

Storage of Information Using Small Organic Molecules

Brian J. Cafferty,^{†,‡} Alexei S. Ten,^{‡,§} Michael J. Fink,^{†,||} Scott Morey,[§] Daniel J. Preston,[†] Milan Mrksich,^{‡,§,||} and George M. Whitesides^{*,†,||,⊥}

[†]Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138, United States

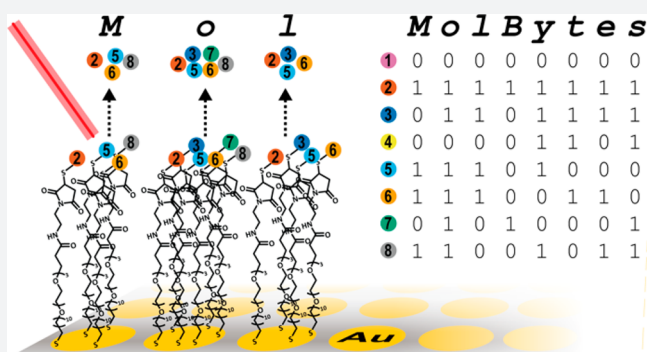
[‡]Department of Chemistry and [§]Department of Biomedical Engineering, Northwestern University, 2145 Sheridan Road, Evanston, Illinois 60208, United States

^{||}Kavli Institute for Bionano Science and Technology, Harvard University, 29 Oxford Street, Cambridge, Massachusetts 02138, United States

[⊥]Wyss Institute for Biologically Inspired Engineering, 3 Blackfan Circle, Boston, Massachusetts 02115, United States

Supporting Information

ABSTRACT: Although information is ubiquitous, and its technology arguably among the highest that humankind has produced, its very ubiquity has posed new types of problems. Three that involve storage of information (rather than computation) include its usage of energy, the robustness of stored information over long times, and its ability to resist corruption through tampering. The difficulty in solving these problems using present methods has stimulated interest in the possibilities available through fundamentally different strategies, including storage of information in molecules. Here we show that storage of information in mixtures of readily available, stable, low-molecular-weight molecules offers new approaches to this problem. This procedure uses a common, small set of molecules (here, 32 oligopeptides) to write binary information. It minimizes the time and difficulty of synthesis of new molecules. It also circumvents the challenges of encoding and reading messages in linear macromolecules. We have encoded, written, stored, and read a total of approximately 400 kilobits (both text and images), coded as mixtures of molecules, with greater than 99% recovery of information, written at an average rate of 8 bits/s, and read at a rate of 20 bits/s. This demonstration indicates that organic and analytical chemistry offer many new strategies and capabilities to problems in long-term, zero-energy, robust information storage.



INTRODUCTION

Technologies from printing with ink on paper, to very sophisticated electronic, optical, and magnetic methods, are used to store information. The importance (across a range of parameters: cost, space, energy use, rate of reading and writing, rate of degradation on storage, potential for corruption through tampering, independence of protocols and hardware for reading) is such that each of these methods has weaknesses in addition to its strengths,^{1–6} and there remains a need to evaluate possible alternatives.^{7–11} New methods would not necessarily replace the amazingly highly engineered methods currently used, but might circumvent some of their weaknesses and perhaps open new applications.

The use of molecules for storing information^{12–16} has in large part been stimulated by the ability of cells to store very large amounts of information in molecules (especially macromolecules: DNA, RNA, proteins, and carbohydrates) and metabolic networks. Most macromolecules use a common strategy of ordering the information along a one-dimensional array of covalently linked monomers (“beads on a string”).

They have also raised enormously interesting questions about the meaning of “information” in living cells; for example, is the cell a Turing machine?

Research is developing strategies that use high-molecular-weight, biologically derived systems, largely based on the sequence of synthetic DNA strands.^{17–19} Our objective is to explore a different strategy—one not modeled on biology—that uses *low* molecular weight molecules. We especially wished to avoid macromolecules that use (often repetitive) organic synthetic steps and that require the synthesis of a unique macromolecule for each separate message. We have instead used sets of oligopeptides having distinguishable molecular weights to store information. Overall, this system requires a set of a maximum of eight oligopeptides, as a mixture, in a microwell, to store one byte, and a mixture of 32 oligopeptides to store four bytes. Using larger mixtures of oligopeptides enables storage of larger sets of data. These

Received: March 3, 2019

Published: May 1, 2019

systems are capable of writing any arbitrary binary information using the same set of small molecules. In this work, reading is accomplished by identifying the masses of the molecules that are immobilized to a self-assembled monolayer (primarily as disulfides from the laser desorption process) using mass spectrometry. Mass spectrometry provides both high precision (enabling accurate determination of the composition of mixtures of oligopeptides in a single sub-millimeter spot of an immobilized array, without separation, and with few errors) and high rates of reading.

