



**UNIVERSIDADE DA INTEGRAÇÃO INTERNACIONAL DA LUSOFONIA
AFRO-BRASILEIRA**

**PRO-REITORIA DE GRADUAÇÃO (PROGRAD)
INSTITUTO DE CIÊNCIAS EXATAS E DA NATUREZA - ICEN
CURSO DE LICENCIATURA PLENA EM QUÍMICA**

SADRACK QUEQUE CABONGO

**PROJETO DE ESTRUTURAS BIOATIVAS *DE NOVO* PARA INIBIDORES DA
ENZIMA *CANDIDA AURIS* POLIMERASE POR MEIO DE INTELIGÊNCIA
ARTIFICIAL**

REDENÇÃO - CE

2023

SADRACK QUEQUE CABONGO

**PROJETO DE ESTRUTURAS BIOATIVAS *DE NOVO* PARA INIBIDORES DA
ENZIMA *CANDIDA AURIS* POLIMERASE POR MEIO DE INTELIGÊNCIA
ARTIFICIAL**

Trabalho de conclusão do curso apresentado ao curso de Licenciatura em Química do Instituto de Ciências Exatas e da Natureza - ICEN da Universidade da Integração Internacional da Lusofonia Afro-Brasileira como requisito parcial para obtenção do título de graduado em licenciatura em Química.

Orientador: Prof. Dr. Alúcio Marques da Fonseca

REDENÇÃO - CE

2023

Cabongo, Sadrack Queque.

C116p

Projeto de estruturas bioativas de novo para inibidores da enzima Candida Auris Polimerase por meio de inteligência artificial / Sadrack Queque Cabongo. - Redenção, 2023.
49fL: il.

Monografia - Curso de Química, Instituto de Ciências Exatas e da Natureza, Universidade da Integração Internacional da Lusofonia Afro-Brasileira, Redenção, 2023.

Orientador: Prof. Dr. Aluísio Marques da Fonseca.

1. Fungo. 2. Dinâmica molecular. 3. Aprendizagem. 4. Estudo ADMET. 5. MM/GBSA. I. Fonseca, Prof. Dr. Aluísio Marques da. II. Título.

CE/UF/BSCA

CDD 664.001

SADRACK QUEQUE CABONGO

**PROJETO DE ESTRUTURAS BIOATIVAS DE NOVO PARA INIBIDORES DA
ENZIMA CANDIDA AURIS POLIMERASE POR MEIO DE INTELIGÊNCIA
ARTIFICIAL**

Trabalho de Conclusão de Curso apresentado à Coordenação do Curso de Licenciatura em Química, do Instituto de Ciências Exatas e da Natureza (ICEN) da Universidade da Integração Internacional da Lusofonia Afro-Brasileira (UNILAB), como requisito para a obtenção do Título de Licenciado em Química.

Aprovado em: 23/01/2023

Banca Examinadora:



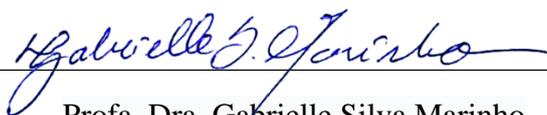
Prof. Dr. Aluísio Marques da Fonseca (Orientador)

Universidade da Integração Internacional da Lusofonia Afro-Brasileira - UNILAB



Prof. Dr. Emmanuel Silva Marinho

Universidade Estadual do Ceará - UECE



Profa. Dra. Gabrielle Silva Marinho

Universidade Estadual do Ceará - UECE

REDENÇÃO - CE

2023

Dedico a minha família, pela nossa história.
Aos meus pais pela dedicação no cuidado
dos filhos, pois investiram fortemente na
nossa formação acadêmica e profissional.
Hoje os frutos do seu trabalho podem ser
colhidos!

AGRADECIMENTOS

Agradeço a Deus Todo-poderoso pela vida e sustento, até aqui Ele tem me guiado, e de forma graciosa tem me dado força e vigor para suportar as dificuldades e superar os obstáculos que tenho encontrado na minha jornada. A Ele seja a glória para sempre.

Agradeço a minha família pelo apoio que tenho recebido desde o momento que dei início nessa empreitada, em especial aos meus pais, Cabongo Pinto e Rosa Fernanda pelo muito que tiveram que trabalhar para sustentar a nossa família, pelos ensinamentos que moldaram a minha vida e por me mostrar o caminho do SENHOR desde pequeno. Sou muito grato pela motivação, encorajamento e muitas alegrias que recebo dos meus irmãos e irmãs, espero estar com vocês logo!

Sou grato a Deus por cada irmão e irmã que me deu no decorrer da minha graduação, em especial aos meus queridos irmãos da Igreja Batista do Calvário de Acarape, pois através de cada um deles tenho aprendido as virtudes da vida.

Sou grato aos meus amigos e colegas do curso de Química (2018.1), pois cada um marcou a minha caminhada na graduação de forma diferente e especial, obrigado por cada momento vivenciado e pelos aprendizados, espero poder desfrutar da nossa amizade por longos anos, especialmente ao “Quarteto”.

Sou muito grato ao coletivo dos professores do curso de Química, pelas vivências, momentos que ficarão marcados na minha memória, pois todos contribuíram significativamente para a minha formação, muito obrigado pelos ensinamentos que levarei para vida, tenho grande admiração e prazer de tê-los como meus mestres e mestra. Ao meu professor e orientador Dr. Alúcio Marques da Fonseca, muito obrigado por me instruir nesses últimos semestres onde tive mais contato com a pesquisa, tens minha admiração, pois tive o privilégio de aprender com um grande professor e pesquisador, muito obrigado!

Sou grato aos meus amigos e colegas de pesquisa do Grupo Interdisciplinar em Química (GIQ) e do Programa Residência Pedagógica (PRP), tem sido um privilégio para mim trabalhar e seguir crescendo com vocês.

“Os únicos que não têm cicatrizes, são aqueles que decidiram não combater.”

C.S. Lewis

RESUMO

Desde a sua descoberta no Japão, a espécie fúngica *Candida auris* tem atraído a atenção dos pesquisadores devido à sua excelente capacidade de infecção e resistência inata a diferentes drogas amplamente utilizadas em casos de infecção por fungos, mostrando claramente a necessidade de desenvolver novas drogas que possam atuar de forma mais eficiente nos tratamentos para esta infecção. Este trabalho apresenta a concepção de novos compostos bioativos que podem ser tomados como possíveis candidatos à inibição da proteína alvo do superfungo. Através do método De novo drug design (DNDD) auxiliado por inteligência artificial (IA), foi possível criar novas estruturas usando um modelo de ligante dentro de um modelo de receptor. Das 1.722 (Hits) novas estruturas derivadas do composto inicial (S)-piperidin-ilmetanol, o composto com maior valor de bioafinidade foi o Hit-960. Considerando as análises tanto da energia de ligação quanto de sua acessibilidade sintética, mostrou-se o inibidor mais potente nos estudos *in silico* com energia de $-9,12 \cdot \text{kcal/mol}$ e viabilidade sintética próxima a 75%, além de apresentar estabilidade na análise dinâmica nos estudos RMSD, RMSF e SASA, em comparação com o fármaco de referência FK-506, com afinidade de ligação de $-7,3 \text{ kcal/mol}$. As análises do ADMET sugerem que o composto ocupa um espaço físico-químico com alinhamento entre seus atributos de permeabilidade passiva, depuração e segurança. No entanto, sugere-se que o composto Hit-960 seja utilizado em ensaios clínicos para testar sua eficácia, contribuindo para o desenvolvimento e produção de novos medicamentos que cheguem com mais eficiência aos tratamentos por infecção por *C. auris*.

Palavras-chave: superfungo; dinâmica molecular; aprendizagem profunda; estudo ADMET; MM/GBSA.

ABSTRACT

Since its discovery in Japan, the fungal species *Candida auris* has attracted the attention of researchers due to its excellent infection capacity and its ability to innately resist the different drugs widely used in cases of yeast infection, clearly showing the need to develop new drugs that can act more efficiently in the treatments for this infection. This paper presents the conception of new bioactive compounds that can be taken as possible candidates for inhibiting the protein target of the superfungus. Through the De novo drug design (DNDD) method aided by artificial intelligence (AI), it was possible to create new structures using a ligand model within a receiver model. Of the 1,722 (Hits) new forms derived from the initial compound (*S*)-piperidin-ylmethanol, the compound with the highest bioaffinity value was Hit-960. Considering the analyses of both the binding energy and its synthetic accessibility, it proved to be the most potent inhibitor in the *in silico* studies with an energy of -9.12·kcal/mol and synthetic viability close to 75%, besides presenting stability in the molecular dynamics analysis in the RMSD, RMSF and SASA studies, compared to the reference drug FK-506, with a binding affinity of -7.3 kcal/mol. ADMET analyzes suggest that the compound occupies a physicochemical space with alignment between its attributes of passive permeability, clearance and safety. However, it is suggested that the Hit-960 compound be used in clinical trials to test its efficacy, contributing to developing and producing new drugs that land more efficiently in treatments for *C. auris* infection.

Keywords: superfungo; molecular dynamics; deep learning; ADMET study; MM/GBSA.

LISTA DE ILUSTRAÇÕES

LISTA DE FIGURAS

Figura 1.	Sobreposição de duas proteínas (azul e magenta) e a cavidade co-cristalizada do ligante FK-506 (laranja).	19
Figura 2.	Estrutura de crescimento inicial (S)-piperidin-2-imetanol	24
Figura 3.	Número de clusters versus energia de afinidade para MolAICaID	24
Figura 4.	Acessibilidade sintética versus energia de afinidade sugerida por MolAiCAID	26
Figura 5.	Energia de afinidade face aos clusters produzidos na simulação	28
Figura 6.	(a) 3D e 2D das interações entre a estrutura mais promissora (Hit 960) e (b) o ligante de referência e os resíduos da polimerase	29
Figura 7.	Perfil do valor de RMSD obtido para MD com o ligante de referência FK-506 (linha preta) e MD com Hit 960 (linha vermelha)	31
Figura 8.	Perfil do valor de RMSF obtido para polimerase em MD com Hit 960 (linha vermelha), MD com FK-506 (linha preta)	32
Figura 9.	Número de ligações de hidrogênio (H-Bond) encontrado entre <i>Candida auris</i> polimerase e MD com FK-506 (linha preta) e MD com Hit 960 (linha vermelha) com software VMD, considerando o comprimento máximo 3,3 Å	33
Figura 10.	Área de superfície acessível ao solvente (SASA) da polimerase <i>Candida auris</i> em função do tempo das simulações MD: MD com Hit 960, estabilizando íons monovalentes (linha preta) e MD com FK-506, estabilizando íons monovalentes (linha vermelha). As curvas são médias dos dados brutos com uma janela de 100 ns	34
Figura 11.	Mapa de potencial lipofílico molecular (MLP) e valor de pKa estimado de Hit 960 (a), gráfico de distribuição de porções químicas em função do pH (b), radar espacial de drogas (c), alinhamento entre TPSA e logP para estimativa de segurança (d) e predição do local de metabolismo e descritores farmacocinéticos (e)	38

LISTA DE TABELAS

Tabela 1:	Dados de estimativa de energia livre de Hit 960 e FK-506 contra polimerase de <i>Candida auris</i>	30
Tabela 2:	Resíduos da polimerase <i>Candida auris</i> que mostraram ligações de H ao longo das MDs	33
Tabela 3:	Propriedades físico-químicas do Hit 960 aplicadas aos critérios de similaridade com drogas Pfizer e MCE-18	35
Tabela 4:	Propriedades farmacocinéticas previstas por modelos ADMET dos servidores ADMETlab 2.0 e ADMETboost – AI Drug Lab	37

LISTA SIGLAS, SÍMBOLOS E ABREVIATURAS

1D	Uma dimensão
2D	Duas dimensões
3D	Três dimensões
5HUA	Código de identificação da proteína polimerase de <i>Candida glabrata</i>
6VSI	Código de identificação da proteína polimerase de <i>Candida auris</i>
ΔE	Variação de energia
ΔE_{ele}	Energia eletrostática
$\Delta E_{interna}$	Energia de ligação
ΔE_{MM}	Energia MM da fase gasosa
ΔE_{vdw}	Energia de Van Der Waals
ΔG	Variação da energia livre de Gibbs
ΔG_{sol}	Energia livre de solvatação
ΔG_{GB}	Contribuição polar da energia livre de solvatação
ΔG_{SA}	Contribuição apolar da energia livre de solvatação
ΔH	Variação de Entalpia
Å	Angstrom (Unidade de medida de comprimento equivalente a 10^{-10} m)
AR	Anéis aromáticos
Arg	Arginina
Asn	Asparagina
Asp	Aspartato
BBB	Permeabilidade da barreira hematoencefálica
$CL_{int,u}$	Intrinsic clearance of the unbound fraction
CYP450	Substratos do citocromo P450
Cys	Cisteína
DILI	Drug-induced liver damage
DL	Aprendizado profundo (Deep Learning)
DNDD	<i>De novo</i> design drug
F%	Fração de biodisponibilidade oral
FDA	Food and Drug Administration
FK-506	Ligante nativo da proteína polimerase de <i>Candida glabrata</i>
Gln	Glutamina

