

UNIVERSIDADE DE BRASÍLIA
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA MOLECULAR

TESE DE DOUTORADO

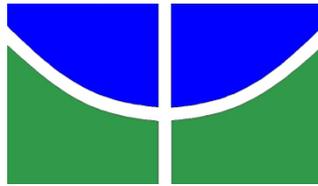
ANÁLISES ESTRUTURAIS E DE MODELOS DE INTERAÇÃO DAS PROTEASES NS3
DOS FLAVIVIRUS ZIKA E DENGUE, E NSP2 DOS ALPHAVIRUS CHIKUNGUNYA E
MAYARO UTILIZANDO FERRAMENTAS *IN SILICO* E *IN VITRO*

ALUNO: JÔNATAS CUNHA BARBOSA LIMA

ORIENTADOR: JOÃO ALEXANDRE RIBEIRO GONÇALVES BARBOSA

14 DE JANEIRO DE 2022

BRASÍLIA- DF



UNIVERSIDADE DE BRASÍLIA
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA MOLECULAR

Jônatas Cunha Barbosa Lima

ANÁLISES ESTRUTURAIS E DE MODELOS DE INTERAÇÃO DAS PROTEASES NS3
DOS FLAVIVIRUS ZIKA E DENGUE, E NSP2 DOS ALPHAVIRUS CHIKUNGUNYA E
MAYARO UTILIZANDO FERRAMENTAS *IN SILICO* E *IN VITRO*

Tese submetida ao programa de Pós-graduação em Biologia Molecular da Universidade de Brasília como requisito parcial à obtenção do título de Doutor em Biologia Molecular.

Orientador: Prof. Dr. João Alexandre Ribeiro Gonçalves Barbosa

Brasília
2022

**À minha maravilhosa esposa Karla e ao pequeno Isaque.
Agradeço a Deus por tê-los em minha vida.**

Agradecimentos

Agradeço em primeiro lugar a Deus, não de forma meramente rotineira, mas por reconhecer que até aqui fui guardado e sustentado por Ele. Tenho sido abençoado com muito mais do que mereço, principalmente em receber a benção da salvação em Cristo Jesus.

Ao meu orientador Dr. João Alexandre, obrigado por me ensinar tanto durante minha iniciação científica, mestrado e doutorado, foram anos muito ricos de aprendizado. Aos demais professores do Laboratório de Biofísica Molecular da Universidade de Brasília, Dra. Sonia Freitas, Dr. Napoleão Valadares e Dr. Aisel Garay, os quais também tanto contribuíram para minha formação. Além de tantos outros professores que durante todas as fases da minha vida me agradeceram com tantos ensinamentos.

Agradeço a minha família. Minha esposa Karla e o recém-chegado Isaque, com os quais tenho tido prazer de dividir a minha vida. A minha mãe Lílian e meu pai Jolson, obrigado por me darem muito mais que um sustento material ao longo dos anos, obrigado por me ensinarem sobre Cristo. Agradeço aos meus irmãos André, Marcos e Mariane e minha cunhada Cássia. A minha animada e amada vovó Cenyr, e aos demais vovôs e vovós que não se encontram mais aqui. A família do meu sogro Josué e minha sogra Lindaura que sempre me acolheram com tanto carinho. Aos meus sobrinhos Dante, Rebeca, Levi e Romeu. A tantos tios e tias, primos e primas, faltariam linhas para agradecer a todos vocês.

Agradeço aos amigos do Laboratório de Biofísica Molecular, a todos que lá estão e que por lá passaram durante esses anos que tive o prazer de trabalhar naquele laboratório. Aos novos amigos do Departamento de Assistência Farmacêutica do Ministério da Saúde, em especial ao meu chefe Sidney e meu amigo João Paulo. Aos tantos outros amigos/irmãos da 1ª Igreja Presbiteriana da Arniqueira, e de tantas outras igrejas e locais diferentes e que me acompanham até o dia de hoje.

Ao professor Dr. Rolf Hilgenfeld, a Dra. Linlin Zhang e toda a equipe do Instituto de Bioquímica da Universidade de Lübeck, Alemanha, pela ótima recepção e ricos ensinamentos durante os meses de trabalho juntos.

Aos prezados membros da banca Dr. Diego Nolasco, Dr. Jorge Lulek, Dra. Izabela Dourado e Dr. Napoleão Valadares pela disponibilidade e preciosos comentários visando aprimorar o presente trabalho.

Ao CNPq, CAPES e FAP-DF pelo apoio financeiro.

A estes e a tantos outros o meu muito obrigado.

“As we conquer peak after peak we see in front of us regions full of interest and beauty, but we do not see our goal, we do not see the horizon; in the distance tower still higher peaks, which will yield to those who ascend them still wider prospects, and deepen the feeling, the truth of which is emphasised by every advance in science, that "Great are the Works of the Lord."”

J. J. THOMSON, The British Association at Winnipeg. *Nature*, 81, 248–263 (1909).

RESUMO

Os arbovírus (Arthropod-borne virus) são uma classificação que compreende grupos virais cuja transmissão é realizada por vetores artrópodes que infectam vertebrados. Dentre esses, alguns arbovírus têm se mostrado de grande relevância, devido ao grande número de infecções, como os Flavivirus dengue (DENV) e o zika (ZIKV), e os Alphavirus chikungunya (CHIK) e o mayaro (MAYV), sendo que todos carecem de estudos estruturais mais aprofundados objetivando tanto um maior conhecimento da biologia molecular destes, quanto visando possíveis alvos para novos fármacos. Dentre as possíveis proteínas alvo, as proteases virais têm se mostrado excelentes alvos. Estas proteínas (NS3 nos Flavivirus e nsP2 nos Alphavirus) possuem a função de clivar determinadas regiões da poliproteína viral sendo essenciais para a replicação. O objetivo do presente trabalho é realizar análises *in silico* e *in vitro* das proteases dos arbovírus ZIKV, DENV, CHIK e MAYV visando uma melhor compreensão estrutural das proteases e da interação com inibidores e substratos. Os genes da NS3 de ZIKV e DENV, e nsP2 de MAYV e CHIK foram expressos em diversas cepas de *E. coli* e passaram por diversas etapas de otimização de purificação. Realizou-se ultracentrifugação analítica e cromatografia por exclusão molecular das proteases de ZIKV e DENV, que apontaram uma possível formação de dímero. Análises de espectroscopia de fluorescência da NS3 de ZIKV em presença de aprotinina demonstraram uma redução na intensidade de fluorescência possivelmente por maior exposição do Trp50 ao solvente, que poderia ser explicada pela dissociação dos dímeros. *Screening* e refinamento da cristalização de das proteases NS3 estudadas apresentaram condições onde havia formação de cristal de proteína, todavia não foram obtidos cristais que permitissem a resolução da estrutura. Foram realizados ensaios de inibição da NS3 de ZIKV com inibidores da protease homóloga DENV, os quais apresentaram capacidade inibitória também para esta proteína de ZIKV. Além disso, realizou-se a modelagem por homologia da protease do MAYV, *docking* com peptídeos que possuem os sítios de clivagem das nsP2 de CHIK e MAYV e dinâmica das principais poses. Neste experimento observou-se que as poses com melhores *scores* apresentavam semelhante posicionamento de uma glicina, altamente conservada na posição P2 de substratos de nsP2 de Alphavirus, em uma fissura estreita. Este posicionamento possivelmente auxilia no posicionamento do carbono da carbonila do substrato próximo ao enxofre da cisteína catalítica desprotonada facilitando o ataque nucleofílico, uma das primeiras etapas para a quebra da cadeia polipeptídica. Além disso,

estabiliza os ângulos diedros χ_1 e χ_2 dos resíduos que compõe o sítio catalítico.

ABSTRACT

The arbovirus (Arthropod-borne virus) are a classification that comprise viral groups whose transmission is performed by Arthropod vectors infecting vertebrate hosts. Among the arbovirus, dengue virus (DENV) and zika virus (ZIKV) of the family Flaviviridae and gender Flavivirus, and chikungunya virus (CHIK) and mayaro virus (MAYV) of the family Togaviridae and gender Alphavirus, have shown great relevance due to the large number of human infecting cases. All these viruses require deeper structural studies aiming at a much bigger knowledge of their molecular biology to develop new drugs able to counteract their diseases. Among the protein targets, the viral proteases have been recognized as excellent candidates. The proteins (NS3 of Flavivirus and nsP2 of Alphavirus) have a cleavage function in specific regions of the viral polyprotein, therefore they are essential to viral replication. The purpose of this work is to perform *in silico* and *in vitro* experiments of ZIKV, DENV, CHIK and MAYV proteases aiming at a better understanding of the structure and their interaction with inhibitors and substrates. The genes of ZIKV and DENV NS3 and, MAYV and CHIK nsP2 were expressed by several *E. coli* strains and followed by the optimization of their purifications. Analytical ultracentrifugation experiments and size exclusion chromatography were performed for the ZIKV and DENV protease, where a protein dimer was encountered. Fluorescence spectroscopy analysis of ZIKV protease in the presence of aprotinin demonstrated fluorescence intensity reduction possibly by the increase of Trp50 solvent accessibility, which could be explained by dimer dissociation. So far, crystal screening and refinement of the proteases yielded crystals with poor quality, preventing structure solution. Inhibition assays with DENV inhibitors for ZIKV NS3 protease were performed and demonstrated inhibitory capacity of the ZIKV protein. In addition, the MAYV protein homology modeling, docking of the CHIK and MAYV nsP2 cleavage site peptides and molecular dynamics of the best docking poses were performed. In docking experiments, it was observed that best scored poses had the substrate P2 glycine, a highly conserved residue in Alphavirus, placed in a narrow cleft at the protein surface. This arrangement possibly helps to position the substrate carbonyl carbon near the sulfur atom of the deprotonated catalytic cysteine, thus facilitating the nucleophilic attack, the first step of polypeptide chain cleavage. Moreover, this contact stabilizes dihedral angles χ_1 and χ_2 of the active site residues.

LISTA DE FIGURAS

Figura 1- Predição da área de distribuição dos vetores (A) *Aedes aegypti* e (B) *Aedes albopictus* baseado em banco de dados de ocorrência (KRAEMER et al., 2015).

Figura 2- Composição da poliproteína dos Flavivirus. A) Representação das regiões transmembrana da poliproteína indicando com setas os locais de clivagem. B) Divisão das proteínas estruturais e não estruturais com exemplos de imagens de estruturas resolvidas do Flavivirus DENV (FAHEEM et al., 2019).

Figura 3- Composição da poliproteína de Alphavirus com esquematização dos momentos de processamento do genoma durante o a infecção viral. As setas coloridas marcam as proteínas responsáveis por processar cada região (HERNANDEZ; BROWN; PAREDES, 2014).

Figura 4- Mecanismo de catálise de serino proteases (WALKER; LYNAS, 2001).

Figura 5- Estrutura da serino protease NS3 de DENV3 (3U1I) (ciano). A NS2b (verde) necessária para a atividade da proteína se posiciona em volta da estrutura em forma de cinto. Os resíduos do sítio ativo são visualizados em *stick* (His51, Asp 75 e Ser135) e com as distâncias entre estes resíduos em Å representadas pelo pontilhado amarelo.

Figura 6- Estrutura da nsP2 de CHIK (3TRK). Domínio protease (ciano) com resíduos da díade catalítica e alguns do sítio ativo em *stick* (roxo), e domínio MTase (verde).

Figura 7- Esquema com ciclos da reação de PCR para amplificação do segmento gênico de correspondente a protease nsP2 de MAYV.

Figura 8- Inibidores da protease de DENV obtidos para ensaio de inibição contra a protease NS2B/NS3 de ZIKV.

Figura 9- Expressão dos genes nsP2 de CHIK e MAYV, e NS3 de ZIKV a 37°C na cepa de *E. coli* (A) BL21(DE3), (B) BL21(DE3) pLysE, (C) BL21(DE3) Lemo21, (d) BL21(DE3) Rosetta. P. I. marca o padrão de expressão antes da indução. P. as proteínas na fração

insolúvel e S. na solúvel.

Figura 10- Expressão e teste de solubilidade das proteína nsP2 de CHIK e MAYV, e NS3 de ZIKV na cepa BL21 (DE3) de *E.coli* a temperatura 19°C a 1 mM de IPTG. P. I. marca o padrão de expressão antes da indução. P. as proteínas na fração insolúvel e S. na solúvel.

Figura 11- (A) Expressão e (B e C) teste de solubilidade das proteína nsP2 de MAYV na cepa BL21 (DE3) de *E.coli* a temperatura 16°C a 0,2 mM de IPTG em meio LB. P. I. marca o padrão de expressão antes da indução. Pel. as proteínas na fração insolúvel e Sol. na solúvel.

Figura 12- Teste de solubilidade da expressão em meio autoindutor da protease nsP2 de MAYV. A) Apresenta a fração insolúvel (Pel.) e solúvel (Sol.) das coletas referente aos tempos 9h, 10h e 11h, B) e das referentes a 12h e 13h.

Figura 13- Análise por gel SDS-PAGE15% da expressão em diversas cepas de *E. coli* da protease nsP2 de MAYV corado por comassie (esquerda) e *western blot* (direita). Este foi realizado usando tampão contendo 20mM Tris-Hcl, 200mM NaCl, 200mM arginina, 200mM ácido glutâmico, pH 7.1 e glicerol 5%. As cepas usadas foram aplicadas no gel na seguinte ordem: BL21-Gold (DE3), C43 (DE3), Turner (DE3), SoluBL21, BL21 (DE3), BL21 (DE3) pLysE e BL21 (DE3) PlyS. P se refere a fração insolúvel da amostra e E a fração solúvel.

Figura 14- Análise da expressão e purificação da protease nsP2 de MAYV com 1 mM de IPTG, 25°C e 14 horas por SDS-PAGE 15%. Este foi realizado usando tampão contendo 20mM Tris-Hcl, 200mM NaCl, 200mM *Arginine*, 200mM *Glutamic Acid*, pH 7.1. A esquerda utilizando a cepa BL21 e a direita pLysE. PC refere-se a amostra antes da indução por IPTG, Pel a fração insolúvel, Sup a fração solúvel, FT a fração que não interagiu com quando aplicado a coluna de níquel, os demais números aos tubos coletados durante gradiente de imidazol.

Figura 15- Análise por SDS-PAGE 15% da remoção de cauda de histidina da proteína nsP2 de MAYV pela HRV3C. PC se refere a amostra antes da indução por IPTG, Pel a

fração insolúvel e Sup a fração solúvel durante a etapa de expressão. Prot+HRV mostra a proteína junto com a protease HRV3C, enquanto Prot se refere a proteína não tratada para remoção da cauda. Os números representam os diferentes tubos coletados após a remoção da HRV3C por coluna de GST.

Figura 16- Análise por SDS-PAGE 15% da expressão da protease nsP2 de MAYV usando diferentes cepas BL21, pLysE e pLysS. Este foi realizado usando tampão contendo 20mM Tris-HCl, 200mM NaCl, 200mM *Arginine*, 200mM *Glutamic Acid*, pH 7.1. Pel se refere a fração insolúvel, Sup a fração solúvel durante a etapa de expressão, e FT com a amostra que não apresentou interação quando aplicada a coluna de níquel. Os valores referem-se a concentração de imidazol na qual a amostra foi eluída.

Figura 17- Clonagem do gene da fração protease nsP2 de MAYV nos vetores pGEX, pET SUMO, pET28a e pET15b. (A) Amplificação de segmentos gênicos com adição de sequências para os sítios de restrição para as endonucleases NdeI, XhoI e BamHI. (B) Digestão dos vetores selecionados com enzimas de restrição. (B) PCR de colônia para demonstrando a inserção do gene nos vetores.

Figure 18- Análise da expressão de nsP2 de MAYV pos SDS-PAGE 15% usando quatro diferentes vetores de expressão em bactéria *E.coli*, pET28a, pET SUMO, pGEX e pET 15b. PC se refere a amostra antes da indução por IPTG, Pel a fração insolúvel, Sup a fração solúvel, FT a fração que não interagiu com quando aplicado a coluna de níquel, seguido pelas eluições a três diferentes concentrações de imidazol (50, 250 e 500 mM).

Figura 19- Espectro obtido pela técnica de CD da proteína nsP2 de MAYV após renaturação por diluição rápida.

Figura 20- Purificação da protease nsP2 de CHIK. (A) Purificação por cromatografia líquida de afinidade a níquel. (B) Purificação por exclusão molecular. P. marca a fração insolúvel localizada no pellet, FT indica o *flow through* coletado. Os números em mM são referentes a concentração de imidazol na qual as amostras foram eluídas da coluna. O p1, p2, p3 e p4 representa frações de um mesmo pico que foi eluído. Purif. Se refere a amostra purificada por cromatografia líquida de afinidade a níquel antes de ser aplicada na coluna de exclusão molecular.

Figura 21- Expressão e purificação da proteína nsP2 de CHIK. A) Perfil cromatográfico da purificação por gradiente de imidazol em coluna de níquel. A linha Azul se refere a absorvância (AU) e a linha vermelha a concentração de imidazol, de 0 a 500 mM. B) Gel de poliacrilamida 15% apresentando a expressão e purificação, onde PC corresponde as células antes da indução com IPTG, Pel e Sup ao pellet e sobrenadante após a expressão e FT a fração que não interagiu com a coluna de níquel. A numeração 1 a 10 se refere as frações eluidas durante o gradiente de imidazol da purificação.

Figura 22- Remoção da cauda de histidina da nsP2 de CHIK. A) Gel SDS-PAGE 15% referente a etapa de encubação com a protease HRV3C. B) Separação da HRV3C por coluna de GST e de nsP2 com cauda de histidina usando coluna de níquel. GST FT 1 a 3 referem-se as proteínas que não interagiram com a coluna de GST, GSH se refere a eluição da proteínas que interagiram com a coluna de GST, e os poços 50, 250 e 500 mM referente a três lavagens com diferentes concentrações de imidazol da coluna de níquel acoplada. C) Reaplicação da nsP2 em coluna de GSH em diferentes volumes (15, 10 e 5 μ L) em coluna de GST de duas diferentes frações.

Figura 23- Expressão do substrato da protease nsP2 de CHIK. A) Perfil cromatográfico da purificação por gradiente de imidazol usando coluna de níquel. O limite de detecção é de 0,5 AU, a linha azul refere-se a absorvância da amostra (AU) e a vermelha a concentração de imidazol (0 a 500 mM). B) SDS-PAGE 15% da purificação por coluna de níquel. Pel e Sup referem-se ao pellet e sobrenadante obtidos durante as etapas de lise e centrifugação, FT as proteínas que não interagiram com a coluna de níquel e os números de 1 a 9 as amostras eluidas em diferentes concentrações de imidazol. C) Perfil cromatográfico da purificação por troca aniônica (HiTrap Q HP). D) SDS-PAGE 15% das amostras purificadas na etapa de purificação usando coluna de troca aniônica usando gradiente de NaCl (0 a 500 mM). Os números de 5 a 12 referem-se as amostras eluidas nas diferentes concentrações de NaCl.

Figura 24- SDS-PAGE 15% contendo o ensaio de atividade da nsP2 de CHIK. Variação de massa de protease de 0 a 4 μ g em valores fixos de substrato. Os poços marcados com No DTT não apresentavam DTT, todavia os dois poços finais apresentam uma concentração de 1 mM de DTT.

Figura 25- Análises para a qualidade do modelo da protease nsP2 de MAYV. (A) Diagrama de Ramachandran. (B e C) Valores de qualidade obtidos pelo software qmean.

Figura 26- Sobreposição das proteases nsP2 dos Alphavirus CHIK (ciano, PDB 3TRK), MAYV (rosa, modelo por homologia) e VEEV (verde, PDB 2HWK). O quadro mostra os resíduos que compõe o sítio de atividade (C478, C/S/A482 e H548). O resíduo 482 é uma C, S e A em MAYV, CHIV e VEEV, respectivamente.

Figura 27- Gráficos de (A) RMSD e (B) RMSF da dinâmica molecular realizada com o modelo apo da nsP2 de MAYV.

Figura 28- (A) Alinhamento múltiplo entre diversos peptídeos substratos de Alphavirus. Glicina da posição P2 em vermelho e posição P3 em amarelo, ocupada preferencialmente por uma alanina. (B) Posicionamento da glicina P2 em uma fissura na superfície da protease. Peptídeo em verde, glicina em laranja com raio de van der Waals.

Figura 29- Modelo da protease nsP2 de MAYV colorido por níveis de conservação dos resíduos. Em vermelho os resíduos conservados em todos os Alphavirus analisados, em vermelho claro resíduos altamente conservados e branco resíduos variáveis no alinhamento múltiplo. Os *stick* verdes representam uma pose calculada do peptídeo contendo o sítio de clivagem nsP2/nsP3 de MAYV, neste somente está representada a cadeia principal do P5-P1'. A) Modelo da região protease da nsP2 de MAYV em interação com o peptídeo. B) Resíduos que compreendem a fissura responsável por reconhecer o substrato. C) Visão da figura B após rotação de 90°.

Figura 30- Variações dos ângulos ϕ e ψ dos resíduos P3, P2, P1, P1' e P2' dos substratos peptídicos de CHIK e MAYV durante as análises de MD. As linhas em preto referem-se aos valores de ϕ e as em vermelho a de ψ .

Figura 31- Gráficos de Ramachandran 2D e 3D do modelo da glicina na posição P2 do substrato peptídico da proteína nsP2 durante análise por dinâmica molecular de (A) CHIK e (B) MAYV.

Figura 32- Valores dos ângulos χ_1 dos resíduos C478, H548 e S/C482, e χ_2 da H548 durante as análises de MD das nsP2 de CHIK e MAYV.

Figura 33- Posições dos resíduos C478 com diferentes valores dos ângulos χ_1 . Em verde o ângulo é aproximadamente 60° , em ciano 180° e em magenta -60° .

Figura 34- Distância em nm entre o carbono da carbonila do resíduo P1 de substrato e o átomo de enxofre da cisteína do modelo da protease Nsp2 de MAYV e da estrutura de CHIK.

Figura 35- Interação entre a nsP2 de CHIK (ciano) e diferentes substratos peptídicos. nsP2/nsP3 (roxo) e nsP3/nsP4 (cinza). Os pontos amarelos mostram a distância entre o átomo de enxofre da cisteína 478 e o carbono da carbonila o qual sofre o ataque nucleofílico.

Figura 36- Purificação da protease NS3 de ZIKV. (A) Purificação por cromatografia líquida de afinidade a níquel. (B) Purificação por exclusão molecular. P. marca a fração insolúvel localizada no pellet, FT indica o flow through coletado. Os números em mM são referentes a concentração de imidazol na qual as amostras foram eluídas da coluna. O p1 e p2, representa frações de um mesmo pico que foi eluído. A. Se refere a amostra purificada por cromatografia líquida de afinidade a níquel antes de ser aplicada na coluna de exclusão molecular. V.1 e V.2 representam as amostras que não adentraram na malha da coluna de exclusão molecular.

Figura 37- Purificação por troca iônica usando a coluna MonoQ (*strong anion*) da proteína NS2b/NS3 de ZIKV.

Figura 38- Gel SDS-PAGE 12% no qual pode ser observado picos referentes a purificação por troca iônica da protease NS2b/NS3 de ZIKV. P1 e P2 referem-se ao pico 1 (eluído a ~ 180 mM de NaCl) e pico 2 (eluído a aproximadamente 230 mM de NaCl) da figura 9. O número 1 refere-se a fração ascendente do pico e 2 a fração descendente.

Figura 39- Análise por ultracentrifugação analítica da proteína NS2b/NS3 de ZIKV. (A) Perfil de sedimentação da protease. (B) Coeficiente de sedimentação calculado em

diferentes concentrações de proteína.

Figura 40- Representação estrutural de duas possíveis conformações do dímero de NS2b/NS3 da protease de ZIKV. A) Dímero observado por Lei e colaboradores (2016). B) Dímero observado por Phoo e colaboradores (2016). Os *sticks* representam os resíduos que compõe a tríade catalítica da protease, His51, Asp75 e Ser135. Na imagem A é observado também dois *sticks* presentes nas extremidades, referentes ao resíduo Cis143 responsável pela oligomerização.

Figura 41- Ensaio de inibição da NS2b/NS3 de ZIKV com diversos inibidores de protease. (A) 1 hora. (B) 2 horas. (C) 3 horas. (D) 16 horas. MW- Marcador de peso molecular. - NS2b/NS3 sozinha desordenada. Z- NS2b/NS3 sozinha. Ap- Protease com aprotinina. Be- Protease com benzamidina. Le- Protease com leupeptina. PM- Protease com PMSF. BT- Protease com BTCl.

Figura 42- Ensaio de fluorescência da protease NS2b/NS3 de ZIKV com aprotinina. (A) Espectro de emissão de fluorescência de triptofano em diferentes concentrações de aprotinina. (B) Emissão de fluorescência máxima da proteína em presença de quantidades crescentes de aprotinina em relação a proteína em ausência do ligante. Curva de inibição da protease.

Figura 43- Interface de interação entre a protease NS2b/NS3 de ZIKV e aprotinina. A proteína NS3 está em cinza, com os resíduos do sítio catalítico em cinza escuro, a NS2b em coloração trigo e a aprotinina em verde. As ligações de hidrogênio estão representadas pelos pontilhados amarelos.

Figura 44- (A) Homodímero da NS2b/NS3 de ZIKV (5LC0) e modelo da interação entre o monómero desta protease obtido através de *docking* proteína-proteína. NS3 (cinza), NS2b (trigo), aprotinina (azul) e TRP50 (vermelho).

Figura 45- Cristal da protease NS2b/NS3 de ZIKV (dioxano 10%, sulfato de amônio 1,2 M, 0,1 M MES pH 6,1).

Figura 46- Expressão e purificação da proteína NS3 de ZIKV (ZIKV-NS3-Lub). A) Perfil

cromatográfico da purificação por gradiente de imidazol usando coluna de níquel. A linha vermelha refere-se a concentração de imidazol em porcentagem (0 a 500 mM). B) Análise por gel SDS-PAGE 15% da expressão e purificação utilizando coluna de níquel. Pel e Sup referem-se ao pellet e sobrenadante obtidos durante as etapas de lise e centrifugação, FT as proteínas que não interagiram com a coluna de níquel, e 1, 2 e 3 a cada um dos picos coletados. C) Perfil cromatográfico da purificação por coluna de troca aniônica. A linha vermelha refere-se a concentração de NaCl em porcentagem (0 a 1 M). D) Análise por gel SDS-PAGE 15% da purificação por coluna de troca aniônica, onde os números 1, 2 e 3 mostram os picos coletados.

Figura 47- Gráfico de atividade específica da protease NS2B/NS3 de ZIKV na presença substrato Benzoyl-Nle-Lys-Arg-Arg-AMC.

Figura 48- Gráfico do ensaio de Michaelis-Menten da protease de NS2B/NS3 de ZIKV na presença substrato Benzoyl-Nle-Lys-Arg-Arg-AMC.

Figura 49- Ensaio de inibição enzimática da protease NS2B/NS3 de ZIKV para cálculo de IC50 na presença das moléculas A) F0174-0048, B) F0816-0342, C) F3165-0782 e D) leupeptina.

Figura 50- Geis de expressão SDS-PAGE 15% da protease NS2b-NS3 de DENV. Expressão após transformação nas cepas Lemo21 (A) e DE3 (B). PI é referente as amostras antes da indução com IPTG, Exp a fração insolúvel e solúvel juntas, enquanto Pel motra a insolúvel e Sob a solúvel separadamente.

Figura 51- (A) Cromatograma da purificação por cromatografia líquida de afinidade a metal (Ni) da protease NS2b e NS3 de DENV. (A) A amostra é referente a expressão com a cepa BL21 DE3. A absorbância (mAU) é observada pela linha azul e a porcentagem de imidazol é vista pela linha amarela, vale lembrar que 100% se refere a concentração de 500 mM no tampão. (B) Gel SDS-PAGE 15% da purificação por coluna de Níquel, Sol é referente a fração solúvel, Pel a fração insolúvel, P1 e P2 são referente ao primeiro e segundo pico enquanto a, b, e c a diferentes frações dos picos.

Figura 52- Purificação por exclusão molecular da protease NS2b/NS3 de DENV. A coluna

utilizada para o experimento foi a S200.

Figura 53- Análise da proteína NS2b/NS3 de DENV em ultracentrífuga analítica por velocidade de sedimentação.

Figura 54- Fotos dos poços com cristais da placa de screening PACT++ feita com a proteína de NS2b/NS3 de DENV a 17 mg/mL em tampão tris 20 mM e NaCl 300 mM. A- 25% w/v Polyethylene glycol 1500; 100 mM SPG Buffer; pH 8,0 (A5) B-25% w/v Polyethylene glycol 1500; 100 mM TBG Buffer; pH 8,0 (A11). C-20% w/v Polyethylene glycol 6000; 100 mM MES; pH 6,0; 10 mM zinc chloride D- 20% w/v Polyethylene glycol 6,000, 100 mM TRIS; pH 8.0, 200 mM Ammonium chloride. E- 20% w/v Polyethylene glycol 6,000, 100 mM TRIS; pH 8.0, 200 mM Lithium chloride. F- 20% w/v Polyethylene glycol 3,350, 100 mM BIS-TRIS propane; pH 7.5, 200 mM Sodium fluoride.

Figura 55- Fotos dos poços com cristais da placa de screening JCSG++ feita com a proteína de NS2b/NS3 de DENV a 17 mg/mL em tampão tris 20 mM e NaCl 300 mM. A- 800 mM di-Sodium succinate; pH 7,0. B- 0,5 % w/v Jeffamine ED-2001; pH 7,0; 100 mM HEPES; pH 7,0; 1,1 M di-Sodium malanate; pH 7,0. C- 1% w/v Polyehylene glycol monomethyl ether 2000; 100 mM HEPES; pH 7,0; 1 M di-sodium succinate; pH 7,0.

Figura 56- Fotos dos poços com cristais da placa de screening JCSG++ feita com a proteína de NS2b/NS3 de DENV a 17 mg/mL em tampão tris 20 mM e NaCl 300 mM. A- 20 % w/v Polyethylene glycol 3,350, 200 mM tri-Potassium citrate; pH 8.3. B- 100 mM MES; pH 6,5; 1 M tri-sodium citrate. C- 2M Ammonium sulfate; 100 mM MES; pH 6,5. D- 1,26 M Ammonium sulfate; 100 mM TRIS; pH 8,5; 200 mM Lithium sulfate. E- 100 mM Sodium acetate; pH 4,5; 1 M di-Ammonium hydrogen phosphate. F- 100 mM MES; pH 6,5; 1,6 M Magnesium sulfate.

LISTA DE TABELAS

Tabela 1- Sequência de aminoácidos dos *primers* utilizados para amplificação da sequência gênica da protease nsP2 de MAYV.

Tabela 2- Características das cepas de E. Coli utilizadas durante a etapa da otimização da expressão das proteases de ZIKV, DENV, CHIK e MAYV.

Tabela 3- Lista de plasmídeos utilizados das proteases arbovirais e suas respectivas marcas de seleção.

Tabela 4 – Sequência dos substratos das proteases nsP2 de MAYV e CHIK.

Tabela 5- Porcentagem de estrutura secundária de Alphavirus e da deconvolução dos dados do espectro da protease nsP2 de MAYV

Tabela 6- Tempo de existência das ligações de hidrogênio da MD de CHIK com substrato CHIKs34.

Tabela 7- Tempo de existência das ligações de hidrogênio da MD de MAYV com substrato MAYVs23.

Tabela 8- Resíduos que compõe os subsítios S4' - S6 para os Alphavirus VEEV, CHIK e MAYV.

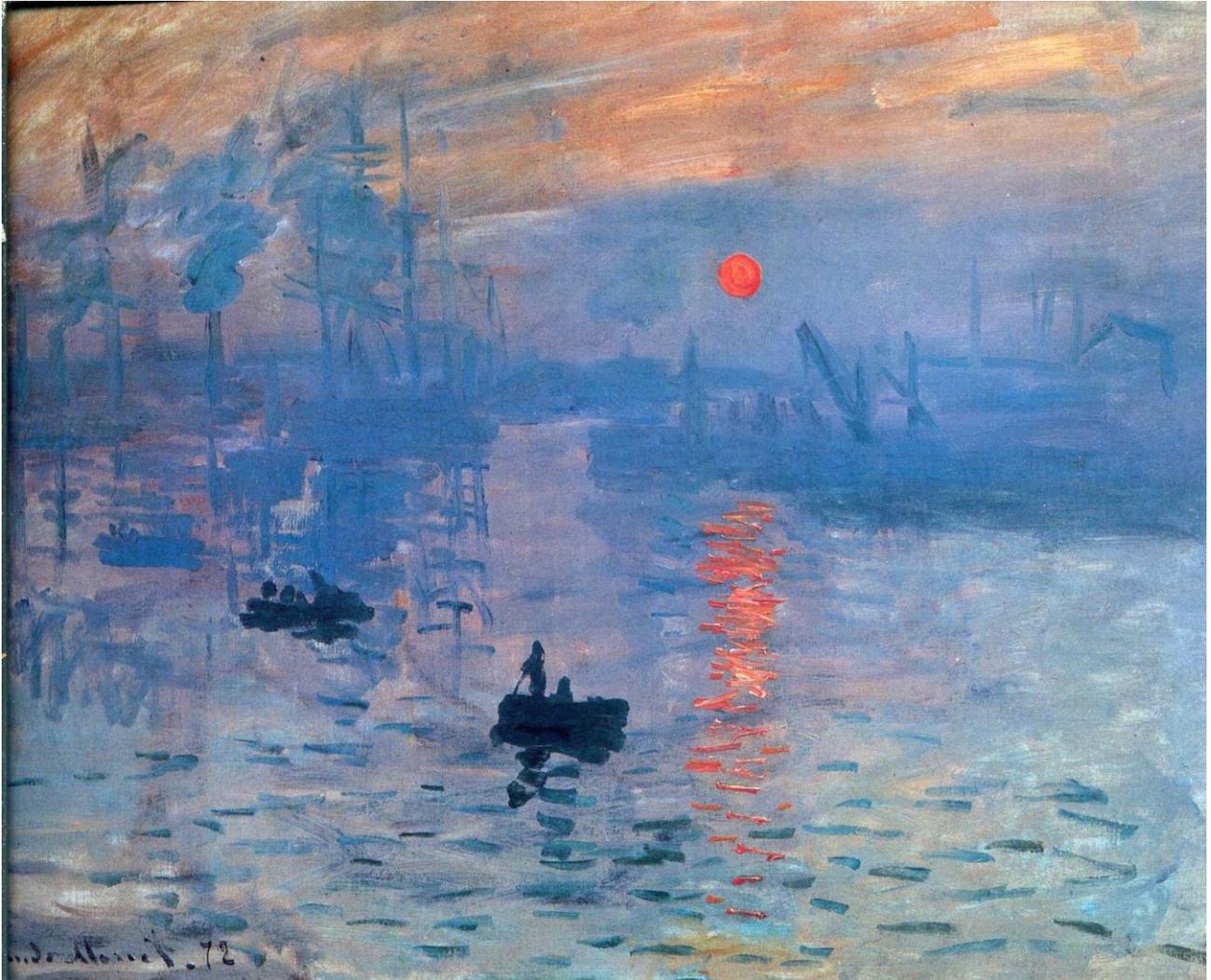
LISTA DE ABREVIACOES

Vrus da dengue (DENV)
Vrus da febre amarela (YFV)
Vrus da encefalite de St Louis (SLEV)
Vrus da encefalite japonesa (JEV)
Febre hemorrgica de Crimeia Congo (CCHF)
Encefalite equina do leste (EEL)
Vrus da zika (ZIKV)
Vrus chikungunya (CHIK)
Vrus mayaro (MAYV)
Open reading frame (ORF)
Vrus da floresta de Semliki (SFV)
Metiltransferase (Mtase)
Interface grfica do usurio (GUI)
Isopropyl β -D-1-thiogalactopyranoside (IPTG)
Dicrosmo circular (CD)
Protena fluorescente ciano (CFP)
Protena fluorescente amarela (YFP)
Ditiotreitol (DTT)
Protena de ligao de maltose (MBP)
Vrus Sindbis (SIN)
Motivo de especificidade a glicina (GSM)
Fluoreto de fenilmetilsulfonil (PMSF)
Black-Eyed Pea Trypsin Chymotrypsin Inhibitor (BTCI)
Vrus do Oeste do Nilo (WNV)

SUMÁRIO

1. Introdução	21
2. Justificativa	34
3. Objetivos	37
3.1. Objetivo geral	38
3.2. Objetivos específicos	38
3.2.1. Alphavirus	38
3.2.2. Flavivirus	38
4. Metodologia	39
4.1. Experimentos <i>in vitro</i>	40
4.1.1. Obtenção dos genes	40
4.1.2. Clonagem da protease nsp2 de MAYV	40
4.1.2.1. Amplificação	40
4.1.2.2. Digestão e ligação	41
4.1.2.3. PCR de colônia	42
4.1.3. Preparo de células competentes	42
4.1.4. Transformação	43
4.1.5. Otimização da expressão e purificação	43
4.1.5.1. NS2b/NS3 de ZIKV	46
4.1.5.2. NS2b/NS3 de DENV	47
4.1.5.3. nsP2 de MAYV	48
4.1.5.4. nsP2 de CHIK	49
4.1.5.5. Substrato de nsP2 de CHIK	50
4.1.6. Renaturação	50
4.1.6.1. Purificação em condições desnaturantes	50
4.1.6.2. Renaturação por diluição rápida	51
4.1.6.3. Análise de estrutura secundária por dicroísmo circular	51
4.1.7. Análise por gel SDS-PAGE 12%	52
4.1.8. Teste de solubilidade	53
4.1.9. <i>Western blot</i>	53
4.1.10. Testes de atividade e inibição	54
4.1.10.1. Chikungunya	54
4.1.10.2. Zika	54

4.1.10.2.1. Teste de inibição por gel SDS-PAGE 15% (zikv-ns3-bra)	54
4.1.10.2.2. Ensaio de atividade e inibição enzimática (zikv-ns3-lub)	54
4.1.10.2.2.1. Influência da concentração de enzima na atividade enzimática	55
4.1.10.2.2.2. Gráfico de Michaelis-Menten	55
4.1.10.2.2.3. Ensaio de inibição da atividade enzimática	56
4.1.11. Ensaio de fluorescência	57
4.1.12. Ultracentrifugação analítica	57
4.1.13. <i>Screening</i> e refinamento de cristalização	58
4.1.13.1. Zika	58
4.1.13.2. Dengue	59
4.1.13.3. Chikungunya	59
4.1.14. Coleta de dados e refinamento	59
4.2. Experimentos <i>in silico</i>	59
4.2.1. Modelagem por homologia	59
4.2.2. <i>Docking</i> proteína-peptídeo	60
4.2.3. Dinâmica molecular (MD)	61
4.2.4. <i>Docking</i> proteína-proteína	62
5. Resultados e discussão	64
5.1. Alphavirus	66
5.1.1. Experimentos <i>in vitro</i>	66
5.1.1.1. Mayaro	66
5.1.1.2. Chikungunya	76
5.1.2. Experimentos <i>in silico</i>	83
5.2. Flavivirus	98
5.2.1. Zika	98
5.2.1. Dengue	113
6. Conclusão	123
7. Referências Bibliográficas	126



Impression, sunrise
1872 48x63cm Oil on canvas
Claude Monet

1. Introdução

1. INTRODUÇÃO

Os arbovírus (*Arthropod-borne virus*) são uma classificação que compreende grupos virais cuja transmissão é realizada por vetores artrópodes infectando vertebrados, sendo na sua grande maioria vírus de RNA. Dentro dessa classificação, que não é taxonômica, algumas famílias que podem ser observadas são: Togaviridae, Flaviviridae, Bunyaviridae, Rhabdoviridae, Reoviridae, Orthomyxoviridae e Asfarviridae (CIOTA; KRAMER, 2010; HERNANDEZ; BROWN; PAREDES, 2014).

Este grupo está relacionado a uma grande quantidade de doenças em todo o mundo como: dengue, febre amarela (YFV), encefalite de St Louis (SLEV), encefalite japonesa (JEV), febre de oropouche, febre hemorrágica de Crimeia Congo (CCHF), meningoencefalites, encefalite equina do leste (EEL), dentre muitas outras (MELTZER, 2012). Especialmente no Brasil, em regiões como a amazônica, em torno de 200 arbovírus já foram isolados, sendo que mais de 30 são patógenos humanos (VASCONCELOS et al., 2001). Dentre esses, alguns arbovírus têm se mostrado de grande relevância, devido ao grande número de casos, como o vírus da dengue (DENV) e o zika (ZIKV), pertencentes ao gênero *Flavivirus* da família *Flaviviridae*, o chikungunya (CHIK) e o mayaro (MAYV), gênero *Alphavirus* da família *Togaviridae*.

Em relação à incidência destes arbovírus, estimativas dos números de infecções no mundo anualmente, por exemplo, somente de dengue, alcançam 390 milhões (BHATT et al., 2013). No Brasil estes vírus também têm se mostrado como um grande problema de saúde pública, de acordo com o boletim epidemiológico do ministério da saúde, no ano de 2019 o número de casos prováveis de DENV chegou a mais de 1,5 milhões, de CHIK a 130 mil e ZIKV a mais de 10 mil, e os números de óbito foram 754, 95 e 3, respectivamente (BRASIL, 2019a). O ZIKV, além do grande número de casos, foi o agente causador de uma epidemia em 2015-2016 no Brasil responsável por diversas alterações congênitas afetando 3483 recém-nascidos constatados entre o período entre 2015 e 2019, sendo que 1927 casos foram confirmados somente no ano de 2016 (BRASIL, 2019b). Em relação ao MAYV, em estudos nas regiões rurais da Amazônia, análises com anticorpo anti-MAYV IgG foram positivas em 119 de 270 amostras de voluntários (ABAD-FRANCH et al., 2012), indicando que este pode apresentar uma grande incidência, todavia com um número real de casos altamente subnotificado. Além disso, as arboviroses apresentam grandes custos econômicos, apenas no ano de 2016 quase R\$ 1,5 bilhões foram aplicados ao combate ao vetor e R\$ 394 milhões em custos

médicos diretos, além de gastos indiretos (TEICH; ARINELLI; FAHHAM, 2017).

Quanto aos artrópodes transmissores dos vírus que estão relacionados com doenças em vertebrados, uma grande diversidade de grupos tem sido estudada, desde as ordens Blattodea, Hymenoptera, Orthoptera, pertencentes à classe Insecta, até ordens como Decapoda da classe Crustacea, entre outros. Entretanto, ordens como Diptera (classe Insecta), representadas pelos mosquitos e moscas, e Ixodida (classe Arachnida), representada pelos carrapatos, apresentam uma grande variedade de espécies transmissoras destes vírus (CALISHER; HIGGS, 2018). Dentre os arbovírus com grande presença no Brasil, DENV, ZIKV e CHIK têm como principal agente transmissor mosquitos do gênero *Aedes*, cujos dois representantes principais são o *Aedes aegyptis* e o *Aedes albopictus*. Como pode ser observado na distribuição destes dois vetores na figura 1, ambos apresentam distribuição em quase todos os continentes (KRAEMER et al., 2015; MESSINA et al., 2016). O MAYV, dentre estes, apresenta maior diferenciação em relação ao vetor, seu principal agente de transmissão é o mosquito *Haemagogus janthinomys*, todavia ele já foi isolado em outros gêneros como *Culex* e *Aedes* (ACOSTA-AMPUDIA et al., 2018), o que tem levantado a discussão de sua capacidade de disseminação e prejuízo à saúde pública.

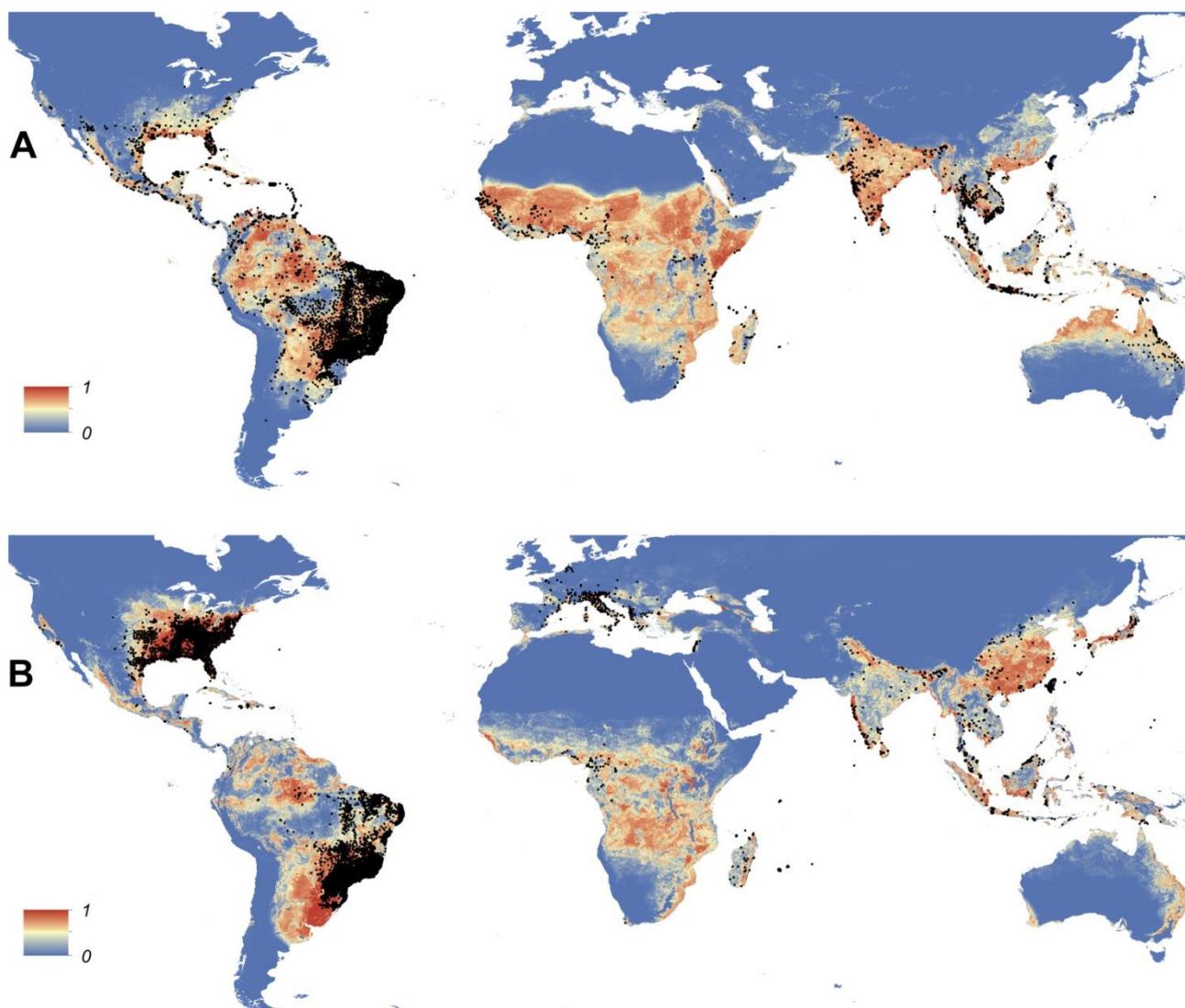


Figura 1- Predição da área de distribuição dos vetores (A) *Aedes aegypti* e (B) *Aedes albopictus* baseado em banco de dados de ocorrência (KRAEMER et al., 2015).

Os *Flavivirus* são vírus envelopados que possuem um genoma de RNA fita simples com polaridade positiva o qual contém cap na região 5' e um tamanho aproximado de 11 kb (BOLLATI et al., 2010; FERNANDEZ-GARCIA et al., 2009). Este é traduzido para uma poliproteína a partir de uma única *open reading frame* (ORF), a qual é composta por um total de 10 genes, sendo 3 estruturais (C, PrM e E) e 7 não estruturais (NS1, NS2A, NS2B, NS3, NS4A, NS4B e NS5). A separação dessas proteínas ocorre pela ação de proteases endógenas, como furinas e signalases, e pela protease viral denominada NS3. A composição dos *Flavivirus* pode ser observada na figura 2 (CHONG et al., 2019; FAHEEM et al., 2019). A montagem final deste vírus apresenta um envelope contendo 90 homodímeros das proteínas estruturais E e M que se organizam em uma simetria icosaédrica. Estas proteínas E são responsáveis pela entrada do vírus na célula através

do processo de endocitose mediada por receptor, tem sido observado que os Flavivirus podem usar múltiplos receptores para a entrada na célula. O pH baixo realiza uma alteração conformacional nesta proteína estrutural formando 60 trímeros, conduzindo para a fusão na membrana da célula hospedeira e liberação do material genético no interior celular (MARZINEK et al., 2016; SMIT et al., 2011).

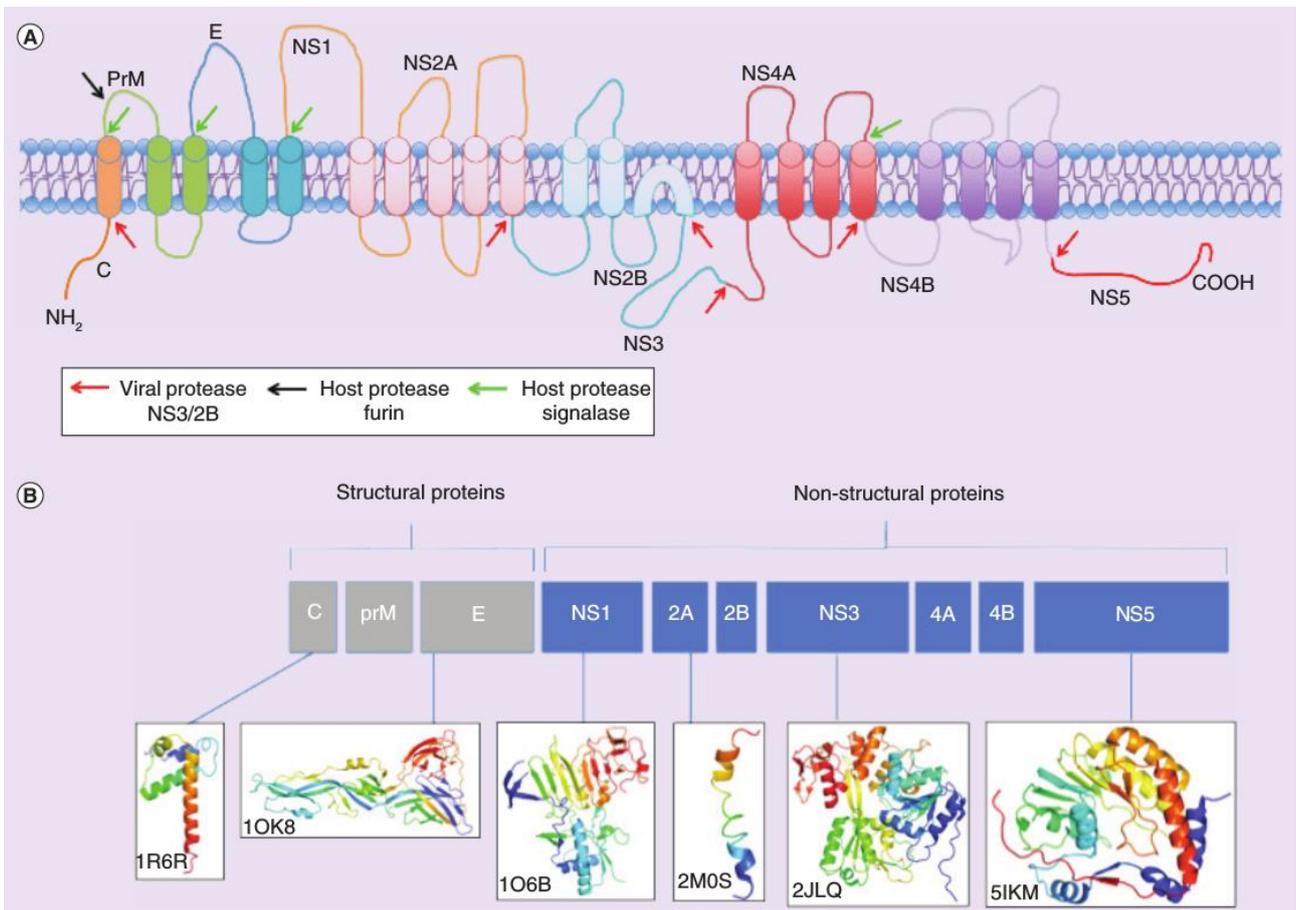


Figura 2- Composição da poliproteína dos Flavivirus. A) Representação das regiões transmembrana da poliproteína indicando com setas os locais de clivagem. B) Divisão das proteínas estruturais e não estruturais com exemplos de imagens de estruturas resolvidas do Flavivirus DENV (FAHEEM et al., 2019).

