ACS central science

http://pubs.acs.org/journal/acscii

Research Article

Storing and Reading Information in Mixtures of Fluorescent Molecules

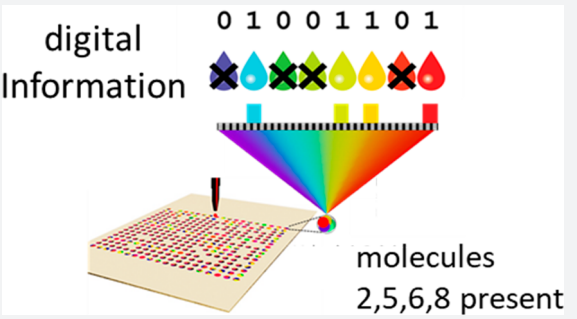
Amit A. Nagarkar, Samuel E. Root, Michael J. Fink, Alexei S. Ten, Brian J. Cafferty, Douglas S. Richardson, Milan Mrksich, and George M. Whitesides*

Cite This: *ACS Cent. Sci.* 2021, 7, 1728–1735

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The rapidly increasing use of digital technologies requires the rethinking of methods to store data. This work shows that digital data can be stored in mixtures of fluorescent dye molecules, which are deposited on a surface by inkjet printing, where an amide bond tethers the dye molecules to the surface. A microscope equipped with a multichannel fluorescence detector distinguishes individual dyes in the mixture. The presence or absence of these molecules in the mixture encodes binary information (i.e., "0" or "1"). The use of mixtures of molecules, instead of sequence-defined macromolecules, minimizes the time and difficulty of synthesis and eliminates the requirement of sequencing. We have written, stored, and read a total of approximately 400 kilobits (both text and images) with greater than 99% recovery of information, written at an average rate of 128 bits/s (16 bytes/s) and read at a rate of 469 bits/s (58.6 bytes/s).



Downloaded via NORTHWESTERN UNIV on November 29, 2021 at 17:48:25 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

INTRODUCTION

In order to preserve information over long periods of time, reduce the energy consumption for storage, and prevent tampering with stored information, new materials and strategies for storage of information would be useful and may be required.[1−5] Current devices used to store information (optical media, magnetic media, and flash memory) have insufficient operational lifetimes for long-term storage—typically less than two decades—and require substantial energy to maintain the stored information.[6]

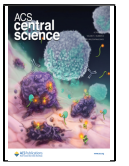
Molecules (including, but not limited to DNA) can be used to store information without power, at high areal density, and are claimed to be stable for thousands of years or more.[7−14] For these systems to be applied to store information, however, several problems must be considered including (i) read/write speeds, (ii) retention of information, (iii) density of information, and (iv) cost.[15]

Here, we demonstrate a write-once-read-many (WORM) molecular information storage approach using mixtures of fluorescent dye molecules covalently bound to an epoxy substrate. An inkjet printer enables writing of information at a rate of 16 bytes/s, and a multichannel fluorescence detector in a confocal microscope enables reading at a rate of 58 kilobytes/s. Using this approach, we have written 14 075 bytes of digital information on a 7.2 mm × 7.2 mm surface (resulting in an aerial information density of 271.5 bytes/mm^2) and read this information over 1000 times without significant loss (less than 20%) in fluorescent signal intensity. This approach enables information storage with high density, fast read/write speeds, and multiple reads of a single set of molecules without loss of information, all at an acceptable cost.

Devices currently used to store digital information—including optical disks, flash drives, and hard disk drives—have operational lifetimes on the order of decades.[16] An alternative approach to such technologies is, in principle, to store information in molecules, as molecule-based storage systems can have very high theoretical storage densities and half-lives that can extend to millions of years.

Sequence-defined polymers have been examined for application in data storage, information processing, and product validation. Inspired by how nature stores genetic information, synthetic DNA has become the most popular molecule to be considered for information storage. While synthetic DNA provides one of the densest methods of data storage (~10^{18} bytes/mm^3),[17] storage of information in long DNA strands suffers from several significant problems: (i) DNA sequencing methods (e.g., Next Gen Sequencing[18]) are slow and, even with massive parallelization, typically require multiple hours to decode a simple message. This slow rate of reading makes this technique impractical for many applications where latency (time to access and read the stored information)

Received: June 17, 2021
Published: October 13, 2021



ACS Publications

© 2021 The Authors. Published by American Chemical Society

is important (e.g., data centers); (ii) information systems that use synthetic DNA typically use polymers that are greater than 100 nucleotides in length, which, due to inefficient monomer coupling, lead to multiple truncation products that decreases the information density of the material.[19,20]

As an alternative to DNA-based systems, several groups have examined nonbiological polymers for molecular information storage. In particular, Lutz and co-workers[11−13] have encoded binary information into several sequence-defined polymers, including non-natural polyphosphates,[14] oligo(alkoxyamine amide)s,[15] and oligo(triazole amide)s[16] and decoded information in these polymers by sequencing them with tandem mass spectrometry. These synthetic polymer systems require extensive synthesis and purification. For these polymers to encode kilobytes of data, the polymer chain must be thousands of units long, but iterative monomer addition suffers from a decrease in the yield of the polymer with each additional coupling.

We have recently demonstrated that information can be stored in the composition of a mixture of oligopeptides, rather than the sequence of a long polymer with individual units covalently bonded to form a chain.[21] The use of smaller fragments, combined with the commercial availability of these units, eliminates the need for time-consuming and expensive synthesis. We have used laser-ionization mass spectrometry to read information stored in molecules on a metal surface. This method has certain limitations: (i) mass spectrometry is a destructive approach, and thus information is destroyed during read-out; (ii) only one location is read at a time, making the process of read-out slow (20 bits/s) and difficult to parallelize; (3) there is limited potential to scale down the feature size, as a decrease in laser spot size leads to an increase in noise.[22]

The objective of this work is to demonstrate the storage of information in a set of *optically* distinguishable molecules (rather than oligopeptides distinguishable by molecular weight using a mass spectrometer). Rather than molecular weight, we use the difference in the wavelength of fluorescent emission of commercially available dyes to design an optochemical molecular information storage system. Information is "written" by inkjet printing of solutions of fluorescent dyes onto a reactive polymeric substrate. Information is "read" using a fluorescence microscope equipped with a multichannel fluorescence detector that can resolve, simultaneously and independently, any combination of the dyes on the substrate. This optical read-out technique uses commercially available technologies and takes advantage of parallelized "reading". The system enabled by this combination of molecules is fundamentally different from other optical storage methods.

