



CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS
UNIVERSIDADE FEDERAL DE SÃO JOÃO DEL-REI

**Utilização de técnicas de desenvolvimento racional de
fármacos no desenvolvimento de Antivirais contra o
COVID19**

**Projeto de extensão apresentado ao CEFET chamada 32/2020 - SELEÇÃO
PÚBLICA PARA APOIO A PROJETOS DE EXTENSÃO EMERGENCIAIS
VISANDO O ENFRENTAMENTO DO COVID-19.**

Proponentes:

Prof. Eduardo Habib Bechelane Maia (CEFET-MG)

Prof. Alisson Marques da Silva (CEFET-MG)

Prof. Tiago Alves de Oliveira (CEFET-MG)

Prof. Alex Gutterres Taranto (UFSJ)

Divinópolis - 2020

Sumário

1. DADOS DA AÇÃO DE EXTENSÃO.....	3
2. Resumo	5
3. Objetivos gerais.....	5
a. Objetivos específicos:	5
4. Justificativa.....	6
5. Fundamentação teórica	6
6. Metodologia.....	10
a. CORONAVIRUS DATA BASE (CODAB-19).....	10
b. DESENVOLVIMENTO DE FÁRMACOS COM BASE NO LIGANTE E PREPARAÇÃO DOS <i>HITS</i>	11
c. DESENVOLVIMENTO DE FÁRMACOS COM BASE NA ESTRUTURA 	12
7. RESULTADOS ESPERADOS	12
8. PROPRIEDADE INTELECTUAL.....	13
9. BOLSISTAS	13
10. Parceiros	15
11. Referências.....	15

1. DADOS DA AÇÃO DE EXTENSÃO

Código: PJ135-2020

Título: Utilização de técnicas de desenvolvimento racional de fármacos no desenvolvimento de Antivirais contra o COVID-19

Ano: 2020

Período de Realização: 27/04/2020 a 26/10/2020

Tipo: PROJETO

Situação: EM EXECUÇÃO

Município de Realização:

Espaço de Realização:

Abrangência: Nacional

Público Alvo: Discengens do curso e professores que atuarão no desenvolvimento racional de um antiviral

Unidade Proponente: DVN - DEPARTAMENTO DE INFORMÁTICA, GESTÃO E DESIGN / CEFETMG

Unidade Orçamentária: -

Outras Unidades Envolvidas:

DVN - CURSO TÉCNICO EM INFORMÁTICA / CEFETMG

DVN - CURSO DE ENGENHARIA DE COMPUTACAO / CEFETMG

Área Principal: Saúde

Área do CNPq: Ciências Exatas e da Terra

Fonte de Financiamento: FINANCIAMENTO INTERNO (Edital 32/2020 - Seleção pública para apoio a PROJETOS de extensão emergenciais visando o enfrentamento do COVID-19)

Convênio Fundação: NÃO

Renovação: NÃO

Nº Bolsas Solicitadas: 3

Nº Bolsas Concedidas: 3

Nº Discentes Envolvidos: 4

Faz parte de Programa de Extensão: NÃO

Grupo Permanente de Arte e Cultura: NÃO

Público Estimado: 156 pessoas

Público Real Atendido: Não informado

Tipo de Cadastro: SUBMISSÃO DE NOVA PROPOSTA

Coordenação: EDUARDO HABIB BECHELANE MAIA

2. Resumo

Coronavírus (CoVs) é uma doença infecciosa causada pelo coronavírus da síndrome respiratória aguda grave 2. Em todo o mundo, mais de 89 mil pessoas morreram em decorrência da doença. Apesar do impacto epidemiológico, existem poucos fármacos para o tratamento desses pacientes. Uma forma de acelerar a descoberta de novos fármacos é através da utilização de técnicas computacionais capazes de explorar informações sobre alvos validados, exclusivos e/ou essenciais ao vírus. Neste contexto, o desenvolvimento de fármacos é uma área muito explorada atualmente. Muitas tecnologias surgem no mercado a fim de tornar esses estudos mais rápidos e com menores custos. Dentre as diversas metodologias utilizadas pela indústria farmacêutica, a triagem virtual (TV) é uma ferramenta de bioinformática que busca direcionar a seleção de moléculas orgânicas para alvos terapêuticos de interesse. Nos últimos anos, o grupo vem aplicando esta abordagem para selecionar moléculas ativas com sucesso contra os arbovírus Dengue, Mayaro e Zika e contra a Malária. Esta abordagem permite identificar compostos bioativos utilizando métodos *in silico*, os quais podem ser aperfeiçoados por síntese orgânica e posteriormente encaminhados para ensaios biológicos. Inicialmente, será construído um banco de alvos terapêuticos do Coronavírus (CoVs) através das estruturas depositadas no *Protein Data Bank* ou construídas por modelagem comparativa. Para cada alvo terapêutico, serão prospectados possíveis inibidores por meio do programa Pharmit, o qual realiza busca farmacofórica de potenciais inibidores, utilizando a abordagem de desenho de fármacos com base no ligante. Os ligantes mais promissores serão filtrados quanto as suas propriedades farmacocinéticas utilizando o programa Data Warrior. A seguir, os ligantes filtrados serão ancorados contra o respectivo alvo e ranqueados pela abordagem de desenvolvimento de fármacos com base na estrutura por meio do programa desenvolvido pelo grupo, denominado MolAr. Assim, os resultados teóricos irão motivar a realização do trabalho experimental específico para cada alvo molecular encontrado. Os *hits* encontrados na etapa de bioinformática poderão, futuramente, serem avaliados por bioensaios quanto a sua toxicidade e possível atividade antiviral. Com isso, busca-se diminuir o tempo e custo no processo de descoberta de um novo antiviral contra o CoVs. Além disso, o presente projeto objetiva promover o maior intercâmbio entre químicos, detentores de moléculas, com biólogos e a computação.