RESULTS

Our initial demonstration has been to write messages in eight-bit American Standard Code for Information Interchange (ASCII), convert them to an equivalent molecular code, store them on an array plate (four bytes per spot), and read them using self-assembled monolayers for matrix-assisted laser desorption/ionization (SAMDI) mass spectrometry.²⁰ ASCII is a look-up table that includes the alphabet, numbers, punctuation, and special characters—a maximum of 256 characters—and is used primarily for alphanumeric text. Table 1 summarizes this strategy for the letter “K,” and

Table 1. Correspondence of an Alphanumeric Character (the Letter “K”) Encoded in ASCII in Binary, and in Four Molbits as Oligopeptides

Byte		K							
I	Letter	0	1	0	0	1	0	1	1
	Binary String in ASCII [†]	0	1	0	0	1	0	1	1
	Molbit Assignment to Binary String		Ac-ZK ^{(me3)C}			Ac-GVK ^{(me3)C}		Ac-ALK ^{(me3)C}	Ac-GFK ^{(me3)C}
	Oligopeptide Number [‡]	–	2	–	–	5	–	7	8
II*		–	10	–	–	13	–	15	16
III		–	18	–	–	21	–	23	24
IV		–	26	–	–	29	–	31	32

Ac: acetyl, A: alanine, G: glycine, C: cysteine, F: phenylalanine, K^(me3): N^ε,N^ε,N^ε-trimethyl lysine, L: leucine, V: valine, Z: 2-aminobutyric acid.
[†]See Table S2 for binary assignment of printable character in ASCII.
[‡]Numbering scheme for the 32 peptides used in this work (Table S1 gives complete sequences).
 *The peptides that correspond to the molbits that represent the letter “K” in bytes II, III, and IV are also shown.

Table S1 summarizes a complete assignment of oligopeptides sufficient to encode four bytes in a single mixture, with their assignments to a binary molecular representation.

To differentiate the two systems with which we are working (*electronic* storage and its theoretic foundation in Boolean algebra, and *molecular* storage), we call the equivalent of a bit, and of an eight-bit byte, of information—in the form of mixtures of molecules—a “molbit” and a “molbyte.” To store information in molecules, we first designed a method that would allow us to encode ASCII in molecules distinguishable by mass spectrometry. For example, the letter “K” in ASCII is represented by one byte (01001011) in binary. We convert that binary representation to a molecular one by assigning an oligopeptide to each of the eight bits in a byte, and include that oligopeptide on the spot if the bit value is “1” and omit it if the bit value is “0” (Table 1).

These oligopeptides were selected to have four characteristics: (i) All were resolvable by mass using SAMDI as components of a common mixture (Figure 1). The different amino acids in each oligopeptide are covalently bonded, but

their order is not relevant—only the total mass. The oligopeptides are *not* covalently bonded to one another and do not form macromolecules. Information is thus stored as mixtures of low molecular weight ($MW < 1000 \text{ g mol}^{-1}$) molecules, in arrays, specifying “1” and “0” in a binary representation, rather than as a sequence of groups in a linear polymer. (ii) All oligopeptides terminate in a cysteine to allow efficient immobilization by Michael addition to the reactive maleimide group present in the 1.25 mm diameter spot of the SAMDI plate. (iii) Each oligopeptide includes a trimethyllysine (K^{Me3}) with a fixed positive charge to aid in mass spectrometry (positive mode). By using the set of 32 peptides listed in Table S1, each of which is distinguishable in a mixture containing the others, we can store the information for four molbytes (e.g., four letters in ASCII) in one spot.

Using this method, the *presence* of a particular peptide in a mixture indicates three parameters: (i) The byte (out of the four bytes, when 32 peptides are used) to which it is contributing information; (ii) its location (which is assigned based on the molecular weight of the corresponding peptide) in the bitstring of that byte; and (iii) its value (“1”). The *absence* of that peptide indicates that that position in the molbyte is “0”. The presence of the four oligopeptides listed in Table 1 is thus assigned to bits with the value 1, and the four oligopeptides absent from the mixture are assigned to bits with the value 0. The one remaining parameter to be defined is the position of this byte (or bytes when more than eight molbits are used per spot) in the sequence of the entire message: this information is provided by the position of the spot in the sequence of spots on the SAMDI array plate. The attractive feature of this method is that only eight oligopeptides allows the specification of all of the characters of one byte, and thus allows an arbitrary message to be written in ASCII (or any character set of 256 members); by using 32 distinguishable oligopeptides, we can specify four bytes in one spot.

Figure 2 outlines the process we used to “write”, “store”, and “read” text using this set of 32 peptides. For a particular byte, the appropriate set of oligopeptides representing “1”s in the bitstring is deposited and mixed in wells of a 384-well plate using an Echo 555 liquid handler. A Tecan liquid handler then transfers these mixtures to an array plate having 1536 gold islands (“spots”), each presenting a self-assembled monolayer. The peptides react covalently with the terminal maleimide groups present on the monolayers of the array plate. Covalent coupling prevents the components of the mixture from spreading on the surface and allows their analysis with SAMDI mass spectrometry. The plate, with the completed text encoded as mixtures of oligopeptides in spots ordered on the plate, is stored. Reading by SAMDI is as described previously.²⁰

This strategy for writing and reading bytes allows a small number of low-molecular-weight molecules to encode many forms of information and, once synthesized, avoids the need for further synthesis to store a new message. (In this demonstration, to order these molbytes, we use an array plate in the format of a conventional microwell plate, but a number of other formats are also possible). The density of information (D) we can put on a plate depends on the representation, but here is given by $D = (\text{molbyte}/\text{cm}^2) = (\text{wells}/\text{cm}^2) (\text{molbyte}/\text{well})$. For the current system, this number is $D = 64 \text{ bytes}/\text{cm}^2$.

Our examples here include text (Supporting Information) and JPEG images (Figure 3). The procedure we use is

