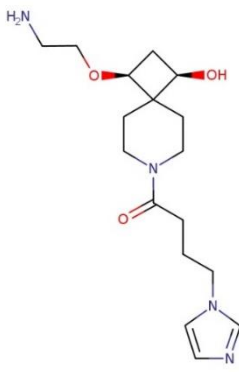
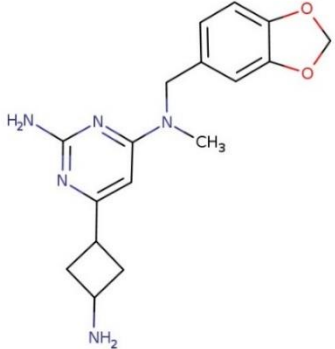
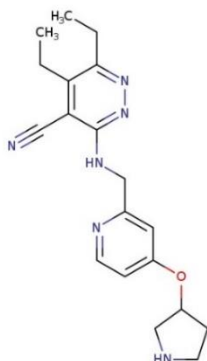
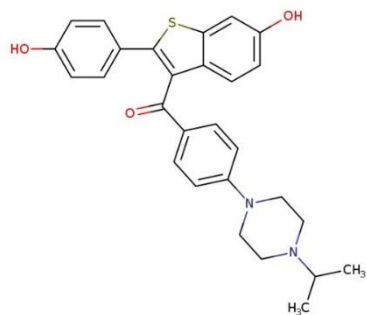
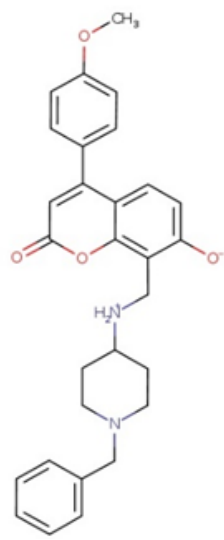
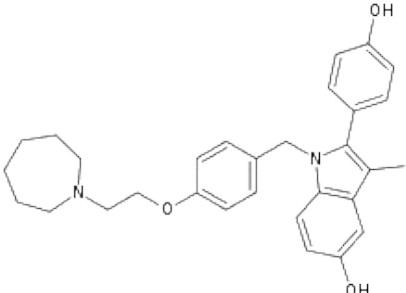
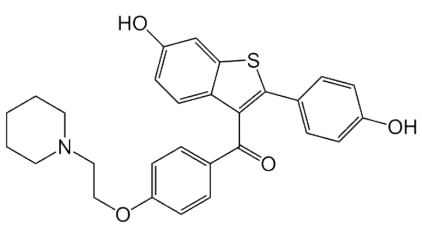
Embora kNN seja um importante modelo de classificação para predição de atividade biológica, sua função é mais lenta que os outros dois modelos, adiando os cálculos das classes até a obtenção dos resultados. Já o modelo de RF combina parâmetros em árvores de predição e avalia individualmente qual melhor categoriza substâncias em classes distintas, podendo ser utilizado para classificação de substâncias ou análises de regressão. SVM, por sua vez, representa cada conjunto de dados (substâncias) como pontos, construindo um hiperplano que divide os candidatos em diferentes categorias de acordo com suas características, aumentando a acurácia da predição, uma vez que evita agrupar amostras de uma mesma natureza (GERTRUDES et al., 2012; MALTAROLLO et al., 2015). Somente SVM foi capaz de prever substâncias classificadas como ativas (**136, 151, 154, 140 e 158**). Finalizadas as predições propostas, as oito substâncias selecionadas foram adquiridas e os ensaios para validação *in vitro* da potencial atividade antiviral foram realizados. Os resultados das atividades biológicas estão dispostos a seguir.

5.3 Ensaios biológicos

5.3.1 Avaliação da citotoxicidade das substâncias

Diversos métodos computacionais foram aplicados para a triagem e obtenção das oito substâncias (**Tabela 23**) a serem testadas contra as três espécies de flavivírus: DENV-2 (PI59), DENV-3 (PI76), YFV (17DD) e ZIKV (PE243).

Tabela 23. Estrutura química das oito substâncias selecionadas na triagem para avaliação da atividade biológica *in vitro*: 134, 151, 154, 156, 136, 128, 140 e 158.

Substância (134)	Substância (151)	Substância (154)
		
Substância (156)	Substância (136)	Substância (128)
		
Bazedoxifeno (140)	Raloxifeno (158)	
		

Nos ensaios de MTT em Vero e BHK-21 (100 até 1,56 μM), as substâncias **134**, **151**, **156** e **128** apresentaram os maiores valores da concentração citotóxica a 50% em Vero ($67,21 \pm 3,04$, >50 , $141,74 \pm 4,64$ e $47,59 \pm 2,9$ μM , respectivamente), enquanto **154**, **136**,

bazedoxifeno e raloxifeno, os menores valores ($<1,56$; $16,87 \pm 1,49$; $21,38 \pm 1,07$ e $17,3 \pm 1,67$ μM , respectivamente). Já na avaliação da CC_{50} em células BHK-21, as substâncias **134**, **151**, **156** e **128** também apresentaram os maiores valores ($48,77 \pm 8,0$; $35,95 \pm 1,33$; $106,71 \pm 7,76$ e $30,05 \pm 4,09$ μM , respectivamente), enquanto **154**, **136**, **140** e **158**, novamente, exibiram os menores valores ($1,67 \pm 0,03$; $9,5 \pm 0,66$; $11,14 \pm 0,34$ e $11,33 \pm 0,07$ μM , respectivamente). No geral, os resultados de CC_{50} em BHK foram similares, mas inferiores àqueles para células Vero.

5.3.2 Avaliação da concentração efetiva a 50% (CE_{50})

Logo após a avaliação da citotoxicidade, buscando avaliar a potencial atividade antiviral das substâncias, outro ensaio de MTT foi realizado sob as mesmas condições, utilizando diluições seriadas a partir dos valores de CC_{50} obtidos no ensaio anterior, adicionando agora uma suspensão viral ($\text{MOI} = 0,1$) de ZIKV, YFV, DENV-2 e DENV-3. Dentre as oito substâncias avaliadas, cinco demonstraram algum efeito de proteção (atividade antiviral/inibição) nas células infectadas, em pelo menos uma das três espécies de flavivírus.

Apenas a substância **136** foi ativa em ZIKV, com uma concentração efetiva a 50% de $11,86 \pm 0,43$ μM e índice de seletividade de 1,42 em células Vero. A avaliação em YFV, por sua vez, resultou em quatro candidatos com atividade antiviral, sendo: novamente a substância **136** ($\text{CE}_{50} = 4,79 \pm 0,17$ e $\text{IS} = 3,52$), além de **128** ($\text{CE}_{50} = 24,15 \pm 1,15$; $\text{IS} = 1,97$), **140** ($\text{CE}_{50} = 6,37 \pm 0,12$; $\text{IS} = 3,36$) e **158** ($\text{CE}_{50} = 4,62 \pm 0,13$; $\text{IS} = 3,74$), com estes índices de seletividade determinados também em células Vero. Nos resultados avaliando os dois sorotipos de DENV, novamente as substâncias **136** ($\text{CE}_{50} = 4,21 \pm 0,14$ e $\text{IS} = 2,26$) e **128** ($\text{CE}_{50} = 9,51 \pm 0,12$ e $\text{IS} = 3,16$) apresentaram atividade, agora em DENV-2. Por outro lado, a substância **156**, foi ativa somente em DENV-3 ($\text{CE}_{50} = 37,51 \pm 0,8$ e $\text{IS} = 2,85$).

Três das cinco substâncias (**136**, **140** e **158**) que apresentaram atividade antiviral nos ensaios biológicos haviam sido preditas como ativas contra a protease no modelo de SVM polinomial. Por outro lado, as substâncias **156** e **128** foram preditas como inativas nos três modelos avaliados (kNNw, RF e SVM), o que contrapõe aos resultados observados nos ensaios de MTT para YFV, DENV-2 e DENV-3. Os resultados de todas as atividades biológicas avaliadas nos ensaios de MTT estão resumidos na **Tabela 24**.

Tabela 24. Avaliação da atividade antiviral das oito substâncias selecionadas na triagem em três *Flavivirus*: *Zika virus* (ZIKV), *Yellow fever virus* (YFV) e *Dengue virus* (DENV).^aÍndice de seletividade. Corresponde à razão entre concentração citotóxica e concentração efetiva (CC₅₀/CE₅₀). NA: não ativa.

Substâncias	ZIKV (PE243)		YFV (17DD)		DENV-2 (PI59)			DENV-3 (PI76)		
	Vero	Vero (MOI 0,1)	Vero (MOI 0,1)	BHK-21	BHK-21 (MOI 0,1)	BHK-21 (MOI 0,1)	BHK-21 (MOI 0,1)	BHK-21 (MOI 0,1)		
	CC ₅₀ (µM)	CE ₅₀ (µM)	IS ^a	CE ₅₀ (µM)	IS ^a	CC ₅₀ (µM)	CE ₅₀ (µM)	IS ^a	CE ₅₀ (µM)	IS ^a
134	67,21 ± 3,04	NA	-	NA	-	48,77 ± 8,0	NA	-	NA	-
151	>50	NA	-	NA	-	35,95 ± 1,33	NA	-	NA	-
154	<1,56	NA	-	NA	-	1,67 ± 0,03	NA	-	NA	-
156	141,74 ± 4,64	NA	-	NA	-	106,71 ± 7,76	NA	-	37,51 ± 0,8	2,85
136	16,87 ± 1,49	11,86 ± 0,43	1,42	4,79 ± 0,17	3,52	9,5 ± 0,66	4,21 ± 0,14	2,26	NA	-
128	47,59 ± 2,9	NA	-	24,15 ± 1,15	1,97	30,05 ± 4,09	9,51 ± 0,12	3,16	NA	-
140	21,38 ± 1,07	NA	-	6,37 ± 0,12	3,36	11,14 ± 0,34	NA	-	NA	-
158	17,3 ± 1,67	NA	-	4,62 ± 0,13	3,74	11,33 ± 0,07	NA	-	NA	-
Ácido micofenólico	>100	NA	-	NA	-	<1,56	NA	-	NA	-
Ribavirina	>100	4,1 ± 0,35	>24,39	40,9 ± 5,25	>2,44	25,08 ± 0,19	NA	-	NA	-

Vero e BHK-21 são células de epitélio renal, embora oriundas de organismos modelo diferentes (macaco-verde africano e hamster-sírio, respectivamente). Os resultados de citotoxicidade foram similares nas duas linhagens celulares, com BHK-21 apresentando valores inferiores de toxicidade àqueles em Vero. Na literatura, resultados similares de CC_{50} foram demonstrados por Lani et al. (2015), nessas mesmas linhagens celulares ao avaliar o efeito citotóxico de flavonoides, evidenciando maior toxicidade destas substâncias em BHK-21 (305,4 e 537,3 $\mu\text{g/mL}$) do que em células Vero (425,1 e >1000 $\mu\text{g/mL}$, respectivamente) (LANI et al., 2015).

Arpornsuwan et al. (2006), também evidenciaram efeitos citotóxicos similares de extratos do fruto noni (*Morinda citrifolia* Linn.) em BHK-21 e Vero. Novamente, as substâncias avaliadas foram mais tóxicas nas células de murino (2,5 $\mu\text{g/mL}$) do que de primata não humano (3,0 $\mu\text{g/mL}$). Kováčik et al. (2012), avaliaram o efeito de diversos antimicrobianos macrolídeos em ambas as células, também evidenciando resultados similares, mas com maior toxicidade em BHK-21 do que em Vero, respectivamente: tilmicosina, 50 $\mu\text{g/mL}$ e 300 $\mu\text{g/mL}$; Tylosin, 500 $\mu\text{g/mL}$ e 750 $\mu\text{g/mL}$; e espiramicina 150 $\mu\text{g/mL}$ e 500 $\mu\text{g/mL}$ (ARPORNSUWAN; PUNJANON, 2006; KOVÁČIK et al., 2012).

A sensibilidade de BHK-21 comparada à Vero também já foi demonstrada em modelos de infecção, por exemplo, efeitos citopáticos observados em infecções por *Mycoplasma* spp., mais intensos no início de cultivo das monocamadas e significativamente maior comparado à Vero, com maiores percentuais de redução do crescimento celular e em diluições maiores da bactéria (NETTO et al., 2014). Esses achados corroboram aos efeitos citotóxicos mais intensos e frequentes das oito substâncias observados nas primeiras 24h de infecção em BHK-21 comparados às células Vero, por exemplo, de arredondamento e afilamento de células (dados não mostrados).

Vero e BHK-21 já haviam demonstrado permissividade similar à infecção por isolados brasileiros de ZIKV, com títulos produzidos de ambas linhagens variando entre 1 a 48 UFP/mL, sendo um isolado de Pernambuco, Brasil, produzindo maior título em ambas células (NIKOLAY et al., 2018). Paralelamente, BHK-21 e Vero foram demonstradas como bons modelos celulares para plataformas de isolamento de DENV-2 e DENV-4, sem apresentar diferenças significativas entre as linhagens (PHANTHANAWIBOON et al., 2018). Esses dados similares de toxicidade e permissividade corroboram a escolha destas linhas celulares como modelos nos ensaios deste trabalho.

Dentre as substâncias ativas contra esses vírus, a substância 136 apresentou atividade em ZIKV e YFV, o que poderia sugerir um possível efeito mult flavivírus. Dentre as quatro substâncias ativas para YFV, 136 também apresentou a segunda maior atividade ($CE_{50} = 4,79 \pm 0,17 \mu\text{M}$), com valor equivalente àquele do raloxifeno ($CE_{50} = 4,62 \pm 0,13 \mu\text{M}$) e semelhante ao obtido para bazedoxifeno ($CE_{50} = 6,37 \pm 0,12 \mu\text{M}$). Como já discutido anteriormente, a substância 136 já foi descrita e validada como modulador seletivo do receptor de estrogênio por Yang et al. (2005), mesmo alvo molecular dos fármacos raloxifeno e bazedoxifeno (DUNN, 2002). Esses dados poderiam sugerir um mecanismo de ação similar dentre essas três substâncias em YFV.

O trabalho citado de Mazzon et al. (2019), corrobora os dados observados, em que raloxifeno apresentou atividade antiviral contra DENV-1 e 3, além de YFV. Neste estudo foram sugeridas inibições em processos de fusão durante o ciclo de multiplicação destes vírus, o que não se descarta nos resultados até então observados nos ensaios de MTT. Por outro lado, não haviam sido relatados atividade antiviral em ZIKV, demonstrado pela primeira vez nos resultados obtidos neste trabalho, em que raloxifeno apresentou algum efeito inibitório na infecção por ZIKV, protegendo a células ao interferir em alguma etapa do ciclo de multiplicação desse vírus.

No entanto, esses resultados não foram equivalentes dentre todos os flavivírus. Possivelmente, a inserção do substituinte dimetil em uma das extremidades da substância pôde ser o diferencial que aferiu atividade da substância 136 em ZIKV e DENV-2, enquanto raloxifeno e bazedoxifeno não foram ativos nestes dois vírus. Já foi demonstrado que a inserção deste grupamento pode levar à formação de agregados moleculares em interações com estruturas de azometinas. Os substituintes terminais dimetil podem, nesse caso, interferir no arranjo de substâncias alinhadas nos eixos de cada interação, variando as conformações estruturais e permitindo novos arranjos moleculares (JINDO et al., 2015). Por outro lado, se evidencia a presença de um éter espaçador entre os grupos terminais azepano (bazedoxifeno) e piperidina (raloxifeno), diferenciando ambos de seu análogo sintético. Amicangelo et al. (2008), já haviam demonstrado que éteres dimetílicos, especialmente direcionados a um anel aromático, aumentam as forças de interação moleculares de uma substância, podendo ser explicado por potenciais eletrostáticos do éter e do anel (AMICANGELO et al., 2008).

Caso a protease fosse confirmada como alvo do mecanismo de ação dessas substâncias, empregando ensaios enzimáticos, os dados de Amicangelo et al. (2008) e Jindo et al. (2015)

poderiam sugerir a diferença de atividade observada entre DENV-2 e ZIKV com YFV, considerando, por exemplo algumas diferenças de aa entre os sítios da protease. Nesse caso, a variação na posição 155, entre Val155 (DENV e ZIKV) e Ile155 (YFV), poderia permitir um deslocamento das interações em YFV, com a presença de um maior potencial eletrostático, devido ao éter presente no bazedoxifeno e raloxifeno. Igualmente, limitaria as interações em ZIKV e DENV, em razão das possíveis variações nas conformações e arranjos estruturais da presença do substituinte dimetil.

Behnam et al. (2015), por outro lado, já haviam demonstrado uma redução significativa do título viral de DENV-2 e WNV ao avaliar a atividade antiviral dos compostos peptidomiméticos utilizados no início deste trabalho. Na tentativa de sustentar a evidência de uma característica mais específica entre a estrutura molecular e a atividade biológica, foi estabelecida uma correlação linear entre os resultados de atividade antiviral e atividade inibitória na protease, apresentando valores robustos do coeficiente de r^2 (0,945). No entanto, nenhuma conclusão foi observada, haja vista a diversidade de radicais nas substâncias e uma baixa variabilidade dos valores de EC_{50} , limitando as análises de QSAR. Uma análise experimental dos tempos de retenção das substâncias, por outro lado, pôde enfim sugerir uma relação do caráter lipofílico das substâncias e de uma sensibilidade à fração éter-benzil com a atividade antiviral, empregando uma técnica tempo de retenção em coluna de HPLC (GIAGNIS, 2007; BEHNAM, 2015). Esse mesmo perfil já havia sido proposto por Noble et al. (2012), em que o caráter lipofílico de ligantes resultava em interações mais favoráveis de ligantes e maiores valores de inibição em uma protease de DENV (GIAGNIS; TSANTILIKAKOULIDOU,).

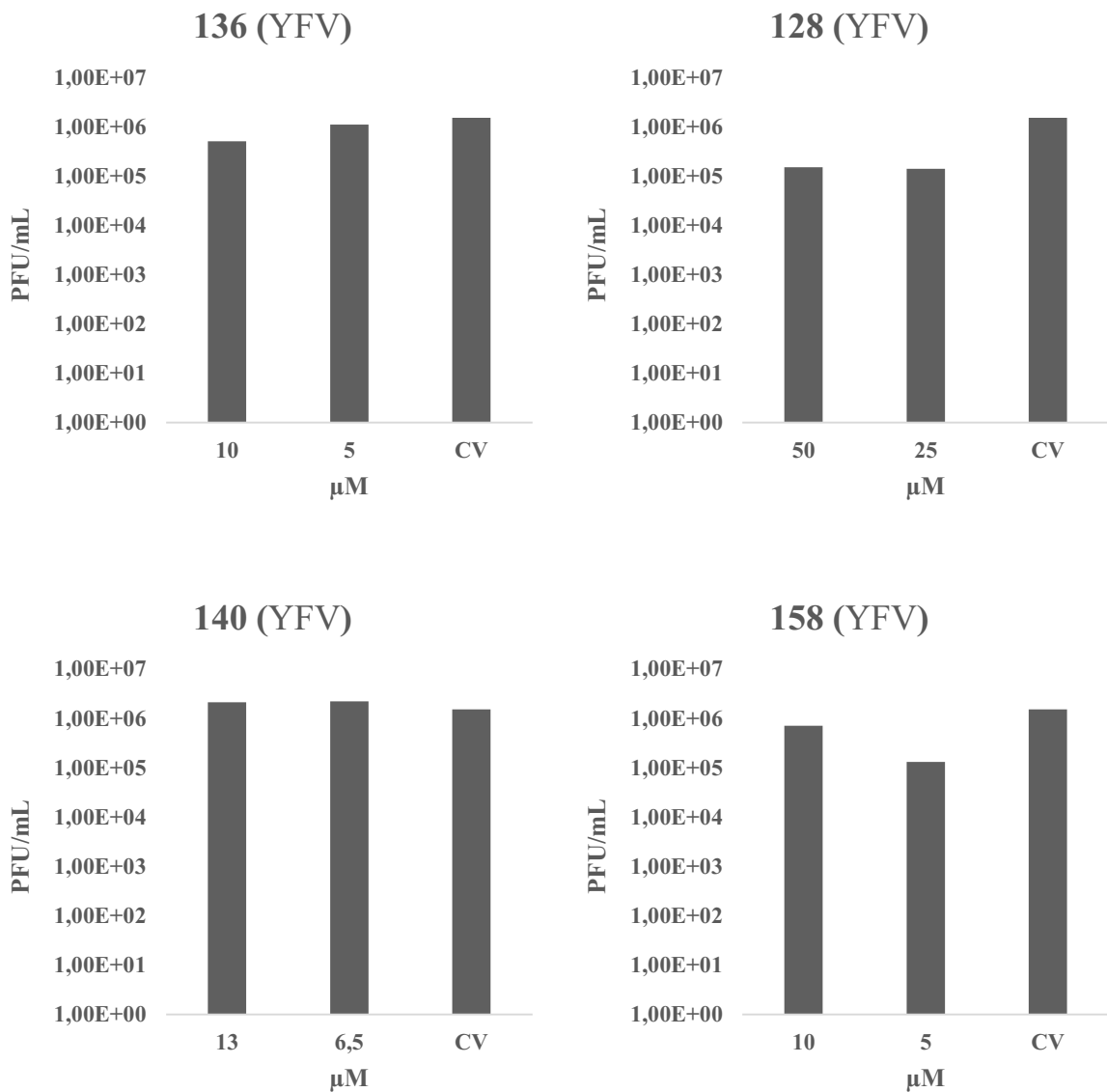
Cinco das oito substâncias avaliadas apresentaram atividade contra um ou mais flavivírus, resultados que corroboram dados da literatura, como a atividade de raloxifeno e o caráter lipofílico das substâncias citados acima. Ademais, essas substâncias também foram preditas positivamente para absorção intestinal e negativamente para biodisponibilidade oral, o que reforça o perfil de lipofilia evidenciado. Posteriormente, todas as substâncias foram avaliadas em ensaios de redução de placas de lise, a fim de se corroborar a atividade observada nos ensaios de MTT.

5.3.3 Ensaios de redução do número de placas de lise

Buscando corroborar a atividade de atividade antiviral/inibição obtidos nos testes de MTT, ensaios de redução de placas de lise foram realizados. Além disso, alguns ensaios foram

propostos para tentar elucidar as possíveis etapas de ação das substâncias, avaliando a potencial atividade virucida e o efeito do pré-tratamento. Inicialmente, os *hits* foram avaliados nos ensaios de placa nos valores da EC_{50} e em duas vezes essas concentrações. A substância **136** reduziu cerca de 1 \log_{10} (UFP/mL) do título viral de ZIKV (PE243) a 12 μM , enquanto na concentração de 24 μM não foi observada uma redução significativa do título viral.

Além disso, as substâncias **136** (10 μM), **128** (50 e 25 μM) e raloxifeno (5 μM) reduziram cerca de 0,5 a 1,5 \log_{10} do título viral YFV (17DD). Bazedoxifeno, por outro lado, não demonstrou redução do título viral em YFV nas concentrações testadas. Já para DENV-2, **136** (20 e 10 μM) e **128** (30 μM) também reduziram cerca de 0,5 a 1,5 \log_{10} do título viral. Finalmente, não houve redução significativa do título de DENV-3 quando avaliada a substância **156**. Esses resultados estão dispostos na **Figura 55**.



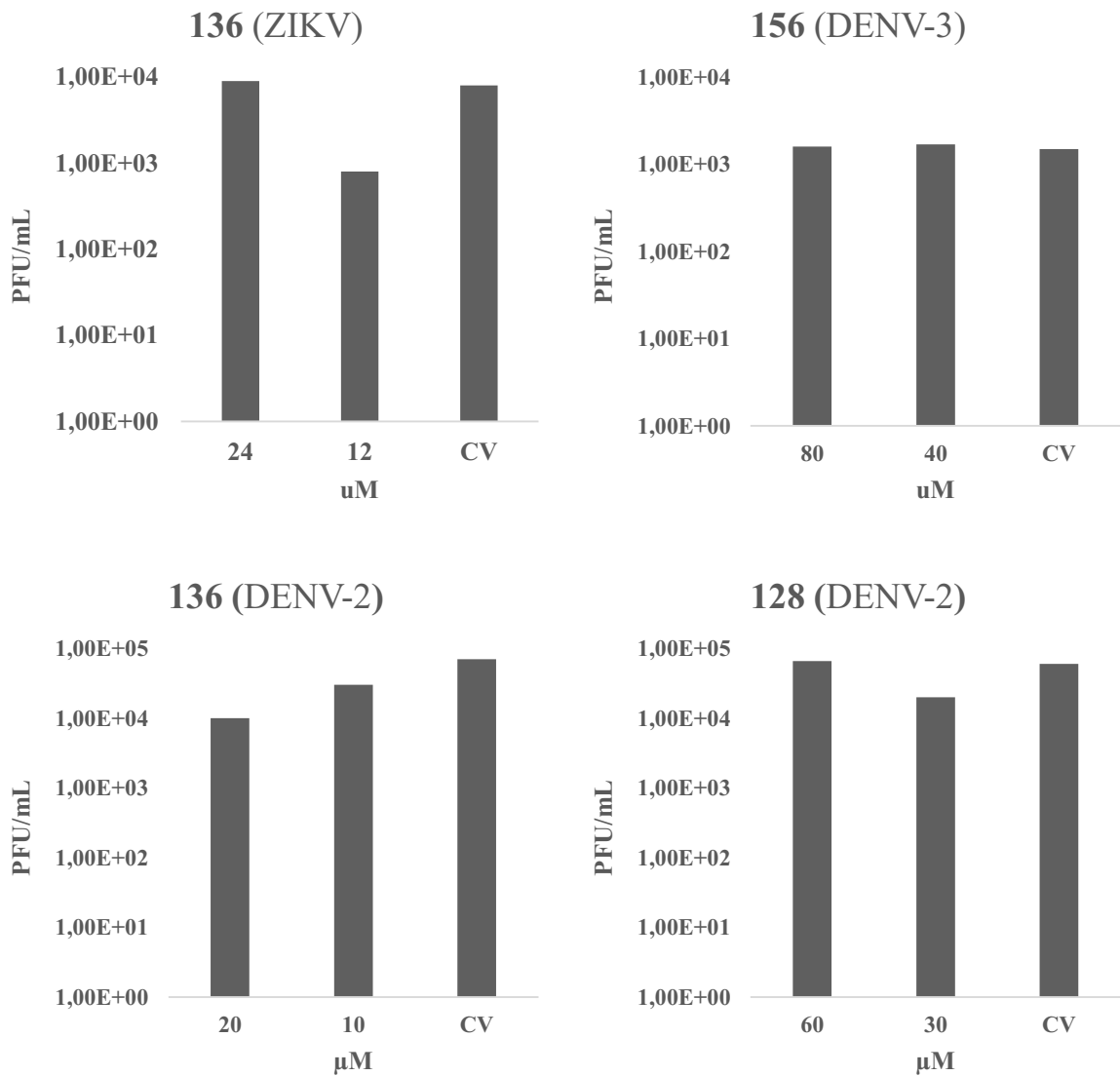


Figura 55. Avaliação da atividade antiviral das cinco substâncias ativas por ensaios de redução de placas de lise. Substâncias **136**, **128**, **140** e **158** avaliadas em YFV. Substância **136** avaliada em ZIKV. Substâncias **128** e **136** avaliadas em DENV-2, além de **156** em DENV-3.

Nas avaliações dos ensaios de redução de placas de lise, foram confirmadas três de oito substâncias com atividade antiviral em ZIKV, YFV e DENV-2. Curiosamente, a substância **156** não apresentou redução dos títulos virais de DENV-3, contrariando os resultados dos ensaios de MTT. Igualmente, basedoxifeno (**140**), também não inibiu títulos de YFV. As substâncias **128** e, especificamente, **136**, apresentaram atividade em dois e três dos flavivírus avaliados, respectivamente, corroborando os resultados obtidos nos ensaios de MTT. Esse resultado já sugere um potencial de ação multi-flavivírus, proposto ao longo de todo trabalho desenvolvido. Isso reflete a importância e relevância das aplicações dos métodos computacionais na

descoberta e desenvolvimento de inibidores e potenciais fármacos, especificamente com uma atividade amplo-espectro.

Nesse cenário, se tratando de DENV, apenas quatro substâncias chegaram às fases clínicas I e II nos últimos anos, sendo uma descontinuada, outra incerta e as duas restantes apresentando atividade inferior à esperada (NORMILE, 2013; TIAN et al., 2018). Essa busca por substâncias bioativas com atividade antiviral, especificamente em flavivírus, requer ainda algumas particularidades que podem dificultar o processo de desenvolvimento. A inibição, nos casos de DENV precisa atingir aos quatro sorotipos do vírus, além de abranger os diferentes e variados genótipos. Idealmente, essas substâncias são mais promissoras caso apresentem um efeito panflavivírus, em razão da similaridade genética entre os membros deste gênero e os problemas associados ao diagnóstico clínico-laboratorial (TAKAHASHI, 2012; NORMILE, 2013; LEONEL, 2018).

Embora os DENV apresentem, assim como os demais flavivírus, variados alvos para o planejamento de fármacos, como NS2B-NS3, NS3 helicase, proteína E, NS5 metiltransferase, RdRp e até fatores do hospedeiro, são poucos aqueles associados a inibidores eficazes ou casos de maior sucesso em estudos *in vitro*, especialmente por variações genéticas dentre as espécies, sorotipos e genótipos, bem como das variações conformacionais das estruturas dos alvos. Ainda assim, o complexo da protease se mostra promissor, especialmente nos casos de sucesso já abordados em HCV e HIV-1 (ZHU, 2012; TIAN, 2018).

Inibidores de protease propostos nesse estudo, especificamente peptídeos, já foram relatados desde o início dos anos 2000. Leung et al. (2001), obtiveram peptídeos lineares com atividade inibitória na protease de DENV. A partir desse trabalho, juntamente ao interesse e incentivo de indústrias farmacêuticas, diversas modificações estruturais destes peptídeos foram realizadas buscando aprimorar a atividade e, potencialmente, a biodisponibilidade das substâncias. Destes partiram trabalhos com modificações de grupos substituintes, como os estudos de Yin et al. (2006), membro da Novartis, que também obteve compostos peptidomiméticos com atividade inibitória da protease de DENV (LEUNG et al., 2001; YIN et al., 2006).

Avanços nessa área identificariam ligantes ainda mais ativas em DENV, a exemplo dos peptídeos de Xu et al. (2012), com valores de $IC_{50} = 2,2 \mu M$. Pouco depois, estudos liderados pela companhia Shionogi empregariam análises de QSAR para melhorar estes mesmos peptídeos, obtendo compostos combinando argininas e aa hidrofóbicos como substituintes nas

novas estruturas. Essa estratégia resultou em duas novas substâncias com atividade inibitória enzimática e confirmação de inibição da multiplicação viral *in vitro*, com valores na faixa de micromolar e índices de seletividade superiores a 10 (XU et al., 2012; TAKAGI et al., 2017).

Os avanços nos estudos com inibidores não peptídicos, por sua vez, foram favorecidos pelas técnicas computacionais baseadas na estrutura de ligantes ou do alvo. Triagens *in silico* também auxiliaram na busca de grandes bibliotecas, permitindo uma avaliação rápida em ensaios *in vitro*. Dentre os métodos utilizados nessas abordagens, policresulen foi identificado com atividade inibitória na protease de DENV-2 ($IC_{50} = 0,48 \mu\text{g/mL}$), além de inibir a multiplicação em células BHK-21 ($EC_{50} = 4,99$). Essa substância ainda apresentou baixa toxicidade em BHK-21 ($CC_{50} = 459,45$), com IS superior a 92 (WU, 2015).

Paralelamente, sulfonamidas, derivados quaternários de amônio, estruturas análogas às quinolonas e até biliverdina também apresentaram alguma atividade anti-DENV, ressaltando a importância dessas abordagens computacionais na descoberta e/ou desenvolvimento de substâncias bioativas (TIAN, 2018). Nesse âmbito, empregando uma triagem virtual, Pambudi et al. (2013), também identificaram um inibidor (SK-12) tendo a interação da NS2B com a NS3 como alvo, apresentando atividade nos quatro sorotipos de DENV, obtendo valores de $EC_{50} = 0,98 \mu\text{M}$ e $IS = 68,63$. Estudos de Mirza et al. (2016), avaliando mais de 18 milhões de substâncias do ZINC, também permitiu identificar algumas substâncias com atividade em um ou mais dos quatro sorotipos de DENV, mais uma vez demonstrando as vantagens das abordagens computacionais (PAMBUDI et al., 2013; MIRZA et al., 2016).

A combinação dessas análises computacionais e experimentais também identificou a aprotinina (NOBLE, 2010) e ribavirina (LIM, 2013) inibidores de DENV. Abordagens mais aprofundadas, como o trabalho de Chu et al. (2015), também identificou substâncias com uma alta permeabilidade de membrana e baixa toxicidade celular, associadas a uma redução da infecção viral. Nesta abordagem, substâncias como a antraquinona, a uma concentração de $1 \mu\text{M}$, reduziram títulos virais de DENV em mais de $1 \log_{10}$ de UFP/mL. Esses estudos corroboram as abordagens empregadas neste trabalho, não somente à diversidade de técnicas computacionais disponíveis para triagem de substâncias bioativas, bem como dos métodos de avaliação da atividade antiviral *in vitro*, resultando em substâncias com atividade em concentrações similares, na faixa de micromolar (CHU et al., 2015).

Igualmente, a combinação de análises *in silico* e *in vitro* também resultaram em algumas substâncias com atividade multi-flavivírus, a exemplo da ivermectina, com atividade em

DENV, JEV e YFV (SWEENEY, 2015). Calleit et al. (2014), também haviam identificado uma série de compostos com atividade em DENV, YFV e WNV, mas descartadas nos testes pré-clínicos, com alta toxicidade em murinos e cães (CALLEIT et al. 2014).

Modificações sintéticas também identificaram novas substâncias com atividade multi-flavivírus, especialmente em DENV, YFV e WNV. Alguns derivados apresentaram atividade inibitória na protease sem estudos subsequentes, enquanto outros foram descontinuados por problemas como baixa biodisponibilidade ou elevada toxicidade em modelos celulares e murinos (GARCÍA; PADILLA; CASTAÑO, 2017). No entanto, poucos são os relatos do amplo-espectro de substâncias anti-flavivírus envolvendo ZIKV. Han et al. (2018) apresentaram inibidores análogos de nucleotídeos e nucleosídeos com atividade *in vitro* (1 a 3,8 μ M) em DENV ou YFV, mas não em ZIKV (HAN et al., 2018).

Entretanto, a atividade de fármacos já aprovados pelo FDA demonstra um potencial de reposicionamento em ZIKV. Novobiocina, sofosbuvir, hidroxicloroquina (HCQ) e niclosamida foram identificados, novamente combinando estratégias de triagem virtual e triagens biológicas *in vitro* e *in vivo*. Novobiocina já foi confirmada também em modelos murinos, enquanto HCQ foi capaz de reduzir títulos virais nas infecções de ZIKV em trofoblastos de placentas de camundongos (BARROWS, 2016; HAN, 2018). No trabalho de Barrows et al. (2016), foram identificadas 20 substâncias dentre 774 fármacos avaliados, incluindo ivermectina, além de daptomicina, mebendazol e ácido micofenólico.

Comparando aos resultados deste trabalho, a substância **128** (IS = 1,97 e 3,16) e especificamente a substância **136** (IS = 1,42 a 3,52) apresentam um potencial multi-flavivírus de interesse, uma vez que apresentaram atividade em ZIKV, DENV e YFV. Embora baixos índices de seletividade tenham sido obtidos, modificações estruturais podem ainda melhorar a atividade ou reduzir a toxicidade celular, especialmente pelo caráter sintético dessas substâncias. Ademais, alguns métodos podem auxiliar na obtenção de maiores informações acerca da atividade e mecanismo de ação das substâncias, como ensaios de pré-tratamento e avaliação da atividade virucida.

5.3.4 Ensaios de pré-tratamento

A substância **128** (25 μ M) e o fármaco raloxifeno (5 μ M) também reduziram pouco mais de 1 log₁₀ do título viral de YFV nos ensaios de pré-tratamento. Já as substâncias **136** e **128** em DENV-2 e **156** em DENV-3 não reduziram significativamente os títulos virais (**Figura 56**).

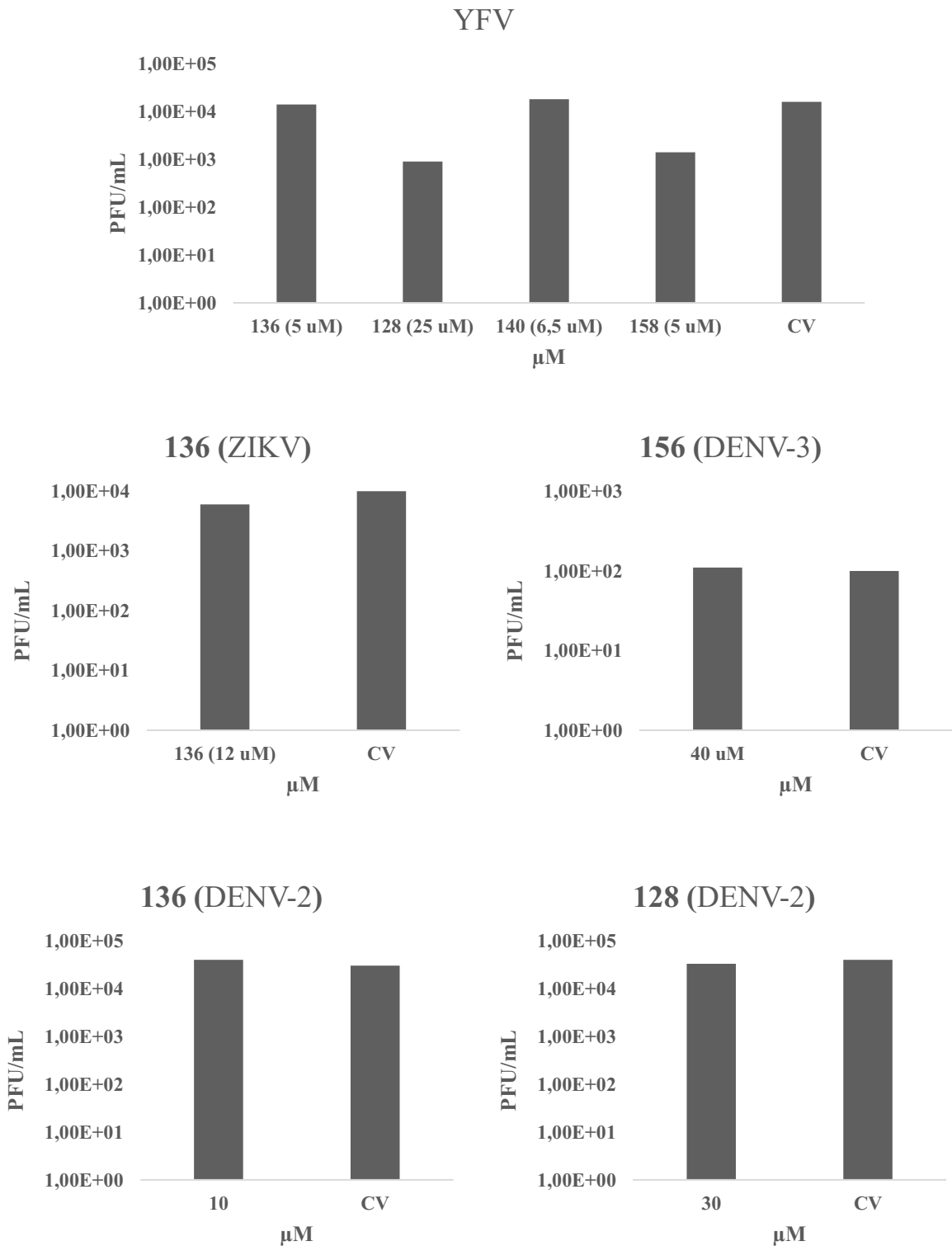


Figura 56. Avaliação do efeito de pré-tratamento das cinco substâncias ativas. Substâncias 136, 128, 140 e 158 avaliadas em YFV. Substância 136 avaliada em ZIKV. Substâncias 128 e 136 avaliadas em DENV-2, além de 156 em DENV-3.

