

UNIVERSIDADE FEDERAL DE SÃO JOÃO DEL-REI
CURSO DE ENGENHARIA DE BIOPROCESSOS

PREDIÇÃO DA ESTRUTURA 3D DA PROTEÍNA nsP3 DO *MAYARO VIRUS* POR
MODELAGEM POR HOMOLOGIA E TRIAGEM VIRTUAL PARA A BUSCA DE
ANTIVIRAIS

Jenyfer Emanuelle Figueiredo

Orientador: Prof. Dr. José Carlos de Magalhães

Co-orientadora: M.^a Priscila Gonçalves Ferreira

Ouro Branco, Minas Gerais

Novembro/2016

PREDIÇÃO DA ESTRUTURA 3D DA PROTEÍNA nsP3 DO *MAYARO VIRUS* POR
MODELAGEM POR HOMOLOGIA E TRIAGEM VIRTUAL PARA A BUSCA DE
ANTIVIRAIS

Jenyfer Emanuelle Figueiredo

Trabalho de conclusão de curso apresentado
em cumprimento às exigências do Curso de
Bacharelado em Engenharia de Bioprocessos
da Universidade Federal de São João Del-Rei.

Orientador: Prof. Dr. José Carlos de Magalhães

Co-orientadora: M.^a Priscila Gonçalves Ferreira

Ouro Branco, Minas Gerais

Novembro/2016

UNIVERSIDADE FEDERAL DE SÃO JOÃO DEL-REI
CURSO DE ENGENHARIA DE BIOPROCESSOS

**PREDIÇÃO DA ESTRUTURA 3D DA PROTEÍNA nsP3 DO MAYARO VIRUS POR
MODELAGEM POR HOMOLOGIA E TRIAGEM VIRTUAL PARA A BUSCA DE
ANTIVIRAIS**

Jenyfer Emanuelle Figueiredo

A banca examinadora, composta pelos membros abaixo,avaliou este TCC.

Dr. José Carlos de Magalhães- Orientador

Dr. Edson Romano Nucci- Examinador 1

M.e Ivan Evangelista do Vale Coelho

Ouro Branco 01/12/2016

“Por vezes sentimos que aquilo que fazemos não é senão uma gota de água no mar. Mas o mar seria menor se lhe faltasse uma gota.”

(Madre Teresa de Calcutá)

DEDICATÓRIA

Dedico este trabalho aos meus pais, Luciene e Manoel, que tanto se esforçaram para me educar e me fazer continuar nessa jornada.

AGRADECIMENTOS

Primeiramente agradeço a Deus por ter me fornecido saúde e força para superar todos os desafios que encontrei até chegar aqui.

Aos meus queridos pais Luciene R. dos Anjos Figueiredo e Manoel A. C. Figueiredo por minha educação, afeto e apoio incondicionais.

À minha Grande Família, Figueiredo e Rafael dos Anjos, pelo apoio e carinho.

As minhas amizades, antigas e adquiridas ao longo do curso. Vocês fizeram essa etapa ser mais divertida e proveitosa. Em especial agradeço a minha querida amiga Rayla Novais pela convivência, conselhos e apoio.

À Universidade Federal de São João del-Rei pela oportunidade de graduação em Engenharia de Bioprocessos, pela estrutura, corpo técnico, discente que foram cruciais para o meu desenvolvimento pessoal e profissional.

A todos os mestres docentes do Campus Alto Paraopeba pelos ensinamentos, em especial, o meu Orientador José Carlos de Magalhães, pelo acolhimento na Microbiologia desde os períodos iniciais da graduação, pelo incentivo, instrução e confiança depositada em meu trabalho.

À Priscila Gonçalves Ferreira pela constante co-orientação, pela disposição em ajudar sempre eu a recorri e acima de tudo pela grande amizade.

Ao pessoal do Laboratório de Microbiologia pelo companheirismo e conhecimentos compartilhados.

Aos membros da banca examinadora Dr. Edson R. Nucci e M.e Ivan E. V. Coelho por aceitarem esse desafio, contribuindo assim para o enriquecimento de minha formação.

Aos órgãos de fomento CNPq e Fapemig pelo apoio financeiro ao projeto.

RESUMO

O *Mayaro vírus* (MAYV), pertencente a família *Togaviridae*, é causador da Febre Mayaro, doença associada a severas artralgias, com sintomas confundidos aos da Febre Dengue e Chikungunya. O vírus é endêmico em áreas rurais, transmitido pelo *Haemagogus janthinomys*, porém há riscos de sua urbanização devido o *Aedes aegypti* ser um vetor potencial. Até o momento não há vacina e/ou terapia disponível, e o desenvolvimento de novos fármacos é limitado pela ausência da estrutura 3D das proteínas de MAYV. A proteína nsP3 está envolvida nas etapas de replicação viral e, portanto, é uma boa candidata a alvo molecular para o desenvolvimento de novos fármacos. Dessa forma, o objetivo desse trabalho foi obter o modelo 3D para a proteína nsP3 a fim de auxiliar a busca de antivirais por meio de triagem virtual. O modelo da nsP3 de MAYV foi construído via modelagem por homologia no SWISS MODEL (SM), e a qualidade avaliada pelo Gráfico de Ramachandran (GR), QMEAN, RMSD e ANOLEA. Para descrever a dinâmica conformacional, confirmar a estabilidade do modelo e refinar a estrutura foi realizada a simulação de dinâmica molecular (DM). Para a triagem virtual foi utilizada a estratégia de ancoragem molecular pelo sítio dirigido no qual o modelo da proteína foi o receptor e os ligantes foram provenientes da literatura e ZINC. Para a construção do modelo foi utilizada a proteína nsP3 de *Sindbis virus* (código:4GUA.1A). O modelo da proteína nsP3 de MAYV consistiu de uma estrutura de 11 alfa-hélices e 10 folhas beta. O modelo, construído, validado e refinado por dinâmica molecular foi empregado como receptor num processo de triagem virtual por meio do *docking* com o sítio dirigido. O sítio de ligação prevaleceu a região onde se encontra onde se encontrava o ligante cristalográfico 2-N-morfolino-etanosulfônico (MES), da 4GUA.1A, obtendo-se energia de ligação para este ligante no sítio da nsP3 de -5,7 kcal/mol. Estudos indicam que este sítio é um macrodomínio também ocupado pela ADP-ribose, que foi ancorada e apresentou maior afinidade do que o MES (-8.7 kcal/mol). Sendo utilizada como padrão na do ancoragem de 296 moléculas provenientes da literatura e da base de dados do ZINC, a menor energia encontrada foi -11,5 kcal/mol da molécula zinc_8442275. De maneira geral cerca de 20% das moléculas testadas se mostraram potenciais antivirais. Entretanto, estudos como a dinâmica molecular do complexo ADP-ribose-nsP3 e também moléculas para o complexo das moléculas candidatas com a nsP3. Além disso, testes *in vitro* de atividade antiviral e citotoxicidade para as moléculas candidatas ainda serão efetuados.

Palavras chave: nsP3, MAYV, modelagem por homologia, *docking*, triagem virtual, antiviral.

LISTA DE FIGURAS

- Figura 1.** Ciclo de transmissão do MAYV. Fonte: adaptado de Caetano, 2016. **2**
- Figura 2.** Esquema representando a organização das nove proteínas codificadas pelo genoma do MAYV, o qual apresenta em sua extremidade 3' uma calda de poliadenilato (Poli-A) na extremidade 5' o resíduo 7-metilguanosina (cap5'). Fonte: Adaptado de Napoleão-Pego *et al.*, 2014. **3**
- Figura 3.** Modelo e molde da proteína nsP3 (a) Modelo construído no SWISS-MODEL da proteína nsP3 do MAYV (NP_740689) usando como molde (b) a poliproteína 4GUA.1A ((Chain A, Alphavirus P23 pro-Zbd) de SINV. A visualização de (a) foi efetuada no programa VMD, modo *new cartoon* e (b) foi obtida no PDB. **14**
- Figura 4.** Gráfico Ramachandram do modelo da proteína nsP3 do MAYV. O gráfico mostra que 97,5% dos resíduos de aminoácidos estão na região favorável e 2,5% na região permitida e 0% na região proibida. **15**
- Figura 5.** Representação gráfica da análise de energia ao longo da sequência de aminoácidos do modelo da nsP3 de MAYV, gerado no ANOLEA. As barras verdes indicam que as energias estão menores que zero ao passo que as vermelhas são utilizadas para energias superiores a zero. **16**
- Figura 6.** Análise de qualidade, representação gráfica da função *Z-score* QMEAN4 gerado no SWISS-MODEL. Mostrando que todos os valores foram negativos e próximos de zero, indicando o quão próximo o modelo está de estruturas resolvidas experimentalmente... **17**
- Figura 7.** Representação gráfica do QMEAN4 normalizado gerado no portal SWISS-MODE. O *Z-score* encontra-se entre 1 e 2. **17**
- Figura 8.** Curva de de RMSD (Å) *versus* tempo (fs) durante a simulação por DM da nsP3 MAYV durante 20ns. **18**
- Figura 9.** Complexo nsP3 MAYV-MES em (a) e em (b) a sobreposição do MES cristalográfico no processo de *re-docking* como validação do procedimento de *docking*. Elaborado no *DS Visualizer*. **19**
- Figura 10.** Interações intermoleculares da ADP-Ribose com a proteína npP3 modelada de MAYV (a), CHIKV (b) Adaptado de Rungrotmongkolet al., 2010 e VEEV Adaptado de Rungrotmongkolet al., 2010 (c). mostram as interações da ADP-Ribose com a proteína nsP3 de CHIKV e VEEV respectivamente, adaptados de Rungrotmongkolet al., 2010. Legenda: Círculos em rosa-resíduos envolvidos em ligações de hidrogênio ou interações apolares; Círculos em verde-resíduos envolvidos em interações de van der Waals; Halo azul em torno do resíduo-interação com o solvente (água); Halo azul em torno do átomo – interação com o solvente (água). Seta tracejada em verde-ligação de hidrogênio com a cadeia principal. Seta tracejada em azul-ligação de hidrogênio com a cadeia lateral; Linha laranja- Interações Pi. **21**
- Figura 11.** Distribuição percentual da energia de ligação (Kcal/mol) das 296 moléculas submetidas ao *docking*, tendo como receptor a nsP3 MAYV e parâmetro energético o valor -8,7 kcal/mol (ADP-ribose). **23**

Figura 12. Interações interatômicas dos aminoácidos da proteína nsP3 MAYV com os ligantes promissores provenientes do *docking*. Em (a) a interação com a molécula Derivado Fenólico ((2S,3S,4S,5S,6S)-4-{[(2S,3R,4S,6R)-3,4-dihydroxy-6-methyloxan-2-yl]oxy}-6-[2-(3,4-dioxocyclohexyl)ethoxy]-5-hydroxy-2(hydroxymethyl)oxan-3-yl]-3-(3,4dioxocyclohexyl)propanoate), em (b) com a molécula molécula Zinc_8442273 (6,6'-bis(2-{[(4-methylphenyl)sulfonyl]amino}-1,3-benzothiazole) e em (c) do com a molécula Zinc_8442275 (6,6'-bis{N-[(4-chlorophenyl)sulfonyl]-1,3-benzoxazol-2-amine}), que apresentou a menor energia de ligação, -11,5 Kcal/mol. Legenda: Círculos em rosa-resíduos envolvidos em ligações de hidrogênio ou interações apolares; Círculos em verde-resíduos envolvidos em interações de van der Waals; Halo azul em torno do resíduo-interação com o solvente (água); Halo azul em torno do átomo – interação com o solvente (água). Seta tracejada em verde-ligação de hidrogênio com a cadeia principal. Seta tracejada em azul-ligação de hidrogênio com a cadeia lateral; Linha laranja-Interações Pi. 26

LISTA DE TABELAS

Tabela 1. Tamanho e função descrita pela literatura das proteínas do MAYV.	4
Tabela 2. Informações estatísticas fornecidas pelo <i>BLASTp</i> tendo como sequência de busca (<i>query</i>) a NP_740689.1.....	12

LISTA DE QUADROS

Quadro 1. Sequência em formato .FASTA da proteína nsP3 de MAYV.....	12
--	-----------

LISTA DE ABREVIATURAS E SIGLAS

3D	Tridimensional
6k	Proteína 6k
ADP-ribose	Adenosina Difosfato Ribose
ANOLEA	do inglês <i>Atomic Non-Local Enviroment Assessment</i>
BLAST	do inglês <i>Basic Local Alignment Search Tool</i>
CDC	do inglês <i>Centers for Disease Control and Prevention</i>
C	Proteína do Capsídeo
CHIKV	do inglês <i>Chikungunya virus</i>
DENV	do inglês <i>Dengue virus</i>
DM	Dinâmica Molecular
DS Visualizer	do inglês <i>Discovery Studio</i>
E1	Glicoproteína E1
E2	Glicoproteína E2
E3	Glicoproteína E3
MAYV	do inglês <i>Mayaro virus</i>
MES	ácido 2-N-morfolino-etanosulfônico
NCBI	do inglês <i>National Centers Bioinformatics Institute</i>
nsP1	do inglês <i>Non structuralprotein 1</i>
nsP2	do inglês <i>Non structuralprotein 2</i>
nsP3	do inglês <i>Non structuralprotein 3</i>
nsP4	do inglês <i>Non structuralprotein 4</i>
ORF	do ingles <i>Open Read Frame</i>
PDB	do inglês <i>Protein Data Bank</i>
ROCV	do inglês <i>Rocio vírus</i>
RMN	Ressonância Magnética Nuclear
RMSD	do ingles <i>Root-Mean-Square Deviation</i>
VEEV	do inglês <i>Venezuelan Encephalite EquineVirus</i>
VMD	do inglês <i>Visual Molecular Dynamics</i>
WNV	do inglês <i>West Nilo virus</i>
ZIKV	do inglês <i>Zica virus</i>

SUMÁRIO

1. INTRODUÇÃO	1
2. OBJETIVO	7
2.1. Objetivo geral	7
2.2. Objetivos específicos;	7
3. METODOLOGIA	7
3.1. Recuperação da sequência primária nsP3	7
3.2. Busca por moldes	7
3.3. Construção do modelo.....	8
3.4. Validação do Modelo	8
3.5. Simulação de Dinâmica Molecular	9
3.6. Triagem virtual de compostos ligantes da nsP3	10
3.6.1 Preparação do receptor e ligantes	10
3.6.2 Ancoragem molecular	11
4. RESULTADOS E DISCUSSÃO	11
4.1. Recuperação da sequência nsP3	11
4.2. Alinhamento Local	12
4.3 Construção do modelo e seleção dos moldes	13
4.4. Validação do modelo construído	14
4.5. Simulação de dinâmica molecular.....	18
5. CONSIDERAÇÕES FINAIS E PERSPECTIVAS	27
6.REFERÊNCIAS BIBLIOGRÁFICAS.....	28
APÊNDICE I	33

1. INTRODUÇÃO

Embora haja a possibilidade de prevenção e erradicação, as doenças infecciosas humanas são de grande preocupação para a Saúde Pública devido ao seu caráter irruptivo em nível global, imprevisibilidade e transmissibilidade (Lima-Camara, 2016 *apud* Fauci, 2012). Acredita-se que a maior parte dos patógenos causadores desse tipo de doença tem origem zoonótica, com ciclo de vida envolvendo um vetor artrópode e um animal silvestre. Entretanto, com as modificações do ecossistema pelo homem, alguns vetores estão se tornando cada vez mais adaptados aos meios urbanos, sinantrópicos, facilitando a transmissão desses patógenos ao homem (Lima-Camara, 2016 *apud* Norris, 2004).

Em especial, vale ressaltar que uma classe de patógenos que vem causando preocupação por estarem emergindo em áreas antes indenes são a dos arbovírus (*Arthropod-born virus*). No mundo inteiro, esta classe vem causando epidemias notáveis em humanos e possui muitos agentes causadores de doenças febris em humanos, embora haja uma prevalência na região dos trópicos (Napoleão-Pego *et al.*, 2014; Forshey *et al.*, 2010). Há relatos em que a disseminação das arboviroses esteja ligada a fatores climáticos, antropogênicos como atividades agrícolas, desmatamento e reflorestamento e mudança nos ecossistemas hídricos (Pedroso & Rocha, 2009).