Glu	Ácido Glutâmico
Gly	Glicina
H-Bond	Ligação de hidrogênio
HIA	Absorção intestinal humana
His	Histidina
HLM	Microsomo hepático humano
HTS	High Throughput Screening
IA	Inteligência Artificial
Ile	Isoleucina
LBDD	Design de drogas baseado no ligante
LC₅₀	Concentração letal
Leu	Leucina
MD	Dinâmica Molecular
MCE18	Medicinal Chemistry Evolution 2018
Met	Metionina
ML	Aprendizado de máquina (Machine Learning)
MLP	Potencial lipofílico molecular
MM/GBSA	Mecânica Molecular combinada com Solvatação Contínua Generalizada de Born e Área de Superfície
NAR	Anéis não aromáticos
ns	Nano segundo
P_{app}	Permeabilidade passiva (Passive permeability)
PDB	Protein Data Bank
PDBQT	Protein Data Bank, Partial Charge (Q) & Atom Type (T)
Pgp	Glicoproteína P
Phe	Fenilalanina
Pro	Prolina
PPB	Ligação às proteínas plasmáticas
RMSD	Desvio Médio Quadrático de Rotação (Root Mean Square Deviation)
RMSF	Flutuação Média Quadrática de Rotação (Root Mean Square Fluctuation)
SA	Acessibilidade sintética
SASA	Área de Superfície Acessível por Solvente
Ser	Serina

SOM	Local do metabolismo (Site of metabolism)
T	Temperatura
$T\Delta S$	Entropia conformacional
TF	Pesca de alvo (Target Fischer)
Trp	Triptófano
Tyr	Tirosina
Val	Valina
VMD	Visual Molecular Dynamics
VS	Triagem Virtual (Virtual Screening)

SUMÁRIO

1.	INTRODUÇÃO	16
2.	METODOLOGIA	19
2.1.	Preparação Estrutural de Proteína	19
2.2.	Modelo de aprendizado profundo e design de drogas <i>De novo</i>	19
2.3.	Coordenadas da grade	20
2.4.	Pontos de Fibonacci	20
2.5.	Filtro Lipinski	20
2.6.	Interferência	20
2.7.	Acessibilidade de síntese	20
2.8.	Processamento	21
2.9.	Docking molecular	21
2.10.	Dinâmica molecular	21
2.11.	Cálculos MM/GBSA	22
2.12.	Método <i>in silico</i> ADMET	22
3.	RESULTADOS E DISCUSSÕES	24
3.1.	Processamento	24
3.2.	Produção de hits	24
3.3.	Disponibilidade sintética	25
3.4.	Energia de afinidade	26
3.5.	Interação com resíduos de proteínas	29
3.6.	Cálculos MM/GBSA	29
3.7.	Dinâmica molecular	31
3.7.1.	Análise RMSD	31
3.7.2.	Análise RMSF	31
3.7.3.	Ligações de H	32
3.7.4.	Área de Superfície Acessível ao Solvente	33
3.8.	Estudo <i>in silico</i> ADMET	34
3.8.1	Estimativa quantitativa da semelhança com drogas	34
3.8.2	Descritores farmacocinéticos previstos	36
3.8.3	Local de metabolismo e previsão de toxicidade	37
4.	CONCLUSÃO	39
5.	REFERÊNCIAS	40
	APÊNDICE	44

1. INTRODUÇÃO

Candida auris é uma espécie fúngica descoberta em 2009 no Japão, tendo sido isolada pela primeira vez no ouvido de um paciente e posteriormente identificada com algumas dificuldades em três hospitais universitários da Coreia do Sul (BASHIR et al., 2020; LEE et al., 2011). Desde sua descoberta, *C. auris* foi identificado em vários países do mundo. Nota-se que a espécie possui excelente capacidade de infecção, principalmente em hospitais. Por ser parente de *C. haemulonii*, é necessário identificar a espécie com mais precisão para que haja tratamento antifúngico adequado nesses ambientes (SCHELENZ et al., 2016). Com base em estudos recentes, se observou uma resistência inata a antifúngicos amplamente utilizados em casos de infecções fúngicas (RHODES; FISHER, 2019). De acordo com o Relatório de Ameaças de Resistência a Antibióticos publicado pelos Centros de Controle e Prevenção de Doenças (CDCP), variantes de *C. auris* tornaram-se os primeiros patógenos fúngicos incluídos entre as cinco ameaças mais urgentes, resistentes a todas as três principais classes de drogas antifúngicas (BASHIR et al., 2020). Nesse contexto, é necessário estudar e desenvolver novos medicamentos que possam interagir de forma mais eficaz com a principal cadeia proteica da espécie *C. auris* para inibi-la, evitando assim sua proliferação e a crescente contaminação que tem chamado a atenção da comunidade científica globalmente.

A inteligência artificial (IA) tem se tornado cada vez mais uma aliada na descoberta de novos fármacos, no desenvolvimento de novos medicamentos e na rápida obtenção de resultados satisfatórios na área de modelagem computacional. Avanços recentes relacionados à inteligência artificial têm contribuído para o desenvolvimento de novas técnicas que facilitam a concepção e o desenvolvimento de novos medicamentos. Assim, permitindo a minimização dos custos de desenvolvimento de medicamentos na pesquisa e concepção e novos agentes terapêuticos (GUPTA et al., 2021). A abordagem computacional tradicional para a descoberta e desenvolvimento de drogas consiste no processo de triagem virtual (VS) e acoplamento molecular. No entanto, esses processos tendem a impor outros obstáculos, como ineficiência e imprecisão (GUPTA et al., 2021). A inteligência artificial (IA) mostra a capacidade de uma máquina de imitar funções cognitivas humanas, mas com a vantagem de gerar novas informações, adquirir maior grau de precisão, simulações e previsões automatizadas e desempenho contínuo (HASSANZADEH; ATYABI; DINARVAND, 2019). A inteligência artificial (IA), incluindo algoritmos de aprendizado profundo (DL) e aprendizado de máquina (ML), surgiu como uma solução viável para superar problemas e obstáculos no processo de

concepção e descoberta de medicamentos (DUCH; SWAMINATHAN; MELLER, 2007; GUPTA et al., 2021). O uso da abordagem de aprendizado de máquina (ML) para prever alvos com base em ligantes possibilitou o surgimento da técnica de pesca de alvo (TF), por meio da qual através do conjunto de informações de um ligante e uma proteína alvo, considerando sua relação pode ser usado para prever novos alvos proteicos de novos compostos com atividades biológicas (AGATONOVIC-KUSTRIN; BERESFORD, 2000; HASSANZADEH; ATYABI; DINARVAND, 2019).

De novo design drug (DNDD) é uma abordagem computacional que visa a projeção de novos compostos químicos que se encaixem em um conjunto de restrições usando o algoritmo de crescimento (MOUCLIS et al., 2021). O método *De novo*, aplicado no design de medicamentos, cria novos compostos químicos com base em informações relacionadas a um alvo biológico (receptor) e seus ligantes ativos (ligantes que possuem boa ligação e têm a capacidade de inibir o receptor) para projetar compostos que são interessantes para a pesquisa e desenvolvimento de novos fármacos levando em consideração o alvo específico, a acessibilidade sintética do composto e as propriedades farmacocinéticas requeridas (NICOLAOU; BROWN, 2013).

O aprendizado profundo (DL) é o ramo da inteligência artificial que se tornou muito útil no design de ligantes 1D e 2D, entretanto, a geração de ligantes tridimensionais 3D no bolso 3D de um alvo (proteína) é mais desafiadora para o design de medicamentos e DL. Com base nas características e méritos do DL e da programação clássica, o software MolAICal é programado para o design de medicamentos 3D no bolso da proteína. Pode projetar ligantes 3D favoráveis no bolso da proteína através do método de acoplamento e *De novo* (BAI et al., 2021).

MolAICal tem dois módulos principais que são escritos pelo programa JAVA. O primeiro é baseado no algoritmo genérico e DL treinado em fragmentos de drogas aprovados pela Food and Drug Administration (FDA), e o segundo é baseado no acoplamento molecular e modelo DL treinado em ligantes semelhantes aos do banco de dados ZINC (BAI et al., 2021). MolAICal pode gerar automaticamente fragmentos válidos e semelhantes aos da FDA; ele também tem uma pontuação de avaliação de ligação de ligante de energia melhor em bolsões 3D de alvos de proteína do que o AutoDock Vina (Wang et al., 2016). MolAICalD é um código derivado do Autodock Vina (TROTT; OLSON, 2010) e está no pacote computacional do MolAICal (BAI et al., 2021). Este código usa a mesma licença do Autodock Vina, APACHE 2.0(FOUNDATION, 2012; VIEIRA; SOUSA, 2019), e pode realizar docking molecular,

considerado um software de acoplamento molecular de código aberto estudado e aplicado no campo da descoberta de drogas. A validação é realizada pelos coeficientes de correlação de Pearson e Spearman, que estão relacionados entre a afinidade de ligação experimental e a afinidade de ligação preditiva, ou seja, quando o ligante com o menor RMSD em cada complexo extraído tiver valores menores que 2,0 Å (YUSUF et al., 2008). A nível de comparação, Autodock Vina tem uma taxa de precisão de acoplamento equivalente a 54,665%, e seus coeficientes de correlação de Pearson e Spearman são 0,5259 e 0,5421, respectivamente, com base na afinidade de ligação experimental. Por outro lado, MolAICaD em MolAICal, a taxa de precisão de acoplamento é de 54,888%, e seus coeficientes de correlação de Pearson e Spearman são 0,5335 e 0,5489, indicando que MolAICal tem um "encaixe" e "poder de classificação" que pode ser comparado ao Autodock Vina e pode ser validado em simulações de docking molecular.

Dentre as aplicações, é fundamental destacar que diversos dockings realizados pelo MolAICaD, entre outros softwares, estão sendo utilizados na pesquisa de inibidores de doenças negligenciadas como Dengue, doença de Chagas, leishmaniose, entre outras, na busca de novos candidatos a fármacos com base no conhecimento da química de produtos naturais, no entendimento da estrutura do protótipo (DIAS et al., 2013).

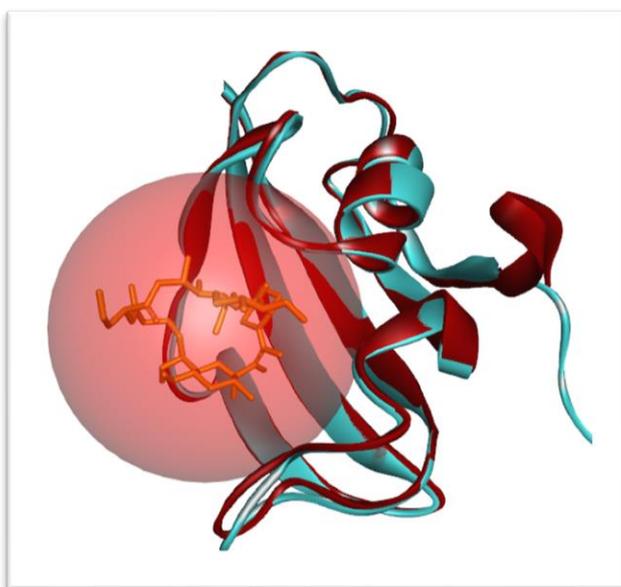
Este trabalho visa a busca de novos inibidores de *Candida auris*, através do método DNDD, com auxílio de inteligência artificial MolAICal e ensaio adicional de dinâmica molecular para verificar a estabilidade dos complexos formados entre o ligante e o alvo molecular.

2. METODOLOGIA

2.1. Preparação Estrutural de Proteína

O banco de dados PDB forneceu a estrutura cristalina da proteína polimerase (PDB ID, 6VSI) (<https://www.rcsb.org/>) (BASHIR et al., 2020). O ligante co-cristalizado da polimerase de *C. glabrata* (5HUA) definiu a cavidade para o planejamento de drogas por sobreposição de proteínas (BIOVIA, 2016). A liga FK-506 co-cristalizada foi excluída do novo desenho de droga e do cristal de polimerase de aprendizado profundo (DL), Figura.1. O arquivo de rede de polimerase *Candida auris* foi produzido para o crescimento de fragmentos pelo pacote MolaICal (<https://molaical.github.io>) (BAI et al., 2021). O fragmento de crescimento inicial da polimerase era um esqueleto básico da estrutura do ligante nativo FK-506. A estrutura da polimerase *Candida auris* foi preparada para triagem virtual usando Autodock Tools (MORRIS et al., 2009). As cargas de hidrogênio gasteiger e polar serão adicionadas à proteína, salvas no formato molecular pdbqt.

Figura 1. Sobreposição de duas proteínas (azul e magenta) e a cavidade co-cristalizada do ligante FK-506 (laranja).



Fonte: Autor (2023)

2.2. Modelo de aprendizado profundo e design de drogas *De novo*

O MolaICal conterá o modelo gerador de aprendizagem profunda do medicamento, que será treinado a partir de 21.064 fragmentos de medicamentos aprovados pela FDA. Os 90 fragmentos gerados por MolaICal e outros 30 fragmentos

primários serão misturados para crescimento de fragmentos na cavidade da enzima polimerase de *Candida auris*.

2.3. Coordenadas da grade

As coordenadas x, y e z do centro da caixa de cavidade da proteína foram definidas como $x = -28,152$, $y = 24,208$, $z = -1,140$ com base em grade. Os comprimentos da caixa de cavidade da polimerase foram ajustados para $30,0 \text{ \AA}$ ao longo das direções x, y e z. Moléculas elitistas foram extraídas para posterior crescimento evoluído de 10% das populações moleculares geradas. As 140 melhores moléculas de populações moleculares geradas serão desenvolvidas como moléculas-mãe. Assim, mais de 60 moléculas serão selecionadas aleatoriamente das populações moleculares geradas para aumentar a diversidade e a novidade dos ligantes de crescimento. A população máxima será fixada em 3000.

2.4. Pontos de Fibonacci

Os 361 pontos de Fibonacci são gerados para a busca de distúrbios de fragmentos. Os operadores de cruzamento e mutação serão definidos como 1,0 e 0,5, respectivamente.

2.5. Filtro Lipinski

De acordo com a regra dos cinco de Lipinski, os valores dos ligantes cristalinos na proteína foram definidos para os valores de XLOGP (5,0), aceptores de hidrogênio (10), doadores de hidrogênio (5), peso molecular (500) e ligações rotativas (10)

2.6. Interferência

Os compostos de interferência de pan-ensaio (PAINS) irão filtrar os ligantes de crescimento indesejados.

2.7. Acessibilidade de síntese

A pontuação de acessibilidade sintética dos ligantes de crescimento será salva no arquivo de resultados estatísticos no final da simulação. Um total de 30 gerações de ciclos serão realizadas para todo o processo de design de drogas. Um total de seis processos paralelos de design de drogas serão realizados na proteína. A polimerase dos ligantes gerados foi salva entre 480 e 785 do peso molecular.

2.8. Processamento

Um total de 30 multicores de CPU serão executados em paralelo para todo o processo de crescimento molecular. A droga como todo o processo de design combinado com modelo de aprendizado profundo e programação clássica será realizado automaticamente pelo pacote projetado do MolAICal.

2.9. Docking molecular

O Algoritmo Genético Lamarckiano foi escolhido para realizar as simulações de docking molecular implementadas no MolAiCAID. A caixa da grade foi centralizada para abranger todas as cadeias de proteínas para determinar o espaço de simulação. O RMDS de até 2,0 Å (HEVENER et al., 2009), e energia de afinidade maior que -7,0 kcal/mol foi usada como fator de exclusão. Os mais favoráveis foram representados pela menor energia de ligação livre (ΔG)(GURUNG et al., 2016). Discovery Studio (BIOVIA, 2015) realizou estudos de análise de visualização 3D/2D de interação, e o Poseview foi adicionado (FRICKER; GASTREICH; RAREY, 2004; STIERAND; MAASS; RAREY, 2006).

2.10. Dinâmica molecular

As simulações de dinâmica molecular (MD) foram realizadas com o programa NAMD (PHILLIPS et al., 2005). As melhores conformações obtidas no docking molecular foram no caso solvatado em água no modelo TIP3P (KATO et al., 2021), no campo de força CHARMM36-mar2019 (HUANG et al., 2016). A preparação do sistema foi realizada em duas etapas. Na primeira etapa, os ligantes foram parametrizados no servidor Charmm-Gui (JO et al., 2013) (<https://www.charmm-gui.org/>), e então eles foram enviados ao servidor CGenFF para identificação de parâmetros para CHARMM36 (VANOMMESLAEGHE et al., 2010). Na segunda etapa, a proteína foi preparada no programa NAMD. 1 íon Na^+ por ligante foi adicionado para neutralizar a carga total do sistema. Este último foi submetido à minimização de energia pelo método Steepest Descent. Em seguida, o sistema foi submetido aos equilíbrios NVT e NPT nas condições descritas por Langevin (FARAGO, 2019). As simulações de produção para estudo do sistema foram realizadas para 100 ns. O N3 foi usado como droga de referência padrão para analisar as interações entre o ligante e a proteína.

A qualidade das estruturas obtidas em MDs foram avaliadas usando os seguintes parâmetros com NAMD: Energia Potencial (kcal/mol)(DIEZ et al., 2014); Energia de

interação proteína-ligante (kcal/mol); Root Mean Square Deviation (RMSD, Å) da proteína, dos ligantes e das distâncias entre eles; Flutuação quadrática média da raiz (RMSF, Å), distâncias mínimas entre proteínas e ligantes observadas em MD (ARSHIA et al., 2021). As ligações de hidrogênio foram avaliadas com Visual Molecular Dynamics (VMD) (HUMPHREY; DALKE; SCHULTEN, 1996). Os gráficos serão gerados usando o programa Qtrace (LIMA et al., 2012).

2.11. Cálculos MM/GBSA

As energias da mecânica molecular combinadas com a solvatação contínua de Born e área de superfície generalizada MM/GBSA foram calculadas por MolAICal (BAI et al., 2021) com base no arquivo de log MD do software NAMD (PHILLIPS et al., 2005). O MM/GBSA é estimado pelas equações 1, 2 e 3.

$$\Delta G_{\text{bind}} = \Delta H - T\Delta S \approx \Delta E_{\text{MM}} + \Delta G_{\text{sol}} - T\Delta S \quad \text{Eq. 1}$$

$$\Delta E_{\text{MM}} = \Delta E_{\text{interno}} + \Delta E_{\text{ele}} + \Delta E_{\text{vdw}} \quad \text{Eq. 2}$$

$$\Delta G_{\text{sol}} = \Delta G_{\text{GB}} + \Delta G_{\text{SA}} \quad \text{Eq. 3}$$

Onde ΔE_{MM} , ΔG_{sol} e $T\Delta S$ representam a energia MM da fase gasosa, energia livre de solvatação (soma da contribuição polar ΔG_{GB} e contribuição apolar ΔG_{SA}) e entropia conformacional, respectivamente. ΔE_{MM} contém energia de Van Der Waals ΔE_{vdw} , eletrostática ΔE_{ele} e $\Delta E_{\text{interna}}$ de ligação, ângulo e energia diedro. Se não houver mudanças estruturais induzidas por ligações no processo de simulações de dinâmica molecular (MD), o cálculo da entropia pode ser omitido.

2.12. Método *in silico* ADMET

Inicialmente, foi usado o Marvin JS versão Iodine 6 suporte de longo prazo (LTS) e software de licença acadêmica, Chemaxon (<https://chemaxon.com/marvin>) para estimar a distribuição relativa de porções químicas calculando a constante de ionização (pKa) e convertendo em notação linear do Simplified molecular-input line-entry system (SMILES) para estimativa quantitativa de semelhança com drogas pelo critério de "Medicinal Chemistry Evolution 2018" (MCE18), conforme a equação 4, presente no servidor ADMETlab 2.0 (<https://admetmesh.scbdd.com/>).

$$\text{MCE18} = \left(\text{AR} + \text{NAR} + \text{Chiral} + \text{Spiro} \frac{\text{sp}^3 + \text{Cyc} - \text{Acyc}}{1 + \text{sp}^3} \right) Q^1 \quad \text{Eq. 4}$$

Onde a distribuição da fração de átomos sp^3 (F_{sp^3}) entre estruturas cíclicas, que incluem anéis aromáticos (AR) e não aromáticos (NAR), e estruturas acíclicas e a presença de centros quirais e grupos espirocíclicos resulta em uma pontuação que expressa o equilíbrio entre complexidade estrutural e grau de similaridade com compostos registrados em patentes nos últimos anos, considerando o limite de idealidade de MCE18 > 45, que expressa o melhor ajuste neste espectro (IVANENKOV; ZAGRIBELNYY; ALADINSKIY, 2019). Os resultados foram comparados ao espectro físico-químico de duas regras da empresa farmacêutica Pfizer: regra do Triângulo Dourado 2/5 (JOHNSON; DRESS; EDWARDS, 2009) e a regra 3/75 da Pfizer (HUGHES et al., 2008).

Para predição dos descritores farmacocinéticos, foram considerados os modelos de absorção, distribuição, metabolismo, excreção e toxicidade (ADMET) de servidores disponíveis para uma análise consensual entre decisões empíricas e dados de testes *in vitro* e *in vivo* (ROCHA et al., 2022). Aqui, usamos um servidor de previsão ADMET baseado em IA – ADMETboost, AI DrugLab (<https://ai-druglab.smu.edu/admet>) – que reúne vários bancos de dados moleculares disponíveis online para uma previsão mais precisa. , classificado em 1º lugar em 18 de seus 23 modelos ADMET, que incluem: permeabilidade passiva (P_{app}), inibição da glicoproteína P (Pgp), absorção intestinal humana (HIA), fração de biodisponibilidade oral (F%), ligação às proteínas plasmáticas (PPB), permeabilidade da barreira hematoencefálica (BBB), substratos do citocromo P450 (CYP450) isoformas, depuração intrínseca da fração não ligada ($CL_{int,u}$) no microsomo hepático humano (HLM), dano hepático induzido por drogas (DILI), mutagenicidade de Ames (XU et al., 2012), e estimativa da concentração letal (LC_{50}), onde os resultados foram comparados com os indicadores farmacocinéticos do servidor ADMETlab 2.0 (<https://admetmesh.scbdd.com/>). Para estimar a toxicidade orgânica, foi realizada uma previsão do local do metabolismo (SOM) no servidor online SMARTcyp (https://smartcyp.sund.ku.dk/mol_to_som), onde os resultados foram comparados com os descritores do metabolismo pelas isoformas CYP450 (CYP2C9, CYP2D6 e CYP3A4) e toxicidade do teste de consenso ADMET, incluindo previsão de fragmentos moleculares reativos da biblioteca e Brenk (BRENK et al., 2008), presente no servidor SwissADME (<http://www.swissadme.ch/>).

3. RESULTADOS E DISCUÇÕES

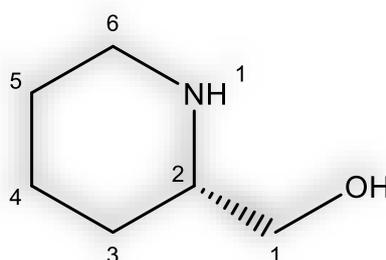
3.1. Processamento

A geração molecular face a proteína alvo foi realizada usando o pacote computacional MolAiCal (BAI et al., 2021) em uma CPU Intel® Core Intel de 10ª geração, até 32 GB de RAM e uma GPU NVIDIA® GeForce® GTX 1660 Ti, com um tempo de varredura definido para 8h. Dez clusters foram usados para gerar as estruturas. Uma média de aproximadamente 11.655 moléculas foram desenvolvidas para a proteína alvo durante o período experimental.

3.2. Produção de hits

Após o novo método de design de drogas, usando software de inteligência artificial MolAiCal, obtivemos 1.722 acertos (substância ativa no sistema), a partir da estrutura de crescimento inicial (S)-piperidin-2-ilmetanol, na Figura 2.

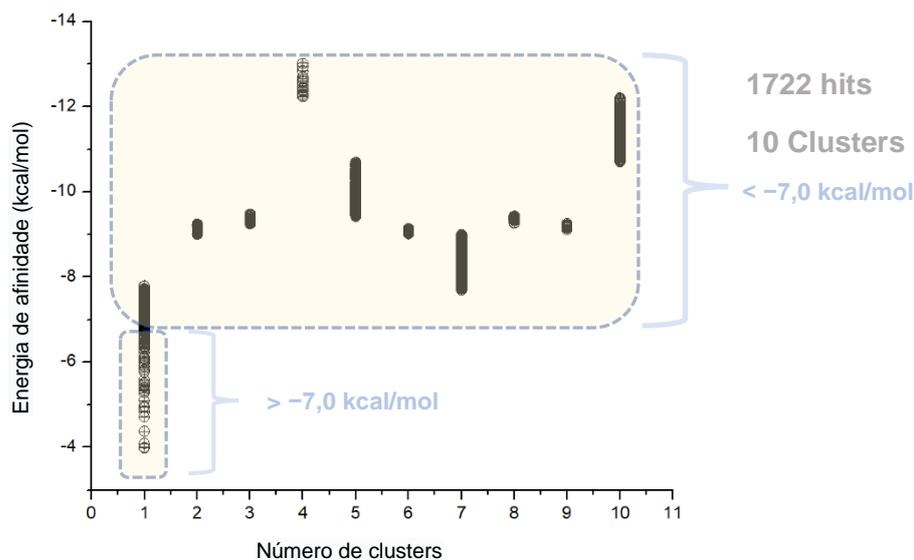
Figura 2. Estrutura de crescimento inicial (S)-piperidin-2-ilmetanol.



Fonte: Autor (2023)

Foram processados até 10 clusters, sendo que todos obtiveram estruturas com variação de energia de afinidade variando de -13,02 a -7,7 kcal/mol, com exceção de boa parte dos ligantes gerados pelo cluster 1, que obtiveram energias menos favoráveis. Conforme mostrado na Figura 3.

Figura 3. Número de clusters versus energia de afinidade para MolAICaID.



