

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Instituto de Ciências Biológicas
Programa Interunidades de Pós-Graduação em Bioinformática

Vivian Morais Paixão

**IDENTIFICAÇÃO DE NOVOS PEPTÍDEOS INIBIDORES DE TMPRSS2: uma
abordagem computacional para prospecção de um tratamento antiviral**

Belo Horizonte
2023

Vivian Morais Paixão

IDENTIFICAÇÃO DE NOVOS PEPTÍDEOS INIBIDORES DE TMPRSS2: uma abordagem computacional para prospecção de um tratamento antiviral

Dissertação apresentada ao Programa Interunidades de Pós-Graduação em Bioinformática da Universidade Federal de Minas Gerais, como requisito à obtenção do título de Mestra em Bioinformática.

Orientadora: Raquel Cardoso de Melo Minardi

Coorientador: Adriano de Paula Sabino

Belo Horizonte

2023

043

Paixão, Vivian Morais.

Identificação de novos peptídeos inibidores de TMPRSS2: uma abordagem computacional para prospecção de um tratamento antiviral [manuscrito] / Vivian Morais Paixão. – 2023.

79 f. : il. ; 29,5 cm.

Orientador: Raquel Cardoso de Melo Minardi. Coorientador: Adriano de Paula Sabino.

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

1. Bioinformática. 2. Simulação de Acoplamento Molecular. 3. Peptídeos. 4. Betacoronavirus. 5. Angiotensina II. I. Minardi, Raquel Cardoso de Melo. II. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. III. Título.

CDU: 573:004



UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA INTERUNIDADES DE PÓS-GRADUAÇÃO EM BIOINFORMÁTICA

ATA DA DEFESA DE DISSERTAÇÃO

VIVIAN MORAIS PAIXÃO

Às treze horas e trinta minutos do dia **05 de dezembro de 2023**, reuniu-se, no aplicativo Zoom, a Comissão Examinadora de Dissertação, indicada pelo Colegiado do Programa, para julgar, em exame final, o trabalho intitulado: "**Identificação de novos peptídeos inibidores de Tmprss2: uma abordagem computacional para prospecção de um tratamento antiviral**", requisito para obtenção do grau de Mestre em **Bioinformática**. Abrindo a sessão, a Presidente da Comissão, **Dra. Raquel Cardoso de Melo Minardi**, após dar a conhecer aos presentes o teor das Normas Regulamentares do Trabalho Final, passou a palavra à candidata, para apresentação de seu trabalho. Seguiu-se a arguição pelos Examinadores, com a respectiva defesa da candidata. Logo após, a Comissão se reuniu, sem a presença da candidata e do público, para julgamento e expedição de resultado final. Foram atribuídas as seguintes indicações:

| Professor(a)/Pesquisador(a) | Instituição | Indicação |
|---------------------------------------|--------------------------------------|-----------------|
| Dra. Raquel Cardoso de Melo Minardi | Universidade Federal de Minas Gerais | Aprovada |
| Dr. Adriano de Paula Sabino | Universidade Federal de Minas Gerais | Aprovada |
| Dr. Lucas Bleicher | Universidade Federal de Minas Gerais | Aprovada |
| Dra. Maria Goreti de Almeida Oliveira | Universidade Federal de Viçosa | Aprovada |
| Dra. Rafaela Salgado Ferreira | Universidade Federal de Minas Gerais | Aprovada |

Pelas indicações, a candidata foi considerada: **Aprovada**

O resultado final foi comunicado publicamente à candidata pela Presidente da Comissão. Nada mais havendo a tratar, a Presidente encerrou a reunião e lavrou a presente ATA, que será assinada por todos os membros participantes da Comissão Examinadora.

Belo Horizonte, 05 de dezembro de 2023.



Documento assinado eletronicamente por **Maria Goreti de Almeida Oliveira, Usuário Externo**, em 06/12/2023, às 07:18, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Rafaela Salgado Ferreira, Professora do Magistério Superior**, em 06/12/2023, às 09:01, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Raquel Cardoso de Melo Minardi, Professora do Magistério Superior**, em 06/12/2023, às 09:27, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Lucas Bleicher, Professor do Magistério Superior**, em 06/12/2023, às 11:07, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Adriano de Paula Sabino, Professor do Magistério Superior**, em 15/12/2023, às 16:42, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **2873347** e o código CRC **962C711C**.

Agradecimentos

Primeiramente, agradeço a Deus pela dádiva da vida, pela proteção e pela saúde que me acompanham ao longo da minha jornada.

À minha família, sou imensamente grata pelo apoio e encorajamento desde os tempos de colégio e durante a graduação, estendendo-se ao período do mestrado. Especialmente aos meus pais, Valéria e Carlos, que não só me deram a vida, mas também têm apoiado as minhas escolhas, oferecendo auxílio e sempre demonstrando orgulho, mesmo diante do que muitas vezes pode ser um território desconhecido para eles. À minha irmã Sylvia, agradeço pelo suporte contínuo e pelos momentos descontraídos que tornaram o caminho mais leve. Ao meu namorado, Pedro, que agradeço pelo apoio incansável, pela crença em minhas capacidades quando eu mesma duvidava, pelos conselhos e pelo suporte na resolução de questões computacionais, especialmente quando surgiam problemas em algum código.

Agradeço calorosamente a todos os amigos e colegas do Laboratório de Bioinformática e Sistemas. O aprendizado, companheirismo, amizade e auxílio no desenvolvimento e escrita do projeto foram inestimáveis. Em especial, agradeço à Luana, cuja paciência para esclarecer minhas dúvidas e ajuda, mesmo nos finais de semana, foram de grande valia. Agradeço também aos amigos feitos ao longo do mestrado, os quais espero levar para a vida.

Não poderia deixar de agradecer à Profª. Raquel Minardi (minha orientadora) e ao Prof. Adriano Sabino (meu co-orientador) por aceitarem orientar e contribuir para a execução deste projeto. Durante esses dois anos, aprendi imensamente e reconheço que nada disso teria sido realizado sem os ensinamentos, o suporte e as oportunidades oferecidas por eles.

Por fim, expresso minha gratidão às agências de fomento à pesquisa: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) e Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). Além disso, agradeço ao Programa de Pós-graduação em Bioinformática da UFMG e à Sheila e ao Tiago.

O presente trabalho contou com o apoio da Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG).

Resumo

Em 2020, uma pandemia causada pelo SARS-CoV-2, um novo coronavírus, impactou a economia global e a saúde pública, resultando em quase 7 milhões de mortes em todo o mundo. Pesquisas revelaram que o vírus requer a ligação da proteína Spike à Enzima Conversora de Angiotensina 2 (ACE2) para infectar as células do hospedeiro. Essa ligação depende da clivagem da proteína Spike, uma ação mediada pelas enzimas Furina e TMPRSS2. Este estudo concentra-se na TMPRSS2, uma serino protease transmembrana encontrada em várias superfícies celulares, incluindo próstata, pulmão e intestino, e associada à patogênese de várias doenças. Embora já existam medicamentos aprovados para o tratamento da COVID-19, eles apresentam desafios, como custos elevados, efeitos colaterais e uma crescente perda de eficácia frente às novas variantes do vírus. Portanto, o objetivo principal deste estudo foi desenvolver uma metodologia para a busca e filtragem de peptídeos como possíveis abordagens terapêuticas, com ênfase na identificação de peptídeos ainda não descritos como inibidores dessa protease. Nossa metodologia incluiu uma revisão da literatura, a pesquisa por peptídeos com características de ligação ao sítio catalítico da TMPRSS2 por meio do banco de dados Propedia, a filtragem das sequências resultantes, simulações de *docking* molecular dos peptídeos selecionados e a aplicação de um código para filtrar os modelos gerados. Isso resultou na identificação de 147 candidatos, dos quais seis se destacaram como os mais promissores. Posteriormente, foi realizada uma análise da estabilidade das interações. Três dos candidatos mais promissores foram otimizados, gerando novas sequências com maior probabilidade de ligação ao sítio ativo. Esses peptídeos otimizados foram posteriormente sintetizados para realização de ensaios *in vitro*. Os resultados demonstraram que nossa metodologia é capaz de identificar poses de *docking* que inicialmente não seriam consideradas para avaliação, destacando aquelas com melhor ajuste ao sítio ativo. Além disso, três peptídeos com histórico de inibição de proteases foram identificados como candidatos potenciais para a inibição da TMPRSS2, sugerindo sua aplicabilidade não apenas para o tratamento da COVID-19, mas também para outras doenças nas quais a TMPRSS2 está envolvida.

Palavras-chave: tmprss2; docking; peptídeos; inibição

Abstract

In 2020, a pandemic caused by SARS-CoV-2, a novel coronavirus, had a global impact on the economy and public health, resulting in nearly 7 million deaths worldwide. The virus requires the binding of its Spike protein to the Angiotensin-Converting Enzyme 2 (ACE2) receptor to infect host cells. This binding relies on the cleavage of the Spike protein, a process mediated by the enzymes Furin and TMPRSS2. This study focuses on TMPRSS2, a transmembrane serine protease found on various cell surfaces, including the prostate, lungs, and intestines, and associated with the pathogenesis of multiple diseases. Although approved drugs for COVID-19 treatment already exist, they present challenges such as high costs, side effects, and a growing loss of efficacy against new virus variants. Therefore, the primary objective of this study was to develop a methodology for the search and filtering of peptides as potential therapeutic approaches, with an emphasis on identifying peptides not previously described as inhibitors of this protease. Our methodology included a literature review, the search for peptides with binding characteristics to the catalytic site of TMPRSS2 using the Propedia database, filtering of the resulting sequences, molecular *docking* simulations of selected peptides, and the application of a code to filter the generated docking models. This led to the identification of 147 candidates, of which six emerged as the most promising. Subsequently, an analysis of interaction stability was performed. Three of the most promising candidates were optimized, generating new sequences with a higher probability of binding to the active site. These optimized peptides were later synthesized for *in vitro* assays. The results demonstrated that our methodology is capable of revealing docking poses that initially would not have been considered for evaluation, highlighting poses with a better fit to the active site. Moreover, three peptides with a history of protease inhibition were identified as potential candidates for TMPRSS2 inhibition, suggesting their applicability not only for COVID-19 treatment but also for other diseases in which TMPRSS2 plays a crucial role.

Keywords: tmprss2; docking; peptides; inhibition

Lista de Figuras

| | |
|---|----|
| Figura 1. Gene TMPRSS2 em localização genômica. | 16 |
| Figura 2. Ectodomínio da TMPRSS2. | 17 |
| Figura 3. Representação da ação da TMPRSS2. | 18 |
| Figura 4. Estruturas da hemaglutinina do vírus Influenza e da proteína spike do coronavírus. | 20 |
| Figura 6. Alinhamento das sequências de aminoácidos dos vírus envolvidas com a TMPRSS2. | 22 |
| Figura 7. Taxa de sucesso da predição do modo de ligação para HPEPDOCK e HADDOCK. | 28 |
| Figura 8. Taxa de sucesso da predição do modo de ligação para HPEPDOCK e pepATTRACT. | 29 |
| Figura 9. Esquema metodológico. | 32 |
| Figura 10. Estrutura da serino protease TMPRSS2 (PDB ID: 7MEQ). | 33 |
| Figura 11. Funcionamento do script desenvolvido no projeto. | 40 |
| Figura 12. Arquivo de saída do script. | 40 |
| Figura 16. Gráfico de análise de flexibilidade da TMPRSS2. | 44 |
| Figura 17. Estrutura da TMPRSS2 gerada pela ferramenta Propedia. | 44 |
| Figura 18. Resultados gerados pelo Propedia. | 45 |
| Figura 19. Melhores modelos de docking dos peptídeos ancorados ao sítio catalítico de TMPRSS2. | 49 |
| Figura 20. Gráfico de análise de flexibilidade para 1FLE. | 53 |
| Figura 21. Gráfico de análise de flexibilidade para 1GL1. | 54 |
| Figura 22. Gráfico de análise de flexibilidade para 3FP7. | 55 |
| Figura 23. Gráfico de análise de flexibilidade para 1LDT. | 56 |
| Figura 24. Gráfico de análise de flexibilidade para 4DOQ. | 57 |
| Figura 25. Gráfico de análise de flexibilidade para 4ZKN. | 58 |
| Figura 26. Melhores modelos de docking dos peptídeos otimizados. | 60 |

Lista de Tabelas

| | |
|--|-----------|
| Tabela 1. Estruturas disponíveis no PDB utilizadas para a etapa de re-docking. | 38 |
| Tabela 2. Tipos de ligação identificados através do script. | 40 |
| Tabela 3. Sequências selecionadas após o alinhamento par-a-par com a aprotinina. | 47 |
| Tabela 4. Sequências selecionadas após a etapa de filtragem (3.2.2). | 47 |
| Tabela 5. Peptídeos utilizados na etapa de docking molecular. | 48 |
| Tabela 6. Resultado do re-docking. | 49 |
| Tabela 7. Peptídeos selecionados utilizando a metodologia do projeto. | 50 |
| Tabela 8. Tabela contendo as informações sobre os peptídeos selecionados no projeto. | 53 |
| Tabela 9. Tabela indicando os dez decoys utilizados como controle negativo para o script desenvolvido no projeto. | 54 |
| Tabela 10. Resultados da Análise de Modos Normais. | 61 |
| Tabela 11. Sequências utilizadas para otimização. | 61 |
| Tabela 12. Matriz de especificidade do MEROPS para TMPRSS2. | 63 |

Lista de Abreviações

ACE2 - *Angiotensin Converting Enzyme 2*

ANVISA - Agência Nacional de Vigilância Sanitária

AR - *Androgen Receptor*

ARE - *Androgen Response Element*

BLAST - *Basic Local Alignment Search Tool*

blastp - *protein-protein BLAST*

BPTI - *Bovine pancreatic trypsin inhibitor*

C - Carbono

CA - Carbono alfa

CG - *Coarse Grained*

ECM - *Extracellular Matrix*

ERG - *Erythroblast-specific-related gene*

FDA - *Food and Drug Administration*

GISAID - *Global Initiative on Sharing Avian Influenza Data*

HMPV - *Metapneumovirus humano*

HA - *Hemaglutinina*

Ha - *Heavy Atoms*

HIV - *Human Immunodeficiency Virus*

HPIV - *Human Parainfluenza Virus*

kDa - *kiloDalton*

LDLR - *Low Density Lipoprotein Receptor*

LDTI - *Leech-derived Tryptase inhibitor*

MERS - *Middle East Respiratory Syndrome coronavirus*

Mpro - *Main protease*

mRNA - RNA mensageiro

MTP - Molnupiravir trifosfato

N - Nitrogênio

NCBI - *National Center for Biotechnology Information*

NMA - *Normal Mode Analysis*

NTD - *N-Terminal Domain*

OMS - Organização Mundial da Saúde

O - Oxigênio

ORF - *Open Reading Frame*

PDB - *Protein Data Bank*

RBD - *Receptor Binding Domain*

RdRp - *RNA dependent RNA polymerase*

RMSD - *Root Mean Square Deviation*

SARS-CoV-2 - *Severe Acute Respiratory Syndrome Coronavirus 2*

SeV - *Sendai Virus*

SLPI - *Secretory Leukocyte Protease Inhibitor*

SP - *Serine Peptidase*

SRCR-2 - *Scavenger Receptor Cysteine Rich 2*

TMPRSS2 - *Transmembrane Protease Serine 2*

TTSP - *Type II Transmembrane Serine Protease*

WAP - *Whey Acidic Protein*

Produção Acadêmica (2021 - 2023)

Artigos publicados e apresentações diretamente relacionados a esta dissertação

1. **Capítulo de livro:** PAIXÃO, V. M.; MINARDI, R. C. *Computational methodology for discovery of potential inhibitory peptides. Advances in Bioinformatics and Computational Biology*, v. 13523, 2022. DOI: 10.1007/978-3-031-21175-1_10.
2. **Apresentação oral** do projeto “*Computational methodology for discovery of potential inhibitory peptides*” no *Brazilian Symposium on Bioinformatics (BSB)* - 2022
3. **Apresentação de pôster** do projeto “*Computational methodology for discovery of potential inhibitory peptides, focusing on the TMPRSS2 protease*” no *X-Meeting* - 2023
4. **Apresentação de pôster** do projeto “*Identifying novel TMPRSS2 inhibitory peptides: a computational approach towards COVID-19 treatment*” no *Congresso Brasileiro de Virologia* - 2023

Outros tipos de produção acadêmica

1. **Organização de livro: Revista BIOINFO:** BASTOS, L. L. et al. *BIOINFO #02 - Revista Brasileira de Bioinformática e Biologia Computacional*, v. 1, 2022. Ed. Alfahelix. DOI: 10.51780/978-65-992753-5-7.
2. **Curso de curta duração ministrado:** Introdução à programação com Python (Curso de Inverno em Bioinformática 2022 - UFMG)
3. **Capítulo de livro:** PAIXÃO, V. M. Por que usar testes estatísticos na sua pesquisa, e quais os mais comuns?. *BIOINFO*, v. 2, p. 39-45, 2022.
4. **Membro da comissão organizadora** da “*Liga Brasileira de Bioinformática - 3ª Edição*”, realizada ao longo do primeiro semestre de 2023, com a última etapa sendo realizada no *X-Meeting 2023*.
5. **Artigo:** MARTINS, P. et al. Propedia v2.3: A novel representation approach for the peptide-protein interaction database using graph-based structural signatures. *Frontiers in Bioinformatics*, v. 3, p. 1-6, 2023. DOI: 10.3389/fbinf.2023.1103103.

6. **E-book publicado:** PAIXÃO, V. M. et al. LBS Tech - Desenvolvimento web, v. 1, 2023. Ed. Alfahelix. DOI: 10.51780/978-65-992753-6-4.
7. **E-book publicado:** PAIXÃO, V. M. et al. Bioinformática com Jupyter Notebook, v. 1, 2023. Ed. Alfahelix. DOI: 10.51780/978-65-992753-7-1.
8. **Capítulo de livro:** Carvalho, F. et al. (2023). *Using Natural Language Processing for Context Identification in COVID-19 Literature*. In: Reis, M.S., de Melo-Minardi, R.C. (eds) *Advances in Bioinformatics and Computational Biology*. BSB 2023. Lecture Notes in Computer Science(), vol 13954. Springer, Cham. https://doi.org/10.1007/978-3-031-42715-2_7
9. **Capítulo de livro:** PAIXÃO, V. M. et al. AlphaFold 2: revolucionando a modelagem de estruturas 3D de macromoléculas. BIOINFO, v. 3, p. 23, 2023.

Sumário

| | |
|--|----|
| 1. INTRODUÇÃO | 17 |
| 1.1 Serino protease transmembrana 2 | 18 |
| 1.2 Doenças relacionadas à TMPRSS2 | 20 |
| 1.2.1 Coronavírus (SARS-CoV-2, SARS-CoV e MERS-CoV) | 20 |
| 1.2.2 Vírus influenza | 22 |
| 1.2.3 Câncer de próstata | 22 |
| 1.3 Medicamentos para COVID-19 | 24 |
| 1.3.1 Potencial da inibição da TMPRSS2 como abordagem terapêutica | 24 |
| 1.3.2 Medicamentos disponíveis e suas limitações | 25 |
| 1.3.3 Peptídeos como novas abordagens terapêuticas | 28 |
| 1.4 Docking molecular como ferramenta para busca de peptídeos inibidores | 29 |
| 1.4.1 Limitações do docking molecular | 30 |
| 1.5 Justificativa do trabalho | 33 |
| 2. OBJETIVOS | 35 |
| 2.1 Objetivo geral | 35 |
| 2.2 Objetivos específicos | 35 |
| 3. MATERIAIS E MÉTODOS | 36 |
| 3.1 Levantamento bibliográfico e análise da estrutura da TMPRSS2 | 36 |
| 3.2 Obtenção dos peptídeos candidatos | 38 |
| 3.2.2 Filtragem das sequências | 39 |
| 3.4 Docking molecular | 40 |
| 3.4.2 Re-docking HPEPDOCK/HDOCK | 41 |
| 3.5 Filtragem dos melhores modelos | 42 |
| 3.5.1 Análise visual das poses selecionadas | 44 |
| 3.5.2 Controle negativo do script | 45 |
| 3.6 Análise da estabilidade dos modos de ligação | 45 |
| 3.7 Otimização e síntese dos peptídeos | 45 |
| 4. RESULTADOS E DISCUSSÃO | 47 |
| 4.1 Levantamento bibliográfico e análise da estrutura da TMPRSS2 | 47 |
| 4.2 Obtenção dos peptídeos candidatos | 48 |
| 4.2.2 Filtragem das sequências | 49 |
| 4.3 Docking molecular | 51 |
| 4.3.1 Re-docking HPEPDOCK/HDOCK | 52 |
| 4.4 Filtragem das melhores poses | 52 |
| 4.4.1 Controle negativo do script | 57 |
| 4.5 Análise da estabilidade dos modos de ligação | 57 |
| 4.5.1 Peptídeo 1FLE | 57 |
| 4.5.2 Peptídeo 1GL1 | 58 |
| 4.5.3 Peptídeo 3FP7 | 59 |

| | |
|--|-----------|
| 4.5.4 Peptídeo 1LDT | 60 |
| 4.5.5 Peptídeo 4DOQ | 61 |
| 4.5.6 Peptídeo 4ZKN | 62 |
| 4.7 Otimização e síntese dos peptídeos | 64 |
| 5. CONCLUSÕES | 67 |
| 6. LIMITAÇÕES E PERSPECTIVAS | 68 |
| 7. REFERÊNCIAS | 69 |

1. Introdução

No final de 2019, o mundo foi apresentado a um novo vírus, o SARS-CoV-2 (*Severe Acute Respiratory Syndrome Coronavirus 2*), que desencadeou a pandemia de COVID-19 (YUAN et al., 2023). Em março de 2020, a Organização Mundial da Saúde (OMS) declarou oficialmente o SARS-CoV-2 como um grave problema de saúde pública em escala global, caracterizando-o como uma pandemia (CUCINOTTA; VANELLI, 2020). Essa pandemia teve um impacto profundo tanto na economia quanto na saúde pública em todo o mundo (MARTINS et al., 2020) e, desde então, tem havido esforços incansáveis para desenvolver novos medicamentos e vacinas para combater o vírus (YUAN et al., 2023).

É nesse contexto que um alvo molecular já conhecido voltou a despertar o interesse da comunidade científica: as Serino Proteases Transmembrana 2, ou TMPRSS2. Esta protease desempenha um papel importante em diversas doenças virais, facilitando a entrada desses vírus nas células humanas (FRASER et al., 2022). Dentre os vírus em que as TMPRSS2 estão envolvidas no mecanismo de infecção, destacam-se o SARS-CoV-2, o SARS-CoV, o metapneumovírus humano (HMPV) (SHIROGANE et al., 2008), o influenza A (incluindo cepas H1N1, H3N2 e H7N9) (LAPORTE; NAESENS, 2017), o vírus Sendai (SeV) e o parainfluenza humano (HPIV) (ABE et al., 2013). Além de seu papel em infecções virais, estudos também apontam que o gene da TMPRSS2 é altamente expresso em casos de câncer de próstata (AFAR et al., 2001), contribuindo para seu desenvolvimento.

Recentemente, alguns medicamentos antivirais foram aprovados pela FDA (*Food and Drug Administration*) para o tratamento da COVID-19, como Remdesivir, Paxlovid e Molnuparavir. Entretanto, a adoção desses medicamentos tem encontrado alguns obstáculos. Apesar de alguns estudos mostrarem a eficácia do Remdesivir (BEIGEL et al., 2020; GOTTLIEB et al., 2022), não houve resultados significativos em outros (ADER et al., 2022; SPINNER et al., 2020), apesar de se mostrar efetivo em conjunto com o antiviral baricitinib (KALIL et al., 2021). O Paxlovid e o Molnuparavir, apesar de supostamente serem eficientes, são encontrados em farmácias por preços que variam de R\$1,5 mil a quase R\$5 mil, ou \$308 a \$1027 dólares (FONTANETTO, 2023). Além disso, medicamentos que possuem a proteína Spike ou outra proteína viral como alvo podem perder a eficácia diante de variantes com mutações significativas (SHEN et al., 2017) e, dessa forma, terapias visando alvos do hospedeiro podem ser mais vantajosas. Um estudo mostrou também que a inibição da

TMPRSS2 reduz a invasão celular e metástase em câncer de próstata (KO et al., 2020), mostrando outro potencial tratamento em que inibidores dessa protease podem ser úteis.

Nos últimos anos, o uso de peptídeos como medicamentos destacou-se devido à sua versatilidade e eficácia em diversas aplicações clínicas. Os peptídeos são moléculas formadas pela união de aminoácidos, sendo menores que proteínas. Essas moléculas têm demonstrado ser uma solução promissora em áreas como tratamentos hormonais, como a insulina no controle da diabetes, e terapias direcionadas para uma variedade de condições médicas (WANG et al., 2022).