Os Alphavirus formam um gênero que compreende vírus envelopados icosaédricos com um genoma de aproximadamente 12 Kb composto de uma fita única de RNA com duas ORFs com cap 5' e cauda poli A. A primeira ORF codifica quatro proteínas não estruturais (nsP1, nsP2, nsP3 e nsP4), produzida nas fases iniciais da infecção. Após a clivagem e separação dessas proteínas não estruturais, cinco proteínas estruturais (C, E3, E2, 6k e E1) são traduzidas a partir de um RNA subgenômico gerado através de uma

segunda ORF na fase tardia da infecção. A protease nsP2 possui um papel fundamental no processamento e separação das proteínas não estruturais, sendo esta uma etapa fundamental no processo de replicação viral. (HERNANDEZ; BROWN; PAREDES, 2014; LEUNG; NG; CHU, 2011; RUPP et al., 2015). Estas informações estão representadas na figura 3. A entrada do vírus na célula se dá também por endocitose mediada por receptor pela formação de poços e vesículas de clatrina, através do contato da proteína E2 tendo como um dos possíveis receptores a proteína transmembrana NRAMP (STILES; KIELIAN, 2011). Entretanto, em estudos têm-se observado a penetração do vírus mesmo na ausência de exposição a pH baixo, além da presença do vírus da floresta de Semliki (SFV) tanto em vesículas cobertas por clatrina quanto em descobertas, o que tem levado à hipótese da entrada também por vias alternativas (HASE; SUMMERS; COHEN, 1989; HERNANDEZ; LUO; BROWN, 2001).

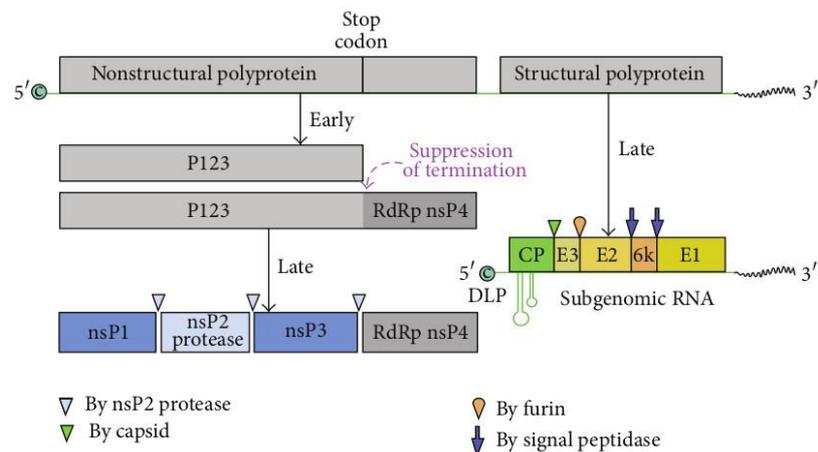


Figura 3- Composição da poliproteína de Alphavirus com esquematização dos momentos de processamento do genoma durante o a infecção viral. As setas coloridas marcam as proteínas responsáveis por processar cada região (HERNANDEZ; BROWN; PAREDES, 2014).

Dentre as proteínas que compõe a poliproteína viral as proteases têm se mostrado excelentes alvos de estudo visando a busca por inibidores para possíveis tratamentos do vírus. De acordo com a classificação da *Enzyme Commission* (EC), são pertencentes as classes das hidrolases, as quais catalisam a quebra de ligações químicas utilizando moléculas de água. Além disso, estas apresentam subclasses que são divididas entre grupos que hidrolisam cadeias peptídicas próximas a região aminoterminal (*aminopeptidases*), carboxiterminal (*carboxypeptidases*), ambas conhecidas como

exoproteases, e as que clivam na porção central (proteínases ou *endoproteases*) (GARCIA- CARREON; TORO; NAVARRETE, 1997). Este último grupo por sua vez é classificado em pelo menos sete classes que apresentam resíduos com diferentes características químicas responsáveis pela clivagem, sendo estas: serino, cisteíno, aspártico, metaloproteases e mais recentemente descritas, asparagino, glutâmico e treonino proteases (RAWLINGS et al., 2018). Entre essas, duas classes de grande diversidade e importância são as serino e cisteíno proteases.

Ambos as classes apresentam mecanismos com várias características em comum. Primeiramente a presença de um resíduo que será responsável pelo ataque nucleofílico e outro resíduo básico, o qual atuará como acceptor do hidrogênio. A serina ou cisteína atacam o carbono da carbonila do substrato formando um estado de transição com um intermediário tetraédrico caracterizado por um oxianion estabilizado por ligações de hidrogênio na cavidade do oxianion. Em seguida, ocorre a quebra da ligação peptídica, a acilação da enzima e a liberação do primeiro produto ($R'NH_2$). Uma molécula de água reage com o carbono da carbonila liberando o segundo produto ($RCOOH$) da enzima em uma etapa chama de desacilação e assim regenera a enzima (CSTORER; MÉNARD, 1994; JAMES C. POWERS et al., 2002).

As serino proteases compreendem um terço de todas as enzimas proteolíticas conhecidas e estão representadas em todos os domínios (*Bacteria*, *Archeae* e *Eukarya*), além de estarem presentes também nos vírus. No banco de dados de proteases MEROPS, este grupo apresenta 12 clãs e 38 famílias, sendo que algumas famílias ainda não foram atribuídas a nenhum clã ("MEROPS - the Peptidase Database", [s.d.]). Este grupo de proteases é caracterizado por ter uma serina como principal resíduo do seu mecanismo catalítico, cujo grupamento hidroxila é responsável pela realização do ataque nucleofílico ao carbono da carbonila da ligação peptídica do substrato (WALKER; LYNAS, 2001). O mecanismo mais conhecido das serino proteases utiliza outros dois resíduos catalíticos, uma histidina, responsável por remover um próton da serina, e um aspartato, que estabiliza e orienta a His através de uma ligação de hidrogênio (HEDSTROM, 2002). Um esquema desse mecanismo pode ser observado na figura 4. Todavia, vale lembrar que existem variações neste mecanismo, tal que algumas serino proteínases apresentam díades catalíticas, outros resíduos catalíticos como lisinas e glutamatos, e variações nos sítios ativos, variações estas que garantem diferentes graus de especificidade (DI CERA, 2009).

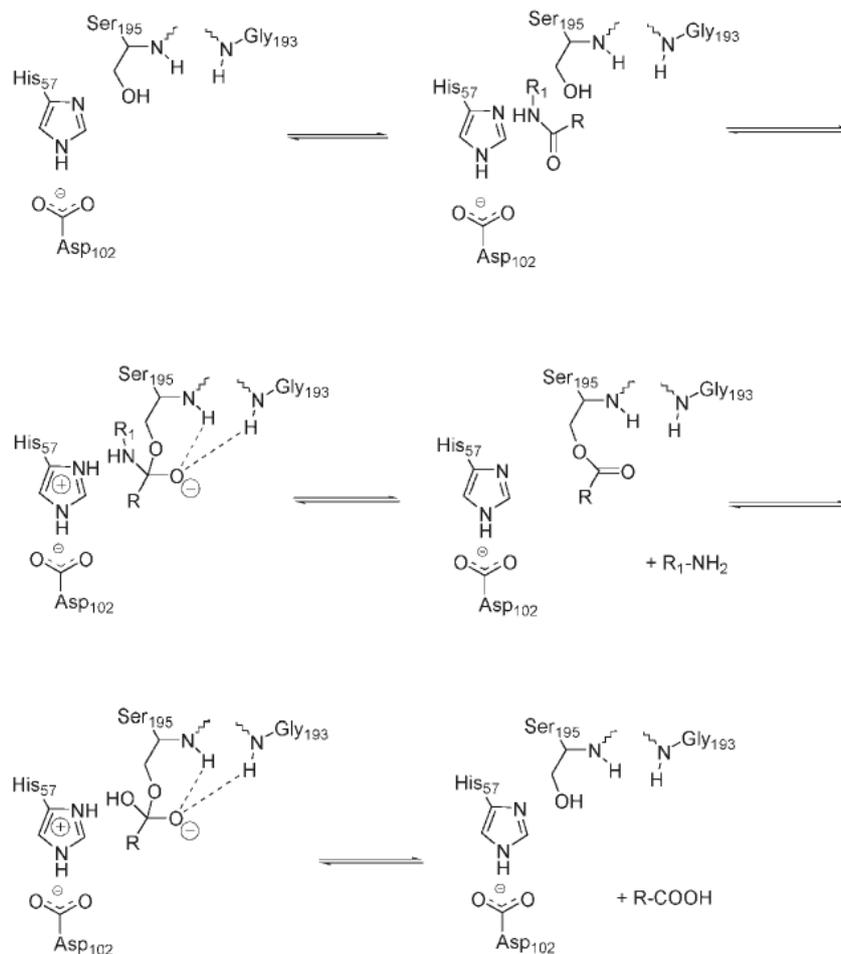


Figura 4- Mecanismo de catálise de serino proteases (WALKER; LYNAS, 2001).

As cisteíno proteases, por sua vez, de acordo com o banco de dados MEROPS apresentam 9 clãs e 78 famílias, mas há famílias ainda não atribuídas a clãs (“MEROPS - the Peptidase Database”, [s.d.]). Este grupo, também presente em todas as classes dos organismos vivos, possui famílias altamente abundantes, como a da papaína e da USP-14 (ubiquitin-specific peptidase 14). Na papaína, cisteíno protease com o mecanismo de catálise melhor compreendido, observa-se uma tríade catalítica com uma cisteína, uma histidina e uma asparagina. Todavia, existe variações em várias famílias dessa proteína (VERMA; DIXIT; PANDEY, 2016).

As proteases virais têm grande importância no ciclo de vida viral, estas são responsáveis por catalisar a quebra da poliproteína viral permitindo a liberação em diversas proteínas e, em alguns casos, o processamento de pré-proteínas do capsídeo para proteínas maduras. Para estas proteases, em geral, não são conhecidos homólogos

celulares, que somente apresentam baixa identidade com sua sequência, mesmo em casos onde o dobramento proteico é similar, o que está relacionado em parte com a sua especificidade distinta aos substratos (TONG, 2002). Este grupo de proteínas tem sido assim alvo de diversos fármacos, sendo que os principais exemplos de êxito são inibidores de proteases do vírus da imunodeficiência humana tipo 1 (HIV-1) para os quais já se tem aprovado uma série de drogas (SIENCZYK, 2013).

Em se tratando de endoproteases arbovirais, a da hepatite C (HCV), a do vírus do oeste do Nilo (WNV) e até mesmo as dos sorotipos da DENV já têm sido utilizadas em testes e experimentos com inibidores de proteases (BEHNAM et al., 2015; HARTMAN et al., 2015). Todavia, algumas destas proteases, como a NS3 dos Flavivirus ZIKV e nsP2 dos Alphavirus CHIK e MAYV, carecem de estudos estruturais mais aprofundados para maior conhecimento de sua biologia molecular visando possíveis alvos para novos fármacos.

As estruturas de NS3 presente nos vírus do gênero Flavivirus, vírus DENV (códigos do PDB 2WHX, 2FOM, 3U1I e 4M9F) e ZIKV (códigos PDB 5GXJ, 5H4I, 5LC0, 5T1V, 5TFN, 5YOD, 5YOF, 6JPW, 5H6V, 5GJ4, 5GPI, 5TFO, 5ZMQ, 5ZMS, 5ZOB e 6UM3), apresentam domínios com funções diferentes. O domínio N-terminal apresenta uma função de serino protease, há uma espécie de *linker* e o C-terminal tem função de RNA helicase, podendo essa região ser dividida em subdomínios com uma helicase dependente de ATP e uma RNA trifosfatase. Na proteína NS3 do vírus de sorotipo 4 (DENV4), por exemplo, que possui um total de 618 resíduos de aminoácidos, o domínio com atividade protease vai dos resíduos 1 ao 168 (esta região é rica em folhas- β), o *linker* do 169 ao 179 e o domínio RNA helicase do 180 ao 618 (apresenta várias estruturas secundárias, tanto hélices α quanto folhas β), sendo que esses domínios são altamente flexíveis entre si, por causa do *linker* (LUO et al., 2008, 2010; NOBLE et al., 2012).

Quanto ao sítio ativo do domínio serino protease da proteína NS3, este é marcado pela presença de uma tríade catalítica, da qual fazem parte três resíduos de aminoácidos: uma histidina na posição 51 (His51), um ácido aspártico na posição 75 (Asp 75) e uma serina na posição 135 (Ser135). Entretanto, para que a proteína fique ativa é necessário também uma ligação não covalente com uma região da NS2B, uma proteína transmembrana que possui provavelmente três hélices fazendo passagem pela membrana do retículo endoplasmático. Essa última proteína parece se enrolar em torno da NS3, fazendo uma estrutura “tipo cinto” (“*belt-like*”), e seu grampo formado por folhas β (*beta hairpin*) atua também no sítio ativo (NOBLE et al., 2012). Além disso, a proteína

NS3 apresenta regiões como GLFG, um loop hidrofóbico que tem se mostrado importante para realização de contatos com a membrana (LUO et al., 2010). Um exemplo dessa estrutura pode ser visualizado na figura 5.

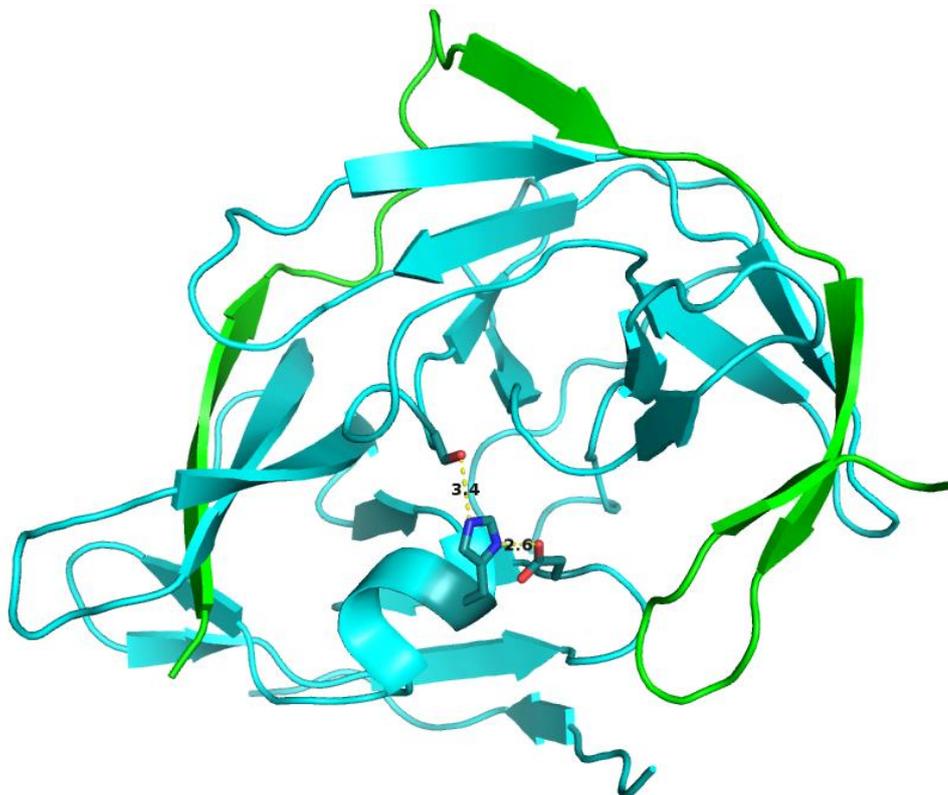


Figura 5- Estrutura da serino protease NS3 de DENV3 (3U1I) (ciano). A NS2b (verde) necessária para a atividade da proteína se posiciona em volta da estrutura em forma de cinto. Os resíduos do sítio ativo são visualizados em *stick* (His51, Asp 75 e Ser135) e com as distâncias entre estes resíduos em Å representadas pelo pontilhado amarelo.

A região do sítio catalítico dessa protease apresenta algumas regiões importantes para o reconhecimento do substrato. Os resíduos P2 e P1 do substrato são compostos por resíduos básicos variando entre os diversos sítios de Flavivirus entre Lisinas e Argininas, resíduos como o D129 e F130 compõe o subsítio S1 o qual interage com o P1, enquanto o N152 e G153 mostram-se importantes no reconhecimento do P2. O P1' em geral é composto por resíduos pequenos como glicina, serina ou treonina, sendo acomodado no subsítio S1' em que ocorre em uma fissura entre a fita β 1 e uma pequena hélice formada pelos resíduos 50 – 53. Outra região importante é a cavidade do oxiânion, composto pelos resíduos G133 e S135, a qual não é sempre formada na estrutura sem o substrato (ALESHIN et al., 2007; NOBLE et al., 2012; SALAEMAE et al., 2010;

SHIRYAEV; STRONGIN, 2010).

Por sua vez, a protease nsP2, presente nos vírus do gênero Alphavirus CHIK (códigos do PDB 3TRK e 4ZTB) e MAYV (não possui estrutura resolvida no banco de dados PDB), apresenta um conjunto de domínios semelhantes a NS3, ou seja, primeiramente, na porção N-terminal, há domínios de RNA helicase, nucleosídeo-trifosfatase e uma RNA trifosfatase. A porção que se localiza no C-terminal apresenta dois subdomínios, o primeiro é a protease que composta por 7 hélices α e 4 fitas β , onde se localizam os resíduos com compõe a díade catalítica (Cys478 e His548). O outro subdomínio apresenta similaridade com a super-família das metiltransferases (MTase) dependentes de S-adenosil-metionina (SAM), todavia acredita-se que esta é biologicamente inativa, ou seja, sua função é pouco entendida. Este último subdomínio apresenta estrutura secundária com 5 hélices α e 1 folha β composta por 7 fitas (NARWAL et al., 2018; RUSSO et al., 2010; RUSSO; WHITE; WATOWICH, 2006).

O estudo da proteína nsP2 de alphavírus havia demonstrado estas proteínas como uma cisteína-protease tipo papaína, com uma díade catalítica no sítio ativo (composto por uma cisteína e uma histidina), e capaz de reconhecer diferentes sítios de clivagem (RUSSO; WHITE; WATOWICH, 2006). Entretanto, estudos estruturais da nsP2 de CHIK parecem indicar de fato uma díade, porém com a possibilidade de intercambiar a cisteína (Cys478) por uma serina (Ser482), ou seja, análises com a proteína mutada para somente um dos resíduos apresentava função normal, mas no caso da mutação de ambos resíduos havia perda de função. Além disso, estes novos dados têm alterado a classificação dessa proteína anteriormente tida como uma cisteína-protease tipo papaína, pois além da díade intercambiável, resíduos importantes como um triptofano com importante função catalítica no mecanismo da papaína, na nsP2 apresentam uma orientação que inviabiliza sua participação na catálise. Todavia, é necessário analisar se as proteases de outros Alphavirus também apresentam essas características da CHIK, sendo assim, estudos posteriores se mostrariam de extrema relevância para corroborar os dados encontrados, pois o entendimento destes mecanismos catalíticos é de importância para obtenção de compostos inibidores dessas proteases (SAISAWANG et al., 2015b). A estrutura dessa proteína pode ser observada na figura 6.

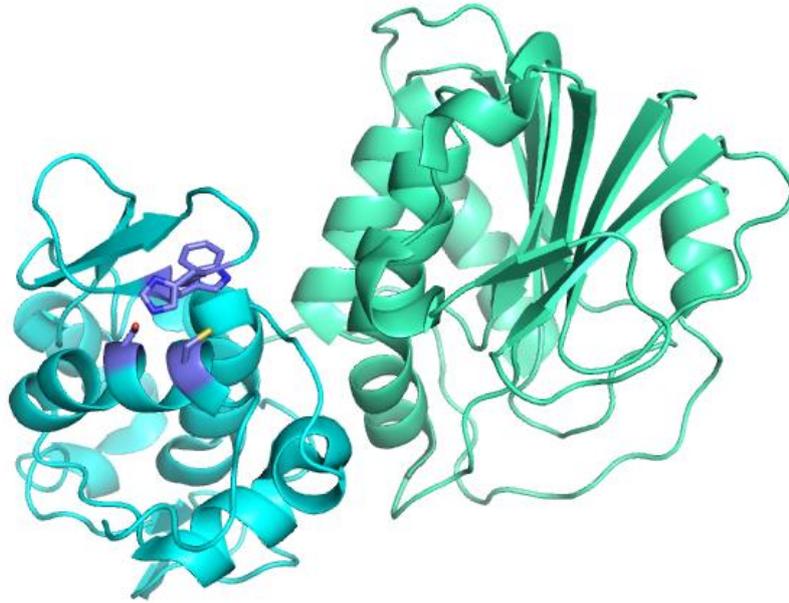


Figura 6- Estrutura da nsP2 de CHIK (3TRK). Domínio protease (ciano) com resíduos da díade catalítica e alguns do sítio ativo em *stick* (roxo), e domínio MTase (verde).

Em relação ao reconhecimento do substrato por esta proteína, tem-se sugerido que o motivo de especificidade a glicina (GSM) tenha um papel importante neste evento. Este motivo está presente em dois grandes clãs de cisteinoproteases: CA, exemplificados nas enzimas Tymovirus endopeptidase e *Ubiquitin-specific peptidase 14*, e CE, *African Swine Fever virus processing peptidase* e *Adenain*. A composição desses motivos em geral apresenta resíduos aromáticos volumosos que interagem com essa glicina. Nos Alphavirus este sítio de contato entre substrato e proteína está posicionado entre os subdomínios protease e Mtase. Outra característica desta região é a presença de um loop flexível entre as fitas $\beta 1$ e $\beta 2$ da protease, o qual atua possivelmente como um portão controlando a entrada do substrato, dentre os resíduos que compõe este está a His548, um dos que formam a díade catalítica (GOLUBTSOV; KÄÄRIÄINEN; CALDENTEY, 2006; NARWAL et al., 2018; RUSSO et al., 2010; RUSSO; WHITE; WATOWICH, 2006).

Como forma de exemplificar o potencial dos inibidores de protease pode-se citar o caso dos inibidores de protease aspártica de HIV, a qual é de grande importância para a maturação viral. A primeira molécula aprovada com esta função foi o saquinavir, que ocorreu no ano de 1995 nos EUA, entretanto atualmente conta-se com um total de 10 moléculas aprovadas, das quais, as que são categorizadas como de segunda geração, tem como base *design* cepas resistentes aos primeiros inibidores (LV, 2015;

WEBER,2021).

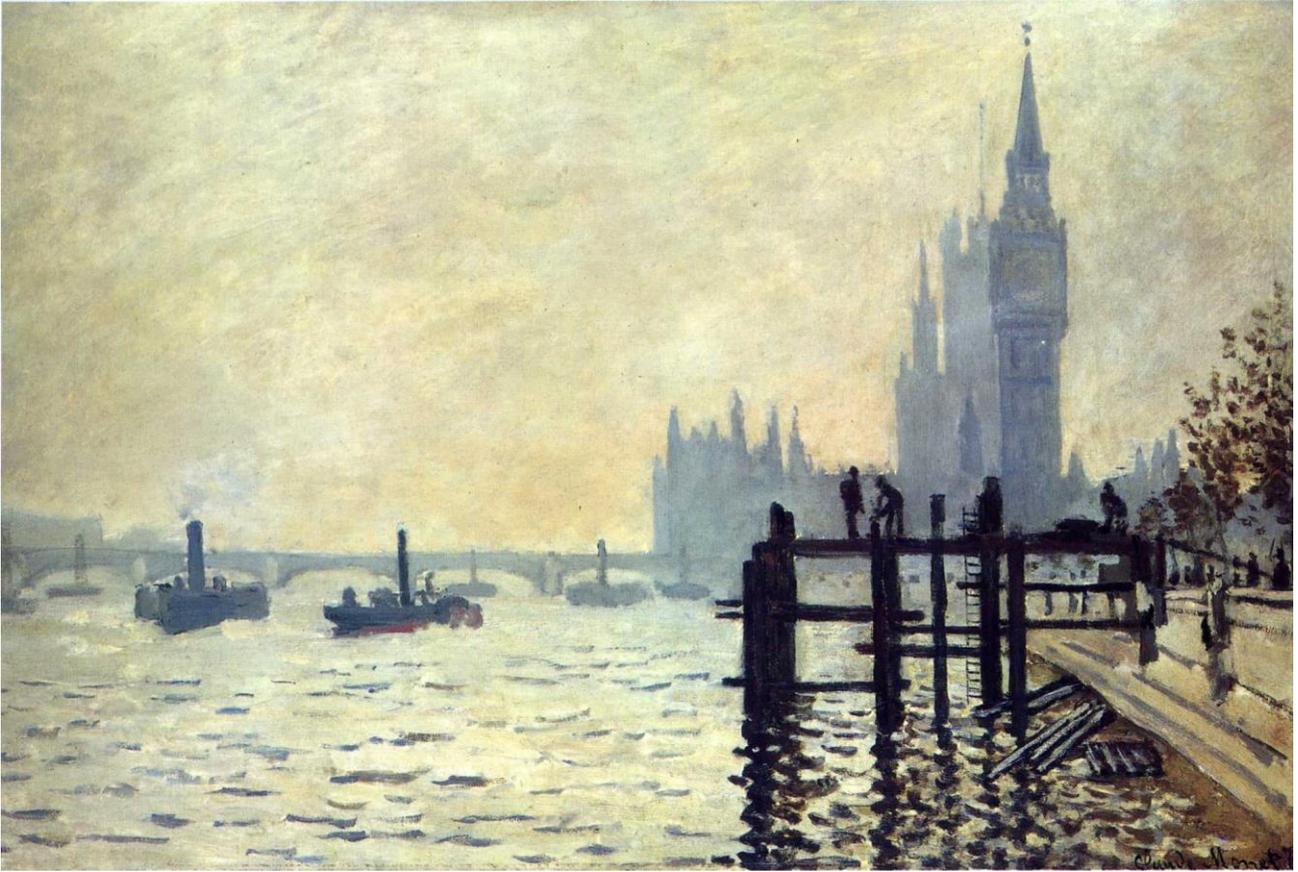
Recentemente a importância dos inibidores de protease tem sido amplamente demonstrada durante a pandemia de COVID-19, a qual já foi responsável por mais de 350 milhões de casos e mais de 5 milhões de morte em todo o mundo. Dentre esta grande lista de compostos estudados buscando inibir diferentes alvos proteicos já se encontram bastante avançados nos testes de aprovação para terapia antiviral como o rendesivir que atua inibindo a RNA polimerases dependentes de RNA (KOKIC. 2021), e o Molnupiravir, que se incorpora no RNA viral impedindo a replicação (WHITLEY, 2021). Entretanto, os inibidores de protease tem sido outro grupo amplamente estudado, apresentando um total de 24 patentes para a protease 3CLpro (CHIA, 2022), tendo como exemplo o PAXLOVID™ que associa um inibidor de protease PF-07321332 e o ritonavir o qual se encontra em fase 2/3 de testes, o qual ainda não possui registro em órgãos reguladores como Agência Nacional de Vigilância Sanitária (Anvisa) e U. S. Food and Drug Administration (FDA), entretanto já tem sido realizados pedidos de uso emergencial deste medicamento (BRASIL, 2021).

Várias são as formas de selecionar novos fármacos que funcionam como inibidores para as proteases virais, alguns exemplos são: abordagens baseadas em estruturas, baseadas em fragmentos, reposicionamento de medicamentos, exploração de produtos naturais, *design* de inibidores híbridos, métodos *in silico*, triagem de alto desempenho (HTS), entre outros (LOPEZ et al., 2016). Um tipo de abordagem que tem apresentado bons resultados na triagem e escolha de moléculas como fármacos é o *fragment screening*. Esta técnica consiste em utilizar uma biblioteca de fragmentos de compostos com baixo peso molecular, em geral até 250 Da, com objetivo de gerar moléculas para serem usados como fármacos (BAKER, 2013; MURRAY; VERDONK; REES, 2012). Dentre as técnicas para análise dos fragmentos em interação com a proteína alvo a cristalografia de raios X tem se mostrado altamente relevante. Sua realização requer o uso desses fragmentos para o *soaking* em cristais da proteína alvo ou para a co-cristalização. Além desta, outras análises biofísicas podem ser utilizadas complementarmente como a ressonância plasmônica de superfície (SPR), e a espectrometria de massas e calorimetria, entre outros (ERLANSON, 2012). Várias moléculas já têm sido testadas através dessa metodologia apresentando bons resultados (MURRAY et al., 2015; SMITH et al., 2015) e muitos desses compostos selecionados a partir dessa metodologia já se encontram em testes clínicos (BAKER, 2013).

Outro método que tem apresentado grande importância na descoberta de novos

fármacos é o de análises *in silico*. Apesar destes métodos não substituírem as avaliações biológicas as quais os compostos precisam ser submetidos, estes incrementam as chances de sucesso de identificar novos compostos por permitir explorar o espaço químico de forma mais completa (LAVECCHIA; DI GIOVANNI, 2013; LOPEZ et al., 2016). Devido a essa característica estas técnicas são amplamente utilizadas nos estágios iniciais dos processos de busca por novas drogas. Além disso, estas análises permitem reduzir o número de compostos que seriam sintetizados e testados por diversas ferramentas *in vitro*, o que reduz significativamente os custos da busca por novas moléculas (GIMENO et al., 2019).

Sendo assim, diversas técnicas *in silico* e *in vitro* têm sido utilizadas conjuntamente visando otimizar processos e/ou diminuir custos. Além disso, elas podem ser usadas conjuntamente para prever modelos biológicos ou criar hipóteses que visem explicar dados observados em experimentos laboratoriais (LI et al., 2018). Dessa forma, o presente trabalho objetiva o uso de ferramentas *in silico* e *in vitro* para análises estruturais das proteases dos arbovírus DENV, ZIKV, CHIK e MAYV, além das interações destas com diversos ligantes, visando uma melhor compreensão das interações dos inibidores com seus respectivos substratos.



The Thames below Westminster

1871 72x47cm oil/canvas

Claude Monet

2. Justificativa

2. JUSTIFICATIVA

Os arbovírus têm apresentado grande incidência em todo o mundo e inclusive no Brasil. Quanto aos vírus levantados para este trabalho, no caso do DENV, estima-se que no mundo ocorram anualmente 22000 mortes em decorrência dele e em nosso país aproximadamente 1,5 milhões de casos foram reportados somente no ano de 2013 (CHRISTO, 2015), no ano de 2021 até a semana epidemiológica 39 ocorreram aproximadamente 500.000. O CHIK tem afetado pessoas na Índia, sudeste de Ásia, Europa e Américas, tanto central quanto sul, alcançando somente na ilha de Reunião 300000 casos, e no Brasil, só em 2021, foram quase 85.794 casos confirmados até a semana epidemiológica 39 (ABRAHAM et al., 2016; GÉRARDIN et al., 2016; BRASIL, 2015a). O ZIKV no Brasil têm sido correlacionado com um grande aumento no número de casos de microcefalia e no ano de 2021 ocorreram 5.361 casos de infecção (BRASIL, 2015b; BRASIL, 2016). Por último, todos os vírus citados anteriormente são transmitidos pelo mesmo vetor, o mosquito *Aedes aegypti*, que está amplamente distribuído em todas as regiões do Brasil.

Esse grande número de casos em várias partes do mundo se deve ao fato de que grande parte da população mundial se encontrar em áreas de risco, ou seja, em locais favoráveis ao desenvolvimento do vetor (BARRETO; TEIXEIRA, 2008). Assim, é de grande relevância estudos que gerem conhecimentos relacionados a esses patógenos virais.

Além disso, estudos estruturais de proteases virais são de grande importância, pois ao analisarmos essas estruturas podemos compreender de modo mais aprofundado como ocorrem as interações dessas proteínas com os seus substratos, o que nos permite entender melhor o processo de ação viral na infecção da célula hospedeira. Outro ponto que demonstra a relevância do presente trabalho é o fato do MAYV não apresentar nenhuma estrutura resolvida e depositada no banco de dados PDB.¹

Quanto ao estudo de interação das proteínas com inibidores ou ligantes, esta é uma área de suma importância, ainda mais se tratando de inibidores de proteínas de um patógeno com tamanha incidência. A partir desses estudos é possível entender como

1 O ZIKV, que a dois anos atrás não apresentava nenhuma estrutura depositada no PDB, devido ao surto e a grande divulgação nas mídias nacionais e internacionais, atualmente conta com 126 estruturas, sendo que 16 são da protease NS3.

ocorre a inibição a nível estrutural, e isso, por sua vez, permite buscar fármacos específicos para combater esse vírus.

Quanto a inibidores virais sendo utilizados como fármacos, isso já é uma realidade contra patógenos do vírus da AIDS (HIV), onde até o ano de 2010, um total de 9 moléculas já haviam sido aprovadas para uso clínico (WENSING; VAN MAARSEVEEN; NIJHUIS, 2010). Isto demonstra que a seleção de inibidores para proteases virais é uma realidade possível.

Estas doenças também apresentam grande incidência em países como o Brasil, não apresentando muitos casos de infecção nos países mais ricos como Estados Unidos, com 435 casos em 2020 (CDC, 2020), e na Europa. Dessa forma, estas doenças negligenciadas precisam ser estudadas por pesquisas nacionais visando o melhor entendimento da doença.

Por fim, a busca de moléculas que realizam a inibição da protease desses vírus, acarretando a parada do ciclo de infecção viral, é de grande importância devido ao fato de atualmente não existir nenhum tipo de tratamento para essas infecções virais, as quais são tratadas somente com medicamentos para redução dos sintomas da contaminação. Dessa forma, este e outros estudos que abarquem este assunto se mostram extremamente relevantes.



The Avenue
1878 oil/canvas
Claude Monet

3. Objetivos

3. OBJETIVOS

3.1. OBJETIVO GERAL

O objetivo do presente trabalho é realizar estudos estruturais e funcionais das proteases dos Flavivirus DENV e ZIKV, e dos Alphavirus CHIK e MAYV, bem como da interação destes com ligantes variados através de ferramentas *in silico* e *in vitro*.

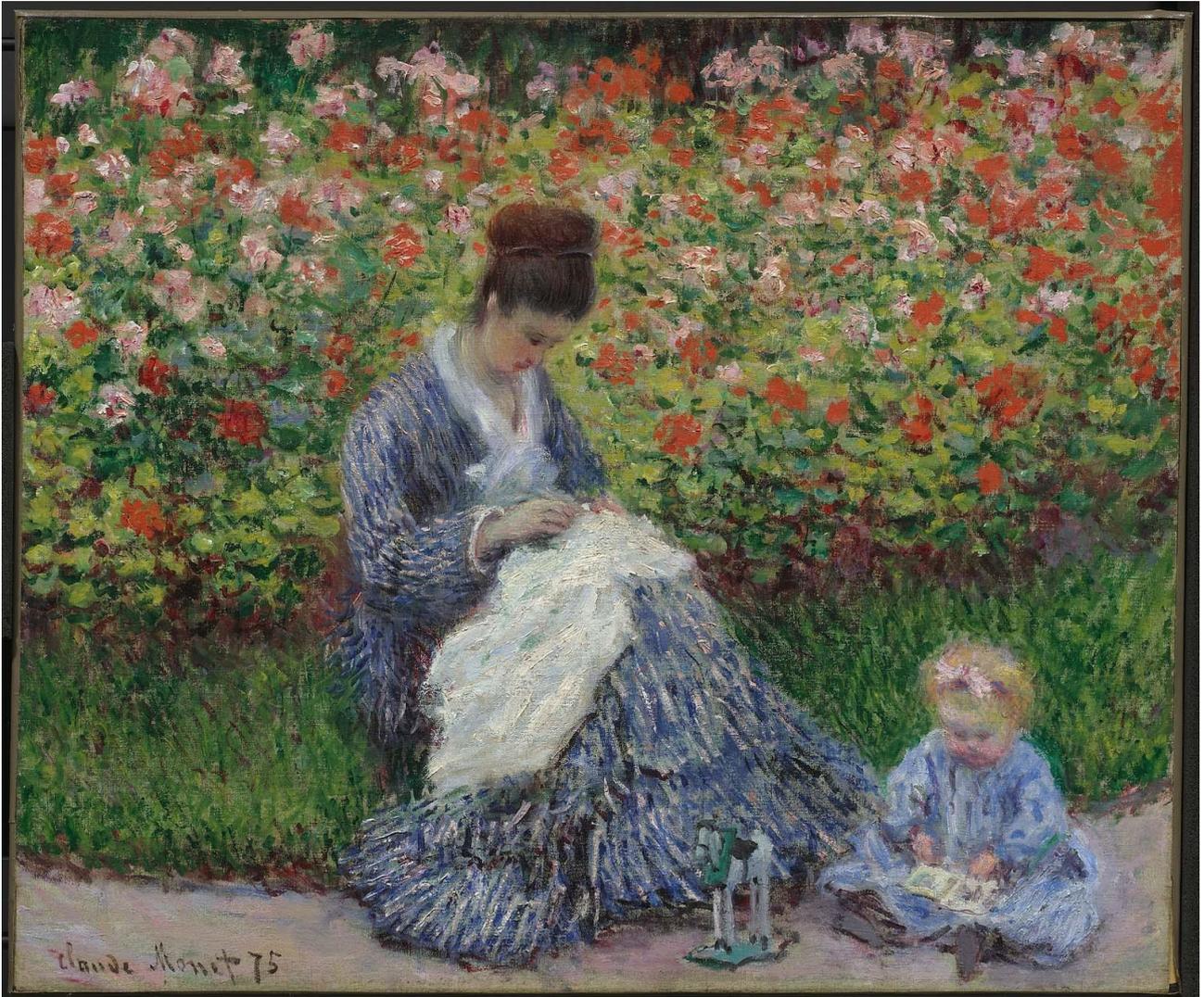
3.2. OBJETIVOS ESPECÍFICOS

3.2.1. Alphavirus

- 1- Obter os plasmídeos das proteases virais nsP2 de CHIK e MAYV.
- 2- Expressar os genes e purificar as proteínas em questão.
- 3- Realizar testes de atividade das proteases purificadas.
- 4- Realizar experimentos biofísicos diversos para análise estrutural.
- 5- Realizar *screening* das condições de cristalização das proteases.
- 6- Fazer a modelagem por homologia da protease nsP2 de MAYV.
- 7- Realizar experimento de *in silico* do contato entre a nsP2 de CHIK e MAYV com peptídeos contendo os sítios de clivagem destas proteases.
- 8- Realizar experimentos de dinâmica molecular das poses com melhores *scores*.
- 9- Analisar interações e sítios de reconhecimento.

3.2.2. Flavivirus

- 1- Obter os plasmídeos das proteases virais NS3 de ZIKV e DENV.
- 2- Expressar os genes e purificar as proteínas em questão.
- 3- Realizar testes de atividade das proteases purificadas.
- 4- Realizar experimentos biofísicos diversos para análise estrutural.
- 5- Realizar *screening* das condições de cristalização das proteases.
- 6- Analisar a inibição da protease em testes com diversas moléculas.



Camille Monet and a Child in the Artist's Garden in Argenteuil

1875 65x74cm oil/canvas

Claude Monet

4. Metodologia

4. METODOLOGIA

4.1. EXPERIMENTOS IN VITRO

4.1.1. Obtenção dos genes

Os genes das proteases virais de CHIK, MAYV e ZIKV selecionadas foram comprados da empresa GenOne em plasmídeos de expressão pET-21a, sendo selecionados os sítios de restrição NdeI e XhoI presentes nos próprios plasmídeos. Na construção da NS3 do Flavivirus ZIKV adicionou-se cauda de histidina na porção N-terminal, seguido de um sítio para a protease de HRV3C, com sequência LEVLFQGP. O plasmídeo da protease NS3 de ZIKV foi construído para expressa uma porção da proteína NS2b (48-95), um linker de sequência GGGGSGGGG e o domínio protease (1-170), construção que será chamada no presente trabalho de ZIKV-NS3-Bra. Os plasmídeos das proteases nsP2 dos Alphavirus de CHIK e MAYV, por sua vez, foram montados com os domínios protease e MTase (469-798), possuindo também cauda de histidina na porção N-terminal, seguido de um sítio para a protease de HRV3C, com sequência LEVLFQGP.

O plasmídeo da protease NS3 de DENV (SILVA, 2012) foi obtido com a professora Dra. Izabela Marques Dourado Bastos, do laboratório de Interação Patógeno-Hospedeiro na Universidade de Brasília. Outros dois plasmídeos foram obtidos com o professor Dr. Rolf Hilgenfeld, no Instituto de Bioquímica da Universidade de Lübeck, Alemanha. O primeiro foi um plasmídeo da protease NS2b/NS3 de ZIKV, com construção similar a realizada no Brasil, todavia apresentando as mutações em dois resíduos Arg95 da fração NS2b para Ala e Arg29 da NS3 para Gly (LEI et al., 2016), que tiveram como objetivo evitar uma possível autoproteólise. Esta montagem será chamada no presente trabalho como ZIKV-NS3-Lub. O segundo plasmídeo foi referente ao substrato de clivagem da nsP2 de CHIK, o qual era constituído por uma tioredoxina seguida do sítio de clivagem localizado entre as proteínas nsP3 e nsP4 de CHIK.

4.1.2. Clonagem da protease nsP2 de MAYV

4.1.2.1. Amplificação

O plasmídeo com a sequência do gene da protease nsP2 de MAYV que foi comprado para expressão em pET-21a foi usado em reação de PCR como material molde para amplificação do segmento gênico desejado. Os primers foram obtidos na empresa Eurofins Scientific, os quais tinham como objetivo amplificar a sequência gênica e,

concomitantemente, adicionar sequências para diferentes sítios de restrição. Foram então confeccionados três *primers*, dois *forward* com adição do sítio de clivagem para BamHI e NdeI, e um *reverse* com XhoI. Sendo assim, foram amplificadas duas amostras diferentes, uma acrescentando os *primers* com os sítios de NdeI e XhoI, e outra com BamHI e XhoI. Os passos da reação completa podem ser observados na figura 7.

- 1- 98°C – 5 min
- 2- 98°C – 30 sec
- 3- 65°C – 30 sec → 30x
- 4- 68°C – 70 sec
- 5- 68°C – 10 min
- 6- 4°C – ∞

Figura 7- Esquema com ciclos da reação de PCR para amplificação do segmento gênico de correspondente a protease nsP2 de MAYV.

O produto gerado pela amplificação teve seu tamanho confirmado através de gel de agarose 1%. Após isso as bandas dos géis referentes ao produto foram cortadas e o material genético extraído com uso do *GeneJET Gel Extraction Kit* (ThermoFisher Scientific) de acordo com o protocolo estabelecido pelo fabricante. A tabela 1 apresenta a sequência dos *primers* utilizados para a reação.

Tabela 1- Sequência de aminoácidos dos *primers* utilizados para amplificação da sequência gênica da protease nsP2 de MAYV.

Nome	Sequência (número de aminoácidos)
MAYV-nsP2-BamHI-F	CGCGGATCCGACGTCTTTCAGAACAAGGCGAAA (33)
MAYV-nsP2-NdeI-F	GGAATTCATATGGACGTCTTTCAGAACAAGGCG (34)
MAYV-nsP2-XhoI-R	CCGCTCGAGTTATTAGGTACCAGCATAGATGCTA (34)

4.1.2.2. Digestão e ligação

As amostras obtidas através da amplificação por PCR e os vetores escolhidos

foram digeridos com o uso das enzimas de restrição segundo os sítios adicionados. Os vetores escolhidos foram pET-15b, pET-28a, pET-sumo e pGEX, onde os dois primeiros foram tratados com NdeI e XhoI, e os dois últimos com BamHI e XhoI. A enzima de digestão NdeI foi incubada com as amostras por 2 horas, enquanto BamHI e XhoI por 40 minutos, sendo a temperatura das reações 37 °C. Após esse tempo estas foram submetidas a uma temperatura de 80 °C por 10 minutos visando a desativação das enzimas de restrição. Após a digestão, os vetores foram submetidos a gel de agarose 1 % e em seguida as bandas do gel foram cortadas e o material genético extraído usando o *GeneJET Gel Extraction Kit* (ThermoFisher Scientific) de acordo com o protocolo estabelecido pelo fabricante.

Os produtos de PCR e os vetores digeridos foram ligados usando a enzima DNA ligase T4. A proporção de vetor para produto de PCR foi de aproximadamente 1/5, onde foram adicionados 80 ng de cada um destes na reação. A reação ocorreu por 30 minutos a 22 °C e, após essa etapa, as amostras foram armazenadas em gelo. As amostras foram utilizadas para a transformação em células competente DH5 α , de acordo com protocolo posteriormente apresentado. As placas de meio YT contendo as células transformadas foram armazenadas em estufa *overnight* a 37°C.

4.1.2.3. PCR de colônia

Cinco colônias de cada uma das placas referentes às diferentes construções (pET-15b, pET-28a, pET-sumo e pGEX) foram selecionadas e tiveram uma parte separada para análise por PCR de colônia. Estas células selecionadas foram colocadas em tubo eppendorf e a elas foram acrescentados 150 pmol dos primers *forward* e *reverse*, 1 % DMSO, 1X *AccuPrime™Pfx Reaction Mix*, 2,5 U *AccuPrime™Pfx DNA Polymerase* (Invitrogen), com o uso de água milli-Q para completar até o volume final de 100 μ L. Os tempos utilizados em cada etapa da PCR são os mesmos previamente citados na seção de amplificação. O resultado foi analisado através de gel de agarose 1 %.

4.1.3. Preparo de células competentes

Em fluxo laminar, adicionou-se 100 μ L da célula em um tubo falcon com 5 mL de meio LB líquido (peptona 10 g/L, extrato de levedura 5 g/L e NaCl 10 g/L, pH 7), e esta foi incubada a 37 °C, *overnight*, em shaker a 200 rpm. No dia seguinte, retirou-se 1 mL do

pré-inóculo que foi adicionado em Erlenmeyer (125 mL) com 25 mL de meio LB, e incubou-se a 37 °C, 200 rpm, até atingir a O.D. (600 nm) de 0,3.

Após isso, colocou-se as células em gelo por 15 minutos seguido de centrifugação a 5000 g por 10 minutos a 4 °C. Descartou-se o sobrenadante e o *pellet* foi ressuscitado em 10 mL CaCl₂ 100 mM (estéril), seguido de mais 15 minutos de incubação em gelo e centrifugação na mesma condição da anterior. O sobrenadante foi descartado e o *pellet* ressuscitado em 1 mL de CaCl₂ com 15 % de glicerol (estéril), que se aliquotou 100 µL em microtubos (1,5 mL).

4.1.4. Transformação

Realizou-se a transformação das células competentes em fluxo laminar, adicionou-se 100 ng/µL do plasmídeo em 100 µL destas células, que foram mantidas por quarenta minutos em gelo. Submeteu-se as células a um choque térmico ao colocá-las por noventa segundos a 42 °C e, após isso, elas foram retornadas ao gelo. Adicionou-se 900 µL de meio LB líquido ao tubo e as mesmas foram deixadas em *shaker*, a 200 rpm, por um período de uma hora a 37 °C.

Preparou-se uma placa de Petri contendo meio LB sólido (peptona 10 g/L, extrato de levedura 5 g/L, NaCl 10 g/L e 10 g/L de Ágar, pH 7) com o antibiótico canamicina 50 µg/mL e 34 µg/mL de cloranfenicol, para células competentes com resistência a este antibiótico. Um total de 100 µL com células foi plaqueado nas placas contendo LB sólido e antibiótico, e estas ficaram a 37°C *overnight*.

4.1.5. Otimização da expressão e purificação

Otimizações da expressão das proteases virais do presente trabalho foram feitas nas cepas de *E. coli* BL21 (DE3), BL21 (DE3) pLysE, BL21 (DE3) PlysS, Lemo21 (DE3), Rosetta (DE3) e pGro (DE3), as quais estavam disponíveis no Laboratório de Biofísica Molecular da Universidade de Brasília. Além disso, para a nsP2 de MAYV foram testadas adicionalmente as cepas SoluBL21, C43 (DE3), Turner (DE3) e BL21-Gold (DE3), no Instituto de Bioquímica da Universidade de Lübeck, Alemanha. As diferentes características das cepas utilizadas podem ser observadas na tabela 2. Foram feitos testes também em diferentes temperaturas (entre 16 °C e 37 °C) e concentrações de IPTG (0,1 mM a 1 mM) para algumas das cepas.

Tabela 2- Características das cepas de *E. coli* utilizadas durante a etapa da otimização da expressão das proteases de ZIKV, DENV, CHIK e MAYV.

Cepas de <i>E. coli</i>	Características
BL21 (DE3)	T7 RNA polimerase sobre controle do promotor lac UV5. Deficiente nas proteases Lon e OmpT.
BL21 (DE3) pLysE	T7 RNA polimerase sobre controle do promotor lac UV5. Deficiente nas proteases Lon e OmpT. Produção da Lisozima T7 (em maior quantidade que a pLysS). Resistência a cloranfenicol.
BL21 (DE3) PlysS	T7 RNA polimerase sobre controle do promotor lac UV5. Deficiente nas proteases Lon e OmpT. Produção da Lisozima T7 (em menor quantidade que a pLysE). Resistência a cloranfenicol.
Lemo21 (DE3)	T7 RNA polimerase sobre controle do promotor lac UV5. Deficiente nas proteases Lon e OmpT. Expressão da T7 RNA polimerase modulada indiretamente por ramnose. Resistência a cloranfenicol.
Rosetta (DE3)	T7 RNA polimerase sobre controle do promotor lac UV5. Deficiente nas proteases Lon e OmpT. Aumento da expressão de proteínas eucarióticas que possuem códons raros para <i>E. coli</i> (AGG, AGA, AUA, CUA, CCC, GGA). Resistência a cloranfenicol.
pGro (DE3)	T7 RNA polimerase sobre controle do promotor lac UV5. Deficiente nas proteases Lon e OmpT. Resistência a cloranfenicol. Expressão das chaperonas groES-groEL modulada por arabinose.
SoluBL21 (DE3)	T7 RNA polimerase sobre controle do promotor lac UV5. Deficiente nas proteases Lon e OmpT. Aumento da produção de proteínas recombinantes solúveis.*

C43 (DE3)	T7 RNA polimerase sobre controle do promotor lac UV5. Deficiente nas proteases Lon e OmpT. Resistência a cloranfenicol. Produção da lisozima T7. Mutações que conferem resistência a morte celular por proteínas tóxicas.*
Turner (DE3)	T7 RNA polimerase sobre controle do promotor lac UV5. Deficiente nas proteases Lon e OmpT. Possui uma deleção na Lac Y que permite a entrada uniforme de IPTG no interior da célula.
BL21-Gold (DE3)	T7 RNA polimerase sobre controle do promotor lac UV5. Deficiente nas proteases Lon e OmpT. Aumento da eficiência de transformação.*

* Não foram encontrados detalhamento dos mecanismos moleculares que conferem as características especificadas.

