

Aprendizado de Profundo Auto Supervisionado para Similaridade de Padrões de Estruturas Químicas de Antibiótico: Metodologia para Auxílio no Combate a Superbactérias

Ana Carolina Borges Monteiro Padilha (CTI – Renato Archer) ana.monteiro@cti.gov.br

Rodrigo Bonacin (CTI – Renato Archer) rbonacin@cti.gov.br

Reinaldo Padilha França (CTI – Renato Archer) reinaldo.franca@cti.gov.br

Resumo

O uso excessivo de antibióticos tem desencadeado o desenvolvimento de bactérias resistentes, as quais são capazes de sobreviver e se multiplicar na presença de antibióticos, facilitando o surgimento de superbactérias bem como dificultando o tratamento de infecções. Com base nesse problema de saúde pública global, o presente estudo visa desenvolver algoritmos de inteligência artificial (IA) capazes de analisar dados referentes a antibióticos de forma a auxiliar no desenvolvimento, descobrimento ou reposicionamento de moléculas com potencial antibacteriano. Para isso foi aplicada uma rede de Aprendizado Profundo (AP) auto supervisionada, Autoencoders, para o descobrimento de padrões moleculares similares entre estruturas moleculares que compõem os medicamentos antibióticos e antimicóticos. Os resultados apontaram 5 estruturas químicas antimicóticas com similaridade com o antibiótico daptomicina, com *score* de similaridade de 0,22, juntamente com função de perda (*loss*) de 0,0034. O emprego de metodologias de AP como esse podem ser vistas como ferramentas crucias para o combate a superbactérias, bem como aliada para mitigar custos e tempo para o combate de infecções bacterianas.

Palavras-chave: Aprendizado Profundo; superbactérias, autoencoder.

1. Introdução

Dados apresentados pela Organização Mundial de Saúde (OMS) destacam que a resistência bacteriana a antibióticos é uma grave ameaça à saúde pública global, pois estima-se que, em 2050, essa resistência resulte em 10 milhões de óbitos anuais. Projeções estatísticas também indicam que as superbactérias poderão levar a óbito uma pessoa a cada três segundos, tornando a resistência bacteriana a antibióticos uma das principais causas de morte no mundo (ICTQ, 2023).

As bactérias que causam doenças em seres humanos são classificadas em dois grupos principais: bactérias patogênicas e oportunistas. Algumas das bactérias mais comuns que causam doenças em humanos incluem: *Streptococcus pneumoniae* (causa pneumonia, meningite e infecções do trato respiratório); *Escherichia coli* (responsável por infecções do trato urinário, infecções do trato digestivo e infecções do sangue), *Salmonella typhi* (febre tifóide); *Mycobacterium tuberculosis* (tuberculose), *Neisseria gonorrhoeae* (gonorreia) e *Treponema pallidum* (sífilis), entre outros (McPherson, 2021).

A resistência bacteriana é desencadeada pelo desenvolvimento de mecanismos bacterianos que possibilitam a sobrevivência e multiplicação das bactérias mesmo na presença de antibióticos. Esses mecanismos podem ser adquiridos naturalmente ou transmitidos de uma bactéria resistente para outra (Morgado, 2020).

Diversos fatores contribuem fortemente para o aumento da resistência bacteriana como: uso inadequado de antibióticos, automedicação, e falta de saneamento básico. O uso excessivo, inadequado ou a automedicação com antibióticos pode resultar em bactérias resistentes [3]. A resistência bacteriana ao uso de antibióticos ocorre por duas maneiras: mutação genética da bactéria ou transferência horizontal de genes. Algumas mutações podem ocorrer de forma natural enquanto outras são decorrentes a fatores externos ambientais, tais como exposição aos antibióticos (Burnham, 2021).

Pesquisas recentes têm apontado que até mesmo as mudanças climáticas estão relacionadas a resistência das bactérias a antibióticos, visto que alteraram os padrões de temperatura e umidade, o que pode criar condições favoráveis ao crescimento de bactérias resistentes (Morgado, 2020).