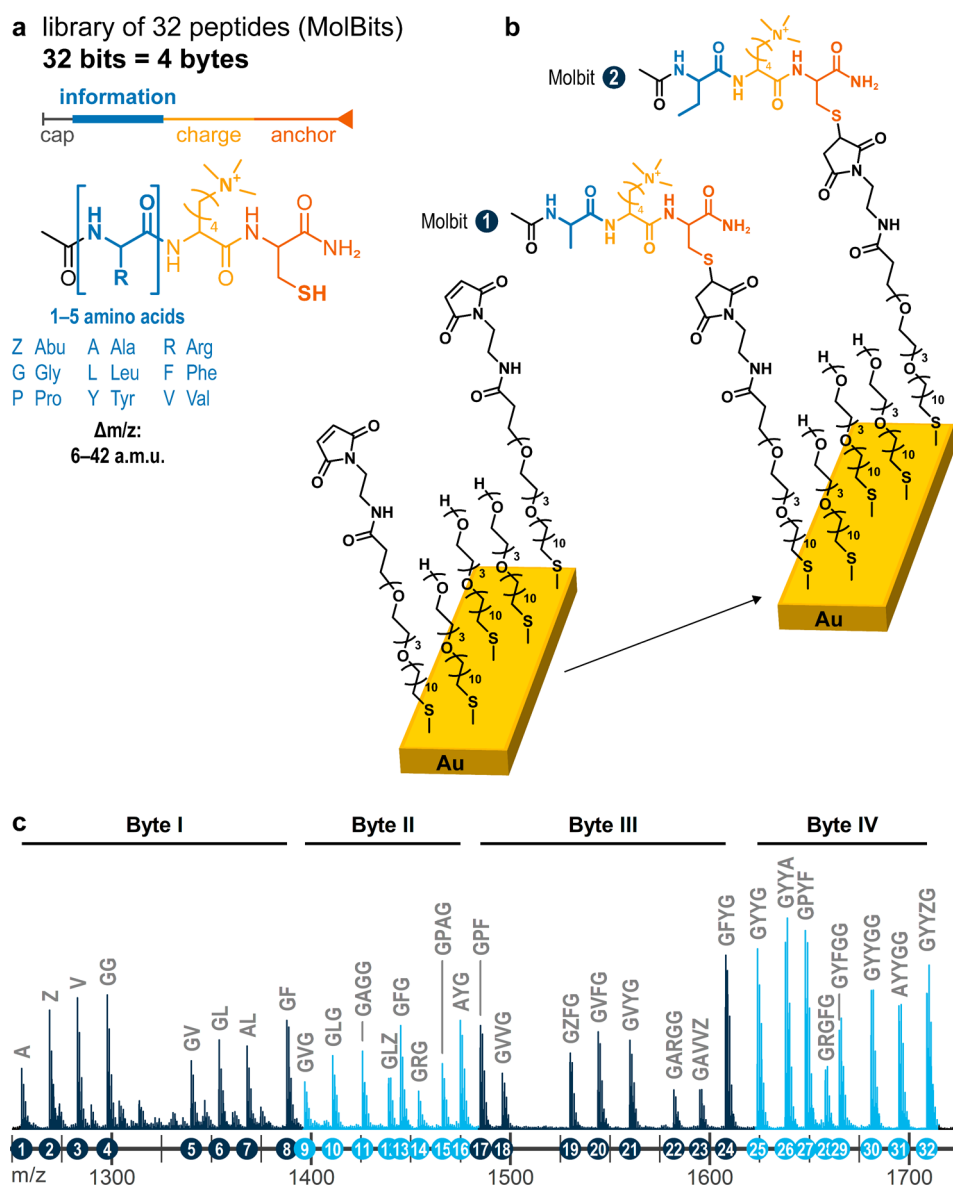


Figure 1. Design of oligopeptide molbits and spectrum of all 32 molbits in a single mixture. (a) Oligopeptide molbits contain an information region that consists of one to five amino acids (chosen from 2-aminobutyric acid, alanine, arginine, glycine, leucine, phenylalanine, proline, tyrosine, valine), which provides a distinguishable mass-to-charge ratio for each peptide (a difference of 6–42 a.m.u.), a charge residue (trimethyl lysine), and an anchor residue (terminal cysteine). The N-terminus is capped by an acetyl group for chemical stability. (b) Schematic showing an example of the immobilization of two oligopeptides (corresponding to molbit 1 and molbit 2 in panel c) to a maleimide-terminated monolayer for storage. Prior to conjugation of oligopeptide(s), the monolayer consists of a mixture of triethyleneglycol undecanethiol (EG₃-capped alkanethiol) terminating in either an alcohol or maleimide. (c) A spectrum of a SAMDI spot containing all 32 molbits; the intensity was normalized to the highest signal. Oligopeptides were grouped by molecular weight into sets of eight, representing a byte of information (4 bytes total). The single-letter codes of residues in the information region are listed above each peak in the mass spectrum (see Table S1 for a full list of peptide sequence and corresponding masses, and Figure S1 for a detailed spectrum). The observed masses are for mixed disulfides derived from a EG₃-capped alkanethiol and the oligopeptide conjugated to a maleimide-terminated EG₃-capped alkanethiol.

operationally simple. The small number of molecules required (within a given set such as oligopeptides) need only be synthesized once (or, more probably, purchased, since there are many custom commercial suppliers) and serves to encode a very wide range of information. The text of Feynman’s famous lecture “There is plenty of room at the bottom” is a demonstration of current capability (the text and errors are shown in the Supporting Information). It was written, stored, and read with 99.9% recovery of information. This text (38 313 bytes or alphanumeric characters) was written and read using one set of devices (Figure 2) in 20 h. The images (Figure 3)

are another. The speed of “writing” is 8 bits/s, and “reading” is 20 bits/s, without parallelization. This process is obviously amenable to simple linear parallelization, particularly since each line of instruments could be writing different information at the same time, using a shared set of molecules for storage: the speed could thus easily be increased by a factor of 10 or more, albeit at 10 times the capital cost.

Safety Statement. No unexpected or unusually high safety hazards were encountered.