Fonte: Autor (2023)

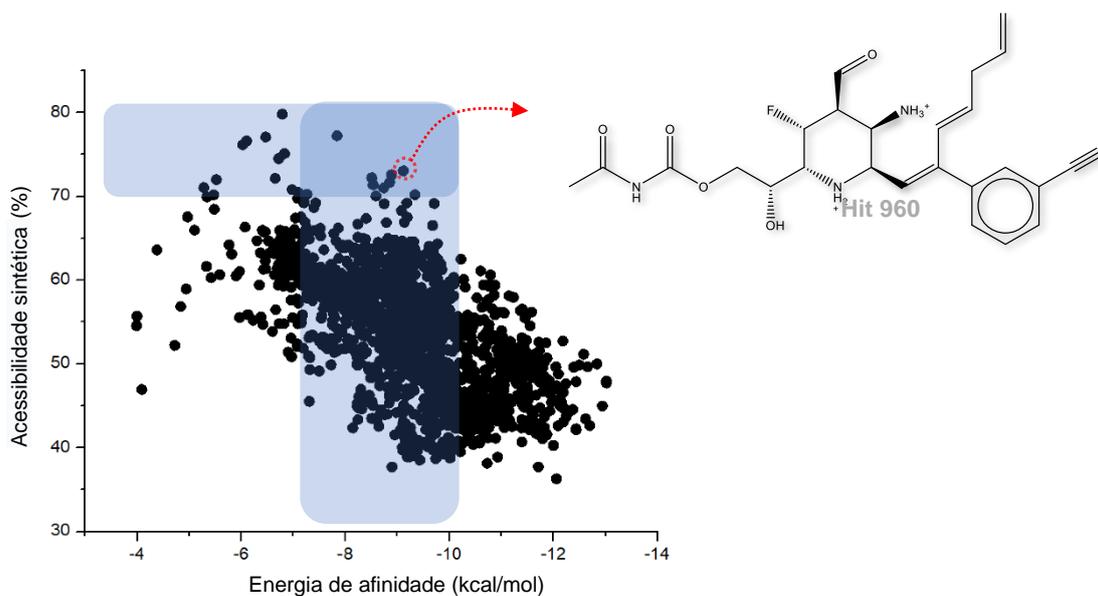
3.3. Disponibilidade sintética

Uma das vantagens de usar uma nova abordagem de design de drogas é que alguns fragmentos bem posicionados e ligações de hidrogênio (lipofílicas neste caso) podem fornecer interação energética significativa. No entanto, à medida que a diversidade de fragmentos plausíveis aumenta, a oportunidade de criar composições terapêuticas se tornará mais difícil (JAIN & AGRAWAL, 2004; WANG et al., 2022). Outra questão está relacionada à limitação desse método, que pode sugerir diretamente testes dos ligantes formados. Uma abordagem para estimar a facilidade de síntese de um ligante é chamada de acessibilidade sintética (SA), usada para gerar moléculas semelhantes a drogas e necessária em muitas áreas no processo de descoberta de drogas. A avaliação do SA de um candidato a principal (hit) é uma tarefa que desempenha um papel na descoberta do lead, independentemente do método pelo qual o candidato a principal é identificado. Quanto mais complexa a síntese do candidato principal, mais tempo e recursos são necessários para explorar essa área específica do espaço químico. No caso de uma molécula projetada em *De novo*, a validação experimental de sua atividade requer a síntese do composto. No caso de exploração prática ou virtual da SAR em torno do hit, também pode ser necessário o acesso sintético ao quimiotipo (SCOTTI et al., 2013).

Quando as estruturas químicas são construídas durante o processo do novo projeto, não se pode presumir que a síntese química de tais compostos seja viável. Portanto, é ainda mais importante estimar se esses compostos podem ser sintetizados com esforço razoável. O padrão de acessibilidade sintético (SA) dos acertos do estudo

apresentou um comportamento bem característico da literatura, onde quando melhor se encaixa a proteína, mais complexa se torna a síntese (ERTL; SCHUFFENHAUER, 2009). Porém, é possível obter acertos com facilidade de 70-80%, com afinidade energética de até -7 a -10 kcal/mol, conforme mostra a Figura 4. O destaque dos ligantes ficou evidenciado pelo composto Hit 960, que teve uma melhor afinidade e acessibilidade sintética.

Figura 4. Acessibilidade sintética versus energia de afinidade sugerida por MolAiCAID.



Fonte: Autor (2023)

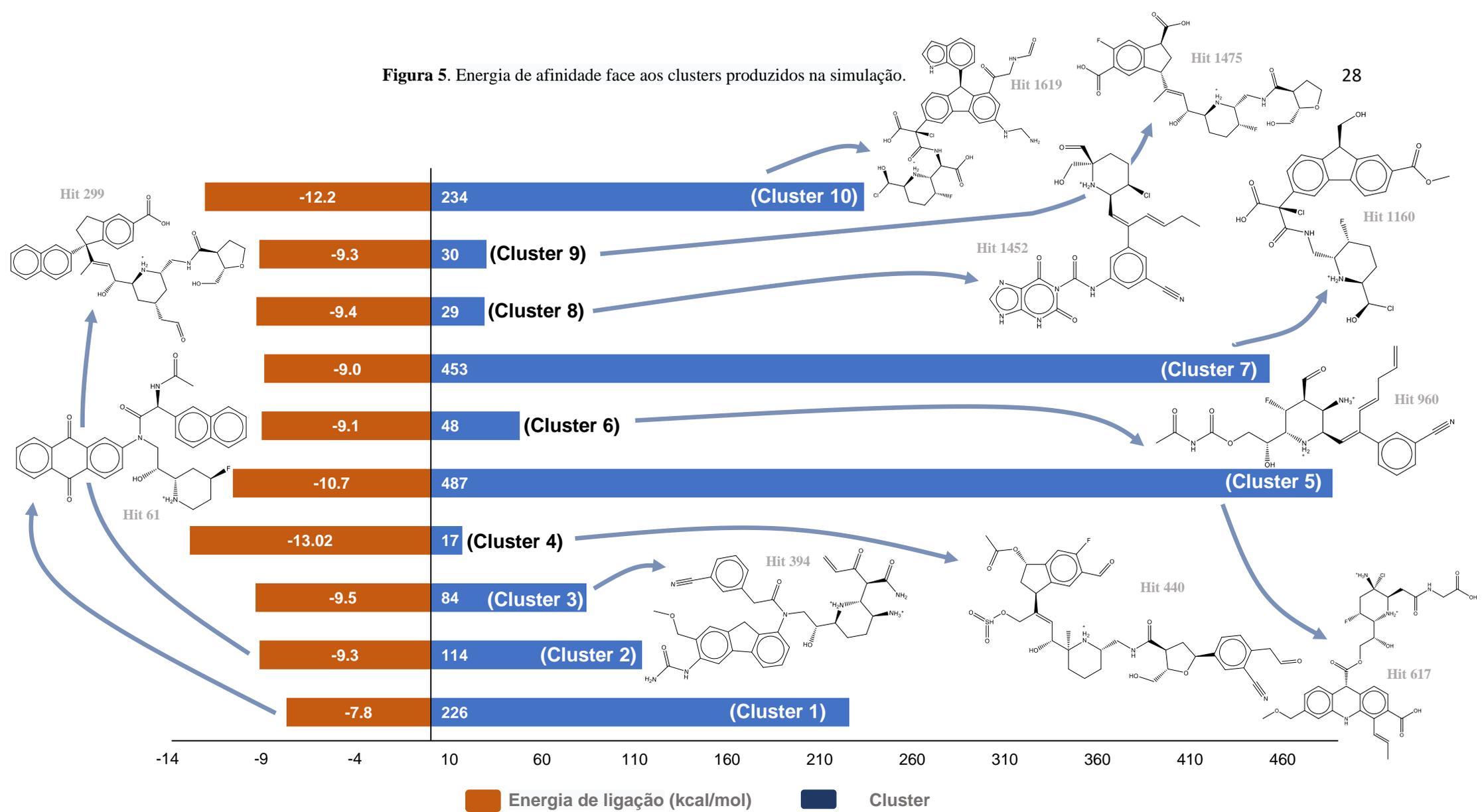
3.4. Energia de afinidade

Em seguida, os 100 melhores resultados mais próximos em geometria e caixas eletrostáticas foram selecionados a partir dos análogos produzidos, acionando a opção dos ligantes serem agrupados em clusters de similaridade. Esses representantes foram tomados como representantes de cada cluster e ordenados de acordo com a afinidade/disponibilidade sintética de energia.

Os experimentos de geração mostraram que as moléculas obtidas posteriormente tiveram uma pontuação de melhor ajuste. Os resultados da proteína demonstram bem essa tendência porque os clusters gerados foram formados proporcionalmente ao tempo. Assim, isso é possível porque valores maiores geram moléculas mais diversas, mas a convergência dos escores de ancoragem tornam-se mais pobre. De acordo com a metodologia, o cluster 1 inicialmente apresentava estruturas que não interagiam muito bem com a proteína. No entanto, uma pequena parcela da população cedeu energias de

até -7,8 kcal/mol, com destaque para o Hit 61. Do Cluster 2 ao Cluster 10, foi observado um aumento na afinidade de ligação com a proteína, apresentando energias de até -13,02 kcal/mol, onde a estrutura mais favorável de cada cluster foi descrita conforme a sequência a seguir; O cluster 2 deu Hit 299 com uma afinidade de ligação de -9,25 kcal/mol; O Cluster 3 apresentou Hit 394 com energia de afinidade de -9,47 kcal/mol; O Cluster 4 apresentou hit 440 com energia de ligação de -13,02 kcal/mol, que foi considerada a melhor energia da simulação. Continuando a discussão, o Cluster 5 apresentou Hit 617, com melhor energia em seu grupo, com -10,7 kcal/mol; O Cluster 6 apresentou o hit 960, com melhor energia de afinidade, -9,12 kcal/mol. Para os Clusters 7, 8, 9 e 10, os seguintes valores de cada representante, Hit 1160 (-9,0 kcal/mol), Hit 1452 (-9,43 kcal/mol), Hit 1475 (-9,26 kcal/mol) e Hit 1619 (-12,21 kcal/mol), respectivamente, como foram apresentados. Observou-se que quanto melhor a energia de ligação apresentada no complexo receptor-ligante, maior a complexidade estrutural dos compostos obtidos. O cluster 6, em particular, apresentou um modelo que melhor se adaptou à acessibilidade sintética 75%, o Hit 960, com uma energia de afinidade de -9,12 kcal/mol, muito mais favorável que o ligante de referência. Esses resultados são apresentados na Figura 5.

Figura 5. Energia de afinidade face aos clusters produzidos na simulação.

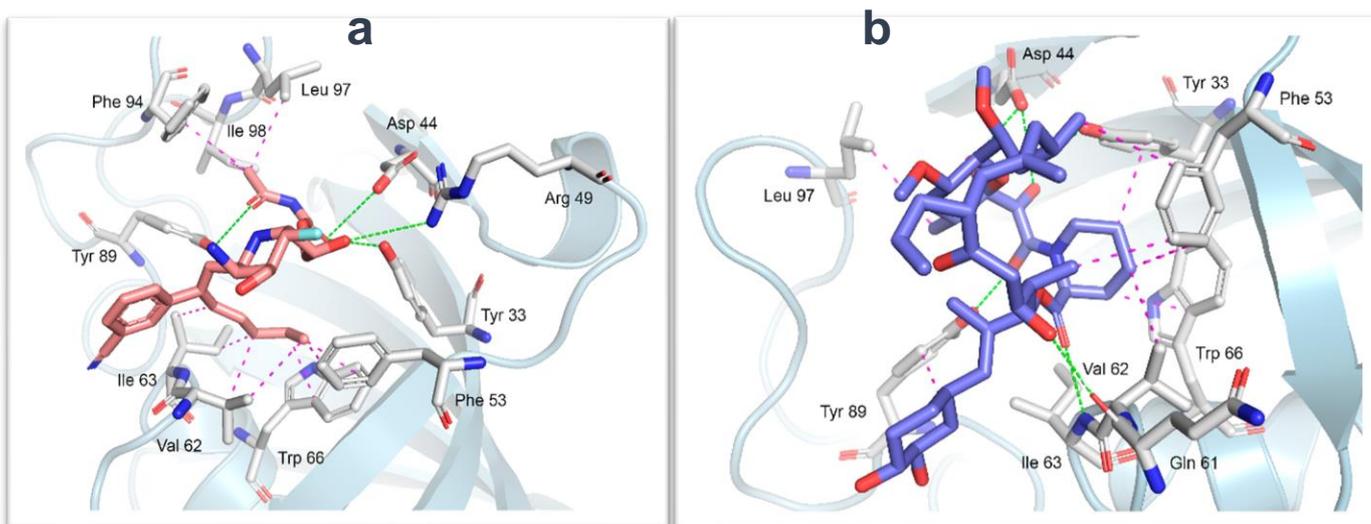


Fonte: Autor (2023)

3.5. Interação com resíduos de proteínas

O Hit 960 forneceu no docking molecular um RMSD de 1,3 Å, com uma energia de afinidade de -9,12 kcal/mol, interagindo na mesma região do ligante nativo FK-506, conforme mostrado na Figura 6. Com interações robustas, o Hit 960 apresentou três pontes de hidrogênio intramoleculares nos resíduos Asp 44, Arg 49 e Tyr 33. Um intermolecular na região oxigenada, acompanhado de interação hidrofóbica do tipo alquil com Phe 94, Leu 97 e Ile 98, que fazem parte da região ativa de a proteína (BASHIR et al., 2020). O ligante nativo FK-506, utilizado como referência positiva na pesquisa, apresentou as interações polares representadas por cinco pontes de hidrogênio nos resíduos Asp 44 (dois), Tyr 89, Val 62 e Gln 61, além das interações hidrofóbicas do tipo alquila com Leu 97, necessária na região ativa da proteína, Tyr 89, Trp 66, Phe 53 e Tyr 33.

