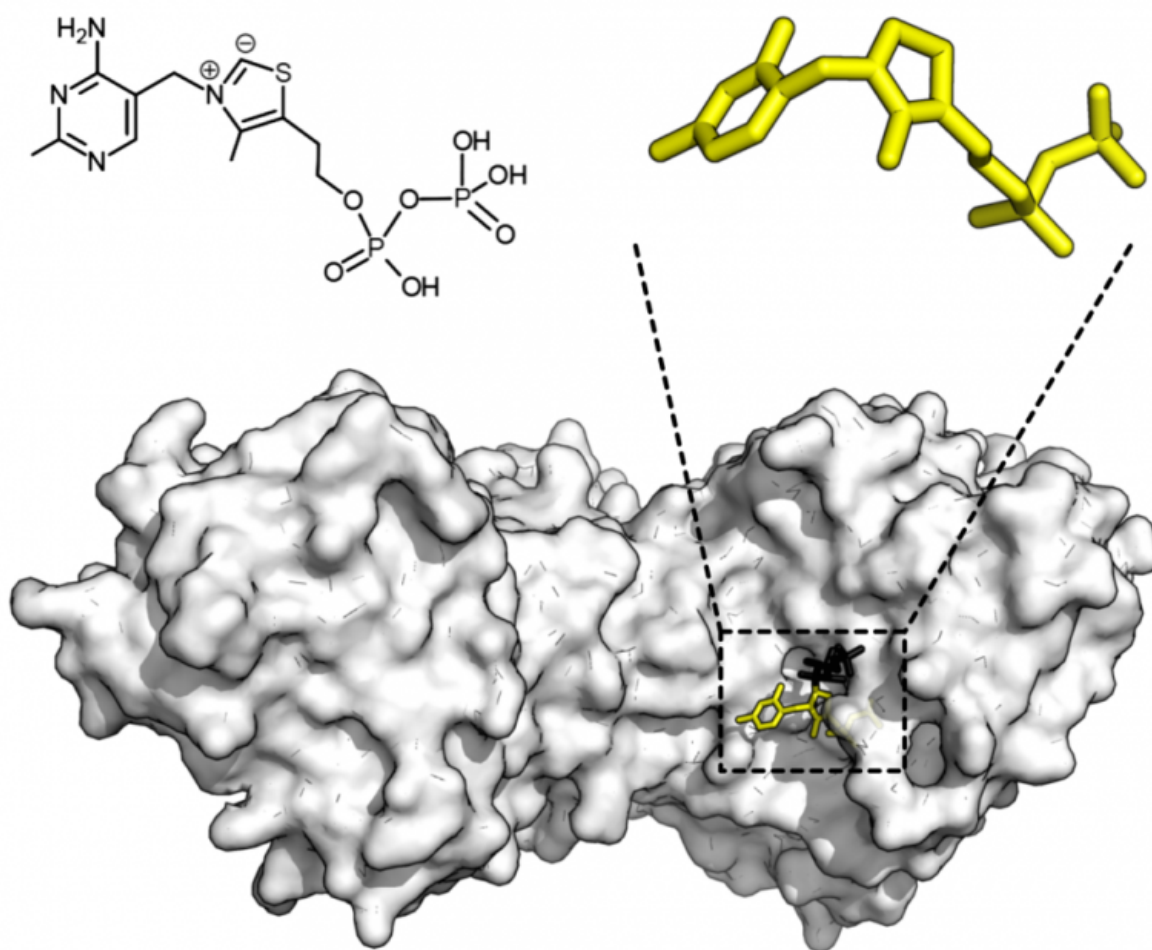


AI predicts enzyme function better than leading tools

March 30 2023



Chemical structure for thiamine pyrophosphate and protein structure of transketolase. Thiamine pyrophosphate cofactor in yellow and xylulose 5-phosphate substrate in black. Credit: Thomas Shafee/Wikipedia

A new artificial intelligence tool can predict the functions of enzymes based on their amino acid sequences, even when the enzymes are unstudied or poorly understood. The researchers said the AI tool, dubbed CLEAN, outperforms the leading state-of-the-art tools in accuracy, reliability and sensitivity. Better understanding of enzymes and their functions would be a boon for research in genomics, chemistry, industrial materials, medicine, pharmaceuticals and more.

"Just like ChatGPT uses data from written language to create predictive text, we are leveraging the language of proteins to predict their activity," said study leader Huimin Zhao, a University of Illinois Urbana-Champaign professor of chemical and biomolecular engineering.

"Almost every researcher, when working with a new protein sequence, wants to know right away what the [protein](#) does. In addition, when making chemicals for any application—biology, medicine, industry—this [tool](#) will help researchers quickly identify the proper enzymes needed for the synthesis of chemicals and materials."

The researchers will publish their findings in the journal *Science* and make CLEAN accessible online March 31.

With advances in genomics, many enzymes have been identified and sequenced, but scientists have little or no information about what those enzymes do, said Zhao, a member of the Carl R. Woese Institute for Genomic Biology at Illinois.

Other computational tools try to predict [enzyme](#) functions. Typically, they attempt to assign an enzyme commission number—an ID code that indicates what kind of reaction an enzyme catalyzes—by comparing a queried sequence with a catalog of known enzymes and finding similar sequences. However, these tools don't work as well with less-studied or uncharacterized enzymes, or with enzymes that perform multiple jobs, Zhao said.

"We are not the first one to use AI tools to predict enzyme commission numbers, but we are the first one to use this new [deep-learning algorithm](#) called contrastive learning to predict enzyme function. We find that this algorithm works much better than the AI tools that are used by others," Zhao said. "We cannot guarantee everyone's product will be correctly predicted, but we can get higher accuracy than the other two or other three methods."

The researchers verified their tool experimentally with both computational and in vitro experiments. They found that not only could the tool predict the function of previously uncharacterized enzymes, it also corrected enzymes mislabeled by the leading software and correctly identified enzymes with two or more functions.

Zhao's group is making CLEAN accessible online for other researchers seeking to characterize an enzyme or determine whether an enzyme could catalyze a desired reaction.

"We hope that this tool will be used widely by the broad research community," Zhao said. "With the web interface, researchers can just enter the sequence in a search box, like a [search engine](#), and see the results."

Zhao said the group plans to expand the AI behind CLEAN to characterize other proteins, such as binding proteins. The team also hopes to further develop the machine-learning algorithms so that a user could search for a desired reaction and the AI would point to a proper enzyme for the job.

"There are a lot of uncharacterized binding proteins, such as receptors and transcription factors. We also want to predict their functions as well," Zhao said. "We want to predict the functions of all proteins so that we can know all the proteins a cell has and better study or engineer the

whole cell for biotechnology or [biomedical applications](#)."

Zhao also is a U. of I. professor of bioengineering, of chemistry, and of biomedical and translational sciences in the Carle Illinois College of Medicine.

The paper is titled "Enzyme function prediction using contrastive learning."

More information: Tianhao Yu et al, Enzyme function prediction using contrastive learning, *Science* (2023). [DOI: 10.1126/science.adf2465](#).
www.science.org/doi/10.1126/science.adf2465

Provided by University of Illinois at Urbana-Champaign

Citation: AI predicts enzyme function better than leading tools (2023, March 30) retrieved 23 April 2024 from <https://phys.org/news/2023-03-ai-enzyme-function-tools.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.