

# PREDIÇÃO DE NOVAS MOLÉCULAS SEMISSINTÉTICAS POR APRENDIZADO DE MÁQUINA PARA O TRATAMENTO DO CARCINOMA MAMÁRIO

AYSLA OLIVEIRA ARAÚJO<sup>1</sup>, MARCOS IVAN DA SILVA FELIX<sup>1</sup>, SEBASTIANA VICENTE BEZERRA<sup>2</sup>, ROBERTA JEANE BEZERRA JORGE<sup>2</sup>.

1 – Alunos Autores; 2 - Professoras Orientadoras.

## INTRODUÇÃO

O câncer de mama é uma das neoplasias mais comuns e uma das principais causas de mortalidade entre mulheres em todo o mundo. Representa aproximadamente 25% dos casos de câncer feminino e é responsável por cerca de 15% das mortes relacionadas ao câncer em mulheres (World Health Organization, 2020). A incidência do câncer de mama varia globalmente, sendo mais alta em países desenvolvidos e aumentando em países em desenvolvimento, refletindo mudanças nos padrões de estilo de vida e melhorias nos métodos de detecção precoce (Jemal et al., 2021). No Brasil, a taxa de incidência é de incacerca de 61,5 casos por 100 mil mulheres por ano, conforme dados do Instituto Nacional de Câncer (INCA, 2023).

## OBJETIVOS

Desenvolver um *script* computacional gerador de análogos de curcumina através de ferramentas de inteligência artificial e simular a interação molecular de curcuminóides inéditos com alvos relevantes do câncer de mama.

## METODOLOGIA

### 1ª FASE FORMAÇÃO DA REDE DE PROCESSOS BIOLÓGICOS



### 2ª FASE SCRIPT GERADOR DE ANÁLOGOS, DOCAGEM MOLECULAR E AVALIAÇÃO ESTATÍSTICA

#### Acessibilidade sintética: uma nova ferramenta

Adição de parâmetro de custo do reagente + Método de impressão digital 2 (FP2) (Propriedades físico-químicas, potencial de reatividade, similaridade molecular)

Preço (g ou mL)	Score
< 50\$	1
< 100\$	2
< 250\$	3
< 500\$	4
< 1000\$	5
< 2500\$	6
< 5000\$	7
< 7500\$	8
≥ 7500\$	9

Preço de reagente Banco de dados = Sigma-Aldrich  
Categorização de valores e combinação  
Desenvolvimento da ferramenta via machine learning com RDKit  
Comparação com SWISSADME  
Mean Absolute Error (MAE), Mean Squared Error (MSE), Regressão Linear, Principal Component Analysis (PCA), e Random Forest.

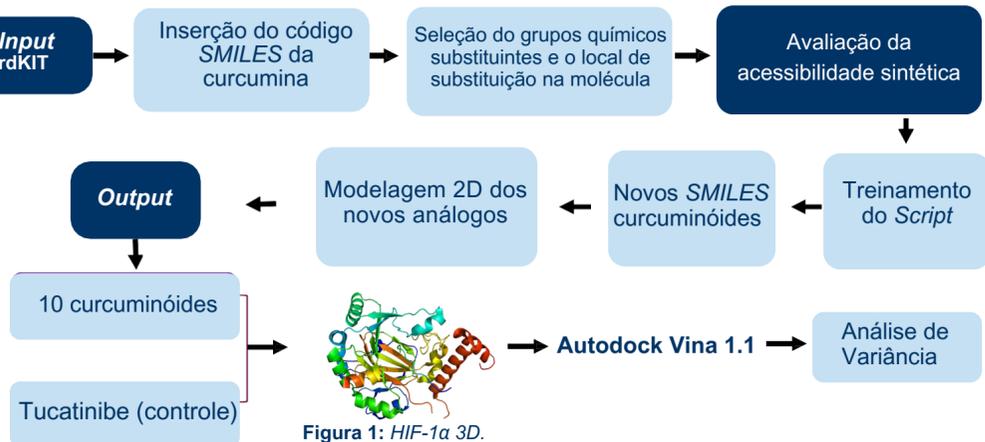
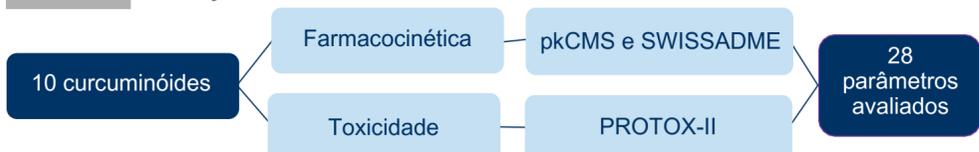
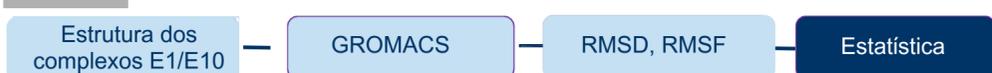


Figura 1: HIF-1α 3D.