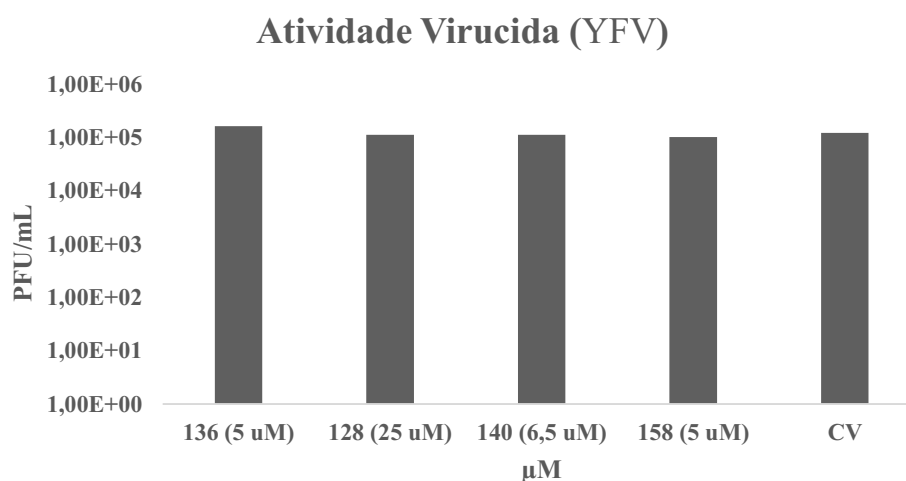
Efeitos de pré-tratamento já foram reportados na literatura e apresentam uma opção favorável para potenciais inibidores. Gómez-Calderón et al. (2017) já demonstraram o efeito de

extratos de frutos de abricó no pré-tratamento em DENV, também após identificar essa atividade antiviral por ensaios de MTT. Similarmente, outros extratos, como os de goiaba, já apresentaram inibição da multiplicação de partículas de DENV em ensaios de pré-tratamento (GÓMEZ-CALDERÓN et al., 2017).

Essa característica também seria interessante na garantia de biodisponibilidade de um potencial fármaco em administrações de diferentes dosagens e períodos. Nesse sentido, Mazzon et al. (2019) apresentaram sucesso em uma estratégia de triagem que indicou 12 substâncias com atividade antiviral em DENV, YFV e ZIKV. Dentre elas, raloxifeno foi sugerido com efeito inibitório em fusão de membranas, durante o ciclo de multiplicação destes vírus, sob condições semelhantes. Esses resultados não divergem deste trabalho, uma vez que raloxifeno também inibiu YFV e promoveu uma pequena redução do título viral em ZIKV. Além disso, a substância **128**, um análogo sintético desse fármaco, também reduziu o título de YFV nos ensaios de pré-tratamento.

5.3.5 Avaliação da atividade virucida

Nenhuma das substâncias apresentou atividade virucida em nenhum dos quatro vírus (**Figura 57**). Esses dados podem sugerir que não há um efeito direto das substâncias sobre a estrutura das partículas, reforçando resultados obtidos nos ensaios de redução de placas de lise e de pré-tratamento.



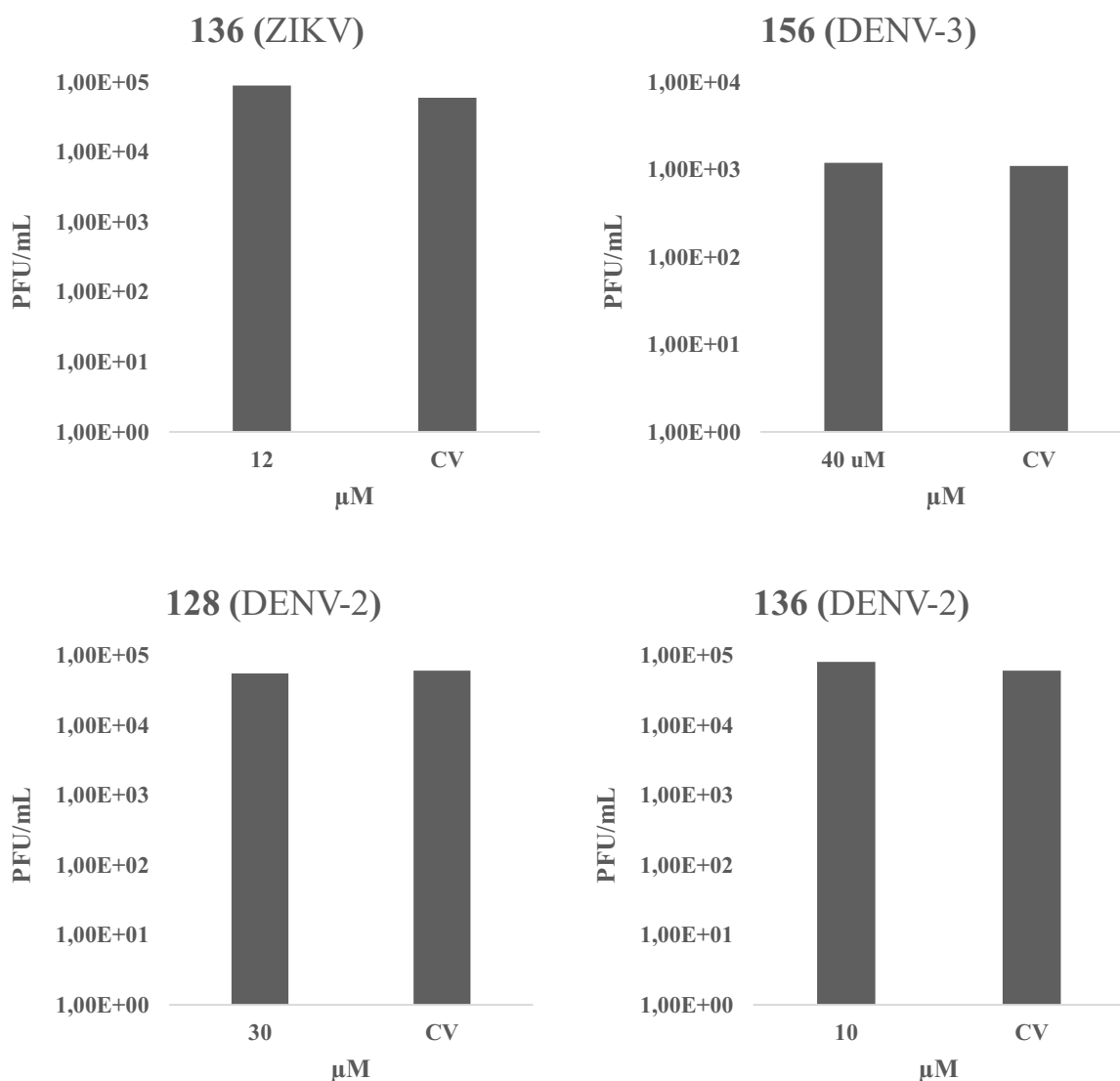


Figura 57. Avaliação da atividade virucida das cinco substâncias ativas. Substâncias 136, 128, 140 e 158 avaliadas em YFV. Substância 136 avaliada em ZIKV. Substâncias 128 e 136 avaliadas em DENV-2, além de 156 em DENV-3.

Uma atividade virucida das substâncias seria interessante, haja visto que os métodos ou substâncias capazes de destruir partículas virais podem cessar uma ameaça de emergência ou re-emergência de um patógeno viral (OSTERHOLM, 2015). A própria OMS já havia proposto nas “Diretrizes de Higiene das Mãos em Cuidados de Saúde” em 2009, o uso de substâncias para desinfecção cirúrgica e higiênica das mãos e superfícies em ambientes de saúde, reduzindo, por exemplo, a transmissão de patógenos, como outros vírus envelopados pelas mãos de indivíduos (WORLD HEALTH ORGANIZATION, 2009; OSTERHOLM et al., 2015).

Pensando em ZIKV, por exemplo, ainda que não seja possível contato direto de soluções ou formulações com a partícula no ambiente, uma substância virucida permitiria maior controle em situações de surtos ou epidemias, reduzindo de maneira mais eficaz eventuais novos casos de infecção e disseminação do vírus (SIDDHARTA et al., 2017; BATISTA et al., 2018). Uma atividade virucida também poderia favorecer um potencial inibidor de amplo espectro, como demonstrado por Batista et al. (2018), com inibidores de HCV, especificamente àqueles capazes de desestabilizar membranas, outro potencial alvo de interesse aos demais flavivírus neste estudo.

Esses dados, no entanto, não descartam a relevância deste e demais estudos, bem como dos potenciais inibidores do complexo da protease viral (NS2B-NS3) avaliados neste trabalho. Ademais, resta confirmar um potencial efeito multi-flavivírus em ensaios, por exemplo, de inibição enzimática, permitindo, dessa forma, corroborar a toda construção e resultados obtidos nesse trabalho e validar o alvo proposto.

Os resultados promissores dos métodos e técnicas *in silico* desenvolvidas, junto à combinação das predições e validações biológicas, fizeram desta triagem virtual e biológica um planejamento de sucesso, apresentando uma atividade antiviral identificada em cinco das oito substâncias adquiridas, sendo três confirmadas em ambos os ensaios de atividade antiviral. Ainda, a possibilidade de modificações estruturais e síntese de novos análogos, buscando melhorias na atividade ou biodisponibilidade no planejamento de inibidores e potenciais fármacos, favorecem a aplicabilidade dessas substâncias como potenciais fármacos com atividade antiviral multi-flavivírus.

6 CONCLUSÕES

6.1. Fragmentos de lisina e arginina nos ligantes são importantes para a ligação dos compostos peptidomiméticos avaliados ao sítio ativo da protease (NS3) de DENV;

6.2. Regiões favoráveis de interação molecular (HBD, HBA, iônica e hidrofóbicas) foram confirmadas por interações em sondas e sítios de ligação em ambas as proteases de DENV-3 (3U1I) e ZIKV (5YOF), corroboradas nas análises de acoplamento molecular;

6.3. A triagem virtual de 7.600.000 moléculas resultou em 18 restantes, submetidas à dinâmica molecular para análise de interação e estabilidade, bem como das predições de toxicidade;

6.4. Foram selecionadas 8 substâncias para validação da atividade biológica: seis sob investigação (denominadas **134**, **151**, **154**, **156**, **136**, **128**) e dois fármacos licenciados como propostas de reposicionamento, bazedoxifeno (**140**) e raloxifeno (**158**). Destas, cinco (**136**, **151**, **154**, **140** e **156**) foram preditas como ativas no modelo de SVM polinomial;

6.5. Nas avaliações biológicas, cinco substâncias apresentaram atividade nos flavivírus, sendo: uma em ZIKV ($CE_{50} = 11,86 \pm 0,43 \mu\text{M}$; IS = 1,42), quatro em YFV ($CE_{50} = 4,79 \pm 0,17$ a $24,15 \pm 1,15 \mu\text{M}$; IS = 1,97 a 3,74), duas em DENV-2 ($CE_{50} = 4,21 \pm 0,14$ e $9,51 \pm 0,12 \mu\text{M}$; IS = 2,26 e 3,16) e uma em DENV-3 ($CE_{50} = 37,51 \pm 0,8 \mu\text{M}$; IS = 2,85).

6.6. Os ensaios de redução do número de placas de lise corroboraram a atividade das substâncias **136** e **128** em ZIKV, YFV e DENV-2, reduzindo cerca de 1 a 1,5 \log_{10} de UFP/mL para os três vírus;

6.7. Nos ensaios de pré-tratamento em YFV, a substância **128** e raloxifeno (**158**) reduziram cerca de 1 \log_{10} de UFP/mL, corroborando dados já observados na literatura;

6.8. Não foi observada atividade virucida para as substâncias avaliadas.

7 PERSPECTIVAS

7.1. Quantificar o RNA genômico para avaliação do potencial efeito das substâncias ativas na replicação do genoma viral.

7.2. Validar o mecanismo de ação destas substâncias (inibição de NS3pro) empregando ensaios de inibição enzimática.

7.3. Avaliar os efeitos de toxicidade em linhagens celulares de diferentes tecidos humanos, por exemplo, HEK293 (renal), Hep G2 ou Huh-7 (hepáticas) e THP-1 (monocítica) ou L929 (linfocítica), além de outros tecidos murinos, como U251 (neuroblastoma) e Raw 264.7 (monocítica), respectivamente.

7.4. Realizar ensaios pré-clínicos para avaliação de biodisponibilidade, dos efeitos de toxicidade e de infecção pelos diferentes vírus *in vivo* em modelo(s) murino(s).

8 REFERÊNCIAS BIBLIOGRÁFICAS

ADACHI, A. et al. Advances in Diagnosis, Surveillance, and Monitoring of Zika Virus: An Update. **Front. Microbiol.**, v. 8, p. 2677, 2018.

AGUILERA-PESANTES, D. et al. Discovering key residues of dengue virus NS2b-NS3-protease: New binding sites for antiviral inhibitors design. **Biochemical and Biophysical Research Communications**, v. 492, n. 4, p. 631–642, 2017.

ALESHIN, A. E. et al. Structural evidence for regulation and specificity of flaviviral proteases and evolution of the *Flaviviridae* fold. **Protein Science**, v. 16, n. 5, p. 795–806, 2007.

ALPHEY, L. et al. Sterile-insect methods for control of mosquito-borne diseases: an analysis. **Vector borne and zoonotic diseases (Larchmont, N.Y.)**, v. 10, n. 3, p. 295–311, abr. 2010.

ALVES, V. M. et al. **Development of web and mobile applications for chemical toxicity prediction** *Journal of the Brazilian Chemical Society* Sociedade Brasileira de Química, 1 maio 2018. .

AMARO, R. E. et al. Mechanism of 150-cavity formation in influenza neuraminidase. **Nature Communications**, v. 2, n. 1, 2011.

AMICANGELO, J. C. et al. Ab initio study of substituent effects in the interactions of dimethyl ether with aromatic rings. **Physical Chemistry Chemical Physics**, v. 10, n. 19, p. 2695–2705, 2008.

AN, J. Surflex: Fully Automatic Flexible Molecular Docking Using a Molecular Similarity-Based Search Engine. **Journal of medicinal chemistry**, v. 46, n. 4, 13 fev. 2003. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/12570372/>>. Acesso em: 25 mar. 2021.

AN, J. Surflex-Dock 2.1: Robust Performance from Ligand Energetic Modeling, Ring Flexibility, and Knowledge-Based Search. **Journal of computer-aided molecular design**, v. 21, n. 5, maio 2007. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/17387436/>>. Acesso em: 25 mar. 2021.

ARPORNSUWAN, T.; PUNJANON, T. Tumor cell-selective antiproliferative effect of the extract from *Morinda citrifolia* fruits. **Phytotherapy Research**, v. 20, n. 6, p. 515–517, jun. 2006.

ARUN, B.; ANTHONY, M.; DUNN, B. **The search for the ideal SERM** *Expert Opinion on Pharmacotherapy*, 2002.

ARUNKUMAR, J.; RAJARAJAN, S. Study on antiviral activities, drug-likeness and molecular docking of bioactive compounds of *Punica granatum* L. to Herpes simplex virus - 2 (HSV-2). **Microbial Pathogenesis**, v. 118, p. 301–309, maio 2018.

AU, J. S.; POCKROS, P. J. Novel therapeutic approaches for hepatitis C. **Clinical Pharmacology and Therapeutics**, v. 95, n. 1, p. 78–88, jan. 2014.

BALAKIN, K. V. DMSO solubility and bioscreening. p. 27–32, 1 ago. 2003.

BALAKIN, K. V. et al. In Silico Estimation of DMSO Solubility of Organic Compounds for Bioscreening. **Journal of Biomolecular Screening**, v. 9, n. 1, p. 22–31, 1 fev. 2004.

BALMASEDA, A. et al. Antibody-based assay discriminates Zika virus infection from other flaviviruses. **Proceedings of the National Academy of Sciences**, v. 114, n. 31, p. 8384–8389, 1 ago. 2017.

BANERJEE, P. et al. ProTox-II: A webserver for the prediction of toxicity of chemicals. **Nucleic Acids Research**, v. 46, n. W1, p. W257–W263, 2 jul. 2018.

BARILLARI, C.; MARCOU, G.; ROGNAN, D. Hot-Spots-Guided Receptor-Based Pharmacophores (HS-Pharm): A Knowledge-Based Approach to Identify Ligand-Anchoring Atoms in Protein Cavities and Prioritize Structure-Based Pharmacophores. **Journal of Chemical Information and Modeling**, v. 48, n. 7, p. 1396–1410, 21 jul. 2008.

BARNETT, E. D. Yellow Fever: Epidemiology and Prevention. **Clinical Infectious Diseases**, v. 44, n. 6, p. 850–856, 15 mar. 2007.

BARROWS, N. J. et al. A Screen of FDA-Approved Drugs for Inhibitors of Zika Virus Infection. **Cell Host and Microbe**, v. 20, n. 2, p. 259–270, 2016.

BATISTA, M. N. et al. GA-Hecate antiviral properties on HCV whole cycle represent a new antiviral class and open the door for the development of broad spectrum antivirals. **Scientific Reports**, v. 8, n. 1, 1 dez. 2018.

BAUD, D. et al. An update on Zika virus infection. **The Lancet**, v. 390, n. 10107, p. 2099–2109, 2017.

BEHNAM, M. A. M. et al. Discovery of Nanomolar Dengue and West Nile Virus Protease Inhibitors Containing a 4 - Benzyloxyphenylglycine Residue. 2015.

BEISEL, U.; BOËTE, C. The Flying Public Health Tool: Genetically Modified Mosquitoes and Malaria Control. **Science as Culture**, v. 22, n. 1, p. 38–60, 8 mar. 2013.

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

BHATT, S. et al. The global distribution and burden of dengue. **Nature**, v. 496, n. 7446, p. 504–507, 2013.

BOLCATO, G. et al. Can We Still Trust Docking Results? An Extension of the Applicability of DockBench on PDBbind Database. **International Journal of Molecular Sciences**, v. 20, n. 14, p. 3558, 20 jul. 2019.

BORHANI, D. W.; SHAW, D. E. The future of molecular dynamics simulations in drug discovery. **Journal of Computer-Aided Molecular Design**, v. 26, n. 1, p. 15–26, jan. 2012.

BORODINA, Y. V. et al. Assessment of conformational ensemble sizes necessary for specific resolutions of coverage of conformational space. **Journal of Chemical Information and Modeling**, v. 47, n. 4, p. 1428–1437, jul. 2007.

BOWERS, K. J. et al. Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. In: SC '06: Proceedings of the 2006 ACM/IEEE Conference on Supercomputing, **Anais**. In: SC '06: PROCEEDINGS OF THE 2006 ACM/IEEE CONFERENCE ON SUPERCOMPUTING. nov. 2006.

BRADY, O. J. et al. Refining the Global Spatial Limits of Dengue Virus Transmission by Evidence-Based Consensus. **PLoS Neglected Tropical Diseases**, v. 6, n. 8, p. e1760, 7 ago. 2012.

BRAGA, R. et al. Tuning hERG Out: Antitarget QSAR Models for Drug Development. **Current Topics in Medicinal Chemistry**, v. 14, n. 11, p. 1399–1415, 10 jun. 2014.

BRAGA, R. C. et al. Pred-hERG: A Novel web-Accessible Computational Tool for Predicting Cardiac Toxicity. **Molecular Informatics**, v. 34, n. 10, p. 698–701, 1 out. 2015.

BRYAN, C. S.; MOSS, S. W.; KAHN, R. J. Yellow fever in the Americas. **Infectious Disease Clinics of North America**, v. 18, n. 2, p. 275–292, jun. 2004.

CALY, L.; WAGSTAFF, K. M.; JANS, D. A. Nuclear trafficking of proteins from RNA viruses: potential target for antivirals? **Antiviral research**, v. 95, n. 3, p. 202–6, set. 2012.

CAMPI-AZEVEDO, A. C. et al. Booster dose after 10 years is recommended following 17DD-YF primary vaccination. **Human Vaccines & Immunotherapeutics**, v. 12, n. 2, p. 491–502, 11 fev. 2016.

CAPEDING, M. R. et al. Clinical efficacy and safety of a novel tetravalent dengue vaccine in healthy children in Asia: a phase 3, randomised, observer-masked, placebo-controlled trial. **Lancet (London, England)**, v. 384, n. 9951, p. 1358–65, 11 out. 2014.

CAUCHEMEZ, S. et al. Association between Zika virus and microcephaly in French Polynesia, 2013–15: a retrospective study. **The Lancet**, v. 387, n. 10033, p. 2125–2132, 21 maio 2016.

CHAI, T.; DRAXLER, R. R. Root mean square error (RMSE) or mean absolute error (MAE)?-Arguments against avoiding RMSE in the literature. **Geosci. Model Dev**, v. 7, p. 1247–1250, 2014.

CHANPRAPAPH, S. et al. Competitive Inhibition of the Dengue Virus NS3 Serine Protease by Synthetic Peptides Representing Polyprotein Cleavage Sites. **Biochemical and Biophysical Research Communications**, v. 330, n. 4, p. 1237–1246, 20 maio 2005.

CHIRICO, N.; GRAMATICA, P. Real External Predictivity of QSAR Models: How To Evaluate It? Comparison of Different Validation Criteria and Proposal of Using the Concordance Correlation Coefficient. **Journal of Chemical Information and Modeling**, v. 51, n. 9, p. 2320–2335, 26 set. 2011.

CHU, J. J. H. et al. Antiviral activities of 15 dengue NS2B-NS3 protease inhibitors using a human cell-based viral quantification assay. **Antiviral Research**, v. 118, 2015.

COELHO, S. V. A. et al. Development of standard methods for Zika virus propagation, titration, and purification. **Journal of Virological Methods**, v. 246, n. February, p. 65–74, 2017.

COUVELARD, A. et al. Report of a fatal case of dengue infection with hepatitis: Demonstration of dengue antigens in hepatocytes and liver apoptosis. **Human Pathology**, v. 30, n. 9, p. 1106–1110, 1999.

CUCUNAWANGSIH; LUGITO, N. P. H.; KURNIAWAN, A. Immunoglobulin G (IgG) to IgM ratio in secondary adult dengue infection using samples from early days of symptoms onset. **BMC Infectious Diseases**, v. 15, n. 1, 21 jul. 2015.

DA PAIXÃO, V. G.; PITA, S. S. da R. In Silico Identification and Evaluation of New Trypanosoma Cruzi Trypanothione Reductase (TcTR) Inhibitors Obtained from Natural Products Database of the Bahia Semi-Arid Region (NatProDB). **Computational Biology and Chemistry**, v. 79, p. 36–47, abr. 2019.

DANEMAN, R.; PRAT, A. The blood–brain barrier. **Cold Spring Harbor Perspectives in Biology**, v. 7, n. 1, 1 jan. 2015.

DARDEN, T.; YORK, D.; PEDERSEN, L. Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems. **Citation: The Journal of Chemical Physics**, v. 98, p. 7694, 1993.

DE ARAÚJO, T. V. B. et al. Association between Zika virus infection and microcephaly in Brazil, January to May, 2016: preliminary report of a case-control study. **The Lancet. Infectious diseases**, v. 16, n. 12, p. 1356–1363, 1 dez. 2016.

DE CLERCQ, E. **Anti-HIV drugs: 25 compounds approved within 25 years after the discovery of HIV** *International Journal of Antimicrobial Agents*, abr. 2009. .

DE LIMA NETO, D. F. et al. Molecular dynamics suggests antiviral compounds active against Dengue Virus show similar binding patterns to Zika Virus proteins. **bioRxiv**, 2018.

DONALD, C. L. et al. Full Genome Sequence and sfRNA Interferon Antagonist Activity of Zika Virus from Recife, Brazil. **PLOS Neglected Tropical Diseases**, v. 10, n. 10, p. e0005048, 5 out. 2016.

D'ORTENZIO, E. et al. Evidence of Sexual Transmission of Zika Virus. **New England Journal of Medicine**, v. 374, n. 22, p. 2195–2198, 2 jun. 2016.

DOWD, K. A. et al. Rapid development of a DNA vaccine for Zika virus. **Science**, v. 354, n. 6309, p. 237–240, 14 out. 2016.

DRWAL, M. N. et al. ProTox: A web server for the in silico prediction of rodent oral toxicity. **Nucleic Acids Research**, v. 42, n. W1, 1 jul. 2014.

DUAN, Y.; KOLLMAN, P. A. Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. **Science**, v. 282, n. 5389, p. 740–744, 23 out. 1998.

DUFFY, M. R. et al. Zika Virus Outbreak on Yap Island, Federated States of Micronesia. **New England Journal of Medicine**, v. 360, n. 24, p. 2536–2543, 11 jun. 2009.

DURRANT, J. D.; MCCAMMON, J. A. Molecular dynamics simulations and drug discovery. **BMC Biology**, v. 9, n. 1, p. 71, 28 dez. 2011.

- EFFERTH, T. et al. The Antiviral Activities of Artemisinin and Artesunate. p. 47, 2008.
- EILLANNY, P.; MARINHO, S. Atividade antiviral de extratos de *Maytenus imbricata* contra Flavivírus. Dissertação de Mestrado. 2015.
- FANG, J. et al. Structure-based discovery of two antiviral inhibitors targeting the NS3 helicase of Japanese encephalitis virus. **Scientific Reports**, v. 6, n. 1, p. 34550, 29 dez. 2016.
- FARIA, N. R. et al. Zika virus in the Americas: Early epidemiological and genetic findings. **Science**, v. 352, n. 6283, p. 345–349, 2016.
- FERNÁNDEZ-SANLÉS, A. et al. Functional information stored in the conserved structural RNA domains of flavivirus genomes. **Frontiers in Microbiology**, v. 8, n. APR, p. 1–16, 2017.
- FERREIRA, R. S.; GLAUCIUS, O.; ANDRICOPULO, A. D. Integração das técnicas de triagem virtual e triagem biológica automatizada em alta escala: oportunidades e desafios em P&D de fármacos. **Química Nova**, v. 34, n. 10, p. 1770–1778, 2011.
- FOY, B. D. et al. Probable non-vector-borne transmission of Zika virus, Colorado, USA. **Emerging infectious diseases**, v. 17, n. 5, p. 880–2, maio 2011.
- FRAGOSO, Y. D. et al. Guillain-Barré Syndrome and Dengue Fever: Report on Ten New Cases in Brazil. **Arquivos de Neuro-Psiquiatria**, v. 74, n. 12, p. 1039–1040, dez. 2016.
- FRECER, V.; MIERTUS, S. Design, structure-based focusing and in silico screening of combinatorial library of peptidomimetic inhibitors of Dengue virus NS2B-NS3 protease. **Journal of Computer-Aided Molecular Design**, v. 24, n. 3, p. 195–212, 21 mar. 2010.
- FRITZ, H.; WUNDERER, G. Biochemistry and Applications of Aprotinin, the Kallikrein Inhibitor from Bovine Organs. **Arzneimittel-Forschung**, v. 33, n. 4, p. 479–494, 1983.
- GAO, M. et al. Chemical Genetics Strategy Identifies an HCV NS5A Inhibitor with a Potent Clinical Effect. **Nature**, v. 465, n. 7294, p. 96–100, 6 maio 2010.
- GARCÍA, L. L.; PADILLA, L.; CASTAÑO, J. C. **Inhibitors compounds of the flavivirus replication process** *Virology Journal* BioMed Central Ltd. 15 maio 2017.
- GARCIA, M. Situação Entomológica das Arboviroses 2017/2018. 2018.
- GARCIA, T. S.; HONÓRIO, K. M. Two-dimensional quantitative structure-activity relationship studies on bioactive ligands of Peroxisome proliferator-activated receptor δ . **Journal of the Brazilian Chemical Society**, v. 22, n. 1, p. 65–72, 2011a.
- GARCIA, T. S.; HONÓRIO, K. M. Two-dimensional quantitative structure-activity relationship studies on bioactive ligands of peroxisome proliferator-activated receptor δ . **Journal of the Brazilian Chemical Society**, v. 22, n. 1, p. 65–72, jan. 2011b.
- GARDNER, C. L.; RYMAN, K. D. Yellow Fever: A Reemerging Threat. **Clinics in Laboratory Medicine**, v. 30, n. 1, p. 237–260, mar. 2010.

GERMOLEC, D. et al. **Immunotoxicology: A brief history, current status and strategies for future immunotoxicity assessment** *Current Opinion in Toxicology* Elsevier B.V. 1 ago. 2017. .

GERTRUDES, J. C. et al. Machine Learning Techniques and Drug Design. **Current Medicinal Chemistry**, v. 19, n. 25, p. 4289–4297, 24 set. 2012.

GIAGINIS, C.; TSANTILI-KAKOULIDOU, A. Current State of the Art in HPLC Methodology for Lipophilicity Assessment of Basic Drugs. A Review. [s.d.]

GOLBRAIKH, A.; TROPSHA, A. Beware of q²! In: *Journal of Molecular Graphics and Modelling*, 4, **Anais...**2002.

GÓMEZ-CALDERÓN, C. et al. Antiviral effect of compounds derived from the seeds of *Mammea americana* and *Tabernaemontana cymosa* on Dengue and Chikungunya virus infections. **BMC Complementary and Alternative Medicine**, v. 17, n. 1, p. 57, 18 dez. 2017.

GRAMATICA, P.; SANGION, A. A Historical Excursus on the Statistical Validation Parameters for QSAR Models: A Clarification Concerning Metrics and Terminology. **Journal of Chemical Information and Modeling**, v. 56, n. 6, p. 1127–1131, 27 jun. 2016.

GROSSI-SOYSTER, E. N. et al. Serological and spatial analysis of alphavirus and flavivirus prevalence and risk factors in a rural community in western Kenya. **PLoS Neglected Tropical Diseases**, 2017.

GUIDO, R. V. C.; ANDRICOPULO, A. D.; OLIVA, G. Planejamento de fármacos, biotecnologia e química medicinal: aplicações em doenças infecciosas. **Estudos Avançados**, v. 24, n. 70, p. 81–98, 2010.

GUZMAN, M. G. et al. Multi-Country Evaluation of the Sensitivity and Specificity of Two Commercially-Available NS1 ELISA Assays for Dengue Diagnosis. **PLoS Neglected Tropical Diseases**, v. 4, n. 8, p. e811, 31 ago. 2010.

GUZMAN, M. G. et al. Dengue infection. **Nature Reviews Disease Primers**, v. 2, p. 1–26, 2016.

GUZMAN, M. G.; HARRIS, E. Dengue. **The Lancet**, v. 385, n. 9966, p. 453–465, 2015.

H. BROOKS, W.; C. GUIDA, W.; G. DANIEL, K. The Significance of Chirality in Drug Design and Development. **Current Topics in Medicinal Chemistry**, v. 11, n. 7, p. 760–770, 28 mar. 2011.

HALSTEAD, S. B. The XXth century dengue pandemic: need for surveillance and research. **World health statistics quarterly. Rapport trimestriel de statistiques sanitaires mondiales**, v. 45, n. 2–3, p. 292–8, 1992.

HAWKINS, P. C. D.; SKILLMAN, A. G.; NICHOLLS, A. Comparison of Shape-Matching and Docking as Virtual Screening Tools. **Journal of Medicinal Chemistry**, v. 50, n. 1, p. 74–82, jan. 2007.

HAZUDA, D. J. et al. A naphthyridine carboxamide provides evidence for discordant resistance between mechanistically identical inhibitors of HIV-1 integrase. **Proceedings of the National**

Academy of Sciences of the United States of America, v. 101, n. 31, p. 11233–11238, 3 ago. 2004.

HEINZ, F. X.; STIASNY, K. Flaviviruses and flavivirus vaccines. **Vaccine**, v. 30, n. 29, p. 4301–4306, 2012.

HERITAGE, T. W.; LOWIS, D. R. Molecular Hologram QSAR. In: [s.l.: s.n.]p. 212–225.

HORTON, D. A.; BOURNE, G. T.; SMYTHE, M. L. Exploring Privileged Structures: The Combinatorial Synthesis of Cyclic Peptides. **Journal of Computer-Aided Molecular Design**, v. 16, n. 5–6, p. 415–430, jun. 2002.

HUEY ET AL., R. “Autodock Vina”. p. 1–12, 2012.

HUNSPERGER, E. A. et al. Evaluation of Commercially Available Anti–Dengue Virus Immunoglobulin M Tests. **Emerging Infectious Diseases**, v. 15, n. 3, p. 436–440, mar. 2009.

HUNSPERGER, E. A. et al. Evaluation of commercially available diagnostic tests for the detection of dengue virus NS1 antigen and anti-dengue virus IgM antibody. **PLoS neglected tropical diseases**, v. 8, n. 10, p. e3171, out. 2014a.

HUNSPERGER, E. A. et al. Evaluation of Commercially Available Diagnostic Tests for the Detection of Dengue Virus NS1 Antigen and Anti-Dengue Virus IgM Antibody. **PLoS Neglected Tropical Diseases**, v. 8, n. 10, p. e3171, 16 out. 2014b.

IKEJEZIE, J. et al. Zika Virus Transmission — Region of the Americas, May 15, 2015–December 15, 2016. **MMWR. Morbidity and Mortality Weekly Report**, v. 66, n. 12, p. 329–334, 31 mar. 2017.

JACOBSON, M. P. et al. A Hierarchical Approach to All-Atom Protein Loop Prediction. **Proteins: Structure, Function and Genetics**, v. 55, n. 2, p. 351–367, 1 maio 2004.

JAIN, A. N. Surflex: Fully automatic flexible molecular docking using a molecular similarity-based search engine. **Journal of Medicinal Chemistry**, v. 46, n. 4, p. 499–511, 13 fev. 2003.

JANI, K. S.; DALAFAVE, D. S. Computational design of targeted inhibitors of Polo-like Kinase 1 (PK1). **Bioinformatics and Biology Insights**, v. 6, p. 23–31, 2012.

JEFFRIES, C. L.; WALKER, T. Wolbachia Biocontrol Strategies for Arboviral Diseases and the Potential Influence of Resident Wolbachia Strains in Mosquitoes. **Current tropical medicine reports**, v. 3, p. 20–25, 2016.

JINDO, T. et al. The effect of terminal dimethyl and diethyl substituents on the J-aggregate-like molecular arrangement of bisazomethine dye molecules. **CrystEngComm**, v. 17, n. 37, p. 7213–7226, 31 jul. 2015.

KADKHODA, K.; GRETCHEN, A.; RACANO, A. Evaluation of a commercially available Zika virus IgM ELISA: specificity in focus. **Diagnostic Microbiology and Infectious Disease**, v. 88, n. 3, p. 233–235, jul. 2017.

KARIM, A. et al. Efficient Toxicity Prediction via Simple Features Using Shallow Neural Networks and Decision Trees. 2019.

KIRALJ, R.; FERREIRA, M. M. C. Basic validation procedures for regression models in QSAR and QSPR studies: Theory and application. **Journal of the Brazilian Chemical Society**, v. 20, n. 4, p. 770–787, 2009.

KOVÁČIK, A. et al. Effect of macrolide antibiotics on various cell cultures in vitro: 2. cell biochemistry. **Journal of Microbiology**. 2019.

KOZAKOV, D. et al. The FTMap family of web servers for determining and characterizing ligand-binding hot spots of proteins. **Nature Protocols**, v. 10, n. 5, p. 733–755, 6 maio 2015.

KRESTEN LINDORFF, S. P. How Proteins Fold-Suppl. material. **Science**, v. 334, n. 6055, p. 464–465, 2011.

KRONENBERGER, T. et al. Studies of *Staphylococcus aureus* FabI inhibitors: fragment-based approach based on holographic structure–activity relationship analyses. **Future Medicinal Chemistry**, v. 9, n. 2, p. 135–151, 2017.

KUNO, G.; CHANG, G.-J. J. Biological Transmission of Arboviruses: Reexamination of and New Insights into Components, Mechanisms, and Unique Traits as Well as Their Evolutionary Trends. **CLINICAL MICROBIOLOGY REVIEWS**, v. 18, n. 4, p. 608–637, 2005.

LAMAS, B.; NATIVIDAD, J. M.; SOKOL, H. **Aryl hydrocarbon receptor and intestinal immunity review-article** *Mucosal Immunology* Nature Publishing Group, , 1 jul. 2018. .

LANI, R. et al. Antiviral activity of silymarin against chikungunya virus. **Scientific Reports**, v. 5, 16 jun. 2015.

LAZAR, by; PREDICTOR SUPANDI, A.; MERDEKAWATI, F. In Silico Study of Pyrazolylaminoquinazoline Toxicity ARTICLE INFO ABSTRACT. **Journal of Applied Pharmaceutical Science**, v. 8, n. 09, p. 119–129, 2018.

LAZEAR, H. M.; DIAMOND, M. S. Zika Virus: New Clinical Syndromes and Its Emergence in the Western Hemisphere. **Journal of Virology**, v. 90, n. 10, p. 4864–4875, 15 maio 2016.

LEE, D. J.; ROBINSON, W. E. Preliminary mapping of a putative inhibitor-binding pocket for human immunodeficiency virus type 1 integrase inhibitors. **Antimicrobial Agents and Chemotherapy**, v. 50, n. 1, p. 134–142, jan. 2006.

LEONEL, C. A. et al. Pharmacophoric characteristics of dengue virus NS2B/NS3pro inhibitors: a systematic review of the most promising compounds. **Archives of Virology**, v. 163, n. 3, p. 575–586, 2018.

LEUNG, D. et al. Activity of Recombinant Dengue 2 Virus NS3 Protease in the Presence of a Truncated NS2B Co-factor, Small Peptide Substrates, and Inhibitors. **Journal of Biological Chemistry**, v. 276, n. 49, p. 45762–45771, 7 dez. 2001.

LI, J. et al. Functional Profiling of Recombinant NS3 Proteases from All Four Serotypes of Dengue Virus Using Tetrapeptide and Octapeptide Substrate Libraries*. **Journal of Biological Chemistry**, v. 280, n. 31, p. 28766–28774, 5 ago. 2005.

LI, L. et al. The Flavivirus Precursor Membrane-Envelope Protein Complex: Structure and Maturation. **Science**, v. 319, n. 5871, p. 1830–1834, 28 mar. 2008.

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

LIMA, M. N. N. et al. QSAR-driven design and discovery of novel compounds with antiplasmodial and transmission blocking activities. **Frontiers in Pharmacology**, v. 9, n. MAR, 6 mar. 2018.

LINDENBACH, B. D.; RICE, C. M. Molecular biology of flaviviruses. **Advances in virus research**, v. 59, p. 23–61, 2003.