Dado o aumento da resistência aos antibióticos, no ano de 2017, a OMS informou que os agentes patogénicos *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae* *Acinetobacter baumannii*, *Pseudomonas aeruginosa* e espécies *Enterobacter spp.* (denominados pela sigla ESKAPE) foram classificados como o mais alto grau de prioridade por representarem grande ameaça para os humanos (Mancuso, 2021).

Diante dessas problemáticas, medidas devem ser adotadas para reduzir e até mesmo resolver a crise da resistência bacteriana a antibióticos, dentre elas deve-se adotar o uso racional de antibióticos, educação e conscientização da população, e investimentos em pesquisa. Em relação ao uso indiscriminado de antibióticos pela população, no Brasil em 2009, a Comissão de Constituição e Justiça e de Cidadania aprovou o Projeto de Lei 6492/06, que veta a comercialização de antibióticos humanos e veterinários sem receita médica (CRF-RJ, 2023). Infelizmente, mesmo com a aprovação dessa lei, a comercialização ilegal de antibióticos continua sendo uma realidade no território brasileiro (Aslam, 2021).

O desenvolvimento e investimento em pesquisas é uma ótima opção para a resolução dos desafios relacionados a resistência bacteriana a antibióticos. Entretanto, em setembro de 2023, a indústria farmacêutica anunciou que está abandonando as pesquisas de novos antibióticos. Esse anuncio foi feito junto ao aumento de 700 mil para 1,2 milhão o número de óbitos relacionadas às superbactérias resistentes a fármacos tradicionais. Algumas das justificativas para esse abandono residem no fato de que as bactérias estão tornando-se resistentes em um ritmo muito acelerado, bem como o fato de que o retorno financeiro nessas pesquisas não cobre os custos do seu desenvolvimento, produção e distribuição dos fármacos (Tardáguila; Reis; Dantas, 2023).

Com base nisso, o presente estudo tem como objetivo o emprego de AP para reconhecimento de padrões moleculares de antibióticos de forma a fornecer *insights* para reposicionamento de fármacos ou associação de moléculas para desenvolvimento de novos antibióticos ou associações para potencialização da ação antimicrobiana de antibióticos.

2. Metodologia

2.1 Banco de dados de moléculas

O DrugBank é uma plataforma *on-line* de acesso aberto que disponibiliza dados sobre todos os medicamentos já aprovados ou ainda sob fase de análise para aprovação de comercialização pelo FDA (*Food and Drug Administration*). Informações farmacoquímicas e farmacocinéticas das moléculas são disponibilizadas por meio de dados sobre a absorção, disponibilidade,

metabolismo, excreção, estrutura molecular, efeitos colaterais, iteração medicamentosa, entre outros. Ao todo existem 388 antibióticos cadastrados no Drugbank, sendo excluídos fármacos de uso veterinário e aqueles ainda em fase de aprovação, gerando um arquivo CSV.

Como o objetivo do estudo é realizar o reposicionamento de fármacos, com ênfase em antibióticos, foram escolhidos medicamentos que assim como os antibióticos eliminam microorganismos, como é o caso dos antimicóticos, responsáveis pelo combate de infecções fúngicas humanas. No Drugbank, estão listados 19 medicamentos. Sendo assim foram criados dois bancos de dados, um com moléculas que compõem antibióticos e outro com moléculas que compõem antimicóticos, no formato de SMILES (do inglês “*Simplified Molecular-Input Line-Entry System*”). Um exemplo é a daptomicina, antibiótico cujo SMILE é representado por CCCCCCCCC(=O)N[C@@H](CC1=CNC2=C1C=CC=C2)C(=O)N[C@H](CC(N)=O)C(=O)N[C@@H](CC(O)=O)C(=O)N[C@H]1[C@@H](C)OC(=O)[C@H](CC(=O)C2=CC=CC=C2N)NC(=O)[C@@H](NC(=O)[C@@H](CO)NC(=O)CNC(=O)[C@H](CC(O)=O)NC(=O)[C@@H](C)NC(=O)[C@H](CC(O)=O)NC(=O)[C@H](CCN)NC(=O)CNC1=O)[C@H](C)CC(O)=O (Drugbank, 2025).

