



UNIVERSIDADE FEDERAL DE JATAÍ

UNIVERSIDADE FEDERAL DE JATAÍ
UNIDADE ACADÊMICA ESPECIAL DE CIÊNCIAS DA SAÚDE
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS APLICADAS À SAÚDE

GABRIELA DE LIMA MENEZES

**Prospecção de inibidores da formação do dímero da NS1 de
três principais *Flavivirus* de importância médica: uma
abordagem por dinâmica molecular e triagem virtual de
compostos**

JATAÍ

2021



UNIVERSIDADE FEDERAL DE GOIÁS
UNIDADE ACADÊMICA ESPECIAL DE CIÊNCIAS EXATAS

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO (TECA) PARA DISPONIBILIZAR VERSÕES ELETRÔNICAS DE TESES

E DISSERTAÇÕES NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a [Lei 9.610/98](#), o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou download, a título de divulgação da produção científica brasileira, a partir desta data.

O conteúdo das Teses e Dissertações disponibilizado na BDTD/UFG é de responsabilidade exclusiva do autor. Ao encaminhar o produto final, o autor(a) e o(a) orientador(a) firmam o compromisso de que o trabalho não contém nenhuma violação de quaisquer direitos autorais ou outro direito de terceiros.

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO (TECA) PARA DISPONIBILIZAR VERSÕES ELETRÔNICAS DE TESES E DISSERTAÇÕES NA BIBLIOTECA DIGITAL DA UFJ

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Jataí (UFJ) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFJ), regulamentada pela Resolução CEPEC no 832/2007, sem ressarcimento dos direitos autorais, de acordo com a Lei 9.610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou download, a título de divulgação da produção científica brasileira, a partir desta data. O conteúdo das Teses e Dissertações disponibilizado na BDTD/UFJ é de responsabilidade exclusiva do autor. Ao encaminhar o produto final, o autor(a) e o(a) orientador(a) firmam o compromisso de que o trabalho não contém nenhuma violação de quaisquer direitos autorais ou outro direito de terceiros.

1. Identificação do material bibliográfico

[x] Dissertação [] Tese

2. Nome completo do autor:

GABRIELA DE LIMA MENEZES

3. Título do trabalho:

PROSPECÇÃO DE INIBIDORES DA FORMAÇÃO DO DÍMERO DA NS1 DE TRÊS PRINCIPAIS FLAVIVIRUS DE IMPORTÂNCIA MÉDICA: UMA ABORDAGEM POR DINÂMICA MOLECULAR E TRIAGEM VIRTUAL DE COMPOSTOS

4. Informações de acesso ao documento (este campo deve ser preenchido pelo orientador)

Concorda com a liberação total do documento [x] SIM [] NÃO

[1] Neste caso o documento será embargado por até um ano a partir da data de defesa. Após esse período, a possível disponibilização ocorrerá apenas mediante:

a) consulta ao(à) autor(a) e ao(à) orientador(a);

b) novo Termo de Ciência e de Autorização (TECA) assinado e inserido no arquivo da tese ou dissertação.

O documento não será disponibilizado durante o período de embargo.

Casos de embargo:

- Solicitação de registro de patente;

- Submissão de artigo em revista científica;

- Publicação como capítulo de livro;
- Publicação da dissertação/tese em livro.

[1] Neste caso o documento será embargado por até um ano a partir da data de defesa. Após esse período, a possível disponibilização ocorrerá apenas mediante:

a) consulta ao(à) autor(a) e ao(à) orientador(a);

b) novo Termo de Ciência e de Autorização (TECA) assinado e inserido no arquivo da tese ou dissertação.

O documento não será disponibilizado durante o período de embargo.

Casos de embargo:

- Solicitação de registro de patente;
- Submissão de artigo em revista científica;
- Publicação como capítulo de livro;
- Publicação da dissertação/tese em livro.

Obs. Este termo deverá ser assinado no SEI pelo orientador e pelo autor.



Documento assinado eletronicamente por **Roosevelt Alves Da Silva, Professor do Magistério Superior**, em 26/08/2021, às 14:59, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **GABRIELA DE LIMA MENEZES, Discente**, em 26/08/2021, às 19:42, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **2307374** e o código CRC **4FC36233**.

GABRIELA DE LIMA MENEZES

**Prospecção de inibidores da formação do dímero da NS1 de
três principais *Flavivirus* de importância médica: uma
abordagem por dinâmica molecular e triagem virtual de
compostos**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciências Aplicadas à Saúde da Universidade Federal de Jataí como requisito parcial para obtenção do Título de Mestre em Ciências Aplicadas à Saúde.

Área de concentração: Mecanismos e processos biológicos e biotecnológicos

Linha de pesquisa: Mecanismos moleculares e funcionais envolvidos na manutenção da saúde.

Orientador: Prof. Dr. Roosevelt Alves da Silva

JATAÍ

2021

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UFJ.

Menezes, Gabriela de Lima

Prospecção de inibidores da formação do dímero da NS1 de três principais Flavivirus de importância médica [manuscrito] : uma abordagem por dinâmica molecular e triagem virtual de compostos / Gabriela de Lima Menezes. - 2021.

145, f.

Orientador: Prof. Dr. Roosevelt Alves da Silva.

Dissertação (Mestrado) - Universidade Federal de Jataí, Unidade Acadêmica Especial de Ciências da Saúde, Jataí, Programa de Pós-Graduação em Ciências Aplicadas a Saúde, Jataí, 2021.

Ilustrações. Apêndice.

Inclui siglas, tabelas, lista de figuras, lista de tabelas.

1. Flavivirus. 2. Desenho racional de fármacos. 3. Biologia computacional. 4. Ancoragem molecular. 5. Dinâmica molecular. I. Silva, Roosevelt Alves da, orient. II. Título.

CDU 61



UNIVERSIDADE FEDERAL DE GOIÁS
COORDENAÇÃO DE PÓS-GRADUAÇÃO - REGIONAL JATAÍ
ATA DE DEFESA DE DISSERTAÇÃO

Ata nº **34** da sessão de Defesa de Dissertação de GABRIELA DE LIMA MENEZES, que confere o título de Mestra em Ciências Aplicadas à Saúde, na área de concentração em Mecanismos e Processos Biológicos e Biotecnológicos.

No dia primeiro de junho de dois mil e vinte e um, a partir das 15h00 horas, realizou-se a sessão pública integralmente por meio de tecnologias de comunicação à distância de Defesa de Dissertação intitulada “PROSPECÇÃO DE INIBIDORES DA FORMAÇÃO DO DÍMERO DA NS1 DE TRÊS PRINCIPAIS FLAVIVIRUS DE IMPORTÂNCIA MÉDICA: UMA ABORDAGEM POR DINÂMICA MOLECULAR E TRIAGEM VIRTUAL DE COMPOSTOS”. Os trabalhos foram instalados pelo Orientador, Professor Doutor Roosevelt Alves da Silva (CIEXA / UFJ) com a participação dos demais membros da Banca Examinadora: Professor Doutor Marcos Lázaro Moreli (CISAU - UFJ), membro titular interno; Professor Doutor Jonas Ivan Nobre Oliveira (Dep. de Biofísica e Farmacologia/UFRN), membro titular externo. Durante a arguição os membros da banca mantiveram o título do trabalho. A Banca Examinadora reuniu-se em sessão secreta a fim de concluir o julgamento da Dissertação, sendo a candidata **aprovada** pelos seus membros. Proclamados os resultados pelo Professor Doutor Roosevelt Alves da Silva, Presidente da Banca Examinadora, foram encerrados os trabalhos e, para constar, lavrou-se a presente ata que é assinada pelos Membros da Banca Examinadora, no dia primeiro de junho de dois mil e vinte e um.

TÍTULO SUGERIDO PELA BANCA



Documento assinado eletronicamente por **Marcos Lázaro Moreli, Professor do Magistério Superior**, em 01/06/2021, às 23:03, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Roosevelt Alves Da Silva, Professor do Magistério Superior**, em 02/06/2021, às 09:20, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Jonas Ivan Nobre Oliveira, Usuário Externo**, em 02/06/2021, às 11:32, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).

A autenticidade deste documento pode ser conferida no site



https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **2084568** e o código CRC **22870C19**.

Referência: Processo nº 23070.024867/2021-65

SEI nº 2084568

*À minha mãe por todo amor, dedicação e
apoio para concretização de mais essa
conquista na minha vida.*

AGRADECIMENTOS

À Deus, por nunca ter me deixado desistir e por iluminar minha mente em momentos importantes.

À minha mãe, por todo amor dedicado traduzido através dos esforços para me dar todas as condições necessárias para a concretização dos meus sonhos. É tudo por você e para você!

Ao meu padrasto Álvaro que teve que partir no meio dessa jornada, mas que sempre torceu e se orgulhou de cada passo que eu dava. Tenho certeza que hoje está celebrando mais essa conquista.

À toda minha família, irmãos, primos, tios e meus avós, que mesmo de longe sempre torceram pelo meu sucesso e compreenderam minha ausência em momentos especiais.

À minha madrinha Vera, por sempre me colocar em suas orações e vibrar pelas minhas vitórias.

Ao meu orientador, Dr. Roosevelt por todo acolhimento desde a iniciação científica até agora. Graças ao seus incentivos eu consegui alçar voos que jamais acreditei que conseguiria. Serei eternamente grata e te levarei como exemplo a ser seguido!

Aos meus colegas que tive a honra de trabalhar junto no NCBios: Luan, Luiza, Thalita, Júlia, Lu e Ricardo. Todos vocês foram fundamentais no meu crescimento como pessoa e pesquisadora.

Às minhas colegas de turma do mestrado, em especial Carol, Natália, Nayara e Suzi. Obrigada por tantas risadas, força e torcida. Vocês estarão sempre no meu coração.

Aos meus professores do PPGCAS que sem dúvidas foram fundamentais nessa trajetória e me ensinaram tanto.

Às minhas amadinhas, por toda torcida e carinho que pude sentir mesmo distante. Obrigada pela compreensão quando não pude estar presente momentos tão importantes na vida de vocês.

A todos que tive o prazer de conviver e conhecer ao longo dessa jornada e que de alguma forma me ajudaram a chegar até aqui, minha eterna gratidão!

Às agências de fomento CAPES e CNPq pelo apoio dado para o desenvolvimento desta pesquisa.

SUMÁRIO

1	INTRODUÇÃO / REVISÃO DA LITERATURA	1
1.1	<i>Flavivirus</i>	1
1.1.1	Doenças e epidemiologia das infecções pelo DENV, YFV e ZIKV	2
1.1.1.1	Vírus da dengue (DENV)	2
1.1.1.2	Vírus da febre amarela (YFV)	4
1.1.1.3	Vírus da Zika (ZIKV)	6
1.1.2	Proteína NS1	8
1.2	Ferramentas de Bioinformática	11
1.2.1	Modelagem de proteínas	11
1.2.2	Dinâmica Molecular	14
1.2.3	Ancoragem molecular	17
1.2.4	Predição <i>in silico</i> de parâmetros ADMET	18
2	JUSTIFICATIVA	23
3	OBJETIVOS	24
3.1	Objetivo Geral	24
3.2	Objetivo Específicos	24
4	METODOLOGIA	25
4.1	Obtenção das estruturas monoméricas da NS1	25
4.2	Simulações de dinâmica molecular	26
4.3	Análise das cavidades de ancoragem	28
4.4	Triagem virtual de compostos	30
4.5	Análise perfil ADMET	31
4.6	Análise estatística	31
5	RESULTADOS	32
5.1	Modelagem por homologia	32
5.2	Simulações de dinâmica molecular	33
5.3	Análise das cavidades de ancoragem da proteína NS1	40
5.4	Triagem virtual de compostos	43

5.4.1 Critério de afinidade.....	43
5.4.1.1 Comparação da triagem virtual entre estruturas da NS1 do DENV...	43
5.4.1.2 Comparação da triagem virtual entre estruturas da NS1 do YFV.....	47
5.4.1.3 Comparação da triagem virtual entre estruturas da NS1 do ZIKV....	54
5.4.1.4 Análise da intersecção de compostos presentes na triagem virtual das três espécies.....	61
5.4.1.5 Ancoragem molecular dos compostos sem risco toxicológico.....	68
5.4.2 Critério de eficiência.....	77
5.4.2.1 Comparação da triagem virtual entre estruturas da NS1 do DENV...	77
5.4.2.2 Comparação da triagem virtual entre estruturas da NS1 do YFV.....	78
5.4.2.3 Comparação da triagem virtual entre estruturas da NS1 do ZIKV....	91
5.4.2.4 Análise da intersecção de compostos presentes na triagem virtual das três espécies.....	95
6 DISCUSSÃO.....	101
7 CONCLUSÃO E PESPECTIVAS FUTURAS.....	104
REFERÊNCIAS.....	106
APÊNDICE A.....	117
APÊNDICE B.....	119
APÊNDICE C.....	120
APÊNDICE D.....	121
APÊNDICE E.....	122
APÊNDICE F.....	125

QUADROS, TABELAS, FIGURAS E ANEXOS

Lista de Tabelas

Tabela 1 - Estruturas selecionadas após a inspeção visual.....	42
Tabela 2 - Interseção de compostos por critério de afinidade presentes nas estruturas da NS1.....	62
Tabela 3 – Compostos por critério de afinidade presentes na interseção entre estruturas da NS1.....	63
Tabela 4 - Predição toxicidade, <i>druglikeness</i> e <i>drug-score</i> de acordo com OSIRIS dos cinco compostos em comum.....	67
Tabela 5 - Interseção de compostos por critério de eficiência presentes nas estruturas da NS1.....	96
Tabela 6 - Compostos por critério de eficiência presentes na interseção entre estruturas da NS1.....	96
Tabela 7 - Seis compostos selecionados como promissores para inibição da formação do dímero da NS1.....	103

Lista de Figuras

Figura 1 – Representação da estrutura geral dos <i>Flavivirus</i>	1
Figura 2 – Estrutura do genoma do <i>Flavivirus</i>	2
Figura 3 – Distribuição geográfica dos casos de dengue reportados em 2020.....	3
Figura 4 – Número de casos de Dengue nas Américas.....	4
Figura 5 – Ciclos do YFV.....	5
Figura 6 – Gráfico do número de casos de infecções por YFV entre 2000 e 2019 no Brasil.....	6
Figura 7 – Casos de DENV, YFV e ZIKV no Brasil.....	8
Figura 8 – Papel da NS1 proposto para o ciclo de vida do vírus da dengue.....	9
Figura 9 – Estrutura 3D do dímero NS1 Zika Vírus (PDB ID: 5K6K).....	10

Figura 10 – Etapas da modelagem por homologia.....	13
Figura 11 – Avaliação da predição teórica de toxicidade pelo OSIRIS.....	19
Figura 12 – Distribuição do valor de druglikeness predito pelo OSIRIS.....	20
Figura 13 – Distribuição do valor de clogP de fármacos comerciais.....	20
Figura 14 – Distribuição do valor de logS de fármacos comerciais.....	21
Figura 15 – Distribuição do peso molecular de fármacos comerciais.....	21
Figura 16 – Fluxograma da metodologia básica utilizada no trabalho.....	25
Figura 17 – Sistema NS1 do ZIKV.....	27
Figura 18 – Ramachandran dos modelos gerados da NS1 do DENV e YFV.....	32
Figura 19 – Estruturas monoméricas da proteína NS1.....	33
Figura 20 – RMSD das réplicas da NS1 do ZIKV.....	34
Figura 21 – Mapa de calor do RMSD por resíduo ao longo da trajetória da NS1 do ZIKV.....	35
Figura 22 – RMSF das simulações de DM das réplicas da NS1 do ZIKV.....	35
Figura 23 – RMSD das simulações de DM das réplicas da NS1 do YFV.....	36
Figura 24 – RMSF das simulações de DM das réplicas da NS1 do YFV.....	37
Figura 25 - RMSD das simulações de DM das réplicas da NS1 do DENV.....	38
Figura 26 - RMSF das simulações de DM das réplicas da NS1 do DENV.....	39
Figura 27 - Alinhamento da sequência de aminoácido das NS1 do ZIKV, YFV e DENV.....	40
Figura 28 - Número de estruturas em comum nas análises do TM-align.....	41
Figura 29 - Inspeção visual da viabilidade da cavidade.....	42
Figura 30 - Alinhamento estrutural das seis conformações selecionadas para a triagem virtual.....	42
Figura 31 - <i>Boxplot</i> das distribuições de energia das 100 simulações exaustivas dos 100 compostos ranqueados de acordo com o critério de energia (afinidade) para as duas estruturas da NS1 do DENV.....	43
Figura 32 - Representação do diagrama 2D de interação dos dois compostos em	

comum às duas estruturas do DENV com seus respectivos escores de energia	45
Figura 33 - Análise fragmento-específico do ligante ZINC000070704746 no sítio da NS1 do DENV	46
Figura 34 - Análise fragmento-específico do ligante ZINC000070699829 no sítio da NS1 do DENV	47
Figura 35 - <i>Boxplot</i> das distribuições de energia das 100 simulações exaustivas dos 100 compostos ranqueados de acordo com o critério de energia (afinidade) para as duas estruturas da NS1 do YFV	48
Figura 36 - Representação do diagrama 2D de interação dos cinco compostos em comum às duas estruturas do YFV com seus respectivos escores de energia	49
Figura 37 - Análise fragmento-específico do ligante ZINC000008792187 no sítio da NS1 do YFV	51
Figura 38 - Análise fragmento-específico do ligante ZINC000070705115 no sítio da NS1 do YFV	52
Figura 39 - Análise fragmento-específico do ligante ZINC000095486217 no sítio da NS1 do YFV	52
Figura 40 - Análise fragmento-específico do ligante ZINC000015956900 no sítio da NS1 do YFV	53
Figura 41 - Análise fragmento-específico do ligante ZINC000070699829 no sítio da NS1 do YFV	54
Figura 42 - <i>Boxplot</i> das distribuições de energia das 100 simulações exaustivas dos 100 compostos ranqueados de acordo com o critério de energia (afinidade) para as duas estruturas da NS1 do ZIKV	55
Figura 43 - Representação do diagrama 2D de interação dos seis compostos em comum às duas estruturas do ZIKV com seus respectivos escores de energia	56
Figura 44 - Análise fragmento-específico do ligante ZINC000524730935 no sítio da NS1 do ZIKV	58
Figura 45 - Análise fragmento-específico do ligante ZINC000070691963	

no sítio da NS1 do ZIKV.....	59
Figura 46 - Análise fragmento-específico do ligante ZINC000008792352 no sítio da NS1 do ZIKV.....	60
Figura 47 - Análise fragmento-específico do ligante ZINC0000085878218 no sítio da NS1 do ZIKV.....	60
Figura 48 - Análise fragmento-específico do ligante ZINC0000085569026 no sítio da NS1 do ZIKV.....	61
Figura 49 - <i>Boxplot</i> da distribuição das energias da triagem virtual obtida para os cinco compostos em comum.....	63
Figura 50 - Similaridade estrutural entre os cinco compostos.....	64
Figura 51 – Diagrama das principais interações* observadas entre proteína e ligante.....	66
Figura 52 – Alteração estrutural das moléculas com média/elevada toxicidade.....	67
Figura 53 – <i>Boxplot</i> da distribuição das energias dos compostos antes e após modificação para cada estrutura da NS1.....	68
Figura 54 – Análise comparativa da distribuição global das energias dos compostos antes e após a modificação estrutural.....	69
Figura 55 – Diagrama das principais interações observadas entre receptor e ligante após modificação estrutural.....	70
Figura 56 – Diagrama 2D da interação entre proteína NS1 e ligante ZINC000008792352 antes e após modificação estrutural.....	71
Figura 57 - Diagrama 2D da interação entre proteína NS1 e ligante ZINC000012879754 antes e após modificação estrutural.....	73
Figura 58 - Diagrama 2D da interação entre proteína NS1 e ligante ZINC000002108995 antes e após modificação estrutural.....	74
Figura 59 - Diagrama 2D da interação entre proteína NS1 e ligante ZINC000008792389 antes e após modificação estrutural.....	76
Figura 60 – <i>Boxplot</i> das distribuições de energia das 100 simulações exaustivas para os 100 compostos ranqueados de acordo com critério de eficiência para as duas estruturas da NS1 do DENV.....	77
Figura 61 – <i>Boxplot</i> das distribuições de energia das 100 simulações exaustivas para os 100 compostos ranqueados de acordo com critério	

de eficiência para as duas estruturas da NS1 do YFV.....	78
Figura 62 - Representação do diagrama 2D de interação dos onze compostos em comum às duas estruturas do YFV com seus respectivos escores de energia.....	80
Figura 63 - Análise fragmento-específico do ligante ZINC000013370090 no sítio da NS1 do YFV.....	83
Figura 64 - Análise fragmento-específico do ligante ZINC000000462723 no sítio da NS1 do YFV.....	84
Figura 65 - Análise fragmento-específico do ligante ZINC000002123259 no sítio da NS1 do YFV.....	84
Figura 66 – Pose de ancoragem dos ligantes ZINC000013370090 e ZINC000002123259.....	85
Figura 67 - Análise fragmento-específico do ligante ZINC000000095489 no sítio da NS1 do YFV.....	85
Figura 68 - Pose de ancoragem dos ligantes ZINC000002123259 e ZINC000000095489.....	86
Figura 69 - Análise fragmento-específico do ligante ZINC000000485867 no sítio da NS1 do YFV.....	86
Figura 70 - Análise fragmento-específico do ligante ZINC000000086259 no sítio da NS1 do YFV.....	87
Figura 71 - Análise fragmento-específico do ligante ZINC000000178240 no sítio da NS1 do YFV.....	88
Figura 72 - Análise fragmento-específico do ligante ZINC000000488503 no sítio da NS1 do YFV.....	88
Figura 73 - Análise fragmento-específico do ligante ZINC000065248686 no sítio da NS1 do YFV.....	89
Figura 74 - Análise fragmento-específico do ligante ZINC000065248870 no sítio da NS1 do YFV.....	90
Figura 75 – Conformações da ancoragem dos compostos	

ZINC000065248870 e ZINC000000095489.....	90
Figura 76 - Análise fragmento-específico do ligante ZINC000065249270 no sítio da NS1 do YFV.....	91
Figura 77 - <i>Boxplot</i> das distribuições de energia das 100 simulações exaustivas para os 100 compostos ranqueados de acordo com critério de eficiência para as duas estruturas da NS1 do ZIKV.....	92
Figura 78 - Representação do diagrama 2D de interação dos dois compostos em comum às duas estruturas do ZIKV com seus respectivos escores de energia.....	93
Figura 79 - Análise fragmento-específico do ligante ZINC000000271839 no sítio da NS1 do ZIKV.....	94
Figura 80 - Análise fragmento-específico do ligante ZINC000006117574 no sítio da NS1 do ZIKV.....	95
Figura 81 – <i>Boxplot</i> da distribuição das energias da triagem virtual obtida para os dois compostos em comum.....	97
Figura 82 – <i>Boxplot</i> da distribuição global das energias dos dois compostos em comum.....	97
Figura 83 – Diagrama 2D da interação entre proteína NS1 e os dois ligantes em comum (ZINC000000271839 e ZINC000000485867).....	99
Figura 84 – Alteração estrutural da região apontada como potencial risco reprodutivo para o composto ZINC000000485867.....	100

SÍMBOLOS, SIGLAS E ABREVIATURAS

3D	Tridimensional
Å	Ångström
β	Beta
ADMET	Absorção, Distribuição, Metabolismo, Excreção e Toxicidade.
ALA	Alanina
ARG	Arginina
ASN	Asparagina
ASP	Aspartato
CASP	<i>Critical assessment of methods of protein structure prediction</i>
CLoop	<i>Conector loop</i>
CYX	Cisteína com ponte dissulfeto
DENV	Vírus da dengue (<i>Dengue virus</i>)
DM	Dinâmica molecular
fs	Fentossegundos
GLU	Ácido glutâmico
GLY	Glicina
GROMACS	<i>GRONingen MACHine for Chemical Simulations</i>
HID	Histidina com hidrogênio no nitrogênio delta
HIP	Histidina com hidrogênio em ambos nitrogênios
ID	<i>Identification</i>
ILE	Isoleucina
IST	Infecção Sexualmente Transmissível
K	Kelvin
kcal	kilocaloria
kJ	kiloJoule
LYS	Lisina
MET	Metionina
mNS1	Forma dimérica da proteína não-estrutural 1
NAG	N-acetilglucosamina
nm	Nanômetro
ns	Nanossegundo
NS1	Proteína não-estrutural 1

OMS	Organização Mundial da Saúde
PDB	<i>Protein Data Bank</i>
PHE	Fenilalanina
ps	Picossegundos
RMSD	Raiz do Desvio Quadrático Médio (<i>Root mean square deviation</i>)
RMSF	Raiz da Flutuação do Quadrático Médio (<i>Root mean square fluctuation</i>)
RNA	Ácido ribonucleico (<i>Ribonucleic acid</i>)
SER	Serina
SLI	<i>Spaghetti Loop Interface</i>
sNS1	Forma secretada da proteína não-estrutural 1
THR	Treonina
TRP	Triptofano
VAL	Valina
WNV	Vírus do Oeste do Nilo (<i>West Nile virus</i>)
YFV	Vírus da febre amarela (<i>Yellow fever virus</i>)
ZIKV	Vírus da Zika (<i>Zika virus</i>)

RESUMO

Prospecção de inibidores da formação do dímero da NS1 de três principais *Flavivirus* de importância médica: uma abordagem por dinâmica molecular e triagem virtual de compostos

As doenças causadas por vírus do gênero *Flavivirus* estão entre as principais enfermidades que atingem o mundo, em especial o Brasil, sendo, portanto, um sério problema de saúde pública. No Brasil, destacam-se três delas: dengue, febre amarela e Zika. A proteína não-estrutural 1 (NS1), codificada por esse gênero viral, na sua forma dimérica exerce importantes funções na patogênese e replicação do RNA desses vírus. Entretanto, não há disponível tratamento farmacológico com atuação anti-viral. Assim sendo, a identificação de químicos com potencial de inibição da formação do dímero da proteína NS1 do DENV, YFV e ZIKV os possibilitaria atuarem como um fármaco multialvo. Para isso, com auxílio do TM-Align selecionou-se conformações da proteína NS1 monômero com a estrutura do domínio β -roll similar (TM-Score ≥ 0.57) entre as três espécies de vírus a partir de conformações obtidas de simulações de dinâmica molecular executadas no GROMACS em 5 réplicas de 150 ns para cada espécie. Essa análise resultou na seleção de duas conformações de NS1 para cada espécie viral, totalizando seis conformações de NS1. Após a seleção das estruturas proteicas, realizou-se a triagem virtual de compostos provenientes do catálogo de produtos naturais do banco de dados do ZINC com as seis estruturas de NS1 através do AutoDock Vina. A partir do resultado da ancoragem para cada proteína NS1, foi realizada a classificação dos 100 melhores compostos de acordo com dois critérios: afinidade e eficiência. Na classificação por afinidade, comparou-se os compostos ranqueados para cada conformação de NS1 e cinco deles foram encontrados em comum à NS1 das três espécies virais. Desses, quatro possuíam risco toxicológico de acordo com o programa OSIRIS Property Explorer. Após remoção do fragmento tóxico, uma nova ancoragem apontou redução da afinidade da molécula pela NS1, porém ainda com escores de energia promissores (entre -7.5 kcal/mol e -9.6 kcal/mol). Já entre os compostos classificados por critério de eficiência, apenas dois foram observados em comum às espécies, sendo que um deles possuía risco toxicológico e não se encaixou em parâmetros de compostos de potencial farmacológico após remoção do fragmento tóxico. Os resultados aqui obtidos demonstram a alta similaridade das proteínas NS1 no gênero *Flavivirus* e afinidade pelos mesmos compostos; justificando, pois, o potencial dessas seis macromoléculas na terapêutica multialvo. Os dados teórico-computacionais apresentados fundamentam/amparam futuros experimentos *in vitro* e *in vivo*, tal como ensaios de avaliação da capacidade replicativa do vírus na presença das moléculas bioativas destacadas. Logo, as informações apresentadas neste trabalho serão relevantes para desenvolvimento de uma molécula com potencial de inibição da formação do dímero da proteína NS1.

Palavras-chaves: *Flavivirus*, desenho racional de fármacos, biologia computacional, ancoragem molecular e dinâmica molecular.

ABSTRACT

Prospecting inhibitors of NS1 dimer formation of three main *Flaviviruses* of medical importance: an approach by molecular dynamics and virtual screening of compounds

Diseases caused by viruses of the *Flavivirus* genus are among the main diseases that affect the world, especially Brazil, and they are a serious public health problem. In Brazil, three of them stand out: dengue, yellow fever and Zika. The non-structural protein 1 (NS1), encoded by this viral genus, in its dimeric form plays important roles in the pathogenesis and RNA replication of these viruses. However, there is no available pharmacological treatment with antiviral action. Therefore, the identification of chemicals with the potential to inhibit the formation of the NS1 protein dimer of DENV, YFV and ZIKV would enable them to act as a multi-target drug. For this, with the aid of TM-Align we selected conformations of the NS1 protein monomer with similar β -roll domain structure (TM-Score ≥ 0.57) among the three virus species from conformations obtained from molecular dynamics simulations performed in GROMACS in 5 replicates of 150 ns for each species. This analysis resulted in the selection of two NS1 conformations for each viral species, totaling six NS1 conformations. After selecting the protein structures, a virtual screening of compounds from the natural products catalog of the ZINC database with the six NS1 structures was performed using AutoDock Vina. From the result of anchoring for each NS1 protein, the 100 best compounds were classified according to two criteria: affinity and efficiency. In the affinity classification, the ranked compounds for each NS1 conformation were compared and five of them were found in common with the NS1 of the three viral species. Of these, four were at toxicological risk under the OSIRIS Property Explorer program. After removing the toxic fragment, a new anchor showed a reduction in the affinity of the molecule for NS1, but still with promising energy scores (between -7.5 kcal/mol and -9.6 kcal/mol). Among the compounds classified by efficiency criteria, only two were observed in common to the species, one of which had toxicological risk and did not fit in parameters of compounds with pharmacological potential after removal of the toxic fragment. The results obtained here demonstrate the high similarity of NS1 proteins in the *Flavivirus* genus and affinity for the same compounds; thus justifying the potential of these six macromolecules in multitarget therapy. The theoretical-computational data presented support future *in vitro* and *in vivo* experiments, such as tests to assess the replicative capacity of the virus in the presence of highlighted bioactive molecules. Therefore, the information presented in this work will be relevant for the development of a molecule with the potential to inhibit the formation of the NS1 protein dimer.

Keywords: Computational biology, *Flavivirus*, molecular docking, molecular dynamics, and rational drug design

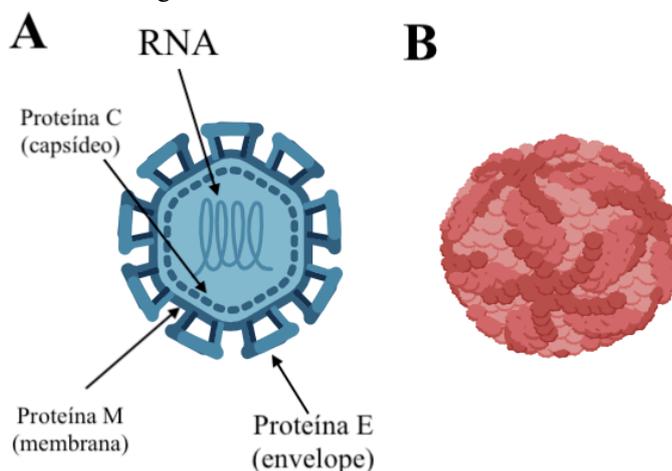
1. INTRODUÇÃO / REVISÃO DA LITERATURA

1.1 *Flavivirus*

Os menores agentes infectantes conhecidos são os vírus, com a habilidade de infectar qualquer forma de célula viva procariota ou eucariota e, ocasionalmente, provocar doenças simples, como resfriados ou doenças letais, como encefalites. Além disso, alguns vírus podem ser responsáveis pelo desencadeamento de processos carcinogênicos (CARTER; SAUNDERS, 2007).

Os vírus são classificados em diferentes famílias e gêneros. O vírus da dengue (*Dengue virus* - DENV), vírus da Zika (*Zika virus* - ZIKV) e vírus da febre amarela (*Yellow fever virus* - YFV) são classificados como entidades da família *Flaviviridae* e gênero *Flavivirus*. Os vírus que compõem esse gênero possuem formato esférico e medem cerca de 40 a 60 nm, apresentam envoltório de constituição lipoproteica cujo ácido nucleico é o RNA de fita simples positiva de aproximadamente 11 kilobases (Figura 1) (CHAMBERS et al., 1990; KORSMAN et al., 2012; VASCONCELOS, 2003).

Figura 1 – Representação da estrutura geral dos *Flavivirus*

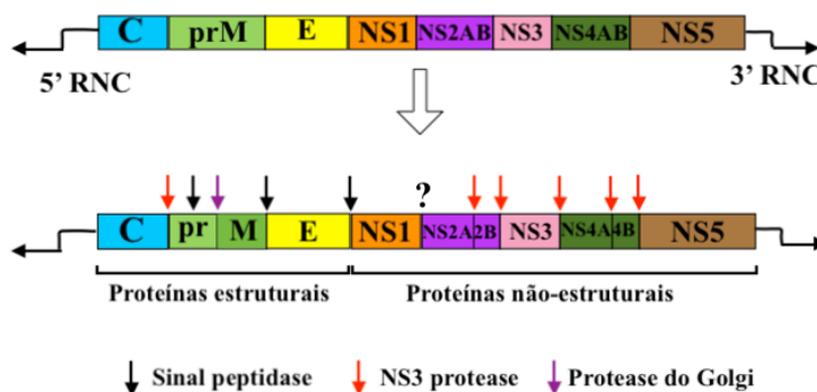


(A) Organização estrutural do vírus, destacando as proteínas estruturais e o código genético. (B) Representação da estrutura esférica dos *Flavivirus*.

O genoma dos *Flavivirus* é inicialmente transcrito em uma poliproteína precursora que é, posteriormente, processada por proteases virais e do hospedeiro em três proteínas estruturais: proteína de capsídeo (C), proteína do Envelope (E) e pré-membrana (prM) ou membrana (M) no vírion imaturo e maduro, respectivamente. Além dessas, há a produção de sete proteínas não-estruturais (NS1, NS2A, NS2B, NS3, NS4A, NS4B e NS5) que não

fazem parte da estrutura do vírus, sendo que a protease responsável pela clivagem da junção NS1/NS2 permanece desconhecida (Figura 2)

Figura 2 – Estrutura do genoma do *Flavivirus*



Processamento da poliproteína em proteínas estruturais e não-estruturais

Os *Flavivirus* podem ser divididos de acordo com o conhecimento ou não do seu vetor. O grupo de vetores conhecidos pode ser separado em vírus transmitidos por carrapatos ou por mosquitos, sendo este último grupo subdividido em dois: vírus transmitido por *Aedes* e por *Culex* (JUNGLEN et al., 2009). Uma interessante observação é o fato de que ao longo do tempo o número de vetores aumentou juntamente com o número de doenças causadas por esse gênero viral (AKEY et al., 2015).

Os vírus, como outros parasitas, precisam se multiplicar para dar continuidade ao processo infeccioso e infectar outros hospedeiros, porém com a peculiar característica que eles se utilizam da maquinaria celular do hospedeiro para realizar seu ciclo de vida (KORSMAN et al., 2012). O processo de infecção da célula se dá por meio de interações da glicoproteína de superfície (proteína E) do *Flavivirus* com receptores específicos presente na superfície celular externa (LINDENBACH; RICE, 2003). Ao adentrar na célula por meio de endocitose, o processo de replicação - no qual as proteínas virais são produzidas - ocorre no citoplasma, seguido pela maturação (processo de montagem completa da partícula viral), que se dá pelo brotamento na membrana do retículo endoplasmático ou complexo de Golgi (WHITE; FENNER, 1994).

1.1.1 Doenças e epidemiologia das infecções pelo DENV, YFV e ZIKV

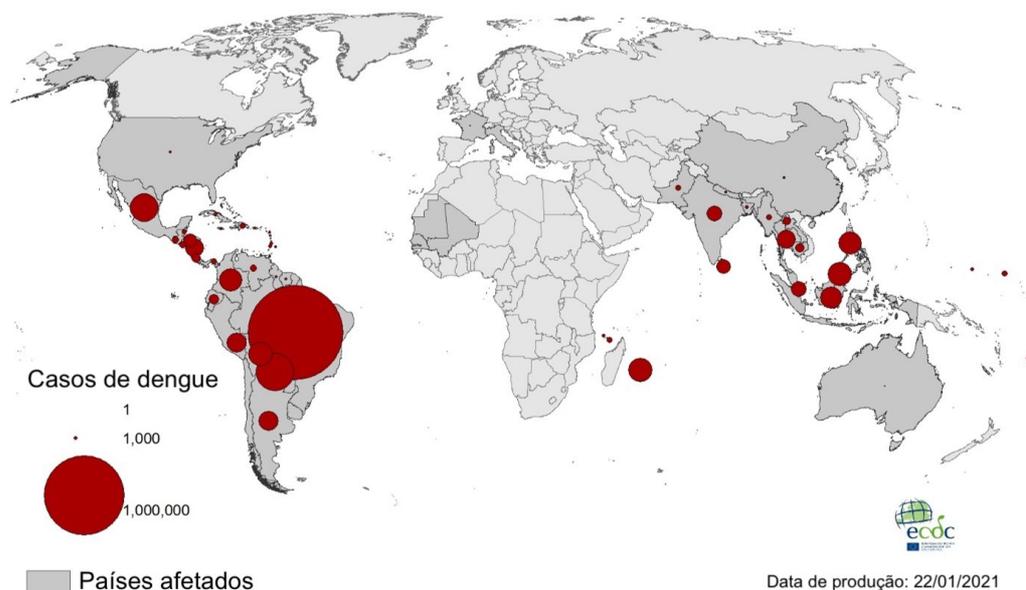
1.1.1.1 Vírus da dengue (DENV)

Até meados de 2013 acreditava-se que o DENV poderia ser encontrado em apenas quatro diferentes sorotipos (DENV-1 ao -4) sendo transmitido por mosquitos do gênero *Aedes*, tendo como principal espécie o *Aedes aegypti* (GUBLER, 2011; GUZMAN;

HARRIS, 2015). Em outubro de 2013 um novo sorotipo (DENV-5) foi anunciado durante um rastreio de amostra viral coletada em 2007 de um fazendeiro da Malásia (NORMILE, 2013). Na infecção causada pelo DENV é comumente observada febre, dores nas articulações com sintomas e sinais similares ao da gripe, sendo reportado fatalidade em alguns casos (SPECIAL PROGRAMME FOR RESEARCH AND TRAINING IN TROPICAL DISEASES; ORGANIZAÇÃO MUNDIAL DA SAÚDE, 2009).

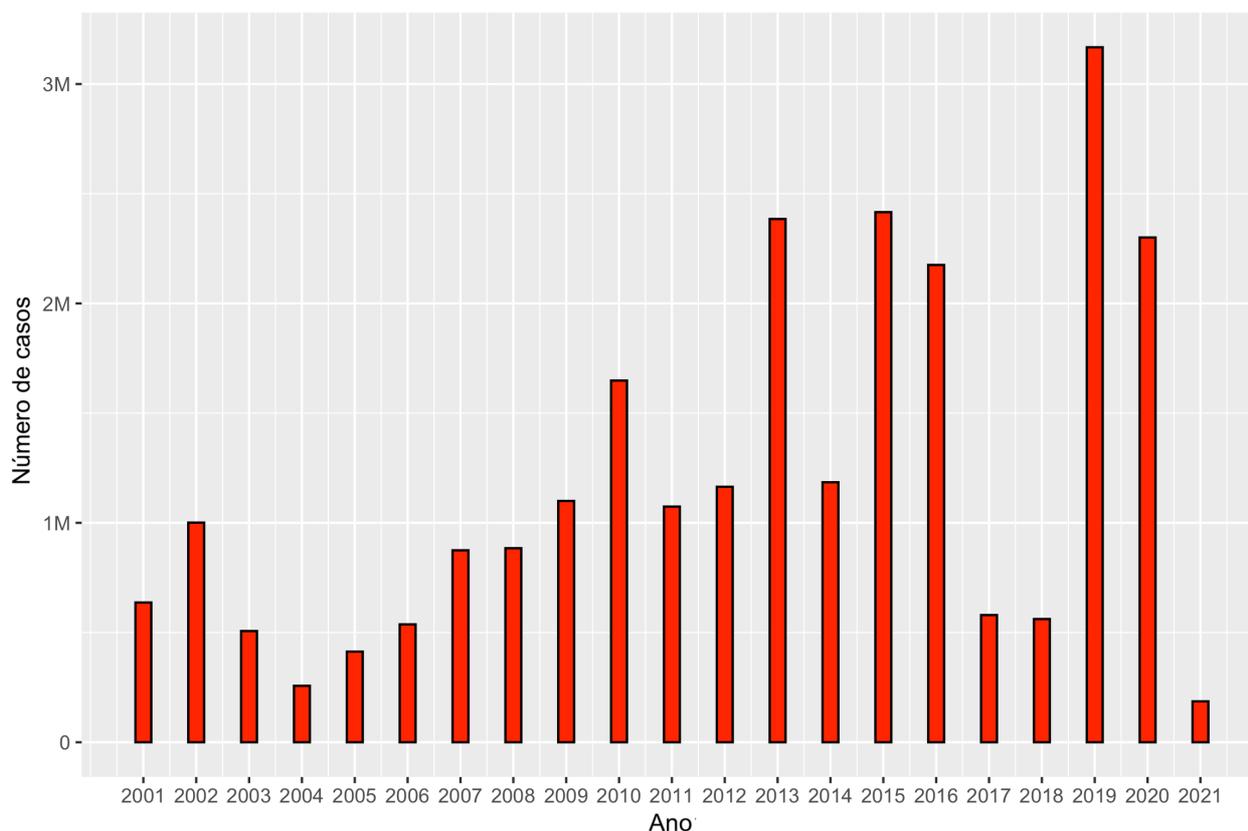
De acordo com os dados da Organização Mundial da Saúde, a doença hoje é considerada endêmica em mais de 100 países e atinge mais de 390 milhões de pessoas anualmente, sendo a América (Figura 3 e Figura 4), sudeste asiático e Pacífico ocidental as regiões mais afetadas (BHATT et al., 2013; CENTRO EUROPEU DE PREVENÇÃO E CONTROLE DE DOENÇAS, 2021; ORGANIZAÇÃO MUNDIAL DA SAÚDE, 2020). De acordo com a Organização Pan-Americana da Saúde, mais de 1.46 milhão de casos foram reportados no Brasil em 2020 (Figura 7) sendo que em 2018 esse número não passou de 300 mil. (ORGANIZAÇÃO PAN-AMERICANA DA SAÚDE, 2021). Um estudo recém realizado prevê o aumento de mais de 2.5 bilhões de pessoas expostas ao risco de dengue em 2080 quando comparado com 2015, totalizando 60 % da população mundial (MESSINA et al., 2019).

Figura 3 – Distribuição geográfica dos casos de dengue reportados em 2020



Número de casos de dengue reportados entre 1 de janeiro de 2020 a 31 de dezembro de 2020. Fonte: ECDC (2021), traduzido.

Figura 4 – Número de casos de Dengue nas Américas



Casos suspeitos ou confirmados de infecções causadas pelo DENV entre 2001 e 2021. Dados de 2021 foram obtidos entre a semana epidemiológica 3 e 12, dependendo do país. Fonte dos dados: Organização Pan-Americana Da Saúde (2021). Disponível em: <https://www.paho.org/data/index.php/en/mnu-topics/indicadores-dengue-en/dengue-nacional-en/252-dengue-pais-ano-en.html?start=1>

Até o presente momento não há uma vacina amplamente disponível para infecções causadas pelos quatro sorotipos do DENV. Recentemente, a vacina tetravalente CYD é a única que foi licenciada para uso e testes clínicos em alguns países sendo que a mesma possui limitação na capacidade de neutralização do DENV-2 (DENG et al., 2020; PROMPETCHARA et al., 2019).

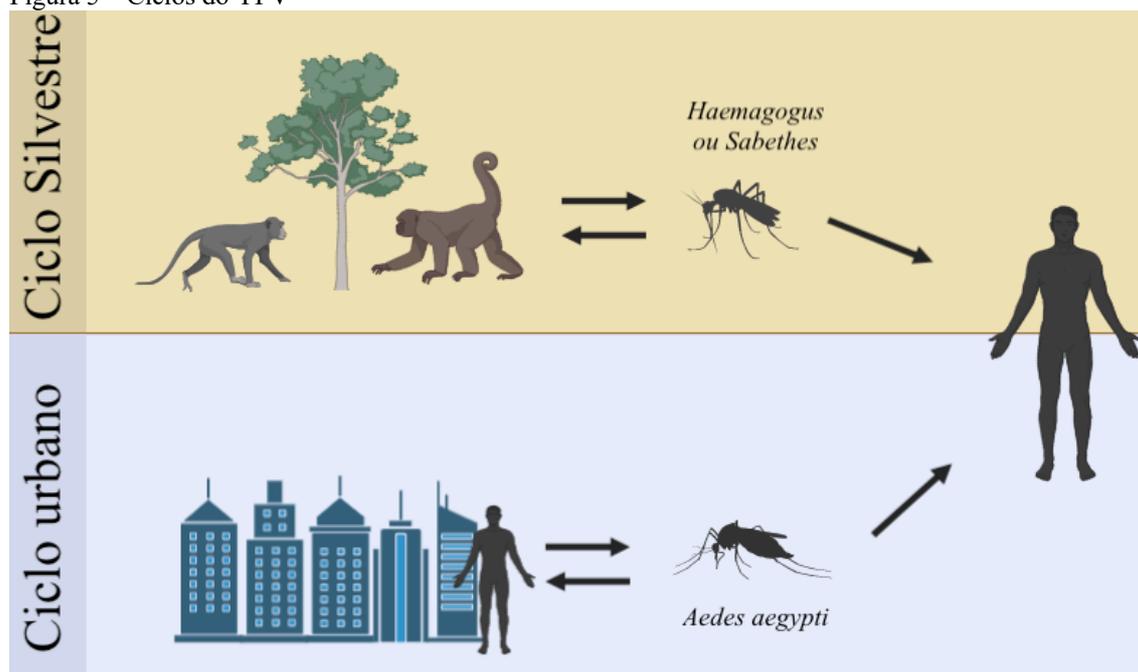
1.1.1.2 Vírus da febre amarela (YFV)

Outro importante *Flavivirus* é o YFV. As infecções causadas pelo YFV podem ser assintomáticas ou apresentar sintomas mais graves como febre, icterícia, doença renal e hemorragia (BARNETT, 2007). O principal tipo celular alvo desse vírus são os hepatócitos, sendo determinante no efeito citopático no fígado, podendo prejudicar a produção de fatores de coagulação e provocar a coagulação intravascular disseminada e trombocitopenia (KORSMAN et al., 2012).

A transmissão do YFV pode ser descrita em dois ciclos principais: o urbano, que ocorre entre *Aedes aegypti* e humanos, e o ciclo silvestre, quando o vírus é transmitido

entre mosquitos e primatas não-humanos, com possível infecção acidental de humanos (Figura 5) (LUTZ, 1930; MONATH; VASCONCELOS, 2015; SOPER et al., 1933). No ocidente, os vetores primários e secundários do ciclo silvestre são mosquitos do gênero *Haemagogus* e *Sabethes*, respectivamente, tendo sido o *Haemagogus* o principal vetor responsável pelo surto de YFV em 2016 a 2018 no Brasil (ABREU et al., 2019; HANLEY et al., 2013).

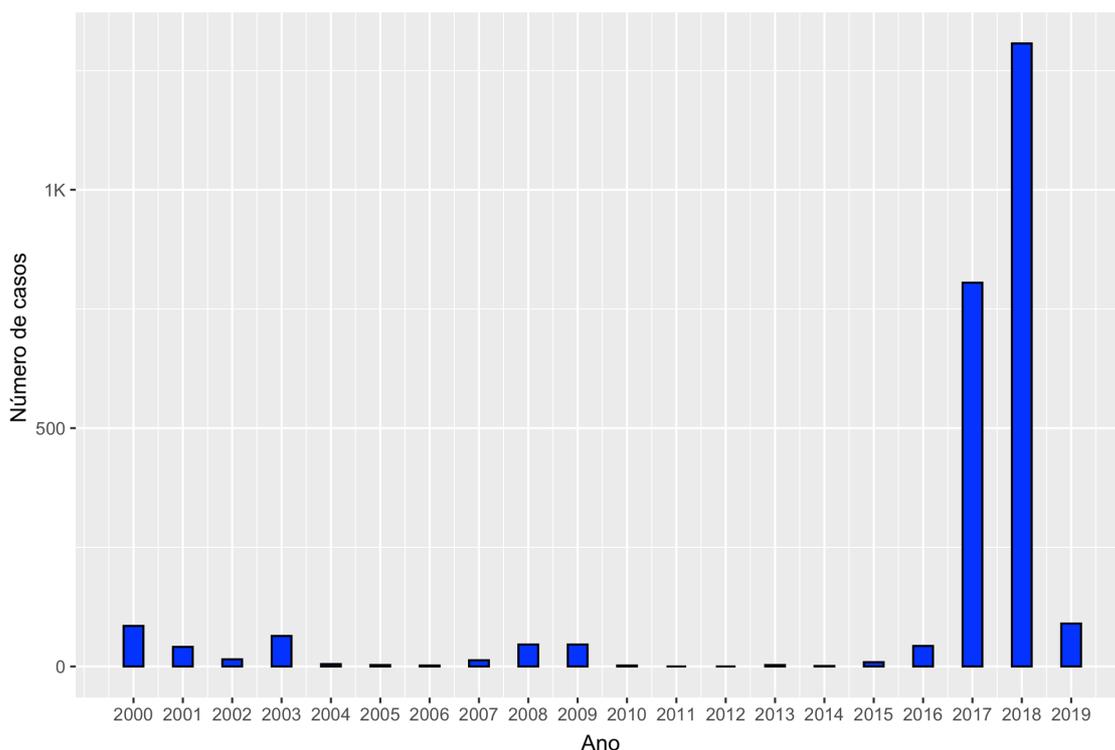
Figura 5 – Ciclos do YFV



Esquema representado os dois ciclos descritos para o YFV no Ocidente. No ciclo silvestres, o humano pode acidentalmente ser infectado, enquanto o ciclo urbano é mantido entre mosquito e humano.

Entre julho de 2017 e junho de 2018, 1376 casos silvestres da doença foram confirmados no Brasil (com 35,1 % de óbitos) e mais de 700 estavam sob investigação (BRASIL, 2018). Diferente do DENV, há vacina disponível contra o YFV desde 1937 contudo, o surto de febre amarela ocorrido entre 2016 e 2018 na região sudoeste do Brasil foi principalmente causado pela baixa cobertura vacinal, já que a última ocorrência de febre amarela urbana no Brasil foi em 1942 (OLIVEIRA et al., 2017; SILVA et al., 2018; THEILER; SMITH, 1937). Além disso, estudos demonstram a persistência do vírus em 2019 no Rio de Janeiro após a sua primeira introdução alertando para uma maior vigilância epidemiológica da doença (GIOVANETTI et al., 2019). A Figura 6 mostra o rápido e alarmante crescimento do número de casos provocado por esse surto.

