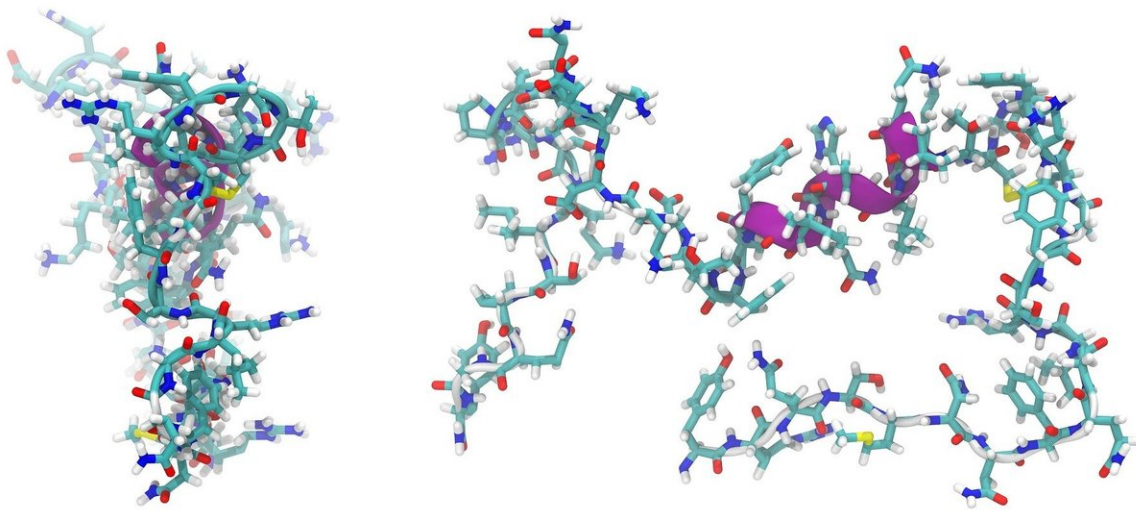


Unlocking the secrets of peptide sequences in cells with AI

November 28 2023



Credit: CC0 Public Domain

Machine learning is now helping researchers analyze the makeup of unfamiliar cells, which could lead to more personalized medicine in the treatment of cancer and other serious diseases.

Researchers at the University of Waterloo developed GraphNovo, a new program that provides a more accurate understanding of the peptide sequences in cells. Peptides are chains of amino acids within cells and are building blocks as important and unique as DNA or RNA.

The study, "Mitigating the missing fragmentation problem in de novo peptide sequencing with a two-stage graph-based deep learning model," was [published](#) in *Nature Machine Intelligence*.

In a healthy person, the immune system can correctly identify the [peptides](#) of irregular or foreign cells, such as [cancer cells](#) or harmful bacteria, and then target those cells for destruction. For people whose immune system is struggling, the promising field of immunotherapy is working to retrain their immune systems to identify these dangerous invaders.

"What scientists want to do is sequence those peptides between the normal tissue and the cancerous tissue to recognize the differences," said Zeping Mao, a Ph.D. candidate in the Cheriton School of Computer Science who developed GraphNovo under the guidance of Dr. Ming Li.

This sequencing process is particularly difficult for novel illnesses or cancer cells, which may not have been analyzed before. While scientists can draw on an existing peptide database when analyzing diseases or organisms that have previously been studied, each person's cancer and [immune system](#) are unique.

To quickly build a profile of the peptides in an unfamiliar cell, scientists have been using a method called de novo peptide sequencing, which uses [mass spectrometry](#) to rapidly analyze a new sample. This process may leave some peptides incomplete or entirely missing from the sequence.

Utilizing [machine learning](#), GraphNovo significantly enhances the accuracy in identifying peptide sequences by filling these gaps with the precise mass of the peptide sequence.

Such a leap in accuracy will likely be immensely beneficial in a variety of medical areas, especially in the [treatment of cancer](#) and the creation

of vaccines for ailments such as Ebola and COVID-19. The researchers achieved this breakthrough due to Waterloo's commitment to advances in the interface between technology and health.

"If we don't have an algorithm that's good enough, we cannot build the treatments," Mao said. "Right now, this is all theoretical. But soon, we will be able to use it in the real world."

More information: Zeping Mao et al, Mitigating the missing-fragmentation problem in de novo peptide sequencing with a two-stage graph-based deep learning model, *Nature Machine Intelligence* (2023).

[DOI: 10.1038/s42256-023-00738-x](https://doi.org/10.1038/s42256-023-00738-x)

Provided by University of Waterloo

Citation: Unlocking the secrets of peptide sequences in cells with AI (2023, November 28)
retrieved 30 April 2024 from

<https://phys.org/news/2023-11-secrets-peptide-sequences-cells-ai.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.