2.2 Aprendizado Profundo Auto Supervisionado

Os bancos de dados de antibióticos e antifúngicos foram submetidos a AP auto supervisionado, para a solucionar do desafio de encontrar padrões moleculares semelhantes entre as estruturas químicas dos antibióticos e dos antimicóticos, visto que essa metodologia dispensa a rotulagem dos dados.

Os autocodificadores (*autoencoders*) são algoritmos que empregam redes neurais. Seu funcionamento consiste basicamente em forçar a rede a aprender os dados mais relevantes presentes na entrada (*input*). A arquitetura de um *autoencoder* é dividida em duas partes principais: (1) Codificador (*Encoder*) que é responsável por receber os dados de entrada e transformá-los em uma representação compacta de dimensão menor, chamada de espaço latente e (2) Decodificador (*Decoder*) responsável por receber essa representação compacta e tentar reconstruir os dados de entrada originais, ou o mais próximo possível deles (Chen, *et al.* 2024).

O presente estudo foi desenvolvido no Jupyter Notebook em linguagem Python. A lógica empregada para reconhecimento de estruturas químicas semelhantes entre antibióticos e antimicóticos por meio de redes auto supervisionadas do tipo *autoencoder*, conforme Figura 1.

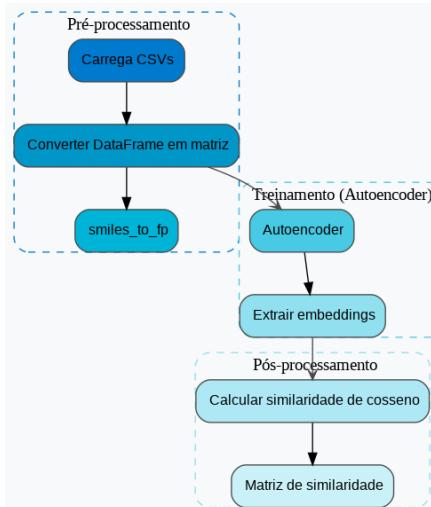


Figura 1: Lógica de reconhecimento de padrões moleculares a partir do SMILES de antibióticos e antimicóticos

3. Resultados e discussão

Diferentemente das redes de aprendizado supervisionado, os resultados de redes auto supervisionadas, como o *autoencoder*, são analisados por meio da análise da função de perda (*loss*). Essa métrica expressa a diferença entre o que o modelo recebeu como entrada (*input*) e o que ele conseguiu reconstruir na saída (*output*), ou seja, quanto menor o valor da função de perda, maior a precisão dos resultados da rede. No caso da rede *autoencoder* empregada para reconhecimento de padrões moleculares entre antibióticos e antimicóticos, a função de perda foi de 0.0034, indicando excelente performance e desempenho, conforme Figura 2, onde 286 e 94 representam a quantidade de estruturas moleculares de antibióticos e antimocóticos, respectivamente.

```
Abx shape: (286, 2048)
Amf shape: (94, 2048)
Usando device: cpu
Epoch 0, Loss: 0.0241
Epoch 5, Loss: 0.0121
Epoch 10, Loss: 0.0089
Epoch 15, Loss: 0.0070
Epoch 20, Loss: 0.0059
Epoch 25, Loss: 0.0051
Epoch 30, Loss: 0.0045
Epoch 35, Loss: 0.0042
Epoch 40, Loss: 0.0037
Epoch 45, Loss: 0.0034
Matriz de similaridade Abx x Amf: (286, 94)
```

Figura 2- Análise de função de perda da rede *autoencoder* para reconhecimento de similaridade de padrões de moleculares entre antibióticos e antimicóticos.

Para testar a similaridade molecular entre antibiótico e antimicóticos foi selecionado um antibiótico injetável denominado de daptomicina, de uso adulto e pediátrico destinado ao combate de infecções de pele e partes moles complicadas, causada pela bactéria gram-positiva *Staphylococcus aureus*. Essa escolha foi feita, pois o *S. aureus* é classificado como uma bactéria de grande risco pela OMS. Com base na lógica apresentada na Figura 1, a rede de AP auto supervisionado apontou as moléculas de *capric acid* (ácido cáprico), *undecylenic acid* (ácido undecilênico), *aminocandin*, *naftifine* (naftifina), *benzoic acid* (ácido benzoico), conforme resultados apresentados na Figura 3.