Figura 6 – Gráfico do número de casos de infecções por YFV entre 2000 e 2019 no Brasil



Fonte: Gráfico gerado a partir de dados obtidos do “WHO vaccine-preventable diseases: monitoring system. 2020 global Summary” – Organização Mundial da Saúde (2020). Disponível em: https://apps.who.int/immunization_monitoring/globalsummary/incidences?c=BRA. Não há dados do ano de 2020 e 2021.

1.1.1.3 Vírus da Zika (ZIKV)

O ZIKV foi primeiro isolado em abril de 1947 do soro de macaco *rhesus* e uma segunda vez foi isolado de mosquitos da espécie *Aedes africanus* quando estudava-se o ciclo silvestre do YFV na floresta Zika (DICK; KITCHEN; HADDOW, 1952). Uma das complicações causadas por essa espécie viral é a síndrome de Guillain-Barré, que foi pela primeira vez associada ao ZIKV depois de um surto ocorrido em 2013 na Polinésia Francesa, onde também foi relatado o risco de documentação por transfusão sanguínea, sêmen, saliva e urina (WEAVER et al., 2016).

No final de 2014 no Brasil, começou-se a observar elevado aumento de pacientes com erupções cutâneas, febre e artralgia na região nordeste do país e em abril de 2015 foi anunciado a provável etiologia desses casos que rapidamente se espalhou pela região: ZIKV (CAMPOS; BANDEIRA; SARDI, 2015). Inicialmente, acreditava-se que o ZIKV havia sido introduzido no Brasil durante Copa do Mundo de Futebol em 2014, ou durante o Campeonato Mundial de Canoagem, ocorrido no mesmo ano com participação de atletas da Polinésia Francesa, mas pesquisadores descartaram as duas hipóteses relatando que o vírus chegou do meio para o final de 2013 (CAMPOS; BANDEIRA; SARDI, 2015;

FARIA et al., 2016; MUSSO, 2015). O ZIKV recentemente provocou surtos da doença na Índia e Angola, o que alerta para o risco de reemergência de ZIKV em regiões com circulação do vetor e presença de pessoas susceptíveis (HAMER; CHEN, 2019; MUSSO; KO; BAUD, 2019).

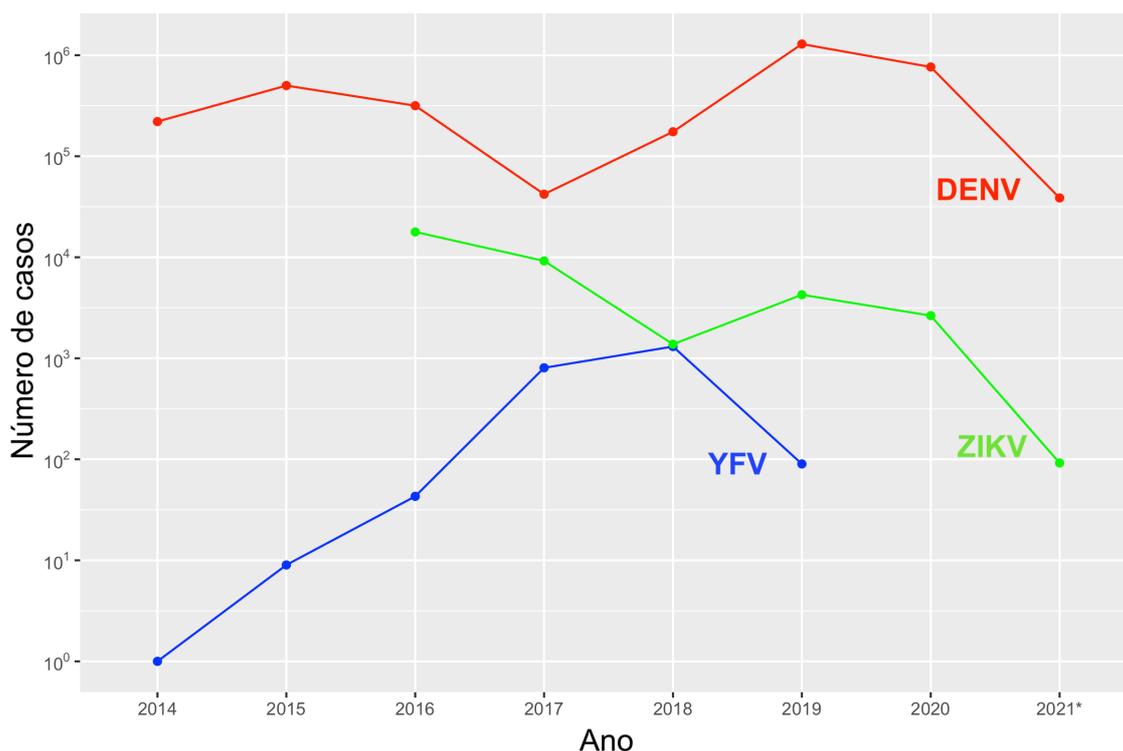
O surto dos casos de ZIKV logo trouxe um aumento preocupante no número de recém-nascidos portadores de microcefalia em especial na região do nordeste brasileiro, e logo em seguida diversos estudos foram publicados relatando a relação da infecção de mulheres grávidas por ZIKV e microcefalia congênita (BRADY et al., 2019; MELO et al., 2016; MLAKAR et al., 2016).

Atualmente, diversas vias de transmissão são conhecidas para o ZIKV. A principal, sem dúvidas, é por meio de vetor artrópode. Na África, destacam-se os mosquitos do subgênero *Aedes (Stegomyia)*, *Aedes (Diceromyia)* e *Aedes (Fredwardsius)*, incluindo *Aedes africanus*, fonte do segundo isolado de ZIKV, *Aedes vittatus* e *Aedes furcifer* como principais vetores (DIAGNE et al., 2015; DIALLO et al., 2014). Já os casos mais recentes de surto de ZIKV tem sido relacionado à transmissão pelo *Aedes aegypti*, como exemplo do primeiro caso confirmado na América em 2015, o que justifica o rápido avanço da doença no Brasil (GUERBOIS et al., 2016).

Outras vias de transmissão do ZIKV incluem transmissão sexual, sendo portanto considerada uma Infecção Sexualmente Transmissível (IST), por via congênita e durante o parto, transfusão sanguínea, e possivelmente pelo leite materno (BARJAS-CASTRO et al., 2016; BESNARD et al., 2014; FOY et al., 2011; SOTELO et al., 2017).

Esses três *Flavivirus* vem se destacando no cenário da saúde pública no Brasil, em especial nos últimos anos com a introdução do ZIKV, e do aumento de casos de YFV observados no surto entre 2016 e 2018 (Figura 7).

Figura 7 – Casos de DENV, YFV e ZIKV no Brasil



Número de casos de DENV, YFV e ZIKV observado nos últimos anos no Brasil. Não há dados de casos de YFV em 2020 e 2021. *Os dados dos casos de ZIKV foram obtidos até a 6ª semana epidemiológica de 2021 e do DENV até a 9ª semana de 2021. Fonte: Organização Pan-Americana Da Saúde (2021).

1.1.2 Proteína NS1

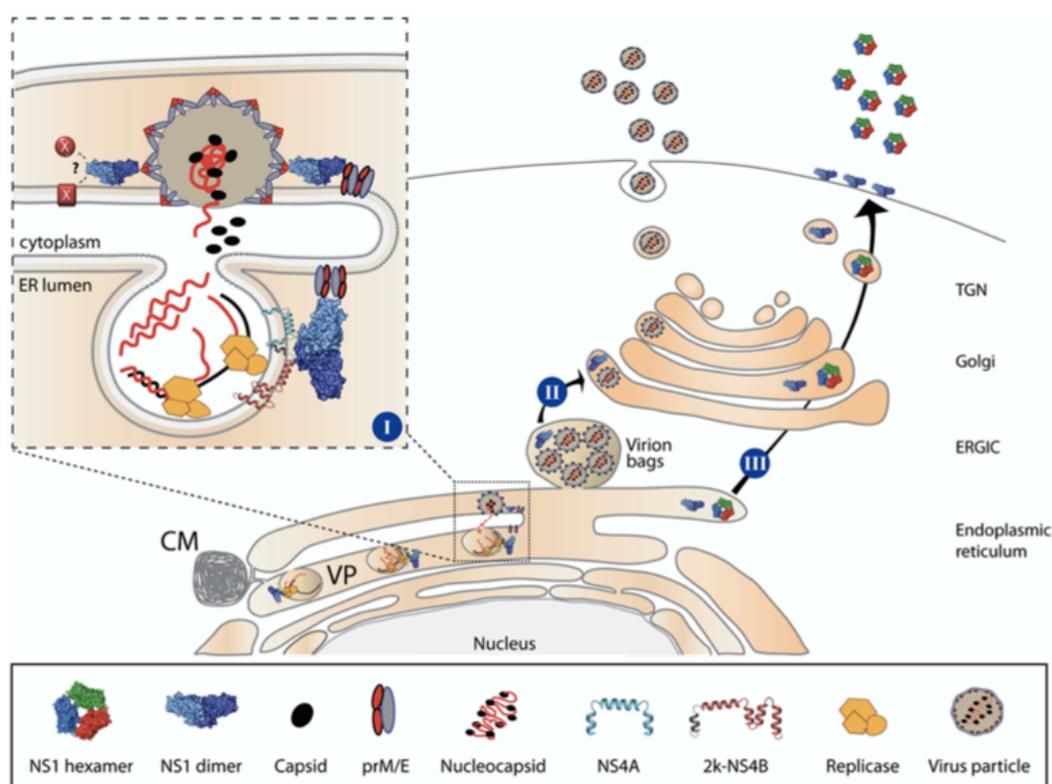
Todos os vírus que fazem parte do gênero *Flavivirus* possuem o gene capaz de produzir a proteína não-estrutural 1 (NS1), a qual apresenta 352 aminoácidos e tem peso molecular de aproximadamente 50 kD, sendo esse gênero o único representante da família *Flaviviridae* que a codifica. Isso sugere, portanto, que essa proteína ajude na replicação e transmissão de dois hospedeiros completamente distintos: mamíferos e insetos (YEN et al., 2016). Na análise da NS1 entre as diferentes espécies de *Flavivirus*, nota-se que há diversas regiões conservadas (RASTOGI; SHARMA; SINGH, 2016).

Inicialmente foi sugerido que a forma madura da proteína era encontrada na forma de dímero, e hoje sabe-se com mais clareza que a forma glicosilada da NS1 pode ser encontrada na forma de homodímero associado à membrana (*membrane-bounded NS1* – mNS1) após translocação no lúmen do retículo endoplasmático (WANG et al., 2017; WINKLER et al., 1988). Outra forma encontrada da NS1 é a forma hexamérica, resultado da união de três homodímeros, que é secretada como uma lipoproteína (*secreted NS1* – sNS1) descrita por interagir com o sistema imune adaptativo e inato, sendo responsável

pela patogênese e evasão do sistema imune do hospedeiro (AKEY et al., 2014; HILGENFELD, 2016).

Atualmente considera-se que a NS1 (esteja ela na membrana plasmática ou no meio extracelular) apresenta funções importantes na patogênese viral ligando-se a proteínas do complemento antagonizando ou modificando suas funções (AVIRUTNAN et al., 2011). Dentro da célula a NS1 apresenta papel importante em eventos modulatórios da replicação do RNA viral uma vez que possui a capacidade de remodelar a estrutura do retículo endoplasmático, sendo esta uma etapa crucial na formação do compartimento de replicação dos *Flavivirus* (CI et al., 2020). Scaturro et al. (2015) em seu estudo demonstrou que interações da NS1 com glicoproteínas do envelope na superfície do virion são fundamentais para eficiência na produção de partículas virais infectantes (Figura 8).

Figura 8 – Papel da NS1 proposto para o ciclo de vida do vírus da dengue

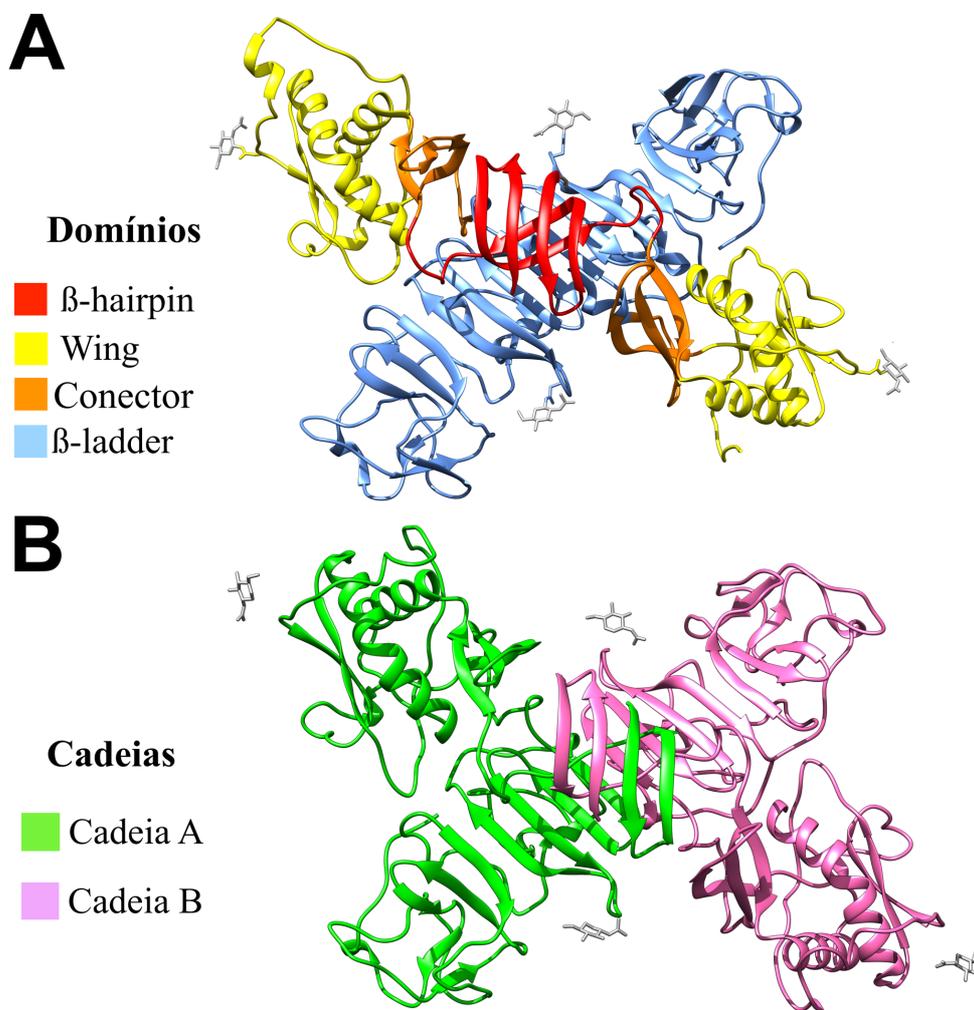


Esquema demonstrando os possíveis papéis da NS1. Em (I) observa-se a NS1 interagindo com a NS4A/NS4B através do domínio β -roll para auxiliar nos eventos iniciais da replicação do RNA viral. (II) Tráfego de partículas virais (do retículo endoplasmático para o compartimento intermediário RE-Golgi) pode ser assistido pela NS1. (III) NS1 é secretada como lipopartícula hexamérica. Fonte: Scaturro et al. (2015).

Cada monômero da proteína NS1 é composto por três domínios bem caracterizados: um N-terminal β -hairpin (abrangendo os resíduos 1 a 30) que forma um

mini-domínio denominado *β-roll* (resíduo 1 ao 25); o domínio *Wing* (aminoácidos 31 ao 180) que possui um sub-domínio descontínuo (31-37 e 152-180) denominado de conector e o domínio *β-ladder* (resíduos 181-352) (Figura 9) (HILGENFELD, 2016).

Figura 9 - Estrutura 3D do dímero NS1 Zika Vírus (PDB ID: 5K6K)



Em *ribbon* estão representadas as cadeias principais do dímero da NS1 do ZIKV. Em *stick* estão os açúcares NAG ligados ao dímero. (A) Destaca-se os domínios da NS1. (B) Destaque para as duas cadeias monoméricas formando o homodímero.

A NS1 possui doze cisteínas bem conservadas em sua estrutura as quais garantem à molécula uma estabilidade pela formação de seis ligações dissulfeto entre os resíduos (BLITVICH et al., 2001; WALLIS et al., 2004). Outra peculiaridade da proteína é a presença de dois sítios de N-glicosilação bem conservados nos resíduos ASN130 e ASN207, cuja função está relacionada com o processo de dimerização e secreção da molécula (CHAMBERS et al., 1990; PRYOR; WRIGHT, 1994). Contudo, acredita-se que a glicosilação do resíduo 207 seja a mais significativa uma vez que Pryor et al. (1994) em seus estudos observaram que a falta do açúcar N-acetilglucosamina (NAG) no resíduo

207, apesar de não impedir, reduziu o processo de dimerização e secreção da proteína. Mais recentemente foi observada que a glicosilação no resíduo 207 da NS1 do DENV, ZIKV e vírus do Oeste do Nilo (*West Nile virus* -WNV) é essencial no papel da disfunção endotelial provocada por essa proteína e, por essa razão, o processo de glicosilação pode ser considerado um interessante alvo para desenvolvimento de fármacos (WANG et al., 2019). Um outro estudo de mutação A188V na proteína NS1 em cepas do ZIKV mostra o aumento de virulência em vírus com essa substituição (XIA et al., 2018). Essa substituição pode estar causando um aumento da flexibilidade da proteína, uma vez que ela ocorre em uma interface de interação e um estudo recém publicado relaciona a flexibilidade da proteína NS1 de diferentes cepas do ZIKV com o aumento da virulência, uma vez que poderia dificultar o reconhecimento pelo sistema imune (POVEDA-CUEVAS; BARROSO DA SILVA; ETCHEBEST, 2021).

Apesar das diversas regiões conservadas na sequência da NS1 entre as espécies de *Flavivirus*, as divergências entre elas podem ser as responsáveis pelas diferenças nas interações entre patógeno e hospedeiro (SONG et al., 2016; XU et al., 2016). Para a NS1 do DENV observou-se que ela sozinha é capaz de induzir a hiperpermeabilidade endotelial de células humanas pulmonares, dérmicas e umbilicais *in vitro*, bem como ao extravasamento vascular no pulmão, fígado e intestino delgado em camundongos (BEATTY et al., 2015; MODHIRAN et al., 2015).

Já a NS1 do ZIKV, que tem como alvo as células do cérebro e placenta humana, foi descrito por elevar a permeabilidade de células endoteliais cerebrais e da veia umbilical, e a NS1 do YFV foi relacionada com a indução *in vitro* da hiperpermeabilidade apenas nas células endoteliais do fígado, sendo esses achados associados com as doenças causadas por essas respectivas espécies (GLASNER et al., 2018).

Em razão dessas diferenças encontradas entre as NS1 dos *Flavivirus* entende-se que estudar o comportamento do monômero das espécies desse gênero viral é de fundamental importância para elucidar os diferentes mecanismos de interação da proteína e processos por trás da patogênese desses vírus.

1.2 Ferramentas de Bioinformática

1.2.1 Modelagem de proteínas

Os avanços na área do sequenciamento, especialmente após o projeto genoma, provocou o surgimento de uma lacuna entre o número de sequências nucleotídicas

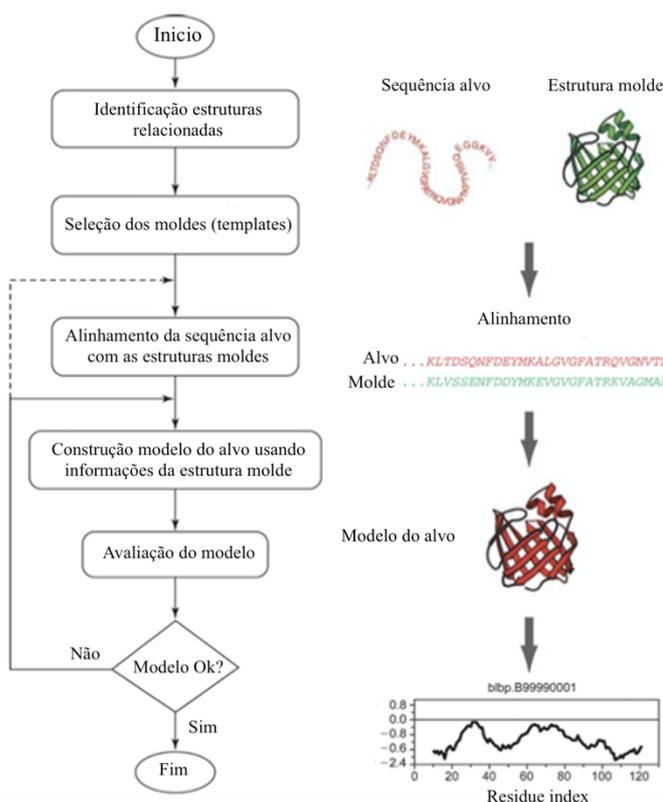
conhecidas e o número de estruturas proteicas resolvidas experimentalmente, uma vez que esta última requer um aparato experimental maior e mais oneroso (RODRIGUEZ et al., 1998; STOESSER et al., 2002). Por essa razão, foi necessário a criação de ferramentas capazes de resolver estruturas proteicas com eficácia e rapidez.

Atualmente, diversos recursos computacionais vêm sendo desenvolvidos e cada vez mais utilizados com intuito de diminuir gastos e poupar tempo nas pesquisas. Uma ferramenta amplamente empregada é modelagem molecular por homologia (ou comparativa) que é capaz de gerar modelos de estruturas proteicas, sendo útil para diversos tipos de estudos incluindo o desenvolvimento de novos fármacos (HILLISCH; PINEDA; HILGENFELD, 2004). A origem dela se deu após a elaboração de programas e servidores capazes de calcular estruturas de moléculas aliado com a diminuição de custo da maquinaria de alta performance necessária para a execução dos programas (BARREIRO et al., 1997).

A premissa da modelagem por homologia reside no fato de que proteínas com estruturas primárias semelhantes à proteínas homólogas (moldes) experimentalmente resolvidas resulta em similaridade na estrutura e função, principalmente se a sequência tiver uma homologia maior que 50 % (CHOTHIA; LESK, 1986). Além disso, proteínas que são consideradas homólogas possuem regiões bem conservadas, sendo que as maiores diferenças estão na parte externa, geralmente compostas por alças (SANTOS FILHO; ALENCASTRO, 2003).

A modelagem por homologia de proteínas pode ser feita seguindo quatro passos sucessivos (Figura 10). O primeiro deles é a identificação e seleção das proteínas-moldes (*templates*), aquelas que possuem sua estrutura resolvida e depositada no banco de dados. Em seguida alinha-se a sequência alvo (sequência da proteína a ser modelada) com os *templates*, e obtém-se um modelo inicial com as coordenadas preditas. Por fim, valida-se esse modelo gerado utilizando servidores de avaliação de qualidade do modelo tais como PROCHECK e MolProbity (CHEN et al., 2010; ESWAR et al., 2014; LASKOWSKI et al., 1993; SANTOS FILHO; ALENCASTRO, 2003; SCHWEDE et al., 2003; WEBB; SALI, 2017).

Figura 10 – Etapas da modelagem por homologia



Fonte: Eswar et al. (2014), traduzido.

Há diversos servidores disponíveis capazes de realizar modelagem de proteínas seguindo diferentes métodos. Os servidores mais conhecidos são o MODELLER (ESWAR et al., 2014), I-TASSER (ROY; KUCUKURAL; ZHANG, 2010; YANG; ZHANG, 2015) e SWISS-MODEL (SCHWEDE et al., 2003). A abordagem dos servidores MODELLER e SWISS-MODEL são descritos como sendo de montagem de fragmentos rígidos, e para uma boa predição é necessário que pelo menos uma estrutura homóloga da proteína em estudo tenha sido resolvida experimentalmente (SCHWEDE et al., 2003).

Por outro lado, o I-TASSER foi, a princípio, desenvolvido para modelagem de estrutura de proteína por simulações de montagem de segmentação iterativa, uma metodologia chamada de predição de enovelamento (*threading*) considerada intermediária entre a modelagem por homologia e a *ab initio*, onde geralmente a escolha dos melhores moldes é dada por uma biblioteca de fragmentos para obtenção do modelo final (VERLI, 2014; YANG; ZHANG, 2015). Por essa razão, o MODELLER e SWISS-

MODEL são servidores que se destacam na modelagem de proteínas homologas e, por outro lado, o I-TASSER sobressai quando a identidade da sequência é menor que 30 %.

Os preditores de estruturas proteicas são avaliados bianualmente em uma competição denominada *Critical assessment of methods of protein structure prediction* (CASP), a qual teve o I-TASSER como melhor servidor desde a 7ª edição até a 14ª, ocorrida em 2020 (YANG; ZHANG, 2015; PROTEIN STRUCTURE PREDICTION CENTER, 2021). Por ter essa reputação, desde seu lançamento ao público em novembro de 2006 até meados de abril do corrente ano, o I-TASSER já gerou mais de 600 000 modelos de 144 000 usuários de 154 países distintos (I-TASSER *web server*, 2021)

Em contraste, o *Protein Data Bank* (PDB) possui atualmente cerca de 177009 estruturas de macromoléculas biológicas depositadas sendo 173325 estruturas proteicas (PROTEIN DATA BANK, 2021) com uma média nos últimos 10 anos de aproximadamente 8600 estruturas sendo depositadas a cada ano. O número quase 3.5 vezes maior de estruturas geradas pelo I-TASSER em apenas 12 anos de lançamento mostra a importância dessa ferramenta em se estudar estruturas proteicas e seu comportamento em meio biológico, uma vez que a obtenção da estrutura cristalográfica é por vezes onerosas e, a depender da proteína, pode ser bastante complexa (BERMAN et al., 2000).

Em suma, os programas de preditores de estruturas proteicas vem se tornando uma ferramenta cada vez mais poderosa na construção de modelos *in silico* que muitas vezes não possuem estruturas cristalográficas disponíveis, e dessa forma permite que diversos estudos sejam realizados e informações pertinentes a respeito da molécula sejam obtidas.

1.2.2 Dinâmica Molecular

A dinâmica molecular (DM) é uma ferramenta capaz de estudar a física local das moléculas tais como distorção, flexibilidade e estabilização, sendo capaz de melhorar a orientação das cadeias laterais da molécula e reproduzir sua estabilidade estrutural tal qual ocorre no sistema biológico (PRONK et al., 2013). Além disso, as simulações de DM podem ser úteis na extração mais facilitada de informações específicas de um sistema modelo em relação à um sistema real por meio de experimentos (KARPLUS; MCCAMMON, 2002).

Na DM, as moléculas são tratadas como uma coleção de partículas unidas por forças elásticas ou harmônicas, descritas por forças newtonianas, baseado na Mecânica

Clássica (NAMBA; SILVA; SILVA, 2008). Como a DM tenta realocar os átomos seguindo as leis do movimento de Newton, essa ferramenta vem sendo amplamente utilizada para refinamento a nível atômico de estruturas de proteínas (ZHANG; LIANG; ZHANG, 2011). A determinação da trajetória seguida pelos átomos do sistema em estudo é dada pela solução da equação de movimento que regem essas leis, a qual pode ser descrita para um átomo por:

Segunda lei de Newton:

$$F_{x_i} = \frac{d^2 x_i}{dt^2} m_i = \frac{\Delta v_i}{\Delta t} m_i = a_i m_i \quad (1)$$

Onde F_{x_i} é a força aplicada à partícula i no instante de tempo t e posição x , e a_i é a aceleração do átomo i de massa igual a m_i e v é a sua velocidade. O fornecimento de calor é o responsável por dar início ao sistema regido pela Segunda Lei de Newton, quando a energia fornecida ocasiona a aceleração e movimentação dos átomos (VERLI, 2014).

Uma questão que deve ser levado em consideração é o fato de, por se tratar de macromoléculas, grande parte dos átomos que constitui o sistema estão ligados à vários outros tipos de átomos, estando submetidos às forças intra-atômicas e intermoleculares. Portanto, se faz necessário a representação desse sistema por uma função matemática capaz de descrever as forças envolvidas nas interações intra-atômicas e intermoleculares. Dessa forma, o cálculo das forças é então realizado por um modelo denominado campo de força, cuja função também reside no ordenamento das geometrias em que uma rede de átomo se situa (VERLI, 2014). Um exemplo é o campo de força Amber (PONDER; CASE, 2003):

$$V(r^N) = \sum_{\text{ligações}} k_b (l - l_0)^2 + \sum_{\text{ângulos}} k_a (\theta - \theta_0)^2 + \sum_{\text{torsões}} \sum_n \frac{1}{2} V_n [1 + \cos(n\omega - \gamma)] + \sum_{j=1}^{N-1} \sum_{i=j+1}^N f_{ij} \left\{ \epsilon_{ij} \left[\left(\frac{r_{0ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{r_{0ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right\} \quad (2)$$

O primeiro termo observado na equação (somatório sobre as ligações) representa as energias envolvidas entre os átomos ligados covalentemente; k_b é a constante elástica das forças em equilíbrio; o comprimento da ligação entre os átomos i e j é dado por l e l_0 é comprimento da ligação de equilíbrio (obtidos a partir de cálculos quânticos). A

descrição dessa energia se dá por um potencial harmônico, cujo comportamento se assemelha à uma mola e é considerada uma boa aproximação do equilíbrio de ligação.

O segundo termo diz respeito dada à geometria dos orbitais eletrônicos compreendidos na ligação covalente, onde k_a é a constante elástica, θ é o ângulo de ligação e θ_0 corresponde ao ângulo de equilíbrio.

O terceiro termo da equação refere-se à energia do torssionamento das ligações, levando em conta a ordem das ligações, na qual V_n é a barreira de energia para torção, n representa os valores máximo/mínimo de uma torção completa, e ω_n é o ângulo do diedro, e γ o ângulo de fase (0° a 180°). Essa energia é expressa por uma série de *Fourier*.

O quarto e último termo trata-se de um somatório duplo da energia envolvida nas interações de pares de átomos não-ligados. Os termos com expoente de ordem 6 e 12 descrevem a energia envolvida nas interações de van der Waals, onde ϵ_{ij} é a distância entre as barreiras atrativas e repulsivas, σ_{ij} representa a distância na qual o potencial entre as partículas é zero. Juntos, esses termos são denominados de potencial de Lennard-Jones. O último termo traduz a energia envolvida nas interações eletrostáticas, onde q_i e q_j retratam a magnitude das cargas pontuais de cada átomo, r_{ij} é a distância entre as cargas i e j , ϵ_r e ϵ_0 representam a constante dielétrica e a permissividade do meio, respectivamente.

As simulações de DM vêm se tornando cada vez mais aprimoradas com o desenvolvimento de melhores campos de forças que permitem estudar escalas de tempo extremamente pequenas, tal como se observa no enovelamento proteico (ABRAHAM et al., 2015). Para proteínas já enoveladas, os campos de forças disponíveis apresentam boa precisão. O campo de força AMBER ff99SB-ILDN (LINDORFF-LARSEN et al., 2010) se destaca por duas importantes razões: melhora a predição de enovelamento *ab initio* além de aperfeiçoar as torções dos resíduos de isoleucina, leucina, ácido aspártico e asparagina, permitindo melhor concordância com dados estruturais de NMR.

Para execução de DM propriamente dita destaca-se o GROMACS, ferramenta que vem amplamente sendo utilizada para estudos de moléculas biológicas (VAN DER SPOEL et al., 2005). O GROMACS é um *software* livre usado principalmente nas simulações de macromoléculas biológicas tanto em membrana como em meio aquoso (VAN DER SPOEL et al., 2005).

1.2.3 Ancoragem molecular

A ancoragem molecular (*molecular docking*, em inglês) baseado em estruturas proteicas conhecidas - ou modelos gerados por programas de modelagem - tem sido uma interessante abordagem nas fases iniciais do desenho de novos fármacos (PAGADALA; SYED; TUSZYNSKI, 2017; SHOICHET et al., 2002). Os programas usados para executar a ancoragem molecular tem a habilidade de predizer conformações de pequenas moléculas em um sítio de ligação definido (FERREIRA et al., 2015).

Diversos programas de ancoragem molecular estão atualmente disponíveis. Uma análise com dez deles conclui que nenhum programa de ancoragem tem uma vantagem dominante sobre os outros, sendo portanto, o ideal a combinação de diferentes ferramentas de ancoragem em uma única plataforma (WANG et al., 2016). A função de pontuação (*função scoring*) do AutoDock Vina, um *software* de código-livre que tem sido utilizado nos trabalhos do grupo com resultados satisfatórios, leva em consideração a soma de interações estéricas, interação hidrofóbica entre átomos hidrofóbicos e pontes de hidrogênio de forma efetiva, bem como o número de ligações rotativas ativas (COSTA et al., 2015; DA SILVA et al., 2019; TROTT; OLSON, 2009).

Uma barreira encontrada na busca de novos fármacos reside na disponibilidade de um banco de dados amplo e gratuito para execução da triagem virtual de compostos (em inglês, *virtual screening*). O banco de dados do ZINC (IRWIN et al., 2012) possui uma grande diversidade de moléculas, tais como metabólitos, fármacos e produtos naturais, com suas respectivas estruturas tridimensionais em formatos para uso nos principais programa de ancoragem (IRWIN; SHOICHET, 2014).

A maioria dos novos fármacos são obtidos a partir de produtos naturais ou de derivados de produtos naturais. Esses compostos vêm sendo usados na busca de potenciais tratamentos tanto para doenças infecciosas ou não infecciosas, como doenças fúngicas, virais, câncer e osteoporose (CHIKHALE et al., 2020; COSTA et al., 2015; FERREIRA et al., 2021; LIU et al., 2021; OPO, 2021).

Existem diversas vantagens em se usar produtos naturais no desenvolvimento de um novo fármaco. Uma dessas vantagens é a enorme gama de estrutural e química que eles possuem. Provavelmente devido a uma estrutura mais complexa, os produtos naturais exibem característica de interação avançadas quando comparadas com produtos sintéticos. Além disso, esses compostos tendem a possuir maiores pesos moleculares,

sendo portanto importantes fontes de novas moléculas terapêuticas (LAHLOU, 2013; MÜLLER-KUHRT, 2003).

1.2.4 Predição *in silico* de parâmetros ADMET

Uma das etapas do desenvolvimento de novos fármacos compreende os estudos dos perfis de Absorção, Distribuição, Metabolização, Excreção e Toxicidade (ADMET) de um potencial fármaco. A indústria farmacêutica gasta milhões de dólares no desenvolvimento de novos medicamentos, e na fase clínica do estudo muitos falham por pK indesejável ou pela toxicidade acima do tolerável (HONORIO; MODA; ANDRICOPULO, 2013; SERTKAYA et al., 2016). Por essa razão, selecionar bons candidatos ainda no início da pesquisa é fundamental.

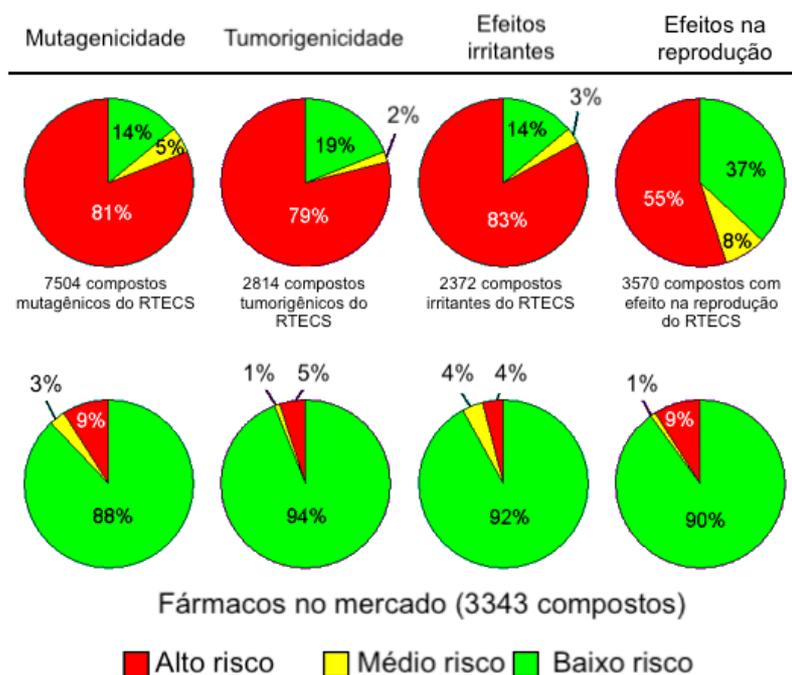
Entretanto, a avaliação *in vitro* das propriedades de ADMET ainda acarreta em elevados gastos de recursos financeiros e tempo (JIA et al., 2020). Por esse motivo, diversas ferramentas *in silico* estão sendo desenvolvidas. Algumas dessas ferramentas se baseiam na relação entre fragmentos de fármacos conhecidos e fragmentos de moléculas que não são fármacos (incluindo compostos tóxicos conhecidos), classificando com baixo potencial aquelas moléculas com muitos fragmentos desse segundo grupo (TIAN et al., 2015; URSU; OPREA, 2010).

O programa *Java-based OSIRIS property explorer* (<https://www.organic-chemistry.org/prog/peo/>) é um preditor *in silico* dos parâmetros de ADMET que usa essa estratégia para definição de uma molécula promissora ou não no desenvolvimento de novos fármacos. Diversos estudos recém realizados utilizaram essa ferramenta de predição com a finalidade busca de potenciais fármacos, indicando ser uma estratégia promissora (DIVYASHRI et al., 2021; OSMAN et al., 2021; RASHID, 2020; SARFRAZ et al., 2021).

O OSIRIS é capaz de avaliar a probabilidade do ligante em se tornar um potencial candidato a fármaco com base na avaliação teórica de riscos toxicológicos e índices de *druglikeness* e *drug-score*. As informações de cada métrica descritas a seguir foram retiradas do próprio site do programa (<https://www.organic-chemistry.org/prog/>).

A avaliação do risco de toxicidade teórica predita pelo programa foi avaliada submetendo um conjunto de compostos tóxicos e um conjunto de compostos provavelmente não-tóxicos. Os resultados do teste podem ser observados na Figura abaixo:

Figura 11 – Avaliação da predição teórica de toxicidade pelo OSIRIS



Na primeira linha está a avaliação de compostos sabidamente com algum risco (mutagenicidade, tumorigenicidade, efeitos irritantes e efeitos na reprodução). Na segunda linha está a predição a partir de fármacos disponíveis no mercado. Adaptado de: <https://www.organic-chemistry.org/prog/peo/tox.html>

É possível inferir uma boa confiabilidade do servidor para os quatro parâmetros de risco avaliados. Por exemplo, para o parâmetro de mutagenicidade o servidor errou apenas na classificação de 14% dos compostos, sendo eficaz na predição em 86% deles (médio e alto risco).

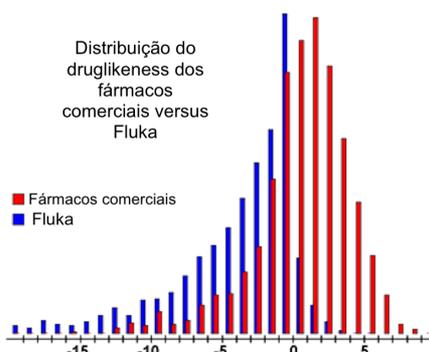
O índice de *druglikeness* é baseado em uma lista de 5300 fragmentos estruturais distintos e com pontuações associadas. A lista foi construída a partir da quebra de ligações rotáveis de 3300 fármacos comerciais e 15000 compostos não semelhantes a fármacos (em inglês, *non-drug-like*) do conjunto Fluka. A pontuação associada a cada fragmento está relacionada ao logaritmo do quociente da frequência em cada um desses fragmentos nesses bancos de compostos. O valor *druglikeness* (d) é dado então pela somatória da pontuação associada a cada fragmento (V_i) dividido pela raiz do número de átomos (n):

$$d = \frac{\sum V_i}{\sqrt{n}} \quad (5)$$

O histograma apresentado na Figura 12 mostra a distribuição dos valores de *druglikeness* calculados do Fluka e dos fármacos comerciais. Ele apresenta cerca de 80% dos fármacos disponíveis no mercado com valores positivos de *druglikeness*, enquanto a

maioria do Fluka estava em valores negativos, sugerindo dessa forma que um potencial fármaco possua valor de *druglikeness* positivo.

Figura 12 – Distribuição do valor de *druglikeness* predito pelo OSIRIS

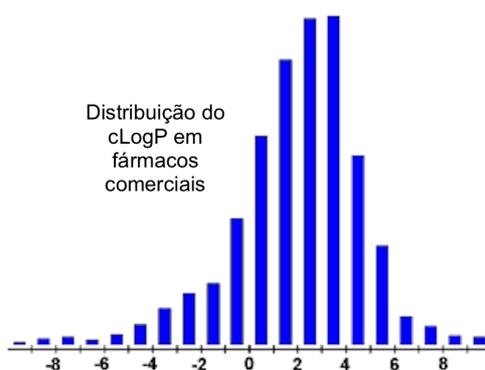


A distribuição dos valores de *druglikeness* calculado pelo servidor OSIRIS indica que moléculas com potencial farmacológico possuem valor de *druglikeness* maior que 0. Adaptado de: <https://www.organic-chemistry.org/prog/peo/druglikeness.html>.

Por fim, o *drug-score* é uma métrica do OSIRIS que combina os resultados do *druglikeness*, cLogP (lipofilicidade), logS (solubilidade), peso molecular e riscos de toxicidade preditos pelo próprio programa. A escala varia de 0 a 1, sendo 1 considerado o melhor cenário para um potencial fármaco.

O valor de logP é uma medida bem estabelecida da hidrofiliicidade do composto. Ele é dado pelo o logaritmo do coeficiente de partição entre n-octanol e água – ($\log c_{\text{octanol}}/c_{\text{água}}$). Baixa hidrofiliicidade e alto valor de logP causa pouca absorção. Assim, tem sido mostrando que compostos para ter uma boa probabilidade de serem bem absorvidos devem ter logP menor que 5.0. No diagrama da Figura 13 podemos ver esse fato:

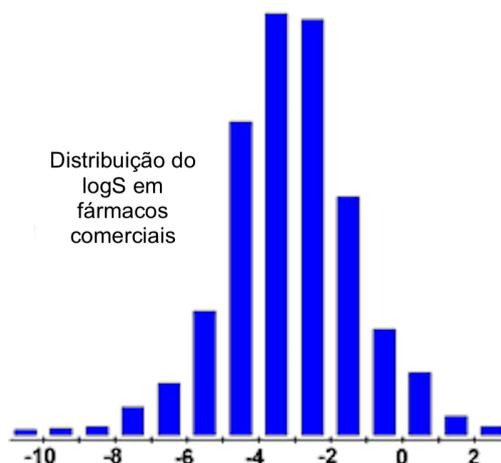
Figura 13 – Distribuição do valor de clogP de fármacos comerciais



Distribuição dos valores de clogP preditos pelo OSIRIS para fármacos comerciais. Nota-se que a maioria dos fármacos possuem clogP < 5. Adaptado de: <https://www.organic-chemistry.org/prog/peo/cLogP.html>

A solubilidade aquosa afeta diretamente a absorção e distribuição de um fármaco pelo organismo. Normalmente, baixa solubilidade está relacionado com má absorção. O valor estimado de $\log S$ é logaritmo da solubilidade medida em mol por litro (mol/L). No diagrama da Figura 14 é possível observar mais de 80% das drogas no mercado tem um valor estimado de $\log S$ maior que -4.

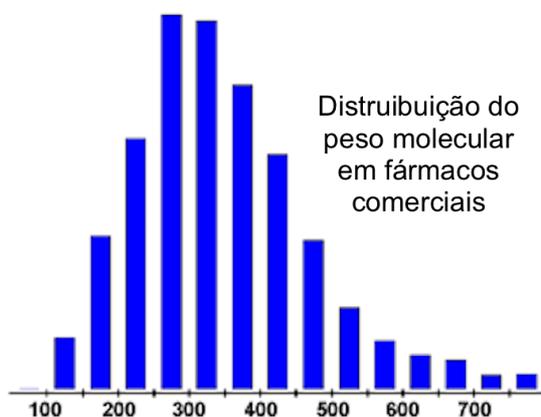
Figura 14 – Distribuição do valor de $\log S$ de fármacos comerciais



Distribuição dos valores de $\log S$ preditos pelo OSIRIS para fármacos comerciais. Nota-se que a maioria dos fármacos possuem $\text{clogP} > -4$. Adaptado de: <https://www.organic-chemistry.org/prog/peo/logS.html>

No desenvolvimento de fármacos, a otimização de compostos com alta atividade biológica se dá geralmente com aumento do peso molecular por meio de adição de fragmentos. Entretanto, compostos com alto peso molecular possuem mais dificuldade de ultrapassar as barreiras da membrana biológica, e, portanto, são menos prováveis de serem absorvidas (Figura 15). Logo, é indicado manter os pesos moleculares dos compostos o mais baixo possível.

Figura 15 – Distribuição do peso molecular de fármacos comerciais



Distribuição dos pesos moleculares de compostos disponíveis comercialmente. Compostos com alto peso molecular não possuem potencial para se tornar fármaco. Adaptado de: <https://www.organic-chemistry.org/prog/peo/mw.html>

Para tanto, diante das informações expostas, percebe-se que o uso das metodologias *in silico* se apresenta como uma grande aliada na busca de fármacos contra doenças ainda sem tratamento farmacológico disponível. Ressalta-se ainda que elas não visam as substituições dos testes *in vitro* e *in vivo*, mas uma etapa a mais no direcionamento dos próximos passos da pesquisa. As ferramentas, além de inovadoras, podem auxiliar na redução de tempo e custos inerentes ao projeto.

2. JUSTIFICATIVA

A dengue, Zika e febre amarela são doenças provocadas por *Flavivirus* que causam problemas de saúde pública no Brasil e em grande parte do mundo, em especial nos países subdesenvolvidos. Por essa razão, há pouco interesse da indústria farmacêutica em desenvolver fármacos para tratamento delas. Dessa forma, essas doenças também podem ser denominadas de doenças negligenciadas.

Nenhuma delas há tratamento farmacológico específico, sendo apenas possível o tratamento sintomático, ou seja, com intuito de amenizar os sintomas e não a causa deles. O problema é agravado para o ZIKV e DENV, pois não há vacina disponível ainda. Apesar da disponibilidade de uma vacina contra YFV, ainda se observa surtos como o que ocorreu em 2017 no estado de Minas Gerais.

A proteína NS1, objeto de estudo deste trabalho, possui importantes funções na patogênese viral. Por estar conservada nos *Flavivirus*, e possuir alta similaridade estrutural entre espécies do gênero, ela se torna um importante alvo para desenvolvimento de uma terapia gênero-específica.

Sabe-se, entretanto, que as diferenças nas sequências proteicas das espécies podem levar a diferentes comportamentos da molécula quando em solução. Nesse contexto, o estudo do comportamento da NS1 do DENV, YFV e ZIKV, utilizando diferentes trajetórias de dinâmica molecular e a busca de conformações em comum (similares) entre espécies, surge como uma poderosa ferramenta para busca de inibidores da NS1 que possam atuar simultaneamente contra as doenças causadas por esses três vírus.

3. OBJETIVOS

3.1 Objetivo Geral

Determinar moléculas com potencial de inibição, por meio de impedimento estérico, da formação do dímero da proteína NS1 do DENV, YFV e ZIKV possibilitando atuar como um fármaco multialvo.

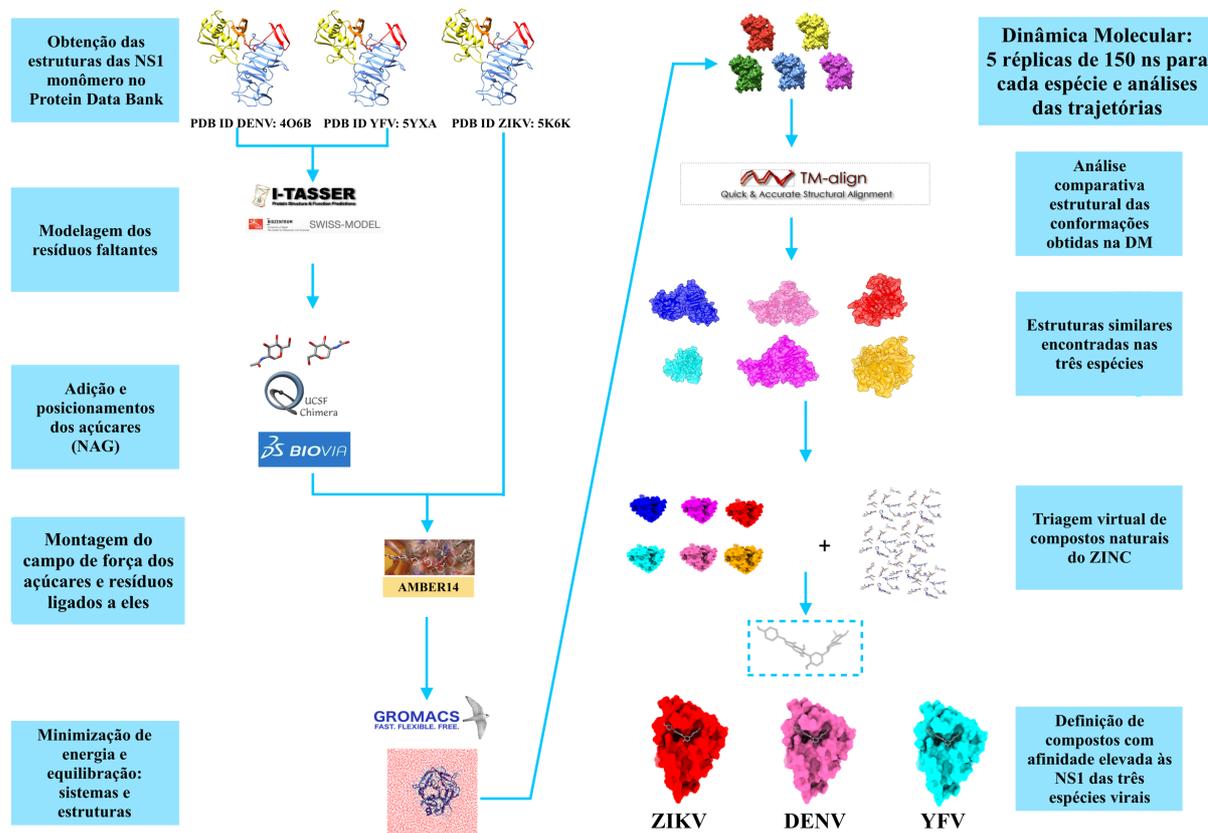
3.2 Objetivos específicos

1. Obtenção das estruturas do monômero da NS1 do DENV, YFV e ZIKV por banco de dados;
2. Modelagem dos resíduos faltantes em servidores de modelagem de proteínas;
3. Execução de dinâmica molecular em réplicas com intuito de aumentar a amostragem de conformações;
4. Agrupamento das conformações obtidas nas trajetórias e seleção das conformações similares encontradas nas três espécies para proceder com a triagem virtual;
5. Realizar triagem virtual com 224 mil compostos naturais do banco de dados do ZINC contra as conformações da NS1 das três espécies;
6. Encontrar compostos que tenham afinidades elevadas pela NS1 das três espécies, a fim de propor um composto com perfil multialvo;
7. Analisar as propriedades *druglikeness*, *drug-score* e predição de toxicidade desses potenciais compostos.

