

# New machine-learning approach is better at spotting enzymatic metals in proteins

September 8 2021

---



Joanna Slusky, associate professor of molecular biosciences and computational biology at the University of Kansas, heads the lab where machine learning improved the precision of identifying enzymatic and non-enzymatic metals in proteins. Credit: Meg Kumin

Last season, Kansas City Chiefs quarterback Patrick Mahomes boasted a

66.3 pass-completion percentage.

But Mahomes's impressive stat pales compared with the accuracy of MAHOMES, or Metal Activity Heuristic of Metalloprotein and Enzymatic Sites, a [machine-learning model](#) developed at the University of Kansas—and named in the quarterback's honor—that could lead to more effective, eco-friendly and cheaper drug therapies and other industrial products.

Instead of targeting wide receivers, MAHOMES differentiates between enzymatic and non-enzymatic metals in proteins with a precision rate of 92.2%. A team at KU recently published results on this machine-learning approach to differentiating enzymes in *Nature Communications*.

"Enzymes are super interesting proteins that do all the chemistry—an [enzyme](#) does a chemical reaction on something to transform it from one thing to another thing," said corresponding author Joanna Slusky, associate professor of molecular biosciences and computational biology at KU. "Everything that you bring into your body, your body breaks it down and makes it into new things, and that process of breaking down and making into new things—all of that is due to enzymes."

Slusky and graduate student collaborators in her lab, Ryan Feehan (the Chiefs fan who named MAHOMES) and Meghan Franklin of KU's Center for Computational Biology, sought to use computers to distinguish between metalloproteins, which don't perform [chemical reactions](#), and metalloenzymes, which facilitate chemical reactions with amazing power and efficiency.

The problem is metalloproteins and metalloenzymes are in many ways identical.

"People don't exactly know how enzymes work," Slusky said. "For any

given enzyme you can say, "OK, you know, it takes off this hydrogen and puts on the -OH group," or whatever it does. But if I gave you a protein you had never seen before and I asked, 'Which end is up? Which side of this does the reaction?', you, as a scientist and even as an enzymologist, could probably not tell me. Now, one of the keys is about 40% of all enzymes use metals for catalysis—so their protein binds a [metal](#) and then whatever is getting changed comes into that [active site](#) and is changed. We see this these metal-binding proteins and metalloenzymes, which are enzymes that are binding metals, as a tremendous opportunity for us because my lab is interested in machine learning that can do a really good job at differentiating enzyme sites from similar but nonenzymatic sites."

As a KU undergraduate, co-lead author Feehan began compiling the world's largest structural dataset of enzymatic and nonenzymatic metalloprotein sites—work that carried on into his career as a graduate student. Then, he made the dataset freely available to other researchers on Github.

"Structural data is very hard to come by," Slusky said. "But if you're interested in what the physics and chemistry are, and where those atoms are, and what can they do within those relationships, you need protein structures. The hard part of this was getting a bunch of structures of enzyme sites, knowing they were enzyme sites, then getting a bunch of nonenzyme sites that were binding metals—and knowing they were not enzymes—and digging those out from a large structural database."

Feehan was able to find thousands of unique active and inactive metal binding sites, then tested machine-learning approaches to distinguish between the two. To accomplish this, Feehan and Franklin trained a computer-learning model (MAHOMES) to examine a cleft in a protein and predict if that cleft could do chemistry (meaning it was an enzyme). By looking at physicochemical features, MAHOMES achieved 92.2%

precision and 90.1% recall in telling apart the active and inactive sites.

Slusky said the approach could be an important step to making enzymes more useful for the production of life-saving drug therapies and a host of other industrial processes. Indeed, the approach pioneered by the KU team even could revolutionize how enzymes are designed.

"I hope that it will change synthesis in general," she said. "I hope that there will be cheaper drugs made with fewer environmental ramifications. Right now, pharmaceutical companies' synthesis has tremendous environmental implications, and it would be great if we could lower those. But there's also synthesis in generally every industry. If you want to make paint, paint needs synthesis. Everything's made of chemicals—for instance, textiles. You can harvest cotton, but ultimately, you're going to give particular material properties to that cotton before you sell it, and that requires chemicals. The more synthesis we can do by enzymes and the easier we can make it for companies to do that synthesis by enzymes, the cheaper it will be, and the greener it will be."

According to Slusky, the machine-learning research would continue along three lines.

"Number one, we're trying to make the [machine-learning approach](#) work a little bit better," she said. "Number two, we're starting to design enzymes with it. And number three is we want to do this for enzymes that don't bind metals. Forty percent of all enzyme active sites have metals bound. Let's do the other 60%, too—and finding the right comparison set for the other 60% is a project another graduate student in my lab is working on."

**More information:** Ryan Feehan et al, Machine learning differentiates enzymatic and non-enzymatic metals in proteins, *Nature Communications* (2021). [DOI: 10.1038/s41467-021-24070-3](https://doi.org/10.1038/s41467-021-24070-3)

Provided by University of Kansas

Citation: New machine-learning approach is better at spotting enzymatic metals in proteins (2021, September 8) retrieved 27 April 2024 from <https://phys.org/news/2021-09-machine-learning-approach-enzymatic-metals-proteins.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.