	Antibiotic	Antifungal	Similarity Score
0	Daptomycin	Capric acid	0.369569
1	Daptomycin	Undecylenic acid	0.273811
2	Daptomycin	Aminocandin	0.263886
3	Daptomycin	Naftifine	0.239026
4	Daptomycin	Benzoic acid	0.228941

Figura 3 - Estruturas moleculares de medicamentos antimicóticos com possibilidade de potencializar a ação do antibiótico daptomicina, sugeridos pela rede autoencoder por similaridade molecular.

O ácido cáprico (*capric acid*), também conhecido como *decanoic acid*, é naturalmente encontrado em óleos de coco e de palma, bem como no leite de vários mamíferos, sendo considerada uma substância capaz de combater fungos, através da supressão da sua capacidade de crescimento ou reprodução. Também foi identificado como o agente antifúngico ativo contra o fungo da doença do olmo holandês (*Ceratocystis ulmi*) e vários outros fungos (NCBI,2025).

O ácido undecilênico (*undecylenic acid*), é naturalmente encontrado no suor humano, apresentando-se na forma de um ácido graxo insaturado com uma ligação dupla terminal, sendo derivado do óleo de rícino. Ácidos graxos orgânicos como esse, exercem ações fungicidas ou fungistáticas. O undecilendato de zinco (*Zinc undecylendate*) atua como um agente fungistático, podendo também ser observada ação fungicida quando há exposição crônica em altas concentrações, sendo eficaz contra o fungo *Candida albicans* (NCBI,2025).

Aminocandina é um membro da família das equinocandinas que apresenta amplo espectro de atividade *in vitro* contra os fungos *Aspergillus* e *Candida spp*. Essa molécula pode potencialmente afetar a morfologia celular e o estado metabólico das células de *Candida albicans* em biofilmes (NCBI,2025).

A naftifina é uma amina terciária cujo nitrogênio é substituído por grupos metil, alfa-naftilmetil e (1E)-cinamil. Empregada no tratamento de infecções fúngicas da pele, sendo um agente antifúngico sintético de amplo espectro, para o tratamento tópico de *Tinea pedis*, *Tinea cruris* e *Tinea corporis*, causadas pelos organismos *Trichophyton rubrum*, *Trichophyton mentagrophytes*, *Trichophyton tonsurans* e *Epidermophyton floccosum* (NCBI,2025).

O ácido benzoico (*benzoic acid*) possui propriedade que destroem fungos, suprimindo sua capacidade de crescimento ou reprodução, pois atuam na defesa contra fungos presentes em tecidos humanos ou animais. O ácido benzoico está presente em medicamentos antifúngicos tópicos, sendo também empregado como composto fungistático amplamente utilizado como conservante de alimentos (NCBI,2025).

Com base nessas estruturas encontradas a partir da similaridade molecular detectadas a partir da comparação entre estruturas químicas de antibióticos e antimicóticos, a rede *autoencoder* desenvolvida pode ser vista como uma ferramenta crucial para o desenvolvimento ou reposicionamento de fármacos, visto que as 5 opções supracitadas podem ser utilizadas em combinações ou isoladas para a criação de um novo fármaco com potencial de combate antimicrobiano, bem como pode ser utilizado em conjunto com a daptomicina para potencialização de sua ação antimicrobiana. Esse estudo pode ser visto como ponto de partida para novas pesquisas, visto que são necessárias simulações moleculares atomísticas, testes *in vitro*, testes em modelos murinos e ensaios clínicos para conhecimento de administração, absorção, toxicidade, excreção, interações medicamentosas e efeitos colaterais, entre outros.

Estudos de simulação computacional com emprego de IA podem acelerar a descoberta de fármacos, reduzir os custos de pesquisa, mitigar o uso de testes em animais bem como diminuir o tempo para que um novo antibiótico seja lançado no mercado. Dessa forma, poderá ser empregada como uma ferramenta para o combate a superbactérias.

