

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA DE SÃO CARLOS

GABRIELA DIAS NOSKE

Caracterização estrutural e busca por inibidores das proteases dos vírus
emergentes: vírus da Febre Amarela e SARS-CoV-2

São Carlos
2023

GABRIELA DIAS NOSKE

Caracterização estrutural e busca por inibidores das proteases dos vírus emergentes: vírus da Febre Amarela e SARS-CoV-2

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

Opção: Física Biomolecular
Área de concentração: Física Biomolecular
Orientador: Prof. Dr. Glaucius Oliva
Co-orientador: Dr. Andre Schützer de Godoy

Versão corrigida
(versão original disponível na Unidade que aloja o Programa)

São Carlos
2023

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

Noske, Gabriela Dias

Caracterização estrutural e busca por inibidores das proteases dos vírus emergentes: vírus da Febre Amarela e SARS-CoV-2 / Gabriela Dias Noske; orientador Glaucius Oliva; co-orientador Andre Schützer de Godoy - versão corrigida -- São Carlos, 2023.

184 p.

Tese (Doutorado - Programa de Pós-Graduação em Física Biomolecular) -- Instituto de Física de São Carlos, Universidade de São Paulo, 2023.

1. Protease. 2. Febre amarela. 3. SARS-CoV-2. I. Oliva, Glaucius, orient. II. Godoy, Andre Schützer de, co-orient. III. Título.

*Dedico este trabalho às três das mulheres mais fortes e exemplos de perseverança que
tive em minha vida (in memoriam), minha amiga Iolanda e minhas avós Lourdes e
Dirce, com todo o meu amor.*

AGRADECIMENTOS

Aos meus pais, obrigada por tudo, em especial por todo o amor e suporte que me deram até hoje.

Ao Prof. Dr. Glaucius Oliva pela orientação, que me deu esta oportunidade ainda durante minha graduação. Agradeço a todos os ensinamentos e conversas, que muito contribuíram com minha formação profissional e pessoal.

Ao Dr. Andre S. Godoy, agradeço por ter me formado a cientista que sou hoje, não somente por ter me ensinado muito dentro do laboratório, mas também pela amizade, carinho e incentivo ao longo destes anos. Agradeço por todo o apoio, ensinamentos e convivência.

Aos meus amigos e colegas de laboratório, em especial Victor, Aline, Rafaela, Igor, Camila, Guilherme. Obrigada por terem que tornado o dia a dia no laboratório mais leve, obrigada pelas conversas, discussões e amizade ao longo destes anos de pós-graduação.

Aos meus amigos Carol, Julia e Higor. Agradeço por estarem sempre comigo.

Ao prof. Dr. Kay Grünewald, os colegas do CSSB (Desy, Hamburgo – Alemanha), e todos os funcionários da *facility* de *cryo-EM*. Obrigada por todos os ensinamentos de microscopia e por terem me recebido em seu laboratório.

Ao CIBFar, Instituto de Física de São Carlos e Universidade de São Paulo, agradeço à oportunidade da realização deste doutorado direto.

Ao Diamond Light Source e CNPEM-Sirius, a equipe da linha MANACA, obrigada pelas horas cedidas para realização deste trabalho, além de todo apoio durante a realização das coletas.

Aos funcionários e técnicos de laboratório do IFSC - campus 2, Humberto, Andressa, Bel, Susana e Lívia, agradeço por todo apoio e assistência na realização dos experimentos.

À Fundação de Amparo à Pesquisa do Estado de São Paulo, pela concessão da bolsa de doutorado direto, pelo apoio financeiro para a realização desta pesquisa (2018/25600-4 e CEPID 2013/07600-3).

O presente trabalho também foi realizado com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES/PROEX - 88887.338499/2019-00).

"Para ser grande, sê inteiro: nada
Teu exagera ou exclui.
Sê todo em cada coisa. Põe quanto és
No mínimo que fazes.
Assim em cada lago a lua toda
Brilha, porque alta vive."

Ricardo Reis (Fernando Pessoa)

RESUMO

NOSKE, G. D. **Caracterização estrutural e busca por inibidores das proteases dos vírus emergentes: vírus da Febre Amarela e SARS-CoV-2.** 184 p. Tese (Doutorado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2023.

Os recentes avanços em pesquisa possibilitaram uma maior compreensão do mecanismo de diversas doenças virais e auxiliaram no desenvolvimento de vacinas e medicamentos. No entanto, as doenças virais ainda representam uma das maiores preocupações de saúde pública e econômicas mundiais, o que evidencia a necessidade constante do desenvolvimento de novos antivirais. Dentre os vírus ainda sem fármacos específicos para seu tratamento, estão os pertencentes ao gênero flavivírus, que inclui os vírus causadores da dengue, zika e febre amarela (YFV). O vírus da febre amarela (YFV) tem genoma composto por uma única fita de RNA que codifica uma única poli proteína contendo três proteínas estruturais e sete não estruturais. A NS3 tem dois domínios: um domínio protease e um domínio helicase/NTP-ase. A NS3 protease age juntamente com uma outra proteína não estrutural como cofator, a NS2B, que auxilia o enovelamento correto da NS3^{pro} e permite que tenha uma forma ativa. O complexo NS2B-NS3^{pro} auxilia na clivagem da poli proteína imatura, liberando as proteínas formadoras do complexo de replicação viral. Considerando a importância deste complexo no ciclo de replicação viral, é evidente que ele representa um importante alvo no planejamento de candidatos antivirais. Sendo assim, o objetivo deste projeto consistiu em determinar a estrutura cristalográfica da enzima NS2B-NS3^{pro} e utilizá-la na busca por inibidores. A sequência codificante da proteína do vírus foi clonada, expressa e purificada. Também foi possível obter cristais da enzima, utilizados para a resolução da estrutura da enzima, resultando num modelo final a 2.9 Å de resolução. Além disso, avaliamos a atividade da proteína utilizando um ensaio baseado em um produto fluorescente. Foi possível determinar as constantes cinéticas da mesma, além de realizar a triagem de inibidores. Dos aproximadamente 1000 compostos testados, 18 *hits* foram encontrados, 6 deles apresentando IC₅₀ com valores inferiores à 1.0 µM. Os ligantes mais promissores tiveram atividade antiviral testada em células contendo o replicon de YFV, no qual 3 compostos apresentaram EC₅₀ em baixo micromolar. Em conclusão, nós conseguimos avançar com a clonagem, expressão e purificação do complexo NS2B-NS3 protease do YFV. Além disso, obtivemos cristais da proteína, que possibilitaram a resolução de sua estrutura cristalográfica. A enzima demonstrou elevada atividade proteolítica contra o peptídeo sintético fluorescente, que foi de suma importância para a triagem de inibidores.

Adicionalmente, diante da pandemia causada pelo novo coronavírus, o SARS-CoV-2, e como parte do projeto colaborativo conduzido pelo CIBFar/CEPID como esforço emergencial para o desenvolvimento de antivirais contra o SARS-CoV-2, estudos também foram desenvolvidos com as proteases deste vírus. Uma das proteases do vírus, a M^{pro}, é uma proteína dimérica amplamente estudada e explorada para o desenvolvimento de antivirais. No entanto, os detalhes sobre seu processo de auto maturação permaneciam desconhecidos. Neste trabalho, a enzima M^{pro} foi expressa e purificada, em três construções diferentes (IMT-M^{pro}, M^{pro} e C145S M^{pro}). Ambas as construções foram cristalizadas e tiveram estruturas cristalográficas resolvidas, assim como foram caracterizadas em solução. Além disso, a enzima C145S M^{pro} teve estrutura determinada a 3.5Å, em complexo com o peptídeo nativo do N-terminal, utilizando crio-microscopia eletrônica. Analisamos ainda a influência de diferentes inibidores no processo de maturação da enzima. Por fim, pudemos inferir que a clivagem do N-terminal não é crítica para a dimerização da enzima, mas sim, as mudanças conformacionais que são ocasionadas pelo ajuste induzido após ligação covalente do substrato. A elucidação das mudanças funcionais e estruturais que ocorrem durante esse processo de maturação obtidas neste trabalho fornecem informações importantes para a compreensão deste mecanismo e para a proposição de inibidores específicos, que tenham como alvo etapas intermediárias do processo de maturação da enzima. Além disso, este trabalho nos ajuda a ter um maior entendimento sobre as proteases dos vírus da febre amarela e SARS-CoV-2, e habilita novas estratégias para o desenvolvimento de novos antivirais.

Palavras-chave: Protease. Febre amarela. SARS-CoV-2.

ABSTRACT

NOSKE, G. D. **Structural characterization and discovery of antiviral candidates targeting the viral proteases from: yellow Fever virus and SARS-CoV-2.** 184p Thesis (Doctor in Science) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2023.

The recent advances in research have enabled a greater understanding of the mechanisms of several viral diseases that supported the development of vaccines and antiviral drugs. However, viral diseases still represent one of the major public health and economic concerns worldwide, highlighting the constant need for the development of new antivirals. Among the viruses that still lack specific drugs for their treatment are those belonging to the flavivirus genus, which includes the viruses that cause dengue, zika and yellow fever. The yellow fever virus (YFV) has a genome composed of a single RNA strand that encodes a polyprotein containing three structural and seven non-structural proteins. NS3 has two domains: a protease domain and a helicase/NTPase domain. The NS3 protease acts together with another non-structural protein as a cofactor, NS2B, which helps the correct folding of NS3^{pro} and allows it to assume its active form. The NS2B-NS3^{pro} complex assists in the cleavage of the immature polyprotein, releasing the individual proteins that form the viral replication complex. Considering the importance of this complex in the viral replication, it is clear that it represents an important target in the design of antiviral candidates. Therefore, the objective of this project was to determine the crystallographic structure of the NS2B-NS3^{pro} enzyme and use it in the search for inhibitors. The coding sequence was cloned, expressed and the protein purified. It was also possible to obtain crystals of the enzyme, used to solve the 3D structure, resulting in a final model at 2.9 Å resolution. Furthermore, we evaluated the activity of the protein using a fluorescence-based assay. It was possible to determine its kinetic constants, in addition to the screening for inhibitors. Among approximately 1000 compounds screened, 18 hits were found, 6 of them with IC₅₀ values below 1.0 μM. The most promising ligands had antiviral activity tested in cells containing the YFV replicon, in which 3 compounds presented EC₅₀ at low micromolar. In conclusion, we were able to move forward with the cloning, expression and purification of the YFV NS2B-NS3 protease complex. Furthermore, we obtained crystals of the protein, which made it possible to resolve its crystallographic structure. The enzyme demonstrated high proteolytic activity against the synthetic fluorescent peptide, which was of paramount importance for screening inhibitors. Additionally, due to the pandemic caused by the new coronavirus, SARS-CoV-2, and as part of the collaborative project led by CIBFar/CEPID as an emergency effort to develop antivirals against SARS-CoV-2, studies were also developed with the coronaviral proteases. One of the virus proteases, M^{pro}, is a dimeric protein widely studied and explored for the development of antivirals. However, details about its self-maturation process remained unknown. In this work, M^{pro} was expressed and purified in three different constructs (IMT-M^{pro}, M^{pro} and C145S M^{pro}). Both constructs were crystallized and had their crystallographic structures solved, as well as characterized in solution. Furthermore, the structure of the C145S M^{pro} enzyme was determined at 3.5Å, in complex with the endogenous N-terminal peptide, using cryo-electron microscopy. We also analyzed the influence of different inhibitors on the enzyme maturation process. Finally, we were able to infer that the N-terminal cleavage is not critical for the dimerization of the enzyme, but rather, the conformational changes are caused by the adjustment induced after covalent binding of the substrate. The elucidation of the functional and structural changes that occur during the maturation process obtained in this work provides important information for understanding this mechanism and for proposing specific inhibitors, which target intermediate stages of the enzyme maturation process. Furthermore, this work helps us gain a greater understanding of the proteases of yellow fever and SARS-CoV-2 viruses and supports the development of new antivirals.

Keywords: Protease. Yellow fever. SARS-CoV-2.

LISTA DE FIGURAS

- Figura 1** - Ciclos de transmissão do vírus da febre amarela. Fonte: Elaborado pela autora com os dados de CDC- Yellow Fever transmission..... 40
- Figura 2** - Estrutura e expressão genômica dos flavivírus. 44
- Figura 3** - Estrutura do complexo NS2B-NS3 protease de dengue tipo 2. Podemos ver o enovelamento do complexo NS2B (azul) com a NS3 (vermelho) na conformação aberta em A (PDB ID 2FOM), e na conformação fechada em B (PDB ID 3U1I). Em ambas temos a visão da tríade catalítica, com os resíduos His51, Asp75 e Ser135 representados em esferas (cinza). 45
- Figura 4** - Representação esquemática do mecanismo de catálise das proteases de Flavivírus. As numerações de 1-6 representam as etapas da reação 46
- Figura 5** - A) Alinhamento das sequências de aminoácidos da proteína NS3 de diferentes linhagens de YFV. Os resíduos mutantes estão destacados em amarelo. B) Gel agarose 1% mostrando o cDNA contendo a sequência codificante das proteínas NS2B e NS3 (3110 pb)..... 51
- Figura 6** - Representação esquemática das construções da proteína NS2B-NS3 protease. A porção transmembrana do cofator NS2B está colorida em laranja e a porção solúvel e essencial para o enovelamento da NS3pro está colorida em verde. 52
- Figura 7** - Representação esquemática das etapas de clonagem utilizadas. A) Clonagem pelo método LIC. B) Clonagem pelo método de Gibson Assembly..... 53
- Figura 8** - Representação esquemática do ensaio enzimático da proteína YFV NS2B-NS3^{pro}. O substrato Bz-nKRR-AMC é clivado pela enzima liberando AMC, cuja fluorescência pode ser monitorada nos comprimentos de onda apresentados. 60
- Figura 9** - Predições servidor TMHMM para as construções da enzima NS2B-NS3 protease de YFV..... 63
- Figura 10** - A) Gel agarose 1% contendo os fragmentos obtidos para a construção com linker. Em M vemos o marcador molecular (em pb), nas amostras de 1-6 vemos cada uma das colônias testadas. As amostras 1-5 contém a sequência com linker e a amostra 6 contém a sequência denominada de construção inicial (NS2B-NS3^{pro} linker = 684 pb; NS2B-NS3^{pro} construção inicial = 774 pb). B) Mapa do plasmídeo pETSUMO/LIC contendo a construção NS2B-NS3^{pro} linked 64
- Figura 11** - A) Western Blot contendo a fração solúvel de diversas linhagens de E. coli para o teste de expressão realizado em dois meios diferentes (LB e TB). A proteína é detectada pela cauda de histidina e aparece em rosa, as setas vermelhas indicam as linhagens com maior taxa de expressão (NS2B-NS3^{pro} + SUMO = 38 kDa). B) SDS-PAGE 15% contendo ambas as frações solúveis e insolúveis obtidas para a linhagem de E. coli Rosetta 2, nos dois meios diferentes (LB e TB)..... 65

- Figura 12** - Cromatogramas contendo os resultados do sequenciamento dos plasmídeos contendo a sequência codificante das enzimas NS2B-NS3^{pro} mutantes D88E, K121R e mutante duplo..... 66
- Figura 13** - A) SDS-PAGE 15% contendo as amostras obtidas nas cromatografias de afinidade e clivagem para a enzima NS2B-NS3^{pro} YFV WT. Em M, vemos o marcador de peso molecular, seguido respectivamente pelas frações obtidas na eluição da primeira afinidade, após clivagem pela TEV protease, proteína já clivada obtida no FT da segunda cromatografia de afinidade e a amostra eluída da segunda afinidade, contendo contaminantes, além de proteína não clivada, TEV protease e a cauda 6-His-SUMO. Ambas as proteínas antes e após clivagem estão indicadas pelas setas, com seu respectivo peso molecular. B) Cromatograma obtido na primeira etapa de cromatografia de afinidade para a enzima NS2B-NS3^{pro} YFV WT. Em vermelho, está representado a absorvância em 280 nm e em azul a concentração do tampão B (contendo 300 mM de imidazol). 67
- Figura 14** - Resultados da purificação das proteínas NS2B-NS3^{pro} linked. A) Cromatogramas da etapa de exclusão molecular para cada uma das enzimas. B) SDS-PAGE 12.5% mostrando a pureza final de cada amostra ao final da purificação. C) Curva de calibração da coluna Superdex 75 16/60 D) Tabela contendo os volumes de retenção de cada uma das proteínas, suas respectivas massas moleculares calculadas e estado oligomérico. 68
- Figura 15** - Cristais da enzima NS2B-NS3^{pro}. A) Cristais da proteína inicialmente obtidos (condição G01 Index HT) sob luz UV. B) Cristais da proteína mutante K121R na condição B10 PACT (0.1 M MES pH 6.0, 0.2 M MgCl₂, 20% PEG 6000). C) Cristais da proteína mutante D88E na condição B10 PACT. D) Cristais otimizados da condição D9 PACT (0.1 M Tris pH 8.5, 0.1 M LiCl, 25% PEG 6000). E-G) Cristais otimizados da condição G1 Index HT (0.1 M NaCl, 0.1 M Tris pH 8.5, 25% PEG 3350) contendo respectivamente 23%, 27% e 29% PEG 3350. H-I) Cristais otimizados da condição G1 Index HT (0.1 M NaCl, 0.1 M Tris pH 8.5, 25% PEG 3350) antes (H) e após (I) utilização de microseeding..... 69
- Figura 16** - Padrão de difração obtido para um dos cristais de YFV da condição contendo 0.1 M Tris pH 8.5, 0.1 M LiCl, 25% PEG 6000, obtidos respectivamente na linha de luz MX2, no LNLS, Campinas, Brasil (A) e na linha I24, no Diamond. 70
- Figura 17** - Estrutura cristalográfica da enzima NS2B-NS3^{pro} de YFV. A) Representação esquemática da estrutura secundária da proteína, indicando cada um dos domínios. As fitas β estão representadas por setas e a hélice α_3 por cilindro. Visão geral da estrutura cristalográfica da enzima NS2B-NS3 protease de YFV. A estrutura está representada em modo cartoon com o cofator NS2B colorido em laranja e o domínio NS3 protease em verde. A superfície de ambos os domínios está representada nas suas respectivas cores. O quadro pontilhado destaca o sítio ativo da enzima. B) Visão geral da estrutura cristalográfica da enzima NS2B-NS3 protease de YFV rotacionada em 180°. C) Visão ampliada do sítio ativo da enzima, mostrando os resíduos catalíticos His53, Asp77 e Ser138, representados como sticks. D) Representação das cargas da superfície da enzima NS2B-

NS3^{pro} com destaque para o bolso de reconhecimento de substrato S1' e S1-S4..... 74

- Figura 18** - Diferentes conformações do complexo NS2B-NS3^{pro} de Flavivirus. Em A-D, estruturas em conformação aberta (open conformation), respectivamente para os vírus DENV2, ZIKV, WNV e MVEV. Em E-H, estruturas em conformação fechada (closed conformation). Para as estruturas de A-H, o domínio protease está representado como superfície, juntamente com o traço da cadeia alpha ambos coloridos em cinza. Já o cofator NS2B, está representado como cartoon, em diferentes cores para cada estrutura. I) Superposição dos complexos NS2B-NS3^{pro} de YFV (closed) e DENV2 (open). Os domínios protease estão coloridos de acordo com o RMSD (azul para baixo RMSD e vermelho para alto) e o cofator NS2B está colorido em amarelo e azul para DENV2 e YFV respectivamente. Ambas as estruturas estão representadas na forma de cartoon. Em destaque vemos o loop que determina o nome dado a conformação. J) Superposição dos complexos NS2B-NS3^{pro} de YFV (closed) e DENV2 (open), ambas representadas como traço da cadeia alpha. 75
- Figura 19** - A) Contatos cristalinos observados nos cristais da enzima NS2B-NS3 protease de YFV (PDBid 6URV). As moléculas formadoras do dímero cristalográfico estão coloridas em verde e laranja, representadas em cartoon. Os quadros vermelhos contêm uma visão ampliada das regiões de contato do dímero cristalográfico, na qual os resíduos diretamente envolvidos estão coloridos em amarelo, e representados como sticks. B) Dímero cristalográfico observado para a NS2B-NS3 protease do vírus Zika. C) Estrutura da enzima NS2B-NS3^{pro} de YFV representada de acordo com os B-fatores de cada resíduo..... 77
- Figura 20** - Ensaio de atividade da enzima NS2B-NS3^{pro} A) Curvas obtidas para o teste inicial de atividade e padronização da concentração de enzima, cada uma das cores representa uma concentração diferente de enzima. B) Curva de calibração de 7-amino-4-metilcumarina (AMC)..... 78
- Figura 21** - Caracterização das enzimas NS2B-NS3^{pro}. A) Curvas de Michaelis-Menten para cada uma das enzimas mutantes e WT. As barras de erros correspondem ao desvio padrão das triplicatas. Os valores de afinidade pelo substrato (K_m), constante catalítica (k_{cat}) e eficiência catalítica (k_{cat}/K_m) de cada uma das enzimas é mostrado acima do gráfico. B) Perfis de desnaturação térmica de cada uma das enzimas. Os valores de T_m para cada uma das enzimas é mostrado acima do gráfico. 79
- Figura 22** - Análise de conservação da enzima NS2B-NS3 protease. A) Visão geral da estrutura da NS2B-NS3 protease de YFV, onde o domínio NS3 protease está colorido de acordo com a variabilidade sequencial comparada a outras proteases de flavivírus. A análise de conservação foi feita utilizando o servidor ConSurf.¹¹⁴ Cada resíduo da NS3^{pro} está representado como esferas. O cofator NS2B está representado na forma de cartoon e colorido em azul escuro. B-C) Visão ampliada das regiões contendo os resíduos mutantes da linhagem circulante da protease de YFV, K121 e D88 em (B) e (C) respectivamente. D) Sequência da protease de YFV com os resíduos

coloridos de acordo com a escala de conservação do Consurf.¹¹⁴ Os resíduos mutantes estão destacados em vermelho. 80

Figura 23 - Análise estrutural das mutações pontuais observadas na protease de YFV da linhagem circulante. A) Estrutura cristalográfica da enzima NS2B-NS3^{pro} representada em formato cartoon e localização das mutações D88 e K121, representadas como esferas (em amarelo). O cofator NS2B está colorido em laranja e a NS3 protease em azul. A região correspondente ao sítio ativo está destacada em cinza. B) Visão ampliada do resíduo K121 de YFV (em azul) (PDB ID 6URV) e sua interação com o resíduo N122 sobreposta com a mesma região da estrutura de WNV (em verde) (PDB ID 2YOL). C) Visão ampliada do resíduo E88 de YFV (em azul) e suas interações com os resíduos K86 e Q169 sobreposta com a mesma região da estrutura de DENV2 (em ciano) (PDB ID 4M9K). D) Alinhamento das sequências da proteína NS3 protease de YFV, DENV3, DENV2, ZIKV e WNV. Os resíduos mutantes estão destacados em verde e o alinhamento foi feito utilizando o servidor ESPrit.¹¹⁵ 81

Figura 24 - Resultados da triagem de compostos do Pandemic Response Box. A) Gráfico contendo os resultados da triagem inicial dos compostos, com suas respectivas porcentagens de atividade no teste pontual. Em azul, estão representados os controles 1% DMSO de cada uma das placas, em vermelho as reações branco de cada uma das placas e em cinza cada um dos compostos avaliados. B) Tabela contendo os resultados referentes aos compostos selecionados, suas posições na placa, códigos MMV, classificação e respectivos IC₅₀ (em µM). C) Gráfico de dose-resposta do composto MMV 1581548. D) Gráfico de dose-resposta do composto MMV 396785 (Alexidina). E) Gráfico de dose-resposta do composto MMV 1580853. 85

Figura 25 - Mecanismo de inibição do composto AG-Hecate na enzima NS2B-NS3^{pro}. A) Curvas de Michaelis-Menten para as concentrações de 0, 150 pM e 300 pM de AG-Hecate. As barras de erros correspondem ao desvio padrão das triplicatas. B) Curvas de Lineweaver-Burk para as concentrações de 0, 150 pM e 300 pM de AG-Hecate, obtidas da linearização dos dados utilizados na curva de Michaelis-Menten. C) Representação esquemática da sequência do peptídeo AG-Hecate, contendo ácido gálico no N-terminal, destacando os potenciais sítios de clivagem pela protease, coloridos na figura em amarelo. 86

Figura 26 - Gráficos de dose-resposta dos peptídeos da série Hecate. A) Dose-resposta do peptídeo Hecate. B) Dose-resposta do peptídeo AG-Hecate. C) Dose-resposta do peptídeo K-Hecate. D) Dose-resposta do peptídeo Bz-Hecate. E) Dose-resposta do peptídeo Acetil-Hecate. F) Dose-resposta do peptídeo PSSct 1905. 88

Figura 27 - Avaliação da atividade antiviral dos peptídeos A) Hecate, B) Bz-Hecate e C) PSSct 1905 em células BHK21-RepYFV-17D e citotoxicidade (D-F) respectivamente. 89

Figura 28 - Mapa do plasmídeo pET_DUET-1 contendo a sequência codificante da enzima NS2B-NS3 protease de YFV. 90

- Figura 29** - Gel de agarose 1% contendo os fragmentos obtidos para o vetor pET_DUET-1. As bandas de 1-6 correspondem a colônias contendo ambos os insertos, NS2B e NS3 protease. A banda indicada por M representa o marcador molecular (em pb). As bandas 1-3 correspondem a PCR realizada com os primers NS2Bduetfw e T7 term (1020 pb), enquanto as bandas de 4-6 correspondem a PCR realizada com os primers NS2Bduetfw e NS3prorv (970 pb). A banda 7 corresponde a uma das colônias positivas da primeira etapa de clonagem contendo apenas o inserto correspondente ao cofator NS2B. A PCR foi realizada com os primers NS2Bduetfw e T7 term (506 pb). B) SDS-PAGE 15% contendo os resultados dos testes de expressão da construção unlinked. No gel, vemos as frações celulares solúveis de cada uma das condições testadas, indicadas na figura. A banda indicada por M representa o marcador molecular (em kDa). A proteína de interesse está indicada pela seta vermelha (NS3 protease: 18 kDa). 90
- Figura 30** - Resultados da purificação da proteína NS2B-NS3^{pro} unlinked. A) Cromatograma obtido na etapa de exclusão molecular. B) SDS-PAGE 15% mostrando a pureza da amostra ao final da purificação, a proteína está indicada pela seta vermelha (NS3 protease: 18 kDa)..... 91
- Figura 31** - A) Curvas de Michaelis-Menten para cada uma das enzimas WT linked e unlinked. As barras de erros correspondem ao desvio padrão das triplicatas. B) Curvas obtidas na DSF para cada uma das construções..... 92
- Figura 32** - Cristais da enzima NS2B-NS3^{pro} YFV unlinked. A) Cristais da proteína na condição B10 PACT (0.2 M cloreto de magnésio, 0.1 M MES pH 6, 20% PEG 6000). B) Cristais da proteína na condição D04 PACT (0.1 M MMT pH 7, 25% PEG 1500). C) Cristais da proteína na condição G05 PACT (0.2 M nitrato de sódio, 20% PEG 3350). 93
- Figura 33** - Representação esquemática do genoma viral do SARS-CoV-2 e da poliproteína viral pp1ab. As proteínas não estruturais (NSPs) estão representadas em verde, as proteínas estruturais em vermelho e as proteínas acessórias em roxo. As setas estão indicando os sítios de clivagem de cada uma das proteases virais. Em vermelho, estão representados os sítios de clivagem da enzima Papain-like-protease ou PL^{pro} e em azul, os sítios da enzima Main-protease ou M^{pro}. 100
- Figura 34** - Representação da estrutura da enzima M^{pro} do vírus SARS-CoV-2. Cada um dos protômeros está colorido em cores diferentes (verde e amarelo). Os resíduos catalíticos estão destacados em vermelho. DI, DII e DIII indicam respectivamente cada um dos domínios da proteína. A direita, visão ampliada do sítio ativo da enzima, composto pela díade dos resíduos Cys145 e His41 (representados como ball and sticks e coloridos em vermelho), mostrando também o sítio de ligação ao substrato com os subsítios S1-S4. 102
- Figura 35** - Representação esquemática das construções da proteína SARS-CoV-2 Main Protease. A porção correspondente a sequência da proteína está colorida em verde, os sítios de clivagem de proteases estão indicados por setas (vermelho para M^{pro} e amarelo para TEV protease) e a cauda de 6xHis está destacada em azul..... 106

- Figura 36** - Representação esquemática do ensaio enzimático da proteína SARS-CoV-2 M^{pro} baseado em FRET. O substrato Dabcyl-KTSAVLQ↓SGFRKME(Edans)-NH₂ é clivado pela enzima liberando dois peptídeos menores, cuja fluorescência pode ser monitorada nos comprimentos de onda apresentados. 111
- Figura 37** - Clonagem das diferentes construções da M^{pro} de SARS-CoV-2. A) Gel de agarose 1% contendo os resultados da PCR de colônia obtidos para as construções IMT-M^{pro} (961 pb) e M^{pro} (982 pb) respectivamente. A amostra indicada por M contém o marcador de peso molecular, com os tamanhos em pb indicados na figura. B) Mapa do plasmídeo pET_M11/LIC obtido na clonagem para as construções de M^{pro}. 115
- Figura 38** - Purificação das 3 construções da enzima M^{pro}, IMT-M^{pro}, M^{pro} e C145S M^{pro}, respectivamente. Para ambas as enzimas estão apresentados os cromatogramas referentes a última etapa de purificação e os respectivos SDS-PAGE (12.5%). 116
- Figura 39** - A) Curvas de fluorescência das diferentes construções de M^{pro} medidas ao longo do tempo. B) Curvas de Michaelis-Menten para cada uma das enzimas mutantes. As barras de erros correspondem ao desvio padrão das triplicatas. C) Perfis de eluição na cromatografia de exclusão molecular para cada uma das enzimas. As curvas correspondem a mudança na intensidade no espalhamento normalizado em 90° (linhas) e massas moleculares calculadas para cada um dos picos (pontos). D) Perfis de desnaturação térmica de cada uma das enzimas. 119
- Figura 40** - Cristais da enzima M^{pro} de SARS-CoV-2. A) Cristais da enzima IMT-M^{pro} obtidos na condição E09 do kit MORPHEUS (0.12 M etileno glicóis, 0.1 M Buffer System 3 pH 8.5, 30% PEG mix 1). B) Cristais da enzima IMT-M^{pro} obtidos na condição D02 do kit BCS (10% 2-PropOH, 0.1 M MES pH 6.5, 22.5% PSL). Em A-B, a imagem da direita corresponde a mesma condição sob luz UV. C) Cristais otimizados da IMT-M^{pro} na condição 0.1 M MES pH 6.7, 5% DMSO, 8% PEG 4K (utilizando seed obtido da condição B). D) Cristais da M^{pro} (ortorrômbico) obtidos na condição 0.1 M MES pH 6.7, 5% DMSO, 8% PEG 4K (utilizando seed da IMT-M^{pro} obtido da condição B). E) Cristais da enzima C145S M^{pro} (tetrâmero), na condição A08 do kit BCS (0.1 M fosfato/citrato pH 5.5, 20% PEG Smear High). 120
- Figura 41** - Padrões de difração de cristais da enzima M^{pro} de SARS-CoV-2, ambos coletados na linha MANACA, no SIRIUS. A) Cristais da enzima IMT-M^{pro} obtidos na condição 0.1 M MES pH 6.7, 5% DMSO, 8% PEG 4K. B) Cristais da enzima M^{pro} no grupo espacial ortorrômbico. 121
- Figura 42** - Mudanças estruturais observadas na forma imatura da enzima M^{pro}, destacando a região do bolsão de ligações ao substrato e a distância entre os dois protômeros. A esquerda, a estrutura da IMT-M^{pro} (PDBid 7KFI) está representada como ribbon e colorida de acordo com o RMSD, comparada a M^{pro} nativa (PDBid 7MBG). Nos quadros a direita, as estruturas da M^{pro} imatura e nativa estão representadas como cartoon, e coloridas respectivamente em vermelho e azul. Resíduos impactados pelas mudanças

estruturais estão representados como sticks, e as distâncias estão indicadas pelos traços em amarelo. 123

Figura 43 - A) Visão geral do domínio DIII da IMT-M^{pro} (cadeia A em amarelo e cadeia B em ciano) superposta com a M^{pro} nativa (cinza). Os resíduos adicionais da região N-terminal estão representados como esferas e coloridos em vermelho. B) Visão rotacionada do domínio DIII da IMT-M^{pro}. C) Sítio ativo e bolsão de ligação ao substrato da IMT-M^{pro} cadeia B (cartoon ciano) superposta com a M^{pro}. Os resíduos catalíticos e importantes na ligação ao substrato estão coloridos em amarelo e representados como sticks. M^{pro} nativa e seus respectivos resíduos estão coloridos em cinza..... 124

Figura 44 - A) Visão geral dos fragmentos identificados para os cristais da enzima IMT-M^{pro}. A estrutura da enzima está representada como superfície, com as cadeias A e B coloridas em amarelo e ciano, respectivamente. Os fragmentos estão representados como esferas e coloridos em vermelho. B) Visão ampliada dos fragmentos identificados em cada um dos sítios, destacando as interações feitas por cada um deles. A estrutura da enzima está representada como cartoon, os fragmentos como sticks (coloridos em amarelo) e os resíduos parte das interações como linhas (coloridos em verde)..... 126

Figura 45 - Detalhes do grupo espacial dos cristais da IMT-M^{pro}. A) Representação de superfície da região onde foi identificado o fragmento f2xg02 na enzima IMT-M^{pro}, mostrada na M^{pro} cristalizada no grupo espacial canônico. B) Representação da mesma região mostrada em A, na enzima IMT-M^{pro} cristalizada no grupo espacial ortorrômbico, contendo o fragmento f2xg02. As cadeias A e B estão representadas como superfície com transparência, juntamente com os resíduos da região representados como sticks, e coloridas em ciano e amarelo respectivamente. C) Visão espacial da simetria presente no grupo espacial canônico da M^{pro}, destacando a região do sítio ativo. D) Visão espacial do grupo espacial da enzima IMT-M^{pro}, destacando a região do sítio ativo e evidenciando o maior espaço para acesso ao bolsão de ligação ao substrato. Cada um dos dímeros está representado como superfície e coloridos em diferentes cores.. 127

Figura 46 - Estrutura cristalográfica da enzima M^{pro} em dois grupos espaciais, P2₁2₁2₁ e C2₁. Ambas as estruturas estão sobrepostas, representadas como cartoon e coloridas em magenta e amarelo. Na esquerda, visão ampliada do sítio ativo e resíduos do bolsão de ligação ao substrato. Os resíduos estão representados como sticks e coloridos de acordo com cada uma das estruturas (magenta para ortorrômbica e amarela para monoclinica). 129

Figura 47 - Estrutura cristalográfica da C145S M^{pro} tetrâmero. À direita, a estrutura do dímero presente na unidade assimétrica está representada como superfície, e as cadeias A e B coloridas de amarelo e azul respectivamente. O dímero relacionado por simetria está representado como superfície transparente e cartoon, seguindo as mesmas cores. Os peptídeos ligados no sítio ativo estão representados como esferas, coloridos em vermelho e verde para o N e C terminal respectivamente. À direita, visão ampliada do sítio ativo de cada uma das cadeias, destacando os peptídeos ligados ao sítio ativo, mostrando a densidade eletrônica 2Fo-Fc de cada um deles ($\sigma = 1$). 131

- Figura 48** - Interações feitas pelos peptídeos do N e C-terminal com os resíduos do sítio ativo da proteína C145S M^{pro}. A) Detalhe das interações para o peptídeo N-terminal (cadeia A). B) Detalhe das interações para o peptídeo do C-terminal (cadeia B). Em ambas, os peptídeos estão representados como ball and stick e coloridos em amarelo e os resíduos parte das interações estão representados como linhas e coloridos em ciano..... 132
- Figura 49** - Superposição dos resíduos do C-terminal da M^{pro} para as enzimas M^{pro} madura, C145S M^{pro} tetrâmero - cadeia B e C145A M^{pro} (PDBid 7JOY). As estruturas estão representadas como cartoon e sticks e coloridas em verde, ciano e magenta, respectivamente. 132
- Figura 50** - A) SDS-PAGE contendo a clivagem do N-terminal da C145S M^{pro} ao longo do tempo. Na parte superior, reação contendo 10 mM C145S M^{pro}, e na parte inferior a mesma reação suplementada com 5 nM de M^{pro}. As setas vermelhas indicam para a banda de clivada da M^{pro}. B) Gráfico de barras mostrando a intensidade relativa das bandas de M^{pro} clivada e não clivada ao longo do tempo. C-D) Perfis de eluição na cromatografia de exclusão molecular da C145S M^{pro} ao longo do tempo para as reações sem e com M^{pro} madura respectivamente. C. As curvas correspondem a mudança na intensidade no espalhamento normalizado em 90° (linhas) e massas moleculares calculadas para cada um dos picos (pontos)..... 134
- Figura 51** - Curvas de fluorescência ao longo do tempo para diferentes tempos de auto clivagem da enzima C145S M^{pro} monômero. 135
- Figura 52** - Esquema do processamento dos dados de crio-microscopia eletrônica da C145S M^{pro}. A) Exemplo de micrografias alinhada. A barra de escala corresponde a 20 nm. B) Anéis de Fourier da micrografia apresentada em A. C) Exemplos de partículas presentes nas micrografias. D) Passo a passo detalhado contendo as etapas de processamento e refinamento dos dados. 136
- Figura 53** - Camadas de correlação de Fourier (FSC) mostrando a resolução global (FSC = 0.143) e mapa de resolução local. A resolução está colorida de acordo com a resolução local, de vermelho para o azul, de acordo com valores crescentes de resolução..... 139
- Figura 54** - A) Visão geral da densidade do volume final reconstruído para a C145S M^{pro} obtido por crio-microscopia eletrônica. Cada uma das cadeias A e B estão coloridas respectivamente em cinza claro e escuro. O peptídeo correspondente a região nsp4-nsp5 está colorido em ciano. B) Visão ampliada do sítio ativo da estrutura da cryo-EM da C145S M^{pro}. A densidade correspondente ao peptídeo nsp4-nsp5 está colorida em ciano e os resíduos pertencentes à este peptídeo estão representados como sticks e coloridos em amarelo. Os resíduos que interagem com o peptídeo através de interações polares estão representados na figura como sticks e coloridos em cinza. C) Visão ampliada do sítio ativo da enzima mostrando a região como superfície, destacando o bolsão de ligação ao substrato e os subsítios S5-S5'. Cada um dos subsítios está indicado em vermelho. e os resíduos pertencentes a este peptídeo estão representados como sticks e coloridos em amarelo. 140

- Figura 55** - Representação esquemática da M^{pro} tetramérica em solução. A M^{pro} está representada como superfície com as cadeias A e B coloridas em azul e ciano, respectivamente. As M^{pro} monoméricas parcialmente enoveladas estão coloridas em vermelho. B) Visão detalhada das partículas de M^{pro} encontradas nas micrografias, onde as moléculas de M^{pro} dimérica estão marcadas em azul e as M^{pro} monômero associadas ao dímero estão marcadas em vermelho. 140
- Figura 56** - Espectrometria de massas nativa da M^{pro} tetramérica. Da esquerda para a direita, os picos mostram monômeros clivados (semicírculo azul) e não clivados (semicírculo vermelho), dímeros formados por partículas clivadas (círculos azuis) ou semi-clivadas (círculo azul-vermelho), trímeros formados por duas partículas clivadas e uma não clivada (dois terços azuis, um terço círculos vermelhos) e tetrâmeros formados por duas partículas clivadas e duas não clivadas (dois quartos de azul, dois quartos de vermelho). 141
- Figura 57** - Comparação entre as diferentes estruturas de M^{pro}, M^{pro} madura (PDBid 7MBG), C145S M^{pro} tetrâmero (PDBid 7N6N) e C145S M^{pro} obtida por cryo-EM (7S82). As estruturas estão representadas como cartoon e coloridas respectivamente em amarelo, rosa e azul. B) RMSD calculado entre a estrutura da C145S M^{pro} tetrâmero e a estrutura obtida por cryo-EM. A proteína está representada como traço da cadeia alpha e colorida de acordo com os valores de RMSD, do azul para o vermelho, de acordo com a legenda. 141
- Figura 58** - Representação esquemática da dinâmica em solução observada no SEC-MALS ao longo do tempo para a amostra 1, contendo C145S M^{pro} monômero. De cima para baixo estão apresentados respectivamente os tempos 0h, 24h e 48h. A) Controle DMSO. B) Amostra contendo o inibidor não covalente MAT-POS-e194df51-1. C) Amostra contendo o inibidor covalente nirmatrelvir. 145
- Figura 59** - Representação esquemática da dinâmica em solução observada no SEC-MALS ao longo do tempo para a amostra 1, contendo C145S M^{pro} tetrâmero. De cima para baixo estão apresentados respectivamente os tempos 0h, 24h e 48h. A) Controle DMSO. B) Amostra contendo o inibidor não covalente MAT-POS-e194df51-1. C) Amostra contendo o inibidor covalente nirmatrelvir. 145
- Figura 60** - Estrutura cristalográfica da C145S M^{pro} monomérica em complexo com nirmatrelvir, alinhada com a estrutura da enzima apo madura. Ambas as proteínas estão representadas como cartoon. A estrutura da C145S M^{pro} e da M^{pro} estão coloridas em amarelo e azul (para as cadeias A e B) e cinza, respectivamente. 147
- Figura 61** - Mudanças estruturais no sítio ativo da enzima M^{pro} durante o processo de clivagem. A) Principais resíduos do sítio ativo (representados em sticks verdes) da M^{pro} em forma apo (topo), representação em cartoon colorido em amarelo (meio) superfície de potencial eletrostático (parte inferior). B) Principais resíduos do sítio ativo (sticks verdes) da C145S M^{pro} (cryo-EM) ligada ao peptídeo não clivado (topo), representação em cartoon colorido em azul (meio) e superfície de potencial eletrostático (parte inferior). C)

Principais resíduos do sítio ativo (sticks verdes) da C145S M^{pro} ligado covalentemente ao peptídeo (complexo intermediário enzima-substrato) (topo), representação em cartoon colorido em verde (meio) e superfície de potencial eletrostático (parte inferior). D) Principais resíduos do sítio ativo (sticks verdes) da C145S M^{pro} ligado não covalentemente ao peptídeo (topo), representação em cartoon colorido em vermelho (meio) e superfície de potencial eletrostático (parte inferior). Os sticks transparentes e cartoon coloridos em cinza nas figuras do topo e central representam a posição estrutural dos elementos na etapa anterior..... 148

Figura 62 - Representação esquemática do processo de maturação da enzima M^{pro} de SARS-CoV-2. Cada um dos monômeros está colorido respectivamente de verde e laranja..... 150

LISTA DE TABELAS

- Tabela 1** - Primers planejados para clonagem da sequência codificantes da proteína NS2B-NS3^{pro} e inserção do *linker*. 53
- Tabela 2** - Sequências dos primers planejados para a clonagem da nova construção no vetor pET_DUET-1. 54
- Tabela 3** - Sequências dos primers planejados para mutagênese sítio dirigida. 55
- Tabela 4** - Condições avaliadas no teste de expressão 56
- Tabela 5** - Coleta de dados e estatísticas do refinamento da enzima NS2B-NS3^{pro}. Os dados entre parênteses correspondem à última camada de resolução. 71
- Tabela 6** - Série de peptídeos derivados do Hecate, suas respectivas sequências, porcentagem de inibição no teste pontual a 20 μ M e valores de IC₅₀ (em nM). 87
- Tabela 7** - Primers utilizados para a clonagem da sequência codificante da enzima M^{pro} em 3 diferentes construções 106
- Tabela 8** - Volumes de retenção de diferentes construções da M^{pro}, massas moleculares calculadas e estados oligoméricos. Ambos os volumes de retenção foram obtidos da filtração em coluna Superdex 75 16/60 e as massas calculadas de acordo com a curva de calibração apresentada no capítulo 1. 117
- Tabela 9** - Parâmetros cinéticos calculados e estabilidade térmica para cada uma das M^{pros}. A eficiência relativa foi calculada com relação a M^{pro} nativa. 117
- Tabela 10** - Massas moleculares para cada pico observado nos gráficos do SEC-MALS de cada uma das construções de M^{pro} 118
- Tabela 11** - Coleta de dados e estatísticas do refinamento referentes a enzima M^{pro} de SARS-CoV-2. Os dados entre parênteses correspondem à última camada de resolução. 122
- Tabela 12** - Coleta de dados e estatísticas do refinamento referentes a enzima IMT-M^{pro} com diferentes fragmentos. Os dados entre parênteses correspondem à última camada de resolução. 125
- Tabela 13** - Massa molecular e razão de recuperação das massas entre as porcentagens dos picos obtidos no SEC-MALS para a C145S M^{pro}. A razão de recuperação da massa foi calculada entre as porcentagens do pico 1 e 2. 134
- Tabela 14** - Coleta de dados de crio microscopia, parâmetros de refinamento e validação e estatísticas. 137
- Tabela 15** - Massa molecular e razão de recuperação das massas entre as porcentagens dos picos obtidos no SEC-MALS para a C145S M^{pro}. A razão de recuperação da massa foi calculada entre as porcentagens do pico 1 e 2. 144

Tabela 16 - Coleta de dados e estatísticas do refinamento referentes a enzima C145S M^{pro} em complexo com nirmatrevir. Os dados entre parênteses correspondem à última camada de resolução.146

LISTA DE ABREVIATURAS E SIGLAS

3CL ^{pro}	<i>3-chymotrypsin-like protease</i>
3DVA	análise de variabilidade 3D
AG	ácido gálico
AMC	7-amino-4-metilcumarina
BSA	albumina de soro bovina
CC	coeficiente de correlação
CC ₅₀	concentração de toxicidade média
COVID-19	<i>coronavirus disease of 2019</i>
cryo-EM	crio microscopia eletrônica
CTF	função de transferência de contraste
DENV	Vírus da Dengue
DMEM	<i>Dulbecco's Modified Eagle Medium</i>
DMSO	dimetil sulfóxido
DNA	ácido desoxirribonucléico
DSF	fluorimetria diferencial de varredura
DTT	ditiotreitól
EC ₅₀	concentração inibitória média
EDTA	ácido etilenodiamino tetra-acético
EMDB	<i>Electron Microscopy Data Bank</i>
FBDD	desenvolvimento de fármacos baseado em fragmentos
FDA	<i>U.S. Food and Drug Administration</i>
FRET	transferência ressonante de energia por fluorescência

FSC	camada de correlação de Fourier
HCV	Vírus da Hepatite C
HPLC	cromatografia líquida de alta eficiência
HRV	Rinovírus humano
IC ₅₀	concentração inibitória média
IPTG	isopropiltio-β-galactosídeo
LB	<i>lysogeny broth</i>
LIC	ligação independente de ligase
MERS-CoV	Síndrome respiratória do Oriente Médio
Mpro	<i>Main Protease</i>
MTT	3- (4,5-dimetiltiazol-2-il) -2,5-brometo de difeniltetrazólio
Native-MS	espectrometria de massas nativa
NSs	proteínas não estruturais
OMS	Organização Mundial de Saúde
ORF	frame aberto de leitura
PCR	reação em cadeia da polimerase
PDB	Protein Data Bank
PL ^{pro}	<i>papain-like-protease</i>
PNK	T4 polinucleotídeo quinase
RMSD	desvio da média quadrática
RNA	ácido ribonucleico
SAM	S-adenosil metionina
SARS-CoV	síndrome respiratória aguda grave do coronavirus

SARS-CoV-2	síndrome respiratória aguda grave do coronavírus 2
SBF	soro bovino fetal
SDS-PAGE	eletroforese em gel de poliacrilamida com dodecil- sulfato de sódio
SEC-MALS	cromatografia por exclusão de tamanho acoplada ao espalhamento de luz em múltiplos ângulos
TB	<i>terrific broth</i>
IMAC	cromatografia de afinidade por íons metálicos imobilizados
TEV	protease do Vírus do mosaico do Tabaco
T _m	temperatura de <i>melting</i>
UTR	<i>untranslated region</i>
WNV	Vírus do leste do Nilo
WT	<i>wild type</i>
YFV	Vírus da febre amarela
ZIKV	Vírus zika

SUMÁRIO

1	INTRODUÇÃO GERAL.....	33
---	-----------------------	----

CAPÍTULO 1: Estudos estruturais e busca por candidatos antivirais contra a Febre Amarela tendo como alvo o complexo NS2B-NS3 protease

1	INTRODUÇÃO	39
1.1	Panorama geral sobre a Febre Amarela.....	39
1.2	Situação atual da febre amarela no Brasil	41
1.3	Biologia estrutural do vírus e do complexo NS2B-NS3 protease	43
1.4	Polimorfismo gênico do surto de 2017 e possíveis implicações estruturais	47
1.5	NS2B-NS3 ^{pro} como alvo no desenvolvimento de antivirais e objetivos do trabalho	48
2	MATERIAIS E MÉTODOS	51
2.1	Aquisição das sequências codificantes de YFV	51
2.2	Clonagem	51
2.2.1	Construções inicial e NS2B-NS3 ^{pro} <i>linked</i>	51
2.2.2	NS2B-NS3 ^{pro} <i>unlinked</i>	53
2.3	Mutação sítio dirigida.....	54
2.4	Testes de expressão e expressão em larga escala	55
2.5	Purificação das enzimas	56
2.6	Cristalização, co-cristalização e coleta de dados.....	57
2.7	Resolução das estruturas	58
2.8	Caracterização das proteínas mutantes.....	59
2.9	Ensaio de atividade	59
2.10	Ensaio de inibição.....	60
2.11	Ensaio antiviral baseado no replicon de YFV	61
3	RESULTADOS	63
3.1	Construção NS2B-NS3 ^{pro} YFV <i>linked</i>	63
3.1.1	Clonagem	63
3.1.2	Testes de expressão	64
3.1.3	Mutagênese sítio dirigida	65
3.1.4	Expressão em larga escala e purificação das enzimas	66
3.1.5	Ensaio de cristalização e coleta de dados	68
3.1.6	Resolução da estrutura cristalográfica da proteína NS2B-NS3 ^{pro} YFV	71
3.1.7	Estrutura cristalográfica da proteína NS2B-NS3 ^{pro} YFV	72

3.1.8	Ensaio de atividade	77
3.1.9	Caracterização das enzimas mutantes.....	78
3.1.10	Ensaio de inibição	82
3.2	Construção NS2B-NS3 ^{pro} <i>unlinked</i>	89
3.2.1	Clonagem e testes de expressão.....	89
3.2.2	Expressão em larga escala e purificação	91
3.2.3	Caracterização da enzima	92
3.2.4	Ensaio de cristalização	93
4	CONCLUSÕES E PERSPECTIVAS	95

CAPÍTULO 2: Caracterização em solução e elucidação estrutural do processo de maturação da enzima Main Protease (M^{pro}) de SARS-CoV-2

1	INTRODUÇÃO.....	99
1.1	SARS-CoV-2 e pandemia de COVID-19	99
1.2	Biologia estrutural do vírus SARS-CoV-2	99
1.3	Vacinação e desenvolvimento de antivirais contra a COVID-19	100
1.4	Enzima Main Protease (M ^{pro})	101
1.5	Processo de maturação da enzima M ^{pro} e objetivos do trabalho	102
2	MATERIAIS E MÉTODOS.....	105
2.1	Clonagem da sequência codificante da enzima M ^{pro} em 3 diferentes construções: IMT-M ^{pro} , M ^{pro} e C145S M ^{pro}	105
2.2	Expressão em larga escala das enzimas.....	106
2.3	Purificação das enzimas.....	106
2.3.1	Purificação da enzima IMT-M ^{pro}	106
2.3.2	Purificação da enzima M ^{pro}	107
2.3.3	Purificação da enzima C145S M ^{pro}	107
2.4	Cristalização	108
2.5	Coleta de dados de difração e processamento	108
2.6	Triagem de Fragmentos contra forma imatura da M ^{pro} (IMT-M ^{pro})	109
2.7	Caracterização das enzimas em solução	110
2.7.1	Ensaio de atividade	110
2.7.2	Differential Scanning Fluorimetry (DSF).....	111
2.7.3	Cromatografia por exclusão de tamanho acoplada ao espalhamento de luz em múltiplos ângulos (SEC-MALS)	111
2.7.4	Espectrometria de massas nativa (Native-MS).....	112
2.8	Obtenção da estrutura da enzima C145S M ^{pro} por crio-microscopia eletrônica (cryoEM)	112

2.8.1	Preparação das grades e coleta de dados	113
2.8.2	Processamento dos dados	113
2.8.3	Modelagem e análises estruturais.....	114
3	RESULTADOS E DISCUSSÃO	115
3.1	Clonagem, expressão e purificação das enzimas.....	115
3.2	Ensaio de atividade e caracterização bioquímica	117
3.3	Cristalização, coleta de dados e resolução das estruturas.....	119
3.4	Estrutura cristalográfica da enzima IMT-M ^{pro}	122
3.5	Triagem de Fragmentos contra a forma imatura da M ^{pro} (IMT-M ^{pro})	124
3.6	Estrutura cristalográfica da enzima M ^{pro}	128
3.7	Estrutura da C145S M ^{pro} em forma tetramérica em complexo com os peptídeos endógenos do N e C-terminal.....	129
3.8	Caracterização em solução da enzima C145S M ^{pro}	133
3.10	Análise da influência de inibidores no processo de maturação da M ^{pro}	142
3.10.1	Inibidor não-covalente: MAT-POS-e194df51-1	142
3.10.2	Inibidor covalente: Nirmatrelvir.....	143
3.11	Estrutura cristalográfica da enzima C145S M ^{pro} (monomérica) em complexo com nirmatrelvir.....	146
3.12	Bases moleculares do processo de auto clivagem da M ^{pro}	147
3.13	Modelo proposto para o auto processamento (maturação) da enzima M ^{pro} e implicações para o desenvolvimento de antivirais.....	149
4	CONCLUSÕES E PERSPECTIVAS.....	151
	CONCLUSÕES GERAIS E PERSPECTIVAS	153
	TRABALHOS PUBLICADOS REFERENTE AOS TRABALHOS DESENVOLVIDOS NESTA TESE.....	155
	OUTROS TRABALHOS PUBLICADOS PELA AUTORA DURANTE A REALIZAÇÃO DESTE DOUTORADO.....	155
	OUTROS TRABALHOS PUBLICADOS DURANTE A REALIZAÇÃO DESTE DOUTORADO DESENVOLVIDOS EM COLABORAÇÃO.....	155
	REFERÊNCIAS.....	159
	ANEXOS.....	171

1 INTRODUÇÃO GERAL

Os vírus são os agentes mais simples e numerosos na Terra. Estima-se que em nosso corpo, para cada célula humana, existem cerca de 100 vírus, dos quais conhecemos apenas uma pequena parte. Embora os primeiros relatos de doenças virais datem do início da idade média, a virologia é uma ciência relativamente recente, dada a dificuldade para a visualização e identificação destes agentes.¹⁻² No século XIX, a publicação de um artigo contendo as primeiras imagens de microscopia de um vírus marcaram a ciência, e a partir do século XX, diversos vírus foram identificados.³ O desenvolvimento de vacinas teve papel fundamental no controle e erradicação de algumas viroses graves, como poliomielite e varíola. No entanto, diversas permanecem sem vacinas, e no geral, existem poucos tratamentos específicos para doenças virais.

A natureza biológica dos vírus, a ausência de metabolismo próprio e a necessidade de serem parasitas intracelulares, dificulta o estudo dos vírus e o desenvolvimento de antivirais. Por não possuírem metabolismo próprio, encontrar um fármaco capaz de afetar exclusivamente o vírus, sem afetar as nossas células, é uma tarefa extremamente difícil. Além da dificuldade da identificação dos vírus, estes não são fáceis de serem cultivados e manuseados em laboratório. A simplicidade genética dos vírus e alta variabilidade genômica e estrutural, impossibilita o desenvolvimento de fármaco antiviral de amplo espectro, e os torna mais suscetíveis a mutações, o que corrobora com a emergência de resistência aos medicamentos já existentes.^{2,4}

Os recentes avanços em pesquisa possibilitaram uma maior compreensão do mecanismo de diversas doenças virais e auxiliaram no desenvolvimento de novos antivirais, no entanto, as doenças virais ainda representam uma das principais causas de enfermidades e mortes entre humanos. Segundo a Organização de Saúde (OMS), 37.7 milhões de pessoas atualmente vivem com HIV, mais de 250 milhões com Hepatite B e 70 milhões de pessoas com Hepatite C.⁵ Além disso, entre 3 e 5 milhões de pessoas são infectadas anualmente por Influenza. As doenças virais totalizaram durante o ano de 2021, de acordo com a OMS, aproximadamente 20% da taxa de mortalidade mundial.⁵

É unânime entre a comunidade científica que doenças virais continuarão a emergir, causando epidemias e potencialmente podendo ocasionar uma nova pandemia. Novas doenças virais são comumente ocasionadas por novos vírus provenientes de ambiente silvestre, mutação de vírus já existentes ou pela reemergência de doenças já não observadas em ambiente urbano.^{2,4} Existem diversas possíveis causas, dentre elas a globalização e a

alta mobilidade internacional de pessoas, aumento populacional, expansão urbana, aquecimento global e desmatamento, que acaba por aproximar a população do ambiente silvestre, o que contribui com a reurbanização de doenças e transmissão de novos vírus provenientes de hospedeiros nativos de ambientes silvestres.⁶⁻⁹

Neste contexto, é evidente que o estudo da biologia estrutural dos vírus é fundamental para investigação do mecanismo molecular das infecções. Uma melhor compreensão da estrutura, função dos vírus e sua interação com o hospedeiro nos permite não somente compreender o mecanismo das doenças em nível fundamental, proporcionando maior entendimento sobre o ciclo de replicação dos vírus, mas também permite a identificação de novos alvos e auxilia no desenvolvimento de ferramentas para controle e prevenção de atuais e futuras epidemias.

A pandemia de SARS-CoV-2 (Síndrome respiratória aguda grave causada por coronavírus 2), causador da COVID-19, uma doença altamente infecciosa e contagiosa, se espalhou rapidamente em proporções mundiais, afetando de forma impactante a economia e os sistemas de saúde pública. Diante das graves consequências geradas por essa pandemia, um rápido desenvolvimento de candidatos vacinais e busca por antivirais, ocorreu de forma nunca vivenciada pela comunidade científica e população mundial. Os desafios impostos pela pandemia de COVID-19 devem estimular a conscientização de cientistas, governos e sociedade da importância das pesquisas básicas em virologia e o desenvolvimento de novos antivirais. Diversas doenças virais permanecem sem tratamento específico, como poxvírus, alphavírus, flavivírus, e alguns vírus de febre hemorrágica. Além disso, conforme previamente mencionado, a constante emergência de resistência a antivirais evidencia a necessidade constante de busca por novas alternativas de tratamento. Uma vez que, dada a grande variedade de vírus existentes e desconhecidos na natureza, epidemias e pandemias virão acontecer, cabe a nós continuarmos nosso trabalho e estarmos mais preparados para enfrentar e tratar as infecções, e talvez em algum momento, prevenir o surgimento de novos surtos.

Considerando a necessidade da obtenção de novos antivirais contra vírus emergentes, este projeto, que teve início em 2018, visava a obtenção de potenciais candidatos antivirais contra a febre amarela, dada a preocupação com a possível reurbanização da doença após a epidemia ocorrida no Brasil entre 2015-2017, através da elucidação da estrutura do complexo NS2B-NS3 protease e utilização deste para a triagem de inibidores. Além disso, após o início da pandemia de COVID-19, diversos grupos de pesquisa voltaram seus esforços para estudos estruturais de proteínas do vírus SARS-CoV-

2 e a busca por inibidores. No Centro de Pesquisa e Inovação em Biodiversidade e Fármacos (CIBFar, CEPID-FAPESP 2013/07600-3), o cenário não foi diferente. Como o laboratório conta com uma ampla infraestrutura para a produção e cristalização de proteínas recombinantes, e possui uma plataforma para a caracterização de inibidores de proteínas para flavivírus e alphavírus, rapidamente redirecionamos nossas pesquisas para o SARS-CoV-2 utilizando toda a infraestrutura já existente.

Desta forma, esta tese será dividida em 2 capítulos:

1. Busca por candidatos antivirais contra o vírus da Febre Amarela baseados na estrutura do complexo NS2B-NS3 protease, que abordará a expressão recombinante da proteína e a obtenção de sua estrutura cristalográfica seguida da triagem e identificação de inibidores potentes.
2. Caracterização em solução e elucidação estrutural do processo de maturação da enzima *Main Protease* (M^{Pro}) de SARS-CoV-2, que abordará a expressão recombinante da proteína e utilização de técnicas biofísicas, bioquímicas e estruturais para compreensão de seu processo de maturação.

CAPÍTULO 1

*Estudos estruturais e busca por candidatos antivirais contra
Febre Amarela tendo como alvo o complexo NS2B-NS3
protease*

1 INTRODUÇÃO

1.1 Panorama geral sobre a Febre Amarela

A febre amarela é uma doença infecciosa grave ocasionada por um vírus de RNA pertencente à família *Flaviviridae* e gênero *Flavivirus*. Recentemente, a epidemia no Brasil de origem silvestre em 2017 trouxe de volta a febre amarela como risco de saúde pública, podendo espalhar-se novamente no ambiente urbano.¹⁰ Apesar de haver controvérsias na determinação da origem do vírus, estudos genômicos indicam que o vírus surgiu há pelo menos 3000 anos no continente africano.¹¹ Durante o final do século XVIII e XIX, o comércio de escravos ajudou o vírus a se espalhar pelo mundo, causando epidemias na América do Norte, América Central e regiões da Ásia.¹² Em 1881, o pesquisador Carlos Finlay demonstrou que a febre amarela era transmitida pelo mosquito *Aedes aegypti*, mas ainda não se sabia qual a era causa da doença.¹³ Apenas em 1927, em Gana, o pesquisador Adrian Stokes foi capaz de isolar o primeiro vírus do sangue de um homem, dando o nome ao grupo dos flavivírus – do latim “flavus” significa amarelo.¹⁴

Atualmente a febre amarela ainda é um problema de saúde pública, afetando principalmente regiões de clima tropical no continente africano e na América do Sul.¹⁵ No ano de 2023, de acordo com a Organização Mundial de Saúde (OMS), 29 países africanos foram classificados como áreas endêmicas da doença e na América do Sul e Central, 13 países encontram-se em áreas de risco, incluindo o Brasil, nos quais a vacinação é altamente recomendada.¹⁶ A transmissão se dá primariamente pela picada do mosquito infectado da família *Culicidae*, principalmente dos gêneros *Aedes* e *Haemagogus* (Figura 1).¹⁷ Do ponto de vista epidemiológico, a febre amarela é classificada em dois tipos: silvestre e urbano, que diferem entre si de acordo com os transmissores e o local de ocorrência (Figura 1).¹⁸ O ciclo silvestre envolve a transmissão entre primatas não-humanos e humanos principalmente por mosquitos do gênero *Haemagogus*, que são comuns em regiões de mata. O ciclo urbano envolve a transmissão do vírus entre humanos, cujo principal vetor é o mosquito *Aedes aegypti*. Em geral, o vírus é trazido para a cidade quando um homem se infecta com a febre amarela do tipo silvestre (Figura 1).

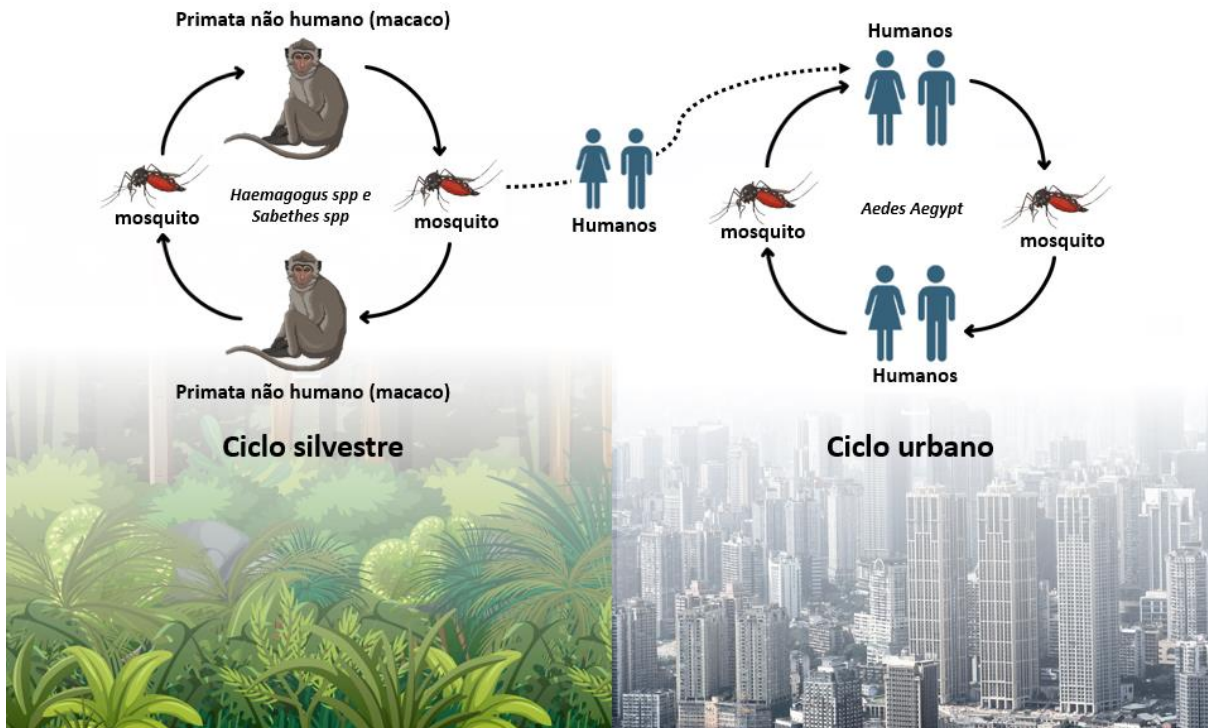


Figura 1 - Ciclos de transmissão do vírus da febre amarela.