3. Objetivos gerais

Descobrir compostos, com razoável atividade antiviral contra o Coronavirus, através de métodos de bioinformática (*in silico*) para que posteriormente eles possam ser enviados para a realização de ensaios biológicos.

a. Objetivos específicos:

- Elaborar banco de alvos moleculares contra o Coronavírus (CoVs) denominado Coronavirus Data Base (CoDaB-19);
- realizar estudos de triagem virtual com base no ligante;
- realizar estudos de triagem virtual com base no receptor;
- prever a atividade antiviral de fármacos, compostos naturais da flora brasileira, e demais compostos presente na base de dados ZINC;

- disponibilizar o CoDaB-19 permitindo o acesso para a comunidade acadêmica;
- realizar pequenas modificações moleculares nos *hits* encontrados para aperfeiçoar as interações ligante-receptor;
- submeter os *hits*, modificados ou não, aos ensaios experimentais;
- publicar um artigo científico e de ampla circulação no meio acadêmico;
- solicitar a propriedade intelectual de moléculas promissoras quando pertinente;
- contribuir com a formação de alunos de IC;
- divulgar os resultados no site www.drugdiscovery.com.br e no canal do YouTube do Laboratório de Bioinformática (<https://www.youtube.com/channel/UCtpnd1JnDAPk4D6aSv7cBZw>).

4. Justificativa

O possível avanço científico e tecnológico desse projeto, somado aos projetos anteriores e posteriores na área de desenvolvimento de fármacos, poderá permitir que o Brasil passe de uma posição de simples consumidor para descobridor de novos agentes terapêuticos. O presente projeto também irá contribuir para a formação do professor do CEFET-MG Tiago Alves de Oliveira que atualmente é doutorando no programa de pós graduação em Biotecnologia da UFSJ (PPGBiotec/UFSJ). Além disso, todos os membros da equipe participarão do projeto com a publicação de artigos e participação em congressos, contribuindo para a uma formação de melhor qualidade.

Este projeto também visa colaborar com outros pesquisadores na área de desenvolvimento de fármacos pelo fornecimento do Coronavírus Data Base (CoDaB-19) através do site www.drugdiscovery.com.br. Conseqüentemente, diferentes grupos poderão prever a atividade de moléculas as quais ainda não foram sintetizadas. Com isso, espera-se uma economia de tempo e custos no desenvolvimento de moléculas com atividade antiviral contra o Coronavírus. Adicionalmente, moléculas biologicamente ativas podem levar ao desenvolvimento de um produto elaborado, com valor agregado, resultando em patente.

5. Fundamentação teórica

Os Coronavírus (CoVs) consistem em um diverso grupo de vírus envelopados pertencentes à família *Coronaviridae* que relacionados com a etiologia de doenças respiratórias, entéricas, hepáticas e neurológicas de gravidade variável em um amplo espectro de espécies de animais (WOO *et al.*, 2009). A diversidade dos CoVs está

relacionada com: (1) altas taxas de mutação durante a replicação do genoma; (2) a incidência de recombinações homólogas no genoma e (3) o tamanho do genoma em relação a outros vírus de RNA (DUFFY; SHACKELTON; HOLMES, 2008; LAI, 1992).