Para a obtenção do pré-inóculo, selecionou-se uma colônia isolada das placas obtidas durante a etapa de transformação, a qual era inserida em um tubo falcon (15 mL) contendo 5 mL de meio LB acrescidos de canamicina (Kan) ou ampicilina (Amp) na concentração de 50 µg/mL e foi ainda adicionado 34 µg/mL de cloranfenicol de acordo com as marcas de seleção específicas de cada plasmídeo e cepas, as quais cresceram *overnight* a 37 °C a 200 rpm. A lista dos plasmídeos utilizados e suas respectivas marcas de seleção podem ser observados na tabela 3.

Tabela 3- Lista de plasmídeos utilizados das proteases arbovirais e suas respectivas marcas de seleção e tamanho da proteína produzida

Proteína	Plasmídeo	Marca de seleção	Tamanho (~kDa)
ZIKV-NS2b/NS3	pET-15b	Amp	26,1
	pET-21a	Amp	25,9
DENV-NS2b/NS3	pET-19b	Amp	29,2
MAYV-nsP2	pET-15b	Amp	38,3

	pET-21a	Amp	38,2
	pET-28a	Kan	38,3
	pET-Sumo	Kan	47,5
	pGEX	Amp	62,3
CHIV-nsP2	pET-21a	Amp	38,7
Substrato CHIK nsP3/nsP4	pCDNA	Amp	19,6

As otimizações ocorreram em duas etapas, primeiramente 1 mL do pré-inóculo foi transferido para um Erlenmeyer (125 mL) com 25 mL de meio LB. Incubou-se as células a 37 °C, 220 rpm, até atingir a O.D. (600nm) de 0,6. Após se atingir os valores de O.D. desejados, a expressão foi induzida pela adição de IPTG, todavia 1 mL era separado para análise posterior em gel SDS-PAGE 12%. A expressão ocorria em incubadoras refrigeradas de acordo com a temperatura da expressão desejada. A segunda etapa, era o escalonamento para frascos Erlenmeyers com volume de 1 L com 200 – 500 mL de meio LB.

Para a preparação das amostras de expressão para a etapa da purificação, centrifugou-se utilizando-se a centrífuga 5430 R (Eppendorf). Sonicou-se as amostras, em gelo, utilizando o equipamento Sonifier 450 (Branson). Filtrou-se as amostras do sobrenadante com um filtro de seringa de 0,45 µm. Água ultrapurificada, etanol 20 %, e tampões foram filtrados em filtro 0,22 µm usando bomba à vácuo, WP6111560 (Millipore), e desaerados em banho ultrassom (Quimis).

Utilizou-se o sistema de cromatografia AKTAprime plus (GE healthcare) e a coluna HisTrap HP 1 mL (GE Healthcare) para a cromatografia de afinidade a metal (níquel) com objetivo de purificar proteínas com cauda de histidina. Passou-se na coluna um total de dez volumes de etanol 20 %, água e tampão de ligação, e em seguida foi injetada a amostra, coletou-se alíquotas da fração não adsorvida. Durante toda a purificação mantevesse um fluxo de 1 mL/minuto, com um limite máximo de pressão de 0,5 MPa. As etapas de purificação e refinamento variaram entre as diferentes proteínas e serão detalhadas a seguir.

4.1.5.1. NS2b/NS3 de ZIKV

As diferentes construções das proteases de ZIKV foram expressas usando diferentes condições. Primeiramente a construção ZIKV-NS3-Bra foi expressa em meio LB usando frascos Erlenmeyer de 1 L preenchidos com 500 mL. Após indução com 1 mM de IPTG, a proteína foi expressa a 19 °C por um tempo total de 16 horas. Centrifugou-se por 15 minutos a 8000 rpm para se formar o *pellet* de células. A amostra foi ressuspendida em 40 mL de tampão A (50 mM de NaH₂PO₄, 300 mM NaCl, pH 7,5) e as células lisadas em sonicador Sonifier 450 (Branson) (*duty cycle* 30, *output control* 3 por 15 minutos).

A purificação usando coluna de afinidade a metal mobilizado (níquel) foi realizada equilibrando-se a coluna HisTrap HP (1 mL) no mesmo tampão da lise. A eluição foi realizada com tampão B contendo 50 mM de NaH₂PO₄, 300 mM NaCl, 500 mM de imidazol, pH 7,5, sendo que foram feitas lavagens contendo 50 mM, 125 mM, 250 mM e 500 mM de imidazol.

Uma segunda etapa de purificação foi realizada utilizando-se duas metodologias de purificação. A primeira foi a coluna Superdex S75 increase (GE Healthcare), para a qual a amostra foi concentrada usando-se os concentradores Vivaspín® Turbo 4 10 kDa (Sartorius). A coluna foi equilibrada em tampão A e a amostra injetada a um fluxo de 0,5 mL/min e os picos coletados para análise posterior. A segunda foi utilizando-se a coluna de troca aniônica (Mono Q 4.6/100 PE– strong anion), na qual a coluna e a amostra foram colocadas em tampão 20 mM Tris-HCl, pH 8,0, utilizando-se os concentradores anteriormente citados. Realizou-se a eluição por gradiente de NaCl (0 – 500 mM) em um volume total de 30 mL a um fluxo de 1 mL/min.

A proteína correspondente à construção ZIKV-NS3-Lub foi expressa de acordo com artigo previamente publicado (LEI et al., 2016), com alterações. A protease foi expressa em frasco Erlenmeyer de 1 L com 500 mL meio LB. Após a cultura ser inoculada e alcançar 0,6 de O.D. adicionou-se 0,5 mM de IPTG, e a expressão ocorreu por 12 horas a 20 °C a 220 rpm. As células foram centrifugadas a 7500 rpm por 15 minutos, a 4 °C e ressuspendidas em tampão contendo 25 mM Tris-HCl, 5 % glicerol, pH 8,5 e lisadas em sonicador por 15 minutos de tempo ativo usando a seguinte configuração: 30 % amplitude, ciclos de 2 segundos de lise e 3 segundos de descanso. A amostra foi centrifugada por 40 minutos a 12000 rpm, 4 °C e purificada por cromatografia de afinidade a metal (níquel) usando a coluna HisTrap HP (1 mL), através gradiente de imidazol de 50 mL em fluxo de 1 mL/min, variando a concentração de imidazol de 0 a 500 mM.

Uma segunda etapa de purificação por troca aniônica foi realizada utilizando a coluna HiTrap Q FF. A proteína passou por uma etapa para remoção do imidazol presente na amostra, e após isso foi aplicada na coluna e eluída por gradiente de NaCl de 0 a 500 mM em fluxo de 1 mL/min por 100 mL.

4.1.5.2. NS2b/NS3 de DENV

A expressão e purificação da protease de DENV foi feita de acordo com trabalho previamente realizado (SILVA, 2012) com modificações. As células ao atingirem uma O.D. de aproximadamente 0,6 foi adicionado 1 mM de IPTG em frascos Erlenmeyer (1 L) preenchidos com um volume total de 500 mL. A expressão ocorreu a 37 °C por três horas.

As amostras foram ressuspensas em tampão Tris-HCl 20 mM, pH 7,9, NaCl 500 mM e lisadas em sonicador com os seguintes parâmetros: amplitude 30, tempo de processamento 2 minutos, 10 segundos de funcionamento por 50 segundos parado. Após isto, esta foi centrifugada a 12000 rpm, a uma temperatura de 4 °C. O sobrenadante foi injetado na coluna de cromatografia líquida de afinidade a níquel HisTrap HP (1 mL). A eluição foi realizada após etapa de equilíbrio com tampão A contendo Tris-HCl 20 mM, pH 7,9, NaCl 500 mM, pH 7,5, sendo que foram feitas lavagens com tampão de eluição (Tris-HCl 20 mM, pH 7,9, NaCl 500 mM, pH 7,5, 500 mM de Imidazol) nas concentrações 50 mM, 125 mM, 250 mM e 500 mM de imidazol.

Seguiu-se para uma etapa de refinamento da purificação, para esta utilizou-se a coluna Hiload 16/600 Superdex 200 (GE Healthcare). A coluna foi equilibrada em tampão A, e a amostra foi aplicada em um fluxo de 0,5 mL/min. Os picos foram coletados e analisados em gel SDS-PAGE 12 %.

4.1.5.3. nsP2 de MAYV

A protease de mayaro teve sua expressão testada em várias cepas de *E. coli* devido a sua grande dificuldade de expressão solúvel. As cepas de *E. coli* foram estas: BL21 (DE3), BL21 (DE3) pLysE, BL21 (DE3) PlysS, Lemo21 (DE3), Rosetta (DE3) e pGro (DE3). Os testes variaram a temperatura entre 37 °C, 25 °C, 19 °C ou 16 °C, e a concentração de IPTG entre 0,1 e 1 mM. Os primeiros testes foram realizados em tampão 50 mM NaH₂PO₄, 300 mM NaCl, com pH 7,5. Além disso, monitorou-se a solubilidade da protease durante as primeiras horas de expressão.

Outro teste foi realizado utilizando-se meio auto-indutor 1 % peptona, 0,5 % extrato de levedura, 25 mM Na₂HPO₄, 25 mM KH₂PO₄, 50 mM NH₄Cl, 5 mM Na₂SO₄, 2 mM MgSO₄, 0,5 % glicerol, 0,05 % glicose e 0,2 % α-lactose. As células cresceram a 37 °C até atingir uma O.D. de 0,6, após isso tiveram a temperatura reduzida para 18 °C. Foram coletadas amostras nos tempos 9, 10, 11, 12 e 13 horas de expressão. As amostras foram ressuspendidas em tampão Tris-HCl 20 mM, 100 mM NaCl e pH 7,5. Estas foram lisadas e submetidas a teste de solubilidade, conforme descrito na seção (4.1.8).

Um segundo teste de expressão dessa protease foi realizada no Instituto de Bioquímica da Universidade de Lübeck, Alemanha. Uma outra condição de tampão foi testada baseada em condições de expressão de outras proteases de Alphavirus homólogas (HU et al., 2016; SAISAWANG et al., 2015a), a composição foi de 20 mM Tris-HCl, 200 mM NaCl, 200 mM arginina, 200 mM ácido glutâmico, pH 7,1 e glicerol 5 %. O meio de cultura utilizado foi o meio YT 1x. Nestas novas condições foram testadas as cepas BL21 (DE3), BL21 (DE3) pLysE, BL21 (DE3) PlysS, SoluBL21, C43 (DE3), Turner (DE3) e BL21-Gold (DE3). As temperaturas testadas foram 25°C e 16°C, usando-se a concentração de 1 mM de IPTG.

As cepas BL21 (DE3), BL21 (DE3) pLysE e BL21 (DE3) PlysS apresentaram uma banda em tamanho próxima ao esperado em gel SDS-PAGE 15 % e foram então escalonadas para expressão em maior quantidade em que foram usados Erlenmeyer de 5 L com 1 L de meio YT 1x. As células cresceram a 37 °C até atingir a O.D. de 0,6, então reduziu-se a temperatura a 25 °C e foram induzidas com 1 mM de IPTG, com um tempo de expressão de 14 horas.

Centrifugaram-se as células a 7500 rpm por 15 minutos, a 4°C e seguiu-se para a etapa de lise em sonicador com os seguintes parâmetros: 30% amplitude, ciclos de 2 segundos de lise e 3 segundos de descanso. As amostras foram ressuspendidas no tampão anteriormente citado. As amostras foram centrifugadas em ultracentrífuga por 1 hora, a 40000 rpm a 4°C. As amostras foram purificadas por cromatografia líquida de afinidade a níquel (Histrap 5 mL, GE Healthcare), onde foi utilizado gradiente de imidazol (0 a 500 mM), com fluxos de 1 mL/min durante a injeção e 3 mL/min durante o gradiente. Os picos foram analisados em gel SDS-PAGE 15 %.

Devido ao insucesso de expressão dessa proteína na fração solúvel, quatro diferentes construções dela foram montadas, como previamente descrito (pET-15b, pET-28a, pET-sumo e pGEX). As quatro construções foram expressas de acordo com os protocolos de expressão e o preparo para purificação foi mantido como anteriormente

citado. A purificação ocorreu por cromatografia líquida de afinidade a níquel (HisTrap 1 mL, GE Healthcare) onde a eluição se deu em três etapas com 50 mM, 250 mM e 500 mM de imidazol. O resultado da purificação foi analisado em gel SDS-PAGE 15%.

4.1.5.4. nsP2 de CHIK

A protease de CHIK foi expressa em meio YT. As células cresceram a 37 °C, até atingirem a densidade ótica de 0,6. A expressão foi induzida por IPTG na concentração final de 1 mM, a 18 °C por 16 horas, sob agitação constante a 150 rpm.

As células foram centrifugadas a 7500 rpm por 15 minutos, a 4 °C e ressuspendidas em tampão contendo 20 mM Tris-HCl, 200 mM NaCl, 200 mM arginina e 200 mM ácido glutâmico, pH 7,1 e lisadas em sonicador por 15 minutos de tempo ativo usando a seguinte configuração: 30 % amplitude, ciclos de 2 segundos de lise e 3 segundos de descanso. A amostra foi centrifugada por 1 hora, em rotação de 40000 rpm a 4 °C em ultracentrífuga e o sobrenadante foi filtrado em filtro de seringa 0,45 µm. A proteína foi purificada por cromatografia de afinidade a metal (níquel) usando-se a coluna HisTrap HP (1 mL), através gradiente de imidazol de 50 mL em fluxo de 1 mL/min, variando-se a concentração de imidazol de 0 a 500 mM.

A proteína purificada foi estocada *overnight* com a protease HRV3C visando a remoção da cauda de histidina, em que a protease de CHIV foi colocada em concentração 10 vezes superior. Adicionou-se 1 mM de DTT na amostra. Após isso a protease foi purificada em coluna de GST acoplada a coluna de níquel. As colunas foram equilibradas em tampão de lise, e após a aplicação foram coletadas as proteínas que não interagiram com ambas as resinas. As proteínas que interagiram com as respectivas colunas foram eluídas separadamente, a coluna de GST usando tampão contendo GSH 20 mM, Tris 50 mM, DTT 1 mM e pH 8, e a coluna de níquel usando-se tampão de lise contendo 500 mM de imidazol. Esta etapa foi realizada duas vezes visando-se a obtenção de maior nível de pureza proteica.

4.1.5.5. Substrato de nsP2 de CHIK

A expressão do substrato da protease de CHIK foi realizada em meio autoindutor com a composição: 928 mL ZY-Medium, 1 mL MgSO₄ 1 M, 20 mL 50x 5052 e 50 mL 20x NPS, onde ZY-Medium era composto de 10 g peptona e 5 g extrato de levedura, o 50x

5052 (1L) por 250 g glicerol, 25 g glicose e 100 g α -lactose, e o 20x NPS por 0,5 M $(\text{NH}_4)_2\text{SO}_4$, 1 M KH_2PO_4 e 1 M Na_2HPO_4 . Realizou-se um pré-inóculo em meio LB *overnight* a 37 °C e 10 mL deste foram inoculados em meio autoindutor. A temperatura foi reduzida para 16 °C quando a densidade celular do meio atingiu um valor de 0,6, e foi mantida a uma rotação de 130 rpm *overnight*.

As células foram centrifugadas a 7500 rpm por 15 minutos, a 4 °C e ressuspendidas em tampão contendo Tris 20 mM, NaCl 150 mM, pH 7,5 e 5 % de glicerol. Estas foram lisadas em sonicador por 15 minutos de tempo ativo usando a seguinte configuração: 30 % amplitude, ciclos de 2 segundos de lise e 3 segundos de descanso. A amostra foi centrifugada por 1 hora, em rotação de 40.000 rpm a 4 °C em ultracentrífuga e o sobrenadante foi filtrado em filtro de seringa 0,45 μm . Esta foi submetida a cromatografia de afinidade a metal (níquel), em que a coluna (HisTrap HP 1 mL) foi equilibrada em tampão de lise e a amostra foi aplicada na coluna e eluída mediante um gradiente de imidazol. Este ocorreu por um volume de 50 mL no fluxo de 1 mL/min, numa variação de 0 até 500 mM de imidazol.

Visando o aumento do nível de pureza proteica, esta foi submetida a uma etapa de purificação por troca aniônica (HiTrap Q HP). A coluna foi equilibrada em tampão contendo 20 mM de Tris, pH 8 e 5 % de glicerol e a proteína foi aplicada após etapa de remoção de sal em amicon ultra-4 10K. A proteína foi eluída após gradiente de NaCl por um volume de 100 mL em fluxo de 1 mL/min, em que ocorreu a variação da concentração de NaCl entre 0 e 500 mM.

4.1.6. Renaturação

4.1.6.1. Purificação em condições desnaturantes

A expressão da protease nsP2 de MAYV foi realizada por 16 horas a 19 °C induzida por 1 mM de IPTG em meio LB. A presença da proteína em corpos de inclusão foi confirmada por gel SDS-PAGE 12%. As células foram então centrifugadas a 8000 rpm durante 20 minutos na centrífuga refrigerada CR22N (Hitachi) para remoção do meio de cultura, ressuspendidas em tampão de lise (Tris 20 mM, NaCl 150 mM, pH 7,5) e congeladas a -20 °C.

A amostra foi lisada em sonicador (Branson Sonifier 450) (*duty cycle* 30, *output control* 3 por 15 minutos), e o pellet foi separado por centrifugação 10.000 rpm durante 30 minutos em centrífuga refrigerada (Hitachi CR22N). Este foi então ressuspendido em

tampão de desnaturação I (ureia 8 M, Tris 10 mM e NaH₂PO₄ 100 mM, pH 8,0) e incubado em gelo por 3 horas. A etapa de centrifugação foi então repetida visando obtenção do sobrenadante para purificação por cromatografia líquida de afinidade a metal (Ni) em condições desnaturantes.

A coluna de níquel foi equilibrada em tampão de desnaturação I e a amostra foi injetada. Após esta etapa realizaram-se sucessivas lavagens com tampão de desnaturação II (ureia 8 M, Tris 10 mM e NaH₂PO₄ 100 mM, pH 8,0, 500 mM de Imidazol). As diferentes frações eluídas foram então analisadas em gel SDS-PAGE 12%.

4.1.6.2. Renaturação por diluição rápida

Com a fração eluída seguiu-se à etapa de renaturação, em que a amostra foi titulada em tampão de renaturação (Tris 50 mM, pH 7,5, L-arginina 0,5 M, CaCl₂ 20 mM, EDTA 1 mM, NaCl 100 mM, Brij 58 a 0,05 %, glutathiona oxidada 0,05 mM e glutathiona reduzida 0,5 mM) no fluxo de 0,1 mL por minuto. Esta foi incubada sob agitação por 72 horas a temperatura ambiente. A amostra foi então concentrada utilizando-se concentrador de centrífuga com membrana vertical Vivaspin Turbo 10000 MW (Sartorius).

4.1.6.3. Análise de estrutura secundária por dicroísmo circular

Realizou-se também a análise estrutural das amostras após a etapa da renaturação através de dicroísmo circular (CD). Utilizou-se o espectropolarímetro Jasco J-815, o qual estava equipado com aparelho Peltier de controle de temperatura. A proteína nsP2 de MAYV foi analisada a uma concentração de 0,11 mg/mL, em cubeta de quartzo de caminho óptico 0,1 cm e na faixa de comprimento de onda de 190-260 nm. A protease estava em tampão tris HCl 2 mM pH 7,5. Obtiveram-se assim valores de elipsidade, os quais foram convertidos para valores de elipsidade molar utilizando-se a equação abaixo:

$$[\theta] = 115 \cdot \alpha \cdot (10 \cdot C \cdot l)^{-1} \quad (1)$$

Na presente equação α referi-se ao sinal dicroico em miligraus, C à concentração da amostra (mg.mL⁻¹) e l, por sua vez, é o caminho óptico da cubeta em cm.

Os gráficos foram gerados utilizando-se o programa OriginPro 8.1. Os valores de porcentagem da estrutura secundária foram obtidos a partir da deconvolução dos dados

de elipsidade molar utilizando o programa CDNN utilizando-se duas bibliotecas de espectros: “23 basespectra” e “33 basespectra”. As demais estruturas e modelos tiveram sua porcentagem de estrutura secundária calculada usando-se o servidor 2struc (KLOSE; WALLACE; JANES, 2010).

4.1.7. Análise por gel SDS-PAGE 12 %

O gel SDS-PAGE 12 % é preparado em duas partes, um gel separador e um concentrador. No primeiro utilizou-se, para cada gel, 2 mL uma solução de poliacrilamida (29 % acrilamida e 1 % bis-acrilamida), 1,7 mL de água destilada, 1,25 mL de tampão Tris-HCl (pH 8,8), 50 µL de PSA 10 %, 3 µL de TEMED e 150 µL de SDS 10 %. No segundo as quantidades de reagente foram 415 µL da solução da mesma solução de poliacrilamida, 1,75 mL de água destilada, 340 µL de tampão Tris-HCl (pH 6,8), 25 µL de PSA 10 %, 3 µL de TEMED e 75 µL de SDS 10 %.

Quanto à preparação das amostras, estas foram centrifugadas, em *mini spin plus* (Eppendorf), a 14500 rpm por 2 minutos. O sobrenadante foi descartado e o pellet foi ressuscitado em tampão de amostra 1x (60 mM Tris-HCl pH 6,8, SDS 2 %, B-mercaptoetanol 5 % e azul de bromofenol 0,01 %), em seguida ferveu-se em banho maria por dez minutos. Realizou-se a corrida do gel na cuba de eletroforese *Mini-PROTEAN Tetra System* (BIO-RAD) com a fonte EPS 301 (GE healthcare) em tampão de corrida (Tris-HCl 25 mM, glicina 200 mM e SDS 0,1 %). Nos primeiros vinte minutos, durante a entrada da amostra no gel e passagem pelo concentrador, utilizou-se a corrente elétrica de 25 mA e após isso, 45 mA durante toda a corrida, com a tensão elétrica de 200 V. Os géis foram corados com uma solução corante (50 % etanol, 25 % ácido acético, 24,5 % água destilada e 0,5 % comassie brilliant blue R-250) por duas horas para marcação das proteínas, e descorados em ácido acético 20 %.

4.1.8. Teste de solubilidade

As amostras foram centrifugadas a 14500 rpm por dois minutos, o sobrenadante foi descartado e as amostras ressuscitadas em 150 µL de tampão de lise. Sonicaram-se as amostras utilizando-se o Branson Sonifier 450 por um minuto (duty cycle 50 % e output control 5). Após isso, as amostras foram mais uma vez centrifugadas conforme os parâmetros anteriormente citados, o sobrenadante foi transferido para outro microtubo

(1,5 mL) e adicionaram-se 50 µL de tampão de amostra 4x, o pellet, por sua vez, foi ressuspenso em 200 µL de tampão de amostra 1x. Em seguida, ambas amostras foram fervidas e analisadas em gel SDS-PAGE 12 %.

4.1.9. Western Blot

Para esta análise um gel SDS-PAGE 15 % foi corrido, mas não foi corado com a solução corante de comassie brilliant blue R-250. Uma membrana de PVDF e seis papéis de filtro foram brevemente embebidos em metanol 100 %, em seguida, em tampão CAPS (10 mM CAPS em metanol 10 % pH11). Para a transferência foram empilhados três papéis de filtro, seguidos pelo gel, a membrana de PVDF e mais três papéis de filtro. Este foi submetido a 35 mA por 40 minutos.

A membrana foi submetida a duas lavagens de 10 minutos em tampão TBS (10 mM Tris-HCl, pH 7,5 e 150 mM NaCl), etapa que foi seguida por 1 hora de imersão em tampão de bloqueio (3 % BSA em TBS). Duas outras lavagens foram feitas com tampão tween/triton (20 mM Tris-HCl, pH 7,5, 500 mM NaCl, 0,05 % tween e 0,2 % triton) e mais duas com tampão TBS, todas estas por 10 minutos cada.

Uma etapa de imersão *overnight* em tampão de bloqueio acrescido do anticorpo tetra-his (Cat No./ID: 34670, Qiagen) em uma proporção de 1:5000 foi realizada. Seguiram-se mais três etapas de lavagem por 10 minutos, as duas primeiras com tampão tween/triton e a última com o TBS. A membrana foi então imersa em uma solução contendo Anti-Mouse IgG (Product No. A4312, Sigma) diluído em tampão de bloqueio em uma proporção de 1:10000 por 1 hora, seguida por quatro lavagens de 10 minutos em tampão tween/triton. A membrana foi então corada usando-se uma solução de 44 µL NBT (7,5 % em 70 % de dimetilformamida) e 33 µL BCIP (5 % em 100 % de dimetilformamida) em 10 mL de solução tampão (100 mM Tris-HCl, pH 9,5, 100 mM NaCl e 5 mM MgCl₂).

4.1.10. Testes de atividade e inibição

4.1.10.1. Chikungunya

A construção para expressão do substrato para o teste de atividade da protease de CHIK foi obtido com o professor Dr. Rolf Hilgenfeld no Instituto de Bioquímica da Universidade de Lübeck, Alemanha. Este consiste em uma tioredoxina seguida do sítio de clivagem entre as proteínas virais nsP3/4 acrescida de cauda de histidina no sítio C-

terminal.

Na realização do ensaio as proteínas tiveram a concentração fixada do substrato (1 µg) mas variada a massa da protease nsP2 (0, 0,2, 0,4, 0,6, 0,8, 1, 1,5, 2, 3 e 4 µg) na presença e ausência de DTT 1 mM (0 e 1 µg). As amostras foram incubadas durante 2 horas a 30 °C, em tampão 50 mM Tris-HCl, pH 7,5, 50 mM NaCl e 1 mM EDTA, e após isso analisadas em gel de agarose 15 %.

4.1.10.2. Zika

4.1.10.2.1. Teste de inibição por gel SDS-PAGE 15 % (ZIKV-NS3-Bra)

Realizou-se um teste de inibição preliminar onde a proteína NS3 foi colocada em proporção 1:2 em relação aos inibidores (Aprotinina, Benzamidina, Leupeptina, PMSF e BTCl), em que a proteína se encontrava em uma concentração de 20 µM. A proteína sozinha e com os inibidores foram mantidos à temperatura ambiente em quatro microtubos que tiveram suas reações paradas em diferentes tempos: 1 hora, 2 horas, 3 horas e 16 horas por adição de tampão de amostra 1x (60 mM Tris-HCl pH 6,8, SDS 2 %, B-mercaptotanol 5% e azul de bromofenol 0,01%) e subsequente aquecimento a 95 °C. Todas as amostras foram analisadas em gel de agarose 15%.

4.1.10.2.2. Ensaios de atividade e inibição enzimática (ZIKV-NS3-Lub)

Os ensaios de atividade e inibição da protease de NS2b/NS3 ZIKV foram feitos utilizando-se o substrato Benzoyl-Nle-Lys-Arg-Arg-AMC (Peptides International), diluído em DMSO 100%. Um segundo estoque foi feito com DMSO 1%, sendo esta concentração usada durante os experimentos, para este processo este foi diluído em água Milli-Q obtendo-se uma concentração final de 200 µM. Os ensaios foram feitos em microplacas de 96 poços utilizando-se tampão Tris-HCl 25 mM, pH 8,5 e 5 % de glicerol. As leituras das microplacas foram feitas no equipamento SpectraMax M2e (Molecular Devices), fixando-se a temperatura a 37°C com agitação inicial de 5 segundos e coleta de dados a cada 30 segundos durante 30 minutos. Para a realização do experimento foi fixado o comprimento de onda de excitação e emissão, 380 nm e 460 nm, respectivamente. Todos os experimentos foram realizados em triplicatas. O processamento e preparação dos gráficos foram realizados com o programa OriginPro v2015 (*OriginLab Corporation*, Northampton, MA, USA). A atividade enzimática foi determinada a partir da determinação

das velocidades iniciais. Para isso, as variações da intensidade de emissão de fluorescência em relação ao tempo foram convertidas a unidades de atividade enzimática *katan* (kat) ou a mólis do produto 7-amino-4-metilcoumarina (AMC) formado utilizando uma curva de padrão (OLIVEIRA, 2018).

4.1.10.2.2.1. Influência da concentração de enzima na atividade enzimática

No ensaio manteve-se fixa a concentração inicial não limitante de substrato (50 μM) e variou-se a concentração da enzima (0, 10, 20, 30, 40, 50, 60, 70 e 80 nM) nas mesmas condições experimentais anteriormente descritas. A atividade enzimática foi determinada e correlacionada linearmente com a concentração da enzima para determinar o intervalo de concentração ótimo da enzima para os experimentos. Da correlação linear foi determinada a atividade específica da enzima como a taxa de variação da atividade enzimática em relação à quantidade de enzima no ensaio.

4.1.10.2.2.2. Gráfico de Michaelis-Menten

Os parâmetros cinéticos da enzima (K_m e $V_{m\acute{a}x}$) foram determinados a partir do gráfico de *Michaelis-Menten*. Para tal fixou-se a concentração de proteína em 20 nM e variou-se a concentração de substrato (0, 10, 20, 30, 40, 50, 60, 80, 100, 120, 140 e 160 μM). Os dados obtidos foram utilizados para construção do gráfico. Os parâmetros cinéticos foram calculados a partir da equação de Michaelis-Menten.

$$V_0 = V_{m\acute{a}x} [S] / (K_m + [S]) \quad (2)$$

onde V_0 é a velocidade inicial da reação, $V_{m\acute{a}x}$ é a velocidade máxima da reação, $[S]$ é a concentração do substrato e K_m é a constante de Michaelis-Menten que pode ser entendida como a concentração do substrato na qual a atividade da enzima é igual à metade da velocidade máxima, ou seja, $V_{m\acute{a}x}/2$.

Além disso, calcularam-se os valores de K_{cat} e de eficiência catalítica (EC) da protease NS3, utilizando-se as equações a baixo:

$$K_{cat} = V_{m\acute{a}x}/[Et] \quad (3)$$

$$EC = K_{cat}/K_m \quad (4)$$

onde [Et] representa o valor da concentração de enzima em nM.

4.1.10.2.3. Ensaios de inibição da atividade enzimática

A atividade inibitória foi calculada mediante a determinação da metade do valor da máxima concentração inibitória (IC₅₀) dos inibidores da NS2b/NS3 de DENV previamente testados (SILVA, 2012). Foram obtidas 4 moléculas (F0174-0048, F0745-1098, F0816-0342 e F3165-0782) da empresa Life Chemicals (Ontário, Canadá) (figura 8). Durante os ensaios foram mantidas as concentrações de proteína e substrato fixas, respectivamente 20 nM e 200 µM, enquanto variou-se a concentração do inibidor de 400 µM a 0,39 µM. Além disso o teste foi realizado com o inibidor Leupeptina nas mesmas concentrações. A partir dos gráficos construídos, foi possível calcular o valor da concentração inibitória média (IC₅₀), a qual representa a metade do valor da máxima concentração inibitória. Os gráficos de atividade enzimática em presença de quantidades crescentes do inibidor foram ajustados ao modelo de função sigmoidal de Boltzmann para calcular o IC₅₀.

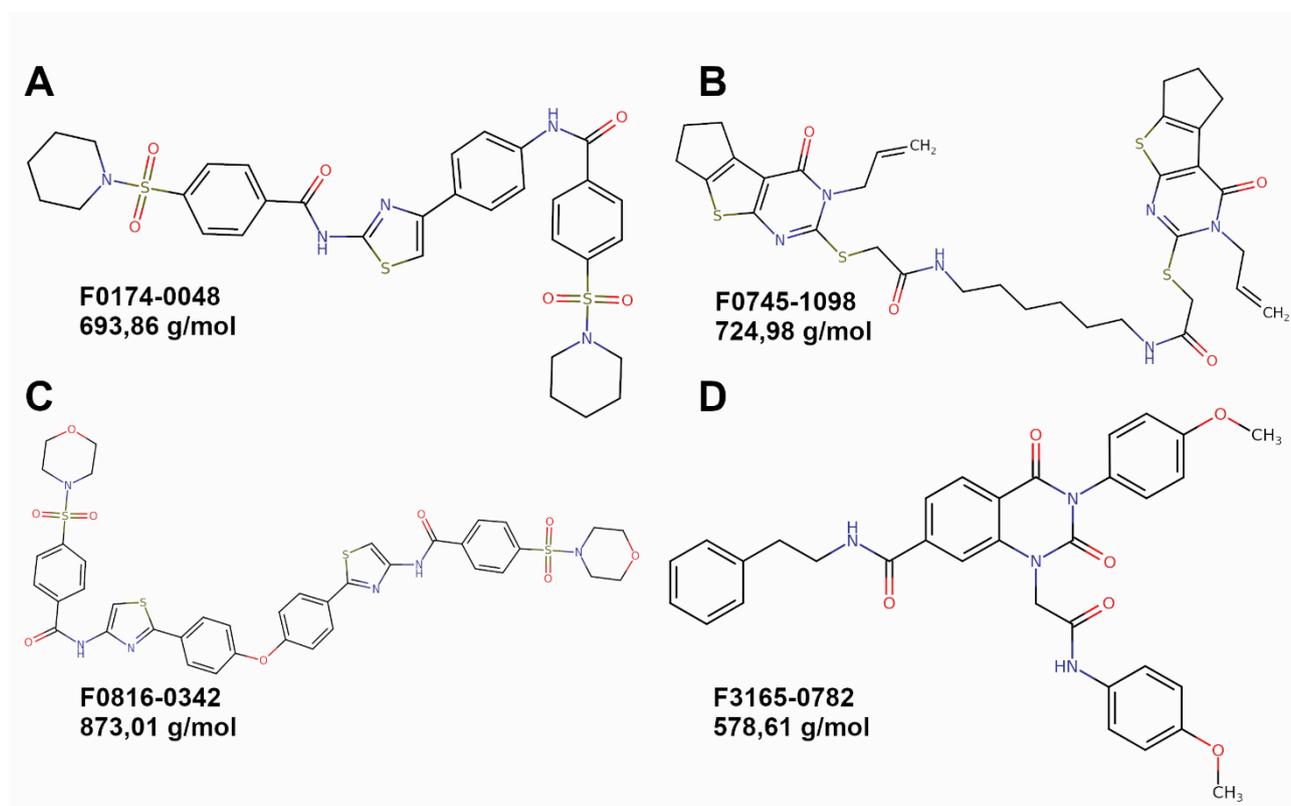


Figura 8- Inibidores da protease de DENV obtidos para ensaio de inibição contra a protease NS2B/NS3 de ZIKV.

4.1.11. Ensaios de fluorescência

Foram feitas várias medidas de fluorescência da protease NS2b/NS3 de ZIKV (construção ZIKV-NS3-Bra) utilizando-se o espectrofluorímetro (Jasco FP-6500) incrementando-se a concentração de aprotinina. As proporções de proteína/ligante foram 1/0,2, 1/0,5, 1/1, 1/2, 1/4, 1/6, 1/8, 1/10, 1/15 e 1/20. Foi feita uma varredura entre os comprimentos de onda entre 300 nm e 450 nm e o comprimento de onda de excitação utilizado foi de 295 nm. Para análise retirou-se a contribuição do tampão e os valores referentes à aprotinina. Os gráficos foram gerados no *software* Origin.

4.1.12. Ultracentrifugação analítica

Nos experimentos de ultracentrifugação analítica utilizou-se a análise de velocidades de sedimentação (UCA-SV) em que foram realizadas centrifugações a uma velocidade de 35.000 rpm em uma ultracentrífuga analítica ProteomeLab XL-A/XL- I (Beckman Coulter, EUA), equipada com um rotor de titânio An-60 Ti de quatro posições (Beckman, EUA). Equilibrou-se o rotor a 20 °C na câmara de vácuo antes do ensaio, e utilizou-se células de alumínio padrão com uma peça central de duplo canal e janelas de quartzo. Adquiriram-se os dados de absorvância a 215 nm e 280 nm com resolução radial de 20 μm , em intervalos de 5 minutos para cada leitura da amostra. Foram realizadas análises para as proteases de Flavivirus ZIKV (ZIKV-NS3-Bra) e DENV. Para a primeira as amostras de proteínas pura foram mantidas em tampão 20 mM Tris-HCl, pH 8,0, 230 mM de NaCl e foram colocadas em três células com diferentes $\text{OD}_{280\text{nm}}$ 0,75, 1,12 e 1,50. Utilizou-se o *software* SEDENTERP (<http://www.rasmb.bbri.org>) para estimar o volume específico parcial da proteína (v) no valor de 0,73014 e 0,726087, a densidade (ρ) no valor de 1,00910 e 1,0029 e a viscosidade (η) da solução calculado em 0,01031 e 0,01002, para ZIKV e DENV respectivamente. As varreduras radiais coletadas foram analisadas usando o modelo de análise de distribuição contínua do Coeficiente de Sedimentação [c(s)] pelo *software* SEDFIT v14.7. A distribuição de MW e c(s) das proteínas foram calculadas utilizando-se o modelo global de análise de espécies do *software* SEDPHAT.

4.1.13. *Screening* e refinamento de cristalização

Os ensaios de cristalização foram iniciados utilizando-se uma metodologia de *screening* já estabelecida no Laboratório de Biofísica Molecular da Universidade de Brasília com o equipamento mosquito (TTP labtech) em placas de 96 poços na metodologia de gota pendurada por difusão de vapor, com três gotas por poço variando-se suas composições misturando-se proteína e as soluções de cristalização nas seguintes proporções 200 nL + 100 nL, 100 nL + 100 nL, e 100 nL + 200 nL. As etapas de refinamento ocorreram em placas de 24 poços na metodologia de gota sentada por difusão de vapor.

4.1.13.1. Zika

Foram realizados testes da proteína de NS2b/NS3 de ZIKV na concentração de 2,1 mg/mL utilizando-se os kits Crystal Screen 1 e 2 (Hampton) e da protease de ZIKV (4 mg/mL) com aprotinina, na proporção 1/1, com os kits JBScreen JCSG++ 1, 2, 3 e 4, todo em placa de 96 poços utilizando-se a técnica de difusão de vapor por gota suspensa (*hanging drop*).

No experimento com somente a protease, obtiveram-se cristais na condição Dioxano 10 %, sulfato de amônio 1,6 M e 0,1 M MES pH 6,5. Assim, realizou-se refinamento da condição utilizando-se placas de 24 poços em que variaram-se as concentrações de sulfato de amônio entre 1,2 M e 2,2 M e o pH entre 6,1 e 6,7.

4.1.13.2. Dengue

As placas da etapa de *screening* para a protease NS2b/NS3 de DENV foram feitas usando-se os *kits* JBScreen JCSG++ e PACT++. Foram utilizadas as concentrações da protease de DENV a 2,5 mg/mL e 17 mg/mL.

4.1.13.3. Chikungunya

O *screening* da condição de cristalização da protease nsP2 de CHIK foi feito com duas concentrações de proteína, 7,33 mg/mL e 12 mg/mL, mediante o uso de amicon ultra 15. Utilizaram-se os kits de cristalização Salt RX, PEG RX e PEG/ion (Hampton Research). Foram usadas placas de 96 poços utilizando-se o equipamento Crystal

Gryphon (Art Robbins Instruments) e a técnica de difusão de vapor por gota pendurada. As placas foram armazenadas em temperatura ambiente.

4.1.14. Coleta de dados e refinamento

As coletas de dados de difração dos cristais de proteína foram realizadas em dois diferentes equipamentos. Na primeira foi utilizado o difratômetro D8 Venture (Bruker), equipado com o detector Photon II e sistema de ânodo rotatório *microfocus* TXS de cobre. Este se encontra localizado no Laboratório de Biofísica Molecular da Universidade de Brasília. Na segunda, foi utilizada a linha MX2 no Laboratório Nacional de Luz Síncrotron (LNLS), situado em Campinas, que está equipada com um detector Pilatus 2M (Dectris).

4.2. EXPERIMENTOS *IN SILICO*

4.2.1. Modelagem por homologia

Realizou-se a modelagem por homologia da proteína nsP2 de MAYV com o programa Modeller (WEBB; SALI, 2016). A seleção das proteínas homólogas foi feita através de alinhamentos contra o banco de dados do PDB (Blast), na qual foram escolhidas as cadeias A, sendo as únicas cadeias disponíveis, das estruturas 2HWK (domínios protease e MTase da nsP2 de VEEV, com resolução 2,45 Å e 42,5% de identidade) e 3TRK (domínios protease e MTase da nsP2 de CHIK, com resolução 2,40 Å e 64,4% de identidade). Para o alinhamento múltiplo foi utilizado o software Clustal Omega. Dentre os modelos gerados selecionou-se um que apresentasse bons valores tanto de MolPDF e DOPE score. Em seguida, foi gerado o diagrama de Ramachandran para análise das posições dos ângulos Φ (phi) e ψ (psi), usando o site RAMPAGE. Além disso, utilizou-se o QMEAN (BENKERT; KÜNZLI; SCHWEDE, 2009) para análise de qualidade do modelo.

4.2.2 *Docking* proteína-peptídeo

Experimentos de *docking* foram realizados com a proteína rígida, ou seja, somente o ligante teve a possibilidade de fazer a rotação dos ângulos torcionais entre seus átomos. Foram selecionadas as a estrutura de CHIK (3TRK) e MAYV (modelo). Foram desenhados peptídeos com os sítios de clivagem que tinham 10 resíduos de aminoácidos,

da posição P6 a p4', sendo que estes possuem os sítios de clivagem específicos para cada protease em questão. Os sítios são referentes às sequências entre as proteínas virais nsP1-nsP2, nsP2-nsP3 e nsP3-nsP4 e podem.

Tabela 4 – Sequência dos substratos das proteases nsP2 de MAYV e CHIK.

Virus	P6P5P4P3P2P1 / P1'P2'P3'P4' sequência do substrato		
	nsP1/nsP2	nsP2/nsP3	nsP3/nsP4
MAYV	EFRAGA/GVVE	LQAAGC/APAY	LGRAGA/YIFS
CHIK	EDRAGA/GIIE	VTRAGC/APSY	LDRAGG/YIFS

O programa utilizado para a execução e as análises de *docking* foi o AutoDock Vina (TROTT; OLSON, 2010). Antes de realizar este experimento *in silico* prepararam-se tanto a proteína quanto o ligante para a execução, para isso utilizou-se o pacote MGLTools, o qual define as protonações dos resíduos. Primeiramente realizou-se a preparação da proteína/receptor, sendo assim, retiraram-se todas as moléculas de água da estrutura, e adicionou-se os hidrogênios polares (se estes não estão presentes no arquivo do PDB), e selecionou-se o volume em que foi feita a procura das poses (*grid box*) (CHIK - center_x = -0.684, center_y = 25.548, center_z = 26.326, size_x = 26, size_y = 34 e size_z = 30 e MAYV - center_x = -5.27, center_y = 0.667, center_z = 50.812, size_x = 26, size_y = 34 e size_z = 30), onde os valores de size são medidos em Å, o qual compreende toda a região do sítio ativo.

Após isso, realizou-se a preparação do ligante, ao carregar o composto no programa, automaticamente é detectada a carga do ligante, além disso, cada átomo presente na molécula é classificado e as ligações passíveis de rotação detectadas. Para realização dos experimentos selecionou-se um valor de *exhaustiveness* fixa de 8, com valor do *num_modes* de 20, relacionado ao número de poses selecionados no arquivo de saída, e o *energy_range* igual a 10. As análises das poses foram feitas utilizando-se o programa Pymol para gerar as imagens.

Os cálculos de afinidade de ligação entre a proteína e os substratos foram feitos pelo próprio AutoDock Vina. Este valor de energia livre de interação, com unidades de kcal/mol, é obtido usando-se somente os dados intermoleculares, conforme a equação (5).

$$s_1 = g(c_1 - c_{intra1}) = g(c_{inter1}) \quad (5)$$

Nesta equação, s é a afinidade (energia livre de interação), g uma função não linear, e c_{intra} e c_{inter} estão relacionados às contribuições obtidas entre as próprias moléculas e entre as diferentes moléculas (ligante e proteína), respectivamente. O valor de c , soma dos valores das interações intermoleculares e intramoleculares, pode ser obtido pela equação (6).

$$c = \sum_{i < j} f_{t_{ij}}(r_{ij}) \quad (6)$$

Aqui, por sua vez, é observada a soma de um par de átomos i e j , sendo t o tipo do átomo, r as distâncias interatômicas e f um conjunto de funções de interação obtido a partir da equação (7).

$$f_{t_{ij}}(r_{ij}) \equiv h_{t_{ij}}(d_{ij}) \quad (7)$$

Nesta última equação, d está relacionado a distância entre as superfícies de dois átomos, i e j , onde r_{ij} é subtraído dos valores dos raios de van der Waals dos átomos i e j . Por sua vez, h é a soma de determinados pesos de três diferentes interações, um termo hidrofóbico (gauss 1 com peso de -0,0356), um termo de ligação de hidrogênio (gauss 2 com peso de -0,00516), e um de repulsão (0,840), conforme as equações (8), (9) e (10).

$$\text{gauss}_1(d) = e^{-(d/0.5\text{\AA})^2} \quad (8)$$

$$\text{gauss}_2(d) = e^{-((d-3\text{\AA})/2\text{\AA})^2} \quad (9)$$

$$\begin{aligned} \text{repulsion}(d) &= d^2, \text{if } d < 0 \\ \text{repulsion}(d) &= 0, \text{if } d \geq 0 \end{aligned} \quad (10)$$

4.2.3. Dinâmica molecular (MD)

A dinâmica molecular foi realizada com o programa Gromacs (PRONK et al., 2013), versão 5.1, com o campo de força gromos 54A7 (SCHMID et al., 2011). As estruturas

utilizadas foram as mesmas usadas nos experimentos de *docking*, ou seja, para a nsP2 de CHIK a 3TRK e para a de MAYV a modelada. Executaram-se algumas dinâmicas das proteínas sem ligante (50 ns) e com os peptídeos referentes aos sítios nsP3-nsP4 de CHIK e nsP2-nsP3 (100 ns), as quais foram selecionadas por serem as poses com maior proximidade do carbono da carbonila do resíduo P1 e o enxofre da cisteína catalítica e por bons resultados de *score* no *docking*.

Após isso, definiu-se a caixa periódica, um dodecaedro com distância mínima de 0,7 nm entre a proteína e as bordas da caixa, e foi feita a solvatação, usando o modelo de água spc216, que levou à presença em torno de 60.000 moléculas de solvente por MD. Foi realizada a adição de íons, em geral oito íons de Na⁺. Realizou-se a minimização de energia com o algoritmo *steepest descent*.

Duas fases de equilíbrio foram feitas, a primeira era relacionada à temperatura (200 ps), na qual o sistema alcançou o valor de 300 K e ficou estável. Em seguida foi feita a segunda etapa de equilíbrio, a de pressão (200 ps). A etapa da MD produtiva foi feita logo depois utilizando-se versão compatível com GPU.

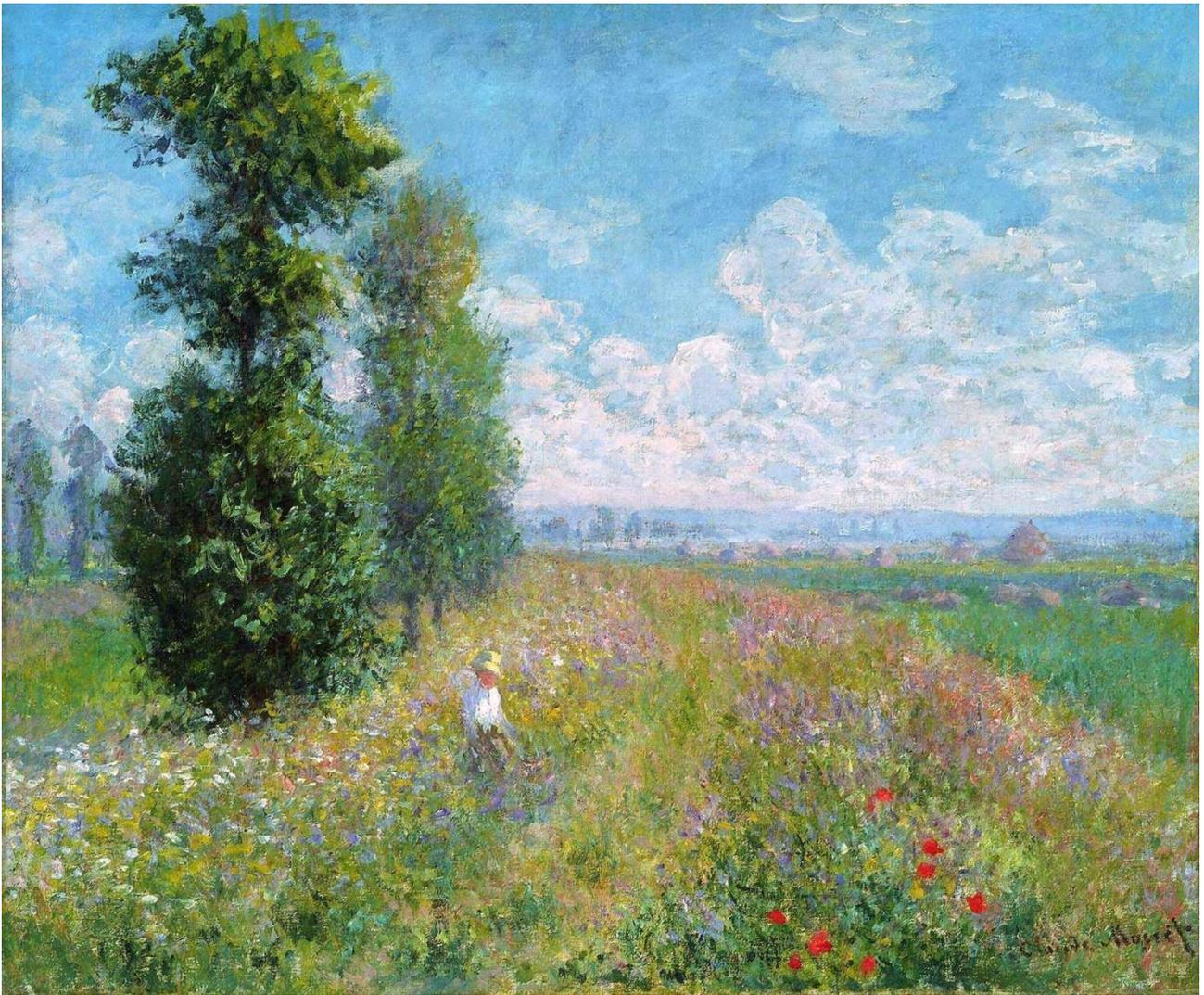
Os dados de RMSD e RMSF foram gerados por programas internos do Gromacs, *gmx rms* e *gmx rmsf*, respectivamente, e os gráficos foram gerados pelo programa Grace, utilizando a interface gráfica do usuário (GUI) *xmgrace*.

Gerou-se também, usando o programa *gmx rama*, os ângulos ϕ e ψ dos resíduos da proteína durante a dinâmica, os quais foram filtrados usando-se o comando *grep* do terminal linux para selecionar somente os valores referentes aos resíduos P3, P2, P1, P1' e P2' do substrato e, assim, um gráfico de Ramachandran foi gerado. Um *script* foi escrito para retirar as casas decimais dos valores de ângulos ϕ e ψ e dar o número de repetições dos mesmos, visando a confecção de um gráfico de Ramachandran 3D feito no programa Origin, este último foi feito somente para a glicina na posição P2. Além disso, foi gerado um gráfico com os valores dos ângulos χ_1 dos resíduos C478, H548 e S/C482, e χ_2 da H548, visando ver a variação dos mesmos na presença e na ausência dos substratos. Estas análises foram feitas tanto para CHIK quanto para MAYV.