LINDORFF-LARSEN, K. et al. How Fast-Folding Proteins Fold. **Science (New York, N.Y.)**, v. 334, n. 6055, p. 517–520, 28 out. 2011.

LIPINSKI, C. A. et al. Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings. **Advanced Drug Delivery Reviews**, In Vitro Models for Selection of Development Candidates. v. 23, n. 1, p. 3–25, 15 jan. 1997.

LIPINSKI, C.; HOPKINS, A. Navigating Chemical Space for Biology and Medicine. **Nature**, v. 432, n. 7019, p. 855–861, 16 dez. 2004.

LIU, C.-J. et al. Syntheses, cytotoxic activity evaluation and HQSAR study of 1,2,3-triazole-linked isosteviol derivatives as potential anticancer agents. **Bioorganic & Medicinal Chemistry Letters**, v. 26, n. 22, p. 5455–5461, 15 nov. 2016.

LO, Y. C. et al. **Machine learning in chemoinformatics and drug discovery** **Drug Discovery Today** Elsevier Ltd. 1 ago. 2018. .

LU, X. et al. Pharmacophore and Molecular Docking Guided 3D-QSAR Study of Bacterial Enoyl-ACP Reductase (FabI) Inhibitors. **International Journal of Molecular Sciences**, v. 13, n. 6, p. 6620–6638, 30 maio 2012.

LUO, D.; VASUDEVAN, S. G.; LESCAR, J. The flavivirus NS2B-NS3 protease-helicase as a target for antiviral drug development. **Antiviral Research**, v. 118, n. April, p. 148–158, 2015.

MAEDA, K. et al. Involvement of the Second Extracellular Loop and Transmembrane Residues of CCR5 in Inhibitor Binding and HIV-1 Fusion: Insights into the Mechanism of Allosteric Inhibition. **Journal of Molecular Biology**, v. 381, n. 4, p. 956–974, 12 set. 2008.

MAGILL, S. S. et al. Multistate Point-Prevalence Survey of Health Care–Associated Infections. **New England Journal of Medicine**, v. 370, n. 13, p. 1198–1208, 2014.

MALTAROLLO, V. G. et al. **Applying machine learning techniques for ADME-Tox prediction: A review** **Expert Opinion on Drug Metabolism and Toxicology** Informa Healthcare, 1 fev. 2015. .

MALTAROLLO, V. G. et al. **Advances with support vector machines for novel drug discovery** **Expert Opinion on Drug Discovery** Taylor and Francis Ltd, , 2 jan. 2019. .

MARSDEN, P. M. et al. Predicting protein–ligand binding affinities: a low scoring game? **Org. Biomol. Chem.**, v. 2, n. 22, p. 3267–3273, 21 nov. 2004.

MAZZON, M. et al. Identification of broad-spectrum antiviral compounds by targeting viral entry. **Viruses**, v. 11, n. 2, 1 fev. 2019.

MIDGLEY, C. M. et al. An in-depth analysis of original antigenic sin in dengue virus infection. **Journal of virology**, v. 85, n. 1, p. 410–21, 1 jan. 2011.

MINISTÉRIO DA SAÚDE. Manual de vigilância epidemiológica de febre amarela. 2019.

MINISTÉRIO DA SAÚDE. Monitoramento de Febre Amarela. Secretaria de Vigilância em Saúde. n 18, 06/2019.

MINISTÉRIO DA SAÚDE. Boletim epidemiológico. Secretaria de Vigilância em Saúde. v 50 n 22, 09/2019.

MIRZA, S. B. et al. Virtual screening of eighteen million compounds against dengue virus: Combined molecular docking and molecular dynamics simulations study. **Journal of Molecular Graphics and Modelling**, v. 66, p. 99–107, 2016.

MIRZA, S. B. et al. Discovery of selective dengue virus inhibitors using combination of molecular fingerprint-based virtual screening protocols, structure-based pharmacophore model development, molecular dynamics simulations and in vitro studies. **Journal of Molecular Graphics and Modelling**, v. 79, p. 88–102, 2018.

MISHRA, B.; BEHERA, B. The Mysterious Zika Virus: Adding to the Tropical Flavivirus Mayhem. **Journal of Postgraduate Medicine**, v. 62, n. 4, p. 249–254, dez. 2016.

MLAKAR, J. et al. Zika Virus Associated with Microcephaly. **New England Journal of Medicine**, v. 374, n. 10, p. 951–958, 10 mar. 2016.

MONATH, T. P. et al. Comparative safety and immunogenicity of two yellow fever 17D vaccines (ARILVAX and YF-VAX) in a phase III multicenter, double-blind clinical trial. **American Journal of Tropical Medicine and Hygiene**, v. 66, n. 5, p. 533–541, 2002.

MONATH, T. P.; VASCONCELOS, P. F. C. Yellow fever. **Journal of clinical virology: the official publication of the Pan American Society for Clinical Virology**, v. 64, p. 160–73, 1 mar. 2015.

MOSMANN, T. Rapid colorimetric assay for cellular growth and survival: Application to proliferation and cytotoxicity assays. **Journal of Immunological Methods**, v. 65, n. 1–2, p. 55–63, 1983.

MOTTA, I. J. F. et al. Evidence for Transmission of Zika Virus by Platelet Transfusion. **New England Journal of Medicine**, v. 375, n. 11, p. 1101–1103, 15 set. 2016.

MUELLER, N. H. et al. Characterization of the West Nile virus protease substrate specificity and inhibitors. **International Journal of Biochemistry and Cell Biology**, v. 39, n. 3, p. 606–614, 2007.

MUKHOPADHYAY, S.; KUHN, R. J.; ROSSMANN, M. G. A structural perspective of the. v. 3, n. January, 2005.

MURRAY, C. L.; JONES, C. T.; RICE, C. M. Architects of assembly: roles of Flaviviridae non-structural proteins in virion morphogenesis. **Nature reviews. Microbiology**, v. 6, n. 9, p. 699–708, 2008.

MURRELL, S.; WU, S.-C.; BUTLER, M. Review of dengue virus and the development of a vaccine. **Biotechnology Advances**, v. 29, n. 2, p. 239–247, mar. 2011.

MUSSO, D.; LANTERI, M. C. Emergence of Zika virus: where does it come from and where is it going to? **The Lancet. Infectious diseases**, v. 17, n. 3, p. 255, 1 mar. 2017.

MUSSO, D.; NILLES, E. J.; CAO-LORMEAU, V.-M. Rapid spread of emerging Zika virus in the Pacific area. **Clinical Microbiology and Infection**, v. 20, n. 10, p. O595–O596, out. 2014.

NETTO, C. et al. Experimental infection of BHK21 and vero cell lines with different Mycoplasma spp. **Brazilian Journal of Microbiology**, v. 45, n. 4, p. 1513–1519, 2014.

NEUFELDT, C. J. et al. Rewiring cellular networks by members of the Flaviviridae family. **Nature Publishing Group**, v. 16, 2018.

NEWMAN, D. J.; CRAGG, G. M. **Natural Products as Sources of New Drugs from 1981 to 2014** *Journal of Natural Products* American Chemical Society. 25 mar. 2016. .

NIKOLAY, A. et al. Propagation of Brazilian Zika virus strains in static and suspension cultures using Vero and BHK cells. **Vaccine**, v. 36, n. 22, p. 3140–3145, 24 maio 2018.

NOBLE, C. G. et al. Ligand-Bound Structures of the Dengue Virus Protease Reveal the Active Conformation. **Journal of Virology**, v. 86, n. 1, p. 438–446, 2012a.

NOBLE, C. G. et al. Ligand-bound structures of the dengue virus protease reveal the active conformation. **Journal of virology**, v. 86, n. 1, p. 438–46, 2012b.

NORMILE, D. Surprising New Dengue Virus Throws a Spanner in Disease Control Efforts. **Science**, v. 342, n. 6157, p. 415–415, 25 out. 2013.

ODE, H. et al. **Molecular dynamics simulation in virus research** *Frontiers in Microbiology* Frontiers Research Foundation, , 2012. .

OLIVEIRA, A. S. de 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.

OSORIO, J. E. et al. Development of a recombinant, chimeric tetravalent dengue vaccine candidate. **Vaccine**, v. 33, n. 50, p. 7112–7120, 10 dez. 2015.

OSTERHOLM, M. T. et al. Transmission of Ebola viruses: What we know and what we do not know. **mBio**, v. 6, n. 2, 19 fev. 2015.

PAMBUDI, S. et al. A small compound targeting the interaction between nonstructural proteins 2B and 3 inhibits dengue virus replication. **Biochemical and Biophysical Research Communications**, v. 440, n. 3, p. 393–398, 2013.

- PASTORINO, B. et al. Role of host cell factors in flavivirus infection: Implications for pathogenesis and development of antiviral drugs. **Antiviral Research**, v. 87, n. 3, p. 281–294, set. 2010.
- PATTERSON, J.; SAMMON, M.; GARG, M. Dengue, Zika and Chikungunya: Emerging Arboviruses in the New World. **Western Journal of Emergency Medicine**, v. 17, n. 6, p. 671–679, 2016a.
- PATTERSON, J.; SAMMON, M.; GARG, M. Dengue, Zika and Chikungunya: Emerging Arboviruses in the New World. **Western Journal of Emergency Medicine**, v. 17, n. 6, p. 671–679, 2016b.
- PETTERSSON, J. H.-O. et al. How Did Zika Virus Emerge in the Pacific Islands and Latin America? **mBio**, v. 7, n. 5, p. e01239-16, 11 nov. 2016.
- PHANTHANAWIBOON, S. et al. Construction of a high-yield dengue virus by replacing nonstructural proteins 3–4B without increasing virulence. **Biochemical and Biophysical Research Communications**, v. 495, n. 1, p. 1221–1226, 1 jan. 2018.
- PHOO, W. W. et al. Structure of the NS2B-NS3 protease from Zika virus after self-cleavage. **Nature Communications**, v. 7, 2016.
- POLIKAR, R. Ensemble based systems in decision making. **IEEE Circuits and Systems Magazine**, v. 6, n. 3, p. 21–45, Third 2006.
- POSSAS, C. et al. Yellow Fever Outbreak in Brazil: The Puzzle of Rapid Viral Spread and Challenges for Immunisation. **Memórias do Instituto Oswaldo Cruz**, v. 113, n. 10, 2018. Disponível em: <http://www.scielo.br/scielo.php?script=sci_abstract&pid=S0074-02762018001000200&lng=en&nrm=iso&tlng=en>. Acesso em: 25 mar. 2021.
- PRECIOSO, A. R. et al. Clinical evaluation strategies for a live attenuated tetravalent dengue vaccine. **Vaccine**, v. 33, n. 50, p. 7121–7125, 10 dez. 2015.
- PRYOR, M. J. et al. Nuclear Localization of Dengue Virus Nonstructural Protein 5 Through Its Importin α/β -Recognized Nuclear Localization Sequences is Integral to Viral Infection. **Traffic**, v. 8, n. 7, p. 795–807, jul. 2007.
- PUBCHEM. **Aprotinin from bovine lung, ca. 140 U/mg**. Disponível em: <<https://pubchem.ncbi.nlm.nih.gov/compound/122403016>>. Acesso em: 26 mar. 2021.
- RAMAKRISHNAIAH, V. et al. Exosome-mediated transmission of hepatitis C virus between human hepatoma Huh7.5 cells. **Proceedings of the National Academy of Sciences**, v. 110, n. 32, p. 13109–13113, 6 ago. 2013.
- RASMUSSEN, A. et al. Antibody-Dependent Enhancement and Zika: Real Threat or Phantom Menace? **Frontiers in Cellular and Infection Microbiology**, v. 8, p. 44, 2018.
- REYES-DEL VALLE, J. et al. Dengue Virus Cellular Receptors and Tropism. **Current Tropical Medicine Reports**, v. 1, n. 1, p. 36–43, 1 mar. 2014.
- RICO-HESSE, R. Microevolution and virulence of dengue viruses. **Advances in virus research**, v. 59, p. 315–41, 2003.

ROTHHAMMER, V.; QUINTANA, F. J. **The aryl hydrocarbon receptor: an environmental sensor integrating immune responses in health and disease** *Nature Reviews Immunology* Nature Publishing Group, 1 mar. 2019. .

ROTHMAN, A. L. **Immunity to dengue virus: A tale of original antigenic sin and tropical cytokine storms** *Nature Reviews Immunology*, ago. 2011. .

ROY, K.; KAR, S.; DAS, R. N. Statistical Methods in QSAR/QSPR. In: [s.l: s.n.]p. 37–59.

SABCHAREON, A. et al. Protective efficacy of the recombinant, live-attenuated, CYD tetravalent dengue vaccine in Thai schoolchildren: a randomised, controlled phase 2b trial. **The Lancet**, v. 380, n. 9853, p. 1559–1567, 3 nov. 2012.

SAIZ, J. C. et al. Zika Virus: What have we learnt since the start of the recent epidemic? **Frontiers in Microbiology**, v. 8, n. AUG, p. 1–25, 2017.

SALUM, L. B.; ANDRICOPULO, A. D. Fragment-based QSAR: perspectives in drug design. **Molecular Diversity**, v. 13, n. 3, p. 277–285, 31 ago. 2009.

SANGUINETTI, M. C.; TRISTANI-FIROUZI, M. **hERG potassium channels and cardiac arrhythmia** *Nature* Nature Publishing Group, , 23 mar. 2006. .

SANTOS, N. Q. et al. Guillain-Barré Syndrome in the Course of Dengue: Case Report. **Arquivos de Neuro-Psiquiatria**, v. 62, n. 1, p. 144–146, mar. 2004.

SCHAMES, J. R. et al. Discovery of a Novel Binding Trench in HIV Integrase. **Journal of Medicinal Chemistry**, v. 47, n. 8, p. 1879–1881, abr. 2004.

SCHNEIDER, G.; FECHNER, U. Computer-Based de Novo Design of Drug-like Molecules. **Nature Reviews Drug Discovery**, v. 4, n. 8, p. 649–663, ago. 2005.

SCREATON, G. et al. **New insights into the immunopathology and control of dengue virus infection** *Nature Reviews Immunology* Nature Publishing Group, , 1 dez. 2015. .

SELISKO, B. et al. Regulation of Flavivirus RNA synthesis and replication. **Current opinion in virology**, v. 9, p. 74–83, dez. 2014.

SHEARER, F. M. et al. Global Yellow Fever Vaccination Coverage from 1970 to 2016: An Adjusted Retrospective Analysis. **The Lancet Infectious Diseases**, v. 17, n. 11, p. 1209–1217, 1 nov. 2017.

SHELLEY, J. C. et al. Epik: A software program for pKa prediction and protonation state generation for drug-like molecules. **Journal of Computer-Aided Molecular Design**, v. 21, n. 12, p. 681–691, 2007.

SHEPARD, D. S. et al. The global economic burden of dengue: a systematic analysis. **The Lancet Infectious Diseases**, v. 16, n. 8, p. 935–941, ago. 2016.

SIDDHARTA, A. et al. Virucidal activity of world health organization-recommended formulations against enveloped viruses, including zika, ebola, and emerging coronaviruses. **Journal of Infectious Diseases**, v. 215, n. 6, p. 902–906, 15 mar. 2017.

SILVA MARINHO, P. E.; KROON, E. G. Flaviviruses as agents of childhood central nervous system infections in Brazil. **New Microbes and New Infections**, v. 30, 1 jul. 2019.

SIMMONDS, P. et al. ICTV ICTV Virus Taxonomy Profile : Flaviviridae. p. 2–3, 2017.

SIRIVICHAYAKUL, C. et al. Safety and Immunogenicity of a Tetravalent Dengue Vaccine Candidate in Healthy Children and Adults in Dengue-Endemic Regions: A Randomized, Placebo-Controlled Phase 2 Study. **Journal of Infectious Diseases**, v. 213, n. 10, p. 1562–1572, 15 maio 2016.

SORAIRES SANTACRUZ, M. C. et al. Synthesis, antiviral evaluation and molecular docking studies of N 4 -aryl substituted/unsubstituted thiosemicarbazones derived from 1-indanones as potent anti-bovine viral diarrhea virus agents. **Bioorganic & Medicinal Chemistry**, v. 25, n. 15, p. 4055–4063, 1 ago. 2017.

SOUTHWOOD, T. R. et al. Studies on the life budget of *Aedes aegypti* in Wat Samphaya, Bangkok, Thailand. **Bulletin of the World Health Organization**, v. 46, n. 2, p. 211–26, 1972.

STANAWAY, J. D. et al. The global burden of dengue: an analysis from the Global Burden of Disease Study 2013. **The Lancet. Infectious diseases**, v. 16, n. 6, p. 712–723, 2016.

STAPLES, J. E. et al. Interim Guidelines for the Evaluation and Testing of Infants with Possible Congenital Zika Virus Infection — United States, 2016. **MMWR. Morbidity and Mortality Weekly Report**, v. 65, n. 3, p. 1–5, 26 jan. 2016.

STAPLES, J. E.; MONATH, T. P. Yellow Fever: 100 Years of Discovery. **JAMA**, v. 300, n. 8, p. 960, 27 ago. 2008.

STEUER, C. et al. Synthesis and biological evaluation of α -ketoamides as inhibitors of the Dengue virus protease with antiviral activity in cell-culture. **Bioorganic and Medicinal Chemistry**, v. 19, n. 13, p. 4067–4074, 2011.

TAKAGI, Y. et al. Discovery of novel cyclic peptide inhibitors of dengue virus NS2B-NS3 protease with antiviral activity. **Bioorganic and Medicinal Chemistry Letters**, v. 27, n. 15, p. 3586–3590, 2017.

TAKAHASHI, H. et al. Establishment of a robust dengue virus NS3-NS5 binding assay for identification of protein-protein interaction inhibitors. **Antiviral Research**, v. 96, n. 3, p. 305–314, 2012.

TAKEUCHI, H.; BABA, M.; SHIGETA, S. An application of tetrazolium (MTT) colorimetric assay for the screening of anti-herpes simplex virus compounds. **Journal of virological methods**, v. 33, n. 1–2, p. 61–71, jun. 1991.

TAY, M. Y. F. et al. The C-terminal 50 amino acid residues of dengue NS3 protein are important for NS3-NS5 interaction and viral replication. **The Journal of biological chemistry**, v. 290, n. 4, p. 2379–94, 23 jan. 2015.

TEN BRINK, T.; EXNER, T. E. Influence of Protonation, Tautomeric, and Stereoisomeric States on Protein–Ligand Docking Results. **Journal of Chemical Information and Modeling**, v. 49, n. 6, p. 1535–1546, 22 jun. 2009.

KOZAKOV, D. The FTMap family of web servers for determining and characterizing ligand-binding hot spots of proteins. **Nat Protoc.** v. 10, n. 5, p. 733-755. 2015.

TIAN, Y.-S. et al. **Dengue Virus and Its Inhibitors: A Brief Review** **Chem. Pharm. Bull.**

TIMIRI, A. K.; SINHA, B. N.; JAYAPRAKASH, V. Progress and prospects on DENV protease inhibitors. **European Journal of Medicinal Chemistry**, v. 117, p. 125–143, 2016.

TORRESI, J.; EBERT, G.; PELLEGRINI, M. Vaccines licensed and in clinical trials for the prevention of dengue. **Human Vaccines and Immunotherapeutics**, v. 13, n. 5, p. 1059–1072, 2017a.

TORRESI, J.; EBERT, G.; PELLEGRINI, M. Vaccines licensed and in clinical trials for the prevention of dengue. **Human Vaccines & Immunotherapeutics**, v. 13, n. 5, p. 1059–1072, 4 maio 2017b.

TROPSHA, A.; GRAMATICA, P.; GOMBAR, V. The Importance of Being Earnest: Validation is the Absolute Essential for Successful Application and Interpretation of QSPR Models. **QSAR & Combinatorial Science**, v. 22, n. 1, p. 69–77, abr. 2003.

TUBOI, S. H. et al. Clinical and epidemiological characteristics of yellow fever in Brazil: analysis of reported cases 1998-2002. **Transactions of the Royal Society of Tropical Medicine and Hygiene**, v. 101, n. 2, p. 169–175, fev. 2007.

VAN DER SCHAAR, H. M.; WILSCHUT, J. C.; SMIT, J. M. Role of antibodies in controlling dengue virus infection. **Immunobiology**, v. 214, n. 7, p. 613–629, jul. 2009.

VASCONCELOS, P. F.; VASCONCELOS, P. F. Single shot of 17D vaccine may not confer life-long protection against yellow fever. **Memórias do Instituto Oswaldo Cruz**, v. 113, n. 2, p. 135–137, 27 nov. 2017a.

VASCONCELOS, P. F.; VASCONCELOS, P. F. Single shot of 17D vaccine may not confer life-long protection against yellow fever. **Memórias do Instituto Oswaldo Cruz**, v. 113, n. 2, p. 135–137, 27 nov. 2017b.

VEBER, D. F. et al. Molecular Properties That Influence the Oral Bioavailability of Drug Candidates. **Journal of Medicinal Chemistry**, v. 45, n. 12, p. 2615–2623, 6 jun. 2002.

VERDONK, M. L. et al. Improved protein-ligand docking using GOLD. **Proteins: Structure, Function and Genetics**, v. 52, n. 4, p. 609–623, 1 set. 2003.

VERMA, P.; TIWARI, M.; TIWARI, V. In silico high-throughput virtual screening and molecular dynamics simulation study to identify inhibitor for AdeABC efflux pump of *Acinetobacter baumannii*. **Journal of Biomolecular Structure and Dynamics**, v. 36, n. 5, p. 1182–1194, 4 abr. 2018.

VILLAR, L. et al. Efficacy of a Tetravalent Dengue Vaccine in Children in Latin America. **New England Journal of Medicine**, v. 372, n. 2, p. 113–123, 8 jan. 2015.

VISWANATHAN, U. et al. Identification of a Novel Inhibitor of Dengue Virus Protease through Use of a Virtual Screening Drug Discovery Web Portal. **Journal of Chemical Information and Modeling**, v. 54, n. 10, p. 2816–2825, 2014.

WAGGONER, J. J.; PINSKY, B. A. Zika Virus: Diagnostics for an Emerging Pandemic Threat. **Journal of clinical microbiology**, v. 54, n. 4, p. 860–7, 1 abr. 2016.

WARREN, G. L. et al. A Critical Assessment of Docking Programs and Scoring Functions. **Journal of Medicinal Chemistry**, v. 49, n. 20, p. 5912–5931, 5 out. 2006.

WAYBRIGHT, T. J.; BRITT, J. R.; MCCLOUD, T. G. Overcoming problems of compound storage in DMSO: solvent and process alternatives. **Journal of biomolecular screening**, v. 14, n. 6, p. 708–15, jul. 2009.

WEAVER, S. C.; BARRETT, A. D. T. Transmission cycles, host range, evolution and emergence of arboviral disease. **Nature Reviews Microbiology**, v. 2, n. 10, p. 789–801, 2004.

WEIGEL, L. F. et al. Phenylalanine and Phenylglycine Analogues as Arginine Mimetics in Dengue Protease Inhibitors. 2015.

WIKAN, N.; SMITH, D. R. Zika virus: history of a newly emerging arbovirus. **The Lancet Infectious Diseases**, v. 16, n. 7, p. e119–e126, jul. 2016.

WINKLER, D. A. The role of quantitative structure--activity relationships (QSAR) in biomolecular discovery. **Briefings in bioinformatics**, v. 3, n. 1, p. 73–86, 2002.

WORLD HEALTH ORGANIZATION. Dengue: guidelines for diagnosis, treatment, prevention, and control. **Special Programme for Research and Training in Tropical Diseases**, p. 147, 2009.

WORLD HEALTH ORGANIZATION. GACVS Statement on Dengvaxia ® (CYD-TDV). **WHO**, 2017.

WORLD HEALTH ORGANIZATION. Zika situation report. **WHO**, 2017.

WU, D. W. et al. Poliresulen, a novel NS2B/NS3 protease inhibitor, effectively inhibits the replication of DENV2 virus in BHK-21 cells. **Acta Pharmacologica Sinica**, v. 36, n. 9, p. 1126–1136, 2015a.

WU, H. et al. Novel dengue virus NS2B/NS3 protease inhibitors. **Antimicrobial Agents and Chemotherapy**, v. 59, n. 2, p. 1100–1109, 2015b.

XU, J.; HAGLER, A. Chemoinformatics and Drug Discovery. **Molecules**, v. 7, n. 8, p. 566–600, 30 ago. 2002.

XU, S. et al. Critical Effect of Peptide Cyclization on the Potency of Peptide Inhibitors against Dengue Virus NS2B-NS3 Protease. **Journal of Medicinal Chemistry**, v. 55, n. 15, p. 6881–6887, 9 ago. 2012.

YAMASHITA, A. et al. Origin and distribution of divergent dengue virus: novel database construction and phylogenetic analyses. **Future Virology**, v. 8, n. 11, p. 1061–1083, 17 nov. 2013.

YANG, J. M.; SHEN, T. W. A pharmacophore-based evolutionary approach for screening selective estrogen receptor modulators. **Proteins: Structure, Function and Genetics**, v. 59, n. 2, p. 205–220, 1 maio 2005.

YANG, S.-Y. Pharmacophore modeling and applications in drug discovery: challenges and recent advances. **Drug Discovery Today**, v. 15, p. 444–450, 2010.

YAP, C. W. PaDEL-descriptor: An open source software to calculate molecular descriptors and fingerprints. **Journal of Computational Chemistry**, v. 32, n. 7, p. 1466–1474, maio 2011.

YIN, Z. et al. Peptide inhibitors of dengue virus NS3 protease. Part 1: Warhead. **Bioorganic and Medicinal Chemistry Letters**, v. 16, n. 1, p. 36–39, 1 jan. 2006.

YOKOYAMA, S.; STARMER, W. T. Possible Roles of New Mutations Shared by Asian and American Zika Viruses. **Molecular Biology and Evolution**, v. 34, n. 3, p. msw270, 12 jan. 2017.

YUAN, S. et al. Structure-based discovery of clinically approved drugs as Zika virus NS2B-NS3 protease inhibitors that potently inhibit Zika virus infection in vitro and in vivo. **Antiviral Research**, v. 145, p. 33–43, 2017.

ZHANG, C. et al. Structure of the NS5 methyltransferase from Zika virus and implications in inhibitor design. **Biochemical and Biophysical Research Communications**, v. 492, n. 4, p. 624–630, 2017.

ZHAO, Y. H. et al. Evaluation of human intestinal absorption data and subsequent derivation of a quantitative structure - Activity relationship (QSAR) with the Abraham descriptors. **Journal of Pharmaceutical Sciences**, v. 90, n. 6, p. 749–784, 2001.

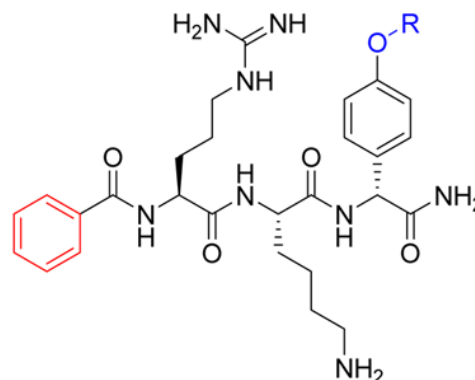
ZHU, J. et al. Insight into the structural requirements of narlaprevir-type inhibitors of NS3/NS4A protease based on HQSAR and molecular field analyses. **Combinatorial chemistry & high throughput screening**, v. 15, n. 6, p. 439–50, jul. 2012.

ZOU, J. et al. Characterization of Dengue Virus NS4A and NS4B Protein Interaction. **Journal of Virology**, v. 89, n. 7, p. 3455–3470, 1 abr. 2015.

9. APÊNDICE A

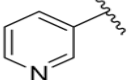
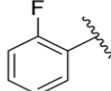
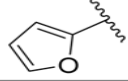
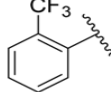
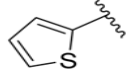
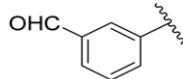
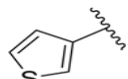
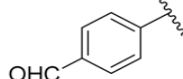
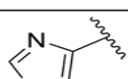
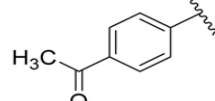
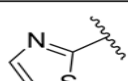
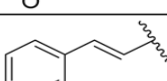
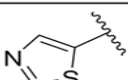
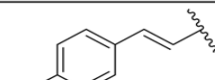
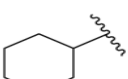
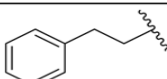
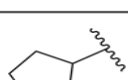
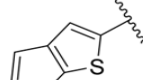

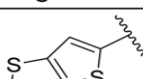

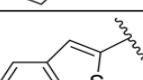
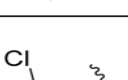
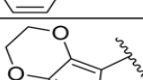
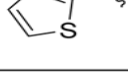
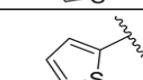
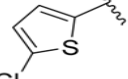
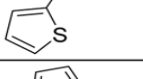
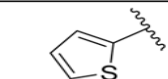
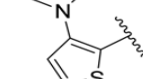
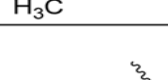
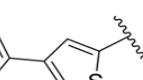
Nº	Estrutura
1	Bz-Arg-Lys-L-Phg-NH ₂
2	Bz-Arg-Lys-D-Phg-NH ₂
3	Bz-Arg-Lys-(4-OH)-L-Phg-NH ₂
4	Bz-Arg-Lys-(4-OH)-D-Phg-NH ₂
5	Bz-Arg-Lys-(4-benzyloxy)-L-Phg-NH ₂
6	Bz-Arg-Lys-(4-benzyloxy)-D-Phg-NH ₂
7	Bz-Lys-(4-benzyloxy)-D-Phg-NH ₂
8	Bz-Arg-(4-benzyloxy)-D-Phg-NH ₂
9	Bz-Lys-Arg-(4-benzyloxy)-D-Phg-NH ₂
10	Bz-D-Lys-D-Arg-(4-benzyloxy)-D-Phg-NH ₂
11	Bz-D-Arg-D-Lys-(4-benzyloxy)-L-Phg-NH ₂
12	Bz-D-Lys-D-Arg-(4-benzyloxy)-L-Phg-NH ₂

Estrutura geral (1 a 30)



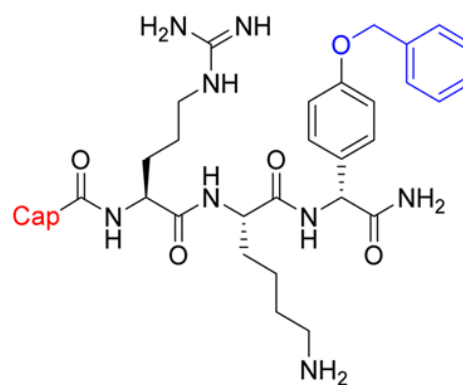
Nº	R
13	
14	
15	
16	
17	
18	
19	
20	
21	

Nº	R
22	
23	
24	
25	
26	
27	
28	
29	
30	
87	

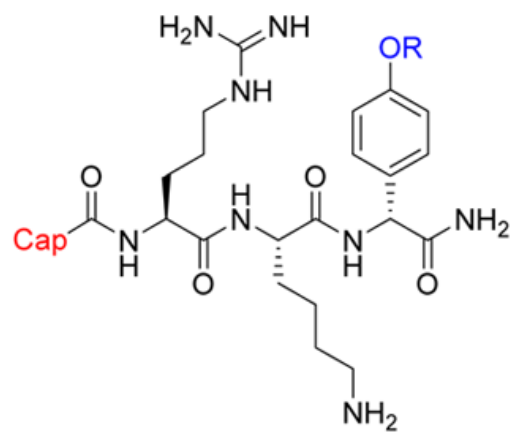
31		47	
32		48	
33		49	
34		50	
35		51	
36		52	
37		53	
38		54	
39		55	
40		56	
41		57	
42		58	
43		59	
44		60	
45		61	
46		62	

63		
64		
65		
66		
67		
68		
69		
70		
71		
72		
73		
74		
75		
76		
77		
78		
79		
80		
81		
82		
83		
84		
85		
86		

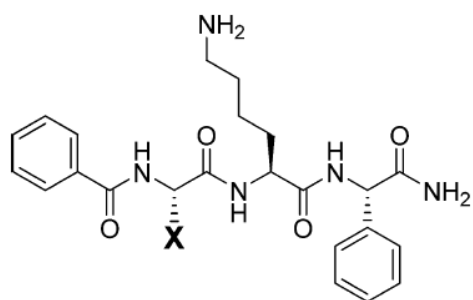
Estrutura geral (31 a 62)



Estrutura geral (63 a 86)



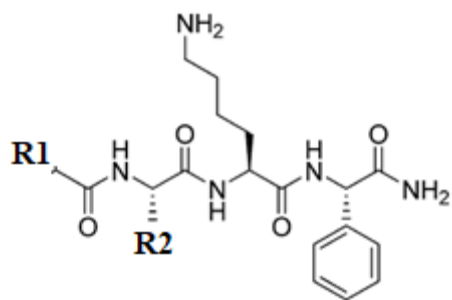
Estrutura Geral (88 a 111)

Bz-X-Lys-Phg-NH₂

88		29a
90		30a
92		31a
94		32a
96		33a
98		34a

89		29b
91		30b
93		31b
95		32b
97		33b
99		34b

100		35
101		39
102		37
103		40



R1-R2-Lys-Phe-NH₂

104		Arg
105		(4-amidino)Phe
106		(3-amidino)Phe
107		(4-guanidino)Phe
108		Arg
109		(4-amidino)Phe
110		(3-amidino)Phe
111		(4-guanidino)Phe

ORIGINAL ARTICLE

Antibacterial activity of synthetic 1,3-bis(aryloxy)propan-2-amines against Gram-positive bacteria

Mateus S. M. Serafim¹ | Stefânia N. Lavorato² | Thales Kronenberger³ |
 Yamara V. Sousa¹ | Grazielle P. Oliveira⁴ | Simone G. dos Santos⁴ | Erna G. Kroon⁴ |
 Vinícius G. Maltarollo² | Ricardo J. Alves² | Bruno E. F. Mota¹

¹Departamento de Análises Clínicas e Toxicológicas, Faculdade de Farmácia, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Brazil

²Departamento de Produtos Farmacêuticos, Faculdade de Farmácia, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Brazil

³Department of Internal Medicine VIII, University Hospital Tübingen, Tübingen, Germany

⁴Departamento de Microbiologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Brazil

Correspondence

Bruno E. F. Mota, Departamento de Análises Clínicas e Toxicológicas, Faculdade de Farmácia, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Brazil.
 Email: brunofmota@gmail.com

Present address

Stefânia N. Lavorato, Centro das Ciências Biológicas e da Saúde, Universidade Federal do Oeste da Bahia (UFOB), Rua Professor José Seabra de Lemos, 316. CEP 47808-021., Barreiras, Bahia, Brazil

Funding information

Fundação de Amparo à Pesquisa do Estado de Minas Gerais; Coordenação de Aperfeiçoamento de Pessoal de Nível Superior; Conselho Nacional de Desenvolvimento Científico e Tecnológico; Pró-Reitoria de Pesquisa, Universidade Federal de Minas Gerais

Abstract

Synthetic 1,3-bis(aryloxy)propan-2-amines have been shown in previous studies to possess several biological activities, such as antifungal and antiprotozoal. In the present study, we describe the antibacterial activity of new synthetic 1,3-bis(aryloxy)propan-2-amines against Gram-positive pathogens (*Streptococcus pyogenes*, *Enterococcus faecalis* and *Staphylococcus aureus*) including Methicillin-resistant *S. aureus* strains. Our compounds showed minimal inhibitory concentrations (MIC) in the range of 2.5–10 µg/ml (5.99–28.58 µM), against different bacterial strains. The minimal bactericidal concentrations found were similar to MIC, suggesting a bactericidal mechanism of action of these compounds. Furthermore, possible molecular targets were suggested by chemical similarity search followed by docking approaches. Our compounds are similar to known ligands targeting the cell division protein *FtsZ*, Quinolone resistance protein *norA* and the Enoyl-[acyl-carrier-protein] reductase *FabI*. Taken together, our data show that synthetic 1,3-bis(aryloxy)propan-2-amines are active against Gram-positive bacteria, including multidrug-resistant strains and can be a promising lead in the development of new antibacterial compounds for the treatment of these infections.

KEYWORDS

3-bis(aryloxy)propan-2-amines, antibacterial, gram-positive bacteria, MRSA, synthetic 1, target prediction

1 | INTRODUCTION

Infections with antibiotic-resistant bacteria are a persistent problem to Public Health worldwide. Only in the US, at least 2 million

people become infected with bacteria that harbor some type of resistance to commercially available antibiotics, of whom 23,000 die each year, ultimately estimating in \$20 billion the increased health care costs (CDC, 2013). This scenario has become even

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2019 The Authors. *MicrobiologyOpen* published by John Wiley & Sons Ltd.

worse if we consider that in the past 40 years only two classes of narrow-spectrum antibiotics (daptomicin and linezolid) were developed (Clatworthy, Pierson, & Hung, 2007). The scarcity of new therapeutic options against antibiotic-resistant strains has led to the return of older drugs previously disregarded due to its significant toxicity, such as colistin (Li et al., 2006). However, resistance mechanisms continue to emerge even for these drugs leading to the appearance of virtually untreatable infections (Malhotra-Kumar et al., 2016).

Among the infections with resistant bacteria, one can highlight the group of pathogens known as ESKAPE (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* species). These infections are associated with longer periods of hospitalization, increases in hospital costs, higher use of antimicrobial drugs and higher mortality rates. The number of deaths caused by infection with methicillin-resistant *S. aureus* (MRSA) strains, for instance, surpassed the number of deaths from HIV/AIDS and tuberculosis combined in the US (Boucher et al., 2009).

The main strategy to overcome the problem of bacterial resistance is the development of new antibacterial agents. Regarding this strategy, the synthesis of new compounds and modification of the existing ones is promising and can extend the options of new drugs with a broader spectrum of activity, lower toxicity and/or reduced sensitivity to resistance mechanisms (Silver, 2011). This approach has resulted in the introduction

of some new antibacterial agents for clinical use, such as retapamulin, a compound derived from pleuromutilin, and some of the classical modifications of penicillins, the aminopenicillins (Gao et al., 2017; Lobanovska & Pilla, 2017).

The 1,3-Bis(aryloxy)propan-2-amines are a class of compounds synthesized by the amination of 1,3-diaryloxypropyl toluenesulfonate, whose biological potential has not yet been extensively studied. We recently reported the trypanocidal (Lavorato, Sales Júnior, Murta, Romanha, & Alves, 2015) and leishmanicidal (Lavorato et al., 2017) activities of several compounds of this class, but their antibacterial action remains to be further studied.

In the present work we have evaluated the antibacterial activity of a series of 1,3-bis(aryloxy)propan-2-amines, several synthetic intermediates and *N*-substituted amines (Figure 1).

2 | MATERIALS AND METHODS

2.1 | Chemistry

Compounds CPD1–CPD21, CPD23–CPD28 and CPD37 were synthesized and characterized by their IR, ¹H and ¹³C NMR spectra and melting points as previously described (Lavorato et al., 2017). Compounds CPD22, CPD29–CPD36, CPD38 and CPD39 were synthesized and fully characterized as described in Appendix.