Em relação aos critérios taxonômicos, os CoVs são subdivididos em quatro gêneros: *Alphacoronavirus*, *Betacoronavirus*, *Gammacoronavirus* e *Deltacoronavirus*. Pertencentes ao gênero *Betacoronavirus*, os vírus causadores da síndrome respiratória do Oriente Médio (MERS-CoV) e síndrome respiratória aguda grave (SARS-CoV) receberam destaque pela sua emergência global nos anos de 2002 e 2012, respectivamente (ANTONIO *et al.*, 2003; ZAKI *et al.*, 2012). No final de dezembro de 2019, um surto de pneumonia de etiologia desconhecida foi detectado na cidade chinesa de Wuhan (BOGOCH *et al.*, 2020; LU, HONGZHOU; STRATTON; TANG, 2020). Em 7 de janeiro de 2020, o agente etiológico foi identificado como um novo tipo de Coronavírus associado à síndrome respiratória aguda grave, o SARS-CoV-2 (LU, ROUJIAN *et al.*, 2020). A ocorrência de surtos e epidemias avançou rapidamente pelo mundo e a Organização Mundial da Saúde declarou situação de pandemia para a doença causada pelo SARS-CoV-2, então denominada de COVID-19.

A principal rota de transmissão SARS-CoV-2 ocorre por meio de secreções respiratórias provenientes de humanos infectados, bem como pelo contato direto com objetos e superfícies contaminadas (YANG; WANG, 2020). Na primeira etapa do ciclo de multiplicação ocorre a ligação de SARS-CoV-2 com a enzima conversora de angiotensina 2 (ACE2) na superfície das células (HOFFMANN *et al.*, 2020). Em seguida, as partículas virais são internalizadas por endocitose, armazenadas no interior de endossomos e o genoma viral é liberado no citoplasma da célula hospedeira (ZUMLA *et al.*, 2016).

A organização genômica do SARS-CoV-2 é similar à de outros membros do gênero *Betacoronavirus*, que se constitui de uma fita simples de RNA de polaridade positiva [ssRNA(+)] contendo cerca de 30 mil nucleotídeos (WU *et al.*, 2020). O ssRNA(+) do SARS-CoV-2 é constituído de uma região não codificante 5', um complexo replicase (orf1ab) que codifica para 16 proteínas não estruturais (NSp), de quatro genes que codificam para quatro proteínas estruturais (spike [S], envelope [E], membrana [M] e nucleocapsídeo [N]), seguidos de mais uma região codificadora 3' e diversas janelas abertas de leitura (ORFs). Entretanto, o genoma do SARS-CoV-2 não possui o gene que codifica para a proteína hemaglutinina-esterase, característico de outras linhagens de

coronavírus (CHAN *et al.*, 2020). Entre as proteínas não estruturais alguns domínios já são caracterizados pela atividade de cisteínoproteases (NSp3 e NSp5), uma RNA polimerase dependente de RNA (NSp12), RNA helicase (NSp13) e RNA 2'-OMetiltransferase (NSp16), além do papel de outros domínios NSp com envolvidos nos processos de transcrição e replicação (CHAN *et al.*, 2020; CHEN; LIU; GUO, 2020).

Atualmente não existem vacinas para prevenção, e tampouco fármacos específicos para o tratamento da COVID-19. Este cenário é desafiador pela existência de lacunas nos mecanismos de patogênese, diagnóstico dificultado e falhas nas principais medidas de distanciamento social. A indústria farmacêutica tem investido em estratégias modernas e inovadoras de pesquisa para aumentar a liberação de novos fármacos no mercado. As principais estratégias empregam métodos de bioinformática tais como triagem virtual, permitindo a identificação e planejamento de compostos bioativos, e tornando o processo de desenvolvimento de fármacos racional e promissor. Recentemente, nosso grupo de pesquisa obteve resultados promissores que permitiram identificar candidatos a antivirais contra o Zika vírus (SANTOS; LIMA; *et al.*, 2020; SANTOS; NUNES; *et al.*, 2020).

Entre os métodos de concepção de fármacos, a pesquisa sistemática consiste em procurar a atividade biológica a partir de um composto desejado ou de um conjunto de compostos, quer sejam de fontes naturais ou sintéticas. Este é um processo de pesquisa exaustivo e demorado. Recentemente, a robótica foi incorporada neste método, que tem sido denominado de triagem de alta produção (HTS, do inglês *High-throughput screening*). O método HTS tem a capacidade de avaliar milhares de compostos simultaneamente em 30-50 diferentes ensaios bioquímicos (POLGAR; M. KESERU, 2011). Apesar do HTS ter contribuído para a aprovação de novos fármacos no mercado, esta abordagem é limitada pelo seu alto custo de trabalho experimental, motivando o desenvolvimento de novas tecnologias semelhantes ao conceito de HTS. Dentre as novas tecnologias, destaca-se a ancoragem molecular (*docking*). Esta é uma abordagem *in silico* que permite a identificação de compostos-alvo a partir de estruturas tridimensionais do receptor de interesse. O programa de ancoragem mede a afinidade de pequenas moléculas contra um alvo molecular para determinar a energia de interação desses complexos. Como resultado, o encaixe pode sugerir os compostos mais promissores para ensaios biológicos, diminuindo assim os custos no processo de desenvolvimento de fármacos. A abordagem de ancoragem molecular pode ser aplicada a uma biblioteca de ligantes contra alvos moleculares específicos, sendo conhecida como triagem virtual baseado na estrutura do