Fonte: Elaborada pela autora com os dados de *CDC- Yellow Fever transmission* (<https://www.cdc.gov/yellowfever/transmission/index.html>), acesso em 10.04.23.

Clinicamente, tal como em outras viroses, a febre amarela apresenta-se de diversas maneiras, podendo ser desde assintomática a severa.¹⁹ Dentre os sintomas característicos estão: febre alta, calafrios, dor de cabeça, dor lombar, dor muscular generalizada, enjoos e vômito.¹⁹ Nos casos mais severos, pode ocasionar problemas renais, danos hepáticos, hemorragia, choque, falência múltipla de órgãos e alto risco de morte. Embora somente cerca de 15% dos casos se apresentem na forma mais grave, nestes casos a taxa de morte é extremamente alta, chegando a 60%.¹⁹ Até o momento não existe nenhum tratamento específico para a febre amarela. Nos casos mais graves, o paciente deve ser hospitalizado para maiores cuidados e observação.¹⁹

O isolamento do vírus possibilitou o desenvolvimento de uma vacina, produzida por Theiler e Smith e denominada 17D, utilizando o vírus atenuado isolado anteriormente, pertencente a linhagem Asib.²⁰ Nas décadas de 40 e 50, foram administradas 40 milhões de doses da vacina contra febre amarela em países do leste da África e o número de casos declinou consideravelmente.²⁰ Nos dias atuais, a vacina contra a febre amarela é uma das mais seguras e eficientes já desenvolvidas utilizando vírus atenuado.²⁰ Tem atuação baseada na imunização a longo prazo, que induz imunidade humoral e celular, com efetiva produção de anticorpos neutralizantes específicos contra a proteína E presente no envelope viral.²¹⁻²² A vacina inicial, do tipo 17D, foi melhorada e os tipos produzidos são derivados de duas sub-linhagens, a 17DD e a 17D-204 desenvolvidas no Instituto Rockefeller.²³

Apesar de raros, a vacina pode apresentar efeitos colaterais, e levando isto em conta, certos grupos da população não são aconselhados a se vacinar, como pacientes com HIV e que estejam utilizando imunossuppressores ou pessoas com alergia a ovos.²⁴ Como forma de prevenção de novas epidemias, a OMS recomenda a vacinação contra febre amarela em certas áreas consideradas de risco ou endêmicas, ou para aqueles que viajam para estas regiões.²⁵ Podemos citar alguns países africanos como Angola, Senegal e Gana e outros países tropicais como o Brasil, Argentina e Equador.²⁵

1.2 Situação atual da febre amarela no Brasil

No Brasil, após os esforços do Ministério da Saúde e uma campanha intensa de vacinação no início do século XX, a febre amarela urbana foi considerada erradicada e desde 1942 nenhum caso foi registrado no país.²⁶ Porém, o aumento da população não imunizada no país, somado ao aumento do contingente populacional circulando entre as áreas endêmicas, além de mudanças climáticas e desmatamento, possibilitaram o ressurgimento de um novo surto de febre amarela no país em 2017.²⁷ Esses fatores levaram pesquisadores a alertar o sistema de saúde brasileiro para o possível retorno da febre amarela em regiões urbanas.¹⁰

Entre dezembro de 2016 e fevereiro de 2017 um surto de febre amarela atingiu o Brasil, afetando inicialmente o estado de Minas Gerais. Esta foi a epidemia mais grave desde as aparições da doença nas décadas de 50 e 60. Neste período foram reportados 778 casos confirmados da doença, levando à morte 262 pessoas, com uma taxa de letalidade de aproximadamente 34%. A vacinação foi feita nas áreas mais afetadas, o que não impediu que o vírus se espalhasse para outras áreas, dentre elas o estado de São Paulo, Espírito Santo e Bahia.²⁸ No período sazonal do ano seguinte (entre dezembro de 2017 e maio de 2018), houve um rápido aumento no número de casos confirmados nos estados do Rio de Janeiro, São Paulo e Minas Gerais. Neste período foram registrados 1376 casos de febre amarela silvestre, e dentre eles 483 mortes.²⁹ No estado de São Paulo, de acordo com a Secretaria de Saúde, o número de casos confirmados até setembro de 2018 foi de 537, um valor cerca de 400% maior comparado aos 103 registrados no ano anterior.³⁰

Durante o monitoramento 2019/2020 foram registrados 881 casos suspeitos e 18 casos confirmados.³⁰ Já no ano de 2023, após 3 anos sem registros de casos da doença no Brasil, 2 mortes por febre amarela foram confirmadas no estado de São Paulo.³¹ Embora o número de casos tenha diminuído significativamente, neste ano, foram registradas 686 epizootias (casos de FA entre macacos), o que evidencia a circulação do vírus no país.³¹ Em conjunto, a baixa cobertura vacinal no país segundo o Ministério de Saúde, em torno de 50% enquanto a recomendada pela OMS para prevenção de novas epidemias é superior a 80%, leva a preocupação de um possível retorno da doença, especialmente no ambiente urbano.

Segundo dados atuais da OMS, estima-se que a cada ano há 200.000 casos de febre amarela, totalizando 30.000 mortes. No entanto, devido à dificuldade no diagnóstico, apenas uma pequena porcentagem dos casos é identificada e o número real de casos pode ser até 50 vezes maior que o notificado. Nesta atual conjuntura, fica evidente que, nos últimos anos, embora haja a existência de uma vacina como profilaxia para a febre amarela, o número de casos cresceu significativamente. Estes fatores, somados a inexistência de um fármaco para o tratamento da doença, levaram a OMS a classificar a febre amarela como situação de alerta de saúde pública deste fevereiro de 2017.

Considerando a gravidade da doença que atinge mais de 50% dos casos graves, a inexistência de um fármaco específico para seu tratamento, a baixa cobertura vacinal, somados a circulação endêmica silvestre do vírus, evidenciada pelos surtos recentes e pela constante aparição de casos de FA entre macacos, fica claro que é urgente a necessidade do desenvolvimento de um fármaco antiviral contra a febre amarela. Com a existência de um fármaco antiviral, aliado ao diagnóstico precoce da doença, seria possível evitar milhares de casos e mortes anualmente, reduzir a circulação viral e consequentemente evitar o aparecimento de variantes, assim como permitiria tratar e prevenir futuras epidemias.

1.3 Biologia estrutural do vírus e do complexo NS2B-NS3^{pro}

O vírus causador da febre amarela, *Yellow Fever Virus* (YFV) pertence à família *Flaviviridae* e ao gênero *Flavivirus*, que inclui também os vírus da dengue, Zika e Encefalite Japonesa, são responsáveis por uma série de enfermidades que acometem os humanos.³²

Os flavivírus tem genoma composto por uma única fita de RNA positiva, de aproximadamente 11.000 bases que codifica uma única fase aberta de leitura (ORF), a qual apresenta regiões que não são traduzidas (responsáveis pelo controle da tradução e replicação viral), e uma região responsável por codificar uma única poliproteína. A poliproteína é composta por três proteínas estruturais, formando o envelope, o capsídeo e a membrana viral; e sete proteínas não estruturais (NS1, NS2A, NS2B, NS3, NS4A, NS4B e NS5). A poliproteína é clivada pela NS2B-NS3 protease (NS2B-NS3^{pro}), juntamente com algumas proteases do hospedeiro, gerando as proteínas virais maduras (Figura 2).³³⁻³⁴

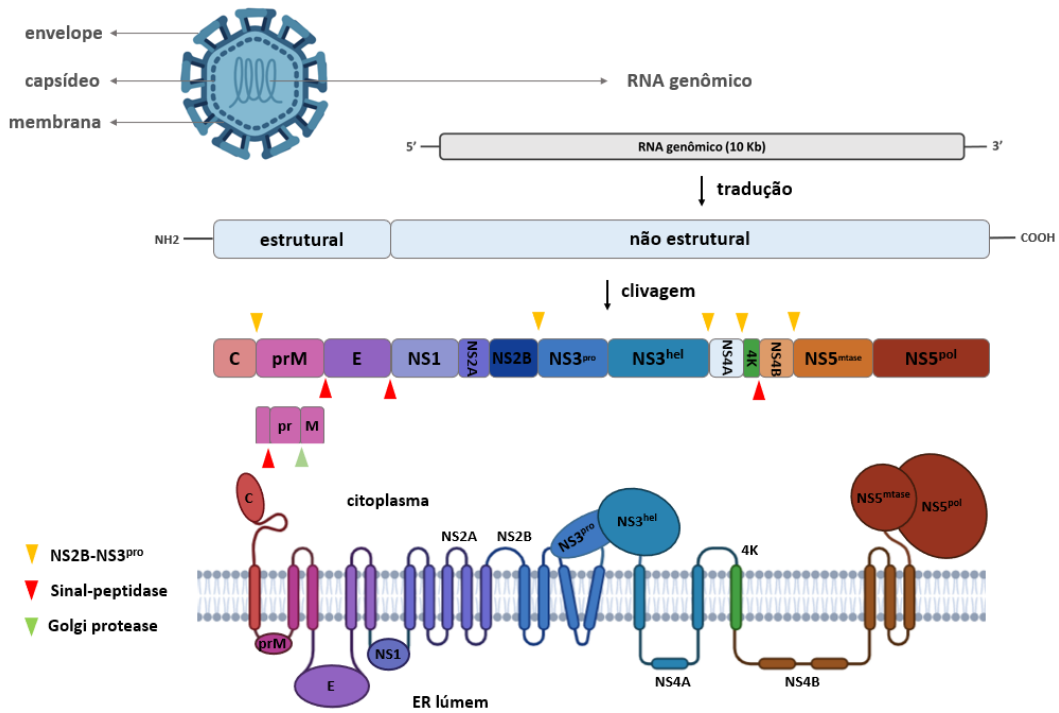


Figura 2 - Estrutura e expressão genômica dos flavivírus.

Fonte: Elaborada pela autora utilizando biorender.com.

A NS3 é uma proteína multifuncional, que contém um domínio serino-protease na região N-terminal e um domínio RNA-helicase/NTPase na região C-terminal.³⁴ Ambas as funções desempenham papel crucial no processo de replicação viral, por isso, é uma das proteínas não estruturais dos flavivírus mais bem caracterizadas e possui várias estruturas cristalográficas resolvidas.^{32,35-40} Hoje a maior parte dos inibidores de replicação viral de flavivírus tem como alvo ou a NS3 protease, ou a NS5 RNA-dependente RNA Polimerase.⁴¹⁻

O enovelamento e atividade catalítica da NS3 protease dependem de uma interação não covalente com outra NS, a NS2B (Figura 3). A NS3 protease é formada por dois barris- β , que contém a tríade de aminoácidos catalíticos (His, Asp e Ser) localizado entre eles. Além disso, a porção terminal da NS2B forma um β -hairpin que se adere a superfície da NS3 protease, formando o complexo NS2B-NS3^{pro} (Figura 3).⁴⁴ A NS2B-NS3 protease é descrita como podendo ser encontrada em duas conformações possíveis: aberta ou fechada. A conformação aberta, típica de estruturas da proteína na forma apo, é caracterizado pelo cofator NS2B não envolvendo a NS3 protease e formando uma região de hélice (Figura 3A). Já a conformação fechada, típica de estruturas da proteína em complexo com um ligante no sítio ativo (substrato ou inibidor), é caracterizado pelo cofator NS2B formando uma região de grampo- β em sua porção C-terminal, cujos resíduos desta volta- β interagem com o ligante presente no sítio ativo, estabilizando o cofator NS2B (Figura 3B).⁴⁴

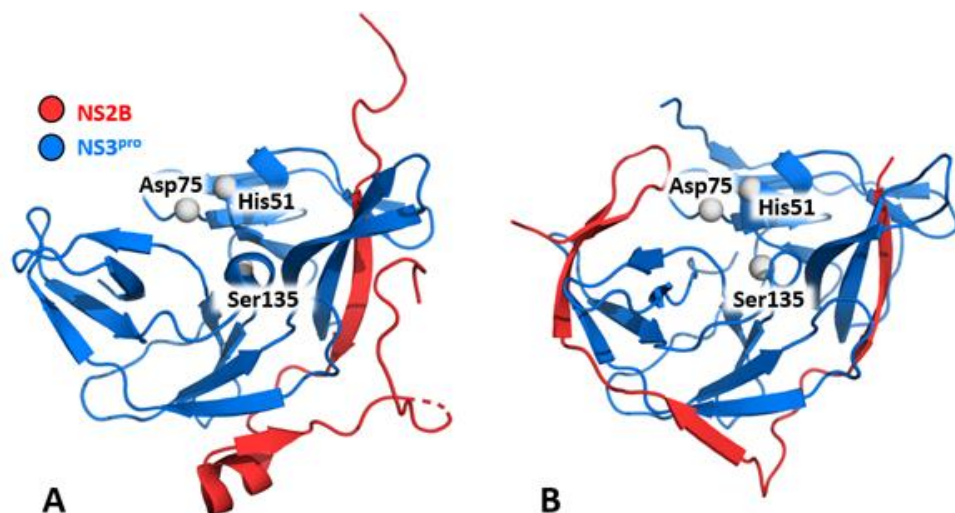


Figura 3 - Estrutura do complexo NS2B-NS3^{pro} de dengue tipo 2. Podemos ver o enovelamento do complexo NS2B (azul) com a NS3 (vermelho) na conformação aberta em A (PDB ID 2FOM), e na conformação fechada em B (PDB ID 3U1I). Em ambas temos a visão da tríade catalítica, com os resíduos His51, Asp75 e Ser135 representados em esferas (cinza).

Fonte: Elaborada pela autora.

O mecanismo de reação das proteases de flavivírus, assim como de outras serino-proteases, pertence ao tipo catálise ácido-básica em que a serina, resíduo que dá nome a família de enzimas, é responsável pelo ataque nucleofílico a carbonila da ligação peptídica.⁴⁵ O potencial nucleofílico da serina (Ser135) é aumentado pela interação com outros resíduos formadores da tríade catalítica, aspartato e histidina (Asp75 e His51) (Figura 4 – etapa 1).

Após o ataque nucleofílico realizado pela hidroxila da serina (Figura 4 – etapa 2), há formação do primeiro complexo intermediário tetraédrico, em que o *oxyanion hole* constituído

pelos agrupamentos amina dos resíduos Ser135, Thr134 e Gly133, estabilizam a carga negativa do oxigênio com formação de ligações de hidrogênio (Figura 4 – etapa 3).⁴⁵⁻⁴⁷ Este estágio de transição é desfeito com a doação de um próton proveniente da His51 para o grupo amino-terminal do produto 1 (Figura 4 – etapas 3 e 4). Simultaneamente com a liberação do produto 1, há formação de um complexo ligado covalentemente entre o substrato e a serina, denominado complexo acil-enzima. Para a liberação da enzima, é necessário o ataque nucleofílico realizado por uma molécula de água (Figura 4 – etapa 4), que gera de forma similar outro complexo intermediário tetraédrico, estabilizado pelo *oxyanion hole* (Figura 4 – etapa 5). Por fim, o complexo é desfeito, novamente com a doação de um próton da His51 para o agrupamento hidroxila da região carboxi-terminal do produto 2 (Figura 4 – etapa 6).⁴⁵

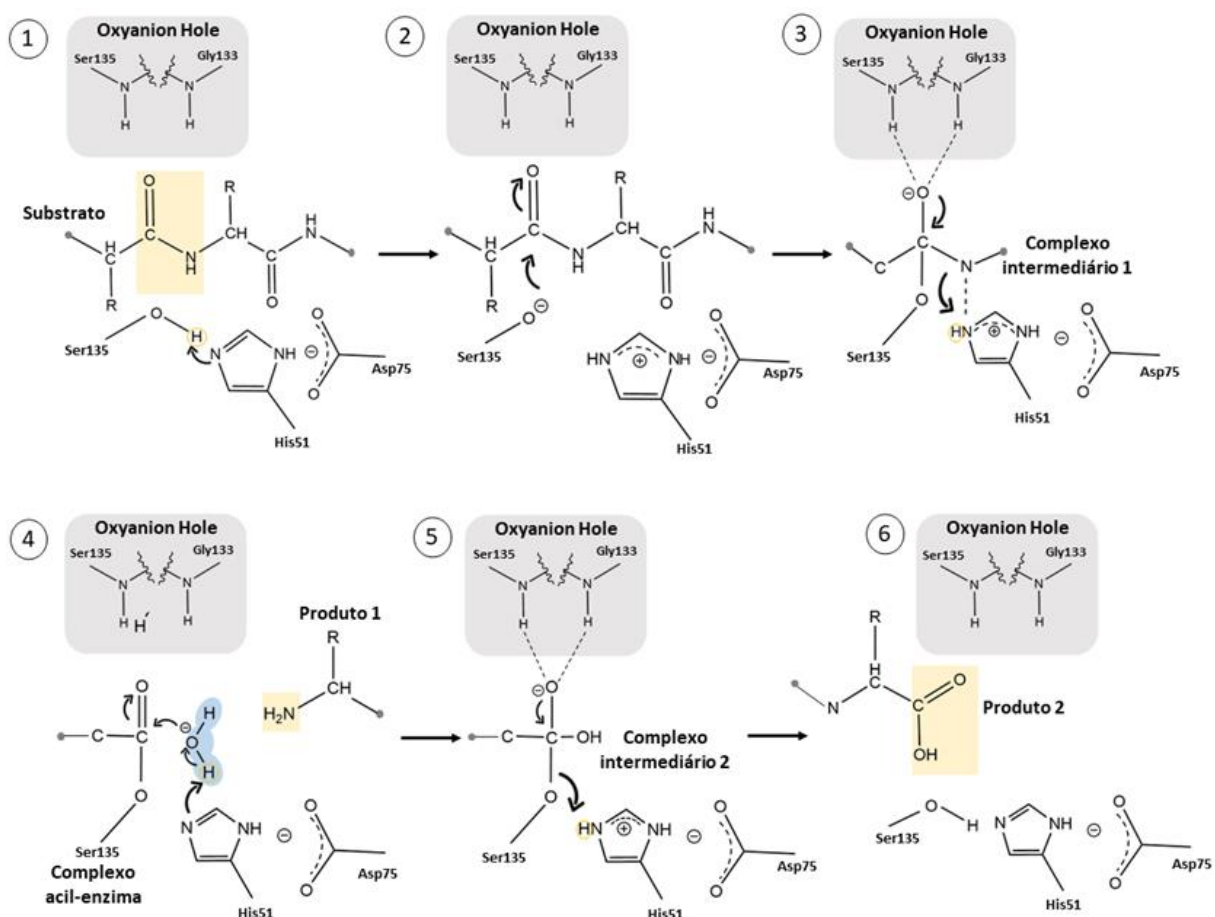


Figura 4 - Representação esquemática do mecanismo de catálise das proteases de Flavivírus. As numerações de 1-6 representam as etapas da reação.

Fonte: Adaptada de COPELAND; SALAEMAE *et al.*; NOBLE *et al.*⁴⁵⁻⁴⁷

Embora a NS3 seja bem caracterizada para vários flavivírus, no início deste projeto ainda não existia estrutura cristalográfica da NS2B-NS3^{pro} do vírus da febre amarela. Tendo isso em vista e a importância deste complexo no ciclo de replicação viral, é notório que esta representa um importante alvo para a produção de possíveis fármacos.

1.4 Polimorfismo gênico do surto de 2017 e possíveis implicações estruturais

A linhagem circulante do YFV durante o surto silvestre de 2017 no Brasil, foi sequenciada pela Dra. Myrna Bonaldo (Instituto Oswaldo Cruz, Brasil) e revelou algumas mutações pontuais, que podem ter relação com o aumento da virulência quando comparada às linhagens antecessoras.⁴⁸ Análises filogenéticas indicam que esta linhagem atual (circulante), originou-se da linhagem de Trinidad e Tobago que se disseminou pela Venezuela e pelo Brasil, e tem sido responsável por surtos não endêmicos detectados na América do Sul desde os anos 2000. O surto de 2017, portanto, foi possivelmente ocasionado por um acúmulo de mutações, que levou a alteração de alguns aminoácidos localizados principalmente nas NSs.⁴⁹

Dentre as mutações identificadas, cinco encontram-se na NS5 e duas na NS3 protease. Modelagens estruturais indicam que uma das mutações na NS3 protease (R121K) está localizada próximo ao local de interação com a NS2B e pode influenciar a interação entre as duas moléculas assim como interferir na afinidade de ligação do complexo NS2B-NS3^{pro} e, portanto, pode influenciar na atividade catalítica da enzima.⁵⁰

O estudo recentemente publicado, realizado com a enzima NS5 domínio metiltransferase, que contém 3 das mutações acima descritas (R101K, I138V e S173G), visou analisar a influência destas na atividade da enzima e como afetam a virulência e propagação do vírus.⁵¹ Os ensaios indicaram que as mutações não afetam significativamente a atividade da enzima, pois localizam-se próximas superfície a proteína e apenas conferem pequeno aumento da afinidade por um dos substratos da enzima (SAM) que não são suficientes para serem detectados em ensaios com modelos celulares de infecção viral.⁵¹ Desta forma, considerando que a NS5 metiltransferase está diretamente envolvida na resposta imune do hospedeiro, as mutações possivelmente implicam em modificações estruturais que interferem na via de produção de interferon tipo 1.⁵¹

Logo, a caracterização estrutural e funcional das NSs pertencentes a linhagem circulante de YFV são de suma importância para melhor compreensão do mecanismo de replicação viral, além de fundamentais no planejamento de candidatos a fármacos específicos e eficazes contra a febre amarela visando tratar e prevenir futuras epidemias.

1.5 NS2B-NS3^{pro} como alvo no desenvolvimento de antivirais e objetivos do trabalho

Nas últimas décadas, o complexo NS2B-NS3^{pro} de flavivírus tem sido amplamente utilizado como alvo para o desenvolvimento de potenciais candidatos a inibidor da replicação viral.⁵² Grande parte dos inibidores foram estudados para DENV, porém diversos estudos mostram o potencial destes inibidores de inibirem proteases de vários flavivírus como ZIKV e WNV, dado a conservação e similaridade estrutural entre as proteínas. Existem duas categorias de inibidores da NS2B-NS3^{pro}, os inibidores competitivos que se interagem com o sítio ativo da enzima e os inibidores não competitivos, que interagem em outro sítio que não o sítio ativo (alostéricos ou ortostérico), gerando alterações conformações após ligação ou interrompendo a interação com o cofator NS2B e assim inibindo a atividade da enzima.⁵²

Um dos primeiros inibidores descritos na literatura foi o trabalho de Leung e colab em 2001, que identificou a aprotinina, uma pequena proteína de aproximadamente 6.5 kDa, como potente inibidor da NS2B-NS3^{pro} de DENV (IC₅₀ de 65 nM).⁵³ A partir de 2010, diversos estudos realizados utilizando grandes bibliotecas de compostos para a triagem de inibidores de protease, dentre eles os peptídeos miméticos, peptídeos cíclicos, os derivados de carbazol, e as hidroxiquinolinas.^{52,54} Peptídeos sintéticos que mimetizam o substrato natural da enzima mostraram ser capazes de interagir com o sítio de ligação com o substrato e bloquear a atividade enzimática, no entanto, devido à instabilidade e baixa solubilidade dos peptídeos, os estudos mais recentes tem focado no desenvolvimento de pequenas moléculas.⁵⁴ Em um dos estudos mais recentes, do ano de 2021, Nie e colab descreveram uma série de compostos contendo indol, capazes de inibir as proteases de DENV, WNV e ZIKV (sendo o mais potente com IC₅₀ de 1.6, 5.7 e 0.32 μM, respectivamente).⁵⁵

Outros trabalhos ainda focaram em utilizar o método de reposicionamento de fármacos, triando bibliotecas de fármacos já existentes, como o trabalho de Li e colab. Utilizando um ensaio capaz de quantificar a interação entre NS2B e NS3^{pro} de ZIKV com a medida da atividade de uma luciferase, foram identificados três inibidores, temoporfina, niclosamida e nitazoxanida, com IC₅₀ de 1.1 μM, 12.3 μM e 15.9 μM. O composto mais potente, a temoporfina, também mostrou atividade inibitória na replicação de outros flavivírus como YFV, WNV e DENV2.⁵⁶

Embora diversos compostos já tenham sido descritos como potentes inibidores da enzima, grande parte deles não demonstra elevada eficácia quando testado em ensaios celulares. Dado a característica altamente negativamente carregada do sítio ativo das proteases de flavivírus, as moléculas capazes de inibir a atividade da enzima tendem a ter elevada carga positiva, e conseqüentemente baixa hidrofobicidade, que por sua vez tornam estes compostos de difícil permeabilidade na membrana plasmática, resultando em baixa biodisponibilidade. Em conjunto, estes fatores tornam a protease, mesmo sendo um alvo extremamente atrativo, um difícil alvo no desenvolvimento de potenciais antivirais, sendo necessário ainda muito avanço na área para obtenção de um candidato a inibidor que seja ao mesmo tempo eficaz contra a enzima e capaz de permear a membrana plasmática.⁵²

Os objetivos gerais deste projeto foram a obtenção da estrutura cristalográfica, a utilização de técnicas biofísicas e bioquímicas para elucidar os aspectos funcionais e estruturais do complexo NS2B-NS3^{pro} de duas linhagens de YFV e subsequentemente realizar a busca por ligantes específicos que possam ser utilizados para o desenvolvimento de um potencial fármaco antiviral.

Os objetivos específicos do projeto foram:

- Obtenção do complexo NS2B-NS3^{pro} da linhagem de YFV circulante no Brasil;
- Obtenção da linhagem antecessora (Venezuela, 2010) aplicando o método de mutagênese sítio-dirigida;
- Ensaios biofísicos de caracterização de ambas as proteínas em solução;
- Ensaios de cristalização e estudos estruturais de ambas as proteínas;
- Padronização e realização de ensaios de atividade para caracterização funcional e determinação de parâmetros cinéticos das proteínas selvagem e mutantes;
- Utilização do método de SBDD para triagem de ligantes utilizando o banco de moléculas do CIBFar;
- Realização de ensaios de inibição para triagem de ligantes utilizando o ensaio de atividade padronizado e determinação da potência inibitória dos hits encontrados;
- Ensaios de co-cristalização com os *hits* obtidos nas triagens realizadas;
- Utilização dos métodos de planejamento de fármacos para otimização dos ligantes obtidos;
- Ensaios de inibição *in-situ* utilizando o clone não-infeccioso de febre amarela (replicon)

2 MATERIAIS E MÉTODOS

2.1 Aquisição das sequências codificantes de YFV

A sequência da linhagem circulante do YFV está depositada no GenBank sob o código KY885001.1. As mutações observadas na sequência da proteína NS3 estão apresentadas na Figura 5A.⁴⁸ O cDNA contendo a sequência das proteínas NS2B-NS3^{pro} (3110 bases), utilizado como molde para PCR neste trabalho, foi obtido de amostras de macacos infectados no ambiente silvestre do Estado do Rio de Janeiro e gentilmente cedido pela Dra. Myrna C. Bonaldo (Instituto Oswaldo Cruz, Brasil) (Figura 5B).

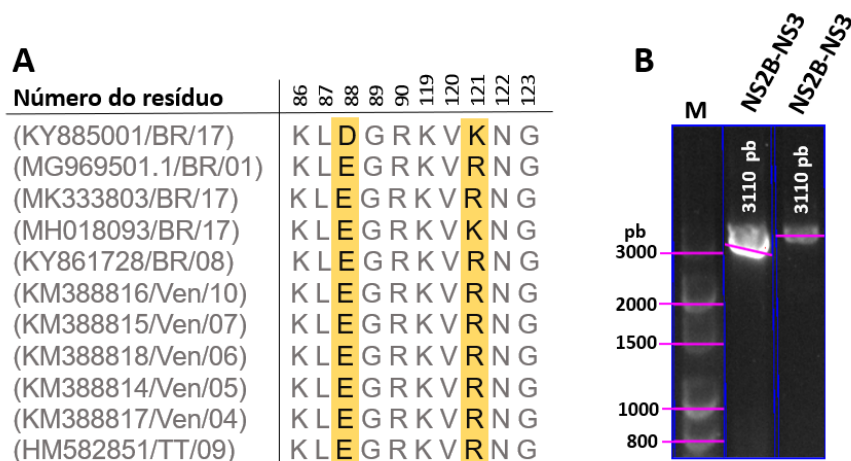


Figura 5 - A) Alinhamento das sequências de aminoácidos da proteína NS3 de diferentes linhagens de YFV. Os resíduos mutantes estão destacados em amarelo. B) Gel agarose 1% mostrando o cDNA contendo a sequência codificante das proteínas NS2B e NS3 (3110 pb).

Fonte: Adaptada de NOSKE *et al.*⁵⁷

2.2 Clonagem

2.2.1 Construções inicial e NS2B-NS3^{pro} linked

A clonagem do fragmento codificante da enzima NS2B-NS3^{pro} foi realizada em duas etapas no vetor pET_SUMO/LIC, que codifica a proteína de interesse fusionada a proteína SUMO e uma cauda de 6xHis na região N-terminal. A primeira delas foi realizada pelo método

LIC⁵⁸ e a segunda parte pelo método de *Gibson Assembly*⁵⁹ utilizando o kit comercial In-Fusion® (Takara Bio) para a adição de um *linker* (G₄SG₄) entre a porção NS2B e a NS3^{pro}, juntamente com a remoção da região transmembrana do cofator NS2B seguida da sequência de auto-clivagem da protease, de modo a obter a construção final (NS2B-NS3^{pro} *linked*). A construção final contém os resíduos 49-95 da NS2B, covalentemente ligados pelo *linker* G₄SG₄ aos resíduos 5-176 da NS3. Um esquema das construções utilizadas neste trabalho está representado na Figura 6.

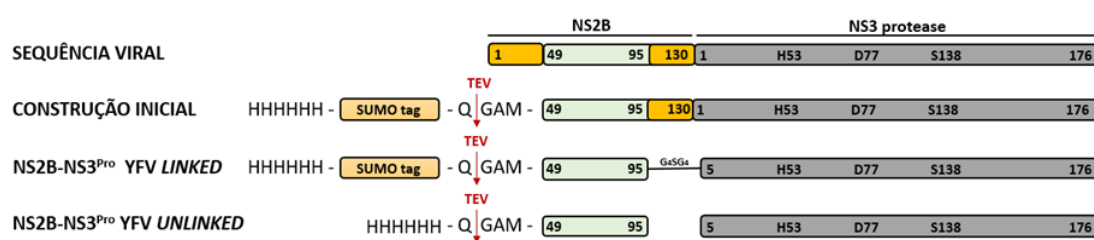


Figura 6 - Representação esquemática das construções da proteína NS2B-NS3^{pro} de YFV. A porção transmembrana do cofator NS2B está colorida em laranja e a porção solúvel e essencial para o enovelamento da NS3^{pro} está colorida em verde. Já o domínio NS3 protease está colorido em cinza.

Fonte: Elaborada pela autora.

Na primeira etapa da clonagem, a sequência correspondente aos resíduos 49-130 do cofator NS2B aos resíduos 1-176 da NS3 protease (construção inicial, Figura 6) foi realizada utilizando o método independente de ligação (LIC – “Ligation-Independent Cloning”).⁵⁸ Esse sistema é uma alternativa aos métodos tradicionais que utilizam enzimas de restrição e se baseia na atividade exonucleásica 3’→5’ da enzima T4 DNA Polimerase para criar regiões de fita simples complementares no inserto e vetor. No planejamento dos primers são adicionadas extensões apropriadas de tal modo que o fragmento amplificado produza extremidades coesivas complementares as do vetor após tratamento com T4 DNA polimerase. O pareamento do vetor com o inserto é então feito na ausência de ligase, simplesmente misturando-se os dois fragmentos de DNA (Figura 7A). Tal mistura foi utilizada para transformar bactérias *E. coli* DH5α competentes com o respectivo plasmídeo.

Ao planejar os primers para a reação, foi incluída a sequência 5’-CAGGGCGCCATG-3’ na extremidade 5’ de cada primer forward, e a sequência 5’-GACCCGACGCGGTTA-3’ na extremidade 5’ de cada primer reverse (Tabela 1). Os genes de interesse foram amplificados por PCR com a enzima Phusion® High-Fidelity DNA Polymerase (NEB) segundo as especificações do fabricante. Após a amplificação, os fragmentos foram tratados com a enzima

T4 DNA Polimerase (NEB) na presença de NEB buffer 2.1 1x, DTT 4.0 mM e dATP 2.5 mM. A reação foi incubada por 30 min a 22 °C e inativada por 20 min a 75 °C. O mesmo tratamento foi feito com o vetor, porém utilizando dTTP 2.5 mM, de forma a gerar um fragmento complementar as extensões dos primers.

Na segunda etapa de clonagem, os plasmídeos resultantes foram amplificados com os primers apresentados na Tabela 1 para inserção do *linker* (G₄SG₄) seguida da utilização do método de *Gibson Assembly*⁵⁹ com o kit comercial In-fusion® para pareamento das regiões complementares (Figura 7B). Em ambas as etapas, as colônias positivas foram selecionadas utilizando antibiótico canamicina 50 µg/mL, seguida de PCR de colônia e confirmadas por sequenciamento.

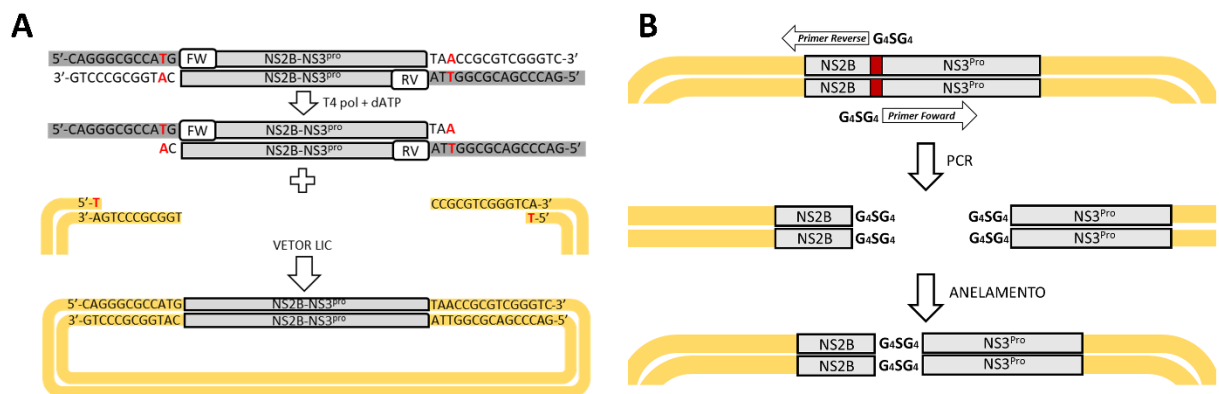


Figura 7 - Representação esquemática das etapas de clonagem utilizadas. A) Clonagem pelo método LIC. B) Clonagem pelo método de *Gibson Assembly*.

Fonte: Elaborada pela autora.

Tabela 1 - Primers planejados para clonagem da sequência codificante da proteína NS2B-NS3^{pro} e inserção do *linker*.

Primer	Sequência	T _m (°C)
ns2bns3lic_fw	5'-CAGGGCGCCATGGATGGACTTGAACCTCAGGAA-3'	57
ns2bns3lic_rv	5'-GACCCGACGCGTTACTCCTCTTACCTCAGTTTG-3'	62
linker_fw	5'-GGTGGTGGCGGCAGCGGTGGTGGCGGCTTTGGGACATCCCACAC-3'	63
linker_rv	5'-GCCGCCACCACCGCTGCCGCCACCACCATCCCATGGCACCTTCTC-3'	63

Fonte: Elaborada pela autora.

2.2.2 NS2B-NS3^{pro} *unlinked*

Baseando-se nas construções observadas nas estruturas depositadas no PDB para a enzima NS2B-NS3^{pro} de ZIKV em alta resolução,⁶⁰⁻⁶¹ foi proposta uma nova construção da proteína contendo os domínios NS2B e NS3 protease ligados de maneira não covalente (NS2B-

NS3^{pro} *unlinked*, Figura 5). Para isto, as sequências codificantes dos domínios NS2B e NS3 protease foram amplificadas a partir do plasmídeo molde (pET_SUMO/LIC) contendo a sequência ligada pelo *linker* covalente utilizando os primers apresentados na Tabela 2.

A clonagem foi realizada em duas etapas, ambas utilizando enzimas de restrição. Na primeira etapa foi realizada a clonagem da sequência que codifica o cofator NS2B no primeiro sítio de clonagem vetor pET_DUET-1. Para isto o vetor e o inserto correspondente a NS2B foram digeridos utilizando as enzimas de restrição BamHI (Cellco Biotec) e NotI (Cellco Biotec) em Universal buffer 1x (Cellco Biotec), a digestão foi verificada utilizando eletroforese em gel de agarose 1% e ambos inserto e vetor foram extraídos e purificados a partir do gel utilizando o kit comercial Agarose Gel Extraction Kit (Cellco Biotec) de acordo com as instruções do fabricante. Após a purificação, foi realizada a ligação através da ação da enzima T4 DNA Ligase (Thermo Fisher Scientific) seguida da transformação em células de *E. coli* DH5 α . As colônias positivas foram selecionadas com antibiótico ampicilina (100 μ g/mL) e a clonagem confirmada utilizando PCR de colônia. A segunda etapa de clonagem foi realizada seguindo o mesmo protocolo, partindo agora do vetor pET_DUET-1 já contendo a sequência do cofator NS2B para clonagem do inserto correspondente a NS3^{pro}. Nesta etapa foram utilizadas as enzimas de restrição NdeI (Jena Bioscience) e XhoI (Thermo Fisher Scientific). Novamente, as colônias positivas foram selecionadas com antibiótico ampicilina (100 μ g/mL) e a clonagem confirmada utilizando PCR de colônia, seguida de sequenciamento.

Tabela 2- Sequências dos primers planejados para a clonagem da nova construção no vetor pET_DUET-1.

Primer	Sequência	T _m (°C)
duetNS2Bfw	5'-CACAGCCAGGATCCGGAGAATCTTTATTTTCAGGGCG-3'	63
duetNS2Nrv	5'-GCATTATGCGGCCGCTTAATCCCATGGCACCTTCTC-3'	63
duetNS3fw	5'-GGAGATATACATATGCTTTGGGACATTCCC-3'	62
duetNS3rv	5'-TTTACCAGACTCGAGTTACTCCTCTTTCACCTCAGTTTG-3'	62

Fonte: Elaborada pela autora.

2.3 Mutação sítio dirigida

Para obtenção das proteínas mutantes K121R e D88E NS2B-NS3^{pro}, assim como o mutante duplo, foi utilizado o protocolo de mutação sítio dirigida. Ambos os mutantes foram obtidos partindo do plasmídeo molde pET_SUMO/LIC contendo a construção NS2B-NS3^{pro} *linked*. Para isto, foram planejados primers contendo o códon mutante, um deles contendo a sequência do códon mutada. Além disso, os primers foram planejados de modo a garantir a amplificação do plasmídeo molde em sentidos opostos pelo método de PCR inversa (Tabela 3).

Durante a reação de PCR, realizada pela enzima FastPol (Cellco biotec), a mutação é amplificada de forma exponencial, de forma que ao final da reação, torna-se o produto predominante. Em seguida foi realizado tratamento com a enzima DpnI (NEB) a 37°C por 5 h para degradação do DNA molde. Na última etapa, foi realizada uma reação contendo o produto tratado da PCR e a enzima T4 Polynucleotide Kinase (PNK) (Thermo Fisher Scientific) a qual realiza a fosforilação das extremidades 5' da fita de DNA, seguida da ação da enzima T4 DNA ligase (Thermo Fisher Scientific) que faz a re-circularização do plasmídeo. O produto obtido foi utilizado para transformar a linhagem de *E. coli* DH5 α . A presença das mutações foi confirmada por sequenciamento e a obtenção do duplo mutante foi realizada segundo o mesmo procedimento.

Tabela 3- Sequências dos primers planejados para mutagênese sítio dirigida.

Primer	Sequência	T _m (°C)
ns3yfvmutd_fw	5'-GGAAGTTGGAAGGTAGATGGGATG-3'	67
ns3yfvmutd_rv	5'-ATGAGCCGCCATAAGCAACA-3'	68
ns3yfvmutk_fw	5'-GCCTATTCCGCGTTAAGAATGG-3'	67
ns3yfvmutk_rv	5'-TTGGCTTTGTTTGGACATTCCTACTAC-3'	67

Fonte: Elaborada pela autora.

2.4 Testes de expressão e expressão em larga escala

Após confirmação da clonagem, os plasmídeos contendo as sequências de interesse foram propagados em *E. coli* DH5- α e isolados utilizando kit comercial Fast n' Easy Mini-Prep Extraction Kit (Cellco Biotec) de acordo com as recomendações do fabricante. Os plasmídeos resultantes foram utilizados para transformar as diferentes linhagens de expressão de *E. coli* *Rosetta* (DE3), *Rosetta 2* (DE3), *BL21* (DE3), *C41* (DE3) e *C43* (DE3), pelo método de choque térmico. Para os testes de expressão, as bactérias foram lisadas, a fração celular solúvel foi separada por centrifugação e a presença da proteína de interesse foi analisada por SDS-PAGE 15% ou por Western Blot, utilizando His Detect Solution (Jena Bioscience). Também foram utilizados outros meios de cultura, como autoindutor ZYM5052 (NH₄)₂SO 0,33%; KH₂PO₄ 0,68%; Na₂PO₄ 0,7%; Glicerol 0,5%; Glicose 0,05%; Lactose 0,2%; Triptona 1%; Extrato de Levedura 0,5%)⁶² e LB (triptona 1%, extrato de levedura 0,5% e NaCl 1%). Também foi variada a concentração de IPTG. As condições utilizadas no teste de expressão estão sumarizadas na Tabela 4.

Tabela 4- Condições avaliadas no teste de expressão

Linagem de <i>E. coli</i>	Meio de Cultura	Concentração de IPTG (mM)
<i>Rosetta</i>	LB	0.1
<i>Rosetta 2</i>		0.2
<i>BL21</i>	TB	0.5
<i>C41</i>	ZYM	0.6
<i>C43</i>		0.8

Fonte: Elaborada pela autora.

Para a expressão em larga escala da NS2B-NS3^{pro} YFV *linked*, foram preparados pré-inóculos de 5 mL em meio LB utilizando as células *E. coli Rosetta 2* (DE3) transformadas com os plasmídeos de interesse, contendo 50 µg/mL canamicina e 37 µg/mL cloranfenicol, e crescidos por 12 horas a 37°C sob agitação de 200 rpm. Cada pré-inóculo foi utilizado para inocular 1 L de meio TB, também contendo 50 µg/mL canamicina e 37 µg/mL cloranfenicol, os quais foram mantidos a 37°C sob agitação de 200 rpm até que atingissem DO₆₀₀ de 1.0. Para indução da expressão, foi adicionado 0.5 mM IPTG, seguido da alteração da temperatura para 18°C e manteve-se a agitação de 200 rpm por cerca de 16 h. Após a expressão, os meios foram centrifugados (5000 x g, 40 min, 8 °C) para sedimentação das células, e em seguida ressuspendidas em tampão PBS (137 mM NaCl, 2.7 mM KCl, 8 mM Na₂HPO₄, 2 mM KH₂PO₄, pH 7.4, glicerol 10%) utilizando-se 50 mL de tampão para cada litro de cultura. As células foram armazenadas em -20°C. Para a expressão da NS2B-NS3^{pro} YFV *unlinked* foi utilizado mesmo protocolo, substituindo o antibiótico canamicina por ampicilina (100 µg/mL)

2.5 Purificação das enzimas

Para a purificação da proteína NS2B-NS3^{pro} *linked* células foram lisadas em solução contendo lisozima (50 µg/mL), benzonase (10 U/mL), DTT (1.0 mM) e imidazol (30 mM). A mistura foi incubada em gelo por 30 min, seguida de sonicação (12 min, 30 s de pulso com 45 s de intervalo, amplitude 40). Adicionou-se benzonase novamente (10 u/mL) e repetiu-se a sonicação com o mesmo ciclo, por mais 4 min. Após a lise, a fração solúvel foi separada por centrifugação (15000 x g, 30 min, 4°C).

A primeira etapa de purificação foi realizada por cromatografia de afinidade (IMAC) passando a fração solúvel por uma coluna HisTrap HP (GE Healthcare) previamente equilibrada com tampão A (300 mM NaCl, 50 mM Tris pH 8.0, 30 mM Imidazol, 10% glicerol). A coluna foi lavada com 10 volumes de coluna de tampão A e a proteína de interesse foi eluída com

gradiente de tampão B (300 mM NaCl, 50 mM Tris pH 8.0, 300 mM Imidazol, 10% glicerol). A proteína de fusão SUMO foi removida realizando diálise por cerca de 16 h, a 4°C na presença de 3 mg de TEV protease e 4 mM de DTT. As frações foram então separadas por uma nova passagem pela coluna HisTrap, e desta vez a proteína foi obtida na fração que atravessou diretamente a coluna.

A última etapa de purificação foi realizada por cromatografia de exclusão molecular (gel filtração) para separação por tamanho. Para isto foi utilizada uma coluna HiLoad Superdex 75 16/60 (GE Healthcare) equilibrada com tampão C (20 mM HEPES, 500 mM NaCl, 5% glicerol, pH 7.5).

Para a purificação das enzimas mutantes, utilizamos o mesmo protocolo descrito acima.

Para a NS2B-NS3^{pro} *unlinked* foi utilizado o mesmo protocolo de purificação, com exceção da primeira afinidade em que ao invés de utilizarmos coluna HisTrap, o extrato celular solúvel foi incubado com 5 mL de resina Ni-NTA Superflow (Qiagen) sob agitação por 1 h a 4°C.

A pureza final das amostras foi analisada por SDS-PAGE 15% e quantificadas utilizando a absorção em 280 nm, medidas em um Nanodrop 1000 (Thermo Fischer Scientific) e calculadas pelo coeficiente de absorção teórico de 47,440 M⁻¹.cm⁻¹. Alíquotas a 1 mg/mL foram congeladas em nitrogênio líquido e armazenadas a -80°C para a realização de ensaios de atividade e ensaios biofísicos.

Para cristalização NS2B-NS3^{pro} *linked*, duas condições para a enzima foram utilizadas. Na primeira, a enzima foi concentrada a 38 mg/mL utilizando um concentrador de 10 kDa MWCO (Vivaspin, Sartorius) e na segunda o tampão da proteína foi trocado utilizando concentrador para 0.1 M Bis-Tris pH 6.5, 250 mM NaCl e então concentrada a 20 mg/mL. Para cristalização de enzima NS2B-NS3^{pro} *unlinked*, a enzima foi concentrada a 20 mg/mL.

2.6 Cristalização, co-cristalização e coleta de dados

Os cristais da proteína NS2B-NS3^{pro} YFV (*linked* e *unlinked*) foram inicialmente obtidos utilizando método de matriz esparsa com os kits comerciais Morpheus (Molecular dimensions), Index HT (Hampton Research), PACT++ (Qiagen), JCGS (Jena Bioscience), Crystal Screen (Hampton Research), PEGs I e II (Qiagen) e o robô de cristalização Crystal Gryphon (ARI®), em seguida os cristais foram otimizados a fim de obter cristais com melhor difração para a obtenção da estrutura cristalográfica da enzima. Para as otimizações foi utilizado o método de *hanging drop* e *sitting drop* em condições derivadas das quais foram obtidos

cristais. Também foi utilizada a técnica de *microseeding* e a cristalização na presença de aditivos. Além disso, realizamos experimentos de cristalização em diferentes temperaturas (25, 18 e 4°C) e diferentes proporções de proteína e solução precipitante.

Os cristais otimizados foram então transferidos para uma solução crio-protetora (solução do poço contendo 8-15% etilenoglicol), e em seguida crio-resfriados para a coleta, que foram realizadas sob fluxo constante de nitrogênio na linha de luz MX2 no Laboratório Nacional de Luz Sincrotron, Campinas, SP e na linha de luz I24 no Diamond Light Source, Oxfordshire, Reino Unido.

Para a obtenção de complexos proteína-ligante, foi utilizada o método de co-cristalização e a construção NS2B-NS3^{pro} *linked*. A proteína foi concentrada até 1 mg/mL, foi adicionado ligante em proporção molar 3:1 e, após incubação a 4°C por 1 h a amostra foi então concentrada para 40 mg/mL. As condições de cristalização foram derivadas da condição contendo 0.1 M Tris pH 8.5, 0.1 M NaCl e 25% PEG 3350. Os cristais foram crio protegidos da mesma forma descrita para a proteína apo. A coleta foi realizada na linha de luz MANACÁ, no SIRIUS, Campinas, Brasil.

2.7 Resolução das estruturas

Os conjuntos de dados coletados foram integrados e escalonados com o programa XDS⁶³, onde os limites de resolução foram estabelecidos pelo método de Karplus.⁶⁴⁻⁶⁵ O programa AIMLESS do CCP4 suite⁶⁶ foi utilizado para conversão das intensidades em amplitudes e cálculo das estatísticas. A qualidade dos dados foi analisada com o programa phenix.Xtriage.⁶⁷

Para a resolução da estrutura da NS2B-NS3^{pro} foi utilizada a substituição molecular para cálculo das fases, utilizando como modelo a NS2B-NS3^{pro} de dengue (PDB ID 2FOM, identidade de 56%). Para isso, utilizamos o programa PHASER,⁶⁸ procurando por quatro cópias do modelo na unidade assimétrica, baseado no coeficiente de Matthews obtido utilizando Matthews_coef do CCP4 suite.⁶⁶ A construção do modelo inicial da estrutura foi conduzida pelo programa Autobuild.⁶⁹ O refinamento foi realizado utilizando phenix.refine⁷⁰⁻⁷¹ e a estrutura foi manualmente reconstruída utilizando Coot.⁷²⁻⁷³ A validação da estrutura foi realizada utilizando o Molprobit.⁷⁴ As figuras foram feitas com uso do programa Pymol (Schrödinger, LLC).

2.8 Caracterização das proteínas mutantes

As enzimas mutantes foram inicialmente caracterizadas pela obtenção de suas respectivas temperaturas de *melting* (T_m). A T_m foi obtida por *Differential Scanning Fluorimetry* (DSF) que foi realizado na presença da sonda fluorescente SYPRO Orange 5x (Sigma Aldrich), a qual se liga em regiões hidrofóbicas de proteínas expostas ao longo da desnaturação térmica. A desnaturação térmica foi realizada no equipamento de PCR em tempo real Stratagene Mx3005P (Agilent Technologies), partindo da temperatura de 25°C até 74°C com acréscimo de 1°C por ciclo. A leitura de fluorescência foi realizada com excitação de 490 nm e 530 nm de emissão. Os dados obtidos foram analisados com o software Origin® 2018 (OriginLab), e a partir do *fitting* das curvas por um ajuste de Boltzmann obtivemos os valores das temperaturas de *melting* (T_m) referente a cada uma das proteínas. As figuras foram geradas no software GraphPad Prism 7 (GraphPad Software Inc.).

2.9 Ensaios de atividade

Os ensaios de atividade foram realizados baseados no protocolo descrito para Zika vírus.⁷⁵ O ensaio se baseia na utilização de um substrato peptídico Bz-nKRR-AMC (International Peptides) que contém a sequência de reconhecimento da enzima, a qual após clivagem libera 7-amino-4-metilcumarina (AMC), permitindo que a velocidade da reação seja monitorada através da medida de fluorescência (Figura 8). Para a padronização deste ensaio, cada reação enzimática foi realizada em tampão 20 mM Tris pH 8.5, 10% glicerol e 0.01% Triton X-100, contendo substrato Bz-nKRR-AMC em concentração final de 30 μ M. As reações foram iniciadas após adicionar protease em concentrações variando de 30 nM à 1.875 nM. Após a padronização, foi fixada a concentração de enzima de 7 nM para os ensaios subsequentes. A reação foi mantida a temperatura constante de 37°C, com a fluorescência sendo monitorada a cada 60 s durante 30 min. As reações foram feitas em placas Corning® brancas de 96 poços. Os comprimentos de onda utilizados para excitação e emissão foram respectivamente 380 nm e 460 nm. Todos os experimentos foram realizados em triplicatas. As medidas de fluorescência foram realizadas no equipamento SpectraMax Gemini EM Microplate Reader (Molecular Devices).

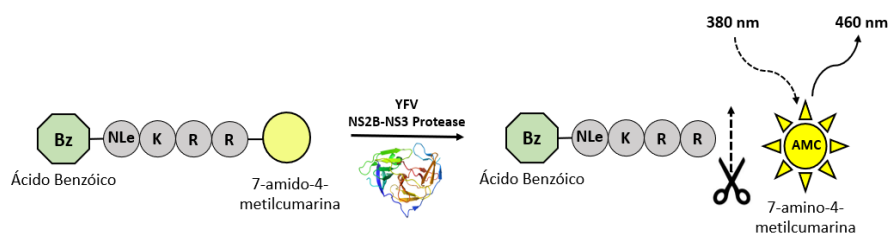


Figura 8 - Representação esquemática do ensaio enzimático da proteína YFV NS2B-NS3^{pro}. O substrato Bz-nKRR-AMC é clivado pela enzima liberando AMC, cuja fluorescência pode ser monitorada nos comprimentos de onda apresentados.

Fonte: Elaborada pela autora.

Para a determinação dos parâmetros cinéticos, a concentração de substrato foi variada (iniciando de 600 μM até 18.8 μM), de modo a se obter uma curva de concentração de substrato versus velocidade inicial de reação, a qual foi utilizada para determinação do K_m , k_{cat} e V_{max} com base na equação de Michaelis-Menten do *fitting* realizado no software Origin® 9 (Origin Lab). Os valores obtidos em fluorescência foram convertidos para concentração de produto a partir de uma curva de calibração a obtida utilizando concentrações conhecidas de AMC. As figuras foram geradas utilizando o software GraphPad Prism 7 (GraphPad Software Inc.).

2.10 Ensaios de inibição

Para a realização da triagem de inibidores, a proteína foi diluída em tampão de reação para concentração final de 7 nM, e foram adicionados ligantes para concentração final de 20 μM 1% DMSO em cada reação, que foi então incubada à 4°C por 30 min. As reações foram iniciadas com adição do substrato Bz-nKRR-AMC para uma concentração final de 30 μM . Todos os outros parâmetros do ensaio foram mantidos como descrito na seção anterior (seção 2.9) e foi utilizada a construção NS2B-NS3^{pro} *linked*. Os inibidores foram selecionados a partir de sua respectiva porcentagem de inibição, calculada utilizando uma reação controle, contendo proteína e 1% DMSO e uma reação branco contendo tampão de reação e 1% DMSO, localizados respectivamente na primeira e última coluna das placas testadas (equação 1).

Os valores utilizados nos cálculos foram obtidos pelos valores de fluorescência de cada uma das reações no tempo de 20 minutos. Este ensaio foi realizado na forma de ponto único. Foram testados cerca de 1000 compostos, dentre eles o Pandemic Response Box (PRB) fornecido pela organização *Medicines for Malaria Venture* (MMV), uma série de peptídeos sintéticos sintetizados e fornecidos pelo colaborador do CIBFar prof. Dr. Eduardo Maffud Cilli (Universidade Estadual Paulista "Júlio de Mesquita Filho", UNESP, Araraquara, Brasil) e uma

série de porfirinas sintetizadas e fornecidas pelo prof. Dr. Kleber Thiago de Oliveira (Universidade Federal de São Carlos, UFSCar, São Carlos, Brasil).

$$\text{Porcentagem de inibição (\%)} = \left(1 - \frac{\text{reação com inibidor (RFU)} - \text{branco (RFU)}}{\text{controle (RFU)} - \text{branco (RFU)}}\right) \times 100 \quad (1)$$

A validação do ensaio foi realizada a partir do cálculo do valor de Z-fator para cada uma das placas testadas, utilizando os desvios padrão das reações controle (σ_c) e branco (σ_b) e suas respectivas médias (μ_c e μ_b) de acordo com a equação 2. Para ensaios enzimáticos, valores de Z-fator acima de 0.6 significam um ensaio confiável.⁷⁶⁻⁷⁷

$$Z - \text{fator} = 1 - \frac{3(\sigma_c + \sigma_b)}{|\mu_c - \mu_b|} \quad (2)$$

Os compostos que apresentaram inibição superior a 80% foram selecionados para obtenção de curvas de dose-resposta. Para isso, os compostos foram diluídos serialmente e o ensaio foi realizado da maneira descrita anteriormente, assim como os cálculos das porcentagens de inibição. Os ensaios foram realizados em triplicata. Os valores de IC₅₀ foram obtidos a partir do *fitting* de Hill utilizando o software Origin 9 (OriginLab).

Ensaio para determinação do mecanismo de inibição, foram obtidos com diluição seriada do substrato (de 600 μM até 18.8 μM) e variação da concentração de inibidor, sendo esta fixa para cada uma das curvas de substrato. O mecanismo de inibição foi obtido pela análise dos parâmetros cinéticos (K_m , k_{cat} e V_{max}), obtidos do fitting de Michaelis-Menten e dos gráficos de Lineweaver-Burk. O ensaio foi realizado em triplicata e as análises gráficas obtidas utilizando o software Origin 9 (OriginLab). As figuras foram geradas utilizando o software GraphPad Prism 7 (GraphPad Software Inc.).

2.11 Ensaio antiviral baseado no replicon de YFV

Os compostos com resultados mais promissores nos ensaios enzimáticos foram testados quanto à atividade antiviral utilizando a linhagem celular BHK21-RepYFV-17D.

Esta linhagem celular, expressando de forma persistente a construção gênica do replicon de YFV contendo a sequência do gene repórter Firefly luciferase (Fluc) e do gene de resistência à geneticina, foi gentilmente cedida pela Profa. Dra. Laura H. V. G. Gil (Instituto Aggeu Magalhães, IAM-Fiocruz, Pernambuco, Brasil). Os ensaios antivirais foram realizados em

colaboração pela pós-doutoranda do grupo Dra. Rafaela Sachetto Fernandes conforme descrito em LI et al, 2018.⁷⁸⁻⁷⁹

Aproximadamente 2×10^4 células por poço foram semeadas em meio DMEM (Dulbecco's Modified Eagle Medium – GIBCO) contendo 10% soro bovino fetal (SFB) em placas de 96 poços (Corning®) transparentes tanto para o ensaio de atividade da luciferase quanto para o teste de citotoxicidade. Após 16 h de incubação à 37 °C 5% CO₂, o meio foi substituído por meio fresco contendo 2% SFB, para evitar a ligação dos compostos às proteínas do soro. Os compostos foram então adicionados às células em concentração única ou diluição seriada (fator 2). As células foram tratadas com 1% DMSO como controle negativo. Após 48 h de incubação, as células foram lisadas e submetidas ao teste de atividade da Firefly luciferase utilizando o kit Luciferase Assay System (Promega) ou tiveram viabilidade celular analisada utilizando o ensaio de proliferação baseado em MTT.⁷⁸ As concentrações de compostos, capazes de inibir 50% da atividade da Fluc e causar 50% de citotoxicidade (EC50 e CC50) foram estimados utilizando o software Origin 9.0 (OriginLab) e as figuras foram geradas utilizando o software GraphPad Prism 7 (GraphPad Software Inc.).

3 RESULTADOS

3.1 Construção NS2B-NS3^{pro} YFV *linked*

3.1.1 Clonagem

Os resultados iniciais obtidos para os testes de expressão e purificação da NS2B-NS3^{pro}, mostraram baixa solubilidade da enzima. Dessa forma, o rendimento na etapa de purificação foi muito baixo e não seria então possível obter proteína suficiente para posteriores ensaios de cristalização. Visando solucionar o problema de auto clivagem, foi realizada a inserção de um *linker* flexível (G₄SG₄) entre os dois domínios. Vale também ressaltar que, no processo de inserção do *linker*, foi possível realizar a retirada da região transmembrana que diminuía a solubilidade da proteína e dificultava a purificação. As predições realizadas pelo servidor TMHMM⁸⁰ mostraram a presença de uma região transmembrana na construção inicial e ausência desta na construção final, que mantém apenas a porção solúvel do cofator NS2B, essencial para enovelamento e atividade catalítica da NS3 protease (Figura 9).

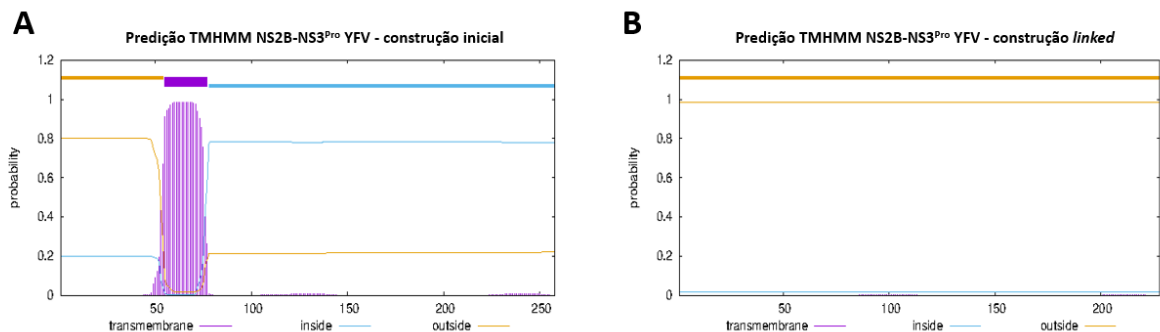


Figura 9 - Predições servidor TMHMM para as construções da enzima NS2B-NS3^{pro} de YFV.

Fonte: Elaborada pela autora.

O resultado da clonagem para a construção final contendo o *linker* foi feito pela análise da PCR de colônia em gel de agarose 1% e através de sequenciamento (Figura 10A). Como pode-se observar no gel de agarose, foi possível obter a construção pretendida, para as colônias 1-5 utilizadas para a PCR. Sendo assim, utilizamos uma das colônias contendo a construção desejada para propagação e isolamento do vetor e posterior transformação bacteriana.

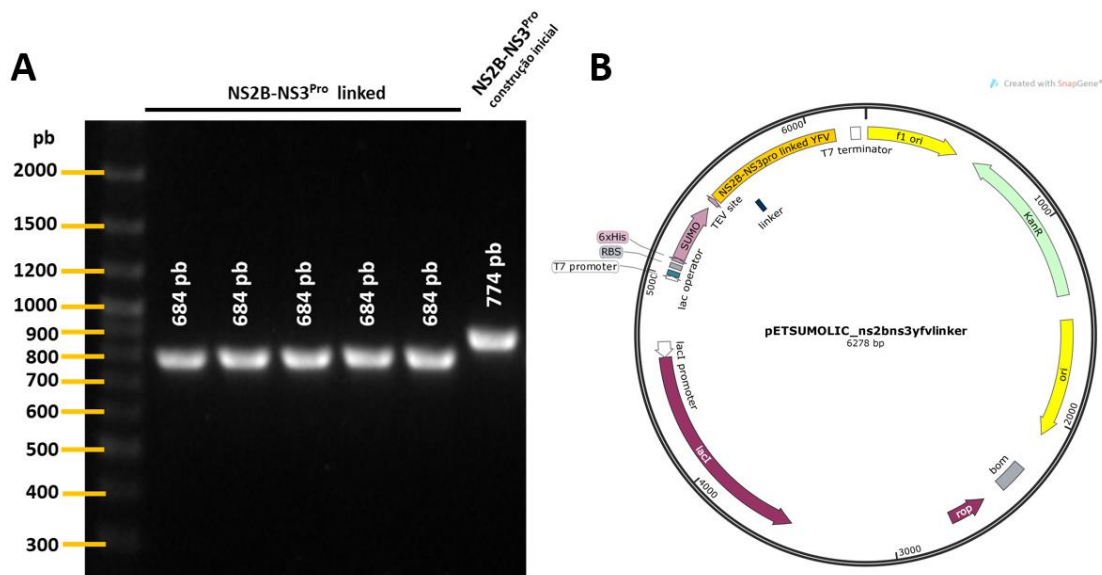


Figura 10 - A) Gel agarose 1% contendo os fragmentos obtidos para a construção com *linker*. Em M vemos o marcador molecular (em pb), nas amostras de 1-6 vemos cada uma das colônias testadas. As amostras 1-5 contêm a sequência com *linker* e a amostra 6 contém a sequência denominada de construção inicial (NS2B-NS3^{Pro} *linker* = 684 pb; NS2B-NS3^{Pro} construção inicial = 774 pb). B) Mapa do plasmídeo pETSUMO/LIC contendo a construção NS2B-NS3^{Pro} *linked*.