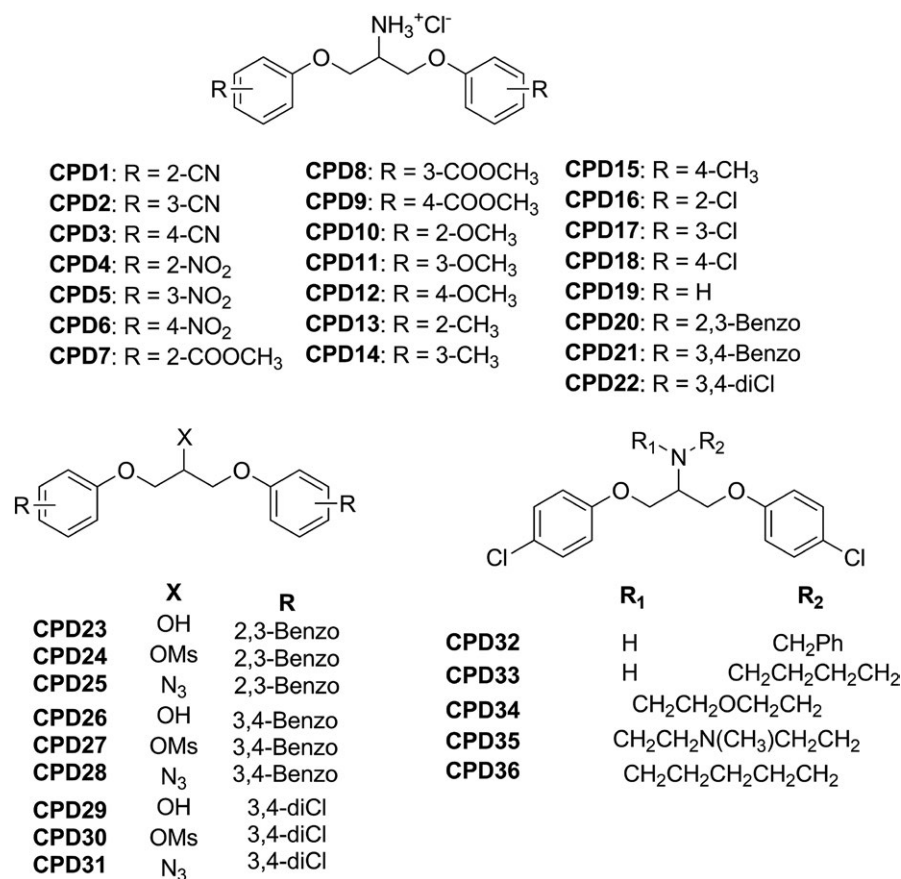


FIGURE 1 Compounds screened for antibacterial activity in the present study

2.2 | Cell lines

Vero cells (ATCC number CCL-81) were maintained in Minimal Essential Medium (MEM; Cultilab, Brazil), while BSC-40 cells (ATCC number CRL-2761) were maintained in Dulbecco's Modified Essential Medium (DMEM; Cultilab, Brazil). Both media were supplemented with 5% fetal bovine serum (Cultilab, Brazil), 200 U/ml of penicillin, 100 µg/ml of streptomycin and 2.5 µg/ml of amphotericin B. The bacteria strains used were *Escherichia coli* (ATCC n° 35218), *K. pneumoniae* (ATCC n° 13883), *P. aeruginosa* (ATCC n° 27853), *E. faecalis* (ATCC n° 29212), *S. aureus* (ATCC n° 29213), *Streptococcus pyogenes* (ATCC n° 19615), MRSA (ATCC n° 43300) and five clinical strains of MRSA (*mecA* positives) (Gomes et al., 2015).

2.3 | Minimal inhibitory concentration determination

The antibacterial activity of the compounds was evaluated using the broth microdilution method in 96-well microplates according to the Clinical and Laboratory Standards Institute protocol (CLSI, 2017). First, synthetic compounds were diluted in Mueller Hinton broth (MHB; Oxoid, Thermo Scientific, UK) to concentrations ranging from 20 to 2.5 µg/ml. The same volume of a bacterial suspension containing 10⁵ CFU/ml was added to each of the previous solutions, resulting in final compound concentrations from 10 to 1.25 µg/ml. After incubation at 35°C for 24 hr, the plates were inspected visually for inhibition of bacterial growth. In each plate was included a viability control (bacterial suspension only), an inhibitory control (MHB containing five times the minimal inhibitory concentration (MIC) of penicillin G for Gram-positive and Gentamicin for Gram-negative bacteria, or a serial dilution of vancomycin ranging from 16 to 2 µg/ml or 5.52 to 0.69 µM for MRSA strains) and a sterility control (medium only). All conditions were tested in triplicate and the results shown are representative of three independent assays.

2.4 | Minimal bactericidal concentration determination

To evaluate the minimal bactericidal concentration (MBC) of tested compounds, the content of wells that showed no visual growth in the previous experiments, plus the well containing the viability control were plated in Mueller Hinton agar plates. After incubation at 35°C for 24 hr, the colonies were counted and the percentage of inhibition was calculated. MBC is defined as the lowest compound's concentration that inhibits at least 99.9% of the bacterial cell count compared to nontreated viability control (Clinical and Laboratory Standards Institute, 1999).

2.5 | Cytotoxicity to mammalian cells

The cytotoxicity of active compounds to mammalian cells was assessed using the MTT reduction assay (Mosmann, 1983). Vero

and BSC-40 cells were seeded in 96-well plates (8 × 10⁴ cells per well) and incubated at 37°C and 5% CO₂ atmosphere. After 24 hr of incubation, 200 µL of fresh medium containing a serial dilution of compounds (10–1.25 µg/ml) were added to the plates. After 48 hr of incubation in the same conditions, 100 µL of MTT solution in MEM or DMEM (5 mg/ml) was added to each well and incubated for 3 hr at 37°C and 5% CO₂ atmosphere. The medium was removed and 100 µL of DMSO was used to solubilize formazan crystals. Absorbance at 570 nm of each well was read using a spectrophotometer (VersaMax, Molecular Devices). The cytotoxic concentration of 50% (CC₅₀) is defined as the lowest concentration of a specific compound that reduces by 50% the viability of cultured cells.

2.6 | Putative molecular target identification by 3D chemical similarity and interaction profiling by molecular docking

First, the lowest energy conformations of tested compounds showing antibacterial activity were obtained by conformational analysis performed on OMEGA 2.5.1.4 software (Hawkins, Skillman, Warren, Ellingson, & Stahl, 2010). Then, the database of compounds with known effects over *S. aureus*, *S. pyogenes* and *E. faecalis* proliferation was retrieved from ChEMBL v23 (Bento et al., 2014). The three obtained databases were filtered to remove entries without experimental activity determined, inactive compounds and mixtures of compounds. For all compounds, the structures had their protonation states calculated according to pH = 7.4 using fixpka software implemented on QUACPAC 1.7.0.2 (OpenEye Scientific Software, 2016) and, then, the lowest energy conformers were generated using OMEGA.

Chemical similarity queries were created for each active compound by considering common chemical features (rings, H-bond donors and acceptors, ions and hydrophobes) and the overall compound shape using the program ROCS 3.2.1.4 (Hawkins, Skillman, & Nicholls, 2007). ROCS software was used to identify the most similar compounds from the database against our queries. ROCS can overlay the library of conformers against a query composed of the shape and colors (representing chemical properties) derived from a compound. The output conformers were ranked according to their similarity with the query using a Tanimoto-combo coefficient (TC, a linear sum of Tanimoto coefficient for molecular shape and colors) and the compounds were considered for further analysis when TC > 1, representing at least 50% of chemical similarity (Rush, Grant, Mosyak, & Nicholls, 2005). Within this chemically similar dataset, compounds with experimental activity against molecular targets were identified and used in docking studies. Those targets were retrieved from the Protein Data Bank (PDB) or constructed using homology modeling.

Identified proteins were prepared by adding the adjusting protonation states of amino acids and fixing missing side-chain atoms (PrepWiz, Maestro v2017.4). Molecular docking was performed around the cocrystallized ligand of the different protein using the default settings of the Glide program (Glide v7.7, Maestro v2017.4) in extraprecision mode, with at least five poses selected for visual

inspection (Friesner et al., 2006). The amino acid residues were considered rigid and structural water molecules were kept during calculation. The employed docking protocol was evaluated with re-docking experiments. Our target prediction protocol was based on the previously published methodology (Vallone et al., 2018).

2.7 | Homology modeling

Homology model of the *S. aureus* NorA (UniProt accession number POA0J7) was inferred using the *E. coli* homolog (PDB code: 4ZP0, resolution: 2.0 Å, sequence similarity: 77.3%) as a template. 3D model of the SaNorA domain was generated using the online server HHPred (Söding, Biegert, & Lupas, 2005) for template identification and alignment followed by Modeller 9v19 (Eswar et al., 2006) for the model construction. The quality of the final structure was accessed by MolProbity (Davis, Murray, Richardson, & Richardson, 2004) showing three residues out of the Ramachandran allowed region, which was then fixed by the protein preparation step prior to docking.

3 | RESULTS

To investigate the antibacterial potential of 1,3-bis(aryloxy)propan-2-amines, 22 compounds of this class, variations in the nature

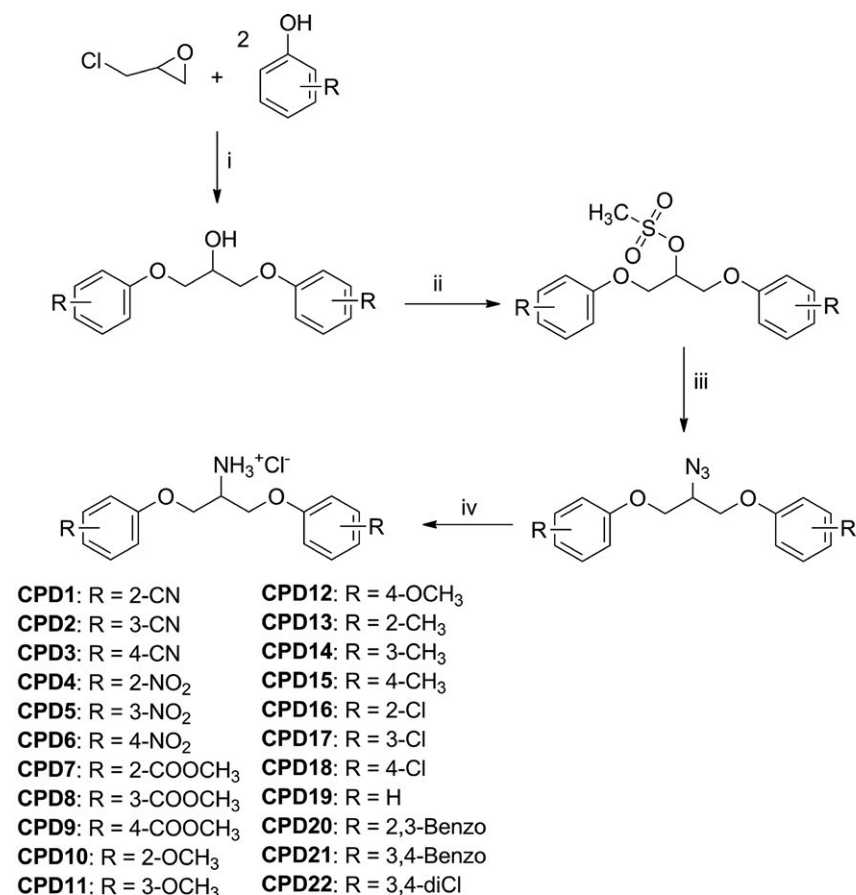
and position of the substituents on the aromatic ring, were evaluated against Gram-positive and Gram-negative bacteria. These compounds, named as CPD1–CPD22, were synthesized in four steps (Figure 2), as previously described by Lavorato et al. (2017).

3.1 | Initial screening for antibacterial activity

Among the compounds initially tested, four—CPD18, CPD20, CPD21 and CPD22—presented antibacterial activity at the concentration of 10 µg/ml. Among the six bacterial species tested (*Escherichia coli*, *K. pneumoniae*, *P. aeruginosa*, *E. faecalis*, *S. aureus* and *S. pyogenes*), the activity was observed only against Gram-positive bacteria. CPD20 and CPD22 inhibited the growth of all Gram-positive bacteria tested (*E. faecalis*, *S. aureus* and *S. pyogenes*), while CPD18 and CPD21 showed activity against *S. aureus* and *S. pyogenes*.

3.2 | Minimal inhibitory and MBC determination

Compounds that showed antibacterial activity in the initial screening were submitted to MIC determination by broth microdilution method. Among the four active compounds in the initial screening, CPD20 showed the best results, with MIC values of 2.5 µg/ml (6.58 µM) against *S. pyogenes* and *S. aureus* and 5 µg/ml (13.16 µM) against *E. faecalis*. CPD22 showed an MIC value of 2.5 µg/ml (5.99 µM) against *S. pyogenes*



Reagents and conditions: i) NaOH, H₂O, 60°C; ii) MsCl, pyr, 0°C-r.t.; iii) NaN₃, DMF, 80°C; iv) a) H₂/Pd-C, THF or DTT/TEA, MeCN; b) conc. HCl, MeOH.

FIGURE 2 Synthesis of 1,3-bis(aryloxy)propan-2-amines

and 5 µg/ml (11.97 µM) against *S. aureus* and *E. faecalis*. On the other hand, both compounds CPD18 and CPD21 showed an MIC of 10 µg/ml (28.68 and 26.32 µM, respectively) against *S. aureus* and *S. pyogenes*.

To evaluate the MBC, the content of each well that showed no visual growth in the previous experiments were plated on MH agar and colony counts were compared to counts obtained from viability controls. Some of the MBC were equivalent to its MIC values suggesting a bactericidal activity of the tested compounds. CPD 21 showed equivalence between MBC and MIC values, against *S. aureus* and *S. pyogenes*. CPD20 also showed equivalence against *S. pyogenes* and CPD22 against *S. aureus*. CPD18 was an exception to this scenario, in which the MBC values were not equivalent to its MIC values against *S. aureus* and *S. pyogenes*. The results of quantitative antibacterial assays are summarized in Table 1.

3.3 | Antibacterial activity against MRSA strains

In order to evaluate the efficacy of compounds against antibiotic resistant strains, we performed a broth microdilution method using MRSA. Corroborating the findings above, compound CPD20 showed MIC values of 2.5 µg/ml (6.58 µM) against all MRSA strains, being the most promising among all compounds tested. Values of MIC ranged from 2.5 to 5 µg/ml (5.99–11.97 µM) for CPD22 and from 5 to 10 µg/ml (13.16–26.32 µM) for CPD21. Compound CPD18 showed activity against four out of six MRSA strains, with MIC values of 10 µg/ml (28.68 µM). The MBC values were equivalent to its MIC values corroborating the hypothesis of a bactericidal activity of the tested compounds. CPD21 showed equivalence between MBC and MIC values against all MRSA strains. CPD20 showed the same values of MBC and MIC against MRSA strains n° 5749, 5912, 6100 and 6613 and CPD22 against MRSA 43300, 5749, 5912 and 6613. For compound CPD18, no bactericidal activity was observed for the concentrations tested. The results of antibacterial assays against MRSA strains are summarized in Table 2.

3.4 | Changes in chemical group in R position abolish the antibacterial activity of tested compounds

To investigate the importance of the amino group to antibacterial activity, compounds CPD23–CPD31, synthetic intermediates of the

most active amines CPD20, CPD21 and CPD22, were selected for biological testing. As shown in Figure 3, they were obtained in one, two or three steps according to the substituent in C-2.

In this second screening, we also prepared a series of secondary and tertiary amines for evaluation. As shown in Figure 4, compounds CPD38 and CPD39 were used as precursors to synthesize *N*-substituted amines CPD32–CPD36 and both compounds were obtained from alcohol CPD37. CPD37 was obtained as previously described by Lavorato et al. (2017). The ketone CPD38 was obtained from CPD37 by Albright–Goldman oxidation using DMSO and acetic anhydride (Fritsche, Elfringhoff, Fabian, & Lehr, 2008), while the tosylate CPD39 was prepared by reacting CPD37 with *p*-toluenesulfonyl chloride in dry pyridine (King & Bigelow, 1952). The secondary amines CPD32 and CPD33 were obtained by reductive amination reaction of CPD38 with benzylamine or butylamine, respectively, in the presence of NaCNBH₃ as reducing agent (Borch, Bernstein, & Durst, 1971). The nucleophilic substitution reaction between CPD39 and the heterocyclic amines morpholine, *N*-methylpiperazine and piperidine under heating at 100°C resulted in the tertiary amines CPD34, CPD35 and CPD36, respectively (Yuxiu, Guiqin, & Guangren, 2000).

None of these compounds presented antibacterial activity, with no complete inhibition of bacterial growth in all concentrations tested (up to 10 µg/ml, data not shown).

3.5 | Cytotoxicity concentration (CC₅₀) of active compounds in mammalian cells

We also evaluated the cytotoxicity in Vero and BSC-40 cell lines of the active compounds using the colorimetric MTT assay. Compounds CPD18 and CPD22 showed higher values of CC₅₀ for Vero (5.99 ± 0.09 µg/ml or 17.18 ± 0.26 µM and 5.47 ± 1.89 µg/ml or 13.1 ± 4.53 µM, respectively) and CPD18 and CPD21 for BSC-40 (5.73 ± 0.18 µg/ml or 16.43 ± 0.52 µM and 5.06 ± 1.13 µg/ml or 13.32 ± 2.97 µM, respectively). The compound CPD20 showed the lowest values of CC₅₀ (3.79 ± 0.60 µg/ml or 9.98 ± 1.58 µM in Vero and 2.50 ± 0.75 µg/ml or 6.58 ± 1.97 µM for BSC-40 cells). In conjunction, CPD22 showed a higher Selective Index (SI) for both cell lines (2.19 in Vero and 1.77 for BSC-40). The results of cytotoxicity assays are summarized in Table 3.

TABLE 1 Minimal inhibitory concentration (MIC) and minimal bactericidal concentration (MBC) range (µM) for compounds that showed antibacterial activity in the initial screening

Compounds	MIC			MBC (MBC/MIC)		
	<i>Staphylococcus aureus</i> ATCC 29212	<i>Streptococcus pyogenes</i> ATCC 19615	<i>Enterococcus faecalis</i> ATCC 29213	<i>S. aureus</i> ATCC 29212	<i>S. pyogenes</i> ATCC 19615	<i>E. faecalis</i> ATCC 29213
CPD18	28.68	28.68	–	>28.68 (>1.0)	>28.68 (>1.0)	–
CPD20	6.58	6.58	13.16	26.32 (4.0)	6.58 (1.0)	26.32 (2.0)
CPD21	26.32	26.32	–	26.32 (1.0)	26.32 (1.0)	–
CPD22	11.97	5.99	11.97	11.97 (1.0)	23.95 (4.0)	>23.95 (>2.0)
Penicillin G	0.06	0.24	5.98	ND	ND	ND

Note. Values presented are representative of at least three independent experiments. ND: not determined.

3.6 | The putative molecular targets of CDP20–22 and binding mode proposal

In order to identify the putative molecular target for the active compounds, we apply a ligand-based similarity approach combined with inverse docking using compounds CPD20, CPD21 and CPD22 as templates, since they presented stronger antibacterial activity in previous assays. Ligand-based similarity searches

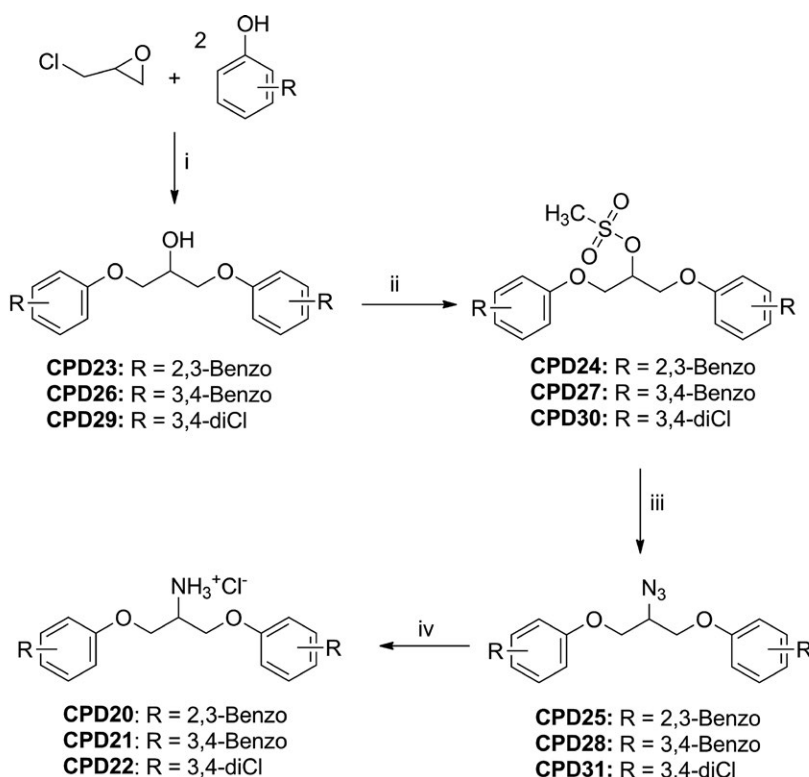
for each active compound were performed against a database of compounds with known effects against *S. aureus*, *S. pyogenes* and *E. faecalis*, resulting in 214, 15 and 30 unique compounds with at least 50% similarity against our hits respectively. Solely, *S. aureus* screening hits had an annotation for specific molecular targets, while the other two resulted in compounds with activity against whole cells or unchecked data (data not shown). The putative molecular targets for our hits in *S. aureus* with know 3D-coordinates

TABLE 2 Minimal inhibitory concentration (MIC) and minimal bactericidal concentration (MBC) range (μM) of the compounds against methicillin-resistant *Staphylococcus aureus* strains

Compounds	ATCC 43300	Strain 5749	Strain 5912	Strain 6100	Strain 6154	Strain 6613
MIC						
CPD18	–	28.68	28.68	–	28.68	28.68
CPD20	6.58	6.58	6.58	6.58	6.58	6.58
CPD21	–	26.32	13.16	26.32	26.32	26.32
CPD22	11.97	5.99	5.99	5.99	5.99	11.97
Vancomycin	1.38	1.38	1.38	1.38	1.38	1.38
MBC						
CPD18	–	>28.68	>28.68	–	>28.68	>28.68
CPD20	13.16	6.58	6.58	6.58	26.32	6.58
CPD21	–	26.32	13.16	26.32	26.32	26.32
CPD22	11.97	5.99	5.99	23.95	>23.95	11.97
Vancomycin	ND	ND	ND	ND	ND	ND

Note. Values presented are representative of at least three independent experiments.

ND: not determined.



Reagents and conditions: i) NaOH, H₂O, 60°C; ii) MsCl, pyr, 0°C-r.t.; iii) NaN₃, DMF, 80°C; iv) a) H₂/Pd-C, THF; b) conc. HCl, MeOH.

FIGURE 3 Synthesis of compounds CPD23–CPD31, synthetic intermediates of amines CPD20–CPD22

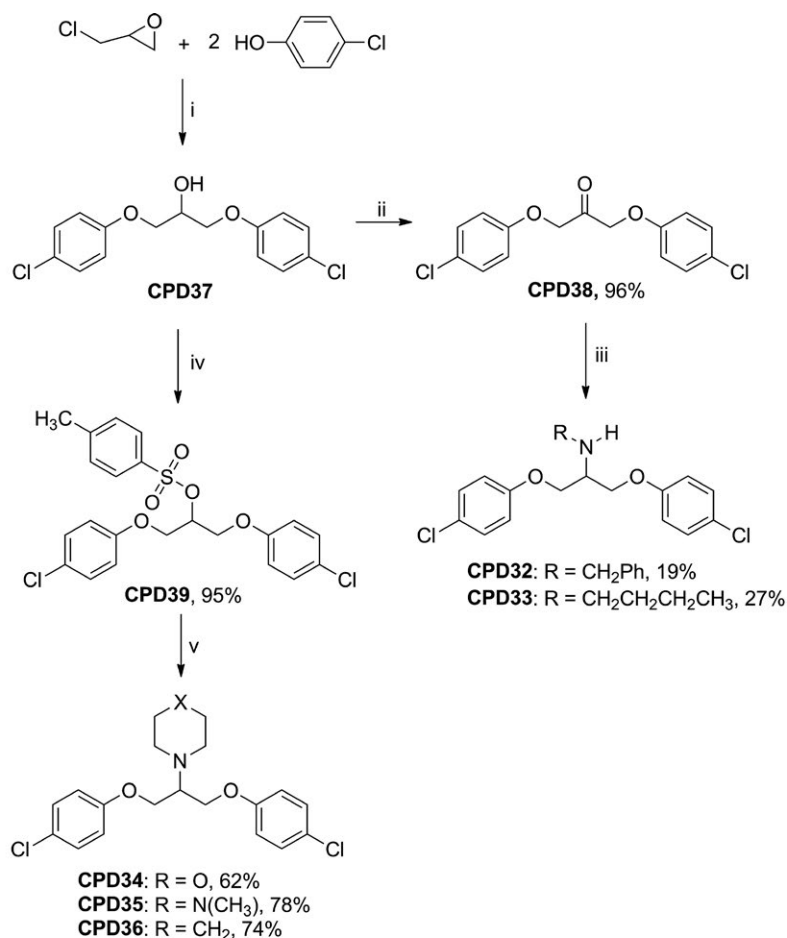


FIGURE 4 Synthesis of *N*-substituted 1,3-bis(aryloxy)propan-2-amines and yields of each synthetic step

Reagents and conditions: i) NaOH, H₂O, 60°C; ii) Ac₂O, DMSO, r.t.; iii) Primary amine, NaCNBH₃, EtOH/THF, pH 6, r.t.; iv) TsCl, pyr, r.t.; v) Heterocyclic amine, 100°C.

are namely the *ftsZ*, sortase and *FabI*, while *norA* has a known *E. coli* homolog (PDB code: 4ZPO, with 77.3% similarity) and underwent homology modeling (Table 4). The similarity search results suggest that CPD22 could have more than one molecular target.

Docking experiments were initially validated by redocking of the original ligand of each PDB file within their own active site. *SaNorA* had the ligand of its homologous structure cross-docked, both for defining the putative binding site, but also for verifying conserved interaction with important residues such as Asp34. Poses derived from the redocking procedure had their heavy-atom root mean squared deviation (RMSD) values calculated against the original co-crystallized conformation and its docking pose (Table A1).

The *SaNorA* homology model shares high structural similarity with *E. coli*, however, cross-docking performed moderately when compared to our redocking results (RMSD 1.20 Å, Figure 5a). CPD21 and CPD22 proposed interaction mode within the *SaNorA* active site (Figure 5b,c) shares hydrophobic interactions mainly with Leu62 and Leu236, but not limited to, with also a large number of hydrophobic side-chains surrounding both ring systems. Thai and collaborators by a comprehensive computational workflow have shown that *SaNorA* has a conserved large binding site within the channel offering more opportunities for binding sites than the one exploited here in this study (Thai et al., 2015).

CPD22 is suggested to have more than one target among the three-hit compounds, with two additional putative molecular targets, besides the *SaNorA*: namely *FabI* and *FstZ*. Comparison between the redocked and co-crystallized conformations of TXA6101, within the *FszT* (PDB 5XDV, Figure 6a), and AFN-1252, within *FabI* (PDB 4F53, Figure 6b), showed very low conformational differences, validating the docking method.

FszT inhibition relies on a set of hydrogen bond interactions between the side-chain of Asn163 and the main-chain atoms of Val207 and Leu209 with TXA6101 amide-moyet but also has some significant apolar contacts with Val310, Ile311, Met225, Ile362, Met219 and Ile197. CPD22 ring shares most of these hydrophobic interactions, despite lacking the hydrogen interaction network (Figure 6c). Previous attempts on the application of virtual screening towards *SaFszT* highlighted the importance of this hydrophobic complementarity with the ring systems of proposed inhibitors (Vijayalakshmi, Nisha, & Rajalakshmi, 2014).

The CPD22 resemblance with AFN-1252 goes beyond the two ring structure, extending towards the compound interactions, both presenting interaction at the Tyr157 pocket (Figure 6d, Mistry, Truong, Ghosh, Johnson, & Mehboob, 2016). However, CPD22 does not have the typical Ala97 interaction previously described in the literature as an important chemical feature for *FabI* inhibition

TABLE 3 Cytotoxic concentration of 50% (CC₅₀) and selectivity index (SI) of the active compounds in Vero and BSC-40 cells

Cell line	Compounds	CC ₅₀ ± SD	SI ^a								
			<i>Enterococcus faecalis</i>	<i>Staphylococcus aureus</i>	<i>Streptococcus pyogenes</i>	MRSA ATCC 43300	Strain 6613	Strain 5912	Strain 5749	Strain 6154	Strain 6100
Vero	CPD18	17.18 ± 0.26	–	0.6	0.6	0.6	–	0.6	0.6	0.6	–
	CPD20	9.98 ± 1.58	0.76	1.52	1.52	1.52	1.52	1.52	1.52	1.52	1.52
	CPD21	11.27 ± 2.58	–	0.43	0.43	0.43	–	0.43	0.43	0.43	0.43
	CPD22	13.1 ± 4.53	1.09	1.09	2.19	1.09	1.09	2.19	2.19	2.19	2.19
BSC-40	CPD18	16.43 ± 0.52	–	0.57	0.57	0.57	–	0.57	0.57	0.57	–
	CPD20	6.58 ± 1.97	0.5	1.0	1.0	0.25	0.25	0.25	0.25	0.5	0.25
	CPD21	13.32 ± 2.97	–	0.51	0.51	0.51	–	0.51	0.51	0.51	0.51
	CPD22	10.59 ± 3.16	0.88	0.88	1.77	0.88	0.88	1.77	1.77	1.77	1.77

Notes: Values presented are representative of at least three independent experiments.

MRSA: methicillin-resistant *S. aureus*.

^aThe SI is calculated by dividing the CC₅₀ with the minimal inhibitory concentration.

(Kronenberger et al., 2017; Mistry et al., 2016), but have a chlorine atom oriented at H-bond region.

Lastly, the cysteine transpeptidase Sortase has been proposed as a putative molecular target for the CPD20. Sortase commonly binds to flexible ligands such as signaling peptides but can also be covalently inhibited by small compound fragments. Redocking in the PDB structure 1QWZ revealed moderate capacity of prediction for this target (Table A1 and Figure 7a), 1QWZ has a large binding site when compared to other sortase structures (Jacobitz et al., 2014). The two double-ring systems of CPD20 were positioned by docking near the aromatic residues Phe114 and Tyr181, however, no pi-pi interactions could be established (Figure 7b).

Structural studies of the SaSortase B complexed with the substrate have shown a substrate-stabilized oxyanion hole involving Arg233 and Glu224 residues, which could accommodate the substrate (Jacobitz et al., 2014). Additionally, they also reported the close proximity of the ligands towards Tyr181, which could have a role in stabilizing the active conformation.

4 | DISCUSSION

The compounds tested in the present study belong to the chemical class of 1,3-bisaryloxypropan-2-amines, which have shown several biological activities in the literature and have easy access by synthesis (Heerding et al., 2003; Yuxiu et al., 2000). Our results showed that four out of 36 compounds presented antibacterial activity against the Gram-positive bacteria tested (*E. faecalis*, *S. aureus* and *S. pyogenes*). The MIC values found in this study were in the low micromolar range, varying from 2.5 to 10 µg/ml (5.99–28.58 µM). The MIC values showed in our study are equal or even higher in comparison to the antimicrobials in clinical use against *S. aureus* and *E. faecalis* strains. For example, one can cite aminoglycosides such as amikacin (1–4 and 64–256 µg/ml for *S. aureus* and *E. faecalis*, respectively) and kanamycin (1–4 and 16–64 µg/ml), some beta-lactams such as carbenicillin (2–8 and 16–64 µg/ml), piperacillin (1–4 µg/ml to both bacteria), methicillin (>16 µg/ml to *E. faecalis*) and ceftazidime (4–16 µg/ml to *S. aureus*), in addition to another important options as linezolid (1–4 µg/ml to both bacteria), chloramphenicol (2–16 and 4–16 µg/ml) and even vancomycin (0.5–2 and 1–4 µg/ml) (CLSI, 2017).

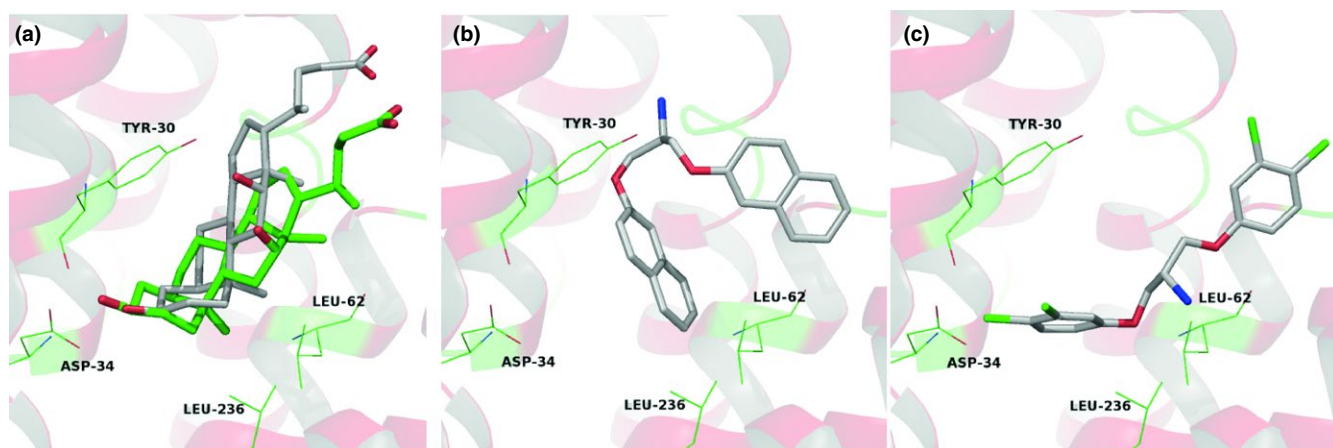
Other studies in the literature regarding the antibacterial activity of synthetic compounds have shown similar results, for example, Heerding et al. (2003) reported that an asymmetric diaryloxypropanamine showed antibacterial activity against *S. aureus*, *E. faecalis* and *S. pneumoniae* strains with MIC values of 16, 32 and 2 µg/ml, respectively (Heerding et al., 2003). Other synthetic compounds derived from pleuromutilin also had similar results, showing MIC values between 0.06 and 32 µg/ml for *S. aureus* and between 1.0 and 32 µg/ml for *E. faecalis* (Gao et al., 2017).

Moreover, the pleuromutilin derivatives mentioned above also had comparable efficacy against MRSA strains, with MIC values ranging

TABLE 4 List of the putative molecular target of CPD20-21-22

Compound ChEMBL-ID	Tanimoto combo			Related target (accession number)	Reference PMID	PDB
	CPD20	CPD21	CPD22			
CHEMBL499196	–	–	1.01	Cell division protein <i>FtsZ</i>	19064318	5XDV
CHEMBL461447	–	–	1.00	POA031	19064318	
CHEMBL1097797	–	–	1.02		20426423	
CHEMBL3098795	–	–	1.07		24287381	
CHEMBL3098796	–	–	1.05		24287381	
CHEMBL3417347	–	1.05	1.22	Quinolone resistance protein <i>norA</i> POA0J7	25817769	Homology model
CHEMBL372191	1.01	–	–	Sortase Q9S446	1615474, 19269184	1QWZ
CHEMBL3623431	–	–	1.00	Enoyl-[acyl-carrier-protein] reductase [NADPH] <i>FabI</i> Q9RMI3	26343826	4FS3

Note. For each compound screened and hit combination the similarity index represented by the Tanimoto combo is presented. Protein accession number refers to the amino-acid sequence code at UniProt database and the Protein Data Bank (PDB) codes refer to the tridimensional structure deposited in the PDB.

**FIGURE 5** Cross-docking of deoxycholate from *EcNorA* (Protein Data Bank code: 4ZP0) ligand within the conserved binding cavity (a). Putative binding mode of CPD21 (b) and CPD22 (c) into the binding site of *SaNorA*

from 0.015 to over 32 $\mu\text{g/ml}$. Other studies have also shown antibacterial activity of synthetic molecules against MRSA, such as synthetic biphenylthiazoles, which presented MIC values ranging from 0.39 to 25 $\mu\text{g/ml}$ against three different MRSA strains (Hagras et al., 2017). In fact, equivalent MIC values were obtained in our study when compounds were tested against MRSA strains ranging from 2.5 to 10 $\mu\text{g/ml}$ (5.99 to 28.58 μM). Those values are also comparable to some antimicrobials used in the clinic for MRSA, such as daptomycin (0.5 $\mu\text{g/ml}$), vancomycin (2 $\mu\text{g/ml}$), oxacillin (16 $\mu\text{g/ml}$) and gentamicin (128 $\mu\text{g/ml}$) (Baltch, Ritz, Bopp, Michelsen, & Smith, 2007).

Regarding the MBC assay, results ranging from 2.5 to 10 $\mu\text{g/ml}$ (5.99 to 26.32 μM) were in the same range of MIC values suggesting the bactericidal mode of action of these compounds. This characteristic is desirable for an antibacterial drug since it is often associated with the capability of inhibiting and preventing bacterial dissemination (Alder & Eisenstein, 2004). For example, synthetic

biphenylthiazoles have presented MBC values ranging from 8 to 32 $\mu\text{g/ml}$ against *S. aureus* strains.