The substrate, onto which information is written, is a thin film of an epoxy polymer, that contains reactive amino groups. The *N*-hydroxysuccinimide (NHS) functionalized dyes react on the substrate to form stable amide bonds. We demonstrate that these covalently immobilized dyes are stable to more than 1000 reads without significant loss of intensity. In this work, we used commercially available fluorescent dyes that have been optimized to reduce the extent of photobleaching.

There are several advantages of our molecular information storage technique as compared to magnetic tape, which is the state-of-the-art for long-term storage technique:[27] (i) information can be stored, presumably, with lower environmental and power requirements (in magnetic tape, the binder that secures the paramagnetic material to the substrate can fail in humid conditions[28]); (ii) information can be stored less expensively than with magnetic tape (see Supporting Information, section S19); (iii) reading of information is parallelized—a single image file can be used to read the information, unlike sequential reading in magnetic tape;[29] (iv) information can be encrypted with novel schemes (see the Registration section).

Table 1. Comparison of Methods for Archival Data Storage[a]

method	cost ($/GB)	stability	write speed (MB/s)	read speed (MB/s)
magnetic tape (LTO-7)	0.016[b]	10−30 years	4×10^{2}[c]	4×10^{2}[c]
DNA[23]	>530,000[d]	up to 2000 years claimed[e]	1×10^{-3}[f]	3×10^{-1}[g]
SAMDI[21]	1[h]	not yet determined	1×10^{-6}	3×10^{-6}
fluorescent imaging (this work)	<0.0001[i]	not yet determined	16×10^{-6}	6×10^{-2}

[a]Magnetic tape is the most common technology used to store data for archival purposes. DNA data storage and self-assembled monolayer-desorption and ionization (SAMDI) data storage are molecular information storage strategies that have received interest in the research community. This work describes storage of information in mixtures of fluorescent molecules. [b]Total revenue divided by total data volume of tapes shipped in a year.[24] [c]Current generation (LTO-9). [d]Reports vary ($530,000−$31,250,000 per GB written). DNA sequencing also incurs cost. [e]Estimated lifetime of DNA encapsulated in silica.[17] [f]Overall throughput is estimated to be on the order of kB/s.[25] Specific values for writing rates are not reported. [g]Using a single state-of-the-art sequencing device.[26] [h]Self-assembled monolayers for matrix-assisted laser desorption/ionization;[21] [i]See Supporting Information. section S19 for detailed calculations.

RESULTS AND DISCUSSION

Choice of Dye Molecules. We chose seven commercially available fluorescent dye molecules with different emission maxima to demonstrate our strategy (Figure 2). The detection technique, a multichannel fluorescence detector, uses a linear array of detection channels to resolve multiple emission bands in parallel and enables spatially resolved information on the presence or absence of the dye molecules to be obtained in a single scan across the substrate. In principle, this technique can be expanded to incorporate more dyes as well (and encode more information in the same amount of area). The dyes are dissolved in dimethyl sulfoxide (DMSO), filtered through a 0.45 μm polysulfone syringe filter, and injected into the inkjet printer cartridge (see Table S1 for concentrations). Figure 2A lists the dyes used in this study. The optimal concentrations of the dyes were determined empirically by observing their fluorescence intensity in a microscope.

"Writing" Information. Inkjet printing is a material deposition technique that has enabled high-resolution microfabrication with specialized materials and has been demonstrated to be applicable to areas such as electronics,[30,31] displays,[32,33] drug discovery[34,35] and others.[36] Inkjet printing has four attractive features: (i) additive operation, where drops are deposited only where needed; (ii) the ability to use a variety of inks (aqueous, organic, nanoparticle composites, biological materials, etc.); (iii) scalability to high throughput and large substrate area; (iv) lower cost than photolithography-based patterning. We use inkjet printing (other technologies like aerosol-jet printing[37] and electrohydrodynamic jet printing[38] provide better printing resolution but are either too expensive or are not commercially available) to print

https://doi.org/10.1021/acscentsci.1c00728
ACS Cent. Sci. 2021, 7, 1728−1735

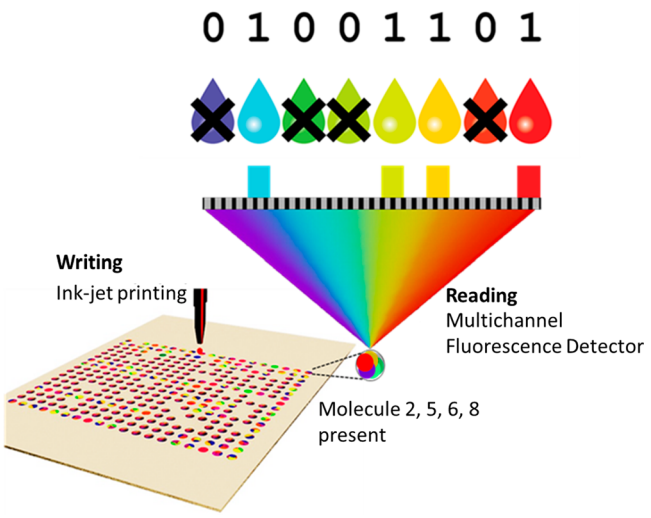


Figure 1. A schematic diagram of the "writing" and "reading" process. ASCII information is converted to a binary bit string which is then encoded into printable patterns and printed with an inkjet printer. The presence or absence of dye molecules at a location represents a byte of data. The information is written on an epoxy substrate which contains free amino groups. Printing of the dyes leads to an amide bond formation between the substrate and the dye and leads to covalent immobilization of the dye onto the substrate at a specific location. Imaging of the printed substrate using a multichannel fluorescence detector represents the "reading" of the written information. The multichannel fluorescence detector can, simultaneously and independently, detect the presence or absence of the dye molecules at a specific location. One very important feature of our approach is that the registration of the dyes with respect to each other is not important for decoding the stored information.