É importante ressaltar que o tratamento utilizando peptídeos como inibidores possui algumas vantagens que se sobrepõem com relação aos medicamentos citados, como: mecanismo de ação específico, pode ser administrado em doses menores e possui menor toxicidade (AGRAHARI; AGRAHARI; MITRA, 2016). Dessa forma, este estudo busca uma abordagem terapêutica baseada em peptídeos visando a inibição da TMPRSS2, de forma a produzir um medicamento de maior especificidade e menos efeitos colaterais.

1.1 Serino protease transmembrana 2

As Serino Proteases Transmembrana 2, ou TMPRSS2, são enzimas que têm atraído considerável atenção na área da saúde devido ao seu papel crucial em uma variedade de contextos fisiológicos e patológicos (FRASER et al., 2022). Essa protease, presente na membrana celular, desempenha funções fundamentais que vão desde a regulação de processos celulares até a sua participação em doenças virais e câncer (THUNDERS; DELAHUNT, 2020).

A TMPRSS2 pertence à família de serino proteases transmembrana tipo 2 (*Type II Transmembrane Serine Protease - TTSP*), composta por 19 serino proteases *trypsin-like* (FRASER et al., 2022), ou seja, que exibem atividade semelhante à da tripsina, uma enzima produzida no pâncreas e envolvida na digestão de proteínas. Os membros dessa família geralmente participam de cascatas proteolíticas para remodelação degradativa da matriz extracelular (MUKAI et al., 2015), ativação proteolítica de proteínas de membrana (HEURICH et al., 2014), entre outras funções. Todas elas são proteínas expressas como proenzimas de cadeia única, denominadas zimogênios. Elas necessitam ser ativadas através de uma clivagem proteolítica em uma ligação peptídica conservada para que, assim, o seu domínio catalítico seja ativado e elas consigam exercer sua função (TSENG et al., 2017). Essa clivagem pode ser realizada pela própria protease ou por alguma outra externa (AFAR et al., 2001; LU et al., 1997). No caso da TMPRSS2, essa clivagem ocorre na ligação peptídica

ARG255-ILE256 em seu motivo de ativação (SRQSR255↓IVGGE) (AFAR et al., 2001; CHEN et al., 2010).

Especificamente sobre a serino protease transmembrana 2, essa protease de 70kDa (*kiloDalton*) é codificada pelo gene TMPRSS2 (BERTRAM et al., 2010b). Este gene está localizado no cromossomo 21q22.3 (Figura 1), possui 15 exons e um ORF (*Open reading frame*) de 492 resíduos de aminoácidos, além de ser altamente conservado em chimpanzés, macaco *Rhesus*, cachorro, vaca, rato, galinha, *zebrafish*, *Caenorhabditis elegans* e no sapo (THUNDERS; DELAHUNT, 2020). Ele possui duas isoformas, ambas possuem o zimogênio como forma inativada e precisam ser ativadas cataliticamente. A isoforma 1 é resultado de um *splicing* alternativo do mRNA (RNA mensageiro), sendo quase idêntica à isoforma 2, porém com um N-terminal citoplasmático estendido, resultando no aumento de mais 37 resíduos de aminoácidos à sequência (THUNDERS; DELAHUNT, 2020). A TMPRSS2 é expressa no pulmão e células dos brônquios, mas também no cólon, estômago, pâncreas, glândulas salivares, próstata e outros tecidos (VAARALA et al., 2001).

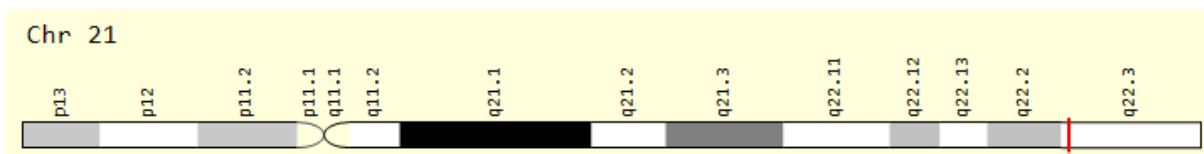


Figura 1. Gene TMPRSS2 em localização genômica, marcado com uma linha vermelha (região 21q22.3).

Fonte: GeneCards. Disponível em:

<https://www.genecards.org/cgi-bin/carddisp.pl?gene=TMPRSS2#pathways_interactions>. Acesso em: 01 de Outubro de 2023.

O gene da TMPRSS2, assim como outros, está sujeito a mutações. David *et al.* (2022) analisaram em sua pesquisa 378 variantes do gene que ocorrem naturalmente na população humana. Nelas, o aminoácido V160 é altamente conservado, sugerindo um papel importante na estrutura ou função. A porção do domínio SRCR-2 (*Scavenger receptor cysteine rich domain 2*) é outra região altamente conservada, cuja função não é completamente compreendida, mas foi proposto por Aruffo et al. (1997) que tenha um papel na interação com o ligante. Além disso, algumas variantes podem influenciar na susceptibilidade ao SARS-CoV-2 e, inclusive, uma delas foi associada a um maior risco de doença grave por vírus influenza H1N1 na pandemia que ocorreu em 2009 (CHENG et al., 2015).

Estruturalmente, a TMPRSS2 é composta por uma região citoplasmática (resíduos 1-84), uma região transmembrana (resíduos 85-105) e uma região extracelular (resíduos

106-492). É composta por três domínios: LDLR (*Low Density Lipoprotein Receptor*) classe A (resíduos 112-149), SRCR-2 (*Scavenger Receptor Cysteine Rich 2*) (resíduos 150-242) e a Peptidase S1 (resíduos 256-489), que contém o sítio ativo (H296, D345 e S441) (DAVID et al., 2022) (Figura 2). O domínio SP (*Serine peptidase*) da TMPRSS2 é altamente conservado com todas TTSPs (PERONA; CRAIK, 1997), e o seu sítio ativo parece capaz de se ligar especificamente a diferentes sequências de substrato, com preferência estrita por Arg e Lys nas posições P1 e P2 (FRASER et al., 2022).

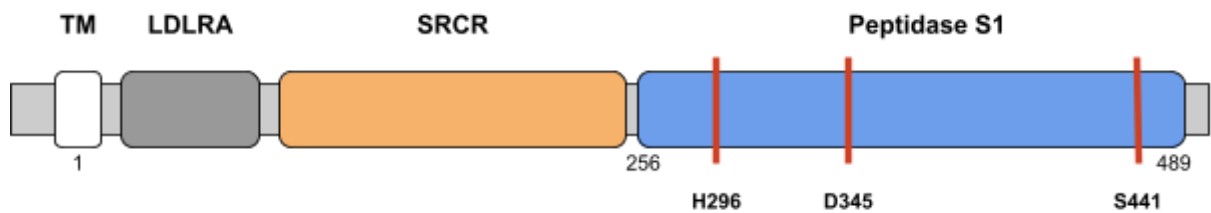


Figura 2. Ectodomínio da TMPRSS2. As cadeias não catalíticas (LDLR e SRCR) estão representadas em cinza e laranja, respectivamente, enquanto a cadeia catalítica (SP) está representada em azul, com destaque para os aminoácidos representantes da tríade catalítica: H296-D345-S441.

Fonte: adaptado de Alessia David et al., 2020.

1.2 Doenças relacionadas à TMPRSS2

Uma das facetas da TMPRSS2 que podemos destacar é seu envolvimento na infecção viral. Em particular, a TMPRSS2 desempenha um papel crucial em doenças virais causadas pelos vírus SARS-CoV-2, SARS-CoV, MERS-CoV e Influenza (THUNDERS; DELAHUNT, 2020). Esses vírus utilizam a TMPRSS2 para clivar proteínas virais, permitindo sua entrada nas células hospedeiras. A compreensão desse mecanismo é essencial para o desenvolvimento de estratégias terapêuticas eficazes.

1.2.1 Coronavírus (*SARS-CoV-2*, *SARS-CoV* e *MERS-CoV*)

O SARS-CoV-2 possui quatro proteínas estruturais primárias: *Spike* (S), de membrana, do envelope e do nucleocapsídeo (SOFI; HAMID; BHAT, 2020). A sua entrada nas células humanas é mediada por uma interação entre a glicoproteína *Spike* e o receptor da Enzima Conversora da Angiotensina 2 (*Angiotensin Converting Enzyme 2* - ACE2), presente na célula do hospedeiro (BROOKE; PRISCHI, 2020).

É importante destacar que a proteína S possui dois domínios funcionais: S1 e S2. O domínio S1 é o RBD (*Receptor Binding Domain*) e o S2 contém elementos funcionais envolvidos na fusão da membrana, parte imprescindível para multiplicação viral

(HOFFMANN; KLEINE-WEBER; PÖHLMANN, 2020). A ligação da proteína S à ACE2 depende da clivagem da proteína em três partes (S1, S2 e S2') por meio de proteases presentes no hospedeiro. Para que a entrada do vírus na célula seja concretizada, a proteína S passa por uma mudança conformacional que expõe o sítio de clivagem entre os domínios S1 e S2 (SLLR667↓), sendo clivada inicialmente pela enzima Furina. A parte S2 é então clivada pela TMPRSS2 para expor o sítio de fusão, possibilitando a entrada endocítica do vírus (STOPSACK et al., 2020) (Figura 3).

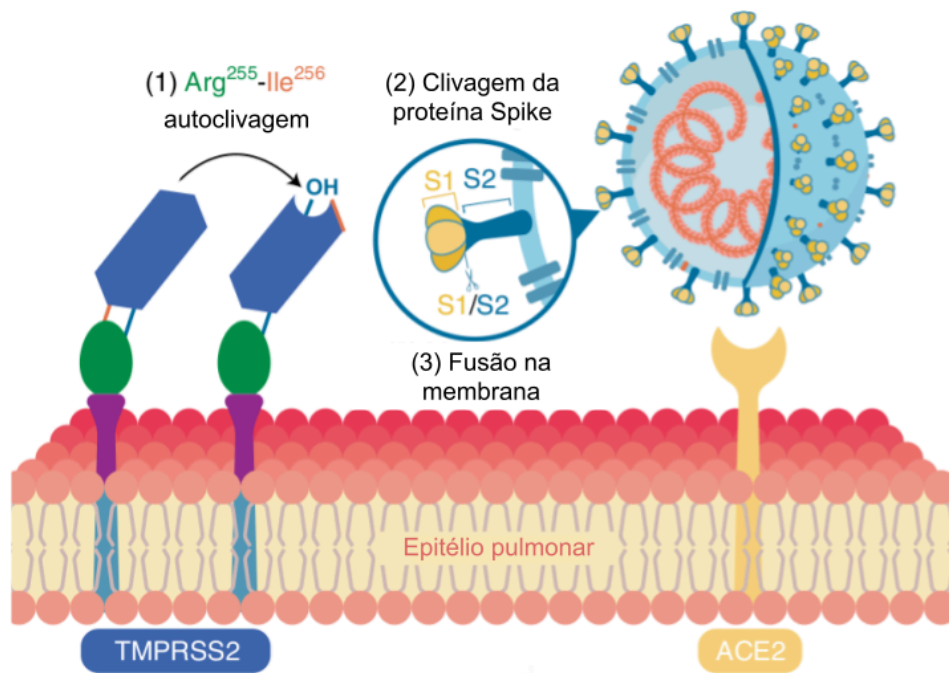


Figura 3. Representação da ação da TMPRSS2. Em (1) ela está sofrendo autoclivagem nos resíduos R255-I256 para que seja ativada; em (2) ocorre a clivagem da proteína Spike do SARS-CoV-2 em S1/S2; em (3) está ocorrendo a fusão do vírion com a membrana da célula do hospedeiro.

Fonte: Adaptado de Fraser, B.J. et al. (2022)

Além disso, outros coronavírus utilizam essa protease como porta de entrada na célula do hospedeiro e início da infecção. O SARS-CoV entra nas células através de duas vias distintas: uma é mediada por TMPRSS2 na superfície celular e a outra feita pelas catepsinas L e B, que são cisteíno-proteases (SIMMONS et al., 2005).

Simmons *et al.* (2013) mostraram que o Camostat, inibidor competitivo de serino protease, pode proteger com eficácia camundongos infectados com o SARS-CoV, evitando o óbito. Entretanto, nesse mesmo estudo, o tratamento com inibidores de serina e catepsina não conseguiu melhorar significativamente a sobrevivência em relação ao alcançado apenas com Camostat, indicando que a propagação *in vivo* de SARS-CoV é mediada preferencialmente

pela TMPRSS2. De maneira semelhante, as proteases envolvidas na clivagem do vírus MERS-CoV (*Middle East Respiratory Syndrome Coronavirus*) incluem TMPRSS2, furina e catepsina (SHEN et al., 2017). Shirato *et al.* (2013) mostraram que o tratamento com Camostat sozinho é tão eficaz quanto o Camostat + EST, um inibidor de catepsina, indicando um papel dominante do TMPRSS2 na entrada do MERS-CoV.

1.2.2 Vírus influenza

No caso do vírus Influenza, é utilizada a proteína Hemaglutinina (HA) para realizar a fusão com a célula do hospedeiro. Ela é uma proteína que possui duas subunidades funcionais, HA1 e HA2, sendo que ela é sintetizada como um precursor HA0 inativo para fusão, assim como a proteína Spike do coronavírus. Estudos mostram que camundongos *knockout* para TMPRSS2, ou seja, mutados para o gene, foram altamente tolerantes à infecção pelos vírus Influenza H7N9 (Influenza aviária) e H1N1 (Influenza suína), mostrando novamente seu papel no processo (TARNOW et al., 2014). A TMPRSS2 tem o papel de clivar a Hemaglutinina em fragmentos HA1 e HA2 (Figura 4), tendo como consequência a mudança de conformação da proteína e a exposição do sítio de fusão, fazendo com que ela consiga realizar a internalização do vírus pela via endocítica (SHEN et al., 2017).

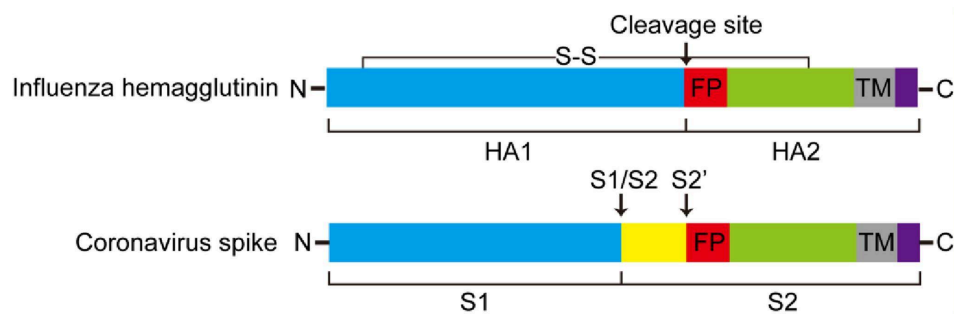


Figura 4. Estruturas da hemaglutinina do vírus Influenza e da proteína *spike* do coronavírus. Seus determinados locais de clivagem estão demarcados utilizando setas.

Fonte: Shen, L. W. et al. (2017)

1.2.3 Câncer de próstata

Além de seu envolvimento em infecções virais, a TMPRSS2 também tem sido associada ao câncer de próstata. Como ressaltado anteriormente, os membros da família das TTSPs precisam ser ativados proteoliticamente, seja pela própria protease ou por alguma outra externa. No caso do câncer de próstata, destaca-se o papel da matriptase, uma serino protease transmembrana expressa em tecidos epiteliais que necessita ser ativada pela TMPRSS2.

A matriptase está envolvida em processos fisiológicos, como a renovação tecidual e a cicatrização de feridas, no entanto, a sua superexpressão e hiperativação estão associadas a várias doenças, incluindo câncer. Segundo Ko *et al.* (2015), o processo se inicia através da ativação da TMPRSS2 pela via de sinalização do receptor de andrógeno (*Androgen Receptor - AR*), levando à ligação do AR ao elemento de resposta de andrógeno (*Androgen Response Element - ARE*) no promotor do gene TMPRSS2, resultando na sua expressão.

Por sua vez, a TMPRSS2 inicia uma cascata proteolítica para ativar a matriptase, levando à invasão celular, crescimento tumoral e metástase (Figura 5). Além disso, a matriptase realiza a degradação de componentes da matriz extracelular (*Extracellular Matrix - ECM*), como nidogênio-1 e laminina $\beta 1$. A degradação desses componentes leva à desestabilização da rede de ECM, tornando o microambiente permissivo à invasão celular e permitindo o crescimento e a metástase do câncer de próstata (WU *et al.*, 2010).

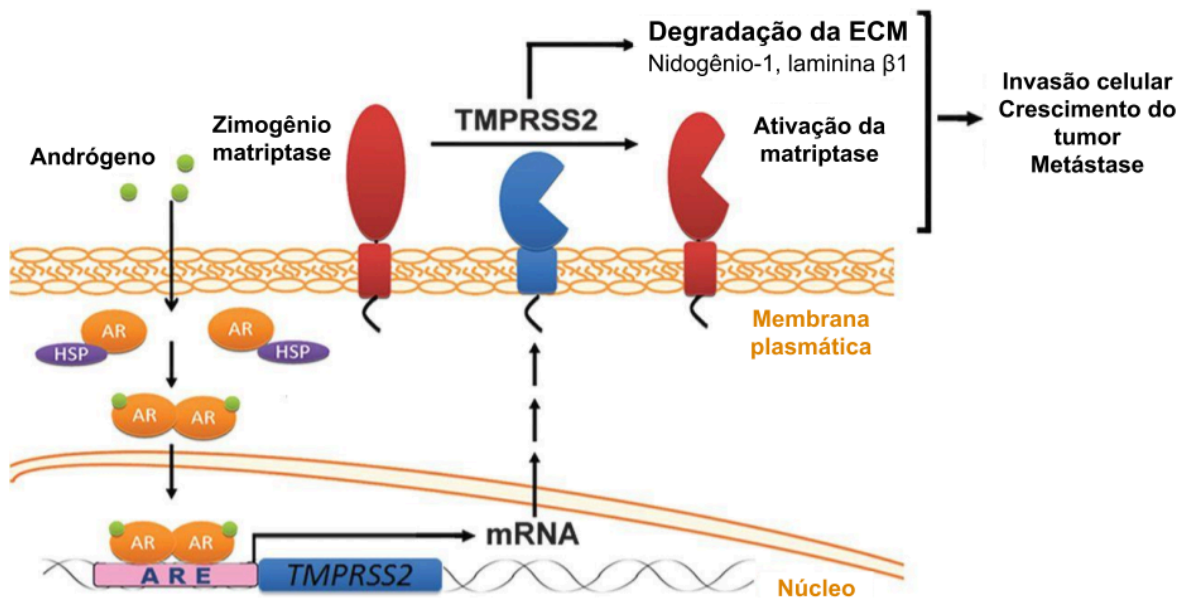


Figura 5. Processo de ativação da TMPRSS2 e suas consequências na progressão do câncer de próstata.

Fonte: Adaptada de Ko C.J., *et al* (2015)

Além disso, uma característica de alguns cânceres de próstata é a fusão gênica da 5'UTR do gene da TMPRSS2 com o gene ERG (*Erythroblast-Specific-Related gene*), um fator de transcrição oncogênico (THUNDERS; DELAHUNT, 2020). Essa fusão gênica leva à superexpressão do ERG e à ativação de vias de sinalização que promovem o crescimento das células cancerosas (GARCÍA-PERDOMO *et al.*, 2018). Foi observado por Zhou *et al.* (2017) que a prevalência dessa fusão é diferente dependendo da etnia da população, sendo maior na

Europa, seguido pela Ásia, e menor na África. Isso pode explicar a origem do câncer de próstata nas diferentes etnias. A TMPRSS2 emerge, assim, como uma molécula de grande importância em várias áreas da saúde, oferecendo oportunidades de intervenção terapêutica com a sua inibição, beneficiando não apenas o campo da virologia.

1.3 Medicamentos para COVID-19

A vacinação tem se destacado como um meio altamente eficaz de controle de doenças, sendo uma das principais ferramentas de prevenção em saúde pública. Contudo, a história nos ensina que a rápida evolução dos vírus, impulsionada por mutações genéticas, pode tornar as vacinas menos eficazes contra cepas emergentes. Assim como ocorreu com a pandemia do vírus Influenza H1N1 (gripe suína) (BERTRAM et al., 2010a), as mutações genéticas do vírus SARS-CoV-2 podem induzir respostas imunológicas menos protetoras, deixando até mesmo as populações vacinadas em risco (KONISHI, 2022; MALIK et al., 2022). Por exemplo, Cele et al. (2021) relataram que as novas variantes do SARS-CoV-2 são menos propensas à neutralização de anticorpos que as variantes iniciais. Um exemplo é a variante 501Y.V2, que apresenta substituições em E484K, K417N e N501Y no RBD e D215G, D80A, L18P, deleções em 242–244 e R246I no NTD (*N-Terminal Domain*).

Além disso, as lições aprendidas com a pandemia de H1N1 também se aplicam aos medicamentos antivirais direcionados às proteínas virais. Assim como a resistência à amantadina e ao oseltamivir emergiu em cepas do vírus Influenza H1N1 em 2009 (INOUE et al., 2010; LIU et al., 2014; WU et al., 2014), existe o risco de que cepas do SARS-CoV-2 possam se tornar resistentes a tratamentos antivirais específicos. Isso destaca a importância da pesquisa contínua e do desenvolvimento de novas abordagens terapêuticas que possam enfrentar a evolução constante dos vírus.

1.3.1 Potencial da inibição da TMPRSS2 como abordagem terapêutica

Diante do contexto apresentado, é fundamental considerar o papel de medicamentos direcionados a alvos hospedeiros, como a TMPRSS2, que podem oferecer vantagens na luta contra vírus mutantes. Essa protease representa um bom alvo para intervenção terapêutica tanto para bloquear infecções virais no início do ciclo de multiplicação quanto para desabilitar cânceres agressivos. Ela pode ser insensível a mutações de variantes devido ao tropismo conservado vírus-hospedeiro por essas proteases, ou seja, a proteína *Spike* tem uma preferência pela clivagem pela TMPRSS2, independente da variante (FRASER et al., 2022). Além disso, alvos dirigidos ao hospedeiro têm um baixo potencial de resistência, uma vez que

o local de clivagem que a TMPRSS2 atua na proteína *Spike* permaneceu muito conservado entre o MERS-CoV, SARS-CoV e variantes do SARS-CoV-2 (Figura 6).

| Virus | S1/S2 site | S2' site |
|----------------------|-----------------------|-------------------------|
| SARS-CoV-2 (Wuhan) | TNSP RRAR ,SVA | PSKPSK R ,SFIEDL |
| SARS-CoV-2 (Alpha) | TNSH RRAR ,SVA | PSKPSK R ,SFIEDL |
| SARS-CoV-2 (Beta) | TNSP RRAR ,SVA | PSKPSK R ,SFIEDL |
| SARS-CoV-2 (Delta) | TNS RRAR ,SVA | PSKPSK R ,SFIEDL |
| SARS-CoV-2 (Omicron) | TKSH RRAR ,SVA | PSKPSK R ,SFIEDL |
| SARS-CoV | S----LL R ,STS | PLKPTK R ,SFIEDL |
| MERS-CoV | TLTP RSVR ,SVP | TG RSAR ,SAIEDL |

Figura 6. Alinhamento das sequências de aminoácidos dos vírus envolvidas com a TMPRSS2. Os motivos de aminoácidos destacados em azul são clivados pela furina, em laranja pela TMPRSS2, e em rosa são provavelmente clivados pela TMPRSS2.

Fonte: Keller C., et al. (2022)

Estudos mostram que camundongos *knockout* para TMPRSS2 não apresentaram anormalidades fenotípicas óbvias, como morte, infertilidade ou sintomas visíveis de doença (S. KIM et al., 2006). A função fisiológica exata de TMPRSS2 *in vivo* permanece desconhecida (THUNDERS; DELAHUNT, 2020). Uma hipótese é de que a protease possa ter uma função altamente especializada e não vital, tornando-se aparente apenas no contexto de estresse, doença ou outro problema sistêmico (S. KIM et al., 2006). Dessa forma, a sua inibição torna-se uma alternativa terapêutica interessante para os casos apresentados. É importante ressaltar que, uma vez que é um alvo extracelular, os medicamentos que o têm como alvo podem não precisar atravessar a membrana plasmática, inibindo a TMPRSS2 sem precisar adentrar na célula. No entanto, devem ser extremamente seletivos para evitar a concentração de proteases no plasma responsáveis por cascatas de coagulação (FRASER et al., 2022).