Figura 6. (a) 3D e 2D das interações entre a estrutura mais promissora (Hit 960) e (b) o ligante de referência e os resíduos da polimerase.



Fonte: Autor (2023)

3.6. Cálculos MM/GBSA

Após o balanceamento da dinâmica de produção, a amostragem das etapas foi realizada de 5 a 5, seguindo o intervalo de amostragem de 10 ns da metodologia de estimativa da variação de energia livre utilizando trajetórias múltiplas. Os cálculos de MMGBSA foram realizados em um campo de solvente implícito simulando uma solução salina 0,15M.

Embora formalmente o cálculo da variação de energia livre nessa técnica passe pela análise da entropia a partir dos modos normais das equações 4, 5 e 6 do sistema, na prática esse cálculo não foi feito.

$$\Delta A_{interação}^{(vac)} = (E_{complex}^{MM} - E_{alvo}^{MM}) - (E_{complex}^{MM} - E_{ligante}^{MM}) + T\Delta S_{NORMODS} \quad \text{Eq. 4}$$

$$\Delta A_{interação}^{(vac)} = E_{complex}^{MM} - E_{alvo}^{MM} - E_{complex}^{MM} + E_{ligante}^{MM} + T\Delta S_{NORMODS} \quad \text{Eq. 5}$$

$$\Delta A_{interação}^{(vac)} = E_{ligante}^{MM} - E_{alvo}^{MM} + T\Delta S_{NORMODS} \quad \text{Eq. 6}$$

Cálculos de modos normais são bastante demorados e computacionalmente caros. Esse tipo acaba tornando os cálculos de triagem virtual, que são o foco deste estudo, não totalmente impossíveis. No entanto, há uma razão mais importante. Foi demonstrado que os cálculos de entropia diminuem a correlação dos valores de afinidade previstos com os valores experimentais quando a análise é feita com alguns microestados amostrados das trajetórias (HOU et al., 2011; RASTELLI et al., 2009). Como as limitações de tempo de cálculo são essenciais, não é recomendável incluir esses cálculos no procedimento.

Energias MM/GBSA são consideradas uma forma de estimar energia livre para estudo *in silico* de ligantes em complexos proteicos (GENHEDEN; RYDE, 2015). Eles são tipicamente baseados em simulações MD e trazem precisão entre pontuação empírica e perturbação alquímica estrita (CHEN et al., 2018). Como na entropia conformacional, é difícil obter um valor concorrente. Principalmente, se os ligantes não tiverem nenhuma alteração estrutural induzida por ligação em simulações MD, a entropia conformacional é geralmente ignorada para calcular pela análise de modo padrão (WANG et al., 2018). MolAICal forneceu uma maneira rápida de avaliar a energia livre de ligação sem entropia do ligante com base na abordagem de três trajetórias. Os valores mostraram que o complexo Hit 960/polimerase obteve um resultado bastante aproximado, baseado em sua energia livre, com -13,67 kcal/mol em comparação com o ligante de referência FK-506, que continuou sendo o mais estável no sistema de estudo, com a livre energia de -16,23 kcal/mol. Portanto, a técnica de decomposição de energia de interação revelou a contribuição do complexo ligante-receptor e sua energia final na Tabela 1.

Tabela 1. Dados de estimativa de energia livre de Hit 960 e FK-506 contra polimerase de *Candida auris*.

Complexo	ΔE (interno)	ΔE (eletrostática) + ΔG (sol)	ΔE (VDW)	ΔG ligação (kcal/mol)	Desvio padrão
----------	-------------------------	--	---------------------	----------------------------------	------------------

FK-506/Polimerase	0,0	22,24	-38,46	-16,23	+/- 0,0293
Hit 960/Polimerase	0,0	23,66	-37,33	-13,67	+/- 0,0236

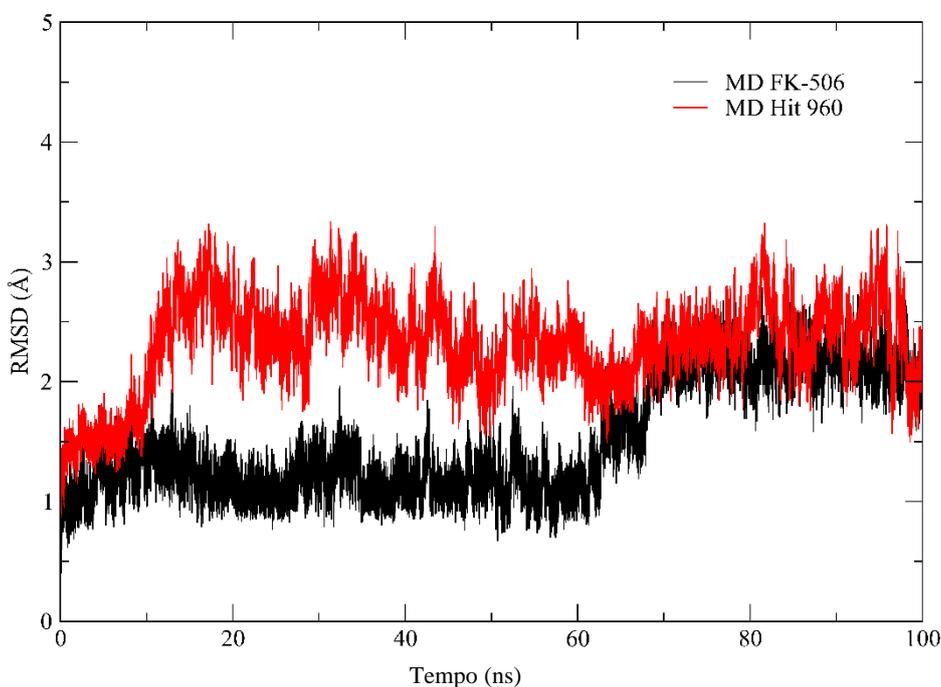
Fonte: Autor (2023)

3.7. Dinâmica molecular

3.7.1 Análise RMSD

Após a análise dos valores de energia, outro parâmetro importante para investigar a qualidade da dinâmica molecular é Desvio Médio Quadrática de Rotação (RMSD) do esqueleto da proteína para Hit 960 e FK-506. Os valores de RMSD obtidos pelo esqueleto da proteína ao longo dos MDs mostram todos os valores entre 1 e 3,50 Å. No MD com o ligante Hit 960, o perfil fica entre 1,5 a 3,5 Å e permanecendo mais estável em 97 ns com RMSD próximo a 2 Å. Logo em seguida no MD do ligante FK-506 de referência apresentou uma situação de adequação próxima a 70 ns o sistema mais favorável durante a trajetória de 100 ns, conforme mostra a Figura 7.

Figura 7. Perfil do valor de RMSD obtido para MD com o ligante de referência FK-506 (linha preta) e MD com Hit 960 (linha vermelha).

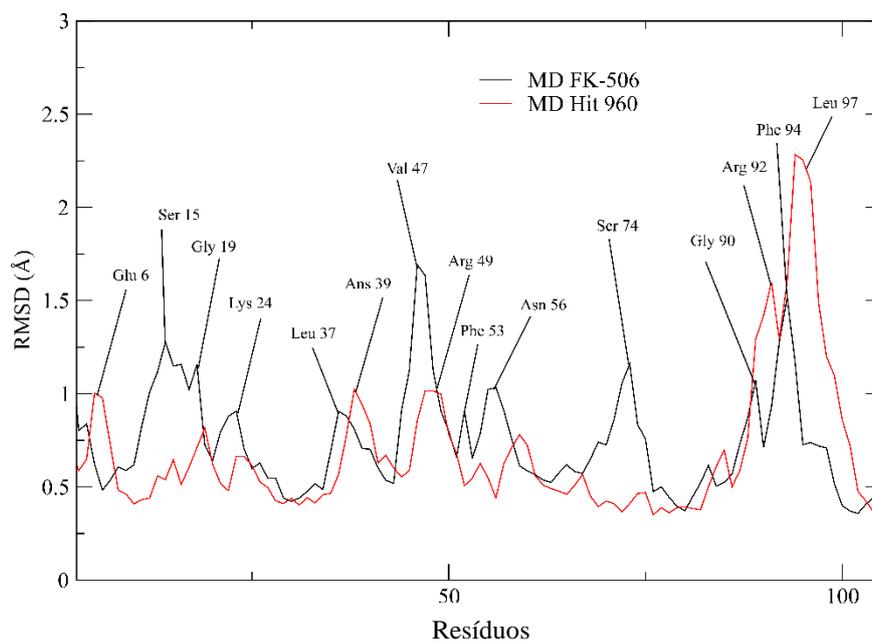


Fonte: Autor (2023)

3.7.2. Análise RMSF

A Flutuação Média Quadrática de Rotação (RMSF) é um parâmetro relacionado à flexibilidade de resíduos de proteínas individuais, servindo para avaliar qualitativamente a progressão da dinâmica molecular (DONG et al., 2018). Considerando a Figura 8, observa-se um perfil semelhante de valores de RMSF independente do ligante em contato.

Figura 8. Perfil do valor de RMSF obtido para polimerase em MD com Hit 960 (linha vermelha), MD com FK-506 (linha preta).

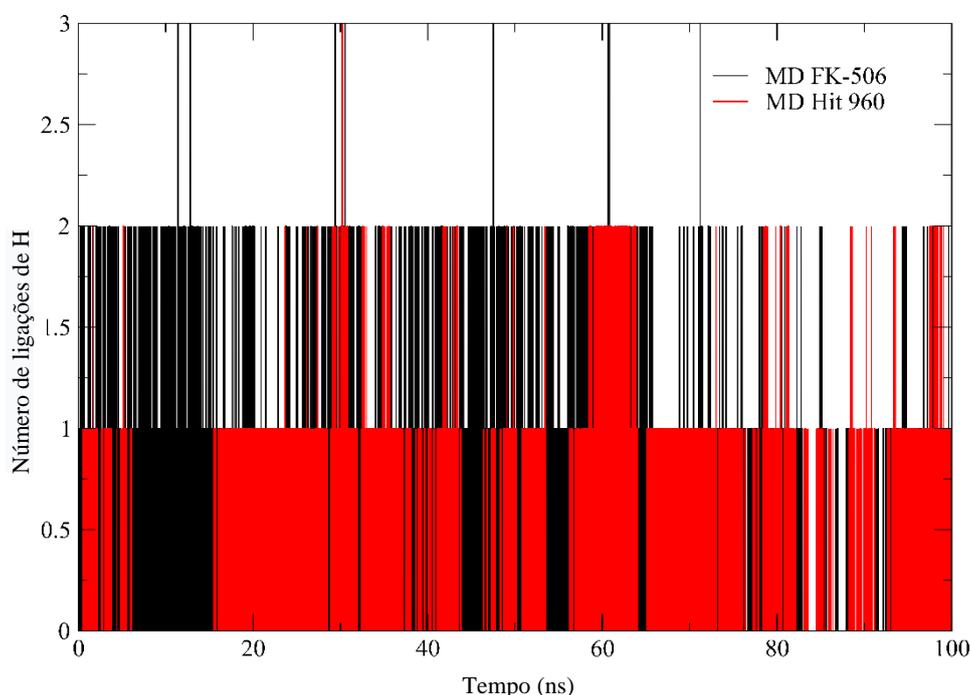


Fonte: Autor (2023)

3.7.3. Ligações de H

O número de ligações de hidrogênio (H-Bond) encontradas durante MDs, considerando o valor máximo de 3,3 Å, são mostrados na Figura 9 e Tabela 2. Pode-se observar que MD com Hit 960 apresenta até 3 ligações de hidrogênio por quadro e vários quadros com duas ligações de hidrogênio. Sobre MD com o ligante de referência FK-506, há apenas dez quadros com três ligações de hidrogênio e vários quadros com duas ligações de hidrogênio durante a dinâmica molecular. Portanto, pode-se inferir que tanto o FK-506 quanto o Hit 960 tendem a interagir com a polimerase, tornando-o um possível antifúngico eficiente contra esse superfungo.

Figura 9. Número de ligações de hidrogênio (H-Bond) encontrado entre *Candida auris* polimerase e MD com FK-506 (linha preta) e MD com Hit 960 (linha vermelha) com software VMD, considerando o comprimento máximo 3,3 Å.



Fonte: Autor (2023)

Tabela 2. Resíduos da polimerase *Candida auris* que mostraram ligação H ao longo das MDs.