In the initial screening for antibacterial activity at the concentration of 10 $\mu\text{g/ml}$, only four compounds showed Gram-positive antibacterial activity. The *p*-chloro-substituted CPD18 was the only monosubstituted aromatic compound, with an MIC of 10 $\mu\text{g/ml}$ (28.68 μM) against *S. aureus*, *S. pyogenes* and several MRSA strains. The introduction of a second chloro-substituent at position 3 on aromatic rings, as observed in CPD22, a 3,4-dichloro-substituted aromatic compound, potentiate the activity against Gram-positive bacteria, including *E. faecalis*, reducing the MIC values to 2.5 to 5 $\mu\text{g/ml}$ (5.99–11.97 μM). Although not essential, the disubstitution pattern seems to be important for antibacterial activity, since three of the four active compounds, namely CPD20, CPD21 and CPD22, have aromatic rings substituted at two positions. The 1-naphthoxy derivative (CPD20) presents an antibacterial activity against all Gram-positive bacteria evaluated, including MRSA strains with

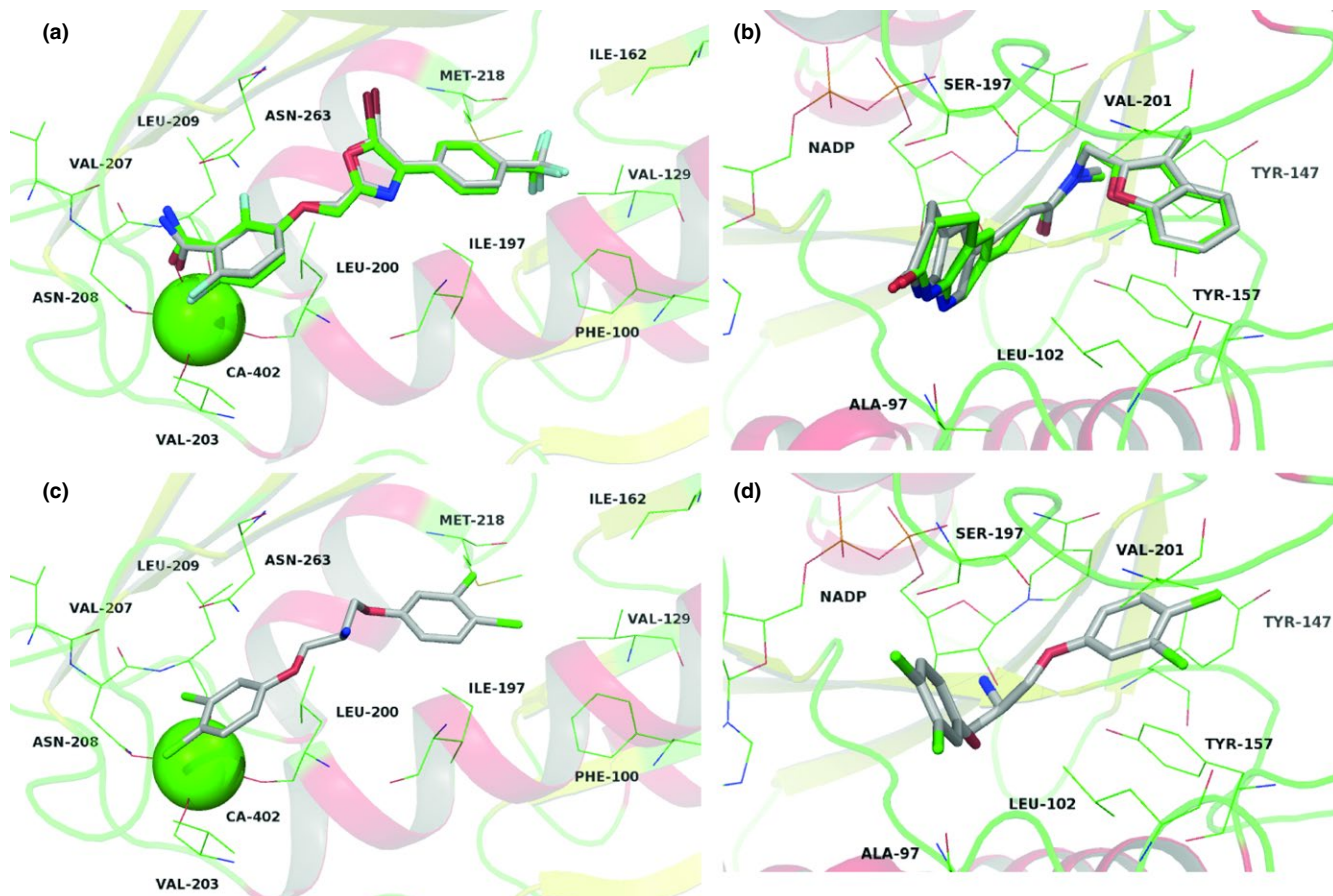


FIGURE 6 Redocking of original ligands TXA6101 from SaFszT (a, Protein Data Bank [PDB] code: 5XDV) and AFN-1252 into SaFabl (b, PDB: 4FS3) into their respective described active site. Suggested binding mode by docking of CPD22 in the binding site of FszT (c) and Fabl (d)

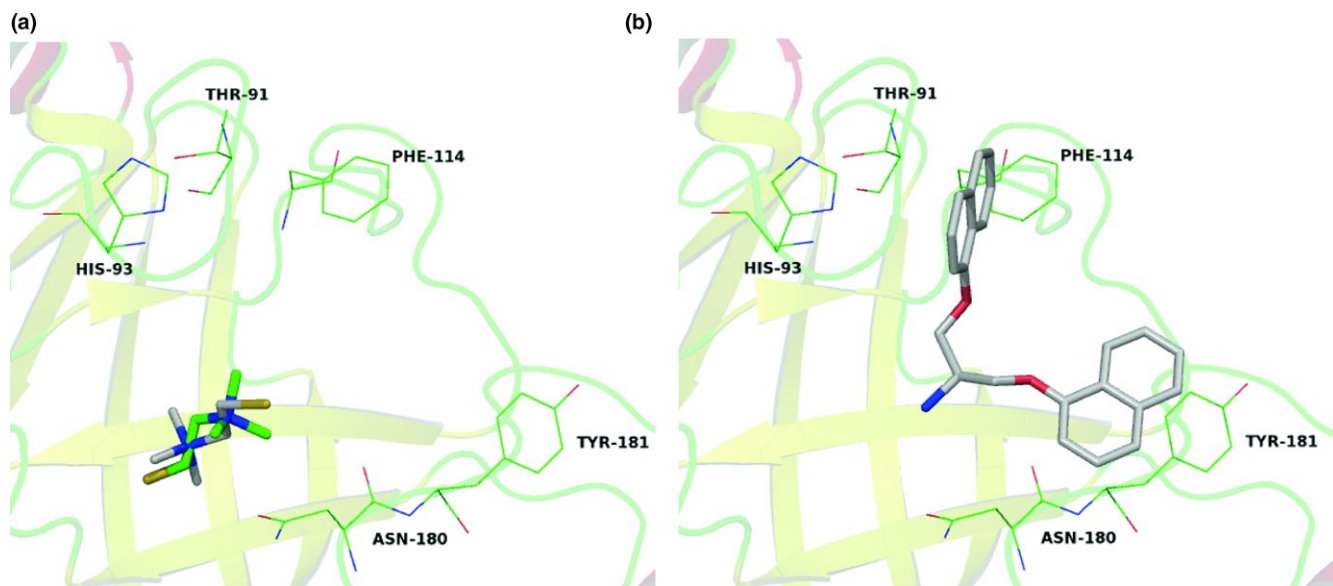


FIGURE 7 Redocking of original ligands 2-(trimethylammonium)ethyl thiol from SaSortase (a, Protein Data Bank code: 1QWZ) and docking of CPD20 within the active site (b)

MIC values ranging from 2.5 to 5 $\mu\text{g/ml}$ (6.58–13.16 μM), similar to the values determined for CPD22. The 2-naphthoxy derivative (CPD21), with MIC values of 10 $\mu\text{g/ml}$ (26.32 μM), demonstrated

antibacterial activity against *S. pyogenes*, *S. aureus* and MRSA strains similar to compound CPD18, however its activity against *E. faecalis* was not observed.

The synthetic intermediates of active compounds CPD20, CPD21 and CPD22 were also evaluated in order to get some insights about the role of an amino group in the aliphatic chain to the antibacterial activity of these compounds. The substitution of the amino group by a hydroxyl, a mesyl or an azide group leads to loss of activity (CPD23–31), indicating that this group is essential for the antibacterial activity. Since we recognize the importance of an amino group placed in the aliphatic chain, we also verified if the substitution of the amino group would interfere with their activity, synthesizing and evaluating secondary and tertiary amines derived from CPD18. The results indicated a loss of activity when the amino group is substituted (CPD32–36), suggesting that these compounds need to be a primary amine to promote antibacterial effects.

Among the evaluated primary amines, CPD18, CPD20, CPD21 and CPD22 are the ones with the highest calculated partition coefficient (ClogP) values (Table A2). The ClogP is a measure of lipophilic character of a compound which increases as the lipophilicity of the compound increases. Thus, this indicates that the highest lipophilicity of these three compounds contributes to a better Gram-positive antibacterial activity. This relationship is similar to that observed to the antibacterial activity of cephalosporins and penicillins against *S. aureus* (Biagi, Guerra, Barbaro, & Gamba, 1970). This trend could indicate that the activity of primary amines to Gram-positive bacteria could be related to the bacterial membrane permeability in which substituent as chloro, dichloro and naphthyl could improve.

Although in most cases the highest lipophilic character may be a limiting factor to the use of these compounds in oral or intravenous formulas, they could favor its topical use, like in ointments and emulsions due to the direct absorption of lipophilic substances. This route of administration may be beneficial for the treatment of surgical site infections, decubitus ulcers, mainly caused by species of *Enterococcus*, *S.* and *Staphylococcus*, especially MRSA strains.

The cytotoxic concentration of 50% (CC₅₀) of active compounds in mammalian cells was generally also in the low micromolar range, corroborating the results obtained by Lavorato et al. (2017). Hence, the selectivity index (SI) values obtained ranged from 0.25 to 2.19. These low SI values can be improved by changing critical chemical groups in the molecule, which can reduce its toxicity or enhance its activity.

Here, experimental validation demonstrated the ability of CPD20, CPD21 and CPD22 to interfere with the growth of *S. aureus*, *S. pyogenes* and *E. faecalis*. However, the molecular target of these drug candidates remains undetermined. In order to identify putative protein targets, we have employed a combination of chemical similarity search with inverse docking approaches. The three-dimensional chemical similarity between our hits and compounds with known activity against the organisms of interest was used to select a set of compounds with known biological targets. The prediction of the binding mode suggests that the compounds can interact with same pockets/regions of known cocrystallized inhibitors, which indicate the possibility of CPD22 to be a multi-target antibacterial.

Furthermore, for the suggested targets, both 3D chemical similarity and parallels in terms the protein–ligand interactions between our compounds with known inhibitors supports this binding mode and encourages further in vitro testing. For instance, the dichloro–benzene groups of CPD22 interacting with the *FabI* hydrophobic pocket (Figure 6d).

Taken together, our data show that some compounds belong to the class of symmetric 1,3-bis(aryloxy)propan-2-amine tested in the present study showed a relevant antibacterial activity against important Gram-positive pathogens (including antibiotic-resistant strains), with minimal inhibitory and bactericidal concentration in the low micromolar range. As a perspective, we intend to investigate the activity of this class of amines against other clinically relevant resistant bacteria, such as *E. faecium*, vancomycin-resistant enterococci and glycopeptide-intermediate *S. aureus*. Through an in silico approach, we identified three putative molecular targets for these compounds and we hope that these data may contribute, in the long-term, to lead these compounds for further optimization towards selectivity, aiming to treat bacterial infections, including those caused by resistant Gram-positive bacteria.

ACKNOWLEDGEMENT

This work was supported by the CNPq, CAPES, FAPEMIG and UFMG intramural funds.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

AUTHORS CONTRIBUTION

M.S.M.S. performed the experiments, analyzed the results and wrote the manuscript; S.N.L. synthesized the compounds and wrote the manuscript; T.K. performed the in silico experiments and wrote the manuscript; Y.V.S. performed the experiments; G.P.O. performed the experiments to evaluate the cytotoxicity of the compounds; S.G.S. provided the clinical MRSA strains; E.G.K. analyzed the results and wrote the manuscript; V.G.M. performed the in silico experiments, analyzed the results and wrote the manuscript; R.J.A. synthesized the compounds, analyzed the results and wrote the manuscript; B.E.F.M. coordinated the Project, design experiments, analyzed the results and wrote the manuscript.

ETHICS STATEMENT

None required.

DATA ACCESSIBILITY

The data that support the findings of this study are available from the corresponding author upon request.

REFERENCES

- Alder, J., & Eisenstein, B. (2004). The advantage of bactericidal drugs in the treatment of infection. *Current Infectious Disease Reports*, 6, 251–253.
- Baltch, A. L., Ritz, W. J., Bopp, L. H., Michelsen, P. B., & Smith, R. P. (2007). Antimicrobial activities of daptomycin, vancomycin, and oxacillin in human monocytes and of daptomycin in combination with gentamicin and/or rifampin in human monocytes and in broth against *Staphylococcus aureus*. *Antimicrobial Agents and Chemotherapy*, 51, 1559–1562. <https://doi.org/10.1128/AAC.00973-06>
- Bento, A. P., Gaulton, A., Hersey, A., Bellis, L. J., Chambers, J., Davies, M., ... Nowotka, M. (2014). The ChEMBL bioactivity database: An update. *Nucleic Acids Research*, 42, D1083–D1090. <https://doi.org/10.1093/nar/gkt1031>
- Biagi, G. L., Guerra, M. C., Barbaro, A. M., & Gamba, M. F. (1970). Influence of lipophilic character on the antibacterial activity of cephalosporins and penicillins. *Journal of Medicinal Chemistry*, 13, 511–516.
- Borch, R. F., Bernstein, M. D., & Durst, H. D. (1971). Cyanohydrinborate anion as a selective reducing agent. *Journal of the American Chemical Society*, 93, 2897–2904. <https://doi.org/10.1021/ja00741a013>
- Boucher, H. W., Talbot, G. H., Bradley, J. S., Edwards, J. E., Gilbert, D., Rice, L. B., ... Bartlett, J. (2009). Bad bugs, no drugs: No ESKAPE! An update from the Infectious Diseases Society of America. *Clinical Infectious Diseases*, 48, 1–12. <https://doi.org/10.1086/595011>
- CDC. (2013). *Antibiotic resistance threats in the United States, 2013*. Atlanta, GA: CDC.
- Clatworthy, A. E., Pierson, E., & Hung, D. T. (2007). Targeting virulence: A new paradigm for antimicrobial therapy. *Nature Chemical Biology*, 3, 541–548. <https://doi.org/10.1038/nchembio.2007.24>
- Clinical and Laboratory Standards Institute. (1999). *Methods for determining bactericidal activity of antimicrobial agents; approved guideline M26-A* (Vol. 19). Villanova, PA: NCCLS.
- CLSI. (2017). *Performance standards for antimicrobial susceptibility testing* (27th ed.). CLSI supplement M100. Wayne, PA: Clinical and Laboratory Standards Institute.
- Davis, I. W., Murray, L. W., Richardson, J. S., & Richardson, D. C. (2004). MOLPROBITY: Structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Research*, 32, W615–W619. <https://doi.org/10.1093/nar/gkh398>
- Eswar, N., Webb, B., Marti-Renom, M. A., Madhusudhan, M. S., Eramian, D., Shen, M. Y., ... Sali, A. (2006). Comparative protein structure modeling with modeller. *Current Protocols in Bioinformatics*, 15, 5–6.
- Friesner, R. A., Murphy, R. B., Repasky, M. P., Frye, L. L., Greenwood, J. R., Halgren, T. A., ... Mainz, D. T. (2006). Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *Journal of Medicinal Chemistry*, 49, 6177–6196. <https://doi.org/10.1021/jm051256o>
- Fritsche, A., Elfringhoff, A. S., Fabian, J., & Lehr, M. (2008). 1-(2-Carboxyindol-5-yloxy)propan-2-ones as inhibitors of human cytosolic phospholipase A2alpha: Synthesis, biological activity, metabolic stability, and solubility. *Bioorganic & Medicinal Chemistry*, 16, 3489–3500. <https://doi.org/10.1016/j.bmc.2008.02.019>
- Gao, M.-L., Zeng, J., Fang, X., Luo, J., Jin, Z., Liu, Y.-H., & Tang, Y.-Z. (2017). Design, synthesis and antibacterial evaluation of novel pleuromutilin derivatives possessing piperazine linker. *European Journal of Medicinal Chemistry*, 127, 286–295. <https://doi.org/10.1016/j.ejmech.2017.01.004>
- Gomes, R. M., Bomfim, M. R. Q., Trindade, M. J., Farias, L. M., Carvalho, M. A. R., Serufo, J. C., & Santos, S. G. (2015). Potential spread of methicillin-resistant *Staphylococcus aureus* recovered from patients with bloodstream infection. *Chemother Open Access*, 4, 1–8. <https://doi.org/10.4172/2167-7700.1000149>
- Hagras, M., Mohammad, H., Mandour, M. S., Hegazy, Y. A., Ghiaty, A., Seleem, M. N., & Mayhoub, A. S. (2017). Investigating the antibacterial activity of biphenylthiazoles against methicillin- and vancomycin-resistant *Staphylococcus aureus* (MRSA and VRSA). *Journal of Medicinal Chemistry*, 60, 4074–4085. <https://doi.org/10.1021/acs.jmedchem.7b00392>
- Hawkins, P. C. D., Skillman, A. G., & Nicholls, A. (2007). Comparison of shape-matching and docking as virtual screening tools. *Journal of Medicinal Chemistry*, 50, 74–82. <https://doi.org/10.1021/jm0603365>
- Hawkins, P. C. D., Skillman, A. G., Warren, G. L., Ellingson, B. A., & Stahl, M. T. (2010). OMEGA 2.5.1.4. Santa Fe, NM: OpenEye Scientific Software.
- Heerding, D. A., Christmann, L. T., Clark, T. J., Holmes, D. J., Rittenhouse, S. F., Takata, D. T., & Venslavsky, J. W. (2003). New benzylidene-thiazolidinediones as antibacterial agents. *Bioorganic & Medicinal Chemistry Letters*, 13, 3771–3773.
- Jacobitz, A. W., Wereszczynski, J., Yi, S. W., Amer, B. R., Huang, G. L., Nguyen, A. V., ... Clubb, R. T. (2014). Structural and computational studies of the *Staphylococcus aureus* sortase B-substrate complex reveal a substrate-stabilized oxyanion hole. *Journal of Biological Chemistry*, 289, 8891–8902. <https://doi.org/10.1074/jbc.M113.509273>
- King, L. C., & Bigelow, M. J. (1952). The structure and configuration of cholesterylpyridinium salts. *Journal of the American Chemical Society*, 74, 3338–3339. <https://doi.org/10.1021/ja01133a036>
- Kronenberger, T., Asse, L. R., Wrenger, C., Trossini, G. H. G., Honorio, K. M., & Maltarollo, V. G. (2017). Studies of *Staphylococcus aureus* FabI inhibitors: Fragment-based approach based on holographic structure-activity relationship analyses. *Future Medicinal Chemistry*, 9, 135–151. <https://doi.org/10.4155/fmc-2016-0179>
- Lavorato, S. N., Duarte, M. C., Lage, D. P., Tavares, C. A. P., Coelho, E. A. F., & Alves, R. J. (2017). Synthesis and antileishmanial activity of 1,3-bis(aryloxy)propan-2-amines. *Medicinal Chemistry Research*, 26, 1052–1072.
- Lavorato, S. N., Sales Júnior, P. A., Murta, S. M. F., Romanha, A. J., & Alves, R. J. (2015). In vitro activity of 1,3-bisaryloxypropanamines against *Trypanosoma cruzi*-infected L929 cultures. *Memorias do Instituto Oswaldo Cruz*, 110, 566–568. <https://doi.org/10.1590/0074-02760150007>
- Li, J., Nation, R. L., Turnidge, J. D., Milne, R. W., Coulthard, K., Rayner, C. R., & Paterson, D. L. (2006). Colistin: The re-emerging antibiotic for multidrug-resistant Gram-negative bacterial infections. *The Lancet Infectious Diseases*, 6, 589–601. [https://doi.org/10.1016/S1473-3099\(06\)70580-1](https://doi.org/10.1016/S1473-3099(06)70580-1)
- Lobanovska, M., & Pilla, G. (2017). Penicillin's discovery and antibiotic resistance: Lessons for the future? *Yale Journal of Biology and Medicine*, 90, 135–145.
- Malhotra-Kumar, S., Xavier, B. B., Das, A. J., Lammens, C., Butaye, P., & Goossens, H. (2016). Colistin resistance gene *mcr-1* harboured on a multidrug resistant plasmid. *The Lancet Infectious Diseases*, 16, 283–284. [https://doi.org/10.1016/S1473-3099\(16\)00012-8](https://doi.org/10.1016/S1473-3099(16)00012-8)
- Mistry, T. L., Truong, L., Ghosh, A. K., Johnson, M. E., & Mehboob, S. (2016). Benzimidazole-based FabI inhibitors: A promising novel scaffold for anti-staphylococcal drug development. *ACS Infectious Diseases*, 3, 54–61. <https://doi.org/10.1021/acsinfecdis.6b00123>
- Mosmann, T. (1983). Rapid colorimetric assay for cellular growth and survival: Application to proliferation and cytotoxicity assays. *Journal of Immunological Methods*, 65, 55–63.
- OpenEye Scientific Software. (2016). QUACPAC 2016. Santa Fe, NM: Author.
- Rush, T. S., Grant, J. A., Mosyak, L., & Nicholls, A. (2005). A shape-based 3-D scaffold hopping method and its application to a bacterial protein-protein interaction. *Journal of Medicinal Chemistry*, 48, 1489–1495. <https://doi.org/10.1021/jm040163o>

- Silver, L. L. (2011). Challenges of antibacterial discovery. *Clinical Microbiology Reviews*, 24, 71–109. <https://doi.org/10.1128/CMR.00030-10>
- Söding, J., Biegert, A., & Lupas, A. N. (2005). The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Research*, 33, W244–W248. <https://doi.org/10.1093/nar/gki408>
- Thai, K. M., Ngo, T. D., Phan, T. V., Tran, T. D., Nguyen, N. V., Nguyen, T. H., & Le, M. T. (2015). Virtual screening for novel *Staphylococcus aureus* nora efflux pump inhibitors from natural products. *Medicinal Chemistry*, 11, 135–155. <https://doi.org/10.2174/1573406410666140902110903>
- Vallone, A., D'Alessandro, S., Brogi, S., Brindisi, M., Chemi, G., Alfano, G., ... Decherig, K. J. (2018). Antimalarial agents against both sexual and asexual parasites stages: Structure-activity relationships and biological studies of the Malaria Box compound 1-[5-(4-bromo-2-chlorophenyl)furan-2-yl]-N-[(piperidin-4-yl)methyl]methanamine (MMV019918) and analogues. *European Journal of Medicinal Chemistry*, 150, 698–718. <https://doi.org/10.1016/j.ejmech.2018.03.024>
- Vijayalakshmi, P., Nisha, J., & Rajalakshmi, M. (2014). Virtual screening of potential inhibitor against FtsZ protein from *Staphylococcus aureus*. *Interdisciplinary Sciences: Computational Life Sciences*, 6, 331–339. <https://doi.org/10.1007/s12539-012-0229-3>
- Yuxiu, L., Guiqin, H., & Guangren, L. (2000). Synthesis and bioactivities of 1,3-diaryloxy-2-propylamines. *Chinese Journal of Synthetic Chemistry*, 8, 511–515.

How to cite this article: Serafim MSM, Lavorato SN, Kronenberger T, et al. Antibacterial activity of synthetic 1,3-bis(aryloxy)propan-2-amines against Gram-positive bacteria. *MicrobiologyOpen*. 2019;e814. <https://doi.org/10.1002/mbo3.814>

APPENDIX

SYNTHESIS AND CHARACTERIZATION DATA OF COMPOUNDS CPD22, CPD29–CPD36, CPD38 AND CPD39

All reagents were obtained from commercial suppliers and used without further purification, unless stated otherwise. IR spectra were obtained using a Spectrum One, Perkin-Elmer ATR system. Melting points were determined on Microquímica MQAPF 301 apparatus. ^1H and ^{13}C NMR spectra were obtained on a Bruker Avance DPX-200 spectrometer and proton and carbon chemical shifts (δ) are reported with respect to Tetramethylsilane (TMS). Column chromatography was carried out on silica gel 60 0.063–0.200 mm/70–230 mesh Merck. Pyridine was dried over KOH pellets. DMSO was dried over 3\AA molecular sieves. Anhydrous DMF was obtained from commercial supplier (Sigma-Aldrich).

1,3-BIS(3,4-DICHLOROPHENOXY)PROPAN-2-OL (CPD29)

Epichlorohydrin (1 ml; 12.8 mmol) was added dropwise to a stirred solution of 3,4-dichlorophenol (38 mmol) and sodium hydroxide (38 mmol) in water (30 ml). The reaction mixture was stirred under heating (60°C) for 14 hr. After this time, the reaction mixture was cooled to room temperature and the product was isolated as a white solid using vacuum filtration in 88% yield; mp: 103.4–105.7°C. IR (ATR) ν_{max} 3,527, 2,948,

1,589, 1,566, 1,478, 1,465, 1,455, 1,227, 1,064, 1,121/cm; ^1H NMR (CDCl_3 , 200 MHz): δ = 7.33 (2H, d, J = 9.0 Hz, H-5'), 7.03 (2H, d, J = 2.6 Hz, H-2'), 6.78 (2H, dd, J = 9.0 Hz, 2.6 Hz, H-6'), 4.37 (1H, s, H-2), 4.11 (4H, d, J = 4.8 Hz, H-1), 2.59 (1H, s, OH); ^{13}C NMR (CDCl_3 , 50 MHz): δ = 157.5 (C, C-1'), 133.2 (C, C-3'), 131.0 (CH, C-5'), 125.0 (C, C-4'), 116.7 (CH, C-2'), 114.7 (CH, C-6'), 69.3 (CH_2 , C-1), 68.6 (CH, C-2).

1,3-BIS(3,4-DICHLOROPHENOXY)PROPAN-2-YL METHANESULFONATE (CPD30)

To a stirred and ice-cooled solution of CPD29 (1.57 mmol) in pyridine (3 ml), methanesulfonyl chloride (6.3 mmol; 0.49 ml) was added dropwise. Crushed ice was added to the flask after 4 hr of reaction, followed by the addition of concentrated HCl until pH 1. The product was extracted with dichloromethane (3×30 ml). The organic layers were combined, washed with water (5×50 ml), dried over anhydrous Na_2SO_4 , filtrated and concentrated under reduced pressure. A white solid was obtained in 83% yield; mp: 83.7–84.4°C. IR (ATR) ν_{max} 3,042, 2,940, 2,883, 1,593, 1,566, 1,478, 1,452, 1,345, 1,231, 1,176, 1,057/cm; ^1H NMR (CDCl_3 , 200 MHz): δ = 7.35 (2H, d, J = 8.8 Hz, H-5'), 7.03 (2H, d, J = 3.0 Hz, H-2'), 6.78 (2H, dd, J = 8.8 Hz, 3.0 Hz, H-6'), 5.21 (1H, qn, J = 5.2 Hz, H-2), 4.29 (4H, d, J = 5.2 Hz, H-1), 3.15 (3H, s, OSO_2CH_3); ^{13}C NMR (CDCl_3 , 50 MHz): δ = 156.9 (C, C-1'), 133.4 (C, C-3'), 131.2 (CH, C-5'), 125.5 (C, C-4'), 116.8 (CH, C-2'), 114.6 (CH, C-6'), 77.1 (CH, C-2), 67.4 (CH_2 , C-1).

1,3-BIS(3,4-DICHLOROPHENOXY)PROPAN-2-YL AZIDE (CPD31)

Sodium azide (10.9 mmol) was added to a stirred solution of CPD30 (1.09 mmol) in DMF (3 ml) at 80°C . Crushed ice was added to the flask after 24 hr of reaction. The product was extracted with dichloromethane (3×30 ml) and the organic layer was washed with water (5×50 ml). The organic layers were combined, dried over anhydrous Na_2SO_4 , filtrated and concentrated under reduced pressure. A white solid was obtained after recrystallization from isopropyl alcohol in 67% yield; mp: 95.6–96.3°C. IR (ATR) ν_{max} 2,940, 2,882, 2,144, 2,094, 1,590, 1,567, 1,478, 1,454, 1,231, 1,023/cm; ^1H NMR (CDCl_3 , 200 MHz): δ = 7.34 (2H, d, J = 8.8 Hz, H-5'), 7.03 (2H, d, J = 2.8 Hz, H-2'), 6.79 (2H, ddd, J = 8.8 Hz, 2.8 Hz, 0.4 Hz, H-6'), 4.20–4.11 (5H, m, H-1, H-2); ^{13}C NMR (CDCl_3 , 50 MHz): δ = 157.2 (C, C-1'), 133.3 (C, C-3'), 131.1 (CH, C-5'), 125.3 (C, C-4'), 116.7 (CH, C-2'), 114.7 (CH, C-6'), 68.0 (CH_2 , C-1), 59.3 (CH, C-2).

1,3-BIS(3,4-DICHLOROPHENOXY)PROPAN-2-AMINIUM CHLORIDE (CPD22)

10% Palladium on activated carbon (20 mg) was added to a solution of CPD31 (0.47 mmol) in Tetrahydrofuran (THF) (10 ml). The reaction was kept under stirring and hydrogen atmosphere for 4 hr. Then, the catalyst was removed using filtration and the filtrate was concentrated. The residue was reconstituted in methanol and concentrated hydrochloric acid was added dropwise until a slight precipitate

is formed. The solvent was evaporated to give a white solid in 41% yield; mp: 224.1–225.6°C. IR (ATR) ν_{\max} 2,882, 1,589, 1,569, 1,511, 1,478, 1,460, 1,231, 1,049/cm; ^1H NMR (DMSO- d_6 , 200 MHz): δ = 8.79 (3H, s, NH_3^+Cl^-), 7.56 (2H, d, J = 8.8 Hz, H-5'), 7.32 (2H, d, J = 2.8 Hz, H-2'), 7.05 (2H, dd, J = 8.8 Hz, 2.8 Hz, H-6'), 3.90 (1H, dd, J = 5.2 Hz, J = 4.6 Hz, H-2), 4.39 (2H, dd, J = 10.5 Hz, J = 4.6 Hz, H-1a), 4.31 (2H, dd, J = 10.5 Hz, J = 5.2 Hz, H-1b); ^{13}C NMR (DMSO- d_6 , 50 MHz): δ = 157.5 (C, C-1'), 131.6 (C, C-3'), 131.1 (CH, C-5'), 123.3 (C, C-4'), 116.9 (CH, C-2'), 115.8 (CH, C-6'), 65.8 (CH_2 , C-1), 49.0 (CH, C-2).

SYNTHESIS OF 1,3-BIS(4-CHLOROPHENOXY)PROPAN-2-ONE (CPD38)

Acetic anhydride (17.67 mmol) was added to anhydrous dimethyl sulfoxide (9 ml; 127 mmol) and the solution was kept under nitrogen atmosphere for 10 min. This solution was then added dropwise to a solution of CPD37 (3.19 mmol). The reaction mixture was stirred for 24 hr. Thereafter, a 1:1 mixture of saturated solutions of NaHCO_3 and NaCl was added to the flask and a white solid was isolated using vacuum filtration in 96% yield; mp: 84.4–87.9°C. IR (ATR) ν_{\max} 3,098, 3,063, 3,045, 2,958, 2,924, 2,890, 1,731, 1,594, 1,585, 1,487, 1,243, 1,223, 1,062, 815/cm; ^1H NMR (acetone- d_6 , 200 MHz): δ = 7.31 (4H, d, J = 9.0 Hz, H-3', H-5'), 7.00 (4H, d, J = 9.0 Hz, H-2', H-6'), 5.05 (4H, s, H-1); ^{13}C NMR (acetone- d_6 , 50 MHz): δ = 201.6 (C, C-2), 158.0 (C, C-1'), 130.2 (CH, C-3', C-5'), 126.7 (C, C-4'), 117.2 (CH, C-2', C-6'), 72.1 (CH_2 , C-1).

SYNTHESIS OF 1,3-BIS(4-CHLOROPHENOXY)PROPAN-2-YL 4-METHYLBENZENESULFONATE (CPD39)

4-Toluenesulfonyl chloride (5.76 mmol) was added to a solution of CPD37 (1.92 mmol) in pyridine (3 ml), under stirring in ice-bath. After 2 hr, crushed ice was added to the flask and then concentrated hydrochloric acid was added until pH 1, leading to the product precipitation. A white solid was isolated by vacuum filtration in 95% yield; mp: 145.3–147.2°C. IR (ATR) ν_{\max} 3,094, 2,930, 1,598, 1,581, 1,491, 1,458, 1,349, 1,171, 1,244, 1,041, 827/cm; ^1H NMR (CDCl_3 , 200 MHz): δ = 7.80 (4H, d, J = 8.0 Hz, H-2'', H-6''), 7.30 (4H, d, J = 8.0 Hz, H-3'', H-5''), 7.20 (2H, d, J = 8.8 Hz, H-3', H-5'), 6.69 (2H, d, J = 8.8 Hz, H-2', H-6'), 5.01 (1H, t, J = 5.0 Hz, H-2), 4.21 (4H, d, J = 5.0 Hz, H-1), 2.44 (1H, s, $\text{OSO}_2\text{PhCH}_3$); ^{13}C NMR (CDCl_3 , 50 MHz): δ = 156.6 (C, C-1'), 145.3 (C, C-1''), 130.0 (CH, C-3'', C-5''), 129.6 (C, CH, C-3', C-5', C-4''), 128.3 (CH, C-2'', C-6''), 126.7 (CH, C-4'), 116.0 (CH, C-2', C-6'), 77.3 (CH, C-2), 66.7 (CH_2 , C-1), 21.9 (CH_3 , $\text{OSO}_2\text{PhCH}_3$).

GENERAL PROCEDURE 1 FOR SYNTHESIS OF N-SUBSTITUTED AMINES

To a solution of the appropriate primary amine in THF and absolute ethyl alcohol (30 ml, 1:1) was added a methanolic solution of hydrochloric acid 5 M until pH 6. Then, CPD38 and NaCNBH_3 were added. The reaction was stirred under room temperature for 72 hr. The reaction was quenched with water (1 ml). The solvent was then removed under

reduced pressure, and the residue was reconstituted in dichloromethane (30 ml) and washed with 0.5 M aqueous hydrochloric acid and water (3 × 30 ml). The organic layers were combined, dried over anhydrous Na_2SO_4 , filtrated and concentrated under reduced pressure.

GENERAL PROCEDURE 2 FOR SYNTHESIS OF N-SUBSTITUTED AMINES

A mixture of CPD39 and the appropriated secondary amine was heated to 100°C for 24 hr without stirring. Then, ethyl acetate (30 ml) was added to the flask and the resulting mixture was transferred to a separatory funnel. This mixture was washed with water (3 × 50 ml). The organic layers were combined, dried over anhydrous Na_2SO_4 , filtrated and concentrated under reduced pressure.

4,4'-((2-(BENZYL OXY)PROPANE-1,3-DIYL)BIS(OXY))BIS(CHLOROBENZENE) (CPD32)

From benzylamine (3.6 mmol), CPD38 (0.6 mmol) and NaCNBH_3 (1.2 mmol) and using the general procedure 1, a colorless oil was isolated using silica gel column chromatography (Hexanes/EtOAc 9:1 v/v) in 19% yield; IR (ATR) ν_{\max} 3,336, 3,063, 3,028, 2,925, 2,875, 2,849, 1,595, 1,581, 1,489, 1,462, 1,235, 1,028, 819/cm; ^1H NMR (CDCl_3 , 200 MHz): δ = 7.37 (5H, s, H-5, H-6, H-7, H-8, H-9), 7.25 (4H, d, J = 6.0 Hz, H-3', H-5'), 6.83 (4H, d, J = 6.0 Hz, H-2', H-6'), 4.10 (4H, s, H-1), 3.98 (2H, s, H-3), 3.40 (1H, s, H-2), 2.04 (1H, s, NH); ^{13}C NMR (CDCl_3 , 50 MHz): δ = 157.6 (C, C-1'), 140.3 (C, C-4), 129.6 (CH, C-3', C-5'), 128.7 (CH, C-6, C-8), 128.3 (CH, C-5, C-9), 127.4 (CH, C-7), 126.2 (C, C-4'), 116.1 (CH, C-2', C-6'), 68.1 (CH_2 , C-1), 55.9 (CH_2 , C-1), 68.6 (CH, C-2).

N-(1,3-BIS(4-CHLOROPHENOXY)PROPAN-2-YL)BUTAN-1-AMINE (CPD33)

From butylamine (1.92 mmol), CPD38 (0.32 mmol) and NaCNBH_3 (0.64 mmol) and using the general procedure 1, a colorless oil was isolated using silica gel column chromatography (Hexanes/EtOAc 9:1 v/v) in 27% yield; IR (ATR) ν_{\max} 2,956, 2,927, 2,871, 1,595, 1,581, 1,489, 1,463, 1,235, 1,031, 1,005, 819/cm; ^1H NMR (CDCl_3 , 200 MHz): δ = 7.22 (4H, d, J = 9.0 Hz, H-3', H-5'), 6.84 (4H, d, J = 9.0 Hz, H-2', H-6'), 4.06 (4H, d, J = 5.6 Hz, H-1), 3.32 (1H, qn, J = 5.6 Hz, H-2), 2.74 (2H, t, J = 6.6 Hz, H-3), 1.96 (1H, s, NH), 1.59–1.26 (4H, m, H-4, H-5), 0.92 (3H, t, J = 7.2 Hz, H-6); ^{13}C NMR (CDCl_3 , 50 MHz): δ = 157.5 (C, C-1'), 129.6 (C, C-3', C-5'), 126.1 (C, C-4'), 116.0 (CH, C-4', C-6'), 67.9 (CH_2 , C-1), 56.7 (CH, C-2), 47.7 (CH_2 , C-3), 32.6 (CH_2 , C-4), 20.6 (CH_2 , C-5), 14.1 (CH_3 , C-6).

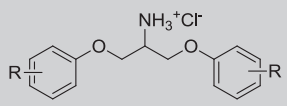
4-(1,3-BIS(4-CHLOROPHENOXY)PROPAN-2-YL)MORPHOLINE (CPD34)

CPD34 was obtained from morpholine (2.1 mmol) and CPD39 (0.21 mmol) using the general procedure 2. The crude residue was triturated with absolute ethyl alcohol and the solid impurities were filtered using vacuum filtration. The filtrate was evaporated under

TABLE A1 Redocking validation protocol (and cross-docking for NorA) and docking pose scores (kcal/mol). RMSD values expressed in Ångström (Å) was calculated from the comparison between redocking results with the cocrystallized conformation and are shown between parenthesis after the calculated energy. Poses are ranked by their GlideScore XP with more negative values representing more energetically stable interactions. Protein structures are described by their Protein Data Bank (PDB) codes, where HM stands for Homology Model

CPD20	Sortase	CPD21	NorA	CPD22	FabI	FstZ	NorA
PDB code	1QWZ		HM	PDB code	4FS3	5XDV	HM
Redocking	-4.98 (1.76 Å)	Cross-docking	-7.03 (1.20 Å)	Redocking	-9.74 (0.63 Å)	-11.59 (0.41 Å)	-7.03 (1.20 Å)
Pose 1	-3.25	Pose 1	-8.50	Pose 1	-9.53	-8.79	-3.66
Pose 2	-2.87	Pose 2	-7.96	Pose 2	-9.08	-8.13	-3.59
Pose 3	-2.66	Pose 3	-3.43	Pose 3	-8.83	-8.07	-1.86
Pose 4	-2.51			Pose 4	-7.46	-7.90	
Pose 5	-2.24			Pose 5	-7.32	-5.20	

TABLE A2 Calculated partition coefficient (ClogP) of compounds CPD1–CPD22



Compound	R	ClogP ^a
CPD1	2-CN	1.95
CPD2	3-CN	1.11
CPD3	4-CN	1.95
CPD4	2-NO ₂	1.67
CPD5	3-NO ₂	2.55
CPD6	4-NO ₂	2.27
CPD7	2-COOCH ₃	2.30
CPD8	3-COOCH ₃	2.38
CPD9	4-COOCH ₃	2.75
CPD10	2-OCH ₃	2.13
CPD11	3-OCH ₃	2.07
CPD12	4-OCH ₃	2.41
CPD13	2-CH ₃	3.40
CPD14	3-CH ₃	3.40
CPD15	4-CH ₃	3.40
CPD16	2-Cl	3.57
CPD17	3-Cl	4.14
CPD18	4-Cl	3.85
CPD19	H	2.48
CPD20	2,3-Benzo	4.94
CPD21	3,4-Benzo	4.94
CPD22	3,4-diCl	5.26

^aCalculated using ACD/ChemSketch software (acdlabs.com).

reduced pressure to give a brown solid in 62% yield; mp: 62.6–64.8°C; IR (ATR) ν_{\max} 2,955, 2,855, 1,595, 1,581, 1,491, 1,468, 1,238, 1,023, 822/cm; ¹H NMR (acetone-*d*₆, 200 MHz): δ = 7.30

(4H, d, *J* = 9.0 Hz, H-3', H-5'), 7.01 (4H, d, *J* = 9.0 Hz, H-2', H-6'), 4.31 (2H, dd, *J* = 10.2 Hz, 5.5 Hz, H-1a), 4.25 (2H, dd, *J* = 10.2 Hz, 5.5 Hz, H-1b), 3.27 (1H, qn, *J* = 5.5 Hz, H-2), 3.60 (4H, t, *J* = 4.4 Hz, H-3), 2.81 (4H, t, *J* = 4.4 Hz, H-4); ¹³C NMR (acetone-*d*₆, 50 MHz): δ = 158.7 (C, C-1'), 130.2 (CH, C-3', C-5'), 126.1 (C, C-4'), 117.2 (CH, C-2', C-6'), 68.2 (CH₂, C-1), 67.3 (CH₂, C-4), 63.4 (CH, C-2), 51.7 (CH₂, C-3).