### 3ª FASE AVALIAÇÃO FARMACOCINÉTICA E DE TOXICIDADE

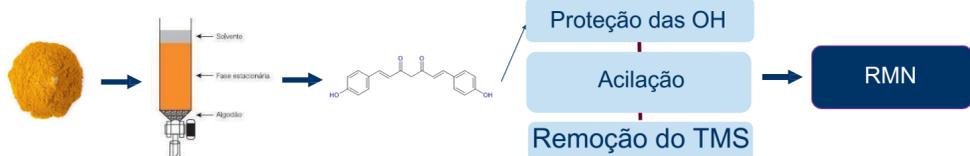


### 4ª FASE DINÂMICA MOLECULAR



### 5ª FASE SÍNTESE QUÍMICA.

Figura 2: Metodologia para de obtenção da bidimetilcurcumina e síntese dos curcuminóides.



\*Todas as imagens sem fonte são de autoria dos finalistas.

## RESULTADOS E DISCUSSÃO

Figura 3: interação proteína-proteína. (A) Interactoma das proteínas NAFSA PIP KRT20 C17orf97 EGFR GPER1 TP53 PGR PIK3R3 ERBB2 ESR1 ARNT NCOA3 HIF1A HEATR6 NCOA1. (B) Diagrama de venn dos genes atrelados aos carcinomas mamários.

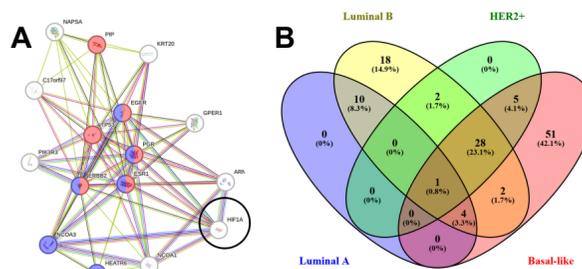


Figura 4: Estruturas bidimensionais da curcumina e seus 10 análogos modelados

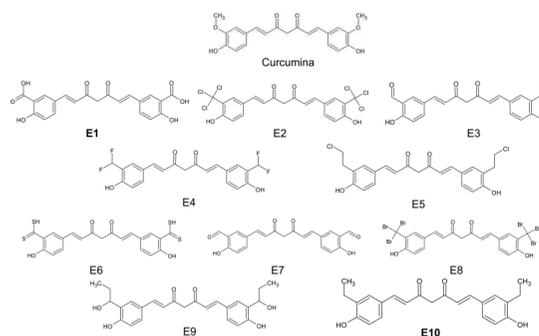


Figura 6: gráfico de afinidade de ligação entre os ligantes testados e o alvo HIF1A.

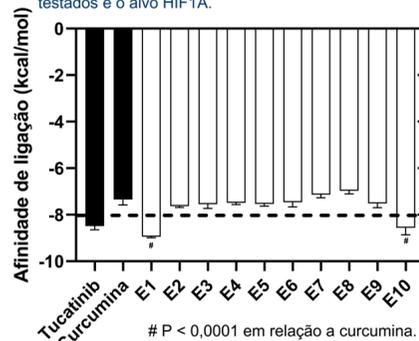


Figura 5: Regressão linear de acessibilidade sintética do treinamento de IA com o preditor SwissADME.

