Bacterial Nanopores Open the Future of Data Storage

2020-12-15
Bioengineers at EPFL have developed a nanopore-based system that can read data encoded into synthetic macromolecules with higher accuracy and resolution than similar methods on the market. The system is also potentially cheaper and longer-lasting, and overcomes limitations that prevent us from moving away from conventional data storage devices that are rapidly maxing out in capacity and endurance.

In 2020, each person in the world is producing about 1.7 megabytes of data every second. In just a single year, that amounts to 418 zettabytes—or 418 billion one-terabyte hard drives.

We currently store data as ones and zeroes in magnetic or optical systems with limited lifespans. Meanwhile, data centers consume massive amounts of energy and produce enormous carbon footprints. Simply put, the way we store our ever-growing volume of data is unsustainable.

DNA as data storage

But there is an alternative: storing data in biological molecules such as DNA. In nature, DNA encodes, stores, and makes readable massive amounts of genetic information in tiny spaces (cells, bacteria, viruses)—and does so with a high degree of safety and reproducibility.

Compared to conventional data-storage devices, DNA is more enduring and compacted, can retain ten times more data, has 1000-fold higher storage density, and consumes 100 million times less energy to store the same amount of data as a drive. Also, a DNA-based data-storage device would be tiny: a year's worth of global data can be stored in just four grams of DNA.

But storing data with DNA also involves exorbitant costs, painfully slow writing and reading mechanisms, and is susceptible to mis-readings.

Nanopores to the rescue

One way is to use nano-sized holes called nanopores, which bacteria often punch into other cells to destroy them. The attacking bacteria use specialized proteins known as "pore-forming toxins" which latch onto the cell's membrane and form a tube-like channel through it.

In bioengineering, nanopores are used for "sensing" biomolecules, such as DNA or RNA. The molecule passes through the nanopore like a string, steered by voltage, and its different components produce distinct electrical signals (an "ionic signature") that can be used to identify them. And because of their high accuracy, nanopores have also been tried out for reading DNA-encoded information.

Nonetheless, nanopores are still limited by low-resolution readouts—a real problem if nanopore systems are ever to be used for storing and reading data.

Aerolysin nanopores

The potential of nanopores inspired scientists at [EPFL's](#) [School of Life Sciences](#) to explore nanopores produced by the pore-forming toxin aerolysin, made by the bacterium Aeromonas hydrophila. Led by Matteo Dal Peraro at EPFL's School of Life Sciences, the researchers show that aerolysin nanopores can be used for decoding binary information.

In 2019, Dal Peraro's lab showed that nanopores can be used for sensing more complex molecules, like proteins. In this study, published in [Science Advances](#), the team joined force with the lab of Alexandra Radenovic (EPFL School of Engineering) and adapted aerolysin to detect molecules tailored-made precisely to be read by this pore. The technology has been filed as a patent.

The molecules, known as digital polymers, were developed in the lab of Jean-François Lutz at the [Institut Charles Sadron of the CNRS](#) in Strasbourg. They are a combination of DNA nucleotides and non-biological monomers designed to pass through aerolysin nanopores and give out an electrical signal that could be read out as a bit of data.

The researchers used aerolysin mutants to systematically design nanopores for reading out signals of their informational polymers. They optimized the speed of the polymers passing

through the nanopore so that it can give out a uniquely identifiable signal. "But unlike conventional nanopore readouts, this signal delivered digital reading with single-bit resolution, and without compromising information density," says Dr. Chan Cao, the first author of the paper.

To decode the readout signals the team used deep learning, which allowed them to decode up to 4 bits of information from the polymers with high accuracy. They also used the approach to blindly identify mixtures of polymers and determine their relative concentration.

The system is considerably cheaper than using DNA for data-storage, and offers longer endurance. In addition, it is "miniaturizable," meaning that it could easily be incorporated into portable data-storage devices.

"There are several improvements we are working on to transform this bio-inspired platform into an actual product for data storage and retrieval," says Matteo Dal Peraro. "But this work clearly shows that a biological nanopore can read hybrid DNA-polymer analytes. We are excited as this opens up new promising perspectives for polymer-based memories, with important advantages for ultrahigh density, long-term storage and device portability."

Read the original article on Swiss Federal Institute of Technology Lausanne (EPFL).