

DESIGN DE NOVOS PEPTÍDEOS ANTICÂNCER USANDO DEEP LEARNING

ISADORA LEITZKE GUIDOTTI¹; FREDERICO SCHMITT KREMER²

¹Universidade Federal de Pelotas – leitzke.gi@gmail.com

²Universidade Federal de Pelotas – fred.s.kremer@gmail.com

1. INTRODUÇÃO

Foi estimado pela *World cancer research fund international* (WCRF) que em 2020 haviam cerca de 18.1 milhões de pessoas diagnosticadas com câncer no mundo, sendo a maior parte nos homens, dado que corresponde a cerca de 9.3 milhões de casos, e cerca de 8.8 milhões de mulheres diagnosticadas com a doença (WCRF INTERNATIONAL, 2022). Atualmente os tratamentos para câncer como a quimioterapia e radioterapia deixam a desejar, já que causam muitos efeitos colaterais prejudicando a qualidade de vida do paciente, além de que muitos pacientes estão apresentando resistência a quimioterapia (BUKOWSKI; KCIUK; KONTEK, 2020), por isso se faz necessário o investimento em terapias biotecnológicas como a imunoterapia e a terapia alvo. Dentro dessas, os peptídeos anticâncer têm sido amplamente utilizados para o desenvolvimento de novas terapias para câncer (ZHANG et al., 2023).

Peptídeos anticâncer (ACPs) são pequenos peptídeos formados de 10 - 60 aminoácidos, fáceis de modificar, específicos e com uma boa capacidade de penetração na célula. São classificados de acordo com a sua estrutura, podendo ser divididos em peptídeos α -helicoidal; folha β -plissada; *random coil* e cíclico (VALENTI et al., 2022). Os peptídeos α -helicoidal são os mais comuns, mais simples e os mais curtos, mas por outro lado eles podem apresentar certa toxicidade e efeitos colaterais nas células saudáveis, esse tipo de peptídeo é encontrado na pele de anfíbios (XIE; LIU; YANG, 2020). Os peptídeos de folha β -plissada tem uma atividade antitumoral menor que os α -helicoidal, porém são menos tóxicos às células saudáveis. Os peptídeos de *random coil*, geralmente são ricos em prolina e glicina e não possuem uma estrutura secundária típica (HWANG et al., 2022). Os peptídeos cíclicos são os mais estáveis por possuírem uma ligação de dissulfeto (ZHANG; CHEN, 2022).

Deep Learning (DL) é uma classe dentro do *Machine Learning* (ML) que visa simular o funcionamento de redes de neurônios profundas (como no cérebro humano) através das *Artificial Neural Network* (ANN). Essas ANN são formadas por um input e um output e entre eles camadas escondidas, as quais possuem pesos que passam por cálculos lineares e não lineares (JING et al., 2018). Dentro do DL existe uma arquitetura chamada *transformers* que é baseada em codificador e um decodificador, isto é, é um modelo baseado em contexto, que leva em consideração outras informações, no caso dos peptídeos anticâncer, ele pode basear a montagem de uma nova sequência de peptídeos anticâncer, com base nos aminoácidos que compõem peptídeos anticâncer já descobertos e usados para a terapia do câncer (CHANDRA et al., 2023).

Dessa forma o objetivo do presente trabalho é desenvolver uma ferramenta usando transformers que possa criar novos peptídeos anticâncer. Além de conseguir prever se um peptídeo é ou não anticâncer.

2. METODOLOGIA

Para desenvolver a ferramenta foi feito o uso do *framework tensorflow*, e para o treino foi utilizado os dados do banco de dados ANTICP2.0 (AGRAWAL et al., 2021). Para avaliar qual arquitetura tinha o melhor resultado foram usadas as métricas de acurácia, *recall*, precisão e *f1 score*, os dados foram treinados com algoritmos como o *transformers*, LSTM e redes convolucionais. Após a construção da ferramenta foram usadas ferramentas de predição ANTICP2.0, ACPred (SCHADUANGRAT et al., 2019) e MLACP2.0 (THI PHAN et al., 2022), para avaliar se os peptídeos gerados. Além disso, um classificador baseado em Transformer também foi treinado para filtrar os peptídeos gerados, adotando-se um critério de pelo menos 95% de probabilidade. Por fim, para prever a estrutura dos peptídeos gerados foi utilizado o PEPFOLD4.0 (REY et al., 2023).

3. RESULTADOS E DISCUSSÃO

Durante a etapa de avaliação dos algoritmos, o modelo de transformers foi o que apresentou o melhor resultado (tabela 1). Após decidir qual o melhor algoritmo foi construir a ferramenta que recebeu o nome TACaPe - *Transformed-based Anti-Cancer Peptide Classification and Generation*.

Tabela 1. Avaliação de algoritmos.

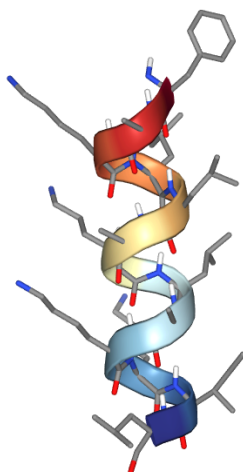
Dataset (AntiCP)	Formato	Algoritmo	Acurácia	Recall	Precisão	F1-Score	ROC AUC
alternado	tabular	LSTM	0.87	0.85	0.85	0.88	0.87
alternado	tabular	Convolutiva	0.87	0.89	0.86	0.87	0.87
alternado	tabular	Transformer	0.88	0.86	0.90	0.88	0.88

Com o TACaPe foram gerados 1000 peptídeos os quais depois de avaliados tiveram taxa de sucesso para a geração de novas sequências de 95%. Os 100 melhores peptídeos de acordo com a validação nos bancos de dados foram preditos na ferramenta PEPFOLD 4.0 (figura 1).