1-(1,3-BIS(4-CHLOROPHENOXY)PROPAN-2-YL)-4-METHYLPIPERAZINE (CPD35)


From N-methylpiperazine (3.2 mmol) and CPD38 (0.32 mmol) and using the general procedure 1, a colorless oil was isolated using silica gel column chromatography (dichloromethane/methanol 98:2 v/v) in 78% yield; IR (ATR) ν_{\max} 2,935, 2,878, 2,839, 1,595, 1,580, 1,489, 1,467, 1,455, 1,235, 1,027, 1,006, 820/cm; ¹H NMR (acetone-*d*₆, 200 MHz): δ = 7.28 (4H, d, *J* = 8.8 Hz, H-3', H-5'), 6.99 (4H, d, *J* = 8.8 Hz, H-2', H-6'), 4.26 (2H, s, H-1a), 4.23 (2H, s, H-1b), 3.28 (1H, s, H-2), 2.81 (4H, s, H-3), 2.34 (4H, s, H-4), 2.16 (3H, s, H-5); ¹³C NMR (acetone-*d*₆, 50 MHz): δ = 158.7 (C, C-1'), 130.2 (CH, C-3', C-5'), 126.0 (C, C-4'), 117.2 (CH, C-2', C-6'), 67.4 (CH₂, C-1), 62.9 (CH, C-2), 56.7 (CH₂, C-3), 50.8 (CH₂, C-4), 46.5 (CH₃, C-5).

1-(1,3-BIS(4-CHLOROPHENOXY)PROPAN-2-YL)PIPERIDINE (CPD36)

From piperidine (3.9 mmol) and CPD38 (0.39 mmol) and using the general procedure 1, a brown oil was isolated using silica gel column chromatography (dichloromethane 100%) in 74% yield; IR (ATR) ν_{\max} 2,933, 2,852, 2,808, 1,595, 1,580, 1,489, 1,467, 1,235, 1,031, 1,017, 819/cm; ¹H NMR (CDCl₃, 200 MHz): δ = 7.23 (4H, d, *J* = 8.6 Hz, H-3', H-5'), 6.86 (4H, d, *J* = 8.6 Hz, H-2', H-6'), 4.18 (4H, d, *J* = 5.4 Hz, H-1), 3.23 (1H, t, *J* = 5.4 Hz, H-2), 2.74 (4H, s, H-3), 1.60 (4H, s, H-4), 1.48 (2H, s, H-5); ¹³C NMR (CDCl₃, 50 MHz): δ = 157.6 (C, C-1'), 129.5 (CH, C-3', C-5'), 126.0 (C, C-4'), 116.2 (CH, C-2', C-6'), 66.8 (CH₂, C-1), 63.0 (CH, C-2), 52.0 (CH₂, C-3), 26.9 (CH₂, C-4), 24.8 (CH₂, C-5).

For reprint orders, please contact: reprints@future-science.com

Virtual screening of antibacterial compounds by similarity search of Enoyl-ACP reductase (FabI) inhibitors

Leonardo Rander Asse Junior^{‡,1}, Thales Kronenberger^{‡,2} , Mateus Sá Magalhães Serafim³, Yamara Viana Sousa⁴, Isabella Drumond Franco¹, Marília Valli⁵, Vanderlan da Silva Bolzani⁵, Gustavo Claro Monteiro⁵, João Lucas Bruno Prates⁵, Erna Geessien Kroon³, Bruno Eduardo Fernandes Mota⁴, Diego dos Santos Ferreira¹, Renata Barbosa de Oliveira¹ & Vinicius Gonçalves Maltarollo^{*,1}

¹Departamento de Produtos Farmacêuticos, Faculdade de Farmácia, Universidade Federal de Minas Gerais (UFMG), Av. Antônio Carlos, 6627 – Pampulha, Belo Horizonte, MG, Brazil 31270-901

²Department of Internal Medicine VIII, University Hospital of Tübingen, Otfried-Müller-Strasse 14, Tübingen, Germany, 72076

³Departamento de Microbiologia, Instituto de Ciências Biológicas, UFMG, Av. Antônio Carlos, 6627 - Pampulha, Belo Horizonte, MG, Brazil 31270-901

⁴Departamento de Análises Clínicas e Toxicológicas, Faculdade de Farmácia, UFMG, Av. Antônio Carlos, 6627 - Pampulha, Belo Horizonte, MG, Brazil 31270-901

⁵Núcleo de Bioensaios, Biossíntese e Ecofisiologia de Produtos Naturais (NuBBE), Departamento de Química Orgânica, Instituto de Química, Universidade Estadual Paulista (UNESP), Araraquara, SP, Brazil 14800-060

*Author for correspondence: maltarollo@ufmg.br

‡ Authors contributed equally

Aim: Antibiotic resistance is an alarming issue, as multidrug-resistant bacteria are growing worldwide, hence the decrease of therapeutic potential of available antibiotic arsenal. Among these bacteria, *Staphylococcus aureus* was pointed by the WHO in the pathogens list to be prioritized in drug development.

Methods: We report the use of chemical similarity models for the virtual screening of new antibacterial with structural similarity to known inhibitors of FabI. The potential inhibitors were experimentally evaluated for antibacterial activity and membrane disrupting capabilities. **Results & conclusion:** These models led to the finding of four new compounds with antibacterial activity, one of which having antimicrobial activity already reported in the literature.

First draft submitted: 18 July 2019; Accepted for publication: 18 October 2019; Published online: 15 November 2019

Keywords: antibacterial • FabI inhibitors • ligand-based virtual screening • MRSA • similarity search • *Staphylococcus aureus*

Staphylococcus aureus is a Gram-positive bacterium which inhabits the human respiratory system and skin, mostly related to opportunistic infections in immunosuppressed hosts [1]. Among the several clinically relevant *S. aureus* strains, the methicillin-resistant *S. aureus* (MRSA) colonize around 2% of the population and is responsible for severe complications ranging from pneumonia, bacteremia and sepsis [2]. Antimicrobial resistance toward clinically used drugs is a burden to the health system and this phenomenon has been intensified in the last years due to the abuse and irrational use of antibiotics [2,3]. WHO has issued a report alerting for the need of research and development of new antibacterial drugs to specific species and strains and included MRSA as a second highest priority category [4].

An ideal antibacterial should be able to selectively target a protein or pathway which is essential to bacteria but absent in the human host. In this sense, the bacterial fatty acid (FA) synthesis pathway, composed of monofunctional individual enzymes is different from the biosynthesis of FAs in mammals, which is composed of a distinct enzymatic complex. Fatty acid synthesis pathway harbors several excellent drug-targets, such as FabI, which catalyzes the limiting step of the FA chain extension [5]. Moreover, inhibiting FabI compromises cell membrane maintenance [6]. FabI is a strategic drug target in the development of narrow-ranged antimicrobial agents [5] since it is known

that not all Gram-positive species are susceptible to FabI inhibition [7] because FAs can be salvaged from host sources. On the other hand, *S. aureus* continues performing FA synthesis even in the presence of serum as a FA source [8–10]. Additionally, FabI inhibitors recently reached clinical trials, indicated for the treatment of acute skin and skin structure infections [11], and are subject of several patents [12–15]. This information altogether indicates their potential as drug candidates and can be considered optimal prototypes for the discovery/design of new antibacterial molecules.

This work focused on specific classes of *S. aureus* FabI inhibitors, namely aminopyridine derivatives [16], indole naphthyridinones [17], diazepinones [18,19], imidazoles [20], 1,2,3,4-tetrahydropyrido[3,4-*b*]indoles [21] and azetidine ene-amides [22]. Regarding the large available amount of compound information, computational techniques for identification of new lead compounds may be applicable. Within the diversity of drug design strategies, virtual screening (VS) is a computational (or *in silico*) approach which allows the study of sizable compound databases for the identification of ligands with potential biological activity, and typically precedes experimental assays. Specifically, ligand-based VS (LBVS) permits the evaluation of numerous compounds by their similarity with known inhibitors. The advantages of applying this strategy in a drug design protocol are due to its high performance, versatility and low cost [23].

In this work, we report a new approach for the development of tridimensional similarity models as well as a comparison of these models with optimized docking protocols, 3D-electrostatic and 2D similarity which resulted in four new hits with validated antibacterial activity.

Materials & methods

Tridimensional similarity models were built based on known FabI inhibitors, individually or in combination, using ROCS (shape/color) queries. Also, electrostatic and 2D fingerprint models were created from the same database. The best models were employed in a VS protocol, and the diversity of hit composition was analyzed.

Dataset selection & libraries preparation

The dataset used for models' generation and validation included 166 known *Sa*FabI inhibitors, of which the activity was reported under the same experimental conditions (Supplementary Table 1). This work focused on some specific classes of *Sa*FabI inhibitors, namely aminopyridine derivatives [16], indole naphthyridinones [17], diazepinones [18,19], imidazoles [20], 1,2,3,4-tetrahydropyrido[3,4-*b*]indoles [21] and azetidine ene-amides [22]. These compounds were grouped into two categories, namely active (74 compounds with IC₅₀ values lower than 1 μM) and inactive (92 compounds with IC₅₀ values higher than 1 μM), applying the suggested cut-off for lead compounds in drug discovery campaigns [24,25]. An additional dataset of decoys were generated using a Database of Useful Decoys: Enhanced, in the proportion of fifty decoys to each active compound, for validation purposes [26]. Clean drug-like (CDL) database from ZINC [27] filtered with 90% similarity unique molecules; US FDA approved drugs from ZINC (both accessed in May 2016); an in-house set of compounds synthesized by Renata Barbosa de Oliveira research group at Pharmaceutical Chemistry Laboratory – UFMG; AfroDb Natural Products (AfroDb) [28]; NuBBE Database (NuBBE) [29,30]; and Traditional Chinese Medicine (TCM, Taiwan) [31] libraries were employed in VS underwent same treatment.

The tridimensional structures used in the models were drawn using Discovery Studio software (Dassault Systèmes BIOVIA, CA, USA). All of them had their protonation and charges calculated using QUACPAC 1.7.0.2 [32] and AM1BCC method [33], using the *fixpka* option and the *AM1bccspt* force-field, followed by conformer generation using OMEGA 2.5.1.4 [34].

Chemical similarity models generation

3D chemical similarity models (CSM) were generated based on the eleven most active molecules as templates (T1-T9, triclosan and AFN-1252; Supplementary Table 2), using vROCS 3.2.1.4 [35]. Structural information, by the means of *Sa*FabI crystal structures, is available for four of the compounds (triclosan, AFN-1252, T7 and T9), which were considered as the active conformation and used for the alignments. Several different combinations of CSM were generated by heuristically excluding chemical features, as well as applying weights aiming to improve active/inactive classification performance. CSM containing more than one molecule as a template (hybrid models) were also constructed to generate an ensemble of the volume occupied by two or more *Sa*FabI inhibitors and chemical features. Molecules employed in most predictive CSM were also used to calculate electrostatic similarities using EON 2.2.0.5 software [36].

Finally, classical 2D similarity and molecular docking protocols were performed for comparison with CSM and electrostatic similarity analysis. 2D queries were created and their PubChem [37] fingerprints were calculated using PaDEL-Descriptor [38]. The validation process consisted of calculating the 2D similarity between each compound in the validation database against the query based on the Tanimoto coefficient. The calculations of PubChem molecular fingerprints, as well as the similarity coefficient, were performed using a workflow implemented on KNIME 3.4.0 (Supplementary Figure 1) [39]. Molecular docking protocols were optimized to better reproduce experimental binding modes of FabI inhibitors by using Surflex Docking software (see Supplementary Data for protocol details) [40].

LBVS protocol

Subsequently to model generation, we performed two VS protocols. In the first one, we filtered the six prepared libraries directly on three best CSMs and substances with TanimotoCombo values above a determined cut-off were selected. Compounds highly ranked among two or three models were visually inspected. Visual inspection aimed to exclude compounds that did not match important features contributing to the biological activity (e.g., an H-bond acceptor which interacts with NADP⁺ and Tyr157 at FabI-binding site). Finally, the remaining compounds underwent prediction of ADMET properties using admetSAR [41] and ProTox [42]. Moreover, the inhibitory activity against SaFabI, expressed in terms of IC₅₀, was predicted using a HQSAR model from the literature [43]. Only compounds with desirable predicted values were selected for further experimental testing. Thereafter, a second VS approach was performed, in which the best 500-ranked compounds from CSM had electrostatic similarities calculated and consensus among models was carried out as the first protocol [44]. Further, visual inspection and prediction of ADMET and SaFabI IC₅₀ values were also carried out.

Cell lines

Vero cells (American Type Culture Collection, ATCC number CCL-81) were maintained in minimal essential medium (MEM; Cultilab, Campinas, Brazil). Media were supplemented with 5% fetal bovine serum (Cultilab), 200 U/ml of penicillin (Cultilab), 100 µg/ml of streptomycin (Sigma-Aldrich, MO, USA) and 2.5 µg/ml of amphotericin B (Cultilab). Bacterial strains used were *Escherichia coli* (ATCC n° 35218), *S. aureus* (ATCC n° 29213), MRSA (ATCC n° 43300) and four clinical isolates of MRSA (*mecA* positives, previously described in [45]).

In vitro screening & minimum inhibitory concentration determination

Antimicrobial activity of the compounds was evaluated using the broth microdilution method in 96-well microplates according to Clinical and Laboratory Standards Institute [46]. First, compounds were diluted in 100 µl of Mueller Hinton broth (MHB; Oxoid, Thermo Scientific, MA, USA) to a final concentration of 500 µM. The same volume of a bacterial suspension containing 10⁵ colony forming units per milliliter (CFU/ml) was added to the previous solution in 96-well microplates, resulting in a final concentration of 250 µM. After incubation at 37° C for 24 h, plates were inspected visually for bacterial growth and the absorbance at 600 nm of each well was read using a spectrophotometer (VersaMax, Molecular Devices, CA, USA). The percentage of inhibition was calculated as the percentage of growth reduction compared with vehicle control (1% v/v DMSO, Sigma-Aldrich,). Wells showing no visual bacterial growth were considered as hits and a serial dilution of each compound (from 250 to 0.39 µM) was further assessed in the same conditions to determinate minimum inhibitory concentration (MIC). In each plate, a viability control (bacterial suspension only), an inhibition control of MHB containing a serial dilution ranging from 0.16 to 0.02 µg/ml (0.48–0.06 µM) of penicillin G for *S. aureus*, a serial dilution ranging from 4 to 0.5 µg/ml (8.38–1.05 µM) of gentamicin for *E. coli*, and a serial dilution of vancomycin ranging from 16 to 2 µg/ml (11.52–1.38 µM) for MRSA strains and a sterility control (medium only) were included. All tests were performed in triplicates.

Mammalian cell viability

Cytotoxicity in mammalian cells was assessed using the 3-[4, 5-dimethylthiazol-2-yl]-2, 5-diphenyl tetrazolium (MTT) reduction assay [47]. Vero cells were seeded in 96-well plates (8 × 10⁴ cells per well) and incubated at 37° C and 5% CO₂ atmosphere. After 24 h of incubation, 200 µl of fresh medium containing serial dilution of compounds (from 125 to 3.9 µM) was added to the plates. A serial dilution of DMSO ranging from 1.25 to 0.04% was used as viability control. After 48 h of incubation in the same conditions, 100 µl of MTT solution in MEM (5 mg/ml) was added to each well. After 3 h of incubation at 37°C and 5% CO₂ atmosphere, the medium

was removed and 100 μ l of DMSO was added to each well to solubilize formazan crystals. Absorbance at 570 nm of each well was read using a spectrophotometer (VersaMax, Molecular Devices). Cytotoxic concentration of 50% (CC_{50}) of each compound was calculated and its selectivity index (SI) was determined, as it is the relation between the CC_{50} and the effective concentration (MIC). All conditions were tested in triplicates.

Zeta potential measurement

Aiming to evaluate the effect of active compounds on the bacteria membrane, in accordance with FabI inhibition, the zeta potential of *S. aureus* and *E. coli* were measured in the presence of selected hit compounds. In order to assess zeta potential measurements bacteria were grown in 4.0 ml of Mueller-Hinton broth overnight in a shaking bath at 37° C and 180 r.p.m. After incubation bacterial surface was cleaned by centrifugation at 5000 \times g and washing three times with phosphate-buffered saline (0.1 M) and resuspended in phosphate-buffered saline to 10⁹ CFU/ml. The resulting suspensions were used to fill clear disposable folded capillary zeta cells (Malvern Instruments, Malvern, UK).

To examine the effect of our hits, 1.0 ml containing 10⁹ CFU/ml of MRSA (ATCC n° 43300), *S. aureus* and *E. coli* cultures were exposed to 1.0 ml of the compounds at its MIC values, diluted in MHB. After that, suspensions were incubated for 2 and 4 h in shaking bath with the same conditions and prepared for zeta potential measurements as described above. The zeta potential measurement (mV) of bacterial cells was done in triplicates at 25° C with a zeta potential analyzer at 50 V (Zetasizer Nano ZS, Malvern Instruments) [48].

Chemical synthesis & acquired reagents

Compound ZINC04052016 was acquired from MolPort (NY, USA); compounds ZINC09664703 and ZINC04425632 from Enamine (Kyiv, Ukraine); compound ZINC05137604 from Specs; donepezil was acquired from AK Scientific (CA, USA); triclosan, nabumetone, flubendazole, crizotinib, topotecan and ampicillin were acquired from Sigma-Aldrich. ZINC000084154603 was kindly provided by NuBBE research group; curcumin was gently provided by Laboratório de Virologia Aplicada from Universidade Federal de Santa Catarina, Brazil; and compounds **8** [49], **28** and **35–37** [50], and **38** [51] were previously described and gently provided by The Laboratory of Pharmaceutical Chemistry of UFMG, Brazil. Compounds **26** and **27** were synthesized in one-step from the 1-methyl-2-pyrrolicarboxylic acid and the corresponding amines in the presence of 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride and N-hydroxysuccinimide in dichloromethane (42–59% yield).

Statistical analysis

The statistical validation process consisted of calculating the ROC curve and its respective AUC value, as well as the enrichment values (0.5, 1.0 and 2.0%). Confusion matrix for the best-generated models (selected based on the AUC) of each approach was generated. So as to determine false positive/negative rates, we set the TanimotoCombo value of the best-ranked decoy as the cut-off (Tanimoto coefficient values for shape plus chemical features, named color, or just Tanimoto, for 2D queries). Also, the Matthews Correlation Coefficient [52] for each evaluated query was calculated from the results of the confusion matrices. Sidak's multiple comparison test quoted to a significant level of $p < 0.05$ was used to determine if the zeta potential of treated cells were significantly different from its blank and antimicrobial control conditions.

Results

Similarity models construction & validation

From all 552 constructed and validated CSMs, we selected the best eleven models with the highest AUC values (Table 1). Those selected models were evaluated according to its ability to distinguish active compounds from decoys by a confusion matrix. From this analysis, we selected queries 433, 531 and 318 which are the only models with a true positive rate (TPR) above than 50% (model 432 also showed TPR >0.5 but this model is very similar to 433) for VS protocols.

Further investigation on the influence of each chemical feature over the AUC and enrichment levels revealed the central hydrogen bond acceptor (HBA) and naphthyridinone hydrogen bond donor (HBD) features as important to improving the AUC (Figure 1A & B). However, no significant change was observed in terms of enrichment values when comparing the presence or absence of HBA/HBD, suggesting that this is not an important feature for model sensitivity (Figure 1C). Other chemical features such as the presence or weight of the ring, hydrophobic

Table 1. Summary of the statistical parameters evaluated in the validation of the ten best ROCS models.

Query	AUC	Cl _{inferior}	Cl _{superior}	TC1FP*	TPR	MCC	ACC
547	0.955	0.930	0.974	0.946	0.405	0.613	0.989
433	0.953	0.934	0.970	1.191	0.649	0.778	0.993
548	0.951	0.929	0.969	1.082	0.405	0.613	0.989
531	0.951	0.921	0.974	1.295	0.581	0.734	0.992
534	0.950	0.922	0.973	1.295	0.459	0.646	0.990
432	0.948	0.925	0.968	1.083	0.622	0.769	0.993
535	0.948	0.918	0.971	1.451	0.297	0.497	0.986
318	0.945	0.921	0.967	1.134	0.649	0.786	0.993
536	0.944	0.913	0.969	1.451	0.203	0.419	0.985
494	0.944	0.913	0.969	0.832	0.473	0.665	0.990
495	0.944	0.922	0.963	1.044	0.000	-0.002	0.982

Information regarding all calculated models is provided in the Supplementary Data.

ACC: Accuracy; AUC: Area under the ROC curve; MCC: Matthews correlation coefficient; TC1FP*: TanimotoCombo value for the first decoy predicted as false positive; TPR: True positive rate.

Table 2. Summary of statistical parameters evaluated in the validation of the electrostatic similarity search.

Query	Molecule	Query	AUC	Tc 1 st FP*	TPR	MCC	ACC
318	T6	318	0.700	1.29	0.160	0.397	0.985
433a	T1	433a	0.719	1.29	0.147	0.380	0.984
433b	T6	433b	0.861	1.33	0.093	0.266	0.983
531a	T1	531a	0.715	1.29	0.147	0.380	0.984
531b	T6	531b	0.853	1.34	0.093	0.266	0.983

ACC: Accuracy; AUC: Area under the receiver operating characteristic curve; MCC: Matthews correlation coefficient; TC1FP*: Tanimoto similarity value for the first decoy predicted as false positive; TPR: True positive rate.

and cation features did not significantly improve the overall model AUC values (Supplementary Figure 2). The significant contribution of the HBA to improve the ability to distinguish actives from decoys, when compared with the respective models without this feature, come as no surprise considering the importance of the central HBA interactions with the Tyr157 and NADPH (Figure 2 & Supplementary Table 3) [43,53,54]. That information guided the visual inspection step of a VS protocol.

As the second VS protocol was defined to carry electrostatic similarity calculations after 3D chemical similarities, we used T1 and T6 template compounds to perform this step (Figure 3). Those templates were employed in the construction of selected ROCS models. Then, the complementary electrostatic similarity models were generated using the three best ROCS models as alignment input, as well as previously aligned datasets were used as input for validation. Since they were generated from the same two template compounds (T1 and T6) and that electrostatic models require a single molecule as input, there are five EON models.

Electrostatic models generated based on the molecule T1 had more accuracy and higher AUC values, despite the similar distribution of electron density. T1-based models agree with previous SAR suggestions that unsubstituted HBA in the 3' position would improve activity, which was developed in electrostatic similarity models from imidazole derivatives [54]. EON implementation opens room for improvements, such as the rigid alignment, which assumes that partial charges are conformationally invariant [55]. However, to integrate it in a solid pipeline, association with other techniques could improve the odds of finding true positives by minimizing the false-positive rate, which is evident by its high AUC values on this study despite the very low TPR (Figure 3 & Table 2).

Complementarily, 2D similarity and docking were carried out and validated using the same validation protocol for comparison with ROCS and EON performances (Supplementary Figure 2 & Supplementary Table 4). Molecular docking optimizations protocol procedures and results are described in the Supplementary Data. In summary, molecular docking simulations were performed by varying several internal parameters of Surflex docking software and evaluated initially by RMSD values comparing against the crystallographic binding modes of known FabI inhibitors. After, the two protocols which best reproduced the experimental binding modes (the most precise model presented a RMSD equal to 0.796 Å for redocking validation using structure PDB ID 4D42 and RMSD varying

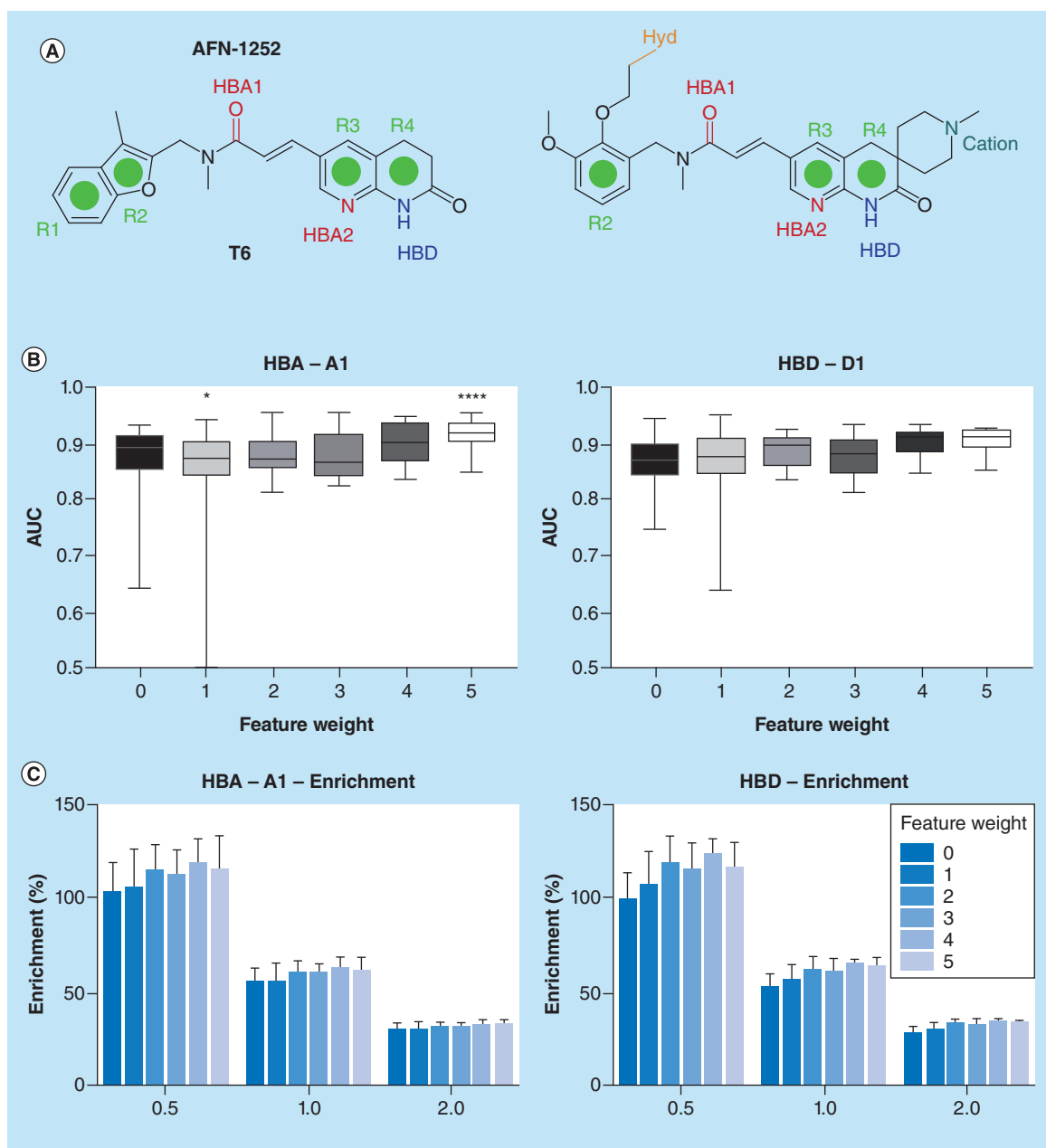


Figure 1. Features employed in best chemical similarity models generation and their impact on classification metrics. (A) Summary of chemical features employed in the primary model's generation. Feature position is highlighted on the T6 template compound and AFN-1252. Influence of hydrogen bond acceptor feature and hydrogen bond donor feature on the ROCS similarity model quality. Models with differently weighted hydrogen bond acceptor/hydrogen bond donor features were compared with models without this feature, both in terms of area under curve (B) and enrichment values (C). Statistical significance between models and the control without the chemical feature was evaluated using Mann-Whitney non-parametric model and the p-value is represented as follow: ****<math><0.0001</math>. HBA: Hydrogen bond acceptor; HBD: Hydrogen bond donor; Hyd: Hydrophobic; R: Ring.

between 0.593 and 2.301 Å for crossdocking simulations) were employed to generate enrichment curves. Despite acceptable values of AUC for all 2D similarities and docking simulations (AUC >0.8 except 2D similarity for triclosan), all TPRs are above 50%. Then, indicating the superior ability of ROCS models to distinguish active compounds from decoys, when compared with all other techniques.

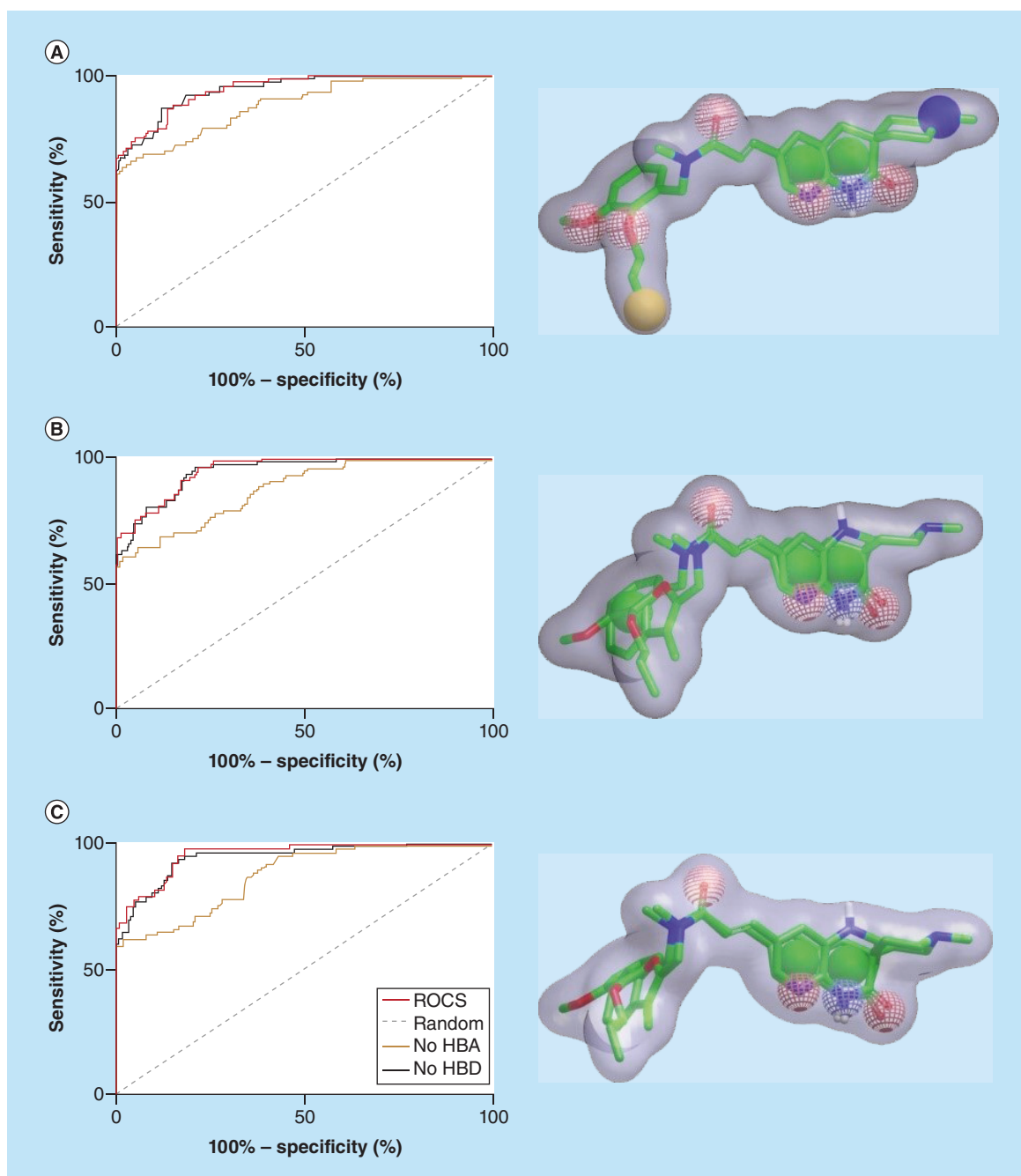


Figure 2. Receiver operating characteristic curve plot for best three chemical similarity models. Best validated models generated by ROCS and complementary validation analyses for (A) model 318 (B) model 433 and (C) model 531. Orange and black curves represent artificial models without the main hydrogen bond acceptor and hydrogen bond donor features respectively. The dashed gray line represents the random model. 3D representation of the queries (right side), where the spheres are color-coded by chemical property: green for rings, dashed-red are hydrogen bond acceptor, dashed-blue is hydrogen bond donor, blue are cations and yellow represents hydrophobic region. HBA: Hydrogen bond acceptor; HBD: Hydrogen bond donor.

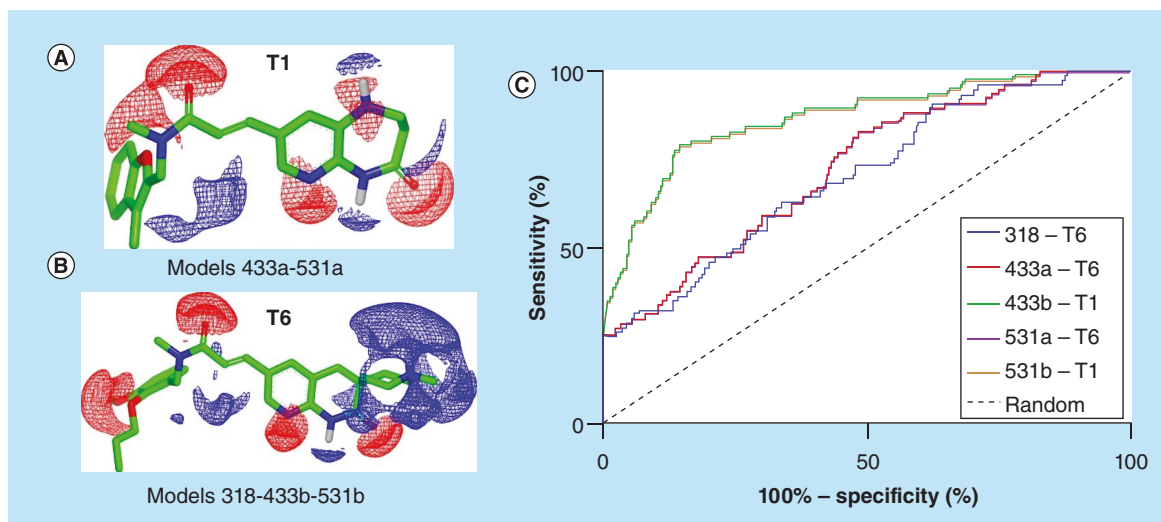


Figure 3. Electrostatic similarity queries calculated based on the two molecules. (A) T1 (models 433a/531a) and (B) T6 (318/433b/531b), which compose the best ROCs models. (C) ROCs for the five electrostatic similarity models dashed gray line represents the random model.

LBVS & broth microdilution antibacterial screening

Both VS protocols found 17 hits each (Supplementary Table 5). Interestingly, it was found that most of the screened molecules have similar functional fused ring systems with a HBD embedded and a carbonyl moiety which aligns with the HBA chemical feature of the models. Among our VS hits, curcumin and ferulic acid derivatives were common scaffolds. The compound ZINC000034006254 (propyl ferulate, as well as other derivatives from this scaffold) has reported antibiotic activity against *E. coli* and *S. aureus* [56,57]. Additionally, curcumin, which contains a carbonyl group acting as a HBA and a complex ring system, has been shown to inhibit FabI and act as a moderately potent antibiotic [58]. Aiming to avoid compounds with high predicted toxicity and higher values of IC_{50} , only fourteen compounds among the virtual hits were selected for experimental validation (Supplementary Table 6). Aiming to represent all the natural products similar to curcumin found in LBVSs (compounds **12–15**, **17**, **33** and **34**), we included curcumin as a control in our broth microdilution experimental assays (Table 3). Selected hits were initially screened against *S. aureus*, MRSA and *E. coli* (at the initial concentration of 250 μ M, Table 3). From this first screening, we found the crizotinib (compound **23**) as a potential hit for all tested organisms and determined its MIC. Compounds **8** and **28** showed interesting activity against *S. aureus* (~70% of growth inhibition) with compound **28** showing a weak inhibition of *E. coli* growth (24.5%).

Specifically, compound **8** is structurally similar to curcumin, which was used as a positive control in this study. Then, we carried a 2D similarity search on our available compounds (in-house and NuBBE databases) aiming to find new compounds structurally similar to curcumin. This new screening retrieved 13 new compounds (Supplementary Table 7) and the most active ones are compounds **35–37** which presented 50, 28.1 and 33% of growth inhibition against *E. coli*. These results support further studies of plants which contain curcumin derivatives as potential antimicrobial agents as well as the isolation of new C_3C_6 arylpropanoids and the design of new ones to evaluate its activities against bacteria. We also tested a compound retrieved from docking studies (Supplementary Table 7; compound **48**) which, together with other eight compounds from similarity search, did not show interesting activities against tested microorganism (Supplementary Table 7).

Compounds are represented by their database code and origin. Experimental initial screening evaluated microdilution broth. Values represent growth inhibition (%) at 250 μ M. The percentage of inhibition was calculated as the percentage of growth reduction compared with vehicle control (1% v/v DMSO).

MIC to reference & clinical strains

In our initial screening, five compounds showed visual inhibition of bacterial growth compounds **8**, **23** (crizotinib) and **28**, as well as curcumin and triclosan. Crizotinib showed the highest inhibitory percentage among them, with over 97% inhibition (Table 3) to all three reference strains initially tested (*E. coli*, MRSA and *S. aureus*). After that,

Table 3. The consensus of virtual screening protocols.