1.3.2 Medicamentos disponíveis e suas limitações

Ao considerar o potencial da inibição da TMPRSS2 como uma abordagem terapêutica promissora no controle de infecções virais, é essencial examinar os medicamentos atualmente disponíveis para COVID-19 e as suas limitações. Como mencionado anteriormente, a TMPRSS2 desempenha um papel crucial na clivagem das proteínas virais, facilitando a entrada do vírus nas células hospedeiras. No entanto, a evolução constante dos vírus, semelhante ao que ocorreu com a pandemia de H1N1 e, mais recentemente, com o SARS-CoV-2, destaca a necessidade de abordagens terapêuticas versáteis e eficazes. Neste

contexto, exploraremos os medicamentos existentes que visam combater infecções virais, bem como suas limitações e desafios associados.

Como tratamento para COVID-19, já foram aprovados medicamentos baseados em anticorpos monoclonais e em antivirais. Em 7 de março de 2023, a ANVISA (Agência Nacional de Vigilância Sanitária) proibiu temporariamente o uso emergencial do Evusheld, composto por dois anticorpos monoclonais: o tixagevimabe e o cilgavimabe (ROCHA, 2023). Dois meses antes, a FDA tomou a mesma decisão, sob justificativa de que ele apresentava perda de eficácia contra as variantes do novo coronavírus então em circulação. Outro anticorpo monoclonal é o Sotrovimabe, cujo uso também foi suspenso nos Estados Unidos em 2022, sob a mesma justificativa, mas continua a ser adotado no Brasil. Ambos os medicamentos atuam na proteína *Spike* do vírus, região em constante evolução. Já como antivirais, no Brasil, há três disponíveis para tratamento: Nirmatrelvir/Ritonavir (Paxlovid), Remdesivir e Molnupiravir.

O Paxlovid é um medicamento composto por Nirmatrelvir, um inibidor da Mpro/3CLpro (*Main protease*) viral envolvida na replicação do SARS-CoV-2 (JIN et al., 2020), e Ritonavir, um inibidor que diminui o metabolismo do Nirmatrelvir, aumentando o tempo de circulação da droga, mecanismo utilizado na terapia de inibição do HIV (*Human Immunodeficiency Virus*) (HULL; MONTANER, 2011). Esse medicamento demonstrou sucesso na inibição do vírus em alguns estudos (AMANI; AMANI, 2023; ZHU et al., 2023), no entanto, há problemas com relação a efeitos de mutações.

Ullrich *et al.* (2020) mostraram que as mutações S284A, T285A e I286A na protease resultam em um aumento de 3,6 vezes na atividade catalítica. Os resultados sugerem que o Nirmatrelvir seria tão eficaz contra as variantes quanto é contra o vírus original, e que as atuais variantes do SARS-CoV-2 não teriam impacto na eficácia do medicamento. No entanto, Iketani *et al.* (2023) conduziram experimentos de passagem *in vitro* do vírus em presença de Nirmatrelvir para estudar a resistência ao medicamento, concentrando-se nas sete mutações pontuais mais comuns. Eles verificaram que, de fato, houve mutações que conferiam resistência e elas foram consistentes com as mutações clinicamente observadas em sequências de SARS-CoV-2 relatadas à iniciativa global COVID-19 CG GISAID (*Global Initiative on Sharing Avian Influenza Data*). Eles também observaram que as mutações encontradas em seus experimentos foram semelhantes às encontradas em outros estudos de resistência a inibidores da protease 3CLpro (HASHEMIAN et al., 2023).

Já o Molnupiravir apresenta um problema diferente. Um estudo recente de Sanderson *et al.* (2023) investigou possíveis consequências do uso do medicamento Molnupiravir. Este

medicamento, amplamente utilizado para tratamento de COVID-19, tem sido associado a mutações no genoma viral durante o processo de replicação. Ele age da seguinte forma: no corpo humano, o Molnupiravir é convertido em um análogo de nucleotídeo, molnupiravir trifosfato (MTP) (ROSENKE et al., 2021). MTP é incorporado no RNA durante a síntese da fita por RNA polimerases RNA-dependentes (*RNA dependent RNA polymerase* - RdRp), resultando em mutações deletérias ao vírus, na grande maioria dos casos.

No entanto, no estudo citado, eles mostraram que bancos de dados com sequências de SARS-CoV-2 contêm evidências de mutagênese não deletéria possivelmente causada pelo medicamento. Usando uma abordagem sistemática, eles descobriram uma classe específica com uma alta proporção de mutações G-A e C-T que aparece quase exclusivamente em sequências de 2022, após a introdução do tratamento. Embora as mutações C-para-T sejam relativamente comuns na evolução do SARS-CoV-2 (MCBROOME et al., 2021), as mutações G-para-A ocorrem com muito menos frequência. Os resultados de ensaios recentes e a aprovação de antivirais mais eficazes levaram desde então vários países a recomendarem contra a sua utilização, ao mesmo tempo que foram levantadas preocupações sobre a potencial atividade mutagênica. O Remdesivir possui o mesmo mecanismo de ação e, como dito anteriormente, não mostrou eficácia em alguns estudos (CHERA; TANCA, 2022).

Camostat é um medicamento aprovado no Japão para o alívio dos sintomas de pancreatite e para uso pós-operatório de refluxo esofágico (GIBO et al., 2005). O mecanismo de ação proposto contra o vírus SARS-CoV-2 envolve a inibição da atividade da protease TMPRSS2, assim como Nafamostat. Apesar do Nafamostat neutralizar muito bem a atividade da TMPRSS2, ele é não-seletivo e inibe serino-proteases *trypsin-like* envolvidas na coagulação, como a plasmina, FXa e FXIIa, assim como outras TTSPs (HITOMI; IKARI; FUJII, 1985; MUTO; IMAI; ASANO, 1995). Além disso, ele é biologicamente instável, como meia-vida de 8 minutos. O Camostat possui o problema de ser hidrolisado em dois produtos adicionais, cada um com potenciais inibitórios menores (HOFFMANN et al., 2021). Um estudo com 205 pacientes hospitalizados analisou a eficácia do Camostat na inibição viral, no entanto, não houve sucesso (GUNST et al., 2021). Uma das explicações propostas foi que a maioria dos pacientes incluídos já apresentavam sintomas 8 dias antes da inclusão no ensaio, fazendo com que a inibição da TMPRSS2 não fosse eficaz.

Em resumo, os medicamentos disponíveis para o tratamento de infecções virais enfrentam desafios significativos, desde a perda de eficácia contra variantes até a preocupação com mutações induzidas. Esses desafios ressaltam a necessidade de identificar compostos mais eficazes e com menos efeitos colaterais. Dessa forma, a transição para abordagens

terapêuticas baseadas em peptídeos, conhecidos por sua especificidade e menor propensão a efeitos colaterais em comparação com pequenas moléculas, surge como uma direção promissora.

1.3.3 Peptídeos como novas abordagens terapêuticas

Os medicamentos atuais são comumente baseados em pequenas moléculas (*small-molecules*), cuja importância é inegável, no entanto, algumas desvantagens relacionadas a essas moléculas criaram a necessidade de proposições terapêuticas alternativas. Dentre as vantagens do uso de pequenas moléculas, podemos citar os benefícios de melhor adesão do paciente e, devido ao seu pequeno tamanho, também conseguem penetrar facilmente na membrana celular para atingir alvos intracelulares (IMAI; TAKAOKA, 2006; LI; LIU; RAO, 2021). No entanto, seu uso também apresenta algumas desvantagens, incluindo seus efeitos adversos significativos e sua baixa especificidade em comparação com peptídeos. Por exemplo, algumas pequenas moléculas inibem não apenas o seu alvo, mas também outros alvos semelhantes, levando à citotoxicidade (SEHDEV, 2016; VUONG et al., 2019).

Para abordar a busca por soluções terapêuticas mais eficazes e com menor probabilidade de efeitos colaterais, os peptídeos surgiram como uma classe promissora de agentes terapêuticos. Eles representam uma classe com algumas propriedades de medicamentos de pequeno porte, bem como atributos de proteínas grandes e outros produtos biológicos (FOSGERAU; HOFFMANN, 2015).

Os peptídeos terapêuticos geralmente atuam como hormônios, fatores de crescimento, neurotransmissores, ligantes de canais iônicos ou agentes anti-infecciosos (WANG et al., 2022). Desde a descoberta da insulina como abordagem terapêutica para o *diabetes mellitus*, em 1921, mais de 80 medicamentos baseados em peptídeos foram aprovados para uma variedade de doenças, incluindo câncer, infecção pelo vírus da imunodeficiência humana (HIV), dor crônica e osteoporose (ANAND et al., 2023; BOURINET; ZAMPONI, 2017).

Dentre as vantagens do uso de peptídeos como abordagem terapêutica em comparação com as pequenas moléculas, podemos citar: maior afinidade com o alvo, levando a uma maior seletividade e menos efeitos colaterais fora do alvo (WANG et al., 2022); propriedades físico-químicas mais potentes, incluindo o seu maior tamanho e estrutura mais flexível (PETTA et al., 2016); capacidade de atingir alvos mais difíceis que apresentem desafios particulares devido à sua localização, conformação estrutural complexa ou características bioquímicas específicas (LEE et al., 2019); e menores taxas de toxicidade e interações

medicamentosas, uma vez que seus produtos metabólicos são aminoácidos (OTVOS; WADE, 2014).

Comparando-se os peptídeos às proteínas maiores, devido ao seu tamanho ser menor, há a possibilidade de se ter um custo de produção mais baixo (CRAIK; KAN, 2021); síntese mais rápida e maior capacidade de penetração nos tecidos (SHARMA et al., 2023) e menor chance de identificação como uma substância estranha pelo corpo, desencadeando uma resposta imune (LAU; DUNN, 2018).

Apesar de todas as vantagens, a abordagem utilizando peptídeos terapêuticos pode enfrentar também algumas desvantagens, como impermeabilidade da membrana e baixa estabilidade *in vivo*. Os peptídeos são geralmente incapazes de atravessar a membrana celular para atingir alvos intracelulares, uma vez que essa capacidade depende de múltiplos fatores, incluindo o tamanho do peptídeo e a composição de aminoácidos (WANG et al., 2022). Com relação à baixa estabilidade, as ligações amida de cadeias de aminoácidos podem ser facilmente hidrolisadas ou destruídas por enzimas após exposição ao meio, conferindo meia-vida curta e rápida eliminação *in vivo* (DIAO; MEIBOHM, 2013).

Para superar essas desvantagens, são utilizados métodos de química sintética para alterar algumas propriedades consideradas ruins para um medicamento. Pode-se modular através de modificações estruturais algumas características farmacocinéticas, como absorção, meia-vida, metabolismo e biodisponibilidade dos peptídeos (GOODWIN; SIMERSKA; TOTH, 2012). Nesse contexto, a utilização de peptídeos terapêuticos mantém-se como uma alternativa altamente promissora para o desenvolvimento de terapias mais eficazes e de menor incidência de efeitos colaterais. Com relação a esses estudos, é importante realizar uma predição das interações entre o alvo e o ligante por métodos *in silico*, utilizando uma abordagem de *docking* molecular. Dessa forma, é possível verificar possíveis efeitos e modos de ligação antes que se realize o ensaio *in vitro*.

1.4 Docking molecular como ferramenta para busca de peptídeos inibidores

A utilização do *docking* molecular emerge como uma poderosa e útil ferramenta na busca por novos inibidores de moléculas. Ele é uma abordagem computacional avançada que permite avaliar a interação entre moléculas, como peptídeos inibidores, e alvos específicos, como a proteína TMPRSS2. Essa técnica desempenha um papel crucial na identificação e desenvolvimento de candidatos a medicamentos da forma mais precisa e eficiente possível (PINZI; RASTELLI, 2019).

O *docking* molecular, ou atracamento molecular, geralmente é realizado prevendo

primeiro a orientação molecular de um ligante dentro de um receptor e, em seguida, estimando sua complementaridade através do uso de uma função de pontuação (KITCHEN et al., 2004). O método de realização do atracamento e o cálculo da pontuação dependem do algoritmo que está sendo utilizado, a depender da finalidade do seu uso. Até agora, foram desenvolvidas várias funções de pontuação baseadas em diferentes algoritmos e conceitos, que podem ser classificados em empíricos, baseados em conhecimento (*knowledge-based*) e baseados em campos de força (COUPEZ; LEWIS, 2006; GUEDES; PEREIRA; DARDENNE, 2018).

Com relação ao seu funcionamento, o *docking* molecular proteína-peptídeo pode ser dividido em três categorias: (I) baseado em *template*, que utiliza o conhecimento baseado em estruturas similares; (II) *docking* global, que realiza uma busca por toda a superfície da proteína para encontrar o sítio de ligação da proteína e o modo de ligação do peptídeo; (III) *docking* local, em que há conhecimento do centro ativo e, portanto, a ferramenta só realiza a busca para saber a conformação do peptídeo.

Graças ao seu poder e velocidade de predição, o *docking* molecular é utilizado para muitos fins, incluindo: identificar locais de ligação a proteínas nos quais os ligantes podem se ligar; encontrar novos alvos moleculares de ligantes conhecidos; analisar potenciais reações adversas a medicamentos (LABUTE et al., 2014), entre outros.

1.4.1 Limitações do docking molecular

Apesar de todas as suas utilidades, essa técnica apresenta algumas limitações. Particularmente, em comparação com o *docking* proteína-ligante e proteína-proteína, o *docking* proteína-peptídeo enfrenta alguns desafios. Primeiro, em muitos casos, a informação do sítio de ligação de peptídeos não está disponível. Dessa forma, é necessária a realização de um *docking* global. Além disso, em comparação com pequenas moléculas e proteínas, os peptídeos são muito mais flexíveis, representando um desafio para geração das conformações corretas (ZHOU et al., 2018). Outro problema tem como causa a pouca quantidade de dados experimentais de estruturas de complexos proteína-peptídeo disponíveis, fazendo com que se tenha poucas informações sobre as moléculas e impedindo a comparação do resultado com dados experimentais, diminuindo a acurácia *in silico* (CIEMNY et al., 2018).

Além disso, para avaliar os modelos gerados, uma estratégia comumente usada consiste no uso do *score* ou pontuação que as ferramentas fornecem. No entanto, esse método pode ser prejudicado, possuindo acurácias ruins em relação à classificação de poses devido ao fato de que as funções de pontuação são apenas aproximadas (COUPEZ; LEWIS, 2006;

GILSON; ZHOU, 2007). Além disso, os algoritmos utilizados empregam uma série de funções matemáticas com aproximações que não consideram com precisão alguns elementos termodinâmicos da energia de ligação, sendo uma aproximação da energia livre de ligação real (PINZI; RASTELLI, 2019). O algoritmo genético, por exemplo, é baseado em princípios de seleção natural e evolução. Ele gera e avalia conformações de ligantes que se encaixam no sítio de ligação do receptor e utiliza uma função de pontuação para avaliar a aptidão de cada solução candidata gerada. Ela leva em consideração a complementaridade geométrica, interações intermoleculares e outras propriedades relevantes (DIAS; DE AZEVEDO, 2008). Outro exemplo é o algoritmo de Monte Carlo, que tenta posicionar o ligante dentro do sítio de ligação do receptor através de muitas posições e rotações aleatórias. A cada movimento aleatório, uma função de energia é calculada para avaliar a aptidão da solução candidata. A função de energia leva em consideração principalmente a complementaridade geométrica e interações intermoleculares. Após cada movimento aleatório, uma minimização de energia é aplicada para melhorar a conformação do ligante, e o processo é repetido até que uma solução satisfatória seja encontrada (DIAS; DE AZEVEDO, 2008).

O artigo do HPEPDOCK (ZHOU et al., 2018), ferramenta de *docking* molecular muito utilizada no *docking* proteína-peptídeo, mostrou que a maioria dos modelos gerados com alta taxa de sucesso estavam classificados a partir do top10 de melhores modelos, tanto para o HPEPDOCK quanto para o outro programa utilizado para comparação de desempenho, o HADDOCK (DOMINGUEZ; BOELEN; BONVIN, 2003) (Figura 7). O mesmo pode ser visto na Figura 8, comparando o desempenho do HPEPDOCK com a ferramenta pepATTRACT (DE VRIES et al., 2017). Ambas as figuras mostram a taxa de sucesso e o número médio de acertos por complexo nas previsões de modo de ligação pelo servidor HPEPDOCK em diferentes benchmarks de *docking* de proteína-peptídeo.

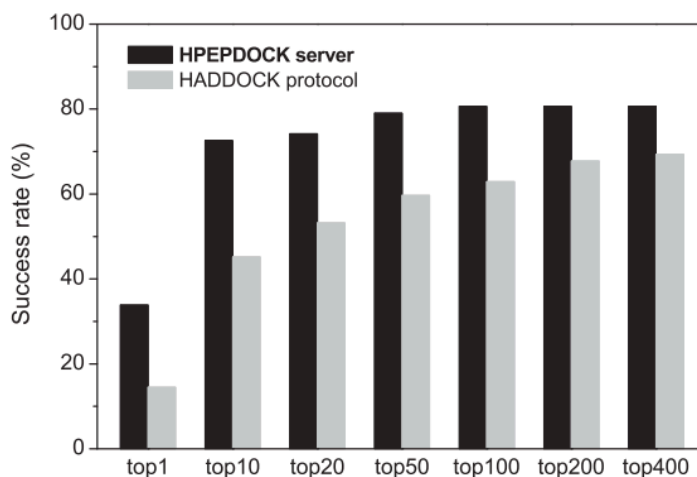


Figura 7. Taxa de sucesso do protocolo de docking local do servidor HPEPDOCK em um benchmark de 62 complexos proteína-peptídeo não ligados, em comparação com o protocolo de docking de peptídeo HADDOCK. O gráfico mostra que os modelos de *docking* gerados após o top10 apresentaram maiores taxas de sucesso.

Fonte: Zhou, P. *et al.* (2018)

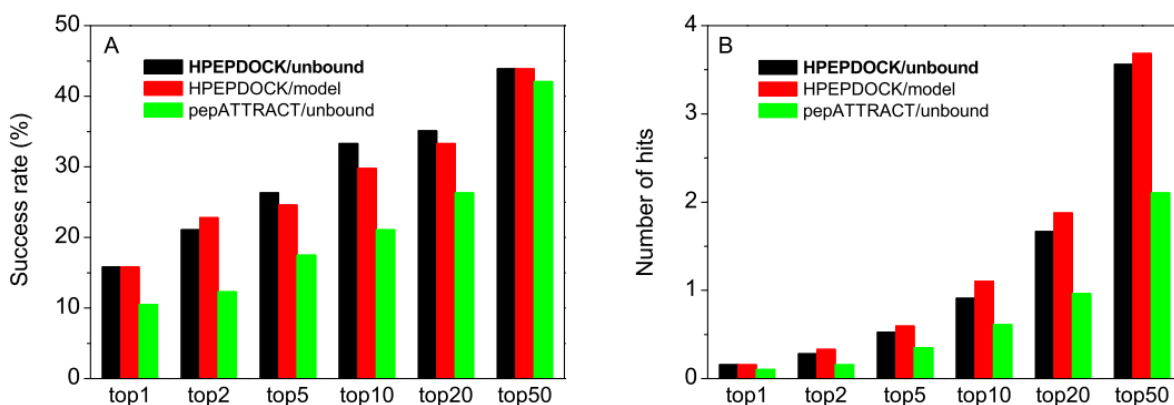


Figura 8. Comparação da taxa de sucesso e do número médio de acertos por complexo entre o servidor HPEPDOCK e o servidor pepATTRACT em um benchmark de 57 complexos proteína-peptídeo não ligados. Os gráficos mostram que os modelos de *docking* gerados após o top10 apresentaram maiores taxas de sucesso.

Fonte: Zhou, P. *et al.* (2018)

Ciemny *et al.* (2018) também mencionam que a identificação bem-sucedida do modelo mais preciso entre um grande número de resultados de *docking* molecular ainda é um desafio. Isso significa que, embora essas ferramentas possam gerar muitos modelos, pode ser difícil determinar qual é o mais preciso. Além disso, o artigo menciona que, na maioria dos casos, os modelos classificados como os melhores pela pontuação de energia são de qualidade inferior aos modelos mais precisos presentes nos resultados. Isso sugere que a pontuação de

energia pode não ser suficiente para identificar o modelo mais preciso e que outras abordagens podem ser necessárias para melhorar a precisão da previsão.

Devido a essas imprecisões, a inspeção visual dos modos de ligação é uma rotina crucial no processo de tomada de decisão sobre o melhor modelo. Alguns dos critérios utilizados envolvem a complementaridade estérica de um composto e seu sítio de ligação e as ligações de hidrogênio entre o ligante e a proteína (devido à sua força e geometria bem definidos, sendo facilmente identificados) (FISCHER et al., 2021).

A inspeção visual pode não ser a melhor das alternativas para escolha das poses geradas devido a limitações da intervenção humana, como o número de poses que podem ser avaliadas com precisão e os diferentes níveis de experiência. Em um estudo de Fischer *et al.* (2021), pesquisadores foram desafiados a encontrar visualmente a pose nativa de algumas estruturas, para analisar os limites da inspeção visual humana. O resultado foi que, apesar de dois de três pesquisadores acertarem a conformação nativa, a inspeção visual deve ser usada em conjunto com outras técnicas computacionais e experimentais para maximizar sua eficácia. Dessa forma, o uso do *docking* molecular apresenta-se como uma alternativa fundamental no estudo *in silico* de novos fármacos. Este método, somado a outras metodologias, poderão, em conjunto, classificar melhor os modelos gerados.

1.5 Justificativa do trabalho

As pandemias de influenza H1N1 e COVID-19 destacaram a necessidade de desenvolver terapias específicas que possam combater a disseminação dos vírus e melhorar o prognóstico dos pacientes. Nesse contexto, o reposicionamento de peptídeos como potenciais inibidores das Serino Proteases Transmembrana 2 (TMPRSS2) surge como uma iniciativa relevante e oportuna.

É importante ressaltar que, até o momento, a maioria dos tratamentos terapêuticos disponíveis concentrou-se em pequenas moléculas, deixando um espaço para a exploração de peptídeos como candidatos a inibidores. Os peptídeos possuem várias vantagens, incluindo alta especificidade e afinidade pelo centro ativo da enzima, propriedades físico-químicas vantajosas, menor probabilidade de efeitos colaterais e menor risco de resistência viral. Dessa forma, o uso de peptídeos como inibidores representa uma estratégia promissora para o desenvolvimento de terapias antivirais mais seguras e eficazes.

Considerando esses aspectos, a identificação de peptídeos capazes de inibir a TMPRSS2 pode potencialmente contribuir para o tratamento da COVID-19 e de outras doenças, demonstrando a relevância desse trabalho. A pesquisa não apenas aborda uma lacuna

na investigação da TMPRSS2 como alvo terapêutico, mas também demonstra a importância de explorar abordagens diferentes para essa finalidade, como o uso de peptídeos.

2. Objetivos

2.1 Objetivo geral

Identificar e caracterizar peptídeos como potenciais inibidores da Serino Protease Transmembrana 2 (TMPRSS2), visando o desenvolvimento de novas abordagens terapêuticas para o combate de infecções virais, incluindo a COVID-19.

2.2 Objetivos específicos

- Coletar dados sobre a estrutura e mecanismo da TMPRSS2.
- Selecionar potenciais peptídeos inibidores da serino protease em bases de dados de complexos proteína-peptídeo.
- Criar métodos computacionais para coletar, filtrar e identificar os melhores ligantes a partir da interação proteína-peptídeo.
- Otimizar os ligantes escolhidos, melhorando sua performance.
- Sintetizar os peptídeos e testar a sua citotoxicidade em células.

3. Materiais e métodos

A Figura 9 resume a metodologia adotada neste trabalho. A metodologia foi dividida em oito etapas: (1) levantamento bibliográfico e análise da estrutura da TMPRSS2; (2) obtenção dos peptídeos candidatos; (3) filtragem das sequências promissoras; (4) *docking* molecular; (5) filtragem das melhores poses geradas no *docking* molecular; (6) análise visual das poses selecionadas; (6) análise da flexibilidade dos modos de ligação; (7) otimização dos peptídeos; (8) síntese dos peptídeos.