Os Arbovírus têm essa denominação devido parte do seu ciclo replicativo ocorrer em insetos e serem transmitidos aos humanos e outros animais devido à picada de artrópodes hematófagos (Lopes, 2014), sendo um grupo, e não uma categoria taxonômica. De acordo com o catálogo do *Centers for Disease Control and Prevention* (CDC), há cerca de 537 espécies relatadas, das quais 134 (25%) estão relacionadas a alguma patogenia confirmadas em humanos (Lopes, 2014; Gubler, 2001). Os arbovírus causadores de doenças em humanos ou animais de sangue quente podem ser encontrados em cinco famílias de vírus, quais sejam: Bunyaviridae, Togaviridae, Flaviviridae, Reoviridae e Rhabdoviridae.

No Brasil, vários desses arbovírus estão em circulação como, por exemplo, Dengue (DENV), Encefalite Equina Venezuelana (VEEV), Rocio (ROCV), Mayaro (MAYV), West Nilo Virus (WNV), Chikungunya (CHIKV) e Zica (ZIKV), sendo um desafio para a saúde pública o diagnóstico e controle dessas novas arboviroses (Lima-Camara, 2016).

O MAYV é endêmico na região amazônica (Spindola, 2014; Coimbra *et al.*, 2007). E por exemplo, sua presença já foi identificada no estado do Mato Grosso durante uma epidemia de Dengue, indicando que o vírus está circulando silenciosamente e ampliando seu espectro geográfico (Vieira, 2015). Além disso, pessoas que costumam visitar essas regiões

podem estar sujeitas à infecção causada pelo MAYV. Em São Paulo, por exemplo, foi identificada a presença do MAYV em três pacientes que haviam regressado de Camapuã, MS (Coimbra *et al.*, 2007).

O MAYV é membro da família *Togaviridae*, gênero *Alphavirus*, os vírus desta família são envelopados com uma estrutura contendo espículas que projetam para a superfície viral, internamente, possuem um capsídeo icosaédrico e ao todo medem cerca de 70nm de diâmetro (Caetano, 2016). Ele é o causador da febre Mayaro, doença debilitante que provoca uma prolongada artralgia, similar à febre Chikungunya. Dentre os outros sintomas relatados, sabe-se que a doença pode causar cefaleia, febre, diarreia, náuseas, vômitos, erupções cutâneas, dor ocular mialgia e prolongada artralgia (Figueiredo, 2014; Halsey, 2013; Coimbra *et al.*, 2007). Sendo este último sintoma similar ao da Febre Chikungunya, responsável por deixar a doença ainda mais debilitante. Estudos mostraram que a artralgia pode permanecer por meses em mais de 50% dos casos da doença (Santiago *et al.*, 2015).

O mosquito *Haemagogus janthinomys* é o principal vetor do MAYV, embora outros mosquitos dos gêneros *Culex* sp, *Psorophora* sp, *Coquillettidia* sp e *Sabethes* sp sejam relatados como vetores. E como vetor potencial há os mosquitos do gênero *Aedes* sp (Caetano, 2016). A **Figura 1** ilustra o ciclo de multiplicação desse vírus, que transita entre áreas enzoóticas, rurais e com a possibilidade de urbanização pelo *A. aegypti*.

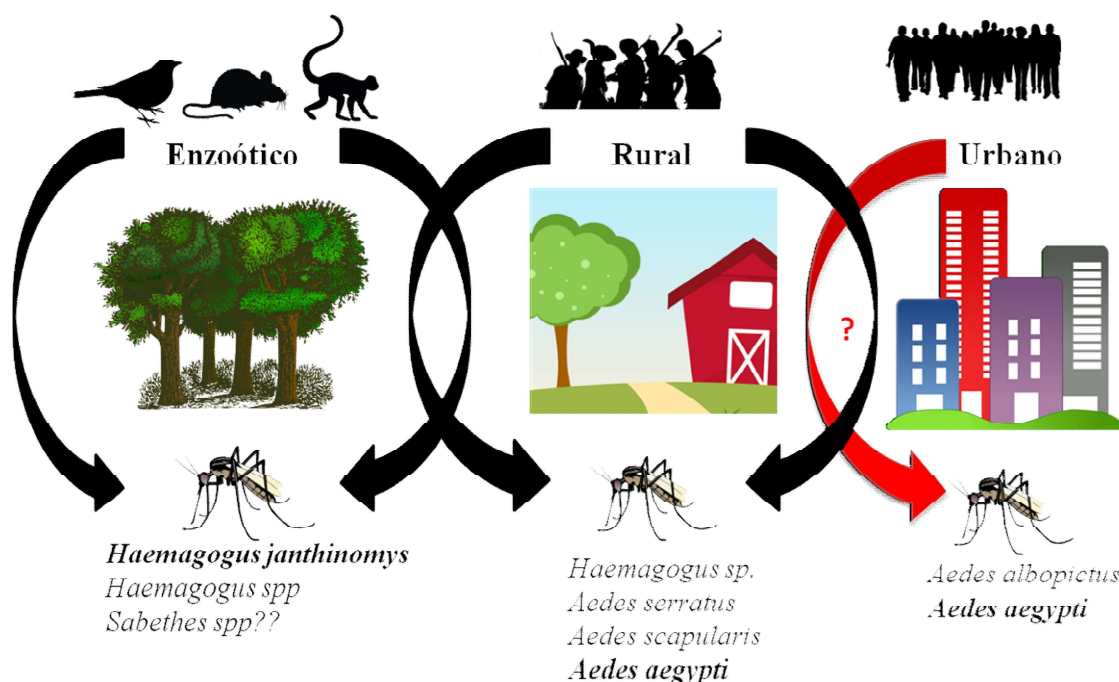


Figura 1. Ciclo de transmissão do MAYV. Fonte: adaptado de Caetano, 2016.

O genoma dos membros da família *Togaviridae* é constituído de uma fita simples linear de RNA de polaridade positiva de cerca de 12.000 nucleotídeos distribuídos em 9 genes (Caetano,2016). A estrutura do RNA viral consiste em duas janelas de leitura aberta (*ORFs*, do inglês *Open Read Frame*) e entre os genes há uma região muito conservada. A primeira e maior ORF, pode ser lida/traduzida diretamente como RNAm, e originar quatro proteínas não estruturais (nsP1-4), sendo essas essenciais para as etapas de transcrição e replicação viral. Já a segunda ORF é responsável por codificar as proteínas estruturais C (do capsídeo), as do envelope viral (E3, E2, E1) e 6K, a partir do RNA subgenômico 26S. Esse RNA é transcrito partindo-se de um RNA intermediário negativo que é complementar ao RNA genômico, que, após ser clivado por enzimas e furinas, dá origem às proteínas estruturais (Caetano, 2016 *apud* Mota *et al.*,2015;Napoleão-Pego *et al.*,2014).A **Figura 2** mostra um esquema da organização genômica do MAYV.

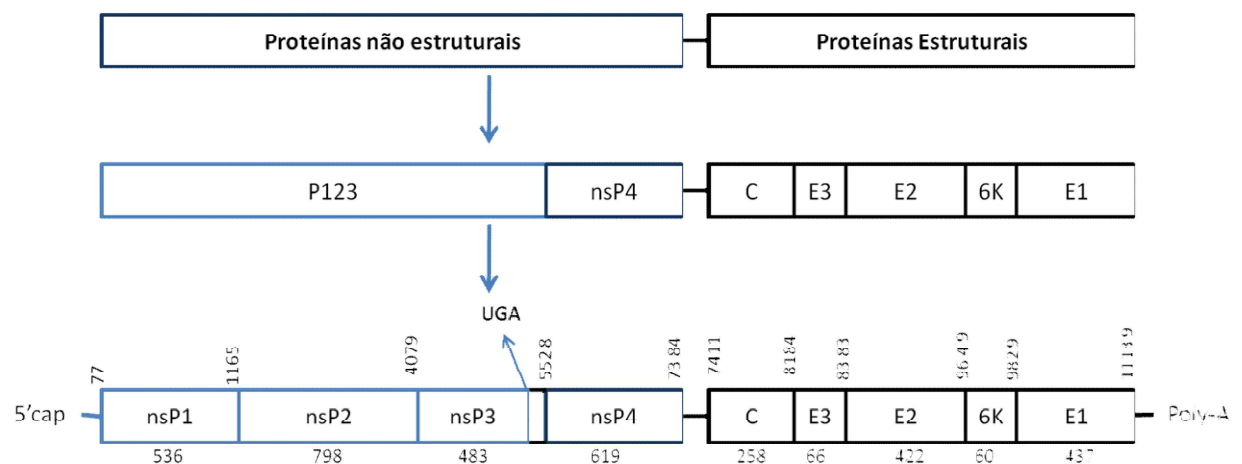


Figura 2. Esquema representando a organização das nove proteínas codificadas pelo genoma do MAYV, o qual apresenta em sua extremidade 3' uma calda de poliadenilato (Poli-A) na extremidade 5' o resíduo 7-metilguanosina (cap5').Fonte: Adaptado de Napoleão-Pego *et al.*,2014.

Já a **Tabela 1** faz um resumo das proteínas do MAYV, bem como suas respectivas funções descritas até o momento pela literatura.

Tabela 1. Tamanho e função descrita pela literatura das proteínas do MAYV.

Proteína	Tamanho (a.a)	Função	Referência
nsP1	536	Capeamento do RNA mensageiro.	LAVERGNE,2006
nsP2	798	Clivagem de sítios específicos na poliproteína não estrutural que é produzida no início da replicação viral Papel na replicação.	FERREIRA,2000. CHENG,1995.
nsP3	498	Desconhecida, mas acredita-se que mesma exerça um papel importante na síntese de mRNAsubgenômico 26s e no transporte dos componentes virais.	POWERS,2006. PARROT,2009. http://www.ncbi.nlm.nih.gov/nucest1885792
nsP4	619	Papel na replicação viral.	http://www.ncbi.nlm.nih.gov/nucest1885792
C	258	Montagem do capsídeo e clivagem do polipeptídio estrutural.	SKOGING,1998. GASPAR,2001. FERREIRA,2000.
E3	66	Acredita-se que esteja envolvida na tradução do genoma viral.	http://www.ncbi.nlm.nih.gov/nucest1885792
E2	422	Envolvida no reconhecimento de adsorção viral.	NAPOLÃO-PEGO <i>et al.</i> ,2014.
E1	437	Responsável pela fusão do envelope viral com a membrana endossomal do hospedeiro.	CHENG,1995
6k	60	Processamento do vírus ,permeabilização celular e Brotamento das partículas virais.	STRAUSS,1994.

Fonte: Adaptado de Napolão-Pego *et al.*,2014.

A função da nsP3 nos alphavirus ainda permanece desconhecida, no entanto tem sido demonstrado que algumas mutações ocorridas no gene que codifica a nsP3 podem afetar diretamente a síntese de RNA específico desses alphavirus e também causar efeitos deletérios na replicação (Karras, 2005; Foy *et al.*, 2013). Ao se basear na conservação das sequências, é possível notar que a nsP3 apresenta 3 domínios, sendo eles um macrodomínio amino-terminal (Domínio X), uma região central específica de alphavirus e uma sequência carboxiterminal de hipervariabilidade (Shin *et al.*, 2012). Há evidências bioquímicas e estruturais que os macrodomínios apresentam alta afinidade com os módulos de ligação da ADP-ribose (Karras *et al.*, 2005).

Recentemente, estudos *in vitro* relataram *A. aegypti*, vetor da febre dengue, como potencial transmissor do MAYV, ampliando a sua possibilidade de maior transmissão urbana, a qual não mais seria feita apenas pelo *Haemagogus* de hábito silvestre (LONG, *et*

al., 2011). Isso é um fator preocupação para a saúde pública, pois uma epidemia poderia ser capaz de abalar economicamente uma região devido à forma mais grave levar à incapacidade de trabalho por um longo período. Sabe-se também que atualmente não existe nenhuma terapia voltada ao tratamento específico da febre Mayaro, sendo este, baseado no alívio da dor e febre com analgésicos (paracetamol) e/ou antiinflamatórios não esteroides (ibuprofeno ou diclofenaco) (Caetano, 2016). Sendo assim, se faz necessário o desenvolvimento de um antiviral contra o MAYV.

É notório que o processo de descoberta de fármacos demanda alto custo financeiro e temporal devido a sua grande complexidade e geralmente está ligado às inovações tecnológicas e científicas (Guido *et al.*, 2008; Guido *et al.*, 2010). O mesmo pode ser dividido em duas fases sendo a primeira a fase pré-clínica, com a etapa de pesquisa básica para o descobrimento do fármaco e a segunda a fase clínica, que seria o seu desenvolvimento (Guido *et al.*, 2010 *apud* Lombardino & Lowe, 2004).

Novas abordagens têm sido empregadas na tentativa de reduzir os custos e otimizar o tempo de desenvolvimento de um fármaco na fase pré-clínica, como é o caso de aplicações da bioinformática e quimioinformática. Além disso, uma das principais promessas seria o aumento da probabilidade de sucesso na fase clínica, visto que atualmente apenas 5% dos compostos testados chegam ao mercado (Domingues, 2013). Nesse contexto, destaca-se a triagem virtual, que é uma das técnicas computacionais mais utilizadas no planejamento racional de fármacos, com a busca *in silico* de compostos bioativos (Domingues, 2013). Ela pode ser definida como o processo de avaliação automática de grandes bibliotecas de compostos por métodos computacionais (Bajorath, 2002; Domingues, 2013). A triagem virtual compreende duas abordagens principais, sendo elas a triagem baseada na estrutura do alvo molecular e a baseada em ligantes (Rodrigues *et al.*, 2012; Domingues, 2013; Verli, 2014). A primeira delas é uma técnica mais robusta e considera que o alvo terapêutico (receptor) apresente a sua estrutura tridimensional (3D) resolvida, para que sejam efetuados os cálculos de ancoragem (*docking*) e seleção de possíveis ligantes. Estes que serão avaliados de acordo com características estruturais, eletrônicas e químicas que favorecem interações com o sítio ligante e o alvo (Rodrigues *et al.*, 2012). Já na triagem baseada em ligantes, os moldes são moléculas orgânicas com atividade biológica conhecida, e faz-se a busca de novas entidades químicas que tenham certo nível de similaridade para compartilharem de propriedades estruturais e também atividade biológica (Rodrigues *et al.*, 2012).

Os fármacos contemporâneos são planejados para atuar por meio da modulação seletiva da resposta biológica de proteínas alvo, aliando características de alta seletividade ou especificidade e propriedades farmacocinéticas adequadas (Rocha, 2009; Ferreira, 2015). Nesse sentido, a triagem baseada na estrutura do alvo molecular é mais empregada e permite a descoberta de ligantes com uma diversidade estrutural maior do que a outra abordagem. Entretanto, uma limitação da técnica é a dependência da disponibilidade da estrutura 3D de um alvo conhecido estar resolvida.

A estrutura 3D das proteínas do MAYV não está disponível em bancos de dados, limitando o uso da abordagem de triagem virtual baseada no alvo molecular. Atualmente, com os avanços na biologia estrutural a predição da estrutura 3D dessas macromoléculas pode ser feita utilizando-se métodos computacionais. De acordo com o legado deixado por Anfinsen, a sequência primária de aminoácidos de muitas proteínas contém informações suficientes para a predição da sua estrutura tridimensional (Anfinsen, 1962).

A modelagem de proteínas por homologia (ou modelagem comparativa) é uma das abordagens de predição estrutural computacional, sendo útil quando as proteínas não foram ou não podem ser resolvidas por métodos experimentais como cristalografia de raios X e ressonância magnética nuclear (RMN) (Filho, 2003). Nesse método, a proteína de interesse ou alvo, terá sua estrutura 3D construída, utilizando-se como referência a estrutura 3D de outra proteína similar (molde), que normalmente tem relação evolutiva com o alvo e tem suas coordenadas cartesianas depositadas no *Protein Data Bank* (PDB) (Verli, 2014). Isso é possível, pois acredita-se que sequências similares gerem estruturas terciárias similares que por sua vez está relacionada com a função biológica.