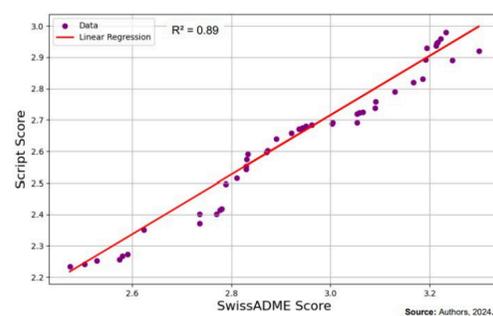
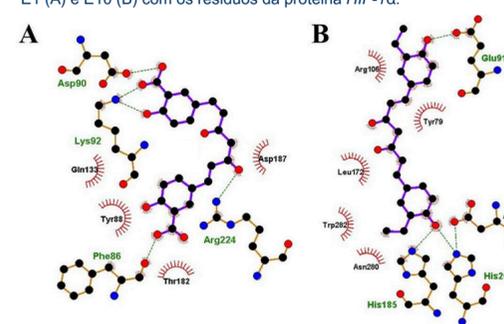


Figura 7: Representação bidimensional das interações entre os ligantes E1 (A) e E10 (B) com os resíduos da proteína HIF-1α.



Os testes de dinâmica molecular realizados para HIF-1α/E1 e HIF-1α/E10 em ambiente celular, utilizando RMSD, RMSF, SASA e ROG, mostraram que ambos mantêm sua conformidade estrutural ao longo do tempo.

Quadro 1: avaliação computacional de toxicidade dos análogos de curcumina.

DL <sub>50</sub>	Classe de toxicidade prevista	Acurácia de predição	Hepatotoxicidade	Carcinogenicidade	Imunotoxicidade	Mutagenicidade	Citotoxicidade
E1 2100 mg/kg	5	68,07%	Inativo	Inativo	Inativo	Inativo	Inativo
E2 2300 mg/kg	5	68,07%	Inativo	Inativo	Inativo	Inativo	Inativo
E3 687 mg/kg	4	68,07%	Inativo	Inativo	Ativo	Inativo	Inativo
E4 687 mg/kg	4	68,07%	Inativo	Inativo	Ativo	Inativo	Inativo
E5 2000 mg/kg	4	68,07%	Inativo	Inativo	Ativo	Inativo	Inativo
E6 2000 mg/kg	4	100%	Inativo	Inativo	Ativo	Inativo	Inativo
E7 2000 mg/kg	4	100%	Inativo	Inativo	Ativo	Inativo	Inativo
E8 2500 mg/kg	5	69,26%	Inativo	Inativo	Ativo	Inativo	Inativo
E9 2100 mg/kg	5	91,67%	Inativo	Inativo	Inativo	Inativo	Inativo
E10 2100 mg/kg	5	98,12%	Inativo	Inativo	Inativo	Inativo	Inativo

Figura 8: Comparação das trajetórias de RMSD (A), RMSF/E1 (B), RMSF/E10 (C) dos complexos com a proteína HIF-1α.

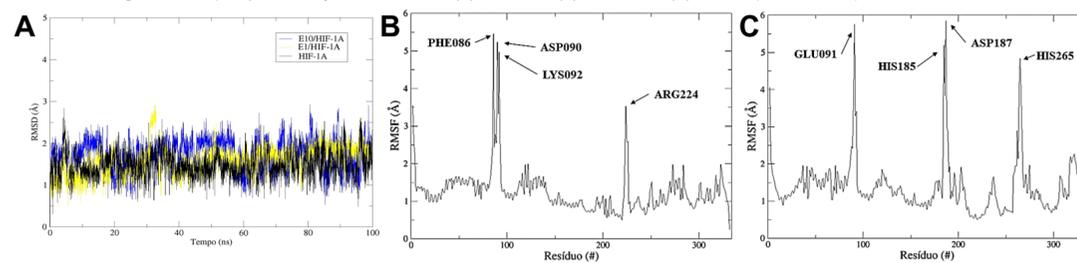


Figura 9: Curcumina, E1 e E10 sintetizados.



## CONCLUSÃO

O estudo proposto permitiu identificar análogos de curcumina como possíveis agentes terapêuticos para o carcinoma mamário. As moléculas E1 e E10 apresentaram resultados promissores, demonstrando uma forte interação com a proteína HIF-1A e baixo risco toxicológico. Neste estudo, também foi possível observar o desenvolvimento de uma metodologia e ferramentas que podem ser úteis para a projeção de diversos fármacos para o tratamento de diversas doenças.

### REFERÊNCIAS

INCA. Estimativa 2023: Incidência de Câncer no Brasil. Instituto Nacional de Câncer José Alencar Gomes da Silva, 2023. JEMAL, A. et al. Global cancer statistics. CA: A Cancer Journal for Clinicians, v. 61, n. 2, p. 69-90, 2021. WORLD HEALTH ORGANIZATION, 2020

Agradecimentos:

