

New technique could streamline drug design

February 27 2020



Credit: CC0 Public Domain

Researchers have developed a process that could sharply reduce the work involved in computational protein design, according to a study in *Proceedings of the National Academy of Sciences*.

The technique uses 3-D structural models to project how novel combinations of molecular blocks might work together to achieve a desired effect.

The advancement, which focuses on a relatively small number of protein

substructures rather than the infinite number of atomic-level combinations, could ease the development of new medications and materials.

"When you design a building, you don't necessarily need to understand how grains of sand interact with each other within one brick," said Gevorg Grigoryan, an associate professor of computer science at Dartmouth and senior researcher on the study. "Because you know what a brick is and what its properties are, you can instead focus on how bricks come together to form the desired shape. That's the same approach we are taking. We only focus on protein sub-structures that we know work."

Proteins are the workhorse of the natural world. Proteins help us sense the world around us, digest food and form the body's natural defenses.

For years, researchers have focused on building custom proteins that can be useful in the human body. For example, custom proteins can be used to develop therapeutic drugs to fight disease. However, while many therapeutics like insulin are produced from naturally occurring proteins, the field has not advanced to allow widespread development of synthetic proteins.

Among the barriers to developing synthetic proteins is the overwhelming number of possible amino acid combinations. Sorting through combinations to find one that would be helpful in any given scenario is a time-intensive and resource-heavy process.

Researchers developing new drugs currently focus on how specific atoms interact. This approach requires labs to build large libraries of variants to find one that will complete the specified task. While this can produce useful results, researchers have found it challenging to build atomic models that have high levels of accuracy.

"The number of sequences is virtually infinite. This really complicates the process of finding a correct combination to fill a specific therapeutic need," said Jianfu Zhou, a Ph.D. student at Dartmouth who co-authored the [research paper](#).

To develop an optimized approach to protein design, the research team scanned a database of the 3-D models of 150,000 known proteins. The team discovered that a small number of structural patterns frequently recurred in proteins, and that much of the diversity in [protein structure](#) comes from how these building blocks are combined.

This basic discovery led the team to hypothesize that rather than modeling proteins as complex networks of interacting atoms, they can instead represent them much more simply as groupings of a limited set of structural building blocks.

With the new method, novel protein structures can be more easily judged against established patterns. The approach allows researchers to easily experiment with more creative designs by affording the chance to check them against a library of known structures.

"This technique takes the challenge away from getting the physics absolutely right at the atomic scale, potentially making computational protein design a much more robust process. Our findings should throw the doors for machine learning in [protein design](#) wide open," said Grigoryan.

The new process focuses on the larger blocks of atoms that occur in proteins, known as tertiary motifs, to design functioning proteins. These are recurring structural arrangements—similar to an archway or column in a building—that can be applied to designing novel proteins without regard to their atomic-level composition.

Since the structures only come together in certain ways, researchers would no longer need to do the atomic-level guesswork. Researchers only focus on the blocks that fit together, ignoring those structures that would not form a functioning protein.

According to the research paper, the results "strongly argue that the Protein Data Bank is now sufficiently large to enable proteins to be designed by using only examples of structural motifs from unrelated proteins."

By applying the new technique, the research team hopes to cut out the redundancy of rediscovering physical principles in [protein](#) structure by simply relying on those principles in the first place.

More information: Jianfu Zhou et al, A general-purpose protein design framework based on mining sequence–structure relationships in known protein structures, *Proceedings of the National Academy of Sciences* (2019). [DOI: 10.1073/pnas.1908723117](https://doi.org/10.1073/pnas.1908723117)

Provided by Dartmouth College

Citation: New technique could streamline drug design (2020, February 27) retrieved 3 May 2024 from <https://phys.org/news/2020-02-technique-drug.html>

<p>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.</p>
--