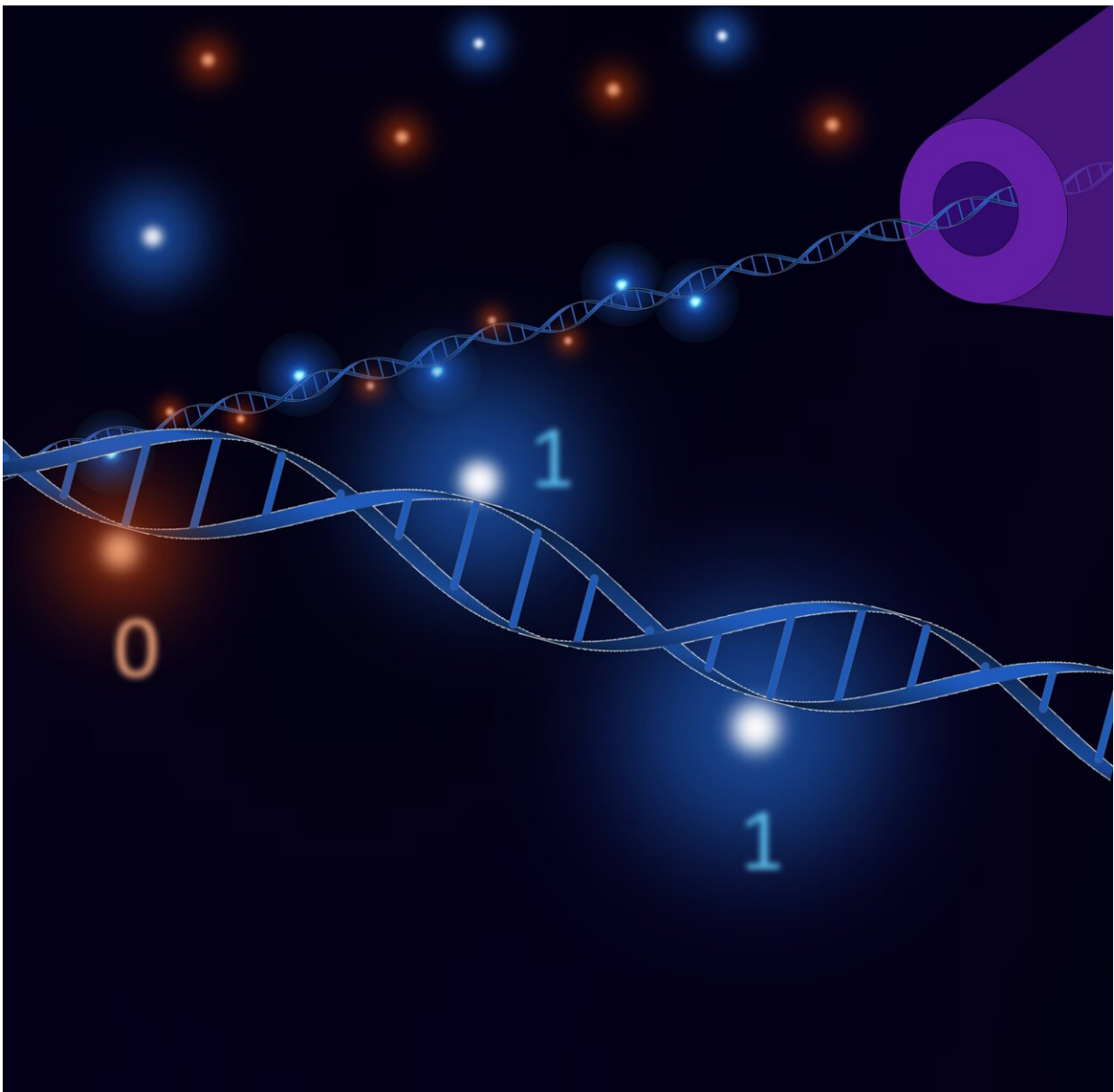


DNA Lego bricks enable fast rewritable data storage

April 7 2020, by Anna Demming



Thinking outside the backbone: overhangs move DNA data technology up a gear.
Credit: Kaikai Chen

DNA data storage may become easier to read and write than before, according to researchers at the University of Cambridge Cavendish Laboratory in the U.K. They report on a technique that can also store encrypted data, as well as re-write data.