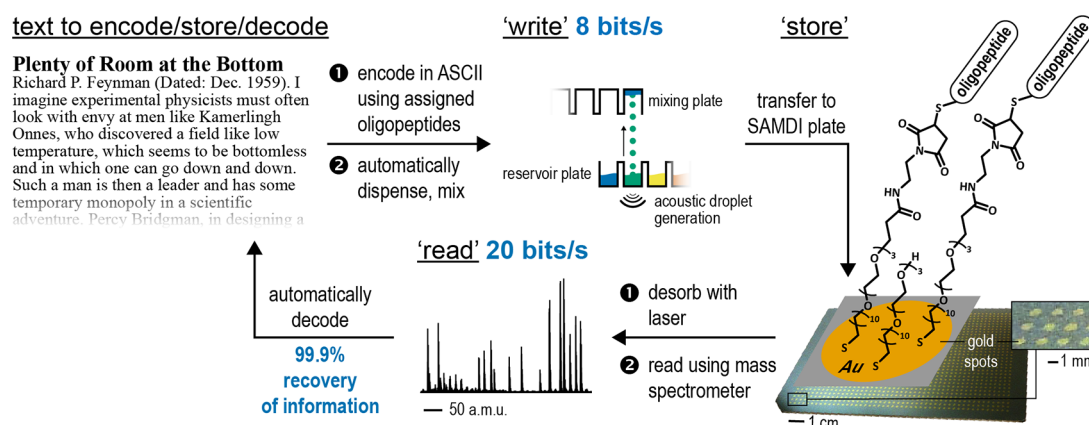


Figure 2. Overview of “writing” using mixtures of molbits and “reading” using SAMDI MS. “Writing” is performed by first translating information (here, the alphanumeric characters of Feynman’s lecture “There is plenty of room at the bottom”) into binary. Binary information is converted to oligopeptides immobilized on a self-assembled monolayer, for storage. A MALDI-TOF mass spectrometer analyzes (“reads”) these plates. A program decodes the information in the spectra and generates a bitstring that is used to regenerate the original text. Recovery of information was determined by (number of correctly identified molbits)/(total number of molbits) \times 100%.

DISCUSSION

This paper demonstrates one method of encoding information for storage in molecules. It represents one of two limiting strategies for this purpose. The first (this method) encodes information in simple, separate molecules that are designed to minimize synthesis and to fit naturally as the molecular equivalent of existing methods of storing digital electronic information. It is not intended (at least at this stage in development) to compete with existing electronic, optical, or magnetic methods of storage. Instead, its immediate objective is to provide an alternative method for archival storage that is stable for long times, does not require energy for storage, and is secure. The molecules are used as mixtures and ordered both by physical location (on an array plate) and by mass of the molecules within each spot of the array. This method is designed for flexibility in writing and use, for simplicity, and for long-term stability in storage. It enables the encoding of *any* binary data, using the same procedures and a constant set of small molecules (which could easily be available on a multi-kilogram scale). No additional synthesis is required for each new message. It depends entirely on simple physical manipulations for sampling, liquid transfer, mixing, separation, and reading at all of its steps. For reading, it uses (in the demonstration shown here) a mass spectrometer—a technique that provides dramatically more information than reading charge on capacitors. The sensitivity in MALDI-TOF MS is highly dependent on the specific analyte and on sample preparation, and has not been rigorously determined for monolayers of alkanethiolates on gold. In favorable cases, however, as few as thousands of analyte molecules are sufficient to produce a spectrum.²¹ The capability of mass spectrometry continues to increase rapidly and thus will enable the further extension of these techniques. The higher the resolution of the spectrometer, the more complex the mixtures that can be analyzed, and thus the greater the amount of information that can be stored per array.

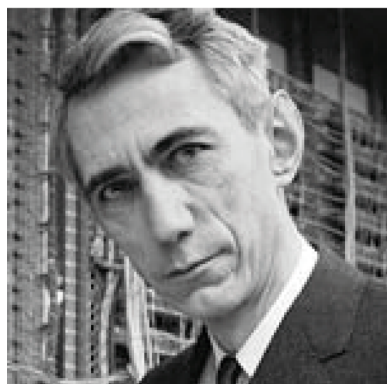
One method of distinguishing different procedures for storing information is by the number and density of “locations” (that is, places where information can be stored: i.e., capacitors in solid-state devices, grooves in optical disks, magnetic domains, visible letters), the amount of information that can be stored in each location, and the cost and time of writing,