receptor (VS, do inglês *Virtual Screening*) (GLAAB, 2016; WESTERMAIER; BARRIL; SCAPOZZA, 2015). Em resumo, os métodos *in silico*, como o acoplamento molecular e triagem virtual, reduzem o tempo e o custo do desenvolvimento de fármacos. O processo de VS tornou-se difundido com a implementação de diversas bibliotecas de compostos, como por exemplo ZINC15 (que significa "ZINC não é comercial") (STERLING; IRWIN, 2015). O ZINC15 é um banco de moléculas curado (manipulado por humanos) com mais de 230 milhões de moléculas disponíveis comercialmente, preparadas para experiências *in silico*. Estas propriedades o tornam uma importante ferramenta de bioinformática para pesquisar novos compostos bioativos.

Por outro lado, o banco de receptores amplamente utilizado pela comunidade acadêmica, *Protein Data Bank* (PDB), apresenta um crescimento bastante acelerado de depósitos de alvos moleculares contra o COVID-19, úteis para o planejamento de fármacos. Dentre os alvos depositados até a presente elaboração deste projeto, estão por exemplo, protease, enzima conversora de angiotensina 2 (do inglês angiotensin converting enzyme 2 (ACE2)) e a glicoproteína *spike*, sob os códigos 6LU7, 6M17/6VW1 e 6LXT, respectivamente. Adicionalmente, os demais alvos podem ser construídos por modelagem comparativa por homologia, uma vez que o genoma do vírus está descrito (WEBB; SALI, 2016). Estes depósitos motivam a criação de um banco previamente preparado para experimentos de VS, de forma semelhante aos bancos Our Own Molecular Targets Data Bank (OOMT) (CARREGAL *et al.*, 2017) e Brazilian Malaria Molecular Targets (BraMMT) (NUNES *et al.*, 2019), os quais são distribuídos junto com o programa Molecular Architect (MolAr) (MAIA *et al.*, 2020) através do site www.drugdiscovery.com.br.

Dessa forma, após as simulações *in silico* empregando o banco de ligantes ZINC contra um banco de alvos moleculares do COVID-19, os compostos mais promissores podem ser adquiridos e modificados em quantidade suficiente para serem utilizados em estudos experimentais.

Portanto, a triagem virtual representa uma forma rápida e válida para pesquisas científicas de novos compostos com potencial farmacológico (LAURO *et al.*, 2012). A modelagem molecular possibilita analisar um grande número de moléculas em curto espaço de tempo, demonstrando a forma de ligação das mesmas com os receptores de interesse farmacológico antes mesmo de sua síntese. Em resumo, o processo racional de fármacos inicia-se com a identificação de alvos moleculares para um determinado

composto (natural ou sintético), seguido de uma cuidadosa análise, no âmbito molecular, das principais forças de reconhecimento molecular entre o ligante e o receptor, desenho de moléculas mais promissoras, e finalmente os ensaios experimentais (NAEEM; HYLANDS; BARLOW, 2012).

6. Metodologia

a. CORONAVIRUS DATA BASE (CODAB-19)

Inicialmente, as estruturas tridimensionais de proteínas do SARS-CoV-2 serão obtidas do banco de dados Protein Data Bank (PDB - <https://www.rcsb.org/>) (BERMAN *et al.*, 2002). Para a busca serão utilizadas as palavras-chave “SARS-CoV-2”, “nCoV 2019” e “COVID-19”. A seguir, as estruturas serão selecionadas quanto a presença de ligantes cristalográficos na região do sítio de ligação e em relação aos parâmetros de resolução e qualidade da estrutura. Modelos tridimensionais da protease viral, proteína *spike* e metiltransferase do SARS-COV-2 já se encontram disponível para o início das análises de bioinformática. Além disso, modelos tridimensionais para as proteínas com atividade de helicase e RNA polimerase dependente de RNA já podem ser construídos através de modelagem por homologia.