1 pL droplets with a 30-μm center-to-center distance between adjacent spots on the substrate (Supporting Information, Figure S4). To demonstrate storage of information at high density, we wrote the first section of one of the most seminal research papers in scientific history: "Experimental researches in electricity" by Michael Faraday, *Phil. Trans. R. Soc. Lond.* **1832**, *122*, 125−162 (Supporting Information, section S18). This text contains 14 075 characters (i.e., 14 075 bytes of information when converted to ASCII).

"Reading" Information. Fluorescence imaging is a powerful tool for high-resolution characterization of biological samples and materials. The availability of a variety of fluorescent dyes enables unprecedented control in the labeling of specific sites on the sample. Recently, the analysis of spectral data sets and the separation of signals by spectral imaging, combined with linear unmixing, have overcome problems of spectral overlap for fluorescent dyes, and is used widely in biological systems.[39]

We used a Zeiss LSM 800 fluorescence microscope, which has one of the most versatile implementations for spectral imaging. In this technique, fluorescence emission passes through a pinhole and is separated by wavelength by a diffraction grating (Supporting Information, Figure S5). The spectrally resolved light is then projected onto a linear array of 34 detection channels in a photomultiplier detector. The wavelength of emitted light is determined by the position of the channel receiving the photons. This system allows very precise determination of the intensity of peaks separated by only a few nanometers and thus the concentration of the dyes responsible. The presence/absence of a specific fluorescent dye molecule at a specific location on the substrate can thus be determined.

Choice of Substrate. Long-term storage of information requires the formation of thermodynamically stable bonds with very long half-lives. An amide bond is one of the most thermodynamically stable bonds available to organic chemists.[40] In our strategy, we used *N*-hydroxysuccinimide-functionalized dye molecules, which spontaneously react with amino groups to form amide bonds. We synthesized a cross-linked epoxy polymer with an excess of the amine curing agent. This substrate contains reactive secondary amino groups (Supporting Information, Figure S2).[41] The epoxy polymer is processed by hot-pressing a mixture of bisphenol A diglycidyl ether and triethylene tetramine at 120 °C between a glass coverslip and a flat PDMS surface (see Experimental Section). We control the pressure (70 psi) to obtain 50-μm thick films (Supporting Information, Figures S2 and S3). It is important to have a flat surface for the substrate because irregularities in thickness lead to blurring of the image and incomplete focusing in the microscope on reading the information (Supporting Information Figure S16 for an example).

Encoding Scheme for "Writing" of Binary Information. A binary representation of ASCII characters consists of eight bits where each bit is either "0" or "1".[42] In our encoding scheme, the binary representation of each ASCII character in the bit string is assigned a position (positions 1−8, Figure 3). This position is assigned to a fluorescent dye molecule (here, we assign dyes to positions in the order of increasing emission maxima). The bit strings for the positions are then used to generate a printable pattern. Here, we generate a square pattern out of the bit strings, but, in principle, any pattern geometry is possible. These square patterns are then sequentially printed on the substrate using an inkjet printer. "0" indicates absence of a dye molecule, and "1" indicates the presence of a dye molecule. For printable ASCII characters, the first binary digit is always "0", and hence, the first square pattern is always a blank pattern. Thus, we require only seven dye molecules for data storage of printable eight-bit ASCII characters.

Registration. Figure 3B shows a schematic representation of the fact that registration of the printed grids of fluorescent molecules is not required. The fluorescent molecules, when deposited onto the substrate, lie on a grid where the presence or absence of the molecule at the intersection of the gridlines determines binary information (i.e., "0" or "1"). When these grids are sequentially printed onto the substrate, any offset between grids of different fluorescent molecules does not make a difference to the output obtained on reading through a multichannel-fluorescence detector. To help in determining the position of the grid, we place three dots that serve as "calibration spots" as shown in Figure 3B.

We used the Fujifilm Dimatix DMP 2831 inkjet printer to deposit the dye molecules onto the substrate. As this inkjet printer can accommodate only one cartridge at a time, we manually changed the cartridges (each containing one fluorescent dye solution) to print the computer-generated images. We needed 7 manual cartridge changes, and it took 116 s on an average to write each pattern for "Experimental researches in electricity" at 30 μm center-to-center spot distance on a 7.2 mm × 7.2 mm substrate area. This time and area translate into a writing speed of 16 bytes/s.

The substrate with the written information was placed in a Zeiss LSM 880 fluorescence microscope in an inverted

https://doi.org/10.1021/acscentsci.1c00728
ACS Cent. Sci. 2021, 7, 1728−1735

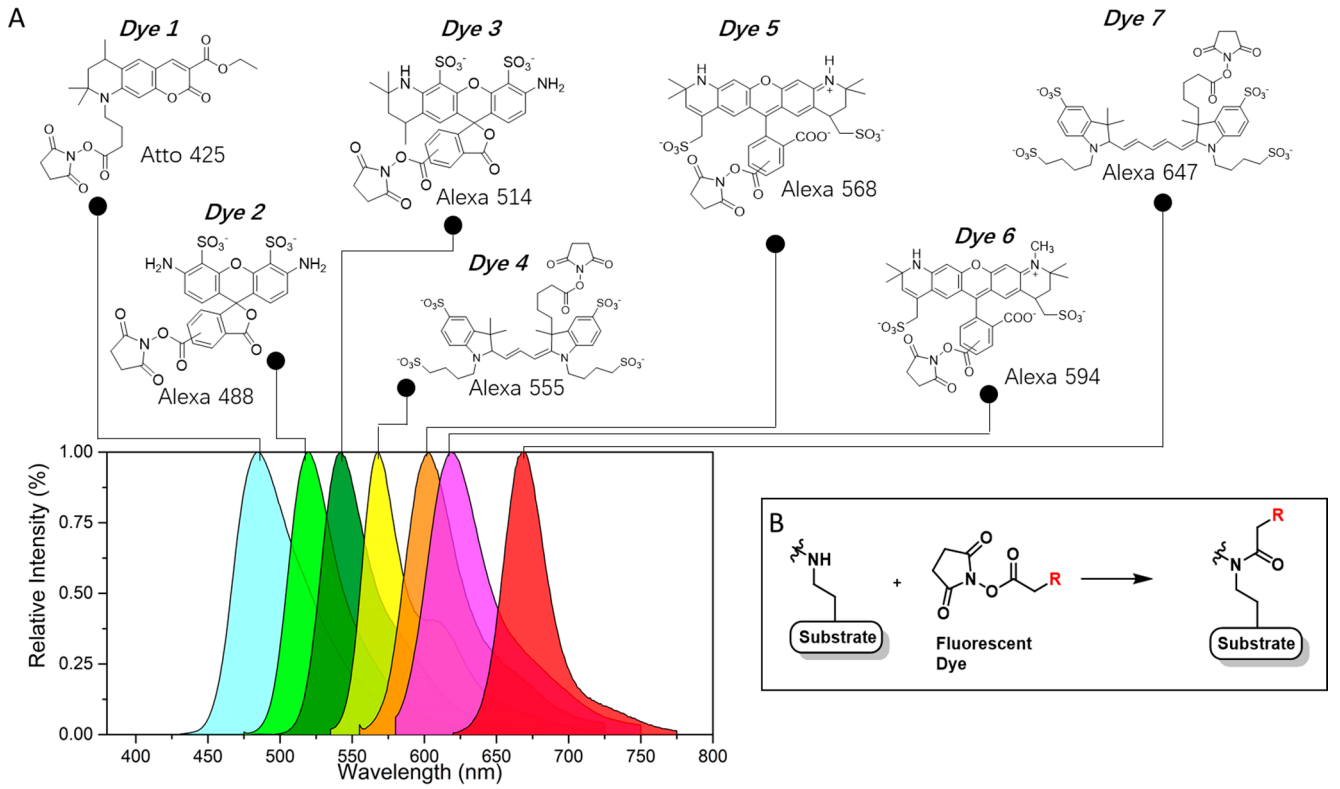


Figure 2. Optically distinguishable fluorescent dyes. (A) Structures of the fluorescent dyes used in this study along with their emission spectra in dimethyl sulfoxide. (B) Reaction scheme for the covalent immobilization of the dye molecule on the substrate. Amino groups in the substrate react with the *N*-hydroxysuccinimide derivatives of the fluorescent dye to link the fluorescent dye to the substrate with an amide bond.

configuration. Four lasers (405 nm, 488 nm, 561 nm, 633 nm) were chosen to excite all the dyes simultaneously in the visible spectrum region (410−695 nm). Using the in-built spectral imaging function, we could resolve all the patterns with very good spatial resolution for each dye. Figure 4B show cropped regions of the unmixed images for all seven dye molecules. It took approximately 240 s to record the image, giving an effective reading speed of 58.64 kilobytes/s (469 kilobits/s). Here we use an inkjet printer and print at the highest resolution, where it is not possible to specify whether the grids are offset or perfectly overlap. The multichannel fluorescence detector is required to show the presence/absence of a dye in the same location as other dyes if they perfectly overlap.

Decoding the Information. It is straightforward to use image analysis to decode the stored information. The individual patterns are read using a simple Python program script using the OpenCV computer vision library.[43] We obtained good accuracy (99.64%) of the recovered information (measured as the number of bits read correctly as a percentage of the total number of bits). This accuracy can be improved with more sophisticated image analysis techniques and error correction codes.[44] The most common reason for inaccuracies during reading were dust particles adhering to the substrate surface.

Stability of the Information. Photobleaching is the attenuation of fluorescence intensity of a fluorophore molecule, primarily due to the cleavage of covalent bonds in the molecule on reaction with oxygen. In our experiments, photobleaching did not significantly affect our recorded information. As compared to traditional biological labeling experiments, we use a high concentration of the fluorescent dye (micromolar quantities). Two benefits of using high concentration of dyes are (i) low laser power is required to excite the fluorescent dyes at a location, (ii) lower laser power also decreases the rate of photobleaching. In our experiments, a 2 mm × 2 mm portion of the information was continuously read 1000 times in air without significant loss in intensity (Figure 5). After 1000 reading cycles, dye 425 showed the largest reduction in intensity (~21%), while all other dyes showed a <15% change in intensity.

Storage of Digital Images. Our strategy of storage of information for ASCII data can be applied to store non-ASCII data as well. As shown in Figure 6, we converted a 3 kilobyte JPEG image of Michael Faraday into a bit string, encoded the bit string to print in seven fluorescent dyes, and inkjet printed the molecules onto the substrate. In this case, as the data are already in a compressed format (JPEG), the quality of recovered data is much more sensitive to errors than when it is in a loss-less image encoding format. An example of the image with 0.4% printing errors (0.4% bits read wrong as compared to the input bit string) is given in the Supporting Information (Figure S17).

Our technology for storage of information in mixtures of fluorescent molecules can be expanded with the use of other fluorophores with narrow emission bandwidths (e.g., quantum dots[45] or J-aggregates[46]). An expanded palette of fluorophores will allow for the use of more fluorophores per location and could also allow simpler, band-pass filter-based reading, eliminating the requirement for an expensive multichannel fluorescence detector. More sophisticated drop-on-demand technologies (electrohydrodynamic inkjet printer commercialized by SIJ corporation, Japan, dip-pen lithography,[47] etc.) can print at much higher resolutions (sub-1 μm spot−spot distance). At 1 μm spot-to-spot distance with eight fluorescent

