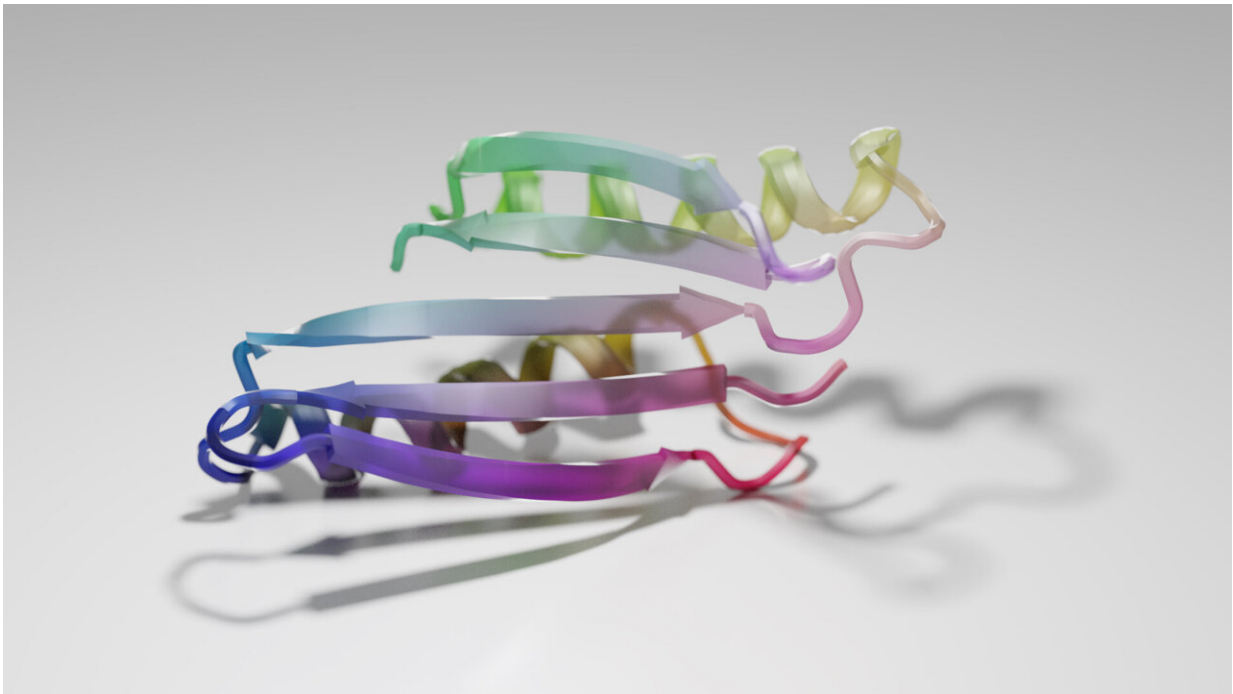


# Deep learning dreams up new protein structures

December 1 2021, by Ian C. Haydon

---



This twisting protein structure is one of hundreds dreamed up by a machine-learning algorithm. Credit: Ian C. Haydon/UW Medicine Institute for Protein Design

Just as convincing images of cats can be created using artificial intelligence, new proteins can now be made using similar tools. In a report in *Nature*, researchers describe the development of a neural network that "hallucinates" proteins with new, stable structures.

Proteins, which are string-like molecules found in every cell, spontaneously fold into intricate three-dimensional shapes. These folded shapes are key to nearly every biological process, including cellular development, DNA repair, and metabolism. But the complexity of [protein](#) shapes makes them difficult to study. Biochemists often use computers to predict how protein strings, or sequences, might fold. In recent years, [deep learning](#) has revolutionized the accuracy of this work.

"For this project, we made up completely random protein sequences and introduced mutations into them until our neural network predicted that they would fold into stable structures," said co-lead author Ivan Anishchenko, He is an acting instructor of biochemistry at the University of Washington School of Medicine and a researcher in David Baker's laboratory at the UW Medicine Institute for Protein Design.

"At no point did we guide the software toward a particular outcome," Anishchenko said, " These new proteins are just what a computer dreams up."

In the future, the team believes it should be possible to steer the [artificial intelligence](#) so that it generates new proteins with useful features.

"We'd like to use deep learning to design proteins with function, including protein-based drugs, enzymes, you name it," said co-lead author Sam Pellock, a postdoctoral scholar in the Baker lab.

The research team, which included scientists from UW Medicine, Harvard University, and Rensselaer Polytechnic Institute (RPI), generated two thousand new protein sequences that were predicted to fold. Over 100 of these were produced in the laboratory and studied. Detailed analysis on three such proteins confirmed that the shapes predicted by the computer were indeed realized in the lab.

"Our NMR [[nuclear magnetic resonance](#)] studies, along with X-ray crystal structures determined by the University of Washington team, demonstrate the remarkable accuracy of protein designs created by the hallucination approach", said co-author Theresa Ramelot, a senior research scientist at RPI in Troy, New York.

Gaetano Montelione, a co-author and professor of chemistry and chemical biology at RPI, noted. "The hallucination approach builds on observations we made together with the Baker lab revealing that protein structure prediction with deep learning can be quite accurate even for a single protein sequence with no natural relatives. The potential to hallucinate brand new proteins that bind particular biomolecules or form desired enzymatic active sites is very exciting".

"This approach greatly simplifies protein design," said senior author David Baker, a professor of biochemistry at the UW School of Medicine who received a 2021 Breakthrough Prize in Life Sciences. "Before, to create a new protein with a particular shape, people first carefully studied related structures in nature to come up with a set of rules that were then applied in the design process. New sets of rules were needed for each new type of fold. Here, by using a deep-learning network that already captures general principles of protein structure, we eliminate the need for fold-specific rules and open up the possibility of focusing on just the functional parts of a protein directly."

"Exploring how to best use this strategy for specific applications is now an active area of research, and this is where I expect the next breakthroughs," said Baker.

**More information:** David Baker, De novo protein design by deep network hallucination, *Nature* (2021). [DOI: 10.1038/s41586-021-04184-w](https://doi.org/10.1038/s41586-021-04184-w).  
[www.nature.com/articles/s41586-021-04184-w](https://www.nature.com/articles/s41586-021-04184-w)

Provided by University of Washington School of Medicine

Citation: Deep learning dreams up new protein structures (2021, December 1) retrieved 29 April 2024 from <https://phys.org/news/2021-12-deep-protein.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.