Para avaliar o potencial das estruturas tridimensionais selecionadas como alvos moleculares será determinado o valor de *Druggability*, que é definido como habilidade em que sítios de ligação do alvo se ligam à compostos com propriedades físico-químicas semelhantes à fármacos. Desse modo, as estruturas selecionadas serão submetidas ao cálculo do *Druggability* utilizando servidores online como o PockDrug (Pock Drugability Prediction - <http://pockdrug.rpbs.univ-paris-diderot.fr/>) (BORREL *et al.*, 2015)

As estruturas dos alvos moleculares serão preparadas para as simulações de ancoragem molecular por meio da reconstrução das regiões de alça com estruturas homólogas através do programa MODELLER (WEBB; SALI, 2016). O estado de protonação das proteínas será ajustado para o pH 7,4, correspondente ao pH de suas localizações celulares. A seguir, o complexo ligante-receptor cristalográfico será refinado por minimização de energia através dos algoritmos steepest descent e gradiente conjugado, com 100 e 10 passos, respectivamente, através do programa Chimera (PETTERSEN *et al.*, 2004).

Em seguida, para validar os alvos moleculares serão utilizadas as estratégias de reancoragem molecular, curva ROC (ROC) e área sob a curva (AUC) (FAWCETT, 2006;

HEVENER *et al.*, 2009). Na validação por reancoragem molecular, os ligantes cristalográficos serão removidos do sítio de ligação com os seus respectivos alvos moleculares e em seguida serão reancorados. As conformações resultantes da reancoragem serão sobrepostos para avaliar a fidelidade conformacional pelo cálculo do desvio quadrático médio (RMSD, do inglês, “Root-mean square deviation”) utilizando o programa Discovery Studio Visualizer. Desse modo, o valor de RMSD resultante da sobreposição dos átomos pesados do ligante reancorado com os do ligante cristalográfico deve ser menor que 2.0Å (WESTERMAIER; BARRIL; SCAPOZZA, 2015).

A curva ROC (ROC) e área sob a curva (AUC) serão determinadas para cada alvo molecular com o objetivo de avaliar a especificidade e sensibilidade da metodologia em diferenciar moléculas ativas de falsos-positivos. A estrutura de compostos ativos para cada alvo molecular com valores de IC₅₀ menores que 1µM será obtida a partir de buscas na literatura. Em seguida, os decoys serão gerados utilizando a plataforma online DUDE (<http://dude.docking.org/>) desenvolvida por MYSINGER e colaboradores (2012). Os decoys consistem em moléculas que apresentam propriedades físico-químicas similares aos compostos ativos, tais como massa molecular, número de ligações rotáveis, lipofilicidade, número de grupos aceptores e doadores de ligação de hidrogênio, mas que se diferem topologicamente de moléculas ativas.

As estruturas dos compostos ativos e dos decoys serão preparadas e ancoradas aos sítios de ligação de seus respectivos alvos moleculares nas coordenadas dos ligantes cristalográficos utilizando o programa MolAr (MAIA *et al.*, 2020), gerando as curvas ROC correspondente para cada alvo.

b. DESENVOLVIMENTO DE FÁRMACOS COM BASE NO LIGANTE E PREPARAÇÃO DOS *HITS*

Os grupamentos farmacofóricos dos ligantes cristalográficos presentes nos alvos moleculares serão utilizados para a busca de *hits* através da plataforma Pharmit (SUNSERI; KOES, 2016). O Pharmit realiza alinhamento estrutural entre ligantes presentes no ZINC contra os grupos farmacofóricos de ligantes cristalográficos. Os *hits* encontrados serão classificados quanto ao cálculo do RMSD. Em seguida, estes ligantes serão selecionados não permitindo violações da regra dos cinco de Lipinski e quanto as propriedades fármaco semelhante (*druglikeness*) (WALTERS; MURCKO; MURCKO, 1999) através do programa DataWarrior (SANDER *et al.*, 2015). Após o ajuste estado de

protonação no pH 7,4 através do programa Open Babel (O'BOYLE *et al.*, 2011), os ligantes selecionados serão refinados pelo método paramétrico 7 (PM7) (STEWART, 2013) implementado no MOPAC2016 usando a rotina run-mopac do programa *Octopus* (MAIA *et al.*, 2017) pós o ajuste do estado de protonação no pH 7,4 através do programa Open Babel (O'BOYLE *et al.*, 2011). As palavras chaves serão: PM7, charge, EF, PDB e PDBOUT.

c. DESENVOLVIMENTO DE FÁRMACOS COM BASE NA ESTRUTURA

As moléculas obtidas na etapa anterior serão ancoradas contra os receptores envolvidos na DP. A fim de automatizar os processos da triagem, esse grupo de pesquisa desenvolveu o programa MolAr (MAIA *et al.*, 2020), o qual é uma plataforma que realiza cálculos de triagem virtual de alta produção. Ele gerencia diversos programas, tais como: MOPAC2016, *Octopus* (MAIA *et al.*, 2017), Modeller (WEBB; SALI, 2016), PROPKA,

PROCHECK, AutoDock Vina (TROTT; OLSON, 2009) e DOCK 6 (Grid e Amber scores) (BROZELL *et al.*, 2012). O MolAr permite a utilização da melhor conformação obtida pelo AutoDock Vina como entrada para o DOCK 6 aperfeiçoando os resultados e ranqueando os ligantes através de consenso score (HOUSTON; WALKINSHAW, 2013).