https://doi.org/10.1021/acscentsci.1c00728
ACS Cent. Sci. 2021, 7, 1728−1735

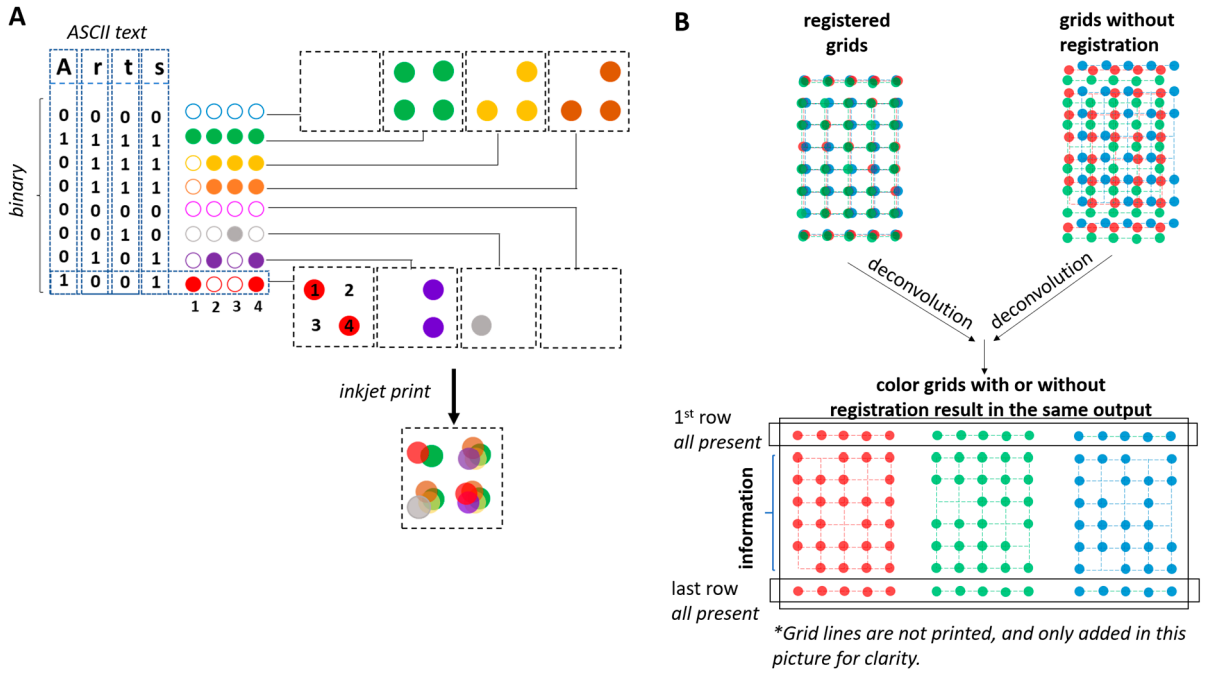


Figure 3. Encoding and registration. (A) Encoding scheme for storage of data for ASCII characters. The algorithm converts the input ASCII string into binary bit strings. Each specific position in the binary data is combined to generate a separate bit string, which corresponds to a specific fluorescent molecule. These eight-bit strings are then converted into a pattern (here, a square pattern, but, in principle, any shape of an array of spots can be used) and printed sequentially onto the substrate with an inkjet printer. (B) A schematic representation demonstrating that the registration of different colors of the printed grids is not required. The grids, when printed onto the substrate, can either be perfectly registered or be printed with an offset. In both cases, as information is read using fluorescence emission at predetermined wavelengths, the patterns can be read independently (1 = presence of the dye and 0 = absence of the dye). Independent read-out from each channel of the fluorescent detector facilitates this "non-registered" information storage.

dyes, the areal storage density will be 5 Gbits/in^2, which is comparable to the latest generation of magnetic tape (LTO-8, areal density: 8 Gbit/in^2). Another area for improvement is the use of error correction codes to decrease error rates (e.g., Reed Solomon error correction codes[44] have been extensively used to decrease error rates in optical media like compact discs, and Blu-ray discs).

CONCLUSION

In conclusion, we report a fundamentally new molecular data storage technology that leverages the optical characteristics of conjugated molecules. A multichannel fluorescence detector enables the simultaneous and independent detection of the presence or absence of a molecule in a mixture on a surface. The "writing" process uses inkjet printing wherein molecules that are deposited onto the surface form an amide bond to link the dye molecules to the substrate. An important characteristic of this information storage method is that registration of the individual molecules is not required. This characteristic is, to our best knowledge, unique; it differentiated this method from existing optical data storage technologies.

We also show that multiple (>1000) readouts of such optical molecular information are possible without significant loss of information through bleaching and other mechanisms. This is also unique as compared to other molecular information storage systems which involve destructive reading (e.g., sequencing of DNA[17] or laser-ablation of oligopeptides[21]). We have also demonstrated the fastest reading speed of any of the molecular information storage methods (0.469 Mbits/s). Access to newer drop-on-demand technologies like electrohydrodynamic inkjet printing would enable commercially competitive areal information density.

This optical molecular information storage technology presents solutions to important problems that are faced by emerging molecular information storage technologies: energy used for storage, cost, and ability to resist corruption.

EXPERIMENTAL SECTION

Safety. Epoxy resins are known skin sensitizers and should be handled carefully with all safety precautions and personal protective equipment in a well-ventilated fume hood. Caution must be taken while handling the reactive fluorescent dyes as their safety hazards are not fully known.

Materials. AlexaFluor dyes were purchased from Thermo Fisher and used without further purification. Atto 425 dye, dry dimethyl sulfoxide (DMSO), bisphenol A diglycidyl ether, and triethylenetetramine were purchased from Sigma-Aldrich and used without further purification.

Fabrication of the Epoxy Substrate. Bisphenol A diglycidyl ether (2.4 g, 7 mmol) was mixed with 0.6 g of triethylene tetramine (4.2 mmol, 3 equiv). This solution was vigorously stirred for 2 min and degassed under a vacuum (80 mbar) for 5 min. This solution (2.6 mL) was poured onto a glass slide and placed inside a heat-press. A PDMS (Sylgard 184) block (10 cm × 10 cm × 0.5 cm) was placed on top of this solution. The PDMS slab plays two roles: (i) it ensures that the top surface is flat, (ii) it does not stick to the epoxy film, and hence it is easy to remove after the epoxy polymer has cured. The polymer was cured under 20 psi pressure at 120 °C for 30 min. The PDMS layer was manually removed, and the