4.2.4. Docking proteína-proteína

O experimento de *docking* proteína-proteína, NS2b/NS3 de ZIKV com o inibidor de protease aprotinina, foi realizado utilizando o *web server* ClusPro (KOZAKOV et al., 2017).

O servidor faz um docking de corpo rígido selecionando 1000 resultados, agrupa os resultados próximos e seleciona os melhores *scores* para uma etapa de refinamento visando remover problemas como possíveis *clashes*, sendo selecionadas ao final dez poses. Estas foram sobrepostas ao PDB 3U1J (NOBLE et al., 2012), o qual apresenta a protease do Flavivirus DENV3 em interação com aprotinina, para se selecionar a pose mais próxima da encontrada interagindo com esta proteína homóloga. Os cálculos de superfície molecular acessível do resíduo W50 foram feitos utilizando-se o servidor WHAT IF (VRIEND, 1990).



Meadow with Poplars
1875 54x65cm oil/canvas
Claude Monet

5. Resultados e Discussão

5. RESULTADOS E DISCUSSÃO

Alguns dos genes das proteases dos diferentes arbovírus aqui estudados (NS3 de ZIKV e DENV, e nsP2 de CHIK e MAYV) foram expressos, e este processo foi otimizado para a obtenção de uma maior quantidade de proteína na fração solúvel. Primeiramente, as construções obtidas foram transformadas em diversas cepas de *E. coli*, e crescidas em placas de petri contendo meio LB com antibióticos para os quais os plasmídeos presentes ofereciam resistência (tanto os inseridos na transformação, quanto os presentes em algumas cepas), etapa na qual se selecionou somente micro-organismos com o plasmídeo para expressar nosso gene de interesse.

Realizou-se a expressão a 37°C e a indução com 1 mM de IPTG visando em primeiro lugar verificar o correto funcionamento da transformação e a presença da proteína heteróloga na fração solúvel. As bandas de expressão das proteínas podem ser observadas na figura 9, todavia, estas estavam localizadas nas na fração insolúvel por estarem presentes nas amostras de *pellet*. A cepa BL21 (DE3) apresentou expressão de todos os genes (figura 9A), entretanto somente o nsP2 de CHIK apresentou expressão em outras cepas, BL21 (DE3) pLysE e Lemo21 (figura 9B e C). Em cepas como a BL21(DE3) Rosetta não foram observadas presenças claras de bandas de expressão das proteases (figura 9D).

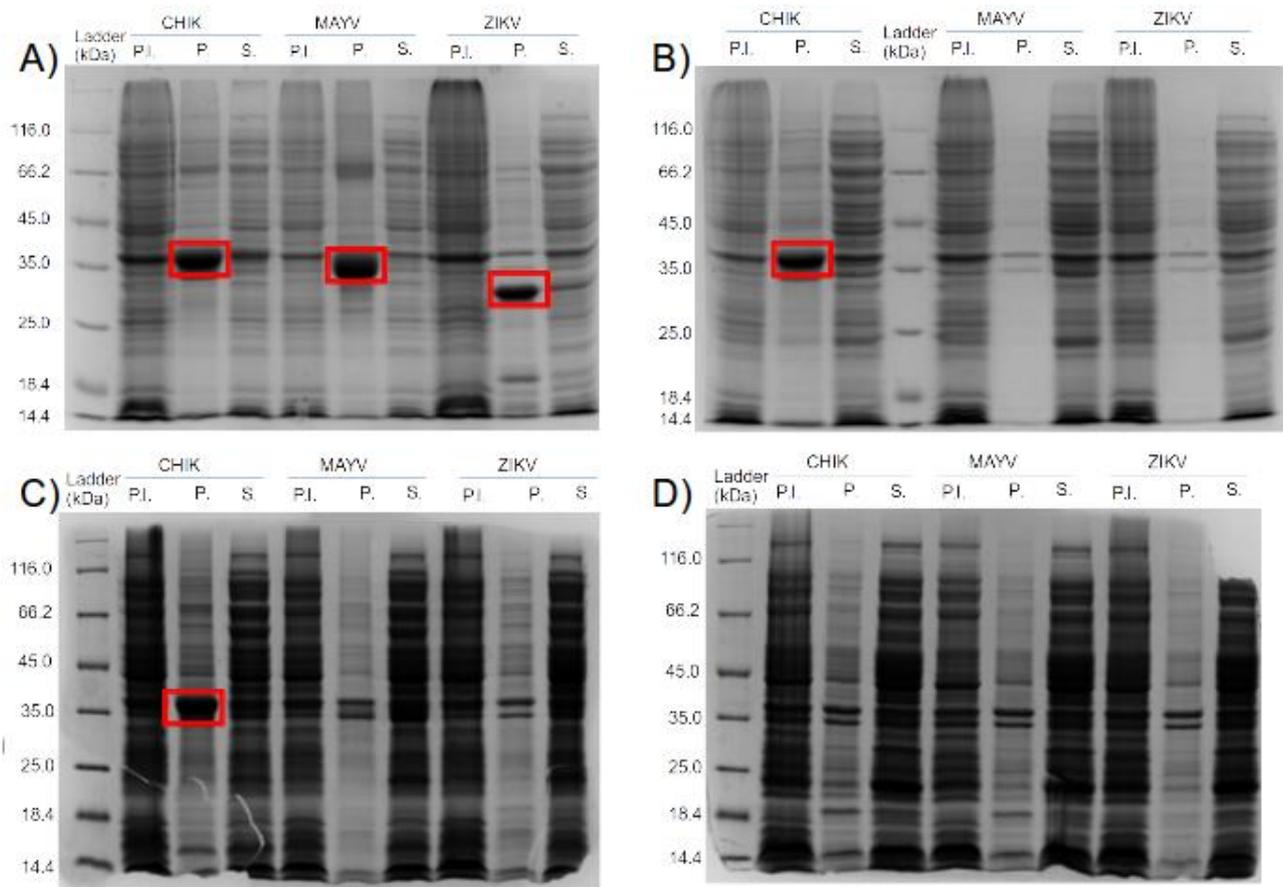


Figura 9- Expressão dos genes nsP2 de CHIK e MAYV, e NS3 de ZIKV a 37°C na cepa de *E. coli* (A) BL21(DE3), (B) BL21(DE3) pLysE, (C) BL21(DE3) Lemo21, (d) BL21(DE3) Rosetta. P. I. marca o padrão de expressão antes da indução. P. as proteínas na fração insolúvel e S. na solúvel.

Objetivando-se obter as proteínas expressas na fração solúvel a temperatura do experimento foi reduzida para 19°C. Observou-se que a protease tanto de CHIK quanto de ZIKV nessa nova condição de expressão puderam ser observadas em solução em quantidades significativas, entretanto a de MAYV não se mostrou com o mesmo sucesso, como se observa na figura 10.

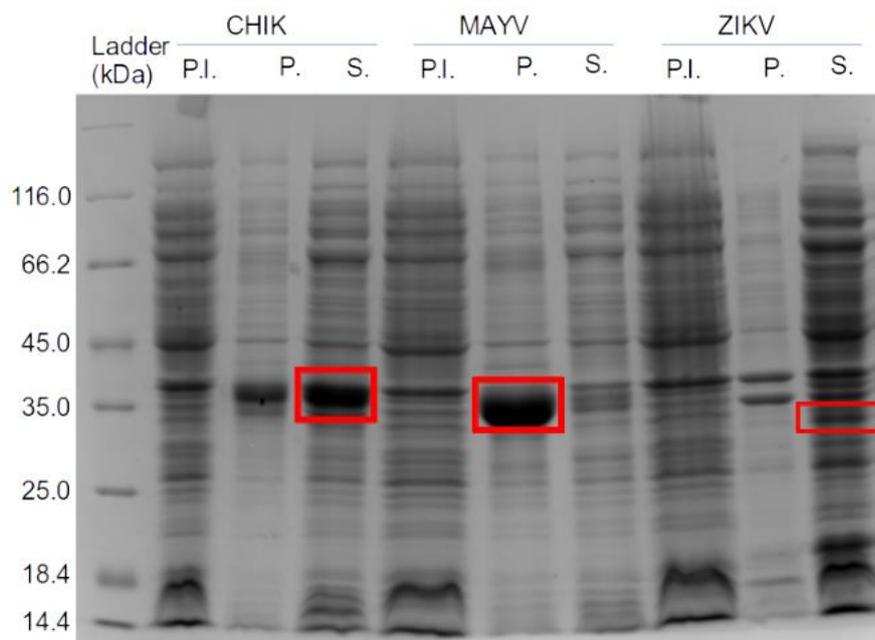


Figura 10- Expressão e teste de solubilidade das proteínas nsP2 de CHIK e MAYV, e NS3 de ZIKV na cepa BL21 (DE3) de *E.coli* à temperatura 19°C a 1 mM de IPTG. P. I. marca o padrão de expressão antes da indução. P. as proteínas na fração insolúvel e S. na solúvel.

As etapas de otimização da expressão e purificação, bem como as diversas análises realizadas para cada uma das proteínas anteriormente citadas, são tratadas de forma específica a seguir.

5.1. ALPHAVIRUS

5.1.1. Experimentos *in vitro*

5.1.1.1. Mayaro

A primeira etapa foi a transformação de diversas cepas de *E. coli* com o plasmídeo pET-21a contendo o gene da nsP2 de MAYV utilizando-se a metodologia de choque térmico. O experimento apresentou um resultado positivo, ou seja, diversas colônias de diferentes cepas cresceram em meio LB sólido contendo antibiótico a uma temperatura de 37°C. Confirmou-se que a protease de MAYV foi então expressa a 37°C durante 6 horas a uma concentração de 1 mM de IPTG, porém, na fração insolúvel, o que ocorreu também a 19°C.

Na otimização da expressão, dentre as tentativas de obtenção dela na fração solúvel realizou-se redução de temperatura para 16°C e de concentração de IPTG para

0,2 mM. A expressão foi realizada por oito horas e alíquotas foram coletadas a cada hora para posterior teste de solubilidade. Este experimento tinha como objetivo observar se nos momentos iniciais de expressão a baixa concentração de IPTG e temperatura era possível encontrar uma condição para produção da proteína solúvel. Observou-se a expressão crescente da proteína heteróloga (figura 11A), todavia a maior parte da amostra ainda foi observada no *pellet* (Figuras 11B e 11C).

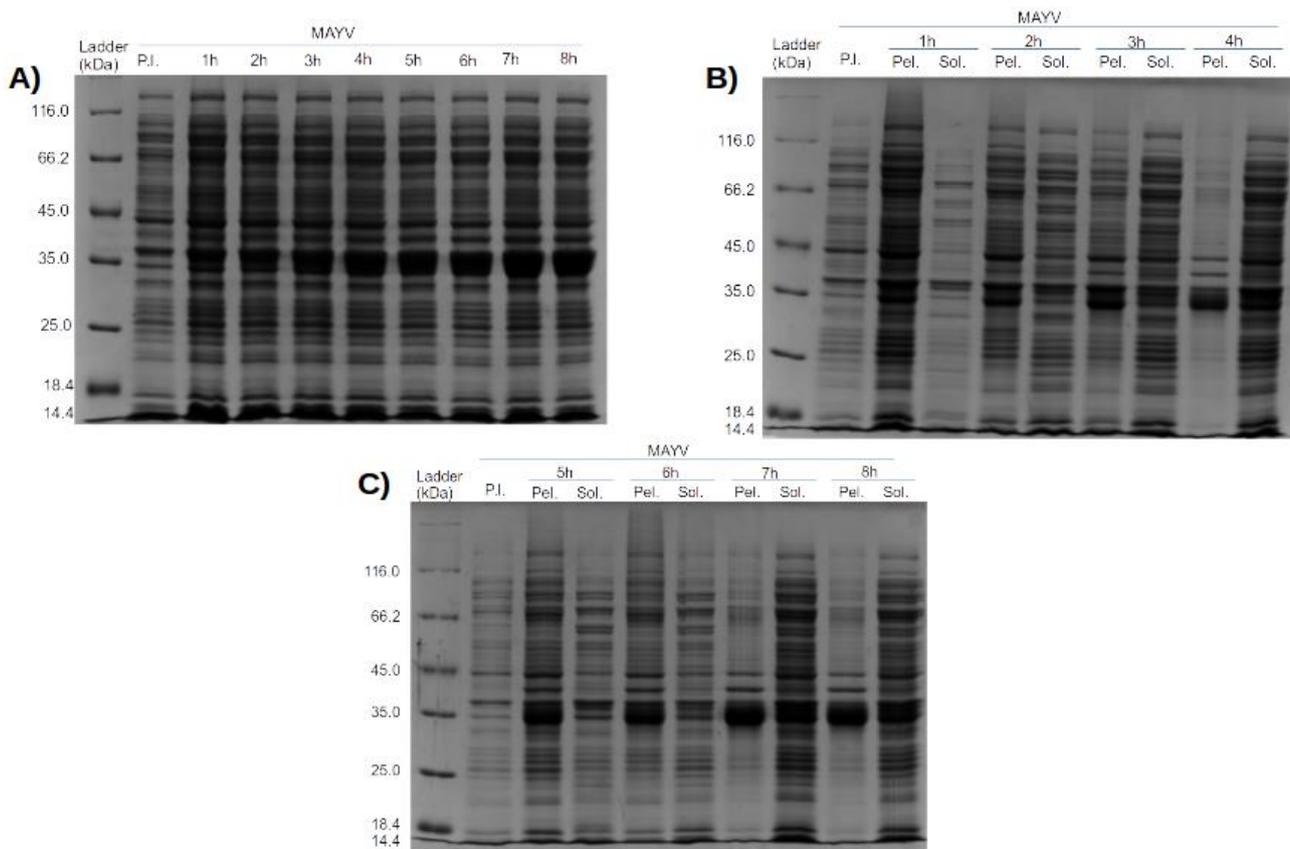


Figura 11- (A) Expressão e (B e C) teste de solubilidade das proteína nsP2 de MAYV na cepa BL21 (DE3) de *E.coli* a temperatura 16°C a 0,2 mM de IPTG em meio LB. P. I. marca o padrão de expressão antes da indução. Pel. as proteínas na fração insolúvel e Sol. na solúvel.

Foram realizados outros testes alterando-se o meio de cultura usado durante a expressão. Utilizou-se também o meio autoindutor visando obter alguma condição de expressão onde a protease fosse expressa de modo solúvel. Este meio de cultura não possui a presença da molécula análoga de alolactose, o IPTG. A expressão ocorre através da redução dos níveis de glicose e a ativação do operon Lac pela presença de

lactose, o que leva ao consequente desacoplamento do repressor Lac que bloqueava a região promotora do operon em questão. Este meio, em geral, apresenta maior produção da proteína alvo e densidade celular por volume de cultura. Além disso, o mesmo possui menor taxa de formação de corpos de inclusão, possivelmente devido a uma taxa de crescimento mais lenta (FAUST; STAND; WEUSTER- BOTZ, 2015; KRAUSE; NEUBAUER; NEUBAUER, 2016; STUDIER, 2005). Todavia, como pode ser observado na figura 12, durante o teste de solubilidade não foram observadas bandas claras de expressão na fração solúvel.

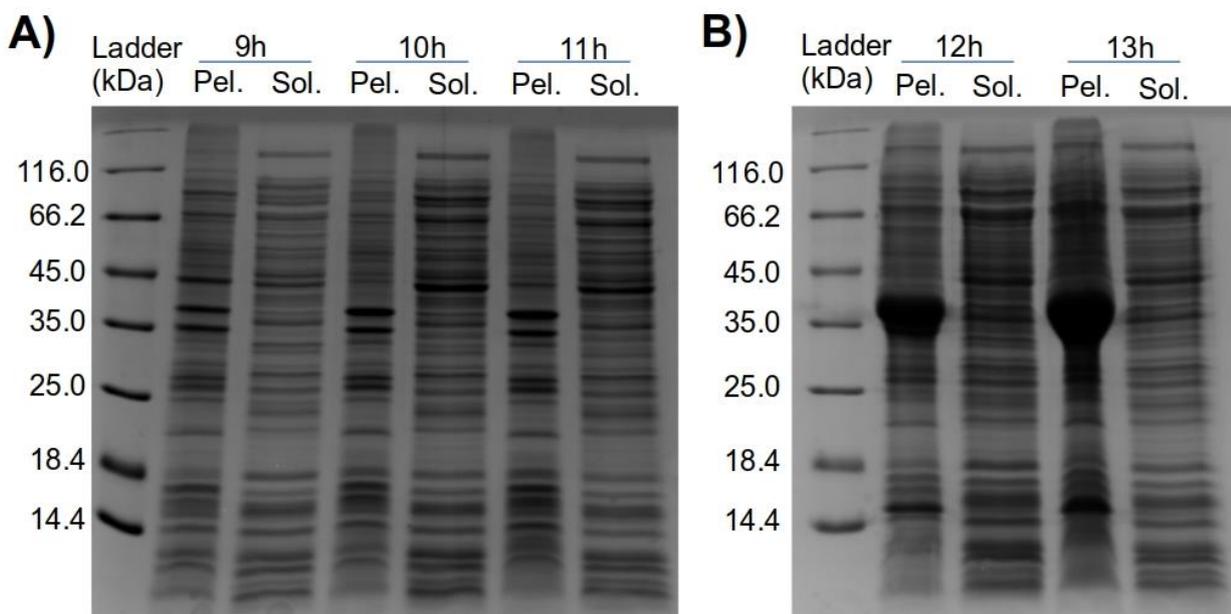


Figura 12- Teste de solubilidade da expressão em meio autoindutor da protease nsP2 de MAYV. A) Apresenta a fração insolúvel (Pel.) e solúvel (Sol.) das amostras coletadas nos tempos 9h, 10h e 11h, B) e das referentes a 12h e 13h.

Devido a não obtenção da proteína na fração solúvel, foram feitos novos testes de expressão utilizando-se diferentes cepas com o mesmo plasmídeo pET-21a anteriormente utilizado, além disso, mudando-se a composição do tampão de lise de acordo com alguns estudos anteriormente publicados (HU et al., 2016; SAISAWANG et al., 2015a). Este novo tampão é composto por 20mM Tris-Hcl, 200mM NaCl, 200mM arginina, 200mM ácido glutâmico, pH 7,1 e glicerol 5%. Foi também alterado o meio de cultura para meio de LB para YT. A protease foi expressa a pela adição de 1 mM de IPTG quando a densidade de células no meio atingiu a O.D. de 0,5, as quais foram mantidas *overnight* por 14 horas a

temperatura de 25°C. Apesar disso, não foi observado nenhuma banda significativa na fração solúvel. Todavia, em análises de *western blot* usando anticorpos com especificidade para cauda da histidina foram observadas bandas na fração solúvel em tamanhos aproximados ao esperado em algumas cepas como: BL21, pLysE, pLysS e C43, como pode ser observado na figura 13.

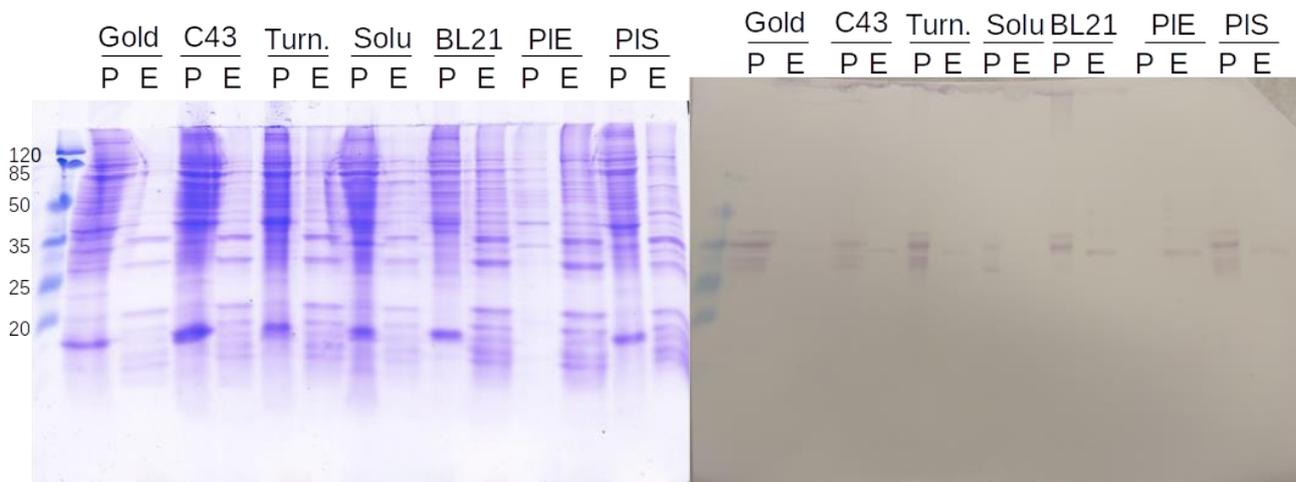


Figura 13- Análise por gel SDS-PAGE 15% da expressão em diversas cepas de *E. coli* da protease nsP2 de MAYV corado por comassie (esquerda) e western blot (direita). Este foi realizado usando-se tampão contendo 20mM Tris-Hcl, 200mM NaCl, 200mM arginina, 200mM ácido glutâmico, pH 7.1 e glicerol 5%. As cepas usadas foram aplicadas no gel na seguinte ordem: BL21-Gold (DE3), C43 (DE3), Turner (DE3), SoluBL21, BL21 (DE3), BL21 (DE3) pLysE e BL21 (DE3) PlysS. P se refere à fração insolúvel da amostra e E à fração solúvel.

Alguns componentes do tampão utilizado poderiam explicar uma possível produção da proteína heteróloga na fração solúvel. Arginina e ácido glutâmico foram alguns dos componentes utilizados no tampão, estes têm sido apresentados em estudos como tendo a capacidade de inibir a agregação e precipitação, aumentarem a estabilidade da amostra, protegem de degradação proteolítica, além de auxiliarem na renaturação de proteínas. Uma das possíveis formas de ação é através da interação com cadeias laterais carregadas na proteína e também por auxiliar a cobrir resíduos hidrofóbicos expostos com porções alifáticas, o que poderia inibir a agregação. Observou-se também que o dipeptídeo ArgGlu em análises com lisozima e BSA gerou efeitos de hidratação preferencial, podendo ser outra possibilidade de explicação para as propriedades

apresentadas por estas moléculas. Todavia, altas concentrações de Arginina, acima de 0,5 M, parecem realizar efeitos inversos, levando a desestabilização e diminuindo a temperatura de *melting* (GOLOVANOV et al., 2004; KITA et al., 1994; SHUKLA; TROUT, 2011).

O glicerol foi outro componente utilizado no tampão de lise, o qual atua induzindo a compactação da proteína, reduz a flexibilidade (GONNELLI; STRAMBINI, 1993) e estabiliza conformações intermediárias da proteína (MISHRA; BHAT; SECKLER, 2007), aumentando assim a estabilidade da proteína. O modo pelo qual essa molécula atua tem sido alvo de diversos estudos que têm indicado que sua atuação como estabilizador se daria através da preservação da hidratação da proteína devido a fenômenos como o da hidratação preferencial e do volume-excluído, o qual excluiria preferencialmente o glicerol da camada de solvatação da proteína criando uma tendência de minimização da superfície proteica (HIRAI et al., 2018). Outra sugestão é que atue através da redução do número e do tamanho de espaços na estrutura proteica que poderiam acomodar moléculas de água (PRIEV et al., 1996). Além disso, este poliálcool inibe a agregação de proteínas impedindo tanto a exposição parcial do núcleo hidrofóbico quanto a redução do contato proteico. Isto foi observado durante análises por dinâmica molecular de proteínas na presença de glicerol, na qual algumas superfícies hidrofóbicas favorecem o contato com orientações anfífilas desta molécula (VAGENENDE; YAP; TROUT, 2009).

Devido à presença dessas bandas foram selecionadas as cepas BL21 e pLysE para expressão em maiores volumes de meio de cultura e purificação por cromatografia líquida de afinidade a níquel, cujos resultados podem ser observados na figura 14. Em ambos os casos, pode ser observada uma banda na fração solúvel após a eluição com imidazol, todavia, a banda não se mostrou no tamanho esperado, visto que o tamanho da proteína é de 38,2 kDa, e as bandas observadas se aproximam mais do tamanho de 25 kDa.

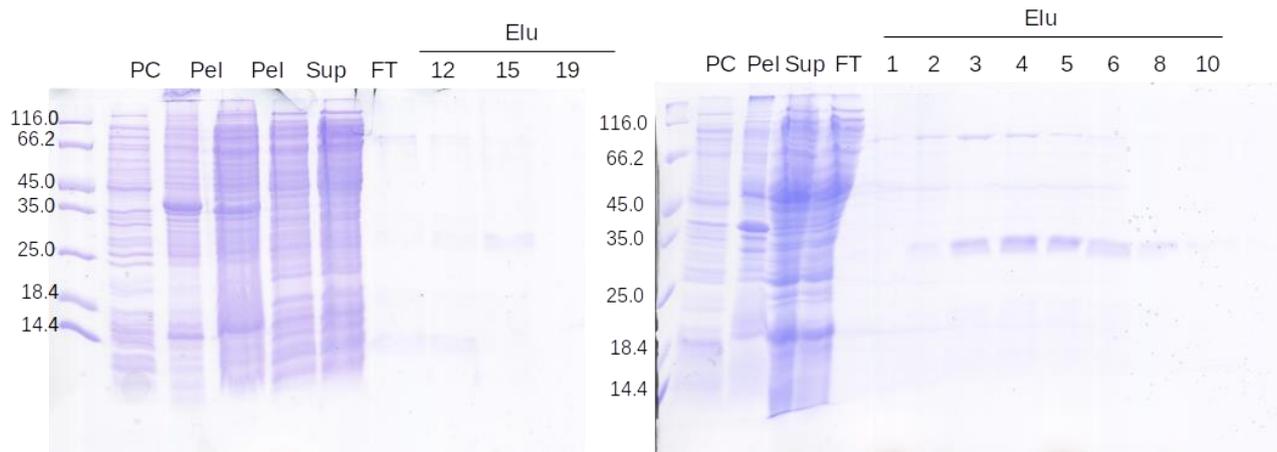


Figura 14- Análise da expressão e purificação da protease nsP2 de MAYV com 1 mM de IPTG, 25°C e 14 horas por SDS-PAGE 15%. Este foi realizado usando tampão contendo 20mM Tris-HCl, 200mM NaCl, 200mM arginina, 200mM ácido glutâmico, pH 7.1. A esquerda utilizando a cepa BL21 e a direita pLysE. PC refere-se a amostra antes da indução por IPTG, Pel a fração insolúvel, Sup a fração solúvel, FT a fração que não interagiu quando aplicado na coluna de níquel, os demais números à frações coletadas durante gradiente de imidazol.

Mesmo com essa diferença de tamanho as amostras foram concentradas e passou-se para a etapa de remoção da cauda de histidina utilizando a protease HRV3C. As amostras conjugadas foram então submetidas a cromatografia líquida usando coluna de GST para remoção da HRV3C. Entretanto, não foi observada uma diminuição do tamanho da banda após essa etapa. Este resultado pode ser observado na figura 15.

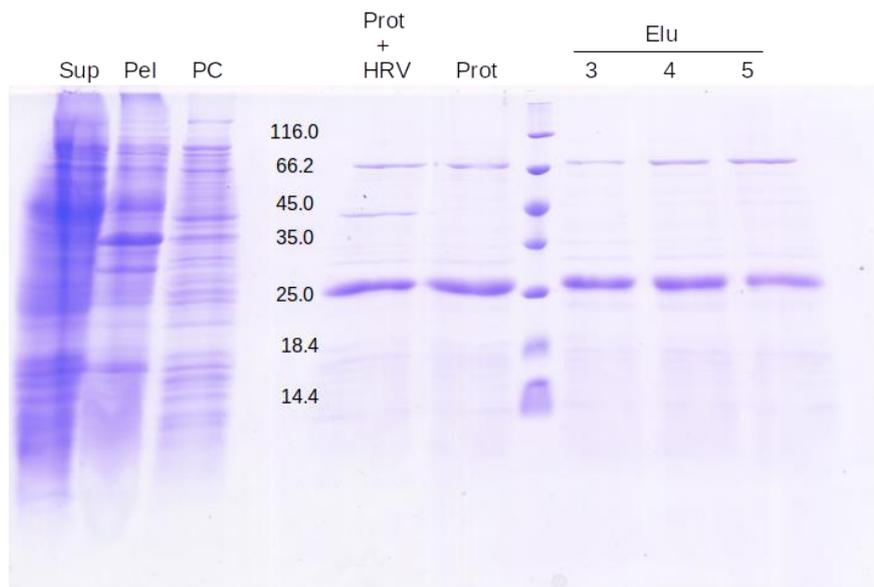


Figura 15- Análise por SDS-PAGE 15% da remoção de cauda de histidina da proteína nsP2 de MAYV pela HRV3C. PC se refere a amostra antes da indução por IPTG, Pel a fração insolúvel e Sup a fração solúvel durante a etapa de expressão. Prot+HRV mostra a proteína junto com a protease HRV3C, enquanto Prot se refere a proteína não tratada para remoção da cauda. Os números representam os diferentes tubos coletados após a remoção da HRV3C por coluna de GST.

O mesmo método de expressão e purificação, em que se utilizou um novo tampão, foi realizado na temperatura de 18°C para as cepas que apresentaram bandas na fração solúvel. Todavia, observou-se padrão muito similar de expressão com a banda mais visível em um tamanho próximo a 25 kDa (figura 16).

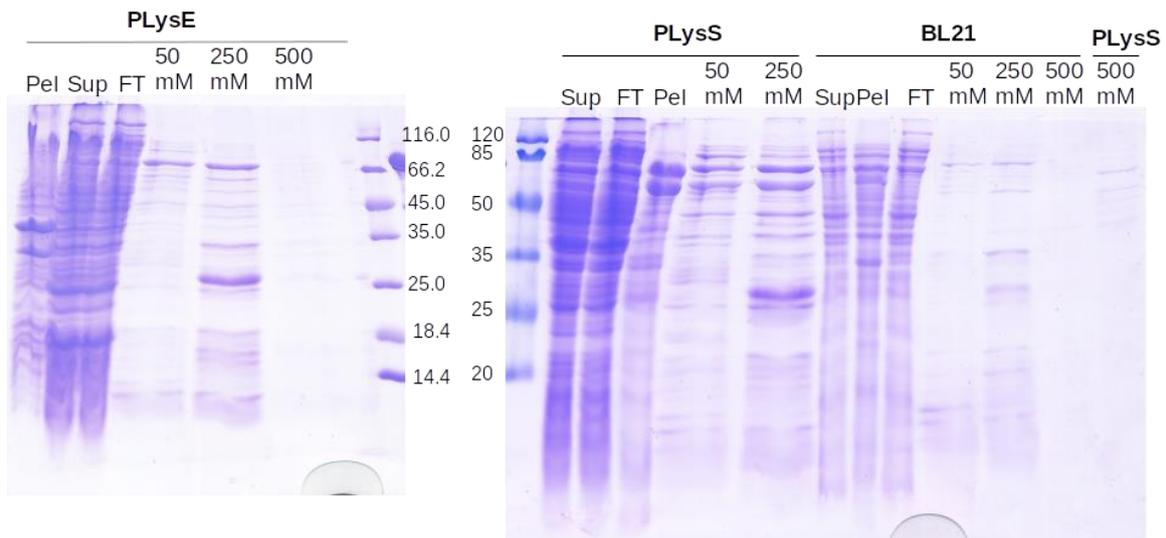


Figura 16- Análise por SDS-PAGE 15% da expressão da protease nsP2 de MAYV usando diferentes cepas BL21, pLysE e pLysS. Este foi realizado em tampão contendo 20mM Tris-HCl, 200mM NaCl, 200mM arginina, 200mM ácido glutâmico, pH 7.1. Pel se refere a fração insolúvel, Sup a fração solúvel durante a etapa de expressão, e FT com a amostra que não apresentou interação quando aplicada a coluna de níquel. Os valores referem-se à concentração de imidazol na qual a amostra foi eluída.

Devido ao fato de não serem obtidos resultados claros da expressão da protease nsP2 de MAYV foi então realizada a confecção de quatro outras construções, na qual foram usados os vetores de *E. coli* pGEX, pET SUMO, pET28a e pET15b. Para isto foram obtidos *primers* objetivando amplificar o gene da protease em questão adicionando-se novos sítios de restrição nas extremidades do gene, sendo estes BamHI e XhoI para pGEX e pet SUMO, e NdeI e XhoI para pET15b e pET28a. Estas foram então submetidas às etapas de amplificação do segmento gênico da protease de MAYV (figura 17A), clivagem pelas enzimas de restrição tanto dos vetores (figura 17B) quanto do material genético, ligação, transformação de cepas de *E. coli*. Selecionaram-se algumas colônias com o objetivo de confirmar a inserção do gene no plasmídeo através da técnica de PCR de colônia (figura 17C).

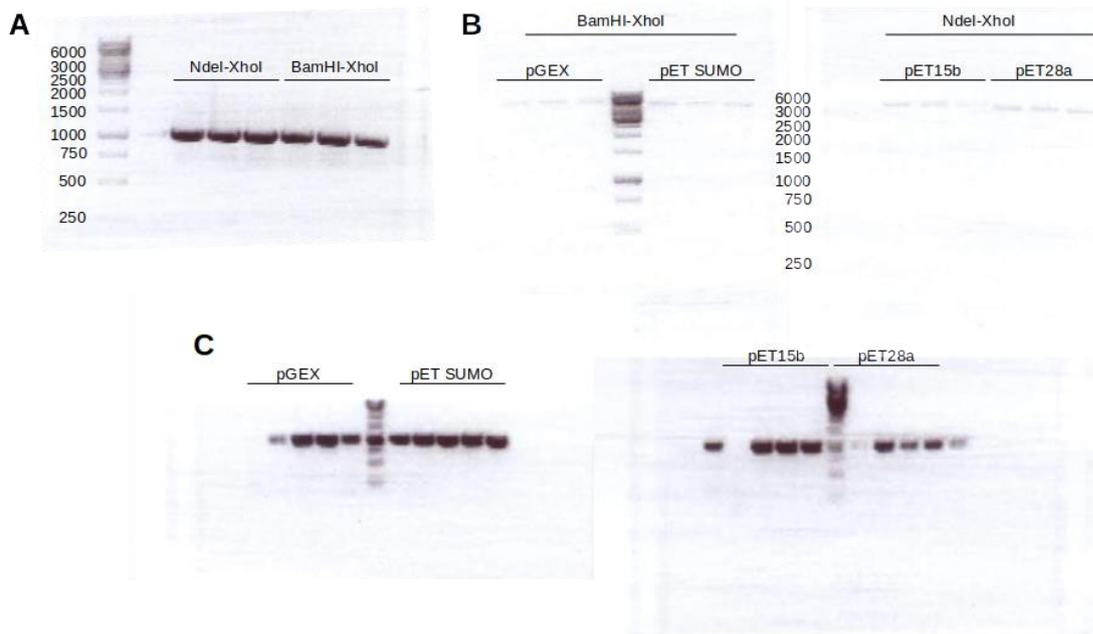


Figura 17- Clonagem do gene da fração protease nsP2 de MAYV nos vetores pGEX, pET SUMO, pET28a e pET15b. (A) Amplificação de segmentos gênicos com adição de sequências para os sítios de restrição para as endonucleases NdeI, XhoI e BamHI. (B) Digestão dos vetores selecionados com enzimas de restrição. (B) PCR de colônia para demonstrar a inserção do gene nos vetores.

As novas construções foram então inseridas em cepa de *E. coli* BL21 DE3 visando testar a expressão da protease, em meio YT a 18°C com 1 mM de IPTG por 16 horas. Todavia, durante estes testes iniciais de otimização não foram observadas em gel SDS-PAGE 25% bandas claras de expressão dessa proteína, figura 18.

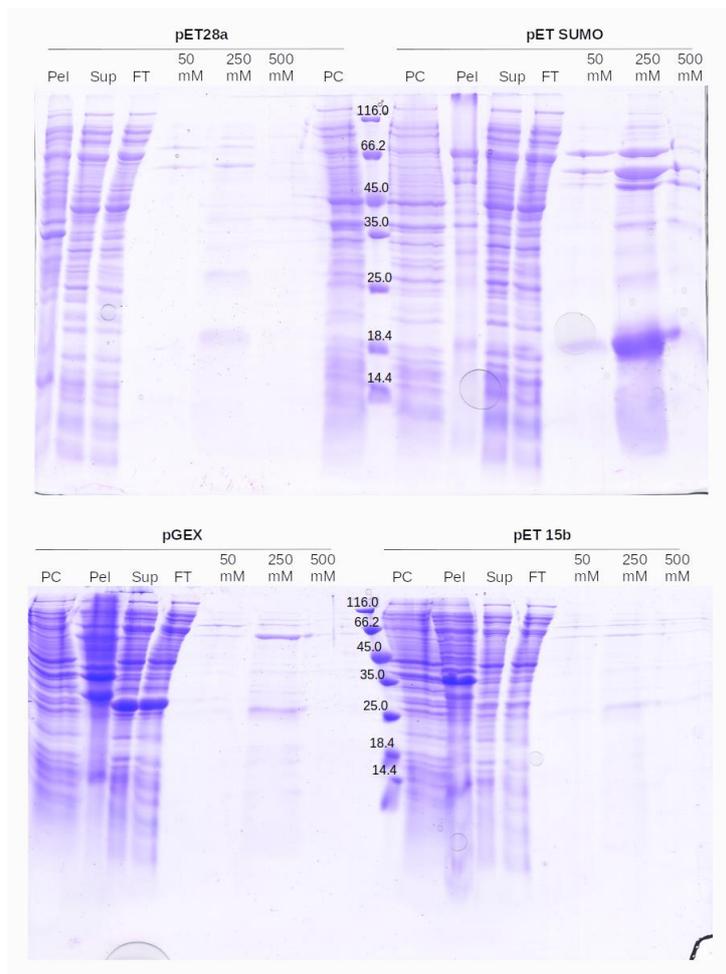


Figure 18- Análise da expressão de nsP2 de MAYV por SDS-PAGE 15% usando quatro diferentes vetores de expressão em bactéria *E.coli*, pET28a, pET SUMO, pGEX e pET 15b. PC se refere a amostra antes da indução por IPTG, Pel a fração insolúvel, Sup a fração solúvel, FT a fração que não interagiu com quando aplicado à coluna de níquel, seguido pelas eluições a três diferentes concentrações de imidazol (50, 250 e 500 mM).

Outra técnica que foi utilizada visando obter a proteína solúvel foi a renaturação. As proteínas presentes na fração insolúvel passaram pelo processo de desnaturação com ureia 8 M seguido de purificação em condições desnaturantes e renaturação por diluição rápida. Após essas etapas, a proteína obtida nesse processo foi submetida a análise pela técnica de CD. Este experimento visava analisar a renaturação proteica por meio da estrutura secundária da nsP2. Estes resultados podem ser vistos na figura 19.

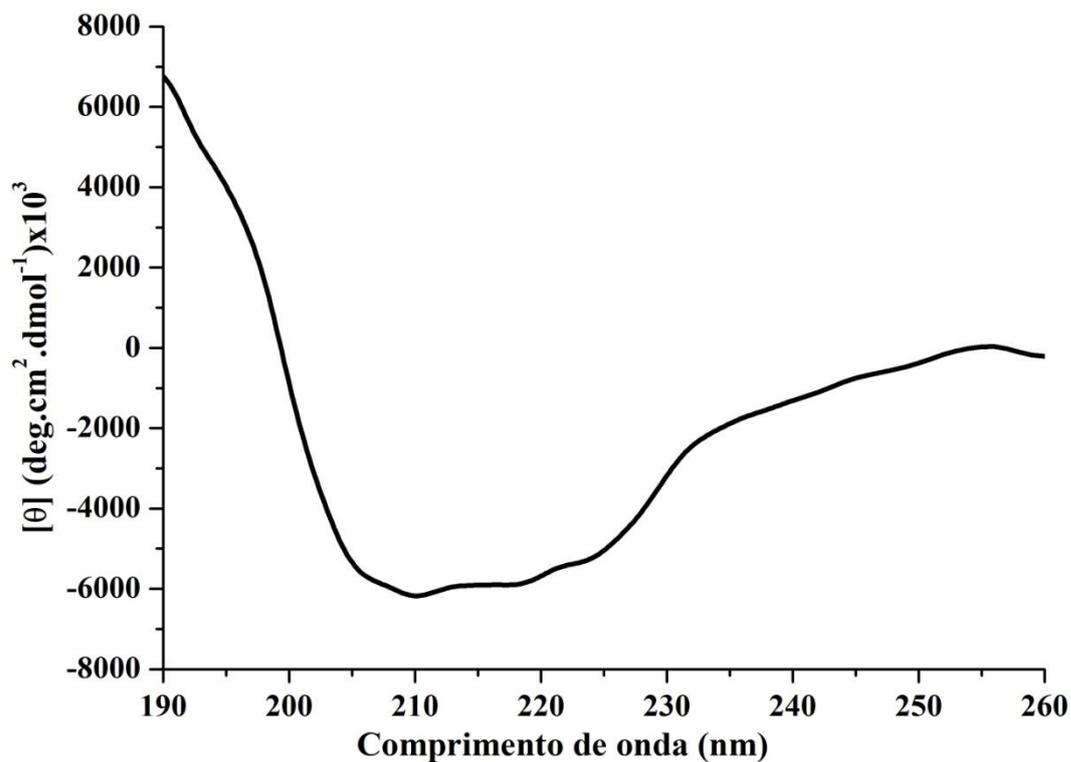


Figura 19- Espectro obtido pela técnica de CD da proteína nsP2 de MAYV após renaturação por diluição rápida.

O espectro da protease nsP2 demonstra a presença de estruturas secundárias na proteína renaturada, com destaque para o comprimento de onda 208 nm, sendo característico de folha- β , e o 222 nm, característico de α -hélice. A partir da deconvolução do espectro foi possível obter a composição de estrutura secundária da proteína. Estes dados e a comparação dos mesmos com proteínas homólogas podem ser observados na tabela 5.

Tabela 5- Porcentagem de estrutura secundária de proteases de Alphavirus e da deconvolução dos dados do espectro da protease nsP2 de MAYV.

Protein	Secondary structure		
	Helix(%)	Sheet(%)	Other(%)
nsP2MAYVmodel	33.7	20.1	46.1
nsP2MAYVcirDic23	17.1	34,2	53.5
nsP2MAYVcirDic33	13.2	31.3	54.7
nsp2CHIK(3trk)	31.2	19.4	49.4
nsP2VEEV(2hwk)	32.2	19.1	48.8

A partir dos dados obtidos por CD, pode observar-se uma considerável diferença entre os padrões de estrutura secundária da protease nsP2 renaturada e outras estruturas de Alphavirus depositadas no banco de dados PDB e para o modelo estrutural calculado neste trabalho. É possível observar um decréscimo na porcentagem de α -hélices, e um acréscimo na de folhas- β e outros. Apesar de serem observadas análises de CD da porção protease ou da proteína nsP2 completa para alguns Alphavirus (COMPTON et al., 2017; RAUSALU et al., 2016) não foram encontradas na literatura valores de estrutura secundária para comparação com os dados obtidos no presente trabalho. Todavia, os dados aqui obtidos parecem indicar que a porção protease em questão não apresentou um dobramento correto após ensaio de renaturação, mas são necessárias outras diferentes análises, como testes de atividade, para confirmação da capacidade catalítica da enzima renaturada.

5.1.1.2. Chikungunya

Na etapa de purificação por cromatografia líquida de afinidade a níquel do gene nsP2 de CHIK, a amostra foi ressuspensa em tampão A (50 mM NaH₂PO₄, 300 mM NaCl e pH 7,5) e aplicada na coluna. Realizaram-se lavagens com soluções com concentração de imidazol crescentes (50 mM, 100 mM e 250 mM), mas a proteína em questão foi eluída somente na lavagem com 500 mM de imidazol. O pico de eluição da proteína foi dividido em quatro tubos falcon, com diferentes concentrações de proteínas (figura 20A). As amostras purificadas foram submetidas a uma etapa de purificação por exclusão molecular, ilustrado na figura 20B.

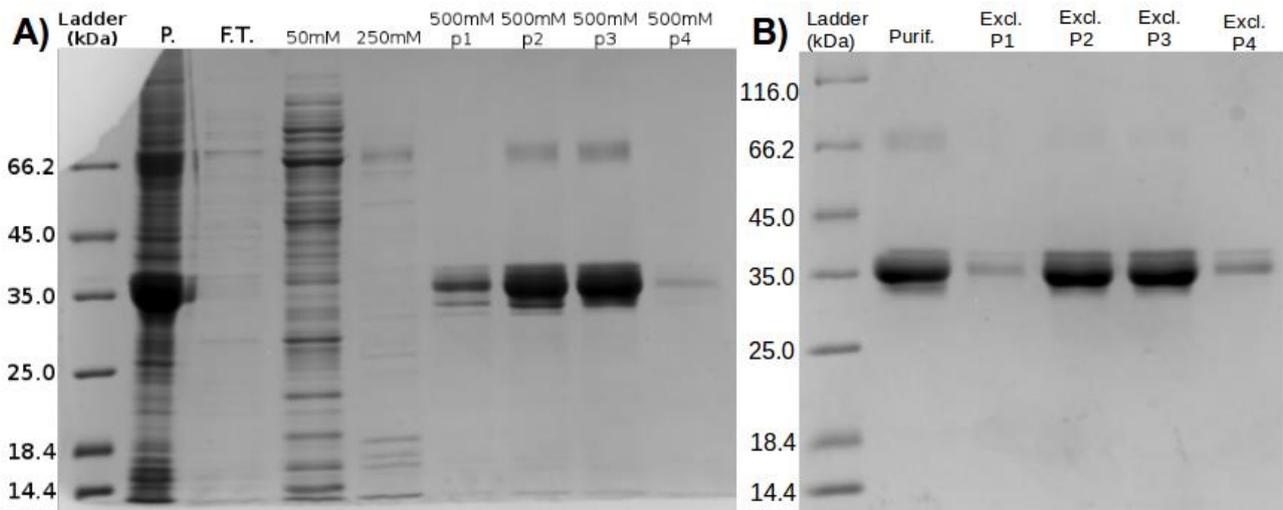


Figura 20- Purificação da protease nsP2 de CHIK. (A) Purificação por cromatografia líquida de afinidade a níquel. (B) Purificação por exclusão molecular. P. marca a fração insolúvel localizada no *pellet*, FT indica o *flow through* coletado. Os números em mM são referentes à concentração de imidazol na qual as amostras foram eluídas da coluna. O p1, p2, p3 e p4 representa frações de um mesmo pico que foi eluído. Purif. Refere-se a amostra purificada por cromatografia líquida de afinidade a níquel antes de ser aplicada na coluna de exclusão molecular.

Nos géis SDS-PAGE 12% das purificações foi observado que bandas mais fracas se localizavam acima da banda da proteína heteróloga, sendo assim essas bandas foram cortadas e levadas para análise por MALDI para confirmação da identidade e pureza da proteína. Todas as quatro bandas analisadas, em todas elas somente peptídeos referentes à proteína nsP2 de CHIK foram identificados.

Entretanto, a mesma metodologia foi testada outras vezes mas não foram obtidos os mesmos resultados de purificação, sendo assim, outra condição foi usada, na qual variou-se o tampão de lise e ligação, a nova composição usada continha 20mM Tris-HCl, 200mM NaCl, 200mM arginina e 200mM ácido glutâmico, 5% glicerol e pH 7,1, o qual se baseou em trabalhos previamente publicados tanto de proteases de chikungunya quanto de proteínas homólogas (HU et al., 2016; SAISAWANG et al., 2015a). Os resultados da expressão e purificação podem ser observados na figura 21.

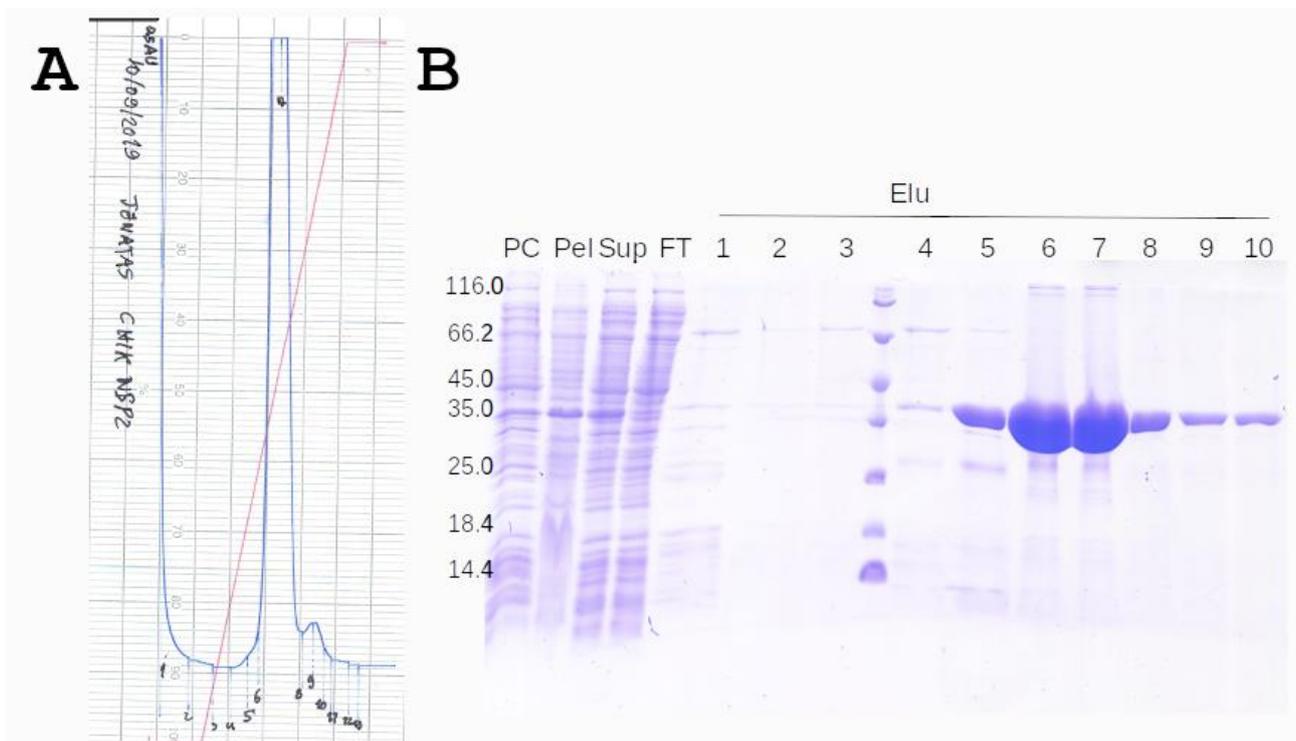


Figura 21- Expressão e purificação da proteína nsP2 de CHIK. A) Perfil cromatográfico da purificação por gradiente de imidazol em coluna de níquel. O limite de detecção é de 0,5 AU, a linha Azul se refere à absorbância (AU) e a linha vermelha a concentração de imidazol, de 0 a 500 mM. B) Gel de poliacrilamida 15% apresentando a expressão e purificação, onde PC corresponde às células antes da indução com IPTG, Pel e Sup ao *pellet* e sobrenadante após a expressão e FT à fração que não interagiu com a coluna de níquel. A numeração 1 a 10 se refere às frações eluídas durante o gradiente de imidazol da purificação.