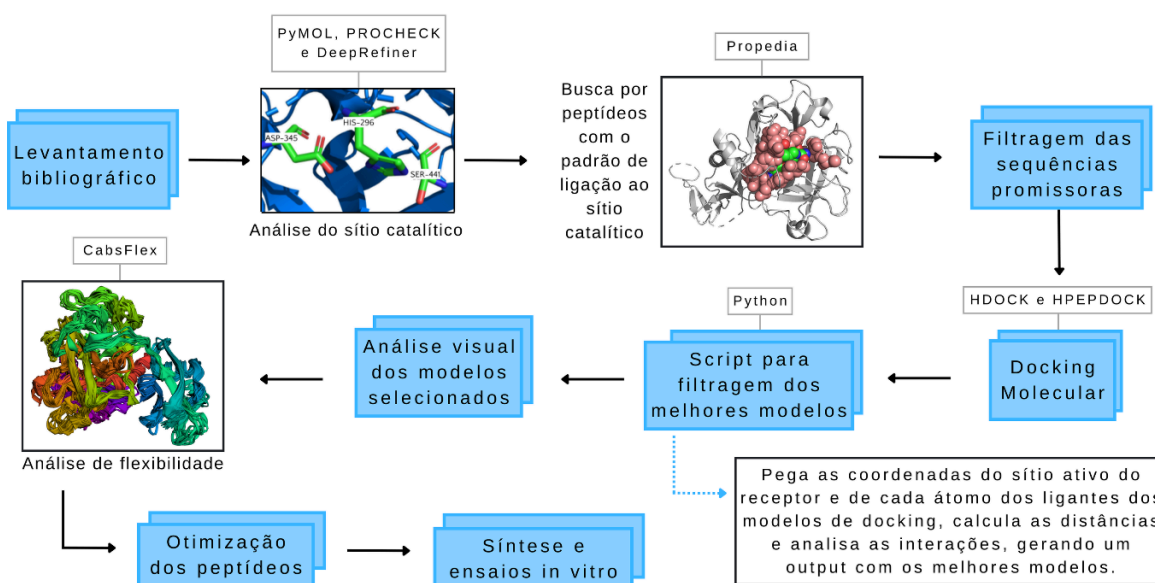


Figura 9. Esquema metodológico.

Fonte: próprio autor.

3.1 Levantamento bibliográfico e análise da estrutura da TMPRSS2

Foi realizado um levantamento bibliográfico para se obter um panorama completo sobre o projeto de pesquisa em geral, as metodologias comumente utilizadas e estudos que já foram publicados na área. Esse levantamento foi realizado utilizando as plataformas PubMed (<https://pubmed.ncbi.nlm.nih.gov/>), SciELO (<https://scielo.org/pt/>), Google Scholar (<https://scholar.google.com/>) e ResearchGate (<https://www.researchgate.net/>), buscando por artigos científicos, livros, revisões da literatura e outros recursos relevantes.

Utilizando as informações obtidas no levantamento bibliográfico, a estrutura da TMPRSS2 disponível no *Protein Data Bank* (PDB; <http://pdb.org>; BERMAN et al., 2000) (PDB ID: 7MEQ; resolução 1.95 Å; FRASER et al., 2022) foi escolhida. Optou-se por essa

estrutura por ser a única resolvida experimentalmente disponível naquele momento, além de possuir uma boa resolução. Utilizando o *software* PyMOL (The PyMOL Molecular Graphics System, Versão 1.2r3pre, Schrödinger, LLC.), a estrutura escolhida foi analisada visualmente, com foco em sua tríade catalítica (S441-H296-D345) (Figura 10).

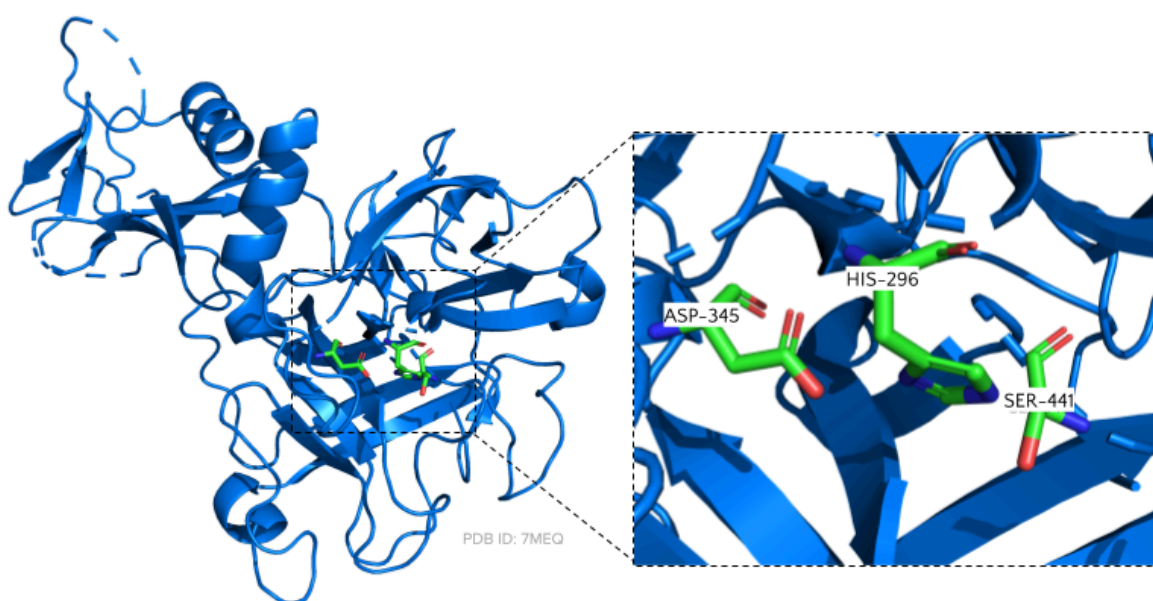


Figura 10. Estrutura da serino protease TMPRSS2 (PDB ID: 7MEQ). Foco em seu sítio ativo (S441-H296-D345), colorido por elementos (carbono em verde, oxigênio em vermelho, nitrogênio em azul). Figura gerada utilizando o software PyMOL.

Fonte: próprio autor.

Por fim, foi realizada uma análise da flexibilidade do sítio ativo do receptor, uma vez que a maioria das estruturas de proteínas estão representadas como estruturas estáticas (resolvidas por Cristalografia de Raio-X). Optamos pelo uso do iMODS (LÓPEZ-BLANCO et al., 2014), uma ferramenta que realiza Análise de Modos Normais (NMA - *Normal Mode Analysis*), devido à falta de recursos computacionais de alto desempenho e da carência de mão de obra qualificada para a realização de dinâmica molecular. Além disso, embora a dinâmica molecular seja uma técnica de simulação mais realista, optamos pelo NMA devido à sua maior eficiência computacional, o que nos possibilita analisar sistemas de maior escala e complexidade.

A NMA é utilizada para representar padrões de movimento coletivo que a proteína pode realizar em torno de sua conformação de equilíbrio, descrevendo pequenas oscilações físicas (BAUER; PAVLOVIĆ; BAUEROVÁ-HLINKOVÁ, 2019). Junto à NMA, essa

ferramenta utiliza um sistema de coordenadas internas baseado em um modelo de mola-haste, onde cada átomo é representado por uma esfera e as ligações químicas são representadas por molas, levando em consideração o tipo de átomo e o tipo de ligação. A energia potencial do sistema é dada pela soma das energias de ligação, ângulo e torção e nas distâncias entre os átomos, que são determinadas a partir da estrutura da molécula.

Segundo os criadores, o iMODS calcula os modos normais de vibração em coordenadas internas a partir da matriz de Hessian, uma matriz simétrica que mostra as curvaturas das superfícies de energia potencial da proteína em torno de sua conformação de equilíbrio. Cada modo normal é caracterizado por uma frequência e um vetor de deslocamento, que descreve a direção e a amplitude da vibração. Como resultado, a ferramenta retorna as frequências de vibração e os deslocamentos atômicos, utilizando fator B do NMA (B-factors) como uma medida de flexibilidade. O fator B é uma medida da amplitude das vibrações térmicas dos átomos em uma molécula e é expresso em unidades de Angstroms quadrados (\AA^2). Quanto maior o valor, maior a flexibilidade.

Dentre os parâmetros disponíveis, estavam: CA (carbono alfa), em que cada resíduo é representado por um único ponto, tornando a simulação mais rápida; C5, que inclui N, C, O e CA do resíduo protéico, permitindo uma simulação um pouco mais realista; Ha (Heavy atoms), em que átomos mais pesados são mantidos e hidrogênios são removidos, permitindo detalhamento intermediário mas reduzindo a complexidade da simulação. Para a nossa análise, foi utilizado o "CA". O uso de CA (carbono alfa) como parâmetro é recomendado em situações em que a análise dos movimentos coletivos de uma proteína é suficiente para entender as mudanças conformacionais relevantes. Isso ocorre porque o CA é um átomo comum a todos os aminoácidos e está localizado próximo ao centro da proteína, o que o torna uma boa representação do movimento global da proteína.

3.2 Obtenção dos peptídeos candidatos

A partir da estrutura escolhida (PDB ID: 7MEQ) e utilizando o software PyMOL, foram selecionados os resíduos que estavam em até 6 angstroms (\AA) de distância do sítio catalítico. Esse valor foi estabelecido pois a faixa de até 6\AA é considerada relevante para interações entre átomos, como ligações de hidrogênio, interações iônicas e interações hidrofóbicas, sendo a distância máxima entre esses tipos de ligações (BICKERTON; HIGUERUELO; BLUNDELL, 2011; MARTINS et al., 2021). Os resíduos selecionados foram então submetidos à ferramenta online Propedia (MARTINS et al., 2021) (<http://bioinfo.dcc.ufmg.br/propedia/>).

Propedia é um banco de dados de busca e visualização de complexos proteína-peptídeo baseado em um algoritmo híbrido de agrupamento, permitindo a comparação dessas moléculas através de sequências, estruturas e sítios de ligação. Ao utilizar como entrada os resíduos próximos do sítio ativo, a ferramenta realiza uma busca por estruturas de complexos cujos receptores possuam sítios de ligação com resíduos similares, em posições similares às selecionadas. Além disso, é possível selecionar na busca entre as opções “*Only CCD*” ou “*Whole database*”. Essas opções dizem respeito à ferramenta realizar a busca utilizando os complexos contidos no “*clusterable complex dataset*” (CCD), composto por 5.971 complexos, ou no bancos de dados inteiro, composto por 19.813 complexos. Nessa pesquisa, optou-se pela opção “*Only CCD*”, permitindo reduzir a redundância dos dados, uma vez que os *clusters* (grupos) são organizados por semelhanças químicas entre as moléculas.

3.2.2 Filtragem das sequências

Foi realizada uma filtragem das sequências com o objetivo de selecionar apenas aquelas que tivessem mais chances de serem peptídeos inteiros (e não pedaços de peptídeos) e de serem inibidores da protease em estudo. Para isso, dois critérios de filtragem foram definidos. Primeiro, a sequência deveria dispor de pelo menos oito resíduos de aminoácidos, isso está relacionado à natureza das assinaturas do MEROPS (*The Peptidase Database*, <https://www.merops.ac.uk/>), uma ferramenta amplamente reconhecida para a classificação e análise de proteases e seus inibidores. No contexto do MEROPS, as assinaturas são essenciais para descrever as características das proteases. As proteases possuem preferências de resíduos específicos em determinadas regiões dos ligantes para que reconheçam o local de clivagem. Sequências muito curtas, com menos de oito resíduos, podem não conter as informações cruciais necessárias para caracterizar uma assinatura de maneira precisa e significativa.

A segunda filtragem consistiu em realizar um alinhamento local par-a-par (*pairwise alignment*) entre cada um dos peptídeos remanescentes e a aprotinina. A aprotinina é um medicamento anticoagulante e conhecido inibidor competitivo de serino-protease, que possui um loop (Lys-15) que se liga à enzima, resultando na inibição (ZHIRNOV; KLENK; WRIGHT, 2011). Há estudos *in vitro* (BOJKOVA et al., 2020), *in vivo* (IVASHCHENKO et al., 2022) e clínico (REDONDO-CALVO et al., 2022) que mostraram sucesso desse medicamento em inibir a TMPRSS2 e, conseqüentemente, reduzir a infecção pelo SARS-CoV-2. Assim, o alinhamento foi realizado utilizando a ferramenta EMBOSS Water (MADEIRA et al., 2022), e a sequência utilizada foi:

“RPDFCLEPPYTGPCKARIIRYFYNAKAGLCQTFVYGGCRAKRNNFKSAEDCMRTCG GA”.

Essa ferramenta utiliza o alinhamento local através do algoritmo *Smith-Waterman* e foi escolhida para encontrar regiões de alta similaridade entre as duas sequências, uma vez que o algoritmo considera todas as possibilidades de *gaps* e substituições de bases ou aminoácidos. Considerando que os peptídeos possuem um baixo número de resíduos de aminoácidos em suas sequências, os alinhamentos que resultassem em pelo menos 3 correspondências com a região da alça de inibição da aprotinina passariam para a próxima etapa. Além disso, foi considerada também a especificidade da protease, de forma que todas as sequências deveriam ter uma Lisina ou Arginina na região. Essa etapa filtrou as sequências de um total de 127 para 11 remanescentes.

3.4 Docking molecular

Para a análise de ligação entre cada uma das 11 moléculas selecionadas na etapa anterior e a TMPRSS2, a estrutura 7MEQ e as 11 sequências foram submetidas ao *docking* molecular. Foram utilizadas as seguintes ferramentas: HPEPDOCK (ZHOU et al., 2018) [<http://huanglab.phys.hust.edu.cn/hpepdock/>], para as duas sequências menores (PDB ID: 4ZKN e PDB ID: 5NHU) e HDOCK (YAN et al., 2017) [<http://hdock.phys.hust.edu.cn/>] para as nove sequências restantes.

HPEPDOCK é um servidor *web* para *docking* proteína-peptídeo através de algoritmo hierárquico que realiza tanto o *docking* global quanto local (com especificação do sítio ativo do receptor). Por padrão, ele considera a flexibilidade de um peptídeo gerando primeiro um conjunto de 1.000 conformações usando o programa MODPEP e, em seguida, acoplando as conformações peptídicas geradas à molécula receptora. Quanto mais conformações peptídicas forem geradas para um peptídeo, mais extensivamente será considerada a flexibilidade do peptídeo, mas mais tempo levará o trabalho de acoplamento. Ele foi escolhido por não precisar especificar os resíduos de ligação, além de ser uma das melhores ferramentas de *docking* proteína-peptídeo disponíveis. Weng et al. (2020) realizaram um estudo de comparação de 14 programas de *docking* proteína-peptídeo e o HPEPDOCK foi considerado o segundo melhor em todas as análises para análise global, dentre elas: *docking* utilizando estruturas *near-native* e de média e alta qualidades, utilizando *subsets* de dificuldade fácil, médio e difícil e considerando diferentes tamanhos de peptídeos.

Já o HDOCK foi utilizado para as nove sequências contendo mais de 30 resíduos de aminoácidos, uma vez que o HPEPDOCK possui um limite no número de resíduos da sequência. Ele é um servidor *web* para *docking* proteína-proteína através de algoritmo de *docking* híbrido que também aceita como entrada a sequência do ligante. Essa ferramenta utiliza modelagem baseada em *template* e *docking* livre, que utiliza um algoritmo de busca global para encontrar a melhor conformação de ligação. Ela foi desenvolvida pelo mesmo laboratório do HPEPDOCK, no entanto, realiza o *docking* molecular em moléculas com tamanho acima de 30 aminoácidos.

3.4.2 Re-docking HPEPDOCK/HDOCK

Para avaliar o desempenho das ferramentas de *docking* molecular utilizadas neste estudo, foi realizado o *re-docking* de dez estruturas de complexos proteína-peptídeo disponíveis no PDB, com suas informações disponíveis na Tabela 1. O *re-docking* consiste no *docking* de estruturas de complexos já resolvidas experimentalmente, seguido da comparação do resultado com a estrutura tridimensional cristalográfica, e então verificar se a ferramenta conseguiu reproduzir o complexo fielmente ao original. As cinco primeiras estruturas (em azul) foram testadas na ferramenta HPEPDOCK e as cinco últimas (em amarelo) na ferramenta HDOCK, de acordo com o tamanho de suas sequências.

Tabela 1. Estruturas disponíveis no PDB utilizadas para a etapa de *re-docking*. As cinco primeiras estruturas (em azul) foram montadas utilizando o HPEPDOCK, e as cinco últimas utilizando o HDOCK (em amarelo). A primeira coluna mostra o ID da molécula no banco de dados, e a segunda coluna mostra o nome do complexo.

| PDB ID | Molécula |
|--------|--|
| 1J2X | RAP74 C-terminal domain complexed with FCP1 C-terminal peptide |
| 2IHS | B30.2/SPRY domain of GUSTAVUS in complex with a 20-residue VASA peptide |
| 2NM1 | BoNT/B in complex with its protein receptor |
| 2PEH | UHM domain of human SPF45 in complex with SF3b155-ULM5 |
| 2PV2 | SurA first peptidyl-prolyl isomerase domain complexed with peptide |
| 2QOS | Complement protein C8 in complex with a peptide containing the C8 binding site on C8 |
| 4DJ9 | Human vinculin head domain Vh1 (residues 1-258) in complex with the talin vinculin binding site 50 (VBS50, residues 2078-2099) |
| 4ETO | S100A4 in complex with non-muscle myosin-IIA peptide |
| 4HTP | DBD domain of human DNA ligase IV bound to Artemis peptide |
| 5F67 | An exquisitely specific PDZ/target recognition revealed by the structure of INAD PDZ3 in complex with TRP channel tail |

3.5 Filtragem dos melhores modelos

Como descrito anteriormente, dentre os desafios do *docking* molecular, podemos incluir a seleção da melhor pose gerada, uma vez que esse experimento gera um grande número de resultados, tornando a escolha visual um processo trabalhoso. Dessa forma, essa etapa visou selecionar as melhores poses, dentre as 100 adquiridas na etapa de *docking* molecular, para cada ligante, com base em métricas de distância do sítio ativo.

Assim, para evitar a análise manual da distância dos átomos dos resíduos ao sítio catalítico de cada modelo de *docking*, para cada peptídeo testado, foi criado um *script* utilizando a linguagem de programação *Python* para uma filtragem final dos modelos de *docking*.

O *script* foi projetado para ser executado após a aquisição dos arquivos e consiste em: abrir todos os arquivos de *docking* simultaneamente, selecionar apenas as linhas referentes ao *backbone* da molécula, ou seja, os átomos de carbono (C), oxigênio (O), nitrogênio (N) e carbono alfa (CA), criar uma entrada para cada modelo com as informações necessárias e extrair as coordenadas de cada átomo (Figura 11, em laranja). Ao mesmo tempo (Figura 11, em azul), o *script* abre o arquivo do receptor, captura as coordenadas do sítio ativo especificado pelo usuário no *script*, e então calcula a distância Euclidiana (Equação 1) do sítio ativo para cada um dos átomos. É importante ressaltar que, ao focar no *backbone*, sem a introdução da cadeia lateral, a análise se concentra nos aspectos fundamentais da interação e na geometria geral da ligação proteína-peptídeo. Além disso, ignorar essas cadeias simplifica a análise, tornando-a mais eficiente e facilitando a compreensão das interações centrais. Em experimentos de *docking*, onde a análise de muitos modelos é necessária, a simplificação pode acelerar significativamente o processo.

No caso da TMPRSS2, o sítio ativo foi especificado pelo átomo de número 2.097, o oxigênio da S441 (principal resíduo do seu sítio ativo), que foi escolhido por ser o mais pesado da cadeia lateral da serina, definindo a posição do centro do sítio ativo. Além disso, o oxigênio na serina desempenha um papel importante, pois está envolvido nas interações de ligação durante o processo de catálise. A serina atua como um nucleófilo, atacando o carbono carbonílico do substrato, o que leva à quebra da ligação peptídica. O oxigênio na serina forma ligações covalentes temporárias com os substratos, facilitando a reação enzimática (NEITZEL, J. J., 2010). A coordenada tridimensional desse átomo é $X = -6.938$, $Y = -4.038$ e $Z = 16.322$.

$$dist = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}$$

Equação 1. Equação representando o cálculo da Distância Euclidiana.

Além disso, a partir dessas informações, o *script* também analisa o tipo de ligação que o resíduo de aminoácido do ligante faz com o do sítio ativo, baseado nas definições do nAPOLI (FASSIO et al., 2020), um servidor *web* de análise de interações proteína-ligante conservadas, em larga escala. Para isso, ele reconhece se o aminoácido é hidrofóbico, aromático, acceptor, doador, positivo ou negativo, e isso é possível devido a um dicionário de contatos contendo essas informações. A partir dessa identificação, ele vai utilizar a informação das distâncias dos átomos, já calculadas anteriormente, e verificar se condiz com as distâncias de determinados tipos de ligação, mostradas na Tabela 2.

Tabela 2. Tipos de ligação identificados através do *script*. Na segunda coluna, estão indicadas as distâncias mínimas e máximas para que ocorra a ligação entre dois átomos e, na terceira coluna, os tipos de átomos necessários para realizar a ligação.

| Tipo de ligação | Distância | Ligação feita entre |
|-----------------------|-----------------------|---|
| Aromática | 2Å < distância < 4Å | Aromático e aromático |
| Ligação de hidrogênio | distância < 3.9Å | Acceptor e doador |
| Hidrofóbica | 2Å < distância < 4.5Å | Hidrofóbico e hidrofóbico |
| Repulsiva | 2Å < distância < 6Å | Positivo e positivo Negativo e negativo |
| Atrativa | 2Å < distância < 6Å | Positivo e negativo |

Por último, é gerado como saída um arquivo contendo as melhores poses (Figura 12), baseado nas menores distâncias, juntamente com as informações do seu *docking score* (pontuação), tipo de ligação, e a resposta se ele é ou não um bom ligante, considerando as métricas analisadas. O *script* foi construído para rodar via terminal, dentro da pasta contendo os arquivos a serem analisados. Um resumo de como ele funciona é mostrado na Figura 11.

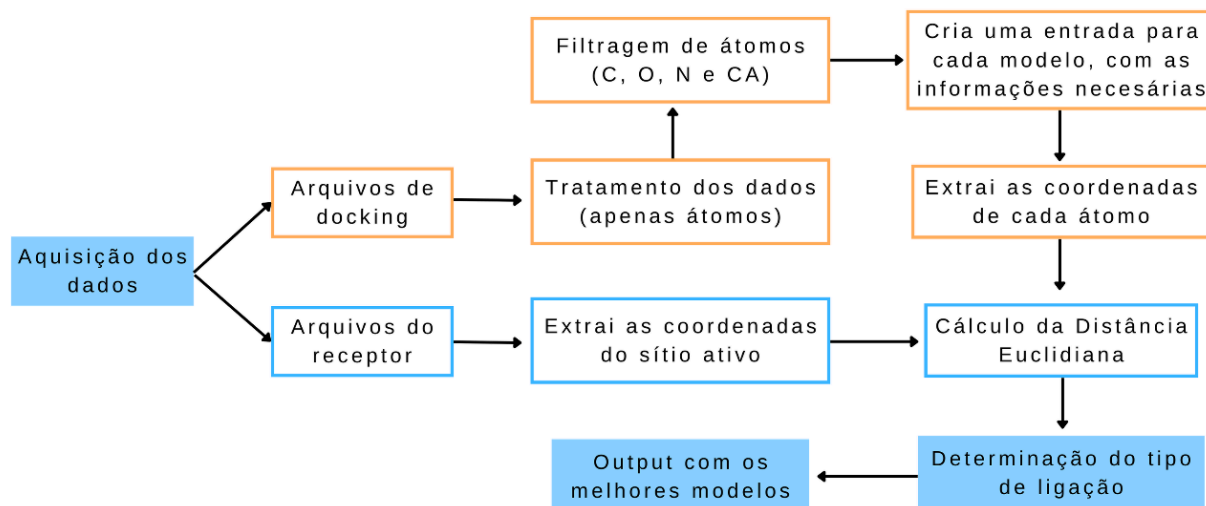


Figura 11. Funcionamento do *script* desenvolvido no projeto. Os retângulos com borda laranja indicam etapas realizadas utilizando os arquivos do *docking*, os retângulos com borda azul indicam etapas utilizando os arquivos do receptor, retângulos preenchidos em azul indicam etapas comuns a ambos os arquivos.

Fonte: próprio autor.

| aminoacid | ligand | remark_score | coordinates | distance | Is it a good candidate? | bond_type |
|-----------|--------|--------------|--------------------------------|--------------------|-------------------------|-----------|
| ALA:O | 41 | -117.67 | [-9.510', '-2.916', '15.249'] | 3.0042298513928656 | true | - |
| ALA:CA | 33 | -123.31 | [-8.237', '-1.935', '14.447'] | 3.102520749326264 | true | - |
| ALA:O | 44 | -114.32 | [-9.943', '-2.806', '14.915'] | 3.5394205740488087 | true | - |
| ALA:O | 96 | -85.22 | [-9.796', '-4.987', '14.301'] | 3.6267348951915404 | true | - |
| ALA:O | 95 | -85.26 | [-9.149', '-1.148', '16.377'] | 3.6391820509559563 | true | - |
| ALA:O | 73 | -100.49 | [-9.952', '-5.192', '14.487'] | 3.712564746910146 | true | - |
| ALA:C | 41 | -117.67 | [-10.471', '-3.645', '14.990'] | 3.7961509453655817 | true | - |
| PRO:O | 97 | -82.28 | [-10.553', '-2.777', '16.741'] | 3.8514811436640852 | true | - |

Figura 12. Arquivo de saída do *script*.

Fonte: próprio autor.

Considerando o desafio de escolher as melhores poses de *docking* e visando uma contribuição à comunidade acadêmica e científica, disponibilizamos via GitHub o *script* utilizado nesta etapa do projeto. Para acessá-lo, basta entrar no link <https://github.com/vivianmp95/ligand-dock-rank.git> e seguir as instruções contidas no arquivo “README”.

3.5.1 Análise visual das poses selecionadas

A partir da metodologia proposta e da saída gerada com os melhores modelos de *docking* entre o receptor e cada ligante, uma análise visual foi realizada utilizando o *software* PyMOL para verificar a distância do ligante em relação ao sítio ativo e confirmar se eram bons modelos. Para este fim, a partir da literatura disponível, foram obtidas informações acerca dos resíduos de aminoácidos dos ligantes que são conhecidamente inibidores. Foi utilizada a

função de medição de distância “*measurement distances*”, presente na ferramenta, para verificar a distância entre os átomos dos resíduos inibidores e do sítio ativo do receptor.

3.5.2 Controle negativo do script

Como controle negativo desta etapa para verificar a eficácia do *script*, foi realizada uma validação utilizando peptídeos da base de dados do Propedia como *decoys*. Para isso, todas as sequências de peptídeos presentes na ferramenta foram baixadas (n=13.618), sendo selecionadas aquelas cujo tamanho médio fosse semelhante aos 11 peptídeos da etapa anterior (entre 35 e 40 resíduos de aminoácidos). Além disso, as sequências também foram filtradas de forma a não possuir a função de inibidor de protease descrita.

Destas sequências, dez foram selecionadas e submetidas ao *docking* com a TMPRSS2 utilizando a ferramenta HDOCK, e então foram realizadas as mesmas análises anteriormente descritas utilizando o *script* desenvolvido.

3.6 Análise da estabilidade dos modos de ligação

A última etapa *in silico* do projeto envolveu a realização de Análise de Modos Normais (NMA) com o objetivo de analisar a estabilidade dos modos de ligação preditos. Apesar de grande parte das pesquisas visando a busca por inibidores da TMPRSS2 utilizar o *software* Desmond (Schrödinger Release 2023-3: Desmond Molecular Dynamics System) para realização da dinâmica molecular (ALZAIN; ELBADWI; ALSAMANI, 2022; KISHK et al., 2020; MANANDHAR et al., 2022), optamos pelo uso do iMODS, utilizado anteriormente.

É interessante destacar que o iMODS também é amplamente utilizado na imunoinformática (SHAWAN et al., 2023) para predição do movimento de um complexo proteico para estudo vacinal, de forma mais rápida e econômica que as simulações mais complexas.

3.7 Otimização e síntese dos peptídeos

A otimização das sequências de peptídeos visa aprimorar a afinidade e a interação dessas sequências com o centro ativo da protease, a fim de maximizar sua eficácia como potenciais agentes terapêuticos. Essa etapa é fundamental para a questão da seletividade do peptídeo, uma vez que os peptídeos escolhidos possuem histórico de inibição de outras serino proteases, também presentes no corpo humano, que não devem ser inibidas. Utilizando a abordagem de otimização das sequências, aumentará a afinidade com o sítio ativo da TMPRSS2, aumentando sua seletividade de inibição.

Este processo de otimização envolveu a utilização de um algoritmo de *design* de peptídeos criado pelo Laboratório de Bioinformática e Sistemas, denominado Potter (ainda não publicado). Basicamente, o algoritmo recebe como entrada uma sequência de peptídeo, uma estrutura do receptor e a especificação do sítio ativo. A cada iteração, mutações aleatórias são modeladas com base na estrutura do peptídeo. As melhores mutações são selecionadas com base em cálculo de contatos entre receptor e ligante (ocupância do sítio). Então, o algoritmo repete o processo por "n" iterações. Essa estratégia é baseada em algoritmos genéticos. Além disso, ele realiza ensaios de *docking* molecular a cada sequência nova criada para avaliar a afinidade entre os objetos de estudo, isso é feito utilizando o *script* do programa PyRosetta (SIDHARTHA et al., 2010).

Para o projeto, foi especificado o sítio ativo H296, D345 e S441, o número de *dockings* realizados por peptídeo foi 10, o número de peptídeos a serem gerados foi 25 e o número de iterações foi 20. Esses parâmetros foram escolhidos de forma a gerar um número considerável de resultados, porém, sem aumentar consideravelmente o tempo de execução para cada peptídeo, uma vez que o programa ainda encontra-se em fase de otimização, buscando a redução do tempo de geração de resultados.

Após a escolha das melhores sequências otimizadas, foi realizada a síntese dos peptídeos pelo Laboratório de Malária e Genômica de Parasitos da UFMG, que sintetizou dois dos três peptídeos escolhidos. O número de peptídeos foi reduzido devido ao tamanho de um deles ser maior do que poderia (n=46) para a síntese. Os dois sintetizados estão em processo de purificação, em parceria com o Laboratório de Biofísica de Macromoléculas da UFMG, para posterior ensaio *in vitro*.

4. Resultados e discussão

4.1 Levantamento bibliográfico e análise da estrutura da TMPRSS2

A partir do levantamento bibliográfico, obtivemos informações de que o sítio ativo da TMPRSS2 possui um ectodomínio de três cadeias, sendo duas não catalíticas (LDLR e SRCR) e a cadeia catalítica SP (Serino-Protease), com uma tríade catalítica canônica S441-H296-D345 (DAVID et al., 2020).

Como mencionado anteriormente na metodologia, foi escolhida a estrutura de TMPRSS2 do *Protein Data Bank*, cujo PDB ID é 7MEQ. Segundo Fraser et al. (2022), a estrutura cristalizada apresentou densidade eletrônica boa para os resíduos 149-491 (contendo SRCR e SP) mas não para os resíduos 109-148, contendo o domínio flexível LDLR-A, responsável por ligar a protease à membrana plasmática. A existência destes resíduos ausentes, no entanto, tem pouca influência sobre a presente pesquisa, uma vez que não afetou a região próxima dos resíduos do sítio catalítico.

Com relação à análise da flexibilidade do sítio ativo da TMPRSS2, é importante ressaltar que, como a estrutura começa no resíduo 148, os números correspondentes aos próximos resíduos de aminoácidos também mudarão. Sendo assim, foi realizada uma conversão em que 148 representa o aminoácido de número 1 nos gráficos a seguir. Dessa forma, o sítio catalítico S441-H296-D345 está representado nos gráficos como S294-H149-D198.

A Figura 16 mostra os resultados da Análise de Modos Normais para avaliar a flexibilidade. Ela foi gerada através da ferramenta RStudio utilizando os resultados gerados pelo iMODS e representa as frequências de vibração dos resíduos de aminoácidos da TMPRSS2, em Angstroms quadrados (\AA^2). Quanto maior o valor, maior a flexibilidade. Verificamos que os resíduos de interesse estão em regiões de baixa flexibilidade, abaixo de 0.4. Com esse resultado, decidimos que não seria necessário adicionar flexibilidade ao sítio ativo na análise de *docking* molecular.

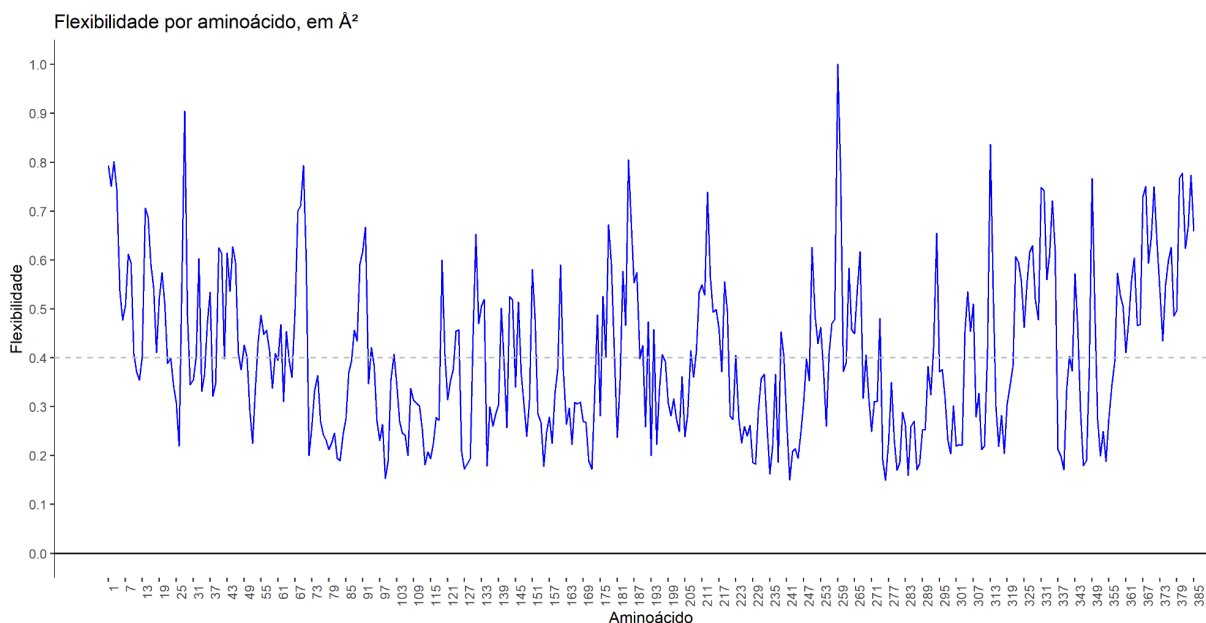


Figura 16. Gráfico de análise de flexibilidade da TMPRSS2. Gráfico gerado pelo RStudio utilizando dados gerados pelo iMODS contendo os resultados da Análise de Modos Normais. No eixo X, estão representados os aminoácidos da TMPRSS2; no eixo Y, estão representadas as frequências de vibração e deslocamento dos resíduos, através da linha azul.

Fonte: próprio autor.

4.2 Obtenção dos peptídeos candidatos

Foram selecionados os resíduos de 7MEQ que estivessem a uma distância de até 6Å do seu sítio ativo (Figura 17). Os resíduos foram: 281, 282, 293, 294, 295, 296, 297, 298, 299, 337, 342, 343, 344, 345, 346, 347, 437, 439, 440, 441, 442, 443, 458, 459, 460, 475 e 480.

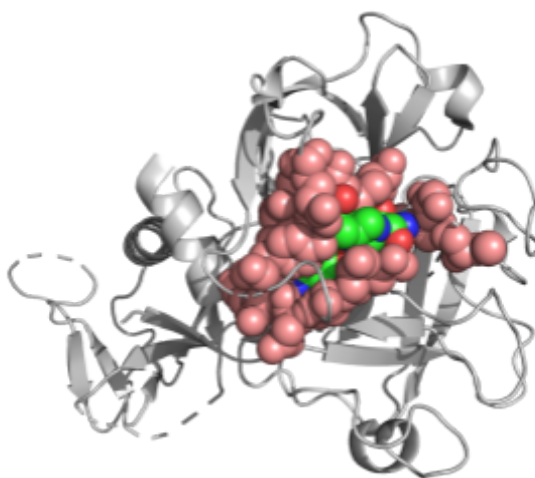


Figura 17. Estrutura da TMPRSS2 gerada pela ferramenta Propedia. O sítio catalítico está colorido por elemento (carbono em verde, oxigênio em vermelho, nitrogênio em azul), e os resíduos de até 6Å de distância do sítio coloridos em rosa.

Fonte: Propedia (<http://bioinfo.dcc.ufmg.br/propedia>).

A seguir, esses resíduos foram submetidos à ferramenta Propedia, que retornou como resultado 127 complexos proteína-peptídeo (Figura 18) cujo receptor possui sítio de ligação com resíduos similares, em posições próximas às selecionadas. Dessa forma, os peptídeos ligados a esses receptores possuem a ligação com o mesmo padrão que um peptídeo pode ter com a TMPRSS2, tornando-os candidatos ao estudo. Um ponto crucial a ser ressaltado é que nós sabemos que esses peptídeos apareceram como resultado como sendo candidatos a inibidores e não por serem clivados pela protease, isso porque os ligantes selecionados através do Propedia já se encontram em arquivos cristalográficos, portanto, não foram clivados pela TMPRSS2.

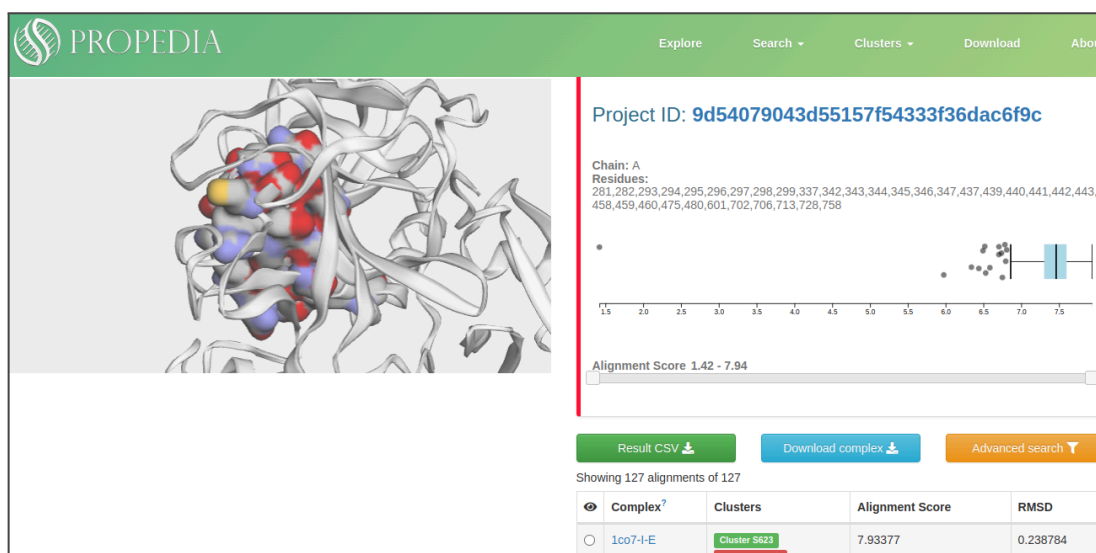


Figura 18. Resultados gerados pelo Propedia.

Fonte: Propedia (<http://bioinfo.dcc.ufmg.br/propedia>).

4.2.2 Filtragem das sequências

A partir da primeira filtragem das sequências resultantes da etapa anterior (n=127), em que foram selecionadas aquelas que tivessem pelo menos oito resíduos de aminoácidos, o número de sequências reduziu para 37. Foi realizada então a segunda filtragem, a partir do alinhamento par-a-par com a aprotinina, que reduziu as sequências de 37 para 11 (Tabela 3). Aquelas que passaram para a próxima etapa podem ser vistas na Tabela 4. É importante ressaltar que, na tabela, foi inserido o *Alignment Score* de cada resultado. Essa pontuação é atribuída pelo algoritmo ProBiS, utilizado na Propedia, a cada alinhamento de pares de sítios de ligação de proteínas. Quanto maior a pontuação de alinhamento, mais semelhantes são os sítios de ligação, em uma pontuação que vai até 10.

| | | | |
|------|-----------------------------|--|---------|
| 5NHU | Prothrombin | <i>Anopheles gambiae</i> anticoagulant | 7.69213 |
| 4AOR | Cationic trypsin | <i>Spinacia oleracea</i> trypsin inhibitor | 7.55852 |
| 1GL1 | Alpha-Chymotrypsin | <i>Locusta migratoria</i> protease inhibitor | 6.79497 |
| 1AN1 | Trypsin | <i>Hirudo medicinalis</i> tryptase inhibitor | 7.44543 |
| 1FLE | Porcine pancreatic elastase | <i>Homo sapiens</i> elafin | 7.48176 |
| 1LDT | Porcine trypsin | <i>Hirudo medicinalis</i> tryptase inhibitor | 7.36271 |
| 2Z7F | Leukocyte elastase | <i>Homo sapiens</i> Antileukoproteinase | 7.46709 |
| 3FP7 | Anionic trypsin-2 | <i>Bos taurus</i> pancreatic trypsin inhibitor | 6.95283 |
| 4DOQ | Porcine pancreatic trypsin | <i>Homo sapiens</i> Antileukoproteinase | 7.25544 |

4.3 Docking molecular

O *docking* molecular, tanto do HDOCK quanto do HPEPDOCK, gerou 100 conformações para cada uma das 11 sequências de peptídeos, obtendo um total de 1.100 arquivos. Como resultado, as ferramentas também geram uma pontuação (quanto menor o valor, melhor o resultado) e um RMSD (*Root Mean Square Deviation*, medida da distância média entre os átomos das duas moléculas sobrepostas) para cada modelo de todos os gerados. O *docking score* é calculado através da função de pontuação iterativa *knowledge-based*. Uma pontuação mais negativa significa um modelo de ligação mais provável, mas ela não deve ser tratada como a verdadeira afinidade de ligação de duas moléculas. Na Tabela 5, é possível visualizar as informações de *docking score* do melhor modelo, segundo as ferramentas, para cada um dos 11 peptídeos utilizados.

Tabela 5. Peptídeos utilizados na etapa de docking molecular. Na primeira coluna, são indicadas as identificações dos peptídeos no PDB; na segunda coluna, são indicadas as pontuações de docking geradas pelas ferramentas.

| PDB ID | Docking Score | Sequência |
|--------|---------------|--|
| 4ZKN | -164.276 | CSARGLENHRMC |
| 4E05 | -165.665 | EDYAAIEASLSETFNNTAADPGRRLGEGSKP |
| 5NHU | -169.840 | EEFDPSLLEEHADAPTARDPGRNPEFLRN |
| 4AOR | -251.130 | CSPSGAICSGFGPPEQCCSGACVPHILRIFVCQ |
| 1GL1 | -191.360 | ISCEPGKTFKDKCNTCRCGADGKSAACTLKACP |
| 1AN1 | -167.690 | KVCACPKILKPVCGSDGRTYANSCARCNGVSIKSEGSCP |
| 1FLE | -233.540 | TKPGSCPIILIRCAMLNPPNRCLKDTDCPGIKKCEGSCGMACFVPQ |
| 1LDT | -175.420 | KKVCACPKILKPVCGSDGRTYANSCARCNGVSIKSEGSCPTGILN |

| | | |
|------|----------|---|
| 2Z7F | -252.290 | RRKPGKCPVTYQGCLMLNPPNFCEMDGQCKRDLKCCMGKSCVSP VKA |
| 3FP7 | -183.900 | ARIIRYFYNAKAGLCQTFVYGGCRAKRNNFKSAEDCMRTCGGA |
| 4DOQ | -252.290 | KPGKCPVTYQGCLMLNPPNFCEMDGQCKRDLKCCMGKSCVSPV K |

4.3.1 Re-docking HPEPDOCK/HDOCK

O *re-docking* mostrou que ambas as ferramentas demonstraram um bom desempenho no encaixe do peptídeo à proteína. Um bom alinhamento resulta em uma pontuação de $RMSD \leq 2.0\text{\AA}$, no entanto, no caso de peptídeos menores, por ser um docking de difícil execução, os valores pouco acima de 2.0\AA obtidos com o HPEPDOCK foram aceitos (ZHOU et al., 2018). Dessa forma, pudemos proceder com o uso dessas ferramentas para a realização do *docking* molecular entre os peptídeos escolhidos e o receptor em estudo. A Tabela 6 mostra as estruturas utilizadas como *decoys* e seus respectivos RMSDs gerados através da comparação do complexo gerado e do complexo presente no PDB.

Tabela 6. Resultado do *re-docking*. As cinco primeiras estruturas (em azul) foram montadas utilizando o HPEPDOCK, e as cinco últimas utilizando o HDOCK (em amarelo). A primeira coluna mostra o ID da molécula no banco de dados, e a segunda coluna indica o RMSD entre a estrutura original e a produzida.

| PDB ID | RMSD do <i>Re-docking</i> | Tamanho do peptídeo |
|--------|---------------------------|---------------------|
| 1J2X | 0.360 | 18 |
| 2IHS | 2.754 | 20 |
| 2NM1 | 0.381 | 17 |
| 2PEH | 2.163 | 10 |
| 2PV2 | 2.842 | 12 |
| 2QOS | 0.327 | 11 |
| 4DJ9 | 0.519 | 29 |
| 4ETO | 0.683 | 16 |
| 4HTP | 0.373 | 11 |
| 5F67 | 0.646 | 19 |

4.4 Filtragem das melhores poses

Dos onze ligantes analisados, oito foram escolhidos por estarem mais próximos ao sítio ativo e podem ser vistos na Tabela 7 e na Figura 19. É importante ressaltar que dois pares dos peptídeos escolhidos são os mesmos ligantes ancorados em moléculas diferentes, ou seja, estão em PDB ID diferentes: 2Z7F e 4DOQ são ambos antileucoproteínas isoladas de *Homo*

sapiens, enquanto 1AN1 e 1LDT são inibidores de triptase isolados de *Hirudo medicinalis*. Dessa forma, para evitar repetição, esses peptídeos são indicados na tabela em amarelo, e os escolhidos dentre eles são indicados em vermelho.

Tabela 7. Peptídeos selecionados utilizando a metodologia do projeto. A primeira coluna indica o ID do peptídeo no Protein Data Bank; a segunda coluna indica o nome do peptídeo; a terceira coluna mostra o melhor modelo de docking, selecionado utilizando o script da etapa 3.5 da metodologia; e a última coluna indica o score de cada peptídeo. As linhas coloridas em amarelo claro indicam um mesmo peptídeo em estruturas diferentes, assim como as linhas em cinza.

| PDB ID | Molécula | Melhor modelo | Score |
|--------|--|---------------|---------|
| 1FLE | Elafin, elastase-specific inhibitor (<i>Homo sapiens</i>) | 41 | -117.67 |
| 1GL1 | Protease Inhibitor LCMI II (<i>Locusta migratoria</i>) | 34 | -99.95 |
| 4ZKN | Upain-1-W3A (murinised human uPA Hydrolase inhibitor) | 19 | -163.96 |
| 3FP7 | Pancreatic trypsin inhibitor | 53 | -119.69 |
| 1AN1 | Leech-derived Tryptase Inhibitor (<i>Hirudo medicinalis</i>) | 7 | -128.41 |
| 1LDT | Leech-derived Tryptase Inhibitor (<i>Hirudo medicinalis</i>) | 12 | -135.51 |
| 2Z7F | Antileukoproteinase (<i>Homo sapiens</i>) | 17 | -134.14 |
| 4DOQ | Antileukoproteinase (<i>Homo sapiens</i>) | 17 | -134.14 |

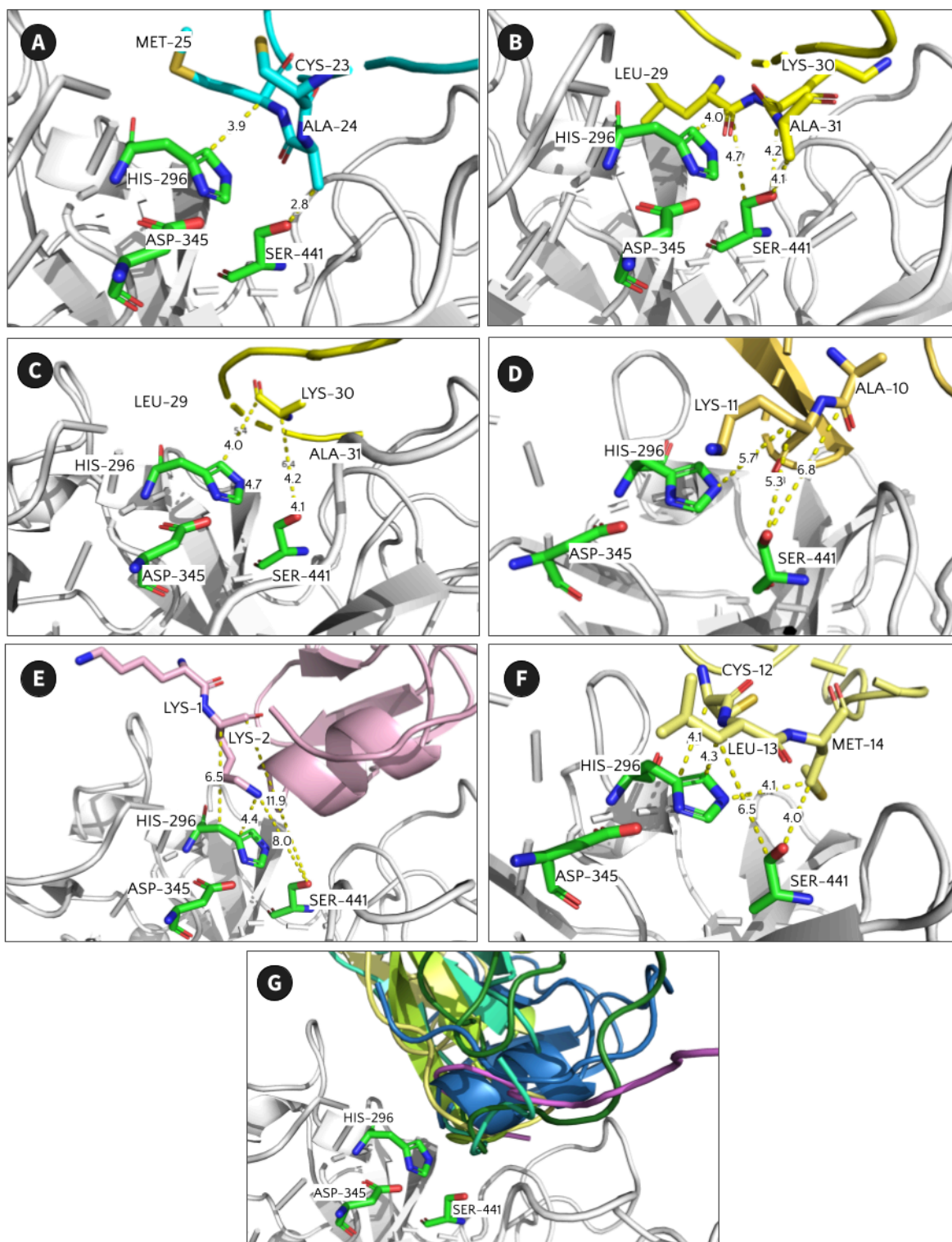


Figura 19. Melhores modelos de *docking* dos peptídeos ancorados ao sítio catalítico de TMPRSS2. (A) PDB ID: 1FLE, peptídeo: cadeia B (em azul); (B) ID do PDB: 1GL1, peptídeo: cadeias D, E, F (em amarelo); (C) PDB ID: 4ZKN, peptídeo: cadeia B (em amarelo); (D) ID do PDB: 3FP7, peptídeo: cadeia B (em amarelo); (E) PDB ID: 1LDT, peptídeo: cadeia B (em rosa); (F) PDB ID: 4DOQ, peptídeo: cadeias B e D (em amarelo); e (G) todos os peptídeos sobrepostos e representados como cartoons. Figuras geradas utilizando o software PyMol.

Fonte: Próprio autor.

Partindo da filtragem de possíveis ligantes, uma etapa crucial desse estudo foi a revisão bibliográfica para adquirir informações sobre os ligantes candidatos à inibição da protease e os resíduos responsáveis pela inibição para cada um deles. Essa revisão nos revelou as seguintes informações, para cada um dos seis ligantes selecionados na etapa 3.5 da metodologia:

- I. 1FLE - Elafina (Figura 19A): Entre os candidatos, destacamos a Elafina, um peptídeo de 57 aminoácidos, pertencente à família das quelonianinas. A Elafina é conhecida por controlar a atividade enzimática de Serino-proteases de neutrófilos, incluindo Elastase, Proteinase 3 e Catepsina G (ZANI et al., 2009). Seu mecanismo de inibição envolve uma alça (C15-C38) e o domínio WAP (*Whey Acidic Protein*), responsáveis pela interação com a protease. O potencial de inibição da Elafina aponta para sua possível eficácia como um inibidor de TMPRSS2.
- II. 1GL1 - Peptídeo de Gafanhoto (Figura 19B): O peptídeo 1GL1 é composto por 36 aminoácidos e foi isolado de gafanhoto (*Locusta migratoria*). Roussel et al. (2001) revelaram que esse peptídeo é um potente inibidor de Quimotripsina bovina, com alguma atividade fraca contra Elastase de leucócitos humanos. A possível origem de sua propriedade inibitória foi associada ao resíduo L30.
- III. 4ZKN - Upain-1-W3A (Figura 19C): O peptídeo 4ZKN, também conhecido como Upain-1-W3A, é um peptídeo cíclico composto por apenas 12 aminoácidos. Foi derivado de um inibidor de uPA peptidil, o upain-1, com uma mutação no W3 para Ala (JIANG et al., 2016). Embora seu mecanismo de inibição não esteja completamente esclarecido, há indicações de que a atividade inibitória possa estar relacionada ao resíduo mutado W3.
- IV. 3FP7 - Variante de Tripsina aniônica S195A (Figura 19D): A variante 3FP7 refere-se a uma Tripsina aniônica S195A de rato (*Rattus norvegicus*) em complexo com o Inibidor de Tripsina pancreática bovina (BPTI). Neste caso, a variante possui o resíduo catalítico Ser195 substituído por Ala (ZAKHAROVA; HORVATH; GOLDENBERG, 2009).
- V. 1LDT - Inibidor de Triptase de Sanguessuga (Figura 19E): O 1LDT é o inibidor de Triptase isolado de sanguessuga (*Hirudo medicinalis*). Composto por 46 aminoácidos, este peptídeo exibe um papel na inibição da Triptase pulmonar humana, onde se acredita que a sequência amino-terminal Lys-Lys desempenhe um papel significativo

(STUBBS et al., 1997). A interação desse inibidor com a TMPRSS2 pode oferecer uma abordagem potencialmente eficaz na inibição da protease.

- VI. 4DOQ - *Secretory Leukocyte Protease Inhibitor* (Figura 19F): Finalmente, o 4DOQ corresponde ao SLPI (*Secretory Leukocyte Protease Inhibitor*), uma proteína de 107 resíduos isolada de *Homo sapiens*. O SLPI inibe várias Serino-proteases, incluindo Elastase, Catepsina G, Quimotripsina e Tripsina, e demonstrou atividades antimicrobianas e anti-HIV (FUKUSHIMA; KAMIMURA; TAKIMOTO-KAMIMURA, 2013). Segundo Shine et al. (2002), ela também mostrou atividades antimicrobianas e anti-HIV. Em estudos, mostrou uma forte afinidade com o sítio S1 pelo resíduo P1 L72.

A Tabela 8, abaixo, contém as informações relevantes resumidas.

Tabela 8. Tabela contendo as informações sobre os peptídeos selecionados no projeto. Na primeira coluna, são indicadas as identificações dos peptídeos no PDB; na segunda coluna, são indicados o que cada peptídeo é; na terceira coluna, é indicada a sua atividade inibidora; na última coluna, são indicados os resíduos de cada peptídeo que participam da inibição.

| PDB ID | O que é | Atividade inibidora | Resíduos inibidores |
|--------|--|--|---------------------|
| 1FLE | Elafin, elastase-specific inhibitor (<i>Homo sapiens</i>) | Serino-proteases de neutrófilos como Elastase, Proteinase 3 e Catepsina G | Alça Cys15-Cys38 |
| 1GL1 | Protease Inhibitor LCMI II (<i>Locusta migratoria</i>) | Potente para Quimotripsina bovina, fraco para Elastase de leucócitos humanos | Leu30 |
| 4ZKN | Upain-1-W3A (murinised human uPA Hydrolase inhibitor) | Enzima proteolítica, converte plasminogênio em plasmina | Trp3 |
| 3FP7 | Pancreatic trypsin inhibitor (BPTI) | Tripsina | Ser195 |
| 1LDT | Leech-derived Tryptase Inhibitor (<i>Hirudo medicinalis</i>) | Triptase | Lys1-Lys2 |
| 4DOQ | Antileukoproteinase (<i>Homo sapiens</i>) | Elastase, Catepsina G, Quimotripsina e Tripsina | Leu72 |

Cada um desses candidatos oferece uma perspectiva única sobre como a TMPRSS2 pode ser inibida, fornecendo uma ampla gama de possibilidades para futuros estudos de desenvolvimento de inibidores da proteína.

4.4.1 Controle negativo do script

Todos os *decoys* apresentaram aminoácidos com, no mínimo, 6Å de distância do sítio ativo, como mostra a Tabela 9. Dessa forma, houveram indícios de que os peptídeos previamente selecionados com o suporte da máquina de busca da Propedia e juntamente com critérios desenvolvidos neste trabalho apresentam maior afinidade e melhor posicionamento no sítio e, portanto, podem ser inibidores mais promissores que os *decoys*, como esperado.

Tabela 9. Tabela indicando os dez *decoys* utilizados como controle negativo para o *script* desenvolvido no projeto. Na primeira coluna está indicada a identificação da molécula no banco de dados PDB. Na segunda coluna estão os nomes dos ligantes de cada molécula. Na terceira coluna, há as menores distâncias encontradas do átomo mais próximo do sítio catalítico da TMPRSS2 com cada ligante.

| PDB ID | Nº de resíduos | Molécula | Menor distância do sítio (Å) |
|--------|----------------|---|------------------------------|
| 1FQJ_C | 38 | Retinal rod rhodopsin-sensitive cGMP 3',5'-cyclic phosphodiesterase subunit gamma | 41.29 |
| 1SSE_A | 35 | AP-1 like transcription factor YAP1 | 17.34 |
| 2DT7_A | 38 | Splicing factor 3A subunit 3 | 22.75 |
| 2K42_B | 36 | ESPFU, an EHEC effector | 27.74 |
| 2OM2_B | 36 | Regulator of G-protein signalling 14 GoLoco motif peptide | 42.12 |
| 2QTV_D | 36 | Protein transport protein SEC31 | 34.71 |
| 1H8E_I | 36 | Bovine mitochondrial F1-ATPASE | 42.97 |
| 2HCJ_A | 35 | Protein chain elongation factor EF-Tu | 38.68 |
| 4N9G_D | 35 | Epitope Scaffold rsv_1isea_FFL_001_C | 31.87 |
| 5W5C_A | 38 | Vesicle-associated membrane protein 2 | 36.42 |

4.5 Análise da estabilidade dos modos de ligação

A análise de estabilidade dos modos de ligação nos permitiu avaliar a flexibilidade e a estabilidade dos resíduos envolvidos nas ligações moleculares, permitindo assim escolher aqueles peptídeos que fossem mais estáveis, de acordo com a região dada como inibidora. A seguir, apresentamos os resultados dessa análise.

4.5.1 Peptídeo 1FLE

O peptídeo corresponde à região dos aminoácidos 1-46, sendo 15-38 os resíduos críticos para inibição, como informado na Tabela 8. Assim, é possível visualizar na Figura 20 que há uma

tendência geral de flexibilidade alta no peptídeo, inclusive na região de inibição. Dessa forma, o peptídeo de identificação 1FLE não se mostrou um bom candidato.

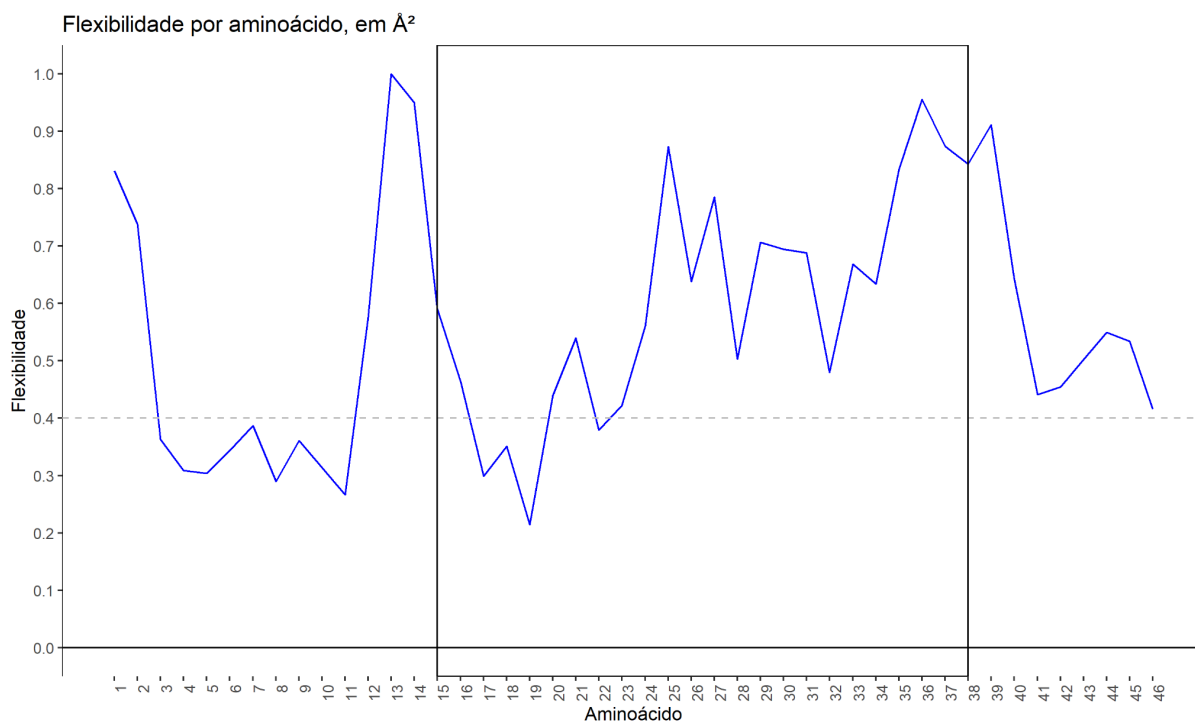


Figura 20. Gráfico de análise de flexibilidade para 1FLE. Gráfico gerado pelo RStudio utilizando dados gerados pelo iMODS contendo os resultados da Análise de Modos Normais. No eixo X, estão representados os aminoácidos do peptídeo (1-46); no eixo Y, estão representadas as frequências de vibração e deslocamento dos resíduos, através da linha azul. A linha pontilhada indica o limite de flexibilidade aceita (0.4 \AA^2) e o retângulo preto representa os aminoácidos responsáveis pela inibição.

Fonte: próprio autor.

4.5.2 Peptídeo 1GL1

O peptídeo corresponde à região dos aminoácidos 1-34, sendo o resíduo 30 crítico para inibição, como informado na Tabela 8. Assim, é possível visualizar na Figura 21 o iMODS mostra uma baixa-média flexibilidade na região analisada, incluindo uma baixa flexibilidade na região inibidora. Dessa forma, o peptídeo de identificação 1GL1 mostrou-se um bom candidato e passou para a próxima etapa da metodologia.

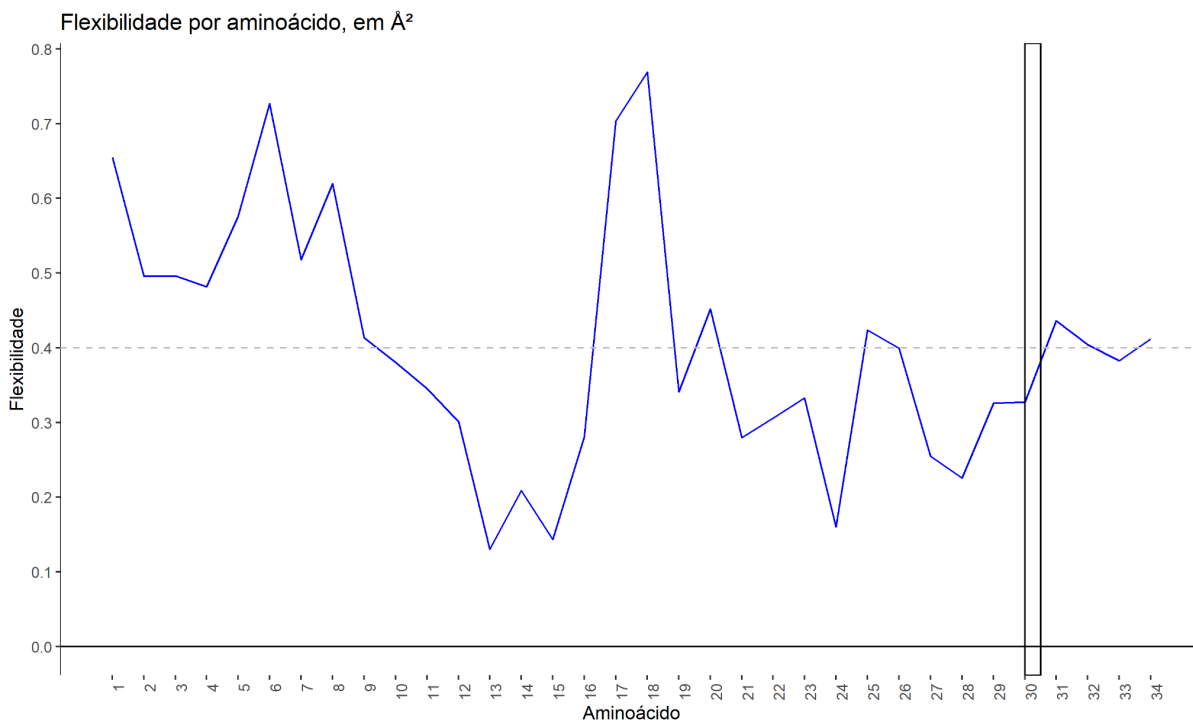


Figura 21. Gráfico de análise de flexibilidade para 1GL1. Gráfico gerado pelo RStudio utilizando dados gerados pelo iMODS contendo os resultados da Análise de Modos Normais. No eixo X, estão representados os aminoácidos do peptídeo (1-34); no eixo Y, estão representadas as frequências de vibração e deslocamento dos resíduos, através da linha azul. A linha pontilhada indica o limite de flexibilidade aceita (0.4 \AA^2) e o retângulo preto representa o aminoácido responsável pela inibição.

Fonte: próprio autor.

4.5.3 Peptídeo 3FP7

O peptídeo corresponde à região dos aminoácidos 1-43, sendo os resíduos 10-11 críticos para inibição, como informado na Tabela 8. Assim, é possível visualizar na Figura 22 que, apesar de apresentar algumas regiões de baixa flexibilidade, o peptídeo oscila muito entre média e alta flexibilidade, inclusive nos resíduos de inibição. Dessa forma, o peptídeo de identificação 3FP7 não se mostrou um bom candidato.

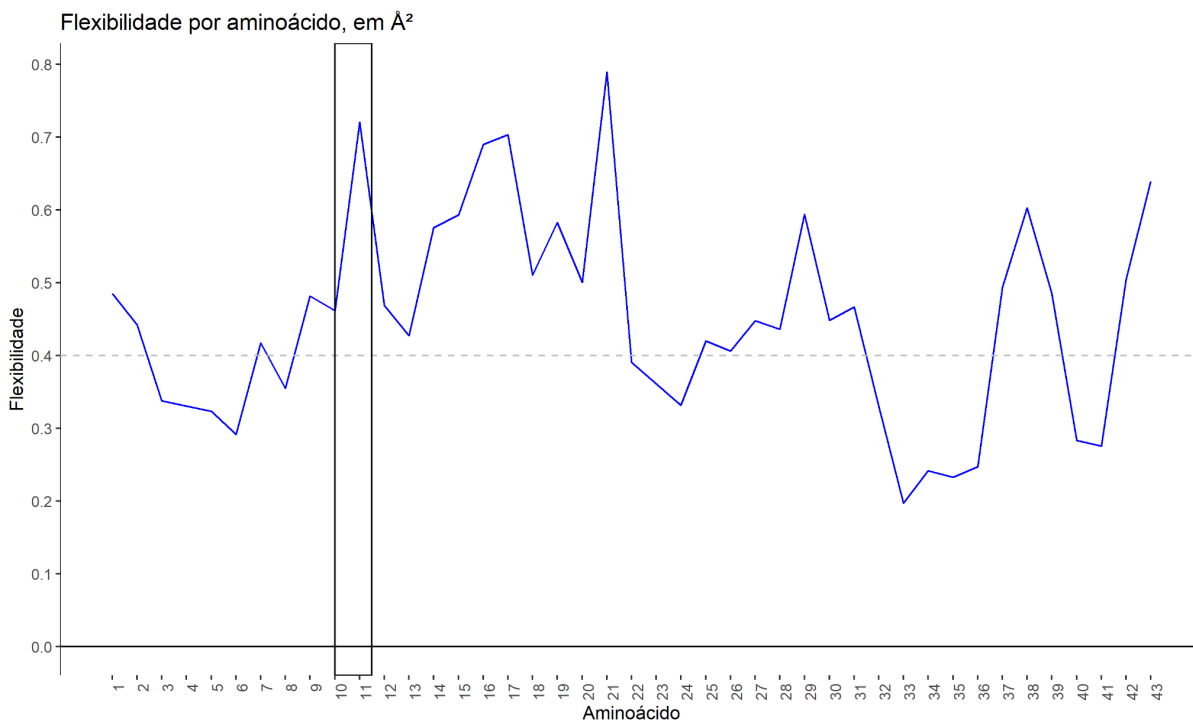


Figura 22. Gráfico de análise de flexibilidade para 3FP7. Gráfico gerado pelo RStudio utilizando dados gerados pelo iMODS contendo os resultados da Análise de Modos Normais. No eixo X, estão representados os aminoácidos do peptídeo (1-43); no eixo Y, estão representadas as frequências de vibração e deslocamento dos resíduos, através da linha azul. A linha pontilhada indica o limite de flexibilidade aceita (0.4 Å²) e o retângulo preto representa os aminoácidos responsáveis pela inibição.

Fonte: próprio autor.

4.5.4 Peptídeo 1LDT

O peptídeo corresponde à região dos aminoácidos 1-46, sendo os resíduos 1-2 críticos para inibição, como informado na Tabela 8. Assim, é possível visualizar na Figura 23 que, apesar de oscilar entre baixa-média-alta flexibilidade, há pouca flexibilidade na região inibidora. Dessa forma, o peptídeo de identificação 1LDT mostrou-se um bom candidato e passou para a próxima etapa da metodologia.

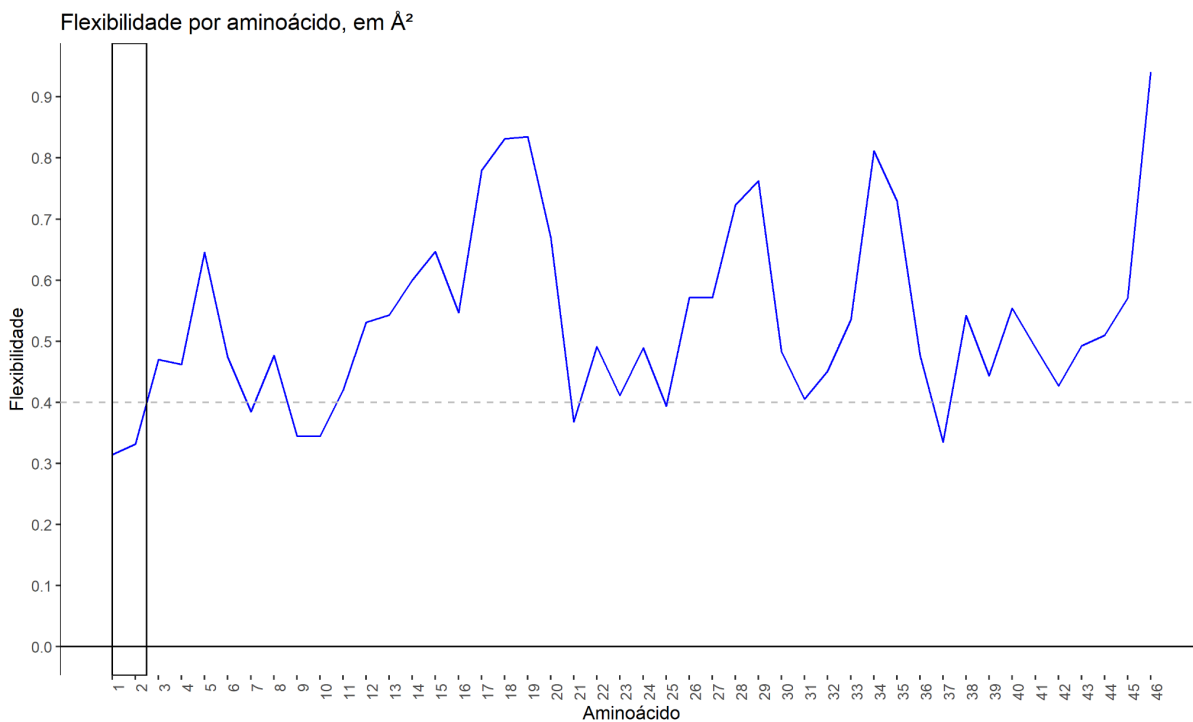


Figura 23. Gráfico de análise de flexibilidade para 1LDT. Gráfico gerado pelo RStudio utilizando dados gerados pelo iMODS contendo os resultados da Análise de Modos Normais. No eixo X, estão representados os aminoácidos do peptídeo (1-46); no eixo Y, estão representadas as frequências de vibração e deslocamento dos resíduos, através da linha azul. A linha pontilhada indica o limite de flexibilidade aceita (0.4 Å²) e o retângulo preto representa os aminoácidos responsáveis pela inibição.