https://doi.org/10.1021/acscentsci.1c00728
ACS Cent. Sci. 2021, 7, 1728−1735

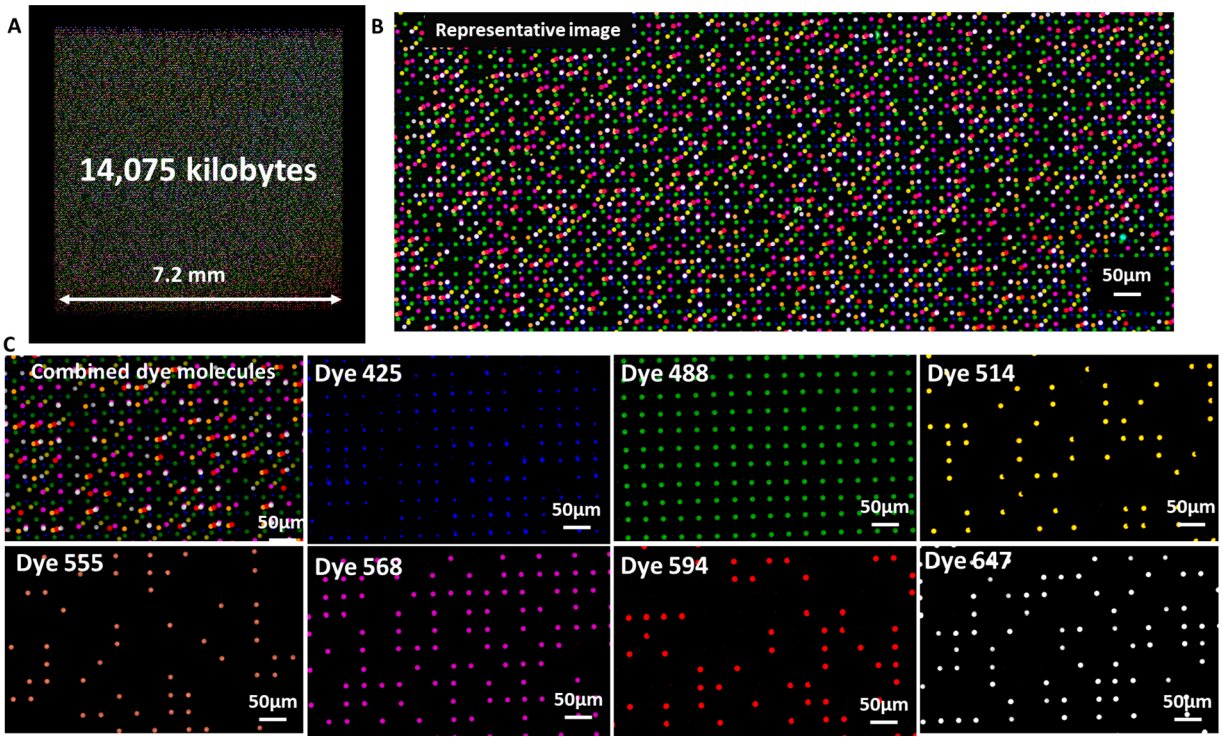
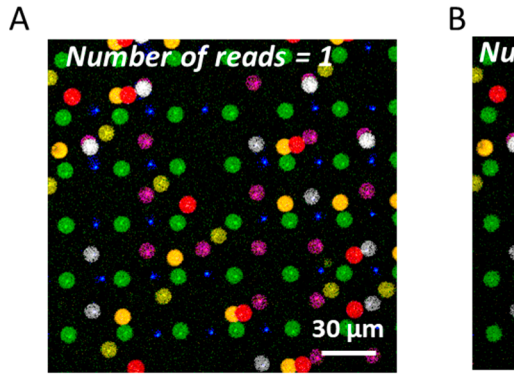


Figure 4. Reading of information. (A) Fluorescence microscope image of the first section of Faraday's "Experimental researches in electricity" on a 50 μm epoxy polymer film written using the encoding scheme shown in Figure 3. The image was recorded with excitation using four lasers simultaneously (405 nm, 488 nm, 561 nm, 633 nm). (B) Zoomed-in image of the printed droplets on the epoxy substrate. (C) Linear unmixing of the fluorescent microscope image leads to independent deconvolution of each dye at a location. The panel shows individual grids of fluorescent dye molecules 425, 488, 514, 555, 568, 594, and 647 obtained by spectral unmixing of the original image using the Zeiss Zen Black software.



C

DYE	% INITIAL INTENSITY AFTER 1000 READS
425	79
488	87
514	86
555	91
568	89
594	90
647	85

Figure 5. A subset of the printed area was continuously read 1000 times in the fluorescence microscope. (A) Image of the printed region on the first read. (B) Image of the same region after 1000 reads. All the patterns of the dyes were easily readable after 1000 cycles of reading. (C) Table showing the percentage of the initial fluorescence intensity remaining after reading the data 1000 times in air.

epoxy film on the glass substrate was cooled down to room temperature. This film was then washed with *n*-hexane three times to remove any potential residue left by the PDMS polymer.

Instrumentation. *Inkjet Printing.* Writing was carried out with a Fujifilm Dimatix DMP 2831 printer with a 1 pL printing volume cartridge. The printing parameters are firing voltage of the active nozzle: 16 V; firing voltage of inactive nozzles: 12 V, printing height: 0.5 mm. Cartridges containing the fluorescent dyes were manually changed for each dye.

Atomic Force Microscopy. A sample consisting of an epoxy film on a glass slide was first sonicated in isopropyl alcohol for 10 min and then dried under a nitrogen stream. Atomic force microscope (AFM) images were obtained using an Asylum Research Cypher AFM in tapping mode with a 300 kHz cantilever. Supporting Information, Figure S2 provides the data.

Profilometry. A sample consisting of an epoxy film on a microscope glass slide (VWR, 1 mm thickness) was sliced with a razor blade to introduce trenches into the film, sonicated in isopropyl alcohol for 10 min, and then dried under nitrogen. Profilometry was performed using a Bruker DektakXT profilometer equipped with a 5-μm radius diamond tip and with 3 mg of applied force.

Supporting Information, Figure S3 provides the data.

Microscopy. Reading was carried out using a Zeiss LSM 880 confocal microscope with an in-built 34 channel photo-multiplier detector. Four lasers were used to excite all the dyes:

https://doi.org/10.1021/acscentsci.1c00728
ACS Cent. Sci. 2021, 7, 1728−1735

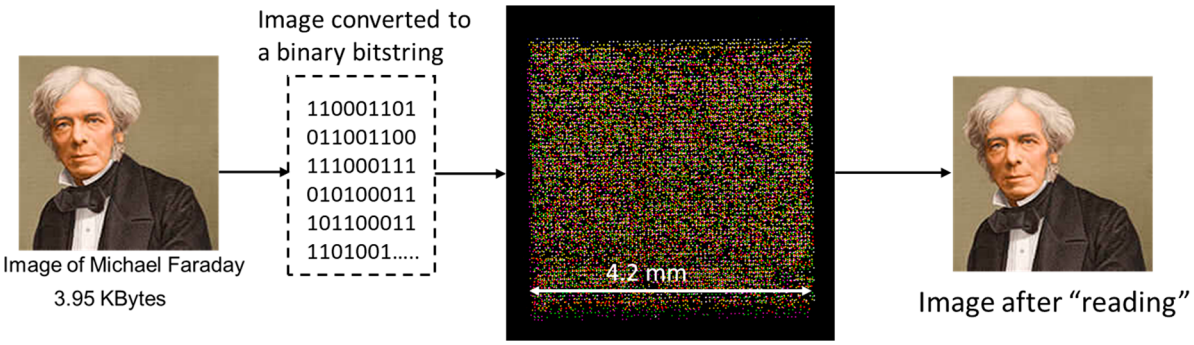


Figure 6. A JPEG image of Michael Faraday (3.95 KB) was converted into patterns for the seven fluorescent dyes used in this study. This information was printed onto a 4.2 mm × 4.2 mm epoxy substrate. An example of a decoded image with 0.4% printing errors is shown in the Supporting Information, Figure S17. The image of Michael Faraday has been reproduced with permission from Getty Images.

405 nm, 488 nm, 561 nm, 633 nm. The in-built multichannel fluorescence detector was calibrated with individual dyes printed on the epoxy substrate.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acscentsci.1c00728.

Absorption and emission spectra of fluorescent dyes, list of concentrations of the fluorescent dyes used, characterization of the substrate, spectrally unmixed images of each dye, computer-generated printing patterns, images of printed droplets, image of problems while focusing in the microscope, stored and decoded image of Michael Faraday containing errors, encoded text from Faraday's "Experimental Researches in Electricity", estimation of the lower bound of cost per GB, typical inkjetting waveform, and an estimation of lifetime of the fluorescent dyes (10^{10} years by extrapolation using the Arrhenius equation) (PDF)

■ AUTHOR INFORMATION

Corresponding Author

George M. Whitesides − *Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, United States;* orcid.org/0000-0001-9451-2442; Email: gwhitesides@gmwgroup.harvard.edu

Authors

Amit A. Nagarkar − *Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, United States*

Samuel E. Root − *Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, United States*

Michael J. Fink − *Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, United States;* orcid.org/0000-0003-0023-8767

Alexei S. Ten − *Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, United States;* orcid.org/0000-0003-0040-9132

Brian J. Cafferty − *Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, United States*

Douglas S. Richardson − *Harvard Center for Biological Imaging, Cambridge, Massachusetts 02138, United States*

Milan Mrksich − *Department of Chemistry and Department of Biomedical Engineering, Northwestern University, Evanston, Illinois 60208, United States;* orcid.org/0000-0002-4964-796X

Complete contact information is available at: https://pubs.acs.org/10.1021/acscentsci.1c00728

Author Contributions

G.M.W. conceived the idea of molecular information storage in mixtures of molecules. A.A.N., B.J.C., S.E.R. and G.M.W. postulated information storage in mixtures of fluorescent molecules. A.A.N., S.E.R., and D.R. conducted the experiments and performed the analysis. A.S.T., M.J.F., and M.M. provided valuable input for improvement of the manuscript. A.A.N., S.E.R., M.J.F., and G.M.W. wrote the manuscript with inputs from all the authors.

Notes

The authors declare the following competing financial interest(s): A.A.N., A.S.T., and M.J.F. acknowledge an equity interest in Datacule Inc. G.M.W. acknowledges an equity interest and a board position in Datacule Inc.

■ ACKNOWLEDGMENTS

This work was supported by Defence Advanced Research Projects Agency (DARPA) under Award No. W911NF-18-2-0030.

■ REFERENCES

(1) Lloyd, S. Ultimate physical limits to computation. *Nature* **2000**, *406*, 1047−1054.

(2) Shulaker, M. M.; et al. Three-dimensional integration of nanotechnologies for computing and data storage on a single chip. *Nature* **2017**, *547*, 74−78.

(3) Pansare, A. V.; et al. In Situ Nanoparticle Embedding for Authentication of Epoxy Composites. *Adv. Mater.* **2018**, *30*, 1801523.

(4) Salahuddin, S.; Ni, K.; Datta, S. The era of hyper-scaling in electronics. *Nature Electronics* **2018**, *1*, 442−450.

(5) Baliga, J.; Ayre, R. W.; 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) Clelland, C. T.; Risca, V.; Bancroft, C. Hiding messages in DNA microdots. *Nature* **1999**, *399*, 533−534.