The original idea behind DNA data storage is to synthesize long molecules of DNA with bespoke sequences of base units that encode [digital data](#). The data density achieved by this approach is orders of magnitude higher than incumbent magnetic or solid-state technologies, and lasts thousands as opposed to tens of years. The longevity and data density of DNA data storage would be particularly useful for data archives were it not for some significant limitations.

"One of the biggest issues is making the DNA," says [Ulrich Keyser](#), professor of applied physics at the University of Cambridge in the U.K. He explains that synthesizing the de nova DNA molecules with prescribed base unit sequences long enough to store data is difficult and requires enzymes. "With our approach, it is just like Lego bricks. You just do it by mixing together, heating up and cooling down."

Reading data stored in the sequence of base pairs is also slow and expensive. Sequencing technology has come a long way, but it still mostly relies on replicating billions of copies of the molecule to amplify signals from protein interactions, and so on.

An alternative sequencing approach passes the DNA molecule through a nanopore and reads the sequence in real time from the changes in ionic current as different base pairs pass through. Although cheaper and more

efficient, reading bits from base pairs in the DNA backbone still takes too long for data storage technologies. However, by storing data on overhangs stuck on the main backbone, Keyser and his team developed an approach that nanopore technology can easily and accurately read, and simple mixing can write.

By incorporating "toe holds" on the overhang-written data, they show it can be easily removed and rewritten. "I was surprised that the re-writing worked and could be so simple, because this is very difficult with any other DNA data technique," says Keyser.

Sensing potential

"The idea that we started with was for sensing amplification," explains [Kaikai Chen](#), the first author of the [Nano Letters](#) paper reporting these results. "Then we came up with the idea for data storage."

Key to the pioneering approach is controlling how overhangs of single-strand DNA are "annealed." While the sequence of base pairs in the DNA backbone is identical from one molecule to another, the researchers anneal specific overhangs with complementary single-stranded DNA that is biotinylated while the rest is annealed with plain single-strand DNA. Where the complementary strand is biotinylated it will bind with streptavidin molecules, which makes an easily detectable change in the ionic current as the DNA passes through a nanopore, reading it as "1." Where the overhang DNA strand has no streptavidin, the data written is "0." The group used recognized techniques based on molecules that home in on specific regions of the molecule to deliver the correct complementary strand to the right address.

The "toehold" that enables re-writing is just a little extra single-stranded DNA that sticks out after functionalizing, making it easy to remove and re-write. Leaving the biotinylated strands off leaves the data encrypted

because only someone who knows the sequence of the single-strand DNA overhangs will know what sequence the complementary strand needs to have to supply the biotinylated strands that will bind with streptavidin, and so distinguish the ones from the zeroes.

Future

The next challenge for the technology would be scaling up. Since they operate a physics lab, Keyser doesn't see this as the focus of their next steps as a team, although it seems straightforward in principle with the use of pipetting robots or microfluidics. "There are already companies that offer the microfluidics products that could be used," adds Chen.

The researchers are now looking at what other [functional groups](#) they can use besides streptavidin. "In principle, our method can adapt to different functionalization," says Chen. They used streptavidin for their proof of principle because it is a functional group they are familiar with. "It is very straightforward and works well," he adds. However, using smaller groups may allow higher-density storage.

No choice of functional group will enable quite the data density achieved by storing the data in the base pair sequence. Keyser suggests this might also explain why no one thought of trying the Lego block approach before. Although work in new technologies tends to follow up on the techniques already demonstrated rather than taking an orthogonal approach, the focus on optimizing data density may have acted as an additional deterrent. Yet, the advantages of faster, simpler reading and writing, and in particular, re-writing, may make the trade-off worth it. Re-writable DNA data storage also opens up opportunities for DNA computations, which could offer an alternative to traditional computing that, although slow, uses very little energy and so has value for some applications.

More information: Kaikai Chen et al. Nanopore-Based DNA Hard Drives for Rewritable and Secure Data Storage, *Nano Letters* (2020) accepted manuscript pubs.acs.org/doi/10.1021/acs.nanolett.0c00755

© 2020 Science X Network

Citation: DNA Lego bricks enable fast rewritable data storage (2020, April 7) retrieved 23 April 2024 from <https://phys.org/news/2020-04-dna-lego-bricks-enable-fast.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.