Fonte: Elaborada pela autora.

3.1.2 Testes de expressão

Após a transformação bacteriana das diversas linhagens de expressão de *E. coli*, foi realizado um teste inicial de expressão com intuito de detectar a presença da proteína de interesse na fração solúvel das amostras. Também foi variado o meio de cultura utilizado na expressão. Utilizamos a técnica de Western Blot revelado com solução capaz de detectar a cauda de histidinas (His-Tag Detect solution, Jena Bioscience). A maior quantidade de proteína na fração solúvel foi identificada na linhagem *Rosetta 2* (Figura 11). Sendo assim, esta linhagem foi selecionada para a expressão em larga escala. Também foi possível observar que havia maior quantidade de proteína expressa no meio de cultura TB, com utilização de 0.5 mM de IPTG para indução a 18°C por 20h. Portanto, como havia quantidade suficiente de proteína presente no extrato celular solúvel, essas condições foram utilizadas na expressão em larga escala.

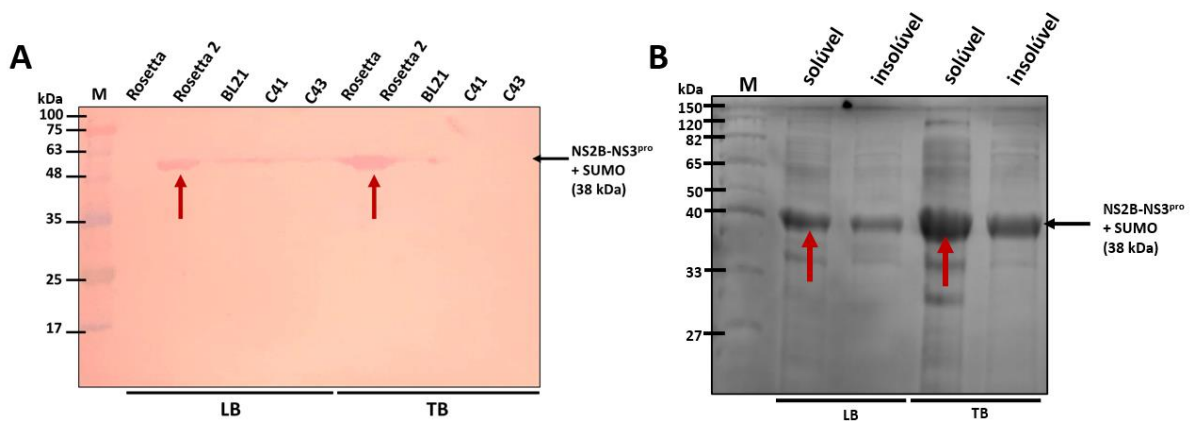


Figura 11 - A) *Western Blot* contendo a fração solúvel de diversas linhagens de *E. coli* para o teste de expressão realizado em dois meios diferentes (LB e TB). A proteína é detectada pela cauda de histidina e aparece em rosa, as setas vermelhas indicam as linhagens com maior taxa de expressão (NS2B-NS3^{pro} + SUMO = 38 kDa). B) SDS-PAGE 15% contendo ambas as frações solúveis e insolúveis obtidas para a linhagem de *E. coli* Rosetta 2, nos dois meios diferentes (LB e TB).

Fonte: Elaborada pela autora.

3.1.3 Mutagênese sítio dirigida

Visando compreender melhor o impacto das mutações presentes na protease da linhagem circulante do vírus da febre amarela, aplicamos mutagênese sítio dirigida para reverter esta linhagem a sua antecessora (Venezuela, 2002). As enzimas mutantes K121R, D88E, assim como o duplo mutante foram obtidos com sucesso utilizando o protocolo de mutagênese sítio dirigida. Para cada um deles cinco colônias foram selecionadas. Após a confirmação da mutação D88E, este plasmídeo foi usado como molde para amplificação com o primer planejado para o mutante K121R a fim de obter-se o mutante duplo. Todas as mutações foram confirmadas por sequenciamento (Figura 12).

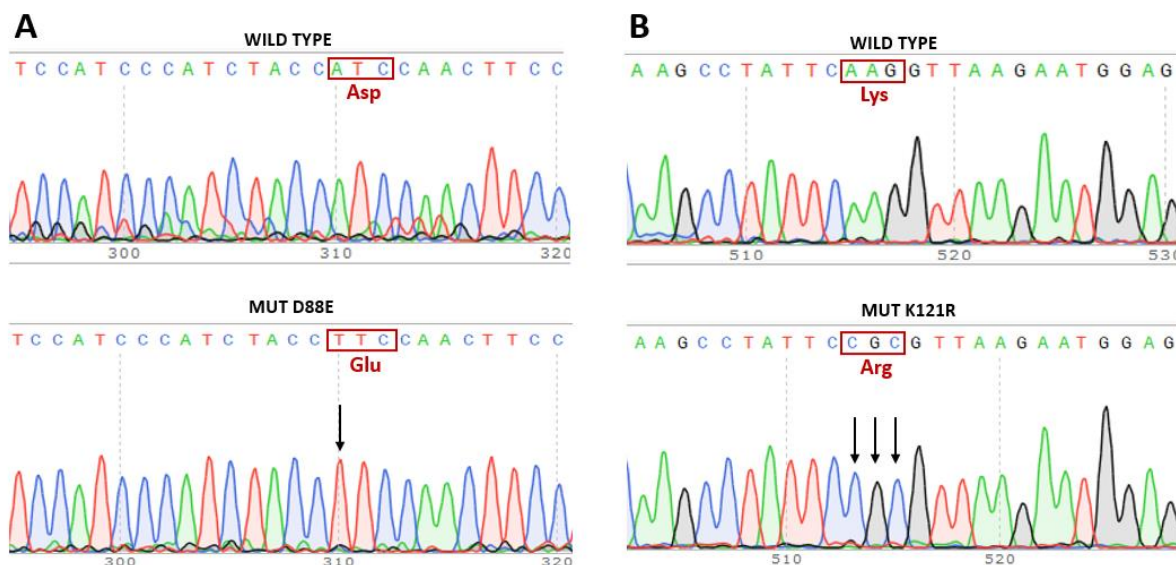


Figura 12 - Cromatogramas contendo os resultados do sequenciamento dos plasmídeos contendo a sequência codificante das enzimas NS2B-NS3^{pro} mutantes D88E, K121R e mutante duplo.
Fonte: Elaborada pela autora.

3.1.4 Expressão em larga escala e purificação das enzimas

As proteínas NS2B-NS3^{pro} WT, K121R, D88E e duplo mutante foram expressas em larga escala utilizando a linhagem de expressão de *E. coli Rosetta 2* em meio TB, de acordo com as condições otimizadas dos testes de expressão. Os resultados da purificação foram analisados por SDS-PAGE e estão apresentados nas Figura 13A e Figura 14B. O primeiro passo de purificação foi realizado por cromatografia de afinidade, onde a proteína de interesse foi eluída com gradiente de tampão B, contendo 300 mM de imidazol. A concentração de tampão B foi fixada em 20% (60 mM imidazol), na qual já era possível obter a proteína, de forma a evitar contaminantes e excesso de imidazol na amostra (Figura 13). Em seguida a proteína foi clivada da 6-His-SUMO, com a enzima TEV protease durante a diálise. Por fim a proteína foi submetida novamente a uma cromatografia de afinidade por níquel, mas desta vez coletada na fração que atravessou a resina.

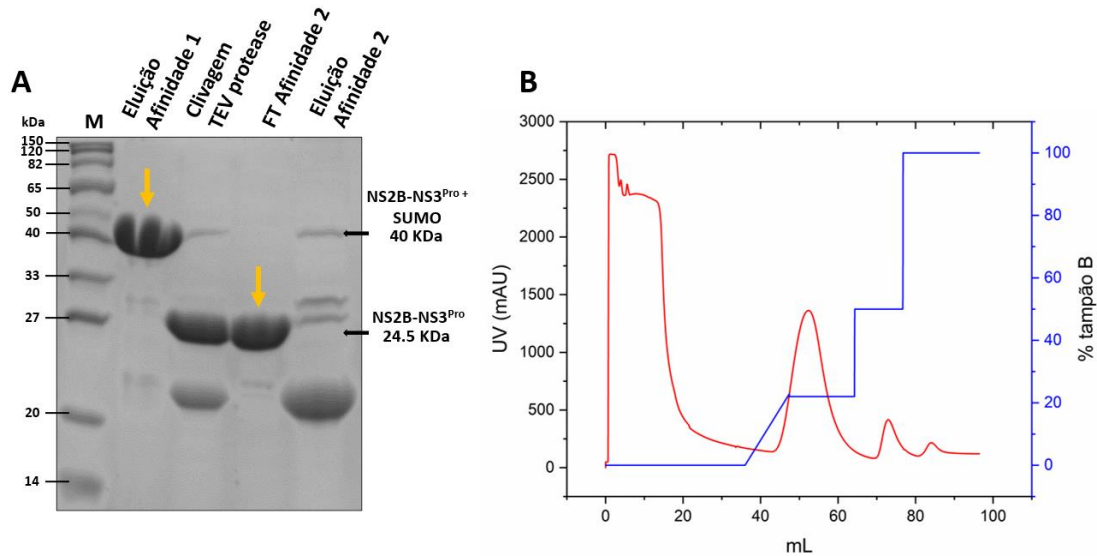


Figura 13 - A) SDS-PAGE 15% contendo as amostras obtidas nas cromatografias de afinidade e clivagem para a enzima NS2B-NS3^{Pro} YFV WT. Em M, vemos o marcador de peso molecular, seguido respectivamente pelas frações obtidas na eluição da primeira afinidade, após clivagem pela TEV protease, proteína já clivada obtida no FT da segunda cromatografia de afinidade e a amostra eluída da segunda afinidade, contendo contaminantes, além de proteína não clivada, TEV protease e a cauda 6-His-SUMO. Ambas as proteínas antes e após clivagem estão indicadas pelas setas, com seu respectivo peso molecular. B) Cromatograma obtido na primeira etapa de cromatografia de afinidade para a enzima NS2B-NS3^{Pro} YFV WT. Em vermelho, está representado a absorbância em 280 nm e em azul a concentração do tampão B (contendo 300 mM de imidazol).

Fonte: Elaborada pela autora.

Embora seja possível observar que a maior parte dos contaminantes foi removida com as etapas de cromatografia de afinidade, para obtermos uma amostra suficientemente pura e homogênea para realização de ensaios de cristalização e atividade, ao final foi realizado uma etapa adicional de purificação para a realização de separação por tamanho molecular (Figura 14A). A fim de determinar a massa molecular da proteína de interesse e seu estado oligomérico, ainda na etapa de cromatografia de exclusão molecular, foi feita uma curva de calibração utilizando um kit comercial Gel Filtration LMW Calibration Kit (GE Healthcare) (Figura 14C). A partir desta curva, foi possível determinar a massa molecular da NS2B-NS3^{Pro} (Figura 14D).

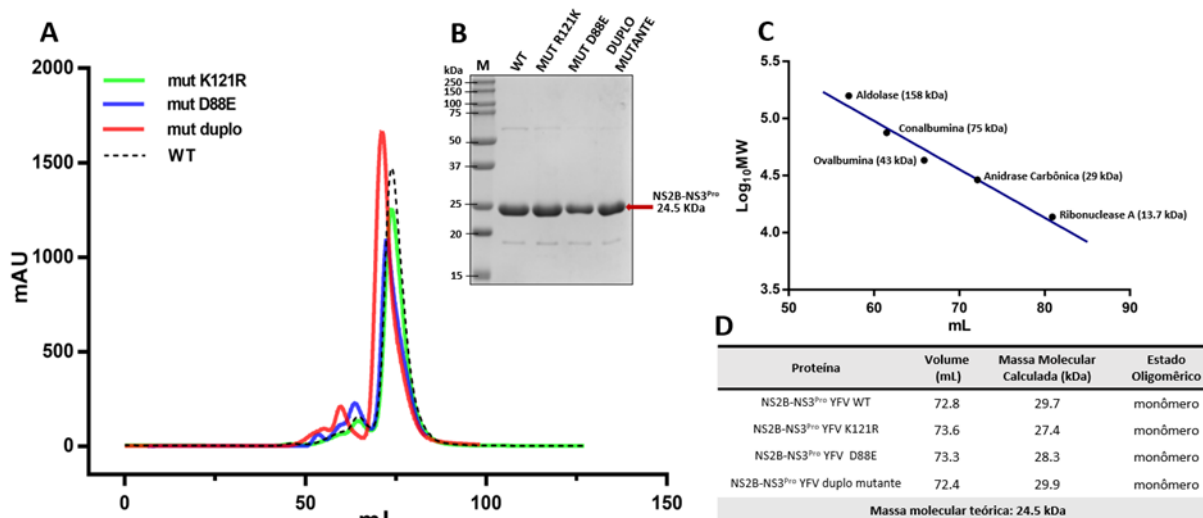


Figura 14 - Resultados da purificação das proteínas NS2B-NS3^{PRO} linked. A) Cromatogramas da etapa de exclusão molecular para cada uma das enzimas. B) SDS-PAGE 12.5% mostrado a pureza final de cada amostra ao final da purificação. C) Curva de calibração da coluna Superdex 75 16/60 D) Tabela contendo os volumes de retenção de cada uma das proteínas, suas respectivas massas moleculares calculadas e estado oligomérico.

Fonte: Elaborada pela autora.

As proteínas NS2B-NS3^{PRO} WT, K121R, D88E e duplo mutante possuem volumes de retenção de 72.8 mL, 73.6 mL, 73.3 mL e 72.4 mL e massas moleculares calculadas de 29.7 kDa, 27.4 kDa, 28.3 kDa e 29.9 kDa respectivamente. Ambas as proteínas NS2B-NS3^{PRO} linked possuem volume de retenção equivalentes ao tamanho de uma molécula no estado monomérico. Além disso, o SDS-PAGE, juntamente com os cromatogramas da exclusão molecular nos permitem inferir que a amostra se encontra homogênea e com elevado grau de pureza. O rendimento aproximado de ambas as proteínas foi de 7 mg por litro de cultura.

3.1.5 Ensaios de cristalização e coleta de dados

Os primeiros cristais foram obtidos utilizando kits comerciais e o método de matriz esparsa. A condição na qual foram observados cristais, condição G01 do kit Index HT® (Hampton Research) (0.1 M NaCl, 0.1 M Tris pH 8.5, 25% PEG 3350) na temperatura de 18°C, foi então utilizada para otimização. Foram testadas várias concentrações de proteína variando entre 10 mg/mL e 38 mg/mL. O aparecimento de cristais foi observado na concentração mais alta de proteína (38 mg/mL) após cerca de um mês. Em outro experimento, realizamos a troca

do tampão da proteína para Bis Tris pH 6.5, 150 mM NaCl e a proteína concentrada a 20 mg/mL. Cristais foram observados no kit PACT® (Quiagen) condição D09 (0.1 M LiCl, 0.1 M Tris pH 8.0, 20% PEG 6000) também após cerca de um mês a 18°C (Figura 15C). Os cristais da NS2B-NS3^{pro} WT foram otimizados manualmente pelo método de *sitting drop* (Figura 15D-F) baseando-se na condição inicial e novamente os cristais apareceram após cerca de um mês. Com a utilização da técnica de *microseeding* na mesma condição (G01 Index) foi possível a obtenção de um menor número de cristais isolados que apareceram em cerca de 10 dias.

Para as proteínas mutantes K121R e D88E o aparecimento de cristais foi observado no kit PACT® (Quiagen) ambos na condição B10 (0.1 M MES pH 6.0, 0.2 M MgCl, 20% PEG 6000) na temperatura de 18°C após cerca de um mês na concentração de 35 e 40 mg/mL respectivamente (Figura 10A-B). Para o mutante duplo não foi possível obter cristais.

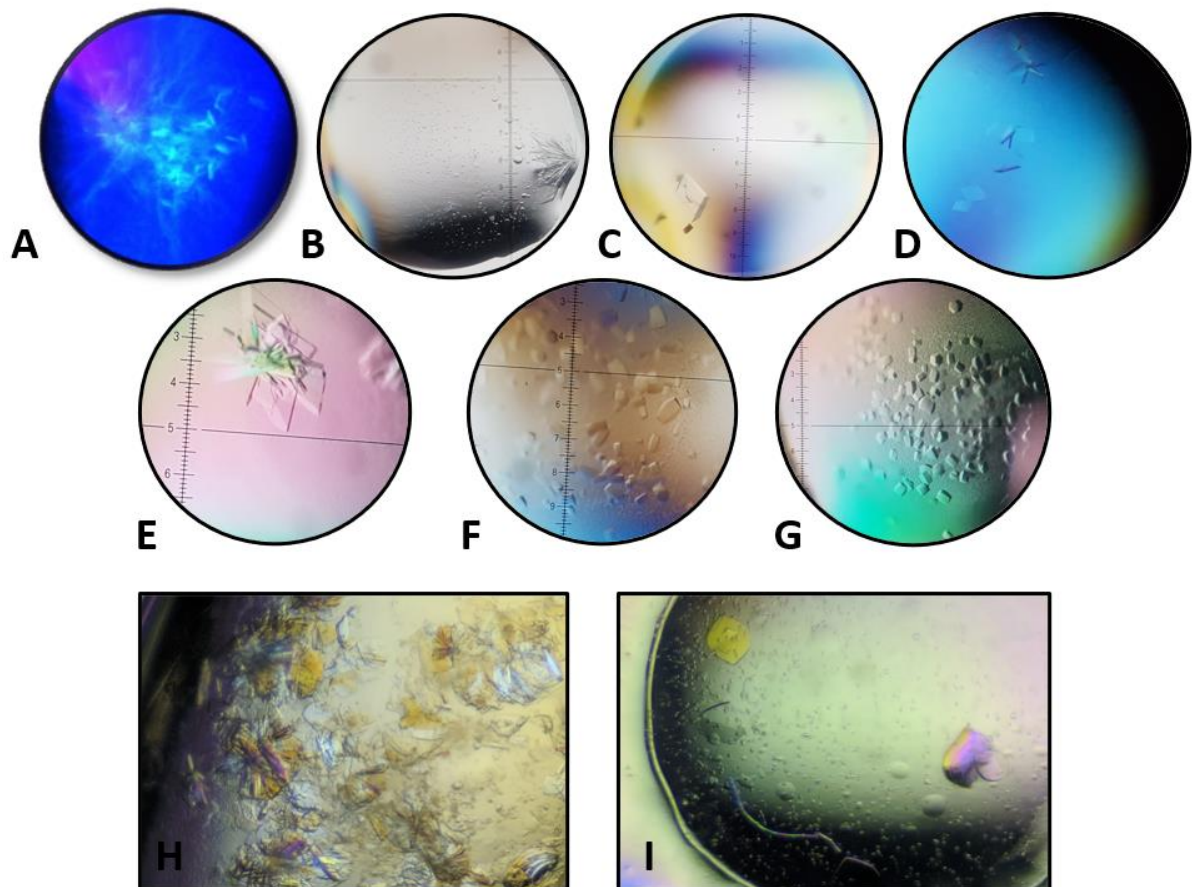


Figura 15 - Cristais da enzima NS2B-NS3^{pro}. A) Cristais da proteína inicialmente obtidos (condição G01 Index HT) sob luz UV. B) Cristais da proteína mutante K121R na condição B10 PACT (0.1 M MES pH 6.0, 0.2 M MgCl, 20% PEG 6000). C) Cristais da proteína mutante D88E na condição B10 PACT. D) Cristais otimizados da condição D9 PACT (0.1 M Tris pH 8.5, 0.1 M LiCl, 25% PEG 6000). E-G) Cristais otimizados da condição G1 Index HT (0.1 M NaCl, 0.1 M Tris pH 8.5, 25% PEG 3350) contendo respectivamente 23%, 27% e 29% PEG 3350. H-I) Cristais otimizados da condição G1 Index HT (0.1 M NaCl, 0.1 M Tris pH 8.5, 25% PEG 3350) antes (H) e após (I) utilização de *microseeding*.

Os mesmos cristais foram armazenados em nitrogênio líquido e coletados na linha de luz I24 no Diamond Light Source, Oxfordshire, Reino Unido. Cerca de 80 cristais foram testados, dos quais somente um difratou à uma resolução de 2.9 Å tornando possível a resolução da estrutura cristalográfica da enzima (Figura 16). Os parâmetros da coleta e estatísticas do processamento podem ser observados na Tabela 5.

Os cristais otimizados da enzima NS2B-NS3^{pro} WT foram crio protegidos com a solução do poço contendo de 8% etilenoglicol, resfriados e utilizados para a coleta de dados de difração. A coleta foi inicialmente realizada no Laboratório Nacional de Luz Síncrotron (LNLS), Campinas – SP, na linha de luz MX2, no entanto os melhores cristais difrataram a baixa resolução (cerca de 6 Å) e foram provenientes da condição otimizada derivada da condição inicial contendo 0.1 M Tris pH 8.5, 0.1 M LiCl, 25% PEG 6000. Os cristais obtidos utilizando a técnica de *microseeding*, embora possibilitaram a obtenção de cristais isolados e com menos tempo (cerca de 10 dias), não proporcionaram melhora na qualidade da difração.

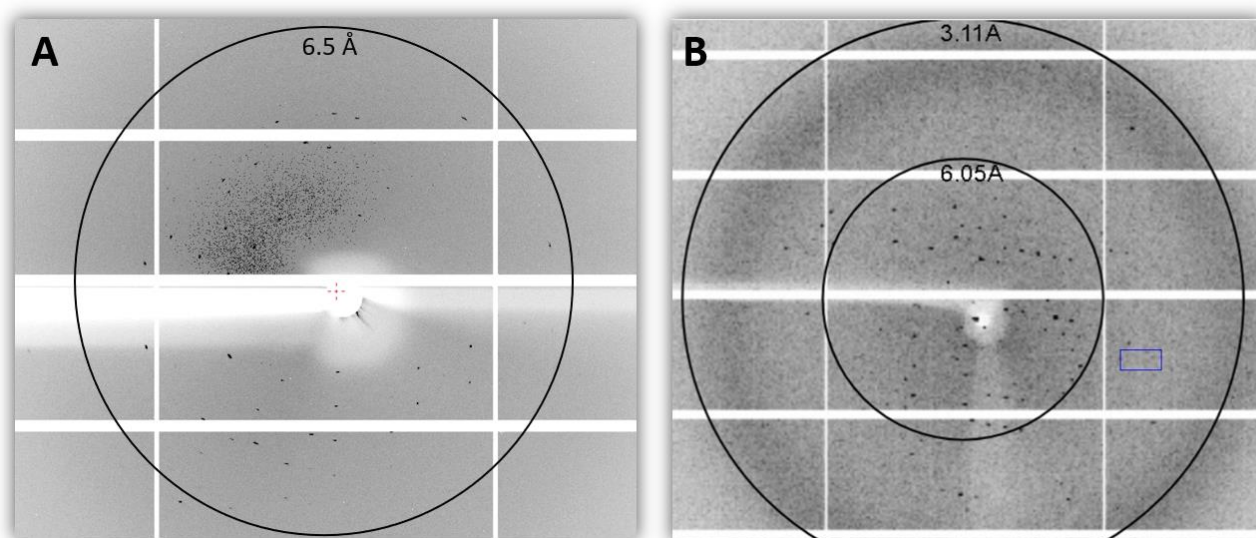


Figura 16 - Padrão de difração obtido para um dos cristais de YFV da condição contendo 0.1 M Tris pH 8.5, 0.1 M LiCl, 25% PEG 6000, obtidos respectivamente na linha de luz MX2, no LNLS, Campinas, Brasil (A) e na linha I24, no Diamond Light Source, Reino Unido (B).

Fonte: Elaborada pela autora.

Tabela 5- Coleta de dados e estatísticas do refinamento da enzima NS2B-NS3^{pro}. Os dados entre parênteses correspondem à última camada de resolução.

		NS2B-NS3 ^{pro} YFV
Grupo especial		<i>P</i> 2 ₁
Dimensões da cela unitária		
	a,b,c (Å)	60.0, 151.8, 60.3
	α,β,γ (°)	90, 111.3, 90
Resolução (Å)		29.2 -2.9 (3.0 - 2.9)
R_{pim}		0.08 (0.90)
Mean I/σ(I)		7.22 (1.0)
CC_{1/2}		0.99 (0.32)
Completeza (%)		98.0 (99.5)
Multiplicidade		6.5 (6.9)
Refinamento		
No. de reflexões		22135 (2232)
Moléculas na unidade assimétrica		4
R_{work}/R_{free}		0.25/0.27
No. de átomos		
	Proteína	5456
	Água	16
B-fatores médios (Å²)		
	Proteína	98.45
	Água	54.95
RMSD		
	Comprimento das ligações (Å ²)	0.007
	Ângulos das ligações (°)	1.29
Ramachandran (%)		
	Favoráveis	97.35
	Não permitidos	0
Código PDB		6URV

Fonte: Elaboradora pela autora, baseada em NOSKE *et al.*⁵⁷

3.1.6 Resolução da estrutura cristalográfica da proteína NS2B-NS3^{pro} YFV

Os dados cristalográficos provenientes da cristalização da enzima NS2B-NS3^{pro} YFV foram integrados no grupo espacial *P*2₁, do qual foi possível obter as dimensões a célula unitária com valores $a = 60.0 \text{ \AA}$, $b = 151.8 \text{ \AA}$, $c = 60.3 \text{ \AA}$, $\alpha = 90^\circ$, $\beta = 111.3^\circ$, $\gamma = 90^\circ$. A partir do coeficiente de Matthews calculado de 2.62 foi possível inferir a presença de quatro moléculas de proteína na unidade assimétrica com 52.98% de solvente. Para a resolução da estrutura da NS2B-NS3^{pro} foi utilizada a substituição molecular para cálculo das fases, utilizando como modelo a NS3 protease de dengue (PDB ID 2FOM, identidade de 56%). A estrutura foi resolvida numa resolução de 2.9 Å determinada pelo valor de CC_{1/2}, com completeza de 99.5% na última camada de resolução.⁶⁴ O modelo final da estrutura contém 37 resíduos do cofator NS2B (49-86) e 153 da NS3^{pro} (14-167) e foi refinada até atingir R_{work}/R_{free} de 0.25/0.27. Os parâmetros e estatísticas finais do refinamento podem ser observadas na Tabela 5.⁵⁷

3.1.7 Estrutura cristalográfica da proteína NS2B-NS3^{pro} YFV

Similar a outras serino-proteases, a porção N-terminal da NS3 protease possui enovelamento característico, e depende de uma região solúvel de cerca de 40 resíduos da NS2B, a qual age como cofator da enzima. A estratégia comumente utilizada para obtenção da enzima para elucidação estrutural e busca por inibidores, consiste em manter apenas a região essencial do cofator NS2B, removendo a sua porção hidrofóbica, a qual é covalentemente ligada através de um *linker* flexível contendo serinas e glicinas ao domínio protease da NS3.⁵⁷

Assim como outras proteases de outros flavivírus, a estrutura cristalográfica da enzima NS2B-NS3^{pro} de YFV tem *folding* similar ao de quimotripsinas, e é caracterizada por dois barris- β com cada um deles contendo 6 fitas β antiparalelas, com o sítio ativo localizado entre eles (Figura 17B). Já o cofator NS2B é composto por 3 fitas β que se enovelam formando um grampo β entre as fitas B2 e B3. A presença do *linker* flexível G₄SG₄ permite que o cofator NS2B circule o domínio NS3^{pro}, com a proteína sendo estabilizada em sua conformação ativa, na qual a fita B1 do cofator NS2B interage com β 1 do domínio NS3 protease. Já as fitas B2 e B3 interagem com as fitas β 6, β 7, β 10 e β 14 da NS3 protease (Figura 17B-C).⁵⁷

O sítio ativo da NS2B-NS3^{pro} é característico de serino-proteases, composto pela tríade catalítica Ser138, His53 e Asp77 situado entre o sítio de reconhecimento e o sítio S1', com a His53 encontrada na hélice 310 A1, D77 na fita β 7 e S138 na fita β 12 (Figura 17B e D). De maneira similar a outras proteases de flavivírus já descritas, a protease de YFV possui sítio ativo carregado negativamente, especialmente nos subsítios S1-S2, que são responsáveis por acomodar os resíduos carregados positivamente (lisinas e argininas) presentes no substrato da enzima. Além disso, o subsítio S1' tem preferência por resíduos pequenos como serina, glicina e treonina. Embora o enovelamento das proteases de flavivírus seja similar ao de muitas outras serino-proteases, o padrão de reconhecimento de substrato nas posições P1'-P2 confere características únicas, que tornam possível o desenvolvimento de inibidores seletivos da enzima. (Figura 17E).⁵⁷

Comparando a estrutura da protease de YFV com estruturas de outros flavivírus, embora a identidade sequencial não seja muito alta (cerca de 50% para os vírus DENV2, DENV3, ZIKV e WNV), as proteínas possuem elevada similaridade estrutural, principalmente na região que compõe o sítio ativo e o bolsão de reconhecimento do substrato. Os valores de RMSD calculados são relativamente baixos, com valores de 2.1 Å para DENV2 (PDB ID: 2FOM, conformação aberta, para 155 resíduos), 1.6 Å para DENV3 (PDB ID: 3U1I, conformação fechada, para 157 resíduos), 1.1 Å para WNV (PDB ID: 2YOL, conformação fechada, para 157

resíduos), e 1.4 Å para ZIKV (PDB ID: 5TFN, conformação aberta, para 158 resíduos) (Figura 18A-I). A região da estrutura que apresentou maior diferença corresponde as fitas β 14 e β 15 e no *hairpin* entre estas. As estruturas cristalográficas da enzima NS2B-NS3^{pro} dos flavivírus na conformação fechada, ou ativa, encontram-se na conformação em que o cofator NS2B circula a NS3 protease, onde a porção C-terminal da protease interage diretamente com o cofator NS2B, formando um dos sítios de reconhecimento do substrato (S2), como é o caso da protease de YFV (Figura 18E-H). Em contraste, as estruturas na forma inativa, ou conformação aberta, encontram-se de modo que o cofator não circula propriamente a NS3^{pro} (Figura 18A-D).⁵⁷

Até hoje, de maneira geral, as estruturas do complexo NS2B-NS3^{pro} de flavivírus depositadas no PDB indicam que na ausência de um substrato ou inibidor ligado ao sítio ativo da enzima, esta encontra-se em conformação aberta, ou inativa, na qual o cofator NS2B não envolve a NS3^{pro}, compondo o sítio ativo da enzima. Em contraste, enzimas em complexo com substrato/inibidor são descritas como estando na conformação fechada e ativa. Além disso as estruturas depositadas até então indicam que a presença do *linker* G₄SG₄ está associada a enzima em conformação aberta.^{52,81-82} Quando o complexo se encontra na presença de um inibidor ou do substrato da enzima ligado em seu sítio ativo, as interações entre substrato/inibidor e o cofator NS2B, são responsáveis por estabilizar e manter a enzima nesta conformação. Surpreendentemente, a estrutura da NS2B-NS3^{pro} de YFV encontra-se em conformação fechada (*closed conformation*), considerada ativa, na qual o cofator NS2B circula o domínio NS3^{pro} mesmo na ausência de qualquer inibidor ou ligante no sítio ativo (Figura 18A-I).

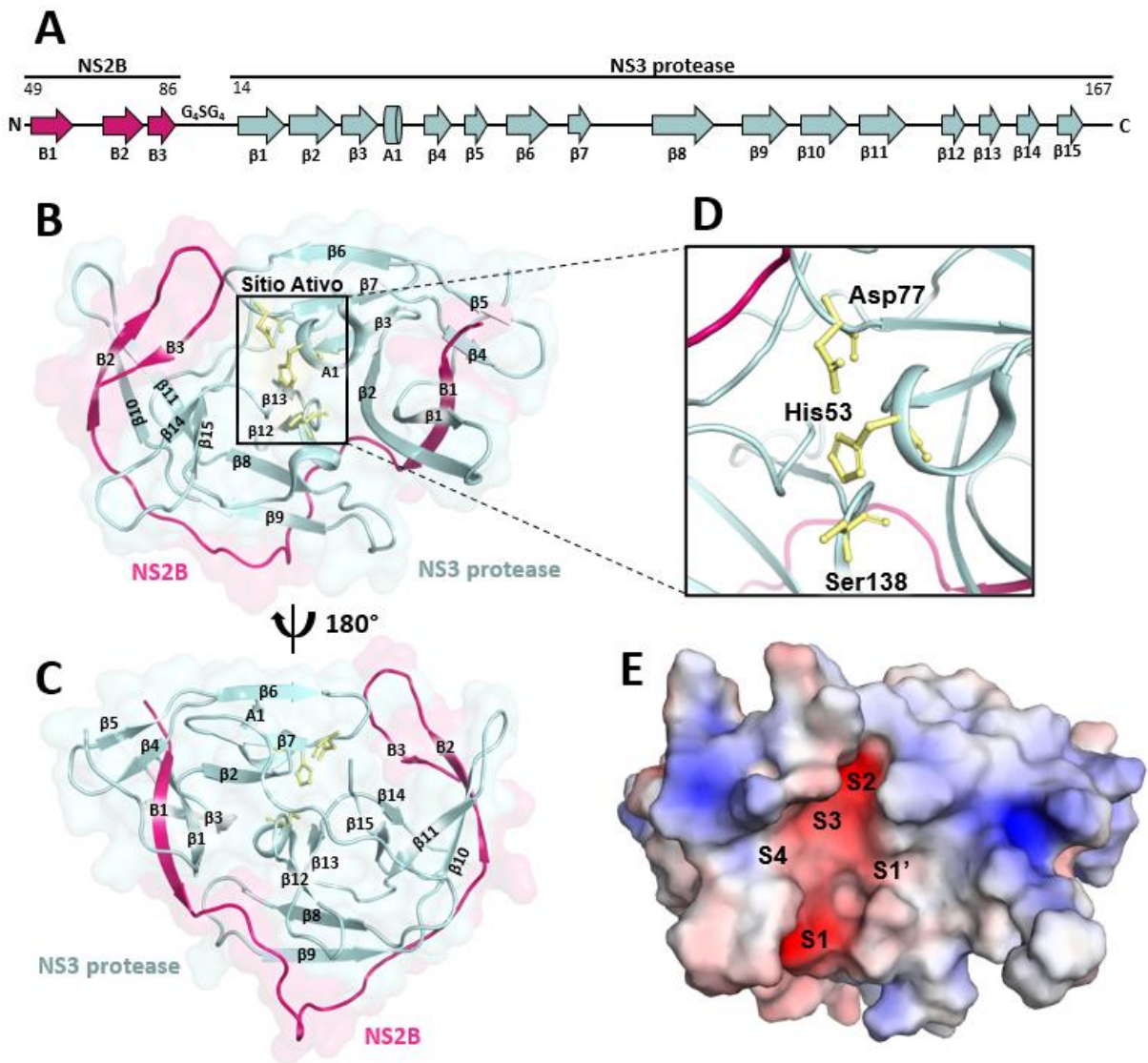


Figura 17 - Estrutura cristalográfica da enzima NS2B-NS3^{pro} de YFV. A) Representação esquemática da estrutura secundária da proteína, indicando cada um dos domínios. As fitas β estão representadas por setas e a hélice 3₁₀ por cilindro. Visão geral da estrutura cristalográfica da enzima NS2B-NS3^{pro} de YFV. A estrutura está representada em modo *cartoon* com o cofator NS2B colorido em laranja e o domínio NS3^{pro} em verde. A superfície de ambos os domínios está representada nas suas respectivas cores. O quadro pontilhado destaca o sítio ativo da enzima. B-C) Visão geral da estrutura cristalográfica da enzima NS2B-NS3^{pro} de YFV rotacionada em 180°. D) Visão ampliada do sítio ativo da enzima, mostrando os resíduos catalíticos His53, Asp77 e Ser138, representados como sticks. E) Representação das cargas da superfície da enzima NS2B-NS3^{pro} com destaque para o bolsa de reconhecimento de substrato S1' e S1-S4.

Fonte: Elaborada pela autora.

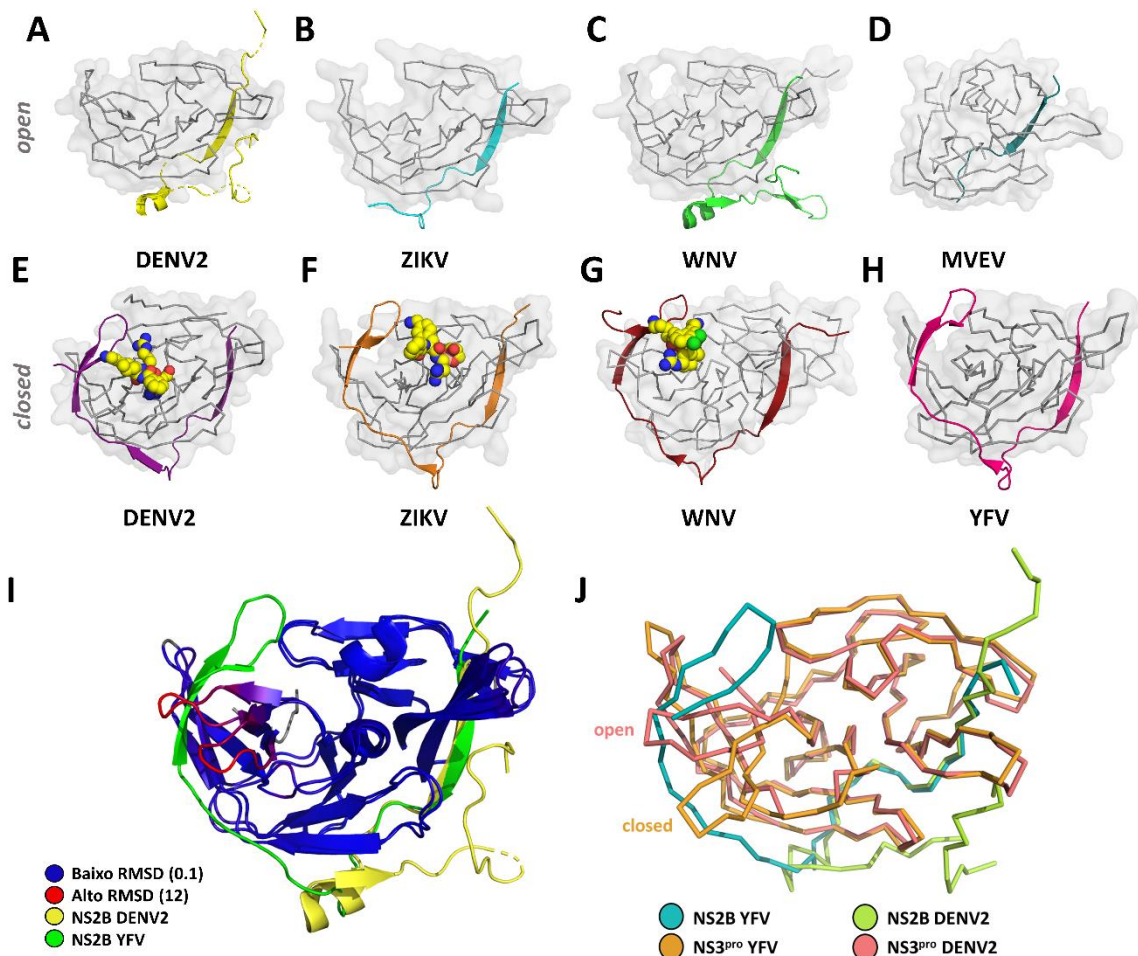


Figura 18 - Diferentes conformações do complexo NS2B-NS3^{pro} de Flavivirus. Em A-D, estruturas em conformação aberta (*open conformation*), respectivamente para os vírus DENV2, ZIKV, WNV e MVEV. Em E-H, estruturas em conformação fechada (*closed conformation*). Para as estruturas de A-H, o domínio protease está representado como superfície, juntamente com o traço da cadeia alpha ambos coloridos em cinza. Já o cofator NS2B, está representado como *cartoon*, em diferentes cores para cada estrutura. I) Superposição dos complexos NS2B-NS3^{pro} de YFV (*closed*) e DENV2 (*open*). Os domínios protease estão coloridos de acordo com o RMSD (azul para baixo RMSD e vermelho para alto) e o cofator NS2B está colorido em amarelo e azul para DENV2 e YFV respectivamente. Ambas as estruturas estão representadas na forma de *cartoon*. Em destaque vemos o loop que determina o nome dado a conformação. J) Superposição dos complexos NS2B-NS3^{pro} de YFV (*closed*) e DENV2 (*open*), ambas representadas como traço da cadeia alpha.

Fonte: Elaborada pela autora baseada em NOSKE *et al.*⁵⁷

Mesmo após uma inspeção cuidadosa das densidades eletrônicas próximas ao sítio ativo da proteína, incluindo a obtenção dos *omit maps* das regiões para ambas as cadeias da unidade assimétrica, não foi possível observar a presença de densidades eletrônicas que poderiam ser de um ligante no sítio ativo da proteína.⁵⁷ Estudos realizados com a proteína NS2B-NS3^{pro} de DENV em solução, pela técnica de ressonância magnética nuclear (RMN) demonstraram que em solução, ambas as conformações *open* e *closed* estão presentes, e que construções não

ligadas covalentemente favorecem a presença da conformação fechada.⁸¹⁻⁸² Conforme ilustrado na Figura 19B, as regiões que definem a conformação da enzima (NS2B, fitas β 14, β 15 e *hairpin* entre estas) possuem elevado B-fator comparado ao restante da estrutura. Em uma estrutura cristalográfica, elevado B-fator em geral é observado em regiões da proteína que possuem alta flexibilidade, assim, para a protease de YFV, possivelmente é resultado da presença de ambas as conformações da proteína, não somente em solução, mas também dentro da estrutura cristalina. Ainda por conta desta elevada flexibilidade, o B-fator geral calculado para todos os átomos da proteína foi de 98.45 Å², valor relativamente alto comparado a maior parte das estruturas cristalográficas em resoluções similares e também comparado ao B-fator obtido na mesma estrutura para as moléculas de solvente, podendo ser proveniente de uma maior instabilidade na estrutura da protease YFV em conformação fechada, quando comparada a estruturas que possuem um ligante no sítio ativo, ajudando a estabilizar a interação com o cofator NS2B. O motivo para a dificuldade em obter cristais que difratassem em alta resolução, pode, portanto, ser ocasionado por esta possível instabilidade, alta flexibilidade da proteína e não homogeneidade da amostra.

Até o momento não existe nenhuma evidência que explique a conformação obtida, podendo ter sido derivada de fatores externos, como a contribuição dos contatos cristalinos para estabilização da proteína em sua conformação fechada.⁵⁷ Os contatos cristalinos, tanto os observados no cristal da protease de YFV quanto os observados para a protease de ZIKV^{61,83} (Figura 19A) sugerem que possa ocorrer dimerização de dois complexos NS2B-NS3^{pro}, embora os resultados obtidos em solução, por gel filtração analítica (Figura 14) e SEC-MALS (dados não apresentados), indicam que a maior parte da proteína se encontra em estado monomérico, este pode ser outro fator que contribuiu na estabilização da conformação ativa. Esse processo também pode ser induzido pela presença de substrato, levando ao aumento da concentração de proteína em sua conformação ativa (fechada).

Levando em conta que a conformação ativa da enzima consiste na conformação fechada, esta é, portanto, ideal para a busca por inibidores. A estrutura obtida, pode então ter suma importância na obtenção de complexos proteína com inibidor e no planejamento de um potencial candidato antiviral. Estudos complementares são necessários para compreender melhor os fatores que levaram a obtenção da estrutura apo da protease de YFV em conformação fechada.

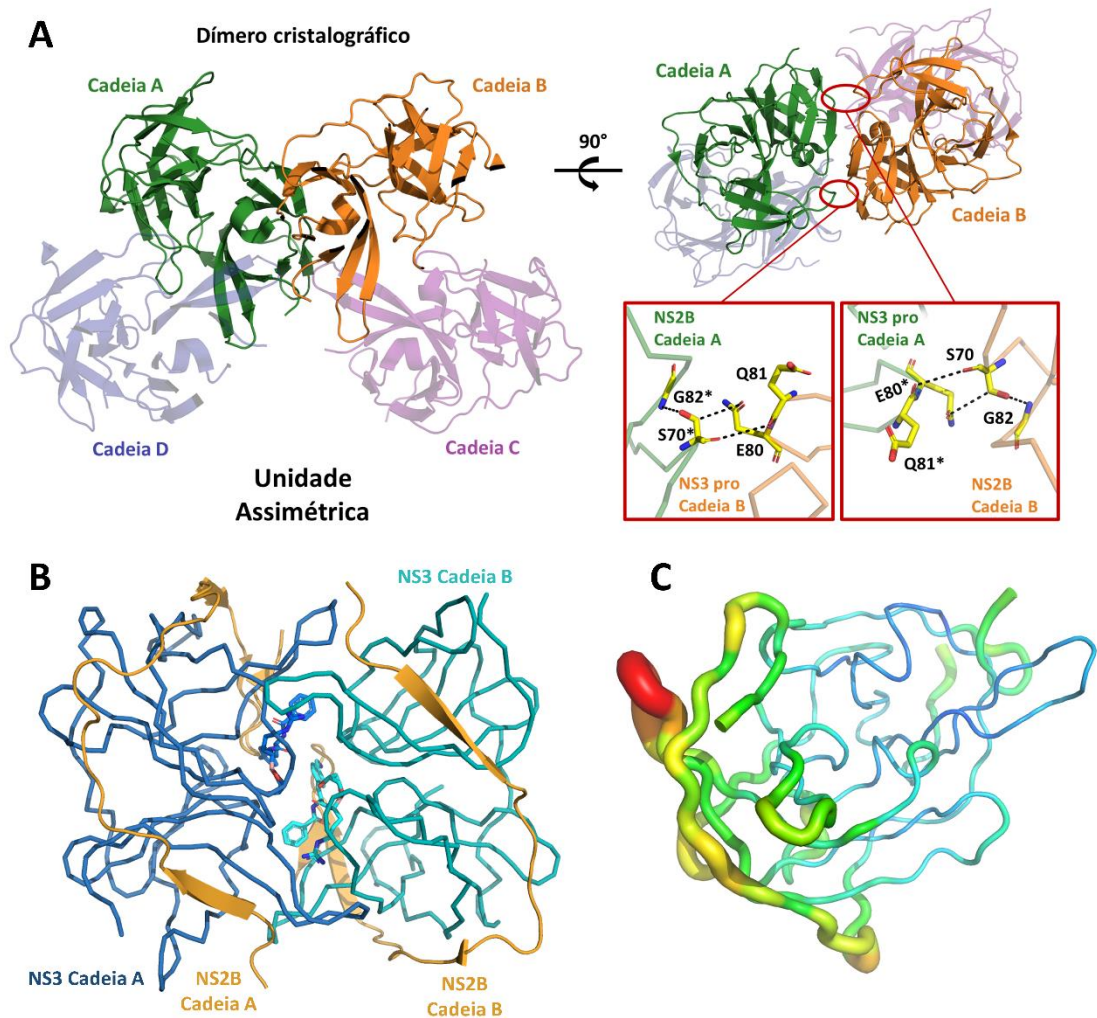


Figura 19 - A) Contatos cristalinos observados nos cristais da enzima NS2B-NS3^{pro} de YFV (PDBid 6URV). As moléculas formadoras do dímero cristalográfico estão coloridas em verde e laranja, representadas em *cartoon*. Os quadros vermelhos contêm uma visão ampliada das regiões de contato do dímero cristalográfico, na qual os resíduos diretamente envolvidos estão coloridos em amarelo, e representados como *sticks*. B) Dímero cristalográfico observado para a NS2B-NS3 protease do vírus Zika. C) Estrutura da enzima NS2B-NS3^{pro} de YFV representada de acordo com os B-fatores de cada resíduo.

Fonte: Elaborada pela autora.

3.1.8 Ensaios de atividade

O aumento da fluorescência observado ao longo do tempo (Figura 20A) demonstra que a proteína se encontra ativa. Como mostrado na Figura 20A, foi realizada a titulação da concentração de enzima a fim de determinar uma concentração que satisfizesse o estado estacionário e, assim, padronizar esta concentração para a realização de experimentos posteriores. Pela análise deste mesmo gráfico, pudemos inferir apenas que para concentrações

abaixo de 7.5 nM estávamos dentro da região linear no período medido. Com isso definimos a concentração de 7 nM como padrão para os ensaios enzimáticos.

Em seguida, obtivemos uma curva de calibração de AMC, através da determinação da fluorescência de concentrações conhecidas de 7-amino-4-metilcumarina (Figura 20B). Nos experimentos subsequentes, para obtenção dos parâmetros cinéticos, os valores de fluorescência foram convertidos para concentração de produto utilizando a regressão linear desta curva como referência.

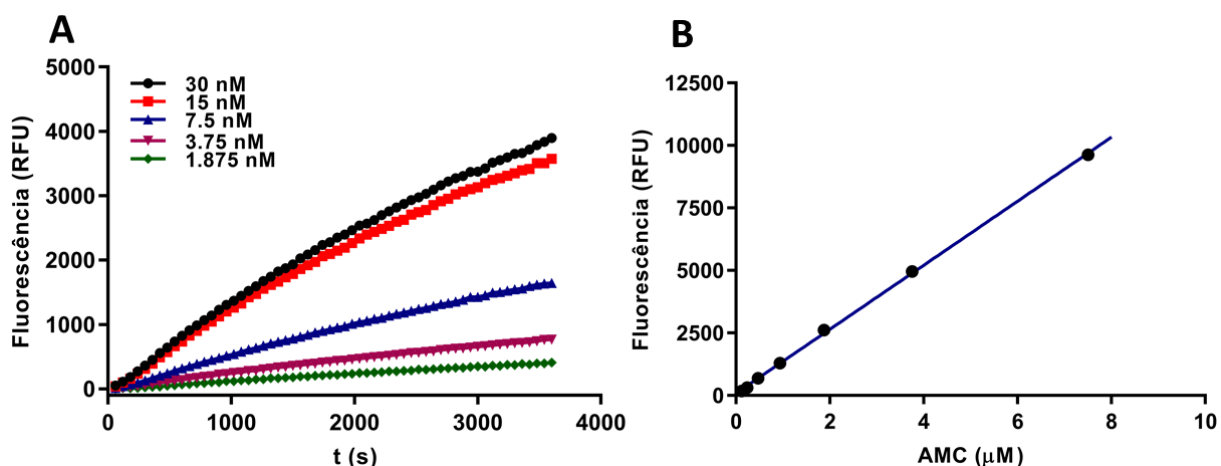


Figura 20 - Ensaios de atividade da enzima NS2B-NS3^{pro} A) Curvas obtidas para o teste inicial de atividade e padronização da concentração de enzima, cada uma das cores representa uma concentração diferente de enzima. B) Curva de calibração de 7-amino-4-metilcumarina (AMC).

Fonte: Elaborada pela autora.

3.1.9 Caracterização das enzimas mutantes

As enzimas mutantes foram entrem caracterizadas através da obtenção de seus parâmetros cinéticos e de sua estabilidade térmica. A determinação dos parâmetros cinéticos, foi realizada realizando uma titulação do substrato (de 600 μM até 18.8 μM) e pelo *fitting* da equação de Michaelis-Menten (Figura 21A). Os valores de afinidade pelo substrato (K_M) e k_{cat} da enzima NS2B-NS3^{pro} WT foram respectivamente $90 \pm 10 \mu\text{M}$ e $0.24 \pm 0.03 \text{ s}^{-1}$, demonstrando assim uma menor afinidade pelo substrato quando comparada a ZIKV (2.9 μM) e WNV (24 μM) com valor de k_{cat} similar a WNV (0.26 s^{-1}) e menor do que ZIKV (1.8 s^{-1}).^{83,84} Os mutantes D88E, K121R e duplo mutante mostraram valores similares de afinidade pelo substrato quando comparados com a enzima WT com respectivos K_M de $80 \pm 10 \mu\text{M}$, $100 \pm 10 \mu\text{M}$ e $110 \pm 10 \mu\text{M}$. Os três mutantes também mostraram diferenças nos valores de eficiência

catalítica (k_{cat}/K_M) quando comparados a enzima WT, com eficiências relativas de 148%, 108% e 78% respectivamente.⁵⁷

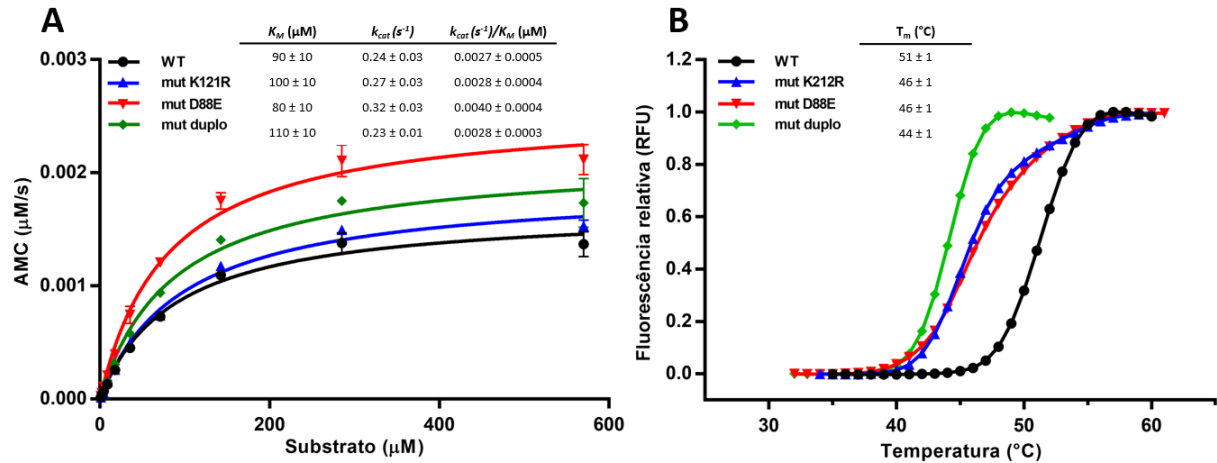


Figura 21 - Caracterização das enzimas NS2B-NS3^{pro}. A) Curvas de Michaelis-Menten para cada uma das enzimas mutantes e WT. As barras de erros correspondem ao desvio padrão das triplicatas. Os valores de afinidade pelo substrato (K_m), constante catalítica (k_{cat}) e eficiência catalítica (k_{cat}/K_m) de cada uma das enzimas é mostrado acima do gráfico. B) Perfis de desnaturação térmica de cada uma das enzimas. Os valores de T_m para cada uma das enzimas é mostrado acima do gráfico.

Fonte: Adaptada de NOSKE *et al.*⁵⁷

Os valores obtidos para as T_m foram $51 \pm 1^{\circ}\text{C}$, $46 \pm 1^{\circ}\text{C}$, $46 \pm 1^{\circ}\text{C}$ and $44 \pm 1^{\circ}\text{C}$, para as enzimas WT, D88E, K121R e duplo mutante. Logo, a protease WT mostrou um aumento aproximado de $5\text{-}6^{\circ}\text{C}$ com relação a seus mutantes, sugerindo ter uma maior estabilidade térmica (Figura 21B).⁵⁷

Embora os mutantes tenham exibido afinidades similares pelo substrato Bz-nKRR-AMC, ambas mostraram ter processividade maior que a enzima WT, resultando numa eficiência catalítica similar. A diferença mais acentuada foi observada na estabilidade térmica, sendo que ambos os mutantes possuem estabilidade térmica significativamente menor do que a enzima WT.

Conforme previamente mencionado no item 3.1.7, as proteases de flavivírus possuem identidade sequencial relativamente baixa (em torno de 50-60% para DENV2, DENV3, WNV e ZIKV), porém possuem elevada similaridade estrutural (Figura 22). A análise realizada pelo servidor Consurf⁸⁵ mostrou que grande parte dos resíduos não são conservados. Dentre as regiões não conservadas, vale destacar a região de interação entre o cofator NS2B e o domínio NS3^{pro}. Devido a esta baixa conservação, existe uma grande dificuldade em planejar candidatos antivirais alostéricos, que inibam a interação NS2B-NS3^{pro}, e que possuam atividade antiviral de amplo espectro contra os flavivírus.⁵⁷ Dentre as regiões conservadas, estão o sítio ativo, e

grande parte dos resíduos formadores do bolsão de ligação ao substrato (S1'-S4), responsáveis pelo reconhecimento dos resíduos carregados positivamente presentes nos sítios da clivagem da poliproteína. Ambos os resíduos mutantes presentes na protease da linhagem circulante de YFV (D88 e K121) são classificados como tendo baixa conservação de acordo com a análise feita pelo Consurf (Figura 22B-D).

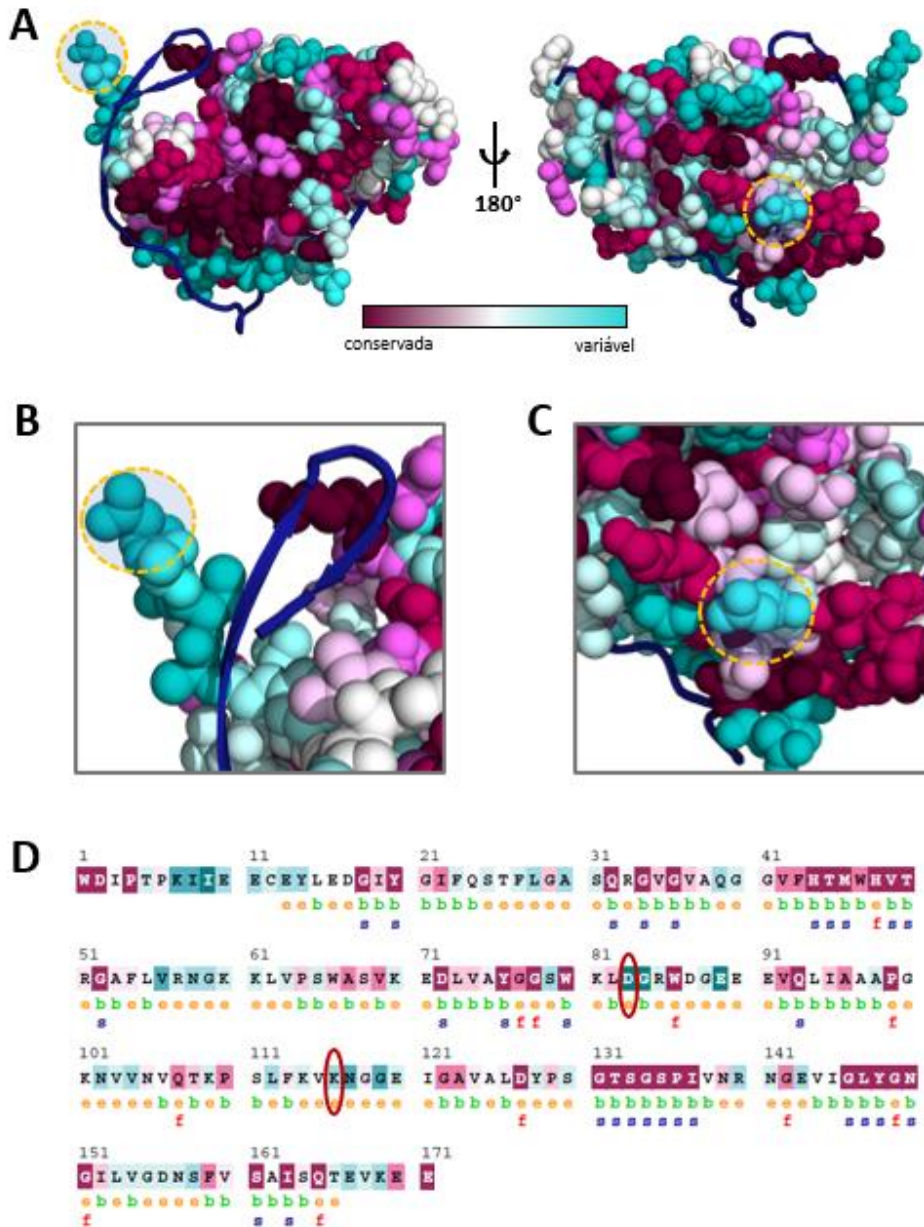


Figura 22 - Análise de conservação da enzima NS2B-NS3^{pro}. A) Visão geral da estrutura da NS2B-NS3^{pro} de YFV, onde o domínio NS3^{pro} está colorido de acordo com a variabilidade sequencial comparada a outras proteases de flavivírus. A análise de conservação foi feita utilizando o servidor Consurf.⁵⁷ Cada resíduo da NS3^{pro} está representado como esferas. O cofator NS2B está representado na forma de *cartoon* e colorido em azul escuro. B-C) Visão ampliada das regiões contendo os resíduos mutantes da linhagem circulante da protease de YFV, K121 e D88 em (B) e (C) respectivamente. D) Sequência da protease de YFV com os resíduos coloridos de acordo com a escala de conservação do Consurf.⁵⁷ Os resíduos mutantes estão destacados em vermelho.

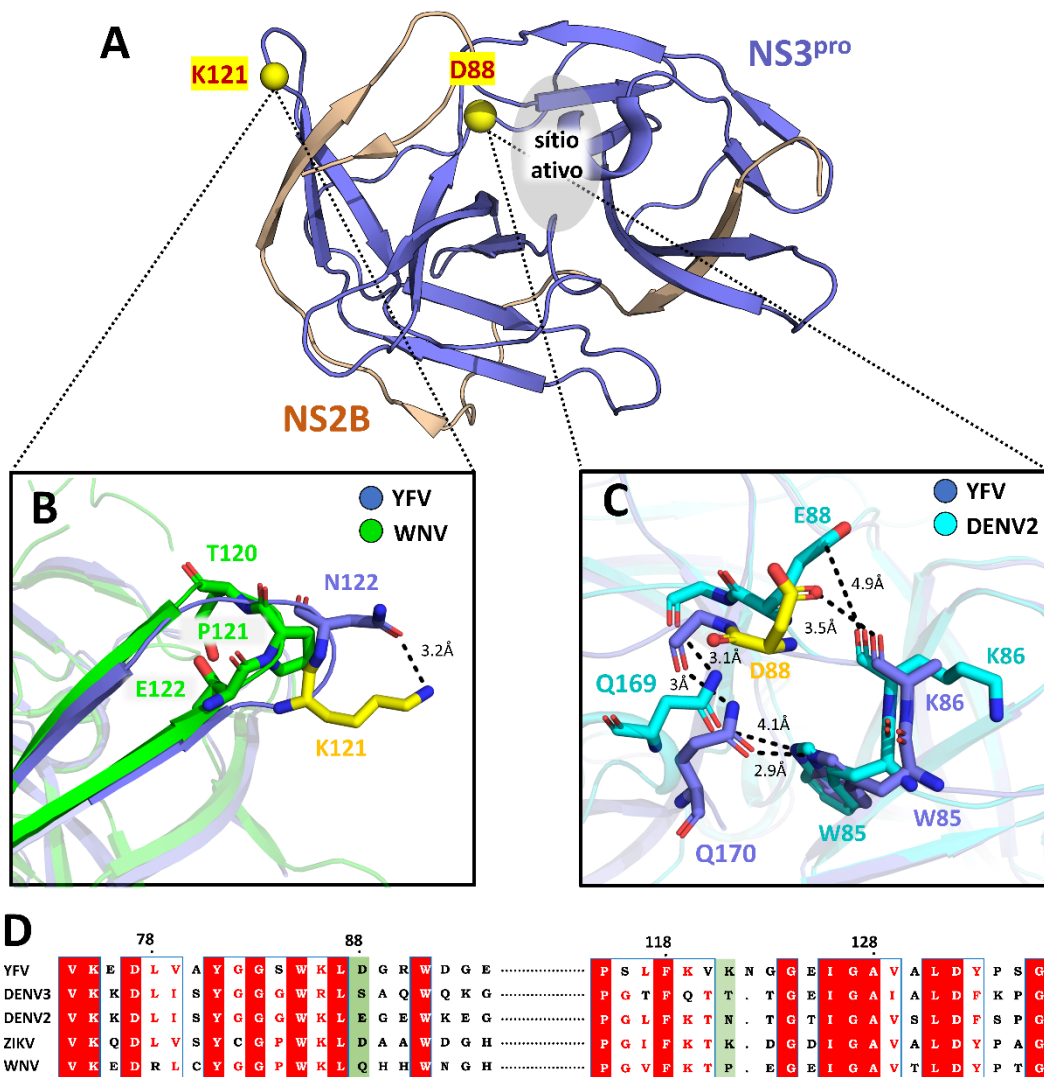


Figura 23 - Análise estrutural das mutações pontuais observadas na protease de YFV da linhagem circulante. A) Estrutura cristalográfica da enzima NS2B-NS3^{pro} representada em formato *cartoon* e localização das mutações D88 e K121, representadas como esferas (em amarelo). O cofator NS2B está colorido em laranja e a NS3 protease em azul. A região correspondente ao sítio ativo está destacada em cinza. B) Visão ampliada do resíduo K121 de YFV (em azul) (PDB ID 6URV) e sua interação com o resíduo N122 sobreposta com a mesma região da estrutura de WNV (em verde) (PDB ID 2YOL). C) Visão ampliada do resíduo E88 de YFV (em azul) e suas interações com os resíduos K86 e Q169 sobreposta com a mesma região da estrutura de DENV2 (em ciano) (PDB ID 4M9K). D) Alinhamento das sequências da proteína NS3 protease de YFV, DENV3, DENV2, ZIKV e WNV. Os resíduos mutantes estão destacados em verde e o alinhamento foi feito utilizando o servidor ESPrit.⁸⁶

Fonte: Adaptada de NOSKE *et al.*^{57,86}

Pela análise da estrutura da protease de YFV e sua comparação com outros flavivírus, inferimos que resíduo D88 tem função de estabilizar uma porção móvel da proteína. A presença do aspartato permite contato com o oxigênio da cadeia principal do resíduo K86. A presença de resíduos mais longos, como a glutamina presente na protease de DENV2, não permite esta

interação, resultando em um achatamento do *loop*, o que pode afetar a estabilidade da proteína de maneira geral (Figura 23C).⁵⁷ Já o resíduo K121, está localizado entre as fitas β 10 e 11, e não somente encontra-se em uma região não conservada, como representa como uma inserção comparado a outros flavivírus (Figura 22B, D e Figura 23D). O resíduo K121 presente em YFV permite um contato polar com a cadeia lateral do resíduo N122, formando um β -*hairpin* mais alongado comparado a outros flavivírus, ilustrado na figura para DENV2, na qual a estabilização no mesmo loop é mediada apenas por contatos hidrofóbicos (Figura 23B). O mutante duplo exibiu uma queda ainda mais acentuada em sua estabilidade térmica, demonstrando que ambas as mutações têm importância na adaptabilidade da enzima. A mutação K121R também parece ser importante para atividade enzimática, pois ambos os mutantes K121R e mutante duplo, apresentaram menor afinidade pelo substrato.⁵⁷ Assim, em conjunto, as mutações podem afetar de forma geral a atividade da enzima, resultando no aumento da estabilidade térmica da enzima WT.