(8) Green, J. E.; et al. A 160-kilobit molecular electronic memory patterned at 1011 bits per square centimetre. *Nature* **2007**, *445*, 414−417.
(9) Colquhoun, H.; Lutz, J.-F. Information-containing macromolecules. *Nat. Chem.* **2014**, *6*, 455−456.
(10) Zhirnov, V.; Zadegan, R. M.; Sandhu, G. S.; Church, G. M.; Hughes, W. L. Nucleic acid memory. *Nat. Mater.* **2016**, *15*, 366−370.
(11) Al Ouahabi, A.; Charles, L.; Lutz, J.-F. Synthesis of non-natural sequence-encoded polymers using phosphoramidite chemistry. *J. Am. Chem. Soc.* **2015**, *137* (16), 5629−5635.
(12) Charles, L.; Laure, C.; Lutz, J.-F.; Roy, R. K. MS/MS sequencing of digitally encoded poly (alkoxyamine amide)s. *Macromolecules* **2015**, *48* (13), 4319−4328.
(13) Amalian, J.-A.; Trinh, T. T.; Lutz, J.-F.; Charles, L. MS/MS digital readout: analysis of binary information encoded in the monomer sequences of poly (triazole amide)s. *Anal. Chem.* **2016**, *88* (7), 3715−3722.
(14) Goodwin, C. A.; Ortu, F.; Reta, D.; Chilton, N. F.; Mills, D. P. Molecular magnetic hysteresis at 60 K in dysprosocenium. *Nature* **2017**, *548*, 439−442.
(15) Bhat, W. A. Bridging data-capacity gap in big data storage. *Future Generation Computer Systems* **2018**, *87*, 538−548.
(16) Gu, M.; Li, X.; Cao, Y. Optical storage arrays: a perspective for future big data storage. *Light: Sci. Appl.* **2014**, *3*, e177.
(17) Grass, R. N.; Heckel, R.; Puddu, M.; Paunescu, D.; Stark, W. J. Robust Chemical Preservation of Digital Information on DNA in Silica with Error-Correcting Codes. *Angew. Chem., Int. Ed.* **2015**, *54*, 2552−2555.
(18) Goodwin, S.; McPherson, J. D.; McCombie, W. R. Coming of age: ten years of next-generation sequencing technologies. *Nat. Rev. Genet.* **2016**, *17*, 333−351.
(19) Kosuri, S.; Church, G. M. Large-scale de novo DNA synthesis: technologies and applications. *Nat. Methods* **2014**, *11*, 499−507.
(20) Organick, L.; et al. Random access in large-scale DNA data storage. *Nat. Biotechnol.* **2018**, *36*, 242−248.
(21) Cafferty, B. J.; Ten, A. S.; Fink, M. J.; Morey, S.; Preston, D. J.; Mrksich, M.; Whitesides, G. M. Storage of Information Using Small Organic Molecules. *ACS Cent. Sci.* **2019**, *5* (5), 911−916.
(22) Feenstra, A. D.; Dueñas, M. E.; Lee, Y. J. Five Micron High Resolution MALDI Mass Spectrometry Imaging with Simple, Interchangeable, Multi-Resolution Optical System. *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 434−442.
(23) Catalogue Press Release, https://techthelead.com/start-up-stores-all-16g-of-wikipedia-onto-dna-strand/.
(24) Fontana, R. E.; Decad, G. M. Moore's law realities for recording systems and memory storage components: HDD, tape, NAND, and optical. *AIP Adv.* **2018**, *8*, 056506.
(25) Ceze, L.; Nivala, J.; Strauss, K. Molecular digital data storage using DNA. *Nat. Rev. Genet.* **2019**, *20*, 456−466.
(26) Meiser, L. C.; Koch, J.; Antkowiak, P. L.; Stark, W. J.; Heckel, R.; Grass, R. N. DNA synthesis for true random number generation. *Nat. Commun.* **2020**, *11*, 5869.
(27) Lantz, M. Why the Future of Data Storage Is (Still) Magnetic Tape. *IEEE Spectrum: Technology Engineering and Science News*, Aug. 2018.
(28) Bertram, H.; Cuddihy, E. Kinetics of the humid aging of magnetic recording tape. *IEEE Trans. Magn.* **1982**, *18*, 993−999.
(29) Sandsta, O.; Midtstraum, R. Analysis of retrieval of multimedia data stored on magnetic tape. *Proceedings International Workshop on Multi-Media Database Management Systems*; Cat. No. 98TB100249; Dayton, OH, USA, 1998; pp 54−63.
(30) Sirringhaus, H.; et al. High-resolution inkjet printing of all-polymer transistor circuits. *Science* **2000**, *290*, 2123−2126.
(31) Shimoda, T.; et al. Solution-processed silicon films and transistors. *Nature* **2006**, *440*, 783−786.
(32) Shimoda, T.; Morii, K.; Seki, S.; Kiguchi, H. Inkjet printing of light-emitting polymer displays. *MRS Bull.* **2003**, *28*, 821−827.
(33) Chang, S. C.; et al. Multicolor organic light-emitting diodes processed by hybrid inkjet printing. *Adv. Mater.* **1999**, *11*, 734−737.
(34) Lemmo, A. V.; Rose, D. J.; Tisone, T. C. Inkjet dispensing technology: Application in drug discovery. *Curr. Opin. Biotechnol.* **1998**, *9*, 615−617.
(35) Heller, M. J. DNA microarray technology: Devices, systems, and applications. *Annu. Rev. Biomed. Eng.* **2002**, *4*, 129−153.
(36) Hiller, J.; Mendelsohn, J. D.; Rubner, M. F. Reversibly erasable nanoporous anti-reflection coatings from polyelectrolyte multilayers. *Nat. Mater.* **2002**, *1*, 59−63.
(37) Wilkinson, N. J.; Smith, M. A.; Kay, R. W.; Harris, R. A. A review of aerosol jet printing—a non-traditional hybrid process for micro-manufacturing. *International Journal of Advanced Manufacturing Technology* **2019**, *105*, 4599−4619.
(38) Park, J.-U.; et al. High-resolution electrohydrodynamic jet printing. *Nat. Mater.* **2007**, *6*, 782−789.
(39) Valm, A. M.; Oldenbourg, R.; Borisy, G. G. Multiplexed Spectral Imaging of 120 Different Fluorescent Labels. *PLoS One* **2016**, *11*, e0158495.
(40) Martin, R. B. Free energies and equilibria of peptide bond hydrolysis and formation. *Biopolymers* **1998**, *45* (5), 351−353.
(41) Pansare, A. V.; et al. Shape-Coding": Morphology-Based Information System for Polymers and Composites. *ACS Appl. Mater. Interfaces* **2020**, *12*, 27555−27561.
(42) Hieronymus, J. L. ASCII phonetic symbols for the world's languages Worldbet. *J. Int. Phonetic Assoc.* **1993**, *23*.
(43) Pulli, K.; Baksheev, A.; Kornyakov, K.; Eruhimov, V. Realtime Computer Vision with OpenCV. *Queue* **2012**, *10*, 40−56.
(44) Reed, I. S.; Solomon, G. Polynomial Codes over Certain Finite Fields. *J. Soc. Ind. Appl. Math.* **1960**, *8* (2), 300−304.
(45) Owen, J.; Brus, L. Chemical Synthesis and Luminescence Applications of Colloidal Semiconductor Quantum Dots. *J. Am. Chem. Soc.* **2017**, *139* (32), 10939−10943.
(46) Würthner, F.; Kaiser, T. E.; Saha-Möller, C. R. J-Aggregates: From Serendipitous Discovery to Supramolecular Engineering of Functional Dye Materials. *Angew. Chem., Int. Ed.* **2011**, *50*, 3376−3410.
(47) Liu, G.; Petrosko, S. H.; Zheng, Z.; Mirkin, C. A. Evolution of Dip-Pen Nanolithography (DPN): From Molecular Patterning to Materials Discovery. *Chem. Rev.* **2020**, *120*, 6009−6047.

https://doi.org/10.1021/acscentsci.1c00728
ACS Cent. Sci. 2021, 7, 1728−1735