Estudo recente realizado pelo grupo de pesquisa do Professor Doutor José Carlos de Magalhães obteve sucesso na construção e refinamento da proteína do capsídeo (C) do MAYV bem como na busca de inibidores por triagem virtual e seleção de uma molécula proveniente desta etapa para ensaios *in vitro*. Seguindo esta linha, buscamos a obtenção de um modelo 3D via modelagem por homologia da proteína nsP3 MAYV para servir como receptor numa triagem virtual para busca de antivirais contra o MAYV.

2. OBJETIVO

2.1. Objetivo geral

Construir um modelo 3D através da modelagem por homologia da proteína nsP3 MAYV para ser utilizado como receptor numa triagem virtual de compostos em busca de antivirais contra o mesmo.

2.2. Objetivos específicos;

- Obter a sequência primária da proteína nsP3 em banco de sequências de proteínas;
- Realizar o alinhamento de sequências para a identificação de moldes;
- Construir e validar um modelo 3D;
- Promover a simulação por dinâmica molecular implícita, para refinar o modelo e avaliar a estabilidade da molécula ao longo do tempo;
- Busca por possíveis fármacos na literatura e moléculas em banco de dados;
- Realizar a ancoragem molecular e análise da energia de ligação.

3. METODOLOGIA

3.1. Recuperação da sequência primária nsP3

O primeiro passo da modelagem molecular por homologia é recuperar a sequência de aminoácidos de uma proteína. Deste modo, a sequência de aminoácidos da proteína nsP3 foi recuperada da base de dados de proteína do NCBI (<http://www.ncbi.nlm.nih.gov/protein>), no formato .FASTA.

3.2. Busca por moldes

Na busca inicial por moldes, o alinhamento de sequências tem um papel fundamental para identificar e informar o grau de similaridade de sequências de proteínas resolvidas experimentalmente com a sequência alvo (Verli, 2014; Souza, 2008). O alinhamento local faz uma busca de pequenas regiões da sequência alvo contra as sequências disponíveis no banco

de dados, avaliando a similaridade, identidade, *gaps* (lacunas) e a cobertura das sequências (Amaral *et al.*, 2007; Altschul *et al.*, 1990). Sendo assim, o alinhamento foi aplicado como um mecanismo preliminar de busca por sequências similares para a identificação de possíveis moldes que tenham suas estruturas resolvidas experimentalmente. Nesse sentido, utilizou-se o algoritmo de código aberto *Basic Local Alignment Search Tool (BLAST)* que está disponível em <http://blast.ncbi.nlm.nih.gov>, mantendo as configurações padrão.

Após a seleção da proteína a ser usada com molde, o modelo foi construído por modelagem por homologia.

3.3. Construção do modelo

A predição da estrutura 3D foi realizada com base nos alinhamentos molde-alvo, utilizando a técnica de modelagem por montagem de corpo rígido, na qual se constrói o modelo baseando-se na conservação das estruturas de proteínas homólogas ou com grau significativo de identidade (Vyas *et al.*, 2012). Para isso, a sequência FASTA das proteínas alvo selecionadas foram exportadas para *workspace* SWISS-MODEL (<https://swissmodel.expasy.org/>) a fim de se construir os modelos tridimensionais.

3.4. Validação do Modelo

A fim de avaliar a qualidade da estrutura construída algumas análises foram efetuadas, tais como: a determinação do desvio médio quadrático (RMSD) entre a nsP3 construída e o molde (4GUA.1A) por meio do *Swiss Pdb Viewer*; análise estereoquímica através do gráfico de Ramachandram gerado no RAMPAGE (<http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>); a qualidade das cadeias laterais pelo ANOLEA e por fim a qualidade global proteica pela função QMEAN.

Em bioinformática o RMSD é a medida da distância média entre os átomos, geralmente do esqueleto proteico, de proteínas sobrepostas (Lesk, 2008; De Carvalho, 2014). Esta análise permite calcular a semelhança entre as estruturas protéicas.

Para se estimar a qualidade absoluta de modelos teóricos, utilizou-se a função QMEAN (Benkert *et al.*, 2011), que está disponível na *workspace* do SWISS-MODEL (<http://swissmodel.expasy.org/workspace>). O QMEAN *Z-score* é uma função na qual se estima a proximidade do modelo com a sua “conformação nativa” (Benkert *et al.*, 2009). E para isso faz uma comparação com as estruturas resolvidas por Cristalografia de raio X em

um banco de dados. Para essas estruturas resolvidas experimentalmente, vale ressaltar que o *Z-score* para os termos da função QMEAN normalmente são próximos a 0 (Benkert *et al.*, 2011). Assim, valores negativos e próximos de 0 são considerados aceitáveis (De Carvalho, 2004).

O *software Atomic Non-Local Environment Assessment* (ANOLEA), disponível em http://swissmodel.expasy.org/workspace/?func=tools_structureassessment1, foi utilizado para avaliar a qualidade das cadeias laterais dos modelos quanto ao contato dos átomos e seu ambiente químico. O ANOLEA obtém essa estimativa por meio do cálculo da energia de conformação de cada resíduo da cadeia lateral da proteína (Melo, 1997). Quando o ambiente está favorável para os átomos o programa assume valores negativos de energia (caracterizadas pela cor verde) já quando o ambiente é desfavorável a energia é positiva e graficamente identificada pela cor vermelha (Junior *et al.*, 2012; Melo, 1997).

Após isso, o modelo então foi submetido ao cálculo por dinâmica molecular para refiná-lo usando o modelo de solvente implícito.

3.5. Simulação de Dinâmica Molecular

A dinâmica molecular consiste na descrição conformacional para uma determinada molécula ou conjunto de moléculas baseada na solução da equação de movimento de Newton, que é descrita abaixo:

$$F_{xi} = a_i m_i$$

Onde F_{xi} é a força aplicada ao átomo i na posição x , t é o tempo, v a velocidade e a_i a aceleração do átomo i .

As equações são resolvidas simultaneamente em todos os passos da dinâmica, enquanto a temperatura do sistema e a pressão são mantidas quase que constante, sendo permitidas somente pequenas flutuações em torno de um valor determinado. As coordenadas dos átomos são escritas em intervalos regulares, e dadas em função do tempo, representando a trajetória do sistema. A análise da trajetória também permitiu obter muitas propriedades macroscópicas de interesse tais como, pressão, energia interna, volume, temperatura, entropia e energia livre.

Inicialmente, o modelo individual foi submetido ao cálculo de minimização de energia, que consiste em reduzir a energia do sistema, alcançando por fim uma conformação mais estável. Para esse cálculo, são empregadas funções de energia potencial (ou campos de

força) da mecânica clássica (Cornell *et al.*, 1995). Em nosso estudo a minimização de energia foi realizada sob gradiente conjugado usando o programa NAMD (Phillips *et al.*, 2005) e o campo de força CHARMM36, que considera todos os átomos explicitamente (*all-atom*). A minimização foi executada em 2500 passos, utilizando o modelo de solvatação implícita de Born generalizado, que se trata de um modelo em que se eliminam os átomos de água incluindo os efeitos do solvente no cálculo das forças inter-atômicas. A partir da estrutura minimizada foi realizada a simulação de dinâmica molecular (DM) a temperatura constante de 310 K para o tempo de simulação de 20 ns. Os resultados da DM foram analisados usando o programa VMD (Humphrey *et al.*, 1996).

3.6. Triagem virtual de compostos ligantes da nsP3

O *docking* molecular é uma técnica computacional que prevê a melhor orientação e detalhes sobre o reconhecimento molecular e afinidade de ligação entre receptor-ligante (Seydi *et al.*, 2016; Verli, 2014). Ele pode ser dividido em dois estágios. O primeiro se trata da investigação e predição da conformação e orientação de uma molécula ligante no sítio de complexação. O segundo da classificação da afinidade das ligações através de uma função *scoring*, que determina a energia de ligação (Seydi *et al.*, 2016; Verli, 2014).

Vale ressaltar que a estrutura utilizada como receptor no *docking* molecular, foi a construída e minimizada antes de ser empregada na dinâmica molecular. Devido a possibilidade da dinâmica molecular alterar irreversivelmente conformação da molécula, normalmente o modelo que passa por ela não é utilizado nas etapas posteriores.

O *docking* molecular foi conduzido utilizando-se o *software* AutoDock Vina. No presente estudo considerou-se a proteína como um corpo rígido ao passo que os ligantes foram totalmente flexíveis. A estratégia de *docking* envolveu duas etapas sendo elas: preparação do receptor e ligantes, e ancoragem molecular.

3.6.1 Preparação do receptor e ligantes

Para a realização da ancoragem molecular realiza-se a preparação das estruturas e arquivos. Devido ao fato do AutoDockVina requerer inputs no formato “pdbqt”, todas as moléculas foram inicialmente convertidas para tal formato. Os ligantes provenientes do banco de dados ZINC (<http://zinc.docking.org/>) foram salvos no formato SMILES “.smi”. O Marvin Sketch versão 15.11.2 foi então utilizado para converter cada uma das moléculas do

formato “.smi” para o “.pdb” além de calcular sua carga para pH fisiológico e identificar possíveis isômeros.

Os ligantes provenientes da literatura foram desenhados no Marvin Sketch versão 15.11.2.tendo sua carga calculada para pH fisiológico e suas formas isoméricas também salvas no formato “.pdb”.Finalmente, as moléculas em “.pdb” foram convertidas para “.pbqt” por meio do AutoDock Tools.

3.6.2 Ancoragem molecular

Os arquivos de entrada para o Autodock Vina foram preparados utilizando-se o Autodock Tools v.15.6. A fim de se definir o sítio de ligação, foi construída uma caixa ao redor do ligante cristalográfico, 2-(*N*-morpholino)-etano sulfônico(MES), que se encontra no sítio de ligação da proteína e definida como um cubo com dimensões de 20 Å x 20 Å x 20 Å , espaçamento de 1 Å e coordenadas x, y e z de 88.217 Å, 82.622 Å e 68.96 Å respectivamente. Para validar o experimento efetuou-se a metodologia *re-docking*, o ligante cristalográfico que estava presente no molde foi transferido para a proteína nsP3 de MAYV mantendo-se as mesmas coordenadas.Por fim, procedeu-se o *docking* com os ligantes preparados anteriormente.

4. RESULTADOS E DISCUSSÃO

4.1. Recuperação da sequência nsP3

A sequência da proteína nsP3 de MAYV composta de 485 aminoácidos foi obtida do NCBI, com identificação NP_740689.1 e está apresentada a seguir em formato .FASTA como mostra o **Quadro 1**.

Quadro 1. Sequência em formato .FASTA da proteína nsP3 de MAYV.

```

>gi|25121715|ref|NP_740689.1|    non    structural    protein    nsP3    [Mayarovirus]
APAYAVKRADIATAIEDAVVNAANHRGQVGDGVCRAVARKWPQAFRNAATPVGTA
KTVKCDEYIIHAVGPNFNNTSEAEGDRDLAAAYRAVAAEINRLSISSVAIPLLSTGIFS
AGKDRVHQSLSHLLAAMDTTEARVTIYCRDKTWEQKIKTVLQNRSA TELVSDELQFE
VNLTRVHPDSSLVGRPGYSTTDGTLYSMEGTFKHQAALDMAEITTLWPRVQDANEHI
CLYALGETMDNIRARCPVEDSDSSTPPKTVPCLCRYAMTPERVTRLRMHHTKDFVVC
SSFQLPKYRIPGVQRVKCEKVMLFDAAPPASVSPVQYL TNQSETTISLSSFSITS DSSSL
STFPDLESAEELDHDSSQSVRPALNEPDDHQPTPTAELATHPVPPPRPNRARRLAAARV
QVQVEVHQPPSNQPTKPIAPRTSLRPVPAPRRYVPRPVVELPWPLETIDVEFGAPTEE
ESDITFGDFSASEWETISNSS

```

Disponível em: <https://www.ncbi.nlm.nih.gov/protein/25121715?report=fasta>.

4.2. Alinhamento Local

Os moldes encontrados por meio do alinhamento local usando o BLAST com a sequência primária da proteína nsP3 de MAYV é apresentada na Tabela 2.

Tabela 2. Informações estatísticas fornecidas pelo *BLASTp* tendo como sequência de busca (*query*) a NP_740689.1.

Acesso	Identidade	Score	Query Cover	e-value
4GUA_A	54%	358	67%	1e ⁻¹¹⁶
3GPG_A	66%	225	32%	1e ⁻⁷¹
3GPO_A	65%	222	32%	1e ⁻⁷⁰
3GQE_A	56%	183	32%	1e ⁻⁵⁵
2VRI_A	28%	40	29%	0.002
3VFQ_A	27%	35	22%	0.14
3Q6Z_A	27%	34.3	22%	0.17
3EWO_A	28%	32.7	20%	0.40
3EJF_A	27%	32.7	26%	0.42

Fonte: NCBI.

Na modelagem por homologia, presume-se que o êxito da técnica é dependente do grau de similaridade entre a sequência primária e o molde (Vygis *et al.*, 2012).

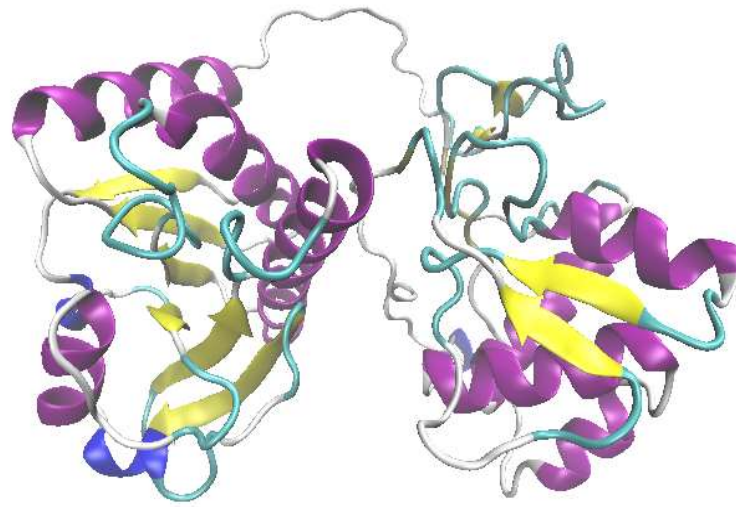
A análise do alinhamento sob o ponto de vista estatístico é muito importante para se inferir homologia entre as sequências (Verli, 2014). Devido à grande variedade de sequências proteicas candidatas a molde, se faz necessário escolher a que for melhor, para assim, gerar

um modelo da proteína alvo com a melhor qualidade possível. Há uma divergência na literatura a respeito da identidade mínima necessária para a geração de um bom modelo, sendo aceitáveis limites superiores a 25-30% (Verli, 2014; Moulton, 2005). Outro parâmetro relevante é o *e-value*, que de modo geral indica a chance de ocorrer um alinhamento da sequência alvo com uma aleatória do banco de dados. Acredita-se que quanto menor o *e-value* mais significativo estatisticamente é o modelo (Verli, 2014). Segundo Pearson (2013) para inferir homologia entre as sequências de aminoácidos o *score* deve ser maior que 50 e o *e-value* menor que 10^{-7} . Desse modo, o melhor molde para a elaboração do modelo da proteína nsP3 MAYV seria o 4GUA_A, pois além de estar de acordo com Pearson, 2013 o mesmo foi o único molde que apresentou um ligante cristalográfico, sendo este necessário para a delimitação de um sítio ligação de importância funcional na etapa de *docking molecular*. Além disso, a presença do ligante é importante, pois a conformação do receptor é alterada pelo mesmo.

4.3 Construção do modelo e seleção dos moldes

Ao inserir a sequência alvo da proteína nsP3 de MAYV o algoritmo do SWISS-MODEL efetuou de forma automática a construção dos possíveis modelos baseando-se na construção por método de corpos rígidos usando moldes encontrados em sua base de dados. Ao todo foram encontrados 35 moldes (*templates*), entretanto somente o modelo construído a partir do *template* 4GUA.1A (Chain A, Alphavirus P23 pro-Zbd), **Figura 3.b**, de *Sindbis virus* (SINV) foi selecionado como mencionado no item 4.2 . O modelo construído consistiu de uma estrutura de 11 alfa-hélices e 10 folhas beta como apresentado na **Figura 3.a**.