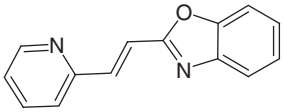
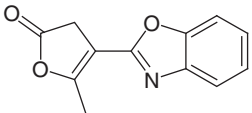
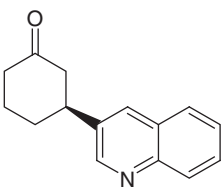
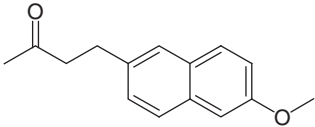
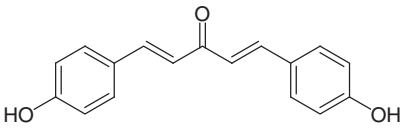
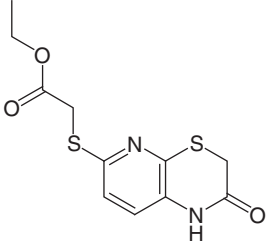
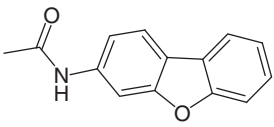
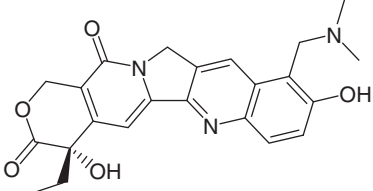
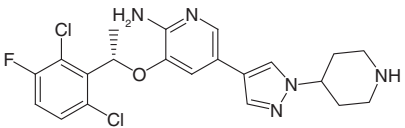
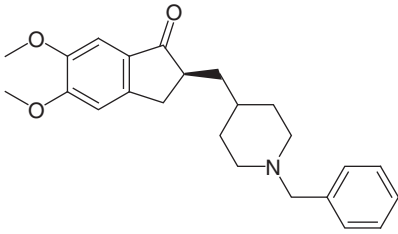
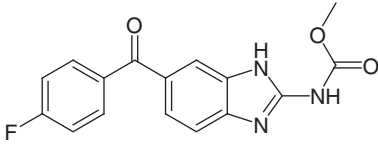
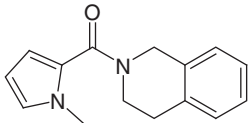
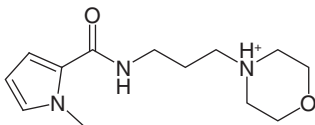
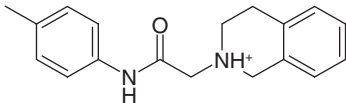
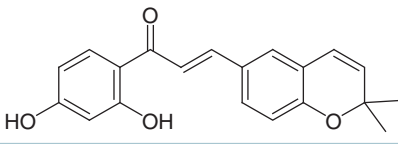
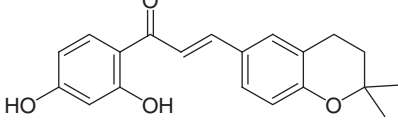
Number (ZINC ID / source)	Database	Chemical structure	<i>Staphylococcus aureus</i>	MRSA	<i>Escherichia coli</i>
Virtual screening - Protocol 1					
Compound 1 (ZINC04052016)	CDL		0	0	31.3
Compound 2 (ZINC09664703)	CDL		0	0	33.3
Compound 3 (ZINC05137604)	CDL		0	0	15.8
Compound 5 (Nabumetone)	FDA		2.3	10.4	14
Compound 8 RS11	In-house		69.8	26.6	20.9
Compound 10 ZINC000084154603	NuBBE		3.1	14.2	27.7
Virtual screening - Protocol 2					
Compound 19 (ZINC04425632)	CDL		0	18	15.8
Compound 21 (topotecan)	FDA		7.2	8.7	9.4
Compound 23 (crizotinib)	FDA		98.6	98.8	97.2
MRSA: Methicillin-resistant <i>Staphylococcus aureus</i> . NT: not tested.					

Table 3. The consensus of virtual screening protocols (cont.).

Number (ZINC ID/source)	Database	Chemical structure	<i>Staphylococcus aureus</i>	MRSA	<i>Escherichia coli</i>
Compound 24 (donepezila)	FDA		0	0	21.3
Compound 25 (flubendazol)	FDA		0	0	0
Compound 26 PATH-m	In-house		0	12.6	26.2
Compound 27 PA3Mo-m	In-house		10.14	18.7	48.8
Compound 28 Thac-M	In-house		98.4	96.7	57.7
Compound 33 (ZINC000043069427; <i>Dorstenia barteri</i> var. <i>subtriangularis</i>)	AfroDb		NT	NT	NT
Compound 34 (ZINC000043121107/semi-synthetic)	AfroDb		NT	NT	NT

MRSA: Methicillin-resistant *Staphylococcus aureus*. NT: not tested.

all five compounds had its MIC values determined, ranging from 250 to 0.39 μM , with highlights to triclosan and crizotinib, which showed activity against all the strains tested, and especially to triclosan with the lowest MIC values between the five compounds (Table 4).

Hereafter, clinical MRSA strains were tested in the same conditions in order to evaluate the compounds antibacterial activity and to obtain its MIC concentration. Found values ranged from 250 to 0.39 μM , and here triclosan also showed the lowest results to all four clinical strains tested, with MIC values ranging from 6.25 to 0.39 μM . Crizotinib also showed activity against all strains, with MIC ranging from 125 to 15.62 μM , comparable with its activity on the MRSA reference strain (31.25 μM). Despite its known activity, curcumin also presented activity against two clinical strains (125 and 62.5 μM). At last, compounds **8** (250 and 125 μM) and **28** (250 μM) showed the lowest activity against two out of four strains. All results are summarized in Table 5.

Table 4. Minimum inhibitory concentration range (μM) for compounds that showed antibacterial activity in the initial screening.

Compounds	<i>Staphylococcus aureus</i> (ATCC 29213)	<i>Escherichia coli</i> (ATCC 35218)	MRSA (ATCC 43300)
28	250	–	250
8	250	–	–
Crizotinib	125	250	31.25
Curcumin	250	–	125
Triclosan	6.25	0.78	0.39
Penicillin G	0.06	NT	NT
Gentamycin	NT	1.05	NT
Vancomycin	NT	NT	1.38

Values presented are representative of at least three independent experiments.

ATCC: American Type Culture Collection; MRSA: Methicillin-resistant *Staphylococcus aureus*; NT: Not tested.

Table 5. Minimum inhibitory concentration range (μM) of the compounds against methicillin-resistant *Staphylococcus aureus*.

Compounds	Strain 5749	Strain 5912	Strain 6100	Strain 6154	CC ₅₀ \pm SD
28	–	250	–	250	131.39 \pm 3.5
8	250	–	125	–	15.67 \pm 0.8
Crizotinib	125	15.62	62.5	62.5	4.23 \pm 0.9
Curcumin	–	125	–	62.5	10.82 \pm 1.5
Triclosan	6.25	0.78	0.39	0.78	>250
Vancomycin	1.38	1.38	1.38	1.38	NT

The CC₅₀ of the active compounds in Vero cells (μM). Values presented are representative of at least three independent experiments.

CC₅₀: Cytotoxic concentration of 50%; NT: Not tested.

Cytotoxicity concentration to mammalian cells (CC₅₀)

Hereafter, knowing its antibacterial activity, cytotoxicity of the compounds was evaluated assessing the MTT tetrazolium reduction assay against Vero cells. The 50% cytotoxic concentration (CC₅₀) refers to a compound concentration that reduces the cell viability in 50%. Results showed the lowest values for crizotinib (4.23 \pm 0.87 μM) and the highest to triclosan (>250 μM). Compounds **28**, **8** and curcumin showed values of 131.39 \pm 3.5, 15.67 \pm 0.8 and 10.82 \pm 1.5 μM respectively. In combination, compounds showed SI values ranging from 0.02 to 641.03, following the lowest and highest CC₅₀ values (crizotinib and triclosan, respectively). Results are summarized in Table 5.

Evaluation of bacterial membrane surface charge disturbance

Following the biological activity evaluation, zeta potential measurement was assessed to obtain information about the compound's disturbance to the bacterial cell membranes charge and its possible mechanisms of action. *Staphylococcus aureus* cells showed no surface charge difference ($p < 0.05$) in any treatment conditions compared with its control's values (2 and 4 h, respectively; Figure 4A). For MRSA cells similar results were obtained, as no surface charge difference ($p < 0.05$) was observed to any conditions compared with its controls (2 and 4 h, respectively; Figure 4B). For *E. coli* only one treatment condition showed a surface charge difference ($p < 0.05$) compared with its controls (2 and 4 h respectively). Crizotinib presented a surface charge of -7.8 ± 0.3 and -6.9 ± 0.2 against its blank control (no-stress condition) surface charge of -18.1 ± 0.4 and -18.2 ± 0.6 , and antimicrobial control (gentamycin) surface charge of -18.7 ± 0.3 and -18.7 ± 0.8 . Triclosan showed a zeta potential of -19.6 ± 1.2 and -19.8 ± 1.4 (Figure 4C).

Discussion

Molecular similarity has been extensively applied as a tool for medicinal chemistry [59–63]. Over the last decade, several molecular representations and similarity metrics have become common for aiding drug discovery, and the quality of the designed compound is highly dependent on the molecular representation and similarity metric used. We have used here a comparison between classical docking and three similarity approaches in the search for new

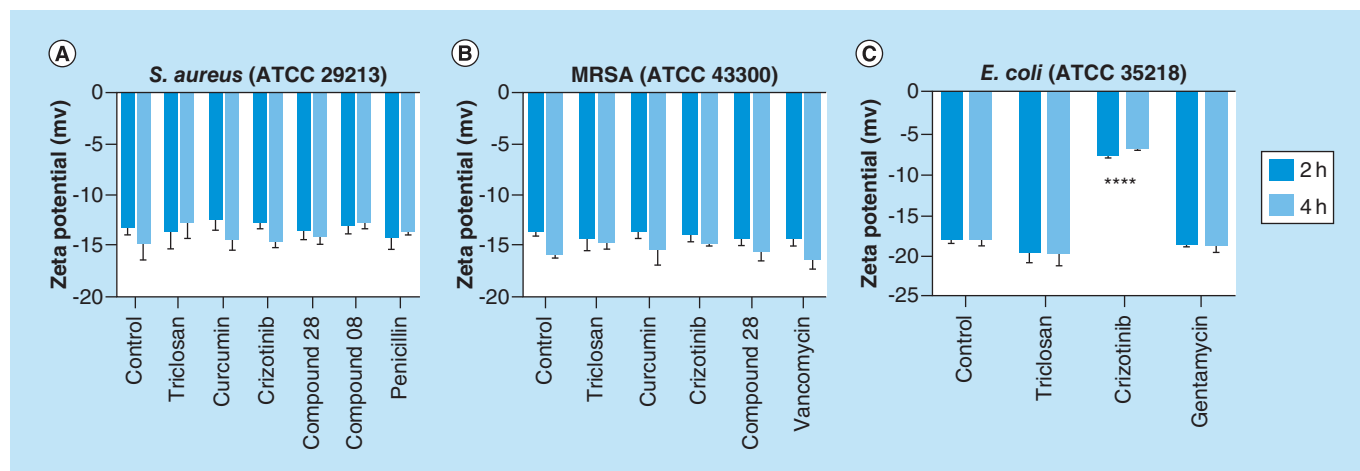


Figure 4. Surface charge measured after 2 and 4-h incubation period of the relevant cells and respective test and control compounds. **(A)** *Staphylococcus aureus* cells during exposure to treatment conditions (triclosan, curcumin, crizotinib, compound **8** and **28**) as well as its blank control and antimicrobial control (penicillin G). **(B)** MRSA cells during exposure to treatment conditions (triclosan, curcumin, crizotinib and compound **28**) and antimicrobial control (vancomycin). **(C)** *Escherichia coli* cells during exposure to treatment conditions (triclosan and crizotinib) and antimicrobial control (gentamycin). Error bars represent standard deviation values and **** represents $p < 0.0001$, on the two-way ANOVA followed by Sidak's multiple comparison test. MRSA: Methicillin-resistant *Staphylococcus aureus*.

antibacterial based on the structure of known FabI inhibitors for *S. aureus*. ROCS similarity models had an overall better performance when compared with the 2D-fingerprint screening and the electrostatic similarity models, as shown by higher AUC, TPR and Matthews correlation coefficient values. Structure-based design using docking toward FabI was previously shown to be inefficient, due to high mobility of the flexible loop covering the active site [54], which motivated us to focus on a tridimensional similarity search with ROCS.

ROCS perform VS by overlaying conformers to a query molecule, which is based on the description of the molecules as atom-centred Gaussian functions that are used to assess volume overlay between a query molecule and a database molecule [35]. ROCS performance on the DUD targets has already been shown as good and comparable with docking programs in terms of enrichment [64]. Additionally, it was reported that edition of weighting factors for ComboScore does not lead to enhanced screening performance unless parameterized specific datasets, which is in agreement with our observations that secondary and tertiary queries were significantly better than primary counterparts [64]. Previous reports that queries generated from multiple aligned active molecules perform better in VS [64,65] are also corroborated by our observation that hybrid queries had better performance during validation, with higher AUC values.

Our group previously reported the importance of two main chemical features for *Sa*FabI inhibition: the HBA from a central carbonyl moiety or imidazole ring, which interacts with both the NADPH cofactor and the amino acid residue Tyr157 and chemical groups accepting or donating hydrogen bond with the oxygen atom of the Ala97 main chain [43]. Other groups reported the importance of HBA from the imidazole ring for FabI inhibition in *Francisella tularensis*, also by generating ROCS and EON similarity models [54], which we confirmed as fundamental for improving overall model quality.

Here, we employed an antibacterial screening and five compounds (**8**, **23**, **28**, curcumin and triclosan) showed activity with complete visual growth inhibition of bacteria tested. The MIC determination results showed values ranging from 250 to 0.39 μM and most of compounds were active against Gram-positive bacteria (all five compounds) than Gram-negative (triclosan and crizotinib). Triclosan and crizotinib were active against all tested ATCC strains and clinical isolates.

Although only *E. coli* was evaluated in this work, the most potent activity of compounds toward Gram-positive may be explained due to its difference in cell-wall composition, as the external membrane of Gram-negative bacteria may hinder penetration and/or justify the retention of molecules outside of the cell [53,54]. Crizotinib showed the lowest CC_{50} values ($4.23 \pm 0.87 \mu\text{M}$) and triclosan was the least cytotoxic compound ($>250 \mu\text{M}$). Vehicle controls with equivalent concentrations also ensured that the biological activity is directly dependent on the action of the

compound. Ideally, to be considered as a prototype or potential drug candidates, a compound SI value must be at least equal or higher than 10 [66]. In this study, triclosan showed the highest results, between 4 and 641.03, with crizotinib, curcumin, compound **8** and **28** showing the lowest SI values, ranging from 0.02 to 0.53. These four compounds SI values are in fact below 10, however structural modifications, analog and hybrid molecules synthesis are some options to reduce its toxicity, changing its size, chemical properties, solubility and polarity.

Crizotinib is a RTK inhibitor, used in the treatment of anaplastic lymphomas and its oncogenic variants. Comparative studies regarding lung cancer treatment, showed that crizotinib had higher activity than other usual treatment options [67]. Also, despite showing high cytotoxicity, this drug is used with certain safety in clinical medicine [68], even more, if we consider that a bacterial infection treatment takes much less time (usually weeks) than a cancer treatment (months/years).

Furthermore, there is no antibacterial activity reported to crizotinib until today, especially against clinically relevant bacteria as shown in this work. The MIC values obtained here for this drug against MRSA strains (56.25 to 7.03 $\mu\text{g/ml}$) are in the same magnitude as clinically approved antibiotics for *S. aureus*, for example, carbenicillin (2–8 $\mu\text{g/ml}$), ceftazidime (4–16 $\mu\text{g/ml}$) and chloramphenicol (2–16 $\mu\text{g/ml}$), hence appearing as a promising candidate for an antibacterial drug [46]. With the possibility of a repositioned drug, many other studies could be done for this purpose, which is favorable considering time and costs reduction to clinical, pharmacological and toxicological trials. In this matter, *in vivo* efficacy assays would be interesting to study this drug for a new purpose of clinical treatment. Crizotinib's experimental cytotoxicity would not favor its repurposing as antibacterial, however the drug repositioning process depends on several further studies, such as determination of cytotoxicity in several cell lines and *in vivo* toxicity in a suitable dosage, a proper route of administration, exposure of the drug to organism (duration of treatment), and formulation and drug delivery studies. In this sense, although we cannot exclude the possibility of repurposing, we would rather consider crizotinib as a potential lead compound for further antimicrobial development.

In this context, crizotinib activity against *E. coli* (ATCC 35218), specially due to its beta-lactamase production, showed interesting results (112.58 $\mu\text{g/ml}$) in comparison with MIC values of some antimicrobials of clinical use, such as ampicillin (>32 $\mu\text{g/ml}$), piperacillin (>64 $\mu\text{g/ml}$), ticarcillin (>128 $\mu\text{g/ml}$), amoxicillin (>256 $\mu\text{g/ml}$) and amoxicillin-clavulanate (4/2–16/8 $\mu\text{g/ml}$). Despite *S. aureus* (ATCC 29213) being susceptible to all five compounds, no equivalent activity (56.29 $\mu\text{g/ml}$, after unity conversion) was shown to some antimicrobials, such as linezolid, amikacin, kanamycin and piperacillin (1–4 $\mu\text{g/ml}$), ceftazidime (4–16 $\mu\text{g/ml}$), chloramphenicol (2–16 $\mu\text{g/ml}$) and even vancomycin (0.5–2 $\mu\text{g/ml}$), interesting results considering its resistant strains were obtained [46].

Regarding resistant strains, MIC values of crizotinib against MRSA (ATCC 43300; 14.07 $\mu\text{g/ml}$, after unity conversion), showed an intermediate result in comparison with those shown by Baltch *et al.* [69], with MIC values of daptomycin (0.5 $\mu\text{g/ml}$), vancomycin (2 $\mu\text{g/ml}$), oxacillin (16 $\mu\text{g/ml}$) and gentamicin (128 $\mu\text{g/ml}$). Against clinical strains, crizotinib (56.29–7.03 $\mu\text{g/ml}$) and curcumin (46.05–23.02 $\mu\text{g/ml}$) also showed equivalent or even lower MIC values of some antimicrobials of clinical use, such as gentamicin (≥ 2 $\mu\text{g/ml}$), levofloxacin (≥ 4 $\mu\text{g/ml}$), chloramphenicol (≥ 8 $\mu\text{g/ml}$), cefoxitin (≥ 32 $\mu\text{g/ml}$), oxacillin (≥ 64 $\mu\text{g/ml}$) and clindamycin (≥ 128 $\mu\text{g/ml}$) [45].

Gunes *et al.* have also shown that curcumin have antibacterial activity against some of the strains used in this work, with MIC values of 219 $\mu\text{g/ml}$ to *S. aureus* (ATCC 29213), 217 $\mu\text{g/ml}$ to MRSA (ATCC 43300) and 163 $\mu\text{g/ml}$ to *E. coli* (ATCC 25922) [70]. These values are higher as compared with the other active compounds, which might also be due to its poor solubility in aqueous media and the low bioavailability of curcumin [71]. In this case, curcumin analog synthesis may contribute to more selective and active compounds and also to revert its high cytotoxicity (10.82 \pm 1.52). As observed in other work from our group for compounds **37** and **38** (Supplementary Table 8) with CC₅₀ values of 400.73 and 202.1 respectively (unpublished data), but there was a loss of its antibacterial activity at the concentrations tested.

Considering triclosan, Assadian *et al.* evaluated its antibacterial activity to ATCC strains of *S. aureus* and *E. coli* obtaining MIC values of 0.125 $\mu\text{g/ml}$ to *S. aureus* (ATCC 29213) and 0.5 $\mu\text{g/ml}$ to *E. coli* (ATCC 25922) [72]. In our study, after unity conversion, triclosan showed MIC values of 1.8 (6.25 μM) and 0.22 $\mu\text{g/ml}$ (0.78 μM) to ATCC strains of *S. aureus* and *E. coli*, respectively. To MRSA strains, Al-Doori and co-workers in 2003 and 2017 reported MIC values ranging from 0.015 to 4 $\mu\text{g/ml}$ [73,74], which is somewhat comparable with the present study, which show MIC values ranged from 1.81 (6.25 μM) to 0.11 $\mu\text{g/ml}$ (0.39 μM).

Among the identified compounds with low inhibitory activity (%), curcumin derivatives and feluric acid were previously described with antibacterial activity [56–58]. Along with the other data, it may also indicate the success of

the VS strategy used in this work, as well as identifying new chemical classes with potential inhibitory activity to the SzFabI. Curcumin itself is also related to different action mechanisms, besides membrane damage, such as generating oxygen reactive species and straight damage to DNA [75–77], which could suggest more than a single target to its antibacterial activity, and corroborate popular knowledge and folk medicine of antimicrobial, anti-inflammatory and many other described activities of medicinal plants and its products, such as curcumin [70].

Finally, zeta potential measurements showed a lower negative cell surface charge to crizotinib against *E. coli* [78]. Halder *et al.* suggested that alteration of zeta potential may have a correlation with the enhancement of membrane permeability and its polydispersity index [79]. Although side effects may occur, as unspecific targeting mammalian cell membrane, taking into consideration how essential this process is for bacteria, we could exploit the bactericidal activity concomitantly triggering less drug resistance [80].

The same observation could be related to curcumin, as in different assays and treatment conditions, its different mechanisms could not be successful in inhibiting bacterial growth. Aside from mechanisms cited above, Kaur *et al.* showed that curcumin inhibits the FtsZ protein polymerization (potential target for antimicrobial drug development) in *Bacillus subtilis* and *E. coli* [81]. Groundwater *et al.* also reported that FtsZ cellular target assembly inhibition by curcumin and its synthetic analogs correlated with antibacterial activity, especially to Gram-positive strains, with lower MIC results for a few curcumin derivatives [82]. Moreover, curcumin antibacterial activity, especially at higher concentrations can change the membrane structure and lead to cell death without evidence or changes to its surface charge potential [76]. Altogether, those data could corroborate and possibly explain our zeta potential results, in which, considering this alternative target (FtsZ) and time needed to the inhibition process, no membrane disturbance or disruption was shown at early stages.

Therefore, establishing a positive relationship between altered zeta potential (membrane disruption) and decreased cell viability may not always be veracious, as membrane surface charge alterations may not lead to cell death, but instead only affect some of its functions at some degree and after a certain time.

Future perspective

CSM can be useful for the design of novel bioactive compounds and the overall approach established can be modified for ligand-based drug-design campaigns. These four selected compounds are active against resistant *S. aureus*, including clinical isolates, although they have low SIs. However, the design, synthesis and biological testing of analogs obtained by molecular modifications can improve toxicity of novel chemicals. Furthermore, synergy effects with other antibacterial/antibiotics could be an interesting strategy since the activity of reported hits (specifically crizotinib) may be related to membrane biosynthesis inhibition.

Executive summary

- We generated three predictive ROCS similarity models and five predictive electrostatic similarity models for antibiotic selection based on the structure of FabI known inhibitors.
- Virtual screening was used to identify four compounds with four new antibacterial lead candidates, and among them the ferulic acid and curcumin derivatives, with reported antimicrobial activity.
- Antibacterial screening results showed the lowest minimum inhibitory concentration values to triclosan and crizotinib to all bacteria tested.
- Zeta potential measurements showed crizotinib with an early and possible direct disturbance and disruption to the external membrane of *Escherichia coli*.

Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at: www.future-science/doi/suppl/10.4155/fmc-2019-0158

Acknowledgments

The authors wish to acknowledge OpenEye Scientific Software for OMEGA, ROCS and QUACPAC academic licenses

Financial & competing interests disclosure

The authors would like to thank Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) grant #456984/2014-3, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Fundação de Amparo à Pesquisa do Estado de São Paulo

(FAPESP) grants #2013/07600-3 (CIBFar-CEPID), #2014/50926-0 (INCT BioNat CNPq/FAPESP) and the Universidade Federal de Minas Gerais for financial support. EG Kroon is a fellow from CNPq. M Valli acknowledges scholarships #2019/05967-3 from FAPESP, #120/2017 from Finatec, and Termo de Execução Descentralizado Arbocontrol #74/2016 for grant support and research fellowship. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

References

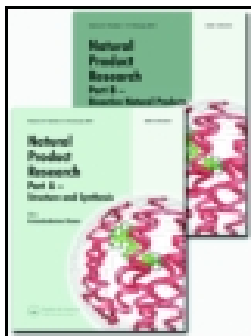
Papers of special note have been highlighted as: ● of interest; ●● of considerable interest

- DeLeo FR, Chambers HF. Reemergence of antibiotic-resistant *Staphylococcus aureus* in the genomics era. *J. Clin. Invest.* 119(9), 2464–2474 (2009).
- Center for Disease Control and Prevention. Methicillin-resistant *Staphylococcus aureus* (MRSA) (2016). www.cdc.gov/mrsa/
- Center for Disease Control and Prevention. US CDC report on antibiotic resistance threats in the United States, 2013 (2013) <http://ecdc.europa.eu/en/news-events/us-cdc-report-antibiotic-resistance-threats-united-states-2013>
- WHO. WHO publishes list of bacteria for which new antibiotics are urgently needed (2017). www.who.int/mediacentre/news/releases/2017/bacteria-antibiotics-needed/en/.
- **WHO report listing the most important bacteria considered priority in the new drugs development.**
- Lu H, Tonge PJ. Inhibitors of FabI, an enzyme drug target in the bacterial fatty acid biosynthesis pathway. *Acc. Chem. Res.* 41(1), 11–20 (2008).
- Janßen H, Steinbüchel A. Fatty acid synthesis in *Escherichia coli* and its applications towards the production of fatty acid based biofuels. *Biotechnol. Biofuels.* 7(1), 7 (2014).
- Kingry LC, Cummings JE, Brookman KW, Bommineni GR, Tonge PJ, Slayden RA. The *Francisella tularensis* FabI enoyl-acyl carrier protein reductase gene is essential to bacterial viability and is expressed during infection. *J. Bacteriol.* 195(2), 351–358 (2013).
- Balemans W, Lounis N, Gilissen R *et al.* Essentiality of FASII pathway for *Staphylococcus aureus*. *Nature* 463(7279), E3–E3 (2010).
- Hafkin B, Kaplan N, Hunt TL. Safety, tolerability and pharmacokinetics of AFN-1252 administered as immediate release tablets in healthy subjects. *Future Microbiol.* 10(11), 1805–1813 (2015).
- Hunt T, Kaplan N, Hafkin B. Safety, tolerability and pharmacokinetics of multiple oral doses of AFN-1252 administered as immediate release (IR) tablets in healthy subjects. *J. Chemother.* 28(3), 164–171 (2016).
- Hafkin B, Kaplan N, Murphy B. Efficacy and safety of AFN-1252, the first *Staphylococcus*-specific antibacterial agent, in the treatment of acute bacterial skin and skin structure infections, including those in patients with significant comorbidities. *Antimicrob. Agents Chemother.* 60(3), 1695–1701 (2016).
- **Describes the clinical trial of AFN-1252, a promising FabI inhibitor to reach the clinical.**
- Christie SM, Ren J, Johnson ME. US20180072666A1 (2018).
- Miller WH, Newlander KA, Seefeld MA. US6762201B1 (2004).
- Tonge P, Pan P. US10071965B2 (2018).
- Takhi M, Hosahalli S, Panigrahi SK. US9062075B2 (2015).
- Miller WH, Seefeld MA, Newlander KA *et al.* Discovery of aminopyridine-based inhibitors of bacterial enoyl-ACP reductase (FabI). *J. Med. Chem.* 45(15), 3246–3256 (2002).
- Seefeld MA, Miller WH, Newlander KA *et al.* Indole naphthyridinones as inhibitors of bacterial enoyl-ACP reductases FabI and FabK. *J. Med. Chem.* 46(9), 1627–1635 (2003).
- Ramnauth J, Surman MD, Sampson PB *et al.* 2,3,4,5-tetrahydro-1H-pyrido[2,3-b and e][1,4]diazepines as inhibitors of the bacterial enoyl ACP reductase, FabI. *Bioorg. Med. Chem. Lett.* 19(18), 5359–5362 (2009).
- Sampson PB, Picard C, Handerson S *et al.* Spiro-naphthyridinone piperidines as inhibitors of *S. aureus* and *E. coli* enoyl-ACP reductase (FabI). *Bioorg. Med. Chem. Lett.* 19(18), 5355–5358 (2009).
- Heerding DA, Chan G, DeWolf WE *et al.* 1,4-Disubstituted imidazoles are potential antibacterial agents functioning as inhibitors of enoyl acyl carrier protein reductase (FabI). *Bioorg. Med. Chem. Lett.* 11(16), 2061–2065 (2001).
- Seefeld MA, Miller WH, Newlander KA *et al.* Inhibitors of bacterial enoyl acyl carrier protein reductase (FabI): 2,9-disubstituted 1,2,3,4-tetrahydropyrido[3,4-b]indoles as potential antibacterial agents. *Bioorg. Med. Chem. Lett.* 11(17), 2241–2244 (2001).
- Takhi M, Sreenivas K, Reddy CK *et al.* Discovery of azetidine based ene-amides as potent bacterial enoyl ACP reductase (FabI) inhibitors. *Eur. J. Med. Chem.* 84, 382–394 (2014).
- Ferreira RS, Glaucius O, Andricopulo AD. [Integration of virtual screening and high-scale automated biological screening techniques: opportunities and challenges in drug R&D]. *Quím. Nova.* 34(10), 1770–1778 (2011).

24. Davis A, Ward SE. The handbook of medicinal chemistry: principles and practice. *Royal Society of Chemistry, Cambridge* 505–528 (2014).
25. Kerns EH, Di L. *Drug-Like Properties: Concepts, Structure Design and Methods*. Academic Press, London (2015).
26. Mysinger MM, Carchia M, Irwin JJ, Shoichet BK. Directory of useful decoys, enhanced (DUD-E): better ligands and decoys for better benchmarking. *J. Med. Chem.* 55(14), 6582–6594 (2012).
27. Irwin JJ, Shoichet BK. ZINC – a free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* 45(1), 177–182 (2005).
28. Ntie-Kang F, Zofou D, Babiaka SB *et al.* AfroDb: a select highly potent and diverse natural product library from African medicinal plants. *PLoS ONE* 8(10), e78085 (2013).
29. Valli M, dos Santos RN, Figueira LD *et al.* Development of a natural products database from the biodiversity of Brazil. *J. Nat. Prod.* 76(3), 439–444 (2013).
30. Pilon AC, Valli M, Dametto AC *et al.* NuBBEDB: an updated database to uncover chemical and biological information from Brazilian biodiversity. *Sci. Rep.* 7(1), 7215 (2017).
31. Chen CY-C. TCM Database@Taiwan: the world's largest traditional Chinese medicine database for drug screening *in silico*. *PLoS ONE* 6(1), e15939 (2011).
32. QUACPAC 2.0.0.3: OpenEye Scientific Software, Santa Fe, NM (2016). www.eyesopen.com
33. Jakalian A, Jack DB, Bayly CI. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comput. Chem.* 23(16), 1623–1641 (2002).
34. OMEGA 2.5.1.4: OpenEye Scientific Software, Santa Fe, NM (2010). www.eyesopen.com
35. Hawkins PCD, Skillman AG, Nicholls A. Comparison of shape-matching and docking as virtual screening tools. *J. Med. Chem.* 50(1), 74–82 (2007).
36. EON 2.3.0.3: OpenEye Scientific Software, Santa Fe, NM (2018). www.eyesopen.com
37. Kim S, Thiessen PA, Bolton EE *et al.* PubChem substance and compound databases. *Nucleic Acids Res.* 44, D1202–D1213 (2016).
38. Yap CW. PaDEL-descriptor: an open source software to calculate molecular descriptors and fingerprints. *J. Comput. Chem.* 32(7), 1466–1474 (2011).
39. Berthold MR, Cebron N, Dill F *et al.* KNIME – the Konstanz information miner: version 2.0 and beyond. *ACM SIGKDD Explor. Newsl.* 11(1), 26 (2009).
40. Jain AN. Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine. *J. Med. Chem.* 46(4), 499–511 (2003).
41. Cheng F, Li W, Zhou Y *et al.* admetSAR: a comprehensive source and free tool for assessment of chemical ADMET Properties. *J. Chem. Inf. Model.* 52(11), 3099–3105 (2012).
42. Drwal MN, Banerjee P, Dunkel M, Wettig MR, Preissner R. ProTox: a web server for the *in silico* prediction of rodent oral toxicity. *Nucleic Acids Res.* 42(W1), W53–W58 (2014).
43. Kronenberger T, Asse LR, Wrenger C, Trossini GHG, Honorio KM, Maltarollo VG. Studies of *Staphylococcus aureus* FabI inhibitors: fragment-based approach based on holographic structure–activity relationship analyses. *Future Med. Chem.* 9(2), 135–151 (2017).
- **A previous *in silico* analysis of known FabI inhibitors that established quantitative structure-relationship analysis and also mapped active site of FabI from *Staphylococcus aureus* and *Escherichia coli* with chemical probes.**
44. Naylor E, Arredouani A, Vasudevan SR *et al.* Identification of a chemical probe for NAADP by virtual screening. *Nat. Chem. Biol.* 5(4), 220–226 (2009).
45. Gomes MFR, Bomfim MRQ, Trindade MJ *et al.* Potential spread of methicillin-resistant *staphylococcus aureus* recovered from patients with bloodstream infection. *Chemother. Open Access* 4, 149 (2015).
46. Clinical and Laboratory Standards Institute. Performance standards for antimicrobial susceptibility testing (2017). https://clsi.org/media/1469/m100s27_sample.pdf
47. Mosmann T. Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. *J. Immunol. Methods* 65(1–2), 55–63 (1983).
48. Gonçalves FDA, de Carvalho CCCR. Phenotypic modifications in *Staphylococcus aureus* cells exposed to high concentrations of vancomycin and teicoplanin. *Front. Microbiol.* 7, 13 (2016).
49. Du Z, Liu R, Shao W *et al.* Alpha-glucosidase inhibition of natural curcuminoids and curcumin analogs. *Eur. J. Med. Chem.* 41(2), 213–218 (2006).
50. Braga SFP, Martins LC, da Silva EB *et al.* Synthesis and biological evaluation of potential inhibitors of the cysteine proteases cruzain and rhodesain designed by molecular simplification. *Bioorg. Med. Chem.* 25(6), 1889–1900 (2017).
51. Lino CI, Gonçalves de Souza I, Borelli BM *et al.* Synthesis, molecular modeling studies and evaluation of antifungal activity of a novel series of thiazole derivatives. *Eur. J. Med. Chem.* 151, 248–260 (2018).

52. Matthews BW. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochim. Biophys. Acta BBA - Protein Struct.* 405(2), 442–451 (1975).
53. Mistry TL, Truong L, Ghosh AK, Johnson ME, Mehboob S. Benzimidazole-based FabI inhibitors: a promising novel scaffold for anti-staphylococcal drug development. *ACS Infect. Dis.* 3(1), 54–61 (2016).
54. Hevener KE, Mehboob S, Su PC *et al.* Discovery of a novel and potent class of *F. tularensis* enoyl-reductase (FabI) inhibitors by molecular shape and electrostatic matching. *J. Med. Chem.* 55(1), 268–279 (2012).
55. Boström J, Grant JA, Fjellström O, Thelin A, Gustafsson D. Potent fibrinolysis inhibitor discovered by shape and electrostatic complementarity to the drug tranexamic acid. *J. Med. Chem.* 56(8), 3273–3280 (2013).
56. Khatkar A, Nanda A, Kumar P, Narasimhan B. Synthesis and antimicrobial evaluation of ferulic acid derivatives. *Res. Chem. Intermed.* 41(1), 299–309 (2015).
- **A previous work describes the activity of ferulic acid and derivatives against bacteria which corroborates our findings.**
57. Nakachi M, Ikemoto S, Yamanishi H *et al.* Antimicrobial activities of synthetic ferulic acid derivatives. *Food Preserv. Sci.* 28(4), 183–188 (2002).
58. Yao J, Zhang Q, Min J, He J, Yu Z. Novel enoyl-ACP reductase (FabI) potential inhibitors of *Escherichia coli* from Chinese medicine monomers. *Bioorg. Med. Chem. Lett.* 20(1), 56–59 (2010).
59. Shave S, McGuire K, Pham NT, Mole DJ, Webster SP, Auer M. Diclofenac identified as a kynurenine 3-monooxygenase binder and inhibitor by molecular similarity techniques. *ACS Omega* 3(3), 2564–2568 (2018).
60. Fradera X, Babaoglu K. Overview of methods and strategies for conducting virtual small molecule screening. *Curr. Protoc. Chem. Biol.* 9(3), 196–212 (2017).
- **A review of virtual screening methods, protocols, its differences, advantages and disadvantages.**
61. Lima AN, Philot EA, Trossini GHG, Scott LPB, Maltarollo VG, Honorio KM. Use of machine learning approaches for novel drug discovery. *Expert Opin. Drug Discov.* 11(3), 225–239 (2016).
62. Meissner KA, Kronenberger T, Maltarollo VG, Trossini GHG, Wrenger C. Targeting the *Plasmodium falciparum* plasmepsin V by ligand-based virtual screening. *Chem. Biol. Drug Des.* 93(3), 300–312 (2019).
63. Serafim MSM, Lavorato SN, Kronenberger T *et al.* Antibacterial activity of synthetic 1,3-bis(aryloxy)propan-2-amines against Gram-positive bacteria. *MicrobiologyOpen.* e814 (2019).
64. Kirchmair J, Distinto S, Markt P *et al.* How to optimize shape-based virtual screening: choosing the right query and including chemical information. *J. Chem. Inf. Model.* 49(3), 678–692 (2009).
- **A very useful description and comparison of several chemical similarity models by applying weights to chemical features and its impact on models predictivity.**
65. Hamza A, Wei NN, Zhan CG. Ligand-based virtual screening approach using a new scoring function. *J. Chem. Inf. Model.* 52(4), 963–974 (2012).
66. Tuberculosis Drug Screening Program. Search for new drugs for treatment of tuberculosis. *Antimicrob. Agents Chemother.* 45(7), 1943–1946 (2001).
67. Shaw AT, Kim DW, Nakagawa K *et al.* Crizotinib versus chemotherapy in advanced ALK-positive lung cancer. *N. Engl. J. Med.* 368(25), 2385–2394 (2013).
68. Kazandjian D, Blumenthal GM, Chen HY *et al.* FDA approval summary: crizotinib for the treatment of metastatic non-small cell lung cancer with anaplastic lymphoma kinase rearrangements. *Oncologist* 19(10), e5–e11 (2014).
69. Baltch AL, Ritz WJ, Bopp LH, Michelsen PB, Smith RP. Antimicrobial activities of daptomycin, vancomycin, and oxacillin in human monocytes and of daptomycin in combination with gentamicin and/or rifampin in human monocytes and in broth against *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* 51(4), 1559–1562 (2007).
70. Gunes H, Gulen D, Mutlu R, Gumus A, Tas T, Topkaya AE. Antibacterial effects of curcumin: an *in vitro* minimum inhibitory concentration study. *Toxicol. Ind. Health* 32(2), 246–250 (2016).
71. Anand P, Kunnumakkara AB, Newman RA, Aggarwal BB. Bioavailability of curcumin: problems and promises. *Mol. Pharm.* 4(6), 807–818 (2007).
72. Assadian O, Wehse K, Hübner N-O *et al.* Minimum inhibitory (MIC) and minimum microbicidal concentration (MMC) of polihexanide and triclosan against antibiotic sensitive and resistant *Staphylococcus aureus* and *Escherichia coli* strains. *GMS Krankenhaushygiene Interdisziplin.* 6(1), Doc06 (2011).
73. Al-Doori Z, Morrison D, Philpott-Howard J. Small colony variants and triclosan resistance in five international clones of methicillin resistant *Staphylococcus aureus*. *J. Mol. Biol. Res.* 7(1), 112 (2017).
74. Al-Doori Z, Morrison D, Edwards G, Gemmell C. Susceptibility of MRSA to triclosan. *J. Antimicrob. Chemother.* 51(1), 185–186 (2003).
75. Rai D, Singh JK, Roy N, Panda D. Curcumin inhibits FtsZ assembly: an attractive mechanism for its antibacterial activity. *Biochem. J.* 410(1), 147–155 (2008).