Fonte: próprio autor.

4.5.5 Peptídeo 4DOQ

O peptídeo corresponde à região dos aminoácidos 1-47, sendo os resíduos 12-14 críticos para inibição, como informado na Tabela 8. Assim, é possível visualizar na Figura 24 que o peptídeo tende a seguir um padrão de média para elevada flexibilidade, inclusive na região inibidora. Dessa forma, o peptídeo de identificação 4DOQ não se mostrou um bom candidato.

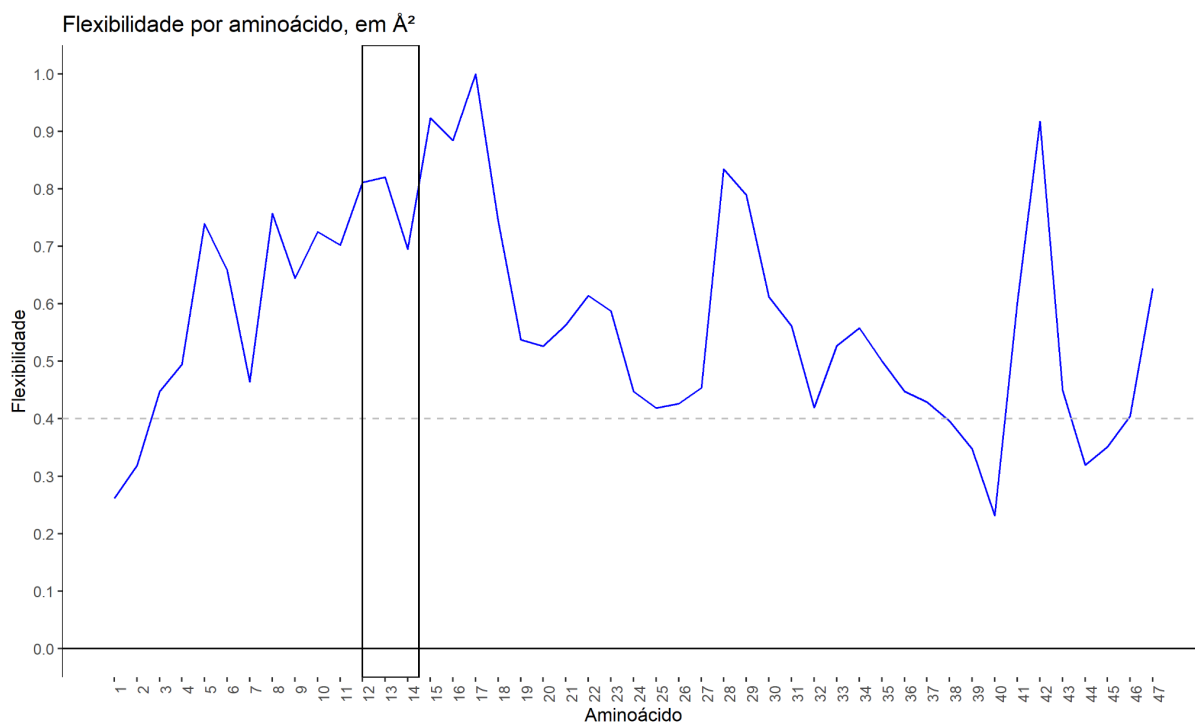


Figura 24. Gráfico de análise de flexibilidade para 4DOQ. Gráfico gerado pelo RStudio utilizando dados gerados pelo iMODS contendo os resultados da Análise de Modos Normais. No eixo X, estão representados os aminoácidos do peptídeo (1-47); no eixo Y, estão representadas as frequências de vibração e deslocamento dos resíduos, através da linha azul. A linha pontilhada indica o limite de flexibilidade aceita (0.4 Å²) e o retângulo preto representa os aminoácidos responsáveis pela inibição.

Fonte: próprio autor.

4.5.6 Peptídeo 4ZKN

O peptídeo corresponde à região dos aminoácidos 1-12, sendo o resíduo 3 crítico para inibição, como informado na Tabela 8. Assim, é possível visualizar na Figura 25 que, apesar de apresentar uma média-alta flexibilidade em algumas regiões, os resíduos 1-5 apresentam baixa flexibilidade, incluindo a região inibidora. Dessa forma, o peptídeo de identificação 4ZKN mostrou-se um bom candidato e passou para a próxima etapa da metodologia.

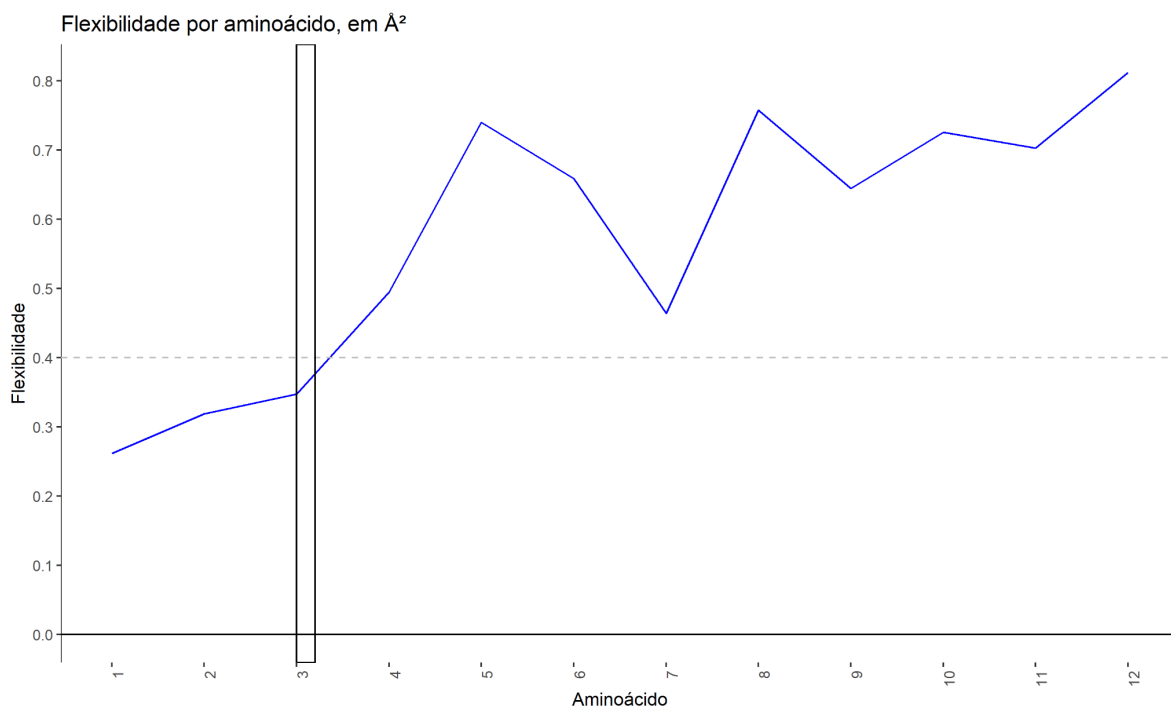


Figura 25. Gráfico de análise de flexibilidade para 4ZKN. Gráfico gerado pelo RStudio utilizando dados gerados pelo iMODS contendo os resultados da Análise de Modos Normais. No eixo X, estão representados os aminoácidos do peptídeo (1-47); no eixo Y, estão representadas as frequências de vibração e deslocamento dos resíduos, através da linha azul. A linha pontilhada indica o limite de flexibilidade aceita (0.4 Å²) e o retângulo preto representa o aminoácido responsável pela inibição.

Fonte: próprio autor.

Na tabela 10, é possível verificar os resíduos de interesse e valor de flexibilidade para os peptídeos analisados. Esses resultados destacam a importância de considerar a flexibilidade dos resíduos nas interações com os peptídeos inibidores. No presente estudo, não foi estabelecido um limite máximo de flexibilidade a ser considerada, desde que fosse pelo menos metade da flexibilidade máxima do ligante. Foram selecionados para a próxima etapa aqueles que apresentaram menor flexibilidade dos resíduos inibidores que se ligam ao receptor em comparação com os outros. Isso permite dar uma maior estabilidade ao inibidor. Foram estes: 1GL1, 4ZKN e 1LDT.

Tabela 10. Resultados da Análise de Modos Normais. Na primeira coluna, são indicadas as identificações dos peptídeos no PDB; na segunda coluna, são indicados os resíduos de interesse; na terceira coluna, são indicados os valores de flexibilidade.

| ID | Resíduos de interesse | Flexibilidade nos resíduos de interesse |
|------|-----------------------|---|
| 1FLE | 15-38 | 0.2 a 1.0 |
| 1GL1 | 30 | 0.2 |
| 3FP7 | 10-11 | 0.8 |
| 1LDT | 1-2 | 0.3 |
| 4DOQ | 12-14 | 0.8 |
| 4ZKN | 3 | 0.3 |

4.7 Otimização e síntese dos peptídeos

Neste tópico, apresentamos os resultados obtidos a partir da otimização das sequências, incluindo as sequências otimizadas e suas respectivas pontuações de *docking*. Na Tabela 11, são destacadas as sequências selecionadas e suas avaliações de afinidade e, na Figura 26, podem ser vistos os complexos tidos como melhores resultantes do *docking* molecular da ferramenta. Essas sequências representam candidatos promissores que merecem análises posteriores e ensaios laboratoriais para confirmar sua eficácia e potencial terapêutico.

Tabela 11. Sequências utilizadas para otimização. Na primeira coluna, são indicadas as identificações dos peptídeos no PDB; na segunda coluna, são indicadas as sequências originais dos peptídeos; na terceira coluna, são indicadas as sequências geradas pela otimização; na quarta coluna, são indicadas as pontuações do *docking* realizado pela ferramenta; na última coluna, é indicada em qual iteração a sequência escolhida foi feita.

| PDB ID | Sequência anterior | Nova sequência | Score | Iteração |
|--------|--|---|-------|----------|
| 4ZKN | CSARGLENHRMC | CTWRGLENRMC | -517 | 1 |
| 1GL1 | ISCEPGKTFKDKCNTCRGADGK SAACTLKACPN | ATAPRADKCNCHSNEEGHSAAA CGG | -470 | 1 |
| 1LDT | KKVCACPKILKPVCGSDGRTYAN SCIARCNGVSIKSEGSCPTGILN | RHACAVGRIVRPVCPADGKIYEN SVLSHVNPVTLKAAPTGANLVN | -337 | 2 |

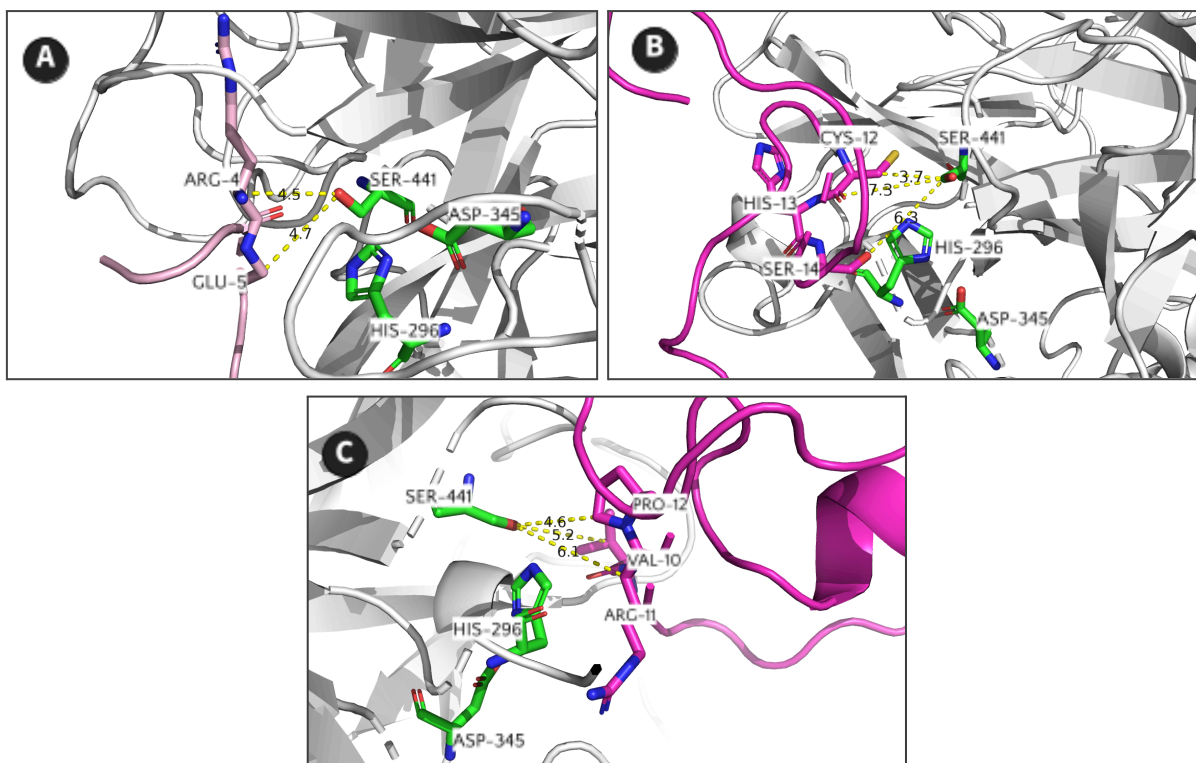


Figura 26. Melhores modelos de *docking* dos peptídeos otimizados. Em verde, encontra-se o sítio catalítico da TMPRSS2 ancorado aos peptídeos: (A) PDB ID: 4ZKN, peptídeo em rosa claro; (B) ID do PDB: 1LDT, peptídeo em magenta; e (C) PDB ID: 1GL1, peptídeo em magenta. Figuras geradas utilizando o software PyMol.

Fonte: Próprio autor.

Segundo Fraser et al. (2022), no geral, o centro ativo da TMPRSS2 parece ser capaz de se ligar a várias sequências de substrato com uma preferência definida nas posições P1 e P2. De acordo com o MEROPS, a posição P1 tem preferência de clivagem por Lisina e Arginina, enquanto a posição P2 tem preferência por Glicina, Prolina, Alanina, Serina, Glutamina, Lisina, Arginina e Histidina (Tabela 12). A partir dos resultados de *docking* molecular dos peptídeos otimizados com a TMPRSS2 (Figura 26), podemos visualizar que alguns desses peptídeos estão representados próximo ao sítio ativo. No caso do ligante correspondente a 4ZKN, os resíduos R4 e E5 encontram-se a menos de 5Å de distância da S441. No ligante 1GL1, são os resíduos C12, H13 e S14 que estão próximos do sítio. Já para 1LDT, os resíduos P12, V10 e R11 estão próximos. Assim, é possível afirmar que os ligantes otimizados podem ter mais chances de inibir a TMPRSS2, uma vez que se encaixam nas posições P1 e P2 da protease, considerando suas preferências de aminoácidos.

| aa | A | R | N | D | C | Q | E | G | H | I | L | K | M | F | P | S | T | W | Y | V |
|----|---|----|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| P1 | | 15 | | | | | | | | | | 6 | | | | | | | | |
| P2 | 2 | 5 | | | | | 1 | 2 | 1 | | | 4 | | | 1 | 5 | | | | |

Tabela 12. Matriz de especificidade do MEROPS para Tmprss2. Na primeira linha (aa) são representados os aminoácidos; na segunda linha são representados os resíduos de preferência da posição P1; na terceira linha são representados os resíduos de preferência da posição P2.

Fonte: Próprio autor.

A última etapa do projeto envolveu a síntese dos peptídeos escolhidos como melhores inibidores. Esses peptídeos encontram-se em processo de purificação, pelo Laboratório de Biofísica de Macromoléculas, para posterior ensaio de citotoxicidade. Dessa forma, será possível avaliar o potencial dos candidatos de causar danos às células, verificando assim a sua segurança e possíveis efeitos colaterais. Esse ensaio tem prioridade sob o ensaio de atividade enzimática pois não foi possível adquirir os materiais necessários para essa experimentação. À medida que prosseguirmos com experimentos laboratoriais adicionais e ensaios, estudaremos com maior clareza o potencial de inibição dessas moléculas.

5. Conclusões

Este trabalho desenvolveu uma metodologia abrangente para a identificação de peptídeos inibidores da TMPRSS2, com foco na interrupção de infecções virais que utilizam dessa protease para iniciar seu ciclo. Além disso, a ferramenta desenvolvida no trabalho com o intuito de filtrar as poses gerados pela ferramenta de *docking* molecular podem repriorizar aquelas que inicialmente não seriam consideradas para avaliação, uma vez que suas pontuações estavam ligeiramente abaixo das pontuações dos dez melhores modelos, mas que mostraram ter uma melhor ligação com o sítio ativo. Essa abordagem representa um passo importante na busca por novas estratégias terapêuticas contra doenças mediadas pela TMPRSS2. Ao final do estudo, obtivemos três candidatos à inibição da TMPRSS2. Esses peptídeos representam candidatos promissores que merecem uma investigação mais aprofundada. À medida que continuamos a explorar esses candidatos em ensaios laboratoriais, esperamos contribuir para a compreensão e o combate desta pandemia e de outras doenças associadas à TMPRSS2.

6. Limitações e perspectivas

Este estudo apresenta contribuições significativas para a investigação de peptídeos como potenciais inibidores da Serino Protease Transmembrana 2 (TMPRSS2). No entanto, é fundamental reconhecer as limitações inerentes ao presente trabalho, bem como delinear as perspectivas futuras para aprofundar a pesquisa.

Uma das limitações presentes neste estudo consiste na ausência de uma validação experimental completa. A análise realizada neste trabalho concentrou-se principalmente em abordagens *in silico*, sendo necessário enfatizar que os resultados obtidos por essas metodologias não constituem evidência definitiva de atividade inibitória. Portanto, é necessário conduzir testes de inibição viral em bancada para confirmar a capacidade dos peptídeos sintetizados de efetivamente inibir o vírus SARS-CoV-2. Outra limitação enfrentada foi a restrição orçamentária. Os ensaios *in vitro* de atividade enzimática demandam recursos financeiros substanciais. A aquisição de *kits* e reagentes específicos para esses experimentos é dispendiosa, o que impediu a realização completa desta fase de validação experimental. Por último, a dinâmica molecular, uma das técnicas computacionais empregadas neste estudo, foi conduzida em um formato simplificado. Esta simplificação resultou da falta de recursos computacionais de alto desempenho e da carência de mão de obra qualificada para a realização de simulações mais complexas. Portanto, é necessário reconhecer que as simulações de análise de estabilidade dos resíduos e ligações realizadas podem não capturar completamente a complexidade dos sistemas biológicos em estudo.

Com relação às perspectivas, pretendemos realizar testes de inibição viral em bancada para avaliar a eficácia dos peptídeos sintetizados contra o vírus SARS-CoV-2, em parceria com o Laboratório de Biologia Integrativa da UFMG. Além disso, pretende-se elaborar um artigo científico que apresentará de forma abrangente os achados e as implicações do estudo. Por último, para promover a colaboração e a replicabilidade da pesquisa, o *script* construído neste trabalho para simulações de *docking* molecular será disponibilizado publicamente. Isso permitirá que outros grupos de pesquisa utilizem essa ferramenta em seus estudos relacionados ao *docking* molecular e à identificação de candidatos terapêuticos.

7. Referências

1. ABE, M. et al. TMPRSS2 Is an Activating Protease for Respiratory Parainfluenza Viruses. **Journal of Virology**, v. 87, n. 21, p. 11930–11935, nov. 2013.
2. ADER, F. et al. Remdesivir plus standard of care versus standard of care alone for the treatment of patients admitted to hospital with COVID-19 (DisCoVeRy): a phase 3, randomised, controlled, open-label trial. **The Lancet. Infectious Diseases**, v. 22, n. 2, p. 209–221, fev. 2022.
3. AFAR, D. E. H. et al. Catalytic Cleavage of the Androgen-regulated TMPRSS2 Protease Results in Its Secretion by Prostate and Prostate Cancer Epithelia. **Cancer Research**, v. 61, n. 4, p. 1686–1692, 2 fev. 2001.
4. AGRAHARI, V.; AGRAHARI, V.; MITRA, A. K. Nanocarrier fabrication and macromolecule drug delivery: challenges and opportunities. **Therapeutic Delivery**, v. 7, n. 4, p. 257–278, abr. 2016.
5. ALZAIN, A. A.; ELBADWI, F. A.; ALSAMANI, F. O. Discovery of novel TMPRSS2 inhibitors for COVID-19 using in silico fragment-based drug design, molecular docking, molecular dynamics, and quantum mechanics studies. **Informatics in Medicine Unlocked**, v. 29, p. 100870, 2022.
6. AMANI, B.; AMANI, B. Efficacy and safety of nirmatrelvir/ritonavir (Paxlovid) for COVID-19: A rapid review and meta-analysis. **Journal of Medical Virology**, v. 95, n. 2, p. e28441, fev. 2023.
7. ANAND, U. et al. Translational aspect in peptide drug discovery and development: An emerging therapeutic candidate. **BioFactors (Oxford, England)**, v. 49, n. 2, p. 251–269, mar. 2023.
8. ARUFFO, A. et al. CD6-ligand interactions: a paradigm for SRCR domain function? **Immunology Today**, v. 18, n. 10, p. 498–504, out. 1997.
9. BAKAN, A.; MEIRELES, L. M.; BAHAR, I. ProDy: Protein Dynamics Inferred from Theory and Experiments. **Bioinformatics**, v. 27, n. 11, p. 1575–1577, 1 jun. 2011.
10. BAUER, J. A.; PAVLOVIĆ, J.; BAUEROVÁ-HLINKOVÁ, V. Normal Mode Analysis as a Routine Part of a Structural Investigation. **Molecules**, v. 24, n. 18, p. 3293, 10 set. 2019.
11. BEIGEL, J. H. et al. Remdesivir for the Treatment of Covid-19 - Final Report. **The New England Journal of Medicine**, v. 383, n. 19, p. 1813–1826, 5 nov. 2020.

12. BERMAN, H. M. et al. The Protein Data Bank. **Nucleic Acids Research**, v. 28, n. 1, p. 235–242, 1 jan. 2000.
13. BERTRAM, S. et al. Novel insights into proteolytic cleavage of influenza virus hemagglutinin. **Reviews in Medical Virology**, v. 20, n. 5, p. 298–310, 2010a.
14. BERTRAM, S. et al. TMPRSS2 and TMPRSS4 Facilitate Trypsin-Independent Spread of Influenza Virus in Caco-2 Cells. **Journal of Virology**, v. 84, n. 19, p. 10016–10025, out. 2010b.
15. BICKERTON, G. R.; HIGUERUELO, A. P.; BLUNDELL, T. L. Comprehensive, atomic-level characterization of structurally characterized protein-protein interactions: the PICCOLO database. **BMC Bioinformatics**, v. 12, p. 313, 29 jul. 2011.
16. BOJKOVA, D. et al. Aprotinin Inhibits SARS-CoV-2 Replication. **Cells**, v. 9, n. 11, p. 2377, 30 out. 2020.
17. BOURINET, E.; ZAMPONI, G. W. Block of voltage-gated calcium channels by peptide toxins. **Neuropharmacology**, v. 127, p. 109–115, dez. 2017.
18. BROOKE, G. N.; PRISCHI, F. Structural and functional modelling of SARS-CoV-2 entry in animal models. **Scientific Reports**, v. 10, n. 1, p. 15917, 28 set. 2020.
19. CELE, S. et al. Escape of SARS-CoV-2 501Y.V2 from neutralization by convalescent plasma. **Nature**, v. 593, n. 7857, p. 142–146, maio 2021.
20. CHEN, Y.-W. et al. TMPRSS2, a serine protease expressed in the prostate on the apical surface of luminal epithelial cells and released into semen in prostasomes, is misregulated in prostate cancer cells. **The American Journal of Pathology**, v. 176, n. 6, p. 2986–2996, jun. 2010.
21. CHENG, Z. et al. Identification of *TMPRSS2* as a Susceptibility Gene for Severe 2009 Pandemic A(H1N1) Influenza and A(H7N9) Influenza. **Journal of Infectious Diseases**, v. 212, n. 8, p. 1214–1221, 15 out. 2015.
22. CHERA, A.; TANCA, A. Remdesivir: the first FDA-approved anti-COVID-19 Treatment for Young Children. **Discoveries**, v. 10, n. 2, p. e151, [s.d.].
23. CIEMNY, M. et al. Protein–peptide docking: opportunities and challenges. **Drug Discovery Today**, v. 23, 1 maio 2018.
24. **Coronavírus: remédios que perderam eficácia contra variantes são suspensos.** Disponível em: <<https://www.uol.com.br/vivabem/noticias/redacao/2023/07/02/uso-de-remedios-que-perderam-eficacia-contra-variantes-da-covid-e-suspenso.htm>>. Acesso em: 21 set. 2023.