A ideia principal dessa técnica de consenso é que a combinação de duas abordagens diferentes na triagem virtual é melhor do que a aplicação de uma única abordagem sozinha (KUKOL, 2011).

7. RESULTADOS ESPERADOS

Espera-se, ao fim desse projeto que sejam realizadas a publicação de artigos em revistas conceituadas pela comunidade acadêmica e a participação em algum congresso e que a base de dados CoDaB-19 seja disponibilizada para a comunidade acadêmica. Além disso, moléculas com características físico-químicas desejadas e, possivelmente, com maior probabilidade de se tornarem um antiviral para covid-19 serão disponibilizadas para que possam ser enviados para a realização de ensaios biológicos. Existe ainda a possibilidade de criação de um medicamento de combate ao Coronavírus que seria desenvolvido posteriormente, a partir dos resultados obtidos nesse trabalho. Por fim, todo o arcabouço para testes computacionais futuros em novas epidemias de saúde por

Coronavírus, ou outra epidemia estarão disponíveis gratuitamente juntamente com o software MolAr.

8. PROPRIEDADE INTELECTUAL

As moléculas mais promissoras poderão ser patenteadas caso o ensaio biológico confirme a viabilidade apresentada nesse trabalho.

9. BOLSISTAS

1) Bolsista 1

A) Pesquisa Bibliográfica: Pesquisa bibliográfica sobre a triagem virtual em Coronavírus e a utilização do Pharmit e do data Warrior nesse processo.

B) Preparação da base de dados de ligantes: Estudo e download da base de dados.

C) Triagem virtual baseada no ligante: Realização da triagem virtual baseada no ligante. Nessa utiliza-se o Pharmit para a realização de uma busca farmacofórica entre os ligantes presentes no ZINC e os ligantes cristalográficos. Em seguida, será utilizado o DataWarrior para filtrar, dentre os selecionados pelo Pharmit, os ligantes mais promissores segundo suas propriedades farmacocinéticas.

D) Refinamento dos ligantes: Os ligantes serão refinados utilizando o MOPAC2016 e ajustados para o pH 7,4.

E) Manutenção do site: www.drugdiscovery.com.br e divulgação dos resultados parciais do projeto

Cronograma de Atividades:

Atividade	Mês					
	1	2	3	4	5	6
A	X					
B	X					
C	X	X	X	X	X	
D	X	X	X	X	X	
E			X	X	X	X

2) Bolsista 2

A) **Pesquisa Bibliográfica:** Pesquisa bibliográfica sobre a triagem virtual em Coronavirus e a utilização do MolAr nesse processo.

B) **Obtenção da curva ROC e Área sobre a curva:** Geração da Curva ROC e cálculo da área sobre a curva para os alvos moleculares selecionados de forma a avaliar a capacidade da metodologia utilizada em diferenciar falsos-positivos de moléculas ativas.

C) **Triagem Virtual baseada na estrutura do alvo molecular:** Utilização do MolAr para a triagem virtual baseada na estrutura do alvo molecular. Nessa etapa, os ligantes selecionados na etapa anterior são ancorados contra os alvos moleculares selecionados utilizando a Triagem virtual de consensus, realizada pelo programa MolAr.

D) **Confecção do artigo:** Confecção do artigo que detalha o que foi feito durante essa fase do projeto

E) **Manutenção do site:** www.drugdiscovery.com.br e divulgação dos resultados parciais do projeto

Cronograma de Atividades:

Atividade	Mês					
	1	2	3	4	5	6
A	X					
B	X	X				
C	X	X	X			
D			X	X	X	X
E			X	X	X	X

3) Bolsista 3 e Bolsista 4 (o bolsista 4 é voluntário)

A) **Leitura do material de referência:** Pesquisa bibliográfica sobre a triagem virtual em coronavirus e a utilização de Triagem Virtual baseada no ligante nesse processo.

B) Estudo de ferramentas *Open Source* de Triagem Virtual Baseada no Ligante: Serão estudadas as ferramentas gratuitas de Triagem Virtual Baseada no ligante Pharmit e DataWarrior, para uma definição de como integrá-las ao MolAr.

C) Integração do software: Integração do software selecionado no passo anterior ao MolAr.

D) Manutenção do site: www.drugdiscovery.com.br e divulgação dos resultados parciais do projeto.

Cronograma de Atividades:

Atividade	Mês					
	1	2	3	4	5	6
A	X					
B		X				
C			X	X	X	X
D			X	X	X	X

10. Parceiros

ALEX Gutterres Taranto (Professor da UFSJ)

11. Referências

ANTONIO, Gregory E. *et al.* Imaging of severe acute respiratory syndrome in Hong Kong. *American Journal of Roentgenology*, v. 181, n. 1, p. 11–17, 2003.