storage, and reading. For electronic storage, transistors are very small (approximately $10^{11}/\text{cm}^2$ at the current 11 nm minimum feature size) and inexpensive, but they store only one bit. Methods for storing information in molbytes—as outlined here—offer a high density of information per location, but—using currently available plates—a density of locations that is modest. Molecular storage by this method will improve rapidly with advances in technology for spotting. Higher density of spots in arrays and faster liquid transfer could be achieved by inkjet printing.^{18,22} For example, inkjet printers can generate drops at rates of ~ 1000 per second and position drops on a surface with center-to-center distances of $10\ \mu\text{m}$.²³ Using a set of 32 molbits (as described in this work), this spotting diameter would give a density of information of $4\ \text{MB}/\text{cm}^2$. Optimization of inkjet printers for the type of molecular ink used and speed would further increase the density and rate of writing information. We have not yet analyzed and redesigned this system for efficiency.

The current demonstrations have used oligopeptides, but many other classes of organic molecules (additional unnatural amino acids, fatty acids, aromatics including heterocycles, saturated terpenes, and others) are also possible: the method thus has broad scope. Although we have designed the current system for simplicity, the combination of molecular design, organic synthesis, and advanced methods of separations and analysis also has the potential to greatly increase the amount of information that can be stored per molecule and per location (e.g., spots, wells).

Choosing classes of molecules for information storage that offer long-term stability, with no energy required for storage, is one long-term objective of this area of research. Long-term stability of appropriate organic molecules with appropriate structures over hundreds of years has not been systematically explored but is commonly assumed. Oligopeptides have stabilities of hundreds or thousands of years under suitable conditions,²⁴ i.e., in the absence of light (or ionizing radiation), oxygen or other oxidants, and high temperatures, and possibly in the absence of water, in inert containers. Importantly, occasional breaks in individual molecules would (unlike breaks in DNA) not significantly damage the fidelity of reading, since they would appear at masses that are not coded by the molbits. Molecular storage of information should be

a



b



Figure 3. JPEG images encoded using mixtures of molbits. (a) Image of Claude Shannon, “the father of information theory,” encoded in 6094 molbytes, and (b) image of the woodblock print entitled *The Great Wave off Kanagawa*, by Hokusai, encoded in 5953 molbytes. Each image was encoded on a single plate and decoded with zero errors (100% recovery). Images were encoded as JPEG files as described in the Supporting Information. See Figure S2 for an example of the resolution of an encoded JPEG file. Photograph of Claude Shannon by Alfred Eisenstaedt/The LIFE Picture Collection/Getty.

especially resistant to tampering electrically, magnetically, or optically, since the only way to read or rewrite the composition of information stored molecularly would be to access the molecules physically and then to do chemistry.

A second strategy for storing information in molecules is using organic polymers (DNA, synthetic polymers, proteins, oligosaccharides).^{14,19} These methods—in principle—have other attractive characteristics. DNA has the specific ability (and the requirement) to order the information in a message in terms of the position of its nucleotide components along a covalently linked chain (“beads on a string”) and thus does not require spatial ordering, but it has the accompanying substantial disadvantage of having to synthesize a new strand and sequence of DNA for each different message. DNA also offers the potential of low cost for reading (using Gen 4 and future methods), and amplification by replication (albeit with

unknown problems, for long unnatural and nonbiological sequences).

Although methods of synthesis are improving,²⁵ synthesis of DNA (particularly of long strands) by chemical synthetic methods remains slow and expensive. Cycle times for the coupling of nucleotides using phosphoramidite chemistry²⁶ (the most common approach used for synthesis in DNA-based storage systems) are on the order of 10 minutes. This amount of time, which does not account for additional processing of the DNA strands (cleavage from support and deprotection of nucleobases), is equivalent to a rate of writing of 0.001 bits/s—a rate that is significantly slower than the (unoptimized) approach described here (8 bits/s). The cost of writing a bit of information using DNA as a storage medium has been reported to be as low as $\$5 \times 10^{-4}$,²⁷ which is $\sim 1 \times 10^8$ times more expensive than the price to store a bit in a hard disk drive.²⁸ The peptides used in this work are relatively expensive ($\sim \$1 \times 10^{-3}$ per bit) because they were custom-synthesized, but they are not intrinsically expensive, even at the multi-kilogram scale. A significant advantage in cost for the approach described here over approaches that use DNA (or other sequence-controlled polymers) is that inexpensive commodity chemicals can be used as information carriers (for example, we estimate that using alkanethiols as molbits would reduce the cost to below $\$1 \times 10^{-10}$ per bit). The price of liquid handlers and analytical instruments (here, MALDI-TOF MS) required for this approach can be expensive, but they need only be purchased once.