A geração de novas terapias envolvendo peptídeos anticâncer é importante pois essa abordagem apresenta como vantagem uma maior especificidade, um custo de produção menor e menos efeitos colaterais. Alguns autores relatam que os principais efeitos colaterais dessas terapias está envolvido diretamente com o coração, nervos (PLUMB et al., 2012), gônadas (GUTIERREZ et al., 2016), e rins (KAMISLI et al., 2015; VAN ACKER et al., 2016). Ao longo dos anos o desenvolvimento de ferramentas para a geração de novos peptídeos, algumas delas como o MLACP2.0 e o ACPred foram usadas para validar o TACaPe, cada uma dessas ferramentas utilizam diferentes abordagens, como por exemplo a

composição de aminoácidos, algoritmos de machine learning variados como o SVM. O uso desses algoritmos mais simples pode gerar alguns problemas como baixa precisão de reconhecimento, capacidade de generalização insuficiente e falta de avaliação de características e modelos de previsão em larga escala, e por isso o uso de transformers no TACaPe, torna a ferramenta confiável (WU et al., 2022).

Figura 1. Predição do peptídeo de código 22.



4. CONCLUSÕES

De acordo com os dados obtidos na validação, concluímos que o TACaPe é uma ferramenta eficaz para o desenho de novos peptídeos anticâncer. Mais estudos serão realizados para melhorar a ferramenta e avaliar os peptídeos gerados para saber como eles se comportam na célula, para isso há perspectivas de testes usando *docking molecular* e dinâmica molecular, bem como testes in vitro.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- AGRAWAL, P. et al. AntiCP 2.0: an updated model for predicting anticancer peptides. **Briefings in Bioinformatics**, v. 22, n. 3, p. bbaa153, 20 maio 2021.
- BUKOWSKI, K.; KCIUK, M.; KONTEK, R. Mechanisms of Multidrug Resistance in Cancer Chemotherapy. **International Journal of Molecular Sciences**, v. 21, n. 9, p. 3233, 2 maio 2020.
- CHANDRA, A. et al. Transformer-based deep learning for predicting protein properties in the life sciences. **eLife**, v. 12, p. e82819, 18 jan. 2023.
- GUTIERREZ, K. et al. Gonadotoxic effects of busulfan in two strains of mice. **Reproductive Toxicology (Elmsford, N.Y.)**, v. 59, p. 31–39, jan. 2016.
- HWANG, J. S. et al. Development of Anticancer Peptides Using Artificial Intelligence and Combinational Therapy for Cancer Therapeutics. **Pharmaceutics**, v. 14, n. 5, p. 997, maio 2022.

JING, Y. et al. Deep Learning for Drug Design: an Artificial Intelligence Paradigm for Drug Discovery in the Big Data Era. **The AAPS journal**, v. 20, n. 3, p. 58, 30 mar. 2018.

KAMISLI, S. et al. Hesperidin protects brain and sciatic nerve tissues against cisplatin-induced oxidative, histological and electromyographical side effects in rats. **Toxicology and Industrial Health**, v. 31, n. 9, p. 841–851, set. 2015.

PLUMB, J. A. et al. Cucurbit[7]uril encapsulated cisplatin overcomes cisplatin resistance via a pharmacokinetic effect. **Metallomics**, v. 4, n. 6, p. 561–567, 1 jun. 2012.

REY, J. et al. PEP-FOLD4: a pH-dependent force field for peptide structure prediction in aqueous solution. **Nucleic Acids Research**, v. 51, n. W1, p. W432–W437, 5 jul. 2023.

SCHADUANGRAT, N. et al. ACPred: A Computational Tool for the Prediction and Analysis of Anticancer Peptides. **Molecules**, v. 24, n. 10, p. 1973, 22 maio 2019.

THI PHAN, L. et al. MLACP 2.0: An updated machine learning tool for anticancer peptide prediction. **Computational and Structural Biotechnology Journal**, v. 20, p. 4473–4480, 2 ago. 2022.

VALENTI, G. E. et al. Antimicrobial Peptides and Cationic Nanoparticles: A Broad-Spectrum Weapon to Fight Multi-Drug Resistance Not Only in Bacteria. **International Journal of Molecular Sciences**, v. 23, n. 11, p. 6108, 29 maio 2022.

VAN ACKER, T. et al. High-resolution laser ablation-inductively coupled plasma-mass spectrometry imaging of cisplatin-induced nephrotoxic side effects. **Analytica Chimica Acta**, v. 945, p. 23–30, 16 nov. 2016.

WCRF INTERNATIONAL. **Worldwide cancer data | World Cancer Research Fund International**. WCRF International, 23 mar. 2022. Disponível em: <<https://www.wcrf.org/cancer-trends/worldwide-cancer-data/>>. Acesso em: 2 set. 2023

WU, X. et al. Anticancer Peptide Prediction via Multi-Kernel CNN and Attention Model. **Frontiers in Genetics**, v. 13, p. 887894, 27 abr. 2022.

XIE, M.; LIU, D.; YANG, Y. Anti-cancer peptides: classification, mechanism of action, reconstruction and modification. **Open Biology**, v. 10, n. 7, p. 200004, 22 jul. 2020.

ZHANG, H.; CHEN, S. Cyclic peptide drugs approved in the last two decades (2001–2021). **RSC Chemical Biology**, v. 3, n. 1, p. 18–31, 5 jan. 2022.

ZHANG, Y. et al. Bioactive peptides for anticancer therapies. **Biomaterials Translational**, v. 4, n. 1, p. 5–17, 28 mar. 2023.