No entanto, as razões para as diferenças observadas, tanto na processividade da enzima quanto em sua estabilidade térmica não são claras e podem ser resultantes da plasticidade observada em proteases de flavivírus sem perdas significativas em sua função.⁵⁷ As mutações descritas podem ainda afetar a infecção viral de outras formas, aumentando a virulência, ou interferindo na resposta imune do hospedeiro, como descrito para as mutações na enzima NS5 metiltransferase pertencente a mesma linhagem viral.⁵¹

3.1.10 Ensaio de inibição

No total foram testados cerca de 1000 compostos, obtendo valor de Z-fator médio de 0.73 para todas as placas testadas. Neste trabalho, o ensaio de inibição foi validado utilizando aprotinina como controle positivo para a inibição (IC₅₀ de 120 nM, dados não apresentados). A plataforma estabelecida tornou possível a realização da triagem de inibidores em placas de 96 poços de forma simples, rápida e confiável. Além disso, o ensaio é reprodutível, relativamente barato e pode ser facilmente escalonado para placas de 384 poços. O maior problema relacionado a este ensaio, e a outros ensaios que utilizam leituras de fluorescência, é que os compostos podem interagir com o fluoróforo utilizado (neste caso AMC), e ocasionar diminuição da fluorescência por meio de *quenching*, levando a obtenção de falsos positivos. No entanto, este problema pode ser facilmente solucionado realizando uma medida adicional com os compostos previamente selecionados, na presença do fluoróforo livre AMC.

Adicionalmente, compostos que possuem absorção no mesmo comprimento de onda do AMC, também não podem ser avaliados neste ensaio.⁷⁶

Dentre os compostos avaliados neste trabalho, estão o *Pandemic Response Box* (PRB) fornecido pela organização *Medicines for Malaria Venture* (MMV), diversos compostos de colaboradores do CIBFar, incluindo uma série de peptídeos sintéticos sintetizados e fornecidos pelo colaborador prof. Dr. Eduardo Maffud Cilli (Universidade Estadual Paulista "Júlio de Mesquita Filho", UNESP, Araraquara, Brasil) e uma série de porfirinas sintetizadas e fornecidas pelo prof. Dr. Kleber Thiago de Oliveira (Universidade Federal de São Carlos, UFSCar, São Carlos, Brasil). De todos os compostos testados, 18 deles apresentaram inibição superior à 80% e foram utilizados para obtenção de curvas de dose-resposta. Os compostos mais potentes foram testados quanto a atividade antiviral em ensaios baseados em replicon utilizando a linhagem celular BHK21-RepYFV-17D.

Os resultados respectivos a duas das séries de compostos estão apresentados nos itens a seguir.

3.1.10.1 Pandemic Response Box (MMV)

Os compostos da biblioteca Pandemic Response Box foram inicialmente avaliados no teste enzimático pontual em concentração de 20 μM . Esta biblioteca, fornecida pela organização *Medicines for Malaria Venture* (MMV) contém diversos compostos *drug-like*, todos eles já descritos por terem algum tipo de atividade e, alguns já são fármacos aprovados enquanto outros encontram-se em fase de desenvolvimento. Dos 400 compostos, 201 são antibacterianos, 153 antivirais e 46 antifúngicos. Destes compostos, quatro deles apresentaram atividade inferior a 20% e foram selecionados para obtenção de curvas dose-resposta. O resultado da triagem inicial está apresentado na Figura 24A-B. Os valores de IC_{50} obtidos foram respectivamente de $10.7 \pm 0.6 \mu\text{M}$, $1.79 \pm 0.06 \mu\text{M}$ e $3.6 \pm 0.2 \mu\text{M}$ para os compostos MMV 1581548, MMV 396785, MMV 1580853 (Figura 24C).

Destes, os compostos MMV 396785 e MMV 1580853 tiveram a maior potência. Apesar de nenhum deles ser descrito como antiviral, dados os valores de IC_{50} destes compostos ambos foram considerados promissores, e foram utilizados em experimentos de co-cristalização. Embora tenha sido possível obter cristais, estes difrataram em baixa resolução (cerca de 4 Å). Em seguida, os compostos tiveram atividade antiviral avaliada em ensaios celulares utilizando o replicon de febre amarela.

O composto MMV 1580853 não apresentou inibição significativa nos ensaios celulares, enquanto o composto MMV 396785 apresentou resultados iniciais promissores, com EC_{50} de $1.5 \pm 0.9 \mu\text{M}$ (gráfico não apresentado). A citotoxicidade do composto não foi avaliada.

O composto MMV 396785, conhecido como Alexidina, é um fármaco aprovado com ação antibacteriana e antifúngica de amplo espectro, comumente presente em soluções para limpeza de lentes e enxaguantes bucais. Além disso também já foi descrito como eficaz contra *S. mansoni*.⁸⁷⁻⁸⁹ O mecanismo de ação sugerido é que devido à natureza da molécula o grupo 2-etil-hexil consegue interagir com os lipídeos formadores das membranas celulares, sendo assim capaz de impedir a produção dos grupos lipídicos presentes nas membranas. Além disso, os grupos biguanidinas são capazes de interagir com os fosfolipídeos já presentes na membrana plasmática, levando a sua desestabilização e consequente ruptura.⁹⁰ Embora a citotoxicidade do composto não tenha sido avaliada neste trabalho, estudos mostram que o valor é em torno de $10 \mu\text{M}$,⁸⁷ o que implicaria num índice de seletividade de ~ 6.6 . Vale ainda ressaltar que embora os resultados iniciais sejam promissores, é necessário otimizar a molécula, seja por QSAR ou por SBDD, de forma a otimizar sua inibição da protease e atividade antiviral, tornando-a um inibidor específico.

Já o composto MMV 1580853, também apresentou um baixo IC_{50} comparado a outros inibidores de proteases de flavivírus, no entanto não foi capaz de inibir a replicação de RNA no ensaio utilizando replicon. A molécula é descrita como tendo eficaz atividade antibacteriana, possivelmente inibindo a via de síntese de isoprenóides. O composto também tem atividade descrita como antimalárico, antifúngico e mais recentemente foi descrito como inibidor da enzima nsp15 endorribonuclease de SARS-CoV-2.⁹¹⁻⁹⁴ Diversos motivos podem explicar a não eficácia do composto no ensaio celular, como baixa permeabilidade ou biodisponibilidade. Tal como o composto anterior, podemos utilizar os métodos de planejamento de fármacos, tendo como enfoque melhorar a eficácia do composto em células, mantendo sua atividade contra a enzima.

O composto MMV 1634402, capaz de inibir a enzima NS2B-NS3^{pro} de ZIKV com elevada potência (IC_{50} de $0.36 \pm 0.08 \mu\text{M}$),⁷⁶ não demonstrou potência similar contra a protease de YFV ($IC_{50} > 10 \mu\text{M}$), sugerindo ser um inibidor específico da protease de ZIKV. Ambos os compostos que apresentaram $IC_{50} > 10 \mu\text{M}$ (MMV 1581548 e MMV 1634402), devido a sua baixa potência, não foram avaliados no ensaio celular e não serão discutidos em detalhes.

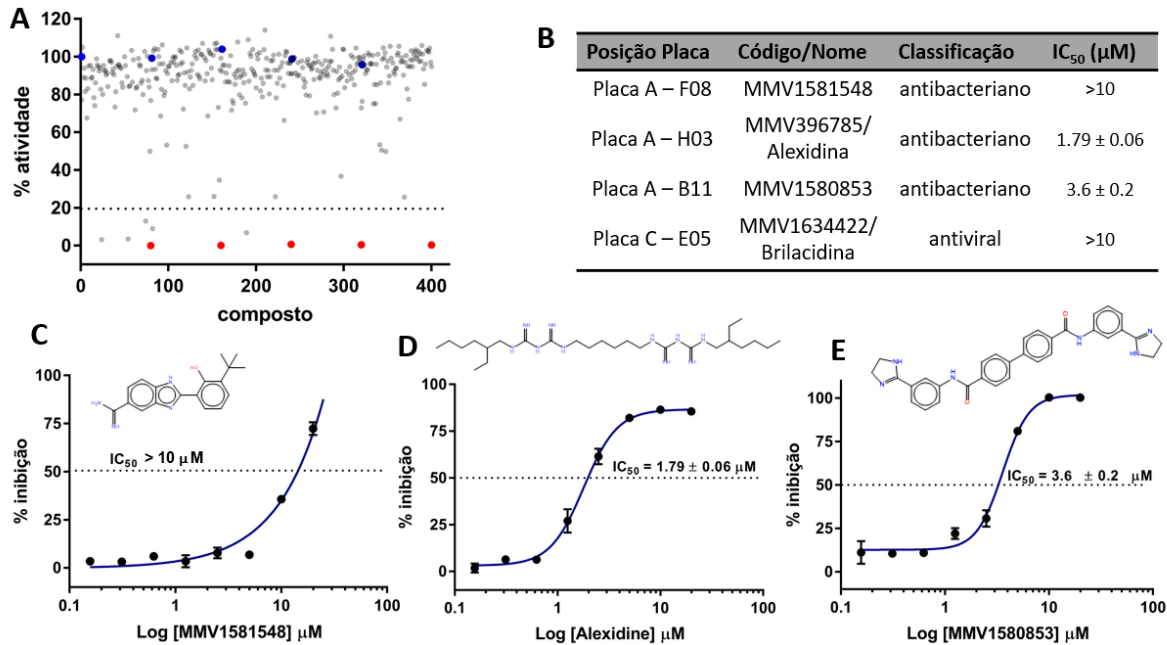


Figura 24 - Resultados da triagem de compostos do Pandemic Response Box. A) Gráfico contendo os resultados da triagem inicial dos compostos, com suas respectivas porcentagens de atividade no teste pontual. Em azul, estão representados os controles 1% DMSO de cada uma das placas, em vermelho as reações branco de cada uma das placas e em cinza cada um dos compostos avaliados. B) Tabela contendo os resultados referentes aos compostos selecionados, suas posições na placa, códigos MMV, classificação e respectivos IC₅₀ (em μM). C) Gráfico de dose-resposta do composto MMV 1581548. D) Gráfico de dose-resposta do composto MMV 396785 (Alexidina). E) Gráfico de dose-resposta do composto MMV 1580853.

Fonte: Elaborada pela autora.

3.1.10.2 Peptídeos da série Hecate

Dentro da triagem dos compostos provenientes da biblioteca do CIBFar, identificamos os peptídeos do colaborador Prof. Dr. Eduardo M. Cilli (IQ-Ar/UNESP), Hecate e AG-Hecate como potentes inibidores da enzima NS2B-NS3^{pro} com valores de IC₅₀ de respectivamente 0.12 ± 0.01 nM e 290 ± 40 nM (Figura 26A-B). Estes peptídeos, além de outros derivados, são descritos como inibidores da replicação do vírus da Hepatite C (HCV), sendo que o peptídeo Hecate tem valor de EC₅₀ de 8.57 μM e o AG-Hecate de 11.18 μM, atuam em diversas etapas do ciclo viral, incluindo a entrada e saída e do vírus nas células e impedindo a replicação do RNA viral.⁹⁵

Na sequência, em colaboração com o Prof. Dr. Eduardo M. Cilli foi sintetizada uma série de peptídeos derivados do Hecate a fim de avaliar a influência de diferentes modificações na região N-terminal do peptídeo (lisina, acetil e ácido benzoico). Os valores de IC₅₀ obtidos foram de 360 ± 30 nM, 160 ± 20 nM e 72 ± 1 nM respectivamente (Figura 26C-E). Pelos valores

obtidos, ficou claro que ambas as modificações não afetam significativamente o valor de IC_{50} , sendo similares ao valor do peptídeo sem nenhuma modificação no N-terminal (Hecate) e que a única modificação capaz de aumentar a potência do composto foi o ácido gálico (AG). As respectivas sequências dos derivados de Hecate, suas porcentagens de inibição no teste pontual a 20 μM e os valores de IC_{50} estão sumarizados na Tabela 6.

Buscando maior compreensão sobre o mecanismo de inibição do peptídeo AG-Hecate na enzima, este foi determinado realizando diluição seriada do substrato (de 300 μM até 18.8 μM) e variação da concentração de inibidor (0, 0.15 e 0.3 nM AG-Hecate), sendo esta fixa para cada uma das curvas. O mecanismo de inibição foi então obtido pela análise dos parâmetros cinéticos (K_M , k_{cat} e V_{max}), obtidos do *fitting* de Michaelis-Menten e dos gráficos de Lineweaver-Burk (Figura 25A-B).

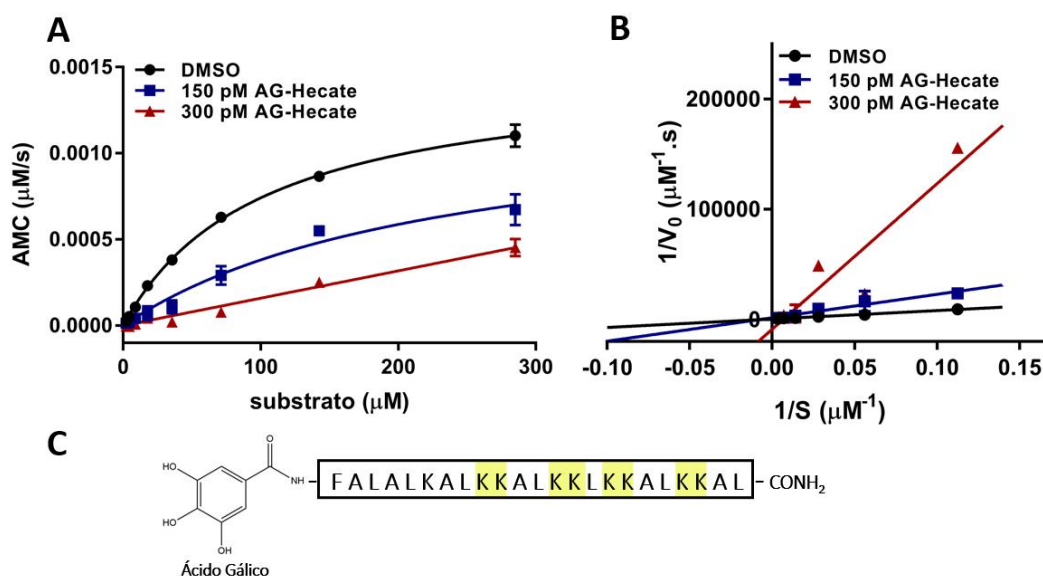


Figura 25 - Mecanismo de inibição do composto AG-Hecate na enzima NS2B-NS3^{pro}. A) Curvas de Michaelis-Menten para as concentrações de 0, 150 pM e 300 pM de AG-Hecate. As barras de erros correspondem ao desvio padrão das triplicatas. B) Curvas de Lineweaver-Burk para as concentrações de 0, 150 pM e 300 pM de AG-Hecate, obtidas da linearização dos dados utilizados na curva de Michaelis-Menten. C) Representação esquemática da sequência do peptídeo AG-Hecate, contendo ácido gálico no N-terminal, destacando os potenciais sítios de clivagem pela protease, coloridos na figura em amarelo.

Fonte: Elaborada pela autora.

Os valores de K_M obtidos foram $102 \pm 6 \mu M$ e $240 \pm 70 \mu M$ para o controle 1% DMSO e 150 pM AG-Hecate respectivamente. Já os valores de V_{max} foram para de $0.00149 \pm 0.00003 \mu M.s^{-1}$, $0.0013 \pm 0.0002 \mu M.s^{-1}$ o controle 1% DMSO e 150 pM AG-Hecate respectivamente. Na concentração de 300 pM AG-Hecate os valores de K_M e V_{max} não puderam ser determinados.

Pelos valores obtidos, é notório que nas diferentes concentrações de inibidor, o valor de V_{max} não é alterado, enquanto o valor de K_m aumenta de acordo com concentração crescente do inibidor, ou seja, quanto maior a concentração de AG-Hecate, menor a afinidade da enzima pelo substrato, indicando que o peptídeo é um inibidor do tipo competitivo. O mecanismo de inibição é evidenciado também pela curva de Lineweaver-Burk (Figura 25B) onde as curvas para diferentes concentrações de inibidor têm ponto de interseção comum no eixo y, caracterizando um inibidor competitivo.

Tipicamente, a sequência preferencial de clivagem observada em proteases de flavivírus é constituída por dois resíduos de carga positiva nas posições P2 e P3 (lisina ou arginina) com um resíduo de cadeia lateral curta ocupando a posição P1' (comumente glicina, leucina, serina e alanina).^{96,97} Pela análise da sequência do peptídeo, considerando a sequência de reconhecimento da enzima existem 4 potenciais regiões de clivagem do peptídeo pela protease no AG-Hecate (Figura 25C). Baseando-se nisso, foram então sintetizados 4 peptídeos contendo diferentes porções do AG-Hecate, definidas pela localização dos possíveis sítios de clivagem a fim de identificar qual delas é importante para o reconhecimento do peptídeo pela enzima (peptídeos PSSct 1901, PSSct 1905, PSSct 1906 e PSSct 1907). Destes peptídeos, o único que apresentou inibição significativa no teste pontual a 20 μM foi o PSSct 1905, com IC_{50} de 9.4 ± 0.5 nM (Figura 26F). Portanto, embora houve uma diminuição expressiva na potência do peptídeo comparado ao AG-Hecate, este ainda continua sendo um potente inibidor da enzima, e o valor de IC_{50} ainda se mantém em baixo nanomolar. Além disso, a diminuição da massa molecular do peptídeo pode ter efeito positivo no aumento de permeabilidade celular em futuros ensaios de inibição *in-situ*.

Tabela 6- Série de peptídeos derivados do Hecate, suas respectivas sequências, porcentagem de inibição no teste pontual a 20 μM e valores de IC_{50} (em nM).

Composto	Sequência	Atividade 20 μM (%)	IC_{50} (nM)
AG-Hecate	Ácido Gálico-FALALKALKKKALKKALKKKAL-CONH ₂	0.0	0.12 ± 0.01
Hecate	FALALKALKKKALKKALKKKAL-CONH ₂	3.2	290 ± 40
K-Hecate	K-FALALKALKKKALKKALKKKAL-CONH ₂	7.4	360 ± 30
Acetil-Hecate	Acetil-FALALKALKKKALKKALKKKAL-CONH ₂	7.1	160 ± 20
Bz-Hecate	Bz-FALALKALKKKALKKALKKKAL-CONH ₂	5.5	72 ± 1
PSSct1905	Ácido Gálico-FALALKALKKKALKKKAL-CONH ₂	22.6	9.4 ± 0.5
PSSct1906	Ácido Gálico-FALALKAL-CONH ₂	59.5	>20
PSSct1907	Ácido Gálico-FALALKALKKKAL-CONH ₂	61.9	>20
PSSct1901	Ácido Gálico-FALALKALKKKALKKAL-CONH ₂	89.8	>20

Fonte: Elaborada pela autora.

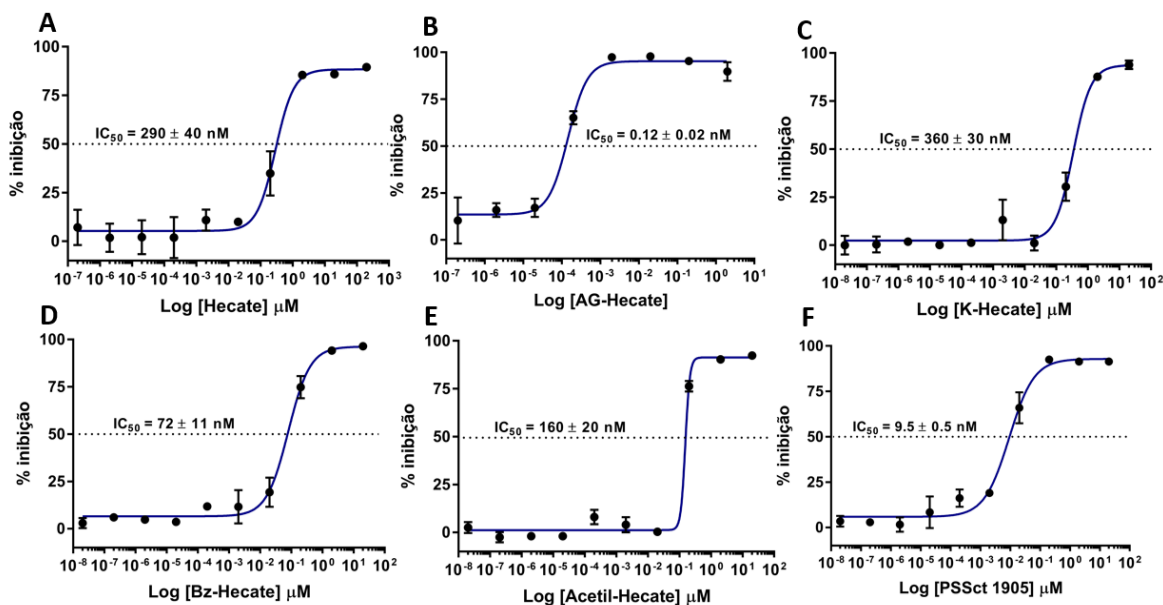


Figura 26 - Gráficos de dose-resposta dos peptídeos da série Hecate. A) Dose-resposta do peptídeo Hecate. B) Dose-resposta do peptídeo AG-Hecate. C) Dose-resposta do peptídeo K-Hecate. D) Dose-resposta do peptídeo Bz-Hecate. E) Dose-resposta do peptídeo Acetil-Hecate. F) Dose-resposta do peptídeo PSSct 1905.

Fonte: Elaborada pela autora.

Em seguida, foi avaliada a potência inibitória dos peptídeos da série Hecate quanto à atividade antiviral utilizando o replicon de YFV. Os peptídeos mais potentes da série, AG-Hecate, Bz-Hecate e PSSct 1905 foram utilizados nos ensaios celulares. Os valores de EC_{50} obtidos foram respectivamente de $1.5 \pm 0.4 \mu\text{M}$, $3.0 \pm 1.0 \mu\text{M}$ e $5.4 \pm 0.7 \mu\text{M}$ (Figura 27A-C) e os valores de CC_{50} foram respectivamente de $7 \pm 1 \mu\text{M}$, $16 \pm 5 \mu\text{M}$ e $> 20 \mu\text{M}$ (Figura 27D-F), resultando em valores de seletividade de 5, 5 e > 4 . Ambos os peptídeos da série apresentaram valores de EC_{50} , CC_{50} e seletividade na mesma faixa, sendo que o peptídeo AG-Hecate foi o mais potente dos três, de forma similar ao observado para os ensaios de atividade enzimática. No entanto, o valor de EC_{50} foi muito maior do que esperado podendo ser resultante da baixa permeabilidade devido ao alto peso molecular das moléculas testadas. Para o peptídeo PSSct 1905, embora tenha apresentado IC_{50} duas ordens de grandeza maior do que o AG-Hecate, teve EC_{50} similar e citotoxicidade menor, sendo o menos tóxico dos três e com maior índice de seletividade, demonstrando que a abordagem de reduzir o tamanho do peptídeo teve resultados satisfatórios nos ensaios celulares. Logo, embora ensaios complementares ainda sejam necessários, considerando os resultados obtidos até o momento, os peptídeos da série aparentam ser promissores inibidores da replicação do RNA viral e podem ser otimizados com base em suas propriedades, visando reduzir a toxicidade e aumentar a permeabilidade celular.

Ensaio de cristalização da enzima NS2B-NS3^{pro} na presença dos peptídeos já foram realizados, porém não foi possível obter cristais.

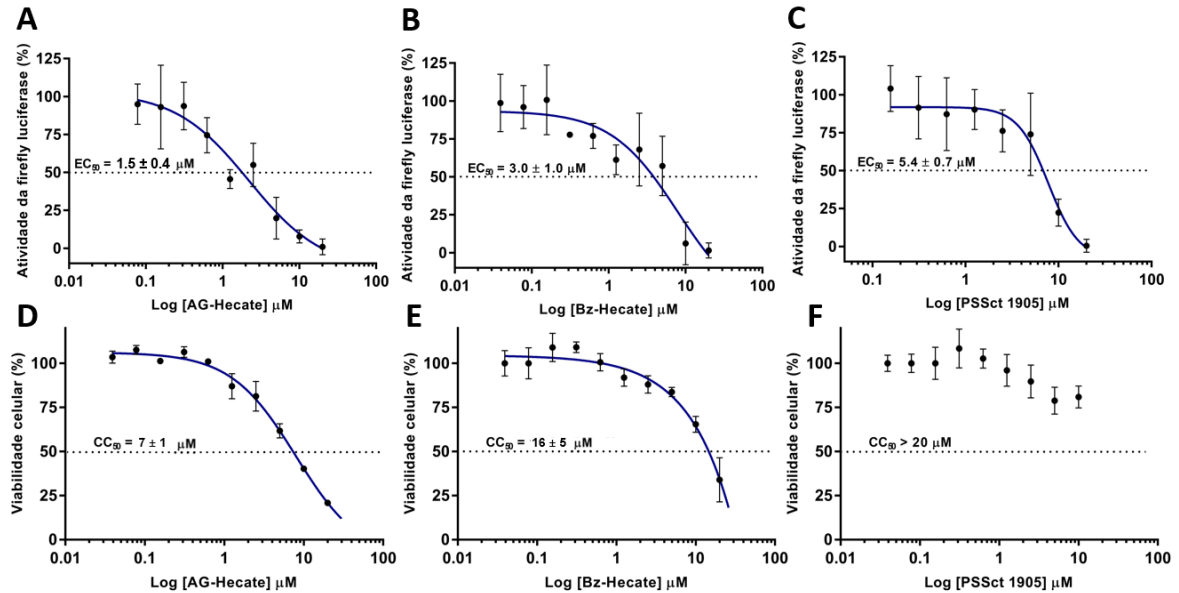


Figura 27 - Avaliação da atividade antiviral dos peptídeos A) Hecate, B) Bz-Hecate e C) PSSct 1905 em células BHK21-RepYFV-17D e citotoxicidade (D-F) respectivamente.

Fonte: Elaborada pela autora.

3.2 Construção NS2B-NS3^{pro} *unlinked*

3.2.1 Clonagem e testes de expressão

Os resultados da clonagem referentes a construção do vetor pET_DUET-1, pode ser observada na Figura 29A, onde é possível observar que ambas as etapas de clonagem foram realizadas com sucesso. Uma das colônias positivas foi isolada, utilizada para propagação do plasmídeo. O mapa do plasmídeo está apresentado na Figura 28.

Após confirmação da clonagem pela PCR de colônia e posterior sequenciamento, o vetor pET_DUET-1 contendo a sequência de interesse foi utilizado para transformar a linhagem *Rosetta 2* de *E. coli*, baseados nos resultados da expressão observados para a construção anterior (*linked*). Na Figura 29B, é possível observar o SDS-PAGE contendo os resultados do teste de expressão. Desta forma, pode-se notar que o nível de expressão é significativamente maior no meio da cultura TB, assim como a concentração de IPTG não provoca alterações significativas.

Assim, foi estabelecido o protocolo de expressão em larga utilizando meio TB e indução com 0.5 mM de IPTG.

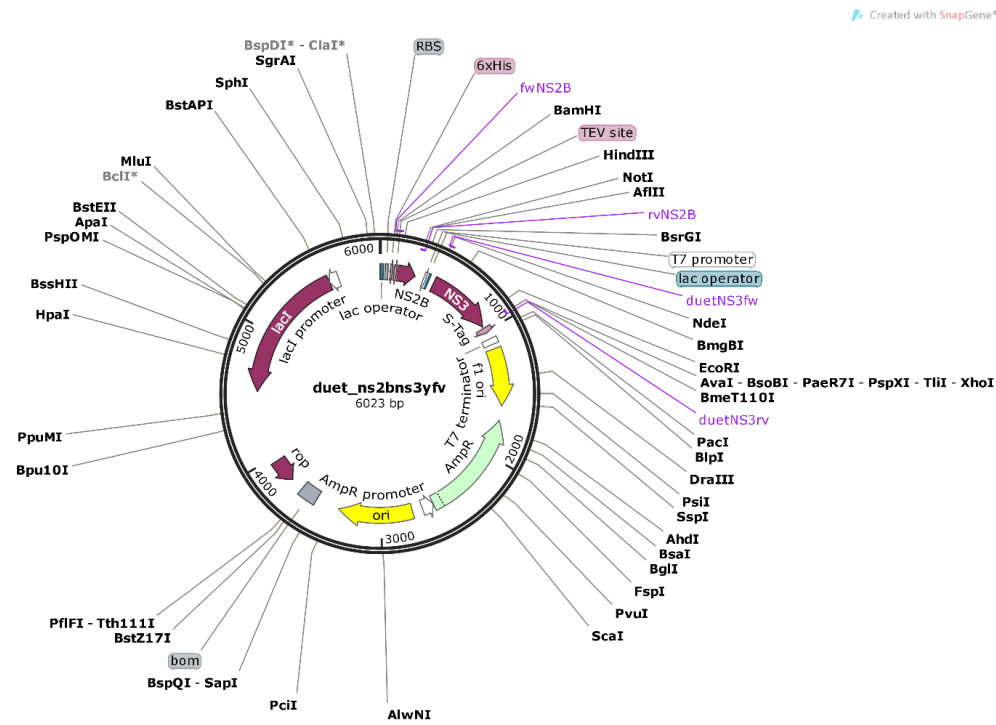


Figura 28 - Mapa do plasmídeo pET_DUET-1 contendo a sequência codificante da enzima NS2B-NS3^{pro} de YFV.

Fonte: Elaborada pela autora.

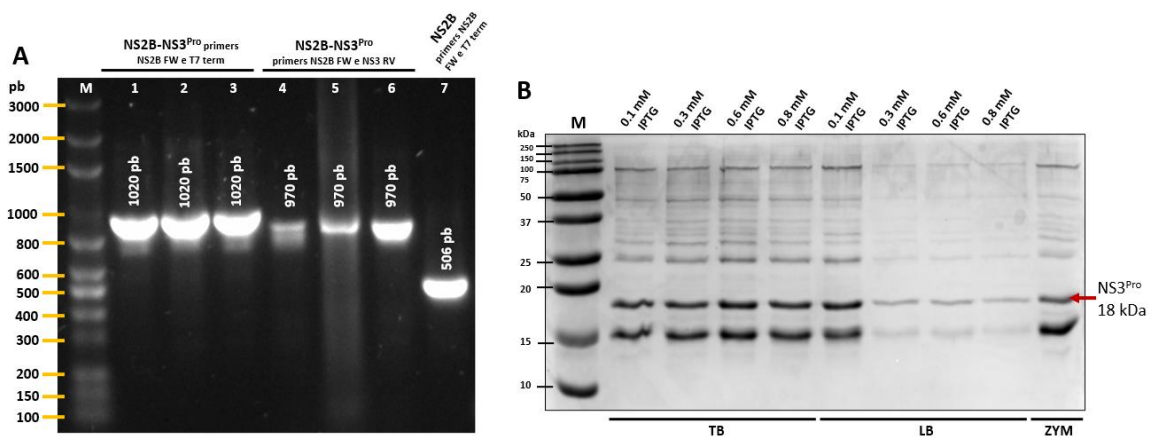


Figura 29 - Gel de agarose 1% contendo os fragmentos obtidos para o vetor pET_DUET-1. As bandas de 1-6 correspondem a colônias contendo ambos os insertos, NS2B e NS3^{pro}. A banda indicada por M representa o marcador molecular (em pb). As bandas 1-3 correspondem a PCR realizada com os primers NS2Bduetfw e T7 term (1020 pb), enquanto as bandas de 4-6 correspondem a PCR realizada com os primers NS2Bduetfw e NS3prorv (970 pb). A banda 7 corresponde a uma das colônias positivas da primeira etapa de clonagem contendo apenas o inserto correspondente ao cofator NS2B. A PCR foi realizada com os primers NS2Bduetfw e T7 term (506 pb). B) SDS-PAGE 15% contendo os resultados dos testes de expressão da construção *unlinked*. No gel, vemos as frações celulares solúveis de cada uma das condições testadas, indicadas na figura. A banda indicada por M representa o marcador molecular (em kDa). A proteína de interesse está indicada pela seta vermelha (NS3 protease: 18 kDa).

Fonte: Elaborada pela autora.

3.2.2 Expressão em larga escala e purificação

A proteína NS2B-NS3^{pro} *unlinked* foi expressa em larga escala utilizando a linhagem de expressão de *E. coli Rosetta 2* em meio TB e indução com 0.5 mM de IPTG, de acordo com as condições otimizadas obtidas dos testes de expressão. Em seguida foram purificadas de maneira análoga ao método empregado para a construção *linked*.

Os resultados da purificação da construção *unlinked* estão representados na Figura 30, na qual podemos observar o cromatograma obtido na etapa de exclusão molecular (Figura 30A) e o SDS-PAGE (Figura 30B). Neste caso, o volume de retenção foi de 74.1 mL, com massa molecular calculada de 26.2 kDa, o que permite inferir que a proteína se encontra em estado monomérico. O SDS-PAGE, juntamente com o cromatograma obtido nos permite inferir que a amostra se encontra homogênea e com elevado grau de pureza. O rendimento aproximado obtido foi de 1.3 mg de proteína por litro de cultura, cerca de 7 vezes menor do que o obtido para a construção ligada covalentemente, demonstrando a eficiência do *linker* na otimização do rendimento da purificação.

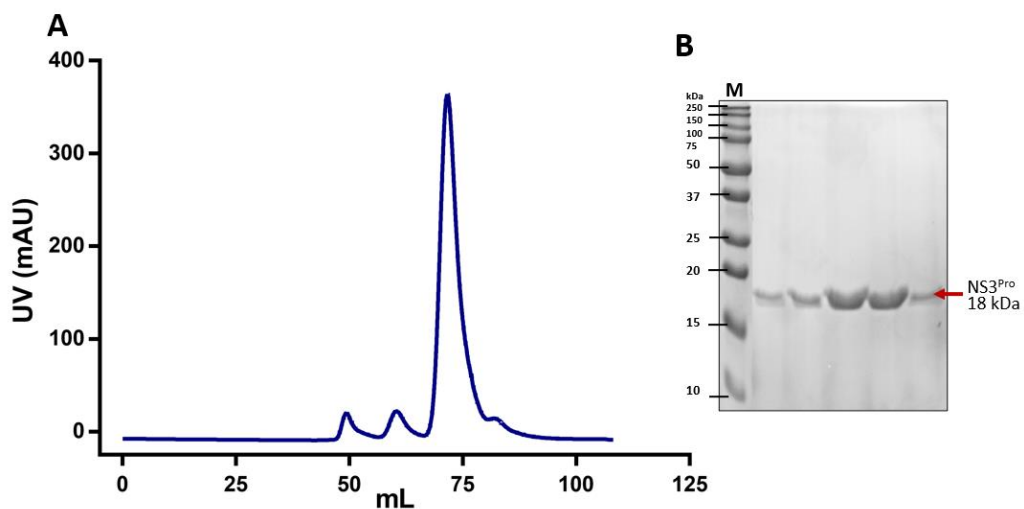


Figura 30 - Resultados da purificação da proteína NS2B-NS3^{pro} *unlinked*. A) Cromatograma obtido na etapa de exclusão molecular. B) SDS-PAGE 15% mostrando a pureza da amostra ao final da purificação, a proteína está indicada pela seta vermelha (NS3^{pro}: 18 kDa).

Fonte: Elaborada pela autora.

3.2.3 Caracterização da enzima

A proteína NS2B-NS3^{pro} construção *unlinked*, foi caracterizada utilizando o ensaio de atividade baseado em fluorescência, e a técnica de DSF. No ensaio de atividade, de maneira similar ao observado para a construção *linked*, foi observado aumento da fluorescência ao longo do tempo demonstrando que a proteína realizou a clivagem do substrato peptídico e encontra-se ativa. Como mostrado na Figura 31A foi realizada a variação da concentração de substrato (de 600 μM até 18.8 μM) para determinação dos parâmetros cinéticos da enzima pelo *fitting* da equação de Michaelis-Menten.

Os valores de afinidade pelo substrato (K_m) e k_{cat} da enzima NS2B-NS3^{pro} *unlinked* foram respectivamente $70 \pm 10 \mu\text{M}$ e $0.19 \pm 0.08 \text{ s}^{-1}$, comparados a $90 \pm 10 \mu\text{M}$ e $0.24 \pm 0.03 \text{ s}^{-1}$ obtidos para a construção *linked*. Portanto, houve aumento da afinidade pelo substrato Bz-nKRR-AMC como esperado, porém, houve diminuição da velocidade máxima da reação, resultando em uma eficiência catalítica apenas 5% maior para a construção *unlinked*. A presença do *linker* é caracterizada por bloquear parcialmente a entrada do sítio ativo e do bolsão de ligação ao substrato (S1-S4)^{60,81} as construções *unlinked* de DENV2 e ZIKV tiveram parâmetros cinéticos determinados e demonstraram ter maior afinidade pelo substrato, com eficiências catalíticas de 2 e 3-5 vezes maior respectivamente.^{60,81,83}

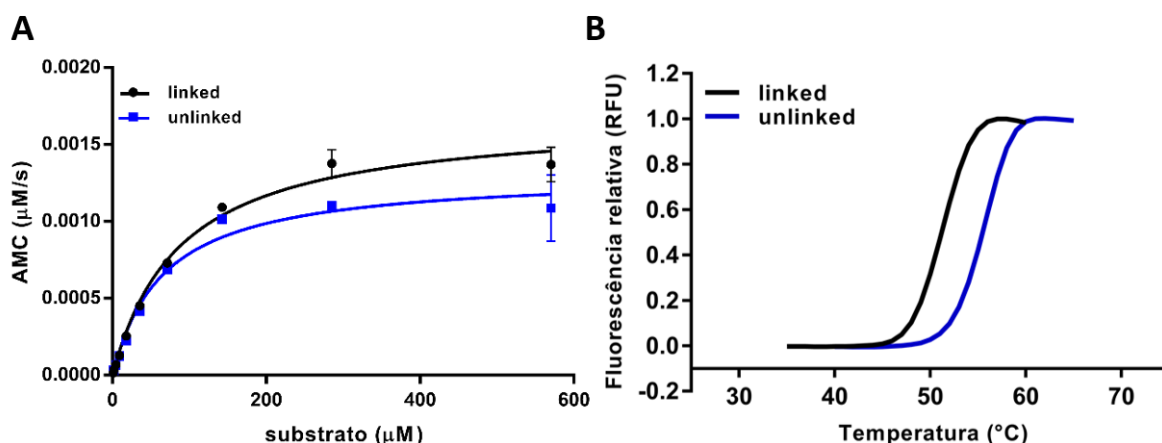


Figura 31 - A) Curvas de Michaelis-Menten para cada uma das enzimas WT *linked* e *unlinked*. As barras de erros correspondem ao desvio padrão das triplicatas. B) Curvas obtidas na DSF para cada uma das construções.

Fonte: Elaborada pela autora.

Através da técnica de DSF (*Differential Scanning Fluorimetry*) foram determinadas as T_m das enzimas NS2B-NS3^{pro} *linked* e *unlinked*. Os valores obtidos foram $51 \pm 1^\circ\text{C}$ e $55.4 \pm 0.5^\circ\text{C}$ respectivamente. Logo, a protease *unlinked* mostrou um aumento aproximado de 3°C com relação a construção *linked*, sugerindo ter uma maior estabilidade térmica (Figura 31B).

3.2.4 Ensaios de cristalização

Num ensaio inicial de cristalização, realizado pelo método de matriz esparsa utilizando kits comerciais, foi possível obter cristais da proteína. Os cristais foram observados em diversas condições do kit PACT (Figura 32), ambos na temperatura de 18°C após cerca de uma semana. Estes cristais não tiveram difração testada, e por conta de seu tamanho, precisariam ser otimizados. O menor tempo necessário para o aparecimento de cristais comparado a construção *linked*, pode significar que a amostra é mais homogênea, considerando que, conforme já previamente mencionado, construções não ligadas covalentemente tendem a aumentar a proporção de enzimas em conformação fechada em solução. No entanto, conforme mencionado na introdução deste trabalho, os experimentos relacionados a NS2B-NS3^{pro} não tiveram continuidade, pois, devido a pandemia, tiveram início os estudos com as proteases de SARS-CoV-2, em particular com a enzima M^{pro}, os quais serão apresentados em detalhes nos próximos capítulos.

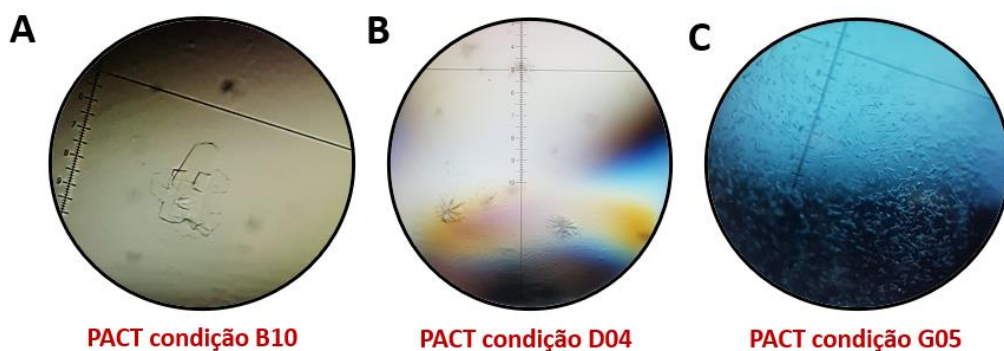


Figura 32 - Cristais da enzima NS2B-NS3^{pro} YFV *unlinked*. A) Cristais da proteína na condição B10 PACT (0.2 M cloreto de magnésio, 0.1 M MES pH 6, 20% PEG 6000). B) Cristais da proteína na condição D04 PACT (0.1 M MMT pH 7, 25% PEG 1500). C) Cristais da proteína na condição G05 PACT (0.2 M nitrato de sódio, 20% PEG 3350).

Fonte: Elaborada pela autora.

4 CONCLUSÕES E PERSPECTIVAS

A febre amarela é uma doença grave, com uma mortalidade que pode ultrapassar 50% nas fases mais severas da doença. Como o único meio de combater a doença é através do tratamento dos sintomas, é urgente o desenvolvimento de compostos com atividade antiviral. A NS2B-NS3^{pro} do YFV é um alvo de grande importância, porém a ausência de informações estruturais sobre essa proteína dificulta o desenvolvimento de novas substâncias com atividade antiviral. Nesse trabalho avançamos na clonagem, expressão e purificação do complexo NS2B-NS3^{pro} do YFV. Desenvolvemos uma construção que possui alto rendimento e grande pureza no final do processo de purificação. Obtivemos cristais da proteína, que possibilitaram a resolução da estrutura cristalográfica da enzima, com modelo final em 2.9 Å. A enzima demonstrou elevada atividade proteolítica contra o peptídeo sintético fluorescente, o qual foi utilizado para triagem de uma biblioteca de compostos. Dos compostos selecionados, três deles apresentaram elevada potência nos ensaios enzimáticos (abaixo de 1 µM), assim como demonstraram atividade antiviral avaliada utilizando replicon de YFV.

Além disso, visando melhorar a qualidade dos cristais obtidos para obtenção de estruturas em complexo com inibidores, produzimos também a enzima NS2B-NS3^{pro} em sua forma não ligada covalentemente, mais similar a enzima nativa. Esta construção, aparentemente mais estável que a construção *linked*, também pode ser cristalizada em tempo significativamente menor. Como perspectivas para o trabalho, em relação aos compostos selecionados, é fundamental a obtenção de estruturas da enzima em complexo com os inibidores, a ser utilizada para otimização dos mesmos utilizando os métodos de planejamento de fármacos, aliada aos ensaios de atividade, ensaios *in-situ*, ensaios fenotípicos antivirais posteriormente ensaios *in-vivo*, de modo a obter um potencial candidato antiviral.

CAPÍTULO 2

Caracterização em solução e elucidação estrutural do processo de maturação da enzima Main Protease (M^{pro}) de SARS-CoV-2

1 INTRODUÇÃO

1.1 SARS-CoV-2 e pandemia de COVID-19

O vírus da síndrome respiratória aguda grave coronavírus 2 (SARS-CoV-2) é o agente causador da COVID-19, uma doença altamente infecciosa que altamente contagiosa e causadora atual pandemia.⁹⁸ O vírus foi identificado pela primeira vez em dezembro de 2019 na província de Wuhan, na China.⁹⁸⁻⁹⁹ Desde então, se espalhou rapidamente em proporções mundiais, afetando de forma impactante a economia e os sistemas de saúde pública. Devido a isso, em março de 2020, a Organização mundial de Saúde declarou a COVID-19 como pandemia.¹⁰⁰

Diante das graves consequências geradas por essa pandemia, um rápido desenvolvimento de candidatos vacinais ocorreu de forma nunca vivenciada pela comunidade científica e população mundial. Atualmente, já existem diversas vacinas aprovadas, sendo as principais aplicadas no Brasil as vacinas da Pfizer/BionTech, Sinovac/Coronovac e Astra-Zeneca/Oxford.¹⁰¹⁻¹⁰² Embora as taxas de cobertura vacinal estão subindo rapidamente, chegando a mais de 80% no Brasil em março de 2023,¹⁰³ a exata eficácia e duração da imunidade vacinal ainda é desconhecida e o surgimento de variantes dificulta o controle do vírus.¹⁰²

Portanto, mesmo no contexto da existência da vacina, faz-se necessária a busca por alternativas terapêuticas, como os antivirais, uma vez que a cobertura e eficácia vacinal sugerem que casos de COVID-19 ainda continuarão a acontecer.

1.2 Biologia estrutural do vírus SARS-CoV-2

O SARS-CoV-2 ou *Severe Acute Respiratory Syndrome Virus 2* é um vírus de RNA envelopado pertencente à linhagem dos β -coronavírus, que inclui também o SARS-CoV e o Vírus Respiratório do Oriente Médio (MERS-CoV), os quais compartilham similaridade sequencial de 79.6% e ~50%, respectivamente.^{99,104-105} O vírus, que possivelmente teve origem proveniente do morcego, teve primeiro caso reportado em dezembro de 2019 na província de Wuhan, China, rapidamente espalhou em escala global, sendo responsável pela atual pandemia de COVID-19.¹⁰⁵

O genoma viral do SARS-CoV-2, é composto por uma fita simples de RNA positiva de aproximadamente 30 Kb que contém a ORF1ab, que codifica duas poliproteínas (pp1a e pp1ab)

contendo as proteínas não estruturais, essenciais para a replicação viral.¹⁰⁶⁻¹⁰⁷ Além de codificar as proteínas estruturais (membrana, nucleocapsídeo, envelope e spike), codifica ainda nove proteínas acessórias (ORF3a, ORF3b, ORF6, ORF7a, ORF7b, ORF8, ORF9b, ORF9c e ORF10) que tem importante papel na interação com o hospedeiro, afetando imunidade inata, assim como a proliferação viral e virulência (Figura 33).¹⁰⁶⁻¹⁰⁸ As poliproteínas pp1a e pp1ab são clivadas pela ação das proteases virais Papain-like-protease (PL^{pro}) e Main protease (M^{pro}), liberando assim as proteínas não estruturais maduras (Figura 33).¹⁰⁹

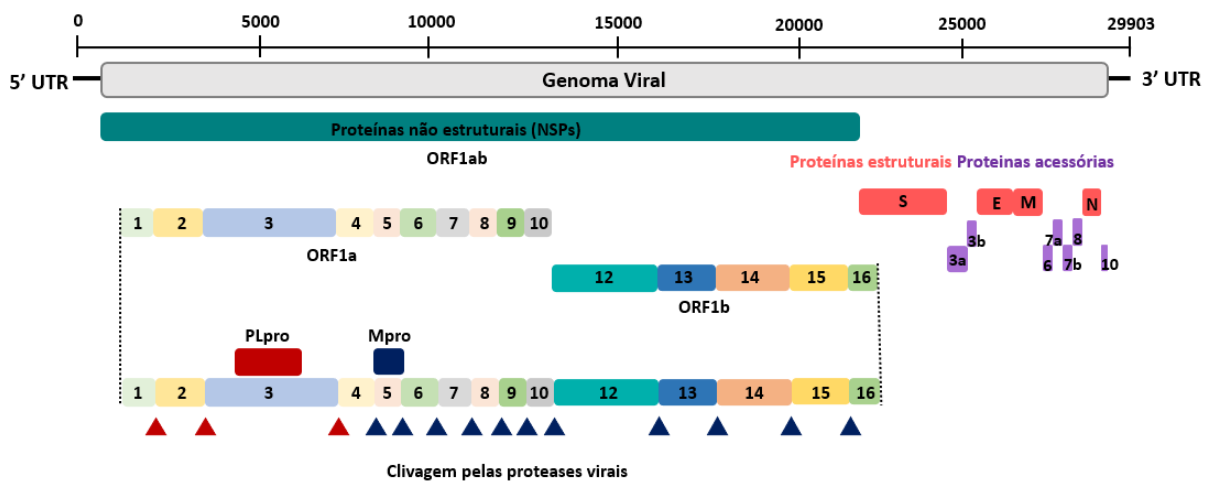


Figura 33 - Representação esquemática do genoma viral do SARS-CoV-2 e da poliproteína viral pp1ab. As proteínas não estruturais (NSPs) estão representadas em verde, as proteínas estruturais em vermelho e as proteínas acessórias em roxo. As setas estão indicando os sítios de clivagem de cada uma das proteases virais. Em vermelho, estão representados os sítios de clivagem da enzima Papain-like-protease ou PL^{pro} e em azul, os sítios da enzima Main-protease ou M^{pro}.

Fonte: Adaptada de YAN *et al.*¹⁰⁸

1.3 Vacinação e desenvolvimento de antivirais contra a COVID-19

Conforme previamente mencionado, diante das graves consequências geradas por essa pandemia, ocorreu um rápido desenvolvimento de candidatos vacinais. Atualmente, já existem diversas vacinas aprovadas, as quais utilizam diversos mecanismos de ação, sendo as principais aplicadas no Brasil as vacinas da Pfizer/BionTech (RNA mensageiro), Sinovac/Coronavac (vírus inativado) e Astra-Zeneca/Oxford (vetor viral).^{101,110} Embora as taxas de cobertura vacinal estão elevadas, chegando a mais de 80% no Brasil em março de 2023,¹⁰³ a exata eficácia e duração da imunidade vacinal ainda é desconhecida e o surgimento de variantes dificulta o controle do vírus.¹⁰¹ Portanto, mesmo no contexto da existência da vacina, permanece

necessária a busca por antivirais, dado que casos de COVID-19 provavelmente ainda continuarão a acontecer.

Neste contexto, a biologia estrutural teve papel extremamente importante ao auxiliar no desenvolvimento de inibidores contra o SARS-CoV-2. A primeira estrutura foi depositada em fevereiro de 2020 (Main Protease), e em pouco tempo, também foi elucidada a estrutura do capsídeo viral, obtida por microscopia eletrônica.^{106,108,113} Das 29 proteínas codificadas pelo genoma viral, apenas 4 delas não tem estrutura resolvida.¹¹³ Atualmente já existem mais de 6000 estruturas de proteínas de SARS-CoV-2 depositadas no PDB, as quais foram obtidas majoritariamente por cristalografia de raios-X e cryo-EM. Além disso, diversas proteínas de SARS-CoV-2 foram utilizadas para realização de *fragment screening*, dentre elas a Nsp15 endorribonuclease, Nsp3 macrodomínio, Nsp13 helicase e Nsp5 M^{pro}.¹¹⁴⁻¹¹⁷ Vale destacar o *fragment screening* realizado para a enzima M^{pro}, no qual mais de 1500 estruturas já foram resolvidas, com um total de 71 *hits* no sítio ativo até março de 2022,¹¹⁷ no qual utilizando o método de FBDD já foi possível obter promissores candidatos a inibidor da enzima, com IC₅₀ na faixa de nanomolar e EC₅₀ em baixo micromolar.¹¹⁸

Dos candidatos a inibidores da replicação do vírus SARS-CoV-2, grande parte tem como alvo as enzimas Nsp12 polimerase e a Nsp5 M^{pro}. Como exemplos, podemos citar remdesevir, favipiravir, molnupiravir (Ns12 polimerase), carmofur, nirmatrelvir, ensitrelvir e MAT-POS-e194df51-1 (M^{pro}) sendo que dentre estes, o remdesevir, molnupiravir e nirmatrelvir já são aprovados pelo FDA para uso emergencial.¹¹⁸⁻¹²² Existem ainda diversos outros candidatos a antivirais que se encontram atualmente em fase clínica.

1.4 Enzima Main Protease (M^{pro})

A Nsp5 Main Protease (ou Protease Principal, M^{pro} ou 3CL^{pro}) é uma proteína classificada como uma cisteína protease, porém possui enovelamento composto por 2 domínios típicos de quimotripsina, em conjunto com um domínio adicional que está envolvido na sua dimerização. A M^{pro} é responsável pela clivagem das poliproteínas virais pp1a e pp1ab em 11 sítios, incluindo seu próprio N e C-terminal (Figura 33). A enzima é ativa em forma dimérica, na qual ambos os protômeros tem importância para a atividade da enzima (Figura 34).

Dada a sua importância, esta enzima é amplamente caracterizada sendo a primeira proteína de SARS-CoV-2 a ter estrutura elucidada, já possui mais de 700 estruturas cristalográficas depositadas no PDB (em setembro de 2023). Além disso, representa um importante alvo para o desenvolvimento de candidatos antivirais e conforme previamente

mencionado, já possui diversos promissores candidatos a inibidor descritos, incluindo o medicamento já aprovado pelo FDA Paxlovid® (Pfizer), contendo nirmatrelvir como um dos princípios ativos.^{106,119,122}

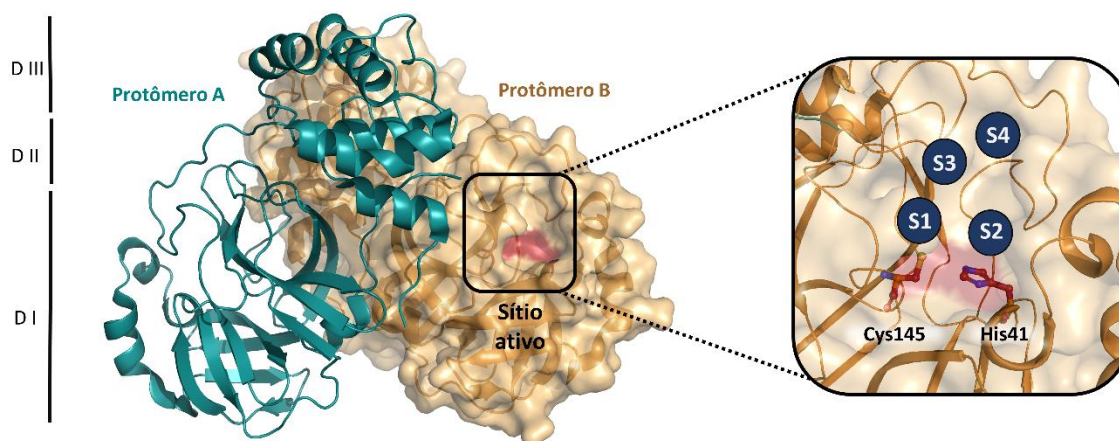


Figura 34 - Representação da estrutura da enzima M^{pro} do vírus SARS-CoV-2. Cada um dos protômeros está colorido em cores diferentes (verde e amarelo). Os resíduos catalíticos estão destacados em vermelho. DI, DII e DIII indicam respectivamente cada um dos domínios da proteína. À direita, visão ampliada do sítio ativo da enzima, composto pela díade dos resíduos Cys145 e His41 (representados como *ball and sticks* e coloridos em vermelho), mostrando também o sítio de ligação ao substrato com os subsítios S1-S4.

Fonte: Elaborada pela autora.

1.5 Processo de maturação da enzima M^{pro} e objetivos do trabalho

Embora a proteína M^{pro} seja extensamente estudada, os detalhes de seu processo de auto-maturação, ou seja, o mecanismo pelo qual a enzima realiza a auto-clivagem de seu N e C-terminal, liberando a forma ativa dimérica, ainda são mal compreendidos.^{109,123-124}

Diversos modelos foram elaborados para explicar a maturação da M^{pro} . O primeiro modelo elaborado em 2005, sugere que pequenas quantidades da enzima em forma ativa e dimérica são formadas, permitindo a clivagem da poliproteína.¹²⁴ Já em 2010, foi observado que a presença de substrato ocasiona um aumento na taxa de dímeros, assim, foi proposto que dois protômeros da enzima formam um dímero transitente, o qual é estabilizado pela ligação ao substrato (N-terminal de outro protômero ainda na poliproteína).^{123,125} Outro modelo sugere ainda que a auto-clivagem do N-terminal necessita apenas de 2 protômeros imaturos na poliproteína (clivagem *cis*), que se associam formando um dímero imaturo intermediário, e este processo é independente da forma madura dimérica da enzima.¹²⁶ Já o estudo mais recente,

publicado em 2020, elucidou a estrutura cristalográfica da enzima em complexo com sua região C-terminal, evidenciando que este processo (clivagem da região C-terminal) ocorre entre dois dímeros na forma madura (clivagem *trans*).¹⁰⁹

O conhecimento dos detalhes moleculares desse mecanismo pode auxiliar no desenvolvimento de inibidores específicos para a M^{pro} baseados em formas intermediárias imaturas da enzima, e impedindo assim que esta atinja a sua forma madura e ativa, consequentemente bloqueando a replicação viral. Desta forma, os objetivos gerais deste projeto foram a utilização de técnicas biofísicas, bioquímicas, juntamente com estudos estruturais, para elucidar o mecanismo molecular do processo de maturação da enzima M^{pro}.

Os objetivos específicos do projeto foram:

- Obtenção da proteína M^{pro} em 3 construções: IMT-M^{pro}, M^{pro} e C145S M^{pro};
- Ensaios biofísicos de caracterização de ambas as proteínas em solução;
- Estudos estruturais de ambas as proteínas utilizando cristalografia e cryo-EM;
- Padronização e realização de ensaios de atividade para caracterização funcional e determinação de parâmetros cinéticos das proteínas;
- Análise estrutural e funcional da influência da presença de inibidores (covalente e não covalente) no processo de maturação da M^{pro};
- Proposição de um modelo para o processo de maturação da proteína

2 MATERIAIS E MÉTODOS

2.1 Clonagem da sequência codificante da enzima M^{pro} em 3 diferentes construções: IMT-M^{pro}, M^{pro} e C145S M^{pro}

A sequência codificante da enzima M^{pro} de SARS-CoV-2, foi amplificada a partir do cDNA (GenBank MT126808.1) produzido a partir do RNA procedente de amostras de SARS-CoV-2 inativadas providas pelo grupo do Prof. Edison Durigon do ICB-USP, foi sintetizado utilizando o kit SCRIPT One-Step RT-PCR kit (Cellco Biotec) e primers hexâmeros randômicos.

Visando a elucidação do processo de auto-maturação da M^{pro}, foram planejadas três diferentes construções para a proteína. Uma chamada imatura ou IMT-M^{pro}, que contém uma inserção de três resíduos não cliváveis na porção N-terminal da proteína (GAM), de forma a mimetizar uma forma imatura ainda não processada da M^{pro} na poliproteína. A segunda construção foi a forma madura da proteína (nativa, ou apenas M^{pro}), que contém a sequência endógena de reconhecimento e auto-clivagem da enzima, que é removida ainda durante a expressão, resultando na forma madura da M^{pro}. Por fim, ainda planejamos um mutante cataliticamente inativo (C145S M^{pro}), substituindo a cisteína catalítica C145 por serina.

Para produção da construção IMT-M^{pro} a sequência correspondente a região codificante da enzima M^{pro} (resíduos 3264–3569) foi amplificada utilizando os primers IMT-Mpro FW e RV (Tabela 7) e inserida no vetor pET_M11-LIC, que codifica uma cauda de 6xHis na porção N-terminal, seguida do sítio de TEV protease (ENLYFQ↓GAM) através do método LIC⁵⁸, resultando no plasmídeo pET_M11-IMT-MPro.

Para obter a forma madura (ou nativa) da enzima, a sequência N-terminal de reconhecimento e auto-clivagem (GAMSAVLQ↓SGFRK) foi inserida no vetor pET_M11-IMT-MPro por PCR inversa utilizando os primers Mpro-FW e Mpro-RV (Tabela 7). O produto da PCR foi tratado com T4 Polynucleotide Kinase (PNK, Thermo Fischer Scientific) e T4 Ligase (Cellco Biotec), formando o plasmídeo pET_M11- Mpro.

Para a construção C145S Mpro, pET_M11-Mpro foi utilizado como molde para PCR inversa com os primers apresentados na Tabela 7 e o produto da PCR foi tratado com T4 Polynucleotide Kinase (PNK, Thermo Fischer Scientific) e T4 Ligase (Cellco Biotec), formando o plasmídeo pET_M11-C145S-Mpro. Todas as PCRs foram realizadas com a enzima FastPol (Cellco Biotec). Todos os plasmídeos foram transformados células de *E. coli* DH5α e

as colônias positivas confirmadas por sequenciamento. A Figura 35 contém uma representação esquemática das construções utilizadas.

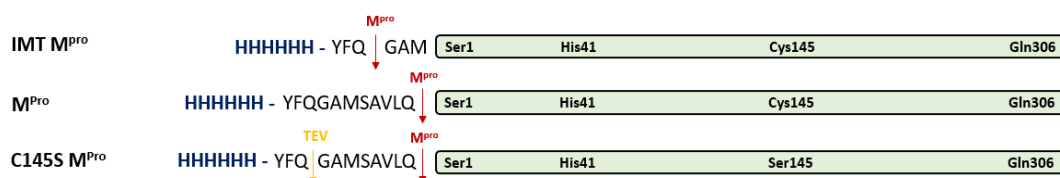


Figura 35 - Representação esquemática das construções da proteína SARS-CoV-2 M^{pro}. A porção correspondente a sequência da proteína está colorida em verde, os sítios de clivagem de proteases estão indicados por setas (vermelho para M^{pro} e amarelo para TEV protease) e a cauda de 6xHis está destacada em azul.

Fonte: Elaborada pela autora.

Tabela 7 - Primers utilizados para a clonagem da sequência codificante da enzima M^{pro} em 3 diferentes construções

Primer	Sequência	T _m (°C)
IMT-Mpro FW	5'- CAGGGCGCCATGAGTGGTTTTAGAAAAATGGCATTG-3'	63
IMT-Mpro RV	5'- GACCCGACGCGGTTATTGGAAAGTAACACCTGAGAC-3'	62
Mpro FW	5'-GCTGCAGAGTGGTTTTAGAAAAATGGCATTG-3'	62
Mpro RV	5'-ACGGCTGACATGGCGCCCTGAAAATA-3'	63
C145S Mpro FW	5'-CCTTAATGGTTCATCTGGTAGTG-3'	61
C145S Mpro RV	5'- AATGAACCCTAATAGTGAAATTGG-3'	62

Fonte: Elaborada pela autora.

2.2 Expressão em larga escala das enzimas

Os plasmídeos resultantes foram utilizados para transformar a linhagem de expressão de *E. coli* BL21. As células contendo os plasmídeos foram crescidas a 37°C, 200 rpm em meio auto indutor ZYM 5052⁶² até atingir DO600 de 0.6. A indução da expressão da proteína foi realizada reduzindo a temperatura para 18°C e crescendo as células por 16 h sob mesma agitação. As células foram obtidas através de centrifugação (5000 x g, 40 min, 4°C) e em seguida ressuspendidas em tampão de lise (20 mM Tris pH 7.8, 150 mM NaCl, 1 mM DTT) e lisadas por sonicação (pulso ON 30s, pulso OFF 30s, 12 min). O extrato celular solúvel foi obtido por centrifugação (12000 x g, 30 min, 4°C).

2.3 Purificação das enzimas

2.3.1 Purificação da enzima IMT-M^{pro}

Após expressão, grande parte da proteína sofre o processo de auto-clivagem resultando na proteína liberada de sua cauda de fusão 6xHIS tag. A pequena fração de enzima ainda com

6xHIS tag é removida do lisado por de cromatografia de afinidade por íon metálico (IMAC) utilizando resina contendo níquel Ni-NTA superflow (Qiagen). A proteína clivada foi purificada utilizando precipitação com sulfato de amônia. Foi adicionado 1 M sulfato de amônia no lisado celular sob agitação e incubação no gelo por 10 min. A proteína precipitada foi removida do extrato celular por centrifugação (12000 x g, 30 min, 4°C). A proteína foi então ressuspendida em tampão de gel filtração contendo 20 mM Tris pH 7.8, 150 mM NaCl, 1 mM EDTA, 1 mM DTT e purificada por exclusão de tamanho molecular utilizando coluna HiLoad 26/100 Superdex 200 (GE Healthcare) pré-equilibrada em tampão de gel filtração. A pureza foi analisada por SDS-PAGE e a concentração determinada usando a absorvância medida em 280 nm e seu coeficiente de extinção teórico de 32890 M⁻¹.cm⁻¹. As frações contendo a proteína pura foram aliqüotadas, congeladas em nitrogênio líquido e armazenadas em -80°C. Para cristalização, a proteína foi concentrada à 14 mg/mL utilizando concentrador de corte 10 kDa (Vivaspin, Sartorius).

2.3.2 Purificação da enzima M^{pro}

A enzima M^{pro} foi purificada de maneira similar, com uma etapa adicional de cromatografia de troca aniônica. Após a cromatografia de exclusão molecular, o tampão foi trocado para 20 mM Tris pH 8.0, 1 mM DTT e a proteína foi injetada em coluna Mono-Q 5/50 GL (GE Healthcare), na qual a proteína de interesse foi eluída com gradiente de tampão B (20 mM Tris pH 8.0, 1 M NaCl, 1 mM DTT). Ao final, o tampão foi novamente trocado para 20 mM Tris pH 7.8, 150 mM NaCl, 1 mM EDTA, 1 mM DTT. As frações contendo proteína pura foram concentradas e quantificadas de maneira similar à realizada para a enzima IMT-M^{pro}.

2.3.3 Purificação da enzima C145S M^{pro}

A enzima C145S M^{pro} foi purificada pelo método de cromatografia de afinidade por íon metálico (IMAC), utilizando a coluna HisTrap FF 5 mL (GE Healthcare) equilibrada em tampão A (20 mM Tris pH 7.8, 150 mM NaCl, 25 mM imidazol). Após lavagem com 10 CV de tampão A, a proteína foi eluída com tampão B suplementado com 250 mM de imidazol. O tampão foi trocado para tampão A utilizando a coluna HiTrap desalting 5 mL (GE Healthcare). Para remoção da cauda de histidinas, foram adicionados 2 mg de TEV protease e 4 mM DTT, e a amostra foi incubada a 4°C por 16 h. Em seguida, a proteína não clivada, a enzima TEV protease e a 6xHIS-tag foram removidas com uma segunda etapa de cromatografia de afinidade em

tampão A. Ao final, a proteína foi purificada por cromatografia de exclusão molecular em coluna HiLoad 16/60 Superdex 75 (GE Healthcare) pré-equilibrada em tampão de gel filtração. As frações contendo proteína pura foram concentradas e quantificadas de maneira similar à realizada para as demais construções.

2.4 Cristalização

Para cristalização, inicialmente utilizamos o método de matriz esparsa, utilizando kits comerciais e robô de cristalização Crystal Gryphon (ARI®), na temperatura de 20°C. Para a enzima M^{pro}, os cristais apareceram após 1 dia na condição contendo 0.1 M Bis-Tris pH 6.5, 25% PEG 3350 e foram crio-protegidos utilizando a solução do poço adicionada de 30% PEG 400. Os cristais da enzima IMT-M^{pro}, apareceram após 1-3 dias em diversas condições. Após otimização, os cristais cresceram em 0.1 M MES pH 6.5, 10% 2-propanol, 20% PEG Smear Low (BCS Screen, Molecular Dimensions) e foram utilizados como *seeds* para o crescimento dos cristais otimizados, na condição 0.1 M MES pH 6.7, 5% DMSO, 8% PEG 4K.^{117,127} Para obtenção dos cristais da M^{pro} no grupo espacial ortorrômbico, foram utilizados *seeds* da IMT-M^{pro} na condição 0.1 M MES pH 6.7, 5% DMSO, 8% PEG 4K,¹¹⁷ na temperatura de 20°C. Os cristais da enzima M^{pro} C145S tetramérica, foram obtidos após 20 dias em 0.1 M fosfato/citrato pH 5.5, 20% PEG Smear High (BCS Screen, Molecular Dimensions) à 20°C. Já os cristais da enzima M^{pro} C145S monomérica, foram obtidos após 20 dias em 0.1 M Bis Tris pH 6.5, 25% PEG 3350 a 20°C.

2.5 Coleta de dados de difração e processamento

Durante a fase de comissionamento (julho a outubro de 2020) da linha de luz MANACA (Macromolecular Micro and NAno CrystAllography) do sincrotron brasileiro SIRIUS (Campinas – SP), a linha adotou um plano de comissionamento emergencial para possibilitar sua utilização para realização da coleta de dados de amostras relacionadas ao SARS-CoV-2. Este projeto foi a primeiro experimento realizado por usuários externos na linha MANACA, a primeira linha operacional do SIRIUS.¹²⁸ Embora ainda fora da sua total capacidade durante a realização deste projeto, a linha MANACA foi fundamental para a realização deste trabalho.

O conjunto de dados coletados da IMT-M^{pro} foi integrado e escalonado com o programa XDS,⁶³ via AutoPROC,¹²⁹ onde os limites de resolução foram estabelecidos pelo método de Karplus.⁶⁴⁻⁶⁵ O programa AIMLESS do CCP4 suite⁶⁶ foi utilizado para conversão das

intensidades em amplitudes e cálculo das estatísticas. A qualidade dos dados foi analisada com o programa phenix.Xtrriage.⁶⁷

Para a resolução da estrutura da IMT-M^{pro} foi utilizada a substituição molecular para cálculo das fases, utilizando como modelo a M^{pro} de SARS-CoV-2 (PDB ID 5RGQ). Para isso, utilizamos o programa PHASER,⁶⁸ procurando por duas cópias do modelo na unidade assimétrica, baseado no coeficiente de Matthews obtido utilizando Matthews_coef do CCP4 suite.⁶⁶ A construção do modelo inicial da estrutura foi conduzida pelo programa Autobuild⁶⁹. O refinamento foi realizado utilizando o BUSTER¹³⁰ e o Coot.⁷²⁻⁷³ A validação da estrutura foi realizada utilizando o Molprobit.⁷⁴

Para a resolução das estruturas da M^{pro} e C145S M^{pro} (monômero e tetrâmero) os conjuntos de dados foram integrados e escalonados e mergeados da mesma forma realizada para IMT-M^{pro}. As estruturas foram resolvidas pelo método de substituição molecular utilizando o programa PHASER⁶⁸ e utilizando como modelos, respectivamente as estruturas dos PDBs 5RGQ, 7KFI e 5R8T. M^{pro} e C145S M^{pro} foram refinadas utilizando Coot⁷²⁻⁷³ e phenix.refine.⁷⁰ As figuras foram feitas com uso do programa Pymol (Schrödinger, LLC).

2.6 Triagem de Fragmentos contra forma imatura da M^{pro} (IMT-M^{pro})

Para a triagem de fragmentos, utilizamos as bibliotecas de fragmentos FragMAXlib¹³¹ e F2Xentry,¹³² totalizando 192 fragmentos testados. Os fragmentos de cada uma das placas foram diluídos em 1 µL de 0.1 M MES pH 6.7, 5% DMSO, 8% PEG 4K, 30% PEG 400, os cristais foram da proteína IMT-M^{pro} obtidos da condição 0.1 M MES pH 6.7, 5% DMSO, 8% PEG 4K e foram manualmente transferidos para cada uma das gotas. O *soaking* foi realizado por 4 h em temperatura ambiente e em seguida os cristais foram manualmente coletados e congelados em nitrogênio líquido.

A coleta de dados foi realizada como descrita previamente em 2.5. 166 destes cristais foram testados, totalizando 77 bons conjuntos (dados para os quais a substituição molecular foi realizada com sucesso e resultou em mapas com resolução <3Å). Para a realização da análise dos dados, uma versão simplificada do *software* FragMAXapp foi instalada localmente em nosso computador.¹³³ Utilizando o FragMAXapp, as restrições para cada composto foram geradas pelo phenix.eLBOW¹³⁴ e os conjuntos foram processados pelo autoPROC/STARANISO ou DIALS via XIA2.^{129,135-136} A substituição molecular e refinamento inicial foi realizada pelo DIMPLE,⁶⁶ tendo PDBid 7KFI como modelo. Para destacar as

densidades eletrônicas referentes aos fragmentos, a análise dos mapas foi conduzida no software PanDDa.¹³⁷ Os modelos das estruturas foram refinados como descrito em 2.5.

2.7 Caracterização das enzimas em solução

As construções da enzima M^{pro} foram caracterizadas pela obtenção de seus parâmetros cinéticos, de acordo com o ensaio de atividade descrito a seguir, pela obtenção de suas respectivas temperaturas de *melting* (T_m), pela determinação de seus estados oligoméricos e por espectrometria de massas nativa (Native-MS).

2.7.1 Ensaio de atividade

Para monitoramento da atividade proteolítica da enzima, foi utilizado o substrato peptídico comercial Dabcyl-KTSAVLQ↓SGFRKM-E(Edans)-NH₂ (Biomatik). Este ensaio utiliza o método FRET (*Fluorescence resonance energy transfer*) no qual é utilizado um peptídeo contendo dois fluoróforos, um deles responsável pela fluorescência, ou doador (EDANS, no caso deste ensaio) e outro responsável por absorver a fluorescência emitida pelo primeiro, em mesmo comprimento de onda, também chamado de acceptor (Dabcyl). Após a clivagem do peptídeo pela ação da protease, dois peptídeos menores são liberados contendo cada um dos fluoróforos, aumentando a distância relativa entre eles. Desta forma, a fluorescência emitida pelo EDANS, não é mais absorvida e pode ser monitorada em tempo real com utilização de um espectrofotômetro nos comprimentos de onda indicados (Figura 36).

Para realização dos ensaios de atividade a enzima foi diluída para uma concentração final de 0.14 μ M, 0.3 μ M e 0.3 μ M, para M^{pro}, IMT M^{pro} e C145S M^{pro} respectivamente, em tampão de reação (20 mM Tris, 1 mM DTT, 1 mM EDTA, pH 7.3) e incubada por 10 minutos a 37°C. A reação foi iniciada com adição de 10 μ M de substrato. A reação foi mantida a temperatura constante a 37°C, com a fluorescência sendo monitorada a cada 30 segundos durante 60 minutos utilizando o equipamento SpectraMax Gemini EM Microplate Reader (Molecular Devices). As reações foram feitas em placas Corning® brancas de 96 poços. Os comprimentos de onda utilizados para excitação e emissão foram respectivamente 360 nm e 460 nm.

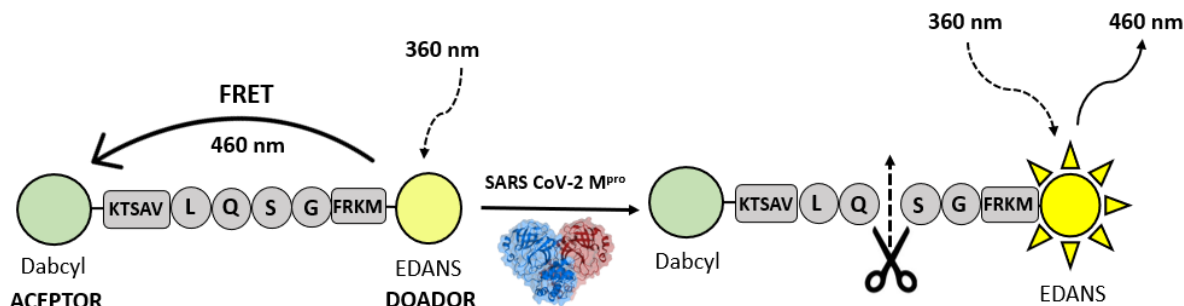


Figura 36 - Representação esquemática do ensaio enzimático da proteína SARS-CoV-2 M^{pro} baseado em FRET. O substrato DabcyL-KTSAVLQ↓SGFRKM-E(Edans)-NH₂ é clivado pela enzima liberando dois peptídeos menores, cuja fluorescência pode ser monitorada nos comprimentos de onda apresentados. Fonte: Elaborada pela autora.

Para a avaliação da atividade de auto-clivagem da enzima C145S M^{pro}, a proteína fusionada a cauda 6-His foi trocada para tampão 20 mM HEPES pH 7.3, 100 mM NaCl, 1 mM DTT. Duas reações simultâneas foram preparadas para comparação, uma contendo 10 µM C145S M^{pro} e outra contendo 10 µM C145S M^{pro} e 5 nM M^{pro}. Alíquotas de cada reação foram coletadas durante um período de 44h, mantendo as amostras em temperatura ambiente (25°C). As amostras foram analisadas por SDS-PAGE em gel 12.5% acrilamida.

2.7.2 Fluorimetria diferencial de varredura (DSF)

Os valores de T_m foram obtidos do experimento de Fluorimetria Diferencial de Varredura (*Differential Scanning Fluorimetry* - DSF) que foi realizado seguindo o mesmo procedimento utilizado para a NS2B-NS3 protease de YFV, conforme descrito no item 2.6 do capítulo 1.

2.7.3 Cromatografia por exclusão de tamanho acoplada ao espalhamento de luz em múltiplos ângulos (SEC-MALS)

O estado oligomérico em solução das amostras foi determinado por cromatografia por exclusão de tamanho acoplada ao espalhamento de luz em múltiplos ângulos (SEC-MALS) em tampão 20 mM Tris-HCl pH 7.8 e 100 mM NaCl. Para isto, 50 µL de cada amostra na concentração de 50 µM foram injetados em um sistema HPLC Waters 600 HPLC system (Waters), acoplado a um detector de UV, a um detector que mede o espalhamento de luz em múltiplos ângulos miniDAWN TREOS (Wyatt Technology), uma coluna Superdex 200 Increase 10/300 GL (GE Healthcare), e um detector de índice de refração Optilab T-rEX (Wyatt Technology). Os detectores foram normalizados com utilização de albumina de soro bovina

(BSA) (Sigma-Aldrich). Os dados foram coletados e analisados com o software ASTRA 7, providenciado pela Wyatt. O fluxo utilizado foi de 0.5 mL/min.

Para monitoramento da formação de dímero, os estados oligoméricos em solução da M^{pro} C145S, correspondente ao pico monomérico foram determinados de maneira similar por SEC-MALS, utilizando a coluna Superdex 75 Increase 10/300 GL (GE Healthcare). Duas reações foram conduzidas simultaneamente uma delas contendo 60 μ M M^{pro} C145S monômero e outra contendo 60 μ M M^{pro} C145S e 10 nM M^{pro} (proporção 6000:1). Ambas foram mantidas em temperatura ambiente e amostras de 50 μ L foram injetadas a cada 24h por um período de 72h.

Para monitoramento da mudança de estados oligoméricos na presença de inibidores, amostras contendo 100 μ M de C145S M^{pro} monômero ou tetrâmero foram incubadas com 1% DMSO (controle) ou ligante por 0h, 24h ou 48h. O inibidor não covalente MAT-POS-e194df51-1 foi gentilmente cedido pelo *COVID Moonshot Consortium*, foi testado na proporção 4:1 e 0.4:1 (ligante:proteína). O inibidor covalente nirmatrelvir foi gentilmente cedido pelo Prof. Dr. Carlos A. Montanari (Instituto de Química de São Carlos, Universidade de São Paulo) e foi testado na proporção 4:1. Devido à dificuldade em distinguir entre trimeros/tetrâmeros, ambos foram considerados como tetrâmeros nos cálculos de porcentagens.

2.7.4 Espectrometria de massas nativa (Native-MS)

Os experimentos de espectrometria de massas foram realizados em colaboração pelo Dr. Rod Chalk e Dra. Lizbé Koekemoer (Centre for Medicines Discovery, Oxford University, UK) utilizando um Agilent 6530 QTOF como descrito em Chalk, R., 2020.¹³⁸ Os resultados foram visualizados utilizando o software QualBrowser (Thermo Fischer Scientific) e OriginPro 2021. Os métodos empregados podem ser encontrados de maneira detalhada em Noske et al, 2023.¹³⁹

2.8 Obtenção da estrutura da enzima C145S M^{pro} por crio-microscopia eletrônica (cryoEM)

Os experimentos de crio-microscopia foram realizados em colaboração pela Dra. Yun Song (eBIC: Electron Bio-Imaging Centre - Diamond Light Source, propostas BI27083 e NT29349). O processamento dos dados foi realizado em colaboração, em conjunto com o pós-doutorando do grupo Dr. Andre S. Godoy.

2.8.1 Preparação das grades e coleta de dados

Para a preparação das grades, foi utilizada a mesma amostra utilizada para cristalização, correspondente a M^{pro} C145S forma tetramérica, purificada da forma descrita anteriormente. A proteína foi descongelada no gelo, centrifugada e diluída em tampão de gel filtração para concentração final de 0.25 mg/mL. 3 μ L de amostra foram aplicados em grade Quantifoil 300 mesh Cu R2/2 previamente *glow discharged*. A grade foi então vitrificada utilizando etano líquido com auxílio de um Vitrobot (FEI), após realização de *blot* por 2.5s, com força 1, humidade 100% a 4°C.

Após a vitrificação da SARS-CoV-2 M^{pro} C145S, a coleta de dados foi realizada em microscópio Titan Krios operando a 300 kV, utilizando detector K3 em modo *counting super resolution*, com filtro de 20 eV, em uma magnificação de 130 kx (correspondente a pixel de tamanho 0.653Å). As grades foram usadas para adquirir 13.770 filmes com faixa de desfocagem entre -0.9 a -2.5 μ m. A aquisição dos dados foi realizada através do software EPU 2.12.

2.8.2 Processamento dos dados

Os filmes foram alinhados usando MotionCor2,¹⁴⁰⁻¹⁴¹ e sua função de transferência de contraste (CTF) foi estimada usando CTFFIND v4.12.¹⁴² Em seguida, 1.956.496 partículas foram encontradas utilizando software SPHIRE-crYOLO¹⁴³ e extraídas com *box* de 72 pixels (reduzidas por fator de 3x) foram submetidas a 3 rodadas de classificação 2D com N=100 para filtragem. As classes selecionadas resultantes mostraram clara semelhança com a forma dimérica da M^{pro} madura. Após a reconstrução *ab-initio* sem simetria, as 569.725 partículas re-extraídas (*box* de 226 pixels) foram submetidas a uma rodada de classificação 3D sem simetria, resultando em uma classe distinguível contendo 39% das partículas. Essa classe selecionada contendo 227.534 partículas foi usada para refinamento sem simetria usando RELION v3.130 Refine3D,¹⁴⁴ resultando em um mapa com resolução FSC "padrão-ouro" (FSC = 0.143) de 4.1 Å.¹⁴⁵

Após o refinamento e polimento do CTF, essas partículas e o volume 3D de alta resolução foram importados para o cryoSPARC v3.2.031¹⁴⁶ para as rodadas finais de refinamento. O mapa final foi obtido após uma rodada de refinamento homogêneo (com um passo final extra e filtro inicial de resolução de 7Å), seguida de refinamento não uniforme¹⁴⁷ (com um passo final extra, filtro inicial de resolução de 8Å e otimização de defocus e CTF por partícula). Por final foi realizado um refinamento local. Os refinamentos foram conduzidos da

mesma forma no cryoSPARC com simetria aplicada C1 e C2 e resultaram em mapas quase idênticos, com a respectiva resolução gsFSC de 3.8 Å e 3.5 Å, portanto o mapa C2 foi utilizado para posteriores análises. A resolução do mapa foi calculada utilizando cryoSPARC Local Resolution e 3DFSC Processing server.^{146,148}

2.8.3 Modelagem e análises estruturais

O mapa *sharpened* do cryoSPARC foi usado para posicionamento inicial manual do modelo cristalográfico da M^{pro} dimérica (PDBid 7N6N) utilizando o ChimeraX v1.2.5.¹⁴⁹ Nas etapas posteriores o mapa *sharpened* gerado com phenix_auto_sharpen,¹⁵⁰ que mostrou conectividade aprimorada quando comparado com ao obtido no cryoSPARC, foi utilizado para a modelagem. O modelo final foi obtido com ciclos iterativos de reconstrução manual utilizando Coot⁷²⁻⁷³ e refinamento no espaço real usando Phenix.⁷¹ A validação foi realizada pelo Molprobability.⁷⁴ As figuras foram geradas utilizando PyMOL e ChimeraX v1.2.5.

3 RESULTADOS E DISCUSSÃO

3.1 Clonagem, expressão e purificação das enzimas

Ambas as construções da enzima M^{pro} de SARS-CoV-2 foram clonadas, expressas e purificadas com sucesso. Inicialmente, as construções IMT-M^{pro}, M^{pro} e C145S M^{pro} foram clonadas conforme os métodos descritos no vetor pET_M11/LIC. Para a IMT-M^{pro} e M^{pro}, a clonagem foi confirmada por PCR de colônia, resultando nas bandas com tamanho condizente com o esperado (Figura 37), seguida de sequenciamento. Já para o mutante inativo C145S M^{pro}, a presença da mutação foi confirmada por sequenciamento.

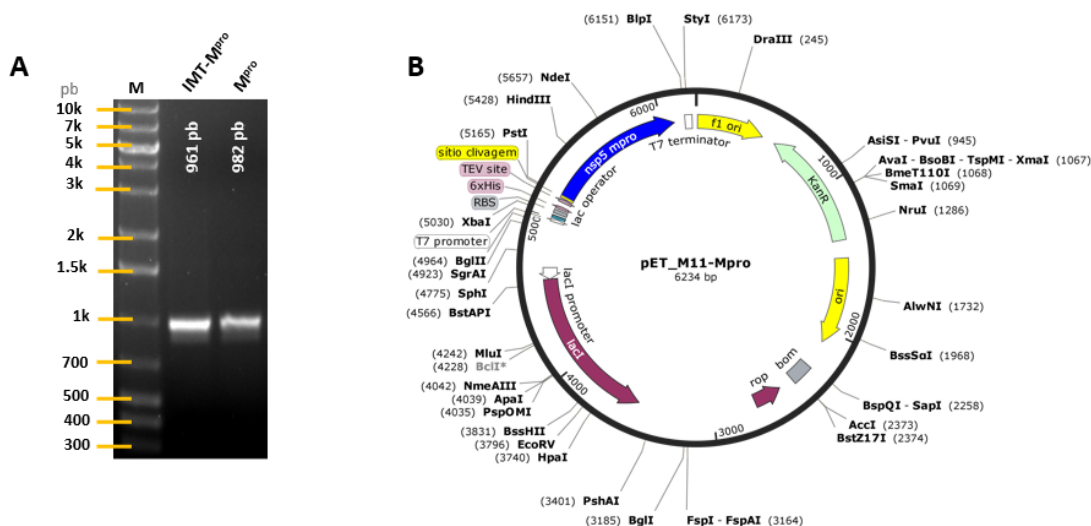


Figura 37 - Clonagem das diferentes construções da M^{pro} de SARS-CoV-2. A) Gel de agarose 1% contendo os resultados da PCR de colônia obtidos para as construções IMT-M^{pro} (961 pb) e M^{pro} (982 pb) respectivamente. A amostra indicada por M contém o marcador de peso molecular, com os tamanhos em pb indicados na figura. B) Mapa do plasmídeo pET_M11/LIC obtido na clonagem para as construções de M^{pro}.

Fonte: Elaborado pela autora.

Após a confirmação da clonagem, ambos os plasmídeos foram isolados utilizados para transformar a linhagem de *E. coli* BL21. As proteínas foram expressas em larga escala utilizando meio auto indutor ZYM 5052, de acordo com as condições otimizadas.

A estratégia comumente utilizada para purificação da M^{pro} nativa,^{106,119,152} consiste em manter a sequência de auto clivagem na região N-terminal, e adicionar um sítio de clivagem da enzima HRV-3C protease, seguida de uma cauda 6xHis na porção C-terminal. Neste trabalho, utilizamos um protocolo simplificado, baseado em precipitação por sulfato de amônio, seguida

de troca de iônica, o qual resulta na obtenção de ambas as enzimas com elevada pureza, levando cerca de 8h.¹²⁷ Para a IMT-M^{pro}, o sítio de TEV protease (Q↓GAM) permite que a enzima realize auto clivagem, assim removendo a cauda de 6xHIS. A pequena porção da proteína que não foi clivada, foi removida da fração solúvel utilizando cromatografia de afinidade por íon imobilizado (IMAC) (Figura 38A). A proteína clivada foi isolada da fração solúvel por precipitação utilizando sulfato de amônio, e em seguida separada por cromatografia de exclusão molecular (Figura 38A). Para a M^{pro} nativa, que contém os resíduos da região de auto clivagem da enzima na poli proteína (nsp4-nsp5:SAVLQ↓SGFRKM) como não foi atingido a pureza desejada após precipitação por sulfato de amônio e cromatografia de exclusão molecular, foi necessária uma etapa adicional de purificação, utilizando troca iônica (Figura 38B).¹²⁷

Para a enzima C145S M^{pro}, que não sofre auto clivagem durante a expressão, utilizamos cromatografia de afinidade por íon imobilizado (IMAC), seguida de cromatografia de exclusão de molecular, com protocolo similar ao descrito no capítulo 1 para a proteína NS2B-NS3 protease de YFV. Ao final, ambas as pudemos concluir que ambas as proteínas foram obtidas com elevada pureza e rendimento (Figura 38).¹²⁷ A pureza final das amostras foi analisada por SDS-PAGE 12.5% e juntamente com os cromatogramas das etapas finais de purificação, podem ser vistos na Figura 38. O rendimento aproximado foi de 80 mg/L de cultura, 2.5 mg/L de cultura e 15 mg/L de cultura para as construções IMT M^{pro}, M^{pro} e C145S M^{pro}, respectivamente.

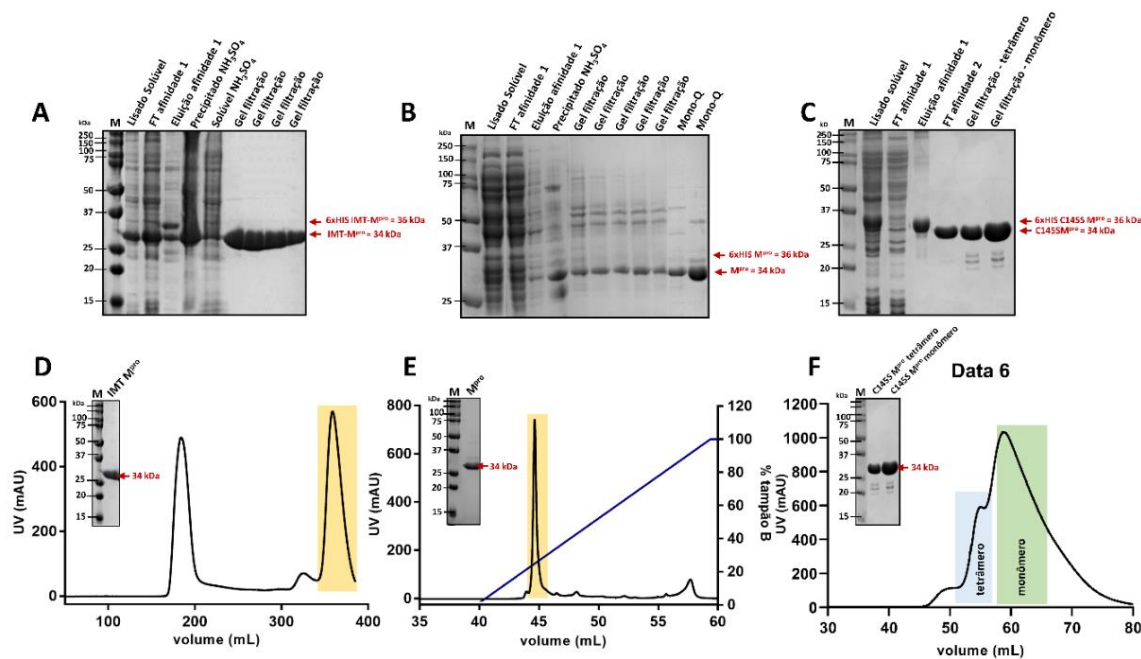


Figura 38 - Purificação das 3 construções da enzima M^{pro}, IMT-M^{pro}, M^{pro} e C145S M^{pro}, respectivamente. Para ambas as enzimas estão apresentados os cromatogramas referentes a última etapa de purificação e os respectivos SDS-PAGE (12.5%).

Fonte: Adaptada de NOSKE *et al.*¹²⁷

Tabela 8 – Volumes de retenção de diferentes construções da M^{pro}, massas moleculares calculadas e estados oligoméricos. Ambos os volumes de retenção foram obtidos da filtração em coluna Superdex 75 16/60 e as massas calculadas de acordo com a curva de calibração apresentada no capítulo 1.

Proteína	Volume de retenção (mL)	Massa calculada (Da)	Estado oligomérico
IMT-M ^{pro}	70.3	37609,7	monômero
C145S M ^{pro} pico 1	56.42	141123,7148	tetrâmero
C145S M ^{pro} pico 2	68.3	42469,77	monômero

Fonte: Elaborada pela autora.

As enzimas IMT-M^{pro} e M^{pro} foram obtidas em forma de um único pico na etapa final de purificação, enquanto a enzima M^{pro} C145S foi obtida na cromatografia de exclusão molecular em 2 diferentes picos, com volumes de retenção de 56.42 e 68.3 mL, correspondentes a proteína em sua forma tetramérica, e monomérica, respectivamente (Tabela 8). Detalhes sobre os diferentes estados de oligomerização desta proteína serão discutidos posteriormente.

3.2 Ensaios de atividade e caracterização bioquímica

Ambas as construções M^{pro} e IMT-M^{pro} demonstraram ser ativas e capaz de reconhecer e clivar o substrato fluorogênico no ensaio FRET (Figura 39A-B), com valores de K_m de $16 \pm 2 \mu\text{M}$ e $34 \pm 2 \mu\text{M}$, respectivamente.¹²⁷ A IMT-M^{pro} exibiu apenas 0.6% da eficiência catalítica em comparação com a M^{pro} madura (Tabela 9). Como relatado anteriormente, o N-terminal da enzima é fundamental para a dimerização e quaisquer resíduos adicionais reduziram ou até aboliram a sua atividade.^{126,152} Como esperado, o mutante C145S M^{pro} possui apenas mostrou atividade residual (Figura 39A). Nos ensaios de DSF, todas as construções de M^{pro} exibiram perfis de estabilidade térmica semelhantes, indicando enovelamento similar (Figura 39D e Tabela 9).¹²⁷

Tabela 9- Parâmetros cinéticos calculados e estabilidade térmica para cada uma das M^{pro}s. A eficiência relativa foi calculada com relação a M^{pro} nativa

Proteína	K_M (μM)	V_{max} (RFU/s ⁻¹)	k_{cat} (s ⁻¹)	Eficiência relativa	T _m (°C)
M ^{pro}	16 ± 2	1.05 ± 0.04	28.0 ± 0.1	1	53.30 ± 0.01
IMT-M ^{pro}	34 ± 2	0.104 ± 0.003	0.35 ± 0.07	0.006	53.01 ± 0.08
C145S M ^{pro}	-	-	-	-	51.31 ± 0.02

Fonte: Adaptada de NOSKE *et al.*¹²⁷

A análise em solução utilizando SEC-MALS sugere que a M^{pro} se comporta como um dímero nas condições testadas, conforme esperado (Figura 39C),¹⁵³ enquanto para a $IMT-M^{pro}$, os resíduos adicionais no N-terminal parecem abolir completamente a dimerização (Figura 39C). Para a $C145S M^{pro}$, no entanto, os resíduos adicionais, juntamente com a presença da mutação, permitem a proteína para adotar múltiplos estados conformacionais variando de monômeros a tetrâmeros (Figura 39C). As massas moleculares de cada uma das construções obtidas no SEC-MALS estão sumarizadas na Tabela 10.¹²⁷

Tabela 10 - Massas moleculares para cada pico observado nos gráficos do SEC-MALS de cada uma das construções de M^{pro}

Proteína	Pico	Volume de eluição (mL)	Massa molar (kDa)	Estado oligomérico
M^{pro}	1	15.22	63.76 ± 0.01	dímero
$IMT M^{pro}$	1	16.30	36.50 ± 1.22	monômero
$C145S M^{pro}$ Monômero	1	16.04	39.06 ± 0.01	monômero
	2	14.34	71.90 ± 0.10	dímero
$C145S M^{pro}$ Tetrâmero	1	16.35	35.84 ± 0.29	monômero
	2	14.45	76.36 ± 1.69	dímero
	3	13.90	93.79 ± 0.93	trímero
	4	12.86	129.36 ± 1.95	tetrâmero

Fonte: Adaptada de NOSKE *et al.*¹²⁷

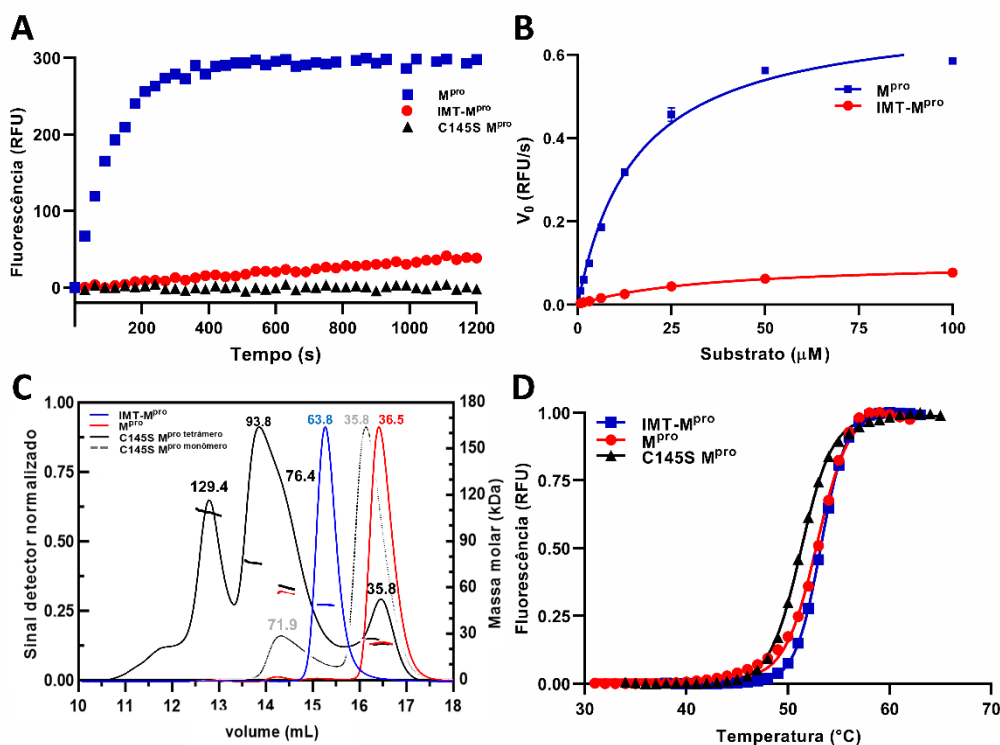


Figura 39 - A) Curvas de fluorescência das diferentes construções de M^{pro} medidas ao longo do tempo. B) Curvas de Michaelis-Menten para cada uma das enzimas mutantes. As barras de erros correspondem ao desvio padrão das triplicatas. C) Perfis de eluição na cromatografia de exclusão molecular para cada uma das enzimas. As curvas correspondem a mudança na intensidade no espalhamento normalizado em 90° (linhas) e massas moleculares calculadas para cada um dos picos (pontos). D) Perfis de desnaturação térmica de cada uma das enzimas.

Fonte: Adaptada de NOSKE *et al.*¹²⁷

3.3 Cristalização, coleta de dados e resolução das estruturas

Para os experimentos de cristalização, inicialmente utilizamos o método de matriz esparsa, utilizando kits comerciais e robô de cristalização Crystal Gryphon (ARI®), na temperatura de 20°C . Para a enzima M^{pro} , os cristais apareceram após 1 dia na condição contendo 0.1 M Bis-Tris pH 6.5, 25% PEG 3350. Os cristais da enzima IMT- M^{pro} , apareceram após 1-3 dias em diversas condições (Figura 40A-B). Após otimização, os cristais cresceram em 0.1 M MES pH 6.5, 10% 2-propanol, 20% PEG Smear Low (BCS Screen, Molecular Dimensions) e foram utilizados como *seeds* para o crescimento dos cristais otimizados, na condição 0.1 M MES pH 6.7, 5% DMSO, 8% PEG 4K (Figura 40C). Para obtenção dos cristais da M^{pro} no grupo espacial ortorrômbico, foram utilizados *seeds* da IMT- M^{pro} na condição 0.1 M MES pH 6.7, 5% DMSO, 8% PEG 4K, na temperatura de 20°C (Figura 40D). Os cristais da enzima M^{pro} C145S tetramérica, foram obtidos após 20 dias em 0.1 M

fosfato/citrato pH 5.5, 20% PEG Smear High (BCS Screen, Molecular Dimensions) à 20°C (Figura 40E).

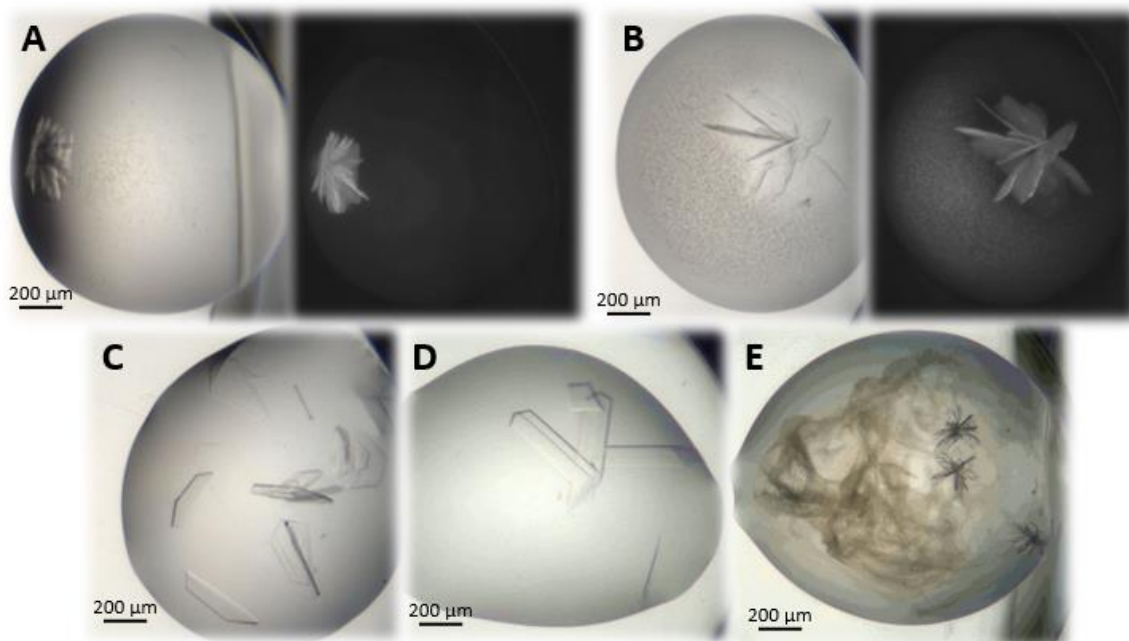


Figura 40 - Cristais da enzima M^{pro} de SARS-CoV-2. A) Cristais da enzima IMT- M^{pro} obtidos na condição E09 do kit MORPHEUS (0.12 M etileno glicóis, 0.1 M Buffer System 3 pH 8.5, 30% PEG mix 1). B) Cristais da enzima IMT- M^{pro} obtidos na condição D02 do kit BCS (10% 2-PropOH, 0.1 M MES pH 6.5, 22.5% PSL). Em A-B, a imagem da direita corresponde a mesma condição sob luz UV. C) Cristais otimizados da IMT- M^{pro} na condição 0.1 M MES pH 6.7, 5% DMSO, 8% PEG 4K (utilizando *seed* obtido da condição B). D) Cristais da M^{pro} (ortorrômbico) obtidos na condição 0.1 M MES pH 6.7, 5% DMSO, 8% PEG 4K (utilizando *seed* da IMT- M^{pro} obtido da condição B). E) Cristais da enzima C145S M^{pro} (tetramero), na condição A08 do kit BCS (0.1 M fosfato/citrato pH 5.5, 20% PEG Smear High)

Fonte: Elaborada pela autora.

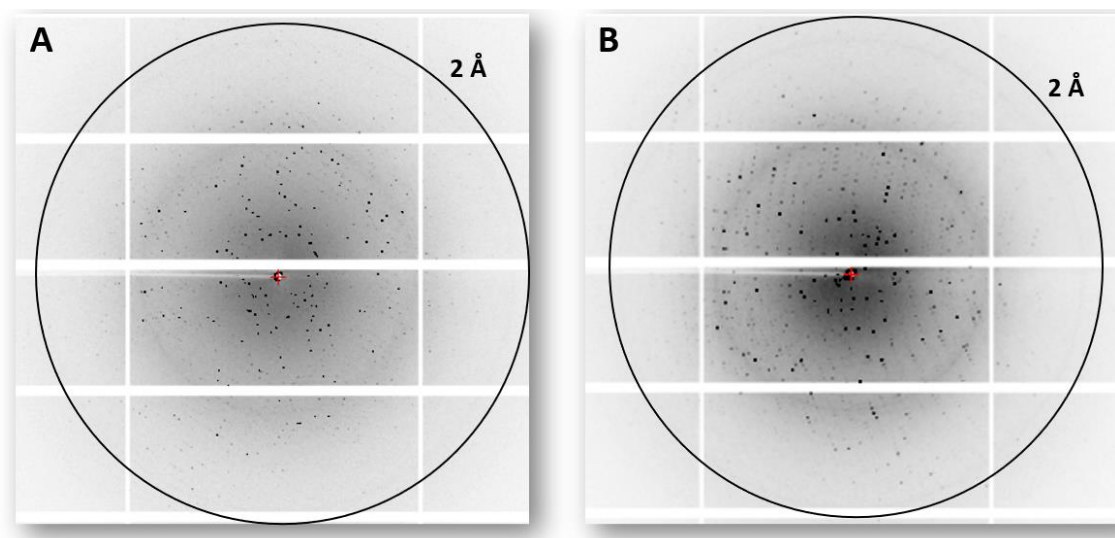


Figura 41 - Padrões de difração de cristais da enzima M^{pro} de SARS-CoV-2, ambos coletados na linha MANACA, no SIRIUS. A) Cristais da enzima IMT- M^{pro} obtidos na condição 0.1 M MES pH 6.7, 5% DMSO, 8% PEG 4K. B) Cristais da enzima M^{pro} no grupo espacial ortorrômbico.

Fonte: Elaborada pela autora.

Os cristais da enzima M^{pro} foram crioprotetidos com a solução do poço contendo 30% PEG 400, e a coleta foi realizada no Laboratório Nacional de Luz Síncrotron (LNLS-SIRIUS), Campinas – SP, na linha de luz MANACÁ. Todos os cristais apresentaram alta qualidade no padrão de difração, difratando em alta resolução (1.5 Å para M^{pro} , IMT- M^{pro} e C145S M^{pro} monômero, e 2.8 Å para C145S M^{pro} tetrâmero) (Figura 41). Detalhes sobre a coleta de dados e estatísticas de refinamento estão apresentados na Tabela 11.

Tabela 11 - Coleta de dados e estatísticas do refinamento referentes a enzima M^{pro} de SARS-CoV-2. Os dados entre parênteses correspondem à última camada de resolução.

	M ^{pro}	M ^{pro} ortorrômbico	IMT M ^{pro}	C145S M ^{pro} tetramero	C145S M ^{pro} monômero
<i>Coleta de dados</i>					
Grupo espacial	C2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁	C2 ₁
Dimensões da cela unitária					
<i>a, b, c</i> (Å)	113.0, 52.8, 44.7	67.2, 98.1, 102.2	67.6, 101.8, 102.2	58.6, 79.3, 62.8	114.8, 53.4, 44.7
<i>α, β, γ</i> (°)	90, 102.8, 90	90, 90, 90	90, 90, 90	90, 106.5, 90	90, 101.8, 90
Resolução (Å)	55.1-1.46 (1.51-1.46)	67.2-1.86 (1.91-1.86)	29.1-1.6 (1.65-1.6)	60.2-2.8 (2.9-2.8)	48.24-1.76 (1.8-1.76)
Reflexões únicas	42249 (5935)	57312 (4147)	93647 (4597)	13623 (1334)	25331 (2588)
Multiplicidade	3.4 (3.3)	6.6 (6.7)	18.4 (15.8)	3.7 (3.3)	3.3 (3.4)
Completeza (%)	94.9 (91.2)	99.6 (99.2)	99.9 (99.7)	99.36 (97.9)	95.76 (99.58)
<i>I/σI</i>	12.8 (2.4)	7.2 (0.7)	8.8 (0.7)	3.3 (1.1)	12.8 (2.7)
<i>R</i> _{p.i.m.} (%)	0.031 (0.298)	0.074 (1.28)	0.023 (0.90)	0.15 (0.83)	0.034 (0.32)
CC _{1/2}	1.0 (0.46)	0.99 (0.31)	1.0 (0.49)	0.96 (0.38)	0.99 (0.84)
<i>Refinamento</i>					
<i>R</i> _{work} / <i>R</i> _{free}	0.16/0.18	0.19/0.22	0.21/0.23	0.20/0.25	0.15/0.19
Número de átomos					
água	328	359	454	-	250
resíduos	306	604	612	617	306
RMS(ligações) (Å)	0.008	0.009	0.014	0.007	0.02
RMS(ângulos) (°)	1.17	1.12	1.68	1.11	1.6
Ramachandran favoráveis (%)	98.68	98.33	96.84	94.9	97.37
Ramachandran não permitidos (%)	0.33	0.17	0.17	0.65	0.33
Clashscore	4.6	9.46	1.59	10.73	2.33
<i>B</i> -fatores médios (Å ²)					
Macromoléculas	24.7	36.4	37.28	48.14	29.57
Solvente	34.9	42.15	50.06	-	39.26
Código PDB	7KPH	7MBG	7KFI	7N6N	7N5Z

Fonte: NOSKE *et al.* ¹²⁷

3.4 Estrutura cristalográfica da enzima IMT-M^{pro}

A estrutura cristalina da IMT-M^{pro} foi determinada usando 3 conjuntos de dados *mergeados*, resultando num mapa final à 1.6 Å de resolução, com 99% de completeza, no grupo espacial P2₁2₁2₁, com duas moléculas na unidade assimétrica, empacotadas de maneira similar à unidade biológica conhecida para a M^{pro}. A estrutura foi refinada resultando em um *R*_{work}/*R*_{free} final de 0.20/0.22, com 97% dos resíduos em regiões favoráveis no gráfico de Ramachandran (Tabela 11). Conforme previamente descrito para a enzima M^{pro}, a estrutura foi obtida em forma dimérica, composta por dois protômeros formados por três domínios (DI, DII e DIII), com sua região catalítica localizada entre os barris β dos domínios DI e DII (Figura 34 e Figura 42).¹²⁷

Diversas mudanças estruturais estão presentes na estrutura da forma imatura da enzima M^{pro} quando comparada a sua forma nativa, destacando as hélices do domínio DIII e o bolsão de ligação ao substrato (Figura 43A-B). Embora a unidade assimétrica presente nos cristais da IMT M^{pro} se assemelhe a conformação do dímero biológico da proteína nativa, análises feitas no PISA¹⁵⁴ indicam que o empacotamento do dímero é instável em solução, com uma área de interface de 1256 Å² (versus 1557 Å² da M^{pro}), ΔG de energia livre de 13.4 kcal/mol (versus 14.9 kcal/mol da M^{pro}) para 26 ligações de hidrogênio (versus 33 da M^{pro}) e 5 pontes salinas (versus 10 da M^{pro}).¹²⁷ O deslocamento da Ser1 é responsável pela quebra de cinco dessas ligações de hidrogênio, mantidas pelos resíduos Phe140, Glu166 e Gly170. As interações entre Gly2 e Ser139, Ala7 e Val124, Ser139 e Gln299 e Phe305 com Pro122 também não estão presentes na IMT M^{pro} . Os domínios DI e DII são menos afetados pela inserção no N-terminal (RMSD de 0.34 Å vs M^{pro} para C α de 1-184), já o domínio DIII parece adotar uma conformação mais aberta em relação a M^{pro} (RMSD de 1.33 Å para C α de 201-301), com a interface dos resíduos Ala285 a uma distância de 10 Å na IMT M^{pro} (versus 5.5 Å na M^{pro}) (Figura 42 e Figura 43A-B). Essa conformação é mais acentuada na cadeia A, onde a densidade eletrônica da inserção N-terminal é claramente visível no mapa obtido após refinamento.¹²⁷

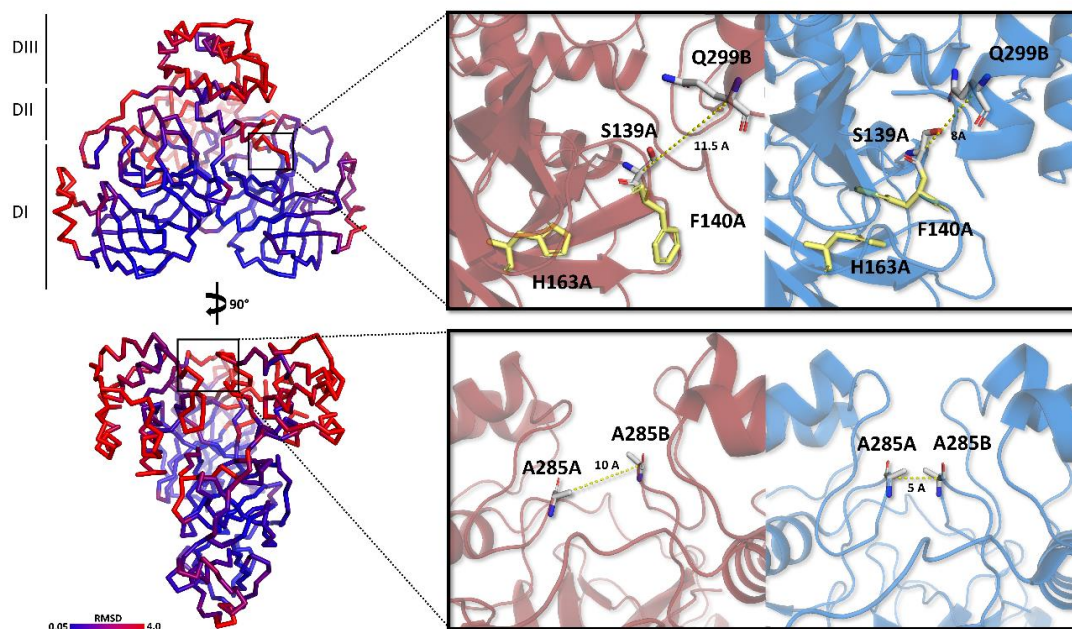


Figura 42 - Mudanças estruturais observadas na forma imatura da enzima M^{pro} , destacando a região do bolsão de ligações ao substrato e a distância entre os dois protômeros. A esquerda, a estrutura da IMT- M^{pro} (PDBid 7KFI) está representada como *ribbon* e colorida de acordo com o RMSD, comparada a M^{pro} nativa (PDBid 7MBG). Nos quadros a direita, as estruturas da M^{pro} imatura e nativa estão representadas como *cartoon*, e coloridas respectivamente em vermelho e azul. Resíduos impactados pelas mudanças estruturais estão representados como *sticks*, e as distâncias estão indicadas pelos traços em amarelo.

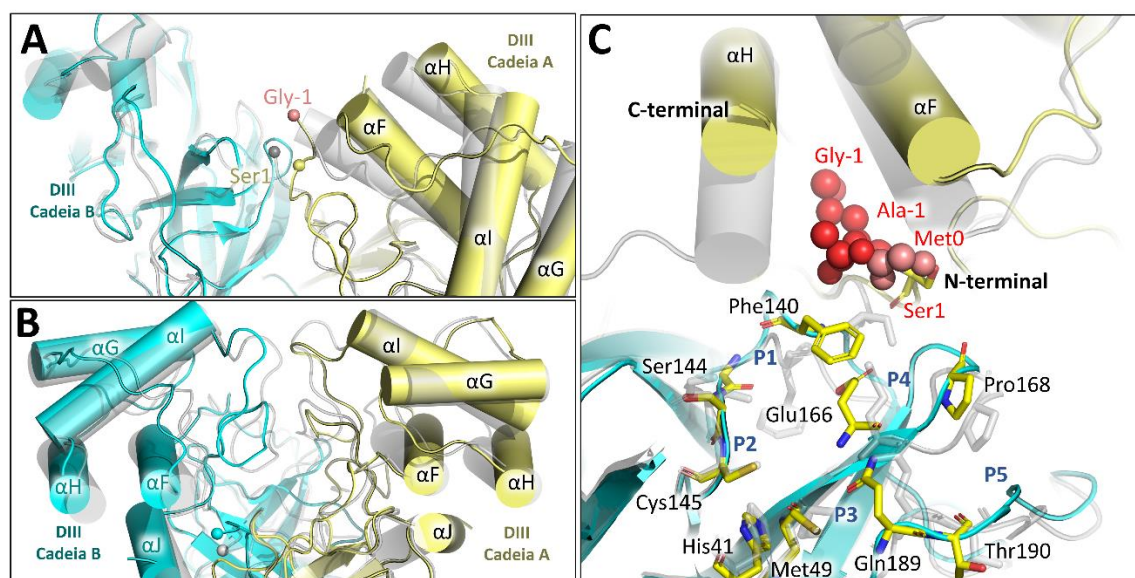


Figura 43 - A) Visão geral do domínio DIII da IMT-M^{pro} (cadeia A em amarelo e cadeia B em ciano) superposta com a M^{pro} nativa (cinza). Os resíduos adicionais da região N-terminal estão representados como esferas e coloridos em vermelho. B) Visão rotacionada do domínio DIII da IMT-M^{pro}. C) Sítio ativo e bolsão de ligação ao substrato da IMT-M^{pro} cadeia B (*cartoon* ciano) superposta com a M^{pro}. Os resíduos catalíticos e importantes na ligação ao substrato estão coloridos em amarelo e representados como *sticks*. M^{pro} nativa e seus respectivos resíduos estão coloridos em cinza.

Fonte: NOSKE *et al.*¹²⁷

3.5 Triagem de Fragmentos contra a forma imatura da M^{pro} (IMT-M^{pro})

Recentemente, uma biblioteca de fragmentos de mais de 1.250 fragmentos foi triada contra os cristais da enzima SARS-CoV-2 M^{pro}, resultando na identificação de 74 *hits*, incluindo 23 *hits* não covalentes, 48 fragmentos covalentemente ligados no sítio ativo, e 3 *hits* próximos a interface de dimerização.¹¹⁷ Aqui, aplicamos a mesma técnica para triar novas cavidades na forma imatura da M^{pro}. Apesar da diferença de escala de nosso experimento (200 fragmentos testados), fomos capazes de identificar cinco diferentes sítios nesta forma da proteína (Figura 44).¹²⁷ O sítio #1 corresponde ao sítio ativo da cadeia A, em que o fragmento f2xe03 foi identificado interagindo com Glu166 N e Cys145 S. Uma cavidade única, denominada como sítio #3 foi identificada em nossos experimentos, contendo o fragmento f2xg02 o qual se encontra ligado pelo oxigênio da cadeia principal da Arg4. Essa cavidade fica entre a interface das cadeias A e B, e não está presente nos cristais de M^{pro} pertencentes ao grupo espacial canônico (monoclínico), que adota uma conformação mais fechada. Este novo local e fragmento podem servir como guia para o desenvolvimento de novos inibidores direcionados ao processo de dimerização de M^{pro}, um modo de ação até então apenas teorizado.¹²⁷ A Figura 44 resume

os fragmentos encontrados, e a Tabela 12 sumariza os dados referentes a coleta de dados e estatísticas do refinamento referentes a enzima IMT-M^{pro} com os diferentes fragmentos.

Tabela 12 - Coleta de dados e estatísticas do refinamento referentes a enzima IMT-M^{pro} com diferentes fragmentos. Os dados entre parênteses correspondem à última camada de resolução.

	<i>f2xe03</i>	<i>f2xe06*</i>	<i>f2xg05</i>	<i>fmaxe01</i>	<i>fmaxe09</i>
Resolução (Å)	51.3- 2.8(2.9 - 2.8)	56.1 - 2.2 (2.3 - 2.2)	37.6 - 2.6 (2.7 - 2.6)	72.09 - 2.09 (2.1 - 2.09)	72.02 - 2.12 (2.19 - 2.12)
Grupo especial	<i>P2₁2₁2₁</i>	<i>P2₁2₁2₁</i>	<i>P2₁2₁2₁</i>	<i>P2₁2₁2₁</i>	<i>P2₁2₁2₁</i>
Cela unitária (a,b,c;Å)	67.7, 102, 102.6	67.3, 101.6, 101.9	67.5, 101.0, 102.1	67.6, 101.9, 101.9	67.6, 101.4, 102.2
Reflexões únicas	96410 (11291)	43094 (499)	40900 (4004)	41113 (4161)	39479 (3970)
Completeza (%)	93.60 (99.60)	61.77 (7.23)	99.84 (100.0)	96.77 (99.93)	97.19 (99.87)
R-work	0.2191	0.2410	0.2141	0.1923	0.2004
R-free	0.2605	0.2700	0.2658	0.2263	0.2244
Número de átomos não hidrogênio	4783	4872	4992	5017	4817
macromoléculas	4675	4663	4681	4680	4643
ligantes	66	77	66	119	69
solvente	42	132	245	218	105
Resíduos de proteína	604	603	605	604	600
RMS (ligações)	0.007	0.007	0.007	0.009	0.014
RMS (ângulos)	1.28	0.87	1.10	1.35	1.81
Ramachandran favoráveis (%)	97.00	95.49	96.84	96.83	97.99
Ramachandran permitidos (%)	3.00	4.34	3.16	2.83	1.85
Ramachandran não permitidos (%)	0.00	0.17	0.00	0.33	0.17
Rotâmeros não permitidos (%)	0.00	1.35	0.58	0.38	1.94
Clashscore	8.31	6.94	7.02	6.40	8.79
B-fator	50.42	29.16	52.56	42.83	51.25
macromoléculas	50.15	29.04	52.31	42.17	51.01
ligantes	74.64	45.08	71.73	63.79	72.19
solvente	41.66	23.86	52.16	45.55	47.99
Código PDB	7LFE	7LDX	7LFP	7KVL	7KVR

Fonte: NOSKE *et al.*¹²⁷

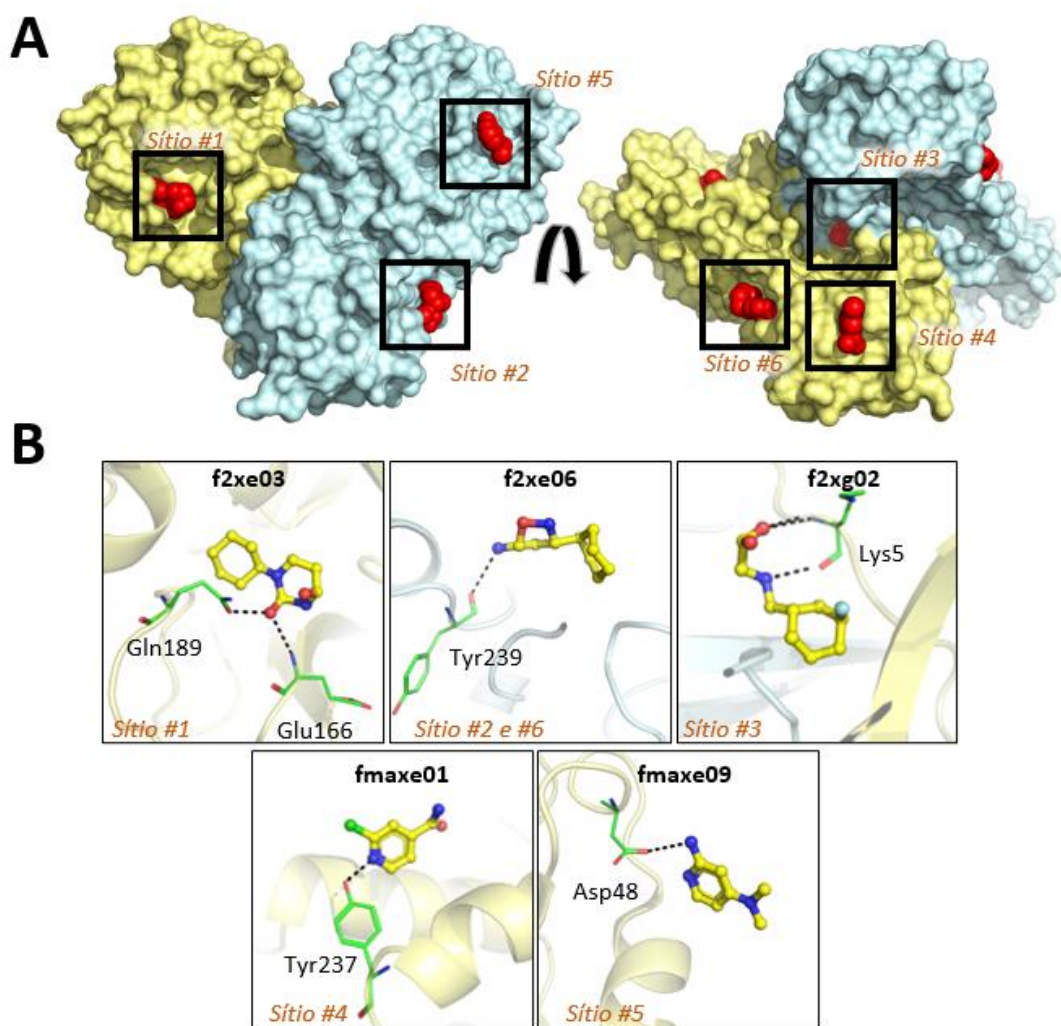


Figura 44 - A) Visão geral dos fragmentos identificados para os cristais da enzima IMT-M^{PrO}. A estrutura da enzima está representada como superfície, com as cadeias A e B coloridas em amarelo e ciano, respectivamente. Os fragmentos estão representados como esferas e coloridos em vermelho. B) Visão ampliada dos fragmentos identificados em cada um dos sítios, destacando as interações feitas por cada um deles. A estrutura da enzima está representada como *cartoon*, os fragmentos como *sticks* (coloridos em amarelo) e os resíduos parte das interações como linhas (coloridos em verde).