Como a proteína apresentava um bom grau de pureza seguiu-se para a etapa da remoção da cauda de histidina utilizando a protease HRV3C, a remoção foi então analisada em gel de agarose 15% (figura 22A). Visando realizar a separação da nsP2 de CHIK, seguiu-se de uma purificação em coluna de GST, para retirada da HRV3C, e de níquel acopladas. Devido à permanência de uma pequena banda referente a protease HRV3C a etapa da aplicação da amostra em coluna de GST teve de ser repetida. Os resultados destas duas últimas etapas podem ser vistos na figura 22B e C.

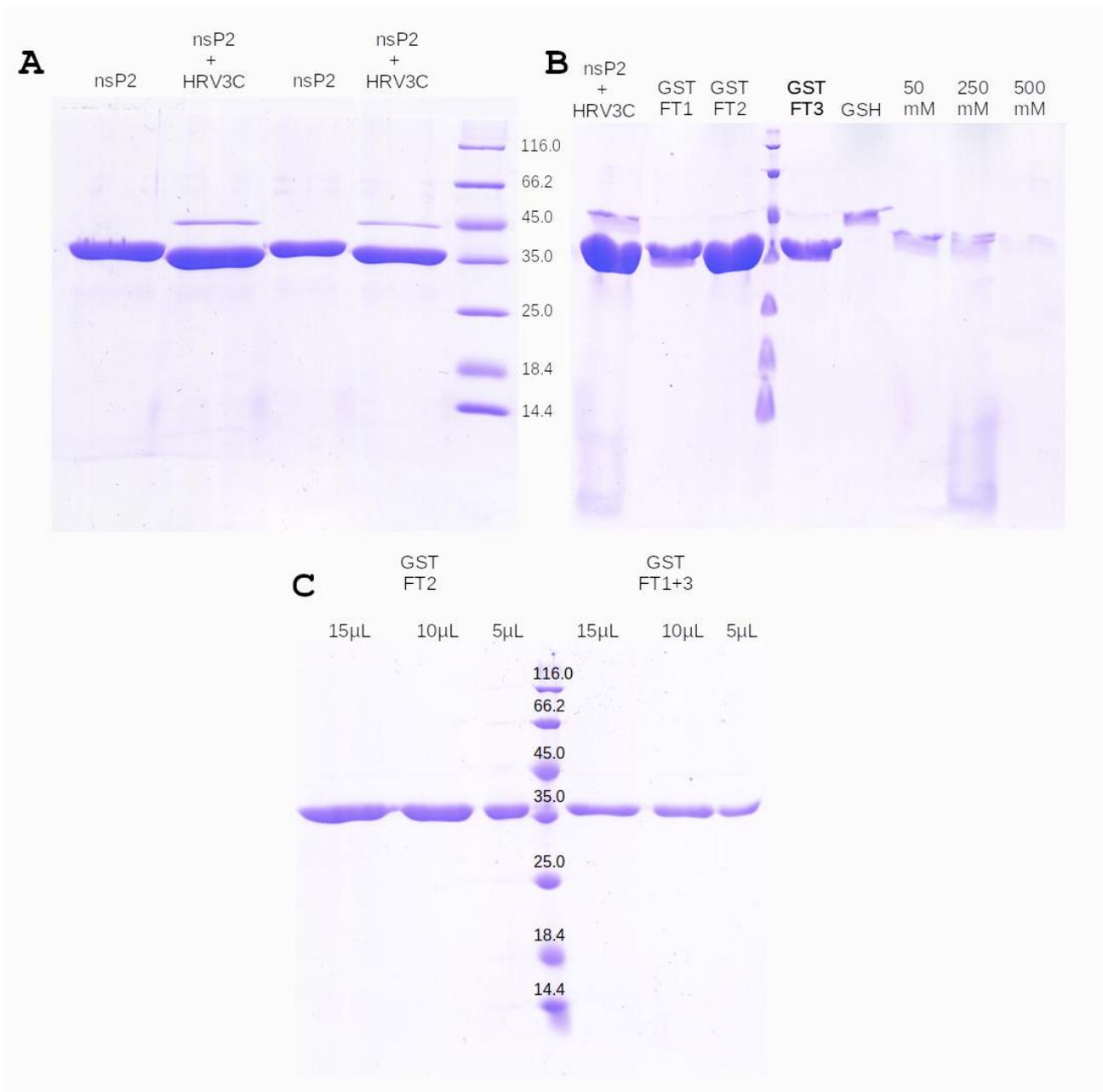


Figura 22- Remoção da cauda de histidina da nsP2 de CHIK. A) Gel SDS-PAGE 15% referente à etapa de incubação com a protease HRV3C. B) Separação da HRV3C por coluna de GST e de nsP2 com cauda de histidina usando coluna de níquel. GST FT 1 a 3 se referem às proteínas que não interagiram com a coluna de GST, GSH se refere à eluição das proteínas que interagiram com a coluna de GST, e os poços 50, 250 e 500 mM se referem a três lavagens com diferentes concentrações de imidazol da coluna de níquel acoplada. C) Reaplicação da nsP2 em coluna de GSH em diferentes volumes (15, 10 e 5 µL) aplicados no gel após purificação em coluna de GST de duas diferentes frações.

Com a purificação da proteína otimizada, seguiu-se para a produção do substrato da mesma objetivando a análise de atividade da protease. O substrato composto por uma tioredoxina seguido do sítio de clivagem entre as proteínas nsP3/nsP4 da protease de CHIK foi purificado inicialmente por cromatografia líquida em coluna de níquel, figura 22A e B. A etapa seguinte foi de refinamento da purificação, na qual foi utilizada uma coluna de troca aniônica (HiTrap Q HP), figura 23C e D.

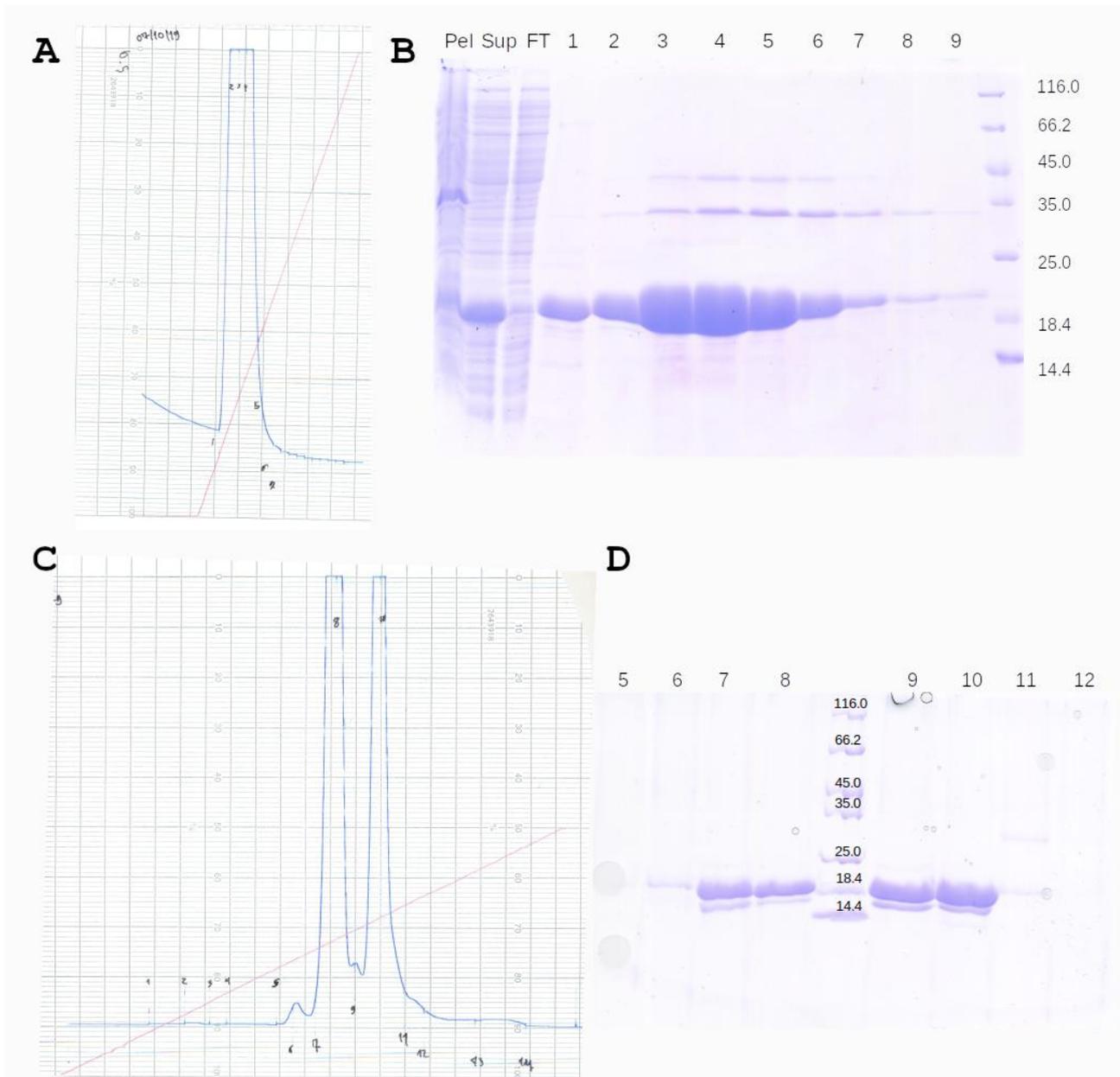


Figura 23- Expressão do substrato da protease nsP2 de CHIK. A) Perfil cromatográfico da purificação por gradiente de imidazol usando-se coluna de níquel. O limite de detecção é de 0,5 AU, a linha azul refere-se à absorvância da amostra (AU) e a vermelha à concentração de imidazol (0 a 500 mM). B) SDS-PAGE 15% da purificação por coluna de

níquel. Pel e Sup referem-se ao *pellet* e ao sobrenadante obtidos durante as etapas de lise e centrifugação, FT às proteínas que não interagiram com a coluna de níquel e os números de 1 a 9 às amostras eluídas em diferentes concentrações de imidazol. C) Perfil cromatográfico da purificação por troca aniônica (HiTrap Q HP). D) SDS-PAGE 15% das amostras purificadas na etapa de purificação pela coluna de troca aniônica usando gradiente de NaCl (0 a 500 mM). Os números de 5 a 12 referem-se as amostras eluídas nas diferentes concentrações de NaCl.

Realizou-se a etapa da análise de atividade da protease nsP2 de CHIK anteriormente purificada, onde variou-se a massa da protease mantendo-se fixa a do substrato na presença e ausência de DTT. Observou-se que a protease em questão apresenta atividade contra o substrato, este que apresenta um dos sítios de clivagem da nsP2 de CHIK, entre as proteínas nsP3 e nsP4. Este resultado indica que a proteína apresenta o dobramento correto, o qual permite a clivagem da cadeia polipeptídica em sítio específico. O gel contendo o resultado do ensaio pode ser observado na figura 24.

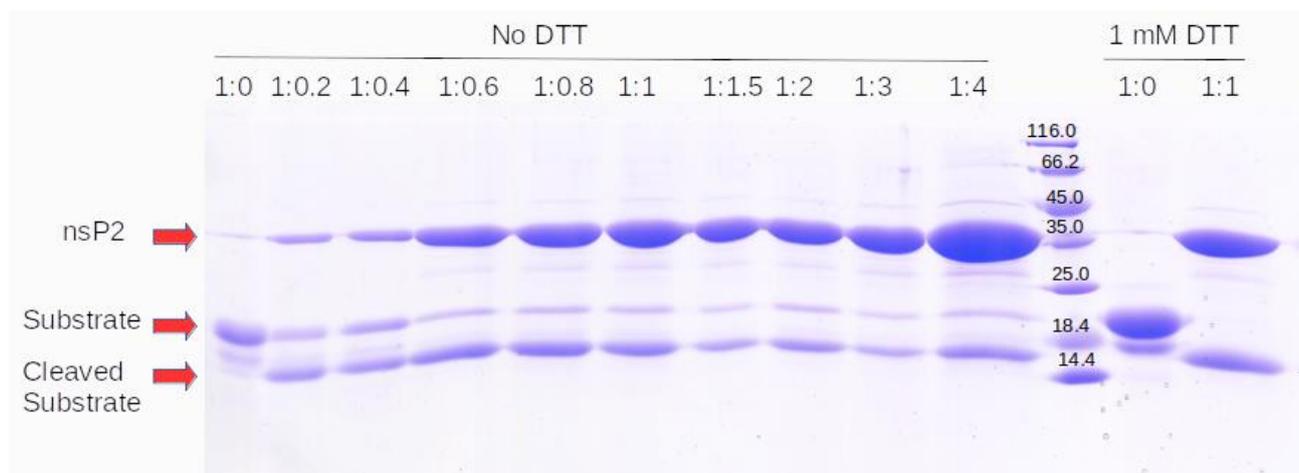


Figura 24- SDS-PAGE 15% mostrando o ensaio de atividade da nsP2 de CHIK. Variação de massa de protease de 0 a 4 μ g com valores fixos de substrato. Os poços marcados com “No DTT” não apresentavam DTT, todavia os dois poços finais apresentam uma concentração de 1 mM de DTT.

Diferentes tipos de testes de atividade já foram utilizados para diversas proteases de Alphavirus, alguns dos quais não apresentaram bons resultados. Em se tratando de substratos peptídicos para CHIK, não foi possível observar a atividade proteolítica da proteína quando substratos eram tripeptídeos (SAISAWANG et al., 2015a). Resultados

semelhantes foram obtidos para SFV, onde construções de substratos com menos que 10 aminoácidos do sítio de clivagem não eram digeridas pela protease. Este trabalho apresentou construções semelhantes a usadas no presente trabalho onde o sítio de clivagem foi acoplado a uma tiorredoxina (MERITS et al., 2001). Além disso, a mesma ideia foi utilizada ao acoplar com a proteína de ligação a maltose (MBP) para as proteases de SFV, vírus da encefalite equina venezuelana (VEEV) e vírus Sindbis (SIN) (ZHANG; TÖZSÉR; WAUGH, 2009). Outros substratos chamados de *FRET substrates* foram construídos com 17 e 25 resíduos correspondendo a região dos sítios de clivagem destas proteases flanqueados por proteínas fluorescentes ciano (CFP) e amarela (YFP) (HU et al., 2016). Estes resultados parecem indicar a grande importância para a atividade desta endo-protease de regiões mais distantes do sítio de clivagem, sendo necessário estudos dos aminoácidos que compõem estes locais e de suas interações com a protease.

Pode ser observado que a protease apresenta melhor atividade na presença de ditioneitol (DTT), pois não é observada mais a banda referente ao substrato enquanto esta é parcialmente encontrada na ausência desta molécula. O DTT é amplamente utilizado com a função de reduzir pontes dissulfeto, auxiliando assim na manutenção da estabilidade e prevenir agregação da proteína (ALLIEGRO, 2000; ULLAH et al., 2016). Outro fator que pode auxiliar na explicação desse resultado é o fato de cisteína proteases poderem ser inativadas durante seu armazenamento devido a oxidação do seu grupo tiol por moléculas como peróxido de hidrogênio (H_2O_2), dióxido de nitrogênio (NO_2), hidroperóxidos orgânicos, ou por alguns radicais como carbonato, peroxil, fenoxil, entre outras moléculas. Sendo assim, a adição de agentes redutores visando uma melhora na atividade dessas proteases parece se mostrar viável, o que é corroborado pelos resultados de nosso trabalho (KLEIN; ANSTEE, 2014; TURELL; ZEIDA; TRUJILLO, 2020).

Outra molécula presente no tampão, todavia presente em todas as amostras analisadas no teste de atividade, era o EDTA. Ele atua como um quelante com alta afinidade para realizar complexos com íons metálicos (OVIEDO; RODRÍGUEZ, 2003). Metais pesados e seus derivados apresentam capacidade de inibir a atividade de tiol-proteases (KLEIN; ANSTEE, 2014; NOLLET et al., 2012). De forma mais específica, testes com a protease de CHIK na presença de alguns íons metálicos já foram realizados. Pastorini e colaboradores mostraram que a atividade da protease foi inibida na presença de 2mM de Zn e Cu, além da resistência a inibição por EDTA. Todavia, Saisawang e colaboradores observaram a atividade da fração protease e da proteína completa, para os

três deferentes sítios de clivagem (nsP1/2, nsP2/3 e nsP3/4) na presença de Co, Cu, Mg, Ni e Zn. Para a grande maioria dos sítios não foi observada diferença entre as duas construções, somente para Zn é que pareceu aumentar a atividade da nsP2 completa e reduzir a atividade somente da fração protease para o sítio nsP3/4, e Co que demonstrou um incremento na atividade de clivagem do sítio nsP1/2 e uma redução no nsP2/3, entretanto não foi observada a inativação da protease (PASTORINO et al., 2008; SAISAWANG et al., 2015a). São necessários mais testes na presença desses íons para verificar a capacidade de inibição dos mesmos para proteases de CHIK e para outros Alphavirus. Devido a esta possibilidade de inibição, a presença de EDTA no experimento visou uma atividade ótima da protease.

O modelo de teste de atividade usado no presente trabalho, apesar de demonstrar a atividade da protease heteróloga, não permite comparações em termos de eficiência de clivagem com resultados anteriormente publicados devido à impossibilidade de calcular os parâmetros cinéticos enzimáticos.

5.1.2. Experimentos *in silico*

Realizou-se uma modelagem por homologia da proteína nsP2 de MAYV, a única dentre as proteases em questão sem nenhuma estrutura depositada no banco de dados PDB. Para este fim selecionaram-se duas estruturas depositadas no PDB, uma de CHIK (3TRK) e outra do VEEV (2HWK). Além das estruturas, para a confecção do modelo, utilizou-se alinhamento múltiplo o qual está presente em anexo.

Vários foram os critérios para seleção do modelo e análise da qualidade do mesmo. Primeiramente utilizaram-se os scores presentes no próprio modeller (molpdf 8613.64062 e *DOPE score* -38694.914062), sendo que o modelo selecionado obteve valores bons de ranqueamento para ambos. Além disso, em relação às análises de qualidade, tanto no diagrama de Ramachandran, 96,9% dos resíduos tiveram seus ângulos Phi (ϕ) e Psi (ψ) em regiões favoráveis, somente (0,3% em regiões desfavoráveis), o resíduo A727 localizado em uma região de *loop* do domínio Mtase distante da região do sítio ativo. Quanto ao qmean, o modelo apresentou bons resultados. Estes dados podem ser vistos na figura 25. Além disso, a figura do modelo sobreposto as estruturas cristalográficas usadas para confecção do modelo pode ser visto na figura 26.

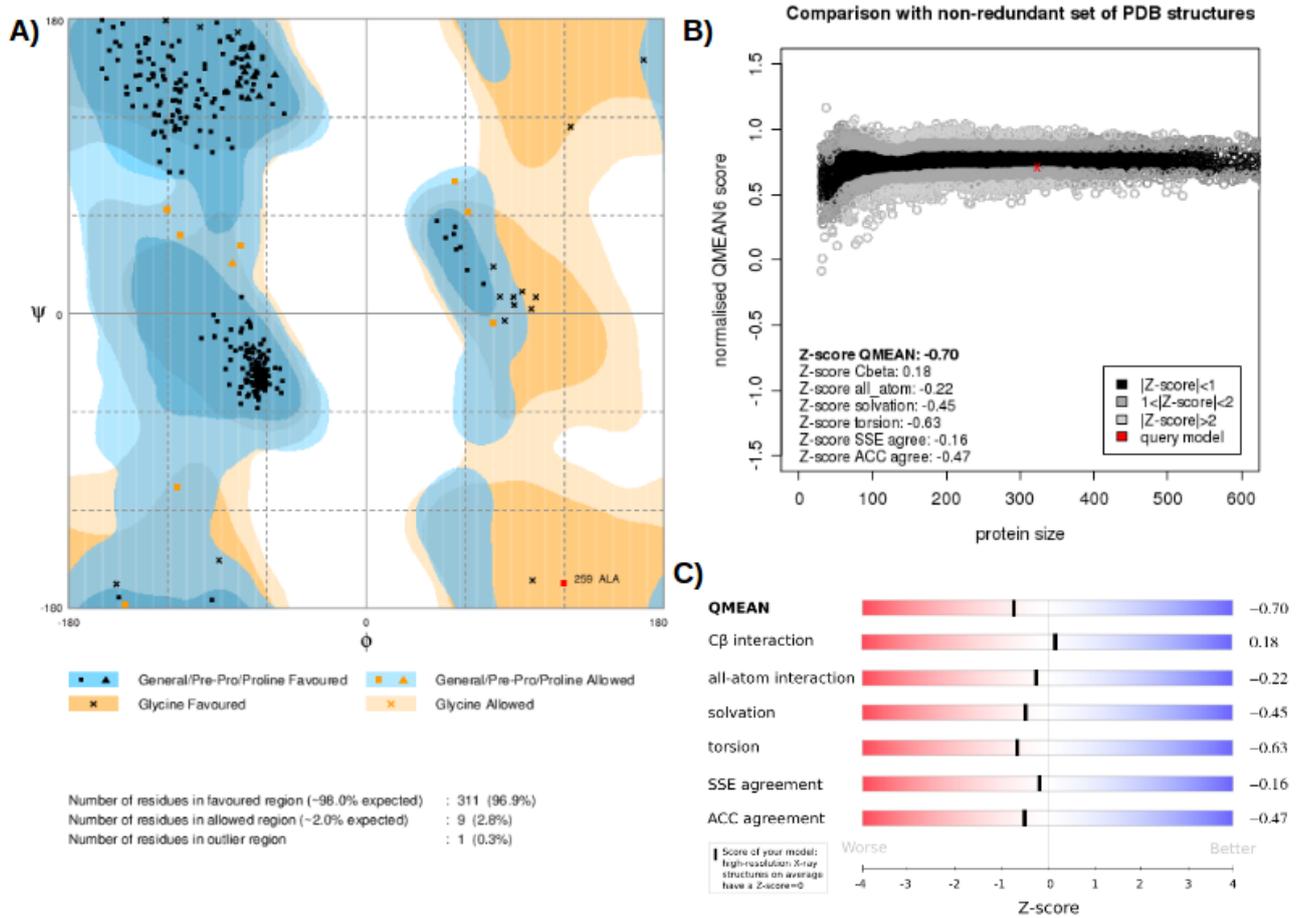


Figura 25- Análises para a qualidade do modelo selecionado da protease nsP2 de MAYV. (A) Diagrama de Ramachandran. (B e C) Valores de qualidade obtidos pelo software qmean.

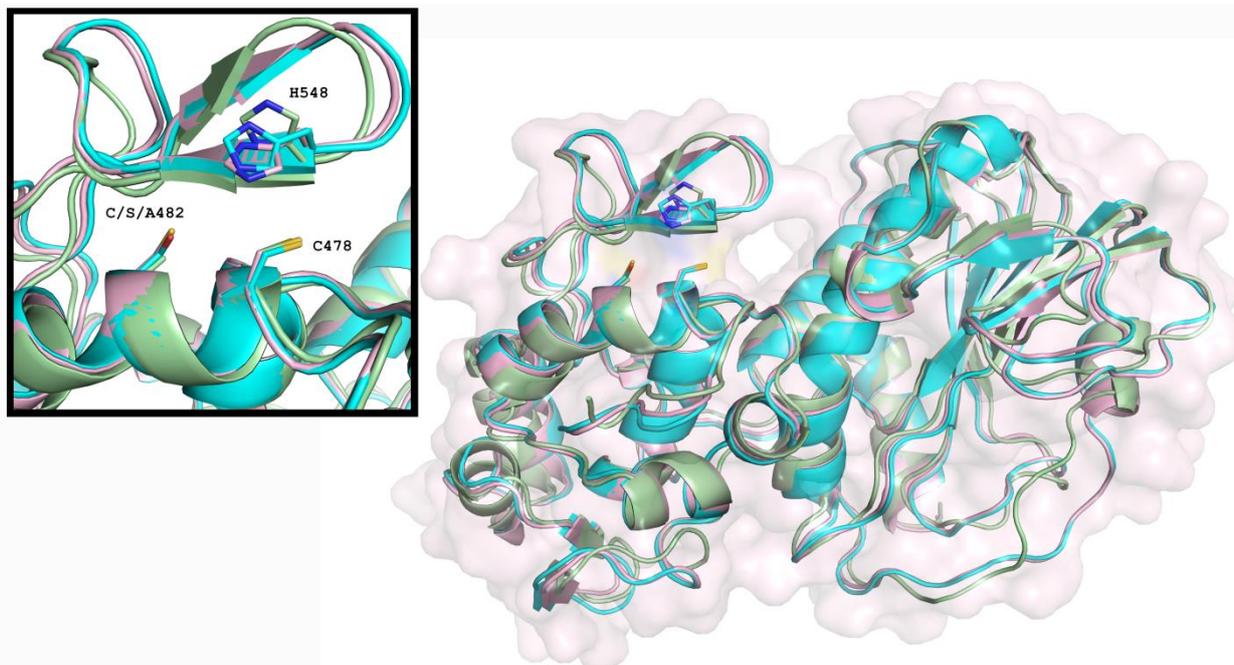


Figura 26- Sobreposição das proteases nsP2 dos Alphavirus CHIK (ciano, PDB 3TRK), MAYV (rosa, modelo por homologia) e VEEV (verde, PDB 2HWK). O quadro destaca os resíduos que compõe o sítio ativo (C478, C/S/A482 e H548). O resíduo 482 é uma C, S e A em MAYV, CHIV e VEEV, respectivamente.

O modelo da protease de MAYV ainda foi submetido a dinâmica molecular, sem nenhum ligante por um tempo de 50 ns. A proteína apresentou uma estabilização dos valores de RMSD durante a dinâmica a partir do tempo de 35 ns quando comparado as posições iniciais da MD (figura 27A). Foi analisada também a movimentação das diversas regiões da proteína através da análise de RMSF (figura 27B), onde constatou-se que, excluindo-se as extremidades que apresentam valores maiores, os resíduos entre 658 a 668, os quais representam um grande *loop* em uma região acessível ao solvente presente no domínio Mtase, apresentaram um valor de RMSF mais alto. Todavia as regiões da protease, as demais importantes para interação no sítio ativo e as estruturas secundárias que compõe esta protease se mantiveram relativamente estáveis.

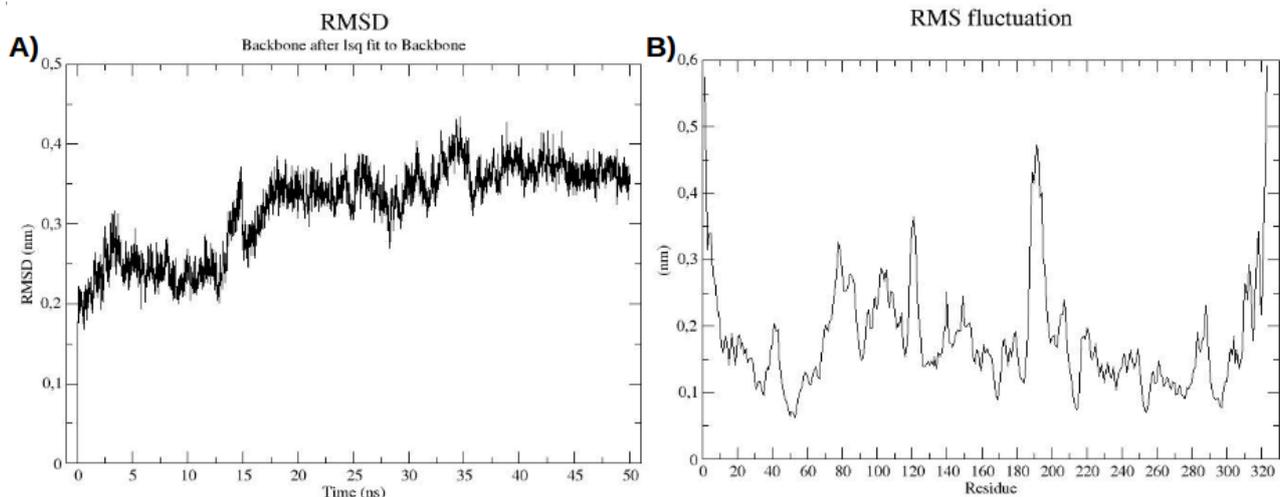


Figura 27- Gráficos de (A) RMSD e (B) RMSF do dinâmica molecular realizada com o modelo apo nsP2 de MAYV.

Realizou-se um alinhamento múltiplo entre diversos peptídeos com sítios de clivagem das proteases nsP2 de variados *Alphavirus*. Nesse alinhamento observou-se que a glicina posicionada na posição P2 dos substratos é conservada em todos os sítios (figura 27A). Além disso, a alanina na posição P3 também apresentou um alto grau de conservação. A importância dessa glicina já foi levantada em artigos anteriores, frisando a alta conservação da posição P2 do substrato em cisteino proteases e apontando que esta glicina se localiza em VEEV entre W478 e I542, o que dificultaria o posicionamento de resíduos maiores (HU et al., 2016).

O modelo da nsP2 de MAYV e a estrutura da CHIK (3TRK) foram submetidos ao experimento de *docking* contra diversos peptídeos com 10 aminoácidos (P6-P4'), os quais apresentavam os sítios de clivagem específicos das proteínas em questão (nsP1/nsP2, nsP2/nsP3 e nsP3/nsP4). Esta quantidade de aminoácidos na composição dos peptídeos foi escolhido devido a resultados experimentais anteriores onde não foi possível observar clivagem dos substratos compostos por tripeptídeos (SAISAWANG et al., 2015a) e nem em construções de substratos com menos que 10 aminoácidos do sítio de clivagem (MERITS et al., 2001). Sendo assim, seria possível observar importantes interações que poderiam ocorrer em regiões mais distantes do resíduo que sofre o ataque nucleofílico. No resultado desse experimento observou-se que poses dentre os maiores valores de *score* alcançados posicionavam a glicina P2 do substrato em uma fissura bastante estreita na superfície dessa proteína, como pode ser observado na figura 28B.

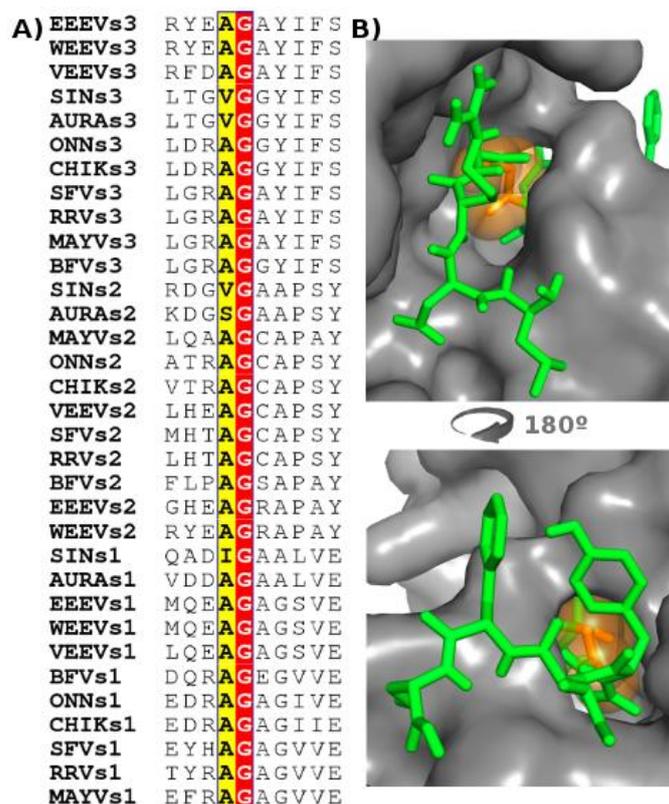


Figura 28- (A) Alinhamento múltiplo entre diversos peptídeos substratos de Alphavirus. Glicina da posição P2 em vermelho e posição P3 em amarelo, ocupada preferencialmente por uma alanina. (B) Posicionamento da glicina P2 em uma fissura na superfície da protease nsP2 de CHIK (3TRK). Peptídeo NSP1/nsP2 em verde, glicina em laranja com raio de van der Waals representado por esferas.

Um alinhamento múltiplo de proteases nsP2 dos Alphavirus foi realizado (anexo), com objetivo de observar regiões altamente conservadas. Observou-se que alguns resíduos que compõe a fissura onde a glicina P2 se localiza são conservados em todas as sequências (C478, W479, S513, H548 e W549), outros altamente conservados (A475, A511, Y512, Y544, L669/670, M702/703 e M706/707) e também por resíduos variáveis nas diferentes sequências (K476/N476, N547 e D668/E669) (figura 28). Estes resíduos que foram selecionados MAYV/CHIK estavam distantes até 5 Å dos resíduos conservados P3 e P2 dos substratos. Observa-se assim que não somente a posição os substratos se mostram conservados, mas o sítio onde esses parecem se posicionar também se apresenta altamente conservado nos Alphavirus. Sendo assim, não somente dois resíduos seriam responsáveis por acoplar a glicina P2, mas outros resíduos altamente conservados são necessários para construção de um arcabouço para o posicionamento do substrato, em especial da glicina conservada P2. Estes dados foram observados

também durante o *docking* e dinâmica molecular.

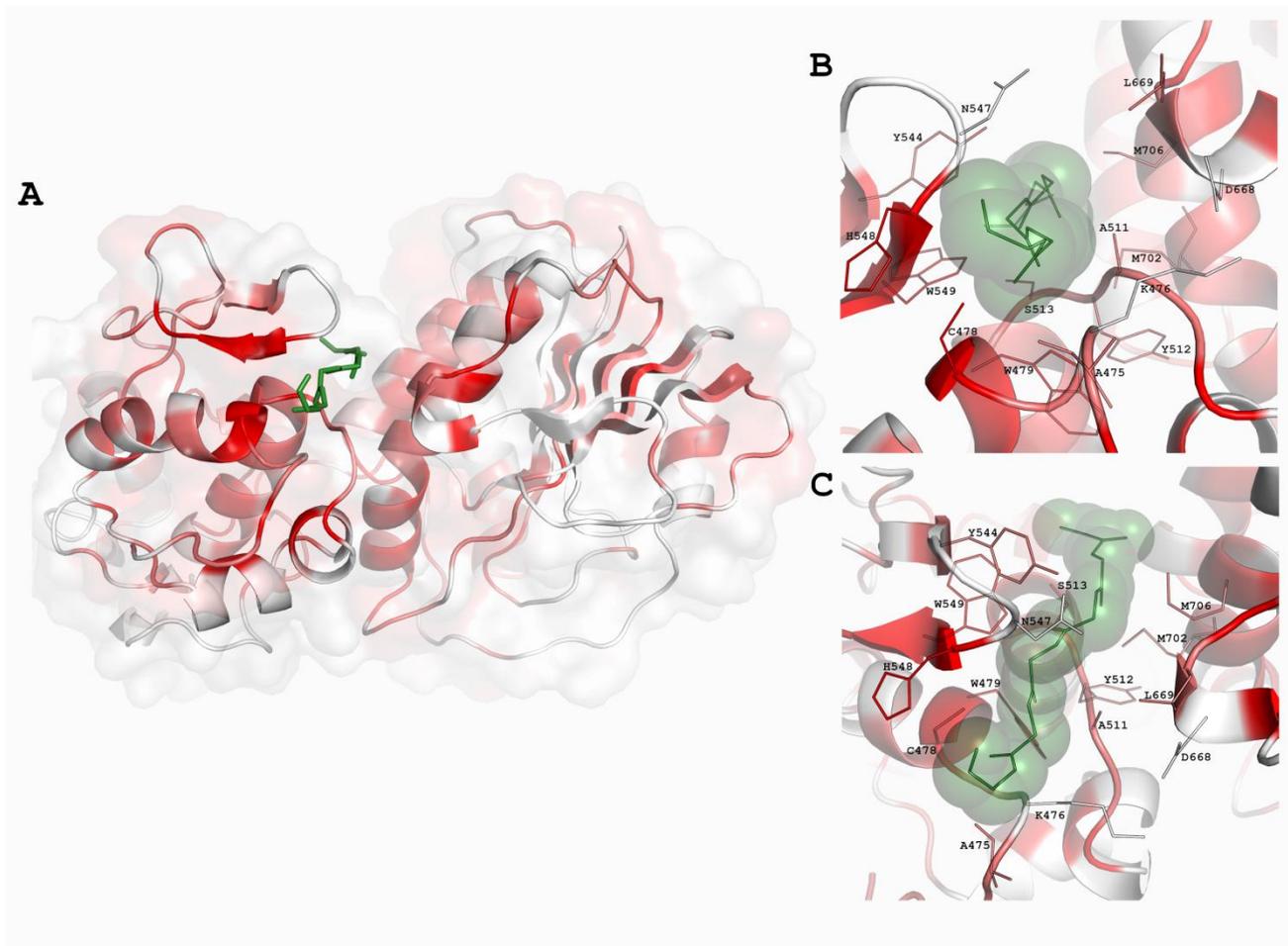


Figura 29- Modelo da protease nsP2 de MAYV colorido por níveis de conservação dos resíduos. Em vermelho os resíduos conservados em todos os Alphavirus analisados, em vermelho claro resíduos altamente conservados e em brancos resíduos variáveis no alinhamento múltiplo. Os *stick* verdes representam uma pose calculada do peptídeo contendo o sítio de clivagem nsP2/nsP3 de MAYV, representada apenas pela cadeia principal do P5-P1'. A) Modelo da região protease da nsP2 de MAYV em interação com o peptídeo. B) Resíduos que compreendem a fissura responsável por reconhecer o substrato. C) Visão da figura B após rotação de 90° em eixo paralelo a cadeia principal do peptídeo.

A lista dos resíduos que compõe cada subsítio responsável por interagir com cada resíduo do substrato para os Alphavirus CHIK, MAYV e VEEV pode ser observado na tabela 8. Esta foi confeccionada tendo como base o levantamento previamente realizado

por Narwal e colaboradores (2018). Foram acrescentados dados referentes aos resíduos do substrato P6, P5, P2', P3' e P4' de CHIK e os dados de MAYV, tanto obtidos a partir do *docking* quanto das análises de dinâmica molecular.

Uma das análises que foram utilizadas para observar possíveis resíduos importantes para o contato entre enzima e substrato foi a de tempo de existência das interações de hidrogênio durante a dinâmica molecular. As durações de interações estão aproximadamente marcadas na tabela 6 e 7. Dentre estas alguns contatos se mostraram bastante representativos durante a análise. Para CHIK observou-se a interação entre o nitrogênio da cadeia principal do resíduo catalítico C478 e o oxigênio da glicina da posição P1, este ocorreu por aproximadamente 70% do tempo analisado. Outras interações importantes foram entre o nitrogênio da cadeia lateral do W549 e o oxigênio da alanina P3, por 75% do tempo, e o oxigênio da cadeia principal da N476 e o nitrogênio da isoleucina P2', 60% do tempo. Em se tratando das análises da nsP2 de MAYV as interações com maior tempo foram, em primeiro lugar, entre o oxigênio da cadeia principal do N547 e o nitrogênio da cisteína P1 e o oxigênio da cadeia principal da Y512 e o nitrogênio da glicina P2, ambos por aproximadamente 91%. Outro contato importante foi entre o oxigênio da cadeia lateral do Q705 e o nitrogênio tanto da cadeia principal quanto lateral da glutamina P5, com tempos superiores a 80%. Por último, o contato ocorreu entre o nitrogênio da cadeia lateral do W549 e o oxigênio da alanina P4, por 76% do tempo analisado. A tabela com todos os tempos de ligação de hidrogênio entre a nsP2 de CHIK e MAYV e os respectivos substratos analisados pode ser encontrada nas tabelas 5 e 6.

Tabela 6- Tempo de existência das ligações de hidrogênio da MD de CHIK com substrato CHIKs34.

CHIKs34(LDRAGG/YIFS)				
Donor		Acceptor		% Exist.
ASN476	ND2	TYR7(P1')	OH	4.52
ASN476	ND2	ILE8(P2')	O	4.66
ASN476	ND2	SER10(P4')	OG	1.10
CYS478	N	GLY6(P1)	O	70.66
TRP479	N	GLY6(P1)	O	10.71
TYR512	N	GLY5(P2)	O	2.01
SER513	N	ARG3(P4)	O	1.35
SER513	OG	ASP2(P5)	O	3.19
TYR544	OH	ASP2(P5)	OD1	5.96
TYR544	OH	ASP2(P5)	OD2	5.31
TYR544	OH	ALA4(P3)	O	9.04
ASN547	ND2	TYR7(P1')	O	6.80
ASN547	ND2	PHE9(P3')	O	1.29
ASN547	ND2	SER10(P4')	OG	1.29
ASN547	ND2	SER10(P4')	O1	2.00
ASN547	ND2	SER10(P4')	O2	3.75
TRP549	NE1	ARG3(P4)	O	1.07
TRP549	NE1	ALA4(P3)	O	75.49
TRP549	NE1	GLY5(P2)	O	6.87
ASN667	ND2	TYR7(P1')	OH	3.80
LEU1(P6)	N	GLU515	OE1	13.69
LEU1(P6)	N	GLU515	OE2	13.10
ARG3(P4)	NE	GLN706	OE1	6.29
ARG3(P4)	NH1	TYR512	O	2.87
ARG3(P4)	NH1	GLN706	OE1	1.88
ARG3(P4)	NH1	THR733	OG1	2.65
ARG3(P4)	NH2	GLU515	OE2	1.00
ARG3(P4)	NH2	GLN706	OE1	3.49
ARG3(P4)	NH2	THR733	OG1	5.35
GLY5(P2)	N	GLU669	O	15.11
GLY6(P1)	N	ASN547	OD1	19.69
GLY6(P1)	N	ASN547	O	4.02
TYR7(P1')	N	ASN547	OD1	14.44
TYR7(P1')	N	ASN547	O	8.37
TYR7(P1')	OH	ASP509	OD1	4.93
TYR7(P1')	OH	ASP509	OD2	2.25
TYR7(P1')	OH	ASN667	OD1	1.48
TYR7(P1')	OH	GLU669	OE2	2.33
ILE8(P2')	N	ASN476	O	59.38
PHE9(P3')	N	HIS548	ND1	42.49
SER10(P4')	N	HIS548	ND1	5.78

Obs. Em rosa resíduos com mais de 50% do tempo, amarelo entre 10% e 49%, e em azul entre 1% e 9%.

Tabela 7- Tempo de existência das ligações de hidrogênio da MD de MAYV com substrato MAYVs23.

MAYVs23(LQAAGC/APAY)				
Donor		Acceptor		% Exist.
LYS476	NZ	ALA7(P1')	O	5.58
TYR512	N	GLY5(P2)	O	64.54
SER513	OG	GLN2(P5)	OE1	23.43
SER513	OG	ALA3(P4)	O	6.84
TYR544	OH	ALA4(P3)	O	22.44
ASN547	ND2	ALA4(P3)	O	3.24
ASN547	ND2	CYS6(P1)	O	4.03
TRP549	NE1	ALA3(P4)	O	76.42
TRP549	NE1	ALA4(P3)	O	8.03
ARG556	NH1	LEU1(P6)	O	1.68
ARG556	NH2	LEU1(P6)	O	1.33
THR657	N	TYR10(P4')	O1	1.55
THR664	OG1	TYR10(P4')	O1	1.70
GLN705	NE2	GLN2(P5)	OE1	71.25
GLN705	NE2	GLN2(P5)	O	9.76
LEU1(P6)	N	GLN705	O	1.39
GLN2(P5)	N	GLN705	OE1	83.48
GLN2(P5)	NE2	GLU515	OE1	31.44
GLN2(P5)	NE2	GLU515	OE2	44.27
GLN2(P5)	NE2	GLN705	OE1	6.19
ALA3(P4)	N	GLU515	OE2	1.22
GLY5(P2)	N	TYR512	O	91.13
CYS6(P1)	N	ASN547	O	91.57
ALA7(P1')	N	LYS476	O	49.30
ALA7(P1')	N	ASN547	O	1.44
ALA9(P3')	N	ALA475	O	4.50

Obs. Em vermelho resíduos acima de 90%, em rosa resíduos entre 89% e 50% do tempo, amarelo entre 10% e 49%, e em azul entre 1% e 9%.

Alguns dos resíduos citados anteriormente têm sido apontados na literatura como importantes para o contato com o substrato: o Y512 e o W549. O primeiro já havia sido determinado como importante na estabilização da glicina P2 conservada e o segundo um resíduo volumoso e conservado nos Alphavirus, este tem se mostrado importante na definição do motivo de especificidade a glicina (GSM) (RUSSO et al., 2010; RUSSO; WHITE; WATOWICH, 2006). O resíduo N547 também se mostrou relevante no contato com o substrato, já que uma mutação do mesmo para alanina demonstrou uma redução da atividade catalítica da enzima. Este já havia sido sugerido como um importante para o contato com o substrato devido apresentar interação com o resíduo da posição P1 (RUSSO et al., 2010). O N547 se posiciona em um importante *loop* flexível entre as fitas β 1 e β 2 compostas pelos resíduos 544 a 549, dentre os quais também possui outros resíduos de grande importância como o H548, o qual compõe a díade catalítica, e o W549. Além disso, ele cobre a região do sítio ativo atuando como uma possível “porta” para o acesso do substrato, apresentando valores de *B factor* mais elevados que as demais regiões do GSM possuindo também variações na posição quando comparada as estruturas de Alphavirus depositadas, demonstrando assim ser uma região de maior

flexibilidade. Sendo assim, este tem se mostrado crucial no auxílio do posicionamento e na interação com o substrato (NARWAL et al., 2018). Os resíduos que interagiram com os sítios P5 e P2' do substratos não haviam sido ainda reportados na literatura já que as análises *in silico* anteriores não abrangiam estas regiões, estes contatos podem ajudar a mostrar a importância de resíduos mais distantes do sítio de ataque nucleofílico para a estabilização do substrato.

Tabela 8- Resíduos que compõe os subsítios S4' - S6 para as proteases nsP2 de Alphavirus VEEV, CHIK e MAYV.

Resíduo do substrato CHIKV NsP2pro	Resíduo do substrato VEEV NsP2pro	Resíduo do substrato MAYV NsP2pro	Subsítio	CHIKV nsP2pro Resíduos	VEEV nsP2pro Resíduos	MAYV nsP2pro Resíduos	Subdomínio
P4'-Ser	-	P4'-Tyr	S4'	Asn 476	-		Protease
						Thr 657	MTase-like
						Thr 664	MTase-like
				Asn 547			Protease
P3'-Phe	-	P3'-Ala	S3'	His 548	-	Ala 475	Protease
				Asn 547			Protease
				Asn 476			Protease
P2'-Ile	-	P1'-Pro	S2'		-		Protease
P1'-Tyr	P1'-Tyr	P1'-Ala	S1'	Ala 475	Ala 474	Ala 475	Protease
				Asn 476	Asn 475	Lys 476	Protease
				Lys 481	Lys 480		Protease
				Cys 478	Cys 477	Cys 478	Protease
				His 548	His 546	His 548	Protease
				Asn 547*		Asn 547	Protease
				Asn 667*			MTase-like
				Asp 509*			Protease
Glu 669*			MTase-like				
P1-Gly	P1-Ala	P1-Cys	S1	Asn 476	Asn 475	Lys 476	Protease
				Cys 478	Cys 477	Cys 478	Protease
				Trp 479	Trp478	Trp 479	Protease
				His 548	His 546	His 548	Protease
				Asn 547	Asn 545	Asn 547	Protease
					Leu 665	Leu 669	MTase-like
P2-Gly	P2-Gly	P2-Gly	S2	Tyr 512	Cys 477	Tyr 512	Protease
				Trp549	Trp 547	Trp549	Protease
				Trp 479	Trp 478	Trp 479	Protease
					Ala 509		Protease
			MTase-like				
P3-Ala	P3-Ala	P3-Ala	S3	Tyr 544	Trp547	Tyr 544	Protease
				Met 707	Met 702	Met 706	MTase-like
				Asn 547	Ile698	Asn 547	MTase-like
				Trp 549*		Trp 549	Protease
				Asn 547	Ser 513	Ser 513	Protease
P4-Arg	P4-Asp	P4-Ala	S4	Trp 549	Trp 547	Trp 549	Protease
				Gln 706	Met 702		MTase-like
				Asp 711	Lys 706		MTase-like
				Ser 513*			Protease
				Glu 515*		Glu 515	Protease
				Tyr 512*			Protease
				Thr 733*			MTase-like
P5-Asp	-	P5-Gln	S5	Ser 513	-	Ser 513	Protease
						Gln 705	MTase-like
				Trp 544		Glu 515	Protease
			Protease				
P6-Leu	-	P6-Leu	S6		-	Arg 556	Protease
				Glu 515		Gln 705	MTase-like
							Protease

Obs1- As marcações com * representam resíduos de CHIK que não estavam incluídos na publicação base para a tabela. (NARWAL et al., 2018)

Obs2- As cores marrons e vermelhas representam tempos calculados de tempo de existência de ligações de hidrogênio, onde a primeira representa interações que ocorreram por menos que 10% do tempo simulado e a segunda porcentagens superiores a esta.

A região de contato entre a protease nsP2 e os seus respectivos substratos têm sido estudada usando-se ferramentas *in silico* para o Alphavirus VEEV com pentapeptídeos e por outros ensaios *in vitro* para outros Alphavirus como CHIK, VEEV, SIN e SFV como anteriormente discutido. Em nossas análises observamos a interação de decapeptídeos para CHIK e MAYV. Tem-se observado que a alta conservação dos

resíduos P3 e P2, alanina e glicina respectivamente, apontam para um sítio de interação estreito na protease no qual resíduos mais volumosos resultariam em choques estéricos (LULLA et al., 2006; RUSSO et al., 2010; RUSSO; WHITE; WATOWICH, 2006). Este sítio de interação tem sido apontado para a mesma região calculada no *docking* do presente trabalho, a qual se localiza em uma fissura estreita entre os domínios protease e MTase, este tem sido denominado motivo de especificidade a glicina (GSM) (GOLUBTSOV; KÄÄRIÄINEN; CALDENTEY, 2006; HU et al., 2016). Sendo assim, com base na literatura atual a explicação para a conservação dessa glicina em P2 e da alanina em P3 baseia-se no tamanho dos resíduos para o encaixe no sítio.

Outra possibilidade para a explicação da alta conservação da glicina seria a necessidade dos ângulos ϕ e ψ do resíduo se manter em conformações não permitidas para outros resíduos. Para analisar essa conjectura realizou-se experimentos de dinâmica molecular com a glicina dos substratos posicionados na fissura, tanto para a estrutura cristalográfica de CHIK (3TRK), quanto para o modelo da nsP2 de MAYV. Com esses dados foram gerados gráficos observando os valores desses ângulos diedros durante as dinâmicas para os resíduos P3, P2, P1, P1' e P2' do substrato (figura 30). Observa-se também a figura 31, onde está representado o gráfico de Ramachandran tanto 2D quanto 3D da glicina P2 na dinâmica molecular.