(a)



(b)

Figura 3. Modelo e molde da proteína nsP3 **(a)** Modelo construído no SWISS-MODEL da proteína nsP3 do MAYV (NP_740689) usando como molde **(b)** a poliproteína 4GUA.1A ((Chain A, Alphavirus P23 pro-Zbd) de SINV. A visualização de **(a)** foi efetuada no programa VMD, modo *new cartoon* e **(b)** foi obtida no PDB.

4.4. Validação do modelo construído

O desvio quadrático médio (RMSD) calculado no *SwissPdb Viewer*, entre as proteínas molde e modelo foi 0,64 Å. Este valor está dentro do aceitável, pois de acordo com Chothia &

Lesk (1986), o valor de referência encontrado entre proteínas estruturalmente similares é menor 2,31 Å.

O modelo também foi validado quanto à qualidade estereoquímica da cadeia principal, por meio do gráfico de Ramachandram, que mostra a relação percentual dos resíduos de aminoácidos em regiões consideradas como favoráveis, permitidas e proibidas. Acredita-se que para um modelo apresentar boa qualidade ele deve ter no mínimo 90% dos aminoácidos nas regiões favoráveis (Junior *et al.*, 2012). O modelo construído apresentou 97,5% dos resíduos de aminoácidos estão na região favorável e 2,5% na região permitida e 0% na região proibida, conforme mostrado na **Figura 4**, indicando que a cadeia principal do modelo apresenta boa qualidade estereoquímica.

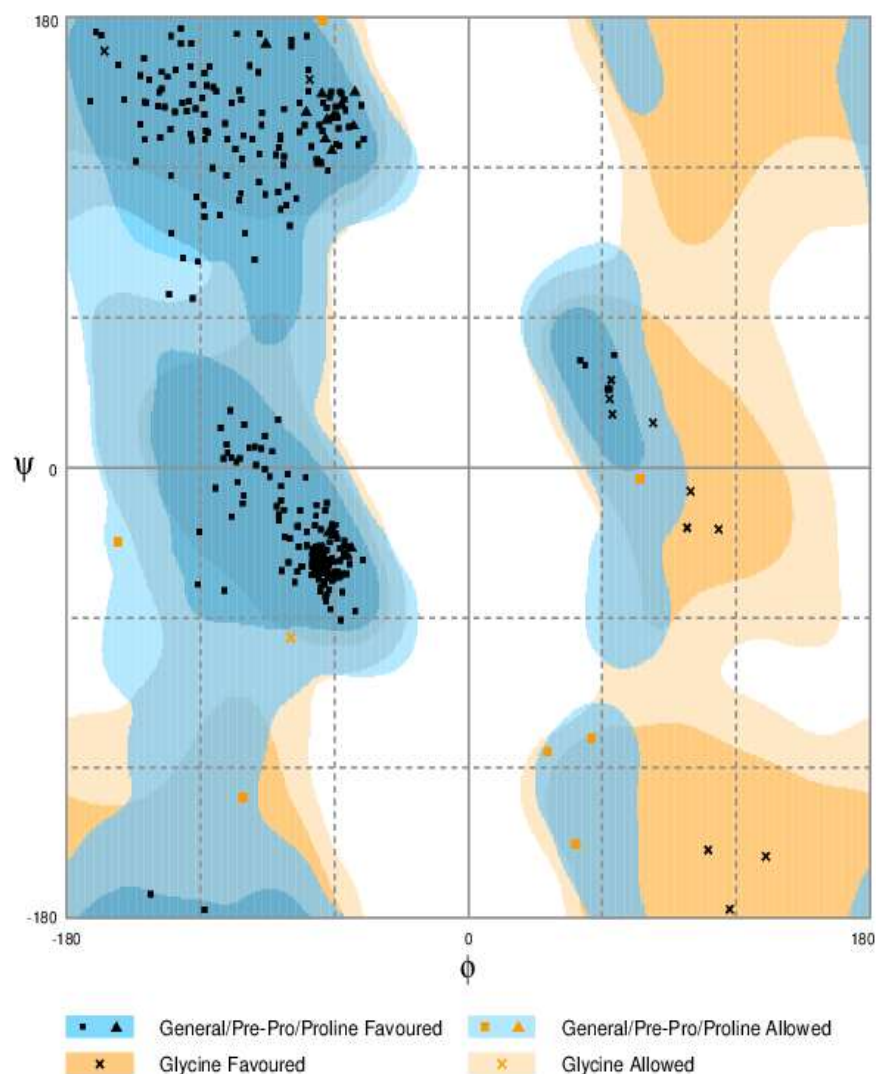


Figura 4. Gráfico Ramachandram do modelo da proteína nsP3 do MAYV. O gráfico mostra que 97,5% dos resíduos de aminoácidos estão na região favorável e 2,5% na região permitida e 0% na região proibida.

A análise de energias ao longo da sequência, feita no ANOLEA, mostrou que o modelo apresentou mais regiões de baixa energia (barras verdes) do que regiões de alta energia (barras vermelhas), como ilustrado na **Figura 5**. O que indica que o modelo possui baixa energia sendo mais um indício que se assemelha a condições reais encontradas para moléculas.

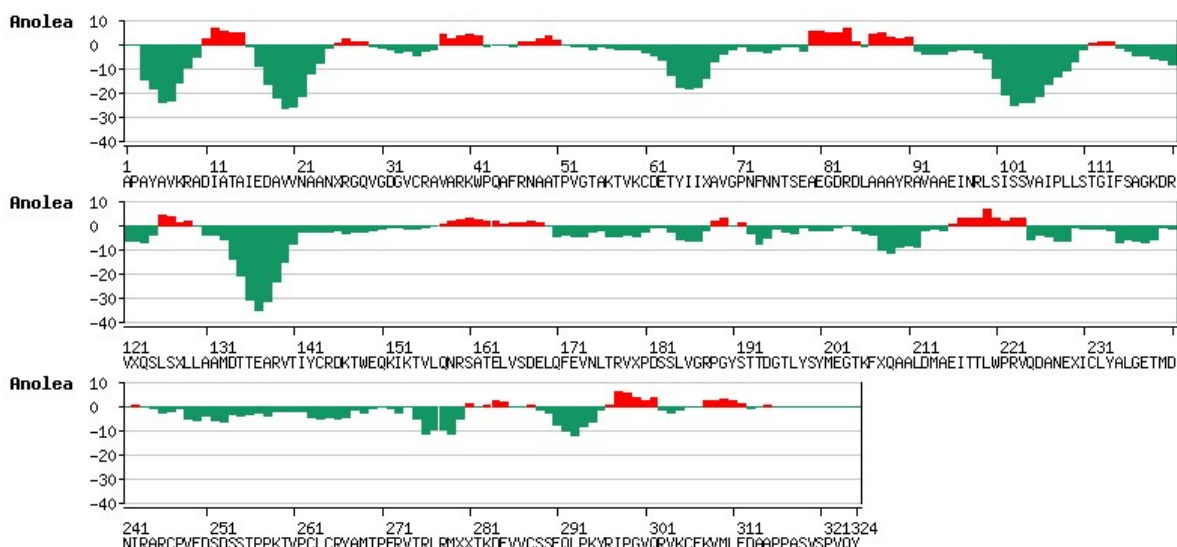


Figura 5. Representação gráfica da análise de energia ao longo da sequência de aminoácidos do modelo da nsP3 de MAYV, gerado no ANOLEA. As barras verdes indicam que as energias estão menores que zero ao passo que as vermelhas são utilizadas para energias superiores a zero.

Adicionalmente, a função de pontuação QMEAN4 (**Figura 6**) também indica a boa qualidade do modelo, uma vez que o valor encontrado para o modelo foi -1.52 , próximo a $Z=0$, estabelecido para estruturas cristalográficas de alta resolução (Benkert *et al.*, 2010).

Na **Figura 7**, é mostrado um gráfico relacionando o QMEAN4 normalizado e o tamanho dos resíduos protéicos mostrando que o Z -score está entre 1 e 2 e os tamanho resíduos estão entre 300 e 400. Os valores normalizados representam a média da “população” comparadas com estruturas de alta resolução cristalografadas. Quanto mais próximo da região cinza escuro melhor a qualidade do modelo.

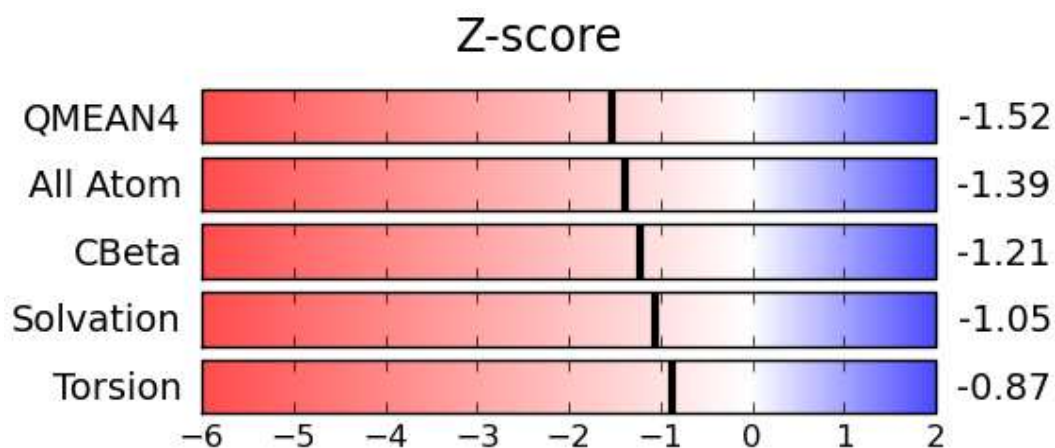


Figura 6. Análise de qualidade, representação gráfica da função *Z-score*QMEAN4gerado no SWISS-MODEL.Mostrando que todos os valores foram negativos e próximos de zero, indicando o quão próximo o modelo está de estruturas resolvidas experimentalmente.

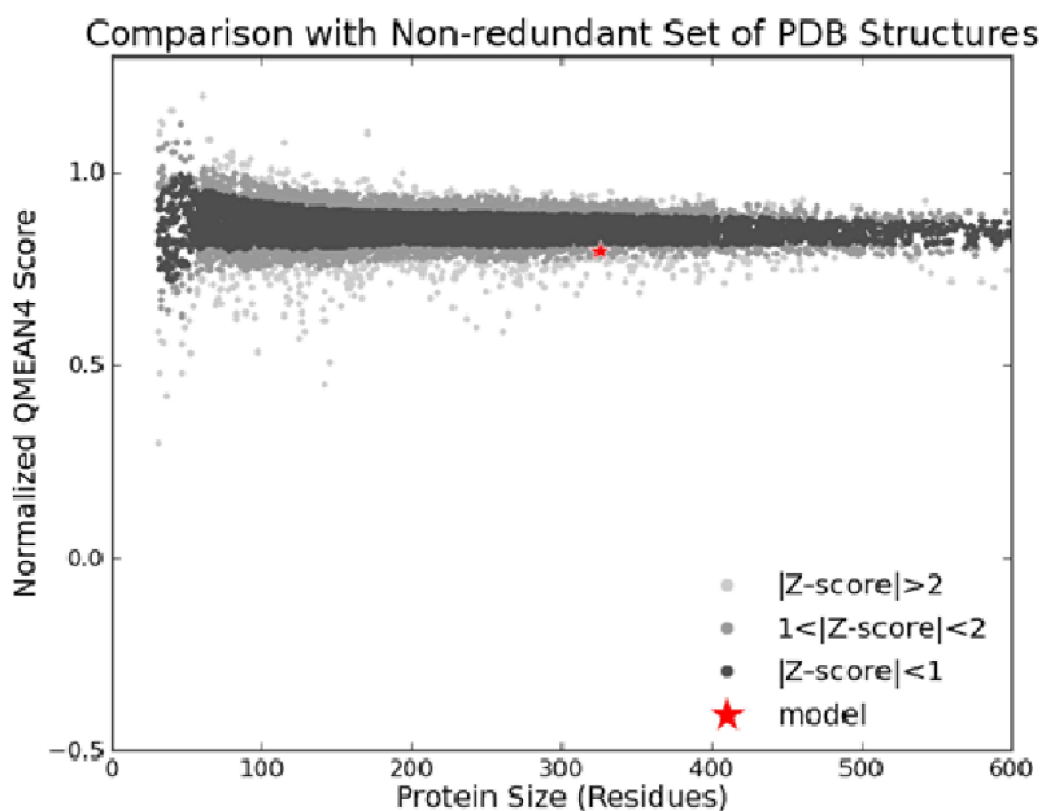


Figura 7. Representação gráfica do QMEAN4 normalizado gerado no portal SWISS-MODE. O *Z-score* encontra-se entre 1 e 2.

4.5. Simulação de dinâmica molecular

Após o cálculo por DM, a análise da trajetória foi realizada usando o RMSD para verificar a estabilidade estrutural durante o tempo de simulação de 20 ns, conforme a **Figura 8**.

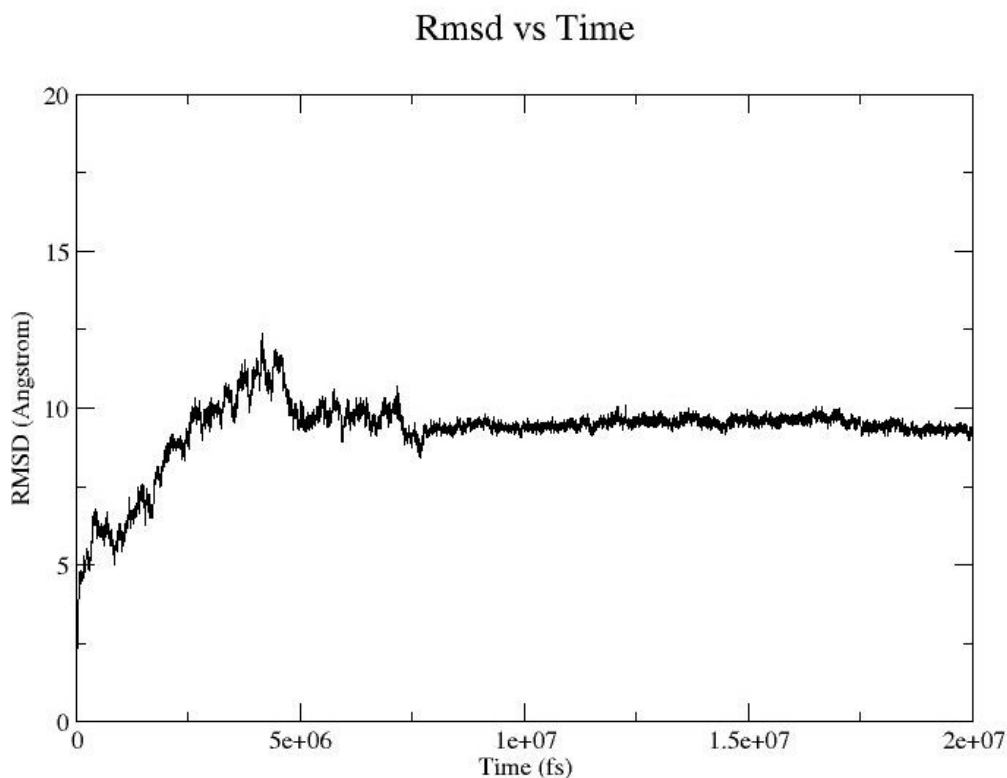


Figura 8. Curva de de RMSD (Å) *versus* tempo (fs) durante a simulação por DM da nsP3 MAYV durante 20ns.

Nota-se que o RMSD teve um pico máximo em 12,5 Å, entretanto a estrutura se estabilizou a partir de 7,5 ns apresentando RMSD equivalente a 9,0 Å. Apesar do módulo do RMSD estar elevado (<2 Å seria o ideal, de acordo com Brozell *et al.*,2012), o comportamento de estabilização da estrutura ao longo do tempo é um ponto favorável. Este valor alto pode estar relacionado à ausência de moléculas de água, responsáveis por estabilizar a estrutura proteica em um sistema biológico. Nesse sentido faz-se necessário a simulação da dinâmica incluindo moléculas de água envolvidas no sistema para averiguar o comportamento da proteína nessas condições e se realmente esse fator contribuiu para a elevação do RMSD. Apesar disso, os resultados indicam que o modelo da nsP3 MAYV construída e refinada pode ser utilizada como receptor na etapa de triagem virtual.