BERMAN, Helen M. *et al.* The protein data bank. *Acta Crystallographica Section D: Biological Crystallography*, v. 58, n. 6 I, p. 899–907, 2002.

BOGOCH, Isaac I *et al.* Pneumonia of Unknown Etiology in Wuhan, China: Potential for International Spread Via Commercial Air Travel. *Journal of Travel Medicine*, p. 1–3, 2020.

BORREL, Alexandre *et al.* PockDrug: A Model for Predicting Pocket Druggability That Overcomes Pocket Estimation Uncertainties. *Journal of chemical information and modeling*, v. 55, n. 4, p. 882–95, 27 abr. 2015. Disponível em:
<<http://www.ncbi.nlm.nih.gov/pubmed/25835082>>. Acesso em: 9 abr. 2020.

BROZELL, Scott R. *et al.* Evaluation of DOCK 6 as a pose generation and database enrichment tool. *Journal of Computer-Aided Molecular Design*. [S.l.]: J Comput Aided Mol Des. , jun. 2012

CARREGAL, Ana Paula *et al.* Docking-based virtual screening of Brazilian natural compounds using the OOMT as the pharmacological target database. *Journal of Molecular Modeling*, v. 23, n. 4, 1 abr. 2017.

CHAN, Jasper Fuk Woo *et al.* Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. *Emerging Microbes and Infections*, v. 9, n. 1, p. 221–236, 2020.

CHEN, Yu; LIU, Qianyun; GUO, Deyin. Emerging coronaviruses: Genome structure, replication, and pathogenesis. *Journal of Medical Virology*, v. 92, n. 4, p. 418–423, 2020.

DUFFY, Siobain; SHACKELTON, Laura A.; HOLMES, Edward C. Rates of evolutionary change in viruses: Patterns and determinants. *Nature Reviews Genetics*, v. 9, n. 4, p. 267–276, 2008.

FAWCETT, Tom. An introduction to ROC analysis. *Pattern Recognition Letters*, v. 27, n. 8, p. 861–874, 1 jun. 2006.

GLAAB, Enrico. Building a virtual ligand screening pipeline using free software: A survey. *Briefings in Bioinformatics*, v. 17, n. 2, p. 352–366, 2016.

HEVENER, Kirk E. *et al.* Validation of molecular docking programs for virtual screening against dihydropteroate synthase. *Journal of Chemical Information and Modeling*, v. 49, n. 2, p. 444–460, 23 fev. 2009.

HOFFMANN, Markus *et al.* SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell*, v. 181, p. 1–10, 2020.

HOUSTON, Douglas R.; WALKINSHAW, Malcolm D. Consensus docking: Improving the reliability of docking in a virtual screening context. *Journal of Chemical Information and Modeling*, v. 53, n. 2, p. 384–390, 25 fev. 2013.

KUKOL, Andreas. Consensus virtual screening approaches to predict protein ligands. *European Journal of Medicinal Chemistry*, v. 46, n. 9, p. 4661–4664, 1 set. 2011.

LAI, M. M.C. RNA recombination in animal and plant viruses. *Microbiological Reviews*, v. 56, n. 1, p. 61–79, 1992.

LAURO, Gianluigi *et al.* Inverse Virtual Screening allows the discovery of the biological activity of natural compounds. *Bioorganic and Medicinal Chemistry*, v. 20, n. 11, p. 3596–3602, 1 jun. 2012.

LU, Hongzhou; STRATTON, Charles W.; TANG, Yi Wei. Outbreak of pneumonia of unknown etiology in Wuhan, China: The mystery and the miracle. *Journal of Medical Virology*, v. 92, n. 4, p. 401–402, 2020.

LU, Roujian *et al.* Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *The Lancet*, v. 395, n. 10224, p. 565–574, 2020. Disponível em: <[http://dx.doi.org/10.1016/S0140-6736\(20\)30251-8](http://dx.doi.org/10.1016/S0140-6736(20)30251-8)>.

MAIA, Eduardo H. B. *et al.* Molecular Architect: A User-Friendly Workflow for Virtual Screening. *ACS Omega*, v. 5, n. 12, p. 6628–6640, 31 mar. 2020. Disponível em: <<https://pubs.acs.org/doi/10.1021/acsomega.9b04403>>.

MAIA, Eduardo H. B. *et al.* Octopus: a platform for the virtual high-throughput screening of a pool of compounds against a set of molecular targets. *Journal of Molecular Modeling*, v. 23, n. 1, p. 1–11, 1 jan. 2017.

