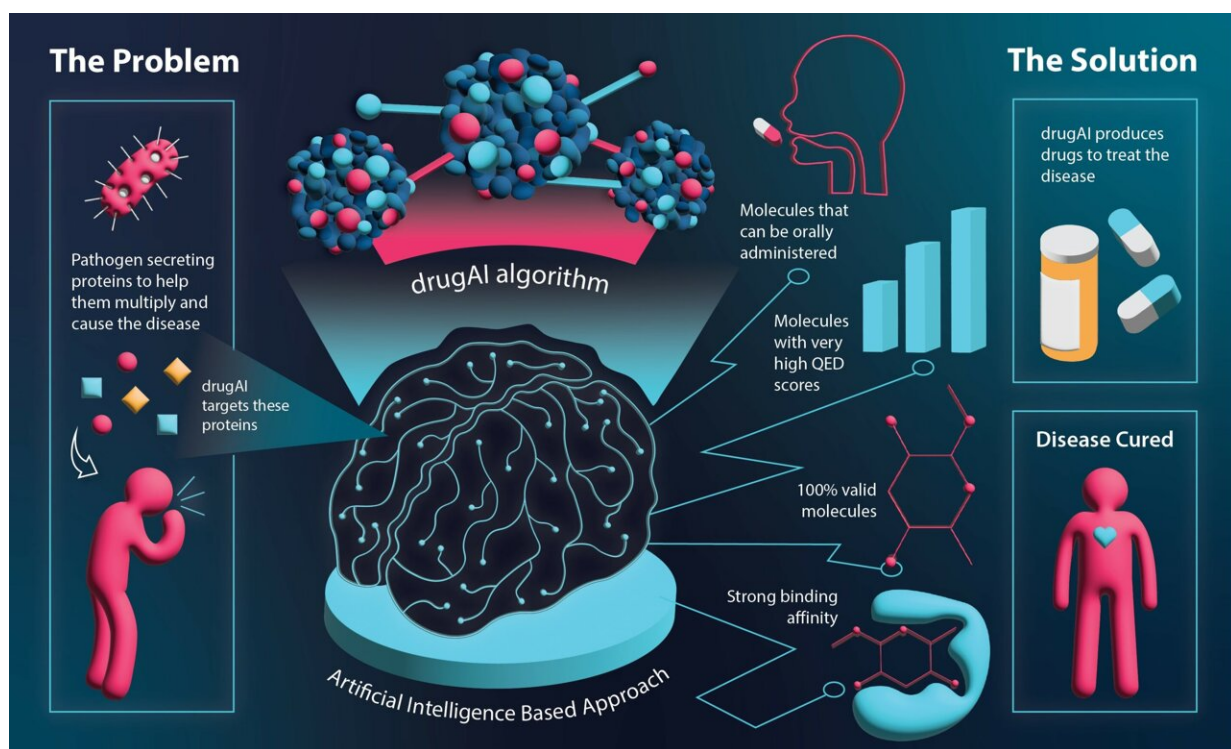


Scientists code ChatGPT to design new medicine

February 7 2024



Graphical abstract. Credit: *Pharmaceuticals* (2024). DOI: 10.3390/ph17020161

Generative artificial intelligence platforms, from ChatGPT to Midjourney, grabbed headlines in 2023. But GenAI can do more than create collaged images and help write emails—it can also design new drugs to treat disease.

Today, scientists use advanced technology to design new synthetic drug compounds with the right properties and characteristics, also known as "de novo drug design." However, current methods can be labor-, time-, and cost-intensive.

Inspired by ChatGPT's popularity and wondering if this approach could speed up the drug design process, scientists in the Schmid College of Science and Technology at Chapman University in Orange, California, decided to create their own GenAI model, detailed in a new paper, "De Novo Drug Design using Transformer-based Machine Translation and Reinforcement Learning of Adaptive Monte-Carlo Tree Search," [appearing](#) in the journal *Pharmaceuticals*.

Dony Ang, Cyril Rakovski, and Hagop Atamian coded a model to learn a massive dataset of known chemicals, how they bind to target proteins, and the rules and syntax of chemical structure and properties writ large.

The end result can generate countless unique molecular structures that follow essential chemical and biological constraints and effectively bind to their targets—promising to vastly accelerate the process of identifying viable drug candidates for a wide range of diseases, at a fraction of the cost.

To create the breakthrough model, researchers integrated two cutting-edge AI techniques for the first time in the fields of bioinformatics and cheminformatics: the well-known "Encoder-Decoder Transformer architecture" and "Reinforcement Learning via Monte Carlo Tree Search" (RL-MCTS). The platform, fittingly named "drugAI," allows users to input a target protein sequence (for instance, a protein typically involved in cancer progression).

DrugAI, trained on data from the comprehensive public database BindingDB, can generate unique molecular structures from scratch, and

then iteratively refine candidates, ensuring finalists exhibit strong binding affinities to respective drug targets—crucial for the efficacy of potential drugs. The model identifies 50-100 new molecules likely to inhibit these particular proteins.

"This approach allows us to generate a potential drug that has never been conceived of," Dr. Atamian said. "It's been tested and validated. Now, we're seeing magnificent results."

Researchers assessed the molecules drugAI generated along several criteria, and found drugAI's results were of similar quality to those from two other common methods, and in some cases, better. They found that drugAI's candidate drugs had a validity rate of 100%—meaning none of the drugs generated were present in the training set.

DrugAI's candidate drugs were also measured for drug-likeness, or the similarity of a compound's properties to those of oral drugs, and candidate drugs were at least 42% and 75% higher than other models. Plus, all drugAI-generated molecules exhibited strong binding affinities to respective targets, comparable to those identified via traditional virtual screening approaches.

Ang, Rakovski and Atamian also wanted to see how drugAI's results for a specific disease compared to existing known drugs for that disease. In a different experiment, screening methods generated a list of natural products that inhibited COVID-19 proteins; drugAI generated a list of novel drugs targeting the same protein to compare their characteristics. They compared drug-likeness and binding affinity between the natural molecules and drugAI's, and found similar measurements in both—but drugAI was able to identify these in a much quicker and less expensive way.

Plus, the scientists designed the algorithm to have a flexible structure

that allows future researchers to add new functions. "That means you're going to end up with more refined drug candidates with an even higher probability of ending up as a real drug," said Dr. Atamian. "We're excited for the possibilities moving forward."

More information: Dony Ang et al, De Novo Drug Design Using Transformer-Based Machine Translation and Reinforcement Learning of an Adaptive Monte Carlo Tree Search, *Pharmaceuticals* (2024). [DOI: 10.3390/ph17020161](https://doi.org/10.3390/ph17020161)

Provided by Chapman University

Citation: Scientists code ChatGPT to design new medicine (2024, February 7) retrieved 28 April 2024 from <https://medicalxpress.com/news/2024-02-scientists-code-chatgpt-medicine.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.