Fonte: NOSKE *et al.*¹²⁷

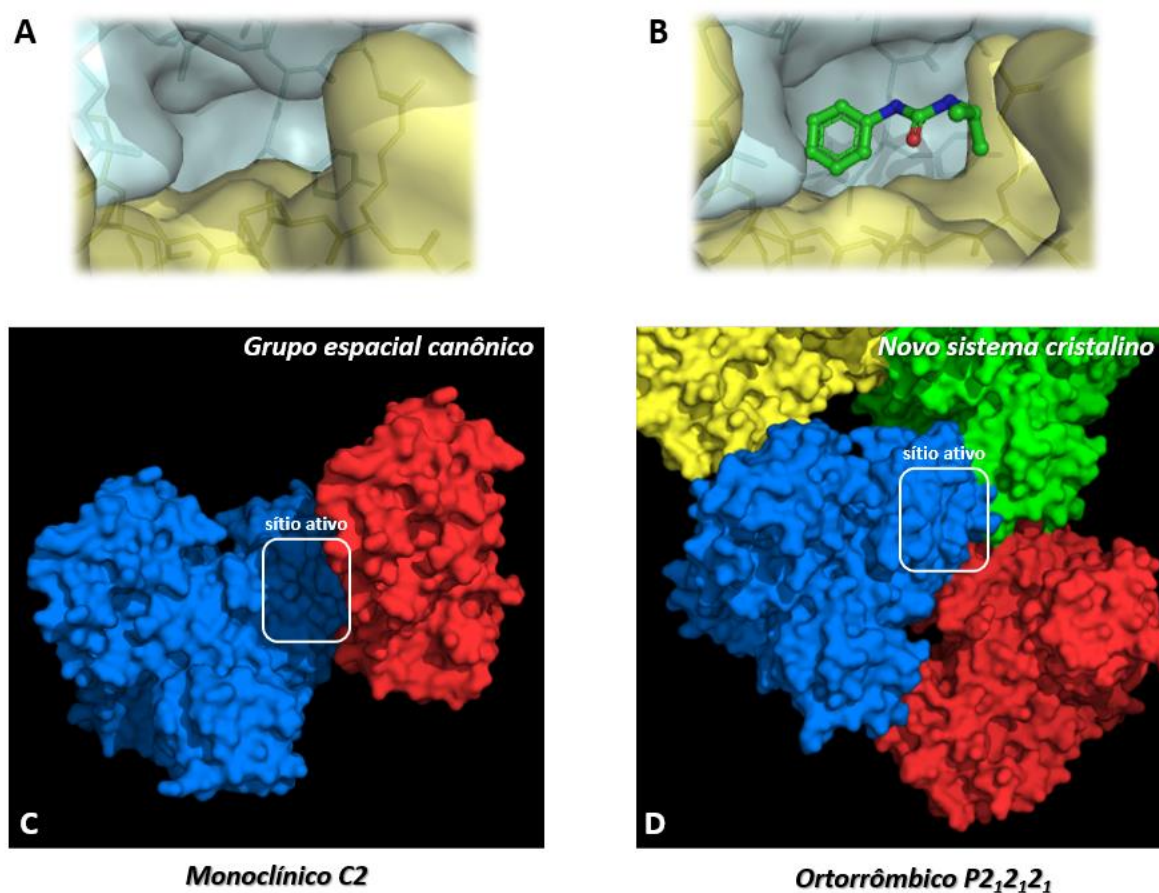


Figura 45 - Detalhes do grupo espacial dos cristais da IMT-M^{pro}. A) Representação de superfície da região onde foi identificado o fragmento f2xg02 na enzima IMT-M^{pro}, mostrada na M^{pro} cristalizada no grupo espacial canônico. B) Representação da mesma região mostrada em A, na enzima IMT-M^{pro} cristalizada no grupo espacial ortorrômbico, contendo o fragmento f2xg02. As cadeias A e B estão representadas como superfície com transparência, juntamente com os resíduos da região representados como *sticks*, e coloridas em ciano e amarelo respectivamente. C) Visão espacial da simetria presente no grupo espacial canônico da M^{pro}, destacando a região do sítio ativo. D) Visão espacial do grupo espacial da enzima IMT-M^{pro}, destacando a região do sítio ativo e evidenciando o maior espaço para acesso ao bolsão de ligação ao substrato. Cada um dos dímeros está representado como superfície e coloridos em diferentes cores.

Fonte: Elaborada pela autora.

Além das mudanças estruturais destacadas acima, a estrutura cristalográfica da enzima IMT M^{pro}, pertence ao sistema cristalino ortorrômbico (grupo espacial P2₁2₁2₁), quando o canônico observado para cristais da M^{pro} é o sistema monoclinico (grupo espacial C2₁). Este sistema, por possuir um empacotamento distinto, com 2 moléculas por unidade assimétrica (comparada a 1 no monoclinico) possui o sítio de ligação ao substrato (S1-S4), em particular os subsítios S3-S4, mais exposto ao solvente (Figura 45), desta forma torna-se ideal para utilização em experimentos de *soaking* para obtenção de estruturas da enzima M^{pro} em complexo com inibidores.¹²⁷ É possível utilizar *seeds* provenientes de cristais da IMT-M^{pro} para

obtenção de cristais da M^{pro} no grupo espacial P2₁2₁2₁ (Figura 40D). Este sistema já tem sido utilizado pelo Diamond Light Source, parte do consórcio Covid-Moonshot e tem se mostrado promissor, possibilitando a obtenção centenas de estruturas da M^{pro} em complexo com inibidores.^{118,127}

3.6 Estrutura cristalográfica da enzima M^{pro}

A enzima M^{pro} foi inicialmente cristalizada no sistema cristalino monoclinico em diversas condições e sua estrutura foi determinada na resolução de 1.46Å, no grupo espacial C2₁, como a maioria das estruturas da proteína em forma apo depositadas no PDB. Todos os 306 resíduos foram refinados na densidade eletrônica, resultando num R_{work}/R_{free} final de 0.16/0.18, com 99% dos resíduos em regiões favoráveis do gráfico de Ramachandran (Tabela 11). A unidade assimétrica contém um monômero de M^{pro} que, utilizando os operadores de simetria, pode ser usado para obter o dímero biológico, seguindo o mesmo padrão da maioria das estruturas conhecidas depositadas no PDB (r.m.s.d de 0.2Å para o PDB 5RGG, para todos os 306 resíduos).¹²⁷

De maneira já esperada para estrutura, os protômeros da M^{pro} são formados por três domínios (DI, DII e DIII), com sua região catalítica localizada entre os barris beta compreendendo DI e DII (Figura 46). Utilizando *seeds* da IMT-M^{pro}, fomos capazes de obter um novo sistema cristalino (ortorrômbico), no grupo espacial P2₁2₁2₁ com resolução final de 1.86Å. A estrutura foi refinada para R_{work}/R_{free} final de 0.19/0.22 e 98.33% dos resíduos em posições favoráveis no gráfico de Ramachandran (Tabela 11).¹²⁷ Conforme mostrado para a IMT-M^{pro} na sessão anterior, este sistema cristalino contém a estrutura completa do dímero na unidade assimétrica e seu empacotamento oferece vantagens para a realização de *soaking* de compostos que se ligam no sítio ativo da enzima, especialmente para compostos que se ligam nos subsítios S3-S4, pois são menos restritos pelo empacotamento cristalino no sistema ortorrômbico.¹²⁷ Mesmo com as diferenças observadas no empacotamento cristalino, a estrutura da enzima no geral é muito similar as estruturas no grupo espacial C2₁, evidenciando que as diferenças descritas para a IMT-M^{pro} se devem exclusivamente a presença dos resíduos adicionais no N-terminal da enzima, e não ao grupo ao empacotamento cristalino. Para confirmar esta hipótese, resolvemos também as estruturas cristalográficas da M^{pro} madura nos dois sistemas cristalinos. Detalhes das estruturas obtidas serão apresentados na sessão 3.8.

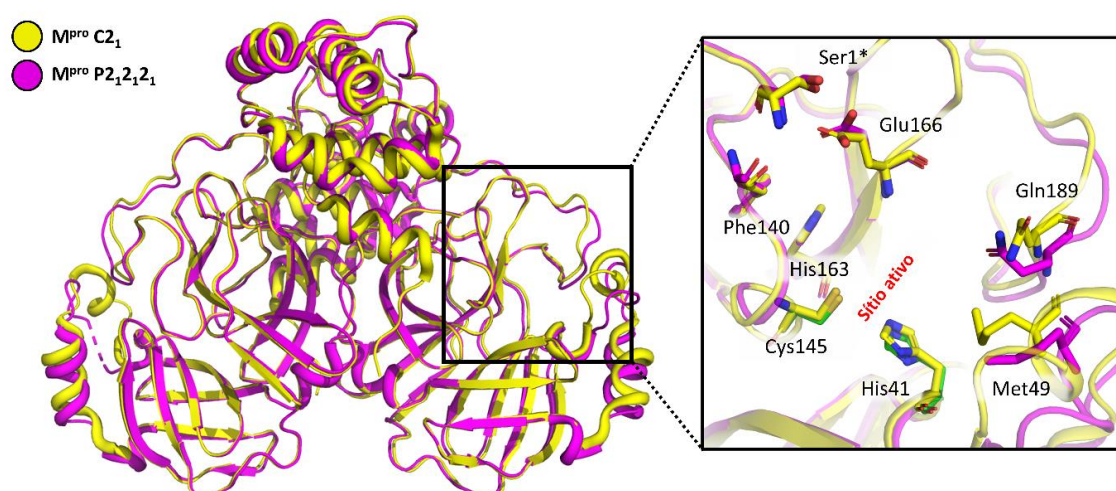


Figura 46 - Estrutura cristalográfica da enzima M^{pro} em dois grupos espaciais, $P2_12_12_1$ e $C2_1$. Ambas as estruturas estão sobrepostas, representadas como *cartoon* e coloridas em magenta e amarelo. Na esquerda, visão ampliada do sítio ativo e resíduos do bolsão de ligação ao substrato. Os resíduos estão representados como *sticks* e coloridos de acordo com cada uma das estruturas (magenta para ortorrômbica e amarela para monoclinica).

Fonte: Elaborado pela autora.

3.7 Estrutura da C145S M^{pro} em forma tetramérica em complexo com os peptídeos endógenos do N e C-terminal

O pico correspondente a C145S M^{pro} em forma tetramérica foi cristalizada e resultou na obtenção sua estrutura cristalográfica numa resolução de 2.8Å, com $R_{\text{work}}/R_{\text{free}}$ de 0.20/0.25 e 94.9% dos resíduos em região favorável no gráfico de Ramachandran (Tabela 11), revelando uma nova forma cristalina na qual observamos os resíduos do N-terminal clivados (Ser-4, Ala-3, Val-2, Leu-1, Gln-0) no sítio ativo da cadeia A, ligado aos subsítios S1-S5 (Figura 47). Apesar da mutagênese sítio dirigida da cisteína catalítica C145 para serina, a densidade eletrônica mostrou que os resíduos Gln0 e Ser1 estão ligados não covalentemente na região amino, indicando claramente que a clivagem do N-terminal foi concluída.¹²⁷

No subsítio S1, o nitrogênio NE2 da Gln0 interage com o oxigênio OE1 do Glu166 através de uma ligação de hidrogênio (2.7 Å), ao mesmo tempo a Gln0 interage com a Ser145 presente no sítio catalítico, na posição do *oxyanion hole* nativo (Figura 48). Para acomodar a cadeia lateral hidrofóbica da Leu-1 na posição P2, o resíduo Met49 e Met165 são afastados um do outro, levando a uma maior exposição do subsítio S2 quando comparado a forma apo da enzima, o que explica a capacidade deste sítio em acomodar uma grande variedade de resíduos com cadeias laterais hidrofóbicas como Leu, Met, Ile, Val e Phe.^{106,155-156} No entanto, vale

destacar que dos onze sítios de reconhecimento endógenos da M^{pro} na poliproteína viral, as sequências que possuem Leu na posição P2 são aquelas em que a M^{pro} apresentam maior eficiência catalítica, ressaltando a importância desta conformação para o design de medicamentos. Nos subsítios S3-S5, as interações da Val-2, Ala-3 e Ser-4 são mantidas principalmente por ligações de hidrogênio ligações entre os resíduos polares da proteína e as cadeias laterais peptídicas (Figura 48), o que explica a capacidade da M^{pro} em reconhecer uma grande variedade de resíduos nessas posições.¹²⁷

A estrutura cristalina da C145S M^{pro} tetrâmero revelou ainda outra etapa importante no processo de maturação de M^{pro}, na qual observamos os resíduos do C-terminal (resíduos 301–306) da cadeia B, rotacionados ~180° em comparação com a conformação canônica para o C-terminal da M^{pro} (Figura 49). Além disso, no sítio ativo da cadeia B, podemos observar a densidade eletrônica proveniente deste C-terminal rotacionado, pertencente a uma cadeia B relacionada por simetria (Figura 47) um fenômeno que também foi recentemente observado por outro grupo no mutante C145S e C145A M^{pro}.¹⁵⁷ Durante este evento, duas C145S M^{pro} em forma dimérica parecem interagir através do C-terminal com um respectivo sítio ativo, resultado em uma estrutura que consiste em um complexo tetramérico composto por uma associação dimérica (Figura 47). A densidade eletrônica deste conjunto de dados indica que o oxigênio OG da Ser145 está ligado covalentemente ao C da Gln306 da cadeia B relacionada por simetria (distância de 1.4Å), com a perda de um oxigênio pela Gln306.¹²⁷

Embora o mutante C145S tenha sido planejado como um mutante inativo, e tenha sido demonstrado como inativo nos ensaios de atividade, a serina confere características químicas similares a cisteína, o que torna possível que a enzima tenha uma lenta atividade residual de auto clivagem. Acreditamos que a atividade reduzida do mutante possibilitou a formação desses cristais depois de quase 20 dias, resultando na obtenção da estrutura neste estado intermediário da maturação. O modelo depositado não representa a ligação covalente, pois não encontramos nenhum programa de refinamento que seja capaz de ligar dois átomos fora da unidade assimétrica.¹²⁷

Resolvemos ainda a estrutura cristalográfica da enzima C145S M^{pro} correspondente ao pico monomérico, resultando num mapa final à 1.76 Å de resolução, com 95.76% de completude, no grupo espacial P2₁2₁2₁, com duas moléculas na unidade assimétrica, empacotadas de maneira similar à unidade biológica conhecida para a M^{pro}. A estrutura foi refinada resultando em um R_{work}/R_{free} final de 0.15/0.19, com 97.4% dos resíduos em regiões favoráveis no gráfico de Ramachandran (Tabela 11). A estrutura obtida tem alta similaridade com a estrutura da forma madura da enzima, com distinção apenas pela cisteína catalítica

mutada para serina, possuindo um RMSD calculado de 0.484 para 3750 átomos. Devido à alta similaridade entre as estruturas, esta não será discutida em detalhes.

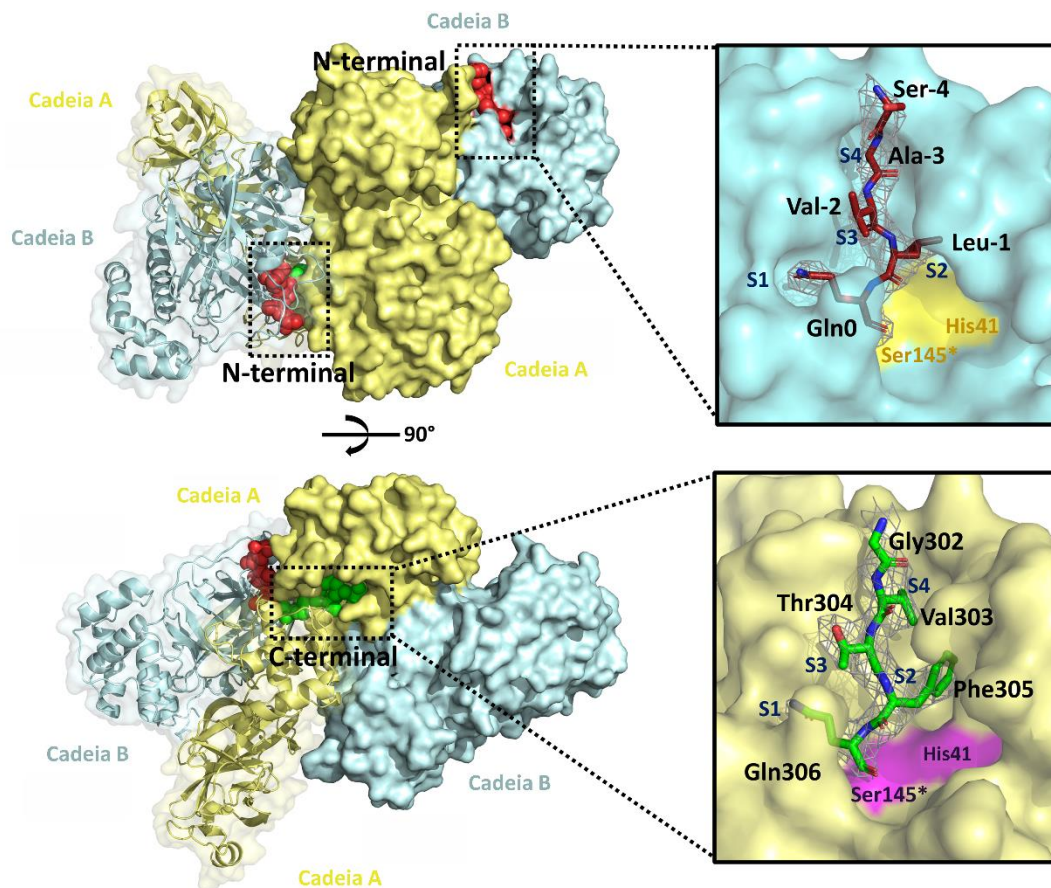


Figura 47 - Estrutura cristalográfica da C145S M^{Pro} tetrâmero. À direita, a estrutura do dímero presente na unidade assimétrica está representada como superfície, e as cadeias A e B coloridas de amarelo e azul respectivamente. O dímero relacionado por simetria está representado como superfície transparente e *cartoon*, seguindo as mesmas cores. Os peptídeos ligados no sítio ativo estão representados como esferas, coloridos em vermelho e verde para o N e C terminal respectivamente. À direita, visão ampliada do sítio ativo de cada uma das cadeias, destacando os peptídeos ligados ao sítio ativo, mostrando a densidade eletrônica 2Fo-Fc de cada um deles ($\sigma = 1$).

Fonte: Adaptada de NOSKE *et al.*¹²⁷

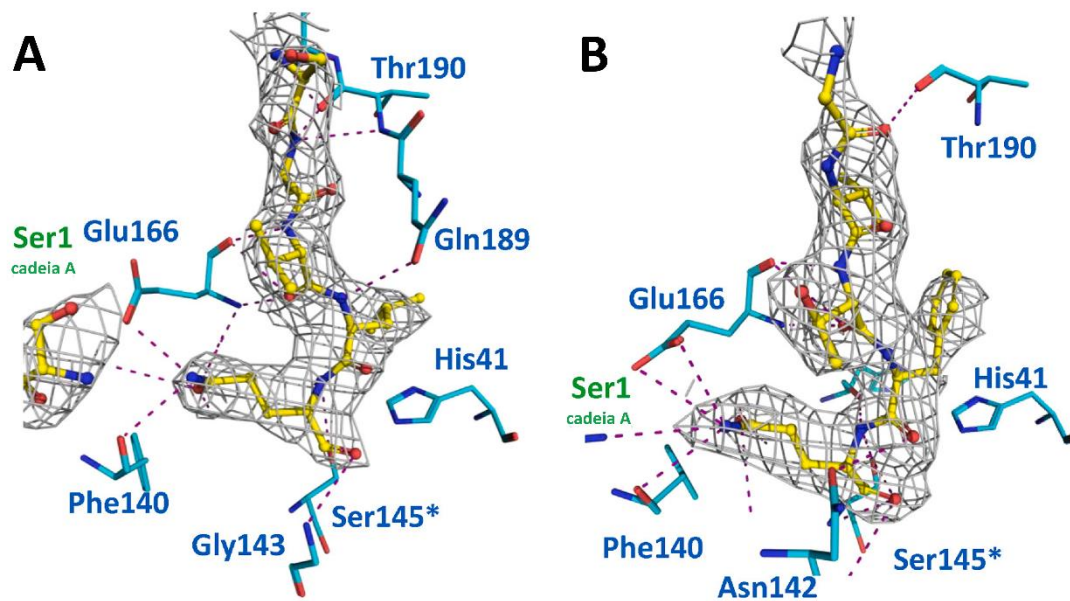


Figura 48 - Interações feitas pelos peptídeos do N e C-terminal com os resíduos do sítio ativo da proteína C145S M^{Pro}. A) Detalhe das interações para o peptídeo N-terminal (cadeia A). B) Detalhe das interações para o peptídeo do C-terminal (cadeia B). Em ambas, os peptídeos estão representados como *ball and stick* e coloridos em amarelo e os resíduos parte das interações estão representados como linhas e coloridos em ciano.

Fonte: NOSKE *et al.*¹²⁷

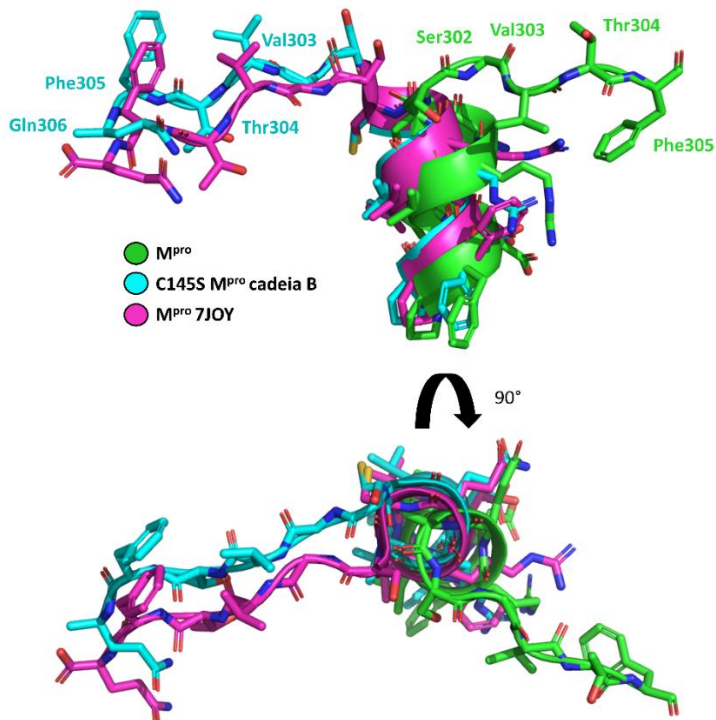


Figura 49 - Superposição dos resíduos do C-terminal da M^{Pro} para as enzimas M^{Pro} madura, C145S M^{Pro} tetrâmero - cadeia B e C145A M^{Pro} (PDBid 7JOY). As estruturas estão representadas como *cartoon* e *sticks* e coloridas em verde, ciano e magenta, respectivamente.

Fonte: Adaptada de NOSKE *et al.*¹²⁷

3.8 Caracterização em solução da enzima C145S M^{pro}

Conforme observado na estrutura cristalográfica da C145S M^{pro}, apesar da mutagênese sítio dirigida da cisteína 145 para serina, esta enzima ainda exibiu atividade proteolítica. Utilizando desta atividade, foi possível observar o auto processamento do pico monomérico da C145S M^{pro} por SDS-PAGE no decorrer de dois dias (Figura 50A).¹²⁷ Após quantificar as intensidades das bandas, estimamos que ao final de dois dias cerca de 30% dos a proteína foi auto clivada após a incubação (Figura 50B). Também utilizamos em conjunto, SEC-MALS, para monitorar a taxa de formação de dímeros a partir dos monómeros ao longo do tempo (após 0h, 24h, 48h e 72h), com incubação à temperatura ambiente (Figura 50C-D) No tempo de 0h, a razão de recuperação de massa entre monômeros/dímeros foi de 14.9, que diminuiu para 1.04 às 24h e 0.09 às 48 horas de incubação, com completa degradação do pico do monômero após 72h (Tabela 13). Esses dados indicaram que a clivagem de o N-terminal é diretamente proporcional a formação de dímeros em solução, destacando a importância do processamento N-terminal para a M^{pro}. Além disso, a taxa de formação de dímeros aparenta ser superior a taxa de clivagem, o que nos permite inferir que o dímero é necessário e precede o evento de clivagem.¹²⁷

Para investigar o efeito da M^{pro} madura no processamento do N-terminal, monitoramos o efeito da adição de M^{pro} às amostras de C145S M^{pro} monômero na proporção de 1:6000. No SDS-PAGE, podemos ver que a amostra contendo M^{pro} mostrou uma taxa de aumento de clivagem de proteínas após 20h quando comparada com o experimento anterior (Figura 50C-D) No SEC-MALS, a taxa de recuperação de massa entre monômeros/dímeros para esta amostra às 0h foi de 12.9, depois 0.5 às 24 horas e 0.02 às 48 horas de incubação, também com degradação completa do pico do monômero após 72h (Figura 50C-D e Tabela 13). Desta forma, os dados mostraram que a adição de M^{pro} madura na C145S M^{pro} monomérica aumentou a velocidade de processamento do N-terminal e formação de dímeros em aproximadamente 50% após 24 horas. Este fenômeno demonstra que a clivagem trans (dímero-dímero) ocorre e auxilia no processamento do N-terminal da enzima.¹²⁷

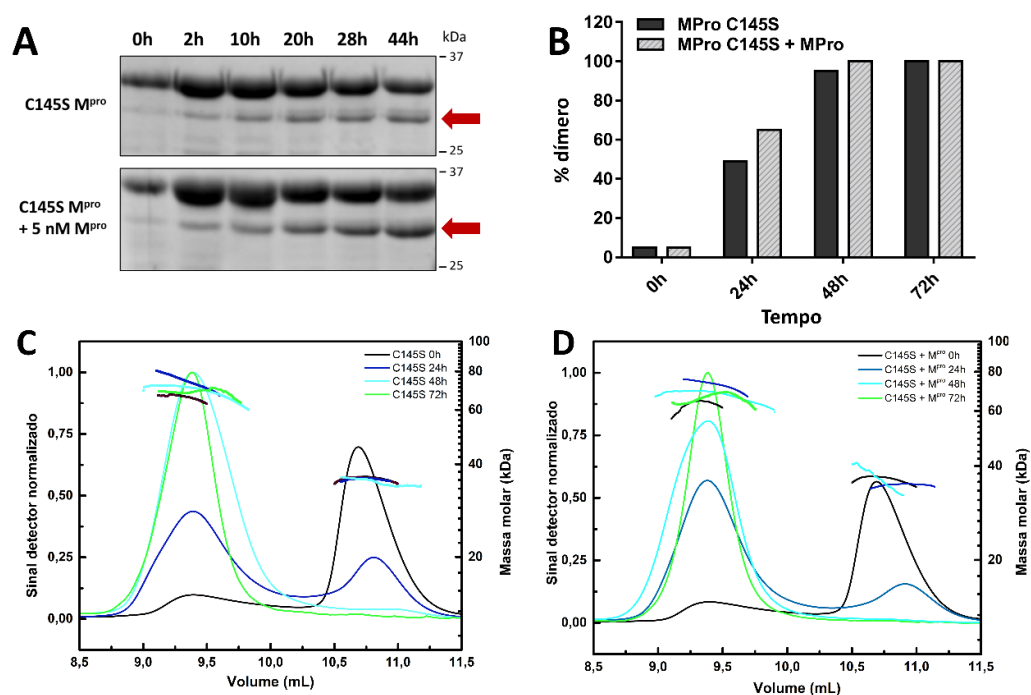


Figura 50 - A) SDS-PAGE contendo a clivagem do N-terminal da C145S M^{Pro} ao longo do tempo. Na parte superior, reação contendo 10 mM C145S M^{Pro}, e na parte inferior a mesma reação suplementada com 5 nM de M^{Pro}. As setas vermelhas indicam para a banda de clivada da M^{Pro}. B) Gráfico de barras mostrando a intensidade relativa das bandas de M^{Pro} clivada e não clivada ao longo do tempo. C-D) Perfis de eluição na cromatografia de exclusão molecular da C145S M^{Pro} ao longo do tempo para as reações sem e com M^{Pro} madura respectivamente. C. As curvas correspondem a mudança na intensidade no espalhamento normalizado em 90° (linhas) e massas moleculares calculadas para cada um dos picos (pontos).

Fonte: Adaptada de NOSKE *et al.*¹²⁷

Tabela 13 - Massa molecular e razão de recuperação das massas entre as porcentagens dos picos obtidos no SEC-MALS para a C145S M^{Pro}. A razão de recuperação da massa foi calculada entre as porcentagens do pico 1 e 2.

Proteína	Tempo (h)	Pico 1 – Monômero		Pico 2 - Dímero		Razão de recuperação da massa (pico 1/pico 2)
		Massa molar (kDa)	Recuperação de massa (%)	Massa molar (kDa)	Recuperação de massa (%)	
C145S M ^{Pro}	0	36.0 ± 0.1	74.5	65.7 ± 0.2	5.0	14.9
	24	35.6 ± 0.1	29.6	73.7 ± 0.1	28.4	1.04
	48	34.9 ± 0.1	7.4	69.1 ± 0.1	84.5	0.09
	72	-	-	68.7 ± 0.2	87.3	-
C145S M ^{Pro} + M ^{Pro}	0	36.4 ± 0.1	75	63.0 ± 0.1	5.8	12.9
	24	34.8 ± 0.2	22.8	72.9 ± 0.1	43.5	0.5
	48	36.7 ± 0.2	2	67.0 ± 0.1	86.7	0.02
	72	-	-	66.1 ± 0.3	91.3	-

Fonte: NOSKE *et al.*¹²⁷

Além disso, avaliamos a atividade residual da enzima C145S M^{pro} ao longo do tempo de auto clivagem utilizando o ensaio FRET. A formação de dímeros ao longo do tempo, aumenta significativamente a atividade residual do mutante C145S (Figura 51), demonstrando que mesmo no caso do sítio ativo contendo serina, a dimerização é fundamental para a atividade da enzima.¹²⁷

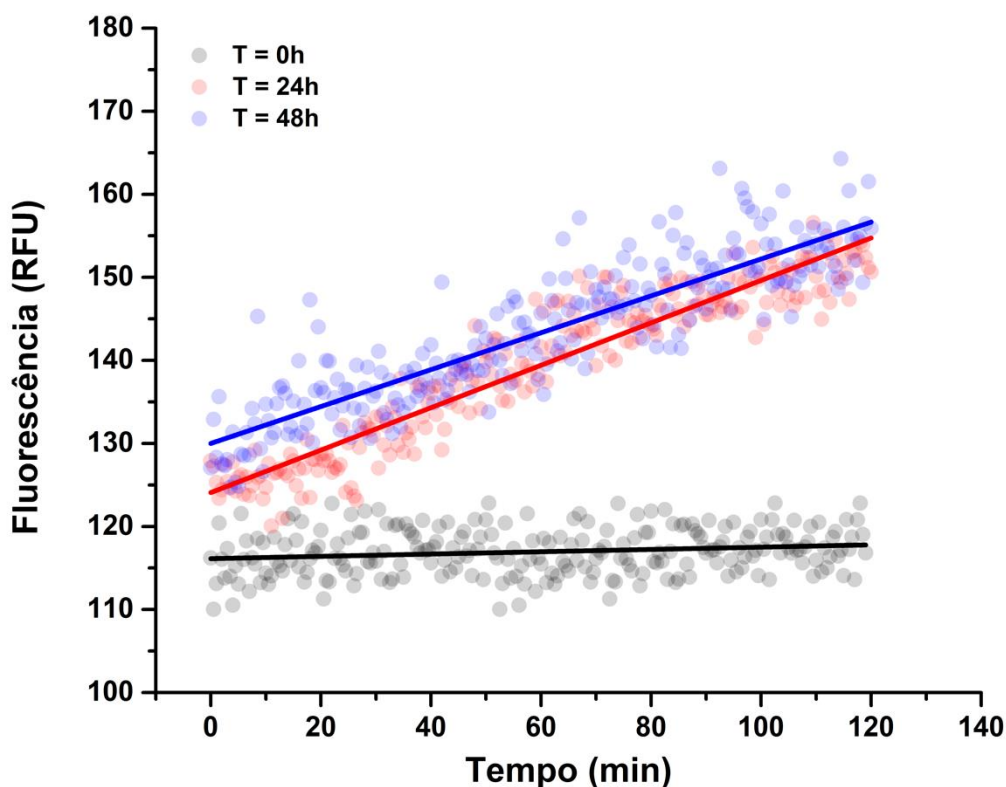


Figura 51 - Curvas de fluorescência ao longo do tempo para diferentes tempos de auto clivagem da enzima C145S M^{pro} monômero.
Fonte: NOSKE *et al.*¹²⁷

3.9 Estrutura da enzima C145S M^{pro} obtida por crio-microscopia eletrônica (cryo-EM)

Após a realização dos ensaios biofísicos, caracterização em solução e obtenção das estruturas cristalográficas, obtivemos também a estrutura da M^{pro} C145S tetrâmero por cryo-EM.¹³⁹ A estrutura foi obtida numa resolução final de 3.5Å (Figura 53) e o modelo final refinado contém todos os 306 resíduos da C145S M^{pro} visíveis para ambas as cadeias, o mapa revelou também a presença de pelo menos 11 resíduos pertencentes a uma cadeia polipeptídica ligada a ambos os sítios ativos (Figura 54). O modelo final obtido não mostrou nenhum rotâmero, C β ou Ramachandran *outliers*, e apresentou coeficiente de correlação (CC) de 0.80 (Tabela 14). O modelo e/ou mapa foram depositados sob os códigos 8EY2 (PDB) e EMD-28666 (EMDB).¹³⁹

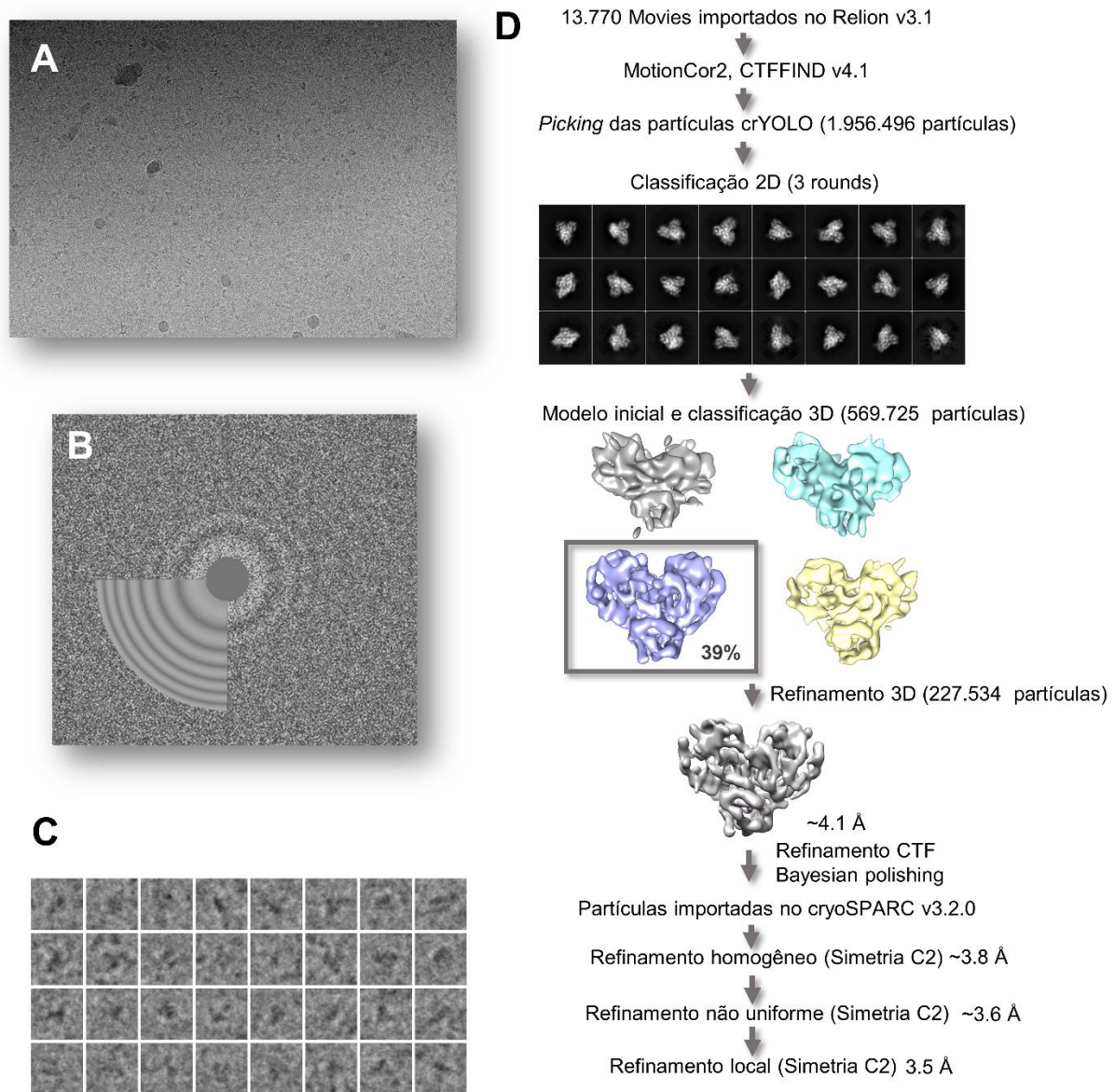


Figura 52 - Esquema do processamento dos dados de crio-microscopia eletrônica da C145S M^{pro}. A) Exemplo de micrografias alinhada. A barra de escala corresponde a 20 nm. B) Anéis de Fourier da micrografia apresentada em A. C) Exemplos de partículas presentes nas micrografias. D) Passo a passo detalhado contendo as etapas de processamento e refinamento dos dados.

Fonte: NOSKE *et al.*¹³⁹

Tabela 14 - Coleta de dados de crio microscopia, parâmetros de refinamento e validação e estatísticas.

<i>Coleta de dados</i>	
Voltagem	300
Magnificação	130.000
Câmera	K3 super resolução
<i>Slit</i> abertura (eV)	20
Pixel Super-resolução (Å)	0.3265
Pixel <i>Binned</i> (Å)	0.653
Dose (e/px/s)	18.5
Tempo de exposição (s)	1.0
Dose total (e/Å ²)	43.385
frações	43
dose por frame (e/Å ² /frame)	1.009
defocus (µm)	-0.5 to -2.5
Movies coletados	13,770
<i>Processamento e refinamento</i>	
Partículas	227.533
Cadeias	4
Átomos	4908
Resíduos	634
Ligações (RMSD)	
Distância (Å)	0.003
Ângulo (°)	0.6
MolProbity score	2.04
Clashscore	13.98
Ramachandran (%)	
Não permitidos	0
Permitidos	5.75
Favoráveis	94.25
Rotâmeros não permitidos (%)	0
Cβ não permitidos (%)	0
CaBLAM não permitidos (%)	1.46
Box	
Tamanho (Å)	72.48, 86.85, 74.44
Ângulos (°)	90, 90, 90
dFSC modelo com máscara (0/0.143/0.5)	3.3/3.5/3.8
Modelo vs. Dados	
CC (máscara)	0.80
CC (box)	0.78
Código PDB	8EY2
EMDB	EMD-28666
EMPIAR	EMPIAR-10810

Fonte: NOSKE *et al.*¹³⁹

A estrutura obtida por cryo-EM é muito semelhante às estruturas obtidas por cristalografia de raios X (Figura 57), com um RMSD de 0.7Å (para 3.909 átomos) quando comparado com a forma dimérica da forma madura da M^{pro} (PDB 7KPH), ou 1.2Å (para 3.822 átomos) quando comparada com a estrutura de raios X do mutante C145S (PDB 7N5Z). (Figura 57). Estatísticas e parâmetros de coleta de dados e processamento estão disponíveis na Tabela 14.¹³⁹

Anteriormente, foi assumido que a amostra tetramérica da C145S M^{pro} observada em solução tinha a mesma organização observada na estrutura cristalina, correspondente a uma associação dímero-dímero.¹²⁵ No entanto, a estrutura de crio microscopia da C145S M^{pro} revelou uma partícula dimérica de M^{pro} (já madura) ligada não covalentemente no peptídeo contendo a região nsp4-nsp5 no momento que antecede a clivagem, com densidade detalhada dos principais resíduos das posições P e P' (Figura 54B-C).¹³⁷ A estrutura na resolução de 3.5 Å fornece informações estruturais sobre uma etapa importante da maturação da M^{pro}, a clivagem do N-terminal, exibindo densidade eletrônica clara para os peptídeos peptídeo em ambas as cadeias A e B do dímero (Figura 54). O mapa de superfície de cargas ao redor da Ser145* do sítio ativo indica uma ligação não covalente do peptídeo nsp4/nps5 (resíduos SAVLQ da nsp4 e resíduos 1-6 da nsp5 M^{pro}, SGFRKM), formando uma folha β cuja interação é mantida principalmente por ligações de hidrogênio com os resíduos Thr24, Gly143, Ser145*, His163, Glu166 e Gln189 (Figura 54C). O posicionamento do substrato no sítio ativo exibiu alta similaridade às estruturas cristalográficas da M^{pro} em complexo com os substratos peptídicos fisiológicos (PDBid 7N6N e 7DVP) com valores de RMSD de respectivamente 0.57 e 0.9, calculados entre 30 e 54 átomos.¹³⁷

No modelo da estrutura de cryo-EM podemos ver a Gln0 ocupando o subsítio S1, com sua porção de cadeia lateral NE2 interagindo com o átomo OE1 do resíduo Glu166 através de uma ligação de hidrogênio (2.7 Å e 2.5 Å para as cadeias A e B, respectivamente). Além disso, o átomo OG da Ser145* forma uma ligação de hidrogênio com o grupo hidroxila da cadeia principal da Gln0. A amida de cadeia principal do resíduo Gly143 atua como doador de ligação de hidrogênio ao oxigênio carbonílico em P1', estabilizando o *oxyanion hole* durante a catálise. A Leu-1 no subsítio S2 interage com a Gln189 através de uma ligação de hidrogênio formada entre a cadeia principal da amida da Leu-1 a cadeia lateral da Gln189 OE1 (Figura 54C).¹³⁷ Os resíduos Met49, Met149 e Gln189 do subsítio S2 assumem uma conformação mais aberta em comparação com formas apo da enzima. No subsítio S3, há uma interação entre os átomos polares da cadeia principal da Val-2 e a cadeia lateral do Glu166 (Figura 54C). O grupo hidroxila da cadeia principal da Ala-3 em P4 doa uma ligação de hidrogênio para o nitrogênio NE2 da Gln189. A ausência de interações polares entre os resíduos formadores dos sítios S2-S4 com os resíduos que se ligam nestas posições, ajuda a explicar a variedade de aminoácidos que estão presentes nestas posições na poli proteína. O mapa de densidade do peptídeo ligado ao sítio ativo da M^{pro} se estende apenas em direção aos resíduos Met6, sugerindo que o restante da partícula é muito móvel para a obtenção desta região no volume obtido após reconstrução.¹³⁷

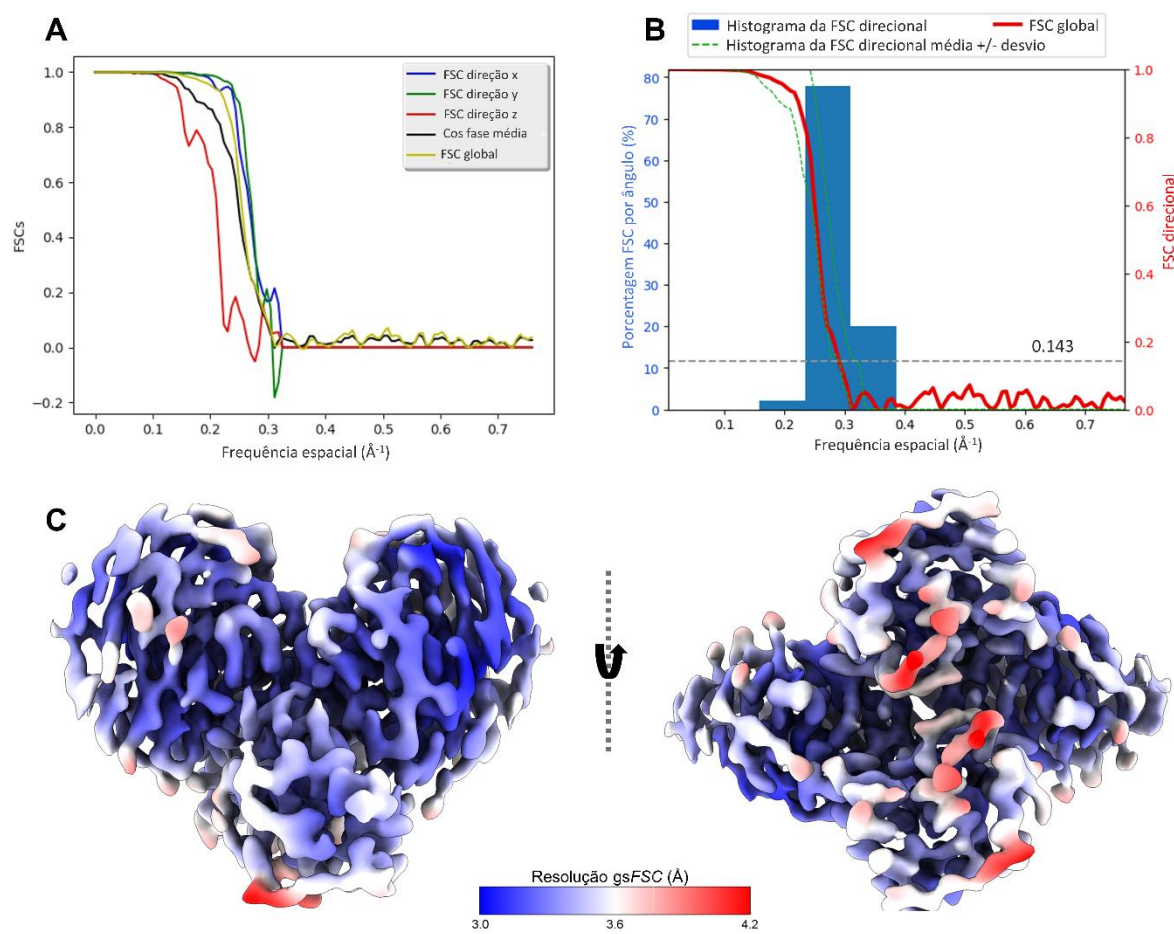


Figura 53 - Camadas de correlação de Fourier (FSC) mostrando a resolução global (FSC = 0.143) e mapa de resolução local. A resolução está colorida de acordo com a resolução local, de vermelho para o azul, de acordo com valores crescentes de resolução.

Fonte: NOSKE *et al.*¹³⁷

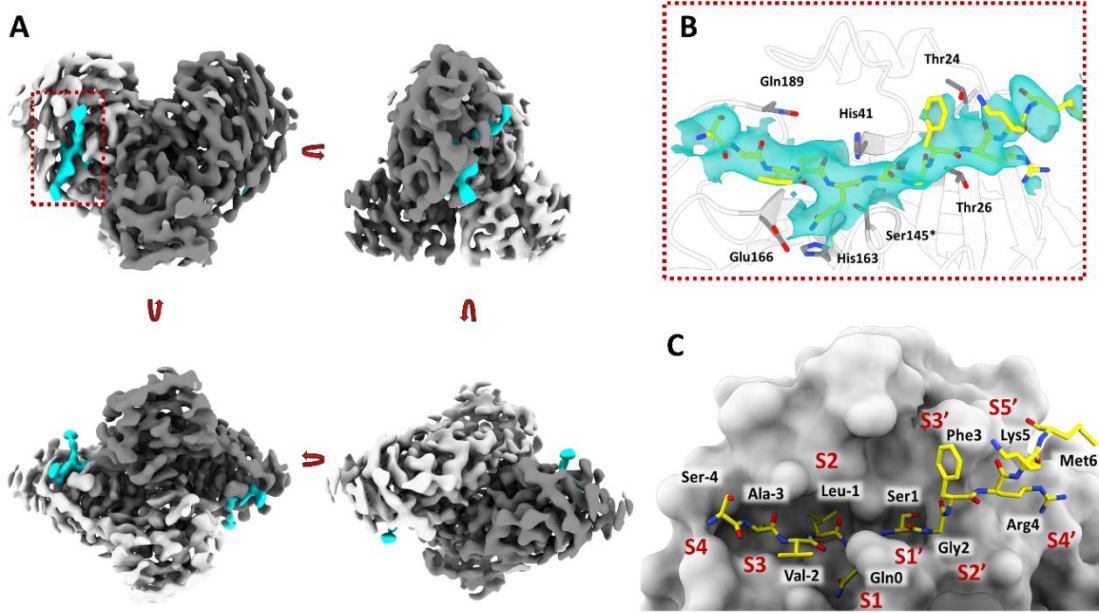


Figura 54 - A) Visão geral da densidade do volume final reconstruído para a C145S M^{PRO} obtido por criomicroscopia eletrônica. Cada uma das cadeias A e B estão coloridas respectivamente em cinza claro e escuro. O peptídeo correspondente a região nsp4-nsp5 está colorido em ciano. B) Visão ampliada do sítio ativo da estrutura da cryo-EM da C145S M^{PRO}. A densidade correspondente ao peptídeo nsp4-nsp5 está colorida em ciano e os resíduos pertencentes à este peptídeo estão representados como *sticks* e coloridos em amarelo. Os resíduos que interagem com o peptídeo através de interações polares estão representados na figura como *sticks* e coloridos em cinza. C) Visão ampliada do sítio ativo da enzima mostrando a região como superfície, destacando o bolsão de ligação ao substrato e os subsítios S5-S5'. Cada um dos subsítios está indicado em vermelho, e os resíduos pertencentes a este peptídeo estão representados como *sticks* e coloridos em amarelo.

Fonte: NOSKE *et al.*¹³⁷

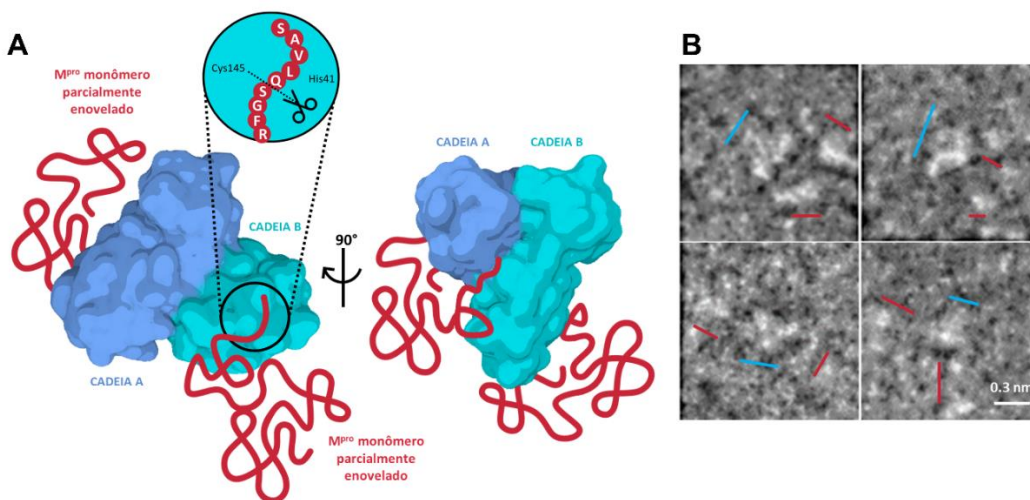


Figura 55 - Representação esquemática da M^{PRO} tetramérica em solução. A M^{PRO} está representada como superfície com as cadeias A e B coloridas em azul e ciano, respectivamente. As M^{PRO} monoméricas parcialmente enoveladas estão coloridas em vermelho. B) Visão detalhada das partículas de M^{PRO} encontradas nas micrografias, onde as moléculas de Mpro dimérica estão marcadas em azul e as M^{PRO} monômero associadas ao dímero estão marcadas em vermelho.

Fonte: Adaptada de: NOSKE *et al.*¹³⁷

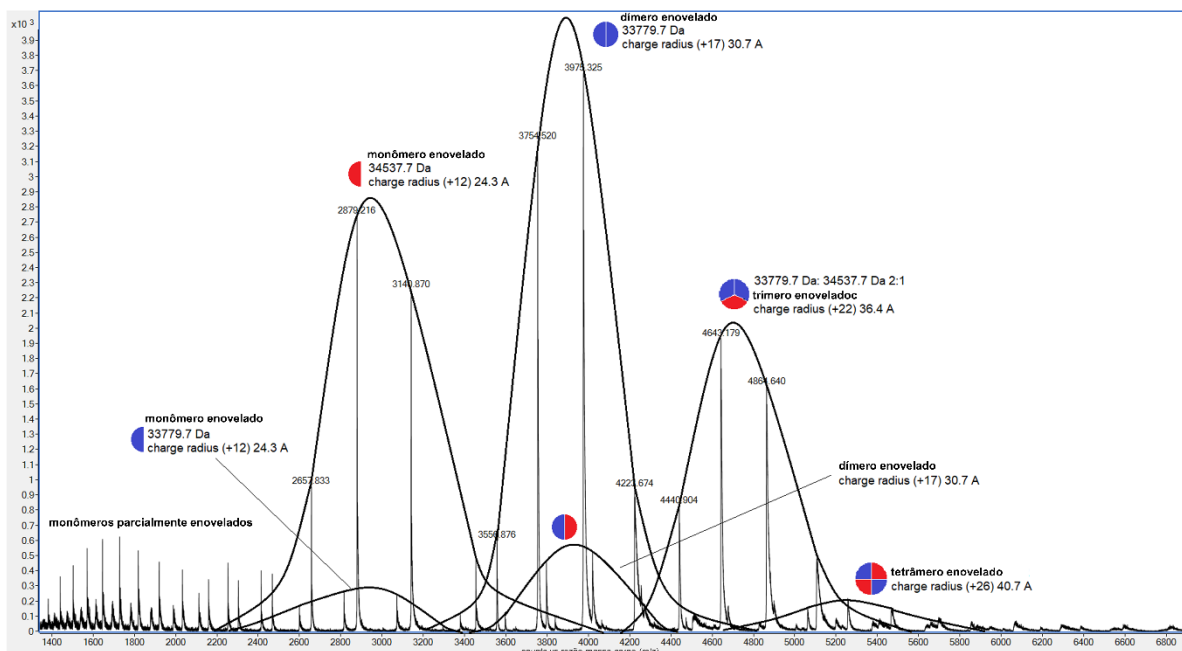


Figura 56 - Espectrometria de massas nativa da M^{pro} tetramérica. Da esquerda para a direita, os picos mostram monômeros clivados (semicírculo azul) e não clivados (semicírculo vermelho), dímeros formados por partículas clivadas (círculos azuis) ou semi-clivadas (círculo azul-vermelho), trímeros formados por duas partículas clivadas e uma não clivada (dois terços azuis, um terço círculos vermelhos) e tetrâmeros formados por duas partículas clivadas e duas não clivadas (dois quartos de azul, dois quartos de vermelho).

Fonte: Adaptada de NOSKE *et al.*¹³⁷

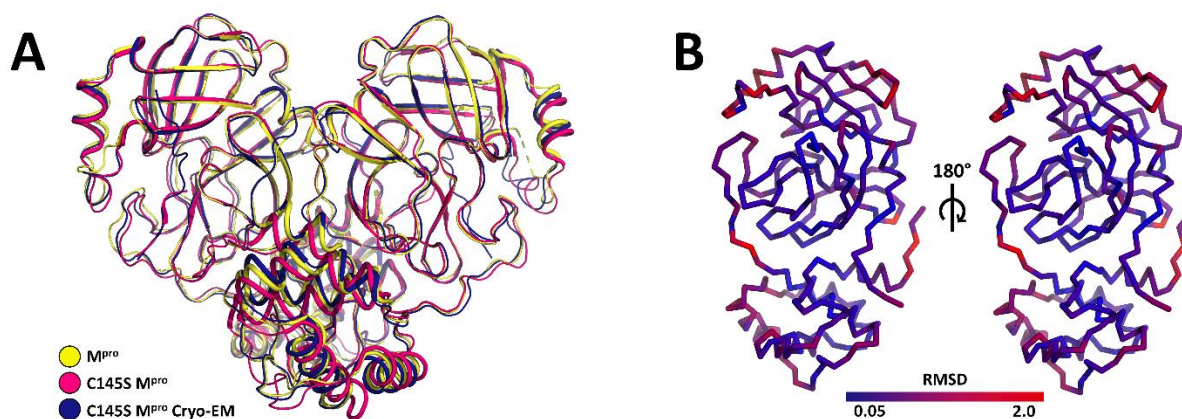


Figura 57 - Comparação entre as diferentes estruturas de M^{pro} , M^{pro} madura (PDBid 7MBG), C145S M^{pro} tetrâmero (PDBid 7N6N) e C145S M^{pro} obtida por cryo-EM (7S82). As estruturas estão representadas como cartoon e coloridas respectivamente em amarelo, rosa e azul. B) RMSD calculado entre a estrutura da C145S M^{pro} tetrâmero e a estrutura obtida por cryo-EM. A proteína está representada como traço da cadeia alpha e colorida de acordo com os valores de RMSD, do azul para o vermelho, de acordo com a legenda.

Fonte: Adaptada de NOSKE *et al.*¹³⁷

3.10 Análise da influência de inibidores no processo de maturação da M^{pro}

3.10.1 Inibidor não-covalente: MAT-POS-e194df51-1

Dada a vantagem em se utilizar a C145S M^{pro} como uma versão da enzima com velocidade de processamento reduzida para o estudo da maturação da proteína, utilizamos o sistema estabelecido para analisar a influência de dois inibidores no processo de maturação da enzima, um não covalente e outro covalente. O primeiro inibidor analisado foi o inibidor não-covalente competitivo MAT-POS-e194df51-1 desenvolvido pelo consórcio *COVID Moonshot*, que possui pIC₅₀ de 7.5.¹¹⁶ Amostras contendo C145S M^{pro} monomérica (amostra 1) ou C145S M^{pro} tetramérica (amostra 2) foram incubadas com duas concentrações de MAT-POS-e194df51-1, e seus estados oligoméricos foram monitorados ao longo de 48 horas por SEC-MALS.¹³⁷

No tempo 0h, a amostra controle apresentou uma razão de 1.0/0.0 (monômero/dímero), progredindo para 0.88/0.12 após 24 horas e 0.39/0.61 após 48 horas (Figura 58A). A amostra 1 contendo 4:1 MAT-POS-e194df51-1/proteína, permaneceu em uma proporção de 1.0/0.0 monômero/dímero entre 0 h e 24 h, progredindo apenas para 0.004/0.996 no período de 48 horas (Figura 58B). Reduzindo a razão molar de MAT-POS-e194df51-1/proteína de 4:1 para 0.4:1, observamos comportamento mais próximo da amostra controle, chegando a uma razão monômero/dímero de 0.89/0.11 após 24 h e 0.74/0.25 após 48 h. A amostra 1 mostrou que os monômeros estão basicamente formando dímeros maduros ao longo do tempo. Ambas as concentrações de MAT-POS-e194df51-1 parecem inibir a maturação do dímero e causar o acúmulo de partículas de monômeros parcialmente enoveladas. Com este experimento conseguimos demonstrar um efeito único do MAT-POS-e194df51-1 sobre o ciclo enzimático com mecanismo nunca descrito, mostrando não só a sua capacidade de competir com o substrato, mas também bloquear o ciclo de maturação da enzima e gerar um acúmulo de poli proteínas não clivadas.¹³⁷

Em 2018, Constant et al demonstraram para a protease NS3 de dengue que inibidores eficazes devem ter como alvo específico os sítios de reconhecimento da enzima que possam acumular precursores de proteínas virais não clivados.¹⁵⁶ Neste caso, este fenótipo pode inibir trans dominantemente a replicação do vírus na célula, possivelmente prevenindo a geração de resistência aos medicamentos. No entanto, mais estudos serão necessários para compreender o efeito de elementos não clivados no metabolismo do SARS-CoV-2.

Para a amostra 2 (tetrâmero), vimos a relação de 0.43/0.34/0.23 entre as massas de monômero/dímero/tetrâmero em 0h, mudando para 0.09/0.89/0.02 em 24 horas e 0.0/0.98/0.02 em 48 horas (Figura 59A). Para amostra 2 contendo 4:1 MAT-POS-e194df51-1/proteína, observamos a proporção de 0.49/0.45/0.05 entre as massas de monômero/dímero/tetrâmero em 0 h, mudando para 0.49/0.51/0.0 em 24h e 0.45/0.55/0.02 em 48h (Figura 59B). Para a amostra 2 contendo 0.4:1 MAT-POS-e194df51-1/proteína, foi obtida a proporção de 0.46/0.47/0.06 monômero/dímero/tetrâmero em 0h, mudando para 0.26/0.73/0.0 em 24h e 0.02/0.96/0.02 em 48h. No controle deste experimento, vimos que a amostra inicia como uma mistura de tetrâmeros, dímeros e monômeros. Ao longo de 48 horas, ambos os monômeros e tetrâmeros são extintos, e os dímeros tornam-se predominante. A presença de MAT-POS-e194df51-1 na amostra 2 parecem causar o mesmo efeito que na amostra 1, bloqueando a conversão de dímeros em monômero, mas também parecem acelerar o consumo de tetrâmeros. Se pensarmos nas amostras de tetrâmeros como sendo o substrato enzimático complexo, é natural que um inibidor competitivo tenha este efeito na amostra, pois ocasiona a remoção dos monômeros ligados ao sítio ativo como substrato. Paralelamente à amostra 1, a conversão de monômeros da amostra 2 em dímeros também foi bloqueada pela presença de MAT-POS-e194df51-1, revelando que o inibidor não só bloqueia a atividade, mas também pode impedir o processo de maturação, possivelmente aumentando seu efeito antiviral.¹³⁷

3.10.2 Inibidor covalente: Nirmatrelvir

O efeito do inibidor covalente de M^{pro} Nirmatrelvir/PF-07321332, que possui pIC₅₀ de 7.710 também foi testado contra as amostras de M^{pro} C145S.¹¹⁷ O efeito desta molécula foi notoriamente diferente do observado para MAT-POS-e194df51-1 (Figura 58C). Para a amostra 1 contendo Nirmatrelvir/proteína 4:1, vimos a relação de 0.65/0.35 entre monômeros/dímeros em 0h, mudando para 0.06/0.94 após 24h e 0.02/0.98 após 48h (Figura 58C). Para a amostra 2 contendo Nirmatrelvir/proteína 4:1, observamos a proporção de 0.31/0.56/0.12 entre monômeros/dímeros/tetrâmeros em 0h, mudando para 0.0/0.88/0.12 após 24h e 0.0/0.98/0.02 após 48h (Figura 58C).¹³⁷

Na amostra 1, a presença de Nirmatrelvir parece não somente ainda permitir a formação de dímeros, mas induz fortemente sua formação a partir de monômeros, desde o momento da adição do composto (t = 0h) e aumenta significativamente a proporção de formação de dímeros ao longo do tempo, com conversão quase completa após 24h (Figura 58C). Na amostra 2, ambos os monômeros e tetrâmeros também parecem estar deslocados para favorecer a formação do

dímero (Figura 59C). O aumento do consumo do tetrâmero pode ser explicado simplesmente pela competição entre o Nirmatrelvir e o complexo substrato-enzima. No entanto, a proporção acelerada de conversão de monômero em dímeros foi completamente inesperada. Ao mesmo tempo, estes resultados podem esclarecer os detalhes da primeira etapa de maturação da proteína, o processamento do N-terminal.¹³⁷

Baseados em nossos resultados anteriores, inicialmente propusemos que nesta etapa de clivagem do N-terminal ocorre uma mistura de eventos cis e trans, e sua clivagem eliminaria o impedimento estérico que impedindo a dimerização e formação da enzima ativa, condizente com modelos anteriores para M^{pro} de SARS-CoV.¹²¹⁻¹²² No entanto, fato de que o inibidor covalente (mas não um não covalente) pode induzir a dimerização em uma amostra não clivada sugere que o fundamental para a dimerização não é o processamento do N-terminal, mas o ajuste induzido causado pela própria ligação covalente. Este modelo também explicaria por que oligômeros podem ser formados pela combinação de partículas não clivados e partículas clivadas em vez de apenas partículas clivadas (Figura 56).¹³⁷

Tabela 15- Massa molecular e razão de recuperação das massas entre as porcentagens dos picos obtidos no SEC-MALS para a C145S M^{pro}. A razão de recuperação da massa foi calculada entre as porcentagens do pico 1 e 2.