4. METODOLOGIA

O resumo da metodologia geral utilizada no desenvolvimento do trabalho pode ser observado na Figura 16.

Figura 16 – Fluxograma da metodologia básica utilizada no trabalho



4.1 Obtenção das estruturas monoméricas da NS1

A estrutura proteica da NS1 para cada uma das três espécies virais foi inicialmente obtida de acordo com sua disponibilidade no PDB. Dentre elas, a única completamente disponível na forma monomérica e preparada para as simulações foi a NS1 do ZIKV (PDB ID: 5K6K, cadeia B) (BROWN et al., 2016).

Já a estrutura do monômero da NS1 do DENV foi obtida a partir do trabalho realizado por Gonçalves et al. (2019). Nele, para obtenção da estrutura da NS1 do DENV, o servidor SWISS-MODEL foi usado para completar as regiões faltantes da estrutura cristalográfica (PDB ID 4O6B) (AKEY et al., 2014). A sequência submetida no servidor de modelagem foi a mesma da cepa usada para obtenção da estrutura do cristal (GenBank AET43250.1). As glicosilações com D-N-Acetilglicosamina (NAG) nas ASP 130 e ASP 207, bem como as 6 pontes de cisteínas foram adicionadas e tiveram suas coordenadas

ajustadas no *software* Discovery Studios Visualizer (SPASSOV; YAN, 2013). A qualidade do modelo gerado foi avaliado no servidor MolProbity (CHEN et al., 2010).

Para a NS1 do YFV foi realizada a modelagem da proteína por meio do servidor de modelagem I-TASSER. O modelo foi gerado tendo como molde a proteína NS1 do ZIKV (PDB ID: 5K6K). Parte da estrutura da proteína (resíduos 172 a 352) já possui estrutura cristalográfica resolvida e depositada no PDB (PDB ID: 5YXA) (WANG et al., 2017), fazendo com que o servidor se concentrasse em modelar essencialmente os resíduos iniciais (N-terminal) da proteína. O genoma utilizado na resolução do cristal foi o da cepa depositada no GenBank (NO KX268355), e essa mesma sequência aminoacídica do monômero da NS1 do YFV foi usada para submissão no servidor I-TASSER.

Após a obtenção do modelo inicial da proteína NS1, a molécula teve sua qualidade avaliada pelo servidor MolProbity e em seguida foi submetida no servidor ModRefiner (XU; ZHANG, 2011) para melhora da estereoquímica. Uma vez que o modelo gerado não possui os açúcares e as pontes de cisteínas formadas, a proteína foi glicosilada nos resíduos 130 e 208, bem como foram feitas as ligações dissulfeto entre seus 6 pares de cisteína (4-15, 55-143, 179-224, 280-329, 291-312 e 313-316) utilizando o *software* UCSF Chimera (PETTERSEN et al., 2004). O NAG utilizado nas glicosilações foi obtido a partir do cristal da estrutura da NS1 do ZIKV. O procedimento de modelagem e ajustes da estrutura da NS1 do YFV foi executado previamente no trabalho de Menezes et al. (2019).

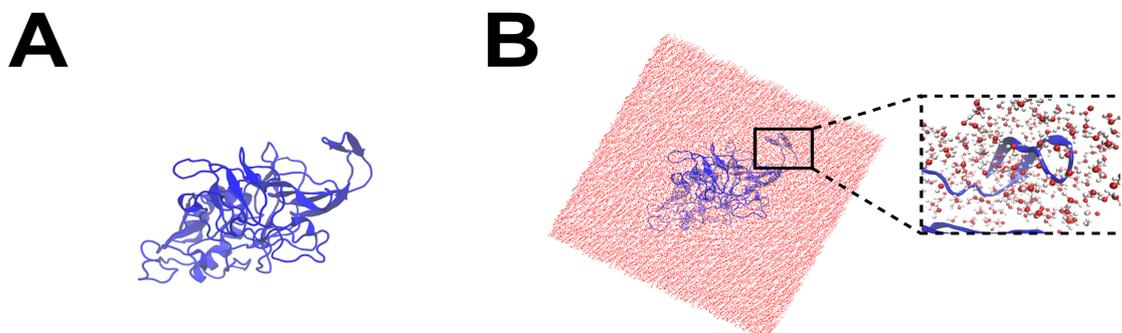
Uma vez obtidas as estruturas da NS1 dos três vírus, a predição da protonação específica da histidina em pH 7.4 foi feita utilizando o servidor PropKa (OLSSON et al., 2011). Em seguida, o *software* AMBER14 (CASE et al., 2014), por meio da ferramenta tleap, foi o responsável por fazer ligação das cisteínas no campo de força AMBERff12SB, assim como foi útil na montagem do campos de força dos açúcares e dos resíduos ligados a ele utilizando o campo de força GLYCAM_06j-1 (KIRSCHNER et al., 2008).

4.2 Simulações de dinâmica molecular

Uma vez glicosilada, as proteínas passaram por simulações de dinâmica molecular usando o *software* GROMACS 5.1.2 (ABRAHAM et al., 2015) no campo de força AMBER99SB-ILDN após ter seu formato convertido utilizando o Acpye (SOUSA DA SILVA; VRANKEN, 2012). O sistema foi neutralizado com íons cloreto (Cl⁻) ou sódio (Na⁺) e uma caixa cúbica foi montada a 10 angstroms (Å) de distância de qualquer átomo

da proteína e preenchida com moléculas de água do tipo TIP3P (Figura 17) (JORGENSEN et al., 1983).

Figura 17 – Sistema NS1 do ZIKV



Sistema montado para simulação da dinâmica molecular da proteína NS1 do ZIKV. (A) Encontra-se a proteína em *cartoon* azul antes da solvatação da caixa. (B) Representação da caixa cúbica com adição da água (*sticks* vermelho e branco) com zoom das moléculas de água em CPK circundando a parte da proteína.

Ligações covalentes do soluto envolvendo todos os átomos foram restringidos através do algoritmo LINCS (HESS, 2008) enquanto o solvente teve sua estrutura rígida interna mantida pelo algoritmo SETTLE (MIYAMOTO; KOLLMAN, 1992). A temperatura do sistema foi ajustada para 310 K (36,85 °C) regulada pelo algoritmo de Berendsen (BERENDSEN et al., 1984) e a pressão de 1 atm pelo barostato Parrinello-Rahman (HUTTER, 2012). Para interações não-covalente um *cutoff* de 1.0 nm foi definido e o método de malha especial da somatória de Ewald (em inglês, *Particle Mesh Ewald*) (ESSMANN et al., 1995) foi usado para o cálculo de interações eletrostáticas de longo alcance. O algoritmo *leap-frog* (HOCKNEY; GOEL; EASTWOOD, 1974) foi usado para integrar as equações de movimento usando 2 fs de tempo de integração.

Inicialmente o sistema foi submetido a duas etapas de minimização de energia. A primeira foi realizada em até 500 etapas ou quando a força máxima fosse menor que 50 kJ/mol/nm usando o algoritmo de minimização do máximo declive (em inglês, *steepest descent minimization*) com a proteína restringida em suas posições. A segunda minimização foi realizada usando o mesmo algoritmo, mas com no máximo 10000 etapas ou quando a força máxima fosse menor que 250 kJ/mol/nm e usando água flexível no lugar de rígida.

Para tanto, depois da etapa de minimização o sistema passou para etapa de equilíbrio. O sistema foi submetido a duas simulações de 100 picossegundos (ps) sendo uma de *ensemble* NVT e outra *ensemble* NPT para o equilíbrio das variáveis

termodinâmicas com a proteína restringida em suas posições. O fim da equilibração se deu com um *ensemble* NPT de 1 nanossegundo (ns) sem restrição da posição da proteína.

A etapa de produção da dinâmica molecular para proteína NS1 foi executada a uma temperatura de 310 K e tempo de 150 ns, sem restrição de conformação. Com o intuito de garantir o perfil de flutuação da proteína NS1, para cada espécie outras quatro réplicas foram submetidas à dinâmica molecular seguindo o mesmo protocolo descrito a partir da etapa de equilibração. Assim, para cada sistema (espécie) de NS1, simulações em quintuplicatas (cinco de cada espécie) foram executadas.

A análise da trajetória foi feita por meio da Raiz do Desvio Quadrático Médio (RMSD) baseada na estrutura inicial da proteína realizada pelo algoritmo gromos como descrito por Daura et al. (1999). Para determinar as conformações que mais estiveram presentes ao longo da trajetória, utilizou-se o programa *g_cluster* do pacote GROMACS.

O perfil da proteína ao longo da simulação de dinâmica molecular foi observado com base na análise de cluster e RMSD. Além disso, a Raiz da Flutuação do Quadrático Médio (RMSF), o qual determina a flutuação média dos resíduos ao longo da trajetória, também foi calculado a partir do pacote de análise do GROMACS. O *software* UCSF Chimera (PETTERSEN et al., 2004) foi utilizado para visualização do comportamento da proteína ao longo da trajetória.

4.3 Análise das cavidades de ancoragem

A definição da cavidade onde as potenciais moléculas inibidoras irão ancorar foi realizada com base em estudos prévios que apontam a região *β -roll* (resíduos 1-28 e 182-216) como uma importante interface de interação monômero-monômero para formação da estrutura funcional dimérica. O trabalho realizado por Gonçalves et al. (2019) apontou essa região como chave no mecanismo de estabilidade da NS1 do ZIKV e DENV. Já o trabalho realizado por Menezes et al. (2019) conferiu a importância dessa região para a NS1 no YFV. Neste trabalho, a mutação de dois aminoácidos (PHE20 e PHE22) alterou significativamente a estabilidade da interface de interação monômero-monômero.

Para encontrar estrutura das cavidades que fossem similares entre as três espécies, as cinco trajetórias em réplicas de cada vírus foram concatenadas, e então uma análise de *cluster* (agrupamento de conformações) apenas da região da cavidade (1-28 e 182-216) com *cutoff* de 0.15 nm foi realizada. As estruturas representativas de cada *cluster*

foram então submetidas para análise de similaridade estrutural no TM-align (ZHANG, 2005).

O parâmetro mais usado para analisar similaridade estrutural é o RMSD. Entretanto, devido o RMSD ponderar as distâncias entre todos os pares de resíduos igualmente, um pequeno número de desvios estruturais locais pode resultar em um alto valor de RMSD, mesmo quando as topologias globais das estruturas comparadas são similares. O TM-score (algoritmo no qual o TM-align se baseia) supera esse problema uma vez que pondera os pares de resíduos em distâncias menores relativamente mais fortes do que aquelas mais distantes. Dessa forma, o TM-score é mais sensível à topologia global do que às variações estruturais. A equação para cálculo do TM-score pode ser descrita como:

$$\text{TM-score} = \text{Max} \left[\frac{1}{L_{\text{Target}}} \sum_i^{L_{\text{ali}}} \frac{1}{1 + \left(\frac{d_i}{d_0(L_{\text{Target}})} \right)^2} \right] \quad (3)$$

Onde L_{Target} é o comprimento da proteína alvo que a outra estrutura está alinhada, L_{ali} é o número de resíduos alinhados e d_i é a distância entre o i -ésimo par de resíduo alinhado. $d_0(L_{\text{Target}}) = 1.24 \sqrt[3]{L_{\text{Target}} - 15} - 1.8$ é o parâmetro de distância que normaliza a distância para que o TM-score não seja dependente do tamanho da proteína (ZHANG, 2005).

Como o programa TM-align faz a análise em pares de estruturas, as estruturas da cavidade da NS1 do ZIKV foram comparadas com as cavidades do YFV, seguida da comparação ZIKV com as estruturas das cavidades do DENV e por fim, foi feita a comparação entre as cavidades da NS1 do YFV e do DENV. Dessa forma, obtiveram-se os valores de TM-score para cada par de análise. Essa análise foi automatizada através de *script* desenvolvido *in house* escrito em linguagem Python.

O valor de TM-score > 0.5 sugere que as proteínas possuem geralmente o mesmo enovelamento. Dessa forma, inicialmente foi usado um *cutoff* de 0.6 para o valor de TM-score para selecionar os pares que resultassem em pelo menos esse valor de similaridade. Como nenhuma análise comparativa gerou esse valor de TM-score, foi-se reduzindo o valor do *cutoff* até encontrar um número considerável de comparações com TM-score igual ou maior que ele. Ao chegar no *cutoff* de 0.57 de TM-score filtrou-se das análises

os pares estruturais da NS1 ZIKV-YFV, ZIKV-DENV e YFV-DENV. As estruturas que estiveram presentes nas três análises após a filtragem foram consideradas similares entre si. Essas estruturas foram então usadas para a triagem virtual de compostos, pois assim aumenta-se a probabilidade de encontrar um fármaco em comum a todas as espécies.

4.4 Triagem virtual de compostos

Os compostos naturais de banco de dados do ZINC foram obtidos através de *download* no site <https://zinc15.docking.org/substances/subsets/natural-products/> em único arquivo de formato SDF. Os compostos foram separados em arquivos únicos e convertidos em formato PDBQT através da automação do programa OpenBabel (O'BOYLE et al., 2011) inserido em um *script* em Python. Em seguida, as estruturas da NS1 foram preparadas usando o *script* `prepare_receptor4.py` do MGLTools (MORRIS et al., 2009) e a obtenção das coordenadas da caixa para delimitação da região de ancoragem foi feita no UCSF Chimera.

A execução da triagem virtual de compostos foi realizada em duas etapas através do programa AutoDock Vina (TROTT; OLSON, 2010). A primeira etapa correspondeu a 01 (uma) simulação por composto com geração de 10 modos de ligação. Já a segunda etapa consistiu em 100 (cem) simulações para cada composto classificado entre os 100 melhores pelo critério de afinidade ou eficiência.

A análise de critério de afinidade se deu com base no valor calculado pela função *escore* do AutoDock Vina. Já o critério de eficiência, proposto por Abad-Zapatero (2007), leva em consideração o número de átomos do ligante. A equação pode ser descrita por:

$$EL = \frac{\Delta G}{n} \quad (4)$$

A função depende do valor de energia de ligação ΔG entre ligante e receptor e o número de átomos total. Conclui-se então que compostos com menores números de átomos e alta afinidade irão possuir maior eficiência. Aqui, o ΔG será então substituído pela função *escore* definido pelo AutoDock Vina assegurando a seleção de ligantes pequenos com estabilidade na cavidade de ligação.

As interações entre ligante e receptor foram visualizadas através do Discovery Studios 2.5 da Accelrys (<https://discover.3ds.com/discovery-studio-visualizer-download>) e a análise de similaridade entre ligantes foi realizada do servidor ChemMine Tools (<http://chemmine.ucr.edu/>) usando a ferramenta de “*Cluster > Hierarchical Clustering*” ou “*Similarity Workbench*” no caso de apenas um par de compostos. Além disso, buscou-

se encontrar fragmentos que interagiram de forma similar no sítio de ancoragem nas diferentes conformações da NS1 de cada espécie. Para isso, foi feita uma análise de distância entre átomo do receptor e átomo do ligante no programa ContPro (FIROZ et al., 2010), e o gráfico da análise foi realizado através do pacote *Alluvial* no RStudio (<https://www.rstudio.com>).

4.5 Análise perfil ADMET

A análise *in silico* das propriedades farmacocinética e toxicológica dos compostos selecionados foi realizada pelo programa OSIRIS Property Explorer (<http://www.organic-chemistry.org/prog/peo/>). Os principais parâmetros utilizados para análise dos compostos foram os riscos toxicológicos (mutagênico, tumorigênico, irritante e reprodutivo) e os escores (*druglikeness* e *drug-score*).

4.6 Análise estatística

Os principais dados avaliados no estudo foram os valores de RMSD das simulações de dinâmica e os escores de energia do AutoDock Vina. Para ambos foram realizados os testes de normalidade: Shapiro-Wilk e Anderson-Darling. O primeiro foi usado para amostras com menos de 5000 dados e o segundo para amostras com mais de 5000. Para todos os casos os dados não estavam em distribuição normal.

Sendo assim, foram aplicados os testes não-paramétricos de Kruskal-Wallis (KRUSKAL; WALLIS, 1952) para comparações múltiplas e Wilcoxon (WILCOXON, 1945) para comparação de duas amostras. As análises foram executadas na versão 3.6.1 do programa R (<https://www.r-project.org>) e os gráficos plotados no RStudio por meio da função *ggboxplot*. As diferenças foram consideradas significantes quando $p < 0.05$.

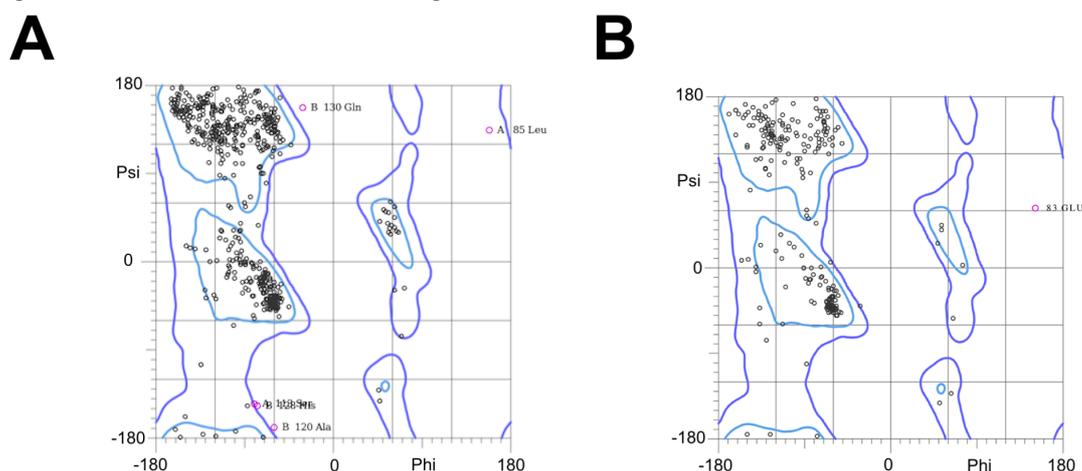
5. RESULTADOS

5.1 Modelagem por homologia

Conforme descrito anteriormente, apenas as proteínas NS1 do DENV e YFV foram submetidas a servidores de modelagem. De acordo com o servidor MolProbity, a modelagem da NS1 do DENV realizada no trabalho de Gonçalves et. al. (2019) no SWISS-MODEL apresentou valores de *clashscore* e MolProbity *score* de 4.29 (96º percentil) e 1.85 (94º percentil), respectivamente. No gráfico de Ramachandran, obteve-se 94.1% dos resíduos em regiões favoráveis e 99.3% em regiões permitidas (Figura 18A).

Já o modelo da NS1 do YFV obtido no trabalho de Menezes et al. (2019) gerado pelo servidor I-TASSER, o mapa de Ramachandran (dado não mostrado) apresentou 82% dos resíduos em regiões favoráveis em relação aos ângulos *phi* e *psi* e 94.3% em regiões permitidas. O modelo então passou por uma etapa de refinamento no servidor MolProbity (CHEN et al., 2010) que permitiu aumentar para 91.7% os resíduos em regiões favoráveis e para 99.1% nas regiões permitidas (Figura 18B). O modelo refinado obteve uma mudança do MolProbity *score* de 2.42 (52º percentil) para 1.00 (100º percentil) e um *clashscore* de 2.21 (99º percentil) para 0 (100º percentil). O RMSD entre as estruturas antes e após refinamento foi de 0.602 Å.

Figura 18 – Ramachandran dos modelos gerados da NS1 do DENV e YFV.

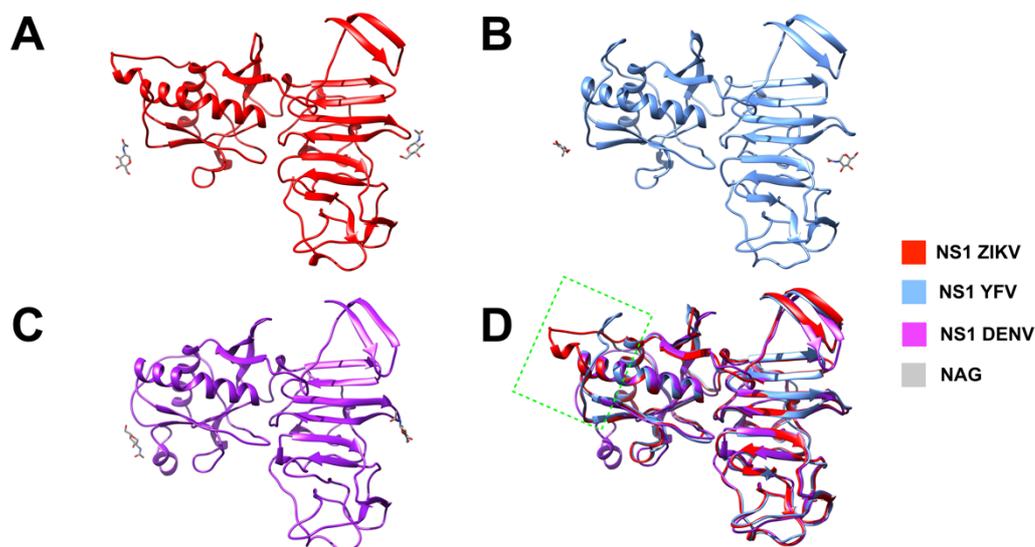


Gráficos de Ramachandran mostrando a posição dos resíduos em relação aos ângulos *phi* e *psi*. As linhas em azul escuro delimitam as regiões permitidas e em azul claro as regiões favoráveis. (A) Avaliação do modelo da estrutura da NS1 do DENV gerado pelo SWISS-MODEL (B) Análise do modelo da NS1 do YFV gerado pelo I-TASSER após refinamento no MolProbity.

As estruturas monoméricas obtidas como descrita anteriormente podem ser visualizadas na Figura 19. Observa-se que as proteínas geradas pelos servidores de modelagem (Figura 19.B-C), quando sobrepostas, mostram-se bastante similares com o

crystal NS1 do ZIKV (Figura 19.D). O RMSD entre a estrutura NS1 de ZIKV sobreposta com YFV e DENV é de 2.137 Å e 4.793 Å, respectivamente. Contudo, nota-se (Figura 19.D) que a maior diferença entre elas está no domínio *wing*, uma região de elevada instabilidade. Quando exclui-se essa região do cálculo do RMSD, estes valores caem para 0.824 Å e 1.603 Å.

Figura 19 – Estruturas monoméricas da proteína NS1

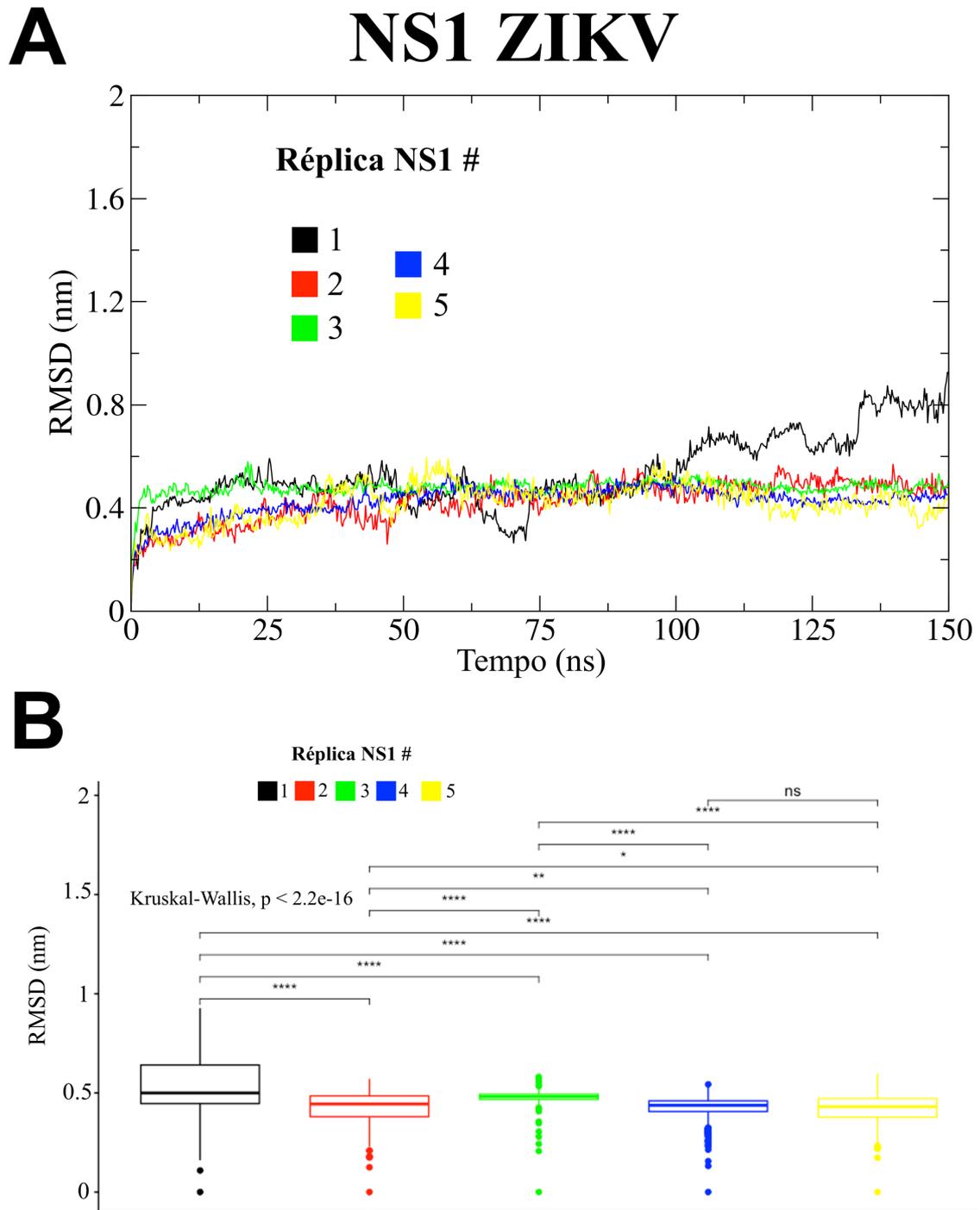


Estruturas monoméricas da proteína NS1 em representação de *ribbon*. Em *stick* cinza está representado o açúcar NAG. Na imagem (D), destaca-se a região do domínio *Wing* com maior divergência entre as estruturas

5.2 Simulações de dinâmica molecular

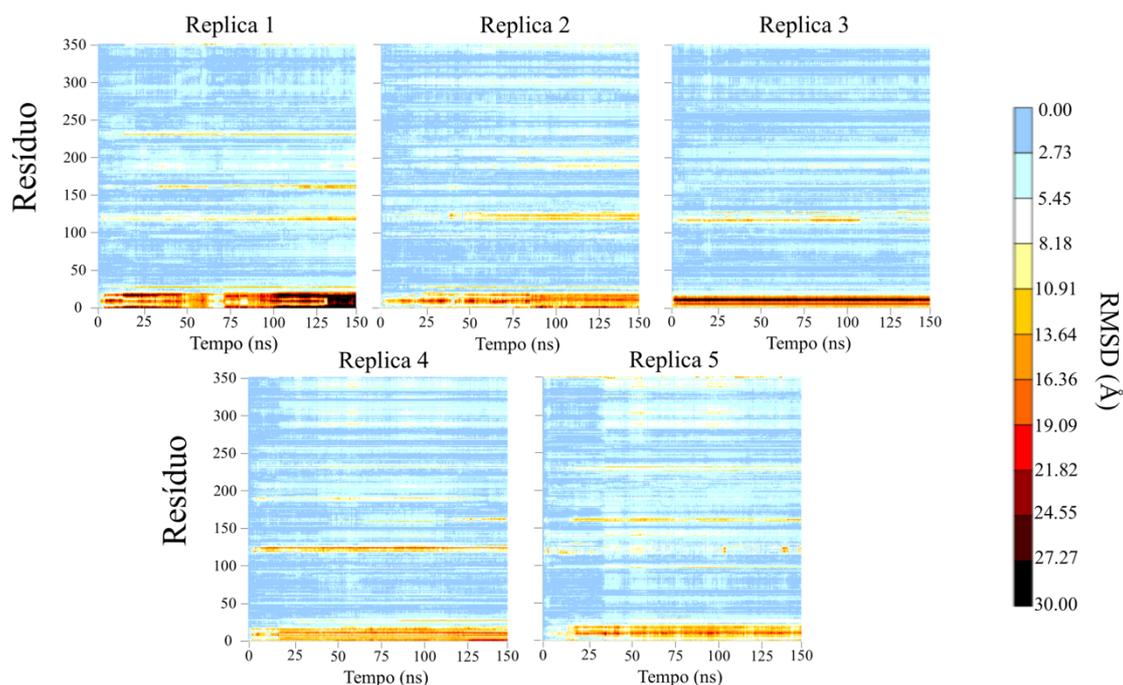
A análise de RMSD entre as réplicas de ZIKV (Figura 20) mostrou diferenças significativas ($p < 0.05$) entre elas, exceto entre as réplicas 4 e 5. Todas as réplicas se estabilizaram em uma faixa entre aproximadamente 0.4 nm e 0.6 nm, com exceção da réplica 1 que teve seu RMSD aumentado após 100 ns para em torno de 0.8 nm causado principalmente pela elevada flutuação da região do β -roll, conforme é mostrado na Figura 21.

Figura 20 – RMSD das réplicas da NS1 do ZIKV



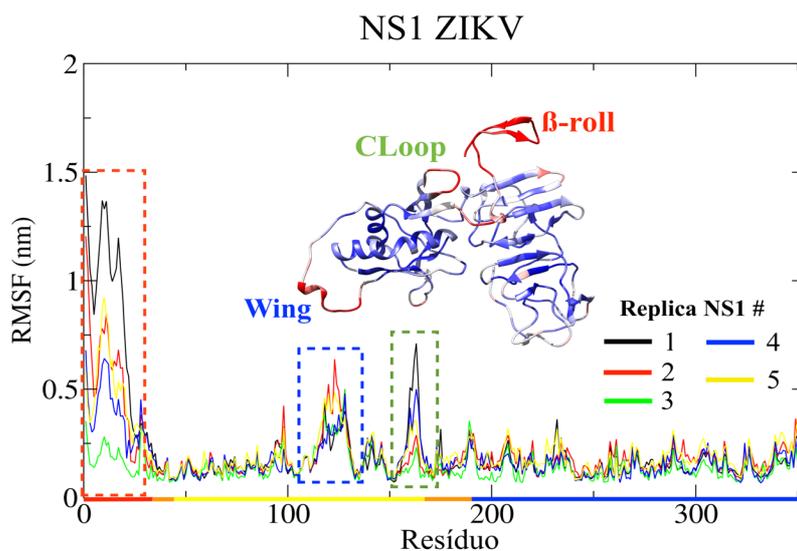
(A) Gráfico RMSD ao longo da simulação de DM em linhas das cinco réplicas da NS1 do ZIKV. (B) Análise dos valores do RMSD em box plot. *: $p < 0.05$; **: $p < 0.01$; ****: $p < 0.0001$; ns: $p > 0.05$

Figura 21 – Mapa de calor do RMSD por resíduo ao longo da trajetória da NS1 do ZIKV



A análise do RMSF (Figura 22) mostra que de todas as regiões instáveis, a única que teve diferença mais aparente entre as réplicas foi na região do β -roll. Certamente, a necessidade do atracamento dessa região com a correspondente do outro monômero, para obtenção da estrutura funcional em forma de dímero, faz com que esse domínio seja mais flexível e instável para facilitar o ajuste da ligação entre eles.

Figura 22 – RMSF das simulações de DM das réplicas da NS1 do ZIKV

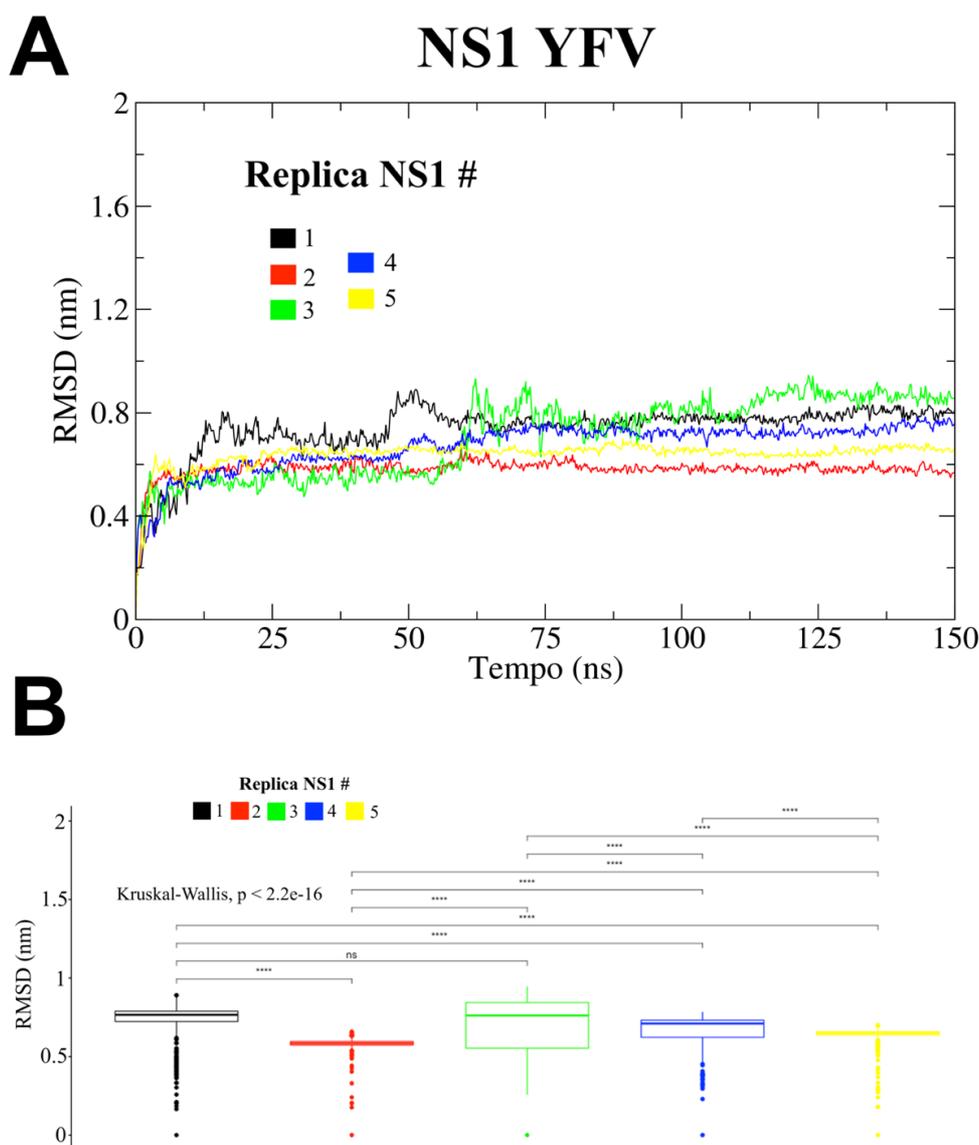


Inserido no gráfico tem-se a estrutura da NS1 do ZIKV apontando para as flutuações observadas no RMSF. As cores dos retângulos pontilhados fazem referência às cores das regiões apontadas na estrutura (β -roll, C-Loop e Wing). Eixo x está colorido de acordo com o domínio da proteína (vermelho – β -hairpin, laranja - conector, amarelo – wing e azul- β -ladder).

As outras regiões que apresentaram maiores flutuações no RMSF compreendem o domínio *wing* e a região denominada de *conector loop* (*CLoop*). Apesar de funções desconhecidas na NS1 monomérica, acredita-se que essas regiões possam desempenhar um papel importante na interação da proteína com membrana do hospedeiro (GONÇALVES et al., 2019)

O perfil de RMSD para as réplicas da NS1 do YFV (Figura 23) apresentou um maior intervalo em relação aos valores de estabilidade, onde variou de 0.55 a 0.85 nm. De forma similar ao observado para a NS1 do ZIKV, duas das réplicas (1 e 3) não apresentaram diferença estatisticamente significativa entre si.

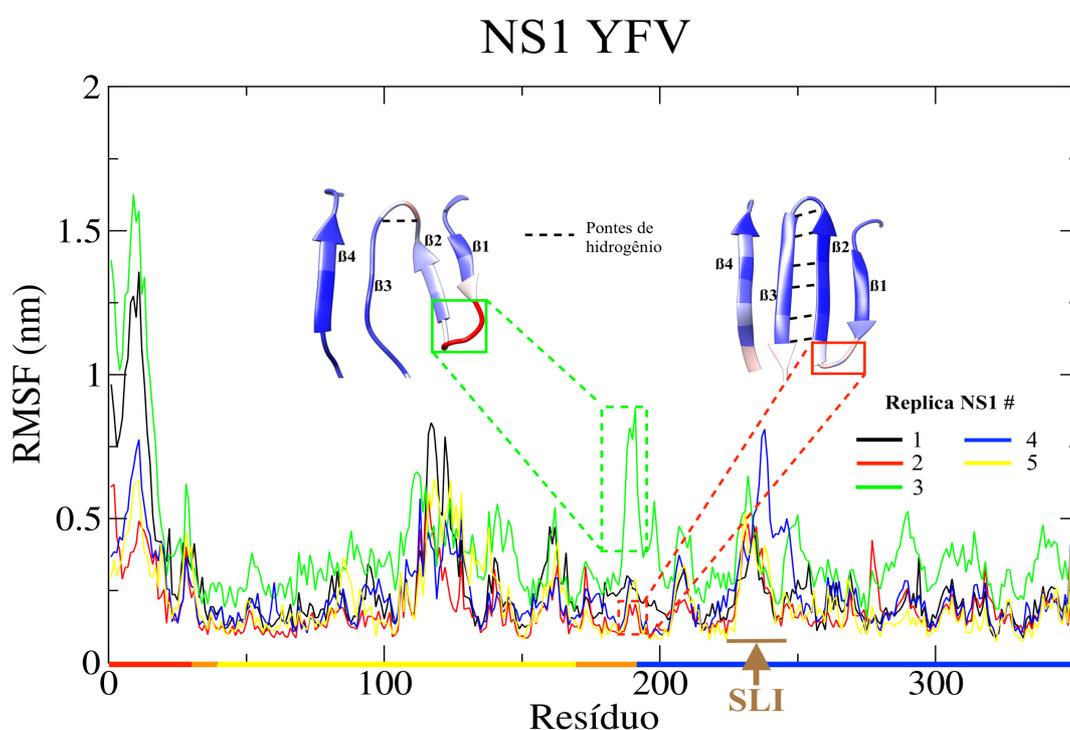
Figura 23 – RMSD das simulações de DM das réplicas da NS1 do YFV



(A) Gráfico RMSD ao longo da simulação de DM em linhas das cinco réplicas da NS1 do YFV. (B) Análise dos valores do RMSD em box plot. ****: $p < 0.0001$; ns: $p > 0.05$

A análise do RMSF mostrou flutuações em regiões similares (Figura 24) ao do ZIKV, como nos domínios β -roll e wing. Por outro lado, foi observado a flutuação elevada na região conhecida como *Spaghetti Loop Interface* (SLI) (resíduos 228 a 238) e não foi observada na região do *CLoop*. Assim como no β -roll, a SLI é uma região de interação intermolecular, o que justificaria a flutuação aumentada. Contudo, deve-se investigar a ausência dessa flutuação na NS1 do ZIKV. Além disso, na réplica 3 foi observada uma flutuação fora do perfil observado para as outras réplicas da NS1 do YFV na região 188-192. Essa região compreende uma das fitas do β -ladder que se estabilizam por meio de pontes de hidrogênio. Provavelmente, a ruptura dessas pontes na replica 3 desencadeou a instabilidade na região.

Figura 24 – RMSF das simulações de DM das réplicas da NS1 do YFV

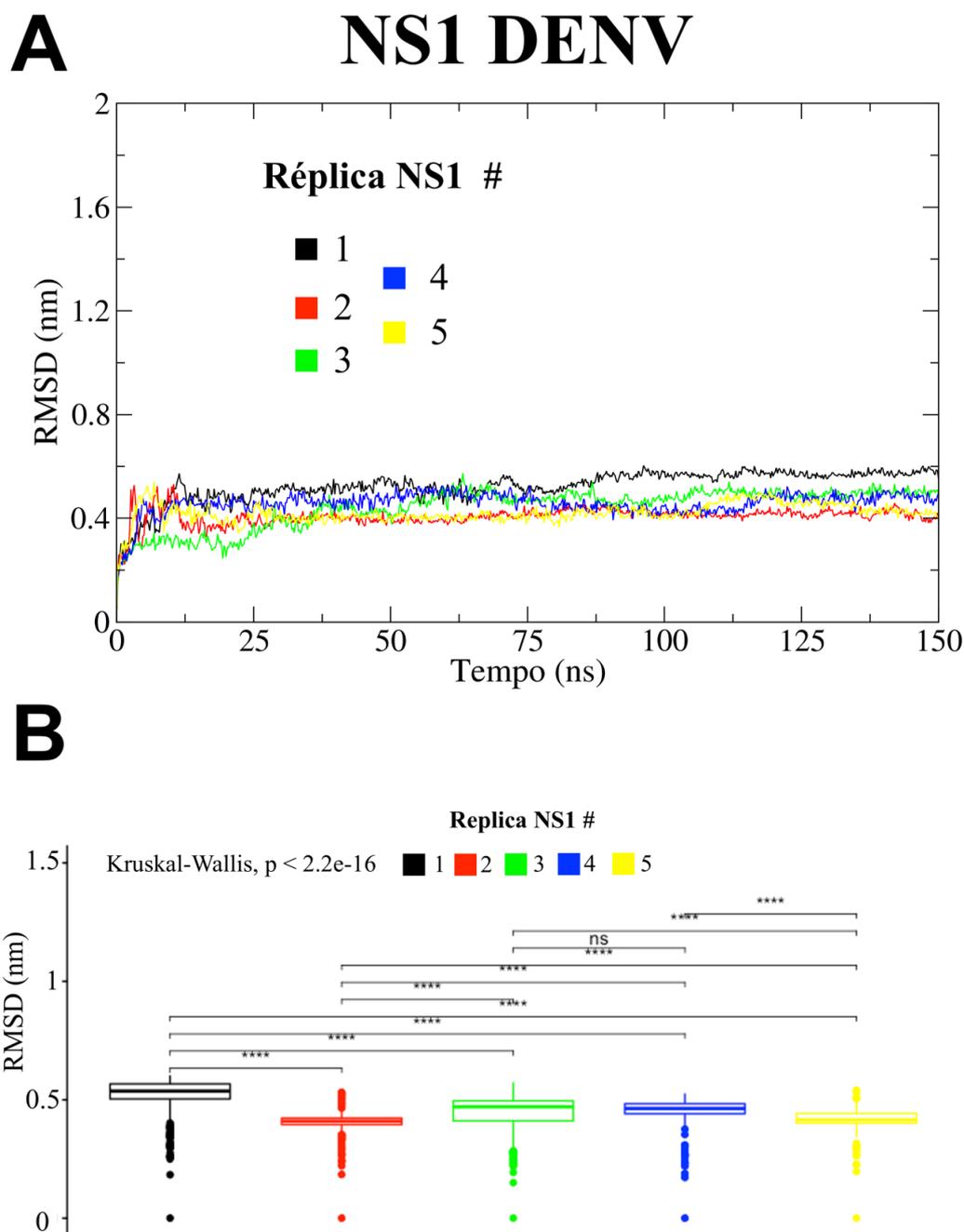


RMSF das réplicas das NS1 do YFV. Representação em *ribbon* apontando a ruptura das ligações de hidrogênio das fitas do β -ladder na replica 3 (verde) em relação a replica 2 (vermelho), causando a instabilidade nessa região. Eixo x está colorido de acordo com o domínio da proteína (vermelho – β -hairpin, laranja - conector, amarelo – wing e azul- β -ladder)

O RMSD das réplicas da NS1 do DENV (Figura 25) se mostrou o mais estável entre as três espécies, com valores após estabilização da proteína entre 0.4 nm e 0.6 nm para todas as réplicas. Como observado para as outras espécies, duas réplicas (3 e 4) não apresentaram diferença nos valores de RMSD estatisticamente significativa entre si.

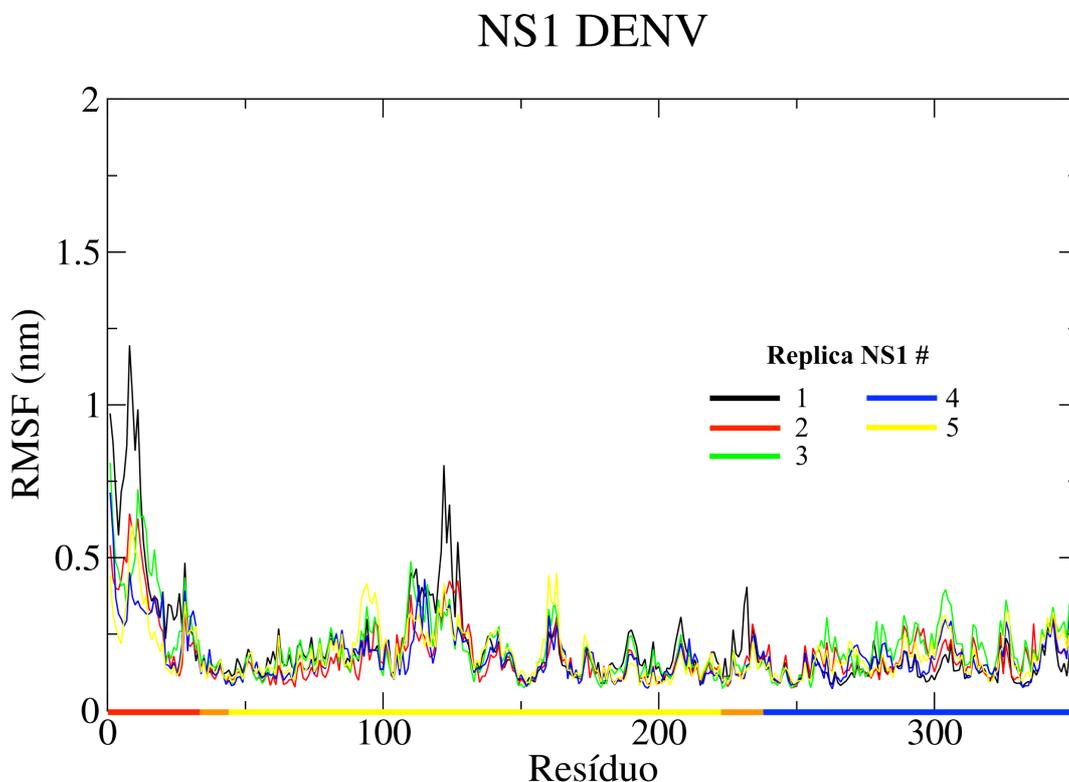
A análise do perfil de RMSF do comportamento da NS1 do DENV em solução (Figura 26), mostrou as flutuações elevadas observadas anteriormente para as duas outras espécies nas regiões do β -roll e domínio *wing*. Curiosamente, uma terceira flutuação presente nas outras espécies (no *CLoop* da NS1 do ZIKV e no SLI da NS1 do YFV) não foi observada para a NS1 do DENV.

Figura 25 – RMSD das simulações de DM das réplicas da NS1 do DENV



(A) Gráfico RMSD ao longo da simulação de DM em linhas das cinco réplicas da NS1 do DENV. (B) Análise dos valores do RMSD em box plot. ****: $p < 0.0001$; ns: $p > 0.05$

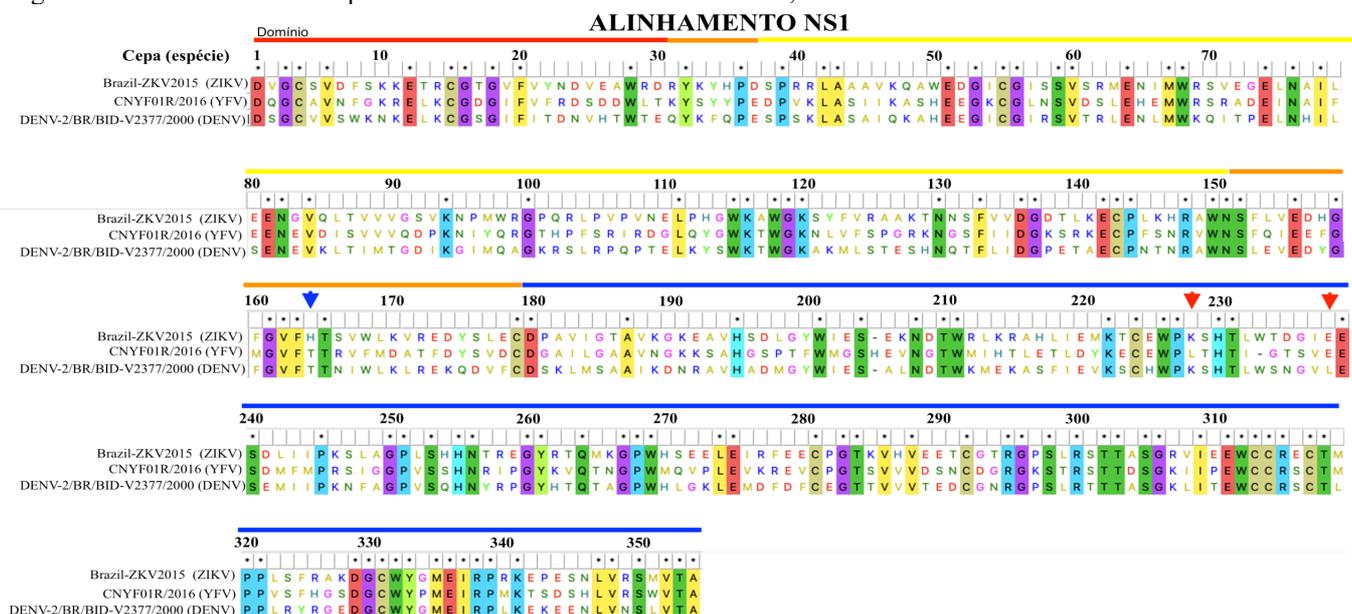
Figura 26 - RMSF das simulações de DM das réplicas da NS1 do DENV



Eixo x está colorido de acordo com o domínio da proteína (vermelho – β -hairpin, laranja - conector, amarelo – wing e azul- β -ladder).

A chave das diferenças observadas entre espécies pode estar na sequência de aminoácido que difere entre elas. Por exemplo, a região do SLI na NS1 YFV difere das proteínas do DENV e ZIKV (Figura 27, delimitado pelas setas vermelhas). Essa composição pode favorecer o aumento da instabilidade nessa região. Já na região do *CLoop* da NS1 do ZIKV, não foi observada nenhuma grande alteração que justifique a diferença, exceto pela substituição da histidina 164 pela treonina no YFV e DENV (Figura 27, seta azul). Análises mais detalhadas devem ser realizadas a fim de confirmar a interferência dessa substituição na NS1 do ZIKV.