4.6. Ancoragem molecular

Para o processo de *docking* sítio-dirigido é necessário se conhecer um sítio de ligação, que normalmente é localizado pelas coordenadas do ligante cristalográfico do molde. Nesse trabalho, o sítio de ligação foi definido pelas coordenadas do ácido 2-(*N*-morpholino)-etano sulfônico(MES) no molde 4GUA_A. Para validar esta metodologia o *re-docking* foi efetuado, o mesmo consiste em retirar o ligante cristalográfico (MES) da estrutura do molde e ancorá-lo ao receptor, sendo posteriormente avaliada a sobreposição do ligante cristalográfico e o ancorado. A sobreposição forneceu um valor de RMSD de 1,16 Å que está dentro do aceitável (Brozell *et al.*, 2012). A **Figura 9** mostra o complexo nsP3 MAYV-MÊS obtido pela transferência das coordenadas do ligante do molde para o modelo e a validação pelo *redocking*.

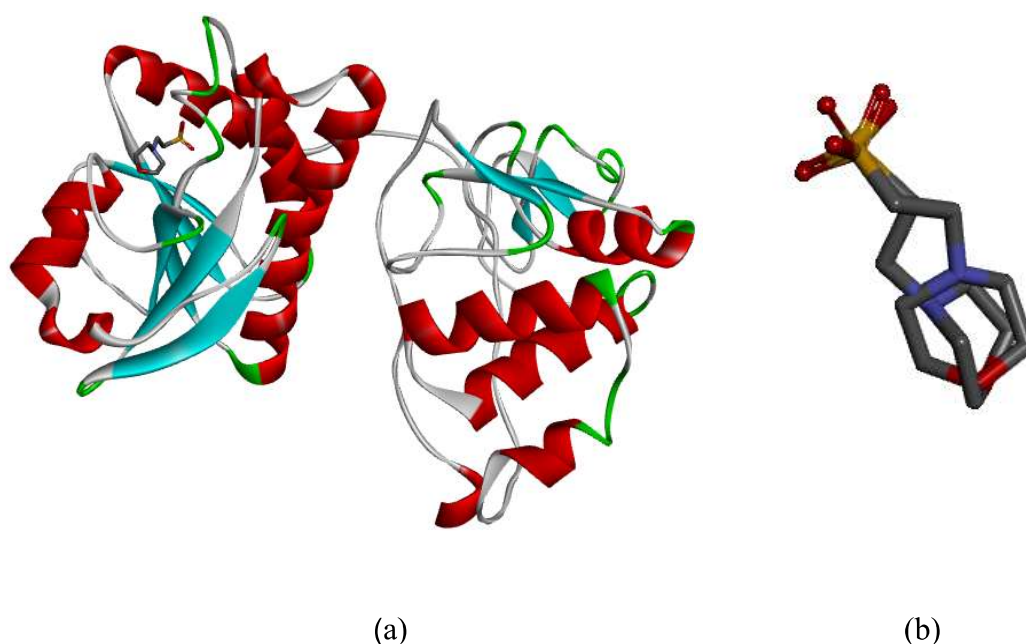


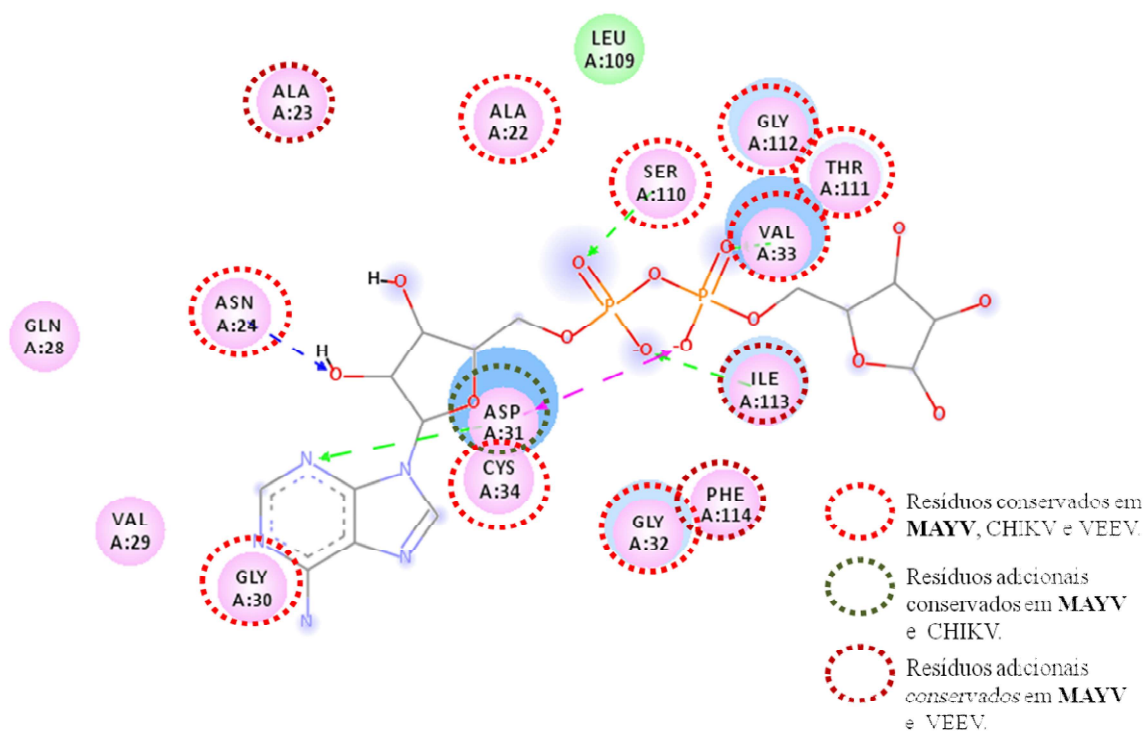
Figura 9.Complexo nsP3 MAYV-MES em (a) e em (b) a sobreposição do MES cristalográfico no processo de *re-docking* como validação do procedimento de *docking*. Elaborado no *DS Visualizer*.

Estudos indicam que o sítio de ligação do MES é o mesmo ocupado pelas moléculas de Adenosina Difosfato Ribose (ADP-ribose) (Karras *et al.*, 2005; Malet *et al.*, 2009; Rungrotmongkol *et al.*, 2010; Shinet *et al.*, 2012; Seyedi *et al.*, 2016). Para verificar esta hipótese, fez-se o *docking* com o ligante cristalográfico e com a ADP-Ribose. As melhores

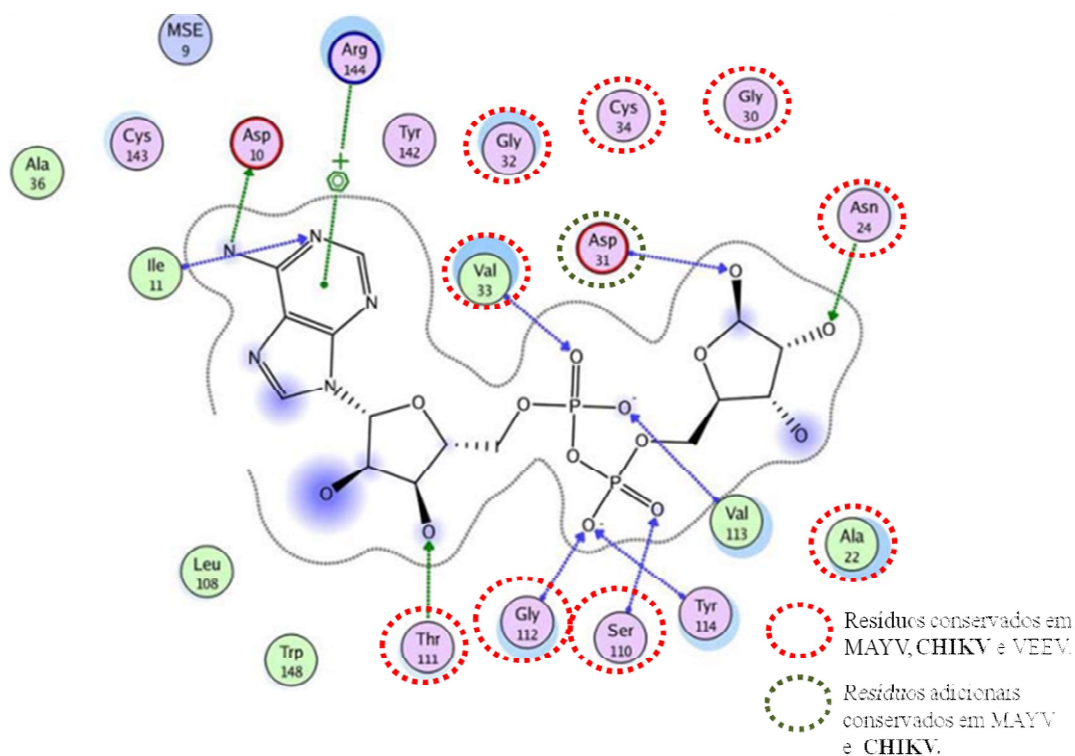
configurações do arquivo de saída mostraram uma energia de -5,7 kcal/mol para o MES e como esperado uma energia ainda menor para a ADP-Ribose, equivalente a -8,7 Kcal/mol.

Recentemente Seyedi e colaboradores (2016) submeteram a nsP3 do CHIKV ao *docking* molecular com a ADP-ribose, e obtiveram a mesma energia de ligação encontrada nesse trabalho (8,7 Kcal/mol para e melhor configuração encontrada). Além disso, um estudo realizado por Rungrotmongkol e colaboradores (2010), verificou por estudos de dinâmica molecular que o macrodomínio da nsP3 de CHIKV e VEEV ao qual se liga a ADP-Ribose é estável e possui resíduos de aminoácidos conservados para a nsP3 dos dois vírus. Visando verificar a conservação desse macrodomínio nos Alphavirus e sustentar a hipótese da interação com a ADP-Ribose, em nosso estudo, foi verificado se o complexo (nsP3-MAYV)-(ADP-Ribose) também compartilha de interações com aminoácidos semelhantes.

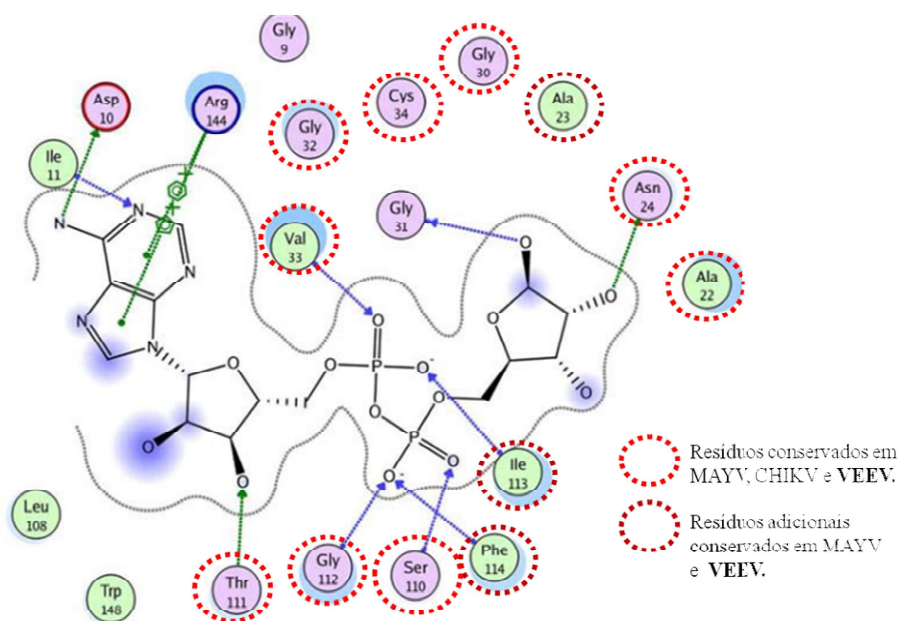
No geral, a ADP-Ribose complexada a nsP3 de MAYV, de CHIKV e de VEEV fez 16, 21 e 19 interações, respectivamente. E dos resíduos da nsP3 de MAYV, nove (56%) foram iguais à proteína dos outros vírus, conservando até mesmo as posições dos aminoácidos: Ala22, Asn24, Cys34, Gly30, Gly32, Val33, Thr111, Ser110 e Gly112. As interações entre a ADP-ribose e a proteína nsP3 de MAYV, CHIV e VEEV bem como as interações que se conservaram são mostradas na Figura 10.a,b,c.



(a)



(b)



(c)

Figura 10. Interações intermoleculares da ADP-Ribose com a proteína npP3 modelada de MAYV (a), CHIKV (b) Adaptado de Rungtongmongkol et al., 2010 e VEEV Adaptado de Rungtongmongkol et al., 2010 (c). mostram as interações da ADP-Ribose com a proteína nsP3 de CHIKV e VEEV respectivamente, adaptados de Rungtongmongkol et al., 2010. Legenda: Círculos em rosa-resíduos envolvidos em ligações de hidrogênio ou interações apolares; Círculos em verde-resíduos envolvidos em interações de van der Waals; Halo azul em torno do resíduo-interação com o solvente (água); Halo azul em torno do átomo – interação com o solvente (água). Seta tracejada em verde-ligação de hidrogênio com a cadeia principal. Seta tracejada em azul-ligação de hidrogênio com a cadeia lateral; Linha laranja- Interações Pi.

A **Figura 10.a** mostra que das 16 interações feitas pelo complexo nsP3 MAYV-ADP-ribose a maioria (93%) foi determinada por interações apolares ou ligações de hidrogênio (círculos rosa). Somente a Leu109 fez interações do tipo van der Waals (círculo verde). Já os aminoácidos que também fazem interação com a água (halo azul em torno do resíduo-interação) são Asp31, Gly32, Val33, Gly112 e Ile11. E 62,5% dos resíduos que desse complexo que se conservaram com a interação da nsP3 de CHIKV com a ADP-ribose, que são: Ala22, Asn24, Cys34, Gly32, Gly30, Val33, Thr111, Ser110 e Gly112 e Asp31. Já para a comparação entre a nsP3 MAYV-ADP-ribose com nsP3 VEEV-ADP-ribose esse índice foi maior, 68,75%, sendo os resíduos conservados: Ala22, Asn24, Cys34, Gly32, Gly30, Val33, Thr111, Ser110, Ile113 e Phe114.

Esses resultados sugerem a conservação desse sítio nos alphavirus avaliados, MAYV, CHIKV e VEEV. Uma vez que a ADP-ribose se liga favoravelmente a nsP3, sua utilização como um precursor de antivirais é interessante, dado que o antiviral desenvolvido terá um amplo espectro de atuação, podendo também atuar contra outros vírus, como CHIKV e VEEV.

A fim de se encontrar novas moléculas que podem se ligar com uma afinidade ainda maior ao complexo nsP3 de MAYV, foi realizada uma busca de análogos da ADP-Ribose e do MES no ZINC. Além disso, algumas moléculas que apresentaram atividade antiviral para outros alphavirus de acordo com a literatura, foram selecionadas sendo realizada também uma nova busca de análogos pelo ZINC para as moléculas consideradas promissoras. Ao todo 296 moléculas foram submetidas individualmente ao processo de *docking*, das quais 89 foram análogas da ADP-ribose, 100 do MES, 28 da literatura e partindo de algumas da literatura, 46 da 4-O-metilepigalocatequina, 12 da cloroquina e 21 da molécula (2S,3S,4S,5S,6S)-4-[[[(2S,3R,4S,6R)-3,4-dihydroxy-6-methyloxan-2-yl]oxy]-6-[2-(3,4-dioxocyclohexyl)ethoxy]-5-hydroxy-2(hydroxymethyl)oxan-3-yl]-3-(3,4-dioxocyclohexyl)propanoate de Teixeira *et. al.* (2014), nomeada como derivado fenólico. Os resultados são expostos em tabela no **ANEXO I**.

De maneira geral, como mostrado na **Figura 11**, a maioria das moléculas (77%) obteve uma energia de ligação superior ao padrão estipulado, entretanto pode-se considerar que a triagem foi satisfatória devido, ainda assim, terem sido encontradas moléculas com energia inferior a -8.7 Kcal/mol.