Proteína	Tempo (h)	C145S M ^{pro} monômero Amostra 1			C145S M ^{pro} tetrâmero Amostra 2			Razão de recuperação da massa (monômero/dímero/tetrâmero)
		Recuperação de massa (%)		Razão de recuperação da massa (monômero/dímero)	Recuperação de massa (%)			
		monômero	dímero		monômero	dímero	tetrâmero	
<i>Controle DMSO</i>	0	100	0	1/0	43	34	23	0.43/0.34/0.23
	24	88	12	0.88/0.12	2	88	10	0.09/0.89/0.02
	48	39	61	0.39/0.61	0	98	2	0/0.98/0.02
<i>1:4 MAT-POSe194 df51-1</i>	0	100	0	1/0	49	45	5	0.49/0.45/0.05
	24	100	0	1/0	49	51	0	0.49/0.51/0
	48	99	1	0.996/0.004	45	55	2	0.45/0.55/0.02
<i>1:0.4 MAT-POSe194 df51-1</i>	0	100	0	1/0	46	37	6	0.46/0.47/0.06
	24	89	11	0.89/0.11	26	73	1	0.26/0.73/0.01
	48	74	25	0.74/0.25	2	96	2	0.02/0.96/0.02
<i>1:4 nirmatrelvir</i>	0	64	35	0.64/0.35	12	56	30	0.12/0.56/0.30
	24	6	94	0.06/0.94	0	89	11	0/0.89/0.11
	48	2	98	0.02/0.98	0	98	2	0/0.98/0.02

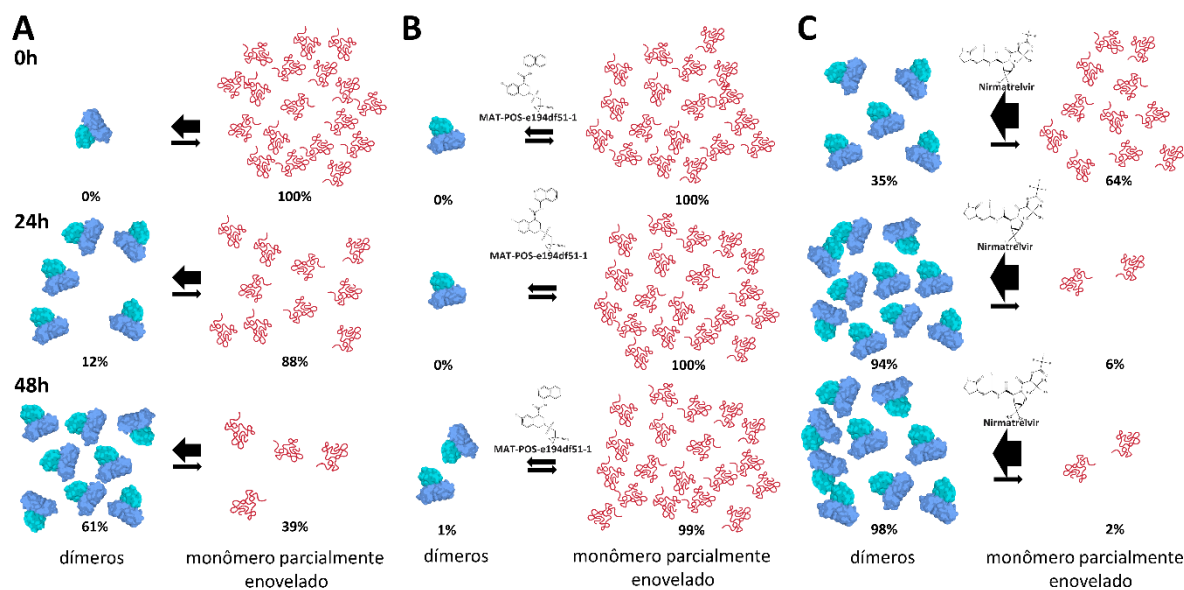


Figura 58 - Representação esquemática da dinâmica em solução observada no SEC-MALS ao longo do tempo para a amostra 1, contendo C145S M^{PRO} monômero. De cima para baixo estão apresentados respectivamente os tempos 0h, 24h e 48h. A) Controle DMSO. B) Amostra contendo o inibidor não covalente MAT-POS-e194df51-1. C) Amostra contendo o inibidor covalente nirmatrelvir.

Fonte: NOSKE *et al.*¹³⁷

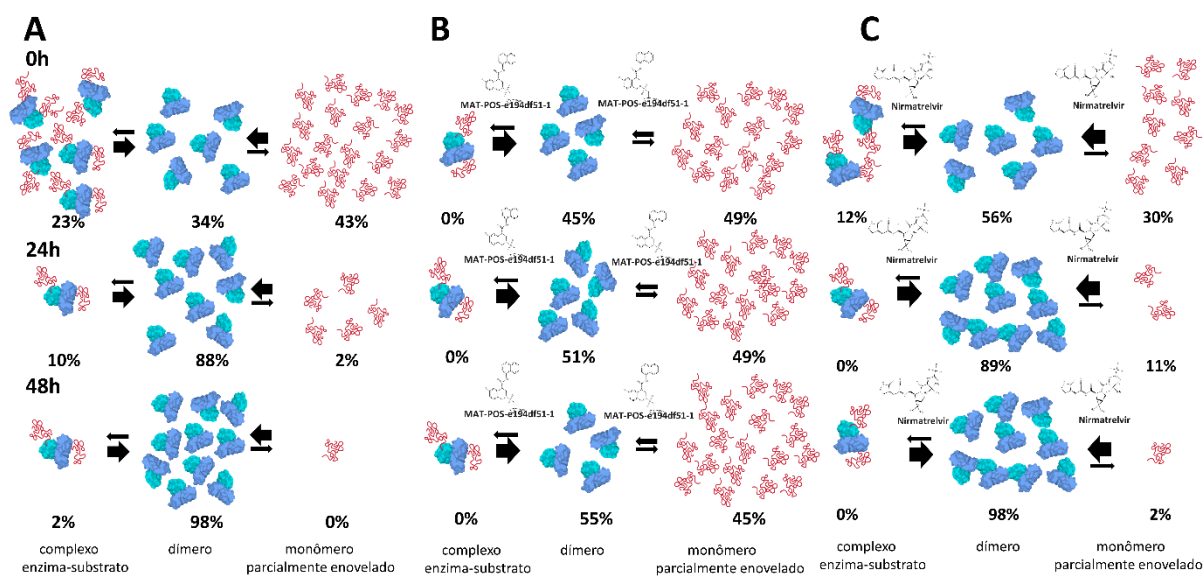


Figura 59 - Representação esquemática da dinâmica em solução observada no SEC-MALS ao longo do tempo para a amostra 1, contendo C145S M^{PRO} tetrâmero. De cima para baixo estão apresentados respectivamente os tempos 0h, 24h e 48h. A) Controle DMSO. B) Amostra contendo o inibidor não covalente MAT-POS-e194df51-1. C) Amostra contendo o inibidor covalente nirmatrelvir.

Fonte: NOSKE *et al.*¹³⁷

3.11 Estrutura cristalográfica da enzima C145S M^{pro} (monomérica) em complexo com nirmatrelvir

Para confirmar a hipótese de que a clivagem do N-terminal não é necessária para dimerização e ativação da enzima, cristalizamos a C145S M^{pro} monomérica na presença de Nirmatrelvir. Os cristais obtidos da condição contendo 100 mM MES pH 6.7, 8% PEG 4K e 5% DMSO (utilizando *seed* da M^{pro} ortorrômbica), foram criados protegidos utilizando 30% PEG 400 e resultaram na obtenção sua estrutura cristalográfica numa resolução de 1.74Å, com R_{work}/R_{free} de 0.19/0.23 e 96.35% dos resíduos em região favorável no gráfico de Ramachandran (Tabela 16).¹³⁷

Tabela 16 - Coleta de dados e estatísticas do refinamento referentes a enzima C145S M^{pro} em complexo com nirmatrelvir. Os dados entre parênteses correspondem à última camada de resolução.

C145S M^{pro} em complexo com Nirmatrelvir	
<i>Coleta de dados</i>	Linha de luz MANACA
Grupo espacial	<i>P</i> ₂ <i>1</i> ₂ <i>1</i>
Dimensões da cela unitária	
<i>a, b, c</i> (Å)	67.64, 101.96, 103.62
<i>α, β, γ</i> (°)	90.0, 90.0, 90.0
Resolução (Å)	72.75 - 1.74
Reflexões únicas	53189 (267)
Multiplicidade	
Completeza (%)	71.5 (3.6)
<i>I</i> / <i>σI</i>	8.9 (1.4)
<i>R</i> _{p.i.m.} (%) ¹	0.07 (0.476)
<i>CC</i> _{1/2} ²	0.995 (0.556)
<i>Refinamento</i>	
<i>R</i> _{work} / <i>R</i> _{free} ³	0.193/0.234
Número de átomos	5489
Água	713
Ligante	70
Proteína	4706
RMS (ligações) (Å)	0.0138
RMS (ângulos) (°)	1.80
Ramachandran favoráveis (%)	96.35
Ramachandran <i>outliers</i> (%)	0
Clashscore ⁴	5.97
<i>B</i> -factores médios (Å ²)	26.31
Macromoléculas	24.92
Ligante	22.62
Solvente	37.23
Código PDB	8EYJ

Fonte: NOSKE *et al.* ¹³⁷

A estrutura cristalina revelou que apesar da mutação C145S, o nirmatrelvir está ligado covalentemente a proteína. Além disso, também observamos o deslocamento das hélices α F, α H e α I do Domínio III adjacente, que parecem ser causados pelo impedimento estérico dos aminoácidos adicionais no N-terminal não clivados (Figura 60), similarmente ao observado para a estrutura da IMT- M^{pro} , com exceção de que não observamos o deslocamento dos resíduos do sítio ativo Phe140, Glu166, Pro168 e Gln189. Isto sugere que a ligação covalente permite a remodelagem adequada do sítio ativo, e o evento de dimerização mesmo na ausência da clivagem do N-terminal. Desta forma, podemos inferir que, para evitar a dimerização, um inibidor deve ser não covalente e inibir a ligação da M^{pro} monomérica aos onze locais de reconhecimento e clivagem na poli proteína viral.¹³⁷

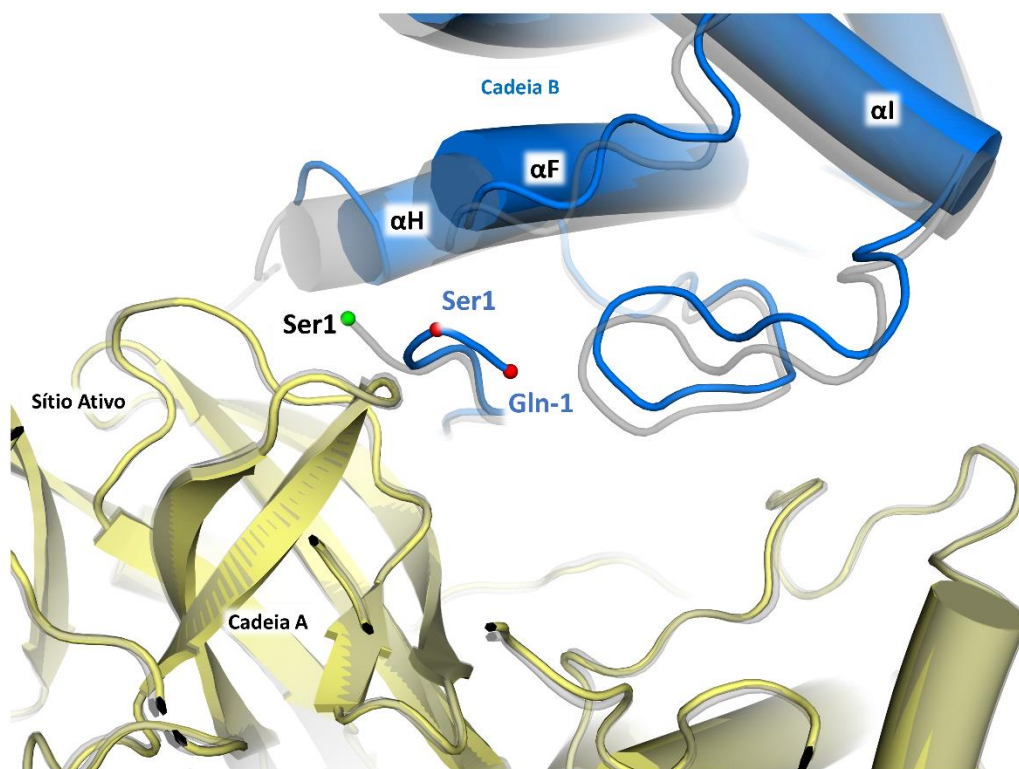


Figura 60 - Estrutura cristalográfica da C145S M^{pro} monomérica em complexo com nirmatrelvir, alinhada com a estrutura da enzima apo madura. Ambas as proteínas estão representadas como *cartoon*. A estrutura da C145S M^{pro} e da M^{pro} estão coloridas em amarelo e azul (para as cadeias A e B) e cinza, respectivamente.

Fonte: NOSKE *et al.*¹³⁷

3.12 Bases moleculares do processo de auto clivagem da M^{pro}

Se combinarmos todas as estruturas obtidas, a estrutura obtida por cryo-EM do mutante C145S ligado ao peptídeo não clivado com a estrutura de raios X da C145S (PDB 7N6N)

cadeias A (peptídeo clivado ligado covalentemente a Ser145) e cadeia B (peptídeo pós clivagem), podemos vislumbrar as modificações estruturais que ocorrem com a M^{pro} durante todas as etapas do processamento enzimático (Figura 61).

Estas estruturas revelam que o processo enzimático começa com resíduos fundamentais para a clivagem posicionados de maneira semelhante à estrutura apo (Figura 61A), mas a fita β 166-171 precisa ser deslocada para acomodar os peptídeos substrato P3-P5 (Figura 61B). Em seguida, durante o complexo intermediário enzima-substrato, a fita β 166-171 é ainda mais deslocada para permitir o encurtamento da distância entre C/S145 e a carbonila C, resultando na formação da ligação covalente e ligação do substrato ao sítio ativo, enquanto as hélices 43-53 e 62-65 exibem uma conformação mais aberta (Figura 61C). Após a clivagem, o produto peptídico permanece ligado ao sítio ativo, porém de maneira não covalente, permitindo que os resíduos do bolsão catalítico retornem a sua conformação apo (Figura 61D). Durante este processo, a superfície do sítio ativo da M^{pro} sofre modificações estruturais e de potencial eletrostático significativas para acomodar todas as conformações intermediárias do complexo enzima-substrato, detalhado nesta série de estruturas (Figura 61).

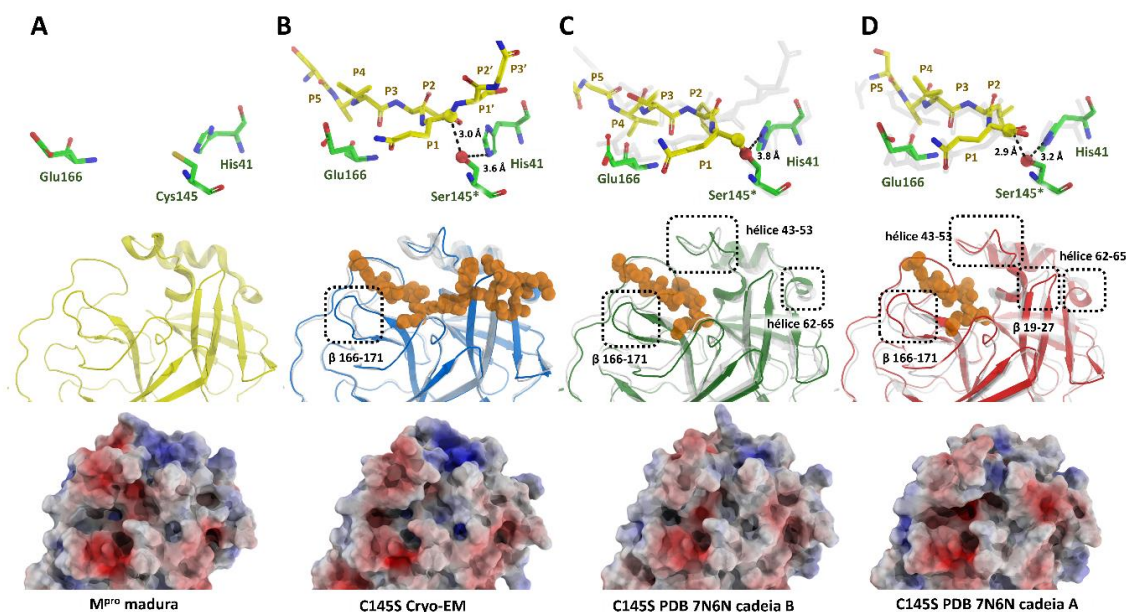


Figura 61 - Mudanças estruturais no sítio ativo da enzima M^{pro} durante o processo de clivagem. A) Principais resíduos do sítio ativo (representados em *sticks* verdes) da M^{pro} em forma apo (topo), representação em *cartoon* colorido em amarelo (meio) superfície de potencial eletrostático (parte inferior). B) Principais resíduos do sítio ativo (*sticks* verdes) da C145S M^{pro} (cryo-EM) ligada ao peptídeo não clivado (topo), representação em *cartoon* colorido em azul (meio) e superfície de potencial eletrostático (parte inferior). C) Principais resíduos do sítio ativo (*sticks* verdes) da C145S M^{pro} ligada covalentemente ao peptídeo (complexo intermediário enzima-substrato) (topo), representação em *cartoon* colorido em verde (meio) e superfície de potencial eletrostático (parte inferior). D) Principais resíduos do sítio ativo (*sticks* verdes) da C145S M^{pro} ligada não covalentemente ao peptídeo (topo), representação em *cartoon* colorido em vermelho (meio) e superfície de potencial eletrostático (parte inferior). Os *sticks* transparentes e *cartoon* coloridos em cinza nas figuras do topo e central representam a posição estrutural dos elementos na etapa anterior.

3.13 Modelo proposto para o auto processamento (maturação) da enzima M^{pro} e implicações para o desenvolvimento de antivirais

Neste trabalho, determinamos diversas estruturas de raios-X de diferentes construções da M^{pro} e realizamos a caracterização em solução destas. Em conjunto com a estrutura obtida por cryo-EM, oferecem informações essenciais para a compreensão de mecanismo de maturação da enzima. Também mostramos que a clivagem do N-terminal da M^{pro} é importante para o enovelamento adequado da proteína, mas não é crítico para a dimerização. Nossos dados sugerem que o evento de dimerização é governado pelo ajuste induzido da proteína durante a formação de uma ligação covalente durante o processo de maturação. Isto provavelmente significa que a ligação em qualquer um dos onze sítios de reconhecimento da enzima na poli proteína serviriam como gatilho para a dimerização, independentemente da clivagem ou não do N-terminal.¹³⁷

Essas observações divergem de modelos propostos anteriormente, nos quais a clivagem cis de N e C-terminais entre dois protômeros são necessariamente a etapa inicial de o processo de maturação.^{122,124} Estudos anteriores mostraram que a dimerização induzida por substrato pode ocorrer em um modelo in vitro da poli proteína viral, onde N e C-terminal estão ligados a outras proteínas.¹⁵⁷ Nosso modelo é consistente com fenômeno descrito para MERS-CoV M^{pro}, em que a dimerização do MERS-CoV M^{pro} é desencadeada pela presença de substrato.

Desta forma, propusemos um mecanismo para a maturação da enzima M^{pro}, que tem início com sua forma monomérica imatura, a qual é capaz de formar um dímero intermediário para clivagem da porção N-terminal, desencadeado pela ligação covalente do substrato, que em teoria também poderia se ligar a qualquer outro sítio da poli proteína viral. Em seguida, a associação de dois dímeros ainda imaturos, forma um complexo tetramérico para clivagem da porção C-terminal, resultando por fim na liberação da enzima na sua forma nativa.^{125,137} Uma representação esquemática deste processo pode ser observada na Figura 62.

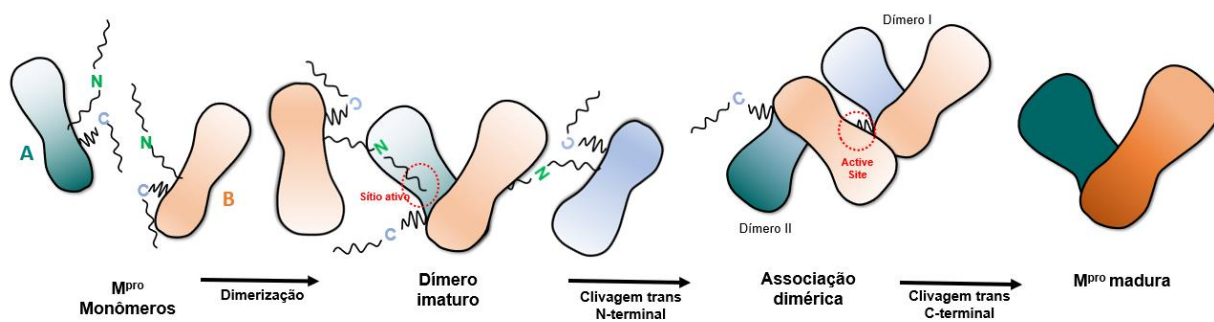


Figura 62 - Representação esquemática do processo de maturação da enzima M^{PRO} de SARS-CoV-2. Cada um dos monômeros está colorido respectivamente de verde e laranja.

Fonte: Adaptada de NOSKE *et al.*¹²⁵

4 CONCLUSÕES E PERSPECTIVAS

A M^{pro} é uma enzima dimérica, fundamental para o ciclo de replicação do SARS-CoV-2, e juntamente com características únicas de seu sítio de ligação ao substrato, consiste em um alvo amplamente estudado e explorado para o desenvolvimento de antivirais. Embora a enzima seja amplamente estudada, ainda existem diversas perguntas com relação ao mecanismo de maturação da proteína. Para elucidar este processo, neste trabalho realizamos a clonagem, expressão e purificação de diferentes construções da enzima: uma forma madura (M^{pro}), uma forma imatura (IMT-M^{pro}) e um mutante inativo (C145S M^{pro}). Além disso, desenvolvemos um método de purificação baseado em precipitação por sulfato de amônio, que permite a obtenção da enzima pura protocolo mais simples e rápido que os previamente utilizados para purificação da M^{pro}.

As proteínas purificadas foram utilizadas para cristalização e elucidação estrutural e caracterização por diversas técnicas em solução como SEC-MALS, ensaios de atividade baseados em FRET e espectrometria de massas nativa. Adicionalmente, determinamos a estrutura da M^{pro} C145S por crio microscopia eletrônica numa resolução de 3.5Å. A estrutura cristalográfica da enzima IMT-M^{pro} exibiu diversas mudanças estruturais comparada à enzima nativa, especialmente no sítio de ligação ao substrato. Além disso, a forma imatura apresentou atividade catalítica reduzida e revelou ser monomérica em solução, em contraste com a forma madura. Determinamos também a estrutura cristalográfica da C145S M^{pro}, a qual foi resolvida em complexo com ambos os peptídeos naturais de reconhecimento da enzima do N e C-terminal. Esta construção revelou alta dinamicidade em solução, com estados oligoméricos variando entre monómeros e tetrameros. A estrutura da C145S M^{pro} tetramérica determinada por cryo-EM, relevou detalhes importantes de outra etapa da maturação, a clivagem do N-terminal, estando em complexo com o peptídeo de reconhecimento da enzima, ainda não clivado. Experimentos de espectrometria de massas nativa relevaram a presença de oligômeros formados entre partículas clivadas e não clivadas. Isto, juntamente com dados de SEC-MALS, que mostraram que inibidores não covalentes impedem a dimerização, enquanto covalentes induzem a dimerização, nos levaram a inferir que a clivagem do N-terminal não é crítica para a dimerização da enzima, mas sim, as mudanças conformacionais são ocasionadas pelo ajuste induzido após ligação covalente do substrato.

A elucidação das mudanças funcionais e estruturais que ocorrem durante esse processo de maturação obtidas neste trabalho fornecem informações importantes para a compreensão deste mecanismo e para a proposição de inibidores específicos, que tenham como alvo etapas

intermediárias do processo de maturação da enzima. Além disso, este trabalho nos ajuda a ter um maior entendimento sobre a replicação do SARS-CoV-2 e suporta o desenvolvimento de novos antivirais contra o vírus tendo como alvo a M^{pro}.

CONCLUSÕES GERAIS E PERSPECTIVAS

Neste trabalho, avançamos com a caracterização estrutural e em solução das proteases virais: NS2B-NS3^{pro} do vírus da febre amarela e M^{pro} do vírus SARS-CoV-2. Para a NS2B-NS3^{pro}, determinamos sua estrutura cristalográfica, e ainda estabelecemos um ensaio de atividade baseado em fluorescência, utilizado para a triagem de inibidores. Para a M^{pro} de SARS-CoV-2, foram obtidas estruturas cristalográficas e de crio microscopia eletrônica para diferentes construções da enzima. Realizamos também a caracterização em solução da proteína. As informações obtidas, em conjunto, foram utilizadas para propor um mecanismo de maturação para a M^{pro}. Portanto, este trabalho nos ajuda a ter um maior entendimento sobre as proteases dos vírus da febre amarela e SARS-CoV-2, e suporta o desenvolvimento de novos antivirais, permitindo assim o desenvolvimento de ferramentas para o controle e prevenção de futuras epidemias.

TRABALHOS PUBLICADOS REFERENTES AOS TRABALHOS DESCRITOS NESTA TESE

NOSKE, G. D.; GAWRILJUK, V. O.; FERNANDES, R. S.; FURTADO, N. D.; BONALDO, M. C.; OLIVA, G.; GODOY, A. S. Structural characterization and polymorphism analysis of the NS2B-NS3 protease from the 2017 Brazilian circulating strain of Yellow Fever virus. **Biochimica et Biophysica Acta - General Subjects**, v. 1864, n. 4, 129521, 2020. <https://doi.org/10.1016/j.bbagen.2020.129521>.

NOSKE, G. D.; NAKAMURA, A. M.; GAWRILJUK, V. O.; FERNANDES, R. S.; LIMA, G. M. A.; ROSA, H. V. D.; PEREIRA, H. D.; ZERI, A. C. M; NASCIMENTO A. F. Z., FREIRE, M. C. L. C.; FEARON, D.; DOUANGAMATH, A.; von DELFT, F.; OLIVA, G.; GODOY, A. S. A Crystallographic Snapshot of SARS-CoV-2 Main Protease Maturation Process. **Journal of Molecular Biology**, 167118, 2021. <https://doi.org/10.1016/j.jmb.2021.167118>.

NOSKE, G. D.; SONG, Y.; FERNANDES, R. S.; CHALK, R.; ELMASSOUDI, H.; KOEKEMOER, L.; OWEN, D.C.; EL-BABA, T. J.; ROBINSON, C. V; Covid Moonshot Consortium; OLIVA, G.; GODOY, A. S. An in-solution snapshot of SARS-COV-2 main protease maturation process and inhibition. **Nature Communications**, 2023 14:1, 14(1), 1–13. <https://doi.org/10.1038/s41467-023-37035-5>.

OUTROS TRABALHOS PUBLICADOS PELA AUTORA DURANTE O PERÍODO DO DOUTORADO

Durante a realização deste trabalho, ainda tive oportunidade de realizar outros trabalhos, como primeira autora e em colaboração, além dos descritos nesta tese, nas áreas de biologia estrutural e busca por inibidores, os quais tiveram suma importância para minha formação como pesquisadora. Visando manter a tese de forma sucinta, optei por não incluir nesta tese todos os trabalhos realizados, pois não havia espaço para descrevê-los de maneira detalhada e adequada.

Como exemplos, participei como co-autora em alguns trabalhos envolvendo a enzima NS2B-NS3^{pro} de ZIKV Durante este período, participei em colaboração em outros três artigos envolvendo o complexo NS2B-NS3 protease de ZIKV. O ensaio de atividade padronizado neste projeto para YFV, foi adaptado para a protease de ZIKV, e este método faz parte dos protocolos descritos no artigo intitulado *High-throughput antiviral assays to screen for inhibitors of Zika virus replication* submetido pelo nosso grupo à revista *JOVE – Journal of Visualized Experiments*, Também participei como co-autora no artigo intitulado *Discovery of an imidazonaphthyridine and a riminophenazine as potent anti-Zika virus agents through a replicon-based high-throughput screening*, da autora Dra. Rafaela S. Fernandes, pós doutoranda do nosso grupo, que descreve dois inibidores da replicação de ZIKV triados

utilizando sistema baseado em replicon, publicado na revista *Virus Research*. Já o artigo intitulado *Natural products from Angelica keiskei inhibit key Zika virus replication proteins e Discovery of New Zika Protease and Polymerase Inhibitors through the Open Science Collaboration Project OpenZika*, foram elaborado em colaboração com a Profa. Dra. Ana Carolina Gomes Jardim (Universidade Federal de Uberlândia – UFU, Minas Gerais) e é referente a inibição da replicação do ZIKV por flavonoides provenientes da planta *Angelica keiskei* e do projeto colaborativo OpenZika, tendo como alvo as enzimas NS2B-NS3 protease e NS5 RdRp.

Com as enzimas M^{pro} e PL^{pro} de SARS-CoV-2, realizei além da expressão, purificação e caracterização, foram padronizados ensaios de atividade baseados em FRET. Estes ensaios já foram utilizados para triagem de diversas bibliotecas de inibidores provenientes de colaboradores do CIBFar e de colaboradores externos, como o Medicines for Malaria Venture (MMV) e da Collaborations Pharmaceuticals, Inc. – Carolina do Norte, Estados Unidos. No total já foi avaliada a atividade de mais de 10.000 compostos para cada uma das proteínas. Dentre eles, já foram obtidos diversos hits promissores com IC50 na faixa baixo micromolar e nanomolar. Estes resultados foram publicados nos artigos intitulados *Pyronaridine Protects against SARS-CoV-2 Infection in Mouse* e *Discovery of PLpro and Mpro Inhibitors for SARS-CoV-2* e outros ainda se encontram em fase de escrita. O consórcio internacional *COVID-Moonshot* do qual faço parte, mencionado nesta tese, resultou em algumas publicações, sendo que a principal delas, o artigo intitulado *COVID Moonshot: Open Science Discovery of SARS-CoV-2 Main Protease Inhibitors by Combining Crowdsourcing, High-Throughput Experiments, Computational Simulations, and Machine Learning* está disponível como *Biorxiv*, e encontra-se aceite para publicação na revista *Science*.

Ainda, trabalhei com a análise de mutações da enzima M^{pro} e como estas afetam a inibição dos compostos nirmatrelvir e ensitrelvir. Foi realizada a caracterização dos mutantes in-vitro e obtidas as estruturas cristalográficas em complexo com os inibidores. Estes resultados foram publicados no artigo intitulado *Structural basis of nirmatrelvir and ensitrelvir activity against SARS-CoV-2 Main Protease naturally occurring polymorphisms*, na revista *Journal of Biological Chemistry*.

Participei também no projeto de *fragment screening* da proteína NS3 helicase de ZIKV, no qual tive oportunidade de passar duas semanas no Diamond Light Source, Reino Unido, onde recebi treinamento para utilização da *facility*, e fui capaz de realizar o *soaking* de mais de 1000 fragmentos e a coleta de dados de difração destes cristais.

Por fim, tive a oportunidade de realizar projeto BEPE-FAPESP, no *Centre for Structural Systems Biology* (CSSB), Desy, em Hamburgo-Alemanha, sob orientação do prof. Dr. Kay Grünewald. Durante o período de um ano, trabalhei com o complexo enzimático Terminase, do Citomegalovírus humano (HCMV), e fui capaz de realizar todo o processo de preparação de amostra, coleta e processamento de dados, de crio microscopia eletrônica. Todos os esforços realizados, em conjunto com um aluno de doutorado do grupo do prof. Dr. Kay Grünewald, resultaram na obtenção de diversas estruturas da enzima em alta resolução, incluindo a forma apo, e em complexo com ATP- γ -S e o inibidor Letermovir. Além disso, tive a oportunidade de realizar alguns experimentos referentes ao complexo NS4B-NS3 helicase de ZIKV, incluindo crio microscopia e crio tomografia. Após retorno ao Brasil, fui capaz de, em conjunto com o pós doutorando Andre S. Godoy, de obter algumas grades do complexo NS4B-NS3 helicase, as primeiras amostras preparadas no recentemente adquirido Vitrobot no IFSC, as quais serão utilizadas para realização de coleta de dados em alta resolução.

NOSKE, G.D.; SILVA, E.S.; GODOY, M.O.; DOLCI, I.; FERNANDES, R.S.; GUIDO, R.V.C.; SJO, P.; OLIVA, G.; GODOY, A.S. Structural basis of nirmatrelvir and ensitrelvir activity against SARS-CoV-2 Main Protease naturally occurring polymorphisms, **Journal of Biological Chemistry**, 2023. doi: <https://doi.org/10.1016/j.jbc.2023.103004>.

FERNANDES, R. S.; GODOY, A. S.; DOS SANTOS I. A.; **NOSKE, G. D.**; OLIVEIRA, K. I .Z.; GAWRILJUK, V. O.; GOMES A. C. G. and OLIVA, G. Discovery of an imidazonaphthyridine and a riminophenazine as potent anti-Zika virus agents through a replicon-based high-throughput screening. **Virus Research**, v.299, 198388, 2021. <https://doi.org/10.1016/j.virusres.2021.198388>.

FREIRE, M. C. L. C. **NOSKE, G. D.**; BITENCOURT, N. V.; SANCHES, P. R. S.; SANTOS-FILHO, N. A.; GAWRILJUK, V. O.; SOUZA, E. P.; NOGUEIRA, V. H. R.; GODOY, M. O.; NAKAMURA, A. M.; FERNANDES, R. S.; GODOY, A. S.; JULIANO, M. A. PERES, B. M.; BARBOSA, C. G.; MORAES, C. B.; FREITAS-JUNIOR, L. H. G.; CILLI, E. M.; GUIDO, R. V. C.; OLIVA, G. Non-Toxic Dimeric Peptides Derived from the Bothropstoxin-I Are Potent SARS-CoV-2 and Papain-like Protease Inhibitors. **Molecules**, 26 (16), 4896, 2021. <https://doi.org/10.3390/molecules26164896>.

ZAIMAN, D.; GEHRTZ, P.; FILEP, M; FEARON, D.; GABIZON, G.; DOUANGAMATH, A.; PRILUSKY, j.; DUBERSTEIN, S.; COHEN, G.; OWEN, C. D.; RESNICK; E.; STRAIN-DAMERELL, C.; LUKACIK, P.; **Covid-Moonshot Consortium**; BARR, H.; WALSH, M. A.; von DELFT, F.; LONDON, N.; An automatic pipeline for the design of irreversible derivatives identifies a potent SARS-CoV-2 Mpro inhibitor, **Cell Chemical Biology**, 28, 1–12, 2021. <https://doi.org/10.1016/j.chembiol.2021.05.018>.

JANSEN, J., REIMER, K.C., NAGAI, J.S., VARGHESE. F.S., OVERHEUL, G.J., DE BEER, M., ROVERTS, R., DAVIRAN, D., FERMIN, LAS., WILLEMSSEN, B., BEUKENBOOM, M., DJUDJAJ, S., von STILLFRIED, S., EIJK, L.E., MASTIK, M., BULTHUIS, M., DUNNEN, W.D., van GOOR, H., HILLEBRANDS, J.L., TRIANA, S.H., ALEXANDROV, T., TIMM, M.C., BERGE, B.T., BROEK,

M., NLANDU, Q., HEIJNERT, J., BINDELS, E.M.J., HOOGENBOEZEM, R.M., MOOREN, F., KUPPE, C., MIESEN, P., GRÜNBERG, K., IJZERMANS, T., STEENBERGEN, E.J., CZOGALLA, J., SCHREUDER, M.F., SOMMERDIJK, N., AKIVA, A., BOOR, P., PUELLES, V.G., FLOEGE, J., HUBER, T.B.; **COVID Moonshot consortium**; RIJ, R.P., COSTA, I.G., SCHNEIDER, R.K., SMEETS, B., KRAMANN, R. SARS-CoV-2 infects the human kidney and drives fibrosis in kidney organoids. *Cell Stem Cell*, 2022 Feb 3;29(2):217-231.e8. doi: 10.1016/j.stem.2021.12.010.

FERNANDES, R.S.; **NOSKE, G.D.**; GAWRILJUK, V.O.; OLIVEIRA, K.I.Z.; GODOY, A.S.; MESQUITA, N.C.M.R.; OLIVA, G. High-throughput Antiviral Assays to Screen for Inhibitors of Zika Virus Replication. *Journal of Visualized Experiments*, 176, e62422, 2021. doi:10.3791/62422.

MOTTIN, M.; CAESAR, L.K.; BRODSKY, D.; MESQUITA, N. C. M. R.; OLIVEIRA, K.I.Z.; **NOSKE, G.D.**; SOUSA, B. K. P.; RAMOS, P. R. P. S.; JARMER, H.; LOH, B.; ZORN, K. M.; FOIL, D. H.; TORRES, P. M.; GUIDO, R. V. C.; OLIVA, G.; SCHOLLE, F.; EKINS, S.; CECH, N. B.; ANDRADE, C. H.; LASTER, S. M. Chalcones from *Angelica keiskei* (ashitaba) inhibit key Zika virus replication proteins, *Bioorganic Chemistry*, 2022. doi: <https://doi.org/10.1016/j.bioorg.2022.105649>.

PUHL, A.C.; GOMES, G.F.; DAMASCENO, S.; GODOY, A.S.; **NOSKE, G.D.**; NAKAMURA, A.M.; GAWRILJUK, V.O.; FERNANDES, R.S.; MONAKHOVA, N.; RIABOVA, O.; LANE, T.R.; MAKAROV, V.; VERAS, F.P.; BATAH, S.S.; FABRO, A.T.; OLIVA, G.; CUNHA, F.Q.; ALVES-FILHO, J.C., CUNHA, T.M., EKINS, S. Pyronaridine Protects against SARS-CoV-2 Infection in Mouse. *ACS Infect Dis*. 2022 Jun 10;8(6):1147-1160. doi: 10.1021/acsinfecdis.2c00091.

FURTADO, N.D.; de MELLO, I.S.; GODOY, A.S.; **NOSKE, G.D.**; OLIVA, G.; CANARD, B.; DECROLY, E.; BONALDO, M.C. Amino Acid Polymorphisms on the Brazilian Strain of Yellow Fever Virus Methyltransferase Are Related to the Host's Immune Evasion Mediated by Type I Interferon. *Viruses*, 2023 Jan 10;15:191. doi: 10.3390/v15010191.

MOTTIN, M.; SOUSA, B.K.P.; MESQUITA, N.C.M.R.; OLIVEIRA, K.I.Z.; **NOSKE, G.D.**; SARTORI, G.; ALBUQUERQUE, A.O.; URBINA, F.; PUHL, A.C.; MOREIRA-FILHO, J.T.; SOUZA, G.E.; GUIDO, R.V.C.; MURATOV, E.; NEVES, B.J.; DA SILVA, J.H.M.; CLARK, A.E.; SIQUEIRA-NETO, J.L.; PERRYMAN, A.L.; OLIVA, G.; EKINS, S.; ANDRADE, C.H. Discovery of New Zika Protease and Polymerase Inhibitors through the Open Science Collaboration Project OpenZika. *J Chem Inf Model*, 2022 Dec 26;62(24):6825-6843. doi: 10.1021/acs.jcim.2c00596.

GODOY, A. S.; NAKAMURA, A. M.; DOUANGAMATH, A. SONG, Y. **NOSKE, G.D.**, GAWRILJUK, V.O.; FERNANDES, R.S.; PEREIRA, H.D.M.; OLIVEIRA, K.I.Z.; FEARON, D.; DIAS, A.; KROJER, T.; FAIRHEAD, M.; POWELL, A.; DUNNET, L. BRANDAO-NETO, J.; SKYNER, R.; CHALK, R.; BAJUSZ, D.; BEGE, M.; BORBÁS, A.; KESERÚ, G.A.; VON DELFT, F.; OLIVA, G. Allosteric regulation and crystallographic fragment screening of SARS-CoV-2 NSP15 endoribonuclease, *Nucleic Acids Research*, Volume 51, Issue 10, 9 June 2023, Pages 5255–5270. doi: <https://doi.org/10.1093/nar/gkad314>.

PUHL, A.C.; GODOY, A.S.; NOSKE, G.D.; NAKAMURA, A.M.; GAWRILJUK, V.O.; FERNANDES, R.S.; OLIVA, G.; EKINS, S. Discovery of PLpro and Mpro Inhibitors for SARS-CoV-2. *ACS Omega*. 2023 Jun 14;8(25):22603-22612. doi: 10.1021/acsomega.3c01110. PMID: 37387790; PMCID: PMC10275482.

REFERÊNCIAS

- 1 LIPKIN, W. I.; ANTHONY, S. J. Virus hunting. **Virology**, v. 479–480, p. 194–199, 2015. DOI: 10.1016/j.virol.2015.02.006.
- 2 PEKOSZ, A.; GLASS, G. E. Emerging viral diseases. **Md Med**, v.9, n. 1, p. 11, 2008.
- 3 LECOQ, H. Discovery of the first virus, Tobacco mosaic virus: 1892 or 1898? **Comptes Rendus de l'Academie des Sciences: serie III**, v. 324, n. 10, p. 929–933, 2001.
- 4 SAGAYA JANSI, R. *et al.* Emerging paradigms of viral diseases and paramount role of natural resources as antiviral agents. **Science of the Total Environment**, v. 759, p. 143539, 2021. DOI: 10.1016/j.scitotenv.2020.143539.
- 5 WORLD HEALTH ORGANIZATION. **World health statistics 2023**. Disponível em: <https://www.who.int/data/gho/publications/world-health-statistics>. Acesso em 22 jun. 2023.
- 6 SILVA, N.I.O. *et al.* Recent sylvatic yellow fever virus transmission in Brazil: the news from an old disease. **Virology Journal**, v. 17, n. 9, 2020. DOI: 10.1186/s12985-019-1277-7
- 7 PIELNAA, P. *et al.* Zika virus-spread, epidemiology, genome, transmission cycle, clinical manifestation, associated challenges, vaccine and antiviral drug development. **Virology**, v. 543, p. 34–42, 2020. DOI: 10.1016/j.virol.2020.01.015.
- 8 COLTER, J. S.; ELLEM, K. A. O. Structure of viruses. **Annual Review of Microbiology**, v. 15, p. 219–244, 2003.
- 9 DOUAM, F.; PLOSS, A. Yellow fever virus: knowledge gaps impeding the fight against an old foe. **Trends in Microbiology**, v. 26, n. 11, p. 913–928, 2018.
- 10 COUTO-LIMA, D. *et al.* Potential risk of re-emergence of urban transmission of yellow fever virus in Brazil facilitated by competent Aedes populations. **Scientific Reports**, v. 7, n. 1, p. 4848, 2017.
- 11 TOMORI, O. Yellow fever: the recurring plague. **Critical Reviews in Clinical Laboratory Sciences**, v. 41, 2004. DOI: 10.1080/10408360490497474.
- 12 BARRETT, A. D. T.; HIGGS, S. Yellow fever: a disease that has yet to be conquered. **Annual Review of Entomology**, v. 52, p. 209–229, 2007. DOI: 10.1146/annurev.ento.52.110405.091454.
- 13 BRYAN, C. S.; MOSS, S. W.; KAHN, R. J. Yellow fever in the Americas. **Infectious Disease Clinics of North America**, v. 18, n. 2, p. 275–292, 2004.
- 14 STAPLES, J. E.; MONATH, T. P. Celebrating 125 years yellow fever: 100 years of discovery the etiology of yellow fever: an additional note. **JAMA**, v. 300, n. 8, p. 1–3, 2008.
- 15 MUTEBI, J. P.; BARRETT, A. D. T. The epidemiology of yellow fever in Africa. **Microbes & Infection**, v. 4, n. 14, p. 1459–1468, 2002.

16 WORLD HEALTH ORGANIZATION. **Countries with risk of yellow fever transmission and countries requiring yellow fever vaccination (July 2020)**. Disponível em: [https://www.who.int/publications/m/item/countries-with-risk-of-yellow-fever-transmission-and-countries-requiring-yellow-fever-vaccination-\(july-2020\)](https://www.who.int/publications/m/item/countries-with-risk-of-yellow-fever-transmission-and-countries-requiring-yellow-fever-vaccination-(july-2020)). Acesso em 10 de jul. 2023.

17 CDC. National Center for Health Statistics – homepage. **Transmission of yellow fever virus**. Disponível em: <https://www.cdc.gov/yellowfever/transmission/index.html>. Acesso em 10 de jul. 2023.

18 FERNANDO, P.; VASCONCELOS, D. C. Yellow fever in Brazil: thoughts and hypotheses on the emergence in previously free areas. **Revista de Saúde Pública**, v. 44, n. 6, p. 1144-1149, 2010.

19 CDC. National Center for Health Statistics: homepage. **Yellow fever - symptoms, diagnosis, & treatment**. Disponível em: <https://www.cdc.gov/yellowfever/symptoms/index.html>. Acesso em 19 jul. 2023.

20 GARDNER, C. L.; RYMAN, K. D. Yellow fever: a reemerging threat. **Clinics in Laboratory Medicine**, v. 30, n. 1, p. 237–260, 2010.

21 DOUAM, F. *et al.* Type III interferon-mediated signaling is critical for controlling live attenuated yellow fever virus infection in vivo. **mBio**, v. 8, n. 4, p. e00819, 2017.

22 WATSON, A. M.; KLIMSTRA, W. B. T cell-mediated immunity towards yellow fever virus and useful animal models. **Viruses**, v. 9, 2017. DOI: 10.3390/v9040077.

23 FRIERSON, J. G. The yellow fever vaccine: a history. **Yale Journal of Biology and Medicine**, v. 83, n. 2, p. 77–85, 2010.

24 MONATH, T. P. Review of the risks and benefits of yellow fever vaccination including some new analyses. **Expert Review of Vaccines**, v.11, p. 427–448, 2012. DOI: 10.1586/erv.12.6.

25 CDC. National Center for Health Statistics: homepage. **Yellow fever vaccine**. Disponível em: <https://www.cdc.gov/yellowfever/vaccine/index.html>. Acesso em 19 jul. 2023.

26 POSSAS, C. *et al.* Yellow fever outbreak in Brazil: the puzzle of rapid viral spread and challenges for immunisation. **Memórias do Instituto Oswaldo Cruz**, v. 113, p. 1–12, 2018. DOI: 10.1590/0074-02760180278.

27 DOUAM, F.; PLOSS, A. Yellow fever virus: knowledge gaps impeding the fight against an old foe. **Trends in Microbiology**, v. 26, n. 11, p. 913–928, 2018.

28 GOLDANI, L. Z. Yellow fever outbreak in Brazil, 2017. **Brazilian Journal of Infectious Diseases**, v. 21, n. 2, p. 123–124, 2017.

29 SÃO PAULO. Secretaria de Estado da Saúde. Centro de Vigilância Epidemiológica. **Situação epidemiológica**. 2018. Disponível em: <https://www.saude.sp.gov.br/resources/cve->

centro-de-vigilancia-epidemiologica/areas-de-vigilancia/doencas-de-transmissao-por-vetores-e-zoonoses/doc/famarela/fa18_boletim_epid_0312.pdf. Acesso em 01 de ago. 2023.

30 BRASIL. Ministério da Saúde. **Boletim epidemiológico da Febre Amarela no Brasil 2019/2020**. 2020. Disponível em: <http://www.rets.epsjv.fiocruz.br/biblioteca/boletim-epidemiologico-da-febre-amarela-no-brasil-20192020>. Acesso em 01 ago. 2023.

31 BRASIL. Ministério da Saúde do Brasil. **Monitoramento dos casos de arboviroses 2022-2023 junho/julho**. 2023. Disponível em: <https://www.gov.br/saude/pt-br/centrais-de-conteudo/publicacoes/boletins/epidemiologicos/edicoes/2023/boletim-epidemiologico-volume-54-no-01/>. Acesso em 01 ago. 2023.

32 ASSENBERG, R. *et al.* Crystal structure of a novel conformational state of the flavivirus NS3 protein: implications for polyprotein processing and viral replication. **Journal of Virology**, v. 83, p. 12895–12906, 2009. DOI:10.1128/JVI.00942-09.

33 KASSAR, T. C. *et al.* Construction and characterization of a recombinant yellow fever virus stably expressing *Gussia luciferase*. **Anais da Academia Brasileira de Ciências**, v. 89, Suppl. 3, p. 2119-2130, 2017.

34 LINDENBACH, B. D.; THIEL, H.-J.; RICE, C. M. Flaviviridae: the viruses and their replication. *In*: KNIPE, D.M.; HOWLEY, O.M. (ed.) **Fields virology**. 6th ed. Philadelphia: Lippincott-Raven, 2007

35 REKHA, M.; KUMAR, A.; MUBEENA, T.; GOPINATH, C. A short review on Zika Virus past, present strategies and future perspectives. **Pharma Tutor**, v. 4, n. 7, p. 21–24, 2016.

36 LUO, D. *et al.* Crystal structure of the NS3 protease-helicase from dengue virus. **Journal of Virology**, v. 82, n. 1, p. 173–183, 2008.

37 HAMMAMY, M. Z. *et al.* Development and characterization of new peptidomimetic inhibitors of the West Nile Virus NS2B-NS3 protease. **ChemMedChem**, v. 8, n. 2, p. 231–241, 2013.

38 Li, Y. *et al.* Structural insights into the inhibition of zika virus NS2B-NS3 protease by a small-molecule inhibitor. **Structure**, v. 26, n. 4, p. 555-564.e3, 2018.

39 LEI, J. *et al.* Crystal structure of Zika virus NS2B-NS3 protease in complex with a boronate inhibitor. **Science**, v. 353, n. 6298, p. 503-505, 2016.

40 HILGENFELD, R.; LEI, J.; ZHANG, L. The structure of the Zika Virus protease, NS2B/NS3pro. **Advances in Experimental Medicine and Biology**, v. 1062, p. 131–145, 2018. DOI: 10.1007/978-981-10-8727-1_10.

41 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. 7, p. 6832–6836, 2019.

- 42 BOLDESCU, V. *et al.* Broad-spectrum agents for flaviviral infections: dengue, Zika and beyond. **Nature Reviews Drug Discovery**, v. 16, n. 8, p. 565–586, 2017.
- 43 GODOY, A. S. *et al.* Structural and mechanistic insight from antiviral and antiparasitic enzyme drug targets for tropical infectious diseases. **Current Opinion in Structural Biology**, v. 59, p. 65, 2019. DOI: 10.1016/j.sbi.2019.02.014
- 44 NITSCHKE, C. Proteases from dengue, West Nile and Zika viruses as drug targets. **Biophysical Reviews**, v. 11, n. 2, p. 157–165, 2019.
- 45 COPELAND, R. A. **Enzymes** - a practical introduction to structure, mechanism and data analysis. 2nd ed. New York: Wiley-VCH, 2000.
- 46 NOBLE, C. G.; SEH, C. C.; CHAO, A. T.; SHI, P. Y. Ligand-bound structures of the dengue virus protease reveal the active conformation. **Journal of Virology**, v. 86, n. 1, p. 438–446, 2012.
- 47 SALAEMAE, W.; JUNAID, M.; ANGSUTHANASOMBAT, C.; KATZENMEIER, G. Structure-guided mutagenesis of active site residues in the dengue virus two-component protease NS2B-NS3. **Journal of Biomedical Science**, v. 68, n. 1, p.1–8, 2010.
- 48 BONALDO, M. C. *et al.* Genome analysis of yellow fever virus of the ongoing outbreak in Brazil reveals polymorphisms. **Memorias do Instituto Oswaldo Cruz**, v. 112, n. 6, p. 447–451, 2017.
- 49 MIR, D. *et al.* Phylodynamics of yellow fever virus in the Americas: new insights into the origin of the 2017 Brazilian outbreak. **Scientific Reports**, v. 7, p. 7385, 2017. DOI: 10.1038/s41598-017-07873-7.
- 50 GÓMEZ, M. M. *et al.* Genomic and structural features of the yellow fever virus from the 2016–2017 Brazilian outbreak. **Journal of General Virology**, v. 99, n. 4, p. 536–548, 2018.
- 51 FURTADO, N. D. *et al.* Amino acid polymorphisms on the Brazilian strain of yellow fever virus methyltransferase are related to the host's immune evasion mediated by type I interferon. **Viruses**, v. 15, n. 1, p. 191, 2023.
- 52 SAMRAT, S. K.; XU, J.; LI, Z.; ZHOU, J.; LI, H. Antiviral agents against flavivirus protease: prospect and future direction. **Pathogens**, v. 11, n. 3, p 293, 2022.
- 53 LEUNG, D. *et al.* Activity of recombinant dengue 2 virus NS3 protease in the presence of a truncated NS2B co-factor, small peptide substrates, and inhibitors. **Journal of Biological Chemistry**, v. 276, n. 49, p. 45762–45771, 2001.
- 54 RASSIAS, G. *et al.* Cell-active carbazole derivatives as inhibitors of the zika virus protease. **European Journal of Medicinal Chemistry**, v. 180, p. 536–542, 2019. DOI: 10.1016/j.ejmech.2019.07.007.
- 55 NIE, S. *et al.* Synthesis, structure-activity relationship and antiviral activity of indole-containing inhibitors of Flavivirus NS2B-NS3 protease. **European Journal of Medicinal Chemistry**, v. 225, p. 113767, 2021. DOI: 10.1016/j.ejmech.2021.113767.

- 56 LI, Z. *et al.* Existing drugs as broad-spectrum and potent inhibitors for Zika virus by targeting NS2B-NS3 interaction. **Cell Research**, v. 27, n. 8, p. 1046–1064, 2017.
- 57 NOSKE, G. D. *et al.* Structural characterization and polymorphism analysis of the NS2B-NS3 protease from the 2017 Brazilian circulating strain of Yellow Fever virus. **Biochimica et Biophysica Acta**: general subject, v. 1864, n. 4, p. 129521, 2020.
- 58 ASLANIDIS, C.; DE JONG, P. J. Ligation-independent cloning of PCR products (LIC-PCR). **Nucleic Acids Research**, v. 18, n. 20, p. 6069–6074, 1990.
- 59 GIBSON, D. G. Enzymatic assembly of overlapping DNA fragments. **Methods in Enzymology**, v. 498, 2011. DOI: 10.1016/B978-0-12-385120-8.00015-2.
- 60 KUIPER, B. D. *et al.* Increased activity of unlinked Zika virus NS2B/NS3 protease compared to linked Zika virus protease. **Biochemical and Biophysical Research Communications**, v. 492, n. 4, p. 668–673, 2017.
- 61 LI, Y. *et al.* Structural insights into the inhibition of Zika virus NS2B-NS3 Protease by a small-molecule inhibitor. **Structure**, v. 26, n. 4, p. 555-564.e3, 2018.
- 62 STUDIER, F. W. Protein production by auto-induction in high density shaking cultures. **Protein Expression and Purification**, v. 41, n. 1, p. 207–234, 2005.
- 63 KABSCH, W. XDS. **Acta Crystallographica Section D: biological crystallography**, v. 66, p. 125–132, 2010.
- 64 KARPLUS, P. A.; DIEDERICHS, K. Assessing and maximizing data quality in macromolecular crystallography. **Current Opinion in Structural Biology**, v. 34, p. 60–68, 2015. DOI: 10.1016/j.sbi.2015.07.003.
- 65 KARPLUS, P. A.; DIEDERICHS, K. Linking crystallographic model and data quality. **Science**, v. 336, n. 6084, p. 1030–1033, 2012.
- 66 WINN, M. D. *et al.* Overview of the CCP4 suite and current developments. **Acta Crystallographica Section D: biological crystallography**, v. 67, n. 4, p. 235–242, 2011.
- 67 ADAMS, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. **Acta Crystallographica Section D**, v. 66, p. 213–221, 2010. DOI: 10.1107/S0907444909052925.
- 68 MCCOY, A. J. *et al.* Phaser crystallographic software. **Journal of Applied Crystallography**, v. 40, p. 658–674, 2007. DOI: 10.1107/S0021889807021206.
- 69 TERWILLIGER, T. C. *et al.* Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. **Acta Crystallographica Section D**, v. 64, p. 61–69, 2008. DOI: 10.1107/S090744490705024X.
- 70 AFONINE, P. V *et al.* Towards automated crystallographic structure refinement with phenix.refine. **Acta Crystallographica Section D**, v. 68, p. 352–367, 2012. DOI: 10.1107/S0907444912001308.

- 71 AFONINE, P. V. *et al.* Real-space refinement in PHENIX for cryo-EM and crystallography. **Acta Crystallographica Section D**, v. 74, p. 531–544, 2018. DOI: 10.1107/S2059798318006551.
- 72 EMSLEY, P.; LOHKAMP, B.; SCOTT, W. G.; COWTAN, K. Features and development of Coot. **Acta Crystallographica Section D**, v. 66, p. 486–501, 2010. DOI: 10.1107/S0907444910007493.
- 73 EMSLEY, P.; COWTAN, K. Coot: model-building tools for molecular graphics. **Acta Crystallographica Section D**, v. 60, p. 2126–2132, 2004. DOI: 10.1107/S0907444904019158.
- 74 CHEN, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. **Acta Crystallographica Section D**, v. 66, p. 12–21, 2010. DOI: 10.1107/S0907444909042073.
- 75 LI, Y. *et al.* Structural characterization of the linked NS2B-NS3 protease of Zika virus. **FEBS Letters**, v. 591, p. 2338–2347, 2017. DOI: 10.1002/1873-3468.12741.
- 76 FERNANDES, R. S. *et al.* High-throughput antiviral assays to screen for inhibitors of Zika virus replication. **Journal of Visualized Experiments**, v. e62422, 2021. DOI: 10.3791/62422.
- 77 ZHANG, J. H.; CHUNG, T. D. Y.; OLDENBURG, K. R. A simple statistical parameter for use in evaluation and validation of high throughput screening assays. **Journal of Biomolecular Screening**, v. 4, p. 67–73, 1999. DOI: 10.1177/108705719900400206.
- 78 LI, J. Q. *et al.* Development of a replicon cell line-based high throughput antiviral assay for screening inhibitors of Zika virus. **Antiviral Research**, v. 150, p. 148–154, 2018. DOI: 10.1016/j.antiviral.2017.12.017.
- 79 SACHETTO, R. *et al.* Discovery of an imidazonaphthyridine and a riminophenazine as potent anti-Zika virus agents through a replicon-based high-throughput screening. **Virus Research**, v. 299, p. 198388, 2021. DOI: 10.1016/j.virusres.2021.198388.
- 80 KROGH, A.; LARSSON, È.; HEIJNE, G.; SONNHAMMER, E. L. L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. **Journal of Molecular Biology**, v. 19, p. 567-80, 2001. DOI: 10.1006/jmbi.2000.4315.
- 81 KIM, Y. M. *et al.* NMR analysis of a novel enzymatically active unlinked dengue NS2B-NS3 protease complex. **Journal of Biological Chemistry**, v. 288, n. 18, p. 12891–12900, 2013.
- 82 NITSCHE, C. Proteases from dengue, West Nile and Zika viruses as drug targets. **Biophysical Reviews**, v. 11, p. 157–165, 2019. DOI: 10.1007/s12551-019-00508-3.
- 83 PHOO, W. W. *et al.* Structure of the NS2B-NS3 protease from Zika virus after self-cleavage. **Nature Communications**, v. 7, p. 13410, 2016. DOI: 10.1038/ncomms13410.

- 84 ARAVAPALLI, S. *et al.* Inhibitors of Dengue virus and West Nile virus proteases based on the aminobenzamide scaffold. **Bioorganic & Medicinal Chemistry**, v. 20, p. 4140–4148, 2012. DOI: 10.1016/j.bmc.2012.04.055.
- 85 ASHKENAZY, H. *et al.* ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. **Nucleic Acids Research**, v. 44, p. 344–50, 2016. DOI: 10.1093/nar/gkw408.
- 86 GOUET, P.; COURCELLE, E.; STUART, D.; METOZ, F. ESPript: analysis of multiple sequence alignments in PostScript. **Bioinformatics**, v. 15, p. 305–308, 1999. DOI: 10.1093/bioinformatics/15.4.305.
- 87 NICK, M. *et al.* Screening the medicines for Malaria Venture (MMV) pandemic response box chemical library on *Caenorhabditis elegans* identifies re-profiled candidate anthelmintic drug leads. **Frontiers in Tropical Diseases**, v. 3, p. 98, 2022. DOI: 10.3389/fitd.2022.1017900.
- 88 KIM, H. S. *et al.* Antimicrobial effect of alexidine and chlorhexidine against *Enterococcus faecalis* infection. **International Journal of Oral Science**, v. 5, p. 26–31, 2013. DOI: 10.1038/ijos.2013.11.
- 89 MAMOUEI, Z. *et al.* Alexidine dihydrochloride has broad-spectrum activities against diverse fungal pathogens. **mSphere**, v. 31, p. e00539-18, 2018. DOI: 10.1128/mSphere.00539-18.
- 90 ALIZADEH, H.; NEELAM, S.; CAVANAGH, H. D. Amoebicidal activities of alexidine against 3 pathogenic strains of *acanthamoeba* NIH public access. **Eye Contact Lens**, v. 35, n. 1, p. 1–5, 2009.
- 91 READER, J. *et al.* Multistage and transmission-blocking targeted antimalarials discovered from the open-source MMV pandemic response box. **Nature Communications**, v. 12, n. 1, p. 269, 2021.
- 92 CHOI, R. *et al.* High-throughput screening of the ReFRAME, pandemic box, and COVID box drug repurposing libraries against SARS-CoV-2 nsp15 endoribonuclease to identify small-molecule inhibitors of viral activity. **PLoS One**, v. 16, n. 4, p. e0250019, 2021.
- 93 ZHU, W. *et al.* Antibacterial drug leads targeting isoprenoid biosynthesis. **Proceedings of the National Academy of Sciences of the United States of America**, v. 110, p. 123–128, 2013. DOI: 10.1073/pnas.1219899110.
- 94 LIM, W. *et al.* Screening the pandemic response box identified benzimidazole carbamates, Olorofim and ravuconazole as promising drug candidates for the treatment of eumycetoma. **PLOS Neglected Tropical Diseases**, v. 16, p. e0010159, 2022. DOI: 10.1371/journal.pntd.0010159.
- 95 NOGUEIRA, M. B. *et al.* GA-Hecate antiviral properties on HCV whole cycle represent a new antiviral class and open the door for the development of broad spectrum antivirals. **Scientific Reports**, v. 25, p. 14329, 2018. DOI: 10.1038/s41598-018-32176-w.

96 GRUBA, N. *et al.* Substrate profiling of Zika virus NS2B-NS3 protease. **FEBS Letters**, v. 590, p. 3459–3468, 2016. DOI: 10.1002/1873-3468.12443.

97 KONDO, M. Y. *et al.* Yellow fever virus NS2B/NS3 protease: hydrolytic properties and substrate specificity. **Biochemical and Biophysical Research Communications**, v. 407, p. 640–644, 2011. DOI: 10.1016/j.bbrc.2011.03.054

98 ZHU, N. A novel coronavirus from patients with pneumonia in China, 2019. **New England Journal of Medicine**, v. 382, p. 727–733, 2020. DOI: 10.1056/NEJMoa2001017

99 WU, F. *et al.* A new coronavirus associated with human respiratory disease in China. **Nature**, v. 579, n. 7798, p. 265–269, 2020.

100 WORLD HEALTH ORGANIZATION. **WHO director-general's opening remarks at the media briefing on COVID-19**. 11 mar. 2020. Disponível em: <https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020>. Acesso em 22 ago. 2023.

101 TREGONING, J. S.; FLIGHT, K. E.; HIGHAM, S. L.; WANG, Z.; PIERCE, B. F. Progress of the COVID-19 vaccine effort: viruses, vaccines and variants versus efficacy, effectiveness and escape. **Nature Reviews Immunology**, v. 21, p. 626–636, 2021. DOI: 10.1038/s41577-021-00592-1.

102 MISTRY, P. *et al.* SARS-CoV-2 variants, vaccines, and host immunity. **Frontiers in Immunology**, v. 12, 2021. DOI: 10.3389/fimmu.2021.809244.

103 WORLD HEALTH ORGANIZATION. **COVID-19 global situation (2023)**. Disponível em: <https://covid19.who.int/#:~:text=Globally%2C%20as%20of%2012%3A20pm,vaccine%20doses%20have%20been%20administered>. Acesso em 23 ago. 2023.

104 ZHU, Z. *et al.* From SARS and MERS to COVID-19: a brief summary and comparison of severe acute respiratory infections caused by three highly pathogenic human coronaviruses. **Respiratory Research**, v. 21, p. 1–14, 2020. DOI: 10.1186/s12931-020-01479-w.

105 ZHOU, P. *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. **Nature**, v. 579, p. 270–273, 2020. DOI: 10.1038/s41586-020-2012-7.

106 JIN, Z. *et al.* Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. **Nature**, v. 582, p. 289–293, 2020. DOI: 10.1038/s41586-020-2223-y.

107 WU, A. *et al.* Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China. **Cell Host & Microbe**, v. 27, p. 325–328, 2020. DOI: 10.1016/j.chom.2020.02.001

108 YAN, W.; ZHENG, Y.; ZENG, X.; HE, B.; CHENG, W. Structural biology of SARS-CoV-2: open the door for novel therapies. **Signal Transduction and Targeted Therapy**, v. 7, 2022. DOI: 10.1038/s41392-022-00884-5.