76. Tyagi P, Singh M, Kumari H, Kumari A, Mukhopadhyay K. Bactericidal activity of curcumin I is associated with damaging of bacterial membrane. *PLoS ONE* 10(3), e0121313 (2015).
77. Yun DG, Lee DG. Antibacterial activity of curcumin via apoptosis-like response in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* 100(12), 5505–5514 (2016).
78. Sahu A, Prabhaskar K, Noronha V, Joshi A, Desai S. Crizotinib: a comprehensive review. *South Asian J. Cancer.* 2(2), 91 (2013).
79. Halder S, Yadav KK, Sarkar R *et al.* Alteration of Zeta potential and membrane permeability in bacteria: a study with cationic agents. *SpringerPlus* 4, 672 (2015).
80. Hurdle JG, O'Neill AJ, Chopra I, Lee RE. Targeting bacterial membrane function: an underexploited mechanism for treating persistent infections. *Nat. Rev. Microbiol.* 9(1), 62–75 (2011).
81. Kaur S, Modi NH, Panda D, Roy N. Probing the binding site of curcumin in *Escherichia coli* and *Bacillus subtilis* FtsZ—a structural insight to unveil antibacterial activity of curcumin. *Eur. J. Med. Chem.* 45(9), 4209–4214 (2010).
82. Groundwater PW, Narlawar R, Liao VWY *et al.* A carbocyclic curcumin inhibits proliferation of Gram-positive bacteria by targeting FtsZ. *Biochemistry* 56(3), 514–524 (2017).



Natural Product Research

Formerly Natural Product Letters

ISSN: 1478-6419 (Print) 1478-6427 (Online) Journal homepage: <https://www.tandfonline.com/loi/gnpl20>


Himatanthus bracteatus stem extracts present anti-flavivirus activity while an isolated sesquiterpene glucoside present only anti-Zika virus activity *in vitro*

Marlene Lourenço da Silva, João Renato Stehmann, Mateus Sá Magalhães Serafim, Valdicley Vieira Vale, Douglas Costa Gontijo, Geraldo Célio Brandão, Erna Geessien Kroon & Alaíde Braga de Oliveira



To cite this article: Marlene Lourenço da Silva, João Renato Stehmann, Mateus Sá Magalhães Serafim, Valdicley Vieira Vale, Douglas Costa Gontijo, Geraldo Célio Brandão, Erna Geessien Kroon & Alaíde Braga de Oliveira (2019): *Himatanthus bracteatus* stem extracts present anti-flavivirus activity while an isolated sesquiterpene glucoside present only anti-Zika virus activity *in vitro*, Natural Product Research, DOI: [10.1080/14786419.2019.1690487](https://doi.org/10.1080/14786419.2019.1690487)

To link to this article: <https://doi.org/10.1080/14786419.2019.1690487>

 View supplementary material 

 Published online: 20 Nov 2019.

 Submit your article to this journal 




 View related articles 

 View Crossmark data 

SHORT COMMUNICATION



Himatanthus bracteatus stem extracts present anti-flavivirus activity while an isolated sesquiterpene glucoside present only anti-Zika virus activity *in vitro*

Marlene Lourenço da Silva^a, João Renato Stehmann^b, Mateus Sá Magalhães Serafim^c, Valdicley Vieira Vale^d, Douglas Costa Gontijo^a , Geraldo Célio Brandão^e , Erna Geessien Kroon^c and Alaíde Braga de Oliveira^a 

^aDepartamento de Produtos Farmacêuticos, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil; ^bDepartamento de Botânica, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil; ^cDepartamento de Microbiologia, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil; ^dPrograma de Pós-Graduação em Ciências Farmacêuticas, Universidade Federal do Pará, Belém, Brazil; ^eEscola de Farmácia, Universidade Federal de Ouro Preto, Ouro Preto, Brazil

ABSTRACT

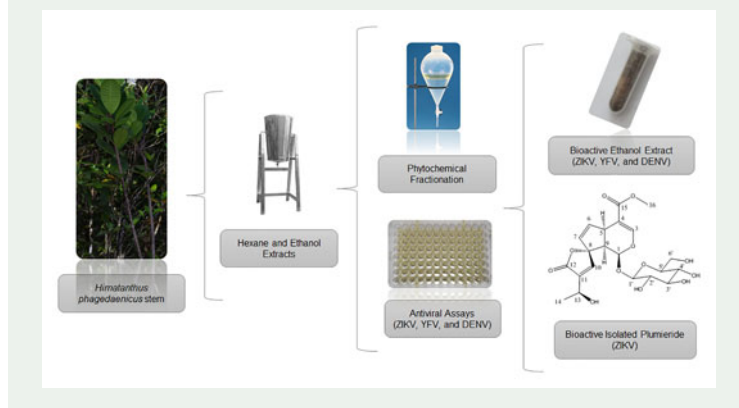
The hexane and ethanol extracts from *Himatanthus bracteatus* (Apocynaceae) stems were evaluated for antiviral activity against Zika virus, yellow fever virus and dengue virus 2 and for cytotoxicity in Vero cells by MTT assay. The ethanol extract showed good antiviral activity against the three viruses with selective indexes (SI) > 10 and its fractionation led to the isolation of the known plumieride that was active only against Zika virus (SI of 15.97).

ARTICLE HISTORY

Received 23 August 2019
Accepted 3 November 2019

KEYWORDS


Himatanthus bracteatus; Zika virus; yellow fever virus; dengue virus; plumieride; spectroscopic analyses



1. Introduction

Emerging and reemerging viral diseases play a significant role in human diseases and have a prominent importance in public health (Vasconcelos and Calisher 2016; Hotez and Murray 2017). Zika, yellow fever, and chikungunya viruses among other

CONTACT Douglas Costa Gontijo  gontijo.douglas@outlook.com; Alaíde Braga de Oliveira  bragalaide@gmail.com

 Supplemental data for this article can be accessed at <https://doi.org/10.1080/14786419.2019.1690487>.

© 2019 Informa UK Limited, trading as Taylor & Francis Group

viral diseases are classified as epidemic and/or pandemic by the World Health Organization (WHO 2019a). No specific drugs are currently available for the treatment of these diseases. The recognized activity of natural products against several viruses (Gandhi et al. 2016; Ferreira et al. 2018; Segun et al. 2019) stimulates the research in this area.

Following up our results on the screening of several Brazilian plant species for anti-viral activity that disclosed the activity of a specimen that was initially identified as *H. phagedaenicus* (Brandão, 2010), further investigation of extracts, fractions and a compound isolated from the plant stems were evaluated against three flaviviruses, Zika virus (ZIKV), yellow fever virus (YFV) and dengue virus (DENV-2). *Himatanthus bracteatus* (A. DC.) Woodson (Apocynaceae) which is endemic to the Atlantic Forest, with a restricted distribution to nine states in the northeastern and southeastern regions of Brazil and is popularly known the janauba, angelica of the forest and parrot banana (Pugas et al., 2018). “Janauba” has a broad spectrum of uses including from medicinal, in the treatment of inflammations, veterinary and food supplement, to latex extractivism, which has been acquiring growing importance in the market (Linhares, 2010).

2. Results and discussion

Extracts were obtained from stems of *Himatanthus bracteatus* affording 5.9 g of the dried hexane extract (HBE HEX, yield 4.1%) and 32.8 g of the dried ethanol extract (HBE EtOH, yield 22.9%). No antiviral activity against the three viruses evaluated was observed for the HBE HEX that was moderately cytotoxic against Vero cells, showing CC_{50} values of $50.36 \pm 1.49 \mu\text{g/mL}$ (Table 1). Low polarity extracts from *Himatanthus* species are reported as terpene-rich (Veloso et al. 1999). On the other hand, the ethanol extract (HBE EtOH) showed lower cytotoxicity on Vero cells (CC_{50} of $348.02 \pm 5.75 \mu\text{g/mL}$) and significant activity against the ZIKV, YFV and DENV-2 (EC_{50} of 31.46 ± 0.72 , 17.43 ± 3.90 and 25.03 ± 1.20 , respectively) with selectivity indexes > 10 , highlighting the higher value for YFV (SI of 19.07) (Table 1). Yellow fever has recently reemerged in Brazil with various seasonal periods (Marcondes and Ximenes 2016;

Table 1. Antiviral activity of the extracts, fractions, and compound isolated from *Himatanthus bracteatus* stems.

Samples	Vero cells CC_{50} ($\mu\text{g/mL}$)	ZIKV (PE243)		YFV (17DD)		DENV-2 (PI59)	
		EC_{50} ($\mu\text{g/mL}$)	SI	EC_{50} ($\mu\text{g/mL}$)	SI	EC_{50} ($\mu\text{g/mL}$)	SI
HBE HEX	50.36 ± 1.49	NA	–	NA	–	NA	–
HBE EtOH	348.02 ± 5.75	31.46 ± 0.72	11.06 ^b	17.43 ± 3.90	19.97 ^a	25.03 ± 1.20	13.90 ^a
HBF CHCl_3	>100.00	NA	–	NA	–	NA	–
HBF H_2O	288.33 ± 23.70	74.37 ± 1.70	3.87 ^c	93.20 ± 1.11	3.09 ^b	82.80 ± 2.60	3.48 ^c
HBI	231.42 ± 37.58	14.49 ± 0.72	15.97 ^a	NA	–	NA	–
Ribavirin	49.00 ± 2.45	3.06 ± 0.25	16.01 ^a	10.02 ± 1.28	4.89 ^b	NA	–
MPA	32.00 ± 6.40	NA	–	NA	–	3.38 ± 0.30	9.47 ^b

Note: HBE: *Himatanthus bracteatus* Extract; HBF: *Himatanthus bracteatus* Fraction; HBI: *Himatanthus bracteatus* Isolated; HEX: Hexane; EtOH: Ethanol; CHCl_3 : chloroform; H_2O : water; MPA: Mycophenolic acid; CC_{50} : 50% Cytotoxic Concentration; EC_{50} : 50% Effective Concentration; SI: Selective Index; NA: Not active. Means (\pm SD, $n = 3$) followed by different letters were significantly different within SI columns (Tukey test's, $\alpha = 0.05$). Statistical analyses of the SI were performed separately.

a, b and c represent statistical difference between the results analyzed, with no corresponding indicators.

Mota et al. 2016; WHO 2019b). Interestingly, in the previous screening, good activity was determined against DENV-2 (EC_{50} of $3.8 \pm 0.6 \mu\text{g/mL}$, SI of 58) (Brandão 2010) while a lower effect was observed in the present assays (EC_{50} of $25.03 \pm 1.20 \mu\text{g/mL}$, SI of 13.90). These differences in results could be due to the use of different virus strains in these experiments. In addition, significant methodological differences are observed in the assays reported here and in the analyzes by Brandão (2010).

Aiming to identify bioactive compounds, HBE EtOH was fractionated by liquid-liquid partition, affording two fractions, HBF H₂O with a decrease of the antiviral activity against ZIKV, YFV and DENV-2 (SI of 3.87, 3.09 and 3.48, respectively), and HBF CHCl₃, which was inactive against all three viruses (Table 1). Interestingly, SI for the HBE EtOH was significantly higher than those observed for ribavirin and MPA controls against the different viruses analyzed. Indeed, despite the lower antiviral activity observed, HBE EtOH showed lower toxicity to Vero cells than the controls used.

The Total Ion Chromatogram (TIC, positive full scan) of HBE EtOH, HBF H₂O and HBF CHCl₃, registered by UPLC-DAD-ESI-MS, showed high similarity between HBE EtOH and HBF H₂O samples, what corroborates the antiviral results (Supplementary material, Figure S1). Furthermore, the different samples showed a major constituent peak with retention time (Rt) of approximately 2.22 min that may be a phytochemical marker for *H. bracteatus* (Supplementary material, Figure S1). HBE HEX and HBF CHCl₃, both of them rich low polarity compounds, were inactive. Therefore, the activity of HBE EtOH and HBF H₂O samples against the three evaluated flavivirus may be attributed to compounds that are present in both of them but are absent or in very low concentration in HBF CHCl₃. Furthermore, an intriguing fact related to the inactivity of HBF CHCl₃ to all the flavivirus assayed although affording the active plumieride after a column chromatography fractionation, may be explained by its profile in the UPLC-DAD-ESI-MS chromatogram (Supplementary material, Figure S1) that shows the presence of peaks corresponding to lower polarity compounds with RTs in the range of 2.5–6.0 min (Supplementary material, Figure S1).

Despite the inactivity of the HBF CHCl₃ fraction, successive chromatographic fractionations on silica gel columns led to the isolation a colourless compound (HBI, *Himatanthus bracteatus* Isolated) that was active only against ZIKV (EC_{50} of 14.49 ± 0.72) disclosing a SI increase to 15.97 compared to HBE EtOH lower activity to this virus (EC_{50} of 31.46 ± 0.72) with a SI of 11.06 (Table 1). The TIC (positive scan) showed HBI as major constituent of the HBE EtOH, HBF CHCl₃ and HBE H₂O of *H. bracteatus* stems (Supplementary material, Figure S1). Despite showing activity only against ZIKV, HBI showed SI similar to the ribavirin control (Table 1).

Physical and spectroscopic data (IR, UV, ¹H and ¹³CNMR, UPLC-DAD-ESI-MS/MS) (Supplementary Material) allowed the identification of HBI as plumieride (Supplementary material, Figure S2), a sesquiterpene glucoside that is reported for different *Himatanthus* species (Vale et al. 2015; Almeida et al. 2017). Plumieride was previously reported for *H. phagedaenicus* but it was isolated as its penta-*O*-acetyl derivative in a mixture with octa-*O*-acetylsucrose obtained after acetylation of an aqueous fraction derived from stem ethanol extract (Vanderlei et al. 1991). Therefore, this is the first report on the isolation of plumieride from *H. bracteatus* showing antiviral activity. Given the high yield of plumieride the *H. bracteatus* species may be an

important source for these bioactive iridoid. In this sense, we reported for the first time with antiviral activity against Zika virus. In addition, due its significant activity against ZIKV, semi-synthetic derivatives will be prepared for investigation of structure-activity relationships, with further enrichment of antiviral activity assays (e.g. pre-treatment of cells and post-penetration of viruses), determining its effects in viral cycle.

3. Experimental section

See [supplementary material](#).

4. Conclusion

Previous report on the *Himatanthus bracteatus* ethanol stem extract activity against DENV-2 was extended to ZIKV and YFV. Plumieride was isolated, fully characterized, and its anti-ZIKV activity was firstly reported. Further investigation of *H. bracteatus* is still of interest aiming to identify the anti-DENV-2 and YFV active natural products.

Disclosure statement

The authors confirm that this article content has no conflict of interest.

Funding

To FAPEMIG Process APQ-01410-14 for financial support to ABO. To CNPq Process 440911/2016, Capes Process 2058/2016 and DECIT/MS for financial support to EGK. To CNPq for scholarship to MLS and fellowship to ABO.

ORCID

Douglas Costa Gontijo  <http://orcid.org/0000-0002-7759-9694>

Geraldo Célio Brandão  <https://orcid.org/0000-0002-7330-051X>

Alaide Braga de Oliveira  <https://orcid.org/0000-0002-0374-8909>

References

- Almeida SCX, Monteiro AB, Costa GM, Viana GSB. 2017. *Himatanthus drasticus*: a chemical and pharmacological review of this medicinal species, commonly found in the Brazilian Northeastern region. *Braz J Pharmacogn.* 27(6):788–793.
- Brandão GC. 2010. Fitoquímica, triagem antiviral e fracionamento biomonitorado de espécies da biodiversidade Brasileira Tese de Doutorado. Universidade Federal de Minas Gerais. Programa de Pós-graduação em Ciências Farmacêuticas, 324 p.
- Ferreira PG, Ferraz AC, Figueiredo JE, Lima CF, Rodrigues VG, Taranto AG, Ferreira JMS, Brandão GC, Vieira-Filho AS, Duarte LP, Magalhães CLB, et al. 2018. Detection of the antiviral activity of epicatechin isolated from *Salacia crassifolia* (Celastraceae) against Mayaro virus based on protein C homology modelling and virtual screening. *Arch Virol.* 163(6):1567–1576.,
- Gandhi GR, Barreto PG, Lima BS, Quintans JSS, Araújo AAS, Narain N, Quintans-Júnior LJ, Gurgel RQ. 2016. Medicinal plants and natural molecules with *in vitro* and *in vivo* activity against rotavirus: a systematic review. *Phytomedicine* 23(14):1830–1842.

- Hotez PJ, Murray KO. 2017. Dengue, West Nile virus, chikungunya, zika and now Mayaro? *Plos Negl Trop Dis*. 11(8):e0005462.
- Linhares JFPL, 2010. Sustentabilidade sócio-ambiental da extração de janaúba (*Himatanthus Willd. ex Schult.*) no município de Alcântara, MA, Brasil. Dissertação de Mestrado. Universidade Federal do Maranhão, São Luís.
- Marcondes CB, Ximenes M. 2016. Zika virus in Brazil and the danger of infestation by *Aedes (Stegomyia)* mosquitoes. *Rev Soc Bras Med Trop*. 49(1):4–10.
- Mota MTO, Terzian AC, Silva M, Estofolete C, Nogueira ML. 2016. Mosquito-transmitted viroses—the great Brazilian challenge. *Braz J Microbiol*. 47:38–50.
- Pugas RM, Paz JRL, Edinaldo Luz das Neves EL, Pigozzo CM, 2018. *Himatanthus bracteatus* (A. DC.) Woodson (Apocynaceae): phenology, morphology, and floral biology of a species endemic to the Atlantic Forest. *Hoehnea* 45:405–412.
- Segun PA, Ogbole OO, Akinleye TE, Faleye TO, Adeniji AJ. 2019. *In vitro* anti-enteroviral activity of stilbenoids isolated from leaves of *Macaranga barteri*. *Nat Prod Res*. 1–5.
- Vale VV, Vilhena TC, Trindade RCS, Ferreira MRC, Percário S, Soares LF, Pereira WLA, Brandão GC, Oliveira AB, Dolabela MF, Vasconcelos F. 2015. Anti-malarial activity and toxicity assessment of *Himatanthus articulatus*, a plant used to treat malaria in the Brazilian Amazonia. *Malar J*. 14(1):132.
- Vanderlei MF, Silva MS, Gottlieb HE, Braz-Filho R. 1991. Iridoids and triterpenes from *Himatanthus phagedaenicus*: the complete assignment of the ¹H and ¹³C NMR spectra of two iridoid glycosides. *J Braz Chem Soc*. 2(2):51–55.
- Vasconcelos PFC, Calisher CH. 2016. Emergence of human arboviral diseases in the Americas, 2000–2016. *Vector Borne Zoonotic Dis*. 16(5):295–301.
- Veloso MP, Nagem TJ, Oliveira TT. 1999. β-dihidroplumericinic acid from *Himatanthus phagedaenicus*. *Biochem Syst Ecol*. 27(6):669–671.
- WHO. 2019a. World Health Organization; [accessed 2019 Aug 2]. <https://www.who.int/emergencies/diseases/en/>.
- WHO. 2019b. World Health Organization; [accessed 2019 Aug 13]. <https://www.who.int/csr/don/18-april-2019-yellow-fever-brazil/en/>.

Supplementary Material

Himatanthus phagedaenicus stem extracts present anti-flavivirus activity while an isolated sesquiterpene glucoside present only anti-Zika virus activity *in vitro*

Marlene Lourenço da Silva^a, João Renato Stehmann^b, Mateus Sá Magalhães Serafim^c, Valdicley Vieira Vale^d, Douglas Costa Gontijo^a, Geraldo Célio Brandão^e, Erna Geessien Kroon^c, Alaíde Braga de Oliveira^{a*}

^aDepartamento de Produtos Farmacêuticos, Universidade Federal de Minas Gerais, Av. Antônio Carlos, 6627, 31270-901, Belo Horizonte, Brazil.

^bDepartamento de Botânica, Universidade Federal de Minas Gerais, Av. Antônio Carlos, 6627, 31270-901, Belo Horizonte, Brazil.

^cDepartamento de Microbiologia, Universidade Federal de Minas Gerais, Av. Antônio Carlos, 6627, 31270-901, Belo Horizonte, Brazil.

^dPrograma de Pós-graduação em Ciências Farmacêuticas, Universidade Federal do Pará, Rua Augusto Corrêa, 1, 68075-110, Belém, Brazil.

^eEscola de Farmácia, Universidade Federal de Ouro Preto, Morro do Cruzeiro, s/n, 35400-000, Ouro Preto, Brazil.

*Corresponding author: Tel.: +55 31 34096950. E-mail address: bragalaide@gmail.com (AB Oliveira).

Abstract

The hexane and ethanol extracts from *Himatanthus phagedaenicus* (Apocynaceae) stems were evaluated for antiviral activity against Zika virus, yellow fever virus and dengue virus 2 and for cytotoxicity in Vero cells by MTT assay. The ethanol extract showed good antiviral activity against the three viruses with selective indexes (SI) > 10 and its fractionation led to the isolation of the known plumieride that was active only against Zika virus (SI of 15.97).

Keywords: *Himatanthus phagedaenicus*; Zika virus; Yellow Fever virus; Dengue virus; Plumieride; Spectroscopic analyses.

Experimental Section

Plant material

Plant samples were collected in the city of Nova Lima, state of Minas Gerais, Brazil. Voucher specimens were deposited in the Herbário BHCB, Universidade Federal de Minas Gerais, under the code BHCB 112116.

Preparation of extracts

After drying in a circulating air oven at 37°C, for 48 h, dried stems were milled and the powder was kept in dark flasks. Stems powder (143 g) was submitted to successive extractions with hexane (HEX) and then with ethanol (EtOH), by percolation at room temperature. The extracts were concentrated in a rotary evaporator and kept in a glass desiccator under vacuum.

Fractionation of Himatanthus phagedaenicus stems ethanol extract (HPE EtOH)

An aliquot of HPE EtOH (30.0 g) was dissolved in a mixture of methanol/water (6/4), and was extracted with chloroform (3X), that resulted in the fractions HPF CHCl₃ (7.4 g, yield 24.5%), and HPF H₂O (20.8 g, yield 69.2%). Subsequently, from the HPF CHCl₃ fraction (6.0 g), after chromatography on a silica gel column (200.0 g), eluted with solvents of increasing polarities, gave 14 different fractions of which F2 (3.0 g) was rechromatographed in silica gel column, obtaining 9 fractions that were

combined according to their TLC profiles, from group 4, *Himatanthus phagedaenicus* Isolated (HPI, 0.5 g) crystallized out.

Phytochemical analyses

Thin layer chromatography (TLC)

Different secondary metabolites classes in the extract and chromatographic fractions were detected by TLC (silica gel F₂₅₄). Mobile phases and visualization spray reagents were used according to Wagner et al. (1984), in comparison with reference compounds.

UPLC-DAD, UPLC-DAD-ESI-MS and UPLC-ESI-MS/MS

Analyses were performed in an ACQUITY UPLC H-Class Core System (Waters) with an ACQUITY UPLC™ PDA DETECTOR in the Laboratório de Fitoquímica, Faculdade de Farmácia, UFMG. The mobile phase consisted of water 0.1% formic acid (solvent A) and acetonitrile 0.1% formic acid (solvent B). The elution protocol was 0-13 min, linear gradient from 5% to 95% B. The flow rate was 0.3 mL min⁻¹, and the sample injection volume was 4.0 µL. The UV spectra were registered from 200 to 400 nm. Mass spectrometry analyses by electron spray ionization (UPLC-PDA-ESI-MS) were performed in an ion trap mass spectrometer equipped with an atmospheric pressure ionization (API) interface operated in the following conditions: positive and negative ion mode; capillary voltage, 3500 V; capillary temperature, 320°C; source voltage, 5 kV; vaporizer temperature, 320°C; corona needle current, 5 mA; and sheath gas, nitrogen, 27 psi. Analyses were run in the full scan mode (100–2000 Da). The ESI-MS/MS analyses were performed in an UPLC Acquity (Waters) equipped with a Tandem Quadrupole Detector (TQD), fitted with an electrospray source in the positive and negative mode; ion spray voltage: -4 kV; orifice voltage: -60 V, with helium as the collision gas; collision energy was set at 30 eV. An ACQUITY UPLC BEH C-18 column (1.7 µm, 50 × 2 mm i.d.) (Waters) was used.

Infrared (IR) Spectroscopy and Melting point

Infrared spectra were recorded on a Spectra One PerkinElmer spectrophotometer, ted with a Paragon ATR accessory and melting point of compound was determined on Microquímica MOAs 301 apparatus.

Nuclear Magnetic Resonance (NMR) Spectroscopy

NMR spectra were registered on a Bruker AvanceIII 200 MHz using TopSpin™ 3.2 pl 6.0 software for acquisition and processing, in the *LAREMAR* at Universidade Federal de Minas Gerais – UFMG. The sample was dissolved in CD₃OD and the chemical shifts (δ , in ppm) are reported relative to tetramethylsilane (TMS). Coupling constants (J) are given in Hz.

Physical and spectroscopic data on plumieride. Colourless crystals (MeOH); mp 227.9 – 230.0 °C, Lit. mp 227 – 229 °C (Singh et al., 2014). IR ν_{\max} 3367, 1754, 1693 and 1633 cm⁻¹; ESI-MS m/z : 469.00 [M – H]⁻ and 471.36 [M + H]⁺ (C₂₁H₂₆O₁₂) (24.0%); ESI-MS/MS m/z : 231.39 [(M – C₁₃H₁₁O₄)]⁺ (26.3%), 291.24 [(M – C₁₅H₁₅O₆)]⁺ (100.0%), 309.27 [(M – C₁₅H₁₇O₇)]⁺ (95.0%); ¹H NMR (200 MHz, CD₃OD): 1.42 (d, J 6.0 Hz, CH₃-14), 2.96 (m, H-2'), 3.20 (dd, J 8.0 and 4.0 Hz, H-9), 3.40 (m, H-6'), 3.50 (m, H-3'), 3.67 (s, CH₃-16), 3.80 (m, H-5'), 3.89 (d, J 12.0 Hz, H-5), 4.44 (m, H-13), 4.50 (dd, H-4'), 4.70 (d, H-1'), 5.27 (d, J 6.0 Hz, H-1), 5.50 (dd, J 6.0 and 4.0 Hz, H-7), 6.40 (dd, J 2.0 and 4.0 Hz, H-6), 7.37 (s, H-3), 7.51 (s, H-10). ¹³C NMR (50 MHz, CD₃OD): 22.50 (C-14), 40.46 (C-5), 50.60 (C-9), 52.10 (C-16), 61.00 (C-6'), 63.60 (C-13), 71.30 (C-4'), 74.70 (C-2'), 77.80 (C-3'), 78.90 (C-5'), 94.20 (C-1), 97.90 (C-8), 100.10 (C-1'), 111.10 (C-4), 130.00 (C-7), 138.60 (C-11), 141.50 (C-6), 150.30 (C-10), 152.60 (C-3), 168.50 (C-15), 172.80 (C-12).

Biological assays

Cell lines

Vero cells (ATCC number CCL-81) were maintained in minimal essential medium (MEM, Cultilab, Brazil). Culture media were supplemented with 5% fetal bovine serum (FBS, Cultilab, Brazil), 200 U/mL of penicillin (ThermoFischer Scientific, USA), 100 µg/mL of streptomycin (ThermoFischer Scientific, USA) and 2.5 µg/mL of amphotericin B (Merck, Germany).

Viruses

Viruses were obtained from the virus collection of the Laboratório de Vírus from Universidade Federal de Minas Gerais (UFMG). Zika virus (PE243) was isolated in Recife, Brazil, from a male patient in 2015, at the Fundação

Oswaldo Cruz (FIOCRUZ) as described before (Donald et al., 2016). Dengue virus 2 (PI59), lineage II, was identified during an outbreak in Piauí, Brazil, as described before (Figueiredo et al., 2014). Vaccine of yellow fever virus (17DD) was kindly provided by Bio-Manguinhos (FIOCRUZ N. 980FB066Z). All viruses were replicated and titrated in Vero cells.

Viral replication

Culture cell flasks of 75 cm² containing Vero cells were incubated with MEM 5% FBS at 37 °C and 5% CO₂ atmosphere. After 24h, with approximately 60 to 80% confluent monolayer the medium was removed and the cells were washed three times with phosphate-buffered saline (PBS) to remove cellular debris and serum residue. Cells of a control flask were counted and cells were infected at a Multiplicity Of Infection (MOI) of 0.01 of each virus in 2 mL MEM, and flasks were gently homogenized every 10 minutes during an 1h incubation at 37 °C and with 5% CO₂. Following adsorption period of 1 hour, 12 mL of fresh MEM with 2% FBS were added to the flasks and incubated in the same conditions, checked daily to evaluate the cytopathic effects. In monolayers with 80% cytopathic effect, the supernatant was removed, centrifuged at 4 °C to remove debris and stored at -70 °C.

Viral titration

Vero cells were seeded in 6-well plates (2.10⁵ cells per well) and incubated at 37 °C and 5% CO₂ atmosphere. After 24h, plates were infected with 300 µL of a serial dilution (1:10) of each virus (10⁻¹ a 10⁻⁷). After 1h adsorption at the same conditions, 2.0 mL of fresh M199 (Gibco, USA) medium with antibiotics, 2% FBS and 1% of carboxymethylcellulose (CMC) (Synth, Brazil) were added to each well. After five (ZIKV) or seven days (DENV-2 and YFV) of incubation at 37 °C and 5% CO₂ atmosphere, plates were fixed with 10% formalin solution overnight, followed by coloration with 1% violet crystal solution for 15 minutes. Plates were washed and set to dry, counting in each well 30 to 100 plate forming units per millilitres (PFU/mL) to determinate the virus titer (Coelho et al. 2017).

Citotoxicity assay

Cytotoxicity was assessed using the MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) reduction assay (Mosmann, 1983). Vero cells were seeded in 96-well plates ($4 \cdot 10^4$ cells per well) and incubated at 37 °C and 5% CO₂ atmosphere. After 24 hours of incubation, 200 µL of fresh MEM medium (1% FBS) containing serial dilution of extracts (from 500 to 1.95 µg/mL) were added to the plates. A serial dilution of DMSO was used as viability control, as well as an inhibition control (10% DMSO solution) and a blank (medium only). After 72 hours of incubation at 37 °C and 5% CO₂ atmosphere, medium was removed and 100 µL of MTT tetrazolium (ThermoFischer Scientific, USA) solution in MEM (0,5 mg/mL) were added to each well. After 3 hours of incubation at 37 °C and 5% CO₂ atmosphere, the medium was removed and 100 µL of DMSO was added to each well to solubilize formazan crystals. After shaking for 15 minutes, absorbance at 570 nm of each well was read using a spectrophotometer (VersaMax, Molecular Devices, USA). Cytotoxic concentration for 50% of the cells (CC₅₀) of each extract was calculated as a percent of cell viability as follow:

$$\frac{(ODe) - (ODb)}{(ODc) - (ODb)} \times 100(\%)$$

ODe, ODc and ODb indicate absorbance of cells with extracts, viability control and blank, respectively. Linear regression was used for analysis, considering R² values superior to 0.9. All conditions were tested in triplicates.

Antiviral activity assay

Vero cells were seeded in 96-well plates ($4 \cdot 10^4$ cells per well) and incubated at 37 °C and 5% CO₂ atmosphere. After 24 hours of incubation, medium was removed and 100 µL of fresh medium containing a serial dilution of extracts (below the CC₅₀ of each extract) were added to the plates. Subsequently, viral suspensions of ZIKV, DENV-2 and YFV were prepared in MEM (1% FBS) to a MOI of 0.1. That is, 100 µL of fresh medium containing $4 \cdot 10^3$ viral particles were added to each well. After 72h of incubation at 37°C and 5% CO₂ atmosphere, medium was removed and 100 µL of MTT (ThermoFischer Scientific, USA) solution in MEM (0.5 mg/mL) were added to

each well. After 3 hours of incubation at 37 °C and 5% CO₂ atmosphere, the medium was removed and 100 µL of DMSO was added to each well solubilize formazan crystals (Takeuchi et al., 1991). After shaking for 15 minutes, absorbance at 570 nm of each well was read using a spectrophotometer (VersaMax, Molecular Devices, USA). Ribavirin and mycophenolic acid (MPA) were used as inhibition controls, and a serial dilution of DMSO was used as viability control. The effective concentration for 50% of the cells (EC₅₀) of each extract was calculated as protection of virus-infected cells, as follow:

$$\frac{(OD_{ve}) - (OD_v)}{(OD_c) - (OD_v)} \times 100(\%)$$

OD_{ve}, OD_v and OD_c indicate absorbance of virus-infected cells with extracts, virus-infected cells only and viability control, respectively. Linear regression was used for analysis, considering R² values superior to 0.9. Selectivity index (SI) to each extract was determined, as the ratio between the cytotoxic concentration (CC₅₀) and the effective concentration of 50% (EC₅₀). All conditions were tested in triplicates.

Statistical analysis

All data were expressed as means ± standard deviations of triplicate measurements. Antiviral activity and cytotoxicity of extracts, fractions and isolated compound were submitted to ANOVA and the Tukey's test was used to compare the means (α = 0.05). All statistical analysis were performed using the software SigmaPlot 12.5 (Systat, 2013).

References

Coelho SVA, Neris RLS, Papa MP, Schnellrath LC, Meuren LM, Tschoeke DA, Leomil L, Verçoza BRF, Miranda M, Thompson FL, Da Poian AT, Souza TML, Carneiro FA, Damaso CR, Assunção-Miranda I, De Arruda LB, 2017. Development of standard methods for Zika virus propagation, titration, and purification. *J Virol Methods* 246:65–74, 2017.

Donald CL, Brennan B, Cumberworth SL, Rezelj VV, Clark JJ, Cordeiro MT, França RFO, Pena LJ, Wilkie GS, Filipe AS, Davis C, Hughes J, Varjak M, Selinger M, Zuvanov L, Owsianka AM, Patel AH, Mclauchlan J, Lindenbach BD, Fall G, Sall AA, Biek R, Rehwinkel J, Schnettler E, Kohl A, 2016. Full Genome Sequence and sfRNA Interferon Antagonist Activity of Zika Virus from Recife, Brazil. *Plos Negl Trop Dis* 10:e0005048.

Figueiredo L B, Sakamoto T, Coelho LFL, Rocha ESO, Gomes Cota MM, Ferreira GP, de Oliveira JG, Kroon EG, 2014. Dengue Virus 2 American-Asian Genotype Identified during the 2006/2007 Outbreak in Piauí, Brazil Reveals a Caribbean Route of Introduction and Dissemination of Dengue Virus in Brazil. *Plos One* 9:e104516.

Mosmann T, 1983. Rapid colorimetric assay for cellular growth and survival: Application to proliferation and cytotoxicity assays. *J Immunol Methods* 65:55-63.

Singh D, Arya PV, Sharma A, Aggarwal VP, Dobhal MP, Gupta RS, 2014. Antioxidant Potential of Plumieride against CCl₄-Induced Peroxidative Damage in Rats. *Antioxidants* 3:798-813.

Systat, 2013. Software, San Jose, CA, USA.

Takeuchi H, Baba M, Shigeta S, 1991. An application of tetrazolium (MTT) calorimetric assay for the screening of anti-herpes simplex virus compounds. *J Virol Methods* 33:61-71.

Wagner H, Blatt S, Zgainski EM. 1984. Plant drug analysis: a thin layer chromatography atlas. Springer Verlag, ed. 1^a, Berlin.

Figure S1

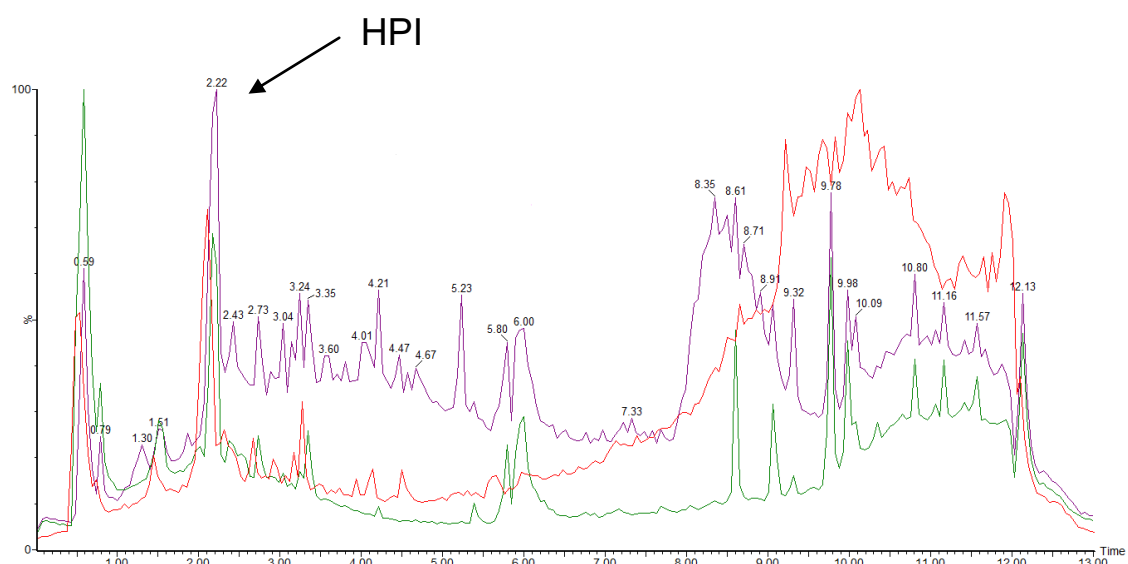


Figure S2

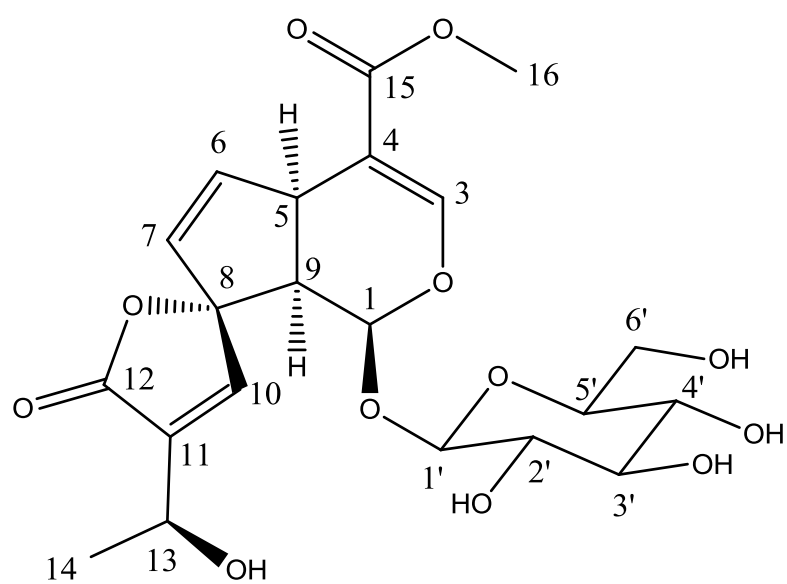


Figure Legend

Figure S1. Full scan of Total Positive Ion Chromatogram (TIC) of the HPE EtOH (red), HPF H₂O (green), and HPF CHCl₃ (purple). Conditions: See Experimental Section.

Figure S2. Chemical structure of the sesquiterpene glucoside plumieride.