Figura 27 – Alinhamento da sequência de aminoácido das NS1 do ZIKV, YFV e DENV



Alinhamento da sequência de aminoácido da NS1 onde colorido estão os aminoácidos conservados, e em branco os aminoácidos divergentes em pelo menos uma espécie. Acima do número dos resíduos tem-se a indicação do domínio de acordo com a cor da linha: vermelho – β -hairpin, laranja - conector, amarelo – wing e azul- β -ladder.

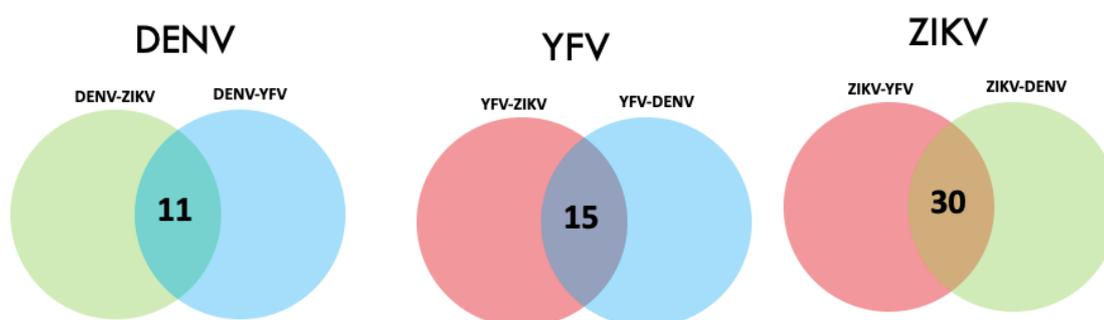
5.3 Análise das cavidades de ancoragem da proteína NS1

Como o objetivo do trabalho é a inibição do processo de dimerização, o foco na busca de cavidades para triagem de compostos foi na região do β -roll. Essa região é a principal entre as superfícies de interação intermolecular na formação do dímero. Sendo assim, com o foco em buscar compostos que se liguem às NS1 das três espécies, foi realizada a análise conformacional dessa região com intuito de buscar conformações similares nas três espécies.

Na análise de *cluster* para um *cutoff* de 0.15 nm com as trajetórias das cinco réplicas concatenadas, e levando em conta apenas a região da cavidade, observou-se 415, 506 e 247 *clusters* para a NS1 do ZIKV, YFV e DENV, respectivamente. As conformações representativas de cada *cluster* foram então submetidas para análise do TM-align. Assim, os pares das conformações da NS1 nas diferentes espécies foram comparados. Inicialmente, foi usado um *cutoff* de 0.6 de valor de TM-score para selecionar os pares com no mínimo esse valor de similaridade. Como nenhuma análise comparativa gerou esse valor de TM-score, foi-se reduzindo o valor do *cutoff* até achar um número considerável de comparações com TM-score igual ou maior que ele. Ao chegar no *cutoff*

de 0.57 de valor de TM-score, encontrou-se 7036 pares de estruturas da NS1 ZIKV e NS1 do DENV, 1338 pares entre NS1 YFV e DENV, e 337 entre NS1 ZIKV e YFV. Nessas análises foram encontradas 11 estruturas da NS1 do DENV concomitantemente similares (TM-score ≥ 0.57) com estruturas dos *clusters* de YFV e ZIKV, 15 conformações da NS1 do YFV foram classificadas similares em comparação com do DENV e ZIKV, e por fim 30 conformações da NS1 do ZIKV similares com DENV e YFV (Figura 28). Esse valor está dentro do valor indicado para proteínas com enovelamento similar (TM-score > 0.5).

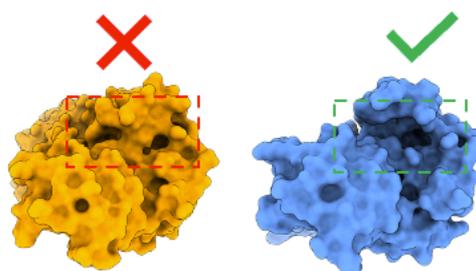
Figura 28 – Número de estruturas em comum nas análises do TM-align



Para cada par de análise no TM-align, buscou-se estruturas que apareceram concomitantemente nos dois pares. Nas análises NS1 DENV-ZIKV e NS1 DENV-YFV, 11 estruturas da NS1 do DENV estavam presentes nas duas análises. Para NS1 YFV-ZIKV e NS1 YFV-DENV, 15 estruturas da NS1 do YFV apareceram nas duas. E na comparação ZIKV-YFV e ZIKV-DENV, 30 estruturas do ZIKV apareceram em comum nas duas análises.

Após essa triagem, uma análise visual da estrutura da NS1 completa para verificar a viabilidade da cavidade foi realizada (Figura 29). Para o DENV, duas conformações foram promissoras, sete para o YFV e quatro para o ZIKV (Tabela 1). Elas foram então submetidas no servidor DoGSiteScorer (VOLKAMER et al., 2012) para análise das suas cavidades. Para os vírus com mais de duas conformações de NS1 selecionadas (YFV e ZIKV), filtrou-se as duas NS1 que apresentaram na cavidade do β -roll maiores *drugscore* de acordo com o servidor. O *drugscore* é uma métrica que varia de 0 a 1 cujo valor leva em consideração o formato, tamanho e hidrofobicidade do sítio, sendo 1 o valor mais promissor para sítio de ancoragem de fármaco. As imagens das estruturas da NS1 selecionadas para a triagem virtual com a cavidade do β -roll em destaque podem ser visualizadas no Apêndice A. As estruturas da NS1 a partir de agora serão descritas somente pela espécie seguido pelo número do *cluster* a qual pertence (por exemplo, DENV 46 ou DENV_46). O alinhamento estrutural entre elas pode ser observado na Figura 30.

Figura 29 – Inspeção visual da viabilidade da cavidade



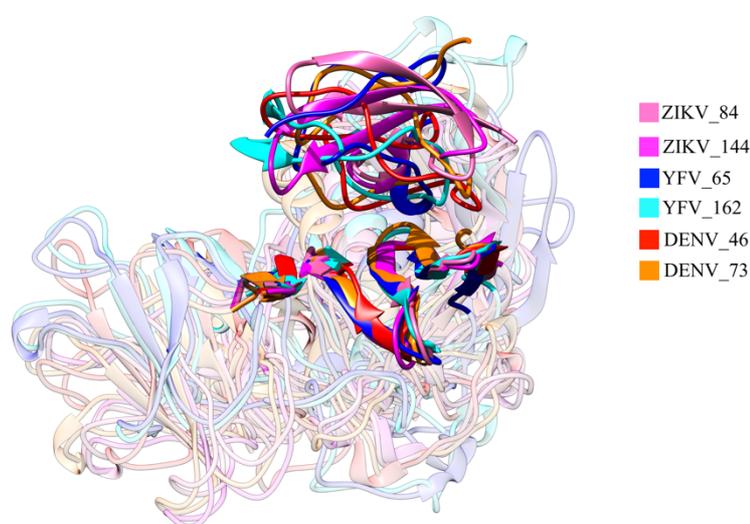
O critério de seleção foi feito com base na viabilidade da cavidade que compreende a região do β -roll. Estrutura sem viabilidade de ancoragem de compostos foram excluídos (estrutura da esquerda, em laranja). Quando a cavidade foi observada como promissora, foi incluída (estrutura da direita, em azul).

Tabela 1 – Estruturas selecionadas após a inspeção visual

Estrutura NS1	Volume (\AA^3)	Área (\AA^2)	<i>DrugScore</i>
DENV 46	681.77	965.98	0.82
DENV 73	2295.81	2628.84	0.81
YFV 9	437.91	768.83	0.7
YFV 11	913.83	1498.04	0.8
YFV 65	796.34	1354.55	0.85
YFV 119	759.95	1218.8	0.81
YFV 135	1422.1	2318.81	0.81
YFV 162	661.45	954.59	0.83
YFV 174	905.15	1317.54	0.8
ZIKV 6	329.02	432.41	0.75
ZIKV 73	1431.81	1595.3	0.81
ZIKV 84	1022.85	1187.18	0.82
ZIKV 144	2877.38	3422.02	0.82

Em negrito estão destacadas as estruturas selecionadas (maiores *drugScores*, sendo duas de cada espécie) para a triagem virtual

Figura 30 – Alinhamento estrutural das seis conformações selecionadas para a triagem virtual



Representação em *ribbon* das seis estruturas da NS1 definidas para a triagem de compostos a ser realizado. Destaque para a região do bolsão (resíduos 1-28 e 182-216)

5.4 Triagem virtual de compostos

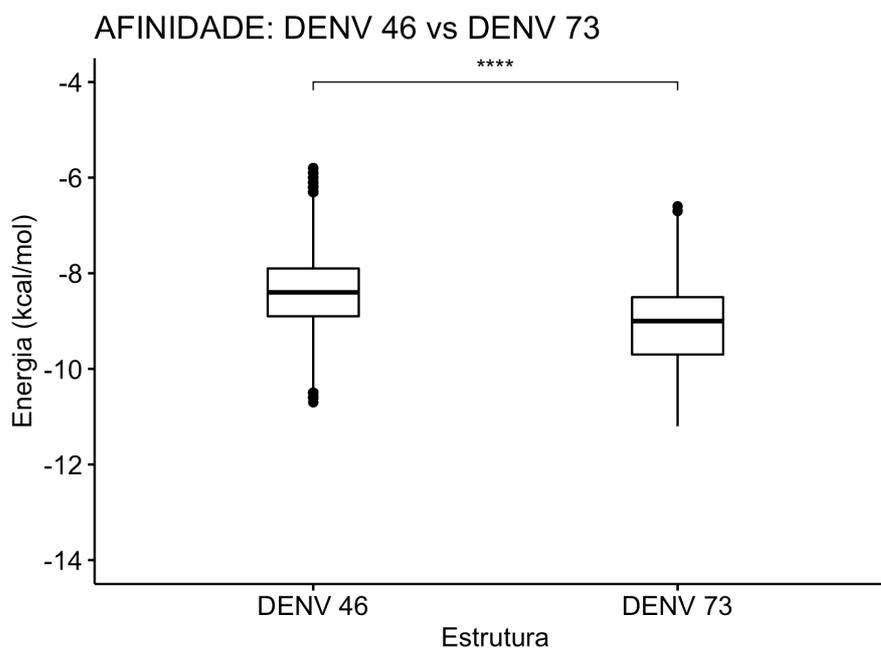
Após a seleção das estruturas da NS1 para execução da triagem virtual, realizou-se a ancoragem dos 224205 mil compostos do catálogo de compostos naturais do banco de dados do ZINC em cada estrutura de NS1 dos *Flavivirus*. Em seguida, esses compostos foram então classificados seguindo dois critérios: afinidade e eficiência.

5.4.1 Critério de afinidade

5.4.1.1 Comparação da triagem virtual entre estruturas da NS1 do DENV

Utilizando critério de afinidade, ou seja, considerando apenas o escore de energia calculado pelo AutoDock Vina como critério de classificação, selecionou-se os 100 compostos com menores energias para cada espécie. Após 100 simulações exaustivas no AutoDock Vina para cada um dos 100 compostos, os 10 modos conformacionais de mais baixa energia de cada composto docado na proteína foram armazenados. Para o DENV, houve diferença estatisticamente significativa ($p < 0.0001$) entre as energias obtidas dos ligantes docados nas duas estruturas selecionadas para a triagem (Figura 31). Para a estrutura do DENV 46 a menor energia observada foi de -10.7 kcal/mol e a maior -5.8 kcal/mol. Já para o DENV 73 esses valores foram de -11.2 kcal/mol e -6.6 kcal/mol. Com TM-score de 0.70576 entre os bolsões, apenas dois compostos dentre os 100 melhores foram comuns às duas estruturas (Apêndice B).

Figura 31 – *Boxplot* das distribuições de energia das 100 simulações exaustivas dos 100 compostos ranqueados de acordo com o critério de energia (afinidade) para as duas estruturas da NS1 do DENV.



**** $p < 0.0001$

O diagrama de interação 2D (Figura 32) apresenta as principais interações entre o ligante e a proteína NS1 do DENV. Nota-se que há uma diferença entre estruturas em relação aos resíduos envolvidos nas interações com cada composto, sendo alguns resíduos chaves (VAL5, SER7, TRP8, LYS9, LYS14, CYX15, GLY16, HID26, ARG192 e GLU203) na interação com todos os compostos. O RMSD entre esses resíduos nas duas estruturas foi de 6.809 Å. Essa divergência nos resíduos envolvidos nas interações se dá pela diferença conformacional entre os sítios de ancoragem de cada estrutura. Nota-se que houve um favorecimento no escore de energia quando envolveu os resíduos SER17 (incluindo formação de ponte de halogênio), ILE19, ILE21, THR22 e ASP23 na estrutura do DENV 73 além de maior número de interações Alkyl ou Pi-alkyl (total de seis) quando comparado com o DENV 46 (total de três). Contudo, nota-se também que muitos resíduos são conservados nas interações independente da estrutura da NS1 do DENV.

observado na Figura 33. Nota-se que os resíduos LYS14, CYX15 e GLY16, que também estavam presentes na análise anterior, apareceram como específicos para uma parte do fragmento do ligante (destacado em vermelho na estrutura da Figura 33). A espessura da linha do diagrama reflete o número de pares observados para cada par de átomo do ligante e resíduo, uma vez que um átomo do composto poderia estar a uma distância $\leq 5\text{\AA}$ de mais de um átomo do aminoácido.

Figura 33 – Análise fragmento-específico do ligante ZINC000070704746 no sítio da NS1 do DENV

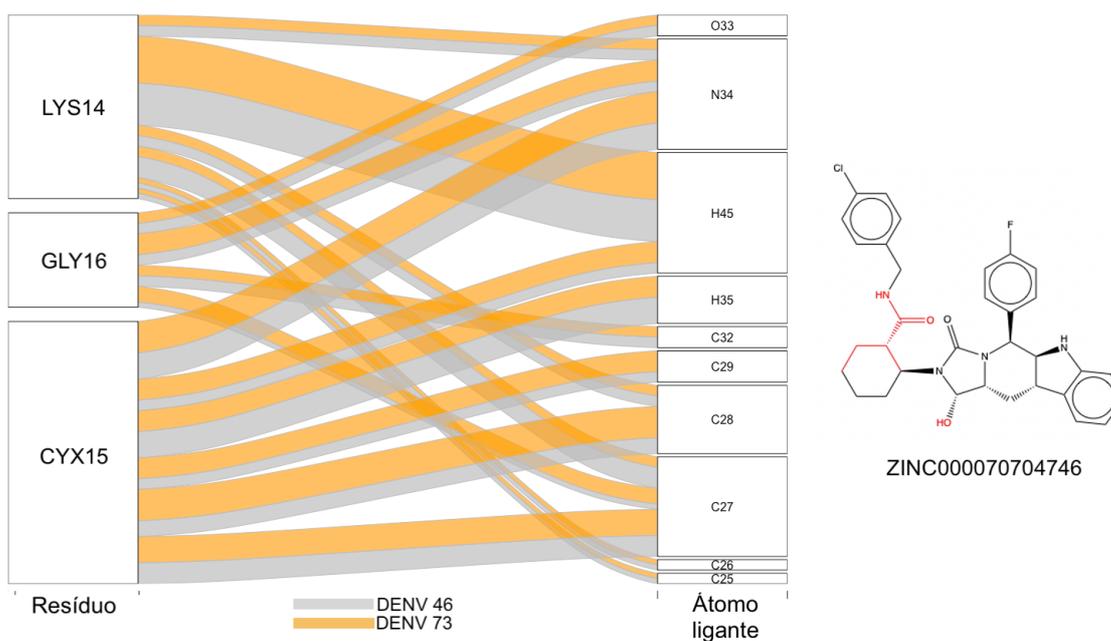


Diagrama ilustrando o local de ancoragem do fragmento ZINC000070704746 em comum nas duas estruturas da NS1. A espessura da linha corresponde ao número de átomos dos resíduos nos quais o átomo do ligante está próximo. Na representação da estrutura do composto, destaca-se (em vermelho) os fragmentos em comum (coluna “Átomo ligante” do diagrama).

Já para o ligante ZINC000070699829 além dos três resíduos destacados anteriormente, mais dois (GLU203 e TRP201) foram observados com interações em comum com átomos do ligante (Figura 34). Percebe-se também que o fragmento destacado nesse ligante difere consideravelmente do fragmento do ZINC000070704746.

Figura 34 – Análise fragmento-específico do ligante ZINC000070699829 no sítio da NS1 do DENV

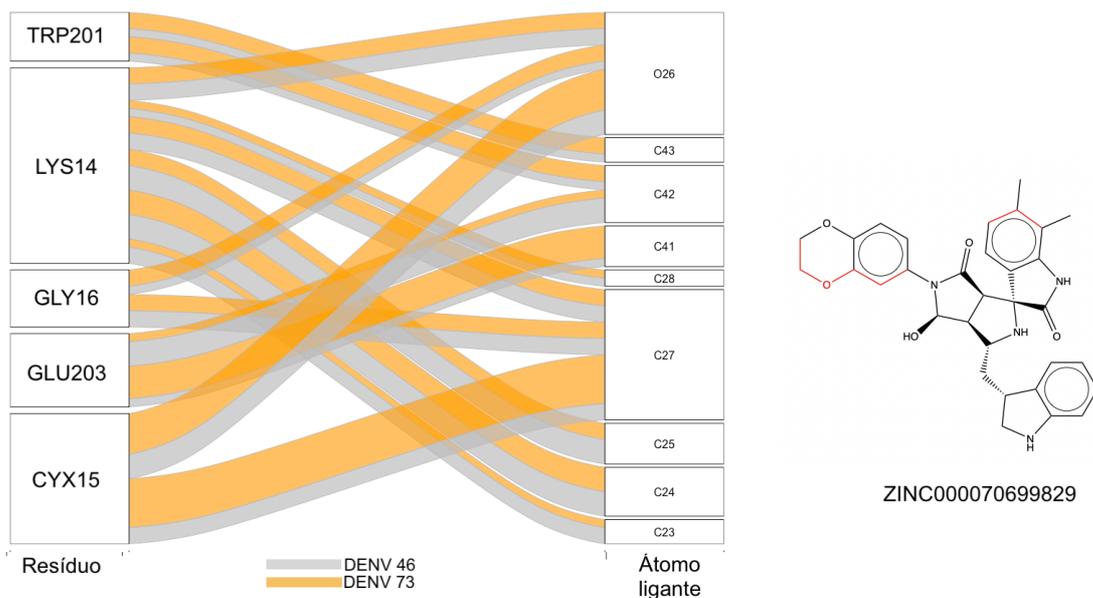
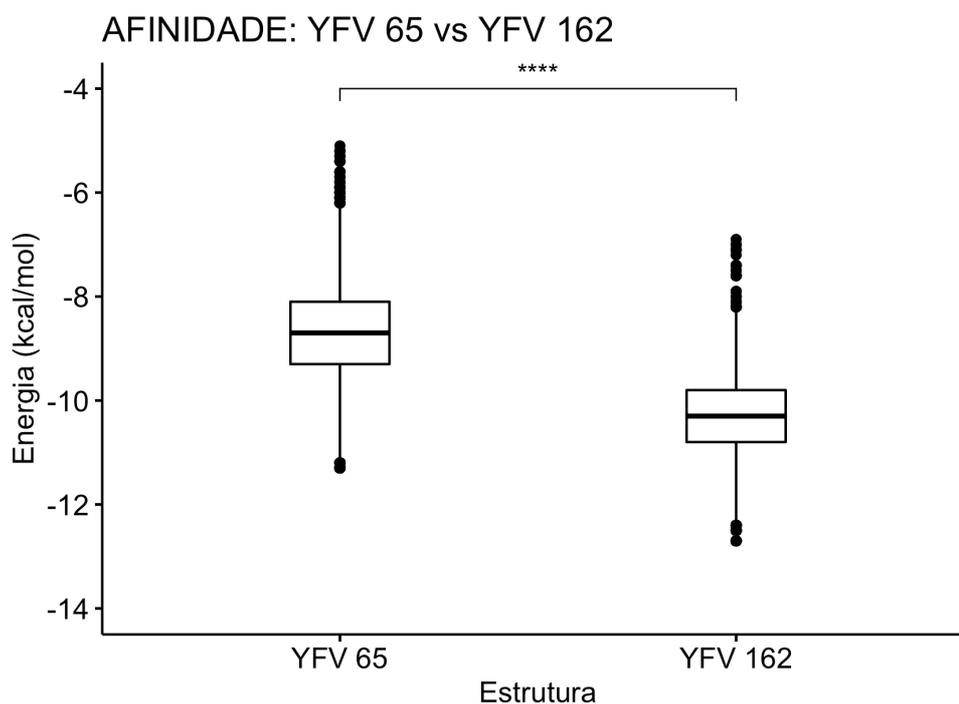


Diagrama ilustrando o local de ancoragem do fragmento ZINC000070699829 em comum nas duas estruturas da NS1. Na representação da estrutura do composto, destaca-se (em vermelho) os fragmentos em comum (coluna “Átomo ligante” do diagrama).

5.4.1.2 Comparação da triagem virtual entre estruturas da NS1 do YFV

Para as estruturas da NS1 do YFV, resultado similar foi observado em relação às distribuições das energias. A estrutura YFV 65 obteve como menor energia -11.3 kcal/mol e maior de -5.1 kcal/mol. Já a estrutura YFV 162 as energias foram -12.7 kcal/mol e -6.9 kcal/mol, com diferença estatisticamente significante ($p < 0.0001$) entre as estruturas (Figura 35). Entre os 100 compostos de maior afinidade para cada estrutura, cinco estavam presentes no ranqueamento de ambas (Apêndice C) apesar de um menor TM-score (0.66) entre os bolsões que o observado para as estruturas do DENV. O composto ZINC000070699829 também esteve entre esses compostos em comum, assim como para as estruturas do DENV, indicando que esse composto aparece no ranqueamento das estruturas de ambas espécies.

Figura 35 – *Boxplot* das distribuições de energia das 100 simulações exaustivas dos 100 compostos ranqueados de acordo com o critério de energia (afinidade) para as duas estruturas da NS1 do YFV.



**** $p < 0.0001$

No diagrama 2D das interações (Figura 36) nota-se que há resíduos (GLY16, ILE19, PHE20, VAL21, ARG23 e TRP201) que estão interagindo com os cinco compostos para ambas as estruturas da NS1 do YFV e DENV. O RMSD entre estruturas levando em conta apenas esses resíduos para o cálculo foi de 4.753 Å, menor que o RMSD entre os resíduos em comum da análise do DENV (6.809 Å), o que pode justificar um maior número de compostos em comum apesar do menor TM-score. Percebe-se também pequenas diferenças em relação aos resíduos envolvidos nas interações inerentes à mudança de conformação entre bolsões do YFV 65 e YFV 162. Nota-se que as cinco maiores energias (-10.2 kcal/mol, -10.3 kcal/mol, -10.5 kcal/mol, -10.9 kcal/mol e -11.1 kcal/mol) pertencem a interação com YFV 65 e os cinco menores (-11.6 kcal/mol, -11.7 kcal/mol, -11.8 kcal/mol, -11.9 kcal/mol e -12.1 kcal/mol) ao YFV 162. A maior divergência entre sítios de ancoragem está na ausência de interação com resíduos LYS10 e GLU12 na estrutura do YFV 162, e presença de interações em regiões envolvendo os resíduos ALA187, GLU203, SER204, HIP214, LYS276, além de possuir maior número de interações (principalmente Alkyl ou Pi-Alkyl e de van der Waals) ocasionando uma menor energia de ligação.

Na análise de especificidade resíduo-átomo para o ligante ZINC000008792187, apenas uma pequena parte do fragmento foi observado em comum para os aminoácidos PHE20, ASP26 e TRP201 (Figura 37). O fragmento compreende três carbonos de um anel aromático e um átomo de nitrogênio (destacados em vermelho na representação da estrutura).

Figura 37– Análise fragmento-específico do ligante ZINC000008792187 no sítio da NS1 do YFV

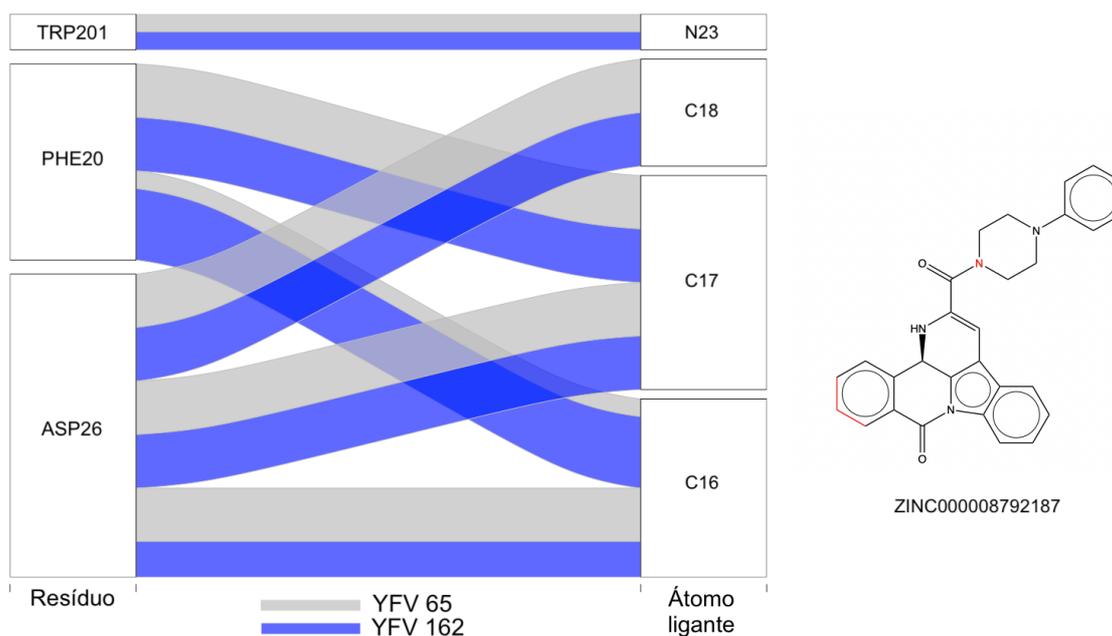


Diagrama ilustrando o local de ancoragem do fragmento ZINC000008792187 em comum nas duas estruturas da NS1. Na representação da estrutura do composto, destaca-se (em vermelho) os fragmentos em comum (coluna “Átomo ligante” do diagrama).

A análise para o ligante ZINC000070705115 também apresentou apenas um pequeno fragmento com relação ao resíduo-átomo ligante específico (Figura 38). O fragmento em destaque também envolveu parte de um anel aromático (C29 a C32), similar ao fragmento do ligante descrito anteriormente. Apesar disso, os resíduos envolvidos foram diferentes.

Figura 38– Análise fragmento-específico do ligante ZINC000070705115 no sítio da NS1 do YFV

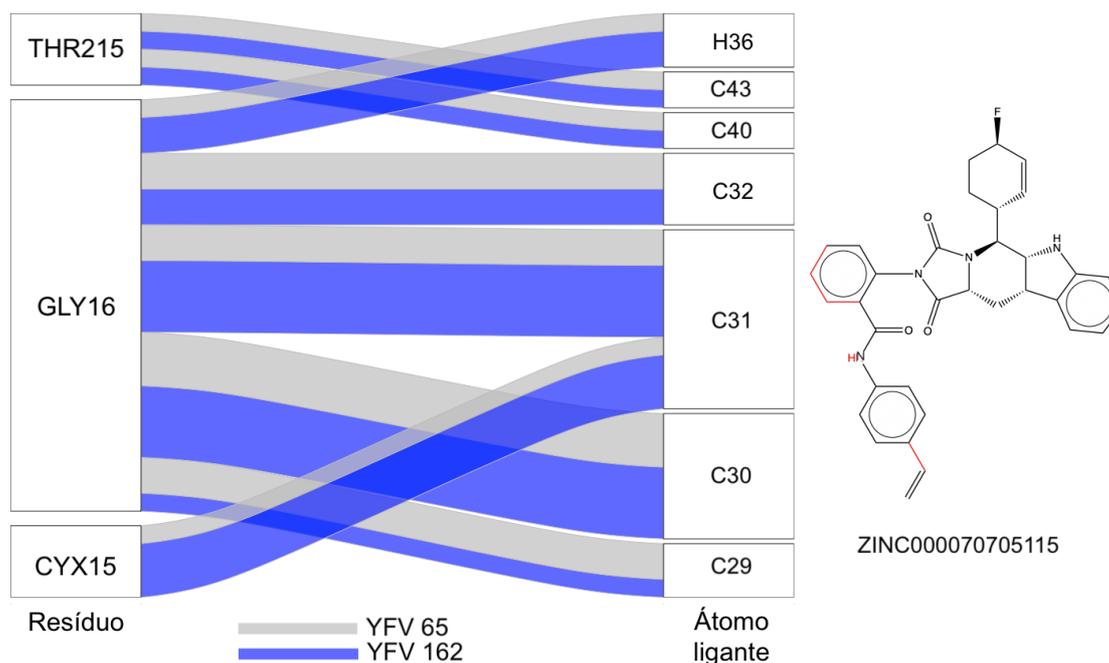


Diagrama ilustrando o local de ancoragem do fragmento ZINC000008792187 em comum nas duas estruturas da NS1. Na representação da estrutura do composto, destaca-se (em vermelho) os fragmentos em comum (coluna “Átomo ligante” do diagrama).

Em relação ao composto ZINC000095486217 apenas quatro átomos não-consecutivos foram encontrados na análise em comum entre as estruturas (Figura 39), não sendo observado um fragmento relevante em comum.

Figura 39– Análise fragmento-específico do ligante ZINC000095486217 no sítio da NS1 do YFV

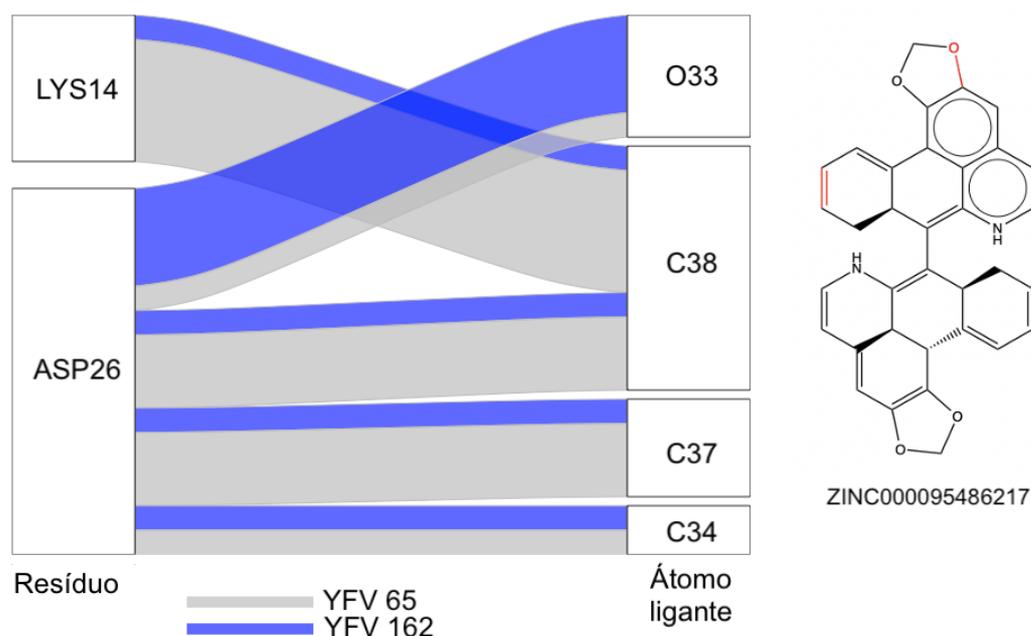


Diagrama ilustrando o local de ancoragem do fragmento ZINC000095486217 em comum nas duas estruturas da NS1. Na representação da estrutura do composto, destaca-se (em vermelho) os fragmentos em comum (coluna “Átomo ligante” do diagrama).

Para o ligante ZINC000015956900 houve maior concordância na proximidade em relação aos átomos do ligante e resíduo (Figura 40). Duas regiões (C21-C29 e C34-C39/N37/H38) foram apontadas em comum para as duas estruturas. Eles compreendem partes de anéis aromáticos, sendo um com a presença de nitrogênio. Além desses fragmentos, átomos de hidrogênios e oxigênio isolados apareceram também na análise.

Figura 40– Análise fragmento-específico do ligante ZINC000015956900 no sítio da NS1 do YFV

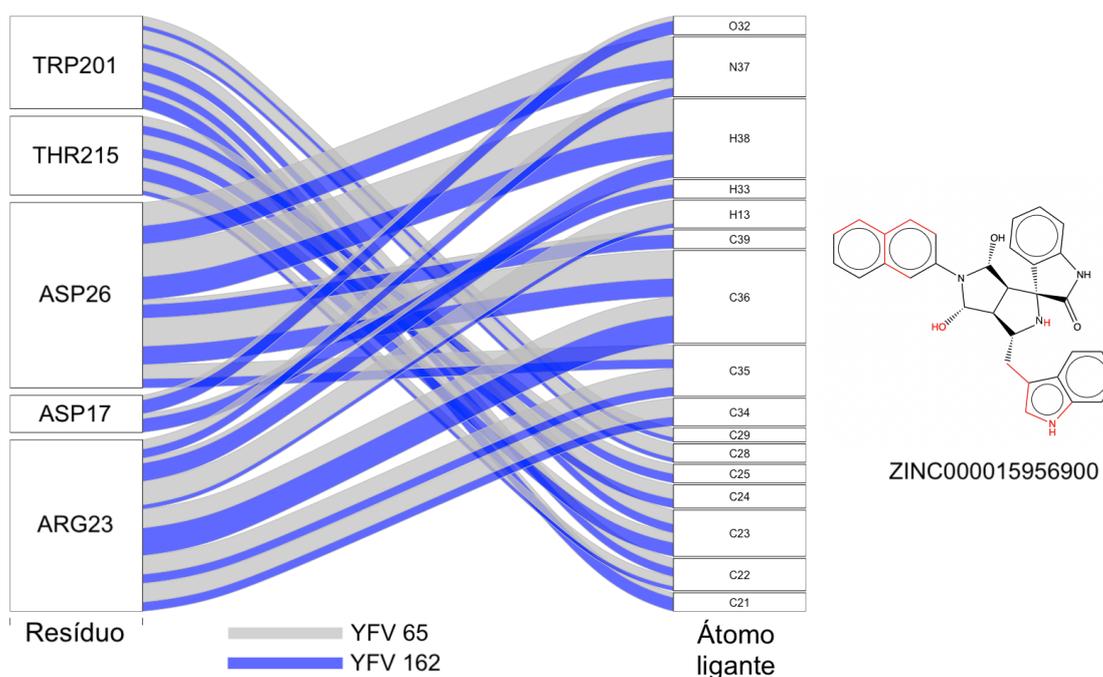


Diagrama ilustrando o local de ancoragem do fragmento ZINC000015956900 em comum nas duas estruturas da NS1. Na representação da estrutura do composto, destaca-se (em vermelho) os fragmentos em comum (coluna “Átomo ligante” do diagrama).

O composto ZINC000070699829 apresentou duas pequenas regiões (C2-C5 e C9/C35/N14/H15) em comum à ambas estruturas (Figura 41). Juntando as informações dessas análises observa-se que alguns resíduos (CYX15, GLY16, ASP26 e TRP201) apareceram em mais de uma análise de comparação resíduo-átomo ligante. Os resíduos GLY16 e TRP201 são dois resíduos que apareceram em todas as interações entre NS1 independente do ligante analisado (diagrama de interação da Figura 36). Nessa última análise de fragmento sítio-específico, para dois compostos, esses resíduos se mostraram também com uma especificidade em relação ao modo de ligação no sítio.

Figura 41 - Análise fragmento-específico do ligante ZINC000070699829 no sítio da NS1 do YFV

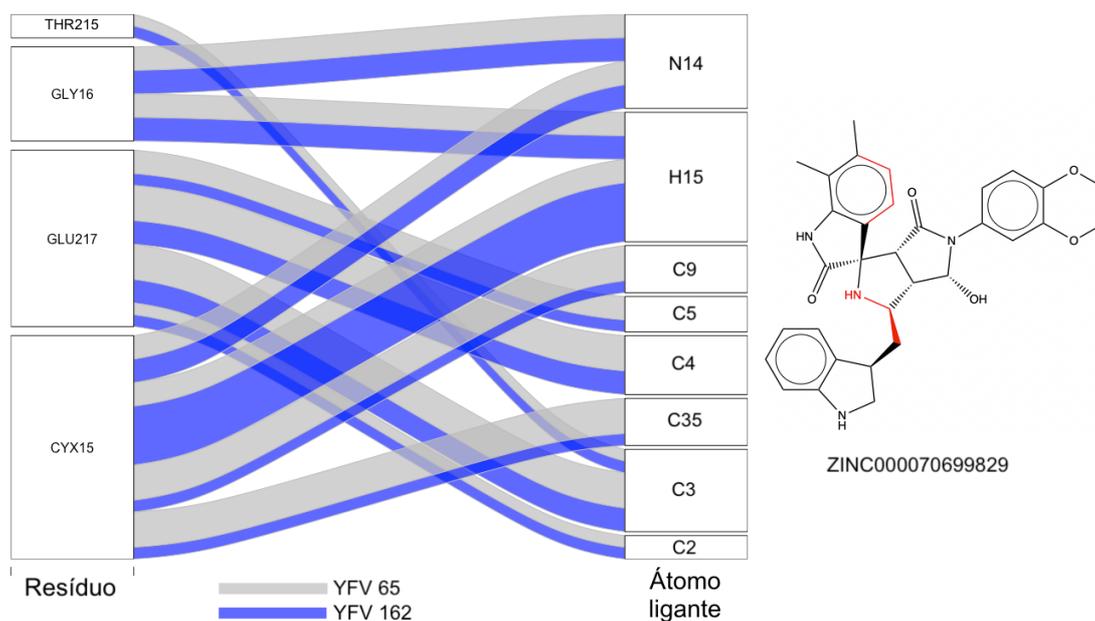
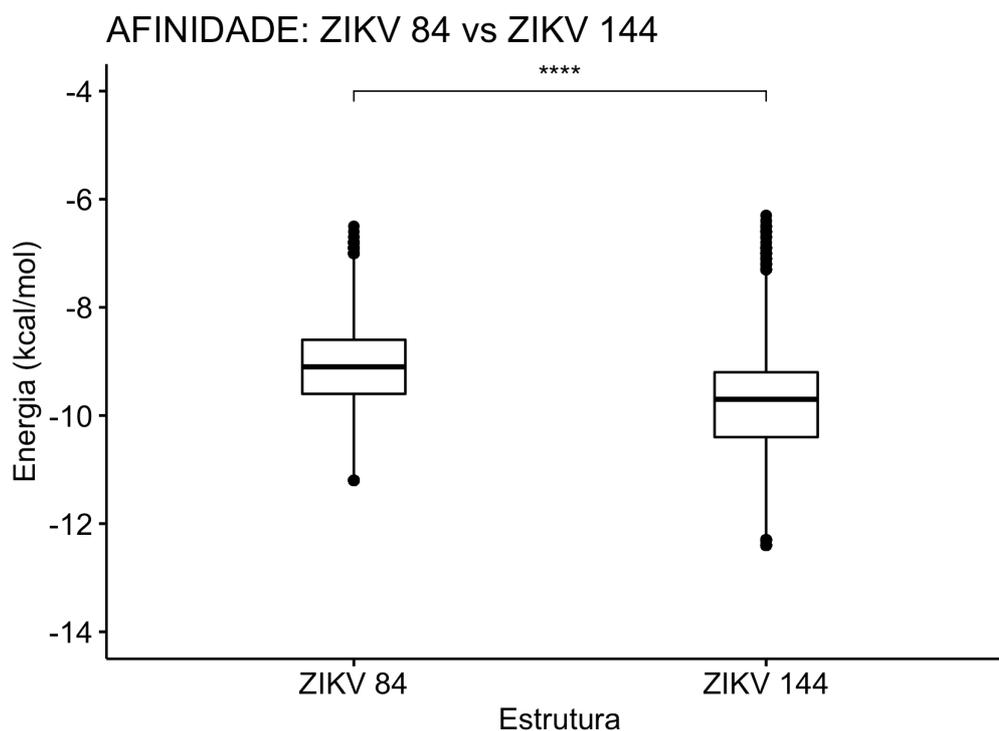


Diagrama ilustrando o local de ancoragem do fragmento ZINC000070699829 em comum nas duas estruturas da NS1. Na representação da estrutura do composto, destaca-se (em vermelho) os fragmentos em comum (coluna “Átomo ligante” do diagrama).

5.4.1.3 Comparação da triagem virtual entre estruturas da NS1 do ZIKV

Para as estruturas da NS1 do ZIKV também houve diferença estatisticamente significativa ($p < 0.0001$) nas distribuições de energias dos compostos ranqueados entre os 100 melhores para cada estrutura. A estrutura do ZIKV 84 obteve como menor energia -11.2 kcal/mol e maior de -6.5 kcal/mol. Já a estrutura do ZIKV 144 as energias foram -12.4 kcal/mol e -6.3 kcal/mol (Figura 42). Entre os 100 compostos de maior afinidade para cada estrutura, seis estavam presentes no ranqueamento de ambas (Apêndice D). A maior quantidade de compostos em comum reflete a maior similaridade entre os bolsões (TM-score = 0.78768). Nenhum desses compostos apareceram nas duas análises realizadas anteriormente para o DENV e YFV.

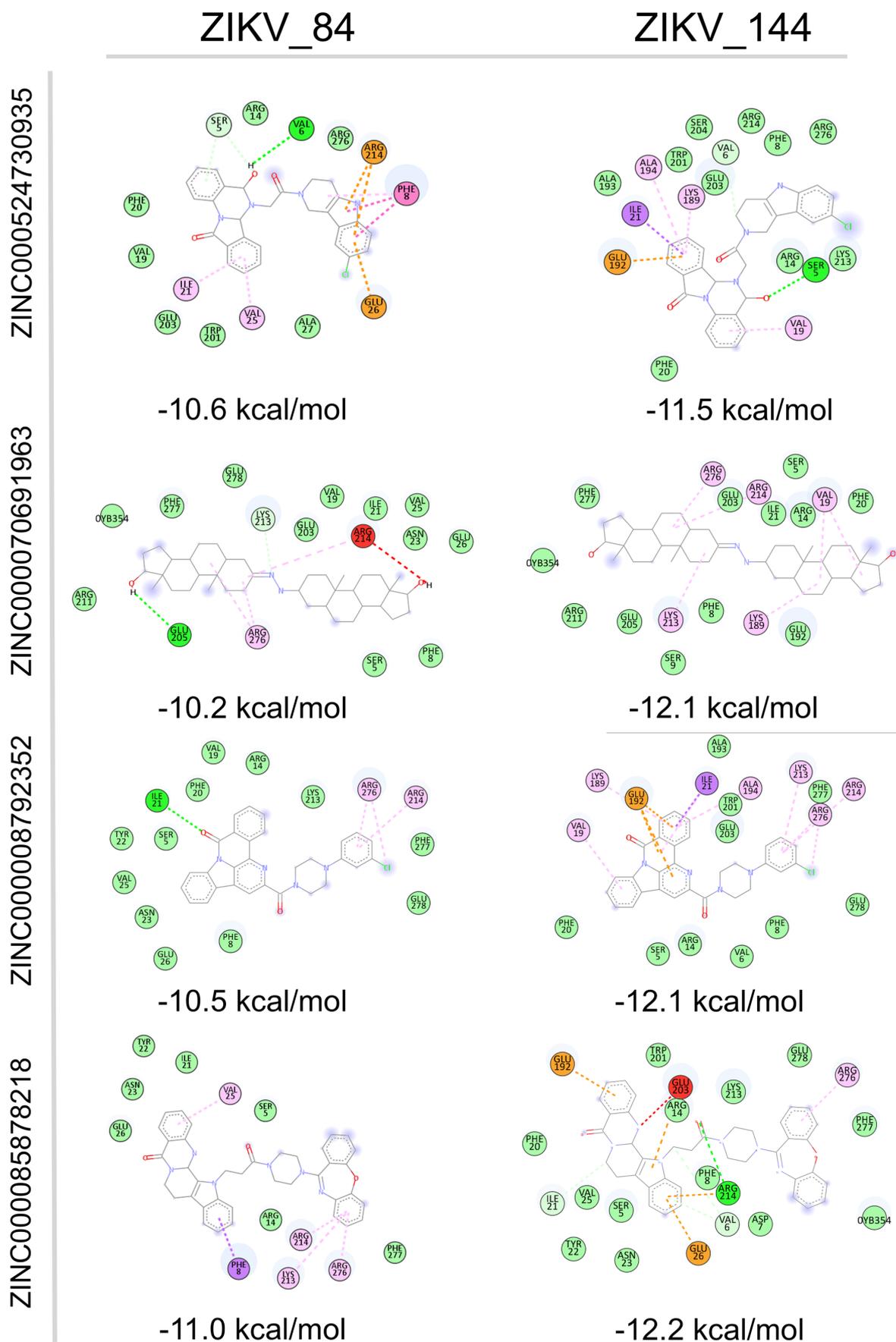
Figura 42 – *Boxplot* das distribuições de energia das 100 simulações exaustivas dos 100 compostos ranqueados de acordo com o critério de energia (afinidade) para as duas estruturas da NS1 do ZIKV.



**** $p < 0.0001$

No diagrama 2D (Figura 43) das interações entre receptor e ligante é possível notar alguns resíduos chaves nas interações com esses seis compostos tais como SER5, PHE8, ILE21, ARG214 e ARG276. Não foi observado nenhum padrão de interação relacionado a estrutura ZIKV 84 ou ZIKV 144, ou seja, não houve interação com resíduos exclusivos de uma estrutura, provavelmente pela elevada similaridade entre as cavidades. Além disso, o RMSD entre os resíduos que apareceram interagindo em todos os compostos foi de 3.108 Å, o menor valor entre as espécies. As menores energias foram todas da estrutura do ZIKV 144 destacando-se pelo maior número de interações de van der Waals e Alkyl/pi-Alkyl.

Figura 43 - Representação do diagrama 2D de interação dos seis compostos em comum às duas estruturas do ZIKV com seus respectivos escores de energia.



A análise entre especificidade do fragmento para o ligante ZINC000524730935 e resíduo da proteína, apresentou duas grandes regiões com similar posição em relação aos resíduos no sítio de ancoragem na entre as duas estruturas da NS1 do ZIKV, indicando que mesmo com uma alteração estrutural no sítio, houve a manutenção da pose em relação aos resíduos (Figura 44).

Figura 44 - Análise fragmento-específico do ligante ZINC000524730935 no sítio da NS1 do ZIKV

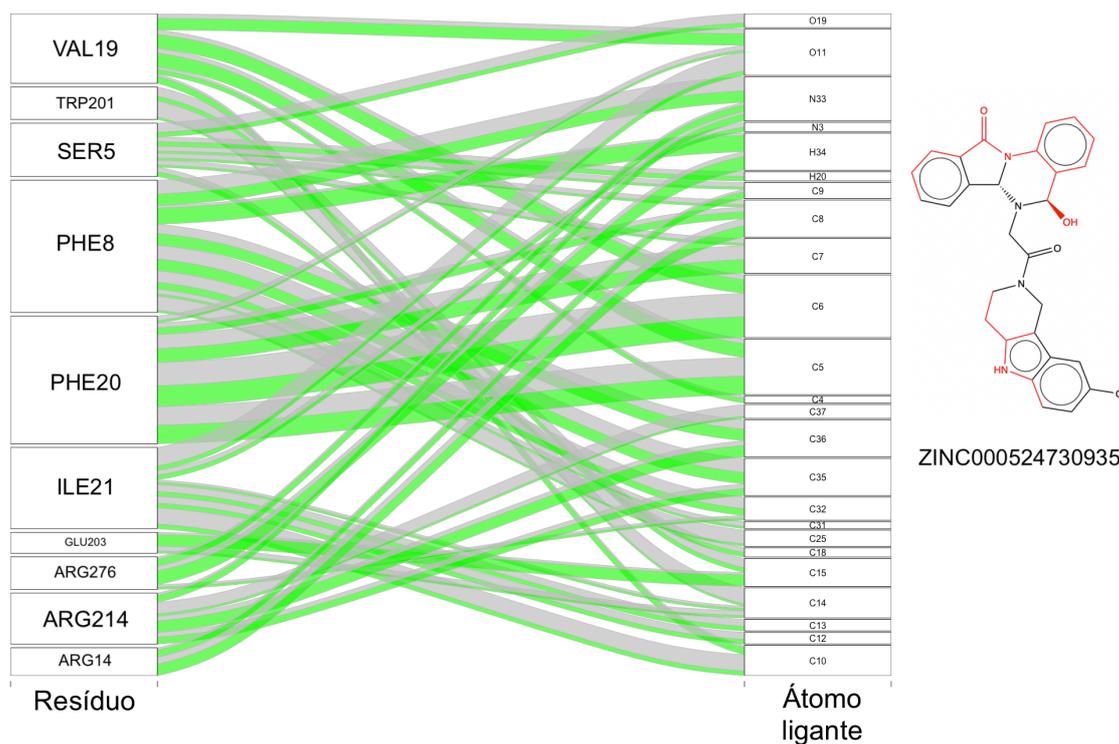


Diagrama ilustrando o local de ancoragem do fragmento ZINC000524730935 em comum nas duas estruturas da NS1. Na representação da estrutura do composto, destaca-se (em vermelho) os fragmentos em comum (coluna “Átomo ligante” do diagrama).

Em relação à análise para o ligante ZINC000070691963, uma grande região da molécula também foi observada em comum às duas estruturas da NS1 do ZIKV (Figura 45) também indicando uma especificidade entre resíduo-fragmento da molécula.