Sistema	Ligação de H
MD com FK-506	Asp 44, Gln 61, Val 62, Tyr 89
MD com Hit 960	Tyr 33, Asp 44, Arg 49

Fonte: Autor (2023)

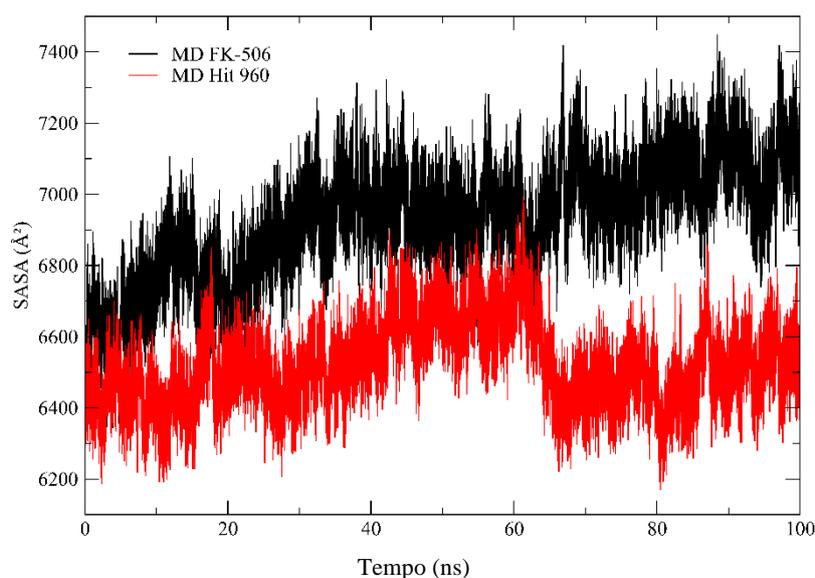
Nos resíduos identificados com ligações de H ao longo da dinâmica molecular, na Tabela 2, observa-se a recorrência dos resíduos Asp 44, demonstrando assim um potencial de interação tanto de Hit 960, quanto da interação entre polimerase e FK-506.

3.7.4. Área de Superfície Acessível ao Solvente

Área de superfície acessível ao solvente (SASA) é definida como a área de superfície de uma proteína que interage com as moléculas de solvente (MAZOLA et al., 2015). Os valores médios SASA livre para FK-506 e complexos Hit 960 foram monitorados durante simulações MD de 100 ns. Os traços para o SASA na Figura 10 mostram um aumento acentuado dentro de 10-20 ns, indicando relaxamento estrutural. Os valores médios de SASA para os complexos FK-506 e Hit 960 livres foram 6,970 Å² e 6,499 Å², respectivamente. Não houve alteração significativa observada nos valores de

SASA devido à ligação dos ligantes. Após esse tempo, os valores flutuam em torno de um valor constante. Assim, assumimos que os tempos de simulação de 100 ns foram suficientes para a amostragem de sistemas equilibrados. O SASA mais alto é encontrado para as moléculas de polimerase com os íons monovalentes estabilizadores. A corrida sem íons monovalentes apresenta uma grande flutuação, enquanto os sistemas com maiores concentrações de íons têm áreas menores e podem estar encolhendo sob a influência da carga de superfície, resultando em estruturas proteicas mais compactas. Uma inspeção mais aprofundada dos dados demonstra que a flutuação ou "respiração" da superfície relaxada se deve principalmente a uma flutuação do SASA da área C-terminal flexível.

Figura 10. Área de superfície acessível ao solvente (SASA) da polimerase *Candida auris* em função do tempo das simulações MD: MD com Hit 960, estabilizando íons monovalentes (linha preta) e MD com FK-506, estabilizando íons monovalentes (linha vermelha). As curvas são médias dos dados brutos com uma janela de 100 ns.



Fonte: Autor (2023)

3.8. Estudo *in silico* ADMET

3.8.1. Estimativa quantitativa da semelhança com drogas

Uma nova singularidade química tem sido frequentemente adotada entre os novos compostos registrados em patentes como agentes terapêuticos nos últimos anos. Essa singularidade se concentra na forma como a fração de átomos sp^3 (F_{sp^3}) se distribui entre estruturas cíclicas e acíclicas (IVANENKOV; ZAGRIBELNYY; ALADINSKIY, 2019),

conforme evidenciado na Equação 4. Esses compostos são ligeiramente maiores e mais polares do que os compostos vendidos comercialmente e compostos ativos no SNC (IVANENKOV; ZAGRIBELNYY; ALADINSKIY, 2019)(WAGER et al., 2016). Para Hughes (2008) (HUGHES et al., 2008), compostos com baixa polaridade, ou seja, com $TPSA \leq 75 \text{ \AA}^2$, podem ser potencialmente tóxicos, numa faixa de alta lipofilicidade ($\log P > 3$), além de serem distribuídos para o cérebro, indicando uma resposta neurotóxica *in vivo*.

No mapa de potencial lipofílico molecular (MLP) da Figura 11a, é possível observar que os centros carregados positivamente de Hit 960, ou seja, os grupos NH_3^+ e NH_2^+ , presentes na estrutura cíclica alifática, contribuem para uma região levemente hidrofílica (espectros de amarelo a vermelho), enquanto a cadeia lateral do polieno e a estrutura aromática do composto formam regiões mais lipofílicas da superfície molecular (espectros de verde a azul). Assim, a diferença $\log P < \log D$ é um produto direto da desprotonação do grupo amina cíclico ($pK_a = 5,14$), onde o equilíbrio químico é deslocado para a formação da espécie monoprótica em $\text{pH} > 5,14$ (Figura 11b) com concentração relativa de 95,93% em $\text{pH} 7,4$ e, conseqüentemente, aumento da lipofilicidade, com valores fisiológicos de $\log D$ na ordem de 0,927 em relação à lipofilicidade intrínseca ($\log P = 0,093$).

Esses valores estão de acordo com o espaço físico-químico da Pfizer, onde residem compostos com excelente absorção e baixa depuração hepática, ou seja, com baixa lipofilicidade fisiológica ($\log D < 5,0$) em uma faixa de PM inferior a 500 g/mol (Figura 11c)(JOHNSON; DRESS; EDWARDS, 2009). Além disso, o F_{sp^3} calculado de 0,385, distribuído entre duas estruturas cíclicas, das quais 1 é um anel aromático (AR), resultou em uma pontuação de $MCE-18 = 55,72$ (Tabela 3), em uma escala em que valores > 45 significam que o composto apresenta estrutura relevante e com grau de complexidade dentro do espectro físico-químico de compostos registrados em patentes nos últimos anos (Figura 11d). O composto atende à singularidade química exigida por ser uma substância mais polar do que os compostos disponíveis comercialmente, com um valor $TPSA = 134,31 \text{ \AA}^2$ como soma das superfícies polares de seus grupos aceitadores de 8 H-bond (HA) que, quando alinhados ao fundo lipofilicidade relativa ($\log P < 3$), ocupa um espaço físico-químico de maior segurança em testes *in vivo* (Figura 11d).

Tabela 3. Propriedades físico-químicas do Hit 960 aplicadas aos critérios de similaridade com drogas Pfizer e MCE-18.

Propriedade	Valor	Ótimo
pKa mais básico	5,14	< 8,0
MW	482,23 g/mol	< 500 g/mol
logP	0,093	< 3
logD	0,927	< 3
TPSA	134,31 Å ²	> 75 Å ²
nHA	8	< 10
nHD	4	< 5
nRing	2	< 6
AR	1	< 3
Alerta Brenk	3 alertas: aldeído, alceno isolado, polieno	
F _{sp} ³	0,85	
Regra 2/5 da Pfizer	Aceito	
	(-2 < logD < 5 e 200 < MW < 500 g/mol)	
Regra 3/75 da Pfizer	Aceito	
	(logP < 3 e TPSA > 75 Å ²)	
MCE-18	55,722	> 45,0

Fonte: Autor (2023)

3.8.2. Descritores farmacocinéticos previstos

Em uma otimização multiparamétrica aperfeiçoada nos últimos anos pela farmacêutica Pfizer, é possível perceber um deslocamento do espaço físico-químico para uma região concentrada por compostos com alto TPSA e baixo logP, que apresentam um alinhamento entre segurança e três atributos fundamentais de ADMET: alta permeabilidade passiva ($P_{app} > 10 \times 10^{-6}$ cm/s), baixo risco de efluxo de Pgp e baixa taxa de depuração hepática intrínseca ($CL_{int,u} < 100$ mL/min/kg)(WAGER et al., 2016)(WAGER et al., 2010), *in vitro* attributes that provide high absorption and metabolic stability to a new drug candidate and that can be predicted early (JOHNSON; DRESS; EDWARDS, 2009).

As decisões empíricas observadas da etapa druglikeness corroboram com os descritores ADMET estimados, onde o valor P_{app} MDCK na ordem de $1,2 \times 10^{-5}$ cm/s sugere uma alta permeabilidade passiva que, quando combinada com a baixa probabilidade de efluxo passivo por Pgp (38,72%), estima uma absorção intestinal de 68,6%. Enquanto isso, a taxa estimada de $CL_{int,u}$ de 40,98 mL/min/kg produz uma fração de biodisponibilidade oral (F%) de 38,43% (Tabela 4).

Além disso, a baixa lipofilicidade dos ligantes permite maior fração molecular não ligada às proteínas séricas, com valor de ligação às proteínas plasmáticas (PPB) em torno de 55,98% (Tabela 4), favorecendo a distribuição do ligante no plasma sanguíneo e em tecidos biológicos (PIRES; KAMINSKAS; ASCHER, 2018). Concordando com o espaço físico-químico de segurança (Figura 11d), foi possível estimar uma baixa probabilidade

de permeabilidade na barreira hematoencefálica (BBB) do Hit 960, indicando uma baixa possibilidade de distribuição do composto no cérebro (Tabela 4).

3.8.3. Local de metabolismo e previsão de toxicidade.

O método de predição do sítio metabólico (SOM) aqui utilizado é capaz de relacionar a sensibilidade dos grupos funcionais a serem biotransformados pelas isoenzimas CYP450 em função do grau de especificidade desses grupos e fragmentos, para prever os possíveis metabólitos secundários formados no ser humano microssomas hepáticos (HLM)(ZHENG et al., 2009). A previsão desses SOMs permite estimar efeitos colaterais de metabólitos secundários reativos, como a formação de epóxidos por hidroxilação aromática ou ligações duplas de estruturas alifáticas, que apresentam um potencial agente eletrofílico capaz de se ligar covalentemente a macromoléculas como proteínas e DNA (HUGHES; MILLER; SWAMIDASS, 2015)(YU et al., 2014).

Aqui, o teste preditivo ADMET mostrou que Hit 960 pode ser um substrato da isoforma CYP2D6, com uma probabilidade estimada de 56,48%. No mapa de similaridade 2D da Figura 11e, é possível observar que o alceno isolado da cadeia lateral do polieno tende a sofrer uma hidroxilação alifática, formando um intermediário epoxidado instável e quimicamente reativo, conforme identificado anteriormente pela biblioteca de Brenk (BRENK et al., 2008) (Tabela 3). A partir dos descritores de toxicidade, foi possível observar que o composto tem uma probabilidade de 54,43% de induzir lesão hepática (DILI), devido à formação de metabólitos secundários, incluindo o risco moderado de resultar em mutagenicidade por interações com estruturas de DNA (40,34%). No entanto, o valor da concentração letal (CL_{50}) da ordem de 5,017 log(mg/L) indica que a dose oral administrada oferece baixo risco de mortalidade em uma população de Fathead Minnow, permitindo o controle da dose oral diária (Tabela 4).

Tabela 4. Propriedades farmacocinéticas previstas por modelos ADMET dos servidores ADMETlab 2.0 e ADMETboost – AI Drug Lab.

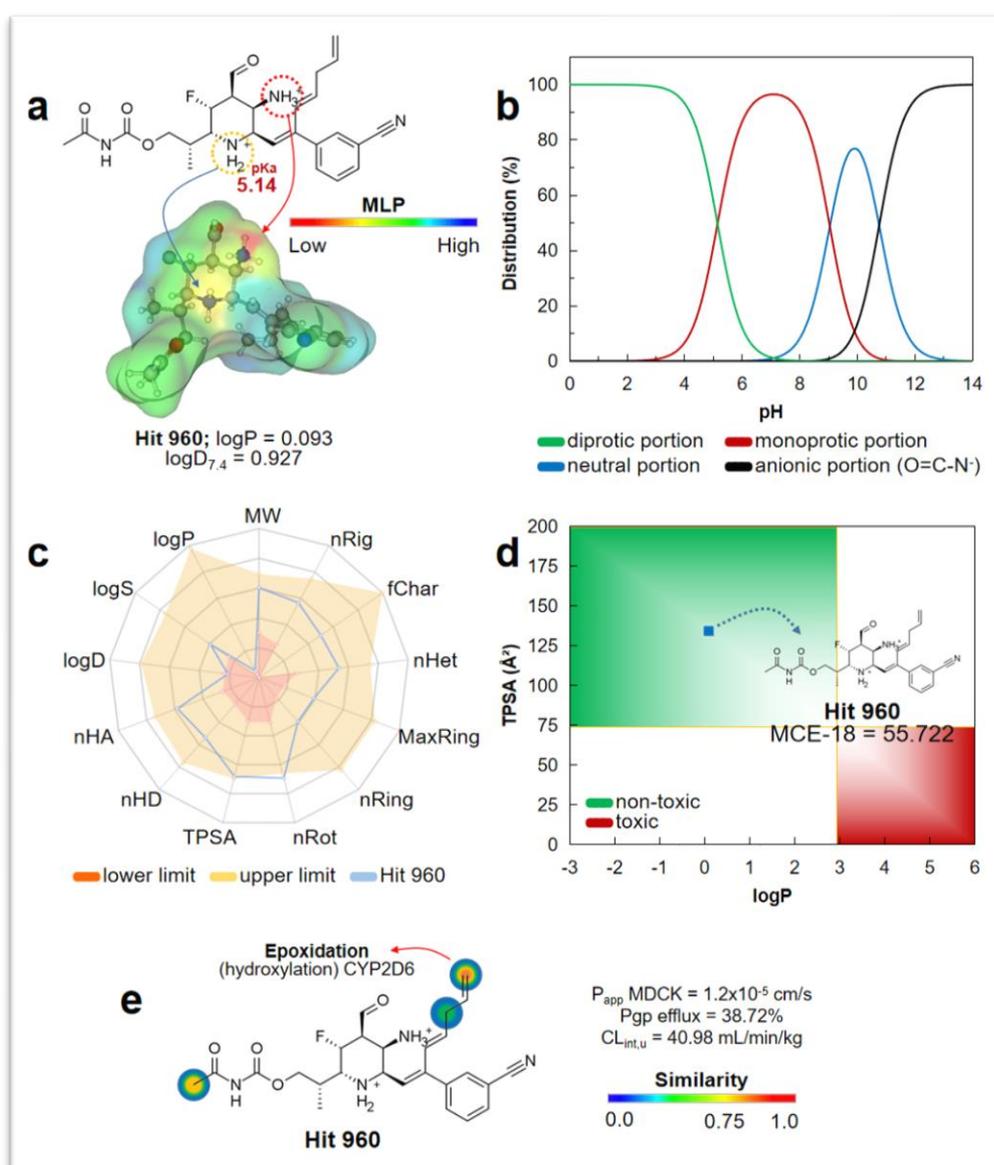
Propriedade	Predição	Fonte
Absorção		
P_{app} MDCK	$1,2 \times 10^{-5}$ cm/s	ADMETlab 2.0
Inibição Pgp	38,72% (Não)	Consenso
HIA	68,6% (Bom)	Consenso
F%	38,43%	ADMETboost
Distribuição		
PPB	55,98%	ADMETboost
BBB	27,21% (Não)	Consenso
Metabolismo		