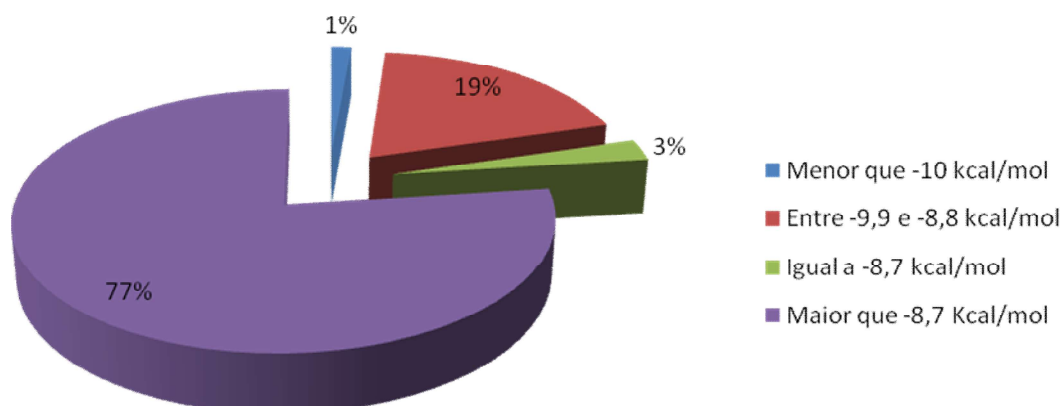


Figura 11. Distribuição percentual da energia de ligação (Kcal/mol) das 296 moléculas submetidas ao *docking*, tendo como receptor a nsP3 MAYV e parâmetro energético o valor -8,7 kcal/mol (ADP-ribose).

Uma vez que esta molécula se liga ao mesmo sítio da ADP-ribose ela pode atuar como um inibidor competitivo e assim, apresentar atividade antiviral. Deste modo, cerca de 20% das moléculas podem ser considerados como potenciais antivirais, uma vez que quanto menor a energia de ligação maior a afinidade pelo complexo molécula-nsP3.

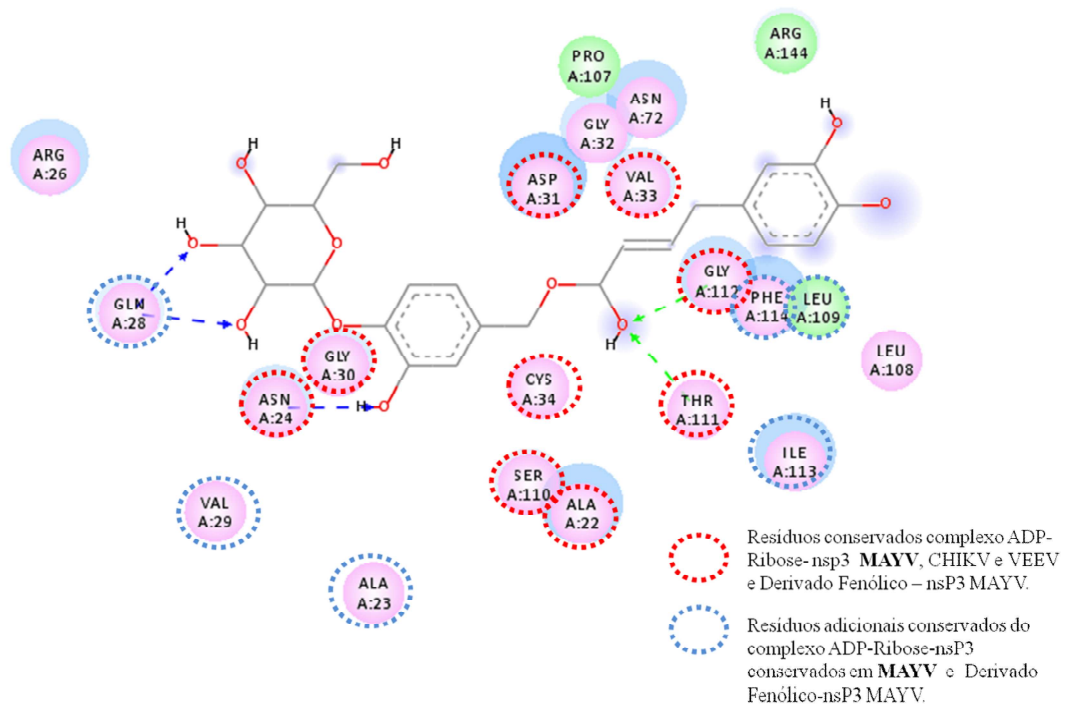
As menores energias encontradas foram provenientes das moléculas obtidas do ZINC zinc_8442273 e zinc_8442275, com energias -11,3e -11,5 kcal/mol, respectivamente. Ambas as moléculas são similares a molécula Derivado Fenólico (-10,1 kcal/mol), Também foi verificado que estas moléculas compartilham de interações com praticamente os mesmos aminoácidos do complexo nsP3-MAYV-ADP-ribose. A **Figura 12.a,b,c** ilustra as interações feitas pelas moléculas citadas acima no sítio de ligação da ADP-Ribose em nsP3-MAYV.

Nota-se Na **Figura 12.a** que a molécula Derivado Fenólico fez 21 interações (Ala22, Ala23, Asn24, Arg26, Gln28, Val29, Gly30, Asp31, Gly32, Val33, Cys34, Asn72, Pro107, Leu109, Leu108, Ser110, Thr111, Gly112, Ile113, Phe114, Arg144), das quais 15 (93,75%) foram com os mesmos aminoácidos do complexo nsP3- MAYV-ADP-ribose (que faz ao todo 16 interações, o único aminoácido desse complexo que não esteve presente foi a Leu108). E 9 (42,85%) são interações conservadas entre o complexo formado pela ADP-ribose e nsP3 de MAYV, CHIKV e VEEV, sendo os aminoácidos e respectivas posições: Ala22, Asn24, Gly30, Asp31, Val33, Cys34, Ser110, Thr111 e Gly112. Tendo em vista a conservação dos resíduos de aminoácidos para a nsP3 de diferentes vírus este pode ser um

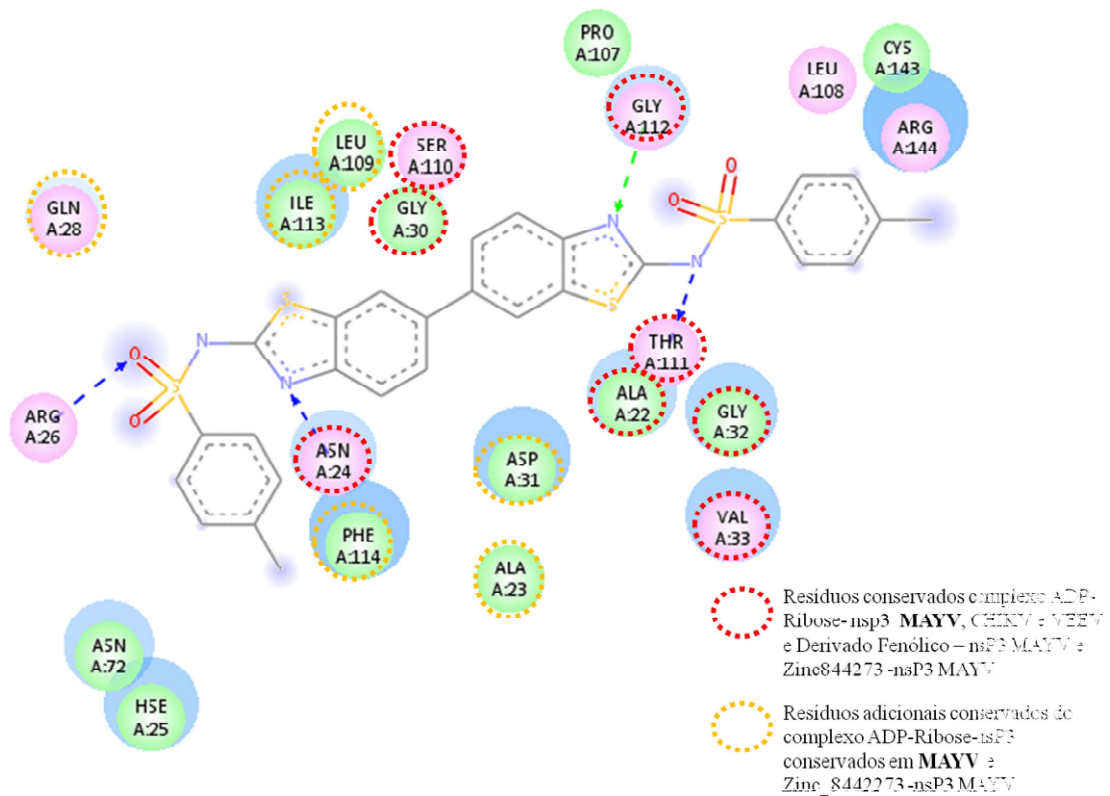
antiviral eficiente para esses outros alphavírus. A maioria (85,71%) das interações do complexo Derivado Fenólico- nsP3 MAYV foi proveniente de ligações de hidrogênio ou interações apolares e o restante por interações de van der Waals (Pro107, Leu109 e Arg 144). Além disso, cerca de 57% dos resíduos também interagem com a água, sendo eles Ala22, Asn24, Arg26, Gly30, Asp31, Gly32, Val33, Asn72, Gly112, Ile113, Phe114, e Arg144.

Já a molécula Zinc_8442273 (**Figura 12.b**) realizou 21 interações (Ala22, Ala23, Asn24, Arg26, Hse25, Gly28, Gly30, Asp31, Gly32, Val33, Asn72, Pro107, Leu108, Leu109, Ser110, Thr111, Gly112, Ile113, Phe114, Cys143 e Arg144), das quais 14 (66,66%) resíduos de aminoácidos se conservaram com o complexo nsP3 MAYV- ADP-ribose, sendo eles Ala22, Ala23, Asn24, Gln28, Gly30, Asp31, Gly32, Val33, Leu109, Ser110, Thr111, Gly112, Ile113 e Phe114. E 8 (38,09%) resíduos se conservaram para nsP3 de MAYV, VEEV e CHIKV (círculo pontilhado vermelho), sendo eles Ala22, Asn24, Gly30, Gly32, Val33, Ser110, Thr11 e Gly112. O tipo de interação predominante (52,38%) foram as ligações de van der Waals (halo verde) e o restante as ligações de hidrogênio ou interações apolares (halo rosa). E 52,31% dos resíduos de aminoácidos também interagiram com a água.

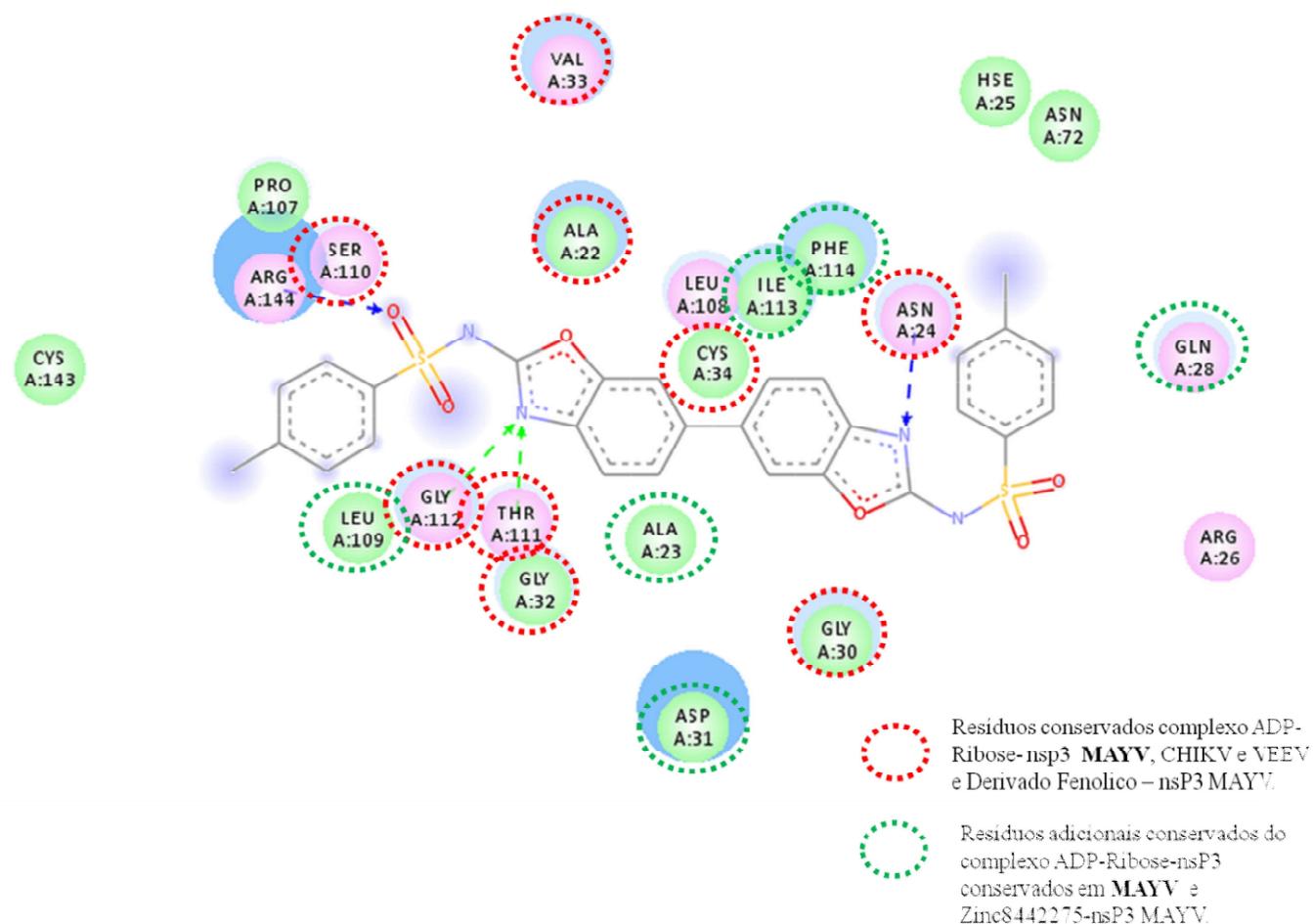
Por fim, a molécula Zinc_8442275 (**Figura 14.c**), realizou 22 interações (Ala22, Ala23, Asn24, Hse25, Arg26, Gln28, Gly30, Asp31, Val33, Cys43, Asn72, Pro107, Leu108, Leu109, Ser110, Thr111, Gly112, Ile113, Phe114, Cys143 e Arg144) com a nsP3 MAYV, sendo 15 (68,18%) delas foram conservadas ao fazer uma comparação com o complexo ADP-ribose-nsP3 MAYV (somente o resíduo de aminoácido Val29 desse complexo não interagiu com Zinc_8442275-nsP3 MAYV). Ao todo 9 (40,9%) interações se conservaram com os complexos ADP-ribose-nsP3 MAYV, CHIKV e VEEV, sendo os resíduos de aminoácidos conservados Ala22, Asn24, Gly30, Gly32, Val33, Cys34, Ser110, Thr111 e Gly112. A maioria (54,54%) das interações foram do tipo van der Waals (halo verde) e o restante ligações de hidrogênio ou interações apolares (halo rosa) ao passo que 54,54% desses resíduos também interagiram com o solvente (água).



(a)



(b)



(c)

Figura 12. Interações interatômicas dos aminoácidos da proteína nsP3 MAYV com os ligantes promissores provenientes do *docking*. Em (a) a interação com a molécula Derivado Fenólico ((2S,3S,4S,5S,6S)-4-[[[(2S,3R,4S,6R)-3,4-dihydroxy-6-methyloxan-2-yl]oxy]-6-[2-(3,4-dioxocyclohexyl)ethoxy]-5-hydroxy-2(hydroxymethyl)oxan-3-yl]-3-(3,4-dioxocyclohexyl)propanoate), em (b) com a molécula molécula Zinc_8442273 (6,6'-bis(2-[[[(4-methylphenyl)sulfonyl]amino]-1,3-benzothiazole]) e em (c) do com a molécula Zinc_8442275 (6,6'-bis{N-[(4-chlorophenyl)sulfonyl]-1,3-benzoxazol-2-amine}), que apresentou a menor energia de ligação, -11,5 Kcal/mol. Legenda: Círculos em rosa-resíduos envolvidos em ligações de hidrogênio ou interações apolares; Círculos em verde-resíduos envolvidos em interações de van der Waals; Halo azul em torno do resíduo-interação com o solvente (água); Halo azul em torno do átomo – interação com o solvente (água). Seta tracejada em verde-ligação de hidrogênio com a cadeia principal. Seta tracejada em azul-ligação de hidrogênio com a cadeia lateral; Linha laranja- Interações Pi.