Figura 45 - Análise fragmento-específico do ligante ZINC000070691963 no sítio da NS1 do ZIKV

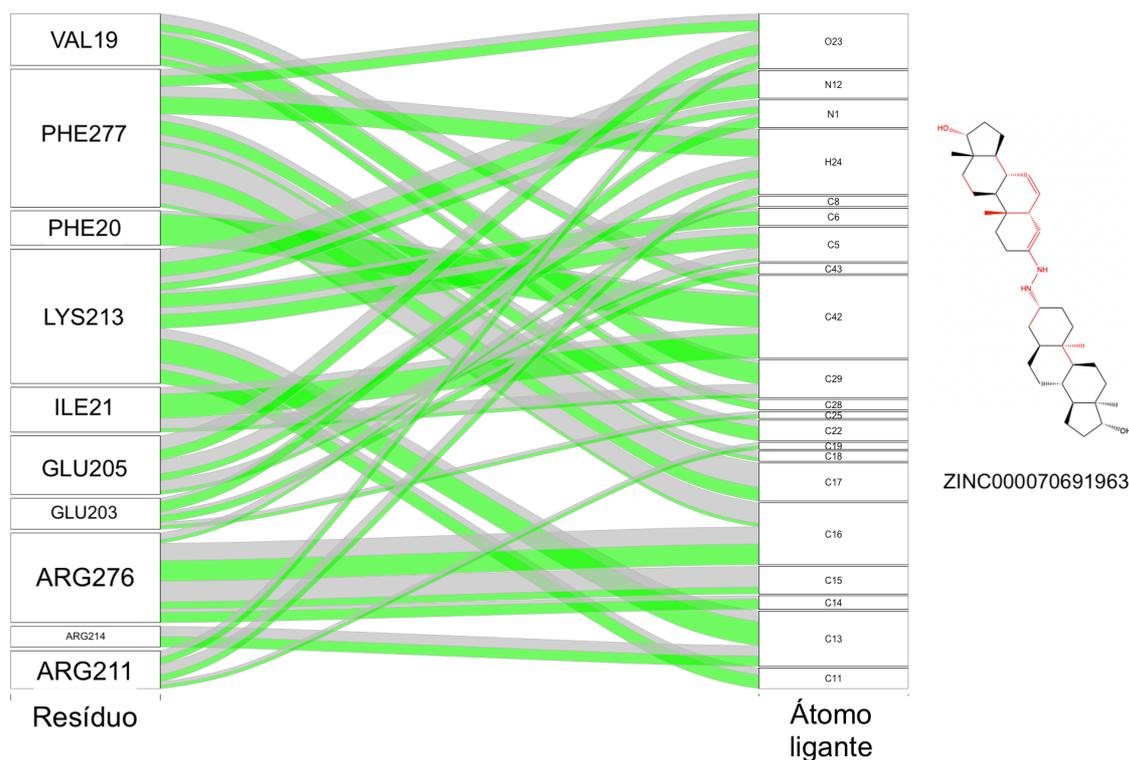


Diagrama ilustrando o local de ancoragem do fragmento ZINC000070691963 em comum nas duas estruturas da NS1. Na representação da estrutura do composto, destaca-se (em vermelho) os fragmentos em comum (coluna “Átomo ligante” do diagrama).

Para o ligante ZINC000008792352 um fragmento contendo um anel aromático ligado a um átomo de cloro e nitrogênio foi observado na análise resíduo-átomo ligante, além de uma pequena região de carbonila ligado ao nitrogênio, e um átomo de oxigênio de outra carbonila (Figura 46).

Duas pequenas regiões do ligante ZINC000085878218 foram apontados com especificidade em relação à distância entre os resíduos da proteína NS1 (Figura 47). Aqui, pode-se concluir que houve pouca especificidade em relação às outras moléculas observadas para a NS1 do ZIKV. Apesar disso, é importante frisar a ocorrência de resíduos que apareceram em análises anteriores como SER5, PHE8, GLU203 e LYS213.

Figura 46 - Análise fragmento-específico do ligante ZINC000008792352 no sítio da NS1 do ZIKV

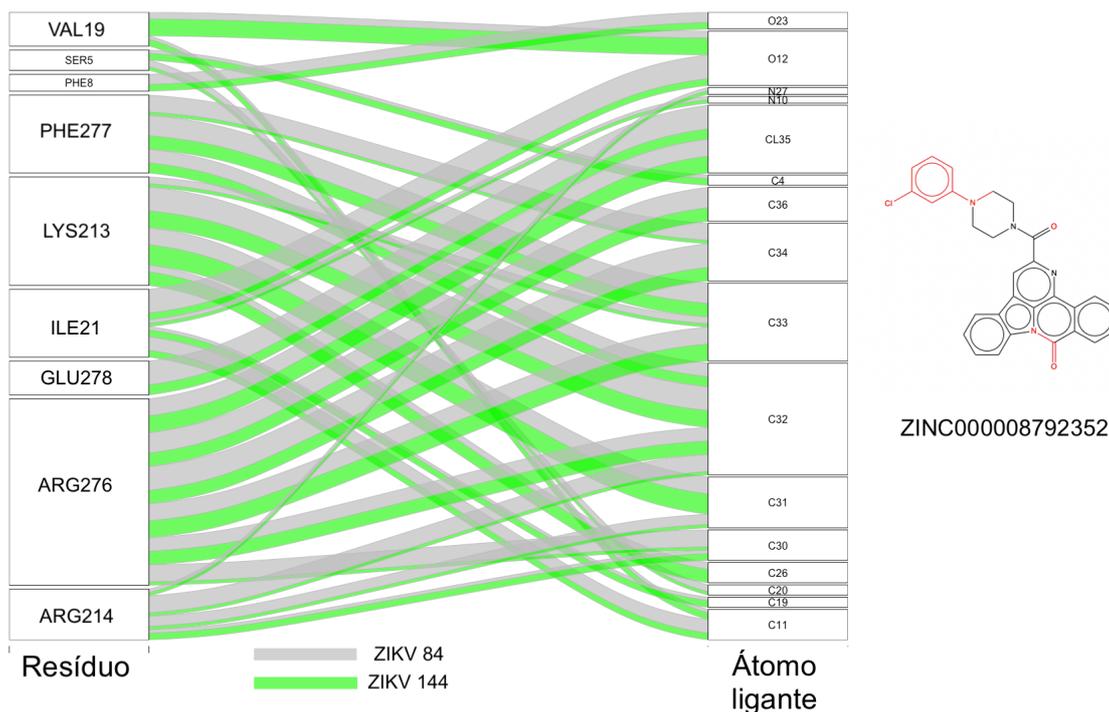


Diagrama ilustrando o local de ancoragem do fragmento ZINC000008792352 em comum nas duas estruturas da NS1. Na representação da estrutura do composto, destaca-se (em vermelho) os fragmentos em comum (coluna “Átomo ligante” do diagrama).

Figura 47 - Análise fragmento-específica do ligante ZINC000085878218 no sítio da NS1 do ZIKV

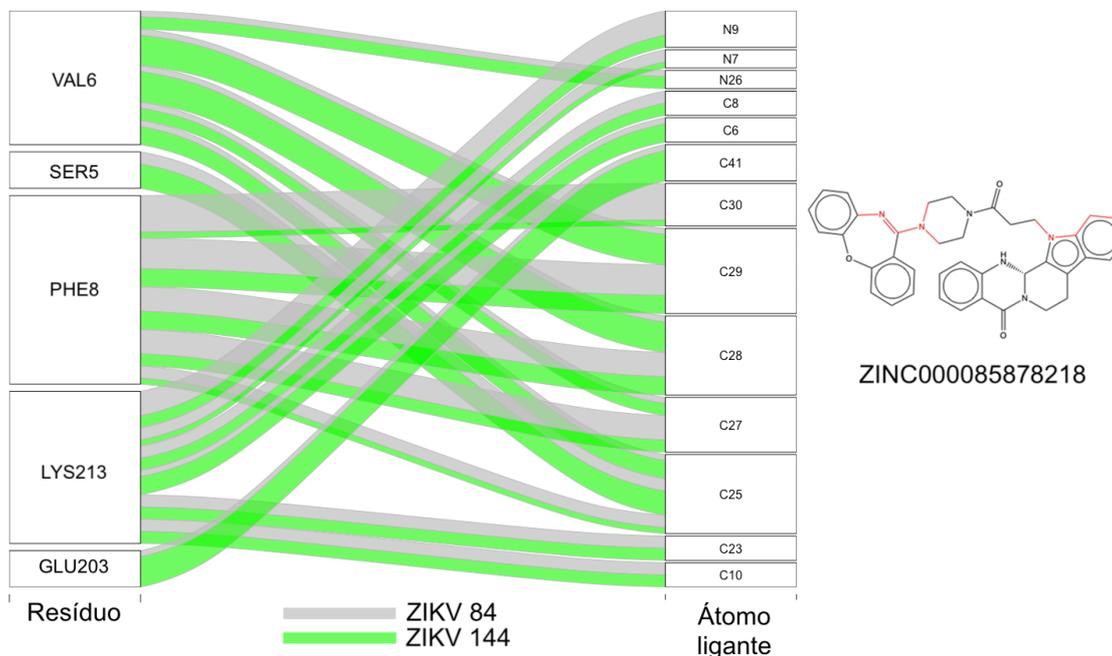


Diagrama ilustrando o local de ancoragem do fragmento ZINC000085878218 em comum nas duas estruturas da NS1. Na representação da estrutura do composto, destaca-se (em vermelho) os fragmentos em comum (coluna “Átomo ligante” do diagrama).

Na análise do ligante ZINC000085569026 apenas houve concordância com um resíduo (LYS213) e alguns átomos da molécula compreendendo quatro regiões com três ou quatro átomos cada (Figura 48). Aqui pode-se também considerar que a mudança na estrutura do sítio de ancoragem afetou a posição da molécula em relação aos resíduos da proteína.

Figura 48 - Análise fragmento-específico do ligante ZINC000085569026 no sítio da NS1 do ZIKV

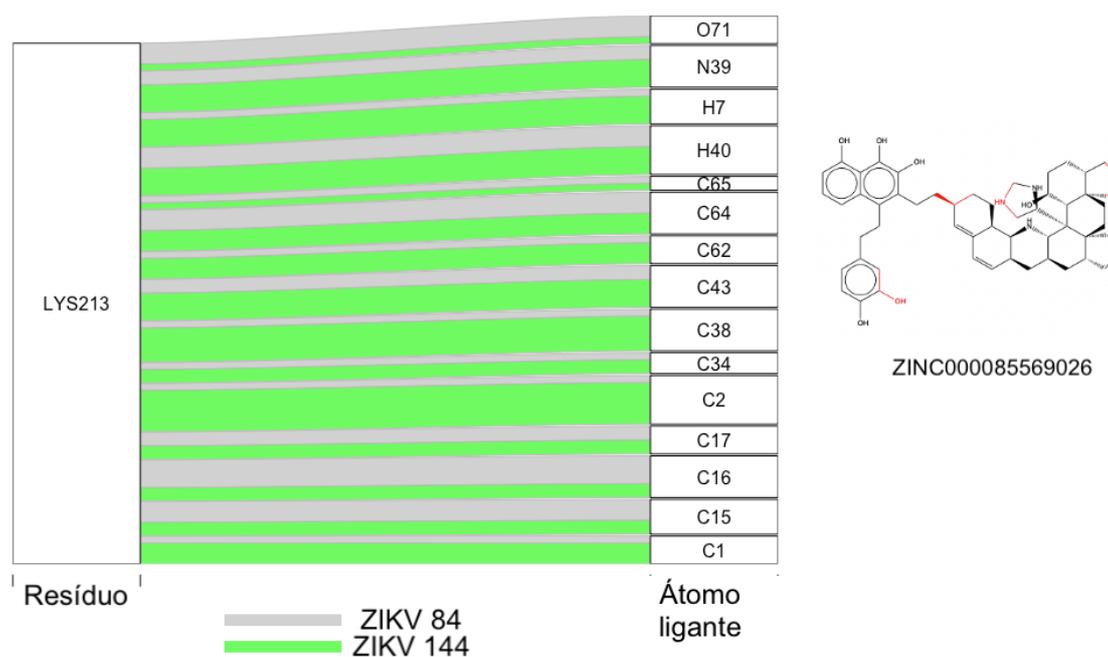


Diagrama ilustrando o local de ancoragem do fragmento ZINC000085569026 em comum nas duas estruturas da NS1. Na representação da estrutura do composto, destaca-se (em vermelho) os fragmentos em comum (coluna “Átomo ligante” do diagrama).

Por fim, para o ligante ZINC000020759145 não foi observado nenhuma interação em comum relacionando resíduo-átomo do ligante, sendo este o ligante mais afetado pela diferença no sítio de ancoragem. Vale salientar que a análise foi feita apenas do modo de ligação de menor energia para cada estrutura do ZIKV, sendo que para as outros modos esse resultado pode ser diferente.

Nessas análises do ZIKV foi possível ver alguns resíduos em destaque, como SER5, PHE8, ARG14, VAL19, PHE20, ILE21, GLU203, LYS213, ARG214, ARG216 e PHE277. Alguns desses resíduos já haviam sido destacados na análise de interação, e além de uma preferência por interação com esses resíduos dos sítios, observou-se também uma preferência relacionada ao fragmento do ligante em análise.

5.4.1.4 Análise da intersecção de compostos presentes na triagem virtual das três espécies

Realizou-se uma análise comparativa dos compostos classificados entre os 100 melhores para cada estrutura usando apenas o critério de energia. Buscou-se compostos que estavam presentes simultaneamente na classificação das três diferentes espécies. Para isso analisou-se inicialmente a interseção de compostos entre as seis estruturas (duas de cada espécie) e assim foi-se reduzindo entre cinco estruturas, quatro e três, sendo que nas análises deveria ter pelo menos uma estrutura de cada espécie. O resultado pode ser visto na Tabela 2 abaixo:

Tabela 2 – Interseção de compostos por critério de afinidade presentes nas estruturas da NS1

ANÁLISE COM 6 ESTRUTURAS						# COMPOSTOS EM COMUM
DENV_46	DENV_73	YFV_65	YFV_162	ZIKV_84	ZIKV_144	0
ANÁLISE COM 5 ESTRUTURAS						# COMPOSTOS EM COMUM
DENV_46	DENV_73	YFV_65	YFV_162	ZIKV_84	0	
DENV_46	DENV_73	YFV_65	YFV_162	ZIKV_144	0	
DENV_46	DENV_73	YFV_65	ZIKV_84	ZIKV_144	0	
DENV_46	DENV_73	YFV_162	ZIKV_84	ZIKV_144	0	
DENV_46	YFV_65	YFV_162	ZIKV_84	ZIKV_144	0	
DENV_73	YFV_65	YFV_162	ZIKV_84	ZIKV_144	0	
ANÁLISE COM 4 ESTRUTURAS						# COMPOSTOS EM COMUM
DENV_46	DENV_73	YFV_65	ZIKV_84	0		
DENV_46	DENV_73	YFV_65	ZIKV_144	0		
DENV_46	DENV_73	YFV_162	ZIKV_84	0		
DENV_46	DENV_73	YFV_162	ZIKV_144	0		
DENV_46	YFV_65	YFV_162	ZIKV_84	0		
DENV_73	YFV_65	YFV_162	ZIKV_84	0		
DENV_73	YFV_65	YFV_162	ZIKV_144	0		
DENV_46	YFV_65	YFV_162	ZIKV_144	0		
DENV_46	YFV_65	ZIKV_84	ZIKV_144	0		
DENV_73	YFV_65	ZIKV_84	ZIKV_144	0		
DENV_46	YFV_162	ZIKV_84	ZIKV_144	0		
DENV_73	YFV_162	ZIKV_84	ZIKV_144	1		
ANÁLISE COM 3 ESTRUTURAS						# COMPOSTOS EM COMUM
DENV_46	YFV_65	ZIKV_84	0			
DENV_46	YFV_162	ZIKV_84	0			
DENV_46	YFV_65	ZIKV_144	0			
DENV_46	YFV_162	ZIKV_144	0			
DENV_73	YFV_65	ZIKV_84	1			
DENV_73	YFV_162	ZIKV_84	1			
DENV_73	YFV_65	ZIKV_144	1			
DENV_73	YFV_162	ZIKV_144	3			

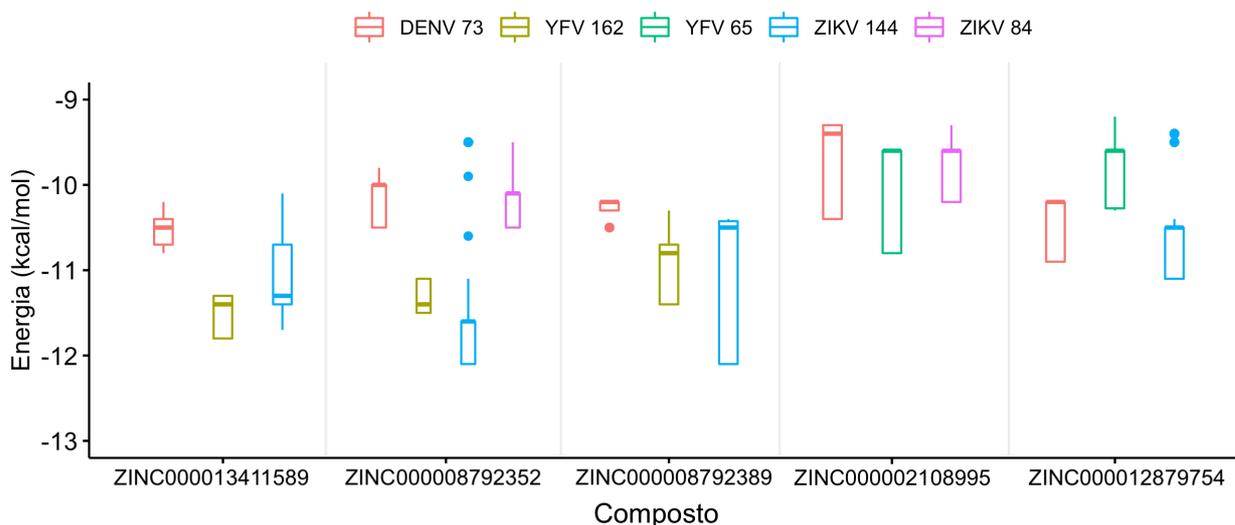
Como é possível observar, apenas nas análises com quatro e três estruturas foi possível achar compostos em comum às três espécies virais. Compilando as informações dos compostos que apareceram nessas análises, obteve-se cinco compostos promissores que possuem afinidade elevada à NS1 das três espécies virais, sendo que o composto ZINC000008792352 apareceu em três das cinco análises (Tabela 3).

Tabela 3 – Compostos por critério de afinidade presentes na interseção entre estruturas da NS1

Composto	Análises
ZINC000008792352	DENV 73∩YFV 162∩ZIKV 84∩ZIKV 144 DENV 73∩YFV 162∩ZIKV 144 DENV 73∩YFV 162∩ZIKV 84
ZINC000012879754	DENV 73∩YFV 65∩ZIKV 144
ZINC000002108995	DENV 73∩YFV 65∩ZIKV 84
ZINC000013411589	DENV 73∩YFV 162∩ZIKV 144
ZINC000008792389	DENV 73∩YFV 162∩ZIKV 144

A Figura 49 apresenta a distribuição das energias dos 250 melhores modos de ligação (menores energias) desses cinco compostos para cada estrutura. Nota-se que os compostos ZINC000013411589, ZINC000008792352 e ZINC000008792389 apresentam médias menores dos valores de energia quando comparado com ZINC000002108995 e ZINC000012879754.

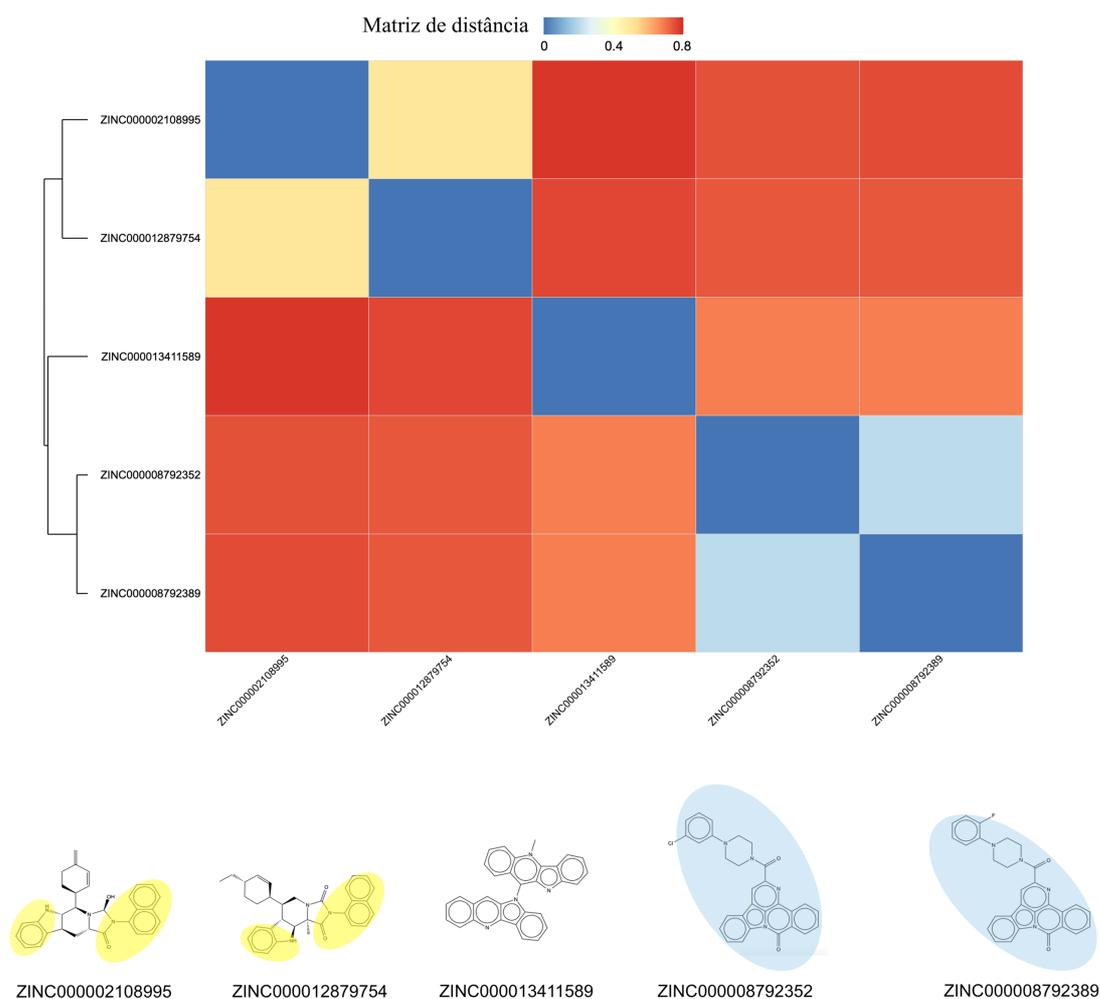
Figura 49 – *Boxplot* da distribuição das energias da triagem virtual obtida para os cinco compostos em comum
Cinco compostos em comum às três espécies - Critério de afinidade



A análise de similaridade usando matriz de distância (onde subtrai-se de 1 o valor do coeficiente de Tanimoto entre compostos) foi realizado para comparar as estruturas dos cinco compostos. Dessa forma, valores próximos de 0 indicam alta similaridade e

valores próximos de 1 indicam baixa similaridade. O mapa de calor (Figura 50) mostra que dois compostos (ZINC000008792352 e ZINC000008792389) apresentaram elevada similaridade (valor da matriz de distância = 0.21; coeficiente de Tanimoto = 0.79). Entre esses dois compostos não foi observado diferença estatisticamente significativa entre as distribuições de energia da Figura 49 (dado não mostrado). Os outros dois compostos com menor distância foram os compostos ZINC000002108995 e ZINC000012879754 (valor matriz de distância = 0.51; coeficiente de Tanimoto = 0.49). Entre eles houve diferença estatisticamente significativa ($p < 0.0001$) em relação a distribuição de energia para as estruturas da NS1 (dado não mostrado). Na Figura 50 os fragmentos em comum desses dois pares de compostos estão destacados em amarelo e azul nas estruturas dos compostos. Esses dados sugerem uma maior afinidade pelo fragmento destacado em azul em relação aos fragmentos destacados em amarelo.

Figura 50 – Similaridade estrutural entre os cinco compostos



Na análise das principais interações entre ligante e proteína é possível ver algumas interações (posição de resíduo) em comum às estruturas a depender do composto (Figura 51). Para o ZINC000008792352 observou-se a conservação da interação nos resíduos LYS/ARG14, ILE/VAL19 e ILE/VAL21. Apesar de resíduos diferentes na mesma posição, eles possuem característica bioquímica em comum: polar de carga positiva para LYS/ARG e apolar alifático no caso do ILE/VAL.

No caso do composto ZINC000012879754 interações com os resíduos LYS/ARG14 e ILE/VAL21 foram conservadas nas três estruturas da NS1. A molécula ZINC000002108995 apresentou, além de interações já citadas (LYS/ARG14 e ILE/VAL19), as interações ARG/ASP/ASN23 e HID/ASP/GLU26 tendo em comum serem aminoácidos de caráter polar, porém com diferentes cargas: ASN – polar não carregado, HID/ARG – polar carga positiva e GLU e ASP – polar carga negativa.

O composto ZINC000013411589 apresentou seis interações em resíduo em comum: resíduos LYS/ARG 14, ILE/VAL19, ILE/VAL21, HID/ASP/GLU26, VAL/ALA194, TRP201 e GLU/GLY203. ASP e GLU possuem caráter polar negativo, diferente do HID que possui carga positiva. VAL/ALA possuem ambos caracteres apolares alifáticos. GLU e GLY são opostos em relação à polaridade. Enquanto a GLY é apolar alifático, a GLU é polar de carga negativa.

Por fim, a molécula ZINC000008792389 foi a que apresentou mais interações em resíduos em comum às estruturas da NS1. Algumas já citadas como LYS/ARG14, ILE/VAL19, ILE/VAL21 e VAL/ALA194. Houveram também interações na PHE20 e TRP201, ambos resíduos conservados nas três espécies. E nos resíduos ASP/LYS/ARG276, também polares com carga, sendo os dois últimos de carga positiva e o primeiro de carga negativa.

A partir dessas observações é possível destacar resíduos chaves nas interações desses cinco compostos como LYS/ARG14 que apareceu em todas as interações, além de outras importantes como ILE/VAL19 e ILE/VAL21. Apesar de serem resíduos distintos, nota-se que eles possuem características bioquímicas conservadas. O resíduo TRP201 apenas não apareceu em três das 16 interações, podendo ser considerado como um importante resíduo principalmente pelo fato de ser conservado nas três espécies, assim como a PHE20 que apareceu em 11 interações.

Figura 51 – Diagrama das principais interações* observadas entre proteína e ligante

Composto	Estrutura	Torções ativas	# Átomos	Energia (kcal/mol)	SER5	VAL5	VAL6	TRP8	PHE8	LYS14	ARG14	CYS15	GLY16	ASP17	SER17	ILE19	VAL19	PHE20	VAL21	ILE21	ARG23	ASP23	ASN23	HID26	ASP26	GLU26	VAL194	ALA194	TRP201	GLY203	GLU203	LYS213	LYS276	ARG276	Total interações
ZINC000008792352	DENV_73	2	36	-10.5																															17
	YFV_162	2	36	-11.5																															17
	ZIKV_84	2	36	-10.5																															15
ZINC000012879754	ZIKV_144	2	36	-12.1																															12
	DENV_73	5	38	-10.9																															16
	YFV_65	5	38	-10.3																															18
ZINC000002108995	ZIKV_144	5	38	-11.1																															15
	DENV_73	3	37	-10.4																															15
	YFV_65	3	37	-10.8																															15
ZINC000013411589	ZIKV_84	3	37	-10.2																															13
	DENV_73	2	35	-10.8																															15
	YFV_162	2	35	-11.8																															19
ZINC000008792389	ZIKV_144	2	35	-11.7																															17
	DENV_73	2	36	-10.5																															19
	YFV_162	2	36	-11.4																															17
ZIKV_144	2	36	-12.1																																19

Interações

■ van der Waals	■ PonteH	■ Pi-Pi Stacked	■ Pi-Cation ou Pi-Anion	■ Doador doador desfavorável
■ LigH Conv.	■ Pi-Sigma	■ Alkyl ou Pi-Alkyl	■ Halogênio	

* apenas resíduos com pelo menos oito das 16 interações possíveis foram incluídos na tabela com intuito de facilitar a visualização das principais interações.

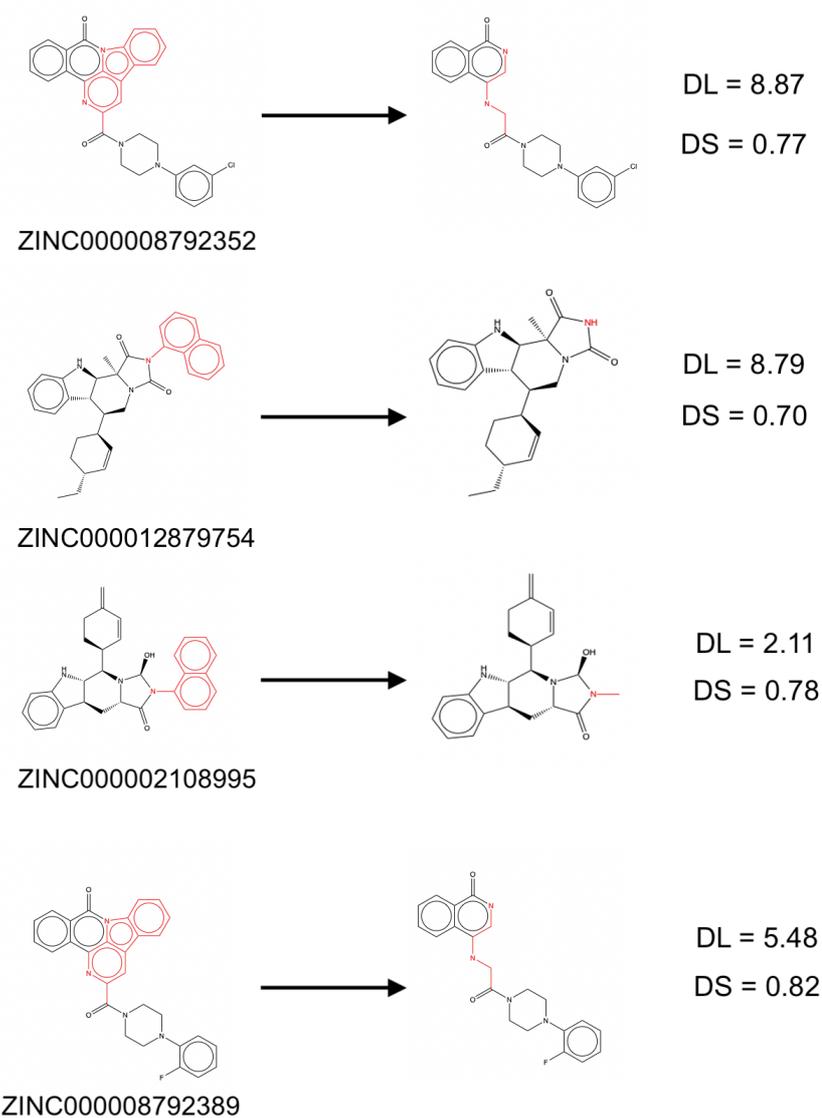
Por fim, realizou a avaliação do perfil ADMET de acordo com o programa OSIRIS. Para tal análise, apenas o composto ZINC000013411589 apresentou baixa toxicidade para os parâmetros analisados conforme apresentado na Tabela 4.

O programa OSIRIS é capaz de apontar qual fragmento da molécula está sendo responsável pelo risco médio ou elevado de toxicidade. A remoção dos fragmentos apontados pelo programa (Figura 52 – destacado em vermelho na estrutura) altera os parâmetros de toxicidade para “baixo” além de aumentar os escores de *druglikeness* (DL) e *drug-score* (DS) como observado na Figura 52.

Tabela 4 – Predição toxicidade, *druglikeness* e *drug-score* de acordo com OSIRIS dos cinco compostos em comum

Composto	Riscos de toxicidade				Scores	
	Risco mutagênico	Risco tumorigênico	Risco irritante	Risco reprodutivo	<i>Druglikeness</i>	<i>Drug-score</i>
ZINC000008792352	Médio	Baixo	Baixo	Baixo	8.84	0.21
ZINC000012879754	Elevado	Elevado	Baixo	Baixo	5.85	0.09
ZINC000002108995	Elevado	Elevado	Baixo	Baixo	1.08	0.1
ZINC000013411589	Baixo	Baixo	Baixo	Baixo	3.78	0.24
ZINC000008792389	Médio	Baixo	Baixo	Baixo	4.96	0.25

Figura 52 – Alteração estrutural das moléculas com média/elevada toxicidade

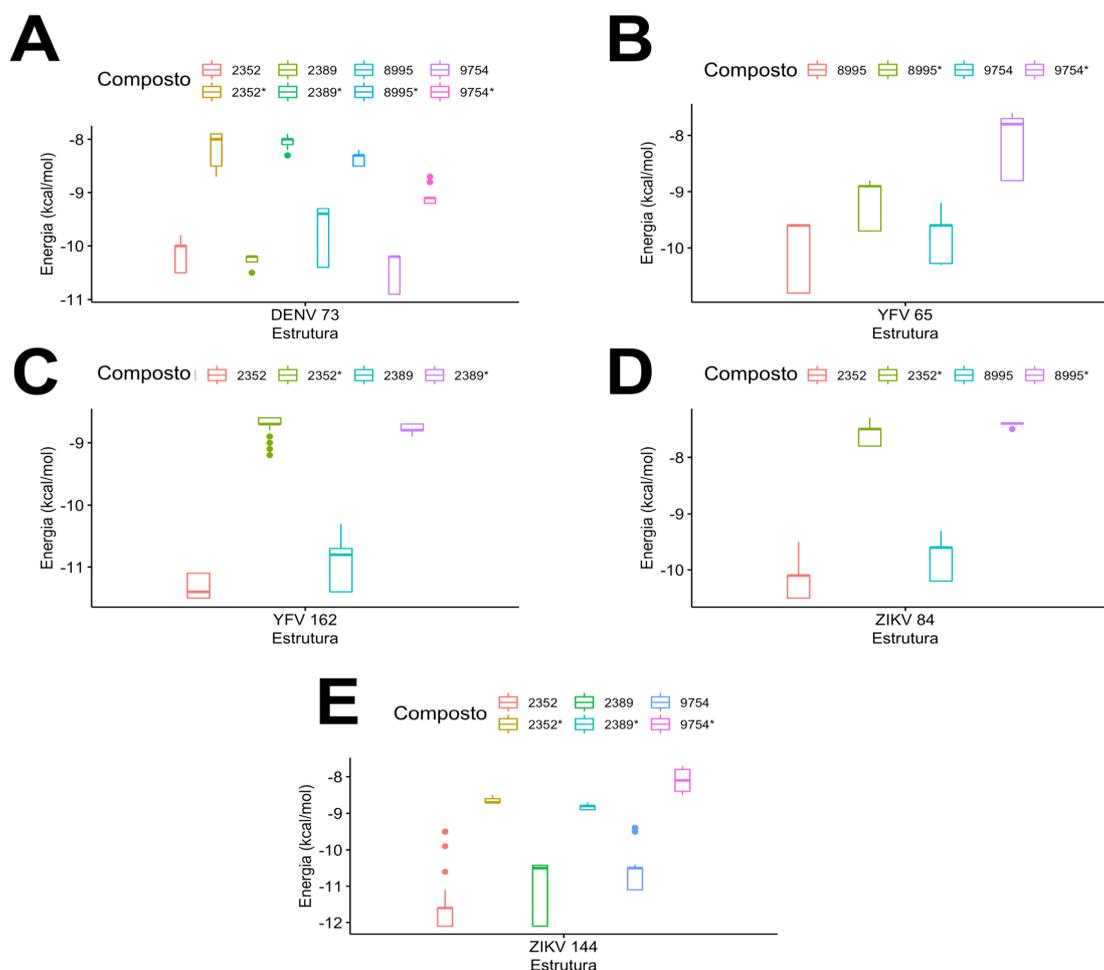


5.4.1.5 Ancoragem molecular dos compostos sem risco toxicológico

Para avaliar a interferência dessas alterações na molécula no escore de afinidade desses compostos, foram realizadas 100 simulações de ancoragem nas estruturas da NS1 onde o composto original foi classificado entre os 100 melhores. Os compostos com alteração serão aqui designados com o código original do ZINC adicionado um * ao final para assim saber a origem da molécula. Nos gráficos, serão todos representados pelos quatro últimos dígitos.

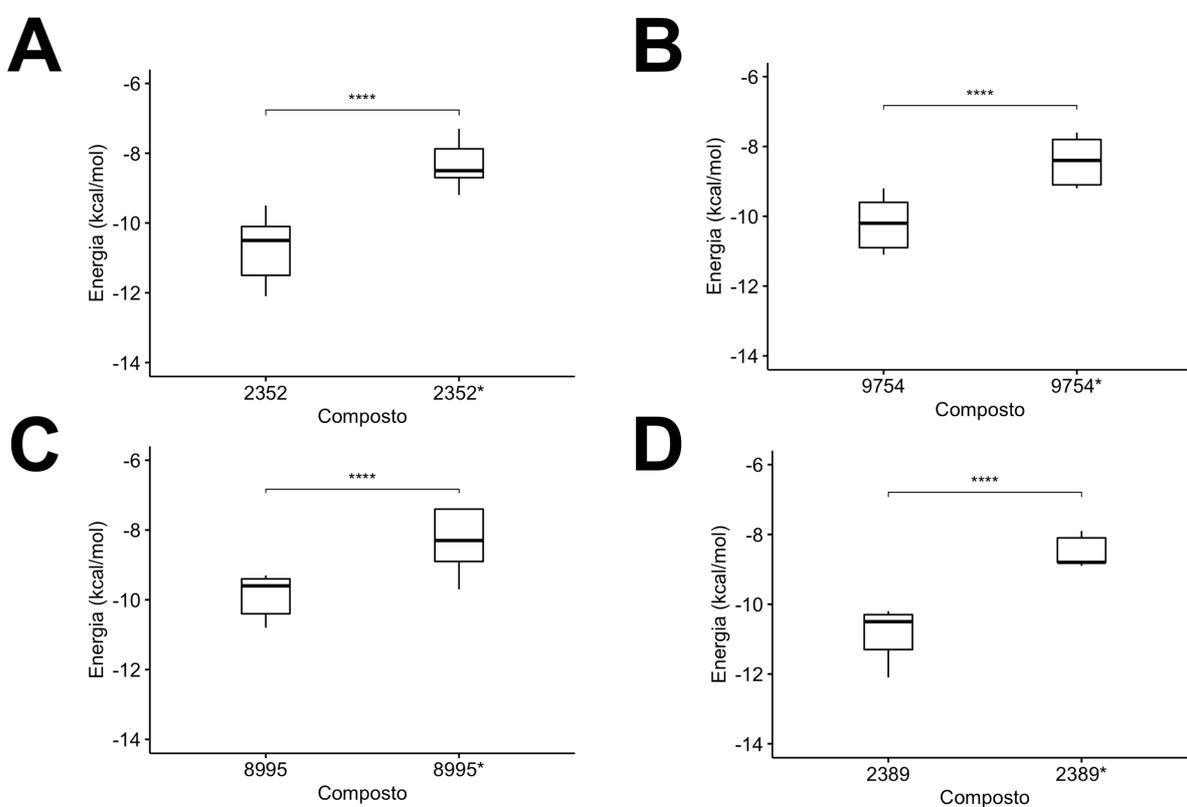
A análise comparativa da distribuição das 250 melhores energias dos compostos por estrutura da NS1 pode ser observada na Figura 53. Nota-se que para todas as estruturas e compostos houve um aumento no valor da energia de afinidade após a modificação na molécula.

Figura 53 – *Boxplot* da distribuição das energias dos compostos antes e após modificação para cada estrutura da NS1



A análise global (sem distinguir a estrutura) da distribuição das 250 melhores energias por composto também apresenta valores maiores de energia para a molécula sem o fragmento potencialmente tóxico quando comparado com a molécula sem alteração ($p < 0.0001$) (Figura 54). Para o composto ZINC000008792352* os valores variaram de -7.3 kcal/mol a -9.2 kcal/mol, com média de -8.28 kcal/mol. Já o composto ZINC000012879754* variou sua energia em -7.6 kcal/mol a -9.2 kcal/mol, com média de -8.43 kcal/mol. O composto ZINC000002108995* teve a energia variável de -7.4 kcal/mol a -9.7 kcal/mol, com média de -8.32 kcal/mol. Por fim, as energias obtidas para interação com o composto ZINC000008792389* foram entre -7.9 kcal/mol e -8.9 kcal/mol. Percebe-se que os valores de energias obtidos pelas novas moléculas ainda assim possuem potencial para uma boa interação ligante-proteína, contudo novas simulações *in silico* envolvendo dinâmica molecular com o ligante poderiam fornecer melhores informações acerca da estabilidade dos compostos na NS1 antes e após as alterações estruturais.

Figura 54 – Análise comparativa da distribuição global das energias dos compostos antes e após a modificação estrutural



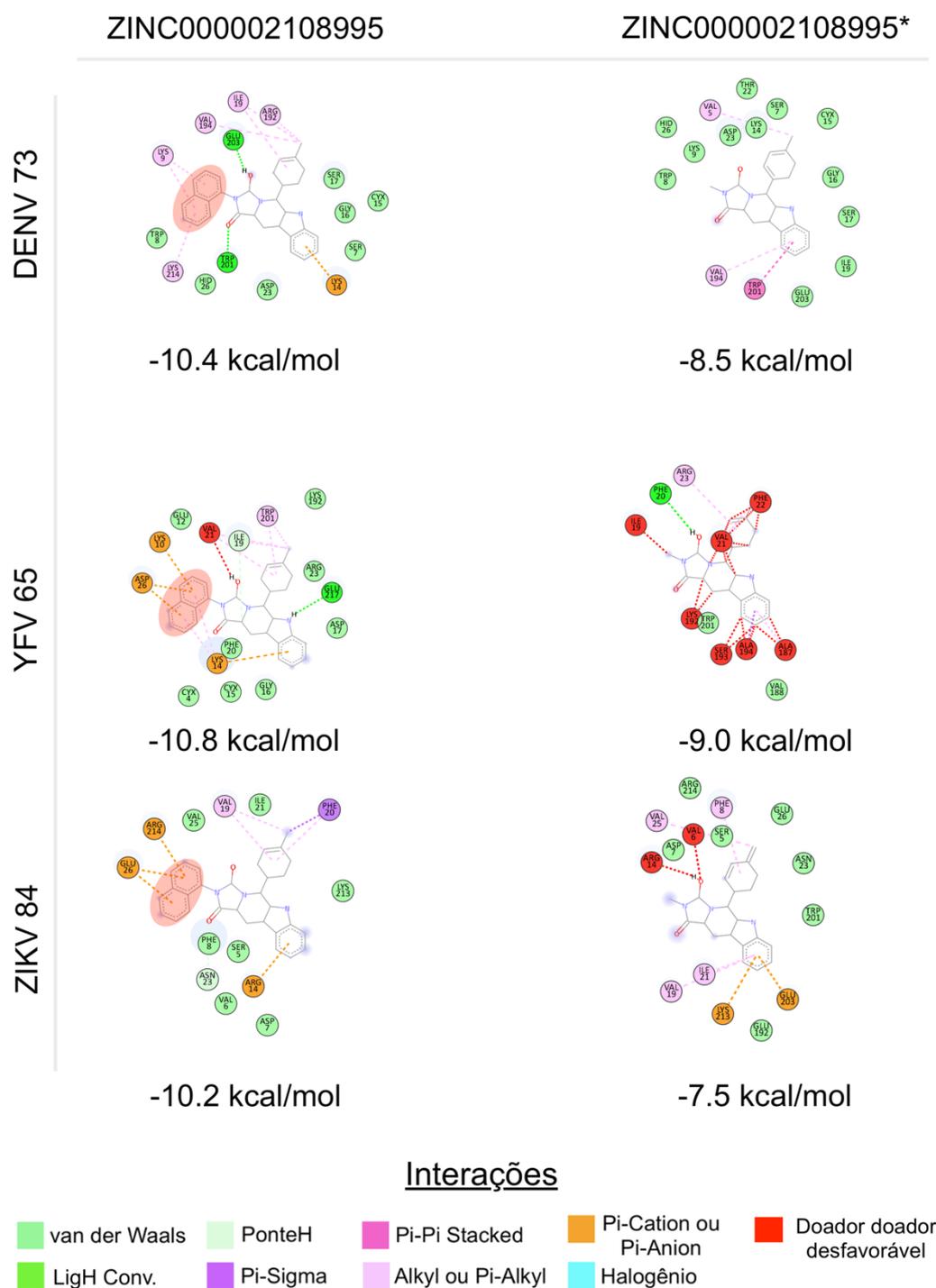
Já para o composto ZINC000012879754 não foi observado resíduos chaves interagindo na região removida da molécula, exceto para a estrutura do YFV 65 que tinha os resíduos PHE20 e VAL21 com interações do tipo van der Waals e LYS14 interagindo em ambos anéis aromáticos (Figura 57). Após a remoção dos anéis aromáticos, PHE20 e VAL21 apresentaram interações não-favoráveis e LYS14 não apareceu no diagrama. Na estrutura do DENV 73, os resíduos ILE19, ILE21 e TRP201 foram mantidos após a alteração na molécula, mas interagindo em diferente localização da molécula, sendo a interação com PHE20 ausente antes e após remoção do fragmento.

Já para o ZIKV144, entre os resíduos chaves que interagem com o composto sem alteração estavam apenas o TRP201 e ILE21. Após a remoção do fragmento tóxico, apenas a interação com o ILE21 se manteve, sendo que VAL19 e PHE20 apareceram na análise. Em resumo, os resíduos chaves não foram tão afetados pela remoção do fragmento, exceto em relação a estrutura do YFV 65.

A análise da interação antes e após a modificação do composto ZINC000002108995 não apresentou resíduos chave (LYS/ARG14, ILE/VAL19, PHE20 e TRP201) na região apontada como toxicológica (Figura 58). Para o DENV 73, antes da modificação dos compostos, os resíduos LYS14, ILE19 e TRP201 estavam presentes interagindo na molécula e se mantiveram após a remoção dos anéis aromáticos. Entretanto, por ajuste da molécula no sítio de ancoragem, a região de interação desses resíduos no composto foi alterada.

Na estrutura do YFV 65 todos os quatro resíduos estavam presentes interagindo na molécula sendo VAL21 de forma desfavorável. Após a alteração do composto, todas se mantiveram, sendo que a ILE19 e VAL21 com interações desfavoráveis, também em regiões diferentes da observada antes da remoção do fragmento. Para as interações na estrutura do DENV 84, o composto antes da modificação interagiu com os resíduos VAL19, PHE20 e ILE21. Após a modificação, apenas a interação com PHE20 foi perdida sendo que a interação com TRP201 foi adicionada.

Figura 58 - Diagrama 2D da interação entre proteína NS1 e ligante ZINC000002108995 antes e após modificação estrutural

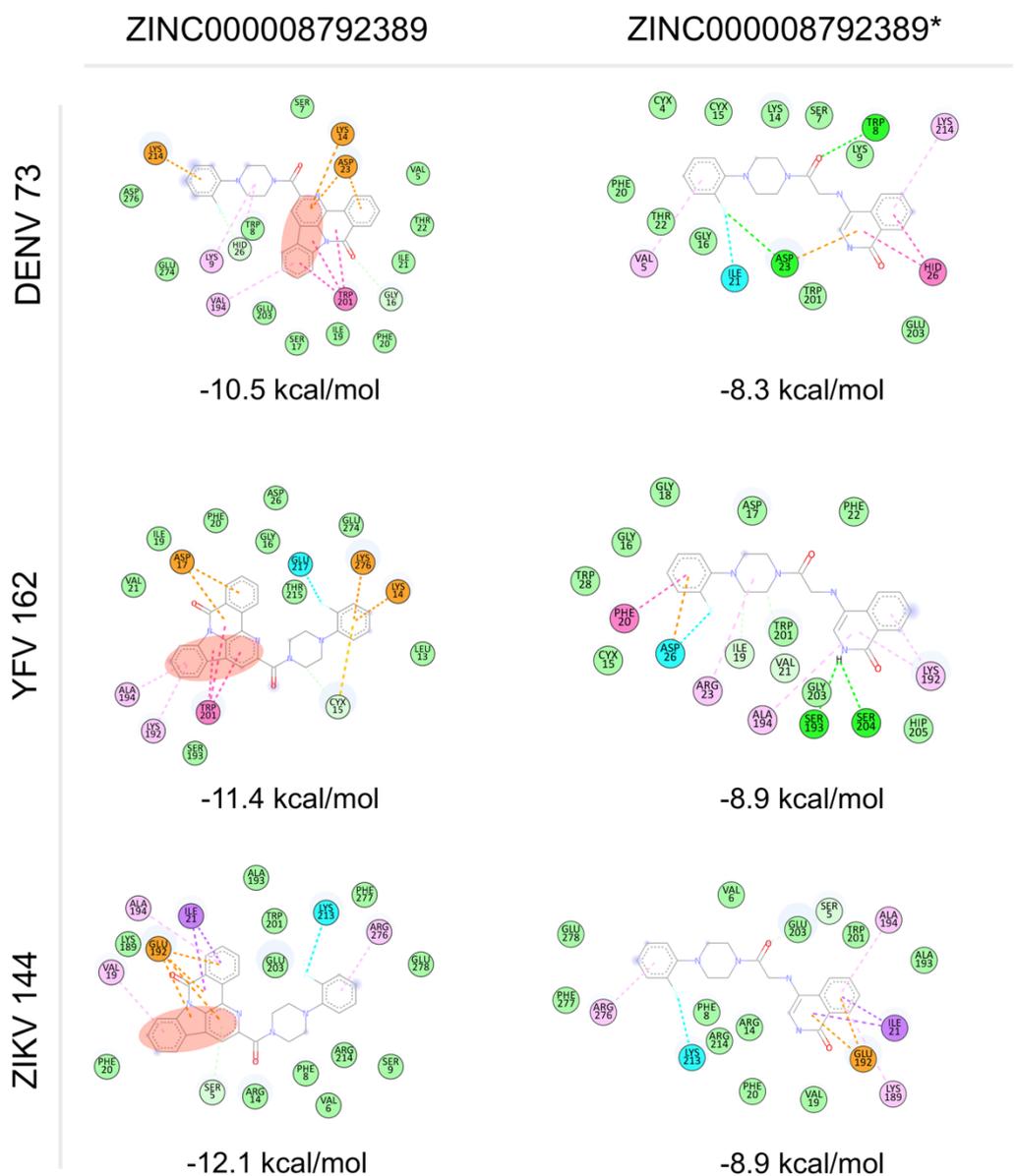


Na primeira coluna está representado o composto sem modificação, e na segunda coluna o composto com modificação. Nas linhas (horizontais, de cima para baixo estão as análises para as estruturas da NS1 do DENV 73, YFV 65 e ZIKV 84. O fragmento removido está destacado em vermelho na primeira coluna.

A análise do composto ZINC000008792389 antes e após a alteração na molécula pode ser observado na Figura 59. Para a estrutura do DENV 73, entre os resíduos chaves já citados anteriormente, o LYS14 e TRP201 interagiram com pelo menos um dos três anéis removidos da molécula, além de interação de van der Waals observadas com resíduos ILE19, PHE20 e ILE21. Após a remoção do fragmento, apenas a interação com ILE19 não foi mantida na análise. Na estrutura do YFV 162 todos os resíduos chaves apareceram na análise antes da remoção do fragmento, sendo o TRP201 o único que interagiu diretamente com a região removida. Após a modificação no composto, não foi observado na análise a interação com LYS14. Para o ZIKV144, também todos os resíduos chaves apareceram na análise, sendo VAL19 e PHE20 que estavam interagindo na região dos anéis removidos. As interações com esses resíduos foram mantidas mesmo após mudança na estrutura do composto.

Percebe-se que apesar do tipo de interação alterar, muitos resíduos se mantiveram interagindo nas duas situações (antes e após a modificação dos compostos). Essa região da cavidade, por ser uma interface de interação que se entrelaça com a mesma região de uma outra unidade monomérica, é rica de aminoácidos hidrofóbicos e, por isso, as estruturas com mais anéis aromáticos (sem modificação estrutural) tiveram o escore de afinidade favorecido. Entretanto, o risco toxicológico aumentado não permite classificar essas moléculas inalteradas como possíveis compostos com propriedades terapêuticas. As moléculas sem os fragmentos tóxicos apontados pelo servidor, apresentaram bom escore de energia e interações com resíduos importantes da proteína, sendo potenciais ligantes para acoplamento na região do *β -roll* da proteína NS1.

Figura 59 - Diagrama 2D da interação entre proteína NS1 e ligante ZINC000008792389 antes e após modificação estrutural



Interações

- | | | | | |
|---------------|----------|-------------------|-----------------------|----------------------------|
| van der Waals | PonteH | Pi-Pi Stacked | Pi-Cation ou Pi-Anion | Doador doador desfavorável |
| LigH Conv. | Pi-Sigma | Alkyl ou Pi-Alkyl | Halogênio | |

Na primeira coluna está representado o composto sem modificação, e na segunda coluna o composto com modificação. Nas linhas (horizontais, de cima para baixo estão as análises para as estruturas da NS1 do DENV 73, YFV 162 e ZIKV 144. O fragmento removido está destacado em vermelho na primeira coluna.

5.4.2 Critério de eficiência

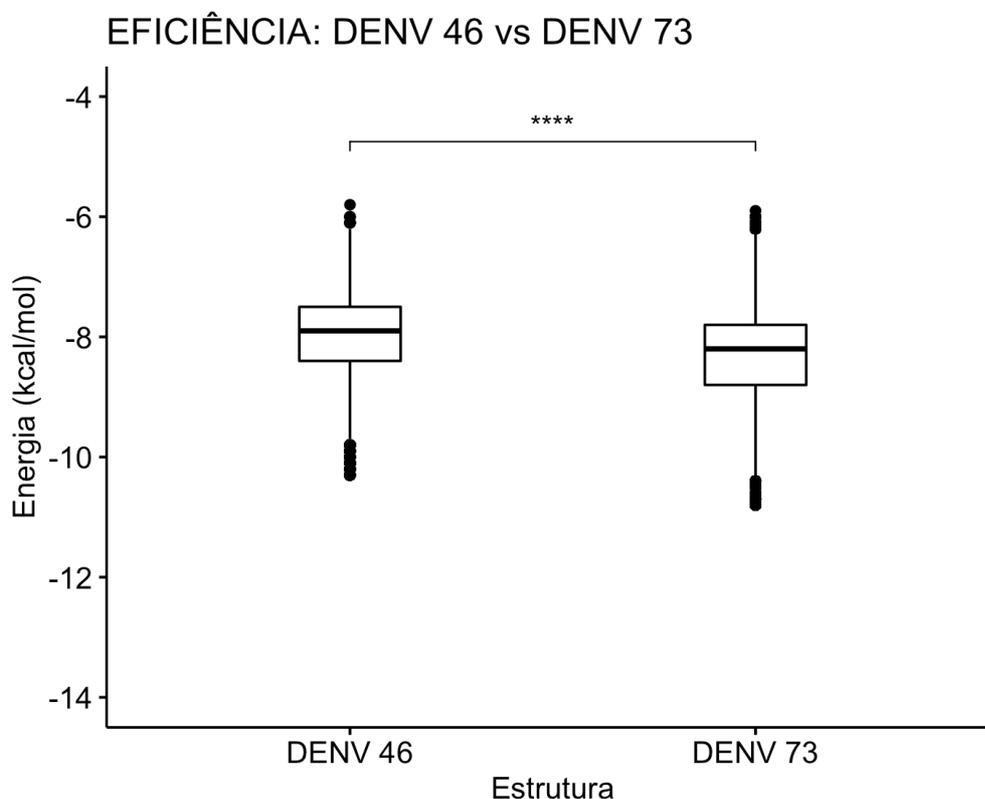
5.4.2.1 Comparação da triagem virtual entre estruturas da NS1 do DENV

Na análise de critério de afinidade, ou seja, considerando o número de átomos do ligante de acordo com a equação 4, selecionou-se os 100 melhores compostos classificados como mais eficientes para cada estrutura de NS1. Após 100 simulações exaustivas para cada um dos 100 compostos ranqueados, a análise das energias de acordo com o escore do AutoDock Vina foi realizada.

Para as estruturas da NS1 do DENV, houve diferença estatisticamente significativa ($p < 0.0001$) entre as energias obtidas na ancoragem desses 100 compostos para cada uma das estruturas (DENV 46 e DENV 73) (Figura 60). Para a estrutura da NS1 do DENV 46 a maior energia foi de -5.8 kcal/mol e menor energia foi de -10.3 kcal/mol. Já para o DENV 73 esses valores foram de -5.9 kcal/mol e -10.8 kcal/mol para maior e menor energia, respectivamente.

Na análise comparativa dos compostos classificados entre os 100 melhores de acordo com o critério de eficiência entre as estruturas do DENV 46 e DENV 73, não foi encontrado nenhum composto em comum, diferente do observado na análise por critério de afinidade.

Figura 60 – *Boxplot* das distribuições de energia das 100 simulações exaustivas para os 100 compostos ranqueados de acordo com critério de eficiência para as duas estruturas da NS1 do DENV.



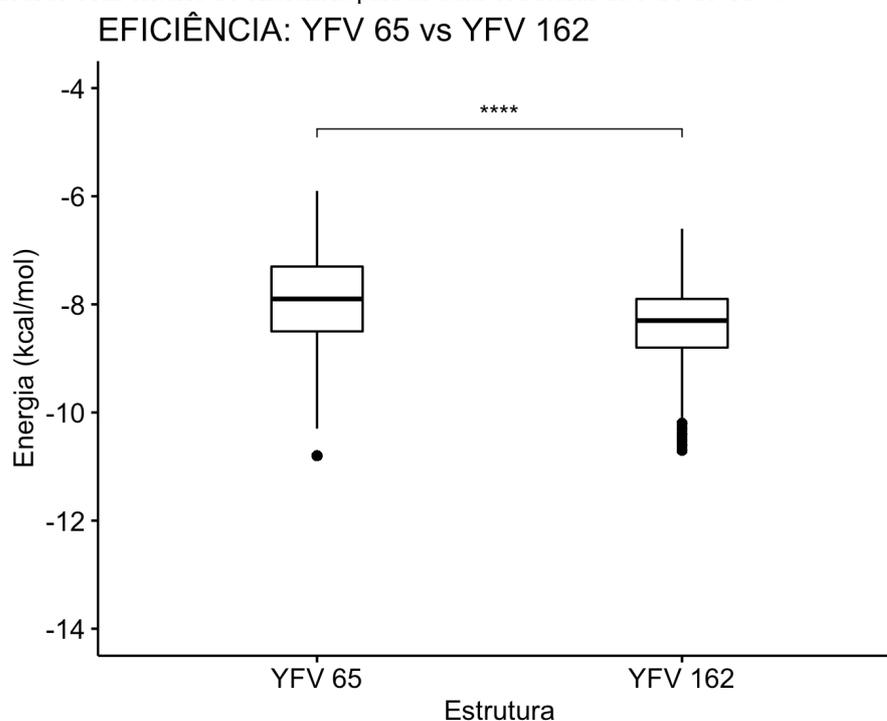
**** $p < 0.0001$

5.4.2.2 Comparação da triagem virtual entre estruturas da NS1 do YFV

Assim como pro DENV, para as estruturas da NS1 do YFV também houve diferença estatisticamente significativa ($p < 0.0001$) entre as energias dos compostos ranqueados entre os 100 melhores (Figura 61). Para a estrutura da NS1 YFV 65 a maior energia observada foi de -5.9 kcal/mol e a menor foi de -10.8 kcal/mol. Já para a estrutura YFV 162 os valores observados de maior energia foi de -6.6 kcal/mol e menor de -10.7 kcal/mol.

Entre os 100 compostos ranqueados como mais eficientes para as duas estruturas do YFV, 11 deles foram encontrados simultaneamente na classificação da NS1 YFV 65 e YFV 162 (Apêndice E), sendo a maior concordância entre estruturas até então observada.

Figura 61 – *Boxplot* das distribuições de energia das 100 simulações exaustivas para os 100 compostos ranqueados de acordo com critério de eficiência para as duas estruturas da NS1 do YFV.



**** $p < 0.0001$

No diagrama 2D das interações (Figura 62), nota-se que há resíduos que estão interagindo com os onze compostos para ambas as estruturas da NS1 do YFV como VAL 21 e TRP201. Esses dois resíduos também estavam presentes análises das interações das estruturas da NS1 do YFV usando o critério de afinidade. Outros resíduos também se caracterizaram como chaves nessas interações como CYX15, GLY16, ASP17, ILE19, PHE20, ARG23 e ASP26 por estarem bastante presentes nas interações. Em comparação à análise da afinidade, algumas regiões da proteína perderam interações (LEU13, LEU

219, GLU274, VAL275 e HIP324). Esses resíduos estão fora da região definida como cavidade (resíduos 1-28 e 182-216), sendo, portanto, aqui evidenciado uma especificidade maior no acoplamento desses compostos menores na região (cavidade) de interesse.

Figura 62 - Representação do diagrama 2D de interação dos onze compostos em comum às duas estruturas do YFV com seus respectivos escores de energia

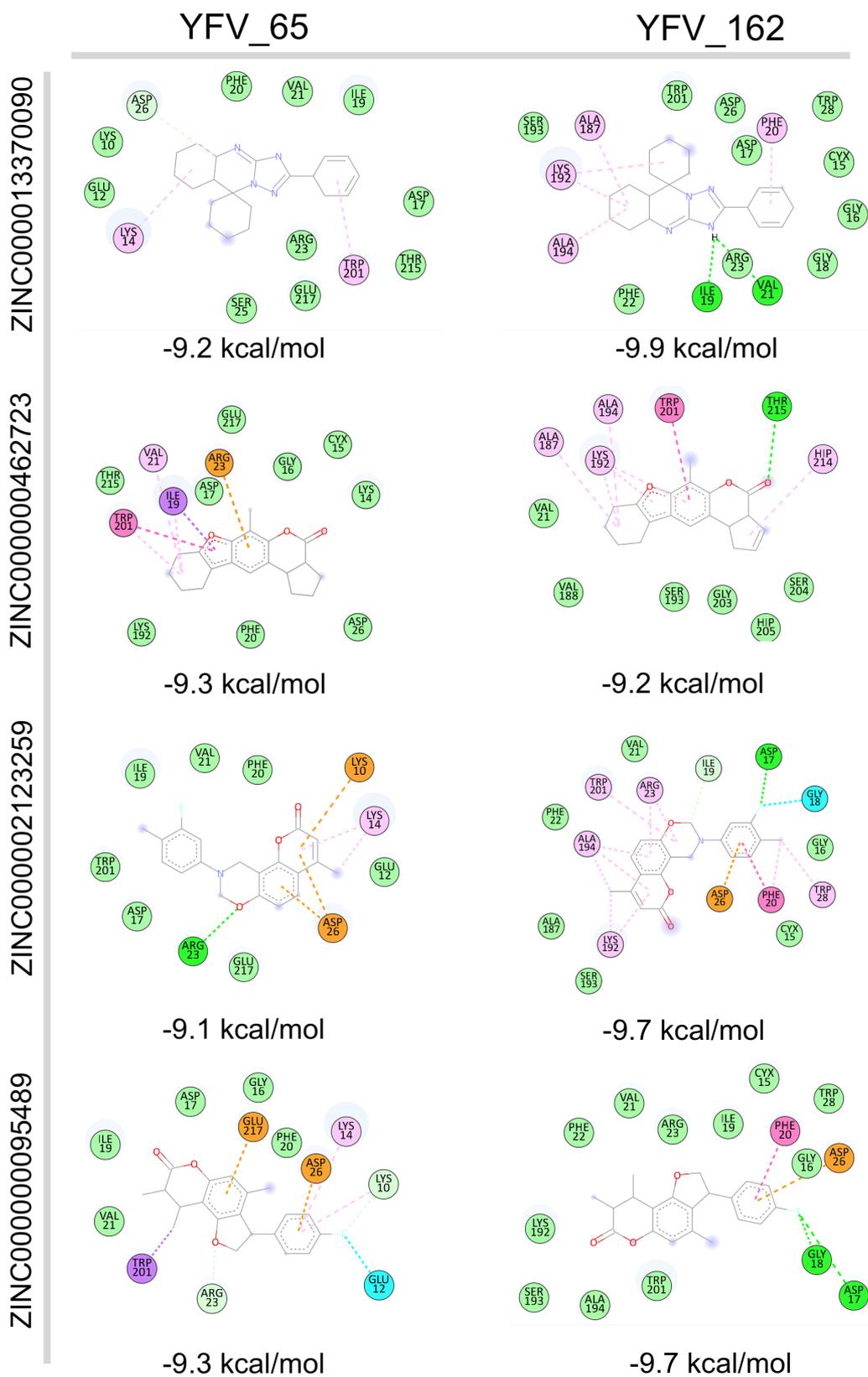
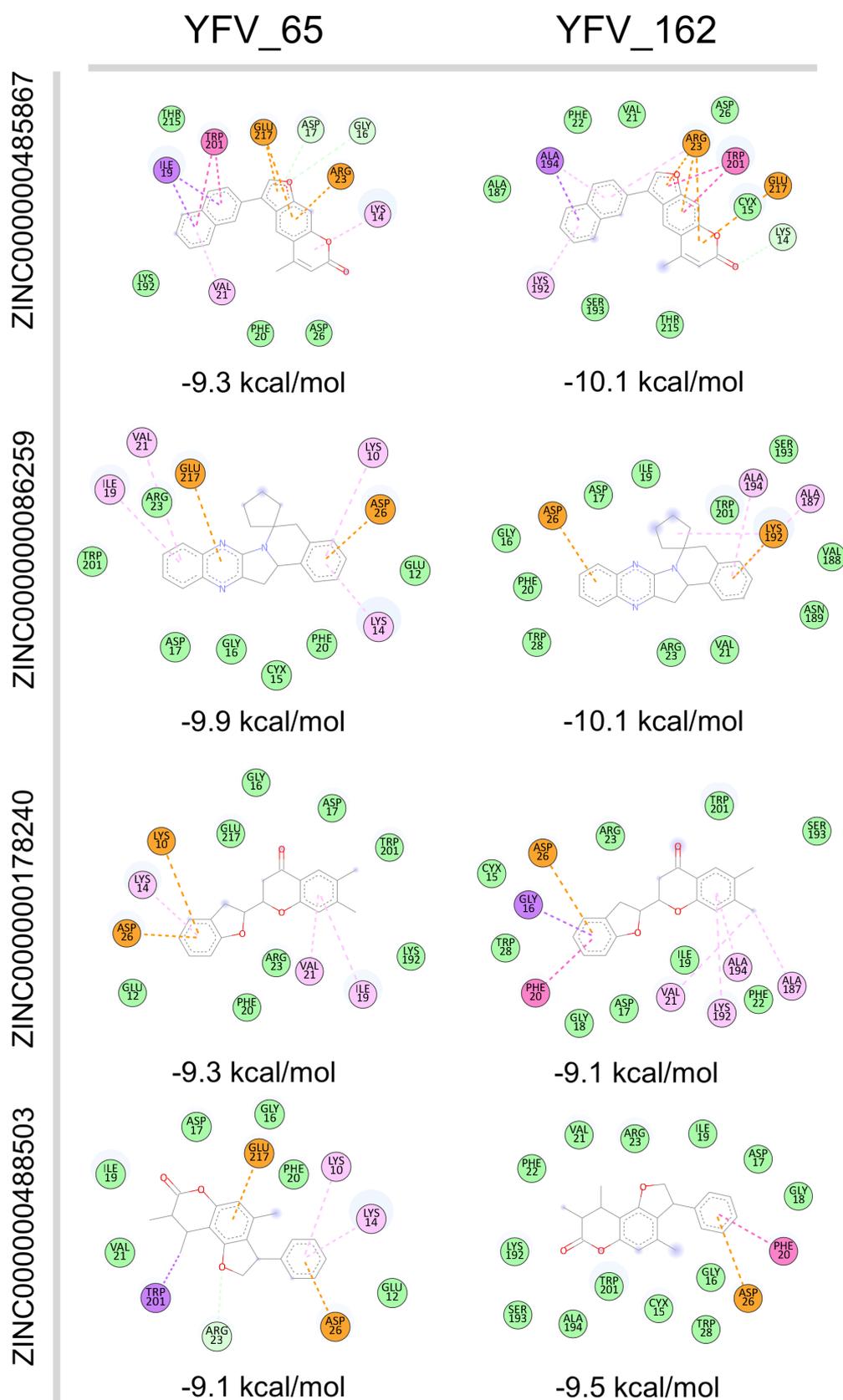


Figura 62 (continuação) - Representação do diagrama 2D de interação dos onze compostos em comum às duas estruturas do YFV com seus respectivos escores de energia



Assim como realizado nas análises dos compostos classificados pelo critério de afinidade, também foi realizada a análise das distâncias entre resíduos da proteína e átomos do ligante. Essa análise nos permite ter uma ideia de especificidade entre fragmento do composto e local de ancoragem na proteína. Para o ligante ZINC000013370090 uma região considerável da molécula apresentou interações em comum com os resíduos ASP17, ILE19, PHE20, VAL21 e ARG23 (Figura 63).

Figura 63 - Análise fragmento-específico do ligante ZINC000013370090 no sítio da NS1 do YFV

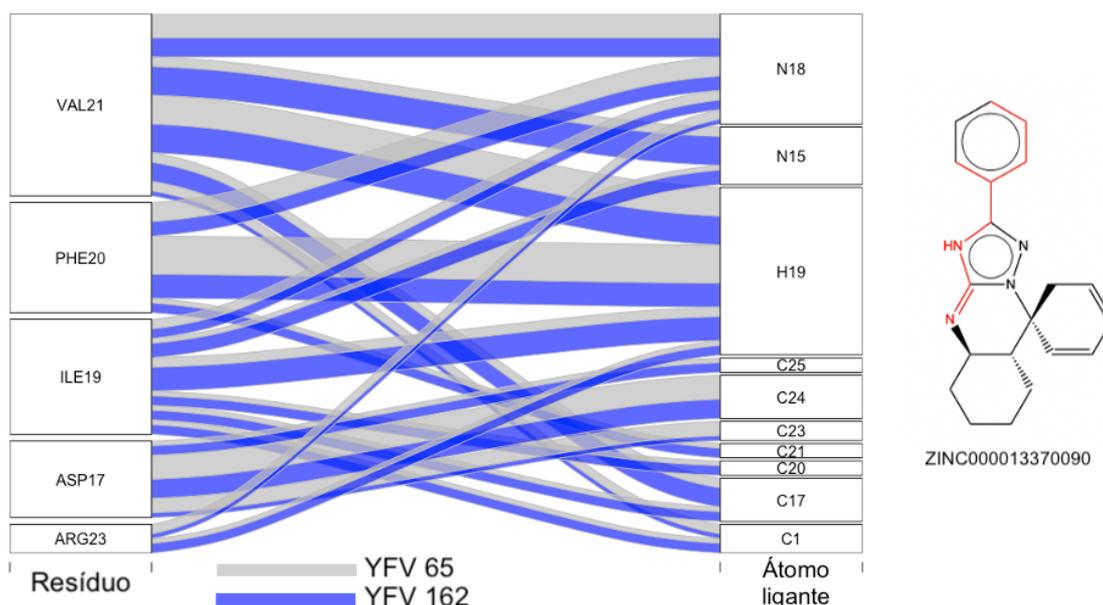


Diagrama ilustrando o local de ancoragem do fragmento ZINC000013370090 em comum nas duas estruturas da NS1. Na representação da estrutura do composto, destaca-se (em vermelho) os fragmentos em comum (coluna “Átomo ligante” do diagrama).

Já a análise realizada para o ligante ZINC000000462723 mostrou dois pequenos fragmentos com especificidade em relação ao resíduo com destaque pra os resíduos ASP17, ILE19, LYS192 e TRP201 (Figura 64).

No ligante ZINC000002123259 um considerável fragmento em comum nas ancoragens com as duas estruturas da NS1 do YFV (Figura 65). Esse fragmento possui bastante similaridade com o fragmento observado na análise para o ZINC000013370090, com destaque para a região do anel aromático. Nota-se também que entre esses dois compostos, quatro resíduos (ASP17, ILE19, PHE20 e ARG23) apareceram nas duas análises, indicando uma possível preferência por esse fragmento nessa região da cavidade (Figura 66).

Figura 64 - Análise fragmento-específico do ligante ZINC000000462723 no sítio da NS1 do YFV

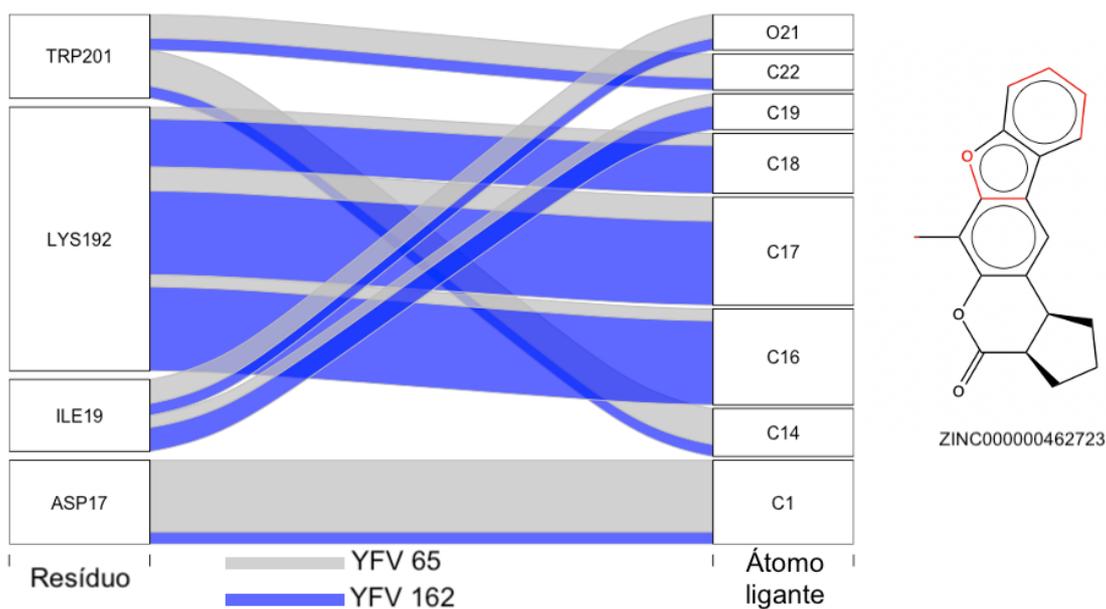


Diagrama ilustrando o local de ancoragem do fragmento ZINC000000462723 em comum nas duas estruturas da NS1. Na representação da estrutura do composto, destaca-se (em vermelho) os fragmentos em comum (coluna “Átomo ligante” do diagrama).

Figura 65 - Análise fragmento-específica do ligante ZINC000002123259 no sítio da NS1 do YFV

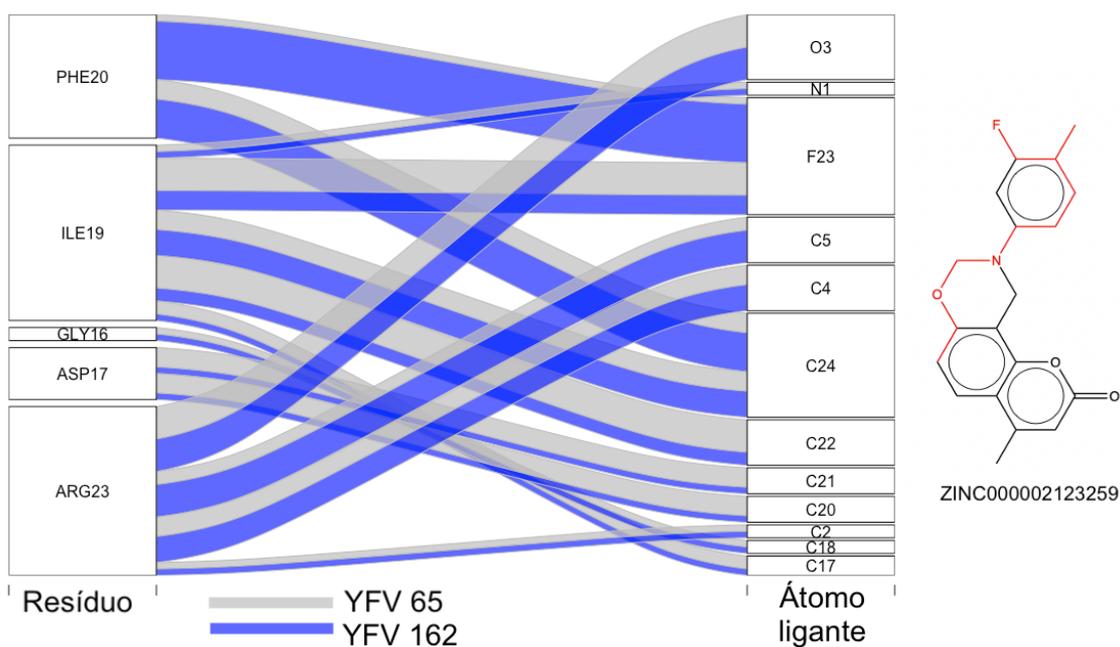


Diagrama ilustrando o local de ancoragem do fragmento ZINC000002123259 em comum nas duas estruturas da NS1. Na representação da estrutura do composto, destaca-se (em vermelho) os fragmentos em comum (coluna “Átomo ligante” do diagrama).

Figura 66 – Pose de ancoragem dos ligantes ZINC000013370090 e ZINC000002123259

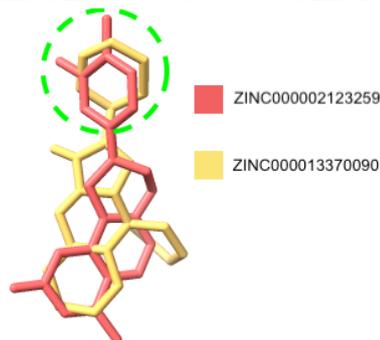


Ilustração dos ligantes ZINC000013370090 e ZINC000002123259 em suas respectivas conformações de ligação na cavidade da NS1 YFV 65 como exemplo. É possível observar que os fragmentos similares (anel aromático com círculo verde destacando) estão de formas semelhantes acoplados ao sítio.

Na análise do composto ZINC000000095489 observou-se um grande fragmento com preferência de proximidade de certos resíduos (CYX15, ILE19, PHE20, VAL21, ARG23, ASP26 e TRP201) independente da estrutura de NS1 do YFV (Figura 67). Nota-se que parte desse fragmento (anel aromático ligado ao flúor) é bastante similar ao do composto anteriormente analisado (ZINC000002123259). Contudo, nota-se que apenas dois resíduos foram em comum nas duas análises (ILE19 e ARG23) sugerindo um diferente modo de ligação entre eles. Ao abrir os modos de ligação desses dois compostos na estrutura da NS1 YFV 65, observa-se que de fato as regiões onde o flúor (átomo verde) se liga em cada composto estão em lados opostos (Figura 68).

Figura 67 - Análise fragmento-específico do ligante ZINC000000095489 no sítio da NS1 do YFV

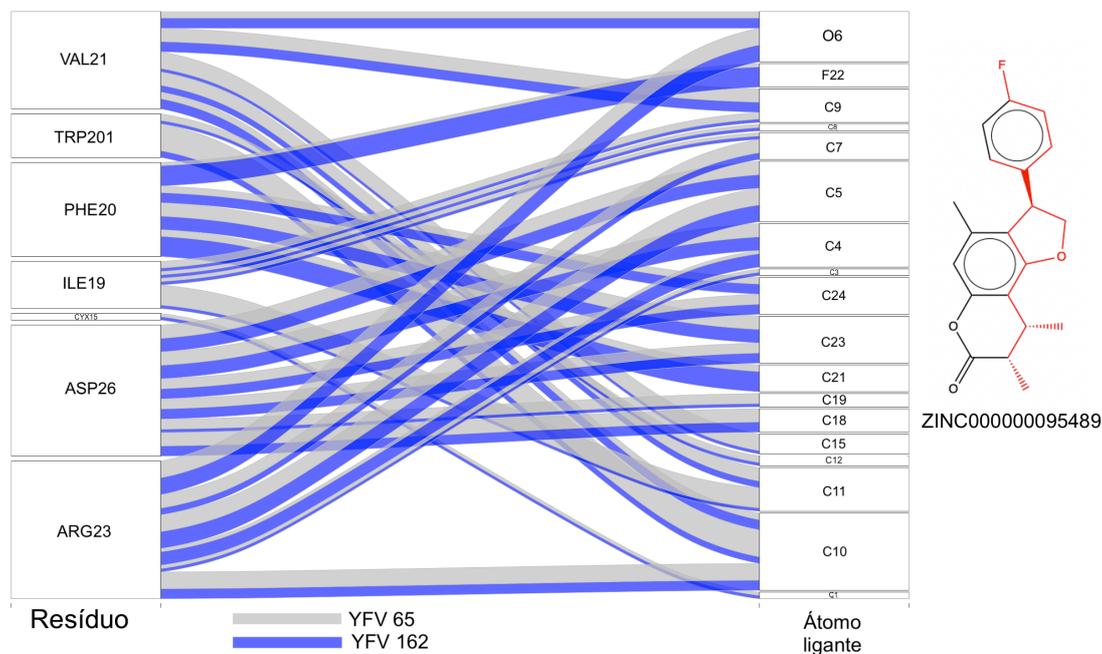


Diagrama ilustrando o local de ancoragem do fragmento ZINC000000095489 em comum nas duas estruturas da NS1. Na representação da estrutura do composto, destaca-se (em vermelho) os fragmentos em comum (coluna “Átomo ligante” do diagrama).

Figura 68 - Pose de ancoragem dos ligantes ZINC000002123259 e ZINC000000095489

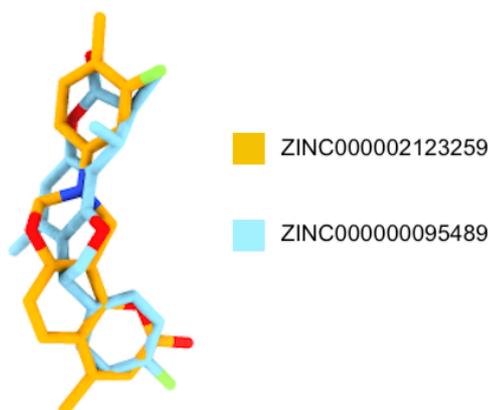


Ilustração dos ligantes ZINC000002123259 e ZINC000000095489 em suas respectivas conformações de ligação na cavidade da NS1 YFV 65 como exemplo. Nota-se que as regiões similares (com átomo de flúor em verde) estão de formas distintas acoplados ao sítio.

Já para a análise do composto ZINC000000485867 uma região composta por dois anéis aromáticos e uma outra composta um grupamento éter e uma cetona, além de mais outros dois átomos de carbono isolados ligados a um outro anel aromático foram observados por estarem próximos aos mesmos resíduos nas duas NS1 (Figura 69). Os resíduos que se destacaram na análise foram LYS14, CYX15, ILE19, ARG23, LYS192 e TRP201.

Figura 69 - Análise fragmento-específico do ligante ZINC000000485867 no sítio da NS1 do YFV

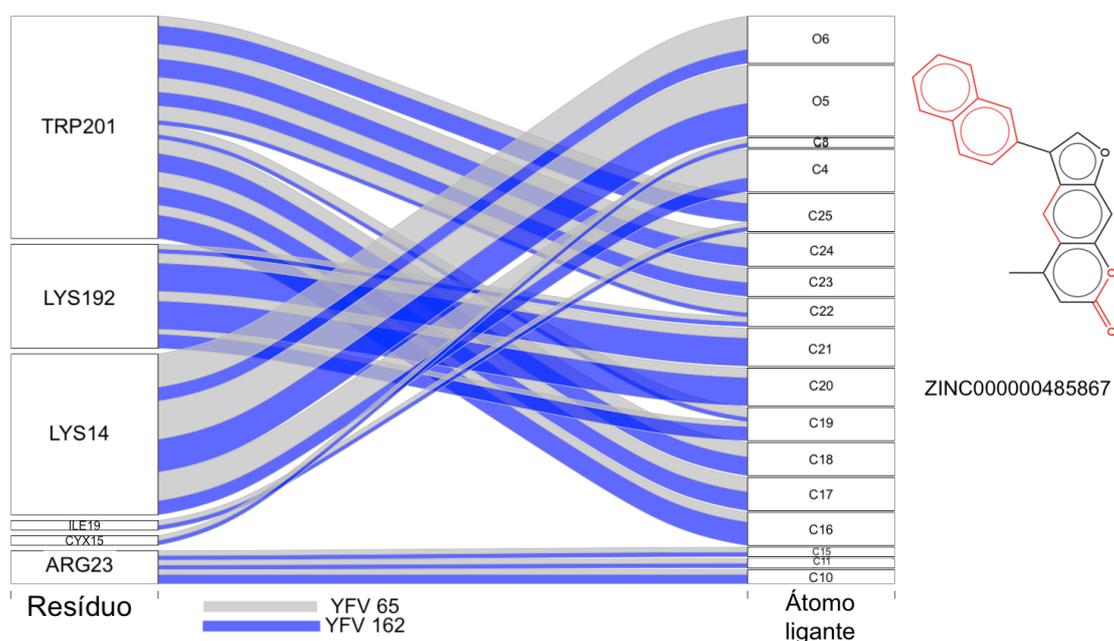


Diagrama ilustrando o local de ancoragem do fragmento ZINC000000485867 em comum nas duas estruturas da NS1. Na representação da estrutura do composto, destaca-se (em vermelho) os fragmentos em comum (coluna “Átomo ligante” do diagrama).

O ligante ZINC000000086259 apresentou uma pequena região em interagindo de forma similar com os resíduos da proteína, sendo uma parte de um anel aromático e outra contendo átomo de nitrogênio (Figura 70). Apenas os resíduos ILE19 e PHE20 apareceram nessa análise

Figura 70 - Análise fragmento-específico do ligante ZINC000000086259 no sítio da NS1 do YFV

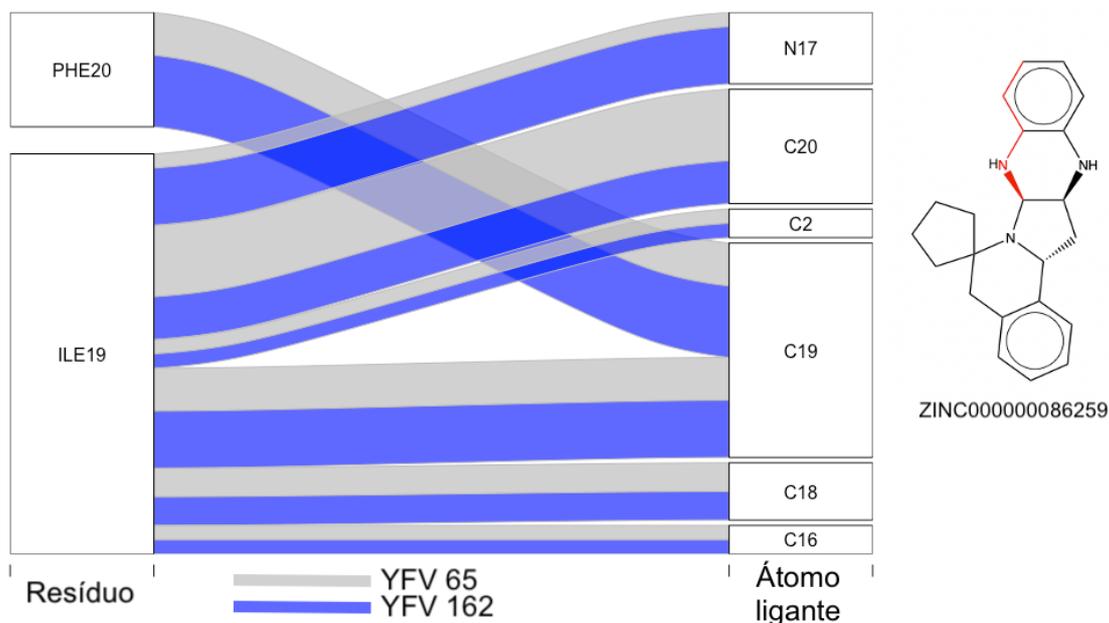


Diagrama ilustrando o local de ancoragem do fragmento ZINC000000086259 em comum nas duas estruturas da NS1. Na representação da estrutura do composto, destaca-se (em vermelho) os fragmentos em comum (coluna “Átomo ligante” do diagrama).

Praticamente toda a estrutura do ligante ZINC000000178240 se manteve a uma distância $< 5 \text{ \AA}$ dos mesmos resíduos da proteína independente da estrutura de NS1 do YFV (Figura 71). Parte dessa região é composta por dois anéis cíclicos contendo um grupamento éter, além de fragmentos de outros dois anéis cíclicos também com um grupamento éter. Na análise destaca-se a presença dos resíduos ILE19, PHE20, VAL21, ARG23, ASP26, LYS192 e TRP201.

Assim como para o ligante anterior, o composto ZINC000000488503 também apresentou uma grande parte da molécula próxima aos mesmos resíduos da NS1 independente da estrutura da proteína, indicando um similar modo de ligação (Figura 72). Destaca-se mais uma vez a presença de um anel aromático e grupamento éter em uma cadeia cíclica. Nessa análise, os resíduos em evidência foram CYX15, ILE19, PHE20, VAL21, ARG23, ASP26 e TRP201 que já estiveram presentes em análises anteriores.

Na análise realizada para o ligante ZINC000065248686, pequenos fragmentos foram em comum distância ($< 5 \text{ \AA}$) à quatro resíduos da proteína (GLY16, ASP17, ILE19 e PHE20) (Figura 73), sugerindo uma maior divergência desse composto em relação à forma de ancoragem nas diferentes estruturas da NS1 do YFV.

Figura 71 - Análise fragmento-específico do ligante ZINC000000178240 no sítio da NS1 do YFV

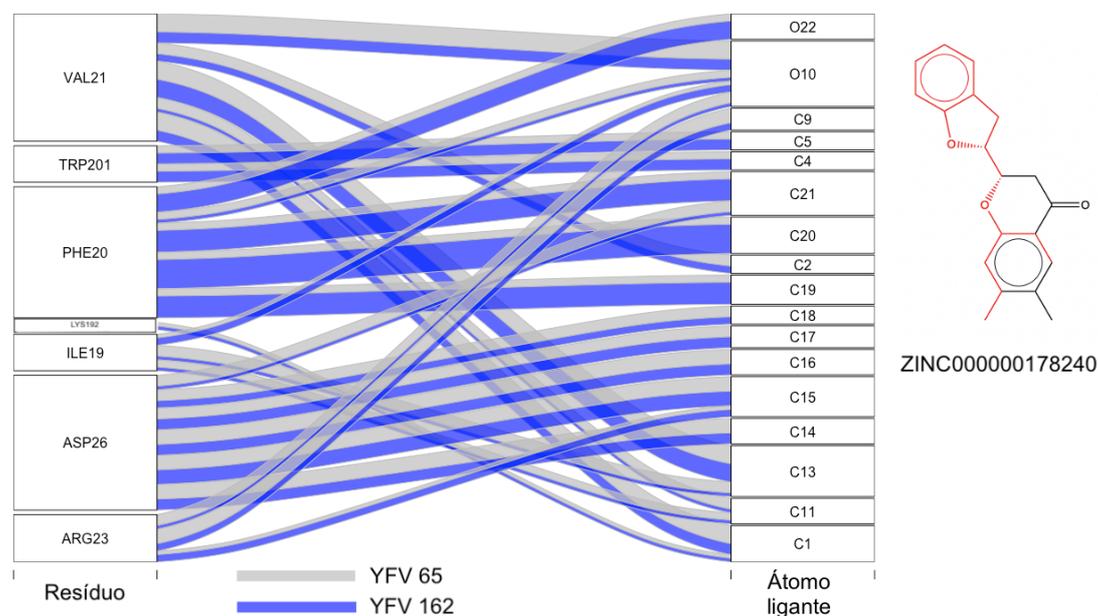


Diagrama ilustrando o local de ancoragem do fragmento ZINC000000178240 em comum nas duas estruturas da NS1. Na representação da estrutura do composto, destaca-se (em vermelho) os fragmentos em comum (coluna “Átomo ligante” do diagrama).

Figura 72 - Análise fragmento-específico do ligante ZINC000000488503 no sítio da NS1 do YFV

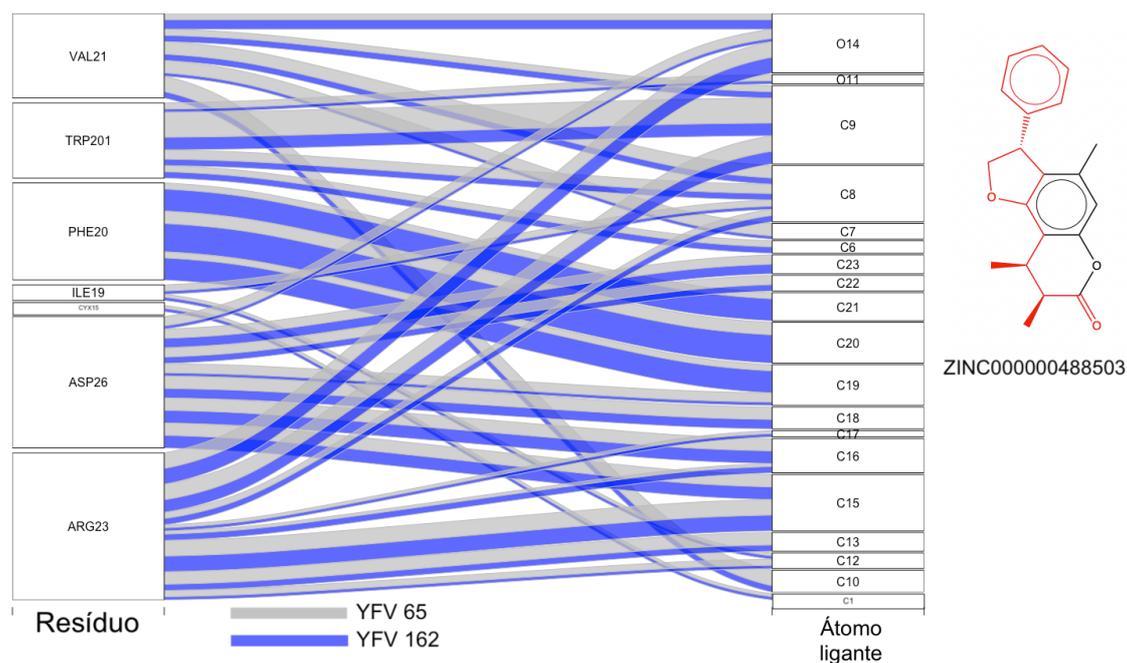


Diagrama ilustrando o local de ancoragem do fragmento ZINC000000488503 em comum nas duas estruturas da NS1. Na representação da estrutura do composto, destaca-se (em vermelho) os fragmentos em comum (coluna “Átomo ligante” do diagrama).

Figura 73 - Análise fragmento-específico do ligante ZINC000065248686 no sítio da NS1 do YFV

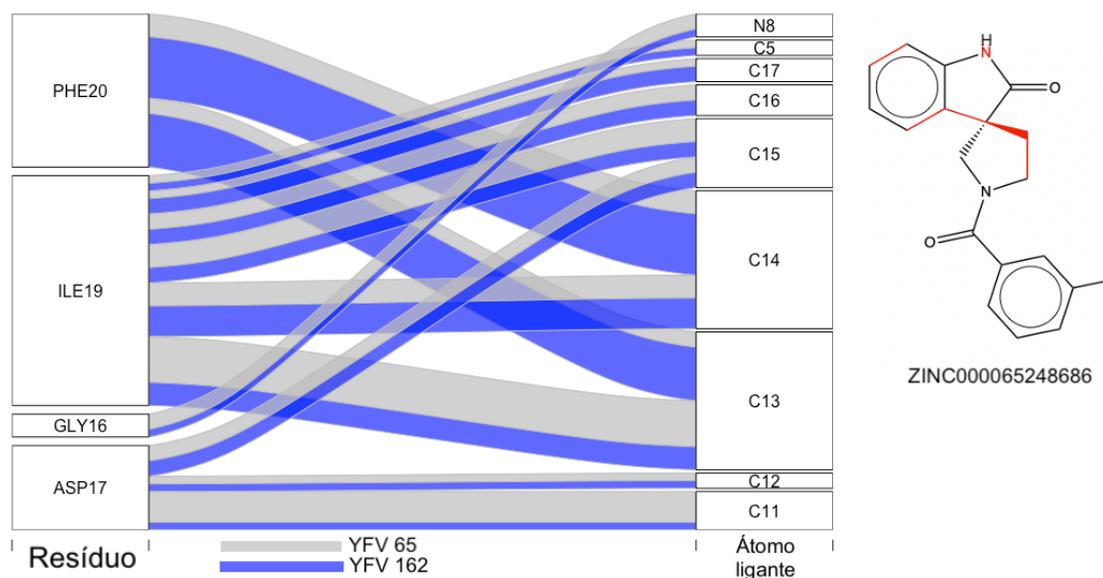


Diagrama ilustrando o local de ancoragem do fragmento ZINC000065248686 em comum nas duas estruturas da NS1. Na representação da estrutura do composto, destaca-se (em vermelho) os fragmentos em comum (coluna “Átomo ligante” do diagrama).

Para o ligante ZINC000065248870 (Figura 74) foi possível observar que um fragmento contendo parte de um anel aromático com o átomo e flúor ligado a ele, com destaque para os resíduos LYS14, GLY16, ILE19, PHE20, VAL21, ARG23, ASP26, LYS192 e TRP201. Esse fragmento com o átomo de flúor é muito similar ao das análises dos compostos ZINC000002123259 e ZINC000000095489 (Figuras 65 e 67, respectivamente). Nota-se, contudo, que há maior concordância de resíduos entre o composto ZINC000065248870 e ZINC000000095489, sugerindo que eles possuem um mesmo modo de ligação no sítio. Na Figura 75, é possível observar que diferente do que aconteceu na análise dos compostos ZINC000002123259 e ZINC000000095489, a orientação do átomo de flúor está similar entre o ZINC000065248870 e ZINC000000095489.

Apesar de uma certa similaridade estrutural entre os ligantes ZINC000065248686 e ZINC000065249270, os resultados nas análises foram distintos. Enquanto que para o primeiro houve pouca concordância entre ligante e distância de resíduo para as duas estruturas de NS1, para o composto ZINC000065249270 grande parte da molécula esteve a uma distância $< 5 \text{ \AA}$ dos mesmos resíduos para as duas estruturas da NS1 do YFV (Figura 76), indicado um similar encaixe no sítio independente da estrutura. Em relação aos resíduos do sítio, se destacaram a LYS14, ILE19, PHE20, VAL21, ARG23, ASP26, LYS192 e TRP201.

Figura 74 - Análise fragmento-específico do ligante ZINC000065248870 no sítio da NS1 do YFV

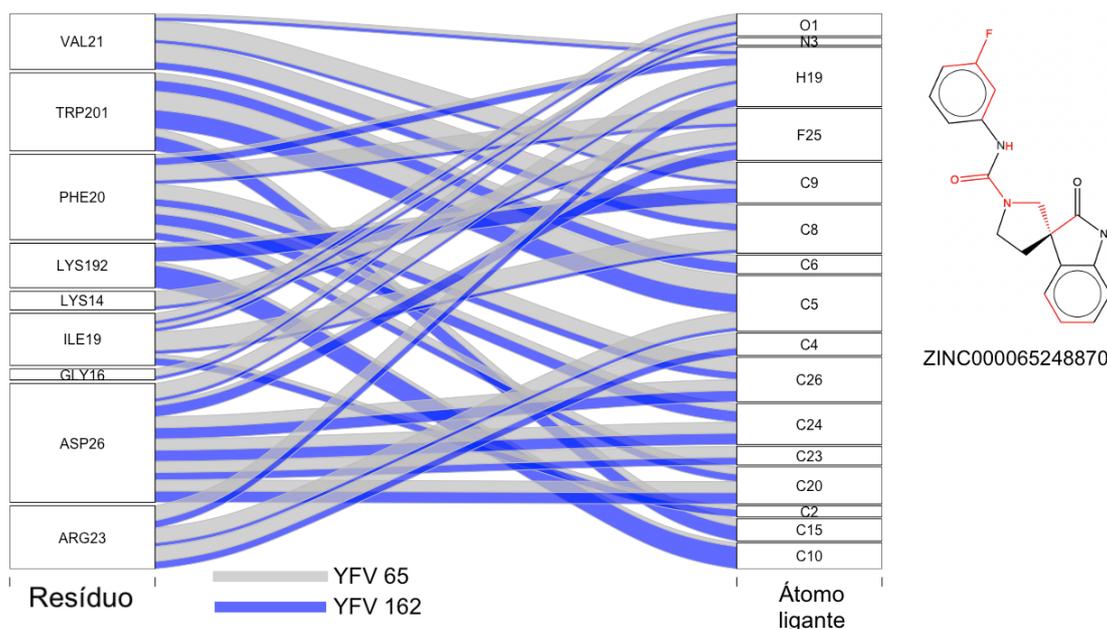


Diagrama ilustrando o local de ancoragem do fragmento ZINC0000652488706 em comum nas duas estruturas da NS1. Na representação da estrutura do composto, destaca-se (em vermelho) os fragmentos em comum (coluna “Átomo ligante” do diagrama).

Figura 75 – Conformações da ancoragem dos compostos ZINC000065248870 e ZINC000000095489

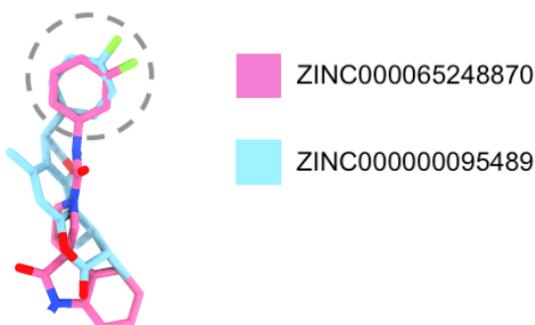


Ilustração dos ligantes ZINC000065248870 e ZINC000000095489 em suas respectivas conformações de ligação na cavidade da NS1 YFV 65 como exemplo. Nota-se que as regiões similares (com átomo de flúor em verde, delimitado pelo círculo tracejado em cinza) estão de formas similares acoplados ao sítio.

Nas análises dos 11 compostos em comum é possível destacar alguns aminoácidos que apareceram em vários resultados como ILE19, PHE20 e ARG23, sendo que o primeiro apareceu em todas. Além desses, LYS14, CYX15, GLY16, ASP17, VAL21, ASP26, LYS192 e TRP201 apareceram em mais de uma análise. Esses resíduos já foram destacados como importantes em outras análises desse estudo. Curiosamente, a PHE22 apesar da proximidade espacial com a PHE20, não se destacou em nenhuma análise realizada para as estruturas da YFV tanto na classificação dos ligantes por critério de eficiência quando afinidade. Esse fato reforça a importância da PHE20, que é um resíduo conservado entre as espécies aqui estudada, diferente da PHE22.

Figura 76 - Análise fragmento-específico do ligante ZINC000065249270 no sítio da NS1 do YFV

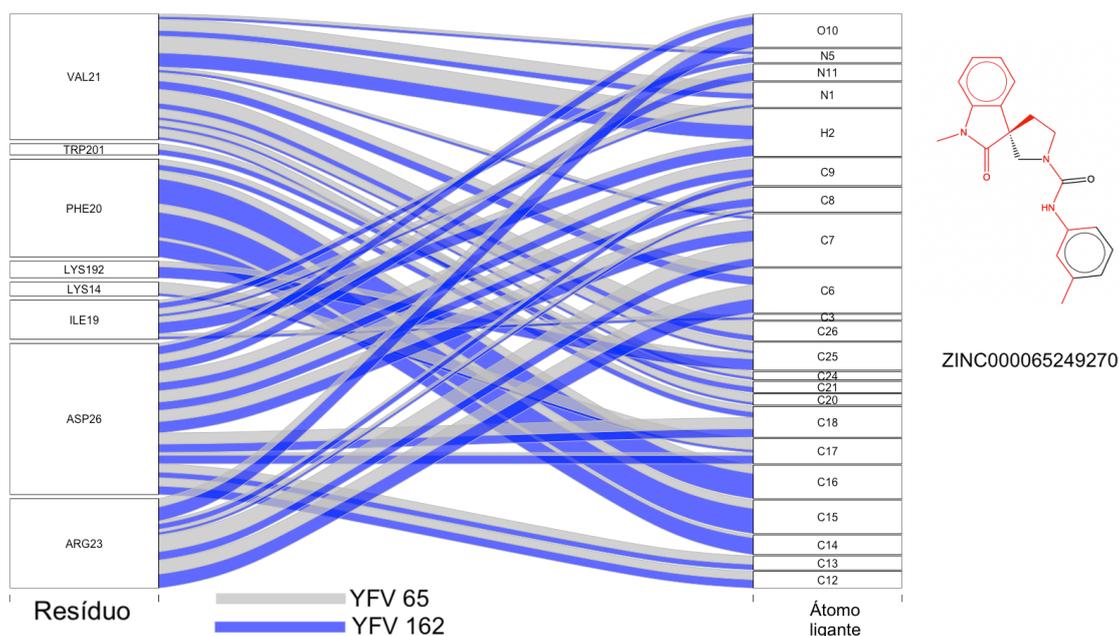


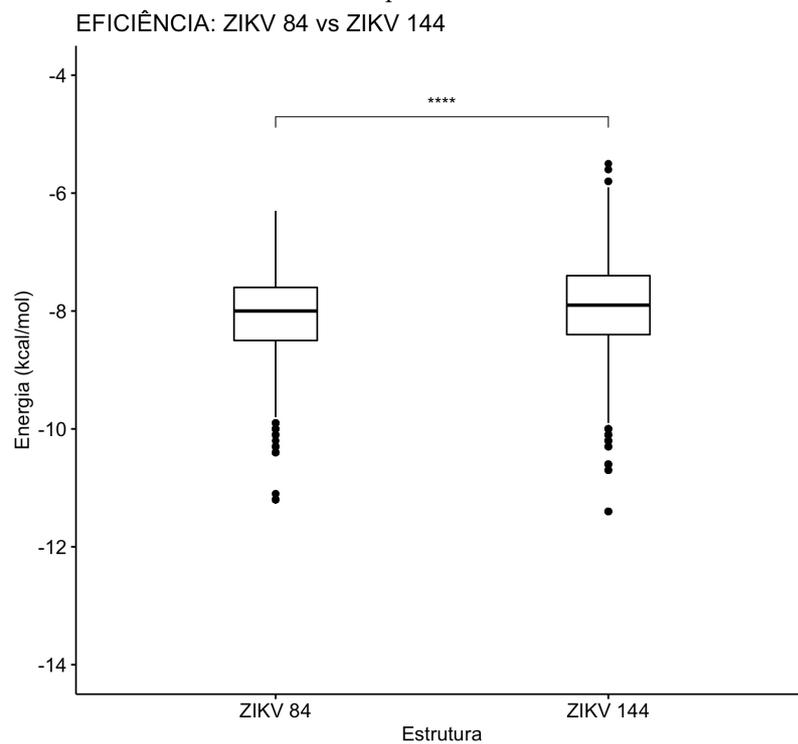
Diagrama ilustrando o local de ancoragem do fragmento ZINC000065249270 em comum nas duas estruturas da NS1. Na representação da estrutura do composto, destaca-se (em vermelho) os fragmentos em comum (coluna “Átomo ligante” do diagrama).

5.4.2.3 Comparação da triagem virtual entre estruturas da NS1 do ZIKV

Similar aos resultados encontrados para o DENV e YFV, a distribuição de energia da triagem virtual dos compostos do ZINC também apresentou diferença estatisticamente significativa ($p < 0.0001$) entre as duas estruturas da NS1: ZIKV 84 e ZIKV144 (Figura 77). Para a estrutura do ZIKV 84 a maior energia observada foi de -6.3 kcal/mol e a menor de -11.2 kcal/mol. Já para a estrutura do ZIKV 144 esses valores foram de -5.5 kcal/mol e -11.4 kcal/mol, respectivamente.

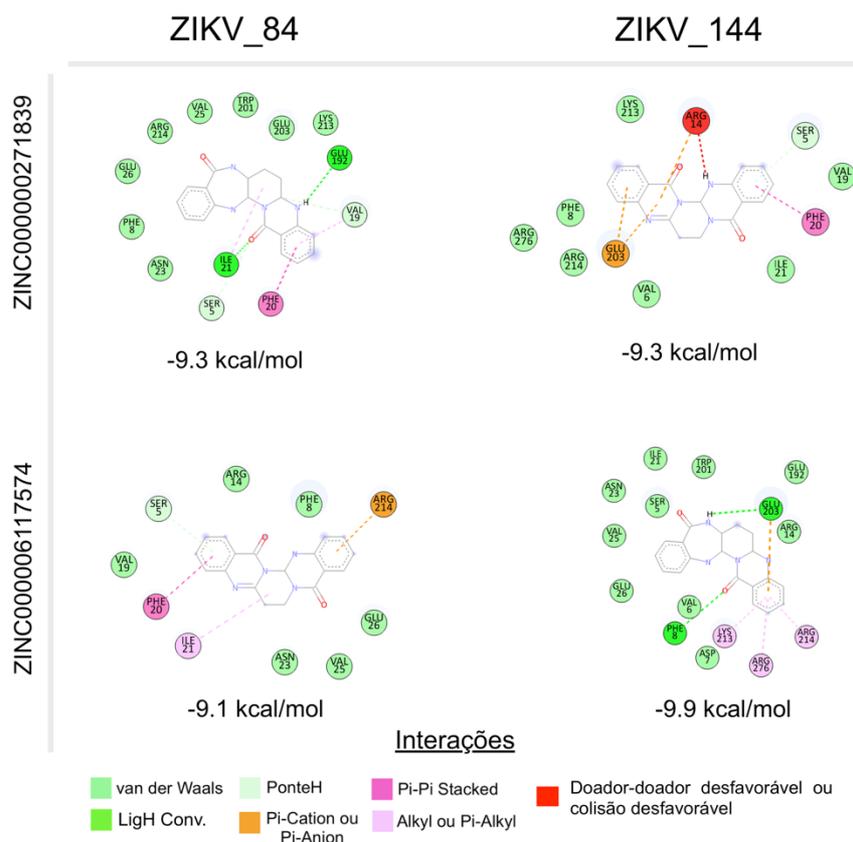
Na comparação dos 100 compostos ranqueados para cada estrutura, observou-se apenas dois compostos em comum às duas NS1 (Apêndice F). Na análise de interação 2D (Figura 78), os aminoácidos envolvidos nas interações apareceram em pelo menos três das quatro interações avaliadas, exceto pra VAL6, ASP7 e ARG276. Os resíduos SER5, ILE21 e ARG214 apareceram em todas as quatro análises das interações.

Figura 77 - *Boxplot* das distribuições de energia das 100 simulações exaustivas para os 100 compostos ranqueados de acordo com critério de eficiência para as duas estruturas da NS1 do ZIKV



**** $p < 0.0001$

Figura 78 - Representação do diagrama 2D de interação dos dois compostos em comum às duas estruturas do ZIKV com seus respectivos escores de energia



Composto	Estrutura	Torções ativas	# Átomos	Energia (kcal/mol)	SER5	VAL6	ASP7	PHE8	ARG14	VAL19	PHE20	ILE21	ASN23	VAL25	GLU26	GLU192	TRP201	GLU203	LYS213	ARG214	ARG276	Total Interações
ZINC000000271839	ZIKV_84	0	25	-9.3																		10
	ZIKV_144	0	25	-9.3																		
ZINC0000006117574	ZIKV_84	0	28	-9.1																		10
	ZIKV_144	0	28	-9.9																		15

Na primeira coluna está representado as interações do ZIKV 84 e na segunda coluna as inerações do ZIKV 144, ambos em superfície rosa claro e rosa mais escuro, respectivamente. Na orientação das linhas (horizontal, de cima para baixo) estão os compostos ZINC000000271839 e ZINC0000006117574 em *sticks*

Na busca de fragmentos próximos aos mesmos aminoácidos nas duas estruturas da NS1, foi possível observar para o ligante ZINC000000271839 que uma região do composto manteve distância $< 5 \text{ \AA}$ aos resíduos SER5, VAL6 e ARG14 independente da estrutura (Figura 79). A maior parte da molécula estava próxima a SER5, indicando uma maior preferência de proximidade dessa região do ligante com esse resíduo. Esse aminoácido também se destacou na análise de interação 2D realizada anteriormente (Figura 78), onde apareceu nas quatro interações analisadas.

Figura 79 - Análise fragmento-específico do ligante ZINC000000271839 no sítio da NS1 do ZIKV

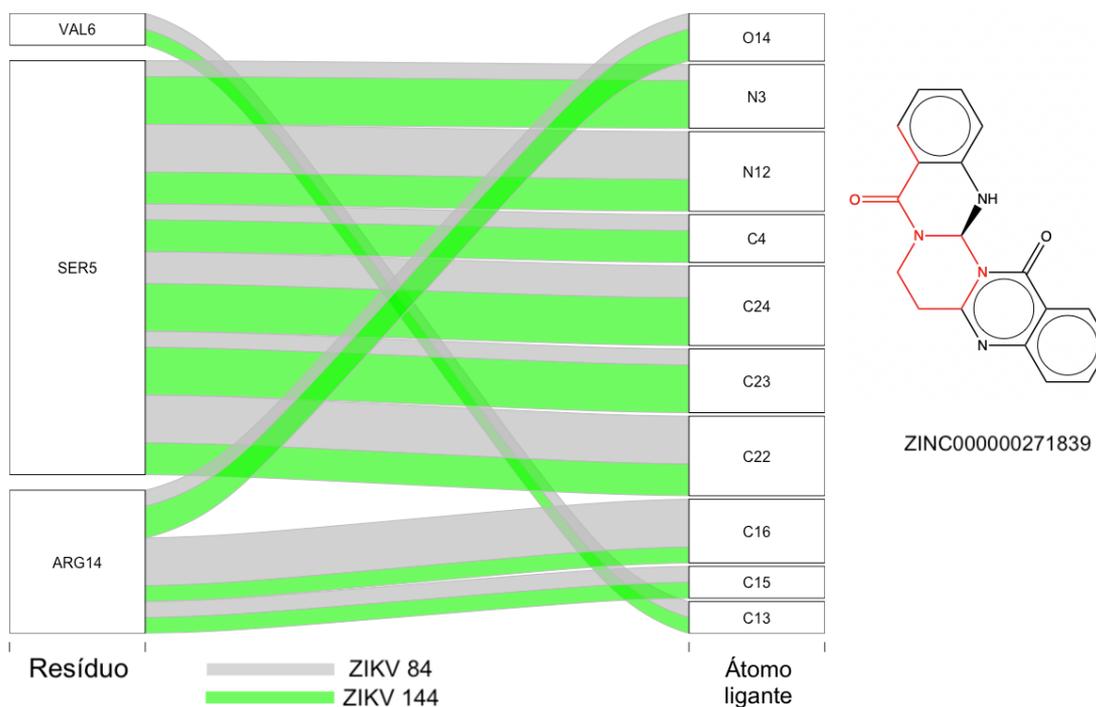


Diagrama ilustrando o local de ancoragem do fragmento ZINC000000271839 em comum nas duas estruturas da NS1. Na representação da estrutura do composto, destaca-se (em vermelho) os fragmentos em comum (coluna “Átomo ligante” do diagrama).

Na análise para o ligante ZINC000006117574, grande parte da molécula estava a uma distância $< 5 \text{ \AA}$ dos mesmos resíduos da proteína NS1 do ZIKV independente da estrutura analisada (ZIKV 84 e ZIKV 144) (Figura 80). Diferente do observado para o ligante ZINC000000271839, vários resíduos se destacaram na análise: SER5, VAL6, PHE8, ASN23, VAL25, GLU26, GLU192, TRP201, GLU203 e LYS213. Esse resultado sugere uma maior concordância entre local de interação da molécula nas estruturas da proteína NS1 do ZIKV.

Figura 80 - Análise fragmento-específico do ligante ZINC000006117574 no sítio da NS1 do ZIKV

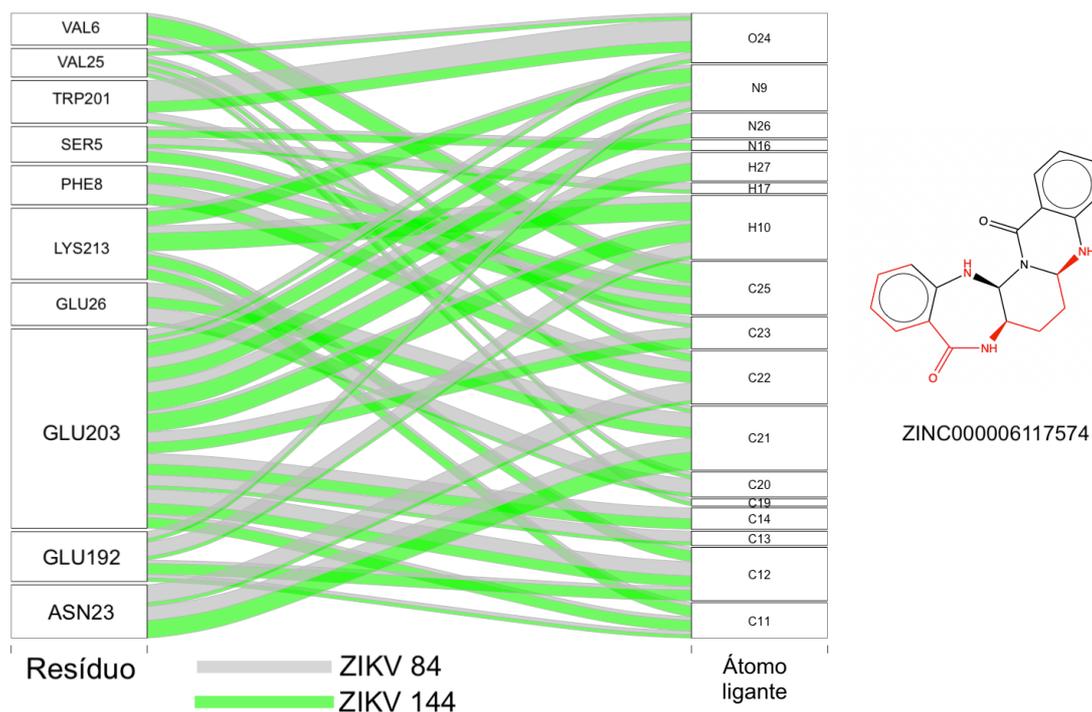


Diagrama ilustrando o local de ancoragem do fragmento ZINC000006117574 em comum nas duas estruturas da NS1. Na representação da estrutura do composto, destaca-se (em vermelho) os fragmentos em comum (coluna “Átomo ligante” do diagrama).

5.4.2.4 Análise da intersecção de compostos presentes na triagem virtual das três espécies

Da mesma forma que foi feita a busca de compostos em comum nas estruturas da NS1 das três espécies nos compostos classificados pelo critério de afinidade, também foi realizada a busca para os compostos classificados por critério de eficiência. O resultado pode ser observado na Tabela 5.

Observou-se intersecção de compostos na análise de quatro estruturas, como pelo menos uma de cada espécie e na análise com três estruturas, sendo uma de cada espécie com um total de quatro análises com pelo menos um composto em comum. Ao avaliar quais compostos eram esses, notou-se que eles se resumiram em dois compostos, como descrito na Tabela 6.

No *boxplot* da distribuição das 250 menores energias resultado da ancoragem desses dois compostos nas estruturas da NS1 (Figura 81), é possível observar que o composto ZINC000000271839 possui valor maior de energia (-7.6 kcal/mol, para estrutura da NS1 do ZIKV 144) e o composto ZINC000000485867 possui menor energia (-10.1 kcal/mol, na estrutura da NS1 YFV 162). Contudo, a distribuição global das

energias é bastante similar para os dois compostos. Ao avaliar essa distribuição globalmente, sem separar por estrutura de NS1, nota-se que não há diferença estatisticamente significante entre elas ($p > 0.05$) (Figura 82). Apesar disso, esses compostos possuem baixa similaridade estrutural, com coeficiente de Tanimoto de 0.22, de acordo com a ferramenta “*Similarity Workbench*” do ChemMine Tools.

Tabela 5 – Interseção de compostos por critério de eficiência presentes nas estruturas da NS1

ANÁLISE COM 6 ESTRUTURAS						# COMPOSTOS EM COMUM
DENV_46	DENV_73	YFV_65	YFV_162	ZIKV_84	ZIKV_144	0
ANÁLISE COM 5 ESTRUTURAS						# COMPOSTOS EM COMUM
DENV_46	DENV_73	YFV_65	YFV_162	ZIKV_84	ZIKV_144	0
DENV_46	DENV_73	YFV_65	YFV_162	ZIKV_144	ZIKV_144	0
DENV_46	DENV_73	YFV_162	ZIKV_84	ZIKV_144	ZIKV_144	0
DENV_46	YFV_65	YFV_162	ZIKV_84	ZIKV_144	ZIKV_144	0
DENV_73	YFV_65	YFV_162	ZIKV_84	ZIKV_144	ZIKV_144	0
ANÁLISE COM 4 ESTRUTURAS						# COMPOSTOS EM COMUM
DENV_46	DENV_73	YFV_65	ZIKV_84	ZIKV_144	ZIKV_144	0
DENV_46	DENV_73	YFV_65	ZIKV_144	ZIKV_144	ZIKV_144	0
DENV_46	DENV_73	YFV_162	ZIKV_84	ZIKV_144	ZIKV_144	0
DENV_46	DENV_73	YFV_162	ZIKV_144	ZIKV_144	ZIKV_144	0
DENV_46	YFV_65	YFV_162	ZIKV_84	ZIKV_144	ZIKV_144	0
DENV_73	YFV_65	YFV_162	ZIKV_84	ZIKV_144	ZIKV_144	1
DENV_73	YFV_65	YFV_162	ZIKV_144	ZIKV_144	ZIKV_144	0
DENV_46	YFV_65	YFV_162	ZIKV_144	ZIKV_144	ZIKV_144	0
DENV_46	YFV_65	ZIKV_84	ZIKV_144	ZIKV_144	ZIKV_144	0
DENV_73	YFV_65	ZIKV_84	ZIKV_144	ZIKV_144	ZIKV_144	0
DENV_46	YFV_162	ZIKV_84	ZIKV_144	ZIKV_144	ZIKV_144	0
DENV_73	YFV_162	ZIKV_84	ZIKV_144	ZIKV_144	ZIKV_144	0
ANÁLISE COM 3 ESTRUTURAS						# COMPOSTOS EM COMUM
DENV_46	ZIKV_84	YFV_65	ZIKV_144	ZIKV_144	ZIKV_144	0
DENV_46	ZIKV_84	YFV_162	ZIKV_144	ZIKV_144	ZIKV_144	0
DENV_46	ZIKV_144	YFV_65	ZIKV_144	ZIKV_144	ZIKV_144	0
DENV_46	ZIKV_144	YFV_162	ZIKV_144	ZIKV_144	ZIKV_144	0
DENV_73	ZIKV_84	YFV_65	ZIKV_144	ZIKV_144	ZIKV_144	2
DENV_73	ZIKV_84	YFV_162	ZIKV_144	ZIKV_144	ZIKV_144	1
DENV_73	ZIKV_144	YFV_65	ZIKV_144	ZIKV_144	ZIKV_144	1
DENV_73	ZIKV_144	YFV_162	ZIKV_144	ZIKV_144	ZIKV_144	0

Tabela 6 – Compostos por critério de eficiência presentes na interseção entre estruturas da NS1

Composto	Análises
ZINC000000271839	DENV_73 ∩ YFV_65 ∩ ZIKV_84 DENV_73 ∩ YFV_65 ∩ ZIKV_144
ZINC000000485867	DENV_73 ∩ YFV_65 ∩ YFV_162 ∩ ZIKV_84 DENV_73 ∩ YFV_65 ∩ ZIKV_84 DENV_73 ∩ YFV_162 ∩ ZIKV_84

Figura 81 – *Boxplot* da distribuição das energias da triagem virtual obtida para os dois compostos em comum

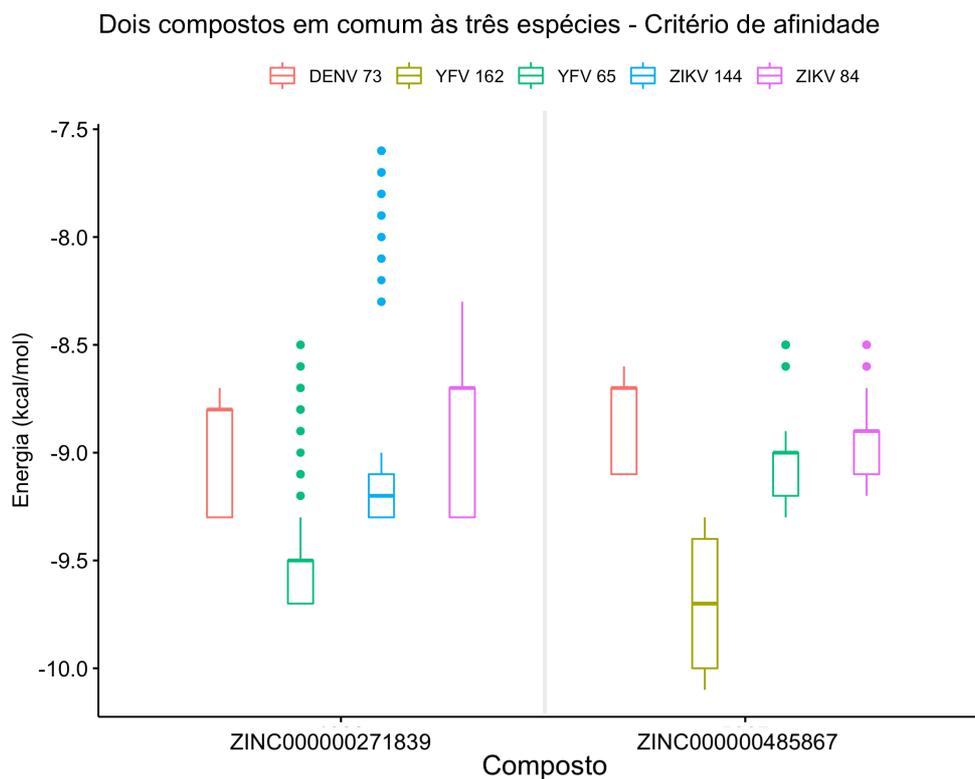
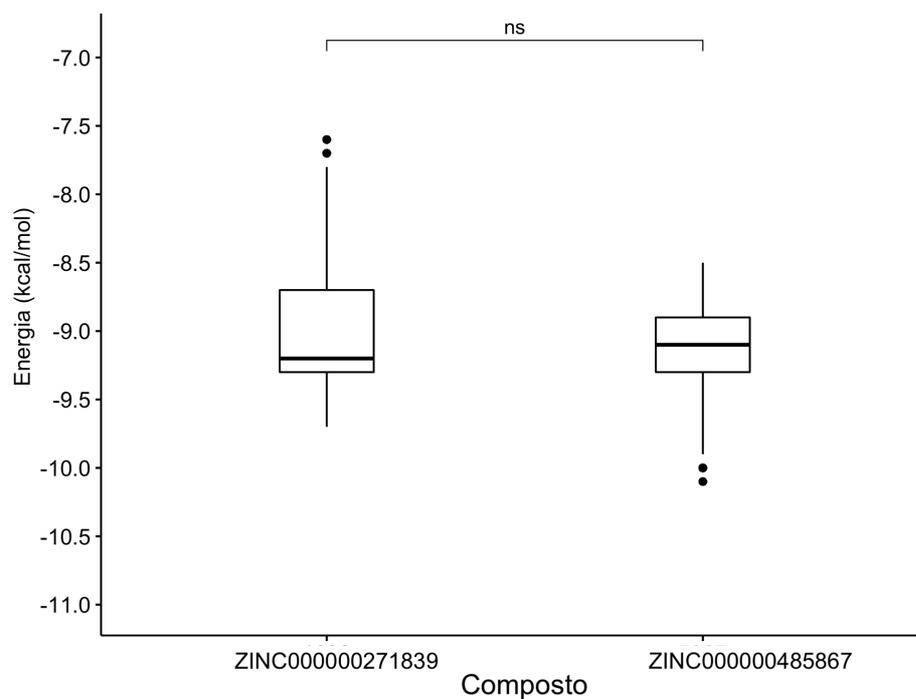


Figura 82 – *Boxplot* da distribuição global das energias dos dois compostos em comum



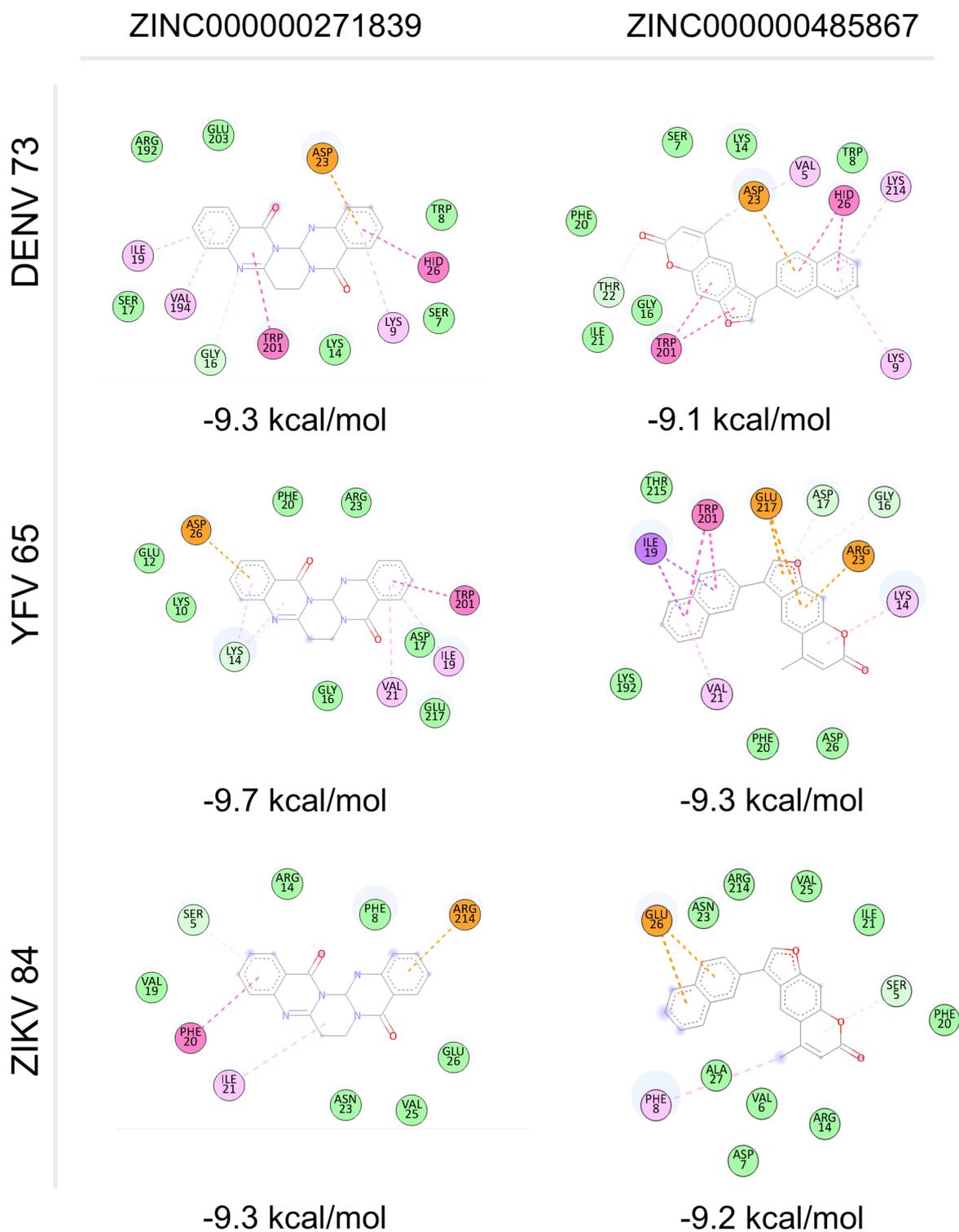
A análise das interações entre o ligante ZINC000000271839 e proteína revelou que foram mantidas as interações nos resíduos LYS/ARG14 e ILE/VAL19 em todas as estruturas da NS1 da análise (DENV 73, YFV 65, ZIKV 84 e ZIKV 144) (Figura 83). Esses resíduos na mesma posição (14 e 19), são classificados como substituição conservativa, uma vez que os aminoácidos possuem características bioquímicas similares como falado anteriormente em outras análises. Em relação ao ligante ZINC000000485867, além da interação LYS/ARG14, também se observou as interações ILE/VAL21, ARG/ASP/ASN23 e HID/ASP/GLU26. Essas cinco interações conservadas para seus respectivos compostos independente da estrutura da NS1 também foram encontradas nas interações dos cinco compostos em comum classificados de acordo com o critério de afinidade. Esses resultados sugerem a importância desses resíduos, não somente em relação à sua posição, mas também às suas características bioquímicas.

Ao unir os resultados obtidos para os dois compostos nota-se que apenas a interação LYS/ARG14 foi mantida para todas as estruturas. Entretanto, VAL/ILE21, ARG/ASP/ASN23 e HID/ASP/GLU26 também se destacaram, aparecendo em sete das oito análises ligante-receptor. Também é importante destacar a presença das interações PHE20 (seis de oito interações), TRP201 e ILE/VAL19 (cinco de oito interações, cada um).

A análise do perfil ADMET *in silico* realizado no programa OSIRIS apresentou bons resultados para o composto ZINC000000271839, com baixo risco de toxicidade para todos os parâmetros analisados e *drug-score* de 0.87 (Tabela 6). Já o ligante ZINC000000485867 apresentou elevado risco de toxicidade na avaliação de risco reprodutivo e *drug-score* de 0.12.

O fragmento apontado como de elevado risco toxicológico pelo programa OSIRIS engloba boa parte da molécula, cuja remoção resulta em um composto com baixo risco toxicológico, porém com baixos valores de *druglikeness* e *drug-score* (-4.24 e 0.39) (Figura 84). Esse comportamento foi diferente do observado nas moléculas classificadas pelo critério de afinidade, uma vez que lá as modificações nas moléculas aumentaram os escores do OSIRES para valores de potenciais fármacos (*druglikeness* > 0) e aqui esse valor se manteve negativo.

Figura 83 – Diagrama 2D da interação entre proteína NS1 e os dois ligantes em comum (ZINC000000271839 e ZINC000000485867)



6. DISCUSSÃO

A proteína NS1 dos *Flavivirus* tem se mostrado um importante fator de virulência, e conseqüentemente um promissor alvo para uso terapêutico contra as viroses causadas por esse gênero. Diversos estudos vêm sendo desenvolvidos almejando a inibição da proteína NS1, seja através de moléculas, peptídeos ou anticorpos Anti-NS1. Entretanto, todos eles focam em apenas uma espécie e na forma dimérica da proteína (AHMAD et al., 2020; RAZA; ABBAS; AZAM, 2019; SONGPRAKHON et al., 2020; WAN et al., 2014).

O presente estudo, contudo, visou a busca de conformações do monômero da NS1 que sejam similares entre as três espécies de *Flavivirus* mais presentes no Brasil na tentativa de propor uma molécula que seja em comum à todas elas. A sugestão pela forma monomérica deve-se ao fato da proteína como dímero possuir importantes funções na patogênese do vírus. Sendo assim, acredita-se que uma interferência na molécula não funcional (monômero), impedindo a formação das estruturas funcionais (dímero e hexâmero) seja uma melhor alternativa.

O estudo realizado em réplicas permitiu o aumento da amostragem conformacional da proteína NS1 das três espécies. A importância dessa metodologia pôde ser observada pela diferença do perfil do RMSD nas análises interespecie, e na quantidade de *clusters* gerados após concatenar as trajetórias.

Um desafio que encontra-se na execução da triagem virtual de compostos é a seleção das conformações do receptor para ancoragem com os ligantes. Muitos pesquisadores têm usado a abordagem conhecida como *ensemble docking* quando a dinâmica de um receptor é caracterizada e um série de conformações são usadas na ancoragem (EVANGELISTA FALCON et al., 2019). Apesar de obter bons resultados, essa metodologia pode ter um elevado custo computacional. Em casos de capacidade computacional limitada, o agrupamento das conformações obtidas de uma única trajetória, e o uso das estruturas representativas de cada grupo, é usada para selecionar as conformações do receptor para a ancoragem (AMARO et al., 2018).

Evangelista et al. (2019) enfatiza que o agrupamento de configurações de uma trajetória não identifica as estruturas proteicas que são mais selecionadas pelos seus ligantes, entretanto, se o agrupamento focar em apenas uma região (um sítio de ligação)

os resultados podem ser aprimorados quando comparados aos agrupamentos focados em grande parte da proteína.

Com base nisso, foi realizada a análise de *cluster* (agrupamento de conformações) apenas da cavidade de ligação dos possíveis ligantes. Como o intuito do trabalho é encontrar moléculas que possuam afinidade elevada à NS1 das três espécies virais concomitantemente, foram selecionadas conformações de cada espécie que se mostraram estruturalmente similares entre si. Àquelas com o sítio de ligação aberto (favorável para a ancoragem do ligante) e com melhores *drug scores* foram selecionados para a execução da triagem virtual. Nota-se no alinhamento das estruturas (Figura 30) que a região com maior similaridade estrutural é a do β -*ladder*, sendo a região do β -*roll* a mais divergente devido à sua flutuação elevada.

Dessa forma, após a seleção das conformações estruturais da NS1 de cada espécie realizou-se a triagem virtual de compostos naturais do banco de dados de ZINC. As análises das ancoragens foram ranqueadas de acordo com dois critérios: afinidade e eficiência. Na análise de critério de afinidade, cinco compostos foram encontrados em comum às estruturas da NS1 sendo pelo menos uma de cada espécie, sendo que quatro ligantes apresentaram risco de toxicidade médio ou elevado de acordo com o programa OSIRIS.

Esses compostos foram então modificados para remoção dos riscos toxicológicos, e uma nova análise de ancoragem foi realizada com eles. Todos tiveram os valores de energia de acordo com o escore do AutoDock Vina aumentados, sendo que ainda assim se mantiveram em uma boa pontuação para se tornar um possível fármaco, incluindo o valor de *druglikeness* do OSIRIS maior que 0. A análise de interação dos ligantes pós-modificação estrutural e do ligante ZINC000013411589 (o único que não apresentou risco toxicológico e não precisou ser modificado) apresentou resíduos importantes (LYS/ARG14, ILE/VAL19, PHE20, ILE/VAL21 e TRP201) que se destacaram nas análises.

Na análise dos compostos classificados por critério de eficiência, apenas dois ligantes foram encontrados em comum às estruturas da NS1 das diferentes espécies. A análise do OSIRIS apontou um composto (ZINC000000485867) com elevado risco reprodutivo, sendo que diferente dos compostos classificados pelo critério de afinidade,

a modificação do fragmento apontado como tóxico não gerou uma molécula promissora para se tornar um potencial fármaco (*druglikeness* < 0). Sendo assim, apenas o composto ZINC000000271839 se mostrou promissor na classificação por critério de eficiência. Na análise de interação receptor-ligante, as interações LYS/ARG14 e ILE/VAL19 se destacaram mais uma vez. No geral, seis potenciais fármacos multialvo foram selecionados. Um resumo deles pode ser observado na Tabela 7.

Tabela 7 – Seis compostos selecionados como promissores para inibição da formação do dímero da NS1

Composto	Critério ^{&}	<i>Druglikeness</i>	<i>Drug-score</i>	Espécie	Energia [†] (kcal/mol)
ZINC000008792352*	A	8.87	0.77	DENV	-8.7
				YFV	-9.2
				ZIKV	-8.7
ZINC000012879754*	A	8.79	0.70	DENV	-9.2
				YFV	-9.6
				ZIKV	-8.5
ZINC000002108995*	A	2.11	0.78	DENV	-8.5
				YFV	-9.0
				ZIKV	-7.5
ZINC000013411589	A	3.78	0.24	DENV	-10.8
				YFV	-11.8
				ZIKV	-11.7
ZINC000008792389*	A	5.48	0.82	DENV	-8.3
				YFV	-8.9
				ZIKV	-8.9
ZINC000000271839	E	4.21	0.87	DENV	-9.3
				YFV	-9.7
				ZIKV	-9.3

*: compostos com modificação estrutural

†: valor do menor escore de energia obtido pelo AutoDock Vina

&: A = afinidade; E = eficiência

7. CONCLUSÃO E PERSPECTIVAS FUTURAS

As ferramentas *in silico* vêm sendo utilizadas como um importante aliado na busca de moléculas promissoras à fármacos permitindo uma diminuição de custo e tempo no processo. Para tanto, novas estratégias e aprimoramento das já existentes estão sendo testadas.

Com base nisso, o presente trabalho obteve a partir de simulações de dinâmica molecular em quintuplicatas da proteína NS1 do DENV, YFV e ZIKV conformações similares entre as espécies com o intuito de encontrar ligantes com elevada afinidade pelo sítio de ancoragem.

A triagem virtual de compostos naturais do banco de dados do ZINC forneceu seis fármacos promissores, sendo cinco classificados pelo critério de afinidade e um pelo critério de eficiência. As análises das interações desses compostos permitiram destacar resíduos chaves para a interação deles.

O presente trabalho foi pioneiro no laboratório de bioinformática do NCBios a trabalhar com diferentes conformações de uma proteína de espécies distintas na busca de um composto multialvo, iniciando uma nova abordagem no desenvolvimento racional de fármacos.

A ideia de que um mesmo composto tem tendência a interagir com conformações similares com elevada afinidade foi mostrado aqui com os compostos em comum às estruturas da NS1 similares entre si. A limitação do trabalho se deu pelo baixo número de estruturas usadas (duas de cada espécie), devido à reduzida capacidade computacional do laboratório. Mais compostos poderiam ser ranqueados com o aumento de conformações da NS1 para cada espécie. Entretanto, os seis compostos aqui apontados são suficientes para dar prosseguimento aos estudos.

Apesar de não ter sido foco do trabalho, também foi realizada a comparação de compostos ranqueados em comum nas estruturas de NS1 da mesma espécie. Os dados gerados e resultados aqui apresentados abrem oportunidades de estudo no desenho racional de fármacos espécie-específico desses *Flavivirus*.

Os resultados até então obtidos se mostraram como um importante passo inicial, sendo necessário ainda análises complementares. Almejamos prosseguir com o estudo de estabilidade dessas moléculas nas estruturas da NS1 dessas três espécies virais por meio de simulações de dinâmica molecular. Espera-se que haja uma estabilidade do sítio de ancoragem (β -hairpin) que é uma região instável sem a interação com a mesma região da outra unidade monomérica. Com a estabilidade garantida pelo ligante, a interface de interação para formação do dímero fica indisponível. Por fim, testes *in vitro* serão realizados para avaliação da capacidade inibitória da reprodução dos vírus por esses compostos.

REFERÊNCIAS

- ABAD-ZAPATERO, C. Ligand efficiency indices for effective drug discovery. **Expert Opinion on Drug Discovery**, v. 2, n. 4, p. 469–488, abr. 2007.
- ABRAHAM, M. J. et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. **SoftwareX**, v. 1–2, p. 19–25, set. 2015.
- ABREU, F. V. S. DE et al. *Haemagogus leucocelaenus* and *Haemagogus janthinomys* are the primary vectors in the major yellow fever outbreak in Brazil, 2016–2018. **Emerging Microbes & Infections**, v. 8, n. 1, p. 218–231, jan. 2019.
- AHMAD, N. et al. Structural insights into the Zika virus NS1 protein inhibition using a computational approach. **Journal of Biomolecular Structure and Dynamics**, p. 1–8, 6 maio 2020.
- AKEY, D. L. et al. Flavivirus NS1 structures reveal surfaces for associations with membranes and the immune system. **Science (New York, N.Y.)**, v. 343, n. 6173, p. 881–885, 21 fev. 2014.
- AKEY, D. L. et al. Structure-guided insights on the role of NS1 in flavivirus infection. **BioEssays**, v. 37, n. 5, p. 489–494, 1 maio 2015.
- AMARO, R. E. et al. Ensemble Docking in Drug Discovery. **Biophysical Journal**, v. 114, n. 10, p. 2271–2278, maio 2018.
- AVIRUTNAN, P. et al. Binding of Flavivirus Nonstructural Protein NS1 to C4b Binding Protein Modulates Complement Activation. **The Journal of Immunology**, v. 187, n. 1, p. 424–433, jul. 2011.
- BARJAS-CASTRO, M. L. et al. Probable transfusion-transmitted Zika virus in Brazil: Transfusion-Transmitted Zika Virus. **Transfusion**, v. 56, n. 7, p. 1684–1688, jul. 2016.
- BARNETT, E. D. Yellow Fever: Epidemiology and Prevention. **Clinical Infectious Diseases**, v. 44, n. 6, p. 850–856, 15 mar. 2007.
- BARREIRO, E. J. et al. Molecular modeling: a tool for rational drug design in medicinal chemistry. **Química Nova**, v. 20, n. 3, p. 300–310, jun. 1997.
- BEATTY, P. R. et al. Dengue virus NS1 triggers endothelial permeability and vascular leak that is prevented by NS1 vaccination. **Science Translational Medicine**, v. 7, n. 304, p. 304ra141-304ra141, 9 set. 2015.
- BERENDSEN, H. J. C. et al. Molecular dynamics with coupling to an external bath. **The Journal of Chemical Physics**, v. 81, n. 8, p. 3684–3690, 15 out. 1984.
- BERMAN, H. M. et al. The Protein Data Bank. **Nucleic Acids Research**, v. 28, n. 1, p. 235–242, 1 jan. 2000.

- BESNARD, M. et al. Evidence of perinatal transmission of Zika virus, French Polynesia, December 2013 and February 2014. **Eurosurveillance**, v. 19, n. 13, p. 20751, 3 abr. 2014.
- BHATT, S. et al. The global distribution and burden of dengue. **Nature**, v. 496, n. 7446, p. 504–507, abr. 2013.
- BLITVICH, B. J. et al. Determination of the intramolecular disulfide bond arrangement and biochemical identification of the glycosylation sites of the nonstructural protein NS1 of Murray Valley encephalitis virus. **The Journal of General Virology**, v. 82, n. 9, p. 2251–2256, set. 2001.
- BRADY, O. J. et al. The association between Zika virus infection and microcephaly in Brazil 2015–2017: An observational analysis of over 4 million births. **PLOS Medicine**, v. 16, n. 3, p. e1002755, 5 mar. 2019.
- BRASIL. Ministério da Saúde. **Monitoramento do Período Sazonal da Febre Amarela Brasil – 2017/2018**. Brasília: Ministério da Saúde, 2018. (Informes técnicos. Ministério da Saúde).
- BROWN, W. C. et al. Extended surface for membrane association in Zika virus NS1 structure. **Nature Structural & Molecular Biology**, v. 23, n. 9, p. 865–867, set. 2016.
- CAMPOS, G. S.; BANDEIRA, A. C.; SARDI, S. I. Zika Virus Outbreak, Bahia, Brazil. **Emerging Infectious Diseases**, v. 21, n. 10, p. 1885–1886, out. 2015.
- CARTER, J.; SAUNDERS, V. **Virology: Principles and Applications**. [s.l.] John Wiley & Sons, 2007.
- CASE, D. A. et al. Amber 14. 2014.
- CENTRO EUROPEU DE CONTROLE E PREVENÇÃO DE DOENÇAS. **Geographical distribution o dengue cases reported worldwide, 2020**. Disponível em: <https://www.ecdc.europa.eu/en/publications-data/geographical-distribution-dengue-cases-reported-worldwide-2020>. Acesso: 01 de abril de 2021.
- CHAMBERS, T. J. et al. Flavivirus Genome Organization, Expression, and Replication. **Annual Review of Microbiology**, v. 44, n. 1, p. 649–688, 1990.
- CHEN, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. **Acta Crystallographica Section D Biological Crystallography**, v. 66, n. 1, p. 12–21, 1 jan. 2010.
- CHIKHALE, R. V. et al. Identification of potential anti-TMPRSS2 natural products through homology modelling, virtual screening and molecular dynamics simulation studies. **Journal of Biomolecular Structure and Dynamics**, p. 1–16, 3 ago. 2020.
- CHOTHIA, C.; LESK, A. M. The relation between the divergence of sequence and structure in proteins. **The EMBO Journal**, v. 5, n. 4, p. 823–826, 1 abr. 1986.
- CI, Y. et al. Zika NS1-induced ER remodeling is essential for viral replication. **The Journal of Cell Biology**, v. 219, n. 2, p. e201903062, 3 fev. 2020.

COSTA, F. G. et al. Alkaloids as Inhibitors of Malate Synthase from *Paracoccidioides* spp.: Receptor-Ligand Interaction-Based Virtual Screening and Molecular Docking Studies, Antifungal Activity, and the Adhesion Process. **Antimicrobial Agents and Chemotherapy**, v. 59, n. 9, p. 5581–5594, set. 2015.

DA SILVA, L. S. et al. Identification of a new antifungal compound against isocitrate lyase of *Paracoccidioides brasiliensis*. **Future Microbiology**, v. 14, p. 1589–1606, dez. 2019.

DAURA, X. et al. Peptide Folding: When Simulation Meets Experiment. p. 5, 1999.

DENG, S.-Q. et al. A Review on Dengue Vaccine Development. **Vaccines**, v. 8, n. 1, p. 63, 2 fev. 2020.

DIAGNE, C. T. et al. Potential of selected Senegalese *Aedes* spp. mosquitoes (Diptera: Culicidae) to transmit Zika virus. **BMC Infectious Diseases**, v. 15, n. 1, p. 492, dez. 2015.

DIALLO, D. et al. Zika Virus Emergence in Mosquitoes in Southeastern Senegal, 2011. **PLoS ONE**, v. 9, n. 10, p. e109442, 13 out. 2014.

DICK, G. W. A.; KITCHEN, S. F.; HADDOW, A. J. Zika Virus (I). Isolations and serological specificity. **Transactions of the Royal Society of Tropical Medicine and Hygiene**, v. 46, n. 5, p. 509–520, set. 1952.

DIVYASHRI, G. et al. In silico approach towards the identification of potential inhibitors from *Curcuma amada* Roxb against *H. pylori*: ADMET screening and molecular docking studies. **BioImpacts : BI**, v. 11, n. 2, p. 119–127, 2021.

ESSMANN, U. et al. A smooth particle mesh Ewald method. **The Journal of Chemical Physics**, v. 103, n. 19, p. 8577–8593, 15 nov. 1995.

ESWAR, N. et al. Comparative Protein Structure Modeling Using Modeller. **Current Protocols in Bioinformatics**, v. 15, n. 1, p. 5.6.1-5.6.30, 2014.

EVANGELISTA FALCON, W. et al. Ensemble Docking in Drug Discovery: How Many Protein Configurations from Molecular Dynamics Simulations are Needed To Reproduce Known Ligand Binding? **The Journal of Physical Chemistry B**, v. 123, n. 25, p. 5189–5195, 27 jun. 2019.

FARIA, N. R. et al. Zika virus in the Americas: Early epidemiological and genetic findings. **Science**, v. 352, n. 6283, p. 345–349, 15 abr. 2016.

FERREIRA, L. et al. Molecular Docking and Structure-Based Drug Design Strategies. **Molecules**, v. 20, n. 7, p. 13384–13421, 22 jul. 2015.

FERREIRA, L. T. et al. QSAR-Based Virtual Screening of Natural Products Database for Identification of Potent Antimalarial Hits. **Biomolecules**, v. 11, n. 3, p. 459, 19 mar. 2021.

FIROZ, A. et al. ContPro: A web tool for calculating amino acid contact distances in protein from 3D -structure at different distance threshold. **open access**, v. 5, n. 2, p. 3, 2010.

- FOY, B. D. et al. Probable Non-Vector-borne Transmission of Zika Virus, Colorado, USA. **Emerging Infectious Diseases**, v. 17, n. 5, p. 880–882, maio 2011.
- GIOVANETTI, M. et al. Yellow Fever Virus Reemergence and Spread in Southeast Brazil, 2016–2019. **Journal of Virology**, v. 94, n. 1, p. e01623-19, /jvi/94/1/JVI.01623-19.atom, 9 out. 2019.
- GLASNER, D. R. et al. The Good, the Bad, and the Shocking: The Multiple Roles of Dengue Virus Nonstructural Protein 1 in Protection and Pathogenesis. **Annual Review of Virology**, v. 5, n. 1, p. 227–253, 29 2018.
- GONÇALVES, R. L. et al. Dynamic Behavior of Dengue and Zika viruses NS1 protein reveals monomer-monomer interaction mechanisms and insights to rational drug design. **Journal of Biomolecular Structure and Dynamics**, p. 1–13, 8 out. 2019.
- GUBLER, D. J. Dengue, Urbanization and Globalization: The Unholy Trinity of the 21st Century. **Tropical Medicine and Health**, v. 39, n. 4SUPPLEMENT, p. S3–S11, 2011.
- GUERBOIS, M. et al. Outbreak of Zika Virus Infection, Chiapas State, Mexico, 2015, and First Confirmed Transmission by *Aedes aegypti* Mosquitoes in the Americas. **Journal of Infectious Diseases**, v. 214, n. 9, p. 1349–1356, 1 nov. 2016.
- GUZMAN, M. G.; HARRIS, E. Dengue. **The Lancet**, v. 385, n. 9966, p. 453–465, jan. 2015.
- HAMER, D. H.; CHEN, L. H. Zika in Angola and India. **Journal of Travel Medicine**, v. 26, n. 5, p. taz012, 11 jun. 2019.
- HANLEY, K. A. et al. Fever versus fever: The role of host and vector susceptibility and interspecific competition in shaping the current and future distributions of the sylvatic cycles of dengue virus and yellow fever virus. **Infection, Genetics and Evolution**, v. 19, p. 292–311, out. 2013.
- HESS, B. P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. **Journal of Chemical Theory and Computation**, v. 4, n. 1, p. 116–122, jan. 2008.
- HILGENFELD, R. Zika virus NS1, a pathogenicity factor with many faces. **The EMBO Journal**, p. e201695871, 27 out. 2016.
- HILLISCH, A.; PINEDA, L. F.; HILGENFELD, R. Utility of homology models in the drug discovery process. **Drug Discovery Today**, v. 9, n. 15, p. 659–669, ago. 2004.
- HOCKNEY, R. W.; GOEL, S. P.; EASTWOOD, J. W. Quiet high-resolution computer models of a plasma. **Journal of Computational Physics**, v. 14, n. 2, p. 148–158, fev. 1974.
- HONORIO, K. M.; MODA, T. L.; ANDRICOPULO, A. D. Pharmacokinetic properties and in silico ADME modeling in drug discovery. **Medicinal Chemistry**, v. 9, n. 2, p. 163–176, 2013.
- HUTTER, J. Car-Parrinello molecular dynamics. **Wiley Interdisciplinary Reviews: Computational Molecular Science**, v. 2, n. 4, p. 604–612, 2012.