CYP2C9 substrato	32,47%	ADMETboost
CYP2D6 substrato	56,48%	ADMETboost
CYP3A4 substrato	34,19%	ADMETboost
Excreção		
CL _{int,u}	40,98 mL/min/kg	ADMETboost
Toxicidade		
DILI	54,43% (Sim)	Consenso
Mutagênico	40,34 (Moderado)	Consenso
LC ₅₀ Fathead Minnow	5,017 mg/L	ADMETlab 2.0

Nota: entre parênteses, a decisão empírica do ADMETlab 2.0.

Fonte: Autor (2023)

Figura 11. Mapa de potencial lipofílico molecular (MLP) e valor de pKa estimado de Hit 960 (a), gráfico de distribuição de porções químicas em função do pH (b), radar espacial de drogas (c), alinhamento entre TPSA e logP para estimativa de segurança (d) e predição do local de metabolismo e descritores farmacocinéticos (e).



Fonte: Autor (2023)

4. CONCLUSÃO

O estudo foi conduzido por meio do método inovador de design de medicamentos auxiliado pelo método computacional *De novo* na busca de novos candidatos a medicamentos para o tratamento do superfungo *Candida auris*. Assim, desempenhou-se um papel importante no desenvolvimento de novos medicamentos, o que foi significativo porque, em meio à limitação de recursos, acelerou o processo de desenvolvimento de medicamentos, reduzindo o tempo e os custos adicionais da triagem tradicional. Alguns novos derivados estruturais de (S)-piperidin-2-ilmetanol foram desenhados e estudados por meio de docking molecular gerando 1.722 hits. O modo de ligação dos compostos propostos com a proteína alvo foi avaliado, e os dados dos estudos de encaixe explicaram que alguns análogos recém-projetados tinham uma afinidade significativamente alta para a proteína alvo em comparação com o FK-506 como um ligante de referência.

O composto com maior valor de bioafinidade foi o Hit 960, que após ser analisado tanto pela energia de ligação quanto pela acessibilidade sintética, provando ser o inibidor mais potente nesta série de estudos *in silico* com uma energia de -9,12 kcal/mol. Sua viabilidade sintética foi próxima a 75%, além de apresentar menor estabilidade na análise de dinâmica molecular em estudos de RMSD, RMSF e SASA, em relação ao fármaco de referência FK-506, com afinidade de ligação de -7,3 kcal/mol.

A partir do estudo *in silico* ADMET, foi possível observar que o Hit 960 ocupa um espaço físico-químico onde residem os compostos químicos terapêuticos registrados em patentes nos últimos anos, bem como um alinhamento entre os atributos farmacocinéticos de permeabilidade, depuração e segurança. Portanto, sugere-se que este composto seja utilizado em ensaios clínicos para testar sua eficácia para benefícios sociais como padrão para projetos futuros, otimização e pesquisa na produção de análogos mais eficazes.

5. REFERÊNCIAS

AGATONOVIC-KUSTRIN, S.; BERESFORD, R. Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. **Journal of Pharmaceutical and Biomedical Analysis**, v. 22, n. 5, p. 717–727, 2000.

ARSHIA, A. H. et al. De novo design of novel protease inhibitor candidates in the treatment of SARS-CoV-2 using deep learning, docking, and molecular dynamic simulations. **Computers in Biology and Medicine**, v. 139, p. 104967, 1 dez. 2021.

BAI, Q. et al. MolAICal: A soft tool for 3D drug design of protein targets by artificial intelligence and classical algorithm. **Briefings in Bioinformatics**, v. 22, n. 3, p. 1–12, 2021.

BASHIR, Q. et al. Crystal structure and transient dimerization for the FKBP12 protein from the pathogenic fungus *Candida auris*. **Biochemical and Biophysical Research Communications**, v. 525, n. 4, p. 1103–1108, 14 maio 2020.

BIOVIA. **Dassault Systemes BIOVIA, Discovery Studio Modelling Environment, Release 4.5. Accelrys Software Inc.** San Diego, 2015.

BIOVIA, D. S. **Discovery Studio Modeling Environment, Release 2017, San Diego. Dassault Systèmes**, 2016.

BRENK, R. et al. Lessons learnt from assembling screening libraries for drug discovery for neglected diseases. **ChemMedChem**, v. 3, n. 3, p. 435–444, 2008.

CHEN, F. et al. Assessing the performance of MM/PBSA and MM/GBSA methods. 8. Predicting binding free energies and poses of protein-RNA complexes. **RNA**, v. 24, n. 9, p. 1183–1194, 2018.

DIAS, L. C. et al. Doenças tropicais negligenciadas: uma nova era de desafios e oportunidades. **Química Nova**, v. 36, n. 10, p. 1552–1556, 2013.

DIEZ, M. et al. Insights into mechanism kinematics for protein motion simulation. **BMC Bioinformatics**, v. 15, n. 1, 2014.

DONG, Y. W. et al. Structural flexibility and protein adaptation to temperature: Molecular dynamics analysis of malate dehydrogenases of marine molluscs. **Proceedings of the National Academy of Sciences of the United States of America**, v. 115, n. 6, p. 1274–1279, 2018.

DUCH, W.; SWAMINATHAN, K.; MELLER, J. Artificial Intelligence Approaches for Rational Drug Design and Discovery. **Current Pharmaceutical Design**, v. 13, n. 14, p. 1497–1508, 2007.

ERTL, P.; SCHUFFENHAUER, A. Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. **Journal of Cheminformatics**, v. 1, n. 1, 2009.

FARAGO, O. Langevin thermostat for robust configurational and kinetic sampling. **Physica A: Statistical Mechanics and its Applications**, v. 534, n. 122210, 2019.

FOUNDATION, T. A. S. **Apache Cordova**. The Apache Software Foundation, , 2012.

FRICKER, P. C.; GASTREICH, M.; RAREY, M. Automated drawing of structural molecular formulas under constraints. **Journal of Chemical Information and**

Computer Sciences, v. 44, n. 3, p. 1065–1078, 2004.

GENHEDEN, S.; RYDE, U. **The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. Expert Opinion on Drug Discovery**, 2015.

GUPTA, R. et al. Artificial intelligence to deep learning: machine intelligence approach for drug discovery. **Molecular Diversity**, v. 25, n. 3, p. 1315–1360, 1 ago. 2021.

GURUNG, A. B. et al. Molecular docking of the anticancer bioactive compound proceraside with macromolecules involved in the cell cycle and DNA replication. **Genetics and Molecular Research**, v. 15, n. 2, p. 1–8, 2016.

HASSANZADEH, P.; ATYABI, F.; DINARVAND, R. **The significance of artificial intelligence in drug delivery system design. Advanced Drug Delivery Reviews** Elsevier, , 1 nov. 2019.

HEVENER, K. 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, 2009.

HOU, T. et al. Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. **Journal of Chemical Information and Modeling**, v. 51, n. 1, p. 69–82, 2011.

HUANG, J. et al. CHARMM36m: An improved force field for folded and intrinsically disordered proteins. **Nature Methods**, v. 14, n. 1, p. 71–73, 2016.

HUGHES, J. D. et al. Physicochemical drug properties associated with in vivo toxicological outcomes. **Bioorganic and Medicinal Chemistry Letters**, v. 18, n. 6, p. 435–444, 2008.

HUGHES, T. B.; MILLER, G. P.; SWAMIDASS, S. J. Modeling epoxidation of drug-like molecules with a deep machine learning network. **ACS Central Science**, v. 1, n. 4, p. 168–180, 2015.

HUMPHREY, W.; DALKE, A.; SCHULTEN, K. VMD: Visual molecular dynamics. **Journal of Molecular Graphics**, v. 14, n. 1, p. 33–38, 1996.

IVANENKOV, Y. A.; ZAGRIBELNYY, B. A.; ALADINSKIY, V. A. Are We Opening the Door to a New Era of Medicinal Chemistry or Being Collapsed to a Chemical Singularity? **Journal of Medicinal Chemistry**, v. 62, n. 22, p. 10026–10043, 2019.

JAIN, S. K.; AGRAWAL, A. **De novo drug design: An overview. Indian Journal of Pharmaceutical Sciences**, 2004.

JO, S. et al. CHARMM-GUI Ligand Binder for Absolute Binding Free Energy Calculations and Its Application. **Journal of Chemical Information and Modeling**, v. 53, n. 1, p. 267–277, jan. 2013.

JOHNSON, T. W.; DRESS, K. R.; EDWARDS, M. Using the Golden Triangle to optimize clearance and oral absorption. **Bioorganic and Medicinal Chemistry Letters**, v. 19, n. 19, p. 5560–5564, 2009.

KATO, K. et al. Molecular dynamics simulations for the protein–ligand complex structures obtained by computational docking studies using implicit or explicit solvents. **Chemical Physics Letters**, v. 781, n. 139022, 2021.

- LEE, W. G. et al. First three reported cases of nosocomial fungemia caused by *Candida auris*. **Journal of Clinical Microbiology**, v. 49, n. 9, p. 3139–3142, 2011.
- LIMA, A. H. et al. Molecular Modeling of *T. rangeli*, *T. brucei gambiense*, and *T. evansi* Sialidases in Complex with the DANA Inhibitor. **Chemical Biology and Drug Design**, v. 80, n. 1, p. 114–120, 2012.
- MAZOLA, Y. et al. A comparative molecular dynamics study of thermophilic and mesophilic β -fructosidase enzymes. **Journal of Molecular Modeling**, v. 21, n. 9, p. 1–15, 2015.
- MORRIS, G. M. et al. Software news and updates AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. **Journal of Computational Chemistry**, v. 30, n. 16, p. 2785–2791, 2009.
- MOUCHLIS, V. D. et al. **Advances in de novo drug design: From conventional to machine learning methods**. **International Journal of Molecular Sciences** Multidisciplinary Digital Publishing Institute, , 7 fev. 2021. Disponível em: <<https://www.mdpi.com/1422-0067/22/4/1676/html>>. Acesso em: 15 dez. 2022
- NICOLAOU, C. A.; BROWN, N. Multi-objective optimization methods in drug design. **Drug Discovery Today: Technologies**, v. 10, n. 3, p. 1–9, 2013.
- PHILLIPS, J. C. et al. **Scalable molecular dynamics with NAMD**. **Journal of Computational Chemistry**, 2005.
- PIRES, D. E. V.; KAMINSKAS, L. M.; ASCHER, D. B. Prediction and optimization of pharmacokinetic and toxicity properties of the ligand. **Methods in Molecular Biology**, v. 1762, p. 271–284, 2018.
- RASTELLI, G. et al. Binding estimation after refinement, a new automated procedure for the refinement and rescoring of docked ligands in virtual screening. **Chemical Biology and Drug Design**, v. 73, n. 3, p. 283–286, 2009.
- RHODES, J.; FISHER, M. C. **Global epidemiology of emerging *Candida auris***. **Current Opinion in Microbiology** Elsevier Ltd, , 1 dez. 2019.
- ROCHA, M. N. et al. Predictive ADMET study of rhodanine-3-acetic acid chalcone derivatives. **Journal of the Indian Chemical Society**, v. 99, Issue 7, 2022.
- SCHELENZ, S. et al. First hospital outbreak of the globally emerging *Candida auris* in a European hospital. **Antimicrobial Resistance and Infection Control**, v. 5, n. 1, p. 1–7, 2016.
- SCOTTI, L. et al. SAR, QSAR and Docking of Anticancer Flavonoids and Variants: A Review. **Current Topics in Medicinal Chemistry**, v. 12, n. 24, p. 2785–2809, 2013.
- STIERAND, K.; MAASS, P. C.; RAREY, M. Molecular complexes at a glance: Automated generation of two-dimensional complex diagrams. **Bioinformatics**, v. 22, n. 14, p. 1710–1716, 2006.
- TROTT, O.; OLSON, A. J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. **Journal of Computational Chemistry**, v. 31, n. 2, p. 455–461, 2010.
- VANOMMESLAEGHE, K. et al. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields.