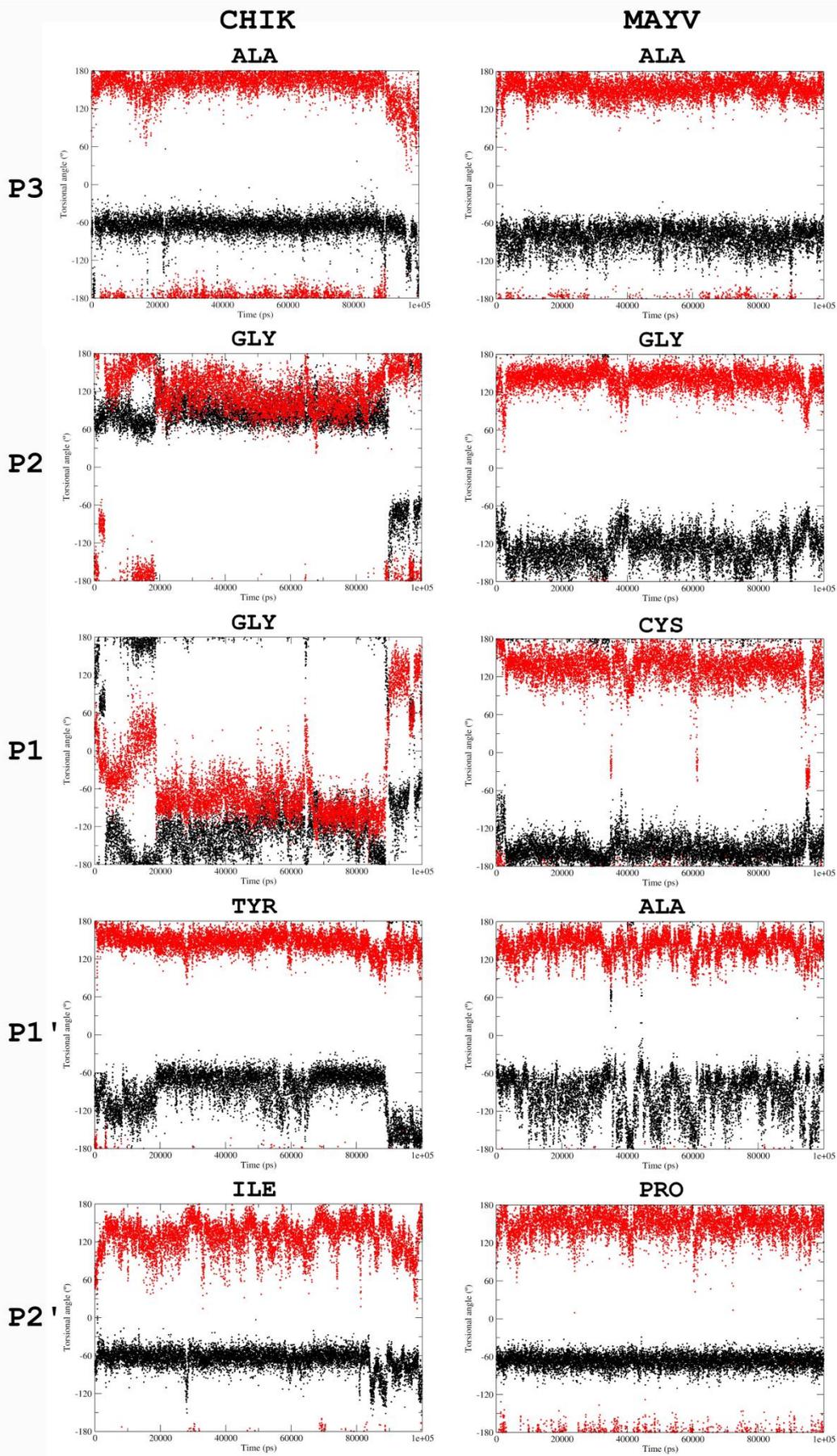


Figura 30- Valores dos ângulos ϕ e ψ dos resíduos P3, P2, P1, P1' e P2' dos substratos peptídicos de CHIK e MAYV durante as análises de MD. As linhas em preto referem-se aos valores de ϕ e as em vermelho a de ψ .

Através da figura 30 é possível ver que em geral o substrato interagindo com a protease de MAYV pareceu se mostrar mais estável no decorrer do tempo. Observou-se também que os resíduos P3 e P2' mantiveram-se mais estáveis durante o decorrer do tempo tanto para CHIK quanto para MAYV. Entretanto, os resíduos P2 e P1 de CHIK apresentaram variações nos valores dos ângulos diedrais analisados, enquanto este mesmo efeito não foi observado para no caso de MAYV. O resíduo P1' de ambos apresentaram comportamentos semelhantes, onde o ϕ se posicionou entre os valores de -60° e -180° , mas o ψ apresentou maior estabilidade.

Valores dos ângulos ϕ e ψ da glicina correspondente as conformações não permitidas para outros resíduos são usualmente apresentados em *loops* ou β -*turns* e no início ou final de hélices e fitas- β possibilitando mudanças bruscas na direção da cadeia polipeptídica (LAKSHMI et al., 2014). Além disso a flexibilidade conferida por este resíduo a um *loop* catalítico é essencial para a atividade proteolítica de proteases aspárticas como a pepsina, mesmo a mutação por aminoácidos ainda pequenos como a alanina reduz significativamente a atividade enzimática (OKONIEWSKA; TANAKA; YADA, 2000). No caso do presente trabalho os ângulos da glicina P2 mantiveram-se em posições não permitidas para os demais resíduos na maior parte do tempo da dinâmica para CHIK, todavia isso não foi observado para MAYV, como pode ser apreciado nas figuras 30 e 31. Estes resultados sugerem que a posição desses ângulos em posições não permitidas pode ser importante para somente alguns dos Alphavirus como CHIK. Além disso, esta possibilidade de flexibilidade dada pela glicina P2 conservada pode ser importante para facilitar a entrada no GSM, ou ainda permitir algumas posições intermediárias do substrato durante as etapas da lise, sendo estas as etapas bastante difíceis de serem simuladas. Vale ressaltar que esta hipótese levantada não exclui a primeira, relacionada ao volume do resíduo, podendo ambas ocorrerem de modo complementar.

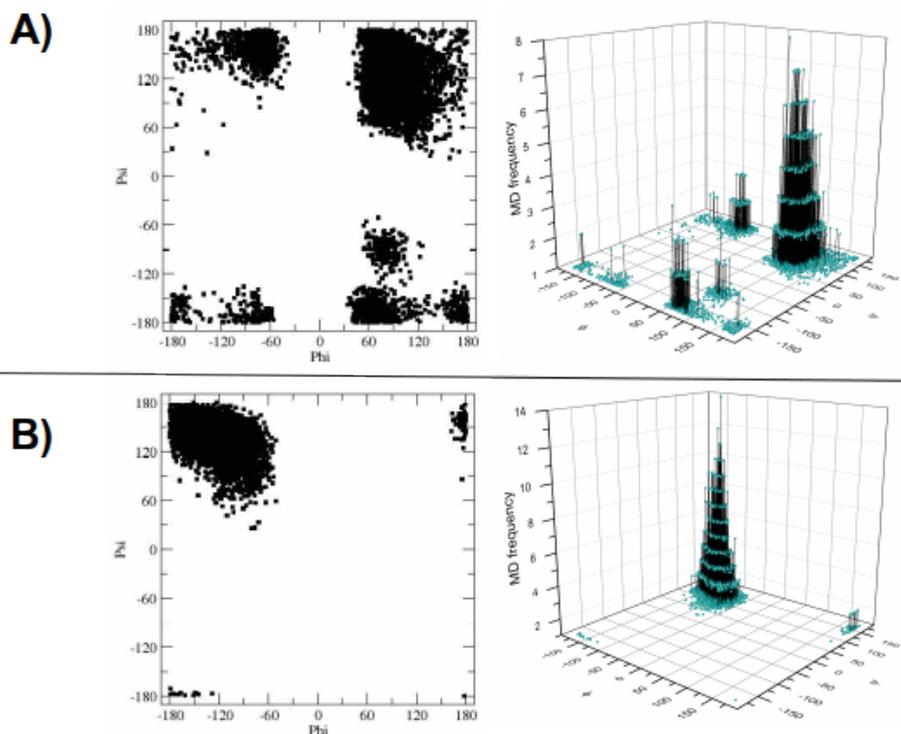


Figura 31- Gráficos de Ramachandran 2D e 3D do modelo da glicina na posição P2 do substrato peptídico da proteína nsP2 durante análise por dinâmica molecular de (A) CHIK e (B) MAYV.

Além dos ângulos ϕ e ψ , foram também analisados os valores dos ângulos χ_1 dos resíduos C478, H548 e S/C482, e χ_2 da H548. Observou-se nessas análises uma diferença entre as dinâmicas moleculares com ou sem os peptídeos/substratos, pois no primeiro caso os valores de χ_1 dos resíduos que compõe o sítio catalítico da enzima C478, H548 e χ_2 da H548 apresentaram maior estabilidade quando comparados com os das dinâmicas onde os substratos estavam ausentes. Os valores foram $\sim 60^\circ$ para χ_1 da C478, $\sim 180^\circ$ para o χ_1 da H548 e $\sim 60^\circ$ para o χ_2 da H548. Por sua vez, quando os peptídeos não estavam presentes os valores dos ângulos apresentaram uma tendência a maior variabilidade. Estes resultados podem ser observados na figura 32.

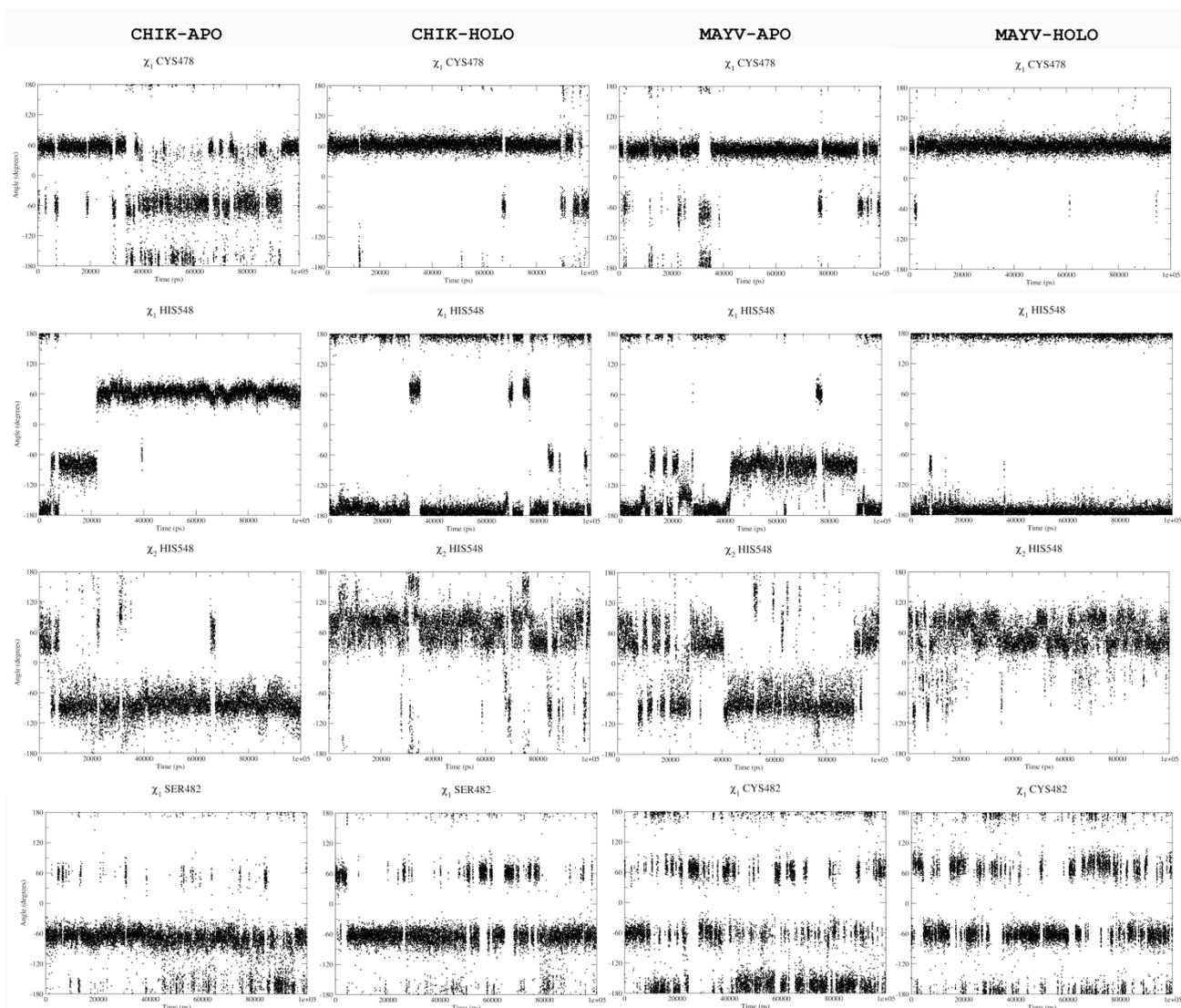


Figura 32- Valores dos ângulos χ_1 dos resíduos C478, H548 e S/C482, e χ_2 da H548 durante as análises de MD das nsP2 de CHIK e MAYV.

Sabe-se que a interação de ligantes com a proteína induzem re-arranjos na estrutura da proteína, isto pode ocorrer de um modo particular através de movimentos e re-organização das cadeias laterais (NAJMANOVICH et al., 2000). Nos resultados do presente trabalho, observou-se que o contato com substratos ordenou e estabilizou as cadeias laterais dos resíduos que compõe o sítio ativo da nsP2. O valor do ângulo χ_1 da C478 em aproximadamente 60° , a qual ficou estável na maior parte do tempo na presença dos substratos peptídicos, posiciona o enxofre da cisteína para a direção da fissura/substrato enquanto os ângulos levariam-no a apontar para outras direções (figura 32). Esta observação, juntamente com outras análises aqui realizadas como a de tempo de existência das ligações de hidrogênio, demonstram que a presença do substrato

estabiliza a cadeia lateral dessa cisteína por intermédio de contatos diretos com o substrato. Já por sua vez, os ângulos χ_1 e χ_2 da H548 na presença dos substratos levam-na a se posicionar de tal forma que o nitrogênio ND1 direciona-se ao C478, o que parece fazer sentido com o mecanismo catalítico da enzima, pois este é o nitrogênio responsável por remover o próton da cisteína para a realização do ataque nucleofílico.

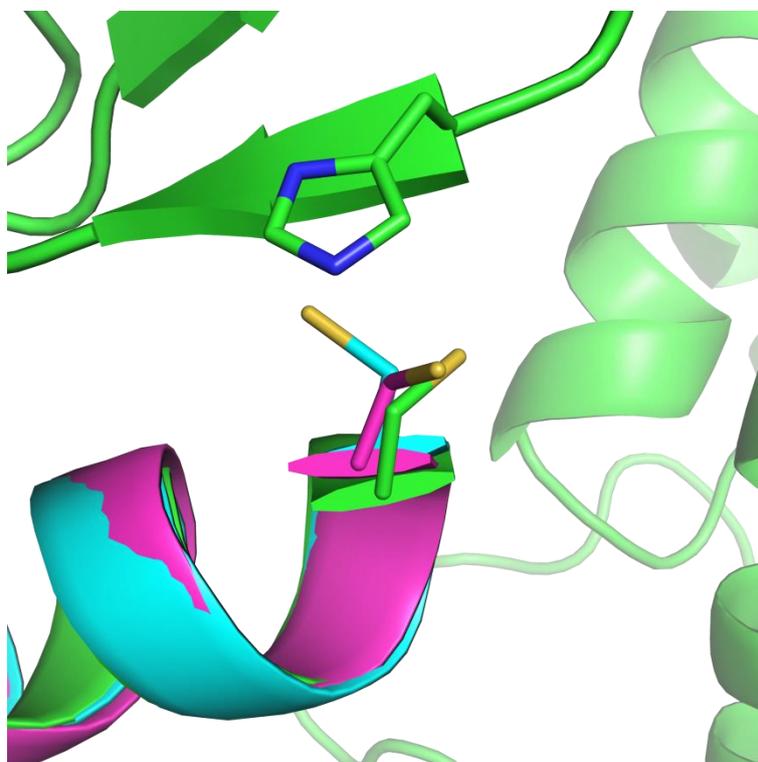


Figura 33- Posições do resíduo C478 com diferentes valores dos ângulos χ_1 . Em verde o ângulo é aproximadamente 60° , em ciano 180° e em magenta -60° .

Entretanto, não foi possível observar a mesma estabilidade no valor dos ângulos χ_1 dos resíduos S/C482 ao se comparar as dinâmicas com e sem o peptídeo contendo o sítio de clivagem dessas proteases. Tem-se discutido atualmente em relação a possível atuação deste resíduo, o qual teria a capacidade de manter a atividade proteolítica da nsP2 de CHIK na ausência da C478, apresentando assim uma díade catalítica intercambiável (SAISAWANG et al., 2015b). Esta poderia ser uma característica também da protease de MAYV, já que este apresenta uma cisteína na posição 482.

Todavia, diferentes análises posteriores não encontraram evidências da capacidade da protease de manter a atividade catalítica de clivagem dos substratos na ausência da C478, ou seja, a mutação do resíduo 478 para uma alanina inibiu completamente a atividade da enzima (NARWAL et al., 2018; RAUSALU et al., 2016).

Nossos resultados parecem corroborar a tese de que esse resíduo não possui atividade proteolítica, tendo em vista que este não apresentou nenhuma diferença de estabilidade do χ_1 na presença dos substratos, algo que ocorreu nos demais resíduos que compõe a díade catalítica. Além disso, caso seja demonstrada essa atividade, serão necessários outros modelos de interação entre proteína e o substrato pois seria preciso uma movimentação da cadeia principal deste para o contato com o resíduo 482, o que provavelmente deslocaria a glicina P2 do contato com o GSM.

Durante o experimento de dinâmica molecular das proteases nsP2 de MAYV e CHIK com o substrato foi calculada também a distância entre o carbono da carbonila do resíduo P1, que sofre o ataque nucleofílico, e o átomo de enxofre do resíduo catalítico durante todo o tempo da dinâmica, como pode ser observado na figura 34. A distância se manteve estável durante o decorrer de ambas dinâmicas, com um valor entre 3 Å e 4 Å durante a maior parte do tempo, aumentando para distâncias de 5 Å no final da dinâmica com a estrutura de CHIK.

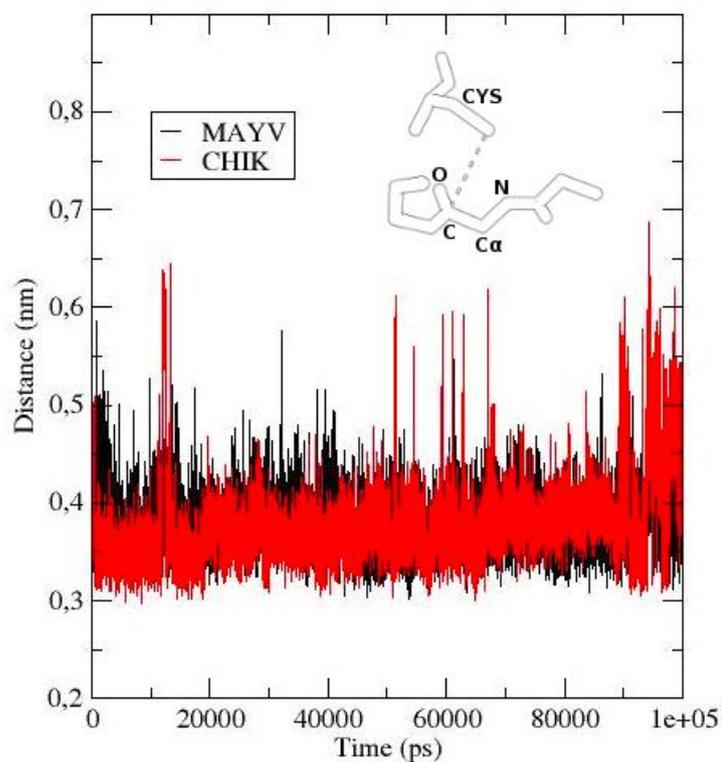


Figura 34- Distância entre o carbono da carbonila do resíduo P1 do substrato e o átomo de enxofre da cisteína do modelo da protease nsP2 de MAYV e da estrutura de CHIK.

Sendo assim, as poses selecionadas do contato entre o substrato no sítio ativo de MAYV e CHIK não somente estavam com distâncias próximas entre o enxofre da C478 e o carbono da carbonila do P1 durante o experimento de *docking*, mas esta proximidade parece se manter no decorrer do tempo.

Outro dado interessante que pode ser analisado a partir dos resultados obtidos no presente trabalho são os diferentes parâmetros cinéticos entre os substratos de CHIK, apresentando preferência pelo sítio nsP2/nsP3 (SAISAWANG et al., 2015). Os nossos resultados de *docking* mostram que a cisteína no resíduo P1 desse sítio pode auxiliar a estabilizar o substrato através de ligações de hidrogênio através da cadeia lateral com resíduos próximos como N476 e N547. Isto não poderia acontecer nos demais substratos por ser a posição P1 ocupada por resíduos apolares, alanina no nsP1/nsP2 e glicina nsP3/nsP4. Esta diferença pode ser observada na figura 35.

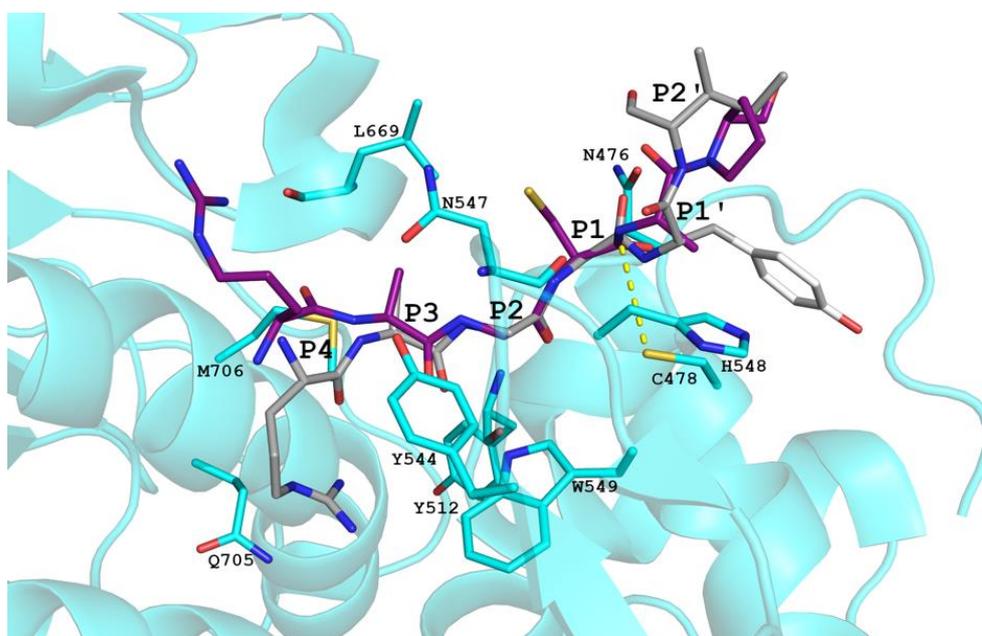


Figura 35- Interação entre a nsP2 de CHIK (ciano) (3TRK) e diferentes substratos peptídicos. nsP2/nsP3 (roxo) e nsP3/nsP4 (cinza). Os traços amarelos mostram a distância entre o átomo de enxofre da cisteína 478 e o carbono da carbonila que sofre o ataque nucleofílico.

5.2. FLAVIVIRUS

5.2.1. Zika

Em se tratando da purificação da proteína NS3 de ZIKV (ZIKV-NS3-Bra), a protease foi eluída com 250 mM de imidazol, após lavagens com 50 mM e 100 mM. A amostra foi estocada na 4°C e após alguns dias submetida a exclusão molecular. Observou-se que a banda da proteína, que tem uma massa de aproximadamente 26 kDa, migrou a ~19 kDa, o que pode indicar que a protease em questão está se auto-clivando. A diferença entre o tamanho da banda original e após alguns dias de estoque, ~7 kDa é muito próxima ao tamanho da região proteína NS2b ligada a NS3 por um *linker* na construção obtida, o que indica que a clivagem pode estar ocorrendo no local do *linker*. Estes dados podem ser observados na figura 36.

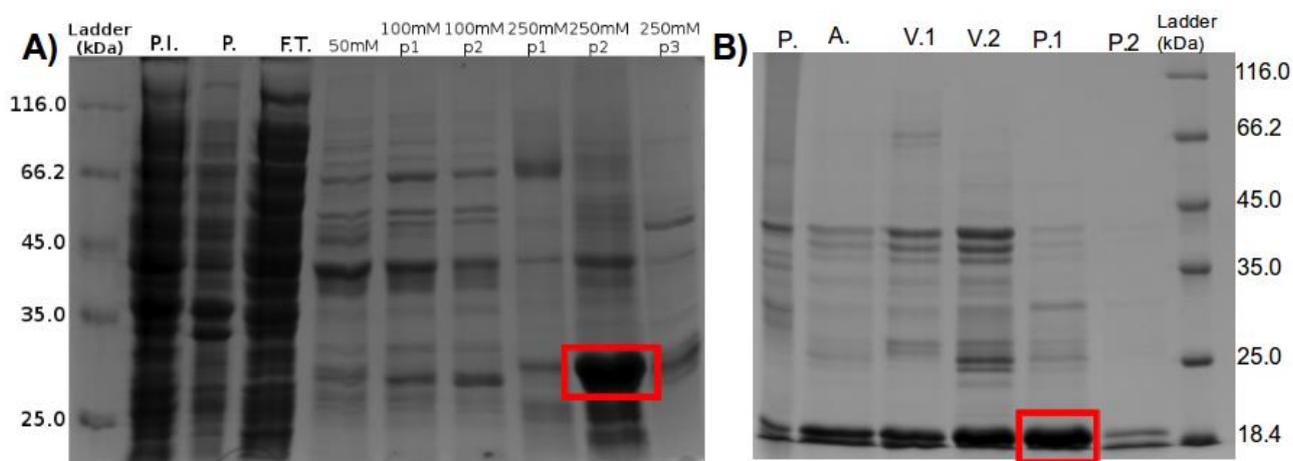


Figura 36- Purificação da protease NS3 de ZIKV. (A) Purificação por cromatografia líquida de afinidade a níquel. (B) Purificação por exclusão molecular. P. marca a fração insolúvel localizada no pellet, FT indica o *flow through* coletado. Os números em mM são referentes a concentração de imidazol na qual as amostras foram eluídas da coluna. O p1 e p2, representa frações de um mesmo pico que foi eluído. A. Refere-se à amostra purificada por cromatografia líquida de afinidade a níquel antes de ser aplicada na coluna de exclusão molecular. V.1 e V.2 representam as amostras que não adentraram a malha da coluna de exclusão molecular.

Em outro lote, após a etapa de purificação por cromatografia de afinidade a níquel as amostras contendo a banda de proteína no tamanho referente à proteína NS2b/NS3 de ZIKV foram purificadas em coluna de troca iônica (*strong anion*). Ao realizar gradiente de concentração de sal dois picos foram eluídos, o primeiro na concentração de aproximada de 180 mM e o segundo 230 mM de NaCl, como pode ser observado na figura 37.

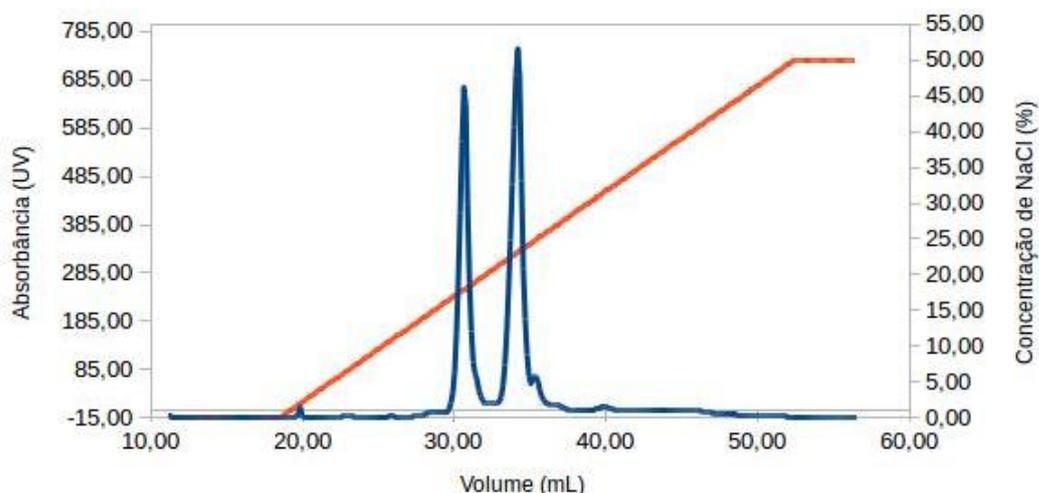


Figura 37- Purificação por troca iônica usando a coluna MonoQ (strong anion) da proteína NS2b/NS3 de ZIKV.

Os picos da purificação por troca iônica foram analisados em gel SDS page 12% (figura 38). É possível observar dois padrões de bandas diferentes entre o pico 1 e 2, onde no segundo é visto uma banda entre os marcadores de 35 kDa e 45 kDa, banda essa com o tamanho aproximado de um dímero.

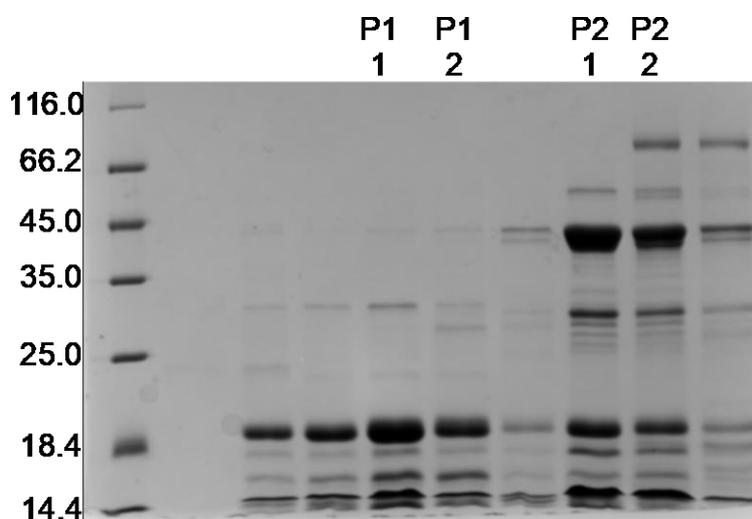


Figura 38- Gel SDS-PAGE 12% no qual pode ser observado picos referentes à purificação por troca iônica da protease NS2b/NS3 de ZIKV. P1 e P2 referem-se ao pico 1 (eluído a ~180 mM de NaCl) e pico 2 (eluído a aproximadamente 230 mM de NaCl) da figura 9. O número 1 refere-se à fração ascendente do pico e 2 à fração descendente.

Devido ao grande número de bandas nos poços referentes à protease de ZIKV,

visando analisar a possível autólise da proteína do pico 2, que apresentava inclusive um possível dímero, este foi analisado por ultracentrifugação analítica (figura 39). O coeficiente de sedimentação observado calculado de mais de 90% da amostra apresentou um valor entre 3,32 e 3,36 S, o que representa uma proteína de massa entre 48,1 e 52,6 kDa, o que corrobora a ideia de que a proteína se encontrava dimerizada e estruturada, pois o tamanho da proteína NS2b/NS3 de ZIKV era de 51,8 kDa, ou seja, dentro da margem de erro calculada pela ultracentrifugação analítica. A presença de homodímeros dessa proteína já tem sido citada na literatura (LEI et al., 2016). As diversas bandas no gel indicam que ela deve estar sofrendo clivagem, todavia parece manter a estrutura em solução e formando bandas em condições desnaturantes.

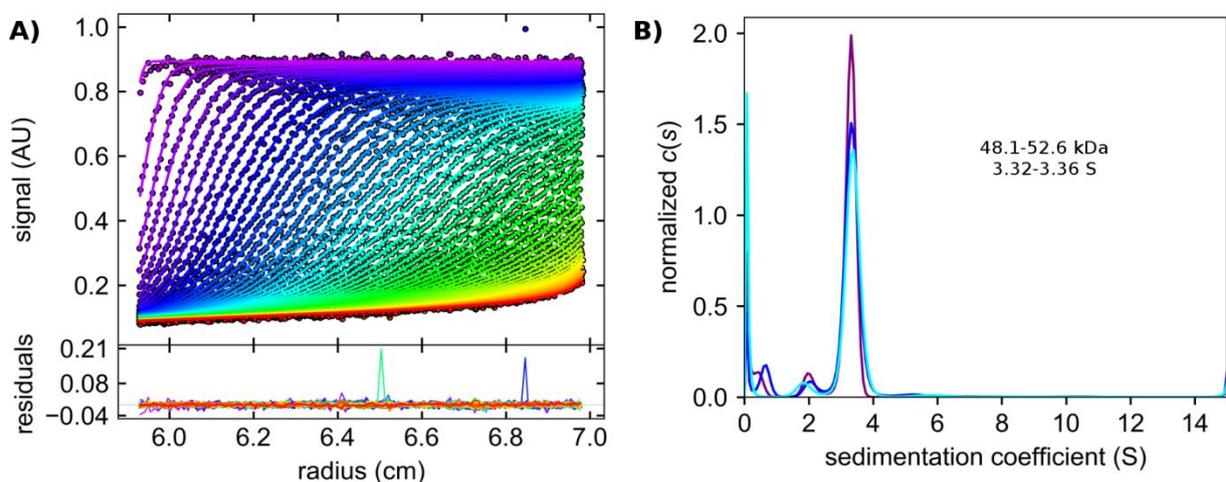


Figura 39- Análise por ultracentrifugação analítica da proteína NS2b/NS3 de ZIKV. (A) Perfil de sedimentação da protease. (B) Coeficiente de sedimentação calculado em diferentes concentrações de proteína.

A formação de dímeros dessa protease já foi anteriormente relatada na literatura, essa oligomerização é observada em diversos cristais da mesma, em que a unidade assimétrica é formada por moléculas nesse estado (5TFN, 5TFO, 6UM3, 5LC0, 5T1V e 5GXJ). Esta conformação dimérica ocorre de tal forma a posicionar a região do sítio ativo dos monômeros frente a frente, reduzindo a acessibilidade da região. Todavia, esta conformação não impossibilita a interação dos monômeros com moléculas inibidoras, ao menos de moléculas menores que 1 kDa, visto que diversas estruturas nessa conformação apresentam ligantes interagindo na região do sítio ativo. Além disso, esta conformação permite a formação de oligômeros, onde os dímeros são conectados entre si através de pontes dissulfeto pelo resíduo de cisteína 143 das diferentes moléculas, o qual

não se mostra presente em outras proteínas homólogas (LEI et al., 2016).

Uma segunda possibilidade de conformação dos dímeros também foi observado em estruturas cristalográficas (5GJ4). O contato acontece entre as regiões de folha β , o qual compreende a 1ª fita da NS2b e a 1ª e 2ª da protease NS3, além de alguns resíduos presentes em dois loops compostos por Val100 até Ala106 e Gli55 até Pro 67. Esta região de dimerização apresenta uma série de contatos polares, por exemplo, entre os resíduos Arg105 e Glu62, Arg55 e Glu104, e Ser33 e Glu54 (PHOO et al., 2016). A figura 40 apresenta os dois modelos citados de dímero.

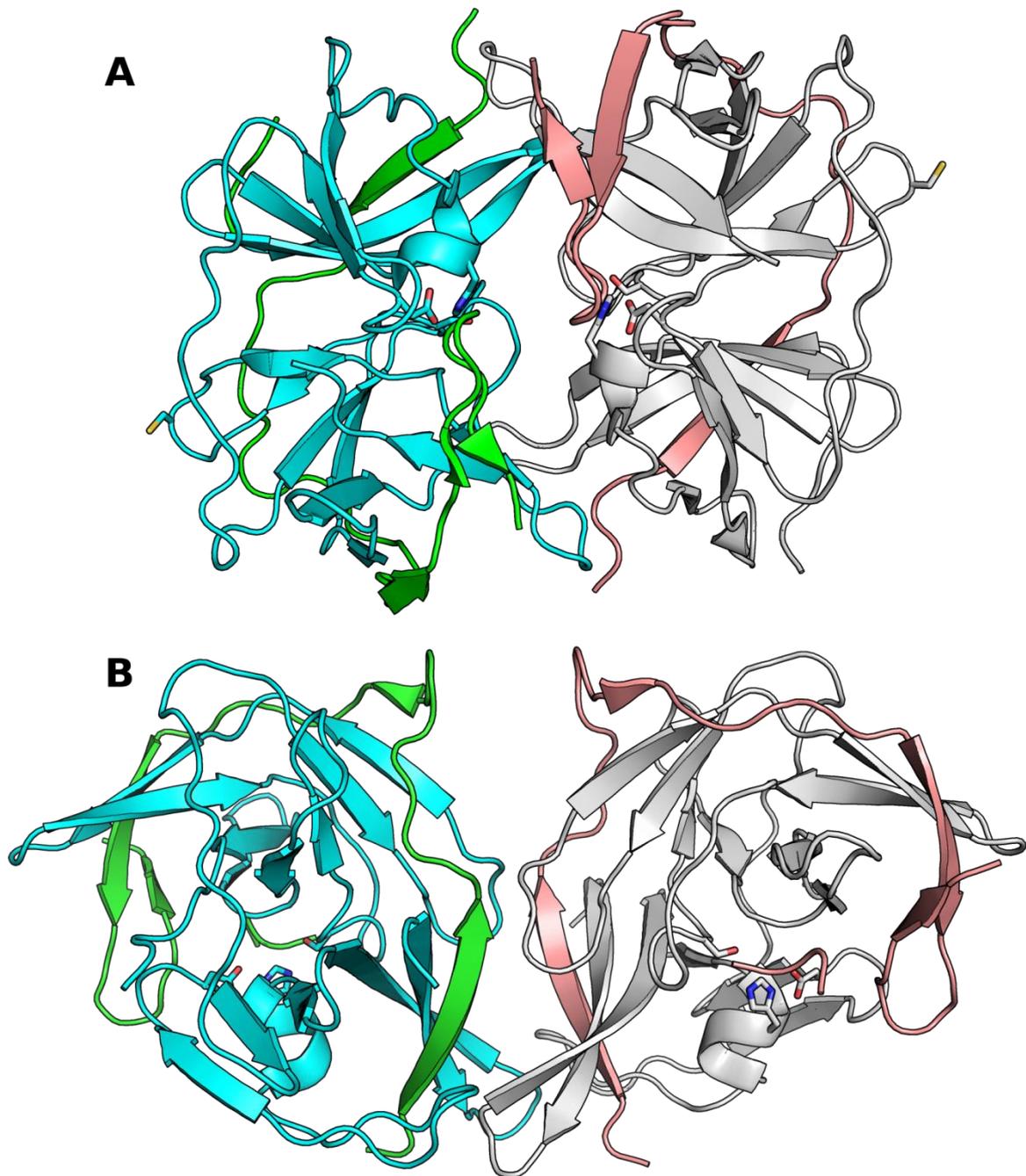


Figura 40- Representação estrutural de duas possíveis conformações de dímero de NS2b/NS3 da protease de ZIKV. A) Dímero observado por Lei e colaboradores (2016). B) Dímero observado por Phoo e colaboradores (2016). Os *sticks* representam os resíduos que compõe a tríade catalítica da protease, His51, Asp75 e Ser135. Na imagem A são observados também *sticks* presentes nas extremidades, referentes ao resíduo Cis143 responsável pela oligomerização.

Lei e colaboradores (2016) demonstraram que estes não são somente artefatos cristalográficos, os dímeros e oligômeros puderam ser observados em análises de

eletroforese nativa e SDS-PAGE. Nossos dados obtidos corroboram a formação de dímeros não cristalográficos da protease de ZIKV. Todavia, é importante ressaltar que a proteína NS3 analisada no presente trabalho apresenta uma construção onde está ausente o domínio helicase, a qual pode influenciar negativa ou positivamente nesse contato. Em se tratando da dimerização da fração protease e de sua importância para ação viral, não há dados claros, todavia estudos relacionados a NS2b, a qual interage com o domínio protease da NS3 para que esta tenha atividade, têm demonstrado que esta proteína atua como viroporina, a qual aumenta a permeabilidade da membrana celular facilitando assim a liberação viral na célula infectada. A ação dessa proteína se dá a partir de oligomerização e formação de poros, esta oligomerização já tem sido relatada na literatura (LEÓN-JUÁREZ et al., 2016). Devido à associação da NS3 com esta viroporina as interfaces de dimerização ou oligomerização podem auxiliar nos contatos durante essas etapas tardias de infecção (PHOO et al., 2016), entretanto, são necessários mais estudos para averiguar se estas se mostram relevantes.

Realizaram-se também testes preliminares de inibição da utilizando para isso a autólise da proteína. Procurou-se observar redução da intensidade da banda da protease NS2b/NS3 (~27 kDa) de ZIKV na presença de diversos inibidores de protease conhecidos: aprotinina, benzamidina, leupeptina, fluoreto de fenilmetilsulfonil (PMSF) e *Black-Eyed Pea Trypsin Chymotrypsin Inhibitor* (BTCl). Quando a proteína estava junto com aprotinina o decaimento se mostrava mais lento que o ensaio com os demais inibidores e que a protease sozinha. Este resultado sugere uma possível ação desse peptídeo como inibidor da NS2b/NS3 de ZIKV. O ensaio pode ser observado na figura 41.

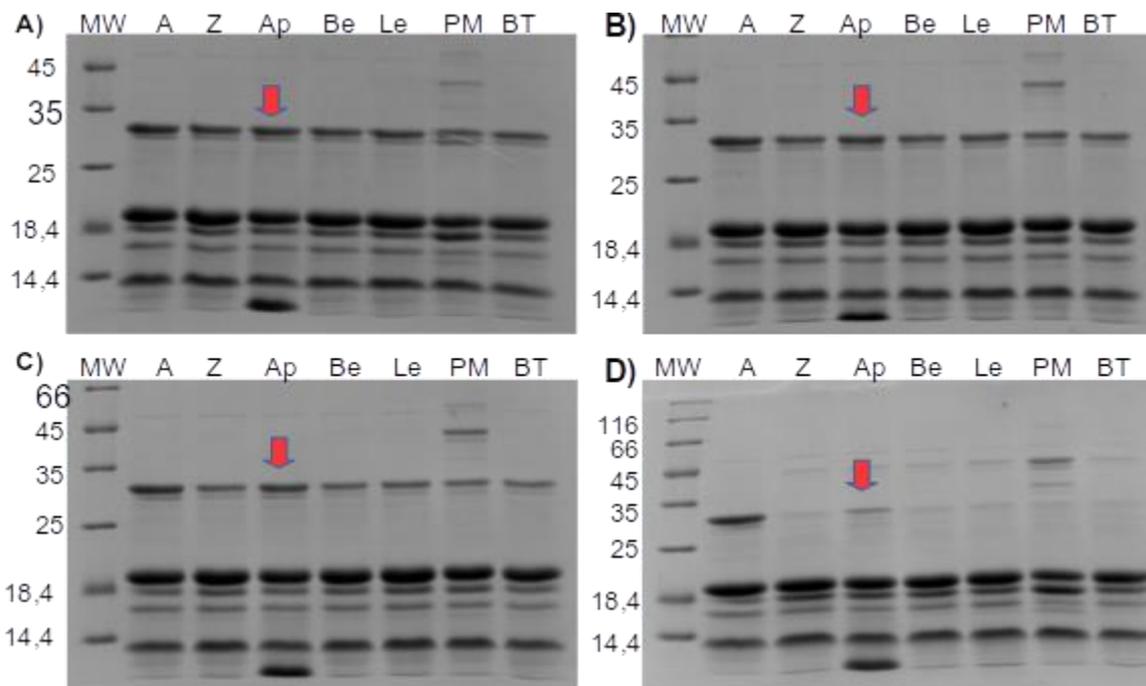


Figura 41- Ensaio de inibição da NS2b/NS3 de ZIKV com diversos inibidores de protease. (A) 1 hora. (B) 2 horas. (C) 3 horas. (D) 16 horas. MW- Marcador de peso molecular. A- NS2b/NS3 sozinha desovelada química e termicamente. Z- NS2b/NS3 sozinha. Ap- Protease com aprotinina. Be- Protease com benzamidina. Le- Protease com leupeptina. PM- Protease com PMSF. BT- Protease com BTCl.

Foram realizados experimentos de fluorescência da protease de ZIKV na presença de aprotinina, nos quais foi observado um aumento da concentração de aprotinina leva a uma redução da intensidade de fluorescência. Além disso, é observado um surgimento de uma segunda banda a direita, esse resultado é mostrado na figura 42.

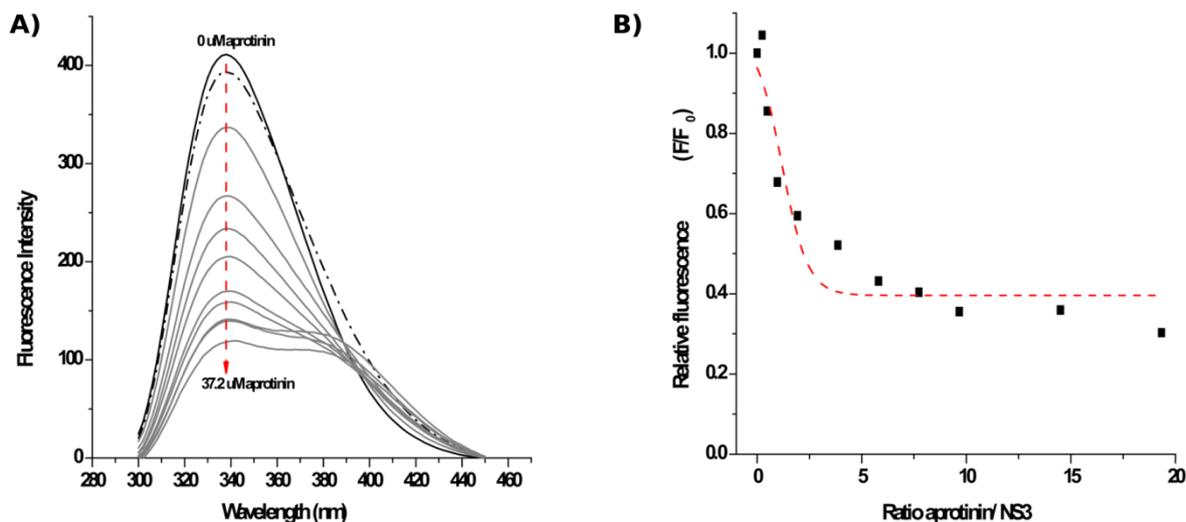


Figura 42- Ensaio de fluorescência da protease NS2b/NS3 de ZIKV com aprotinina. (A) Espectro de emissão de fluorescência de triptofano em diferentes concentrações de aprotinina. (B) Emissão de fluorescência máxima da proteína em presença de quantidades crescentes de aprotinina em relação a proteína em ausência do ligante. Curva de inibição da protease.

Esses resultados são corroborados por dados da literatura, pois sabe-se que outras proteases homólogas de flavivirus, por exemplo WNV (TOMLINSON; WATOWICH, 2008) e os quatro sorotipos da dengue (BRECHER; ZHANG; LI, 2013; LI et al., 2005), são fortemente inibidos por aprotinina, sendo que já há a estrutura cristalográfica dessa interação com a WNV (ALESHIN et al., 2007) e DENV3 (NOBLE et al., 2012). Essa inibição já foi também vista para ZIKV através de ensaios enzimáticos (CHEN et al., 2016).

Um experimento *in silico*, *docking* proteína-proteína, visando entender o modo de interação entre a NS2b/NS3 de ZIKV e aprotinina foi realizado. A pose escolhida realizava contatos, majoritariamente polares, na mesma região de interação que ocorria entre o inibidor e proteína homóloga da DENV3. Esta região de contatos engloba também o sítio catalítico da enzima, onde há contatos diretos inclusive com resíduos do sítio ativo e com resíduos da proteína NS2b. Esta interação pode ser observada na figura 43. Os contatos dos resíduos são bastante semelhantes, todavia o contato da R39 da aprotinina com resíduos da NS2b ainda não haviam sido observado na estrutura resolvida de dengue, todavia uma conformação parecida é observada na estrutura da protease de WNV com aprotinina, em que o R39 é encontrado posicionado próximo aos resíduos D80, D82, N84 e Q86, sendo estes os resíduos que compõem o loop da folha- β antiparalela que se

posiciona próximo ao sítio catalítico da NS3. De acordo com o modelo, a inibição ocorre pelo mesmo mecanismo que as demais proteínas homólogas anteriormente abordadas, em que é necessária a presença de um substrato autêntico (P2 e P1 resíduos básicos e P1' preferencialmente Gly, Ser, ou Thr) para a conformação correta do sítio oxianion (ALESHIN et al., 2007; NOBLE et al., 2012).

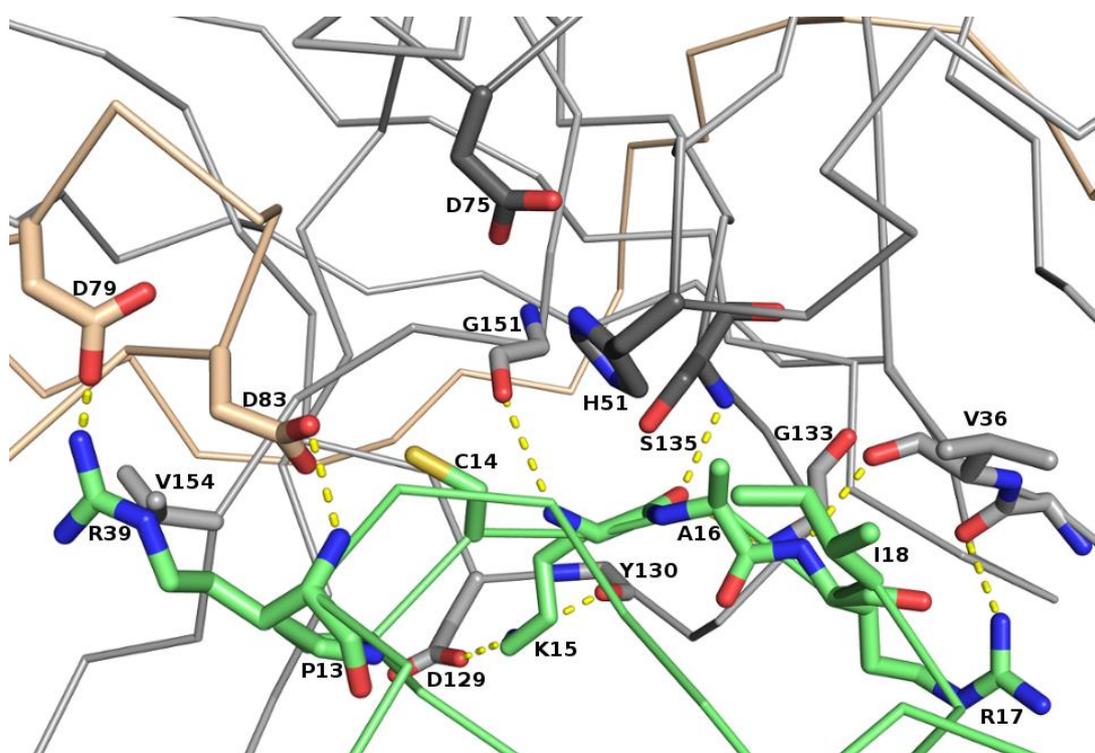


Figura 43- Interface de interação entre a protease NS2b/NS3 de ZIKV e aprotinina. A proteína NS3 está em cinza claro, com os resíduos do sítio catalítico em cinza escuro, a NS2b em coloração trigo e a aprotinina em verde. As ligações de hidrogênio estão representadas pelo tracejado amarelo.