4.0 Conclusão

A IA apresenta tem potencial para redução de custos em pesquisas científicas para antibióticos, pois por meio dessas técnicas pode haver a aceleração do processo de descoberta, já que a IA pode ser empregada na identificação de novos compostos com potencial antibiótico de forma mais rápida e eficiente do que os métodos tradicionais. Outro benefício é a redução do número de testes, pois a IA pode ser usada para previsões e classificações mais assertivas

quanto a eficácia de um antibiótico em testes pré-clínicos. Com base nisso os ensaios clínicos podem ser melhorados por meio da redução do tempo e dos custos resultantes do processo de desenvolvimento de novos fármacos. Trabalhos futuros envolverão a descoberta de similaridade com outros antibióticos em relação a antimicóticos, bem como poderá ser ampliada para outras classes de medicamentos, tais como anti-hipertensivos, antitérmicos, entre outros, e também auxiliar no desenvolvimento de medicamentos contra patógenos classificados como de alto risco pela OMS.

Referências

- Aslam, A. et al.** Evidence of the practice of self-medication with antibiotics among the lay public in low-and middle-income countries: a scoping review. *Antibiotics*, v. 9, n. 9, p. 597, 2020
- Burnham, J.P.** Climate change and antibiotic resistance: a deadly combination. *Therapeutic Advances in Infectious Disease*, v. 8, 2021.
- Chen X. et al. Context autoencoder for self-supervised representation learning. *International Journal of Computer Vision*, v. 132, n. 1, p. 208-223, 2024.
- CRF-RJ.** Crescem denúncias de venda de antibióticos sem receita. Rio de Janeiro: Disponível em: <https://crf-rj.org.br/portal/noticias/1643-crescem-denuncias-de-venda-de-antibioticos-sem-receita.html>. Acesso em: 29 set. 2023.
- Drugbank.** *DrugBank Online*. Edmonton: University of Alberta,
- ICTQ.** OMS alerta: em 2050 superbactérias devem matar uma pessoa a cada três segundos. Disponível em: <https://ictq.com.br/farmacia-clinica/3166-oms-alerta-em-2050-superbacterias-matarao-uma-pessoa-a-cada-tres-segundos>. Acesso em: 29 set. 2023.
- Mancuso, G. et al.** Bacterial antibiotic resistance: The most critical pathogens. *Pathogens*, v. 10, 2021.
- McPherson, R.A, et al.** *Henry's clinical diagnosis and management by laboratory methods*. Elsevier, 2021.
- Morgado, A.M.M.** Causas e consequências da resistência aos antibióticos. 2020. Dissertação de Mestrado.
- National Center for Biotechnology Information (NCBI).** Aminocandin. In: PUBCHEM. Bethesda (MD): NCBI. Disponível em: <https://pubchem.ncbi.nlm.nih.gov/#query=aminocandin>. Acesso em: 23 out. 2025.
- National Center for Biotechnology Information (NCBI).** Benzoic Acid. In: PUBCHEM. Bethesda (MD): NCBI, 16 set. 2004. CID 243. Disponível em: <https://pubchem.ncbi.nlm.nih.gov/compound/243>. Acesso em: 23 out. 2025.
- National Center for Biotechnology Information (NCBI).** Clindamycin. In: PUBCHEM. Bethesda (MD): NCBI, 16 set. 2004. CID 47641. Disponível em: <https://pubchem.ncbi.nlm.nih.gov/compound/47641#section=ClinicalTrials-gov>. Acesso em: 23 out. 2025.
- National Center for Biotechnology Information (NCBI).** Decanoic Acid. In: PUBCHEM. Bethesda (MD): NCBI, 16 set. 2004. CID 2969. Disponível em: <https://pubchem.ncbi.nlm.nih.gov/compound/2969>. Acesso em: 23 out. 2025.
- National Center for Biotechnology Information (NCBI).** Undecylenic Acid. In: PUBCHEM. Bethesda (MD): NCBI, 16 set. 2004. CID 5634. Disponível em: <https://pubchem.ncbi.nlm.nih.gov/compound/5634>. Acesso em: 23 out. 2025.
- Tardáguila, C.; Reis, F.; Dantas, C.** Superbactérias: por que o mercado não cria novos antibióticos para frear resistência? Disponível em: <https://noticias.uol.com.br/saude/ultimas-noticias/redacao/2023/09/27/superbacterias-resistencia-antibioticos-industria-farmaceutica.htm>. Acesso em: 29 set. 2023.