5. CONSIDERAÇÕES FINAIS E PERSPECTIVAS

Acredita-se que a proteína nsP3 MAYV construída a partir da proteína NP_740689.1 foi bem modelada devido aos resultados apresentados pela análise de qualidade e dinâmica molecular. Embora o RMSD da dinâmica molecular tenha sido elevado, pôde ser demonstrado que a estrutura se estabilizou sem alteração conformacional significativa ao longo do tempo diminuiu.

De maneira geral, a triagem virtual apresentou bons resultados, pois para cerca de 20% das moléculas encontradas, obteve-se energia de ligação igual ou inferior ao padrão estipulado ADP-ribose (-8,7 Kcal/mol). A molécula zinc_8442275, obteve maior afinidade com a proteína nsP3 de MAYV apresentando uma energia de -11,5 Kcal/mol.

A ancoragem da ADP-ribose foi condizente com a literatura, sustentando a hipótese de que o macrodomínio possui o sítio de ligação dessa molécula sendo conservado na nsP3 modelada. Além disso, foi verificado que a interação de 56% dos aminoácidos do complexo nsP3-MAYV-ADP-ribose foram conservadas para esta proteína de outros alphavirus, o que pode ampliar o espectro de atuação das moléculas candidatas a antiviral.

Em relação a nsP3 como alvo molecular, ainda não se sabe se ela é essencial para bloquear o processo de multiplicação viral. Sendo assim, apesar dos resultados aparentemente serem promissores, testes *in vitro* como atividade antiviral e análises de citotoxicidade devem ser realizadas.

Este trabalho de conclusão de curso está em consonância com os pilares do curso de Engenharia de Bioprocessos da UFSJ/CAP, cujo perfil do egresso deve focar, entre outros, a Indústria Farmacêutica. A prospecção de novas drogas para fins terapêuticos significa, além de um prazer pelo desafio inerente à pesquisa experimental, uma grande satisfação em poder contribuir com alguma tecnologia para o bem-estar social e utilização sustentável de recursos ambientais. As perspectivas levantadas acima representam desafio para os próximos anos desse grupo de pesquisa do qual pude fazer parte.

6.REFERÊNCIAS BIBLIOGRÁFICAS

- ALTSCHUL, S. F. *et al.* Basic local alignment search tool. *Journal of Molecular Biology*, v. 215, p. 403-410, 1990.
- AMARAL A. M. *et al.* O programa BLAST: guia prático de utilização. Embrapa Recursos Genéticos e Biotecnologia. Brasília, DF, 2007.
- ANFISEN, C. B. Some observations on the basic principles of design in protein molecules. *Comparative Biochemistry and Physiology*, v. 4, p. 229-240, 1962.
- BAJORATH, J. Integration of virtual and high-throughput screening. *Nature Reviews Drug Discovery*, v. 1, p. 82-94, 2002.
- BENKERT P. *et al.* QMEAN clust: estimation of protein model quality by combining a composite scoring function with structural density information. *BMC Structural Biology*, v. 9, 2009.
- BENKERT P. *et al.* Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics*, v. 27, n. 3, p. 343-50, 2011
- BROZELL, S. R. *et al.* Evaluation of DOCK 6 as a pose generation and database enrichment tool. *Journal of Computer-Aided Molecular Design*, v. 26, p. 749-773, 2012
- CAETANO, C. C. S. *Avaliação do estresse oxidativo e defesas antioxidantes em macrófagos murinos após infecção pelo Mayaro virus (Togaviridae)*. 2016. 74f. Dissertação (Mestrado em Ciências Biológicas). Instituto de Ciências Exatas e Biológicas, Universidade Federal de Ouro Preto, Ouro Preto, 2016.
- CHENG R. H. *et al.* Nucleocapsid and glycoprotein organization in an enveloped virus. *Cell*, v. 80, n. 4, p. 621-630, 1995.
- COIMBRA, T. L. *et al.* Mayaro virus: imported cases of human infection in São Paulo State, Brazil. *Revista do Instituto de Medicina Tropical de São Paulo*, São Paulo, v. 49, n. 4, p. 221-224, ago. 2007.
- CORNELL, W. *et al.* A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *Journal of the American Chemical Society*, p. 5179-5197, 1995.
- 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, 1986.
- DE CARVALHO, P. V. S. D. Modelagem Molecular de Novos Derivados Triazólicos Inibidores de Glicosidasas com Potencial Terapêutico em Diabetes do Tipo 2. 2014. 130f. Dissertação (Mestrado em ciências). Instituto de Ciências Biológicas do Instituto Oswaldo Cruz. Rio de Janeiro, 2014.

DOMINGUES.3D-Pharma: uma ferramenta para triagem virtual baseada em fingerprints de farmacóforos. Tese (Doutorado em Bioinformática)-*Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais*, Belo Horizonte,2013.

FAUCI A. S, MORENS D. M.The perpetual challenge of infectious diseases.*The New England Journal of Medicine*, v. 366,n.5;p.454-61,2012.

FERREIRA D. F. *et al.* Weak bases affect late stages of Mayaro virus replication cycle in vertebrate cells. *Journal of Medical Microbiology*,v.49,n.4,p.313-318,2000.

FERREIRA, L. G. *et al.* Molecular Docking and Structure-Based Drug Design Strategies.*Molecules*,v.20,p.13384-13421,2015.

FIGUEIREDO, M. L. G.; FIGUEIREDO, L. T. M. Emerging alphaviruses in the Americas: Chikungunya and Mayaro.*Revista da Sociedade Brasileira de Medicina Tropical*.v. 47, n. 6, p. 677-683,2014 .

FILHO, S. A. O.; ALENCASTRO, R. B.Modelagem de Proteínas por Homologia. *Química Nova*, v. 26, n. 2. p.253-259,2003.

FORSHEYB. M. *et al.* Arboviral etiologies of acute febrile illnesses in Western South America, 2000–2007.*PLOS Neglected Tropical Diseases*, v.4,n.8,p.787,2010.

FOY N. J. *et al.* Hypervariable Domain of Nonstructural Protein nsP3 of Venezuelan Equine Encephalitis Virus Determines Cell-Specific Mode of Virus Replication. *American Society for Microbiology Journals*,v. 83,n. 13, jul. 2013.

GASPAR L. P. *et al.*The metastable state of nucleocapsids of enveloped viruses as probed by high hydrostatic pressure.*Journal of Biological Chemistry*,v.276,n.10,p.7415-7421,2001.

GUIDO, R. V. C. *et al.*Planejamento de fármacos, biotecnologia e química medicinal: aplicações em doenças infecciosas. *Instituto de Estudos Avançados da Universidade de São Paulo*, São Paulo , v. 24, n. 70, p. 81-98,2010 .

GUIDO, R. V. C. *et al.* Virtual screening and its integration with modern drug design technologies. *Current Medicinal Chemistry*, v.15, n.1, p.37-46, 2008.

GUBLER D. J. Human arbovirus infections worldwide.*Annals of the New York Academy of Sciences*, v.951,n.1,p.13-24, 2001.

HALSEY, E. S. *et al.* “Mayaro Virus Infection, Amazon Basin Region, Peru, 2010–2013.” *Emerging Infectious Diseases*,v.19,n.11,p. 1839–1842,2013.

HUMPHREY, W., *et al.*"VMD - Visual Molecular Dynamics",*Journal of Molecular Graphics*,v. 14, p. 33-38,1996.

JUNIOR J. J. S. *et al.* Construção do Receptor CYSLTR1 por Modelagem Comparativa.*Revista de Ciências Farmacêuticas Básica e Aplicada*,v.33,n.4,p.591-596,2012.

KARRAS G. I. *et al.* The macro domain is an ADP-ribose binding module. *Journal EMBO*, v.24,n.11,p.1911,2005.

LAVERGNE A. *et al.* Mayaro Virus: complete nucleotide sequence and phylogenetic relationships with other alphaviruses. *VirusResearch*, v.117,n.2, p.461-469,2006.

LESK A. M. Introdução à Bioinformática. 2nd ed. Porto Alegre: *Artmed*, 2008.

LIMA-CAMARA T. N. Arboviroses emergentes e novos desafios para a saúde pública no Brasil. *Revista de Saúde Publica*, v.50,p.36,2016.

LOPES N.; NOZAWA C.; LINHARES R. E.C. Características gerais e epidemiologia dos arbovírus emergentes no Brasil. *Revista Pan-Amazônica de Saúde*, v.5,n.3,p.55-64,2014.

LOMBARDINO, J. G.; LOWE, J. A. 3rd. The role of the medicinal chemist in drug discovery – then and now. *Nature Reviews Drug Discovery*, v.3, n.10, p.853-62, 2004.

LONG K. C. *et al.* Experimental transmission of *Mayaro virus* by *Aedes aegypti*. *American Journal of Tropical Medicine and Hygiene*, v.85,p.750–757,2011.

LOVELL S. C. *et al.* Structure validation by Calpha5 RAMPAGE: Ramachandran Plot Assessment geometry: phi,psi and Cbeta deviation. *Proteins: Structure, Function & Genetics*, v.50,p.437.- 2002.

MALET H. *et al.* The crystal structures of Chikungunya and Venezuelan equine encephalitis virus nsP3 macro domains define a conserved adenosine binding pocket. *Journal of Virology*, v.83,n.13,p.6534 –6545,2009.

MELO, F.; FEYTMANS, E. "Novel knowledge-based mean force potential at atomic level". *Journal of Molecular Biology*. n. 267, p. 207-222, 1997

MOULT, J. A decade of CASP: progress, bottlenecks, and prognosis in protein structure prediction. *Current Opinion in Structural Biology*, v.15, n.3, p. 285-289,2005.

MOTA, M. T. O. *et al.* *Mayaro virus*: a neglected arbovirus of the Americas. *Future Virology*, v. 10, n. 9, p. 1109-1122, 2015.

NAPOLÊÃO-PEGO, P., *et al.* Mayaro Virus Disease. *Journal of Human Virology & Retrovirology*, v. 1, n. 3, p.00018, 2014.

NORRIS DE. Mosquito-borne diseases as a consequence of land use change. *Journal EcoHealth*, v.1,n.1,p.19-24, 2004.

PARROT M. M. *et al.* Role of conserved cysteines in the alphavirus E3 protein. *Journal of Virology*, v.86,n.6, p.2584-2591, 2009.

PASSAMANI, F. Modelagem Molecular e Avaliação da Relação Estrutura-Atividade Acoplados a Estudos Farmacocinéticos e Toxicológicos In Silico de Derivados Heterocíclicos com Atividade Antiviral.2009,103fs. *Dissertação (Mestrado em Ciências Farmacêuticas)*– Universidade Federal do Rio de Janeiro, Faculdade de Farmácia, Rio de Janeiro, 2009.

PEDROSO, E. R. P.; ROCHA, M. O. C. Infecções emergentes e reemergentes. *Revista Médica de Minas Gerais*, v.19, n.2, p.140-150, 2009.

PEARSON, W. R. An Introduction to Sequence Similarity (“Homology”) Searching. *Current Protocols in Bioinformatics*, p. 1-8, 2013.

PHILLIPS, J. C. *et al.* Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, v. 26, n. 16, p. 1781–1802, 2005.

POWERS A. M. *et al.* Genetic relationships among Mayaro and UNA viruses suggest distinct patterns of transmission. *American Journal of Tropical Medicine and Hygiene*, v.75, n.3, p.461-469, 2006.

ROCHA, K. B.; Júnior. W. F. A. Virtual screening para identificação de inibidores da corismatosintase de *Mycobacterium tuberculosis*. *IV Mostra de Pesquisa da Pós-Graduação*. Pontífica Universidade Católica, Rio Grande do Sul, 2009.

RODRIGUES, R. P. *et al.* Estratégias de Triagem Virtual no Planejamento de Fármacos. *Revista Virtual Química*, v.4, n. 6, p.739-776, 2012.

RUNGROT MONGKOL, T. *et al.* Molecular insight into the specific binding of ADP-ribose to the nsP3 macro domains of *chikungunya* and *venezuelan equine encephalitis* viruses: Molecular dynamics simulations and free energy calculation. *Journal of Molecular Graphics and Modelling*, v.29, p 347-353, 2010.

SANTIAGO, F. W., *et al.* Long-Term Arthralgia after Mayaro Virus Infection Correlates with Sustained Pro-inflammatory Cytokine Response. *PLOS Neglected Tropical Diseases*, v.9, n.10, 2015.

STRAUSS J. H. STRAUSS E. G. The alphaviruses: gene expression, replication, and evolution. *Microbiological Reviews*, v.58, n.3, p.865-872. 1998.

SEYEDI, S. S. *et al.* Computational Approach Towards Exploring Potential Anti-Chikungunya Activity of Selected Flavonoids. *Scientific Reports- Nature*. v.240, n.27, 2016.

SHIN, G. *et al.* Structural and functional insights into alphavirus polyprotein processing and pathogenesis. *Proceedings of the National Academy of Sciences*, v. 109, p. 16534-16539, 2012.

SKOGING U.; LILJESTROM P. Role of the C-terminal tryptophan residue for the structure-function of the alphavirus capsid protein. *Journal of Molecular Biology*, v.279, n.4, p. 865-872, 1998.

SOUZA, M. L. Modelagem por Homologia de Estruturas Protéicas do Veneno de *Apis mellifera*. Monografia. Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto Faculdade de Medicina de Ribeirão Preto. Universidade de São Paulo. 2008. 72p.

SPINDOLA *et al.* Anti-Mayaro virus activity of *Cassia australis* extracts (Fabaceae, Leguminosae). *Parasite and vectors*, v.537, n.7, 2014.

VERLI, H. Bioinformática da Biologia à flexibilidade molecular. *Organização Hugo Verli*, Porto Alegre, v. 1 ,282p,2014.

VIEIRA, G. C. Modelagem Molecular e Imunodeteção de DNA Metiltransferases2 de Drosophilídeos: Uma Abordagem Evolutiva da Enigmática DNMT2.2015.131f.Dissertação (Mestrado em Genética e Biologia Molecular)-*Programa de Pós-Graduação em Genética e Biologia Molecular*, Universidade Federal do Rio Grande do Sul, Porto Alegre.2015.

VYAS, V. K. et al. “Homology Modeling a Fast Tool for Drug Discovery: Current Perspectives.” *Indian Journal of Pharmaceutical Sciences*,v.74,n.1,p.1-17,2012.