MYSINGER, Michael M. *et al.* Directory of useful decoys, enhanced (DUD-E): Better ligands and decoys for better benchmarking. *Journal of Medicinal Chemistry*, v. 55, n. 14, p. 6582–6594, 26 jul. 2012.

NAEEM, Sadaf; HYLANDS, Peter; BARLOW, David. Construction of an Indonesian herbal constituents database and its use in Random Forest modelling in a search for inhibitors of aldose reductase. *Bioorganic and Medicinal Chemistry*, v. 20, n. 3, p. 1251–1258, 1 fev. 2012.

NUNES, Renata Rachide *et al.* Brazilian malaria molecular targets (BraMMT): Selected receptors for virtual high-throughput screening experiments. *Memorias do Instituto Oswaldo Cruz*, v. 114, n. 2, 1 fev. 2019.

O'BOYLE, Noel M. *et al.* Open Babel: An Open chemical toolbox. *Journal of Cheminformatics*, v. 3, n. 10, p. 33, 7 out. 2011. Disponível em: <<https://jcheminf.biomedcentral.com/articles/10.1186/1758-2946-3-33>>. Acesso em: 9 abr. 2020.

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

POLGAR, Timea; M. KESERU, Gyorgy. Integration of Virtual and High Throughput Screening in Lead Discovery Settings. *Combinatorial Chemistry & High Throughput Screening*, v. 14, n. 10, p. 889–897, 2011.

SANDER, Thomas *et al.* DataWarrior: An open-source program for chemistry aware data visualization and analysis. *Journal of Chemical Information and Modeling*, v. 55, n. 2, p. 460–473, 23 fev. 2015.

SANTOS, Felipe R.S.; LIMA, William G.; *et al.* Identification of a Potential Zika Virus Inhibitor Targeting NS5 Methyltransferase Using Virtual Screening and Molecular Dynamics Simulations. *Journal of Chemical Information and Modeling*, v. 60, n. 2, p. 562–568, 2020.

SANTOS, Felipe R.S.; NUNES, Damiana A.F.; *et al.* Identification of Zika Virus NS2B-NS3 Protease Inhibitors by Structure-Based Virtual Screening and Drug Repurposing Approaches. *Journal of Chemical Information and Modeling*, v. 60, n. 2, p. 731–737, 2020.

STERLING, Teague; IRWIN, John J. ZINC 15 - Ligand Discovery for Everyone. *Journal of Chemical Information and Modeling*, v. 55, n. 11, p. 2324–2337, 23 nov. 2015.

STEWART, James J.P. Optimization of parameters for semiempirical methods VI: More modifications to the NDDO approximations and re-optimization of parameters. *Journal of Molecular Modeling*, v. 19, n. 1, p. 1–32, 28 jan. 2013.

SUNSERI, Jocelyn; KOES, David Ryan. Pharmit: interactive exploration of chemical space. *Nucleic acids research*, v. 44, n. W1, p. W442–W448, 2016. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4987880/>>. Acesso em: 9 abr. 2020.

TROTT, Oleg; OLSON, Arthur J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, v. 31, n. 2, p. NA-NA, 2009.

WALTERS, W. Patrick; MURCKO, Ajay; MURCKO, Mark A. Recognizing molecules with drug-like properties. *Current Opinion in Chemical Biology*, v. 3, n. 4, p. 384–387, 1 ago. 1999.

WEBB, Benjamin; SALI, Andrej. Comparative protein structure modeling using MODELLER. *Current Protocols in Bioinformatics*, v. 2016, p. 5.6.1-5.6.37, 2016.

WESTERMAIER, Yvonne; BARRIL, Xavier; SCAPOZZA, Leonardo. Virtual screening: An in silico tool for interlacing the chemical universe with the proteome. *Methods*, v. 71, n. C, p. 44–57, 1 jan. 2015.

WOO, Patrick C.Y. *et al.* Coronavirus diversity, phylogeny and interspecies jumping. *Experimental Biology and Medicine*, v. 234, n. 10, p. 1117–1127, 2009.

WU, Aiping *et al.* Genome Composition and Divergence of the Novel Coronavirus (2019-nCoV) Originating in China. *Cell Host and Microbe*, v. 27, n. 3, p. 325–328, 2020. Disponível em: <<https://doi.org/10.1016/j.chom.2020.02.001>>.

YANG, Penghui; WANG, Xiliang. COVID-19: a new challenge for human beings. *Cellular & molecular immunology*, n. March, p. 19–21, 2020. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/32235915>>.

ZAKI, Ali Moh *et al.* Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *New England Journal of Medicine*, v. 367, n. 19, p. 1814–1820, 2012.

ZUMLA, Alimuddin *et al.* Coronaviruses-drug discovery and therapeutic options. *Nature Reviews Drug Discovery*, v. 15, n. 5, p. 327–347, 2016.