25. COUPEZ, B.; LEWIS, R. A. Docking and scoring--theoretically easy, practically impossible? **Current Medicinal Chemistry**, v. 13, n. 25, p. 2995–3003, 2006.
26. CRAIK, D. J.; KAN, M.-W. How can we improve peptide drug discovery? Learning from the past. **Expert Opinion on Drug Discovery**, v. 16, n. 12, p. 1399–1402, dez. 2021.
27. CUCINOTTA, D.; VANELLI, M. WHO Declares COVID-19 a Pandemic. **Acta Biomedica Atenei Parmensis**, v. 91, n. 1, p. 157–160, 19 mar. 2020.
28. DAVID, A. et al. **Structure, function and variants analysis of the androgen-regulated TMPRSS2, a drug target candidate for COVID-19 infection.** bioRxiv, , 26 maio 2020. Disponível em: <<https://www.biorxiv.org/content/10.1101/2020.05.26.116608v1>>. Acesso em: 3 ago. 2023
29. DAVID, A. et al. A common TMPRSS2 variant has a protective effect against severe COVID-19. **Current Research in Translational Medicine**, v. 70, n. 2, p. None, maio 2022.
30. DE VRIES, S. J. et al. The pepATTRACT web server for blind, large-scale peptide-protein docking. **Nucleic Acids Research**, v. 45, n. W1, p. W361–W364, 3 jul. 2017.
31. DIAO, L.; MEIBOHM, B. Pharmacokinetics and Pharmacokinetic–Pharmacodynamic Correlations of Therapeutic Peptides. **Clinical Pharmacokinetics**, v. 52, n. 10, p. 855–868, 1 out. 2013.
32. DIAS, R.; DE AZEVEDO, W. F. Molecular docking algorithms. **Current Drug Targets**, v. 9, n. 12, p. 1040–1047, dez. 2008.
33. DOMINGUEZ, C.; BOELEN, R.; BONVIN, A. M. J. J. HADDOCK: A Protein–Protein Docking Approach Based on Biochemical or Biophysical Information. **Journal of the American Chemical Society**, v. 125, n. 7, p. 1731–1737, 19 fev. 2003.
34. FASSIO, A. V. et al. nAPOLI: A Graph-Based Strategy to Detect and Visualize Conserved Protein-Ligand Interactions in Large-Scale. **IEEE/ACM transactions on computational biology and bioinformatics**, v. 17, n. 4, p. 1317–1328, 2020.
35. FISCHER, A. et al. Decision Making in Structure-Based Drug Discovery: Visual Inspection of Docking Results. **Journal of Medicinal Chemistry**, v. 64, n. 5, p. 2489–2500, 11 mar. 2021.
36. FOSGERAU, K.; HOFFMANN, T. Peptide therapeutics: current status and future

- directions. **Drug Discovery Today**, v. 20, n. 1, p. 122–128, 1 jan. 2015.
37. FRASER, B. J. et al. Structure and activity of human TMPRSS2 protease implicated in SARS-CoV-2 activation. **Nature Chemical Biology**, v. 18, n. 9, p. 963–971, set. 2022.
38. FUKUSHIMA, K.; KAMIMURA, T.; TAKIMOTO-KAMIMURA, M. Structure basis 1/2SLPI and porcine pancreas trypsin interaction. **Journal of Synchrotron Radiation**, v. 20, n. Pt 6, p. 943–947, 1 nov. 2013.
39. GARCÍA-PERDOMO, H. A. et al. Association between TMPRSS2:ERG fusion gene and the prostate cancer: systematic review and meta-analysis. **Central European Journal of Urology**, v. 71, n. 4, p. 410–419, 2018.
40. GIBO, J. et al. Camostat mesilate attenuates pancreatic fibrosis via inhibition of monocytes and pancreatic stellate cells activity. **Laboratory Investigation; a Journal of Technical Methods and Pathology**, v. 85, n. 1, p. 75–89, jan. 2005.
41. GILSON, M. K.; ZHOU, H.-X. Calculation of protein-ligand binding affinities. **Annual Review of Biophysics and Biomolecular Structure**, v. 36, p. 21–42, 2007.
42. GOODWIN, D.; SIMERSKA, P.; TOTH, I. Peptides As Therapeutics with Enhanced Bioactivity. **Current Medicinal Chemistry**, v. 19, n. 26, p. 4451–4461, [s.d.].
43. GOTTLIEB, R. L. et al. Early Remdesivir to Prevent Progression to Severe Covid-19 in Outpatients. **The New England Journal of Medicine**, v. 386, n. 4, p. 305–315, 27 jan. 2022.
44. GUEDES, I. A.; PEREIRA, F. S. S.; DARDENNE, L. E. Empirical Scoring Functions for Structure-Based Virtual Screening: Applications, Critical Aspects, and Challenges. **Frontiers in Pharmacology**, v. 9, p. 1089, 2018.
45. GUNST, J. D. et al. Efficacy of the TMPRSS2 inhibitor camostat mesilate in patients hospitalized with Covid-19-a double-blind randomized controlled trial. **EClinicalMedicine**, v. 35, p. 100849, maio 2021.
46. HASHEMIAN, S. M. R. et al. Paxlovid (Nirmatrelvir/Ritonavir): A new approach to Covid-19 therapy? **Biomedicine & Pharmacotherapy**, v. 162, p. 114367, jun. 2023.
47. HENNINOT, A.; COLLINS, J. C.; NUSS, J. M. The Current State of Peptide Drug Discovery: Back to the Future? **Journal of Medicinal Chemistry**, v. 61, n. 4, p. 1382–1414, 22 fev. 2018.
48. HEURICH, A. et al. TMPRSS2 and ADAM17 Cleave ACE2 Differentially and Only Proteolysis by TMPRSS2 Augments Entry Driven by the Severe Acute Respiratory Syndrome Coronavirus Spike Protein. **Journal of Virology**, v. 88, n. 2, p. 1293–1307,

- 2014.
49. HITOMI, Y.; IKARI, N.; FUJII, S. Inhibitory effect of a new synthetic protease inhibitor (FUT-175) on the coagulation system. **Haemostasis**, v. 15, n. 3, p. 164–168, 1985.
 50. HOFFMANN, M. et al. Camostat mesylate inhibits SARS-CoV-2 activation by TMPRSS2-related proteases and its metabolite GBPA exerts antiviral activity. **EBioMedicine**, v. 65, p. 103255, mar. 2021.
 51. HOFFMANN, M.; KLEINE-WEBER, H.; PÖHLMANN, S. A Multibasic Cleavage Site in the Spike Protein of SARS-CoV-2 Is Essential for Infection of Human Lung Cells. **Molecular Cell**, v. 78, n. 4, p. 779- 784.e5, 21 maio 2020.
 52. HULL, M. W.; MONTANER, J. S. G. Ritonavir-boosted protease inhibitors in HIV therapy. **Annals of Medicine**, v. 43, n. 5, p. 375–388, 1 ago. 2011.
 53. IKETANI, S. et al. Multiple pathways for SARS-CoV-2 resistance to nirmatrelvir. **Nature**, v. 613, n. 7944, p. 558–564, 2023.
 54. IMAI, K.; TAKAOKA, A. Comparing antibody and small-molecule therapies for cancer. **Nature Reviews. Cancer**, v. 6, n. 9, p. 714–727, set. 2006.
 55. INOUE, M. et al. Emergence of Oseltamivir-Resistant Pandemic (H1N1) 2009 Virus within 48 Hours. **Emerging Infectious Diseases**, v. 16, n. 10, p. 1633–1636, out. 2010.
 56. IVASHCHENKO, A. A. et al. The Efficacy of Aprotinin Combinations with Selected Antiviral Drugs in Mouse Models of Influenza Pneumonia and Coronavirus Infection Caused by SARS-CoV-2. **Molecules**, v. 27, n. 15, p. 4975, 5 ago. 2022.
 57. JIANG, L. et al. Insights into the serine protease mechanism based on structural observations of the conversion of a peptidyl serine protease inhibitor to a substrate. **Biochimica Et Biophysica Acta**, v. 1860, n. 3, p. 599–606, mar. 2016.
 58. JIN, Z. et al. Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. **Nature**, v. 582, n. 7811, p. 289–293, jun. 2020.
 59. KALIL, A. C. et al. Baricitinib plus Remdesivir for Hospitalized Adults with Covid-19. **The New England Journal of Medicine**, v. 384, n. 9, p. 795–807, 4 mar. 2021.
 60. KISHK, S. M. et al. Molecular Insights into Human Transmembrane Protease Serine-2 (TMPS2) Inhibitors against SARS-CoV2: Homology Modelling, Molecular Dynamics, and Docking Studies. **Molecules**, v. 25, n. 21, p. 5007, jan. 2020.
 61. KITCHEN, D. B. et al. Docking and scoring in virtual screening for drug discovery:

- methods and applications. **Nature Reviews. Drug Discovery**, v. 3, n. 11, p. 935–949, nov. 2004.
62. KO, C.-J. et al. Androgen-Induced TMPRSS2 Activates Matriptase and Promotes Extracellular Matrix Degradation, Prostate Cancer Cell Invasion, Tumor Growth, and Metastasis. **Cancer Research**, v. 75, n. 14, p. 2949–2960, 14 jul. 2015.
63. KO, C.-J. et al. Inhibition of TMPRSS2 by HAI-2 reduces prostate cancer cell invasion and metastasis. **Oncogene**, v. 39, n. 37, p. 5950–5963, set. 2020.
64. KONISHI, T. Mutations in SARS-CoV-2 are on the increase against the acquired immunity. **PLOS ONE**, v. 17, n. 7, p. e0271305, 11 jul. 2022.
65. LABUTE, M. X. et al. Adverse drug reaction prediction using scores produced by large-scale drug-protein target docking on high-performance computing machines. **PloS One**, v. 9, n. 9, p. e106298, 2014.
66. LAPORTE, M.; NAESENS, L. Airway proteases: an emerging drug target for influenza and other respiratory virus infections. **Current Opinion in Virology**, v. 24, p. 16–24, jun. 2017.
67. LASKOWSKI, R. A. et al. PROCHECK: a program to check the stereochemical quality of protein structures. **Journal of Applied Crystallography**, v. 26, n. 2, p. 283–291, 1 abr. 1993.
68. LAU, J. L.; DUNN, M. K. Therapeutic peptides: Historical perspectives, current development trends, and future directions. **Bioorganic & Medicinal Chemistry, Peptide Therapeutics**. v. 26, n. 10, p. 2700–2707, 1 jun. 2018.
69. LEE, A. C.-L. et al. A Comprehensive Review on Current Advances in Peptide Drug Development and Design. **International Journal of Molecular Sciences**, v. 20, n. 10, p. 2383, 14 maio 2019.
70. LI, X.-F.; LIU, C.-F.; RAO, G.-W. Monoclonal Antibodies, Small Molecule Inhibitors and Antibody-drug Conjugates as HER2 Inhibitors. **Current Medicinal Chemistry**, v. 28, n. 17, p. 3339–3360, 2021.
71. LIU, Q. et al. Emergence of a novel drug resistant H7N9 influenza virus: evidence based clinical potential of a natural IFN- α for infection control and treatment. **Expert Review of Anti-Infective Therapy**, v. 12, n. 2, p. 165–169, fev. 2014.
72. LÓPEZ-BLANCO, J. R. et al. iMODS: internal coordinates normal mode analysis server. **Nucleic Acids Research**, v. 42, n. Web Server issue, p. W271–W276, 1 jul. 2014.
73. LU, D. et al. Bovine proenteropeptidase is activated by trypsin, and the specificity of

- enteropeptidase depends on the heavy chain. **The Journal of Biological Chemistry**, v. 272, n. 50, p. 31293–31300, 12 dez. 1997.
74. MADEIRA, F. et al. Search and sequence analysis tools services from EMBL-EBI in 2022. **Nucleic acids research**, v. 50, n. W1, p. W276–W279, 1 jul. 2022.
75. MALIK, J. A. et al. The SARS-CoV-2 mutations versus vaccine effectiveness: New opportunities to new challenges. **Journal of Infection and Public Health**, v. 15, n. 2, p. 228–240, fev. 2022.
76. MANANDHAR, S. et al. Identification, virtual screening and molecular dynamic analysis of novel TMPRSS2 inhibitors from natural compound database as potential entry-blocking agents in SARS-CoV-2 therapy. **Structural Chemistry**, v. 33, n. 5, p. 1609–1617, 1 out. 2022.
77. MARTINS, M. M. et al. CLINICAL AND LABORATORY CHARACTERISTICS OF SARS-COV-2 INFECTION IN CHILDREN AND ADOLESCENTS. **Revista Paulista De Pediatria: Orgao Oficial Da Sociedade De Pediatria De Sao Paulo**, v. 39, p. e2020231, 2020.
78. MARTINS, P. M. et al. Propedia: a database for protein–peptide identification based on a hybrid clustering algorithm. **BMC Bioinformatics**, v. 22, n. 1, p. 1, 2 jan. 2021.
79. MCBROOME, J. et al. A Daily-Updated Database and Tools for Comprehensive SARS-CoV-2 Mutation-Annotated Trees. **Molecular Biology and Evolution**, v. 38, n. 12, p. 5819–5824, 1 dez. 2021.
80. MUKAI, S. et al. Matriptase and MET are prominently expressed at the site of bone metastasis in renal cell carcinoma: immunohistochemical analysis. **Human Cell**, v. 28, n. 1, p. 44–50, 1 jan. 2015.
81. MUTO, S.; IMAI, M.; ASANO, Y. Mechanisms of hyperkalemia caused by nafamostat mesilate. **General Pharmacology**, v. 26, n. 8, p. 1627–1632, dez. 1995.
82. NEITZEL, J. J. (2010) Enzyme Catalysis: The Serine Proteases . Nature Education 3(9):21. Disponível em: <https://www.nature.com/scitable/topicpage/enzyme-catalysis-the-serine-proteases-nbsp-14398894/>. Acesso em: 22 dez. 2023.
83. OTVOS, L.; WADE, J. D. Current challenges in peptide-based drug discovery. **Frontiers in Chemistry**, v. 2, p. 62, 2014.
84. PERONA, J. J.; CRAIK, C. S. Evolutionary Divergence of Substrate Specificity within the Chymotrypsin-like Serine Protease Fold *. **Journal of Biological Chemistry**, v. 272, n. 48, p. 29987–29990, 28 nov. 1997.

85. PETTA, I. et al. Modulation of Protein-Protein Interactions for the Development of Novel Therapeutics. **Molecular Therapy: The Journal of the American Society of Gene Therapy**, v. 24, n. 4, p. 707–718, abr. 2016.
86. PINZI, L.; RASTELLI, G. Molecular Docking: Shifting Paradigms in Drug Discovery. **International Journal of Molecular Sciences**, v. 20, n. 18, p. 4331, 4 set. 2019.
87. RAMACHANDRAN, G. N.; RAMAKRISHNAN, C.; SASISEKHARAN, V. Stereochemistry of polypeptide chain configurations. **Journal of Molecular Biology**, v. 7, n. 1, p. 95–99, 1 jul. 1963.
88. REDONDO-CALVO, F. J. et al. Aprotinin treatment against SARS-CoV-2: A randomized phase III study to evaluate the safety and efficacy of a pan-protease inhibitor for moderate COVID-19. **European Journal of Clinical Investigation**, v. 52, n. 6, p. e13776, jun. 2022.
89. ROCHA, L. (2023, 07 de março). Anvisa suspende temporariamente autorização do medicamento Evusheld para Covid-19. CNN Brasil. <https://www.cnnbrasil.com.br/saude/anvisa-suspende-temporariamente-autorizacao-do-medicamento-evusheld-para-covid-19/>. Acesso em: 21 set. 2023.
90. ROSENKE, K. et al. Orally delivered MK-4482 inhibits SARS-CoV-2 replication in the Syrian hamster model. **Nature Communications**, v. 12, n. 1, p. 2295, 16 abr. 2021.
91. ROUSSEL, A. et al. Complexation of two proteic insect inhibitors to the active site of chymotrypsin suggests decoupled roles for binding and selectivity. **The Journal of Biological Chemistry**, v. 276, n. 42, p. 38893–38898, 19 out. 2001.
92. S. KIM, T. et al. **Phenotypic Analysis of Mice Lacking the Tmprss2-Encoded Protease: Molecular and Cellular Biology: Vol 26, No 3**. Disponível em: <<https://www.tandfonline.com/doi/full/10.1128/MCB.26.3.965-975.2006>>. Acesso em: 29 set. 2023.
93. SANDERSON, T. et al. A molnupiravir-associated mutational signature in global SARS-CoV-2 genomes. **Nature**, p. 1–3, 25 set. 2023.
94. SEHDEV, S. Sunitinib toxicity management - a practical approach. **Canadian Urological Association Journal = Journal De l'Association Des Urologues Du Canada**, v. 10, n. 11-12Suppl7, p. S248–S251, 2016.
95. SHARMA, K. et al. Peptide-based drug discovery: Current status and recent advances. **Drug Discovery Today**, v. 28, n. 2, p. 103464, 1 fev. 2023.
96. SHAWAN, M. M. A. K. et al. Advances in Computational and Bioinformatics Tools

- and Databases for Designing and Developing a Multi-Epitope-Based Peptide Vaccine. **International Journal of Peptide Research and Therapeutics**, v. 29, n. 4, p. 60, 2023.
97. SHEN, L. W. et al. TMPRSS2: A potential target for treatment of influenza virus and coronavirus infections. **Biochimie**, v. 142, p. 1–10, 1 nov. 2017.
98. SHINE, N. R. et al. Secretory leukocyte protease inhibitor: inhibition of human immunodeficiency virus-1 infection of monocytic THP-1 cells by a newly cloned protein. **Bioorganic Chemistry**, v. 30, n. 4, p. 249–263, ago. 2002.
99. SHIRATO, K.; KAWASE, M.; MATSUYAMA, S. Middle East respiratory syndrome coronavirus infection mediated by the transmembrane serine protease TMPRSS2. **Journal of Virology**, v. 87, n. 23, p. 12552–12561, dez. 2013.
100. SHIROGANE, Y. et al. Efficient Multiplication of Human Metapneumovirus in Vero Cells Expressing the Transmembrane Serine Protease TMPRSS2. **Journal of Virology**, v. 82, n. 17, p. 8942–8946, set. 2008.
101. SHUVO, M.; GULFAM, M.; BHATTACHARYA, D. DeepRefiner: High-accuracy protein structure refinement by deep network calibration. **Nucleic Acids Research**, v. 49, 17 maio 2021.
102. SIMMONS, G. et al. Inhibitors of cathepsin L prevent severe acute respiratory syndrome coronavirus entry. **Proceedings of the National Academy of Sciences of the United States of America**, v. 102, n. 33, p. 11876–11881, 16 ago. 2005.
103. SIMMONS, G. et al. Proteolytic activation of the SARS-coronavirus spike protein: Cutting enzymes at the cutting edge of antiviral research. **Antiviral Research**, v. 100, n. 3, p. 605–614, dez. 2013.
104. SOFI, M. S.; HAMID, A.; BHAT, S. U. SARS-CoV-2: A critical review of its history, pathogenesis, transmission, diagnosis and treatment. **Biosafety and Health**, v. 2, n. 4, p. 217–225, dez. 2020.
105. SOMMERHOFF, C. P. et al. A Kazal-type inhibitor of human mast cell tryptase: isolation from the medical leech *Hirudo medicinalis*, characterization, and sequence analysis. **Biological Chemistry Hoppe-Seyler**, v. 375, n. 10, p. 685–694, out. 1994.
106. SPINNER, C. D. et al. Effect of Remdesivir vs Standard Care on Clinical Status at 11 Days in Patients With Moderate COVID-19: A Randomized Clinical Trial. **JAMA**, v. 324, n. 11, p. 1048–1057, 15 set. 2020.
107. STOPSACK, K. H. et al. TMPRSS2 and COVID-19: Serendipity or Opportunity for Intervention? **Cancer Discovery**, v. 10, n. 6, p. 779–782, jun. 2020.

108. STUBBS, M. T. et al. The three-dimensional structure of recombinant leech-derived trypsin inhibitor in complex with trypsin. Implications for the structure of human mast cell trypsin and its inhibition. **The Journal of Biological Chemistry**, v. 272, n. 32, p. 19931–19937, 8 ago. 1997.
109. TARNOW, C. et al. TMPRSS2 is a host factor that is essential for pneumotropism and pathogenicity of H7N9 influenza A virus in mice. **Journal of Virology**, v. 88, n. 9, p. 4744–4751, maio 2014.
110. THUNDERS, M.; DELAHUNT, B. Gene of the month: TMPRSS2 (transmembrane serine protease 2). **Journal of Clinical Pathology**, v. 73, n. 12, p. 773–776, dez. 2020.
111. TSENG, C.-C. et al. Matriptase shedding is closely coupled with matriptase zymogen activation and requires de novo proteolytic cleavage likely involving its own activity. **PLoS ONE**, v. 12, n. 8, 2017.
112. ULLRICH, S.; NITSCHKE, C. The SARS-CoV-2 main protease as drug target. **Bioorganic & Medicinal Chemistry Letters**, v. 30, n. 17, p. 127377, 1 set. 2020.
113. VAARALA, M. H. et al. Expression of transmembrane serine protease TMPRSS2 in mouse and human tissues. **The Journal of Pathology**, v. 193, n. 1, p. 134–140, jan. 2001.
114. VUONG, H. G. et al. Efficacy and toxicity of sorafenib in the treatment of advanced medullary thyroid carcinoma: A systematic review and meta-analysis. **Head & Neck**, v. 41, n. 8, p. 2823–2829, ago. 2019.
115. WANG, L. et al. Therapeutic peptides: current applications and future directions. **Signal Transduction and Targeted Therapy**, v. 7, p. 48, 14 fev. 2022.
116. WENG, G. et al. Comprehensive Evaluation of Fourteen Docking Programs on Protein–Peptide Complexes. **Journal of Chemical Theory and Computation**, 23 abr. 2020.
117. WU, S.-R. et al. Matriptase Is Involved in ErbB-2-Induced Prostate Cancer Cell Invasion. **The American Journal of Pathology**, v. 177, n. 6, p. 3145–3158, dez. 2010.
118. WU, Y. et al. Flipping in the Pore: Discovery of Dual Inhibitors That Bind in Different Orientations to the Wild-Type versus the Amantadine-Resistant S31N Mutant of the Influenza A Virus M2 Proton Channel. **Journal of the American Chemical Society**, v. 136, n. 52, p. 17987–17995, 31 dez. 2014.
119. YAN, Y. et al. HDOCK: a web server for protein–protein and protein–DNA/RNA

- docking based on a hybrid strategy. **Nucleic Acids Research**, v. 45, n. Web Server issue, p. W365–W373, 3 jul. 2017.
120. YUAN, Y. et al. The development of COVID-19 treatment. **Frontiers in Immunology**, v. 14, 2023.
121. ZAKHAROVA, E.; HORVATH, M. P.; GOLDENBERG, D. P. Structure of a serine protease poised to resynthesize a peptide bond. **Proceedings of the National Academy of Sciences of the United States of America**, v. 106, n. 27, p. 11034–11039, 7 jul. 2009.
122. ZANI, M.-L. et al. Protease inhibitors derived from elafin and SLPI and engineered to have enhanced specificity towards neutrophil serine proteases. **Protein Science : A Publication of the Protein Society**, v. 18, n. 3, p. 579–594, mar. 2009.
123. ZHANG, S. et al. ProDy 2.0: increased scale and scope after 10 years of protein dynamics modelling with Python. **Bioinformatics**, v. 37, n. 20, p. 3657–3659, 25 out. 2021.
124. ZHIRNOV, O. P.; KLENK, H. D.; WRIGHT, P. F. Aprotinin and similar protease inhibitors as drugs against influenza. **Antiviral Research**, v. 92, n. 1, p. 27–36, out. 2011.
125. ZHOU, C. K. et al. TMPRSS2:ERG Gene Fusions in Prostate Cancer of West African Men and a Meta-Analysis of Racial Differences. **American Journal of Epidemiology**, v. 186, n. 12, p. 1352–1361, 15 dez. 2017.
126. ZHOU, P. et al. HPEPDOCK: a web server for blind peptide-protein docking based on a hierarchical algorithm. **Nucleic Acids Research**, v. 46, n. W1, p. W443–W450, 2 jul. 2018.
127. ZHU, C.-T. et al. Appraisal of evidence reliability and applicability of Paxlovid as treatment for SARS-COV-2 infection: A systematic review. **Reviews in Medical Virology**, p. e2476, 14 ago. 2023.