It is too early in the development of either strategy, or of others, to compare them in specific applications. Their ultimate applications may also be quite different. It is also impossible to compare them with the well-established electrical, optical, and magnetic methods, which are the products of exceptionally successful, multidecade programs in technology development. It is, however, clear that the chemistry of small molecules, and the analytical and synthetic methods that have been developed for synthesizing, separating, and identifying them, offer an exceptionally rich array of scientific and technological methods to apply in new approaches to information storage.

■ ASSOCIATED CONTENT

§ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: [10.1021/acscentsci.9b00210](https://doi.org/10.1021/acscentsci.9b00210).

Peptide synthesis; liquid handling; preparation of SAMs; analysis by MALDI-TOF MS; programs for writing, reading, and compression of data; encoded text; supplementary figures and tables (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: gwhitesides@gmwhgroup.harvard.edu.

ORCID

Michael J. Fink: [0000-0003-0023-8767](https://orcid.org/0000-0003-0023-8767)

Milan Mrksich: [0000-0002-4964-796X](https://orcid.org/0000-0002-4964-796X)

George M. Whitesides: [0000-0001-9451-2442](https://orcid.org/0000-0001-9451-2442)

Author Contributions

B.J.C., M.J.F., and G.M.W. conceived the research. B.J.C., A.S.T., and M.M. designed the experiments. A.S.T., B.J.C., and M.J.F. performed experiments and analyzed results. S.M. and

D.J.P. wrote programs for importing data and analyzing mass spectra. B.J.C., M.J.F., and G.M.W. wrote the paper. All authors edited and commented on the manuscript and contributed to later iterations.

Author Contributions

#B.J.C. and A.S.T. contributed equally to this work.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Sergey V. Ten for writing the spectral analysis program and Sara Fernandez Dunne of the High Throughput Analysis Laboratory at Northwestern University, who helped with liquid handling robots. This work was supported by Defense Advanced Research Projects Agency under Award No. W911NF-18-2-0030.