Outro resultado interessante observado no *docking* foi a sobreposição da região de interação da aprotinina com a de dimerização da protease. Este local de inibição levanta a possibilidade de que o dímero seja desfeito com a interação com o inibidor (figura 44). Essa forma de interação parece corroborar com o resultado dos ensaios de fluorescência, pois a interação monômero e aprotinina deixaria o resíduo Trp50 mais acessível ao solvente, o que poderia explicar a redução da intensidade de fluorescência. Essa maior acessibilidade foi observada no triptofano na estrutura de homodímero em comparação ao modelo de monômero em interação com aprotinina por meio do cálculo de superfície

molecular acessível no servidor WHAT IF (VRIEND, 1990).

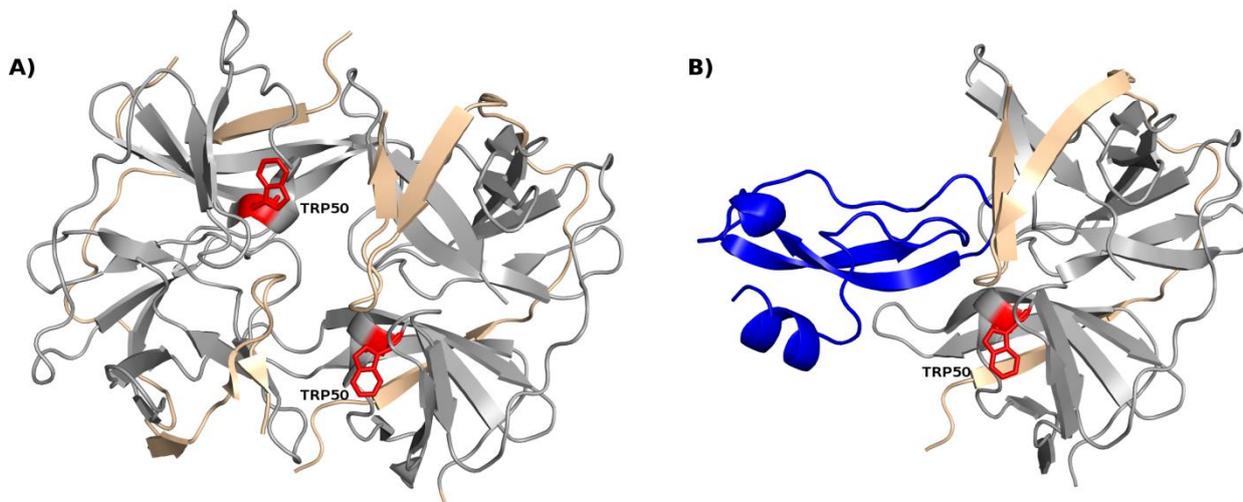


Figura 44- (A) Homodímero da NS2b/NS3 de ZIKV (5LC0) e (B) modelo da interação entre o monômero desta protease com inibidor obtido através de *docking* proteína-proteína. NS3 (cinza), NS2b (trigo), aprotinina (azul) e TRP50 (vermelho).

Diversas condições de cristalização foram testadas durante os experimentos de *screening*. Uma condição de cristalização que apresentou cristais foi dioxano 10%, sulfato de amônio 1,6 M, 0,1 M MES pH 6,5. Essa condição foi refinada e foram obtidos cristais nas concentrações dos reagentes dioxano 10%, sulfato de amônio 1,2 M, 0,1 M MES pH 6,1. A foto dos cristais pode ser vista na figura 45.

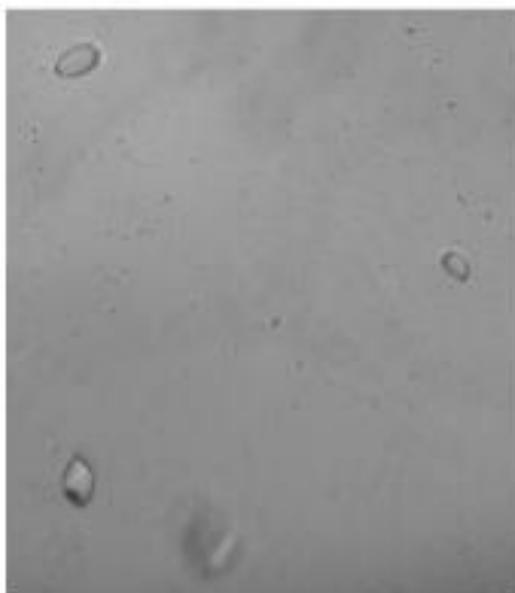


Figura 45- Cristal da protease NS2b/NS3 de ZIKV (dioxano 10%, sulfato de amônio 1,2 M, 0,1 M MES pH 6,1).

Esta condição de cristalização ainda não foi reportada na literatura. Todavia, não foi possível obter padrões de difração para coleta dos dados, definição do grupo espacial e resolução da estrutura da proteína. Esta ausência de padrão de difração do cristal coletado se dá por uma desorganização interna do cristal, esta poderia ocorrer devido à proteína utilizada apresentar diversas bandas no gel, o que poderia indicar diversas populações proteicas dentro da mesma amostra, apesar de que os ensaios de ultracentrifugação apresentam um pico único. Além disso, a própria possibilidade de regiões clivadas na proteína poderia dificultar a formação coesa e organizada o suficiente para gerar padrões de difração coletáveis.

Uma segunda construção da protease NS2b/NS3 de ZIKV foi obtida com o Professor Dr. Rolf Hilgenfeld, denominada ZIKV-NS3-Lub no presente trabalho. A expressão e purificação da mesma ocorreram de acordo com os dados publicados (LEI et al., 2016). Devido às mutações adicionadas nessa construção visando prevenir a autoclivagem, não é observada uma grande quantidade de bandas nas análises por gel SDS-PAGE como a construção ZIKV-NS3-Bra, o que pode ser ilustrado na figura 46.

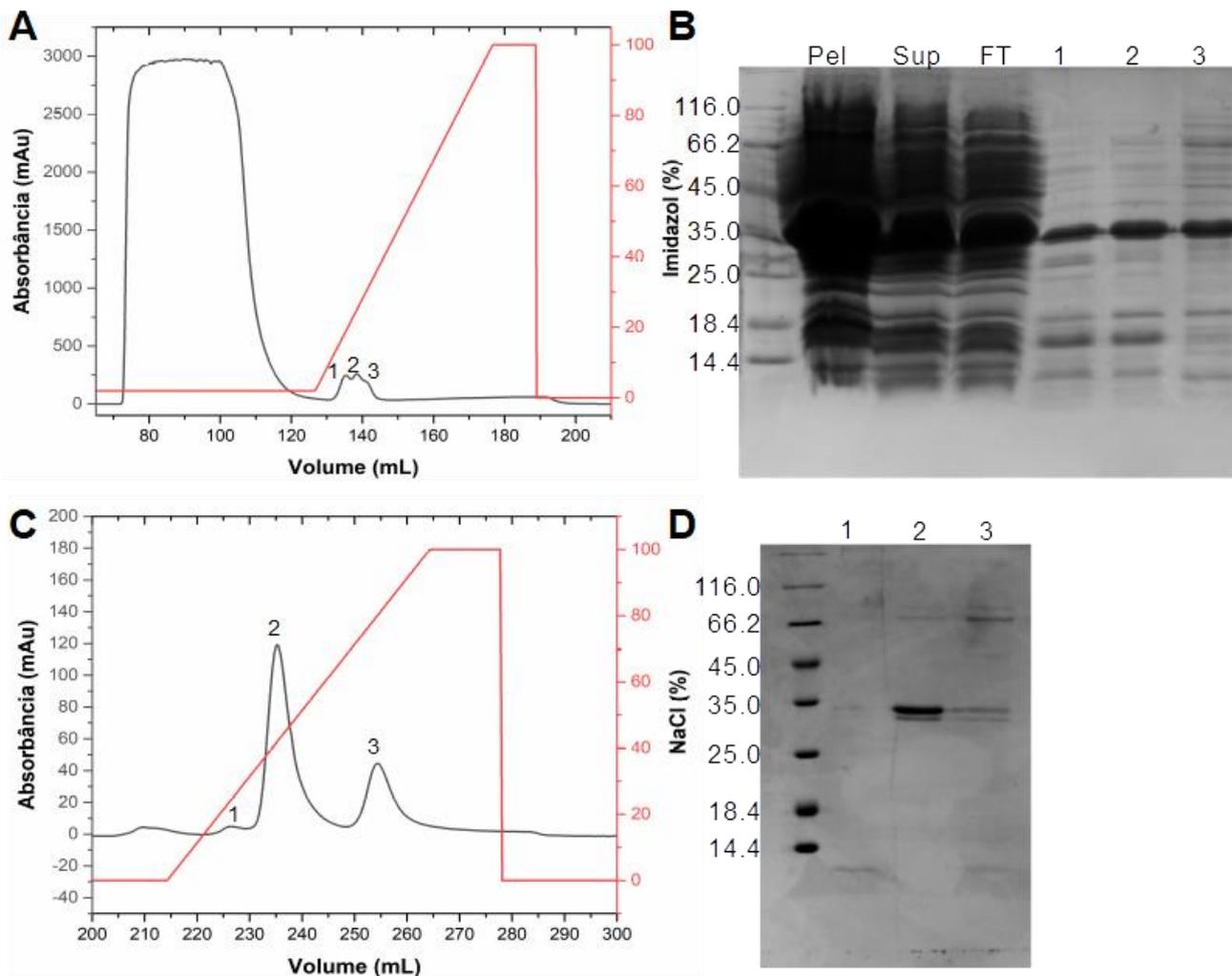


Figura 46- Expressão e purificação da proteína NS3 de ZIKV (ZIKV-NS3-Lub). A) Perfil cromatográfico da purificação por gradiente de imidazol usando coluna de níquel. A linha vermelha refere-se à concentração de imidazol em porcentagem (0 a 500 mM). B) Análise por gel SDS-PAGE 15% da expressão e purificação utilizando-se coluna de níquel. Pel e Sup referem-se ao pellet e sobrenadante obtidos durante as etapas de lise e centrifugação, FT às proteínas que não interagiram com a coluna de níquel, e 1, 2 e 3 a cada um dos picos coletados. C) Perfil cromatográfico da purificação por coluna de troca aniônica. A linha vermelha refere-se à concentração de NaCl em porcentagem (0 a 1 M). D) Análise por gel SDS-PAGE 15% da purificação por coluna de troca aniônica, onde os números 1, 2 e 3 mostram os picos coletados.

A proteína novamente pareceu apresentar duas bandas com tamanhos próximos a monômero e a dímero. Na figura 46C é possível observar dois picos principais, os quais quando analisados em SDS-PAGE (figura 46D) parecem apresentar dois padrões diferentes, o 2 apresenta a banda forte referente ao monômero enquanto o 3 a banda

referente a dímero se mostra mais intensa. Esse padrão de monômero e dímero já havia sido observada com essa construção (LEI et al., 2016). Com a proteína apresentando um bom nível de pureza foram então realizadas as etapas dos ensaios enzimáticos e de inibição.

Primeiramente foi estudada a influência da concentração da enzima na atividade enzimática e calculada sua atividade específica. Observou-se que até valores próximos a $1,3 \times 10^{-4}$ mg da enzima é observada uma linearidade entre a atividade enzimática e a concentração da enzima, figura 47. Com concentrações de enzima superiores ocorre uma perda do comportamento linear provavelmente por perda da condição inicial não limitante do substrato ou por agregação proteica. A quantidade de $0,5 \times 10^{-4}$ mg da NS2B/NS3 presente na região linear da curva foi selecionada para a realização dos ensaios enzimáticos na determinação dos parâmetros cinéticos. Ao mesmo tempo o valor de atividade específica da enzima foi obtido da inclinação da linha de ajuste sendo igual a $2,15 \times 10^{-7}$ kat/mg de enzima.

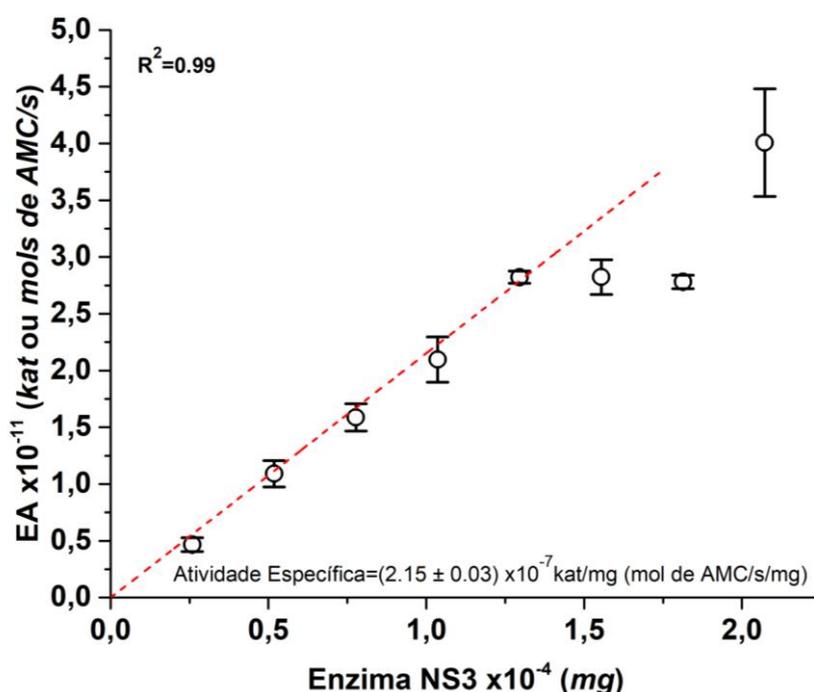


Figura 47- Gráfico da influência da concentração de enzima na atividade enzimática e determinação da atividade específica da protease NS2B/NS3 de ZIKV com o substrato Benzoyl-Nle-Lys-Arg-Arg-AMC.

Em seguida foi calculado a cinética de Michaelis-Menten, o gráfico construído pode

ser observado na figura 48. O valor da constante K_m calculado a partir do gráfico foi de $11,02 \mu\text{M}$, o qual é superior ao anteriormente publicado que foi de $5,9 \mu\text{M}$ bastante próximo ao valor referente ao monômero $18,3 \mu\text{M}$ (LEI et al., 2016). A diferença entre os valores obtidos poderia ser explicada por variações nos componentes do tampão utilizado para realização do ensaio, como a concentração de glicerol de 5%, usado no presente trabalho, para 20%, ou pela presença de 1 mM de CHAPS. Os valores de K_{cat} e EC foram respectivamente de $2,245 \text{ s}^{-1}$ e $95735 (\text{s}^{-1} \text{ M}^{-1})$.

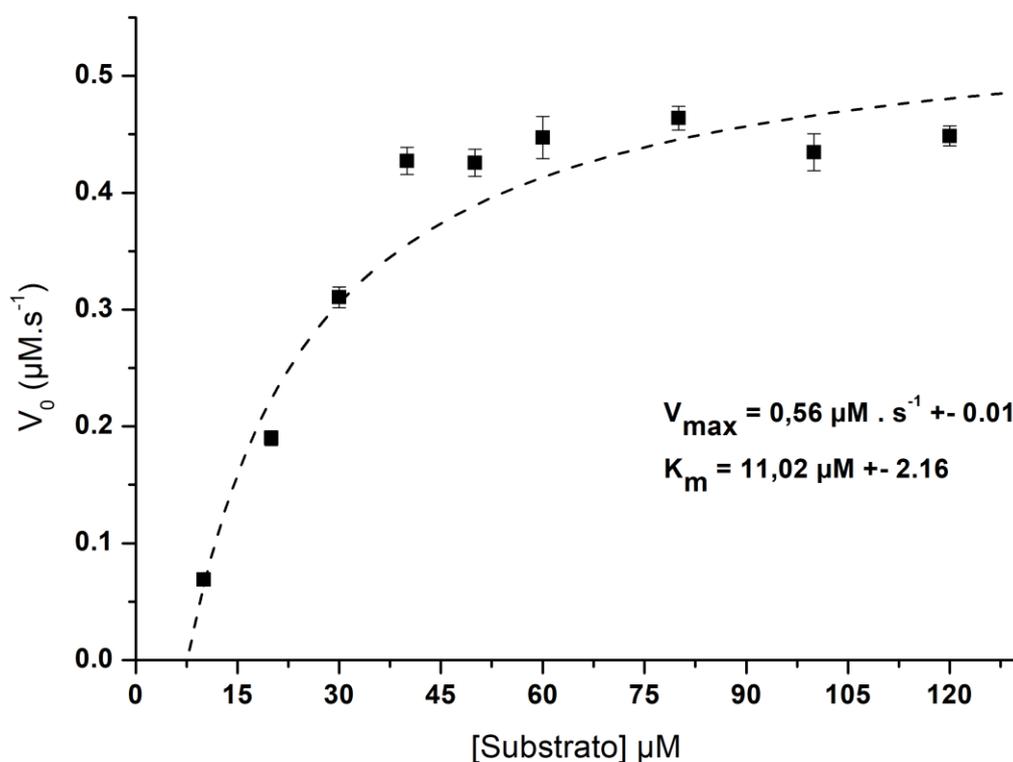


Figura 48- Gráfico do ensaio de Michaelis-Menten da protease de NS2B/NS3 de ZIKV na presença substrato Benzoyl-Nle-Lys-Arg-Arg-AMC.

Outro perfil cinético que poderia ser obtido a partir gráfico do ensaio de Michaelis-Menten seria um perfil cinético sigmoidal, o que difere do perfil hiperbólico clássico. Este padrão está relacionado com cooperatividade e alosteria onde a interação de um ligante, sendo este um substrato, um inibidor ou um ativador, em um sítio da enzima pode afetar a afinidade com que outros sítios interagem com seu respectivo substrato (CÁRDENAS, 2015). Este resultado poderia ser explicado através da formação de dímeros desta

protease, onde as subunidades cooperam durante a ação enzimática na qual a interação do substrato Benzoyl-Nle-Lys-Arg-Arg-AMC em uma das subunidades modifica a afinidade de ligação na outra. Outra possibilidade é a existência de um sítio alostérico não catalítico de interação do próprio substrato nas subunidades que modula a atividade enzimática. Porém, um sítio dessa natureza ainda não foi descrito para esta enzima ou enzimas homólogas apesar desse substrato tem sido amplamente utilizado para caracterização destas proteínas.

Foram realizados ensaios de inibição para cálculo de IC₅₀ com moléculas que apresentaram inibição para proteases homólogas como a NS3 de DENV (SILVA, 2012). Dentre as quatro moléculas selecionadas (F0174-0048, F0745-1098, F0816-0342 e F3165-0782) a F0745-1098 não pode ser testada devido a problemas relacionados a solubilização do composto. Além destes foi testado também a leupeptina, um conhecido inibidor de variadas proteases. O resultado deste ensaio pode ser visto na figura 49.

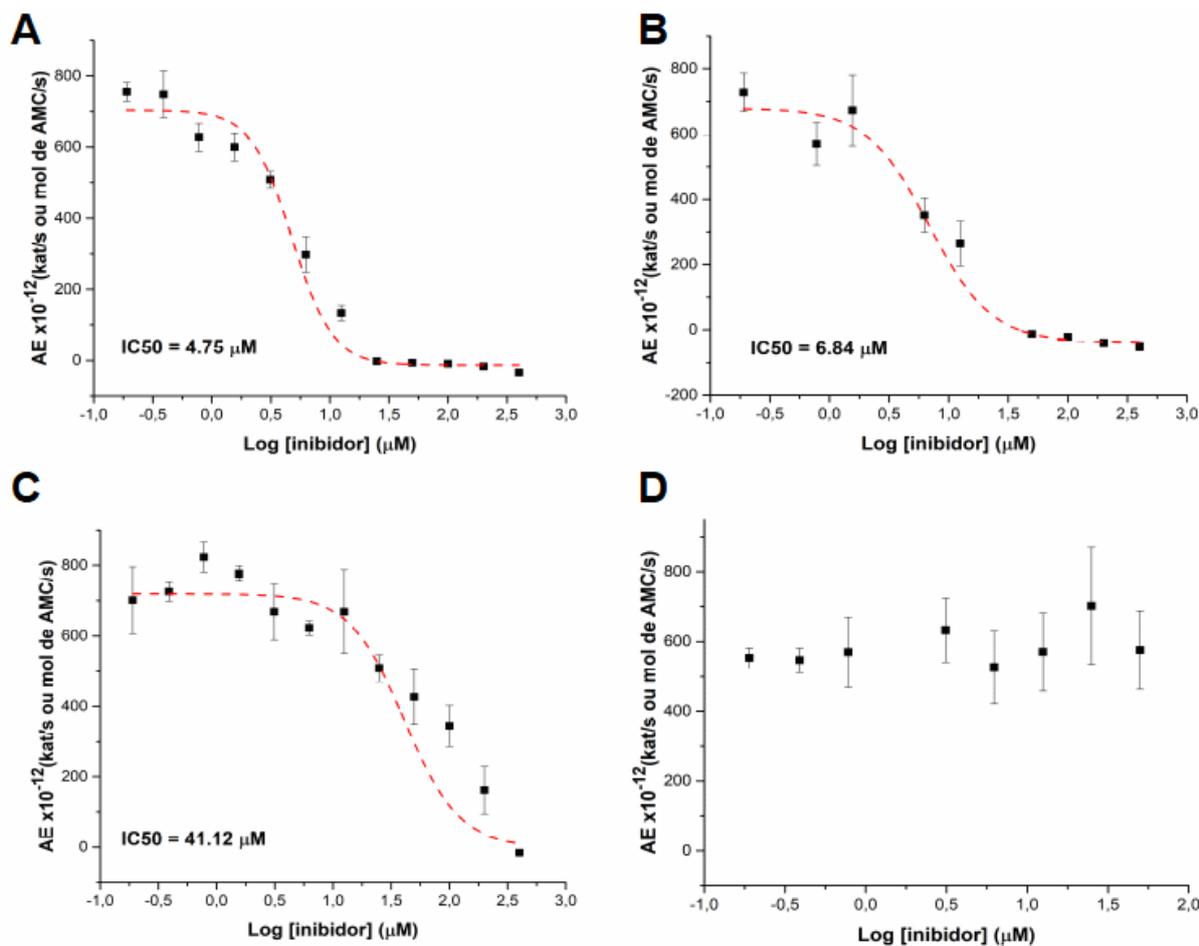


Figura 49- Ensaio de inibição enzimática da protease NS2B/NS3 de ZIKV para cálculo de IC₅₀ na presença das moléculas A) F0174-0048, B) F0816-0342, C) F3165-0782 e D) leupeptina.

As moléculas testadas apresentaram atividade inibitória superior para a protease de ZIKV do que quando testadas para DENV. A F0174-0048, a qual havia apresentado um valor de 78,73 µM, obteve um valor de 4,75 µM, por sua vez a F0816-0342 e F3165-0782 obtiveram os valores de 6,84 µM e 41,12 µM enquanto tinham apresentado os 93,17 µM e 94,95 µM, respectivamente (SILVA, 2012). Outras drogas testadas para protease de ZIKV como *Asunaprevir* e *Simeprevir*, selecionadas a partir de modelos de busca baseados em farmacóforos, também apresentaram valores similares aos obtidos para os compostos F0174-0048 e F0816-0342 com IC₅₀ de 6,0 µM e 2,6 µM, respectivamente (PATHAK et al., 2020). Outro trabalho, no qual se analisou o reposicionamento de dois inibidores da proteína NS3/NS4A do Vírus da Hepatite C como inibidores da NS2B/NS3 de ZIKV apresentaram valores similares de IC₅₀ de 5,2 µM e 4,1 µM (LEE et al., 2017). Sendo assim, as moléculas aqui estudadas mostram atividade inibitória para a protease

de ZIKV, com valores inibitórios na ordem do micromolar e similares a outros inibidores efetivos descritos na literatura.

Além disso, outros compostos com bons resultados de inibição da NS2B/NS3 de ZIKV como o 5-amino-1-((4-methoxyphenyl)sulfonyl)-1H-pyrazol-3-yl benzoate (composto 1), que apresentou valor de IC50 de 1,5 μM , que foi utilizado para a produção de derivados gerando outras moléculas, como o composto 5, com valor aprimorado de IC50 de 0,1 μM (LI et al., 2018). Uma abordagem similar foi realizada para outros inibidores da protease de ZIKV atingindo também potências inibitórias em nível de nanomolar, 0,2 μM (YAO et al., 2019). As moléculas estudadas no presente trabalho, devido às suas capacidades inibitórias contra a protease em questão mostram-se como excelentes alvos para aperfeiçoamento do contato com a proteína em questão, para o aumento da afinidade de ligação e conseqüentemente melhores resultados em ensaios de inibição.

Os ligantes em questão tiveram sua interação analisada a partir de análises *in silico*, que sugeriu a interação do F0174-0048 e F3165-0782 no mesmo sítio de interação ao co-fator da DENV, realizando interações de hidrogênio com os resíduos Met59 e a Ile22, além de interações hidrofóbicas com vários resíduos que compõe a região como Ile25, Ala56, Val52, entre outros. Este contato poderia impedir interação com a NS2B, a qual é necessária para que a protease realize sua atividade. Por sua vez o F0816-0342 parece interagir diretamente no sítio ativo da proteína através de ligações de hidrogênio com os resíduos Gly133 e Ser135, o que impediria o contato com o substrato desta enzima (SILVA, 2012). Devido ao tamanho dos ligantes, outra abordagem seria analisar se seus fragmentos seriam capazes de manter a atividade inibitória. Todavia, são necessários experimentos *in vitro* visando corroborar os modelos de interação sugeridos anteriormente.

5.2.1. Dengue

A protease NS2b/NS3 de DENV foi expressa de acordo com o protocolo anteriormente descrito (SILVA, 2012). Na figura 50 observa-se a expressão solúvel em diversas cepas de *E. coli*.

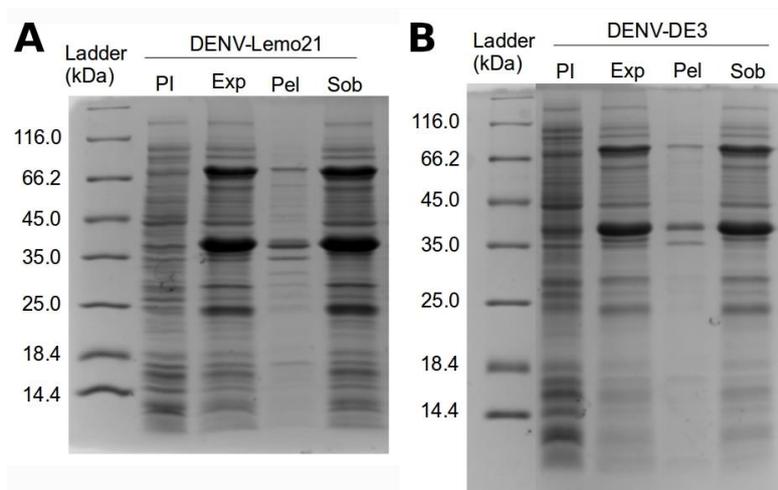


Figura 50- Géis de expressão SDS-PAGE 15% da protease NS2b-NS3 de DENV. Expressão após transformação nas cepas Lemo21 (A) e DE3 (B). PI é referente as amostras antes da indução com IPTG, Exp a fração insolúvel e solúvel juntas, enquanto Pel mostra a insolúvel e Sob a solúvel separadamente.

A protease apresentou um tamanho na análise por gel SDS-PAGE 15% superior a 35 kDa, todavia o tamanho real é de 29,256 kDa. Entretanto, essa variação no tamanho da proteína já foi observada anteriormente com este tipo de construção (SILVA, 2012) e em outros trabalhos na literatura (LI et al., 2005).

Após a constatação da presença da proteína heteróloga na fração solúvel realizou-se purificação desta por meio de cromatografia líquida de afinidade a metal (Ni). Realizou-se uma etapa de gradiente até 100 mM de IPTG seguida por duas etapas de lavagem com 250 mM e 500 mM. A diferença observada entre os dois picos foi em relação a ao maior tamanho da banda de 66 kDa, que pode estar relacionada a uma estrutura dimérica da protease em questão.

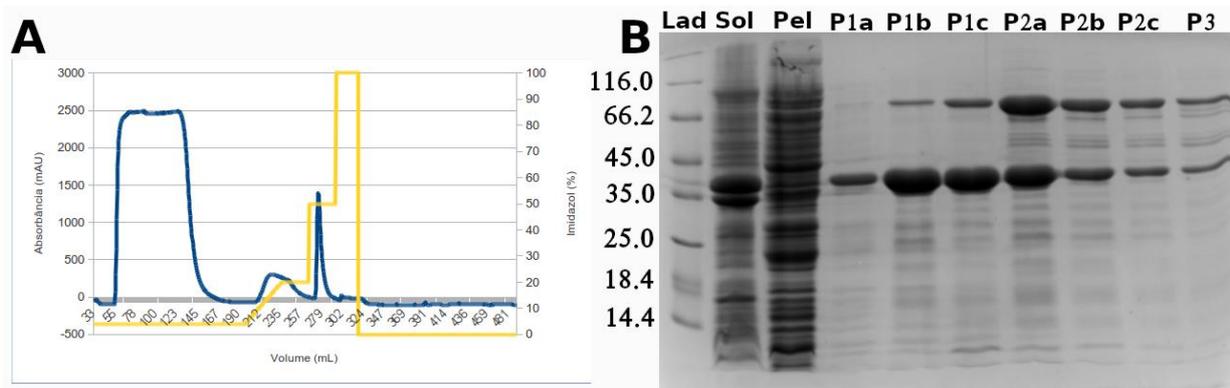


Figura 51- (A) Cromatograma da purificação por cromatografia líquida de afinidade a metal (Ni) da protease NS2b e NS3 de DENV. (A) A amostra é referente a expressão na cepa BL21 DE3. A absorbância (mAU) é observada pela linha azul e a porcentagem de imidazol é observada pela linha amarela, 100% refere-se à concentração de 500 mM no tampão. (B) Gel SDS-PAGE 15% da purificação por coluna de Níquel, Sol é referente à fração solúvel, Pel a fração insolúvel, P1 e P2 são referentes ao primeiro e segundo picos enquanto a, b, e c à diferentes frações dos picos.

A amostra do pico 2 foi então analisada por exclusão molecular, o resultado pode ser observado na figura 52. O cromatograma mostra um pico único onde seu ápice de aproximadamente 330 mAU se encontra no volume eluído aproximado de 80 mL. Para obter uma estimativa do tamanho da proteína em questão, este valor foi comparado com a curva de calibração da coluna, que indicou um tamanho de próximo a 60 kDa, o qual é próximo de um dímero desta protease.

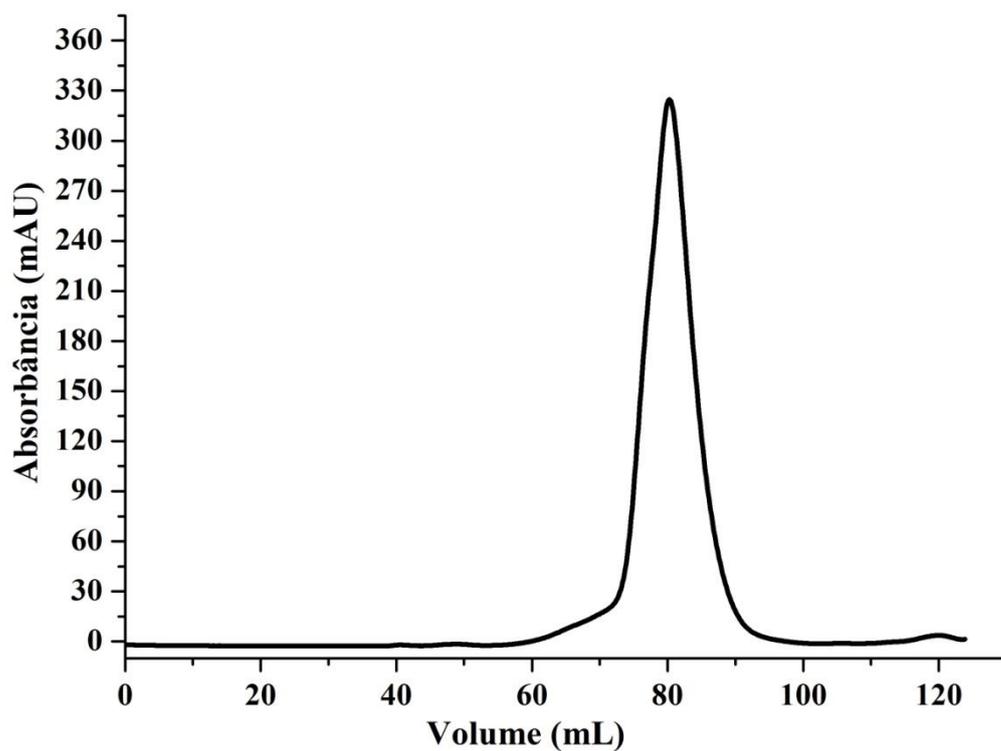


Figura 52- Purificação por exclusão molecular da protease NS2b/NS3 de DENV. A coluna utilizada para o experimento foi a S200.

Esta amostra foi também analisada por meio de ultracentrifugação analítica usando-se o método de velocidade de sedimentação. Foram observados dois picos nas amostras analisadas, o primeiro possui um coeficiente de sedimentação de 2,691 S, com uma massa de 33,982 kDa, correspondendo a 33,41% da amostra analisada. O segundo com um coeficiente de sedimentação de 3,667 S, para o qual foi estimada uma massa molecular de 54, 064 kDa, representando 63,19% da amostra, portanto um tamanho próximo a um dímero. Estes resultados podem ser observados na figura 53.

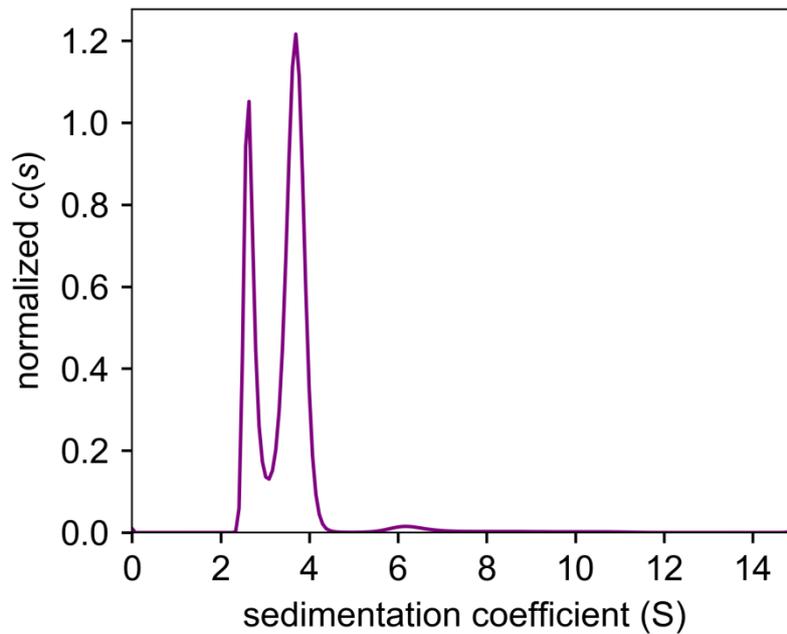


Figura 53- Análise da proteína NS2b/NS3 de DENV em ultracentrífuga analítica por velocidade de sedimentação.

A protease NS2b/NS3 de DENV tem sido relatada na literatura como monômero, por análise de cromatografia por exclusão molecular e por espalhamento dinâmico de luz (DLS) para o sorotipo 2 (YILDIZ et al., 2013), e por exclusão molecular e MALDI-TOF para o sorotipo 1 (CHANDRAMOULI et al., 2010). Nos artigos anteriormente citados não foram observados na composição dos tampões a presença de agentes redutores de pontes dissulfeto. Nossos resultados parecem sugerir que esta protease pode se organizar em dímeros, além disso, existem estruturas cristalográficas que apresentam dímeros em suas unidades assimétricas como as dos códigos de PDB 3U1I (NOBLE et al., 2012) e 6MO1 (YAO et al., 2019), sendo estas diferentes dos modelos anteriormente apresentados para os homólogos de ZIKV. Todavia, como anteriormente ressaltado para ZIKV, a construção expressa no presente trabalho não contém o domínio helicase, o qual poderia influenciar nesse processo. Sendo assim, são necessários mais estudos visando analisar mais profundamente a organização dessa protease em dímeros.

Após essa etapa foram então realizados *screening* de condições de cristalização, com alguns *kits* de cristalização visando a obtenção de cristais da protease NS2b/NS3. Diversas condições apresentaram cristais pequenos como pode ser observado nas figuras 54, 55 e 56. Todavia, algumas apresentaram cristais um pouco maiores como 0,5 % w/v *Jeffamine* ED-2001; pH 7,0; 100 mM HEPES; pH 7,0; 1,1 M malonato dissódico; pH 7,0, figura 55B, e 100 mM acetato de sódio; pH 4,5; 1 M fosfato de diamônio, figura 56F,

condições que não foram encontradas na literatura. Todavia, os cristais obtidos foram submetidos à difração por raios-x mas não se observou padrão de difração.

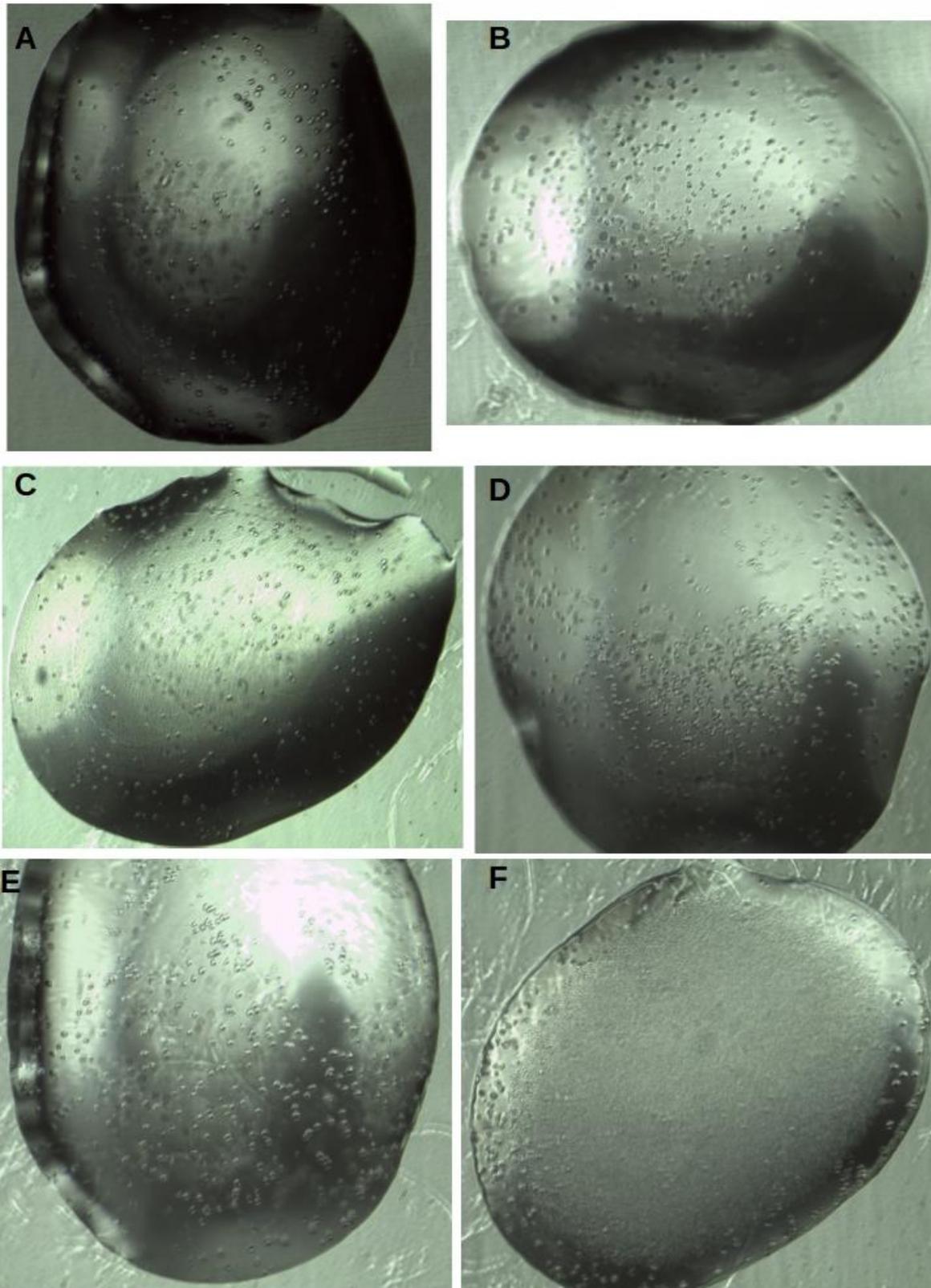


Figura 54- Fotos das gotas com cristais da placa de *screening* PACT++ feita com a

proteína de NS2b/NS3 de DENV a 17 mg/mL em tampão tris 20 mM e NaCl 300 mM. A- 25% w/v polietilenoglicol 1500; 100 mM SPG Buffer; pH 8,0 (A5) B- 25% w/v polietilenoglicol 1500; 100 mM TBG Buffer; pH 8,0 (A11). C- 20% w/v polietilenoglicol 6000; 100 mM MES; pH 6,0; 10 mM Cloreto de zinco. D- 20% w/v polietilenoglicol 6,000, 100 mM TRIS; pH 8,0, 200 mM Cloreto de Amônio. E- 20% w/v polietilenoglicol 6,000, 100 mM TRIS; pH 8,0, 200 mM Cloreto de lítio. F- 20% w/v polietilenoglicol 3,350, 100 mM BIS-TRIS propano; pH 7,5, 200 mM Fluoreto de sódio.

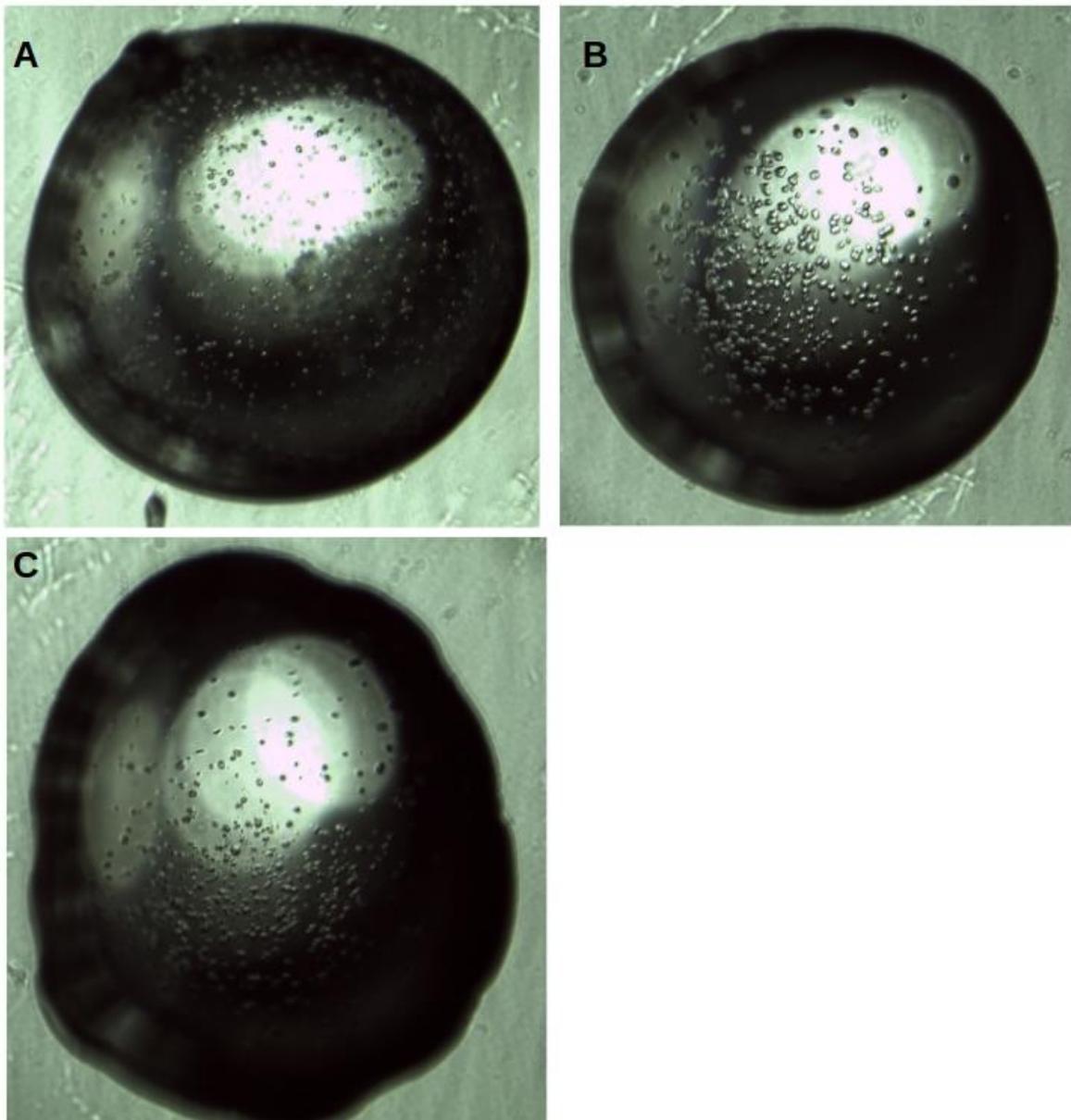


Figura 55- Fotos das gotas com cristais da placa de *screening* JCSG++ feita com a proteína de NS2b/NS3 de DENV a 17 mg/mL em tampão tris 20 mM e NaCl 300 mM. A- 800 mM Succinato dissódico; pH 7,0. B- 0,5 % w/v *Jeffamine* ED-2001; pH 7,0; 100 mM HEPES; pH 7,0; 1,1 M malonato dissódico; pH 7,0. C- 1% w/v éter monoetílico de polietilenoglicol 2000; 100 mM HEPES; pH 7,0; 1 M Succinato dissódico; pH 7,0.

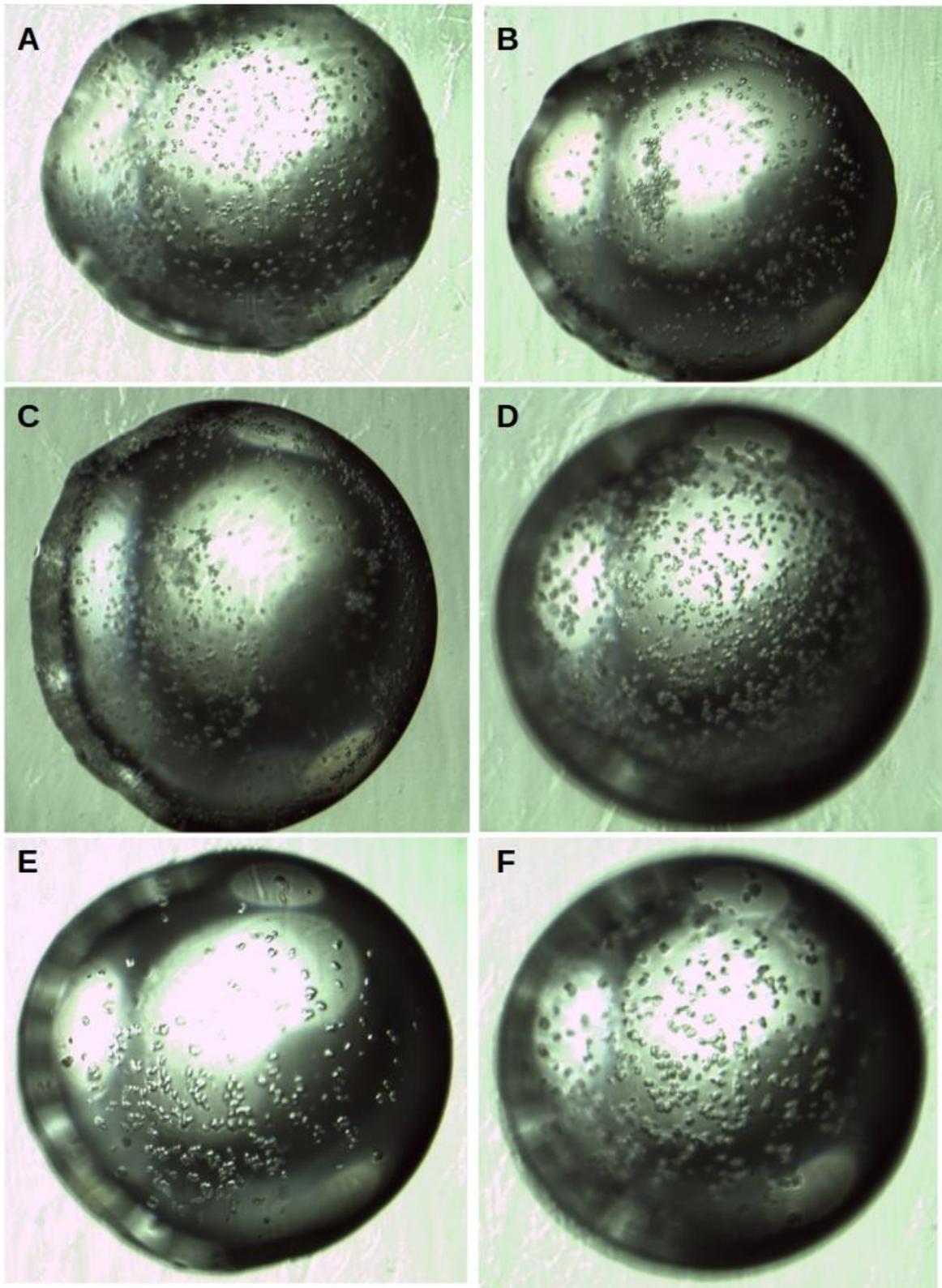


Figura 56- Fotos das gotas com cristais da placa de *screening* JCSG++ feitas com a proteína de NS2b/NS3 de DENV a 17 mg/mL em tampão tris 20 mM e NaCl 300 mM. A- 20 % w/v polietilenoglicol 3,350, 200 mM citrato tripotássico; pH 8,3. B- 100 mM MES; pH

6,5; 1 M citrato trissódico. C- 2M Sulfato de amônio; 100 mM MES; pH 6,5. D- 1,26 M Sulfato de amônio; 100 mM TRIS; pH 8,5; 200 mM Sulfato de lítio. E- 100 mM acetate de sódio; pH 4,5; 1 M fosfato de diamônico. F- 100 mM MES; pH 6,5; 1,6 M sulfato de magnésio.

Uma possibilidade para explicar a ausência de cristais com bons padrões de difração é a presença de *tags*. Apesar de muitas vezes pequenas *tags* não influenciarem no empacotamento cristalino ou gerarem efeitos adversos quando se trata da estrutura ou atividade proteica, o fato da obtenção dos cristais que não apresentam difração mostrando uma desorganização da rede cristalina poderia estar relacionado à não remoção da mesma (BOOTH et al., 2018; CARSON et al., 2007; DELLER; KONG; RUPP, 2016), além da presença do sítio de clivagem de enterokinase após a sequência da cauda de histidina.



Self Portrait in his Atelier

1884 54x85cm oil/canvas

Claude Monet

7. Conclusão

7. CONCLUSÃO

O presente trabalho buscou levantar dados estruturais e de interação com inibidores e substratos das proteases dos arbovírus DENV, ZIKV, CHIK e MAY, os quais tem grande incidência no Brasil. Estes novos dados gerados visam auxiliar no processo de entendimento da estrutura e ação das proteases que compõe estes vírus, o que pode beneficiar etapas futuras como a de desenvolvimento de novos inibidores.

