



PRESS RELEASE | PARIS | OCTOBER 13 2017

WARNING: Embargoed until October 17, 2017, at 0500 US Eastern Time / 1000 London time (BST) / 1100 Paris time

World first for reading digitally encoded synthetic molecules

For the first time ever, using mass spectrometry, researchers have successfully read several bytes¹ of data recorded on a molecular scale using synthetic polymers. Their work, conducted under the aegis of the Institut Charles Sadron (CNRS) in Strasbourg and the Institute of Radical Chemistry (CNRS / Aix-Marseille University), sets a new benchmark for the amount of data—stored as a sequence of molecular units (*monomers*)—that may be read using this routine method. It also sets the stage for data storage on a scale 100 times smaller than that of current hard drives. The scientists' findings are published in *Nature Communications* on October 17, 2017.

Will hard drives one day use polymers for data storage? For several years now, researchers have faced the challenge of trying to record information with these long molecules. Polymers have great potential since, to record a bit,¹ their component monomers require 100 times less space than current hard drives. Their use could therefore dramatically reduce the size of computer storage media. Yet researchers have been thwarted in their efforts to effectively read polymer data. Now a team of CNRS and Aix-Marseille University scientists has demonstrated that a mass spectrometer may be used to read long sequences of data recorded on a molecular scale. This is a breakthrough because mass spectrometers, popular among chemists, are fast and easy to use.

To pull it off, the team used synthetic molecules, simpler to work with than natural molecules like DNA. Their structure was optimized for sequencing by mass spectrometry. The polymers are made up of two kinds of monomers (with phosphate groups)—corresponding to 0 and 1 respectively. After every eight of these monomer “bits,” a molecular separator was added. The number of bytes represented by the complete polymer equals the number of eight-bit groups. The first step in reading the encoded information is to divide the polymer into molecular bytes by snapping it apart at the separator sites; the next is to break the phosphate bonds, for sequencing of each byte.

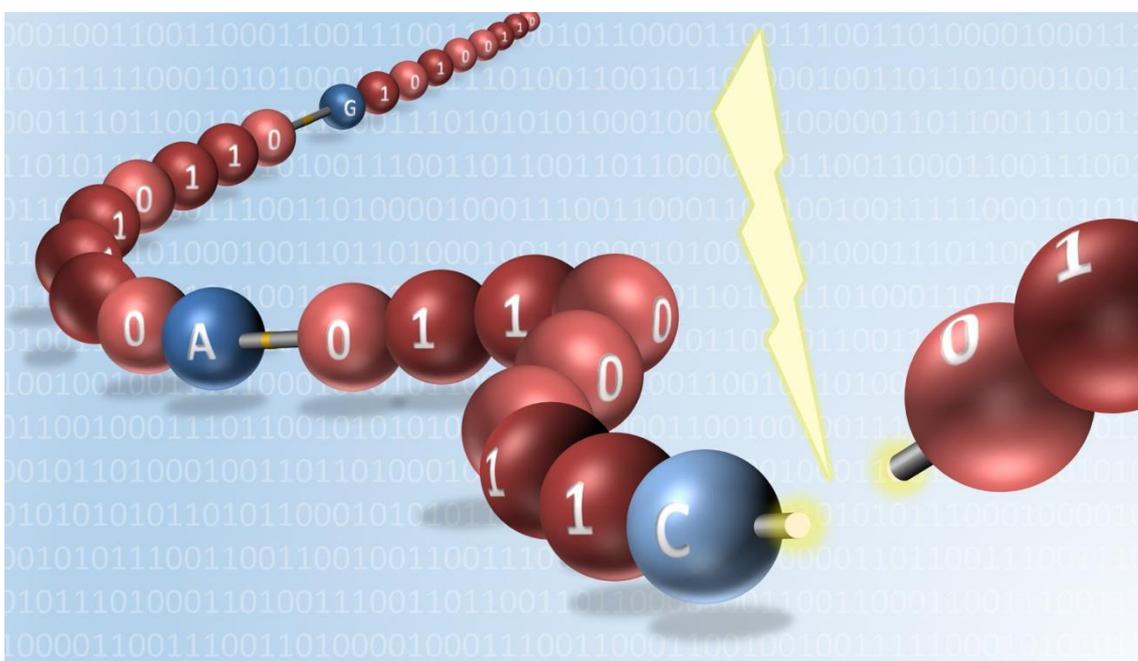
The team of chemists managed to synthesize polymers that can store up to eight bytes. Thus, they were able to record the word “Sequence” in ASCII code, which assigns a unique byte to each letter and punctuation mark. By successfully decoding this word using mass spectrometry, they set a new record for the length of a molecule that may be read using this technique. Although manual analysis of the digital data does take a few hours, it should be possible to reduce the time needed to a few milliseconds by developing software to perform this task. By associating short read times with current automated methods

¹ A *byte* is a segment of information consisting of eight subunits called *bits*. Bits, which represent the lowest level of data, can each have only one of two possible values: 0 or 1.



www.cnrs.fr

for writing data, this work paves the way for synthetic polymer storage of several kilobytes of data, roughly equivalent to a page of text—just like the very first floppy disks.



By inserting fragile bonds between each molecular "byte," digital polymers may be easily read through mass spectrometry. © Jean-François Lutz, Institut Charles Sadron (CNRS)

Bibliography

Mass spectrometry sequencing of long digital polymers facilitated by programmed inter-byte fragmentation | Ouahabi A, Amalian JA, Charles L, Lutz JF. *Nat Commun.* 2017, October 17. DOI : 10.1038/s41467-017-01104-3

Contacts

CNRS researcher | Jean-François Lutz | T (+33) (0)3 88 41 40 16 | jflutz@unistra.fr
CNRS press officer | Alexiane Agullo | T (+33) (0)1 44 96 43 90 | alexiane.agullo@cnrs-dir.fr