IRWIN, J. J. et al. ZINC: A Free Tool to Discover Chemistry for Biology. **Journal of Chemical Information and Modeling**, v. 52, n. 7, p. 1757–1768, 23 jul. 2012.

IRWIN, J. J.; SHOICHET, B. K. ZINC - A Free Database of Commercially Available Compounds for Virtual Screening. v. 45, n. 1, p. 177–182, 22 set. 2014.

I-TASSER *web server*. **Protein Structure & Function Predictions**. Disponível em: <https://zhanglab.dcmf.med.umich.edu/I-TASSER/>. Acesso: 25 de abril de 2021.

JIA, C.-Y. et al. A drug-likeness toolbox facilitates ADMET study in drug discovery. **Drug Discovery Today**, v. 25, n. 1, p. 248–258, 1 jan. 2020.

JORGENSEN, W. L. et al. Comparison of simple potential functions for simulating liquid water. **The Journal of Chemical Physics**, v. 79, n. 2, p. 926–935, jul. 1983.

JUNGLEN, S. et al. A New Flavivirus and a New Vector: Characterization of a Novel Flavivirus Isolated from Uranotaenia Mosquitoes from a Tropical Rain Forest. **Journal of Virology**, v. 83, n. 9, p. 4462–4468, 1 maio 2009.

KARPLUS, M.; MCCAMMON, J. A. Molecular dynamics simulations of biomolecules. **nature structural biology**, v. 9, n. 9, p. 7, 2002.

KIRSCHNER, K. N. et al. GLYCAM06: A Generalizable Biomolecular Force Field. Carbohydrates. **Journal of computational chemistry**, v. 29, n. 4, p. 622–655, mar. 2008.

KORSMAN, S. N. J. et al. **Virology: An Illustrated Colour Text**. [s.l.] Churchill Livingstone Elsevier, 2012.

KRUSKAL, W. H.; WALLIS, W. A. Use of Ranks in One-Criterion Variance Analysis. **Journal of the American Statistical Association**, v. 47, n. 260, p. 583–621, dez. 1952.

LAHLOU, M. The Success of Natural Products in Drug Discovery. **Pharmacology & Pharmacy**, v. 04, n. 03, p. 17–31, 2013.

LASKOWSKI, R. A. et al. PROCHECK: a program to check the stereochemical quality of protein structures. **Journal of Applied Crystallography**, v. 26, n. 2, p. 283–291, 1 abr. 1993.

LINDENBACH, B. D.; RICE, C. M. Molecular Biology of Flaviviruses. v. 59, p. 23–61, 2003.

LINDORFF-LARSEN, K. et al. Improved side-chain torsion potentials for the Amber ff99SB protein force field. **Proteins: Structure, Function, and Bioinformatics**, p. 1950–1958, 2010.

LIU, Z. et al. Deep learning enables discovery of highly potent anti-osteoporosis natural products. **European Journal of Medicinal Chemistry**, v. 210, p. 112982, jan. 2021.

LUTZ, A. Reminiscências da febre amarela no Estado de São Paulo. **Memórias do Instituto Oswaldo Cruz**, v. 24, n. 3, p. 127–160, 1930.

- MELO, A. S. DE O. et al. Congenital Zika Virus Infection: Beyond Neonatal Microcephaly. **JAMA Neurology**, v. 73, n. 12, p. 1407–1416, 1 dez. 2016.
- MENEZES, G. DE L. et al. Mutation of critical residues reveals insights of yellow fever virus nonstructural protein 1 (NS1) stability and its formation. **Journal of Biomolecular Structure and Dynamics**, p. 1–10, 27 nov. 2019.
- MESSINA, J. P. et al. The current and future global distribution and population at risk of dengue. **Nature Microbiology**, v. 4, n. 9, p. 1508–1515, set. 2019.
- MIYAMOTO, S.; KOLLMAN, P. A. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. **Journal of Computational Chemistry**, v. 13, n. 8, p. 952–962, out. 1992.
- MLAKAR, J. et al. Zika Virus Associated with Microcephaly. **New England Journal of Medicine**, v. 374, n. 10, p. 951–958, 10 mar. 2016.
- MODHIRAN, N. et al. Dengue virus NS1 protein activates cells via Toll-like receptor 4 and disrupts endothelial cell monolayer integrity. **Science Translational Medicine**, v. 7, n. 304, p. 304ra142-304ra142, 9 set. 2015.
- MONATH, T. P.; VASCONCELOS, P. F. C. Yellow fever. **Journal of Clinical Virology**, v. 64, p. 160–173, 1 mar. 2015.
- MORRIS, G. M. et al. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. **Journal of Computational Chemistry**, v. 30, n. 16, p. 2785–2791, dez. 2009.
- MÜLLER-KUHRT, L. Putting nature back into drug discovery. **Nature Biotechnology**, v. 21, n. 6, p. 602–602, jun. 2003.
- MUSSO, D. Zika Virus Transmission from French Polynesia to Brazil. 2015.
- MUSSO, D.; KO, A. I.; BAUD, D. Zika Virus Infection — After the Pandemic. **New England Journal of Medicine**, v. 381, n. 15, p. 1444–1457, 10 out. 2019.
- NAMBA, A. M.; SILVA, V. B. DA; SILVA, C. H. T. P. DA. Dinâmica molecular: teoria e aplicações em planejamento de fármacos. **Eclética Química**, v. 33, n. 4, 2008.
- NORMILE, D. Surprising New Dengue Virus Throws a Spanner in Disease Control Efforts. **Science**, v. 342, n. 6157, p. 415–415, 25 out. 2013.
- O'BOYLE, N. M. et al. Open Babel: An open chemical toolbox. **Journal of Cheminformatics**, v. 3, n. 1, p. 33, dez. 2011.
- OLIVEIRA, A. M. DE et al. A AMEAÇA DA FEBRE AMARELA E O RECENTE SURTO NO BRASIL. **REVISTA DE TRABALHOS ACADÊMICOS – UNIVERSO BELO HORIZONTE**, v. 1, n. 2, jul. 2017.
- OLSSON, M. H. M. et al. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pK_a Predictions. **Journal of Chemical Theory and Computation**, v. 7, n. 2, p. 525–537, 8 fev. 2011.

OPO, F. A. D. Structure based pharmacophore modeling, virtual screening, molecular docking and ADMET approaches for identification of natural anti-cancer agents targeting XIAP protein. **Scientific Reports**, p. 17, 2021.

ORGANIZAÇÃO MUNDIAL DA SAÚDE, 2020. **Dengue and severe dengue**. Disponível em: <https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue>. Acesso: 25 de setembro de 2020.

ORGANIZAÇÃO PAN-AMERICANA DA SAÚDE, 2021. **Dengue fever in The Americas**. Disponível em: <https://www.paho.org/data/index.php/en/mnu-topics/indicadores-dengue-en/dengue-nacional-en/252-dengue-pais-ano-en.html?start=1>. Acesso: 01 de abril de 2021.

OSMAN, W. et al. In silico assessment of potential leads identified from *Bauhinia rufescens* Lam. as α -glucosidase and α -amylase inhibitors. **Journal of Receptors and Signal Transduction**, v. 41, n. 2, p. 159–169, 4 mar. 2021.

PAGADALA, N. S.; SYED, K.; TUSZYNSKI, J. Software for molecular docking: a review. **Biophysical Reviews**, v. 9, n. 2, p. 91–102, abr. 2017.

PETTERSEN, E. F. et al. UCSF Chimera - a visualization system for exploratory research and analysis. **Journal of Computational Chemistry**, v. 25, n. 13, p. 1605–1612, out. 2004.

PONDER, J. W.; CASE, D. A. Force Fields for Protein Simulations. In: **Advances in Protein Chemistry**. [s.l.] Elsevier, 2003. v. 66p. 27–85.

POVEDA-CUEVAS, S. A.; BARROSO DA SILVA, F. L.; ETCHEBEST, C. How the Strain Origin of Zika Virus NS1 Protein Impacts Its Dynamics and Implications to Their Differential Virulence. **Journal of Chemical Information and Modeling**, v. 61, n. 3, p. 1516–1530, 22 mar. 2021.

PROMPETCHARA, E. et al. Dengue vaccine: Global development update. **Asian Pacific Journal of Allergy and Immunology**, 2019.

PRONK, S. et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. **Bioinformatics**, v. 29, n. 7, p. 845–854, 1 abr. 2013.

PROTEIN DATA BANK. **RCSB Protein Data Bank**. Disponível em: <http://www.rcsb.org>. Acesso: 25 de abril de 2021.

PROTEIN STRUCTURE PREDICTION CENTER. **14th Community Wide Experiment on the Critical Assessment of Techniques for Protein Structure Prediction**. Disponível em: <https://www.predictioncenter.org/casp14/index.cgi>. Acesso: 25 de abril de 2021.

PRYOR, M. J.; WRIGHT, P. J. Glycosylation Mutants of Dengue Virus NS1 Protein. **Journal of General Virology**, v. 75, n. 5, p. 1183–1187, 1994.

RASHID, M. Design, synthesis and ADMET prediction of bis-benzimidazole as anticancer agent. **Bioorganic Chemistry**, v. 96, p. 103576, 1 mar. 2020.

RASTOGI, M.; SHARMA, N.; SINGH, S. K. Flavivirus NS1: a multifaceted enigmatic viral protein. **Virology Journal**, v. 13, p. 131, 29 jul. 2016.

RAZA, S.; ABBAS, G.; AZAM, S. S. Screening pipeline for Flavivirus based Inhibitors for Zika Virus NS1. **IEEE/ACM Transactions on Computational Biology and Bioinformatics**, p. 1–1, 2019.

RODRIGUEZ, R. et al. Homology modeling, model and software evaluation: three related resources. **Bioinformatics**, v. 14, n. 6, p. 523–528, 1 jan. 1998.

ROY, A.; KUCUKURAL, A.; ZHANG, Y. I-TASSER: a unified platform for automated protein structure and function prediction. **Nature Protocols**, v. 5, n. 4, p. 725–738, abr. 2010.

SANTOS FILHO, O. A.; ALENCASTRO, R. B. DE. Modelagem de proteínas por homologia. **Química Nova**, v. 26, n. 2, p. 253–259, mar. 2003.

SARFRAZ, M. et al. N,N'-dialkyl-2-thiobarbituric acid based sulfonamides as potential SARS-CoV-2 main protease inhibitors. **Canadian Journal of Chemistry**, v. 99, n. 3, p. 330–345, 12 jan. 2021.

SCATURRO, P. et al. Dengue Virus Non-structural Protein 1 Modulates Infectious Particle Production via Interaction with the Structural Proteins. **PLOS Pathogens**, v. 11, n. 11, p. e1005277, 12 nov. 2015.

SCHWEDE, T. et al. SWISS-MODEL: an automated protein homology-modeling server. **Nucleic Acids Research**, v. 31, n. 13, p. 3381–3385, 1 jul. 2003.

SERTKAYA, A. et al. Key cost drivers of pharmaceutical clinical trials in the United States. **Clinical Trials**, v. 13, n. 2, p. 117–126, abr. 2016.

SHOICHET, B. K. et al. Lead discovery using molecular docking. **Current Opinion in Chemical Biology**, v. 6, n. 4, p. 439–446, ago. 2002.

SILVA, C. P. DA et al. Um estudo bibliográfico acerca dos surtos de Febre Amarela no Brasil. **Revista Brasileira de Meio Ambiente**, v. 2, n. 1, 5 set. 2018.

SONG, H. et al. Zika virus NS1 structure reveals diversity of electrostatic surfaces among flaviviruses. **Nature Structural & Molecular Biology**, v. 23, n. 5, p. 456–458, maio 2016.

SONGPRAKHON, P. et al. Peptides targeting dengue viral nonstructural protein 1 inhibit dengue virus production. **Scientific Reports**, v. 10, n. 1, p. 12933, dez. 2020.

SOPER, F. L. et al. YELLOW FEVER WITHOUT AËDES AEGYPTI. STUDY OF A RURAL EPIDEMIC IN THE VALLE DO CHANAAN, ESPIRITO SANTO, BRAZIL, 1932*. **American Journal of Epidemiology**, v. 18, n. 3, p. 555–587, nov. 1933.

SOTELO, J. R. et al. Persistence of Zika Virus in Breast Milk after Infection in Late Stage of Pregnancy. **Emerging Infectious Diseases**, v. 23, n. 5, p. 856–857, maio 2017.

SOUSA DA SILVA, A. W.; VRANKEN, W. F. ACPYPE - AnteChamber PYthon Parser interfacE. **BMC Research Notes**, v. 5, n. 1, p. 367, 2012.

SPASSOV, V. Z.; YAN, L. Accelrys Software Inc., Discovery Studio Modeling Environment, Release 4.0. **Journal of Biological Chemistry**, v. 81, n. 4, p. 704–714, abr. 2013.

SPECIAL PROGRAMME FOR RESEARCH AND TRAINING IN TROPICAL DISEASES; WORLD HEALTH ORGANIZATION (EDS.). **Dengue: guidelines for diagnosis, treatment, prevention, and control**. New ed ed. Geneva: TDR: World Health Organization, 2009.

STOESSER, G. et al. The EMBL Nucleotide Sequence Database. **Nucleic Acids Research**, v. 30, n. 1, p. 21–26, 1 jan. 2002.

THEILER, M.; SMITH, H. H. The Use of Yellow Fever Virus Modified by in Vitro Cultivation for Human Immunization. **Journal of Experimental Medicine**, v. 65, n. 6, p. 787–800, 1 jun. 1937.

TIAN, S. et al. The application of in silico drug-likeness predictions in pharmaceutical research. **Advanced Drug Delivery Reviews**, v. 86, p. 2–10, jun. 2015.

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

URSU, O.; OPREA, T. I. Model-Free Drug-Likeness from Fragments. **Journal of Chemical Information and Modeling**, v. 50, n. 8, p. 1387–1394, 23 ago. 2010.

VAN DER SPOEL, D. et al. GROMACS: Fast, flexible, and free. **Journal of Computational Chemistry**, v. 26, n. 16, p. 1701–1718, dez. 2005.

VASCONCELOS, P. F. DA C. Yellow fever. **Revista da Sociedade Brasileira de Medicina Tropical**, v. 36, n. 2, p. 275–293, abr. 2003.

VERLI, H. **Bioinformática da Biologia à flexibilidade molecular**. [s.l.: s.n.].

VOLKAMER, A. et al. DoGSiteScorer: a web server for automatic binding site prediction, analysis and druggability assessment. **Bioinformatics**, v. 28, n. 15, p. 2074–2075, 1 ago. 2012.

WALLIS, T. P. et al. Determination of the Disulfide Bond Arrangement of Dengue Virus NS1 Protein. **Journal of Biological Chemistry**, v. 279, n. 20, p. 20729–20741, 14 maio 2004.

WAN, S.-W. et al. Protection against Dengue Virus Infection in Mice by Administration of Antibodies against Modified Nonstructural Protein 1. **PLoS ONE**, v. 9, n. 3, p. e92495, 21 mar. 2014.

WANG, C. et al. Endocytosis of flavivirus NS1 is required for NS1-mediated endothelial hyperpermeability and is abolished by a single N-glycosylation site mutation. **PLOS Pathogens**, v. 15, n. 7, p. e1007938, 29 jul. 2019.

WANG, H. et al. Crystal structure of the C-terminal fragment of NS1 protein from yellow fever virus. **Science China Life Sciences**, v. 60, n. 12, p. 1403–1406, 1 dez. 2017.

WANG, Z. et al. Comprehensive evaluation of ten docking programs on a diverse set of protein–ligand complexes: the prediction accuracy of sampling power and scoring power. **Physical Chemistry Chemical Physics**, v. 18, n. 18, p. 12964–12975, 2016.

WEAVER, S. C. et al. Zika virus: History, emergence, biology, and prospects for control. **Antiviral Research**, v. 130, p. 69–80, jun. 2016.

WEBB, B.; SALI, A. Protein Structure Modeling with MODELLER. In: KAUFMANN, M.; KLINGER, C.; SAVELSBERGH, A. (Eds.). . **Functional Genomics**. New York, NY: Springer New York, 2017. v. 1654p. 39–54.

WHITE, D. O.; FENNER, F. **Medical virology**. 4th ed ed. San Diego: Academic Press, 1994.

WILCOXON, F. Individual Comparisons by Ranking Methods. **Biometrics Bulletin**, v. 1, n. 6, p. 80, dez. 1945.

WINKLER, G. et al. Evidence that the mature form of the flavivirus nonstructural protein NS1 is a dimer. **Virology**, v. 162, n. 1, p. 187–196, 1 jan. 1988.

XIA, H. et al. An evolutionary NS1 mutation enhances Zika virus evasion of host interferon induction. **Nature Communications**, v. 9, n. 1, p. 414, dez. 2018.

XU, D.; ZHANG, Y. Improving the Physical Realism and Structural Accuracy of Protein Models by a Two-Step Atomic-Level Energy Minimization. **Biophysical Journal**, v. 101, n. 10, p. 2525–2534, nov. 2011.

XU, X. et al. Contribution of intertwined loop to membrane association revealed by Zika virus full-length NS 1 structure. **The EMBO Journal**, v. 35, n. 20, p. 2170–2178, 17 out. 2016.

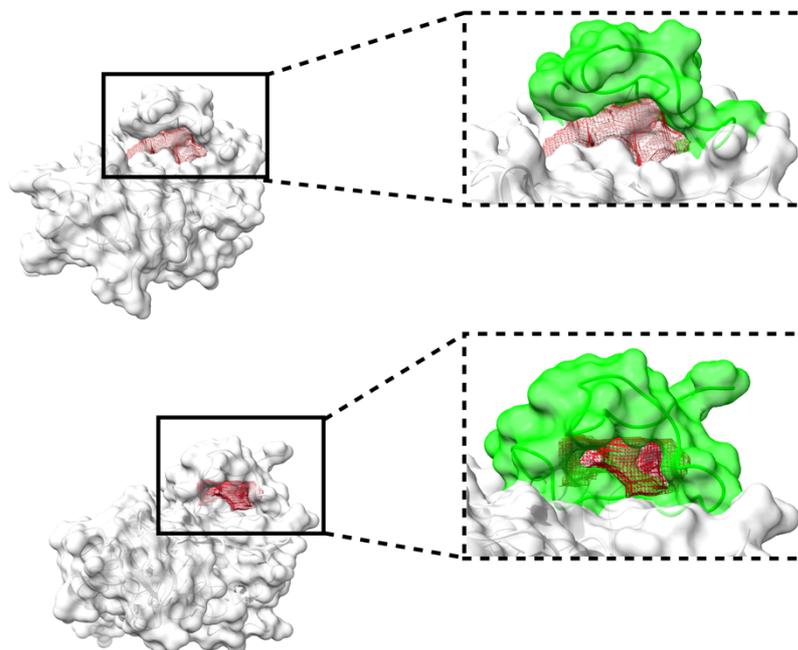
YANG, J.; ZHANG, Y. I-TASSER server: new development for protein structure and function predictions. **Nucleic Acids Research**, v. 43, n. 1, p. 174–181, 1 jul. 2015.

YEN, L.-C. et al. The C Terminus of the Core β -Ladder Domain in Japanese Encephalitis Virus Nonstructural Protein 1 Is Flexible for Accommodation of Heterologous Epitope Fusion. **Journal of Virology**, v. 90, n. 3, p. 1178–1189, 2 jan. 2016.

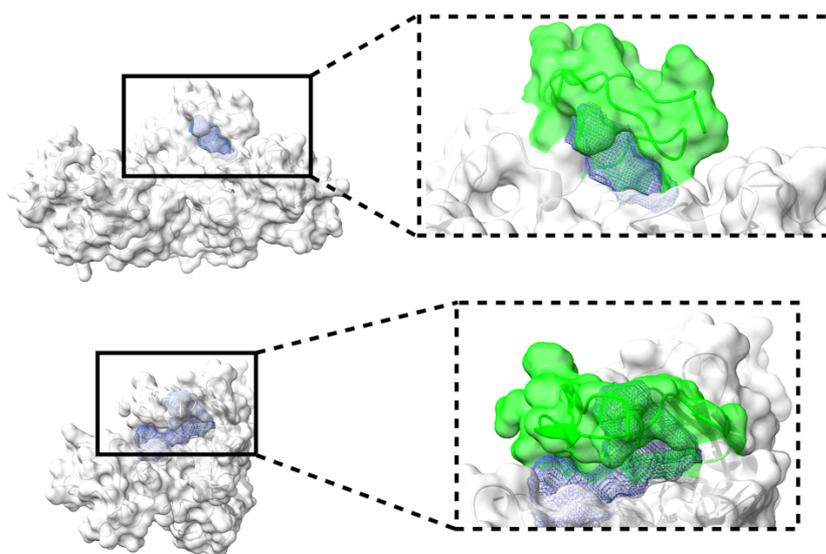
ZHANG, J.; LIANG, Y.; ZHANG, Y. Atomic-Level Protein Structure Refinement Using Fragment-Guided Molecular Dynamics Conformation Sampling. **Structure**, v. 19, n. 12, p. 1784–1795, dez. 2011.

ZHANG, Y. TM-align: a protein structure alignment algorithm based on the TM-score. **Nucleic Acids Research**, v. 33, n. 7, p. 2302–2309, 11 abr. 2005.

APÊNCIDE A – Estruturas da NS1 selecionadas para triagem virtual de compostos

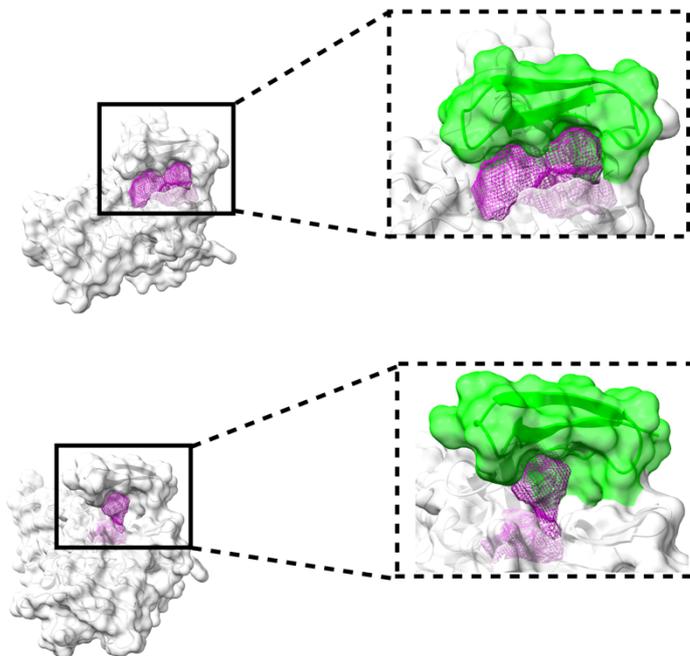
NS1 DENV

No *chart* superior está a estrutura DENV 46 e no inferior encontra-se DENV 73 em superfície branca. Em tom de vermelho e em superfície *mesh* está destacado o *pocket*. Em superfície verde, está selecionado os resíduos do β -roll (1 o 25).

NS1 YFV

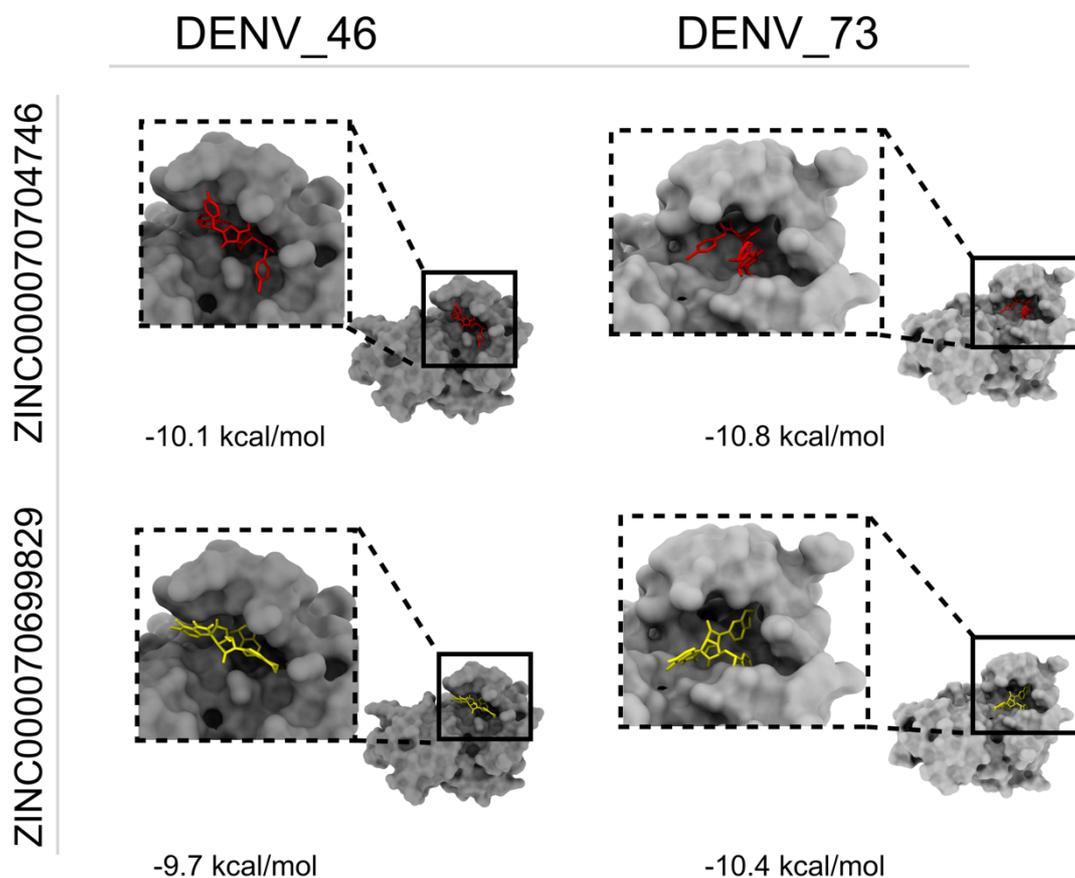
No *chart* superior está a estrutura YFV 65 e no inferior encontra-se YFV 162 em superfície branca. Em tom de azul e em superfície *mesh* está destacado o *pocket*. Em superfície verde, está selecionado os resíduos do β -roll (1 o 25).

NS1 ZIKV



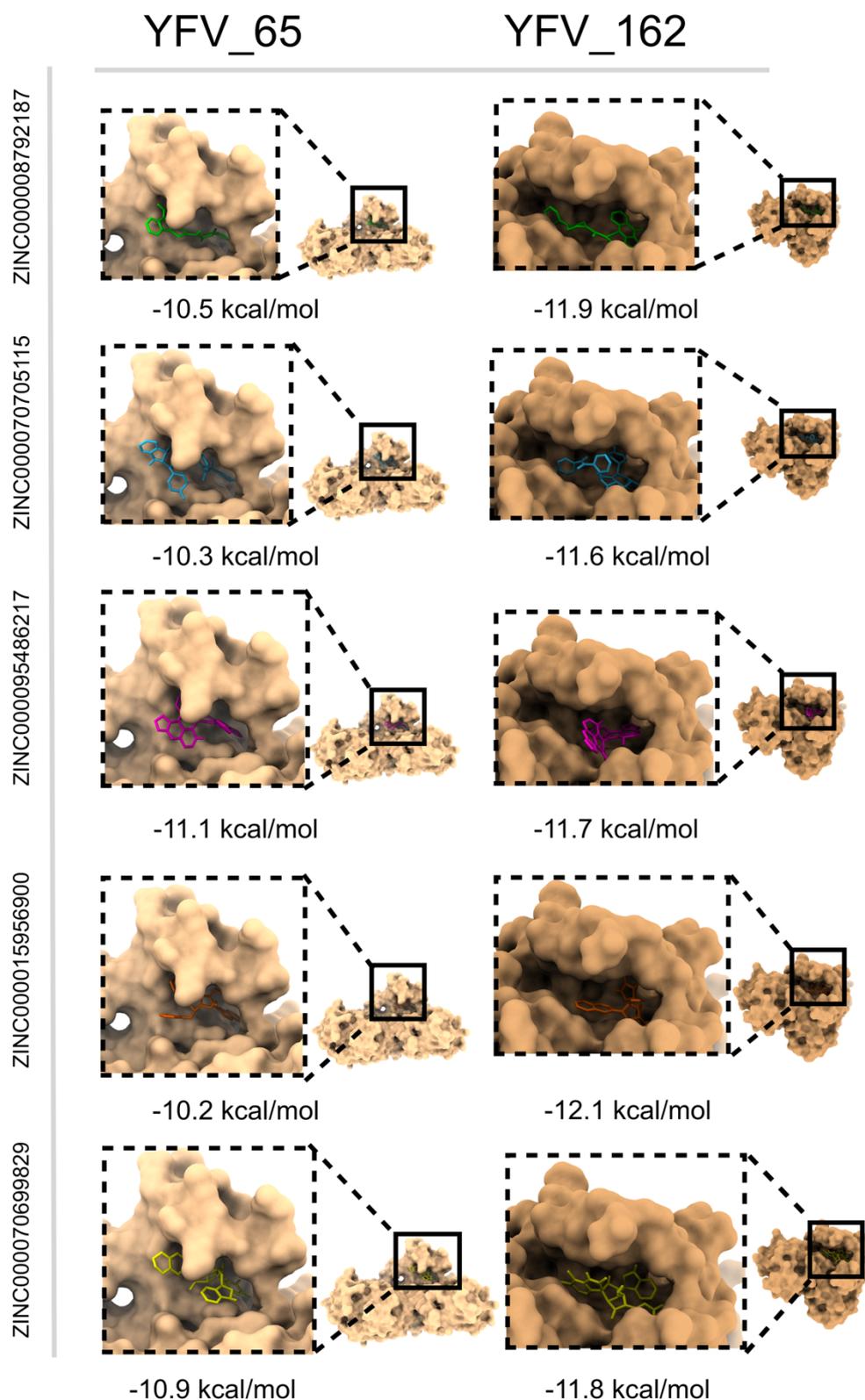
No *chart* superior está a estrutura ZIKV 84 e no inferior encontra-se ZIKV 144 em superfície branca. Em tom de magenta e em superfície *mesh* está destacado o *pocket*. Em superfície verde, está selecionado os resíduos do β -roll (1 o 25).

APÊNDICE B – Modos de ligação dos compostos em comum nas cavidades da NS1
DENV_46 e DENV_73 por critério de afinidade



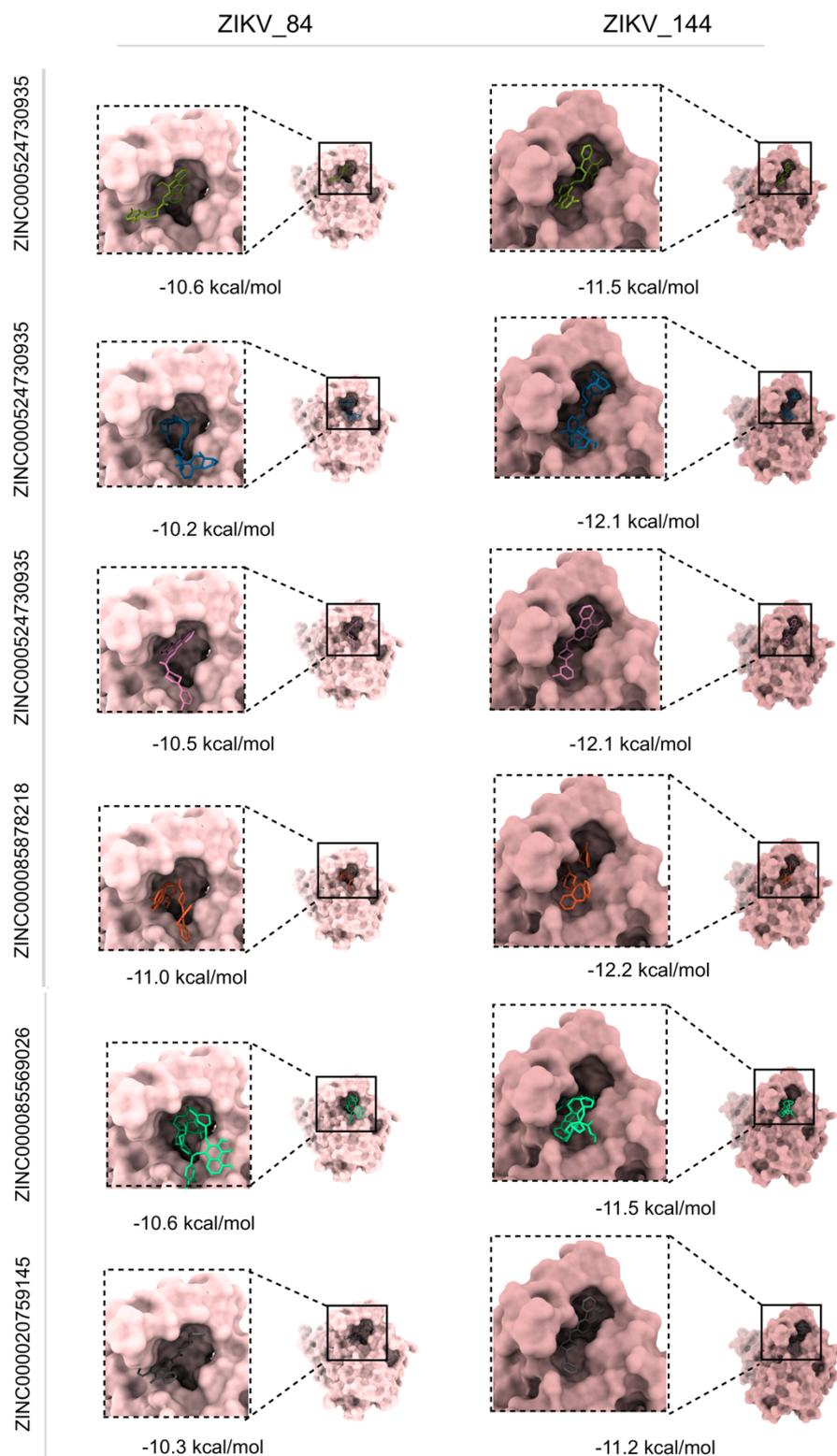
Na primeira coluna está representado as estruturas do DENV 46 e na segunda coluna as estruturas do DENV 73, ambos em superfície cinza escuro e cinza claro, respectivamente. Na primeira linha tem-se o composto ZINC000070704746 em *stick* vermelho e na segunda linha o composto ZINC000070699829 em *stick* amarelo com as respectivas energias do melhor modo de ligação.

APÊNDICE C – Modos de ligação dos compostos em comum nas cavidades da NS1
YFV_65 e YFV_162 por critério de afinidade



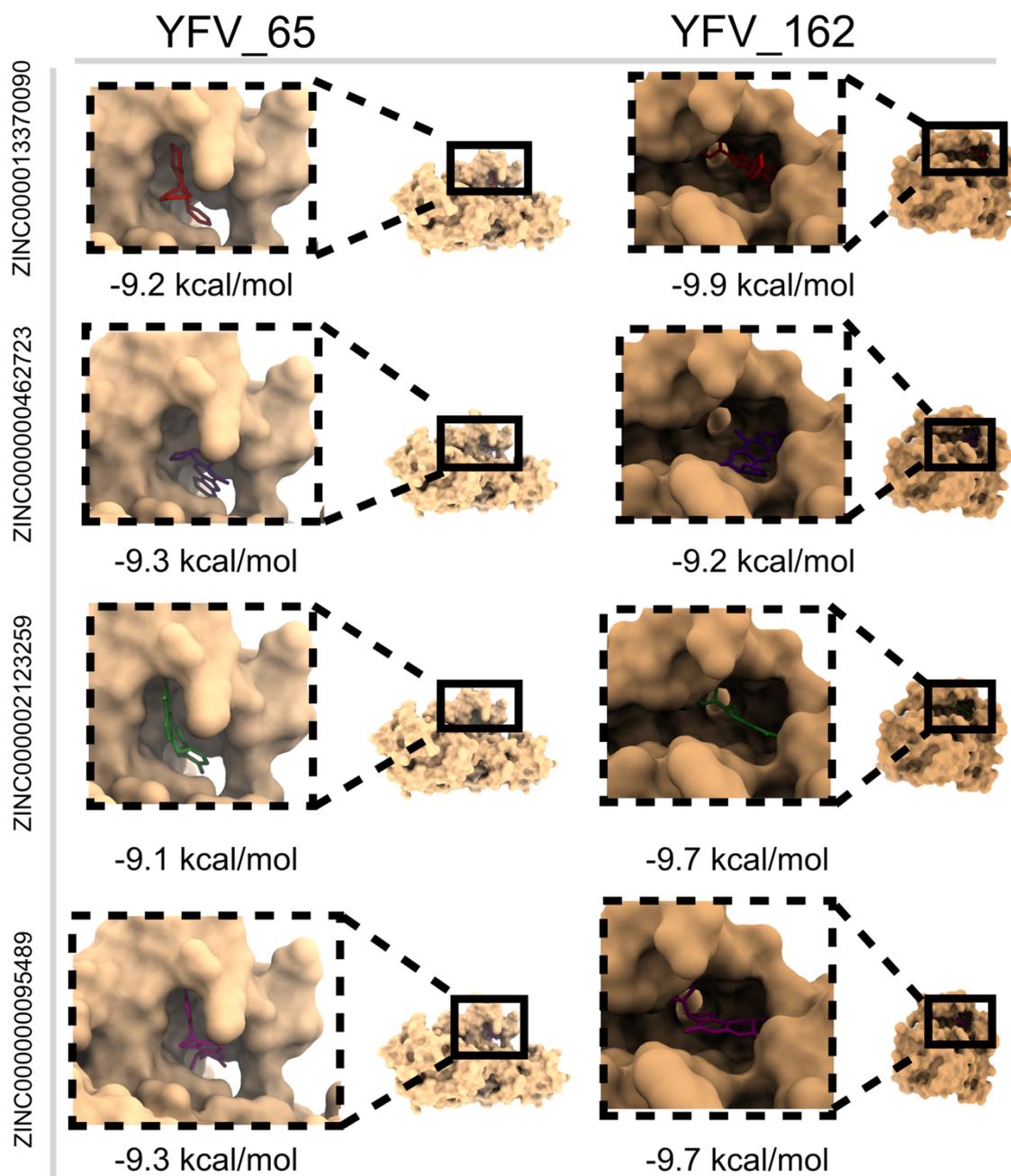
Na primeira coluna está representado as estruturas do YFV 65 e na segunda coluna as estruturas do YFV 162, ambos em superfície bege claro e bege mais escuro, respectivamente. Na orientação das linhas (horizontal, de cima para baixo) estão os compostos ZINC000008792187, ZINC000070705115, ZINC000095486217, ZINC000015956900 e ZINC000070699829 representados em *sticks*.

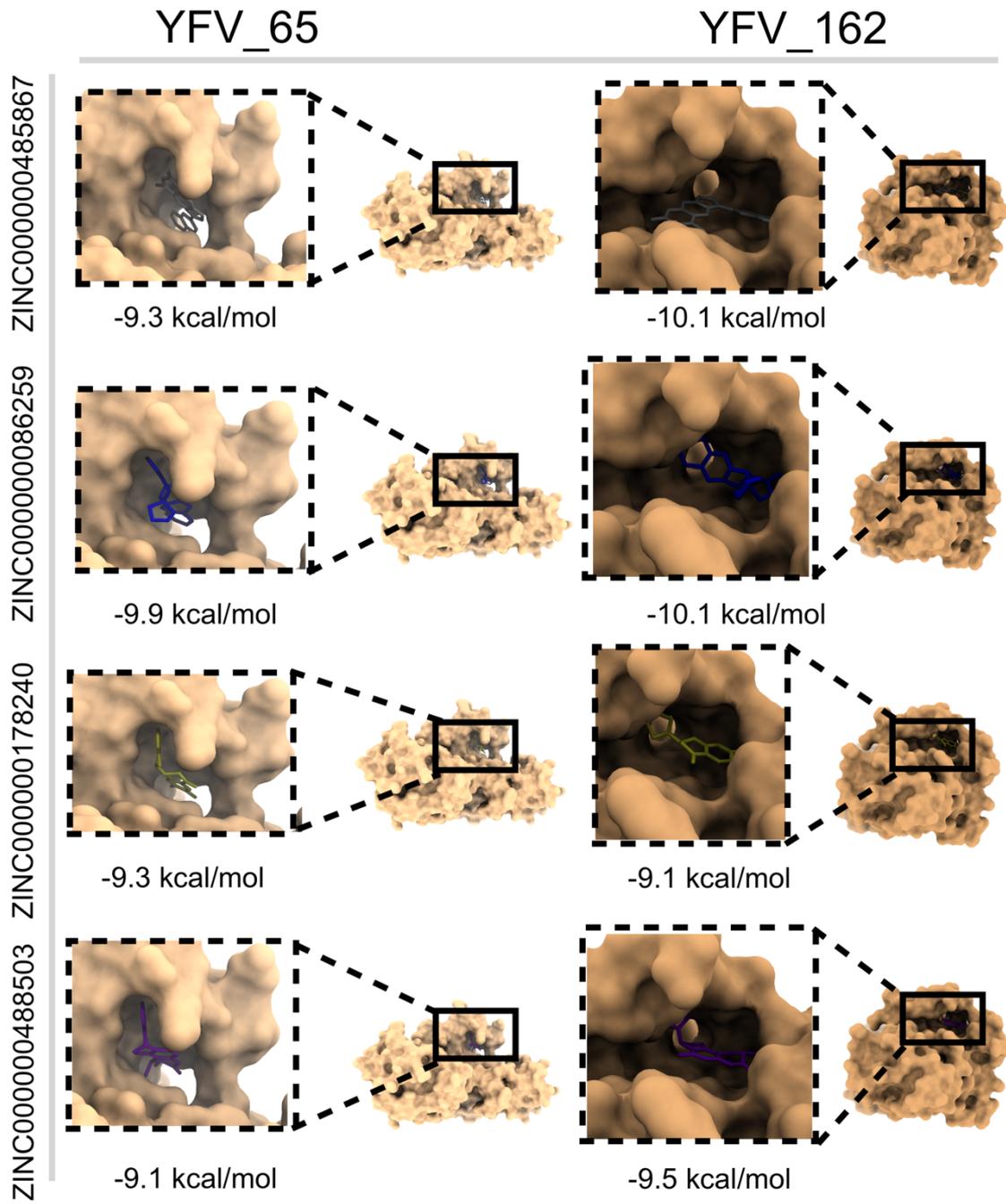
APÊNDICE D - Modos de ligação dos compostos em comum nas cavidades da NS1
ZIKV_84 e ZIKV_144 por critério de afinidade

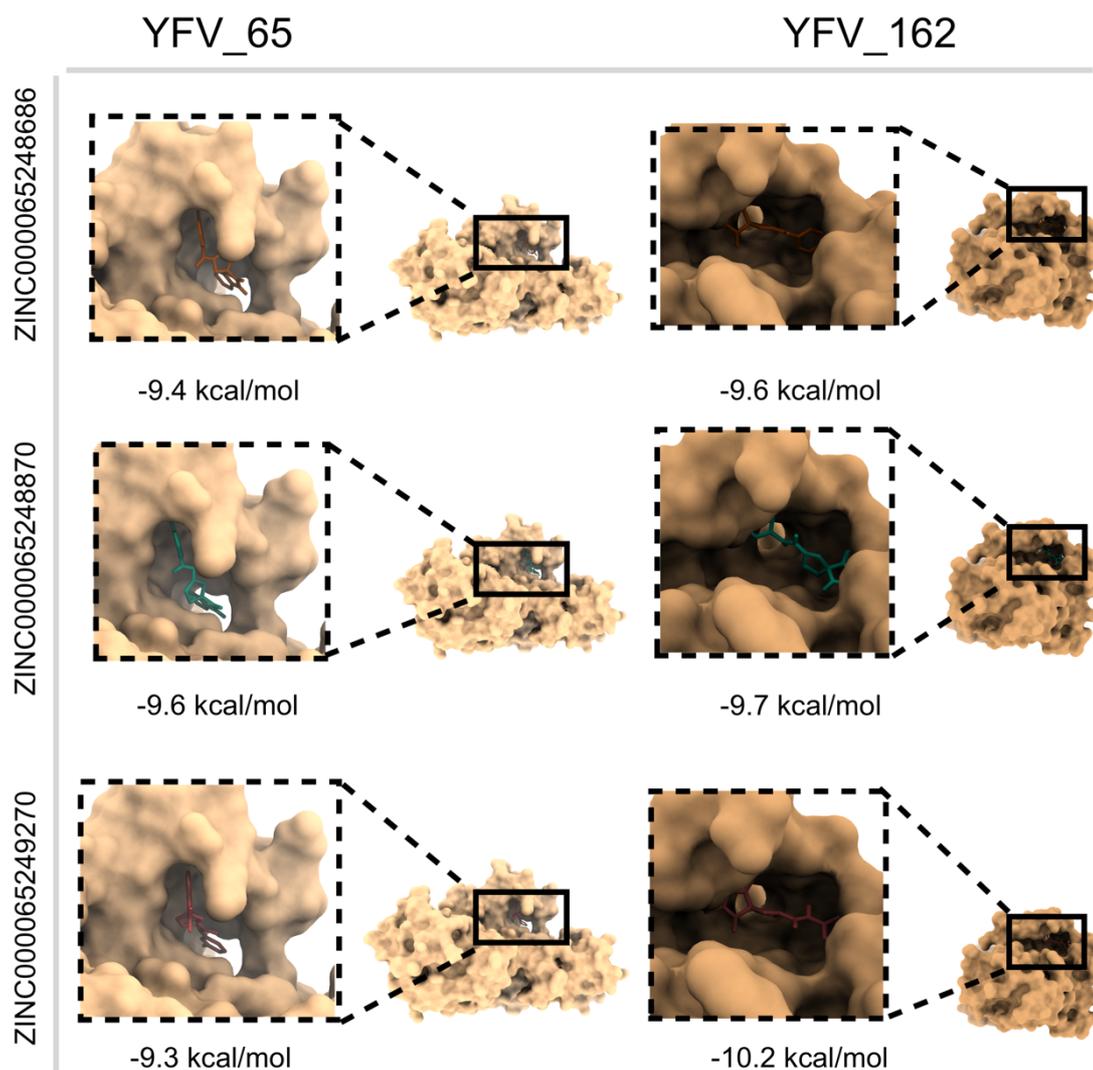


Na primeira coluna está representado as estruturas do ZIKV 84 e na segunda coluna as estruturas do ZIKV 144, ambos em superfície rosa claro e rosa mais escuro, respectivamente. Na orientação das linhas (horizontal, de cima para baixo) estão os compostos ZINC000524730935, ZINC000070691963, ZINC00008792352, ZINC000085878218, ZINC000085569026 e ZINC000020759145 representados em *sticks*.

APÊNDICE E - Modos de ligação dos compostos em comum nas cavidades da NS1
YFV_65 e YFV_162 por critério de eficiência

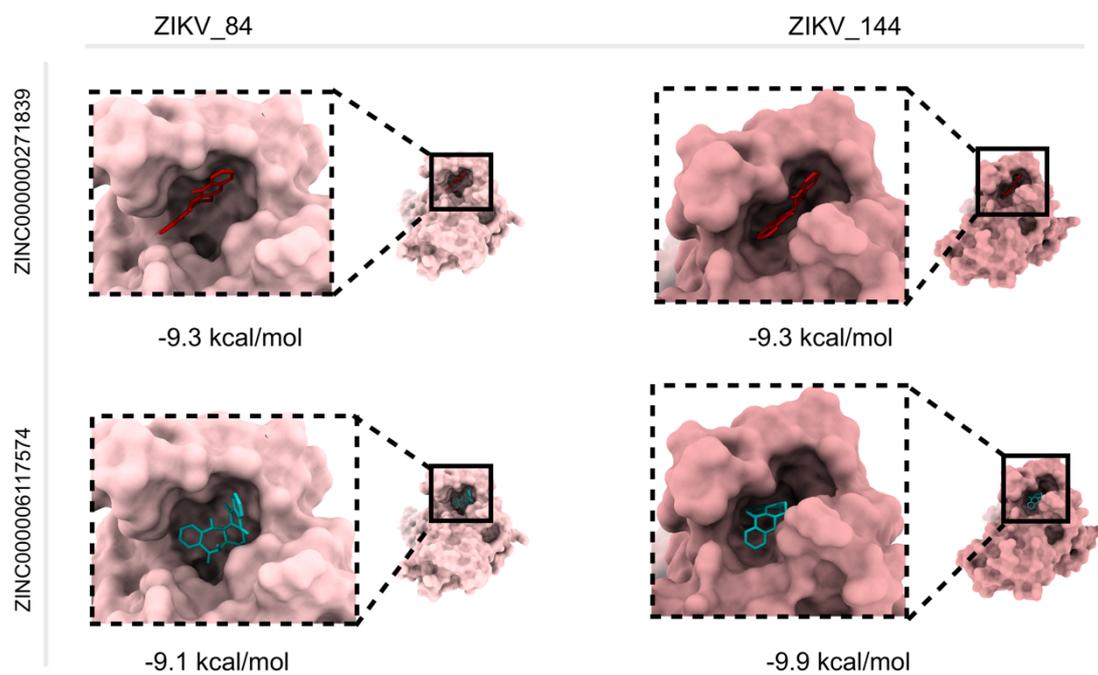






Na primeira coluna está representado as estruturas do YFV 65 e na segunda coluna as estruturas do YFV 162, ambos em superfície bege claro e bege mais escuro, respectivamente. Na orientação das linhas (horizontal, de cima para baixo) estão os compostos ZINC000013370090, ZINC000000462723, ZINC000002123259, ZINC000000095489, ZINC000000485867, ZINC000000086259, ZINC000000178240, ZINC000000488503, ZINC000065248686, ZINC000065248870 e ZINC000065249270 representados em *sticks*.

APÊNCIDE F - Modos de ligação dos compostos em comum nas cavidades da NS1
ZIKV_84 e ZIKV_144 por critério de eficiência



Na primeira coluna está representado as estruturas do ZIKV 84 e na segunda coluna as estruturas do ZIKV 144, ambos em superfície rosa claro e rosa mais escuro, respectivamente. Na orientação das linhas (horizontal, de cima para baixo) estão os compostos ZINC000000271839 e ZINC000006117574 em *sticks*.