Foram otimizadas as etapas de expressão, purificação e remoção da cauda de histidina para a protease nsP2 de CHIK. Além disso, foi expressa uma construção contendo o sítio de clivagem desta proteína, por meio da qual foi possível realizar ensaio de atividade da mesma.

A protease de MAYV, por sua vez, não apresentou bons resultados em relação à purificação, e apesar de mostrar uma grande quantidade na expressão não foi obtida na fração solúvel de forma clara nas diversas tentativas de otimização deste trabalho. Novas construções foram montadas para tentativa de expressão em pET-15b, pET-28a, pET-sumo e pGEX, todavia os testes iniciais de expressão não apresentaram resultados claros de produção desta proteína sendo necessário a busca da otimização dessas novas condições. Além disso, a análise por CD da amostra obtida através dos experimentos de renaturação não indicaram que ela havia sido renaturada com o dobramento correto. São necessários outros experimentos visando tentar obter a proteína estruturada na fração solúvel.

Quanto às análises *in silico* dessas proteases, foi feito um modelo por homologia para a protease nsP2 de MAYV, o qual não apresenta estrutura resolvida e depositada no banco de dados PDB. Analisamos também a interação da protease de MAYV e CHIK com substratos decapeptídicos (P6-P4'), pela qual foi proposto tanto os resíduos importantes do substrato quanto os que compõe os subsítios da protease responsáveis por acomodá-los. Estes foram pela primeira vez descritos para MAYV e foram ampliados para CHIK. Além disso, observou-se que os ângulos diedrais ϕ e ψ da glicina P2 do substrato de CHIK se mantiveram na maior parte do tempo da MD em conformações não permitidas para outros resíduos, indicando que este pode ser um dos motivos para a conservação deste resíduo em todos os substratos de todos os Alphavirus aqui estudados. Outro dado aqui analisado foi o da estabilização dos ângulos χ_1 dos resíduos catalíticos C478 e H548, e χ_2 da H548, tanto para MAYV quanto para CHIK, quando estes se encontravam na presença do substrato, mas o mesmo fato não foi observado para C/S482, que tem sido

estudada como capaz de manter a capacidade catalítica da proteína na ausência de C478.

Em se tratando dos Flavivirus, a protease NS2b/NS3 de ZIKV foi estudada em duas construções, a primeira ausente de mutações teve sua expressão e purificação otimizadas, apesar de apresentar autólise. A proteína demonstrou se organizar em dímero, o que foi confirmado por análise de ultracentrifugação analítica. Testes prévios com inibidores de protease conhecidos demonstraram a capacidade de inibição por aprotinina e análises de fluorescência titulando esse inibidor mostraram redução na intensidade do sinal, o que foi hipotetizado estar ocorrendo por causa da interação desfazer o dímero e expor o W150 ao solvente.

A construção mutada, que não apresentou autólise, foi utilizada para ensaios de atividade e inibição proteica. Calcularam-se assim os parâmetros cinéticos da enzima através do gráfico de Michaelis-Menten. Observou-se que a proteína foi inibida por alguns inibidores anteriormente testados para a protease de DENV1, sendo estes F01740048, F08160342 e F31650782.

No caso da NS2b/NS3 de DENV1 a protease foi expressa e purificada, que demonstrou se organizar também dímeros através da análise via purificação por exclusão molecular. Nas análises por ultracentrifugação analítica apresentou duas populações, sendo a maior delas a de dímero.

8. REFERÊNCIAS BIBLIOGRÁFICAS

- ABAD-FRANCH, F. et al. Mayaro Virus Infection in Amazonia: A Multimodel Inference Approach to Risk Factor Assessment. **PLoS Neglected Tropical Diseases**, v. 6, n. 10, 11 out. 2012.
- ABRAHAM, R. et al. Correlation of phylogenetic clade diversification and in vitro infectivity differences among Cosmopolitan genotype strains of Chikungunya virus. **Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases**, v. 37, p. 174–184, jan. 2016.
- ACOSTA-AMPUDIA, Y. et al. Mayaro: an emerging viral threat? **Emerging Microbes & Infections**, v. 7, 26 set. 2018.
- ALESHIN, A. E. et al. Structural evidence for regulation and specificity of flaviviral proteases and evolution of the Flaviviridae fold. **Protein Science: A Publication of the Protein Society**, v. 16, n. 5, p. 795–806, maio 2007.
- ALLIEGRO, M. C. Effects of dithiothreitol on protein activity unrelated to thiol-disulfide exchange: for consideration in the analysis of protein function with Cleland's reagent. **Analytical Biochemistry**, v. 282, n. 1, p. 102–106, 15 jun. 2000.
- BAKER, M. Fragment-based lead discovery grows up. **Nature Reviews. Drug Discovery**, v. 12, n. 1, p. 5–7, jan. 2013.
- BARRETO, M. L.; TEIXEIRA, M. G. Dengue in Brazil: epidemiological situation and contribution to a research agenda. **Estudos Avançados**, v. 22, n. 64, p. 53–72, dez. 2008.
- BEHNAM, M. A. M. et al. Discovery of Nanomolar Dengue and West Nile Virus Protease Inhibitors Containing a 4-Benzoyloxyphenylglycine Residue. **Journal of Medicinal Chemistry**, v. 58, n. 23, p. 9354–9370, 10 dez. 2015.
- BENKERT, P.; KÜNZLI, M.; SCHWEDE, T. QMEAN server for protein model quality estimation. **Nucleic Acids Research**, v. 37, n. Web Server issue, p. W510-514, jul. 2009.
- BHATT, S. et al. The global distribution and burden of dengue. **Nature**, v. 496, n. 7446, p. 504–507, abr. 2013.
- BOLLATI, M. et al. Structure and functionality in flavivirus NS-proteins: Perspectives for drug design. **Antiviral Research**, v. 87, n. 2, p. 125–148, ago. 2010.
- BOOTH, W. T. et al. Impact of an N-terminal Polyhistidine Tag on Protein Thermal Stability. **ACS Omega**, v. 3, n. 1, p. 760–768, 31 jan. 2018.
- BRASIL. Comissão nacional de incorporação de tecnologias no SUS. Monitoramento do horizonte tecnológico. PAXLOVID para o tratamento de COVID-19. Novembro, 2021.
- BRECHER, M.; ZHANG, J.; LI, H. The Flavivirus Protease As a Target for Drug Discovery. **Virologica Sinica**, v. 28, n. 6, p. 326–336, dez. 2013.
- CALISHER, C. H.; HIGGS, S. The Discovery of Arthropod-Specific Viruses in Hematophagous Arthropods: An Open Door to Understanding the Mechanisms of Arbovirus and Arthropod

- Evolution? **Annual Review of Entomology**, v. 63, n. 1, p. 87–103, 2018.
- CÁRDENAS, M. L.. Understanding mechanisms of enzyme co-operativity: The importance of not being at equilibrium. **Perspectives in Science**. v. 4, p. 10-16, 2015.
- CARSON, M. et al. His-tag impact on structure. **Acta Crystallographica. Section D, Biological Crystallography**, v. 63, n. Pt 3, p. 295–301, mar. 2007.
- CHANDRAMOULI, S. et al. Serotype-Specific Structural Differences in the Protease-Cofactor Complexes of the Dengue Virus Family. **Journal of Virology**, v. 84, n. 6, p. 3059–3067, 15 mar. 2010.
- CHEN, X. et al. Mechanisms of activation and inhibition of Zika virus NS2B-NS3 protease. **Cell Research**, v. 26, n. 11, p. 1260–1263, nov. 2016.
- CHIA, C. S. B.; et al. A Patent Review on SARS Coronavirus Main Protease (3CLpro) Inhibitors. **ChemMedChem** 2022, 17, e202100576.
- CHONG, H. Y. et al. Flavivirus infection—A review of immunopathogenesis, immunological response, and immunodiagnosis. **Virus Research**, v. 274, p. 197770, 1 dez. 2019.
- CHRISTO, P. P. Encephalitis by dengue virus and other arboviruses. **Arquivos De Neuro-Psiquiatria**, v. 73, n. 8, p. 641–643, ago. 2015.
- CIOTA, A. T.; KRAMER, L. D. Insights into arbovirus evolution and adaptation from experimental studies. **Viruses**, v. 2, n. 12, p. 2594–2617, 2010.
- COMPTON, J. R. et al. Mutation of Asn-475 in the Venezuelan Equine Encephalitis Virus nsP2 Cysteine Protease Leads to a Self-Inhibited State. **Biochemistry**, v. 56, n. 47, p. 6221–6230, 28 nov. 2017.
- CSTORER, A.; MÉNARD, R. [33] Catalytic mechanism in papain family of cysteine peptidases. In: **Methods in Enzymology**. Proteolytic Enzymes: Serine and Cysteine Peptidases. [s.l.] Academic Press, 1994. v. 244p. 486–500.
- DELLER, M. C.; KONG, L.; RUPP, B. Protein stability: a crystallographer's perspective. **Acta Crystallographica Section F Structural Biology Communications**, v. 72, n. 2, p. 72–95, 1 fev. 2016.
- DI CERA, E. Serine Proteases. **IUBMB life**, v. 61, n. 5, p. 510–515, maio 2009.
- ERLANSON, D. A. Introduction to fragment-based drug discovery. **Topics in Current Chemistry**, v. 317, p. 1–32, 2012.
- FAHEEM, M. et al. An insight into dengue virus proteins as potential drug/vaccine targets. **Future Virology**, v. 14, n. 10, p. 671–691, 1 out. 2019.
- FAUST, G.; STAND, A.; WEUSTER-BOTZ, D. IPTG can replace lactose in auto-induction media to enhance protein expression in batch-cultured Escherichia coli. **Engineering in Life Sciences**, v. 15, n. 8, p. 824–829, 2015.
- FERNANDEZ-GARCIA, M.-D. et al. Pathogenesis of flavivirus infections: using and abusing the

host cell. **Cell Host & Microbe**, v. 5, n. 4, p. 318–328, 23 abr. 2009.

GARCIA-CARREON, F. L.; TORO, D.; NAVARRETE, M. A. Classification of proteases without tears. **Biochemical Education**, v. 25, n. 3, p. 161–167, 1 jul. 1997.

GÉRARDIN, P. et al. Chikungunya virus-associated encephalitis: A cohort study on La Réunion Island, 2005–2009. **Neurology**, v. 86, n. 1, p. 94–102, 5 jan. 2016.

GIMENO, A. et al. The Light and Dark Sides of Virtual Screening: What Is There to Know? **International Journal of Molecular Sciences**, v. 20, n. 6, 19 mar. 2019.

GOLOVANOV, A. P. et al. A simple method for improving protein solubility and long-term stability. **Journal of the American Chemical Society**, v. 126, n. 29, p. 8933–8939, 28 jul. 2004.

GOLUBTSOV, A.; KÄÄRIÄINEN, L.; CALDENTEY, J. Characterization of the cysteine protease domain of Semliki Forest virus replicase protein nsP2 by in vitro mutagenesis. **FEBS Letters**, v. 580, n. 5, p. 1502–1508, 2006.

GONNELLI, M.; STRAMBINI, G. B. Glycerol effects on protein flexibility: a tryptophan phosphorescence study. **Biophysical Journal**, v. 65, n. 1, p. 131–137, jul. 1993.

HARTMAN, J. et al. Re-re-treatment of hepatitis C virus: Eight patients who relapsed twice after direct-acting-antiviral drugs. **World Journal of Gastroenterology**, v. 21, n. 43, p. 12430–12438, 21 nov. 2015.

HASE, T.; SUMMERS, P. L.; COHEN, W. H. A comparative study of entry modes into C6/36 cells by Semliki Forest and Japanese encephalitis viruses. **Archives of Virology**, v. 108, n. 1, p. 101–114, 1 mar. 1989.

HEDSTROM, L. Serine Protease Mechanism and Specificity. **Chemical Reviews**, v. 102, n. 12, p. 4501–4524, 1 dez. 2002.

HERNANDEZ, R.; BROWN, D. T.; PAREDES, A. Structural differences observed in arboviruses of the alphavirus and flavivirus genera. **Advances in Virology**, v. 2014, p. 259382, 2014.

HERNANDEZ, R.; LUO, T.; BROWN, D. T. Exposure to Low pH Is Not Required for Penetration of Mosquito Cells by Sindbis Virus. **Journal of Virology**, v. 75, n. 4, p. 2010–2013, 15 fev. 2001.

HIRAI, M. et al. Direct Evidence for the Effect of Glycerol on Protein Hydration and Thermal Structural Transition. **Biophysical Journal**, v. 115, n. 2, p. 313–327, 17 jul. 2018.

HU, X. et al. Kinetic, Mutational, and Structural Studies of the Venezuelan Equine Encephalitis Virus Nonstructural Protein 2 Cysteine Protease. **Biochemistry**, v. 55, n. 21, p. 3007–3019, 31 2016.

JAMES C. POWERS, * et al. **Irreversible Inhibitors of Serine, Cysteine, and Threonine Proteases**. research-article. Disponível em: <<http://pubs.acs.org/doi/abs/10.1021/cr010182v>>. Acesso em: 7 mar. 2018.

KITA, Y. et al. Contribution of the surface free energy perturbation to protein-solvent interactions. **Biochemistry**, v. 33, n. 50, p. 15178–15189, 20 dez. 1994.

- KLEIN, H. G.; ANSTEE, D. J. **Mollison's Blood Transfusion in Clinical Medicine**. [s.l.] John Wiley & Sons, 2014.
- KLOSE, D. P.; WALLACE, B. A.; JANES, R. W. 2Struc: the secondary structure server. **Bioinformatics**, v. 26, n. 20, p. 2624–2625, 15 out. 2010.
- KOKIC, G., HILLEN, H.S., TEGUNOV, D. et al. Mechanism of SARS-CoV-2 polymerase stalling by remdesivir. *Nat Commun* 12, 279 (2021). <https://doi.org/10.1038/s41467-020-20542-0>
- KOZAKOV, D. et al. The ClusPro web server for protein–protein docking. **Nature Protocols**, v. 12, n. 2, p. 255–278, fev. 2017.
- KRAEMER, M. U. et al. **The global distribution of the arbovirus vectors *Aedes aegypti* and *Ae. albopictus***. Disponível em: <<https://elifesciences.org/articles/08347/figures>>. Acesso em: 26 mar. 2020.
- KRAUSE, M.; NEUBAUER, A.; NEUBAUER, P. The fed-batch principle for the molecular biology lab: controlled nutrient diets in ready-made media improve production of recombinant proteins in *Escherichia coli*. **Microbial Cell Factories**, v. 15, n. 1, p. 110, 17 jun. 2016.
- LAKSHMI, B. et al. Ramachandran analysis of conserved glycol residues in homologous proteins of known structure. **Protein Science : A Publication of the Protein Society**, v. 23, n. 6, p. 843–850, jun. 2014.
- LAVECCHIA, A.; DI GIOVANNI, C. Virtual screening strategies in drug discovery: a critical review. **Current Medicinal Chemistry**, v. 20, n. 23, p. 2839–2860, 2013.
- LEE, H.; et al. Identification of novel small molecule inhibitors against NS2B/NS3 serine protease from Zika virus. **Antiviral Res.** V. 139, p. 49-58, 2017.
- LEI, J. et al. Crystal structure of Zika virus NS2B-NS3 protease in complex with a boronate inhibitor. **Science (New York, N.Y.)**, v. 353, n. 6298, p. 503–505, 29 jul. 2016.
- LEÓN-JUÁREZ, M. et al. Recombinant Dengue virus protein NS2B alters membrane permeability in different membrane models. **Virology Journal**, v. 13, n. 1, p. 1, 4 jan. 2016.
- LEUNG, J. Y.-S.; NG, M. M.-L.; CHU, J. J. H. Replication of Alphaviruses: A Review on the Entry Process of Alphaviruses into Cells. **Advances in Virology**, v. 2011, 2011.
- LI, H. et al. Combining in vitro and in silico Approaches to Find New Candidate Drugs Targeting the Pathological Proteins Related to the Alzheimer's Disease. **Current Neuropharmacology**, v. 16, n. 6, p. 758–768, jul. 2018.
- LI, J. et al. Functional profiling of recombinant NS3 proteases from all four serotypes of dengue virus using tetrapeptide and octapeptide substrate libraries. **The Journal of Biological Chemistry**, v. 280, n. 31, p. 28766–28774, 5 ago. 2005.
- LI, Y.; et al. Structural Insights into the Inhibition of Zika Virus NS2B-NS3 Protease by a Small-Molecule Inhibitor. **Structure**. V. 26, p. 555-564, 2018.
- LOPEZ, M. et al. Chapter 3 - Drug Discovery Methods. In: EGGER, G.; ARIMONDO, P. (Eds.). . **Drug Discovery in Cancer Epigenetics**. Boston: Academic Press, 2016. p. 63–95.

- LULLA, A. et al. Molecular determinants of substrate specificity for Semliki Forest virus nonstructural protease. **Journal of Virology**, v. 80, n. 11, p. 5413–5422, jun. 2006.
- LUO, D. et al. Crystal structure of the NS3 protease-helicase from dengue virus. **Journal of Virology**, v. 82, n. 1, p. 173–183, jan. 2008.
- LUO, D. et al. Flexibility between the protease and helicase domains of the dengue virus NS3 protein conferred by the linker region and its functional implications. **The Journal of Biological Chemistry**, v. 285, n. 24, p. 18817–18827, 11 jun. 2010.
- LV, Z.; CHU, Y.; WANG, Y. HIV protease inhibitors: a review of molecular selectivity and toxicity. **HIV AIDS (Auckl)**. 2015;7:95-104. Published 2015 Apr 8. doi:10.2147/HIV.S79956
- MARZINEK, J. K. et al. Pushing the Envelope: Dengue Viral Membrane Coaxed into Shape by Molecular Simulations. **Structure**, v. 24, n. 8, p. 1410–1420, 2 ago. 2016.
- MELTZER, E. Arboviruses and viral hemorrhagic fevers (VHF). **Infectious Disease Clinics of North America**, v. 26, n. 2, p. 479–496, jun. 2012.
- MERITS, A. et al. Proteolytic processing of Semliki Forest virus-specific non-structural polyprotein by nsP2 protease. **The Journal of General Virology**, v. 82, n. Pt 4, p. 765–773, abr. 2001.
- MEROPS - the Peptidase Database**. Disponível em:
<<https://www.ebi.ac.uk/merops/index.shtml>>. Acesso em: 2 mar. 2018.
- MESSINA, J. P. et al. Mapping global environmental suitability for Zika virus. **eLife**, v. 5, p. e15272, 19 abr. 2016.
- MISHRA, R.; BHAT, R.; SECKLER, R. Chemical chaperone-mediated protein folding: stabilization of P22 tailspike folding intermediates by glycerol. **Biological Chemistry**, v. 388, n. 8, p. 797–804, ago. 2007.
- MURRAY, C. W. et al. Fragment-Based Discovery of Potent and Selective DDR1/2 Inhibitors. **ACS medicinal chemistry letters**, v. 6, n. 7, p. 798–803, 9 jul. 2015.
- MURRAY, C. W.; VERDONK, M. L.; REES, D. C. Experiences in fragment-based drug discovery. **Trends in Pharmacological Sciences**, v. 33, n. 5, p. 224–232, maio 2012.
- NAJMANOVICH, R. et al. Side-chain flexibility in proteins upon ligand binding. **Proteins: Structure, Function, and Bioinformatics**, v. 39, n. 3, p. 261–268, 2000.
- NARWAL, M. et al. Crystal structure of chikungunya virus nsP2 cysteine protease reveals a putative flexible loop blocking its active site. **International Journal of Biological Macromolecules**, v. 116, p. 451–462, set. 2018.
- NOBLE, C. G. et al. Ligand-bound structures of the dengue virus protease reveal the active conformation. **Journal of Virology**, v. 86, n. 1, p. 438–446, jan. 2012.
- NOLLET, L. M. L. et al. **Food Biochemistry and Food Processing**. [s.l.] John Wiley & Sons, 2012.
- OKONIEWSKA, M.; TANAKA, T.; YADA, R. Y. The pepsin residue glycine-76 contributes to

active-site loop flexibility and participates in catalysis. **Biochemical Journal**, v. 349, n. Pt 1, p. 169–177, 1 jul. 2000.

OLIVEIRA, G. M. de. Caracterização estrutural e enzimática da serino protease transmembrana II humana (TMPRSS2), um alvo potencial para o tratamento do câncer de próstata. 2018

OVIEDO, C.; RODRÍGUEZ, J. EDTA: the chelating agent under environmental scrutiny. **Química Nova**, v. 26, n. 6, p. 901–905, dez. 2003.

PASTORINO, B. A. M. et al. Expression and biochemical characterization of nsP2 cysteine protease of Chikungunya virus. **Virus Research**, v. 131, n. 2, p. 293–298, fev. 2008.

PATHAK, N.; et al. Zika Virus NS3 Protease Pharmacophore Anchor Model and Drug Discovery. **Scientific reports**. 2020.

PHOO, W. W. et al. Structure of the NS2B-NS3 protease from Zika virus after self-cleavage. **Nature Communications**, v. 7, n. 1, p. 13410, 15 nov. 2016.

PRIEV, A. et al. Glycerol decreases the volume and compressibility of protein interior. **Biochemistry**, v. 35, n. 7, p. 2061–2066, 20 fev. 1996.

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

RAUSALU, K. et al. Chikungunya virus infectivity, RNA replication and non-structural polyprotein processing depend on the nsP2 protease's active site cysteine residue. **Scientific Reports**, v. 6, 15 nov. 2016.

RAWLINGS, N. D. et al. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. **Nucleic Acids Research**. Volume 46, Issue D1, 4 January 2018, Pages D624–D632, <https://doi.org/10.1093/nar/gkx1134>

RUPP, J. C. et al. Alphavirus RNA synthesis and non-structural protein functions. **The Journal of General Virology**, v. 96, n. 9, p. 2483–2500, set. 2015.

RUSSO, A. T. et al. Structural basis for substrate specificity of alphavirus nsP2 proteases. **Journal of Molecular Graphics & Modelling**, v. 29, n. 1, p. 46–53, 24 ago. 2010.

RUSSO, A. T.; WHITE, M. A.; WATOWICH, S. J. The Crystal Structure of the Venezuelan Equine Encephalitis Alphavirus nsP2 Protease. **Structure**, v. 14, p. 1449–1458, 2006.

SAISAWANG, C. et al. Full length and protease domain activity of chikungunya virus nsP2 differ from other alphavirus nsP2 proteases in recognition of small peptide substrates. **Bioscience Reports**, v. 35, n. 3, 22 abr. 2015a.

SAISAWANG, C. et al. Chikungunya nsP2 protease is not a papain-like cysteine protease and the catalytic dyad cysteine is interchangeable with a proximal serine. **Scientific Reports**, v. 5, p. 17125, 24 nov. 2015b.

SCHMID, N. et al. Definition and testing of the GROMOS force-field versions 54A7 and 54B7. **European biophysics journal: EBJ**, v. 40, n. 7, p. 843–856, jul. 2011.

- SHUKLA, D.; TROUT, B. L. Understanding the synergistic effect of arginine and glutamic acid mixtures on protein solubility. **The Journal of Physical Chemistry. B**, v. 115, n. 41, p. 11831–11839, 20 out. 2011.
- SIENCZYK, M. S. AND M. **Viral Proteases as Targets for Drug Design**. Disponível em: <<http://www.eurekaselect.com/105745/article>>. Acesso em: 6 mar. 2018.
- SILVA, H. de A.. Ferramentas virtuais contra doenças tropicais: Dengue e doença de chagas como alvo. 2012
- SMIT, J. M. et al. Flavivirus Cell Entry and Membrane Fusion. **Viruses**, v. 3, n. 2, p. 160–171, 22 fev. 2011.
- SMITH, C. R. et al. Fragment-Based Discovery of a Small Molecule Inhibitor of Bruton’s Tyrosine Kinase. **Journal of Medicinal Chemistry**, v. 58, n. 14, p. 5437–5444, 23 jul. 2015.
- STILES, K. M.; KIELIAN, M. Alphavirus entry: NRAMP leads the way. **Cell host & microbe**, v. 10, n. 2, p. 92–93, 18 ago. 2011.
- STUDIER, F. W. Protein production by auto-induction in high-density shaking cultures. **Protein Expression and Purification**, v. 41, n. 1, p. 207–234, 1 maio 2005.
- TEICH, V.; ARINELLI, R.; FAHHAM, L. Aedes aegypti e sociedade: o impacto econômico das arboviroses no Brasil. **Jornal Brasileiro de Economia da Saúde**, v. 9, n. 3, p. 267–276, dez. 2017.
- TOMLINSON, S. M.; WATOWICH, S. J. Substrate Inhibition Kinetic Model for West Nile Virus NS2B-NS3 Protease. **Biochemistry**, v. 47, n. 45, p. 11763–11770, 11 nov. 2008.
- TONG, L. Viral Proteases. **Chemical Reviews**, v. 102, n. 12, p. 4609–4626, 1 dez. 2002.
- 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, 30 jan. 2010.
- TURELL, L.; ZEIDA, A.; TRUJILLO, M. Mechanisms and consequences of protein cysteine oxidation: the role of the initial short-lived intermediates. **Essays in Biochemistry**, v. 64, n. 1, p. 55–66, 17 fev. 2020.
- ULLAH, R. et al. Activity of the Human Rhinovirus 3C Protease Studied in Various Buffers, Additives and Detergents Solutions for Recombinant Protein Production. **PLoS ONE**, v. 11, n. 4, 19 abr. 2016.
- VAGENENDE, V.; YAP, M. G. S.; TROUT, B. L. Mechanisms of protein stabilization and prevention of protein aggregation by glycerol. **Biochemistry**, v. 48, n. 46, p. 11084–11096, 24 nov. 2009.
- VASCONCELOS, P. F. et al. Inadequate management of natural ecosystem in the Brazilian Amazon region results in the emergence and reemergence of arboviruses. **Cadernos De Saude Publica**, v. 17 Suppl, p. 155–164, 2001.
- VERMA, S.; DIXIT, R.; PANDEY, K. C. Cysteine Proteases: Modes of Activation and Future Prospects as Pharmacological Targets. **Frontiers in Pharmacology**, v. 7, 25 abr. 2016.

VRIEND, G. WHAT IF: a molecular modeling and drug design program. **Journal of Molecular Graphics**, v. 8, n. 1, p. 52–56, 29, mar. 1990.

WALKER, B.; LYNAS, J. F. Strategies for the inhibition of serine proteases. **Cellular and Molecular Life Sciences CMLS**, v. 58, n. 4, p. 596–624, 1 abr. 2001.

WEBB, B.; SALI, A. Comparative Protein Structure Modeling Using MODELLER. **Current Protocols in Bioinformatics**, v. 54, p. 5.6.1-5.6.37, 20 jun. 2016.

WEBER, I.T.; WANG, Y.F.; HARRISON, R.W. HIV Protease: Historical Perspective and Current Research. **Viruses**. 2021 May 6;13(5):839. doi: 10.3390/v13050839. PMID: 34066370; PMCID: PMC8148205.

WENSING, A. M. J.; VAN MAARSEVEEN, N. M.; NIJHUIS, M. Fifteen years of HIV Protease Inhibitors: raising the barrier to resistance. **Antiviral Research**, v. 85, n. 1, p. 59–74, jan. 2010.

WHITLEY, R. Molnupiravir. A Step toward Orally Bioavailable Therapies for Covid-19. **The new england journal of medicine**. Editorial. December 16, 2021. DOI: 10.1056/NEJMe2117814

YAO, Y. et al. Discovery, X-ray Crystallography and Antiviral Activity of Allosteric Inhibitors of Flavivirus NS2B-NS3 Protease. **Journal of the American Chemical Society**, v. 141, n. 17, p. 6832–6836, 01 2019.

YILDIZ, M. et al. Allosteric Inhibition of the NS2B-NS3 Protease from Dengue Virus. **ACS Chemical Biology**, v. 8, n. 12, p. 2744–2752, 20 dez. 2013.

ZHANG, D.; TÖZSÉR, J.; WAUGH, D. S. Molecular cloning, overproduction, purification, and biochemical characterization of the p39 nsp2 protease domains encoded by three alphaviruses. **Protein expression and purification**, v. 64, n. 1, p. 89–97, mar. 2009.

ANEXOS

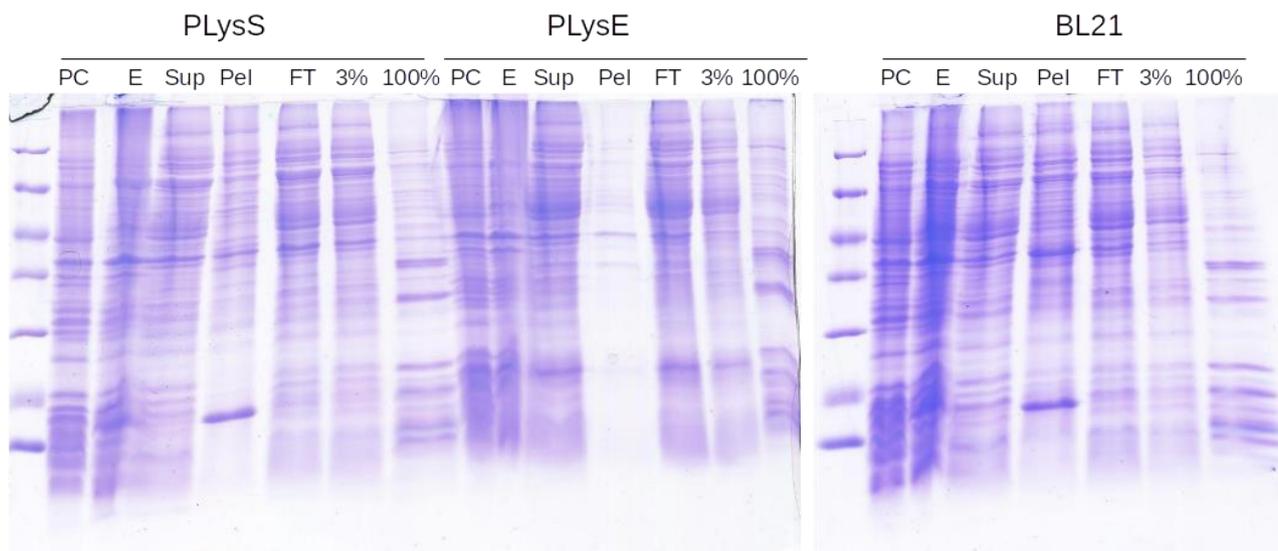


Figura 1- Gel SDS-PAGE 15% da protease nsP2 de MAYV da expressão a 25°C, 1 mM IPTG nas cepas PLYS, PLYSE e BL21. PC se refere ao pré-inóculo, E ao final da expressão, Sup ao sobrenadante, Pel a fração insolúvel, FT a amostra que não interagiu com a coluna de níquel, 3% as amostras eluídas com 15 mM de imidazol e 100% com as amostras eluídas com 500 mM de imidazol.

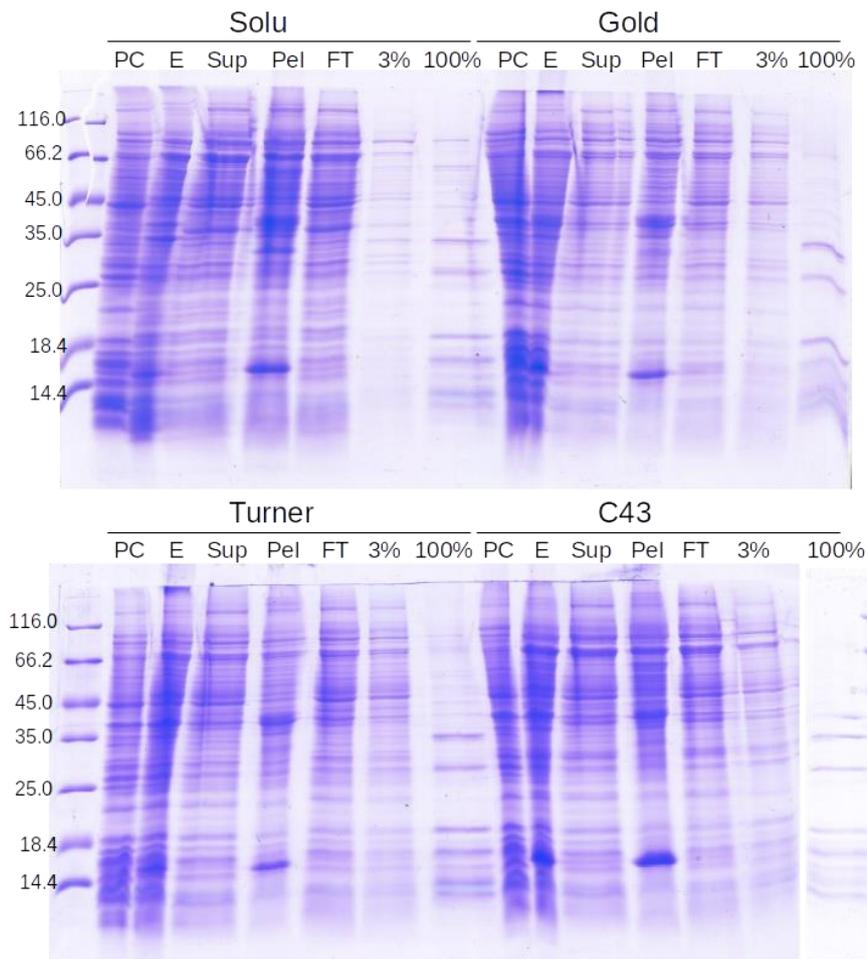


Figura 2- Gel SDS-PAGE 15% da protease nsP2 de MAYV da expressão a 25°C, 1 mM IPTG nas cepas Solu, Gold, Turner e C43. PC se refere ao pré-inóculo, E ao final da expressão, Sup ao sobrenadante, Pel a fração insolúvel, FT a amostra que não interagiu com a coluna de níquel, 3% as amostras eluídas com 15 mM de imidazol e 100% com as amostras eluídas com 500 mM de imidazol.

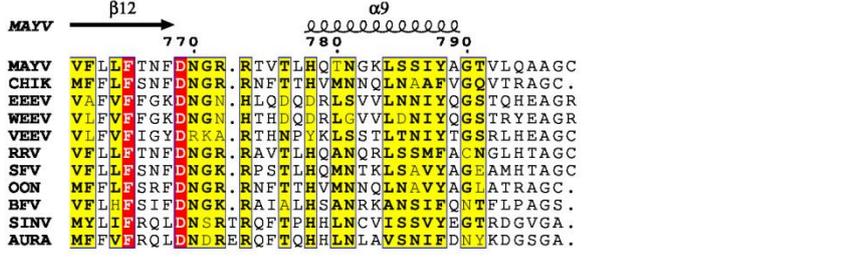
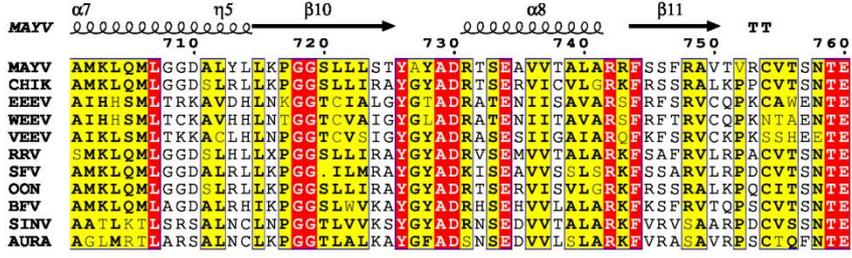
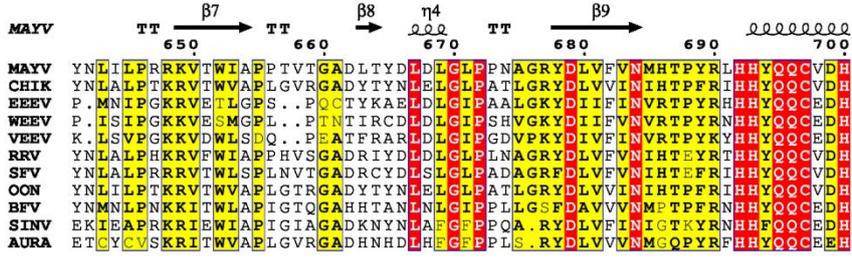
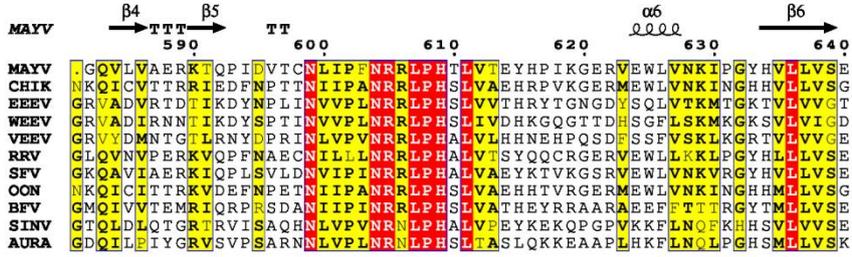
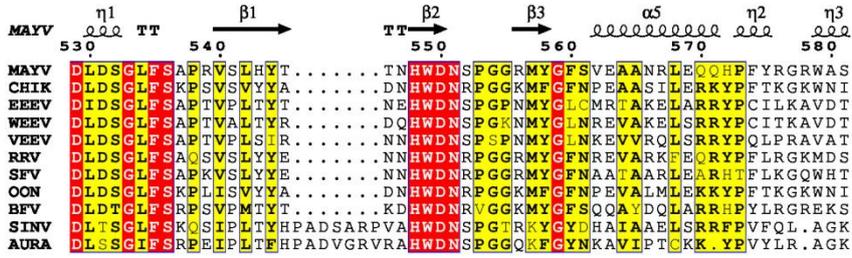
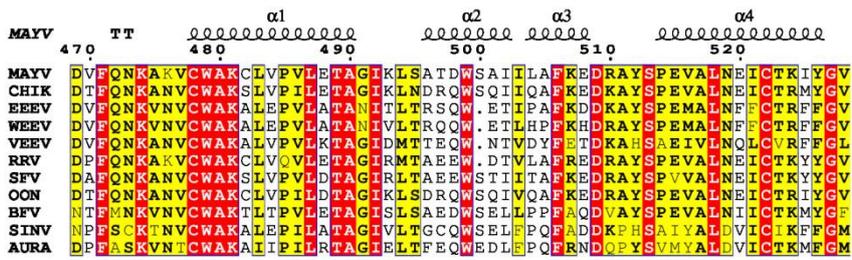


Figura 4- Alinhamento múltiplo de nsP2 de Alphavirus. Em cima é possível ver a estrutura

secundária baseada no modelo por homologia calculado para a protease MAYV. As marcações em vermelho mostram os resíduos conservados em todas as proteínas analisadas e amarelo as altamente conservadas. Os vírus usados no alinhamento vírus da encefalite equina do leste (EEEV), vírus encefalite equina do oeste (WEEV), vírus da encefalite equina venezuelana (VEEV), Sindbis vírus (SIN), Aura vírus (AURA), o'nyong-nyong vírus (ONN), chikungunya vírus (CHIK), vírus da floresta Semliki (SFV), Rio Negro vírus (RRV), Mayaro vírus (MAYV) e vírus da floresta Barmah (BFV).

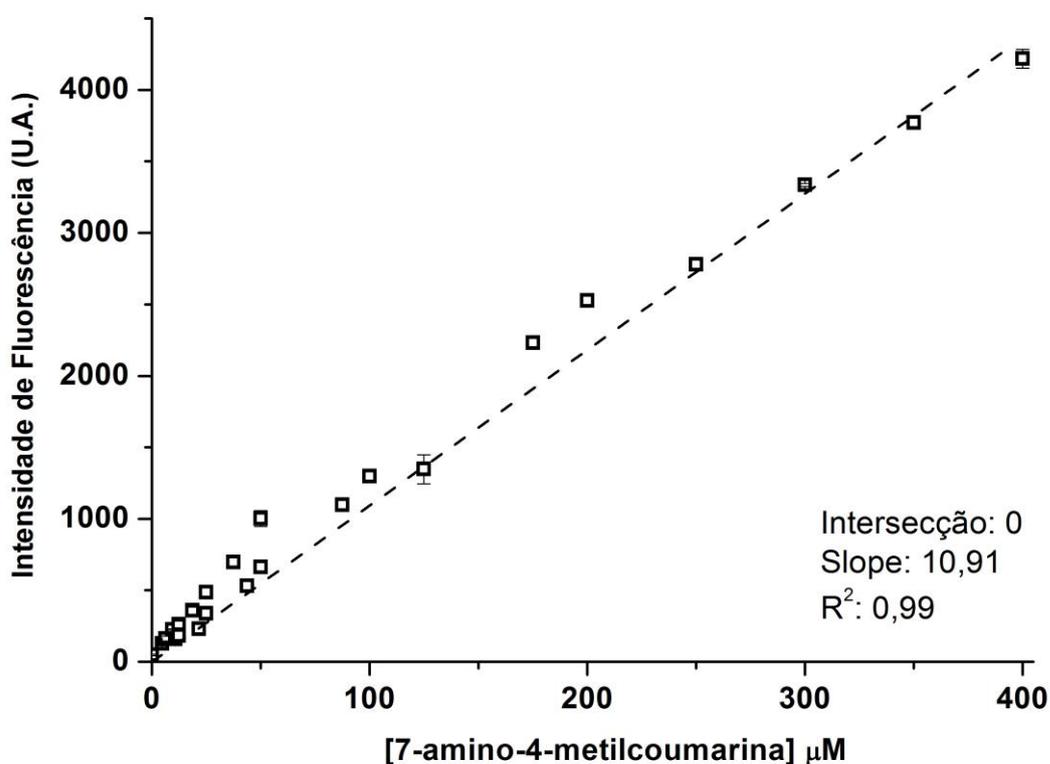


Figura 5- Curva de calibração da emissão de fluorescência durante a formação do substrato AMC (7-amino-4-metilcoumarina) (OLIVEIRA, 2018).

SCIENTIFIC REPORTS

OPEN

Self-homodimerization of an actinoporin by disulfide bridging reveals implications for their structure and pore formation

Received: 4 January 2018

Accepted: 4 April 2018

Published online: 26 April 2018

Aisel Valle^{1,2}, Luis Benito Pérez-Socas², Liem Canet², Yadira de la Patria Hervis², German de Armas-Guitart², Diogo Martins-de-Sa², Jônatas Cunha Barbosa Lima¹, Adolfo Carlos Barros Souza¹, João Alexandre Ribeiro Gonçalves Barbosa¹, Sonia Maria de Freitas² & Isabel Fabiola Pazos²

The Trp111 to Cys mutant of sticholysin I, an actinoporin from *Stichodactyla helianthus* sea anemone, forms a homodimer via a disulfide bridge. The purified dimer is 193 times less hemolytic than the monomer. Ultra centrifugation, dynamic light scattering and size-exclusion chromatography demonstrate that monomers and dimers are the only independent oligomeric states encountered. Indeed, circular dichroism and fluorescence spectroscopies showed that Trp/Tyr residues participate in homodimerization and that the dimer is less thermostable than the monomer. A homodimer three-dimensional model was constructed and indicates that Trp147/Tyr137 are at the homodimer interface. Spectroscopy results validated the 3D-model and assigned 85° to the disulfide bridge dihedral angle responsible for dimerization. The homodimer model suggests that alterations in the membrane/carbohydrate-binding sites in one of the monomers, as result of dimerization, could explain the decrease in the homodimer ability to form pores.

Stichodactyla helianthus is an abundant anemone in the Caribbean Sea and in 1979 cytolytic activity due to transmembrane pore formation originated from toxins partially purified from the same organism was first demonstrated^{1,2}. In 1988, the purification of four hemolytic polypeptides of 17–20 kDa were first reported³, and in 2001 were purified two monomeric small toxins (20 kDa) with pore-forming activity, named sticholysin I (StI) and sticholysin II (StII)⁴. StI and StII have been functionally and structurally characterized^{4–6} and both toxins belong to the actinoporin protein family due to their purification from the order *Actiniaria* and its ability to form pores in membranes⁷. Actinoporins show high affinity for sphingomyelin (SM)-containing membranes^{8–15} and are classified as α -type pore-forming proteins (PEP) based on their mode of membrane integration through α -helical elements¹⁴. The actinoporin structure presents a typical β -sandwich fold flanked on each side by one short α -helix, where the N-terminal helix α_1 is amphipathic and participates in the formation of the pore walls^{15–17}. These toxins lack of post-translational modifications and have been successfully expressed in *E. coli* by different authors¹⁸. In addition, five phosphorylcholine (POC)-binding sites formed by conserved residues within the actinoporin family were identified, suggesting a multivalence binding mode during interaction with membrane lipids^{15,17}. Recently, it has been demonstrated that Fragaceatoxin C (FraC), an actinoporin from *Actinia fragacea*, employs one of these lipid-binding sites for recognizing carbohydrates¹⁹. Many other physicochemical and structural aspects of the actinoporins such as, soluble structures, phospholipids-binding sites, multistep pore-formation mechanism, experimental approaches and theoretical tools for structure–function relationship studies, and the pore models toroidal, 9-mer conical α -helical-bundle and 8-mer conical hybrid α -helical-bundle/lipid have been described and reviewed recently^{14,18,20–24}.

¹Laboratory of Molecular Biophysics, Institute of Biological Sciences, University of Brasília (IB-CEL/UnB), Campus Darcy Ribeiro, Asa Norte, Brasília, DF-70910-900, Brazil. ²Biochemistry Department and Center for Protein Studies, Faculty of Biology, Havana, University (Bq-CEP/UH), Universidad de la Habana, Calle 25 No. 455, Plaza de la Revolución, La Habana, CP-10 400, Cuba. Correspondence and requests for materials should be addressed to A.V. (email: aiselvalle@gmail.com)



UNIVERSIDAD
DE LA HABANA
FACULTAD DE BIOLOGÍA

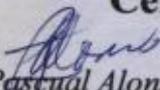
PREMIO DE INVESTIGACION 2018

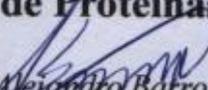
ARTICULO CIENTÍFICO DISTINGUIDO

A: Self-homodimerization of an actinoporin by disulfide bridging reveals implications for their structure and pore formation. Scientific Reports 8:6614

Aisel Valle, Luis Benito Pérez-Socas, Liem Canet, Yadira de la Patria Hervis, German de Armas-Guitart, Diogo Martins-de-Sa, Jônatas Cunha Barbosa Lima, Adolfo Carlos Barros Souza, João Alexandre Ribeiro Gonçalves Barbosa, Sonia Maria de Freitas, Isabel Fabiola Pazos.

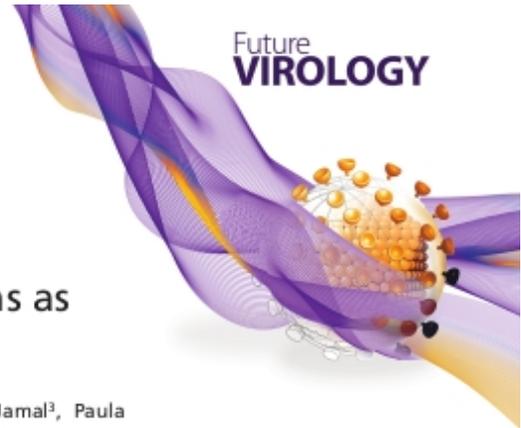
Centro de Estudios de Proteínas


Dra. Isel Pascual Alonso
Presidente Consejo Científico
Facultad de Biología


Dr. Alejandro Barro Cañamero
Vicedecano de Investigaciones
Facultad de Biología




Dra. Maday Alonso del Rivero Antigua
Decana
Facultad de Biología



An insight into dengue virus proteins as potential drug/vaccine targets

Muhammad Faheem^{*,1,2}, Jônatas Cunha Barbosa Lima¹, Syed Babar Jamal³, Paula Andreia Silva² & João Alexandre Ribeiro Gonçalves Barbosa^{*,1,2}

¹Laboratory of Biophysics, Department of Cellular Biology, University of Brasilia, Brasilia-DF 70910-900, Brazil

²Post-graduate program of Genomics Sciences & Biotechnology, Catholic University of Brasilia, Brasilia-DF 70790-160, Brazil

³Department of Biological Sciences, National University of Medical Sciences, The Mall road, Rawalpindi, Punjab 46000, Pakistan

*Author for correspondence: Tel.: +92 (31 5) 9016588; faheem08@live.com

**Author for correspondence: Tel.: +55 (61) 991010205; jpao.argb@gmail.com

Dengue virus (DENV) is an arbovirus that belongs to family *flaviviridae*. Its genome is composed of a single stranded RNA molecule that encodes a single polyprotein. The polyprotein is processed by viral and cellular proteases to generate ten viral proteins. There are four antigenically distinct serotypes of DENV (DENV1, DENV2, DENV3 and DENV4), which are genetically related. Although protein variability is a major problem in dengue treatment, the functional and structural studies of individual proteins are equally important in treatment development. The data accumulated on dengue proteins are significant to provide detailed understanding of viral infection, replication, host-immune evasion and pathogenesis. In this review, we summarized the detailed current knowledge about DENV proteins.

First draft submitted: 21 August 2019; Accepted for publication: 31 October 2019; Published online: 17 December 2019

Keywords: dengue virus • flaviviridae • polyprotein • RNA • serotypes

Dengue virus (DENV) is a mosquito-borne viral pathogen that is transmitted by infected *Aedes* mosquitoes, primarily *Aedes aegypti*, a day time biting highly urbanized mosquito that breeds in stored water [1], responsible for causing dengue fever. Dengue is one the most important human arboviruses disease in terms of morbidity and mortality [1,2]. Currently 2.5 billion people are estimated to be at risk of dengue infection. Until 2015, around 3.2 million cases were reported annually [2,3]. Recently, it is estimated that over 390 million people are infected annually with one of the four serotypes of DENV [4]. DENV primary infection is mainly associated with an asymptomatic or moderate disease, accompanying with high fever, headache, myalgia and arthralgia, while severe disease conditions arise with secondary infection. It includes dengue hemorrhagic fever (DHF) and dengue shock syndrome (DSS) that can be fatal without supportive medical treatment [5]. Till the date, no effective antiviral agents or vaccines to DENV are available to control the infection. In order to develop new antiviral strategies against this life threatening disease, detailed information about the functions of DENV proteins is essential. Thus, in this article we have reviewed functions and structural features of DENV proteins.

There are four serotypes of DENV which are antigenically variant and each one elicit serotype specific immunity [6]. DENV also upregulates the expression of MHC class I molecules [7], which inhibit the action of innate immunity, specifically NK cells. DENV evolves and escapes immune system responses due to mutations in its RNA genome during the replication [8,9]. It is a member of the family *Flaviviridae*, genus *flavivirus*, presenting an enveloped icosahedral symmetry capsid with a diameter of 500 Å. The viral genome is a non-segmented positive-stranded RNA of approximately 11 kb in length [10–12]. Upon cell infection, the viral RNA is released in the cytoplasm, the 5'UTR of the positive strand RNA is directed to the ribosome and polyprotein is translated from single open reading frame (ORF) [13]. This polypeptide protein originates ten viral proteins. The polypeptide is co- and post-translationaly processed by viral serine protease and host signalases into three structural C, prM and E and seven non-structural proteins in order of C-prM-E-NS1-NS2A-NS2B-NS3-NS4A-NS4B-NS5 shown in Figure 1 [14]. Only when the proteins are cleaved and released in mature form are they able perform their respective functions (Table 1).

