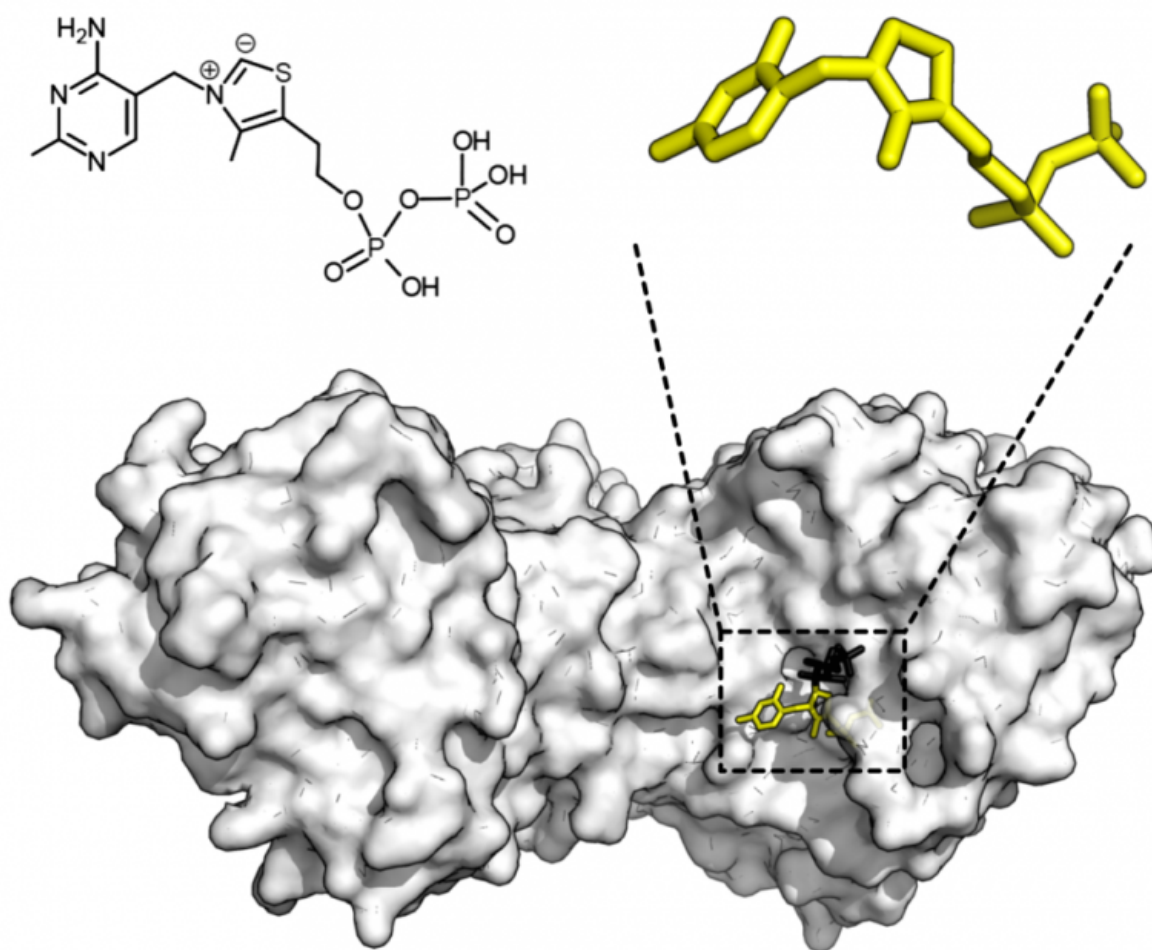# Researchers teach artificial intelligence about frustration in protein folding

August 20 2024, by Marcy de Luna

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

Scientists have found a new way to predict how proteins change their shape when they function, which is important for understanding how they work in living systems. While recent artificial intelligence (AI) technology has made it possible to predict what proteins look like in their resting state, figuring out how they move is still challenging because there is not enough direct data from experiments on protein motions to train the neural networks.

In a new study published in the *Proceedings of the National Academy of Sciences*, Rice University's Peter Wolynes and his colleagues in China combined information about protein energy landscapes with deep-learning techniques to predict these movements.

Their method improves AlphaFold2 (AF2), a tool that predicts static protein structures by teaching it to focus on "energetic frustration." Proteins have evolved to minimize energetic conflicts between their parts, so they can be funneled toward their static structure. Where conflicts persist, there is said to be frustration.

"Starting from predicted static ground-state structures, the new method generates alternative structures and pathways for protein motions by first finding and then progressively enhancing the energetic frustration features in the input multiple sequence alignment sequences that encode the protein's evolutionary development," said Wolynes, the D.R. Bullard-Welch Foundation Professor of Science and study co-author.

The researchers tested their method on the protein adenylate kinase and found that its predicted movements matched experimental data. They also successfully predicted the functional movements of other proteins that change shape significantly.

"Predicting the three-dimensional structures and motions of proteins is integral to understanding their functions and designing new drugs,"

Wolynes said.

The study also examined how AF2 works, showing that combining physical knowledge of the energy [landscape](#) with AI not only helps predict how proteins move but also explains why the AI overpredicts structural integrity, leading only to the most stable structures.

The energy landscape theory, which Wolynes and his collaborators have worked with over the decades, is a key part of this method, but recent AI codes were trained to predict only the most stable protein structures and ignore the different shapes proteins might take when they function.

The energy landscape theory suggests that while evolution has sculpted the protein's energy landscape where they can fold into their optimal structures, deviations from a perfectly funneled landscape that otherwise guides the folding, called local frustration, are essential for protein functional movements.

By pinpointing these frustrated regions, the researchers taught the AI to ignore those regions in guiding its predictions, thereby allowing the code to predict alternative protein structures and functional movements accurately.

Using a frustration analysis tool developed within the energy landscape framework, researchers identified frustrated and therefore flexible regions in proteins.

Then, by manipulating the evolutionary information in the aligned protein family sequences used by AlphaFold and in accordance with the frustration scores, the researchers taught the AI to recognize these frustrated regions, enabling accurate predictions of alternative structures and pathways between them, said Wolynes.

"This research underscores the significance of not forgetting or abandoning physics-based methods in the post-AlphaFold era, where the emphasis has been on agnostic learning from experimental data without any theoretical input," Wolynes said. "Integrating AI with biophysical insights will significantly impact future practical applications, including drug design, enzyme engineering and understanding disease mechanisms."

Other authors include Xingyue Guana, Wei Wanga, and Wenfei Lia at the Department of Physics at Nanjing University; Qian-Yuan Tang at the Department of Physics at Hong Kong Baptist University; Weitong Ren at the Wenzhou Key Laboratory of Biophysics at the University of Chinese Academy of Sciences; and Mingchen Chen at the Changping Laboratory in Beijing.

More information: Xingyue Guan et al, Predicting protein conformational motions using energetic frustration analysis and AlphaFold2, *Proceedings of the National Academy of Sciences* (2024). DOI: 10.1073/pnas.2410662121

Provided by Rice University