REFERENCES

- (1) Lloyd, S. Ultimate physical limits to computation. *Nature* **2000**, *406*, 1047.
- (2) Shulaker, M. M.; Hills, G.; Park, R. S.; Howe, R. T.; Saraswat, K.; Wong, H. S. P.; Mitra, S. Three-dimensional integration of nanotechnologies for computing and data storage on a single chip. *Nature* **2017**, *547*, 74–78.
- (3) Salahuddin, S.; Ni, K.; Datta, S. The era of hyper-scaling in electronics. *Nat. Electron.* **2018**, *1* (8), 442–450.
- (4) Fettweis, G.; Zimmermann, E. ICT Energy Consumption - Trends and Challenges. The 11th International Symposium on Wireless Personal Multimedia Communications (WPMC 2008), Saariselkä, Finland, 2008.
- (5) Baliga, J.; Ayre, R.; Hinton, K.; Tucker, R. S. Green cloud computing: balancing energy in processing, storage, and transport. *Proc. IEEE* **2011**, *99*, 149–167.
- (6) Brandner, R.; Pordesch, U.; Wallace, C. *Long-Term Archive Service Requirements*; The IETF Trust: Internet Requests for Comments, 2007.
- (7) Kalff, F. E.; Rebergen, M. P.; Fahrenfort, E.; Girovsky, J.; Toskovic, R.; Lado, J. L.; Fernandez-Rossier, J.; Otte, A. F. A kilobyte rewritable atomic memory. *Nat. Nanotechnol.* **2016**, *11* (11), 926–929.
- (8) Ahn, J. Information storage and retrieval through quantum phase. *Science* **2000**, *287* (5452), 463–465.
- (9) Zhang, J.; Gecevicius, M.; Beresna, M.; Kazansky, P. G. Seemingly unlimited lifetime data storage in nanostructured glass. *Phys. Rev. Lett.* **2014**, *112* (3), 033901.
- (10) Gu, M.; Zhang, Q.; Lamon, S. Nanomaterials for optical data storage. *Nat. Rev. Mater.* **2016**, *1* (12), 16070.
- (11) Begtrup, G. E.; Gannett, W.; Yuzvinsky, T. D.; Crespi, V. H.; Zettl, A. Nanoscale reversible mass transport for archival memory. *Nano Lett.* **2009**, *9* (5), 1835–1838.
- (12) Clelland, C. T.; Risca, V.; Bancroft, C. Hiding messages in DNA microdots. *Nature* **1999**, *399* (6736), 533–534.
- (13) Green, J. E.; Choi, J. W.; Boukai, A.; Bunimovich, Y.; Johnston-Halperin, E.; DeIonno, E.; Luo, Y.; Sheriff, B. A.; Xu, K.; Shin, Y. S.; Tseng, H. R.; Stoddart, J. F.; Heath, J. R. A 160-kilobit molecular electronic memory patterned at 10(11) bits per square centimetre. *Nature* **2007**, *445* (7126), 414–417.
- (14) Colquhoun, H.; Lutz, J. F. Information-containing macromolecules. *Nat. Chem.* **2014**, *6* (6), 455–456.
- (15) Zhirnov, V.; Zadegan, R. M.; Sandhu, G. S.; Church, G. M.; Hughes, W. L. Nucleic acid memory. *Nat. Mater.* **2016**, *15* (4), 366–370.
- (16) Goodwin, C. A. P.; Ortu, F.; Reta, D.; Chilton, N. F.; Mills, D. P. Molecular magnetic hysteresis at 60 K in dysprosocenium. *Nature* **2017**, *548* (7668), 439–442.
- (17) Church, G. M.; Gao, Y.; Kosuri, S. Next-generation digital information storage in DNA. *Science* **2012**, *337* (6102), 1628.
- (18) Goldman, N.; Bertone, P.; Chen, S.; Dessimoz, C.; LeProust, E. M.; Sipos, B.; Birney, E. Towards practical, high-capacity, low-maintenance information storage in synthesized DNA. *Nature* **2013**, *494* (7435), 77–80.
- (19) Organick, L.; Ang, S. D.; Chen, Y. J.; Lopez, R.; Yekhanin, S.; Makarychev, K.; Racz, M. Z.; Kamath, G.; Gopalan, P.; Nguyen, B.; Takahashi, C. N.; Newman, S.; Parker, H. Y.; Rashtchian, C.; Stewart, K.; Gupta, G.; Carlson, R.; Mulligan, J.; Carmean, D.; Seelig, G.; Ceze, L.; Strauss, K. Random access in large-scale DNA data storage. *Nat. Biotechnol.* **2018**, *36* (3), 242–248.
- (20) Mrksich, M. Mass spectrometry of self-assembled monolayers: a new tool for molecular surface science. *ACS Nano* **2008**, *2* (1), 7–18.
- (21) Keller, B. O.; Li, L. Detection of 25,000 molecules of Substance P by MALDI-TOF mass spectrometry and investigations into the fundamental limits of detection in MALDI. *J. Am. Soc. Mass Spectrom.* **2001**, *12* (9), 1055–1063.
- (22) Singh, M.; Haverinen, H. M.; Dhagat, P.; Jabbour, G. E. Inkjet printing-process and its applications. *Adv. Mater.* **2010**, *22* (6), 673–685.
- (23) Derby, B. Inkjet printing of functional and structural materials: fluid property requirements, feature stability, and resolution. *Annu. Rev. Mater. Res.* **2010**, *40* (1), 395–414.
- (24) Martin, R. B. Free energies and equilibria of peptide bond hydrolysis and formation. *Biopolymers* **1998**, *45* (5), 351–353.
- (25) Hughes, R. A.; Ellington, A. D. Synthetic DNA synthesis and assembly: putting the synthetic in synthetic biology. *Cold Spring Harbor Perspect. Biol.* **2017**, *9* (1), No. a023812.
- (26) Beaucage, S. L.; Caruthers, M. H. Deoxynucleoside phosphoramidites—A new class of key intermediates for deoxy-polynucleotide synthesis. *Tetrahedron Lett.* **1981**, *22* (20), 1859–1862.
- (27) Erlich, Y.; Zielinski, D. DNA Fountain enables a robust and efficient storage architecture. *Science* **2017**, *355* (6328), 950–954.
- (28) Panda, D.; Molla, K. A.; Baig, M. J.; Swain, A.; Behera, D.; Dash, M. DNA as a digital information storage device: hope or hype? *3 Biotech* **2018**, *8* (5), 239.