- 109 LEE, J. *et al.* Crystallographic structure of wild-type SARS-CoV-2 main protease acyl-enzyme intermediate with physiological C-terminal autoprocessing site. **Nature Communications**, v. 11, p. 1–9, 2020. DOI: 10.1038/s41467-020-19662-4
- 110 SINGH, D. D., PARVEEN, A.; YADAV, D. K. SARS-CoV-2: emergence of new variants and effectiveness of vaccines. **Frontiers in Cellular and Infection Microbiology**, v. 11, 2021. DOI: 10.3389/fcimb.2021.777212.
- 111 ARYA, R. *et al.* Structural insights into SARS-CoV-2 proteins. **Journal of Molecular Biology**, v. 433, p. 16672, 2021. DOI: 10.1016/j.jmb.2020.11.024.
- 112 SCHULLER, M. *et al.* SARS-CoV-2 identified through crystallographic screening and computational docking. **Science Advances**, v. 7, p. 25, 2021. DOI: 10.1126/sciadv.abf8711.
- 113 NEWMAN, J. A. *et al.* Structure, mechanism and crystallographic fragment screening of the SARS-CoV-2 NSP13 helicase. **Nature Communications**, v. 12, p. 1–11, 2021. DOI: 10.1038/s41467-021-25166-6.
- 114 GODOY, A. *et al.* Allosteric regulation and crystallographic fragment screening of SARS-CoV-2 NSP15 endoribonuclease. **Nucleic Acids Research**, v. 51, p. 5255-5270, 2023.
- 115 DOUANGAMATH, A. *et al.* Crystallographic and electrophilic fragment screening of the SARS-CoV-2 main protease. **Nature Communications**, v. 11, 2020. DOI: 10.1038/s41467-020-18709-w.
- 116 COVID Moonshot: open science discovery of SARS-CoV-2 main protease inhibitors by combining crowdsourcing, high-throughput experiments, computational simulations, and machine learning. DOI:10.1101/2020.10.29.339317.
- 117 OWEN, D. R. *et al.* An oral SARS-CoV-2 Mpro inhibitor clinical candidate for the treatment of COVID-19. **Science**, v. 374, p. 586-1593, 2021. DOI: 10.1126/science.abl4784.
- 118 UNOH, Y. *et al.* Discovery of S-217622, a noncovalent oral SARS-CoV-2 3CL protease inhibitor clinical candidate for treating COVID-19. **Journal of Medicinal Chemistry**, v. 65, p. 6499–6512, 2022. DOI: 10.1021/acs.jmedchem.2c00117.
- 119 JIN, Z. *et al.* Structural basis for the inhibition of SARS-CoV-2 main protease by antineoplastic drug carmofur. **Nature Structural & Molecular Biology**, v. 27, p. 529–532, 2020. DOI: 10.1038/s41594-020-0440-6.
- 120 TAO, K. *et al.* SARS-CoV-2 antiviral therapy. **Clinical Microbiology Reviews**, v. 34, 2021. DOI: 10.1128/CMR.00109-21.
- 121 XIA, B.; KANG, X. Activation and maturation of SARS-CoV main protease. **Protein and Cell**, v. 2, p. 282–290, 2011. DOI: 10.1007/s13238-011-1034-1.
- 122 HSU, M. F. *et al.* Mechanism of the maturation process of SARS-CoV 3CL protease. **Journal of Biological Chemistry**, v. 280, p. 31257–31266, 2005. DOI: 10.1074/jbc.M502577200.

- 123 LI, C. *et al.* Maturation mechanism of severe acute respiratory syndrome (SARS) coronavirus 3C-like proteinase. **Journal of Biological Chemistry**, v. 285, n. 36, p. 28134, 2010. DOI: 10.1074/jbc.M109.095851.
- 124 CHEN, S.; JONAS, F.; SHEN, C.; HIGENFELD, R. Liberation of SARS-CoV main protease from the viral polyprotein: N-terminal autocleavage does not depend on the mature dimerization mode. **Protein & Cell**, v. 1, p. 59–74, 2010. DOI: 10.1007/s13238-010-0011-4.
- 125 NOSKE, G. D. *et al.* A crystallographic snapshot of SARS-CoV-2 main protease maturation process. **Journal of Molecular Biology**, v. 433, p. 167118, 2021. DOI: 10.1016/j.jmb.2021.167118.
- 126 SÃO PAULO. Agência FAPESP. **Primeiro experimento realizado no SIRIUS busca desenvolver fármaco contra a COVID-19 (2020)**. Disponível em: <https://agencia.fapesp.br/primeiro-experimento-realizado-no-sirius-busca-desenvolver-farmaco-para-covid-19/34396/>. Acesso em 27 ago. 2023.
- 127 VONRHEIN, C. *et al.* Data processing and analysis with the autoPROC toolbox. **Acta Crystallographica Section D**, v. 67, p. 293–302, 2011. DOI: 10.1107/S0907444911007773.
- 128 SMART, O. S. *et al.* Exploiting structure similarity in refinement: Automated NCS and target-structure restraints in BUSTER. **Acta Crystallographica Section D**, v. 68, p. 368–380, 2012. DOI: 10.1107/S0907444911056058.
- 129 LIMA, G. M. A. *et al.* FragMAX: the fragment-screening platform at the MAX IV Laboratory. **Acta Crystallographica Section D**, v. 76, p. 771–777, 2020. DOI: 10.1107/S205979832000889X.
- 130 WOLLENHAUPT, J. *et al.* F2X-universal and F2X-entry: structurally diverse compound libraries for crystallographic fragment screening. **Structure**, v. 28, p. 694-706.e5, 2020. DOI: 10.1016/j.str.2020.04.019.
- 131 LIMA, G. M. A. *et al.* FragMAXapp: crystallographic fragment-screening data-analysis and project-management system. **Acta Crystallographica Section D**, 77, 2021. DOI: 10.1107/S2059798321003818.
- 132 MORIARTY, N. W.; GROSSE-KUNSTLEVE, R. W.; ADAMS, P. D. Electronic ligand builder and optimization workbench (eLBOW): a tool for ligand coordinate and restraint generation. **Acta Crystallographica Section D**, v. 65, p. 1074–1080, 2009. DOI: 10.1107/S0907444909029436.
- 133 VONRHEIN, C. *et al.* Advances in automated data analysis and processing within autoPROC, combined with improved characterisation, mitigation and visualisation of the anisotropy of diffraction limits using STARANISO. **Acta Crystallographica Section A**, v. 74, p. a360–a360, 2018.
- 134 WINTER, G. *et al.* DIALS: implementation and evaluation of a new integration package. **Acta Crystallographica Section D**, v. 74, p. 85–97, 2018. DOI: 10.1107/S2059798317017235

- 135 PEARCE, N. M. *et al.* A multi-crystal method for extracting obscured crystallographic states from conventionally uninterpretable electron density. **Nature Communications**, v. 8, p. 1–8, 2017. DOI: 10.1038/ncomms15123.
- 136 CHALK, R. Mass spectrometric analysis of proteins. *In*: BURGESS-BROWN, N. A. (ed.). **Heterologous gene expression in E. coli: methods and protocols**. New York: Humana Press, 2017. p. 373-395.
- 137 NOSKE, G. *et al.* An in-solution snapshot of SARS-COV-2 main protease maturation process and inhibition. **Nature Communications**, v. 14, p. 1–13, 2023. DOI: 10.1038/s41467-023-37035-5.
- 138 ZHENG, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. **Nature Methods**, v. 14, p. 331–332, 2017. DOI: 10.1038/nmeth.4193.
- 139 ZIVANOV, J.; NAKANE, T.; SCHERES, S.H.W. A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. **IUCrJ**, v. 6, p. 5–17, 2019. DOI: 10.1107/S205225251801463X.
- 140 ROHOU, A.; GRIGORIEFF, N. CTFFIND4: fast and accurate defocus estimation from electron micrographs. **Journal of Structural Biology**, v. 192, n. 2, p. 216–221, 2015.
- 141 WAGNER, T. *et al.* SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. **Communications Biology**, v. 2, p. 1–13, 2019. DOI: 10.1038/s42003-019-0437-z.
- 142 ZIVANOV, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in RELION-3. **eLife**, v. 7, 2018. DOI: 10.7554/eLife.42166.
- 143 ROSENTHAL, P. B.; HENDERSON, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. **Journal of Molecular Biology**, v. 333, n. 4, p. 721–745, 2003.
- 144 PUNJANI, A.; RUBINSTEIN, J. L.; FLEET, D. J.; BRUBAKER, M. A. CryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. **Nature Methods**, v. 14, p. 290–296, 2017. DOI: 10.1038/nmeth.4169.
- 145 PUNJANI, A.; ZHANG, H.; FLEET, D. J. Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. **Nature Methods**, v. 17, p. 1214–1221, 2020. DOI: 10.1038/s41592-020-00990-8.
- 146 ZI, Y.T. *et al.* Addressing preferred specimen orientation in single-particle cryo-EM through tilting. **Nature Methods**, v. 14, p. 793–796, 2017. DOI: 10.1038/nmeth.4347.
- 147 PETTERSEN, E. F. *et al.* UCSF ChimeraX: structure visualization for researchers, educators, and developers. **Protein Science**, v. 30, p. 70–82, 2021. DOI: 10.1002/pro.3943.

148 TERWILLIGER, T. C.; SOBOLEV, O. V.; AFONINE, P. V.; ADAMS, P. D. Automated map sharpening by maximization of detail and connectivity. **Acta Crystallographica Section D**, v. 74, p. 545–559, 2018. DOI: 10.1107/S2059798318004655

149 PUNJANI, A.; FLEET, D. J. 3D variability analysis: resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM. **Journal of Structural Biology**, v. 213, n. 2, 2021. DOI: 10.1016/j.jsb.2021.107702.

150 XUE, X. *et al.* Production of authentic SARS-CoV Mpro with enhanced activity: application as a novel Tag-cleavage endopeptidase for protein overproduction. **Journal of Molecular Biology**, v. 366, p. 965–975, 2007. DOI: 10.1016/j.jmb.2006.11.073.

151 MA, C. *et al.* Boceprevir, GC-376, and calpain inhibitors II, XII inhibit SARS-CoV-2 viral replication by targeting the viral main protease. **Cell Research**, v. 30, n. 8, p. 678–692, 2020.

152 KRISSINEL, E.; HENRICK, K. Inference of macromolecular assemblies from crystalline state. **Journal of Molecular Biology**, v. 372, n. 3, p. 774–797, 2007.

153 RUT, W. *et al.* SARS-CoV-2 Mpro inhibitors and activity-based probes for patient-sample imaging. **Nature Chemical Biology**, v. 17, p. 222–228, 2020. DOI: 10.1038/s41589-020-00689-z.

154 KNELLER, D. W. *et al.* Structural plasticity of SARS-CoV-2 3CL Mpro active site cavity revealed by room temperature X-ray crystallography. **Nature Communications**, v. 11, p. 7–12, 2020. DOI: 10.1038/s41467-020-16954-7.

155 LEE, J. *et al.* Crystallographic structure of wild-type SARS-CoV-2 main protease acyl-enzyme intermediate with physiological C-terminal autoprocessing site. **Nature Communications**, v. 11, n. 1, p. 5877, 2020. DOI: 10.1038/s41467-020-19662-4.

156 CONSTANT, D. A.; MATEO, R.; NAGAMINE, C. M.; KIRKEGAARD, K. Targeting intramolecular proteinase NS2B/3 cleavages for trans-dominant inhibition of dengue virus. **Proceedings of the National Academy of Sciences of the United States of America**, v. 115, n. 40, p. 10136–10141, 2018.

157 HO, B. L. *et al.* Critical assessment of the important residues involved in the dimerization and catalysis of MERS Coronavirus Main Protease. **PLoS One**, v. 10, p. 1–18, 2015. DOI: 10.1371/journal.pone.0144865.

ANEXOS

A. ARTIGOS PUBLICADOS REFERENTES AO TRABALHO DESCRITO NESTA TESE

BBA - General Subjects 1864 (2020) 129521



Contents lists available at ScienceDirect

BBA - General Subjects

journal homepage: www.elsevier.com/locate/bbagen

Structural characterization and polymorphism analysis of the NS2B-NS3 protease from the 2017 Brazilian circulating strain of Yellow Fever virus



Gabriela Dias Noske^a, Victor Oliveira Gawriljuk^a, Rafaela Sachetto Fernandes^a,
Nathalia Dias Furtado^b, Myrna Cristina Bonaldo^b, Glaucius Oliva^a, Andre Schutzer Godoy^{a,*}

^a Physics Institute of Sao Carlos, University of Sao Paulo, Av. Joao Dagnone, 1100 - Jardim Santa Angelina, São Carlos 13563-120, Brazil

^b Laboratório de Biologia Molecular de Flavivírus, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil

ARTICLE INFO

Keywords:
Yellow Fever virus
Flavivirus
Protease
NS3
NS2B

ABSTRACT

Background: The Yellow Fever virus (YFV) is transmitted by mosquitos and causes an infection with symptoms including fever, headaches and nausea. In 20–50% of the cases, the disease may evolve to a visceral stage, reaching high mortality rates. YFV NS2B-NS3 protease has been identified as an important drug target.

Methods: Herein, we describe the crystal structure of the NS2B-NS3 protease from the 2017 YFV Brazilian circulating strain using X-ray crystallography. Furthermore, we used a combination of biochemical and biophysical assays to characterize the enzyme and investigate the impact of the polymorphisms observed in different YFV circulating strains.

Results: Surprisingly, the crystal structure of YFV protease seems to adopt the closed conformation without the presence of a binding partner. Although D88E and K121R mutants exhibited a lower affinity for the substrate, both revealed to be more processive, resulting in a similar catalytic efficiency in relation to the WT protease. Still, both mutants showed an accentuated decrease in stability when compared with the WT.

Conclusions: The crystal structure of YFV NS2B-NS3 in closed conformation might be an important tool for the development of new drugs, as well as understanding the activation mechanism of viral proteases. Biochemical analyses indicate that the NS2B-NS3 protease of the circulating strain of YFV is more stable than previous strains.

General significance: The YFV NS2B-NS3 protease is the first flaviviral structure described in its closed conformation when in a free form, implying that external factors might induce the activation of the enzyme.

1. Introduction

Yellow fever (YF) is one of the most severe mosquito-borne infections in African and South American countries, responsible for 29,000–60,000 deaths annually [1,2]. The 2016/2019 YF outbreak in Brazil is the most severe one in nearly eight decades, already resulting in more than 2000 cases of infection with 757 confirmed deaths [3]. In South America, the Yellow Fever virus (YFV) is transmitted by *Aedes aegypti* (urban cycle) and *Haemagogus* spp. (sylvatic cycle), causing the disease with a large spectrum of symptoms, ranging from mild to a severe deadly hemorrhagic fever [2,4,5]. The rapid spread of YFV to Southeastern Brazil and the increasing numbers of cases in rural areas

near densely populated cities constitute a potential risk of urban YF re-emergence [6,7]. To date, there is no antiviral treatment for YFV and prevention strategy relies on vaccination of people living or travelling to risk areas [8]. Despite the highly effectiveness and long-lasting immunity provided by a single dose, the use of 17D live-attenuated vaccine in low risk populations is not indicated due to the potential occurrence of severe adverse effects [8,9]. Therefore, the constant resurgence of YF in endemic countries is a major public health concern.

YFV is a member of the genus *Flavivirus* within the family *Flaviviridae*. The flavivirus genome is a single-stranded positive-sense RNA of ~11 kb in length consisting of a 5' untranslated region (UTR), a single long open reading frame (ORF), and a 3' UTR. The long ORF

Abbreviations: AMC, 4-amine-7-methylcoumarin; AU, asymmetric unit; DENV, Dengue virus; DSF, differential scanning fluorimetry; DTT, dithiothreitol; IMAC, immobilized metal chromatography; IPTG, β-D-1-thiogalactopyranoside; LIC, ligation independent cloning; NS, nonstructural protein; OD₆₀₀, optical density of 600 nm; ORF, open reading frame; PCR, polymerase chain reaction; PDB, Protein Data Bank; PNK, T4 Polynucleotide Kinase; RMSD, root mean square deviation; SUMO, Small Ubiquitin-like Modifier; TB, terrific broth; TEV, Tobacco Etch Virus; UTR, untranslated region; WNV, West Nile virus; WT, wild type; YFV, Yellow Fever virus; YF, yellow fever; ZIKV, Zika virus

* Corresponding author.

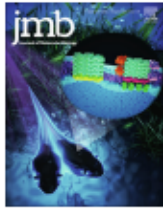
E-mail address: andregodoy@ifsc.usp.br (A.S. Godoy).

<https://doi.org/10.1016/j.bbagen.2020.129521>

Received 8 September 2019; Received in revised form 3 January 2020; Accepted 9 January 2020

Available online 10 January 2020

0304-4165/ © 2020 Elsevier B.V. All rights reserved.



A Crystallographic Snapshot of SARS-CoV-2 Main Protease Maturation Process

G. D. Noske^{1†}, A. M. Nakamura^{1†}, V. O. Gawriljuk¹, R. S. Fernandes¹,
G. M. A. Lima^{2†}, H. V. D. Rosa¹, H. D. Pereira¹, A. C. M. Zeri³, A. F. Z. Nascimento³,
M. C. L. C. Freire¹, D. Fearon^{4,5}, A. Douangamath^{4,5}, F. von Delft^{4,5,6,7},
G. Oliva^{1*} and A. S. Godoy^{1*}

1 - Institute of Physics of Sao Carlos, University of Sao Paulo, Av. Joao Dagnone, 1100, Jardim Santa Angelina, Sao Carlos 13563-120, Brazil

2 - BioMAX, MAX IV Laboratory, Fotogatan 2, Lund 224 84, Sweden

3 - Brazilian Synchrotron Light Laboratory (LNLS), Brazilian Center for Research in Energy and Materials (CNPEM), Zip Code 13083-970 Campinas, Sao Paulo, Brazil

4 - Diamond Light Source Ltd., Harwell Science and Innovation Campus, Didcot OX11 0QX, UK

5 - Research Complex at Harwell, Harwell Science and Innovation Campus, Didcot OX11 0FA, UK

6 - Centre for Medicines Discovery, University of Oxford, Old Road Campus, Roosevelt Drive, Headington OX3 7DQ, UK

7 - Department of Biochemistry, University of Johannesburg, Auckland Park 2006, South Africa

Correspondence to G. Oliva and A.S. Godoy: oliva@ifsc.usp.br (G. Oliva), andregodoy@ifsc.usp.br (A.S. Godoy)
<https://doi.org/10.1016/j.jmb.2021.167118>

Edited by Eric O. Freed

Abstract

SARS-CoV-2 is the causative agent of COVID-19. The dimeric form of the viral M^{pro} is responsible for the cleavage of the viral polyprotein in 11 sites, including its own N and C-terminus. The lack of structural information for intermediary forms of M^{pro} is a setback for the understanding its self-maturation process. Herein, we used X-ray crystallography combined with biochemical data to characterize multiple forms of SARS-CoV-2 M^{pro}. For the immature form, we show that extra N-terminal residues caused conformational changes in the positioning of domain-three over the active site, hampering the dimerization and diminishing its activity. We propose that this form precludes the cis and trans-cleavage of N-terminal residues. Using fragment screening, we probe new cavities in this form which can be used to guide therapeutic development. Furthermore, we characterized a serine site-directed mutant of the M^{pro} bound to its endogenous N and C-terminal residues during dimeric association stage of the maturation process. We suggest this form is a transitional state during the C-terminal trans-cleavage. This data sheds light in the structural modifications of the SARS-CoV-2 main protease during its self-maturation process.

© 2021 Elsevier Ltd. All rights reserved.

Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the causative agent of COVID-19, a highly infectious disease that rapidly spreads causing a global pandemic. SARS-CoV-2 is an enveloped RNA virus belonging to the β -lineage of

coronaviruses, which includes SARS-CoV and Middle East (MERS-CoV) respiratory viruses.^{1–3} The viral genome is a single-stranded positive RNA comprising about 30,000 nucleotides, that shares 82% sequence identity with SARS-CoV.⁴ The replicase gene (ORF1ab) encodes two over-



An in-solution snapshot of SARS-CoV-2 main protease maturation process and inhibition

Received: 1 August 2022

Accepted: 28 February 2023

Published online: 20 March 2023

Check for updates

Gabriela Dias Noske^{1,6}, Yun Song^{2,6}, Rafaela Sachetto Fernandes¹, Rod Chalk³, Haitem Elmassoudi³, Lizbé Koekemoer³, C. David Owen², Tarick J. El-Baba^{4,5}, Carol V. Robinson^{4,5}, The COVID Moonshot Consortium*, Glaucius Oliva¹ & Andre Schutzer Godoy¹✉

The main protease from SARS-CoV-2 (M^{pro}) is responsible for cleavage of the viral polyprotein. M^{pro} self-processing is called maturation, and it is crucial for enzyme dimerization and activity. Here we use C145S M^{pro} to study the structure and dynamics of N-terminal cleavage in solution. Native mass spectroscopy analysis shows that mixed oligomeric states are composed of cleaved and uncleaved particles, indicating that N-terminal processing is not critical for dimerization. A 3.5 Å cryo-EM structure provides details of M^{pro} N-terminal cleavage outside the constrains of crystal environment. We show that different classes of inhibitors shift the balance between oligomeric states. While non-covalent inhibitor MAT-POS-e194df51-1 prevents dimerization, the covalent inhibitor nirmatrelvir induces the conversion of monomers into dimers, even with intact N-termini. Our data indicates that the M^{pro} dimerization is triggered by induced fit due to covalent linkage during substrate processing rather than the N-terminal processing.

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the causative agent of COVID-19¹. Like SARS-CoV and Middle East Respiratory Syndrome coronavirus (MERS-CoV), SARS-CoV-2 is single-strand RNA virus (ssRNA) that belongs to the genera of the beta-coronaviruses^{1,2}. The SARS-CoV-2 genome is composed of nearly 30,000 nucleotides which contains the *ORF1ab* gene, a large open-reading-frame (ORF) responsible for encoding 16 non-structural proteins (nsp's) as two polyproteins after ribosomal frameshifting, named 1a and 1b³. Proteolytic processing of viral polyproteins is essential for the viral life cycle, and it's performed by two SARS-CoV-2 encoded cysteine proteases: the papain-like protease (PL^{pro}), which is one of the domains of nsp3, and the viral main protease (M^{pro} or $3CL^{pro}$), encoded by the nsp5^{4,5}. PL^{pro} cleaves the viral polyprotein at three sites, while M^{pro} is responsible for cleaving the polyprotein at eleven distinct sites^{4,5}, including its own N- and C-termini⁶. M^{pro} is one of the most

promising targets for drug development against SARS-CoV-2^{5,7}. All the gathered structural and biochemical information on this target has been crucial for the rapid development of new antivirals^{8–11}, including the recent drug Paxlovid/nirmatrelvir¹⁰, approved for emergency use in USA, Europe and China.

To obtain heterologous expressed mature M^{pro} , researchers adopted the general strategy of adding nsp4 C-terminal portion to the N-terminal portion of nsp5 constructs, allowing the self-cleavage and dimerization of M^{pro} in-vitro⁶. Previously, we described the activity and biochemical profile of SARS-CoV-2 M^{pro} C145S mutant containing the C-terminal portion of nsp4 at its N-termini⁶. This serine mutation generated an active but much slower version of M^{pro} , a valuable tool for studying the biochemical aspects of this enzyme. Differently than the dimeric M^{pro} , this sample behaved as a dynamic mix of monomers, dimers, trimers and tetramers in solution⁶. The residual activity of this

¹Sao Carlos Institute of Physics, University of Sao Paulo, Av. Joao Dagnone, 1100 - Jardim Santa Angelina, Sao Carlos 13563-120, Brazil. ²Electron Bio-imaging Centre, Diamond Light Source Ltd., Harwell Science and Innovation Campus, Didcot OX11 0QX, UK. ³Centre for Medicines Discovery, Oxford University, OX1 3QU Oxford, UK. ⁴Physical and Theoretical Chemistry Laboratory, Department of Chemistry, University of Oxford, South Parks Road, OX1 3TA Oxford, UK. ⁵The Kavli Institute for Nanoscience Discovery, Dorothy Crowfoot Hodgkin Building, South Parks Road, OX1 3QU Oxford, UK. ⁶These authors contributed equally: Gabriela Dias Noske, Yun Song. *A list of authors and their affiliations appears at the end of the paper. ✉ e-mail: andregodoy@ifsc.usp.br



Structural basis of nirmatrelvir and ensitrelvir activity against naturally occurring polymorphisms of the SARS-CoV-2 main protease

Received for publication, October 13, 2022, and in revised form, December 15, 2022. Published, Papers in Press, February 10, 2023.
<https://doi.org/10.1016/j.jbc.2023.103004>

Gabriela Dias Noske¹, Ellen de Souza Silva¹, Mariana Ortiz de Godoy¹, Isabela Dolci¹,
 Rafaela Sachetto Fernandes¹, Rafael Victório Carvalho Guido¹, Peter Sjö², Glaucius Oliva^{1,*}, and
 Andre Schutzer Godoy^{1,*}

From the ¹Sao Carlos Institute of Physics, University of Sao Paulo, Sao Carlos, Brazil; ²Drugs for Neglected Diseases Initiative (DNDI), Geneva, Switzerland

Reviewed by members of the JBC Editorial Board. Edited by Craig Cameron

SARS-CoV-2 is the causative agent of COVID-19. The main viral protease (M^{pro}) is an attractive target for antivirals. The clinically approved drug nirmatrelvir and the clinical candidate ensitrelvir have so far showed great potential for treatment of viral infection. However, the broad use of antivirals is often associated with resistance generation. Herein, we enzymatically characterized 14 naturally occurring M^{pro} polymorphisms that are close to the binding site of these antivirals. Nirmatrelvir retained its potency against most polymorphisms tested, while mutants G143S and Q189K were associated with diminished inhibition constants. For ensitrelvir, diminished inhibition constants were observed for polymorphisms M49I, G143S, and R188S, but not for Q189K, suggesting a distinct resistance profile between inhibitors. In addition, the crystal structures of selected polymorphisms revealed interactions that were critical for loss of potency. In conclusion, our data will assist the monitoring of potential resistant strains, support the design of combined therapy, as well as assist the development of the next generation of M^{pro} inhibitors.

SARS-CoV-2 is a highly transmissible β -coronavirus (1, 2) with a genome composed of a single RNA positive strand that comprises about 30 kb encoding for 16 nonstructural, four structural, and six accessory proteins (3). The viral replicase codifies two frame shifting open reading frames, ORF1a/ORF1ab, containing 16 nonstructural proteins required for viral replication (4). The SARS-CoV-2 Main protease (M^{pro}) or 3C-like protease ($3CL^{pro}$) is a dimeric cysteine protease responsible for the cleavage of the viral polyproteins 1a and 1ab in 11 sites, including its own N and C terminal (5–7). M^{pro} substrate recognition has unique features and is specific for Gln residue at P1, hydrophobic residues at P2, and small side chains such as Ser and Ala at P1' (8). The absence of similar sites in human proteases together with the importance of the enzyme for the viral replication makes M^{pro} a primary target for antiviral discovery and development.

Several small molecules were identified as M^{pro} inhibitors that exhibited efficacy in cellular culture including boceprevir, carmofur, MAT-POS-e194df51-1, PF-07321332 (nirmatrelvir), and S-217622 (ensitrelvir) (6, 9–14). The first oral COVID-19 antiviral from Pfizer, Paxlovid, is a combination of reversible covalent M^{pro} inhibitor nirmatrelvir and ritonavir, a CYP3A4 inhibitor, with safety and efficacy demonstrated in clinical trials, and approved for use by the US Food and Drug Administration in December 2021 (14, 15). In addition, the compound ensitrelvir from Shionogi is a promising non-covalent inhibitor of M^{pro} (11). Currently in phase 3 clinical trials, the compound has shown exciting pharmacokinetics properties, with potential for therapeutic doses to be reached without requirement of CYP inhibitors (11).

Although most polymorphisms are not expected to generate a variant of concern (16), World Health Organization is constantly monitoring the emergence of SARS-CoV-2 mutations, since recent variants have exhibited more transmissible and infectious properties and can affect vaccines effectiveness (17, 18). Moreover, amino acid replacements in the viral protein can impact the catalytic activity of the enzyme and modify the efficacy of inhibitors (19, 20). The *in vitro* effectiveness of nirmatrelvir against variants Alpha (B.1.1.7), Beta (B.1.351), Delta (B.1.617.2), Gamma (P.1), Lambda (B.1.1.1.37/C37), and Omicron (B.1.1.529) has already been demonstrated (21). However, none these variants contain mutations that are in the vicinity of the active site. Herein, we evaluated the effect of single mutations in the active site of M^{pro} from circulating polymorphisms on the kinetics and inhibitory constants of nirmatrelvir and ensitrelvir. We also used X-ray crystallography to characterize the structural features of selected polymorphisms. These findings provided key information for predicting and avoiding resistance, designing the next generation of inhibitors, and raised important considerations for combination therapies.

Results

For this study, we selected active site polymorphic versions of M^{pro} that have been already identified in circulation. For

* For correspondence: Andre Schutzer Godoy, andregodoy@ifsc.usp.br; Glaucius Oliva, oliva@ifsc.usp.br.



B. ARTIGOS PUBLICADOS EM COLABORAÇÃO

bioRxiv preprint doi: <https://doi.org/10.1101/2020.10.29.339317>; this version posted October 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

COVID Moonshot: Open Science Discovery of SARS-CoV-2 Main Protease Inhibitors by Combining Crowdsourcing, High-Throughput Experiments, Computational Simulations, and Machine Learning

The COVID Moonshot Consortium¹

¹ Full list of authors: <https://tinyurl.com/y3r7redd>

***For correspondence:**

john.chodera@choderalab.org (John Chodera); alpha.lee@postera.ai (Alpha Lee); nir.london@weizmann.ac.il (Nir London); frank.von-delft@damond.ac.uk (Frank von Delft)

Abstract Herein we provide a living summary of the data generated during the COVID Moonshot project focused on the development of SARS-CoV-2 main protease (Mpro) inhibitors. Our approach uniquely combines crowdsourced medicinal chemistry insights with high throughput crystallography, exascale computational chemistry infrastructure for simulations, and machine learning in triaging designs and predicting synthetic routes. This manuscript describes our methodologies leading to both covalent and non-covalent inhibitors displaying protease IC50 values under 150 nM and viral inhibition under 5 μ M in multiple different viral replication assays. Furthermore, we provide over 200 crystal structures of fragment-like and lead-like molecules in complex with the main protease. Over 1000 synthesized and ordered compounds are also reported with the corresponding activity in Mpro enzymatic assays using two different experimental setups. The data referenced in this document will be continually updated to reflect the current experimental progress of the COVID Moonshot project, and serves as a citable reference for ensuing publications. All of the generated data is open to other researchers who may find it of use.

Since early 2020, when the coronavirus SARS-CoV-2 was implicated as the cause of the pandemic soon to engulf the world, the need for targeted therapeutic development has been evident due to the dearth of existing antiviral treatments. Realizing this need, researchers worked rapidly to determine the structure of the novel SARS-CoV-2 main protease (MPro) [1], a promising antiviral drug target involved in SARS-CoV-2 viral replication. Other researchers quickly followed with a crystallographic and electrophilic fragment-screen of the structure [2].

Recognizing the necessity for a large-scale, international effort to further develop antiviral compounds, the COVID Moonshot project was created to build on these early results and advance the development of Mpro inhibitors [3]. The project brought together academic and industrial partners from across the world



Contents lists available at ScienceDirect

Virus Research

journal homepage: www.elsevier.com/locate/virusres



Discovery of an imidazonaphthyridine and a riminophenazine as potent anti-Zika virus agents through a replicon-based high-throughput screening

Rafaela Sachetto Fernandes^{a,*}, Andre Schutzer de Godoy^a, Igor Andrade Santos^b, Gabriela Dias Noske^a, Kettlyn Irene Zagato de Oliveira^a, Victor Oliveira Gawriljuk^a, Ana Carolina Gomes Jardim^b, Glaucius Oliva^a

^a Physics Institute of Sao Carlos, University of Sao Paulo, Av. Joao Dagnone, 1100, Jardim Santa Angelina, São Carlos, 13563-120, Brazil

^b Institute of Biomedical Sciences, Federal University of Uberlândia, Av. Amazonas, 1700, bloco 4C sala 216, Umuarama, Uberlândia, 38405-317, Brazil

ARTICLE INFO

Keywords:

Zika
Replicon
Drug discovery
Antiviral screening
RO8191
Clofazimine

ABSTRACT

The 2015/16 Zika virus (ZIKV) epidemic led to almost 1 million confirmed cases in 84 countries and was associated to the development of congenital microcephaly and Guillain-Barré syndrome. More recently, a ZIKV African lineage was identified in Brazil raising concerns about a future outbreak. The long-term consequences of viral infection emphasizes the need for the development of effective anti-ZIKV drugs. In this study, we developed and characterized a ZIKV replicon cell line for the screening of viral replication inhibitors. The replicon system was developed by engineering the IRES-Neo cassette into the 3' UTR terminus of the ZIKV Rluc DNA construct. After *in vitro* transcription, replicon RNA was used to transfect BHK-21 cells, that were selected with G418, thus generating the BHK-21-RepZIKV_IRES-Neo cell line. Through this replicon-based cell system, we identified two molecules with potent anti-ZIKV activities, an imidazonaphthyridine and a riminophenazine, both from the MMV/DNDi Pandemic Response Box library of 400 drug-like compounds. The imidazonaphthyridine, known as RO8191, showed remarkable selectivity against ZIKV, while the riminophenazine, the antibiotic Clofazimine, could act as a non-nucleoside analog inhibitor of viral RNA-dependent RNA polymerase (RdRp), as evidenced both *in vitro* and *in silico*. The data showed herein supports the use of replicon-based assays in high-throughput screening format as a biosafe and reliable tool for antiviral drug discovery.

1. Introduction

The Zika virus (ZIKV) is a mosquito-borne virus member of the genus *Flavivirus*, within the family *Flaviviridae*. Since its first isolation in Uganda, in 1947, ZIKV has caused sporadic infections in tropical Africa and the south of Asia, mostly associated with mild symptoms (Baud et al., 2017). However, the number of ZIKV infection cases unexpectedly increased in 2007, initially in Micronesia, then in Pacific Ocean Islands, finally reaching South America in 2015 (Ferraris et al., 2019). In Brazil, the viral outbreak attracted international attention, especially due to its association with microcephaly and congenital malformations in newborns of ZIKV infected pregnant women (de Araújo et al., 2018) and Guillain-Barré syndrome in adults (Cao-Lormeau et al., 2016). Although the number of reported cases drastically dropped after 2017, the potential long-term consequences of viral infection, as well as the emergence in 2019 of a ZIKV African lineage Uganda subtype that has not

been previously reported in Brazil (Kasprzykowski et al., 2020), highlight the urgent need to develop effective antiviral therapies to aid in the treatment and control of future outbreaks (Loe et al., 2019).

Like other flaviviruses, ZIKV possesses a single-stranded positive-sense RNA genome in a small enveloped virion. The genome, of approximately 11 kb in length, consists of a 5' untranslated region (5' UTR), a large open reading frame (ORF) and a 3' UTR. The ORF is translated into a polyprotein further processed by cellular and viral proteases to produce three structural proteins: capsid (C), precursor membrane (prM) and envelop (E); and seven nonstructural (NS) proteins: NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5. The structural proteins form the viral particle while the NS proteins participate in viral genome replication, virion assembly and evasion of the host immune response (Chambers et al., 1990).

A biosafe and reliable strategy to screen large compound libraries for their antiviral activity is the use of replicon-based assays. Replicons are

* Corresponding author.

E-mail address: rafaela.fernandes@usp.br (R.S. Fernandes).

<https://doi.org/10.1016/j.virusres.2021.198388>

Received 6 January 2021; Received in revised form 19 February 2021; Accepted 6 March 2021

Available online 20 April 2021

0168-1702/© 2021 Elsevier B.V. All rights reserved.

High-throughput Antiviral Assays to Screen for Inhibitors of Zika Virus Replication

Rafaela S. Fernandes¹, Gabriela D. Noske¹, Victor O. Gawriljuk¹, Kettlyn I. Z. de Oliveira¹, Andre S. Godoy¹, Nathalya C. M. R. Mesquita¹, Glaucius Oliva¹

¹ São Carlos Institute of Physics, University of Sao Paulo

Corresponding Author

Glaucius Oliva
oliva@ifsc.usp.br

Citation

Fernandes, R.S., Noske, G.D., Gawriljuk, V.O., de Oliveira, K.I.Z., Godoy, A.S., Mesquita, N.C.M.R., Oliva, G. High-throughput Antiviral Assays to Screen for Inhibitors of Zika Virus Replication. *J. Vis. Exp.* (176), e62422, doi:10.3791/62422 (2021).

Date Published

October 30, 2021

DOI

10.3791/62422

URL

jove.com/video/62422

Abstract

Antiviral drug discovery requires the development of reliable biochemical and cellular assays that can be performed in high-throughput screening (HTS) formats. The flavivirus non-structural (NS) proteins are thought to co-translationally assemble on the endoplasmic reticulum (ER) membranes, forming the replication complex (RC). The NS3 and NS5 are the most studied enzymes of the RC and constitute the main targets for drug development due to their crucial roles in viral genome replication. NS3 protease domain, which requires NS2B as its cofactor, is responsible for the cleavage of the immature viral polyprotein into the mature NS proteins, whereas NS5 RdRp domain is responsible for the RNA replication. Herein, we describe in detail the protocols used in replicon-based screenings and enzymatic assays to test large compound libraries for inhibitors of the Zika virus (ZIKV) replication. Replicons are self-replicating subgenomic systems expressed in mammalian cells, in which the viral structural genes are replaced by a reporter gene. The inhibitory effects of compounds on viral RNA replication can be easily evaluated by measuring the reduction in the reporter protein activity. The replicon-based screenings were performed using a BHK-21 ZIKV replicon cell line expressing *Renilla* luciferase as a reporter gene. To characterize the specific targets of identified compounds, we established *in-vitro* fluorescence-based assays for recombinantly expressed NS3 protease and NS5 RdRp. The proteolytic activity of the viral protease was measured by using the fluorogenic peptide substrate Bz-nKRR-AMC, while the NS5 RdRp elongation activity was directly detected by the increase of the fluorescent signal of SYBR Green I during RNA elongation, using the synthetic biotinylated self-priming template 3'UTR-U30 (5'-biotin-U30-ACUGGAGAUCGAUCUCCAGU-3').

Resource

An automatic pipeline for the design of irreversible derivatives identifies a potent SARS-CoV-2 M^{pro} inhibitor

Daniel Zaidman,^{1,8} Paul Gehrtz,^{1,8} Mihajlo Filep,¹ Daren Fearon,² Ronen Gabizon,¹ Alice Douangamath,² Jaime Prilusky,³ Shirly Duberstein,⁴ Galit Cohen,⁴ C. David Owen,^{2,5} Efrat Resnick,¹ Claire Strain-Damerell,^{2,5} Petra Lukacik,^{2,5} Covid-Moonshot Consortium, Haim Barr,⁴ Martin A. Walsh,^{2,5} Frank von Delft,^{2,5,6,7} and Nir London^{1,9,*}

¹Department of Chemical and Structural Biology, Weizmann Institute of Science, 7610001 Rehovot, Israel

²Diamond Light Source Ltd., Harwell Science and Innovation Campus, Didcot OX11 0QX, UK

³Life Sciences Core Facilities, Weizmann Institute of Science, 7610001 Rehovot, Israel

⁴Wohl Institute for Drug Discovery of the Nancy and Stephen Grand Israel National Center for Personalized Medicine, The Weizmann Institute of Science, 7610001 Rehovot, Israel

⁵Research Complex at Harwell, Harwell Science and Innovation Campus, Didcot OX11 0FA, UK

⁶Structural Genomics Consortium, University of Oxford, Old Road Campus, Roosevelt Drive, Headington OX3 7DQ, UK

⁷Department of Biochemistry, University of Johannesburg, Auckland Park 2006, South Africa

⁸These authors contributed equally

⁹Lead contact

*Correspondence: nir.london@weizmann.ac.il

<https://doi.org/10.1016/j.chembiol.2021.05.018>

SUMMARY

Designing covalent inhibitors is increasingly important, although it remains challenging. Here, we present *covalentizer*, a computational pipeline for identifying irreversible inhibitors based on structures of targets with non-covalent binders. Through covalent docking of tailored focused libraries, we identify candidates that can bind covalently to a nearby cysteine while preserving the interactions of the original molecule. We found ~11,000 cysteines proximal to a ligand across 8,386 complexes in the PDB. Of these, the protocol identified 1,553 structures with covalent predictions. In a prospective evaluation, five out of nine predicted covalent kinase inhibitors showed half-maximal inhibitory concentration (IC₅₀) values between 155 nM and 4.5 μM. Application against an existing SARS-CoV M^{pro} reversible inhibitor led to an acrylamide inhibitor series with low micromolar IC₅₀ values against SARS-CoV-2 M^{pro}. The docking was validated by 12 co-crystal structures. Together these examples hint at the vast number of covalent inhibitors accessible through our protocol.

INTRODUCTION

Covalent irreversible inhibitors have become increasingly popular over the last decade as chemical probes and drugs. Most often these inhibitors target a cysteine residue to form the covalent bond. Several rationally designed irreversible inhibitors targeting cysteines were approved by the FDA in recent years, with notable examples, such as ibrutinib (Burger and Buggy, 2013), afatinib (Sequist et al., 2013), and osimertinib (Oxnard et al., 2016). Irreversible inhibitors offer a variety of advantages over non-covalent ones. These include: (1) prolonged residence time (Bradshaw et al., 2015), (2) an ability to compete with high-affinity natural substrates (Lonsdale and Ward, 2018; Michalczyk et al., 2008; Yun et al., 2008), (3) their improved selectivity when targeting non-conserved cysteine residues (Cohen et al., 2005; Ghosh et al., 2019), (4) targeting shallow binding sites (Sutanto et al., 2020), and (5) covalent binding can enable targeting of especially challenging targets such as the G12C oncogenic

K-Ras mutation (Canon et al., 2019; Nnadi et al., 2018; Zeng et al., 2017).

Historically, most covalent inhibitors were designed by the addition of an electrophile to an already known reversible inhibitor that suitably binds next to a cysteine residue (Angst et al., 2020; Dubiella et al., 2015; Hagel et al., 2015; Vazquez-Rodriguez and Wright, 2019; Ward et al., 2015; Weisner et al., 2015). More recently, covalent inhibitors are also being discovered by empirical screening of covalent fragment libraries (Backus et al., 2016; Craven et al., 2018; Johansson et al., 2019; Kathman et al., 2014, 2015; Parker et al., 2017; Resnick et al., 2019) and by covalent virtual screening (Bensinger et al., 2019; Chowdhury et al., 2019; London et al., 2014; Nnadi et al., 2018; Rachman et al., 2019; Scarpino et al., 2018; Shraga et al., 2019; Toledo Warshaviak et al., 2014). While covalent fragment and virtual screening can potentially discover new scaffolds, the binding affinity of primary hits may be relatively low, and often require laborious medicinal chemistry to reach suitable potency.

Article

Non-Toxic Dimeric Peptides Derived from the Bothropstoxin-I Are Potent SARS-CoV-2 and Papain-like Protease Inhibitors

Marjorie C. L. C. Freire ¹, Gabriela D. Noske ¹, Natália V. Bitencourt ², Paulo R. S. Sanches ², Norival A. Santos-Filho ², Victor O. Gawriljuk ¹, Eduardo P. de Souza ³, Victor H. R. Nogueira ¹, Mariana O. de Godoy ¹, Aline M. Nakamura ¹, Rafaela S. Fernandes ¹, Andre S. Godoy ¹, Maria A. Juliano ⁴, Bianca M. Peres ⁵, Cecília G. Barbosa ⁵, Carolina B. Moraes ⁶, Lucio H. G. Freitas-Junior ⁵, Eduardo M. Cilli ², Rafael V. C. Guido ^{1,*} and Glaucius Oliva ^{1,*}

- ¹ São Carlos Institute of Physics, University of Sao Paulo, Avenida João Dagnone, 1100, São Carlos 13563-120, SP, Brazil; marjorie_freire@hotmail.com (M.C.L.C.F.); gabriela.noske@usp.br (G.D.N.); victor.gawriljuk@gmail.com (V.O.G.); victor.nogueira@usp.br (V.H.R.N.); marianaort@gmail.com (M.O.d.G.); alinemnk@gmail.com (A.M.N.); rafaela.fernandes@usp.br (R.S.F.); andre_godoy@yahoo.com.br (A.S.G.)
- ² Department of Biochemistry and Organic Chemistry, Institute of Chemistry, São Paulo State University (UNESP), Araraquara 14800-060, SP, Brazil; natalia.bitencourt@unesp.br (N.V.B.); paulo.sanches@unesp.br (P.R.S.S.); norival.santos-filho@unesp.br (N.A.S.-F.); eduardocilli@gmail.com (E.M.C.)
- ³ Department of Genetics and Evolution, Federal University of São Carlos, Rodovia Washington Luis km 235, São Carlos 13565-905, SP, Brazil; edupsouza96@gmail.com
- ⁴ The Sao Paulo School of Medicine, Federal University of São Paulo, Rua Três de Maio, 100, São Paulo 04044-020, SP, Brazil; ma.juliano@unifesp.br
- ⁵ Department of Microbiology, Institute of Biomedical Sciences, University of Sao Paulo, Av. Prof. Lineu Prestes, 1374, São Paulo 05508-900, SP, Brazil; biaperes04@gmail.com (B.M.P.); cecigomes.barbosa@gmail.com (C.G.B.); luciofreitasjunior@gmail.com (L.H.G.F.-J.)
- ⁶ Department of Pharmaceutical Sciences, Federal University of São Paulo, Rua São Nicolau, 210, Diadema 09913-030, SP, Brazil; cbmoraes@unifesp.br
- * Correspondence: rvcguido@usp.br (R.V.C.G.); oliva@ifsc.usp.br (G.O.); Tel.: +55-16-3373-8673 (R.V.C.G.); +55-16-3373-6664 (G.O.)



Citation: Freire, M.C.L.C.; Noske, G.D.; Bitencourt, N.V.; Sanches, P.R.S.; Santos-Filho, N.A.; Gawriljuk, V.O.; de Souza, E.P.; Nogueira, V.H.R.; de Godoy, M.O.; Nakamura, A.M.; et al. Non-Toxic Dimeric Peptides Derived from the Bothropstoxin-I Are Potent SARS-CoV-2 and Papain-like Protease Inhibitors. *Molecules* **2021**, *26*, 4896. <https://doi.org/10.3390/molecules26164896>

Academic Editors: Jean-Marc Sabatier and Simona Rapposelli

Received: 26 March 2021

Accepted: 7 August 2021

Published: 12 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



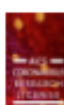
Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: The COVID-19 outbreak has rapidly spread on a global scale, affecting the economy and public health systems throughout the world. In recent years, peptide-based therapeutics have been widely studied and developed to treat infectious diseases, including viral infections. Herein, the antiviral effects of the lysine linked dimer des-Cys¹¹, Lys¹²,Lys¹³-(pBthTX-I)₂K ((pBthTX-I)₂K) and derivatives against SARS-CoV-2 are reported. The lead peptide (pBthTX-I)₂K and derivatives showed attractive inhibitory activities against SARS-CoV-2 (EC₅₀ = 28–65 μM) and mostly low cytotoxic effect (CC₅₀ > 100 μM). To shed light on the mechanism of action underlying the peptides' antiviral activity, the Main Protease (M^{pro}) and Papain-Like protease (PL^{pro}) inhibitory activities of the peptides were assessed. The synthetic peptides showed PL^{pro} inhibition potencies (IC₅₀s = 1.0–3.5 μM) and binding affinities (K_d = 0.9–7 μM) at the low micromolar range but poor inhibitory activity against M^{pro} (IC₅₀ > 10 μM). The modeled binding mode of a representative peptide of the series indicated that the compound blocked the entry of the PL^{pro} substrate toward the protease catalytic cleft. Our findings indicated that non-toxic dimeric peptides derived from the Bothropstoxin-I have attractive cellular and enzymatic inhibitory activities, thereby suggesting that they are promising prototypes for the discovery and development of new drugs against SARS-CoV-2 infection.

Keywords: COVID-19; SARS-CoV-2; inhibitors; Papain-like protease; peptides

1. Introduction

The Severe Acute Respiratory Syndrome virus 2 (SARS-CoV-2) is the causative agent of the Coronavirus disease 2019 (COVID-19), which was first reported in patients in Wuhan,



This article is made available via the ACS COVID-19 subject for unrestricted RESEARCH re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for the duration of the World Health Organization (WHO) declaration of COVID-19 as a global pandemic.

Article

Pyronaridine Protects against SARS-CoV-2 Infection in Mouse

Ana C. Puhl,^{*,#} Giovanni F. Gomes,[#] Samara Damasceno, Andre S. Godoy, Gabriela D. Noske, Aline M. Nakamura, Victor O. Gawriljuk, Rafaela S. Fernandes, Natalia Monakhova, Olga Riabova, Thomas R. Lane, Vadim Makarov, Flavio P. Veras, Sabrina S. Batah, Alexandre T. Fabro, Glaucius Oliva, Fernando Q. Cunha, José C. Alves-Filho, Thiago M. Cunha,^{*} and Sean Ekins[#]

Cite This: *ACS Infect. Dis.* 2022, 8, 1147–1160

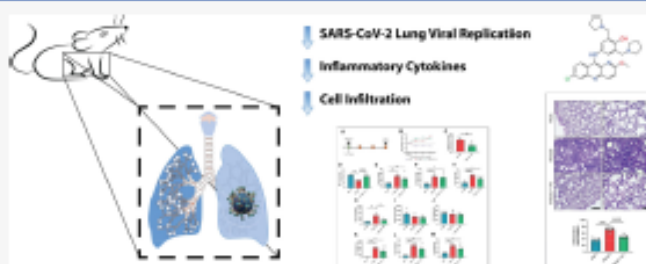
Read Online

ACCESS |

Metrics & More

Article Recommendations

Supporting Information



ABSTRACT: There are currently relatively few small-molecule antiviral drugs that are either approved or emergency-approved for use against severe acute respiratory coronavirus 2 (SARS-CoV-2). One of these is remdesivir, which was originally repurposed from its use against Ebola. We evaluated three molecules we had previously identified computationally with antiviral activity against Ebola and Marburg and identified pyronaridine, which inhibited the SARS-CoV-2 replication in A549-ACE2 cells. The *in vitro* efficacy of pyronaridine has now been assessed in a K18-hACE2 transgenic mouse model of COVID-19. Pyronaridine treatment demonstrated a statistically significant reduction of viral load in the lungs of SARS-CoV-2-infected mice, reducing lung pathology, which was also associated with significant reduction in the levels of pro-inflammatory cytokines/chemokine and cell infiltration. Pyronaridine inhibited the viral PL^{pro} activity *in vitro* (IC₅₀ of 1.8 μM) without any effect on M^{pro}, indicating a possible molecular mechanism involved in its ability to inhibit SARS-CoV-2 replication. We have also generated several pyronaridine analogs to assist in understanding the structure activity relationship for PL^{pro} inhibition. Our results indicate that pyronaridine is a potential therapeutic candidate for COVID-19.

KEYWORDS: antiviral, SARS-CoV-2, spike protein, pyronaridine

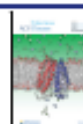
There is currently intense interest in discovering small molecules with direct antiviral activity against the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Pyronaridine, an antiviral drug with *in vitro* activity against Ebola, Marburg, and SARS-CoV-2, has now statistically significantly reduced the viral load in mice along with TNF-α, CXCL1, and CCL3 and restored levels of IFN-1β, ultimately demonstrating a protective effect against lung damage by infection to provide a new potential treatment for testing clinically.

At the time of writing, we are in the midst of a major global health crisis caused by the virus severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) that was originally reported in Wuhan, China in late 2019.^{1,2} Infection with this virus leads to extensive morbidity, mortality, and a very broad range of clinical symptoms such as cough, loss of smell and taste, respiratory distress, pneumonia, and extrapulmonary

events characterized by a sepsis-like disease collectively called 2019 coronavirus disease (COVID-19).³ In the United States, there are currently three vaccines available, one of which has recently obtained with full approval from the Food and Drug Administration (FDA) to protect against SARS-CoV-2.^{4–6} There are however few small-molecule drugs approved for COVID-19,⁷ including remdesivir,⁸ which originally demonstrated activity in Vero cells,^{9,10} human epithelial cells, and in Calu-3 cells¹⁰ infected with SARS-CoV-2 prior to clinical

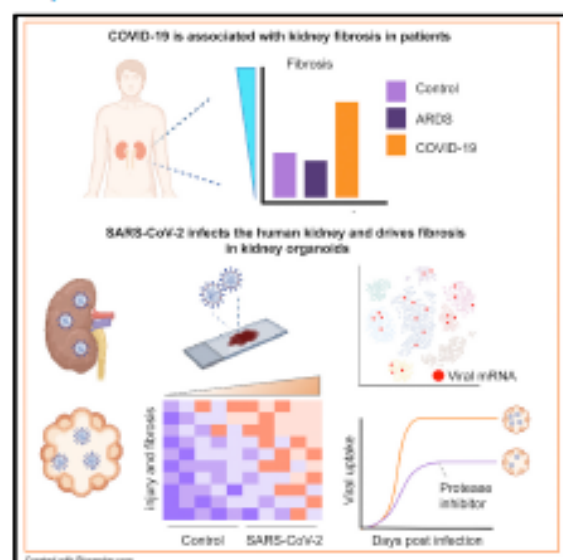
Received: February 11, 2022

Published: May 24, 2022



SARS-CoV-2 infects the human kidney and drives fibrosis in kidney organoids

Graphical abstract



Authors

Jitske Jansen, Katharina C. Reimer, James S. Nagai, ..., Rebekka K. Schneider, Bart Smeets, Rafael Kramann

Correspondence

r.kramann@gmx.net

In brief

Jansen, Reimer, Nagai, et al. report that SARS-CoV-2 infects kidney cells and is associated with kidney fibrosis in patients. Using single-cell transcriptomics of infected kidney organoids, they show that SARS-CoV-2 causes kidney injury and stimulates profibrotic signaling. Viral infection in organoids was inhibited by a recently developed protease blocker.

Highlights

- COVID-19 patients present tubulo-interstitial kidney fibrosis compared with controls
- SARS-CoV-2 infection stimulates profibrotic signaling in human kidney organoids
- SARS-CoV-2 infection can be inhibited by a protease blocker in human kidney organoids



Jansen et al., 2022, Cell Stem Cell 29, 217–231
 February 3, 2022 © 2021 The Author(s). Published by Elsevier Inc.
<https://doi.org/10.1016/j.stem.2021.12.010>



Discovery of New Zika Protease and Polymerase Inhibitors through the Open Science Collaboration Project OpenZika

Melina Mottin,[∞] Bruna Katiele de Paula Sousa,[∞] Nathalya Cristina de Moraes Roso Mesquita, Kettlyn Irene Zagato de Oliveira, Gabriela Dias Noske, Geraldo Rodrigues Sartori, Aline de Oliveira Albuquerque, Fabio Urbina, Ana C. Puhl, José Teófilo Moreira-Filho, Guilherme E. Souza, Rafael V. C. Guido, Eugene Muratov, Bruno Junior Neves, João Herminio Martins da Silva, Alex E. Clark, Jair L. Siqueira-Neto, Alexander L. Perryman, Glaucius Oliva, Sean Ekins, and Carolina Horta Andrade*

Cite This: <https://doi.org/10.1021/acs.jcim.2c00596>

Read Online

ACCESS |

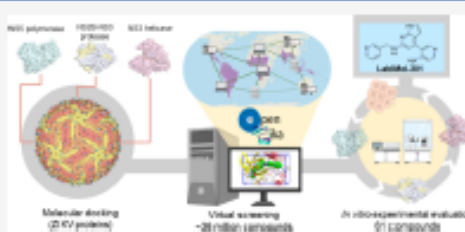
Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: The Zika virus (ZIKV) is a neurotropic arbovirus considered a global threat to public health. Although there have been several efforts in drug discovery projects for ZIKV in recent years, there are still no antiviral drugs approved to date. Here, we describe the results of a global collaborative crowdsourced open science project, the OpenZika project, from IBM's World Community Grid (WCG), which integrates different computational and experimental strategies for advancing a drug candidate for ZIKV. Initially, molecular docking protocols were developed to identify potential inhibitors of ZIKV NS5 RNA-dependent RNA polymerase (NS5 RdRp), NS3 protease (NS2B-NS3pro), and NS3 helicase (NS3hel).

Then, a machine learning (ML) model was built to distinguish active vs inactive compounds for the cytoprotective effect against ZIKV infection. We performed three independent target-based virtual screening campaigns (NS5 RdRp, NS2B-NS3pro, and NS3hel), followed by predictions by the ML model and other filters, and prioritized a total of 61 compounds for further testing in enzymatic and phenotypic assays. This yielded five non-nucleoside compounds which showed inhibitory activity against ZIKV NS5 RdRp in enzymatic assays (IC_{50} range from 0.61 to 17 μ M). Two compounds thermally destabilized NS3hel and showed binding affinity in the micromolar range (K_d range from 9 to 35 μ M). Moreover, the compounds LabMol-301 inhibited both NS5 RdRp and NS2B-NS3pro (IC_{50} of 0.8 and 7.4 μ M, respectively) and LabMol-212 thermally destabilized the ZIKV NS3hel (K_d of 35 μ M). Both also protected cells from death induced by ZIKV infection in *in vitro* cell-based assays. However, while eight compounds (including LabMol-301 and LabMol-212) showed a cytoprotective effect and prevented ZIKV-induced cell death, agreeing with our ML model for prediction of this cytoprotective effect, no compound showed a direct antiviral effect against ZIKV. Thus, the new scaffolds discovered here are promising hits for future structural optimization and for advancing the discovery of further drug candidates for ZIKV. Furthermore, this work has demonstrated the importance of the integration of computational and experimental approaches, as well as the potential of large-scale collaborative networks to advance drug discovery projects for neglected diseases and emerging viruses, despite the lack of available direct antiviral activity and cytoprotective effect data, that reflects on the assertiveness of the computational predictions. The importance of these efforts rests with the need to be prepared for future viral epidemic and pandemic outbreaks.



INTRODUCTION

The Zika virus (ZIKV) is a flavivirus that belongs to the *Flaviviridae* family, transmitted mainly by mosquito bites^{1,2} but also through perinatal³ and sexual⁴ contacts and blood transfusions.⁵ Because of the last ZIKV outbreaks in 2014 in French Polynesia and in 2015 in Brazil and the increasing reports of severe neurological disorders, such as microcephaly^{6–8} and Guillain-Barré syndrome,^{9,10} ZIKV has been included as a world public health concern by the World Health Organization (WHO).¹¹ Moreover, in 2018, Zika was added to






the WHO Blueprint List, a list of priority diseases for research and development in an emergency context. In a recent study (2022), researchers mimicked the natural transmission cycle of

Special Issue: Advancing Women in Chemistry

Received: May 10, 2022

Article

Amino Acid Polymorphisms on the Brazilian Strain of Yellow Fever Virus Methyltransferase Are Related to the Host's Immune Evasion Mediated by Type I Interferon

Nathália Dias Furtado ¹, Iasmim Silva de Mello ¹, Andre Schutzer de Godoy ², Gabriela Dias Noske ², Glaucius Oliva ², Bruno Canard ³, Etienne Decroly ³ and Myrma C. Bonaldo ^{1,*}

¹ Laboratório de Biologia Molecular de Flavivirus, Instituto Oswaldo Cruz-Fiocruz, Rio de Janeiro 21040-900, Brazil

² Centro de Pesquisa e Inovação em Biodiversidade e Fármacos, Instituto de Física de São Carlos-USF, São Paulo 13563-120, Brazil

³ Architecture et Fonction des Macromolécules Biologiques, Aix-Marseille Université, CNRS, UMR7257, 13009 Marseille, France

* Correspondence: mbonaldo@ioc.fiocruz.br

Abstract: Since late 2016, a yellow fever virus (YFV) variant carrying a set of nine amino acid variations has circulated in South America. Three of them were mapped on the methyltransferase (MTase) domain of viral NS5 protein. To assess whether these changes affected viral infectivity, we synthesized YFV carrying the MTase of circulating lineage as well as its isoform with the residues of the previous strains (NS5 K101R, NS5 V138L, and NS5 G173S). We observed a slight difference in viral growth properties and plaque phenotype between the two synthetic YFVs. However, the MTase polymorphisms associated with the Brazilian strain of YFV (2016–2019) confer more susceptibility to the IFN-I. In addition, *in vitro* MTase assay revealed that the interaction between the YFV MTase and the methyl donor molecule (SAM) is altered in the Brazilian MTase variant. Altogether, the results reported here describe that the MTase carrying the molecular signature of the Brazilian YFV circulating since 2016 might display a slight decrease in its catalytic activity but virtually no effect on viral fitness in the parameters comprised in this study. The most marked influence of these residues stands in the immune escape against the antiviral response mediated by IFN-I.

Keywords: yellow fever virus outbreak; methyltransferase; virulence; type I interferon



Citation: Furtado, N.D.; Mello, I.S.d.; Godoy, A.S.d.; Noske, G.D.; Oliva, G.; Canard, B.; Decroly, E.; Bonaldo, M.C. Amino Acid Polymorphisms on the Brazilian Strain of Yellow Fever Virus Methyltransferase Are Related to the Host's Immune Evasion Mediated by Type I Interferon. *Viruses* **2023**, *15*, 191. <https://doi.org/10.3390/v15010191>

Academic Editor: Sérgio Mario Best

Received: 17 November 2022

Revised: 3 January 2023

Accepted: 7 January 2023

Published: 10 January 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Yellow fever virus (YFV) is the causative agent of an acute febrile disease endemic to Sub-Saharan Africa and South America. Since the beginning of the 21st century, small events of YFV spreading were described in Brazil, which resulted in a significant and unprecedented outbreak in late 2016, peaking between 2017 and 2018 [1–3]. Phylogenetic analyses described the clustering of all the YFV sequences determined since 2004 in a sub-clade 1E of the modern lineage belonging to the South American I clade [4–6]. During the latest Brazilian outbreak, complete sequenced genomes of YFV from samples of naturally infected mosquitoes, non-human primates, and humans displayed a molecular signature consisting of polymorphisms associated with the mutation of nine amino acids [2,6]. Three amino acid markers are located in the methyltransferase (MTase) domain of viral protein NS5 at positions R101K, I138V, and S173G [6].

In eukaryotes, mRNA capping and methylation is a pivotal post-transcriptional modification, promoting the mRNA translocation from the nucleus to the cytoplasm and limiting RNA degradation [7]. These post-transcriptional modifications are essential for viral RNA translation, replication, and stability. YFV belongs to the genus *Flavivirus* and family *Flaviviridae*, displaying a positive single-stranded RNA composed of a unique open reading

Allosteric regulation and crystallographic fragment screening of SARS-CoV-2 NSP15 endoribonuclease

Andre Schutzer Godoy^{1,*}, Aline Minalli Nakamura¹, Alice Douangamath^{2,3}, Yun Song⁴, Gabriela Dias Noske¹, Victor Oliveira Gawriljuk¹, Rafaela Sachetto Fernandes¹, Humberto D. Muniz Pereira¹, Kettlyn Irene Zagato Oliveira¹, Daren Fearon^{2,3}, Alexandre Dias^{2,3}, Tobias Krojer⁵, Michael Fairhead⁶, Alisa Powell^{2,3}, Louise Dunnet^{2,3}, Jose Brandao-Neto^{2,3}, Rachael Skyner^{2,3}, Rod Chalk⁶, Dávid Bajusz⁷, Miklós Bege^{8,9}, Anikó Borbás^{8,10}, György Miklós Keserű⁷, Frank von Delft^{2,3,6,11} and Glaucius Oliva^{1,*}

¹Sao Carlos Institute of Physics, University of Sao Paulo, Av. Joao Dagnone, 1100 - Jardim Santa Angelina, Sao Carlos, 13563-120, Brazil, ²Diamond Light Source Ltd, Harwell Science and Innovation Campus, Didcot OX11 0QX, UK, ³Research Complex at Harwell, Harwell Science and Innovation Campus, Didcot OX11 0FA, UK, ⁴Electron Bio-imaging Centre, Diamond Light Source Ltd., Harwell Science and Innovation Campus, Didcot OX11 0QX, UK, ⁵BioMAX, MAX IV Laboratory, Fotongatan 2, Lund 224 84, Sweden, ⁶Centre for Medicines Discovery, Oxford University, Oxford OX1 3QU, UK, ⁷Medicinal Chemistry Research Group, Research Centre for Natural Sciences, Magyar tudósok krt. 2, 1117 Budapest, Hungary, ⁸Department of Pharmaceutical Chemistry, University of Debrecen, Egyetem tér 1, 4032 Debrecen, Hungary, ⁹MTA-DE Molecular Recognition and Interaction Research Group, University of Debrecen, Egyetem tér 1, 4032 Debrecen, Hungary, ¹⁰National Laboratory of Virology, University of Pécs, Ifjúság útja 20, H-7624 Pécs, Hungary and ¹¹Department of Biochemistry, University of Johannesburg, PO Box 524, Auckland Park 2006, South Africa

Received September 26, 2022; Revised April 11, 2023; Editorial Decision April 12, 2023; Accepted April 13, 2023

ABSTRACT

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the causative agent of coronavirus disease 2019 (COVID-19). The NSP15 endoribonuclease enzyme, known as NendoU, is highly conserved and plays a critical role in the ability of the virus to evade the immune system. NendoU is a promising target for the development of new antiviral drugs. However, the complexity of the enzyme's structure and kinetics, along with the broad range of recognition sequences and lack of structural complexes, hampers the development of inhibitors. Here, we performed enzymatic characterization of NendoU in its monomeric and hexameric form, showing that hexamers are allosteric enzymes with a positive cooperative index, and with no influence of manganese on

enzymatic activity. Through combining cryo-electron microscopy at different pHs, X-ray crystallography and biochemical and structural analysis, we showed that NendoU can shift between open and closed forms, which probably correspond to active and inactive states, respectively. We also explored the possibility of NendoU assembling into larger supramolecular structures and proposed a mechanism for allosteric regulation. In addition, we conducted a large fragment screening campaign against NendoU and identified several new allosteric sites that could be targeted for the development of new inhibitors. Overall, our findings provide insights into the complex structure and function of NendoU and offer new opportunities for the development of inhibitors.