APÊNDICE I

Resultado da triagem virtual inversa, por meio do *docking* molecular dirigido, considerando a proteína nsP3 de MAYV modelada como receptor. Ao todo 296 moléculas derivadas da literatura ou do banco de dados ZINC foram testadas. A energia de ligação Kcal/mol é apresentada abaixo:

Nº	Molécula	Derivado	Energia de ligação (Kcal/mol)
1	zinc_8442275	Derivado Fenólico	-11,5
2	zinc_8442273	Derivado Fenólico	-11,3
3	zinc_8442279	Derivado Fenólico	-10,1
4	DERIVADOFENOL ICO_44_98I	*	-10,1
5	zinc_8442293	Derivado Fenólico	-9,9
6	zinc_13507177	ADP-Ribose	-9,9
7	zinc_33974212	ADP-Ribose	-9,8
8	zinc_8216005	ADP-Ribose	-9,6
9	zinc_13434884	ADP-Ribose	-9,6
10	zinc_18456332	ADP-Ribose	-9,6
11	zinc_30320776	ADP-Ribose	-9,6
12	zinc_13434883	ADP-Ribose	-9,5
13	zinc_16051708	ADP-Ribose	-9,5
14	zinc_16546167	ADP-Ribose	-9,5
15	zinc_8586022	ADP-Ribose	-9,4
16	zinc_4806450	ADP-Ribose	-9,3
17	zinc_8664605	ADP-Ribose	-9,3
18	FISSETINA	*	-9,2
19	zinc_3871614_lig2	ADP-Ribose	-9,2
20	zinc_4806450d	ADP-Ribose	-9,2
21	zinc_12360704	ADP-Ribose	-9,2
22	zinc_12502232	ADP-Ribose	-9,2
23	zinc_13434875	ADP-Ribose	-9,2
24	zinc_37256278	ADP-Ribose	-9,2
25	QUERCETINA	*	-9,1
26	zinc_85476338	4-O-Metil- epigallocatequina	-9,1
27	zinc_13434881	4-O-Metil- epigallocatequina	-9,1
28	zinc_3871615_lig1	ADP-Ribose	-9,1
29	zinc_4228234	ADP-Ribose	-9,1
30	zinc_4806446	ADP-Ribose	-9,1
31	zinc_12360703	ADP-Ribose	-9,1
32	zinc_13434879	ADP-Ribose	-9,1
33	zinc_3871612_lig1	ADP-Ribose	-9,1
34	zinc_3871613_lig1	ADP-Ribose	-9,1
35	zinc_85476338	MÊS	-9,1
36	zinc_13434881	MÊS	-9,1

Nº	Molécula	Derivado	Energia de ligação (Kcal/mol)
37	zinc_19990070	Derivado Fenólico	-9
38	zinc_968223	4-O-Metil- epigallocatequina	-9
39	zinc_87606977	4-O-Metil- epigallocatequina	-9
40	zinc_13434873	ADP-Ribose	-9
41	zinc_13434878	ADP-Ribose	-9
42	zinc_13527911	ADP-Ribose	-9
43	zinc_13546979	ADP-Ribose	-9
44	zinc_968223	MÊS	-9
45	zinc_87606977	MÊS	-9
46	zinc_982962	Derivado Fenólico	-8,9
47	zinc_3873977	ADP-Ribose	-8,9
48	zinc_4806433	ADP-Ribose	-8,9
49	zinc_13434876	ADP-Ribose	-8,9
50	zinc_37256281	ADP-Ribose	-8,9
51	BAICALEINA	*	-8,8
52	zinc_14643053	4-O-Metil- epigallocatequina	-8,8
53	zinc_4514067	ADP-Ribose	-8,8
54	zinc_8215728	ADP-Ribose	-8,8
55	zinc_8586020	ADP-Ribose	-8,8
56	zinc_12360002	ADP-Ribose	-8,8
57	zinc_13473520	ADP-Ribose	-8,8
58	zinc_18126793	ADP-Ribose	-8,8
59	zinc_30320773	ADP-Ribose	-8,8
60	zinc_37256276	ADP-Ribose	-8,8
61	zinc_14643053	MÊS	-8,8
62	ADP-RIBOSE	*	-8,7
63	COUMESTROL	*	-8,7
64	HESPERETINA	*	-8,7
65	zinc_4806442	ADP-Ribose	-8,7
66	zinc_12958381	ADP-Ribose	-8,7
67	zinc_13539741	ADP-Ribose	-8,7
68	zinc_15637617	ADP-Ribose	-8,7
69	zinc_16346988	ADP-Ribose	-8,7
70	zinc_9365179	Derivado Fenólico	-8,6
71	zinc_18141403	Derivado Fenólico	-8,6
72	zinc_19799526	Derivado Fenólico	-8,6
73	zinc_4096994	ADP-Ribose	-8,6
74	zinc_8586021	ADP-Ribose	-8,6
75	zinc_24608957	ADP-Ribose	-8,6
76	zinc_1631259	ADP-Ribose	-8,6
77	zinc_4119504	Derivado Fenólico	-8,5
78	zinc_8215662	ADP-Ribose	-8,5
79	zinc_13516496	ADP-Ribose	-8,5
80	zinc_13518964	ADP-Ribose	-8,5
81	zinc_36377954	ADP-Ribose	-8,5

Nº	Molécula	Derivado	Energia de ligação (Kcal/mol)
82	zinc_3201889	ADP-Ribose	-8,5
83	zinc_3869247	ADP-Ribose	-8,5
84	zinc_3869250	ADP-Ribose	-8,5
85	zinc_3869248	ADP-Ribose	-8,5
86	zinc_8575396	Derivado Fenólico	-8,4
87	ZOASTERICACID	*	-8,4
88	7-O-METILGLABRANI NA	*	-8,4
89	GLABRANINA	*	-8,4
90	zinc_3870336	4-O-Metil- epigalocatequina	-8,4
91	zinc_14642744	4-O-Metil- epigalocatequina	-8,4
92	zinc_14642747	4-O-Metil- epigalocatequina	-8,4
93	zinc_8586019	ADP-Ribose	-8,4
94	zinc_8613167	ADP-Ribose	-8,4
95	zinc_12502230	ADP-Ribose	-8,4
96	zinc_13473519	ADP-Ribose	-8,4
97	zinc_13543323	ADP-Ribose	-8,4
98	zinc_24424495	ADP-Ribose	-8,4
99	zinc_33902336	ADP-Ribose	-8,4
100	zinc_1532515	ADP-Ribose	-8,4
101	zinc_1842158	ADP-Ribose	-8,4
102	zinc_3201891	ADP-Ribose	-8,4
103	zinc_3830180	ADP-Ribose	-8,4
104	zinc_3870336	MÊS	-8,4
105	zinc_14642744	MÊS	-8,4
106	zinc_14642747	MÊS	-8,4
107	GENISTEINA	*	-8,3
108	zinc_119978	4-O-Metil- epigalocatequina	-8,3
109	zinc_119983	4-O-Metil- epigalocatequina	-8,3
110	zinc_119985	4-O-Metil- epigalocatequina	-8,3
111	zinc_119988	4-O-Metil- epigalocatequina	-8,3
112	zinc_14642741	4-O-Metil- epigalocatequina	-8,3
113	zinc_14642912	4-O-Metil- epigalocatequina	-8,3
114	zinc_4514065	ADP-Ribose	-8,3
115	zinc_4514075	ADP-Ribose	-8,3
116	zinc_119978	MÊS	-8,3
117	zinc_119983	MÊS	-8,3
118	zinc_119985	MÊS	-8,3

Nº	Molécula	Derivado	Energia de ligação (Kcal/mol)
119	zinc_119988	MÊS	-8,3
120	zinc_14642741	MÊS	-8,3
121	zinc_14642912	MÊS	-8,3
122	COMPOSTOFENOL ICO	*	-8,2
123	4-O-METIL- EPIGALOCATEQUINA	*	-8,2
124	CELGOSIVIN	*	-8,2
125	zinc_1721694	4-O-Metil- epigalocatequina	-8,2
126	zinc_3870337	4-O-Metil- epigalocatequina	-8,2
127	zinc_3870338	4-O-Metil- epigalocatequina	-8,2
128	zinc_3870339	4-O-Metil- epigalocatequina	-8,2
129	zinc_3977897	ADP-Ribose	-8,2
130	zinc_13546985	ADP-Ribose	-8,2
131	zinc_1571045	ADP-Ribose	-8,2
132	zinc_2126310	ADP-Ribose	-8,2
133	zinc_3860156	ADP-Ribose	-8,2
134	zinc_1721694	MÊS	-8,2
135	zinc_3870337	MÊS	-8,2
136	zinc_3870338	MÊS	-8,2
137	zinc_3870339	MÊS	-8,2
138	zinc_968078	4-O-Metil- epigalocatequina	-8,1
139	zinc_8681596	4-O-Metil- epigalocatequina	-8,1
140	zinc_8681597	4-O-Metil- epigalocatequina	-8,1
141	zinc_14643106	4-O-Metil- epigalocatequina	-8,1
142	zinc_4245698	ADP-Ribose	-8,1
143	zinc_16082015	ADP-Ribose	-8,1
144	zinc_3201893	ADP-Ribose	-8,1
145	zinc_968078	MÊS	-8,1
146	zinc_8681596	MÊS	-8,1
147	zinc_8681597	MÊS	-8,1
148	zinc_14643106	MÊS	-8,1
149	zinc_12378847	Derivado Fenólico	-8
150	zinc_1721693	4-O-Metil- epigalocatequina	-8
151	zinc_3978494	4-O-Metil- epigalocatequina	-8
152	zinc_14642684	4-O-Metil- epigalocatequina	-8

Nº	Molécula	Derivado	Energia de ligação (Kcal/mol)
153	zinc_1721693	MÊS	-8
154	zinc_3978494	MÊS	-8
155	zinc_14642684	MÊS	-8
156	zinc_1240782	Derivado Fenólico	-7,9
157	zinc_19166762	Derivado Fenólico	-7,9
158	DAIDZEIN	*	-7,9
159	zinc_2525302	4-O-Metil- epigallocatequina	-7,9
160	zinc_14642730	4-O-Metil- epigallocatequina	-7,9
161	zinc_14642733	4-O-Metil- epigallocatequina	-7,9
162	zinc_14642736	4-O-Metil- epigallocatequina	-7,9
163	zinc_14642739	4-O-Metil- epigallocatequina	-7,9
164	zinc_3581269	ADP-Ribose	-7,9
165	zinc_2525302	MÊS	-7,9
166	zinc_14642730	MÊS	-7,9
167	zinc_14642733	MÊS	-7,9
168	zinc_14642736	MÊS	-7,9
169	zinc_14642739	MÊS	-7,9
170	zinc_3901268	Derivado Fenólico	-7,8
171	EPIGALLOCATEQUINA	*	-7,8
172	zinc_33363601	ADP-Ribose	-7,8
173	zinc_16940491	ADP-Ribose	-7,8
174	zinc_3871612_lig2	ADP-Ribose	-7,8
175	zinc_3871613_lig2	ADP-Ribose	-7,8
176	zinc_20031600	Derivado Fenólico	-7,7
177	zinc_5998626	4-O-Metil- epigallocatequina	-7,7
178	zinc_8681654	ADP-Ribose	-7,7
179	zinc_5998626	MÊS	-7,7
180	zinc_19990034	Derivado Fenólico	-7,6
181	zinc_5850183	Cloroquina	-7,6
182	zinc_12243850	Cloroquina	-7,6
183	zinc_3197625	4-O-Metil- epigallocatequina	-7,5
184	zinc_3978506	4-O-Metil- epigallocatequina	-7,5
185	zinc_6092138	4-O-Metil- epigallocatequina	-7,5
186	zinc_59727703	4-O-Metil- epigallocatequina	-7,5
187	zinc_5615258	ADP-Ribose	-7,5
188	zinc_5850156	Cloroquina	-7,5
189	zinc_12243849	Cloroquina	-7,5

Nº	Molécula	Derivado	Energia de ligação (Kcal/mol)
190	zinc_3197625	MÊS	-7,5
191	zinc_3978506	MÊS	-7,5
192	zinc_6092138	MÊS	-7,5
193	zinc_59727703	MÊS	-7,5
194	RIBAVIRINA	*	-7,4
195	zinc_1845185	4-O-Metil- epigalocatequina	-7,4
196	zinc_3197613	4-O-Metil- epigalocatequina	-7,4
197	zinc_4096941	4-O-Metil- epigalocatequina	-7,4
198	zinc_4096942	4-O-Metil- epigalocatequina	-7,4
199	zinc_4096944	4-O-Metil- epigalocatequina	-7,4
200	zinc_19594730	MÊS	-7,4
201	zinc_1845185	MÊS	-7,4
202	zinc_3197613	MÊS	-7,4
203	zinc_4096941	MÊS	-7,4
204	zinc_4096942	MÊS	-7,4
205	zinc_4096944	MÊS	-7,4
206	zinc_8442272	Derivado Fenólico	-7,3
207	zinc_5556455	Derivado Fenólico	-7,3
208	ALKALOID	*	-7,3
209	zinc_2170446	4-O-Metil- epigalocatequina	-7,3
210	zinc_4096943	4-O-Metil- epigalocatequina	-7,3
211	zinc_3954454oc	ADP-Ribose	-7,3
212	zinc_3869249	ADP-Ribose	-7,3
213	zinc_6424763	Cloroquina	-7,3
214	zinc_2170446	MÊS	-7,3
215	zinc_4096943	MÊS	-7,3
216	zinc_419	4-O-Metil- epigalocatequina	-7,2
217	zinc_4096940	4-O-Metil- epigalocatequina	-7,2
218	zinc_17013937	4-O-Metil- epigalocatequina	-7,2
219	zinc_71404929	MÊS	-7,2
220	zinc_419	MÊS	-7,2
221	zinc_4096940	MÊS	-7,2
222	zinc_17013937	MÊS	-7,2
223	zinc_19989886	Derivado Fenólico	-7,1
224	ACACETIN-7-O- RUTINOSIDE	*	-7,1
225	CLOROQUINA1+	*	-7,1
226	zinc_3591028	4-O-	-7,1

Metilepigalocatequina			
Nº	Molécula	Derivado	Energia de ligação (Kcal/mol)
227	Cloroquine	Cloroquina	-7,1
228	zinc_1712475	Cloroquina	-7,1
229	zinc_19144231	Cloroquina	-7,1
230	zinc_19364835	MÊS	-7,1
231	zinc_36161963	MÊS	-7,1
232	zinc_3591028	MÊS	-7,1
233	zinc_5751278	Cloroquina	-7
234	zinc_19144226	Cloroquina	-7
235	zinc_70190038	MÊS	-7
236	TINGENONA	*	-6,9
237	zinc_4514071	ADP-Ribose	-6,9
238	CLOROQUINA2+	*	-6,8
239	zinc_1530860	Cloroquina	-6,8
240	zinc_1700294	Derivado Fenólico	-6,7
241	zinc_1530861	Cloroquina	-6,7
242	zinc_69730149	MÊS	-6,7
243	AMINOGYCOSIDE GENTICIN	*	-6,6
244	zinc_1710230	MÊS	-6,6
245	zinc_1713409	MÊS	-6,6
246	zinc_2895765	MÊS	-6,6
247	METILGALATO	*	-6,4
248	zinc_19203136	MÊS	-6,4
249	zinc_85438601	MÊS	-6,4
250	zinc_86855108	MÊS	-6,4
251	zinc_2819766	MÊS	-6,3
252	zinc_36162188	MÊS	-6,3
253	zinc_44594695	MÊS	-6,3
254	zinc_86855066	MÊS	-6,3
255	zinc_86855110	MÊS	-6,3
256	zinc_90694025	MÊS	-6,3
257	zinc_94300963	MÊS	-6,3
258	TRIPTAMINA1+	*	-6,2
259	zinc_4343272	MÊS	-6,2
260	zinc_37652461	MÊS	-6,2
261	zinc_71288110	MÊS	-6,2
262	zinc_34936425	MÊS	-6,1
263	zinc_36162219	MÊS	-6,1
264	zinc_71288108	MÊS	-6,1
265	zinc_72194498	MÊS	-6,1
266	zinc_72194499	MÊS	-6,1
267	zinc_86855487	MÊS	-6,1
268	zinc_5848508	MÊS	-6
269	zinc_19419111	MÊS	-6
270	zinc_52480351	MÊS	-6
271	zinc_94300941	MÊS	-6

Nº	Molécula	Derivado	Energia de ligação (Kcal/mol)
272	PETCOLINARIN	*	-5,9
273	zinc_1583444	MÊS	-5,9
274	zinc_1659810	MÊS	-5,9
275	zinc_5849378	MÊS	-5,9
276	zinc_69730476	MÊS	-5,9
277	zinc_78774919	MÊS	-5,9
278	zinc_86854295	MÊS	-5,9
279	zinc_94300838	MÊS	-5,9
280	zinc_37743370	MÊS	-5,8
281	zinc_52480349	MÊS	-5,8
282	zinc_70596246	MÊS	-5,8
283	MÊS	*	-5,7
284	zinc_19369410	MÊS	-5,7
285	zinc_49580718	MÊS	-5,7
286	zinc_52480350	MÊS	-5,7
287	zinc_86855085	MÊS	-5,7
288	zinc_85891957	MÊS	-5,3
289	zinc_3130601	MÊS	-5
290	zinc_38615690	MÊS	-5
291	zinc_1671858	MÊS	-4,8
292	zinc_1706239	MÊS	-4,8
293	zinc_2913677	MÊS	-4,8
294	zinc_282347	MÊS	-4,7
295	zinc_5189619	MÊS	-4,5
296	SQUALAMINE	*	-3