Journal of Computational Chemistry, v. 31, n. 4, p. 671–690, 2010.

VIEIRA, T. F.; SOUSA, S. F. Comparing AutoDock and Vina in ligand/decoy discrimination for virtual screening. **Applied Sciences (Switzerland)**, v. 9, n. 21, p. 1–18, 2019.

WAGER, T. T. et al. Moving beyond rules: The development of a central nervous system multiparameter optimization (CNS MPO) approach to enable alignment of druglike properties. **ACS Chemical Neuroscience**, v. 1, n. 6, p. 435–449, 2010.

WAGER, T. T. et al. Central Nervous System Multiparameter Optimization Desirability: Application in Drug Discovery. **ACS Chemical Neuroscience**, v. 7, n. 6, p. 767–775, 2016.

WANG, C. et al. **Recent developments and applications of the MMPBSA method. Frontiers in Molecular Biosciences**, 2018.

WANG, M. et al. **Deep learning approaches for de novo drug design: An overview. Current Opinion in Structural Biology**, 2022.

WANG, Z. et al. Comprehensive evaluation of ten docking programs on a diverse set of protein-ligand complexes: The prediction accuracy of sampling power and scoring power. **Physical Chemistry Chemical Physics**, v. 18, n. 18, p. 12964–12975, 2016.

XU, C. et al. In silico prediction of chemical ames mutagenicity. **Journal of Chemical Information and Modeling**, v. 52, n. 11, p. 2840–2847, 2012.

YU, K. et al. High daily dose and being a substrate of cytochrome P450 enzymes are two important predictors of drug-induced liver injury. **Drug Metabolism and Disposition**, v. 42, n. 4, p. 744–750, 2014.

YUSUF, D. et al. An alternative method for the evaluation of docking performance: RSR vs RMSD. **Journal of Chemical Information and Modeling**, v. 48, n. 7, p. 1411–1422, 2008.

ZHENG, M. et al. Site of metabolism prediction for six biotransformations mediated by cytochromes P450. **Bioinformatics**, v. 25, n. 10, p. 1251–1258, 2009.

6. APÊNDICE

Apêndice 1: Comprovante de submissão do artigo do autor

19/01/2023 18:04 Review : Submission : Journal of Computer-Aided Molecular Design : springer.com

 **SNAPP** Journal of Computer-Aided Molecular Design
[My account](#) ▾

Review

The search for new efficient *Candida auris* inhibitors through developed drug design New drug design method (DNDD) aided by artificial intelligence

- [1. Files](#)
- [2. Details](#)
- [3. Authors](#)
- [4. Declarations](#)
- [5. Review](#)

Saved

Review

Files

[Edit](#)

1 files added

[Manuscript AMF_23_manuscript.docx](#)

<https://submission.springernature.com/submission/a3da4f04-7f11-458e-b2f7-fce826edc7ac/review> 1/4

Apêndice 2: Comprovante de publicação de artigo do autor



Journal of Biomolecular Structure and Dynamics

ISSN: (Print) (Online) Journal homepage: <https://www.tandfonline.com/loi/tbsd20>

The search for new efficient inhibitors of SARS-COV-2 through the *De novo* drug design developed by artificial intelligence

Aluisio Marques da Fonseca, Sadrack Queque Cabongo, Bernardino Joaquim Caluaco, Regilany Paulo Colares, Carla Freire Celedonio Fernandes, Hécio Silva dos Santos, Pedro de Lima-Neto & Emmanuel Silva Marinho

To cite this article: Aluisio Marques da Fonseca, Sadrack Queque Cabongo, Bernardino Joaquim Caluaco, Regilany Paulo Colares, Carla Freire Celedonio Fernandes, Hécio Silva dos Santos, Pedro de Lima-Neto & Emmanuel Silva Marinho (2022): The search for new efficient inhibitors of SARS-COV-2 through the *De novo* drug design developed by artificial intelligence, Journal of Biomolecular Structure and Dynamics, DOI: [10.1080/07391102.2022.2148128](https://doi.org/10.1080/07391102.2022.2148128)

To link to this article: <https://doi.org/10.1080/07391102.2022.2148128>

 Published online: 24 Nov 2022.

 Submit your article to this journal [↗](#)

 View related articles [↗](#)

 View Crossmark data [↗](#)

Full Terms & Conditions of access and use can be found at
<https://www.tandfonline.com/action/journalInformation?journalCode=tbsd20>

Apêndice 3: Certificado de atuação como bolsista do Programa Residência Pedagógica RP/CAPES



Ministério da Educação - MEC
Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES
Setor Bancário Norte, Quadra 2, Bloco L, Lote 06
CEP 70.040-031 - Brasília, DF

CERTIDÃO

Declaro, para os devidos fins, que o(a) Sr(a). **SADRACK QUEQUE CABONGO**, CPF nº **093.555.831-41**, atuou no **RESIDÊNCIA PEDAGÓGICA (RESIDENCIA)**, na modalidade **Residente**, da Universidade/Instituto **UNIVERSIDADE DA INTEGRAÇÃO INTERNACIONAL DA LUSOFONIA AFRO-BRASILEIRA**, vinculado ao projeto/subprojeto **UNILAB - Química - 13231** no período de **Novembro de 2020 a Abril de 2022**.

As bolsas do **RESIDÊNCIA PEDAGÓGICA** são isentas de imposto de renda de acordo com a legislação vigente (Lei nº 9.250, de 26 de dezembro de 1995, art. 26; Decreto nº 3.000, de 26 de março de 1999 - Regulamento do Imposto sobre a Renda (RIR/1999), art. 39, inciso VII; Instrução Normativa SRF nº 15, de 2001, art. 5º, inciso XVII), não caracterizando vínculo empregatício e não cabendo recolhimento de contribuição previdenciária.

Brasília, 5 de Fevereiro de 2023.

A fim de validar a autenticação deste documento, por favor acesse <http://validadocumentos.capes.gov.br/> e informe o seguinte código: **29xPynwxLAOK**

Apêndice 4: Outras produções científicas do autor



VII SEMANA UNIVERSITÁRIA
2024 - 2025

**DESAFIOS DA UNIVERSIDADE
EM TEMPOS DE ISOLAMENTO SOCIAL**

**MEMÓRIAS DE FORMAÇÃO: UM OLHAR SOBRE A CONSTRUÇÃO DA
IDENTIDADE PROFISSIONAL DOS DOCENTES DO ICEN/UNILAB**

**Bernardino Joaquim Caluaco¹, Sadrack Queque Cabongo², Silvério Ferreira da Silva Filho³,
Regilany Paulo Colares⁴**

Resumo: O estudo do tema proposto no presente trabalho surgiu da necessidade de reacender nos subprojetos do Programa Residência Pedagógica (PRP) as discussões sobre memórias de formação e saberes da experiência. Objetivou-se identificar aspectos da vida e da trajetória que contribuíram para formação da identidade profissional dos docentes investigados. O estudo foi desenvolvido numa abordagem qualitativa, através da revisão da literatura sobre a temática e análise dos relatos autobiográficos de três docentes do Instituto de Ciências Exatas e da Natureza - ICEN/UNILAB, compreendendo as áreas: Química, Física e pedagógicas. Após o estudo, verificou-se que aspectos da vida dos professores do ICEN como o nível de identificação/afinidade com a profissão docente desde muito cedo, os saberes da experiência adquiridos em outros contextos da vida (não apenas acadêmico), dificuldades e facilidades encontradas na relação formação-trabalho-vida, os desafios da profissão e a constante busca pela experiência, têm influenciado diretamente na construção das suas identidades profissionais.

Palavras-chave: Memórias de formação. Saberes da experiência. Identidade docente.

¹ Universidade da Integração Internacional da Lusofonia Afro-Brasileira, Instituto de Ciências Exatas e da Natureza, e-mail: bernardinocaluaco@aluno.unilab.edu.br
² Universidade da Integração Internacional da Lusofonia Afro-Brasileira, Instituto de Ciências Exatas e da Natureza, e-mail: kekecabongo@gmail.com
³ Escola Estadual de Ensino Profissionalizante Adolfo Ferreira de Sousa, e-mail: silverio_f.dasilvafilho@hotmail.com
⁴ Universidade da Integração Internacional da Lusofonia Afro-Brasileira, Instituto de Ciências Exatas e da Natureza, e-mail: regilany@unilab.edu.br





Apêndice 5: Outras produções científicas do autor



ELABORAÇÃO DE UMA VÍDEO-AULA RESOLVENDO EXERCÍCIOS DE TERMOQUÍMICA: UM RELATO DE EXPERIÊNCIA

Bernardino Joaquim Caluaco¹
 Sadrack Queque Cabongo²
 Silvério Ferreira da Silva Filho³
 Regilany Paulo Colares⁴

RESUMO

O presente relato é fruto de uma atividade enquadrada na segunda unidade do módulo II do Ambiente Virtual de Aprendizagem (AVA). A atividade consistia no planejamento, produção e aplicação de uma vídeo-aula de resolução e explicação dos exercícios aplicados como prova trimestral de química da segunda série na escola campo, EEEP Adolfo Ferreira de Sousa – Redenção/CE, com o tema termoquímica. Também envolvia a elaboração e aplicação de um formulário de avaliação contendo exercícios sobre o tema trabalhado na prova e na vídeo-aula. A escolha por relatar essa experiência deve-se ao fato de ter sido uma das mais marcantes do programa, pois, é uma das poucas que envolveu contato direto (virtualmente) dos residentes com os alunos através de uma regência. Desse modo, traçou-se como objetivo relatar a experiência, apontando as principais dificuldades, os aprendizados e refletir sobre o seu papel na consolidação do aprendizado dos alunos, para assim se estabelecer parâmetros de avaliação da atuação do subprojeto de Química do PRP/Unilab na escola-campo. A dupla de residentes resolveu e explicou as questões pares (três no total) da prova, mais uma questão extra que trabalha o aspecto da contextualização do ensino. Essa atividade despertou a reflexão de muitas questões pertinentes no processo de formação de professores, como é a questão da necessidade dos professores se familiarizarem de forma efetiva com as TDICs, a questão da contextualização no ensino e a aplicação de uma abordagem pedagógica voltada para o aluno, incentivando-o a participar de forma ativa e autônoma no seu processo formativo.

Palavras-chave: Formação de professores. Ensino contextualizado. Ensino aprendizagem. Recursos tecnológicos.

¹ Graduando do Curso de Licenciatura em Química da Universidade da Integração Internacional da Lusofonia Afro-Brasileira – CE, bernardinocaluaco@gmail.com;

² Graduando do Curso de Licenciatura em Química na Universidade da Integração Internacional da Lusofonia Afro-Brasileira – CE, kekecabongo@gmail.com;

³ Mestre pelo Curso de Ciências dos Materiais da Universidade Federal do Maranhão – MA, silverio_fdsilva@ufma.br;

⁴ Doutora pelo Curso de Química da Universidade Federal do Ceará – CE, regilany@unilab.edu.br;

Apêndice 6: Outras produções científicas do autor



Congresso Internacional de Ensino e Formação Docente

**A FORMAÇÃO DE PROFESSORES PARA UMA EDUCAÇÃO AUXILIADA
PELAS TECNOLOGIAS DA INFORMAÇÃO E COMUNICAÇÃO (TICs)**

Pedro João Samuel
Estudante do Curso de Licenciatura em Química (UNILAB)
E-mail: pis267@outlook.com

Erika Farias Silvestre
Estudante do Curso de Licenciatura em Matemática (UNILAB)
E-mail: erikasilvestre89@gmail.com

Sadrack Queque Cabongo
Estudante do Curso de Licenciatura em Química (UNILAB)
E-mail: kekecabongo@gmail.com

Vitor Hugo de Araújo Lima
Estudante do Curso de Licenciatura em Química (UNILAB)
E-mail: limavitorhugo6@gmail.com

Orientadora:
Elisângela André da Silva Costa
Professora dos cursos de licenciatura do ICEN/UNILAB
E-mail: elisangelaandre@unilab.edu.br

Resumo

O presente trabalho partiu da discussão sobre os impactos que as tecnologias da informação e comunicação (TICs) têm trazido para a sociedade, gerando mudanças de paradigmas. Assim, é importante considerar o papel que as TICs devem exercer na formação dos futuros docentes, tendo em vista ser incontestável a forma como a sociedade é influenciada pelas tecnologias em todos os contextos. Assim, objetivamos, com o presente estudo, discutir a utilização das tecnologias da informação e comunicação na atividade docente, bem como os impactos do uso dessas tecnologias na educação. Para tanto, realizamos uma breve revisão de literatura, com abordagem qualitativa, para problematizar o assunto e embasar teórica e criticamente a discussão acerca do uso das tecnologias da informação na formação dos docentes como uma ferramenta que sirva de vantagem aos professores e alunos nas escolas. Complementarmente, foi realizado um levantamento quantitativo de teses e dissertações produzidas em instituições de ensino superior brasileiras, nos últimos cinco anos, através da Biblioteca Digital de Teses e Dissertações (BDTD) para visualizarmos a presença deste debate nos meios acadêmicos. Podemos verificar que o uso das TICs na educação tem a sua vantagem a partir do momento em que o docente seja capacitado ao uso crítico e equilibrado das mesmas. A

